

T.C.
ORDU ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
MATEMATİK ANABİLİM DALI

**KORELASYON KATSAYISININ FARKLI GEOMETRİK YORUMLARI,
İSTATİSTİKTE LİNEER MODELLERİN GEOMETRİSİ, LİNEER
MODELLERDE LİNEER KISITLAMALAR ALTINDA PARAMETRE
TAHMİNLERİ VE HİPOTEZ TESTİ**

FATMA BUĞLEM YALÇIN

DOKTORA TEZİ

ORDU 2018

TEZ ONAY

Ordu Üniversitesi Fen Bilimleri Enstitüsü öğrencisi **Fatma Buğlem YALÇIN** tarafından hazırlanan ve **Prof. Dr. Cemil YAPAR** danışmanlığında yürütülen “**KORELASYON KATSAYISININ FARKLI GEOMETRİK YORUMLARI, İSTATİSTİKTE LİNEER MODELLERİN GEOMETRİSİ, LİNEER MODELLERDE LİNEER KISITLAMALAR ALTINDA PARAMETRE TAHMİNLERİ VE HİPOTEZ TESTİ**” adlı bu tez, jürimiz tarafından 20 / 11 / 2018 tarihinde oy birliği ile Matematik Anabilim Dalında Doktora tezi olarak kabul edilmiştir.

Danışman : Prof. Dr. Cemil YAPAR

II. Danışman : Prof. Dr. Vedat Suat ERTÜRK, Ondokuz Mayıs Üniversitesi

Başkan : Prof. Dr. Cemil YAPAR
Matematik, Ordu Üniversitesi

İmza :

Üye : Prof. Dr. Selahattin MADEN
Matematik, Ordu Üniversitesi

İmza :

Üye : Dr. Öğr. Üyesi Nurgül OKUR BEKAR
İstatistik, Giresun Üniversitesi

İmza :

Üye : Dr. Öğr. Üyesi Mehmet KORKMAZ
Matematik, Ordu Üniversitesi

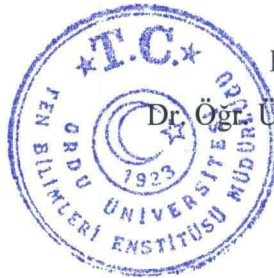
İmza :

Üye : Dr. Öğr. Üyesi Murat GÜL
İstatistik, Giresun Üniversitesi

İmza :

ONAY:

04 / 12 / 2018.. tarihinde enstitüye teslim edilen bu tezin kabulü, Enstitü Yönetim Kurulu'nun 04 / 12 / 2018.. tarih ve 2018... / 566. sayılı kararı ile onaylanmıştır.



Enstitü Müdürü

Dr. Öğr. Üyesi Mehmet Sami GÜLER

TEZ BİLDİRİMİ

Tez yazım kurallarına uygun olarak hazırlanan ve kullanılan intihal tespit programının sonuçlarına göre; bu tezin yazılmasında bilimsel ahlak kurallarına uyulduğunu, başkalarının eserlerinden yararlanılması durumunda bilimsel normlara uygun olarak atıfta bulunulduğunu, tezin içerdiği yenilik ve sonuçların başka bir yerden alınmadığını, kullanılan verilerde herhangi bir tahrifat yapılmadığını, tezin herhangi bir kısmının bu üniversite veya başka bir üniversitedeki başka bir tez çalışması olarak sunulmadığını beyan ederim.

Fatma Buğlem YALÇIN

Buğp

ÖZET

KORELASYON KATSAYISININ FARKLI GEOMETRİK YORUMLARI, İSTATİSTİKTE LİNEER MODELLERİN GEOMETRİSİ, LİNEER MODELLERDE LİNEER KISITLAMALAR ALTINDA PARAMETRE TAHMİNLERİ VE HİPOTEZ TESTİ

Fatma Buğlem YALÇIN

Ordu Üniversitesi
Fen Bilimleri Enstitüsü
Matematik Anabilim Dalı, 2018
Doktora Tezi, 195 s.

Danışman: Prof. Dr. Cemil YAPAR

II. Danışman: Prof. Dr. Vedat Suat ERTÜRK

Bu tez beş bölümden oluşmaktadır. Birinci bölümde teze giriş verilmiştir. İkinci bölümde matrisler, vektörler ve istatistikle ilgili bazı temel bilgilerden söz edilmiştir. Üçüncü bölümde istatistikte lineer modeller ve lineer modellerin geometrisi açıklanmıştır. Sonra, lineer olmayan regresyon modellerinden bahsedilmiştir. Dördüncü bölümde lineer modellerde lineer kısıtlamalar altında parametre tahminleri ve hipotez testi ele alınmıştır. Ayrıca, korelasyon katsayısının farklı geometrik yorumları ortaya koyulmuştur. Özellikle bu geometrik yorumlar Pearson korelasyon katsayısı üzerinde yoğunlaşmıştır. Verilen geometrik yorumlar, iki veri vektörü ve onların regresyon doğruları üzerine kurulan paralelkenarların alanları yardımıyla doğrulanmıştır. Beşinci bölümde sonuç ve öneriler getirilmiştir.

Anahtar Kelimeler: Ortogonallik, Genelleştirilmiş ters, Test istatistikleri, En çok olabilirlik fonksiyonu, En çok olabilirlik oranı, Alışılmış en küçük kareler tahmini, Kısıtlanmış en küçük kareler tahmini.

ABSTRACT

DIFFERENT GEOMETRIC INTERPRETATIONS OF CORRELATION COEFFICIENT, THE GEOMETRY OF THE LINEAR MODELS IN STATISTICS, PARAMETER ESTIMATIONS AND HYPOTHESIS TESTING UNDER LINEAR CONSTRAINTS IN THE LINEAR MODELS

Fatma Buğlem YALÇIN

University of Ordu

Institute for Graduate Studies in Natural and Technology

Department of Mathematics, 2018

PhD Thesis, 195 p.

Supervisor: Prof. Dr. Cemil YAPAR

II. Supervisor: Prof. Dr. Vedat Suat ERTÜRK

This thesis consists of five chapters. In the first chapter, an introduction to thesis has been given. In the second chapter, some basic information about matrices, vectors and associated with statistics have been mentioned. In the third chapter, the linear models and the geometry of the linear models in statistics have been illustrated. Then, non-linear regression models have been mentioned. In the fourth chapter, parameter estimations and hypothesis testing under linear constraints in the linear models have been discussed. Furthermore, different geometric interpretations of correlation coefficient have been produced. These geometric interpretations have been centered especially upon Pearson correlation coefficient and have also been expressed in terms of the areas of parallelograms constructed on two data vectors and their regression lines. In the fifth chapter, conclusions and suggestions have been presented.

Key Words : Orthogonality, Generalized inverse, Pearson Correlation coefficient, Test statistics, Maximum likelihood function, Maximum likelihood ration, Ordinary least squares estimation, Restricted least squares estimation.

TEŐEKKÜR

Tüm alıőmalarım süresince kıymetli zamanlarını bana ayırarak engin bilgi, deneyim ve hoşgörülerıyla tez danışmanlığımı yürüten değerli danışman hocam Sayın Prof. Dr. Cemil YAPAR 'a en içten teşekkürlerimi sunarım.

Bu zorlu ve uzun süreçte ideallerimi gerçekleőtirmemi sağlayan ve bana her konuda destek olan anneme ve rahmetli babama, Ordu Üniversitesi Fen Edebiyat Fakültesi Matematik Bölümü öğretim üyeleri ve öğretim elemanlarına yürekten teşekkürü bir borç bilirim.

İÇİNDEKİLER

	<u>Sayfa</u>
TEZ BİLDİRİMİ	I
ÖZET	II
ABSTRACT	III
TEŞEKKÜR	IV
İÇİNDEKİLER	V
ŞEKİLLER LİSTESİ	VIII
ÇİZELGELER LİSTESİ	IX
SİMGELER ve KISALTMALAR	X
1. GİRİŞ	1
2. GENEL BİLGİLER	2
2.1. Matris ve Vektör Cebiri	2
2.1.1. Tanımlar ve Temel İşlemler	2
2.1.2. Vektör Uzayları ve Altuzaylar	11
2.1.3. Lineer Dönüşüm ve Operatörler	14
2.1.4. Lineer Sistemlerin Çözümü	18
2.1.5. Özdeğerler ve Özvektörler	32
2.1.6. Matris Normları	41
2.1.7. Bir Matrisin Şart Sayısı.....	43
2.1.8. Duyarlılık Analizi : Lineer Sistemlerin Çözümü	45
2.2. Optimizasyon	46
2.2.1. Matris ve Determinantların Diferensiyellenmesi.....	46
2.2.2. Determinantların Diferensiyellenmesi	46
2.2.3. Bir Fonksiyonun Bir Vektöre Göre Türevi	47
2.2.4. Bir Fonksiyonun Bir Matrise Göre Diferensiyeli	47
2.2.5. Bir Fonksiyonun Optimizasyonu	47
2.3. İstatistiksel Gerçekler.....	49
2.3.1. Tahmin	49
2.3.2. Hipotezlerin Testleri ve Güven Aralıkları	51
3. MATERYAL ve YÖNTEM	53
3.1. İstatistikte Lineer Modeller	53
3.1.1. Basit Lineer Regresyon Modeli	53
3.1.2. Çoklu Lineer Regresyon Modeli.....	53
3.1.2.1. β nın ve σ^2 nin Tahmini.....	57
3.1.2.2. Merkezileştirilmiş Biçimdeki Model	59
3.1.3. En Küçük Karelerin Geometrisi.....	61
3.1.4. Genel Lineer Model için Veri	62
3.1.5. Uyum Testlerinin İyiliği (Mükemmelliği)	66
3.1.6. Model Seçimi ve Karşılaştırma.....	67
3.1.7. Çok Değişkenli Genelleştirme	69
3.1.8. Genel Lineer Modelin Önemi	70
3.1.9. Farklı Değişkenlik.....	71
3.1.10. Gaussian Olmayan Dağılımlar ve Genelleştirilmiş Lineer Model.....	72
3.1.11. Bağımlı Gözlemler.....	74
3.1.12. Eşanlı Testetme	75

3.2. Lineer Olmayan Regresyon Modeline Giriş	78
3.2.1. Regresyon Modelleri.....	78
3.2.2. Lineer Olmayan Modellerde Alışılmış En Küçük Kare Yöntemi ve Uygulaması	80
3.2.2.1. Lineer Olmayan Modeller için Minimumlaştırmanın En Küçük Kareler Kriteri	80
3.2.2.2. Dönüştürülmüş Modele OLS nin Grafiksel Uygulaması.....	83
3.2.3. En Küçük Kare Uyum Unsuru	84
3.2.3.1. En Küçük Kareler En İyi Uyum Doğrusu	88
3.2.3.2. En Küçük Kare En İyi Uyum Düzlemi	89
3.2.3.3. En Küçük Kare En İyi Uyum Çemberi	90
3.2.3.4. En Küçük Kareler Yöntemini Kullanarak En İyi Çember Uydurma	92
3.2.3.5. En Küçük Kareler Küresi	96
3.2.3.6. En Küçük Kareler En İyi Uyum Silindirleri	99
3.3. Son Dönemlerde Kuadratik Eğrilerin ve Yüzeylerin En Küçük Kareler Uyumunu Üzerinde Yapılan Bazı Önemli Çalışmalar.....	101
3.3.1. Çemberlerin ve Elipslerin En Küçük Kareler Uyumunu	101
3.3.2. Kuadratik Eğrilerin ve Yüzeylerin En Küçük Kareler Uyumunu.....	106
3.3.2.1. Konikler Üzerine İzdüşüm	107
3.3.2.2. Kuadratikler Üzerine İzdüşüm	113
4. BULGULAR ve TARTIŞMA.....	118
4.1. Lineer Modellerde Lineer Kısıtlamalar Altında Parametre Tahminleri ve Hipotez Testleri.....	118
4.1.1. Lineer Modellerde Kesin Lineer Eşitlik Kısıtlamaları.....	118
4.1.1.1. Lineer Eşitlik Kısıtlamaları Altında β nın En Küçük Kareler Tahmini..	119
4.1.1.2. $\hat{\beta}_c$ yı Elde Etmek İçin Bir Başka Yöntem.....	123
4.1.1.3. β nın Kısıtlanmış En Çok Olabilirlik Tahmini	129
4.1.1.4. Tahmin Edilen Hataları Kullanarak Model Üzerindeki Kısıtlamaları Test Etme	134
4.1.1.5. Bir Olabilirlik Oran (LR) Testi Kullanarak Model Üzerindeki Kısıtlamaları Test Etme	136
4.1.2. Lineer Modellerde Stokastik Lineer Kısıtlamalar Altında Parametre Tahmini ve Hipotez Testi	141
4.1.2.1. Sade (Pure-Saf) ve Karma Regresyon Tahmini.....	143
4.1.2.2. Hipotezin Test Edilmesi.....	146
4.1.3. Lineer Modellerde Lineer Eşitsizlik Kısıtlamaları	147
4.2. Örneklem Geometrisi, Rasgele Örneklem ve Korelasyon Katsayısının Farklı Geometrik Yorumları	148
4.2.1. Örneklem Geometrisi.....	149
4.2.1.1. Örneklem Geometrisel Yorumu.....	158
4.2.2. Rasgele Örneklem, Örneklem Ortalamasının ve Kovaryans Matrisinin Beklenen Değerleri.....	159
4.2.2.1. (Yansız) Örneklem Varyans-Kovaryans Matrisi	163
4.2.3. Genelleştirilmiş Varyans.....	163
4.2.4. Korelasyon Katsayısının Farklı Geometrik Yorumları.....	166
4.2.4.1. Standartlaştırılmış Vektörlerin Kullanımı.....	170
4.2.4.2. \mathbf{X} ve \mathbf{Y} Vektörlerinin Geometrisi	176

4.2.4.3. x ve y Standartlaştırılmış Vektörlerinin Geometrisi	178
4.2.4.4. Basit Lineer Regresyon Doğrularına Dayanan Bir Genişletme	180
4.2.4.5. Korelasyon Katsayısına Farklı Bir Cebirsel Bakış	182
4.2.4.6. Korelasyon Katsayısına Başka Bir Cebirsel Bakış	187
5. SONUÇ ve ÖNERİLER	189
6. KAYNAKLAR	190
ÖZGEÇMİŞ	194

ŞEKİLLER LİSTESİ

Sayfa

Şekil 2.1. V vektör uzayının W ye dönüşümü kavramı	14
Şekil 2.2. Ek (adjoint) Dönüşümün Tanımı	17
Şekil 2.3. Lineer Sistemlerin Sınıflandırılması.....	20
Şekil 2.4. A^+ in A nın Moore-Penrose Genelleştirilmiş Tersine Olduğu Çeşitli Durumlarda Lineer Sistemlerin Çözümü	21
Şekil 2.5. Eğik İzdüşümler.....	28
Şekil 3.1. y nin X sütun uzayına izdüşümü.....	62
Şekil 3.2. $y = X\beta + \varepsilon$ modelinin geometrik gösterimi.....	65
Şekil 3.3. Sekiz olası modelin bir latisi (örgüsü).....	68
Şekil 3.4. $C(X)$ ve $C(Y)$ uzayları tarafından üretilen u ve v vektörlerinin gösterimi	70
Şekil 3.5. Nüfus miktarının katı atık miktarı üzerindeki etkisi örneğinin grafiği.....	84
Şekil 3.6. Genel Düzlem Geometrisi	86
Şekil 3.7. Newton Yönteminin İterasyonu	87
Şekil 3.8. $y = x^2 + 4x + 4$ parabolünden geçen noktalara uydurulan çember	96
Şekil 3.9. Dönüş matrislerini nasıl geliştirileceğinin örneği.....	100
Şekil 3.10. En İyi Uyuma Karşı Cebrik Uyum	102
Şekil 3.11. Öklid-Değişimsiz Algoritmalar	103
Şekil 3.12. Değişimsiz-Olmayan Cebrik Algoritma.....	103
Şekil 3.13. En İyi Uyuma Karşı Cebrik Uyum	104
Şekil 3.14. $t > -b^2$ için $F(t)$ nin Bir Tipik Grafiği ve Newton İterasyonlarının Köke Doğru İlerlemesi.....	110
Şekil 3.15. $F(t)$ nin $-a^2 < t < b^2$ Aralığı Üzerindeki İki Mümkün Görünümü	112
Şekil 4.1. \hat{y} nin iki ortogonal \hat{y}_c ve $\hat{y}_{\perp c}$ parçalarına ayrışımı	140
Şekil 4.2. X veri matrisinin $p = 2$ boyutlu uzayda $n = 3$ noktaları olarak grafiği.151	
Şekil 4.3. X veri matrisinin $n = 3$ boyutlu uzayda $p = 2$ vektörler olarak grafiği. 152	
Şekil 4.4. y_i nin bir \bar{x}_i ortalama bileşeni ve bir $d_i = y_i - \bar{x}_i$ sapma bileşenine ayrışımı.. ..	154
Şekil 4.5. Şekil 4.4 deki d_i sapma vektörleri	156
Şekil 4.6. d_1 ve d_2 sapma vektörleri.....	158
Şekil 4.7. L_{a_1} ve L_{a_2} ile ikizkenar yamuğun gösterimi.....	165
Şekil 4.8. X , Y , $X - j\bar{X}$, $Y - j\bar{Y}$ vektörlerinin ve θ , θ_1, θ_2 , φ açılarının n -boyutlu uzayda şekilsel gösterimleri	167
Şekil 4.9. X ve Y vektörleri üzerine kurulan paralelkenar.....	170
Şekil 4.10. \hat{y} ve \hat{x} vektörleri ve φ , ψ , η ve ξ açıları.....	172
Şekil 4.11. X üzerinde Y nin ve Y üzerinde X in regresyon doğruları.....	174
Şekil 4.12. X ve Y üzerine kurulan ve \dot{X} ve \dot{Y} üzerine kurulan paralelkenarlar.....	177
Şekil 4.13. x , y ve \hat{x} , \hat{y} vektörleri üzerine kurulan eşkenar dörtgenler	179

ÇİZELGELER LİSTESİ

Sayfa

Çizelge 3.1. Açıklayıcı verinin bir minyatür kümesi.....	63
Çizelge 3.2. Hipotez testinde varılan karar.....	76
Çizelge 3.3. Ardışık test etmede varılan karar.....	78
Çizelge 3.4. İlçelerin nüfus miktarı ve toplam ortalama katı atık miktarı.....	83
Çizelge 3.5. $y = x^2 + 4x + 4$ parabolünden geçen birkaç nokta	95
Çizelge 3.6. 6 noktalı veri noktaları kümesi örneği.....	102
Çizelge 3.7. 8 noktalı veri noktaları kümesi örneği.....	103
Çizelge 3.8. Cebirsel Çemberin Başlangıç Parametreler ile Geometrik Uyumu.....	105
Çizelge 3.9. Cebirsel Elipsin Başlangıç Parametreleri ile Geometrik Uyumu	105
Çizelge 3.10. İki Elips Uydurma Yönteminin Karşılaştırması	106
Çizelge 3.11. İki Elips Uydurma Yönteminin Karşılaştırması (G: Geometrik parametreler, K: Kepler parametreleri).....	107
Çizelge 4.1. Fabrikada çalışan işçi sayısı ve kişi başına düşen kazanç.....	164

SİMGELER ve KISALTMALAR

\mathbf{A}'	:	\mathbf{A} matrisinin transpozesi
$\ \mathbf{A}\ _F$:	\mathbf{A} matrisinin Frobenius normu
\mathbf{A}^{-1}	:	\mathbf{A} matrisinin tersi
A_{ij}	:	a_{ij} elemanının eşçarpanı
\mathbf{A}^*	:	\mathbf{A} nın eki
\mathbf{A}_S	:	\mathbf{A} nın simetrik parçası
\mathbf{A}_{SS}	:	\mathbf{A} nın anti-simetrik parçası
B	:	Birden fazla lineer dönüşümlerinin kümesi
\mathcal{B}_1	:	Öklid uzayında standart taban
$\text{boy}[C(\mathbf{A})]$:	$C(\mathbf{A})$ uzayının boyutu
$C(\mathbf{A})$:	\mathbf{A} nın sütun uzayı
$C(\mathbf{A})^\perp$:	$C(\mathbf{A})$ nın ortogonal tümleyeni
$\det(\mathbf{A})$ (veya $ \mathbf{A} $)	:	\mathbf{A} matrisinin determinantı
\mathbf{d}_i	:	Sapma veya düzeltilmiş ortalama vektörü
E	:	Beklenen değer
$\exp\langle \cdot \rangle$:	Üstel
GLR	:	Genelleştirilmiş Olabilirlik Oranı
H_0	:	Sıfır hipotezi
H_a	:	Alternatif hipotez
\mathbf{H}_f	:	Hessian matris
HKO	:	Hata Kareler Ortalaması
\mathbf{I}_m	:	m -mertebeli birim matris
$\text{iz}(\mathbf{A})$:	\mathbf{A} matrisinin izi

J	:	Jakobiyen matris
$\mathbf{j} = [1, 1, \dots, 1]'$:	Uzayın eksenleriyle eşit açı yapan vektör
KKO	:	Kısıtlı Kareler Ortalaması
kov	:	Kovaryans
$köşeg(\cdot)$:	Köşegen matrisi
$\kappa(\mathbf{A})$:	Şart sayısı
$L(\cdot)$:	Olabilirlik fonksiyonu
$L_{(\cdot)}$:	Uzunluk
LR	:	Olabilirlik Oran Testi
M_{ij}	:	a_{ij} nin minörü
$N(\mathbf{A})$:	\mathbf{A} nın sıfır uzayı (ya da çekirdeği)
$N_n(\mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I})$:	$\mathbf{X}\boldsymbol{\beta}$ ortalamalı, σ^2 varyanslı, n - değişkenli çoklu normal dağılım
$OLSE$:	Alışılmış En Küçük Kareler Tahmin Edicisi
OO	:	Olabilirlik Oran Testi
$ÖB(\cdot)$:	Öz-Bağımlı Model
P	:	Olasılık
$\text{rank}(\mathbf{A})$:	\mathbf{A} matrisinin rankı
$R(\mathbf{A})$:	\mathbf{A} nın satır uzayı
$\mathbb{R}^{m \times n}$:	Tüm reel $m \times n$ matrislerinin kümesi
\mathbb{R}^m	:	$m \times 1$ tipindeki matrislerin kümesi
r	:	Korelasyon Katsayısı
\mathbf{R}	:	Örneklem korelasyonu
R^2	:	Determinasyon Katsayısı

$ S $: Genelleştirilmiş varyans
S_n	: Örneklem varyans-kovaryansı
S_E^2	: Tahmin Edilen Hata Varyansı
SSE	: Hata Kareler Toplamı
Sup	: Supremum
T	: Lineer dönüşüm operatörü
\mathbf{u}^\perp	: \mathbf{u} ya ortogonal olan tüm vektörlerin kümesi
var	: Varyans
$\mathbf{1}$: Elemanlarının tümü 1 lerden oluşan sütun vektörü
$(1 - \alpha)$: Güven Katsayısı
$\alpha_{kümelenme}$: Kümelenme hata oranı
$\alpha_{bireysel}$: Bireysel hata oranı
$\hat{\beta}$: β nın alışılmış en küçük kareler tahmini
$\tilde{\beta}$: β nın en çok olabilirlik tahmini
$\hat{\beta}_c$: β nın kısıtlanmış en küçük kareler tahmini
$\tilde{\beta}_c$: β nın kısıtlanmış en çok olabilirlik tahmini
$\hat{\beta}_M$: β nın karma regresyon tahmini
$\hat{\beta}_f$: β nın uygulanabilir karma regresyon tahmini
$\bar{\mathbf{x}}$: Örneklem ortalaması
$\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}$: \mathbf{X}' nün bir g - tersi
ε	: Hata vektörü
μ	: Ortalama
$\nabla_{\mathbf{x}}$: \mathbf{x} e göre gradient
$\pi(\cdot)$: İstatistiksel testin güç fonksiyonu
$\langle \cdot, \cdot \rangle$: İç çarpım

$\|\cdot\|_2$: 2-normu veya Öklidyen normu

1. GİRİŞ

Bu tez beş bölümden oluşur. İkinci bölümde diğer bölümlere yardımcı olması açısından matrisler, vektörler ve istatistikle ilgili bazı temel bilgilerden söz edeceğiz. Öncelikle matris ve vektör cebirini tanıtacağız ve bunlarla ilgili bir takım tanımlardan ve özelliklerden bahsedeceğiz. Sonra en küçük kareler yöntemine geçebilmek için optimizasyon konusunu ifade edeceğiz. Lineer modellerde lineer kısıtlamalar altında parametre tahminleri ve hipotez testi konusunu daha iyi anlayabilmek için istatistiksel gerçekler konu başlığı altında bazı tanımları vereceğiz.

Üçüncü bölümde istatistikte lineer modeller ve lineer modellerin geometrisini ifade edip bunu bir örnekle ve görsellerle destekleyeceğiz. Dördüncü bölüme hazırlık olması için lineer modellerle ilgili önemli bilgiler vereceğiz. Ardından lineer olmayan regresyon modellerine geçeceğiz. Kuadratik eğrilerin ve yüzeylerin en küçük kareler uyumundan kısaca bahsedeceğiz. Böylece lineer ve lineer olmayan modellerle bölümü sonlandıracağız.

Dördüncü bölümün ilk kısmında lineer modellerde lineer kısıtlamalar altında parametre tahminleri ve hipotez testini ele alacağız. Lineer modellerde olabilirlik fonksiyonu vasıtasıyla kısıtlanmış ve kısıtlanmamış parametre tahminlerini ifade edeceğiz. Sonra başka bir yöntemle kısıtlanmış parametre tahminlerini bulup bir önceki tahminle bunu karşılaştıracacağız. Dördüncü bölümün ikinci kısmında önce örnekleme geometrisini vereceğiz. Bu konuyla ilgili birkaç örnek yapacağız. Sonra da korelasyon katsayısına farklı geometrik yorumlar getireceğiz. Özellikle bu geometrik yorumlar, Pearson korelasyon katsayısı üzerinde yoğunlaşmıştır.

Beşinci bölümde sonuç ve öneriler ile tezi bitireceğiz.

2. GENEL BİLGİLER

2.1. Matris ve Vektör Cebiri

Bu bölümde matrislerle ve vektörlerle ilgili bazı kavramlara ve özelliklere değineceğiz.

2.1.1. Tanımlar ve Temel İşlemler

Bir \mathbf{A} $m \times n$ reel matrisi m -sayıda satır ve n -sayıda sütunda düzenlenen mn -tane reel sayının

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \quad (2.1)$$

biçimindeki bir dikdörtgen düzenidir. Bir matris genellikle, satır indisi denilen $1 \leq i \leq m$ ve sütun indisi denilen $1 \leq j \leq n$ ile, $\mathbf{A} = [a_{ij}]$ olarak gösterilir. Tüm reel elemanlı $m \times n$ matrislerinin kümesi $\mathbb{R}^{m \times n}$ ile gösterilir. $m = n$ olduğunda bu matrise m -mertebeli bir kare matris denir. \mathbf{A} nın elemanlarının tümü sıfır olduğunda \mathbf{A} ya sıfır matris denir. \mathbf{A} nın i -yinci satırı a_{i*} ve \mathbf{A} nın j -yinci sütunu a_{*j} ile gösterilir. Bu nedenle, \mathbf{A} matrisi:

$$\mathbf{A} = [a_{ij}] = [a_{*1} \quad a_{*2} \quad \cdots \quad a_{*n}] = \begin{bmatrix} a_{1*} \\ a_{2*} \\ \vdots \\ a_{m*} \end{bmatrix} \quad (2.2)$$

olarak gösterilebilir. $\{a_{11}, a_{22}, \dots, a_{mm}\}$ elemanlarının kümesine \mathbf{A} nın ana veya asıl köşegeni denir.

❖ Matrisler Üzerinde İşlemler :

$\mathbf{A}, \mathbf{B}, \mathbf{C}$; $\mathbb{R}^{m \times n}$ deki matrisler, p, q ; skalarlar, $\mathbf{y} \in \mathbb{R}^m$ ve $\mathbf{x} \in \mathbb{R}^n$, (yani $m \times 1$ ve $n \times 1$ matrisler) olsun. \mathbf{y} ve \mathbf{x} sütun matrislerine sütun vektörleri olarak bakacağız.

a) **Bir Matrisinin Transpozesi :** Eğer $\mathbf{A} \in \mathbb{R}^{m \times n}$ ise, bu takdirde \mathbf{A} nın sütunlarını ve satırlarını değiştirmek suretiyle elde edilen $\mathbf{A}' \in \mathbb{R}^{n \times m}$ matrisine \mathbf{A} matrisinin transpozesi denir. $(\mathbf{A}')' = \mathbf{A}$ olduğu gerçektelebilir. Elemanlarının tümü sıfır olan matrise sıfır matris denir ve bu $\mathbf{0}$ ile gösterilir.

b) **Toplam/Fark :** $c_{ij} = a_{ij} \pm b_{ij}$ olmak üzere, $\mathbf{C} = \mathbf{A} \pm \mathbf{B}$ ye \mathbf{A} ve \mathbf{B} matrislerinin elemanlara bağlı toplamı/farkı denir. $(\mathbf{A} \pm \mathbf{B})' = \mathbf{A}' \pm \mathbf{B}'$ olduğu gerçektelebilir.

c) **Skalarla Çarpma :** $c_{ij} = pa_{ij}$ olmak üzere, $\mathbf{C} = p\mathbf{A}$ ya \mathbf{A} nın p skalarıyla çarpımı denir.

d) **İki Matrisin Çarpımı :** $\mathbf{A} \in \mathbb{R}^{m \times n}$ ve $\mathbf{B} \in \mathbb{R}^{n \times r}$ olsun. Bu takdirde, $\mathbf{C} = \mathbf{A}\mathbf{B} \in \mathbb{R}^{m \times r}$ çarpımı üç eşdeğer yöntemle tanımlanabilir.

I. **İç çarpım :** c_{ij} elemanı, \mathbf{A} nın i -yinci satırı ile \mathbf{B} nin j -yinci sütununun iç çarpımıdır. Şöyle ki,

$$c_{ij} = \sum_{k=1}^n a_{ik}b_{kj}, \quad 1 \leq i \leq m, \quad 1 \leq j \leq r \quad (2.3)$$

dir.

II. **Bir skalar \times bir vektör :** \mathbf{C} nin j -yinci c_{*j} sütunu, \mathbf{B} nin j -yinci sütununun elemanlarını katsayılar olarak kullanan, \mathbf{A} nın sütunlarının lineer kombinasyonudur, yani

$$c_{*j} = \sum_{i=1}^n a_{*i}b_{ij}, \quad 1 \leq j \leq r \quad (2.4)$$

dir.

III. Dış çarpım : \mathbf{C} çarpım matrisi, \mathbf{A} nın i -yinci sütunu ve \mathbf{B} nin j -yinci satırıyla elde edilen n -tane dış çarpım matrisinin toplamı olarak da elde edilebilir.

Yani

$$\mathbf{C} = \sum_{j=1}^n a_{*j} b_{j*} \quad (2.5)$$

dır. $(\mathbf{AB})' = \mathbf{B}'\mathbf{A}'$ olduğu gerçekleştirilebilir.

Bu işlemler aşağıdaki özellikleri kolayca gerçekleştirir.

- i. $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$
- ii. $p\mathbf{A} = \mathbf{A}p$
- iii. $\mathbf{AB} \neq \mathbf{BA}$ (genel olarak)
- iv. $\mathbf{A}(\mathbf{x} + \mathbf{y}) = \mathbf{Ax} + \mathbf{Ay}$
- v. $(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C})$
- vi. $(p + q)\mathbf{A} = p\mathbf{A} + q\mathbf{A}$
- vii. $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$
- viii. $\mathbf{A} + (-\mathbf{A}) = \mathbf{0}$
- ix. $p(\mathbf{A} + \mathbf{B}) = p\mathbf{A} + p\mathbf{B}$

Ayrıca $\mathbf{x}, \mathbf{y} \in \mathbb{R}^m$ olmak üzere, $\mathbf{x}'\mathbf{y} = \mathbf{y}'\mathbf{x}$ çarpımına \mathbf{x} ve \mathbf{y} vektörlerinin iç çarpımı (veya skalar çarpımı) denir. Bu çarpımın bir skalar olduğu görülür.

$\|\mathbf{x}\| = \sqrt{\mathbf{x}_1^2 + \mathbf{x}_2^2 + \dots + \mathbf{x}_m^2} = \sqrt{\|\mathbf{x}'\mathbf{x}\|}$ skalarına \mathbf{x} vektörünün normu denir.

e) Matris-Vektör Çarpımı : \mathbf{y} nin i -yinci elemanı, \mathbf{A} nın i -yinci satırının \mathbf{x} vektörü (sütun matrisi) ile iç çarpımı olmak üzere, $\mathbf{y} = \mathbf{Ax}$,

$$y_i = \sum_{j=1}^n a_{ij}x_j, \quad 1 \leq i \leq m \quad (2.6)$$

olarak tanımlanır. Bir başka yolla \mathbf{y} ;

$$\mathbf{y} = \sum_{j=1}^m a_{*j} x_j \quad (2.7)$$

olarak \mathbf{A} nın sütunlarının bir lineer kombinasyonu olarak da tanımlanabilir. Burada \mathbf{x} vektörünün elemanları, lineer kombinasyonunun katsayıları olarak kullanılır.

❖ **Özel Matrisler, Diğer İşlemler ve Özellikler :**

a) **Köşegen Matris** : Eğer $i \neq j$ için, $a_{ij} = 0$ ise, \mathbf{A} matrisine bir köşegen matris denecektir. \mathbf{A} köşegen matrisi

$$\mathbf{A} = \text{köşeg}(a_{11}, a_{22}, \dots, a_{mm}) \quad (2.8)$$

ile gösterilir.

b) **Birim Matris** : \mathbf{I}_m , köşegen elemanları 1 e ve köşegen dışındaki tüm elemanları 0 a eşit olan m -mertebeli bir köşegen matristir. \mathbf{I}_m ye m -mertebeli bir birim matris denir. Bu nedenle

$$\mathbf{I}_m = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix} \quad (2.9)$$

dir.

c) **Üst/Alt Üçgen Matris** : Eğer $i > j$ için $a_{ij} = 0$ ise, \mathbf{A} bir üst üçgen matristir. Benzer şekilde, eğer $i < j$ için $a_{ij} = 0$ ise, \mathbf{A} bir alt üçgen matristir.

d) **Bir Matrisin İzi** : Eğer $\mathbf{A} \in \mathbb{R}^{m \times m}$ ise, $\text{iz}(\mathbf{A})$ ile gösterilen \mathbf{A} nın izi, onun köşegen elemanlarının toplamıyla tanımlanan bir sayı (skalar) dır. Bu nedenle,

$$\text{iz}(\mathbf{A}) = \sum_{i=1}^m a_{ii} \quad (2.10)$$

dir.

- i. $\text{iz}(\mathbf{A}) = \text{iz}(\mathbf{A}')$
- ii. $\text{iz}(\mathbf{A} + \mathbf{B}) = \text{iz}(\mathbf{A}) + \text{iz}(\mathbf{B})$
- iii. $\text{iz}(\alpha \mathbf{A}) = \alpha \text{iz}(\mathbf{A})$
- iv. $\text{iz}(\mathbf{AB}) = \text{iz}(\mathbf{BA})$
- v. $\text{iz}(\mathbf{ABC}) = \text{iz}(\mathbf{BCA}) = \text{iz}(\mathbf{CAB})$

olduğu gerçektelebilir.

e) Bir Matrisin Determinantı : $\mathbf{A} \in \mathbb{R}^{m \times m}$ olsun. \mathbf{A} nın, $\det(\mathbf{A})$ (veya $|\mathbf{A}|$) ile gösterilen determinantı, aşağıdaki gibi tekrarlı şekilde tanımlanan bir skaldardır. Herhangi bir sabit i -sattır indisi için,

$$\det(\mathbf{A}) = \sum_{j=1}^m a_{ij} A_{ij} \quad (2.11)$$

dir. Burada $A_{ij} = (-1)^{i+j} M_{ij}$ ile verilen A_{ij} ye a_{ij} elemanın eşçarpmanı denir. a_{ij} nin minörü denilen M_{ij} , \mathbf{A} nın i -yinci satırının ve j -yinci sütununun silinmesiyle elde edilen $(m-1) \times (m-1)$ matrisinin determinantıdır.

- i. $\det(\mathbf{A}') = \det(\mathbf{A})$
- ii. $\det(\mathbf{AB}) = \det(\mathbf{A}) \det(\mathbf{B})$

olduğu gerçektelebilir.

f) Ek Matris : Bir $\mathbf{A} \in \mathbb{R}^{m \times m}$ kare matrisi için

$$\mathbf{A}^* = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1m} \\ A_{21} & A_{22} & \cdots & A_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m1} & A_{m2} & \cdots & A_{mm} \end{bmatrix} \quad (2.12)$$

matrisine \mathbf{A} nın eki denir.

g) Bir Matrisin Rankı : $\mathbf{A} \in \mathbb{R}^{m \times n}$ olsun. \mathbf{A} nın lineer bağımsız sütunlarının (satırlarının) sayısına \mathbf{A} nın sütun (satur) rankı denir. Bu nedenle \mathbf{A} nın sütun (satur) rankı n (m) ye eşit ya da ondan küçüktür. Verilen herhangi bir \mathbf{A} matrisi için, onun sütun ve satur rankları daima eşittir ve bu ortak tamsayı değerine \mathbf{A} nın rankı denir ve $\text{rank}(\mathbf{A})$ ile gösterilir, yani,

$$0 \leq \text{rank}(\mathbf{A}) \leq \min \{m, n\} \quad (2.13)$$

dir. Eğer $\text{rank}(\mathbf{A}) = \min \{m, n\}$ ise, bu takdirde \mathbf{A} ya tam ranklı matris denecektir, aksi takdirde eksik-ranklı matris adını alacaktır.

Şimdi rankın birkaç önemli özelliğini sıralayalım.

- i. $\text{rank}(\mathbf{A}') = \text{rank}(\mathbf{A})$
- ii. $\text{rank}(\mathbf{A} + \mathbf{B}) \leq \text{rank}(\mathbf{A}) + \text{rank}(\mathbf{B})$
- iii. $\text{rank}(\mathbf{A} - \mathbf{B}) \geq |\text{rank}(\mathbf{A}) - \text{rank}(\mathbf{B})|$
- iv. Eğer $\mathbf{A} \in \mathbb{R}^{m \times n}$ ve $\mathbf{B} \in \mathbb{R}^{n \times r}$ ise, bu takdirde $\text{rank}(\mathbf{AB}) \leq \min \{\text{rank}(\mathbf{A}), \text{rank}(\mathbf{B})\}$ dir.
- v. Eğer $\mathbf{x}, \mathbf{y} \in \mathbb{R}^m$ ise, bu takdirde \mathbf{xy}' iç çarpım matrisinin rankı 1 dir, yani $\text{rank}(\mathbf{xy}') = 1$ dir.

h) Bir Matrisin Singüler (Tekil) Olmaması : Eğer aşağıdaki şartlar gerçekleşirse, bir $\mathbf{A} \in \mathbb{R}^{m \times m}$ kare matrisine singüler (tekil) değildir denir.

- i. $\det(\mathbf{A}) \neq 0$
- ii. $\text{rank}(\mathbf{A}) = m$
- iii. \mathbf{A} nın tüm sütunları (satırları) lineer bağımsızdır.

i) Bir Matrisin Tersi : $\mathbf{A} \in \mathbb{R}^{m \times m}$ tekil olmasın. Bu takdirde

$$\mathbf{AA}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}_m \quad (2.14)$$

olacak şekilde $\mathbf{A}^{-1} \in \mathbb{R}^{m \times m}$ ile gösterilen bir çarpımsal ters vardır.

- i. $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$
- ii. $\mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|} \mathbf{A}^*$

- iii. $(\mathbf{AB})^{-1} = (\mathbf{B})^{-1}(\mathbf{A})^{-1}$
- iv. $(\mathbf{A}')^{-1} = (\mathbf{A}^{-1})' = \mathbf{A}'^{-1}$
- v. $\text{iz}(\mathbf{A}^{-1}\mathbf{BA}) = \text{iz}(\mathbf{B})$
- vi. $\det(\mathbf{A}^{-1}) = \frac{1}{\det(\mathbf{A})}$

olduğu gerçektebilir (Meyer, 2000).

j) Sherman-Morrison-Woodbury Formülü : $\mathbf{c}, \mathbf{d} \in \mathbb{R}^m$ olsun. Bu takdirde \mathbf{cd}' bir “rank–bir (1 ranklı) dış-çarpım” matrisidir. Bir dış çarpım matrisini tekil olmayan bir matrise eklemeye “rank–bir sarsımı” denir. \mathbf{A} ve $(\mathbf{A} + \mathbf{cd}')$ nün tersi arasında ilginç bir ilişki vardır.

I. Bir Özdeşlik : Aşağıdaki özdeşlik kolayca gerçektir:

$$(\mathbf{I}_m + \mathbf{cd}')^{-1} = \mathbf{I}_m - \frac{\mathbf{cd}'}{1 + \mathbf{d}'\mathbf{c}}. \quad (2.15)$$

II. Sherman-Morrison Formülü :

$$(\mathbf{A} + \mathbf{cd}')^{-1} = \mathbf{A}^{-1} - \frac{\mathbf{A}^{-1}\mathbf{cd}'\mathbf{A}^{-1}}{1 + \mathbf{d}'\mathbf{A}^{-1}\mathbf{c}}. \quad (2.16)$$

III. Woodbury Genişlemesi : $\mathbf{C} \in \mathbb{R}^{m \times k}$ ve $\mathbf{D} \in \mathbb{R}^{m \times k}$, k -ranklı matrisler olsun. Bu takdirde,

$$(\mathbf{A} + \mathbf{CD}')^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{C}[\mathbf{I}_k + \mathbf{D}'\mathbf{A}^{-1}\mathbf{C}]^{-1}\mathbf{D}'\mathbf{A}^{-1} \quad (2.17)$$

dir. Benzer şekilde eğer, \mathbf{A} ve \mathbf{B} tersi alınabilir (tersinir) matrisler ise, bu takdirde

$$(\mathbf{A} + \mathbf{CBD}')^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{C}[\mathbf{B}^{-1} + \mathbf{D}'\mathbf{A}^{-1}\mathbf{C}]^{-1}\mathbf{D}'\mathbf{A}^{-1} \quad (2.18)$$

dir (Basilevsky, 1983; Meyer, 2000).

k) Başka Bir Yararlı Matris Özdeşliği :

$$[\mathbf{A}'\mathbf{B}^{-1}\mathbf{A} + \mathbf{D}^{-1}]\mathbf{A}'\mathbf{B}^{-1} = \mathbf{D}\mathbf{A}'[\mathbf{B} + \mathbf{ADA}']^{-1} \quad (2.19)$$

dir (Basilevsky, 1983).

l) Bir Matrisin Genelleştirilmiş Tersi : Bir $\mathbf{A} \in \mathbb{R}^{m \times n}$ matrisi için $\mathbf{AA}^{-}\mathbf{A} = \mathbf{A}$ özelliğini sağlayan $\mathbf{A}^{-} \in \mathbb{R}^{n \times m}$ matrisine \mathbf{A} nın Genelleştirilmiş tersi (g-tersi) denir.

m) **Moore-Penrose Genelleştirilmiş Ters** : $\mathbf{A} \in \mathbb{R}^{m \times n}$ olsun. Aşağıdaki şartları sağlayan $\mathbf{A}^+ \in \mathbb{R}^{n \times m}$ matrisine, \mathbf{A} nın Moore-Penrose genelleştirilmiş tersi denir.

- i. $\mathbf{AA}^+\mathbf{A} = \mathbf{A}$
- ii. $\mathbf{A}^+\mathbf{AA}^+ = \mathbf{A}^+$
- iii. $(\mathbf{AA}^+)' = \mathbf{AA}^+$ (\mathbf{AA}^+ bir simetrik matristir.)
- iv. $(\mathbf{A}^+\mathbf{A})' = \mathbf{A}^+\mathbf{A}$ ($\mathbf{A}^+\mathbf{A}$ bir simetrik matristir.)

\mathbf{A}^- matrisi birçok sayıda vardır. Fakat \mathbf{A}^+ matrisi tektir. Eğer \mathbf{A} tam sütun ranklı ise, $\mathbf{A}^- = \mathbf{A}^+ = (\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'$ ve \mathbf{A} tam satır ranklı ise, $\mathbf{A}^- = \mathbf{A}^+ = \mathbf{A}'(\mathbf{AA}')^{-1}$ dir.

n) **Parçalanmış Matrisler** : Bazen bir matrisi altmatrislere ayırmak uygundur. Örneğin, bir \mathbf{A} matrisinin uygun boyutlu dört (kare veya dikdörtgen) altmatrise parçalanması, sembolik olarak aşağıdaki gibi gösterilebilir:

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix}.$$

Bunu açıklamak için \mathbf{A} , 4×5 matrisi, yani

$$\mathbf{A} = \begin{pmatrix} 7 & 2 & 5 & \vdots & 8 & 4 \\ -3 & 4 & 0 & \vdots & 2 & 7 \\ \dots & \dots & \dots & \vdots & \dots & \dots \\ 9 & 3 & 6 & \vdots & 5 & -2 \\ 3 & 1 & 2 & \vdots & 1 & 6 \end{pmatrix} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix},$$

olarak parçalansın. Burada

$$\mathbf{A}_{11} = \begin{pmatrix} 7 & 2 & 5 \\ -3 & 4 & 0 \end{pmatrix}, \mathbf{A}_{12} = \begin{pmatrix} 8 & 4 \\ 2 & 7 \end{pmatrix}, \mathbf{A}_{21} = \begin{pmatrix} 9 & 3 & 6 \\ 3 & 1 & 2 \end{pmatrix}, \mathbf{A}_{22} = \begin{pmatrix} 5 & -2 \\ 1 & 6 \end{pmatrix}$$

dır.

Eğer iki \mathbf{A} ve \mathbf{B} matrisi çarpım için uyumlu ise ve \mathbf{A} ve \mathbf{B} nin altmatrisleri uygun bir şekilde çarpıma uyumlu olacak şekilde parçalanırsa, bu takdirde \mathbf{AB} çarpımı, sanki tek elemanlar gibi alışlagelen satırın sütunla çarpımı kalıbını kullanarak altmatrislerle bulunabilir.

Örneğin,

$$\begin{aligned}\mathbf{AB} &= \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix} \begin{pmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{A}_{11}\mathbf{B}_{11} + \mathbf{A}_{12}\mathbf{B}_{21} & \mathbf{A}_{11}\mathbf{B}_{12} + \mathbf{A}_{12}\mathbf{B}_{22} \\ \mathbf{A}_{21}\mathbf{B}_{11} + \mathbf{A}_{22}\mathbf{B}_{21} & \mathbf{A}_{21}\mathbf{B}_{12} + \mathbf{A}_{22}\mathbf{B}_{22} \end{pmatrix}\end{aligned}\quad (2.20)$$

dir. Eğer \mathbf{B} yerine elemanların iki kümesine parçalanmış bir \mathbf{b} vektörü konursa ve \mathbf{A} buna bağlı olarak sütunların iki kümesine ayrılırsa, bu takdirde (2.20)

$$\mathbf{Ab} = (\mathbf{A}_1, \mathbf{A}_2) \begin{pmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \end{pmatrix} = \mathbf{A}_1\mathbf{b}_1 + \mathbf{A}_2\mathbf{b}_2 \quad (2.21)$$

olur. Burada \mathbf{A}_1 in sütun sayısı \mathbf{b}_1 in eleman sayısına eşittir ve aynı şekilde \mathbf{A}_2 ve \mathbf{b}_2 de çarpıma uyumludur. $\mathbf{A} = (\mathbf{A}_1, \mathbf{A}_2)$ deki parçalanmanın bir virgül ile gösterildiğine dikkat edelim.

(2.21) deki parçalanmış çarpım \mathbf{A} nın ayrı ayrı sütunlarına ve \mathbf{b} nin ayrı ayrı satırlarına genişletilebilir:

$$\mathbf{Ab} = (\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_p) \begin{pmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \\ \vdots \\ \mathbf{b}_p \end{pmatrix} = \mathbf{A}_1\mathbf{b}_1 + \mathbf{A}_2\mathbf{b}_2 + \dots + \mathbf{A}_p\mathbf{b}_p \quad (2.22)$$

olur.

Bu nedenle \mathbf{Ab} , \mathbf{A} nın sütunlarının bir lineer kombinasyonu olarak ifade edilebilir ki burada kombinasyon katsayıları \mathbf{b} nin elemanlarıdır.

Bir satır vektör ve bir matrisin $\mathbf{a}'\mathbf{B}$ çarpımı, \mathbf{B} nin satırlarının bir lineer kombinasyonudur ki oradaki katsayılar \mathbf{a}' nün elemanlarıdır.

$$\mathbf{a}'\mathbf{B} = (\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n) \begin{pmatrix} \mathbf{B}_1' \\ \mathbf{B}_2' \\ \vdots \\ \mathbf{B}_n' \end{pmatrix} = \mathbf{a}_1\mathbf{B}_1' + \mathbf{a}_2\mathbf{B}_2' + \dots + \mathbf{a}_n\mathbf{B}_n' \quad (2.23)$$

dür. Son olarak, bir \mathbf{A} matrisi $\mathbf{A} = (\mathbf{A}_1, \mathbf{A}_2)$ olarak parçalanırsa, bu takdirde

$$\mathbf{A}' = (\mathbf{A}_1, \mathbf{A}_2)' = \begin{pmatrix} \mathbf{A}_1' \\ \mathbf{A}_2' \end{pmatrix} \quad (2.24)$$

olacağını belirtelim (Rencher ve Schaalje, 2007).

2.1.2. Vektör Uzayları ve Altuzaylar

Aşağıdaki şartları sağlayan boş olmayan bir \mathcal{S} kümesine bir vektör uzayı denir.

- i. \mathcal{S} deki herhangi \mathbf{x}, \mathbf{y} için $\mathbf{x} + \mathbf{y}$ tanımlıdır ve \mathcal{S} dedir. Ayrıca,
 $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$ (değişme özelliği),
 $\mathbf{x} + (\mathbf{y} + \mathbf{z}) = (\mathbf{x} + \mathbf{y}) + \mathbf{z}$ (birleşme özelliği) dir.
- ii. \mathcal{S} de her \mathbf{x} için $\mathbf{x} + \mathbf{0} = \mathbf{x}$ olacak şekilde $\mathbf{0}$ ile gösterilen bir eleman mevcuttur.
- iii. \mathcal{S} deki herhangi bir \mathbf{x} için, $\mathbf{x} + \mathbf{y} = \mathbf{0}$ olacak şekilde \mathcal{S} de bir \mathbf{y} elemanı mevcuttur.
- iv. \mathcal{S} deki herhangi bir \mathbf{x} için ve herhangi bir c reel sayısı için $c\mathbf{x}$ tanımlıdır ve \mathcal{S} dedir; herhangi bir \mathbf{x} için $1\mathbf{x} = \mathbf{x}$ dir.
- v. \mathcal{S} deki herhangi bir $\mathbf{x}_1, \mathbf{x}_2$ için ve c_1, c_2 reel sayıları için
 $c_1(\mathbf{x}_1 + \mathbf{x}_2) = c_1\mathbf{x}_1 + c_1\mathbf{x}_2$, $(c_1 + c_2)\mathbf{x}_1 = c_1\mathbf{x}_1 + c_2\mathbf{x}_1$ ve $c_1(c_2\mathbf{x}_1) = (c_1c_2)\mathbf{x}_1$ dir.

\mathcal{S} deki elemanlara vektörler denir. Eğer \mathbf{x}, \mathbf{y} vektör ise, bu takdirde $\mathbf{x} + \mathbf{y}$ toplamını alma işlemine vektör toplamı olarak bakılır. (ii) deki vektöre sıfır vektör denir. (iv) deki işleme skalarla çarpım denir. Bir vektör uzayı herhangi bir cisme göre tanımlanır. Amacımız için yeterli olacak olan reel sayılar cismini cisim olarak aldık. n - mertebeli sütun vektörlerin kümesi (veya $n \times 1$ matrisler) bir vektör uzayıdır. Bu nedenle n - mertebeli satır vektörlerin kümesi de bir vektör uzayıdır. Çoğu zaman göz önüne aldığımız vektör uzayları bu iki vektör uzayıdır. \mathbb{R} reel sayıların kümesi olmak üzere, \mathbb{R}^n ; $\underbrace{\mathbb{R} \times \mathbb{R} \times \dots \times \mathbb{R}}_{n\text{-tane}}$ kümesi olsun. \mathbb{R}^n in elemanlarını, verilen bir durumda hangisinin uygun olduğuna bağlı olarak ya sütun vektörleri ya da satır vektörleri olarak yazacağız.

Eğer \mathcal{S}, T vektör uzayları ve $\mathcal{S} \subset T$ ise, bu takdirde \mathcal{S} ye T nin bir alt uzayı denir. \mathbb{R}^3 ün tüm mümkün altuzaylarını tanımlayalım. Açık olarak, \mathbb{R}^3 bir vektör uzayıdır ve bu nedenle sadece sıfır vektöründen ibaret olan yani tüm sıfırların vektörü de bir vektör uzayıdır. c_1, c_2, c_3 reel sayılar olsun.

$$c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + c_3\mathbf{x}_3 = \mathbf{0}$$

eşitliğini sağlayan tüm $\mathbf{x} \in \mathbb{R}^3$ vektörlerinin kümesi \mathbb{R}^3 ün bir alt uzayıdır (Burada $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$; \mathbf{x} in koordinatlarıdır.). Geometrik olarak bu küme, orijinden geçen bir düzlemi temsil eder. Orijinden geçen iki farklı düzlemin arakesiti orijinden geçen bir düz doğrudur ve o da bir düz doğrudur. Bunlar sadece \mathbb{R}^3 ün mümkün olabilir altuzaylarıdır (Bapat, 1993).

a) Taban ve Boyut

$\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m$ vektörlerinin lineer gereni (veya $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m$ vektörleri tarafından gerilen uzay), c_1, c_2, \dots, c_m reel sayılar olmak üzere, tüm $c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \dots + c_m\mathbf{x}_m$ lineer kombinasyonlarının kümesi olarak tanımlanır. Lineer geren bir altuzaydır. Bu; tanımdan görülür. Eğer en az bir c_i sıfırdan farklı ve $c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \dots + c_m\mathbf{x}_m = \mathbf{0}$ olacak şekilde c_1, c_2, \dots, c_m reel sayıları varsa, $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m$ vektörlerinin bir kümesine lineer bağımlıdır denir. Eğer bir küme lineer bağımlı değilse lineer bağımsızdır. Bir vektör uzayının gerindeki bağımsız vektörlere vektör uzayının tabanı denir.

b) Ortonormal Taban : $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m\}$, bir \mathcal{S} vektör uzayının tabanı olsun. Eğer

$$\mathbf{x}_i' \mathbf{x}_j = \mathbf{x}_j' \mathbf{x}_i = \begin{cases} 1, & \forall i = j \text{ için} \\ 0, & i \neq j \text{ için} \end{cases} \quad (2.25)$$

ise, $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m\}$ kümesine \mathcal{S} için bir ortonormal taban denir.

c) Bir Matrisin Sütun Uzayı : Bir $\mathbf{A} \in \mathbb{R}^{m \times n}$ matrisinin sütun uzayı \mathbf{A} nın sütunları tarafından gerilen vektör uzayıdır. Bu uzayı $C(\mathbf{A})$ ile göstereceğiz.

d) Bir Matrisin Satır Uzayı : Bir $\mathbf{A} \in \mathbb{R}^{m \times n}$ matrisinin satır uzayı \mathbf{A} nın satırları tarafından gerilen vektör uzayıdır. Bu uzayı $R(\mathbf{A})$ ile göstereceğiz.

➤ Buna göre bir \mathbf{A} matrisinin rankı, satır veya sütun uzaylarının boyutuna eşittir.

e) Bir Matrisin Sütun ve Sıfır Uzayı : $V \subseteq \mathbb{R}^n$ ve $W \subseteq \mathbb{R}^m$ ve $\mathbf{A} \in \mathbb{R}^{m \times n}$ olacak şekilde, $\mathbf{A} : V \rightarrow W$ olsun. \mathbf{A} nın sütun uzayını $C(\mathbf{A})$ ile göstermiştik. Buna göre

$$C(\mathbf{A}) = \{\mathbf{y} \in \mathbb{R}^m \mid \text{her } \mathbf{x} \in V \text{ için, } \mathbf{y} = \mathbf{A}\mathbf{x}\} \subseteq W \quad (2.26)$$

olarak tanımlanır. Bu uzay görüldüğü gibi \mathbf{A} nın sütunlarının lineer kombinasyonlarının kümesidir (bkz. Şekil 2.1). Yani $C(\mathbf{A})$, \mathbf{A} nın sütunları tarafından üretilen lineer vektör uzayıdır ve bu nedenle $C(\mathbf{A})$ ya \mathbf{A} altında V nin görüntüsü (resmi) veya basitçe \mathbf{A} nın görüntü uzayı denir. Benzer şekilde \mathbf{A} nın satır uzayı da denilen \mathbf{A}' nün sütun uzayı,

$$C(\mathbf{A}') = \{\mathbf{x} \in \mathbb{R}^n \mid \text{her } \mathbf{y} \in \mathbb{R}^m \text{ için, } \mathbf{x} = \mathbf{A}'\mathbf{y}\} \subseteq V \quad (2.27)$$

olarak tanımlayabiliriz.

$N(\mathbf{A})$ ile gösterilen \mathbf{A} nın sıfır uzayı

$$\mathcal{N}(\mathbf{A}) = \{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} = 0 \} \subseteq V \quad (2.28)$$

olarak tanımlanır. Bu nedenle $\mathcal{N}(\mathbf{A})$; $\mathbf{A}\mathbf{x} = 0$ sisteminin tüm çözümlerinden oluşur. $\mathcal{N}(\mathbf{A})$ nın bir vektör uzayı olduğu gerçekleştirilebilir ve $\mathcal{N}(\mathbf{A})$ ya \mathbf{A} nın çekirdeği de denir. Benzer şekilde \mathbf{A}' nün sıfır uzayını

$$\mathcal{N}(\mathbf{A}') = \{ \mathbf{y} \in \mathbb{R}^m \mid \mathbf{A}'\mathbf{y} = 0 \} \subseteq W \quad (2.29)$$

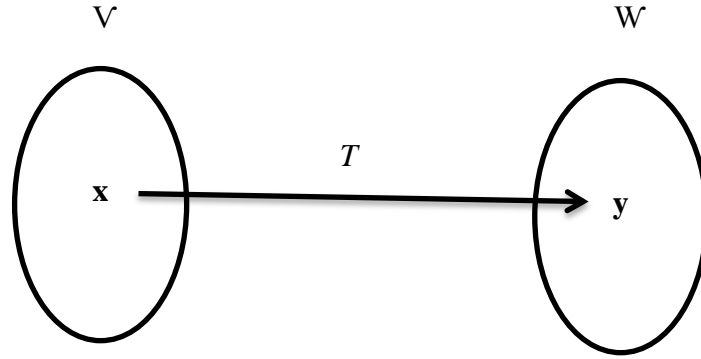
olarak tanımlayabiliriz.

Gerçekten, verilen bir $\mathbf{A} \in \mathbb{R}^{m \times n}$ matrisi için \mathbf{A} matrisi ile ilgili $C(\mathbf{A}), C(\mathbf{A}'), \mathcal{N}(\mathbf{A})$ ve $\mathcal{N}(\mathbf{A}')$ olarak ifade edilen dört vektör uzayı vardır. Aşağıda bu dört vektör uzayı arasındaki ilişkileri ve birkaç ana özelliği yazacağız.

- i. $\mathcal{N}(\mathbf{A}) = C^\perp(\mathbf{A}')$, yani \mathbf{A} nın sıfır uzayı, \mathbf{A}' nün sütun uzayına ortogondur. Aynı şekilde, $\mathcal{N}(\mathbf{A}') = C^\perp(\mathbf{A})$ dür.
- ii. \mathbf{A} bir kare simetrik matris olduğunda, $C(\mathbf{A}) = C(\mathbf{A}')$, $\mathcal{N}(\mathbf{A}) = \mathcal{N}(\mathbf{A}')$ ve $\mathcal{N}(\mathbf{A}) = C^\perp(\mathbf{A}')$ dür.
- iii. $\text{boy}[C(\mathbf{A})] = \text{rank}(\mathbf{A}) = \text{boy}[C(\mathbf{A}')]$ dür.
- iv. $\text{boy}[C(\mathbf{A})] = n - r$ dir. Burada $r = \text{rank}(\mathbf{A})$ dir. $\text{boy}[C(\mathbf{A})]$ ya \mathbf{A} nın sıfırlığı denir.
- v. Her $\mathbf{A} \in \mathbb{R}^{m \times n}$ için $\text{boy}[C(\mathbf{A})] + \text{boy}[\mathcal{N}(\mathbf{A})] = n$ dir.
- vi. Kesin olarak \mathbf{A} kare ve tekil olmadığında, $\text{boy}[\mathcal{N}(\mathbf{A})] = 0$ dir.
- vii. $\text{boy}[\mathcal{N}(\mathbf{A}')] = m - r$ dir.

Hatırlatma 2.1. Bir simetrik matris olarak $\mathbf{A}\mathbf{A}^+ \in \mathbb{R}^{m \times m}$; \mathbf{A} nın sütunları tarafından gerilen altuzaya (bu uzaya \mathbf{A} nın sütun uzayı da denir) ortogonal izdüşüm matrisini gösterir. Bir simetrik matris olarak $\mathbf{A}^+\mathbf{A} \in \mathbb{R}^{n \times n}$ da, \mathbf{A} nın satırları tarafından gerilen \mathbf{A}' nün sütun uzayı denilen $(C(\mathbf{A}'))$ altuzay üzerine ortogonal izdüşüm matrisini gösterir. $\mathbf{A} \in \mathbb{R}^{m \times m}$ tekil olmadığında, $\mathbf{A}^+ = \mathbf{A}^{-1}$ ve $\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}_m$ dir.

2.1.3. Lineer Dönüşüm ve Operatörler



Şekil 2.1. V vektör uzayının W ye dönüşümü kavramı.

V ve W iki vektör uzayı ve $T : V \rightarrow W$ bir $\mathbf{x} \in V$ vektörünü bir $\mathbf{y} \in W$ vektörüne götüren bir dönüşüm, yani $\mathbf{y} = T(\mathbf{x})$ olsun (bkz. Şekil 2.1). V vektör uzayına T nin tanım bölgesi ve $\{\mathbf{x} \in V \text{ için } \mathbf{y} | \mathbf{y} = T(\mathbf{x})\}$ kümesine de T nin değer bölgesi denir. Bu nedenle bu operatör, bir vektör uzayını kendi içine resmeden bir dönüşümdür.

T nin sıfır uzayı $T(\mathbf{x}) = \mathbf{0}$ olacak şekilde $\mathbf{x} \in V$ nin kümesidir. Eğer $\mathbf{x}_1 = \mathbf{x}_2$ olduğunda kesin olarak $T(\mathbf{x}_1) = T(\mathbf{x}_2)$ ise, T dönüşümüne “bire-bir” dir denir ve eğer $T(\mathbf{x}_1 + \mathbf{x}_2) = T(\mathbf{x}_1) + T(\mathbf{x}_2)$ (toplanabilirlik) ve herhangi bir $a \in \mathbb{R}$ ve her $\mathbf{x} \in V$ için $T(a\mathbf{x}) = aT(\mathbf{x})$ (homojenlik) gerçekleşirse, T ye lineer denir.

Bir vektör uzayı verildiğinde, onun için bir taban ve bir koordinatlar sistemi bulabiliriz. V ve W için seçilen bir tabana göre sonlu boyutlu V ve W vektör uzayları arasındaki her lineer dönüşümün bir $\mathbf{A} \in \mathbb{R}^{m \times n}$ matrisiyle gösterilebilmesi gerçeği esastır. Burada n ve m sırasıyla V ve W vektör uzaylarının boyutlarıdır. Bu nedenle, lineer dönüşüm ve onu gösteren matris arasında ayırım yapmayacağız. Böylece, $\mathbf{y} = T(\mathbf{x})$ yerine $\mathbf{y} = \mathbf{A}\mathbf{x}$ yazacağız. Burada $\mathbf{A} \in \mathbb{R}^{m \times n}$ dir. Buna göre lineer dönüşümler ile ilgili çıkarılacak sonuçları, matrislerin özelliklerinden çıkarabiliriz (Meyer, 2000).

Hatırlatma 2.2. V vektör uzayından W vektör uzayına birden fazla lineer dönüşümlerin kümesi $B(V, W)$ ile gösterilir. $V=W$ olduğunda $B(V)$, V üzerindeki tüm lineer operatörlerin kümesini gösterir. V bir sonlu-boyutlu vektör uzayı olduğunda lineer operatörlerin özelliklerini kare matrislerin özellikleri vasıtasıyla analiz edebiliriz.

❖ **Dönüşümün /Operatörün Örnekleri :**

a) Öteleme : Herhangi sabit $\mathbf{a} \in V$ için eğer her $\mathbf{x} \in V$ için $T(\mathbf{x}) = \mathbf{x} + \mathbf{a}$ ise, $T : V \rightarrow W$ operatörüne (dönüşümüne) bir öteleme denir.

$$T(\mathbf{x} + \mathbf{y}) = \mathbf{a} + \mathbf{x} + \mathbf{y} \neq T(\mathbf{x}) + T(\mathbf{y}) = 2\mathbf{a} + \mathbf{x} + \mathbf{y}$$

olduğundan, ötelemenin bir lineer dönüşüm (operatör) olmadığı görülür. \mathbb{R}^m de çok iyi bilinen bir öteleme dönüşümüne

$$T(\mathbf{x}) = \mathbf{x} - \bar{\mathbf{x}}\mathbf{1}$$

ile tanımlanan merkezileştirme denir. Burada $\bar{\mathbf{x}} = \frac{1}{m} \sum_{i=1}^m x_i$ ve $\mathbf{1} = (1 \ 1 \ 1 \ \dots \ 1)$ elemanlarının tümü 1 lerden oluşan bir vektör (sütun vektörü) dür. Öteleme, vektörler arasındaki uzaklığı değiştirmez, yani $\|\mathbf{x} - \mathbf{y}\|_2 = \|T(\mathbf{x}) - T(\mathbf{y})\|_2$ dir.

Hatırlatma 2.3. Vektör normlarının özel bir hali olan 2-normu veya Öklidyen normu, $\|\mathbf{x}\|_2 = \sqrt{x_1^2 + x_2^2 + \dots + x_m^2} = \|\mathbf{x}'\mathbf{x}\|_2$ şeklindedir.

b) Dönme : Her $m \times m$ ortogonal matris (Eğer $\mathbf{Q}^{-1} = \mathbf{Q}'$, yani $\mathbf{Q}'\mathbf{Q}^{-1} = \mathbf{Q}^{-1}\mathbf{Q}' = \mathbf{I}$ ise, \mathbf{Q} ya ortogonal dendiğini hatırlayalım.), \mathbb{R}^m de bir katı-cisim dönmesini gösterir. Bir örnek olarak

$$\mathbf{Q} = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix},$$

saat yelkovanı yönünde bir θ açısı kadar dönmeyi gösteren bir 2×2 ortogonal matristir.

$\mathbf{Q}' = \mathbf{Q}^{-1}$ olduğundan,

$$\mathbf{Q}' = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$$

saat yelkovanının tersi yönünde bir θ açısı kadar dönmeyi verir. Ortogonal dönme, vektörlerin uzunluklarını değiştirmez, yani

$$\|\mathbf{Q}(\mathbf{x})\|_2^2 = ((\mathbf{Q}(\mathbf{x}))' (\mathbf{Q}(\mathbf{x}))) = \mathbf{x}' \mathbf{Q}' \mathbf{Q} \mathbf{x} = \mathbf{x}' \mathbf{x} = \|\mathbf{x}\|_2^2$$

dir (Meyer, 2000).

c) **Genel Koordinat Dönüşümü** : $\mathcal{B}_1 = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_m\}$, \mathbb{R}^m (m -boyutlu Öklid uzayı) için standart taban olsun. $\mathcal{B}_2 = \{\mathbf{g}_1, \mathbf{g}_2, \dots, \mathbf{g}_m\}$, \mathbb{R}^m için yeni bir taban olsun. İlk olarak yeni tabanın her bir üyesini standart tabana bağlı olarak aşağıdaki gibi ifade edelim:

$$\mathbf{g}_i = t_{i1}\mathbf{e}_1 + t_{i2}\mathbf{e}_2 + \dots + t_{im}\mathbf{e}_m, \quad 1 \leq i \leq m. \quad (2.30)$$

Burada $(t_{1i}, t_{2i}, \dots, t_{mi})$, standart tabanda \mathbf{g}_i vektörünün gösterimidir. Bu m -bağıntıyı matris notasyonuna göre yeniden yazarak,

$$[\mathbf{g}_1 \quad \mathbf{g}_2 \quad \dots \quad \mathbf{g}_m] = [\mathbf{e}_1 \quad \mathbf{e}_2 \quad \dots \quad \mathbf{e}_m] \begin{bmatrix} t_{11} & t_{12} & \dots & t_{1m} \\ t_{21} & t_{22} & \dots & t_{2m} \\ \vdots & \vdots & & \vdots \\ t_{m1} & t_{m2} & \dots & t_{mm} \end{bmatrix} \quad (2.31)$$

veya

$$\mathbf{G} = \mathbf{E}\mathbf{T} \quad (2.32)$$

elde ederiz. Burada $\mathbf{G} = [\mathbf{g}_1 \quad \mathbf{g}_2 \quad \dots \quad \mathbf{g}_m] \in \mathbb{R}^{m \times m}$, yani yeni \mathcal{B}_2 tabanının üyeleri tarafından biçimlendirilen matris, $\mathbf{E} = [\mathbf{e}_1 \quad \mathbf{e}_2 \quad \dots \quad \mathbf{e}_m] \in \mathbb{R}^{m \times m}$, yani standart \mathcal{B}_1 tabanının üyeleri tarafından biçimlendirilen matris ve $\mathbf{T} \in \mathbb{R}^{m \times m}$ elemanları standart tabana göre yeni tabanın koordinatları olan matristir. \mathbf{T} nin tekil olmayan bir matris olduğu kolaylıkla gerçekleştirilebilir.

Şimdi bir $\mathbf{x} \in \mathbb{R}^m$ vektörünü gözönüne alalım.

$$\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \dots + x_m\mathbf{e}_m, \quad (2.33)$$

\mathbf{x} in \mathcal{B}_1 standart tabanına göre gösterimi olsun ve

$$\mathbf{x} = x_1^*\mathbf{g}_1 + x_2^*\mathbf{g}_2 + \dots + x_m^*\mathbf{g}_m, \quad (2.34)$$

aynı \mathbf{x} in yeni \mathcal{B}_2 tabanına göre gösterimi olsun. Şimdi (2.34) ile (2.30) un birleşimi,

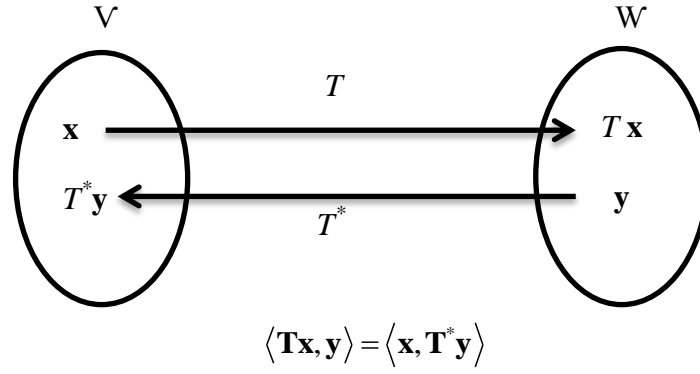
$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} = \begin{bmatrix} t_{11} & t_{12} & \dots & t_{1m} \\ t_{21} & t_{22} & \dots & t_{2m} \\ \vdots & \vdots & & \vdots \\ t_{m1} & t_{m2} & \dots & t_{mm} \end{bmatrix} \begin{bmatrix} x_1^* \\ x_2^* \\ \vdots \\ x_m^* \end{bmatrix} \quad (2.35)$$

veya

$$\mathbf{x} = \mathbf{T}\mathbf{x}^* \quad (2.36)$$

olduğunu ortaya koyar. Bu nedenle, $\mathbf{T}: \mathbb{R}^m \rightarrow \mathbb{R}^m$ vektörlerin standart ve yeni tabanlara göre gösterimi arasındaki bir köprü (geçişi) sağlayan bir matristir. Bu \mathbf{T} matrisi non-singülerdir (tekil değildir) fakat ortogonal değildir ve bu nedenle bazen eğik dönme adımı alır (Meyer, 2000).

d) Benzerlik Dönüşümü : \mathbf{x} ve \mathbf{y} standart tabanda \mathbb{R}^m deki iki vektörü gösterebiliriz. \mathbf{A} , $\mathbf{y} = \mathbf{A}\mathbf{x}$ olacak şekilde standart tabanda bir lineer dönüşümün matrisini gösterebiliriz. \mathbf{T} , \mathcal{B}_1 standart tabanının yeni \mathcal{B}_2 tabanına dönüşümünü gösteren tekil olmayan matris olsun. \mathbf{x}^* ve \mathbf{y}^* , $\mathbf{x} = \mathbf{T}\mathbf{x}^*$ ve $\mathbf{y} = \mathbf{T}\mathbf{y}^*$ olmak üzere, \mathbf{x} ve \mathbf{y} nin yeni \mathcal{B}_2 tabanındaki gösterimi olsun.



Şekil 2.2. Ek (adjoint) Dönüşümün Tanımı.

$\mathbf{y} = \mathbf{A}\mathbf{x}$ olduğundan, $\mathbf{y} = \mathbf{T}\mathbf{y}^*$ de yerine koyarak,

$$\mathbf{y} = \mathbf{T}\mathbf{y}^* = \mathbf{A}\mathbf{x} = \mathbf{A}\mathbf{T}\mathbf{x}^* \quad (2.37)$$

veya

$$\mathbf{y}^* = (\mathbf{T}^{-1}\mathbf{A}\mathbf{T})\mathbf{x}^* \quad (2.38)$$

elde ederiz. Başka bir deyişle, $\mathbf{T}^{-1}\mathbf{A}\mathbf{T}$ yeni \mathcal{B}_2 tabanına göre \mathbf{A} matrisinin gösterimidir. \mathbf{A} matrisinin $\mathbf{T}^{-1}\mathbf{A}\mathbf{T}$ matrisine dönüşümüne “benzerlik dönüşümü” denir. Benzerlik dönüşümü özdeğerleri korur. Yani, \mathbf{A} ve $\mathbf{T}^{-1}\mathbf{A}\mathbf{T}$ aynı özdeğerler kümesine sahiptir (Golub ve Van Loan 1989; Meyer, 2000).

e) Eşleşim (Kongrüans) Dönüşümü : $\mathbf{A} \in \mathbb{R}^{m \times m}$ olsun. Eğer $\mathbf{B} \in \mathbb{R}^{m \times m}$ tekil olmayan bir matris ise, bu takdirde $\mathbf{B}^t\mathbf{A}\mathbf{B}$ ye \mathbf{A} nın eşleşim dönüşümü denir. Yani, \mathbf{A} ve $\mathbf{B}^t\mathbf{A}\mathbf{B}$ ye denk matrisler denir. Eşleşim dönüşümü \mathbf{A} nın birçok özelliklerini korur. Bu nedenle \mathbf{A} simetrik, anti-simetrik veya pozitif (yarı) tanımlı olduğunda, $\mathbf{B}^t\mathbf{A}\mathbf{B}$ simetrik, anti-simetrik, pozitif (yarı) tanımlıdır.

f) Ek (Adjoint) Dönüşüm : V ve W , her biri kendi iç çarpımına sahip iki vektör uzayı olsun. $\langle \mathbf{x}, \mathbf{y} \rangle_V$, V deki \mathbf{x} ve \mathbf{y} vektörlerinin iç çarpımını gösterebilir. Benzer şekilde $\langle \mathbf{x}, \mathbf{y} \rangle_W$ tanımlanır. $T: V \rightarrow W$, $\mathbf{x} \in V$ ve $\mathbf{y} \in W$ olmak üzere, $\mathbf{y} = T\mathbf{x}$ olacak şekilde bir lineer dönüşüm olsun. Eğer

$$\langle T\mathbf{x}, \mathbf{y} \rangle_W = \langle \mathbf{x}, T^*\mathbf{y} \rangle_V \quad (2.39)$$

olacak şekilde, bir T lineer dönüşümü varsa, bu takdirde T^* a T nin “eki” denir. Gösterim için Şekil 2.2 ye başvururuz.

