T.C.

ORDU ÜNİVERSİTESİ

FEN BİLİMLERİ ENSTİTÜSÜ

LİNEER MODELLERDE BAZI KAVRAMLARIN GEOMETRİK YORUMLARI, PARAMETRE TAHMİNLERİ VE HİPOTEZ TESTLERİ

AYSUN KÖR

Bu tez,

Matematik Anabilim Dalında

Yüksek Lisans

derecesi için hazırlanmıştır.

ORDU 2013

TEZ ONAY

Ordu Üniversitesi Fen Bilimleri Enstitüsü öğrencisi Aysun KÖR tarafından ve Prof. Dr. Cemil YAPAR danışmanlığında hazırlanan "Lineer Modellerde Bazı Kavramların Geometrik Yorumları, Parametre Tahminleri ve Hipotezler Testleri" adlı bu tez, jürimiz tarafından 09/07/2013 tarihinde oy birliği ile Matematik Anabilim Dalında Yüksek Lisans tezi olarak kabul edilmiştir.

Danışman : Prof. Dr. Cemil YAPAR

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

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

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

ONAY:

Bu tezin kabulü, Enstitü Yönetim Kurulu'nun 06/09/2013 tarih ve 2013/236 sayılı kararı ile onaylanmıştır.

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

TEZ BİLDİRİMİ

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

İmza

Aysun KÖR

ÖZET

LİNEER MODELLERDE BAZI KAVRAMLARIN GEOMETRİK YORUMLARI, PARAMETRE TAHMİNLERİ VE HİPOTEZ TESTLERİ

Aysun KÖR

Ordu Üniversitesi

Fen Bilimleri Enstitüsü

Matematik Anabilim Dalı, 2013

Yüksek Lisans Tezi, 93s.

Danışman: Prof. Dr. Cemil YAPAR

Bu çalışma beş bölümden ibarettir. Birinci bölümde lineer modellerde kullanılan matris cebiri ve özellikle matrislerin genelleştirilmiş tersleri ele alınmıştır. İkinci bölümde önceki çalışmalar ele alındı. Üçüncü bölümde lineer modeller ele alındı. Dördüncü bölümde lineer modellerde en küçük karelerin geometrisi, vektör uzayı geometrisi, Gauss-Markov ve FWLT teoremi verildi. Beşinci bölümde en küçük kareler yöntemi, en küçük karelerin geometrisi, sabit-x regresyonunda R^2, lineer modellerde en iyi lineer yansız tahmin edicinin (EİLYTE)'nin bir geometrik görünümü şekillerle beraber ele alındı.

Anahtar Kelimeler: Varyans analizi, Ortogonal izdüşüm, Genelleştirilmiş ters, Rank, Boyut, İdempotent matris, En iyi lineer yansız tahmin edici (EİLYTE), FWLT.

ABSTRACT

GEOMETRİC COMMENTS OF SOME CONCEPTS, PARAMETER ESTİMATİONS AND HYPOTESİS TESTS IN LİNEAR MODELS.

Aysun KÖR

University of Ordu

Institute for Graduate Studies in Science and Technology

Department, 2013

MSc. Thesis, 93p.

Supervisor : Prof. Dr. Cemil YAPAR

This dissertation consists of five chapters. The first chapter covers matrix algebra used in linear models and especially generalized inverses of matrices. The previous studies related to the topic have been given in the second part. In the third chapter, linear models have been discussed and the fourth chapter have been allocated with geometry of least squares in linear models, geometry of vector space, Gauss-Markov theorem and FWLT theorem.

Lastly, least squares method, geometry of least squares, R^2 in the fixed x-regression and a geometric view of best linear unbiased estimator in linear models have been considered together with figures in the fifth chapter.

Key Words: Analysis of variance, Orthogonal projection, Generalized inverse, Rank, Dimension, İdempotent matrix, Best linear unbiased estimator (BLUE), FWLT.

TEŞEKKÜR

Tüm çalışmalarım boyunca her zaman bilgi ve deneyimleriyle yolumu açan değerli hocam Sayın Prof. Dr. Cemil YAPAR' a en samimi duygularım ile teşekkürlerimi sunarım.

Ayrıca, çalışmalarım boyunca desteklerini esirgemeyen Matematik Bölümü öğretim üyeleri Sayın Doç. Dr. Selahattin MADEN' e Sayın Yrd. Doç. Dr. Erdal ÜNLÜYOL hocalarıma en içten şükranlarımı sunuyorum.

İÇİNDEKİLER

ŞEKİLLER ve ÇİZELGELER LİSTESİ

Şekil No: **Sayfa:**

SİMGELER VE KISALTMALAR

$\mathbf{A}^{-1}$	: **A** matrisinin tersi
$\mathbf{A}^{+}$	: **A** ' nın Moore -Penrose g-tersi
$\mathbf{A}^{-}$	: **A** ' nın g-tersi
boy	: boyut
$C(\mathbf{X})$	: **X** 'in sütun uzayı
$Cr(\mathbf{M})$	: **M** 'nin sütun rankı
D	: Olasılık yoğunluk fonksiyonu
$d(\mathbf{X},\mathbf{Y})$	: **X** ile **Y** arasındaki öklid uzaklığı
E	: Beklenen değer
EILYTE	: En İyi Lineer Yansız Tahmin Edici
izd	: izdüşüm
kov	: kovaryans
GKT	: Genel Kareler Toplamı
HKT	: Hata Kareler Toplamı
KO	: Kareler Ortalaması
KT	: Kareler Toplamı
LPF	: Lineer parametrik fonksiyon
LSF	: Lineer sıfır fonksiyonu

$N(\mathbf{M})$	: $\mathbf{M}$ 'nin sıfır uzayı
$\mathbf{P}_{\mathbf{X}}$	: $\mathbf{X}$ 'in sütun uzayı üzerine ortogonal izdüşüm
RKT	: Regresyon Kareler Toplamı
$R(\mathbf{X})$	: $\mathbf{X}$ 'in satır uzayı
Sd	: Serbestlik Derecesi
S	: Geren
var	: varyans
$\mathbf{V}^{\perp}$	: $\mathbf{V}$ 'nin ortogonal tümleyeni
$\hat{\mathbf{Y}}$	: $\mathbf{Y}$ 'nin tahmin edilen değeri
μ	: Ortalama
Ω	: Örneklem uzay
$\bar{\mathbf{X}}$	: Ortalamalardan Sapmalar
$\mathbf{1}$	: boyutu $n \times 1$ olan tüm elemanları 1 olan bir sütun matris
$\langle \mathbf{X}, \mathbf{Y} \rangle$	: $\mathbf{X}$ ile $\mathbf{Y}$ 'nin iç çarpımı
$\lVert \mathbf{V} \rVert$	: $\mathbf{V}$ vektörünün normu

1.GİRİŞ

1.1.MATRİS CEBİRİ

Tanım 1.1.1: $\mathbf{A}$ matrisinin sütun uzayı, $R(\mathbf{A})$ ile gösterilir ve $\mathbf{A}$'nın sütun vektörleri tarafından gerilen vektör uzayıdır; yani,

$$R(\mathbf{A}) = \left\{ \mathbf{z} : \mathbf{z} = \mathbf{Ax} = \sum_{i=1}^{p} \mathbf{a}_{(i)} \mathbf{x}_i , \mathbf{x} \in R^p \right\} \subset R^n \qquad (1.1.1)$$

dir. Burada, $\mathbf{a}_{(1)}, \mathbf{a}_{(2)}, ..., \mathbf{a}_{(p)}$ $\mathbf{A}$'nın sütun vektörleridir. $\mathbf{A}$ matrisinin sıfır uzayı $N(\mathbf{A})$ ile gösterilir ve

$$N(\mathbf{A}) = \left\{ \mathbf{x} \in R^p \ ve \ \mathbf{Ax} = \mathbf{0} \right\} \subset R^p \qquad (1.1.2)$$

olarak tanımlanan vektör uzayıdır. Bir $\mathbf{V}$ vektör uzayını geren lineer bağımsız vektörler kümesine $\mathbf{V}$'nin tabanı denir. Bir vektör uzayının birden fazla tabanı olabilir.

Teorem 1.1.1:

i) *boy* $\mathbf{V}$; bir $\mathbf{V}$ vektör uzayının taban vektörlerinin sayısını göstermek üzere

rank ($\mathbf{A}$)= *boy* $R(\mathbf{A})$ dır.

ii) *boy* $R(\mathbf{A})$+*boy* $N(\mathbf{A})$=p

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

iv) $R(\mathbf{AA}') = R(\mathbf{A})$ dır.

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

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

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

Tanım 1.1.2: Eğer $\mathbf{AB}=\mathbf{I}=\mathbf{BA}$ ise, $\mathbf{B}: n\times n$ matrisine $\mathbf{A}$ 'nın tersi (normal tersi) denir. Böyle bir $\mathbf{B}$ varsa $\mathbf{A}^{-1}$ ile gösterilir. $\mathbf{A}^{-1}$ 'in mevcut olması için gerek ve yeter şartın $\mathbf{A}$ ' nın singüler olmaması olduğu kolayca görülür. $\mathbf{A}^{-1}$ mevcut ise, $\mathbf{A}^{-1}\mathbf{A}=\mathbf{AA}^{-1}=\mathbf{I}$ dır.

Tanım 1.1.3: $\{\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_n\}$ n sayıda sütun vektörünün oluşturduğu küme olsun.

$$\sum_{i=1}^{n}\mathbf{a}_i\mathbf{x}_i=\mathbf{0} \qquad (1.1.3)$$

eşitliğini sağlayan en az biri sıfırdan farklı olan $\mathbf{a}_1,\mathbf{a}_2,...,\mathbf{a}_n$ reel sayıları varsa, $\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_n$ vektörleri lineer bağımlıdır. Aksi halde, $\{\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_n\}$ kümesi lineer bağımsızdır. $\mathbf{A}: n\times n$ matrisinin rankı lineer bağımsız satır veya sütunlarının sayısıdır, yani, $R(\mathbf{A})$ 'nın boyutudur. $\mathbf{A}$ 'nın rankı $rank(\mathbf{A})$ ile gösterilir. $n\geq p$ olmak üzere $rank(\mathbf{A})=p$ ise, $\mathbf{A}$ 'ya tam sütun ranklı ve $rank(\mathbf{A})=n$ ise, $\mathbf{A}$ 'ya tam satır ranklıdır denir.

Tanım 1.1.4: $\mathbf{A}: n\times n$ matrisinin ortogonal (dik) olması için gerek ve yeter şart $\mathbf{A}^{-1}\mathbf{A}=\mathbf{I}$ olmasıdır. Her ortogonal matris bir kare matristir. $\mathbf{A'A}=\mathbf{I}$ olduğundan $\mathbf{A}^{-1}=\mathbf{A'}$ dür. Ortogonal 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 1999).

Tanım 1.1.5: $\mathbf{x}$ ve $\mathbf{y}: n\times 1$ boyutlu vektörler olsun.

$$\mathbf{x'y}=\sum\mathbf{x}_i\mathbf{y}_i=\mathbf{0} \qquad (1.1.4)$$

ise, $\mathbf{x}$ ve $\mathbf{y}$ vektörleri ortogonaldirler.

Tanım 1.1.6: $\mathbf{A}: n\times n$ matrisinin sütun vektörleri ortonormal vektörler kümesi oluşturuyorsa, $\mathbf{A}$ matrisi ortogonaldir.

Tanım 1.1.7: Eğer $\mathbf{AA}=\mathbf{A}$ ise, $\mathbf{A}: n \times n$ matrisine idempotent matris denir. $\mathbf{A}^2=\mathbf{0}$ ise, $\mathbf{A}$ matrisi nilpotent, $\mathbf{A}^2=\mathbf{I}$ ise, $\mathbf{A}$ matrisi unipotent matris olarak isimlendirilir.

Tanım 1.1.8: $\mathbf{A}: m \times n$ matrisi için, eğer aşağıdaki dört şartı sağlayan ve $\mathbf{A}^+$ ile gösterilen bir matris varsa, $\mathbf{A}^+$, $\mathbf{A}$ 'nın bir Moore -Penrose genelleştirilmiş tersi (Moore -Penrose g-tersi) olarak tanımlanacaktır.

i) $\mathbf{AA^+A}=\mathbf{A}$

ii) $\mathbf{A^+AA^+}=\mathbf{A^+}$

iii) $\mathbf{AA^+}$ simetriktir.

iv) $\mathbf{A^+A}$ simetriktir. (1.1.5)

Teorem 1.1.1: Her $\mathbf{A}$ matrisi için (1.1.5) bağıntılarını sağlayan bir tek $\mathbf{A}^+$ matrisi vardır, yani her matris bir Moore -Penrose g-terse sahiptir.

Teorem 1.1.2: Herhangi bir $\mathbf{A}: m \times n$ matrisi ve $\mathbf{AA^-A}=\mathbf{A}$ şartını sağlayan herhangi bir $\mathbf{A}^-: n \times m$ g-tersi için,

i) $\mathbf{A^-A}$ ve $\mathbf{AA^-}$ idempotenttir.

ii) $rank(\mathbf{A})=rank(\mathbf{AA^-})=rank(\mathbf{A^-A})$ dir.

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

Tanım 1.1.9: Eğer $\mathbf{X}=\begin{bmatrix}\mathbf{X}_1\\ \mathbf{X}_2\\ \vdots\\ \mathbf{X}_n\end{bmatrix}$ bir rasgele vektör ise, $\mathbf{E(X)}=\boldsymbol{\mu}=\begin{bmatrix}\mu_1\\ \mu_2\\ \vdots\\ \mu_n\end{bmatrix}=\begin{bmatrix}\mathbf{E(X_1)}\\ \mathbf{E(X_2)}\\ \vdots\\ \mathbf{E(X_n)}\end{bmatrix}$ olmak üzere, $\mathbf{X}$ 'in varyansını

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

olarak tanımlarız. Eğer $\mathbf{X}$ ve $\mathbf{Y}$ bir ortak olasılık yoğunluk fonksiyonuna sahip iki rasgele değişken ise, $\mathbf{X}$ ve $\mathbf{Y}$ arasındaki kovaryans

$$Kov(\mathbf{X},\mathbf{Y}) = E\left[(\mathbf{X}-E(\mathbf{X}))(\mathbf{Y}-E(\mathbf{Y}))\right]$$

olarak tanımlanır. Eğer $Kov(\mathbf{X},\mathbf{Y}) = 0$ ise, $\mathbf{X}$ ve $\mathbf{Y}$ istatistiksel olarak ilişkisizdir (bağımsızdır).

Teorem 1.1.3: $\mathbf{X}$ ve $\mathbf{Y}: n\times 1$ rasgele vektörler ve $\mathbf{Z}: m\times 1$ 'de herhangi bir rasgele vektör olsun. $\mathbf{A}: p\times n$ ve $\mathbf{B}: q\times m$ reel sayıların matrisleri olsunlar. Bu taktirde,

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

dir.

$Kov(\mathbf{AX},\mathbf{BZ}) = \mathbf{A}Kov(\mathbf{X},\mathbf{Z})\mathbf{B}'$ ve eğer $\mathbf{X}$ ve $\mathbf{Y}$ ilişkisiz iseler,

$Var(\mathbf{X}+\mathbf{Y}) = Var(\mathbf{X}) + Var(\mathbf{Y})$ dir (Magnus 1990).

Teorem 1.1.4: $\mathbf{X}$, $E(\mathbf{X}) - \mu$ ortalamalı ve $Var(\mathbf{X}) = V$ varyanslı $n\times 1$ tipinde rasgele bir vektör olsun. Bu taktirde, $E(\mathbf{X}'\mathbf{AX}) = iz(\mathbf{A}V) + \mu'\mathbf{A}\mu$ dir (Rencher ve Schaalje 2007).

1.2. M^- GENELLEŞTİRİLMİŞ TERSLERİ

Bazen bunlara (pseuda-yalancı) terslerde denir. İlk olarak $\mathbf{M}$' nin tekil olmayan bir kare matris olduğunu farz edelim. Bu taktirde $\mathbf{M}^{-} = \mathbf{M}^{-1}$ dir. Eğer $\mathbf{M}$ tekil ise;

$$\mathbf{M}^{-}\ ;\ \mathbf{M}\mathbf{M}^{-}\mathbf{M} = \mathbf{M} \tag{1.2.1}$$

bağıntısını sağlayan herhangi bir matristir. Bir genelleştirilmiş tersin bu kavramı ters matris kavramının tekil matrislere ilginç bir genişlemesidir. $\mathbf{M}\mathbf{M}^{-}\mathbf{M} = \mathbf{M}$ ' yi sağlayan bir genelleştirilmiş ters daima mevcuttur ve tek değildir. (Not: $\mathbf{M}$ kare olduğunda, $\mathbf{M}^{-}$ 'yi aynı tarzda tanımlamak da mümkündür. Regresyon amaçları için bu genişlemeye ihtiyacımız yoktur.)

Moore-Penrose Ters

$\mathbf{M}^{-}$ üzerinde ısrarla durarak fazladan üç şartı; yani

$$\begin{aligned} \mathbf{M}^{-}\mathbf{M}\mathbf{M}^{-} &= \mathbf{M}^{-} \\ (\mathbf{M}\mathbf{M}^{-})' &= \mathbf{M}\mathbf{M}^{-} \\ (\mathbf{M}^{-}\mathbf{M})' &= \mathbf{M}^{-}\mathbf{M} \end{aligned} \tag{1.2.2}$$

şartlarını sağlayan bir tanım (bu nedenle Moore Penrose ters denen) elde edilebilir. Moore Penrose ters için bazen $\mathbf{M}^{-}$ yazarız.

Bir Genelleştirilmiş Tersi Elde Etme

$\mathbf{M}$' nin bir kare matris olduğunu ve regresyon uygulamaları için, $\mathbf{X'X}$ simetrik formuna sahip olduğunu kabul edelim. $\mathbf{M}$ $p \times p$ boyutlu ve r<p ranklı (satır veya sütun rank) olsun. $\mathbf{M}^{-}$'yi elde etmenin muhtemelen en kolay yöntemi aşağıda verilmiştir.

$\mathbf{M}^{-}$'yi Elde Etmek İçin Bir Yöntem

$\mathbf{M}$ içinde $r \times r$ boyutlu, r ranklı tekil olmayan bir alt matris buluruz. Eğer bu alt matris üst sol köşede yer alıyorsa, bu taktirde $\mathbf{M}_{11}$ tekil olmamalı ve bu nedenle $\mathbf{M}_{11}^{-1}$ mevcut olmak üzere

$$\mathbf{M} = \begin{bmatrix} \mathbf{M}_{11} & \mathbf{M}_{12} \\ \mathbf{M}_{21} & \mathbf{M}_{22} \end{bmatrix} \tag{1.2.3}$$

dir. Bu taktirde $\mathbf{M}^{-}$ $p \times p$ boyutlu bir matris olmak üzere,

$$\mathbf{M}^{-} = \begin{bmatrix} \mathbf{M}_{11}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \tag{1.2.4}$$

$\mathbf{M}$ için bir g-ters olacaktır.

