T.C.

ORDU ÜNİVERSİTESİ

FEN BİLİMLERİ ENSTİTÜSÜ

LİNEER MODELLERDE PARAMETRE TAHMİNLERİ VE KANONİK KORELASYONLAR

ZEYNEP ALBAYRAK

YÜKSEK LİSANS TEZİ

MATEMATİK ANABİLİM DALI

AKADEMİK DANIŞMAN

Prof. Dr. Cemil YAPAR

ORDU-2010

T.C.
ORDU ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

Bu çalışma jürimiz tarafından 13/08/2010 tarihinde yapılan sınav ile Matematik Anabilim Dalı'nda YÜKSEK LİSANS tezi olarak kabul edilmiştir.

Başkan : Prof. Dr. Cemil YAPAR

Üye : Prof. Dr. İhsan ÜNVER

Üye : Yrd. Doç. Dr. Süleymen ŞENYURT

ONAY :

Yukarıdaki imzaların adı geçen öğretim üyelerine ait olduğunu onaylarım.

..../..../2010

Yrd. Doç. Dr. BeyhanTAŞ
Fen Bilimleri Enstitüsü Müdürü

ÖZET

Bu tezde, lineer modellerde parametre tahminleri ve kanonik korelasyonlar ele alındı. Bu tez çalışması üç bölümden oluşmaktadır. Birinci bölümde, tez çalışmasının alt yapısı için gerekli görülen temel tanım ve teoremler verilmiştir. İkinci bölümde, $Y : nx1$ gözlenebilir rasgele bir vektör, $X: nxp(n > p)$ reel sayılar matrisi, $\beta: px1$ bilinmeyen parametrelerin bir vektörü ve ε, $E(\varepsilon) = 0$, $Cov(\varepsilon) = \sigma^2 V$ olacak şekilde, rasgele değişkenlerin gözlenebilir olmayan bir hata vektörü olmak üzere $Y = X\beta + \varepsilon$ lineer modelinin tanımı yapılarak, bu modelin ε'nun dağılımına, V varyans-kovaryans matrisine ve X'in rankına bağlı durumları incelenmiştir. Bununla beraber, β parametre vektörü üzerine $R\beta = r$ tutarlı kesin lineer kısıtlaması altında β'nın en iyi lineer yansız tahmin edicisi orjinal olarak, çeşitli yöntemlerle ve farklı bakış açılarıyla elde edilmiştir. Kesin lineer kısıtlamaların adım adım hesaba katılması ile parametre tahminleri karşılaştırmalı olarak incelenmiştir. β parametre vektörü üzerinde hipotez testi test edilmiş ve β nın bileşenleri ve bunların lineer parametrik fonksiyonları için güven aralıkları oluşturulmuştur.

Üçüncü bölümde, korelesyon ölçüleri ve çok değişkenli bir istatistik analiz yöntemi olan, birden fazla değişkeni iki alt kümeye ayırıp doğrusal bileşenlerine indirgeyerek değişkenler arasındaki ilişkinin yorumlanmasında birçok kolaylık sağlayan kanonik korelasyon analizi ve Kernel kanonik korelasyon analizi incelenmiştir. Bu bölümde, ayrıca en küçük kareler uygun değerleri ve artıklar(hatalar) arasındaki kanonik korelasyonların özellikleri ele alınmıştır.

Anahtar Kelimeler: *Lineer Modeller, Parametre Tahmini, Kesin Lineer Kısıtlama, Korelasyon, Kanonik Korelasyon Analizi, Kernel Kanonik Korelasyon.*

ABSTRACT

In this thesis, parameter estimations in linear models and canonical correlations have been considered. This thesis consists of three chapters. In the first chapter, fundamental definitions and theorems used in the thesis have been given.

In the second chapter, linear model is defined as follows:

$$Y = X\beta + \varepsilon,$$

where Y is an $nx1$ vector of observable random variables, X is an $nxp(n > p)$ matrix of known explanatory variables, β is a $px1$ vector of unknown parameters (nonstochastic) and ε is an nx1 unobservable random vector such that $E(\varepsilon) = 0$, $Cov(\varepsilon) = \sigma^2 V$. This model has been examined according to distribution of ε, variance- covariance matrix V and rank of the matrix X. Furthermore, when exact linear restriction $R\beta = r$ (this is consistent) is imposed on the parameter vector β, the best linear unbiased estimator of β has been orjinally obtained by several methods and has been considered in the different forms.

Parameter estimations have been obtained comparatively with stepwise inclusion of exact linear restrictions. Various hypotheses have been developed on the parameter vector β and theses hypotheses have been tested. For the components of the parameter vector β and the linear parametric functions of these components have been constructed confidence intervals.

In the third chapter, corelation measurements, canonical correlation analysis which is one of multivariate statistical analysis methods and Kernel canonical correlation analysis have been examined. In this chapter, properties of the canonical correlations between the least squares fitted values and the residuals also have been considered.

Key words : *Linear Models, Parameter Estimation, Exact Linear Restriction, Correlation, Canonical Corelation Analysis, Kernel Canonical Correlation.*

TEŞEKKÜR

Bu tezin hazırlanmasında kıymetli zamanlarını bana ayırarak büyük katkılarda bulunan danışmanım Sayın Prof. Dr. Cemil YAPAR'a, çalışmalarım sırasında bilimsel yaklaşımı kendisinden öğrenmeye çalıştığım değerli hocalarım Sayın Yrd. Doç. Dr. Selahattin MADEN'e ve Sayın Yrd. Doç. Dr. Süleyman ŞENYURT'a, maddi ve manevi desteğini esirgemeyen AİLEME sonsuz teşekkürlerimi ve saygılarımı sunarım.

İÇİNDEKİLER

GİRİŞ

Matrisler ve istatistiksel kavramlar ile ilgili temel tanım ve teoremler çeşitli kaynaklardan derlenerek genel bilgiler kısmında ele alınmıştır.

Lineer modeller, bağımlı değişken ile bir veya daha fazla bağımsız değişken arasındaki ilişkiyi incelemek amacıyla kullanılan bir istatistiksel modeldir. Bir tek bağımsız değişkenin kullanıldığı model "basit lineer model"; birden fazla bağımsız değişkenin kullanıldığı model de "çok değişkenli lineer model" olarak adlandırılır. Bu çalışmada, çok değişkenli lineer model matris formunda yazılmıştır. Yani, y; $x_1, x_2, \dots, x_p$ değişkenlerinin $i-yinci$ gözlem değerleri için $y_i, x_{i1}, x_{i2}, \dots, x_{ip}$ olsun. $y_1, y_2, \dots, y_n$ gözlenebilen rasgele değişkenler için model

$$y_i = \beta_0 + \sum_{j=1}^{p} x_{ij}\,\beta_j + \varepsilon_i, \qquad i = 1,2,\dots n$$

dir. Burada $\beta_0, \beta_1, \dots, \beta_p$ bilinmeyen parametreleri, ε_i ler ortalamaları 0, bilinmeyen varyansları σ^2V olan rasgele değişkenlerdir. Bu model, aşağıdaki gibi yazılabilir.

$$\mathrm{y}_1 = \beta_0 + \mathrm{x}_{11}\beta_1 + \mathrm{x}_{12}\beta_2 + \cdots + \mathrm{x}_{1p}\beta_p + \varepsilon_1,$$

$$\mathrm{y}_2 = \beta_0 + \mathrm{x}_{21}\beta_1 + \mathrm{x}_{22}\beta_2 + \cdots + \mathrm{x}_{2p}\beta_p + \varepsilon_2,$$

$$\vdots \qquad \vdots \qquad \vdots \qquad \vdots \qquad \vdots \qquad \vdots$$

$$\mathrm{y}_n = \beta_0 + \mathrm{x}_{n1}\beta_1 + \mathrm{x}_{n2}\beta_2 + \cdots + \mathrm{x}_{np}\beta_p + \varepsilon_n.$$

Yukarıdaki modeli matris formunda

$$\begin{bmatrix} \mathrm{y}_1 \\ \mathrm{y}_2 \\ \vdots \\ \mathrm{y}_n \end{bmatrix} = \begin{bmatrix} 1 & \mathrm{x}_{11} & \cdots & \mathrm{x}_{1p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \mathrm{x}_{n1} & \cdots & \mathrm{x}_{np} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

veya kısaca,

$$Y = \begin{bmatrix} \mathrm{y}_1 \\ \mathrm{y}_2 \\ \vdots \\ \mathrm{y}_n \end{bmatrix}, X = \begin{bmatrix} \mathrm{x}_{11} & \cdots & \mathrm{x}_{1p} \\ \vdots & \ddots & \vdots \\ \mathrm{x}_{n1} & \cdots & \mathrm{x}_{np} \end{bmatrix}, \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}, \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{bmatrix}$$

olmak üzere,

$$Y = X\beta + \varepsilon$$

olarakta yazılır. Literatürde ε vektörünün, V ve X matrislerinin durumlarına göre parametre tahminleri ele alınmıştır.

Bu çalışmada, tahminler çeşitli yöntemler ve farklı bakış açılarıyla incelendi. Yapar, C. (1979), β parametre vektörü üzerine tutarlı kesin lineer kısıtlamalara bağlı $Y = X\beta + \varepsilon$ lineer modelini göz önüne alarak en iyi lineer yansız tahmin ediciyi farklı bir yöntemle elde etmiştir. Burada, kesin lineer kısıtlama tanıtıldı ve kesin lineer kısıtlama altında parametre tahmini ele alınarak bu tahminin çeşitli yöntemlerle nasıl elde edilebileceği ortaya konuldu ve bazı karşılaştırmaları incelendi. Stokastik kısıtlamalar altında tahminler incelendi. Ayrıca, β parametre vektörü üzerine konulan çeşitli hipotezler test edildi ve β parametre vektörünün bileşenleri ve bunların lineer parametrik fonksiyonları için güven aralıkları oluşturuldu.

Kanonik korelasyon analizi, çok değişkenli lineer modelin bir uzantısıdır. Çok değişkenli lineer modelde X bağımsız değişkeni bir veya birden fazla değişken, Y bağımlı değişkeni bir tane değişken içerirken, kanonik korelasyon analizinde Y değişkeni birden fazla değişkeni içermektedir.

Kanonik korelasyon analizi çok değişkenli bir istatistik analiz yöntemidir ve 1935–1936 yıllarında Hotelling tarafından geliştirilmiştir. Hotelling kanonik korelasyon analizini, psikolojide zeka testleri ve fiziksel değişkenler arasındaki ilişkilerin ölçülmesinde kullanmıştır. İlk anda iki değişken kümesi arasındaki ilişkiyi tanımlayan karmaşık bir yol olarak görünmesine rağmen çok sayıda değişkeni iki alt kümeye ayırıp az sayıda doğrusal bileşenlerine indirgeyerek değişkenler arasındaki ilişkinin yorumlanmasında birçok kolaylık sağlamaktadır ve iki değişken küme arasındaki en büyük ilişkiyi araştırmaya çalışır.

Barıtçı,İ. (2001) 3 ve 6 aylık kilis keçisi oğlaklarında doğumdaki vücut ölçüleri arasındaki ilişkileri kanonik korelasyon yöntemiyle araştırmış. Çankaya,S. (2005) kanonik korelasyon analizini ve hayvancılıkta kullanımını ve Mirtaghizadeh,H. (1990), Türkiye'nin sosyo-ekonomik yapısını kanonik korelasyon analizi ile belirlemeye çalışmıştır. Özel,H. (1984), araştırmasında, Birleşmiş Milletler ve bağlı kuruluşlarına ait yayınlardan derlenen 33 ülkeye ilişkin 32 adet değişken verilerini değerlendirerek ekonomik kalkınma ve eğitim arasındaki ilişkiyi kanonik korelasyon analizi kullanarak incelemiştir.

Bu çalışmada, kanonik korelasyon analizi matematiksel yaklaşım ile ele alındı. Aynı zamanda çok boyutlu bir kitleden seçilmiş olan iki veri kümesi arasındaki kanonik korelasyonun elde edilmesi ve anlamlılık testleri anlatıldı. Welling tarafından verilen Kernel kanonik korelasyon analizi incelendi.

Son olarak, Puntanen,S. (1985), $Y = X\beta + \varepsilon$ modelinde $E(Y) = X\beta$, $Kov(Y) = \sigma^2 V$ ve H, X in sütun uzayı üzerine ortogonal izdüşürücü ve $M = I - H$ olduğunda, HY uygun değerleri ve MY artıkları arasındaki kanonik korelasyonların özelliklerini inceledi. Bu çalışmada ise, bu makaledeki bilgiler esas alınarak bazı genişletmeler yapıldı.

BÖLÜM 1

GENEL BİLGİLER

1.1. Matris Cebiri

Tanım 1.1.1: V boş olmayan bir küme olsun ve V üzerinde toplama ile skalerle çarpma işlemleri,

$$\oplus : VxV \to V \quad \text{ve} \quad \odot : \mathbb{R}xV \to V \tag{1.1.1}$$

aşağıdaki özellikleri sağlarsa, V ye $\mathbb{R}$ cismi üzerinde bir vektör uzayı denir.

i. $(V,\oplus)$ bir abel grubu,

ii. $\forall u, v \in V$ ve $a, b \in \mathbb{R}$ için

- **a.** $(a.b) \odot v = a \odot (b \odot v)$
- **b.** $(a + b) \odot v = (a \odot v) \oplus (b \odot v)$
- **c.** $a \odot (u \oplus v) = (a \odot u) \oplus (a \odot v)$
- **d.** $1 \odot v = v$.

Tanım 1.1.2: V, W vektör uzayları olmak üzere,

$$L: V \to W$$

$$v \to L(v) \tag{1.1.2}$$

fonksiyonu her $u, v \in V$ ve $a, b \in \mathbb{R}$ için

$$L(av + bu) = aL(v) + bL(u) \tag{1.1.3}$$

özelliğine sahip ise, L fonksiyonuna V den W ye bir lineer dönüşüm,

$$R(L) = \{w : w \in W, \exists v \in V \; için \; w = L(v)\} \tag{1.1.4}$$

kümesine L lineer dönüşümün değer kümesi ve

$$N(L) = \{v: v \in V, L(v) = 0\} \tag{1.1.5}$$

kümesine L lineer dönüşümün çekirdeği denir.

Tanım 1.1.3: A matrisinin sütun uzayı, $\mathcal{R}(A)$, A'nın sütun vektörleri tarafından gerilen vektör uzayıdır, yani,

$$\mathcal{R}(A) = \left\{z: z = Ax = \sum_{i=1}^{p} a_{(i)} x_i, x \in \mathbb{R}^p\right\} \subset \mathbb{R}^n \tag{1.1.6}$$

dir. Burada, $a_{(1)}, a_{(2)}, \dots, a_{(p)}$ A'nın sütun vektörleridir. A matrisinin sıfır uzayı $\mathcal{N}(A)$,

$$\mathcal{N}(A) = \{x \in \mathbb{R}^p \; ve \; Ax = 0\} \subset \mathbb{R}^p \tag{1.1.7}$$

ile tanımlanan vektör uzayıdır.

Bir V vektör uzayını geren lineer bağımsız vektörler kümesine V'nin tabanı denir. Bir vektör uzayının birden fazla tabanı olabilir.

Teorem 1.1.1: V boştan farklı bir vektör uzayı, A, B ve C uygun boyutlu matrisler ve $\mathcal{R}(A)$, A nın sütun uzayı olsun. Bu takdirde,

i) $boyV$, bir V vektör uzayının taban vektörlerinin sayısını göstermek üzere $rank(A) = boy\mathcal{R}(A)$ dır.

ii) $\mathcal{N}(A) = \{\mathcal{R}(A')\}^{\perp}$ dir. ($V^{\perp}$, $V^{\perp} = \{x: x'y = 0, \; \forall y \in V\}$ ile tanımlanan bir V vektör uzayının ortogonal tümleyenidir.)

iii) $\mathcal{R}(AA') = \mathcal{R}(A)$ dır.

iv) Bir C matrisi için $AC = B \Leftrightarrow \mathcal{R}(B) \subset \mathcal{R}(A)$ dır.

v) Her A ve B için $\mathcal{R}(AB) \subset \mathcal{R}(A)$ dır. Eğer B singüler(tekil) değilse, $\mathcal{R}(AB) = \mathcal{R}(A)$ dır.

vi) Her $A \geq 0$ ve her B için $\mathcal{R}(BAB') = \mathcal{R}(BA)$ dır. Burada $A \geq 0$ gösterimi yani, A nın pozitif kararlılığı veya negatif olmayan kararlılığı, daha ileride tanımlanacaktır.(Rao, C. R. and Toutenburg, H.,1999)

Tanım 1.1.4: $A: nxp$ matrisi, n satır ve p sütunda elemanların bir dikdörtgensel dizilimidir. Genellikle bu elemanlar reel sayılar olarak seçilir. $A: nxp$ matrisi

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1p} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{np} \end{bmatrix} = (a_{ij}) \tag{1.1.8}$$

ile gösterilir. (1.1.8) matrisinde a_{ij} reel sayısına A'nın $ij-yinci$ elemanı (yani,$i-yinci$ satır, $j-yinci$ sütun elemanı) denir. Satır ve sütun sayıları eşit olan bir matrise kare matris denir. Eğer $A: nx1$ matris ise, A ya n-boyutlu bir vektör denir. $A=(a_{ij})$ ve $B=(b_{ij}):nxp$ matrisler olsun. A ve B nin toplamı $A+B$ şeklinde yazılan $C=(c_{ij})$ matrisidir. Burada $c_{ij}=a_{ij}+b_{ij}$ dir. s reel bir sayı (skaler) olsun. s ile A'nın çarpımı sA şeklinde yazılan $D=(d_{ij})$ matrisidir. Burada $d_{ij}=sa_{ij}$ dir. Toplama ve skalerle çarpımların bazı özellikleri aşağıda verilmiştir

$$A+B=B+A, \qquad (A+B)+C=A+(B+C), \qquad (s+t)A=sA+tA,$$
$$s(A+B)=sA+sB, \qquad (st)A=s(tA)$$

$A: nxp$ ve $B: pxr$ matrislerinin çarpımı AB şeklinde yazılan $C=(c_{ij}): nxr$ matrisidir. Burada $c_{ij}=\sum_{k=1}^{p} a_{ik}b_{kj}$ dir. A'nın sütunlarının sayısı B'nin satırlarının sayısı ile aynı ise, çarpım o zaman tanımlıdır. Aynı zamanda her iki çarpım tanımlı olsa bile, genel olarak $AB \neq BA$ olabilir. Çarpımın önemli özellikleri aşağıda verilmiştir: A, B ve C çarpılabilir matrisler olmak üzere,

$$A(B+C)=(AB)+(AC), \qquad (A+B)C=(AC)+(BC).$$

$A: nxp$ matrisinin transpozesi A' şeklinde yazılan $C=(c_{ij}): pxn$ matrisidir. Burada $c_{ij}=a_{ji}$ dir. Transpozenin özellikleri aşağıda verilmiştir:

$$(A')'=A, \qquad (A+B)'=A'+B', \qquad (AB)'=B'A'.$$

$A: nxn$ kare matris ise, $A'=A$, yani $a_{ij}=a_{ji}$ ise, A matrisine simetrik matris denir. Köşegenleri dışındaki elemanlar sıfır olan bir $A: nxn$ kare matrisine köşegen matris denir. Bu durum aşağıdaki şekilde gösterilir.

$$A=k\ddot{o}\c{s}(a_{ii})=\begin{bmatrix} a_{11} & & 0 \\ & \ddots & \\ 0 & & a_{nn} \end{bmatrix} \tag{1.1.9}$$

Köşegen üzerindeki elemanları 1, köşegen dışıdaki elemanları 0 olan nxn matrise birim matris denir ve $\mathrm{I_n}$ ile gösterilir. Sadece köşegenin altındaki(üstündeki) elemanları sıfır olan A kare matrisine bir üst(alt) üçgensel matris denir.

Tanım 1.1.5: X, n satır ve p sütundan oluşan nxp boyutlu bir matris olsun. X matrisinin elemanları x_{ij} ile gösterilirken X matrisinin sütun vektörleri (değişkenler) $X_1, X_2, \dots, X_j, \dots, X_p$ ile gösterilsin. Değişkenlerin satırlarda, gözlemlerin ise sütunlarda olduğu X veri matrisi aşağıdaki gibi gösterilir. (Alpar, R.,2003)

Çizelge 1.1.1

	DEĞİŞKENLER					
GÖZLEM NO	X_1	X_2	...	X_j	...	X_p
1	x_{11}	x_{12}	...	x_{1j}	...	x_{1p}
2	x_{21}	x_{22}	...	x_{2j}	...	x_{2p}
3	x_{31}	x_{32}	...	x_{3j}	...	x_{3p}
.	.	.	...	.	...	.
i	x_{i1}	x_{i2}	...	x_{ij}	...	x_{ip}
.	.	.	...	.	...	.
n	x_{n1}	x_{n2}	...	x_{nj}	...	x_{np}

Tanım 1.1.6: Bir A matrisinin elemanları alt matrisler halinde düzenlenirse, A matrisine parçalanmış matris denir. $r + s = p$ olmak üzere,

$$\underset{nxp}{A} = \left[\underset{nxr}{A_1} \vdots \underset{nxs}{A_2}\right] \tag{1.1.10}$$

veya

$$\underset{nxp}{A} = \begin{bmatrix} \underset{rxp-s}{A_{11}} & \underset{rxs}{A_{12}} \\ \underset{(n-r)x(p-s)}{A_{21}} & \underset{(n-r)xs}{A_{22}} \end{bmatrix} \tag{1.1.11}$$

ile gösterilecektir.

Parçalanmış matrisler için transpozeleri sırasıyla,

$$A' = \begin{bmatrix} A'_1 \\ \cdots \\ A'_2 \end{bmatrix}, \ A' = \begin{bmatrix} A'_{11} & A'_{21} \\ A'_{12} & A'_{22} \end{bmatrix} \tag{1.1.12}$$

olarak elde edilir.

Tanım 1.1.7: $A : pxq$ ve $B : mxn$ matrislerinin Kronecker(direkt) çarpımı

$$A \otimes B = (a_{ij}B) \tag{1.1.13}$$

ile verilen $pmxqn$ boyutlu bir matristir. Kronecker çarpımın özellikleri,

i) $A \otimes (B \otimes C) = (A \otimes B) \otimes C$ dır.

ii) $(A \otimes B)(C \otimes D) = AC \otimes BD$ dir.

iii) $(A+B) \otimes (C+D) = A \otimes C + A \otimes D + B \otimes C + B \otimes D$ dir.

iv) $(A \otimes B)' = A' \otimes B'$ dır.

v) $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$ dır. **(**Magnus, J. R.,1990)

Tanım 1.1.8: A: nxn kare matrisinin izi köşegen elemanlarının toplamıdır. Yani,

$$izA = \sum_{i=1}^{n} a_{ii} \tag{1.1.14}$$

dir.

Teorem 1.1.2: A ve B: nxn matrisler ve c bir skaler olsun. Bu takdirde aşağıdaki eşitlikler geçerlidir:

i) $iz(A+B) = iz(A) + iz(B)$.

ii) $iz(A') = iz(A)$.

iii) $iz(cA) = c.iz(A)$.

iv) (Devirli özellik) Her iki çarpımın tanımlı olduğunu kabul ederek $iz(AB) = iz(BA)$ dır.

v) $iz(ABc) = iz(cAB) = iz(BcA)$.

vi) A: nxk matrisi olmak üzere, $iz(A'A) = iz(AA') = \sum_{j=1}^{k}\sum_{i=1}^{n} a_{ij}^2$ dir.

vii) Herhangi iki $nx1$ boyutlu a ve b vektörleri için $a'b = \sum a_i b_i = izab'$ dür.

viii) $iz(A \otimes B) = iz(A)iz(B)$ dir.(Rao, C. R. and Toutenburg, H.,1999)

Tanım 1.1.9: $n > 1$ pozitif bir tamsayı olsun. $A{:}\, nxn$ kare matrisinin determinantı

$$|A| = \sum_{i=1}^{n}(-1)^{i+j}a_{ij}\ |M_{ij}| \quad \text{(herhangi bir j için, j sabit)} \qquad (1.1.15)$$

dir. Burada $|M_{ij}|$, a_{ij} elemanın minörüdür. $|M_{ij}|$, A'nın $i-yinci$ satır ve $j-yinci$ sütunu silindiğinde geriye kalan $(n-1)x(n-1)$ matrisinin determinantıdır. $A_{ij} = (-1)^{i+j}|M_{ij}|$ bağıntısına a_{ij} nin kofaktörü (eş çarpanı) denir.

Tanım 1.1.10: Bir A kare matrisin determinantı sıfırdan farklı ise, A matrisine regüler veya singüler(tekil) olmayan matris denir. Aksi halde, A matrisine singüler denir.

Teorem 1.1.3: A ve B : nxn kare matrisler ve c bir skaler olsun. Bu takdirde, aşağıdaki eşitlikler geçerlidir:

i. $|A'| = |A|$.

ii. $|cA| = c^n|A|$.

iii. $|AB| = |A||B|$.

iv. $|A^2| = |A|^2$.

v. Eğer A köşegen veya üçgensel ise, bu takdirde $|A| = \prod_{i=1}^{n} a_{ii}$.

vi. $D = \begin{bmatrix} \underset{nxn}{A} & \underset{nxm}{C} \\ \underset{mxn}{0} & \underset{mxm}{B} \end{bmatrix}$ için $\begin{vmatrix} \underset{nxn}{A} & \underset{nxm}{C} \\ \underset{mxn}{0} & \underset{mxm}{B} \end{vmatrix} = |A|\ |B|$ veya $\begin{vmatrix} A' & 0' \\ C' & B' \end{vmatrix} = |A|\ |B|$.

vii. Eğer A, $A_{11}{:}\, pxp$ ve $A_{22}{:}\, qxq$ kare ve singüler olmayan matrisler olacak şekilde parçalanırsa, bu takdirde

$$\begin{vmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{vmatrix} = |A_{11}||A_{22} - A_{21}{A_{11}}^{-1}A_{12}|$$
$$= |A_{22}||A_{11} - A_{12}{A_{22}}^{-1}A_{21}|$$

dir.

viii. $x{:}\, nx1$ vektör olmak üzere, $\begin{vmatrix} A & x \\ x' & c \end{vmatrix} = |A|\ (\mathrm{c} - \mathrm{x}'\mathrm{A}^{-1}\mathrm{x})$ dir.

ix. $B{:}\, pxn$ ve $C{:}\, nxp$ herhangi matrisler ve $A{:}\, pxp$ singüler olmayan bir matris olsun. Bu takdirde $|A + BC| = |A|\ |I_p + A^{-1}BC| = |A|\ |I_p + A^{-1}BC|$ dir.

x. Eğer A singüler olmayan bir matris ise, $|A + aa'| = |A|\ |1 + a'A^{-1}a|$ dır.

xi. Eğer $B{:}\, pxn$ ve $C{:}\, nxp$ ise, $|I_p + BC| = |I_p + CB|$ dir. (Rao, C. R. and Toutenburg, H.,1999)

Tanım 1.1.11: Eğer $AB = BA = I_n$ ise, $B: nxn$ matrisine A nın tersi(normal tersi) denir. Böyle bir B matrisi varsa, A^{-1} ile gösterilir. A^{-1} in mevcut olması için gerek ve yeter şartın A nın singüler olmaması olduğu kolayca görülür. A^{-1} mevcut ise, $\mathrm{A}^{-1}\mathrm{A} = \mathrm{AA}^{-1} = I_n$ dır.

Teorem 1.1.4: Eğer tüm tersler mevcut ise, aşağıdaki eşitlikler vardır:

i. $(cA)^{-1} = \mathrm{c}^{-1}\,\mathrm{A}^{-1}$.

ii. $(\mathrm{AB})^{-1} = \mathrm{B}^{-1}\,\mathrm{A}^{-1}$.

iii. Eğer $A: pxp$, $B: pxn$, $C: nxn$ ve $D: nxp$ ise, bu takdirde
$(A + BCD)^{-1} = \mathrm{A}^{-1} - \mathrm{A}^{-1}B'(\mathrm{C}^{-1} + D\mathrm{A}^{-1}B)^{-1}D\mathrm{A}^{-1}$.

iv. Eğer $1 + b'\mathrm{A}^{-1}a \neq 0$ ise, bu takdirde bu teoremin **iii.** den

$$(A + ab')^{-1} = \mathrm{A}^{-1} - \frac{\mathrm{A}^{-1}\mathrm{ab}'\mathrm{A}^{-1}}{1 + b'\mathrm{A}^{-1}a}$$

elde edilir.

v. $|\mathrm{A}^{-1}| = |\mathrm{A}|^{-1}$.

vi. (Bir parçalanmış matrisin tersi) E ve $D = H - GE^{-1}F$ singüler olmayacak şekilde $E: n_1xn_1, F: n_1xn_2, G: n_2xn_1$ ve $H: n_2xn_2$ matrisler olmak üzere singüler olmayan

$$A = \begin{bmatrix} E & F \\ G & H \end{bmatrix}$$

parçalanmış matrisinin parçalanmış tersi:

$$\mathrm{A}^{-1} = \begin{bmatrix} E^{-1}(I_{n_1} + FD^{-1}GE^{-1}) & -E^{-1}FD^{-1} \\ -D^{-1}GE^{-1} & D^{-1} \end{bmatrix} = \begin{bmatrix} A^{11} & A^{12} \\ A^{21} & A^{22} \end{bmatrix}$$

dir. (Rao, C. R. and Toutenburg, H.,1999)

Tanım 1.1.12: $\{x_1, x_2, \ldots x_n\}$ n sayıda sütun vektörünün oluşturduğu küme olsun.

$$\sum_{i=1}^{n} a_i x_i = 0 \qquad (1.1.16)$$

eşitliğini sağlayan en az biri sıfırdan farklı olan $a_1, a_2 \ldots, a_n$ reel sayıları varsa, $x_1, x_2, \ldots x_n$ vektörleri lineer bağımlıdır. Aksi halde, $\{x_1, x_2, \ldots x_n\}$ kümesi lineer bağımsızdır. $A: nxn$ matrisinin rankı lineer bağımsız satır veya sütunlarının sayısıdır. Yani, $\mathcal{R}(A)$ nın boyutudur. A nın rankı $rank(A)$ ile gösterilir. $n \geq p$ olmak üzere $rank(A) = p$ ise, A ya tam sütun ranklı ve $rank(A) = n$ ise, A ya tam satır ranklıdır denir.

Teorem 1.1.5: (Ranklar için Kurallar)

i) $n \geq k$ olmak üzere, $A: nxk$ matrisinin rankı k olsun.

$rank(A) = rank(A') = rank(A'A) = k$ dır.

ii) Bir $A: nxk$ matris, $P: nxn$ ve $Q: kxk$ tam ranklı matrisler olmak üzere,

$rank(A) = rank(PA) = rank(AQ)$ dur.

iii) Bir köşegen matrisin rankı elemanları sıfırdan farklı olan sütun vektörlerinin sayısına ya da sıfırdan farklı olan köşegen elemanlarının sayısına eşittir.

iv) $rank(AB) \leq min[rank(A), rank(B)]$ dir.

v) Eğer A ve C singüler değilseler, bu takdirde $rank(ABC) = rank(B)$ dir.

vi) Eğer $rank(A) = rank(A_1)$ ve $A = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix}$ ise, bu takdirde $A_2 = BA_1$ olacak şekilde bir B matrisi vardır. (Rao, C. R. and Toutenburg, H.,1999)

Tanım 1.1.13: x ve y: $nx1$ boyutlu vektörler olsun.

$$x'x = 1,\ y'y = 1 \text{ ve } x'y = 0 \tag{1.1.17}$$

ise, x ve y vektörleri ortonormaldir.

Tanım 1.1.14: $A: nxn$ matrisinin ortogonal(dik) olması için gerek ve yeter şart $A'A = AA' = I_n$ olmasıdır. Her ortogonal matris bir kare matristir. $A'A = I_n$ olduğundan $A^{-1} = A'$ dür. Ortoganal matrisler için aşağıdakiler geçerlidir.

i) Satır(sütun) vektörleri ortonormaldir.

ii) İki ortogonal matrisin çarpımı yine bir ortogonal matristir.

iii) Determinant değerleri $+1$ ve -1 dir.

iv) Singüler değildirler. (Ekni, M.,1999)

Tanım 1.1.15: x ve y : $nx1$ boyutlu vektörler olsun.

$$x'y = \sum x_i y_i = 0 \tag{1.1.18}$$

ise, x ve y ortogonaldir.

Tanım 1.1.16: $x : nx1$ bir vektör olsun ve $\|x\|, x$ in uzunluğunu(normunu) göstersin. Bu takdirde

$$\|x\| = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2} \tag{1.1.19}$$

dir.

Tanım 1.1.17: $\{x_1, x_2 \ldots x_n\}$ $nx1$ boyutlu vektörler kümesi olsun. Her bir vektör birim uzunlukta ve karşılıklı olarak ortogonal ise, bu küme bir ortonormal vektörler kümesidir.

Tanım 1.1.18: $A: nxn$ matrisinin sütun vektörleri ortonormal vektörler kümesi oluşturuyorsa, A matrisi ortogonaldir.

Tanım 1.1.19: Eğer $AA = A$ ise, $A: nxn$ matrisine idempotent matris denir. $A^2 = 0$ ise, A matrisi nilpotent, $A^2 = I_n$ ise, A matrisi unipotent matris olarak isimlendirilir.

Teorem 1.1.6: $rank(A) = r \leq n$ olmak üzere, $A : nxn$ idempotent bir matris olsun. Bu takdirde,

i) $rank(A) = n$ ise, $A = I_n$ dır.
ii) A' de idempotenttir.
iii) P ortogonal ise, $P'AP$ (simetrik) idempotenttir.
iv) P singüler değil ise, PAP^{-1} idempotenttir.
v) A matrisinin rankı izine eşittir.
vi) $AA' = A'A$ ise, $A'A$ ve AA'matrisleri simetrik idempotenttir.
vii) $I_n - A$ da idempotenttir.
viii) Birim matris dışında tüm idempotent matrisler tekildirler. Yani eksik ranklıdırlar.
ix) Tüm idempotent matrisler köşegenleştirilebilir. (Rao, C. R. and Toutenburg, H.,1999)

Tanım 1.1.20: $A : nxn$ bir matris, $x : nx1$ sıfırdan farklı bir vektör olsun. $Ax = \lambda x$ denklemini sağlayan $\lambda \in \mathbb{R}$ sayısına A matrisinin özdeğeri, λ ya karşılık gelen x vektörüne de A nın bir özvektörü denir.

$Ax = \lambda x$ ise, $Ax - \lambda x = 0$ eşitliğinden

$$(A - \lambda I_n)x = 0 \tag{1.1.20}$$

lineer homojen sistemi ortaya çıkar. Bu homojen sistemin bir çözümü $x = 0$ dır ve bu bilinen çözümdür. $x \neq 0$ ın bir çözüm olabilmesi için

$$q(\lambda) = |A - \lambda I_n| = 0 \tag{1.1.21}$$

olması gerekir. $q(\lambda)$ denklemine A kare matrisinin karakteristik denklemi denir. $q(\lambda)$ karakteristik denkleminin katlı, kompleks(karmaşık) ya da reel n tane $\lambda_1, \lambda_2, \ldots, \lambda_n$ kökleri, A nın özdeğerleridir .

Teorem 1.1.7:

i) Bir tekil matris, en az bir sıfır öz değere sahiptir.

ii) A: nxn bir simetrik matris olsun. A matrisinin özdeğerleri reeldir.

iii) A: nxn bir matris ve C : nxn ortogonal bir matris olsun. $C'AC$ matrisinin özdeğerleri ile A matrisinin özdeğerleri aynıdır.

iv) A : nxn bir matris ve G: nxn singüler olmayan bir matris ise, bu takdirde A matrisinin özdeğerleri ile $G^{-1}AG$ matrisinin özdeğerleri aynıdır.

v) A: mxn matris ve B : nxm matris ($n \geq m$)olsun. Bu takdirde BA nın özdeğerleri; AB nin özdeğerleri + $n - m$ tane sıfırdan oluşur.

vi) A: nxn simetrik(reel) bir matris olsun. Bu takdirde A matrisi n tane lineer bağımsız(reel) özvektöre sahiptir.

vii) Simetrik matrislerin sıfırdan farklı özdeğerinin sayısı, matrisin rankını verir.

viii) İdempotent bir matrisin özdeğerleri ya sıfır ya da birdir.

ix) A matrisi idempotent bir matris ise, 1 e eşit olan özdeğer sayısı $rank(A)$ dır. $(n - rank(A))$ kadar özdeğer sıfırdır.

x) A: nxn özdeğerleri $\lambda_1, \lambda_2 \ldots \lambda_n$ olan bir matris ise,

$izA = \sum_{i=1}^{n} \lambda_i$ ve $izA^k = \sum_{i=1}^{n} \lambda_i^k$ dır.

xi) λ, ortogonal bir matrisin özdeğeri ise, $\frac{1}{\lambda}$ da bu matrisin özdeğeridir.

xii) $A: nxn$ bir simetrik matris olsun. A nın birbirinden bağımsız özdeğerleri için elde edilen özvektörler ortogonaldirler.

xiii) Eğer λ, $A : nxn$ matrisinin özdeğeri ve x, λ özdeğerine karşılık gelen A'nın bir özvektörü ise, bu takdirde λ^k, A^k nın bir özdeğeri ve x, λ^k özdeğerine karşılık gelen A^k nın bir özvektörüdür.

Teorem 1.1.8 (Spektral ya da Tayfi Ayrışım Teoremi): Herhangi bir $A: nxn$ simetrik matrisi

$$A = P\Lambda P' = \sum_{i=1}^{n} \lambda_i \, e_i e_i' \tag{1.1.22}$$

olarak yazılabilir. Burada $\Lambda = köş(\lambda_1, \lambda_2, \ldots, \lambda_p)$, A nın özdeğerlerinin köşegen matrisidir ve $P = (e_1, e_2, \ldots, e_p)$ dir. e_i ise, A nın $\lambda_i, i = 1,2 \ldots., p$ özdeğerlerine karşılık gelen standartlaştırılmış özvektörlerdir. P nin ortogonal bir matris olduğu da açıktır.