Şimdi bu tanımları özelleştireceğiz. $V \subseteq \mathbb{R}^n$ ve $W \subseteq \mathbb{R}^m$ olsun. Bu takdirde T ve T^* in her ikisi de lineer operatörler olarak, V ve W için seçilen tabanda bir matris gösterimine sahiptir. $\mathbf{A}: V \rightarrow W$ ve $\mathbf{A}^*: W \rightarrow V$ olmak üzere, $\mathbf{A} \in \mathbb{R}^{m \times n}$ ve $\mathbf{A}^* \in \mathbb{R}^{n \times m}$, T ve T^* a karşılık gelen matrisler olsun. Bu takdirde, yukarıdaki tanımdan hemen,

$$\langle \mathbf{Ax}, \mathbf{y} \rangle = (\mathbf{Ax})' \mathbf{y} = \mathbf{x}' \mathbf{A}' \mathbf{y} = \langle \mathbf{x}, \mathbf{A}' \mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{A}^* \mathbf{y} \rangle \quad (2.40)$$

elde ederiz ki, buradan $\mathbf{A}^* = \mathbf{A}'$ olduğu görülür. Yani V ve W sonlu boyutlu vektör uzayları olduğunda, \mathbf{A} nın transpozesi gerçekten, \mathbf{A} nın “eki” dir.

$V=W=\mathbb{R}^m$ olduğunda $\mathbf{A} \in \mathbb{R}^{m \times m}$ bir lineer operatördür ve $\mathbf{A}^* = \mathbf{A}'$ ye ek operatör denir. Özel durumda, \mathbf{A} bir simetrik matris olduğunda, $\mathbf{A}^* = \mathbf{A}' = \mathbf{A}$ dır ve \mathbf{A} nın kendisine bir ek operatör denir.

Aşağıdaki özellikler kolayca gerçekleştirilebilir.

- i. Bir \mathbf{A} lineer dönüşümünün \mathbf{A}^* eki de lineerdir.
- ii. $(\mathbf{A}^*)^* = \mathbf{A}$ dır.
- iii. $(a\mathbf{A})^* = a\mathbf{A}^*$ dır.
- iv. $(\mathbf{A} + \mathbf{B})^* = \mathbf{A}^* + \mathbf{B}^*$ dır.
- v. $(\mathbf{AB})^* = \mathbf{B}^* \mathbf{A}^*$ dır.
- vi. Eğer \mathbf{A} tersinir ise, bu takdirde $(\mathbf{A}^{-1})^* = (\mathbf{A}^*)^{-1}$ dir.

2.1.4. Lineer Sistemlerin Çözümü

$\mathbf{A} \in \mathbb{R}^{m \times n}$ ve $\mathbf{b} \in \mathbb{R}^m$ olsun. \mathbf{A} ve \mathbf{b} verildiğinde, $\mathbf{Ax} = \mathbf{b}$ olacak şekilde $\mathbf{x} \in \mathbb{R}^n$ vektörünü bulma problemi, \mathbf{x} in bileşenlerinin bilinmeyenler olduğu n -bilinmeyenli m -denklemlilik bir sistemi çözme problemi olarak bilinir. Eğer \mathbf{b} vektörü $\in C(\mathbf{A})$ ise, bu takdirde \mathbf{b} , \mathbf{A} nın sütunlarının bir lineer kombinasyonu olarak ifade

edilebilmelidir ve bu lineer kombinasyonun katsayıları \mathbf{x} vektörünün bulmak istediğimiz bileşenleri olmalıdır. Bu durumda, $\mathbf{Ax}=\mathbf{b}$ lineer sistemine tutarlı bir sistem denir. Eğer $\mathbf{AA}^{-}\mathbf{b}=\mathbf{b}$ veya $\mathbf{AA}^{+}\mathbf{b}=\mathbf{b}$ ya da $\text{rank}(\mathbf{A}:\mathbf{b})=\text{rank}(\mathbf{A})$ ise, $\mathbf{Ax}=\mathbf{b}$ lineer denklemler sistemi tutarlıdır denir. Burada $(\mathbf{A}:\mathbf{b})$ ye sistemin genelleştirilmiş katsayılar matrisi denir. Eğer \mathbf{b} vektörü $C(\mathbf{A})$ ya, yani \mathbf{A} nın sütun uzayına, ait değilse, bu takdirde $\mathbf{Ax}=\mathbf{b}$ olacak şekilde bir $\mathbf{x} \in \mathbb{R}^n$ mevcut değildir ve bu durumda sistem tutarsız olarak adlandırılır. $\mathbf{b}=0$ olduğunda, $\mathbf{Ax}=0$ lineer sistemine bir homojen sistem denir ve bu sistem daima tutarlıdır.

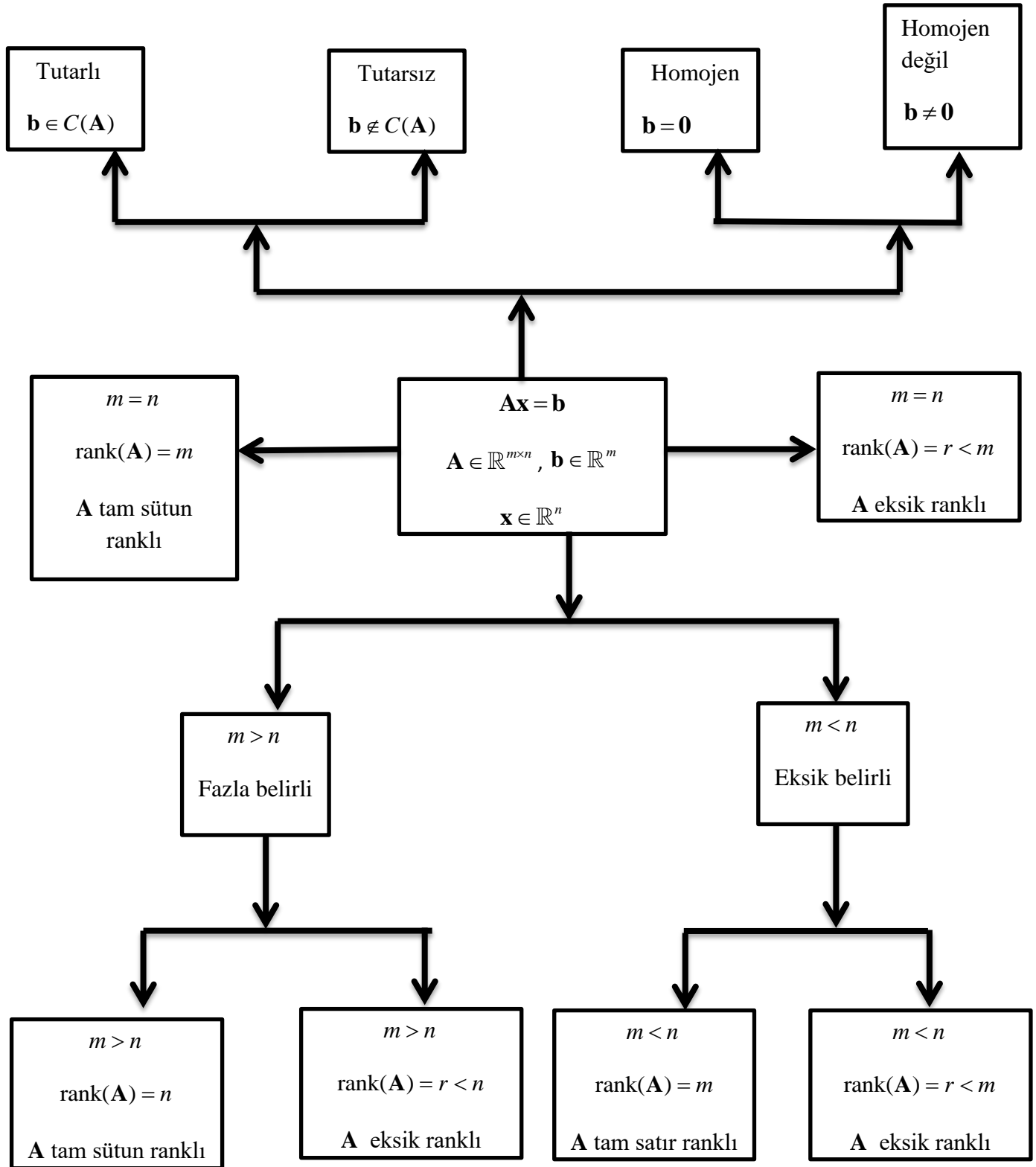
$\text{rank}(\mathbf{A}) \leq \min(m, n)$ olduğunu hatırlayalım. Eğer $\text{rank}(\mathbf{A}) = \min(m, n)$ ise, lineer sistemin tam ranklı olduğu söylenir, aksi takdirde sisteme “eksik ranklı” denir. Eğer denklemlerin sayısı m , bilinmeyenlerin sayısı n ye eşit ise, sisteme belirli (kararlı) sistem denir. Eğer $m > n$ ise, sistem fazla belirli sistem olarak bilinir ve eğer $m < n$ ise, sistem eksik belirli sistem olarak bilinir.

Tutarlı/tutarsız, homojen/homojen olmayan, tam ranklı/eksik ranklı ve belirli/fazla veya eksik belirli gibi çeşitli nitelikleri bir araya getirerek, Şekil 2.3 de gösterilen bir tür şematik kombinasyonları elde ederiz.

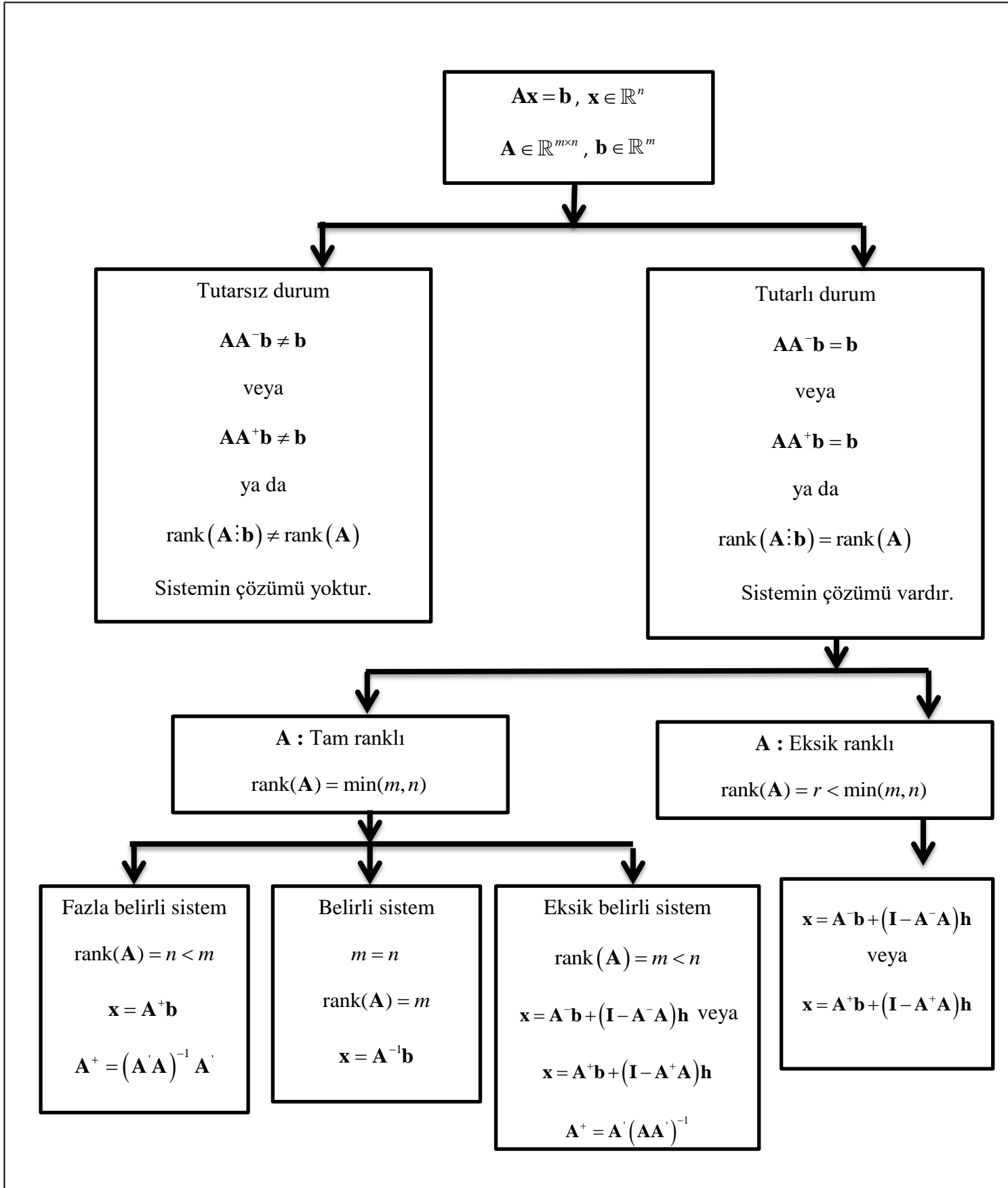
- i. Eğer $m = n$ ve \mathbf{A} tam ranklı ise, sistem daima tutarlıdır, çözüm mevcuttur ve tektir. Eğer $\mathbf{b} \neq 0$ ise, bu takdirde $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$ sistemin bir tek çözümüdür ve eğer $\mathbf{b} = 0$ ise, $\mathbf{x} = 0$ yegane çözümdür.
- ii. Sistemin tutarlı ve eksik belirli veya fazla belirli olması durumunda genel çözüm, \mathbf{h} keyfi bir vektör olmak üzere, sistemin genel çözümü

$$\mathbf{x} = \mathbf{A}^{-}\mathbf{b} + (\mathbf{I} - \mathbf{A}^{-}\mathbf{A})\mathbf{h} \quad \text{veya} \quad \mathbf{x} = \mathbf{A}^{+}\mathbf{b} + (\mathbf{I} - \mathbf{A}^{+}\mathbf{A})\mathbf{h}$$

olarak belirlenir. $\mathbf{h}=0$ olduğunda, $\mathbf{x} = \mathbf{A}^{-}\mathbf{b}$ veya $\mathbf{x} = \mathbf{A}^{+}\mathbf{b}$ ye sistemin özel bir çözümü olarak bakılır. Bu çözümlerin bir özeti Şekil 2.4 de verilmiştir (Bellman, 1960).



Şekil 2.3. Lineer Sistemlerin Sınıflandırılması.



Şekil 2.4. A^{+} ın A nın Moore-Penrose Genelleştirilmiş Tersisi Olduğu Çeşitli Durumlarda Lineer Sistemlerin Çözümü.

❖ **Özel Matrisler :**

a) **Simetrik Matris :** Eğer $\mathbf{A}' = \mathbf{A}$ ise, $\mathbf{A} \in \mathbb{R}^{m \times m}$ matrisi simetriktir. Aşağıdaki özellikler kolayca gerçekleşir:

- i. Eğer \mathbf{A} simetrik ise, \mathbf{A}^{-1} de simetriktir.
- ii. Eğer \mathbf{A} ve \mathbf{B} simetrik iseler, bu takdirde \mathbf{A} ve \mathbf{B} değişimli, yani $\mathbf{AB} = \mathbf{BA}$ olduğunda, \mathbf{AB} de simetriktir.

Eğer $\mathbf{A} \in \mathbb{R}^{m \times n}$ ise, bu takdirde $\mathbf{AA}' \in \mathbb{R}^{m \times m}$ ve $\mathbf{A}'\mathbf{A} \in \mathbb{R}^{n \times n}$ her ikisi de Gramian denilen simetrik matrislerdir (Meyer, 2000).

b) **Anti-Simetrik Matris :** Eğer $\mathbf{A}' = -\mathbf{A}$ ise, $\mathbf{A} \in \mathbb{R}^{m \times m}$ matrisi anti-simetriktir. Bu nedenle, bir anti-simetrik matrisin köşegen elemanları sıfırdır. Verilen herhangi bir $\mathbf{A} \in \mathbb{R}^{m \times m}$ matrisi için,

- i. $\mathbf{A}_S = (\mathbf{A} + \mathbf{A}') / 2$ bir simetrik matristir ve
- ii. $\mathbf{A}_{SS} = (\mathbf{A} - \mathbf{A}') / 2$ bir anti-simetrik matristir.

\mathbf{A}_S ve \mathbf{A}_{SS} ye sırasıyla \mathbf{A} nın simetrik ve anti-simetrik parçaları denir (Meyer, 2000).

c) **Bilineer Form :** $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{x} \in \mathbb{R}^m$ ve $\mathbf{y} \in \mathbb{R}^n$ olsun. Bu takdirde $f_A = \mathbf{x}'\mathbf{A}\mathbf{y}$ ile tanımlanan $f_A(\mathbf{x}, \mathbf{y}) : \mathbb{R}^m \times \mathbb{R}^n \rightarrow \mathbb{R}$ (iki vektörün skalar değerli bir fonksiyonu) ye, \mathbf{x} ve \mathbf{y} vektörlerinin bileşenleri $f_A(\mathbf{x}, \mathbf{y})$ de onların birinci derecesiyle yer aldığından, bilineer form denir.

d) **Kuadratik (Karesel) Form :** $\mathbf{A} \in \mathbb{R}^{m \times m}$, $\mathbf{x} \in \mathbb{R}^m$ olsun. Bu takdirde $\mathbf{Q}_A : \mathbb{R}^m \rightarrow \mathbb{R}$ (bir vektörün skalar değerli fonksiyonu)

$$\mathbf{Q}_A(\mathbf{x}) = \mathbf{x}'\mathbf{A}\mathbf{x}$$

(2.41)

olarak tanımlanır ve \mathbf{x} e göre bir kuadratik form adını alır. $\mathbf{Q}_A(\mathbf{x}) = \mathbf{x}'\mathbf{A}\mathbf{x}$ kuadratik formu \mathbf{x} in elemanlarına göre ikinci dereceden bir çok değişkenli polinomdur. Bir örnek olarak, $n = 2$ olduğunda,

$$\mathbf{Q}_A(\mathbf{x}) = a_{11}x_1^2 + (a_{12} + a_{21})x_1x_2 + a_{22}x_2^2$$

dir. $\mathbf{Q}_A(\mathbf{x})$ bir skalar olduğundan,

$$[\mathbf{Q}_A(\mathbf{x})]' = \mathbf{x}'\mathbf{A}'\mathbf{x} = \mathbf{Q}_A(\mathbf{x}) = \mathbf{x}'\mathbf{A}\mathbf{x}$$

yazarız. Bu nedenle, \mathbf{A}_s ; \mathbf{A} nın bir simetrik kısmı olmak üzere,

$$\mathbf{Q}_A(\mathbf{x}) = \frac{1}{2}[\mathbf{x}'\mathbf{A}'\mathbf{x} + \mathbf{x}'\mathbf{A}\mathbf{x}] = \mathbf{x}'\left(\frac{\mathbf{A} + \mathbf{A}'}{2}\right)\mathbf{x} = \mathbf{x}'\mathbf{A}_s\mathbf{x}$$

dir. Böylece, genelliği bozmaksızın, $\mathbf{Q}_A(\mathbf{x})$ deki \mathbf{A} matrisinin bir simetrik matris olduğunu kabul ederiz ve

$$\mathbf{Q}_A(\mathbf{x}) = a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2$$

yazarız.

e) Pozitif /Negatif Tanımlı Matris: Bir reel simetrik $\mathbf{A} \in \mathbb{R}^{m \times m}$ matrisi için

$$\mathbf{x}'\mathbf{A}\mathbf{x} \begin{cases} > 0, & \text{eğer } \mathbf{x} \neq \mathbf{0} \text{ ise} \\ = 0, & \text{kesin olarak } \mathbf{x} = \mathbf{0} \text{ olduğunda} \end{cases} \quad (2.42)$$

olduğunda, \mathbf{A} simetrik matrisine pozitif tanımlıdır denir.

Bir reel simetrik \mathbf{A} matrisi için

$$\text{her } \mathbf{x} \in \mathbb{R}^m \text{ için, } \mathbf{x}'\mathbf{A}\mathbf{x} \geq 0 \quad (2.43)$$

olduğunda, \mathbf{A} simetrik matrisine pozitif yarı-tanımlıdır denir. Eğer \mathbf{A} pozitif yarı-tanımlı ise, bunu $\mathbf{A} \succeq \mathbf{0}$ ile gösteririz. Eğer \mathbf{A} pozitif tanımlı ise bunu $\mathbf{A} \succ \mathbf{0}$ ile gösteririz.

Pozitif yarı-tanımlılığın aşağıdaki gibi birçok eşdeğer karakterizasyonu vardır (Harvey, 2011).

❖ **Yararlı bir şart :**

$$\mathbf{A} \succeq \mathbf{0} \Leftrightarrow \forall \mathbf{x} \in \mathbb{R}^d \text{ için } \mathbf{x}'\mathbf{A}\mathbf{x} \geq 0. \quad (2.44a)$$

Benzer şekilde,

$$\mathbf{A} \succ \mathbf{0} \Leftrightarrow \forall \mathbf{x} \in \mathbb{R}^d \text{ için } \mathbf{x}'\mathbf{A}\mathbf{x} > 0. \quad (2.44b)$$

❖ **Başka bir şart :** $\mathbf{A} \succeq \mathbf{0}$ olması için gerek ve yeter şart $\mathbf{A} = \mathbf{V}\mathbf{V}'$ olacak şekilde \mathbf{V} matrisinin var olmasıdır.

Pozitif yarı-tanımlı şartı tüm simetrik matrisler üzerinde bir kısmi sıralamayı tanımlamak için kullanılabilir. Bu sıralamaya **Loewner sıralaması** veya **pozitif yarı-**

tanımlı sıralama denir. Herhangi iki simetrik \mathbf{A} ve \mathbf{B} matrisi için, eğer $\mathbf{A} - \mathbf{B} \succcurlyeq \mathbf{0}$ ise, $\mathbf{A} \succcurlyeq \mathbf{B}$ yazarız.

Olgu 2.1. $\mathbf{A} \succcurlyeq \mathbf{B}$ olması için gerek ve yeter şart her \mathbf{x} vektörü için $\mathbf{x}'\mathbf{A}\mathbf{x} \geq \mathbf{x}'\mathbf{B}\mathbf{x}$ olmasıdır.

İspat. $\mathbf{x}'\mathbf{A}\mathbf{x} \geq \mathbf{x}'\mathbf{B}\mathbf{x}$ nin sağlanması için gerek ve yeter şartın $\mathbf{V}'(\mathbf{A} - \mathbf{B})\mathbf{V} \geq \mathbf{0}$ eşitsizliğinin sağlanması gerektiğine dikkat edelim. (2.44a) ya göre, bu $\mathbf{A} \succcurlyeq \mathbf{B}$ nin tanımını olan $\mathbf{A} - \mathbf{B} \succcurlyeq \mathbf{0}$ eşitsizliğine denktir. \square

Pozitif yarı-tanımlı matrislerin kümesi, toplama ve negatif olmama ölçeklendirmesi altında kapalıdır (Böyle bir kümeye bir **konveks koni** denir).

Olgu 2.2. \mathbf{A} ve \mathbf{B} $d \times d$ boyutlu pozitif yarı-tanımlı matrisler olsun. α, β negatif olmayan skalarlar olsun. Bu takdirde $\alpha\mathbf{A} + \beta\mathbf{B} \succcurlyeq \mathbf{0}$ dır.

İspat. Bu, (2.44a) dan kolayca görülür. \square

Uyarma 2.1. Loewner sıralaması reel sayıların alışılmış sıralamasının güzel özelliklerinin tümüne sahip değildir. Örneğin, eğer $\mathbf{A} \succcurlyeq \mathbf{B} \succcurlyeq \mathbf{0}$ ise, bu takdirde $\mathbf{A}^2 \succcurlyeq \mathbf{B}^2$ olması gerekmez.

Olgu 2.3. \mathbf{A} ve \mathbf{B} $d \times d$ boyutlu simetrik matrisler olsun ve \mathbf{C} herhangi bir $d \times d$ boyutlu matris olsun. Eğer $\mathbf{A} \succcurlyeq \mathbf{B}$ ise, bu takdirde $\mathbf{C}\mathbf{A}\mathbf{C}' \succcurlyeq \mathbf{C}\mathbf{B}\mathbf{C}'$ dür. Bundan başka, eğer \mathbf{C} tekil değilse, bu takdirde “gerek şart” sahiden “gerek ve yeter şart” tır.

İspat. $\mathbf{A} \succcurlyeq \mathbf{B}$ olduğundan, bu durumda $\mathbf{A} - \mathbf{B} \succcurlyeq \mathbf{0}$ dır, bu nedenle, $\mathbf{A} - \mathbf{B}; \mathbf{V}\mathbf{V}'$ olacak şekilde bir \mathbf{V} matrisi mevcuttur. Bu takdirde,

$$\mathbf{C}(\mathbf{A} - \mathbf{B})\mathbf{C}' = \mathbf{C}\mathbf{V}\mathbf{V}'\mathbf{C}' = \mathbf{C}\mathbf{V}(\mathbf{C}\mathbf{V})'$$

olur ki, bu ise, $\mathbf{C}(\mathbf{A} - \mathbf{B})\mathbf{C}'$ nün pozitif yarı-tanımlı olduğunu gösterir. \mathbf{C} nin tekil olmadığını ve $\mathbf{C}\mathbf{A}\mathbf{C}' \succcurlyeq \mathbf{C}\mathbf{B}\mathbf{C}'$ olduğunu farz edelim. $\mathbf{A} \succcurlyeq \mathbf{B}$ yi elde etmek için \mathbf{C}^{-1} ile soldan ve $(\mathbf{C}^{-1})' = (\mathbf{C}')^{-1}$ ile sağdan çarparız. \square

Bu nedenle, $\mathbf{x}'\mathbf{A}\mathbf{x}$ kuadratik formu sadece orijinde sifıra eşit olduğunda \mathbf{A} pozitif tanımlıdır ve eğer orijine ilaveten en az bir $\mathbf{x} \neq \mathbf{0}$ için, $\mathbf{x}'\mathbf{A}\mathbf{x}$ sıfır oluyorsa, \mathbf{A} matrisi pozitif yarı-tanımlıdır. Eğer yukarıdaki özellikler tersine çevrilen eşitsizlikler ile gerçekleşirse, bir reel simetrik \mathbf{A} matrisine negatif tanımlı ve negatif yarı-tanımlı denir. Eğer herhangi $\mathbf{x} \in \mathbb{R}^m$ için $\mathbf{x}'\mathbf{A}\mathbf{x} \geq 0$ ve \mathbf{x} in herhangi diğer değerleri için $\mathbf{x}'\mathbf{A}\mathbf{x} < 0$ ise, \mathbf{A} matrisine tanımsız denir.

Şimdi pozitif ve pozitif yarı-tanımlı matrislerin birkaç önemli özelliğini yazacağız.

- i. Bir pozitif tanımlı matrisin köşegen elemanları, her mertebeden asıl minörleri ve determinantı pozitiftir.
- ii. Pozitif tanımlılık, matrisin elemanlarının pozitif olduklarını ifade etmez.

Örneğin,

$$\mathbf{A} = \begin{bmatrix} 4 & -2 \\ -2 & 2 \end{bmatrix}$$

pozitif tanımlıdır, fakat

$$\mathbf{B} = \begin{bmatrix} 2 & 2 \\ 3 & 1 \end{bmatrix}$$

pozitif tanımlı değildir.

- iii. \mathbf{A} pozitif tanımlı ise, \mathbf{A}^{-1} de pozitif tanımlıdır.
- iv. Eğer \mathbf{A} pozitif tanımlı ve \mathbf{B} tekil değilse, bu takdirde $(\mathbf{B}^{-1})' \mathbf{A} \mathbf{B}^{-1}$ pozitif tanımlıdır.

f) Köşegen Olarak Baskın (Etkin) Matris : Bir $\mathbf{A} \in \mathbb{R}^{m \times m}$ kare matrisine, eğer

$$|a_{ii}| \geq \sum_{j \neq i} |a_{ij}| \quad (2.45)$$

ise, köşegen olarak etkin matris denir. Eğer her $i = 1, 2, \dots, m$ için eşitsizlik sağlanırsa, kesinlikle köşegen olarak etkin matris denir. Köşegen olarak etkin matrisin tekil olmadığı gösterilebilir (Trefethen ve Bau, 1997).

g) Ortogonal (Dik) Matris : Bir $\mathbf{Q} \in \mathbb{R}^{m \times m}$ matrisine eğer $\mathbf{Q}^{-1} = \mathbf{Q}'$ ise (yani $\mathbf{Q}'\mathbf{Q} = \mathbf{Q}\mathbf{Q}' = \mathbf{I}_m$ ise), bir ortogonal matris denir. Bir ortogonal matris aşağıdaki özelliklere sahiptir:

- i. Eğer \mathbf{Q} ortogonal ise, \mathbf{Q} tekil değildir,
- ii. $\mathbf{Q} \in \mathbb{R}^{m \times m}$ nin sütunları (satırları) \mathbb{R}^m için bir tam ortonormal taban oluşturur.
- iii. Eğer $\mathbf{x} \in \mathbb{R}^m$ ise, bu takdirde $\|\mathbf{x}\|_2 = \|\mathbf{Q}\mathbf{x}\|_2$ dir, yani Öklid normu bir ortogonal dönüşüm altında değişmez; bu nedenle, ortogonal dönüşümlere “izometrik” dönüşümler denir.
- iv. $\det(\mathbf{Q}'\mathbf{A}\mathbf{Q}) = \det(\mathbf{A})$ dir.

- v. Eğer \mathbf{Q}_1 ve \mathbf{Q}_2 ortogonal ise, bu takdirde onların $\mathbf{Q}_1\mathbf{Q}_2$ çarpımı da ortogondur.
- vi. Bir lineer dönüşüm olarak, ortogonal dönüşüm bir katı cisim dönmesini temsil eder. Bu nedenle

$$\mathbf{Q} = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix}$$

saat yelkovanı yönünde bir θ açısı kadar bir dönmeyi gösteren bir ortogonal matristir (Meyer, 2000).

h) P Permütasyon (Değişim) Matrisi : Eğer $\mathbf{p}_{ij} \in \{0,1\}$ ve her bir satır ve her bir sütunda sadece 1 varsa, $\mathbf{P} = [\mathbf{p}_{ij}] \in \mathbb{R}^{m \times m}$ ye bir permütasyon matrisi denir. Örneğin,

$$\mathbf{P} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

bir permütasyon matrisidir. Permütasyon matrislerinin diğer bazı özellikleri aşağıdaki gibidir :

- i. \mathbf{I} birim matrisi bir permütasyon matrisidir.
- ii. Eğer $\mathbf{A} \in \mathbb{R}^{m \times m}$ herhangi bir matrisi ve $\mathbf{P} \in \mathbb{R}^{m \times m}$ bir permütasyon matrisi ise, bu takdirde \mathbf{PA} ; \mathbf{A} nın satırlarının ve \mathbf{AP} ; \mathbf{A} nın sütunlarının sırasını değiştirir.
- iii. Her permütasyon matrisi ortogondur.
- iv. İki permütasyon matrisinin çarpımı yine bir permütasyon matrisidir.
- v. $\mathbf{PP}^T = \mathbf{I}$ olduğundan, $\det(\mathbf{P}) = \pm 1$ elde ederiz ve bu nedenle her permütasyon matrisi tekil değildir.

i) İdempotent (Eş Güçlü) Matris : Bir $\mathbf{P} \in \mathbb{R}^{m \times m}$ matrisine eğer $\mathbf{P}^2 = \mathbf{P}$ ise, idempotenttir denir. Örneğin; $\mathbf{j}_n = [1 \ 1 \ \dots \ 1]^T$ tüm elemanları 1 olan $n \times 1$ boyutlu

bir matris ve $\mathbf{J}_n = \mathbf{j}_n \mathbf{j}_n^T = \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & 1 & \dots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \dots & 1 \end{bmatrix}$ tüm elemanları 1 lerden ibaret olan bir $n \times n$

kare matris olmak üzere, $\mathbf{J}_n \mathbf{J}_n = \mathbf{j}_n \mathbf{j}_n^T \mathbf{j}_n \mathbf{j}_n^T = n \mathbf{j}_n \mathbf{j}_n^T = n \mathbf{J}_n$ dir. Buna göre, $\frac{1}{n} \mathbf{J}_n$ bir simetrik-idempotent matristir. Gerçekten,

$$\left(\frac{1}{n}\mathbf{J}_n\right)\left(\frac{1}{n}\mathbf{J}_n\right) = \frac{1}{n^2}\mathbf{J}_n^2 = \frac{1}{n^2}\underset{=n}{\mathbf{j}_n\mathbf{j}_n'\mathbf{j}_n\mathbf{j}_n'} = \frac{1}{n}\mathbf{J}_n$$

dır. Böylece istenen ispatlanmış olur.

Ayrıca, $\mathbf{A}^-\mathbf{A}$, $\mathbf{A}\mathbf{A}^-$, $\mathbf{A}^+\mathbf{A}$ ve $\mathbf{A}\mathbf{A}^+$ matrisleri idempotent matrislerdir ve $\mathbf{A}^+\mathbf{A}$ ve $\mathbf{A}\mathbf{A}^+$ simetrik-idempotent matrislerdir.

İdempotent matrislerin özellikleri:

- i. Birim matris hariç olmak üzere, her idempotent matris tekildir.
- ii. Eğer \mathbf{P} idempotent ise, bu takdirde $(\mathbf{I} - \mathbf{P})$ de idempotenttir.
- iii. Eğer \mathbf{P}_1 ve \mathbf{P}_2 idempotent iseler, bu takdirde $\mathbf{P}_1\mathbf{P}_2$ sadece değişmeli olduğunda idempotenttir ve $\mathbf{P}_1 + \mathbf{P}_2$; $\mathbf{P}_1\mathbf{P}_2 = \mathbf{P}_2\mathbf{P}_1 = 0$ olduğunda idempotenttir.
- iv. Eğer $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{B} \in \mathbb{R}^{n \times m}$ ve $\text{rank}(\mathbf{A}) = n = \text{rank}(\mathbf{B})$ ise, bu takdirde

$$\mathbf{P} = \mathbf{B}(\mathbf{A}'\mathbf{B})^{-1}\mathbf{A}'$$

idempotenttir. $\mathbf{A} = \mathbf{B}$ olduğunda

$$\mathbf{P} = \mathbf{A}(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'$$

matrisinin idempotent olduğu açıktır. \mathbf{P} matrisi, $C(\mathbf{A})$ üzerine ortogonal izdüşüm matrisidir, yani ortogonal izdüşürücüdür.

Hatırlatma 2.4. İdempotent matrisler, her izdüşüm matrisinin idempotent olması anlamında birçok temel rol oynar. Bunun ötesinde, her simetrik, idempotent matris lineer en küçük kareler problemlerinin çözümünde kritik bir rol oynayan bir ortogonal izdüşüm tanımlar.

j) Nilpotent Matris : Bir $\mathbf{A} \in \mathbb{R}^{m \times m}$ matrisine eğer $\mathbf{A}^k = 0$ ve $\mathbf{A}^{k-1} \neq 0$ ise, k -indeksli bir nilpotent matris denir. Örneğin, alt üçgen matrisler kesin olarak nilpotenttirler.

Örneğin, $\mathbf{L} = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 2 & 3 & 0 \end{bmatrix}$ 3-indeksli bir nilpotent matristir (Meyer, 2000).

k) Grammian Matris : Eğer $\mathbf{A} \in \mathbb{R}^{m \times n}$ ise, bu takdirde $\mathbf{A}'\mathbf{A} \in \mathbb{R}^{n \times n}$ ve $\mathbf{A}\mathbf{A}' \in \mathbb{R}^{m \times m}$ matrislerine Grammian matrisler denir. Grammian matrisler aşağıdaki özelliklere sahiptir.

- i. Grammian matrisler simetriktir.

- ii. Eğer \mathbf{A} tam ranklı ise, bu takdirde onun Grammianı simetrik ve pozitif tanımlıdır,
çünkü her $\mathbf{x} \neq 0$ için, $\mathbf{x}'(\mathbf{A}'\mathbf{A})\mathbf{x} = (\mathbf{Ax})'\mathbf{Ax} = \|\mathbf{Ax}\|^2 > 0$ dir.
- iii. Eğer \mathbf{A} tam ranklı değilse, bu takdirde Grammian matrisler simetrik pozitif yarı-tanımlı matrislerdir (Meyer, 2000).

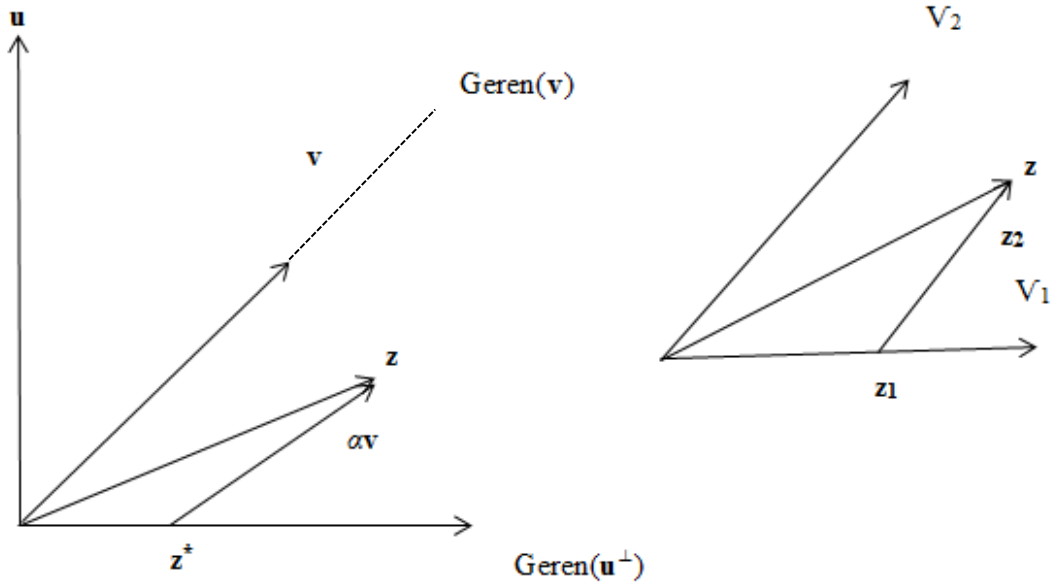
l) **İzdüşüm matrisi** : \mathbf{u} ve \mathbf{v} , $\mathbf{u}'\mathbf{v} \neq 0$ olacak şekilde, \mathbb{R}^m de iki vektör olsun. Bu takdirde, \mathbf{u}^\perp ; \mathbf{u} ya ortogonal olan tüm vektörlerin kümesi olmak üzere,

$$\text{Geren}(\mathbf{v}) \oplus \text{Geren}(\mathbf{u}^\perp) = \mathbb{R}^m \quad (2.46)$$

olduğu gerçekte olabilir. Bu nedenle,

$$\text{Geren}(\mathbf{v}) \cap \text{Geren}(\mathbf{u}^\perp) = \{0\} \quad (2.47)$$

olduğundan, \mathbb{R}^m ; bu altuzayların direkt toplamıdır ($\text{Geren}(\mathbf{v})$ ve $\text{Geren}(\mathbf{u}^\perp)$ altuzaylarının birbirinin diğerine ortogonal olması gerektiğine dikkat edelim.). Bunu görmek için bkz. Şekil 2.5.



Şekil 2.5. Eğik İzdüşümler.

Bir \mathbf{z} vektörünün izdüşümünün sezgisel kavramı \mathbf{v} ye paralel doğrultuda, \mathbf{z} üzerindeki ışık kavramına bağlıdır ve \mathbf{z}^* ; \mathbf{z} nin $\text{Geren}(\mathbf{u}^\perp)$ üzerine düşen gölgesi olsun. Bu takdirde, \mathbf{z}^* a \mathbf{z} nin $\text{Geren}(\mathbf{v})$ boyunca ($\text{Geren}(\mathbf{v})$ ye paralel

olarak) $\text{Geren}(\mathbf{u}^\perp)$ üzerindeki eğik izdüşümü denir. Şekil 2.5 e başvurarak, yorumlamak suretiyle, herhangi bir α reel sayısı için

$$\mathbf{u} \perp \mathbf{z}^* = \mathbf{z} - \alpha \mathbf{v} \quad (2.48)$$

yazar ve buradan da

$$\mathbf{u}'(\mathbf{z} - \alpha \mathbf{v}) = 0 \quad \text{veya} \quad \alpha = \frac{\mathbf{u}'\mathbf{z}}{\mathbf{u}'\mathbf{v}} \quad (2.49)$$

elde ederiz. Bu nedenle,

$$\mathbf{z}^* = \mathbf{z} - \frac{\mathbf{u}'\mathbf{z}}{\mathbf{u}'\mathbf{v}} \mathbf{v} = \mathbf{z} - \frac{\mathbf{v}\mathbf{u}'}{\mathbf{u}'\mathbf{v}} \mathbf{z} = \left[\mathbf{I} - \frac{\mathbf{v}\mathbf{u}'}{\mathbf{u}'\mathbf{v}} \right] \mathbf{z} = \mathbf{P}\mathbf{z} \quad (2.50)$$

dir. Burada

$$\mathbf{P}\mathbf{z} = \left[\mathbf{I} - \frac{\mathbf{v}\mathbf{u}'}{\mathbf{u}'\mathbf{v}} \right] \mathbf{z} \quad (2.51)$$

matrisine, $\mathbf{z} \in \mathbb{R}^m$ vektörlerini $\text{Geren}(\mathbf{v})$ boyunca $\text{Geren}(\mathbf{u}^\perp)$ üzerine izdüşüren, (eğik) izdüşüm matrisi denir.

- i. Bu \mathbf{P} matrisi idempotenttir.
- ii. $(\mathbf{I} - \mathbf{P}) = \frac{1}{\mathbf{u}'\mathbf{v}} \mathbf{v}\mathbf{u}'$ bir ranklı bir dış çarpım matrisidir ve $\mathbf{z} \in \mathbb{R}^m$ vektörlerini $\text{Geren}(\mathbf{u}^\perp)$ boyunca $\text{Geren}(\mathbf{v})$ üzerine izdüşüren izdüşüm matrisidir.
- iii. $(\mathbf{I} - \mathbf{P})$ bir idempotent matristir.
- iv. \mathbf{P} ve bu nedenle $(\mathbf{I} - \mathbf{P})$ simetrik değildirler.

Şimdi yukarıdaki gelişimi genelleştireceğiz. $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m\}$, \mathbb{R}^m de lineer bağımsız vektörlerin bir kümesi olsun. $V_1 = \text{Geren}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ ve $V_2 = \text{Geren}\{\mathbf{v}_{n+1}, \mathbf{v}_{n+2}, \dots, \mathbf{v}_m\}$ olsun.

$$\mathbb{R}^m = V_1 + V_2 \quad \text{ve} \quad V_1 \cap V_2 = \{0\} \quad (2.52)$$

olduğu, yani \mathbb{R}^m nin V_1 ve V_2 nin direkt toplamı olduğu açıktır. Herhangi bir $\mathbf{z}_1 \in V_1$ ve $\mathbf{z}_2 \in V_2$ olmak üzere, herhangi bir $\mathbf{z} \in \mathbb{R}^m$; $\mathbf{z} = \mathbf{z}_1 + \mathbf{z}_2$ olarak ifade edilebilir. $\mathbf{z}_1(\mathbf{z}_2)$ ye \mathbf{z} nin $V_2(V_1)$ boyunca $V_1(V_2)$ üzerine eğik izdüşümü denir (bkz. Şekil 2.5). \mathbf{P}_1 ve \mathbf{P}_2 sırasıyla izdüşüm matrisleri olsun, yani,

$$\mathbf{z}_1 = \mathbf{P}_1 \mathbf{z} \quad \text{ve} \quad \mathbf{z}_2 = \mathbf{P}_2 \mathbf{z} \quad (2.53)$$

olsun. Bu takdirde $\mathbf{P}_1(\mathbf{P}_2)$ ye $V_2(V_1)$ boyunca V_1 (V_2) üzerine eğik izdüşüm matrisleri denir.

$\mathbf{V}_1 = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\} \in \mathbb{R}^{m \times n}$ ve $\mathbf{V}_2 = \{\mathbf{v}_{n+1}, \mathbf{v}_{n+2}, \dots, \mathbf{v}_m\} \in \mathbb{R}^{m \times (m-n)}$ iki matris olsun. Tekil olmayan $\mathbf{V} = [\mathbf{V}_1, \mathbf{V}_2] \in \mathbb{R}^{m \times m}$ matrisini elde ederiz ve

$$\begin{aligned} i = 1, 2, \dots, n \quad \text{için,} \quad \mathbf{P}_1 \mathbf{v}_i &= \mathbf{v}_i, \quad j = n+1, n+2, \dots, m \quad \text{için,} \quad \mathbf{P}_1 \mathbf{v}_j = \mathbf{0} \\ j = n+1, n+2, \dots, m \quad \text{için,} \quad \mathbf{P}_2 \mathbf{v}_j &= \mathbf{v}_j, \quad i = 1, 2, \dots, n \quad \text{için} \quad \mathbf{P}_2 \mathbf{v}_i = \mathbf{0} \end{aligned} \quad (2.54)$$

elde ederiz.

Bu durumda,

$$\mathbf{P}_1 \mathbf{V} = \mathbf{P}_1 [\mathbf{V}_1, \mathbf{V}_2] = [\mathbf{P}_1 \mathbf{V}_1, \mathbf{P}_1 \mathbf{V}_2] = [\mathbf{V}_1, \mathbf{0}] \quad (2.55)$$

ve

$$\mathbf{P}_1 = [\mathbf{V}_1, \mathbf{0}] \mathbf{V}^{-1} = [\mathbf{V}_1, \mathbf{V}_2] \begin{bmatrix} \mathbf{I}_n & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \mathbf{V}^{-1} = \mathbf{V} \Sigma_1 \mathbf{V}^{-1} \quad (2.56)$$

dir. Burada, $\Sigma_1 = \begin{bmatrix} \mathbf{I}_n & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$ dir.

Aynı şekilde

$$\mathbf{P}_2 \mathbf{V} = \mathbf{P}_2 [\mathbf{V}_1, \mathbf{V}_2] = [\mathbf{P}_2 \mathbf{V}_1, \mathbf{P}_2 \mathbf{V}_2] = [\mathbf{0}, \mathbf{V}_2] \quad (2.57)$$

ve

$$\mathbf{P}_2 = [\mathbf{0}, \mathbf{V}_2] \mathbf{V}^{-1} = [\mathbf{V}_1, \mathbf{V}_2] \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{m-n} \end{bmatrix} \mathbf{V}^{-1} = \mathbf{V} \Sigma_2 \mathbf{V}^{-1} \quad (2.58)$$

dir. Burada, $\Sigma_2 = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{m-n} \end{bmatrix}$ dir.

\mathbf{P}_1 ve \mathbf{P}_2 (eğik) izdüşüm matrislerinin aşağıdaki özellikleri kolaylıkla sağlanır.

- i. \mathbf{P}_1 ve \mathbf{P}_2 idempotenttirler.
- ii. $\mathbf{I} = \mathbf{P}_1 + \mathbf{P}_2$ dir.

iii. \mathbf{P}_1 ve \mathbf{P}_2 simetrik değildirler. (Meyer, 2000)

Sonuç 2.1. Her (eğik) izdüşüm matrisi idempotenttir ve tersine olarak her idempotent matris bir eğik izdüşümü tanımlar.

m) Ortogonal İzdüşüm Matris : Şekil 2.5 de $\mathbf{v} = \mathbf{u}$ olduğunda ışık $\mathbf{v} = \mathbf{u}$ ya paralel ışır ve \mathbf{z}^* gölgesine \mathbf{z} nin \mathbf{u}^\perp üzerine ortogonal izdüşümü denir. Bu durumda,

$$\mathbf{P} = \left[\mathbf{I} - \frac{\mathbf{u}\mathbf{u}'}{\mathbf{u}\mathbf{u}} \right] \quad (2.59)$$

olmak üzere,

$$\mathbf{z}^* = \left[\mathbf{I} - \frac{\mathbf{u}\mathbf{u}'}{\mathbf{u}\mathbf{u}} \right] \mathbf{z} = \mathbf{P}\mathbf{z} \quad (2.60)$$

dır.

- i. \mathbf{P} ; Geren(\mathbf{u}^\perp) üzerine bir ortogonal izdüşümdür.
- ii. $\mathbf{P}^2 = \mathbf{P}$ dir.
- iii. $(\mathbf{I} - \mathbf{P})\mathbf{z} = \frac{\mathbf{u}\mathbf{u}'}{\mathbf{u}\mathbf{u}} \mathbf{z}$; Geren(\mathbf{u}) üzerine bir ortogonal izdüşümdür.
- iv. \mathbf{P} bir simetrik matristir.

Buradan çıkan önemli bir sonuç her ortogonal izdüşüm matrisinin idempotent ve simetrik olduğudur. Gerçekten, tersi de doğrudur.

Şimdi ortogonal izdüşüm kavramını genelleştirelim. $V_1 \cap V_2 = \{0\}$ ve $V_1 \perp V_2$,yani V_1 deki her vektör V_2 deki her vektöre, dik olmak üzere, $\mathbb{R}^m = V_1 + V_2$ olsun. Bu durumda,

$V_1 = \text{Geren}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ ye ve $V_2 = \text{Geren}\{\mathbf{v}_{n+1}, \mathbf{v}_{n+2}, \dots, \mathbf{v}_m\}$ ye her biri diğerinin ortogonal tümleyenleridir denir.

Yine $\mathbf{V}_1 = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\} \in \mathbb{R}^{m \times n}$ ve $\mathbf{V}_2 = \{\mathbf{v}_{n+1}, \mathbf{v}_{n+2}, \dots, \mathbf{v}_m\} \in \mathbb{R}^{m \times (m-n)}$ göstererek,

- i. $\mathbf{V}_1' \mathbf{V}_1 \in \mathbb{R}^{n \times n}$ tekil değildir.
- ii. $\mathbf{V}_2' \mathbf{V}_2 \in \mathbb{R}^{(m-n) \times (m-n)}$ tekil değildir.
- iii. $\mathbf{V}_1' \mathbf{V}_2 = 0 = \mathbf{V}_2' \mathbf{V}_1$ dir.

Ne $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ nin ne de $\{\mathbf{v}_{n+1}, \mathbf{v}_{n+2}, \dots, \mathbf{v}_m\}$ nin ortogonal olmasına ihtiyaç duymadığımızı dikkat edelim. İhtiyacımız olan tek şey onların bağımsız ve $V_1 \perp V_2$ olmasıdır. Örneğin,

$$\mathbf{V}_1 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \text{ ve } \mathbf{V}_2 = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}$$

gibi. Eğer \mathbf{P}_1 ve \mathbf{P}_2 ; V_1 ve V_2 altuzayları üzerine ortogonal izdüşüm matrisleri ise, bu takdirde aşağıdaki özellikleri yazabiliriz.

- i. $\mathbf{P}_1 = \mathbf{V}_1 (\mathbf{V}_1' \mathbf{V}_1)^{-1} \mathbf{V}_1' \in \mathbb{R}^{m \times m}$.
- ii. $\mathbf{P}_2 = \mathbf{V}_2 (\mathbf{V}_2' \mathbf{V}_2)^{-1} \mathbf{V}_2' \in \mathbb{R}^{m \times m}$.
- iii. $\mathbf{I} = \mathbf{P}_1 + \mathbf{P}_2$.
- iv. $\mathbf{P}_1^2 = \mathbf{P}_1$ ve $\mathbf{P}_2^2 = \mathbf{P}_2$.
- v. \mathbf{P}_1 ve \mathbf{P}_2 simetriktir.
- vi. Ayrıca, $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ taban vektörleri ortonormal vektörler ve aynı şekilde $\mathbf{v}_{n+1}, \mathbf{v}_{n+2}, \dots, \mathbf{v}_m$ taban vektörleri de ortonormal vektörler ise, bu takdirde $\mathbf{V}_1' \mathbf{V}_1 = \mathbf{I}_n$, $\mathbf{V}_2' \mathbf{V}_2 = \mathbf{I}_{m-n}$, $\mathbf{P}_1 = \mathbf{V}_1 \mathbf{V}_1'$ ve $\mathbf{P}_2 = \mathbf{V}_2 \mathbf{V}_2'$ dır (Meyer, 2000).

Sonuç 2.2. Her ortogonal izdüşüm matrisi simetrik ve idempotenttir ve tersine olarak, her simetrik ve idempotent matris bir ortogonal izdüşüm tanımlar.

2.1.5. Özdeğerler ve Özvektörler

$\mathbf{A} \in \mathbb{R}^{m \times m}$ olsun. Eğer $\mathbf{x} \in \mathbb{R}^m$,

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x} \quad (2.61)$$

olacak şekilde sıfır olmayan bir vektör ve bir λ skaları (reel veya kompleks) var ise, bu takdirde (λ, \mathbf{x}) ikilisine \mathbf{A} nın özdeğeri ve öz vektörü (basitçe öz çifti) denir. Yukarıda sağ yandaki lineer homojen denklem $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$ olduğunda sıfırdan farklı bir çözüme sahiptir. $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$ denkleminde \mathbf{A} matrisinin karakteristik denklemini denir. $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$, λ ya göre m -yinci dereceden bir polinomdur ve karakteristik polinom adını alır. Bu polinomun reel ya da kompleks katlı veya katsız m -tane kökü vardır. Bir kök kompleks ise, onun eşleniği de köktür. Bu nedenle $\mathbf{A} \in \mathbb{R}^{m \times m}$ m -tane özdeğere sahiptir.

Eğer

$$\det(\mathbf{A} - \lambda \mathbf{I}) = (\lambda - \lambda_1)^{m_1} (\lambda - \lambda_2)^{m_2} \dots (\lambda - \lambda_k)^{m_k} \quad (2.62)$$

ise bu takdirde, m_i , λ_i nin katlılığı ve $\sum_{i=1}^k m_i = m$ olmak üzere, λ_1 den λ_k ya kadar k -tane farklı kök vardır. Burada m_i ye λ_i nin cebrik katlılığı denir. \mathbf{A} nın tüm özdeğerlerinin kümesine \mathbf{A} nın spektrumunu (tayfı-izgesi) denir. En büyük özdeğerin büyüklüğüne spektrol (izgesel) yarıçap denir ve $\rho(\mathbf{A})$ ile gösterilir. Şimdi birkaç özelliği yazalım.

- i. Farklı özdeğerler farklı özvektörlere sahiptir. Yani, eğer $(\lambda_1, \mathbf{x}_1)$ ve $(\lambda_2, \mathbf{x}_2)$, $\lambda_1 \neq \lambda_2$ olacak şekilde iseler, bu takdirde \mathbf{x}_1 ve \mathbf{x}_2 lineer bağımsızdırlar.
- ii. Aynı özdeğere karşılık gelen birden daha fazla özvektör olabilir. Örneğin \mathbf{I}_m birim matrisi, \mathbf{I}_m için, 1 yegane özdeğerdir, fakat \mathbf{I}_m m -tane lineer bağımsız özvektöre sahiptir.
- iii. Eğer (λ, \mathbf{x}) ; \mathbf{A} nın bir öz çifti ise, bu takdirde $\mathbf{x} \in \mathcal{N}(\mathbf{A} - \lambda \mathbf{I})$ dır. Burada $\mathcal{N}(\mathbf{A} - \lambda \mathbf{I})$, $(\mathbf{A} - \lambda \mathbf{I})$ nın sıfır uzayıdır. Öz değere karşılık gelen lineer bağımsız \mathbf{x} vektörlerinin maksimal m -sayısı, yani farklılık sayısı, geometrik katlılık olarak bilinir. Bu nedenle geometrik katlılık $\mathcal{N}(\mathbf{A} - \lambda \mathbf{I})$ nın boyutunu gösterir. $\text{boy}[\mathcal{N}(\mathbf{A} - \lambda \mathbf{I})] = \text{rank}(\mathbf{A} - \lambda \mathbf{I})$ olduğunu hatırlayalım. Bu değer $(\mathbf{A} - \lambda \mathbf{I})$ nın sıfırlılığı olarak bilinir.

λ nın geometrik katlılığı \leq λ nın cebrik katlılığı olduğu bir gerçektir.

Yani, kısaca bir λ öz değeri için, onun cebrik katlılığı, karakteristik polinomun bir kökü olarak λ nın katlılığıdır. Geometrik katlılığı ise, ona karşılık gelen lineer bağımsız vektörlerin maksimal sayısıdır.

Örneğin, $\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ matrisi için λ iki katlı bir özdeğerdir. Bu özdeğere

karşılık gelen özvektör $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ olduğundan, $\mathbf{x}_1 = 0$ ve \mathbf{x}_2

herhangi bir vektör olabilir. Bu nedenle, lineer bağımsız sadece 1 vektörü

vardır. Bu vektör de $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ özvektörüdür. $\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ matrisini düşünelim. Bu

matris için de $\lambda = 1$ 2-katlı bir öz değerdır. Fakat, sıfırdan farklı herhangi bir vektör onun özvektörü olarak hizmet görür, herhangi bir $\mathbf{x} \in \mathbb{R}^2$ için

$\mathbf{Ax} = \mathbf{x} = \mathbf{1x}$ yazarız. Bu nedenle, bu matris $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ ve $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ gibi iki lineer

bağımsız özvektöre sahiptir.

Burada her iki matris için $\lambda = 1$, iki cebrik katlı yegane özdeğerdir. Fakat bu özdeğerin geometrik katlılığı $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ matrisi için 1 ve $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ matrisi için 2 dir. Buna göre birinci matris için geometrik katlılık (yani, 1) \leq cebrik katlılık (yani, 2) ve ikinci matris için geometrik katlılık = cebrik katlılık olduğu görülür. Genel olarak

$$\lambda \text{ nın geometrik katlılığı} \leq \lambda \text{ nın cebrik katlılığı}$$

eşitsizliği vardır.

iv. Eğer $\mathbf{B} \in \mathbb{R}^{m \times m}$ tekil olmayan bir matris ise, bu takdirde $\mathbf{B}^{-1}\mathbf{AB}$ matrisi \mathbf{A} ya benzerdir. Benzer matrislerin aynı özdeğerlere sahip oldukları gösterilebilir. Bu nedenle, eğer (λ, \mathbf{x}) ; $\mathbf{B}^{-1}\mathbf{AB}$ nin bir öz çifti ise, bu takdirde (λ, \mathbf{Bx}) ; \mathbf{A} için karşılık gelen öz çifttir.

v. m_i , λ_i nin cebrik katlılığı olmak üzere, $\text{iz}(\mathbf{A}) = m_1\lambda_1 + m_2\lambda_2 + m_k\lambda_k$ dir.

vi. $\det(\mathbf{A}) = \lambda_1^{m_1} \lambda_2^{m_2} \dots \lambda_k^{m_k}$ dir.

vii. **Gerschgorin Çemberleri** : $\mathbf{A} = [a_{ij}]$ olsun. Bir diski veya kompleks düzlemde a_{ii} de merkezleşen ve $\sum_{j \neq i} |a_{ij}|$ yarıçaplı D_i çemberini tanımlayalım. Burada

$$D_i = \left\{ \mathbf{z} \mid |a_{ii} - \mathbf{z}| \leq \sum_{j \neq i} |a_{ij}| \right\}, \quad i = 1, 2, \dots, m \text{ dir. Bu takdirde } \mathbf{A} \text{ nın her özdeğeri,}$$

$S = \bigcup_{i=1}^m D_i$ ile gösterilen D_i disklerinin birleşiminde uzanır.

viii. p cebrik katlılık ve q geometrik katlılık ($q \leq p$) olmak üzere, $\lambda = 0$; $\mathbf{A} \in \mathbb{R}^{m \times m}$ nin bir özdeğeri olsun, bu takdirde

$$\text{rank}(\mathbf{A}) = m - q \geq m - p$$

yani, bir matrisin rankı sıfırdan farklı özdeğerlerinin sayısını geçemez. Örneğin,

$$\mathbf{A} = \begin{bmatrix} 0 & a_1 & a_2 & a_3 \\ 0 & 0 & b_1 & b_2 \\ 0 & 0 & 0 & c_1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

nilpotent matrisine bakılabilir.

ix. Eğer (λ, \mathbf{x}) ; \mathbf{A} için bir öz çift ise, bu takdirde (λ^k, \mathbf{x}) , herhangi bir $k \geq 1$ tamsayısı için, \mathbf{A}^k için karşılık gelen öz çifttir.

- x. Eğer (λ, \mathbf{x}) ; \mathbf{A} nın bir öz çifti ise, bu takdirde $(\lambda + a, \mathbf{x})$; herhangi bir a skaları için $(\mathbf{A} + a\mathbf{I})$ nın karşılık gelen öz çiftidir.
- xi. $p(x) = a_0 + a_1x + a_2x^2$ olsun. Bu takdirde $p(\mathbf{A}) = a_0\mathbf{I} + a_1\mathbf{A} + a_2\mathbf{A}^2$ ye \mathbf{A} ya göre 2. mertebeden bir matris polinomu denir. Eğer (λ, \mathbf{x}) ; \mathbf{A} nın bir öz çifti ise, bu takdirde $(p(\lambda), \mathbf{x})$; $p(\mathbf{A})$ matrisinin bir öz çiftidir.
- xii. Bir reel simetrik matrisin özdeğerleri reeldir.
- xiii. Bir reel simetrik ve pozitif tanımlı matrisin özdeğerlerinin tümü reel ve pozitifdir ve bir pozitif yarı-tanımlı matrisin özdeğerleri reel ve negatif değildir.
- xiv. Eğer \mathbf{A} bir simetrik ve pozitif tanımlı bir matris ise, bu takdirde, $\mathbf{Q}'\mathbf{A}\mathbf{Q} = \Lambda = köşeg(\lambda_1, \lambda_2, \dots, \lambda_m)$ olacak şekilde bir ortogonal \mathbf{Q} matrisi vardır. Burada $\lambda_i > 0$; \mathbf{A} nın özdeğerleridir ve \mathbf{Q} nun sütunları, karşılık gelen özvektörlerin ortonormal kümesidir.
- xv. $\mathbf{Q}'\mathbf{A}\mathbf{Q} = \Lambda$ eşitliğinden, hemen $\mathbf{A} = \mathbf{Q}\Lambda\mathbf{Q}' = \mathbf{Q}\Lambda^{1/2}\Lambda^{1/2}\mathbf{Q}' = (\mathbf{Q}\Lambda^{1/2})(\mathbf{Q}\Lambda^{1/2})' = \mathbf{C}\mathbf{C}'$ elde ederiz. Bu nedenle, eğer \mathbf{A} pozitif tanımlı ise, önemli bir $\mathbf{A} = \mathbf{C}\mathbf{C}'$ parçalanmasını (çarpanlamasını) elde ederiz (Meyer, 2000).

Şimdi özel matrislerin özdeğerlerinin özelliklerini ifade edelim.

a) **Simetrik Matrisler** : $\mathbf{A} \in \mathbb{R}^{m \times m}$ bir simetrik matris olsun.

- i. \mathbf{A} nın özdeğerleri reeldir.
- ii. \mathbf{A} nın sıfırdan farklı olan özdeğerlerinin sayısı = rank(\mathbf{A}) dır.
- iii. Bir reel simetrik matrisin özvektörleri ortogonaldır. Genelliği yitirmeksizin, onların normalleştirildiklerini de kabul edebiliriz (her öz vektörü normuna bölerek).
- iv. $\mathbf{P} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m]$; \mathbf{A} nın m -tane özvektörünün ortonormal matrisini gösterebiliriz. Bu takdirde $\mathbf{A}\mathbf{P} = \mathbf{P}\Lambda$ dır. Burada $\Lambda = köşeg(\lambda_1, \lambda_2, \dots, \lambda_n)$ köşegen matrisi \mathbf{A} nın özdeğerlerinin köşegen matrisidir.
- v. **Spektral (İzgesel) Ayırışım** : \mathbf{P} ortogonal olduğundan, \mathbf{A} yı

$$\mathbf{A} = \mathbf{P}\Lambda\mathbf{P}' = [\mathbf{x}_1 \quad \mathbf{x}_2 \quad \dots \quad \mathbf{x}_m] \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_m \end{bmatrix} \begin{bmatrix} \mathbf{x}'_1 \\ \mathbf{x}'_2 \\ \vdots \\ \mathbf{x}'_m \end{bmatrix}$$

$$= \sum_{i=1}^m \lambda_i \mathbf{x}_i \mathbf{x}'_i = \sum_{i=1}^m \lambda_i \mathbf{P}_i$$

olarak yeniden yazabiliriz. Burada $\mathbf{P} = \mathbf{x}_i \mathbf{x}_i'$, 1 ranklı dış-çarpım matrisleridir.

$$\mathbf{P} = \frac{\mathbf{x}_i \mathbf{x}_i'}{\mathbf{x}_i' \mathbf{x}_i} \quad (\mathbf{x}_i \text{ ler normalleştirildiğinden) nin } \text{Geren}(\mathbf{x}_i) \text{ ne ortogonal}$$

izdüşümü gösterdiğini hatırlayalım. Ortogonal izdüşüm matrislerinin lineer kombinasyonu olarak \mathbf{A} nın bu açılımına \mathbf{A} nın spektral (izgesel) ayrışımı denir (Meyer, 2000).

b) Eşanlı Köşegenleştirilebilirlik : Eğer $\mathbf{A} \in \mathbb{R}^{m \times m}$ ve $\mathbf{B} \in \mathbb{R}^{m \times m}$ iki simetrik matris ise, bu takdirde, \wedge ve Σ sırasıyla \mathbf{A} ve \mathbf{B} nin özdeğerlerinin köşegen matrisleri olmak üzere, kesin olarak $\mathbf{AB} = \mathbf{BA}$ olduğunda,

$$\mathbf{AP} = \mathbf{P} \wedge \quad \text{ve} \quad \mathbf{BP} = \mathbf{P} \Sigma \quad (2.63)$$

olacak şekilde bir ortak \mathbf{P} ortogonal matrisi vardır.

c) Krylov Altuzayı : $\mathbf{y} \in \mathbb{R}^m$ herhangi bir vektör olsun. Bu takdirde, \mathbf{A} nın herhangi $r \geq 0$ tamsayısı için, Krylov altuzayı denilen

$$\mathcal{K}(\mathbf{A}, \mathbf{y}) = \{\mathbf{y}, \mathbf{A}\mathbf{y}, \mathbf{A}^2\mathbf{y}, \dots, \mathbf{A}^{r-1}\mathbf{y}\} \quad (2.64)$$

vektör uzayına ait olan bir \mathbf{x} özvektörü vardır (Golub ve Van Loan 1989; Meyer, 2000).

d) Rayleigh Katsayısı : $\mathbf{A} \in \mathbb{R}^{m \times m}$ bir simetrik matris olsun. Bu takdirde

$$r(\mathbf{x}) = \frac{\mathbf{x}' \mathbf{A} \mathbf{x}}{\mathbf{x}' \mathbf{x}} \quad (2.65)$$

ye Rayleigh katsayısı denir.

Rayleigh Katsayısının Özellikleri :

1) $\mathbf{A} \in \mathbb{R}^{n \times n}$ ve $\mathbf{B} \in \mathbb{R}^{n \times n}$ iki simetrik ve pozitif tanımlı matris olsun.

$$\mathbf{A}\mathbf{x} = \lambda \mathbf{B}\mathbf{x} \quad (2.66)$$

olacak şekilde, $\lambda \in \mathbb{R}$ ve $\mathbf{x} \in \mathbb{R}^n$ olsun. Bu takdirde λ ya genelleştirilmiş özdeğer denir ve \mathbf{x} e (\mathbf{A}, \mathbf{B}) çiftinin karşılık gelen genelleştirilmiş özvektörü denir. (2.66) bağıntısının her iki yanını \mathbf{B}^{-1} ile çarparak, $(\mathbf{B}^{-1}\mathbf{A})\mathbf{x} = \lambda \mathbf{x}$ olduğundan, λ ve \mathbf{x} in $(\mathbf{B}^{-1}\mathbf{A})$ nın özdeğeri ve özvektörü olduğu görülür. Ayrıca, (2.66) nın her iki yanını \mathbf{x}' ile çarparak ve yeniden düzenleyerek, λ nın

$$\lambda = \frac{\mathbf{x}' \mathbf{A} \mathbf{x}}{\mathbf{x}' \mathbf{B} \mathbf{x}} \quad (2.67)$$

Rayleigh katsayısı ile verildiği kolayca anlaşılır.

2) \mathbf{A} ve \mathbf{B} iki simetrik pozitif matris olsun.

$$\begin{aligned} g(\lambda) &= \|(\mathbf{A} - \lambda\mathbf{B})\mathbf{x}\|_{\mathbf{B}^{-1}}^2 \\ &= \mathbf{x}' [(\mathbf{A} - \lambda\mathbf{B})\mathbf{B}^{-1}(\mathbf{A} - \lambda\mathbf{B})] \mathbf{x} \\ &= \mathbf{x}' (\mathbf{A}\mathbf{B}^{-1}\mathbf{A})\mathbf{x} - 2\lambda(\mathbf{x}'\mathbf{A}\mathbf{x}) + \lambda^2(\mathbf{x}'\mathbf{B}\mathbf{x}) \end{aligned} \quad (2.68)$$

fonksiyonunu göz önüne alalım.

$g(\lambda)$ nın sabit (max-min) noktası

$$0 = \frac{dg}{d\lambda} = -2(\mathbf{x}'\mathbf{A}\mathbf{x}) + 2\lambda(\mathbf{x}'\mathbf{B}\mathbf{x})$$

denklemini çözmek suretiyle elde edilir. Yani,

$$\lambda = \frac{\mathbf{x}'\mathbf{A}\mathbf{x}}{\mathbf{x}'\mathbf{B}\mathbf{x}}$$

dir.

$$\frac{d^2g}{d\lambda^2} = \mathbf{x}'\mathbf{B}\mathbf{x} > 0 \quad (2.69)$$

olduğundan, (2.69) daki λ nın $g(\lambda)$ nın minimumlaştırıcısı olduğu görülür. Yani, (2.69) Rayleigh katsayısı (2.68) deki normu minimum yapar.

$$3) \quad r(\mathbf{x}) = \frac{\mathbf{x}'\mathbf{A}\mathbf{x}}{\mathbf{x}'\mathbf{B}\mathbf{x}} \quad (2.70)$$

olsun. Bu takdirde

$$\nabla r(\mathbf{x}) = \frac{2}{(\mathbf{x}'\mathbf{B}\mathbf{x})} [\mathbf{A}\mathbf{x} - r(\mathbf{x})\mathbf{B}\mathbf{x}] \quad (2.71)$$

olduğu gerçeklenebilir. Kesin olarak \mathbf{x} ; $r(\mathbf{x})$, (\mathbf{A}, \mathbf{B}) çiftinin genelleştirilmiş özdeğeri olacak şekilde

$$\mathbf{A}\mathbf{x} = r(\mathbf{x})\mathbf{B}\mathbf{x} \quad (2.72)$$

olduğunda, bu gradyent sıfır olur.

4) $r(\mathbf{x})$; (2.70) ile verilmiş olsun. $\mathbf{B} = \mathbf{G}\mathbf{G}'$, \mathbf{B} nin Cholesky çarpansallaması olsun. Bu

takdirde $\mathbf{y} = \mathbf{G}'\mathbf{x}$ olarak ve (2.70) de yerine koyarak,

$$r(\mathbf{y}) = \frac{\mathbf{y}'\mathbf{A}_1\mathbf{y}}{\mathbf{y}'\mathbf{y}} \quad (2.73)$$

elde ederiz. Burada,

$$\mathbf{A}_1 = \mathbf{G}^{-1}\mathbf{A}(\mathbf{G}')^{-1} \quad (2.74)$$

simetrik ve pozitif tanımlı bir matristir.

(2.70) deki $r(\mathbf{x})$ ve (2.73) deki $r(\mathbf{y})$ arasında bu denklik verildiğinde, genelliği bozmaksızın, (2.73) deki daha basit şekli analiz edebiliriz.

5) $\mathbf{A} \in \mathbb{R}^{n \times n}$ bir simetrik ve pozitif tanımlı matris olsun. Bu takdirde, $\hat{\mathbf{x}}; \mathbf{x}$ in doğrultusundaki birim vektör olmak üzere,

$$r(\mathbf{x}) = \frac{\mathbf{x}'\mathbf{A}\mathbf{x}}{\mathbf{x}'\mathbf{x}} = \frac{\mathbf{x}'\mathbf{A}\mathbf{x}}{\|\mathbf{x}\|^2} = \left(\frac{\mathbf{x}}{\|\mathbf{x}\|} \right)' \mathbf{A} \left(\frac{\mathbf{x}}{\|\mathbf{x}\|} \right) = (\hat{\mathbf{x}})' \mathbf{A}(\hat{\mathbf{x}}) \quad (2.75)$$

dir. Bu nedenle genelliği bozmaksızın,

$$\|\mathbf{x}\| = 1 \text{ olduğunda, } r(\mathbf{x}) = \mathbf{x}'\mathbf{A}\mathbf{x} \quad (2.76)$$

kuadratik formunu göz önüne alalım.

6) $(\lambda_i, \mathbf{P}_i), i = 1, 2, \dots, n$; (2.76) daki \mathbf{A} matrisinin özdeğer/özvektör çifti olsun. Burada

$$\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \dots \geq \lambda_n > 0 \quad (2.77)$$

olduğu kabul edilir. Eğer $\mathbf{P} = [\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_n]$ ve $\Lambda = \text{köşeg}(\lambda_1, \lambda_2, \dots, \lambda_n)$ ise, bu takdirde

$$\mathbf{A}\mathbf{P} = \mathbf{P}\Lambda$$

elde ederiz. \mathbf{P} nin ortogonal ve $\mathbf{P}'\mathbf{P} = \mathbf{P}\mathbf{P}' = \mathbf{I}$ olduğu gerçekleştirilebilir. Bu gerçeği kullanarak,

$$\mathbf{A} = \mathbf{P}\Lambda\mathbf{P}' \text{ ve } \mathbf{P}'\mathbf{A}\mathbf{P} = \Lambda \quad (2.78)$$

olduğu görülür. Bu nedenle, herhangi bir $\mathbf{x} \in \mathbb{R}^n$ vektörü, \mathbf{P} nin sütunlarının bir lineer kombinasyonu olarak ifade edilebilir, yani,

$$\mathbf{x} = \mathbf{P}\mathbf{y} \quad (2.79)$$

dur. (2.79) u yerine koyarak ve (2.78) i kullanarak

$$r(\mathbf{x}) = \mathbf{x}'\mathbf{A}\mathbf{x} = \mathbf{y}'(\mathbf{P}'\mathbf{A}\mathbf{P})\mathbf{y} = \mathbf{y}' \wedge \mathbf{y} = r(\mathbf{y}) \quad (2.80)$$

elde ederiz. Yani

$$r(\mathbf{y}) = \sum_{i=1}^n \lambda_i y_i^2 \quad (2.81)$$

dir. Burada $\mathbf{y} = (y_1, y_2, \dots, y_n)'$

$$\sum_{i=1}^n y_i^2 = 1 \quad (2.82)$$

olacak şekilde bir vektördür. Bunu (2.77) ile birleştirerek

$$\lambda_n \leq \sum_{i=1}^n \lambda_i y_i^2 \leq \lambda_1 \quad (2.83)$$

elde ederiz. Daha kesin olarak,

$$\lambda_1 = \max_{\|\mathbf{y}\|=1} \{\mathbf{y}'\mathbf{A}\mathbf{y}\} = \max_{\|\mathbf{x}\|=1} \{\mathbf{x}'\mathbf{A}\mathbf{x}\} \quad (2.84)$$

ve

$$\lambda_n = \min_{\|\mathbf{y}\|=1} \{\mathbf{y}'\mathbf{A}\mathbf{y}\} = \min_{\|\mathbf{x}\|=1} \{\mathbf{x}'\mathbf{A}\mathbf{x}\} \quad (2.85)$$

elde ederiz.