İspat: Üç uygun alt matrise sahip olan

$$\begin{bmatrix} \mathbf{M}_{11} & \mathbf{M}_{12} \\ \mathbf{M}_{21} & \mathbf{M}_{21}\mathbf{M}_{11}^{-1}\mathbf{M}_{12} \end{bmatrix} \tag{1.2.5}$$

matrisini elde etmek için $\mathbf{MM^{-}M}$ çarpımını yapalım. İfadeden anlaşılan varsayımdan dolayı, $\mathbf{M}$' nin son $(p-r)$ tane sütunu ilk r tane sütuna bağlıdır. O halde

$$\mathbf{M}_{11}\boldsymbol{\theta} = \mathbf{M}_{12} \text{ ve } \mathbf{M}_{21}\boldsymbol{\theta} = \mathbf{M}_{22} \tag{1.2.6}$$

olacak şekilde, $\boldsymbol{\theta}$ diyeceğimiz bir $r \times (b-r)$ matrisi vardır. Bunların ilkinden, $\boldsymbol{\theta}$ için çözüm $\boldsymbol{\theta} = \mathbf{M}_{11}^{-1}\mathbf{M}_{12}$ dir. Bu nedenle $\mathbf{M}_{21}\mathbf{M}_{11}^{-1}\mathbf{M}_{12} = \mathbf{M}_{21}\boldsymbol{\theta} = \mathbf{M}_{22}$ dir. $\mathbf{M}_{11}$ tekil olmayan matrisinin elemanları nerede bulunursa bulunsun, bu yöntem kesinlikle aynı şekilde çalışır. Regresyon uygulamalarımız aşağıdadır. Eğer $\mathbf{X'X}$ tekil ise, ve $\mathbf{(X'X)}$ herhangi bir genelleştirilmiş ters ise, $\mathbf{X'Xb = X'Y}$ normal denklemleri $\mathbf{b = (X'X)^{-}X'Y}$ ile sağlanır (Seber 1977). $\mathbf{(X'X)^{-}}$'nin farklı seçimleri farklı $\mathbf{b}$ tahminleri üretse de, geometriye göre $\mathbf{\hat{Y} = Xb}$ ' nın değişmediğine dikkat edelim.

Örnek 1.2.1: Aynı $\mathbf{X = X^*}$ noktasındaki n tane veri noktasının hepsine bir $\mathbf{Y} = \beta_0 + \beta_1\mathbf{X} + \varepsilon$ doğrusunu uydurmayı düşünelim. Yalnız iki veya daha çok $\mathbf{X}$ -sitesi var olduğunda ve burada yalnız birine sahibiz; bir tek çözüm elde edildiğinden en küçük kareler çözümü $(\mathbf{\overline{X}, \overline{Y}}) = (\mathbf{X^*, \overline{Y}})$ noktasından geçen herhangi bir doğrudur. Böylece b_0' ın herhangi bir seçimi için genel çözüm

$$\mathbf{\hat{Y}} = b_0 + (\mathbf{\overline{Y}} - b_0)(\mathbf{X / X^*}) \tag{1.2.7}$$

biçimindedir. $\mathbf{X = X^*}$ olduğunda $\mathbf{\hat{Y} = \overline{Y}}$, yani $\mathbf{\hat{Y}} = (\mathbf{\overline{Y}, \overline{Y}, ..., \overline{Y}})'$ olduğunu ve $\mathbf{\hat{Y}}$ ' nın bir tek olduğuna dikkat edelim. Bunun böyle olmak zorunda olduğunu geometriden biliyoruz. Bu problem için normal denklemler

$$\begin{bmatrix} n & n\mathbf{X}^* \\ n\mathbf{X}^* & n\mathbf{X}^{*2} \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \end{bmatrix} = \begin{bmatrix} \sum \mathbf{Y}_i \\ \mathbf{X}^* \sum \mathbf{Y}_i \end{bmatrix} \tag{1.2.8}$$

dır ve bir tek çözüme sahip değildirler. Şimdi $\mathbf{b = (X'X)^{-}X'Y}$ değeri belirlendiğinde $\mathbf{(X'X)^{-}}$ ' nin özel seçimleriyle neyin başarıldığına bakalım.

Seçim 1:

$$\mathbf{(X'X)} = \begin{bmatrix} n^{-1} & 0 \\ 0 & 0 \end{bmatrix} \tag{1.2.9}$$

olsun. Bu taktirde $b_0 = \mathbf{\overline{Y}}$ ve $b_1 = 0$ dır. $(\mathbf{X^*, \overline{Y}})$ den geçen yatay bir doğru elde ederiz.

Seçim 2:

$$(\mathbf{X'X}) = \begin{bmatrix} 0 & 0 \\ 0 & (n\mathbf{X}^{*2})^{-1} \end{bmatrix} \tag{1.2.10}$$

olsun. Bu taktirde $b_0 = 0$, $b_1 = \overline{\mathbf{Y}} / \mathbf{X}^*$ dır. Orijini $(\mathbf{X}^*, \overline{\mathbf{Y}})$ noktasına birleştiren bir doğru elde ederiz.

Seçim 3:

$$(\mathbf{X'X}) = \begin{bmatrix} 0 & (n\mathbf{X}^{*})^{-1} \\ 0 & 0 \end{bmatrix} \tag{1.2.11}$$

olsun. Bu taktirde $b_0 = \overline{\mathbf{Y}}, b_1 = 0$ dır ki bu, seçim 1 deki aynı çözümdür.

Seçim 4:

$$(\mathbf{X'X}) = \begin{bmatrix} 0 & 0 \\ (n\mathbf{X}^{*})^{-1} & 0 \end{bmatrix} \tag{1.2.12}$$

olsun. Bu taktirde $b_0 = 0$, $b_1 = \overline{\mathbf{Y}} / \mathbf{X}^*$ dır. Seçim 2 deki çözümün aynıdır.

Bu (bir dereceye kadar sınırlı) örnekten iki özelliği not edelim.

1: $(\mathbf{X'X})^{-}$ 'nin herhangi bir özel seçimi sadece (1.2.7) ile gösterilen sonsuz çözümlerden birine götürür (Bu nokta genel durumda da doğrudur).

2: $b_0 = \overline{\mathbf{Y}}$ veya $b_0 = 0$ olduğu yalnız en fazla iki aşikar çözüm ortaya çıkar. Diğer çözümleri (tümü hala (1.2.7) yi sağlayan) elde etmek için $(\mathbf{X'X})^{-}$ 'nin diğer seçimleri gerekir. Bununla beraber, b 'ler hakkında diğer varsayımları yapmak için zaten bize diğer seçimler soruldukça bu noktaya kadarı olanın peşinden gitmek anlamsızdır. Aşikar olarak; hiçbir genelleştirilmiş ters kullanmadan (1.2.7) genel çözümünü doğrudan doğruya tercih edebileceğimiz herhangi bir varsayımı uygulayabilirdik.

Ne yapmalıyız?

Genel tavsiyemiz uygulanabilir bir regresyon problemi için bir genelleştirilmiş tersi kullanma, genellikle bir zaman kaybıdır. Dört değişik seçim:

1) Orijinal veriyi muhafaza ediniz, fakat yeni **X'X** tekil olmayan matrisini oluşturmak için modeli değiştiriniz.
2) Orijinal modeli muhafaza ediniz, fakat yeni **X'X** tekil olmayan matrisi oluşturmak için daha fazla veri elde ediniz.
3) Veri ve modelin her ikisini de muhafaza ediniz ve yeni **X'X** tekil olmayan matrisi oluşturmak için parametreler üzerinde akla uygun lineer sınırlamaları yürürlüğe koymaya karar veriniz. [Bilgisayar programları (bilgisayara verilen sırada) orijinal X matrisinin daha sonraki bağımlı sütunlarına ilişkin tüm parametreleri sıfıra eşitler.]
4) Lineer olmayan kısıtlamalar ekleyiniz ve onlara bağlı en küçük kareler problemini çözünüz. Ridge regresyonu bunun bir örneğidir. 3) seçimi genellikle en uygulanabiliri olacaktır.

2.ÖNCEKİ ÇALIŞMALAR

Saville ve Wood (1991) bazı istatistiksel kavramlara geometrik yaklaşımları ve bu kavramların geometrik yorumlarını geniş bir şekilde ele almıştır. Ekni (1990) ortogonal (dik) matrislerin singüler olmadıklarını ispat etmiştir. Magnus (1990) $\mathbf{X}$ ve $\mathbf{Y}$: $n \times 1$ rasgele vektörlerinin varyansları hakkında bazı sonuçları ortaya koymuştur. Seber (1977) normal denklemlerin çözümünde genelleştirilmiş tersi kullandı ve $\mathbf{b} = (\mathbf{X'X})^{-}\mathbf{X'Y}$ çözümünü elde etti.

Rasgele değişkenlerin vektör uzayında X rasgele vektörünün normu

$$\|\mathbf{X}\| = \sqrt{(\mathbf{X}, \mathbf{X})} = \sqrt{\mathbf{E}(\mathbf{X}^2)}$$

ile verilir. $\mathbf{X}$ ve $\mathbf{Y}$ rasgele vektörleri arasındaki uzaklık ise, $d(\mathbf{X}, \mathbf{Y}) = \|\mathbf{X} - \mathbf{Y}\|$ ile verilir. Grimmet ve Stirzaker (2004) rasgele değişkenler üzerinde skalar çarpımı ortaya koymadan buna benzer bir yaklaşımı kullanırlar.

Bring (1996) bir regresyon modelinde değişkenleri karşılaştırmak için bir geometrik yaklaşım sunmuştur. Çekirdek yumuşatmasını içeren bir kaç ileri istatistiksel yöntemi (Eubank ve Eubank (1999)), Fourier analizi (Bloomfield (2000)) ve dalgacık analizi (Odgen (1997)), bu geometrik yaklaşımın genelleştirmeleri olarak yorumlanabilir. Lineer modellere geometrik yaklaşım, ilk olarak Fisher Mahalanaobis (1964) tarafından önerildi. Lineer istatistiksel modeli nerdeyse tamamen geometrik açıdan Christensen (1996) ve Jammalamadaka ve Sengupta (2003) ele aldılar.

3. GENEL BİLGİLER

3.1.LİNEER MODELLER VE LİNEER MODELLERDE PARAMETRE TAHMİNİ

$\mathbf{Y}$ rasgele değişkenlerin bir $n \times 1$ gözlenebilir vektörü olsun, $\mathbf{X}$ reel sayıların (bilinenlerin, açıklayıcı değişkenlerin) bir $n \times p$ $(n > p)$ matrisi ($\mathbf{X}$'in elemanları rasgele değişkenler olmamak üzere) olsun, $\boldsymbol{\beta}$ bilinmeyen, fakat tahmin edilebilen parametrelerin bir $p \times 1$ vektörü olsun, $E(\boldsymbol{\varepsilon}) = 0$, $Kov(\boldsymbol{\varepsilon}) = \sigma^2 \mathbf{V}$ olmak üzere, $\boldsymbol{\varepsilon}$, rasgele değişkenlerin bir gözlenebilir olmayan $n \times 1$ vektörü (hataların vektörü) olsun, ve bu nicelikler

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{3.1.1}$$

bağıntısı ile bağlanmış olsun.(3.1.1) bağıntısı bir genel lineer modeli tanımlar. $\mathbf{X}$ matrisi, tam sütun ranklı ise, yani $rank(\mathbf{X}_{n \times p}) = p \leq n$ ise, (3.1.1) modeline tam ranklı lineer model, $rank(\mathbf{X}_{n \times p}) = r \leq p$ yani, $\mathbf{X}$ matrisi tam ranklı değilse, (3.1.1) modeline eksik ranklı bir lineer model denir. Ayrıca, $\sigma^2 > 0$ bilinmeyen fakat tahmin edilebilir bir parametredir. Bir lineer model, $\boldsymbol{\varepsilon}$'nun dağılımına, $\mathbf{V}$ varyans-kovaryans matrisine, $\mathbf{X}$'in yapısına ve rankına bağlı olarak ayrı ayrı incelenebilir.

3.2. İZDÜŞÜM KAVRAMI, ORTOGONAL VE EĞİK İZDÜŞÜM

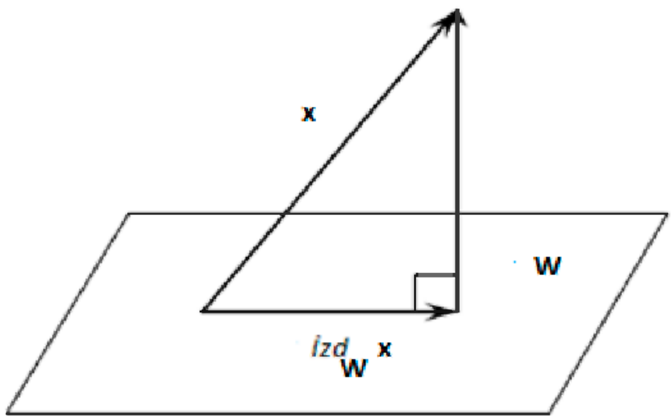
Ortogonal İzdüşüm Hakkında Geometrik Gerçekler

Bu kısımda bir $\mathbf{L}$ öklit uzayını tanıtacağız.

Tanım 3.2.1.a) $\mathbf{X}$'in $\mathbf{L}$ de bir vektör olduğunu $\mathbf{W}$' nun da $\mathbf{L}$' nin bir altuzayı olduğunu farz edelim. Eğer;

1) $\mathbf{Z} \in \mathbf{W}$,

2) $(\mathbf{X}-\mathbf{Z}) \perp \mathbf{W}$ ise $\mathbf{Z}$ vektörüne $\mathbf{X}$'in $\mathbf{W}$ üzerine bir ortogonal izdüşümü denir.



Şekil 3.1. x in w üzerine bir ortogonal izdüşümü

Bu taktirde $\mathbf{Z}$, $\mathbf{izd_w X}$ ile gösterilir. Aşağıdakiler lineer cebirde herkesçe bilinen bir gerçektir.

Teorem 3.2.1: $\mathbf{izd_w X}$ vektörü $\mathbf{W}$ da $\mathbf{X}$' e en yakın vektördür, ve bu vektör bu özelliğe sahip olan yegane vektördür.

Teorem 3.2.2: $\mathbf{v_1, v_2, ..., v_n}$ $\mathbf{W}$ da bir ortogonal taban ise; bu taktirde, $\mathbf{(u, v)}$; $\mathbf{u}$ ve $\mathbf{v}$ vektörlerinin skalar (nokta) çarpımını göstermek üzere,

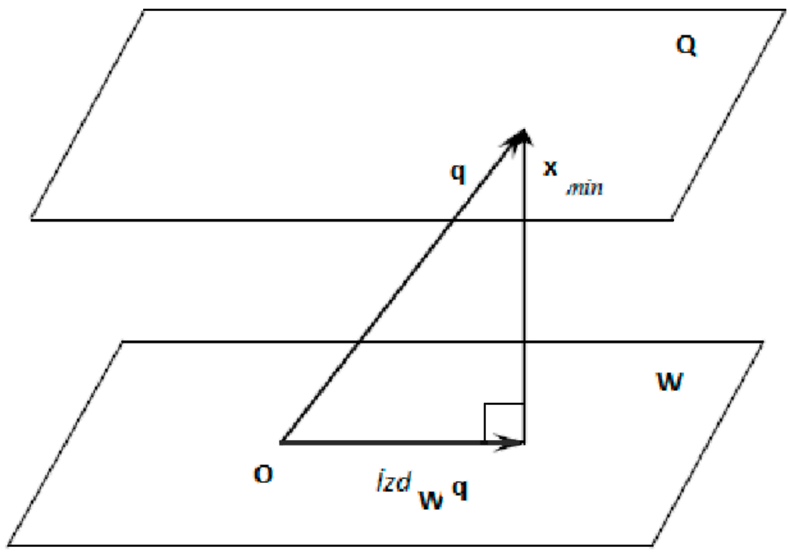
$$\mathbf{izd_w X} = \frac{(\mathbf{x}, \mathbf{v_1})}{(\mathbf{v_1}, \mathbf{v_1})}\mathbf{v_1} + ... + \frac{(\mathbf{x}, \mathbf{v_n})}{(\mathbf{v_n}, \mathbf{v_n})}\mathbf{v_n}$$

dir.

Tanım 3.2.1 b) Eğer $\mathbf{q} \in \mathbf{Q}$ ve $\mathbf{Q} = \{\mathbf{q} + \mathbf{w} \mid \mathbf{w} \in \mathbf{W}\}$ olacak şekilde $\mathbf{B}$'nin bir $\mathbf{W}$ lineer altuzayı varsa, $\mathbf{B}$ lineer uzayının bir $\mathbf{Q}$ alt kümesine, $\mathbf{B}$'nin bir afin altuzayı denir. Bu taktirde $\mathbf{W}$'ya, karşılık gelen lineer altuzay denir. $\mathbf{Q}$ daki herhangi bir vektörün $\mathbf{q}$ olarak alınabileceğini kontrol etmek kolaydır.

Teorem 3.2.3: $\mathbf{Q} = \{\mathbf{q} + \mathbf{w} \mid \mathbf{w} \in \mathbf{W}\}$, $\mathbf{L}$'nin bir afin altuzayı olsun. Bu taktirde $\mathbf{Q}$ da en küçük uzunluğa sahip olan vektör tektir ve $\mathbf{X_{min}} = \mathbf{q} - \mathbf{izd_w q}$ formülü ile verilir.

Teorem 3.2.3' ün ispatı:



Şekil 3.2. $\mathbf{X}_{\mathbf{min}} = \mathbf{q} - \mathbf{Z}$ vektörünün gösterimi

$\mathbf{Z} = \mathbf{izd}_{\mathbf{w}}\mathbf{q}$ olsun. Bu taktirde $\mathbf{X}_{\mathbf{min}} = \mathbf{q} - \mathbf{Z}$ dir. Herhangi bir $\mathbf{Y} \in \mathbf{Q}$ vektörünü düşünelim. Herhangi $\mathbf{w} \in \mathbf{W}$ için, $\mathbf{Y} = \mathbf{q} + \mathbf{w}$ dır.

Teorem 3.2.1' e göre, $\mathbf{Z}$, $\mathbf{W}$ da $\mathbf{q}$ ' ya en yakın vektördür ve $\mathbf{-w} \in \mathbf{W}$ dir. Bu nedenle, $\|\mathbf{Y}\| = \|\mathbf{q} - (\mathbf{-w})\| > \|\mathbf{q} - \mathbf{Z}\| = \|\mathbf{X}_{\mathbf{min}}\|$ elde ederiz. Eşitlik yalnız $\mathbf{-w} = \mathbf{Z}$ yani $\mathbf{Y} = \mathbf{q} + \mathbf{w} = \mathbf{q} - \mathbf{Z} = \mathbf{X}_{\mathbf{min}}$ olduğunda, sağlanır. $\mathbf{X}_{\mathbf{min}}$ tek olduğundan, $\mathbf{X}_{\mathbf{min}}$, $\mathbf{q}$ 'nun seçimine bağlı değildir.

3.3. RASGELE DEĞİŞKENLERE GEOMETRİK YAKLAŞIM

Rasgele Değişkenlerin Vektör Uzayı

Aynı olasılık uzayı üzerindeki rasgele değişkenleri göz önüne alacağız. Sonlu varyanslı tüm rasgele değişkenlerin $\mathbf{H}$ kümesi bir lineer uzaydır (Aşikar olarak, toplama ve bir skalarla (sayıyla) çarpma işlemleriyle). Bir $\mathbf{X}$ rasgele değişkeninin beklenen değerini $\boldsymbol{\mu}_{\mathbf{x}} = E(\mathbf{X})$ ile, $\mathbf{X}$ 'in varyansını $\sigma_{\mathbf{X}}^2 = Var(\mathbf{X})$ ile, ve $\mathbf{X}$ ve $\mathbf{Y}$ rasgele değişkenlerinin kovaryansı da $kov(\mathbf{X},\mathbf{Y})$ ile göstereceğiz.

$(\mathbf{X},\mathbf{Y}) = E(\mathbf{X}.\mathbf{Y})$ ile verilen bir skalar çarpım $\mathbf{H}$ uzayını bir öklit uzayı yapar. Bu uzayda bir $\mathbf{X}$ vektörünün uzunluğu $\|\mathbf{X}\| = \sqrt{(\mathbf{X},\mathbf{X})} = \sqrt{E(\mathbf{X}^2)}$ ile verilir ve $\mathbf{X}$ ve $\mathbf{Y}$ vektörleri arasındaki uzaklık $d(\mathbf{X},\mathbf{Y}) = \|\mathbf{X}-\mathbf{Y}\|$ ile verilir.

