T.C

ORDU ÜNİVERSİTESİ

FEN BİLİMLER ENSTİTÜSÜ

LİNEER MODELLERDE EŞİTLİK VE EŞİTSİZLİK KISITLAMALARI ALTINDA PARAMETRE TAHMİNLERİ VE BAZI HİPOTEZLERİN TESTLERİ

CANAN KOÇAK

Bu tez,

Matematik Anabilim Dalında

Yüksek Lisans

derecesi için hazırlanmıştır.

ORDU – 2013

TEZ ONAY

Ordu Üniversitesi Fen Bilimleri Enstitüsü öğrencisi Canan KOÇAK tarafından ve Prof. Dr. Cemil YAPAR danışmanlığında hazırlanan "Lineer Modellerde Eşitlik ve Eşitsizlik Kısıtlamaları Altında Parametre Tahminleri ve Bazı Hipotezlerin Testleri" adlı bu tez, jürimiz tarafından 09 /07/2013 tarihinde oy birliği ile Matematik Anabilim Dalında Yüksek Lisans tezi olarak kabul edilmiştir.

Danışman : Prof. Dr. Cemil YAPAR

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

Üye : Doç. Dr. Selahattin MADEN
Matematik, Ordu Üniversitesi
İmza:

Üye : Yrd. Doç. Dr. Nurgül OKUR BEKAR
İstatistik, Giresun Üniversitesi
İmza:

ONAY:

Bu tezin kabulü, Enstitü Yönetim Kurulu'nun 06 /09/2013 tarih ve 2013/237. sayılı kararı ile onaylanmıştır.

06/09/2013
Enstitü Müdürü
Doç. Dr. M. Fikret BALTA

TEZ BİLDİRİMİ

Tez yazım kurallarına uygun olarak hazırlanan 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.

İmza

Canan KOÇAK

Not: Bu tezde kullanılan özgün ve başka kaynaktan yapılan bildirimlerin, çizelge, şekil ve fotoğrafların kaynak gösterilmeden kullanımı, 5846 sayılı Fikir ve Sanat Eserleri Kanunundaki hükümlere tabidir.

ÖZET

LİNEER MODELLERDE EŞİTLİK VE EŞİTSİZLİK KISITLAMALARI ALTINDA PARAMETRE TAHMİNLERİ VE BAZI HİPOTEZLERİN TESTLERİ

Canan KOÇAK

Ordu Üniversitesi
Fen Bilimler Enstitüsü
Matematik Anabilim Dalı, 2013
Yüksek Lisans Tezi, 83s.

Danışman: Prof. Dr. Cemil YAPAR

Bu tez üç bölümden ibarettir. Giriş bölümünde Lineer Modellerde kullanılan matris cebiri ele alınmıştır. İkinci bölümde, materyal ve yöntem bölümünde, Lineer Modeller ve Tahmin Edilebilirlik ile ilgili temel kavramlara yer verilmiştir.

Son bölümde Lineer Modellerde Eşitlik ve Eşitsizlik Kısıtlamaları Altında Parametre Tahminleri ve Bazı Hipotezlerin Testlerine yer verilmiştir.

Anahtar Kelimeler: Lineer *Modeller, Parametre Tahmini, Kısıtlanmış model, Kısıtlanmamış Model*

ABSTRACT

PARAMETER ESTIMATIONS AND TESTS OF SOME HYPOTHESES IN LINEAR MODELS SUBJECT TO EQUALITY AND INEQUALITY CONSTRAITS

Canan KOÇAK

University of Ordu
Institute for Graduate Studies in Science and Technology
Department of Mathematic, 2013
MSc. Thesis, 83p.

Supervisor: Prof. Dr. Cemil YAPAR

This thesis consist three chapters. In the instroduction part, Matrix algebra used in linear models has been discussed. In second chapter, entitled materials and methods have been inclueded with the concepts of linear models and estimability.

Finally, parameter estimation and hypothesis tests subject to equality and inequality constraits in linear models have been given in the last chapter.

Key Words: *Linear models, parameter estimation, Restricted linear model, unrestricted linear model.*

TEŞEKKÜR

Tüm çalışmalarım boyunca her zaman bilgi ve deneyimleriyle yolumu açan değerli hocam Sayın Prof. Dr. Cemil YAPAR' a ve Matematik lisans eğitimim esnasında yetişmemde büyük emekleri olan Matematik Bölümü öğretim üyelerine ve tüm bölüm çalışanlarına en samimi duygularım ile teşekkürlerimi sunarım.

İÇİNDEKİLER

SİMGELER VE KISALTMALAR

E^n	: n-boyutlu Öklid Uzayı
$\lVert\ \rVert$	: Norm
$\Re(\ )$	: Matrisinin satır uzayı
BLUE	: En İyi Lineer Yansız Tahmin
$C(\ )$	: Matrisin sütun uzayı
Con	: Kısıtlanmış
GM	: Gauss- Markov
HKT	: Hata Kareler Toplamı
köşeg	: Köşegen
$N(\ )$	: Sıfır uzayı
R( )	: Rank
Unc	: Kısıtlanmamış
$D(\ )$	: Dağılım Matrisi
kov( )	: Kovaryans
var( )	: Varyans

ÖNCEKİ ÇALIŞMALAR

Hallmos, (1958), Hoffman ve Kunze (1971) yaptıkları çalışmalarda vektör uzaylarını ve izdüşümleri ayrıntılı bir şekilde ele almışlar ve önemli sonuçlar ortaya koymuşlardır.

Burdick ve ortakları (1974) kısıtlamış model ile ilgili yaptıkları çalışmalarda ömenli sonuçlar ortaya koymuşlar ve $\mathbf{X(X'X)^{-1}, X'}$ ve $\mathbf{X}$, $\mathbf{(X'X)^{-1}X'}$ dönüşüm çiftlerini sırasıyla "en iyi tahmin köprüsü" ve "en iyi uyum köprüsü" olarak adlandırmışlardır.

Coleman, (1977) yaptığı bir çalışmada geometrik yaklaşımın, hem kısıtlanmış hem de kısıtlanmamış modellere uygulanan tahmin ve hipotez testi için tutarlı bir teoriyi verdiğini incelemiş ve önemli sonuçlar ortaya koymuştur.

Charles J. Monlezun, F.M. Speed. (1980), kısıtlanmış lineer modelde tahmin ve hipotez testinin geometrisini tam-ranklı durum için incelemiş ve önemli sonuçları ortaya koymuştur.

R.B. Bapat, (2000), Lineer modeller ve lineer modellerde matris cebirini ayrıntılı bir şekilde incelemiş ve birçok teoremi ortaya koyarak ispatlamıştır.

1. GİRİŞ

1.1. Matris Cebiri

Bu kısımda belirli temel kavramları ele alacağız. Sadece reel matrisleri göz önüne alacağız. İşlemimiz gerekli kısımları kapsamakla beraber, okuyucunun matrisler üzerindeki temel işlemleri bildiği kabul edilir. Determinantın basit özelliklerinin bilgisine de sahip olduğumuzu kabul edeceğiz.

Tanım 1.1.1. Bir mxn matris, m satır ve n sütunda düzenlenen mn tane reel sayıdan ibarettir. Matrisleri koyu harflerle göstereceğiz. **A** matrisinin i-yinci satır, j-yinci sütunundaki elemanı; a_{ij} ile gösterilir. Bir mx1 matrisine m mertebeli bir sütun vektör; aynı şekilde bir 1xn matrisine de n mertebeli bir *satır vektör* denir. Eğer $m=n$ ise bir mxn matrisine bir kare matris denir.

Eğer **A, B** mxn matrisler ise, bu takdirde, **A+B** ; (i, j)-yinci elemanı $a_{ij}+b_{ij}$ olan mxn matris gibi tanımlanır. Eğer **A** bir matris ve c bir reel sayı ise, bu takdirde, c**A**, **A**'nın her bir elemanını c ile çarpmak suretiyle elde edilir.

Eğer **A** bir mxp ve **B** bir pxn matris ise, bu takdirde, **C=AB** çarpımı (i,j)-yinci elemanı

$$c_{ij} = \sum_{k=1}^{p} a_{ik} b_{kj}$$

ile verilen bir mxn matristir. Aşağıdaki özellikleri sağlanır:

(AB)C=A(BC)

A(B+C)=AB+AC

(A+B)C=AC+BC

A mxn matrisinin $\mathbf{A}'$ ile gösterilen transpozesi, (i,j)-yinci elemanı a_{ji} olan bir nxm matristir. $(\mathbf{A}')' = \mathbf{A}, (\mathbf{A}+\mathbf{B})' = \mathbf{A}'+\mathbf{B}', (\mathbf{AB})' = \mathbf{B}'\mathbf{A}'$ olduğu gerçeklenebilir.

Matris çarpımının tanımının iyi anlaşılması oldukça faydalıdır. Sık sık gerekli olan bazı basit gerçekleri belirtelim. Burada ortaya çıkan tüm çarpımların, matrislerin mertebelerinin onları çarpım için uygun kılabilmesi anlamında, tanımlı olduklarını kabul edeceğiz.

(i) **AB**'nin j-yinci sütunu, **B**'nin j-yinci sütunu ile **A**'nın çarpımının aynısıdır.

(ii) **AB**'nin i-yinci satırı, **A**'nın i-yinci satırı ile **B**'nin çarpımının aynısıdır.

(iii) **ABC**'nin (i,j)-yinci elemanı, $(x_1, x_2, ..., x_p)$, **A**'nın i-yinci satırı ve $(y_1, y_2, ..., y_q)$, **C**'nin j-yinci sütunu olmak üzere,

$$(x_1, x_2, ..., x_p)\mathbf{B}\begin{bmatrix} y_1 \\ \vdots \\ y_q \end{bmatrix}$$

olarak elde edilir.

(iv) $\mathbf{a_i}$'ler **A**'nın sütunlarını ve $\mathbf{b'_j}$'ler de **B**'nin satırlarını göstermek üzere, eğer

$$\mathbf{A} = [\mathbf{a_1}, ..., \mathbf{a_n}] \text{ ve } \mathbf{B} = \begin{bmatrix} \mathbf{b'_1} \\ \vdots \\ \mathbf{b'_n} \end{bmatrix}$$

ise, bu takdirde,

$$\mathbf{AB} = \mathbf{a_1 b'_1} + ... + \mathbf{a_n b'_n}$$

dür.

Tanım 1.1.2. Bir köşegen matris, $i \neq j$ için $a_{ij} = 0$ olacak şekilde bir **A** kare matrisidir.

$$\begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

köşegen matrisini, köşeg($\lambda_1, ..., \lambda_n$) ile göstereceğiz. Her i için λ_i=1 olduğunda, bu matris, $\mathbf{I_n}$ ile göstereceğimiz veya mertebe içerikten açık ise basit olarak **I** ile göstereceğimiz n-mertebeli birim matrise indirgenir. Herhangi bir **A** kare matrisi için **AI=IA=A** olduğunu belirtelim.

Tanım 1.1.3. $a_{11}, ..., a_{nn}$ elemanlarına **A**'nın asıl köşegen elemanları denir. **A**'nın izi, iz(**A**)= $a_{11}+,\ldots,+a_{nn}$

olarak tanımlanır. Eğer **A** ve **B** matrisleri hem **AB** hem de **BA** tanımlı olacak şekilde matrisler ise, bu takdirde,

iz(**AB**)=iz(**BA**)

olduğu bu tanımdan görülür.

Tanım 1.1.4. Bir **A** nxn matrisinin, $|\mathbf{A}|$ ile gösterilen determinantı, toplam, $\{1,\ldots,n\}$ 'in tüm $\{\sigma(1),\ldots,\sigma(n)\}$ permütasyonları üzerinden olmak üzere ve $\varepsilon(\sigma)$, σ 'nın çift veya tek oluşuna göre 1 veya -1 olmak üzere,

$$|\mathbf{A}| = \sum_{\sigma} \varepsilon(\sigma) a_{1\sigma(1)} \ldots a_{n\sigma(n)}$$

olarak tanımlanır.

Determinantın bazı temel özelliklerini ispatsız ifade edeceğiz.

(i) Determinant, bir satır veya bir sütun boyunca açmak suretiyle değerlendirilebilir. Bu nedenle $\mathbf{A}_{1j}$, **A**'nın birinci satır ve j-yinci sütununu silmek suretiyle elde edilen alt matris olmak üzere, birinci satır boyunca açma,

$$|\mathbf{A}| = \sum_{j=1}^{n} (-1)^{1+j} a_{1j} |\mathbf{A}_{1j}|$$

dır.

$$\sum_{j=1}^{n} (-1)^{1+j} a_{ij} |\mathbf{A}_{1j}| = 0, \quad i=2,\ldots,n$$

olduğuna da dikkat edelim.

(ii) Eğer iki satırın (veya sütunun) yerleri değiştirilirse, determinant işaret değiştirir.

(iii) Eğer bir satırın bir sabit katına başka bir satır eklenirse, determinant değişmez. Benzer özellik sütun için de geçerlidir.

(iv) Diğer tüm sütunlar (satırlar) sabit tutulduğunda, determinant herhangi bir sütununun (satırının) bir lineer fonksiyonudur.

(v) $|\mathbf{AB}| = |\mathbf{A}||\mathbf{B}|$ dir.

Tanım 1.1.5. Eğer $i > j$ için $a_{ij} = 0$ ise, bu takdirde, **A** matrisi üst üçgenseldir. Bir üst üçgensel matrisin transpozesi alt üçgenseldir.

Tanım 1.1.6. Her bir $\mathbf{A}_{ij}, \mathbf{B}_{ij}$ 'nin kendisi bir matris olmak üzere,

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix} \quad \textbf{ve} \quad \mathbf{B} = \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{bmatrix}$$

iki matris olsun. Eğer matris çarpımı için uygunluk tümüyle kabul edilirse, bu durumda, matrislere uygun şekilde parçalanmışlardır diyeceğiz. Bu takdirde,

$$\mathbf{AB} = \begin{bmatrix} \mathbf{A}_{11}\mathbf{B}_{11} + \mathbf{A}_{12}\mathbf{B}_{21} & \mathbf{A}_{11}\mathbf{B}_{12} + \mathbf{A}_{12}\mathbf{B}_{22} \\ \mathbf{A}_{21}\mathbf{B}_{11} + \mathbf{A}_{22}\mathbf{B}_{21} & \mathbf{A}_{21}\mathbf{B}_{12} + \mathbf{A}_{22}\mathbf{B}_{22} \end{bmatrix}$$

yazabiliriz.

1.2. Vektör Uzayları ve Altuzaylar

Tanım 1.2.1. Eğer aşağıdaki şartlar sağlanırsa, boş olmayan bir S kümesine bir vektör uzayı denir.

(i) S'deki herhangi **x, y** elemanı için, **x+y** tanımlıdır ve S'dedir. Ayrıca,

x+y=y+x (değişme özelliği)

ve

x+(y+z)=(x+y)+z (birleşme özelliği)

vardır.

(ii) Her **x** için **x+0=x** olacak şekilde S'de, **0** ile gösterilen bir eleman mevcuttur.

(iii) S'deki herhangi bir **x** elemanı için **x+y=0** olacak şekilde, S'de bir **y** elemanı mevcuttur.

(iv) S'deki herhangi bir **x** elemanı ve herhangi bir c reel sayısı için c**x** tanımlıdır ve S'dedir. Ayrıca herhangi bir **x** elemanı için 1**x**=**x** dir.

(v) S'deki herhangi $\mathbf{x}_1$, $\mathbf{x}_2$ elemanları ve c_1, c_2 reel sayıları için,

$$c_1(\mathbf{x}_1 + \mathbf{x}_2) = c_1\mathbf{x}_1 + c_1\mathbf{x}_2$$

$$(c_1 + c_2)\mathbf{x}_1 = c_1\mathbf{x}_1 + c_2\mathbf{x}_1$$

$$c_1(c_2\mathbf{x}_1) = (c_1c_2)\mathbf{x}_1$$

dir. S'deki elemanlara vektörler denir. Eğer **x, y** vektör iseler, bu takdirde onların **x+y** toplamını alma işlemine vektör toplamı olarak başvurulur. (ii)'deki vektöre(yani **0**'a) sıfır vektör denir. (iv)'deki işleme skalarla çarpım denir. Bir vektör uzay herhangi bir cisme başvurarak tanımlanabilir. Amacımız için yeterli olacağından, cismi reel sayılar cismi olarak aldık.

n mertebeli sütun vektörlerin kümesi (veya nx1 matrisler) bir vektör uzayıdır. Bu nedenle n mertebeli satır vektörlerin kümesi de bir vektör uzaydır. Bu iki vektör uzayı çoğu zaman ele alacağımız vektör uzaylarıdır.

R reel sayılar kümesi olmak üzere, R^n, n kere çarpılan, R x R x ...x R kümesi olsun. R^n'nin elemanlarını, verilen bir durumdaki uygunluğa bağlı olarak ya satır vektörleri ya da sütun vektörlerinden biri olarak yazacağız.

Tanım 1.2.2. Eğer S, T vektör uzayları ve $S \subset T$ ise, bu takdirde, S'ye T'nin bir altuzayı denir.

R^3'ün tüm mümkün olabilir altuzaylarını tanımlayalım. Açık olarak R^3 bir vektör uzayıdır ve bu nedenle yalnız sıfır vektöründen ibaret olan uzay da bir vektör uzaydır. c_1, c_2, c_3 reel sayılar olsun.

$$c_1x_1 + c_2x_2 + c_3x_3 = 0$$

bağıntısını sağlayan tüm $\mathbf{x} \in R^3$ vektörlerinin kümesi, R^3'ün bir altuzayıdır (burada x_1, x_2, x_3'ler **x**'in koordinatlarıdır). Geometrik olarak, bu küme orijinden geçen bir düzlemi gösterir. Orijinden geçen iki farklı düzlemin ara kesiti orijinden geçen bir doğrudur ve bu doğru da bir altuzaydır. Bunlar R^3'ün yegane mümkün olabilir altuzaylarıdır.

1.3. Taban ve Boyut

Tanım 1.3.1 $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_m$ vektörlerinin lineer gereni (veya bu vektörlerle gerilen uzay) $c_1, c_2, ..., c_m$ reel sayılar olmak üzere, tüm $c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + ... + c_m\mathbf{x}_m$ lineer kombinasyonlarının kümesi olarak tanımlanır. Lineer gerenin bir altuzayı olduğu tanımdan görülür.

Tanım 1.3.2. Eğer en az bir c_i sıfırdan farklı ve $c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + ... + c_m\mathbf{x}_m = \mathbf{0}$ olacak şekilde $c_1, c_2, ..., c_m$ reel sayıları varsa, $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_m$ vektörlerinin kümesine lineer bağımlıdır denecektir. Bir küme lineer bağımlı değilse, lineer bağımsızdır. Açıkçası yukarıdaki iki tanımda, bir vektörler kümesinden ziyade, bir vektörler demetine başvurmalıyız. Bu nedenle lineer bağımlı veya bağımsız olan $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_m$ vektörlerinden söz ettiğimizde, ister istemez farklı olmayan vektörlerin mümkün olabilirliğine izin veririz.

Aşağıdaki ifadeler kolayca ispatlanır;

(i) Sadece sıfır vektöründen ibaret olan küme lineer bağımlıdır.

(ii) Eğer $X \subset Y$ ve X lineer bağımlı ise, bu takdirde, Y de lineer bağımlıdır.

(iii) Eğer $X \subset Y$ ve Y lineer bağımsız ise, bu takdirde, X de lineer bağımsızdır.

Bir vektörler kümesine bir S vektör uzayı için bir taban oluşturur denecektir, eğer bu küme lineer bağımsız ve onun gereni S'ye eşitse.

$\mathbf{e}_i$, nxn birim matrisinin i-yinci sütunu olsun. $\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n$ kümesi, R^n için, standart taban denilen bir taban oluşturur. Eğer $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_m$ S için bir taban ise, bu takdirde, S'deki herhangi bir $\mathbf{x}$ vektörü bir $c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + ... + c_m\mathbf{x}_m$ lineer kombinasyonu olarak tek bir gösterimi kabul eder. Eğer

$$\mathbf{x} = c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + ... + c_m\mathbf{x}_m = d_1\mathbf{x}_1 + d_2\mathbf{x}_2 + ... + d_m\mathbf{x}_m$$

şeklinde iki türlü gösterilseydi, bu takdirde,

$$(c_1 - d_1)\mathbf{x}_1 + (c_2 - d_2)\mathbf{x}_2 + ... + (c_m - d_m)\mathbf{x}_m = \mathbf{0}$$

dır ve $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_m$'ler lineer bağımsız olduklarından, her bir i için $c_i = d_i$ olacaktır. Bir vektör uzayı eğer sonlu bir çok vektörden ibaret olan bir tabana sahip ise, bu vektör uzayına sonlu-boyutlu denir. Yalnız sıfır vektörünü ihtiva eden vektör uzayı da sonlu-boyutludur. Sadece sonlu boyutlu vektör uzaylarını göz önüne alacağız. Genellikle sıfır vektöründen başka diğer vektörleri de içeren vektör uzaylarını ele alacağız.

Teorem 1.3.1. S bir vektör uzayı olsun. Bu takdirde S'nin herhangi iki tabanı aynı sayıda vektöre sahiptir.

İspat: $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_p$ ve $\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_q$'nin S için taban olduklarını farz edelim. Mümkünse, $p > q$ olsun. Her $\mathbf{x}_i$'yi $\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_q$'nun bir lineer kombinasyonu olarak ifade edebiliriz. Bu nedenle,

$$\mathbf{x}_i = \sum_{j=1}^{q} a_{ij}\mathbf{y}_j, \qquad i=1,...,p \tag{1.3.1}$$

olacak şekilde bir $\mathbf{A} = (a_{ij})$, pxq matrisi vardır. Aynı şekilde,

$$\mathbf{y}_j = \sum_{j=1}^{p} b_{jk}\mathbf{x}_k, \qquad j=1,...,q \tag{1.3.2}$$

olacak şekilde bir şekilde bir $\mathbf{B} = (b_{ij})$, q x p matrisi vardır.

C=AB olmak üzere, (1.3.1) ve (1.3.2) bağıntılarından,

$$\mathbf{x}_i = \sum_{k=1}^{p} c_{ik}\mathbf{x}_k, \qquad i=1,...,p \tag{1.3.3}$$

olduğunu görürüz. (1.3.3) bağıntısından ve Teorem 1.3.1'den evvel yapılan gözlemden **AB=I** olduğu görülür. Burada **I**, p boyutlu bir birim matristir. **U** pxp matrisini elde etmek için, **A**'ya p-q tane sıfır sütun ekleyelim. Aynı şekilde **V** pxp matrisini elde etmek için de **B**'ye p-q tane sıfır satır ekleyelim. Bu takdirde **UV=AB=I** dır. Bu nedenle $|\mathbf{UV}| = 1$dir. Bununla beraber, **U** bir sıfır sütuna ve **V** de bir sıfır satıra sahip olduğundan, $|\mathbf{U}| = |\mathbf{V}| = 0$ dir. Böylece bir çelişkiye sahip oluruz ve bu nedenle $p \le q$ dur. Aynı şekilde $q \le p$ olduğunu da gösterebiliriz. O halde $q = p$ dir.

Teorem 1.3.1'in ispatlanması süresince faydalı olacak olan aşağıdaki ifadeyi ispatladık. S bir vektör uzayı olsun $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_p$'nin S için bir taban olduğunu ve $\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_q$ kümesinin de S'yi gerdiğini farz edelim. Bu takdirde $p \le q$ dur.

Boy(S) ile gösterilen, S vektör uzayının boyutu, S'nin bir tabanını oluşturan bağımsız vektörlerin sayısıdır. Sadece sıfır vektöründen ibaret olan uzayın boyutu sıfır kabul edilir.

Tanım 1.3.3. S ve T vektör uzayları olsun. f, lineer, yani, S'deki her **x, y** elemanı ve reel c sayıları için $f(\mathbf{x}+\mathbf{y})=f(\mathbf{x})+f(\mathbf{y})$ ve $f(c\mathbf{x})=cf(\mathbf{x})$ olacak şekilde, eğer üzerine bire-bir $f:S\to T$ dönüşümü varsa S, T'ye izomorftur denir.

Teorem 1.3.2. S, T vektör uzayları olsun. Bu takdirde S, T'nin izomorf olmaları için gerek ve yeter şart, boy(S)=boy(T) olmasıdır.

İspat: İlk olarak yeter kısmını ispatlayacağız. $f:S\to T$'nin bir izomorfizm olduğunu farz edelim. Eger $\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_k$ S için bir taban ise, bu takdirde, $f(\mathbf{x}_1), f(\mathbf{x}_2),...,f(\mathbf{x}_k)$'nın T için bir taban olduğunu göstereceğiz. İlk olarak $c_1 f(\mathbf{x}_1)+c_2 f(\mathbf{x}_2)+...+c_k f(\mathbf{x}_k)=\mathbf{0}$ olduğunu farz edelim. İzomorfizmin tanımından $f(c_1\mathbf{x}_1+c_2\mathbf{x}_2+...+c_k\mathbf{x}_k)=0$ olduğu görülür ve buradan da $c_1\mathbf{x}_1+c_2\mathbf{x}_2+...+c_k\mathbf{x}_k=\mathbf{0}$ dır. $\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_k$'lar lineer bağımsız olduklarından, $c_1=c_2=...=c_k=0$ dır ve böylece $f(\mathbf{x}_1), f(\mathbf{x}_2),...,f(\mathbf{x}_k)$'lar lineer bağımsızdır. Eğer $\mathbf{v}\in T$ ise, bu takdirde, $f(\mathbf{u})=\mathbf{v}$ olacak şekilde $\mathbf{u}\in S$ mevcuttur. Herhangi $d_1,d_2,....d_k$ için $\mathbf{u}=d_1\mathbf{x}_1+d_2\mathbf{x}_2+...+d_k\mathbf{x}_k$ yazabiliriz. Şimdi $\mathbf{v}=f(\mathbf{u})=d_1 f(\mathbf{x}_1)+...+d_k f(\mathbf{x}_k)$ dır. Bu nedenle $f(\mathbf{x}_1), f(\mathbf{x}_2),...,f(\mathbf{x}_k)$, T'yi gerer ve böylece T için bir taban oluşturur. Bu da gösterir ki boy(T)=k dır.

Tersini ispatlamak için, $\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_k$; $\mathbf{y}_1,\mathbf{y}_2,...,\mathbf{y}_k$ sırasıyla S, T için taban olsunlar (boy(S)=boy(T) olduğundan, tabanlar aynı sayıda bağımsız vektörlere sahiptirler). S'deki herhangi bir **x** vektörü, bir tek

$$\mathbf{x}=c_1\mathbf{x}_1+c_2\mathbf{x}_2+...+c_k\mathbf{x}_k$$

gösterimine sahiptir. $\mathbf{y}=c_1\mathbf{y}_1+c_2\mathbf{y}_2+...+c_k\mathbf{y}_k$ olmak üzere, $f(\mathbf{x})=\mathbf{y}$'yi tanımlayalım. f'nin izomorfizmin tanımını sağladığı gerçeklenebilir.

Teorem 1.3.3. S bir vektör uzayı olsun ve S'nin $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_m$ vektörlerinin lineer gereni olduğunu farz edelim. Eğer herhangi bir $\mathbf{x}_i$, $\mathbf{x}_1, ..., \mathbf{x}_{i-1}, \mathbf{x}_{i+1}, ..., \mathbf{x}_m$'in bir lineer kombinasyonu ise, bu takdirde, bu son vektörler de S'yi gerer.

Teorem 1.3.4. S, n boyutlu bir vektör uzayı olsun ve $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_m$ S'deki lineer bağımsız vektörler olsun. Bu takdirde S için $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_m$'i içeren bir taban mevcuttur.

İspat: $\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_n$ S için bir taban olsun. $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_m, \mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_n$ kümesi lineer bağımlıdır. Bu nedenle, bazı c_i veya d_i'ler sıfırdan farklı olmak üzere bir

$$c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + ... + c_m\mathbf{x}_m + d_1\mathbf{y}_1 + ... + d_n\mathbf{y}_n = \mathbf{0}$$

lineer kombinasyonu mevcuttur. Bununla beraber , $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_m$'ler lineer bağımsız olduklarından herhangi bir d_i'nin sıfırdan farklı olduğunun doğru olması gerekir. Bu nedenle herhangi bir $\mathbf{y}_i$, geri kalan vektörlerin bir lineer kombinasyonudur. Teorem 1.3.3'e göre,

$$\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_m,\ \mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_{i-1}, \mathbf{y}_{i+1}, ..., \mathbf{y}_n$$

kümesi de S'yi gerer. Eğer küme lineer bağımsız ise, bu takdirde, istenen tabanı elde ederiz. Aksi takdirde, işlemi $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_m$'yi içeren bir taban elde edinceye kadar sürdürürüz.

Teorem 1.3.5. R^n'de $n \cdot 1$ sayıda vektörün herhangi bir kümesi lineer bağımlıdır.

ispat: Eğer küme lineer bağımsız ise, bu taktirde, Teorem 1.3.4'e göre R^n için kümeyi ihtiva eden bir taban bulabiliriz. R^n için her tabanın kesin olarak n sayıda vektör içermesi gerektiğinden bu bir çelişmedir.

Teorem 1.3.6. R^n'in herhangi bir S altuzayı bir tabana sahiptir.

İspat: S'de her bir aşamada lineer bağımsız olacak şekilde ardışık olarak $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_m$ vektörlerini seçelim. Herhangi bir aşamada eğer vektörler S'yi gererse, bu takdirde bir taban elde ederiz. Aksi takdirde S'de $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_m$'in lineer gerenin de olmayan bir $\mathbf{x}_{m+1}$ vektörü vardır ve lineer bağımsız olan $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_m$, $\mathbf{x}_{m+1}$ kümesine ulaşırız. Teorem 1.3.5.'e göre, R^n 'deki vektörler lineer bağımlı olduklarından işlem sona ermelidir.

Teorem 1.3.7. Eğer S, T'nin bir altuzayı ise, bu takdirde, boy(S)$\leq$boy(T) dir. Ayrıca eşitliğin sağlaması için gerek ve yeter şart S=T olmasıdır.

Ispat: Yalnız sonlu-boyutlu vektör uzaylarını ele aldığımızı hatırlayalım. boy(S)=p, boy(T)=q olduğunu varsayalım ve $\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_p$ ve $\mathbf{y}_1,\mathbf{y}_2,...,\mathbf{y}_q$ sırasıyla S ve T için tabanlar olsun. Teorem 1.3.6'nın ispatındakine benzer bir muhakemeyi kullanarak, eğer $r>q$ ise, bu takdirde, T'deki r tane vektörün herhangi bir kümesinin lineer bağımlı olduğunu gösterebiliriz. $\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_p$, $S\subset T$'deki vektörlerin bir lineer bağımsız kümesi olduğundan, $p\leq q$ elde ederiz.

İkinci kısmı ispatlamak için $p=q$ ve $S\neq T$ olduğunu varsayalım. Bu takdirde $\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_p$'nin gereninde olmayan bir $\mathbf{z}\in T$ vektörü mevcuttur. Bu durumda $\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_p,\mathbf{z}$ kümesi lineer bağımsızdır. Daha evvel yapılan uyarmaya göre T'deki herhangi p+1 sayıdaki vektör lineer bağımlı olması gerektiğinden bu bir çelişmedir. Böylece S, T'nin bir altuzayı ve boy(S)=boy(T) ise, bu takdirde, $S=T$ olduğunu gösterdik. Tersine olarak eğer $S=T$ ise, bu takdirde, boy(S)=boy(T) olduğu açıktır ve ispat tamamlanır.

1.4. Rank

Tanım 1.4.1. **A** bir mxn matris olsun. **A**'nın sütun vektörleri tarafından gerilen R^m 'in bir altuzayına, **A**'nın sütun uzayı veya sütun gereni denir ve $C(\mathbf{A})$ ile gösterilir. Aynı şekilde **A**'nın satırları tarafından gerilen R^n 'nin altuzayına **A**'nın satır uzayı denir ve $\Re(\mathbf{A})$ ile gösterilir. Aşikâr olarak, $\Re(\mathbf{A})$, $C(\mathbf{A}')$'ne izomorftur. Sütun uzayının boyutuna matrisin sütun rankı ve satır uzayının boyutuna da matrisin satır rankı denir. Aşağıda göstereceğimiz gibi iki rank daima eşit olduğundan bu iki tanım herhangi bir lineer cebir kitabında çok kısa olarak ortaya çıkar. İkisi beraber çok nadir kullanılır.