$r(\mathbf{x})$ ve diğer özdeğerler arasındaki ilişkiyi anlamak için $S_1 = \text{Geren}\{\mathbf{p}_1\}$ ve S_1^\perp ; S_1 e ortogonal tüm vektörlerin kümesini gösterebiliriz. Bu takdirde, $\mathbf{x} \in S_1^\perp$ için, $\mathbf{e}_1 = (1, 0, 0, \dots, 0)'$ olmak üzere,

$$0 = \mathbf{p}_1'\mathbf{x} = \mathbf{p}_1'\mathbf{P}\mathbf{y} = \mathbf{e}_1'\mathbf{P}'\mathbf{P}\mathbf{y} = \mathbf{e}_1'\mathbf{y} \quad (2.86)$$

elde ederiz. Yani \mathbf{y} , birinci bileşeni $y_1 = 0$ olan bir vektördür. Bu nedenle,

$$\lambda_2 = \max_{\substack{\|\mathbf{y}\|=1 \\ y_1=0}} \left\{ \sum_{i=2}^n \lambda_i y_i^2 \right\} = \max_{\substack{\|\mathbf{y}\|=1 \\ y_1=0}} \{\mathbf{y}' \wedge \mathbf{y}\} = \max_{\substack{\|\mathbf{x}\|=1 \\ \mathbf{x} \in S_1^\perp}} \{\mathbf{x}'\mathbf{A}\mathbf{x}\} \quad (2.87)$$

dir. Benzer şekilde, eğer

$$S_j = \text{Geren} \{ \mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_n \}$$

ise, bu takdirde

$$\lambda_j = \max_{\substack{\|\mathbf{x}\|=1 \\ \mathbf{x} \in S_j^\perp}} \{ \mathbf{x}' \mathbf{A} \mathbf{x} \} \quad (2.88)$$

dir. Bu özellikler \mathbf{A} nın özdeğerlerinin nümerik hesabında çok önemli bir rol oynar (Golub ve Van Loan 1989; Meyer, 2000; Trefethen ve Bau, 1997).

e) Ortogonal Matrislerin Özdeğerleri :

i. Eğer $\mathbf{Q} \in \mathbb{R}^{m \times m}$ ortogonal ise, bu takdirde \mathbf{Q} sadece iki $\{+1, -1\}$ özdeğerine sahip olabilir. Gerçekten $\mathbf{Q}\mathbf{x} = \lambda\mathbf{x}$ (λ , \mathbf{Q} nun özdeğeri ve \mathbf{x} de ona karşılık gelen özvektör) olsun.

1) $\mathbf{x}'\mathbf{Q}\mathbf{x} = \lambda\mathbf{x}'\mathbf{x} > 0$ dir. Çünkü $\mathbf{x}'\mathbf{x} > 0$ dir, o halde $\lambda > 0$ dir.

2) $|\lambda|^2 \mathbf{x}'\mathbf{x} = (\mathbf{Q}\mathbf{x})'\mathbf{Q}\mathbf{x} = \mathbf{x}'\mathbf{x}$ dir. Bu nedenle $|\lambda| = 1$ dir. O halde $\lambda = -1$ veya $\lambda = 1$ dir.

ii. \mathbf{Q} nun özvektörleri reeldir ve ortogondur.

f) İzdüşüm Matrislerinin Özdeğerleri :

i. Bir izdüşüm matrisinin özdeğerleri sadece iki $\{0, 1\}$ değerini alabilir.

ii. \mathbf{P} bir izdüşüm matrisi olmak üzere, $\text{iz}(\mathbf{P}) = \text{rank}(\mathbf{P})$ dir.

g) Simetrik ve Pozitif Tanımlı Matrislerin Özdeğerleri : $\mathbf{A} \in \mathbb{R}^{m \times m}$ olsun.

i. Eğer \mathbf{A} simetrik ve pozitif (negatif) tanımlı ise, bu takdirde tüm özdeğerler pozitif (negatif) dir.

ii. Eğer \mathbf{A} simetrik ve pozitif (negatif) yarı-tanımlı ise, bu takdirde özdeğerler negatif değil (pozitif değil) dir.

iii. Eğer özdeğerler eşanlı olarak pozitif/sıfır/negatif değerleri alırsa, \mathbf{A} matrisine tanımsızdır denir.

iv. $\mathbf{A} \in \mathbb{R}^{m \times m}$ simetrik ve pozitif tanımlı bir matris olsun. Bu takdirde

$$\mathbf{Q}(\mathbf{x}) = \mathbf{x}' \mathbf{A} \mathbf{x}$$

denklemini \mathbb{R}^m de bir hiper elipsoidi gösterir. Eğer

$$\wedge = \text{köşeg}(\lambda_1, \lambda_2, \dots, \lambda_n)$$

özdeğerlerin köşegen matrisi ve $\mathbf{P} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m]$, \mathbf{A} nın karşılık gelen özvektörlerinin ortogonal matrisi ise, bu takdirde

$$\mathbf{A}\mathbf{P} = \mathbf{P}\mathbf{\Lambda} \quad \text{ve} \quad \mathbf{A} = \mathbf{P}\mathbf{\Lambda}\mathbf{P}'$$

dür. Eğer $\mathbf{x} = \mathbf{P}'\mathbf{y}$ veya $\mathbf{P}\mathbf{x} = \mathbf{y}$ (\mathbf{P} ortogonal matrisini kullanarak standart koordinat sisteminin bir katı-cisim dönmesi), bu takdirde

$$\mathbf{Q}(\mathbf{x}) = \mathbf{Q}(\mathbf{P}'\mathbf{y}) = \mathbf{y}'\mathbf{\Lambda}\mathbf{y} = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \dots + \lambda_m y_m^2$$

dir. Başka bir deyişle, m -tane ana eksen olan hiper elipsoidde $\mathbf{Q}(\mathbf{x})$; \mathbf{A} nın m -tane ortogonal özvektörü ile rastlaşır ve $1/\sqrt{\lambda_i}$; i -yinci ana doğrultuda yarı eksenin uzunluğunu gösterir (Golub ve Van Loan 1989; Meyer, 2000).

2.1.6. Matris Normları

$\mathbf{A} \in \mathbb{R}^{m \times m}$ olsun. $\mathbf{A} = [a_{ij}]$ diyelim. Kavramsal olarak, matris normunun tanımına yaklaşabileceğimiz iki yol vardır. Bir matematiksel kavram olarak düşünüldüğünde, bunun büyüklüğünü göstermek için normu tanımlayabiliriz. Bir matrisin Frobenius normu aşağıdaki gibi tanımlanır:

$$\|\mathbf{A}\|_F = \left[\sum_{i=1}^m \sum_{j=1}^m a_{ij}^2 \right]^{1/2}. \quad (2.89)$$

Bu, vektörlerin Öklid normu kavramının bir genişletmesidir. Çift yönlü olarak, vektörlerden vektörlere bir dönüşüm olarak bakıldığında, onun vektör normlarını kullanan hareketiyle matris normunu vektör normunu tanımlayabiliriz. p ; 1, 2 veya ∞ u gösterebilir. $\mathbf{x} \in \mathbb{R}^n$ ve $\mathbf{y} = \mathbf{A}\mathbf{x}$ olsun. Bu takdirde \mathbf{y} ye \mathbf{A} altında \mathbf{x} in görüntüsü denir.

$$\|\mathbf{A}\|_p = \sup_{\|\mathbf{x}\|_p \neq 0} \frac{\|\mathbf{A}\mathbf{x}\|_p}{\|\mathbf{x}\|_p} = \sup_{\mathbf{x} \neq 0} \left\| \mathbf{A} \left(\frac{\mathbf{x}}{\|\mathbf{x}\|_p} \right) \right\| = \sup_{\|\mathbf{x}\|_p = 1} \|\mathbf{A}\mathbf{x}\|_p \quad (2.90)$$

tanımını yapalım. Böylece max tüm birim vektörler üzerinden alınmak üzere, \mathbf{A} nın p -normu, \mathbf{y} nin p -normunun \mathbf{x} in p -normuna oranının maksimumudur. Bu nedenle, bu, indüklenmiş norm olarak da bilinir. Eğer $\|\mathbf{A}\mathbf{x}\|_p > 1$ (< 1) ise, bu takdirde \mathbf{A} , \mathbf{x} vektörünün uzunluğunu büyültür. $p = 1, 2$ veya ∞ olduğunda, \mathbf{A} nın 1, 2 veya ∞ normunu elde ederiz. Yukarıdaki tanımdan aşağıdaki yararlı eşitsizlikleri elde ederiz:

$$\begin{aligned}\|\mathbf{Ax}\| &\leq \|\mathbf{A}\|\|\mathbf{x}\| \\ \|\mathbf{AB}\| &\leq \|\mathbf{A}\|\|\mathbf{B}\|.\end{aligned}\tag{2.91}$$

Burada $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m \times m}$ dir. Daha ziyade matris normlarını aşağıdaki gibi doğrudan doğruya hesaplayabiliriz.

N1) $\|\mathbf{A}\|_1 = \max_j \left\{ \sum_i |a_{ij}| \right\}$ -sütun normu denir.

N2) $\|\mathbf{A}\|_\infty = \max_i \left\{ \sum_j |a_{ij}| \right\}$ -sattır normu denir.

N3) $\|\mathbf{A}\|_2 = \sigma_1$, burada σ_1^2 ; \mathbf{A} nın en büyük özdeğeridir. σ_1 in \mathbf{A} nın en büyük tekil değeri olduğunu hatırlayalım.

\mathbf{A} simetrik olduğunda, $\mathbf{A}'\mathbf{A} = \mathbf{A}^2$ ve λ , \mathbf{A} nın bir özdeğeri olmak üzere, λ^2 nin \mathbf{A}^2 nin özdeğeri olması gerçeğinden, $\|\mathbf{A}\|_2 = |\lambda_{\max}|$ olduğu görülür. Burada λ_{\max} , \mathbf{A} nın maksimum özdeğeridir. Ayrıca, bu matris normları eşdeğerdirler ve aşağıdaki eşitsizlikler gerçekleşir:

- i. $\|\mathbf{A}\|_2 \leq \left[\|\mathbf{A}\|_1 \|\mathbf{A}\|_\infty \right]^{1/2}$
- ii. $\frac{1}{\sqrt{m}} \|\mathbf{A}\|_\infty \leq \|\mathbf{A}\|_1 \leq \sqrt{m} \|\mathbf{A}\|_\infty$
- iii. $\frac{1}{\sqrt{m}} \|\mathbf{A}\|_1 \leq \|\mathbf{A}\|_2 \leq \sqrt{m} \|\mathbf{A}\|_1$
- iv. $\|\mathbf{A}\|_2 \leq \|\mathbf{A}\|_F \leq \sqrt{m} \|\mathbf{A}\|_2$.

Bundan başka, $\rho(\mathbf{A})$ izgel (spektral) yarıçapı $\leq \mathbf{A}$ nın herhangi bir matris normu, yani $\rho(\mathbf{A}) \leq \|\mathbf{A}\|$ dir. Bu gerçek, iteratif yöntemlerin yakınsaklığını göstermede çok faydalıdır (Golub ve Van Loan 1989; Meyer, 2000).

Örnek 2.1. $\mathbf{A} = \begin{bmatrix} 5.0 & -2.0 \\ -3.0 & 4.0 \end{bmatrix}$ olsun. Bu durumda $\|\mathbf{A}\|_1 = \max \{8.0, 6.0\} = 8.0$, $\|\mathbf{A}\|_\infty = \max \{7.0, 7.0\} = 7.0$ ve $\|\mathbf{A}\|_F = 7.348$ dir. $\mathbf{A}'\mathbf{A}$ nın özdeğerleri

$$\det[\mathbf{A}'\mathbf{A} - \lambda\mathbf{I}] = \lambda^2 - 54\lambda + 196 = 0$$

denkleminin çözümüyle elde edilir. Bu nedenle $\sigma_1^2 = 50.087$ ve $\sigma_2^2 = 3.913$ ve böylece, $\sigma_1 = 7.07$ ve $\sigma_2 = 1.978$ dir. Bundan dolayı, $\|\mathbf{A}\|_2 = 7.07$ dir. Yine \mathbf{A} nın özdeğerlerinin $(\mathbf{A} - \lambda\mathbf{I}) = \lambda^2 - 9\lambda + 14 = 0$ denkleminin kökleriyle verildikleri

gerçeklenebilir. Bu nedenle, özdeğerler $\lambda_1 = 7.0$ ve $\lambda_2 = 2.0$ ile verilir. Yukarıda verilen tüm eşitsizlikler kolaylıkla gerçekleşir.

Tüm matris normları altında birim matrisin normunun 1 olduğu kolaylıkla gerçekleştirilebilir. Açık olarak, $\|\mathbf{A}\|_1$ ve $\|\mathbf{A}\|_\infty$ normlarının herbiri $(n^2 - n)$ -sayıda toplamı ve $(n-1)$ -sayıda karşılaştırmayı gerektirir. Fakat $\|\mathbf{A}\|_2$ spektral (tayfi) normu hesaplamak oldukça zordur. Bu son hesaplama $O(n^3)$ işlemi gerektirebilir. Bununla beraber spektral normlar, iteratif yöntemlerin yakınsaklığının teorik analizinde çok yararlıdır (Golub ve Van Loan 1989; Meyer, 2000).

Neuman Serileri: Eğer x , $|x| < 1$ olacak şekilde bir reel sayı ise, bu takdirde

$$(1-x)^{-1} = \frac{1}{1-x} = 1 + x + x^2 + x^3 + \dots + \quad (2.92)$$

meşhur bir özdeşliktir. Eğer $\|\mathbf{A}\| < 1$ olacak şekilde, $\mathbf{A} \in \mathbb{R}^{m \times m}$ ise, bu takdirde aynı biçimde

$$(\mathbf{I} - \mathbf{A})(\mathbf{I} + \mathbf{A} + \mathbf{A}^2 + \mathbf{A}^3 + \dots) = \mathbf{I} \quad (2.93)$$

olduğu gerçekleştirilebilir ki, buradan $(\mathbf{I} - \mathbf{A})$ nın tersinin Neuman seri açılımı denilen aşağıdaki eşitliğin varlığı görülür (Meyer, 2000):

$$(\mathbf{I} - \mathbf{A})^{-1} = \mathbf{I} + \mathbf{A} + \mathbf{A}^2 + \dots + \mathbf{A}^k + \dots \quad (2.94)$$

2.1.7. Bir Matrisin Şart Sayısı

$\mathbf{A} \in \mathbb{R}^{m \times m}$ bir matris olsun. Bir matrisin $\kappa(\mathbf{A})$ ile gösterilen şart sayısı

$$\kappa(\mathbf{A}) = \|\mathbf{A}\| \|\mathbf{A}^{-1}\| \quad (2.95)$$

olarak tanımlanır. Burada \mathbf{A}^{-1} , \mathbf{A} nın tersidir. Bu nedenle şart sayısının değeri norma bağlıdır. $\mathbf{I} = \mathbf{A}\mathbf{A}^{-1}$ olduğundan, matris normlarının özelliğini kullanarak,

$$1 = \|\mathbf{I}\| = \|\mathbf{A}\mathbf{A}^{-1}\| \leq \|\mathbf{A}\| \|\mathbf{A}^{-1}\| = \kappa(\mathbf{A}) \quad (2.96)$$

elde ederiz. Bu nedenle, herhangi bir matris için

$$1 \leq \kappa(\mathbf{A}) \leq \infty \quad (2.97)$$

dir. Küçük (büyük) şart sayılı matrislere iyi (zayıf) şartlı denir. \mathbf{A}^{-1} in özdeğerleri \mathbf{A} nın özdeğerlerinin tersleri olduklarından,

$$\|\mathbf{A}^{-1}\|_2 = \max_i \left\{ \frac{1}{|\lambda_i|} \right\} = \frac{1}{\min_i \{|\lambda_i|\}} \quad (2.98)$$

elde ederiz ve spektral şart sayısı

$$\kappa_2(\mathbf{A}) = \|\mathbf{A}\|_2 \|\mathbf{A}^{-1}\|_2 = \frac{\max_i \{|\lambda_i|\}}{\min_i \{|\lambda_i|\}} \quad (2.99)$$

olduğunu görürüz. Çeşitli normlar arasındaki bağıntıyı kullanarak,

$$\begin{aligned} \frac{1}{m} \kappa_2(\mathbf{A}) &\leq \kappa_1(\mathbf{A}) \leq m \kappa_2(\mathbf{A}) \\ \frac{1}{m} \kappa_\infty(\mathbf{A}) &\leq \kappa_2(\mathbf{A}) \leq m \kappa_\infty(\mathbf{A}) \\ \frac{1}{m} \kappa_1(\mathbf{A}) &\leq \kappa_\infty(\mathbf{A}) \leq m^2 \kappa_1(\mathbf{A}) \end{aligned} \quad (2.100)$$

elde ederiz. Matris normlarında olduğu gibi teorik olarak yararlı spektral şart sayısını hesaplamak oldukça zordur. Bununla beraber, $\kappa_1(\mathbf{A})$ ve $\kappa_\infty(\mathbf{A})$ kolayca hesaplanabilir ve bu eşitsizlikleri kullanarak, $\kappa_2(\mathbf{A})$ tahmin edilebilir. $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m \times m}$ ve $\alpha \neq 0$ olmak üzere,

$$\kappa(\mathbf{AB}) \leq \kappa(\mathbf{A})\kappa(\mathbf{B}) \quad (2.101)$$

ve

$$\kappa(\alpha\mathbf{A}) \leq \kappa(\mathbf{A}) \quad (2.102)$$

olduğu gerçekleştirilebilir. Bu nedenle, şart sayısı matrisin ölçeklenmesinden bağımsızdır. Sonuç olarak, $\det(\mathbf{A})$ ve $\kappa(\mathbf{A})$ arasında bağlantı veya ilişkinin olmadığını söyleyelim. $\mathbf{A} \in \mathbb{R}^{m \times m}$ her bir köşegen elemanı $\frac{1}{2}$ olan bir köşegen matrisi olsun. Bu takdirde

$$n \rightarrow \infty \text{ iken } \det(\mathbf{A}) = \frac{1}{2^n} \rightarrow 0 \quad (2.103)$$

dür. Fakat

$$p = 1, 2 \text{ veya } \infty \text{ için } \kappa_p(\mathbf{A}) = 1 \quad (2.104)$$

dür. Aynı şekilde eğer

$$a_{ij} = \begin{cases} 1 & i = j \\ -1 & i > j \\ 0 & i < j \end{cases} \quad (2.105)$$

olacak şekilde $\mathbf{A} \in \mathbb{R}^{m \times m}$ bir üst üçgen matris ise bu takdirde $\det(\mathbf{A}) = 1$ ve $n \rightarrow \infty$ iken $\kappa_\infty(\mathbf{A}) = n2^{n-1} \rightarrow \infty$ dur (Meyer, 2000).

2.1.8. Duyarlılık Analizi : Lineer Sistemlerin Çözümü

Şimdi lineer sistemlerde duyarlılık analizine ilişkin bir temel sonucu ispatsız olarak tekrar edelim.

$$\mathbf{Ax} = \mathbf{b}$$

verilen sistem olsun ve $\kappa(\mathbf{A})$, \mathbf{A} nın şart sayısı olsun. $f \in \mathbb{R}^m$ ve $\mathbf{B} \in \mathbb{R}^{m \times m}$ olsun ve $\varepsilon > 0$ bir reel parametre olsun. \mathbf{y} ,

$$(\mathbf{A} + \varepsilon\mathbf{B})\mathbf{y} = \mathbf{B} + \varepsilon f \quad (2.106)$$

bozuk sisteminin çözümü olsun. $\varepsilon\|\mathbf{B}\|/\|\mathbf{A}\|$ ve $\varepsilon\|f\|/\|\mathbf{b}\|$ nin sırasıyla \mathbf{A} ve \mathbf{b} deki $\varepsilon\mathbf{B}$ ve εf bozulmalarının sebep olduğu göreceli hataları gösterdiğine dikkat edelim. \mathbf{y} çözümündeki göreceli hatanın

$$\frac{\|\mathbf{y} - \mathbf{x}\|}{\|\mathbf{x}\|} \leq \kappa(\mathbf{A}) \left(\varepsilon \frac{\|\mathbf{B}\|}{\|\mathbf{A}\|} + \varepsilon \frac{f}{\|\mathbf{b}\|} \right) \quad (2.107)$$

dir. $\kappa(\mathbf{A}) \geq 1$ olduğunda, \mathbf{A} ve \mathbf{b} deki göreceli hatalar büyür. Bu nedenle \mathbf{A} ve \mathbf{b} üzerinde verilen göreceli hatalar için çözümdeki göreceli hatalar doğrudan doğruya $\kappa(\mathbf{A})$ yla orantılıdır. Bu noktada $\kappa(\mathbf{A})$ nın daha büyük değeri, çözümün duyarlılığıdır (Meyer, 2000).

2.2. Optimizasyon

Lineer eşitlik kısıtlamalarına bağlı çok değişkenli fonksiyonların optimizasyonundan bahsedeceğiz. Değişkenler; fonksiyonlarda vektörlerin, matrislerin ve matrislerin determinantlarının elemanları olarak gözükebilir. Diferensiyelleme için kurallar vereceğiz ve kısıtlanmış/kısıtlanmamış optimizasyon üzerinde klasik sonuçları ifade edeceğiz.

2.2.1. Matris ve Determinantların Diferensiyellenmesi

\mathbf{A} ; a_{ij} elemanları bir θ parametresine bağlı olan bir matris olsun. Bu takdirde,

$$\frac{d\mathbf{A}}{d\theta} = \left[\frac{da_{ij}(\theta)}{d\theta} \right] = \left(\frac{da_{ij}}{d\theta} \right) \quad (2.108)$$

bağıntısını tanımlayalım. Yani bir matrisin türevi, elemanlarının türevlerinin matrisi olarak tanımlanır.

2.2.2. Determinantların Diferensiyellenmesi

A_{ij} , a_{ij} nin eş çarpanını (kofaktörünü) gösterebiliriz. Bu takdirde, herhangi bir i için

$$|\mathbf{A}| = \sum_{i=1}^p a_{ij} A_{ij} \quad (2.109)$$

dur. \mathbf{A} nın elemanlarının fonksiyonel olarak bağımsız olduklarını kabul ederek, aşağıdaki sonuçları elde ederiz:

$$\text{i. } \frac{d|\mathbf{A}|}{da_{ij}} = A_{ij}$$

$$\begin{aligned} \text{ii. } \frac{d|\mathbf{A}|}{d\theta} &= \sum_i \sum_j \left(\frac{d|\mathbf{A}|}{da_{ij}} \right) \frac{da_{ij}}{d\theta} \\ &= \sum_i \sum_j A_{ij} \frac{da_{ij}}{d\theta} \\ &= \text{iz} \left[(A_{ij})' \left(\frac{da_{ij}}{d\theta} \right) \right] \end{aligned}$$

$$\text{iii. Eger } \mathbf{A} \text{ simetrik ise, } \frac{d|\mathbf{A}|}{da_{ij}} = \begin{cases} A_{ii} & i = j \\ A_{ij} & i \neq j \end{cases}$$

$$\text{iv. } \frac{d \ln |\mathbf{A}|}{d\theta} = \left(\frac{1}{|\mathbf{A}|} \right) \frac{d|\mathbf{A}|}{d\theta} = \text{iz} \left(\frac{\mathbf{A}^{-1} d\mathbf{A}}{d\theta} \right)$$

$$\text{v. } \frac{d \text{iz}(\mathbf{A}\mathbf{B})}{d\theta} = \text{iz} \left[\left(\frac{d\mathbf{A}}{d\theta} \right) \mathbf{B} \right] + \text{iz} \left[\mathbf{A} \left(\frac{d\mathbf{B}}{d\theta} \right) \right]$$

$$\text{vi. } \frac{d\mathbf{A}^{-1}}{d\theta} = -\mathbf{A}^{-1} \left(\frac{d\mathbf{A}}{d\theta} \right) \mathbf{A}^{-1}.$$

2.2.3. Bir Fonksiyonun Bir Vektöre Göre Türevi

$\mathbf{t}' = (t_1, t_2, \dots, t_n)$ olsun. Bu durumda eğer, $f(\mathbf{t})$ \mathbf{t} nin bir fonksiyonu ise,

$$\nabla f = \frac{df(\mathbf{t})}{d\mathbf{t}} = \begin{bmatrix} df/dt_1 \\ \vdots \\ df/dt_n \end{bmatrix} \quad (2.110)$$

gradyent vektörünü ve

$$\frac{df(\mathbf{t})}{d\mathbf{t}'} = \left[\frac{df(\mathbf{t})}{d\mathbf{t}} \right]' \quad (2.111)$$

bağıntısını tanımlarız. Aynı zamanda Hessian matrisini-yani, ikinci türevlerin matrisini- de

$$\mathbf{H}_f = \frac{d^2 f(\mathbf{t})}{d\mathbf{t}' d\mathbf{t}} = d \left(\begin{bmatrix} \frac{df(\mathbf{t})}{d\mathbf{t}} \end{bmatrix} \right) = \left(\frac{\partial^2 f}{\partial t_j \partial t_i} \right) \quad (2.112)$$

olarak tanımlarız. \mathbf{H}_f simetriktir (Hocking, 1985).

2.2.4. Bir Fonksiyonun Bir Matrise Göre Diferensiyeli

$\mathbf{X} \quad x_{ij}, i=1,2,\dots,n; j=1,2,\dots,m$ elemanlı $n \times m$ matris olsun. Bu takdirde eğer $f(\mathbf{X})$, \mathbf{X} in elemanlarının bir fonksiyonu ise,

$$\frac{df}{d\mathbf{X}} = \frac{df}{dx_{ij}}, i=1,2,\dots,n; j=1,2,\dots,m \quad (2.113)$$

yi tanımlarız.

2.2.5. Bir Fonksiyonun Optimizasyonu

a) Kısıtlanmamış Durum :

$f(\mathbf{t})$ fonksiyonunun iki kere diferensiyellenebildiğini kabul edeceğiz. $f(\mathbf{t})$ nin sabit (maksimum veya minimum) noktaları

$$\nabla f = 0 \quad (2.114)$$

denklemler sisteminin çözümleridir.

Sonuçlar :

- i. Eğer \mathbf{t}^* , $f(\mathbf{t})$ nin bir göreceli (bağlı) minimumda veya maksimumda aldığı bir noktayı gösterirse, bu takdirde \mathbf{t} , $f(\mathbf{t})$ nin bir sabit noktasıdır.
- ii. Eğer $\mathbf{H}_f(\mathbf{t}^*)$ pozitif (negatif) tanımlı ise, bu takdirde \mathbf{t}^* , $f(\mathbf{t})$ nin bir lokal minimumunu (maksimumunu) tanımlar.

b) Kısıtlanmış Durum

\mathbf{G} q -ranklı ve q -boyutlu olmak üzere, sadece

$$\mathbf{G}\mathbf{t} = \mathbf{g} \quad (2.115)$$

biçiminde yazacağımız lineer kısıtlamaları göz önüne alalım.

c) **İndirgeme Yöntemi** : \mathbf{G}_1 tekil olmamak üzere, $\mathbf{G} = (\mathbf{G}_1/\mathbf{G}_2)$ olsun ve $\mathbf{t} = (t_1, t_2)$ olsun. Bu takdirde kısıtlamaların t_1 için çözümü

$$t_1 = \mathbf{G}_1^{-1}(\mathbf{g} - \mathbf{G}_2 t_2) \quad (2.116)$$

yı verir.

$$f^*(t_2) = f[\mathbf{G}_1^{-1}(\mathbf{g} - \mathbf{G}_2 t_2), t_2] \quad (2.117)$$

olsun. Bu takdirde $f(\mathbf{t})$ nin $\mathbf{G}\mathbf{t} = \mathbf{g}$ ye bağlı sabit noktaları, kısıtlamaları tarafından belirlenen t_1 in değerlerine bağlı $f^*(t_2)$ nin sabit noktaları vasıtasıyla belirlenir. $f(\mathbf{t})$ nin $\mathbf{G}\mathbf{t} = \mathbf{g}$ ye bağlı kısıtlanmış maksimizasyonu (minimizasyonu) $f^*(t_2)$ nin kısıtlanmamış maksimizasyonu (minimizasyonu) na denktir (Hocking, 1985).

d) **Lagrange Çarpanı Yöntemi** : λ Lagrange çarpanlarının bir q vektörü olmak üzere, Lagrange fonksiyonu

$$F(\mathbf{t}, \lambda) = f(\mathbf{t}) + \lambda'(\mathbf{G}\mathbf{t} - \mathbf{g}) \quad (2.118)$$

ile tanımlanır.

$F(\mathbf{t}, \lambda)$ nin sabit noktaları

$$\nabla_{\mathbf{t}} F = \nabla f + \mathbf{G}'\lambda = 0 \quad (2.119)$$

ve

$$\nabla_{\lambda} F = \mathbf{G}\mathbf{t} - \mathbf{g} = 0 \quad (2.120)$$

ile verilir. Eğer $(\mathbf{t}^*, \lambda^*)$; $F(\mathbf{t}, \lambda)$ nin bir sabit noktası ve $\mathbf{H}_f(\mathbf{t}^*)$ pozitif (negatif) tanımlı ise, bu takdirde \mathbf{t}^* , $f(\mathbf{t})$ nin $\mathbf{G}\mathbf{t} = \mathbf{g}$ ye bağlı bir lokal (gerçek) minimumunu (maksimumunu) tanımlar.

2.3. İstatistiksel Gerçekler

Gelecek bölümlere yardımcı olması açısından, tahmin ve hipotez test etmenin klasik teorisindeki bazı temel sonuçların kısa bir özetini verelim.

2.3.1. Tahmin

a) Olabilirlik Fonksiyonu

\mathbf{y} yoğunluğu bir θ parametre vektörüne bağlı olan bir rasgele vektör olsun. \mathbf{y} nin $f(\mathbf{y})$ diyeceğimiz yoğunluk fonksiyonuna θ nın bir fonksiyonu olarak bakıldığında, $L(\theta)$ olabilirlik fonksiyonudur.

b) En Çok Olabilirlik Tahmini

$L(\theta)$ nın θ ya göre maksimumlaştırılması

$$\max_{\theta} L(\theta) \quad (2.121)$$

ile gösterilir. Bu problemin çözümleri \mathbf{y} nin $\theta(\mathbf{y})$ diyeceğimiz fonksiyonlarıdır ve $\theta(\mathbf{y})$ ler en çok olabilirlik tahmin edicileridir. Bir özel vektör için $\theta(\mathbf{y})$ nin aldığı değer, $\hat{\theta}$ veya $\tilde{\theta}$ ile göstereceğimiz bir en çok olabilirlik tahminini verir. Maksimum değer $L(\theta)$ nın sabit noktalarında veya buna eşdeğer olarak logaritmik dönüşümün, yani $\ln L(\theta)$ nın sabit noktalarında ortaya çıkar. Genel olarak, maksimum ediciler bir kapalı form ifadeye sahip olmayabilirler. Fakat sayısal olarak belirlenebilmektedirler.

Böyle durumlarda mutlak dağılımsal sonucu belirlemek zordur.

\mathbf{I} , (i, j) -yinci elemanı

$$\mathbf{I}_{ij} = -E \left[\frac{\partial^2 \ln L(\theta)}{\partial \theta_i \partial \theta_j} \right] \quad (2.122)$$

olan bilgi matrisi olmak üzere, en çok olabilirlik tahmin edicisinin asimptotik olarak

$$\hat{\theta} \sim N(\theta, \mathbf{I}^{-1}) \quad (2.123)$$

biçiminde dağıldığı bilinmektedir. θ_i nin yansız tahminlerinin varyansı üzerindeki bir alt sınırı, \mathbf{I}^{-1} in i -yinci köşegen elemanı ile verilir. Gramer-Rao alt sınırı \mathbf{I}_{ii} nin tersiyle verilir (Hocking, 1985).

c) Kısıtlanmış Tahmin

Bazı durumlarda, θ nin tüm değerleri maksimumlayıcılar olarak kabul edilemezler, bu nedenle, Θ , Öklid uzayının bir alt kümesi olmak üzere,

$$\max_{\theta \in \Theta \text{ kısıtlaması altında}} L(\theta) \quad (2.124)$$

kısıtlanmış maksimumlaştırma problemini çözmemiz gerekir. Bazen bu kısıtlamalar kendiliğinden sağlanır. Örneğin, $N(\boldsymbol{\mu}, \boldsymbol{\sigma}^2)$ den bir rasgele örneğe bağlı $\boldsymbol{\sigma}^2$ nin tahmininin negatif olmayacağı garanti edilir. Diğer durumlarda maksimumlaştırma kısıtlanmış olmalıdır. En basit durum, uygun değerlerin modeldeki parametrelerin sayısını azaltmak için kullanılabilen eşitlik kısıtlamaları ile tanımlandığı durumdur. Kısıtlamalar lineer olduklarından, durum özellikle basittir.

Diğer durumlarda kısıtlama uzayı eşitsizliklerle- örneğin, $\theta_i \geq 0$, $\theta_i \geq \theta_j$ v.s., tanımlanabilir.

d) Tam, Yeterli İstatistikler

$T_i(\mathbf{y})$, $i = 1, 2, \dots, p$, $T_i(\mathbf{y})$ verildiğinde \mathbf{y} nin şartlı dağılımı $\theta_1, \theta_2, \dots, \theta_n$ parametrelerine bağlı olmayacak şekilde, istatistiklerin bir kümesi olsun. Bu takdirde $T_i(\mathbf{y})$, $i = 1, 2, \dots, p$; $\theta_1, \theta_2, \dots, \theta_n$ için birlikte yeterlidir. T nin sıfır beklenen değerli fonksiyonunun olmaması anlamında tam olan yeterli istatistiklerin en küçüğü ile ilgileneceğiz. Hemen hemen her yerde sıfır olan fonksiyonları kapsam dışında bırakıyoruz.

e) Üstel Aile

θ parametrelerin bir \mathbf{p} vektörü olmak üzere, \mathbf{y} nin yoğunluğunun

$$f(\mathbf{y}, \theta) = a(\theta)b(\mathbf{y}) \exp \left[\sum_{j=1}^p c_j(\theta)T_j(\mathbf{y}) \right] \quad (2.125)$$

biçiminde yazılabildiğini ve değişkenler üzerindeki aralığın θ ya bağlı olmadığını farz edelim. Bu takdirde yoğunluk üstel aileye aittir ve bazı şartlar altında, $T_j(\mathbf{y})$, $j = 1, 2, \dots, p$ istatistikleri θ için tam, yeterli istatistiklerdir. Bu sonun bir genişlemesi, Gautschi (1950)'ye göre, dağılımların ailesinin

$$f(\mathbf{y}, \theta, \boldsymbol{\mu}) = a(\theta, \boldsymbol{\mu})b(\mathbf{y}) \exp \left[\sum_{j=1}^p c_j(\theta) T_j(\mathbf{y}) + T_{p+1}(\mathbf{y}) d(\theta, \boldsymbol{\mu}) + T_{p+1}^2(\mathbf{y}) h(\theta) \right] \quad (2.126)$$

ile verildiğini ve aynı zamanda tam olduğunu gösterir.

Aşağıdaki iki teorem bu özelliğe sahip istatistiklerin önemini vurgular.

f) Rao-Blackwell Teoremi

Eğer $g(\mathbf{y})$, $h(\theta)$ diyeceğimiz bir skalar parametre fonksiyonunun bir yansız tahmin edicisi ve T_i , $i=1,2,\dots,p$; θ için birlikte yeterli ise, $E[P(T)] = h(\theta)$ ve $\text{var}[P(T)] \leq \text{var}[g(\mathbf{y})]$ olacak şekilde yalnız yeterli istatistikler vasıtasıyla veriye bağlı olan bir $P(T)$ tahmin edicisi vardır. Ayrıca, bu tahmin edici $g(\mathbf{y})$ nin T_i üzerindeki şartlı beklenen değerini, yani, $P(T) = E[g(\mathbf{y})/T_1, T_2, \dots, T_p]$ yi belirlemek suretiyle oluşturulabilir (Hocking, 1985).

g) Lehmann-Scheffe Teoremi

Eğer T_i , $i=1,2,\dots,p$ ler θ için tam geçerli istatistikler iseler, bu takdirde Rao-Blackwell Teoreminden $P(T)$ tektir ve bu nedenle $h(\theta)$ nın minimum varyans, yansız tahminini verir (Hocking, 1985).

2.3.2. Hipotezlerin Testleri ve Güven Aralıkları

Bu tezde testler genellikle olabilirlik oranı prensibini kullanarak ortaya konacaktır. θ , $\theta \in \Theta$ ile kısıtlanmış parametrelerin vektörü olmak üzere, $L(\theta)$ olabilirlik fonksiyonu olsun. Θ_1 , Θ nın bir alt kümesi olmak üzere, $H_A: \theta \notin \Theta_1$ alternatif hipotezine karşı $H_0: \theta \in \Theta_1$ hipotezini test etmek istediğimizi varsayalım. Bu takdirde test istatistiği

$$\gamma(\mathbf{y}) = \frac{\max_{\theta \in \Theta_1} L(\theta)}{\max_{\theta \in \Theta} L(\theta)} \quad (2.127)$$

oranı ile verilir. γ nın payı hipotezin kısıtlamaları ve model ifadesinde içerilen parametreler üzerindeki doğal kısıtlamalar altında olabilirlik fonksiyonunun maksimum değeridir. Payda maksimizasyonu sadece son söylenen kısıtlamayla sınırlıdır. γ nın H_0 in reddedilmesini öneren küçük değerleri ile $0 \leq \gamma \leq 1$ olduğu görülür. $\gamma(\mathbf{y}) \leq \gamma^*$ olduğunda eğer H_A nın lehine H_0 ı reddedersek, bu takdirde testin boyutu, yani α ,

$$P[\gamma(\mathbf{y}) \leq \gamma^* | \theta \in \Theta_1] \leq \alpha \quad (2.128)$$

ile tanımlanır. θ nın bir fonksiyonu olarak H_0 ı reddetmenin olasılığına testin gücü denir ve

$$\Pi(\theta) = P[\gamma(\mathbf{y}) \leq \gamma^* | \theta] \quad (2.129)$$

ile belirlenir. Genel olarak, $\gamma(\mathbf{y})$ nin mutlak küçük örneklem dağılımını elde etmek zordur. Büyük örneklem teorisine dayanan bir yaklaşım sonucu H_0 altında

$$-2Ln\gamma(\mathbf{y}) \sim \chi^2(\nu) \quad (2.130)$$

ile verilir. Burada ν , hipotezdeki bağımsız kısıtlamaların sayısını gösterir.

Test istatistiğinin tersini almak suretiyle θ üzerinde (veya θ nın fonksiyonları üzerinde) güven bölgelerini ortaya koyacağız, yani θ üzerinde - $R(\mathbf{y})$ diyeceğimiz bir güven bölgesi \mathbf{y} verisiyle tanımlanan, eğer $\theta^* \in R(\mathbf{y})$ ise, $H_0 : \theta = \theta^*$ hipotezinin \mathbf{y} ye dayanarak kabul edilebilmesi özelliğine sahip olan bir bölge olacaktır (Hocking, 1985).

3. MATERYAL ve YÖNTEM

3.1. İstatistikte Lineer Modeller

Bilimsel yöntem genellikle, öğrenmeye götüren bir yaklaşım olarak kullanılır. Lineer istatistiksel yöntemler bu öğrenme yöntemlerinin parçası olarak yaygın bir şekilde kullanılır. İşletme ve mühendisliğin yanı sıra, biyolojik, fiziksel ve sosyal bilimlerde lineer modeller, hem araştırmanın planlanma aşamalarında hem de sonuçta ortaya konan verinin analizinde yararlıdırlar.

Burada basit ve çoklu lineer regresyon modellerini kısaca tanıtacağız.

3.1.1. Basit Lineer Regresyon Modeli

Basit lineer regresyonda, iki değişken arasındaki ilişkiyi, örneğin, gelir ve eğitim yıllarının sayısı, insanların boyu ve ağırlığı, zarfların uzunluğu ve genişliği, sıcaklık ve bir sanayi sürecinin çıktısı, yükseklik ve suyun kaynama noktası veya bir ilacın dozu ve etkisi arasındaki ilişkiyi modellemek için çalışırız. Bir lineer ilişki için,

$$\mathbf{y} = \beta_0 \mathbf{j}_n + \beta_1 \mathbf{x} + \boldsymbol{\varepsilon} \quad (3.1)$$

biçiminde bir modeli kullanabiliriz. Burada \mathbf{y} bağımlı veya tepki değişkenlerinin $(n \times 1)$ vektörüdür ve \mathbf{x} bağımsız veya açıklayıcı değişkenlerin bir $(n \times 1)$ vektörüdür. $\boldsymbol{\varepsilon}$ modeldeki hata terimlerinin bir $(n \times 1)$ rasgele hata vektörüdür. Bu kapsamda hata yanlış anlamına gelmez, fakat rasgele fonksiyonların, yani ölçüm hatalarının veya kontrolümüzün dışındaki faktörlerin (etmenlerin) etkisini gösteren bir istatistiksel terimdir. Genel olarak $\boldsymbol{\varepsilon}$ hata vektörünün $\mathbf{0}$ ortalamalı ve $\sigma^2 \mathbf{I}_n$ varyans-kovaryans matrisli bir normal dağılıma sahip olduğu, yani, $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_n)$ olduğu kabul edilir. β_0 ve β_1 , modelin tahmin edilecek parametreleridir. β_0 a regresyon sabiti (intercept) ve β_1 e eğim parametresi denir.

(3.1) deki modelin lineerliği (doğrusallığı) bir varsayımdır. Hata terimlerinin dağılımı, \mathbf{y} nin gözlenen değerlerinin bağımsızlığı vs. hakkındaki diğer varsayımları genel anlamda ekleriz. \mathbf{x} ve \mathbf{y} nin gözlenen değerlerini kullanarak, β_0 ve β_1 i tahmin ederiz ve β_0 ve β_1 için güven aralıkları ve hipotezlerin testleri gibi çıkarımlar yaparız. \mathbf{x} in özel bir değeri için \mathbf{y} de tahmin edilebilir.

3.1.2. Çoklu Lineer Regresyon Modeli

y_i cevap değişkenleri çoğu kez bir açıklayıcı değişkenden daha fazlası tarafından etkilenir. Örneğin, bir mahsulün verimi, kullanılan azot, potasyum ve fosfor gübrelerinin miktarına bağlı olabilir. Bu değişkenler deneyci tarafından kontrol

edilir, fakat verim havayla ilişkili kontrol edilemeyen değişkenlere de bağlı olabilir. y_i cevap değişkenlerini birçok açıklayıcıya bağlayan bir lineer model:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip} + \varepsilon_i, \quad i = 1, 2, \dots, n \quad (3.2)$$

biçimine sahiptir. $\beta_0, \beta_1, \beta_2, \dots, \beta_p$ parametrelerine regresyon katsayıları denir. (3.1) deki gibi, ε_i ; y_i de \mathbf{x}_k ($k = 1, 2, \dots, p$) değişkenleriyle açıklanamayan değişimi verir. Bu ε_i rasgele değişkeni, y_i ye etki eden fakat bilinen veya gözlenemeyen diğer değişkenlere kısmen bağlı olabilir.

(3.2) deki model $\beta_0, \beta_1, \beta_2, \dots, \beta_p$ parametrelerine göre lineerdir; \mathbf{x}_k değişkenlerine göre lineer olması gerekmez.

Bu nedenle,

$$\mathbf{y} = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i1}^2 + \beta_3 x_{i2} + \beta_4 \sin x_{i2} + \varepsilon_i, \quad (i = 1, 2, \dots, n) \quad (3.3)$$

gibi modeller, lineer model gösterimin kapsamına alınır.

Bir model, bir doğal olayın daha iyi anlaşılması için bir teorik çatı temin eder. Bu nedenle bir model, eldeki gözlemleri üreten mekanizmayı temsil edebildiğine inandığımız bir matematiksel kurgudur. Kabul edilen model karmaşık gerçek dünya durumunun ideal duruma getirilmiş bir aşırı basitleştirilmesi olabilir, fakat böyle birçok durumlarda deneysel modeller, değişkenler arasındaki ilişkilerin yararlı yaklaşımlarını verir. Bu ilişkiler çağrışımsal veya nedensel olabilir.

(3.2) deki gibi regresyon modelleri, aşağıda içerilen çeşitli amaçlar için kullanılır:

1. **Tahmin.** Tahmin (öngörü) için ayrı ayrı β_l ($l = 0, 1, 2, \dots, p$) parametrelerinin ayrı ayrı tahminleri \mathbf{x}_k değişkenlerinin y_i üzerindeki toplu etkisinden daha az öneme sahiptir. Bununla beraber, iyi öngörü performansını elde etmek için iyi tahminlere ihtiyaç vardır.
2. **Veri Tanımlaması veya Anlamlandırma.** Bilim insanı veya mühendis, gözlenen verileri özetlemek veya tanımlamak için tahmin edilen modeli kullanır.
3. **Parametre Tahmini.** Kabul edilen bir model için tahmin edilen parametrelerin değerleri teorik çıkarımsallara (sonuçlara) sahip olabilir.
4. **Değişken Seçimi veya Ayırma.** Buradaki vurgu, y_i deki değişimi modellemede her bir öngörü değişkeninin belirlenmesi hakkındadır. y_i deki değişimin önemli bir miktarıyla ilişkili olan öngörücülere ulaşılır, daha az katkıda bulunanlar çıkarılır.
5. **Çıktının Kontrolü.** y_i ve \mathbf{x}_k değişkenleri arasında bir sebep-ve-etki ilişkisi kabul edilir. Bu takdirde tahmin edilen model, girdileri değiştirmek

suretiyle bir işlemin çıktısını kontrol etmek için kullanılabilir. Sistematik deneye göre, optimal çıktıya ulaşmak mümkün olabilir.

1-5 amaçları arasında bir temel fark vardır. Öngörü için, öngörülerin yapılacağı yerde veriler sürekli toplandığında da sadece baskın gelen aynı ilişkileri kabul etmemiz gerekir. (3.2) deki y_i ve \mathbf{x}_k değişkenleri arasında anlamlı bir ilişkinin varlığını gösterme, nedensel olan ilişkiyi ispatlamayı gerektirmez. Çıktıyı kontrol etmek için nedenselliği saptamada, araştırmacı modeldeki \mathbf{x}_k değişkenlerinin değerlerini seçmeli ve hesaba katılmayan diğer mümkün değişkenlerin etkilerinden sakınmak için rasgeleleştirmeyi kullanmalıdır. Diğer bir deyişle, \mathbf{x}_k değişkenleri değiştirildiğinde, \mathbf{x}_k değişkenlerinin y_i üzerindeki etkisini belirlemek için, onları değiştirmek gerekir. Yukarıda listelenen beş amaca katkı yapan tahmin ve çıkarımsal yöntemler Rencher ve Schaalje, 2007 de ele alınmıştır.

β lara göre lineer olan fakat \mathbf{x}_k lere göre lineer olmayan bir modelin örneği;

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i1}^2 + \beta_4 \sin x_{i2}^2 + \varepsilon_i, \quad (i = 1, 2, \dots, n) \quad (3.4)$$

ikinci mertebeden tepkime yüzeyi modelidir.

(3.2) deki β ları tahmin etmek için y_i ve ilişkili \mathbf{x}_k değişkenleri üzerindeki n - gözleme sahip olan bir örnekleme kullanacağız. ε_i ya da y_i için varsayımlar aşağıda verilmiştir.

1. $i = 1, 2, \dots, n$ için, $E(\varepsilon_i) = \mathbf{0}$ ya da eşdeğer olarak $E(y_i) = \hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2} + \dots + \hat{\beta}_p x_{ip}$ dir.
2. $i = 1, 2, \dots, n$ için, $\text{var}(\varepsilon_i) = \sigma^2$ ya da eşdeğer olarak $\text{var}(y_i) = \sigma^2$ dir.
3. Her $i \neq j$ için $\text{kov}(\varepsilon_i, \varepsilon_j) = \mathbf{0}$ ya da eşdeğer olarak $\text{kov}(y_i, y_j) = \mathbf{0}$ dir.

Varsayım 1, modelin doğru olduğunu, diğer bir deyişle konuyla ilgili tüm \mathbf{x}_k ların hesaba katıldığını ve modelin gerçekten lineer olduğunu ifade eder. Varsayım 2, y_i nin varyansının sabit olduğunu ve bu nedenle \mathbf{x}_k lara bağlı olmadığını iddia eder. Varsayım 3, y_i lerin diğer her biri ile ilişkisiz olduklarını ifade eder ki, bu; bir rasgele örneklemede genel olarak gerçekleşir (Bir zaman serisinde veya tekrarlı ölçümler bir tek bitki ya da hayvan üzerinde yapıldığında gözlemler genellikle ilişkili olacaktır.). Şimdi bunlara, bir normallik varsayımını da ekleyeceğiz.

Buraya kadar y_1, y_2, \dots, y_n rasgele değişkenleri hakkında normallik varsayımı yapmadık. Kısım 3.1.2. deki varsayımlara şimdi,

$$\mathbf{y}, N(\mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I}_n) \text{ veya } \boldsymbol{\varepsilon}, N(\mathbf{0}, \sigma^2\mathbf{I}_n) \quad (3.5)$$

varsayımını ekleyeceğiz. Normallik varsayımını altında $\sigma_{ij} = 0$, ilişkisizliğin yanı sıra \mathbf{y} (veya $\boldsymbol{\varepsilon}$) değişkenlerinin bağımsız olduklarını ifade eder.

Normallik varsayımını altında \mathbf{y} değişkeni, ilişkisiz olmanın yanı sıra bağımsız olur.

(3.2) bağıntısını, n -tane gözlemin her birisi için yazarak

$$\begin{aligned} y_1 &= \beta_0 + \beta_1 x_{11} + \beta_2 x_{12} + \dots + \beta_p x_{1p} + \varepsilon_1 \\ y_2 &= \beta_0 + \beta_1 x_{21} + \beta_2 x_{22} + \dots + \beta_p x_{2p} + \varepsilon_2 \\ &\vdots \\ y_n &= \beta_0 + \beta_1 x_{n1} + \beta_2 x_{n2} + \dots + \beta_p x_{np} + \varepsilon_n \end{aligned} \quad (3.6)$$

elde ederiz. Bu n -denklemin matris formunda,

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1p} \\ 1 & x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{np} \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix} \quad (3.7)$$

veya

$$\mathbf{y} = \mathbf{X}_1 \boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (3.8)$$

olarak yazılabilir. ε_i ya da y_i üzerinde yukarıda verilen üç varsayım, (3.8) deki modele göre aşağıdaki şekilde ifade edilebilir.

1. $E(\boldsymbol{\varepsilon}) = \mathbf{0}$ ya da $E(\mathbf{y}) = \mathbf{X}\boldsymbol{\beta}$ dır.
2. $kov(\boldsymbol{\varepsilon}) = \sigma^2\mathbf{I}$ ya da $kov(\mathbf{y}) = \sigma^2\mathbf{I}$ dır.

$kov(\boldsymbol{\varepsilon}) = \sigma^2\mathbf{I}$ varsayımının, önceki $var(\varepsilon_i) = \sigma^2$ ve $kov(\varepsilon_i, \varepsilon_j) = \mathbf{0}$ varsayımlarının her ikisini de içerdiğine dikkat edelim.

(3.8) deki \mathbf{X}_1 matrisi, $n \times (p+1)$ dir. $n > (p+1)$ ve $rank(\mathbf{X}_1) = p+1$ olduğunu kabul edeceğiz. Eğer $n < (p+1)$ ya da \mathbf{X}_1 matrisinin sütunları arasında bir lineer ilişki varsa, bu takdirde \mathbf{X}_1 tam sütun ranklı olmayacaktır. Eğer x_{ij} lerin değerleri planlanır (araştırmacı tarafından seçilir) ise bu takdirde \mathbf{X}_1 matrisi özde deneysel tasarımı içerir ve \mathbf{X}_1 e bazen tasarım matrisi de denir.

(3.2) deki ya da (3.8) deki $\boldsymbol{\beta}$ parametrelerine regresyon katsayıları denir. Onların ortak etkisini vurgulamak için onlara bazen kısmi regresyon katsayıları olarak

başvurulur. “ Kısmi” sözü hem matematiksel hem de istatistiksel anlam taşır. Matematiksel olarak örneğin, $E(\mathbf{y}) = \hat{\beta}_0 + \hat{\beta}_1\mathbf{x}_1 + \hat{\beta}_2\mathbf{x}_2 + \dots + \hat{\beta}_p\mathbf{x}_p$ nin \mathbf{x}_1 e göre türevi $\hat{\beta}_1$ dir. Bu nedenle $\hat{\beta}_1$; $\mathbf{x}_2, \mathbf{x}_3, \dots, \mathbf{x}_p$ sabit tutulduğunda, \mathbf{x}_1 deki bir birim artma ile $E(\mathbf{y})$ deki değişimi gösterir. İstatistiksel olarak, $\hat{\beta}_1$; diğer \mathbf{x} lerin varlığında \mathbf{x}_1 in $E(\mathbf{y})$ üzerindeki etkisini gösterir. Eğer \mathbf{x} ler modelde mevcut olmasa idi, bu etki genel olarak \mathbf{x} in $E(\mathbf{y})$ üzerindeki etkisinden farklı olacaktı. Bu nedenle, örneğin,

$$\mathbf{y} = \beta_0 + \beta_1\mathbf{x}_1 + \beta_2\mathbf{x}_2 + \boldsymbol{\varepsilon} \quad (3.9a)$$

modelindeki β_0 ve β_1 genel olarak

$$\mathbf{y} = \beta_0^* + \beta_1^*\mathbf{x}_1 + \boldsymbol{\varepsilon}^* \quad (3.9b)$$

modelindeki β_0^* ve β_1^* dan farklı olur.

[\mathbf{x}_1 ve \mathbf{x}_2 , \mathbf{X}_1 matrisindeki sütunlar olmak üzere, eğer \mathbf{x}_1 ve \mathbf{x}_2 ortogonal iseler, yani $\mathbf{x}_1'\mathbf{x}_2 = 0$ ya da $(\mathbf{x}_1 - \bar{\mathbf{x}}_1\mathbf{j})'(\mathbf{x}_2 - \bar{\mathbf{x}}_2\mathbf{j}) = 0$ ise bu takdirde $\beta_0 = \beta_0^*$ ve $\beta_1 = \beta_1^*$ dir (Rencher ve Schaalje, 2007).]

3.1.2.1. $\boldsymbol{\beta}$ nin ve σ^2 nin Tahmini

\mathbf{X} bilinenlerin p ranklı bir $n \times p$ matrisi, \mathbf{y} gözlenebilir bir $n \times 1$ rasgele vektör, $\boldsymbol{\beta}$ tahmin edilecek parametrelerin bir $p \times 1$ vektörü, $\boldsymbol{\varepsilon}$, $E(\boldsymbol{\varepsilon}) = \mathbf{0}$ ortalamalı ve $kov(\boldsymbol{\varepsilon}) = \sigma^2\mathbf{I} = \Sigma$ kovaryanslı bir rasgele hata vektörü ve σ^2 bilinmeyen pozitif bir parametre olmak üzere $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ genel lineer modelini göz önüne alalım.

Bilindiği gibi $\boldsymbol{\beta}$ nin alışılmış en küçük kareler tahmin edicisi,

$$\begin{aligned} SSE(\boldsymbol{\beta}) &= \boldsymbol{\varepsilon}'\boldsymbol{\varepsilon} = (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) \\ &= \mathbf{y}'\mathbf{y} - \boldsymbol{\beta}'\mathbf{X}'\mathbf{y} - \mathbf{y}'\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta} \\ &= \mathbf{y}'\mathbf{y} - 2\boldsymbol{\beta}'\mathbf{X}'\mathbf{y} + \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta} \end{aligned} \quad (3.10)$$

ile tanımlanan hata kareler toplamını $\boldsymbol{\beta}$ ya göre minimumlaştırarak elde edilir. $SSE(\boldsymbol{\beta})$ nin minimum olması için gerek şart

$$\begin{aligned} \frac{\partial SSE(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}'} &= 0 \\ \Rightarrow -2\mathbf{X}'\mathbf{y} + 2\mathbf{X}'\mathbf{X}\boldsymbol{\beta} &= 0 \end{aligned} \quad (3.11)$$

olmasıdır. Bu durumda $\hat{\beta}$ en küçük kareler tahmin edicisi için, aşağıdaki normal denklemler sistemini elde ederiz:

$$\mathbf{X}'\mathbf{X}\hat{\beta} = \mathbf{X}'\mathbf{y}. \quad (3.12)$$

\mathbf{X} in rankı ne olursa olsun (3.12) normal denklemler sistemi daima tutarlıdır. \mathbf{X} in tam sütun ranklı olması durumunda, yani, $rank(\mathbf{X}_{n \times p}) = p \leq n$ olması durumunda, (3.12) denklemler sistemi için $\hat{\beta}$ çözümü

$$\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} \quad (3.13)$$

ile verilir. $E(\hat{\beta}) = E\left[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}\right] = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'E[\mathbf{y}] = \beta$ dır yani $\hat{\beta}$; β nın bir yansız tahmin edicisidir. $var(\hat{\beta})$ ya gelince:

$$\begin{aligned} var(\hat{\beta}) &= E\left[\left((\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} - \beta\right)\left((\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} - \beta\right)'\right] \\ &= E\left[\left((\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathbf{X}\beta + \varepsilon) - \beta\right)\left((\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathbf{X}\beta + \varepsilon) - \beta\right)'\right] \\ &= E\left[\left(\beta + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\varepsilon - \beta\right)\left(\beta + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\varepsilon - \beta\right)'\right] \\ &= E\left[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\varepsilon\varepsilon'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\right] \\ &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\underbrace{E(\varepsilon\varepsilon')}_{=\sigma^2\mathbf{I}_n}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} \\ &= \sigma^2(\mathbf{X}'\mathbf{X})^{-1} \end{aligned} \quad (3.14)$$

dir. Tahmin edilen hata $e = \mathbf{y} - \mathbf{X}\hat{\beta}$ olmak üzere,

$$\hat{\sigma}^2 = \frac{e'e}{n - rank(\mathbf{X})} = \frac{(\mathbf{y} - \mathbf{X}\hat{\beta})'(\mathbf{y} - \mathbf{X}\hat{\beta})}{n - p} = s^2, \quad (3.15)$$

σ^2 nin bir yansız tahmin edicisidir. Böylece $var(\hat{\beta})$ nın tahmini $\hat{\sigma}^2(\mathbf{X}'\mathbf{X})^{-1}$ dir.

Normal dağılan hatalar ile \mathbf{y} nin ortak yoğunluğu (ve bu nedenle olabilirlik fonksiyonu)

$$L(\mathbf{y}; \boldsymbol{\mu}, \boldsymbol{\Sigma} = \sigma^2\mathbf{I}) = \frac{e^{-\frac{1}{2}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})'\boldsymbol{\Sigma}^{-1}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})}}{(2\pi)^{n/2}|\boldsymbol{\Sigma}|^{1/2}} = \frac{e^{-\frac{1}{2\sigma^2}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})'(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})}}{(2\pi)^{n/2}|\sigma^2\mathbf{I}|^{1/2}} = \frac{e^{-\frac{1}{2\sigma^2}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})'(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})}}{(2\pi)^{n/2}(\sigma^2)^{n/2}} \quad (3.16)$$

dir. Olabilirlik fonksiyonunun doğal logaritması

$$\begin{aligned}\ell = \ln L &= -\frac{(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})}{2\sigma^2} - \frac{n}{2} \ln 2\pi - \frac{n}{2} \ln \sigma^2 \\ &= -\frac{\mathbf{y}'\mathbf{y} - 2\boldsymbol{\beta}'\mathbf{X}'\mathbf{y} + \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta}}{2\sigma^2} - \frac{n}{2} \ln 2\pi - \frac{n}{2} \ln \sigma^2\end{aligned}\quad (3.17)$$

ile verilir.

Burada $\boldsymbol{\beta}$ ve σ^2 model parametrelerinin sırasıyla $\tilde{\boldsymbol{\beta}}$ ve $\tilde{\sigma}^2$ en çok olabilirlik tahmin edicileri; (3.17) nin $\boldsymbol{\beta}$ ve σ^2 ye göre türevlerini sıfıra eşitleyerek ve sonuçta oluşan $p+1$ sayıda denklemi, p sayıda $\boldsymbol{\beta}$ lar ve σ^2 için çözerek elde edilir.

En çok olabilirlik tahmin edicileri için birinci mertebe şartları

$$\begin{aligned}\frac{\partial \ell}{\partial \boldsymbol{\beta}} &= -\frac{1}{2\sigma^2} (-2\mathbf{X}'\mathbf{y} + 2(\mathbf{X}'\mathbf{X})\boldsymbol{\beta}) = 0 \\ \frac{\partial \ell}{\partial \sigma^2} &= \frac{1}{2} \left(\frac{1}{\sigma^2} \right)^2 (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) - \frac{n}{2} \frac{1}{\sigma^2} = 0\end{aligned}\quad (3.18)$$

dır. Bu denklemleri aynı anda çözerek, aşağıdaki $\tilde{\boldsymbol{\beta}}$ ve $\tilde{\sigma}^2$ en çok olabilirlik tahmin edicilerini elde ederiz:

$$\begin{aligned}\tilde{\boldsymbol{\beta}} &= (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{y} \\ \tilde{\sigma}^2 &= \frac{1}{n} (\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}})'(\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}}) \\ &= \frac{\boldsymbol{\delta}'\boldsymbol{\delta}}{n}.\end{aligned}\quad (3.19)$$

Burada $\boldsymbol{\delta} = \mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}}$ tahmin edilen hatayı verir. $\tilde{\sigma}^2$ tahmin edicisi, σ^2 için yansız değildir, fakat $\tilde{\boldsymbol{\beta}}$ tahmin edicisi, $\boldsymbol{\beta}$ için yansızdır.

3.1.2.2. Merkezleştirilmiş Biçimdeki Model

Her bir y_i için (3.2) deki model merkezleştirilen x değişkenlerine göre,

$$\begin{aligned}y_i &= \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip} + \varepsilon_i, \quad i = 1, 2, \dots, n \\ &= \alpha + \beta_1 (x_{i1} - \bar{x}_1) + \beta_2 (x_{i2} - \bar{x}_2) + \dots + \beta_p (x_{ip} - \bar{x}_p) + \varepsilon_i\end{aligned}\quad (3.20)$$

olarak yazılabilir. Burada

$$\alpha = \beta_0 + \beta_1 \bar{x}_1 + \beta_2 \bar{x}_2 + \dots + \beta_p \bar{x}_p \quad (3.21)$$

ve $\bar{x}_j = \sum_{i=1}^n x_{ij}/n$, $j=1,2,\dots,p$ dir. Modelin merkezleştirilmiş biçimi, etkili gözlemler için bir araştırmada ve diğer sezışleri temin etmede, belirli hipotezlerin ifade edilmesinde yararlıdır (Rencher ve Schaalje, 2007).

y_1, y_2, \dots, y_n için (3.20) merkezleştirilmiş modeli matris formunda,

$$\mathbf{y} = (\mathbf{j}, \mathbf{X}_c) \begin{pmatrix} \boldsymbol{\alpha} \\ \boldsymbol{\beta}_1 \end{pmatrix} + \boldsymbol{\varepsilon} \quad (3.22)$$

olur. Burada, $\boldsymbol{\beta}_1 = (\beta_1, \beta_2, \dots, \beta_p)'$ ve $\mathbf{X} = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{pmatrix}$ olmak üzere,

$$\mathbf{X}_c = \left(\mathbf{I} - \frac{1}{n} \mathbf{J} \right) \mathbf{X} = \begin{pmatrix} x_{11} - \bar{x}_1 & x_{12} - \bar{x}_2 & \cdots & x_{1p} - \bar{x}_p \\ x_{21} - \bar{x}_1 & x_{22} - \bar{x}_2 & \cdots & x_{2p} - \bar{x}_p \\ \vdots & \vdots & & \vdots \\ x_{n1} - \bar{x}_1 & x_{n2} - \bar{x}_2 & \cdots & x_{np} - \bar{x}_p \end{pmatrix} \quad (3.23)$$

dir. $\mathbf{I} - (1/n)\mathbf{J}$ matrisine bazen “merkezleme matrisi” de denir.

(3.12) de olduđu gibi, (3.22) deki model için normal denklemler

$$(\mathbf{j}, \mathbf{X}_c)' (\mathbf{j}, \mathbf{X}_c) \begin{pmatrix} \hat{\boldsymbol{\alpha}} \\ \hat{\boldsymbol{\beta}}_1 \end{pmatrix} = (\mathbf{j}, \mathbf{X}_c)' \mathbf{y} \quad (3.24)$$

dir. Kısım 2.1.1. deki (2.20) ve (2.24) bağıntılarına göre (3.24) ün sol yanındaki $(\mathbf{j}, \mathbf{X}_c)' (\mathbf{j}, \mathbf{X}_c)$ çarpımı,

$$\begin{aligned} (\mathbf{j}, \mathbf{X}_c)' (\mathbf{j}, \mathbf{X}_c) &= \begin{pmatrix} \mathbf{j}' \\ \mathbf{X}_c' \end{pmatrix} (\mathbf{j}, \mathbf{X}_c) = \begin{pmatrix} \mathbf{j}\mathbf{j} & \mathbf{j}\mathbf{X}_c \\ \mathbf{X}_c'\mathbf{j} & \mathbf{X}_c'\mathbf{X}_c \end{pmatrix} \\ &= \begin{pmatrix} n & \mathbf{0}' \\ \mathbf{0} & \mathbf{X}_c'\mathbf{X}_c \end{pmatrix} \end{aligned} \quad (3.25)$$

olur. Burada \mathbf{X}_c nin sütunları toplamı sıfır olduğundan $\mathbf{j}'\mathbf{X}_c = \mathbf{0}$ dir. (3.24) ün sağ yanı

$$(\mathbf{j}, \mathbf{X}_c)' \mathbf{y} = \begin{pmatrix} \mathbf{j}' \\ \mathbf{X}_c' \end{pmatrix} \mathbf{y} = \begin{pmatrix} n\bar{y} \\ \mathbf{X}_c'\mathbf{y} \end{pmatrix}$$

olarak yazılabilir. Bu durumda en küçük-kareler tahmin edicileri

$$\begin{aligned} \begin{pmatrix} \hat{\alpha} \\ \hat{\beta}_1 \end{pmatrix} &= [(\mathbf{j}, \mathbf{X}_c)' (\mathbf{j}, \mathbf{X}_c)]^{-1} (\mathbf{j}, \mathbf{X}_c)' \mathbf{y} = \begin{pmatrix} n & \mathbf{0}' \\ \mathbf{0} & \mathbf{X}_c' \mathbf{X}_c \end{pmatrix}^{-1} \begin{pmatrix} n\bar{y} \\ \mathbf{X}_c' \mathbf{y} \end{pmatrix} \\ &= \begin{pmatrix} 1/n & \mathbf{0}' \\ \mathbf{0} & (\mathbf{X}_c' \mathbf{X}_c)^{-1} \end{pmatrix} \begin{pmatrix} n\bar{y} \\ \mathbf{X}_c' \mathbf{y} \end{pmatrix} = \begin{pmatrix} \bar{y} \\ (\mathbf{X}_c' \mathbf{X}_c)^{-1} \mathbf{X}_c' \mathbf{y} \end{pmatrix} \end{aligned}$$

veya

$$\hat{\alpha} = \bar{y} \quad (3.26)$$

$$\hat{\beta}_1 = (\mathbf{X}_c' \mathbf{X}_c)^{-1} \mathbf{X}_c' \mathbf{y} \quad (3.27)$$

ile verilir. Bu tahmin ediciler, (3.21) deki α nın bir tahmin edicisinden elde edilen

$$\hat{\beta}_0 = \alpha - \hat{\beta}_1 \bar{x}_1 - \hat{\beta}_2 \bar{x}_2 - \dots - \hat{\beta}_p \bar{x}_p = \bar{y} - \beta_1' \bar{\mathbf{x}} \quad (3.28)$$

ayarlaması ile birlikte (3.13) deki $\hat{\beta} = (\mathbf{X}' \mathbf{X})^{-1} \mathbf{X}' \mathbf{y}$ alışılmış en küçük-kareler tahmin edicisiyle aynıdır (Rencher ve Schaalje, 2007).

\hat{y} yı,

$$\hat{y} = \hat{\alpha} + \hat{\beta}_1 (\mathbf{x}_1 - \bar{x}_1) + \hat{\beta}_2 (\mathbf{x}_2 - \bar{x}_2) + \dots + \hat{\beta}_p (\mathbf{x}_p - \bar{x}_p)$$

merkezleştirilmiş formunda ifade ettiğimizde, uygun regresyon düzleminin $(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_p, \bar{y})$ noktasından geçtiği açıktır. Bu durumda genel lineer modeli aşağıdaki gibi ifade edebiliriz:

$$\mathbf{y}_{n \times 1} = \mathbf{X}_{n \times p} \boldsymbol{\beta}_{p \times 1} + \boldsymbol{\varepsilon}_{n \times 1}, \quad \boldsymbol{\varepsilon} \sim N(0, \sigma^2 \mathbf{I}_n). \quad (3.29)$$

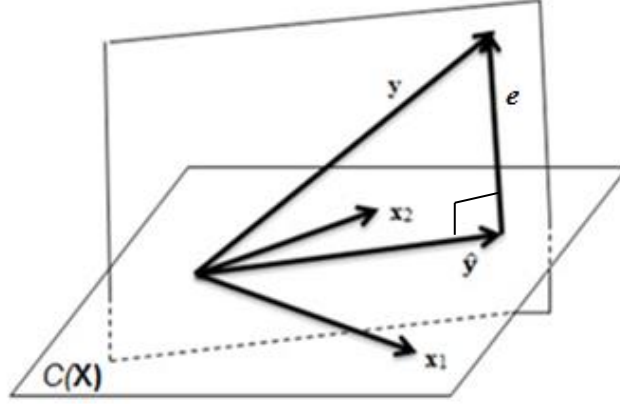
3.1.3. En Küçük Karelerin Geometrisi

Genel lineer modelde, tek bir y hedef değişkeni $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_p$ sütun vektörlerinin bir lineer kombinasyonu ile tahmin edilir. Yani, $\hat{\beta} = (\mathbf{X}' \mathbf{X})^{-1} \mathbf{X}' \mathbf{y}$ olmak üzere, matris-vektör notasyonu ile

$$\hat{y} = \mathbf{X} \hat{\beta} \quad (3.30)$$

dır.

Geometrik olarak bu gösterim y yi X in sütun uzayına izdüşürmek suretiyle elde edilir.



Şekil 3.1. y nin X sütun uzayına izdüşümü.

Cebirsel olarak bu izdüşüm için $\hat{\beta}$ katsayı vektörü, $(X'X)^{-1}X'y$ matrisiyle y yi çarparak belirlenir, yani (3.13) deki gibidir.

Bu denklem sıkça görülecektir. Yukarıdaki şekilde görüldüğü gibi y ; biri \hat{y} tahmini diğeri e hata tahmini vektörleri olmak üzere iki ortogonal kısma ayrılır. Bu vektörler $|\hat{y}|$ maksimal (en büyük) olacak şekilde seçilir ve sonuç olarak $|e| = |\hat{e}|$ minimal (en küçük) dir. Burada \hat{e} tahmin edilen hata vektörüdür ve $\hat{e} = e = y - X\hat{\beta}$ dır.

3.1.4. Genel Lineer Model için Veri

Şimdi Çizelge 3.1 de sunulan veriler için genel lineer modeli göz önüne alalım.

Çizelge 3.1. Açıklayıcı verinin bir minyatür kümesi. İki uyarıcı, farklı zamanlarda sunulmuş ve tepkimeler dört konumdan kaydedilmiştir.

Zaman (t)	Uyarıcı	Açıklayıcı değişkenler (şartlar)			Tepkimeler			
		\mathbf{x}_0	\mathbf{x}_1	\mathbf{x}_2	\mathbf{y}_1	\mathbf{y}_2	\mathbf{y}_3	\mathbf{y}_4
1	0	1	0	0	12.15	13.45	16.45	17.95
2	0	1	0	0	11.56	14.33	16.63	16.42
3	1	0	1	1	15.50	19.39	11.23	11.13
4	1	0	1	1	11.65	12.45	17.54	15.50
5	2	1	0	1	11.04	13.25	12.85	11.09
6	2	1	0	1	9.45	10.27	15.12	17.21
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

Analizin temel amacı, y_k yı \mathbf{x} lerin bir fonksiyonu olarak tanımlamaktır. Çizelge 3.1, içerilen veri tipinin bir örneğini gösterir. 6 tanesi gösterilen bir tepkime zaman noktalarının bir dizisinde dört yerleşimde (\mathbf{y}_1 den \mathbf{y}_4 e) ölçülmüştür.

Çizelge 3.1 deki örneğe, ayrı ayrı (3.13) formülünü uygulayalım. Burada \mathbf{X} ,

çizelgenin üç \mathbf{x} sütununda verilen 3x6 matrisidir ve $\hat{\boldsymbol{\beta}} = \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix}$ 3x1 vektördür (Veri

Çizelge 3.1 de gösterilen 6 gözlemden daha fazlasını içermektedir.).

$$\mathbf{y}_1 = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

$$\begin{bmatrix} 12.15 \\ 11.56 \\ 15.50 \\ 11.65 \\ 11.04 \\ 9.45 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix} \boldsymbol{\beta} + \boldsymbol{\varepsilon} \Rightarrow \hat{\boldsymbol{\beta}} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} 11.855 \\ 15.185 \\ -1.610 \end{bmatrix},$$

$$\mathbf{y}_2 = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

$$\begin{bmatrix} 13.45 \\ 14.33 \\ 19.39 \\ 12.45 \\ 13.25 \\ 10.27 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix} \boldsymbol{\beta} + \boldsymbol{\varepsilon} \Rightarrow \hat{\boldsymbol{\beta}} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} 13.890 \\ 18.050 \\ -2.130 \end{bmatrix},$$

$$\mathbf{y}_3 = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

$$\begin{bmatrix} 16.45 \\ 16.63 \\ 11.23 \\ 17.54 \\ 12.85 \\ 15.12 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix} \boldsymbol{\beta} + \boldsymbol{\varepsilon} \Rightarrow \hat{\boldsymbol{\beta}} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} 16.560 \\ 15.940 \\ -3.035 \end{bmatrix},$$

ve

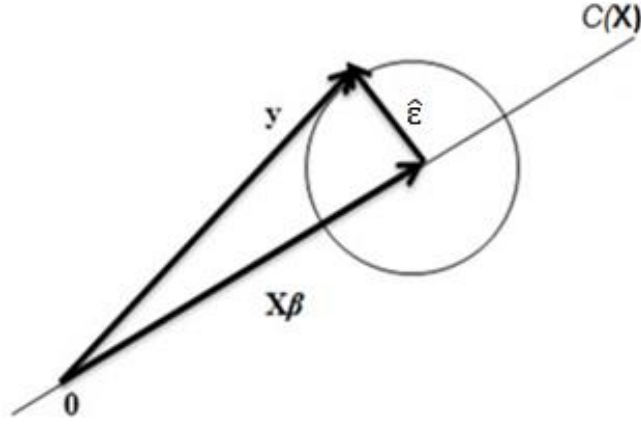
$$\mathbf{y}_4 = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

$$\begin{bmatrix} 17.95 \\ 16.42 \\ 11.13 \\ 15.50 \\ 11.09 \\ 17.21 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix} \boldsymbol{\beta} + \boldsymbol{\varepsilon} \Rightarrow \hat{\boldsymbol{\beta}} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} 17.185 \\ 16.350 \\ -3.035 \end{bmatrix}$$

elde ederiz. Sonuç dört tahminin kümesidir. Bu küme aşağıda verilmiştir.