Benzer bir yaklaşım Grimmett ve Stirzaker (2004) tarafından kullanılır. Fakat onlar rasgele değişkenler üzerinde skalar çarpım ortaya koymazlar. Halbuki, skalar çarpım ortogonal izdüşümlere çok uygundur ve ispatları daha da kısaltır. Basit durumlarda $\mathbf{H}$ uzayının bir tabanını oluşturabiliriz. Aşağıdaki örnek bunu açıklar.

Örnek 3.3.1: Sonuçların $\mathbf{p}_i = \mathbf{P}(\omega_i) > 0,\ \mathbf{i} = 1,\dots,\mathbf{n}$ olasılıkları ile sonlu bir $\boldsymbol{\Omega} = \{\omega_1, \omega_2, \dots, \omega_n\}$ örneklem uzayını göz önüne alalım. Bu durumda $\mathbf{H}$ öklid uzayında sonlu bir ortogonal taban ortaya koyabiliriz.

Her bir $\boldsymbol{i}$ için bir $\mathbf{F}_i$; rasgele değişkenini aşağıdaki gibi tanımlayalım:

$$\mathbf{F}_i(\omega_i) = \begin{cases} 1, & eğer\ \mathbf{j} = \mathbf{i}\ ise \\ 0, & eğer\ \mathbf{j} \neq \mathbf{i}\ ise \end{cases}$$

olarak tanımlayalım. Bu taktirde $\mathbf{H}$ deki herhangi bir $\mathbf{X}$ rasgele değişkeni için $\mathbf{x}_i = \mathbf{X}(\omega_i)$ olmak üzere,

$$\mathbf{X} = \sum_{i=1}^{n} \mathbf{x}_i \mathbf{F}_i \tag{3.3.1}$$

dir. Herhangi bir $\mathbf{i} \neq \mathbf{j}$ için $\mathbf{F}_i.\mathbf{F}_j = \mathbf{0}$ ve $(\mathbf{F}_i,\mathbf{F}_j) = \mathbf{E}(\mathbf{F}_i.\mathbf{F}_j) = \mathbf{0}$ dır. Bu nedenle,

$$\mathbf{F}_i \perp \mathbf{F}_j \tag{3.3.2}$$

dir. (3.3.1) ve (3.3.2), $\mathbf{F}_j,\dots,\mathbf{F}_n$ 'nin $\mathbf{H}$ de bir ortogonal taban oluşturduğunu ve $\mathbf{H}$' nın boyutunun $\mathbf{n}$ olduğunu ifade eder. $\mathbf{y}_i = \mathbf{Y}(\omega_i)$ olmak üzere $\mathbf{H}$ deki herhangi bir $\mathbf{X},\mathbf{Y}$ için, onların skalar (nokta) çarpımı $(\mathbf{X},\mathbf{Y}) = \sum_{i=1}^{n} \mathbf{p}_i \mathbf{x}_i \mathbf{y}_i$ dir.

Bağımsızlık Kavramları

Yazarlar, farklı şartlar ve çevrede öğrencilerin karşı karşıya geldikleri bağımsızlık ve bağımlılık kavramlarının incelemeye değer olduklarını düşünürler. Olayların bağımsızlığı da vardır ki; onu biz burada kullanmayacağız. Hali hazırdaki konular için, öğrencilerin rasgele değişkenlerin bağımsızlığı ve bağımlılığı arasındaki farkı görmesi yani onların lineer bağıntı ve lineer bağımlılığı arasındaki farkı $\mathbf{H}$'daki vektörler gibi görmesi önemlidir. Kısalık için, yalnız iki değişken durumundaki tanımları hatırlatacağız.

Tanım 3.3.1 a) $\mathbf{X}$ ve $\mathbf{Y}$' nin, $\mathbf{H}$' nin elemanları olduklarını farz edelim.

1)Eğer herhangi $\mathbf{x,y}$ sayıları için $\mathbf{P(X \le x, Y \le y) = P(X \le x).P(Y \le y)}$ ise $\mathbf{X}$ ve $\mathbf{Y}$ rasgele değişkenlerine bağımsızdırlar denir. Aksi taktirde bu rasgele değişkenlere bağımlıdırlar denir.

2)Herhangi $\boldsymbol{\alpha,\beta}$ sayıları için $\mathbf{Y = \boldsymbol{\alpha} + \boldsymbol{\beta} X}$ veya $\mathbf{X = \boldsymbol{\alpha} + \boldsymbol{\beta} Y}$ ise, $\mathbf{X}$ ve $\mathbf{Y}$ rasgele değişkenleri bir lineer bağıntıya sahiptir denecektir. Eşdeğer olarak, lineer bağıntı, her ikisi birlikte sıfır olmayan $\mathbf{a}$ ve $\mathbf{b}$ sayıları ve bir $\mathbf{c}$ sayısının olduğunu ifade eder ki bunlar için $\mathbf{aX + bY = c}$ dir.

3)Eğer $\mathbf{aX + bY = 0}$ olacak şekilde her ikisi birlikte sıfır olmayan $\mathbf{a}$ ve $\mathbf{b}$ sayıları varsa, bir lineer uzaydaki $\mathbf{X}$ ve $\mathbf{Y}$ vektörlerine lineer bağımlıdırlar denir. Aksi taktirde bu vektörlere lineer bağımsızdırlar denir. Aşikar olarak, 3) lineer bağımlılığı 2) lineer ilişkisini, ve 2) lineer bağıntısı 1) lineer bağımlılığını ifade eder.

Örnek 3.3.2: Örnek 3.3.1 de tanımlanan $\mathbf{F_1, F_2, ..., F_n}$ değişkenlerine bakalım. $\mathbf{P(F_1 = 1, F_2 = 1) = 0}$ ve $\mathbf{P(F_1 = 1).P(F_2 = 1) = P(\boldsymbol{\omega}_1).P(\boldsymbol{\omega}_2) = P_1.P_2 \ne 0}$ olduğundan, $\mathbf{F_1}$ ve $\mathbf{F_2}$ rasgele değişkenler olarak, bağımlıdırlar. (Tanım 3.3.1 den)

Sonra da $\mathbf{F_1 + F_2 + ... + F_n = 1}$ olduğundan, $\mathbf{F_1, F_2, ..., F_n}$ rasgele değişkenleri bir lineer bağıntıya (Tanım3.2.1 e göre) sahiptir.Gerçekten, herhangi bir $\boldsymbol{\omega}_i$ için, $\mathbf{(F_1 + ... + F_n)(\boldsymbol{\omega}_i) = F_i(\boldsymbol{\omega}_i) = 1}$ dir.

Son olarak, eğer $\mathbf{a_1F_1 + ... + a_nF_n = 0}$ ise, bu taktirde herhangi bir $\mathbf{i = 1,2,3,...,n}$ için, $\mathbf{0 = (a_1F_1 + a_2F_2 + ... + a_nF_n)(\omega_i) = a_iF_i(\omega_i) = a_i}$ olduğundan $\mathbf{F_1, F_2, ..., F_n}$ vektörler olarak lineer bağımsızdırlar. (Tanım 3.3.1 den)

Öğrencilerin şaşkınlığına ilaveten, istatistikte kesin bir şekilde tanımlanamayan, fakat sezgisel olarak anlaşılan "bağımsız değişken" ve "bağımlı değişken" terimleri de vardır. İstatistikte, rasgele değişkenlerin, onların $\boldsymbol{\rho}_{xy}$ korelasyon (ilişki) katsayısıyla ölçülen lineer bağımlılığı hakkında da söz ederiz. Bu "lineer bağımlılık" terimi asla doğru olarak tanımlanamaz. Fakat teori $\mathbf{X}$ ve $\mathbf{Y}$ arasındaki lineer bağımlılığın $\boldsymbol{\rho}_{xy}$ 'nın $\mathbf{0}$ ' a (sıfıra) yaklaşırken azaldığını ve $\boldsymbol{\rho}_{xy} = \mathbf{0}$ olduğunda mevcut olmadığını ifade eder.

Bu lineer bağımlılık $\mathbf{0}$ beklenen değerine sahip (bekleyen değerli) değişkenler için, enteresan bir geometrik benzerliğe sahiptir. Eğer $\mathbf{X}$ ve $\mathbf{Y}$, $\mathbf{H}$ 'da böyle değişkenler ise, bu taktirde $\mathbf{(X,Y)} = kov\mathbf{(X,Y)}$, $\|\mathbf{X}\| = \sigma_X$, $\|\mathbf{Y}\| = \sigma_Y$ dır ve $\mathbf{X}$ ve $\mathbf{Y}$ arasındaki $\boldsymbol{\theta}$ açısı için

$$\cos\boldsymbol{\theta} = \frac{\mathbf{(X,Y)}}{\|\mathbf{X}\|.\|\mathbf{Y}\|} = \frac{kov\mathbf{(X,Y)}}{\sigma_X\sigma_Y} = \boldsymbol{\rho}_{XY}$$

dir.

Korelasyon katsayısının özelliklerini rasgele değişkenler arasında lineer bağımlılığın bir ölçüsü olarak kullanarak, aşağıdaki gerçekleri ispatlamak kolaydır.

Teorem 3.3.1: Beklenen değerleri $\mathbf{0}$ ve aralarındaki açı $\boldsymbol{\theta}$ olan $\mathbf{X}$ ve $\mathbf{Y}$ rasgele değişkenlerinin $\mathbf{H}$ 'da olduklarını farzedelim. Bu taktirde,

1) $\cos\boldsymbol{\theta} = \boldsymbol{\rho}_{xy}$ dır.

2) $\mathbf{X}$ ve $\mathbf{Y}$ rasgele değişkenlerinin bir lineer bağıntıya sahip olmaları için gerek ve yeter şart $\boldsymbol{\theta}$ açısının 0° ya da 180° olmasıdır.

3) Eğer $\mathbf{X}$ ve $\mathbf{Y}$ bağımsızsalar, bu taktirde onlar ortogonaldırlar.

4) Eğer $\mathbf{X}$ ve $\mathbf{Y}$ ortogonal (dik) vektörler ise, $\mathbf{X}$ ve $\mathbf{Y}$ rasgele değişkenleri arasında lineer bağımlılık yoktur.

5) $\boldsymbol{\theta}$ açısı 90° ye yaklaşırken, $\mathbf{X}$ ve $\mathbf{Y}$ arasındaki lineer bağımlılık azalır.

3.4. REGRESYONUN KİTLE MODELİ VE REGRESYON ÖĞRETİMİNDE GEOMETRİK YAKLAŞIM

Rasgele değişkenlere uygulandığında ve örneklemler olmayan, bir kitleye bağlı olarak ifade edildiğinde, regresyon kavramı açık ve basittir. Regresyon, görünüşte, bir $\mathbf{X}$ rasgele değişkenine göre erişilemez bir $\mathbf{Y}$ rasgele değişkenini tahmin etmeyi, yani, $\mathbf{Y}$ 'ye en yakın bir $\mathbf{f(X)}$ fonksiyonunu bulmayı ifade eder (Hsu 1997).Buna "regresyonun kitle modeli" diyeceğiz. $\mathbf{f(X)}$ en geneli lineer fonksiyonların sınıfı olan, belirli bir fonksiyonlar sınıfına sınırlanabilir "en yakınlığı" daha önce tanımlanan d uzaklığına göre ifade ederiz.

Teorem 3.2.2, $\mathbf{izd_w Y}$ ' nin $\mathbf{W}$ da $\mathbf{Y}$ 'ye en yakın vektör olduğunu ifade eder. Farklı $\mathbf{W}$ ' ları seçerek, basit lineer regresyon, çoklu lineer regresyon, kuadratik regresyon, polinom regresyon vs. gibi farklı regresyon tiplerini elde edebiliriz.

Teorem 3.4.1: $E(\mathbf{Y}/\mathbf{X})$ şartlı beklenen değeri $\mathbf{X}$ 'in $\mathbf{Y}$ 'ye en yakın fonksiyonudur.

Bu aşağıdaki gerçeğe dayanır:

$\mathbf{W}=\{\mathbf{f(X)} \mid \mathbf{f}:\mathbf{R}\rightarrow\mathbf{R}\ \mathbf{ve}\ \mathbf{f(X)}\in\mathbf{H}\}$ için $E(\mathbf{Y}/\mathbf{X}) = \mathbf{izd_w Y}$ dir Grimmet ve Stirzaker (2004), $E(\mathbf{Y}\setminus\mathbf{X})\in\mathbf{W}$ olduğunu ve herhangi bir $\mathbf{h(X)}\in\mathbf{W}$ için, $E[(\mathbf{Y}-E(\mathbf{Y}\setminus\mathbf{X})).\mathbf{h(X)}]=\mathbf{0}$ olduğunu yani, $\mathbf{Y}-E(\mathbf{Y}\setminus\mathbf{X})\perp\mathbf{h(X)}$ olduğunu göstererek, bu gerçeğin doğruluğunu kanıtlar.

Teorem 3.4.2: (Basit lineer regresyon). Eğer $\sigma_{\mathbf{X}}\neq\mathbf{0}$ ise, bu taktirde $\mathbf{X}$ 'in $\mathbf{Y}$ 'ye en yakın lineer fonksiyonu

$$\begin{cases}\boldsymbol{\beta}=\dfrac{Kov(\mathbf{Y},\mathbf{X})}{\sigma_{\mathbf{X}}^2} \\ \boldsymbol{\alpha}=\boldsymbol{\mu}_{\mathbf{Y}}-\boldsymbol{\beta}\boldsymbol{\mu}_{\mathbf{X}}\end{cases} \quad \text{olmak üzere,} \tag{3.4.1}$$

$\boldsymbol{\alpha}+\boldsymbol{\beta}\mathbf{X}$ ile verilir.

Sonuç: Eğer $\hat{\mathbf{Y}} = \boldsymbol{\alpha} + \boldsymbol{\beta}\mathbf{X}$ Teorem 3.4.3'e göre $\mathbf{Y}$ ' nin en iyi lineer tahmin edici ise, bu taktirde $\boldsymbol{\varepsilon} = \mathbf{Y} - \hat{\mathbf{Y}}$ hata tahmini aşağıdaki özelliklere sahiptir.

1) $\boldsymbol{\mu}_{\varepsilon} = \mathbf{0}$, 2) $kov(\boldsymbol{\varepsilon}, \mathbf{X}) = \mathbf{0}$

Bu nedenle, sonuca göre, hata tahminleri $\mathbf{0}$ ortalamaya sahiptir ve $\mathbf{X}$ açıklayıcısı ile ilişkisizdirler; bu, $\hat{\mathbf{Y}}$ ' nın $\mathbf{Y}$ 'nin en iyi lineer tahmini olduğunun başka bir kanıtıdır.

Teorem 3.4.2 ve onun sonucunun geometrik ispatları aşağıda verilmiştir.

Teorem 3.4.2' nin ispatı: $\mathbf{W} = \{\mathbf{a} + \mathbf{bX} \mid \mathbf{a}, \mathbf{b} \in \mathbf{R}\}$ olduğunu gösterelim. $\mathbf{izd}_{\mathbf{w}}\mathbf{Y} \in \mathbf{W}$,bu nedenle bazı $\boldsymbol{\alpha}, \boldsymbol{\beta}$ için $\mathbf{Y} = \boldsymbol{\alpha} + \boldsymbol{\beta}\mathbf{X}$ dir. Şimdi $\boldsymbol{\alpha}$ ve $\boldsymbol{\beta}$ ' nın (3.4.1) formülüyle verildiğini göstermemiz gerekir. $\boldsymbol{\varepsilon} = \mathbf{Y} - \mathbf{izd}_{\mathbf{w}}\mathbf{Y} = \mathbf{Y} - (\boldsymbol{\alpha} + \boldsymbol{\beta}\mathbf{X})$ için, $\mathbf{1}, \mathbf{X} \in \mathbf{W}$ olduğundan $\boldsymbol{\varepsilon} \perp \mathbf{1}$ ve $\boldsymbol{\varepsilon} \perp \mathbf{X}$ olduğu görülür.(Burada $\mathbf{1}$, 1'lerin bir sütun vektörüdür.) Böylece $(\boldsymbol{\varepsilon}, \mathbf{1}) = \mathbf{0}$ ve $(\boldsymbol{\varepsilon}, \mathbf{x}) = \mathbf{0}$, $(\boldsymbol{\alpha} + \boldsymbol{\beta}\mathbf{X}, \mathbf{1}) = (\mathbf{Y}, \mathbf{1})$ ve $(\boldsymbol{\alpha} + \boldsymbol{\beta}\mathbf{X}, \mathbf{X}) = (\mathbf{Y}, \mathbf{X})$ ifadeleri bir

$$\begin{cases} E(\boldsymbol{\alpha} + \boldsymbol{\beta}\mathbf{X}) = E(\mathbf{Y}) \\ E(\boldsymbol{\alpha}\mathbf{X} + \boldsymbol{\beta}\mathbf{X}.\mathbf{X}) = E(\mathbf{Y}.\mathbf{X}) \end{cases} \text{veya} \begin{cases} \boldsymbol{\alpha} + \boldsymbol{\beta}\boldsymbol{\mu}_{X} = \boldsymbol{\mu}_{Y} \\ \boldsymbol{\alpha}\boldsymbol{\mu}_{X} + \boldsymbol{\beta}E(\mathbf{X}^2) = E(\mathbf{Y}.\mathbf{X}) \end{cases}$$

lineer denklemler sistemine götürür. Sistemin çözümü (3.4.1) ile verilir. Teorem 3.4.2'nin aşağıdaki sonucu daha önce de ifade edildi.

Sonuç: Eğer $\hat{\mathbf{Y}} = \boldsymbol{\alpha} + \boldsymbol{\beta}\mathbf{X}$ Teorem 3.4.2 ye göre , $\mathbf{Y}$ ' nin en iyi lineer tahmin edicisi ise, bu taktirde $\boldsymbol{\varepsilon} = \mathbf{Y} - \hat{\mathbf{Y}}$ hatası aşağıdaki özelliklere sahiptir.

1) $\boldsymbol{\mu}_{\varepsilon} = \mathbf{0}$, 2) $kov(\boldsymbol{\varepsilon}, \mathbf{X}) = \mathbf{0}$

Sonucun ispatı: 1) $\boldsymbol{\varepsilon} \perp \mathbf{1}$ bu nedenle $E(\boldsymbol{\varepsilon}) = (\mathbf{1}, \boldsymbol{\varepsilon}) = \mathbf{0}$ dır.

2) $\boldsymbol{\varepsilon} \perp \mathbf{X}$, bu nedenle $E(\boldsymbol{\varepsilon}.\mathbf{X}) = (\boldsymbol{\varepsilon}, \mathbf{X}) = \mathbf{0}$ ve $kov(\boldsymbol{\varepsilon}, \mathbf{X}) = E(\boldsymbol{\varepsilon}.\mathbf{X}) - E(\boldsymbol{\varepsilon}).E(\mathbf{X})$ dir.

Sonuç, regresyonun farklı bir yorumuna götürür. Bu yorum ortogonal izdüşüm kavramıyla aşina olmayan öğrencilerin eğitiminde faydalıdır. $\mathbf{Y}$ rasgele değişkeni iki kısma bölünür.

$\mathbf{Y = f(X) + \varepsilon}$ ve $\boldsymbol{\mu}_\varepsilon = \mathbf{0}$ ve $kov(\boldsymbol{\varepsilon}, \mathbf{X}) = \mathbf{0}$ olması istenir. $\mathbf{f(X)}$ kısmına $\mathbf{Y}$ ' nin regresyon tahmini denir.