Teorem 1.4.1. Bir matrisin sütun rankı, onun satır rankına eşittir.

İspat: **A**, r sütun ranklı bir mxn matris olsun. Bu takdirde $C(\mathbf{A})$ $\mathbf{b_1, b_2, ..., b_r}$ diyeceğimiz r tane vektörden oluşan bir tabana sahiptir. **B**, mxr boyutlu $[\mathbf{b_1, b_2, ..., b_r}]$ matrisi olsun. **A**'nın her sütunu $\mathbf{b_1, b_2, ..., b_r}$'nin bir lineer kombinasyonu olduğundan, herhangi bir rxn boyutlu **C** matrisi için **A=BC** yazabiliriz. Bu takdirde **A**'nın her satırı **C**'nin satırlarının bir lineer kombinasyonudur ve bu nedenle $\Re(\mathbf{A}) \subset \Re(\mathbf{C})$ dır. Teorem 1.3.7'den görülür ki **A**'nın satır rankı olan $\Re(\mathbf{A})$'nın boyutu en fazla r dir. Aynı şekilde sütun rankı satır rankını aşmayacağını ve bu nedenle de ikisinin eşit olması gerektiğini gösterebiliriz.

A'nın sütun rankının ve satır rankının ortak değerine, bundan böyle **A**'nın rankı denecektir ve R(**A**) ile gösterilecektir. Bu notasyon **A**'nın satır uzayını göstermek için kullanılan notasyonla, yani $\Re(\mathbf{A})$ ile karıştırılmamalıdır.

$\mathrm{R}(\mathbf{A}) = \mathrm{R}(\mathbf{A}')$ olduğu açıktır. **A**'nın rankının sıfır olması için gerek ve yeter şart **A**'nın sıfır matrisi olmasıdır.

Teorem 1.4.2. **A** ve **B** matrisleri **AB** çarpımı tanımlı olacak şekilde matrisler olsunlar. Bu takdirde

$$\mathrm{R}(\mathbf{AB}) < \min\{\mathrm{R}(\mathbf{A}), \mathrm{R}(\mathbf{B})\}$$

dir.

İspat: $C(\mathbf{AB})$'deki bir vektör herhangi bir **x** vektörü için **ABx** biçimindedir ve bu nedenle bu vektör $C(\mathbf{A})$'ya aittir. Böylece, $C(\mathbf{AB}) \subset C(\mathbf{A})$ ve bu nedenle Teorem 1.3.7'ye göre,

$$\mathrm{R}(\mathbf{AB}) = \mathrm{boy}C(\mathbf{AB}) < \mathrm{boy}C(\mathbf{A}) = \mathrm{R}(\mathbf{A})$$

dır. Şimdi bu gerçeği kullanarak,

$$\mathrm{R}(\mathbf{AB}) = \mathrm{R}(\mathbf{A}'\mathbf{B}') < \mathrm{R}(\mathbf{B}') = \mathrm{R}(\mathbf{B})$$

elde ederiz.

Teorem 1.4.3. **A,** $r \neq 0$ olmak üzere, r ranklı bir mxn matris olsun. Bu taktirde R(**B**)=R(**C**)=r ve **A=BC** olacak şekilde sırasıyla mxr ve rxn mertebeli **B, C** matrisleri mevcuttur. Bu ayrışıma **A**'nın bir rank ayrışımı denir.

İspat: İspat, Teorem 1.4.1'in ispatındaki aynı yolu izler. Bu nedenle **B**, mxr ve **C**, rxn matrisler olmak üzere **A=BC** yazabiliriz. **B**'nin sütunları lineer bağımsız olduklarından, R(**B**)= r dir. **C**, r tane satıra sahip olduğundan, R(**C**)$\leq$ r dir. Bununla beraber Teorem 1.4.2'ye göre r = R(**A**)$\leq$ R(**C**), bu nedenle R(**C**)= r dir.

Bu çalışmada baştan sona her ne zaman bir matrisin rank ayrışımından söz edildiğinde, onun dolaylı olarak sıfır olmayan bir matris olduğu kabul edilir.

Teorem 1.4.4. **A** ve **B** mxn matrisler olsun. Bu taktirde R(**A+B**)$\leq$R(**A**)+R(**B**) dir.

İspat: **A=XY**, **B=UV** **A, B**'nin rank ayrışımları olsun. Bu takdirde

$$\mathbf{A+B=XY+UV}=[\mathbf{X},\mathbf{U}]\begin{bmatrix}\mathbf{Y}\\ \mathbf{V}\end{bmatrix}$$

dir. Bu nedenle Teorem 1.4.2'ye göre

$$\mathrm{R}(\mathbf{A+B})\leq \mathrm{R}[\mathbf{X},\mathbf{U}]$$

olur. $\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_p$ ve $\mathbf{u}_1,...,\mathbf{u}_q$, sırasıyla $C(\mathbf{X})$, $C(\mathbf{U})$ için taban olsunlar. $[\mathbf{X},\mathbf{U}]$'nun sütun uzayındaki herhangi bir vektör $p+q$ sayıda vektörün lineer kombinasyonu olarak elde edilebilir. Böylece,

$$\mathrm{R}[\mathbf{X},\mathbf{U}]\leq \mathrm{R}(\mathbf{X})+\mathrm{R}(\mathbf{U})=\mathrm{R}(\mathbf{A})+\mathrm{R}(\mathbf{B})$$

olur ve ispat tamamdır.

Tanım 1.4.2. Bir **A** matrisi üzerinde yapılan aşağıdaki işlemlere elementer sütun işlemleri denir;

(i) **A**'nın iki sütununu yer değiştirme,

(ii) **A**'nın bir sütununu '0' olmayan bir skalarla çarpma,

(iii) Bir sütununu bir skalarla çarpıp başka bir sütuna ekleme.

Bu işlemler açık olarak C(**A**)'yı değiştirmez. Bu nedenle matrisin rankı değişmez. Aynı şekilde elementer satır işlemlerini tanımlayabiliriz. Bu elementer satır ve sütun işlemleri, özellikle hesaplamalarda kullanılır. Bu nedenle bir matrisin rankını bulmak için ilk olarak onu bu işlemlerle birçok elemanı '0' olan bir matrise indirgeriz ve bundan sonra ortaya çıkan matrisin rankını hesaplarız.

1.5. Diklik (Ortogonallik)

Tanım 1.5.1. S bir vektör uzayı olsun. S'deki her **x, y** vektör çiftine, bir reel $\langle \mathbf{x},\ \mathbf{y} \rangle$ sayısı karşılık getiren bir fonksiyona iç çarpım denecektir, eğer o aşağıdaki şartları sağlarsa:

(i) $\langle \mathbf{x},\ \mathbf{y} \rangle = \langle \mathbf{y},\ \mathbf{x} \rangle$

(ii) $\langle \mathbf{x},\ \mathbf{x} \rangle \geq 0$ ve eşitliğin sağlanması için gerek ve yeter şart **x=0** olmasıdır.

(iii) $\langle c\mathbf{x},\ \mathbf{y} \rangle = c\langle \mathbf{x},\ \mathbf{y} \rangle$

(iv) $\langle \mathbf{x}+\mathbf{y},\ \mathbf{z} \rangle = \langle \mathbf{x},\ \mathbf{z} \rangle + \langle \mathbf{y},\ \mathbf{z} \rangle$ dir.

R^n de, $\langle \mathbf{x},\ \mathbf{y} \rangle = \mathbf{x}'\mathbf{y} = x_1 y_1 + \ldots + x_n y_n$'nin bir iç çarpım olacağı kolayca görülür. Aksi belirtilmedikçe R^n ve onun altuzaylarıyla ilgilenirken bu iç çarpım ile çalışacağız.

Tanım 1.5.2. Herhangi bir **x** vektörü için $\langle \mathbf{x},\ \mathbf{x} \rangle$ iç çarpımının pozitif kareköküne **x**'in normu denir ve $\|\mathbf{x}\|$ ile gösterilir. Eğer $\langle \mathbf{x},\ \mathbf{y} \rangle = 0$ ise, bu takdirde, **x, y** vektörlerine ortogonal ya da diktir denecektir. Bu durumda $\mathbf{x} \perp \mathbf{y}$ yazarız.

Teorem 1.5.1. Eğer $\mathbf{x}_1, \ldots, \mathbf{x}_m$ ikişer ikişer dik, sıfır olmayan vektörler ise, bu takdirde, onlar lineer bağımsızdırlar.

İspat: $c_1\mathbf{x}_1 + \ldots + c_m\mathbf{x}_m = \mathbf{0}$ olduğunu varsayalım. Bu takdirde,

$$\langle c_1\mathbf{x}_1 + \ldots + c_m\mathbf{x}_m,\ \mathbf{x}_1 \rangle = 0$$

ve bu nedenle,

$$\sum_{i=1}^{m} c_i \langle \mathbf{x}_i,\ \mathbf{x}_1 \rangle = 0$$

dır. $\mathbf{x}_1, \ldots, \mathbf{x}_m$ vektörleri karşılıklı ikişer ikişer ortogonal olduklarından, $c_1\langle \mathbf{x}_1,\ \mathbf{x}_1 \rangle = 0$ dır ve $\mathbf{x}_1$ sıfırdan farklı olduğundan, $c_1 = 0$ dır. Aynı şekilde, her bir c_i'nin sıfır olduğunu gösterebiliriz. Bu nedenle, bu vektörler lineer bağımsızdırlar.

Tanım 1.5.3. Bir $\mathbf{x}_1,\ldots,\mathbf{x}_m$ vektörlerinin kümesine S vektör uzayı için bir ortonormal taban oluşturuyor denir, eğer bu küme S için bir taban ise ve bundan başka eğer $i \neq j$ olduğunda $\langle \mathbf{x}_i,\ \mathbf{x}_j \rangle=0$ ise, $i=j$ olduğunda $\langle \mathbf{x}_i,\ \mathbf{x}_j \rangle=1$ ise.

Şimdi verilen bir $\mathbf{x}_1,\ldots,\mathbf{x}_m$ tabanıyla başlayarak bir ortonormal taban üreten Gram-Schmidt yöntemini tanımlayacağız.

$\mathbf{y}_1=\mathbf{x}_1$ alalım. Tanımlanan $\mathbf{y}_1,\ldots,\mathbf{y}_{i-1}$'e sahip olarak, $a_{i,i-1},\ldots,a_{i,1}$'ler $\mathbf{y}_i$; $\mathbf{y}_1,\ldots,\mathbf{y}_{i-1}$'e ortogonal (dik) olacak şekilde seçilmek üzere,

$$\mathbf{y}_i=\mathbf{x}_i-a_{i,i-1}\mathbf{y}_{i-1}-\ldots-a_{i,1}\mathbf{y}_1$$

bağıntısını tanımlarız. Bu nedenle $\langle \mathbf{y}_i,\ \mathbf{y}_j \rangle=0,\quad j=1,\ldots,i-1$ denklemini çözmeliyiz. Bu bizi,

$$\langle \mathbf{x}_i,\mathbf{y}_j \rangle-\sum_{k=1}^{i-1} a_{ik}\langle \mathbf{y}_k,\mathbf{y}_j \rangle=0,\qquad j=1,\ldots,i-1$$

bağıntısını veren,

$$\langle \mathbf{x}_i-a_{i,i-1}\mathbf{y}_{i-1}-\ldots-a_{i,1}\mathbf{y}_1,\mathbf{y}_j \rangle=0,\qquad j=1,\ldots,i-1$$

bağıntısına götürür. Şimdi $\mathbf{y}_1,\ldots,\mathbf{y}_{i-1}$ bir ortogonal küme olduğundan,

$$\langle \mathbf{x}_i,\mathbf{y}_j \rangle-a_{ij}\langle \mathbf{y}_j,\mathbf{y}_j \rangle=0$$

dır ve bu nedenle de,

$$a_{ij}=\frac{\langle \mathbf{x}_i,\mathbf{y}_j \rangle}{\langle \mathbf{y}_j,\mathbf{y}_j \rangle},\qquad j=1,\ldots,i-1$$

bağıntılarını elde ederiz. Bu işlem ikişer ikişer ortogonal vektörlere sahip olan $\mathbf{y}_1,\ldots,\mathbf{y}_n$ tabanını elde etmek için sürdürülür. $\mathbf{x}_1,\ldots,\mathbf{x}_n$'ler lineer bağımsız olduklarından, her bir $\mathbf{y}_i$ sıfırdan farklıdır. Şimdi, eğer $\mathbf{z}_i=\dfrac{\mathbf{y}_i}{\|\mathbf{y}_i\|}$ koyarsak, bu takdirde $\mathbf{z}_1,\ldots,\mathbf{z}_n$ bir ortonormal tabandır. Her bir i için, $\mathbf{z}_1,\ldots,\mathbf{z}_i$'nin lineer gereninin $\mathbf{x}_1,\ldots,\mathbf{x}_i$'nin lineer gerenine eşit olduğuna dikkat edelim. $\mathbf{x}_1,\ldots,\mathbf{x}_m$ lineer bağımsız vektörlerinin verilen bir kümesi için yukarıda ifade edilen Gram-Schmidt yönteminin $\mathbf{y}_i$'ler $\mathbf{x}_1,...,\mathbf{x}_{i-1}$, $i=1,...,m$'nin bir lineer kombinasyonu olacak şekilde ikişer ikişer ortogonal bir $\mathbf{y}_1,...,\mathbf{y}_m$ kümesini üretmek için kullanılabildiğini hatırlatalım. Bu gerçek bir sonraki sonucun ispatında kullanılır.

W bir S vektör uzayındaki vektörlerin bir kümesi (bir altuzay olması gerekmez) olsun.

$$W^{\perp} = \{\mathbf{x} : \mathbf{x} \in S, \langle \mathbf{x}, \mathbf{y} \rangle = 0 \text{ her } \mathbf{y} \in W \text{ } i\c{c}in\}$$

kümesini tanımlayalım. Tanımlardan $W^{\perp}$'nin S'nin bir altuzayı olduğu görülür.

Teorem 1.5.2. S, T vektör uzayının bir altuzayı olsun ve $\mathbf{x} \in T$ olsun. Bu takdirde $\mathbf{u} \in S$ ve $\mathbf{v} \in S^{\perp}$ olacak şekilde bir tek **x=u+v** ayrışımı mevcuttur. **u** vektörüne **x**'in S vektör uzayı üzerindeki ortogonal izdüşümü denir.

İspat: $\mathbf{x} \in S$ ise, bu takdirde, **x=x+0** istenen ayrışımdır. Aksi takdirde, $\mathbf{x}_1,...,\mathbf{x}_m$, S için bir taban olsun. İkişer ikişer ortogonal vektörlerin $\mathbf{y}_1,...,\mathbf{y}_m,\mathbf{v}$ dizisini elde etmek için $\mathbf{x}_1,...,\mathbf{x}_m,\mathbf{x}$ kümesi üzerinde Gram-Schmidt yöntemini kullanalım **v** her bir $\mathbf{y}_i$'ye dik olduğundan ve $\mathbf{y}_1,...,\mathbf{y}_m$'in lineer gereni $\mathbf{x}_1,...,\mathbf{x}_m$'in lineer gerenine eşit olduğundan $\mathbf{v} \in S^{\perp}$ olur. Aynı zamanda Gram-Schmidt yöntemine göre, **x-v** de $\mathbf{y}_1,...,\mathbf{y}_m$'in bir lineer kombinasyonudur ve bu nedenle $\mathbf{x}-\mathbf{v} \in S$ dir. Şimdi, **x=(x-v)+v** istenen ayrışımdır. Geriye tekliği göstermek kalır.

Eğer $\mathbf{x} = \mathbf{u}_1 + \mathbf{v}_1 = \mathbf{u}_2 + \mathbf{v}_2$, $\mathbf{u}_1 \in S$, $\mathbf{u}_2 \in S$, $\mathbf{v}_1 \in S^{\perp}$, $\mathbf{v}_2 \in S$ sağlayan iki ayrışım ise, bu takdirde,

$$(\mathbf{u}_1 - \mathbf{u}_2) + (\mathbf{v}_1 - \mathbf{v}_2) = \mathbf{0}$$

dir. $\langle \mathbf{u}_1 - \mathbf{u}_2, \mathbf{v}_1 - \mathbf{v}_2 \rangle = 0$ olduğundan, $\langle \mathbf{u}_1 - \mathbf{u}_2, \mathbf{u}_1 - \mathbf{u}_2 \rangle = 0$ olduğu yukarıdaki bağıntıdan görülür. Bu takdirde $\mathbf{u}_1 - \mathbf{u}_2 = \mathbf{0}$ ve buradan da $\mathbf{u}_1 = \mathbf{u}_2$ olduğu görülür. $\mathbf{v}_1 = \mathbf{v}_2$ olduğu da kolaylıkla görülür. Böylece ayrışım tektir.

Teorem 1.5.3. W, T vektör uzayının bir altuzayı olsun ve S, W'nin lineer gereni olsun. Bu takdirde,

$$\text{boy}(S)+\text{boy}(W^{\perp})=\text{boy}(T)$$

dir.

İspat: boy(S)=m, boy($W^{\perp}$)=n, boy(T)=p olduğunu farz edelim. $\mathbf{x}_1,...,\mathbf{x}_m$ ve $\mathbf{y}_1,...,\mathbf{y}_n$ sırasıyla S, W için taban olsunlar.

$$c_1\mathbf{x}_1+...+c_m\mathbf{x}_m+d_1\mathbf{y}_1+...+d_n\mathbf{y}_n=0$$

olduğunu farz edelim. $\mathbf{u}=c_1\mathbf{x}_1+...+c_m\mathbf{x}_m$, $\mathbf{v}=d_1\mathbf{y}_1+...+d_n\mathbf{y}_n$ olsun. Her bir i, j için $\mathbf{x}_i,\mathbf{y}_j$ ortogonal olduklarından **u, v** ortogonaldırlar. Bununla beraber **u+v=0** ve bu nedenle **u=v=0** dir. Her bir i, j için $c_i=0$, $d_j=0$ olduğu görülür ve bu nedenle $\mathbf{x}_1,...,\mathbf{x}_m,\mathbf{y}_1,...,\mathbf{y}_n$ bir lineer bağımsız kümedir. Böylece m+n$\leq p$ dir. Eğer m+n<p ise, bu takdirde, $\mathbf{x}_1,...,\mathbf{x}_m,\mathbf{y}_1,...,\mathbf{y}_n$, **z** bir lineer bağımsız küme olacak şekilde bir **z** vektörü mevcuttur. M, $\mathbf{x}_1,...,\mathbf{x}_m,\mathbf{y}_1,...,\mathbf{y}_n$'nin lineer gereni olsun. Teorem 1.5.2'ye göre $\mathbf{u}\in M$, $\mathbf{v}\in M^{\perp}$ olacak şekilde bir **z=u+v** ayrışımı mevcuttur. Bu takdirde her bir i için, **v**, $\mathbf{x}_i$'e ortogonaldır ve bu nedenle $\mathbf{v}\in W$ dir. Aynı zamanda her bir i için **v**, $\mathbf{y}_i$'ye ortogonaldır ve bu nedenle $\langle\mathbf{v},\mathbf{v}\rangle$=0 ve buradan da **v=0** dır ve **z=u** olduğu görülür. Bu, **z**'nin $\mathbf{x}_1,...,\mathbf{x}_m,\mathbf{y}_1,...,\mathbf{y}_n$'den lineer bağımsız olduğu gerçeğiyle çelişir. Bu nedenle $m \cdot n \;\; p$ dir.

Teorem 1.5.4. Eğer $S_1\subset S_2\subset T$ vektör uzayları ise bu takdirde, (i) $(S_2)^{\perp}\subset(S_1)^{\perp}$; (ii) $(S_1^{\perp})^{\perp}=S_1$ dir.

Tanım 1.5.4. **A** bir mxn matris olsun **Ax=0** olacak şekilde tüm $\mathbf{x}\in R^n$ vektörlerinin kümesinin R^n'in bir altuzayı olacağı kolayca görülür. Bu altuzaya **A**'nın *sıfır uzayı* denir ve N(**A**) ile gösterilir.

Teorem 1.5.5. **A** bir mxn matris olsun. Bu takdirde N(**A**)= $C(\mathbf{A}')^{\perp}$ dir.

İspat: Eğer $\mathbf{x}\in N(\mathbf{A})$ise, bu takdirde, **Ax=0** dir ve bu nedenle her $\mathbf{y}\in R^m$ için, $\mathbf{y}'\mathbf{Ax}=\mathbf{0}$ dır. Bundan dolayı, **x**, $C(\mathbf{A}')$'deki herhangi bir vektöre ortogonaldır.

Tersine olarak eğer, $\mathbf{x} \in C(\mathbf{A}')^{\perp}$ ise, bu takdirde, **x**, $\mathbf{A}'$'nün her sütununa ortogonaldir ve bu nedenle **Ax=0** dır.

Teorem 1.5.6. **A**, r ranklı bir mxn matris olsun. Bu takdirde boy(N(**A**))= n-r dir.

İspat: boy$N$(**A**)=boy$C(\mathbf{A}')^{\perp}$ (Teorem 1.5.5'e göre)

$= n - \text{boy}\,C(\mathbf{A}')$ (Teorem 1.5.3'e göre)

$= n - r$

elde ederiz. Bu da ispatı tamamlar.

Tanım 1.5.5. **A**'nın sıfır uzayının boyutuna **A**'nın sıfırlığı denir. Bu nedenle Teorem 1.5.6, rank+sıfırlık=sütunların sayısı olduğunu sağlar.

1.6 Tekil Olmama

Tanım 1.6.1. $x_1, x_2 ..... x_n$ bilinmeyenli m tane lineer denkleme sahip olduğumuzu varsayalım. **A** bir mxn matris olmak üzere, denklemler tek bir **Ax=b** matris denklemiyle uygun bir şekilde ifade edilebilir. **Ax=b** denklemine tutarlı denecektir eğer bu denklem en az bir çözüme sahipse, aksi takdirde denklem tutarsızdır. Eğer **b=0** ise denklem homojendir. **Ax=0** homojen denkleminin çözümlerinin kümesi aşikâr olarak **A**'nın sıfır uzayıdır.

Eğer **Ax=b** denklemi tutarlı ise, bu takdirde, $\mathbf{a}_1 ..... \mathbf{a}_n$ **A**'nın sütunları olmak üzere herhangi $x_1^0, x_2^0, ... x_n^0$ için

$$\mathbf{b}= x_1^0\mathbf{a}_1 + x_2^0\mathbf{a}_2 +, ... + x_n^0\mathbf{a}_n$$

yazabiliriz. Bu nedenle $\mathbf{b} \in C(\mathbf{A})$ dır. Tersine olarak eğer $\mathbf{b} \in C(\mathbf{A})$ ise, bu takdirde, **Ax=b** tutarlı olmalıdır. Eğer denklem tutarlı ise ve $\mathbf{x}^0$ denklemin bir çözümü ise bu takdirde denklemin tüm çözümlerinin kümesi,

$$\{\mathbf{x}^0 + \mathbf{x} : \mathbf{x} \in N(\mathbf{A})\}$$

ile verilir. Açık olarak, **Ax=b** denklemi ya çözüme sahip değildir ya bir tek çözüme sahiptir ya da sonsuz birçok çözüme sahiptir. Eğer R(**A**)= n ise, nxn mertebeli bir **A** matrisine tekil değildir denir; aksi halde, **A** matrisi tekildir.

Teorem 1.6.1. **A** bir nxn matris olsun. Aşağıdaki şartlar eşdeğerdirler.

(i) **A** tekil değildir yani R(A)= n dir.

(ii) Herhangi bir $\mathbf{b} \in R^n$ için **Ax=b** bir tek çözüme sahiptir.

(iii) **AB=BA=I** olacak şekilde bir tek **B** matrisi vardır.

İspat: (i)$\Rightarrow$(ii): R(A)= n olduğundan C(**A**)=R^n yazarız ve bu nedenle **Ax=b** bir çözüme sahiptir. Eğer **Ax=b** ve **Ay=b** ise bu takdirde **A(x-y)=0** dır**.** Teorem 1.5.6'ya göre boy(N(**A**))=0 ve böylece **x=y** dir. Bu, tekliği gösterir.

(ii)$\Rightarrow$(iii): (ii) şıkkına göre, $\mathbf{e_i}$, birim matrisin i-yinci sütunu olmak üzere, $\mathbf{Ax=e_i}$, $\mathbf{b_i}$ diyeceğimiz, bir tek çözüme sahiptir. Bu takdirde $\mathbf{B=[b_1,b_2,...,b_n]}$, **AB=I** eşitliğini sağlayan yegane matristir. Aynı muhakemeyi $\mathbf{A}'$'ne uygulayarak, **CA=I** olacak şekilde bir tek **C** matrisinin varlığı sonucunu çıkarırız. Şimdi **B=(CA)B=C(AB)=C** dir.

(iii)$\Rightarrow$(i)**:** (iii) şıkkının doğru olduğunu varsayalım. Bu takdirde herhangi bir $\mathbf{x} \in R^n$, **x=A(Bx)** olarak ifade edebilir ve buradan da C(**A**)=R^n dir. Bu nedenle tanımdan, C(**A**)'nın boyutu olan R(**A**) n olmalıdır.

Teorem 1.6.1'in (iii) şıkkının **B** matrisine **A**'nın tersi denir ve $\mathbf{A^{-1}}$ ile gösterilir. Eğer **A**, **B** nxn matrisler ise, bu takdirde, $\mathbf{(AB)(B^{-1}A^{-1}) = I}$ dır ve bu nedenle $\mathbf{(AB)^{-1} = B^{-1}A^{-1}}$ dır. Özellikle, iki tekil olmayan matrisin çarpımı tekil değildir.

Teorem 1.6.2. **A** bir nxn matris olsun. **A**'nın i-yinci satır ve j-yinci sütununu silmek suretiyle **A**'dan elde edilen alt matrisi $\mathbf{A_{ij}}$ ile göstereceğiz. a_{ij}'nin kofaktörü (eşçarpan) $(-1)^{i+j}\left|\mathbf{A_{ij}}\right|$ olarak tanımlanır. ek(**A**) ile gösterilen, **A**'nın eki, (i,j)-yinci elemanı a_{ji}'nin eş çarpanı olan nxn matristir.

Determinantların teorisinden,

$$\sum_{j=1}^{n} a_{ij}(-1)^{i+j}\left|\mathbf{A_{ij}}\right| = \left|\mathbf{A}\right|$$

ve $i \neq k$ için,

$$\sum_{j=1}^{n} a_{ij}(-1)^{j+k}\left|\mathbf{A}_{kj}\right| = 0$$

denklemlerini elde ederiz. Bu denklemler Aek(A)=$|\mathbf{A}|\,\mathbf{I}$ olarak yorumlanabilir. Bu nedenle eğer $|\mathbf{A}| \neq 0$ ise, bu takdirde, $\mathbf{A}^{-1}$ mevcuttur ve,

$$\mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|} ek(\mathbf{A})$$

dır. Tersine olarak eğer **A** tekil değil ise, bu takdirde, $\mathbf{AA}^{-1} = \mathbf{I}$ eşitliğinden $\left|\mathbf{AA}^{-1}\right| = |\mathbf{A}|\left|\mathbf{A}^{-1}\right| = \mathbf{I}$ olduğu sonucunu çıkarırız ve bu nedenle $|\mathbf{A}| \neq 0$ dır. Böylece aşağıdaki sonucu ispatladık.

Bir kare matrisin tekil olmaması için gerek ve yeter şart onun determinantının sıfırdan farklı olmasıdır.

Bir matrisin bir rxr minörü **A**'nın bir rxr alt matrisinin determinantı olarak tanımlanır. **A** r ranklı bir mxn matris olsun, $s > r$ olsun ve **A**'nın, $i_1, i_2, \ldots, i_s$ satırları ve $j_1, j_2, \ldots, j_s$ sütunları tarafından oluşturulan, **A**, bir sxs minörünü düşünelim. $j_1, j_2, \ldots, j_s$ sütunlarının lineer bağımlı olması gerektiğinden yukarıdaki ifade edilen sonuca göre minör sıfır olmalıdır. Tersine olarak eğer **A**, r ranklı ise, bu takdirde, **A,** $i_1, i_2, \ldots, i_r$ satırları diyeceğimiz, r tane lineer bağımsız satıra sahiptir. **B**, bu r tane satır tarafından oluşturulan alt matris olsun. Bu takdirde **B**, r rankına sahiptir ve böylece **B,** r sütun rankına sahiptir. Bu nedenle **B**'nin rxr boyutlu bir **C** alt matrisi vardır ve bu nedenle **A**'nın rankı da r dir. Yine yukarıda ifade edilen sonuca göre **C**, sıfırdan farklı bir determinanta sahiptir. Böylece rankın minörlere dayanan aşağıdaki tanımını elde ederiz:

(i) Sıfırdan farklı bir rxr minör varsa,

(ii) Her sxs minör ($s > r$) sıfır ise,

A matrisinin rankı r dir. Daha önce hatırlatıldığı gibi, rankın sıfır olması için gerek ve yeter şart **A**'nın sıfır matrisi olmasıdır.

1.7. Frobenius Eşitsizliği

Teorem 1.7.1. **B** r ranklı bir mxr matris olsun. Bu takdirde **XB=I** olacak şekilde **B**'nin sol tersi denilen bir **X** matrisi vardır.

İspat: Eğer $m=r$ ise, bu takdirde, **B** tekil değildir ve bir ters kabul eder. Bu nedenle $r<m$ olduğunu farz edelim. **B**'nin sütunları lineer bağımsızdır. Bu nedenle **B**'nin sütunlarıyla birlikte R^m için bir taban oluşturan m-r tane sütunun bir kümesini bulabiliriz. Diğer bir değişle $[\mathbf{B},\mathbf{U}]$ tekil olmayacak şekilde mx$(m$-$r)$ mertebeli bir **U** matrisi bulabiliriz. **X**, rxm mertebeli bir matris olmak üzere, $[\mathbf{B},\mathbf{U}]$'nin tersi $\begin{bmatrix}\mathbf{X}\\ \mathbf{V}\end{bmatrix}$ olarak parçalanmış olsun,

$$\begin{bmatrix}\mathbf{X}\\ \mathbf{V}\end{bmatrix}[\mathbf{B},\mathbf{U}]=\mathbf{I}$$

olduğundan, **XB=I** elde ederiz.

Aynı şekilde r ranklı bir **C** rxn matrisinin bir sağ terse sahip olduğunu gösteririz. Yani **CY=I** olacak şekilde bir **Y** matrisi buluruz. Matris kare ve tekil olmadıkça sol ve sağ tersin tek olmadığına dikkat edelim.

Teorem 1.7.2. **B**, r ranklı bir m x r matris olsun. Bu takdirde

$$\mathbf{PB}=\begin{bmatrix}\mathbf{I}\\ \mathbf{0}\end{bmatrix}$$

olacak şekilde tekil olmayan bir **P** matrisi mevcuttur.

İspat: İspat Teorem 1.7.1'in aynısıdır. Eğer $\mathbf{P}=\begin{bmatrix}\mathbf{X}\\ \mathbf{V}\end{bmatrix}$ alırsak, bu takdirde **P** istenilen şartı sağlar. Aynı şekilde eğer **C**, r ranklı rxn matris ise bu takdirde $\mathbf{CQ}=[\mathbf{I},\mathbf{0}]$ olacak şekilde tekil olmayan bir **Q** matrisi mevcuttur. Bu iki sonuç ve rank ayrışımı derhal bizi aşağıdaki sonuca götürür.

Teorem 1.7.3. **A**, r ranklı bir mxn matris olsun. Bu takdirde

$$\mathbf{PAQ}=\begin{bmatrix}\mathbf{I}_r & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{bmatrix}$$

olacak şekilde tekil olmayan **P, Q** matrisleri mevcuttur.

Teorem 1.7.4. **A, B** sırasıyla mxn ve nxp mertebeli matrisler olsun. Eğer R(**A**)=n ise bu takdirde R(**AB**)=R(**B**) dir. Eğer R(**B**)=n ise bu takdirde R(**AB**)=R(**A**) dır.