	\mathbf{y}_1	\mathbf{y}_2	\mathbf{y}_3	\mathbf{y}_4
$\hat{\beta}_0$	11.855	13.890	16.560	17.185
$\hat{\beta}_1$	15.185	18.050	16.940	16.350
$\hat{\beta}_2$	-1.610	-2.130	-2.555	-3.035

(3.29) modelinde ε_i hataları bağımsızdırlar. $\mathbf{0}$ ortalama ve σ^2 varyans ile bir normal dağılıma sahip olarak aynı dağılırlar. Bu varsayım, gerçek \mathbf{y} gözleminin $\mathbf{X}\boldsymbol{\beta}$ gerçek model değeri etrafında bir küresel simetrik dağılıma sahip olduğunu ifade eder.



Şekil 3.2. $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ modelinin geometrik gösterimi.

Hatanın izotropisi, $\mathbf{X}\boldsymbol{\beta}$ yı $\hat{\boldsymbol{\varepsilon}}$ hatası ile ilişkisiz tutar ve $\hat{\boldsymbol{\beta}}$ nın izdüşümle tahmin edilmesine olanak tanır. Bu, Gauss dağılımının varsayımının önemli bir sonucudur.

(3.29) modeli altında, $\boldsymbol{\beta}$ nın $\hat{\boldsymbol{\beta}}$ tahmini bir dereceye kadar farklı iki yoruma sahiptir.

Bu yorumlardan birincisi, $\hat{\mathbf{y}}$ nın \mathbf{y} ye bir en küçük kareler uyumu gibidir. $\hat{\boldsymbol{\beta}}$ tahmini $\hat{\boldsymbol{\varepsilon}}$ hatasının uzunluğunu minimum yapan $\hat{\mathbf{y}} = \mathbf{X}\hat{\boldsymbol{\beta}}$ lineer kombinasyonunu ortaya koyar. Özellikle bu, \mathbf{y} ve $\hat{\mathbf{y}}$ arasındaki farkın kareleri toplamı için olası en küçük değeri, yani

$$|\mathbf{y} - \hat{\mathbf{y}}|^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n (y_i - \mathbf{x}_i' \hat{\boldsymbol{\beta}})^2 \quad (3.31)$$

bağıntısını verir. \mathbf{x}_i' ; \mathbf{X} açıklayıcı değişkenler matrisinin i -yinci satır vektörüdür.

İkinci yorum maksimum olabilirliğin yorumudur ve \mathbf{y}_i nin olasılık dağılımına dayanır. Bu yorumu kısaca ifade edeceğiz.

1. Genel lineer model altında y_i gözleminin olasılığı $\boldsymbol{\beta}$ parametrelerine ve σ^2 ye bağlıdır ve bu,

$$P(y_i; \boldsymbol{\beta}, \sigma^2) \quad (3.32)$$

olarak yazılabilir.

2. Gözlemler bağımsız olduklarından, \mathbf{y} verisinin tam kümesinin olasılığı onun ayrı ayrı gözlemlerinin olasılıklarının çarpımıdır. Yani,

$$P(\mathbf{y}; \boldsymbol{\beta}, \sigma^2) = \prod_{i=1}^n P(y_i; \boldsymbol{\beta}, \sigma^2) \quad (3.33)$$

dir.

3. Bir tahmin probleminde verinin bir kümesi elde edilmiştir, bu nedenle o artık bir rasgele nicelik olarak işlenemez. Bunun yerine yukarıdaki olasılığı parametrelerin bir fonksiyonu olarak işleriz ve olabilirliği yardıma çağırırız. Değişik nedenlerden dolayı olabilirliğin kendisinden ziyade olabilirliğin logaritması ile çalışmak uygundur, bundan dolayı

$$L(\boldsymbol{\beta}, \sigma^2) = \log P(y_i; \boldsymbol{\beta}, \sigma^2) \quad (3.34)$$

bağıntısını tanımlarız.

4. $\boldsymbol{\beta}$ ve σ^2 parametrelerini tahmin etmek için, $L(\hat{\boldsymbol{\beta}}, s^2)$ yi maksimum yapan $\hat{\boldsymbol{\beta}}$ ve s^2 değerlerini buluruz. Bunlar en çok olabilirlik tahminleridir.

Tahminlerin en küçük kareler ve en çok olabilirlik yorumları; gözlemlerin bağımsız, aynı dağılımlı ve Gaussian olmaları varsayımları altında eşdeğerdirler (aynıdırılar). Bu varsayımlar olmadığında iki yöntem, farklı tahminleri verecektir (Wickens, 2004).

3.1.5. Uyum Testlerinin İyiliği (Mükemmelliği)

(3.29) genel lineer modelinde σ^2 parametresinin tahmini (3.15) de verilmiştir. Buradan parametre tahminlerinin kovaryans matrisi

$$\mathbf{S}_{\hat{\boldsymbol{\beta}}} = \hat{\sigma}^2 (\mathbf{X}'\mathbf{X})^{-1}$$

olarak elde edilmiştir. Modelin uyumunun iyiliği;

$$\mathbf{R}^2 = \frac{\text{açıklanan değişim}}{\text{toplam değişim}} = \frac{|\hat{\mathbf{y}}|^2}{|\mathbf{y}|^2} \quad (3.35)$$

ile ölçülebilir. Bu nicelik, korelasyon oranı olarak bilinir. (3.29) genel lineer modeli altında (Şekil 3.1 den de görüldüğü gibi)

$$\mathbf{y} = \hat{\mathbf{y}} + \mathbf{e} = \mathbf{X}\hat{\boldsymbol{\beta}} + (\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}) \quad (3.36)$$

olarak yazılabilir.

$H_0 : \boldsymbol{\beta} = 0$ hipotezinin bir testi, \mathbf{y} nin iki bileşeninin boyut başına uzunluk karelerini karşılaştırarak elde edilir. Burada $\hat{\mathbf{y}}$ nin uzandığı uzayın boyutu p ve $\hat{\boldsymbol{\epsilon}} = \mathbf{e}$ nin

uzandıđı uzayın boyutu $n - p$ dir. Modelin açıklayıcı deđiřkenler matrisi \mathbf{X} in $n \times p$ boyutlu ve $\text{rank}(\mathbf{X}) = p < n$ olması durumunda bu hipotez

$$F = \frac{|\hat{\mathbf{y}}|^2 / p}{|\hat{\boldsymbol{\varepsilon}}|^2 / (n - p)} = \frac{\text{Model kareler ortalaması}}{\text{Hata kareler ortalaması}} \quad (3.37)$$

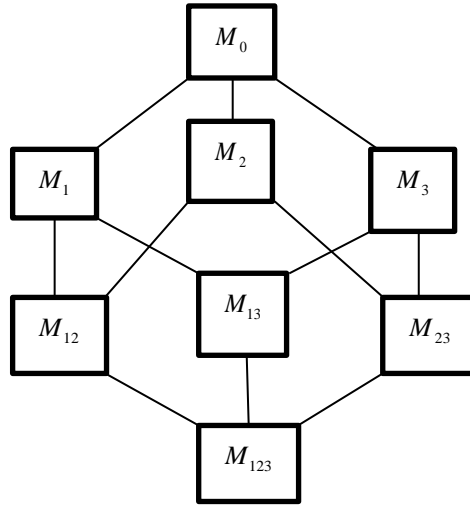
test istatistiđi ile test edilir. Bu istatistik; payı p , paydası $n - p$ serbestlik dereceli merkezi olmayan bir F -dađılımlına sahiptir. H_0 hipotezi dođru olduđunda bu istatistik merkezi $F_{1-\alpha} = (p, n - p)$ -dađılımlına sahiptir. Burada α testin anlam düzeyidir (I.Tip hatadır). Bu biçimdeki test, $\boldsymbol{\beta} = 0$ in modeli basit bir řekilde çok bilgilendirici olmayan $y_i = \varepsilon_i$ ye indirgediđinin iddiasıdır. Parametrelerin özel deđerleri hakkındaki hipotezlerin testleri daha faydalıdır. Bunlar 4. Bölümün ışıđı altında daha iyi anlaşılır.

3.1.6. Model Seçimi ve Karřılařtırma

Veri için bir gösterimin veya modelin seçimi konusu çok önemlidir ve onu genel olarak ele almak için genel modelden kısa bir süreliđine uzaklařacađız. Olası modellerin bir Latisini (örgüsünü) kurmak suretiyle problemi grafiksel olarak düşünelim. Örneđin $\mathbf{x}_1, \mathbf{x}_2$ ve \mathbf{x}_3 açıklayıcı deđerkenleri ile, sekiz olası modeli ortaya koyabiliriz. Bunlar:

$$\begin{aligned} M_0 &: y = b_0 \\ M_1 &: y = b_0 + b_1x_1 \\ M_2 &: y = b_0 + b_2x_2 \\ M_3 &: y = b_0 + b_3x_3 \\ M_{12} &: y = b_0 + b_1x_1 + b_2x_2 \\ M_{13} &: y = b_0 + b_1x_1 + b_3x_3 \\ M_{23} &: y = b_0 + b_2x_2 + b_3x_3 \\ M_{123} &: y = b_0 + b_1x_1 + b_2x_2 + b_3x_3 \end{aligned}$$

dir. Bu gösterimler tepede en basit M_0 modeli ve dipte en karışık M_{123} modeli ile hiyerarşik bir örgüye uyar.



Şekil 3.3. Sekiz olası modelin bir latisi (örgüsü).

Her birisi diğerinin kısıtlamaları olan modeller bu örgüye göre bir doğru ile bağlanır. Örgü, açıklayıcı değişkenlerin bazısının katsayısını eşit yapanlar (örneğin; $y = b + 0 + b_{12}x_1 + b_{12}x_2$) veya onların değerinin kuvvetlerini içeren (örneğin; $y = b_0 + b_{11}x_1 + b_{12}x_1^2$) gibi diğer birçok modeli dahil etmek için genişletilebilir.

Bu örgünün dibi tarafına doğru modeller, tepesine doğru olanlardan daha fazla parametre içerir ve sonuç olarak onlar genel olarak daha iyi uyarlar. Bir doğru (veya inen veya çıkan doğruların dizisi) ile bağlanan modellere gelince, onların her biri diğeri ile hiyerarşik bir ilişkiye sahiptir ve daha üstte olan kesinlikle daha alttakinde içerilir. Örneğin, $M_1 \subset M_{13} \subset M_{123}$ dür. Bu dizide daha alttaki modeller daha üstteki modeller kadar iyi uyacaktır.

İlişkili bağlantılı olmayan modellere gelince- ilişki hiyerarşik değildir- daha alttaki modelin daha iyi uyması gerekmez.

Örneğin, M_{12} modeli üç parametreye ve M_3 modeli iki parametreye sahip olmasına rağmen, M_3 modelinin verilen bir veri kümesi için daha iyi uyması gerekmez.

İki model arasındaki ilişkinin hiyerarşik olup olmadığı, onları karşılaştırılabilir yapma yöntemindeki farkı yaratır.

Hiyerarşik modellerin bir karşılaştırması için genel olarak elde mevcut iyi testler vardır. Bir test istatistiği (yukarıdaki F testi gibi) kısıtlanmamış model (örgüde daha altta olan) ve onun yukarısındaki kısıtlanmış modelin uyumu arasındaki farka bakarak oluşturulabilir.

$$Test\ istatistiği = Uyum(M_{kısıtlanmamış}) - Uyum(M_{kısıtlanmış}) \quad (3.38)$$

dir. Bu istatistiğin büyüklüğü, modelleri ayıran kısıtlamaların sayısıyla belirlenen bir dağılıma karşılığında değerlendirilir. Log olabilirlik hesaplanabildiğinde fark, ki-kare dağılımına başvurabilir:

m tane kısıtlamanın bir testi için,

$$2\left[L(M_{\text{kısıtlanmamış}}) - L(M_{\text{kısıtlanmış}})\right] \sim \chi_m^2 \quad (3.39)$$

dir. Hiyerarşik olmayan modellerin karşılaştırılması daha zordur. Daha fazla parametrelili bir modelin daha az parametrelili bir modelden daha iyi uyacağına güvencesi yoktur ve uyumdaki farkın dağılımı genel olarak mevcut değildir. Bir modelin kullanılan parametrelerin sayısı için bir ceza ile uyumun bir ölçüsünü birleştiren başarısının bir ölçüsünü inşa etmek akla uygundur:

$$\text{Başarı}(M) = \text{Uygun}(M) - f[\text{Parametreler}(M)]. \quad (3.40)$$

Birkaç modelin bir karşılaştırmasında, en başarılı olan tercih edilebilir. Buradaki zorluk, uygun ceza fonksiyonunu belirlemektir. Onlar genel olarak uyumu ölçmek için olabilirliğe bağlıdır ve gözlemlerin ve parametrelerin sayısına bağlı olan bir $f(n, p)$ ceza fonksiyonunu kullanırlar. Bunların en çok bilinenlerinden ikisi Akaike Bilgi Ölçütü (ABÖ),

$$\text{ABÖ} = 2L(M) - 2p \quad (3.41)$$

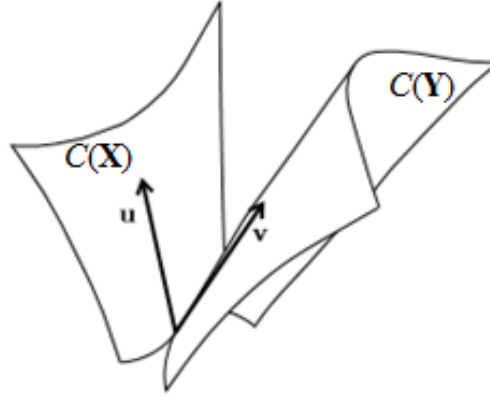
ve Bayesci Bilgi Ölçütü (BBÖ),

$$\text{BBÖ} = 2L(M) - \log(n)p \quad (3.42)$$

dir (Wickens, 2004).

3.1.7. Çok Değişkenli Genelleştirme

Genel lineer modellerin önemli bir özelliği onun, hatanın bir değişkenli Gaussian dağılımı yerine sadece bir çok değişkenli dağılımı koyarak, çok değişkenli gözlemlere kolayca genelleştirilmesidir. Genel lineer modelinkine benzer problem iki $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_p\}$ ve $\{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_q\}$ değişkenler kümesi ile ilgili olmaktadır. Bu, \mathbf{X} vektörleri tarafından üretilen $C(\mathbf{X})$ uzayından $\mathbf{u} = \mathbf{X}\mathbf{b}$ ve \mathbf{Y} vektörleri tarafından üretilen $C(\mathbf{Y})$ uzayından $\mathbf{v} = \mathbf{Y}\mathbf{c}$ gibi, iki küme ile ilgili değişkenlerin lineer kombinasyonları olan yeni değişkenleri ortaya koymak suretiyle yapılır.



Şekil 3.4. $C(\mathbf{X})$ ve $C(\mathbf{Y})$ uzayları tarafından üretilen \mathbf{u} ve \mathbf{v} vektörlerinin gösterimi.

\mathbf{b} ve \mathbf{c} katsayıları; \mathbf{u} ve \mathbf{v} vektörleri arasındaki korelasyon mümkün olabildiği kadar büyük olacak şekilde, yani, bu nedenle onlar arasındaki açı minimal olacak şekilde seçilirler (İki vektör arasındaki korelasyon Bölüm 4 de ayrıntılı bir şekilde verilmiştir.).

$C(\mathbf{X})$ ve $C(\mathbf{Y})$ uzaylarının her ikisi de çok değişkenli olduklarından, ilave vektörler seçilebilir. Her bir küme içindeki değişkenler daha önce seçilenlere ortogondur ve kümeler arasında korelasyonu büyüyecek şekilde seçilir. Bu sonuç sırayla daha büyük boyutlu bağlanmış uzayların bir serisidir. Yani,

$$\begin{aligned}
 \mathbf{u}_1 &\leftrightarrow \mathbf{v}_1 \\
 \{\mathbf{u}_1, \mathbf{u}_2\} &\leftrightarrow \{\mathbf{v}_1, \mathbf{v}_2\} \\
 \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\} &\leftrightarrow \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\} \\
 \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4\} &\leftrightarrow \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\} \quad \text{v.s.}
 \end{aligned} \tag{3.43}$$

dür. Bu yöntem içerilen değişkenlerin yorumuna bağlı olarak, çeşitli isimlere sahiptir (çok değişkenli varyans analizi, çok değişkenli çoklu regresyon, kanonik korelasyon, v.s. gibi) (Wickens, 2004).

3.1.8. Genel Lineer Modelin Önemi

Genel lineer modelin bir özeti olarak, onu cazip yapan özelliklere dikkat edelim.

Hem tanımlamak hem de yorumlamak kavranabilirdir. Özellikle (3.29) lineer modeli hem basittir hem de çok güçlüdür ve hata dağılımı, etkilerden açık bir şekilde ayırır. Bu şekildeki $C(\mathbf{X})$ ve $C(\mathbf{Y})$ uzayları gerçekten lineer uzaylardır, Şekil 3.4 dört boyutlu bir kurulumu iki boyutlu bir sayfa üzerine izdüşürme teşebbüsüdür ve \mathbf{u} , \mathbf{v} nin ilişkisini gösterir.

Rasgele yapı basittir. Gaussian dağılımı; onun ilk iki momenti, yani μ ve σ^2 (veya çok değişkenli durumda, $\boldsymbol{\mu}$ ortalama vektörü ve Σ kovaryans matrisiyle) ile tam olarak tanımlanır. Bu momentler stokastik olarak bağımsızdır, bu ise onları tahmin etmeyi daha da kolaylaştırır. Hata tarafından tetiklenen geometrik uzay Öklid uzayıdır ve eş yönlü, yani yönden bağımsızdır.

Model hesapsal olarak işlenebilir. Parametrelerin kapalı-form tahminleri vardır ve ortalama ve varyans ayrı ayrı tahmin edilebilir. Tahmin edilen lineer parametreler izdüşüm vasıtasıyla elde edilirler, yani $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{y}$ dir ve orijinal \mathbf{y} verisinin lineer kombinasyonlarıdır. Bu ve diğer hesaplamalar için lineer cebirin gücü kullanılabilir. Hesaplamalar elle bile yapılabilir.

Tahminlerin (en küçük kareler ve en çok olabilirlik tahminleri) farklı yorumu ve testlerin farklı biçimleri aynı sonuçları ortaya koyar.

Model bir değişkenliden çok değişkenliye çabucak genelleşir. Çok değişkenli Gaussian dağılımın hem şartlı hem de marjinal dağılımları yine Gaussiandır, bu nedenle bu dağılımlarla ilgili testler tam çok değişkenli modelle ilgili olanlarla aynı biçime sahiptir.

Genel lineer modelin varsayımları gevşetildiğinde, bu karakteristiklerin bazıları ona uyar. Bununla beraber, bazı durumlarda, model gereğinden çok kısıtlıdır. Onun genişlemelerinden bazısını özellikle,

- Tam gözlemler aynı ağırlığa sahiptir.
- Hata dağılımı Gaussiandır.
- Gözlemler bağımsızdır.

varsayımlarını kısaca ele alacağız.

3.1.9. Farklı Değişkenlik

Verinin bazı kümelerinde, \mathbf{y}_i nin varyansı farklıdır; yani,

$$i \neq k \quad \text{için} \quad \text{var}(\mathbf{y}_i) = \sigma_i^2 \neq \text{var}(\mathbf{y}_k) = \sigma_k^2 \quad (3.44)$$

dür. Bu değişkenlikler, açık olarak bir tek gözlem üzerinde ölçülemez fakat sayılar ölçüm hataları içeren bir kaynaktan elde edildiğinde mevcut olabilir. Varyanstaki farklılıklar gözlemleri farklı bilgi verici yapar. Bu eşitsizlik gözlemleri onların varyansının tersiyle ağırlıklandırmak suretiyle analizde barındırılır. Örneğin, ağırlıklı

ortalama $w_i = \frac{1}{\sigma_i^2}$ olmak üzere,

$$\bar{y} = \frac{\sum w_i y_i}{\sum w_i} \quad (3.45)$$

olarak hesaplanır. Bir $\hat{y} = \mathbf{X}\hat{\beta}$ lineer modelinin parametrelerinin tahminleri, en küçük kareler tahmin denklemini ağırlıklandırarak elde edilir. Buna göre,

$$\hat{\beta} = (\mathbf{X}'\mathbf{W}\mathbf{X})^{-1} (\mathbf{X}'\mathbf{W}\mathbf{y}) \quad (3.46)$$

dır. Burada \mathbf{W} ağırlık matrisi; varyansların tersini içerir, yani,

$$\mathbf{W} = \begin{bmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \sigma_n^2 \end{bmatrix}^{-1} \quad (3.47)$$

dir.

3.1.10. Gaussian Olmayan Dağılımlar ve Genelleştirilmiş Lineer Model

Bazı veri üretme yöntemleri Gaussian dağılımlara sahip düzlemleri vermez. Bir Gaussian dağılımı varsayma, bir Bernoulli dağılımına göre dağılan ikili gözlemleri düşünmek suretiyle hemen görülen birkaç sonuca sahiptir.

a) Ortalama ve varyans ilişkilidirler.

Örneğin, $P(y=1) = \pi$ ve $P(y=0) = 1 - \pi$ olan bir y Bernoulli rasgele değişkeni için ortalama ve varyans

$$E(y) = \pi \quad \text{ve} \quad \text{var}(y) = \pi(1 - \pi) \quad (3.48)$$

ile verilir. Bir sabit ortalamayı tahmin etmeyen herhangi bir model, değişen bir varyansı da tahmin etmez. Bernoulli den başka dağılımlarla (örneğin, poisson hesapları), dağılımlar ortalama büyük olduğunda, yaklaşık olarak Gaussian olabilirler, fakat varyansın ortalamaya bağımlılığı devam eder.

b) Bir lineer model, veriyi modellemek için uygun bir yöntem olmayabilir.

Örneğin, Bernoulli verisi ile bir $\hat{\pi} = \mathbf{X}\hat{\beta}$ lineer modeli, $[0,1]$ aralığının dışında uzanan $\hat{\pi}$ nın değerlerini verebilir, bu nedenle olasılıkları veremeyebilir.

c) Değişen varyans en küçük kareler tahminlerini (yani, basit izdüşümü) bir tahmin tekniği olarak uygun kılmaz.

Genelleştirilmiş lineer model olarak bilinen bir yöntem bu özelliklere uyum sağlamak için genel lineer modeli genişletir. Bu genelleştirme üç karakteristiğe sahiptir.

1. Gözlemlerin dağılımı üstel aile olarak bilinen olasılıkların büyük bir kümesinden çekilir. Bu dağılımların yoğunluğu

$$f(\mathbf{y}) = \exp \left[\frac{\theta \mathbf{y} - b(\theta)}{a(\phi)} + c(\mathbf{y}, \phi) \right] \quad (3.49)$$

biçimine sahiptir. Burada θ , dağılımın konumunu belirleyen bir parametredir. ϕ , değişkenliği ölçen bir parametredir ve $a(\cdot)$, $b(\cdot)$ ve $c(\cdot)$ fonksiyonlardır. Bu aile birçok (fakat tümü değil) yararlı dağılımı içerir :

- Sürekli dağılımlar arasında, Gaussiyen, Log-normal, Ters-Gaussian, Üstel, Gamma (ki-kare), Pareto, Gompers, Weibull, v.s. dir.
- Kesikli dağılımlar arasında, Bernoulli, Binom, Poisson, Geometrik, Negatif binom v.s. dir.

2. Gözlemlerin varyansı doğal olarak ortalama ile bağlantılıdır. Örneğin, Poisson dağılımlı veri ile $\text{var}(\mathbf{y}) = [E(\mathbf{y})]$ ve Gamma dağılımlı veri ile $\text{var}(\mathbf{y}) = [E(\mathbf{y})]^2$ dir. $a(\phi)$ niceliği (genellikle tamamen ϕ), bu ilişki değişmeksizin yukarı veya aşağı değişkenliği, bu nedenle modeli az veya çok dağılmış gözlemlere uydurmaya izin vermeyi ölçer.

μ ortalaması, bir η konum parametresi ve lineer olmayan bir bağ fonksiyonu, yani

$$g(\boldsymbol{\mu}) = \boldsymbol{\eta} = \mathbf{x}'\boldsymbol{\beta} \quad (3.50)$$

fonksiyonu ile bir lineer modele bağlanır. Bağ için herhangi bir fonksiyon kullanılmakla beraber, bağ fonksiyonunun konum parametresi dağılımın θ konum parametresiyle aynı olduğunda, model en kolaydır. Bu takdirde bağ fonksiyonu dağılımın kendisi tarafından belirlenen bir kanonik forma sahiptir. Örneğin, Bernoulli verisi ile kanonik bağ fonksiyonu lojistik (onu lojistik regresyonun modeli yaparak) ve Gamma verisi için, o iki taraflıdır (terstir).

Genelleştirilmiş lineer model kapalı-form, tahmin denklemlerine sahip değildir. Bununla beraber, onun uydurulabildiği birçok iteratif (tekrarlı) algoritma vardır. Örneğin, en çok olasılık uyumları modelden, varyansları ürettikten sonra yukarıda verilen ağırlıklı en küçük kareleri tekrarlı olarak uygulamak suretiyle elde edilebilir. Bunları kolayca tahmin eden yazılım vardır ve başlıca istatistiksel paket programlarında gözüktürler.

Genelleştirilmiş lineer modelin uyumu, log olabilirliğin tam negatif 2 katı olan, sapma olarak bilinen bir nicelikle ölçülür. Bu nicelik

$$sapma = -2 \log L \quad (3.51)$$

dir. Model ϕ parametresini tahmin etmek suretiyle aşırı dağılım veya az dağılıma izin verdiğinde, sapma uyumun iyiliğinin bir ölçümü olarak yorumlanabilir-genel lineer model için var iken hatanın bağımsız tahmini yoktur. Bununla beraber, hiyerarşik modeller için testler, sapmadaki farkla yani,

$$sapma(M_{kısıtlanmış}) - sapma(M_{kısıtlanmamış}) \sim \chi_r^2 \quad (3.52)$$

ile ifade edilir.

Genel lineer model ile bazı özellikleri paylaşmakla birlikte, genelleştirilmiş lineer model kolayca yorumlanamaz. Bir Gaussian (Gauss) dağılımı belirlenmedikçe y nin gözlemlerinin uzayı izotropik ve Öklidyen değildir. Bu gösterime dayanan önseziler dikkatle kullanılmalıdır. Öklidyen olmayan hata uzayı aynı zamanda değişkenliğe dayalı testlere ve model karşılaştırmasına dayalı testlere farklı sonuçlar da veririr. Verilen bir hipotez çoğu kez azıcık farklı sonuçlarla bir veya daha çok yöntemle test edilebilir (olabilirlik oran testleri, skorlar testi, wald testi vs gibi) (Wickens, 2004).

3.1.11. Bağımlı Gözlemler

Hem genel lineer modelde hem de genelleştirilmiş lineer modelde, y_1, \dots, y_n gözlemlerinin bağımsız oldukları kabul edilir. Bununla beraber, birçoğu görüntü verisi ile birlikte ortaya çıkan y_i ler arasındaki bağımlılıklar için birçok neden vardır. Bağımlılığın asıl nedeni vaktiyle değerler daha evvel ki değerlerle etkilenmiş olduğu zaman üzerinde olagelen gözlemlerle ilgilidir. Bu ilişkinin en yaygın gösterimi bir gözlemin bir veya daha evvel ki denemedeki değerle lineer olarak bağlı olduğu bir öz-bağımlı (ÖB) modeli ile olanıdır:

$$\begin{aligned} \text{ÖB}(1) : y_t &= \mathbf{x}_t' \mathbf{b} + a_1 y_{t-1}, \\ \text{ÖB}(2) : y_t &= \mathbf{x}_t' \mathbf{b} + a_1 y_{t-1} + a_2 y_{t-2}, \\ \text{ÖB}(3) : y_t &= \mathbf{x}_t' \mathbf{b} + a_1 y_{t-1} + a_2 y_{t-2} + a_3 y_{t-3}, \text{ v.s.} \end{aligned} \quad (3.53)$$

Öz-bağımlı modeller, gözlemler arasındaki sabit korelasyonları farklı zaman aralıklarının sabit bir sayısına indükler.

Bağımlılığın başka bir kaynağı, kümelenen gözlemlerin varlığıdır. Her bir denekten birkaç gözlem yapıldığında, bir takım gözlemler farklı sonuçların her biri esnasında toplandıklarında veya denekler, aileler veya bir batında doğan yavrular diyeceğimiz gruplar olabildiğinde, kümelerin yaygın örnekleri ortaya çıkar. Her bir bağımlılık

tipi, kümeler içinde bir karakteristik kovaryans yapısına götürür. Genel olarak, kümelerin bağımsız olacakları kabul edilir. Birtakım bağımlılık kaynağı var olduğunda bağımlılığın daha karmaşık kaynakları ortaya çıkar. Örneğin, etkilenimli gözlemler aynı zamanda birkaç ilişkili zeka bölgesinde yapılabilir. Konumsal yapı, bir korelasyon yapısını, yani başka zaman serilerini tetikleyecektir.

Kümelenmiş gözlemleri daha ayrıntılı olarak düşünelim. Kümelenmenin varlığı veri için hiyerarşik bir yapıyı ortaya koyar. Genellikle bu yapı iki (veya daha fazla) örnekleme aşamasıyla temsil edilebilir. Bu hiyerarşik, rasgele-etkiler yapısının basit bir versiyonu, parametre vektörlerini örneklenmiş gibi işleme tabi tutarak genel lineer modeli genişletir.

Aşama 1: Her bir kümedeki sayılar bağımsız hatalara sahip bir lineer model ve bir

$$\mathbf{y}_1 = \begin{bmatrix} y_{11} \\ \vdots \\ y_{1c} \end{bmatrix} = \mathbf{X}_1 \boldsymbol{\beta}_1 + \boldsymbol{\varepsilon}_1, \quad \mathbf{y}_2 = \begin{bmatrix} y_{21} \\ \vdots \\ y_{2c} \end{bmatrix} = \mathbf{X}_2 \boldsymbol{\beta}_2 + \boldsymbol{\varepsilon}_2, \quad \text{v.s.} \quad (3.54)$$

küme-özel parametre vektörü ile belirlenir. Burada $\boldsymbol{\varepsilon}_j \sim N(0, \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}})$ ve $\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}$ köşgendir.

Aşama 2 : Model parametreleri, $\boldsymbol{\beta}_j \sim N(\boldsymbol{\mu}_{\boldsymbol{\beta}}, \boldsymbol{\Sigma}_{\boldsymbol{\beta}})$ Gaussian dağılımdan örneklenir. Bu nedenle model, $\boldsymbol{\beta}$ nın dağılımının $\boldsymbol{\mu}_{\boldsymbol{\beta}}$ ortalamasına ve iki $\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}$ ve $\boldsymbol{\Sigma}_{\boldsymbol{\beta}}$ kovaryans matrisine bağlıdır.

Bir rasgele etkiler modeli, bir hiyerarşik lineer modeli, bir deneysel Bayes modeli gibi kapsama dahil ederek veya genelleştirilmiş tahmin etme denklemleri olarak hangisinin bilindiğinin kullanılmasıyla, bu modeli uydurabilmenin çeşitli yorumları ve yöntemleri vardır. Bu yaklaşımlar durumdan duruma onların basitlik veya uygulanabilirliğinde bir dereceye kadar farklıdır. Buna rağmen onlar genellikle denktirler ve genel olarak benzer sonuçları verirler (Wickens, 2004).

3.1.12. Eşanlı Testetme

Son olarak, çoklu testler yapıldığında, ortaya çıkan problemleri, yani görüntü verisinde kayda değer önem konusunu göz önüne almak istiyoruz. Sadece geleneksel çözümlerden bahsedeceğiz.

Standart hipotezi test etme çerçevesinde ilişkinin var olmadığı bir H_0 sıfır hipotezi, bilinmeyen ilişkinin var olduğu H_a olumsuzlanması ile karşılaştırılır. Bu

alternatifler arasında karar vermek için herhangi bir şemanın özellikleri iki şartlı hata oranı ile ölçülür.

1.Tip : $\alpha = P (H_0 \text{ in yanlış olduğuna karar verme } \setminus H_0 \text{ doğru})$

2.Tip : $\beta = P (H_a \text{ in yanlış olduğuna karar verme } \setminus H_a \text{ doğru})$ (3.55)

dir. Bu olasılıklar olası doğru ifadeye ve aşağıdaki tanıdık çizelgeye bağlı olur.

Çizelge 3.2. Hipotez testinde varılan karar.

Varılan Karar

Doğru ifade	Etki Var	Etki Yok
Etki Var	$1 - \beta$	β
Etki Yok	α	$1 - \alpha$

Bu hata oranları arasında bir değiş-tokuş vardır. Verilen şu veya bu şekilde daha fazla bilgi çıkarılmadıkça, birini artıran bir yöntem diğerini azaltacaktır. H_0 özel bir hipotez olmakla birlikte, H_a birçok özel alternatiflerin bir birleşimidir. Bu nedenle α genellikle çabucak hesaplanırken, β öyle değildir. Bu yüzden diğeri yok ise, hipotezi test etmeye genel yaklaşım, sadece 1.Tip hatayı kontrol eder.

Bir takım testler yapıldığında, kümelenme veya grup (aile) tarzı hata oranları (yani, bir veya daha fazla hata yapmanın olasılığı) artar. Bir düzeltme formu ortaya konmadıkça, bu artış her iki hatayı da etkiler. Eşanlı-test etme yöntemlerinin amacı bu hata oranlarını rasyonel bir şekilde yönetmektir.

Özellikle, farklı 2. Tip hata oranının ilginç bulguların kaybolduğu noktaya düşmesine bir arada izin vermeksizin, kontrol 1. Tip kümelenme hata oranı üzerinde uygulanacaktır.

Bu işi başarmak için her biri kendi avantajlarına ve sınırlamalarına sahip birçok yöntem geliştirilmiştir. Onlar kabaca ara sıra örtüşen üç sınıfa ayrılabilir.

a) Bonferroni Testleri : Bonferroni eşitsizliği olarak bilinen temel bir olasılık ilişkisi k -tane testin herhangi bir kümesi için kümelenme ve ayrı ayrı (bireysel) hata oranları arasındaki ilişkinin

$$\alpha_{\text{kümeleme}} \leq 1 - (1 - \alpha_{\text{bireysel}})^k < k\alpha_{\text{bireysel}} \quad (3.56)$$

kümelenme hata oranı daima $k \times \text{bireysel}$ hata oranından küçüktür, bu nedenle o, bireysel hata oranını istenen değerin bir oranıyla yazmak suretiyle sınırlanabilir.

$$\alpha_{\text{bireysel}} = \frac{\text{İstenen } \alpha_{\text{kümelenme}}}{k} \quad (3.57)$$

dir. Bonferroni eşitsizliği bir eşitsizlik olduğundan, gerçek kümelenme hatası onun nominal değerinden küçüktür (Rencher ve Schaalje, 2007).

b) Birleşim-Kesişim Prensibi : Birçok eşanlı test etme durumlarında, tüm bireysel sıfır hipotezlerini sağlayan birleşik H_{C0} sıfır hipotezi için, çok amaçlı bir test mevcuttur. Örneğin, genel lineer modelde $\beta = 0$ hipotezinin genel testi her bir bireysel $\beta_j = 0$ testlerini birleştirir. Bu nedenle onun alternatifi onun yokluğunun birleşimi iken, birleşik sıfır hipotezi bu basit hipotezlerin kesişimidir. Örneğin, eğer β üç bileşene sahipse, birleşik hipotez ve alternatifi,

$$\begin{aligned} H_{C0} &: (\beta_1 = 0) \cap (\beta_2 = 0) \cap (\beta_3 = 0), \\ H_{Ca} &: (\beta_1 \neq 0) \cup (\beta_2 \neq 0) \cup (\beta_3 \neq 0) \end{aligned} \quad (3.58)$$

dir. Genel olarak, bireysel hipotezleri H_{01}, H_{02}, \dots olarak yazarak,

$$\begin{aligned} H_{C0} &: H_{01} \cap H_{02} \cap \dots \\ H_{Ca} &: H_{a1} \cup H_{a2} \cup \dots \end{aligned} \quad (3.59)$$

olur. Bu durumda kümelenme hata oranını kontrol etmek için, H_{C0} in çok amaçlı testinin istenen $\alpha_{\text{kümelenme}}$ de reddedilmesine de sebep olacak olan H_{0j} bileşen testlerinin herhangi biri için yeteri derecede büyük olması gereken delilin düzeyini belirlemek yeterlidir. Varyans analizinde Scheffe'in testi bu prensibi kullanır (Rencher ve Schaalje, 2007).

c) Ardışık (sıralı) Test Etme : Bu yaklaşımlar birçok aşama test etme yöntemini benimser. Birçok amaçlı test (veya testlerinin serisi) ilk olarak genel hata oranını $\alpha_{\text{kümelenme}}$ ye göre sınırlamak için uygulanır. Başlangıç testi genel 1. Tip hata oranını kontrol ettiğinden, ardışık testler genel hata oranını riske almaksızın az uyulması gereken standartları benimser. Ardışık testlerin bilinen örnekleri varyans analizindeki Fisher ve Newmann-Keuls yöntemleridir. Bu yöntemlerin tümü 1. Tip hatalarının kontrolü üstünde durur ve daha fazla uyulması gereken ölçütleri benimsemeye onlar 2. Tip hata oranını artırırlar.

Bu stratejilerin her biri onların güçlü ve zayıf noktalarına sahiptir. Kısaca, Bonferroni yöntemi testlerin sayısı artarken, özellikle onlar bağımsız olmadığında, giderek artan bir şekilde tutucu (konservatif) olur. Birleşim-Kesişim ilişkisine dayanan testler de

epey konservatif olmaya meyleder ve onlar sadece bir tatmin edici çok amaçlı istatistik mevcut olduğunda kullanılabilir. Ardışık yöntemler çoğu kez diğerlerinden daha güçlüdür, fakat onlar gücünü onların hipotezlerine göre daha karmaşık bir ilişkiden alırlar. Sıfır hipotezi kısmen yanlış olduğunda onlar özel olarak yetersizdir. Çünkü sadece H_{C0} in düzeyinde uygulanır.

Böyle veri ile, potansiyel althipotezlerin sayısı, althipotezler hayli yüksek derecede bağımlıdır ve birleşik sıfır hipotezi kısmen (veya tam olarak bile) yanlıştır. Yukarıdaki çizelge yerine aşağıdaki çizelgeyi elde ederiz (Wickens, 2004).

Çizelge 3.3. Ardışık test etmede varılan karar.

Varılan Karar

Etkinin Büyüklüğü	Teşhis Edilen Bağlantı	Teşhis Edilmeyen Bağlantı
Enteresan	Yararlı Bulgu	Kaçırılan Fırsat
Enteresan Değil	Yanlıştır Öncü	Doğru İptal

3.2. Lineer Olmayan Regresyon Modeline Giriş

Kısım 3.1 de lineer modellerle ilgili ayrıntılı bilgiler verdik. Şimdi de lineer olmayan modellerden bahsedeceğiz. Lineer olmayan regresyon modeline geçmeden önce regresyon modellerinden bazılarını kısaca tanıtaacağız. Sonra lineer olmayan modellerde alışılmış en küçük kare yöntemine değineceğiz ve bir uygulama yapacağız. Ayrıca regresyon analizinde, bir düzlem üzerindeki noktaların bir kümesine kuadratik eğrilerin ve yüzeylerin lineer ve lineer olmayan en küçük kareler yöntemiyle en iyi uyumundan kısaca bahsedeceğiz.

3.2.1. Regresyon Modelleri

a) Regresyon Modeli : Regresyon; ilgili bir y değişkeni ve bir veya daha fazla açıklayıcı \mathbf{x}_k ($k = 1, 2, \dots, p$) değişkeni arasındaki ilişkiyi inceler. Genel model

$$y_i = h(x_{i1}, x_{i2}, \dots, x_{ip} ; \beta_1, \beta_2, \dots, \beta_p) + \varepsilon_i \quad (i = 1, 2, \dots, n) \quad (3.60)$$

dir. Burada h ; $\mathbf{x} = [x_{i1}, x_{i2}, \dots, x_{ip}]$ ve $\boldsymbol{\beta} = [\beta_1, \beta_2, \dots, \beta_p]'$ vektörleri ile kısıtlamayı istediğimiz açıklayıcı değişkenlere ve parametrelere bağlı olan uygun bir fonksiyondur. h fonksiyonundan planlanmamış sapmalar, ε_i rasgele hatalarıyla ifade edilir. Bu rasgele hatanın dağılımı için normal dağılım kabul edilir, bu nedenle

$$\varepsilon_i \sim N(0, \sigma^2) \quad (3.61)$$

dir, yani hatalar bağımsızdır.

b) Lineer Regresyon Modeli : (Çoklu) Lineer regresyonda, h lineer regresyon fonksiyonlarının β_k parametrelerine göre lineer oldukları, yani

$$h(x_{i1}, x_{i2}, \dots, x_{ip}; \beta_1, \beta_2, \dots, \beta_p) = \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_p x_{ip} \quad (3.62)$$

olduğu düşünülür.

c) Lineer Olmayan Regresyon Modeli : Lineer olmayan regresyonda, h fonksiyonlarının parametrelere göre lineer olarak yazılamadığı düşünülür. Genellikle böyle fonksiyonlar teoriden ortaya konur. Prensipinde, modelin deterministik kısmını belirlemek için sınırsız olasılıklar vardır. Bu esneklik, genellikle istatistiksel ifadelerle daha büyük etkiyi ifade eder.

d) Lineerleştirilebilir Regresyon Fonksiyonları : Bazı lineer olmayan regresyon fonksiyonları, ilgilenilen değişken ve açıklayıcı değişkenlerin lineer dönüşümleri vasıtasıyla lineerleştirilebilir. Örneğin, bir

$$h(\mathbf{x}; \boldsymbol{\beta}) = \beta_1 \mathbf{x}^{\beta_2} \quad (3.63)$$

kuvvet (güç) fonksiyonu, $\theta_0 = \ln(\beta_1)$, $\theta_1 = \beta_2$ ve $\tilde{x} = \ln(x)$ olmak üzere, bir lineer (parametrelere göre)

$$\ln \langle h \langle x; \theta \rangle \rangle = \ln \langle \theta_1 \rangle + \theta_2 \ln \langle x \rangle = \beta_0 + \beta_1 \tilde{x} \quad (3.64)$$

fonksiyonu için dönüştürülebilir. Eğer onu bağımsız değişkenlerin dönüşümleri ve sonucun bir monoton fonksiyonuyla (bilinmeyen) parametreler cinsinden bir lineer fonksiyona dönüştürebilsek, h regresyon fonksiyonu lineerleştirilebilir deriz.

Birkaç lineerleştirilebilir fonksiyon örnekleri (Daniel ve Wood, 1980):

$$\begin{aligned}
h\langle x; \theta \rangle &= 1/(\theta_1 + \theta_2 \exp\langle -x \rangle) & \leftrightarrow & \quad 1/h\langle x; \theta \rangle = \theta_1 + \theta_2 \exp\langle -x \rangle \\
h\langle x; \theta \rangle &= \theta_1 x / (\theta_2 + x) & \leftrightarrow & \quad 1/h\langle x; \theta \rangle = 1/\theta_1 + \theta_2 / \theta_1 x \\
h\langle x; \theta \rangle &= \theta_1 x^{\theta_2} & \leftrightarrow & \quad \ln\langle h\langle x; \theta \rangle \rangle = \ln\langle \theta_1 \rangle + \theta_2 \ln\langle x \rangle \\
h\langle x; \theta \rangle &= \theta_1 \exp\langle \theta_2 g\langle x \rangle \rangle & \leftrightarrow & \quad \ln\langle h\langle x; \theta \rangle \rangle = \ln\langle \theta_1 \rangle + \theta_2 g\langle x \rangle \\
h\langle x; \theta \rangle &= \exp\langle -\theta_1 x^{(1)} \exp\langle -\theta_2 / x^{(2)} \rangle \rangle & \leftrightarrow & \quad \ln\langle \ln\langle h\langle x; \theta \rangle \rangle \rangle = \ln\langle -\theta_1 \rangle + \ln\langle x^{(1)} \rangle - \theta_2 / x^{(2)} \\
h\langle x; \theta \rangle &= \theta_1 (x^{(1)})^{\theta_2} (x^{(2)})^{\theta_3} & \leftrightarrow & \quad \ln\langle h\langle x; \theta \rangle \rangle = \ln\langle \theta_1 \rangle + \theta_2 \ln\langle x^{(1)} \rangle + \theta_3 \ln\langle x^{(2)} \rangle.
\end{aligned} \tag{3.65}$$

e) İstatistiksel Olarak Tam Model : E_i rasgele hatalarının tümü aynı normal dağılıma sahip olmak üzere, başvurulan örnekteki lineerleştirilen regresyon ile bir lineer regresyon

$$\ln\langle Y_i \rangle = \beta_0 + \beta_1 \tilde{x}_i + E_i \tag{3.66}$$

modeline dayanır. Bu modeli geri dönüştürür ve böylece $\tilde{E}_i = \exp\langle E_i \rangle$ olmak üzere,

$$Y_i = \theta_1 . x^{\theta_2} . \tilde{E}_i \tag{3.67}$$

elde ederiz. Şimdi \tilde{E}_i , $i = 1, \dots, n$ hataları çarpımsal olarak katkıda bulunur ve lognormal dağılırlar. Bu nedenle rasgele sapmalar hakkındaki varsayımlar toplamsal olarak katkıda bulunan ve özel bir normal dağılıma sahip olan alışılabilir E_i^* rasgele hatalı, doğrudan doğruya h a dayanan bir

$$Y_i = \theta_1 . x^{\theta_2} + E_i^* \tag{3.68}$$

modeli için olandan kesin bir şekilde farklıdır. Bu nedenle regresyon fonksiyonunun bir lineerleştirilmesi akla uygundur.

3.2.2. Lineer Olmayan Modellerde Alışılmış En Küçük Kare Yöntemi ve Uygulaması

3.2.2.1. Lineer Olmayan Modeller için Minimumlaştırmanın En Küçük Kareler Kriteri

Aşağıdaki (3.69) regresyon modelini göz önüne alalım:

$$y_i = \beta_1 . \beta_2^{x_i} . E_i \tag{3.69}$$

(3.69) bağıntısına benzer modeller doğası itibarıyla uygun (log) dönüşümle modellerin β_1 ve β_2 parametrelerine göre lineer yapılabilmesi anlamında lineer (parametreye göre) regresyon modelleridir (Modelin β_1 ve β_2 ye göre lineer olmadığına dikkat edelim.). Bu durumda, hata değişkenliğinin tüm x_i de sabit olmadığını, yani hatanın toplamsal olmadığını kabul ederiz. Büyük bir olasılıkla, x_i nin farklı düzeylerinde daha rasgele ve önceden kestirilemeyen dalgalanma vardır.

$$\prod E_i = 1 \quad \text{veya} \quad \prod \text{Ln}E_i = 0 \quad (3.70)$$

olduğunu kabul ederiz. (3.69) bağıntısının her iki yanının logaritmalarını alarak,

$$\text{Lny}_i = \text{Ln}\beta_1 + x_i \text{Ln}\beta_2 + \text{Ln}E_i \quad (3.71)$$

bağıntısını elde ederiz.

$$\sum (\text{Lny}_i - \text{Ln}b_1 - x_i \text{Ln}b_2)^2 \quad (3.72)$$

minimum kriteri altında örneklem regresyonu için OLS (Alışılmış En Küçük Kareler) yi kullanarak diferensiyelleme yöntemi β_1 ve β_2 yi tahmin etmek için aşağıdaki normal denklemleri ortaya koyar:

$$\begin{cases} \sum \text{Lny}_i = n \cdot \text{Ln}b_1 + \text{Ln}b_2 \cdot \sum x_i \\ \sum x_i \cdot \text{Lny}_i = \text{Ln}b_1 \cdot \sum x_i + \text{Ln}b_2 \cdot \sum x_i^2 \end{cases} \quad (3.73)$$

Normal denklemleri aynı anda çözerek

$$\text{Ln}b_1 = \frac{\sum \text{Lny}_i}{n} - \text{Ln}b_2 \cdot \frac{\sum x_i}{n} = \frac{\sum \text{Lny}_i}{n} - \text{Ln}b_2 \cdot \bar{x} \quad (3.74)$$

bağıntısını elde ederiz.

$\bar{x} = 0$ olsun, bu takdirde

$$\text{Ln}b_1 = \frac{1}{n} \sum \text{Lny}_i \quad \text{veya} \quad b_1 = \sqrt[n]{y_1 \cdot y_2 \cdots y_n} \quad (3.75)$$

olur.

Bu, β_1 parametresinin y değişkeninin geometrik ortalaması olduğunu söyler. Klasik normal lineer regresyon modelini kullanmak için $\text{Ln}E_i \sim N(0, \sigma^2)$ olduğunu kabul etmeliyiz. Bu nedenle, (3.71) bağıntısını çalıştırdığımızda, bu regresyondan

elde edilen hatalara (hata tahminlerine) normallik testleri uygulamalıyız (Gujarati, 1995).

Şimdi aşağıdaki (3.76) regresyon modelini göz önüne alalım. Bu model aslında parametreye göre lineer değildir. Bu modelde hata değişkenliğinin x_i den bağımsız olduğunu, yani, tahmin edilen hata değişkenliğinin miktarının herhangi bir x_i de aynı olduğunu kabul ederiz. Modelimiz

$$y_i = \beta_1 \cdot \beta_2^{x_i} + E_i \quad (3.76)$$

dir.

$$\sum (y_i - b_1 \cdot b_2^{x_i})^2 \quad (3.77)$$

minimum kriteri altında örneklem regresyon fonksiyonu için OLS yi kullanarak, diferensiyelleme yöntemi β_1 ve β_2 yi tahmin etmek için aşağıdaki normal denklemleri ortaya koyar:

$$\begin{cases} \sum 2(y_i - b_1 \cdot b_2^{x_i}) \cdot b_2^{x_i} = 0 \\ \sum 2(y_i - b_1 \cdot b_2^{x_i}) \cdot b_1 \cdot x_i \cdot b_2^{x_i} / b_2 = 0. \end{cases} \quad (3.78)$$

Bu normal denklemleri analitik olarak çözmek mümkün değildir ve lineer olmayan durumda lineer olmayan alışılmış en küçük kareler tahmini modelin parametrelere göre bir lineerleştirilmesini kullanarak iteratif olarak yapılabilir. (3.77) minimum kriterinden elde edilen β_1 ve β_2 parametreler tahmininin, (3.72) minimum kriterinden elde edilen β_1 ve β_2 parametrelerinin tahmininden yanlı oldukları açıktır. Bu modele giren stokastik hata teriminin özellikleri hakkında dikkatli olmalıyız. Hipotezi test etmek için (3.76) modeli için E_i stokastik hata teriminin normal dağılım izlediğini kabul etmeliyiz. Ancak (3.69) modeli ve onun (3.71) emsali için, $LnE_i \sim N(0, \sigma^2)$ ve sırasıyla E_i stokastik hata teriminin aşağıda verilen (3.79) ortalamalı ve (3.80) varyanslı log-normal dağılıma uyduğunu kabul etmeliyiz:

$$e^{\sigma^2/2}. \quad (3.79)$$

$$e^{\sigma^2/2} (e^{\sigma^2/2} - 1). \quad (3.80)$$

3.2.2.2. Dönüştürülmüş Modele OLS nin Grafiksel Uygulaması

Ordu Büyükşehir Belediyesi'ndeki nüfus miktarının katı atık miktarı üzerindeki etkisini belirleyelim. Analiz Ordu ilinin 7 farklı ilçesinde 2015 yılı kış dönemi için ilçe belediyelerince toplanan atığa ve ilçelerdeki nüfusa dayanır.

Regresyon analizi için aşağıdaki hipotez seçilir:

Katı atık miktarı nüfus miktarına bağlıdır.

y_i ; 7 belediyedeki 2015 yılında kış döneminde toplam ortalama katı atık miktarı ve x_i ; 7 belediyedeki nüfus olmak üzere, nüfusa bağlı atık modellemesi için aşağıdaki model seçildi:

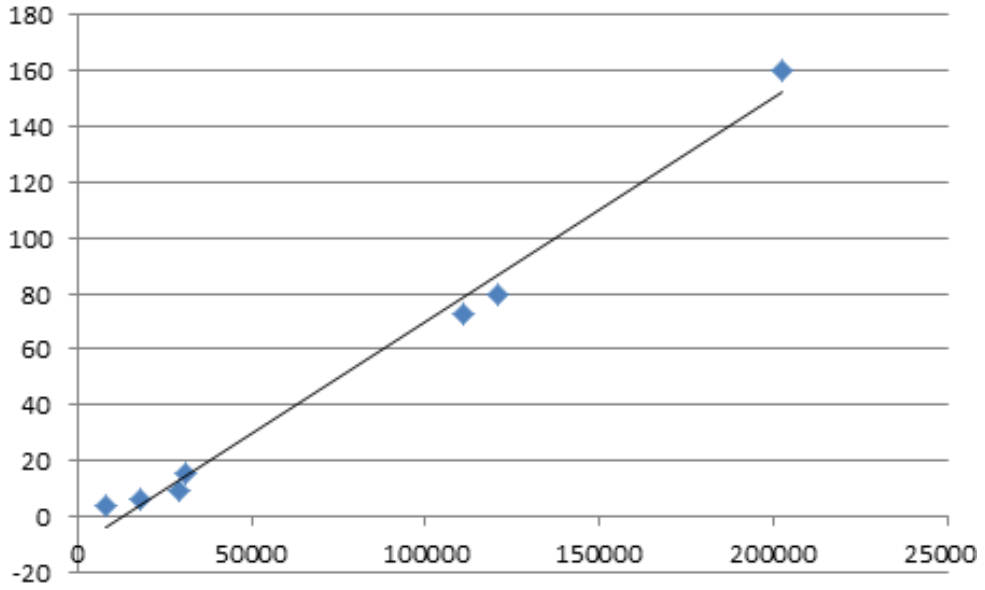
$$y_i = \beta_1 \cdot \beta_2^{x_i} \cdot E_i$$

$$b_1 = \sqrt[n]{y_1 \cdot y_2 \dots y_n} = 22.62 \quad \text{ve} \quad \text{Ln}b_1 = \frac{\sum \text{Ln}y_i}{n} - \text{Ln}b_2 \cdot \bar{x} \Rightarrow b_2 = 1.000$$

elde edilir. Regresyon sonuçları $\text{Ln}\beta_1$ ve β_2 parametrelerinin her ikisinin de anlamlı olduklarını (p-değeri<0.05) ve nüfus miktarının atılan atığa bağlı olduğunu gösterir (bkz Çizelge 3.4 ve Şekil 3.5).

Çizelge 3.4. İlçelerin nüfus miktarı ve toplam ortalama katı atık miktarı.

İLÇELER	NÜFUS	TOP. ORT. KATI ATIK MİKTARI (TON/GÜN)
Altınordu	202310	160
Fatsa	111070	73
Ünye	120720	80
Perşembe	31094	15
Ulubey	18239	6
Korgan	29349	9
Gülyalı	7994	4



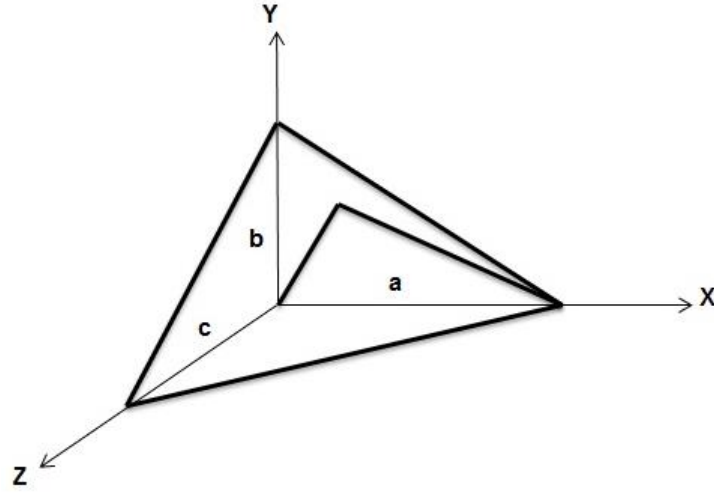
Şekil 3.5. Nüfus miktarının katı atık miktarı üzerindeki etkisi örneğinin grafiği.

3.2.3. En Küçük Kare Uyum Unsuru

En küçük kare ölçütünün uygulaması eğri uydurma problemlerinin geniş bir yelpazesine uygulanabilir. Veriye en küçük kare en iyi uyum unsuru, veriyi bir düzleme uydurma problemi ele alarak açıklanır. Bu; bir parametreleme problemidir. En iyi düzlem, düzlem üzerindeki bir $C(x_0, y_0, z_0)$ noktası ve düzleme göre normalin (a, b, c) doğrultu kosinüsü vasıtasıyla belirlenebilir. Düzlem üzerindeki herhangi bir (x, y, z) noktası

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0 \quad (3.81)$$

bağıntısını sağlar.



Şekil 3.6. Genel Düzlem Geometrisi.

Herhangi bir $A(x_i, y_i, z_i)$ noktasından yukarıda belirlenen düzleme uzaklık

$$d_i = a(x_i - x_0) + b(y_i - y_0) + c(z_i - z_0) \quad (3.82)$$

ile verilir. Her bir noktanın düzlemden uzaklıklarının kareleri toplamı

$$F = \sum_{i=1}^n d_i^2 \quad (3.83)$$

dir. Bu nedenle problemimiz, F toplamını minimum yapan (x_0, y_0, z_0) ve (a, b, c) parametrelerini bulmaktır.

➤ **Optimizasyon :**

$\mathbf{u} = (u_1, \dots, u_n)$ parametrelerine göre minimumlaştırılacak olan bir

$$E(\mathbf{u}) = \sum_{i=1}^n d_i^2(\mathbf{u}) \quad (3.84)$$

fonksiyonunu göz önüne alalım. Burada d_i , veri noktasının \mathbf{u} ile parametrelenen geometrik unsura (ögeye) uzaklığını gösterir.

➤ **Gauss-Newton Algoritması :**

$f(x)$ lineer olmayan bir fonksiyon olmak üzere, bir $f(x) = 0$ kök bulma problemini çözmek için Newton'un yöntemi en güçlü ve meşhur nümerik yöntemlerden biridir. Bu teknik; çember, küre, silindir ve konilerin lineer en küçük kareler modelinin uygulanmasında kullanılır. Böyle bir algoritmanın nasıl çalıştığını anlamak için ilk olarak Newton yöntemini ifade edelim.

Bir f fonksiyonunun verildiğini ve onun tanım bölgesinin $[a, b]$ olduğunu farz edelim ve $f(\bar{x}) = 0$ olacak şekilde $\bar{x} \in [a, b]$ olsun. $f(x)$ için \bar{x} civarında Taylor polinomu açılımı $\chi(x) \in [x, \bar{x}]$ olmak üzere,

$$f(x) = f(\bar{x}) + (x - \bar{x})f'(\bar{x}) + \frac{(x - \bar{x})^2}{2} f''(\chi(x)) \quad (3.85)$$

dir. $q = x$ olmak üzere $f(q) = 0$ olduğundan (3.85),

$$0 = f(\bar{x}) + (x - \bar{x})f'(\bar{x}) + \frac{(x - \bar{x})^2}{2} f''(\chi(x))$$

olduğunu ortaya koyar.

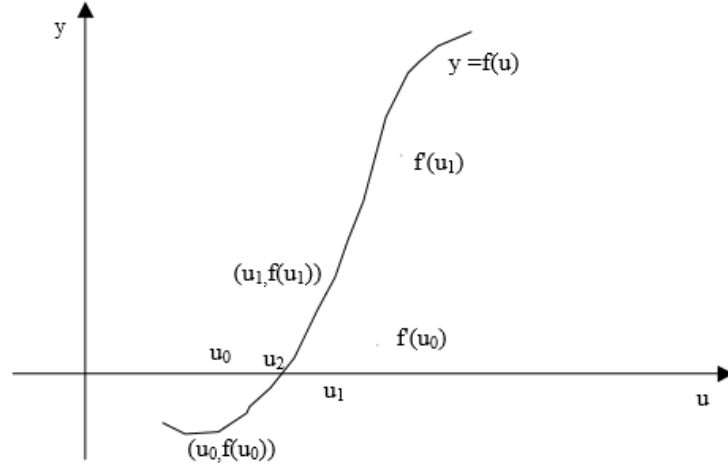
Newton'un yöntemi $|q - \bar{x}|$ küçük olduğunda kabul edilebilir. Bu nedenle $(p - \bar{x})^2$ çok küçüktür. Bu nedenle,

$$\begin{aligned} 0 &= f(\bar{x}) + (q - \bar{x})f'(\bar{x}) \\ (q - \bar{x}) &= -\frac{f(\bar{x})}{f'(\bar{x})} \end{aligned}$$

dir. $q = u_1$, $\bar{x} = u_0$ ve $p = -\frac{f(\bar{x})}{f'(\bar{x})}$ olsun. Böylece,

$$u_1 = u_0 + p \quad \text{ve} \quad p = -\frac{f(u_0)}{f'(u_0)} \quad (3.86)$$

dır.



Şekil 3.7. Newton Yönteminin İterasyonu.

Şekil 3.7 den $y = f(u)$ kök bulma problemi aşağıdaki adımlarla çözülebilir.

- 1) $f(u_0)$ ve $f'(u_0)$ değerleri bulunur.
- 2) $(u_0, f(u_0))$ da grafiğe bir teğet doğru bulunur. O, eksen u_1 de keser.
- 3) u_1 noktası bulunur ve u_2 noktasını bulmak için grafiğe u_1 noktasında bir teğet çizilir.
- 4) Yukarıdaki bu adımlar yakınsayınca kadar (u^* a yakın) tekrar edilir.

Algoritmanın ana içerikleri şunlardır:

- a) Bir yaklaşık çözüm verildiğinde, problem lineerleştirilir.
- b) Problemin lineer versiyonu çözülür.
- c) Lineer problemin çözümü, çözümün tahminini güncelleştirmek (yenilemek) için kullanılır.
- d) Lineer en küçük kareler probleminde, geometrik ögenin veri noktası kümesine ne kadar iyi uyduğunun ölçümü

$$E = \sum_1^n d_i^2 \quad (3.87)$$

ile belirlenebilir. Burada d_i noktadan geometrik ögeye bir uzaklıktır.

Bu durumda, d_i lineer fonksiyon değildir; bir E kareler toplamının minimumunu bulmak için Gauss-Newton algoritması kullanılır.

Bir u^* başlangıç tahmininin var olduğunu kabul ederek,

$$Jp = -d$$

biçimindeki bir lineer en küçük kareler sistemi çözülür. Burada J ; i -yinci satırı u parametresine göre d_i nin gradiyenti olan $m \times n$ Jakobiyen matristir, yani,

$$J_{ij} = \frac{\partial d_i}{\partial u_j} \quad (3.88)$$

dir. O; u da değerlendirilir ve d nin i -yinci bileşeni $d_i(u)$ dur. Parametre

$$u := u + p \quad (3.89)$$

olarak yenilenir.

➤ Yakınsayan Şartlar :

Newton algoritmasının adımları, yakınsama noktasına varıncaya kadar tekrarlanır. Burada 3 kriter, yakınsamayı test etmek için uygulanabilir:

- 1) E değişimi küçük olmalıdır.
- 2) Güncelleme büyüklüğü, örneğin $(p'p)^{1/2}$ küçük olmalıdır.
- 3) E nin optimizasyon parametrelerine göre kısmi türetimi, örneğin $(gg')^{1/2}$ küçük olmalıdır. Burada $g = J'd$ dir.

3.2.3.1. En Küçük Kareler En İyi Uyum Doğrusu

Doğrular belirlenen bir düzlemde (2 boyutta) veya 3 boyutta olabilir. $m \geq 2$ olmak üzere, m sayıda (x_i, y_i, z_i) noktalarına bir doğru uydurmak için yöntem aşağıda verilir.

Doğru üzerindeki herhangi bir (x, y, z) noktası, herhangi bir t değeri için

$$(x, y, z) = (x_0, y_0, z_0) + t(a, b, c) \quad (3.90)$$

bağıntısını sağlar. 3 boyutta bir noktadan bir doğruya uzaklık,

$$d_i = \sqrt{[u_i^2 + v_i^2 + w_i^2]} \quad (3.91)$$

olarak bilinir.

➤ Parametreleme :

Bir doğru

- i) üzerindeki bir (x_0, y_0, z_0) noktası ve
- ii) (a, b, c) doğrultu kosinüsleri

ile belirlenebilir.

➤ **Algoritma Tanımı :**

En iyi uyum doğrusu verinin $(\bar{x}, \bar{y}, \bar{z})$ kitle merkezinden geçer ve bu L (doğru) üzerindeki bir noktayı aynı zamanda doğrultu kosinüslerinin bulunabileceğini belirtir.

- i) Birinci adım x , y ve z noktalarının ortalamasını bulmaktır.

$$\bar{x} = \sum x_i / n$$

$$\bar{y} = \sum y_i / n$$

$$\bar{z} = \sum z_i / n$$

dir.

- ii) Birinci sütunu $x_i - \bar{x}$, ikinci sütunu $y_i - \bar{y}$ ve üçüncü sütunu $z_i - \bar{z}$ olacak şekilde **A** matrisi formülendir.
- iii) Bu **A** matrisi singüler değer ayrışımı ile çözülür. **A** nın en küçük singüler değeri matristen seçilir ve karşılık gelen singüler vektör seçilir ki onun doğrultu kosinüsleri (a, b, c) dir.
- iv) En iyi uyum düzlemi $\bar{x}, \bar{y}, \bar{z}, a, b$ ve c ile belirlenir.

3.2.3.2. En Küçük Kare En İyi Uyum Düzlemi

$m \geq 3$ olmak üzere, m sayıda (x_i, y_i, z_i) noktasına bir düzlem uydurmak için yöntem aşağıda verilir.

Düzlem üzerindeki herhangi bir (x, y, z) noktası

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0$$

denklemini sağlar. Bir (x_i, y_i, z_i) noktasından, x_0, y_0, z_0, a, b ve c ile belirlenen bir düzleme uzaklığın

$$d_i = a(x - x_0) + b(y - y_0) + c(z - z_0)$$

ile verildiği bilinir.

➤ **Parametreleme :**

Bir düzlem

- i) düzlem üzerindeki bir (x_0, y_0, z_0) noktası ve
- ii) düzleme göre normalin (a, b, c) doğrultu kosinüsleri ile belirlenebilir.

➤ **Algoritma Tanımı :**

En iyi uyum düzlemi verinin $(\bar{x}, \bar{y}, \bar{z})$ kitle merkezinden geçen ve bu P düzlemi üzerindeki bir noktayı aynı zamanda doğrultu kosinüslerinin bulunabileceğini belirler.

Bunun için (a, b, c) , $\mathbf{B} = \mathbf{A}'\mathbf{A}$ nın en küçük özdeğeri ile ilgili özvektördür.

- i) İlk adım, x, y ve z noktalarının

$$\begin{aligned}\bar{x} &= \sum x_i/n \\ \bar{y} &= \sum y_i/n \\ \bar{z} &= \sum z_i/n\end{aligned}$$

ortalamasını bulmaktır.

- ii) İlk sütunu $x_i - \bar{x}$, ikinci sütunu $y_i - \bar{y}$ ve üçüncü sütunu $z_i - \bar{z}$ olacak şekilde \mathbf{A} matrisi formülendir.
- iii) Bu \mathbf{A} matrisi singüler değer ayrışımı ile çözülür. \mathbf{A} nın en küçük singüler değeri matristen seçilir ve karşılık gelen singüler vektör seçilir ki onun doğrultu kosinüsleri (a, b, c) dir.
- iv) En iyi uyum düzlemi $\bar{x}, \bar{y}, \bar{z}, a, b$ ve c ile belirlenir.

3.2.3.3. En Küçük Kare En İyi Uyum Çemberi

$m \geq 3$ olmak üzere, m sayıda (x_i, y_i) veri noktasına bir çember uydurmak için yöntem aşağıda verilir.

Çember üzerindeki herhangi bir (x, y) noktası

$$(x_i - x_0)^2 + (y_i - y_0)^2 = r^2 \quad (3.92)$$

denklemini sağlar. Bir (x_i, y_i) noktasından x_0, y_0 ve r ile belirlenen bir çembere uzaklık, $r_i = \sqrt{(x_i - x_0)^2 + (y_i - y_0)^2}$ olmak üzere,

$$d_i = r_i - r \quad (3.93)$$

ile verilir ve J Jakobiyen matrisinin elemanları, d_i nin x_0, y_0 ve r ye göre kısmi türevinden bulunur ve

$$\begin{aligned}\frac{\partial d_i}{\partial x_0} &= -(x_i - x_0)^2 / r_i \\ \frac{\partial d_i}{\partial y_0} &= -(y_i - y_0)^2 / r_i \\ \frac{\partial d_i}{\partial r} &= -1\end{aligned}\tag{3.94}$$

ile verilir.

➤ **Parametreleme :**

Bir çember

- i) onun (x_0, y_0) merkezi ve
- ii) onun r yarıçapı ile belirlenir.

➤ **Algoritma Tanımı :**

En uygun çemberi bulmak için kullanılan algoritma daha evvel açıklanan Gauss-Newton algoritmasıdır. İlk olarak giriş tahminleri bulunur ve sonra da Gauss-Newton algoritması düzenlenir. Çemberin merkezi ve yarıçapının giriş tahminleri, bir lineer en küçük kare modeli olarak problemi çözümlenir.

İzleyen adımlar aşağıdaki gibidir.

- i) $f_i = r_i^2 - r^2$ olmak üzere,

$$F = \sum_{i=1}^m f_i^2$$

olarak ifade edilen F minimumlaştırılır.

- ii) Bu, ρ

$$x_0^2 + y_0^2 - r^2$$

olmak üzere, x_0, y_0 ve ρ ya göre

$$\begin{aligned}f_i &= (x_i - x_0)^2 + (y_i - y_0)^2 - r^2 \\ &= -2x_i x_0 - 2y_i y_0 + (x_0^2 + y_0^2 - r^2) + (x_i^2 + y_i^2)\end{aligned}$$

olarak kısaltılabilir.

- iii) F yi minimumlaştırmak için,

$$\mathbf{A} \begin{bmatrix} x_0 \\ y_0 \\ r \end{bmatrix} = \mathbf{b}$$

lineer en küçük kareler sistemi çözülür. Burada \mathbf{A} nın i -yinci satırının elemanları $(2x_i, 2y_i, -1)$ katsayılarıdır ve \mathbf{b} nin i -yinci elemanı $x_i^2 + y_i^2$ dir.

iv) r nin bir tahmini ρ nun denkleminde elde edilir. Bir kez giriş tahminleri elde edildiğinde, \mathbf{d} sağ yan vektörü ve J jakobiyen matrisi oluşturulur. Bundan sonra

$$J \begin{bmatrix} \rho_{x_0} \\ \rho_{y_0} \\ \rho_r \end{bmatrix} = -\mathbf{d}$$

lineer en küçük kareler sistemi çözülür. Parametrelerin değerleri

$$x_0 := x_0 + \rho_{x_0}$$

$$y_0 := y_0 + \rho_{y_0}$$

$$r := r + \rho_r$$

tahminlerine göre yenilenir.

Yukarıdaki adımlar algoritma yakınsayınca kadar tekrarlanır.

3.2.3.4. En Küçük Kareler Yöntemini Kullanarak En İyi Çember Uydurma

\mathbb{R}^2 de $\{(x_i, y_i) | 0 \leq i < N\}$ diyeceğimiz noktaların sonlu kümesi verilsin. Noktalara (bir en küçük kareler anlamında) en iyi uyan çemberi bulmak istiyoruz.

$$\bar{x} = \frac{1}{N} \sum_i x_i \quad \text{ve} \quad \bar{y} = \frac{1}{N} \sum_i y_i$$

eşitliklerini tanımlayalım ve $0 \leq i < N$ için $u_i = x_i - \bar{x}$, $v_i = y_i - \bar{y}$ olsun. İlk olarak problemi (u, v) koordinatları cinsinden çözeriz ve sonra da (x, y) ye geri dönüşüm yaparız.

Çember; (u_c, v_c) merkezine ve R yarıçapına sahip olsun. $g(u, v) = (u - u_c)^2 + (v - v_c)^2 - \alpha$ ve $\alpha = R^2$ olmak üzere $S = \sum_i (g(u_i, v_i))^2$ yi minimumlaştırmayı istiyoruz. Bunu yapmak için, $S = (\alpha, u_c, v_c)$ yi diferensiyelleriz.

$$\frac{\partial S}{\partial \alpha} = 2 \sum_i g(u_i, v_i) \frac{\partial g}{\partial \alpha}(u_i, v_i) = -2 \sum_i g(u_i, v_i)$$

dir. Bu nedenle, $\partial S / \partial \alpha = 0$ olması için gerek ve yeter şart

$$\sum_i g(u_i, v_i) = 0 \quad (3.95)$$

olmasıdır.

Devam ederek,

$$\begin{aligned} \frac{\partial S}{\partial u_c} &= 2 \sum_i g(u_i, v_i) \frac{\partial g}{\partial u_c}(u_i, v_i) \\ &= 2 \sum_i g(u_i, v_i) 2(u_i - u_c)(-1) \\ &= -4 \sum_i (u_i - u_c) g(u_i, v_i) \\ &= -4 \sum_i u_i g(u_i, v_i) + 4u_c \underbrace{\sum_i g(u_i, v_i)}_{(3.95) \text{ den dolayı } = 0} \end{aligned}$$

elde ederiz.

Bu nedenle, (3.95) bağıntısının varlığında $\partial S / \partial u_c = 0$ ın gerçekleşmesi için gerek ve yeter şart

$$\sum_i u_i g(u_i, v_i) = 0 \quad (3.96)$$

bağıntısının gerçekleşmesidir.

Aynı şekilde, $\partial S / \partial v_c = 0$ gerektirmesi

$$\sum_i v_i g(u_i, v_i) = 0 \quad (3.97)$$

olduğunu ortaya koyar.

(3.96) bağıntısını açarak,

$$\sum_i u_i [u_i^2 - 2u_i u_c + u_c^2 + v_i^2 - 2v_i v_c + v_c^2 - \alpha] = 0$$

elde ederiz.

$S_u = \sum_i u_i$, $S_{uu} = \sum_i u_i^2$ vs. tanımlayarak, bunu

$$S_{uuu} - 2u_c S_{uu} + u_c^2 S_u + S_{uvv} - 2v_c S_{uv} + v_c^2 S_u - \alpha S_u = 0$$

olarak yeniden yazabiliriz.

$S_u = 0$ olduğundan, bu

$$u_c S_{uu} + v_c S_{uv} = \frac{1}{2}(S_{uuu} + S_{uvv}) \quad (3.98)$$

ifadesine sadeleşir.

Benzer biçimde (3.97) bağıntısını açarak ve $S_v = 0$ olduğunu kullanarak

$$u_c S_{uv} + v_c S_{vv} = \frac{1}{2}(S_{vvv} + S_{vuu}) \quad (3.99)$$

olduğu görülür.

(3.98) ve (3.99) bağıntısını eşanlı olarak çözmeye (u_c, v_c) yi verir. Bu takdirde, çemberin orijinal koordinat sistemindeki (x_c, y_c) merkezi; $(x_c, y_c) = (u_c, v_c) + (\bar{x}, \bar{y})$ dir.

R yarıçapını bulmak için, (3.95) bağıntısını açarız ve

$$\sum_i [u_i^2 - 2u_i u_c + u_c^2 + v_i^2 - 2v_i v_c + v_c^2 - \alpha] = 0$$

olduğunu görürüz. Yine $S_u = S_v = 0$ eşitliğini kullanarak,

$$N(u_c^2 + v_c^2 - \alpha) + S_{uu} + S_{vv} = 0$$

elde ederiz. Böylece,

$$\alpha = u_c^2 + v_c^2 + \frac{S_{uu} + S_{vv}}{N} \quad (3.100)$$

dir ve şüphesiz $R = \sqrt{\alpha}$ dır.

Örnek 3.1. $y = x^2 + 4x + 4$ parabolünden birkaç noktayı alalım ve bu noktalara bir çember uyduralım. Kullanılan noktaları veren çizelge aşağıda verilmiştir.