Teorem 3.4.3: $\mathbf{X}$ ' in tüm lineer fonksiyonlarının sınıfında $\mathbf{Y}$ ' nin regresyon tahmini 3.4.1 formülüyle verilir.

İspat: $\mathbf{Y = \boldsymbol{\alpha} + \boldsymbol{\beta} X + \boldsymbol{\varepsilon}}.kov(\mathbf{Y,X}) = kov(\boldsymbol{\alpha}, \mathbf{X}) + \boldsymbol{\beta}.kov(\mathbf{X,X}) + kov(\boldsymbol{\varepsilon}, \mathbf{X}) = 0 + \boldsymbol{\beta}\sigma_X^2$ dir. Bu nedenle, $\boldsymbol{\beta} = \dfrac{kov(\mathbf{Y,X})}{\sigma_X^2}$ dir. $\boldsymbol{\mu}_Y = \boldsymbol{\alpha} + \boldsymbol{\beta\mu}_X + \boldsymbol{\mu}_\varepsilon = \boldsymbol{\alpha} + \boldsymbol{\beta\mu}_X$ dir. Bu nedenle $\boldsymbol{\alpha} = \boldsymbol{\mu}_Y - \boldsymbol{\beta\mu}_X$ dir.

Teorem 3.4.2 den farklı olarak, Teorem 3.4.3 en yakın nesne kavramına göre ifade edilmez. Fakat bu çok kolay bir ispattır.

Örneklemler İçin Lineer Regresyon

Kitle regresyon modelini tanıttıktan sonra, istatistik regresyon modelini bir örneklem tahmini olarak ortaya koyacağız. Bir kitle nesnesi, bir örneklemden tahmin edildiğinde, tahmin teorisindeki genel şablonu izleriz. Örneğin, $\boldsymbol{\mu}$ kitle ortalaması bir $\overline{\mathbf{X}} = \dfrac{\sum_{i=1}^{n} \mathbf{X}_i}{\mathbf{n}}$ ile tahmin edilir. Benzer şekilde, basit lineer regresyonun $\mathbf{Y} = \boldsymbol{\alpha} + \boldsymbol{\beta}\mathbf{X} + \boldsymbol{\varepsilon}$ denklemi (bağıntısı); $\mathbf{a, b}$ ve $\mathbf{e}$ sırasıyla $\boldsymbol{\alpha}, \boldsymbol{\beta}$ ve $\boldsymbol{\varepsilon}$ nun örneklem tahminleri olmak üzere, örneklemden $\mathbf{Y = a + bX + e}$ denklemiyle(bağıntısıyla) tahmin edilir.(3.4.1) deki parametreler için karşılık gelen örneklem tahminlerini yerlerine koyarak, $\mathbf{a}$ ve $\mathbf{b}$ katsayıları için formülleri elde ederiz. Böylece, $\overline{\mathbf{X}}, \overline{\mathbf{Y}}, \mathbf{S}_X^2$ ve $\mathbf{S}_{YX}$ sırasıyla, $\boldsymbol{\mu}_X, \boldsymbol{\mu}_Y, \boldsymbol{\sigma}_X^2$ ve $kov(\mathbf{Y,X})$ 'in örneklem tahminleri olmak üzere,

$$\begin{cases} \mathbf{b} = \dfrac{\mathbf{S}_{YX}}{\mathbf{S}_X^2} \\ \mathbf{a} = \overline{\mathbf{Y}} - \mathbf{b}\overline{\mathbf{X}} \end{cases} \quad \text{dir.}$$

4.MATERYAL VE YÖNTEM

EN KÜÇÜK KARELERİN GEOMETRİSİ

4.1. VEKTÖR UZAYI GEOMETRİSİ

Burada daha önce verdiğimiz bazı kavramları tekrarlayarak konunun daha iyi anlaşılmasını sağlayacağız.

Tanım 4.1.1: Bir lineer model, $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ olarak tanımlanır. Burada $\mathbf{Y}_{n\times 1}$, gözlenebilir değişkenlerin yani (bağımlı değişkenlerin) gözlenen bir vektörü, $\mathbf{X}_{n\times k}$, bilinen bir matrisi ve $\boldsymbol{\varepsilon}_{n\times 1}$, hataların bir vektörüdür ve $\boldsymbol{\varepsilon} \sim \mathbf{N}(\mathbf{0}, \sigma^2\mathbf{I})$ dağılımına sahiptir.

Bir $\mathbf{S}$ vektör uzayı: $\mathbf{S}$ üzerinde bir toplama ve bir skalar çarpım ile birlikte, değişim, birleşim vs. gibi bazı özellikleri sağlayan bir kümedir.

$\mathbf{R}^n$ Öklid uzayı, vektörlerin alışılmış toplamı ve skalar çarpım ile $\boldsymbol{R}$ reel sayıları göstermek üzere, $\boldsymbol{R}^n$ deki tüm vektörler tarafından oluşturulan vektör uzayıdır. Aslında bir vektör uzayını oluşturmak için gerekenlerden daha fazla yapı koyacağız.

4.2. BAZI TANIMLAR VE NOTASYONLAR

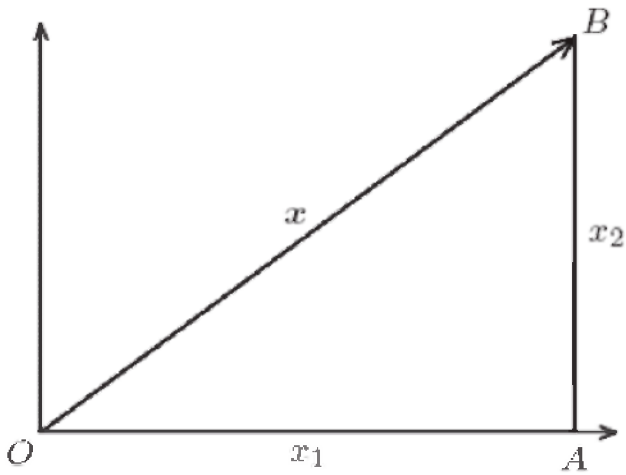
İç çarpım: $\langle \mathbf{x}, \mathbf{y} \rangle \equiv \mathbf{x}'\mathbf{y}$ dir.

Norm: $\|\mathbf{x}\| = (\mathbf{x}'\mathbf{x})^{\frac{1}{2}} = (\sum_{i=1}^{n} \mathbf{x}_i^2)^{\frac{1}{2}}$ dir.

Ortogonallik: $\mathbf{x}$ ve $\mathbf{y}$ vektörlerinin ortogonal olmaları için gerek ve yeter şart $\langle \mathbf{x}, \mathbf{y} \rangle \equiv \mathbf{x}'\mathbf{y} = \mathbf{0}$ olmasıdır.

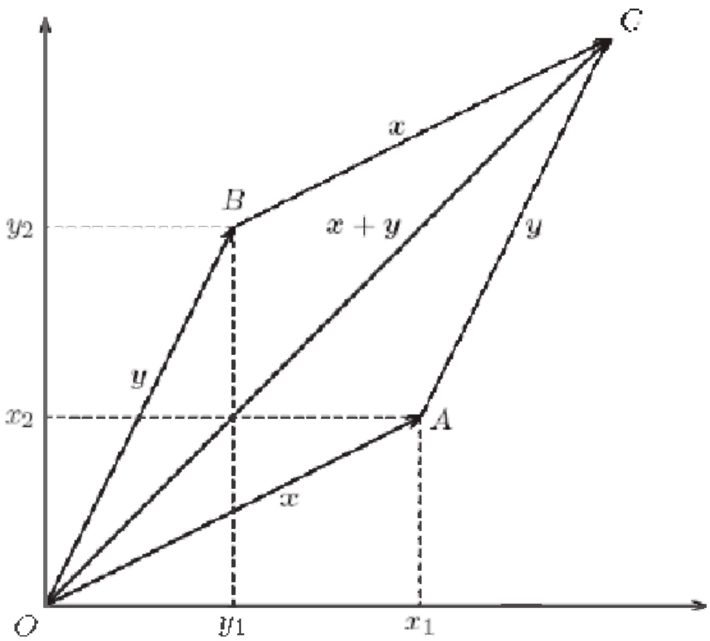
Lineer bağımlılık: Eğer $\mathbf{x}_j = \sum_{i \neq j} \mathbf{c}_i \mathbf{x}_i$ olacak şekilde $\mathbf{x}_j$, $i \le j \le k$, ve $\mathbf{c}_i$ katsayıları varsa, $\mathbf{x}_1, \ldots, \mathbf{x}_k$ vektörleri lineer bağımlıdır.

$\mathbf{R}^2$ de bir vektörün geometrik gösterimi aşağıdaki şekilde verilmiştir.



Şekil 4.1. $\mathbf{R}^2$ de bir vektörün geometrik gösterimi

$\mathbf{R}^2$ de iki vektörün toplamı için paralelkenar kuralının geometrik gösterimi aşağıdaki şekilde verilmiştir.



Şekil 4.2. $\mathbf{R}^2$ de iki vektörün toplamı için paralelkenar kuralı

4.3. ÖKLİD UZAYI VE ÖKLİD UZAYININ ALTUZAYLARI

Bir vektör altuzayı; bir vektör uzayının, kendisi de bir vektör uzayı olan herhangi bir alt kümesidir.

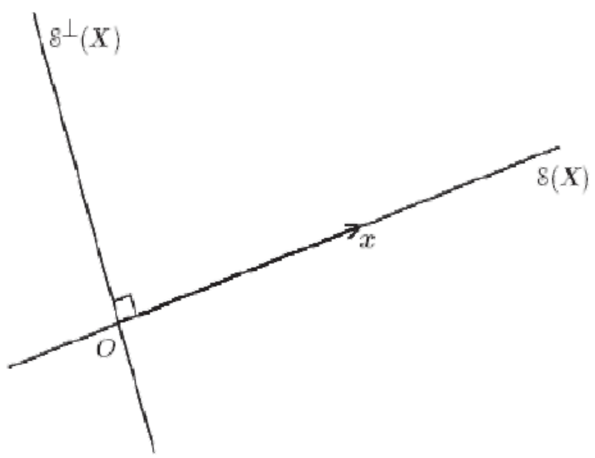
Geren: Bir $\mathbf{S(x_1,\ldots,x_k)} \equiv \left\{ \mathbf{z} \in \mathbf{E^n} \mid \mathbf{z} = \sum_{i=1}^{k} \mathbf{b_i x_i}, \mathbf{b_i} \in \mathbf{R} \right\}$ gereni, $\mathbf{x_1,\ldots,x_k}$ tarafından gerilen Öklid vektör altuzayıdır yani ($\mathbf{x_1,\ldots,x_k}$)' nın tüm lineer kombinasyonlarının kümesidir.

Bir başka şekilde, $\mathbf{X} = \{ \mathbf{x_1,\ldots,x_k} \}$, yani $\mathbf{S(X)} = \left\{ \mathbf{z} \in \mathbf{E^n} \mid \mathbf{z} = \mathbf{X}_\gamma \right\}, \mathbf{X}$ in sütunları tarafından üretilen altuzay ($\mathbf{X}$ ' in gereni) dir. Bu altuzayın tüm vektörleri $\mathbf{X}$ ' in sütunlarının bir lineer kombinasyonu olarak oluşturulabilirler.

Ortogonal Tümleyen : $\mathbf{S(X)}$ in ortogonal tümleyeni,
$\mathbf{S^\perp(X)} = \left\{ \mathbf{w} \in \mathbf{E^n} \mid \mathbf{w'z} = \mathbf{0}, \text{ her } \mathbf{z} \in \mathbf{S(X)} \text{ için} \right\}$ dır. Ortogonal tümleyenin tüm vektörleri $\mathbf{X}$ ' in sütunlarına ortogonaldırlar.

Boyut: Bir $\mathbf{X}$ matrisinin sütun uzayının boyutu herhangi bir tabanın vektörlerinin sayısıdır ($= rank(\mathbf{X})$). Boy $\mathbf{S(X)} \equiv rank(\mathbf{X})$ olduğuna dikkat edelim.

Sonuç: $\mathbf{X}_{n\times k}$ için boy $\mathbf{S(X)} = \mathbf{k}$ ise boy $\mathbf{S^\perp(X)} = \mathbf{n - k}$ dır.



Şekil 4.3. $\mathbf{S(X)}$ in ortogonal tümleyeni $\mathbf{S^\perp(X)}$

Şekil 4.3’ den de görüldüğü gibi, $\mathbf{X}$, $\mathbf{R}^2$ de bir vektördür. $\mathbf{S(X)}$, $\mathbf{X}$ tarafından gerilen altuzaydır. $\mathbf{S}^{\perp}\mathbf{(X)}$ onun ortogonal tümleyenidir. Bu uzayların her biri $\mathbf{1}$ boyutludur.

4.4. EKSENE GÖRE DEĞİŞKENLER VE GÖZLEMLER

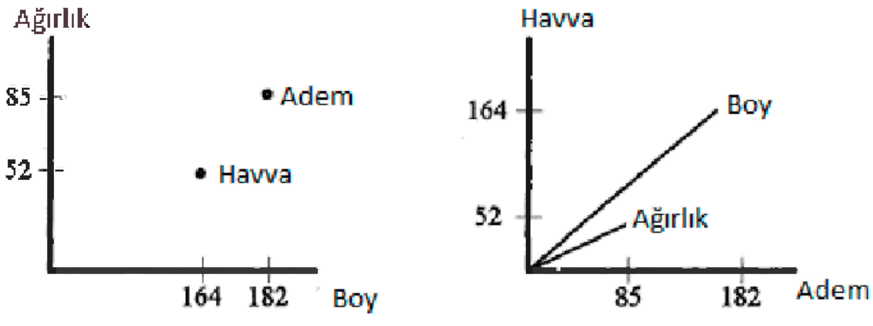
Amaç veriyi ve OLS (ordinary least squares – basit en küçük kareler) tahmin edicisini temsil etmek (göstermek)tir. “Nokta” kavramımızı değiştirmemiz gerekir. Bir saçılım grafiği (noktasal diyagram) her gözlemi bir nokta olarak alır. Şimdi $\mathbf{Y}$ ’ yi ve $\mathbf{X}$ ’in sütunlarını $\mathbf{R}^n$ de $k+1$ tane nokta olarak düşünmemiz gerekir.

$$\begin{pmatrix} y_1 & x_{11}\cdots & x_{1k} \\ y_2 & x_{21}\cdots & x_{2k} \\ y_n & x_{n1}\cdots & x_{nk} \end{pmatrix}$$

Her bir sütun bir noktadır.

Çizelge 4.1. Adem ve Havva için Varsayımsal Veri

	Boy(cm)	Ağırlık (kg)
Adem	182	85
Havva	164	52



Şekil 4.4. Bir regresyon modelinde değişkenleri karşılaştırmak için bir geometrik yaklaşım.(Bring 1996)

Eğer bu şekle üçüncü bir şahsı, dördüncü bir şahsı vs. eklersek ne olmasını bekleriz?

4.5. GAUSS-MARKOV TEOREMİ

$\mathbf{Y} = \mathbf{X\beta} + \boldsymbol{\varepsilon}$ lineer modelini düşünelim.

Burada $\mathbf{X}$ 'e bazen tasarım matrisi olarak başvurulur. $\mathbf{X}$, açıklayıcı değişkenlere karşılık gelen sütunlara sahip olan, sabitlerin bir $n \times p$ matrisidir. $\boldsymbol{\beta}$, $\mathbf{R}^p$ de bir bilinmeyen parametre vektörüdür.

Tasarım matrisinin sütun uzayı

$\mathbf{X\beta}$, $\mathbf{X}$ 'in sütunlarının bir lineer kombinasyonudur.

Yani, $\mathbf{X\beta} = \left[\mathbf{x}_1, ..., \mathbf{x}_p\right]\begin{bmatrix} \boldsymbol{\beta}_1 \\ \cdot \\ \cdot \\ \cdot \\ \boldsymbol{\beta}_P \end{bmatrix} = \boldsymbol{\beta}_1\mathbf{x}_1 + ... + \boldsymbol{\beta}_p\mathbf{x}_p$ dır. Bir başka deyişle $\mathbf{X}$ 'in sütunlarının tüm mümkün olabilen lineer kombinasyonlarının kümesine $\mathbf{X}$ 'in sütun uzayı denir ve $C(\mathbf{X}) = \left\{\mathbf{Xa} : \mathbf{a} \in \mathbf{R}^P\right\}$ ile gösterilir.

Gauss –Markov lineer modeli, $\mathbf{Y}$ 'nin; ortalaması $\mathbf{X}$ 'in sütun uzayı olan, ve bir σ^2 pozitif reel sayısı için, varyansı $\sigma^2\mathbf{I}$ olan, yani, $E(\mathbf{Y}) \in C(\mathbf{X})$ ve $Var(\mathbf{Y}) = \sigma^2\mathbf{I}$, $\sigma^2 \in \mathbf{R}^+$ olan, bir rasgele vektör olduğunu söyler.

$\mathbf{X} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \Rightarrow C(\mathbf{X}) = \left\{\mathbf{Xa} : \mathbf{a} \in \mathbf{R}\right\} = \left\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}\left[\mathbf{a}_1\right] : \mathbf{a}_1 \in \mathbf{R}\right\} = \left\{\begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_1 \end{bmatrix} : \mathbf{a}_1 \in \mathbf{R}\right\}$ bir sütun uzayı örneğidir.

Başka bir sütun uzayı örneği;

$$\mathbf{X} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \Rightarrow \mathbf{C}(\mathbf{X}) = \left\{\begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \end{bmatrix} : \mathbf{a} \in \mathbf{R}^2\right\}$$

$$= \left\{ \mathbf{a}_1 \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + \mathbf{a}_2 \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} : \mathbf{a}_1, \mathbf{a}_2 \in \mathbf{R} \right\}$$

$$= \left\{ \begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_1 \\ \mathbf{a}_2 \\ \mathbf{a}_2 \end{bmatrix} : \mathbf{a}_1, \mathbf{a}_2 \in \mathbf{R} \right\}$$

dir.

Bu sütun uzayı $\mathbf{R}^4$ de uzanan bir düzlemdir. Üçüncü bir sütun uzayı örneği;

$$\mathbf{X}_2 = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix}, \ \mathbf{x} \in \mathcal{C}(\mathbf{X}_2) \Rightarrow \text{herhangi } \mathbf{a} \in \mathbf{R}^3 \text{ için } \mathbf{x} = \mathbf{X}_2\mathbf{a} \Rightarrow \text{ herhangi } \mathbf{a} \in \mathbf{R}^3 \text{ için}$$

$$\Rightarrow \mathbf{x} = \mathbf{a}_1 \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} + \mathbf{a}_2 \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + \mathbf{a}_3 \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}$$

$$\Rightarrow \text{herhangi } \mathbf{b}_1, \mathbf{b}_2 \in \mathbf{R} \text{ için } \mathbf{x} = \begin{bmatrix} \mathbf{a}_1 + \mathbf{a}_2 \\ \mathbf{a}_1 + \mathbf{a}_2 \\ \mathbf{a}_1 + \mathbf{a}_3 \\ \mathbf{a}_1 + \mathbf{a}_3 \end{bmatrix} = \begin{bmatrix} \mathbf{b}_1 \\ \mathbf{b}_1 \\ \mathbf{b}_2 \\ \mathbf{b}_2 \end{bmatrix} \text{dir.}$$

Bu da $\mathbf{R}^4$ de bir düzlemdir. Aynı düzlem midir?