İspat: İlk olarak R(**A**)=n olduğunu farz edelim. Teorem 1.7. 1'e göre **XA=I** olacak şekilde bir **X** matrisi mevcuttur. Bu takdirde

$$R(\mathbf{B})=R(\mathbf{XAB})\leq R(\mathbf{AB})\leq R(\mathbf{B})$$

dır ve bu nedenle R(**AB**)= R(**B**) dir. İkinci kısım benzer şekilde görülür. Teorem 1.7.4'ün bir sonucu olarak rankın tekil olmayan bir matrisle çarpımdan etkilenmediğini görürüz.

Teorem 1.7.5. **A**, r ranklı bir mxn matris olsun. Bu takdirde **A+Z**, n rankına sahip olacak şekilde n-r ranklı bir **Z** mxn matrisi vardır.

İspat: Teorem 1.7. 3'e göre,

$$\mathbf{PAQ}=\begin{bmatrix}\mathbf{I}_r & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{bmatrix}$$

olacak şekilde tekil olmayan **P, Q** matrisleri vardır. **W**, n-r ranklı herhangi bir matris olmak üzere

$$\mathbf{Z}=\mathbf{P}^{-1}\begin{bmatrix}\mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{W}\end{bmatrix}\mathbf{Q}^{-1}$$

alalım. Bu takdirde **P(A+Z)Q**'nun rankının n olduğu kolayca gerçeklenir. **P, Q** tekil olmadıklarından, uyarmayla hemen görülür ki, **(A+Z)**'nin rankı n dir. (Teorem 1.7.5'in rank ayrışımını kullanarak ispatlanabildiğini de belirtelim.)

Teorem 1.7.6. (Frobenius Eşitsizliği) **A** ve **B** sırasıyla mxn ve nxp mertebeli matrisler olsun. Bu takdirde

$$R(\mathbf{AB})\geq R(\mathbf{A})+R(\mathbf{B})-n$$

dir.

İspat: Teorem 1.7.5'e göre **A+Z**, n ranklı olacak şekilde n- R(**A**) ranklı bir **Z** matrisi vardır.

$$R(\mathbf{B})=R((\mathbf{A}+\mathbf{Z})\mathbf{B}) \qquad \text{(Teorem 1.7.4'e göre)}$$

$$=R(\mathbf{AB}+\mathbf{ZB})$$

$\leq$ R(**AB**)+R(**ZB**) (Teorem 1.4.4'e göre)

$\leq$ R(**AB**)+R(**Z**)

= R(**AB**)+n-R(**A**)

dır ve bu nedenle R(**AB**)$\geq$ R(**A**)+ R(**B**)-n olur.

1.8. Özdeğerler ve Spektral (Görünge) Teoremi

Tanım 1.8.1. **A** bir nxn matris olsun. $|\mathbf{A}-\lambda\mathbf{I}|$ determinantı λ (kompleks de olabilir) değişkenine göre n-inci dereceden bir polinomdur. Bu polinoma **A**'nın karakteristik polinomu denir.

$$|\mathbf{A}-\lambda\mathbf{I}|=0$$

denklemine **A**'nın karakteristik denklemi denir. Cebirin temel teoremine göre bu denklem n tane köke sahiptir ve bu köklere **A**'nın özdeğerleri denir.(Cebirin temel teoremi n-inci dereceden bir $P(\lambda)=0$ denklemi reel ya da kompleks n tane köke sahiptir. Köklerden birisi kompleks ise eşleniği de köktür.)

Özdeğerlerin hepsi farklı olmayabilir. Karakteristik denklemin bir kökü olarak ortaya çıkan bir özdeğerin tekrar sayısına özdeğerin cebirsel katlılığı denir. Karakteristik polinomu,

$$|\mathbf{A}-\lambda\mathbf{I}|=(\lambda_1-\lambda)...(\lambda_n-\lambda) \qquad (1.8.1)$$

gibi çarpanlara ayırırız. (1.8.1) bağıntısında λ=0 koyarak $|\mathbf{A}|$'nın **A**'nın özdeğerlerinin bir çarpımı olduğunu hemen görürüz. Aynı şekilde (1.8.1) bağıntısının her iki tarafında λ^{n-1}'in katsayısını eşitleyerek **A**'nın izinin özdeğerlerinin toplamına eşit olduğunu görürüz.

Bir kare matrisin bir esas alt matrisi, satırların bir kümesi ve sütunların karşılık gelen kümesiyle oluşturulan bir alt matristir. **A**'nın bir esas minörü bir esas alt matrisin determinantıdır.

Tanım 1.8.2. Eğer **A**= $\mathbf{A}'$ ise, **A** kare matrisine simetrik denir. Bir **A** nxn matrisine pozitif-tanımlı denecekti, eğer **A** simetrik ve sıfırdan farklı herhangi bir **x** vektörü için $\mathbf{x'Ax} > 0$ ise. Birim matrisin pozitif-tanımlı olduğu açıktır ve bu nedenle köşegen elemanları pozitif olan bir köşegen matris de pozitif-tanımlıdır.

Teorem 1.8.1. Eğer **A** pozitif-tanımlı ise, bu takdirde, **A** tekil değildir.

İspat: Eğer **Ax=0** ise, bu takdirde, $\mathbf{x'Ax} = 0$ dır ve **A** pozitif-tanımlı olduğundan **x=0** dır. Bu nedenle **A** tekil olmamalıdır. Bir sonraki sonuç tanımdan açıktır.

Teorem 1.8.2. Eğer **A, B** pozitif-tanımlı ve $\alpha+\beta>0$ olmak üzere, $\alpha\geq 0$, $\beta\geq 0$ ise, bu takdirde, $\alpha\mathbf{A}+\beta\mathbf{B}$ pozitif-tanımlıdır.

Teorem1.8.3. Eğer **A** pozitif-tanımlı ise, bu takdirde, $|\mathbf{A}|>0$ dır.

İspat: $0\leq\alpha\leq 1$ için,

$$f(\alpha)=|\alpha\mathbf{A}+(1-\alpha)\mathbf{I}|$$

fonksiyonunu tanımlayalım Teorem 1.8.2'ye göre $\alpha\mathbf{A}+(1-\alpha)\mathbf{I}$ pozitif-tanımlıdır ve bu nedenle Teorem 1.8.1'e göre $f(\alpha)\neq 0$, $0\leq\alpha\leq 1$ dir. Aşikar olarak $f(0)=1$ dir. f sürekli olduğundan $f(1)=|\mathbf{A}|>0$ dır.

Teorem 1.8.4. **A** pozitif-tanımlı ise, bu takdirde, **A**'nın herhangi bir esas alt matrisi pozitif-tanımlıdır.

İspat: **A** pozitif-tanımlı olduğundan, her $\mathbf{x}\neq\mathbf{0}$ için $\mathbf{x'Ax}>0$ dır. Bu şartı $j_1,\ldots,j_s$ koordinatlarında sıfırlar olan vektörler kümesine uygulayalım. Böyle bir **x** vektörü için $\mathbf{x'Ax}$, $\mathbf{y'By}$ tipinden bir ifadeye indirgenir. Burada **B**, **A**'nın $j_1,\ldots,j_s$ satır ve sütunlarını silmek suretiyle oluşturulan esas alt matristir. **B**'nin ve aynı şekilde **A**'nın herhangi bir esas alt matrisinin pozitif-tanımlı olduğu görülür.

Tanım 1.8.3. Eğer her $\mathbf{x}\in R^n$ için $\mathbf{x'Ax}\geq 0$ ise, simetrik bir **A** nxn matrisine pozitif yarı-tanımlı denir.

Teorem 1.8.5. Eğer **A** bir simetrik matris ise, bu takdirde, **A**'nın özdeğerlerinin tümü reeldir.

İspat: μ'nün **A**'nın bir özdeğeri olduğunu farz edelim ve α, β reel sayılar v $i=\sqrt{-1}$ olmak üzere, $\mu=\alpha+i\beta$ olsun. $|\mathbf{A}-\mu\mathbf{I}|=0$ olduğundan,

$$|(\mathbf{A}-\mu\mathbf{I})-i\beta\mathbf{I}|=0$$

elde ederiz. Yukarıdaki determinantın eşleniğini alarak ve ikisini çarparak,

$$|(\mathbf{A}-\mu\mathbf{I})-i\beta\mathbf{I}||(\mathbf{A}-\mu\mathbf{I})+i\beta\mathbf{I}|=0$$

elde ederiz. Bu nedenle,

$$|(\mathbf{A}-\mu\mathbf{I})^2+\beta^2\mathbf{I}|=0 \qquad (1.8.2)$$

dir. **A** simetrik olduğundan, **A**'nın karesinin pozitif yarı-tanımlı olduğu gerçektir (tanımdan herhangi bir **B** matrisi için $\mathbf{BB}'$nün pozitif yarı-tanımlı olduğu görülür). Bu nedenle eğer $\beta \neq 0$ ise, bu takdirde, $|(\mathbf{A}-\mu\mathbf{I})^2+\beta^2\mathbf{I}|$ pozitif-tanımlıdır ve bu takdirde Teorem 1.8.1.'e göre (1.8.2) bağıntısı sağlanamaz. Bu nedenle $\beta=0$ ve μ reel olmalıdır. Eğer **A** bir simetrik nxn matris ise, **A**'nın özdeğerlerini ara sıra $\lambda_1(\mathbf{A})\geq ...\geq \lambda_n(\mathbf{A})$ ile ve eğer karıştırma olasılığı yoksa $\lambda_1 \geq ... \geq \lambda_n$ ile göstereceğiz.

Tanım 1.8.4. **A** bir simetrik nxn matris olsun. Bu takdirde herhangi bir i için $|\mathbf{A}-\lambda_i\mathbf{I}|=0$ dır ve bu nedenle $\mathbf{A}-\lambda_i\mathbf{I}$ tekildir. Bu nedenle $\mathbf{A}-\lambda_i\mathbf{I}$'nin sıfır uzayı en az 1 boyutuna sahiptir. Bu sıfır uzayına **A**'nın λ_i'ye karşılık gelen özuzayı denir ve özuzayda sıfır olmayan herhangi bir vektöre **A**'nın λ_i'ye karşılık gelen bir özvektörü denir. Sıfır uzayının boyutuna λ_i'nin geometrik katlılığı denir.

Teorem 1.8.6. **A** bir simetrik nxn matris olsun ve **x, y** sırasıyla λ ve μ'ye karşılık gelen özvektörler olmak üzere $\lambda \neq \mu$, **A**'nın özdeğerleri olsun. Bu takdirde $\mathbf{x}'\mathbf{y}=0$ dır.

İspat: **Ax=** λ**x** ve **Ay=** μ**y** elde ederiz. Bu nedenle, $\mathbf{y}'\mathbf{Ax}=\mathbf{y}'(\mathbf{Ax})=\lambda\mathbf{y}'\mathbf{x}$ dir. Aynı zamanda $\mathbf{y}'\mathbf{Ax}=(\mathbf{y}'\mathbf{A})\mathbf{x}=\lambda\mathbf{y}'\mathbf{x}$ dir. Böylece, $\lambda\mathbf{y}'\mathbf{x}=\mu\mathbf{y}'\mathbf{x}$ dir. $\lambda\neq\mu$ olduğundan, $\mathbf{x}'\mathbf{y}=0$ olduğu görülür.

Tanım 1.8.5. Eğer $\mathbf{P}^{1}=\mathbf{P}'$ ya da $\mathbf{PP}'=\mathbf{P}'\mathbf{P}=\mathbf{I}$ ise, bir **P** kare matrisine ortogonal matris denecektir, eğer bir nxn kare matrisin satırları (ya da sütunları) R^n için bir ortonormal taban oluşturursa. Birim matrisin ortogonal olduğu açıktır. Birim matristen, onun satırlarının (veya sütunlarının) permütasyonu ile elde edilen bir matrise bir permütasyon matris denir ve bu matris ortogonaldır. Ortogonal matrislerin çarpımının ortogonal olacağı kolayca görülür.

Teorem 1.8.7. Spektral Teoremi: **A** simetrik bir nxn matris olsun. Bu takdirde

$$\mathbf{P'AP} = k\ddot{o}\c{s}eg(\lambda_1,...,\lambda_n) \tag{1.8.3}$$

olacak şekilde ortogonal bir **P** matrisi mevcuttur. Burada $\lambda_1,...,\lambda_n$'ler **A**'nın özdeğerleridir.

İspat: $n=1$ için sonuç açıktır. $n-1$ mertebeli matrisler için sonucu kabul edelim ve tümevarımla ilerleyelim. **x**, $\|\mathbf{x}\|=1$ olmak üzere, λ_1'e karşılık gelen bir özvektör olsun. **Q**, birinci sütunu **x** olan bir ortogonal matris olsun (böyle bir **Q** matrisi mevcuttur. İlk olarak **x**'i R^n için bir tabana genişletiriz ve sonra Gram-Schmidt yöntemini uygularız). Bu takdirde

$$\mathbf{Q'AQ} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & \mathbf{B} & \\ 0 & & & \end{bmatrix}$$

dır. $\mathbf{Q'AQ}$'nün özdeğerleri aynı zamanda $\lambda_1,...,\lambda_n$ dir. Bu nedenle **B**'nin özdeğerleri de $\lambda_2,...,\lambda_n$ dir. $\mathbf{Q'AQ}$ simetrik olduğundan **B**'nin de simetrik olduğu açıktır. İndüksiyon varsayımına göre,

$$\mathbf{R'BR} = k\ddot{o}\c{s}eg(\lambda_2,...,\lambda_n)$$

olacak şekilde bir ortogonal **R** matrisi mevcuttur. Şimdi,

$$\mathbf{P} = \mathbf{Q}\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & \mathbf{R} & \\ 0 & & & \end{bmatrix}$$

alalım. Bu takdirde $\mathbf{P'AP} = k\ddot{o}\c{s}g(\lambda_1,...,\lambda_n)$ dir.

Teorem 1.8.7'deki **P** matrisinin $\mathbf{x_1,...,x_n}$ sütunlarına sahip olduğunu farz edelim. Bu takdirde

$$\mathbf{AP} = \mathbf{P}k\ddot{o}\c{s}g(\lambda_1,\dots,\lambda_n)$$

olduğundan $\mathbf{Ax}_i = \lambda_i \mathbf{x}_i$ elde ederiz. Bir başka ifadeyle $\mathbf{x}_i$, **A'**nın λ_i'ye karşılık gelen bir özvektörüdür. (1.8.3) bağıntısını yazmanın başka bir biçimi,

$$\mathbf{A} = \lambda_1 \mathbf{x}_1 \mathbf{x}_1' + \dots + \lambda_n \mathbf{x}_n \mathbf{x}_n'$$

dir. Bu **A**'nın spektral ayrışımı olarak bilinir.

Teorem 1.8.8. **A** bir simetrik nxn matris oldun. Bu takdirde **A**'nın pozitif-tanımlı olması için gerek ve yeter şart **A**'nın özdeğerlerinin tümünün pozitif olmasıdır.

İspat: Spektral teoremine göre bir ortogonal **P** matrisi için $\mathbf{P}'\mathbf{AP} = k\ddot{o}\c{s}eg(\lambda_1,\dots,\lambda_n)$ dir. Sonuç, **A**'nın pozitif-tanımlı olması için gerek ve yeter şart $\mathbf{P}'\mathbf{AP}$'nin de pozitif-tanımlı olması gerçeğinden görülür. Aynı şekilde bir simetrik matrisin pozitif yarı-tanımlı olması için gerek ve yeter şart onun özdeğerlerinin her birinin negatif olmamasıdır.

Teorem 1.8.9. Eğer **A** pozitif yarı-tanımlı ise, bu takdirde, $\mathbf{B}^2 = \mathbf{A}$ olacak şekilde bir tek pozitif yarı-tanımlı **B** matrisi vardır. **B** matrisine **A**'nın karekökü denir ve $\mathbf{A}^{1/2}$ ile gösterilir.

İspat: (1.8.3) bağıntısını gerçekleyecek şekilde bir ortogonal **P** matrisi vardır. **A** pozitif yarı-tanımlı olduğundan $\lambda_i \geq 0,\ i = 1,\dots,n$ dir.

$$\mathbf{B} = \mathbf{P}k\ddot{o}\c{s}eg(\lambda_1^{\frac{1}{2}},\dots,\lambda_n^{\frac{1}{2}})\mathbf{P}'$$

alalım. Bu takdirde $\mathbf{B}^2 = \mathbf{A}$ dır.

Tekliği ispatlamak için, eğer **B** ve **C**, $\mathbf{A} = \mathbf{B}^2 = \mathbf{C}^2$ bağıntısını sağlayan pozitif yarı-tanımlı matrisler ise, bu takdirde, **B=C** olduğunu göstermeliyiz. $\mathbf{D} = \mathbf{B} - \mathbf{C}$ olsun. Spektral teoremine göre $\mathbf{Z} = \mathbf{QDQ}'$ bir köşegen matris olacak şekilde bir **Q** ortogonal matrisi mevcuttur. $\mathbf{E} = \mathbf{QBQ}'$, $\mathbf{F} = \mathbf{QCQ}'$ olsun. **E=F** olduğunu göstermek yeterli olacaktır. **Z=E-F** bir köşegen matris olduğundan, $e_{ij} = f_{ij},\ i \neq j$ dir. Aynı zamanda,

$$\mathbf{EZ} + \mathbf{ZF} = \mathbf{E}(\mathbf{E} - \mathbf{F}) + (\mathbf{E} - \mathbf{F})\mathbf{F} = \mathbf{E}^2 - \mathbf{F}^2 = \mathbf{Q}(\mathbf{B}^2 - \mathbf{C}^2)\mathbf{Q}' = \mathbf{0}$$

ve bu nedenle,

$$(e_{ii}+f_{ii})z_{ii}=0, \qquad i=1,...,n$$

dir. Eğer $z_{ii}=0$ ise, bu takdirde, $e_{ii}=f_{ii}$ dir. Eğer $z_{ii}\neq 0$ ise, bu takdirde, $e_{ii}+f_{ii}=0$ dır. Bununla beraber **E, F** pozitif yarı-tanımlı olduklarından, $e_{ii}\geq 0, f_{ii}\geq 0$ dir ve $e_{ii}=f_{ii}=0$ olduğu görülür. Bu nedenle $e_{ii}=f_{ii}$, $i=1,...,n$ dir ve ispat tamamdır.

Tanım 1.8.6. Eğer $\mathbf{A}^2=\mathbf{A}$ ise, **A** matrisine idempotent denir.

Teorem 1.8.10. Eğer **A** idempotent ise, bu takdirde, **A**'nın her bir özdeğeri ya 0 ya da 1 dir.

İspat: $\lambda_1,...,\lambda_n$ **A**'nın özdeğerleri olsun. Bu takdirde $\lambda_1^2,...,\lambda_n^2$, $\mathbf{A}^2$'nin özdeğerleridir. $\mathbf{A}^2=\mathbf{A}$ olduğundan, $\{\lambda_1^2,...,\lambda_n^2\}=\{\lambda_1,...,\lambda_n\}$ dir ve her bir i için λ_i =0 ya da 1 olduğu görülür.

Tersine olarak, eğer **A** simetrik ve **A**'nın her bir özdeğeri 0 ya da 1 ise, bu takdirde, **A** idempotenttir. Bu spektral teoreminin bir uygulamasıyla görülür. Bir matrise tam satır(veya sütun) ranklıdır diyeceğiz, eğer onun rankı, satırlarının(veya sütunlarının) sayısına eşitse.

Teorem 1.8.11. Eğer **A** idempotent ise, bu takdirde, **R(A)=izA** dır.

İspat: **A=BC** bir rank ayrışımı olsun. **B** tam sütun ranklı olduğundan Teorem 1.7.1'e göre **B** bir sol terse sahiptir. Aynı şekilde **C** de bir sağ terse sahiptir. B_l, C_r sırasıyla B, C'nin sol ve sağ tersi olsun. Bu takdirde $\mathbf{A}^2=\mathbf{A}$ bağıntısı,

$$\mathbf{B_l^- BCBCC_r^-}=\mathbf{B_l^- BCC_r^-}=\mathbf{I}$$

Olduğunu gösterir ve bu nedenle **CB=I** dır. Burada **I** birim matrisinin R(**A**) dır. Böylece ,

$$\text{iz}\mathbf{A}=\text{iz}\mathbf{BC}=\text{iz}\mathbf{CB}=\text{R}(\mathbf{A})$$

dır.

1.9. Genelleştirilmiş Tersler

Tanım 1.9.1 **A** bir mxn matris olsun. Eğer **AGA=A** ise nxm mertebeli bir **G** matrisine **A**'nın genelleştirilmiş tersi (veya bir g-tersi) denir.

Eğer **A** kare ve tekil değilse bu takdirde $\mathbf{A}^{-1}$ **A**'nın yegane g-tersidir. Aksi takdirde kısaca göreceğimiz gibi **A** sonsuz birçok g-terse sahiptir.

Teorem 1.9.1. **A, G** sırasıyla mxn ve nxm mertebeli matrisler olsun. Bu takdirde aşağıdaki şartlar eşdeğerdirler

(i) **G, A**'nın bir g-tersidir.

(ii) Herhangi bir $\mathbf{y} \in C(\mathbf{A})$ için $\mathbf{x} = \mathbf{G}\mathbf{y}$, $\mathbf{A}\mathbf{x} = \mathbf{y}$ 'nin bir çözümüdür.

İspat:(i) $\Rightarrow$(ii): Herhangi bir $\mathbf{y} \in C(\mathbf{A})$, herhangi bir **z** için **y=Az** biçimindedir. Bu takdirde, **A(Gy)=AGAz=Az=y** dir.

(ii) $\Rightarrow$(i): Herhangi bir $\mathbf{y} \in C(\mathbf{A})$ için **AGy=y** olduğundan her **z** için **AGAz=Az** elde ederiz. Özellikle eğer **z**'yi birim matrisin i-inci sütunu olarak alırsak, bu takdirde **AGA** ve **A**'nın i-yinci sütunlarının aynı olduklarını görürüz. Bu nedenle **AGA=A** dır.

A=BC bir rank ayrışımı olsun. **B**'nin bir $\mathbf{B}_\ell^-$ sol tersine ve **C** nin $\mathbf{C}_\mathbf{r}^-$ sağ tersine sahip olduğunu görürüz. Bu takdirde,

$$\mathbf{AGA=BC(C_r^- B_\ell^-)BC=BC=A}$$

olduğundan, $\mathbf{G} = \mathbf{C}_\mathbf{r}^-\mathbf{B}_\ell^-$ **A'**nın bir g-tersidir. Bunun yerine, eğer **A** r ranklı ise, bu takdirde, Teorem 1.7.3'e göre,

$$\mathbf{A} = \mathbf{P}\begin{bmatrix} \mathbf{I}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\mathbf{Q}$$

olacak şekilde tekil olmayan **P, Q** matrisleri vardır. Uygun boyutlara sahip herhangi **U, V, W** için, gösterilebilir ki,

$$\begin{bmatrix} \mathbf{I}_r & \mathbf{U} \\ \mathbf{V} & \mathbf{W} \end{bmatrix}$$

matrisi,

$$\begin{bmatrix} \mathbf{I}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$$

matrisinin bir g-tersidir. Bu takdirde,

$$\mathbf{G} = \mathbf{Q}^{-1}\begin{bmatrix} \mathbf{I}_r & \mathbf{U} \\ \mathbf{V} & \mathbf{W} \end{bmatrix}\mathbf{P}^{-1}$$

A'nın bir g-tersidir. Bu aynı zamanda kare ve tekil olan herhangi bir matrisin sonsuz g-terse sahip olduğunu gösterir. Bir g-tersi hesaplamak için özellikle uygun olan başka bir yöntem aşağıdaki gibidir. **A**, r ranklı olsun. **A**'nın rxr boyutlu tekil olmayan herhangi bir alt matrisini seçelim. Uygunluk için,

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

olduğunu kabul edelim. Burada $\mathbf{A}_{11}$, rxr ve tekil değildir. **A**, r ranklı olduğundan, $\mathbf{A}_{12} = \mathbf{A}_{11}\mathbf{X}$, $\mathbf{A}_{22} = \mathbf{A}_{21}\mathbf{X}$ olacak şekilde bir **X** matrisi vardır. Şimdi,

$$\mathbf{G} = \begin{bmatrix} \mathbf{A}_{11}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$$

olarak tanımlanan **G** nxm matrisinin **A**'nın bir g-tersi olduğu gerçeklenebilir.(Bu gerekli çarpımı yapmak suretiyle gerçeklenir.) **A**'nın bir g-tersini göstermek için sık sık **A** notasyonunu kullanacağız.

Teorem 1.9.2. Eğer **G**, **A** nın bir g-tersi ise R(**A**)=R(**AG**)=R(**GA**) dır.

İspat: R(**A**)=R(**AGA**) $\leq$R(**AG**) $\leq$R(**A**) dır. İkinci kısım aynı şekilde görülür.

A'nın bir g-tersine bir reflexive(dönüşlü) bir g-ters denir, eğer o ve aynı zamanda **GAG=G** eşitliğini de sağlarsa. Eğer **G, A**'nın herhangi bir g-tersi ise, bu takdirde, **GAG**'nin **A**'nın bir reflexive g-tersi olduğuna dikkat edelim.

Teorem 1.9.3. G, **A**'nın bir g-tersi olsun. Bu takdirde R(**A**) $\leq$ R(**G**) dır. Ayrıca eşitliğin sağlanması için gerek ve yeter şart **G**'nin reflexive olmasıdır.

İspat: Herhangi bir **G** g-ters için, R(**A**)=R(**AGA**) $\leq$ R(**G**) elde ederiz. Eğer **G** reflexive ise, bu takdirde, R(**G**)=R(**GAG**) $\leq$ R(**A**) ve bu nedenle R(**A**)=R(**G**) dir.

Tersine olarak R(**A**)=R(**G**) olduğunu farz edelim. İlk olarak $\mathcal{C}(\mathbf{G}) \subset \mathcal{C}(\mathbf{GA})$ olduğuna dikkat edelim. Teorem 1.9.2'ye göre R(**A**)=R(**GA**) ve bu nedenle $\mathcal{C}(\mathbf{G}) = \mathcal{C}(\mathbf{GA})$ dır. Bundan dolayı herhangi **X** için **G=GAX** dir. Şimdi,

$$\mathbf{GAG=GAGAX=GAX=G}$$

dir ve **G** reflexivedir.

Teorem 1.9.4. **A** bir mxn matris olsun, **G, A**'nın bir g-tersi olsun ve $\mathbf{y} \in \mathcal{C}(\mathbf{A})$ olsun. Bu takdirde **Ax=y**'nin çözümler sınıfı $\mathbf{Gy} + (\mathbf{I} - \mathbf{GA})\mathbf{z}$ ile verilir. Burada **z** keyfidir.($\mathbf{z} \in R^n$)

İspat: $\mathbf{y} \in \mathcal{C}(\mathbf{A})$ olduğundan herhangi bir **z** için,

$$\mathbf{A}\{\mathbf{Gy} + (\mathbf{I} - \mathbf{GA})\mathbf{z}\} = \mathbf{AGy} = \mathbf{y}$$

dir ve bu nedenle $\mathbf{Gy} + (\mathbf{I} - \mathbf{GA})\mathbf{z}$ bir çözümdür. Tersine olarak, eğer **u** bir çözüm ise, bu takdirde, $\mathbf{z} = \mathbf{u} - \mathbf{Gy}$ koyarız.

$$\mathbf{Gy} + (\mathbf{I} - \mathbf{GA})\mathbf{z} = \mathbf{u}$$

Olduğunu gerçekleriz. Bu da ispatı tamamlar.

Eğer **AGA=A**'ya ek olarak, $(\mathbf{GA})' = \mathbf{GA}$ sağlanırsa **A**'nın bir **G** g-tersine **A**'nın minimum norm g-tersi denecektir. Bu terminolojinin sebebi aşağıdaki sonuçtan açık olacak.

Teorem 1.9.5. **A** bir mxn matris olsun. Bu takdirde aşağıdaki şartlar eşdeğerdirler.

(i) **G, A**'nın bir minimum norm g-tersidir.

(ii) Herhangi bir $\mathbf{y} \in \mathcal{C}(\mathbf{A})$ için **x=Gy, Ax=y**'nin minimum norma sahip bir çözümüdür.

İspat:(i) $\Rightarrow$(ii): Teorem 1.9.4 'e göre herhangi bir $\mathbf{y} \in \mathcal{C}(\mathbf{A})$ ve herhangi bir **z** için,

$$\|\mathbf{Gy}\| \le \|\mathbf{Gy} + (\mathbf{I} - \mathbf{GA})\mathbf{z}\| \qquad (1.9.1)$$

olduğunu göstermeliyiz.

$$\|\mathbf{Gy} + (\mathbf{I} - \mathbf{GA})\mathbf{z}\|^2 = \|\mathbf{Gy}\|^2 + \|(\mathbf{I} - \mathbf{GA})\mathbf{z}\|^2 + 2\mathbf{y}'\mathbf{G}'(\mathbf{I} - \mathbf{GA})\mathbf{z} \qquad (1.9.2)$$

yazarız. $\mathbf{y} \in \mathcal{C}(\mathbf{A})$ olduğundan herhangi **u** için **y=Au** dur. Bu nedenle,

$$\begin{aligned}\mathbf{y'G'(I-GA)z} &= \mathbf{u'A'G'(I-GA)z}\\ &= \mathbf{u'GA(I-GA)z}\end{aligned}$$

dir. Bunu (1.9.2) bağıntısında yerine koyarak (1.9.1) bağıntısını elde ederiz.

(ii) $\Rightarrow$(i): Herhangi bir $\mathbf{y} \in \mathcal{C}(\mathbf{A})$ için **x=Gy, Ax=y**'nin bir çözümü olduğundan Teorem 1.9.1'e göre **G**, **A**'nın bir g-tersidir. Şimdi her **z** için (1.9.1) bağıntısını elde ederiz ve bu nedenle her **u**, **z** için,

$$0 \leq \|\mathbf{(I-GA)z}\|^2 + \mathbf{2u'A'G'(I-GA)z} \qquad (1.9.3)$$

dir. (1.9.3) bağıntısında **u**'nun yerine α **u** koyalım. Eğer $\mathbf{u'A'G'(I-GA)z} < 0$ ise, bu takdirde, α'yı büyük ve pozitif seçerek (1.9.3) bağıntısına bir çelişme elde ederiz. Aynı şekilde $\mathbf{u'A'G'(I-GA)z} > 0$ ise, bu takdirde, α'yı büyük ve negatif seçerek bir çelişme elde ederiz. Böylece her **u**, **z** için,

$$\mathbf{u'A'G'(I-GA)z = 0}$$

ve bu nedenle $\mathbf{A'G'(I-GA)z = 0}$ olduğu sonucuna varırız. Bundan dolayı $\mathbf{A'G'}$, simetrik olan, $\mathbf{(GA)'GA}$'ya eşittir.

Tanım 1.9.2. **A**'nın bir **G** g-tersine **A**'nın bir en küçük kareler g-tersi denir. Eğer o **AGA=A**'ya ek olarak $\mathbf{(AG)' = AG}$ bağıntısını da sağlarsa.

Teorem 1.9.6. **A** bir *m*x*n* matris olsun. Aşağıdaki şartlar eşdeğerdirler.

(i) **G, A**'nın bir en küçük kareler g-tersidir.

(ii) Herhangi bir **x, y** için $\|\mathbf{AGy-y}\| \leq \|\mathbf{Ax-y}\|$ dir.

İspat:(i) $\Rightarrow$(ii): $\mathbf{x-Gy=w}$ olsun. Bu takdirde,

$$\|\mathbf{AGy-y}\| \leq \|\mathbf{AGy-y+Aw}\| \qquad (1.9.4)$$

olduğunu göstermeliyiz.

$$\|\mathbf{AGy-y+Aw}\|^2 = \|\mathbf{(AG-I)y}\|^2 + \|\mathbf{Aw}\|^2 + \mathbf{2w'A'(AG-I)y} \quad (1.9.5)$$

elde ederiz. Fakat $\mathbf{(AG)' = AG}$ olduğundan,

$$\mathbf{w'A'(AG - I)y = w'(A'G'A' - A')y = 0}$$

dır bunu (1.9.5) bağıntısında yerine koyarak (1.9.4) bağıntısını elde ederiz.