Çizelge 3.5. $y = x^2 + 4x + 4$ parabolünden geçen birkaç nokta.

i	x_i	y_i	u_i	v_i
0	-2.000	0.000	-2.875	-10.938
1	-1.000	1.000	-1.875	-9.938
2	0.000	4.000	-0.875	-6.938
3	1.000	9.000	0.125	-1.938
4	1.500	12.250	0.625	1.312
5	2.000	16.000	1.125	5.062
6	2.500	20.250	1.625	9.312
7	3.000	25.000	2.125	14.062

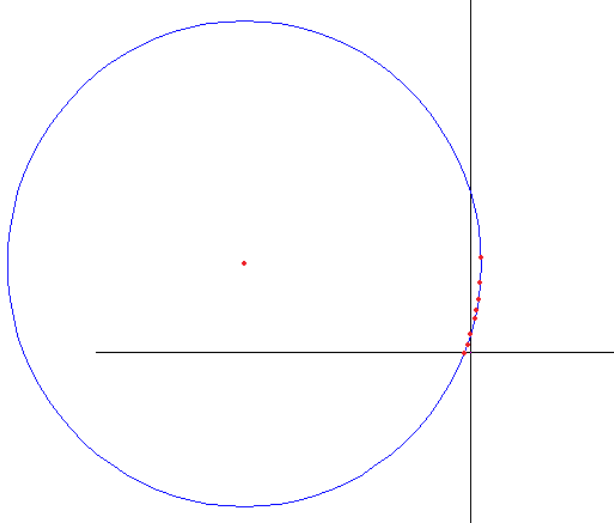
Burada $N = 8$, $\bar{x} = 0.875$ ve $\bar{y} = 10.938$ dir. Aynı zamanda

$$S_{uu} = 21.375, \quad S_{uv} = 107.438, \quad S_{vv} = 582.094, \quad S_{uuu} = -15.469, \\ S_{vvv} = 1088.678, \quad S_{uvv} = 20.213, \quad S_{vuu} = -35.683$$

dür. Bu nedenle, ((3.98) ve (3.99) u kullanarak) (u_c, v_c) için aşağıdaki 2×2 lineer sistemi elde ederiz:

$$\begin{bmatrix} 21.375 & 107.438 \\ 107.438 & 582.094 \end{bmatrix} \begin{bmatrix} u_c \\ v_c \end{bmatrix} = \begin{bmatrix} 2.372 \\ 526.498 \end{bmatrix}.$$

Bu sistemin çözümü $(u_c, v_c) = (-61.362, 12.230)$ olduğunu ortaya koyar ve bu nedenle, $(x_c, y_c) = (-60.487, 23.168)$ dir. Bu değerleri (3.100) bağıntısında yerine koyarak, $\alpha = 3990.345$ olduğu görülür ve buradan $R = 63.170$ dir. Bu örneğin grafiği aşağıda verilmiştir.



Şekil 3.8. $y = x^2 + 4x + 4$ parabolünden geçen noktalara uydurulan çember.

3.2.3.5. En Küçük Kareler Küresi

➤ Parametreleme :

Bir küre (x_0, y_0, z_0) noktası ve r_0 yarıçapıyla belirlenir. Küre üzerindeki herhangi bir nokta,

$$(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2 = r^2 \quad (3.101)$$

denklemini sağlar.

➤ Merkez ve yarıçap için giriş tahminleri :

Bir minimumlaştırma fonksiyonunun seçimi :

Merkez ve yarıçap için bir başlangıç tahmini elde etmek için bir minimumlaştırma fonksiyon tanımlanmalıdır.

$r_i = \sqrt{(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2}$ olmak üzere, $f_1 = r_i - r$ fonksiyonunu göz önüne alalım. Bu fonksiyonu x_0, y_0, z_0 ve r_0 a göre diferensiyelleme, çözmek için zor olan karmaşık denklemleri ortaya koyar. Bu nedenle $f_2 = r_i^2 - r^2$ fonksiyonunu göz önüne alalım. $r_i + r$ ye $2r$ olarak yaklaşılabilirdiğinden, bu fonksiyon $f_2 = (r_i - r)(r_i + r) \approx 2r(r_i - r)$ olarak yazılabilir. Merkez ve yarıçap giriş tahminlerini elde etmek için, bu fonksiyonu x_0, y_0, z_0 ve r_0 a göre diferensiyelleyelim. Böylece bir küre için başlangıç tahminleri elde etmek için minimumlaştırma fonksiyonu $f = r_i^2 - r^2$ dir.

Giriş tahmini matematiği

$f = r_i^2 - r^2$ yi açarak, $\rho = (x_0^2 + y_0^2 + z_0^2) - r^2$ olmak üzere,

$$f = (x_i - x_0)^2 + (y_i - y_0)^2 + (z_i - z_0)^2 - r^2 = -(2x_i x_0 + 2y_i y_0 + 2z_i z_0) + \rho + (x_i^2 + y_i^2 + z_i^2)$$

yi elde ederiz. ρ değişkeni, denklemi lineer yapmak için ortaya konur.

Veri noktalarının n kümesi için yukarıdaki denklemlerin kümesi matris formunda sunulur:

$$\begin{pmatrix} -2x_1 & -2y_1 & -2z_1 & 1 \\ -2x_2 & -2y_2 & -2z_2 & 1 \\ \dots & \dots & \dots & \dots \\ -2x_n & -2y_n & -2z_n & 1 \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \\ z_0 \\ \rho \end{pmatrix} - \begin{pmatrix} x_1^2 + y_1^2 + z_1^2 \\ x_2^2 + y_2^2 + z_2^2 \\ \dots \\ x_n^2 + y_n^2 + z_n^2 \end{pmatrix} = \begin{pmatrix} f_1 \\ f_2 \\ \dots \\ f_n \end{pmatrix}. \quad (3.102)$$

Bir en küçük kare çözümü için, $f_i = 0$ dır. Giriş matris notasyonu

$$\mathbf{A} = \begin{pmatrix} -2x_1 & -2y_1 & -2z_1 & 1 \\ -2x_2 & -2y_2 & -2z_2 & 1 \\ \dots & \dots & \dots & \dots \\ -2x_n & -2y_n & -2z_n & 1 \end{pmatrix}, \quad \mathbf{P} = \begin{pmatrix} x_0 \\ y_0 \\ z_0 \\ \rho \end{pmatrix} \quad \text{ve} \quad \mathbf{B} = \begin{pmatrix} x_1^2 + y_1^2 + z_1^2 \\ x_2^2 + y_2^2 + z_2^2 \\ \dots \\ x_n^2 + y_n^2 + z_n^2 \end{pmatrix} \quad (3.103)$$

dir.

$\mathbf{AP} - \mathbf{B} = \mathbf{0}$ elde ederiz. \mathbf{P} yi elde etmek için bu denklemi en küçük kare anlamında çözelim. Bu; \mathbf{P} nin $\mathbf{A}^T \mathbf{AP} = \mathbf{A}^T \mathbf{B}$ denklemini sağlar. x_0, y_0, z_0 ve ρ için giriş tahminleri \mathbf{P} için yukarıdaki çözümden elde edilir. r yarıçapı için giriş tahmini $\rho = (x_0^2 + y_0^2 + z_0^2) - r^2$ ilişkisinden elde edilebilir.

➤ Gauss-Newton yöntemi :

Merkez ve r yarıçapı için giriş tahminleri elde ettikten sonra, merkez ve yarıçap için son değerlere ulaşmak için Gauss-Newton yöntemi kullanılır.

Minimumlaştırma fonksiyonu

$r_i = \sqrt{(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2}$ olmak üzere, minimumlaştırma fonksiyonu $d_i = r_i - r$ ile verilir.

i) Jakobiyen matrisini inşa etme : Jakobiyen matrisinin elemanları

$$\begin{pmatrix} \frac{\partial d_1}{\partial x_0} & \frac{\partial d_1}{\partial y_0} & \frac{\partial d_1}{\partial z_0} & \frac{\partial d_1}{\partial r_0} \\ \frac{\partial d_2}{\partial x_0} & \frac{\partial d_2}{\partial y_0} & \frac{\partial d_2}{\partial z_0} & \frac{\partial d_2}{\partial r_0} \\ \dots & \dots & \dots & \dots \\ \frac{\partial d_n}{\partial x_0} & \frac{\partial d_n}{\partial y_0} & \frac{\partial d_n}{\partial z_0} & \frac{\partial d_n}{\partial r_0} \end{pmatrix}$$

ile verilir. Jakobiyenin çeşitli bileşenlerini değerlendirerek ve matriste yerine koyarak,

$$J = \begin{pmatrix} \frac{-(x_1 - x_0)}{r_1} & \frac{-(y_1 - y_0)}{r_1} & \frac{-(z_1 - z_0)}{r_1} & -1 \\ \frac{-(x_2 - x_0)}{r_2} & \frac{-(y_2 - y_0)}{r_2} & \frac{-(z_2 - z_0)}{r_2} & -1 \\ \dots & \dots & \dots & \dots \\ \frac{-(x_n - x_0)}{r_n} & \frac{-(y_n - y_0)}{r_n} & \frac{-(z_n - z_0)}{r_n} & -1 \end{pmatrix}$$

elde ederiz.

ii)

$$\mathbf{P} = \begin{bmatrix} \rho_{x_0} \\ \rho_{y_0} \\ \rho_{z_0} \\ \rho_r \end{bmatrix}$$

olmak üzere, $\mathbf{JP} = -\mathbf{d}$ lineer en küçük kare sistemidir.

iii)

$$\begin{aligned} x_0 &:= x_0 + \rho_{x_0}; \\ y_0 &:= y_0 + \rho_{y_0}; \\ z_0 &:= z_0 + \rho_{z_0}; \\ r &:= r + \rho_r; \end{aligned}$$

ye göre artırma parametreleridir.

iv)

Yakınsama şartı : Algoritma yakınsayınca kadar adımları tekrarlanır. Yakınsama şartı $\mathbf{g} = J'\mathbf{d}$ nin minimum olmasıyla verilir.

3.2.3.6. En Küçük Kareler En İyi Uyum Silindirleri

$m \geq 5$ olmak üzere, m sayıda (x_i, y_i, z_i) noktalarına bir silindir uydurulduğunu kabul edelim.

➤ Silindirler için Gauss-Newton Stratejisi :

Bir doğru; doğru üzerinde verilen bir nokta ve (a, b, c) doğrultu kosinüsü ile belirlenebilir. Bu nedenle bir doğruyu belirlemek için 6 sayı gerekir. Onun kısıtlaması $a^2 + b^2 + c^2 = 1$ dir. Bu nedenle bileşenlerden ikisi verilir ve üçüncü belirlenebilir. $c=1$, yani, bir düşey doğru olmak üzere, bir doğru üzerindeki bir sınırlamayı göz önüne alalım. Bu sınırlama iki a ve b doğrultu kosinüsünü belirlemek için yeterlidir.

z ;

$$ax + by + cz = 0 \quad (3.104)$$

ilişkisinden belirlenebilir. $c=1$ olduğundan,

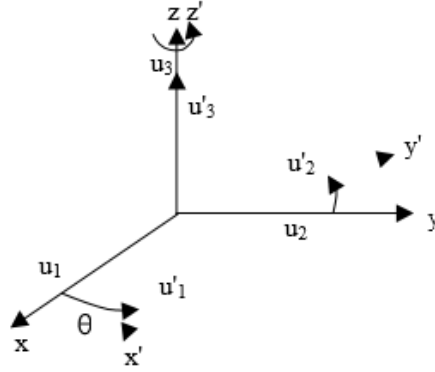
$$z = -ax - by \quad (3.105)$$

dir. Bu nedenle bu kısıtlamanın avantajı, 6 parametreyi 4 tane a, b, x_0 ve y_0 parametresine göre minimumlaştırmaktır. Bir Gauss-Newton yöntemi kullanıldığında, uzaklığın türevleri hesaplanabildiğinden, o; jakobiyen matrisini çözmenin ve onu değerlendirme süresinin karmaşıklığını azaltır.

Gauss-Newton algoritması bir silindirin durumu için değiştirilir. Bu değişiklik şudur:

- 1) Her bir iterasyonun başlangıcında koordinat sistemi çevrilir. Bu yüzden eksen üzerindeki nokta, koordinat sisteminin orijinidir.
- 2) Eksenin doğrultusu z - eksenini boyunca olacak şekilde koordinat sistemi döndürülür.

➤ **Dönüş Kavramı :**



Şekil 3.9. Dönüş matrislerini nasıl geliştirileceğinin örneği.

Yukarıdaki şekli göz önüne alalım. $\mathbf{B} = \{u_1, u_2, u_3\}$, yeni $\mathbf{B}' = \{u'_1, u'_2, u'_3\}$ tabanına döndürülür. Burada u_1, u_2, u_3 ve u'_1, u'_2, u'_3 birim vektörlerdir. u'_1, u'_2 ve u'_3

$$\begin{bmatrix} u'_1 \end{bmatrix}_{\mathbf{B}} = \begin{bmatrix} \cos(\theta) \\ \sin(\theta) \\ 0 \end{bmatrix} \quad \begin{bmatrix} u'_2 \end{bmatrix}_{\mathbf{B}} = \begin{bmatrix} -\sin(\theta) \\ \cos(\theta) \\ 0 \end{bmatrix} \quad \begin{bmatrix} u'_3 \end{bmatrix}_{\mathbf{B}} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

olarak ifade edilebilir. Bu nedenle, \mathbf{B}' den \mathbf{B} ye geçiş matrisi

$$\mathbf{P} = \begin{bmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

dir ve \mathbf{B} den \mathbf{B}' ye geçiş matrisi

$$\mathbf{P}' = \begin{bmatrix} \cos(\theta) & \sin(\theta) & 0 \\ -\sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

dir. Bu nedenle yeni (x', y', z') koordinat sistemi,

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} \cos(\theta) & \sin(\theta) & 0 \\ -\sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

vasıtasıyla eski (x, y, z) koordinatlarından hesaplanabilir.

➤ **Parametreleme :**

Bir silindir

- 1) onun ekseni üzerindeki bir (x_0, y_0, z_0) noktasıyla,
- 2) bir uzun ekseni işaret eden bir (a, b, c) vektörü ve
- 3) onun r yarıçapıyla

belirlenebilir.

3.3. Son Dönemlerde Kuadratik Eğrilerin ve Yüzeylerin En Küçük Kareler Uyumu Üzerinde Yapılan Bazı Önemli Çalışmalar

3.3.1. Çemberlerin ve Elipslerin En Küçük Kareler Uyumu

Gander, Golub ve Strebel, 1994 yılında çemberlerin ve elipslerin en küçük kareler uyumu üzerine çalışma yaptı. Verilen noktalara uzaklıkların karelerinin toplamının minimal olduğu elipslere ve çemberlere “en iyi uyum” veya “geometrik uyum” olarak başvurdular ve algoritmaları “geometrik” dediler. En küçük kareler anlamında $F(\mathbf{x})=0$ cebirsel denkleminin parametrelerini belirlemeyi “cebrik uyum” ile gösterdiler ve algoritmaları “cebirsal” dediler. Lineer olmayan en küçük kareler problemini çözmek için ünlü Gauss-Newton yöntemini kullandılar.

İlk olarak, $a \neq 0$ ve $\mathbf{x}, \mathbf{b} \in R^2$ olmak üzere, düzlemde çemberin bir

$$F(\mathbf{x}) = a\mathbf{x}'\mathbf{x} + b'\mathbf{x} + c = 0 \quad (3.106)$$

cebirsal gösterimini ele aldılar. Çemberi uydurmak için verilen veri noktalarından a , \mathbf{b} ve c katsayılarını hesapladılar. Bunun için $a \neq 0$ olmak üzere (3.106) denklemini

$$\left(x_1 + \frac{b_1}{2a}\right) + \left(x_2 + \frac{b_2}{2a}\right) = \frac{\|\mathbf{b}\|^2}{4a^2} - \frac{c}{a} \quad (3.107)$$

denkleminde dönüştürdüler. (3.107) denkleminin sağ yanının pozitif olması şartıyla

$\mathbf{z} = (z_1, z_2) = \left(-\frac{b_1}{2a}, -\frac{b_2}{2a}\right)$ merkezli, $r = \sqrt{\frac{\|\mathbf{b}\|^2}{4a^2} - \frac{c}{a}}$ yarıçaplı çember elde ettiler ve

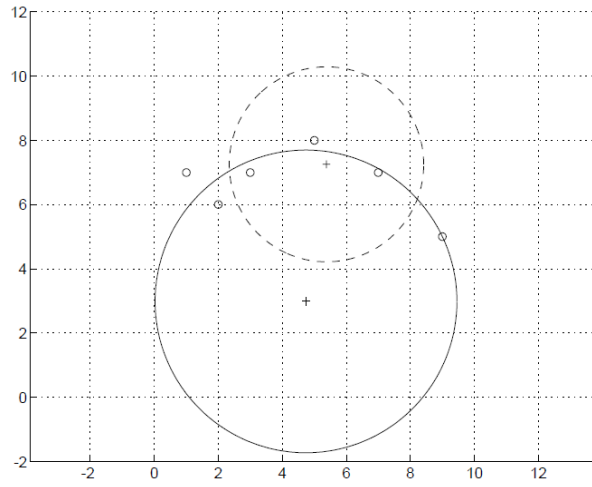
buna “cebrik uyum” adını verdiler.

$d_i^2 = (\|\mathbf{z} - \mathbf{x}_i\| - r)^2$ uzaklıklar kareler toplamını minimumlaştırmak için bir lineer olmayan en küçük kareler problemini çözmek için Gauss-Newton yöntemini kullandılar. Uydurulan çembere “en iyi uyum (geometrik uyum)” adını verdiler. Daha iyi anlaşılması açısından aşağıdaki 6 noktalı veri noktaları kümesi örneği üzerinde uygulama yaptılar.

Çizelge 3.6. 6 noktalı veri noktaları kümesi örneği.

x	1	2	5	7	9	3
y	7	6	8	7	5	7

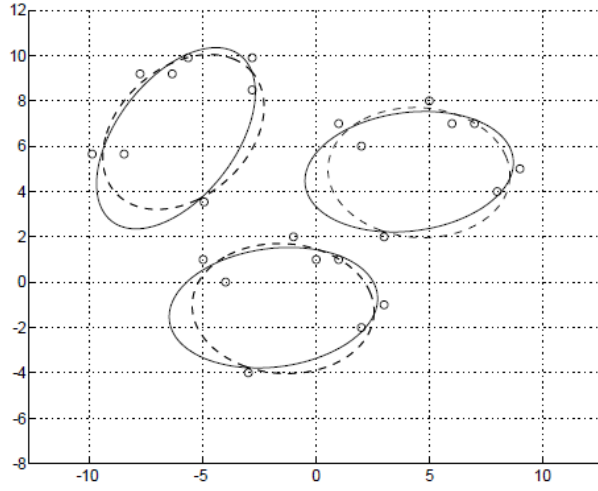
Cebrik uzaklığı minimumlaştırarak, $r=3.0370$ yarıçaplı ve $\mathbf{z}=(5.3794,7.2532)$ merkezli kesik çizgili çemberi elde ettiler. Cebrik çözümden elde edilen değerlerle iterasyonu başlatarak 11 Gauss-Newton adımından sonra düzeltme vektörünün normu $2.05E-6$ buldular. $\mathbf{z}=(4.7398, 2.9835)$ merkezli ve $r=4.7142$ yarıçaplı “en iyi uyum çemberi” ni elde ettiler.



Şekil 3.10. En İyi Uyuma Karşı Cebrik Uyum.

————— En iyi uyum
- - - - - Cebrik uyum

Elipsler için “cebrik uyum” u bulmak için Bookstein kısıtlaması ile en küçük kareler problemini çözdüler.



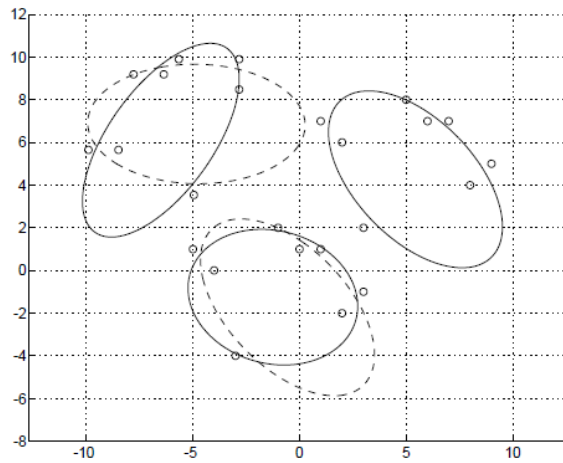
Şekil 3.11. Öklid-Değişimsiz Algoritmalar.

————— $\lambda_1^2 + \lambda_2^2 = 1$ kısıtlaması
 - - - - - $\lambda_1 + \lambda_2 = 1$ kısıtlaması

Çizelge 3.7. 8 noktalı veri noktaları kümesi örneği.

x	1	2	5	7	9	6	3	8
y	7	6	8	7	5	7	2	4

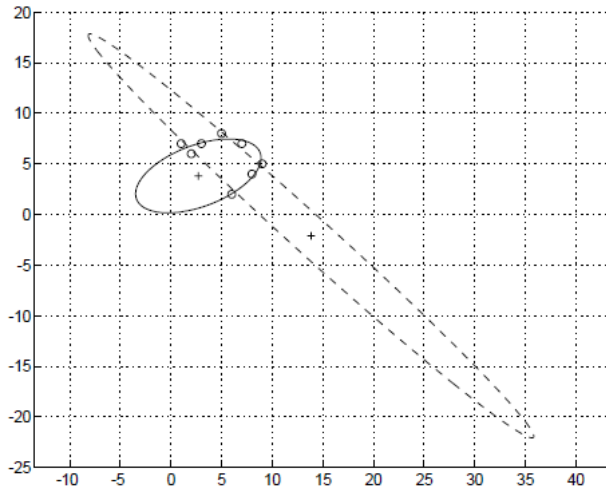
Farklı koordinat sistemlerinin etkisini göstermek için yukarıdaki 8 noktalı veri kümesi için, ilk olarak (-6,-6) kadar ve sonra da (-4,4) kadar kaydırılan ve $\pi/4$ kadar döndürülen, elips uyumunu hesapladılar (Elipsler, hiperboller ve paraboller için $\lambda_1^2 + \lambda_2^2 \neq 0$ olduğundan, Bookstein kısıtlaması bunlardan herhangi birini uydurmak için uygundur.).



Şekil 3.12. Değişimsiz-olmayan cebrik algoritma.

_____ Uygun elipsler
 - - - - - Dönüşümden sonra özgül bir biçimde uydurulan elipsler

$\|\mathbf{u}\|=1$ ile cebrik uzaklığı minimumlaştırarak, $\mathbf{z}=(13.8251, -2.1099)$, $a=29.6437$, $b=1.8806$ ve $r=1.80$ hata normlu kesik çizgili büyük puro şeklindeki elipsi elde ettiler. Uzaklıkların kareleri toplamını minimumlaştırarak, $\mathbf{z}=(2.6996, 3.8160)$, $a=6.5187$, $b=3.0319$ ve $r=1.17$ li koyu çizgili elipsi elde ettiler. Lineer olmayan en küçük kareler problemi için başlangıç değerlerini elde etmek için en iyi çember uyumu ile elde edilen merkezi kullandılar. Jakobiyen $b=a$ için tekil olduğundan $b_0 = a_0 = r$ yaklaşımı kullanılmadığından, bir başlangıç değeri olarak $b_0 = r/2$ yi kullandılar. Şekil 3.13 de gösterilen “en iyi elipsi” hesaplamak için $a_0 = 0$ ile 71 iterasyon adımına ihtiyaç duydular.



Şekil 3.13. En İyi Uyuma Karşı Cebrik Uyum.

_____ En iyi uyum
 - - - - - Cebrik uyum ($\|\mathbf{u}\|=1$)

Bilinen birkaç lineer olmayan en küçük kareler algoritmasını yerine getirdiler ve bunların karşılaştırmasını yaptılar.

1. Gauss-Newton (gauss)
2. Newton (newton)
3. Marquardt değişikliği ile Gauss-Newton (marq)
4. Değişken izdüşümü (varpro)
5. Ortogonal uzaklık regresyonu (odr)

Algoritmalar aşağıdaki problemler için her biri 8 noktadan ibaret olan örnek veri ile test edilmiştir.

1. Çizelge 3.7 deki noktaların özel kümesi
2. Bir kare içinde düzgün dağılan veri
3. Bir çember üzerindeki noktalar
4. $a/b=2$ ile bir elips üzerindeki noktalar
5. Hiperbol kolu üzerindeki noktalar

Çizelge 3.8, ilgili algoritma ve problem için başarısızlıkların sayısını (1000 e göre) gösterir; en küçük sayı üstünde durulur. Eğer algoritma 100 adımdan sonra sona erdirmeyi ise, yakınsama olmadığı kabul edilir ve bunun yerine bir “ \diamond ” gösterilir.

Çizelge 3.8. Cebirsel çemberin başlangıç parametreleri ile geometrik uyumu.

	gauss	newton	marq	varpro	odr
Özel	146	<u>85</u>	468	1146	\diamond
Rasgele	\diamond	\diamond	\diamond	<u>2427</u>	\diamond
Çember	22	22	22	36	<u>7</u>
Çember +	86	<u>67</u>	189	717	69
Elips	<u>30</u>	37	67	143	41
Elips +	186	\diamond	633	1977	<u>103</u>
Hiperbol	22	22	22	36	<u>10</u>
Hiperbol	\diamond	\diamond	\diamond	\diamond	\diamond

Başarısızlık /1000 , minimum belirtilir.

' \diamond ' Eğer yakınsama yoksa

Çizelge 3.9. Cebirsel elipsin başlangıç parametreleri ile geometrik uyumu.

	gauss	newton	marq	varpro	odr
Özel	<u>165</u>	896	566	1506	\diamond
Rasgele	\diamond	\diamond	\diamond	<u>2819</u>	\diamond
Çember	32	32	32	102	<u>7</u>
Çember +	76	<u>63</u>	145	574	66
Elips	22	22	22	112	<u>7</u>
Elips +	161	<u>40</u>	435	1870	74
Hiperbol	\diamond	\diamond	\diamond	<u>1747</u>	\diamond
Hiperbol +	\diamond	\diamond	\diamond	<u>2986</u>	\diamond

Başarısızlık /1000 , minimum belirtilir.

' \diamond ' Eğer yakınsama yoksa

Eğer başlangıç parametreleri Bookstein algoritmasından elde edildi ise, Çizelge 3.10 sonuçları gösterir. Çizelge 3.10, kesin (tam) konikler için daha doğru başlangıç verisi ile tüm algoritmaların daha hızlı yakınsadıklarını gösterir. Kaygılı elips verisi için, çözüme yakın başlangıç değerlerinden yararlanan newton algoritması her şeyden öncedir. Ayrıca başlangıç yaklaşımı olarak hizmet eden cebirsel yaklaşım geometrik çözümden tamamen farklı olduğundan, newton algoritmasının özel veri için doğru çözümü bulmadığına dikkat edelim.

Küçük test serilerinden buldukları genel sonuçlar şunlardır:

- Tüm algoritmalar basit cebirsel çözüme kıyasla (faktör 10-100) yasak edilecek derecede masraflıdır.
- Eğer problem iyi konumlanmış ve sonucun doğruluğu yüksek olmalı ise, parametrelenmiş algoritmaya uygulanan newton yöntemi çok daha etkindir.
- Uyumu optimize eden basit bir genel amaç olsa bile, odr algoritması genellikle elips uydurma problemi için yazılan algoritmalarla yarışma mahiyetindedir. Ayrıca yüksek derecede optimize eden bir odr yönteminin kullanılmadığı göz önünde bulundurulursa, çözümün yöntemi şaşırtıcı derecede basit ve etkindir.
- varpro algoritması en pahalı olarak görünür. Onun etkinliğinin olmamasının nedenleri, çoğu parametrelerin lineer olmamaları ve algoritmanın bu problem için özel matris yapısını kullanmamasıdır.

3.3.2. Kuadratik Eğrilerin ve Yüzeylerin En Küçük Kareler Uyumu

- Yakın geçmişte Sturm ve Gargallo çeşitli biçimlerde GGS (Gander, Golub ve Strebel) nin yöntemini değiştirdi. Özellikle onlar, 5 parametrelili aynı küme ile, tüm elips tipli (elipsler, hiperboller, parabol) konikleri tanımlamaya imkan veren bir projektif matrisi kullandılar.
- 1990 ların sonunda, bilgisayar görüntüsü uygulamalarında, cebirsel uzaklıkların minimumlaştırıldığını dile getiren çeşitli uyum planları geliştirildi. Chernov ve Ma onları yeniden gözden geçirdi.
- 2000 lerin başında, geometrik uzaklıklarının minimumlaştırılması ile ilgili Ahn ve ortaklarına ait başka bir yaklaşım çıktı. Chernov ve Ma onun bir genel biçimini ifade ettiler. GGS'nin 8 noktalı veri örneğini kullanarak iki elips uydurma problemini karşılaştırdılar.

Çizelge 3.10. İki elips uydurma yönteminin karşılaştırması.

	Başarısızlık Oranı	Ortalama İterasyon	İterasyon Başına Maliyet (Başarısızlıklar)
Kapalı	% 11	20	1640
GGS	% 26	60	1710

Çizelge 3.11. İki elips uydurma yönteminin karşılaştırması (G: Geometrik parametreler, K: Kepler parametreleri).

Başlangıç Elipsleri			
	En İyi Çember	Doğrudan Uyum	Taubin Uyumu
Kapalı (G)	16	17	17
Kapalı (K)	14	16	16
GGs	50	54	54

3.3.2.1. Konikler Üzerine İzdüşüm

Elipsler üzerine izdüşürmeye dikkat çekici bir yaklaşım, 2004 de D. Eberly tarafından bulundu. Chernov ve Ma, Elberly'nin yöntemini diğer kuadratik eğrilere (elipsler, hiperboller, parabol) ve yüzeylere (elipsoidler, hiperbolik parabolidler, hiperbolidler, vs.) uyarladılar. Her bir durumda yakınsaklığın bir teorik ispatını ortaya koydular. Şimdi bunu inceleyelim.

3.1. Yardımcı Teorem. (x, y) verilen bir nokta olsun ve (x', y') onun $P(x, y; \Theta) = 0$ (bu takdirde x', y', Θ ye bağlıdır) eğrisi üzerine izdüşümünü gösterebilir. $g = x - x', h = y - y'$ ve $d^2 = g^2 + h^2$ koyalım. Bu durumda,

$$g_{\Theta} = \frac{P_{\Theta} P_x^2 - g P_y (P_x P_{y\Theta} - P_y P_{x\Theta})}{P_x (P_x^2 + P_y^2)}, \quad (3.108)$$

$$h_{\Theta} = \frac{P_{\Theta} P_y^2 + h P_x (P_x P_{y\Theta} - P_y P_{x\Theta})}{P_y (P_x^2 + P_y^2)} \quad (3.109)$$

ve

$$d_{\Theta} = \frac{P_{\Theta}}{\sqrt{P_x^2 + P_y^2}} \quad (3.110)$$

elde ederiz. Burada P_{Θ}, P_x, P_y sırasıyla P nin Θ, x, y ye göre birinci mertebeden kısmi türevlerini ve $P_{x\Theta}$ ve $P_{y\Theta}$ karşılık gelen ikinci mertebeden kısmi türevlerini gösterir; tüm türevler (x', y') iz düşüm noktasında alınır.

İspat. $(x - x', y - y')$ eğriye ortogondur. Herhangi bir t skaleri için

$$g = x - x' = tP_x \quad \text{ve} \quad h = y - y' = tP_y \quad (3.111)$$

dir. Bu doğrudan doğruya

$$d^2 = g^2 + h^2 = t^2(P_x^2 + P_y^2) \quad (3.112)$$

olduğunu ortaya koyar.

Bunun ardından, Θ ya göre diferensiyeli kullanırız. $P(x', y'; \Theta) = 0$ özdeşliğini diferensiyelleme

$$P_\Theta = -P_x x'_\Theta - P_y y'_\Theta = (gg_\Theta + hh_\Theta) / t \quad (3.113)$$

bağıntısını verir ve $d^2 = g^2 + h^2$ özdeşliğini diferensiyelleme

$$dd_\Theta = gg_\Theta + hh_\Theta = tP_\Theta \quad (3.114)$$

bağıntısını verir. Şimdi (3.110), (3.114) ve (3.112) den görülür. (3.111) in diferensiyellenmesi

$$g_\Theta = t_\Theta P_x + tP_{x\Theta}, \quad h_\Theta = t_\Theta P_y + tP_{y\Theta}$$

bağıntılarını verir. Bu iki denklemden t_Θ yı yok etme

$$g_\Theta P_y - h_\Theta P_x = -t(P_x P_{y\Theta} - P_y P_{x\Theta}) \quad (3.115)$$

bağıntısını verir. (3.114) ve (3.115) i g_Θ ve h_Θ için çözerek (3.108) i ve (3.109) u elde ederiz.

a) Elipsler :

Bir (u, v) noktasını onun kanonik koordinatları cinsinden bir

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - 1 = 0 \quad (3.116)$$

elipsi üzerine izdüşürmek yeterli olacaktır.

Gerçekten, diğer elipsler (3.116) kanonik formuna çevrilebilir ve döndürülebilir ve bundan sonra izdüşüm noktası orijinal elipse geri çevrilebilir ve döndürülebilir.

Simetri nedeniyle bunu $u > 0, v > 0$ çemberin ilk dörtte birinde çalıştırmak yeterli olacaktır, bu takdirde (x, y) izdüşüm noktası da birinci çeyrekte olacaktır, yani $x > 0, y > 0$ olacaktır (Diğer noktalar eksenler etrafında birinci çeyreğe yansıtılabilir

ve bu takdirde izdüşüm noktası geriye yansıtılabilir). Aynı zamanda $u=0$ veya $v=0$ olduğu dejenere durumu hariç tutarız, onlar oldukça basittir ve ayrı ayrı ele alınabilir.

Şimdi elips üzerindeki (x, y) izdüşüm noktası (3.111) ortogonallik şartını sağlar, bu nedenle, herhangi bir reel t için (elips içindeki noktalar için $t < 0$ ve elips dışındaki noktalar için $t > 0$ olduğuna dikkat edelim)

$$u - x = tx/a^2 \text{ ve } v - y = ty/b^2 \quad (3.117)$$

dir. (3.117) den,

$$x = \frac{a^2 u}{t + a^2} \text{ ve } y = \frac{b^2 v}{t + b^2} \quad (3.118)$$

buluruz. $x > 0, y > 0$ olduğundan, $t > -a^2$ ve $t > -b^2$ kısıtlamasına sahibiz. Alışıla geldiği üzere, $a \geq b$ kabul ederek, bir tek $t > -b^2$ kısıtlaması elde ederiz. (3.118) i (3.116) da yerine koyarak, bulmamız gereken köke sahip olan (x, y) , elips üzerinde olmak zorunda olduğundan bir

$$F(t) = \frac{a^2 u^2}{(t + a^2)^2} + \frac{b^2 v^2}{(t + b^2)^2} - 1 \quad (3.119)$$

fonksiyonunu elde ederiz. (3.119) denklemini t için çözdüğümüzde, (3.118) bağıntısı vasıtasıyla (x, y) izdüşüm noktasını hesaplayabiliriz.

$$\lim_{t \rightarrow -b^2+} F(t) = +\infty \text{ ve } \lim_{t \rightarrow \infty} F(t) = -1$$

dir. F nin türevlerini alarak,

$$F'(t) = -\frac{2a^2 u^2}{(t + a^2)^3} - \frac{2b^2 v^2}{(t + b^2)^3} \quad (3.120)$$

ve

$$F''(t) = \frac{6a^2 u^2}{(t + a^2)^4} + \frac{6b^2 v^2}{(t + b^2)^4} \quad (3.121)$$

olduğunu görürüz.

Böylece, $(-b^2, \infty)$ aralığı üzerinde $F' < 0$ ve $F'' < 0$ dır, yani F monoton azalan ve konkavdır; bkz Şekil 3.14. Bu nedenle, $F(t_0) > 0$ olmak üzere, herhangi bir t_0 noktasında başlayan Standart Newton Yöntemi F nin bir tek köküne

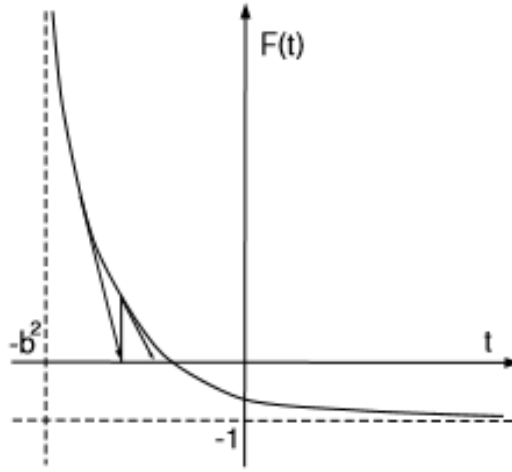
yakınsayacaktır. $F(t_0) > 0$ (3.119) ile garanti edildiğinden, Eberly $t_0 = bv - b^2$ ile başlatmayı önerir.

Chernov ve Ma,

$$t_0 = \max \{ay - a^2, bv - b^2\} \quad (3.122)$$

ile başlatmayı daha yararlı buldu. Bu takdirde, tüm uygulama durumlarında Newton yöntemi 3-5 iterasyonda yakınsar ve 7-8 anlamlı basamak ile kökü bulur.

Bu, dördüncü dereceden denklemi çözmekten veya genel sezgiselleri kullanmaktan ortalama olarak 2-3 defa daha hızlıdır (Ahn ve ark., 2001; Ahn, 2004).



Şekil 3.14. $t > -b^2$ için $F(t)$ nin Bir Tipik Grafiği ve Newton İterasyonlarının Köke Doğru İlerlemesi.

b) Hiperboller :

Şimdi bir hiperbol üzerine bir (u, v) noktasını izdüşürelim. Yine, hiperbol onun kanonik koordinatları ile tanımlanabilir, yani

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} - 1 = 0 \quad (3.123)$$

dır. Simetri nedeniyle, yöntemi $u > 0, v > 0$ ile sınırlar, sonra da $x > 0, y > 0$ elde ederiz. Şimdi (3.111) ortogonalite şartları

$$u - x = tx/a^2 \text{ ve } v - y = -ty/b^2 \quad (3.124)$$

olduğunu gösterir ki buradan

$$x = \frac{a^2 u}{t + a^2} \quad \text{ve} \quad y = \frac{b^2 v}{-t + b^2} \quad (3.125)$$

olduğu görülür. $x > 0, y > 0$ olduğundan, $-a^2 < t < b^2$ kısıtlamalarını elde ederiz. (3.125) i (3.123) de yerine koyarak, bulmamız gereken köke sahip olan bir

$$F(t) = \frac{a^2 u^2}{(t + a^2)^2} - \frac{b^2 v^2}{(-t + b^2)^2} - 1 \quad (3.126)$$

fonksiyonunu elde ederiz.

$$\lim_{t \rightarrow -a^2+} F(t) = +\infty \quad \text{ve} \quad \lim_{t \rightarrow b^2-} F(t) = -\infty$$

olduğuna dikkat edelim. F nin türevlerini alarak

$$F'(t) = -\frac{2a^2 u^2}{(t + a^2)^3} - \frac{2b^2 v^2}{(-t + b^2)^3} \quad (3.127)$$

ve buradan her $t \in (-a^2, b^2)$ için $F' < 0$ olduğunu görürüz. Daha sonra,

$$F''(t) = \frac{6a^2 u^2}{(t + a^2)^4} - \frac{6b^2 v^2}{(-t + b^2)^4} \quad (3.128)$$

elde edilir.

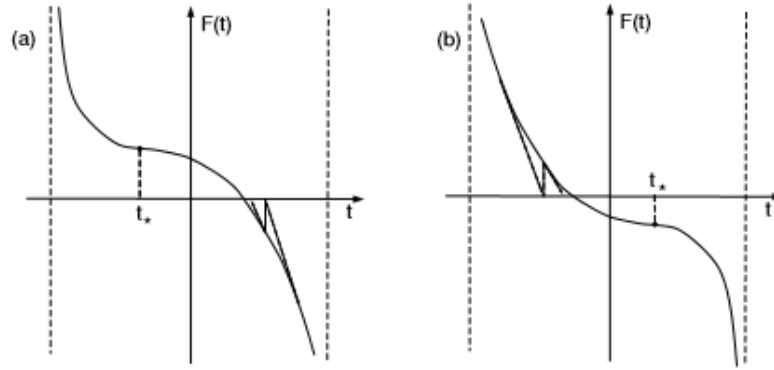
Şimdi F'' , $+\infty$ dan (yaklaşık olarak $-a^2$ den), $-\infty$ a (yaklaşık b^2 ye) azalır ve monotondur (Çünkü, kolayca doğrulanabileceği gibi, $F'' < 0$ dir.). Bu nedenle F , $(-a^2, b^2)$ aralığı içinde bir tek dönüm noktasına, yani t_* noktasına sahiptir (bkz. Şekil 3.15). Şekil 3.15 de iki mümkün durum gösterilir : (a) Dönüm noktası x eksenini yukarısında bulunur, yani $F(t_*) > 0$ dir ve (b) dönüm noktası x ekseninin aşağısında bulunur. Dönüm noktası $F'' = 0$ denklemini çizerek bulunur, bu nedenle

$$t_* = \frac{b^2 \sqrt{au} - a^2 \sqrt{bv}}{\sqrt{au} + \sqrt{bv}}$$

dir. Şimdi $F(t_*)$ hesaplayarak, (a) veya (b) durumunu belirleyebilir ve el altına alırız. Standart Newton yöntemi $F(t) = 0$ in köküne yakınsayacaktır, fakat t_0 başlama noktası akıllıca seçilmelidir. (a) durumunda, $F(t_0) < 0$ ve (b) durumunda $F(t_0) > 0$ olacak şekilde t_0 seçmemiz gerekir. (a) durumunda F nin bir negatif olduğu değeri buluncaya kadar, $k = 1, 2, \dots$ için $t_k = b^2 - (b^2 - t_*)/2^k$ noktalarını deneyebiliriz. (b) durumunda F nin bir pozitif olduğu değeri buluncaya kadar,

$k = 1, 2, \dots$ için, $t_k = -a^2 + (t_* + a^2)/2^k$ noktalarını deneyebiliriz. t_0 seçiminin bu yöntemi nispeten masrafsızdır ve F nin türevlerini içermez. Bu; hiperboller için izdüşümümüzü tamamlar.

Uygulamada bu, t_0 in daha dikkatle hazırlanmış bir seçimi nedeniyle, elipsler için izdüşüm algoritmasından hemen hemen 1.5 kat daha uzun çalışır, bu nedenle bunun hızı sezgisel olanına yakındır (Ahn ve ark., 2001; Ahn, 2004). Ancak, analizimize göre o, yakınsamayı garanti eder.



Şekil 3.15. $F(t)$ nin $-a^2 < t < b^2$ Aralığı Üzerindeki İki Mümkün Görünümü (Oklar Newton İterasyonlarının Köke Doğru İlerlemelerini Gösterir).

c) Paraboller :

Şimdi bir (u, v) noktasını bir parabol üzerine izdüşürelim. Yine, parabol onun kanonik koordinatları cinsinden tanımlanabilir. Yani, $p > 0$ odaktan doğrultmana uzaklık olmak üzere

$$y^2 - 2px = 0 \quad (3.129)$$

denklemleriyle tanımlanır.

Simetri nedeniyle yöntemi $v > 0$ a kısıtlarız, bu takdirde yine $y > 0$ elde ederiz. Şimdi (3.111) ortogonallik şartları

$$u - x = -pt \quad \text{ve} \quad v - y = yt \quad (3.130)$$

olduğunu ortaya koyar ki buradan

$$x = u + pt \quad \text{ve} \quad y = \frac{v}{t+1} \quad (3.131)$$

dir.

$y > 0$ olduğundan $t > -1$ kısıtlamasına sahip oluruz. (3.131) i (3.129) da yerine koyarak, bulmamız gereken köke sahip olan bir

$$F(t) = \frac{v^2}{(t+1)^2} - 2pu - 2p^2t \quad (3.132)$$

fonksiyonunu elde ederiz.

$$\lim_{t \rightarrow -1^+} F(t) = +\infty \quad \text{ve} \quad \lim_{t \rightarrow \infty} F(t) = -\infty$$

olduğuna dikkat edelim.

F nin türevlerini alarak, $F'(t) = -\frac{2v^2}{(t+1)^3}$ ve $F''(t) = \frac{6v^2}{(t+1)^4}$ olduğunu görürüz.

Bu nedenle, $(-1, \infty)$ aralığı üzerinde $F' < 0$ ve $F'' > 0$ elde ederiz, yani F monoton azalandır ve konkavdır. Şimdi $F(t_0) > 0$ olmak üzere, herhangi bir t_0 ile başlayan Standart Newton yöntemi F nin bir tek köküne yakınsayacaktır. F nin pozitif olduğu bir değeri buluncaya kadar, $k=1,2,\dots$ için $t_k = -1 + 2^{-k}$ noktalarını deneyebiliriz.

3.3.2.2. Kuadratikler Üzerine İzdüşüm

Burada, özel bir (u, v, w) noktasının çeşitli tiplerden kuadratik yüzeyler üzerine izdüşümünü belirleriz.

a) Elipsoidler :

$a \geq b \geq c > 0$ elipsoidin yan eksenleri olmak üzere, kanonik koordinatları ile tanımlanan bir elipsoid aşağıdaki gibidir.

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} - 1 = 0 \quad (3.133)$$

(Diğer elipsoidler durumu (3.133) kanonik formuna çevrilebilir ve döndürülebilir ve bu takdirde izdüşüm noktası orijinal elipsoide geri çevrilebilir ve döndürülebilir). Simetri nedeniyle, $u > 0, v > 0, w > 0$ a kısıtlar, bundan sonra yine $x > 0, y > 0, z > 0$ elde ederiz.

Şimdi ortogonallik şartı, herhangi bir t skaları için

$$u - x = tx/a^2, \quad v - y = ty/b^2, \quad w - z = tz/c^2 \quad (3.134)$$

olduğunu ortaya koyar ki, buradan

$$x = \frac{a^2 u}{t+a^2}, \quad y = \frac{b^2 v}{t+b^2}, \quad z = \frac{c^2 w}{t+c^2} \quad (3.135)$$

elde edilir. $x, y, z > 0$ olduğundan, $t > \max\{-a^2, -b^2, -c^2\} = -c^2$ kısıtlamasını elde ederiz. (3.135) i (3.133) de yerine koyarak, bulmamız gereken köke sahip olan bir

$$F(t) = \frac{a^2 u^2}{(t+a^2)^2} + \frac{b^2 v^2}{(t+b^2)^2} + \frac{c^2 w^2}{(t+c^2)^2} - 1 \quad (3.136)$$

fonksiyonunu elde ederiz.

$$\lim_{t \rightarrow -c^2+} F(t) = +\infty \quad \text{ve} \quad \lim_{t \rightarrow \infty} F(t) = -1$$

olduğuna dikkat edelim.

F nin türevlerini alarak,

$$F'(t) = -\frac{2a^2 u^2}{(t+a^2)^3} - \frac{2b^2 v^2}{(t+b^2)^3} - \frac{2c^2 w^2}{(t+c^2)^3} \quad (3.137)$$

ve

$$F''(t) = \frac{6a^2 u^2}{(t+a^2)^4} + \frac{6b^2 v^2}{(t+b^2)^4} + \frac{6c^2 w^2}{(t+c^2)^4} \quad (3.138)$$

olduğunu görürüz.

Bu nedenle, $(-c^2, \infty)$ aralığı üzerinde $F' < 0$ ve $F'' > 0$ elde ederiz, yani, yukarıda Şekil 3.14 deki gibi, F monoton azalan ve konkavdır. Bu nedenle, $F(t_0) > 0$ olmak üzere herhangi bir t_0 da başlayan standart Newton yöntemi F nin bir tek köküne yakınsayacaktır ve

$$t_0 = \max\{au - a^2, bv - b^2, cw - c^2\}$$

seçeriz (Kanatani, 2008).

b) Hiperbolik Parabolidler :

Şimdi bir (u, v, w) noktasını onun kanonik koordinatları cinsinden

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} - z = 0 \quad (3.139)$$

olarak tanımlanan hiperbolik paraboloid (semer) üzerine izdüşürelim.

Simetri nedeniyle, yöntemi $u > 0$, $v > 0$ ile sınırlarız, bu durumda da $x > 0$, $y > 0$ elde ederiz. Şimdi ortogonalite şartları, herhangi bir t skaları için,

$$u - x = tx/a^2, \quad v - y = -ty/b^2, \quad w - z = -t/2 \quad (3.140)$$

olduğunu ve buradan,

$$x = \frac{a^2 u}{t + a^2}, \quad y = \frac{b^2 v}{-t + b^2}, \quad z = w + \frac{t}{2} \quad (3.141)$$

olacağını ortaya koyar. $x, y > 0$ olduğundan, $-a^2 < t < b^2$ kısıtlamalarına sahip oluruz. (3.141) i (3.139) da yerine koyarak, bulmamız gereken köke sahip olan bir

$$F(t) = \frac{a^2 u^2}{(t + a^2)^2} - \frac{b^2 v^2}{(-t + b^2)^2} - w - \frac{t}{2} \quad (3.142)$$

fonksiyonunu elde ederiz.

$$\lim_{t \rightarrow -a^2+} F(t) = +\infty \quad \text{ve} \quad \lim_{t \rightarrow b^2-} F(t) = -\infty$$

olduğuna dikkat edelim. F nin türevlerini alarak

$$F'(t) = -\frac{2a^2 u^2}{(t + a^2)^3} - \frac{2b^2 v^2}{(-t + b^2)^3} - \frac{1}{2} \quad (3.143)$$

elde edilir. Buradan da her $t \in (-a^2, b^2)$ için $F' < 0$ olduğunu görürüz. Bundan sonraki türevi

$$F''(t) = \frac{6a^2 u^2}{(t + a^2)^4} - \frac{6b^2 v^2}{(-t + b^2)^4} \quad (3.144)$$

dür.

Şimdi F'' , $+\infty$ dan $(-a^2)$ yakınından) $-\infty$ a (b^2) yakınına) azalır ve monotondur (çünkü, kolayca doğrulanacağı gibi, $F''' < 0$ dır). Bu nedenle F , $(-a^2, b^2)$ aralığı içinde bir tek t_* dönüm noktasına sahiptir. İleriki analizimiz önceki kısımdaki hiperboller için yapılanı tekrarlardır.

c) Hiperboloidler :

Şimdi bir (u, v, w) noktasını $a \geq b$ olmak üzere, onun kanonik koordinatları cinsinden

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} - 1 = 0 \quad (3.145)$$

olarak tanımlanan bir hiperboloid (bir çarşaf) üzerine izdüşürelim. Simetriden dolayı yöntemi $u > 0, v > 0, w > 0$ a kısıtlarız, bundan sonra da $x > 0, y > 0, z > 0$ elde ederiz. Ortogonallik şartları şimdi herhangi bir t skaları için

$$u - x = tx/a^2, \quad v - y = ty/b^2, \quad w - z = -tz/c^2 \quad (3.146)$$

olduğunu ortaya koyar ki buradan,

$$x = \frac{a^2 u}{t + a^2}, \quad y = \frac{b^2 v}{t + b^2}, \quad z = \frac{c^2 w}{-t + c^2} \quad (3.147)$$

dir. $x, y, z > 0$ olduğundan, $-b^2 < t < c^2$ sınırlarına sahip oluruz. (3.147) yi (3.145) de yerine koyarak, bulmamız gereken köke sahip olan bir

$$F(t) = \frac{a^2 u^2}{(t + a^2)^2} + \frac{b^2 v^2}{(t + b^2)^2} - \frac{c^2 w^2}{(-t + c^2)^2} - 1 \quad (3.148)$$

fonksiyonunu elde ederiz.

$$\lim_{t \rightarrow -b^2+} F(t) = +\infty \quad \text{ve} \quad \lim_{t \rightarrow c^2-} F(t) = -\infty$$

olduğuna dikkat edelim. F nin türevlerini alarak,

$$F'(t) = -\frac{2a^2 u^2}{(t + a^2)^3} - \frac{2b^2 v^2}{(t + b^2)^3} - \frac{2c^2 w^2}{(-t + c^2)^3} \quad (3.149)$$

olduğunu, buradan da her $t \in (-b^2, c^2)$ için $F' < 0$ olduğunu görürüz. Bundan sonraki türev

$$F''(t) = \frac{6a^2 u^2}{(t + a^2)^4} + \frac{6b^2 v^2}{(t + b^2)^4} - \frac{6c^2 w^2}{(-t + c^2)^4} \quad (3.150)$$

dir.

Önceden olduğu gibi, yine F'' , $+\infty$ dan ($-b^2$ yakınından) $-\infty$ a (c^2 yakınına) azalır ve monotondur (çünkü, kolayca doğrulanacağı gibi, $F''' < 0$ dır). Bu nedenle, F'' , $(-b^2, c^2)$ içinde bir tek t_* dönüm noktasına sahiptir. Onun grafiği Şekil 3.15 de gösterildiği gibidir.

Ancak şimdi, Şekil 3.15 nin (a) veya (b) kısmında elde ettiğimiz durumu belirlemek kolay değildir. Bununla beraber, hiperboller durumunda tanımlanan iki iteratif yöntemden biri (yani, biri soldan diğeri sağdan çalışan Newton yöntemi) çalışmalıdır.

Bu nedenle, sadece bu iki yöntemden biri rasgele seçilebilir ve bunun yakınsama umudu sonra gelir. Eğer bu başarısız ise, yani, eğer bir iterasyon $(-b^2, c^2)$ dışına düşerse, bu takdirde, diğere yönteme geçilir ve o kesin olarak yakınsar.

Ters tarafından başlasak bile, Newton iterasyonunun doğru tarafa düşebildiğine ve yakınsadığına dikkat edelim. İki taraftan birini rasgele doğru olarak seçmenin 50% şansı var olduğu gibi, halihazırdaki izdüşüm yöntemi öncekinden, ortalama olarak belki 1.5 kat civarında daha yavaştır. Analizin, hala yöntemin tüm durumlarında doğru izdüşüm noktasına yakınsadığı garanti edilir.

4. BULGULAR ve TARTIŞMA

4.1. Linear Modellerde Linear Kısıtlamalar Altında Parametre Tahminleri ve Hipotez Testleri

(3.29) modelinde parametrelerinin tahmini ve birçok hipotezin testi, β parametre vektörü üzerine lineer kısıtlamalar koyarak da yapılır. Bu kısıtlamalar, \mathbf{R} bilinenlerin bir $m \times p$ matrisi, $rank(\mathbf{R}_{m \times p}) = m$ ve \mathbf{r} bilinenlerin bir $m \times 1$ vektörü olmak üzere, birtakım \mathbf{R} sınırlama matrisleri için katsayılar üzerine koyulan m - sayıda

$$\mathbf{R}\beta = \mathbf{r} \quad (4.1)$$

biçimindeki kesin lineer eşitlik,

$$\mathbf{R}\beta + \mathbf{v} = \mathbf{r} \quad (4.2)$$

biçimindeki stokastik lineer eşitlik veya

$$\mathbf{R}\beta \geq \mathbf{r} \quad (4.3)$$

biçimindeki lineer eşitsizlik kısıtlamalarıdır.

4.1.1. Linear Modellerde Kesin Linear Eşitlik Kısıtlamaları

(4.1) kesin lineer eşitlik kısıtlamasının bazı örnekleri aşağıdaki gibidir.

1. $p = 6$ olmak üzere, eğer

$$\begin{aligned} \beta_2 &= \beta_4 \\ \beta_3 + 2\beta_4 + \beta_5 &= 1 \end{aligned}$$

gibi iki kısıtlama varsa, bu takdirde,

$$\mathbf{r} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \mathbf{R} = \begin{bmatrix} 0 & 1 & 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 2 & 1 & 0 \end{bmatrix}$$

dır.

2. Eğer $p = 3$ ise, $\beta_2 = 3$ olduğunu farz edersek, bu takdirde,

$$\mathbf{r} = [3], \mathbf{R} = [0 \ 1 \ 0]$$

dır.

3. Eğer $p = 3$ ve $\beta_1 : \beta_2 : \beta_3 :: ab : b : 1$ olduğunu farz edersek, bu takdirde,

$$\mathbf{r} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \mathbf{R} = \begin{bmatrix} 1 & -a & 0 \\ 0 & 1 & -b \\ 1 & 0 & -ab \end{bmatrix}$$

dir.

$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$ alışılmış en küçük kareler tahmin edicisi ön bilgi kullanmaz. O, $\mathbf{R}\boldsymbol{\beta} \neq \mathbf{r}$ anlamında kısıtlamalara uymaz. Bu nedenle sorun, $\boldsymbol{\beta}$ nın geliştirilen tahmin edicisini bulmada örneklem bilgisi ve ön bilgiyi birlikte nasıl kullanacağımızdır.

4.1.1.1. Lineer Eşitlik Kısıtlamaları Altında $\boldsymbol{\beta}$ nın En Küçük Kareler Tahmini

Kısıtlanmış en küçük kareler tahmini, aşağıdaki amaç fonksiyonunu

$$\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon} = (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) \quad (4.4)$$

(4.1) kısıtlaması altında minimumlaştırarak elde edilir. Bu durumda, λ_1' Lagrange çarpanlarının bir $m \times 1$ vektörü olmak üzere,

$$\mathfrak{L}_1 = \mathbf{y}'\mathbf{y} - 2\boldsymbol{\beta}'\mathbf{X}'\mathbf{y} + \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta} - \lambda_1'(\mathbf{r} - \mathbf{R}\boldsymbol{\beta}) \quad (4.5)$$

Lagrange fonksiyonunu oluştururuz.

\mathbf{a} ve \mathbf{b} vektörler ve \mathbf{A} uygun bir şekilde tanımlanan bir matris olmak üzere,

$$\frac{\partial(\mathbf{a}'\mathbf{A}\mathbf{a})}{\partial\mathbf{a}} = (\mathbf{A} + \mathbf{A}')\mathbf{a}, \quad (\mathbf{A} \text{ simetrik olduğunda, } = 2\mathbf{A}\mathbf{a})$$

ve

$$\frac{\partial(\mathbf{a}'\mathbf{b})}{\partial\mathbf{a}} = \mathbf{b} \quad (4.6)$$

olduğu sonucunu kullanarak, (4.5) Lagrange fonksiyonunu $\boldsymbol{\beta}$ ya ve λ_1 e göre diferensiyelleme

$$\begin{aligned} -2\mathbf{X}'\mathbf{y} + 2\mathbf{X}'\mathbf{X}\boldsymbol{\beta} + \mathbf{R}'\lambda_1 &= 0 \\ \mathbf{R}\boldsymbol{\beta} - \mathbf{r} &= 0 \end{aligned} \quad (4.7)$$

şartlarını verir.

Şimdi, (4.7) ifadesindeki birinci bağıntıyı $\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}$ ile önden çarpma

$$-2\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} + 2\mathbf{R}\boldsymbol{\beta} + \mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'\lambda_1 = 0 \quad (4.8)$$

bağıntısını verir. $\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'$ matrisi pozitif tanımlı olduğundan, $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$ alışılmış en küçük kareler (kısıtlanmamış) tahmin edicisini kullanarak,

$$\begin{aligned}
-2\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} + 2\mathbf{R}\boldsymbol{\beta} + \mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'\boldsymbol{\lambda}_1 &= 0 \\
\Rightarrow \mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'\boldsymbol{\lambda}_1 &= 2\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} - 2\mathbf{R}\boldsymbol{\beta} \\
&= 2\mathbf{R}\hat{\boldsymbol{\beta}} - 2\mathbf{R}\boldsymbol{\beta} \quad (4.9) \\
\Rightarrow \boldsymbol{\lambda}_1 &= (\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1} (2\mathbf{R}\hat{\boldsymbol{\beta}} - 2\mathbf{R}\boldsymbol{\beta}) \\
&= -2(\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1} (\mathbf{r} - \mathbf{R}\hat{\boldsymbol{\beta}})
\end{aligned}$$

bağıntısını elde ederiz. $\boldsymbol{\lambda}_1$ in bu değerini (4.7) ifadesindeki birinci bağıntıda yerine koyarak, $\boldsymbol{\beta}$ nın kısıtlanmış en küçük kareler tahmin edicisi $\hat{\boldsymbol{\beta}}_c$ için aşağıdaki denk ifadeleri elde ederiz:

$$\hat{\boldsymbol{\beta}}_c = \hat{\boldsymbol{\beta}} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}' (\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1} (\mathbf{r} - \mathbf{R}\hat{\boldsymbol{\beta}}). \quad (4.10a)$$

$$\hat{\boldsymbol{\beta}}_c = \hat{\boldsymbol{\beta}} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}' (\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1} (\mathbf{R}\hat{\boldsymbol{\beta}} - \mathbf{r}). \quad (4.10b)$$

$\hat{\boldsymbol{\beta}}_c$ kısıtlanmış en küçük kareler tahmin edicisi (4.1) bağıntısını sağlar. Gerçekten de,

$$\begin{aligned}
\mathbf{R}\hat{\boldsymbol{\beta}}_c &= \mathbf{R} \left[\hat{\boldsymbol{\beta}} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}' (\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1} (\mathbf{r} - \mathbf{R}\hat{\boldsymbol{\beta}}) \right] \\
&= \mathbf{R}\hat{\boldsymbol{\beta}} + \mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}' (\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1} (\mathbf{r} - \mathbf{R}\hat{\boldsymbol{\beta}}) = \mathbf{R}\hat{\boldsymbol{\beta}} + \mathbf{r} - \mathbf{R}\hat{\boldsymbol{\beta}} = \mathbf{r}
\end{aligned}$$

dir. $\hat{\boldsymbol{\beta}}_c$ nın yansızlığına gelince:

$$\begin{aligned}
E(\hat{\boldsymbol{\beta}}_c) &= E(\hat{\boldsymbol{\beta}}) + E \left[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}' (\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1} (\mathbf{r} - \mathbf{R}\hat{\boldsymbol{\beta}}) \right] \\
&= \boldsymbol{\beta} + \left[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}' (\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1} (\mathbf{r} - \mathbf{R}E(\hat{\boldsymbol{\beta}})) \right] \\
&= \boldsymbol{\beta} + \left[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}' (\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1} (\mathbf{r} - \mathbf{R}\boldsymbol{\beta}) \right] \quad (4.11) \\
&\quad (\text{eğer (4.1) kısıtlaması doğruysa}) \\
&= \boldsymbol{\beta}
\end{aligned}$$

olduğundan, $\hat{\boldsymbol{\beta}}_c$ kısıtlanmış en küçük kareler tahmin edicisi $\boldsymbol{\beta}$ için bir yansız tahmin edicidir. $\hat{\boldsymbol{\beta}}_c$ nin varyans-kovaryans matrisine gelince:

$$\begin{aligned}
\text{Var}(\hat{\beta}_c) &= E\left[(\hat{\beta}_c - \beta)(\hat{\beta}_c - \beta)'\right] \\
E(\hat{\beta}_c - \beta)(\hat{\beta}_c - \beta)' &= E\left[\mathbf{M}^c(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}(\mathbf{M}^c)'\right] \\
&= \mathbf{M}^c(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'E(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}')\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}(\mathbf{M}^c)' \\
&= \mathbf{M}^c(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\sigma^2\mathbf{I}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}(\mathbf{M}^c)' \\
&= \sigma^2\mathbf{M}^c(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}(\mathbf{M}^c)' \\
&= \sigma^2\mathbf{M}^c(\mathbf{X}'\mathbf{X})^{-1}(\mathbf{M}^c)'
\end{aligned} \tag{4.12}$$

elde ederiz.

Burada $\mathbf{M}^c = \left[\mathbf{I} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'(\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1}\mathbf{R}\right]$ matrisi idempotent fakat simetrik değildir. Gerçekten

$$\begin{aligned}
\mathbf{M}^c\mathbf{M}^c &= \left[\mathbf{I} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'(\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1}\mathbf{R}\right] \left[\mathbf{I} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'(\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1}\mathbf{R}\right] \\
&= \mathbf{I} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'(\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1}\mathbf{R} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'(\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1}\mathbf{R} \\
&\quad + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'(\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1}\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'(\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1}\mathbf{R} \\
&= \mathbf{I} - 2(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'(\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1}\mathbf{R} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'(\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1}\mathbf{R} \\
&= \mathbf{I} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'(\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1}\mathbf{R}
\end{aligned} \tag{4.13}$$

olduğundan, \mathbf{M}^c idempotenttir. Transpoz olarak simetrik olmadığı görülür. Aynı zamanda $\text{Var}(\hat{\beta}_c)$ için aşağıdaki denk ifadeyi de yazabiliriz:

$$\begin{aligned}
Var(\hat{\beta}_c) &= \sigma^2 (\mathbf{M}^c)' (\mathbf{X}\mathbf{X})^{-1} (\mathbf{M}^c) \\
&= \sigma^2 \left[\mathbf{I} - (\mathbf{X}\mathbf{X})^{-1} \mathbf{R}' (\mathbf{R}(\mathbf{X}\mathbf{X})^{-1} \mathbf{R}')^{-1} \mathbf{R} \right] (\mathbf{X}\mathbf{X})^{-1} \left[\mathbf{I} - \mathbf{R}' (\mathbf{R}(\mathbf{X}\mathbf{X})^{-1} \mathbf{R}')^{-1} \mathbf{R}(\mathbf{X}\mathbf{X})^{-1} \right] \\
&= \sigma^2 \left[(\mathbf{X}\mathbf{X})^{-1} \right] \\
&\quad - \sigma^2 \left[(\mathbf{X}\mathbf{X})^{-1} \mathbf{R}' (\mathbf{R}(\mathbf{X}\mathbf{X})^{-1} \mathbf{R}')^{-1} \mathbf{R}(\mathbf{X}\mathbf{X})^{-1} \right] \\
&\quad - \sigma^2 \left[(\mathbf{X}\mathbf{X})^{-1} \mathbf{R}' (\mathbf{R}(\mathbf{X}\mathbf{X})^{-1} \mathbf{R}')^{-1} \mathbf{R}(\mathbf{X}\mathbf{X})^{-1} \right] \\
&\quad + \sigma^2 \left[(\mathbf{X}\mathbf{X})^{-1} \mathbf{R}' (\mathbf{R}(\mathbf{X}\mathbf{X})^{-1} \mathbf{R}')^{-1} \mathbf{R}(\mathbf{X}\mathbf{X})^{-1} \mathbf{R}' (\mathbf{R}(\mathbf{X}\mathbf{X})^{-1} \mathbf{R}')^{-1} \mathbf{R}(\mathbf{X}\mathbf{X})^{-1} \right] \\
&= \sigma^2 \left[(\mathbf{X}\mathbf{X})^{-1} - 2(\mathbf{X}\mathbf{X})^{-1} \mathbf{R}' (\mathbf{R}(\mathbf{X}\mathbf{X})^{-1} \mathbf{R}')^{-1} \mathbf{R}(\mathbf{X}\mathbf{X})^{-1} + (\mathbf{X}\mathbf{X})^{-1} \mathbf{R}' (\mathbf{R}(\mathbf{X}\mathbf{X})^{-1} \mathbf{R}')^{-1} \mathbf{R}(\mathbf{X}\mathbf{X})^{-1} \right] \\
&= \sigma^2 \left[(\mathbf{X}\mathbf{X})^{-1} - (\mathbf{X}\mathbf{X})^{-1} \mathbf{R}' (\mathbf{R}(\mathbf{X}\mathbf{X})^{-1} \mathbf{R}')^{-1} \mathbf{R}(\mathbf{X}\mathbf{X})^{-1} \right] \\
&= \sigma^2 \left[\mathbf{I} - (\mathbf{X}\mathbf{X})^{-1} \mathbf{R}' (\mathbf{R}(\mathbf{X}\mathbf{X})^{-1} \mathbf{R}')^{-1} \mathbf{R} \right] (\mathbf{X}\mathbf{X})^{-1} \\
&= \sigma^2 \mathbf{M}^c (\mathbf{X}\mathbf{X})^{-1}. \tag{4.14}
\end{aligned}$$

(4.15) bağıntısını aşağıdaki gibi de yazabiliriz:

$$\begin{aligned}
Var(\hat{\beta}_c) &= \sigma^2 \left[(\mathbf{X}\mathbf{X})^{-1} - (\mathbf{X}\mathbf{X})^{-1} \mathbf{R}' (\mathbf{R}(\mathbf{X}\mathbf{X})^{-1} \mathbf{R}')^{-1} \mathbf{R}(\mathbf{X}\mathbf{X})^{-1} \right] \\
&= \sigma^2 (\mathbf{X}\mathbf{X})^{-1} - \sigma^2 (\mathbf{X}\mathbf{X})^{-1} \mathbf{R}' (\mathbf{R}(\mathbf{X}\mathbf{X})^{-1} \mathbf{R}')^{-1} \mathbf{R}(\mathbf{X}\mathbf{X})^{-1}. \tag{4.15}
\end{aligned}$$

Ayrıca $\hat{\beta}$ için

$$E(\hat{\beta}) = (\mathbf{X}\mathbf{X})^{-1} \mathbf{X}' E(\mathbf{y}) = (\mathbf{X}\mathbf{X})^{-1} \mathbf{X}\mathbf{X}\beta = \beta, \tag{4.16}$$

yani, $\hat{\beta}$ nın β için yansız olduğunu ve

$$Var(\hat{\beta}) = (\mathbf{X}\mathbf{X})^{-1} \mathbf{X}' \underbrace{Var(\mathbf{y})}_{\sigma^2 \mathbf{I}} \mathbf{X}(\mathbf{X}\mathbf{X})^{-1} = \sigma^2 (\mathbf{X}\mathbf{X})^{-1} \tag{4.17}$$

olduğunu görürüz.

Böylece, kısıtlanmış en küçük kareler tahmin edicisinin varyansı, alışılmış en küçük kareler tahmin edicisinin varyansından bir pozitif yarı tanımlı matris kadar eksiktir. Yani, $Var(\hat{\beta}) \succcurlyeq Var(\hat{\beta}_c)$ olduğu, bu nedenle $\hat{\beta}_c$ nın; kovaryans matrislerinin kriteri veya σ^2 nin biliniyor olması şartı ile Loewner sıralaması altında $\hat{\beta}$ dan daha etkin olduğu ispatlanmış olur (Chow, 1960; Fisher, 1970).

4.1.1.2. $\hat{\beta}_c$ yi Elde Etmek için Bir Başka Yöntem

Şimdi burada $\hat{\beta}_c$ yi elde etmek için başka bir yöntemi ortaya koyacağız. Bunun için aşağıdaki tanım ve teoremi verelim.

Kısıtlanmış Parametre Uzayı :

Şimdi bir (4.1) lineer denklemler sistemiyle sınırlanmış parametre uzayına sahip lineer modeli tekrar göz önüne alalım. Bu incelemeyi hem tahmin edilebilir hem de tahmin edilemez kısıtlamaları göz önüne almak suretiyle yapacağız. Kısıtlamaların tahmin edilebilirliğe nasıl etki edeceği ve bu kısıtlamaya bağlı hata kareler toplamının nasıl minimum yapılacağı ile ilgileneceğiz.

β ; $\mathbf{R}\beta = \mathbf{r}$ tutarlı lineer denklemler sistemini sağlamak üzere, β için parametre uzayının, \mathbb{R}^p nin bir \mathcal{T} alt kümesine, yani $\mathcal{T} = \{\beta : \mathbf{R}\beta = \mathbf{r}\}$ ya kısıtlandığı, beklenen değeri $\mathbf{X}\beta$ olan \mathbf{y} yi gözleriz.

\mathbf{R} matrisinin tam ranklı $m \times p$ olduğunda ısrar edeceğiz; aksi takdirde gereksiz denklemlere sahip olacağız, fakat denklemlerin tutarlı olmalarını garantilemek için $\mathbf{r} \in C(\mathbf{R})$ olması yeterlidir.

Tanım 4.1. $\lambda \beta$ fonksiyonunun kısıtlanmış modelde tahmin edilebilir olması için gerek ve yeter şart $\mathbf{R}\beta = \mathbf{r}$ denklemini sağlayan her β için $E(c + \mathbf{a}'\mathbf{y}) = \lambda \beta$ olacak şekilde bir c skaları ve bir \mathbf{a} ($n \times 1$) vektörünün mevcut olmasıdır.

Her β için $E(c + \mathbf{a}'\mathbf{y}) = \lambda \beta$ olmak üzere, eğer $\lambda \beta$ kısıtlanmamış modelde tahmin edilebilseydi, $\beta \in \mathcal{T}$ alt kümesine kısıtlandığı için $\lambda \beta$ nin tahmin edilebilir olması gerektiğine dikkat edelim.

Sonuç 4.1. Kesin eşitlik kısıtlamalı modelde, $(c + \mathbf{a}'\mathbf{y})$ nin $\lambda \beta$ için yansız olması için $\Leftrightarrow \lambda = \mathbf{X}'\mathbf{a} + \mathbf{R}'\mathbf{d}$ ve $c = \mathbf{d}'\mathbf{r}$ olacak şekilde \mathbf{d} nin mevcut olmasıdır.

İspat. (\Rightarrow) Eğer $\mathbf{R}\beta = \mathbf{r}$ denklemini sağlayan her β için $E(c + \mathbf{a}'\mathbf{y}) = \lambda \beta$ ise, bu takdirde bu denklemlere tüm çözümleri- yani \mathcal{T} daki tüm β ları- \mathbf{z} keyfi bir vektör olmak üzere $\beta = \mathbf{R}^+\mathbf{r} + \underbrace{(\mathbf{I} - \mathbf{R}^+\mathbf{R})}_{=\mathbf{W}}\mathbf{z}$ biçiminde veririz. $\mathbf{z} = \mathbf{0}$ olduğunda, $\beta_* = \mathbf{R}^+\mathbf{r}$,

\mathcal{T} da özel bir çözümdür. Bu durumda,

$$C(\mathbf{I} - \mathbf{R}^+\mathbf{R}) = C(\mathbf{W}) = \{\mathbf{w} : \mathbf{w} = \mathbf{W}\mathbf{t}, \quad \mathbf{t} \in \mathbb{R}^p\}$$

olarak alalım. Burada, $\mathbf{R}\mathbf{w} = \mathbf{R}\mathbf{W}\mathbf{t} = \mathbf{R}(\mathbf{I} - \mathbf{R}^+\mathbf{R})\mathbf{t} = \mathbf{R}\mathbf{t} - \mathbf{R}\mathbf{t} = \mathbf{0} \Rightarrow \mathbf{w} \in \mathcal{N}(\mathbf{R})$ dir.

$\mathbf{y} \in \mathcal{N}(\mathbf{R})$ olsun. Bu takdirde,

$$\mathbf{R}\mathbf{y} = \mathbf{0} \Rightarrow \mathbf{y} = \mathbf{R}^+\mathbf{0} + (\mathbf{I} - \mathbf{R}^+\mathbf{R})\mathbf{t}_1 = (\mathbf{I} - \mathbf{R}^+\mathbf{R})\mathbf{t}_1, \quad \mathbf{t}_1 \in \mathbb{R}^p \text{ biçimindedir. Yani,}$$

$\mathbf{y} \in \mathcal{C}(\mathbf{W})$ dir. Bu nedenle $\mathcal{N}(\mathbf{R}) = \mathcal{C}(\mathbf{W})$ dir.

Bu durumda her \mathbf{z} için

$$E(c + \mathbf{a}'\mathbf{y}) = c + \mathbf{a}' E(\mathbf{y}) = c + \mathbf{a}' \underbrace{\mathbf{X}(\boldsymbol{\beta}_* + \mathbf{W}\mathbf{z})}_{=\mathbf{x}\boldsymbol{\beta}} = \lambda' \underbrace{(\boldsymbol{\beta}_* + \mathbf{W}\mathbf{z})}_{=\boldsymbol{\beta}} \quad (4.18)$$

bağıntısını elde ederiz. Şimdi aşağıdaki faydalı teoremi verelim.