İki Sütun Uzayının, yani $\mathcal{C}(\mathbf{X}_1)$ ve $\mathcal{C}(\mathbf{X}_2)$'nin, Aynı Olduklarını Gösterme

Kavram: Eğer $\mathbf{x} \in \mathcal{C}(\mathbf{X}_1)$ ile başlayabilir ve $\mathbf{x} = \mathbf{X}_2\mathbf{b}$ alırsak, bu, $\mathbf{x} \in \mathcal{C}(\mathbf{X}_2)$ ve $\mathcal{C}(\mathbf{X}_1) \subseteq \mathcal{C}(\mathbf{X}_2)$ olduğunu gösterir. En azından henüz $\mathcal{C}(\mathbf{X}_1) = \mathcal{C}(\mathbf{X}_2)$ değildir. Çünkü $\mathbf{X}_2\mathbf{b}$ 'nin $\mathcal{C}(\mathbf{X}_1)$ de olmadığı bazı b' ler mevcut olabilir. Eğer aynı zamanda

$C(\mathbf{X}_2) \subseteq C(\mathbf{X}_1)$ olduğunu gösterebilirsek, bu taktirde $C(\mathbf{X}_1) \subseteq C(\mathbf{X}_2)$ ve $C(\mathbf{X}_2) \subseteq C(\mathbf{X}_1) \Rightarrow C(\mathbf{X}_1) = C(\mathbf{X}_2)$ olduğunu gösterir.

Örnek: $\mathbf{X}_1 = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}$, $\mathbf{X}_2 = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix}$ olsun.

$\mathbf{x} \in C(\mathbf{X}_1) \Rightarrow \mathbf{x} = \mathbf{X}_1\mathbf{a} \Rightarrow \mathbf{x} = \mathbf{X}_2 \begin{bmatrix} 0 \\ \mathbf{a} \end{bmatrix} \Rightarrow$ bazı $\mathbf{b} \in \mathbf{R}^3$ ler için $\mathbf{x} = \mathbf{X}_2\mathbf{b} \Rightarrow \mathbf{x} \in C(\mathbf{X}_2)$ olduğunu gösterir. Bu nedenle $C(\mathbf{X}_1) \subseteq C(\mathbf{X}_2)$ dir. Yine $\mathbf{X}_1 = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}$, $\mathbf{X}_2 = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix}$ olsun. $\Rightarrow \mathbf{x} \in C(\mathbf{X}_2)$ bazı $\mathbf{a} \in \mathbf{R}^3$ ler için $\mathbf{x} = \mathbf{X}_2\mathbf{a} \Rightarrow$ bazı $\mathbf{a} \in \mathbf{R}^3$ ler için

$\mathbf{x} = \mathbf{a}_1 \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} + \mathbf{a}_2 \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + \mathbf{a}_3 \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} \Rightarrow$ bazı $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3 \in \mathbf{R}$' ler için $\mathbf{x} = \begin{bmatrix} \mathbf{a}_1 + \mathbf{a}_2 \\ \mathbf{a}_1 + \mathbf{a}_2 \\ \mathbf{a}_1 + \mathbf{a}_3 \\ \mathbf{a}_1 + \mathbf{a}_3 \end{bmatrix} \Rightarrow$ bazı $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3 \in \mathbf{R}$ ler için $\mathbf{x} = \mathbf{X}_1 \begin{bmatrix} \mathbf{a}_1 + \mathbf{a}_2 \\ \mathbf{a}_1 + \mathbf{a}_3 \end{bmatrix} \Rightarrow$ bazı $\mathbf{b}_1, \mathbf{b}_2 \in \mathbf{R}$ için $\mathbf{x} = \mathbf{X}_1 \begin{bmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \end{bmatrix} \Rightarrow$ bazı $\mathbf{b} \in \mathbf{R}$ ler için $\mathbf{x} = \mathbf{X}_1\mathbf{b} \Rightarrow \mathbf{x} \in C(\mathbf{X}_1)$ olduğunu gösterir. Bu nedenle, $C(\mathbf{X}_2) \subseteq C(\mathbf{X}_1)$ dır, $C(\mathbf{X}_1) \subseteq C(\mathbf{X}_2)$ olduğunu daha önce gösterdik. Böylece $C(\mathbf{X}_1) = C(\mathbf{X}_2)$ olduğu görülür.

$E(\mathbf{Y})$'nin Tahmini

Lineer model analizinin temel amaçlarından biri $E(\mathbf{Y})$'yi tahmin etmektir. Şüphesiz, $E(\mathbf{Y})$ yi tahmin etmek için $\mathbf{Y}$'yi kullanabilirdik. $\mathbf{Y}$ aşikar olarak, $E(\mathbf{Y})$'nin bir yansız tahmin edicisidir. Ancak o, çoğu kez çok makul edici bir tahmin edici değildir.

Örneğin,

$$\begin{bmatrix} \mathbf{Y}_1 \\ \mathbf{Y}_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \boldsymbol{\mu} + \begin{bmatrix} \boldsymbol{\varepsilon}_1 \\ \boldsymbol{\varepsilon}_2 \end{bmatrix}$$

olduğunu ve

$$\mathbf{Y} = \begin{bmatrix} 6.1 \\ 2.3 \end{bmatrix}$$

gözlemini yaptığımızı farz edelim.

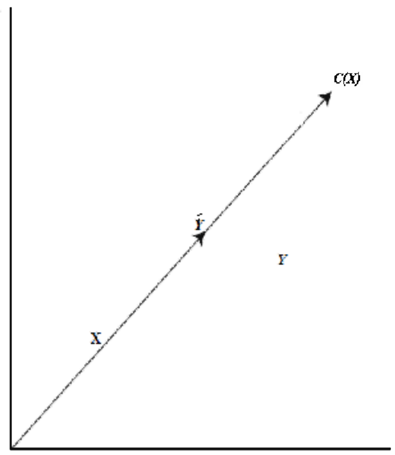
$E(\mathbf{Y}) = \begin{bmatrix} \boldsymbol{\mu} \\ \boldsymbol{\mu} \end{bmatrix}$ yi, $\mathbf{Y} = \begin{bmatrix} 6.1 \\ 2.3 \end{bmatrix}$ ile tahmin edebilir miyiz?

$E(\mathbf{Y})$' nin Tahmini

Gauss–Markov lineer modeli $E(\mathbf{Y}) \in C(\mathbf{X})$ olduğunu söyler, bu nedenle $E(\mathbf{Y})$' yi tahmin ederken bu bilgiyi kullanmalıyız.

$C(\mathbf{X})$' de $\mathbf{Y}$ 'ye en yakın olan noktayla, $E(\mathbf{Y})$' yi tahmin etmeyi düşünelim (uzaklık, alışılmış Öklid uzaklığı ile ölçülsün). Bu tek noktaya, $\mathbf{Y}$ 'nin $E(\mathbf{X})$ üzerindeki ortogonal izdüşümü denir ve bu izdüşüm $\hat{\mathbf{Y}}$ ile gösterilir.(Bununla beraber $E(\mathbf{Y})$ daha iyi bir notasyondur). Tanıma göre, $\|\mathbf{a}\| = \sqrt{\sum_{i=1}^{n} \mathbf{a}_i^2}$ olmak üzere, $\|\mathbf{Y} - \hat{\mathbf{Y}}\| = \min_{z \in C(X)} \|\mathbf{Y} - \mathbf{Z}\|$ dir.

Bu durumu bir şekil 4.5 ile gösterelim $\mathbf{X} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ ve $\mathbf{Y} = \begin{bmatrix} 6.1 \\ 2.3 \end{bmatrix}$ olduğunu farz edelim.



Şekil 4.5. $\mathbf{Y}$ 'nin $\mathbf{C(X)}$ üzerindeki ortogonal izdüşümü: $\hat{\mathbf{Y}}$

Geometri Bir İstatikçiye Ne Söyler?

$\hat{\mathbf{Y}}, C(\mathbf{X})$ ' de $\left\|\mathbf{Y}-\hat{\mathbf{Y}}\right\|^2 = \sum_{i=1}^{n}(\mathbf{Y}-\hat{\mathbf{Y}})^2$ 'yi minimumlaştıran noktadır. Bu nedenle en küçük kareler (E.K.K) tahmini yapıyoruz. Geometrik olarak, $\left\|\mathbf{Y}-\hat{\mathbf{Y}}\right\|$, $\hat{\mathbf{Y}}$ vektörü ve $\mathbf{Y}-\hat{\mathbf{Y}}$ vektörü arasındaki açı 90° olduğunda minimumlaşır. $\hat{\mathbf{Y}}$ ve $\mathbf{Y}-\hat{\mathbf{Y}}$ vektörleri ortogonaldir. $\hat{\mathbf{Y}}$ ve $\mathbf{Y}-\hat{\mathbf{Y}}$ arasındaki kolerasyon sıfırdır, yani tahmin edilmiş $\hat{\mathbf{Y}}$ değerleri ve $\mathbf{Y}-\hat{\mathbf{Y}}=\mathbf{e}$ hata tahminleri ilişkisizdirler.

Teorem 4.5.1 (Pisagor teoremi): $\left\|\mathbf{Y}\right\|^2 = \left\|\hat{\mathbf{Y}}\right\|^2 + \left\|\mathbf{Y}-\hat{\mathbf{Y}}\right\|^2$, aslında kareler toplamlarının ANOVA (varyans analizi) ayrışımı $= SS_{model} + SS_{hata}$ =model kareler toplamı +hata kareler toplamıdır. Ancak, doğruları çizmeden $\hat{\mathbf{Y}}$ 'yı nasıl hesaplarız?

Ortogonal İzdüşüm (dik izdüşüm) Matrisleri

$\hat{\mathbf{Y}}$ 'yı matris çarpımıyla bulabilir miyiz? Her $\mathbf{Y} \in \mathbf{R}^n$ için $\hat{\mathbf{Y}} = \mathbf{P}_\mathbf{X}\mathbf{Y}$ dir. Burada $\mathbf{P}_\mathbf{X}$ bir ortogonal izdüşüm matrisi olarak bilinen bir tek $n \times n$ matristir. $\mathbf{P}_\mathbf{X}$ nasıl bir matristir? Eğer $\mathbf{(X'X)}^{-1}$ mevcutsa, yani, $\mathbf{X'X}$ tam ranklı ($\mathbf{X}$ tam sütun ranklı) ise, $\mathbf{P}_\mathbf{X} = \mathbf{X(X'X)}^{-1}\mathbf{X'}$ dir.

Eğer $\mathbf{X'X}$ tam ranklı değilse, $\mathbf{P_X = X(X'X)\ X'}$ dür. Burada $\mathbf{(X'X)}$, $\mathbf{X'X}$ 'in herhangi bir genelleştirilmiş tersidir. $\mathbf{P_X X = X}$ ve $\mathbf{X'P_X = X'}$ $\mathbf{P_X P_X = P_X}$,yani $\mathbf{P_X}$ simetrik-idempotent, oldukları gösterilebilir.

Niçin $\mathbf{P_X X = X}$ dir?

Geometrik olarak: $\mathbf{X}$ 'in her bir sütunu $C\mathbf{(X)}$ de bir noktayı gösterir (tanımdan). $\mathbf{P_X Y}$, $\mathbf{Y} \in \boldsymbol{R}^n$ dik izdüşümü $C\mathbf{(X)}$' de $\mathbf{Y}$ 'ye en yakın noktadır.

Cebirsel olarak: ($\mathbf{X}$ 'in tam ranklı olduğunu kabul ederek)

$\mathbf{P_X X = X(X'X)^{-1} X'\ X = X}$ dir.

$\mathbf{P_X}$ Niçin İdempotentdir?

Geometrik olarak: Eğer $\mathbf{X}$ bir $n \times p$ matris ise, herhangi bir $\mathbf{A}$ $n \times k$ matrisini düşünelim. $\mathbf{P_X A}$ 'nın her bir sütunu $C\mathbf{(X)}$ de bir noktayı temsil eder (tanımdan). İkinci kere bir izdüşüm $\mathbf{P_X A}$ yı hareket ettirmez. Bu, herhangi bir $\mathbf{A}$ için doğru olduğundan, $\mathbf{P_X P_X X = P_X X}$ dir. Yani $\mathbf{P_X^2 = P_X}$ dir. Sonuç olarak $\mathbf{P_X}$ idempotenttir.

Cebirsel olarak: ($\mathbf{X}$ matrisini $n \times p$ boyutlu ve tam ranklı kabul ederek): $n \times k$ boyutlu herhangi bir $\mathbf{A}$ matrisini düşünelim.

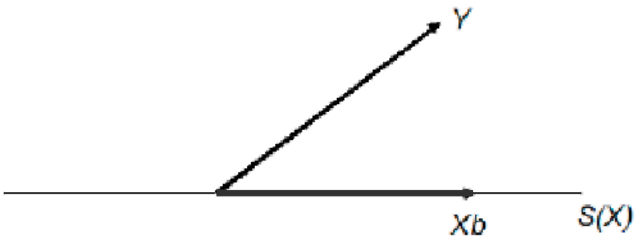
$$\begin{aligned}\mathbf{P_X P_X A} &= \left[\mathbf{X(X'X)^{-1}X'}\right]\left[\mathbf{X(X'X)^{-1}X'}\right]\mathbf{A}\\ &= \mathbf{X(X'X)^{-1}(X'X)(X'X)^{-1}X'A = X(X'X)^{-1}X'A = P_X A}\end{aligned}$$

olur.

4.6. ALIŞILMIŞ EN KÜÇÜK KARELER (OLS) GEOMETRİSİ

$\mathbf{S(X)} \equiv \left\{\mathbf{z} \in \mathbf{E}^n \mid \mathbf{z} = \mathbf{X}_\gamma\right\}$, yani $\mathbf{x_1,...,x_k}$ tarafından gerilen Öklid vektör altuzayı ya da $\mathbf{X}$` in gereni veya $\left[\mathbf{x_1, x_2, ... x_k}\right]$'nın tüm lineer kombinasyonları olarak olmak üzere, $\mathbf{S(X)}$ deki herhangi bir nokta, $\boldsymbol{\beta} \in \mathbf{R}^k$ olmak üzere, $\mathbf{X}\boldsymbol{\beta}$ olarak ifade edilebilir.

En küçük kareler: Verilen $\mathbf{X}$ ve $\mathbf{Y}$ için $\mathbf{S(X)}$ de $\mathbf{Y}$ ’ ye mümkün olabildiği kadar yakın olan noktayı bulalım.



Şekil 4.6. $\mathbf{Y}$ nin $\mathbf{S(X)}$ e olan en yakın noktası

Bir lineer model, daha önceden olduğu gibi, $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ olarak tanımlanır. Burada;

Problem: $min_{\beta} \|\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}\| \Leftrightarrow min_{\beta} \|\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}\|^2$ dır.

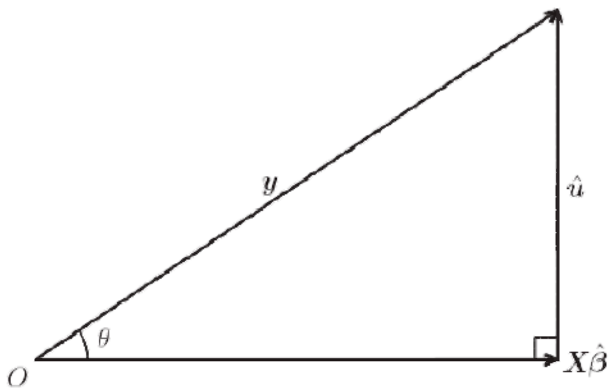
Tanım: $\hat{\boldsymbol{\beta}}$ probleme bir çözümü ise, $\hat{\mathbf{Y}} = \mathbf{X}\hat{\boldsymbol{\beta}}$ (gözlemlerin tahmin edilen vektörü), ve $\mathbf{e} = \mathbf{Y} - \hat{\mathbf{Y}}$ (tahmin edilen hata vektörü) olacaktır.

Bazı özellikler:

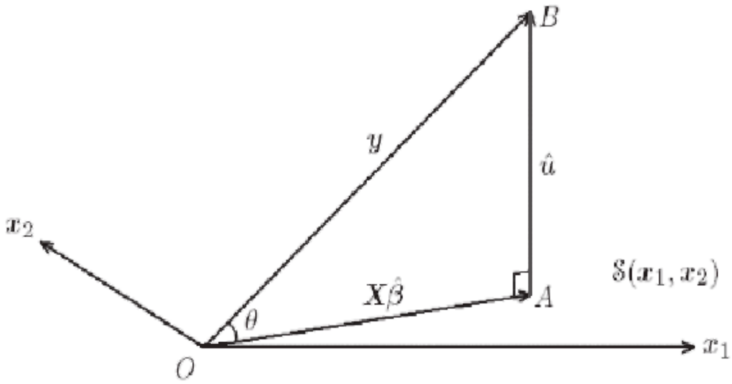
$\mathbf{e}$, $\mathbf{S(X)}$ deki herhangi bir noktaya, özellikle $\mathbf{X}$ ’e veya $\mathbf{X}\hat{\boldsymbol{\beta}}$ ’ya diktir.

$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$ ’, yani $\mathbf{X}'\mathbf{X}\boldsymbol{\beta} = \mathbf{X}'\mathbf{Y}$ normal denklemlerinin çözümüdür.

Ortogonallik şartından $\mathbf{X'(Y - X\hat{\boldsymbol{\beta}}) = X'e = 0}$ dır.



Şekil 4.7. $\mathbf{Y}$ nin $\mathbf{X}\hat{\boldsymbol{\beta}}$ üzerine ortogonal izdüşümü



Şekil 4.8. $\mathbf{Y}$ nin $\mathbf{S(x_1, x_2)}$ üzerine ortogonal izdüşümü

4.7. İZDÜŞÜMLER

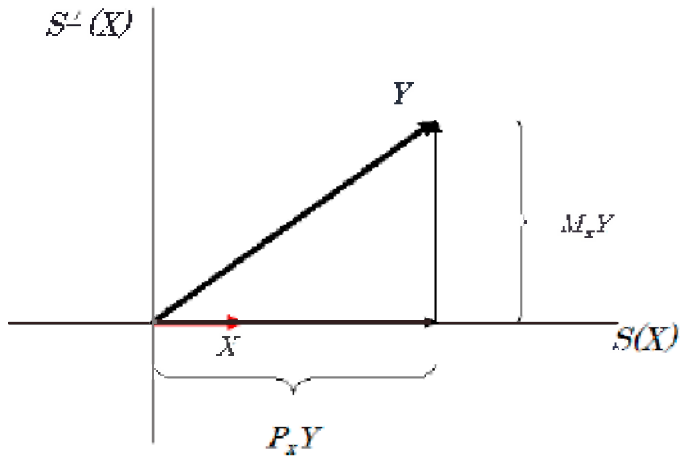
Bir izdüşüm, $\mathbf{E}^n$ deki herhangi bir noktayı $\mathbf{E}^n$ 'nin bir altuzayındaki bir noktaya götüren bir dönüşümdür.

Bir ortogonal izdüşüm herhangi bir noktayı altuzayın ona en yakın noktasına dönüştürür.

$\mathbf{\hat{Y} = X\hat{\beta} = X(X'X)^{-1}X'Y = P_X Y}$, $\mathbf{Y}$ 'nin $\mathbf{S(X)}$ üzerindeki ortogonal izdüşümüdür. $\mathbf{P_X = X(X'X)^{-1}X'}$, $\mathbf{Y}$ 'yi $\mathbf{S(X)}$ üzerine ortogonal olarak izdüşüren izdüşüm matrisidir. $\mathbf{e = Y - \hat{Y} = Y - X\hat{\beta} = (I - X(X'X)^{-1}X')Y = M_X Y}$, $\mathbf{Y}$ 'nin $\mathbf{S(X)}$ ' in ortogonal tümleyeni üzerine yani $\mathbf{S^{\perp}(X)}$ üzerine izdüşümüdür.

$\mathbf{M_X = I - P_X = I - X(X'X)^{-1}X'}$, $\mathbf{Y}$ 'yi $\mathbf{S^{\perp}(X)}$ üzerine ortogonal olarak izdüşüren izdüşüm matrisidir.