(ii) $\Rightarrow$(i): Herhangi bir **x** vektörü için, (ii)'de **y=Ax** koyalım. Bu takdirde,

$$\|\mathbf{AGAx - Ax}\| \leq \|\mathbf{Ax - Ax}\| = \mathbf{0}$$

ve buradan da **AGAx=Ax** olduğunu görürüz. **x** keyfi olduğundan, **AGA=A** dır ve bu nedenle **G, A**'nın bir g-tersidir. İspatın geriye kalan kısmı Teorem 1.9.5. de (i) $\Rightarrow$(ii) ispatına paraleldir.

Tutarlı olmayan **Ax=y** denklemine sahip olduğumuzu ve $\|\mathbf{Ax - y}\|$ minimum olacak şekilde bir **x** çözümü bulmayı istediğimizi varsayalım. Bu takdirde, teorem 1.9.6'ya göre bu çözüm, **A**'nın herhangi bir en küçük kareler g-tersi için **x=Gy** almak suretiyle elde edilir.

Tanım 1.9.3. Eğer **G, A**'nın bir reflexive g-tersi ise yani hem minimum norm hem de en küçük kareler g-tersi ise, bu takdirde, **G**'ye **A**'nın bir Moore-Penrose tersi denir. Başka bir deyişle eğer **G**,

$$\mathbf{AGA=A}, \qquad \mathbf{GAG=G}, \qquad \mathbf{(AG)' = AG}, \qquad \mathbf{(GA)' = GA} \quad (1.9.6)$$

bağıntısını sağlarsa **G**, **A**'nın bir Moore-Penrose tersidir. Böyle bir **G**'nin mevcut ve gerçekten tek olduğunu göstereceğiz. İlk olarak tekliği gösterelim. $\mathbf{G_1}$ ve $\mathbf{G_2}$'nin her ikisinin de (1.9.6) bağıntısını sağladığını varsayalım. Bu takdirde $\mathbf{G_1 = G_2}$ olduğunu göstermeliyiz. Aşağıdaki adımların her biri (1.9.6) bağıntısını uygulamak suretiyle görülür. Altları çizilen terimler her bir kerede bir sonraki adımı elde etmek için yeniden yorumlanacaktır.

$$
\begin{aligned}
\mathbf{G}_1 &= \mathbf{G}_1\underline{\mathbf{AG}_1} \\
&= \mathbf{G}_1\mathbf{G}_1'\underline{\mathbf{A}'} \\
&= \mathbf{G}_1\mathbf{G}_1'\mathbf{A}'\underline{\mathbf{G}_2'\mathbf{A}'} \\
&= \mathbf{G}_1\underline{\mathbf{G}_1'\mathbf{A}'}\mathbf{AG}_2 \\
&= \mathbf{G}_1\underline{\mathbf{AG}_1}\mathbf{AG}_2 \\
&= \mathbf{G}_1\underline{\mathbf{A}}\mathbf{G}_2 \\
&= \mathbf{G}_1\mathbf{A}\underline{\mathbf{G}_2\mathbf{A}}\mathbf{G}_2 \\
&= \underline{\mathbf{G}_1\mathbf{A}}\mathbf{A}'\mathbf{G}_2'\mathbf{G}_2 \\
&= \underline{\mathbf{A}'\mathbf{G}_1'\mathbf{A}'}\mathbf{G}_2'\mathbf{G}_2 \\
&= \underline{\mathbf{A}'\mathbf{G}_2'}\mathbf{G}_2 \\
&= \mathbf{G}_2\mathbf{AG}_2 \\
&= \mathbf{G}_2
\end{aligned}
$$

dir. **A**'nın Moore-Penrose tersini $\mathbf{A}^{+}$ ile göstereceğiz. Şimdi varlığı göstereceğiz. **A=BC**, **A**'nın bir rank ayrışımı olsun. Bu takdirde,

$$\mathbf{B}^{+} = (\mathbf{B}'\mathbf{B})^{-1}\mathbf{B}', \qquad \mathbf{C}^{+} = \mathbf{C}'(\mathbf{CC}')^{-1}$$

ve sonrada,

$$\mathbf{A}^{+} = \mathbf{C}^{+}\mathbf{B}^{+}$$

olacağı kolaylıkla gerçeklenebilir.

2. MATERYAL ve YÖNTEM

2.1. Lineer Modeller

Tanım 2.1.1. **y**, $y_1, y_2, ...y_n$ bileşenli bir sütun vektör olsun. Eğer her bir y_i bir rastgele değişken ise **y**'ye bir rastgele vektör diyeceğiz. **y**'nin beklenen değeri E(**y**) ile gösterilen i-yinci bileşeni $E(y_i)$ olan sütun vektördür.

x, y rastgele vektör ve **B**, **C** rastgele olmayan sabit matrisler olmak üzere

$$E(\mathbf{Bx}+\mathbf{Cy}) = \mathbf{B}E(\mathbf{x})+\mathbf{C}E(\mathbf{y})$$

olduğu açıktır.

Tanım 2.1.2. Eğer **x, y** sırasıyla m, n mertebeli rastgele vektörler ise, bu takdirde, **x, y** arasındaki kov(**x, y**) ile gösterilen kovaryans matrisi (i, j)-yinci elemanı kov(x_i, y_i) olan bir mxn matristir. **y**'nin, D(**y**) ile gösterilen, dağılım matrisi veya varyans-kovaryans matrisi kov(**y, y**) olarak tanımlanır. Dağılım matrisinin simetrik olduğu açıktır. Eğer **b**, **c** sabit vektörler ise, bu takdirde,

$$\text{kov}(\mathbf{b}'\mathbf{x}, \mathbf{c}'\mathbf{y}) = \text{kov}(b_1x_1 + ...+b_mx_m, c_1y_1 + ...+c_ny_n)$$

$$= \sum_{i=1}^{m}\sum_{j=1}^{n} b_ic_j \text{kov}(x_i, y_j)$$

$$= \mathbf{b}'\text{kov}(\mathbf{x}, \mathbf{y})\mathbf{c}$$

dir. Eğer **B, C** sabit matrisler ise bu taktirde

$$\text{kov}(\mathbf{Bx}, \mathbf{Cy}) = \mathbf{B}\text{kov}(\mathbf{x},\mathbf{y})\mathbf{C}'$$

olduğu görülür. **x=y** ve **b=c** alarak

$$\text{var}(\mathbf{b}'\mathbf{x}) = \mathbf{b}'D(\mathbf{x})\mathbf{b}$$

elde edilir. Varyans negatif olmadığından, D(**x**)'in pozitif yarı-tanımlı olduğu sonucunu çıkarırız. Kesinlikle sabit olan bir $\mathbf{b}'\mathbf{x}$ lineer kombinasyonu mevcut olmadıkça D(**x**)'in pozitif tanımlı olduğuna dikkat edelim. Şimdi bir lineer model kavramını ortaya koyacağız. $y_1, y_2, ...y_n$ rastgele değişkenlerinin ortaya koyan bir deneyi yürüttüğümüzü farz edelim. Rastgele değişkenlerin dağılımının bazı bilinmeyen parametrelerle kontrol edildiğini varsayalım. Bir lineer modelde, temel varsayım, $E(y_i)$'nin, bilinen katsayılar ile $\beta_1,\ldots,\beta_p$ parametrelerinin bir lineer fonksiyonu olduğudur. Matris notasyonu ile bu

$$E(\mathbf{y})=\mathbf{X}\boldsymbol{\beta}$$

olarak ifade edilebilir. Burada **y**, $y_1, y_2, \ldots, y_n$ bileşenli nx1 vektördür; **X**, nxp mertebeli rastgele olmayan matristir ve β, $\beta_1,\ldots,\beta_p$ parametrelerinin px1 vektörüdür. $y_1, y_2, \ldots, y_n$ lerin ilişkisiz olduklarını ve her i için var(y_i)=σ^2 olduğunu da kabul ederiz. Bu özelliğe homoscedasticity(eşit varyanslılık durumu) denir. Bu nedenle

$$D(\mathbf{y})=\sigma^2\mathbf{I}$$

dır. Modeli yazmak için başka bir yol

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

dır. Burada ε vektörü $E(\varepsilon)=\mathbf{0}$, $D(\varepsilon)=\sigma^2\mathbf{I}$ olan bir rastgele vektördür. Şimdilik **y**'nin dağılımı hakkında herhangi başka varsayımlar yapmayacağız. Birinci amacımız $\beta_1,\ldots,\beta_p$'nın ve onların lineer kombinasyonlarının tahminlerini bulmaktır. Aynı zamanda σ^2'nin de bir tahminin araştıracağız.

2.2. Tahmin Edilebilirlik

y nx1 vektör, **X**, nxp matris ve $\boldsymbol{\beta}$, px1 vektör olmak üzere

$$E(\mathbf{y})=\mathbf{X}\boldsymbol{\beta}, \qquad D(\mathbf{y})=\sigma^2\mathbf{I} \qquad (2.2.1)$$

lineer modellerini göz önüne alalım. Her $\boldsymbol{\beta}\in R^p$ için $E(\mathbf{c}'\mathbf{y})=\ell'\boldsymbol{\beta}$ olacak şekilde gözlemlerin bir $\mathbf{c}'\mathbf{y}$ lineer fonksiyonu varsa, $\ell'\boldsymbol{\beta}$ lineer parametrik fonksiyonuna tahmin edilebilir denecektir. $E(\mathbf{c}'\mathbf{y})=\ell'\boldsymbol{\beta}$ şartı $\mathbf{c}'\mathbf{X}\boldsymbol{\beta}=\ell'\boldsymbol{\beta}$ ya denktir ve bu R^p deki her $\boldsymbol{\beta}$ için sağlanması gerektiğinden $\mathbf{c}'\mathbf{X}=\ell'$ yazmalıyız. Bu nedenle $\ell'\boldsymbol{\beta}$ nın tahmin edilebilir olması için gerek ve yeter şart $\ell'\in\Re(\mathbf{X})$ olmasıdır. Genelleştirilmiş tersi ilgilendiren aşağıdaki gerçekler bu ve bundan sonraki bölümde sık sık kullanılacaktır.

(i) Herhangi bir $\mathbf{X}$ matrisi için $\Re(\mathbf{X}) = \Re(\mathbf{X'X})$ dir. Bu aşağıda gösterilmiştir. $\Re(\mathbf{X'X}) \subset \Re(\mathbf{X})$ olduğu açıktır. Bununla beraber $\mathbf{X'X}$ ve $\mathbf{X}$ aynı ranka sahiptirler ve bu nedenle onların satır uzayları aynı boyuta sahiptir. Bu, uzayların aynı uzaylar (eşit uzaylar) olmaları gerektiğini ifade eder. Bir sonuç olarak herhangi bir $\mathbf{M}$ matrisi için $\mathbf{X}=\mathbf{M}\ \mathbf{X'X}$ yazabiliriz.

(ii) Eğer $\mathcal{C}(\mathbf{B}) \subset \mathcal{C}(\mathbf{C})$ ve $\Re(\mathbf{A}) \subset \Re(\mathbf{C})$ ise, bu takdirde, **C**'nin $\mathbf{C}^-$ g-tersinin seçimi altında $\mathbf{AC}^-\mathbf{B}$ değişmez. Bu, aşağıdaki gibi görülür. Bazı **U, V** matrisleri için $\mathbf{B}=\mathbf{CU}$ ve $\mathbf{A}=\mathbf{VC}$ yazabiliriz. Bu takdirde

$$\mathbf{AC}^-\mathbf{B}=\mathbf{VCC}^-\mathbf{CU}=\mathbf{VCU}$$

dur. Yani $\mathbf{AC}^-\mathbf{B}$ g-tersin seçimine bağlı değildir.(**U, V** matrislerinin ister istemez tek olmadıklarına dikkat edelim. Bununla beraber, eğer $\mathbf{B}=\mathbf{CU}_1$, $\mathbf{A}=\mathbf{V}_1\mathbf{C}$ başka bir gösterim ise, bu taktirde,

$$\mathbf{V}_1\mathbf{CU}_1=\mathbf{V}_1\mathbf{CC}^-\mathbf{CU}_1=\mathbf{AC}^-\mathbf{B}=\mathbf{VCC}^-\mathbf{CU}=\mathbf{VCU}$$

dur.)

(iii) $\mathbf{X(X'X)^-X'}$ matrisi g-tersin seçimi altında değişmez. $\Re(\mathbf{X}) = \Re(\mathbf{X'X})$ olduğundan bu hemen görülür.

(iv) $\mathbf{X(X'X)^-X'X}=\mathbf{X}$, $\mathbf{X'X(X'X)^-X'}=\mathbf{X'}$ dür. Bu, $\mathbf{X}=\mathbf{MX'X}$ yazarak kolayca ispatlanır.

Teorem 2.2.1. $\ell'\boldsymbol{\beta}$ bir tahmin edilebilir lineer parametrik fonksiyon olsun ve **G**, **X** in en küçük kareler g-tersi olsun. Bu takdirde $\ell'\mathbf{Gy}$, $\ell'\boldsymbol{\beta}$'nın tüm yansız lineer tahminleri arasında minimum varyanslı bir yansız lineer tahminidir. $\ell'\mathbf{Gy}$'ye $\ell'\boldsymbol{\beta}$'nın en iyi lineer yansız tahmini diyeceğiz ve bundan böyle en iyi lineer yansız tahmin yerine "**BLUE**" yi kullanacağız. $\ell'\mathbf{Gy}$'nin varyansı $\sigma^2\ell'(\mathbf{X'X})^-\ell$ dir.

İspat: $\ell'\boldsymbol{\beta}$ tahmin edilebilir olduğundan herhangi **u** için $\ell' = \mathbf{u'X}$ dir. Bu takdirde

$$E(\ell'\mathbf{Gy}) = \mathbf{u'XGX}\boldsymbol{\beta} = \mathbf{u'X}\boldsymbol{\beta} = \ell'\boldsymbol{\beta}$$

dır ve bu nedenle $\ell'\mathbf{Gy}$, $\ell'\boldsymbol{\beta}$ için yansızdır. $\mathbf{w'X}=0$ olmak üzere diğer herhangi bir lineer yansız tahmin $(\ell'\mathbf{G}+\mathbf{w}')\mathbf{y}$ formuna sahiptir. Şimdi,

$$\begin{aligned}\operatorname{var}\left\{(\ell'\mathbf{G}+\mathbf{w}')\mathbf{y}\right\} &= \sigma^2(\ell'\mathbf{G}+\mathbf{w}')(\mathbf{G}'\ell+\mathbf{w}) \\ &= \sigma^2(\mathbf{u'XG}+\mathbf{w}')(\mathbf{G'X'u}+\mathbf{w})\end{aligned}$$

dur. **G, X**'in bir en küçük kareler g-tersi olduğundan,

$$\mathbf{u'XGw}=\mathbf{u'G'X'w}=0$$

ve bu nedenle,

$$\begin{aligned}\operatorname{var}\left\{(\ell'\mathbf{G}+\mathbf{w})\mathbf{y}\right\} &= \sigma^2(\mathbf{u}'(\mathbf{XG})(\mathbf{XG})'\mathbf{u}+\mathbf{w'w}) \\ &\geq \sigma^2\mathbf{u}'(\mathbf{XG})(\mathbf{XG})'\mathbf{u} \\ &= \operatorname{var}(\ell'\mathbf{Gy})\end{aligned}$$

dir. Böylece $\ell'\mathbf{Gy}$, $\ell'\boldsymbol{\beta}$'nın **BLUE**'sidir. $\ell'\mathbf{Gy}$'nin varyansı $\sigma^2\ell'\mathbf{GG}'\ell$ dir. g-tersin herhangi bir seçimi için $(\mathbf{X'X})^{-}\mathbf{X}'$'nün **X**'in bir en küçük kareler g-tersi olduğu kolayca görülür. Özellikle $\ell'(\mathbf{X'X})^{-}\ell=u'\mathbf{X}(\mathbf{X'X})^{-}\mathbf{X}'u$, g-tersin seçimine göre değişmez olduğundan, Moore-Penrose tersi kullanarak,

$$\begin{aligned}\ell'\mathbf{GG}'\ell &= \ell'(\mathbf{X'X})^{+}\mathbf{X'X}(\mathbf{X'X})^{-}\ell \\ &= \ell'(\mathbf{X'X})^{+}\ell \\ &= \ell'(\mathbf{X'X})\ \ell\end{aligned}$$

dir.

2.3. Hata (Tahmin Edilmiş Hata) Kareler Toplamı

Tanım 2.3.1. $\mathbf{X'X}\boldsymbol{\beta}=\mathbf{X'y}$ denklemlerine "normal denklemler" denir.

$\mathcal{C}(\mathbf{X}')=\mathcal{C}(\mathbf{X'X})$ olduğundan, denklemler tutarlıdırlar(çözümlüdürler). $\hat{\boldsymbol{\beta}}$ normal denklemin bir çözümü olsun. Bu takdirde g-tersin herhangi bir seçimi için $\hat{\boldsymbol{\beta}}=(\mathbf{X'X})\ \mathbf{X'y}$ dir. Hata kareler toplamı (HKT)

$$(\mathbf{y}-\mathbf{X}\hat{\boldsymbol{\beta}})'(\mathbf{y}-\mathbf{X}\hat{\boldsymbol{\beta}})$$

olarak tanımlanır. Her ne kadar $\hat{\boldsymbol{\beta}}$ g-tersin seçimine bağlı olmakla beraber HKT, $(\mathbf{X'X})^{-}$ g-tersinin seçimi altında değişmez. Bu nedenle $\hat{\boldsymbol{\beta}}$ tek değildir ve herhangi bir istatistiksel yorumu kabul etmez. "Uygun Model" ile genel olarak ilgili parametrik fonksiyonların **BLUE'** lerini hesaplamayı ve HKT'yi hesaplamayı ifade ederiz.

Teorem 2.3.1. $(\mathbf{y}-\mathbf{X\beta})'(\mathbf{y}-\mathbf{X\beta})$ 'nın minimumuna $\hat{\boldsymbol{\beta}}$ 'da ulaşılır.

İspat: $\mathbf{X}'(\mathbf{y}-\mathbf{X}\hat{\boldsymbol{\beta}})=\mathbf{X}'(\mathbf{y}-\mathbf{X}(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'\mathbf{y})=\mathbf{0}$ olduğundan,

$$(\mathbf{y}-\mathbf{X\beta})'(\mathbf{y}-\mathbf{X\beta})$$

$$=(\mathbf{y}-\mathbf{X}\hat{\boldsymbol{\beta}}+\mathbf{X}\hat{\boldsymbol{\beta}}-\mathbf{X\beta})'(\mathbf{y}-\mathbf{X}\hat{\boldsymbol{\beta}}+\mathbf{X}\hat{\boldsymbol{\beta}}-\mathbf{X\beta})$$

$$=(\mathbf{y}-\mathbf{X}\hat{\boldsymbol{\beta}})'(\mathbf{y}-\mathbf{X}\hat{\boldsymbol{\beta}})+(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})'\mathbf{X}'\mathbf{X}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})$$

dır. Buradan

$$(\mathbf{y}-\mathbf{X\beta})'(\mathbf{y}-\mathbf{X\beta})\geq(\mathbf{y}-\mathbf{X}\hat{\boldsymbol{\beta}})'(\mathbf{y}-\mathbf{X}\hat{\boldsymbol{\beta}})$$

olduğu görülür. Bu da ispatı tamamlar.

Teorem 2.3.2. R(**X**)=r olsun. Bu takdirde $E(\mathbf{y}-\mathbf{X}\hat{\boldsymbol{\beta}})'(\mathbf{y}-\mathbf{X}\hat{\boldsymbol{\beta}})=(n-r)\sigma^2$ dir.

İspat: $E(\mathbf{y}-\mathbf{X\beta})(\mathbf{y}-\mathbf{X\beta})'=D(\mathbf{y})=\sigma^2\mathbf{I}$ elde ederiz. Bu nedenle,

$$E(\mathbf{yy}')=E(\mathbf{y})\boldsymbol{\beta}'\mathbf{X}'+\mathbf{X\beta}E(\mathbf{y}')-\mathbf{X\beta\beta}'\mathbf{X}'+\sigma^2\mathbf{I}$$

$$=\mathbf{X\beta\beta}'\mathbf{X}'+\sigma^2\mathbf{I} \qquad (2.3.1)$$

dır. Bu ve sonraki bölümde,

$$\mathbf{P}=\mathbf{I}-\mathbf{X}(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'$$

notasyonunu kullanacağız. **P'**nin, bir simetrik-idempotent matris ve **PX=0** olduğuna dikkat edelim. Bu özellikler yararlı olacaktır. Şimdi, (2.3.1) bağıntısından ve **PX=0** gerçeğinden,

$$E(\mathbf{y}-\mathbf{X}\hat{\beta})'(\mathbf{y}-\mathbf{X}\hat{\beta})=E(\mathbf{y}-\mathbf{X}(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'\mathbf{y})'(\mathbf{y}-\mathbf{X}(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'\mathbf{y})$$

$$=E\mathbf{y}'\mathbf{P}\mathbf{y}$$

$$=E\ \mathrm{iz}(\mathbf{y}'\mathbf{P}\mathbf{y})$$

$$=E\ \mathrm{iz}(\mathbf{P}\mathbf{y}\mathbf{y}')$$

$= \text{iz } \mathbf{P}E(\mathbf{yy}')$

$= \sigma^2 \text{ iz}\mathbf{P}$

dir ve son olarak $\mathbf{(X'X)^{-}X'X}$ idempotent olduğundan,

$$\begin{aligned} iz(\mathbf{P}) &= n - iz\mathbf{X(X'X)^{-}X'} \\ &= n - iz\mathbf{(X'X)^{-}X'X} \\ &= n - \mathrm{R}(\mathbf{(X'X)^{-}X'X}) \end{aligned}$$

dır. Halbuki

$$\mathrm{R}(\mathbf{(X'X)^{-}X'X}) = \mathrm{R}(\mathbf{X'X}) = \mathrm{R}(\mathbf{X}) = r$$

dir ve ispat tamamdır. Teorem 2.4.2'den $\mathrm{HKT}/(n-r)$'nin σ^2'nin bir yansız tahmin edicisi olduğu sonucunu çıkarırız, yani $E\left\{\frac{\mathrm{HKT}}{n-r}\right\} = \frac{E\{\mathrm{HKT}\}}{n-r} = \frac{(n-r)\sigma^2}{n-r} = \sigma^2$ dir.

Hesaplamalar için, HKT=$\mathbf{y'y} - \hat{\boldsymbol{\beta}}'\mathbf{X'X}\hat{\boldsymbol{\beta}} = \mathbf{y'y} - \mathbf{y'X}\hat{\boldsymbol{\beta}}$ ifadelerini kullanmak daha uygundur.

2.4. Kısıtlamalara Bağlı Tahmin

$\mathbf{y}$, bir $nx1$ vektör $\mathbf{X}$, bir nxp matris olmak üzere bilinen $E(\mathbf{y}) = \mathbf{X}\boldsymbol{\beta}$, $D(\mathbf{y}) = \sigma^2\mathbf{I}$ modelini göz önüne alalım. Parametreler üzerinde bir ön lineer $\mathbf{L}\boldsymbol{\beta} = \mathbf{z}$ kısıtlamasına sahip olduğumuzu varsayalım. $\Re(\mathbf{L}) \subset \Re(\mathbf{X})$ olduğunu ve $\mathbf{L}\boldsymbol{\beta} = \mathbf{z}$'nin tutarlı olduğunu kabul edelim. Sabit bir $\mathbf{(X'X)^{-}}$ g-tersi için $\hat{\boldsymbol{\beta}} = \mathbf{(X'X)^{-}X'y}$ olsun ve

$$\tilde{\boldsymbol{\beta}} = \hat{\boldsymbol{\beta}} - \mathbf{(X'X)^{-}L'(L(X'X)^{-}L')^{-}}(\mathbf{L}\hat{\boldsymbol{\beta}} - \mathbf{z})$$

olsun.

Teorem 2.4.1. $(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$'nın $\mathbf{L}\boldsymbol{\beta} = \mathbf{z}$'ye bağlı minimumuna $\boldsymbol{\beta} = \tilde{\boldsymbol{\beta}}$'da ulaşılır.

İspat: $\Re(\mathbf{L}) \subset \Re(\mathbf{X})$ ve $\mathrm{R}(\mathbf{X}) = \mathrm{R}(\mathbf{X'X})$ olduğundan herhangi bir $\mathbf{W}$ için $\mathbf{L} = \mathbf{WX'X}$ dir. $\mathbf{T} = \mathbf{WX'}$ olsun. Şimdi,

$$
\begin{aligned}
\mathbf{L(X'X)^{-}L'} &= \mathbf{WX'X(X'X)^{-}X'XW'} \\
&= \mathbf{WX'XW'} \\
&= \mathbf{TT'}
\end{aligned}
$$

dür. $\mathbf{L\beta = z}$ tutarlı olduğundan, herhangi $\mathbf{v}$ için $\mathbf{Lv=z}$ dir. Böylece,

$$
\begin{aligned}
\mathbf{L(X'X)^{-}L'(L(X'X)^{-}L')z} &= \mathbf{L(X'X)^{-}L'(L(X'X)^{-}L')^{-}Lv} \\
&= \mathbf{TT'(TT')TXv} \\
&= \mathbf{TXv} \\
&= \mathbf{Lv} \\
&= \mathbf{z}
\end{aligned}
\tag{2.4.1}
$$

dir. Benzer şekilde,

$$
\begin{aligned}
\mathbf{L(X'X)^{-}L'(L(X'X)^{-}L')L\hat{\beta}} &= \mathbf{TT'(TT')^{-}WX'X(X'X)^{-}X'y} \\
&= \mathbf{TT'(TT')^{-}WX'v} \\
&= \mathbf{TT'(TT')^{-}Ty} \\
&= \mathbf{Ty}
\end{aligned}
\tag{2.4.2}
$$

ve

$$
\begin{aligned}
\mathbf{L\hat{\beta}} &= \mathbf{L(X'X)^{-}X'y} \\
&= \mathbf{WX'X(X'X)^{-}X'y} \\
&= \mathbf{WX'y} \\
&= \mathbf{Ty}
\end{aligned}
\tag{2.4.3}
$$

dir. (2.4.1), (2.4.2) ve (2.4.3) bağıntılarını kullanarak, $\mathbf{L\tilde{\beta} = z}$ olduğunu ve böylece $\tilde{\boldsymbol{\beta}}$ nın $\mathbf{L\beta = z}$ kısıtlamasını sağladığını görürüz.

Şimdi, $\mathbf{(\tilde{\beta}-\beta)'X'(y-X\tilde{\beta}) = 0}$ olduğunu aşağıdaki gibi gösterebileceğimizden, yani, $\mathbf{L' = X'XW'}$ olduğundan;

$$
\begin{aligned}
\mathbf{X'X\tilde{\beta}} &= \mathbf{X'X\hat{\beta} - X'X(X'X)^{-}L'(L(X'X)^{-}L')^{-}(L\hat{\beta} - z)} \\
&= \mathbf{X'y - L'(L(X'X)^{-}L')^{-}(L\hat{\beta} - z)}
\end{aligned}
$$

bağıntısını elde ederiz. Bu nedenle

$$
\mathbf{X'(y - X\tilde{\beta}) = L'(L(X'X)^{-}L')^{-}(L\hat{\beta} - z)}
$$

dır ve $\mathbf{L\tilde{\beta} = L\beta = z}$ olduğundan,

$$(\tilde{\boldsymbol{\beta}}-\boldsymbol{\beta})'\mathbf{X}'(\mathbf{y}-\mathbf{X}\tilde{\boldsymbol{\beta}})=(\tilde{\boldsymbol{\beta}}-\boldsymbol{\beta})'\mathbf{L}'(\mathbf{L}(\mathbf{X}'\mathbf{X})^{-}(\mathbf{L}\tilde{\boldsymbol{\beta}}-\mathbf{z})=\mathbf{0}$$

olduğu görülür. $\mathbf{L}\boldsymbol{\beta}=\mathbf{z}$ denklemini sağlayan herhangi bir $\boldsymbol{\beta}$ için,

$$\begin{aligned}&(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})'(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})\\&=(\mathbf{y}-\mathbf{X}\tilde{\boldsymbol{\beta}}+\mathbf{X}(\tilde{\boldsymbol{\beta}}-\boldsymbol{\beta}))'(\mathbf{y}-\mathbf{X}\tilde{\boldsymbol{\beta}}+\mathbf{X}(\tilde{\boldsymbol{\beta}}-\boldsymbol{\beta}))\\&=(\mathbf{y}-\mathbf{X}\tilde{\boldsymbol{\beta}})'(\mathbf{y}-\mathbf{X}\tilde{\boldsymbol{\beta}})+(\tilde{\boldsymbol{\beta}}-\boldsymbol{\beta})'\mathbf{X}'\mathbf{X}(\tilde{\boldsymbol{\beta}}-\boldsymbol{\beta})\end{aligned}\tag{2.4.4}$$

elde edilir. (2.4.4) bağıntısından, eğer $\mathbf{L}\boldsymbol{\beta}=\mathbf{z}$ ise,

$$(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})'(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})\geq(\mathbf{y}-\mathbf{X}\tilde{\boldsymbol{\beta}})'(\mathbf{y}-\mathbf{X}\tilde{\boldsymbol{\beta}})$$

olduğu açıktır ve ispat tamamlanır.

Teorem 2.4.2. R(L)=R(T)=R(L($\mathbf{X}'\mathbf{X}$)$^{-}$ $\mathbf{L}'$) dir.

İspat: $\mathbf{L}(\mathbf{X}'\mathbf{X})^{-}\mathbf{L}'=\mathbf{T}\mathbf{T}'$ olduğundan $\mathrm{R}(\mathbf{L}(\mathbf{X}'\mathbf{X})^{-}\mathbf{L}')=\mathrm{R}(\mathbf{T}\mathbf{T}')=\mathrm{R}(\mathbf{T})$ dır. $\mathrm{R}(\mathbf{L})=\mathrm{R}(\mathbf{T}\mathbf{X})\leq\mathrm{R}(\mathbf{T})$ olduğu açıktır. $\mathrm{R}(\mathbf{X})=\mathrm{R}(\mathbf{X}'\mathbf{X})$ olduğundan herhangi bir $\mathbf{M}$ için $\mathbf{X}=\mathbf{M}\mathbf{X}'\mathbf{X}$ dır. Böylece $\mathbf{T}=\mathbf{W}\mathbf{X}'=\mathbf{W}\mathbf{X}'\mathbf{X}\mathbf{M}'=\mathbf{L}\mathbf{M}'$ dir. Bu nedenle $\mathrm{R}(\mathbf{T})\leq\mathrm{R}(\mathbf{L})$ dir ve buradan $\mathrm{R}(\mathbf{T})=\mathrm{R}(\mathbf{L})$ dir.

Eğer ek varsayımlar yapılırsa bazı sadeleştirmelerin ortaya çıktığına dikkat edelim. R(X)=p olduğunu, bu nedenle bir tam-ranklı modele sahip olduğumuzu varsayalım. Aynı zamanda **L**'nin m ranklı bir mxp matris olduğunu kabul edelim. Bu takdirde Teorem 2.4.2'ye göre

$$\mathrm{R}(\mathbf{L}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{L}')=\mathrm{R}(\mathbf{L})=m$$

dir ve bundan dolayı $\mathbf{L}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{L}'$ tekil değildir. Eğer m=1 ise $\mathbf{L}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{L}'$ bir skalar olur.

2.5. Lineer Hipotezlerin Testleri

2.5.1. Schur Tümleyenleri

Eğer **A** pozitif tanımlı ise, bu takdirde, **A**'nın tüm esas alt matrisleri pozitif tanımlıdır. Bu nedenle **A**'nın tüm esas minörleri pozitiftir. Şimdi tersini ispatlayacağız. Determinantı bir sütun boyunca birkaç defa açmak suretiyle ispatı ortaya konan aşağıdaki sonuç kullanılacak.