Tanım 1.1.21 : A, $A = \sum_{i=1}^{n} \lambda_i \, e_i e_i'$ spektral ayrışıma sahip nxn pozitif kararlı(belirli) bir matris olsun. A nın standartlaştırılmış öz vektörleri, $P = [e_1, e_2, \ldots, e_n]$ matrisin sütunları olsun. Bu takdirde, $PP' = P'P = I_n$ ve Λ köşegen matris,

$$\underset{nxn}{\Lambda} = \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix}, \qquad \lambda_i > 0$$

olmak üzere

$$A = \sum_{i=1}^{n} \lambda_i \, e_i e_i' = P\Lambda P'$$

dür. Böylece $(P\Lambda^{-1}P')P\Lambda P' = P\Lambda P'(P\Lambda^{-1}P') = PP' = I_n$ olduğundan,

$$A^{-1} = P\Lambda^{-1}P' = \sum_{i=1}^{n} \frac{1}{\lambda_i} e_i e_i'$$

dür. $\Lambda^{1/2}$, $i - yinci$ köşegen elemanı $\sqrt{\lambda_i}$ olan köşegen matrisi göstersin.

$$A^{1/2} = \sum_{i=1}^{n} \sqrt{\lambda_i}\, e_i e_i{}' = P\Lambda^{1/2}P' \tag{1.1.23}$$

matrisine A matrisinin karekökü denir ve bu matris aşağıdaki özelliklere sahiptir:

i. $(A^{1/2})' = (A')^{1/2}$

ii. $A^{1/2}A^{1/2} = A$

iii. $(A^{1/2})^{-1} = \sum_{i=1}^{n} \frac{1}{\sqrt{\lambda_i}} e_i e_i{}' = P\Lambda^{-1/2}P'$, $\Lambda^{-1/2}$ i-yinci köşegen eleman $\frac{1}{\sqrt{\lambda_i}}$ olan bir köşegen bir matristir.

iv. $A^{1/2}A^{-1/2} = A^{-1/2}A^{1/2} = I$ ve $A^{-1/2}A^{-1/2} = A^{-1}$, $A^{-1/2} = (A^{1/2})^{-1}$ dir. (Johnson, R. A. and Wichern D. W., 1982).

Tanım 1.1.22: A: mxn matrisi için, eğer aşağıdaki dört şartı sağlayan ve A^+ ile gösterilen bir matris varsa, A^+, A nın bir Moore-Penrose genelleştirilmiş tersi (Moore-Penrose g-tersi) olarak tanımlanır.

i. $AA^+A = A$

ii. $A^+AA^+ = A^+$ (1.1.24)

iii. AA^+ simetriktir.

iv. A^+A simetriktir.

Herhangi bir A: mxn matrisi için A^+ matrisi tektir. Fakat, A^+ yerine Moore-Penrose koşullarından sadece **i.** ($AA^+A = A$) koşulunu sağlayan ve lineer denklemlerin çözümü için yeterli olabilen birden fazla A^+ matrisi, A matrisinin herhangi bir genelleştirilmiş tersi(g-tersi) olarak tanımlanır ve bu kısaca A^{g_1} veya yaygın olarak A^- ile gösterilir. Sadece **i.** ve **ii.** koşullarını sağlayan g-terse iki şartlı g-ters ($g_2 - ters$) g-ters($g_3 - ters$) denir ve bu A^{g_2} ile gösterilir. **i., ii.** ve **iii.** koşullarını sağlayan g-terse üç şartlı denir ve bu A^{g_3} ile gösterilir. **i., ii.** ve **iv.** koşularını sağlayan g-terse de üç şartlı g-ters($g_3 - ters$) denir ve bu da $A^{g_3{}^*}$ ile gösterilir.

Teorem 1.1.9: Her A matrisi için (1.1.24) bağıntılarını sağlayan bir A^+ matrisi vardır, yani her matris bir Moore-Penrose g-terse sahiptir.

Teorem 1.1.10: Herhangi bir A: mxn matrisi ve A^-: nxm g-tersi için,

i) A^-A ve AA^- idempotenttir

ii) $rank(A) = rank(AA^-) = rank(A^-A)$ dir.

iii) $rank(A) \leq rank(A^-)$ dir.

Teorem 1.1.11: Herhangi bir $A{:}\, mxn$ matrisi için

i. $A'A(A'A)^-A'A = A'A$

ii. $A(A'A)^-A'A = A$

iii. $A'A(A'A)^-A' = A'$ bağıntıları geçerlidir.(Rao, C. R. and Toutenburg, H.,1999)

Teorem 1.1.12: Herhangi bir $A{:}\, mxn$ matrisi için aşağıdakiler geçerlidir:

i) A singüler olmayan bir matris ise, $A^+ = A^{-1}$ dir.

ii) $(A^+)^+ = A$ dır.

iii) $(A')^+ = (A^+)'$ dür.

iv) Eğer A simetrik ve idempotent ise, $A^+ = A$ dır.

v) AA^+ ve A^+A idempotenttir.

vi) A, A^+, AA^+, A^+A aynı ranka sahiptir. $A^+ = (A'A)^+A' = A'(AA')^+$, $(AA')^+ = (A^+)'A^+$ dır.

vii) $A'AA^+ = A' = A^+AA'$ dür.

viii) $A'(A^+)'A^+ = A^+ = A^+(A^+)'A'$ dür.

ix) $(A'A)^+ = A^+(A^+)'$, $(AA')^+ = (A^+)'A^+$ dır.

x) $A(A'A)^+A'A = A = AA'(AA')^+A$ dır.

xi) $A^+ = (A'A)^+A' = A'(AA')^+$ dır.

xii) Eğer A tam sütun ranklı ise, $A^+ = (AA')^{-1}A'$ ve $A^+A = I_n$ dır.

xiii) Eğer A tam satır ranklı ise, $A^+ = A'(AA')^{-1}$ ve $AA^+ = I_n$ dır.

xiv) $A = 0 \Leftrightarrow A^+ = 0$.

xv) $AB = 0 \Leftrightarrow B^+A^+ = 0$.

xvi) $A^+B = 0 \Leftrightarrow A'B = 0$.

xvii) $(A \otimes B)^+ = A^+ \otimes B^+$ dır. (Pringle, R. and Rayner, A.,1971)

Teorem 1.1.13: Eğer A blok köşegen bir matris ise, bu takdirde A^+ da blok köşegen bir matristir.

Teorem 1.1.14: A, r ranklı mxn tipinde bir matris olsun. Bu takdirde, A nın Moore-Penrose g-tersi aşağıdaki adımlarla hesaplanabilir:

i) $C_1 = I_n$ olmak üzere, $B = AA'$ hesaplanır.

ii) $i = 1,2, \ldots, r-1$ için $C_{i+1} = I_n \left(\frac{1}{i}\right) iz(C_i B) - C_i B$ hesaplanır.

iii) $\frac{rC_r A'}{iz(C_r B)}$ hesaplanıldığı takdirde, bu A^+ dır. Bununla beraber, $C_{r+1} = 0$ ve $iz(C_r B) \neq 0$ dır. **(**Graybill, F. A.,1983).

Herhangi bir r ranklı A matrisinin g-tersi aşağıda verilen algoritma ile bulunabilir.

i) A matrisinin r rankına sahip herhangi bir alt matrisi W ile gösterilsin.

ii) W matrisi için W^{-1} ve $(W^{-1})'$ hesaplanır.

iii) Bulunan $(W^{-1})'$ matrisini A matrisindeki yerine koyarken A nın diğer elemanları sıfır alınır.

iv) Bulunan matrisin transpozesini alınır.

v) Sonuç, A nın $AA^{-}A = A$ koşulunu sağlayan bir genelleştirilmiş tersidir. **(**Ekni, M.,1999)

Teorem 1.1.15: Elemanları 1 sayılarından oluşan nxm tipinde bir matris

$$J_{nxm} = \begin{bmatrix} 1 & 1 & \ldots & 1 \\ 1 & 1 & \ldots & 1 \\ \ldots & \ldots & \ldots & \ldots \\ 1 & 1 & \ldots & 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ \ldots \\ 1 \end{bmatrix}_{nx1} \otimes \begin{bmatrix} 1 & 1 & \ldots & 1 \end{bmatrix}_{1xm} \text{ olmak üzere,}$$

$$J^{+}{}_{mxn} = \begin{bmatrix} 1 \\ 1 \\ \ldots \\ 1 \end{bmatrix}^{+}_{nx1} \otimes \begin{bmatrix} 1 & 1 & \ldots & 1 \end{bmatrix}^{+}_{1xm} = \frac{1}{n}\begin{bmatrix} 1 & 1 & \ldots & 1 \end{bmatrix}_{1xn} \otimes \frac{1}{m}\begin{bmatrix} 1 \\ 1 \\ \ldots \\ 1 \end{bmatrix}_{mx1} = \frac{1}{nm} J_{mxn} \tag{1.1.25}$$

dir.

Teorem 1.1.16: a : $nx1$ tipinde bir vektör ise,

$$\alpha^{+} = (\alpha'\alpha)^{+}a' \qquad (1.1.26)$$

dür.

Teorem 1.1.17: $a \in R$ reel sayısı için

$$a^{+} = \begin{cases} 0, a = 0 \\ \frac{1}{a}, a \neq 0 \end{cases} \qquad (1.1.27)$$

dır.

Teorem 1.1.18: $M \geq 0$:nxn ve N:mxn herhangi matrisler olsun. Bu takdirde,

$$M - N'(NM^{+}N')^{+}N \geq 0 \qquad (1.1.28)$$

olması için gerek ve yeter şart $\mathcal{R}(N'NM) \subset \mathcal{R}(M)$ olmasıdır. (Rao, C. R. and Toutenburg, H.,1999)

Teorem 1.1.19: A : mxn ve g: $mx1$ bilinen matrisler, x: $nx1$ bilinmeyenlerin bir vektörü ve $Ax = g$ lineer denklem sistemi olsun.

a) Bu lineer denklem sisteminin tutarlı olması için gerek ve yeter şart

$$AA^{+}g = g \qquad (1.1.29)$$

olmasıdır.

b) Denklem sistemi tutarlı olsun. $x_0 = Gg$ nin bir çözüm olması için gerek ve yeter şart G matrisinin A nın herhangi bir genelleştirilmiş tersi olmasıdır.
c) Verilen denklem sistemi tutarlı olsun. w uygun boyutlu keyfi bir vektör olmak üzere, $x = A^{+}g + (I_n - A^{+}A)w$, bu denklem sisteminin bir çözümüdür.

Tanım 1.1.23: $x' = (x_1, x_2, ..., x_n)$ ve A: nxn tipinde simetrik bir matris olmak üzere, $x_1, x_2, ..., x_n$ değişkenlerine göre

$$Q = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_i x_j$$

$$= \sum_{i=j=1}^{n} x_i^2 a_{ii} + \sum_{i \neq j}^{n} a_{ij} x_i x_j$$

$$= x'Ax \tag{1.1.30}$$

homojen kuadratik fonksiyonuna x'e göre bir karesel form (kuadratik form) denir. A'ya bu karesel formun matrisi denir. Bir karesel formun matrisi daima simetrik olacak şekilde yazılabilir. Eğer karesel form simetrik değil ise, $x'[\frac{1}{2}(A+A')]x$ dönüşümü ile simetrik duruma getirilebilir. Simetrik olması durumunda $a_{ij} = a_{ji}$ olduğunda

$$x'Ax = \sum x_i^2 a_{ii} + 2\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}\ x_i x_j \tag{1.1.31}$$

şeklinde yazılabilir ve

$$x'Ax = iz(x'Ax) = iz(Axx') \tag{1.1.32}$$

dür.

Tanım 1.1.24: $a{:}\,nx1$ bir vektör ve $x{:}\,nx1$ bir vektör olsun. Bu durumda $a'x$ ifadesine x'e göre bir lineer form denir.

Tanım 1.1.25: $A{:}\,nxn$ simetrik bir matris ve her $x \neq 0{:}nx1$ bir vektör olmak üzere, $Q = x'Ax > 0$ sağlanıyorsa, bu takdirde A matrisine pozitif kararlıdır (pozitif belirlidir) denir. Bu durum kısaca $A > 0$ şeklinde yazılır.

Teorem 1.1.20: Eğer bir $A{:}\,nxn$ matrisi, $A > 0$ ise, bu takdirde

i) A matrisi tam ranka sahiptir. Yani, A tekil değildir.

ii) $izA = \sum_{i=1}^{n} a_{ii} > 0$ dır.

iii) $\det A = |A| > 0$ dır.

Teorem 1.1.21: $A > 0$ ve $P{:}\,nxm$, m ranklı $(m \leq n)$ bir matris olsun. Bu takdirde $P'AP > 0$ ve özellikle $A = I_n$ seçmek suretiyle $P'P > 0$ dır.

Teorem 1.1.22: Eğer A pozitif kararlı ise, bu takdirde A^{-1} de pozitif kararlıdır. Bunu $A > 0 \Leftrightarrow A^{-1} > 0$ şeklinde ifade edilebilir.

Teorem 1.1.23: Bir $A{:}\,nxn$ matrisi, $A > 0$ ise, bu takdirde A matrisinin tüm özdeğerleri pozitiftir.

Tanım 1.1.26: A: nxn simetrik bir matris ve x'in sıfırdan farklı tüm değerleri için $Q = x'Ax = 0$ sağlanıyorsa, A matrisine pozitif yarı-kararlıdır denir.

Teorem 1.1.24: Bir $A{:}\,nxn$ matrisi pozitif yarı kararlı ise, A nın özdeğerleri negatif değildir ve en az biri sıfırdır.

Tanım 1.1.27: Eğer $Q(x) = x'Ax$ pozitif kararlı veya pozitif yarı kararlı ise, yani her x için $Q(x) = x'Ax \geq 0$ ise, $Q(x) = x'Ax$ karesel formuna (ve aynı zamanda A matrisine) negatif olmayan kararlıdır (n.o.k.) denir. Bu durum kısaca $A \geq 0$ şeklinde gösterilir.

Teorem 1.1.25: Eğer $A \geq 0$ ise, bu takdirde $\lambda_i \geq 0$, $i = 1,2, \ldots n$ dir.

Teorem 1.1.26: A negatif olmayan kararlı, nxn tipinde, r ranklı simetrik bir matris olsun. Bu takdirde $A = P'P$ olacak şekilde rxn boyutlu ve $rank(P) = r$ olan bir P matrisi vardır.

Teorem 1.1.27: $A{:}\,nxn$ pozitif kararlı bir matris olsun. $A = P'P$ eşitliğini sağlayan tekil olmayan nxn tipinde bir P matrisi vardır.

Teorem 1.1.28: Eğer A pozitif kararlı ise, bu takdirde her P için $P'AP$ negatif olmayan kararlıdır. Eğer P kare ve tekil değil ise, bu takdirde $P'AP$ pozitif kararlıdır.

Teorem 1.1.29: r ranklı herhangi bir nxp tipinde P matrisi için $P'P \geq 0$ ve $PP' \geq 0$ dır.

Teorem 1.1.30: Herhangi bir $A \geq 0$ matrisi için $0 \leq \lambda_i \leq 1, i = 1,2, \ldots n,$ olması için gerek ve yeter şart tüm negatif olmayan kararlı C matrisi için $izCD \geq 0$ olmasıdır. Yani $D \geq 0 \Leftrightarrow$ Her $C \geq 0$ için $izCD \geq 0$ dır.

Teorem 1.1.31: Pozitif yarı kararlı simetrik herhangi bir matris tekildir.

Teorem 1.1.32: A: nxn idempotent bir matris ve $rank(A) < n$ ise, bu takdirde A pozitif yarı kararlı bir matristir.

Teorem 1.1.33: A: nxm, $rank(A) = m \leq n$ ve B: mxm herhangi bir simetrik matris olsun. Bu takdirde $ABA' \geq 0$ olması için gerek ve yeter şart $B \geq 0$ olmasıdır.

Teorem 1.1.34: $MM' - NN' \geq 0$ kabul edelim. Bu takdirde, $N = MH$ olacak şekilde bir H matrisi vardır. (Rao, C. R. and Toutenburg, H.,1999)

Tanım 1.1.28: $f : R^n \to R$

$x \to f(x) = f(x_1, x_2, ..., x_n)$ olmak üzere

$$\frac{\partial f}{\partial x} = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \cdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix} \tag{1.1.33}$$

fonsiyonuna f nin x vektörüne göre türevi denir. Bu türev f nin gradyent vektörüdür.

Tanım 1.1.29: $X = (x_{ij})$ elemanları reel sayı olan bir matris ve

f: $R^{nxm} \to R$,

$$X \to f(X) \tag{1.1.34}$$

reel değerli bir fonksiyon olmak üzere, $\frac{\partial f}{\partial X} = (\frac{\partial f}{\partial x_{ij}})_{nxm}$ fonksiyonuna f fonksiyonunun X matrisine göre türevi denir.

Teorem 1.1.35:

i. a :$nx1$ sabitlerin bir vektörü olmak üzere,

$$\frac{\partial(a'x)}{\partial x} = \frac{\partial(x'a)}{\partial x} = a \, , \frac{\partial^2 a'x}{\partial x^2} = 0 \tag{1.1.35}$$

ii. A: nxn sabitlerin matrisi olmak üzere,

$$\frac{\partial(x'Ax)}{\partial x} = (A + A')x \tag{1.1.36}$$

iii. A: nxn simetrik bir matris olmak üzere,

$$\frac{\partial(x'Ax)}{\partial x} = 2Ax, \quad \frac{\partial^2(x'Ax)}{\partial x^2} = 2A \tag{1.1.37}$$

dır.

1.2 İstatistiksel Ön Hazırlık

Tanım 1.2.1: Değeri bir deney sonucuyla belirlenen bir değişkene rasgele değişken denir. X bir rasgele değişken olmak üzere, X'in alabileceği değerlerin sayısı sonlu veya sayılabilir sonsuzlukta ise, X'e kesikli rasgele değişken; X bir aralıkta ya da bir aralıklar topluluğunda her değeri alabiliyorsa X'e sürekli rasgele değişken denir.

Tanım 1.2.2: Bir X sürekli rasgele değişkeninin herhangi bir g fonksiyonunun beklenen değeri, eğer integral mevcut ise ,

$$Eg(x) = \int_{-\infty}^{+\infty} g(\xi) f(\xi) d\xi \tag{1.2.1}$$

olarak tanımlanılır. Daha genel olarak $x = (x_1, x_2, ..., x_n)'$ f ortak olasılık yoğunluk fonksiyonuna sahip $nx1$ boyutlu bir rasgele vektör olsun. Bu takdirde x'in herhangi bir g fonksiyonunun beklenen değeri n-katlı integral mevcut ise,

$$Eg(x) = \int_{-\infty}^{+\infty} ... \int_{-\infty}^{+\infty} g(\xi_1, \xi_2, ..., \xi_n) d\xi_1 ... d\xi_n \tag{1.2.2}$$

şeklinde tanımlanır.

Elemanları rasgele değişkenler olan $nx1$ boyutlu x rasgele vektörü $x = \begin{bmatrix} \mathrm{x}_1 \\ \vdots \\ \mathrm{x}_\mathrm{n} \end{bmatrix}$ olsun. Burada $E(\mathrm{x}_\mathrm{i}) = \mu_\mathrm{i}$, $i = 1,2, ..., n$, x_i rasgele değişkenlerinin beklenen değerleridir. Bu nedenle, x rasgele vektörünün beklenen değeri,

$$E(x) = \mu = \begin{bmatrix} E(\mathrm{x}_1) \\ \vdots \\ E(\mathrm{x}_\mathrm{n}) \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_n \end{bmatrix} \tag{1.2.3}$$

yazılabilir. a reel sayılar vektörü olmak üzere,

$$E(a) = a \tag{1.2.4}$$

dır. A ve B sabitlerin matrisleri ve $x : px1$ boyutlu rasgele bir vektör olmak üzere,

$$E(AxB) = A(E(x))B \tag{1.2.5}$$

dir. $a{:}\, nx1$ sabitlerin bir vektörü ve $x : px1$ boyutlu rasgele bir vektör olmak üzere

$$E(a'x) = a'E(x) \tag{1.2.6}$$

dir.

Tanım 1.2.3: Eğer X bir rasgele değişken ise, varyansı

$$Var(X) = E(X - E(X))^2 \tag{1.2.7}$$

şeklinde tanımlanır. Eğer X ve Y bir ortak olasılık fonksiyonuna sahip iki rasgele değişken ise, X ve Y arasındaki kovaryans

$$Kov(X,Y) = E[(X - E(X))(Y - E(Y))] \tag{1.2.8}$$

şeklinde tanımlanır. Eğer $Kov(X,Y) = 0$ ise, X ve Y istatistiksel olarak ilişkisizdirler.

X ve Y rasgele değişkenleri ve α , β sabitleri hakkında aşağıdaki özellikler vardır:

$$Var(X+\alpha) = Var(X), \tag{1.2.9}$$

$$Var(\alpha X) = \alpha^2 Var(X), \tag{1.2.10}$$

$$Var(X+Y) = Var(X) + Var(Y) + 2Kov(X,Y) \tag{1.2.11}$$

ve

$$Kov(\alpha X, \beta Y) = \alpha Kov(X,Y)\beta \tag{1.2.12}$$

dir. Eğer X ve Y ilişkisiz iseler, (1.2.11) bağıntısının özel bir durumu olarak

$$Var(X+Y) = Var(X) + Var(Y) \tag{1.2.13}$$

elde edilir.

Varyansın çok değişkenli analizdeki karşılığı ise, nxn boyutlu simetrik varyans-kovaryans matrisidir. $x: nx1$ bir rasgele vektörünün varyansı

$$Var(x) = E[(x - E(x)(x - E(x))'] \qquad (1.2.14)$$

matrisi ile tanımlanır.

Tanım 1.2.4: Eğer $x: nx1$ ve $y: mx1$ rasgele vektörler ise, bu takdirde x ve y arasındaki varyans-kovaryans matrisi nxm tipindeki

$$Var(x, y) = E[(x - E(x)(y - E(y))'] \qquad (1.2.15)$$

matrisi olarak tanımlanır. x ve y rasgele vektörlerinin varyans- kovaryans matrisi

$$V = \begin{bmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{bmatrix} \qquad (1.2.16)$$

ile gösterilsin. Burada V_{11}, V_{12}, V_{21} ve V_{22} alt matrisleri sırasıyla nxn, nxm, mxn, mxm boyutludur.

$$V_{11} = \begin{bmatrix} \sigma_{X_1X_1} & \cdots & \sigma_{X_1X_n} \\ \vdots & \ddots & \vdots \\ \sigma_{X_nX_1} & \cdots & \sigma_{X_nX_n} \end{bmatrix}, \; V_{12} = \begin{bmatrix} \sigma_{X_1Y_1} & \cdots & \sigma_{X_1Y_m} \\ \vdots & \ddots & \vdots \\ \sigma_{X_nY_1} & \cdots & \sigma_{X_nY_m} \end{bmatrix},$$

$$V_{22} = \begin{bmatrix} \sigma_{Y_1Y_1} & \cdots & \sigma_{Y_1Y_m} \\ \vdots & \ddots & \vdots \\ \sigma_{Y_mY_1} & \cdots & \sigma_{Y_mY_m} \end{bmatrix}.$$

V_{21} matrisi de V_{12} matrisinin transpozesidir. Varyans değerleri

$$\sigma_{ij} = \frac{1}{n}\left[\sum_{r=1}^{n}(x_{ir} - \bar{x}_i)(x_{jr} - \bar{x}_j)\right] \qquad (1.2.17)$$

şeklinde hesaplanır. Buna göre, X ve Y değişken kümelerine ait varyans-kovaryans matrisinde, $i = j$ iken varyans değerleri (σ_{ii}), $i \neq j$ iken kovaryans değerleri hesaplanmış olmaktadır. V, varyans-kovaryans matrisinde köşegende yer alan varyans değerleri, değişkenlerin dağılışı hakkında bilgi verir iken, köşegen dışındaki kovaryans değerleri, değişken çiftleri arasındaki birlikte değişimi vermektedir. Genellikle kitle değerini bilmek mümkün olmadığından bunların yerine örnekten hesaplanan istatistiklerin kullanılması ile,

$$\hat{\mu} = \begin{bmatrix} \hat{\mu}_1 \\ \hat{\mu}_2 \end{bmatrix}, \quad S = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \qquad (1.2.18)$$

yazılabilir. (Alpar R., 2003) Burada $\hat{\mu}$, μ nün tahminini, S ise V varyans-kovaryans matrisinin tahminidir.

Eğer $Kov(x, y) = 0$ ise, iki x ve y vektörleri ilişkisizdirler, denir.

Teorem 1.2.1: Her varyans-kovaryans matrisi simetrik ve en azından pozitif-yarı kararlıdır. Yani, x: $nx1$ rasge bir vektör olmak üzere,

$$Var(x) \geq 0 \tag{1.2.19}$$

Teorem 1.2.2: $x : nx1$ tipinde bir rasgele vektör olsun ve A: mxn bir reel sayılar matrisi ve b: $mx1$ reel sayılar vektörü olmak üzere $y = Ax + b$ vektörünü tanımlansın. Bu takdirde

$$E(y) = AE(x) + b \tag{1.2.20}$$

ve

$$Var(Ay) = AVar(y)A' \tag{1.2.21}$$

dir.

Teorem 1.2.3: x ve y: $nx1$ rasgele vektörler ve z: $mx1$ de bir rasgele vektör olsun. A: pxn ve B: qxm reel sayıların matrisleri olsunlar. Bu takdirde,

$$Var(x+y) = Var(x) + Var(y) + K\text{ov}(x, y) + K\text{ov}(y, x), \tag{1.2.22}$$

$$K\text{ov}(Ax, Bz) = AK\text{ov}(x, z)B' \tag{1.2.23}$$

x ve y ilişkisiz iseler,

$$Var(x+y) = Var(x) + Var(y) \tag{1.2.24}$$

dir. **(**Magnus, J. R., 1990)

Teorem 1.2.4: x, $E(x) = \mu$ ortalamalı ve $Var(x) = V$ varyanslı $nx1$ tipinde rasgele bir vektör olsun. Bu takdirde,

$$E(x'Ax) = iz(AV) + \mu' A\mu \tag{1.2.25}$$

dir. (Rencher, A.C. and Schaalje, B.G.,2007)

İspat 1:

$$E(x'Ax) = E(izx'Ax) = E(izAxx')$$

$$= izE(Axx') = izAE(xx')$$

$$= izA(V + \mu\mu') = izAV + \mu' A\mu$$

Bu ispat aşağıdaki yöntemle de yapılabilir.

İspat 2:

$$E(x'Ax) = E(\sum_{ij} a_{ij} x_i x_j)$$

$$= \sum_{ij} a_{ij} E(x_i x_j) = \sum a_{ij} \left[K\,\text{ov}(x_i, x_j) + E(x_i) E(x_j) \right]$$

$$= \sum_{ij} a_{ij} K\,\text{ov}(x_i, x_j) + \sum_{ij} a_{ij} E(x_i) E(x_j)$$

$$= iz(AK\,\text{ov}(x)) + E(x)' AE(x)$$

dir.

Sonuç 1.2.1: $K\,\text{ov}(x) = \sigma^2 I$ ve $E(x) = 0$ ise,

$$E(x'Ax) = \sigma^2 iz(A) \qquad (1.2.26)$$

dır.

Tanım 1.2.5: Eğer $x: nx1$ ve $y: mx1$ rasgele vektörler ise, bu takdirde x ve y arasındaki korelasyon matrisi,

$$Kor(x_i, y_i) = \frac{Kov(x_i, y_i)}{\sqrt{Var(x_i) Var(y_i)}} \qquad (1.2.27)$$

olmak üzere, nxm tipindeki $R = \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix}$ matrisidir. Burada, R, $n + m$ tane değişkenin kendi aralarındaki ilişki katsayılarını veren ilişki(korelasyon) matrisidir. Bu matrisin, köşegen elemanları 1 ve köşegen dışı elemanları ise -1 ile +1 arasında değerler alabilen ilişki katsayılarını, R_{11}, n tane değişkenin kendi aralarındaki ilişki katsayılarını veren ilişki matrisini, R_{22}, m tane değişkenin kendi aralarındaki ilişki katsayılarını veren ilişki matrisini, R_{21}, n ve m tane değişkenlerin kendi aralarındaki ilişki katsayısını veren ilişki matrisini, R_{12} ise R_{21} matrisinin transpozesini göstermektedir. Bu korelasyon matrisinin alt matrisleri sırası ile,

$$R_{11} = \begin{bmatrix} \rho_{x_1 x_1} & \cdots & \rho_{x_1 x_n} \\ \vdots & \ddots & \vdots \\ \rho_{x_n x_1} & \cdots & \rho_{x_n x_n} \end{bmatrix}, R_{22} = \begin{bmatrix} \rho_{y_1 y_1} & \cdots & \rho_{y_1 y_m} \\ \vdots & \ddots & \vdots \\ \rho_{y_m y_1} & \cdots & \rho_{y_m y_m} \end{bmatrix},$$

$$R_{12} = \begin{bmatrix} \rho_{x_1 y_1} & \cdots & \rho_{x_1 y_n} \\ \vdots & \ddots & \vdots \\ \rho_{x_n y_1} & \cdots & \rho_{x_n y_m} \end{bmatrix}$$

dir. Burada $\rho_{x_i x_i} = 1, i = 1,2,\dots,n$ ve $\rho_{y_i y_i} = 1, i = 1,2,\dots,m$ dir. $D^{1/2}$, köşegen elemanları V nin köşegen elemanlarının karekökü olan bir köşegen matris olmak üzere,

$$R = (D^{1/2})^{-1} S (D^{1/2})^{-1}$$

ile hesaplanabilir. (Çankaya, S., 2005)

Tanım 1.2.6: Bir veri matrisinin her bir elemanı x_{ij}: $j-yinci$ değişkenin $i-yinci$ gözlem değerini, $\bar{X}_j$: $j-yinci$ değişkenin ortalamasını ve s_{jj}: $j-yinci$ değişkenin varyansını (S matrisinin köşegen elemanlarını) göstermek üzere her bir X_j değişkeni değerleri,

$$Z = \frac{\sum_{i=1}^{n}(x_{ij}-\bar{X}_j)}{s_{jj}} \qquad j = 1,2,\dots,p, \tag{1.2.28}$$

ile standartlaştırılabilir. Bu standartlaştırılmış değişkenlerin ortalaması 0, varyansı 1'dir. Kovaryans değeri ise, $+1$ ile -1 arasında değişir. Standartlaştırılmış iki değişken arasındaki kovaryans değeri korelasyon katsayısına eşittir.

Tanım 1.2.7: Sürekli bir X rasgele değişkeni aşağıdaki $f(x)$ olasılık yoğunluk fonksiyonuna sahip ise, bu rasgele değişken tek değişkenli normal dağılıma sahiptir denir.

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}} \quad , -\infty < x < \infty, \ -\infty < \mu < \infty, \ \sigma^2 > 0. \tag{1.2.29}$$

Eğer X yukarıdaki gibi bir dağılıma sahip ise, bu durum $X \sim N(\mu, \sigma^2)$ şeklinde yazılır. Eğer $\mu = 0$ ve $\sigma^2 = 1$ ise, X standard normal dağılımlıdır denir.

Teorem 1.2.5: $X \sim N(\mu, \sigma^2)$ normal dağılıma sahip ise, $Z = \frac{X-\mu}{\sigma}$ değişkeni $N(0,1)$ normal dağılımına, yani, standart normal dağılıma sahiptir. (Alpar, R., 2003)

Tanım 1.2.8: (X, Y) iki-boyutlu rasgele değişken aşağıdaki olasılık yoğunluk fonksiyonuna sahip ise, bu rasgele değişken iki değişkenli normal dağılıma sahiptir denir.

$$f_{x,y}(x,y) = f(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} e^{\left\{-\frac{1}{2(1-\rho^2)}\left[(\frac{x-\mu_x}{\sigma_x})^2 - 2\rho\frac{x-\mu_x}{\sigma_x}.\frac{y-\mu_y}{\sigma_y} + (\frac{y-\mu_y}{\sigma_y})^2\right]\right\}}$$

(1.2.30)

Burada, $-1 < \rho < 1$, $0 < \sigma_x$, $0 < \sigma_y$, $-\infty < \mu_x < \infty$ ve $-\infty < \mu_y < \infty$,ρ, σ_x, σ_y, μ_x ve μ_y sabit sayılar ve $-\infty < x < \infty$, $-\infty < y < \infty$ dur.

Tanım 1.2.9: $x' = (x_1, x_2, \ldots, x_n)$, $n-$boyutlu bir rasgele vektör olsun. Eğer, x in olasılık yoğunluk fonksiyonu

$$f(x) = \frac{1}{(2\pi)^{n/2}|\mathrm{V}|^{n/2}} e^{-\frac{1}{2}(\mathrm{x}-\mathrm{\mu})'(\mathrm{V})^{-1}(\mathrm{x}-\mathrm{\mu})} \quad (1.2.31)$$

ise, bu takdirde x rasgele vektörü, $\mu = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{bmatrix}$ beklenen değerli ve $V > 0$ kovaryans matrisli $n-$boyutlu bir normal dağılıma sahiptir denir. Böyle bir durumda, $x \sim N_n(\mu, \mathrm{V})$ şeklinde yazılır.

Teorem 1.2.6: x: $nx1$, $x \sim N_n(\mu, \mathrm{V})$ ise, bu takdirde

$$V\,\mathrm{ar}(x'Ax) = 2iz\left[(AV)^2\right] + 4\mu' AVA\mu$$

dür. (Rencher, A.C. and Schaalje, B.G.,2007)

Teorem 1.2.7: x: $nx1$, $x \sim N_n(\mu, \mathrm{V})$ ise, bu takdirde

$$K\,\mathrm{ov}(x, x'Ax) = 2VA\mu$$

dür. (Rencher, A.C. and Schaalje, B.G.,2007)

Sonuç 1.2.2: $B: kxn$ reel sayılar matrisi olsun. Bu takdirde

$$Kov(Bx, x'Ax) = 2BVA\mu$$

dür.

Teorem 1.2.8: $x \sim N_n(\mu, \mathrm{V})$, $A : nxn$ ve $rank(A) = n$ olan rasgele olmayan bir matris ve $b: nx1$ rasgele olmayan bir vektör olsun. Bu takdirde,

$$Y = Ax + b \sim N_n(A\mu + b, A\mathrm{VA}') \tag{1.2.32}$$

dür.

Teorem 1.2.9: Y nin çoklu normal dağılıma sahip olması için gerek ve yeter şart sıfırdan farklı her a reel değerli vektörü için $a'Y$ nin tekli normal dağılmış olmasıdır. (Seber, G.A.F., 1977)

Teorem 1.2.10: Eğer, $x \sim N_n(0, \mathrm{I})$ ise, bu takdirde

$$x'x \sim \chi^2_n \tag{1.2.33}$$

dir, yani $x'x$, n serbestlik dereceli merkezi χ^2 dağılımına sahiptir. (Rao, C. R. and Toutenburg, H.,1999)

Teorem 1.2.11: Eğer $x \sim N_n(\mu, \mathrm{I})$ ise, bu takdirde

$$x'x \sim \chi^2_{n,\lambda} \tag{1.2.34}$$

dir. Yani, $x'x$, merkezi olmama parametresi $\lambda = \mu'\mu = \sum_{i=1}^{n} \mu_i^2/2$ ve serbestlik derecesi n olan merkezi olmayan bir χ^2 dağılımına sahiptir. (Rao, C. R. and Toutenburg, H.,1999)

Teorem 1.2.12: $x : nx1$ rasgele vektörü, $x \sim N_n(\mu, I_n)$ dağılımına ahip ise ve $A : nxn$ bir simetrik, idempotent matris ise, bu takdirde

$$x'Ax \sim \chi^2_{rank(A),\lambda} \tag{1.2.35}$$

dır. Yani, $x'Ax$, merkezi olmama parametresi $\lambda = \frac{\mu' A\mu}{2}$ ve serbestlik derecesi $rank(A)$ olan merkezi olmayan bir χ^2 dağılımına sahiptir. (Ekni,M.,1999)

İspat: A matrisi simetrik, idempotent ve $rank(A) = k$ olsun. Öyle bir dik P matrisi vardır ki, $\lambda_1, \lambda_2, ..., \lambda_n$ A'nın özdeğerleri olmak üzere

$$P'AP = \begin{bmatrix} \lambda_1 & . & . & 0 \\ 0 & \lambda_2 & . & 0 \\ . & . & . & . \\ 0 & . & . & \lambda_n \end{bmatrix} \tag{1.2.36}$$

dir. A matrisi simetrik ve idempotent olduğunda, özdeğerleri 0 veya 1 dir. Ayrıca $rank(A) = iz(A)$ dır.

$$iz\left(P'AP\right) = iz(PP'A) = iz(A) = k \tag{1.2.37}$$

dır. Zira $PP' = I$ dır. $iz(P'AP) = \sum_{i=1}^{k} \lambda_i = k$ olur. k tane özdeğer 1 ve $(n-k)$ tane özdeğer 0 dır.

$$P'AP = \begin{bmatrix} I_k & 0 \\ 0 & 0 \end{bmatrix} \tag{1.2.38}$$

biçiminde yazılabilir. $z = P'x$ olan yeni bir z vektörü

$$z = \begin{bmatrix} \underset{kx1}{z_1} \\ \cdots \\ \underset{(n-k)x1}{z_2} \end{bmatrix} \tag{1.2.39}$$

şeklinde tanımlansın.

$$\underset{nxp}{P} = \begin{bmatrix} \underset{nxk}{P_1} & \vdots & \underset{nx(n-k)}{P_2} \end{bmatrix} \tag{1.2.40}$$

ve

$$P' = \begin{bmatrix} \underset{kxn}{P'_1} \\ \cdots \\ \underset{(n-k)xn}{P'_2} \end{bmatrix} \tag{1.2.41}$$

olsun. Bu takdirde,

$$P'x = z = \begin{bmatrix} \underset{kx1}{P'_1 x} \\ \cdots \\ \underset{(n-k)x1}{P'_2 x} \end{bmatrix} \tag{1.2.42}$$

olur. Ayrıca $z_1 = P'_1 x$ dir. z_1 bağımsız ve normal dağılıma sahip $x_1, x_2, \ldots, x_n$ rasgele değişkenlerinin bir lineer kombinasyonudur. Karesel form $x'Ax$ tir. $P'x = z$ idi. P dik matris olduğundan $PP'x = Pz, x = Pz$ ve

$$x'Ax = z'P'APz$$

$$= z'\begin{bmatrix} I_k & 0 \\ 0 & 0 \end{bmatrix} z = [z_1' \vdots z_2'] \begin{bmatrix} I_k & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} z_1 \\ \cdots \\ z_2 \end{bmatrix}$$

$$= [z_1' I_k \vdots 0] \begin{bmatrix} z_1 \\ \cdots \\ z_2 \end{bmatrix} = z_1' z_1 \tag{1.2.43}$$

olur.

$x'Ax = z_1' z_1$ elde edilir. $x'Ax$ karesel formunun dağılımını belirlemek için $z_1' z_1$ in dağılımının bilinmesi yeterli olacaktır

$$E(z_1) = P_1' E(x), \ Var(z_1) = P_1' Var(x) P_1 = P_1' P_1 = I$$

olduğundan $z_1 \sim N(P_1'\mu, I_k)$ bulunur. $z_1' z_1$ dağılımı, $z_1' z_1 \sim \chi^2_{k, \lambda = \frac{(P_1'\mu)'(P_1'\mu)}{2} = \frac{\mu' P_1 P_1' \mu}{2}}$ olur. Burada

$$P'AP = \begin{bmatrix} I_k & 0 \\ 0 & 0 \end{bmatrix} \Rightarrow A = P \begin{bmatrix} I_k & 0 \\ 0 & 0 \end{bmatrix} P' = [P_1 \vdots P_2] \begin{bmatrix} I_k & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} P_1' \\ \cdots \\ P_2' \end{bmatrix} = P_1 P_1' \tag{1.2.44}$$

dır. Bu nedenle, $z_1' z_1 \sim \chi^2_{k, \lambda = \frac{\mu' A \mu}{2}}$ elde edilir.