Özellikler: $\mathbf{M_X}$ ve $\mathbf{P_X}$ 'i cebirsel olarak kontrol etmek kolaydır, geometriksel olarak anlamak ise daha iyidir. $\mathbf{M_X}$ ve $\mathbf{P_X}$ simetrik matrislerdir. $\mathbf{M_X + P_X = I}$ 'dır. Bu, $\mathbf{Y = M_X Y + P_X Y}$ ortogonal ayrışımını ortaya koyar.



Şekil 4.9. $\mathbf{Y}$ nin $\mathbf{S(X)}$'e ve $\mathbf{S}^{\perp}\mathbf{(X)}$' e göre durumu

$\mathbf{P_X}$ ve $\mathbf{M_X}$ idempotenttirler: yani $\mathbf{P_X P_X = P_X}$, $\mathbf{M_X M_X = M_X}$ dir.

Seziş: Bir vektör önceden $\mathbf{S(X)}$ de ise, ayrıca onu $\mathbf{S(X)}$' e izdüşürme bir anlam ifade etmez.

$$\mathbf{P_X . M_X = 0}$$

Birinci izdüşümü yaptıktan sonra diğerini (herhangi bir sırada) nasıl yapacağımızı düşünelim. $\mathbf{P_X}$ ve $\mathbf{M_X}$ biri diğerini yok eder. $\mathbf{0}$ hem $\mathbf{S(X)}$'e hem de $\mathbf{S}^{\perp}\mathbf{(X)}$' e ait olan yegane noktadır.

$\mathbf{M_X}$ $\mathbf{S(X)}$ deki herhangi bir noktayı yok eder, yani $\mathbf{M_X X\beta = 0}$ dır.

$\mathbf{P_X}$ $\mathbf{S}^{\perp}\mathbf{(X)}$ deki herhangi bir noktayı yok eder; yani $\mathbf{P_X M_X X\beta = 0}$ dır. Eğer $\mathbf{A}$ bir tekil olmayan $k \times k$ matris ise, $\mathbf{P_{XA} = P_X}$ dir.

$rank(\mathbf{X}) = rank(\mathbf{P_X})$ dir.

4.8. UYUMUN İYİLİĞİ

Ortogonal ayrışımdan,

$$\mathbf{Y = PY + MY}$$

dir. Bu taktirde

$$\mathbf{Y'Y} = \mathbf{Y'PY} + \mathbf{Y'MY} \ldots.(1)$$

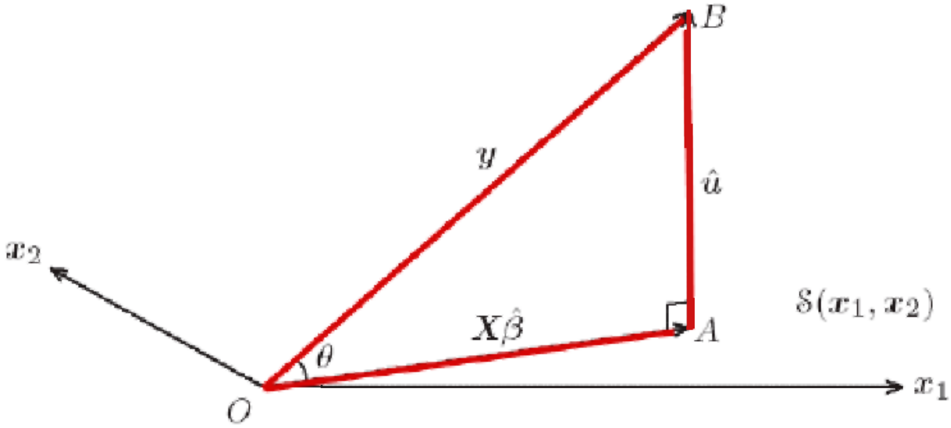
$$= \mathbf{Y'P'PY} + \mathbf{Y'M'MY} \ldots(2)$$

$$\|\mathbf{Y}\|^2 = \|\mathbf{PY}\|^2 + \|\mathbf{MY}\|^2 \ldots(3)$$

dır. $\mathbf{R}^2$ de bu, basit olarak Pisagor Teoremidir. Bu taktirde $\boldsymbol{\theta}$; $\mathbf{Y}$ ve $\mathbf{PY}$ tarafından oluşturulan açı olmak üzere,

$$\mathbf{R}^2 = \frac{\|\mathbf{PY}\|^2}{\|\mathbf{Y}\|^2} = \mathbf{cos}^2\boldsymbol{\theta}$$

dır. Gerçekten bu merkezlenmemiş $\mathbf{R}^2$ dir.



Şekil 4.10. $\cos\theta$ nın Pisagor Teoremi ile bulunması

4.9. FRİSCH-WAUGH-LOVELL TEOREMİ

$\mathbf{Y} = \mathbf{X\beta} + \mathbf{u}$ lineer modelini düşünelim ve onu aşağıdaki gibi parçalayalım. $\mathbf{Y} = \mathbf{X_1\beta_1} + \mathbf{X_1\beta_2} + \mathbf{u}$. Burada $\mathbf{X_1}$ ve $\mathbf{X_2}$ sırasıyla k_1 ve k_2 boyutlu $\mathbf{X_1}$ ve $\mathbf{X_2}$ açıklayıcı değişkenler matrisleridir. Bu taktirde $\mathbf{X} = [\mathbf{X_1\,X_1}]$, $\boldsymbol{\beta} = (\boldsymbol{\beta}_1', \boldsymbol{\beta}_2)'$ ve $\mathbf{k} = \mathbf{k_1} + \mathbf{k_2}$ dır. $\mathbf{M_1} = \mathbf{I} - \mathbf{X_1}(\mathbf{X_1'X_1})^{-1}$, $\mathbf{R}^n$ deki herhangi bir vektörü $\mathbf{X_1}$' in gereninin ortogonal tümleyeni içine izdüşürür.

$\mathbf{Y}^* = \mathbf{M}_1\mathbf{Y}$, $\mathbf{X}_2^* = \mathbf{M}_1\mathbf{X}_2$ sırasıyla $\mathbf{Y}$'nin $\mathbf{X}_1$ üzerinde regresyonunun alışılmış OLS tahminleri ve $\mathbf{X}_2$' nin $\mathbf{X}_1$ üzerindeki tüm sütunlarıdır.

$\boldsymbol{\beta}_2$'yi tahmin etmeyle ilgilendiğimizi varsayalım ve aşağıdaki değişik yöntemleri göz önüne alalım.

1.Yöntem: $\hat{\boldsymbol{\beta}} = (\hat{\boldsymbol{\beta}}_1', \hat{\boldsymbol{\beta}}_2')' = (\mathbf{X'X})^{-1}\mathbf{X'Y}$ elde ederken alışıldığı gibi ilerleyelim $\mathbf{Y}$' yi $\mathbf{X}$ üzerinde regresyonlayalım. $\hat{\boldsymbol{\beta}}_2$ istenen tahmin olmalıdır.

2.Yöntem: $\mathbf{X}_2^*$ üzerinde $\mathbf{Y}^*$'yı regresyonlayalım ve tahmin olarak $\tilde{\boldsymbol{\beta}}_2 = (\mathbf{X}_2^{*'}\mathbf{X}_2^*)^{-1}\mathbf{X}_2^{*'}\mathbf{Y}^*$ yı elde edelim.

$\mathbf{e}_1$ ve $\mathbf{e}_2$ sırasıyla 1. Yöntem ve 2.Yöntemdeki regresyonların hata tahmin vektörleri olsunlar.

Teorem 4.9.1: (Frisch ve Waugh 1933, Lovell 1963)(FWLT Teoremi): $\hat{\boldsymbol{\beta}}_2 = \tilde{\boldsymbol{\beta}}_2$ (birinci kısım) ve $\mathbf{e}_1 = \mathbf{e}_2$ (ikinci kısım).

İspat: İspata $\mathbf{Y} = \mathbf{PY} + \mathbf{MY} = \mathbf{X}_1\hat{\boldsymbol{\beta}}_1 + \mathbf{X}_2\hat{\boldsymbol{\beta}}_2 + \mathbf{MY}$ ortogonal ayrışımı ile başlayalım. Birinci kısmı göstermek için, $\mathbf{Y}$'yi önden $\mathbf{X}_2'\mathbf{M}_1$ ile çarparak $\mathbf{X}_2'\mathbf{M}_1\mathbf{Y} = \mathbf{X}_2'\mathbf{M}_1\mathbf{X}_1\hat{\boldsymbol{\beta}}_1 + \mathbf{X}_2'\mathbf{M}_1\mathbf{X}_2\hat{\boldsymbol{\beta}}_2 + \mathbf{X}_2'\mathbf{M}_1\mathbf{MY}$ elde ederiz. Burada $(\mathbf{I} - \mathbf{X}_1(\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1')\mathbf{X}_1 = 0$ olduğundan $\mathbf{M}_1\mathbf{X}_1 = \mathbf{0}$ dır. $\mathbf{X}_2'\mathbf{M}_1\mathbf{M} = \mathbf{X}_2'(\mathbf{I} - \mathbf{P}_1)\mathbf{M} = \mathbf{X}_2'\mathbf{M} - \mathbf{X}_2'\mathbf{P}_1\mathbf{M} = \mathbf{0}$ dır (Önceki aynı nedenlerden). Burada, $\mathbf{M}_1 = \mathbf{I} - \mathbf{P}_1$, $\mathbf{R}^n$ deki herhangi bir vektörü $\mathbf{X}_1$'in sütunları tarafından gerilen lineer uzayın ortogonal tümleyeni üzerine izdüşüren bir ortogonal izdüşüm matrisi olduğundan, $\mathbf{P}_1\mathbf{P} = \mathbf{P}_1$ dir, bu nedenle $\mathbf{P}_1\mathbf{M} = \mathbf{M}$ dir.

Bu taktirde; $\mathbf{X}_2'\mathbf{M}_1\mathbf{Y} = \mathbf{X}_2'\mathbf{M}_1\mathbf{X}_2\hat{\boldsymbol{\beta}}_2$ dır. Bu nedenle $\hat{\boldsymbol{\beta}}_2 = (\mathbf{X}_2'\mathbf{M}_1\mathbf{X}_2)^{-1}\mathbf{X}_2'\mathbf{M}_1\mathbf{Y}$. İkinci kısmı göstermek için ortogonal ayrışımı $\mathbf{M}_1$ ile önden çarparız ve

$\mathbf{M}_1\mathbf{Y} = \mathbf{M}_1\mathbf{X}_1\hat{\boldsymbol{\beta}}_1 + \mathbf{M}_1\mathbf{X}_2\hat{\boldsymbol{\beta}}_2 + \mathbf{M}_1\mathbf{MY}$ bağıntısını elde ederiz, yine $\mathbf{M}_1\mathbf{X}_1 = \mathbf{0}$ dır.

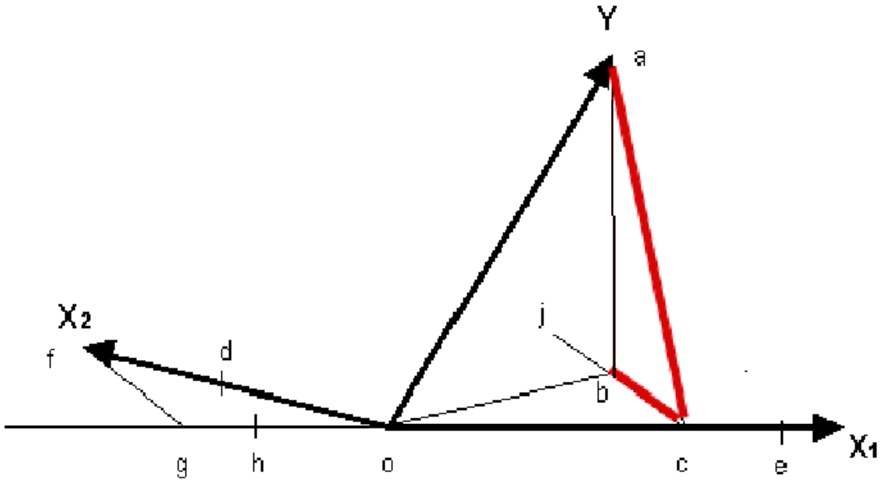
$\mathbf{MY}$ $[\mathbf{X}_1 \mathbf{X}_2]$' nin ortogonal tümleyenine aittir, bu nedenle ayrıca onu $\mathbf{X}_1$' in ortogonal tümleyeni üzerine izdüşürme anlamsız olacaktır, bu nedenle $\mathbf{M}_1\mathbf{MY} = \mathbf{MY}$ dir. (Yapılması gereken şey $\mathbf{M}_1$ ile önden çarpmadır), bu;

$$\mathbf{M}_1\mathbf{Y} - \mathbf{M}_1\mathbf{X}_2\hat{\boldsymbol{\beta}}_2 = \mathbf{MY}$$
$$\mathbf{Y}^* - \mathbf{X}_2^*\hat{\boldsymbol{\beta}}_2 = \mathbf{M}_1$$
$$\mathbf{e}_2 = \mathbf{e}_1$$

bağıntılarını verir.

4.10. (FWLT) NİN GEOMETRİK AÇIKLAMASI

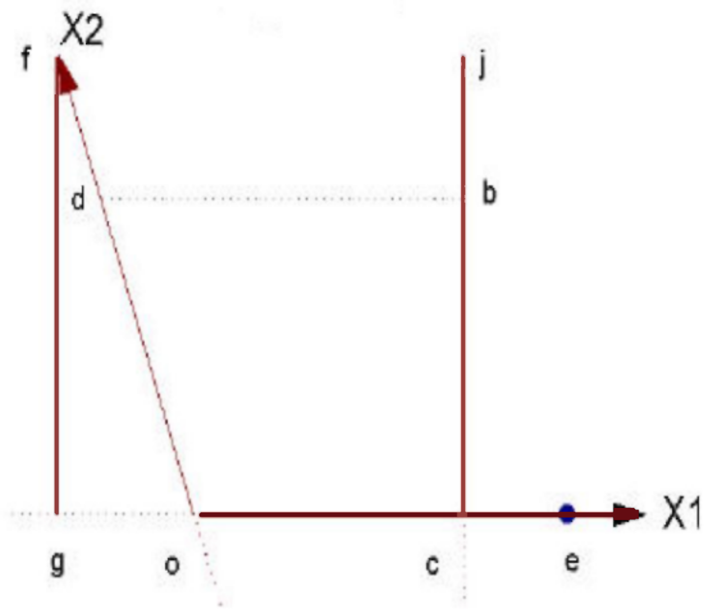
Burada sunulan geometrik gösterim (Davidson ve Mackinon 1993) dakini genişletir. Sadelik için, $k_1 = k_2 = 1$, yani sadece iki açıklayıcı değişkenin var olduğu durumu düşünelim. Şekil 4.11 OLS tahmininde içerilen üç temel vektörü gösterir. $\mathbf{Y}, \mathbf{X}_1$ ve $\mathbf{X}_2$ bir üç boyutlu Öklid vektör uzayındaki vektörlerdir. Veri vektörleri oklarla gösterilirler ve koyu harflerle isimlendirilirler. Küçük harfler noktaları gösterir. OLS, $\mathbf{Y}$ yi $\mathbf{X}_1$ ve $\mathbf{X}_2$ ile gerilen bu durumda iki boyutlu olan, uzay üzerine izdüşürür, OLS izdüşümü, $\mathbf{P} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$, $\mathbf{Y}$ 'yi $\mathbf{X}$ 'in gereni üzerine ortogonal olarak izdüşüm matris olmak üzere, $\mathbf{ob} = \mathbf{PY}$ vektörüyle gösterilir. Tahmin edilen hata vektörü $\mathbf{ab} = \mathbf{MY}$ dir. Burada $\mathbf{M} = \mathbf{I} - \mathbf{P}$, $\mathbf{Y}$ 'yi $\mathbf{X}$ 'ingereninin ortogonal tümleyeni üzerine izdüşüren matristir. Belirlenen $\mathbf{PY} = \mathbf{X}_1\hat{\boldsymbol{\beta}}_1 + \mathbf{X}_2\hat{\boldsymbol{\beta}}_2$ için, $\mathbf{oe} = \mathbf{X}_1\hat{\boldsymbol{\beta}}_1$ ve $\mathbf{od} = \mathbf{X}_2\hat{\boldsymbol{\beta}}_2$ vektörlerinin koordinatları paralelkenar kuralını kullanmak suretiyle kolayca bulunabilir. Bu, 1.yöntemde içerilen tüm unsurların geometrik gösterimini ortaya koyar.



Şekil 4.11.FWLT nin geometrik gösterimi

İkinci yöntemin geometrisini açıklamak için, ilk olarak $\mathbf{Y}$'yi $\mathbf{X}_1$'rin gereni üzerine ortogonal olarak izdüşürelim. Bu izdüşüm $\mathbf{oc} = \mathbf{P}_1\mathbf{Y}$ ile gösterilir ve karşılık gelen hata tahmini vektörü $\mathbf{ac} = \mathbf{Y}^{\cdot\cdot} = \mathbf{M}_1\mathbf{Y}$ dir. Şimdi aynısını $\mathbf{X}_2$ ile yapalım. $\mathbf{X}_2$'nin $\mathbf{X}_1$ üzerindeki izdüşümü $\mathbf{og} = \mathbf{P}_1\mathbf{X}_2$ ile gösterilir ve karşılık gelen hata tahmini vektörü $\mathbf{fg} = \mathbf{X}_2^* = \mathbf{M}_1\mathbf{X}_2$ dir. Belki FWL teoreminin pek çok ilginç sonuçlarından biri çok değişkenli durumun tüm ilgili görünümlerinin iki değişkenli duruma indirgenebilmesidir.

İkinci yöntem, $\mathbf{Y}^*$'yı $\mathbf{c}$ de olacak şekilde basit olarak $\mathbf{fg}$ kadar ötelenen, $\mathbf{cg}$ doğru parçasını ihtiva eden doğruyla gösterilen, $\mathbf{X}_2^*$ zın gereni üzerinde regresyonlar. Bu izdüşüm $\mathbf{cb}$ vektörünü ve karşılık gelen hata tahminleri vektörünü, önemsiz bir şekilde $\mathbf{ab}$ vektörünü, verir. Bu teoremin ikinci kısmını açıklar; her iki yöntemin OLS hata tahminleri tamamen aynıdır.



Şekil 4.12.FWLT nin kısa bir geometrik gösterimi

Teoremin birinci kısmı aynı şekil içinde kolayca açıklanabilmekle beraber, Şekil 4.11 haddinden fazla karıştırmaktan kaçınmak için yukarıdan basit olarak görülen Şekil 4.12 ye bakalım. 1. Yöntemden, $\mathbf{X_2\hat{\beta}_2 = od = of.\hat{\beta}_2}$ ve 2. Yöntemden, $\mathbf{X_2\tilde{\beta}_2 = cb = cj\tilde{\beta}_2}$ dır. Şimdi Thales teoremine göre $\mathbf{od \backslash of. = cb \backslash cj}$ dir ve yerine koyarak, teoremin birinci kısmını, yani $\mathbf{\hat{\beta}_2 = \tilde{\beta}_2}$ yı elde ederiz.