Teorem 2.5.1.1. **A** bir nxn matris olsun. Bu takdirde herhangi bir μ için

$$|\mathbf{A}+\mu\mathbf{I}|=\sum_{i=0}^{n}\mu^{n-i}s_i \qquad (2.5.1.1)$$

dir. Burada s_i, **A**'nın tüm ixi ana minörlerinin toplamıdır. $s_0=1$ alalım. $s_n=|\mathbf{A}|$ olduğuna dikkat edelim. Eğer **A** bir nxn simetrik matris ise ve **A**'nın tüm ana minörleri pozitif ise, bu takdirde, Teorem 2.5.1.1'e göre herhangi bir $\mu\geq 0$ için ($\mu=0$ olduğunda, $|\mathbf{A}|>0$ gerçeğini kullanırız.) $|\mathbf{A}+\mu\mathbf{I}|>0$ dır. Böylece **A** pozitif olamayan bir özdeğere sahip olamaz ve bu nedenle **A** pozitif tanımlıdır. Bu gözlemi Teorem 1.8.3 ve Teorem 1.8.4 ile bir araya getirerek aşağıdaki ifadeyi elde ederiz.

Teorem 2.5.1.2. **A** bir nxn matris olsun. Bu takdirde **A**'nın pozitif tanımlı olması için gerek ve yeter şart **A**'nın tüm ana minörlerinin pozitif olmasıdır.

Teorem 2.5.1.3. Aynı şekilde bir simetrik matrisin pozitif yarı tanımlı olması için gerek ve yeter şart onun tüm ana minörlerinin negatif olmamasıdır.

Tanım 2.5.1.1. **B, D** kare matrisler olmak üzere,

$$\mathbf{A}=\begin{bmatrix}\mathbf{B} & \mathbf{C}\\ \mathbf{C}' & \mathbf{D}\end{bmatrix} \qquad (2.5.1.2)$$

olarak parçalanmış olan bir simetrik matris olsun. Eğer **B** tekil değilse bu takdirde **B**'nin **A**'daki Schur tümleyeni $\mathbf{D}-\mathbf{C}'\mathbf{B}^{-1}\mathbf{C}$ olarak tanımlanır. Aynı şekilde eğer **D** tekil değilse, bu takdirde **D**'nin **A**'daki Schur tümleyeni $\mathbf{B}-\mathbf{C}\mathbf{D}^{-1}\mathbf{C}'$ dir. **B** tekil olmasın ve $\mathbf{X}=-\mathbf{C}'\mathbf{B}^{-1}$ olsun. Aşağıdaki özdeşlik basit matris çarpımıyla gerçeklenebilir;

$$\begin{bmatrix}\mathbf{I} & \mathbf{0}\\ \mathbf{X} & \mathbf{I}\end{bmatrix}\begin{bmatrix}\mathbf{B} & \mathbf{C}\\ \mathbf{C}' & \mathbf{D}\end{bmatrix}\begin{bmatrix}\mathbf{I} & \mathbf{X}'\\ \mathbf{0} & \mathbf{I}\end{bmatrix}=\begin{bmatrix}\mathbf{B} & \mathbf{0}\\ \mathbf{0} & \mathbf{D}-\mathbf{C}'\mathbf{B}^{-1}\mathbf{C}\end{bmatrix} \qquad (2.5.1.3)$$

dır. Bir çok faydalı gerçekler (2.5.1.3) bağıntısı kullanılarak gösterilebilir.

Teorem 2.5.1.4. Aşağıdaki iddialar doğrudur.

(i) **A** pozitif tanımlı ise, bu takdirde, $\mathbf{D}-\mathbf{C}'\mathbf{B}^{-1}\mathbf{C}$ pozitif tanımlıdır.

(ii) **A** simetrik olsun. Eğer **A**'nın bir ana alt matrisi ve onun **A**'daki Schur tümleyeni pozitif tanımlı ise bu takdirde **A** da pozitif tanımlıdır.

(iii) $|\mathbf{A}| = |\mathbf{B}||\mathbf{D}-\mathbf{C}'\mathbf{B}^{-1}\mathbf{C}|$ dir.

İspat: (i): Açık olarak eğer **A** pozitif tanımlı ise, bu takdirde, herhangi bir tekil olmayan **S** matrisi için $\mathbf{SAS}'$ pozitif tanımlıdır. Önceden olduğu gibi $\mathbf{X} = -\mathbf{C}'\mathbf{B}^{-1}$ olmak üzere eğer,

$$\mathbf{S} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{X} & \mathbf{I} \end{bmatrix}$$

ise, bu takdirde, $|\mathbf{S}| = \mathbf{1}$ dir ve bundan dolayı **S** tekil değildir. Böylece $\mathbf{SAS}'$ pozitif tanımlıdır (bkz. 2.5.1.3), ve $\mathbf{D}-\mathbf{C}'\mathbf{B}^{-1}\mathbf{C}$, $\mathbf{SAS}'$'nın bir esas alt matrisi olduğundan bu matris pozitif tanımlıdır.

(ii): **A**'nın (2.5.1.2) bağıntısındaki gibi parçalandığını ve **B** ve $\mathbf{D}-\mathbf{C}'\mathbf{B}^{-1}\mathbf{C}$'nin pozitif tanımlı olduğunu farz edelim. Bu takdirde (2.5.1.3) bağıntısının sağ-yanı pozitif tanımlıdır ve (i)'de tanımlanan **S** tekil olmadığından, **A**'nın pozitif tanımlı olduğu görülür.

(iii)**:** Bu, (2.5.1.3) bağıntısında her iki yanın determinantını almak suretiyle hemen görülür. (2.5.1.2) bağıntısında **A**'nın bir *nxn* matris olduğunu ve **B**'nin de *(n-1)*x*(n-1)* matris olduğunu farz edelim. Bu takdirde **C** bir sütun vektördür ve **D** *1*x*1* matristir, yani, bir skalardır. (2.5.1.2) bağıntısını;

$$\mathbf{A} = \begin{bmatrix} \mathbf{B} & \mathbf{c} \\ \mathbf{c}' & \mathbf{d} \end{bmatrix} \tag{2.5.1.4}$$

olarak yazalım. **B**'nin **A**'daki Schur tümleyeni bir skalar olan, $\mathbf{d}-\mathbf{c}'\mathbf{B}^{-1}\mathbf{c}$ dir. Teorem 2.5.1.4 kısım (iii)'e göre,

$$\mathbf{d}-\mathbf{c}'\mathbf{B}^{-1}\mathbf{c} = \frac{|\mathbf{A}|}{|\mathbf{B}|} \tag{2.5.1.5}$$

dır.

Tanım 2.5.1.2. Herhangi bir *k* için *1,...,k* satır ve *1,...,k* sütun tarafından oluşturulan bir esas alt matrise bir baş esas alt matris denir ve onun determinantına da bir baş esas minör denir. Pozitif tanımlı matrislerin başka bir tanımını elde etmek için şimdi hazırız.

Teorem 2.5.1.5. **A**, bir nxn simetrik matris olsun. Bu takdirde **A**'nın pozitif tanımlı olması için gerek ve yeter şart **A**'nın tüm baş esas minörlerinin pozitif olmasıdır.

İspat: Aşikar olarak eğer **A** pozitif tanımlı ise, bu takdirde, onun baş esas minörleri pozitiftir. Tersini indüksiyonla ispatlarız. n=1 için sonuç açıktır. *(n-1)*x*(n-1)* matrisleri için sonucun doğru olduğunu kabul edelim. **A**, (2.5.1.4) bağıntısındaki gibi parçalanmış olsun. **B**'nin herhangi bir baş esas minörü pozitif olmak zorunda olduğundan, indüksiyon varsayımına göre **B** pozitif tanımlıdır. Aynı zamanda $|\mathbf{A}| > 0$' dır ve bu nedenle (2.5.1.5) bağıntısına göre **B**' nin **A**'daki Schur tümleyeni, yani, $\mathbf{d}-\mathbf{c}'\mathbf{B}^{-1}\mathbf{c}$, pozitiftir. Bu nedenle Teorem 2.5.1.4 kısım (iii)'e göre **A** pozitif tanımlıdır ve ispat tamamlanır.

A ve **B** nxn matrisler olsunlar. **A, B** ve **A-B**'nin tümünün pozitif yarı-tanımlı matrisler oldukları gerçeğini göstermek için $\mathbf{A} \geq \mathbf{B}$ yazarız. Eğer $\mathbf{B} \geq \mathbf{A}$ doğru ise $\mathbf{A} \leq \mathbf{B}$ yazarız.

Teorem 2.5.1.6. **A, B** , $\mathbf{A} \geq \mathbf{B}$ olacak şekilde pozitif tanımlı matrisler olsun. Bu takdirde $\mathbf{A}^{-1} \leq \mathbf{B}^{-1}$ dir.

İspat: İlk olarak, **B=I** olduğunu farz edelim. Bu takdirde $\mathbf{A} \geq \mathbf{I}$ olması $\mathbf{A}-\mathbf{I}$'nın pozitif yarı-tanımlı olduğunu ifade eder. Bu nedenle **A**'nın her bir öz değeri 1 den büyük veya 1 e eşittir. Böylece, $\mathbf{A}^{-1}$' in her bir öz değeri 1 den küçük veya 1 e eşittir ve $\mathbf{A}^{-1} \leq \mathbf{I}$ dır. Genel olarak $\mathbf{A} \geq \mathbf{B}$;

$$\mathbf{B}^{-1/2}\mathbf{A}\mathbf{B}^{-1/2} \geq \mathbf{I}$$

olduğunu ifade eder ve artık birinci kısım ispatı bitirmek için kullanılabilir.

2.5.2 Çok Değişkenli Normal Dağılım

Tanım 2.5.2.1. **u**, $u_1,...,u_n$ bileşenleri bağımsız standart normal dağılıma sahip olan, n mertebeli bir rastgele vektör olsun. **X** bir rxn matris olsun ve μ sabit bir rx1 vektör olsun. $\mathbf{y} = \mathbf{X}\mathbf{u} + \mu$ vektörüne bir r-boyutlu çok değişkenli normal dağılıma sahiptir denir. Açık olarak $E(\mathbf{y}) = \mathbf{X}E(\mathbf{u}) + \mu = \mu$ ve $D(\mathbf{y}) = \mathbf{X}D(\mathbf{u})\mathbf{X}' = \mathbf{X}\mathbf{X}'$ dir. $\Sigma = \mathbf{X}\mathbf{X}'$ olsun. Şimdi **y**'nin,

$$\phi_{\mathbf{y}}(\mathbf{t}) = E(\exp(i\mathbf{t}'\mathbf{y}))$$

olarak tanımlanan, $\phi_{\mathbf{y}}(\mathbf{t})$ karakteristik fonksiyonunu elde edeceğiz.

İlk olarak,

$$\begin{aligned}\phi_{\mathbf{u}}(\mathbf{t}) &= E(\exp(i\mathbf{t}'\mathbf{u})) = \prod_{j=1}^{n} E(\exp(it_j u_j)) \\ &= \prod_{j=1}^{n} \exp(-\frac{t_j^2}{2}) = \exp(-\frac{\mathbf{t}'\mathbf{t}}{2})\end{aligned}$$

elde ederiz. Şimdi,

$$\begin{aligned}\phi_{\mathbf{y}}(\mathbf{t}) &= E(\exp(i\mathbf{t}'\mathbf{y})) = E(\exp(i\mathbf{t}'(\mathbf{X}\mathbf{u}+\mu)) \\ &= \exp(i\mathbf{t}'\mu)E(\exp(i\mathbf{t}'\mathbf{X}\mathbf{u})) = \exp(i\mathbf{t}'\mu)\phi_{\mathbf{u}}(\mathbf{t}'\mathbf{X}) \\ &= \exp(i\mathbf{t}'\mu)\exp(-\frac{1}{2}\mathbf{t}'\mathbf{X}\mathbf{X}'\mathbf{t}) = \exp(i\mathbf{t}'\mu-\frac{1}{2}\mathbf{t}'\Sigma\mathbf{t})\end{aligned} \qquad (2.5.2.1)$$

dır. Bu nedenle $\mathbf{y}$'nin dağılımı yalnız μ ve Σ'ya bağlıdır. Bundan dolayı $\mathbf{y}\sim N(\mu,\ \Sigma)$ notasyonunu kullanacağız. Şimdi, Σ tekil olmadığında $\mathbf{y}$'nin yoğunluk fonksiyonunun

$$f(\mathbf{y}) = \frac{1}{(2\pi)^{\frac{n}{2}}|\Sigma|^{\frac{1}{2}}}\exp\left(-\frac{1}{2}(\mathbf{y}-\mu)'\Sigma^{-1}(\mathbf{y}-\mu)\right) \qquad (2.5.2.2)$$

ile verildiğini göstereceğiz. Eğer bir $\mathbf{y}$ rastgele vektörü (2.5.2.2) ile verilen yoğunluk fonksiyonuna sahipse bu takdirde $\mathbf{y}$'nin karakteristik fonksiyonunun (2.5.2.1) ile verildiğini göstereceğiz. Bu takdirde bir karakteristik fonksiyona karşılık gelen dağılımın tekliğiyle, Σ tekil olmamak üzere eğer y, $N(\mu,\ \Sigma)$ ise, bu takdirde, $\mathbf{y}$' nin yoğunluk fonksiyonunun (2.5.2.2) olduğu görülecektir.

İlk olarak (2.5.2.2)'deki fonksiyonun integralinin 1'e eşit olduğunu ve bu nedenle onun bir yoğunluk fonksiyonu olduğunu gerçekleyelim. $\mathbf{z} = \Sigma^{-\frac{1}{2}}(\mathbf{y}-\mu)$ dönüşümünü yapalım. Dönüşümünün jakobiyeni $\left|\left(\frac{\partial z_i}{\partial y_j}\right)\right|$ ' nun determinantının mutlak değeridir ve jakobiyeninin kolaylıkla $\left|\Sigma^{\frac{1}{2}}\right|$ olacağı görülür. Bu nedenle, çarpımdaki her bir terim bir standart normal yoğunluğun tam integrali olduğundan,

$$\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f(\mathbf{y})dy_1...dy_n$$

$$=\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} \frac{1}{(2\pi)^{\frac{n}{2}}|\Sigma|^{\frac{1}{2}}} \text{x} \exp\left(-\frac{1}{2}\mathbf{z}'\mathbf{z}\right)|\Sigma|^{\frac{1}{2}} dz_1...dz_n$$

$$=\frac{1}{(2\pi)^{\frac{n}{2}}}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}\sum_{j=1}^{n} z_j^2\right)dz_1...dz_n$$

$$=\prod_{j=1}^{n}\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}z_j^2\right)dz_j$$

$$=1$$

dir. $\mathbf{y}$' nin karakteristik fonksiyonu,

$$\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} \frac{1}{(2\pi)^{\frac{n}{2}}|\Sigma|^{\frac{1}{2}}}\exp\left(-\frac{1}{2}(\mathbf{y}-\mu)'\Sigma^{-1}(\mathbf{y}-\mu)\right)\exp(i\mathbf{t}'\mathbf{y})dy_1...dy_n \qquad (2.5.2.3)$$

ile verilir. (2.5.2.3) bağıntısında $\mathbf{z}=\mathbf{y}-\mu$ dönüşümü yapalım. Jakobiyen aşikar olarak 1 dir. Bu nedenle (2.5.2.3) bağıntısındaki integral,

$$\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} \frac{1}{(2\pi)^{\frac{n}{2}}|\Sigma|^{\frac{1}{2}}}\exp\left(-\frac{1}{2}\mathbf{z}'\Sigma^{-1}\mathbf{z}\right)\exp(i\mathbf{t}'(\mathbf{z}+\mu))dz_1...dz_n \qquad (2.5.2.4)$$

dir. Bu ise,

$$\Delta=\frac{1}{(2\pi)^{\frac{n}{2}}|\Sigma|^{\frac{1}{2}}}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\exp\left(-\frac{1}{2}(\mathbf{z}-\Sigma\mathbf{t})'\Sigma^{-1}(\mathbf{z}-\Sigma\mathbf{t})\right)dz_1...dz_n$$

olmak üzere

$$\exp\left(i\mathbf{t}'\mu-\frac{1}{2}t'\Sigma\mathbf{t}\right)\text{x}\Delta$$

ile aynıdır. Δ'da $\mathbf{u}=\mathbf{z}-\Sigma\mathbf{t}$ dönüşümünü yapalım. Bu takdirde Δ bir standart normal yoğunluğun integraline indirgenir. Bu nedenle 1'e eşittir. Böylece $\mathbf{y}$' nin karakteristik fonksiyonunun,

$$\exp\left(i\mathbf{t}'\mu-\frac{1}{2}\mathbf{t}'\Sigma\mathbf{t}\right)$$

ile verildiği sonucunu çıkarırız.

Teorem 2.5.2.1. $\mathbf{y} \sim N(\mu,\Sigma)$ olsun ve $\mathbf{y}$, μ ve Σ' nin

$$\mathbf{y} = \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{pmatrix},\quad \mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix},\quad \Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \tag{2.5.2.5}$$

olarak uygun bir şekilde parçalandıklarını farz edelim. Bu takdirde $\mathbf{y}_1 \sim N(\mu_1,\Sigma_{11})$ ve $\mathbf{y}_2 \sim N(\mu_2,\Sigma_{22})$ dir.

İspat: $\mathbf{t} = \begin{pmatrix} \mathbf{t}_1 \\ \mathbf{t}_2 \end{pmatrix}$, **t**' nin karşılık gelen parçalanması olmak üzere, $\mathbf{y}_1$' nin karakteristik fonksiyonu $\mathbf{t}_2$=**0** koyarak elde edilir. Böylece,

$$\phi_{\mathbf{y}_1}(\mathbf{t}_1) = \exp\left(i\mathbf{t}_1'\mu_1 - \frac{1}{2}\mathbf{t}_1'\Sigma_{11}\mathbf{t}_1 \right)$$

dir ve bu nedenle $\mathbf{y}_1 \sim N(\mu_1,\Sigma_{11})$ dir. Aynı şekilde $\mathbf{y}_2 \sim N(\mu_2,\Sigma_{22})$ dir.

Teorem 2.5.2.2. $\mathbf{y} \sim N(\mu,\Sigma)$ olsun ve $\mathbf{y}$, μ ve Σ'nin (2.5.2.5) bağıntısındaki gibi uygun bir şekilde parçalandıklarını ifade edelim. Bu takdirde $\mathbf{y}_1$ ve $\mathbf{y}_2$' nin bağımsız olmaları için gerek ve yeter şart Σ_{12}=0 olmasıdır.

İspat: Eğer $\mathbf{y}_1$ ve $\mathbf{y}_2$ bağımsız iseler, bu takdirde kov($\mathbf{y}_1,\mathbf{y}_2$) =$\Sigma_{12}$=0 dır. Şimdi tersinin de doğru olduğunu göstereceğiz. Bu nedenle Σ_{12}=0 olduğunu farz edelim. Bu takdirde,

$$\mathbf{t}'\Sigma\mathbf{t} = \mathbf{t}_1'\Sigma_{11}\mathbf{t}_1 + \mathbf{t}_2'\Sigma_{22}\mathbf{t}_2$$

dir. Böylece

$$\phi_{\mathbf{y}}(\mathbf{t}) = \phi_{\mathbf{y}_1}(\mathbf{t}_1)\phi_{\mathbf{y}_2}(\mathbf{t}_2)$$

dır ve $\mathbf{y}_1$ ve $\mathbf{y}_2$ bağımsızdırlar.

Teorem 2.5.2.3. $\mathbf{y} \sim N(\mu,\sigma^2\mathbf{I})$ olsun ve **A**, **B**; $\mathbf{A}\mathbf{B}'$=0 olacak şekilde matrisler olsunlar. Bu takdirde **Ay**, **By** bağımsızdırlar.

İspat: Teorem 2.5.2.1' e göre

$$\begin{bmatrix} \mathbf{A} \\ \mathbf{B} \end{bmatrix} \mathbf{y} = \begin{bmatrix} \mathbf{Ay} \\ \mathbf{By} \end{bmatrix}$$

'nın çok değişkenli normal dağılıma sahip olduğuna dikkat edelim. Bu nedenle Teorem 2.5.2.2'ye göre **Ay, By'** nin bağımsız olmaları için kov**(Ay, By)=** $\mathbf{A}\mathbf{B}'=0$ olması gerekir.

Şimdi Σ'nın tekil olmadığını farz edelim ve $\mathbf{y}_1$ verildiğinde $\mathbf{y}_2$'nin şartlı dağılımını elde edeceğiz. (2.5.1.3) özdeşliğinin Σ'ya uygulandığını düşünelim. Bu takdirde $\mathbf{X} = -\Sigma_{21}\Sigma_{11}^{-1}$ dir.

$$\mathbf{S} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{X} & \mathbf{I} \end{bmatrix}$$

olsun bu takdirde,

$$\mathbf{S}^{-1} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{-X} & \mathbf{I} \end{bmatrix}$$

dir.

$$\Sigma = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{-X} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \Sigma_{11} & \mathbf{0} \\ \mathbf{0} & \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{-X}' \\ \mathbf{0} & \mathbf{I} \end{bmatrix}$$

olduğu sonucunu çıkarırız. Böylece $\tilde{\Sigma}_{22} = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}$ Σ' da Σ_{11}' nın Schur tümleyeni olmak üzere

$$\Sigma^{-1} = \mathbf{S}' \begin{bmatrix} \Sigma_{11}^{-1} & 0 \\ 0 & \tilde{\Sigma}_{22}^{-1} \end{bmatrix} \mathbf{S} \qquad (2.5.2.6)$$

dir. Şimdi (2.5.2.6) bağıntısını kullanarak,

$$(\mathbf{y} - \mu)'\Sigma^{-1}(\mathbf{y} - \mu)$$

$$=((\mathbf{y}_1 - \mu_1)', (\mathbf{y}_2 - \mu_2)')\Sigma^{-1} \begin{bmatrix} \mathbf{y}_1 - \mu_1 \\ \mathbf{y}_2 - \mu_2 \end{bmatrix}$$

$$=(\mathbf{y}_1 - \mu_1)'\Sigma_{11}^{-1}(\mathbf{y}_1 - \mu_1) + ((\mathbf{y}_2 - \mu_2)' + (\mathbf{y}_1 - \mu_1)'\mathbf{X}')\tilde{\Sigma}_{22}^{-1}((\mathbf{y}_2 - \mu_2) + \mathbf{X}(\mathbf{y}_1 - \mu_1))$$

olduğu görülür. Aynı zamanda,

$$|\Sigma| = |\Sigma_{11}||\tilde{\Sigma}_{22}|$$

dir. $\mathbf{y}_1$ verildiğinde $\mathbf{y}_2$'nin şartlı yoğunluğunu elde etmek için, (2.5.2.2) bağıntısında verilen $\mathbf{y}$'nin yoğunluk fonksiyonunda bu ifadeleri yerine koyarız ve sonra da $\mathbf{y}_1$'in marjinal yoğunluğu ile yani $N(\mu_1, \Sigma_{11})$ yoğunluğu ile böleriz. Bu, $\mathbf{y}_1$ verildiğinde $\mathbf{y}_2$'nin şartlı dağılımının

$$\mu_2 - \mathbf{X}(\mathbf{y}_1 - \mu_1) = \mu_2 + \Sigma_{21}\Sigma_{11}^{-1}(\mathbf{y}_1 - \mu_1)$$

ortalama vektörü ve $\tilde{\Sigma}_{22}$ dağılım matrisine sahip çok değişkenli normal dağılım olduğunu ortaya koyar.

2.5.3 Kuadratik (Karesel) Formlar ve Cochran Teoremi

Teorem 2.5.3.1. $\mathbf{y} \sim N(\mathbf{0}, \mathbf{I}_n)$ olsun ve **A** nxn boyutlu bir simetrik matris olsun. Bu takdirde $\mathbf{y}'\mathbf{A}\mathbf{y}$' nin r serbestlik dereceli ki-kare (χ_r^2) dağılıma sahip olması için gerek ve yeter şart **A**' nın idempotent ve R(**A**)=r olmasıdır.

İspat: Eğer **A**, r ranklı idempotent bir matris ise, bu takdirde,

$$\mathbf{A} = \mathbf{P}'\begin{bmatrix} \mathbf{I}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\mathbf{P}$$

olacak şekilde bir ortogonal ($\mathbf{P}'\mathbf{P} = \mathbf{P}\mathbf{P}' = \mathbf{I}$) $\mathbf{P}$ matrisi mevcuttur.

z=Py olsun. Bu takdirde $\mathbf{z} \sim N(\mathbf{0}, \mathbf{I}_n)$ dir. Böylece,

$$\begin{aligned}
\mathbf{y}'\mathbf{A}\mathbf{y} &= \mathbf{y}'\mathbf{P}'\begin{bmatrix} \mathbf{I}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\mathbf{P}_y \\
&= \mathbf{z}'\begin{bmatrix} \mathbf{I}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\mathbf{z} \\
&= z_1^2 + \ldots + z_r^2 \\
&= \chi_r^2
\end{aligned}$$

elde ederiz.

Tersine olarak $\mathbf{y}'\mathbf{A}\mathbf{y} \sim \chi_r^2$ olduğunu farz edelim. Bu takdirde,

$$\mathbf{A} = \mathbf{P}'k\ddot{o}\c{s}eg(\lambda_1, \ldots, \lambda_n)\mathbf{P}$$

olacak şekilde bir ortogonal **P** matrisi vardır. Burada $\lambda_1,\ldots,\lambda_n$'ler **A**'nın özdeğerleridir. Yine, **z=Py** olsun. Bundan dolayı $\mathbf{z} \sim N(\mathbf{0},\mathbf{I_n})$ dir. $z_j^2 \sim \chi_1^2$ olduğundan, **y**'nin karakteristik fonksiyonu,

$$\begin{aligned}
\phi(t) &= E(\exp(it\mathbf{y}'\mathbf{A}\mathbf{y})) \\
&= E(\exp(it\sum_{j=1}^{n}\lambda_j z_j^2)) \\
&= \prod_{j=1}^{n} E(\exp(it\lambda_j z_j^2)) \\
&= \prod_{j=1}^{n} (1-2it\lambda_j)^{-\frac{1}{2}}
\end{aligned} \qquad (2.5.3.1)$$

ile verilir. Bununla beraber $\mathbf{y}'\mathbf{A}\mathbf{y} \sim \chi_r^2$ olduğundan onun karakteristik fonksiyonu

$$\phi(t) = (1-2it)^{-r/2} \qquad (2.5.3.2)$$

dir. (2.5.3.1) ve (2.5.3.2) bağıntılarını eşitleyerek her t için,

$$(1-2it)^r = \prod_{j=1}^{n}(1-2it\lambda_j) \qquad (2.5.3.3)$$

elde ederiz. (2.5.3.3) bağıntısının sol yanı t'ye göre bir polinomdur. Tümü $\frac{1}{2i}$'ye eşit olan r tane köke sahiptir. Bu nedenle sağ yan da aynı köklere sahip olmalıdır. λ_i'lerin r tanesi 1'e, geri kalanı 0'a eşit olduğunda bu kesin olarak mümkün olabilir. Bu nedenle, **A**, r ranklı idempotent bir matristir.

Teorem 2.5.3.2. $\mathbf{y} \sim N(\mathbf{0},\mathbf{I_n})$ olsun ve $\mathbf{A}_1$, $\mathbf{A}_2$ simetrik idempotent matrisler olsunlar. Bu takdirde $\mathbf{y}'\mathbf{A}_1\mathbf{y}$, $\mathbf{y}'\mathbf{A}_2\mathbf{y}$ karesel formlarının bağımsız olmaları için gerek ve yeter şart $\mathbf{A}_1.\mathbf{A}_2 = \mathbf{0}$ olmasıdır.

İspat: $\mathbf{A}_1.\mathbf{A}_2 = \mathbf{0}$ olduğunu farz edelim. Bu takdirde Teorem 2.5.2.3'e göre $\mathbf{A}_1\mathbf{y}$, $\mathbf{A}_2\mathbf{y}$ bağımsızdırlar. Bağımsız rastgele değişkenlerin fonksiyonları olduklarından bu nedenle;

$$\mathbf{y}'\mathbf{A}_1\mathbf{y} = (\mathbf{A}_1\mathbf{y})'(\mathbf{A}_1\mathbf{y}), \quad \mathbf{y}'\mathbf{A}_2\mathbf{y} = (\mathbf{A}_2\mathbf{y})'(\mathbf{A}_2\mathbf{y})$$

bağımsızdırlar.

Tersine olarak $\mathbf{y}'\mathbf{A}_1\mathbf{y}$, $\mathbf{y}'\mathbf{A}_2\mathbf{y}$ bağımsız olsunlar. Teorem 2.5.3.1'e göre $\mathbf{y}'\mathbf{A}_1\mathbf{y}$, $\mathbf{y}'\mathbf{A}_2\mathbf{y}$ ki-kare değişkenlerine sahiptirler ve bu nedenle,

$$\mathbf{y}'\mathbf{A}_1\mathbf{y} + \mathbf{y}'\mathbf{A}_2\mathbf{y} = \mathbf{y}'(\mathbf{A}_1 + \mathbf{A}_2)\mathbf{y}$$

ki-kare değişkenine sahip olmalıdır. Yine Teorem 2.5.3.1'e göre $\mathbf{A}_1 + \mathbf{A}_2$ idempotent dir. olur. Bu nedenle,

$$\begin{aligned}\mathbf{A}_1 + \mathbf{A}_2 &= (\mathbf{A}_1 + \mathbf{A}_2)^2 \\ &= \mathbf{A}_1^2 + \mathbf{A}_2^2 + \mathbf{A}_1\mathbf{A}_2 + \mathbf{A}_2\mathbf{A}_1 \\ &= \mathbf{A}_1 + \mathbf{A}_2 + \mathbf{A}_1\mathbf{A}_2 + \mathbf{A}_2\mathbf{A}_1\end{aligned}$$

dir. Buradan,

$$\mathbf{A}_1\mathbf{A}_2 + \mathbf{A}_2\mathbf{A}_1 = 0$$

dır. Bu eşitliği $\mathbf{A}_2$ ile sondan çarparak

$$\mathbf{A}_1\mathbf{A}_2 + \mathbf{A}_2\mathbf{A}_1\mathbf{A}_2 = 0 \qquad (2.5.3.4)$$

bağıntısı elde edilir. (2.5.3.4) bağıntısını $\mathbf{A}_2$ ile önden çarparak

$$\mathbf{A}_2\mathbf{A}_1\mathbf{A}_2 + \mathbf{A}_2^2\mathbf{A}_1\mathbf{A}_2 = 2\mathbf{A}_2\mathbf{A}_1\mathbf{A}_2 = 0$$

elde ederiz. Bu nedenle, $\mathbf{A}_2\mathbf{A}_1\mathbf{A}_2 = 0$ dır. Bunu (2.5.3.4) bağıntısında yerine koyarak $\mathbf{A}_1\mathbf{A}_2 = 0$ elde ederiz.

Teorem 2.5.3.3. $\mathbf{y} \sim N(\mathbf{0}, \mathbf{I}_n)$ olsun. $\mathbf{A}$ bir simetrik idempotent matris olsun ve $\ell \in R^n$, sıfırdan farklı bir vektör olsun. Bu takdirde $\mathbf{y}'\mathbf{A}\mathbf{y}$ ve $\ell'\mathbf{y}$'nin bağımsız olmaları için gerek ve yeter şart $\mathbf{A}\ell = 0$ olmasıdır.

İspat: Genelliği kaybetmeksizin, $\|\ell\| = 1$ olduğunu kabul edelim. Bu takdirde $\mathbf{B} = \ell\ell'$ bir simetrik idempotent matristir.

İlk olarak $\mathbf{y}'\mathbf{A}\mathbf{y}$ ve $\ell'\mathbf{y}$' nin bağımsız olduklarını farz edelim. Bu takdirde önceden olduğu gibi, bağımsız rastgele (ölçülebilir) değişkenlerin fonksiyonlarının bağımsız oldukları gerçeğini kullanarak $\mathbf{y}'\mathbf{A}\mathbf{y}$ ve $\mathbf{y}'\mathbf{B}\mathbf{y}$'nin bağımsız olduklarını görürüz.

Teorem 2.5.3.2'den **AB=0** olduğu görülür ve bu takdirde $\mathbf{A}\ell = 0$ olduğunu göstermek kolaydır. Tersine olarak, $\mathbf{A}\ell = 0$ ise, bu takdirde, Teorem 2.5.2.3'den $\mathbf{A}y$ ve $\ell'\mathbf{y}$ bağımsızdırlar. Bundan dolayı $\mathbf{y}'\mathbf{A}\mathbf{y} = (\mathbf{A}\mathbf{y})'(\mathbf{A}\mathbf{y})$ ve $\ell'\mathbf{y}$ bağımsızdırlar. Bu ispatı tamamlar. Şimdi, Cochran teoreminin bir uyarlamasının bir matris teorisi formüllemesini ifade ve ispat edeceğiz.