Teorem 1.2.13: $x \sim N_n(\mu, V)$, A_i, $i = 1,2, \ldots, k$ için r_i ranklı simetrik matrisler ve $A = \sum_{i=1}^{k} A_i$ r ranklı simetrik matris olmak üzere, $x'Ax = \sum_{i=1}^{k} x'A_i x$ olsun. Bu takdirde,

i. $x'A_i x / \sigma^2 \sim \chi^2_{r_i, \lambda = \frac{\mu' A_i \mu}{2\sigma^2}}$, $i = 1,2, \ldots k$.

ii. Her $i \neq j$ için $x'A_i x$ ve $x'A_j x$ bağımsızdırlar.

iii. $x'Ax/\sigma^2 \sim \chi^2_{r,\lambda=\frac{\mu' A\mu}{2\sigma^2}}$ dir.

Bu sonuçların elde edilmesi için gerek ve yeter şart aşağıdaki üç ifadenin herhangi ikisinin doğru olmasıdır:

a. Her A_i idempotenttir.
b. Her $i \neq j$ için $A_i A_j = 0$ dır.
c. $A = \sum_{i=1}^{k} A_i$ idempotenttir.

veya,

d. $r = \sum_{i=1}^{k} r_i$

olmak üzere, **c** ve **d** nin sağlanmasıdır. (Rencher, A.C. and Schaalje, B.G.,2007)

Sonuç 1.2.3: $x \sim N_n(\mu, V)$, A_i, $i = 1,2,\dots,k$ için $\mathrm{r_i}$ ranklı simetrik matrisler ve $x'Ax = \sum_{i=1}^{k} x'A_i x$ olsun. Bu takdirde, **i.** her $x'A_i x/\sigma^2 \sim \chi^2_{r_i,\lambda=\frac{\mu' A_i \mu}{2\sigma^2}}$ ve **ii.** $x'A_i x$ terimleri karşılıklı olarak bağımsız olması için gerek ve yeter şart aşağıdaki ifadelerin herhangi birinin sağlanmasıdır:

a. Her A_i idempotenttir.
b. Her $i \neq j$ için $A_i A_j = 0$ dır.
c. $n = \sum_{i=1}^{k} r_i$ dir.

Teorem 1.2.14: $x \sim N_n(\mu, V)$ ve V pozitif kararlı bir matris olsun. Bu takdirde, $x'Ax$ karesel formunun, $\lambda = \frac{\mu' A\mu}{2}$ olmak üzere, $\chi^2_{rank(A)=p,\lambda}$ dağılımına sahip olması için gerek ve yeter şart aşağıdaki üç şarttan herhangi birinin sağlanmasıdır.

i. AV, p ranklı idempotent bir matristir.
ii. VA, p ranklı idempotent bir matristir.
iii. V, A nın bir $g-tersidir$ ve A matrisi p ranklıdır. **(**Graybill, F. A.,1976).

Teorem 1.2.15: Eğer, $x \sim N_n(\mu, \Sigma)$ ise, bu takdirde

i) $x'\Sigma^{-1}x \sim \chi^2_{n,\lambda=\mu'\Sigma^{-1}\mu}$ dür.
ii) $(x-\mu)'\Sigma^{-1}(x-\mu) \sim \chi^2_n$ dir. (Rao, C. R. and Toutenburg, H.,1999)

Teorem 1.2.16: $A: nxn$ simetrik, idempotent bir matris ve $rank(A) = r$ olsun. rxn boyutlu bir R matrisi için $A = R'R$ ise, $RR' = I$ dır. (Ekni,M.,1999)

İspat: $A = R'R$ eşitliğini sağlayan rxn tipindeki matris R olsun. Bu durumda

$$A^2 = A \Rightarrow R'RR'R = R'R$$

$$\Rightarrow R(R'RR'R)R' = R(R'R)R'$$

$$\Rightarrow (RR')^{-1}(RR'RR'RR')(RR')^{-1} = (RR')^{-1}(RR'RR')(RR')^{-1}$$

$$\Rightarrow RR' = I$$

olur.

Teorem 1.2.17: V n ranklı olmak üzere, $x: nx1$ rasgele vektör $N_n(\mu, V)$ dağılımlı ve $B: qxn$ bir matris olsun. $x'Ax$ karesel formun, Bx lineer formundan bağımsız olması için gerek ve yeter şartın $BVA = 0$ olmasıdır. (Rencher, A.C. and Schaalje, B.G.,2007)

Sonuç 1.2.4: $x : nx1$ rasgele vektör $N_n(\mu, I_n)$ dağılımlı ve $B: qxn$ bir matris olsun. $Bx : qx1$ vektörü, $x'Ax$ karesel formundan bağımsız olması için gerek ve yeter şart $BA = 0$ olmasıdır.

Teorem 1.2.18: V n ranklı olmak üzere, $x: nx1$ rasgele vektör $N_n(\mu, V)$ dağılımlı, B ve A simetrik matrisler olsun. $x'Ax$ ve $x'Bx$ iki karesel formun bağımsız olması için gerek ve yeter şart $AVB = 0$ olmasıdır. (Rencher, A.C. and Schaalje, B.G.,2007)

Sonuç 1.2.5: $x : nx1$ rasgele vektör $N_n(\mu, I_n)$ dağılımlı olsun. $x'Bx$ ve $x'Ax$ karesel formların bağımsız olması için gerek ve yeter şart $BA = 0$(denk olarak $AB = 0$) olmasıdır.

Teorem 1.2.19: X ve Y bağımsız olmak üzere $X \sim N_n(\mu, \mathrm{I})$ ve $\mathrm{Y} \sim \chi_n^2$ dağılımlı olsun. Bu takdirde,

$$t = \frac{X}{\sqrt{Y/n}} \tag{1.2.45}$$

n serbestlik dereceli ve μ merkezi olmama parametreli, merkezi olmayan bir t dağılımına sahiptir.

$$\frac{X}{\sqrt{Y/n}} \sim t_{n,\mu} \tag{1.2.46}$$

şeklinde yazılır. Eğer $\mu = 0$ ise, merkezi t dağılımına sahiptir denir. (Rencher, A.C. and Schaalje, B.G.,2007)

Teorem 1.2.20: Eğer, $\mathrm{Q}_1 \sim \chi^2_{m,\lambda}$ ve $\mathrm{Q}_2 \sim \chi^2_n$ ve Q_1 ve Q_2 bağımsız ise, bu takdirde

i) $F = \frac{\mathrm{Q}_1/\mathrm{m}}{\mathrm{Q}_2/\mathrm{n}}$ oranı merkezi olmayan bir $F_{m,n,\lambda}$ dağılımna sahiptir.

ii) Eğer λ=0 ise, bu takdirde $F \sim F_{m,n}$ merkezi F dağılımına sahiptir.

iii) Eğer $m = 1$ ise, bu takdirde $\sqrt{F}$ merkezi olmayan bir $\mathrm{t}_{\mathrm{n},\sqrt{\lambda}}$ dağılımına sahiptir veya eğer λ=0 ise, merkezi bir t_n dağılıma sahiptir. (Rao, C. R. and Toutenburg, H.,1999)

Tanım 1.2.10: $\mathrm{x}_1, \mathrm{x}_2, \dots, \mathrm{x}_\mathrm{n}$ rasgele vektörleri karşılıklı bağımsız ve her biri $N_n(\mu, \mathrm{V})$ dağılımına sahip olsunlar. Bu takdirde,

$$L(\mu, V) = \frac{1}{(\sqrt{2\pi})^n |V|^{n/2}} e^{-\frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)' V^{-1}(x_i-\mu)} \tag{1.2.47}$$

fonksiyonuna olabilirlik (Likelihood) fonksiyonu denir. Bu fonksiyonu maksimumlaştıran parametreleri bulma tekniğine en çok olabilirlik (maksimum likelihood) tahmini ve maksimumlaştırmayı sağlayan parametrelere de en çok olabilirlik tahminleri denir. (Seber, G.A.F. ,1977)

Tanım 1.2.11 : Bir θ $(\theta \in \Omega)$ parametresi için önerilen bir T tahmin edicisi her $\theta \in \Omega$ için

$$E(T) = \theta \tag{1.2.48}$$

özelliğine sahipse, bu tahmin ediciye yansız tahmin edici denir. Burada Ω parametre kümesidir(parametre uzayıdır).

Tanım 1.2.12: n değişkenli reel değerli bir fonksiyon,

$$f: R^n \to R$$

$$(x_1, x_2, \dots, x_n) \to f(x_1, x_2, \dots, x_n)$$

olmak üzere bu fonksiyonun birinci dereceden $\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots, \frac{\partial f}{\partial x_n}$ türevleri sürekli olduğunda,

$$\frac{\partial f(x_1, x_2, \dots, x_n)}{\partial x_1} = 0, \quad \frac{\partial f(x_1, x_2, \dots, x_n)}{\partial x_2} = 0, \quad \frac{\partial f(x_1, x_2, \dots, x_n)}{\partial x_n} = 0 \qquad (1.2.49)$$

denklem sisteminin bir

$$(x_1{}^*, x_2{}^*, \dots, x_n{}^*) \qquad (1.2.50)$$

çözümü(kökü) f nin ekstremum (maksimum, minimum) noktasıdır. Ayrıca, ikinci dereceden kısmi türevler de sürekli ve $(x_1{}^*, x_2{}^*, \dots, x_n{}^*) \epsilon R^n$ noktasında,

$$H = \frac{\partial f(x_1{}^*, x_2{}^*, \dots, x_n{}^*)}{\partial x_i \partial x_j} \qquad (1.2.51)$$

matris pozitif kararlı (kesin-belirli) ise, $(x_1{}^*, x_2{}^*, \dots, x_n{}^*)$ noktasında fonksiyon minimum (lokal minimum) değerine sahiptir. H matrisi negatif kararlı ise, bu noktada fonksiyon maksimum değerine sahiptir.

$g: R^n \to R$ olmak üzere f fonksiyonunun

$$g(x_1, x_2, \dots, x_n) = 0 \qquad (1.2.52)$$

kısıtlaması altında maksimum ve minimum noktaları,

$$\ell(x_1, x_2, \dots, x_n, \lambda) = f(x_1, x_2, \dots, x_n) - \lambda g(x_1, x_2, \dots, x_n) \qquad (1.2.53)$$

fonksiyonunun $\frac{\partial \ell}{\partial x_1}, \frac{\partial \ell}{\partial x_2}, \dots, \frac{\partial \ell}{\partial x_n}, \frac{\partial \ell}{\partial \lambda}$ birinci türevlerini sıfır yapan noktalardır. ℓ fonksiyonuna Langrange fonksiyonu, λ ya Langrange çarpanı ve bu yönteme de Langrange çarpanları yöntemi denir.

Kısıtlama sayısı birden çok olduğunda $i = 1,2, \dots, k$ için,

$$g_i: R^n \to R \qquad (1.2.54)$$

olmak üzere f fonksiyonunun,

$$g_1(x_1, x_2, \dots, x_n) = 0$$

$$g_2(x_1, x_2, \dots, x_n) = 0$$

$$\vdots$$

$$g_k(x_1, x_2, \dots, x_n) = 0$$

kısıtlamaları altında maksimum veya minimum noktaları,

$$\ell(x_1, x_2, \dots, x_n, \lambda_1, \lambda_2, \dots, \lambda_k) = f(x_1, x_2, \dots, x_n) - \sum_{i=1}^{k} \lambda_i g_i(x_1, x_2, \dots, x_n) \tag{1.2.55}$$

fonksiyonunun $\frac{\partial \ell}{\partial x_1}, \frac{\partial \ell}{\partial x_2}, \dots, \frac{\partial \ell}{\partial x_n}, \frac{\partial \ell}{\partial \lambda_1}, \frac{\partial \ell}{\partial \lambda_2}, \dots, \frac{\partial \ell}{\partial \lambda_k}$ birinci türevlerini sıfır yapan noktalardır**. (**Akdeniz F. and Öztürk F.,1996)

BÖLÜM 2

LİNEER MODELLER

2.1 Lineer Modeller ve Lineer Modellerde Parametre Tahmini

Y rasgele değişkenlerin bir $nx1$ gözlenebilir vektörü, X reel sayıların (bilinenlerin, açıklayıcı değişkenlerin) bir $nxp(n > p)$ matrisi, β bilinmeyen, fakat tahmin edilebilen parametrelerin bir $px1$ vektörü ve ε, $E(\varepsilon) = 0$, $Kov(\varepsilon) = \sigma^2 V$ olmak üzere, rasgele değişkenlerin bir gözlenebilir olmayan $nx1$ vektörü olsun, bu nicelikler

$$Y = X\beta + \varepsilon \tag{2.1.1}$$

ile bağlanmış ise, bu takdirde (2.1.1) bağıntısı bir genel lineer modeli tanımlar.

X matrisi, tam sütun ranklı ise, yani $rank(X_{nxp}) = p \leq n$ ise, (2.1.1) modeline tam ranklı lineer model, $rank(X_{nxp}) = r \leq p$ yani, X matrisi tam ranklı değilse, (2.1.1) modeline eksik ranklı bir lineer model denir. Ayrıca, $\sigma^2 > 0$ bilinmeyen fakat tahmin edilebilir bir parametredir.

Bir lineer model ε'nun dağılımına, V varyans-kovaryans matrisine, X'in yapısına ve rankına bağlı olarak ayrı ayrı incelenebilir.

1. **Durum**: $\varepsilon \sim N(0, \sigma^2 I_n)$ dağılımına sahiptir.
2. **Durum:** ε bilinmeyen bir dağılıma sahiptir. Ortalaması ve varyans-kovaryans matrisi sırasıyla $E(\varepsilon) = 0, Kov(\varepsilon) = \sigma^2 I_n$ ile gösterilir.
3. **Durum:** $E(\varepsilon) = 0$ ve $Kov(\varepsilon) = V$ dir. Burada V bilinen pozitif kararlı bir matristir.

1. durum: Her bir ε_i, beklenen değeri sıfır ve σ^2 varyanslı normal dağılıma sahiptir ve ε_i, $i = 1,2, \ldots, n$, ler bağımsızdırlar. (2.1.1) modelinde parametreler hakkında nokta tahmini, aralık tahmini ve hipotez testleri düşünülebilir. Nokta tahmini aşağıda verilmiştir. Aralık tahmini ve hipotez testlerini ileriki kısımlarda ele alınacaktır.

(2.1.1) modelinde 1. durum sağlansın. Modelin parametre kümesi

$$\Omega = \{(\beta, \sigma^2): \beta \in \mathbb{R}^p, \sigma^2 > 0\} \tag{2.1.2}$$

olsun. Bu durumda $Y \sim N(X\beta, \sigma^2 I_n)$ dir. β ve σ^2'nin en çok olabilirlik (maximum likelihood) tahmin edicilerini bulmak için (1.2.47) bağıntısından,

$$L(\beta, \sigma^2; Y) = \frac{1}{(\sqrt{2\pi})^n (\sigma^2)^{n/2}} e^{-\frac{1}{2\sigma^2}(Y - X\beta)'(Y - X\beta)} \tag{2.1.3}$$

dır. $lnL(\beta, \sigma^2; Y)$, β ve σ^2'ye göre türevleri alınır ve sıfıra eşitlenirse,

$$\frac{\partial}{\partial\beta}\left(lnL(\beta, \sigma^2; Y)\right) = -\frac{1}{2\sigma^2}(-2X'Y + 2X'X\beta) = 0 \tag{2.1.4}$$

$$\frac{\partial}{\partial\sigma^2}\left(lnL(\beta, \sigma^2; Y)\right) = -\frac{n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2}(Y - X\beta)'(Y - X\beta) = 0 \tag{2.1.5}$$

denklemleri elde edilir. (2.1.4) ve (2.1.5) bağıntılarından,

$$X'X\beta = X'Y \tag{2.1.6}$$

$$\sigma^2 = \frac{(Y - X\beta)'(Y - X\beta)}{n} \tag{2.1.7}$$

olur. (2.1.6) denklemine modelin normal denklemi denir. Normal denklemde, X tam sütun ranklı ise ($X'X$ tekil değilse), (2.1.6) nın tek bir çözümü

$$\hat{\beta} = (X'X)^{-1}X'Y \tag{2.1.8}$$

dir. Bu değer (2.1.7)'de yerine yazılırsa,

$$\begin{aligned}
\tilde{\sigma}^2 &= \frac{(Y - X\hat{\beta})'(Y - X\hat{\beta})}{n} \\
&= \frac{1}{n}[Y'(I_n - X(X'X)^{-1}X')Y] \\
&= \frac{1}{n}[Y'(I_n - XX^-)Y]
\end{aligned} \tag{2.1.9}$$

elde edilir.

Eğer X tam ranklı değil (eksik ranklı) ise,

$$X'X\beta = X'Y \tag{2.1.10}$$

normal denklem sistemi daima tutarlı olduğundan, w keyfi bir reel vektör olmak üzere,

$$\beta^* = (X'X)^- X'Y + (I_n - (X'X)^- X'X)w = X^+Y + \left(I_p - X^+X\right)w, \tag{2.1.11}$$

çözümüne sahiptir. Burada $w = 0$ alınırsa,

$$\beta^* = (X'X)^- X'Y \tag{2.1.12}$$

özel çözümü elde edilir. Bu değer

$$\sigma^2 = \frac{(Y-X\beta)'(Y-X\beta)}{n} \tag{2.1.13}$$

bağıntısında yerine yazılırsa,

$$\sigma^{*2} = \frac{(Y-X\beta^*)'(Y-X\beta^*)}{n}$$

$$= \frac{1}{n}[Y'(I_n - X(X'X)^- X'\ ) Y] = \frac{1}{n}[Y'(I_n - XX^-) Y] \tag{2.1.14}$$

elde edilir.

Teorem 2.1.1: X tam ranklı olduğunda 1. durum için aşağıdaki özellikler verilebilir:

i) $\forall\, \beta \in \mathbb{R}^p$ için $E(\hat{\beta})$=β olup $\hat{\beta}$, β'nın yansız tahmin edicisidir. $Var\left(\hat{\beta}\right) = \sigma^2(X'X)^{-1}$ dir. Bu durumda $\hat{\beta}\sim N(\beta, \sigma^2(X'X)^{-1})$ olmaktadır.

ii) $E(\tilde{\sigma}^2) = \frac{n-p}{n}\sigma^2$ olmak üzere $\tilde{\sigma}^2, \sigma^2$ için yansız bir tahmin edici değildir. Ancak, $\tilde{\sigma}^2 = \frac{1}{n-p}[Y'(I_n - X(X'X)^{-1}X'\ ) Y]$ alındığında

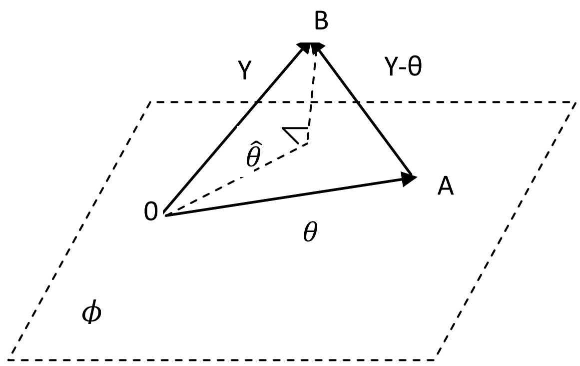
$$E(\tilde{\sigma}^2) = E\left(\frac{n}{n-p}\tilde{\sigma}^2\right) = \frac{n}{n-p}E(\tilde{\sigma}^2) = \sigma^2 \tag{2.1.15}$$

olduğundan $\tilde{\sigma}^2, \sigma^2$ için yansız bir tahmin edicidir. $\tilde{\sigma}^2$ tahmin edicisine yansızlık için düzeltilmiş en çok olabilirlik tahmin edicisi denir. $\tilde{\sigma}^2$ tahmininin dağılımı için $Q = [Y'(I_n - X(X'X)^{-1}X'\ ) Y]$ karesel formunu göz önüne alarak $\frac{Q}{\sigma^2}\sim\chi^2_{(\mathrm{n-p})}$ olduğu görülür.

iii) $\hat{\beta}$ ve $\tilde{\sigma}^2$ en çok olabilirlik tahmin edicileri bağımsızdır.

iv) $\hat{\beta}$ ve $\tilde{\sigma}^2$, β ve σ^2 için yeterli istatistiklerdir. (Graybill, F.A.,1976,)

2.Durum: Her bir ε_i'nin beklenen değeri sıfır, ε_i'ler ilişkisiz ve ε_i'ler bilinmeyen ortak σ^2 varyansına sahiptirler. Yani, $\varepsilon \sim (0, \sigma^2 I_n)$ ve $rank(X_{\text{nxp}}) = p$ olduğu kabul edilirse, (2.1.1) modelinde Y gözlem vektörünün dağılımı bilinmez. Bundan dolayı β ve σ^2 parametrelerinin en çok olabilirlik tahmin edicileri söz konusu değildir. β nın bir tahminini elde etmenin bir yöntemi (bilinen) en küçük kareler yöntemidir. Bu yöntem, $\theta = X\beta$ koyarak, β'ya göre $\sum_i \varepsilon_i^2$ nin minimumlaştırılmasından oluşur, yani, ϕ , X'in sütun uzayı olmak üzere $\theta \in \mathcal{R}(X) = \phi$ ye bağlı $\varepsilon'\varepsilon = ||Y - \theta||^2$ ifadesini minimumlaştırırız. Eğer, ϕ'deki θ değişkenini değiştirirsek, $||Y - \theta||^2$ ($Y - \theta$ nın uzunluğunun(normunun) karesi) $(Y - \hat{\theta}) \perp \perp \phi$ olduğunda, $\theta = \hat{\theta}$ için minimum olacaktır.



Şekil 2.1

Böylece,

$$X'(Y - \hat{\theta}) = 0 \tag{2.1.16}$$

veya

$$X'\hat{\theta} = X'Y \tag{2.1.17}$$

dir. Burada $\hat{\theta}$ tek olarak belirlenir, yani, $\hat{\theta}$, ϕ üzerinde Y'nin tek ortogonal izdüşümüdür. X'in verilen sütunları bağımsızdır ve $\hat{\theta} = X\hat{\beta}$ olacak şekilde bir tek $\hat{\beta}$ vektörü vardır. Bu nedenle, $\hat{\theta}$, (2.1.17) bağıntısında yerine yazılırsa,

$$X'X\hat{\beta} = X'Y \tag{2.1.18}$$

normal denklemi elde edilir. X tam sütun ranklı olduğundan, $X'X$ pozitif kararlıdır ve bu nedenle singüler değildir. Böylece (2.1.18) denkleminin çözümü

$$\hat{\beta} = (X'X)^{-1}X'Y \tag{2.1.19}$$

olacaktır.

$\hat{\beta}$, $(\beta'X'Y = (\beta'X'Y)' = Y'X\beta$ gerçeğini kullanarak) $\varepsilon'\varepsilon$ nun β'ya göre türevini alarakta çıkarılabilir. Bunun için

$$\varepsilon'\varepsilon = (Y - X\beta)'(Y - X\beta)$$

$$= Y'Y - 2Y'X\beta + \beta'X'X\beta \tag{2.1.20}$$

yazılır. Böylece, $\frac{\partial\varepsilon'\varepsilon}{\partial\beta} = 0$ eşitliğinden

$$-2X'Y + 2X'X\beta = 0 \qquad (2.\ 1.21)$$

veya

$$X'X\beta = X'Y \tag{2.1.22}$$

elde edilir. Bu denklemin β için çözümü bize $\varepsilon'\varepsilon$ nun bir sabit değerini(yani, $\varepsilon'\varepsilon$ nun 1. mertebeden diferansiyelinin sıfır olduğu noktayı) verir ve basit bir cebirsel özdeşlik $\varepsilon'\varepsilon$ nun $\hat{\beta}$ da minimum olacağını gösterir. Burada

$$\hat{\beta} = (X'X)^{-1}X'Y \tag{2.1.23}$$

dir. Uygun $X\hat{\beta}$ regresyonu $\hat{Y} = (= [(\hat{Y}_i)])$ ve

$$e = Y - \hat{Y} = Y - X\hat{\beta}$$

$$= (I_n - X(X'X)^{-1}X')Y$$

$$= (I_n - P)Y \tag{2.1.24}$$

elemanlarına artıklar denir. Burada $P = X(X'X)^{-1}X'$ dür. $\varepsilon'\varepsilon$ nun minimum değeri

$$e'e = (Y - X\hat{\beta})'(Y - X\hat{\beta})$$

$$= Y'Y - 2\hat{\beta}X'Y + \hat{\beta}'X'X\hat{\beta}$$

$$= Y'Y - \hat{\beta}X'Y + \hat{\beta}'\left(\underbrace{X'X\hat{\beta} - X'Y}_{0}\right)$$

$$= Y'Y - \hat{\beta}X'Y$$

$$= Y'(I_n - X(X'X)^{-1}X')\mathrm{Y} \qquad (2.1.25)$$

olacaktır. (2.1.25) bağıntısına artık kareler toplamı (AKT) denir.

$\varepsilon'\varepsilon$ kareler toplamını minimumlaştırmak σ^2 parametresinin bir tahminini elde etmez; ancak β'nın en küçük kareler tahminine($\hat{\beta}$ ya) bağlı olan σ^2'nin yansız bir tahmini

$$\hat{\sigma}^2 = \frac{AKT}{n-p}$$

ile elde edilir. Burada $E(\hat{\sigma}^2) = \sigma^2$ olduğu kolayca görülür. (Seber, G.A.F., 1977)

3. Durum: (2.1.1) modelinde $Kov(\varepsilon) = V$, V bilinen pozitif kararlı, simetrik, n ranklı ve singüler olmayan bir matris olduğundan $P'VP = \Lambda$ olacak şekilde ortogonal bir P matrisi bulunabilir. Burada Λ, V matrisinin pozitif öz değerlerinin nxn boyutlu bir köşegen matrisidir.

$P'VP = \Lambda$ eşitliğinden $V = P\Lambda P'$ ve buradan $V^{-1} = P'\Lambda^{-1}P$ elde edilir. Böylece $\Lambda^{-1/2}P'$ dönüşüm matrisi olarak seçildiğinde, (2.1.1) modeli aşağıdaki modele dönüştürülür.

$$\underbrace{\Lambda^{-1/2}P'Y}_{Z} = \underbrace{\Lambda^{-1/2}P'X}_{W}\beta + \underbrace{\Lambda^{-1/2}P'\varepsilon}_{\delta}. \qquad (2.1.26)$$

Gerçekten bu modelde

$$E(\delta) = \Lambda^{-1/2}P'E(\varepsilon)$$

$$= \Lambda^{-1/2}P'0 = 0 \qquad (2.1.27)$$

$$Var(\delta) = \Lambda^{-1/2}P'VP\Lambda^{-1/2}$$

$$= \Lambda^{-1/2}\Lambda\Lambda^{-1/2}$$

$$= \Lambda^{1/2}\Lambda^{-1/2}$$

$$= I_n \qquad (2.1.28)$$

dır.

O halde, β'nın en küçük kareler tahmin edicisi

$$\breve{\beta} = ((\Lambda^{-1/2}P'X)'\,(\Lambda^{-1/2}P'X))^{-1}(\Lambda^{-1/2}P'X)'\,(\Lambda^{-1/2}P'Y)$$

$$= (X'P\Lambda^{-1/2}\Lambda^{-1/2}P'X)^{-1}X'P\Lambda^{-1/2}\Lambda^{-1/2}P'Y$$

$$= (X'V^{-1}X)^{-1}X'V^{-1}Y \qquad (2.1.29)$$

olacaktır.

Teorem 2.1.2: (2.1.1) lineer modelini ele alalım. $\ell: px1$ herhangi bir verilmiş sabit sayılar vektörü olmak üzere β nın belirtilmiş bir lineer fonksiyonu, $\ell'\beta$ nın tahmin edilebilir bir fonksiyon olması için gerek ve yeter şart aşağıdaki şartların herhangi birinin sağlanmasıdır:

i. ℓ, X' matrisinin sütunlarının bir lineer kombinasyonudur.(bu ifadenin başka şekilleri aşağıdaki gibidir: ℓ, X' matrisinin sütun uzayındadır; ℓ', X' matrisinin satırlarının bir lineer kombinasyonudur; veya, ℓ', X' matrisinin satır uzayındadır.)

ii. $rank[X' \vdots \ell] = rank[X']$

iii. $rank[X'X] = rank[X'X \vdots \ell]$

iv. $X'Xr = \ell$ denklemleri için bir r çözüm vektörü vardır.

v. X in herhangi bir g-tersi için $\ell'X^{-}X = \ell'$ dür.

vi. X in herhangi bir g-tersi için $X'(X')^{-}\ell = \ell$ dir.

vii. $X'X$ in herhangi bir g-tersi için $(X'X)(X'X)^{-}\ell = \ell$ dir.

viii. $X'X$ in herhangi bir g-tersi için $\ell'(X'X)^{-}(X'X) = \ell'$ dür. **(**Graybill, F. A., 1976)

2.2 Gauss Markov, Aitken ve Rao En Küçük Kareler Tahmin Edicileri

$Y: nx1$, $X: nxp\ (p < n)$, $\beta: px1$ ve $\varepsilon: nx1$ matrisleri olmak üzere,

$$Y = X\beta + \varepsilon,$$

$$E(\varepsilon) = 0, \quad E(\varepsilon\varepsilon') = \sigma^2 V \qquad (2.2.1)$$

lineer modeli gözönüne alınsın. X ve V'nin rankları üzerinde çeşitli varsayımlar altında $(Y - X\beta)$'nın bir kuadratik fonksiyonunun minimumlaştırılması suretiyle σ^2 ve β'nın tahmini yeniden gözden geçirilsin.

Tanım 2.2.1: $V = I_n$ ve $rank(X) = p$ (yani, X, p tam rankına sahip) olsun. Bu şartlar altında, β'nın minimum varyans lineer tahmin edicisinin

$$\hat{\beta} = \min_{\beta}(Y - X\beta)'\,(Y - X\beta), \tag{2.2.2}$$

bunun açık bir şekli

$$Var(\hat{\beta}) = \sigma^2(X'X)^{-1} \tag{2.2.3}$$

olmak üzere,

$$\hat{\beta} = (X'X)^{-1}X'Y \tag{2.2.4}$$

olduğu ve σ^2 nin bir yansız tahmin edicisinin de

$$\hat{\sigma}^2 = \frac{(Y-X\hat{\beta})'(Y-X\hat{\beta})}{n-p} \tag{2.2.5}$$

olduğu daha önce gösterildi. Bu yönteme Gauss-Markov en küçük kareler yöntemi denir.

$V = I_n$ ve $rank(X) = k < p$ (yani, X eksik ranka sahiptir ya da tam ranklı değildir.) olsun. L: pxk ve $\mathcal{R}(L) \subset \mathcal{R}(X')$, yani, L nin sütunları ile gerilen lineer uzay, X' nün sütunları ile gerilen lineer uzayda ihtiva edilmek üzere, $L'\beta$ nın minimum varyans lineer tahmin edicisi, bu şartlar altında

$$Var(L'\hat{\beta}) = \sigma^2 L'(X'X)^{-}L \tag{2.2.6}$$

ve

$$\hat{\beta} = \min_{\beta}(Y - X\beta)'\,(Y - X\beta) \tag{2.2.7}$$

olmak üzere,

$$L'\hat{\beta} = L'(X'X)^{-}X'Y \tag{2.2.8}$$

dir. σ^2 nin bir yansız tahmin edicisi ise,

$$\hat{\sigma}^2 = \frac{(Y-X\hat{\beta})'(Y-X\hat{\beta})}{n-k} \tag{2.2.9}$$

dır. (Rao, C. R., Toutenburg, H.,1999)

Tanım 2.2.2 : V pozitif kararlı ve $rank(X) = p$ olsun. Bu durumda, β'nın minimum varyans lineer yansız tahmin edicisinin

$$Var(\hat{\beta}) = \sigma^2(X'V^{-1}X)^{-1} \tag{2.2.10}$$

olmak üzere, açık bir şekli

$$\hat{\beta} = (X'V^{-1}X)^{-1}X'V^{-1}Y \tag{2.2.11}$$

olan,

$$\hat{\beta} = \min_{\beta}(Y - X\beta)'\,V^{-1}(Y - X\beta) \tag{2.2.12}$$

olduğu da daha önce gösterildi. Bu durumda, σ^2 nin bir yansız tahmin edicisi

$$\hat{\sigma}^2 = \frac{(Y-X\hat{\beta})'\,V^{-1}(Y-X\hat{\beta})}{n-p} \tag{2.2.13}$$

dır. Bu yönteme de Aitken en küçük kareler yöntemi denir.

V pozitif kararlı ve $rank(X) = k < p$ olsun. Bu durumda, L yukarıdaki aynı şartları sağlamak üzere, $L'\beta$ nın minimum varyans lineer tahmin edicisi

$$\hat{\beta} = \min_{\beta}(Y - X\beta)'\,V^{-1}(Y - X\beta) \tag{2.2.14}$$

ve

$$L'\hat{\beta} = L'(X'V^{-1}X)^{-}X'V^{-1}Y \tag{2.2.15}$$

dir. $L'\hat{\beta}$ nın varyans-kovaryans matrisi

$$Var(L'\hat{\beta}) = \sigma^2\, L'(X'V^{-1}X)^{-}L \tag{2.2.16}$$

dir. σ^2 nin bir yansız tahmin edicisi ise,

$$\hat{\sigma}^2 = \frac{(Y-X\hat{\beta})'V^{-1}(Y-X\hat{\beta})}{n-k} \tag{2.2.17}$$

dir. Böylece, Aitken en küçük kareler yöntemi, X'in eksik ranklı olması duruma genişletilebilir.

En küçük kareler teorisi, $(Y - X\beta)$'nın uygun bir kuadratik fonksiyonunun minimumlaştırılmasıyla hem X'in hem de V'nin eksik ranklı olabilmesi durumu, Rao, C.R. (1973a) tarafından incelenmiştir ve çözümü aşağıdaki gibidir:

$rank(V) = s \leq n$, $rank(X) = k \leq p$ olsun. İlk olarak bir teorem ifade edelim:

Teorem 2.2.1: U, $\mathcal{R}(V) \subset \mathcal{R}(R)$ ve $\mathcal{R}(X) \subset \mathcal{R}(R)$ olacak şekilde negatif kararlı olmayan bir matris olmak üzere, $R = V + XUX'$ olsun. Bu takdirde,

i. $X(X'R^{-}X)^{-}X'RX = X$,

ii. Eğer $X'M = 0$ ise, $X(X'R^{-}X)^{-}X'R^{-}RM = 0$,

iii. $iz(R^{-}R - X(X'R^{-}X)^{-}X') = rank[V \vdots X] - rank(X)$ dir.