Teorem 4.1. \mathbf{A} bir $m \times n$ matris ve \mathbf{b} bir $n \times 1$ sabit vektör olsun. Eğer her $\mathbf{x} \in \mathbb{R}^n$ için $\mathbf{A}\mathbf{x} + \mathbf{b} = \mathbf{0}$ ise, bu takdirde $\mathbf{A} = \mathbf{0}$ ve $\mathbf{b} = \mathbf{0}$ dır (Monahan, 2008, Sonuç A.8).

(4.18) bağıntısı aşağıdaki biçimde düzenlenebilir:

$$\underbrace{(\lambda' - \mathbf{a}'\mathbf{X})\mathbf{W}}_{=\mathbf{A}} \mathbf{z} + \underbrace{\lambda'\boldsymbol{\beta}_* - \mathbf{a}'\mathbf{X}\boldsymbol{\beta}_* - c}_{=\mathbf{b}} = \mathbf{0}. \quad (4.19)$$

Şimdi Teorem 4.1. e göre $(\lambda' - \mathbf{a}'\mathbf{X})\mathbf{W} = \mathbf{0}$ veya $(\lambda - \mathbf{X}\mathbf{a})'\mathbf{W} = \mathbf{0}$ ya da $\mathbf{W}'(\lambda - \mathbf{X}\mathbf{a}) = \mathbf{0}$ yazılır. Bu nedenle, $(\lambda - \mathbf{X}\mathbf{a}) \in \mathcal{N}(\mathbf{W}')$ dür.

\mathbf{W} bir simetrik-idempotent bir matris olduğundan, $\mathcal{N}(\mathbf{W}) = \mathcal{N}(\mathbf{W}')$ dir.

$$\mathbf{W} = (\mathbf{I} - \mathbf{R}^+\mathbf{R}) \Rightarrow \mathbf{W}' = (\mathbf{I} - \mathbf{R}^+\mathbf{R})$$

$$\mathbf{t} \in \mathcal{N}(\mathbf{W}') \Rightarrow (\mathbf{I} - \mathbf{R}^+\mathbf{R})\mathbf{t} = \mathbf{0} \Rightarrow \mathbf{t} = \mathbf{R}^+\mathbf{R}\mathbf{t} = \mathbf{R}'\mathbf{R}^+\mathbf{t} \Rightarrow \mathbf{t} \in \mathcal{C}(\mathbf{R}') \quad =\boldsymbol{\mu}$$

$$\Rightarrow (\lambda - \mathbf{X}\mathbf{a}) \in \mathcal{N}(\mathbf{W}') = \mathcal{C}(\mathbf{R}')$$

dür.

$$\lambda - \mathbf{X}\mathbf{a} = \mathbf{R}'\mathbf{d} \Rightarrow \lambda = \mathbf{X}\mathbf{a} + \mathbf{R}'\mathbf{d} \text{ ve } c + \mathbf{a}'\mathbf{X}\boldsymbol{\beta} = \lambda' \boldsymbol{\beta} = \mathbf{a}'\mathbf{X}\boldsymbol{\beta} + \mathbf{d}'\mathbf{R}\boldsymbol{\beta} \Rightarrow c = \mathbf{d}'\mathbf{r} \quad =\mathbf{a}'\mathbf{X}+\mathbf{d}\mathbf{R} \quad =\mathbf{r}$$

olacak şekilde bir \mathbf{d} vektörü mevcuttur.

(\Leftarrow) $\mathbf{R}\boldsymbol{\beta} = \mathbf{r}$ denklemini sağlayan her $\boldsymbol{\beta}$ için $\lambda = \mathbf{X}\mathbf{a} + \mathbf{R}'\mathbf{d}$ ve $c = \mathbf{d}'\mathbf{r}$ olacak şekilde \mathbf{d} mevcut olsun.

Buradan,

$$\begin{aligned}\lambda' &= \mathbf{a}'\mathbf{X} + \mathbf{d}'\mathbf{R} \Rightarrow \lambda' \boldsymbol{\beta} = \mathbf{a}'\mathbf{X}\boldsymbol{\beta} + \mathbf{d}'\mathbf{R}\boldsymbol{\beta} \Rightarrow \mathbf{a}'\mathbf{X}\boldsymbol{\beta} = \lambda' \boldsymbol{\beta} - \mathbf{d}'\mathbf{R}\boldsymbol{\beta} \\ E(c + \mathbf{a}'\mathbf{y}) &= c + \mathbf{a}'\mathbf{X}\boldsymbol{\beta} \stackrel{=\mathbf{d}'\mathbf{r}}{=} \mathbf{d}'\mathbf{r} + \lambda' \boldsymbol{\beta} - \mathbf{d}'\mathbf{R}\boldsymbol{\beta} = \lambda' \boldsymbol{\beta} + \mathbf{d}' \left(\mathbf{r} - \mathbf{R}\boldsymbol{\beta} \right) \\ &\stackrel{=\lambda' \boldsymbol{\beta} - \mathbf{d}'\mathbf{R}\boldsymbol{\beta}}{=} \lambda' \boldsymbol{\beta}\end{aligned}$$

dir. Yani $(c + \mathbf{a}'\mathbf{y})$, $\lambda' \boldsymbol{\beta}$ için yansızdır (Monahan, 2008).

$\mathbf{R}\boldsymbol{\beta} = \mathbf{r}$ kısıtlaması ve $rank(\mathbf{R}_{m \times p}) = m \leq p$ şartı altında, \mathbf{R} nin $\mathbf{R}^+ = \mathbf{R}'(\mathbf{R}\mathbf{R}')^{-1}$ Moore-Penrose genelleştirilmiş tersini (\mathbf{R}^+ tektir.) kullanarak (4.1) lineer denklemler sisteminden

$$\boldsymbol{\beta}^* = \mathbf{R}^+\mathbf{r} + (\mathbf{I} - \mathbf{R}^+\mathbf{R})\mathbf{h}, \quad \mathbf{h} \in \mathbb{R}^p \quad (4.20)$$

çözümünü elde ederiz. Burada $\mathbf{h} \in \mathbb{R}^p$ keyfi bir vektördür ve $\mathbf{R}\mathbf{R}^+ = \mathbf{R}\mathbf{R}'(\mathbf{R}\mathbf{R}')^{-1} = \mathbf{I}$ dir. Şimdi optimum $\mathbf{h} \in \mathbb{R}^p$ değerini arayalım ve buna $\hat{\mathbf{h}}$ diyelim. $\boldsymbol{\beta}^*$ değerini (3.29) modelinde yerine koyarak,

$$\underbrace{\mathbf{y} - \mathbf{X}\mathbf{R}^+\mathbf{r}}_{=\boldsymbol{\omega}} = \mathbf{X} \underbrace{(\mathbf{I} - \mathbf{R}^+\mathbf{R})\mathbf{h}}_{=\boldsymbol{\gamma}} + \boldsymbol{\varepsilon} \quad (4.21)$$

bağıntısını elde ederiz. Bu takdirde $\boldsymbol{\gamma}$ nın $\hat{\boldsymbol{\gamma}}$ en küçük kareler tahmin edicisi,

$\mathbf{X}'\mathbf{X}\boldsymbol{\gamma} = \mathbf{X}'\boldsymbol{\omega}$ normal denklemleri sağlar. Buradan

$$\hat{\boldsymbol{\gamma}} = (\mathbf{I} - \mathbf{R}^+\mathbf{R})\hat{\mathbf{h}} = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'(\mathbf{y} - \mathbf{X}\mathbf{R}^+\mathbf{r}) = \hat{\boldsymbol{\beta}} - \mathbf{R}^+\mathbf{r} \quad (4.22)$$

olduğu görülür. (4.22) denkleminin $\hat{\mathbf{h}}$ için çözümünden

$$\begin{aligned}\hat{\mathbf{h}} &= (\mathbf{I} - \mathbf{R}^+\mathbf{R})(\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'(\mathbf{y} - \mathbf{X}\mathbf{R}^+\mathbf{r}) + \left[\mathbf{I} - \underbrace{(\mathbf{I} - \mathbf{R}^+\mathbf{R})'(\mathbf{I} - \mathbf{R}^+\mathbf{R})}_{=(\mathbf{I} - \mathbf{R}^+\mathbf{R})} \right] \boldsymbol{\zeta} \\ &= (\mathbf{I} - \mathbf{R}^+\mathbf{R})(\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'(\mathbf{y} - \mathbf{X}\mathbf{R}^+\mathbf{r}) + \left[\underbrace{\mathbf{I} - (\mathbf{I} - \mathbf{R}^+\mathbf{R})}_{=\mathbf{R}^+\mathbf{R}} \right] \boldsymbol{\zeta} \\ &= (\mathbf{I} - \mathbf{R}^+\mathbf{R})(\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'(\mathbf{y} - \mathbf{X}\mathbf{R}^+\mathbf{r}) + \mathbf{R}^+\mathbf{R}\boldsymbol{\zeta}\end{aligned} \quad (4.23)$$

elde edilir. Burada $\boldsymbol{\zeta} \in \mathbb{R}^p$ keyfi bir vektördür.

(4.20) bağıntısında \mathbf{h} yerine $\hat{\mathbf{h}}$ koyarak $\hat{\boldsymbol{\beta}}_c$ yı elde ederiz. Buna göre,

$$\begin{aligned}
\hat{\beta}_c &= \mathbf{R}^+ \mathbf{r} + (\mathbf{I} - \mathbf{R}^+ \mathbf{R}) \left\{ (\mathbf{I} - \mathbf{R}^+ \mathbf{R}) (\mathbf{X}' \mathbf{X})^{-1} \mathbf{X}' (\mathbf{y} - \mathbf{X} \mathbf{R}^+ \mathbf{r}) + \mathbf{R}^+ \mathbf{R} \zeta \right\} \\
&= \mathbf{R}^+ \mathbf{r} + (\mathbf{I} - \mathbf{R}^+ \mathbf{R}) (\mathbf{X}' \mathbf{X})^{-1} \mathbf{X}' (\mathbf{y} - \mathbf{X} \mathbf{R}^+ \mathbf{r}) + \underbrace{(\mathbf{I} - \mathbf{R}^+ \mathbf{R}) \mathbf{R}^+ \mathbf{R} \zeta}_{=0} \\
&= \mathbf{R}^+ \mathbf{r} + (\mathbf{I} - \mathbf{R}^+ \mathbf{R}) (\mathbf{X}' \mathbf{X})^{-1} \mathbf{X}' (\mathbf{y} - \mathbf{X} \mathbf{R}^+ \mathbf{r}) \\
&= \mathbf{R}^+ \mathbf{r} + (\mathbf{I} - \mathbf{R}^+ \mathbf{R}) (\hat{\beta} - \mathbf{R}^+ \mathbf{r}) \\
&= \mathbf{R}^+ \mathbf{r} + (\mathbf{I} - \mathbf{R}^+ \mathbf{R}) \hat{\beta}
\end{aligned} \tag{4.24}$$

olur.

Eğer \mathbf{R}^+ yerine $(\mathbf{X}' \mathbf{X})^{-1} \mathbf{R}' \left[\mathbf{R} (\mathbf{X}' \mathbf{X})^{-1} \mathbf{R}' \right]^{-1}$ i kullanırsak, bu takdirde bu genelleştirilmiş ters, Moore-Penrose genelleştirilmiş tersin ilk üç özelliğini sağlar, yani,

(i) $\mathbf{R} \mathbf{R}^+ \mathbf{R} = \mathbf{R}$, (ii) $\mathbf{R}^+ \mathbf{R} \mathbf{R}^+ = \mathbf{R}^+$ ve (iii) $(\mathbf{R} \mathbf{R}^+)' = \mathbf{R} \mathbf{R}^+$ dir.

Bu nedenle \mathbf{R}^+ yerine $(\mathbf{X}' \mathbf{X})^{-1} \mathbf{R}' \left[\mathbf{R} (\mathbf{X}' \mathbf{X})^{-1} \mathbf{R}' \right]^{-1}$ Moore-Penrose tersini kullanırsak,

$$\hat{\beta}_c = \left((\mathbf{X}' \mathbf{X})^{-1} \mathbf{R}' \left[\mathbf{R} (\mathbf{X}' \mathbf{X})^{-1} \mathbf{R}' \right]^{-1} \right) \mathbf{r} + \left(\mathbf{I} - \left((\mathbf{X}' \mathbf{X})^{-1} \mathbf{R}' \left[\mathbf{R} (\mathbf{X}' \mathbf{X})^{-1} \mathbf{R}' \right]^{-1} \right) \mathbf{R} \right) \hat{\beta},$$

$$\hat{\beta}_c = \hat{\beta} + \left((\mathbf{X}' \mathbf{X})^{-1} \mathbf{R}' \left[\mathbf{R} (\mathbf{X}' \mathbf{X})^{-1} \mathbf{R}' \right]^{-1} \right) (\mathbf{r} - \mathbf{R} \hat{\beta}), \tag{4.25a}$$

$$\hat{\beta}_c = \hat{\beta} - \left((\mathbf{X}' \mathbf{X})^{-1} \mathbf{R}' \left[\mathbf{R} (\mathbf{X}' \mathbf{X})^{-1} \mathbf{R}' \right]^{-1} \right) (\mathbf{R} \hat{\beta} - \mathbf{r}) \tag{4.25b}$$

olduğunu görürüz. (4.25a) ve (4.25b) bağıntılarının (4.10a) ve (4.10b) bağıntılarının aynısı olduğu görülür. Yine, (4.24), (4.25a) veya (4.25b) deki tahmin ediciler (4.1) denklemini, yani, $\mathbf{R} \hat{\beta} = \mathbf{r}$ yi sağlar.

(4.24) de verilen $\hat{\beta}_c$ yansızdır, gerçekten

$$\begin{aligned}
E(\hat{\beta}_c) &= \mathbf{R}^+ \mathbf{r} + (\mathbf{I} - \mathbf{R}^+ \mathbf{R}) E(\hat{\beta}) \\
&= \mathbf{R}^+ \mathbf{r} + \beta - \mathbf{R}^+ \mathbf{R} \beta
\end{aligned}$$

dır. Eğer kısıtlama sağlanırsa,

$$E(\hat{\beta}_c) = \mathbf{R}^+ \mathbf{r} + \beta - \mathbf{R}^+ \mathbf{R} \beta = \beta$$

=r

yani, $\hat{\beta}_c$ nin β nin bir yansız tahmin edicisi olduğu görülür.

(4.24) deki kısıtlanmış en küçük kareler tahmin edicisinin varyans-kovaryansına gelince,

$$\begin{aligned} \text{Var}(\hat{\beta}_c) &= \sigma^2 (\mathbf{I} - \mathbf{R}^+ \mathbf{R}) (\mathbf{X} \mathbf{X})^{-1} (\mathbf{I} - \mathbf{R}^+ \mathbf{R}) \\ &= \sigma^2 \left((\mathbf{X} \mathbf{X})^{-1} - \mathbf{R}^+ \mathbf{R} (\mathbf{X} \mathbf{X})^{-1} \right) (\mathbf{I} - \mathbf{R}^+ \mathbf{R}) \\ &= \underbrace{\sigma^2 (\mathbf{X} \mathbf{X})^{-1}}_{=\text{Var}(\hat{\beta})} + \sigma^2 \left(-(\mathbf{X} \mathbf{X})^{-1} \mathbf{R}^+ \mathbf{R} - \mathbf{R}^+ \mathbf{R} (\mathbf{X} \mathbf{X})^{-1} + \mathbf{R}^+ \mathbf{R} (\mathbf{X} \mathbf{X})^{-1} \mathbf{R}^+ \mathbf{R} \right) \quad (4.26) \\ \text{Var}(\hat{\beta}) - \text{Var}(\hat{\beta}_c) &= \sigma^2 \left((\mathbf{X} \mathbf{X})^{-1} \mathbf{R}^+ \mathbf{R} + \mathbf{R}^+ \mathbf{R} (\mathbf{X} \mathbf{X})^{-1} - \mathbf{R}^+ \mathbf{R} (\mathbf{X} \mathbf{X})^{-1} \mathbf{R}^+ \mathbf{R} \right) \end{aligned}$$

bağıntısını elde ederiz. $\mathbf{A} = \left((\mathbf{X} \mathbf{X})^{-1} \mathbf{R}^+ \mathbf{R} + \mathbf{R}^+ \mathbf{R} (\mathbf{X} \mathbf{X})^{-1} - \mathbf{R}^+ \mathbf{R} (\mathbf{X} \mathbf{X})^{-1} \mathbf{R}^+ \mathbf{R} \right)$ diyelim. \mathcal{T} parametre uzayından $\mathbf{z} \in C(\mathbf{R}') = \mathcal{N}(\mathbf{I} - \mathbf{R}^+ \mathbf{R}) = \mathcal{N}(\mathbf{I} - \mathbf{R} \mathbf{R}^+)$ vektörünü seçelim. $\mathbf{z} = \mathbf{R}' \mathbf{t}$, $\mathbf{t} \in \mathbb{R}^m$ ve $\mathbf{z} \in \mathbb{R}^p$ olacaktır. O halde, her $\mathbf{z} \in \mathbb{R}^p$ için

$$\begin{aligned} \mathbf{z}' \mathbf{A} \mathbf{z} &= \mathbf{t}' \mathbf{R}' \left[(\mathbf{X} \mathbf{X})^{-1} \mathbf{R}^+ \mathbf{R} + \mathbf{R}^+ \mathbf{R} (\mathbf{X} \mathbf{X})^{-1} - \mathbf{R}^+ \mathbf{R} (\mathbf{X} \mathbf{X})^{-1} \mathbf{R}^+ \mathbf{R} \right] \mathbf{R}' \mathbf{t}, \\ &= \mathbf{t}' \left[\mathbf{R} (\mathbf{X} \mathbf{X})^{-1} \underbrace{\mathbf{R}^+ \mathbf{R} \mathbf{R}'}_{=\mathbf{R}'} + \underbrace{\mathbf{R} \mathbf{R}^+ \mathbf{R}'}_{=\mathbf{R}} (\mathbf{X} \mathbf{X})^{-1} \mathbf{R}' - \underbrace{\mathbf{R} \mathbf{R}^+ \mathbf{R}'}_{=\mathbf{R}} (\mathbf{X} \mathbf{X})^{-1} \underbrace{\mathbf{R}^+ \mathbf{R} \mathbf{R}'}_{=\mathbf{R}'} \right] \mathbf{t} \quad (4.27) \\ &= \mathbf{t}' \left[\mathbf{R} (\mathbf{X} \mathbf{X})^{-1} \mathbf{R}' + \mathbf{R} (\mathbf{X} \mathbf{X})^{-1} \mathbf{R}' - \mathbf{R} (\mathbf{X} \mathbf{X})^{-1} \mathbf{R}' \right] \mathbf{t} \\ &= \mathbf{t}' \left[\mathbf{R} (\mathbf{X} \mathbf{X})^{-1} \mathbf{R}' \right] \mathbf{t} \end{aligned}$$

dir. Burada $\text{rank}(\mathbf{X}_{n \times p}) = p$ olduğundan, $(\mathbf{X} \mathbf{X})$, $(\mathbf{X} \mathbf{X})^{-1}$ ve $\mathbf{R} (\mathbf{X} \mathbf{X})^{-1} \mathbf{R}'$ matrisleri pozitif tanımlıdır.

Böylece, $\hat{\beta}_c$ kısıtlanmış en küçük kareler tahmin edicisinin varyansı, $\hat{\beta}$ alışımlı en küçük kareler tahmin edicisinin varyansından bir pozitif tanımlı matris kadar eksiktir. Yani, $\text{Var}(\hat{\beta}) \succeq \text{Var}(\hat{\beta}_c)$ olduğu, bu nedenle $\hat{\beta}_c$ nin; kovaryans matrislerinin kriteri veya σ^2 nin biliniyor olması şartı ile Loewner sıralaması altında $\hat{\beta}$ dan daha etkin olduğu ispatlanmış olur.

Bu arada $\mathbf{R}^+ = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{R}' [\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1} \mathbf{R}']^{-1}$ aldığımızda, \mathbf{R}^+ in Moore-Penrose un sadece dördüncü özelliğini sağlamadığını ve bu durumda ortaya koyduğumuz yöntemle elde ettiğimiz $\hat{\boldsymbol{\beta}}_c$ nin (4.10a) veya (4.10b) ile verilen kısıtlanmış en küçük kareler tahmin edicisine denk olduğunu daha önce belirtmiştik. (4.24) bağıntısını soldan $\mathbf{X}'\mathbf{X}$ ile çarpalım.

$$\begin{aligned}
\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}}_c &= \mathbf{X}'\mathbf{X}\mathbf{R}^+\mathbf{r} + \mathbf{X}'\mathbf{X} \underbrace{(\mathbf{I} - \mathbf{R}^+\mathbf{R})}_{=\hat{\boldsymbol{\gamma}}} \hat{\boldsymbol{\beta}} \\
&= \mathbf{X}'\mathbf{X}\mathbf{R}^+\mathbf{r} + \mathbf{X}'\mathbf{X}\hat{\boldsymbol{\gamma}} = \mathbf{X}'\mathbf{X}\mathbf{R}^+\mathbf{r} + \mathbf{X}'\boldsymbol{\omega} \\
&= \mathbf{X}'\mathbf{X}\mathbf{R}^+\mathbf{r} + \mathbf{X}'(\mathbf{y} - \mathbf{X}\mathbf{R}^+\mathbf{r}) \\
&= \mathbf{X}'\mathbf{X}\mathbf{R}^+\mathbf{r} + \mathbf{X}'\mathbf{y} - \mathbf{X}'\mathbf{X}\mathbf{R}^+\mathbf{r} \\
&= \mathbf{X}'\mathbf{y}
\end{aligned} \tag{4.28}$$

olur. Bu bağıntıyı soldan $(\mathbf{I} - \mathbf{R}^+\mathbf{R})$ ile çarparak

$$(\mathbf{I} - \mathbf{R}^+\mathbf{R})\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}}_c = (\mathbf{I} - \mathbf{R}^+\mathbf{R})\mathbf{X}'\mathbf{y} \tag{4.29a}$$

veya

$$(\mathbf{I} - \mathbf{R}^+\mathbf{R})(\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}}_c - \mathbf{X}'\mathbf{y}) = 0 \tag{4.29b}$$

elde ederiz. Buradan $(\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}}_c - \mathbf{X}'\mathbf{y}) \in \mathcal{N}(\mathbf{I} - \mathbf{R}^+\mathbf{R})$ olduğu yani,

$(\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}}_c - \mathbf{X}'\mathbf{y}) \in \mathcal{C}(\mathbf{R}')$ veya $(\mathbf{X}'\mathbf{X})\hat{\boldsymbol{\beta}}_c = \mathbf{X}'\mathbf{y} + \mathbf{R}'\boldsymbol{\lambda}$ olduğu görülür. Bu nedenle $\hat{\boldsymbol{\beta}}_c$ tektir.

Şimdi $\text{Var}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})$ yı hesaplayalım:

$$E(\hat{\boldsymbol{\beta}}) = \boldsymbol{\beta} \text{ ve}$$

$$\hat{\boldsymbol{\beta}} - \boldsymbol{\beta} = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{y} - \boldsymbol{\beta} = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'(\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}) - \boldsymbol{\beta} = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\boldsymbol{\varepsilon} \tag{4.30}$$

olduğundan,

$$\begin{aligned}
\text{Var}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) &= E\left[(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})'\right] \\
&= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' \underbrace{E[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}']}_{=\sigma^2\mathbf{I}_n} \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} \\
&= \sigma^2(\mathbf{X}'\mathbf{X})^{-1} = \text{Var}(\hat{\boldsymbol{\beta}})
\end{aligned} \tag{4.31}$$

dır. $\text{Var}(\hat{\boldsymbol{\beta}}_c - \hat{\boldsymbol{\beta}})$ ya gelince; bu durumda (4.25a) dan $\hat{\boldsymbol{\beta}}_c - \hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}']^{-1}(\mathbf{r} - \mathbf{R}\hat{\boldsymbol{\beta}})$ ve $E(\hat{\boldsymbol{\beta}}_c - \hat{\boldsymbol{\beta}}) = 0$ olduğu görülür. Buna göre,

$$\begin{aligned}
\text{Var}(\hat{\boldsymbol{\beta}}_c - \hat{\boldsymbol{\beta}}) &= E\left[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}']^{-1}(\mathbf{r} - \mathbf{R}\hat{\boldsymbol{\beta}})(\mathbf{r} - \mathbf{R}\hat{\boldsymbol{\beta}})'[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}']^{-1}\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\right] \\
&= \left[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}']^{-1} E[(\mathbf{r} - \mathbf{R}\hat{\boldsymbol{\beta}})(\mathbf{r} - \mathbf{R}\hat{\boldsymbol{\beta}})'] [\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}']^{-1}\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\right] \\
&= \left[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}']^{-1} \underbrace{\text{Var}[(\mathbf{r} - \mathbf{R}\hat{\boldsymbol{\beta}})]}_{=\sigma^2[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}']} [\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}']^{-1}\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\right] \\
&= \sigma^2 \left[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}']^{-1} [\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'] [\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}']^{-1}\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\right] \\
&= \sigma^2(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}']^{-1}\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}
\end{aligned} \tag{4.32}$$

dir.

4.1.1.3. $\boldsymbol{\beta}$ nın Kısıtlanmış En Çok Olabilirlik Tahmini

σ^2 üzerinde kısıtlama olmadığında, $\boldsymbol{\beta}$ nın kısıtlanmış tahmin edicisinin bir fonksiyonu olarak σ^2 nin bir tahmin edicisini elde etmek için (3.17) bağıntısını σ^2 ye göre diferensiyelleyebiliriz. Bunu yaparak,

$$\begin{aligned}
\frac{\partial \ell}{\partial \sigma^2} &= -\frac{n}{2\sigma^2} + \frac{1}{2} \left(\frac{1}{\sigma^2}\right)^2 (\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}}_c)'(\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}}_c) = 0 \\
\Rightarrow \tilde{\sigma}_c^2 &= \frac{1}{n} (\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}}_c)'(\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}}_c)
\end{aligned} \tag{4.33}$$

bağıntısını elde ederiz. Burada $\tilde{\boldsymbol{\beta}}_c$ kısıtlanmış en çok olabilirlik tahmin edicisidir. Şimdi σ^2 için elde edilen $\tilde{\sigma}_c^2$ tahmin edicisini (3.17) log olabilirlik denkleminde yerine koyalım ve

$$\begin{aligned}
\ell = \ln L &= -\frac{n}{2} \ln 2\pi - \frac{n}{2} \ln \sigma^2 - \frac{(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})}{2\sigma^2} \\
&= -\frac{n}{2} \ln 2\pi - \frac{n}{2} \ln \left[\frac{1}{n} (\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}}_c)'(\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}}_c) \right] - \frac{(\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}}_c)'(\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}}_c)}{2 \left[\frac{1}{n} (\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}}_c)'(\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}}_c) \right]} \quad (4.34) \\
&= -\frac{n}{2} \ln 2\pi - \frac{n}{2} \ln \left[\frac{1}{n} (\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}}_c)'(\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}}_c) \right] - \frac{n}{2}
\end{aligned}$$

bağıntısını elde etmek için sadeleştirelim. Şimdi aşağıdaki ifadeye dikkat edelim:

$$\begin{aligned}
\ln L &= -\frac{n}{2} \ln 2\pi - \frac{n}{2} \ln \left[\frac{1}{n} (\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}}_c)'(\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}}_c) \right] - \frac{n}{2} \\
\Rightarrow L &= (2\pi)^{-\frac{n}{2}} \left[\frac{1}{n} (\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}}_c)'(\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}}_c) \right]^{-\frac{n}{2}} e^{-\frac{n}{2}}. \quad (4.35)
\end{aligned}$$

Kısıtlanmış tahmin ediciyi tanımlayan maksimumlaştırma problemi

$$\mathbf{R}\tilde{\boldsymbol{\beta}}_c = \mathbf{r} \quad (4.36)$$

altında

$$\max_{\tilde{\boldsymbol{\beta}}_c} \ell = -\frac{n}{2} \ln 2\pi - \frac{n}{2} \ln \left[\frac{1}{n} (\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}}_c)'(\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}}_c) \right] - \frac{n}{2} \quad (4.37)$$

olarak ifade edilebilir.

Belli ki, bu olabilirlik fonksiyonunu, $(\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}}_c)'(\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}}_c)$ hata kareler toplamını minimumlaştırmak suretiyle maksimumlaştırırız. Kısıtlamaya bağlı olarak bunu maksimumlaştırmak için, λ_2 Lagrange çarpanlarının bir $m \times 1$ vektörü olmak üzere,

$$\mathfrak{S}_2 = \mathbf{y}'\mathbf{y} - 2\tilde{\boldsymbol{\beta}}_c'\mathbf{X}\mathbf{y} + \tilde{\boldsymbol{\beta}}_c'\mathbf{X}\mathbf{X}\tilde{\boldsymbol{\beta}}_c - \lambda_2'(\mathbf{r} - \mathbf{R}\tilde{\boldsymbol{\beta}}_c) \quad (4.38)$$

Lagrange fonksiyonunu oluştururuz.

$\tilde{\boldsymbol{\beta}}_c$ ve λ_2 ya göre diferensiyelleme

$$\begin{aligned}
-2\mathbf{X}'\mathbf{y} + 2\mathbf{X}'\mathbf{X}\tilde{\boldsymbol{\beta}}_c + \mathbf{R}'\lambda_2 &= 0 \\
\mathbf{R}\tilde{\boldsymbol{\beta}}_c - \mathbf{r} &= 0 \quad (4.39)
\end{aligned}$$

şartlarını verir.

Şimdi,

$$-2\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} + 2\mathbf{R}\tilde{\boldsymbol{\beta}}_c + \mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'\lambda_2 = 0 \quad (4.40)$$

bağıntısını elde etmek için (4.39) daki birinci bağıntıyı $\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}$ ile önden çarpalım.

$\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}' > 0$ matrisi pozitif tanımlı olduğundan, $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$ yi kullanarak

$$\begin{aligned} -2\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} + 2\mathbf{R}\boldsymbol{\beta} + \mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'\lambda_2 &= 0 \\ \Rightarrow \mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'\lambda_2 &= 2\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} - 2\mathbf{R}\boldsymbol{\beta} \\ &= 2\mathbf{R}\hat{\boldsymbol{\beta}} - 2\mathbf{R}\boldsymbol{\beta} \\ \Rightarrow \lambda_2 &= (\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1} (2\mathbf{R}\hat{\boldsymbol{\beta}} - 2\mathbf{R}\boldsymbol{\beta}) \\ &= -2(\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1} (\mathbf{r} - \mathbf{R}\hat{\boldsymbol{\beta}}) \end{aligned} \quad (4.41)$$

elde ederiz.

λ_2 nin bu değerini (4.39) daki birinci bağıntıda yerine koyarsak, $\boldsymbol{\beta}$ nın kısıtlanmış en çok olabilirlik tahmin edicisi aşağıdaki gibi elde edilir:

$$\tilde{\boldsymbol{\beta}}_c = \hat{\boldsymbol{\beta}} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}' (\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1} (\mathbf{r} - \mathbf{R}\hat{\boldsymbol{\beta}}), \quad (4.42a)$$

$$\tilde{\boldsymbol{\beta}}_c = \hat{\boldsymbol{\beta}} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}' (\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1} (\mathbf{R}\hat{\boldsymbol{\beta}} - \mathbf{r}). \quad (4.42b)$$

(4.42a) ve (4.42b) ifadeleri denktir.

Böylece, modelde normal dağılılan hatalar ile, $\tilde{\boldsymbol{\beta}}_c = \hat{\boldsymbol{\beta}}_c$ ve $Var(\tilde{\boldsymbol{\beta}}_c) = Var(\hat{\boldsymbol{\beta}}_c)$ olduğu görülür. Bu nedenle, $Var(\tilde{\boldsymbol{\beta}}_c) \leq Var(\hat{\boldsymbol{\beta}})$ eşitsizliğini de elde ederiz.

(4.10a), (4.10b), (4.42a) ve (4.42b) ifadeleri aşağıdaki gibi yararlı biçimde yeniden yazılabilir:

$$(\mathbf{X}'\mathbf{X})(\hat{\boldsymbol{\beta}}_c - \hat{\boldsymbol{\beta}}) = (\mathbf{X}'\mathbf{X})(\tilde{\boldsymbol{\beta}}_c - \hat{\boldsymbol{\beta}}) = \mathbf{R}' (\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1} (\mathbf{r} - \mathbf{R}\hat{\boldsymbol{\beta}}). \quad (4.43)$$

(4.42a) bağıntısından

$$\tilde{\boldsymbol{\beta}}_c - \hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}' (\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1} (\mathbf{r} - \mathbf{R}\hat{\boldsymbol{\beta}}) \quad (4.44)$$

elde ederiz. Bu nedenle,

$$(\tilde{\boldsymbol{\beta}}_c - \hat{\boldsymbol{\beta}})' = (\mathbf{r} - \mathbf{R}\hat{\boldsymbol{\beta}})' (\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1} \mathbf{R}(\mathbf{X}'\mathbf{X})^{-1} \quad (4.45)$$

ve

$$(\tilde{\beta}_c - \hat{\beta})'(\mathbf{X}'\mathbf{X}) = (\mathbf{r} - \mathbf{R}\hat{\beta})'(\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1}\mathbf{R} \quad (4.46)$$

ifadelerini elde ederiz. (4.44) ü ve (4.46) yı kullanarak,

$$\begin{aligned} (\tilde{\beta}_c - \hat{\beta})'(\mathbf{X}'\mathbf{X})(\tilde{\beta}_c - \hat{\beta}) &= (\mathbf{r} - \mathbf{R}\hat{\beta})'(\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1}\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'(\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1}(\mathbf{r} - \mathbf{R}\hat{\beta}) \\ &= (\mathbf{r} - \mathbf{R}\hat{\beta})'(\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1}(\mathbf{r} - \mathbf{R}\hat{\beta}) \end{aligned} \quad (4.47)$$

veya

$$(\hat{\beta}_c - \hat{\beta})'(\mathbf{X}'\mathbf{X})(\hat{\beta}_c - \hat{\beta}) = (\mathbf{r} - \mathbf{R}\hat{\beta})'(\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1}(\mathbf{r} - \mathbf{R}\hat{\beta}) \quad (4.48)$$

olduğu görülür.

(4.42a) denkleminin her iki yanını \mathbf{X} ile çarparak ve sonra da her iki yanını \mathbf{y} den çıkararak (4.42a) denklemini başka bir yararlı biçimde yeniden düzenleyebiliriz. Bunu yaparak,

$$\begin{aligned} \mathbf{X}\tilde{\beta}_c &= \mathbf{X}\hat{\beta} + \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'(\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1}(\mathbf{r} - \mathbf{R}\hat{\beta}) \\ \Rightarrow \mathbf{y} - \mathbf{X}\tilde{\beta}_c &= \mathbf{y} - \mathbf{X}\hat{\beta} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'(\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1}(\mathbf{r} - \mathbf{R}\hat{\beta}) \\ \Rightarrow e_c &= e - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'(\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1}(\mathbf{r} - \mathbf{R}\hat{\beta}) \end{aligned} \quad (4.49)$$

bağıntısını elde ederiz. Burada e_c , kısıtlanmış regresyondan tahmin edilen hatadır. Kısıtlanmış regresyondan tahmin edilen hataların kareleri toplamı olan $e_c'e_c$ yi göz önüne alalım. $e = \mathbf{y} - \mathbf{X}\hat{\beta}$ kısıtlanmamış modelden tahmin edilen hata vektörü olmak üzere,

$$\begin{aligned} e_c'e_c &= \left(e' - (\mathbf{r} - \mathbf{R}\hat{\beta})'(\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1}\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' \right) \left(e - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'(\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1}(\mathbf{r} - \mathbf{R}\hat{\beta}) \right) \\ &= e'e - e'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'(\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1}(\mathbf{r} - \mathbf{R}\hat{\beta}) - (\mathbf{r} - \mathbf{R}\hat{\beta})'(\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1}\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'e \\ &\quad + (\mathbf{r} - \mathbf{R}\hat{\beta})'(\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1}\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}\mathbf{X}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'(\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1}(\mathbf{r} - \mathbf{R}\hat{\beta}) \\ &= e'e - e'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'(\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1}(\mathbf{r} - \mathbf{R}\hat{\beta}) - (\mathbf{r} - \mathbf{R}\hat{\beta})'(\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1}\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'e \\ &\quad + (\mathbf{r} - \mathbf{R}\hat{\beta})'(\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1}(\mathbf{r} - \mathbf{R}\hat{\beta}) \end{aligned} \quad (4.50)$$

dır. Şimdi, (3.12) bağıntısını aşağıdaki gibi yeniden yazmak suretiyle, görülebildiği gibi alışılmış en küçük karelerde $\mathbf{X}'\mathbf{e} = 0$ olduğunu hatırlayalım.

$$\begin{aligned} 2\mathbf{X}'\mathbf{y} - 2\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} &= 0 \\ \mathbf{X}'(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}) &= \mathbf{X}'\mathbf{e} = 0 \end{aligned} \quad (4.51)$$

elde edilir. Bu bilgiyi (4.50) bağıntısında kullanarak,

$$\begin{aligned} e'_c e_c &= e'e - e'\mathbf{X}(\mathbf{X}\mathbf{X})^{-1}\mathbf{R}'(\mathbf{R}(\mathbf{X}\mathbf{X})^{-1}\mathbf{R}')^{-1}(\mathbf{r} - \mathbf{R}\hat{\boldsymbol{\beta}}) - (\mathbf{r} - \mathbf{R}\hat{\boldsymbol{\beta}})'(\mathbf{R}(\mathbf{X}\mathbf{X})^{-1}\mathbf{R}')^{-1}\mathbf{R}(\mathbf{X}\mathbf{X})^{-1}\mathbf{X}'e \\ &\quad + (\mathbf{r} - \mathbf{R}\hat{\boldsymbol{\beta}})'(\mathbf{R}(\mathbf{X}\mathbf{X})^{-1}\mathbf{R}')^{-1}(\mathbf{r} - \mathbf{R}\hat{\boldsymbol{\beta}}) \\ &= e'e + (\mathbf{r} - \mathbf{R}\hat{\boldsymbol{\beta}})'(\mathbf{R}(\mathbf{X}\mathbf{X})^{-1}\mathbf{R}')^{-1}(\mathbf{r} - \mathbf{R}\hat{\boldsymbol{\beta}}) \\ \Rightarrow e'_c e_c - e'e &= (\mathbf{r} - \mathbf{R}\hat{\boldsymbol{\beta}})'(\mathbf{R}(\mathbf{X}\mathbf{X})^{-1}\mathbf{R}')^{-1}(\mathbf{r} - \mathbf{R}\hat{\boldsymbol{\beta}}) \end{aligned} \quad (4.52)$$

bağıntısını elde ederiz.

Bu nedenle, kısıtlanmış ve kısıtlanmamış modellerde tahmin edilen hata kareler toplamındaki fark; $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$ kısıtlanmamış alışlagelen en küçük kareler tahmini olmak üzere, $\mathbf{R}\hat{\boldsymbol{\beta}}$ ve \mathbf{r} arasındaki farka göre bir kuadratik form olarak yazılabilir.

(4.42a) bağıntısı yine, kısıtlanmış tahmin edicinin varyansını bulmada yararlı olacak başka bir tarzda yeniden düzenlenebilir. İlk olarak, alışılmış en küçük kareler tahmin edicisini $\boldsymbol{\beta}$ nın ve $\boldsymbol{\varepsilon}$ un bir fonksiyonu olarak aşağıdaki gibi yazalım:

$$\begin{aligned} \hat{\boldsymbol{\beta}} &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} \\ &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}) \\ &= \boldsymbol{\beta} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon}. \end{aligned} \quad (4.53)$$

Bu takdirde (4.42a) bağıntısında $\hat{\boldsymbol{\beta}}$ için bu ifadeyi aşağıdaki gibi yerine koyalım.

$$\begin{aligned} \tilde{\boldsymbol{\beta}}_c &= \hat{\boldsymbol{\beta}} + (\mathbf{X}\mathbf{X})^{-1}\mathbf{R}'(\mathbf{R}(\mathbf{X}\mathbf{X})^{-1}\mathbf{R}')^{-1}(\mathbf{r} - \mathbf{R}\hat{\boldsymbol{\beta}}) \\ &= \boldsymbol{\beta} + (\mathbf{X}\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon} + (\mathbf{X}\mathbf{X})^{-1}\mathbf{R}'(\mathbf{R}(\mathbf{X}\mathbf{X})^{-1}\mathbf{R}')^{-1}(\mathbf{r} - \mathbf{R}(\boldsymbol{\beta} + (\mathbf{X}\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon})) \\ &= \boldsymbol{\beta} + (\mathbf{X}\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon} + (\mathbf{X}\mathbf{X})^{-1}\mathbf{R}'(\mathbf{R}(\mathbf{X}\mathbf{X})^{-1}\mathbf{R}')^{-1}(\mathbf{r} - \mathbf{R}\boldsymbol{\beta} - \mathbf{R}(\mathbf{X}\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon}) \\ \Rightarrow \tilde{\boldsymbol{\beta}}_c - \boldsymbol{\beta} &= (\mathbf{X}\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon} + (\mathbf{X}\mathbf{X})^{-1}\mathbf{R}'(\mathbf{R}(\mathbf{X}\mathbf{X})^{-1}\mathbf{R}')^{-1}(-\mathbf{R}(\mathbf{X}\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon}), \quad \mathbf{r} = \mathbf{R}\boldsymbol{\beta} \\ \Rightarrow \tilde{\boldsymbol{\beta}}_c - \boldsymbol{\beta} &= (\mathbf{X}\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon} - (\mathbf{X}\mathbf{X})^{-1}\mathbf{R}'(\mathbf{R}(\mathbf{X}\mathbf{X})^{-1}\mathbf{R}')^{-1}\mathbf{R}(\mathbf{X}\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon} \end{aligned} \quad (4.54)$$

$\mathbf{M}^c = \left[\mathbf{I} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'(\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1}\mathbf{R} \right]$ matrisini daha önce tanımlamıştık.

$\tilde{\beta}_c - \beta$ yi,

$$\begin{aligned}\tilde{\beta}_c - \beta &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'(\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1}\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon} \\ &= \left[\mathbf{I} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'(\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1}\mathbf{R} \right] (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon} \\ &= \mathbf{M}^c (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon} \\ &= \mathbf{M}^c \left[\hat{\beta} - \beta \right]\end{aligned}\quad (4.55)$$

olarak yazabiliriz (Chow, 1960; Fisher, 1970).

4.1.1.4. Tahmin Edilen Hataları Kullanarak Model Üzerindeki Kısıtlamaları Test Etme

$$\frac{(\mathbf{r} - \mathbf{R}\hat{\beta})' \left[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}' \right]^{-1} (\mathbf{r} - \mathbf{R}\hat{\beta})}{\frac{\frac{m}{e'e}}{n-p}} \sim F(m, n-p) \quad (4.56)$$

dağılımını daha önceki bilgilerimizi kullanarak gösterelim. Pay için,

$$\begin{aligned}\mathbf{r} - \mathbf{R}\hat{\beta} &= \mathbf{r} - \mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = \mathbf{R}\beta - \mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} \\ &= \mathbf{R}(\beta - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}) = -\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathbf{y} - \mathbf{X}\beta) \\ (\mathbf{r} - \mathbf{R}\hat{\beta})' &= -(\mathbf{y} - \mathbf{X}\beta)' \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'\end{aligned}$$

$$(\mathbf{r} - \mathbf{R}\hat{\beta})' \left[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}' \right]^{-1} (\mathbf{r} - \mathbf{R}\hat{\beta}) = (\mathbf{y} - \mathbf{X}\beta)' \left[\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'(\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1}\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' \right] (\mathbf{y} - \mathbf{X}\beta) \quad (4.57)$$

olduğunu görürüz.

$\mathbf{A} = \left[\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'(\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1}\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' \right]$ olsun. \mathbf{A} simetrik bir matristir. O

halde, $(\mathbf{r} - \mathbf{R}\hat{\beta})' \left[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}' \right]^{-1} (\mathbf{r} - \mathbf{R}\hat{\beta}) = (\mathbf{y} - \mathbf{X}\beta)' \mathbf{A} (\mathbf{y} - \mathbf{X}\beta)$, $(\mathbf{y} - \mathbf{X}\beta)$ ya göre bir karesel formdur.

Payda için,

$$(\mathbf{y} - \mathbf{X}\hat{\beta})' (\mathbf{y} - \mathbf{X}\hat{\beta}) = \mathbf{y}' [\mathbf{I} - \mathbf{P}_X] \mathbf{y} = (\mathbf{y} - \mathbf{X}\beta)' [\mathbf{I} - \mathbf{P}_X] (\mathbf{y} - \mathbf{X}\beta) \quad (4.58)$$

dır. Burada $\mathbf{P}_X = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$ dür.

$\mathbf{B} = [\mathbf{I} - \mathbf{P}_X] = \mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$ olsun. \mathbf{B} de simetrik bir matristir. $(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})' \mathbf{B}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$ de $(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$ ya göre bir karesel formdur. $V = \sigma^2 \mathbf{I}$ olduğundan ve

$$\begin{aligned} \mathbf{A}\mathbf{V}\mathbf{B} &= \left[\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'(\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1}\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' \right] \sigma^2 \mathbf{I} \left[\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' \right] \\ &= \sigma^2 \left[\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'(\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1}\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' \right. \\ &\quad \left. - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'(\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1}\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' \right] \\ &= \mathbf{0} \end{aligned} \quad (4.59)$$

olduğundan bu iki karesel formların dağılımları bağımsızdır. $(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})' \mathbf{A}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) \sim \chi_{m,0}^2$ ve $(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})' \mathbf{B}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) \sim \chi_{n-p,0}^2$ olduğundan (4.56) daki ifadenin doğru olduğu ispatlanmış olur.

(4.56) bağıntısındaki payı göz önüne alalım. Bu, (4.52) bağıntısını kullanarak kısıtlanmış ve kısıtlanmamış modellerden tahmin edilen hatalara göre,

$$\begin{aligned} e_c' e_c &= e' e + (\mathbf{r} - \mathbf{R}\hat{\boldsymbol{\beta}})' (\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1} (\mathbf{r} - \mathbf{R}\hat{\boldsymbol{\beta}}) \\ \Rightarrow e_c' e_c - e' e &= (\mathbf{r} - \mathbf{R}\hat{\boldsymbol{\beta}})' (\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1} (\mathbf{r} - \mathbf{R}\hat{\boldsymbol{\beta}}) \\ \Rightarrow \frac{e_c' e_c - e' e}{n-p} &\sim F(m, n-p) \end{aligned} \quad (4.60)$$

yazılabilir. Bir özel modelden tahmin edilen hata kareler toplamını $SSE(\hat{\boldsymbol{\beta}})$ ile göstererek,

$$\frac{SSE(\hat{\boldsymbol{\beta}}_c) - SSE(\hat{\boldsymbol{\beta}})}{n-p} \sim F(m, n-p) \quad (4.61)$$

bağıntısını elde ederiz.

Hipotez testini $[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}']$ matrisinin tersini alarak ve sonra da $\mathbf{r} - \mathbf{R}\hat{\boldsymbol{\beta}}$ ile ön ve son çarpımları yaparak yürütmekten ziyade, basit olarak farklı iki regresyonu çalıştırır ve kısıtlanmış ve kısıtlanmamış hata tahminlerinden F istatistiğini hesaplarız.

(4.61) bağıntısındaki test istatistiği biçimine amaç fonksiyonundaki değişime dayanan bir test istatistiği olarak başvurulur. Bir takım testler bu kategoriye düşer. Buradaki düşünce kısıtlanmış ve kısıtlanmamış modellerdeki kareler toplamlarını karşılaştırmaktır. Eğer kısıtlama kısıtlanmamışinkinden anlamlı derecede daha büyük olacak SSE ye sebep oluyorsa bu, verinin kısıtlamayı sağlamadığının kanıtıdır ve kısıtlamanın geçerli olduğu hipotezini reddederiz. Böyle testler için genel yöntem, iki regresyonu aşağıdaki gibi çalıştırmaktadır.

1) β vektörü üzerine hiçbir kısıtlama koymaksızın regresyon modelini tahmin ederiz. İlgili hata kareler toplamı ve hata kareler toplamının serbestlik derecesi sırası ile $SSE(\hat{\beta})$ ve $(n - p)$ ile gösterilmiş olsun.

2) β hipotezle belirtildiği gibi sınırlanmış olmak üzere, aynı regresyon modelini tahmin edelim. İlgili hata kareler toplamı ve hata kareler toplamının serbestlik derecesi sırası ile $SSE(\hat{\beta}_c)$ ve $m + (n - p) = (n - p)_c$ ile gösterilmiş olsun.

3) $m = (n - p)_c - (n - p)$ olmak üzere, hipotez ile β üzerine koyulan bağımsız kısıtlamaların sayısı olmak üzere, testi aşağıdaki

$$\frac{\frac{SSE(\hat{\beta}_c) - SSE(\hat{\beta})}{m}}{\frac{SSE(\hat{\beta})}{n - p}} \sim F(m, n - p)$$

istatistiğini kullanarak yürütürüz. Örneğin, (3.29) modelinde eğer hipotez $H_0 : \beta_2 + \beta_3 = 4, \beta_4 = 0$ olsaydı, bu takdirde payın serbestlik derecesi 2 ye eşittir. Eğer hipotez geçerli ise bu takdirde $SSE(\hat{\beta}_c)$ ve $SSE(\hat{\beta})$ her biri diğerinden anlamlı derecede farklı olmamalıdır. Bu nedenle, eğer F değeri büyükse kısıtlamayı reddederiz. Testin bu tipi üzerinde iki yararlı kaynak Chow (1960) ve Fisher (1970) dir.

4.1.1.5. Bir Olabilirlik Oran (LR) Testi Kullanarak Model Üzerindeki Kısıtlamaları Test Etme

a) Fikir ve Tahmin

Olabilirlik oran (LR) testi klasik istatistikte istatistiksel çıkarımın yaygın bir yöntemidir. LR test istatistiği, kısıtlanmış ve kısıtlanmamış olabilirlik fonksiyonlarının kıyaslanması vasıtasıyla verinin bir örnekleme ve sıfır hipotezi arasındaki uygunluğunu yansıtır. LR , tahmin işleminde θ parametresi üzerinde varsayılan sınırlamaları koymanın bir sonucu olarak, olabilirliğin değerinde (veya log-olabilirlik değerinde) anlamlı bir azalmanın var olup olmadığını belirlemeye

dayanır. Usulen (x_1, x_2, \dots, x_n) rasgele örnekleme $f(x_1, x_2, \dots, x_n; \theta)$ ortak olasılık fonksiyonuna sahip olsun ve ilgili olabilirlik fonksiyonu $L(\theta; x_1, x_2, \dots, x_n)$ olsun. Genelleştirilmiş olabilirlik oranı (*GLR*), “sup” sıfır hipotezini (H_0) sağlayan parametrelerin kümesi üzerinden fonksiyonun supremumunu veya sıfır veya alternatif hipotezi (H_a) sağlaması gereken parametrelerin kümesi üzerinden fonksiyonun supremumunu göstermek üzere,

$$\lambda(x_1, x_2, \dots, x_n) = \left(\frac{\sup_{\theta \in H_0} L(\theta; x_1, x_2, \dots, x_n)}{\sup_{\theta \in H_0 \cup H_a} L(\theta; x_1, x_2, \dots, x_n)} \right) \quad (4.62)$$

olarak tanımlanır. H_a alternatif hipotezine karşı H_0 ı test etmek için bir genelleştirilmiş olabilirlik oran testi aşağıdaki kuralla verilir. c yine belirlenecek bir dağılımdan kritik değer olmak üzere, H_0 ın reddedilmesi için gerek ve yeter şart aşağıdaki

$$\lambda(x_1, x_2, \dots, x_n) \leq c \quad (4.63)$$

kuralı ile verilir. $\pi(\theta)$ istatistiksel testin güç fonksiyonu olmak üzere, bir α testi için c sabiti

$$\sup_{\theta \in H_0} \pi(\theta) = \sup_{\theta \in H_0} P(\lambda(x_1, x_2, \dots, x_n) \leq c; \theta) = \alpha \quad (4.64)$$

bağıntısını sağlayacak şekilde seçilir.

Orana göre iki olabilirlik fonksiyonunun farklı bölgeler üzerinden θ ya göre “en uygun” olduğunu göstermek için olabilirlik oranını karşılaştırırken “genelleştirilmiş” olabilirlik oranı terimini kullanırız. Yazılı kaynaklar genellikle buna, genelleştirilen dönüştürücü olmaksızın bir olabilirlik oranı olarak başvurur ve genellikle o geleneği izler (Chow, 1960; Fisher, 1970).

b) Klasik Normal Lineer Regresyon Modeli İçin Olabilirlik Oranı

Klasik normal lineer regresyon modelinde $\mathbf{r} = \mathbf{R}\boldsymbol{\beta}$ sıfır hipotezini göz önüne alalım. (4.35) bağıntısından, kısıtlanmış en küçük kareler tahminlerinde değerlendirilen olabilirlik fonksiyonu

$$L^{c*} = (2\pi)^{-\frac{n}{2}} \left[\frac{1}{n} (\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}}_c)' (\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}}_c) \right]^{-\frac{n}{2}} e^{-\frac{n}{2}} \quad (4.65)$$

dir. Benzer şekilde, *OLS* tahminlerinde değerlendirilen olabilirlik fonksiyonunu

$$L^* = (2\pi)^{-\frac{n}{2}} \left[\frac{1}{n} (\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})' (\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}) \right]^{\frac{n}{2}} e^{-\frac{n}{2}} \quad (4.66)$$

olarak yazabiliriz.

Bu takdirde genelleştirilmiş olabilirlik oran istatistiği

$$\begin{aligned} \lambda(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n) &= \frac{\left((2\pi)^{-\frac{n}{2}} \left[\frac{1}{n} (\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}}_c)' (\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}}_c) \right]^{\frac{n}{2}} e^{-\frac{n}{2}} \right)}{\left((2\pi)^{-\frac{n}{2}} \left[\frac{1}{n} (\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})' (\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}) \right]^{\frac{n}{2}} e^{-\frac{n}{2}} \right)} \\ &= \frac{\left[(\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}}_c)' (\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}}_c) \right]^{\frac{n}{2}}}{\left[(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})' (\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}) \right]^{\frac{n}{2}}} \\ &= \left(\frac{(\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}}_c)' (\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}}_c)}{(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})' (\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})} \right)^{\frac{n}{2}} \end{aligned} \quad (4.67)$$

dir. λ nın küçük değerleri veya (4.67) nin sağ yanının küçük değerleri için sıfır hipotezini reddederiz. Zorluk, (4.67) bağıntısının sağ yanının dağılımını bilmediğimizdir. Onu, tahmin edilen hatalar cinsinden

$$\begin{aligned} \lambda &= \left(\frac{(\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}}_c)' (\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}}_c)}{(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})' (\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})} \right)^{\frac{n}{2}} \\ &= \left(\frac{\boldsymbol{\delta}'_c \boldsymbol{\delta}_c}{\mathbf{e}' \mathbf{e}} \right)^{\frac{n}{2}} \end{aligned} \quad (4.68)$$

olarak yazabildiğimize dikkat edelim. Bu takdirde bu,

$$\lambda = \left(\frac{\boldsymbol{\delta}'_c \boldsymbol{\delta}_c}{\mathbf{e}' \mathbf{e}} \right)^{\frac{n}{2}} = \left(\frac{SSE(\tilde{\boldsymbol{\beta}}_c)}{SSE(\hat{\boldsymbol{\beta}})} \right)^{\frac{n}{2}} \quad (4.69)$$

olarak yazılabilir. Burada $\boldsymbol{\delta}_c = \mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}}_c$ tahmin edilen hatayı verir. Bu nedenle eğer,

$$\begin{aligned} \lambda &= \left(\frac{SSE(\tilde{\boldsymbol{\beta}}_c)}{SSE(\hat{\boldsymbol{\beta}})} \right)^{\frac{n}{2}} \leq c \\ \Rightarrow \lambda^{\frac{2}{n}} &= \left(\frac{SSE(\tilde{\boldsymbol{\beta}}_c)}{SSE(\hat{\boldsymbol{\beta}})} \right) \geq c^{\frac{2}{n}} \end{aligned} \quad (4.70)$$

ise; $\mathbf{r} = \mathbf{R}\boldsymbol{\beta}$ sıfır hipotezini reddederiz.

Şimdi $\frac{SSE(\hat{\boldsymbol{\beta}})}{SSE(\hat{\boldsymbol{\beta}})} = 1$ ifadesini, (4.70) bağıntısının her iki yanından çıkaralım ve sadeleştirelim. Böylece,

$$\begin{aligned}\lambda^{\frac{-2}{n}} &= \frac{SSE(\tilde{\boldsymbol{\beta}}_c)}{SSE(\hat{\boldsymbol{\beta}})} \geq c^{\frac{-2}{n}} \\ \Rightarrow \lambda^{\frac{-2}{n}} - \frac{SSE(\hat{\boldsymbol{\beta}})}{SSE(\hat{\boldsymbol{\beta}})} &= \frac{SSE(\tilde{\boldsymbol{\beta}}_c)}{SSE(\hat{\boldsymbol{\beta}})} - \frac{SSE(\hat{\boldsymbol{\beta}})}{SSE(\hat{\boldsymbol{\beta}})} \geq c^{\frac{-2}{n}} - \frac{SSE(\hat{\boldsymbol{\beta}})}{SSE(\hat{\boldsymbol{\beta}})} \\ \Rightarrow \lambda^{\frac{-2}{n}} - 1 &= \frac{SSE(\tilde{\boldsymbol{\beta}}_c) - SSE(\hat{\boldsymbol{\beta}})}{SSE(\hat{\boldsymbol{\beta}})} \geq c^{\frac{-2}{n}} - 1\end{aligned}\quad (4.71)$$

elde ederiz.

(4.71) bağıntısının her iki yanını $\frac{n-p}{m}$ ile çarparak,

$$\begin{aligned}\lambda^{\frac{-2}{n}} - 1 &= \frac{SSE(\tilde{\boldsymbol{\beta}}_c) - SSE(\hat{\boldsymbol{\beta}})}{SSE(\hat{\boldsymbol{\beta}})} \geq c^{\frac{-2}{n}} - 1 \\ \frac{n-p}{m} \left(\lambda^{\frac{-2}{n}} - 1 \right) &= \frac{n-p}{m} \left(\frac{SSE(\tilde{\boldsymbol{\beta}}_c) - SSE(\hat{\boldsymbol{\beta}})}{SSE(\hat{\boldsymbol{\beta}})} \right) \geq \frac{n-p}{m} \left(c^{\frac{-2}{n}} - 1 \right)\end{aligned}\quad (4.72)$$

bağıntısı elde edilir. Eğer (4.72) bağıntısındaki değer herhangi keyfi $\mathbf{b} = \left(\frac{n-p}{m} \right) \left(c^{\frac{-2}{n}} - 1 \right)$ değerinden daha büyük ise, sıfır hipotezini reddederiz. Bu takdirde sorun, (4.72) bağıntısındaki test istatistiğinin dağılımını bulmadır. Test istatistiğini,

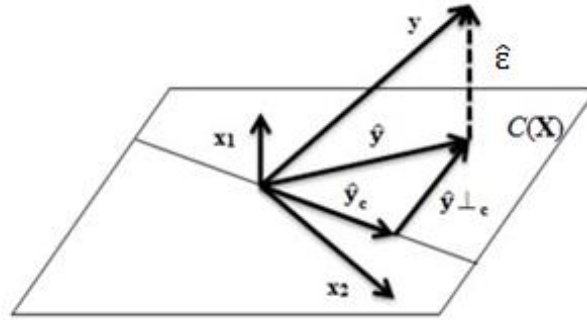
$$\frac{n-p}{m} \left(\lambda^{\frac{-2}{n}} - 1 \right) = \frac{\frac{SSE(\tilde{\boldsymbol{\beta}}_c) - SSE(\hat{\boldsymbol{\beta}})}{m}}{\frac{SSE(\hat{\boldsymbol{\beta}})}{n-p}}\quad (4.73)$$

olarak da yazabiliriz. (4.73) deki payın, (4.56) ve (4.61) deki pay ile aynı olduğunu daha evvel gösterdik. Bu nedenle, bu istatistik o istatistiklere eşdeğerdir ve bir F olarak dağılır. Özellikle,

$$\frac{n-k}{m} \left(\lambda^{\frac{-2}{n}} - 1 \right) = \frac{\frac{SSE(\tilde{\beta}_c) - SSE(\hat{\beta})}{m}}{\frac{SSE(\hat{\beta})}{n-k}} \sim F(m, n-k) \quad (4.74)$$

dır. Böylece, klasik normal lineer regresyon modelinde (4.1) lineer kısıtlamalarının bir kümesi için olabilirlik oran testi ve F testi eşdeğerdir (Chow, 1960; Fisher, 1970).

Şimdi Çizelge 3.1 deki örneği kullanarak, \mathbf{y} nin ikinci uyarıcı tarafından etkilenmediğinin bir testi \mathbf{x}_2 nin β_2 katsayısını sıfır olmasını sağlayan, $\mathbf{R} = [0 \ 0 \ 1]$ matrisiyle verilir. Tepkilerin iki uyarıcı için aynı olup olmadığının bir testi β_1 ve β_2 yi eşit yapan, $\mathbf{R} = [0 \ 1 \ -1]$ matrisiyle verilir. Birleşik hipotezler mümkün olabilir: Modelin her iki uyarıcı parametrelerini sıfıra eşitleyen bir test $\mathbf{R} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ den elde edilebilir. Geometrik olarak böyle herhangi bir lineer hipotez, $\hat{\mathbf{y}}$ vektörünü, biri hipotezle tutarlı olan $\hat{\mathbf{y}}_c$ vektörü ve diğeri ona dik olan bir $\hat{\mathbf{y}}_{\perp c}$ vektörü olmak üzere, iki ortogonal parçaya ayırarak test edilir.



Şekil 4.1. $\hat{\mathbf{y}}$ nin iki ortogonal $\hat{\mathbf{y}}_c$ ve $\hat{\mathbf{y}}_{\perp c}$ parçalarına ayrışımı.

Eğer $\mathbf{R}\boldsymbol{\beta} = \mathbf{0}$ ise, bu takdirde $\hat{\mathbf{y}}_{\perp c}$ tümüyle hataya yüklenebilir ve onun büyüklüğü $\hat{\boldsymbol{\varepsilon}}$ nun büyüklüğü ile karşılaştırılmalıdır. Eğer $\mathbf{R}\boldsymbol{\beta} \neq \mathbf{0}$ ise, bu takdirde $\hat{\mathbf{y}}_{\perp c}$ nin her-boyut uzunluğu $\hat{\boldsymbol{\varepsilon}}$ nunkinden çok daha büyüktür.

Bir Gaussian hata dağılımlı genel lineer model altında tümü karşılaştırılabilen bu hipotezler için, bir test istatistiğini elde etmek için üç yöntem vardır:

1. Geniş kapsamlı test için ifade edildiği gibi bir F istatistiğini kullanarak, $\hat{\mathbf{y}}_{\perp c}$ ve $\hat{\boldsymbol{\varepsilon}}$ vektörlerinin uzunluklarının kareleri ortalamasını karşılaştırarak test edilir. Eğer m tane lineer kısıtlama (\mathbf{R} nin satırları) varsa, bu takdirde bu istatistik,

$$F = \frac{|\hat{\mathbf{y}}_{\perp c}|^2 / m}{|\hat{\boldsymbol{\varepsilon}}|^2 / (n-p)} = \frac{\text{Kısıtlı Kare Ortalaması}}{\text{Hata Kare Ortalaması}} = \frac{KKO}{HKO} \quad (4.75)$$

dur.

2. Hipotez tek boyutlu ise, yani hipotez sadece bir kısıtlamayı ihtiva ederse, bu takdirde, onun uyumsuzluğunun gözlenen büyüklüğünü ve bu niceliğin standart hatasını, yani

$$\mathbf{V} = \mathbf{R}\hat{\boldsymbol{\beta}} \quad \text{ve} \quad \mathbf{s}_v^2 = \mathbf{R}\mathbf{S}_{\hat{\boldsymbol{\beta}}}\mathbf{R}' \quad (4.76)$$

yü hesaplayarak test edilebilir. Hipotez yukarıdaki F nin karekökü olan bir t istatistiğiyle, yani $t = \mathbf{V} / \mathbf{s}_v$ istatistiğiyle ile test edilir.

3. Biri test edilen parametreleri içeren, biri içermeyen iki modeli uydurarak test edilebilir. Eğer hipotez sağlanmazsa, bu takdirde hipotezin konulmadığı kısıtlanmamış model azıcık daha küçük bir hataya sahip olmakla beraber bu iki modelin hatası aynı büyüklük civarında olacaktır. Kısıtlamanın konulmasıyla hata artış miktarı

$$F = \frac{\left(|\boldsymbol{\varepsilon}_{\text{kısıtlanmamış}}|^2 - |\boldsymbol{\varepsilon}_{\text{kısıtlanmış}}|^2 \right) / m}{|\boldsymbol{\varepsilon}_{\text{kısıtlanmamış}}|^2 / (n-p)} \quad (4.77)$$

istatistiği vasıtası ile genel hatayla karşılaştırılır (Wickens, 2004).

4.1.2. Lineer Modellerde Stokastik Lineer Kısıtlamalar Altında Parametre Tahmini ve Hipotez Testi

Kesin lineer kısıtlamalar yardımcı veya ön bilgide içerilen rasgeleliğin olmadığını kabul eder. Bu varsayım birçok pratik durumda doğru olmayabilir ve bir takım rasgelelik mevcut olabilir. Böyle durumlarda ön bilgi (4.2) olarak formüllenebilir. Burada \mathbf{v} rasgele hataların bir $m \times 1$ vektörüdür. \mathbf{r} ve \mathbf{R} deki elemanlar bilinir. \mathbf{v} terimi $\mathbf{r} = \mathbf{R}\boldsymbol{\beta}$ ön bilgisinde içerilen rasgeleliği yansıtır. $\boldsymbol{\Psi}$ bilinen bir $m \times m$ pozitif tanımlı matris ve $\boldsymbol{\varepsilon}$; $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ modelindeki hata terimi olmak üzere,

$$\begin{aligned} E(\mathbf{v}) &= \mathbf{0} \\ E(\mathbf{v}\mathbf{v}') &= \boldsymbol{\Psi} \end{aligned}$$

ve

$$E(\boldsymbol{\varepsilon}\mathbf{v}') = \mathbf{0} \quad (4.78)$$

olduğunu kabul edelim. $E(\mathbf{r}) = \mathbf{R}\boldsymbol{\beta}$ olduğuna dikkat edelim.

Böyle lineer stokastik kısıtlama için olası nedenler aşağıdaki gibidir:

1. Stokastik lineer kısıtlamalar; tahminlerin istikrarsızlığını ortaya koyar. Standart hatalı bir yansız tahmin; istikrarı (kararlılığı) ortaya koyar. Örneğin, tekrarlı incelemelerde (çalışmalarda) araştırmalar her yıl yürütülür. β_1 regresyon katsayısının, birkaç yıl için değişmediğini farz edelim. Onun tahmininin, standart hatası ile beraber tahmin edildiğini farz edelim. Onun değerinin 2 standart hata ile 0.5 civarında değişmediğini kabul edelim. Bu bilgi, $r=0.5$, $E(\mathbf{v}_1)=\mathbf{0}$, $E(\mathbf{v}_1^2)=2^2$ olmak üzere,

$$\mathbf{r} = \beta_1 + \mathbf{v}_1$$

olarak ifade edilebilir. Şimdi Ψ bu veri ile formüllenebilir. Tam regresyon katsayıları için bilgiye sahip olmamız gerekmez fakat sadece regresyon katsayılarının bazısı hakkında bilgiye sahip olabiliriz.

2. Bazen kısıtlamalar eşitsizlik biçimindedir. Böyle kısıtlamalar teorik düşüncelerden meydana gelebilir. Örneğin,

$$\mathbf{y} = \beta_0 + \beta_1 \mathbf{x} + \varepsilon$$

basit lineer regresyon modelini göz önüne alalım. Burada \mathbf{y} yiyecek üzerindeki tüketim harcamasını ve \mathbf{x} geliri gösterir. Bu takdirde tüketim yapmak için marjinal eğilim

$$\frac{dy}{dx} = \beta_1$$

dir, yani, eğer maaş 1 TL kadar artarsa, bu takdirde birinin yiyecek üzerinde 1 TL lik miktarı, yani β_1 i, harcaması veya $(1-\beta_1)$ lik miktarı tasarruf etmesi beklenir. β_1 üzerine ya bir TL nin tümünü harcamadığı ya da bir TL den hiç harcama yapmadığı bir sınırı konabilir. Bu nedenle, $0 < \beta_1 < 1$ dir. Bu, teorik düşünceden ortaya çıkan bir doğal kısıtlamadır.

Bu sınırlar p -sigma limitleri, diyelim ki 2-sigma limitleri veya güven aralıkları olarak işleme sokulabilir. Bu nedenle,

$$\begin{aligned} \mu - 2\sigma &= 0 \\ \mu + 2\sigma &= 1 \\ \Rightarrow \mu &= \frac{1}{2}, \sigma = \frac{1}{4} \end{aligned}$$

dür. Bu değerler

$$\beta_1 + v_1 = \frac{1}{2}$$

$$E(v_1^2) = \frac{1}{16}$$

olarak yorumlanabilir.

3. Bazen $\mathbf{r} = \mathbf{R}\beta$ kesin lineer kısıtlamasının doğruluğundan şüphe edilebilir ve bu nedenle bir belirsizlik unsuru öne sürülebilir. Örneğin, kısıtlamaların %95'inin gerçekleştiği söylenebilir. Bu nedenle, herhangi bir belirsizlik unsuru geçerli olur (Srivastava, 2018).

4.1.2.1. Sade (Pure-Saf) ve Karma Regresyon Tahmini

Kısım 3.1.2'deki n -gözlemlili $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_p$ gibi p -açıklayıcı değişkene sahip (3.29) modelini göz önüne alalım. β nın alışımlı en küçük kareler tahmin edicisi $\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$, sade tahmin edici olarak adlandırılır. $\hat{\beta}$ sade tahmin edicisi (4.2) kısıtlamasını sağlamaz. Bu nedenle amaç; sonuçta oluşan tahmin edici stokastik kısıtlamaları da sağlayacak şekilde, stokastik kısıtlamaları kullanarak β nın bir tahminini elde etmektir. Ön bilgi ve örneklem bilgisi arasındaki uyumsuzluktan kaçınmak için onları aşağıdaki gibi birleştirebiliriz:

$$\mathbf{y} = \mathbf{X}\beta + \boldsymbol{\varepsilon}, \quad E(\boldsymbol{\varepsilon}) = \mathbf{0}, \quad E(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}') = \sigma^2\mathbf{I}_n$$

$$\mathbf{R}\beta + \mathbf{v} = \mathbf{r}, \quad E(\mathbf{v}) = \mathbf{0}, \quad E(\mathbf{v}\mathbf{v}') = \Psi, \quad E(\boldsymbol{\varepsilon}\mathbf{v}') = \mathbf{0}. \quad (4.79)$$

veya

$$\begin{bmatrix} \mathbf{y} \\ \mathbf{r} \end{bmatrix} = \begin{bmatrix} \mathbf{X} \\ \mathbf{R} \end{bmatrix} \beta + \begin{bmatrix} \boldsymbol{\varepsilon} \\ \mathbf{v} \end{bmatrix} \quad (4.80)$$

gibi ya da $\mathbf{a} = (\mathbf{y} \ \mathbf{r})'$, $\mathbf{A} = (\mathbf{X} \ \mathbf{R})'$, $\boldsymbol{\omega} = (\boldsymbol{\varepsilon} \ \mathbf{v})'$ olmak üzere,

$$\mathbf{a} = \mathbf{A}\beta + \boldsymbol{\omega} \quad (4.81)$$

olarak yazarız. Burada

$$E(\boldsymbol{\omega}) = \begin{bmatrix} E(\boldsymbol{\varepsilon}) \\ E(\mathbf{v}) \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \quad \boldsymbol{\Omega} = E(\boldsymbol{\omega}\boldsymbol{\omega}') = \begin{bmatrix} \boldsymbol{\varepsilon}\boldsymbol{\varepsilon}' & \boldsymbol{\varepsilon}\mathbf{v}' \\ \mathbf{v}\boldsymbol{\varepsilon}' & \mathbf{v}\mathbf{v}' \end{bmatrix} = \begin{bmatrix} \sigma^2\mathbf{I}_n & \mathbf{0} \\ \mathbf{0} & \Psi \end{bmatrix} \quad (4.82)$$

olduğuna dikkat ederiz. Bu, $\boldsymbol{\omega}$ hatalarının küresel olmadığını veya aynı varyanslı olmadığını (değişen varyanslı olduğunu) gösterir. Bu nedenle, genelleştirilmiş en

küçük kareler tahmininin uygulanması alışılmış en küçük kareler tahmininden daha etkin tahmin ediciyi verir. Böylece

$$\mathbf{a} = \mathbf{A}\boldsymbol{\beta} + \boldsymbol{\omega}, \quad E(\boldsymbol{\omega}) = \mathbf{0}, \quad \text{Var}(\boldsymbol{\omega}) = \boldsymbol{\Omega} \quad (4.83)$$

modeline, genelleştirilmiş en küçük kareleri uygulayarak, $\boldsymbol{\beta}$ nın en küçük kareler tahmin edicisi

$$\hat{\boldsymbol{\beta}}_M = (\mathbf{A}'\boldsymbol{\Omega}^{-1}\mathbf{A})^{-1} \mathbf{A}'\boldsymbol{\Omega}^{-1}\mathbf{A} \quad (4.84)$$

ile verilir. Bu tahmin edicinin açık biçimi aşağıdaki gibi elde edilir:

$$\begin{aligned} \mathbf{A}'\boldsymbol{\Omega}^{-1}\mathbf{a} &= \begin{bmatrix} \mathbf{X}' & \mathbf{R}' \end{bmatrix} \begin{bmatrix} \frac{1}{\sigma^2} \mathbf{I}_n & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Psi}^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{y} \\ \mathbf{r} \end{bmatrix} \\ &= \frac{1}{\sigma^2} \mathbf{X}'\mathbf{y} + \mathbf{R}'\boldsymbol{\Psi}^{-1}\mathbf{r}, \\ \mathbf{A}'\boldsymbol{\Omega}^{-1}\mathbf{A} &= \begin{bmatrix} \mathbf{X}' & \mathbf{R}' \end{bmatrix} \begin{bmatrix} \frac{1}{\sigma^2} \mathbf{I}_n & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Psi}^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{X} \\ \mathbf{R} \end{bmatrix} \\ &= \frac{1}{\sigma^2} \mathbf{X}'\mathbf{X} + \mathbf{R}'\boldsymbol{\Psi}^{-1}\mathbf{R}. \end{aligned} \quad (4.85)$$

Bu nedenle, σ^2 yi bilinmeyen olarak kabul ederek,

$$\hat{\boldsymbol{\beta}}_M = \left(\frac{1}{\sigma^2} \mathbf{X}'\mathbf{X} + \mathbf{R}'\boldsymbol{\Psi}^{-1}\mathbf{R} \right)^{-1} \left(\frac{1}{\sigma^2} \mathbf{X}'\mathbf{y} + \mathbf{R}'\boldsymbol{\Psi}^{-1}\mathbf{r} \right) \quad (4.86)$$

elde ederiz. Bu, karma regresyon tahmin edicisi olarak adlandırılır.

Eğer σ^2 bilinmiyorsa, bu takdirde σ^2 yerine onun

$$\hat{\sigma}^2 = \mathbf{s}^2 = \frac{1}{n-k} (\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})' (\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}) \quad (4.87)$$

tahmin edicisi koyulabilir ve $\boldsymbol{\beta}$ nın uygulanabilir karma regresyon tahmin edicisi

$$\hat{\boldsymbol{\beta}}_f = \left(\frac{1}{\mathbf{s}^2} \mathbf{X}'\mathbf{X} + \mathbf{R}'\boldsymbol{\Psi}^{-1}\mathbf{R} \right)^{-1} \left(\frac{1}{\mathbf{s}^2} \mathbf{X}'\mathbf{y} + \mathbf{R}'\boldsymbol{\Psi}^{-1}\mathbf{r} \right) \quad (4.88)$$

olarak elde edilir. Bu da tahmin edilmiş veya işlevselleştirilmiş en küçük kareler tahmin edicisi olarak adlandırılır (Srivastava, 2018).