Teorem 2.5.3.4. $\sum_{i=1}^{k} \mathbf{A}_i = \mathbf{I}$ olmak üzere, $\mathbf{A}_1, \ldots \mathbf{A}_k$ nxn matrisler olsun. Bu takdirde aşağıdaki şartlar denktirler.

(i) $\sum_{i=1}^{k} \mathrm{R}(\mathbf{A}_i) = n$

(ii) $\mathbf{A}_i^2 = \mathbf{A}_i, \quad i = 1, \ldots, k$

(iii) $\mathbf{A}_i\mathbf{A}_j = \mathbf{0}, \quad i \neq j$

İspat: (i) $\Rightarrow$(ii): $\mathbf{A}_i = \mathbf{B}_i\mathbf{C}_i$, $i=1,\ldots,k$, bir rank ayrışımı olsun. Bu takdirde,

$$\mathbf{B}_1\mathbf{C}_1 + \ldots + \mathbf{B}_k\mathbf{C}_k = \mathbf{I}$$

ve bu nedenle,

$$[\mathbf{B}_1 \ldots \mathbf{B}_k] = \begin{bmatrix} \mathbf{C}_1 \\ \cdot \\ \cdot \\ \cdot \\ \mathbf{C}_k \end{bmatrix} = \mathbf{I}$$

dır. $\sum_{i=1}^{k} \mathrm{R}(\mathbf{A}_i) = n$ olduğundan $[\mathbf{B}_1 \ldots \mathbf{B}_k]$ bir kare matristir ve bu nedenle,

$$\begin{bmatrix} \mathbf{C}_1 \\ \vdots \\ \mathbf{C}_k \end{bmatrix} [\mathbf{B}_1 \ldots \mathbf{B}_k] = \mathbf{I}$$

dır. Böylece $\mathbf{C}_i\mathbf{B}_j = \mathbf{0}$, $i \neq j$ dir. $i \neq j$ için,

$$\mathbf{A}_i\mathbf{A}_j = \mathbf{B}_i\mathbf{C}_i\mathbf{B}_j\mathbf{C}_j = \mathbf{0}$$

olduğu görülür.

(iii) $\Rightarrow$ (ii): $\sum_{i=1}^{k} \mathbf{A}_i = \mathbf{I}$ olduğundan,

$$\mathbf{A}_j\left(\sum_{i=1}^{k} \mathbf{A}_i\right) = \mathbf{A}_j, \qquad j=1,...,k,$$

dir. $\mathbf{A}_j^2 = \mathbf{A}_j$ olduğu görülür.

(ii) $\Rightarrow$(i): $\mathbf{A}_i$ idempotent olduğundan, $R(\mathbf{A}_i) = iz\mathbf{A}_i$ dir. Şimdi,

$$\sum_{i=1}^{k} R(\mathbf{A}_i) = \sum_{i=1}^{k} iz\mathbf{A}_i = iz\left(\sum_{i=1}^{k} \mathbf{A}_i\right) = n$$

dir. Bu da ispatı tamamlar.

2.5.4. Bir-Yönlü ve İki-Yönlü Sınıflandırmalar

Tanım 2.5.4.1. Gözlemlerin vektörünün $\mathbf{y} \sim N(\mathbf{0},\mathbf{I}_n)$ olduğunu farz edelim. $\mathbf{y}'\mathbf{y} = \sum_{i=1}^{n} y_i^2$ 'ye ham kareler toplamı denir. Eğer, $\mathbf{A}_i$'ler simetrik ve $\mathbf{A}_i\mathbf{A}_j = \mathbf{0}$, $i \neq j$ olmak üzere $\mathbf{y}'\mathbf{y}$ 'yi

$$\mathbf{y}'\mathbf{y} = \mathbf{y}'\mathbf{A}_1\mathbf{y} + ... + \mathbf{y}'\mathbf{A}_k\mathbf{y}$$

olarak ayrıştırabilirsek, bu takdirde Cochran teoremine göre $\mathbf{y}'\mathbf{A}_i\mathbf{y}$ lerin bağımsız ki-kare rastgele değişkenleri oldukları sonucunu çıkarabiliriz. Bu takdirde serbestlik dereceleri $\mathbf{A}_i$'nin ranklarıyla verilirler.

İlk olarak, daha önce incelenen bir-yönlü sınıflandırma modeline bir uygulamayı açıklayalım. Model;

$$y_{ij} = \mu + \alpha_i + \varepsilon_{ij}, \qquad i=1,...,k, \qquad j=1,..., n_i$$

dir. Şimdi burada ε_{ij}'lerin $N(0,\sigma^2)$ dağılımlı bağımsız rastgele değişkenler olduklarını kabul edeceğiz.

$$H_0 : \alpha_1 = ... = \alpha_k$$

hipotezini test etmek istediğimizi farz edelim.

$$z_{ij} = \frac{y_{ij} - \mu - \alpha}{\sigma},$$

olsun. Eğer H_0 doğru ise, z_{ij}'ler standart normal dağılıma sahiptirler ve burada α, $\alpha_1, \ldots, \alpha_k$' ların ortak değerini gösterir. **z**,

$$(z_{11}, \ldots, z_{1n_1}; z_{21}, \ldots, z_{2n_2}; z_{k1}, \ldots, z_{kn_k})$$

vektörü olsun ve $n = \sum_{i=1}^{k} n_i$ olsun. Nokta notasyonunu kullanacağız. Böylece,

$$z_{i\cdot} = \sum_{j=1}^{n_i} z_{ij}, \quad \bar{z}_{i\cdot} = \frac{z_{i\cdot}}{n_i}$$

dir ve benzer bir notasyon ikiden fazla alt indis olduğunda da kullanılır. Çapraz çarpım terimleri 0'a eşit olduğundan,

$$\sum_{i=1}^{k}\sum_{j=1}^{n_i} z_{ij}^2 = \sum_{i=1}^{k}\sum_{j=1}^{n_i} (z_{ij} - \bar{z}_{i\cdot} + \bar{z}_{i\cdot} - \bar{z}_{\cdot\cdot} + \bar{z}_{\cdot\cdot})^2 \tag{2.5.4.1}$$

$$= \sum_{i=1}^{k}\sum_{j=1}^{n_i} (z_{ij} - \bar{z}_{i\cdot})^2 + \sum_{i=1}^{k} n_i (\bar{z}_{i\cdot} - \bar{z}_{\cdot\cdot})^2 + n\bar{z}_{\cdot\cdot}^2 \tag{2.5.4.2}$$

özdeşliğini elde ederiz. Örneğin;

$$\sum_{i=1}^{k}\sum_{j=1}^{n_i} (z_{ij} - \bar{z}_{i\cdot})(\bar{z}_{i\cdot} - \bar{z}_{\cdot\cdot}) = \sum_{i=1}^{k} (\bar{z}_{i\cdot} - \bar{z}_{\cdot\cdot}) \sum_{j=1}^{n_i} (z_{ij} - \bar{z}_{i\cdot}) = 0 \tag{2.5.4.3}$$

dır.

$$\mathbf{z}'\mathbf{A}_1\mathbf{z}, \qquad \mathbf{z}'\mathbf{A}_2\mathbf{z}, \qquad \mathbf{z}'\mathbf{A}_3\mathbf{z},$$

sırasıyla (2.5.4.2) bağıntısındaki üç forma eşit kuadratik formlar olacak şekilde $\mathbf{A}_1$, $\mathbf{A}_2$, $\mathbf{A}_3$ simetrik matrisler olsun. Her bir form bir kareler toplamı olduğundan $\mathbf{A}_i$' ler gerçekten pozitif yarı-tanımlıdır. Bir moment düşüncesi (2.5.4.3) bağıntısının $\mathbf{A}_1\mathbf{A}_2 = 0$ ifade ettiğini gösterecek. Aynı şekilde,

$$\mathbf{A}_1\mathbf{A}_3 = \mathbf{A}_2\mathbf{A}_3 = 0$$

dır. Cochran teoremine göre $\mathbf{z}'\mathbf{A}_1\mathbf{z}$, $\mathbf{z}'\mathbf{A}_2\mathbf{z}$ 'nin bağımsız ki-kare değişkenleri olduğu sonucunu çıkarırız. Geriye serbestlik derecelerini bulmak kalır. Bunlar sırasıyla, $R(\mathbf{A_1})$, $R(\mathbf{A_2})$ dir. $\mathbf{A}_1, \mathbf{A}_2$ idempotent(yine Cochran teoreminden görülür) olduğundan iz $\mathbf{A}_1$, iz $\mathbf{A}_2$ 'yi bulmak yeterlidir.

i $1,...,k$, j $1,...,$ n_i için, $\mathbf{z}^{ij}$, (i,j) koordinatında 1 ve diğer yerlerde 0'lar olan üzere n mertebeli sütun vektör olsun. Bu takdirde,

$$\mathrm{iz}\mathbf{A_1} = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(\mathbf{z^{ij}})'\mathbf{A_1}\mathbf{z^{ij}}$$

dir. Bununla beraber,

$$\begin{aligned}(\mathbf{z^{ij}})'\mathbf{A_1}\mathbf{z^{ij}} &= \left(1-\frac{1}{n_i}\right)^2 + \frac{n_i-1}{n_i^2}\\ &= \frac{n_i-1}{n_i^2}xn_i = \frac{n_i-1}{n_i}\end{aligned}$$

dir. Bu nedenle,

$$\mathrm{iz}\mathbf{A}_1 = \sum_{i=1}^{k} n_i \mathrm{x} \frac{n_i-1}{n_i} = n-k$$

dir. Aynı şekilde $\mathrm{iz}\mathbf{A_2} = k-1$ 'dir. Bir başka şekilde her bir i için, $\sum_{j\ 1}^{n_i}(z_{ij}-\overline{z}_{i.})^2$ 'nin n_i 1 serbestlik dereceli ki-kare olarak dağıldığına ve bağımsız ki-karelerin bir toplamı olan, $\sum_{i\ 1}^{k}\sum_{j\ 1}^{n_i}(z_{ij}-\overline{z}_{i.})^2$ nin de,

$$\sum_{i\ 1}^{k}(n_i-1) = n-k$$

serbestlik dereceli ki-kare olarak dağıldığına dikkat edelim. Aynı şekilde $z'A_2z$ 'de k-1 serbestlik dereceli ki-kare dağılımına sahiptir. Böylece, H_0 hipotezi altında,

$$\frac{\sum_{i=1}^{k} n_i(\bar{z}_{i.}-\bar{z}_{..})^2/k-1}{\sum_{i=1}^{k}\sum_{j=1}^{n_i}(z_{ij}-\bar{z}_{i.})^2/n-k} \sim F(k-1,n-k)$$

dir. y_{ij}'lere bağlı olarak,

$$\frac{\sum_{i=1}^{k} n_i(\bar{y}_{i.}-\bar{y}_{i.})^2/k-1}{\sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{ij}-\bar{y}_{i.})^2/n-k} \sim F(k-1,n-k)$$

olarak yazabiliriz ve bu oran H_0 hipotezini test etmek için kullanılabilir. Bu test istatistiği sezgisel olarak doğrulanabilir. Eğer her bir kitledeki farkı karşılaştırmada kitleler arasındaki fark büyükse bu takdirde, istatistik büyük olacak ve H_0 reddedilecektir. Şimdi etkileşimsiz bir yönlü sınıflandırmayı tanımlayacağız. Aynı zamanda bu istatistiğin belirli optimal özelliklere sahip olduğunu göstereceğiz.

Tanım 2.5.4.2. Etkileşimsiz Bir Yönlü Sınıflandırma: Biri a seviyesinde ve diğeri b seviyesinde iki faktörün(etmen) var olduğunu farz edelim. Birinci faktörün bir seviyesinin ve ikinci faktörün bir seviyesinin her bir kombinasyonu için bir gözleme sahibiz. ε_{ij} ler özdeş bağımsız $N(0,\sigma^2)$ dağılımlı rastgele değişkenler olmak üzere, model,

$$y_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij}, \qquad i-1,\ldots,a, \qquad j-1,\ldots,b,$$

dir. Burada α_i, birinci faktörün i-yinci seviyesinin etkisini ve β_j ikinci faktörün j-yinci seviyesinin etkisini gösterir. $H_0:\alpha_1=\ldots=\alpha_a$ hipotezini test etmek istediğimizi varsayalım. Bundan sonraki inceleme H_0' ın doğru olması varsayımı altındadır. α, $\alpha_1,\ldots,\alpha_a$'nın ortak değeri olsun.

$$z_{ij} = \frac{y_{ij}-\mu-\alpha-\beta_j}{\sigma}$$

olsun. Bu takdirde z_{ij}'ler özdeş bağımsız $N(0,1)$ dağılımlıdırlar. z,

$$(z_{11},\ldots,z_{1b};z_{21},\ldots,z_{2b};z_{a1},\ldots,z_{ab})$$

vektör olsun.

$$z_{ij} = (z_{ij} - \bar{z}_{i.} - \bar{z}_{.j} + \bar{z}_{..}) + (\bar{z}_{i.} - \bar{z}_{..}) + (\bar{z}_{.j} - \bar{z}_{..}) + \bar{z}_{..}$$

elde ederiz ve çapraz çarpım terimleri sıfır olduğundan önceki gibi,

$$\sum_{i=1}^{a}\sum_{j=1}^{b} z_{ij}^2 = \sum_{i=1}^{a}\sum_{j=1}^{b} (z_{ij} - \bar{z}_{i.} - \bar{z}_{.j} + \bar{z}_{..})^2 + b\sum_{i=1}^{a} (\bar{z}_{i.} - \bar{z}_{..})^2 + a\sum_{j=1}^{b} (\bar{z}_{.j} - \bar{z}_{..})^2 + ab\bar{z}_{..}^2$$

dir. Bu nedenle, $\mathbf{A_i A_j} = \mathbf{0}$, $i \neq j$ olacak şekilde, $\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3, \mathbf{A}_4$ simetrik (gerçekten pozitif yarı-tanımlı) matrisler olmak üzere,

$$\mathbf{z'z} = \mathbf{z'A_1z} + \mathbf{z'A_2z} + \mathbf{z'A_3z} + \mathbf{z'A_4z}$$

yazabiliriz. Cochran teoremine göre $\mathbf{z'A_iz}$'ler bağımsız ki-kare rastgele değişkenleridir. R($\mathbf{A}_1$)=(a-1)(b-1), R($\mathbf{A}_2$)=a-1, olduğunun kontrolünü bir alıştırma olarak bırakıyoruz.

$$\sum_{i=1}^{a}\sum_{j=1}^{b} (z_{ij} - \bar{z}_{i.} - \bar{z}_{.j} + \bar{z}_{..})^2 = \frac{1}{\sigma^2}\sum_{i=1}^{a}\sum_{j=1}^{b} (y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..})^2$$

ve

$$\sum_{i=1}^{a} (\bar{z}_{i.} - \bar{z}_{..})^2 = \frac{1}{\sigma^2}\sum_{i=1}^{a} (\bar{y}_{i.} - \bar{y})^2$$

olduklarına dikkat edelim. Bu nedenle,

$$\frac{b\sum_{i=1}^{a} (\bar{y}_{i.} - \bar{y}_{..})^2 / (a-1)}{\sum_{i=1}^{a}\sum_{j=1}^{b} (y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{i.})^2 / (a-1)(b-1)},$$

$(a-1,\ (a-1)(b-1))$ serbestlik dereceli F dağılımına sahiptir ve H_0 ı test etmek için kullanılabilir. Aynı şekilde $H_0 : \beta_1 = ... = \beta_b$'nin bir testi kurulur. Eğer her bir seviye kombinasyonları başına, gözlemlerin sayısını birden daha çok fakat eşit alırsak bu takdirde, n, her bir seviye kombinasyonu başına gözlemlerin sayısını göstermek üzere model,

$$y_{ijk} = \mu + \alpha_i + \beta_j + \varepsilon_{ijk}, \quad i = 1,...,a, \quad j = 1,...,b, \quad k = 1,...,n$$

dir. Bu durumda analiz benzerdir ve $H_0 : \alpha_1 = ... = \alpha_a$ hipotezi altında,

$$\frac{bn\sum_{i=1}^{a}(\bar{y}_{i..} - \bar{y}_{...})^2 / (a-1)}{n\sum_{i=1}^{a}\sum_{j=1}^{b}(y_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{...})^2 / (abn - a - b + 1)}$$

oranının $(a-1, abn-a-b+1)$ serbestlik dereceli F dağılımına sahip olduğu gösterilebilir. Eğer k=1,..., n_{ij} ise, bu takdirde, genel olarak istatistik kapalı bir biçimde ifade edilemez. Bununla beraber eğer n_{ij},

$$n_{ij} = \frac{n_{i.}n_{.j}}{n_{..}}$$

bağıntısını sağlarsa, bu takdirde, F-istatistiği n_{ij} lerin eşit olduğu duruma benzer bir şekilde elde edilebilir.

Tanım 2.5.4.3. Etkileşimli İki-Yönlü Sınıflandırma: $n > 1$ ve ε_{ijk}'ların her biri $N(0,\sigma^2)$ dağılımına sahip bağımsız rastgele değişkenler olmak üzere,

$$y_{ij} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \varepsilon_{ij}, \quad i=1,...,a, \quad j=1,...,b, \quad k=1,...,n$$

modelini göz önüne alalım. Bu takdirde, γ_{ij}'lerin tümünün eşit oldukları hipotezi altında,

$$\frac{n\sum_{i=1}^{a}\sum_{j=1}^{b}(\bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{...})^2}{\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(y_{ijk} - \bar{y}_{ij.})^2} x \frac{ab(n-1)}{(a-1)(b-1)}$$

istatistiğinin $F((a-1)(b-1), ab(n-1))$ dağılımına sahip olduğu gösterilebilir

2.5.5. Genel Lineer Hipotez

Şimdi lineer modelimize bir normallik varsayımı getireceğiz ve **y**, bir *n*x*1* vektör, **X**, *nxp* matris ve **β** bir *px1* vektör olmak üzere,

$$\mathbf{y} \sim N(\mathbf{X\beta}, \sigma^2\mathbf{I_n})$$

olduğunu kabul edeceğiz. R(**X**)=*r* olsun.

$$\mathrm{HKT} = \min_{\boldsymbol{\beta}}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$$

olduğunu ve

$$\hat{\boldsymbol{\beta}} = (\mathbf{X'X})^{-}\mathbf{X'y}$$

da minimumuna ulaştığını görmüştük.

Teorem 2.5.5.1. $\dfrac{\mathrm{HKT}}{\sigma^2} \sim \chi^2_{n-r}$ dır.

İspat: Önceki gibi, $\mathbf{P} = \mathbf{I} - \mathbf{X}(\mathbf{X'X})^{-}\mathbf{X}'$ olsun. $\mathbf{PX} = 0$ olduğundan Teorem 2.3.2'nin ispatından $\mathrm{HKT} = \mathbf{y'Py} = (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'\mathbf{P}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$ dir. Bu nedenle Teorem 2.5.3.1'den HKT/σ^2 nin χ^2 dağılımına sahip olduğu görülür. Serbestlik derecesi **R(P)** dir ve bunun (*n-r*) olduğu görülür.

Şimdi $H: \mathbf{L}\boldsymbol{\beta} = \mathbf{z}$ hipotezini düşünelim. Önceki gibi $\Re(\mathbf{L}) \subset \Re(\mathbf{X})$ ve $\mathbf{L}\boldsymbol{\beta} = \mathbf{z}$ denkleminin tutarlı olduğu varsayımı yapacağız. Kısıtlamalara bağlı tahmin bölümünü izleyerek, $\mathbf{L} = \mathbf{WX'X}$, $\mathbf{WX'} = \mathbf{T}$ olsun. Bu takdirde,

$$\tilde{\boldsymbol{\beta}} = \hat{\boldsymbol{\beta}} - \mathbf{X'XL'}(\mathbf{TT'})^{-}(\mathbf{L}\hat{\boldsymbol{\beta}} - \mathbf{z})$$

olmak üzere,

$$\mathrm{HKT}_H = \min_{\boldsymbol{\beta}:\mathbf{L}\boldsymbol{\beta}=\mathbf{z}}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$$

$\tilde{\boldsymbol{\beta}}$ de minimumunu alır. Böylece,

$$\begin{aligned} \mathbf{X}\tilde{\boldsymbol{\beta}} &= (\mathbf{1} - \mathbf{P})\mathbf{y} - \mathbf{T'}(\mathbf{TT'})^{-}(\mathbf{Ty} - \mathbf{z}) \\ &= (\mathbf{1} - \mathbf{P})\mathbf{y} - \mathbf{T'}(\mathbf{TT'})^{-}(\mathbf{Ty} - \mathbf{TX}\boldsymbol{\beta} + \mathbf{TX}\boldsymbol{\beta} - \mathbf{z}) \end{aligned} \qquad (2.5.5.1)$$

dir. Eğer H doğru ise, bu takdirde, $\mathbf{TX}\boldsymbol{\beta} = \mathbf{L}\boldsymbol{\beta} = \mathbf{z}$ dır ve böylece $\mathbf{U} = \mathbf{T'}(\mathbf{TT'})^{-}\mathbf{T}$ olmak üzere (2.5.5.1) bağıntısına göre,

$$\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}} = \mathbf{Py} + \mathbf{U}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$$

dır. Bu nedenle $\mathbf{PU} = \mathbf{0}$ iken,

$$\mathrm{HKT}_H = (\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}})'(\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}}) = \mathbf{y'Py} + (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'\mathbf{U}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$$

dır. **U** idempotent olduğundan,

$$\frac{\text{HKT}_H - \text{HKT}}{\sigma^2} \sim \chi^2_{R(U)} \tag{2.5.5.2}$$

sonucunu çıkarırız. Aynı zamanda $\mathbf{PU}=\mathbf{0}$ olduğundan HKT ve $\text{HKT}_H - \text{HKT}$ bağımsız olarak dağılırlar.

$$\begin{aligned} \text{R}(\mathbf{U}) &= \text{R}(\mathbf{T}'(\mathbf{TT}')^- \mathbf{T}) = \text{iz}\mathbf{T}'(\mathbf{TT}')^- \mathbf{T} \\ &= \text{iz}(\mathbf{TT}')^- \mathbf{TT}' = \text{R}((\mathbf{TT}')^- \mathbf{TT}') \\ &= \text{R}(\mathbf{TT}') = \text{R}(\mathbf{T}) \end{aligned}$$

yazarız. Teorem 2.4.2'den $\text{R}(\mathbf{U}) = \text{R}(\mathbf{L})$ olduğu görülür. Böylece,

$$\frac{(\text{HKT}_H - \text{HKT}) / \text{R}(\mathbf{L})}{\text{HKT} / (n-r)} \sim F(\text{R}(\mathbf{L}), n-r)$$

sonucunu elde ederiz ve bu oran H hipotezini test etmek için kullanılabilir.

2.5.6. Kuadratik Formların Maksimum ve Minimumları

Teorem 2.5.6.1. **A** bir nxn simetrik matris olsun. Bu takdirde,

$$\max_{\mathbf{x}\neq\mathbf{0}} \frac{\mathbf{x}'\mathbf{A}\mathbf{x}}{\mathbf{x}'\mathbf{x}} = \lambda_1$$

dir ve maksimuma **A**'nın λ_1'e karşılık gelen herhangi bir özvektöründe ulaşır.

İspat: **P** ortogonal ve alışıldığı gibi, $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$ ler **A** nın özdeğerleri olmak üzere, $\mathbf{A} = \mathbf{P}'\textit{köşeg}(\lambda_1, \ldots, \lambda_n)\mathbf{P}$ yazarız (çünkü **A** reel simetrik matristir) ve bu nedenle özdeğerleri de reeldir. Bu takdirde, $\mathbf{x} \neq \mathbf{0}$ (sıfırdan farklı **x** vektörü) için,

$$\begin{aligned} \frac{\mathbf{x}'\mathbf{A}\mathbf{x}}{\mathbf{x}'\mathbf{x}} &= \frac{\mathbf{y}'\textit{köşg}(\lambda_1, \ldots, \lambda_n)\mathbf{y}}{\mathbf{y}'\mathbf{y}}, \qquad (\mathbf{y} = \mathbf{Px}) \\ &= \frac{\lambda_1 y_1^2 + \ldots + \lambda_n y_n^2}{y_1^2 + \ldots + y_n^2} \\ &\leq \frac{\lambda_1 y_1^2 + \ldots + \lambda_1 y_n^2}{y_1^2 + \ldots + y_n^2} = \lambda_1 \end{aligned}$$

dir. Böylece,

$$\max_{\mathbf{x}\neq\mathbf{0}} \frac{\mathbf{x}'\mathbf{A}\mathbf{x}}{\mathbf{x}'\mathbf{x}} \leq \lambda_1$$

dir. Açık olarak eğer $\mathbf{x}$, λ_1' e karşılık gelen bir özvektör ise bu takdirde,

$$\frac{\mathbf{x'Ax}}{\mathbf{x'x}} = \lambda_1$$

dir ve sonuç ispatlanmıştır.

Teorem 2.5.6.2. $\mathbf{A}$ ve $\mathbf{B}$ simetrik ve $\mathbf{B}$ pozitif tanımlı olmak üzere nxn matrisler olsunlar. Bu takdirde, μ, $\mathbf{AB}^{-1}$ in en büyük özdeğeri olmak üzere,

$$\max_{\mathbf{x} \neq 0} \frac{\mathbf{x'Ax}}{\mathbf{x'Bx}} = \mu$$

dür.

İspat:

$$\max_{\mathbf{x} \neq 0} \frac{\mathbf{x'Ax}}{\mathbf{x'Bx}} = \max_{\mathbf{x} \neq 0} \frac{\mathbf{x'Ax}}{(\mathbf{x'B}^{1/2})(\mathbf{B}^{1/2}\mathbf{x})}$$
$$= \max_{\mathbf{y} \neq 0} \frac{\mathbf{y'B}^{-1/2}\mathbf{AB}^{-1/2}\mathbf{y}}{\mathbf{y'y}}$$

dir. Teorem 2.5.6.1'e göre bu son ifade $\mathbf{B}^{-1/2}\mathbf{AB}^{-1/2}$'nin en büyük özdeğerine eşittir. Bununla beraber $\mathbf{B}^{-1/2}\mathbf{AB}^{-1/2}$ ve $\mathbf{AB}^{-1}$ aynı özdeğerlere sahiptir ve ispat tamamdır. Teorem 2.5.6.2'deki maksimumuna $\mathbf{B}^{-1/2}\mathbf{AB}^{-1/2}$'nın μ özdeğerine karşılık gelen herhangi bir özvektöründe ulaşılır.

Teorem 2.5.6.3. $\mathbf{B}$ bir pozitif tanımlı nxn matris olsun ve $\mathbf{y} \in R^n$ olsun. Bu takdirde,

$$\max_{\mathbf{x} \neq 0} \frac{(\mathbf{x'y})^2}{\mathbf{x'Bx}} = \mathbf{y'B}^{-1}\mathbf{y} \text{ dir.}$$

İspat: $$\max_{\mathbf{x} \neq 0} \frac{(\mathbf{x'y})^2}{\mathbf{x'Bx}} = \max_{\mathbf{x} \neq 0} \frac{\mathbf{x'yy'x}}{\mathbf{x'Bx}}$$

yazarız. Bu ise, Teorem 2.5.6.1'e göre $\mathbf{yy'B}^{-1}$ in en büyük özdeğeridir. Yine $\mathbf{yy'B}^{-1}$ ve $\mathbf{B}^{-1/2}\mathbf{AB}^{-1/2}$ nin özdeğerleri eşittirler. Son matris simetrik ve rankı 1 olduğundan katlılık içeren sıfır olmayan yalnız bir özdeğere sahiptir. Bu özdeğer;

$$\text{iz}\,\mathbf{B}^{-1/2}\mathbf{yy'B}^{-1/2} = \text{iz}\,\mathbf{y'B}^{-1}\mathbf{y} = \mathbf{y'B}^{-1}\mathbf{y}$$

'ye eşittir ve ispat tamamlanır.

Teorem 2.5.6.3'ü herhangi bir pozitif tanımlı **B** matrisi için,

$$(\mathbf{x}'\mathbf{y})^2 \le (\mathbf{x}'\mathbf{B}\mathbf{x})(\mathbf{y}'\mathbf{B}^{-1}\mathbf{y}) \qquad (2.5.6.1)$$

olarak yazabiliriz.

(2.5.6.1) bağıntısının değişik bir ispatı,

$$\left(\sum u_i v_i\right)^2 \le \left(\sum u_i^2\right)\left(\sum v_i^2\right)$$

Cauchy-Schwarz eşitsizliğini kullanarak verilebilir.

2.5.7. Çoklu Korelasyon ve Regresyon modelleri

$$(y, x_1, \ldots, x_p)'$$

($p+1$ mertebeli) rastgele vektörünün,

$$\mathbf{V} = \begin{bmatrix} \sigma^2 & \mathbf{u}' \\ \mathbf{u} & \Sigma \end{bmatrix} \qquad (2.5.7.1)$$

dağılım matrisine sahip olduğunu varsayalım. Burada Σ, p-inci mertebeden pozitif-tanımlıdır.

$$\boldsymbol{\alpha}'\mathbf{x} = \alpha_1 x_1 + \ldots + \alpha_p x_p, \qquad \boldsymbol{\alpha} \ne \mathbf{0}$$

lineer kombinasyonunun y ile maksimum korelasyona sahip olduğunu bulmak istiyoruz. Maksimum değere y ve $x_1, \ldots, x_p$ arasındaki çoklu korelasyon katsayısı denir ve $r_{y(x_1,\ldots,x_p)}$ ile gösterilir. Bu nedenle Teorem 2.5.6.3'e göre,

$$\begin{aligned}
r^2_{y(x_1\ldots x_p)} &= \max_{\boldsymbol{\alpha} \ne \mathbf{0}} \{korelasyon(y, \boldsymbol{\alpha}'\mathbf{x})\}^2 \\
&= \max_{\boldsymbol{\alpha} \ne \mathbf{0}} \frac{(kov(y, \boldsymbol{\alpha}'\mathbf{x}))^2}{\text{var}(y)\,\text{var}(\boldsymbol{\alpha}'\mathbf{x})} \\
&= \max_{\boldsymbol{\alpha} \ne \mathbf{0}} \frac{(\boldsymbol{\alpha}'\mathbf{u})^2}{\sigma^2 \boldsymbol{\alpha}'\Sigma\boldsymbol{\alpha}} \\
&= \frac{\mathbf{u}'\Sigma^{-1}\mathbf{u}}{\sigma^2}
\end{aligned}$$

dir. Maksimuma $\boldsymbol{\alpha} = \Sigma^{-1}\mathbf{u}$ da ulaşır. $r^2_{y(x_1\ldots x_p)}$ için başka bir ifadeyi aşağıdaki gibi elde ederiz.

$$\frac{1}{z_{11}} = \sigma^2 - \mathbf{u}'\Sigma^{-1}\mathbf{u}$$

dur. Bu nedenle,

$$r^2_{y(x_1,\ldots,x_p)} = 1 - \frac{1}{\sigma^2 z_{11}}$$

dir. $(y, x_1, \ldots, x_p)$ vektörünün $(\tau, \mu')'$ ortalama vektörlü ve (2.5.7.1) bağıntısındaki gibi parçalanmış ve dağılım matrisli çok değişkenli normal dağılıma sahip olduğunu farz edelim. $x_1, \ldots, x_p$ verildiğinde y' nin şartlı dağılımı,

$$N(\tau + \mathbf{u}'\Sigma^{-1}(\mathbf{X} - \mu), \sigma^2 - \mathbf{u}'\Sigma^{-1}\mathbf{u})$$

dir. Bu nedenle, $\mathbf{x} = (x_1, \ldots, x_p)'$ verildiğinde y' nin şartlı varyansı ,

$$\sigma^2\left(1 - r^2_{y(x_1,\ldots,x_p)}\right)$$

dir. **x** verildiğinde **y'** nin şartlı beklenen değeri (x üzerinde y'nin regresyon doğrusu olarak da bilinen),

$$\mathbf{u}'\Sigma^{-1}\mathbf{x} + \tau - \mathbf{u}'\Sigma^{-1}\mu$$

dür ve $\mathbf{u}'\Sigma^{-1}\mathbf{x}$ in kesin olarak y ile maksimum korelasyona sahip olan $x_1, \ldots, x_p$ 'nin lineer kombinasyonu olduğunu hatırlayalım. Bu nedenle, eğer değişkenlerin dağılımı çok değişkenli normal dağılım ise, çoklu korelasyon katsayısı özel yorum kabul eder.