Sonuçları, g-terslerin özelliklerini kullanarak ispatlamak kolaydır. **i. ii. iii.** deki tüm ifadelerin g-tersin herhangi bir seçimi için değişmez kaldığına dikkat edin ve sonuçları ispatlamak için Moore-Penrose g-tersi kullanmak uygundur. (Rao, C. R., Toutenburg, H.,1999)

Teorem 2.2.2: $\hat{\beta}$, içerilen g-terslerin herhangi bir seçimi için

$$\hat{\beta} = \min_{\beta}(Y - X\beta)'\, R^{-}(Y - X\beta) \qquad (2.\ 2.18)$$

nın bir çözümü olan

$$\hat{\beta} = (X'R^{-}X)^{-}X'R^{-}Y \qquad (2.\ 2.19)$$

olsun. Bu takdirde

i. $L: pxk$ ve $\mathcal{R}(L) \subset \mathcal{R}(X')$ olmak üzere $L'\beta$ nın minimum varyans lineer tahmin edicisi,

$$Var(L'\hat{\beta}) = \sigma^2\, L'\{(X'R^{-}X)^{-} - U\}L \qquad (2.2.20)$$

varyans kovaryans matrisine sahip olan $L'\hat{\beta}$ dır.

ii. σ^2 nin bir yansız tahmin edicisi ise, $f = rank[V \vdots X] - rank(X)$ olmak üzere,

$$\hat{\sigma}^2 = \frac{(Y-X\hat{\beta})'R^-(Y-X\hat{\beta})}{f} \tag{2.2.21}$$

dir. (Rao, C. R., Toutenburg, H.,1999)

İspat: $\mathcal{R}(L) \subset \mathcal{R}(X')$ olduğundan $L = X'C$ olsun. Bu takdirde, Teorem 2.2.1'in **i.** den,

$$E(L'\hat{\beta}) = C'X(X'R^-X)^-X'R^-E(Y)$$

$$= C'X(X'R^-X)^-X'R^-X\beta$$

$$= C'X\beta = L'\beta \tag{2.2.22}$$

dır. Bundan dolayı $L'\hat{\beta}$, $L'\beta$ için yansızdır. $M'Y$'yi $E(M'Y) = 0$, yani $M'X = 0$, olacak şekilde alınırsa, Teorem 2.2.1(**ii.**) den

$$Kov\big(L'\hat{\beta}, M'Y\big) = \sigma^2 C'X(X'R^-X)^-X'R^-WM$$

$$= \sigma^2 C'X(X'R^-X)^-X'R^-RM$$

$$= \sigma^2 C'X(X'R^-X)^-X'R^-RM = 0 \tag{2.2.23}$$

olur. Bu $E(M'Y) = 0$ olacak şekilde, her M için doğrudur, bu nedenle, $L'\hat{\beta}$, $L'\beta$'nın bir yansız tahmin edicisi olarak minimum varyans-kovaryans matrisine sahiptir. $L'\hat{\beta}$ nın varyans-kovaryans matrisi için aşağıdaki ifade elde edilir:

$$Var(L'\hat{\beta}) = \sigma^2\, C'X(X'R^-X)^-X'R^-V(C'\,X(X'R^-X)^-X'\,R^-)$$

$$= \sigma^2\, C'[(X'R^-X)^- - U]X'C$$

$$= \sigma^2\, L'[(X'R^-X)^- - U]L. \tag{2.2.24}$$

Son olarak, Teorem 2.2.1 (**iii.**) den, σ^2'nin yansız tahmin edicisini veren

$$E(Y - X\hat{\beta})'R^-(Y - X\hat{\beta} = E(Y - X\beta)'\,R^-(Y - X\beta) - E(Y - X\beta)'\,R^-(X\beta - X\hat{\beta})$$

$$= \sigma^2 iz[\,R^-V - \,R^-X(X'R^-X)^-X'R^-V]$$

$$= \sigma^2 iz\,R^-[I_n - X(X'R^-X)^-X'R^-]R$$

$$= \sigma^2 iz[\,R^-R - izX(X'R^-X)^-X'R^-]$$

$$= \sigma^2(rank[V \vdots X] - rank(X)) \tag{2.2.25}$$

elde edilir. (Rao, C. R., Toutenburg, H.,1999)

2.3. $\boldsymbol{R\beta = r}$ Kesin Lineer Kısıtlaması Altında Parametre Tahmini

X'in tam ranklı olduğu durum: $X: nxp$, rankı p olan bir matris olmak üzere, (2.1.1) modeli ve 1. Durum ele alınsın. R_{qxp} $(q < p)$ bilinen bir matris, $rank(R) = q$ (yani, R tam satır ranklı) ve r: $qx1$ de bilinen bir vektör olmak üzere, β parametre vektörü üzerinde $R\beta = r$ tutarlı kesin lineer kısıtlaması altında β'nın en iyi lineer yansız tahmin edicisini bulmak için kullanılan bir yöntem, Langrange çarpanları yöntemidir. Langrange fonksiyonu

$$L(\beta,\lambda) = (Y - X\beta)'(Y - X\beta) + \lambda'(R\beta - r) \tag{2.3.1}$$

dir. Burada $\lambda: qx1$ Langrange çarpanları vektörüdür. Langrange fonksiyonunun β'ya göre türevi alınırsa,

$$\frac{\partial L(\beta,\lambda)}{\partial \beta} = -2X'Y + 2X'X\beta + R'\lambda = 0 \tag{2.3.2}$$

denklemi elde edilir. Buradan

$$X'X\beta = X'Y - \frac{1}{2}R'\lambda$$

bağıntısı bulunur ve bu denklemin çözümünden

$$\tilde{\beta} = (\mathrm{X'X})^{-1}\mathrm{X'Y} - \frac{1}{2}(\mathrm{X'X})^{-1}R'\lambda \tag{2.3.3}$$

elde edilir. Ayrıca, (2.3.1) fonksiyonunun λ'ya göre türevi alınırsa,

$$\frac{\partial L(\beta,\lambda)}{\partial \lambda} = R\beta - r = 0$$

bağıntısı veya

$$R\beta = r \tag{2.3.4}$$

olur. β'nın (2.3.3) bağıntısındaki değeri (2.3.4) bağıntısında yerine yazılırsa,

$$R((X'X)^{-1}X'Y - \frac{1}{2}(X'X)^{-1}R'\lambda) = r\ ,$$

$$R\hat{\beta} - \frac{1}{2}R(X'X)^{-1}R'\lambda = r,$$

$$R\hat{\beta} - r = \frac{1}{2}R(X'X)^{-1}R'\lambda,$$

$$\frac{1}{2}\lambda = (R(X'X)^{-1}R')^{-1}(R\hat{\beta} - r) \qquad (2.3.5)$$

elde edilir. Bu ifade (2.3.3) bağıntısında yerine yazılırsa,

$$\tilde{\beta} = \hat{\beta} - (X'X)^{-1}R'(R(X'X)^{-1}R')^{-1}(R\hat{\beta} - r) \qquad (2.3.6)$$

veya

$$\tilde{\beta} = \hat{\beta} + (X'X)^{-1}R'(R(X'X)^{-1}R')^{-1}(r - R\hat{\beta}) \qquad (2.3.7)$$

elde edilir. (Toutenburg, H.,1982)

Burada, $\tilde{\beta}$ tahmininin $\hat{\beta}$'dan farkının, $(R\hat{\beta} - r)$ bağıntısının bir lineer fonksiyonu olduğu görülmektedir. Ayrıca, $\tilde{\beta}$ kısıtlanmış en küçük kareler tahmin edicisi, kısıtlanmamış en küçük kareler tahmin edicisi $\hat{\beta}$ ve bir düzeltme teriminin toplamıdır ki, β'nın bu tahmin edicisi için $R\beta = r$ kesin kısıtlamasını sağlar. Şöyle ki;

$$R\tilde{\beta} = R\left[\hat{\beta} + (X'X)^{-1}R'(R(X'X)^{-1}R')^{-1}\left(r - R\hat{\beta}\right)\right]$$

$$= r \qquad (2.3.8)$$

dir. $\tilde{\beta}$ tahmini değişik şekillerde ifade edilebilir. Bu ifadeler aşağıda verilmiştir:

R tam satır ranklı olduğunda, $RR^{+} = I_q$ olduğu göz önünde bulundurulursa,

$$\tilde{\beta} = \hat{\beta} - (X'X)^{-1}R'(R(X'X)^{-1}R')^{-1}R\hat{\beta} \ + (X'X)^{-1}R'(R(X'X)^{-1}R')^{-1}RR^{+}r \qquad (2.3.9)$$

dır ve $M = (X'X)^{-1}R'(R(X'X)^{-1}R')^{-1}R$ alınırsa, bu durumda

$$\tilde{\beta} = (I - M)\hat{\beta} + MR^{+}r$$

elde edilir. Bu tahmin edicinin beklenen değeri

$$E(\tilde{\beta}) = \hat{\beta} + (X'X)^{-1}R'(R(X'X)^{-1}R')^{-1}(r - R\beta) \tag{2.3.10}$$

dır ve $R\beta = r$ olduğundan $\tilde{\beta}$ yansızdır. Bununla beraber, $\tilde{\beta}$ nın varyans-kovaryans matrisi,

$$Var(\tilde{\beta}) = \sigma^2[(X'X)^{-1} - (X'X)^{-1}R'(R(X'X)^{-1}R')^{-1}R(X'X)^{-1}] \tag{2.3.11}$$

dir. Aşağıdaki ifade $\tilde{\beta}$ tahmin edicisinin, $\hat{\beta}$ tahmin edicisine göre daha küçük bir varyansa sahip olduğunu gösterir.

$$Var(\hat{\beta}) - Var(\tilde{\beta}) = \sigma^2[(X'X)^{-1}R'(R(X'X)^{-1}R')^{-1}R(X'X)^{-1}] \geq 0 \tag{2.3.12}$$

dır. (Toutenburg, H.,1982) Aynı zamanda $\tilde{\beta}$ yı gerekli işlemler yaparak,

$$\tilde{\beta} = \left(I_p - (X'X)^{-1}R'(R(X'X)^{-1}R')^{-1}R\right)\hat{\beta}$$

$$+(X'X)^{-1}R'(R(X'X)^{-1}R')^{-1}r \tag{2.3.13}$$

şeklinde de yazılabilir. Şimdi bu son eşitlikte, R nin Moore-Penrose g-tersi $R^+ = R'(RR')^{-1}$ (R tam satır ranklı olduğundan) olacağından Moore-Penrose g-tersin ilk üç şartını sağlayan g-ters($g_3 -$ ters) $R^{g_3} = (X'X)^{-1}R'(R(X'X)^{-1}R')^{-1}$ olarak alınabilir. Gerçekten $RR^{g_3}R = R, R^{g_3}RR^{g_3} = R^{g_3}$, $(RR^{g_3})' = RR^{g_3}$ dır. Bu nedenle (2.3.13) denkleminde R^{g_3} yerine yazılırsa,

$$\tilde{\beta} = (I - R^{g_3}R)\hat{\beta} + R^{g_3}r \tag{2.3.14}$$

elde edilir.

Şimdi, $R\beta = r$ kesin lineer kısıtlaması altında parametre tahminine değişik bir şekilde bakılabilir. Şöyle ki,

$$Y = X\left(I_q - R^+R + R^+R\right)\beta + \varepsilon$$

$$= X\left(I_q - R^+R\right)\beta + XR^+R\beta + \varepsilon$$

$$= X\left(I_q - R^+R\right)\beta + XR^+r + \varepsilon$$

$$Y - XR^+r = X\left(I_q - R^+R\right)\beta + \varepsilon \tag{2.3.15}$$

elde edilir. $Q_R = \left(I_q - R^+R\right)$ alınsın. Bu durumda,

$$Y - XR^{+}r = X\mathrm{Q_R}\beta + \varepsilon \tag{2.3.16}$$

modeli elde edilir. (Yapar, C., 1979) $\mathrm{Q_R} = \left(I_q - R^{+}R\right)$ simetrik-idempotent olduğundan $\mathrm{Q_R} = QQ'$ olacak şekilde $rank(Q'_{(p-q)xp}) = p - q$ matrisi vardır. QQ' nün simetrik-idempotent olması nedeniyle $Q'Q = I_{p-q}$ dur ve Q' satır ranklı olduğundan $(Q')^{+} = Q(Q'Q)^{-1} = Q$ ve buradan $\mathrm{Q}^{+} = \mathrm{Q}'$ dür. Gerçekten $Q'QQ' = Q'$ dür.

$$R\left(I_q - R^{+}R\right) = 0 \tag{2.3.17}$$

olduğundan

$$R\mathrm{Q_R} = 0 \Longrightarrow RQQ' = 0 \Longrightarrow RQ = 0 \tag{2.3.18}$$

olacaktır. O halde Q', R'nin ortoganal tümleyenidir. Böylece, (2.3.16) modeli

$$Y - XR^{+}r = X\mathrm{QQ}'\beta + \varepsilon \tag{2.3.19}$$

modeline dönüşür. Burada, $\mathrm{QQ}'\beta = \gamma$ olsun. Böylece (2.3.19) modeli

$$Y - XR^{+}r = X\gamma + \varepsilon \tag{2.3.20}$$

şeklinde yazılabilir. Bu durumda, γ'nın en küçük kareler tahmini

$$\hat{\gamma} = (\mathrm{X'X})^{-1}\mathrm{X}'(\mathrm{Y} - XR^{+}r) \tag{2.3.21}$$

olur. (2.3.19) modelinde $\mathrm{QQ}'\beta = \gamma$ olduğundan β'nın tahmini $\tilde{\beta} = QQ'\hat{\gamma}$ dır.

Böylece,

$$\begin{aligned}\tilde{\beta} &= QQ'\hat{\gamma} = QQ'(\mathrm{X'X})^{-1}\mathrm{X}'(\mathrm{Y} - XR^{+}r)\\ &= \left(I_q - R^{+}R\right)(\mathrm{X'X})^{-1}\mathrm{X}'(\mathrm{Y} - XR^{+}r)\\ &= \left(I_q - R^{+}R\right)(\mathrm{X'X})^{-1}\mathrm{X'Y}\end{aligned} \tag{2.3.22}$$

yazılabilir. (Akdeniz, F., 1980)

$R\beta = r$ denkleminden

$$\beta = R^{+}r + \left(I_q - R^{+}R\right)t \qquad (t \text{ keyfi bir vektör}) \tag{2.3.23}$$

elde edilir. Şimdi de $R\tilde{\beta} = r$ denklemini sağlayan t tahmini bulunsun. (2.3.23) denkleminde β yerine $\tilde{\beta}$ ve $\left(I_q - R^+R\right)$ yerine QQ' alınırsa,

$$\tilde{\beta} = R^+r + \left(I_q - R^+R\right)t$$

$$= R^+r + QQ't \tag{2.3.24}$$

denklemi elde edilir. (2.3.24) denklemi t'ye göre çözülürse, $R\tilde{\beta} = r$ elde edilir.

Bu durumda,

$$QQ't = \tilde{\beta} - R^+r$$

$$\tilde{t} = QQ'\tilde{\beta} - QQ'R^+r$$

$$= QQ'\tilde{\beta} - \left(I_q - R^+R\right)R^+r$$

$$= QQ'\tilde{\beta} - (R^+ - R^+RR^+)r$$

$$= QQ'\tilde{\beta} - (R^+ - R^+)r$$

$$= QQ'\tilde{\beta} \tag{2.3.25}$$

dır. (2.3.22) denkleminden $QQ'\tilde{\beta} = QQ'\hat{\beta}$ elde edilir. t, (2.3.23) denkleminde yerine yazılırsa,

$$\tilde{\beta} = (\mathrm{I} - R^+R)\hat{\beta} + R^+r \tag{2.3.26}$$

elde edilir.

Şimdi de, $XQ = B$ ve $\mathrm{Q}'\beta = \gamma$ alınsın ve β nın tahmini bulunsun. Bu durumda (2.3.19) modeli

$$Y - XR^+r = B\gamma + \varepsilon \tag{2.3.27}$$

modeli şeklinde yazılabilir. Böylece γ'nın en küçük kareler tahminini

$$\hat{\gamma} = ((\mathrm{XQ})'\mathrm{XQ})^{-1}(\mathrm{XQ})'(\mathrm{Y} - XR^+r)$$

$$=(\mathrm{X}'\mathrm{Q}'\mathrm{XQ})^{-1}\mathrm{Q}'\mathrm{X}'(\mathrm{Y} - XR^+r)$$

$$=\mathrm{Q}^+(\mathrm{X}'\mathrm{X})^{-1}(\mathrm{Q}')^+Q'X'(\mathrm{Y} - XR^+r)$$

$$=\mathrm{Q}^{+}(\mathrm{X}'\mathrm{X})^{-1}\mathrm{Q}Q'X'(\mathrm{Y}-XR^{+}r)$$

$$=\mathrm{Q}'(\mathrm{X}'\mathrm{X})^{-1}\mathrm{Q}Q'X'(\mathrm{Y}-XR^{+}r) \qquad (2.3.28)$$

şeklinde bulunur. Yani, $\hat{\gamma}=\mathrm{Q}'\tilde{\beta}=\mathrm{Q}'(\mathrm{X}'\mathrm{X})^{-1}\mathrm{Q}Q'X'(\mathrm{Y}-XR^{+}r)$ ve $(Q')^{+}=Q$ olduğundan $\tilde{\beta}=\mathrm{QQ}'(\mathrm{X}'\mathrm{X})^{-1}\mathrm{Q}Q'X'(\mathrm{Y}-XR^{+}r)$ olur. t uygun boyutlu bir rasgele vektör olmak üzere, $R\beta=r$ denkleminden

$$\beta=R^{+}r+\left(I_q-R^{+}R\right)t$$

$$=R^{+}r+Q_Rt$$

$$=R^{+}r+QQ't \qquad (2.3.29)$$

elde edilir. Bu ifade modelde yerine yazılırsa,

$$Y=X(R^{+}r+QQ't)+\varepsilon$$

$$=XR^{+}r+XQQ't+\varepsilon$$

$$Y-XR^{+}r=XQQ't+\varepsilon \qquad (2.3.30)$$

bulunur. t tahmin edilebilir olduğundan t nin tahmini

$$\tilde{t}=\left((XQQ')XQQ'\right)^{-1}(XQQ')'(Y-XR^{+}r)$$

$$=(\mathrm{QQ'X'XQQ'})^{-1}\mathrm{QQ'X'}(Y-XR^{+}r)$$

$$=\mathrm{QQ}'(\mathrm{X}'\mathrm{X})^{-1}\mathrm{QQ'QQ'X'}(Y-XR^{+}r)$$

$$=Q\hat{\gamma} \qquad (2.3.31)$$

olur. t tahmini (2.3.29) denkleminde yerine yazılırsa,

$$\tilde{\beta}=R^{+}r+\mathrm{QQ}'(\mathrm{X}'\mathrm{X})^{-1}\mathrm{QQ'QQ'X'}(Y-XR^{+}r)$$

$$=R^{+}r+Q_R(\mathrm{X'X})^{-1}Q_R\mathrm{X}'(Y-XR^{+}r) \qquad (2.3.32)$$

elde edilir. Burada,

$$Kov\left(\tilde{\beta}\right)=Q_R(\mathrm{X'X})^{-1}Q_R \qquad (2.3.33)$$

ve

$$E(\tilde{\beta}) = \beta \tag{2.3.34}$$

dır.

***X*'in eksik ranklı olduğu durum:** $X{:}\,nxp$, $rank(X) = k < p$ olmak üzere, (2.1.1) genel lineer model ve 1. Durum ele alınsın. R_{qxp} $(q < p)$ bilinen bir matris ve $r{:}\,qx1$ de bilinen bir vektör olsun. β parametre vektörü üzerinde $R\beta = r$ kesin lineer kısıtlaması altında, β nın en iyi lineer yansız tahmin edicisi bulunsun. Bunun için Langrange fonksiyonu

$$L(\beta, \lambda) = (Y - X\beta)'(Y - X\beta) + \lambda'(r - R\beta). \tag{2.3.35}$$

dir. Langrange fonksiyonu β ya göre türevi alınır ve bu türev sıfıra eşitlenirse,

$$\frac{\partial L(\beta,\lambda)}{\partial \beta} = -2X'Y + 2X'X\beta - R'\lambda = 0$$

denklemi elde edilir. Buradan

$$X'X\beta = X'Y + \frac{1}{2}R'\lambda$$

bağıntısı elde edilir ve denklem çözümünden m_1 uygun boyutlu keyfi bir vektör olmak üzere,

$$\breve{\beta} = (\mathrm{X'X})^{-}\mathrm{X'Y} + \frac{1}{2}(\mathrm{X'X})^{-}R'\lambda + (I_q - X^{-}X)m_1 \tag{2.3.36}$$

bulunur. ((2.3.36) bağıntısında $m_1 = 0$ yazılırsa, son tahminimiz yansız en küçük kareler olacaktır ve sadece böyle bir çözüme ihtiyaç vardır.) Bu durumda

$$\breve{\beta} = (\mathrm{X'X})^{-}\mathrm{X'Y} + \frac{1}{2}(\mathrm{X'X})^{-}R'\lambda$$

$$= \hat{\beta} + \frac{1}{2}(\mathrm{X'X})^{-}R'\lambda \tag{2.3.37}$$

olur. (2.3.35) fonksiyonunun λ ya göre türevi alınırsa,

$$\frac{\partial L(\beta,\lambda)}{\partial \lambda} = r - R\beta = 0 \tag{2.3.38}$$

bağıntısını veya $R\beta = r$ eşitliği elde edilir. $R\beta = r$ bağıntısında β nın (2.3.37) deki değeri yerine yazılırsa,

$$R(\hat{\beta} + \frac{1}{2}(X'X)^{-}R'\lambda) = r,$$

$$-\frac{1}{2}R(X'X)^{-}R'\lambda = R\hat{\beta} - r,$$

$$-\frac{1}{2}(X'X)^{-}R'\lambda = R^{+}(R\hat{\beta} - r) + (I_q - R^{+}R)m_2 \qquad (2.3.39)$$

bağıntısı elde edilir. Burada, m_2 uygun boyutlu keyfi bir vektördür. (2.3.39) bağıntısında $m_2 = 0$ yazılırsa,

$$-\frac{1}{2}(X'X)^{-}R'\lambda = R^{+}(R\hat{\beta} - r) \qquad (2.3.40)$$

özel çözümü elde edilir. Buradan

$$\check{\beta} = \hat{\beta} - R^{+}(R\hat{\beta} - r) \qquad (2.3.41)$$

olduğu görülür. Artık, (2.3.41) tahmin edicisi yansızdır. **(**Pore, M.D., 1969)

X'in eksik ranklı olduğu durum için $E(\varepsilon'\varepsilon) = \sigma^2 V$ durumu ele alınsın:

$$\sigma^2 = \varepsilon' V^{-1}\varepsilon = (Y - X\beta)' V^{-1}(Y - X\beta) \qquad (2.3.42)$$

dır. (2.3.42) bağıntısının β ya göre türevi alınır ve sıfıra eşitlenirse,

$$\frac{\partial \varepsilon' V^{-1}\varepsilon}{\partial \beta} = -2X'V^{-1}Y + 2X'V^{-1}X\beta = 0$$

$$X'V^{-1}Y = X'V^{-1}X\beta \qquad (2.3.43)$$

denklemi elde edilir. Buradan

$$\hat{\beta} = (X'V^{-1}X)^{-}X'V^{-1}Y + (I_q - (X'V^{-1}X)^{-}X'V^{-1}X)m_3 \qquad (2.3.44)$$

dur. (2.3.44) bağıntısında $m_3 = 0$ yazılırsa,

$$\hat{\beta} = (X'V^{-1}X)^{-}X'V^{-1}Y$$

özel çözümü elde edilir.

$R\beta = r$ kesin lineer kısıtlaması altında, β nın en iyi lineer yansız tahmin edicisini bulmak için Langrange problemi ilk durumdaki gibi aynıdır. Çözüm aşağıdaki gibidir:

$$L(\beta, \lambda) = (Y - X\beta)'V^{-1}(Y - X\beta) + \lambda'(r - R\beta) \qquad (2.3.45)$$

Langrange fonksiyonunun β ya göre türevi alınır ve bu türev sıfıra eşitlenirse,

$$\frac{\partial L(\beta,\lambda)}{\partial \beta} = -2X'V^{-1}Y + 2X'V^{-1}X\beta - R'\lambda = 0$$

$$X'V^{-1}X\beta = X'V^{-1}Y - \frac{1}{2}R'\lambda \qquad (2.3.46)$$

denklemi elde edilir. Buradan

$$\beta^{*} = (X'V^{-1}X)^{-}X'V^{-1}Y - \frac{1}{2}(X'V^{-1}X)^{-}R'\lambda$$

$$+(I - (X'V^{-1}X)^{-}X'V^{-1}X)m_4 \qquad (2.3.47)$$

dir. (2.3.47) bağıntısında $m_4 = 0$ yazılırsa,

$$\beta^{*} = \hat{\beta} - \frac{1}{2}(X'V^{-1}X)^{-}R'\lambda \qquad (2.3.48)$$

özel çözümü elde edilir. (2.3.45) fonksiyonu λ ya göre türev alınır ve sıfıra eşitlenirse,

$$\frac{\partial L(\beta,\lambda)}{\partial \lambda} = R\beta - r = 0 \qquad (2.3.49)$$

bağıntısı veya $R\beta = r$ olur. Burada, β nın yerine (2.3.48) değeri yazılırsa,

$$R\hat{\beta} - \frac{1}{2}R(X'V^{-1}X)^{-}R'\lambda - r = 0$$

$$\frac{1}{2}R(X'V^{-1}X)^{+}R'\lambda = R\hat{\beta} - r$$

elde edilir.

$$\frac{1}{2}(X'V^{-1}X)^{-}R'\lambda = R^{+}\left(R\hat{\beta} - r\right) + \left(I_q - R^{+}R\right)m_5 \qquad (2.3.50)$$

dir. (2.3.50) bağıntısında $m_5 = 0$ alınırsa,

$$\beta^{*} = \hat{\beta} - R^{+}\left(R\hat{\beta} - r\right) \qquad (2.3.51)$$

özel çözümü elde edilir. **(**Pore, M.D., 1969)

2.4. Kesin Lineer Kısıtlamaların Adım Adım Hesaba Katılması

Lineer kısıtlamaların $r = R\beta$ kümesi q $(q < p)$ tane lineer bağımsız

$$r_j = R_j'\beta, \quad j = 1,2, \dots q \tag{2.4.1}$$

kısıtlamalarına sahiptir. Burada ya iki içi içe (yani, lineer bağımlı) ya da iki ayrık (lineer bağımsız) kısıtlamalar kümesi için kısıtlanmış en küçük kareler tahmin edicileri arasındaki ilişkiler araştırılacaktır.

$r_1 = R_1\beta$ ve $r_2 = R_2\beta$'nın sırasıyla q_1 ve q_2 tane kesin lineer kısıtlamanın ayrık kümeleri oldukları kabul edilsin. Burada $\mathrm{q}_1 + \mathrm{q}_2 = \mathrm{q}$ dur. Kısıtlamaların tam kümesi

$$r = \begin{bmatrix} r_1 \\ \cdots \\ r_2 \end{bmatrix} = \begin{bmatrix} R_1 \\ \cdots \\ R_2 \end{bmatrix} \beta = R\beta \tag{2.4.2}$$

ile gösterilir. R_1, R_2 ve R'nin tam sütun ranklı, yani, $rank(R_1) = \mathrm{q}_1$, $rank(R_2) = \mathrm{q}_2$ ve $rank(R) = q$ oldukları da kabul edilsin. Eğer $\tilde{\beta}_1$, $\tilde{\beta}_2$ ve $\tilde{\beta}$ sırasıyla R_1, R_2 ve R kısıtlama matrislerine karşılık gelen kısıtlanmış en küçük kareler tahmin ediciler ve $\hat{\beta}$ bilinenen küçük kareler tahmin edicisi ise,

$$Var(\tilde{\beta}) \leq Var(\tilde{\beta}_i) \leq Var(\hat{\beta}), \ \ i = 1,2. \tag{2.4.3}$$

elde edilir. (Bu bağıntı iki varyans-kovaryans matrisinin farkının negatif kararlı olmaması anlamındadır.)

$$Var(\hat{\beta}) - Var(\tilde{\beta}_i) \geq 0$$

ve

$$Var(\hat{\beta}) - Var(\tilde{\beta}) \geq 0 \tag{2.4.4}$$

bağıntıları (2.3.12) eşitliğinin bir sonucudur. Bu nedenle, bir kısıtlamalar kümesine diğer başka kısıtlamaları eklemenin genel olarak etkinlikte bir kazanca gidileceğini belirten

$$Var(\tilde{\beta}_1) - Var(\tilde{\beta}) \geq 0 \tag{2.4.5}$$

eşitsizliğinin doğru olduğu kontrol edilmelidir.

(2.4.2) bağıntısının yapısını kullanarak $r = R\beta$ tam kümesi için kısıtlanmış en küçük kareler tahmini aşağıdaki gibi yeniden yazılabilir.

$$\tilde{\beta} = \hat{\beta} + [\, R_1' \quad \vdots \quad R_2'] \begin{bmatrix} R_1(\mathrm{X'X})^{-1}R_1' & R_1(\mathrm{X'X})^{-1}R_2' \\ R_2(\mathrm{X'X})^{-1}R_1' & R_2(\mathrm{X'X})^{-1}R_2' \end{bmatrix}^{-1} \begin{bmatrix} r_1 - R_1 b \\ \cdots\cdots\cdots\cdots \\ r_2 - R_2 b \end{bmatrix} \tag{2.4.6}$$

$$A = R(\mathrm{X'X})^{-1}\mathrm{R}' = \begin{bmatrix} E & F \\ F' & G \end{bmatrix} \tag{2.4.7}$$

$$R_1(\mathrm{X'X})^{-1}R_1' = E, \qquad R_1(\mathrm{X'X})^{-1}R_2' = F, \tag{2.4.8}$$

$$R_2(\mathrm{X'X})^{-1}R_2' = G\ , \qquad G - F'E^{-1}F = H$$

kısaltmalarıyla ve Teorem 1.1.4(**vi)**’den

$$Var(\tilde{\beta}) = \sigma^2[(\mathrm{X'X})^{-1} - (\mathrm{X'X})^{-1}R'(\mathrm{R}(\mathrm{X'X})^{-1}R')^{-1}R(\mathrm{X'X})^{-1}] \tag{2.4.9}$$

elde edilir. Varyans-kovaryans matrisinin parçalanmış yapısı

$$\sigma^{-2}Var(\tilde{\beta}) =$$

$$(\mathrm{X'X})^{-1} - (\mathrm{X'X})^{-1}[\, R_1' \quad \vdots \quad R_2'] \begin{bmatrix} \mathrm{E}^{-1} + \mathrm{E}^{-1}\mathrm{F}\mathrm{H}^{-1}\mathrm{F'}\mathrm{E}^{-1} & -\mathrm{E}^{-1}\mathrm{F}\mathrm{H}^{-1} \\ -\mathrm{H}^{-1}\mathrm{F'}\mathrm{E}^{-1} & \mathrm{H}^{-1} \end{bmatrix} \times$$

$$\begin{bmatrix} R_1 \\ \cdots \\ R_2 \end{bmatrix} (\mathrm{X'X})^{-1} \tag{2.4.10}$$

dir. Ayrıca, $\tilde{\beta}_1$ ve $\tilde{\beta}$’nın kovaryansı

$$E(\, \tilde{\beta}_1 \; - \mathrm{\beta})(\, \tilde{\beta} \; - \beta)' = Kov(\, \tilde{\beta}_1 \,, \tilde{\beta}) \tag{2.4.11}$$

dır.

$$[I \quad \vdots \quad \mathrm{E}^{-1}\mathrm{F}]\mathrm{A}^{-1} = [\mathrm{E}^{-1} \,\vdots \quad 0] \tag{2.4.12}$$

$$R_1'[I \quad \vdots \quad \mathrm{E}^{-1}\mathrm{F}]\mathrm{A}^{-1} \begin{bmatrix} R_1 \\ \cdots \\ R_2 \end{bmatrix} = R_1'\mathrm{E}^{-1}R_1 \tag{2.4.13}$$

ile birlikte

$$\tilde{\beta}_1 - \beta = (X'X)^{-1}(I - R_1{}'E^{-1}R_1(X'X)^{-1})X'\varepsilon \tag{2.4.14}$$

$$\tilde{\beta} - \beta = (X'X)^{-1}(I - [\,R_1' \quad \vdots \quad R_2{}']A^{-1}\begin{bmatrix} R_1 \\ \cdots \\ R_2 \end{bmatrix}(X'X)^{-1})X'\varepsilon \tag{2.4.15}$$

bağıntılar kullanılarak aşağıdaki sonuca ulaşılır:

$$Kov(\,\tilde{\beta}_1, \tilde{\beta}) = Var(\tilde{\beta}) \tag{2.4.16}$$

Teorem 1.1.29'a göre,

$$E(\,\tilde{\beta}_1 - \tilde{\beta})(\,\tilde{\beta}_1 - \tilde{\beta})' \geq 0 \tag{2.4.17}$$

eşitsizliği herhangi bir örneklem için ve beklenen değer için de sağlanır.

(2.4.16) bağıntısı kullanılarak, (2.4.5) bağıntısı aşağıdaki gibi de elde edilebilir:

$$E[\tilde{\beta}_1 - \beta - (\,\tilde{\beta} - \beta)][\tilde{\beta}_1 - \beta - (\,\tilde{\beta} - \beta)]' = Var(\tilde{\beta}_1) + Var(\tilde{\beta}) - 2Kov(\,\tilde{\beta}_1, \tilde{\beta}) = Var(\tilde{\beta}_1) - Var(\tilde{\beta}) \geq 0 \tag{2.4.18}$$

Böylece aşağıdaki teorem bulunur:

Teorem 2.4.1: $rank(R_1) = q_1$ olmak üzere bir $r_1 = R_1\beta$ kesin lineer kısıtlamalar kümesinin elde mevcut olduğu kabul edilsin. $rank(R_2) = q_2$ ve $rank(R) = q = q_1 + q_2$ olmak üzere, başka bir bağımsız $r_2 = R_2\beta$ kümesi eklenirse, bu takdirde

$$Var(\tilde{\beta}_1\,) - Var(\tilde{\beta}) \geq 0 \tag{2.4.19}$$

olmak üzere kısıtlanmış $\tilde{\beta}_1$ ve $\tilde{\beta}$ en küçük kareler tahmin edicileri yansızdırlar.

Bu nedenle, bağımsız kısıtlamaların eklenmesiyle bir kesin kısıtlamalar kümesinin bir adım adım artışı, (2.4.19) bağıntısına göre varyansın bir adım adım azalışını ortaya koyar.

İspat: İspat değişik bir şekilde aşağıdaki gibi verilebilir:

$P = [I \quad \vdots \quad 0]$ olmak üzere, R_1 ve R matrisleri aşağıdaki lineer dönüşümlerle bağlanırlar.

$$R_1 = PR \tag{2.4.20}$$

(2.4.10) bağıntısından parçalanmış A matrisini kullanarak, kovaryans matrislerin farkı

$$\sigma^{-2}\left[Var\left(\tilde{\beta}_1\right) - Var\left(\tilde{\beta}\right)\right] \tag{2.4.21}$$

$$= (\mathrm{X}'\mathrm{X})^{-1}R'(R(\mathrm{X}'\mathrm{X})^{-1}\mathrm{R})^{-1}R(\mathrm{X}'\mathrm{X})^{-1} - (\mathrm{X}'\mathrm{X})^{-1}R_1'(R_1(\mathrm{X}'\mathrm{X})^{-1}R_1')^{-1}R_1(\mathrm{X}'\mathrm{X})^{-1}$$

$$= (\mathrm{X}'\mathrm{X})^{-1}R'\,(\mathrm{A}^{-1} - \mathrm{P}'(\mathrm{PAP}')^{-1}\mathrm{P})\mathrm{R}(\mathrm{X}'\mathrm{X})^{-1} \tag{2.4.22}$$

olarak yazılabilir. Varsayıma göre, $rank(R) = q$ dur. Bu takdirde (bakınız Teorem 1.1.33) bu farkın negatif olmayan kararlı olması için gerek ve yeter şart

$$(\mathrm{A}^{-1} - \mathrm{P}'(\mathrm{PAP}')^{-1}\mathrm{P}) \geq 0 \tag{2.4.23}$$

veya eş değer olarak (Teorem 1.1.18)

$$\mathcal{R}(\mathrm{P}'\mathrm{PA}^{-1}) \subset \mathcal{R}(\mathrm{A}^{-1}) \tag{2.4.24}$$

bağıntısının açık olarak gerçekleşmesidir.

Aynı zamanda, $\tilde{\beta}_1$ ve $\tilde{\beta}_2$ tahminleri karşılaştırılsın. İki

$$r_j = R_j\beta, \quad rank(R_i) = q_j\,, \ (j = 1{,}2) \tag{2.4.25}$$

kısıtlamalar kümesi için kısıtlanmış en küçük kareler tahmin edicileri arasındaki bağıntı araştırılsın. Karşılık gelen tahmin ediciler $(j = 1{,}2)$

$$\tilde{\beta}_j = \hat{\beta} + (\mathrm{X}'\mathrm{X})^{-1}R_j{}'\left(R_j(\mathrm{X}'\mathrm{X})^{-1}R_j{}'\right)^{-1}(r_j - R_j\hat{\beta}) \tag{2.4.26}$$

dır.

$$A_j = R_j(\mathrm{X}'\mathrm{X})^{-1}R_j{}' \tag{2.4.27}$$

$$G_j = (\mathrm{X}'\mathrm{X})^{-1}R_j{}'A_j^{-1}R_j(\mathrm{X}'\mathrm{X})^{-1} \tag{2.4.28}$$

kısaltmalarıyla (2.3.11) bağıntısına benzer şekilde,

$$Var(\tilde{\beta}_j\,) = \sigma^2((\mathrm{X}'\mathrm{X})^{-1} - G_j) \tag{2.4.29}$$

elde edilir.

Eğer

$$C = Var(\tilde{\beta}_1) - \mathrm{Var}(\tilde{\beta}_2) = \sigma^2(G_2 - G_1)$$

$$= \sigma^2(\mathrm{X}'\mathrm{X})^{-1}({R_2}'A_2^{-1}R_2 - {R_1}'A_1^{-1}R_1)(\mathrm{X}'\mathrm{X})^{-1} \geq 0 \quad (2.4.30)$$

veya eş değer olarak,

$$({R_2}'A_2^{-1}R_2 - {R_1}'A_1^{-1}R_1) \geq 0 \quad (2.4.31)$$

ise, $\tilde{\beta}_2$ kısıtlanmış en küçük kareler tahmin edicisi $\tilde{\beta}_1$ kısıtlanmış en küçük kareler tahmin edicisinden daha iyidir. (Rao, C. R., Toutenburg, H., 1999)

Teorem 2.4.2: (2.4.25) varsayımları altında

$${R_2}'A_2^{-1}R_2 - {R_1}'A_1^{-1}R_1 \geq 0 \quad (2.4.32)$$

eşitsizliğini elde etmek için gerek ve yeter şart

$$R_1 = PR_2 \quad (2.4.33)$$

olacak şekilde bir P, $\mathrm{q}_1\mathrm{x}\mathrm{q}_2$ matrisinin var olmasıdır.

İspat: Teorem 1.1.34 kullanılsın ve $M = {R_2}'A_2^{-1/2}$ ve $N = {R_1}'A_1^{-1/2}$ olarak tanımlansın.

i) (2.4.32) bağıntısının varlığı kabul edilsin ve Teorem 1.1.34 kullanılsın. Bu durumda,

$$N = MH$$

olacak şekilde bir H matrisi vardır. Böylece, ${R_1}'A_1^{-1/2} = {R_2}'A_2^{-1/2}H$ veya eş değer olarak $P = A_1^{1/2}H'A_2^{-1/2}$: $q_1 \times \mathrm{q}_2$ matrisi olmak üzere $R_1 = A_1^{1/2}H'A_2^{-1/2}R_2 = PR_2$ elde edilir.

ii) $R_1 = PR_2$ olduğu kabul edilsin. Bu takdirde (2.4.32) farkı

$${R_2}'A_2^{-1/2}(I - F)A_2^{-1/2}R_2 \quad (2.4.34)$$

olarak yazılabilir. Burada F matrisi

$$F = A_2^{1/2}\mathrm{P}'(\mathrm{PA_2P}')^{-1}\mathrm{P}A_2^{1/2} \tag{2.4.35}$$

ile tanımlanan simetrik ve idempotent bir matristir. Bu nedenle, $(I-F)$'de idempotenttir. $B = R_2{}'A_2^{-1/2}(I-F)$ kısaltmasını kullanılırsa, (2.4.34) farkı $BB' \geq 0$ olmaktadır. (Bakınız Teorem 1.1.4)

Sonuç 2.4.1: Eğer, $rank(R_1) = \mathrm{q}_1$ olmak üzere $R_1 = PR_2$ gerçeklenirse, $\mathrm{q}_1 \leq \mathrm{q}_2$ ve $rank(P) = \mathrm{q}_1$ olması gerekir. Ayrıca, $r_1 = Pr_2$ olur.

İspat: Teorem 1.1.5 **(iv)**'den genel olarak

$$rank(AB) \leq min(rank(A), rank(B)) \tag{2.4.36}$$

olduğu biliniyor. Bu probleme uygulanırsa,

$$rank(PR_2) \leq \min\,(rank(P), rank(R_2)) \tag{2.4.37}$$

$$\mathrm{q}_1 = \min(rank(P), \mathrm{q}_2) = rank(P) \tag{2.4.38}$$

elde edilir.

$$rank(R_1) = rank(PR_2) = \mathrm{q}_1 \tag{2.4.39}$$

olduğunda $\mathrm{q}_1 = rank(P)$ olması $\mathrm{q}_1 \leq \mathrm{q}_2$ olmasını gerektirir.

$r_1 = R_1\beta$ ve $R_1 = PR_2$ eşitliklerinden

$$r_1 = PR_2\beta = Pr_2 \tag{2.4.40}$$

sonucu çıkarılabilir.

Not 2.4.1: $\mathrm{q}_1 = \mathrm{q}_2$ durumu karşılık gelen tahmin edicilerin eşitliğine ilaveten

$$r_1 = R_1\beta \quad \text{ve} \quad r_2 = R_2\beta \tag{2.4.41}$$

kısıtlamaların eşitliğine yol açtığından durum $\mathrm{q}_1 < \mathrm{q}_2$ durumuyla sınırlanabilir. Bu gerçek aşağıdaki gibi görülür:

$rank(P) = \mathrm{q}_1 = \mathrm{q}_2$ olmak üzere $R_1 = PR_2$ bağıntısı P^{-1} in varlığını gerektirir. Bundan dolayı $R_2 = P^{-1}R_1$ ve $r_2 = P^{-1}r_1$ gerçeklenir. Böylece, $r_2 = R_2\beta$, $P^{-1}(r_1 - R_1\beta) = 0$ eşitliğine (yani, $r_1 = R_1\beta$ eşitliğine) denktir. P , $\mathrm{q}_1\mathrm{x}\mathrm{q}_2$

ve $rank(P) = q_1 = q_2$ olmak üzere $R_1 = PR_2$ için tahmin edicilerin denkliği hemen kontrol edilebilir: Böylece,

$$\tilde{\beta}_2 = \hat{\beta} + (X'X)^{-1}{R_1}'P^{-1}\left(P^{-1}R_1(X'X)^{-1}{R_1}'P^{-1}\right)^{-1}(P^{-1}r_1 - P^{-1}R_1\hat{\beta})$$

$$= \tilde{\beta}_1 \tag{2.4.42}$$

dır.