Karma Regresyon Tahmin Edicisinin Özellikleri :

a) Yansızlık :

$$\begin{aligned}\hat{\beta}_M - \beta &= (\mathbf{A}'\mathbf{\Omega}^{-1}\mathbf{A})^{-1} \mathbf{A}'\mathbf{\Omega}^{-1}\mathbf{a} - \beta \\ &= (\mathbf{A}'\mathbf{\Omega}^{-1}\mathbf{A})^{-1} \mathbf{A}'\mathbf{\Omega}^{-1}(\mathbf{A}\beta + \omega) - \beta \\ &= (\mathbf{A}'\mathbf{\Omega}^{-1}\mathbf{A})^{-1} \mathbf{A}'\mathbf{\Omega}^{-1}\omega\end{aligned}$$

dır.

$$\begin{aligned}E(\hat{\beta}_M - \beta) &= (\mathbf{A}'\mathbf{\Omega}^{-1}\mathbf{A})^{-1} \mathbf{A}'\mathbf{\Omega}^{-1}E(\omega) \\ &= 0\end{aligned}$$

dır. Bu nedenle, karma regresyon tahmin edicisi β nın bir yansız tahmin edicisidir. $\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$ sade tahmin edicisinin de β nın bir yansız tahmin edicisi olduğuna dikkat edelim.

b) Kovaryans Matrisi :

$\hat{\beta}_M$ nin kovaryans matrisi

$$\begin{aligned}Var(\hat{\beta}_M) &= E\left[(\hat{\beta}_M - \beta)(\hat{\beta}_M - \beta)'\right] \\ &= (\mathbf{A}'\mathbf{\Omega}^{-1}\mathbf{A})^{-1} \mathbf{A}'\mathbf{\Omega}^{-1}E(\mathbf{v}\mathbf{v}')\mathbf{\Omega}^{-1}\mathbf{A}(\mathbf{A}'\mathbf{\Omega}^{-1}\mathbf{A})^{-1} \\ &= (\mathbf{A}'\mathbf{\Omega}^{-1}\mathbf{A})^{-1} \\ &= \left(\frac{1}{\sigma^2} \mathbf{X}'\mathbf{X} + \mathbf{R}'\mathbf{\Omega}^{-1}\mathbf{R}\right)^{-1}\end{aligned}$$

dir.

c) $\hat{\beta}_M$ nin Tahmin Edicisi

$$\begin{aligned}\mathbf{r} &= \mathbf{R}\hat{\beta}_M + \mathbf{v} \\ E(\mathbf{r}) &= \mathbf{R}E(\hat{\beta}_M) + E(\mathbf{v}) \\ &= \mathbf{R}\beta\end{aligned}$$

olması anlamında stokastik lineer kısıtlamaları sağlar.

d) OLSE (Alışılmış En Küçük Kareler Tahmin Edicisi) ile Karşılaştırma :

İlk olarak, $\hat{\beta}_M$ nin $\hat{\beta}$ üzerinde üstünlüğünü saptamak için ileride kullanılan bir sonucu ifade edelim.

Sonuç 4.2. Eğer $(\mathbf{A}_2 - \mathbf{A}_1)$ pozitif tanımlı ise, matrislerin $(\mathbf{A}_1^{-1} - \mathbf{A}_2^{-1})$ farkı da pozitif tanımlıdır.

$$\mathbf{A}_1 = \text{Var}(\hat{\beta}) = \sigma^2 (\mathbf{X}'\mathbf{X})^{-1}$$
$$\mathbf{A}_2 = \text{Var}(\hat{\beta}_M) = \left(\frac{1}{\sigma^2} \mathbf{X}'\mathbf{X} + \mathbf{R}'\mathbf{\Omega}^{-1}\mathbf{R} \right)^{-1}$$

olsun. Bu takdirde,

$$\begin{aligned} \mathbf{A}_1^{-1} - \mathbf{A}_2^{-1} &= \frac{1}{\sigma^2} \mathbf{X}'\mathbf{X} + \mathbf{R}'\mathbf{\Psi}^{-1}\mathbf{R} - \frac{1}{\sigma^2} \mathbf{X}'\mathbf{X} \\ &= \mathbf{R}'\mathbf{\Psi}^{-1}\mathbf{R} \end{aligned} \quad (4.89)$$

dur ki bu, pozitif tanımlı bir matristir. Bu,

$$\mathbf{A}_1 - \mathbf{A}_2 = \text{Var}(\hat{\beta}) - \text{Var}(\hat{\beta}_M) \quad (4.90)$$

ın pozitif tanımlı bir matris olduğunu belirtir. Bu nedenle kovaryans matrislerinin ölçütü (kriteri) veya σ^2 nin bilinmesi şartıyla Löwner sıralaması altında, $\hat{\beta}_M$; $\hat{\beta}$ dan daha etkindir (Srivastava, 2018).

4.1.2.2. Hipotezin Test Edilmesi

(4.2) stokastik kısıtlamasıyla belirlenen ön bilgide, örneklem bilgisi ve ön bilgi arasında yakın ilişkinin var olup olmadığını test etmek isteriz. Örneklem ve ön bilginin uyumluluğu için test, σ^2 nin bilindiğini ve $\hat{\beta}$ nın $\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{y}$ olduğunu kabul ederek,

$$\chi^2 = \frac{1}{\sigma^2} (\mathbf{r} - \mathbf{R}\hat{\beta})' \left[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1} \mathbf{R}' + \mathbf{\Psi} \right]^{-1} (\mathbf{r} - \mathbf{R}\hat{\beta}) \quad (4.91)$$

ile verilen χ^2 -test istatistiği ile test edilir. Bu istatistik m serbestlik dereceli bir χ^2 -dağılımını izler. Eğer $\mathbf{\Psi} = \mathbf{0}$ ise, bu takdirde dağılım dejenere olmuştur ve bu nedenle, \mathbf{r} bir sabit nicelik olur. (4.88) karma regresyon tahmin edicisinin lineerlik, yansızlık ve/veya minimum varyans gibi optimal özellikleri geçerli kalmaz. Bu nedenle, ön bilginin birleştirmesinin etkinlikte kayıba götürebildiği durumlar altında, kusursuz regresyon tahmin edicisini kullanmak daha iyidir. Ön bilginin kullanımının

daha iyi tahmin ediciye götürüp götürmeyeceğini bilmek için, $H_0 : E(\mathbf{r}) = \mathbf{R}\boldsymbol{\beta}$ sıfır hipotezi test edilebilir.

σ^2 bilindiğinde, $H_0 : E(\mathbf{r}) = \mathbf{R}\boldsymbol{\beta}$ sıfır hipotezini test etmek için,

$$\mathbf{s}^2 = \frac{1}{n-p} (\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})' (\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})$$

olmak üzere,

$$F = \frac{(\mathbf{r} - \mathbf{R}\hat{\boldsymbol{\beta}})' [\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}' + \boldsymbol{\Psi}]^{-1} (\mathbf{r} - \mathbf{R}\hat{\boldsymbol{\beta}}) / m}{\mathbf{s}^2} \quad (4.92)$$

ile verilen F -istatistiğini kullanırız ve H_0 altında F , m ve $n-p$ serbestlik dereceli F -dağılımına uyar (Srivastava, 2018).

4.1.3. Lineer Modellerde Lineer Eşitsizlik Kısıtlamaları

Bazen regresyon parametreleri üzerindeki kısıtlama veya eşdeğer olarak, regresyon parametreleri hakkındaki ön bilgi eşitsizlik biçiminde mevcuttur. Örneğin, $1 < \beta_1 < 2$, $5 < \beta_3 < 6$, $2 < \beta_1 + 2\beta_2 < 5$ v.s. Böyle bilginin $\mathbf{R}\boldsymbol{\beta} \geq \mathbf{r}$ biçiminde ifade edilebildiğini farz edelim. $\mathbf{R}\boldsymbol{\beta} \geq \mathbf{r}$ kısıtlamasına bağlı $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ modelinde $\boldsymbol{\beta}$ regresyon katsayısını (regresyon katsayıları vektörünü) tahmin etmeyi istiyoruz.

$\boldsymbol{\beta}$ nın bir tahminini elde etmek için $(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})' (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$, $\mathbf{R}\boldsymbol{\beta} \geq \mathbf{r}$ kısıtlaması altında minimumlaştırılabilir. Bu, bir kuadratik programlama problemi olarak formüllenebilir ve uygun bir algoritmayı, örneğin; simpleks algoritmasını kullanarak çözülebilir ve bir nümerik çözüm elde edilir. Bu yöntemin avantajı, şartı yerine getiren bir $\hat{\boldsymbol{\beta}}$ çözümünün bulunmasıdır. Dezavantajı, tahminin istatistiksel özelliklerinin kolay belirlenememesi ve üstünlük hakkındaki genel sonuçların ortaya konamamasıdır.

$\boldsymbol{\beta}$ nın eşitsizlik kısıtlamalarına bağlı bir tahmin edicisini elde etmek için başka bir seçenek, eşitsizlik kısıtlamalarını stokastik lineer kısıtlamalar biçimine, örneğin, p -sigma sınırları biçimine, çevirmektir.

Eşitsizlik kısıtlamaları altında $\boldsymbol{\beta}$ nın tahmin edicisini elde etmek için minimaks tahmini de kullanılabilir. Minimaks tahmini, $\hat{\boldsymbol{\beta}}$ tahmini için kuadratik risk fonksiyonunun tüm parametre üzerinden minimumlaştırılmadığı aynı zamanda ön bilgi ile sınırlanan bir alan üzerinden veya tahminle ilgili kısıtlamalar üzerinden minimumlaştırıldığı düşüncesine dayanır.

Eğer tüm kısıtlama bir konveks alanı tanımlarsa bu alan orijini merkezde

$$B(\boldsymbol{\beta}) = \{ \boldsymbol{\beta} : \boldsymbol{\beta}' \mathbf{T} \boldsymbol{\beta} \leq p \} \quad (4.93)$$

biçimli bir elipsoid içinde veya $\boldsymbol{\beta}_0$ merkez nokta vektörlü

$$B(\boldsymbol{\beta}, \boldsymbol{\beta}_0) = \{ \boldsymbol{\beta} : (\boldsymbol{\beta} - \boldsymbol{\beta}_0)' \mathbf{T} (\boldsymbol{\beta} - \boldsymbol{\beta}_0) \leq p \} \quad (4.94)$$

biçimindeki bir elipsoid içinde kapsanır. Burada $p > 0$ verilen bir sabittir ve \mathbf{T} pozitif tanımlı olması kabul edilen, bilinen bir $p \times p$ matristir. Burada B , bir yoğunlaşma elipsoidini tanımlar (Krotterus, 2016).

4.2. Örneklem Geometrisi, Rasgele Örneklem ve Korelasyon Katsayısının Farklı Geometrik Yorumları

$$\mathbf{X}_{(n \times p)} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1k} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2k} & \cdots & x_{2p} \\ \vdots & \vdots & & \vdots & & \vdots \\ x_{j1} & x_{j2} & \cdots & x_{jk} & \cdots & x_{jp} \\ \vdots & \vdots & & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nk} & \cdots & x_{np} \end{bmatrix} \quad (4.95)$$

matrisini göz önüne alalım. p -örneklem ortalaması ve p -örneklem varyansı,

$$\bar{x}_k = \frac{1}{n} \sum_{j=1}^n x_{jk}, \quad k = 1, 2, \dots, p \quad (4.96)$$

ve

$$s_k^2 = \frac{1}{n} \sum_{j=1}^n (x_{jk} - \bar{x}_k)^2, \quad k = 1, 2, \dots, p \quad (4.97)$$

şeklindedir. $\bar{\mathbf{x}}$ örneklem ortalaması, \mathbf{S}_n örneklem varyans-kovaryansı ve \mathbf{R} örneklem korelasyonu sırasıyla

$$\bar{\mathbf{x}} = \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \vdots \\ \bar{x}_p \end{bmatrix}, \quad (4.98)$$

$$\mathbf{S}_n = \begin{bmatrix} s_{11} & s_{12} & \cdots & s_{1p} \\ s_{21} & s_{22} & \cdots & s_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ s_{n1} & s_{n2} & \cdots & s_{np} \end{bmatrix} \quad (4.99)$$

ve

$$\mathbf{R} = \begin{bmatrix} 1 & r_{12} & \cdots & r_{1p} \\ r_{21} & 1 & \cdots & r_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ r_{n1} & r_{n2} & \cdots & 1 \end{bmatrix} \quad (4.100)$$

şeklinde gösterilir. Önceki bölümde ortaya konan matrisler ve vektör kavramları ile, şimdi $\bar{\mathbf{x}}$, \mathbf{S}_n ve \mathbf{R} betimleyici istatistiklerinin geometrik yorumlarını daha derin araştırabiliriz. Bunu Kısım 4.2.1. de yapacağız. Açıklamalarımızın birçoğu \mathbf{X} in sütunlarını, n -boyutta p -vektörler olarak kullanır. Kısım 4.2.2. de, gözlemlerin bir rasgele örnekleme oluşturduğu varsayımını ortaya koyacağız.

Basit olarak ifade edilen rasgele örneklem, (1) farklı öğeler (veya denemeler) üzerinde alınan ölçümlerin birinin diğeri ile ilişkisiz olduğunu ve (2) p -değişkeninin tümünün ortak dağılımının tüm öğeler için aynı kaldığını belirtir. Sonuç olarak, uzaklığın özel bir seçimini meşrulaştıran ve verinin n -boyutlu gösterimi için geometriyi dikte eden, rasgele örneklemin bu yapısıdır. Bundan başka veri, bir rasgele örneklem olarak ele alınabildiğinde, istatistiksel çıkarımlar (sonuçlar) sağlam bir zemine dayandırılır.

Geometrik yorumlara dönerek Kısım 4.2.3. de, değişkenliği ifade etmek için genelleştirilmiş varyans denilen bir tek sayıyı ortaya koyacağız. Varyansın bu genelleştirmesi, çok değişkenli ortalamaların karşılaştırılmasının ayrılmaz bir parçasıdır. Daha sonraki kısımlarda $\bar{\mathbf{x}}$ yi ve \mathbf{S}_n yi, \mathbf{X} veri matrisinden doğrudan doğruya hesaplamamıza imkan veren matris çarpımları ve toplamları için özlü ifadeleri ortaya koymak için matris cebirini kullanacağız. $\bar{\mathbf{x}}$ ve \mathbf{S}_n arasındaki ilişki ve değişkenlerin lineer kombinasyonları için ortalamalar ve kovaryanslar da matris çarpımları kavramını kullanarak açık olarak resmedilir.

Kısım 4.2.4. de ise, korelasyon katsayısına farklı geometrik yorumlar getireceğiz.

4.2.1. Örneklemin Geometrisi

Bir tek çok değişkenli gözlem, aynı öge veya deneme üzerinde alınan p -sayıda farklı değişken üzerindeki ölçümlerin biriktirisidir. Eğer n -sayıda gözlem elde

edilmiş ise, tüm veri kümesi $n \times p$ matris (düzen) içine koyulabilir. Böylece (4.95) veri matrisi ortaya çıkar.

\mathbf{X} in her bir satırı bir çok değişkenli gözlemi gösterir. Ölçümlerin tüm kümesi çoğu kez gözlenmiş olan şeyin özel bir gerçekleşmesi olduğundan, verinin p -değişkenli bir “kitleden” n -büyüklüğe sahip bir örneklem olduğunu söyleriz. Bu durumda örneklem, her birinin p -sayıda bileşene sahip olduğu n -sayıda ölçümden oluşur.

Gördüğümüz gibi, veri iki farklı biçimde grafiklenebilir. p -boyutlu saçılım grafiği için, \mathbf{X} in satırları p -boyutlu uzayda n -sayıda noktayı gösterir.

$$\mathbf{X}_{(n \times p)} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1k} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2k} & \cdots & x_{2p} \\ \vdots & \vdots & & \vdots & & \vdots \\ x_{j1} & x_{j2} & \cdots & x_{jk} & \cdots & x_{jp} \\ \vdots & \vdots & & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nk} & \cdots & x_{np} \end{bmatrix} = \begin{bmatrix} \mathbf{x}'_1 \\ \mathbf{x}'_2 \\ \vdots \\ \mathbf{x}'_j \\ \vdots \\ \mathbf{x}'_n \end{bmatrix} \begin{matrix} \leftarrow 1\text{-inci (çok değişkenli) gözlem} \\ \\ \\ \\ \\ \leftarrow n\text{-yinci (çok değişkenli) gözlem} \end{matrix} \quad (4.101)$$

j -yinci gözlemi gösteren \mathbf{x}'_j satır vektörü, bir noktanın koordinatlarını içerir.

p -boyutlu uzayda n -noktanın saçılım grafiği, noktaların konumları ve değişkenliği hakkında bilgiyi ortaya koyar. Eğer noktalara katı yuvar gözüyle bakılırsa, (4.98), (4.99) ve (4.100) ile verilen $\bar{\mathbf{x}}$ örneklem ortalaması vektörü, denge merkezidir. Değişkenlik bir doğrultudan daha fazlasında ortaya çıkar ve \mathbf{S}_n örneklem varyans-kovaryans matrisiyle ölçülür. p ; 3 ten daha büyük olduğunda bu saçılım grafiği gösterimi sahiden grafiklenemez. Yine de, verinin p -boyutdaki n nokta olarak düşünülmesi cebirsel ifadelerden kolayca elde edilemeyen sezgileri verir. Ayrıca, $p = 2$ veya $p = 3$ için açıklanan kavramlar diğer durumlar için geçerli kalır.

Örnek 4.2. Ortalama Vektörünü Hesaplama

$$\mathbf{X} = \begin{bmatrix} 2 & 5 \\ -3 & 4 \\ 7 & 3 \end{bmatrix}$$

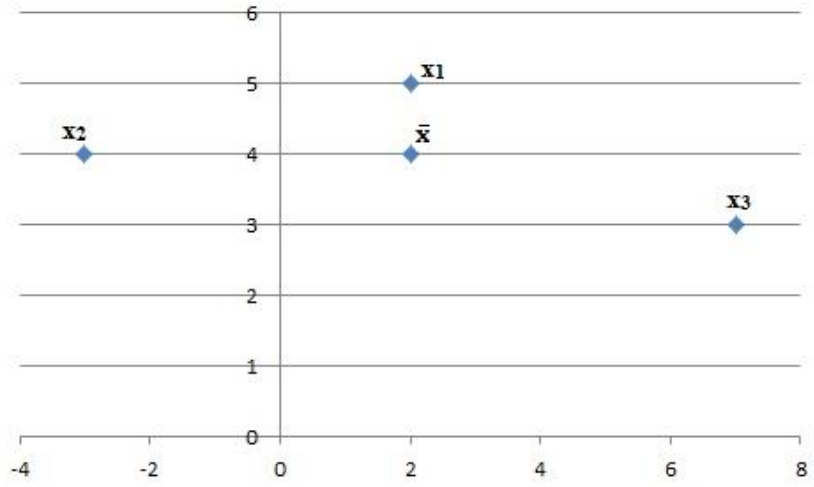
veri matrisinden $\bar{\mathbf{x}}$ ortalama vektörünü hesaplayalım.

$p = 2$ boyutlu uzayda $n = 3$ veri noktasını gösterelim ve sonuçta ortaya çıkan diyagram üzerinde $\bar{\mathbf{x}}$ yi yerleştirelim.

İlk nokta \mathbf{x}_1 , $\mathbf{x}_1' = [2, 5]$ koordinatlarına sahiptir. Benzer şekilde geri kalan iki nokta $\mathbf{x}_2' = [-3, 4]$ ve $\mathbf{x}_3' = [7, 3]$ tür. Sonuç olarak

$$\bar{\mathbf{x}} = \begin{bmatrix} \frac{2-3+7}{3} \\ \frac{5+4+3}{3} \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$$

dür.



Şekil 4.2. \mathbf{X} veri matrisinin $p = 2$ boyutlu uzayda $n = 3$ sayıda nokta olarak grafiği.

Şekil 4.2, $\bar{\mathbf{x}}$ nin saçılım grafiğinin denge noktası (ağırlık merkezi) olduğunu gösterir.

Farklı bir geometrik gösterim, n -boyutlu uzayda p -vektörler olarak veriyi düşünmek suretiyle oluşturulur. Burada veri matrisinin sütunlarının elemanlarını, vektörlerin koordinatları olarak alırız.

$$\mathbf{X}_{(n \times p)} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1k} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2k} & \cdots & x_{2p} \\ \vdots & \vdots & & \vdots & & \vdots \\ x_{j1} & x_{j2} & \cdots & x_{jk} & \cdots & x_{jp} \\ \vdots & \vdots & & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nk} & \cdots & x_{np} \end{bmatrix} = [\mathbf{y}_1 \quad \vdots \quad \mathbf{y}_2 \quad \vdots \quad \cdots \quad \vdots \quad \mathbf{y}_p] \quad (4.102)$$

olsun. Bu durumda $\mathbf{y}_1' = [x_{11}, x_{21}, \dots, x_{n1}]$ noktasının koordinatları, birinci değişken üzerindeki ölçümdür. Genel olarak, i -yinci $\mathbf{y}_i' = [x_{1i}, x_{2i}, \dots, x_{ni}]$; i -yinci değişken üzerindeki tüm ölçümlerin n -lisi ile belirlenir. Bu geometrik gösterimde,

y_1, y_2, \dots, y_p leri p -boyutlu saçılım grafiğinde noktalardan ziyade vektörler olarak gösteririz.

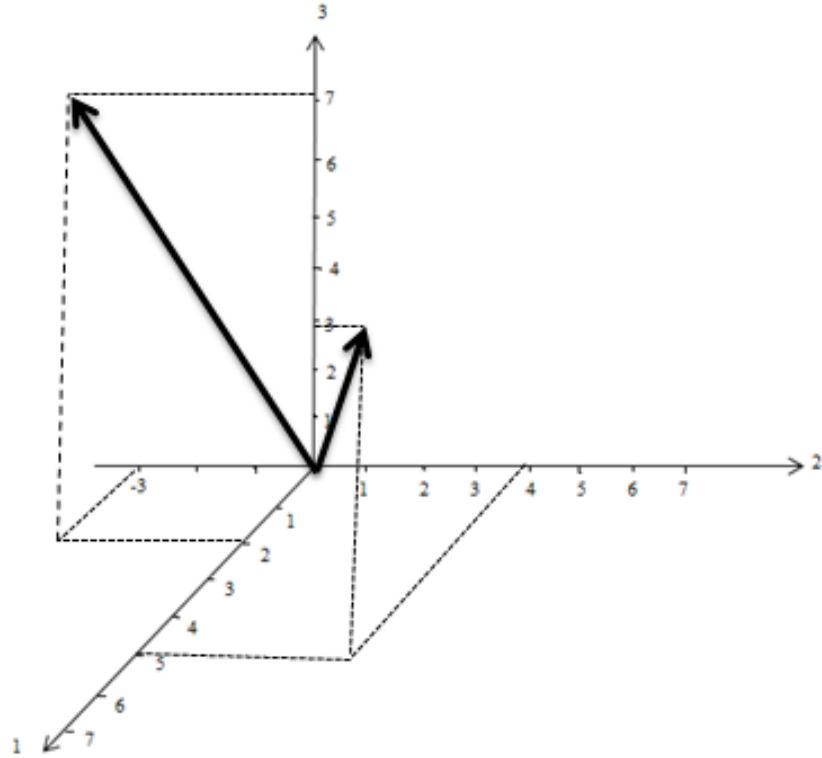
Örnek 4.3. n -Boyutta p -Vektörler Olarak Veri

Aşağıdaki veriyi $n = 3$ boyutlu uzayda $p = 2$ vektörler olarak grafikleyelim.

$$\mathbf{X} = \begin{bmatrix} 2 & 5 \\ -3 & 4 \\ 7 & 3 \end{bmatrix}$$

Burada, $y_1' = [2, -3, 7]$ ve $y_2' = [5, 4, 3]$ dir. Bu vektörler Şekil 4.3 de gösterilir.

Çok değişkenli analizde karşılaştığımız cebirsel ifadelerin birçoğu; uzunluk, açı ve hacimin geometrik anlamlarına bağlanabilir. Geometrik gösterimler, anlamayı kolaylaştırdığından önemlidir.



Şekil 4.3. \mathbf{X} veri matrisinin $n = 3$ boyutlu uzayda $p = 2$ vektörler olarak grafiği.

Maalesef nesnelere 3 boyutta görsellemeye tabi tutulamaz ve sonuç olarak, \mathbf{X} veri matrisinin n -boyutlu gösterimi özellikle $n > 3$ için gözükemeyebilir. Bununla beraber, herhangi üç vektör için geometrik ilişkileri ifade eden ilgili istatistiksel kavramlar, onların boyutlarına bakmaksızın geçerlidir. n -boyutlu olsa bile, üç vektör herhangi sayıda bileşene sahip iki vektör bir düzlemde uzandığı anda, üç vektör bir

üç boyutlu uzaydan daha fazlasını geremediğinden, bu görülür. Uygun bir üç boyutlu perspektifi (bakışı) seçerek, yani, ilgili üç vektörü ihtiva eden n -boyutlu uzayın bir kısmını seçerek, hem uzunlukları hem de açıları koruyan bir görüş elde edilir. Bu nedenle, eksenlerin dik açılı seçimi ile, herhangi n -boyuta sahip yalnız iki veya üç vektöre dayanarak belirli istatistiksel kavramları açıklamak mümkündür.

Bir örneklem ortalaması bulma yönteminin bir geometrik yorumunu vermek mümkündür. $\mathbf{j}'_n = [1, 1, \dots, 1]$ vektörünü tanımlayarak işe başlayalım (Notasyonu basitleştirmek için \mathbf{j}_n vektörünün boyutu içerikten açık olduğundan, n -altındisini yazmayacağız). \mathbf{j} vektörü n -sayıda koordinat ekseninin her biri ile eşit açılar oluşturur, bu nedenle $(1/\sqrt{n})\mathbf{j}$ eşit-açılı doğrultuda birim uzunluğa sahiptir. $\mathbf{y}'_i = [x_{1i}, x_{2i}, \dots, x_{ni}]$ vektörünü göz önüne alalım.

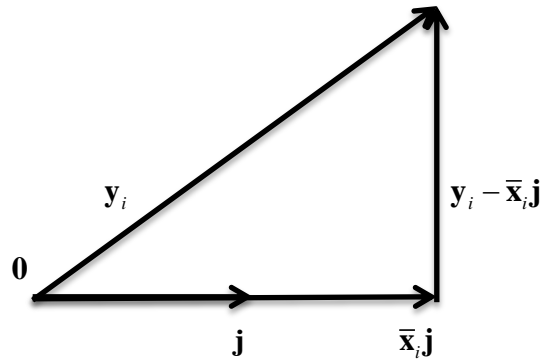
$$\mathbf{x} \text{ vektörünün } \mathbf{y} \text{ üzerindeki izdüşümü} = \frac{(\mathbf{x}'\mathbf{y})}{\mathbf{y}'\mathbf{y}} \mathbf{y} \quad (4.103)$$

dir. \mathbf{y}_i nin $(1/\sqrt{n})\mathbf{j}$ birim vektörü üzerindeki izdüşümü (4.104) ile verilir. Bu izdüşüm;

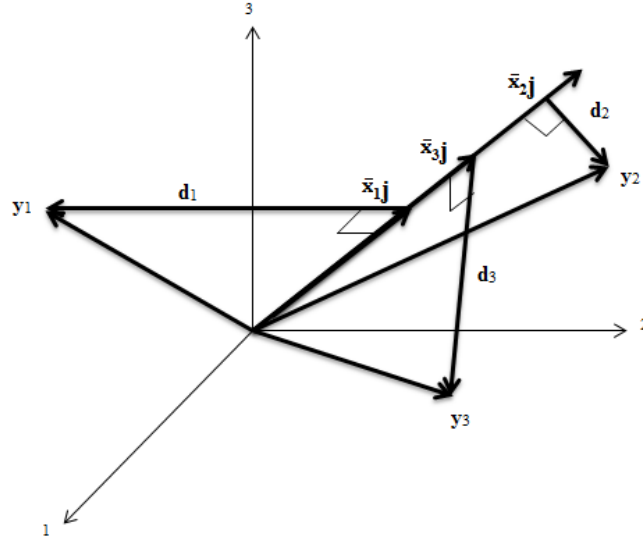
$$\mathbf{y}'_i \left(\frac{1}{\sqrt{n}} \mathbf{j} \right) \frac{1}{\sqrt{n}} \mathbf{j} = \frac{x_{1i} + x_{2i} + \dots + x_{ni}}{n} \mathbf{j} = \bar{x}_i \mathbf{j} \quad (4.104)$$

dir. Yani, $\bar{x}_i = (x_{1i} + x_{2i} + \dots + x_{ni})/n = \mathbf{y}'_i \mathbf{j} / n$, \mathbf{y}_i nin \mathbf{j} ile belirlenen doğru üzerine izdüşümünü vermek için gereken \mathbf{j} nin katına karşılık gelir.

Ayrıca, her bir \mathbf{y}_i için



ayrışımını elde ederiz. Burada $\bar{x}_i \mathbf{j}$; $\mathbf{y}_i - \bar{x}_i \mathbf{j}$ ye diktir.



Şekil 4.4. y_i nin bir \bar{x}_i,j ortalama bileşeni ve bir $d_i = y_i - \bar{x}_i,j$ sapma bileşenine ayrışımı.

Sapma veya düzeltilmiş ortalama vektörü

$$d_i = y_i - \bar{x}_i,j = \begin{bmatrix} x_{1i} - \bar{x}_i \\ x_{2i} - \bar{x}_i \\ \vdots \\ x_{ni} - \bar{x}_i \end{bmatrix} \quad (4.105)$$

dır. d_i nin elemanları, i -yinci değişken üzerindeki ölçümlerin onların örneklem ortalamasından sapmalarıdır. y_i vektörlerinin ortalama bileşenlere ve ortalama bileşenlerden sapmaya ayrışımı, $p = 3$ ve $n = 3$ için, Şekil 4.4 de gösterilir.

Örnek 4.4. Bir Vektörü Ortalamasına ve Sapma Bileşenlerine Ayırma

Örnek 4.3 de verilen veri için, y_i nin \bar{x}_i,j ve $d_i = y_i - \bar{x}_i,j$, $i = 1, 2$ ayrışımını hesaplayalım. Burada görüldüğü üzere,

$$X = \begin{bmatrix} 2 & 5 \\ -3 & 4 \\ 7 & 3 \end{bmatrix}$$

dir. Buna göre, $\bar{x}_1 = 2$ ve $\bar{x}_2 = 4$ dür, bu nedenle,

$$\bar{\mathbf{x}}_1 \mathbf{j} = 2 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \quad \text{ve} \quad \bar{\mathbf{x}}_2 \mathbf{j} = 4 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

dir. Sonuç olarak,

$$\mathbf{d}_1 = \mathbf{y}_1 - \bar{\mathbf{x}}_1 \mathbf{j} = \begin{bmatrix} 2 \\ -3 \\ 7 \end{bmatrix} - \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ -5 \\ 5 \end{bmatrix}$$

ve

$$\mathbf{d}_2 = \mathbf{y}_2 - \bar{\mathbf{x}}_2 \mathbf{j} = \begin{bmatrix} 5 \\ 4 \\ 3 \end{bmatrix} - \begin{bmatrix} 4 \\ 4 \\ 4 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$$

dir.

$$(\bar{\mathbf{x}}_1 \mathbf{j})' (\mathbf{y}_1 - \bar{\mathbf{x}}_1 \mathbf{j}) = [2 \quad 2 \quad 2] \begin{bmatrix} 0 \\ -5 \\ 5 \end{bmatrix} = 0 - 10 + 10 = 0$$

olduğundan, $\bar{\mathbf{x}}_1$ ve $\mathbf{d}_1 = \mathbf{y}_1 - \bar{\mathbf{x}}_1 \mathbf{j}$ nin dik olduğuna dikkat edelim. Benzer sonuç $\bar{\mathbf{x}}_2 \mathbf{j}$ ve $\mathbf{d}_2 = \mathbf{y}_2 - \bar{\mathbf{x}}_2 \mathbf{j}$ için de geçerlidir. Ayrışım,

$$\mathbf{y}_1 = \begin{bmatrix} 2 \\ -3 \\ 7 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix} + \begin{bmatrix} 0 \\ -5 \\ 5 \end{bmatrix}, \quad \mathbf{y}_2 = \begin{bmatrix} 5 \\ 4 \\ 3 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \\ 4 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$$

dir.

Şimdilik, $\mathbf{d}_i = \mathbf{y}_i - \bar{\mathbf{x}}_i \mathbf{j}$ sapma (hata) vektörleri ile ilgileneceğiz. Şekil 4.4 ün sapma vektörlerinin grafiği Şekil 4.5 de verilir. Sapma vektörlerini, uzunluklarını veya yönelimlerini değiştirmeksizin orijine taşıdık.

Şimdi sapma vektörlerinin uzunluklarının karelerini göz önüne alalım. Bir \mathbf{x} vektörünün uzunluğu

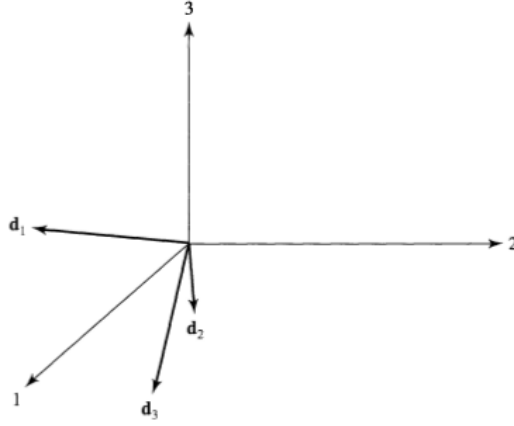
$$L_x = \mathbf{x} \text{ in uzunluğu} = \sqrt{\mathbf{x}' \mathbf{x}} \quad (4.106)$$

ile gösterilir. (4.105) i ve (4.106) yı kullanarak,

$$L_{\mathbf{d}_i}^2 = \mathbf{d}_i' \mathbf{d}_i = \sum_{j=1}^n (x_{ji} - \bar{x}_i)^2 \quad (4.107)$$

(sapma vektörünün uzunluğu)² = sapmaların karelerinin toplamı

elde ederiz.



Şekil 4.5. Şekil 4.4 deki \mathbf{d}_i sapma vektörleri.

(4.97) den uzunluk karesinin, i -yinci değişken üzerindeki ölçümlerin varyansı ile orantılı olduğunu görürüz. Eşdeğer olarak uzunluk, standart sapma ile orantılıdır. Daha uzun vektörler, daha kısa vektörlerden daha fazla değişkenliği temsil eder.

Herhangi iki \mathbf{d}_i ve \mathbf{d}_k sapma vektörleri için,

$$\mathbf{d}_i' \mathbf{d}_k = \sum_{j=1}^n (x_{ji} - \bar{x}_i)(x_{jk} - \bar{x}_k) \quad (4.108)$$

dır. θ_{ik} ; \mathbf{d}_i ve \mathbf{d}_k vektörleri tarafından oluşturulan açığı gösterebilir. Herhangi \mathbf{x} ve \mathbf{y} vektörleri arasındaki açı

$$\cos(\theta) = \frac{\mathbf{x}' \mathbf{y}}{L_x L_y} = \frac{\mathbf{x}' \mathbf{y}}{\sqrt{\mathbf{x}' \mathbf{x}} \sqrt{\mathbf{y}' \mathbf{y}}} \quad (4.109)$$

dir. (4.109) dan,

$$\mathbf{d}_i' \mathbf{d}_k = L_{\mathbf{d}_i} L_{\mathbf{d}_k} \cos(\theta_{ik}) \quad (4.110)$$

dır veya (4.107) yi ve (4.108) i kullanarak,

$$\sum_{j=1}^n (x_{ji} - \bar{x}_i)(x_{jk} - \bar{x}_k) = \sqrt{\sum_{j=1}^n (x_{ji} - \bar{x}_i)^2} \sqrt{\sum_{j=1}^n (x_{jk} - \bar{x}_k)^2} \cos(\theta_{ik}) \quad (4.111)$$

ve bu nedenle,

$$r_{ik} = \frac{S_{ik}}{\sqrt{S_{ii}} \sqrt{S_{kk}}} \cos(\theta_{ik}) \quad (4.112)$$

elde ederiz. Açının kosinüsü örneklem korelasyon katsayısıdır. Bu nedenle eğer, iki sapma vektörü yaklaşık olarak aynı yönelimli ise, örneklem korelasyon katsayısı 1 e yakın olacaktır. Eğer bu iki vektör yaklaşık olarak dik ise, örneklem korelasyonu yaklaşık olarak sıfır olacaktır. Eğer iki vektör yaklaşık olarak zıt yönde yönelimli ise örneklem korelasyon katsayısı -1 e yakın olacaktır. Korelasyon katsayısına Kısım 4.2.4. de yeniden değineceğiz.

Örnek 4.5. Sapma Vektörlerinden S_n yi ve R yi Hesaplama

Örnek 4.4 deki sapma vektörleri verilsin, biraz önce tanıtılan geometrik kavramları kullanarak, S_n örneklem varyans-kovaryans matrisini ve R örneklem korelasyon matrisini hesaplayalım.

Örnek 4.4 den,

$$\mathbf{d}_1 = \begin{bmatrix} 0 \\ -5 \\ 5 \end{bmatrix} \quad \text{ve} \quad \mathbf{d}_2 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$$

dir. Orijine taşınan bu vektörler Şekil 4.6 da gösterilmiştir.

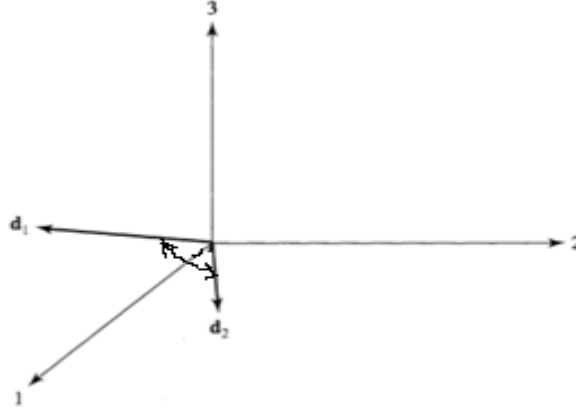
Şimdi

$$\mathbf{d}_1' \mathbf{d}_1 = [0 \quad -5 \quad 5] \begin{bmatrix} 0 \\ -5 \\ 5 \end{bmatrix} = 50 = 3s_{11}$$

veya $s_{11} = \frac{50}{3}$ dür. Aynı zamanda,

$$\mathbf{d}_2' \mathbf{d}_2 = [1 \quad 0 \quad -1] \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} = 2 = 3s_{22}$$

veya $s_{22} = \frac{2}{3}$ dür.



Şekil 4.6. \mathbf{d}_1 ve \mathbf{d}_2 sapma vektörleri.

Son olarak,

$$\mathbf{d}_1' \mathbf{d}_2 = [0 \quad -5 \quad 5] \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} = -5 = 3s_{12}$$

veya $s_{12} = \frac{-5}{3}$ dür. Sonuç olarak

$$r_{12} = \frac{s_{12}}{\sqrt{s_{11}} \sqrt{s_{22}}} = \frac{-5/3}{\sqrt{50/3} \sqrt{2/3}} = -0.5$$

ve

$$\mathbf{S}_n = \begin{bmatrix} 50/3 & -5/3 \\ -5/3 & 2/3 \end{bmatrix}, \quad \mathbf{R} = \begin{bmatrix} 1 & -0.5 \\ -0.5 & 1 \end{bmatrix}$$

dir. Uzunluk, açı ve izdüşüm kavramları bize örneklemin bir geometrik yorumunu sağladı. Bunları aşağıdaki gibi özetleriz.

4.2.1.1. Örneklemin Geometrik Yorumu

1. \mathbf{X} veri matrisinin bir \mathbf{y}_i sütununun \mathbf{j} eşit açılı vektörü üzerine izdüşümü $\bar{\mathbf{x}}_i \mathbf{j}$ vektörüdür. $\bar{\mathbf{x}}_i \mathbf{j}$ vektörü $\sqrt{n} |\bar{\mathbf{x}}_i|$ uzunluğuna sahiptir. Bu nedenle i -yinci örneklem ortalaması $\bar{\mathbf{x}}_i$, \mathbf{y}_i nin \mathbf{j} üzerine izdüşümünün uzunluğu ile ilgilidir.

2. S_n yi kapsayan bilgi $\mathbf{d}_i = \mathbf{y}_i - \bar{\mathbf{x}}_i \mathbf{j} = [x_{1i} - \bar{x}_i, x_{2i} - \bar{x}_i, \dots, x_{ni} - \bar{x}_i]^T$ sapma vektörlerinden elde edilir. \mathbf{d}_i nin uzunluğunun karesi ns_{ii} dir ve \mathbf{d}_i ve \mathbf{d}_k arasındaki (iç) çarpım ns_{ik} dir.
3. r_{ik} örneklem korelasyonu, \mathbf{d}_i ve \mathbf{d}_k arasındaki açının kosinüsüdür.

Örneklem varyansı ve kovaryansının tanımlarında $(n-1)$ böleni kullanıldığında, uzunluğun karesi ve iç çarpım sırasıyla $(n-1)s_{ii}$ ve $(n-1)s_{ik}$ dir.

4.2.2. Rasgele Örneklem, Örneklem Ortalamasının ve Kovaryans Matrisinin Beklenen Değerleri

Nihai çıkarımları yapmak amacıyla $\bar{\mathbf{x}}$ ve S_n gibi istatistiklerin örneklem değişkenliğini incelemek için, \mathbf{X} veri kümesini oluşturan gözlenen değerlerin değişkenleri hakkında varsayımlar yapmamız gerekir.

Bu durumda, verinin henüz gözlenmemiş olduğunu, fakat p -sayıda değişken hakkında ölçümlerin n -sayıda kümesini biriktirdiğimizi farz edelim. Ölçümler yapılmadan önce, genel olarak onların değerleri kesin olarak tahmin edilemez. Sonuç olarak, onları rasgele değişkenler olarak işleme tabi tutarız. Bu bağlamda, veri matrisindeki (j, k) -yüncü eleman x_{jk} rasgele değişkeni olsun. p -sayıda değişken üzerinde \mathbf{x}_j ölçümlerinin her biri bir rasgele vektördür ve (4.101) rasgele matrisine sahip oluruz. Şimdi bir rasgele örneklem tanımlanabilir.

Eğer (4.101) deki $\mathbf{x}'_1, \mathbf{x}'_2, \dots, \mathbf{x}'_n$ satır vektörleri $f(\mathbf{x}) = f(x_1, x_2, \dots, x_p)$ yoğunluk fonksiyonuna sahip bir birleşik müşterek dağılımdan bağımsız gözlemleri gösterirse, bu takdirde, $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ ye $f(\mathbf{x})$ den bir rasgele örneklemi oluşturmaktadır denir. Matematiksel olarak, $f(\mathbf{x}_j) = f(x_{j1}, x_{j2}, \dots, x_{jp})$, j -yüncü satır vektörü için yoğunluk fonksiyonu olmak üzere, eğer onların yoğunluk fonksiyonu $f(\mathbf{x}_1)f(\mathbf{x}_2)\dots f(\mathbf{x}_n)$ çarpımıyla verilirse, $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ bir rasgele örneklem oluşturur.

Rasgele örneklemin tanımı ile ilgili iki önemli nokta vardır:

1. Tek bir denemede, $\mathbf{x}'_j = [x_{j1}, x_{j2}, \dots, x_{jp}]$ gibi, p -sayıda değişkenin ölçümleri genellikle ilişkili olacaktır. Gerçekten bu durumun olmasını bekleriz. Bununla beraber, farklı denemelerden ölçümler bağımsız olmalıdır.
2. Değişkenler p -sayıda stok fiyatı veya p -sayıda ekonomik göstergenin kümelerinde olduğu gibi, zamanla büyük bir olasılıkla yığıldığında, denemeden denemeye ölçümlerin bağımsızlığı geçerli olmayabilir. Bağımsızlığın deneme ve tecrübe niteliğinden varsayımının ihlalleri, istatistiksel sonuçların kalitesi üzerinde ciddi bir etkiye sahip olabilir.

İstatistiksel bağımsızlık kavramı, uzaklık ölçümü için önemli çıkarımlara sahiptir. Eğer bir vektörün bileşenleri bağımsız ve aynı varyanslara sahipse, öklid uzaklığı uygun gözükür. \mathbf{X} in k -yıncı sütunu $\mathbf{y}'_k = [x_{1k}, x_{2k}, \dots, x_{nk}]$ nın konumuna, n -boyutta bir nokta olarak baktığımızı farz edelim. Bu noktanın konumu (yeri) $f(\mathbf{y}_k) = f(x_{1k}, x_{2k}, \dots, x_{nk})$ ortak olasılık dağılım fonksiyonu ile belirlenir. $x_{1k}, x_{2k}, \dots, x_{nk}$ ölçümleri rasgele örneklemeler olduğundan, $f(\mathbf{y}_k) = f(x_{1k}, x_{2k}, \dots, x_{nk}) = f_k(x_{1k})f_k(x_{2k}) \cdots f_k(x_{nk})$ dır ve sonuç olarak, her bir x_{jk} koordinatı aynı (özdeş) $f_k(x_{jk})$ marjinal dağılımları vasıtasıyla konuma eşit katkılarda bulunur.

$\bar{\mathbf{x}}$ nin ve \mathbf{S}_n nin örneklem dağılımları ile ilgili belirli sonuçlarına ulaşılabilir. Özellikle, $\bar{\mathbf{x}}$ nin ve \mathbf{S}_n nin karşılık gelen $\boldsymbol{\mu}$ kitle ortalama vektörünün ve $\boldsymbol{\Sigma}$ kitle kovaryans matrisinin nokta tahmin edicileri olarak nasıl üstesinden geldiğini görebiliriz.

Sonuç 4.3. $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$; $\boldsymbol{\mu}$ ortalama vektörü ve $\boldsymbol{\Sigma}$ kovaryans matrisine sahip olan bir ortak dağılımdan bir rasgele örneklem olsun. Bu takdirde $\bar{\mathbf{x}}$; $\boldsymbol{\mu}$ nün bir yansız tahmin edicisidir ve onun kovaryans matrisi

$$\frac{1}{n} \boldsymbol{\Sigma}$$

dır. Yani,

$$E(\bar{\mathbf{x}}) = \boldsymbol{\mu} \text{ (kitle ortalaması)} \quad (4.113)$$

$$Kov(\bar{\mathbf{x}}) = \frac{1}{n} \boldsymbol{\Sigma} \text{ (örneklem büyüklüğüne bölünmüş kitle varyans-kovaryans matrisi)}$$

dir. \mathbf{S}_n kovaryans matrisi için,

$$E(\mathbf{S}_n) = \frac{n-1}{n} \boldsymbol{\Sigma} = \boldsymbol{\Sigma} - \frac{1}{n} \boldsymbol{\Sigma}$$

dır. Bu nedenle,

$$E\left(\frac{n}{n-1} \mathbf{S}_n\right) = \boldsymbol{\Sigma} \quad (4.114)$$

dır, bundan dolayı, \mathbf{S}_n ; $(yan) = E(\mathbf{S}_n) - \boldsymbol{\Sigma} = -(1/n)\boldsymbol{\Sigma}$ olmak üzere, bir yanlı tahmin edici iken, $\left[\frac{n}{(n-1)}\right] \mathbf{S}_n$, $\boldsymbol{\Sigma}$ nın bir yansız tahmin edicisidir.

İspat. $\bar{\mathbf{x}} = (\mathbf{x}_1 + \mathbf{x}_2 + \dots + \mathbf{x}_n)/n$ dir. \mathbf{X} ve \mathbf{Y} aynı boyutlu rasgele matrisler ve \mathbf{A} ve \mathbf{B} sabitlerin çarpılır matrisleri olsun.

$$\begin{aligned} E(\mathbf{X} + \mathbf{Y}) &= E(\mathbf{X}) + E(\mathbf{Y}) \\ E(\mathbf{AXB}) &= \mathbf{A}E(\mathbf{X})\mathbf{B} \end{aligned} \quad (4.115)$$

özelliği vardır. (4.115) deki beklenen değer açılımının tekrarlı kullanımı,

$$\begin{aligned} E(\bar{\mathbf{x}}) &= E\left(\frac{1}{n}\mathbf{x}_1 + \frac{1}{n}\mathbf{x}_2 + \dots + \frac{1}{n}\mathbf{x}_n\right) \\ &= E\left(\frac{1}{n}\mathbf{x}_1\right) + E\left(\frac{1}{n}\mathbf{x}_2\right) + \dots + E\left(\frac{1}{n}\mathbf{x}_n\right) \\ &= \frac{1}{n}E(\mathbf{x}_1) + \frac{1}{n}E(\mathbf{x}_2) + \dots + \frac{1}{n}E(\mathbf{x}_n) \\ &= \frac{1}{n}\boldsymbol{\mu} + \frac{1}{n}\boldsymbol{\mu} + \dots + \frac{1}{n}\boldsymbol{\mu} \\ &= \boldsymbol{\mu} \end{aligned} \quad (4.116)$$

yü verir. Bundan sonra,

$$\begin{aligned} (\bar{\mathbf{x}} - \boldsymbol{\mu})(\bar{\mathbf{x}} - \boldsymbol{\mu})' &= \left(\frac{1}{n}\sum_{j=1}^n(\mathbf{x}_j - \boldsymbol{\mu})\right)\left(\frac{1}{n}\sum_{l=1}^n(\mathbf{x}_l - \boldsymbol{\mu})'\right) \\ &= \frac{1}{n^2}\sum_{j=1}^n\sum_{l=1}^n(\mathbf{x}_j - \boldsymbol{\mu})(\mathbf{x}_l - \boldsymbol{\mu})' \end{aligned} \quad (4.117)$$

dür. Bu nedenle,

$$Kov(\bar{\mathbf{x}}) = E(\bar{\mathbf{x}} - \boldsymbol{\mu})(\bar{\mathbf{x}} - \boldsymbol{\mu})' = \frac{1}{n^2}\left(\sum_{j=1}^n\sum_{l=1}^n E(\mathbf{x}_j - \boldsymbol{\mu})(\mathbf{x}_l - \boldsymbol{\mu})'\right) \quad (4.118)$$

dür.

\mathbf{x}_j nin her bir bileşeni ve \mathbf{x}_l nin her bir bileşeni arasındaki kovaryans elemanı $j \neq l$ için $E(\bar{\mathbf{x}} - \boldsymbol{\mu})(\bar{\mathbf{x}} - \boldsymbol{\mu})'$ deki her bir eleman sıfırdır ve bunlar bağımsızdır. Bu nedenle,

$$Kov(\bar{\mathbf{x}}) = \frac{1}{n^2}\left(\sum_{j=1}^n E(\mathbf{x}_j - \boldsymbol{\mu})(\mathbf{x}_j - \boldsymbol{\mu})'\right) \quad (4.119)$$

dür. $\Sigma = E(\mathbf{x}_j - \boldsymbol{\mu})(\mathbf{x}_j - \boldsymbol{\mu})'$ her bir \mathbf{x}_j için ortak kitle kovaryans matrisi olduğundan,

$$\begin{aligned}
Kov(\bar{\mathbf{x}}) &= \frac{1}{n^2} \left(\sum_{j=1}^n E(\mathbf{x}_j - \boldsymbol{\mu})(\mathbf{x}_j - \boldsymbol{\mu})' \right) = \frac{1}{n^2} \underbrace{(\Sigma + \Sigma + \dots + \Sigma)}_{n\text{-tane terim}} \\
&= \frac{1}{n^2} (n \Sigma) = \left(\frac{1}{n} \right) \Sigma
\end{aligned} \tag{4.120}$$

elde ederiz.

\mathbf{S}_n nin beklenen deęerini elde etmek için, ilk olarak, $(x_{ji} - \bar{x}_i)(x_{jk} - \bar{x}_k)$ nın, $(\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_j - \bar{\mathbf{x}})'$ nün (i, k) -yıncı elemanı olduęuna dikkat ederiz. Bu durumda kareler ve çarpaz çarpımların toplamlarını gösteren matris,

$$\sum_{j=1}^n (\mathbf{x}_j - \bar{\mathbf{x}}) = \mathbf{0} \quad \text{ve} \quad n\bar{\mathbf{x}} = \sum_{i=1}^n (\mathbf{x}_i) \tag{4.121}$$

olduęundan,

$$\begin{aligned}
\sum_{j=1}^n (\mathbf{x}_j - \bar{\mathbf{x}})(\mathbf{x}_j - \bar{\mathbf{x}})' &= \sum_{j=1}^n (\mathbf{x}_j - \bar{\mathbf{x}})\mathbf{x}_j' + \left(\sum_{j=1}^n (\mathbf{x}_j - \bar{\mathbf{x}}) \right) (-\bar{\mathbf{x}})' \\
&= \sum_{j=1}^n \mathbf{x}_j \mathbf{x}_j' - n\bar{\mathbf{x}}\bar{\mathbf{x}}'
\end{aligned} \tag{4.122}$$

olarak yazılabilir. Bu nedenle, onun beklenen deęeri

$$E \left(\sum_{j=1}^n \mathbf{x}_j \mathbf{x}_j' - n\bar{\mathbf{x}}\bar{\mathbf{x}}' \right) = \sum_{j=1}^n E(\mathbf{x}_j \mathbf{x}_j') - nE(\bar{\mathbf{x}}\bar{\mathbf{x}}') \tag{4.123}$$

dür. $E(\mathbf{V}) = \boldsymbol{\mu}_V$ ve $Kov(\mathbf{V}) = \Sigma_V$ olmak üzere, herhangi bir \mathbf{V} rasgele vektörü için, $E(\mathbf{V}\mathbf{V}') = \Sigma_V + \boldsymbol{\mu}_V \boldsymbol{\mu}_V'$ elde ederiz. Bu sonuçları kullanarak,

$$\sum_{j=1}^n E(\mathbf{x}_j \mathbf{x}_j') - nE(\bar{\mathbf{x}}\bar{\mathbf{x}}') = n \Sigma - n\boldsymbol{\mu}\boldsymbol{\mu}' - n \left(\frac{1}{n} \Sigma + \boldsymbol{\mu}\boldsymbol{\mu}' \right) = (n-1) \Sigma \tag{4.124}$$

ve bu nedenle, $\mathbf{S}_n = (1/n) \left(\sum_{j=1}^n \mathbf{x}_j \mathbf{x}_j' - n\bar{\mathbf{x}}\bar{\mathbf{x}}' \right)$ olduęundan,

$$E(\mathbf{S}_n) = \frac{(n-1)}{n} \Sigma \tag{4.125}$$

olduęu hemen görülür.

Sonuç 4.3; $[n/(n-1)]\mathbf{S}_n$ nin (i, k) -yıncı elemanının, yani $(n-1)^{-1} \sum_{j=1}^n (x_{ji} - \bar{x}_i)(x_{jk} - \bar{x}_k)$ nin, σ_{ik} nin bir yansız tahmin edicisi olduğunu gösterir. Bununla beraber bir bölen olarak n ya da $n-1$ ile hesaplanan bireysel (farklı) $\sqrt{s_{ii}}$ örneklem standart sapmaları, karşılık gelen $\sqrt{\sigma_{ik}}$ kitle niceliklerinin yansız tahmin edicileri değildir. Bundan başka, r_{ik} korelasyon katsayıları ρ_{ik} kitle niceliklerin yansız tahmin edicileri değildir. Bununla beraber, $yanE(\sqrt{s_{ii}}) - \sqrt{\sigma_{ik}}$ veya $E(r_{ik}) - \rho_{ik}$, eğer n örneklem büyüklüğü bir dereceye kadar büyükse, genellikle göz ardı edilebilir.

yan 'ın nedeni, örneklem varyans-kovaryans matrisinin azıcık değiştirilmiş tanımını harekete geçirir. Sonuç 4.3 bize, Σ nın yansız bir \mathbf{S} tahmin edicisini verir.

4.2.2.1. (Yansız) Örneklem Varyans-Kovaryans Matrisi

$$\mathbf{S} = \left(\frac{n}{n-1} \right) \mathbf{S}_n = \frac{1}{n-1} \sum_{j=1}^n (\mathbf{x}_j - \bar{\mathbf{x}})(\mathbf{x}_j - \bar{\mathbf{x}})' \quad (4.126)$$

Burada bir altindissiz \mathbf{S} , $(n-1)^{-1} \sum_{j=1}^n (x_{ji} - \bar{x}_i)(x_{jk} - \bar{x}_k)$ (i, k) -yıncı elemana sahiptir.

Örneklem kovaryansının bu tanımı birçok çok değişkenli test istatistiğinde yaygın bir şekilde kullanılır. Bu nedenle, bundan böyle örneklem kovaryans matrisi olarak \mathbf{S}_n kullanılacaktır.

4.2.3. Genelleştirilmiş Varyans

Bir tek değişkenle örneklem varyansı, çoğu kez o değişken üzerindeki ölçümlerde değişimin miktarını tanımlamak için kullanılır. Her bir birim üzerinde p -sayıda değişken gözlemlendiğinde değişim,

$$\mathbf{S}_n = \begin{bmatrix} s_{11} & s_{12} & \cdots & s_{1p} \\ s_{21} & s_{22} & \cdots & s_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ s_{n1} & s_{n2} & \cdots & s_{np} \end{bmatrix} = \left\{ s_{ik} = \frac{1}{n-1} \sum_{j=1}^n (x_{ji} - \bar{x}_i)(x_{jk} - \bar{x}_k) \right\} \quad (4.127)$$

örneklem varyans-kovaryans matrisiyle tanımlanır. Örneklem kovaryans matrisi p -sayıda varyansı ve $\frac{1}{2}p(p-1)$ -sayıda potansiyel olarak farklı kovaryansı içerir.

Bazen \mathbf{S} ile ifade edilen değişim için bir tek sayısal değeri tahsis etmek istenebilir. Bir değer için bir seçim $p=1$ olduğunda, bir tek karakteristiğin alışılmış örneklem

varyansına indirgenen S nin determinantıdır. Bu determinanta genelleştirilmiş örneklem varyansı denir.

$$\text{Genelleştirilmiş Örneklem Varyansı} = |S| \quad (4.128)$$

dir.

Örnek 4.6. Bir Genelleştirilmiş Varyansı Hesaplama

Ordu ilinde bulunan bir fındık fabrikasında 2013-2017 yıllarının aralık ayında çalışan işçi sayısı ve çalışan başına kazanç aşağıda Çizelge 4.1 de verilmiştir.

Çizelge 4.1. Fabrikada çalışan işçi sayısı ve kişi başına düşen kazanç.

<u>YIL</u>	<u>CALIŞAN SAYISI (KİŞİ)</u>	<u>KİŞİ BAŞINA DÜŞEN KAZANÇ (TL)</u>
2017	78	1,033,132.457 TL
2016	57	1,062,882.890 TL
2015	81	1,324,114.990 TL
2014	55	824,509.375 TL
2013	75	536,710.989 TL

Verilerden elde edilen örneklem kovaryans matrisi,

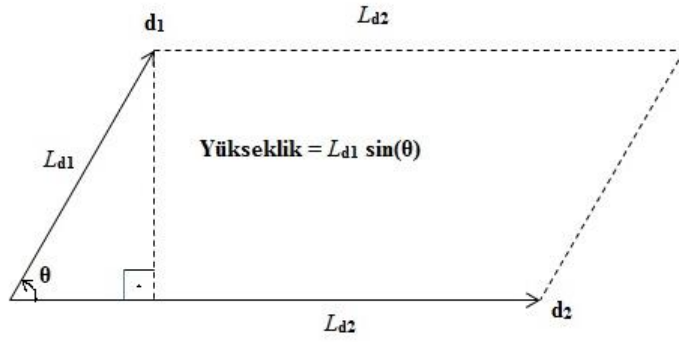
$$S = \begin{bmatrix} 120.016 & 8020732.2 \\ 8020732.2 & 345973546 \times 10^{11} \end{bmatrix}$$

dir. Genelleştirilmiş varyansın değerini bulalım. Bu durumda,

$$|S| = 4.152 \times 10^{21}$$

hesaplarız.

Genelleştirilmiş örneklem varyansı tüm varyans ve kovaryanslar üzerindeki bilgiyi bir tek sayı olarak yazmanın bir yöntemini ortaya koyar. Şüphesiz, $p > 1$ olduğunda, örneklem hakkındaki bir takım bilgi yöntemde kaybolur. $|S|$ nin bir geometrik yorumu betimleyici (tanımlayıcı) bir özet olarak bize onun sağlamlığını ve zayıflığını değerlendirmek için yardımcı olur. Basit geometriden,



Şekil 4.7. L_{d_1} ve L_{d_2} ile ikizkenar yamuğun gösterimi.

Şekil 4.7 yi elde ederiz ve ikizkenar yamuğun alanı $|L_{d_1} \sin(\theta)|L_{d_2}$ dir. $\cos^2 \theta + \sin^2 \theta = 1$ olduğundan bu alanı,

$$Alan = L_{d_1} L_{d_2} \sqrt{1 - \cos^2(\theta)}$$

olarak ifade edebiliriz. (4.107) ve (4.112) den

$$L_{d_1} = \sqrt{\sum_{j=1}^n (x_{j1} - \bar{x}_1)^2} = \sqrt{(n-1)s_{11}}$$

$$L_{d_2} = \sqrt{\sum_{j=1}^n (x_{j2} - \bar{x}_2)^2} = \sqrt{(n-1)s_{22}}$$

ve

$$\cos \theta = r_{12}$$

dir.

Bu nedenle,

$$Alan = (n-1) \sqrt{s_{11}} \sqrt{s_{22}} \sqrt{1 - r_{12}^2} = (n-1) \sqrt{s_{11} s_{22} (1 - r_{12}^2)} \quad (4.129)$$

dir. Aynı zamanda,

$$\begin{aligned} |\mathbf{S}| &= \begin{bmatrix} s_{11} & s_{12} \\ s_{12} & s_{22} \end{bmatrix} = \begin{bmatrix} s_{11} & \sqrt{s_{11}} \sqrt{s_{22}} r_{12} \\ \sqrt{s_{11}} \sqrt{s_{22}} r_{12} & s_{22} \end{bmatrix} \\ &= s_{11} s_{22} - s_{11} s_{22} r_{12}^2 = s_{11} s_{22} (1 - r_{12}^2) \end{aligned} \quad (4.130)$$

dir. Eğer (4.129) u ve (4.130) u karşılaştırırsak,

$$|\mathbf{S}| = (\text{alan})^2 / (n-1)^2 \quad (4.131)$$

olduğunu görürüz.

4.2.4. Korelasyon Katsayısının Farklı Geometrik Yorumları

Bu kısımda, korelasyon katsayısının farklı geometrik yorumları elde edilmiştir. Özellikle bu geometrik yorumlar Pearson korelasyon katsayısı üzerinde yoğunlaşmıştır. Verilen geometrik yorumlar, iki veri vektörü ve onların regresyon doğruları üzerine kurulan paralelkenarların alanları yardımıyla doğrulanmıştır. Bu geometrik yorumların tümü rasgele üretilen veri vektörlerini kullanarak açıklanmıştır.

$\mathbf{X} = [x_i]$, $\mathbf{Y} = [y_i]$, $i=1,2,\dots,n$ olmak üzere,

$$\mathbf{X} = [x_1, x_2, \dots, x_n]' \text{ ve } \mathbf{Y} = [y_1, y_2, \dots, y_n]' \quad (4.132)$$

veri vektörlerini göz önüne alalım.

\mathbf{X} ve \mathbf{Y} n -boyutlu uzayda orijinden, uzayın sırasıyla, (x_1, x_2, \dots, x_n) ve (y_1, y_2, \dots, y_n) noktalarına yönlendirilen vektörleri gösterebiliriz.

$\mathbf{j} = [1, 1, \dots, 1]'$ vektörü, bu uzayın eksenleriyle eşit açı yapan bir vektör olsun. θ ; \mathbf{X} ve \mathbf{Y} veri vektörleri arasında pozitif yöndeki açı, yani, saatin dönme yönünün ters yönündeki açı olsun. \mathbf{X} ve \mathbf{Y} vektörlerinin \mathbf{j} vektörü ile pozitif yönde yaptığı açılar sırasıyla θ_1 ve θ_2 olsun. \mathbf{X} ve \mathbf{Y} vektörlerinin \mathbf{j} vektörü üzerindeki izdüşümleri sırasıyla aşağıdaki gibi verilir:

$$\mathbf{j}(\mathbf{j}\mathbf{j})^{-1}\mathbf{j}'\mathbf{X} = \mathbf{j}\bar{\mathbf{X}} \text{ ve } \mathbf{j}(\mathbf{j}\mathbf{j})^{-1}\mathbf{j}'\mathbf{Y} = \mathbf{j}\bar{\mathbf{Y}}. \quad (4.133)$$

Bu durumda, $s_{\mathbf{XY}}$ örneklem kovaryansı ve $s_{\mathbf{X}}^2$, $s_{\mathbf{Y}}^2$ örneklem varyansları:

$$s_{\mathbf{XY}} = \frac{(\mathbf{X} - \mathbf{j}\bar{\mathbf{X}})'(\mathbf{Y} - \mathbf{j}\bar{\mathbf{Y}})}{n-1}, \quad (4.134)$$

$$s_{\mathbf{X}} = \frac{\|\mathbf{X} - \mathbf{j}\bar{\mathbf{X}}\|}{\sqrt{n-1}} \quad (4.135)$$

ve

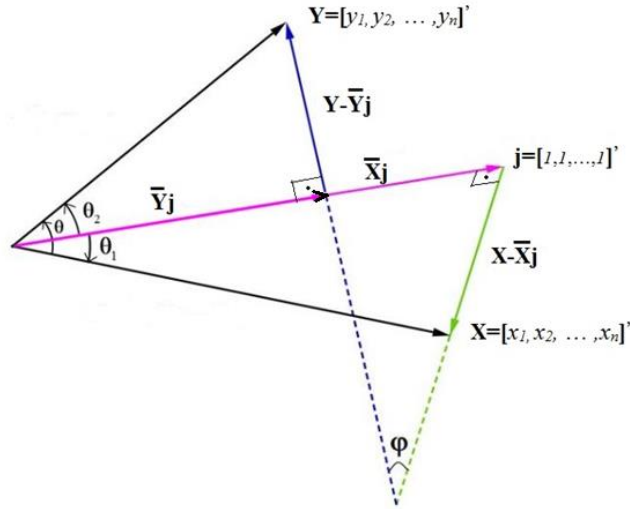
$$s_Y = \frac{\|Y - j\bar{Y}\|}{\sqrt{n-1}} \quad (4.136)$$

dir. Bu takdirde, ρ_{XY} kitle korelasyon katsayısının tahmini olan Pearson r_{XY} örneklem korelasyon katsayısı

$$r_{XY} = \frac{s_{XY}}{s_X s_Y} = \frac{(X - j\bar{X})'(Y - j\bar{Y})}{\|X - j\bar{X}\| \|Y - j\bar{Y}\|} = \frac{X'Y - n\bar{X}\bar{Y}}{\|X - j\bar{X}\| \|Y - j\bar{Y}\|} \quad (4.137)$$

olarak hesaplanabilir.

Bu nedenle, Şekil 4.8 de görüldüğü gibi, r_{XY} örneklem korelasyon katsayısı $(X - j\bar{X})$ ve $(Y - j\bar{Y})$ vektörleri arasındaki φ açısının kosinüsü, yani $r_{XY} = \cos \varphi$ dir.



Şekil 4.8. X , Y , $X - j\bar{X}$, $Y - j\bar{Y}$ vektörlerinin ve θ , θ_1, θ_2 , φ açılarının n -boyutlu uzayda şekilsel gösterimleri.

Şimdi,

$$T = X - j\bar{X}, \text{ yani, } X = T + j\bar{X} \quad (4.138)$$

ve

$$Z = Y - j\bar{Y}, \text{ yani, } Y = Z + j\bar{Y} \quad (4.139)$$

olsun.

Bu durumda,

$$\begin{aligned}\mathbf{X}'\mathbf{Y} &= (\mathbf{T} + \mathbf{j}\bar{\mathbf{X}})'(\mathbf{Z} + \mathbf{j}\bar{\mathbf{Y}}) = \mathbf{T}'\mathbf{Z} + n\bar{\mathbf{T}}\bar{\mathbf{Y}} + n\bar{\mathbf{X}}\bar{\mathbf{Z}} + n\bar{\mathbf{X}}\bar{\mathbf{Y}} \\ &= \|\mathbf{X}\|\|\mathbf{Y}\|\cos\theta\end{aligned}\quad (4.140)$$

dır.

$$\bar{\mathbf{T}} = \frac{\mathbf{j}'(\mathbf{X} - \mathbf{j}\bar{\mathbf{X}})}{n} = \frac{\mathbf{j}'\mathbf{X} - \mathbf{j}'\mathbf{j}\bar{\mathbf{X}}}{n} = \frac{n\bar{\mathbf{X}} - n\bar{\mathbf{X}}}{n} = 0 \quad (4.141)$$

ve

$$\bar{\mathbf{Z}} = \frac{\mathbf{j}'(\mathbf{Y} - \mathbf{j}\bar{\mathbf{Y}})}{n} = \frac{\mathbf{j}'\mathbf{Y} - \mathbf{j}'\mathbf{j}\bar{\mathbf{Y}}}{n} = \frac{n\bar{\mathbf{Y}} - n\bar{\mathbf{Y}}}{n} = 0 \quad (4.142)$$

olduğundan, (4.140) ifadesi aşağıdaki biçimi alır:

$$\mathbf{T}'\mathbf{Z} + n\bar{\mathbf{X}}\bar{\mathbf{Y}} = \|\mathbf{X}\|\|\mathbf{Y}\|\cos\theta. \quad (4.143)$$

(4.138), (4.139) ve (4.140) bağıntılarını kullanarak $\cos\theta$ için aşağıdaki ifadeyi elde ederiz:

$$\begin{aligned}\cos\theta &= \frac{\mathbf{T}'\mathbf{Z}}{\|\mathbf{X}\|\|\mathbf{Y}\|} + \frac{n\bar{\mathbf{X}}\bar{\mathbf{Y}}}{\|\mathbf{X}\|\|\mathbf{Y}\|} \\ &= \frac{(\mathbf{X} - \mathbf{j}\bar{\mathbf{X}})'(\mathbf{Y} - \mathbf{j}\bar{\mathbf{Y}})}{\|\mathbf{X}\|\|\mathbf{Y}\|} + \frac{n\bar{\mathbf{X}}\bar{\mathbf{Y}}}{\|\mathbf{X}\|\|\mathbf{Y}\|}.\end{aligned}\quad (4.144)$$

(4.137) ye göre,

$$r_{xy} \|\mathbf{X} - \mathbf{j}\bar{\mathbf{X}}\|\|\mathbf{Y} - \mathbf{j}\bar{\mathbf{Y}}\| = (\mathbf{X} - \mathbf{j}\bar{\mathbf{X}})'(\mathbf{Y} - \mathbf{j}\bar{\mathbf{Y}}) \quad (4.145)$$

olduğu görülür.

$$\sqrt{n}\bar{\mathbf{X}} = \|\mathbf{j}\bar{\mathbf{X}}\| \quad \text{ve} \quad \sqrt{n}\bar{\mathbf{Y}} = \|\mathbf{j}\bar{\mathbf{Y}}\| \quad (4.146)$$

olduğundan, (4.144) bağıntısı

$$\cos\theta = \frac{r_{xy} \|\mathbf{X} - \mathbf{j}\bar{\mathbf{X}}\|\|\mathbf{Y} - \mathbf{j}\bar{\mathbf{Y}}\|}{\|\mathbf{X}\|\|\mathbf{Y}\|} + \frac{\|\mathbf{j}\bar{\mathbf{X}}\| \|\mathbf{j}\bar{\mathbf{Y}}\|}{\|\mathbf{X}\| \|\mathbf{Y}\|} \quad (4.147)$$

olarak yeniden yazılabilir.

Şekil 4.8 den,

$$\begin{aligned}
-\sin \theta_1 &= \frac{\|\mathbf{X} - \mathbf{j}\bar{\mathbf{X}}\|}{\|\mathbf{X}\|}, & \sin \theta_2 &= \frac{\|\mathbf{Y} - \mathbf{j}\bar{\mathbf{Y}}\|}{\|\mathbf{Y}\|} \\
\cos \theta_1 &= \frac{\|\mathbf{j}\bar{\mathbf{X}}\|}{\|\mathbf{X}\|}, & \cos \theta_2 &= \frac{\|\mathbf{j}\bar{\mathbf{Y}}\|}{\|\mathbf{Y}\|}
\end{aligned} \tag{4.148}$$

elde ederiz.

Bu nedenle, (4.147) bağıntısı θ_1, θ_2 açılara ve $r_{\mathbf{X}\mathbf{Y}}$ ye göre yeniden biçimlendirilebilir:

$$\cos \theta = -r_{\mathbf{X}\mathbf{Y}} \sin \theta_1 \sin \theta_2 + \cos \theta_1 \cos \theta_2. \tag{4.149}$$

Bilindiği gibi, $r_{\mathbf{X}\mathbf{Y}} = 1$ olduğunda, \mathbf{X} ve \mathbf{Y} gözlem vektörleri arasında mükemmel bir pozitif ilişki (lineer ilişki) vardır. $r_{\mathbf{X}\mathbf{Y}} = -1$ olduğunda, \mathbf{X} ve \mathbf{Y} gözlem vektörleri arasında mükemmel bir negatif lineer ilişki vardır. $r_{\mathbf{X}\mathbf{Y}} = 0$ olduğunda, \mathbf{X} ve \mathbf{Y} gözlem vektörleri arasında bir lineer ilişki yoktur.

$r_{\mathbf{X}\mathbf{Y}} = 1$ ise, (4.149) da $r_{\mathbf{X}\mathbf{Y}}$ yi yerine koyma

$$\cos \theta = -\sin \theta_1 \sin \theta_2 + \cos \theta_1 \cos \theta_2 = \cos(\theta_1 + \theta_2) \tag{4.150}$$

olduğunu ortaya koyar.

(4.150) nin her iki tarafına arccos ü uygulayarak, $\theta = \theta_1 + \theta_2$ olduğu görülebilir. Tersine olarak, eğer $\theta = \theta_1 + \theta_2$ ise, bu takdirde $\cos \theta = -\sin \theta_1 \sin \theta_2 + \cos \theta_1 \cos \theta_2 = \cos(\theta_1 + \theta_2)$ elde ederiz. Bu nedenle, (4.149) dan, $r_{\mathbf{X}\mathbf{Y}} = 1$ yazarız, yani \mathbf{X} ve \mathbf{Y} gözlem vektörleri arasında mükemmel bir pozitif ilişki (lineer ilişki) vardır.

Benzer şekilde, $r_{\mathbf{X}\mathbf{Y}} = -1$ ise, bu takdirde (4.160) aşağıdaki biçimi alır:

$$\cos \theta = \sin \theta_1 \sin \theta_2 + \cos \theta_1 \cos \theta_2 = \cos(\theta_1 - \theta_2). \tag{4.151}$$

Yine, (4.151) in her iki tarafına arccos ü uygulayarak, $\theta = \theta_1 - \theta_2$ elde ederiz. Tersine olarak, eğer $\theta = \theta_1 - \theta_2$ ise, bu takdirde $\cos \theta = \sin \theta_1 \sin \theta_2 + \cos \theta_1 \cos \theta_2 = \cos(\theta_1 - \theta_2)$ yazarız. Bu nedenle, (4.149) dan, $r_{\mathbf{X}\mathbf{Y}} = -1$ yazarız, yani, \mathbf{X} ve \mathbf{Y} gözlem vektörleri arasında mükemmel bir negatif ilişki (lineer ilişki) vardır.

Eğer $r_{\mathbf{X}\mathbf{Y}} = 0$, bu takdirde (4.149) da $r_{\mathbf{X}\mathbf{Y}}$ yi yerine koyma,

$$\cos \theta = \cos \theta_1 \cos \theta_2 = \frac{\cos(\theta_1 + \theta_2) + \cos(\theta_1 - \theta_2)}{2} \quad (4.152)$$

olduğunu ortaya koyar.