Verilen bir durumda $y, x_1, \ldots, x_p$ rastgele değişkenlerinin var olduğunu ve y ve x_i ler arasındaki ilişkiyi incelemek istediğimizi farz edelim. Özellikle x_i lerin değerleri verildiğinde y nin değerini tahmin etmeyi isteyebiliriz. İlk olarak, $x_1, \ldots, x_p$ ye dikkatle bakalım. Bu takdirde onları sabit gibi işleme tabi tutarak, gerekli olabilen süreçte, herhangi bir denemeye bağladıktan sonra y üzerinde bir gözlem alırız. Şimdi, eğer y üzerindeki alışılmış varsayımlarla,

$$E(y) = \beta_0 + \beta_1 x_1 + \ldots + \beta_p x_p$$

modelini öngörebilirsek, bu takdirde bir lineer model elde ederiz. Modeldeki E(y) terimi $E(y|x_1,...,x_p)$ şartlı beklenen değeri olarak yorumlanacaktır. Eğer değişkenler üzerinde n tane veri noktasına sahipsek model,

$$E(y_i) = \beta_0 + \beta_1 x_{i1} + ... + \beta_p x_{ip} \qquad i\text{-}1,...,n, \qquad (2.5.7.2)$$

olarak yazılabilir ve lineer model için geliştirilen yöntemlerle analiz edilebilir. Böyle bir modele bir regresyon modeli denir. x_{ij} ler bir rastgele değişken üzerindeki gözlemler olduklarından, modeli tam ranklı olarak kabul edebiliriz(yani x in rankı sütun sayısına eşittir). Gerçekten $x_1,...,x_p$' nin dağılımı üzerindeki ılımlı varsayımlar altında modelin katsayı matrisinin bir olasılıkla(kesin olarak) tam ranka sahip olduğu gösterilebilir. Bu nedenle tam ranklı model ve regresyon modeli terimleri birbirinin yerine kullanılır. $y_1,...,y_n$'lerin bağımsız $N(0,\sigma^2)$ olduğu (2.5.7.2) modelini göz önüne alalım. **e**, elemanlarının tümü 1 olan bir model,

$$E(\mathbf{y}) = [\mathbf{e} \quad \mathbf{X}]\begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix}$$

olarak ifade edilebilir. Modelin tam ranklı olduğunu kabul ederiz ve böylece, $\beta_0,...,\beta_p$'nin **BLUE**'leri,

$$\begin{bmatrix} \hat{\boldsymbol{\beta}}_0 \\ \hat{\boldsymbol{\beta}}_1 \\ \vdots \\ \hat{\boldsymbol{\beta}}_p \end{bmatrix} = \left(\begin{bmatrix} \mathbf{e}' \\ \mathbf{X}' \end{bmatrix}[\mathbf{e} \quad \mathbf{X}]\right)^{-1}\begin{bmatrix} \mathbf{e}' \\ \mathbf{X}' \end{bmatrix}\mathbf{y}$$

$$= \begin{bmatrix} n & \mathbf{e}'\mathbf{X} \\ \mathbf{X}'\mathbf{e} & \mathbf{X}'\mathbf{X} \end{bmatrix}^{-1}\begin{bmatrix} \mathbf{e}'\mathbf{y} \\ \mathbf{X}'\mathbf{y} \end{bmatrix}$$

ile verilir. $\mathbf{M} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}', \mathbf{Q} = \mathbf{I} - \frac{1}{n}\mathbf{e}\mathbf{e}'$ ve $\mathbf{z} = -\frac{1}{n}(\mathbf{X}'\mathbf{Q}\mathbf{X})^{-1}\mathbf{X}'\mathbf{e}$ olmak üzere,

$$\begin{bmatrix} n & \mathbf{e'X} \\ \mathbf{X'e} & \mathbf{X'X} \end{bmatrix}^{-1} = \begin{bmatrix} \dfrac{1}{n - \mathbf{e'Me}} & \mathbf{z'} \\ \mathbf{z} & (\mathbf{X'QX})^{-1} \end{bmatrix}$$

olduğu gösterilebilir. Bu nedenle, $\beta_1, ..., \beta_p$'nin **BLUE**'leri,

$$\begin{bmatrix} \hat{\boldsymbol{\beta}}_1 \\ \vdots \\ \hat{\boldsymbol{\beta}}_p \end{bmatrix} = (\mathbf{X'QX})^{-1}\mathbf{X'Qy} \qquad (2.5.7.3)$$

ile verilir. $(y, x_1, ..., x_p)$ değişkenlerinin örneklem dağılım matrisi aşağıdaki gibi hesaplanır;

$$\mathbf{S} = \begin{bmatrix} y_1 & \\ \vdots & \mathbf{X} \\ y_n & \end{bmatrix}$$

olsun. Bu takdirde dağılım matrisi,

$$\mathbf{S'S} - \frac{1}{n}\mathbf{S'ee'S} = \mathbf{S'QS}$$
$$= \begin{bmatrix} \mathbf{y'Qy} & \mathbf{y'Qx} \\ \mathbf{X'Qy} & \mathbf{X'QX} \end{bmatrix}$$

dir. Böylece $x_1, ..., x_p$'nin y ile maksimum korelasyona sahip olan lineer fonksiyonu,

$$\boldsymbol{\alpha} = (\mathbf{X'QX})^{-1}\mathbf{X'Qy}$$

olarak elde edilir ve bu (2.5.7.3) ile çakışır. Kısaca **y** ile maksimum korelasyona sahip olan $\beta_1 x_1 + ... + \beta_p x_p$ lineer fonksiyonu $\hat{\beta}_i$, β_i'nin en küçük kareler tahmini olmak üzere, $\beta_i = \hat{\beta}_i$ alarak elde edilir.

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_{i1} + ... + \hat{\beta}_p x_{ip}, \qquad i \;\; 1, ..., n,$$

y_i'nin tahmin edilen değeri olsun. $\hat{\mathbf{y}} = (\hat{y}_1, ..., \hat{y}_n)' \sum_{i\;1}^{n} (X_i - \bar{X})^2$ ve $\hat{y} = \frac{1}{n}\mathbf{e'\hat{y}}$ olsun. Bu takdirde,

$$
\begin{aligned}
r^2_{y(x_1,\ldots,x_p)} &= \left\{korelasyon(y, \hat{\beta}_1 x_1 + \ldots + \hat{\beta}_p x_p)\right\}^2 \\
&= \left\{korelasyon(y, \hat{y})\right\}^2 \\
&= \frac{\left\{\sum_{i=1}^{n}(y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})\right\}^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2 \sum_{i=1}^{n}(\hat{y}_i - \bar{\hat{y}})^2}
\end{aligned}
$$

dir. Yukarıdaki ifadenin karekökü bir örneklemden hesaplanan çoklu korelasyon katsayısıdır ve determinasyon katsayısı olarak bilinir.

Şimdi $H : \beta_1 = \ldots = \beta_p = 0$ hipotezi için F-istatistiğini ortaya koyacağız.

$$
\text{HKT} = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2
$$

elde ederiz ve karşılık gelen serbestlik derecesi $n - p - 1$ dir. HKT_H ’ı bulmak için,

$$
\sum_{i=1}^{n}(y_i - \beta_0)^2
$$

toplamını minimum yapmalıyız ve bu $\beta_0 = \bar{y}$ olduğunda minimum olur. Şimdi serbestlik derecesi p’ dir. Bu nedenle,

$$
\frac{(\text{HKT}_H - \text{HKT})/p}{\text{HKT}/(n-p-1)} = \frac{\sum_{i=1}^{n}(y_i - \bar{y})^2 - \sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \times \frac{n-p-1}{p} \qquad (2.5.7.4)
$$

test istatistiği eğer H hipotezi doğru ise, $F(p, n-p-1)$ dağılımına sahiptir.

3. BULGULAR

3.1. Kısıtlanmış Lineer Modelde Tahmin ve Hipotez Testinin Geometrisi-Tam Ranklı Durum

Lineer modellere geometrik yaklaşımda dikkat, tahmin ve hipotez testinin bilinen vektör uzayı tekniklerini kullanarak başarılabildiği; gözlem uzayı üzerinde toplanır. Eğer **Y**, n-boyutlu $\mathbf{E}^n$ öklid uzayının bir q-boyutlu alt uzayı olan, Ω da olduğu bilinen, $E(\mathbf{Y})$ ortalama vektörüne sahip olan gözlemlerin bir $nx1$ vektörü ise, bu takdirde, ortalama vektör $\hat{\mathbf{Y}} = \mathbf{P}_\Omega \mathbf{Y}$ ile tahmin edilir. Burada $\mathbf{P}_\Omega$, $\mathbf{E}^n$'nin Ω üzerindeki ortogonal (dik) izdüşümüdür. $\boldsymbol{\omega}$, Ω'nın r-boyutlu bir altuzayı olmak üzere, sıfır hipotezleri $E(\mathbf{Y}) \in \boldsymbol{\omega}$ olarak ifade edilirler. γ, $\boldsymbol{\omega}$'nın Ω'daki ortogonal tümleyeni alınarak, $\omega \oplus \gamma = \Omega$ olarak yazılır (γ, $\mathbf{v} \in \gamma$ ve $\mathbf{w} \in \omega$ için $\mathbf{v'w} = \mathbf{0}$ olacak şekilde, Ω'nın q-r boyutlu bir altuzayıdır. Burada $\mathbf{v}'$, $\mathbf{v}$' nin transpozesini gösterir).

$$\left(\left\|\mathbf{P}_\gamma \mathbf{Y}\right\|^2 : (q-r)\right) / \left(\left\|\mathbf{P}_{\Omega^\perp} \mathbf{Y}\right\|^2 : (n-q)\right) = \left(\mathbf{Y'P}_\gamma \mathbf{Y} : (q-r)\right) / \left(\mathbf{Y'P}_{\Omega^\perp} \mathbf{Y} : (n-q)\right)$$

dur. $E(\mathbf{Y}) \in \boldsymbol{\omega}$ sıfır hipotezini test etmek için alışılmış orandır. Bununla beraber eğer η, γ'nın herhangi bir altuzayı ve δ, $\Omega^\perp$'nin herhangi bir altuzayı ise, bu takdirde, sıfır hipotezi doğru olduğunda,

$$\left(\| \mathbf{P}_\eta \mathbf{Y} \|^2 : b\right) / \left(\| \mathbf{P}_\delta \mathbf{Y} \|^2 : d\right)$$

bir merkezi F-dağılımına sahiptir (**b** ve **d** sırasıyla η ve δ'nın boyutlarıdır. $\mathbf{v} \in \mathbf{E}^n$ için, $\|\mathbf{v}\|^2 = \mathbf{v'v}$ dir) (bkz. Scheffe 1959 ss. 8-12, 38, 42-45; Kruskal 1960; Seber 1977 ss.43-45, 96, 97, 120).

Klasik yaklaşım (tamamen cebirsel olan) $E(\mathbf{Y}) = \mathbf{X}\boldsymbol{\beta}$ parametrelendirilmiş modeli ile başlar (bir ANOVA, yani, varyans analizi modelinde, $\boldsymbol{\beta}$ pekâlâ göze ortalamalarının vektörü olabilir). Sıfır hipotezleri, $H_o : \mathbf{H'}\boldsymbol{\beta} = \mathbf{0}$ olarak parametre vektörüne göre ifade edilirler. $\boldsymbol{\beta}$'nın $\hat{\boldsymbol{\beta}}$ tahmini $\boldsymbol{\beta}$'ya göre bir kuadratik (karesel) formu minimumlaştırarak tahmin edilir ve $E(\mathbf{Y})$, $\hat{\mathbf{Y}} = \mathbf{X}\hat{\boldsymbol{\beta}}$ ile tahmin edilir. Sıfır hipotezinin test edilmesi için pay kareler toplamı, $\hat{\boldsymbol{\beta}}'\mathbf{H}\left(\mathbf{H}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{H}\right)^{-1}\mathbf{H}'\hat{\boldsymbol{\beta}}$ ile verilir. Kısıtlanmış modelde parametre vektörünün $\mathbf{C}'\boldsymbol{\beta} = \mathbf{0}$ şartını sağlaması istenir. $\boldsymbol{\beta}$'nın $\hat{\boldsymbol{\beta}}_C$ tahminini hesaplamak için $\boldsymbol{\beta}$'ya göre bir kuadratik form $\mathbf{C}'\boldsymbol{\beta} = \mathbf{0}$ kısıtlamasına bağlı olarak minimumlaştırılır. Bu, Lagrange çarpanları tekniğini kullanarak başarılır. (bkz. Graybill 1976 s. 222; Searle1971 ss. 112, 113; Seber 1977 s. 84). Ortalama vektörü $\hat{\mathbf{Y}}_C = \mathbf{X}\hat{\boldsymbol{\beta}}_C$ ile tahmin edilir. Kısıtlanmış modelde $\mathbf{H}'\boldsymbol{\beta} = \mathbf{0}$ hipotezini test etmek için, **C** ve **H** matrisleri yeni bir **T=(C, H)** matrisi oluşturmak için bitiştirilirler. Bu test için pay kareler toplamı,

$$\mathbf{SS}(\mathbf{C},\mathbf{H}) = \hat{\boldsymbol{\beta}}'\left[\mathbf{T}\left\{\mathbf{T}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{T}\right\}^{-1}\mathbf{T}' - \mathbf{C}\left\{\mathbf{C}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{C}\right\}^{-1}\mathbf{C}'\right]\hat{\boldsymbol{\beta}} \qquad (3.1.1)$$

dır. Sıfır hipotezi doğru olduğunda, kısıtlanmış modelde $\boldsymbol{\beta}$'nın $\hat{\boldsymbol{\beta}}_{C,H}$'nın tahmini $\mathbf{T}'\boldsymbol{\beta} = \mathbf{0}$ kısıtlamasına bağlı $E(\mathbf{Y}) = \mathbf{X}\boldsymbol{\beta}$ kısıtlanmış modelinde, sanki $\boldsymbol{\beta}$'nın tahminiymiş gibi hesaplanır. Hesaplama formülü,

$$\hat{\boldsymbol{\beta}}_{C,H} = \left[\mathbf{I} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{T}\left\{\mathbf{T}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{T}\right\}^{-1}\mathbf{T}'\right]\hat{\boldsymbol{\beta}} \qquad (3.1.2)$$

dır (bkz. Searle 1971 s. 207; Seber1977 ss. 121-123). Sıfır hipotezi doğru olduğunda kısıtlanmış modelde $\hat{\mathbf{Y}}_{C,H} = \mathbf{X}\hat{\boldsymbol{\beta}}_{C,H}$, $E(\mathbf{Y})$'yi tahmin eder. Eğer **C** ve **H** sırasıyla qxt ve qxs boyutlu matrisler ise, bu takdirde, **T**, qx$(t+s)$ boyutludur ve $(t+s)$x$(t-s)$ boyutlu $\mathbf{T}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{T}$ matrisinin (3.1.1) ve (3.1.2) deki tersi var olmalıdır.

Bu kısımda kısıtlanmış lineer modele ilişkin tüm cebirsel hesaplamaları geometrik olarak harekete geçireceğiz. Başlangıç olarak model ve ilgili hipotez,

$$\mathbf{C}'\boldsymbol{\beta} = \mathbf{0} \text{ ile kısıtlanmış } E(\mathbf{Y}) = \mathbf{X}\boldsymbol{\beta}$$

$$H_0 : \mathbf{H}'\boldsymbol{\beta} = \mathbf{0}$$

gibi parametrik olarak ifade edilirler ve sonra da geometrik analiz için gözlem uzayına dönüştürülürler.

Geometrik yaklaşım hem kısıtlanmış hem de kısıtlanmamış modellere uygulanan tahmin ve hipotez testi için tutarlı bir teoriyi verir (Lagrange çarpanları yöntemi kısıtlanmış modelde kullanılmaz). Bundan başka, geometrik düşünceler **SS(C,H)** için bu kuadratik form, sadece $\mathbf{X'X}$ hariç txt ve sxs matrislerinin terslerini içerdiğinden, (3.1.1) bağıntısı ile verilen, $\hat{\boldsymbol{\beta}}$'ya göre iki kuadratik formun farkından, hesaplama bakımından daha etkin olan, $\hat{\boldsymbol{\beta}}_C$'ya göre tek bir kuadratik form olarak daha doğal bir ifadeye götürür. (Coleman:1977)

Tahminin geometrisi uygun bir ortogonol izdüşüm yoluyla herhangi bir öncelden ($\hat{\mathbf{Y}}$'nın önceli $\mathbf{Y}$'dir) elde edilebilen üç ardışık $\hat{\mathbf{Y}}, \hat{\mathbf{Y}}_C$ ve $\hat{\mathbf{Y}}_{C,H}$ tahminlerinden herhangi birini açıklar. Aynı sonuç, ortogonal olmayan izdüşümler dışında, parametrenin tahminleri için doğrudur. Altı tahminin (ortalamanın ve parametre vektörlerinin) tümünü hesaplama için etkin bir hesaplama şeması da verilir.

$$\left.\begin{array}{c} \mathbf{C'\boldsymbol{\beta}} = \mathbf{m} \neq \mathbf{0} \text{ ile kısıtlanmış } E(\mathbf{Y}) = \mathbf{X}\boldsymbol{\beta} \\ H_0 : \mathbf{H'\boldsymbol{\beta}} = \mathbf{h} \neq \mathbf{0} \end{array}\right\} \tag{3.1.3}$$

durumu zorluk arz etmez. **C** ve **H**' nin sütunları lineer bağımsız iseler, bu takdirde, $\mathbf{C'\boldsymbol{\beta}} = \mathbf{m}$ ve $\mathbf{H'\boldsymbol{\beta}} = \mathbf{h}$ bir ortak çözüme sahiptirler (Aslında ortak çözümlerin kümesi **C** ve **H**'ye göre tanımlanabilen gerçekten dönüştürülmüş bir altuzaydır) ve veri, (3.1.3) bağıntısı,

$$\left.\begin{array}{c} \mathbf{C'\tilde{\boldsymbol{\beta}}} = \mathbf{0} \text{ ile kısıtlanan } E(\tilde{\mathbf{Y}}) = \mathbf{X}\tilde{\boldsymbol{\beta}} \\ H_0 : \mathbf{H'\tilde{\boldsymbol{\beta}}} = \mathbf{0} \end{array}\right\} \tag{3.1.4}$$

olarak yazılabilecek şekilde dönüştürülebilir.

Notasyon

Bu kısımda başından sonuna kadar $\mathbf{X}$ tam sütun ranka sahip bir nxq matris, $\boldsymbol{\beta}$ bir qx1 sabit vektör (parametre vektörü) ve $\mathbf{I}$ nxn birim matris olmak üzere, $\mathbf{Y}$, $E(\mathbf{Y}) = \mathbf{X}\boldsymbol{\beta}$ ortalama vektörüne ve $\sigma^2\mathbf{I}$ kovaryans matrisine sahip olan nx1 boyutlu çok değişkenli bir normal rastgele vektördür; Ω, X'in sütunları tarafından gerilen $\mathbf{E}^n$'nin altuzayını gösterecek, qxt kısıtlama matrisi $\mathbf{C}$ ile gösterilecek ve $t+s<q$ olmak üzere hipotez matrisi qxs boyutlu $\mathbf{H}$ matrisi ile gösterilecek. $\mathbf{C}$ ve $\mathbf{H}$'nin sütunları birlikte $\mathbf{E}^q$ da $t+s$ tane lineer bağımsız vektörün bir kümesini oluştururlar.

$\mathbf{P}_A$ sembolü $\mathbf{A}$ üzerine ortogonal izdüşümü ve $\mathbf{Q}_A = \mathbf{I} - \mathbf{P}_A$, $\mathbf{A}^{\perp}$ üzerine ortogonal izdüşümü gösterecek. Burada $\mathbf{A}^{-}$, $\mathbf{A}$'nın sütun uzayının ortogonal tümleyenidir ($\mathbf{I}$, $\mathbf{A}$'nın, $\mathbf{E}^n$'nin veya $\mathbf{E}^q$'nun bir altuzayı olup olmamasına bağlı olarak n veya q boyutlu birim matristir).

3.2. Kısıtlanmamış Model

3.2.1. Tahmin

Eğer $E(\mathbf{Y}) = \mathbf{X}\boldsymbol{\beta}$ ise, bu takdirde, $E(\mathbf{Y})$, Ω da uzanır. $\mathbf{P}_\Omega = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$, $\mathbf{E}^n$'nin Ω üzerine ortogonal izdüşümü olmak üzere, $E(\mathbf{Y})$'nin GM (Gauss-Markov) tahmini $\hat{\mathbf{Y}} = \mathbf{P}_\Omega\mathbf{Y}$ ile verilir. Ortalama vektörü tahmin edildiğinde, $\boldsymbol{\beta} = \hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\hat{\mathbf{Y}}$, $\mathbf{X}\boldsymbol{\beta} = \hat{\mathbf{Y}}$ denklemine tek bir çözümü verir. $\hat{\mathbf{Y}}$ için $\mathbf{P}_\Omega\mathbf{Y}$'yi $\hat{\boldsymbol{\beta}}$ için tanımlanan denklemde yerine koyma daha fazla bilinen $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$ sonucunu verir (bkz. Ek 2, Teorem 1; Herr 1979; Kruskal 1960; Scheffe 1959 s. 11; Seber 1977 ss. 43, 44).

3.2.2. Hipotez Testi

$\mathbf{H}'\boldsymbol{\beta} = \mathbf{0}$ hipotezini test etmek için , $\boldsymbol{\gamma}_H = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{H}$ olmak üzere, $\mathbf{H}'\boldsymbol{\beta} = \mathbf{0}$ olması için gerek ve yeter şartın $\boldsymbol{\gamma}_H' E(\mathbf{Y}) = \mathbf{0}$ olduğunu belirtelim.

$\boldsymbol{\omega}_H$, $\boldsymbol{\gamma}_H$'ın $\boldsymbol{\Omega}$'daki ortogonal tümleyeni olsun. Bu takdirde $\boldsymbol{\omega}_H \oplus \boldsymbol{\gamma}_H = \boldsymbol{\Omega}$'dır (kısım 3.2.2'de buraya kadar, $\boldsymbol{\gamma}_H$ hem $\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{H}$ matrisini hem de $\mathbf{E}^n$'nin $\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{H}$ matrisinin sütunları tarafından gerilen altuzayını göstermek için kullanılmıştır). Eğer ortalama vektörü $\gamma_H'E(\mathbf{Y})=\mathbf{0}$ bağıntısını sağlar ise, bu takdirde, $E(\mathbf{Y})$, ω_H'da uzanır ve bu nedenle, sıfır hipotezi $E(\mathbf{Y})\in\omega_H$ olarak ifade edilir. Sıfır hipotezinin testi için pay kareler toplamı $\|\mathbf{P}_{\gamma(\mathbf{H})}\mathbf{Y}\|^2$ dir. Burada $\mathbf{P}_\gamma(\mathbf{H})$, $\gamma(\mathbf{H})=\gamma_H$ üzerine ortogonal izdüşümdür. Cebiri aşağıda yürüteceğiz ($\mathbf{P}_\gamma(\mathbf{H})$'ın hesabı için bkz. Ek 2; Teorem 1).

$$\begin{aligned}\|\mathbf{P}_{\gamma(\mathbf{H})}\mathbf{Y}\|^2 &= \mathbf{Y}'\mathbf{P}_{\gamma(\mathbf{H})}\mathbf{Y}\\ &= \mathbf{Y}'\gamma_H(\gamma_H'\gamma_H)^{-1}\gamma_H'\mathbf{Y}\\ &= \mathbf{Y}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{H}\left\{\mathbf{H}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{H}\right\}^{-1}\mathbf{H}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}\\ &= \hat{\boldsymbol\beta}'\mathbf{H}\left\{\mathbf{H}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{H}\right\}^{-1}\mathbf{H}'\hat{\boldsymbol\beta}\end{aligned} \tag{3.2.2.1}$$

dır. Hata kareler toplamı, yani, $\mathbf{Q}_\Omega$, $\mathbf{E}^n$ nin $\boldsymbol\Omega$ üzerine ortogonal izdüşümü olmak üzere, HKT(*unc*)=$\|\mathbf{Q}_\Omega\mathbf{Y}\|^2=\mathbf{Y}'\mathbf{Q}_\Omega\mathbf{Y}$ ile verilir (Burada "unc." *Kısıtlanmamış* sözcüğünü ifade eder). $\mathbf{Q}_\Omega=\mathbf{I}-\mathbf{P}_\Omega$ olduğundan, bilinen HKT(*unc*)$=\mathbf{Y}'\left\{\mathbf{I}-\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\right\}\mathbf{Y}$ formülünü elde ederiz. (bkz. Searle 1971 s. 92; Seber 1977 s. 120).

3.3. Kısıtlanmış Model

3.3.1. Tahmin

Parametre vektörü, $\mathbf{C}'\boldsymbol\beta=\mathbf{0}$ ile kısıtlandığında $\mathbf{C}'\boldsymbol\beta=\mathbf{0}$ olması için gerek ve yeter şart $\gamma_C'\mathbf{E}(\mathbf{Y})=\mathbf{0}$ olduğundan, $E(\mathbf{Y})=\mathbf{X}\boldsymbol\beta$ ortalama vektörü $\gamma_C=\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{C}$'ye ortogonal olmalıdır. γ_C'nın $\boldsymbol\Omega$'daki ortogonal tümleyenini $\boldsymbol\Omega_C$ alarak (bu nedenle, $\gamma_C\oplus\boldsymbol\Omega_C=\boldsymbol\Omega$ alarak), $\mathbf{C}'\boldsymbol\beta=\mathbf{0}$ kısıtlaması, $E(\mathbf{Y})$ ortalama vektörünün $\boldsymbol\Omega_C$'de uzanması kısıtlamasını koyar. Bu nedenle $\boldsymbol\Omega(\mathbf{C})=\boldsymbol\Omega_C$ olmak üzere, $E(\mathbf{Y})$'nin GM tahmini $\hat{\mathbf{Y}}_C=\mathbf{P}_{\Omega(C)}\mathbf{Y}$ dir.

Eğer $\mathbf{C}'\boldsymbol\beta=\mathbf{0}$ ise, bu taktirde, $\boldsymbol\beta$, C'nin sütunlarına ortogonal olmalıdır, yani, $\boldsymbol\beta$, $\mathbf{C}^\perp$'de uzanmalıdır. Burada $\mathbf{C}^\perp$, C'nin $\mathbf{E}^q$'daki ortogonal tümleyenidir. Bu nedenle, γ_C

ile $\mathbf{C}$ ve $\boldsymbol{\Omega}_C$ ile $\mathbf{C}^{\perp}$ altuzaylarının doğal bir eşlemesi vardır. Altuzay çiftlerinin bağlanması için uygun dönüşümler aşağıda verilirler.

$$\begin{aligned} \mathbf{X(X'X)^{-1}} \; : \; \mathbf{C} &\rightleftarrows \gamma_C \; : \mathbf{X'} \\ \mathbf{X} \; : \; \mathbf{C}^{\perp} &\rightleftarrows \boldsymbol{\Omega}_C \; : \mathbf{(X'X)^{-1}X'} \end{aligned}$$

dür (bkz. Ek 2; Teorem 2). Burdick ve ortakları (1974), $\mathbf{X(X'X)^{-1}, X'}$ ve $\mathbf{X}$, $\mathbf{(X'X)^{-1}X'}$ dönüşüm çiftlerini sırasıyla "en iyi tahmin köprüsü" ve "en iyi uyum köprüsü" olarak adlandırırlar. $\mathbf{P}_{\Omega(C)}$'yi hesaplamak için, $\gamma(\mathbf{C}) = \gamma_C$ olmak üzere, $\mathbf{P}_{\Omega(C)} = \mathbf{P}_{\Omega} - \mathbf{P}_{\gamma(C)}$ yazalım.

$$\begin{aligned} \mathbf{P}_{\Omega} &= \mathbf{X(X'X)^{-1}X'} \text{ ve} \\ \mathbf{P}_{\gamma(C)} &= \gamma_C(\gamma_C'\gamma_C)^{-1}\gamma_C' \\ &= \mathbf{X(X'X)^{-1}C\left\{C'(X'X)^{-1}C\right\}^{-1}C'(X'X)^{-1}X'} \end{aligned}$$

olduğundan,

$$\mathbf{P}_{\Omega(C)} = \mathbf{XA}_C\mathbf{(X'X)^{-1}X'} \tag{3.3.1.1}$$

dür ve burada,

$$\mathbf{A}_C = \mathbf{I} - \mathbf{(X'X)^{-1}C\left\{C'(X'X)^{-1}C\right\}^{-1}C'} \tag{3.3.1.2}$$

, $\mathbf{E}^q$'nun $\mathbf{(X'X)^{-1}C}$ boyunca veya $\mathbf{(X'X)^{-1}C}$'ye paralel olarak $\mathbf{C}^{\perp}$ üzerine izdüşümü değildir (bkz. Ek 2; Teorem 5). (Bundan böyle $\mathbf{A}_C$ sembolü, (3.3.1.2) bağıntısı ile tanımlandığı gibi daima $\mathbf{E}^q$'nun $\mathbf{(X'X)^{-1}C}$ boyunca $\mathbf{C}^{\perp}$ üzerine izdüşümünü gösterecektir.) $\hat{\mathbf{Y}}_C = \mathbf{P}_{\Omega(C)}\mathbf{Y} = \mathbf{XA}_C\mathbf{(X'X)^{-1}X'Y}$ hesaplanmış olduğundan, $\boldsymbol{\beta}$,

$$\hat{\boldsymbol{\beta}}_C = \mathbf{(X'X)^{-1}X'}\hat{\mathbf{Y}}_C = \mathbf{A}_C\mathbf{(X'X)^{-1}X'Y} \tag{3.3.1.3}$$

ile tahmin edilir. $\boldsymbol{\beta} = \hat{\boldsymbol{\beta}}_C$'nın $\mathbf{X}\boldsymbol{\beta} = \hat{\mathbf{Y}}_C$ denkleminin bir tek çözüm olduğu ve $\hat{\boldsymbol{\beta}}_C \in \mathbf{C}$ olduğu kolayca gösterilebilir.

3.3.2. Hipotez Testi

$\boldsymbol{\Omega}_C$'nin, $E(\mathbf{Y}) \in \boldsymbol{\Omega}_C$ olması için gerek ve yeter şartın, $E(\mathbf{Y}) = \mathbf{X}\boldsymbol{\beta}$ ve $\mathbf{C}'\boldsymbol{\beta} = \mathbf{0}$ denklemlerinin aynı anda sağlaması özelliğine sahip olan $\mathbf{E}^n$'nin bir altuzayı olduğunu hatırlayalım. İlk işimiz $E(\mathbf{Y}) \in \boldsymbol{\Omega}_C$ ve $\mathbf{H}'\boldsymbol{\beta} = \mathbf{0}$ olması için gerek ve yeter şart $E(\mathbf{Y}) \in \boldsymbol{\omega}_{C,H}$ olacak şekilde $\boldsymbol{\Omega}_C$'nin $\boldsymbol{\omega}_{C,H}$ altuzayını bulmaktır. Bu takdirde, $\boldsymbol{\omega}_{C,H}$'ın $\boldsymbol{\Omega}_C$'deki $\Pi_{C,H}$ ortogonal tümleyenini oluşturacağız. Kısıtlanmış modelde $\mathbf{H}'\boldsymbol{\beta} = \mathbf{0}$ hipotezini test etme için pay ve payda kareler toplamları sırasıyla $\mathbf{Y}'\mathbf{P}_{\Pi(C,H)}\mathbf{Y}$ ve $\mathbf{Y}'\mathbf{Q}_{\Omega(C)}\mathbf{Y}$ ile verilirler. Burada $\Pi(\mathbf{C},\mathbf{H}) = \Pi_{C,H}$ dır.