$q_1 < q_2$ durumu: Herhangi bir lineer kısıtlama tekil olmayan bir C, $q_2 x q_2$ matrisiyle çarpmaya göre değişmezdir, yani,

$$r_2 = R_2\beta \text{ ve } Cr_2 = CR_2\beta \tag{2.4.43}$$

şartları denktirler. Bu denklik kullanılsın ve C'nin özel bir seçimi yapılsın. P, q_1 ranklı bir $q_1 x q_2$ matris olmak üzere, $R_1 = PR_2$ olduğu kabul edilsin.

$C' = [Q' \vdots P']$, q_2 ranklı olacak şekilde, $(q_2 - q_1)xq_2$ mertebeli $q_2 - q_1$ ranklı bir Q matrisi seçilsin. (Q matrisine, P matrisinin ortogonal tümleyeni denir.) $Qr_2 = r_3$ ve $QR_2 = R_3$ alınırsa,

$$Cr_2 = \begin{bmatrix} Qr_2 \\ \cdots \\ Pr_2 \end{bmatrix} = \begin{bmatrix} r_3 \\ \cdots \\ r_1 \end{bmatrix} \tag{2.4.44}$$

$$CR_2 = \begin{bmatrix} QR_2 \\ \cdots \\ PR_2 \end{bmatrix} = \begin{bmatrix} R_3 \\ \cdots \\ R_1 \end{bmatrix} \tag{2.4.45}$$

elde edilir. Eğer iki $r_1 = R_1\beta$ ve $r_2 = R_2\beta$ lineer kısıtlaması bir $R_1 = PR_2$ lineer dönüşümü ile bağlıysalar, bu takdirde $r_1 = R_1\beta$'nın $r_2 = R_2\beta$'da tamamıyla ihtiva edildiği durumunu farkına varmak ilginçtir. Bu nedenle genelliği kaybetmeksizin $P = [I \ \vdots 0]$ matrisi seçilebilir.

Sonuç 2.4.2: $Qr_2 = r_3$, $QR_2 = R_3$ ve Q, P'ye tümleyen olmak üzere

$$r_1 = R_1\beta, \quad r_2 = R_2\beta, \quad R_1 = PR_2\,, \quad r_1 = Pr_2, rank(P) = q_1 < q_2 \tag{2.4.46}$$

ve

$$r_1 = R_1\beta, r_2 = \begin{bmatrix} r_1 \\ \cdots \\ r_3 \end{bmatrix} = \begin{bmatrix} R_1 \\ \cdots \\ R_3 \end{bmatrix} \beta = R_2\beta \tag{2.4.47}$$

kısıtlamalar kümesi denktirler. Böylece Teorem 2.4.2'den iki kesin lineer kısıtlamanın onların karşılık gelen kısıtlanmış en küçük kareler tahmin edicileriyle karşılaştırılabilir olmaları için gerek ve yeter şartın $R_1 = PR_2$ ve $rank(P) = q_1 < q_2$ olması gerektiği sonucu çıkarılabilir. $P = [I \ \vdots \ 0]$ özel durumu

$$\mathrm{R}_2 = \begin{bmatrix} R_1 \\ \cdots \\ R_3 \end{bmatrix}, \quad r_2 = \begin{bmatrix} r_1 \\ \cdots \\ r_3 \end{bmatrix} \tag{2.4.48}$$

iç içe veya lineer bağımlılık durumunu ifade eder. (Rao, C. R., Toutenburg, H., 1999)

2.5 Stokastik Lineer Eşitlik Kısıtlaması

$R\beta = r$ kısıtlaması tam olmadığında ve bir ϑ hatası içerdiğinde, yani; ψ pozitif kararlı, tekil olmayan bilinen bir matris olmak üzere,

$$r = R\beta + \vartheta, \ \vartheta \sim (0, \psi), \tag{2.5.1}$$

olduğunda, bu kısıtlamayı ϑ ve ε vektörlerinin bağımsız olduğu varsayımı altında

$$\begin{bmatrix} Y \\ \cdots \\ r \end{bmatrix} = \begin{bmatrix} X \\ \cdots \\ R \end{bmatrix} \beta + \begin{bmatrix} \varepsilon \\ \cdots \\ \vartheta \end{bmatrix}, \tag{2.5.2}$$

$$Kov\left(\begin{bmatrix} \varepsilon \\ \cdots \\ \vartheta \end{bmatrix}\right) = E\left(\begin{bmatrix} \varepsilon \\ \cdots \\ \vartheta \end{bmatrix} [\varepsilon \ \vdots \ \vartheta]'\right) = \begin{bmatrix} \sigma^2 I & 0 \\ 0 & \psi \end{bmatrix} = \Omega \tag{2.5.3}$$

biçiminde modele eklensin. Burada Ω tekil değildir ve $\Omega^{-1} = \begin{bmatrix} \frac{1}{\sigma^2} I & 0 \\ 0 & \psi^{-1} \end{bmatrix}$ dir.

Şimdi $w = \begin{bmatrix} X \\ \cdots \\ R \end{bmatrix}$ alınsın.

Bu takdirde Gauss-Markov Teoreminden

$$\tilde{\beta} = (w'\Omega^{-1}w)^{-1}w'\Omega^{-1}\begin{bmatrix} Y \\ \dots \\ r \end{bmatrix}$$

$$= \left([X' \quad \vdots \quad R'] \begin{bmatrix} \frac{1}{\sigma^2}I & 0 \\ 0 & \psi^{-1} \end{bmatrix} \begin{bmatrix} X \\ \dots \\ R \end{bmatrix} \right)^{-1} [X' \quad \vdots \quad R'] \begin{bmatrix} \frac{1}{\sigma^2}I & 0 \\ 0 & \psi^{-1} \end{bmatrix} \begin{bmatrix} Y \\ \dots \\ r \end{bmatrix}$$

$$= \left(\left[\frac{1}{\sigma^2}X' \quad \vdots \quad R'\psi^{-1}\right] \begin{bmatrix} X \\ \dots \\ R \end{bmatrix} \right)^{-1} \left[\frac{1}{\sigma^2}X' \quad \vdots \quad R'\psi^{-1}\right] \begin{bmatrix} Y \\ \dots \\ r \end{bmatrix}$$

$$= (\tfrac{1}{\sigma^2}X'X + R'\psi^{-1}R)^{-1}(\tfrac{1}{\sigma^2}X'Y + R'\psi^{-1}r) \tag{2.5.4}$$

elde edilir. Bu tahmin edici için, $\tilde{\beta}$ yansız tahmin edicidir. Gerçekten

$$E(\tilde{\beta}) = \mathrm{E}\left(\left(\frac{1}{\sigma^2}X'X + R'\psi^{-1}R\right)^{-1}\left(\frac{1}{\sigma^2}X'Y + R'\psi^{-1}r\right)\right)$$

$$= \left(\frac{1}{\sigma^2}X'X + R'\psi^{-1}R\right)^{-1}\left(\frac{1}{\sigma^2}X'X\beta + R'\psi^{-1}R\beta\right)$$

$$= \left(\frac{1}{\sigma^2}X'X + R'\psi^{-1}R\right)^{-1}\left(\frac{1}{\sigma^2}X'X + R'\psi^{-1}R\right)\beta = \beta \tag{2.5.5}$$

dır.

$$Kov(\tilde{\beta}) = \left(\frac{1}{\sigma^2}X'X + R'\psi^{-1}R\right)^{-1} \tag{2.5.6}$$

olur. **(**Pore, M.D.,1969)

2.6 İndirgenmiş Model

$Y = X\beta + \varepsilon,\ \ \varepsilon \sim N(0, \sigma^2 I_n)$ lineer modelinde $Y{:}\,nx1$ gözlenebilir rasgele vektör, $X{:}\,nxp$ tam sütun ranklı reel sayılar matrisi, $\beta{:}\,px1$ bilinmeyen parametrelerin bir vektörü ve ε gözlenemeyen $N(0, \sigma^2 I_n)$ dağılımlı varsayılan bir rasgele hata vektördür.

Bu modelin parametre kümesi

$$\Omega = \{(\beta, \sigma^2) : \beta \in \mathrm{E}^{\mathrm{p}}, \sigma^2 > 0\} \tag{2.6.1}$$

dır. Bu modelde $H\beta = r$ kısıtlaması altında β'nın tahminini bulmak için $H\beta = r$ kısıtlamasını $Y = X\beta + \varepsilon$ modelinde yerine yazarak kısıtlamalı indirgenmiş model elde edilir. $HQ = 0$ (burada, Q', H'ın ortogonal tümleyeni) olacak şekilde, rankı $p - q$ olan $(p-q)xp$ mertebeli Q' matrisi ele alınsın. $rank(Q') = p - q$ tam satır ranklıdır. H matrisinin satır vektörlerine Q' matrisinin satır vektörlerinin eklenmesiyle elde edilen pxp mertebeli matris için,

$$rank\begin{bmatrix} H_{qxp} \\ \cdots\cdots \\ Q'_{\,p-qxp} \end{bmatrix} = p \tag{2.6.2}$$

tam satır ranklıdır, bu nedenle

$$\begin{bmatrix} H_{qxp} \\ \cdots\cdots \\ Q'_{\,p-qxp} \end{bmatrix}^{-1} \tag{2.6.3}$$

vardır. Gerçekten,

$$\begin{aligned}
\begin{bmatrix} H_{qxp} \\ \cdots\cdots \\ Q'_{\,p-qxp} \end{bmatrix}^{-1} &= [H' \;\vdots\; Q]\left(\begin{bmatrix} H \\ \cdots \\ Q' \end{bmatrix}[H' \;\vdots\; Q]\right)^{-1} \\
&= [H' \;\vdots\; Q]\begin{bmatrix} HH' & HQ \\ Q'H' & Q'Q \end{bmatrix}^{-1} \\
&= [H' \;\vdots\; Q]\begin{bmatrix} HH' & 0 \\ 0 & Q'Q \end{bmatrix}^{-1} \\
&= [H' \;\vdots\; Q]\begin{bmatrix} (HH')^{-1} & 0 \\ 0 & (Q'Q)^{-1} \end{bmatrix} \\
&= [H'(HH')^{-1} \quad Q(Q'Q)^{-1}] \\
&= [H^{+} \;\vdots\; (Q')^{+}]
\end{aligned}$$

dır. Buna göre,

$$\begin{bmatrix} H_{qxp} \\ \cdots\cdots \\ Q'_{\,p-qxp} \end{bmatrix}^{-1}\begin{bmatrix} H_{qxp} \\ \cdots\cdots \\ Q'_{\,p-qxp} \end{bmatrix} = [H^{+} \;\vdots\; (Q')^{+}]\begin{bmatrix} H \\ \cdots \\ Q' \end{bmatrix}$$

$$= H^{+}H + (Q')^{+}Q' = I_p \qquad (2.6.4)$$

olur. Buradan $Y = X\beta + \varepsilon,\ \varepsilon \sim N(0, I_n)$ lineer modeli,

$$Y = X(H^{+}H + (Q')^{+}Q')\beta + \varepsilon$$

$$= XH^{+}\underbrace{H\beta}_{r} + X(Q')^{+}Q'\beta + \varepsilon$$

$$= XH^{+}r + X(Q')^{+}Q'\beta + \varepsilon$$

$$Y - XH^{+}r = X(Q')^{+}Q'\beta + \varepsilon \qquad (2.6.5)$$

biçiminde yazılarak,

$$Z = Y - XH^{+}r \qquad (2.6.6)$$

$$B = X(Q')^{+} \qquad (2.6.7)$$

$$\gamma = Q'\beta \qquad (2.6.8)$$

değişken değiştirmesi yapılırsa,

$$Z = B\gamma + \varepsilon \qquad (2.6.9)$$

modeli elde edilir. Bu modele, verilen kısıtlama altında indirgenmiş model denir. Bu modelde $Z: nx1$ gözlenebilir rasgele vektör, $B: nx(p-q)$ sabitlerin tam sütun ranklı bir matrisi, $\gamma: (p-q)x1$ bilinmeyen parametrelerin bir vektörü ve ε gözlenemeyen $N(0, \sigma^2 I_n)$ dağılımlı, bir rasgele hata vektördür. Bu modelin parametre kümesi

$$\Omega^{*} = \{(\gamma, \sigma^2) : \gamma \in \mathrm{R}^{p-q}, \sigma^2 > 0\} \qquad (2.6.10)$$

dır. Ω parametre kümesinde β ve σ^2 nin sırasıyla $\hat{\beta}_{\Omega}$ ve $\hat{\sigma}^2_{\Omega}$ ile gösterilen en çok olabilirlik tahmin edicileri

$$\hat{\beta}_{\Omega} = (X'X)^{-1}X'Y \qquad (2.6.11)$$

$$\hat{\sigma}^2_{\Omega} = \frac{1}{n}Y'(I_n - X(X'X)^{-1}X')Y \qquad (2.6.12)$$

olarak elde edilmiştir.

(2.6.9) modelinde γ ve σ^2 nin tahmin edicileri de, (2.6.11) ve (2.6.12) bağıntılarında Y yerine Z ve X yerine B yazılırsa, aşağıdaki gibi elde edilir.

$$\gamma_{\Omega^*} = (\mathrm{B}'\mathrm{B})^{-1}\mathrm{B}'\mathrm{Z}, \tag{2.6.13}$$

$$\hat{\sigma}^2_{\Omega^*} = \frac{1}{\mathrm{n}} Z'(I - (\mathrm{B}'\mathrm{B})^{-1}\mathrm{B}')\mathrm{Z}. \tag{2.6.14}$$

2.7. Hipotez Testi

(2.1.1) modelinde $rank(X) = p \leq n,\ \varepsilon \sim N(0, \sigma^2 I_n)$ ve parametre kümesi

$$\Omega = \{(\beta, \sigma^2): \beta \in \mathrm{E}^{\mathrm{p}}, \sigma^2 > 0\} \tag{2.7.1}$$

olsun.

$$H_0: H\beta = r \text{ veya } H\beta - r = 0$$

$$H_a: H\beta \neq r \text{ veya } H\beta - r \neq 0 \tag{2.7.2}$$

hipotezleri test edilsin. Burada $H,\ qxp$ boyutlu q ranklı bir matris ve $r,\ qx1$ boyutlu bir vektördür. $H^+,\ H$'ın Moore-Penrose g-tersidir ve H tam satır ranklı olduğundan $H^+ = H'(HH')^{-1}$ dir. Bu durumda $Y \sim N(X\beta, \sigma^2 I_n)$ ve $\hat{\beta} \sim N(\beta, \sigma^2 (X'X)^{-1})$ oldukları verilmiştir. $(H\hat{\beta} - r) \sim N(H\beta - r, H(X'X)^{-1}H'\sigma^2)$ dir.

$$Q = (H\hat{\beta} - r)'[H(X'X)^{-1}H']^{-1}(H\hat{\beta} - r) \tag{2.7.3}$$

$(H\hat{\beta} - r)$ nin bir karesel formudur ve $\frac{Q}{\sigma^2} \sim \chi^2_{q,\lambda}$ dağılımına sahiptir. $q = rank(H)$ olduğundan ve bu karesel formun dağılımının merkezi olmama parametresinin $\lambda = \frac{(H\beta - r)'\left[H(X'X)^{-1}H'\right]^{-1}(H\beta - r)}{2\sigma^2}$ olduğu gösterilebilir. Ayrıca, $\frac{AKT}{\sigma^2} \sim \chi^2{}_{\mathrm{n-p}}$ dağılımına sahiptir. Q ve AKT karesel formları bağımsızdır. Bu nedenle,

$$F_{q,n-p,\lambda} = \frac{Q/q}{AKT/(n-p)}$$

$$= \frac{(Y - XH^+r)'X(X'X)^{-1}H'\left[H(X'X)^{-1}H'\right]^{-1}H(X'X)^{-1}X(Y - XH^+r)/q}{(Y - XH^+r)'(I_n - X(X'X)^{-1}X')(Y - XH^+r)/(n-p)} \tag{2.7.4}$$

oranı, payı q ve paydası $n-p$ serbestlik dereceli ve λ merkezi olmama parametreli bir F dağılımına sahiptir. Bu oran $H\beta = r$ hipotezinin test edilmesi için kullanılır. H_0 hipotezi doğru olduğunda,

$$F_{q,n-p} = \frac{(\hat{\beta}-\beta)H'\left[H(X'X)^{-1}H'\right]^{-1}H(\hat{\beta}-\beta)/q}{Y'(I_n-X(X'X)^{-1}X')Y/(n-p)} \tag{2.7.5}$$

oranı, payı q ve paydası $n-p$ serbestlik dereceli, merkezi F dağılımına sahiptir. α anlam düzeyinde $F_{1-\alpha,q,n-p}$ tablo değeri hesaplanır. Bu değer (2.7.5) test istatistiğinin hesaplanan değerinden büyük olduğunda H_0 hipotezi kabul edilir.

Aynı zamanda, H^+ için, Moore Penrose'un ilk üç şartını sağlayan

$$H^{g_3} = X(X'X)^{-1}H'[H(X'X)^{-1}H']^{-1}H(X'X)^{-1}X \tag{2.7.6}$$

değeri alınırsa ve (2.7.5) oranının paydası

$$P = (Y - XH^{g_3}r)'[I - (X'X)^{-1}H'[H(X'X)^{-1}H']^{-1}H(X'X)^{-1}X](Y - XH^{g_3}r) \tag{2.7.7}$$

olur. Q ile P bağımsız ve $\frac{Q}{\sigma^2}$ ve $\frac{P}{\sigma^2}$ iki bağımsız ki-kare dağılımına sahip olduğundan, P de, H_0: $H\beta = r$ hipotezini test etmede kullanılabilir.

$$F_{q,n-p,\lambda} = \frac{(Y-XH^{g_3}r)'X(X'X)^{-1}H'\left[H(X'X)^{-1}H'\right]^{-1}H(X'X)^{-1}X(Y-XH^{g_3}r)/q}{(Y-XH^{g_3}r)'[I_p-(X'X)^{-1}H'[H(X'X)^{-1}H']^{-1}H(X'X)^{-1}X](Y-XH^{g_3}r)/(n-p)} \tag{2.7.8}$$

$rank(X) = k < p$ olsun. $H\beta$ nın tahmin edilebilir olması için $H(X'X)^{-}(X'X) = H$ olmalı ve $H\beta = r$ denkleminin tutarlı olması için de $HH^{+}r = r$ olmalıdır. $H\beta = r$ hipotezinin test edilmesi için,

$$\begin{aligned} F_{q,n-p,\lambda} &= \frac{(H\hat{\beta}-r)'[H(X'X)^{-}H']^{-1}(H\hat{\beta}-r)}{Y'(I-X(X'X)^{-}X')Y} \\ &= \frac{(Y-XH^{+}r)'X(X'X)^{-}H'[H(X'X)^{-}H']^{-1}H(X'X)^{-}X(Y-XH^{+}r)}{(Y-XH^{+}r)'(I-XX^{-})(Y-XH^{+}r)} \end{aligned} \tag{2.7.9}$$

oranı H_0: $H\beta = r$ nin testinde test istatistiği olarak kullanılır. H_0 hipotezi kabul edilirse, $\tilde{\beta} = \hat{\beta}$ olur. H_0 hipotezi reddedilirse,

$$\tilde{\beta} = \hat{\beta} - H^{+}\left(H\hat{\beta} - r\right)$$

olur..

H_0 hipotezinin bazı özel hallerini göz önüne alalım:

$$H_0 : \beta = 0 \tag{2.7.10}$$

Burada, $H = I_p$ve $r = 0$ olmak üzere,

$$F_{p,n-p} = \frac{\hat{\beta}'(X'X)\hat{\beta}/p}{Y\prime(I_n - X(X'X)^{-1}X'\,)\mathrm{Y}/n-p} \tag{2.7.11}$$

dir.

$$H_0 : \beta = b \tag{2.7.12}$$

b, $px1$ bilinen bir vektör ve $H = I_p$ ve $r = b$ olmak üzere,

$$F_{p,n-p} = \frac{(\hat{\beta}-b)'(X'X)(\hat{\beta}-b)/p}{Y\prime(I - X(X'X)^{-1}X'\,)\mathrm{Y}/(n-p)} \tag{2.7.13}$$

olur. (2.7.13) oranı payı p ve paydası $n - p$ serbestlik dereceli F dağılımına sahiptir.

2.8 Parametreler İçin Aralık Tahmini

$\hat{\beta}$~N$(\beta, \sigma^2 c_{ii}\,)$ dir. Verilen model altında β_j,$j = 0,1,2, \dots, k,$ katsayıları için güven aralığı oluşturulsun. $\hat{\beta}$ tahmin edicisinin varyans-kovaryans matrisini,

$$(X'X)^{-1}\sigma^2 = \begin{bmatrix} c_{00} & \cdots & c_{0k} \\ \vdots & \ddots & \vdots \\ c_{k0} & \cdots & c_{kk} \end{bmatrix} \sigma^2 \tag{2.8.1}$$

biçiminde yazılsın. $\{\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_k\}$ ların varyansları köşegen üzerindeki negatif olmayan değerlerdir. $\hat{\beta}_j$'nın varyansı $c_{jj}\sigma^2$ olarak gösterilebilir. $\hat{\beta}_{\mathrm{j}}$~N$(\beta_j, \sigma^2 c_{jj}\,)$ dir. $c_{jj} \neq 0$ değerleri için

$$\frac{(\hat{\beta}_{\mathrm{j}} - \beta_j)}{\sigma\sqrt{c_{jj}}} \sim \mathrm{N}(0,1) \tag{2.8.2}$$

dir. Diğer yandan $\frac{AKT}{\sigma^2}$ ve $\hat{\beta}$ bağımsızdır. $j = 0,1,2,\dots,k$ için $\hat{\beta}$ vektörünün bileşenleri için, (2.8.2) de σ^2 yerine onun $\hat{\sigma}^2$ tahminini kullanarak $n-p$ serbestlik dereceli,

$$t = \frac{(\hat{\beta}_j - \beta_j)/\sigma\sqrt{c_{jj}}}{\sqrt{\frac{(n-p)\hat{\sigma}^2}{\sigma^2}/(n-p)}} = \frac{(\hat{\beta}_j - \beta_j)}{\hat{\sigma}\sqrt{c_{jj}}} \sim t_{(n-p)} \qquad (2.8.3)$$

rasgele değişkeni elde edilir. Bu durumda

$$P\left[-t_{1-\frac{\alpha}{2}} \le \frac{(\hat{\beta}_j - \beta_j)}{\hat{\sigma}\sqrt{c_{jj}}} \le t_{1-\frac{\alpha}{2}}\right] = 1-\alpha \qquad (2.8.4)$$

yazılabilir. Gerekli işlemler yapılırsa,

$$P\left[\hat{\beta}_j - t_{n-p,1-\frac{\alpha}{2}}\hat{\sigma}\sqrt{c_{jj}} \le \beta_j \le \hat{\beta}_j + t_{n-p,1-\frac{\alpha}{2}}\hat{\sigma}\sqrt{c_{jj}}\right] = 1-\alpha \qquad (2.8.5)$$

elde edilir. Burada, $\hat{\sigma} = s = \frac{AKT}{n-p}$ dir.

2.9.Lineer Parametrik Fonksiyonlar İçin Güven Aralığı

Verilen (2.1.1) lineer modeli ve 1. Duruma göre, $\{\beta_0, \beta_1, \dots, \beta_k\}$ parametrelerinin bir lineer fonksiyonu $\lambda'\beta$ olsun. $\lambda'\beta$ nın en iyi lineer yansız tahmin edicisi $\lambda'\hat{\beta}$ dır. Burada $\hat{\beta} = (X'X)^{-1}X'$ Y, β nın en küçük kareler tahmin edicisidir. $\lambda'\beta$ için $1-\alpha$ güven katsayısılı güven aralığı oluşturulsun. $Y \sim N(X\beta, \sigma^2 I_n)$ varsayımı altında $\lambda'\hat{\beta} = \lambda'(X'X)^{-1}X'$ Y olduğundan $\lambda'\hat{\beta}$ bağımsız ve normal dağılıma sahip $Y_1, Y_2 \dots, Y_k$ rasgele değişkenlerinin bir lineer fonksiyonudur.

$$E(\lambda'\hat{\beta}) = \lambda'\beta \qquad (2.9.1)$$

ve

$$\frac{\lambda'\hat{\beta} - \lambda'\beta}{\sigma\sqrt{\lambda'(X'X)^{-1}\lambda}} \sim N(0,1) \qquad (2.9.2)$$

dir. $\frac{AKT}{\sigma^2}$ ve $\lambda'\hat{\beta}$ bağımsızdır.

$$\frac{\frac{\lambda'\hat{\beta}-\lambda'\beta}{\sigma\sqrt{\lambda'(X'X)^{-1}\lambda}}}{\sqrt{\frac{(n-p)\hat{\sigma}^2}{\sigma^2}/(n-p)}} = \frac{\lambda'\hat{\beta}-\lambda'\beta}{s\sqrt{\lambda'(X'X)^{-1}\lambda}} \sim t_{(n-p)} \tag{2.9.3}$$

olur. Gerekli işlemler yapılırsa,

$$P\left[-t_{1-\frac{\alpha}{2}} \leq \frac{\lambda'\hat{\beta}-\lambda'\beta}{S\sqrt{\lambda'(X'X)^{-1}\lambda}} \leq t_{1-\frac{\alpha}{2}}\right] = 1-\alpha \tag{2.9.4}$$

olur. Böylece $\lambda'\beta$ için $1-\alpha$ güven katsayılı güven aralığı,

$$P\left[\lambda'\hat{\beta} - t_{n-p,1-\frac{\alpha}{2}} s\sqrt{\lambda'(X'X)^{-1}\lambda} \leq \lambda'\beta \leq \lambda'\hat{\beta} + t_{n-p,1-\frac{\alpha}{2}} s\sqrt{\lambda'(X'X)^{-1}\lambda}\right] = 1-\alpha \tag{2.9.5}$$

veya

$$\lambda'\hat{\beta} \pm t_{n-p,1-\frac{\alpha}{2}} s\sqrt{\mathrm{Var}(\lambda'\hat{\beta})} \tag{2.9.6}$$

şeklinde yazılabilir. Burada $t_{n-p,1-\frac{\alpha}{2}}$ t −dağılım tablosundan bulunan değerdir. Ve $\hat{\sigma}^2 = s^2$, σ^2 nin bu durumdaki yansız tahminidir. Aynı şartlar altında, X in eksik ranklı olması durumunda varyansları sıfırdan farklı olan yani $(X'X)^-$ matrisini sıfırdan farklı köşegen elemanları, c_{jj} ler, bazı bilinmeyen parametreler için lineer parametrik fonksiyonları seçilerek güven aralıkları kurulabilir.

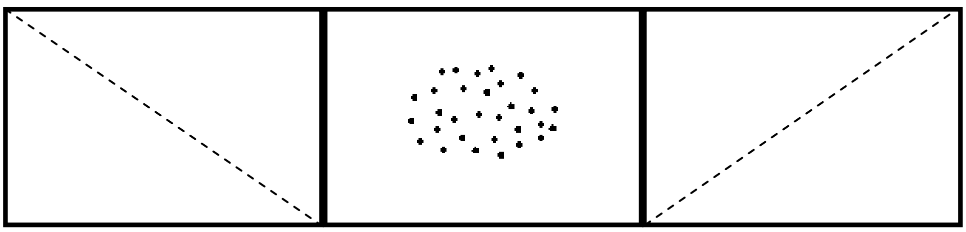
BÖLÜM 3

KANONİK KORELASYON ANALİZİ

3.1. İlişki(Korelasyon) Ölçüleri

Bazı bilimsel araştırmalarda iki dizi arasında belirli bir ilişki sezinlenir ve bu ilişkinin somutlaştırılması, ölçülmesi ve ortaya koyduğu gerçeğe göre yorumlar getirilmesi gerekmektedir.

İki değişken normal dağılım gösterdiğinde arasındaki ilişkinin derecesini, yönünü ve önemini göstermek amacıyla en yaygın kullanılan katsayı Pearson korelasyon katsayısıdır. Korelasyon, iki değişken arasındaki lineerlik derecesinin bir ölçüsüdür. Pearson korelasyon katsayısı $-1 \leq \rho \leq +1$ değerleri arasında değişim göstermektedir. ρ'nun -1 olması, X ile Y değişkenleri arasında negatif tam bir doğrusal ilişki, $+1$ olması durumu ise X ve Y değişkenleri arasında pozitif tam bir doğrusal ilişki olduğunu göstermektedir. ρ değerinin 0 olması durumu ise, iki değişken arasında ilişkinin olmadığını göstermektedir. ρ'nun pozitif değerleri X artarken Y'nin de artacağına, ρ negatif değerleri ise X artarken Y'nin de azalacağını ifade etmektedir.



(a) $\rho = -1$ Mükemmel Negatif İlişki

(b) $\rho = 0$ İlişki Yok

(c) $\rho = +1$ Mükemmel Pozitif İlişki

Şekil 3.1.1: X ve Y değişkenleri arasındaki ilişki

Pearson Korelasyon Katsayısı;

$$\rho = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}} \tag{3.1.1}$$

olarak tanımlanır. Korelasyona ilişkin varsayımlar;

1) X'in her değeri için Y değerlerinin normal dağılan bir alt kümesi vardır.
2) Y'nin her değeri için X değerlerinin normal dağılan bir alt kümesi vardır.
3) X ve Y'nin bileşik dağılımı iki değişkenli normal dağılım gösterir.(İki değişkenin bileşik dağılımı normal dağılım gösteriyorsa, bu dağılıma iki değişkenli normal dağılım denir.)
4) Y değerlerinin alt kümeleri eşit varyansa sahiptir.
5) X değerlerinin alt kümeleri eşit varyansa sahiptir.(Alpar,R., 2003)

Değişken sayısı ikiden fazla ve normal dağılım gösteriyorsa aralarındaki ilişki, kısmi korelasyon katsayısı ile bulunmaktadır. Değişken sayısı üç ise, Z rasgele değişkeni sabit tutulduğunda, X ve Y rasgele değişkenleri arasındaki kısmi korelasyon katsayısı

$$r_{XY.Z} = \frac{r_{XY} - (r_{XZ} r_{YZ})}{\sqrt{(1 - r_{XZ}^2)(1 - r_{YZ}^2)}} \tag{3.1.2}$$

dir. Burada $r_{XY.Z}$, Z değişkeni sabit tutulduğunda, X ve Y değişkenleri arasındaki kısmi korelasyon katsayısını göstermektedir. Değişken sayısı dört ise, 3. ve 4. değişkenler sabit tutularak 1. ve 2. değişkenler arasındaki kısmi korelasyon katsayısı,

$$r_{XY.ZT} = \frac{r_{XY.Z} - (r_{XT.Z} . r_{YT.Z})}{\sqrt{(1 - r_{XT.Z}^2)(1 - r_{YT.Z}^2)}} \tag{3.1.3}$$

dir. Değişken sayısı beş ise, 1.ve 2. değişkenlere diğerlerini sabit tutarak bakmak gerekirse;

$$r_{XY.ZTP} = \frac{r_{XY.ZT} - (r_{XP.ZT} . r_{YP.ZT})}{\sqrt{(1 - r_{XP.ZT}^2)(1 - r_{YP.ZT}^2)}} \tag{3.1.4}$$

formülünü kullanırız. Örneğin; Belli bir yaş aralığı için, yaş-zeka düzeyi arasındaki korelasyon $r_{XY} = 0{,}74$; yaş-kafa çevresi arasındaki korelasyon $r_{XZ} = 0{,}68$; kafa çevresi- zeka düzeyi arasındaki korelasyonda $r_{YZ} = 0{,}62$; yaş etkeni ortadan kaldırıldığında gerçek zeka düzeyi- kafa yüzeyi çevresi arasındaki korelasyon;

$$r_{XY.Z} = \frac{0{,}62-(0{,}74*0{,}68)}{\sqrt{(1-0{,}74^2)(1-0{,}68^2)}} = 0{,}234$$

şeklinde bulunmuştur. (Şenocak,M., 1990)

Verilerin sınıflayıcı ölçekle ölçülmüş olması durumunda kullanılan ya da normallik koşulunu sağlamaması durumunda kullanılan Spearman ya da Kendall-Tau türü ilişki katsayıları Pearson ilişki katsayısına alternatif olmaktadır.

Sperman'ın sıra korelasyon katsayısı,

$$r = 1 - \frac{6\sum d_i^2}{n(n^2-1)} \tag{3.1.5}$$

dir.

Burada $d_i = x_i - y_i$ farkından hesaplanmakta olup, X ve Y değişkenleri içerisinde yer alan gözlem değerlerinin almış olduğu sıra puanları arasındaki farkı ve n toplam gözlem sayısını ifade etmektedir. Sperman korelasyon katsayısı Pearson korelasyon katsayısı ile şu şekilde ispatlanabilir:

İki karakteristik A ve B' ye karşılık gelen sıraları gözönüne alalım. x_i, y_i sırasıyla A ve B de i-yinci gözlemin sırası olsun. Bu takdirde x ve y arasındaki korelasyon katsayısına gözlem grubu A ve B karakteristiklerin de sıra korelasyon katsayısı denir. Değişkenlerin her biri $1,2,\ldots,n$ değerini alsın ve bu nedenle

$$\bar{x} = \frac{1+2+3+\cdots+n}{n} = \frac{n(n+1)}{2n} = \frac{n+1}{2} = \bar{y} \tag{3.1.6}$$

Bir kural olarak x_i, y_i'ye eşit değildir.

$$d_i = x_i - y_i \tag{3.1.7}$$

olacak şekilde farkı göstersin. Eğer, x' ve y' değişkenlerin ortalamalardan sapmalarını gösteriyorsa,

$$d_i = x_i' - y_i' \tag{3.1.8}$$

de elde edilir. Değişkenler arasındaki korelasyon katsayısı

$$r = \frac{\sum x_i' y_i'}{\sqrt{\sum x_i'^2 \sum y_i'^2}} \tag{3.1.9}$$

ile verilir. $x_i' = x_i - \bar{x}$, $y_i' = y_i - \bar{y}$ dir.

$$\sum x_i'^2 = \sum y_i'^2 = \sum x_i^2 - 2x_i\bar{x} + \bar{x}^2$$

$$= \sum x_i^2 - n\,\bar{x}^2$$

$$= \frac{n.(n+1).2n+1)}{6} - n.(\frac{n+1}{2})^2$$

$$= \frac{n.(n^2-1)}{12} \tag{3.1.10}$$

$$\sum d_i^2 = \sum(x_i' - y_i')^2 = \sum x_i'^2 + \sum y_i'^2 - 2\sum x_i' y_i' \tag{3.1.11}$$

ve böylece

$$\sum x_i' y_i' = \frac{1}{12}(n^3 - n) - \frac{1}{2}\sum d_i^2 \tag{3.1.12}$$

elde edilen veriler değişkenler arasındaki korelasyon katsayısı denklemi (3.1.9) bağıntısında yerlerine konulursa, Sperman korelasyon katsayısı,

$$r = 1 - \frac{6\sum d_i^2}{n(n^2 - 1)} \tag{3.1.13}$$

şeklinde elde edilir. (Weatherburn,C.,E., 1968) Örneğin; Matematik ve Fizikte aynı 16 öğrencinin notları Tablo 1'deki gibi olsun.

Çizelge 3.1.1: Matematik ve Fizikte Aynı 16 Öğrencinin Notları

Öğrenci Sayısı	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Matematik Notu	96	90	85	80	75	70	65	60	58	56	54	48	45	40	35	30
Fizik Notu	90	56	85	80	75	65	90	70	60	54	35	58	40	48	30	45

Buradan yine aynı 16 öğrencinin Matematikte ve Fizikte sıraları sırasıyla:(1,1), (2,10), (3,3), (4,4), (5,5), (6,7), (7,2), (8,6), (9,8), (10,11), (11,15,) (12,9), (13,14), (14,12), (15,16), (16,13) şeklinde olur. Buradaki d_i, $i - yinci$ çiftler arasındaki farktır.

$\sum_{i=1}^{n} d_i^2 = 136$ $\qquad$ $n^3 - n = n(n^2 - 1) = 16.255$

Sonuç olarak Sperman'ın sıra korelasyon katsayısı

$$r = 1 - \frac{6\sum d_i^2}{n(n^2-1)}$$

$$r = 1 - \frac{6.136}{16.255} = 1 - \frac{1}{5} = 0,8$$

dir.