4.11. TEOREMİN KISA BİR CEBİRSEL İSPATI

Bütünlük için, teoremin bir standart cebirsel ispatını vereceğiz. Başlama noktası

$$\mathbf{Y = PY + MY = X_1\hat{\beta}_1 + X_2\hat{\beta}_2 + MY}$$

ortogonal ayrışımıdır. Birinci kısmı ispatlamak her iki yanı $\mathbf{X_2'M_1}$ ile çarparız ve

$$\mathbf{X_2'M_1Y = X_2'M_1X_1\hat{\beta}_1 + X_2'M_1X_2\hat{\beta}_2 + X_2'M_1MY}$$

elde ederiz. Tanıma göre, $\mathbf{M_1}$, $\mathbf{X_1}$'ri onun tümleyeni üzerine izdüşürdüğünden ve bu nedenle $\mathbf{M_1X_1 = 0}$ olduğundan, sağ yanın ilk terimi sıfırdır. Önceki gibi aynı nedenlerle, $\mathbf{X_2'M_1M = X_2'M - P_1X_2'M}$ ve $\mathbf{X_2'M = 0}$ olduğundan, üçüncü terimde sıfırdır. O halde

$\mathbf{X_2 M_1 Y = X_2' M_1 X_2 \hat{\beta}_2}$ elde ederiz. Bu denklemin $\hat{\boldsymbol{\beta}}_2$ için çözümü teoremin birinci kısmını ispatlar. İkinci kısmı ispatlamak için, ortogonal ayrışımı $\mathbf{M_1}$ ile çarparız ve

$$\mathbf{M_1 Y = M_1 X_1 \hat{\beta}_1 + M_1 X_2 \hat{\beta}_2 + M_1 M Y}$$

elde ederiz. Yine sağ yandaki birinci terim sıfırdır.

Şimdi üçüncü terim için, $\mathbf{MY}$, $[\mathbf{X_1\ X_2}]$'nin ortogonal tümleyenine aittir, bu nedenle onu $\mathbf{X_1}$'rin ortogonal tümleyeni üzerine ayrıca izdüşürme etkiye sahip değildir, bu nedenle $\mathbf{M_1 Y - M_1 X_2 \hat{\beta}_2 = MY}$ olduğunu gösterir. teoremin birinci kısmından, sol yan $\mathbf{Y}^*$ zın $\mathbf{X}_2^*$ üzerine izdüşümünün hata tahminleridir ve tanıma göre, sağ yan, $\mathbf{Y}$'nin $[\mathbf{X_1, X_2}]$üzerine izdüşümünün hata tahminleridir. Böylece teoremin ikinci kısmı da ispatlanmış olur.

4.12. YORUMLAR VE SEZGİLER

$\mathbf{X_1}$ için kontrol etme fikri: Modelde ya onu yerine koyarız ya da onun etkisini çıkarmak suretiyle ilk olarak ondan kurtuluruz.

Eğer $\mathbf{X_1}$ ve $\mathbf{X_2}$ ortogonal iseler ne olur? Bu taktirde tüm çalışmalarda $\mathbf{X_1' X_2 = 0}$ olduğu görülür.

FWLT' nin uygulamaları

Bu uygulamalar şunlardır:

1) Ortalamalardan sapmalar

2) Caydırma

3) Mevsimsel etkiler

Daha sonra: Çoklu eş lineerlik, yanlı değişkeni silme, panel veri, sabit etkiler tahmini, deneysel değişkenler.

Ortalamadan sapma

Sabit terimli (intercept) basit model

$$\mathbf{Y} = \mathbf{X\beta} + \mathbf{u} = \boldsymbol{\beta}_1\mathbf{1} + [\mathbf{X}_2\mathbf{X}_3...\mathbf{X}_K]\boldsymbol{\beta}_{-1}$$

dir. Burada $\mathbf{1} = (1,1,...,1)'$, $\boldsymbol{\beta}_{-1} = (\boldsymbol{\beta}_2, \boldsymbol{\beta}_3, ..., \boldsymbol{\beta}_K)'$ $\mathbf{X}_k$, $k = 2,...,K$ $\mathbf{X}$'in karşılık gelen sütunlarıdır. $\boldsymbol{\beta}_{-1}$, tahmin etmenin iki yöntemi:

1.Yöntem: $\mathbf{X} = [1', \mathbf{X}_2, ..., \mathbf{X}_K]$ üzerinde $\mathbf{Y}$'yi regresyonlarız.

2.Yöntem: $\mathbf{X}_k$, $k = 2,...,K$ **1** üzerinde izdüşümünün hata tahminlerini elde edelim. Onlara $\mathbf{X}_k^*$ diyelim. $\mathbf{Y}$ ile aynı şeyi yapalım ve onlara $\mathbf{Y}^*$diyelim. $\mathbf{P}_1 = 1(1'1)^{-1}1' = n^{-1}\mathbf{1}$ olduğuna dikkat edelim. Burada $\mathbf{1}$, 1 lerin bir $n \times n$ matrisidir. Burada

$$\mathbf{P}_1\mathbf{X}_1 = \frac{1}{n}\mathbf{1X}_k = (\bar{\mathbf{X}}_k, \bar{\mathbf{X}}_k, ..., \bar{\mathbf{X}}_k)'$$

böylece $\mathbf{X}_k^* = \mathbf{M}_1\mathbf{X}_k = (\mathbf{I} - \mathbf{P}_1)\mathbf{X}_k = \mathbf{X}_k - (\bar{\mathbf{X}}_k, \bar{\mathbf{X}}_k, ..., \bar{\mathbf{X}}_k)'$ tipik $\mathbf{X}_{ik}^* = \mathbf{X}_{ik} - \bar{\mathbf{X}}_k$ elemanlı bir $n \times 1$ vektördür.

Bu nedenle ikinci yöntem:

1) Tüm değişkenleri onların örnek ortalamalarından sapmalar olarak yeniden ifade etmeden,
2) Bu hata tahminlerinin sabit terimsiz standart regresyonu ile çalışmadan ibarettir çalışmaktır.

5.BULGULAR

EN KÜÇÜK KARELER YÖNTEMİ

ve

EN KÜÇÜK KARELERİN GEOMETRİSİ

"Bin söz söyleyeceğine bir resim yap göster."

Regresyon, geometrisiz mükemmel olarak düşünülebilir iken, geometrik kavrayış, örneğin singüler veya neredeyse singüler regresyonlara ve R^2 istatistiğinin yorumlarına ilişkin zorlukların çok daha iyi anlaşılmasını sağlar. En küçük karelerin daha ileri bir yorumu için geometri bilgisi önemlidir.

5.1. TEMELİ OLUŞTURAN GEOMETRİ

$\mathbf{Y}$ ve $\boldsymbol{\varepsilon}$ her ikisi de $n \times 1$, $\mathbf{X}$ $n \times p$ ve $\boldsymbol{\beta}$ $p \times 1$ olmak üzere, en küçük kareler yöntemiyle

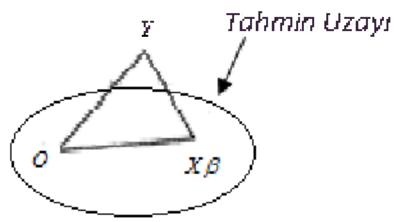
$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \qquad (5.1.1)$$

modelini uydurmak istiyoruz. S diyeceğimiz n-boyutlu bir Öklid ("alışılmış") uzayını düşünelim. $\mathbf{Y}$ vektörünün n tane $Y_1, Y_2 ..., Y_n$ bileşeni bu uzayda bir nokta tanımlar. Onlar aynı zamanda genellikle O(0,0,…,0) orijinini (başlangıcını) Y noktasına birleştirilen doğruyla gösterilen bir vektörü (uzunluk ve yöne sahip doğru) de tanımlar. (Gerçekten aynı uzunluğa sahip olan herhangi bir paralel doğruya da $\mathbf{Y}$ vektörü olarak bakılabilir, fakat daha ziyade O başlangıçlı vektörü düşünürüz.) $\mathbf{X}$ ' in sütunları da n-boyutlu uzaydaki vektörleri tanımlar. Şimdilik $\mathbf{X}$ ' in p tane sütununun tümünün lineer bağımsız olduklarını; yani, onların hiçbirinin diğerlerinden herhangi birinin bir lineer kombinasyonu olarak gösterilemediğini kabul edelim. Bu, $\mathbf{X'X}$ matrisinin tekil olmadığını gösterir. Bu taktirde $\mathbf{X}$ ' in p sayıda sütunu, S ' nin p (<n)-boyutlu bir alt uzayını (ona tahmin uzayı diyeceğiz) tanımlar. Her bir $\mathbf{x}_i$ bir $n \times 1$ vektör ve β_i ler de skalarlar (genellikle $\mathbf{x}_0 = \mathbf{1}$ dir, burada $\mathbf{1}$ tüm elemanları 1 ler olan bir $n \times 1$ vektördür) olmak üzere,

$$\mathbf{X}\boldsymbol{\beta} = \left[\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_{p-1}\right] \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_{p-1} \end{bmatrix} = \beta_0 \mathbf{x}_0 + \beta_1 \mathbf{x}_1 + \ldots + \beta_{p-1} \mathbf{x}_{p-1} \tag{5.1.2}$$

çarpımını düşünelim. Bu, $\mathbf{x}_i{}'$ lerin bir lineer kombinasyonu olarak şekillenmiş bir vektörü tanımlar ve bu nedenle $\mathbf{X}\boldsymbol{\beta}$ tahmin uzayındaki bir vektördür. Kesin olarak $\mathbf{X}\boldsymbol{\beta}$ vektörü β_i ler için seçilen değerlere bağlıdır.

Şimdi aşağıdaki şekli çizebiliriz.



Şekil 5.1. Tahmin Uzayında Genel bir $X\beta$ noktası

İşaretlenen O, $\mathbf{Y}$, $\mathbf{X}\boldsymbol{\beta}$ noktalarının n-boyutlu uzayda bir üçgen oluşturduğunu görüyoruz. Genel olarak, $\mathbf{Y}$ ve $\mathbf{X}$ verildiğinde, açılar β_i lerin değerleriyle belirlenecektir. Üçgenin üç kenarı $\mathbf{Y}$, $\mathbf{X}\boldsymbol{\beta}$ ve $\boldsymbol{\varepsilon} = \mathbf{Y} - \mathbf{X}\boldsymbol{\beta}$ vektörleridir. Böylece (5.1.1) modeli basit olarak, $\mathbf{Y}$ vektörünün, biri tamamen tahmin uzayında olan $\mathbf{X}\boldsymbol{\beta}$ ve diğeri, genel olarak, kısmen tahmin uzayında olmayan $\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}$ gibi iki vektöre ayrılabileceğini söyler. $\boldsymbol{\beta}$ ' yı en küçük kareler yöntemi ile tahmin ettiğimizde $\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$ çözümü

$$S(\boldsymbol{\beta}) = (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}) \tag{5.1.3}$$

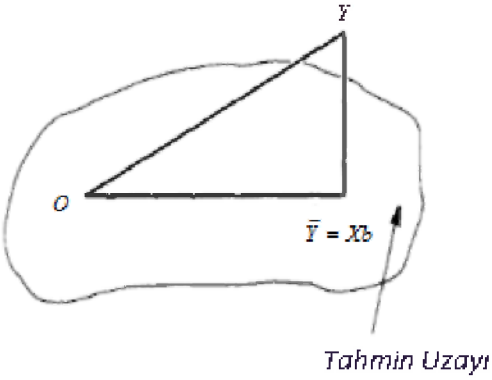
kareler toplamı fonksiyonunu en küçük kılan (minimum yapan) bir çözümdür. Bu noktada eğer $\mathbf{z}$ herhangi bir vektör ise, $\mathbf{z}'\mathbf{z}$ ' nin bu vektörün uzunluğunun karesi olduğu gerçeğini bilmemiz gerekir. Bu nedenle $S(\boldsymbol{\beta})$ kareler toplamı fonksiyonu Şekil 5.1 de $\mathbf{Y}$ ve $\mathbf{X}\boldsymbol{\beta}$ ' yı birleştiren vektörün uzunluğunun karesidir. Ne zaman bu minimumdur? $\boldsymbol{\beta}$,

$\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$ ye eşit küme olduğunda o minimumdur. $\mathbf{Xb} = \hat{\mathbf{Y}}$ vektörünün, 0 orijinini $\mathbf{Y}$' den tahmin uzayına inilen dikmenin ayağına birleştiren vektör olduğu anlaşılır (bkz Şekil (5.2) ve onu Şekil (5.1) ile karşılaştırın). $\boldsymbol{\beta}$ özel b değerini aldığında en küçük kareler değeri yani Şekil (5.1) in üçgeni Şekil (5.2) deki gibi bir dik üçgen olmaktadır. Eğer bu doğru ise, $\hat{\mathbf{Y}} = \mathbf{bX}$ ve $\mathbf{e} = \mathbf{Y} - \mathbf{Xb}$ vektörleri ortogonal (dik) olmalıdır. İki $\mathbf{p}$ ve $\mathbf{q}$ vektörü için, eğer $\mathbf{p'q} = 0 = \mathbf{q'p}$ ise; bu iki vektör ortogonaldir. Burada

$$
\begin{aligned}
\mathbf{0} = (\mathbf{Xb})'\mathbf{e} &= \mathbf{b}'\mathbf{X}'\mathbf{e} \\
&= \mathbf{b}'\mathbf{X}'(\mathbf{Y} - \mathbf{Xb}) \qquad (5.1.4) \\
&= \mathbf{b}'(\mathbf{X}'\mathbf{Y} - \mathbf{X}'\mathbf{Xb})
\end{aligned}
$$

dir. Böylece (5.1.4) deki ortogonallik ihtiyacı ya $\mathbf{b} = \mathbf{0}$ olmasını, ki bu durumda $\mathbf{Y}$ tahmin uzayına ortogonaldir ve $\hat{\mathbf{Y}} = \mathbf{0}$ dır, ya da $\mathbf{X'Xb} = \mathbf{X'Y}$ normal denklemlerinin sağlanmasını gerektirir. $\mathbf{e} = \mathbf{Y} - \mathbf{Xb}$' nın içinde yer aldığı uzaya hata uzayı denir ve hata uzayı $(n-p)$-boyuta sahiptir. p-boyutlu tahmin uzayı ve hata uzayı ikisi birlikte S uzayını oluştururlar. Bu nedenle regresyon modelinin en küçük kareler uyumu S uzayını iki ortogonal uzaya böler ve böylece tahmin uzayındaki her vektör hata uzayındaki her vektöre ortogonaldir.

Örnek 5.1.1: $\mathbf{Y} = (3.1, 2.3, 5.4)', \mathbf{x}_0 = (1,1,1)'$ ve $\mathbf{x}_1 = (2,1,3)'$ olsun. Bu nedenle en küçük kareler yöntemi ile $Y = \beta_0 + \beta_1 X_1 + \varepsilon$ doğrusunu uyduracağız. İlk olarak, geometriyi düşünmeksizin regresyonu uyduralım. Bu regresyon doğrusu $\hat{\mathbf{Y}} = 0.50\mathbf{x}_0 + 1.55\mathbf{x}_1$ olacaktır. Bu taktirde $\hat{\mathbf{Y}} = \mathbf{Xb} = (3.60, 2.05, 5.15)'$ ve $\mathbf{e} = (-0.50, 0.25, 0.25)'$ olduğu gösterilebilir, bu iki vektörün ortogonal oldukları ve bundan başka $\mathbf{e}$, $\mathbf{X}$' in ayrı ayrı her iki sütununa ortogonal olduğundan, (genelde hepsine ortogonal olduğundan) $\mathbf{e}$ her $\boldsymbol{\beta}$ için, $\mathbf{e}$' nin $\mathbf{X}\boldsymbol{\beta} = \beta_0\mathbf{x}_0 + \beta_1\mathbf{x}_1$ biçimindeki herhangi bir vektöre ortogonal olduğu saptanır. Şekil (5.2)' ye benzer bir diyagram üzerinde çeşitli noktaların koordinatlarını yazabiliriz.



Şekil 5.2. $\mathbf{Y}$, $\mathbf{Xb}$ ve $\mathbf{e = Y - Xb}$ vektörlerinin dik üçgeni

5.2. PİSAGOR VE VARYANS ANALİZİ

Bir regresyondan ortaya çıkan her varyans analizi tablosu bir dik üçgenin hipotenüsünün karesinin diğer iki kenarın karelerinin toplamına eşit olduğu hakkındaki Pisagor Teoreminin bir uygulamasıdır. Genellikle Pisagor sonucunun ard arda uygulanmasına ihtiyaç vardır. Şekil (5.2)' yi yeniden göz önüne alalım. Bu şekil

$$\mathbf{Y'Y = \hat{Y}'\hat{Y} + (Y - Xb)'(Y - Xb)} \tag{5.2.1}$$

olduğunu veya genel kareler toplamı= regresyona bağlı kareler toplamı+ hata (kalıntı) kareler toplamı olduğunu ifade eder. (5.2.1) toplamına uygun gelen

$$n = p + (n - p) \tag{5.2.2}$$

serbestlik dereceleri denklemi, $S^{\cdot}$ nin sırasıyla, p ve $(n-p)$ boyutlu iki ortogonal uzaya boyutsal ayrışımına karşılık gelir. Genel kareler toplamının açık bir ayrışımı için $\hat{\mathbf{Y}}$ ve $\mathbf{e}$' nin ortogonalliği esastır. Doğal bir ortogonal ayrışıma sahip olmayan özel bir kareler toplamının bir ayrışımı istendiğinde, ortogonallik ortaya konmalıdır.

Örnek 5.2.1: Alıştırma (5.1.1) deki veriler için (5.2.1) ve (5.2.2) denklemlerinin sırasıyla,

44.06=43.685+0.375

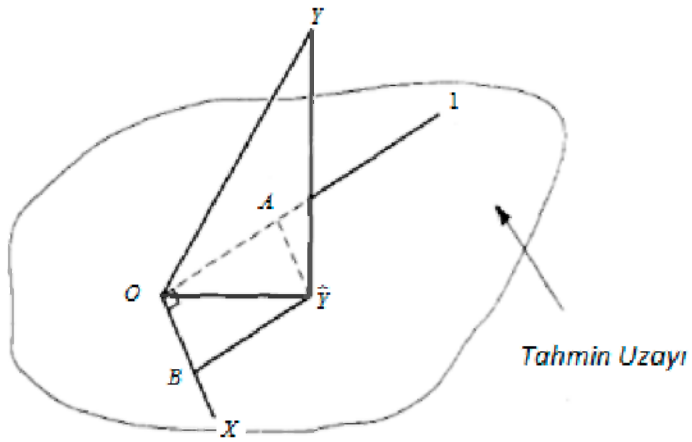
3=2+1

denkliklerine karşılık geldikleri gösterilebilir.

Bir Regresyon Kareler Toplamının Ayrıca Ayrışımı

Eğer tahmin uzayını tanımlayan $\mathbf{X}_{:}$ vektörleri bir başkasına tümden ortogonal olursa, regresyon kareler toplamı hemen iki ortogonal parçaya ayrılabilir. Örneğin, $\mathbf{X = (1, x)}$ ve

$\mathbf{1'x} = \mathbf{0} = \mathbf{x'1}$ olduğunu farz edelim. $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ modeli x' e karşı bir Y doğrusudur. Şekil (5.3) iki ortogonal $\mathbf{1}$ ve $\mathbf{x}$ vektörleriyle tanımlanan tahmin uzayını gösterir.



Şekil 5.3. Model fonksiyonundaki ortogonal vektörler

Bu vektörler ortogonal olduklarından, $\hat{\mathbf{Y}}$ ' dan bu doğrulara çizilen dikmeler (dik doğrular) bir dikdörtgen oluştururlar, bu nedenle

$$O\hat{Y}^2 = OA^2 + OB^2 \tag{5.2.3}$$

(veya $OA^2 + A\hat{Y}^2$, vs.) dir. $\mathbf{1}$ ve $\mathbf{x}$ taban vektörlerinin ortogonalliğinden dolayı bölünme tektir; bu sonuç ortogonal taban vektörlerinin herhangi bir sayısına açık bir şekilde genişletilir ki; o durumda Şekil (5.4)' ün dikdörtgeni p-boyutlu tahmin uzayında bir dikdörtgen blok olur.