Tersine olarak, eğer (4.152) bağıntısı mevcutsa, bu takdirde (4.149) dan $r_{XY} = 0$ elde ederiz, yani, \mathbf{X} ve \mathbf{Y} gözlem vektörleri arasında bir lineer ilişki yoktur.

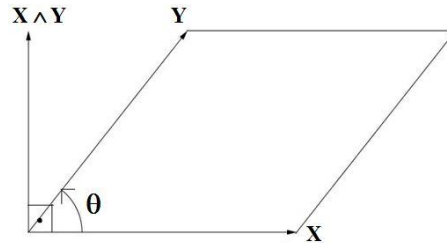
Sonuç olarak, \mathbf{X} ve \mathbf{Y} arasındaki lineer ilişki r_{XY} ile açıklandığı gibi, θ_1 ve θ_2 nin trigonometrisi ile de açıklanır.

Not: θ_1 ve θ_2 den en az biri $\pi/2$ veya $3\pi/2$ olduğunda, \mathbf{X} ve \mathbf{Y} nin biri diğerine dik olmalıdır, yani $\cos \theta = 0$, bu nedenle $r_{XY} = 0$ olmalıdır.

Şimdi, \mathbf{X} ve \mathbf{Y} vektörleri arasındaki lineer ilişkiyi, bu iki vektör üzerine kurulan paralelkenarın alanıyla ifade edelim. " \wedge " çapraz çarpımı (veya vektörel çarpımı) göstermek üzere, bu paralelkenarın alanı,

$$\|\mathbf{X} \wedge \mathbf{Y}\| = \|\mathbf{X}\| \|\mathbf{Y}\| \sin \theta = \|\mathbf{X}\| \|\mathbf{Y}\| \sqrt{1 - \cos^2 \theta}$$

olarak hesaplanabilir. Eğer $\theta = \pi/2$ (veya $3\pi/2$) ise, bu takdirde bu paralelkenarın alanı $\|\mathbf{X}\| \|\mathbf{Y}\|$ dir. Bir başka deyişle paralelkenar; kenarları $\|\mathbf{X}\|$ ve $\|\mathbf{Y}\|$ olan bir dikdörtgen olacaktır. Bu durumda, $\|\mathbf{X}\|$ ve $\|\mathbf{Y}\|$ nin her birinin diğerine dik olması $r_{XY} = 0$ olduğunu söyler.



Şekil 4.9. \mathbf{X} ve \mathbf{Y} vektörleri üzerine kurulan paralelkenar.

4.2.4.1. Standartlaştırılmış Vektörlerin Kullanımı

Şimdi, $\mathbf{x} = \frac{\mathbf{X} - j\bar{\mathbf{X}}}{\|\mathbf{X} - j\bar{\mathbf{X}}\|}$ ve $\mathbf{y} = \frac{\mathbf{Y} - j\bar{\mathbf{Y}}}{\|\mathbf{Y} - j\bar{\mathbf{Y}}\|}$ vektörlerini göz önüne alalım. Burada

$\|\mathbf{x}\| = 1$ ve $\|\mathbf{y}\| = 1$ dir. Bu durumda (4.137) bağıntısından

$$r = \mathbf{x} \cdot \mathbf{y} = \cos \varphi \quad (4.153)$$

olduğu görülür. Burada φ ; standartlaştırılmış x ve y vektörleri arasındaki açıdır. y nin x üzerindeki (dik) izdüşümü ve x in y üzerindeki (dik) izdüşümü sırasıyla \hat{y} ve \hat{x} olsun. Bu takdirde,

$$\hat{y} = x(x'x)^{-1}x'y = x \frac{1}{\|x\|^2} x'y = xr \quad \text{ve} \quad \hat{x} = y(y'y)^{-1}y'x = y \frac{1}{\|y\|^2} y'x = yr \quad (4.154)$$

elde ederiz.

$$\|\hat{x}\| = r \quad \text{ve} \quad \|\hat{y}\| = r \quad (4.155)$$

ve \hat{x} ile \hat{y} arasındaki açı x ile y arasındaki açıya eşit olduğundan, aşağıdaki bağıntı sağlanır:

$$\cos \varphi = \frac{\hat{x}'\hat{y}}{\|\hat{x}\|\|\hat{y}\|} = \frac{r^2 x'y}{rr} = x'y = r. \quad (4.156)$$

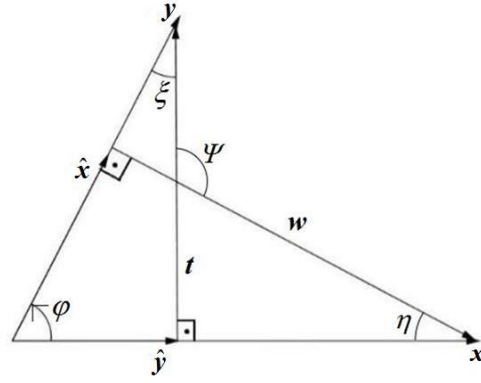
$x - \hat{x} = w$ ve $y - \hat{y} = t$ olsun. Bu durumda,

$$\begin{aligned} \|w\| &= \|x - \hat{x}\| = [(x - \hat{x})'(x - \hat{x})]^{1/2} = [x'x - x'\hat{x} - \hat{x}'x + \hat{x}'\hat{x}]^{1/2} \\ &= (1 - 2 \underbrace{x'\hat{x}}_{=r^2} + \underbrace{\hat{x}'\hat{x}}_{=r^2})^{1/2} = \sqrt{1 - r^2} \end{aligned} \quad (4.157)$$

ve

$$\begin{aligned} \|t\| &= \|y - \hat{y}\| = [(y - \hat{y})'(y - \hat{y})]^{1/2} = [y'y - y'\hat{y} - \hat{y}'y + \hat{y}'\hat{y}]^{1/2} \\ &= (1 - 2 \underbrace{y'\hat{y}}_{=r^2} + \underbrace{\hat{y}'\hat{y}}_{=r^2})^{1/2} = \sqrt{1 - r^2} \end{aligned} \quad (4.158)$$

dir. Aşağıdaki şekli göz önüne alalım.



Şekil 4.10. \hat{y} ve \hat{x} vektörleri ve φ , ψ , η ve ξ açıları.

$$\begin{aligned} t'w &= w't = (x - \hat{x})'(y - \hat{y}) = x'y - x'\hat{y} - \hat{x}'y + \hat{x}'\hat{y} \\ &= r - r - r + r^3 = r^3 - r = r(r^2 - 1) \end{aligned}$$

olduğundan,

$$\cos \psi = \frac{r(r^2 - 1)}{\sqrt{1 - r^2} \sqrt{1 - r^2}} = \frac{-r(1 - r^2)}{(1 - r^2)} = -r \quad (4.159)$$

elde ederiz.

$\varphi + \psi = \pi$ olduğundan,

$$\begin{aligned} \cos \varphi &= \cos(\pi - \psi) = \cos \pi \cos \psi + \underbrace{\sin \pi \sin \psi}_{=0} = -\cos \psi \\ \rightarrow \cos \psi &= -\cos \varphi = -r \end{aligned}$$

yazarız.

$x - \hat{x}$ ve x , $y - \hat{y}$ ve y arasındaki açılar sırasıyla η ve ξ olsun. Bu takdirde, $\sin \eta = \cos \varphi = r$ ve $\sin \xi = \cos \varphi = r$ dir. $\|x\| = \|y\| = 1$, $\|\hat{x}\| = \|\hat{y}\| = r$ olduğundan, Şekil 4.10 daki dik üçgenler eşittir.

\mathbf{X} ve \mathbf{Y} orijinal gözlem vektörleri için \mathbf{Y} nin \mathbf{X} üzerindeki basit lineer regresyonu

$$\mathbf{Y}_{n \times 1} = \alpha_1 \mathbf{j} + \beta_1 \mathbf{X}_{n \times 1} + (\varepsilon_1)_{n \times 1}, \quad \varepsilon_1 : N(0, \sigma^2 \mathbf{I}_n) \quad (4.160)$$

ve \mathbf{X} in \mathbf{Y} üzerindeki basit lineer regresyonu

$$\mathbf{X} = \alpha_2 \mathbf{j} + \beta_2 \mathbf{Y} + \varepsilon_2, \quad \varepsilon_2 : N(0, \sigma^2 \mathbf{I}_n) \quad (4.161)$$

dir. (4.160) ve (4.161) e alışılmış en küçük kareler yöntemini uygulayarak, aşağıdaki bağıntıları elde ederiz:

$$\hat{\mathbf{Y}} = \mathbf{j}\bar{\mathbf{Y}} - b_1\bar{\mathbf{X}}\mathbf{j} + b_1\mathbf{X} \rightarrow \hat{\mathbf{Y}} - \mathbf{j}\bar{\mathbf{Y}} = b_1(\mathbf{X} - \mathbf{j}\bar{\mathbf{X}}), \quad (4.162)$$

$$\hat{\mathbf{X}} - \mathbf{j}\bar{\mathbf{X}} = b_2(\mathbf{Y} - \mathbf{j}\bar{\mathbf{Y}}). \quad (4.163)$$

Burada $\hat{\mathbf{Y}}$ ve $\hat{\mathbf{X}}$ sırasıyla \mathbf{Y} ve \mathbf{X} in (4.160) ve (4.161) regresyonlarından elde edilen en küçük kareler tahminleri ve b_1 ve b_2 de β_1 ve β_2 nin en küçük kareler tahminleridir.

Bu tahminde, b_1 ve b_2 sırasıyla $r \frac{s_Y}{s_X} = r \frac{\|\mathbf{Y} - \mathbf{j}\bar{\mathbf{Y}}\|}{\|\mathbf{X} - \mathbf{j}\bar{\mathbf{X}}\|}$ ve $r \frac{s_X}{s_Y} = r \frac{\|\mathbf{X} - \mathbf{j}\bar{\mathbf{X}}\|}{\|\mathbf{Y} - \mathbf{j}\bar{\mathbf{Y}}\|}$ olarak hesaplanabilir.

(4.162) den,

$$\frac{\hat{\mathbf{Y}} - \mathbf{j}\bar{\mathbf{Y}}}{\|\mathbf{Y} - \mathbf{j}\bar{\mathbf{Y}}\|} = r \frac{\mathbf{X} - \mathbf{j}\bar{\mathbf{X}}}{\|\mathbf{X} - \mathbf{j}\bar{\mathbf{X}}\|} \quad (4.164)$$

ve benzer şekilde (4.163) den,

$$\frac{\hat{\mathbf{X}} - \mathbf{j}\bar{\mathbf{X}}}{\|\mathbf{X} - \mathbf{j}\bar{\mathbf{X}}\|} = r \frac{\mathbf{Y} - \mathbf{j}\bar{\mathbf{Y}}}{\|\mathbf{Y} - \mathbf{j}\bar{\mathbf{Y}}\|} \quad (4.165)$$

elde ederiz. $\mathbf{x} = \frac{\mathbf{X} - \mathbf{j}\bar{\mathbf{X}}}{\|\mathbf{X} - \mathbf{j}\bar{\mathbf{X}}\|}$ ve $\mathbf{y} = \frac{\mathbf{Y} - \mathbf{j}\bar{\mathbf{Y}}}{\|\mathbf{Y} - \mathbf{j}\bar{\mathbf{Y}}\|}$ olduğundan, yine $\hat{\mathbf{x}} = r\mathbf{y}$ ve $\hat{\mathbf{y}} = r\mathbf{x}$ olduğu kolayca görülebilir.

$\hat{\mathbf{x}}$ ile $\hat{\mathbf{y}}$ arasındaki açı $\hat{\mathbf{x}} = \frac{\hat{\mathbf{X}} - \mathbf{j}\bar{\mathbf{X}}}{\|\mathbf{X} - \mathbf{j}\bar{\mathbf{X}}\|}$ ile $\hat{\mathbf{y}} = \frac{\hat{\mathbf{Y}} - \mathbf{j}\bar{\mathbf{Y}}}{\|\mathbf{Y} - \mathbf{j}\bar{\mathbf{Y}}\|}$ vektörleri arasındaki açı ile aynıdır.

Bu açı, ya

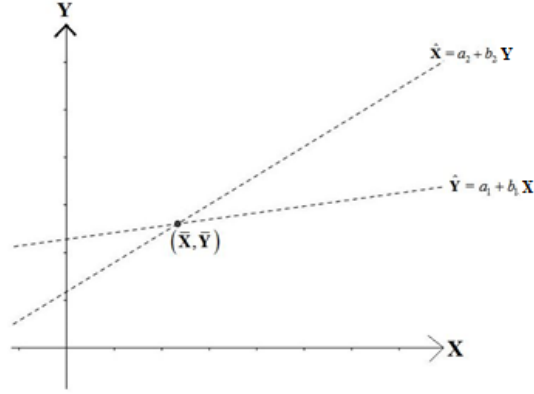
$$r \frac{\mathbf{X} - \mathbf{j}\bar{\mathbf{X}}}{\|\mathbf{X} - \mathbf{j}\bar{\mathbf{X}}\|} \text{ ve } r \frac{\mathbf{Y} - \mathbf{j}\bar{\mathbf{Y}}}{\|\mathbf{Y} - \mathbf{j}\bar{\mathbf{Y}}\|} \text{ ya da } r \frac{\mathbf{X} - \mathbf{j}\bar{\mathbf{X}}}{s_X} \text{ ve } r \frac{\mathbf{Y} - \mathbf{j}\bar{\mathbf{Y}}}{s_Y}$$

vektörleri arasındaki açı olmalıdır.

Bu durumda aşağıdaki bağıntı elde edilebilir:

$$\cos \varphi = \frac{\hat{\mathbf{x}} \cdot \hat{\mathbf{y}}}{\|\hat{\mathbf{x}}\| \|\hat{\mathbf{y}}\|} = \frac{\left[r^2 \frac{(\mathbf{X} - \mathbf{j}\bar{\mathbf{X}})' (\mathbf{Y} - \mathbf{j}\bar{\mathbf{Y}})}{s_{\mathbf{X}} s_{\mathbf{Y}}} \right]}{\left[r^2 \left\| \frac{(\mathbf{X} - \mathbf{j}\bar{\mathbf{X}})}{s_{\mathbf{X}}} \right\| \left\| \frac{(\mathbf{Y} - \mathbf{j}\bar{\mathbf{Y}})}{s_{\mathbf{Y}}} \right\| \right]} = \frac{r^3}{r^2} = r. \quad (4.166)$$

\mathbf{X} in \mathbf{Y} üzerindeki ve \mathbf{Y} nin \mathbf{X} üzerindeki regresyon doğruları $(\bar{\mathbf{X}}, \bar{\mathbf{Y}})$ noktasında kesişir. \mathbf{X} üzerinde \mathbf{Y} nin regresyon doğrusunun tahmini $\hat{\mathbf{Y}} = a_1 + b_1 \mathbf{X}$ ve \mathbf{Y} üzerinde \mathbf{X} in regresyon doğrusunun tahmini $\hat{\mathbf{X}} = a_2 + b_2 \mathbf{Y}$ olduğundan, aşağıdaki gösterim ortaya çıkar. Burada a_1 ve a_2 sırasıyla, (4.160) daki ve (4.161) deki α_1 in ve α_2 nin en küçük kareler tahminleridir.



Şekil 4.11. \mathbf{X} üzerinde \mathbf{Y} nin ve \mathbf{Y} üzerinde \mathbf{X} in regresyon doğruları.

Burada $b_1 = \frac{s_{\mathbf{XY}}}{s_{\mathbf{X}}^2}$ ve $a_1 = \bar{\mathbf{Y}} - b_1 \bar{\mathbf{X}}$ dir. $s_{\mathbf{XY}}$ ve $s_{\mathbf{X}}^2$ sırasıyla kovaryansın ve varyansın tahminleridir. Bunlar, \mathbf{X} ve \mathbf{Y} veri vektörlerini kullanarak

$$b_1 = \frac{\sum x_i y_i - n \bar{\mathbf{X}} \bar{\mathbf{Y}}}{\sum x_i^2 - n \bar{\mathbf{X}}^2} \quad \text{ve} \quad a_1 = \bar{\mathbf{Y}} - b_1 \bar{\mathbf{X}} \quad (4.167)$$

formüllerinden doğrudan doğruya elde edilir. Açıklayıcı değişkenler kontrol edildiğinde, en küçük kareler tahminleri; parametre tahmini ve gelecekteki tahmin için en iyi uyumun doğrusunu verir.

\mathbf{X} üzerinde \mathbf{Y} nin regresyon doğrusu ile \mathbf{Y} üzerinde \mathbf{X} in regresyon doğrusu aynıdır değildir. Şekil 4.11 iki \mathbf{X} ve \mathbf{Y} rasgele vektör değişkenleri için bu durumu açık olarak gösterir. \mathbf{X} bir rasgele vektör değişkeni olsa bile, notasyonu basitleştirmek için değişkenlere \mathbf{X} ve \mathbf{Y} (y_1 ve y_2 yerine) denilmesini

sürdürürüz. Bununla beraber, s_{XY} ifadesi; (b_1) ve (b_2) regresyon katsayıları için aynıdır. \mathbf{X} ; \mathbf{Y} üzerinde regresyonlandığında eğimin paydası s_Y^2 iken, \mathbf{Y} ; \mathbf{X} üzerinde regresyonlandığında eğimin paydası s_X^2 dir. Bundan başka, iki α_1 ve α_2 intercepti (integrasyon sabitleri) tahmin edildiğinde, $\bar{\mathbf{X}}$ ve $\bar{\mathbf{Y}}$ ortalamaları ters rolleri oynar. Regresyon yürütüldüğünde bu, açıklayıcı ve cevap değişkenlerini tanımlamanın önemini açık bir şekilde vurgular.

En küçük kareler regresyon doğruları sadece tüm gözlem noktaları aynı doğru üzerine düştüğünde (korelasyon=1) çakışabilir. $r_{XY} = \frac{s_{XY}}{s_X s_Y}$ olduğunu biliyoruz. Bu nedenle, eğer $r=1$ ise, bu takdirde $s_{XY} = s_X s_Y$ dir. Benzer şekilde, yeri değiştirilen grafikte (eksenlerin yerleri değiştirildiğinde) aynı doğruyu tanımlayan b_2 eğimi $s_X / s_Y = 1/b_1$ dir. Hesaplamalar vasıtasıyla,

$$b_1 = r \frac{s_Y}{s_X} \quad \text{ve} \quad b_2 = r \frac{s_X}{s_Y} \quad (4.168)$$

elde ederiz. Bu nedenle bu iki regresyon doğrusu arasındaki açı Θ_1 ise, bu takdirde m_1 ve m_2 regresyon doğrularının bilinen eğimleri olmak üzere, Θ_1 in tanjantı

$$\text{tg}\Theta_1 = \frac{m_2 - m_1}{1 + m_1 m_2}$$

formülünü kullanarak hesaplanır.

Bu durumda Θ_1 in tanjantı için,

$$\text{tg}\Theta_1 = \frac{1 - r^2}{r} \left(\frac{s_X s_Y}{s_X^2 + s_Y^2} \right) \quad (4.169)$$

ifadesini elde ederiz.

\mathbf{X} ve \mathbf{Y} vektörleri yukarıdaki gibi standartlaştırıldığında,

$$\mathbf{x} = \frac{\mathbf{X} - \mathbf{j}\bar{\mathbf{X}}}{\|\mathbf{X} - \mathbf{j}\bar{\mathbf{X}}\|} \quad \text{ve} \quad \mathbf{y} = \frac{\mathbf{Y} - \mathbf{j}\bar{\mathbf{Y}}}{\|\mathbf{Y} - \mathbf{j}\bar{\mathbf{Y}}\|}, \quad s_x^2 = 1, \quad s_y^2 = 1,$$

$\hat{\mathbf{x}} = r\mathbf{y}$ ve $\hat{\mathbf{y}} = r\mathbf{x}$ olduğunu gördük. Bu iki doğruyu aynı koordinat sistemi üzerinde çizdiğimizde, onlar (0,0) noktasında kesişir ve aralarındaki Θ_2 açısının tanjantı için,

$$m_2 = \frac{1}{r} \quad \text{ve} \quad m_1 = r \quad \text{olduğundan,}$$

$$\operatorname{tg}\Theta_2 = \frac{\frac{1}{r} - r}{1 + \frac{1}{r}} = \frac{1 - r^2}{2r} \quad (4.170)$$

elde ederiz. Bu durumda, $\operatorname{tg}\Theta_2$ nin değeri: $s_x = s_y = 1$ için $\operatorname{tg}\Theta_1$ e eşit olur.

Her iki formülden de görüldüğü gibi, eğer $r = 1$ ise, bu takdirde $\operatorname{tg}\Theta_1 = 0$, yani, $\Theta_1 = 0$ dır. Bu durumda, regresyon doğruları aynı yönde çakışır. Eğer $r = -1$ ise, bu takdirde regresyon doğruları aynı doğrultu ve ters yönde çakışır. Eğer $r = 0$ ise, bu takdirde regresyon doğruları birbirine dik ve eksnlere paraleldir. Buna göre,

1) $r^2 = 1$ ise \mathbf{X} ve \mathbf{Y} değişkenleri arasında mükemmel bir korelasyon vardır ve regresyon doğruları çakışır.

2) $r = 0$ ise, yani hiç korelasyon (lineer ilişki) yoksa, bu takdirde $\operatorname{tg}\Theta_1 = \operatorname{tg}\Theta_2 = \infty$ veya $\Theta_1 = \Theta_2 = \pi/2$ dir. Regresyon doğrusunun her biri diğerine dik ve her biri eksnlere paraleldir. Bu durumda, \mathbf{Y} nin ortalama değeri \mathbf{X} in tüm değerleri için aynıdır ve \mathbf{X} in ortalama değeri \mathbf{Y} nin tüm değerleri için aynıdır.

Bu durumda, $\hat{\mathbf{Y}} - \mathbf{j}\bar{\mathbf{Y}} = b_1(\mathbf{X} - \mathbf{j}\bar{\mathbf{X}})$ ile $\hat{\mathbf{X}} - \mathbf{j}\bar{\mathbf{X}} = b_2(\mathbf{Y} - \mathbf{j}\bar{\mathbf{Y}})$ vektörleri dik olduğundan kolayca

$$(\hat{\mathbf{Y}} - \mathbf{j}\bar{\mathbf{Y}})'(\hat{\mathbf{X}} - \mathbf{j}\bar{\mathbf{X}}) = b_1 b_2 (\mathbf{X} - \mathbf{j}\bar{\mathbf{X}})'(\mathbf{Y} - \mathbf{j}\bar{\mathbf{Y}}) = 0,$$

yani $\operatorname{kov}(\hat{x}, \hat{y}) = 0$ veya

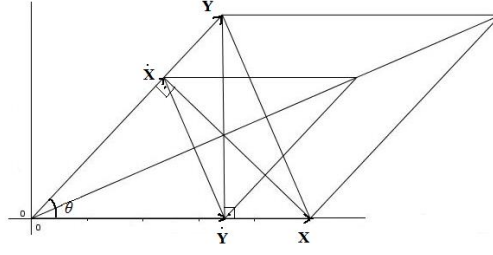
$$\frac{(\hat{\mathbf{Y}} - \mathbf{j}\bar{\mathbf{Y}})'(\hat{\mathbf{X}} - \mathbf{j}\bar{\mathbf{X}})}{\|\mathbf{Y} - \mathbf{j}\bar{\mathbf{Y}}\| \|\mathbf{X} - \mathbf{j}\bar{\mathbf{X}}\|} = r \frac{s_Y}{s_X} \frac{s_X}{s_Y} r \frac{(\mathbf{X} - \mathbf{j}\bar{\mathbf{X}})'(\mathbf{Y} - \mathbf{j}\bar{\mathbf{Y}})}{\|\mathbf{X} - \mathbf{j}\bar{\mathbf{X}}\| \|\mathbf{Y} - \mathbf{j}\bar{\mathbf{Y}}\|} = 0$$

$$= r^2 r = 0$$

yani \hat{y} ile \hat{x} in dik olacağını ifade ederiz.

4.2.4.2. \mathbf{X} ve \mathbf{Y} Vektörlerinin Geometrisi

Şimdi, \mathbf{Y} nin \mathbf{X} üzerindeki ve \mathbf{X} in \mathbf{Y} üzerindeki (dik) izdüşüm vektörlerini göz önüne alınız. Bu izdüşüm vektörleri sırasıyla $\hat{\mathbf{X}}$ ve $\hat{\mathbf{Y}}$ olarak gösterilsin (bkz. Şekil 4.12).



Şekil 4.12. \mathbf{X} ve \mathbf{Y} üzerine kurulan ve $\dot{\mathbf{X}}$ ve $\dot{\mathbf{Y}}$ üzerine kurulan paralelkenarlar.

Buna göre,

$$\dot{\mathbf{X}} = \mathbf{Y}(\mathbf{Y}'\mathbf{Y})^{-1}\mathbf{Y}'\mathbf{X} = \frac{\mathbf{Y}}{\|\mathbf{Y}\|^2} \|\mathbf{Y}\| \|\mathbf{X}\| \cos \theta = \frac{\mathbf{Y}}{\|\mathbf{Y}\|} \|\mathbf{X}\| \cos \theta \quad (4.171)$$

ve

$$\dot{\mathbf{Y}} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} = \frac{\mathbf{X}}{\|\mathbf{X}\|^2} \|\mathbf{X}\| \|\mathbf{Y}\| \cos \theta = \frac{\mathbf{X}}{\|\mathbf{X}\|} \|\mathbf{Y}\| \cos \theta \quad (4.172)$$

elde ederiz.

Şekil 4.12 den görüldüğü gibi,

$$\cos \theta = \frac{\|\dot{\mathbf{X}}\|}{\|\mathbf{X}\|} \rightarrow \|\dot{\mathbf{X}}\| = \|\mathbf{X}\| \cos \theta \quad (4.173)$$

ve

$$\cos \theta = \frac{\|\dot{\mathbf{Y}}\|}{\|\mathbf{Y}\|} \rightarrow \|\dot{\mathbf{Y}}\| = \|\mathbf{Y}\| \cos \theta \quad (4.174)$$

dir. Böylece (4.173) ve (4.174) ün doğrulanması sırasıyla, (4.171) ve (4.172) den de yapılabilir.

S ve W sırasıyla \mathbf{X} ve \mathbf{Y} vektörleri ve $\dot{\mathbf{X}}$ ve $\dot{\mathbf{Y}}$ vektörleri üzerine kurulan paralelkenarın alanı olsun. Bu durumda,

$$S = \|\mathbf{X} \wedge \mathbf{Y}\| = \|\mathbf{X}\| \|\mathbf{Y}\| \sin \theta \quad (4.175)$$

ve

$$W = \|\dot{\mathbf{X}} \wedge \dot{\mathbf{Y}}\| = \|\dot{\mathbf{X}}\| \|\dot{\mathbf{Y}}\| \sin \theta = \underbrace{\|\mathbf{X}\| \cos \theta}_{\|\dot{\mathbf{X}}\|} \underbrace{\|\mathbf{Y}\| \cos \theta}_{\|\dot{\mathbf{Y}}\|} \sin \theta = \|\mathbf{X}\| \|\mathbf{Y}\| \cos^2 \theta \sin \theta \quad (4.176)$$

dir. Eğer W nun S ye oranı λ ise, bu takdirde

$$\lambda = \frac{W}{S} = \frac{\|\mathbf{X}\|\|\mathbf{Y}\|\cos^2\theta\sin\theta}{\|\mathbf{X}\|\|\mathbf{Y}\|\sin\theta} = \cos^2\theta \quad (4.177)$$

dir. $\sqrt{\lambda}$ değeri; θ açısının kosinüsü olduğundan, $-1 \leq \sqrt{\lambda} \leq 1$ eşitsizliğini elde ederiz.

(i) Eğer $\sqrt{\lambda} = 0$ ise, yani, $\cos\theta = 0$ ise, bu takdirde $\theta = \pi/2$ veya $3\pi/2$ olur. Bu durumda $W = 0$ dir, yani, $\dot{\mathbf{X}}$ ve $\dot{\mathbf{Y}}$ sıfır vektörlerdir. S ; kenarları $\|\mathbf{X}\|$ ve $\|\mathbf{Y}\|$ olan bir dikdörtgenin alanı olur. Bu nedenle, \mathbf{X} ve \mathbf{Y} vektörleri lineer bağımsızdır.

(ii) Eğer $\sqrt{\lambda} = +1$ ise, yani, $\cos\theta = +1$ ise, bu takdirde $\theta = 0$ olur. \mathbf{X} ve \mathbf{Y} aynı doğrultudadır.

(iii) Eğer $\sqrt{\lambda} = -1$ ise, yani, $\cos\theta = -1$ ise, bu takdirde $\theta = \pi$ dir. Şimdi bu durumda, \mathbf{X} ve \mathbf{Y} ters doğrultudadır.

(ii) ve (iii) durumlarında \mathbf{X} ve \mathbf{Y} arasında mükemmel bir lineer ilişki vardır.

$$r s_{\mathbf{X}} s_{\mathbf{Y}} = \|\mathbf{X}\|\|\mathbf{Y}\|\cos\theta - n\bar{\mathbf{X}}\bar{\mathbf{Y}} \quad (4.178)$$

olduğunu biliyoruz.

Eğer $\theta = 0$ ise, bu takdirde $S = 0$; \mathbf{X} ve \mathbf{Y} nin aynı doğrultuda olduklarını ve aralarında mükemmel bir pozitif lineer ilişkinin var olduğunu belirtir. Bu durumda $r = 1$ ve $s_{\mathbf{X}} s_{\mathbf{Y}} = \|\mathbf{X}\|\|\mathbf{Y}\| - n\bar{\mathbf{X}}\bar{\mathbf{Y}}$ olur.

Eğer $\theta = \pi$ ise, bu takdirde $S = 0$; \mathbf{X} ve \mathbf{Y} nin aynı doğrultuda ve ters yönde olduklarını ve bu nedenle aralarında mükemmel bir negatif lineer ilişkinin var olduğunu belirtir. Bu durumda $r = -1$ ve $s_{\mathbf{X}} s_{\mathbf{Y}} = \|\mathbf{X}\|\|\mathbf{Y}\| + n\bar{\mathbf{X}}\bar{\mathbf{Y}}$ olur.

Eğer $\theta = \pi/2$ veya $3\pi/2$ ise, bu takdirde $S = \|\mathbf{X}\|\|\mathbf{Y}\|$ dir, yani paralelkenar bir dikdörtgendir. Bu nedenle \mathbf{X} ve \mathbf{Y} birbirine diktir. Bu durumda $r = 0$ ve $\bar{\mathbf{X}}\bar{\mathbf{Y}} = 0$ olur.

4.2.4.3. x ve y Standartlaştırılmış Vektörlerinin Geometrisi

$$x = \frac{\mathbf{X} - j\bar{\mathbf{X}}}{\|\mathbf{X} - j\bar{\mathbf{X}}\|} \quad \text{ve} \quad y = \frac{\mathbf{Y} - j\bar{\mathbf{Y}}}{\|\mathbf{Y} - j\bar{\mathbf{Y}}\|}$$

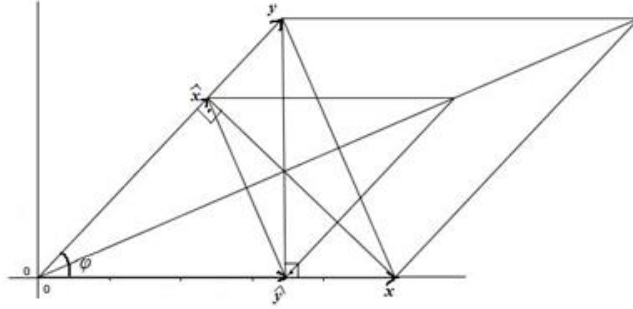
standartlaştırılmış vektörler ile çalıştığımızda, $r = \frac{(\mathbf{X} - j\bar{\mathbf{X}})'(\mathbf{Y} - j\bar{\mathbf{Y}})}{\|\mathbf{X} - j\bar{\mathbf{X}}\|\|\mathbf{Y} - j\bar{\mathbf{Y}}\|} = x'y$,

$$\|\mathbf{x}\| = \|\mathbf{y}\| = 1, \quad \hat{\mathbf{y}} = r\mathbf{x}, \quad \hat{\mathbf{x}} = r\mathbf{y}, \quad \cos\psi = -r = -\cos\varphi,$$

$$\|\hat{\mathbf{x}}\| = \|\hat{\mathbf{y}}\| = r, \quad \|\mathbf{x} - \hat{\mathbf{x}}\| = \sqrt{1-r^2}, \quad \|\mathbf{y} - \hat{\mathbf{y}}\| = \sqrt{1-r^2} \quad \text{ve} \quad \sin\eta = \sin\xi = \cos\varphi = r$$

olduğunu biliyoruz (bkz. Şekil 4.10).

Şimdi, \mathbf{x} , \mathbf{y} ve $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$ vektörleri üzerine kurulan eşkenar dörtgenleri göz önüne alınız (bkz. Şekil 4.13). \mathbf{x} ve \mathbf{y} vektörleri üzerine kurulan eşkenar dörtgenin alanı S_1 ve $\hat{\mathbf{x}}$ ve $\hat{\mathbf{y}}$ vektörleri üzerine kurulan eşkenar dörtgenin alanı W_1 olsun.



Şekil 4.13. \mathbf{x} , \mathbf{y} ve $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$ vektörleri üzerine kurulan eşkenar dörtgenler.

Bir eşkenar dörtgenin alanı köşegen uzunluklarının çarpımının yarısına veya üzerine kuruldukları vektörlerin vektörel çarpımının normuna eşittir. Bu nedenle,

$$S_1 = \|\mathbf{x} \wedge \mathbf{y}\| = \sin\varphi = \sqrt{1 - \cos^2\varphi} = \sqrt{1 - r^2} \quad (4.179)$$

veya

$$\frac{\|\mathbf{x} - \mathbf{y}\| \|\mathbf{x} + \mathbf{y}\|}{2} = \frac{\sqrt{2}(\sqrt{1-r})\sqrt{2}(\sqrt{1+r})}{2} = \sqrt{1-r^2} \quad (4.180)$$

yazarız.

W_1 e gelince,

$$W_1 = \|\hat{\mathbf{x}} \wedge \hat{\mathbf{y}}\| = \|\hat{\mathbf{x}}\| \|\hat{\mathbf{y}}\| \sin\varphi = rr \sin\varphi = r^2 \sin\varphi = r^2 \sqrt{1 - \cos^2\varphi} = r^2 \sqrt{1 - r^2} \quad (4.181)$$

olduğu görülür.

Eğer W_1 in S_1 e oranı λ_1 ise, bu takdirde

$$\lambda_1 = \frac{W_1}{S_1} = r^2 \quad (\text{determinasyon katsayısı}) \quad (4.182)$$

elde ederiz.

- Eğer $\lambda_1 = 1 = r^2$ ise, yani $W_1 = S_1$ ise, bu takdirde x ve y arasında mükemmel bir lineer ilişki vardır.
- Eğer $\lambda_1 = 0$ (veya $r^2 = 0$) ise, bu takdirde $\cos \varphi = 0$, yani x ve y vektörleri diktir ve bu nedenle eşkenar dörtgen bir birim karedir, yani, \hat{x} ve \hat{y} vektörleri sıfır vektörlerdir. \mathbf{X} ve \mathbf{Y} vektörleri lineer bağımsızdır.

4.2.4.4. Basit Lineer Regresyon Doğrularına Dayanan Bir Genişletme

\mathbf{Y}_1 ile \mathbf{X}_1 vektörleri arasındaki basit lineer regresyon bağıntısını göz önüne alınız. Bu basit lineer regresyon bağıntısını aşağıdaki biçimde verelim.

$$\mathbf{Y}_1 = \mathbf{j}\alpha_1 + \beta_1 \mathbf{X}_1 + \boldsymbol{\varepsilon}_1, \quad \boldsymbol{\varepsilon}_1 : N(\mathbf{0}, \sigma^2 \mathbf{I}_n)$$

En küçük kareler yöntemini kullanarak, \mathbf{Y}_1 in en küçük kareler tahmini aşağıdaki gibi elde edilir :

$$\hat{\mathbf{Y}}_1 = \mathbf{X}\hat{\boldsymbol{\beta}}_1 = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{Y}_1 \quad (\text{burada } \mathbf{X} = [\mathbf{j}; \mathbf{X}_1] \text{ dir.}) \quad (4.183)$$

φ_1 ; \mathbf{Y}_1 ve $\hat{\mathbf{Y}}_1$ arasındaki açı olduğunda,

$$\cos \varphi_1 = \frac{\mathbf{Y}_1' \hat{\mathbf{Y}}_1}{\|\mathbf{Y}_1\| \|\hat{\mathbf{Y}}_1\|} = \frac{\mathbf{Y}_1' \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{Y}_1}{\|\mathbf{Y}_1\| \|\hat{\mathbf{Y}}_1\|} = \frac{\|\hat{\mathbf{Y}}_1\|^2}{\|\mathbf{Y}_1\| \|\hat{\mathbf{Y}}_1\|} = \frac{\|\hat{\mathbf{Y}}_1\|}{\|\mathbf{Y}_1\|} \quad (4.184)$$

elde ederiz. Bu nedenle,

$$\sin \varphi_1 = \sqrt{1 - \frac{\|\hat{\mathbf{Y}}_1\|^2}{\|\mathbf{Y}_1\|^2}} \quad (4.185)$$

olduğu görülür. S_2 ; \mathbf{Y}_1 ve $\hat{\mathbf{Y}}_1$ vektörleri üzerine kurulan paralelkenarın alanı olsun.

$$S_2 = \|\mathbf{Y}_1 \wedge \hat{\mathbf{Y}}_1\| \sin \varphi_1 = \|\mathbf{Y}_1\| \|\hat{\mathbf{Y}}_1\| \sqrt{1 - \frac{\|\hat{\mathbf{Y}}_1\|^2}{\|\mathbf{Y}_1\|^2}} \quad (4.186)$$

yazabiliriz.

Eğer $\|\hat{\mathbf{Y}}_1\| = \|\mathbf{Y}_1\|$ ise, bu takdirde $\sin \varphi_1 = 0$ ve $S_2 = 0$ dir. Bu nedenle, \mathbf{Y}_1 ve $\hat{\mathbf{Y}}_1$ lineer bağımlıdır (ya da \mathbf{Y}_1 ve $\hat{\mathbf{Y}}_1$ aynı düzlemde olup lineer bağımlıdır).

Eğer $\varphi_1 = \pi / 2$ ise, bu takdirde $\sin \varphi_1 = 1$ dir. Bu paralelkenar kenarları $\|\mathbf{Y}_1\|$ ve $\|\hat{\mathbf{Y}}_1\|$ olan bir dikdörtgendir. Bu durumda \mathbf{Y}_1 ve $\hat{\mathbf{Y}}_1$ lineer bağımsızdır.

Aynı şekilde \mathbf{X}_1 in \mathbf{Y}_1 üzerindeki basit lineer regresyonu düşünüldüğünde, yukarıda verilen basit lineer regresyon bağıntısı aşağıdaki biçimi alır :

$$\mathbf{X}_1 = \mathbf{j}\alpha_2 + \beta_2 \mathbf{Y}_1 + \boldsymbol{\varepsilon}_2, \quad \boldsymbol{\varepsilon}_2 : N(\mathbf{0}, \sigma^2 \mathbf{I}_n).$$

Bu nedenle, $\mathbf{Y} = [\mathbf{j} : \mathbf{Y}_1]$ olmak üzere,

$$\hat{\mathbf{X}}_1 = \mathbf{Y}(\mathbf{Y}'\mathbf{Y})^{-1}\mathbf{Y}'\mathbf{X}_1 \quad (4.187)$$

elde ederiz.

$\hat{\mathbf{X}}_1$ ve \mathbf{X}_1 vektörleri arasındaki açı φ_2 olduğunda, \mathbf{X}_1 ve \mathbf{Y}_1 in rolleri değişeceğinden

$$\sin \varphi_2 = \sqrt{1 - \frac{\|\hat{\mathbf{X}}_1\|^2}{\|\mathbf{X}_1\|^2}} \quad (4.188)$$

yazabiliriz.

\mathbf{Y}_1 ve $\hat{\mathbf{Y}}_1$ hakkında verilen yorumlar \mathbf{X}_1 ve $\hat{\mathbf{X}}_1$ için de verilebilir.

$\hat{\mathbf{X}}_1$ ve $\hat{\mathbf{Y}}_1$ aynı grafikte gösterildiğinde, eğer onların arasındaki açı w ise,

$$tgw = \frac{1-r^2}{r} \left(\frac{s_{\mathbf{X}}s_{\mathbf{Y}}}{s_{\mathbf{X}}^2 + s_{\mathbf{Y}}^2} \right) \quad (4.189)$$

bu nedenle,

$$\sin w = \frac{tgw}{m\sqrt{1+tg^2w}} \quad (4.190)$$

elde ederiz.

Böylece, $\hat{\mathbf{X}}_1$ ve $\hat{\mathbf{Y}}_1$ üzerine kurulan paralelkenarın alanı

$$\|\hat{\mathbf{X}}_1\| \|\hat{\mathbf{Y}}_1\| \|\sin w\| = \|\hat{\mathbf{X}}_1\| \|\hat{\mathbf{Y}}_1\| \frac{tgw}{m\sqrt{1+tg^2w}} \quad (4.191)$$

olur.

Eğer $r^2 = 1$ ise, yani $r = \pm 1$ bu takdirde $\hat{\mathbf{X}}_1$ ve $\hat{\mathbf{Y}}_1$ arasında mükemmel bir lineer ilişki vardır. Bu durumda, $\hat{\mathbf{X}}_1$ ve $\hat{\mathbf{Y}}_1$ aynı doğrultu aynı yönde ya da aynı doğrultu ters yöndedir.

Eğer $r = 0$ ise, bu takdirde $tgw = \infty$ dur. $w = \pi/2$ veya $3\pi/2$ elde ederiz ki bu durumda $\sin w = \pm 1$ dir.

Bu nedenle, $\hat{\mathbf{X}}_1$ ve $\hat{\mathbf{Y}}_1$ üzerine kurulan paralelkenar $\|\hat{\mathbf{X}}_1\|$ ve $\|\hat{\mathbf{Y}}_1\|$ kenarlı bir dikdörtgen olur.

4.2.4.5. Korelasyon Katsayısına Farklı Bir Cebirsel Bakış

Şimdi, \mathbf{X} ve \mathbf{Y} reel gözlem vektörleri arasındaki bir lineer ilişkiyi tanımlamak için, niçin \mathbf{x} ve \mathbf{y} standartlaştırılmış vektörlerini kullandığımızı düşünelim. Aşağıdaki vektörleri ve matrisi göz önüne alalım:

$$\mathbf{X} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \mathbf{Y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}, \mathbf{j}_{n \times 1} = (1, 1, \dots, 1)', \mathbf{J} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \dots & 1 \end{bmatrix}_{n \times n} = \mathbf{j}\mathbf{j}' \quad (\mathbf{j}\mathbf{j}' = n) \quad (4.192)$$

ve

$$\tilde{\mathbf{x}} = \begin{bmatrix} x_1 - \bar{\mathbf{X}} \\ x_2 - \bar{\mathbf{X}} \\ \vdots \\ x_n - \bar{\mathbf{X}} \end{bmatrix} = \left(\mathbf{I} - \frac{1}{n} \mathbf{J} \right) \mathbf{X} = \mathbf{X} - \mathbf{J}\bar{\mathbf{X}}, \quad \tilde{\mathbf{y}} = \begin{bmatrix} y_1 - \bar{\mathbf{Y}} \\ y_2 - \bar{\mathbf{Y}} \\ \vdots \\ y_n - \bar{\mathbf{Y}} \end{bmatrix} = \left(\mathbf{I} - \frac{1}{n} \mathbf{J} \right) \mathbf{Y} = \mathbf{Y} - \mathbf{J}\bar{\mathbf{Y}}. \quad (4.193)$$

Bilindiği gibi, $\bar{\mathbf{X}} = \sum_{i=1}^n x_i$ ve $\bar{\mathbf{Y}} = \sum_{i=1}^n y_i$ olmak üzere, \mathbf{X} ve \mathbf{Y} gözlem vektörleri arasındaki Pearson örneklem korelasyon katsayısı

$$r_{\mathbf{XY}} = \frac{\sum_{i=1}^n (x_i - \bar{\mathbf{X}})(y_i - \bar{\mathbf{Y}})}{\sqrt{\left(\sum_{i=1}^n (x_i - \bar{\mathbf{X}})^2 \right) \left(\sum_{i=1}^n (y_i - \bar{\mathbf{Y}})^2 \right)}} \quad (4.194)$$

dür.

Açıklama 4.1. (4.192) ve (4.193) bağıntılarının ışığı altında, $\lambda_1, \lambda_2 \in \mathbb{R}$ olmak üzere, aşağıdaki bağıntıyı elde ederiz:

$$\lambda_1 \mathbf{X} + \lambda_2 \mathbf{Y} = 0 \rightarrow \lambda_1 \bar{\mathbf{X}} + \lambda_2 \bar{\mathbf{Y}} = 0 \rightarrow \lambda_1 \underbrace{(\mathbf{X} - \mathbf{j}\bar{\mathbf{X}})}_{\tilde{\mathbf{x}}} + \lambda_2 \underbrace{(\mathbf{Y} - \mathbf{j}\bar{\mathbf{Y}})}_{\tilde{\mathbf{y}}} = 0. \quad (4.195)$$

Eğer, $\lambda_1 \tilde{\mathbf{x}} + \lambda_2 \tilde{\mathbf{y}} = 0$ denklemini sağlayan en az bir $\lambda_i \neq 0$, ($i=1,2$) varsa, bu takdirde $\tilde{\mathbf{x}}$ ve $\tilde{\mathbf{y}}$ vektörleri lineer bağımlıdır. Bu nedenle \mathbf{X} ve \mathbf{Y} gözlem vektörleri de lineer bağımlıdır (aynı λ_1, λ_2 kullanıldığından). Eğer bu denklem yalnız $\lambda_1 = \lambda_2 = 0$ için gerçekleşirse, $\tilde{\mathbf{x}}$ ve $\tilde{\mathbf{y}}$ vektörleri lineer bağımsızdır (bu nedenle, \mathbf{X} ve \mathbf{Y} gözlem vektörleri de lineer bağımsızdır).

(4.195) denklemi matris formunda aşağıdaki gibi yazılabilir:

$$\lambda_1 \tilde{\mathbf{x}} + \lambda_2 \tilde{\mathbf{y}} = 0 \Rightarrow \begin{bmatrix} x_1 - \bar{\mathbf{X}} & y_1 - \bar{\mathbf{Y}} \\ x_2 - \bar{\mathbf{X}} & y_2 - \bar{\mathbf{Y}} \\ \vdots & \vdots \\ x_n - \bar{\mathbf{X}} & y_n - \bar{\mathbf{Y}} \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ \vdots \\ 0 \end{bmatrix}. \quad (4.196)$$

(4.196) lineer denklemler sistemi, bir homojen lineer denklem sistemidir. Bu nedenle (4.196), $\mathbf{A}\boldsymbol{\lambda} = \mathbf{0}$ olarak yazılabilir. Burada \mathbf{A} matrisi ve $\boldsymbol{\lambda}$ vektörü sırasıyla aşağıdaki gibidir:

$$\mathbf{A}_{n \times 2} = [\tilde{\mathbf{x}} \quad \tilde{\mathbf{y}}] = \begin{bmatrix} x_1 - \bar{\mathbf{X}} & y_1 - \bar{\mathbf{Y}} \\ x_2 - \bar{\mathbf{X}} & y_2 - \bar{\mathbf{Y}} \\ \vdots & \vdots \\ x_n - \bar{\mathbf{X}} & y_n - \bar{\mathbf{Y}} \end{bmatrix}_{n \times 2} \quad \text{ve} \quad \boldsymbol{\lambda} = \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix}. \quad (4.197)$$

Bu lineer denklemler sistemi daima tutarlıdır, yani, sistem bir çözüme sahiptir. Eğer \mathbf{A} katsayılar matrisi tam sütun ranklı ise, yani $r(\mathbf{A}) = r(\mathbf{A}'\mathbf{A}) = 2$ ise, bu takdirde $r(\mathbf{A})$; \mathbf{A} matrisinin rankını göstermek üzere, $\tilde{\mathbf{x}}$ ve $\tilde{\mathbf{y}}$ vektörleri lineer bağımsızdır. Bu durumda $\lambda_1 = \lambda_2 = 0$ çözümü hariç başka bir çözüm yoktur. Eğer \mathbf{A} katsayılar matrisi tam sütun ranklı değilse, bu takdirde (4.196) sistemi için en az bir $\lambda_i \neq 0$, ($i=1,2$) çözümü vardır. Bu durumda $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{y}}$ vektörleri lineer bağımlıdır ve \mathbf{X} and \mathbf{Y} vektörleri de lineer bağımlıdır.

$$\mathbf{A}'\mathbf{A} = \begin{bmatrix} \sum_{i=1}^n (x_i - \bar{\mathbf{X}})^2 & \sum_{i=1}^n (x_i - \bar{\mathbf{X}})(y_i - \bar{\mathbf{Y}}) \\ \sum_{i=1}^n (x_i - \bar{\mathbf{X}})(y_i - \bar{\mathbf{Y}}) & \sum_{i=1}^n (y_i - \bar{\mathbf{Y}})^2 \end{bmatrix} \quad (4.198)$$

olduğundan, $\det(\mathbf{A}'\mathbf{A})$; $\mathbf{A}'\mathbf{A}$ matrisinin determinantını göstermek üzere,

$$\det(\mathbf{A}'\mathbf{A}) = \sum_{i=1}^n (x_i - \bar{\mathbf{X}})^2 \sum_{i=1}^n (y_i - \bar{\mathbf{Y}})^2 - \left[\sum_{i=1}^n (x_i - \bar{\mathbf{X}})(y_i - \bar{\mathbf{Y}}) \right]^2 \quad (4.199)$$

elde ederiz.

$$\det(\mathbf{A}'\mathbf{A}) = \sum_{i=1}^n (x_i - \bar{\mathbf{X}})^2 \sum_{i=1}^n (y_i - \bar{\mathbf{Y}})^2 - \left[\sum_{i=1}^n (x_i - \bar{\mathbf{X}})(y_i - \bar{\mathbf{Y}}) \right]^2 = (n-1)^2 [s_{\mathbf{X}}^2 s_{\mathbf{Y}}^2 - s_{\mathbf{XY}}^2] = k \in \mathbb{R} \quad (4.200)$$

alalım. k bir pozitif yarı-tanımlı matrisin determinantı olduğundan, $k \geq 0$ dır.

(4.200) den,

$$\sum_{i=1}^n (x_i - \bar{\mathbf{X}})^2 \sum_{i=1}^n (y_i - \bar{\mathbf{Y}})^2 - k = \left[\sum_{i=1}^n (x_i - \bar{\mathbf{X}})(y_i - \bar{\mathbf{Y}}) \right]^2 \quad (4.201)$$

veya

$$1 - \frac{k}{(n-1)^2 s_{\mathbf{X}}^2 s_{\mathbf{Y}}^2} = \left[\frac{\sum_{i=1}^n (x_i - \bar{\mathbf{X}})(y_i - \bar{\mathbf{Y}})}{s_{\mathbf{X}} s_{\mathbf{Y}}} \right]^2 = r_{\mathbf{XY}}^2 \rightarrow r_{\mathbf{XY}} = \pm \sqrt{1 - \frac{k}{(n-1)^2 s_{\mathbf{X}}^2 s_{\mathbf{Y}}^2}}, \quad (r_{\mathbf{XY}} = \cos(\mathbf{x}, \mathbf{y})) \quad (4.202)$$

elde ederiz.

$$0 \leq r_{\mathbf{XY}}^2 \leq 1 \text{ olduğundan, } 0 \leq 1 - \frac{k}{(n-1)^2 s_{\mathbf{X}}^2 s_{\mathbf{Y}}^2} \leq 1, \text{ yani, } 0 \leq \frac{k}{(n-1)^2 s_{\mathbf{X}}^2 s_{\mathbf{Y}}^2} \leq 1 \text{ elde}$$

ederiz.

Eğer $k=0$ ise, yani $r_{\mathbf{XY}}=1$ ise, \mathbf{X} ve \mathbf{Y} gözlem vektörleri arasında mükemmel bir lineer ilişki vardır.

$$\text{Eğer } 1 - \frac{k}{(n-1)^2 s_{\mathbf{X}}^2 s_{\mathbf{Y}}^2} = 0 \text{ ise, yani } \frac{k}{(n-1)^2 s_{\mathbf{X}}^2 s_{\mathbf{Y}}^2} = 1 \text{ veya } k = (n-1)^2 s_{\mathbf{X}}^2 s_{\mathbf{Y}}^2 \neq 0 \text{ ise,}$$

bu takdirde \mathbf{X} ve \mathbf{Y} gözlem vektörleri arasında lineer ilişki yoktur.

k sabitini farklı bir yöntemle yeniden inceleyelim. Bunun için de Gram-Schmidt ve determinant ile ilgili aşağıdaki iki teoremi ve sonucu verelim.

Teorem 4.2. $\vec{v}_1, \dots, \vec{v}_n$ sütunlarına sahip bir A kare matrisi için, $A = QR$ bir QR çarpansallaması olsun. Bunun, Q nun bir ortogonal matris ve R nin

$\left\| \begin{matrix} \vec{v}_1 \\ \vec{v}_2 \\ \dots \\ \vec{v}_n \end{matrix} \right\|$ köşegen elemanlarına sahip bir üst üçgen matris olduğunu

söylediğini hatırlayalım. Bu takdirde $|\det(A)| = \left\| \begin{matrix} \vec{v}_1 \\ \vec{v}_2 \\ \dots \\ \vec{v}_n \end{matrix} \right\| = \det(R)$ dir.

İspat. $\det(A) = \det(Q)\det(R)$ olduğunu biliyoruz. $\det(Q) = \pm 1$ olduğundan, bundan sonra mutlak değer alma, $|\det(A)| = |\det(R)|$ olduğunu ortaya koyar. Şimdi R ;

$\left\| \begin{matrix} \vec{v}_1 \\ \vec{v}_2 \\ \dots \\ \vec{v}_n \end{matrix} \right\|$ köşegen elemanlarına sahip üst üçgen olduğundan,

$\det(R) = \left\| \begin{matrix} \vec{v}_1 \\ \vec{v}_2 \\ \dots \\ \vec{v}_n \end{matrix} \right\|$ ve bu köşegen elemanlar pozitif olduklarından

$|\det(R)| = \det(R)$ elde ederiz. Bu istenilen sonucu verir (Bu durumda, sadece tersinir matrisler için QR çarpansallamasından söz ettik, bu nedenle teknik olarak yukarıdaki teoremden ispat ettiğimiz hepsi bu kadardır. Aslında, velevki sonuç tersinir olmayan matrisler için de doğrudur. Bu durumda $\det(A) = 0$ dir ve aslına bakılırsa, R köşegeni üzerinde bir sıfır elemana sahip olacak, böylece $\det(R)$ de sıfıra eşit olacak.). $\det(A)$ nın mutlak değeri olarak belirlediğiniz sayı tanıdık gelmelidir, bu sayı (nicelik) $\vec{v}_1, \dots, \vec{v}_n$ tarafından tanımlanan paralel yüzölçümünün n - hacmidir. Bu aşağıdakini ispatlar.

Sonuç 4.4. A ; $\vec{v}_1, \dots, \vec{v}_n$ sütunlarına sahip bir $n \times n$ matris olsun. Bu takdirde $|\det(A)| = V(\vec{v}_1, \dots, \vec{v}_n)$ dir.

Bunun da ötesinde, paralel yüzölçümü, A dönüşümü altında (yani, matrisi A olan lineer dönüşüm altında) birim n -kübün görüntüsüdür. Bu kısmı determinantları kullanarak ($m < n$ olduğunda bile) bir paralel yüzölçümünün hacmini hesaplamak için bir yöntemi vererek sonlandıracağız.

Teorem 4.3. $\vec{v}_1, \dots, \vec{v}_m \in \mathbb{R}^n$ bir A matrisinin sütunları olsun. Bu takdirde

$$V(\vec{v}_1, \dots, \vec{v}_m) = \sqrt{\det(A'A)}$$

dır. $m=n$ olduğunda bu tamı tamına $|\det(A)|$ dır.

Sonucu ispatlamadan önce bu teoremin bize neyi satın aldığını bilmek yararlıdır.

Bundan önce m -hacmi hesaplamak için, $\left\| \begin{matrix} \vec{v}_1 \\ \vec{v}_2 \\ \vdots \\ \vec{v}_m \end{matrix} \right\|$ niceliklerini ve bunların

ancak Gram-Schmidt algoritmasını kullanmak için elde mevcut olmasını bilmemiz gerekmektedir. Bununla beraber, bu teorem bir ortonormal tabanı üretmeksizin

$V(\vec{v}_1, \dots, \vec{v}_m)$ yi hesaplamamıza izin verir-bu; Gram-Schmidt algoritmasını kullanmaya ihtiyacımızın olmadığını söyler ki bu muhteşemdir!

İspat. $\left\{ \vec{v}_1, \dots, \vec{v}_m \right\}$ in lineer bağımlı olduğu durumla başlayınız. Bu durumda $\vec{v}_1, \dots, \vec{v}_m$

tarafından üretilen m -paralel yüzü azami $m-1$ boyutludur ve bu nedenle onun m -

hacmi 0 dır. Öte yandan, bu durumda, $\ker(A)$ (A'nın çekirdeği - sıfırlığı) $\neq \left\{ \vec{0} \right\}$

olduğunu da biliyoruz; o halde buradan $\ker(A'A) \neq \left\{ \vec{0} \right\}$ elde ederiz. Fakat bu takdirde

$A'A$ tersinir olmayan bir kare matristir ve bundan dolayı $\det(A'A) = 0$ dır. Bu nedenle bu durumda

$$V(\vec{v}_1, \dots, \vec{v}_m) = 0 = \det(A'A)$$

elde ederiz.

Şimdi $\left\{ \vec{v}_1, \dots, \vec{v}_m \right\}$ nin lineer bağımsız olduğunu farz edelim. A nin QR

çarpansallamasına (ayrışımına) ihtiyaç duyacağız, bu nedenle (herzamanki gibi) Q

nun ortogonal ve R nin $\left\| \vec{v}_1 \right\|, \left\| \vec{v}_2 \right\|, \dots, \left\| \vec{v}_m \right\|$ köşegen elemanlarına sahip üst üçgen

olduğunu yeniden hatırlayalım. Şimdi, Q ortonormal sütunlara sahip olduğundan * eşitliği ve herhangi bir A matrisi için $\det(A) = \det(A')$ olduğundan ** eşitliği anlaşılacak üzere,

$$\begin{aligned} \det(A'A) &= \det((QR)'QR) = \det(R'Q'QR) = \det(R'R) \\ &= \det(R') \det(R) = \det(R)^2 = \left(\left\| \vec{v}_1 \right\| \left\| \vec{v}_2 \right\| \dots \left\| \vec{v}_m \right\| \right)^2 \end{aligned}$$

elde ederiz.

O halde Teorem 4.2, Teorem 4.3. ve Sonuç 4.4 den $A(\bar{x}, \bar{y}) = \sqrt{\det(A'A)} = \sqrt{k}$ olduğu görülür. Yani \sqrt{k} , \bar{x} ve \bar{y} tarafından üretilen paralel yüzünün alanına eşittir. Bu durumda paralel yüzümüz bir paralelkenardır. Bu ise, $k \geq 0$ olması durumuna göre incelememizin bir başka geometrik yorumudur.

4.2.4.6. Korelasyon Katsayısına Başka Bir Cebirsel Bakış

α ve $\beta \neq 0$ sırasıyla regresyon doğrusunun integrasyon sabiti ve eğimi olmak üzere, \mathbf{X} ve \mathbf{Y} arasındaki lineer ilişki, \mathbf{Y} nin \mathbf{X} üzerindeki

$$y_i = \alpha + \beta x_i + \varepsilon_i, \quad \varepsilon_i : N(0, \sigma^2), \quad i = 1, 2, \dots, n \quad (4.203)$$

basit lineer regresyonunu kullanarak da belirlenebilir.

$\tilde{x} = \mathbf{X} - \mathbf{j}\bar{\mathbf{X}}$ üzerinde $\tilde{y} = \mathbf{Y} - \mathbf{j}\bar{\mathbf{Y}}$ nin basit lineer regresyon doğrusunun β eğimi, (4.203) deki β eğimi ile aynıdır.

(4.203) den, $\tilde{y} = \mathbf{Y} - \mathbf{j}\bar{\mathbf{Y}} = \beta(\mathbf{X} - \mathbf{j}\bar{\mathbf{X}}) + \varepsilon = \beta\tilde{x} + \varepsilon$ elde ederiz. Görüldüğü gibi β ; (4.203) deki β dir.

$x^* = \frac{\tilde{x}}{\|\tilde{x}\|}$ ve $y^* = \frac{\tilde{y}}{\|\tilde{y}\|}$ alalım. Burada, x^* ve y^* vektörleri birim vektörlerdir.

Kısım 4.2.4.5 deki aynı düşüncüyü kullanarak, $\mu_1, \mu_2 \in \mathbb{R}$ olacak şekilde $\lambda_1 = \frac{\mu_1}{\|\tilde{x}\|}$

ve $\lambda_2 = \frac{\mu_2}{\|\tilde{y}\|}$ alalım. Bu takdirde, aşağıdaki lineer homojen denklemler sistemini elde ederiz:

$$\begin{bmatrix} x_1 - \bar{\mathbf{X}} \\ x_2 - \bar{\mathbf{X}} \\ \vdots \\ x_n - \bar{\mathbf{X}} \end{bmatrix}, \begin{bmatrix} y_1 - \bar{\mathbf{Y}} \\ y_2 - \bar{\mathbf{Y}} \\ \vdots \\ y_n - \bar{\mathbf{Y}} \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} = \begin{bmatrix} \left(\frac{1}{\|\tilde{x}\|}\right) x_1 - \bar{\mathbf{X}} \\ \left(\frac{1}{\|\tilde{x}\|}\right) x_2 - \bar{\mathbf{X}} \\ \vdots \\ \left(\frac{1}{\|\tilde{x}\|}\right) x_n - \bar{\mathbf{X}} \end{bmatrix}, \begin{bmatrix} \left(\frac{1}{\|\tilde{y}\|}\right) y_1 - \bar{\mathbf{Y}} \\ \left(\frac{1}{\|\tilde{y}\|}\right) y_2 - \bar{\mathbf{Y}} \\ \vdots \\ \left(\frac{1}{\|\tilde{y}\|}\right) y_n - \bar{\mathbf{Y}} \end{bmatrix} \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}. \quad (4.204)$$

\mathbf{B} , (4.204) ün katsayılar matrisi olsun. Bu takdirde \mathbf{B} aşağıdaki gibi yazılabilir:

$$\mathbf{B} = \begin{bmatrix} \left(\frac{1}{\|\tilde{x}\|}\right) \begin{bmatrix} x_1 - \bar{\mathbf{X}} \\ x_2 - \bar{\mathbf{X}} \\ \vdots \\ x_n - \bar{\mathbf{X}} \end{bmatrix}, \left(\frac{1}{\|\tilde{y}\|}\right) \begin{bmatrix} y_1 - \bar{\mathbf{Y}} \\ y_2 - \bar{\mathbf{Y}} \\ \vdots \\ y_n - \bar{\mathbf{Y}} \end{bmatrix} \end{bmatrix}. \quad (4.205)$$

$$\begin{aligned}
\det(\mathbf{B}\mathbf{B}) &= \begin{vmatrix} 1 & \frac{1}{\|\tilde{\mathbf{x}}\|\|\tilde{\mathbf{y}}\|} \sum (x_i - \bar{\mathbf{X}})(y_i - \bar{\mathbf{Y}}) \\ \frac{1}{\|\tilde{\mathbf{x}}\|\|\tilde{\mathbf{y}}\|} \sum (x_i - \bar{\mathbf{X}})(y_i - \bar{\mathbf{Y}}) & 1 \end{vmatrix} \\
&= 1 - \frac{1}{\|\tilde{\mathbf{x}}\|^2 \|\tilde{\mathbf{y}}\|^2} \underbrace{\left[\sum (x_i - \bar{\mathbf{X}})(y_i - \bar{\mathbf{Y}}) \right]^2}_{=r_{\mathbf{X}\mathbf{Y}}^2} = 1 - r_{\mathbf{X}\mathbf{Y}}^2 = t \in \mathbb{R},
\end{aligned} \tag{4.206}$$

olduğundan, k için yapılan irdeleme t için de sürdürülür. Eğer $t=0$ ise, yani, $r_{\mathbf{X}\mathbf{Y}}^2=1$ ise bu takdirde \mathbf{X} ve \mathbf{Y} vektörleri arasında mükemmel bir lineer ilişki vardır. Eğer $t=1$ ise, yani, $\det(\mathbf{B}\mathbf{B}) \neq 0$ \mathbf{X} ve \mathbf{Y} vektörleri arasında lineer ilişki yoktur.

5. SONUÇ VE ÖNERİLER

Beş bölümden oluşan bu çalışmada lineer cebirle ve istatistikle ilgili olan bazı temel bilgiler tanıtıldıktan sonra lineer modeller ve lineer modellerin geometrisi ele alınmıştır. Lineer modellerde lineer kısıtlamalar altında parametre tahmini farklı bir yöntemle elde edilmiştir. Aynı bölümü takiben korelasyon katsayısının farklı geometrik yorumları ifade edilmiştir.

Geometrik bakış açısı ile soyut kavramlar somutlaştırılmıştır. Geniş bir literatür taraması yapılmış, bu konuda yazılan son makaleler ele alınarak mevcut gelişmelere yeni katkılar yapılmaya çalışılmıştır. Açıklayıcı örnekler, şekiller ve grafikler ile çalışmanın daha iyi anlaşılması sağlanmıştır.

Bu tezde ele alınan konular çoklu regresyon alanında çalışan araştırmacılara ve bu konuda doktora yapacak olanlara azda olsa katkı sağlayacaktır.

6. KAYNAKLAR

- Ahn, S. J., Rauh, W. & Warnecke, H. J. 2001. Least-squares orthogonal distances fitting of circle, sphere, ellipse, hyperbola, and parabola, *Pattern Recog.*, 34, 2283–2303.
- Ahn, S. J., Rauh, W. & Recknagel, M. 2001. Least squares orthogonal distances fitting of implicit curves and surfaces, In: LNCS 2191, 398–405.
- Ahn, S. J., Rauh, W. & Cho, H. S. 2002. Orthogonal distances fitting of implicit curves and surfaces, *IEEE trans. PAMI*, 24, 620–638.
- Ahn, S. J. 2004. Least Squares Orthogonal Distance Fitting of Curves and Surfaces in Space, In: LNCS 3151, Springer, Berlin.
- Aigner, M. & Jüttler, B. 2008. Gauss-Newton type techniques for robustly fitting implicitly defined curves and surfaces to unorganized data points, In: *Shape Modeling International*, pp. 121–130.
- Aigner, M. & Jüttler, B. 2005. Robust computation of foot points on implicitly defined curves, In: Editors Daehlen, M. et al., *Mathematical Methods for Curves and Surfaces*, Tromso 2004, Nashboro Press, pp. 1–10.
- Ajmani, V. B. 2009. *Applied Econometrics*.
- Alistair, B. 1989. Least Squares Best-Fit geometric elements, NPL report April.
- Anonim, 2017. Sphere Fitting, <https://www.google.com/url?sa=t&source=web&rct=j&url=http://caves>.
- Anonim, 2018. Geometry of Determinants; Cramer's Rule. <http://www.math.uiuc.edu/-acs/w10/math416>.
- Baksalary, J.K. and Pordzik, P.R. 1990. A Note on Comparing the Unrestricted and Least Squares Estimators.
- Bapat, R. P. 1993. *Linear Algebra and Linear Models*.
- Basilevsky, A. 1983. *Applied Matrix Algebra in the Statistical Sciences*.
- Batyrshin, I. & Kreinovich, V. 2015. *Thailand Statistician*, 13(1), 125-126.
- Bellman, R. 1960, *Introduction to Matrix Analysis*.
- Bullock, R. 2011. Least Squares Circle Fit.
- Buse, A. 1982. The Likelihood Ratio, Wald and Lagrange Multiplier Tests : An Expository Note. *The American Statistician*.
- Bookstein, F.L. 1979. *Fitting Conic Sections to Scattered Data*, *Computer Graphics and Image Processing* 9, pp. 56-71.
- Casella, G., Berger, R.L. 2002. *Matrix Differential Calculus with Applications in Statistics and Econometrics Statistical Inference*.
- Chernov, N. & Lesort, C. 2005. Least squares fitting of circles, *J. Math. Imag. Vision* 23, 239–251.
- Chernov, N. & Ma, H. 2011. Least Squares Fitting of Quadratic Curves and Surfaces.

- Chernov, N., Ososkov, G. & Silin, I. 2000. Robust fitting of ellipses to non-complete and contaminated data, *Czech. J. Phys.* 50, 347–354.
- Chernov, N. & Lesort, C. 2004. Statistical efficiency of curve fitting algorithms, *Comp. Stat. Data Anal.*, 47, pp. 713–728.
- Chernov, N. 2007. On the convergence of fitting algorithms in computer vision, *J. Math. Imag. Vision* 27, 231–239.
- Chernov, N. 2010. Circular and linear regression: Fitting circles and lines by least squares, *Chapman & Hall/CRC Monographs on Statistics and Applied Probability* 117.
- Chojnacki, W., Brooks, M. J. & van den Hengel, A. 2001. Rationalising the renormalisation method of Kanatani, *J. Math. Imag. Vision*, 14, 21–38.
- Chojnacki, W., Brooks, M. J., Van den Hengel, A. & Gawley, D. 2005. FNS, CFNS and HEIV: A unifying approach, *J. Math. Imag. Vision*, 23, 175–183.
- Chow, G. C. 1960. Test of Equality Between Subsets of Coefficients in Two Linear Regressions, *Econometrica*, 591-605.
- Doran, H. E., O'Donnell, C. J., Rambaldi, A. N. 2003. Imposing Observation-Varying Equality Constraints Using Generalised Restricted Least Squares. *ISSN 1446-5523*: 323.
- Draper, N.R. and Smith, H. 2014. *Applied Regression Analysis*.
- Eberly, D. H. 2007. *3D Game Engine Design*, 2nd ed., Morgan Kaufmann Publishers, San Francisco, CA.
- Eberly, D. H. 1999. *Least Squares Fitting of Data*.
- Gander, W., Golub, G.H. & Strelbel, R. 1994. *Least Squares Fitting of Circles and Ellipses*.
- Gujarati, D.N. 1995. *Basic Econometrics*, United States Military Academy, West Point, Fourth Edition.
- Kanatani, K. 1994. Statistical bias of conic fitting and renormalization, *IEEE Trans. PAMI*, 16, 320–326.
- Knopov, P. S. and Korkhin, A. S. 2012. *Regression Analysis Under A. Priori Parameter Restrictions A.S.*
- Linnik, Y. V. 1961. *Method of least squares and principles of the theory of observations*, New York, Pergamon Press.
- Fisher, F. M. 1970. Test of Equality Between Subsets of Coefficients in Two Linear Regressions: An Expository Note, *Econometrica*, 361-66.
- Golub, G. H. and Van Loan, C. F. 1989. *Matrix Computations*.
- Graybill, F.A. 1969. *Introduction to Matrices with Applications in Statistics*. Wadsworth, Belmont CA.
- Irina, A., Sergejs, A. 2007. Application of Ordinary Least Square Method in Nonlinear Models, *International Statistical Institute, 56th Session*.

- Harvey, N. 2011. Notes on Symmetric Matrices, University of British Columbia.
- Hocking, R. R. 1985. The Analysis of Linear Models.
- Johnson, R. A. and Wichern D. W. 1982. Applied Multivariate Statistical Analysis.
- Krottnerus, P. 2016. On New Variance Approximations for Linear Models with Inequality Constraints, *Statistica Neerlandica* Vol. 70, nr. 1, pp. 26-46.
- Machireddy, R. 1992. Sampling techniques for plane geometry for coordinate measuring machines, MS Thesis, UNCC.
- Meyer, C. D. 2000. Matrix Analysis and Applied Linear Algebra.
- Monahan, J. F. 2008. A Primer on Linear Models.
- Plata, S. 2006. *Applied Mathematics Letters*, 19, 499-502.
- Pratt, V. 1987. Direct Least Squares Fitting of Algebraic Surfaces, *ACM J. Computer Graphics*, Volume 21, Number 4.
- Puntanen, S., Styan, G. P. H., Isotalo, J. 2011. Matrix Tricks for Linear Statistical Models, Springer-Verlag Berlin Heidelberg.
- Rao, C.R. 1973. Linear Statistical Inference and Its Applications, John Wiley & Sons, Inc.
- Rencher, A. C. 2007. Schaalje G. Bruce, Linear Models in Statistics.
- Richard, A.J. and Dean W.W. 2002. Applied Multivariate Statistical Analysis.
- Rorres, A. 1991. Element Linear Algebra Application Version.
- Sampson, P. D. 1982. Fitting conic sections to very scattered data: an iterative refinement of the Bookstein algorithm, *Comp. Graphics Image Proc.* 18, 97–108.
- Srivastava, V. K. Regression Analysis Under Linear Restrictions and Preliminary Test Estimation. <http://home.iitk.ac.in/~shalab/econometrics/Chapter6-Econometrics-RegressionAnalysisUnderLinearRestrictions.pdf>-(Erişim tarihi: 20.02.2018).
- Shapiro, A. 1988. Towards a Unified Theory of Inequality Constrained Testing in Multivariate Analysis, University South Africa.
- Sturm, P. & Gargallo, P. 2007. Conic fitting using the geometric distance, *Proc. Asian Conf. Comp. Vision*, Tokyo, Japan, 2, pp. 784–795.
- Taubin, G. 1991. Estimation of planar curves, surfaces and nonplanar space curves defined by implicit equations, with applications to edge and range image segmentation, *IEEE Trans. PAMI*, 13, 1115–1138.
- Tebbs, J.M. 2010. Linear Statistical Models, Department of Statistics, The University of South Carolina.
- Trefethen, L. N. and Bau, D. 1997. Numerical Linear Algebra.
- Werner, H.J. 1990. Linear Algebra and Its Application 127, 379-392.
- Wickens, T. D. 2004. The General Linear Model, University Press of California, Los Angeles.

- Wijewickrema, S., Paplinski, A. P. & Esson, C. 2006. Orthogonal distance fitting revisited, Tech. report, Clayton School Inf. Technol., Monash U., Melbourne, 2006/205.
- Zhang, Z. 1997. Parameter Estimation Techniques: A Tutorial with Application to Conic Fitting, Intern. J. Image Vision Comput., 15, 59–76.

ÖZGEÇMİŞ

Adı Soyadı : Fatma Buğlem YALÇIN
Doğum Yeri : Ordu
Doğum Tarihi : 24.02.1988
Yabancı Dili : İngilizce
E-mail : yalcinfatmabuglem@gmail.com -
bug_leeem@hotmail.com
Telefon : 05377161414 – 05456017614

Öğrenim Durumu :

Derece	Bölüm/ Program	Üniversite	Yıl
Lisans	Matematik	Ondokuz Mayıs Üniversitesi	2011
Y. Lisans	Matematik	Ordu Üniversitesi	2014
Lisans	Endüstri Mühendisliği	Ondokuz Mayıs Üniversitesi	2018

Uluslararası Bildiriler :

1. Equality Constrained Least Squares and Maximum Likelihood Estimators in Classical Linear Regression Model, International Conference on Advances in Natural and Applied Sciences, 18-21 April 2017, Antalya, Turkey.
2. Different Geometric Interpretations of Correlation Coefficient, International Conference on Advances in Natural and Applied Sciences, 18-21 April 2017, Antalya, Turkey.

3. Easy Correlation Tricks, Geometric View on Partial Correlation and Correlation Between Two Random Events, International Conference on Advances in Natural and Applied Sciences, 18-21 April 2017, Antalya, Turkey.

4. Recently Some Important Studies Delivered on the Least Squares Fitting of Quadratic Curves and Surfaces, International Conference on Mathematics and Mathematics Education, 27-29 June 2018, Ordu, Turkey.