Değişkenler normal dağılım göstermiyorsa, değişkenler arasındaki ilişki Kendall'ın τ sıra korelasyon katsayısı

$$\hat{\tau} = \frac{2S}{n(n-1)} \tag{3.1.14}$$

ile hesaplanır. Burada $S = P - Q$ dur. Burada, X değişkenine ait gözlem değerleri kendi içerisinde küçükten büyüğe sağa doğru sıralandıktan sonra buna karşılık gelen Y değişkeni içerisindeki gözlem değerleri için Y_i'nin sağındaki kendinden büyük Y_j'lerin sayısı P, kendinden küçük Y_j'lerin sayısına da Q adı verilmektedir ve n toplam gözlem sayısını ifade etmektedir. Örneğin;

Çizelge 3.1.2: 10 domuz için sütten kesilme ağırlığı (1 libre=0,45359237 kg) ve kesilmesine kadar geçecek süre

Domuz sayısı	1	2	3	4	5	6	7	8	9	10
Domuzun sütten kesildiğindeki ağırlığı(libre)	59	56	46	50	48	41	49	57	52	39
Domuzun kesilmesine kadar geçecek süre	105	114	121	117	115	147	119	106	111	142

Çizelge 3.1.3: Çizelge 3.1.2'deki gözlemler için Kendall'ın τ sıra korelasyon katsayısının hesaplanması

X	39	41	46	48	49	50	52	56	57	59
Y	142	147	121	115	119	117	111	114	106	105
P	1	0	0	2	0	0	1	0	0	
Q	8	8	7	4	5	4	2	2	1	
$P - Q$	-7	-8	-7	-2	-5	-4	-1	-2	-1	

$S = \sum(P - Q) = -37 \qquad\qquad n = 10$

$\hat{\tau} = \frac{2S}{n(n-1)} = \frac{-2.37}{10.9} = -\frac{74}{90} = -0{,}822$

bulunmaktadır. (Campbell,R.,C., 1974)

Değişkenler normal dağılım göstermiyorsa ikiden fazla değişken arasındaki kısmi ilişki miktarını belirlemek için Kendall'ın τ kısmi korelasyon katsayısı aşağıdaki gibi hesaplanır.

$$\hat{\tau}_{xy.z} = \frac{\hat{\tau}_{xy} - \hat{\tau}_{xz}\hat{\tau}_{yz}}{\sqrt{(1-\hat{\tau}_{xz}^2)(1-\hat{\tau}_{yz}^2)}} \qquad (3.1.15)$$

Değişken sayısının $(y, x_1, x_2, ..., x_p)$ biçimde p+1 tane olması durumunda değişkenlerden biri ile geriye kalan p değişken arasındaki ilişki aranacak olursa bulunacak ilişkiye çoklu ilişki denir. $x' = (y, x_1, x_2, \dots, x_p)$ rasgele vektör değişkeni, ortalaması µ, varyans-kovaryans matrisi Σ olan $p + 1$ değişkenli normal dağılıma sahip olsun. Bu vektörü $x_1 = y$ ve $x_2 = (x_1, x_2, \dots, x_p)$ biçiminde alt vektörlerine ayrılsın. µ vektörü ve Σ varyans-kovaryans matrisi de parçalara ayrılırsa, y ile $x_2{}'$ arasındaki regresyon bağıntısı aşağıdaki gibi yazılabilir.

$$y = x_2'\beta + e \qquad (3.1.16)$$

Burada, y ile x_2 arasındaki korelasyon katsayısı,

$$\rho_{y,x_2} = \frac{\mathrm{Kov}(y, x_2'\beta)}{\sqrt{\mathrm{Var}(y)\mathrm{Var}(x_2'\beta)}} \qquad (3.1.17)$$

ile bulunur. Buradaki $Kov(y, x_2'\beta) = \sum_{12}\beta$, $\mathrm{Var}(y) = \sigma_{11}$ ve $Var(x_2'\beta) = \beta'\sum_{22}\beta$'dır. Bu nedenle,

$$\rho_{y,x_2} = \frac{\Sigma_{12}\beta}{\sqrt{\sigma_{11}\beta'\Sigma_{22}\beta}} \qquad (3.1.18)$$

dır.

β' regresyon katsayıları matrisi $\sum_{12}\sum_{22}^{-1}$ olduğundan,

$$\rho_{y,x_2} = \frac{\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}}{\sqrt{\sigma_{11}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{22}\Sigma_{22}^{-1}\Sigma_{21}}} = \frac{\sqrt{\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}}}{\sqrt{\sigma_{11}}} \qquad (3.1.19)$$

yazılabilir. Elde edilen değer, y değişkeni ile x_2 değişken vektörünün $x_2'\beta$ biçimindeki bir doğrusal bağıntısı(regresyon) arasındaki çoklu korelasyon katsayısıdır. Çoklu korelasyon katsayısının karesi ise,

$$\rho^2{}_{y,x_2} = \frac{\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}}{\sigma_{11}} \tag{3.1.20}$$

çoklu belirtme katsayısı olarak bilinir. (Tatlıdil, 1996)

En gelişmiş ve en karmaşık ilişki analizi olan kanonik korelasyon analizi ise, çok boyutlu kitleden çekilmiş iki ya da daha çok değişken kümesi arasındaki ilişki ile ilgilenir. Rastgele değişkenler kümesinin doğrusal fonksiyonları arasındaki maksimum korelasyonları bulmaya çalışan kanonik korelasyon analizinde tüm formüller iki rastgele değişken kümesi için geliştirilmiş olup küme sayısının ikiden çok olması durumlarında bu formüller geliştirilerek kullanılmaktadır.

Kanonik korelasyonda her bir kümenin rastgele değişkenlerinin, maksimum korelasyonlu ve birim varyanslı birer doğrusal bileşimleri elde edilmektedir. Daha sonra, bulunan bu çiftten bağımsız, maksimum korelasyonlu ve birim varyanslı ikinci bir doğrusal bileşim çifti bulunur. Bu işlemlere küçük değişken kümesindeki değişken sayısı kadar yeni doğrusal bileşim çifti elde edilinceye kadar devam edilir.

Kanonik korelasyon analiz yöntemine ait istatistik varsayımlar aşağıdaki gibidir.

a) Değişken kümeleri arasında ilişki doğrusal olmalıdır.

b) Her bir değişken kümesinin çok değişkenli normal dağılım göstermesi gerekir.

c) İki grup değişken kümesinde yer alan değişkenlerin eşit sayıda olma zorunluluğu yoktur.

d) Değişkenler arasındaki korelasyonu önemli düzeyde etkilemesi nedeni ile veri kümesinde aykırı değerlerin analiz öncesinde saptanarak gerekli düzeltme ya da elimine edilmesi gerekmektedir.

e) Her değişken kümesindeki değişkenler arasında çoklu bağlantı veya çoklu birlikte değişim (multicollinearity) bulunmamalıdır. nxk boyutlu X matrisinin rankının k'ya eşit $(k < n)$ olması, bağımsız değişkenler arasında doğrusal bir bağımlılığın olmadığını ifade etmektedir.

3.2. Kanonik Değişkenler ve Kanonik Korelasyonların Elde Edilmesi ve Tanımı

X değişken kümesinde p tane ve Y değişken kümesinde q tane $(p \leq q)$ değişken var ise, bu iki değişken kümesindeki değişkenlerin doğrusal kombinasyonları alınarak, bunlar arasındaki korelasyon hesaplanabilir. Bu şekilde doğrusal kombinasyonlardan büyük korelasyona ilk kanonik korelasyon, değişken kümelerinden oluşan doğrusal kombinasyonlara ise kanonik değişken adı verilir. X değişken kümesi $px1$ boyutlu μ_1 ortalama vektörüne, Y değişken kümesi ise $qx1$ boyutlu μ_2 ortalama vektörüne sahip olsun. Teorem 1.2.4'e göre bu değişken kümelerine ait ortalama ve kovaryans matrisleri

$$\mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \qquad \Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \tag{3.2.1}$$

olsun.

X ve Y değişken kümelerinin keyfi doğrusal bileşimleri sırasıyla u ve v olmak üzere,

$$u = \alpha' X \text{ ve } v = \gamma' Y \tag{3.2.2}$$

şeklinde verilsin. Burada α ve γ katsayıları sırasıyla $px1$ ve $qx1$'lik vektörlerdir. Bir araştırmaya konu olan X ve Y değişken kümeleri çok değişkenli normal dağılıma sahip ise, (3.2.2) bağıntısında verilen u ve v kanonik değişkenleri de normal dağılıma sahiptir ve kanonik değişkenleri arasındaki doğrusal ilişki maksimize edilebilmektedir. Eğer, X değişken kümesi bağımsız değişken, Y değişken kümesi bağımlı değişken olarak ifade edilirse, yani X, Y'nin sebebi olarak yorumlanırsa, bu durumda u "en iyi tahmin edici",v' de "en iyi tahmin edilebilir kriter" olarak isimlendirilebilir. u doğrusal bileşenleri

$$\begin{aligned} u_1 &= \alpha_1 x_{11} + \alpha_2 x_{12} + \cdots + \alpha_p x_{1p} \\ u_2 &= \alpha_1 x_{21} + \alpha_2 x_{22} + \cdots + \alpha_p x_{2p} \\ &\vdots \\ u_N &= \alpha_1 x_{N1} + \alpha_2 x_{N2} + \cdots + \alpha_p x_{Np} \end{aligned} \tag{3.2.3}$$

ve v doğrusal bileşenleri

$$\begin{aligned} v_1 &= \gamma_1 y_{11} + \gamma_2 y_{12} + \cdots + \gamma_q y_{1p} \\ v_2 &= \gamma_1 y_{21} + \gamma_2 y_{22} + \cdots + \gamma_q y_{2p} \\ &\vdots \\ v_N &= \gamma_1 y_{N1} + \gamma_2 Y_{N2} + \cdots + \gamma_q y_{Nq} \end{aligned} \tag{3.2.4}$$

dir $(N = p - 1)$.

u ve v değişkenleri arasındaki korelasyonun maksimum olmasının sağlanabilmesi ve u ve v değişkenlerinin birim varyanslı olabilmeleri için α ve γ vektörlerinin özel seçilmeleri gerekmektedir. Ayrıca bu vektörlerin normlandırılması yorum yönünden de kolaylık sağlayacaktır. Bu düşünceler ışığında,

$$\begin{aligned} Var(u) = Var(\alpha' X) &= E[(\alpha' X - \mathrm{E}(\alpha' X))(\alpha' X - \mathrm{E}(\alpha' X))'] \\ &= \alpha' E[(\mathrm{X} - \mathrm{E}(\mathrm{X}))(\mathrm{X} - \mathrm{E}(\mathrm{X}))']\alpha \\ &= \alpha' \mathrm{Var}(\mathrm{X})\alpha \\ &= \alpha' \textstyle\sum_{11} \alpha \\ &= 1 \end{aligned} \qquad (3.2.5)$$

Benzer şekilde

$$\begin{aligned} Var(v) = Var(\gamma' Y) &= \mathrm{E}\big[\big(\gamma' Y - \mathrm{E}(\gamma' Y)\big)(\gamma' Y - \mathrm{E}(\gamma' Y)')\big] \\ &= \gamma' \Sigma_{22} \gamma \\ &= 1 \end{aligned} \qquad (3.2.6)$$

$$Kor(u, v) = \frac{Kov(u,v)}{\sqrt{Var(u).Var(v)}} = \frac{Kov(u,v)}{\sqrt{1.1}} = Kov(u, v) \qquad (3.2.7)$$

$$\begin{aligned} Kov(u, v) &= E[(u - E(u)(v - E(v))'] \\ &= \alpha' \textstyle\sum_{12} \gamma \\ &= \gamma' \textstyle\sum_{21} \alpha \\ &= \rho \end{aligned} \qquad (3.2.8)$$

eşitliklerinden bulunur. Bu durumda amaç,

$$F_{u,v} = \underbrace{max}_{\alpha,\gamma} Kor(u, v) = max\ \alpha' \textstyle\sum_{12} \gamma \ = \rho_1 \qquad (3.2.9)$$

fonksiyonunu (3.2.5) ve (3.2.6) kısıtları altında maksimize etmektir. Böylece, u ve v kanonik değişken çifti arasındaki maksimum korelasyona birinci kanonik korelasyon adı verilir. Fonksiyon bu kısıtlar altında katsayıların maksimizasyon problemi olarak düşünülüp ortaya koymak için, λ_1 ve λ_2 langrange çarpanları olmak üzere bir langrange fonksiyonu biçiminde ifade edilebilir.

$$L = \alpha' \textstyle\sum_{12} \gamma - \frac{1}{2}\lambda_1\big(\alpha' \Sigma_{11} \alpha - 1\big) - \frac{1}{2}\lambda_2(\gamma' \Sigma_{22} \gamma - 1) \qquad (3.2.10)$$

Bu fonksiyonun α ve γ vektörlerine göre türevleri alınıp sıfıra eşitlendiğinde elde edilen değerler yukarıda sıralanan koşulları sağlayacaktır.

$$\frac{\partial L}{\partial \alpha} = \Sigma_{12}\gamma - \frac{1}{2}\lambda_1(2\Sigma_{11}\alpha)$$
$$= \Sigma_{12}\gamma - \lambda_1\Sigma_{11}\alpha$$
$$= 0 \qquad (3.2.11)$$

$$\frac{\partial L}{\partial \gamma} = \alpha'\Sigma_{12} - \frac{1}{2}\lambda_2(2\Sigma_{22}\gamma)$$
$$= \Sigma_{21}\alpha - \lambda_2\Sigma_{22}\gamma$$
$$= 0 \qquad (3.2.12)$$

Yukarıdaki ilk eşitlik soldan α' ile ikinci eşitlik yine soldan γ' vektörleri ile çarpılırsa,

$$\alpha'\Sigma_{12}\gamma - \lambda_1(\alpha'\Sigma_{11}\alpha) = 0$$
$$\alpha'\Sigma_{12}\gamma = \lambda_1(\alpha'\Sigma_{11}\alpha)$$
$$\lambda_1 = \alpha'\Sigma_{12}\gamma \qquad (3.2.13)$$

$$\gamma'\Sigma_{21}\alpha - \lambda_2\gamma'\Sigma_{22}\gamma = 0$$
$$\gamma'\Sigma_{21}\alpha = \lambda_2\gamma'\Sigma_{22}\gamma$$
$$\lambda_2 = \gamma'\Sigma_{21}\alpha \qquad (3.2.14)$$

olduğu ve bunun her ikisinin de (3.2.8) nolu gösterimden korelasyon katsayısına (ρ) ya eşit oldukları görülür. Yani,

$$\lambda_1 = \lambda_2 = \alpha'\Sigma_{12}\gamma = \gamma'\Sigma_{21}\alpha = \rho \qquad (3.2.15)$$

Bu bilgiler ışığında (3.2.11) ve (3.2.12) nolu gösterim,

$$-\rho\Sigma_{11}\alpha + \Sigma_{12}\gamma = 0$$
$$\Sigma_{21}\alpha - \rho\Sigma_{22}\gamma = 0 \qquad (3.2.16)$$

olup

$$\begin{bmatrix} -\rho\Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & -\rho\Sigma_{22} \end{bmatrix}\begin{bmatrix} \alpha \\ \gamma \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad (3.2.17)$$

matris biçiminde yazılabilmektedir.

Bu denklem sisteminde α ve γ vektörlerinin elemanları sıfırdan farklı olacaktır. Yazılan eşitliğin sağlanabilmesi(sıfır olabilmesi) için ilk matrisin tekil, yani determinant değerinin sıfır olması gerekir. Bu matrisin determinant değerinin sıfır yapacak ρ değerinin elde edilmesi için,

$-\rho\Sigma_{22} > 0$ ve tekil olduğundan

$$\begin{vmatrix} -\rho\Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & -\rho\Sigma_{22} \end{vmatrix} = 0 \Rightarrow |-\rho\Sigma_{22}|\left|-\rho\Sigma_{11} + \frac{1}{\rho}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\right|$$

$$\Rightarrow |-\rho\Sigma_{22}|\left|\frac{1}{\rho}(-\rho^2\Sigma_{11} + \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})\right|$$

$$\Rightarrow \rho^q|\Sigma_{22}|\frac{1}{\rho}^p|-\rho^2\Sigma_{11} + \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}|$$

$$\Rightarrow \rho^{q-p}|\Sigma_{22}||-\rho^2\Sigma_{11} + \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}|$$

$$\Rightarrow |-\rho^2\Sigma_{11} + \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}| = 0 \qquad (3.2.18)$$

ya da

$$\begin{vmatrix} -\rho\Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & -\rho\Sigma_{22} \end{vmatrix} = 0 \Rightarrow |-\rho\Sigma_{22}|\left|-\rho\Sigma_{11} + \frac{1}{\rho}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\right|$$

$$\Rightarrow |-\rho\Sigma_{22}|\left|\frac{1}{\rho}(-\rho^2\Sigma_{11} + \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})\right|$$

$$\Rightarrow |-\rho\Sigma_{22}|\left|\Sigma_{11}(-\rho^2\mathrm{I_n} + \Sigma_{11}{}^{-1}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})\right|$$

$$\Rightarrow |-\rho\Sigma_{22}||\Sigma_{11}|\left|(\Sigma_{11}{}^{-1}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} - \rho^2\mathrm{I_n})\right|$$

$$\Rightarrow \left|(\Sigma_{11}{}^{-1}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} - \rho^2\mathrm{I_n})\right| = 0 \qquad (3.2.19)$$

dır.

Aynı zamanda,

$-\rho\Sigma_{11} > 0$ ve tekil olduğundan

$$\begin{vmatrix} -\rho\Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & -\rho\Sigma_{22} \end{vmatrix} = 0 \Rightarrow |-\rho\Sigma_{11}|\left|-\rho\Sigma_{22} + \frac{1}{\rho}\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\right|$$

$$\Rightarrow \rho^p|\Sigma_{11}|\frac{1}{\rho}^q|-\rho^2\Sigma_{22} + \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}|$$

$$\Rightarrow \rho^{p-q}|\Sigma_{11}||-\rho^2\Sigma_{22} + \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}|$$

$$\Rightarrow |-\rho^2\Sigma_{22} + \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}| = 0 \qquad (3.2.20)$$

ya da

$$\begin{vmatrix} -\rho\Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & -\rho\Sigma_{22} \end{vmatrix} = 0 \Rightarrow |-\rho\Sigma_{11}|\left|-\rho\Sigma_{22} + \frac{1}{\rho}\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\right|$$

$$\Rightarrow \left|\Sigma_{22}(-\rho^2\mathrm{I_n} + \Sigma_{22}{}^{-1}\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\right|$$

$$\Rightarrow |\Sigma_{22}|\left|(\Sigma_{22}{}^{-1}\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12} - \rho^2\mathrm{I_n}\right|$$

$$\Rightarrow \left|(\Sigma_{22}{}^{-1}\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12} - \rho^2\mathrm{I_n}\right| = 0 \qquad (3.2.21)$$

ya da

$$\begin{vmatrix} -\rho\Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & -\rho\Sigma_{22} \end{vmatrix} = 0 \Rightarrow |-\rho^2\Sigma_{11}\Sigma_{22} + \Sigma_{12}\Sigma_{21}| = 0 \quad (3.2.22)$$

işlemlerinden biri yapılır. Bulunan ρ^2 değeri yerine konarak α ve γ vektörleri aşağıdaki (karakteristik denklemler olarak adlandırılan) denklemlerden elde edilir.

$$(-\rho^2\Sigma_{11} + \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})\alpha = \left(\Sigma_{11}{}^{-1}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} - \rho^2 I_n\right)\alpha = 0$$

$$(-\rho^2\Sigma_{22} + \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12})\gamma = \left(\Sigma_{22}{}^{-1}\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12} - \rho^2 I_n\right)\gamma = 0 \quad (3.2.23)$$

Yukarıda verilen (3.2.18), (3.2.19), (3.2.20), (3.2.21), (3.2.22) bağıntılardan maksimum p adet korelasyon katsayısı elde edilir. Bunun nedeni, kovaryans matrisinde $p \leq q$ olduğundan, verilen kovaryans matrisinin rankı maksimum p olacaktır. Bu sebeple, eşitlikten p tane sıfırdan farklı ρ^2 elde edilebilir. Bulunan bu değerlerin pozitif kareköklerine kanonik korelasyon denir. Elde edilen kanonik korelasyon katsayıları büyükten küçükten doğru sıralanır ($\rho_1 > \rho_2 > \cdots > \rho_p$) ve en büyükten küçüğe doğru olmak koşuluyla, tek tek (3.2.18), (3.2.19), (3.2.20), (3.2.21), (3.2.22) nolu eşitliklerinden birinde yerlerine konularak, u ve v kanonik değişkenleri elde edilir. En büyük kanonik korelasyon katsayısının denklemde yerine konulması ile elde edilen kanonik değişkenlere, birinci kanonik değişken çifti (u_1,v_1) adı verilmektedir. Bu arada, elde edilen p tane kanonik değişken çiftinin birbirinden bağımsız olması gerektiğinin hatırlatılmasında yarar vardır. Öteki u_i ve v_i değişkenleri de benzer biçimde yorumlanır.

Bu aşamada bir başka sorunun da göz önüne alınması gerekir, o da küme içindeki kümeler arasındaki kanonik değişkenlerin birbirinden bağımsız olmalarının sağlanmasıdır.

$$Kov(u_i, u_j) = E\left[(u_i - E(u_i)(u_j - E(u_j))'\right] = \alpha_i'\Sigma_{11}\alpha_j = 0$$

$$Kov(v_i, v_j) = E\left[(v_i - E(v_i)(v_j - E(v_j))'\right] = \gamma_i'\Sigma_{22}\gamma_j = 0 \quad (3.2.24)$$

$$Kov(u_i, v_j) = E\left[(u_i - E(u_i)(v_j - E(v_j))'\right] = \alpha_i'\Sigma_{12}\gamma_j = 0$$

(3.2.11) nolu gösterim j-yinci durum için yazılır ve soldan α_i' ile çarpılırsa,

$$\alpha_i' \Sigma_{12}\gamma_j = \lambda_j\alpha_i' \Sigma_{11}\alpha_j \Rightarrow \alpha_i'\Sigma_{11}\alpha_j = \frac{1}{\rho_j}\alpha_i' \Sigma_{12}\gamma_j = 0 \quad (3.2.25)$$

sonucuna ulaşılır. Bu eşitlikte $\rho_j \neq 0$ olacağından bu eşitliğin sıfıra eşit olabilmesi için $\alpha_i' \Sigma_{11}\alpha_j = \gamma_i'\Sigma_{22}\gamma_j = \alpha_i' \Sigma_{12}\gamma_j = 0$ olması gerekir. Bu durumda (3.2.5) ve (3.2.6) nolu gösterimlerdeki birim varyans kısıtı ile birlikte $\alpha_i' \Sigma_{11}\alpha_j = 0$ ve $\gamma_i'\Sigma_{22}\gamma_j = 0$ kısıtlarını da α ve γ katsayılar vektörlerinin bulunmasında dikkate alınması gerekmektedir. O halde $\lambda_1\, ve\, \lambda_2$ langrange katsayıları ile birlikte $\vartheta_1, \vartheta_2 \dots \vartheta_p$ ve $\theta_1, \theta_2 \dots \theta_p$ katsayıları da denkleme katıldığında herhangi bir i-yinci durum için denklem

$$L_i = \alpha_i' \sum_{12}\gamma_i - \frac{1}{2}\lambda_1(\alpha_i'\Sigma_{11}\alpha_i - 1) - \frac{1}{2}\lambda_2(\gamma_i'\Sigma_{22}\gamma_i - 1) + \sum_{j=1}^{p}\vartheta_j\alpha_i'\Sigma_{11}\alpha_j + + \sum_{j=1}^{p}\theta_j\gamma_i'\Sigma_{22}\gamma_j\,, \quad i = 1,2 \dots p \tag{3.2.26}$$

biçiminde yazılır. Bu fonksiyonun da α_i ve γ_i vektör değişkenlerine göre kısmi türevlerinin alınıp sıfıra eşitlenmesinden

$$\frac{\partial L_i}{\partial \alpha_i} = \sum_{12}\gamma_i - \lambda_1(\Sigma_{11}\alpha_i) + \sum_{j=1}^{p}\vartheta_j\Sigma_{11}\alpha_j = 0 \tag{3.2.27}$$

$$\frac{\partial L_i}{\partial \gamma_i} = \Sigma_{21}\alpha_i - \lambda_2(\Sigma_{22}\gamma_i) + \sum_{j=1}^{p}\theta_j\Sigma_{22}\gamma_j = 0 \tag{3.2.28}$$

denklemleri elde edilir. Yukarıdaki eşitliklerden ilki soldan α_j' ile ikinci soldan γ_j' ile çarpılacak olursa (3.2.5), (3.2.6) ve yukarıda verilen kısıtlardan yararlanılarak,

$$\alpha_j'\sum_{12}\gamma_i - \lambda_1\left(\alpha_j'\Sigma_{11}\alpha_i\right) + \sum_{j=1}^{p}\vartheta_j\alpha_j'\Sigma_{11}\alpha_j = 0 \tag{3.2.29}$$

$$\gamma_j'\Sigma_{21}\alpha_i - \lambda_2\left(\gamma_j'\Sigma_{22}\gamma_i\right) + \sum_{j=1}^{p}\theta_j\gamma_j'\Sigma_{22}\gamma_j = 0 \tag{3.2.30}$$

$$\Rightarrow \sum_{j=1}^{p}\vartheta_j\alpha_j'\Sigma_{11}\alpha_j = \vartheta_j = 0 \tag{3.2.31}$$

$$\Rightarrow \sum_{j=1}^{p}\theta_j\gamma_j'\Sigma_{22}\gamma_j = \theta_j = 0 \tag{3.2.32}$$

sonuçlarına ulaşılır. Yani fonksiyon yeni eklenen kısıtlardan etkilenmemektedir. O halde kanonik değişkenler elde edilirken yeni kısıtların denkleme katılmasına gerek yoktur. Bu durumda,

$$Kov(u_i, u_j) = Kov(v_i, v_j) = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

$$Kov(u_i, v_j) = Kov(u_j, v_i) = \begin{cases} \rho, & i = j \\ 0, & i \neq j \end{cases} \tag{3.2.33}$$

sonucuna ulaşılır. Bu sonuçları özet olarak,

$$Kov(u, v) = \begin{bmatrix} I_p & \rho \\ \rho & I_p \end{bmatrix} \tag{3.2.34}$$

matris formunda göstermek mümkündür.

Veri kümesindeki değişkenlerin ölçü birimlerinin ve varyanslarının farklı olması durumunda, ya değişkenlerin standartlaştırılması ya da korelasyon matrisine göre kanonik korelasyon analizi yapılması gerekir. Çünkü varyansları farklı veri kümelerinin kovaryans matrisine göre elde edilen çözümler ile korelasyon matrisine göre elde edilen çözümler, farklılıklar gösterirken, verilerin standartlaştırılması ile iki yöntem arasındaki çözüm farklılıkları ortadan kalkmaktadır. Ayrıca, kovaryans matrisinde, değişken çiftleri arasındaki değişim, yani, değişkenler arasındaki ilişki ortaya konulmaktadır. Korelasyon matrisi sayesinde, değişken çiftleri arasındaki ilişkinin büyüklüğünün ve yönünün ortaya konulmasıyla, aralarındaki ilişki daha iyi yorumlanabilmektedir. Bu sebeple, uygulamada genellikle Teorem 1.2.5 deki korelasyon matrisini kullanarak kanonik korelasyon analizi yapılması analizdeki hesaplama süreci bakımından kolaylık sağlamaktadır.

Korelasyon matrisinde köşegen elemanlar 1, köşegen dışı elemanlar ise -1 ile +1 arasında olduğundan Teorem 1.2.5 de verilen korelasyon matrisi aracılığıyla, kanonik değişkenlerin ve kanonik korelasyonların elde edilmesi için P_1 ve P_2 matrislerinden yararlanılır.

$$\begin{aligned} P_1 &= R_{11}^{-1}R_{12}R_{22}^{-1}R_{21} \\ P_2 &= R_{22}^{-1}R_{21}R_{11}^{-1}R_{12} \end{aligned} \qquad (3.2.35)$$

(3.2.35) bağıntısında verilen, $R_{11}^{-1}R_{12}R_{22}^{-1}R_{21}$ veya $R_{22}^{-1}R_{21}R_{11}^{-1}R_{12}$ matrislerinin özdeğerleri, ilgili kanonik değişken çifti arasındaki kanonik korelasyon katsayılarının karesini vermektedir.

$$r_{u_i v_i} = \sqrt{\lambda_i}, \qquad i = 1,2,\dots,p \qquad (3.2.36)$$

Kanonik değişken çiftleri arasındaki maksimum kanonik korelasyon katsayısı

$$r_{u_1 v_1} = \sqrt{\lambda_1} \qquad (3.2.37)$$

dir. X değişken kümesine ait kanonik katsayının belirlenmesinde P_1 matrisinin, Y kümesi için ise, P_2 matrisinin öz değer-öz vektör çiftlerinden yararlanılır.

X değişken kümesine ait en büyük öz değer λ_1 için standartlaştırılmamış kanonik katsayılar öz vektör elemanı, e_1;

$$(P_1 - \lambda_1 I)e_1 = 0 \qquad (3.2.38)$$

eşitliği sayesinde hesaplanabilmektedir.

Y değişken kümesine ait en büyük öz değer λ_1 için standartlaştırılmamış kanonik katsayılar öz vektör elemanı, f_1;

$$(P_2 - \lambda_1 I)f_1 = 0 \qquad (3.2.39)$$

eşitliğinden hesaplanabilmektedir. Buradan elde edilen öz vektörler (standartlaştırılmamış kanonik katsayılar), orijinal değişkende meydana gelen bir standart sapmalık artışa karşılık, kanonik değişkende standart sapma cinsinden meydana gelen değişim miktarını göstermektedir. Bir başka değişle, bu katsayılar, bir kümedeki kanonik değişkenin oluşmasında, o kümede yer alan orijinal değişkenlerin etki miktarlarını göstermektedir.

X ve Y değişken kümeleri için, standardart kanonik katsayılar sırası ile (3.2.40) bağıntısı yardımı ile hesaplanabilmektedir.

$$\alpha_i = \frac{e_i}{\sqrt{Var(u_i)}} = \frac{e_i}{\sqrt{(e_{i1},e_{i2},\ldots e_{ip})\Sigma_{11}(e_{i1},e_{i2},\ldots,e_{ip})'}}$$

$$\gamma_i = \frac{f_i}{\sqrt{Var(v_i)}} = \frac{f_i}{\sqrt{(f_{i1},f_{i2},\ldots f_{ip})\Sigma_{22}(f_{i1},f_{i2},\ldots,f_{ip})'}} \qquad (3.2.40)$$

Eğer orijinal değişkenler

$$Z^1=(Z_1^1,Z_2^1,\ldots,Z_p^1)'$$
$$Z^2=(Z_1^2,Z_2^2,\ldots,Z_q^2)' \qquad (3.2.41)$$

şeklinde standartlaştırılırsa, kanonik değişkenler,

$$u_i=\alpha_i'Z^{(1)}=e_i'\rho_{11}^{-1/2}Z^{(1)}$$
$$=\gamma_i'Z^{(2)}=f_i'\rho_{22}^{-1/2}Z^{(2)} \qquad (3.2.42)$$

şeklinde ifade edilir. Burada, $Kov(Z^{(1)}) = \rho_{11}$, $Kov(Z^{(2)}) = \rho_{22}, Kov(Z^{(1)},Z^{(2)}) = \rho_{12} = \rho_{21})$ iken e_i ve f_i değerleri $\rho_{11}^{-1/2}\rho_{12}\rho_{22}^{-1}\rho_{21}\rho_{11}^{-1/2}$ ve $\rho_{22}^{-1/2}\rho_{21}\rho_{11}^{-1}\rho_{12}\rho_{22}^{-1/2}$ nin özvektörleri olarak kabul edilebilir. Kanonik korelasyon ρ_i^*, $Kor(u_i, v_i) = \rho_i^*$, $i = 1,2,\ldots,p$ şartını sağlar (Johnson and Wichern, 2007).

Burada oluşturulan u_1 doğrusal bileşenine X'in birinci kanonik değişkeni, v_1 doğrusal bileşenine ise Y'nin birinci kanonik değişkeni denir. Bunun yanında u_1 ve v_1 arasındaki ilişki, birinci kanonik korelasyon, birinci kanonik korelasyonun karesi de birinci özdeğer adını alır.

Oluşturulan her bir (u, v) değişken çiftinin değerlerini α ve γ katsayılarını hesaplayabiliriz. u ve v çifti arasındaki ilişkinin sonucu α ve γ katsayılarına bağlıdır. Bu nedenle kanonik korelasyon analizinde u ve v arasındaki ilişkiyi maksimum yapan α ve γ katsayılarının değerleri seçilir.

Sonuç olarak, birinci kanonik korelasyon, oluşturulan bileşenler arasında en yüksek ilişkiyi mümkün kılmaktadır. Genel olarak bu süreç, diğer kümelerin kanonik değişken çiftlerinin oluşturulması ile devam eder.

3.3. Kanonik Değişkenlerle Orijinal Değişkenler Arasındaki Korelasyonlar ve Yorumları

X ve Y değişken kümelerinden elde edilen u ve v kanonik değişkenler, hem kendi değişken kümeleri içerisindeki (yani, u ile $x_1, x_2, \ldots, x_p$ arasında), hem de diğer kümenin orijinal değişkenleri (yani, u ile $\mathrm{y}_1, \mathrm{y}_2, \ldots, \mathrm{y}_\mathrm{q}$ arasında) ile bir ilişkinin olması ve bunun yorumu ile kanonik değişkene herhangi bir orijinal değişkenin ne ölçüde katkı sağladığını ortaya koyma açısından oldukça önemlidir. Bu bilgiler doğrultusunda $\mathrm{u_i}$ kanonik değişkeni ile kendi kümesindeki (X) orijinal değişkenler arasındaki korelasyonlar ve $\mathrm{u_i}$ kanonik değişkeni ile Y değişken kümesindeki orijinal değişkenler arasındaki korelasyonlar ve $\mathrm{v_i}$ kanonik değişkeni ile X değişken kümesindeki orijinal değişkenler arasındaki korelasyonlar ve $\mathrm{v_i}$ kanonik değişkeni ile Y değişken kümesindeki orijinal değişkenler arasındaki korelasyonlar aşağıda verilmiştir.

$$Kor(u_i, X) = \frac{Kov(u_i, X_i)}{\{[Köş(Var(u_i)(Var(X_i)]\}^{1/2}}$$

$$= \frac{\alpha_i' \Sigma_{11}}{[Köş(\Sigma_{11})]^{1/2}} \tag{3.3.1}$$

Buna göre diğer eşitlikler sırası ile,

$$Kor(u_i, Y) = \frac{\alpha_i' \Sigma_{12}}{[Köş(\Sigma_{22})]^{1/2}} \tag{3.3.2}$$

$$Kor(v_i, X) = \frac{\gamma_i' \Sigma_{21}}{[Köş(\Sigma_{11})]^{1/2}} \tag{3.3.3}$$

$$Kor(v_i, Y) = \frac{\gamma_i' \Sigma_{22}}{[Köş(\Sigma_{22})]^{1/2}} \tag{3.3.4}$$

olarak verilebilmektedir. Burada, $Köş(.)$ matrisin köşegen elemanıdır. Elde edilen korelasyon katsayıları arasında da aşağıdaki gibi bir ilişki söz konusudur.

$$Kor(u_i, Y) = \rho_i Kor(v_i, Y) \tag{3.3.5}$$

$$Kor(v_i, X) = \rho_i Kor(u_i, X) \tag{3.3.6}$$

$$-\rho_\mathrm{i} \Sigma_{11} \alpha_\mathrm{i} + \Sigma_{12} \gamma_i = 0$$

$$\Sigma_{21} \alpha_i - \rho_i \Sigma_{22} \gamma_i = 0 \tag{3.3.7}$$

İlk eşitlik soldan $\Sigma_{11}{}^{-1}$ ile, ikinci eşitlik yine soldan $\Sigma_{22}{}^{-1}$ ile çarpılıp düzenlenecek olursa,

$$\begin{gathered} -\rho_\mathrm{i} \Sigma_{11}{}^{-1} \Sigma_{11} \alpha_\mathrm{i} + \Sigma_{11}{}^{-1} \Sigma_{12} \gamma_i = 0 \\ \Sigma_{22}{}^{-1} \Sigma_{21} \alpha_i - \rho_i \Sigma_{22}{}^{-1} \Sigma_{22} \gamma_i = 0 \end{gathered} \tag{3.3.8}$$

$$\Rightarrow -\rho_i \alpha_i = \Sigma_{11}{}^{-1}\, \Sigma_{12}\gamma_i$$

$$\rho_i \gamma_i = \Sigma_{22}{}^{-1} \Sigma_{21}\alpha_i \qquad (3.3.9)$$

$$\alpha_i = \frac{1}{\rho_i}\Sigma_{11}{}^{-1}\, \Sigma_{12}\gamma_i \qquad (3.3.10)$$

$$\gamma_i = \frac{1}{\rho_i}\Sigma_{22}{}^{-1}\, \Sigma_{21}\alpha_i \qquad (3.3.11)$$

sonuçları bulunur ve buna göre (3.3.6) bağıntısı gerçekleşmiş olur. Eşitlikte, ρ_i, $i-yinci$ kanonik değişken çifti arasındaki kanonik korelasyon katsayısını göstermektedir.

3.4. Kanonik Korelasyon İçin Özel Durumlar

a) Her iki kümede de değişken sayısının bir olduğu durumda ($p = q = 1$); varyans-kovaryans matrisi,

$$\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix} \qquad (3.4.1)$$

ve (3.2.21) bağıntısı

$$-\rho^2\sigma_{11} + \sigma_{12}\sigma_{22}^{-1}\sigma_{21} = 0 \qquad (3.4.2)$$

forma dönüşür. Buradan ρ^2 yalnız bırakılırsa

$$\rho^2 = \frac{\sigma_{12}^2}{\sigma_{11}\sigma_{22}} \qquad (3.4.3)$$

elde edilir. Bunun pozitif karekökü, Kanonik korelasyon katsayısını vermektedir. Bu ise, Pearson korelasyon katsayısına eşittir.

b) X ve Y değişken kümelerinden, birinde sadece bir değişken ve diğerinde ise, birden fazla değişken (p = 1, q > 1) söz konusu ise, varyans-kovaryans matrisi,

$$\Sigma = \begin{bmatrix} \sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \qquad (3.4.4)$$

$$-\rho^2\sigma_{11} + \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} = 0 \qquad (3.4.5)$$

Buradan ρ^2 yalnız bırakılırsa

$$\rho^2 = \frac{\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}}{\sigma_{11}} \qquad (3.4.6)$$

olarak elde edilir. Bu ρ^2 değerlerinin pozitif karekökü, kanonik korelasyon katsayısını vermektedir. Buradaki değişken kümelerinden biri bağımlı ($p = 1$) diğeri bağımsız

$(q > 1)$ olmak koşulu ile, elde edilen kanonik korelasyon katsayısının karesi (ρ^2), çoklu regresyon denklemin çoklu belirtme katsayına eşit olduğu (3.4.6) bağıntısı yardımıyla görülmektedir. (Çankaya, S., 2005)

3.5. Kanonik korelasyon katsayılarının önemlilik testi

Kanonik korelasyon analizi sonucunda elde edilen kanonik değişken çiftlerinden kaç tanesinin önemli olduğu, yani değişken grupları arasındaki ilişkinin kaç tanesi ile büyük ölçüde açıklanabileceğine karar vermek gerekir. Bu yöntemde amaç, bulunan kanonik korelasyon çiftlerinin kaç tanesi arasındaki ilişkinin önemli sayılıp sayılmayacağını test etmektir. Wilk's Lambda yaklaşımında tüm kanonik korelasyonların sıfıra eşit olduğu hipotezi alternatif hipoteze karşı test edilir.

$\mathrm{H_0} = \Sigma_{12} = 0$ ya da $\rho_1 = \rho_2 = \cdots = \rho_\mathrm{p} = 0$

$H_A = en\ az\ bir\ \rho_\mathrm{i} \neq 0$ (3.5.1)

$\mathrm{H_0}$ hipotezinin reddedilmesi durumunda değeri en büyük olan katsayı hipotezden çıkarılacak ve işlemler $\mathrm{H_0}$ hipotezi kabul edilinceye kadar tekrarlanacaktır. Wilk's Lambda test istatistiği aşağıdaki gibi elde edilir.

$$\Lambda = \prod_{i=1}^{p}(1-\rho_i^2) \quad (3.5.2)$$

Bu katsayı kullanılarak χ_h^2 test istatistik değeri,

$$\chi_\mathrm{h}^2 = -[(n-1)-(p+q+1)/2]\log_\mathrm{e}(\Lambda) \quad (3.5.3)$$

şeklinde hesaplanır. Bu eşitlikte n, örnek hacmini; p, birinci setteki değişken sayısını; q, ikinci setteki değişken sayısını; ρ_i, kanonik korelasyonları; k ise kanonik korelasyon sayısını belirtir.