Kısıtlanmış modelde sıfır hipotezi doğru olduğunda, $\mathbf{C}'\boldsymbol{\beta} = \mathbf{0}$ ve $\mathbf{H}'\boldsymbol{\beta} = \mathbf{0}$ dır. Bu iki denklem, **T=(C, H)** olmak üzere, bir tek $\mathbf{T}'\boldsymbol{\beta} = \mathbf{0}$ denklemi olarak yeniden yazılabilir. Ancak $\mathbf{T}'\boldsymbol{\beta} = \mathbf{0}$ olması için gerek ve yeter şart $\gamma_T' E(\mathbf{Y}) = \mathbf{0}$ olmasıdır. Burada $\gamma_T = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{T}$ dir. Bu nedenle, kısıtlanmış modelde sıfır hipotezi doğru olduğunda, $E(\mathbf{Y}) \in \boldsymbol{\omega}_{C,H}$, γ_T'nin, $\boldsymbol{\Omega}$($\boldsymbol{\omega}_{C,H} \oplus \gamma_T = \boldsymbol{\Omega}$)'daki ortogonal tümleyenidir. $\boldsymbol{\omega}_{C,H}$'ın $\boldsymbol{\Omega}_C$'deki ortogonal tümleyenini $\Pi_{C,H}$ alma, bu nedenle $\boldsymbol{\omega}_{C,H} \oplus \Pi_{C,H} = \boldsymbol{\Omega}_C$ alma,

$$\mathbf{P}_{\Pi(C,H)} = \mathbf{P}_{\Omega(C)} - \mathbf{P}_{\omega(C,H)} \tag{3.3.2.1}$$

olduğunu gösterir. (3.3.2.1) bağıntısında $\mathbf{P}_{\Omega(C)}$ ve $\mathbf{P}_{\omega(C,H)}$ yerine sırasıyla $\mathbf{P}_{\Omega} - \mathbf{P}_{\gamma(C)}$ ve $\mathbf{P}_{\Omega} - \mathbf{P}_{\gamma(T)}$ koyma,

$$\begin{aligned}\mathbf{P}_{\Pi(C,H)} &= \mathbf{P}_{\gamma(T)} - \mathbf{P}_{\gamma(C)} \\ &= \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\left[\mathbf{T}\left\{\mathbf{T}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{T}\right\}^{-1}\mathbf{T}' - \mathbf{C}\left\{\mathbf{C}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{C}\right\}^{-1}\mathbf{C}'\right](\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\end{aligned}$$

bağıntısını ortaya koyar. $\mathbf{P}_{H(C,H)}$ için bu formül $\mathbf{T}$ genişletilmiş matrisini içerir ve $\mathbf{Y}'\mathrm{P}_{H(C,H)}\mathbf{Y}$'de yerine konulduğunda (3.1.1) bağıntısını üretir. $\mathbf{P}_{H(C,H)}$ için yalnız $\mathbf{H}$ ve $\mathbf{C}$ ye göre ($\mathbf{T}$'nin olmadığı) bir ifadeyi elde etmek için,

$$\gamma_T = \gamma_C + \gamma_H = \gamma_C \oplus \mathbf{Q}_{\gamma(C)}\gamma_H \tag{3.3.2.2}$$

yazarız (bkz. Ek 2; Teorem 4). (3.3.2.2) bağıntısı $\mathbf{\Omega}_C = \mathbf{\omega}_{C,H} \oplus \mathbf{Q}_{\gamma(C)}\gamma_H$ olduğunu ifade eder. Bu nedenle, $\Pi_{C,H} = \mathbf{Q}_{\gamma(C)}\gamma_H = \mathbf{X}\mathbf{A}_C(\mathbf{X}'\mathbf{X})^{-1}\mathbf{H}$ dır ve $\mathbf{A}_C(\mathbf{X}'\mathbf{X})^{-1}\mathbf{H}$ ve $\Pi_{C,H}$ arasında doğal bir tekabül, yani, her ikisinin de birebir ve üzerine olan,

$$\mathbf{X}:\mathbf{A}_C(\mathbf{X}'\mathbf{X})^{-1}\mathbf{H} \rightleftarrows \Pi_{C,H} : (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$$

dönüşümleri vardır (bkz. Ek 2; Teorem 2).($\mathbf{A}_C(\mathbf{X}'\mathbf{X})^{-1}\mathbf{H}$, $\mathbf{C}$'ye ortogonal olmakla beraber $\mathbf{T}$'de ihtiva edilmez ve bu nedenle $\mathbf{T}$ 'ne ortogonal değildir). $\mathbf{E}^n$'nin $\Pi_{C,H}$ üzerine ortogonal izdüşümü,

$$\mathbf{P}_{H(C,H)} = \mathbf{Q}_{\gamma(C)}\gamma_H(\gamma_H'\mathbf{Q}_{\gamma(C)}\gamma_H)^{-1}\gamma_H'\mathbf{Q}_{\gamma(C)} \tag{3.3.2.3}$$

ile verilir.

$$\gamma_H'\mathbf{Q}_{\gamma(C)} = \mathbf{H}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Q}_{\gamma(C)} = \mathbf{H}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Q}_{\gamma(C)} = \gamma_H'\mathbf{P}_\Omega\mathbf{Q}_{\gamma(C)} = \gamma_H'\mathbf{P}_{\Omega(C)}$$

olduğundan (3.3.2.3) bağıntısı,

$$\mathbf{P}_{H(C,H)} = \mathbf{P}_{\Omega(C)}\gamma_H(\gamma_H'\mathbf{P}_{\Omega(C)}\gamma_H)^{-1}\gamma_H'\mathbf{P}_{\Omega(C)} \tag{3.3.2.4}$$

$$= \mathbf{X}\mathbf{A}_C(\mathbf{X}'\mathbf{X})^{-1}\mathbf{H}\left\{\mathbf{H}'\mathbf{A}_C(\mathbf{X}'\mathbf{X})^{-1}\mathbf{H}\right\}^{-1}\mathbf{H}'\mathbf{A}_C(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' \tag{3.3.2.5}$$

olarak yeniden yazılabilir. Test için pay kareler toplamı,

$$\begin{gathered}\mathbf{Y}'\mathbf{P}_{H(C,H)}\mathbf{Y} = \mathbf{Y}'\mathbf{P}_{\Omega(C)}\gamma_H(\gamma_H'\mathbf{P}_{\Omega(C)}\gamma_H)^{-1}\gamma_H'\mathbf{P}_{\Omega(C)}\mathbf{Y} \\ \hat{\boldsymbol\beta}_C'\mathbf{H}\left\{\mathbf{H}'\mathbf{A}_C(\mathbf{X}'\mathbf{X})^{-1}\mathbf{H}\right\}^{-1}\mathbf{H}'\hat{\boldsymbol\beta}_C\end{gathered} \tag{3.3.2.6}$$

ifadesine eşittir. Burada $\gamma_H'\mathbf{P}_{\Omega(C)}\mathbf{Y} = \gamma_H'\hat{\mathbf{Y}}_C = \mathbf{H}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\hat{\mathbf{Y}}_C = \mathbf{H}'\hat{\boldsymbol{\beta}}_C$ gerçeğini kullandık. Bu nedenle $\mathbf{Y}'\mathbf{P}_{\Pi(C,H)}\mathbf{Y}$, $\hat{\boldsymbol{\beta}}_C$'ya göre bir kuadratik form olarak yazılabilir ve bu formun $\mathbf{H}\left\{\mathbf{H}'\mathbf{A}_C(\mathbf{X}'\mathbf{X})^{-1}\mathbf{H}\right\}^{-1}\mathbf{H}'$ çekirdeği sadece *txt* ve *sxs* matrislerinin ($\mathbf{X}'\mathbf{X}$ hariç) terslerini ihtiva eder. (3.3.2.6) ve (3.2.2.1) arasındaki biçim benzerliğine dikkat edelim. Bunların her ikisi de model için parametre vektörünün tahminine göre kuadratik formlardır.

Kısıtlanmış hata kareler toplamı, yani, HKT(con) için aşağıdaki eşitlik ortaya çıkar.

$$\text{HKT(con)}= \left\|\mathbf{Q}_{\Omega(C)}\mathbf{Y}\right\|^2 = \mathbf{Y}'\mathbf{Q}_{\Omega(C)}\mathbf{Y}$$

$$\boldsymbol{\Omega}(\mathbf{C})^{\perp} = \gamma_C \oplus \boldsymbol{\Omega}^{\perp}$$

olduğundan,

$$\text{HKT(con)}= \left\|\mathbf{Q}_{\Omega(C)}\mathbf{Y}\right\|^2 = \left\|\mathbf{P}_{\gamma(C)}\mathbf{Y}\right\|^2 + \left\|\mathbf{Q}_{\Omega}\mathbf{Y}\right\|^2 \tag{3.3.2.7}$$

$$= \mathbf{Y}'\mathbf{P}_{\gamma(C)}\mathbf{Y} + \text{HKT(unc)}$$

dır. $\mathbf{Y}'\mathbf{P}_{\gamma(C)}\mathbf{Y} = \mathbf{Y}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{C}\left\{\mathbf{C}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{C}\right\}^{-1}\mathbf{C}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$ 'nin $E(\mathbf{Y}) = \mathbf{X}\boldsymbol{\beta}$ kısıtlanmamış modelinde, $H_0 : \mathbf{C}'\boldsymbol{\beta} = \mathbf{0}$ hipotezini test etmek için pay kareler toplamı olduğuna dikkat edelim.

3.4. Sıfır Hipotezi Doğru Olduğunda Kısıtlanmış Modelde Tahmin

Bu durumda ortalama (gözlem ortalaması) vektörünün tahmini $\mathbf{Y}$'yi, $\boldsymbol{\omega}(\mathbf{C},\mathbf{H})$ üzerine izdüşürmek suretiyle elde edilir. $\boldsymbol{\omega}(\mathbf{C},\mathbf{H}) \oplus \Pi(\mathbf{C},\mathbf{H}) = \boldsymbol{\Omega}(\mathbf{C})$ olduğundan,

$$\mathbf{P}_{\omega(C,H)} = \mathbf{P}_{\Omega(C)} - \mathbf{P}_{\Pi(C,H)} \tag{3.4.1}$$

dır. (3.3.1.1) ve (3.3.2.5) ifadelerini (3.4.1)'de yerine koyarak $\mathbf{P}_{\omega(C,H)} = \mathbf{X}\left[\mathbf{I} - \mathbf{A}_C(\mathbf{X}'\mathbf{X})^{-1}\mathbf{H}\left\{\mathbf{H}'\mathbf{A}_C(\mathbf{X}'\mathbf{X})^{-1}\mathbf{H}\right\}^{-1}\mathbf{H}'\right]\mathbf{A}_C(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$ elde ederiz. Bu nedenle, $\mathbf{P}_{\omega(C,H)} = \mathbf{X}\left[\mathbf{I} - \mathbf{A}_C(\mathbf{X'X})^{-1}\mathbf{H}\left\{\mathbf{H'A}_C(\mathbf{X'X})^{-1}\mathbf{H}\right\}\mathbf{H'}\right]\mathbf{A}_C(\mathbf{X'X})^{-1}\mathbf{X'Y}$ sıfır hipotezi doğru olduğunda kısıtlanmış modelde ortalama vektörünü (ve $\hat{\boldsymbol{\beta}}_{C,H} = (\mathbf{X'X})^{-1}\mathbf{X'}\hat{\mathbf{Y}}_{C,H}$ parametre vektörünü) tahmin eder. $\boldsymbol{\beta} = \hat{\boldsymbol{\beta}}_{C,H}$'nin $\mathbf{X}\boldsymbol{\beta} = \hat{\mathbf{Y}}_{C,H}$ denkleminin tek çözümü olduğuna ve $\hat{\boldsymbol{\beta}}_{C,H} \in \mathbf{T}^-$ olduğuna dikkat edelim.

3.5. Tahminler Arasındaki İlişkiler

Ortalama vektörünün tahminlerinin her biri önceden hesaplananlardan elde edilebilir. Örneğin,

$$\hat{\mathbf{Y}}_C = \mathbf{P}_{\Omega(C)}\mathbf{Y} = \mathbf{Q}_{\gamma(C)}\mathbf{P}_{\Omega}\mathbf{Y}$$
$$= \mathbf{Q}_{\gamma(C)}\hat{\mathbf{Y}} \qquad \text{(Ek 2 Teorem 3'ün (b) kısmından)}$$

dır. $\hat{\mathbf{Y}}_{C,H} = \mathbf{Q}_{\Pi(C,H)}\hat{\mathbf{Y}}_C$ ve $\hat{\mathbf{Y}}_{C,H} = \mathbf{Q}_{\gamma(C,H)}\hat{\mathbf{Y}}$ olduğu da gösterilebilir. Parametre vektörü için benzer sonuçlar geçerlidir. Bununla beraber, izdüşümler ortogonal değildirler. Örneğin,

$$\hat{\boldsymbol{\beta}}_C = (\mathbf{X'X})^{-1}\mathbf{X'}\hat{\mathbf{Y}}_C = (\mathbf{X'X})^{-1}\mathbf{X'Q}_{\gamma(C)}\hat{\mathbf{Y}} = (\mathbf{X'X})^{-1}\mathbf{X'Q}_{\gamma(C)}\mathbf{X}\hat{\boldsymbol{\beta}}$$

dır. Cebirsel olarak, $(\mathbf{X'X})^{-1}\mathbf{X'Q}_{\gamma(C)}\mathbf{X}$ dönüşümünün $\mathbf{A}_C$'ye, yani, $\mathbf{E}^q$'nun $(\mathbf{X'X})^{-1}\mathbf{C}$ boyunca $\mathbf{C}^-$ üzerine izdüşümüne, eşit olduğu kolayca gösterilebilir. Ayrıca,

$$\hat{\boldsymbol{\beta}}_{C,H} = (\mathbf{X'X})^{-1}\mathbf{X'Q}_{\Pi(C,H)}\mathbf{X}\hat{\boldsymbol{\beta}}_C$$

ve

$$\hat{\boldsymbol{\beta}}_{C,H} = (\mathbf{X'X})^{-1}\mathbf{X'P}_{\omega(C,H)}\mathbf{X}\hat{\boldsymbol{\beta}}$$

dır. **T=(C,H)** olmak üzere, $(\mathbf{X'X})^{-1}\mathbf{X'P}_{\omega(C,H)}\mathbf{X}$, $(\mathbf{X'X})^{-1}\mathbf{T}$ boyunca, $\mathbf{T}$ üzerine izdüşümken, $\mathbf{D}_{C,H} = (\mathbf{X'X})^{-1}\mathbf{X'Q}_{\Pi(C,H)}\mathbf{X} = \mathbf{I} - \mathbf{A}_C(\mathbf{X'X})^{-1}\mathbf{H}\left\{\mathbf{H'A}_C(\mathbf{X'X})^{-1}\mathbf{H}\right\}^{-1}\mathbf{H'A}_C$ dönüşümü $\mathbf{A}_C(\mathbf{X'X})^{-1}\mathbf{H}$ boyunca $(\mathbf{X'X})^{-1}\mathbf{C}+\mathbf{T}^-$ üzerine izdüşüm (ortogonal değil) dür (Bu izdüşümlerin daha öte incelemesi için bkz. Herr 1978). Altı tahminin hesabı için etkin bir şema aşağıda görülür. ($\mathbf{A}_C\hat{\boldsymbol{\beta}}_C = \hat{\boldsymbol{\beta}}_C$ olduğundan, $\mathbf{F}_{C,H} = \mathbf{I} - \mathbf{A}_C(\mathbf{X'X})^{-1}\mathbf{H}\left\{\mathbf{H'A}_C(\mathbf{X'X})^{-1}\mathbf{H}\right\}^{-1}\mathbf{H'}$ olmak üzere $\mathbf{D}_{C\,H}\hat{\boldsymbol{\beta}}_C = \mathbf{F}_{C\,H}\hat{\boldsymbol{\beta}}_C$ olduğuna dikkat edelim.)

$$\begin{tikzcd}
 & \hat{\mathbf{Y}} & & \\
\mathbf{Y} \arrow[r, "(\mathbf{X'X})^{-1}\mathbf{X'}"] & \hat{\boldsymbol{\beta}} \arrow[u, "\mathbf{X}"'] \arrow[r, "\mathbf{A}_C"] & \hat{\boldsymbol{\beta}}_C \arrow[r, "\mathbf{X}"] \arrow[d, "\mathbf{F}_{C,H}"'] & \hat{\mathbf{Y}}_C \\
 & & \hat{\boldsymbol{\beta}}_{C,H} \arrow[r, "\mathbf{X}"] & \hat{\mathbf{Y}}_{C,H}
\end{tikzcd}$$

1. $\mathbf{C'\boldsymbol{\beta}} = \mathbf{m} \neq \mathbf{0},\ \mathbf{H'\boldsymbol{\beta}} = \mathbf{h} \neq \mathbf{0}$ DURUM

Bu kısımda (3.1.3) sisteminin (3.1.4) gibi yeniden nasıl yazılabileceğini göstereceğiz.

$\boldsymbol{\alpha}_n(\boldsymbol{\alpha}_m)$, $\mathbf{H'\boldsymbol{\beta}} = \mathbf{h}$ $(\mathbf{C'\boldsymbol{\beta}} = \mathbf{m})$ denklemi herhangi bir çözüm olsun. Bu denklemin tüm çözümlerinin kümesi aşağıdaki gibi tanımlanabilir: $\mathbf{H'\boldsymbol{\beta}} = \mathbf{h}$ $(\mathbf{C'\boldsymbol{\beta}} = \mathbf{m})$ olması için gerek ve yeter şart $\boldsymbol{\beta} \in \boldsymbol{\alpha}_n + \mathbf{H}^{\perp}(\boldsymbol{\alpha}_m + \mathbf{C}^{\perp})$ olmasıdır. Bu nedenle,

$$S = \{\boldsymbol{\beta} : \mathbf{C'\boldsymbol{\beta}} = \mathbf{m} \ \ ve \ \ \mathbf{H'\boldsymbol{\beta}} = \mathbf{h}\} = \left(\boldsymbol{\alpha}_m + \mathbf{C}^{\perp}\right) \bigcap (\boldsymbol{\alpha}_h + \mathbf{H}^{\perp})$$

dır. Ek 2'nin Teorem 6'sına göre S, boş değildir. Gerçekten, herhangi bir $\boldsymbol{\alpha}_0 \in S$ için $S = \boldsymbol{\alpha}_0 + (\mathbf{C}^{\perp} \cap \mathbf{H}^{\perp})$ ile tanımlanabilir. $\boldsymbol{\beta}_0$, S'de herhangi bir vektör olsun. Bu taktirde $\mathbf{c'\boldsymbol{\beta}}_0 = \mathbf{m}$ ve $\mathbf{H'\boldsymbol{\beta}}_0 = \mathbf{h}$ dır. (3.1.3) bağıntısında $\tilde{\boldsymbol{\beta}} = \boldsymbol{\beta} - \boldsymbol{\beta}_0$ ve $\tilde{\mathbf{Y}} = \mathbf{Y} - \mathbf{X}\boldsymbol{\beta}_0$ koyarak (3.1.4) bağıntısını elde ederiz.

Ek 1 :Ortogonal Ayrışımlar Ve Ortogonal İzdüşümler

M ve **N** k-boyutlu $\mathbf{E}^k$ öklid uzayının yalnız sıfır vektörü ortak olan altuzayları olsun. Eğer her bir $\mathbf{v} = \mathbf{E}^k$ için $\mathbf{v} = \mathbf{m} + \mathbf{n}$ olacak şekilde $\mathbf{m} \in \mathbf{M}$ ve $\mathbf{n} \in \mathbf{N}$ vektörleri varsa (diğer bir deyişle **M** ve N'nin boyutları toplamı **k** ise), bu taktirde, $\mathbf{E^k} = \mathbf{M} + \mathbf{N}$'ye $\mathbf{E^k}$'nın bir ayrışımı denecektir.

(') transpozeyi göstermek üzere, eğer $\mathbf{v'u} = \mathbf{u'v} = \mathbf{0}$ ise, $\mathbf{E^k}$'nın **u** ve **v** vektörlerine ortogonaldir denir ve bu durum $\mathbf{u} \perp \mathbf{v}$ şeklinde yazılacaktır.

Eğer her $\mathbf{m} \in \mathbf{M}$ ve $\mathbf{n} \in \mathbf{N}$ için $\mathbf{m} \perp \mathbf{n}$ (veya $\mathbf{m'n} = \mathbf{n'm} = \mathbf{0}$) ise, bu takdirde, **M** ve **N** altuzayları ortogonaldırlar. Bu durumda $\mathbf{M} \perp \mathbf{N}$ yazılır.

Eğer $\mathbf{E^k} = \mathbf{M} + \mathbf{N}$, $\mathbf{E^k}$'nın bir ayrışımı ve $\mathbf{M} \perp \mathbf{N}$ ise, bu taktirde, $\mathbf{E^k} = \mathbf{M} \oplus \mathbf{N}$ yazarız ve $\mathbf{M} \oplus \mathbf{N}$'ye ortogonal ayrışım durumu olarak bakarız. **N**'ye **M**'nin ortogonal tümleyeni denir ve $\mathbf{N} = \mathbf{M}^{\perp}$ olarak yazılır.

Eğer $\mathbf{P}^2 = \mathbf{P}$ (yani **P** idempotent) ise, bu takdirde, **P** *k*x*k* matrisi bir izdüşümdür (izdüşürücüdür). Her bir **P** izdüşümü için , $\mathbf{m} \in \mathbf{M}$ için $\mathbf{P_m} = \mathbf{m}$ ve $\mathbf{n} \in \mathbf{N}$ için $\mathbf{P_n} = \mathbf{0}$ olacak şekilde bir $\mathbf{E^k} = \mathbf{M} + \mathbf{N}$ ayrışımı vardır. **M**, **P**'nin sütun uzayıdır ve **N**, **P**'nın sıfır uzayıdır (**M**, **P**'nin sütunları tarafından gerilir; S, **P**'nin satırları tarafından gerilen uzay olmak üzere, $\mathbf{N} = S^{-}$ dir) Bu nedenle her izdüşüm, kendi sütun uzayı üzerinde birim ve kendi sıfır uzayı üzerinde **0** rolünü oynar. **P**'ye **N** boyunca **M** üzerine izdüşüm denir. Eğer **P** bir ortogonal izdüşüm ise (yani, $\mathbf{P}^2 = \mathbf{P} = \mathbf{P'}$ yani **P** idempotent ve simetrik ise), bu taktirde, $\mathbf{m} \in \mathbf{M}$ için $\mathbf{P_m} = \mathbf{m}$ ve $\mathbf{n} \in \mathbf{N}$ için $\mathbf{P_n} = \mathbf{0}$ olacak şekilde bir ortogonal $\mathbf{E^k} = \mathbf{M} \oplus \mathbf{N}(\mathbf{N} = \mathbf{M}^{\perp})$ ayrışımı vardır ve **P**'ye **M** üzerine ortogonal izdüşüm denir. ($\mathbf{M}^{\perp}$ boyunca **M** üzerine ortogonal izdüşüm) Her izdüşüm kendi sıfır uzayı ve kendi sütun uzayı ile tam olarak tanımlanır.

Ek 2: Teoremlerin İfadeleri

Teorem 1: Eğer S, $\mathbf{E^n}$'nin bir *m*-boyutlu altuzayı ve **M**, sütunları S için bir taban oluşturan, bir *nxm* matris ise, bu taktirde, $\mathbf{E^n}$nin S üzerine ortogonal izdüşümü $\mathbf{M(M'M)^{-1}M'}$ dür.

Teorem2: Kısım 3.1' nin $\mathbf{C'\beta} = \mathbf{0}$, $H_0 : \mathbf{H'\beta} = \mathbf{0}$, $\mathbf{\Omega_C} \oplus \gamma_C = \mathbf{\Omega}$, $\omega_{C,H} \oplus \Pi_{C,H} = \mathbf{\Omega_C}$ ile beraber $E(\mathbf{Y}) = \mathbf{X\beta}$ varsayımlarına bağlı kısıtlanmış modelde hipotez testi problemine yön vermektir. Belirtilen altuzaylara kısıtlandığında, aşağıdaki altı dönüşümün tümü bire bir ve üzerinedir. Bundan başka, bir çiftteki her bir dönüşüm diğerinin tersidir.

$$\mathbf{X}' : \gamma_{C} \rightleftarrows \mathbf{C} : \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}$$

$$(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' : \omega_{C,H} \rightleftarrows (\mathbf{C}+\mathbf{H})^{\perp} : \mathbf{X}$$

$$(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' : \prod_{C,H} \rightleftarrows \mathbf{A}_{C}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{H} : \mathbf{X}$$

dir.

Teorem 3: $\mathbf{D} = \mathbf{A} \oplus \mathbf{B}$ olmak üzere **D** keyfi bir **V** vektör uzayının bir altuzayı olsun. Bu takdirde,

a) $\mathbf{P}_{D} = \mathbf{P}_{A} + \mathbf{P}_{B}$ ve

b) $\mathbf{Q}_{A}\mathbf{P}_{D} = \mathbf{P}_{D}\mathbf{Q}_{A} = \mathbf{P}_{B}$

dır.

Teorem 4: **D=A+B** olmak üzere, (**A** ve **B** ortogonal değildirler.) **D** keyfi herhangi bir **V** vektör uzayının bir altuzayıdır. Bu takdirde $\mathbf{Q}_{A}(\mathbf{B}) = \{\mathbf{Q}_{A}\mathbf{b} \mid \mathbf{b} \in \mathbf{B}\}$ olmak üzere, $\mathbf{D} = \mathbf{A} \oplus \mathbf{Q}_{A}(\mathbf{B})$ dir ve $\mathbf{Q}_{A}$, V nin $\mathbf{A}^{\perp}$ üzerine ortogonal izdüşümüdür.

Teorem 5: **X** ve **C** sırasıyla her biri tam sütun ranklı *n*x*q* ve *q*x*t* boyutlu matrisler olsun. Bu takdirde *q*x*q* boyutlu $\mathbf{A}_{C} = \mathbf{I} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{C}\{\mathbf{C}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{C}\}^{-1}\mathbf{C}'$ matrisi sütun uzayı $\mathbf{C}^{\perp}$ ve sıfır uzayı $(\mathbf{X}'\mathbf{X})^{-1}\mathbf{C}$ olan izdüşümdür (ortogonal değil).

Teorem 6: **C** ve **H**, **C**'nin sütunları **H**'nın sütunlarından lineer bağımsız olmak üzere, sırasıyla, *q*x*t* ve *q*x*s* boyutlu matrisler olsun. **m** ve **h** sırasıyla **C** ve **H**'nin sütun uzaylarındaki vektörler olsun. $\boldsymbol{\alpha}_{m}$ ve $\boldsymbol{\alpha}_{h}$, $\mathbf{C}\boldsymbol{\alpha}_{m} = \mathbf{m}$ ve $\mathbf{H}\boldsymbol{\alpha}_{h} = \mathbf{h}$ olacak şekilde vektörler olsun. Bu taktirde

(a) $\mathbf{C}^{\perp} + \mathbf{H}^{\perp} = (\mathbf{C} \cap \mathbf{H})^{\perp} = \mathbf{E}^{q}$,

(b) $\mathbf{S} = (\boldsymbol{\alpha}_{m} + \mathbf{C}^{\perp}) \cap (\boldsymbol{\alpha}_{h} + \mathbf{H}^{\perp})$ boş değildir ve herhangi bir $\boldsymbol{\alpha}_{0} \in \mathbf{S}$ için,

(c) $\mathbf{S} = \boldsymbol{\alpha}_{0} + (\mathbf{C}^{\perp} \cap \mathbf{N}^{\perp})$ dir.

SONUÇ VE ÖNERİLER

Üç bölümden oluşturduğumuz çalışmamızda Lineer modellerde kullanılan matris cebiri ve özellikle matrislerin genelleştirilmiş terslerini ele aldık. Daha sonra Lineer Modeller ve Tahmin Edebilirlik ile ilgili temel kavramlara yer vererek bir ön bilgi oluşturduk. Çalışmamızın asıl konusu olan Lineer Modellerde Eşitlik ve Eşitsizlik Kısıtlamaları Altında Parametre Tahminleri ve Bazı Hipotezlerin Testlerini ayrıntılı bir şekilde incelemeye çalıştık. Kısıtlanmış Lineer Modelde Tahmin ve Hipotez Testinin Geometrisi-Tam Ranklı Durum mevzusunun incelenmiş olduğu yabancı dilde bir makale dilimize çevrilerek bilim dünyasına kazandırılmıştır. Yine bu makalede göze çarpan Kısıtlanmış ve Kısıtlanmamış Lineer Modelde tahminler konusunu inceleyerek çalışmaya ekledik. Tahminler arasındaki ilişkilere de yer vererek çalışmadan çıkabilecek bulgular ortaya konuldu.

KAYNAKLAR

Burdick, D. S., Herr, D.G., O' Fallon, W.M., and O'Neill, B.V. (1974). "Exact Medhods in the Unbalanced, Two-Way Analysis of Variance-A Geometric Approach," Commun. Statist. 3, 581-594

Charles J. Monlezun, F.M. Speed. (1980). "The Geometry of Estimation and Hypothesis Testing in the constrained linear model- the full-rank case", Communications in Statistics-Theory and Medhods 9:2 213-230.

Coleman, A.T. (1977). "The Cell Means Approach to Analysis of Variance of Large Experiment." Master's thesis, Mississippi State University, Mississippi State, Mississippi.

Drazin, M.P. (1978). Natural structures on semigroups with involutions, Bull. Amer. Math. Soc. 84:139–141.

Graybill, Fraklin A. (1976). Theory and Application of the Linear Model. North Scituate: Duxbury Press.

Halmos, Paul R. (1958). Finite Dimensional Vektor Spaces. Princeton: D. Van Nostrand Co., Inc.

Hartwig, R.E. (1980). How to order regular elements? Mathematica Japonica, 25:1–13.

Haynsworth, E.V. (1968). Determination of the inertia of a partitioned Hermitian matrix, Linear Algebra and Its Applications, 1:73–82.

Hedayat, A. andWallis,W.D. (1978). Hadamard matrices and their applications, Annals of Statistics, 6:1184–1238.

Herr, David G. (1978). "A Version of the Geometry of Estimation and Hypothesis Testing in the Constrained Linear Model-the Full Rank Case", Personel communication.

Herr, David G. (1979). "On the History of the Use of Geometry in the General Linear Model", The American Statistician.

Hoffman, Kenneth and Kunza, Ray (1971). Linear Algebra. Englewood Cliffs: Prentice-Hall, Inc.

Ikebe, Y., Inagaki, T., and Miyamoto, S. (1987). The monotonicity theorem, Cauchy interlace theorem and the Courant–Fischer theorem, American Mathematical Monthly, 94(4):352–354.

Kruskal, William (1960). “The Coordinate-free Approach to Gaoss-Markoff Estimation, and its Application to missing and Extra Observation”, Proc. Fourth Berkeley Symp. Math. Stat. And Prob., 1, 435-451

Mitra, S.K. (1982). Simultaneous diagonalization of rectangular matrices, Linear Algebra and Its Applications, 47:139–150.

Nordstr¨om, K. (1989). Some further aspects of the L¨owner—ordering antitonicity of the Moore–Penrose inverse, Communications in Statistics—Theory and Methods, 18(12):4471–4489.

R.B. Bapat, (2000). “Linear Algebra and Linear Models”, Second Edition Springer Verlag New York, Inc., and Hindustan Book Agency.

Scheffe, Henry (1959). The Analysis of Variance. New York: John Wiley and Sons.

Searle, S.R. (1971). Linear Models. New York: John Wiley and Sons.

Seber, G.A.F. (1977). Linear Regression Anayysis. New York: John Wiley and Sons.

CANAN KOÇAK

Doğum Tarihi : 01.03.1988

Doğum Yeri : Yozgat

2011 - (Devam ediyor)

Ordu Üniversitesi Fen Bilimler Enstitüsü Matematik Bölümü Uygulamalı Matematik Anabilim Dalı Öğrencisi (Ordu)

2006 – 2010 Ordu Üniversitesi Fen-Edebiyat Fakültesi Matematik Bölümü Lisans (Ordu)