Test istatistiğinin hesaplanan değeri χ_h^2 ile $\chi^2_{\mathrm{pxq;1-\alpha}}$ tablo değeri ile karşılaştırılır. $\chi_\mathrm{h}^2 > \chi^2_{\mathrm{pxq;1-\alpha}}$ ise, $\mathrm{H_0}$ hipotezi reddedilir. Yani birinci kanonik korelasyonun anlamlı olduğu söylenir. İlk hesaplanan test istatistiği χ_h^2 önemli ise birinci kanonik korelasyon test dışı bırakılır ve diğer kanonik korelasyonlar ile test yinelenir. Bu defa Wilk's Lambda istatistiği $i = 2,3,\ldots,k$ değerleri için hesaplanır.

$$\Lambda^* = \prod_{i=2}^{p}(1-\rho_i^2) \quad (3.5.4)$$

ve

$$\chi_\mathrm{h}^2 = -[(n-1)-(p+q+1)/2]\log(\Lambda^*) > \chi^2_{\mathrm{(p-1)x(q-1)1-\alpha}} \quad (3.5.5)$$

Bu işlemler önemsiz χ_h^2 değerine kadar devam eder. Ayrıca Wilk's Lambda katsayısı sıfıra yaklaştıkça, H_0 hipotezinin reddedileceği (kanonik korelasyon katsayısının anlamlı olduğunu), χ^2 değeri ile korelasyon katsayılarının sıfırdan farklı (anlamlı) olacağı söylenebilir.

3.6. Kısmi Kanonik Korelasyon Analizi

q tane değişkene sahip K değişken kümesi, r tane değişkene sahip L değişken kümesi ve p tane değişkene sahip Y değişken kümesinden bir tanesinin diğer iki değişken kümeleri üzerinden etkisinin kaldırılması durumunda, iki değişken kümesi arasındaki ilişki kısmi kanonik korelasyon analizi ile ortaya konmaktadır. Bu durumda yeni veri matrisi $x \sim N(\mu, \Sigma); \delta = p + q + r$ iken rastgele değişken vektörü ve kovaryans matrisi aşağıdaki gibidir:

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}; \qquad \Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} & \Sigma_{13} \\ \Sigma_{21} & \Sigma_{22} & \Sigma_{23} \\ \Sigma_{31} & \Sigma_{32} & \Sigma_{33} \end{bmatrix} \tag{3.6.1}$$

Burada x_1, x_2, x_3 alt vektörleri sırasıyla p,q ve r elemanlı alt matrisler Σ_{11}:pxp, Σ_{12}:pxq, Σ_{13}:pxr, Σ_{21}:qxp, Σ_{22}:qxq, Σ_{23}:qxr, Σ_{31}:rxp, Σ_{32}:rxq ve Σ_{33}:rxr boyutlu olarak tanımlıdır. Üçüncü değişkenin etkisi ortadan kaldırıldığında birinci ve ikinci değişken kümelerinden elde edilebilecek kanonik değişkenler,

$$u = \alpha' x_1$$

$$v = \gamma' x_2 \tag{3.6.2}$$

biçiminde gösterilmiş olsun. Ayrıca $t' = (u, v)$ biçimindeki vektörle de bu değişkenler gösterildiğinde x_3'ün etkisi sabit tutulduğunda t vektörünün varyans kovaryans matrisi

$$Var(t/x_3) = Var(t) - Kov(t, x_3)[Var(x_3)]^{-1} Kov(x_3, t)$$

$$= \begin{bmatrix} \alpha'(\Sigma_{11} - \Sigma_{13}\Sigma_{33}{}^{-1}\Sigma_{31})\alpha & \alpha'(\Sigma_{12} - \Sigma_{13}\Sigma_{33}{}^{-1}\Sigma_{32})\gamma \\ \gamma'(\Sigma_{21} - \Sigma_{23}\Sigma_{33}{}^{-1}\Sigma_{31})\alpha & \gamma'(\Sigma_{22} - \Sigma_{23}\Sigma_{33}{}^{-1}\Sigma_{32})\gamma \end{bmatrix}$$

$$= \begin{bmatrix} \alpha'\Sigma_{11.3}\alpha & \alpha'\Sigma_{12.3}\gamma \\ \gamma'\Sigma_{21.3}\alpha & \gamma'\Sigma_{22.3}\gamma \end{bmatrix} \tag{3.6.3}$$

biçiminde olacaktır. Bu sonuçlara göre,

$$Var(u/x_3) = \alpha' \Sigma_{11.3} \alpha = 1 \quad (3.6.4)$$

$$Var(v/x_3) = \gamma' \Sigma_{22.3} \gamma = 1 \quad (3.6.5)$$

$$Var(u/x_3, v/x_3) = \alpha' \Sigma_{12.3} \gamma = 1 \quad (3.6.6)$$

olduğu açıktır. Bu bilgiler ışında (3.2.20) nolu eşitliğe benzer biçimdeki

$$\begin{bmatrix} -\rho_{.3}\Sigma_{11.3} & \Sigma_{12.3} \\ \Sigma_{21.3} & -\rho_{.3}\Sigma_{22.3} \end{bmatrix} \begin{bmatrix} \alpha \\ \gamma \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (3.6.7)$$

denklem sistemine ulaşılır. α ve γ vektörlerinin katsayıları sıfır olamayacağı için yukarıdaki detreminant değerini sıfır yapacak $\rho_{.3}$ değeri

$$|-\rho_{.3}{}^2\Sigma_{11.3} + \Sigma_{12.3}\Sigma_{22.3}^{-1}\Sigma_{21.3}| = 0 \quad (3.6.8)$$

$$|-\rho_{.3}{}^2\Sigma_{22.3} + \Sigma_{21.3}\Sigma_{11.3}^{-1}\Sigma_{12.3}| = 0 \quad (3.6.9)$$

denklemlerinin birinin çözmünden elde edilir. Bulunan $\rho_{.3}{}^2$ değerlerinin eşitlikte yerine konması ile de kanonik değişken katsayıları bulunur.

$$(-\rho_{.3}{}^2\Sigma_{11.3} + \Sigma_{12.3}\Sigma_{22.3}^{-1}\Sigma_{21.3})\alpha_{.3} = 0 \quad (3.6.10)$$

$$(-\rho_{.3}{}^2\Sigma_{22.3} + \Sigma_{21.3}\Sigma_{11.3}^{-1}\Sigma_{12.3})\gamma_{.3} = 0 \quad (3.6.11)$$

Yukarıdaki (3.6.8) ve (3.6.9) bağıntılardan elde edilen $\rho_{i.3}{}^2$ değerlerinin kareköklerine x_3 ün etkisi ortadan kaldırıldırıldığında x_1 ve x_2 değişken kümesi arasındaki kısmi kanonik katsayıları adı verilir. $\alpha_{i.3}$ ve $\gamma_{i.3}$ vektörlerine ise, sırasıyla x_1 ve x_2 uzayındaki kısmi kanonik vektörler adı verilir. Bunların sonucu olarakta $u_{i.3} = \alpha'_{i.3}x_1$ ve $v_{i.3} = \gamma'_{i.3}x_2$ değişkenlerine x_1 ve x_2 uzaylarındaki kısmi kanonik değişkenler adı verilir.

3.7. Kernel Kanonik Korelasyon Analizi

Kernel kanonik korelasyon analizini basit kanonik korelasyon analizine farklı bir açıklama getirerek ifade edilecektir: x ve y iki rasgele vektör olsun. $u = a'x$ ve $v = b'y$ izdüşümlerini düşünelim. Ayrıca x ve y iki rasgele vektörün sıfır ortalamaya sahip olduğunu kabul edilsin. Aşağıdaki ifadeyi ele alalım.

$$\rho = \frac{E(uv)}{\sqrt{E(u^2)E(v^2)}} \tag{3.7.1}$$

u ve v bir boyutlu rasgele değişkenler arasındaki korelasyonu maksimumlaştırmak olduğundan, (3.7.1) bağıntısı maksimumlaştırlmalıdır. Ölçü bağımlılığından uzaklaşmak için izdüşümleri standard sapmalarıyla bölünür: Örneğin; verinin merkezi olmama durumlarına nasıl genelleştirildiği bulunabilir. Aşağıdaki ‘’hileyi’’ ortaya koyalım. Problemi değiştirmeksizin a ve b yeniden ölçeklendirildiğinden onlar 1’e eşit olarak alınır. Bu takdirde bu, problemi aşağıdaki gibi şekilde yazılmaya imkan verir. a ve b değişken vektörler olmak üzere,

$$E(u^2) = 1, \qquad E(v^2) = 1$$

şartları altında

$$\underbrace{max}_{a,b}\ \rho = E(u,v) \tag{3.7.2}$$

veya bir Langrange fonksiyonu inşa eder ve beklenen değerleri tam olarak yazılırsa,

$$min_{a,b} max_{\lambda_1,\lambda_2} \sum_i a' x_i y_i' b - \frac{1}{2}\lambda_1\left(\sum_i a'\, x_i x_i' a - N\right)$$

$$-\frac{1}{2}\lambda_2\left(\sum_i b' y_i y_i' b - N\right) \tag{3.7.3}$$

elde edilir. Burada fonksiyon N ile çarpılmıştır. a ve b ye göre türevleri alınarak aşağıdaki denklem elde edilir.

$$\sum_i x_i y_i' b - \lambda_1 \sum_i x_i x_i' a = 0$$

$$\sum_i y_i x_i' a - \lambda_2 \sum_i y_i y_i' b = 0 \tag{3.7.4}$$

İlk olarak, birinci denklemi a' ve ikinci denklemi b' ile çarpar ve kısıtlamaları kullanarak birinci denklemden ikinci denklemi çıkarırsak λ_1=$\lambda_2 = \lambda$ olduğu görülür. Şimdi, S_x, S_y, S_{xy} yeni adlandırmaları kullanılsın ve aşağıdaki daha büyük matrisler tanımlansın. S_D, köşegeni üzerinde S_x ve S_y, köşegen dışında sıfır bulunan blok köşegen matristir. Aynı zamanda S_0, köşegen dışında S_{xy} olan köşegen olmayan matris olarak tanımlansın. Son olarak $c' = [a, b]$ olsun. İki denklem bileşik olarak aşağıdaki gibi yazılabilir:

$$S_0 c = \lambda S_D c \Rightarrow S_D^{-1} S_0 c = \lambda c \Rightarrow S_0^{1/2} S_D^{-1} S_0^{1/2} (S_0^{1/2}\ c) = \lambda (S_0^{1/2} c) \tag{3.7.5}$$

Bu yine $c' = (S_0^{1/2} c)$ için düzenli bir öz değer denklemidir.

Kernel metodları son zamanlarda pozitif tanımlı kerneller tarafından lineer olmayan veri analizi için bir metod olarak geliştirilmiştir. Kernel kanonik korelasyon analizi pozitif tanımlı kerneller ile kanonik korelasyon analizinin lineer olmayan bir genişlemesi olarakta ifade edilir. X ve Y iki rasgele değişken verilsin. Kernel kanonik korelasyon analizi iki rasgele değişkenler tarafından paylaşılan bilginin özünü çıkartmayı amaçlar. Daha kesin bir ifade ile, Kernel kanonik korelasyon analizinin amacı korelasyonları maksimum olacak şekilde uzaya ait olan lineer olmayan $\phi(x_i)$ ve $\psi(y_i)$ fonksiyonlarını sağlayabilmektir.(Bach ,Fukumizu and Gretton, 2007)

Genel olarak, Kernel Kanonik Korelasyon başlama noktası, verileri(data-cases;veri matrisini) $\phi(x_i)$ ve $\psi(y_i)$ özellik vektörlerine dönüştürmektedir. Uzayın boyutu işleme sokulan kümenin veri matrisinin sayısından daha büyük ise, bu takdirde çözüm verilerin aralığında olmalıdır. Yani,

$$a = \sum_i \alpha_i{}' \phi(x_i)\,, \qquad b = \sum_i \beta_i \psi(y_i) \tag{3.7.6}$$

dir.

Bu bağıntıdan Langrange denklemi kullanılarak aşağıdaki bağıntı elde edilir.

$$L = \alpha' K_x K_y \beta - \frac{1}{2}\lambda\left(\alpha' {K_x}^2 \alpha - N\right) - \frac{1}{2}\lambda(\beta' {K_y}^2 \beta - N) \tag{3.7.7}$$

olup, burada α örneğin bir D-boyutlu uzayda bulunan a'dan hariç bir N-boyutlu uzaydaki bir vektördür ve $\mathrm{K_x} = \sum_i \phi(x_i)' \phi(x_i)$ ve K_y'de benzer şekilde tanımlanabilir. α ve β ya göre türevler alınarak,

$$K_x K_y \beta = \lambda {K_x}^2 \alpha$$

$$K_y K_x \alpha = \lambda {K_y}^2 \beta \tag{3.7.8}$$

bulunur.

K_x'in tam ranklı olduğunu kabul etmek suretiyle bu denklemleri çözelim.

$$\lambda^{-1} K_x^{-1} K_y \beta = {K_x}^2 \alpha \tag{3.7.9}$$

ve bu nedenle ${K_y}^2 \beta = \lambda^2 {K_y}^2 \beta$ dır. Bu ise, λ=1 için daima bir çözüme sahiptir.

$$\rho = \frac{1}{N} \sum_i a' S_{xy} b = \sum_i \lambda a' S_x a = \lambda \tag{3.7.10}$$

olduğunu hatırlayarak bunun maksimal korelasyonu göstereceğine ve bu nedenle tercih edilecek çözüm olduğuna dikkat edelim. Bu, Kernel yöntemlerinde yine düzenleme gerektiğini vurgulayan aşırı uyumun tipik bir durumudur. Bu düzenleme Langrange fonksiyonunda kısıtlamaya (veya eşdeğer olarak orijinal amaç fonksiyonunun paydasına) yeni bir köşegen terim eklemek suretiyle yapılabilir. Bu ise, bizi aşağıdaki Langrange fonksiyonuna götürür.

$$L = \alpha' K_x K_y \beta - \frac{1}{2}\lambda\left(\alpha' {K_x}^2 \alpha + \eta\|\alpha\|^2 - N\right) - \frac{1}{2}\lambda(\beta' {K_y}^2 \beta + \eta\|\beta\|^2 - N)$$

(3.7.11)

Bunun α ve β nın normu üzerinde quadratik bir amaç olarak hareket ettiği görülebilir. Sonuçlanan denklemler

$$K_x K_y \beta = \lambda({K_x}^2 + \eta I)\alpha$$

$$K_y K_x \alpha = \lambda({K_y}^2 + \eta I)\beta \quad (3.7.12)$$

dir. İlk probleme benzer, köşegen üzerinde olmayan bloklarda sıfır ve köşegen üzerinde bloklar olarak $({K_x}^2 + \eta I)$ ve $({K_y}^2 + \eta I)$ olan K_D ve köşegen blok olmayan sol altta $K_y K_x$ ve köşegen blok olmayan sağ üstte $K_x K_y$ matrisleri olan K_o büyük matrisleri tanımlansın. Ayrıca, $\gamma = [\alpha, \beta]$ tanımlansın. Bu yine düzenli bir öz değer denklemi olan

$$K_0\gamma=\lambda K_D\gamma \Rightarrow {K_D}^{-1}K_0\gamma=\lambda\gamma \Rightarrow {K_0}^{1/2}{K_D}^{-1}{K_0}^{1/2}({K_0}^{1/2}\gamma)=\lambda({K_0}^{1/2}\gamma) \quad (3.7.13)$$

denklemine neden olur. Düzenleme de sıfırdan başka en küçük öz değer etkilenir ve bu nedenle tersi sayısal olarak daha sabit olur. η için değer çapraz geçerlilik ölçütü ve başka ölçümler kullanılarak seçilmesi gerekir. Bu daha büyük öz değer problemini kullanan denklemleri çözmek çok gerekli değildir ve daha etkin yöntemler vardır.

Çözümlerin dağınık olması beklenmez çünkü öz vektörlerin dağınık olması beklenmez. Aralıklı olmayı elde etmek için L_1 normu L_2 normunun yerini almalıdır. (Welling, M.)

3.8. Kanonik Korelasyon Analizi İle İlgili Bir Uygulama

Standartlaştırılmış ölçüm değerleri için, *Y* değişken kümesi ile *X* değişken kümesine ait korelasyon matris

$$R = \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix}$$

$$= \begin{bmatrix} 1.000 & 0.504 & 0.464 & 0.367 & 0.448 & 0.412 & 0.312 & 0.456 & 0.480 & 0.455 & 0.371 \\ 0.504 & 1.000 & 0.863 & 0.753 & 0.809 & 0.828 & 0.711 & 0.839 & 0.593 & 0.657 & 0.633 \\ 0.464 & 0.863 & 1.000 & 0.769 & 0.791 & 0.826 & 0.716 & 0.836 & 0.635 & 0.742 & 0.659 \\ 0.367 & 0.753 & 0.769 & 1.000 & 0.833 & 0.739 & 0.529 & 0.739 & 0.470 & 0.629 & 0.634 \\ 0.448 & 0.809 & 0.791 & 0.833 & 1.000 & 0.803 & 0.550 & 0.806 & 0.498 & 0.692 & 0.597 \\ 0.412 & 0.828 & 0.826 & 0.739 & 0.803 & 1.000 & 0.682 & 0.871 & 0.575 & 0.644 & 0.569 \\ 0.312 & 0.711 & 0.716 & 0.529 & 0.550 & 0.682 & 1.000 & 0.762 & 0.546 & 0.552 & 0.437 \\ 0.456 & 0.839 & 0.836 & 0.739 & 0.806 & 0.871 & 0.762 & 1.000 & 0.600 & 0.680 & 0.602 \\ 0.480 & 0.593 & 0.635 & 0.470 & 0.498 & 0.575 & 0.546 & 0.600 & 1.000 & 0.746 & 0.575 \\ 0.455 & 0.657 & 0.742 & 0.629 & 0.692 & 0.644 & 0.552 & 0.680 & 0.746 & 1.000 & 0.794 \\ 0.371 & 0.633 & 0.659 & 0.634 & 0.597 & 0.569 & 0.437 & 0.602 & 0.575 & 0.794 & 1.000 \end{bmatrix}$$

olarak verilsin. Y ve X değişken kümesinin çözüm matrisleri,

$$P_1 = R_{11}^{-1} R_{12} R_{22}^{-1} R_{21}$$

$$P_2 = R_{22}^{-1} R_{21} R_{11}^{-1} R_{12}$$

olarak hesaplanmaktadır. Buna göre, korelasyon matrisleri P_1 çözüm matrisinde yerine konur ise;

$$P_1 = \begin{bmatrix} 1.000 & 0.504 & 0.464 \\ 0.504 & 1.000 & 0.863 \\ 0.464 & 0.863 & 1.000 \end{bmatrix}^{-1} \begin{bmatrix} 0.367 & 0.448 & 0.412 & 0.312 & 0.456 & 0.480 & 0.455 & 0.371 \\ 0.753 & 0.809 & 0.828 & 0.711 & 0.839 & 0.593 & 0.657 & 0.633 \\ 0.769 & 0.791 & 0.826 & 0.716 & 0.836 & 0.635 & 0.742 & 0.659 \end{bmatrix}$$

$$\begin{bmatrix} 1.000 & 0.833 & 0.739 & 0.529 & 0.739 & 0.470 & 0.629 & 0.634 \\ 0.833 & 1.000 & 0.803 & 0.550 & 0.806 & 0.498 & 0.692 & 0.597 \\ 0.739 & 0.803 & 1.000 & 0.682 & 0.871 & 0.575 & 0.644 & 0.569 \\ 0.529 & 0.550 & 0.682 & 1.000 & 0.762 & 0.546 & 0.552 & 0.437 \\ 0.739 & 0.806 & 0.871 & 0.762 & 1.000 & 0.600 & 0.680 & 0.602 \\ 0.470 & 0.498 & 0.575 & 0.546 & 0.600 & 1.000 & 0.746 & 0.575 \\ 0.629 & 0.692 & 0.644 & 0.552 & 0.680 & 0.746 & 1.000 & 0.794 \\ 0.634 & 0.597 & 0.569 & 0.437 & 0.602 & 0.575 & 0.794 & 1.000 \end{bmatrix}^{-1} \begin{bmatrix} 0.367 & 0.753 & 0.769 \\ 0.448 & 0.809 & 0.791 \\ 0.412 & 0.828 & 0.826 \\ 0.312 & 0.711 & 0.716 \\ 0.456 & 0.839 & 0.836 \\ 0.480 & 0.593 & 0.635 \\ 0.455 & 0.657 & 0.742 \\ 0.371 & 0.633 & 0.659 \end{bmatrix}$$

$$= \begin{bmatrix} 0.109 & 0.003 & 0.017 \\ 0.155 & 0.469 & 0.365 \\ 0.246 & 0.393 & 0.494 \end{bmatrix}$$

Bu matrisin öz vektörleri, α vektörlerini vermektedir. Öz vektörlerin hesaplanabilmesi için, P_1 matrisinin öz değerleri hesaplanıp eşitlikte yerine konulması gerekmektedir. $P_1 - \lambda I = 0$ yardımıyla öz değerler (λ değerleri) dolayısı ile ρ değerleri hesaplanabilir.

$$\left\| \begin{bmatrix} 0.109 & 0.003 & 0.017 \\ 0.155 & 0.469 & 0.365 \\ 0.246 & 0.393 & 0.494 \end{bmatrix} - \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix} \right\| = 0$$

Buradan

$\lambda_1 = 0.866$; $\lambda_2 = 0.127$; $\lambda_3 = 0.080$ olarak hesaplanmaktadır. Bu üç özdeğerin karekökleri kanonik değişken çiftleri arasından hesaplanan kanonik korelasyon katsayılarına eşit olmaktadır. Buna göre, üç öz değere karşılık gelen öz vektörler ise, $(P_1 - \lambda_i I)e_i = 0$ eşitliğinden hesaplanmaktadır. Eşitlikte, λ değerleri yerine konularak öz vektörler hesaplanabilmektedir. Buna göre,

$\lambda_1 = 0.866$ için e_1 özvektörü $e_1' = (0.020 \quad 0.681 \quad 0.732)$, $\lambda_2 = 0.127$ için e_2 öz vektörü $e_2' = (-0.403 \quad 0.746 \quad -0.530)$, $\lambda_3 = 0.080$ için e_3 öz vektörü $e_3' = (0.373 \quad 0.552 \quad -0.746)$ olarak hesaplanmıştır

v kanonik değişkenlerine ait kanonik ağırlıkları bulmak için, korelasyon matrisleri P_2 çözüm matrisine ait eşitlikte yerine konur ise;

$$P_2 = \begin{bmatrix} 1.000 & 0.833 & 0.739 & 0.529 & 0.739 & 0.470 & 0.629 & 0.634 \\ 0.833 & 1.000 & 0.803 & 0.550 & 0.806 & 0.498 & 0.692 & 0.597 \\ 0.739 & 0.803 & 1.000 & 0.682 & 0.871 & 0.575 & 0.644 & 0.569 \\ 0.529 & 0.550 & 0.682 & 1.000 & 0.762 & 0.546 & 0.552 & 0.437 \\ 0.739 & 0.806 & 0.871 & 0.762 & 1.000 & 0.600 & 0.680 & 0.602 \\ 0.470 & 0.498 & 0.575 & 0.546 & 0.600 & 1.000 & 0.746 & 0.575 \\ 0.629 & 0.692 & 0.644 & 0.552 & 0.680 & 0.746 & 1.000 & 0.794 \\ 0.634 & 0.597 & 0.569 & 0.437 & 0.602 & 0.575 & 0.794 & 1.000 \end{bmatrix}^{-1} \begin{bmatrix} 0.367 & 0.753 & 0.769 \\ 0.448 & 0.809 & 0.791 \\ 0.412 & 0.828 & 0.826 \\ 0.312 & 0.711 & 0.716 \\ 0.456 & 0.839 & 0.836 \\ 0.480 & 0.593 & 0.635 \\ 0.455 & 0.657 & 0.742 \\ 0.371 & 0.633 & 0.659 \end{bmatrix}$$

$$\begin{bmatrix} 1.000 & 0.504 & 0.464 \\ 0.504 & 1.000 & 0.863 \\ 0.464 & 0.863 & 1.000 \end{bmatrix}^{-1} \begin{bmatrix} 0.367 & 0.448 & 0.412 & 0.312 & 0.456 & 0.480 & 0.455 & 0.371 \\ 0.753 & 0.809 & 0.828 & 0.711 & 0.839 & 0.593 & 0.657 & 0.633 \\ 0.769 & 0.791 & 0.826 & 0.716 & 0.836 & 0.635 & 0.742 & 0.659 \end{bmatrix}$$

$$= \begin{bmatrix} 0.109 & 0.088 & 0.108 & 0.104 & 0.101 & 0.065 & 0.106 & 0.086 \\ 0.161 & 0.207 & 0.192 & 0.153 & 0.202 & 0.138 & 0.112 & 0.137 \\ 0.182 & 0.174 & 0.194 & 0.179 & 0.186 & 0.100 & 0.139 & 0.140 \\ 0.172 & 0.167 & 0.186 & 0.173 & 0.176 & 0.078 & 0.111 & 0.126 \\ 0.092 & 0.112 & 0.103 & 0.078 & 0.114 & 0.120 & 0.114 & 0.093 \\ 0.060 & 0.091 & 0.073 & 0.046 & 0.089 & 0.112 & 0.088 & 0.071 \\ -0.003 & -0.038 & -0.025 & 0.018 & -0.023 & 0.052 & 0.093 & 0.018 \\ 0.097 & 0.113 & 0.114 & 0.099 & 0.112 & 0.046 & 0.044 & 0.071 \end{bmatrix}$$

Bu matrisin öz vektörleri, γ vektörlerini vermektedir. Öz vektörlerin hesaplanabilmesi için, P_2 matrisinin öz değerleri hesaplanıp eşitlikte yerine konulması gerekmektedir. $P_2 - \lambda I = 0$ yardımıyla öz değerler (λ değerleri) dolayısı ile ρ değerleri hesaplanabilir. λ_1= 0.866; λ_2= 0.127 ; λ_3= 0.080; λ_4= 0.001; λ_5= -0.0001 + 0.0003i ;λ_6= = -0.0001 - 0.0003i ; λ_7= -0.0002 ; λ_8= 0.001 olarak hesaplanmaktadır. Çalışmada ele alınan X ve Y değişken kümelerinden en fazla, değişken sayısı en küçük olanın sayısı kadar kanonik değişken çifti elde edileceğinden çözüm matrisinin en büyük üç öz değeri dikkate alınarak öz vektörler hesaplanmaktadır. Bu üç en büyük özdeğerin karekökleri kanonik değişken çiftleri arasından hesaplanan kanonik korelasyon katsayılarına eşit olmaktadır.Buna göre, ilk üç öz değere karşılık gelen öz vektörler ise, $(P_2 - \lambda_i I)f_i = 0$ eşitliğinden hesaplanmaktadır. Eşitlikte λ değerleri yerine konularak öz vektörler hesaplanabilmektedir. Buna göre, $(P_2 - \lambda_1 I)f_1 = 0$ eşitliğinde;

λ_1= 0.866 için,

$f_1' = (0.277 \quad 0.504 \quad 0.492 \quad 0.465 \quad 0.288 \quad 0.215 \quad -0.040 \quad 0.285)$

λ_2= 0.127 için,

$f_2' = (-0.036 \quad 0.264 \quad 0.222 \quad 0.361 \quad -0.224 \quad -0.256 \quad -0.717 \quad 0.346)$

λ_3 = 0.080 için,

$f_3' = (-0.364 \quad 0.545 \quad -0.225 \quad -0.171 \quad 0.208 \quad 0.376 \quad -0.525 \quad 0.173)$

olarak hesaplanmıştır.

u_i ve v_i arasındaki kanonik korelasyon katsayıları sırası ile;

a) $u_1$ve v_1 arasındaki korelasyon I. kanonik korelasyonu

$r_{u_1 v_1} = \sqrt{\lambda_1} = \sqrt{0.866} = 0.93$

b) u_2 ve v_2 arasındaki korelasyon II. kanonik korelasyonu;

$r_{u_2 v_2} = \sqrt{\lambda_2} = \sqrt{0.127} = 0.356$

c) u_3 ve v_3 arasındaki korelasyon III. kanonik korelasyonu;

$$\mathrm{r}_{\mathrm{u}_3\mathrm{v}_3} = \sqrt{\lambda_3} = \sqrt{0.080} = 0.283$$

olarak hesaplanır.

Kanonik korelasyonun önem kontrolleri ise şu şekilde yapılır:

$$\mathrm{H}_0 = \rho_1 = \rho_2 = 0$$

$$\mathrm{H}_1 = \rho_1 \neq \rho_2 \neq 0$$

Λ=$\prod_{i=1}^{p}(1-\rho_i^2)$ olmak üzere

$$\chi_{\mathrm{h}}^2 = -[(n-1)-(p+q+1)/2]\log_{\mathrm{e}}(\Lambda)$$

$$=-[(86-1)-(3+8+1)/2]\log_{\mathrm{e}}[(1-0{,}93^2)(1-0{,}356^2)(1-0{,}283^2]$$

$$=(-79).log_e(0{,}108)$$

$$=(-79).(-2{,}229)$$

$$=170{,}10$$

$\chi^2_{\mathrm{pq},1-\alpha} = \chi^2_{24,1-0,01} = 42{,}98$ ise,170.10 > 42.98 olduğundan H_0 red edilir. Böylece, I. kanonik korelasyonun önemli olduğuna karar verilir.

II. Kanonik korelasyon değerinin önemli olup olmadığına bakılabilmesi için I. kanonik korelasyon değeri eşitlikten çıkartılır.

$$\mathrm{H}_0 = \rho_2 = 0$$

$$\mathrm{H}_1 = \rho_2 \neq 0$$

Λ=$\prod_{i=1}^{p}(1-\rho_i^2)$ olmak üzere,

$$\chi_{\mathrm{h}}^2 =-[(86-1)-(3+8+1)/2]\log_{\mathrm{e}}[(1-0{,}356^2)(1-0{,}283^2]$$

$$=(-79).log_e(0{,}803)$$

$$=17{,}32$$

$\chi^2_{(\mathrm{p}-1)(\mathrm{q}-1),1-\alpha} = \chi^2_{14,1-0,01} = 29{,}14$ $i$se, 17.32 < 29.14 olduğundan H_0 kabul edilir. Böylece, II. kanonik korelasyonun önemsiz olduğuna karar verilir. Dolayısı ile üçüncü kanonik korelasyon katsayısının önem kontrolü yapılmamaktadır. Bundan dolayı, birinci kanonik değişkenlerin bulunması (u_1 ve v_1) ve bu değişkenler ile orijinal değişkenler arasındaki korelasyona bakılması ve hesaplanan kanonik değişkenlerin dahil olmadıkları değişken kümelerinde görülen toplam varyasyonun ne kadarını izah edebildiği tespit edilmelidir. Çalışmada birinci kanonik korelasyon katsayısı önemli bulunduğuna göre birinci kanonik değişkenleri için standart kanonik katsayılar ve kanonik yüklerinin belirlenmesi ve buna göre orijinal değişkenler ile kanonik değişkenler arasındaki ilişkinin ortaya konması gerekmektedir. Buna göre standart kanonik katsayılar,

$$\alpha_1 = \frac{e_{11}}{\sqrt{e_1{}'R_{11}e_1}},$$

$$\alpha_1 = \frac{0{,}020}{\sqrt{(0{,}020 \quad 0{,}681 \quad 0{,}732)\begin{bmatrix} 1.000 & 0.504 & 0.464 \\ 0.504 & 1.000 & 0.863 \\ 0.464 & 0.863 & 1.000 \end{bmatrix}\begin{bmatrix} 0.020 \\ 0.681 \\ 0.732 \end{bmatrix}}},$$

$$\alpha_1 = \frac{0{,}020}{\sqrt{1.887}} = 0.014$$

Buna göre diğerleri sırası ile,

$$\alpha_2 = \frac{0{,}681}{1{,}887} = 0{,}496\ ; \qquad \alpha_3 = 0{,}533$$

olarak tespit edilmiştir.

X değişken kümesi için standart kanonik katsayılar;

$$\gamma_1 = \frac{f_{11}}{\sqrt{f_1{}'R_{22}f_1}},$$

$$\gamma_1 = \frac{0.277}{\sqrt{4.387}} = 0.132$$

$\gamma_2 = 0.241$; $\gamma_3 = 0.235$; $\gamma_4 = 0.222$; $\gamma_5 = 0.137$; $\gamma_6 = 0.103$; $\gamma_7 = -0.019$; $\gamma_8 = 0.136$ $o$larak hesaplanmıştır. Buna göre X değişken kümesi için standart kanonik katsayılar vektörü,

$$\gamma_1' = (0.132 \quad 0.241 \quad 0.235 \quad 0.222 \quad 0.137 \quad 0.103 \quad -0.109 \quad 0.136)$$

olarak tespit edilmiştir. Buna göre birinci kanonik değişken çifti,

$$\mathrm{v}_1 = 0.132y_{11} + 0.214y_{12} + 0.235y_{13} + 0.222y_{14} + 0.137y_{15} + 0.103y_{16} - 0.109y_{17} + 0.136y_{18}$$

$$\mathrm{u}_1 = 0.014x_{11} + 0.496x_{12} + 0.532x_{13}$$

şeklinde yazılabilmektedir.

Bu bilgiler doğrultusunda u_1 kanonik değişkeni ile X değişken kümesi arasındaki kanonik yükler (veya korelasyonlar) (3.3.1) bağıntısından, v_1 kanonik değişkeni ile Y değişken kümesi arasındaki korelasyonlar ise (3.3.4) bağıntısından;

$$Kor(u_1, X) = \frac{\begin{bmatrix}0.132 & 0.241 & 0.235 & 0.222 & 0.137 & 0.103 & -0.109 & 0.136\end{bmatrix}\begin{bmatrix}1.000 & 0.833 & 0.739 & 0.529 & 0.739 & 0.470 & 0.629 & 0.634\\ 0.833 & 1.000 & 0.803 & 0.550 & 0.806 & 0.498 & 0.692 & 0.597\\ 0.739 & 0.803 & 1.000 & 0.682 & 0.871 & 0.575 & 0.644 & 0.569\\ 0.529 & 0.550 & 0.682 & 1.000 & 0.762 & 0.546 & 0.552 & 0.437\\ 0.739 & 0.806 & 0.871 & 0.762 & 1.000 & 0.600 & 0.680 & 0.602\\ 0.470 & 0.498 & 0.575 & 0.546 & 0.600 & 1.000 & 0.746 & 0.575\\ 0.629 & 0.692 & 0.644 & 0.552 & 0.680 & 0.746 & 1.000 & 0.794\\ 0.634 & 0.597 & 0.569 & 0.437 & 0.602 & 0.575 & 0.794 & 1.000\end{bmatrix}}{\left[K\ddot{o}\c{s}\begin{bmatrix}1.000 & 0.833 & 0.739 & 0.529 & 0.739 & 0.470 & 0.629 & 0.634\\ 0.833 & 1.000 & 0.803 & 0.550 & 0.806 & 0.498 & 0.692 & 0.597\\ 0.739 & 0.803 & 1.000 & 0.682 & 0.871 & 0.575 & 0.644 & 0.569\\ 0.529 & 0.550 & 0.682 & 1.000 & 0.762 & 0.546 & 0.552 & 0.437\\ 0.739 & 0.806 & 0.871 & 0.762 & 1.000 & 0.600 & 0.680 & 0.602\\ 0.470 & 0.498 & 0.575 & 0.546 & 0.600 & 1.000 & 0.746 & 0.575\\ 0.629 & 0.692 & 0.644 & 0.552 & 0.680 & 0.746 & 1.000 & 0.794\\ 0.634 & 0.597 & 0.569 & 0.437 & 0.602 & 0.575 & 0.794 & 1.000\end{bmatrix}\right]^{\frac{1}{2}}}$$

$$Kor(\mathrm{u}_1, \mathrm{x}) = (0.848 \quad 0.892 \quad 0.921 \quad 0.794 \quad 0.933 \quad 0.687 \quad 0.783 \quad 0.721)$$

$$Kor(v_1, y) = \frac{(0.014 \quad 0.496 \quad 0.532)\begin{bmatrix}1.000 & 0.504 & 0.464\\ 0.504 & 1.000 & 0.863\\ 0.464 & 0.863 & 1.000\end{bmatrix}}{\left[K\ddot{o}\c{s}\begin{bmatrix}1.000 & 0.504 & 0.464\\ 0.504 & 1.000 & 0.863\\ 0.464 & 0.863 & 1.000\end{bmatrix}\right]^{1/2}},$$

$$Kor(v_1, y) = \frac{(0.511 \quad 0.962 \quad 0.967)}{1},$$

$$Kor(v_1, y) = (0.511 \quad 0.962 \quad 0.967)$$

elde edilir. (Çankaya, S., 2005)

3.9. En Küçük Kareler Uygun Değerleri ve Artıkları(Hataları) Arasındaki Kanonik Korelasyonların Özellikleri

Puntanen, S. (1985), (2.1.1) modelinde $E(Y) = X\beta, Kov(Y) = \sigma^2 V$ ve H, X in sütun uzayı üzerine ortogonal izdüşürücü ve $M = I_n - H$ olduğunda, HY uygun değerleri ve MY artıkları arasındaki kanonik korelasyonların özelliklerini inceledi. Bu çalışmada ise, bu makaledeki bilgiler esas alınarak bazı genişletmeler yapıldı.

Şimdi bu modeli $(Y, X\beta, \sigma^2 V\ )$ üçlüsü ile gösterelim. Burada, genel olarak X eksik ranklıdır.

Bu modelde, $X\beta$ nın bilinen en küçük kareler tahmin edicisi$(BEKKT)$,

$$\text{BEKKT } (X\beta) = X\hat{\beta}$$

$$= X(X'X)^{-}X'Y = XX^{+}Y = HY$$

olarak yazılabilir. Burada, $H = XX^{+}$ dır. Eğer AY nin varyans-kovaryans matrisi yansızlık şartına bağlı olarak minimum (negatif kararlı olmama anlamında) ise, yani, $X\beta$ için yansız olan her BY için

$$Kov(AY) \leq Kov(BY)$$

olduğu zaman $X\beta$ nın en iyi yansız lineer tahmin edicisi$(E\dot{I}YLT)$ AY dir.

Gauss-Markov teoreminin temelinde $(Y, X\beta, \sigma^2 I)$ modelinde, $BEKKT(X\beta) = HY = XX^{+}Y$, $X\beta$ nın $E\dot{I}YLT$ si olduğu 3. kısımda görüldü.

Sonuç olarak V pozitif kararlı olmak üzere $(Y, X\beta, V)$ modelinde $X\beta$ nın $E\dot{I}YLT$ si

$$E\dot{I}YLT(X\beta) = X\beta^{*}$$

$$= X(X'V^{-1}X)^{-}X'V^{-1}Y$$

dir.

Bununla beraber, V nin pozitif yarı kararlı olduğu durumda $|V| = 0$ olduğunu göstererek, özel durumlar dışında $X\beta$ nın en iyi lineer yansız tahmin edicisi için açık ifade V^{-1} yerine V^{-} veya V^{+} koyarak yazılamaz. $\mathcal{R}(T) = \mathcal{R}(X, V)$ olacak şekilde

$T = V + XUX'$ ve U simetrik bir matris olmak üzere, bir ifade (örneğin; bakınız Rao,C.,R., 1973b, s.301) V^{-1} yerine T^{-} koymak suretiyle elde edildi. Burada bilindiği gibi $\mathcal{R}(T)$, T nin sütun uzayını gösterir.

V, I nın bir katı olmasa bile, $BEKKT(X\beta\ )$, $EİYLT(X\beta\ )$ ya eşit olacak şekilde V matrislerinin var olabileceği görülür. Problem ilk defa Anderson,T.W. (1948) tarafından ortaya konmuştur ve hala birçok yazar tarafından ele alınmaktadır Bununla beraber $EİYLT$ ile $BEKKT$ karşılaştırıldığında hangisinin iyi veya kötü olduğu sorusunu sormak doğaldır. Bu iyiliği nasıl ölçmeliyiz? Watson,G.S. (1955) tarafından ifade edildiği gibi bunu yapmanın yolu tek değildir.

Watson, G.S. (1955) (X ve V nin tam ranklı olduğunu kabul ederek) $BEKKT$ nin göreceli iyiliğini genelleştirilmiş varyansların oranı ile, yani,

$$\varphi = BEKKT(\beta\ )\ nin\ etkinliği = \frac{\mathrm{detCov}\ (EİYLT(\beta))}{\mathrm{detCov}\ (\mathrm{BEKKT}(\beta))} \tag{3.9.1}$$

ile ölçtü. Buna etkinlik oranı dedi.

Şimdi,

$$\begin{cases} Kov(\hat{\beta}) = (X'X)^{-1}X'VX(X'X)^{-1} \\ \quad Kov(\beta^{*}) = (X'V^{-1}X)^{-1} \end{cases}$$

elde edilir ve böylece etkinliğin değeri

$$\varphi = \frac{|X'V^{-1}X|^{-1}}{|(X'X)^{-1}X'VX(X'X)^{-1}|}$$

$$= \frac{|X'X|^{2}}{|X'VX||X'V^{-1}X|} \tag{3.9.2}$$

dir. $0 < \varphi \leq 1$ olduğu ve $\varphi = 1$ olması için gerek ve yeter şartın $B\mathrm{EKKT}(\beta)$ nın, $EİYLT(\beta)$ ya eşit olması gerektiği gösterilebilir. En son ifade,

$$Kov(\hat{\beta} - \beta^{*}) = Kov(\hat{\beta}) - Kov(\beta^{*})$$

olması gerçeğine dayanır. Gerçekten,

$$Kov(\varepsilon) = Var(\varepsilon) = E(Y - X\beta)(Y - X\beta)' = \sigma^{2}V$$

$$E(YY' - Y\beta'X' - X\beta Y' + X\beta\beta'X') = \sigma^2 V$$

$$E(YY') - E(Y)\beta'X' - X\beta(Y') + X\beta\beta'X' = \sigma^2 V$$

$$E(YY') - X\beta\beta'X' - X\beta\beta'X' + X\beta\beta'X' = \sigma^2 V$$

$$E(YY') - X\beta\beta'X' = \sigma^2 V$$

$$E(YY') = \sigma^2 V + X\beta\beta'X'$$

dir.

$$Kov\big(\hat{\beta} - \beta^*\big) = E\big\{\big(\hat{\beta} - \beta^* - E(\hat{\beta} - \beta^*)\big)\big(\hat{\beta} - \beta^* - E(\hat{\beta} - \beta^*)\big)'\big\}$$

$$= E\big\{\big(\hat{\beta} - \beta^*\big)\big(\hat{\beta} - \beta^*\big)'\big\}$$

$$= E\{((X'X)^{-1}X'Y - (X'V^{-1}X)^{-1}X'V^{-1}Y)((X'X)^{-1}X'Y - (X'V^{-1}X)^{-1}X'V^{-1}Y)'\}$$

$$= E\{((X'X)^{-1}X'Y - (X'V^{-1}X)^{-1}X'V^{-1}Y)(Y'X(X'X)^{-1} - Y'V^{-1}X(X'V^{-1}X)^{-1})\}$$

$$=E\{((X'X)^{-1}X'YY'X(X'X)^{-1} - (X'X)^{-1}X'YY'V^{-1}X(X'V^{-1}X)^{-1} - (X'V^{-1}X)^{-1}X'V^{-1}YY'X(X'X)^{-1} + (X'V^{-1}X)^{-1}X'V^{-1}YY'V^{-1}X(X'V^{-1}X)^{-1})\}$$

$$= (X'X)^{-1}X'E(YY')X(X'X)^{-1} - (X'X)^{-1}X'E(YY')V^{-1}X(X'V^{-1}X)^{-1} - (X'V^{-1}X)^{-1}X'V^{-1}E(YY')X(X'X)^{-1} + (X'V^{-1}X)^{-1}X'V^{-1}E(YY')V^{-1}X(X'V^{-1}X)^{-1}$$

$$= (X'X)^{-1}X'(\sigma^2 V + X\beta\beta'X')X(X'X)^{-1} - (X'X)^{-1}X'(\sigma^2 V + X\beta\beta'X')V^{-1}X(X'V^{-1}X)^{-1} - (X'V^{-1}X)^{-1}X'V^{-1}(\sigma^2 V + X\beta\beta'X')X(X'X)^{-1} + (X'V^{-1}X)^{-1}X'V^{-1}(\sigma^2 V + X\beta\beta'X')V^{-1}X(X'V^{-1}X)^{-1}$$

$$= \sigma^2(X'X)^{-1}X'VX(X'X)^{-1} + (X'X)^{-1}X'X\beta\beta'X'X(X'X)^{-1} - \sigma^2(X'X)^{-1}X'VV^{-1}X(X'V^{-1}X)^{-1} - (X'X)^{-1}X'X\beta\beta'X'V^{-1}X(X'V^{-1}X)^{-1} - \sigma^2(X'V^{-1}X)^{-1}X'V^{-1}VX(X'X)^{-1} - \sigma^2(X'V^{-1}X)^{-1}X'V^{-1}X\beta\beta'X'X(X'X)^{-1} + (X'V^{-1}X)^{-1}X'V^{-1}VV^{-1}X(X'V^{-1}X)^{-1} + (X'V^{-1}X)^{-1}X'V^{-1}X\beta\beta'X'V^{-1}X(X'V^{-1}X)^{-1}$$

$$= \sigma^2(X'X)^{-1}X'VX(X'X)^{-1} + \beta\beta' - (X'V^{-1}X)^{-1} - \beta\beta' - (X'V^{-1}X)^{-1} - \beta\beta' + (X'V^{-1}X)^{-1} + \beta\beta'$$

$$= \sigma^2(X'X)^{-1}X'VX(X'X)^{-1} - (X'V^{-1}X)^{-1}$$

$= Kov(\hat{\beta}) - Kov(\beta^*)$

dır. Bu bir varyans-kovaryans matris olduğundan negatif kararlı olmayan bir matristir. Bu durumda

$$Kov(\hat{\beta} - \beta^*) = Kov(\hat{\beta}) - Kov(\beta^*) \geq 0$$

dır. Bu eşitsizlik $0 < \varphi \leq 1$ olmasını doğrular. Gerçekten, tanımdan, pozitif kararlı bir matrisin determinantı pozitif olduğundan, iki pozitif sayının oranı pozitiftir ve $0 < \varphi$ dır. Determinantlar için Cauchy eşitsizliğinden $\varphi \leq 1$ dir. Bu ifadeyi ispatlamak için, BB' singüler olmayan matris olmak üzere, A, B: nxk matrislerini göz önüne alalım. Bu takdirde, $A'[I - B(B'B)^{-1}B']A$ negatif kararlı değildir veya bilinen bir notasyonda

$$A'A \geq A'B(B'B)^{-1}B'A$$

dır. Bu nedenle

$$|A'A| \geq |A'B(B'B)^{-1}B'A|,$$

yani,

$$|A'B|^2 \leq |A'A||B'B|$$

olduğu görülür. Bu eşitsizlikte $A = V^{1/2}X$, $B = V^{-1/2}X$ alarak φ nin iddia edildiği gibi 1'i aşamadığı aşağıda gösterilmiştir.

$$\left|X'V^{1/2}V^{-1/2}X\right|^2 \leq \left|X'V^{1/2}V^{1/2}X\right|\left|X'V^{-1/2}V^{-1/2}X\right|$$

$$|X'IX|^2 \leq |X'VX||X'V^{-1}X|$$

$$\frac{|X'X|^2}{|X'VX||X'V^{-1}X|} \leq 1 \ .$$

Bartmann, F.C. ve Bloomfield, P. (1981), φ etkinliğinin, $M = I - H$ olmak üzere, HY en küçük kareler uygun değerleri ve $MY = (I - H)Y$ hataları arasındaki kanonik korelasyonların bir fonksiyonu olarak ifade edilebildiğini belirtti. Bundan başka,

$$\rho_i = \text{kan.}\,kor._i\,(HY, MY),$$

yani, ρ_i, HY ve MY arasındaki $i.yinci$ en büyük kanonik korelasyondur. Bu takdirde ($q \leq n/2$ olduğu kabul edilerek)

$$\varphi = \prod_{i=1}^{q}(1-\rho_i^2) \tag{3.9.3}$$

elde edilir.

φ etkinliği (3.9.1) ve (3.9.2) de tanımlanıldığı gibi sadece X ve V tam ranklı; ρ_i kanonik korelasyonlar olduğunda tanımlanmıştır ve böylece (3.9.3) X in rankı veya V nin rankı hesaba katılmaksızın tanımlanmış olur. V pozitif kararlı olduğunda, bu takdirde tüm $\rho_i < 1$ dir, fakat V singüler olduğunda, bir veya daha fazla ρ_i nin 1 e eşitliği mümkündür. Burada, V idempotent yani $V^2 = V$ olduğunda, bu takdirde tüm sıfır olmayan kanonik korelasyonların 1'e eşit olduğu gösterilecektir.

3.10. *V* Pozitif Kararlı(Kesin) Olduğunda Kanonik Korelasyonlar

Watson, G.S. (1967), etkinliğin $E\dot{I}YLT(\beta)$ ve $BEKKT(\beta)$ arasındaki kanonik korelasyonların bir fonksiyonu olacağını gösteren $BEKKT(\beta)$ nın $E\dot{I}YLT(\beta)$ ya bağlı etkinliği için bir gösterim verdi. Daha sonra Bartmann,F.C. ve Bloomfield,P. (1981) HY ve $(I-H)Y$ arasındaki kanonik korelasyonları kullanarak etkinliği daha detaylı olarak ele aldılar. Hem Watson,G.S. (1967) ve hem de Bartmann,F.C. ve Bloomfield,P. (1981), X in tam ranklı ve V nin pozitif kararlı olduğu durumu ele aldılar.

İlerde Bartmann,F.C. ve Bloomfield,P. (1981) makalelerinin kısa bir özeti verilecektir.

X in tam ranklı, $X'X = I$ ve V nin pozitif kararlı olduğunu kabul ederek, $E\dot{I}YLT(\beta)$ ve $BEKKT(\beta)$ nın varyans-kovaryans matrisleri aşağıda verilmiştir:

$$\begin{cases} Kov(\hat{\beta}) = X'VX = Kov\, B\text{EKKT}(\beta) \\ Kov(\beta^*) = (X'V^{-1}X)^{-1} = Kov\, E\dot{I}YLT(\beta) \end{cases}$$

Rao, C.R. (1967), eğer Z

$$\mathcal{R}(X)^{\perp} = \mathcal{R}(Z), \quad Z'Z = I_{n-q} \tag{3.10.1}$$

olarak tanımlanırsa, bu takdirde aşağıdaki eşitliğin ($X'X = I_q$ olduğunu da kabul ederek) sağlandığını gösterdi.

$$(X'V^{-1}X)^{-1}X'V^{-1} = X' - X'VZ(Z'VZ)^{-1}Z' \tag{3.10.2}$$

Burada $\mathcal{R}(X)^{\perp}$, $\mathcal{R}(X)$ in ortogonal tümleyenini gösterir.

(3.10.2) yi sondan VX ile çarparak,

$$(X'V^{-1}X)^{-1} = X'VX - X'VZ(Z'VZ)^{-1}Z'VX \tag{3.10.3}$$

bulunur. Bu eşitlik

$$Kov(\beta^{*}) = Kov(\hat{\beta}) - X'VZ(Z'VZ)^{-1}Z'VX \tag{3.10.4}$$

olduğunu ifade eder.

(3.10.4) deki iki varyans-kovaryans matrisi arasındaki eşitliğin sağlanması için gerek ve yeter şart (3.10.4) ün sağ tarafındaki son terimin sıfıra eşit olmasıdır. Bu

$$Kov(\beta^{*}) = Kov(\hat{\beta}) \Leftrightarrow X'VZ = 0 \tag{3.10.5}$$

olduğunu ifade eder.

$H = XX'$ ve $M = ZZ' = I - H$ ve böylece Marsaglia, G. ve Styan, G.P.H. (1974) rank kısaltma kurallarını kullanarak

$$X'VZ = 0 \Leftrightarrow HVM = 0 \tag{3.10.6}$$

bağıntısını elde etmişlerdir. Bu nedenle, $EİYLT(\beta)$ ve $BEKKT(\beta)$ nın eşitliği için Zyskind,G. (1967) ve Rao nun şartı (Alalouf,I.S.ve Styan G.P.H. 1983) bulunur.

$$Kov(HY, MY) = HVM \tag{3.10.7}$$

olduğuna ve bu nedenle $\hat{\beta} = \beta^{*}$ olması için gerek ve yeter şartın HY uygun değerlerinin MY artıkları ile ilişkisiz olması gerektiğine dikkat edelim.

Yukarıdaki gözlemlerden sonra Bartmann,F.C. ve Bloomfield,P. (1981), iki varyans-kovaryans matris arasındaki farkı, yani

$$X'VX - (X'V^{-1}X)^{-1} \tag{3.10.8}$$

incelediler.

Onlar: “Sadece etkinliğin koordinat serbest ölçümlerinin $(X'V^{-1}X)^{-1}$ in özdeğerlerinin $X'VX$ e göre fonksiyonları olduğunu” söylediler. Bu ifade aşağıdaki nedene dayanır.

Etkinliğin bir koordinat serbest ölçümü ile, modelin singüler olmayan yeniden parametrelenmesi altında değişmeyen $Kov(\hat{\beta})$ ve $Kov(\beta^*)$ varyans-kovaryans matrislerinin böyle bir fonksiyonu olduğu ifade edilir. Yani, eğer A singüler olmayan bir matris ise, bu takdirde $etkinlik(T,U)$ notasyonu T ve U varyans-kovaryans matrislerine bağlı bir etkinlik fonksiyonunu ifade etmek üzere etkinliğin değerini

$$etkinlik(Kov(\hat{\beta});\ Kov(\beta^*) = etkinlik(Kov(A\hat{\beta})\ ; Kov(A\beta^*) \qquad (3.10.9)$$

şartı sağlaması istenilir. Çünkü $Kov(\hat{\beta})$ ve $Kov(\beta^*)$ pozitif kararlıdırlar,

$$\begin{cases} Kov(A\hat{\beta}) = I_q \\ Kov(A\beta^*) = köşegen(\theta_1, \theta_2, \dots, \theta_q) \end{cases} \qquad (3.10.10)$$

yani, $Kov(A\beta^*)$, $\theta_1, \theta_2, \dots, \theta_q$ köşegen elemanları ile bir köşegen matris olacak şekilde bir A matrisi vardır. Bu nedenle,

$$etkinlik(Kov(\hat{\beta})\ ;Kov(\beta^*)) = etkinlik(I_q\ ;köşegen(\theta_1, \theta_2, \dots, \theta_q)) \qquad (3.10.11)$$

sonucunu, yani, eğer $Kov(\hat{\beta})$ ve $Kov(\beta^*)$ ın bir fonksiyonu olan etkinliğin bir koordinat serbest ölçümüne sahip olmak istenirse, bu takdirde o fonksiyonun

$$\left|Kov(\beta^*) - \theta Kov(\hat{\beta})\right| = 0 \qquad (3.10.12)$$

ın köklerinin bir fonksiyonu olması sonucu çıkarılabilir.

Fakat doğal olarak, etkinliğin $Kov(\hat{\beta})$ ve $Kov(\beta^*)$ ın fonksiyonları olmayan bazı sezilebilir ölçümleri olabilir ve bu nedenle bu fonksiyonların (3.10.12) deki θ_i köklerinin fonksiyonları olması gerekmez.

$(X'V^{-1}X)^{-1}$ in $X'VX$ e göre özdeğerleri

$$|(X'V^{-1}X)^{-1} - \theta X'VX| = 0 \qquad (3.10.13)$$

denkleminin kökleridir veya, denk olarak

$$|X'VX - X'VZ(Z'VZ)^{-1}Z'VX - \theta X'VX| = 0 \qquad (3.10.14)$$

$$\Leftrightarrow \left|\, I_q - (X'VX)^{-1/2}X'VZ(Z'VZ)^{-1}Z'VX(X'VX)^{-1/2} - \theta I\right| = 0. \qquad (3.10.15)$$

Ancak, chi(.) karşılık gelen matrisin $i.yinci$ en büyük özdeğerini göstermek üzere

$$chi\left[(X'VX)^{-1/2}X'VZ(Z'VZ)^{-1}Z'VX(X'VX)^{-1/2}\right]$$

$$= kan.kor._i^2\,(X'Y, Z'Y)$$

$$= \rho_i^2 \tag{3.10.16}$$

olduğu biliniyor.(Bakınız Rao,C.,R., 1973b, s.583)

Ayrıca,

$$\prod_{i=1}^{q}(1-\rho_i^2) = \left|I_q - (X'VX)^{-1/2}X'VZ(Z'VZ)^{-1}Z'VX(X'VX)^{-1/2}\right|$$

$$= |X'VX - X'VZ(Z'VZ)^{-1}Z'VX|/|X'VX|$$

$$= |(X'V^{-1}X)|^{-1}/|X'VX|$$

$$= \frac{1}{|(X'V^{-1}X)||X'VX|}$$

$$= etkinlik\big(\hat{\beta}\big) = \varphi \tag{3.10.17}$$

olduğuna dikkat edelim. (3.10.16) dan

$$\text{kan}.kor._\text{i}\,(X'Y, Z'Y) = \rho_\text{i}$$

olduğu gösterildi . $i = 1,2,\ldots,q$ $(q \leq n/2)$ olduğunda,

$$\rho_\text{i} = \text{kan}.kor._\text{i}\,(X'Y, Z'Y)$$

$$= \text{kan}.kor._\text{i}\,(HY, MY) \tag{3.10.18}$$

olduğunu göstermek kolaydır.

3.11. V Singüler Olduğunda Kanonik Korelasyonlar

Bu kısımda, V varyans-kovaryans matrisinin singüler olmasına izin verildiğinde HY ve MY arasındaki kanonik korelasyonların bazı özellikleri ele alınacaktır. (3.10.17) formülünü (tam rank durumu için) göz önüne alarak eğer bir ρ_i, $1'$e eşit ise, bu takdirde diğer ρ_i nin değerleri ne olursa olsun $etkinlik(\beta) = 0$ olduğunu dikkat edilir. Bu nedenle birim kanonik korelasyonların özelliklerini incelemek mantıklı görülmektedir.

$$\begin{cases} u = (HY, MY) arasındaki\ birim\ kanonik\ korelasyonların\ sayısı \\ p = (HY, MY) arasındaki\ sıfır\ olmayan\ kanonik\ korelasyonların\ sayısı \end{cases}$$

$$cov\begin{bmatrix} HY \\ MY \end{bmatrix} = \Sigma = \begin{bmatrix} HVH & HVM \\ MVH & MVM \end{bmatrix} \quad (3.11.1)$$

olsun. Seshadri, V. ve Styan, G.P.H. (1980), $p = rank(HVM)$ nin

$$p = 0 \Leftrightarrow B\text{EKKT}(X\beta) = E\dot{I}YLT(X\beta) \quad (3.11.2)$$

ifade ettiğini gösterdi.

Aşağıdaki Lemma 3.11.1, u nun bir takım gösterimlerini bir araya getirir. İspatta bir çarpımın rankı için genel kural kullanılacak. (Marsaglia, G., Styan, G.P.H., 1974)

$$r(AB) = r(A) - boy\mathcal{R}(A') \cap \mathcal{R}(B)^{\perp}$$

$$= r(B) - boy\mathcal{R}(A')^{\perp} \cap \mathcal{R}(B), \quad (3.11.3)$$

Burada $r(.)$ rankı gösterir. Aşağıdaki gösterimler de daha sonra kullanılacak:

$$r(AB) = r(A', B^{\perp}) - r(B^{\perp}) \quad (3.11.4)$$

$$= r(A', I - B^{+}B) - r(I - B^{+}B) \quad (3.11.5)$$

Lemma 3.11.1: $(Y, X\beta, V)$ modelinde HY ve MY arasındaki birim kanonik korelasyonların sayısı aşağıdaki gibi yazılabilir:

$$u = r(HVH) + r(MVM) - r(\Sigma) \quad (3.11.6)$$

$$= r(VH) + r(VM) - r(V) \quad (3.11.7)$$

$$= boy\mathcal{R}(VH) \cap \mathcal{R}(VM) \quad (3.11.8)$$

$$= r(V) - boy\mathcal{R}(V) \cap \mathcal{R}(M) - boy\mathcal{R}(V) \cap \mathcal{R}(H) \quad (3.11.9)$$

$$= r(VH) - boy\mathcal{R}(V) \cap \mathcal{R}(H) \quad (3.11.10)$$

İspat: Seshadri, V., Styan, G.P.H. ve Rao, C.R. (3.11.6) yı ortaya koyan bir sonucu ifade ve ispat ettiler. (3.11.7) gösterimi

$$r(\Sigma) = r[(H, M)'V(H, M)]$$

$$= r[V(H, M)]$$

$$= r(V) - boy\mathcal{R}(V) \cap \mathcal{R}(H, M)^{\perp}$$

$$= r(V) \qquad (3.11.11)$$

olduğunu göz önüne alarak ortaya koyar. (3.11.11) de (3.11.3) formülü kullanıldı. u için (3.11.8) gösterimi

$$r(\Sigma) = r[V(H, M)]$$

$$= r(VH) + r(VM) - boy\mathcal{R}(VH) \cap \mathcal{R}(VM) \qquad (3.11.12)$$

ifadesini (3.11.6) da yerine koymak suretiyle elde edilir. Bundan başka, (3.11.9) aşağıdaki formüllerin bir sonucudur.

$$r(VH) = r(V) - boy\mathcal{R}(V) \cap \mathcal{R}(M) \qquad (3.11.13)$$

$$r(VM) = r(V) - boy\mathcal{R}(V) \cap \mathcal{R}(H) \qquad (3.11.14)$$

(3.11.10) u elde etmek için (3.11.14) den elde edilen

$$r(V) - r(VM) = boy\mathcal{R}(V) \cap \mathcal{R}(H) \qquad (3.11.15)$$

bağıntısı (3.11.7) de yerine koyulabilir.

u nun üst limitinin $\mathcal{R}(V)$ nin sıfır uzayının boyutuna denk olduğundan bahsedelim. Bu böyledir, çünkü

$$u = r(VH) + r(VM) - r(V)$$

$$\leq r(H) + r(M) - r(V)$$

$$= n - r(V) = boyN(V) = r(V^{\perp}) \qquad (3.11.16)$$

Bu nedenle, eğer V pozitif kararlı ise, birim kanonik korelasyonlar yoktur.

Lemma 3.11.2: $(Y, X\beta, V)$ modelinde HY ve MY arasındaki birim kanonik korelasyonların sayısının sıfır olması için gerek ve yeter şart

$$VV^{+}H = HVV^{+} \qquad (3.11.17)$$

olmasıdır.

İspat: (3.11.7) bağıntısını kullanarak $u = 0$ için şart

$$r(V) = r(VH) + r(VM) \tag{3.11.18}$$

olarak yazılabilir. $V = VH + VM$ olduğundan, (3.11.18),

$$r(VH + VM) = r(VH) + r(VM) \tag{3.11.19}$$

olduğunu ifade eder.

Anderson T.W. ve Styan G.P.H. (1982), (3.11.19) bağıntısının sağlanması için gerek ve yeter şartın

$$MV^{+}VHV = 0 \tag{3.11.20}$$

ile denk olan

$$MV(VH + VM)^{+}HV = MVV^{+}HV = 0 \tag{3.11.21}$$

olması gerektiğini gösterir. $r(VHV) = r(VH)$ olduğundan, (3.11.20) den son V yi iptal edebiliriz, bu nedenle

$$HV^{+}VH = V^{+}VH \tag{3.11.22}$$

elde edilir. Şimdi (3.11.22), (3.11.17) ile denktir.

Gerçekten (3.11.22) nin tanspozu alınırsa,

$$HVV^{+}H = HV^{+}VH = V^{+}VH = VV^{+}H$$

bağıntısı elde edilir. Burada $W = VV^{+}H$ olsun. Şimdi,

$$HW = HVV^{+}$$

denklemini çözelim. Bu denklemin genel çözümü,

$$W = HHVV^{+} + Z - H^{+}HZ$$

dir. Burada Z keyfi bir matristir. $Z = 0$ koyarak özel çözüm alınırsa,

$$W = HVV^{+}$$

elde edilir. Buradan

$$HHVV^{+} = HVV^{+} = VV^{+}H$$

olduğu, yani (3.11.22) nin (3.11.7) e denk olduğu görülür ve böylece Lemma ispatlanır.

Ben-Israel,A. ve Greville,T.N.E (1974) in sonucunu uygulanırsa, Lemma 3.11.2 yi ispatlamanın mümkün olur.

Lemma 3.11.2 bazı ilginç sonuçlara sahiptir. Örneğin, V nin idempotent olduğu durumda

$$V^{+} = V = V^{2}$$

elde edilir ve böylece

$$u = 0 \Leftrightarrow HVV^{+} = VV^{+}H \Leftrightarrow HV = VH \Leftrightarrow HVM = 0 \qquad (3.11.23)$$

dır. (3.11.23) ifadesi, eğer $V^{2} = V$ ise, bu takdirde $u = 0$ olması için gerek ve yeter şartın $B\text{EKKT}(X\beta) = E\dot{I}YLT(X\beta)$ olması gerektiğini ifade eder. Bununla birlikte, genel olarak $u = 0$ şartı $HV = VH$ olduğu anlamına gelmez.

$X\beta$ nın ağırlıklı en küçük kareler tahmin edicisi, AEKKT,

$$A\text{EKKT}(X\beta) = X(X'V^{+}X)^{-}X'V^{+}Y \qquad (3.11.24)$$

ile tanımlanır.

Zyskind, G. ve Martin, F.B. (1969) a göre , $A\text{EKKT}(X\beta)$ nın $E\dot{I}YLT(X\beta)$ ya eşit olması için gerek ve yeter şart

$$\mathcal{R}(X) \subset \mathcal{R}(V) \qquad (3.11.25)$$

olmasıdır. Bunun sağlanması için gerek ve yeter şart ise,

$$VV^{+}H = H \qquad (3.11.26)$$

dır.

(3.11.26) şartı

$$VV^{+}H = HVV^{+}$$

şartını hatırlatır. Bununla beraber

$$VV^{+}H = H \Leftrightarrow VV^{+}H = HVV^{+} \tag{3.11.27}$$

yani,

$$A\text{EKKT}(X\beta) = E\dot{I}YLT(X\beta) \Longrightarrow \text{u} = 0 \tag{3.11.28}$$

olduğu açıktır, fakat özel durumlar dışında tersi doğru değildir. Eğer V idempotent ise, bu takdirde

$$A\text{EKKT}(X\beta) = E\dot{I}YLT(X\beta) \Longrightarrow BEKKT(X\beta) = E\dot{I}YLT(X\beta)$$

ifadesinin yazılabilir. V idempotent ise, bu takdirde $u = 0$ olması için gerek ve yeter şartın $p = 0$ olması gerektiği yukarıda ileri sürüldü. Bu gözlem, doğal olarak, idempotent durumda p ve u arasındaki 'tam eşitliği' garanti etmez. Bununla beraber, aşağıdaki sonuç p ve u arasında bir bağıntıyı gösterir.

Teorem 3.11.1: V idempotent ise, bu takdirde $(Y, X\beta, V)$ modelinde HY ve MY arasındaki tüm korelasyonlar 0 veya 1 dir, yani, $p = u$ dur.

İspat: V nin bir idempotent matris olduğunu kabul edelim. İlk olarak, VM nin rankının

$$\begin{aligned} r(VM) &= r(V \vdots H) - r(H) \\ &= r(V) + r[(I - VV^{+})H] - r(H) \\ &= r(V) + r[(I - V)H] - r(H) \end{aligned} \tag{3.11.29}$$

şeklinde yazılabildiğini göz önüne alalım. (3.11.29)'un

$$u = r(VH) + r(VM) - r(V)$$

formülünde yerine konması

$$u = r(VH) + r[(I - V)H] - r(H) \tag{3.11.30}$$

sonucunu ortaya koyar.

Daha sonra ispatı sonuçlandıracak olan p yi, yani,

$$\begin{aligned} p &= r(HVM) \\ &= r(VH \vdots H) - r(H) \end{aligned}$$

$$= r(VH) + r[(I - VH(VH)^{+})H] - r(H)$$

$$= r(VH) + r[(I - VH(HVH)^{+})HV] - r(H)$$

$$= r(VH) + r[(H - VH(HVH)^{+})HVH] - r(H)$$

$$= r(VH) + r(H - VH) - r(H)$$

$$= r(VH) + r[(I - V)H] - r(H)$$

$$= u \tag{3.11.31}$$

ele alınır.

p ve u için formüller aşağıdaki gibi yazılabildiğinden

$$p = r(VX) - boy\mathcal{R}(VX) \cap \mathcal{R}(X)$$

$$u = r(VX) - boy\mathcal{R}(V) \cap \mathcal{R}(X),$$

p ve u nun eşitliği için bir genel şartın aşağıdaki gibi yazılabilir:

$$p = u \Leftrightarrow boy\mathcal{R}(VX) \cap \mathcal{R}(X) = boy\mathcal{R}(V) \cap \mathcal{R}(X) \tag{3.11.32}$$

$$\Leftrightarrow \mathcal{R}(VX) \cap \mathcal{R}(X) = \mathcal{R}(V) \cap \mathcal{R}(X) \tag{3.11.33}$$

$$\Leftrightarrow r\{[I - VX(VX)^{+}]X\} = r[(I - VV^{+})X]. \tag{3.11.34}$$

SONUÇ VE ÖNERİLER

Bu tezde ele alınan konular temel istatistiksel kavramların birtakımı olup özellikle lineer modeller alanında çalışan araştırmacılara ve bu konuda doktora ve benzeri araştırmalar yapacaklara azda olsa bir katkıda bulunacaktır.

KAYNAKLAR

Akdeniz, F., 1980. "Linear Restrictions on Parametres in Linear Regression Model With Non-Spherical Disturbances", Universty of Ankara, Faculty of science Department of Mathematics Ankara, Turkey.

Akdeniz F., Öztürk F.,**1996.** *Lineer Modeller*. Ankara.

Alalouf, I.S., Styan, G.P.H., 1983. "Characterizations of the conditions fort he ordinary least squares estimator to be besl linear unbiased". Topics in Applied Statistics (Procedings of the Canadian conference in Applied Statistics,Montreal,1981) Mcgill University, Deparment of Mathematics and Statistics, Report No. 83-3

Alpar, R., **2003.** *Uygulamalı Çok Değişkenli İstatistiksel Yöntemlere Giriş 1*. Nobel, Ankara.

Albayrak, A. S., **2006.** *Uygulamalı Çok Değişkenli İstatistik Teknikleri*. Asil Yayın, 1.Baskı.

Anderson, T.W., 1948. "On the theory of testing serial correlation". Skandinavisk Aktuarietidskrift, 31 , 88-116.

Anderson, T.W., Styan, G.P.H., 1982. "Cochran's theorem, rank addivity and tripotent matrices". Statistics and Probability: Essays in Honor of C.R. Rao , G. Kallianpur,P.R. Krishnaiah and J.K. Ghosh, eds., North-Holland, Amsterdam,1-23.

Bach, F., Fukumizu, K. , Gretton, A., 2007. "Statistical Consistency of Kernel Canonical Correlation Analysis". Journal of Machine Learnings Research 8, 361-383.

Barıtçı, İ., **2001.** "Kilis Keçisi Oğlaklarında Doğumda 3 ve 6 Aylık Yaşta Vücut Ölçüleri Arasındaki İlişkilerin Kanonik Korelasyon Metodu İle Araştırılması". Yüksek Lisans Tezi, 48s, Ankara.

Bartmann, F.C. and Bloomfield, P.,1981. "Inefficiency and correlation." Biometrika, 68 , 67-71.

Ben-Israel, A., Greville, T.N.E., 1974. *Generalized Inverses: Theory and Aplications.*Wiley, New York.

Campbell, R.C., **1974.** *Statistics for biologists*. Second Edition, Cambridge University Press.

Çankaya, S., **2005.** "Kanonik Korelasyon Analizi ve Hayvancılıkta Kullanımı". Çukurova Üniversitesi Fen bilimleri Enstitüsü Doktora Tezi, 135s., Adana.

Ekni, M., **1999.** *Lineer Modeller*. Ankara.

Graybill, F. A., **1976.** *Theory and Application of the Linear Model.* Duxbury Press, Boston, Massachusetts.

Graybill, F. A., **1983.** *Matrices with Applications in Statistics*. Second Edition, Wadsworth, Inc., California.

Johnson, R. A., Wichern D. W.,**1982.** *Applied Multivariate Statistical Analysis*. Prentice Hall, Inc, New Jersey.

Johnson, R. A., **Wichern D. W.** ,**2007.** *Applied Multivariate Statistical Analysis*. Sixth Edition, Pearson Educational International.

Magnus, J. R., **1990.** *Matrix Differential Calculus with Applications in Statistics and Econometrics.* John Wiley&Sons Ltd.

Marsaglia, G. ve Styan, G.P.H., 1974. "Equalities and inequalities for rank of matrices." Linear and Multilinear Algebra, 2, 269-292.

Mirtaghizadeh, H., **1990.** "Kanonik Korelasyon Üzerine Bir Deneme". Gazi Üniversitesi Fen Bilimleri Enstitüsü Yüksek Lisans Tezi,52s.,Ankara.

Özel, M. H., **1984.** "Ekonomik Kalkınma ve Eğitim Arasındaki İlişkinin Kanonik Korelasyon Yardımıyla İncelenmesi". Gazi Üniversitesi Fen Bilimleri Enstitüsü Yüksek Lisans Tezi, 39s., Ankara.

Pore, M.D., 1969. “Quadratic Programming with Linear İnequality Constraints”. Yüksek Lisans Tezi.

Pringle, R., Rayner, A., 1971. *Generalized Inverse Matrices with Applications to Statistics*”. Hafner Publishing Co. Newyork.

Puntanen,S., 1985. “Properties of the canonical correlations between the least squares fitted values and the residuals”. Dept. of Math. Sci., University of Tampere , Finland.

Rao, C. R., 1967. “Least squares theory using an estimated dispersion matrix and its application to measurement of signals.” Proc. Fifth Berkeley Symposium on Math. Statist. And Probab., 1, L.Lecam and J. Neyman, eds.,Univ of California Press, Berkeley, 355-372.

Rao, C. R., 1973a. “Unified theory of least squares”.Communications in Statistics, Part A-Theory and Methods1: 1-8.

Rao, C. R., 1973b. *Linear Statistical Inference and Its Applications*. Second Edition, Wiley,Newyork.

Rao, C. R., Toutenburg, H., 1999. *Linear Models: Least Squares and Alternatives*. Second edition, Wadsworth, Inc., California.

Rencher, A.C., Schaalje, B.G., 2007. *Linear Models in Statistics*. Second Edition, A. John Wiley&Sons.Inc., Publication, Hoboken, New Jersey.

Seber, G.A.F., 1977. *Linear Regression Analysis*. John Wiiley&Sons, Inc.

Seshadri, V., Styan, G.P.H., 1980. “Canonical correlations, rank addivity and characterizations of multivariate normality.” Colloquia Math.Soc., Janos Bolyai, vol.21: Analytic Function Methods in Probablity Theory (Debrecen, Hungary, August,1977). Janos Bolyai, Budapest and North Holland, Amsterdam. 331-344.

Şenocak, M., 1990. *Temel Biyoistatistik*. Çağlayan Kitabevi,Beyoğlu,İstanbul.

Tatlıdil, H., 1996. *Uygulamalı Çok Değişkenli İstatistiksel Analiz*, Cem Web Ofset Ltd, 424s., Ankara.

Toutenburg, H.,**1982.** *Prior İnformation in Linear Models*. John Wiiley&Sons.

Watson, G.S.,1955. “Serial correlation in regression analysis”. Biometrika, 42 , 327-341.

Watson, G.S.,1967. “Linear least squares regression.” Ann. Math. Statist., 38, 1679-1699.

Weatherburn, C.E., **1968.** *A First Course in Mathematical Statistics*, Cambridge University.

Welling, M. “Kernel Canonical Correlation Analysis”, Depertment of Computer Science University of Toronto, Canada

Yapar, C., 1979. “Lineer Modellerde Lineer Parametrik Kısıtlamalar Altında Parametre Kestirimleri ve Hipotez Testleri”. Doktora Tezi, K.T.Ü., Trabzon.

Zyskind, G., 1967. “On canonical forms, non-negative covariance matrices and best and simple least squres linear estimatiors in linear models.” Ann. Math. Statist., 38, 1092-1109.

Zyskind, G. And Martin, F.B., 1969. “On best linear estimation and a general Gauss-Markoff theorem in linear models with arbitrary non-negative covariance structure.” SIAM J. Appl. Math., 17, 1190-1202.

ÖZGEÇMİŞ

Adı Soyadı : Zeynep ALBAYRAK

Doğum Yeri : Kırıkkale

Doğum Tarihi : 10.05.1986

Medeni Hali : Bekar

Bildiği Yabancı Diller : İngilizce

Eğitim Durumu

Lise : Mehmet Akif Ersoy Yabancı Dil Ağırlıklı Lise

Lisans : Ondokuz Mayıs Üniversitesi

Yüksek Lisans : Ordu Üniversitesi

İletişim Bilgileri

Adres : Bağlarbaşı Mah. 814.Sok. Timur Cad. Mimar Sinan Yapı Koop. 15/2 Kırıkkale/Merkez

Telefon : 0 554 871 52 86

Elektronik Posta: zeynep_albayrak2000@hotmail.com