Örnek 5.2.2:

$$\mathbf{Y'} = (6, 9, 17, 18, 26) \text{ ve } \mathbf{X'} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ -2 & -1 & 0 & 1 & 2 \end{bmatrix}$$

olmak üzere, en küçük kareler regresyon problemi için $\mathbf{X'Xb} = \mathbf{X'Y}$ normal denklemlerinin, iki ayrı parçaya ayrıştığını ve Şekil (5.4)' deki A ve B' nin, her bir parçaya ayrı ayrı bakmaktan gelen iki ayrı $\hat{\mathbf{Y}}$ olduğunu gösterelim. Böylece bu veriler için (5.2.1) ve (5.2.2) bağıntıları 1406=1395+10.7 ve 4=2+2 ortogonal değerlerini ortaya koyar.

Normal Denklemlerin Bir Parçalanışı

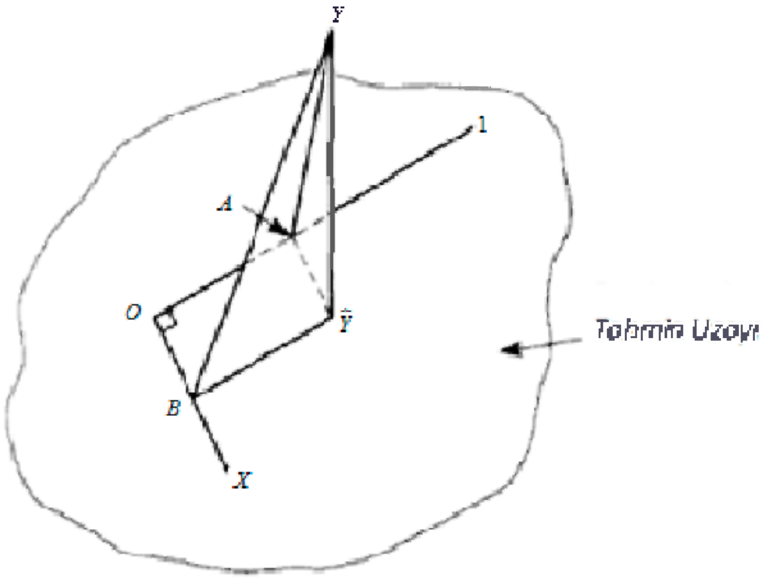
Bir ek ayrıntı ile Şekil (5.3)' ü yeniden çizelim. Şekil (5.4)' de OY ve $O\hat{Y}$ doğrularının silindiğini fakat YA ve YB doğrularının çizilmiş olduklarını görelim. Diğer hususlarda diyagramlar aynı olmaya niyetlidir. İşaretlenen O, A, $\hat{Y}$ ve B noktaları bir dikdörtgen oluştururlar, ve bu nedenle görüldüğü gibi, $O\hat{Y}^2 = OA^2 + OB^2$ dir. Aynı zamanda, YA,

OA'ya ve YB de OB'ye diktirler. Bu nedenle OA, "$\mathbf{Y}$' nin yalnız $\mathbf{1}$ üzerindeki regresyonu için $\hat{\mathbf{Y}}$' dır." ve OB de "$\mathbf{Y}$' nin yalnız $\mathbf{x}$ üzerindeki regresyonu için $\hat{\mathbf{Y}}$' dır." (burada) $\mathbf{1}$ ve $\mathbf{x}$ ortogonal olduklarından, bu olur. Sonuç, karşılıklı olarak ortogonal olmaları şartıyla herhangi sayıdaki $\mathbf{X}$ vektörlerine de genişler.

Örnek 5.2.3: Alıştırma (5.2.2) için gösterebiliriz ki, $OA^2 = 1155.2, OB^2 = 240.1$ dir ve onların toplamı $O\hat{Y}^2 = 1395.3$ dür. Aynı zamanda gösterilebilir ki, $OA = (15.2, 15.2, 15.2, 15.2, 15.2)'$ vektörüdür, OB' de $(-9.8, -4.9, 0, 4.9, 9.8)'$ vektörüdür ve $O\hat{Y}$ bu ortogonal vektörlerin toplamıdır.

Genel Olarak X' in Vektörlerini (Sütunlarını) Ortogonalleştirme

$\mathbf{X}$' in vektörleri (sütunları) karşılıklı olarak ortogonal olmadıkları zaman bile sık sık bir regresyon kareler toplamının bileşenlere ayrılması istenir.



Şekil 5.4. Model vektörleri ortogonal olmadıklarında normal denklemlerin ortogonal ayrışımı

Böyle bir ayırma dizide her biri daha önce ondan önceki elemanlara ek olarak bir kareler toplamı olan, bir dizisel kareler toplamı kümesini temin eder. Geometrik olarak bunu yapmak için bir ardışık ortogonalleştirme yapmamız gerekir. Şekil (5.5) $\mathbf{1'x} \neq \mathbf{0}$ olmak üzere iki $\mathbf{1} = \mathbf{x}_0$ ve $\mathbf{x}$ vektörü için düşünceyi açıklar. Başlangıç olarak bir vektörü seçeriz ($\mathbf{1}$ vektörünü seçeceğiz fakat bir diğeri de olabilirdi.) ve bundan sonra $\mathbf{x}$' in $\mathbf{1}$' ye ortogonal olan kısmını (2. Vektörü) yani $\mathbf{x}_0$ vektörünü ortaya koyarız. Bu nasıl olacak? Regresyonun standart sonucunu yani kalanların (rezidülerin-hataların tahmininin) uygun değerlere ortogonal olması durumunu kullanırız. $\mathbf{x}'$ i $\mathbf{1}$ üzerine (onu bir "Y" gibi düşünerek) uygun bir hale getiririz ve $b = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} = (\mathbf{1}'\mathbf{1})^{-1}\mathbf{1}'\mathbf{x} = \bar{x}$ olmak üzere $\hat{\mathbf{x}} = b\mathbf{1}$ doğruluğunu kabullenerek hata tahminlerini ortaya koyarız. Böylece $\mathbf{x}_{.0} = \mathbf{x} - \bar{x}\mathbf{1}$ dir. $\mathbf{x}'_{.0}\mathbf{1} = \mathbf{0}$ olduğunu bu nedenle de iki vektörün ortogonal olduğunu kontrol etmek kolaydır. Şimdi

$$\begin{pmatrix} b_0 \\ b_1 \end{pmatrix} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$$

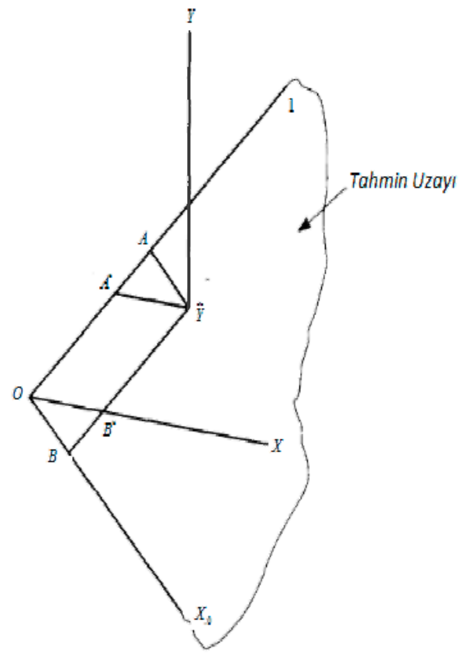
orijinal regresyon denklemi

$$\hat{\mathbf{Y}} = b_0\mathbf{1} + b_1\mathbf{x} \tag{5.2.4}$$

biçimindeydi. Ortogonalleştirilmiş regresyon denklemi

$$\begin{aligned} \hat{\mathbf{Y}} &= \bar{Y}\mathbf{1} + b_1(\mathbf{x} - \bar{x}\mathbf{1}) \\ &= \bar{Y}\mathbf{1} + b_1\mathbf{x}_{.0} \end{aligned} \tag{5.2.5}$$

dır. Değerleri aynı olacağından (5.2.4) ve (5.2.5) denklemlerinin her ikisinde de bu sembolü kullandık. Şekil (5.4) ve (5.5) de elde ettiğimiz şey Şekil (5.5) deki aynı $O\hat{Y}$ vektörünün farklı iki tanımıdır. Bu vektör, $\hat{\mathbf{Y}}$ ' ya ortogonal olmayan OA^* ve OB^* vektörlerinin toplamı olarak ya da OA ve OB ortogonal vektörlerinin toplamı olarak gösterilebilir. Her iki tanım da geçerlidir. Ancak ikincisi $O\hat{Y}^2$ kareler toplamının Pisagor sonucu yoluyla $OA^2 + OB^2$ olarak ayrıştırmaya izin verir. $\mathbf{Y}$ ve $\hat{\mathbf{Y}}$ ' nın yaptığımız şeyden etkilenmediğine dikkat edelim. Yaptığımızın tümü uzaydaki vektörlerin bir lineer kombinasyonuna göre, $\hat{\mathbf{Y}}$ ' nın değişik tanımıdır. Örnek 5 ve 6 ilk olarak iki vektörden birini seçebileceğimiz gerçeğini açıklar.



Şekil 5.5. $\mathbf{1}$ ’ ye göre ikinci bir açıklayıcıyı ortogonalleştirme

Örnek 5.2.4: Örnek (5.1.1) in verileri için, yani,

$\mathbf{Y}=(3.1,2.3,5.4)', \mathbf{x}_0=(1,1,1)', \mathbf{x}=(2,1,3)'$ ve $\hat{\mathbf{Y}}=0.50\mathbf{1}+1.55\mathbf{x}$ için $\mathbf{x}_{.0}=(0,-1,1)'$

olduğunu ve $\mathbf{Y}=3.5\ 1+1.55\mathbf{x}_{.0}$ ’ nun aynı uygun değerler ve hata tahminlerini ortaya koyduğunu gösteririz. Uygunluğunu elde etmek için (çok ya da az) onun önceden $OA^*, OA, OB, OB^*, O\hat{Y}$ ve $Y\hat{Y}$ uzunluklarını bularak Şekil 5.5 biçiminde bir doğruluk diyagramı çizelim. YA ’ nın, $\mathbf{Y}$ ’ nın $\mathbf{1}$ üzerindeki bir izdüşümü olduğuna, YB ’ nin, $\mathbf{Y}$ ’ nin $\mathbf{x}_{.0}$ üzerindeki bir izdüşümü olduğuna dikkat edelim.

$$O\hat{Y}^2 = OA^2 + OB^2 \neq OA^{*2} + OB^{*2}$$

$$43.685 = 38.88 + 4.805 \neq 0.125 + 33.635 \qquad (5.2.6)$$

$OB^2 = SS(b_1 / b_0) = 4.805$ iken, OA^2 niceliği $OA^2 = SS(b_0) = n\overline{Y}^2 = 38.88$ dır. Yukarıda ilk olarak $\mathbf{x}_0 = \mathbf{1}$ seçtiğimize ve $\mathbf{x}_{.0}$ ’ ı bulduğumuza dikkat edelim.

Örnek 5.2.5: Örnek 5.2.4’ün verileri için ilk olarak $\mathbf{x}$ ’ i seçelim ve $\mathbf{x}_1'\mathbf{x}=0$ olacak şekilde $\mathbf{x}_1$ ’ i belirleyelim. $\hat{\mathbf{Y}}$ ’ dan $\mathbf{x}_1$ ’ e çizilen dikmelerin ayağı C ve $\hat{\mathbf{Y}}$ ’ dan $\mathbf{x}'$ e çizilen dikmelerin ayağı D olsun ve A yerine C, B yerine D koyalım. $43.685 = O\hat{Y}^2 = OC^2 + OD^2 = OA^2 + OB^2$ fakat ayrışmanın iki durumda farklı olduğunu yani $OC \neq OA$ ve $OB \neq OD$ olduğunu gösterelim. Özellikle $0.107 = OC^2 \neq OA^2 = 38.88$ ve $43.578 = OD^2 \neq OB^2 = 4.805$ dır. Yukarıda yaptıklarımızda önceden cebirsel olarak bildiğimiz bir gerçeği geometriksel olarak gördük. **X** deki tüm vektörler karşılıklı olarak ortogonal olmadıkça dizisel kareler toplamı açıklayıcı değişkenin elemanları için seçilen diziye bağlı olacaktır. İlk olarak değişkenin bir elemanı seçildikten sonra diğerleri onlardan önce kaydedilenlerin tümüne sırasıyla ortogonaldirler. Bunda büyük bir öneme sahip olan noktayı yeniden vurgulayalım. Kayıt dizisi ne olursa olsun $\hat{\mathbf{Y}}$ tektir ve sabittir. Ortogonalleştirilen dizide seçtiğimiz vektörlere göre sadece farklı $\hat{\mathbf{Y}}$ tanımları veririz. Hepsi bu kadar.

5.3. GENEL REGRESYON İÇİN VARYANS ANALİZİ VE F TESTİ

Modelin $\mathbf{Y} = \beta_0 \mathbf{1} + \beta_1 \mathbf{x} + \boldsymbol{\varepsilon}$ ve $\mathbf{1'x} \neq 0$ olduğu Şekil (5.5) durumunu yeniden göz önüne alalım. Standart varyans analizi tablosu çizelge 5.1 şeklini alır.

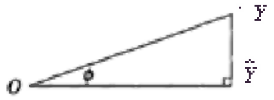
Çizelge 5.1. Bir doğru uyumu için varyans analizi tablosu

Kaynak	*KT*	*Sd*	*KO*	*F*
b_0	$OA^2 = n\bar{Y}^2$	*1*	--	--
b_1 / b_0	$OB^2 = S_{XY}^2 / S_{XX}$	*1*	S_{XY}^2 / S_{XX}	$F = S_{XY}^2 / (s^2 S_{XX})$ $\{OB^2/1\}/\{Y\hat{Y}^2/(n-2)\}$
Hata tahmini	$Y\hat{Y}^2 = \text{çıkarmayla}$	*n-2*	s^2	
Toplam	$OY^2 = \mathbf{Y'Y}$	*n*		

Varyans analizi tablosunun basit olarak $OA^2 + OB^2 + Y\hat{Y}^2 = OY^2$ ile izlenen bir

$$O\hat{Y}^2 = OA^2 + OB^2$$

pisagor ayrışımı olduğuna ve $\beta_1 \neq 0$ ' a karşı $H_0 : \beta_1 = 0$ için F testinin basit olarak "$Y\hat{Y}$ ' nın serbestlik derecesi başına uzunluk karesi" ne karşı "OB' nin serbestlik derecesi başına uzunluk karesi" nin bir karşılaştırması olduğuna dikkat edelim.



Şekil 5.6. β_0 olmadığında bile, regresyonun hiç olmaması için F testinin basit geometrisi

Bu taktirde OB^2 ' nin daha küçükten ziyade daha büyük değerleri H_0' ın reddine sebep olur. Bu nedenle aslında B, s boyuta kıyasla, O' ya yakın mıdır? (H_0' ı reddetme) sorusunu soruyoruz veya B s boyutta O' ya yakın mıdır? (H_0' ı reddet) sorusunu soruyoruz. [H_0 doğru olduğunda F- oranının F$(1, n-2)$ dağılımını gösterdiği cebirsel olarak tespit edilebilir, fakat o aynı zamanda, aşağıda verilen bir geometrik yoruma da sahiptir.]

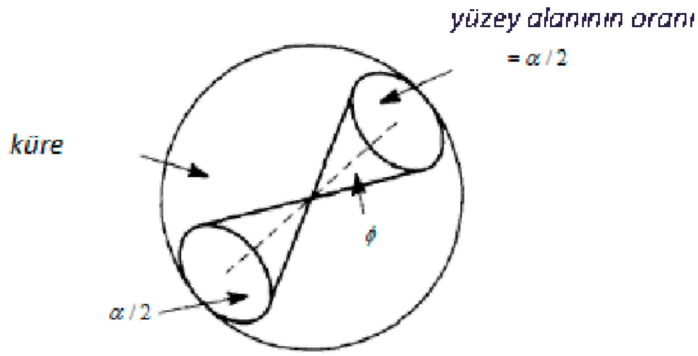
$H_1 : \beta_0 = \beta_1 \neq 0$ hipotezine karşı $H_0 : \beta_0 = \beta_1 = 0$ hipotezi için bir F testi aynı şekilde

$$F = \left\{(OA^2 + OB^2)/\sqrt{2}\right\} / \left\{Y\hat{Y}^2/(n-2)\right\}$$

olmasını gerektirecektir. Bu, $O\hat{Y}^2$ nin "serbestlik derecesi başına uzunluğunun" $Y\hat{Y}^2$ ile bir karşılaştırmasını gerektirdiğine dikkat edelim (bkz. Şekil 5.7). Anlamlı regresyon $Y\hat{Y}$ 'ya kıyasla $O\hat{Y}$ ' nın daha büyük olması durumunda olacak. Bu Şekil 5.7 deki θ açısının küçük olmasını gerektirir. Eğer OY yi bir sabit eksen olarak ve $O\hat{Y}$ ' yı OY ile aynı θ açısında herhangi bir yerde olabilen, (pozitif veya negatif yönde), uygun vektörün bir mümkün olabilen durumu olarak düşünürsek, F-testinin bir geometrik yorumuna sahip oluruz (bkz. Şekil 5.8). F-istatistiğinin kuyruk olasılığı herhangi bir yarıçaplı bir küre üzerinde, Y etrafında dönen $\hat{Y}$ olarak tanımlanan iki dairesel şapkanın oransal alanıdır. Eğer test β_0 ve β_1' in 0 olmayan değerleri üzerinde olsaydı ne olurdu? $H_0 : \beta_0 = \beta_{00}$ ve $\beta_1 = \beta_{10}$ hipotezini test etmek istediğimizi varsayalım. Bu taktirde $\mathbf{Y}_0 = \beta_{00}\mathbf{1} + \beta_{10}\mathbf{x}$ ile tanımlanan O^* noktası yukarıda O ile yer değiştirecek. Fakat geometrinin diğer elemanları esas itibariyle korunacaktır (bkz. Şekil 5.9).

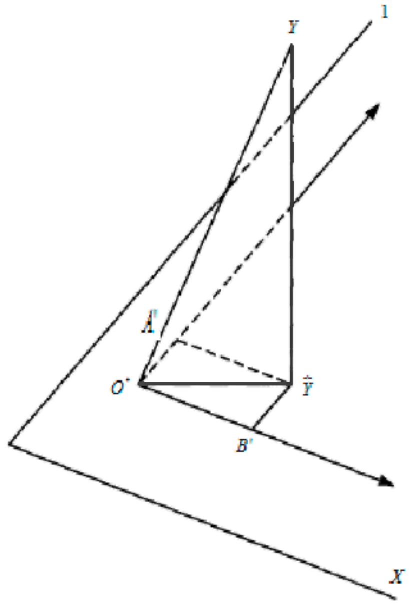
5.4. $X'X$ ' İN SİNGÜLER OLDUĞU DURUM: BİR ÖRNEK

Bir regresyonda $\mathbf{X'X}$ singüler olduğunda, yani $\det(\mathbf{X'X}) = 0$ olduğunda geometri ne olur? Cevap kısadır ve bu cevabın anlaşılması uygulamada böyle problemlerle karşılaşma korkusunu azaltacaktır. Şekil 5.7' ye bakalım. Üç boyutlu çatıda tartışılmış olmakla beraber, tartışılan sonuçlar geneldir. Tahmin uzayının gösterilen düzlem olduğunu fakat $\mathbf{X} = (\mathbf{1}, \mathbf{x}_1, \mathbf{x}_2)$ matrisinin üç vektörden oluştuğunu farz edelim. **X'** deki üç vektörün bağımlı olmaları nedeniyle, **X** matrisinin singüler olduğu açıktır. Eğer bu vektörlerden herhangi ikisinin düzlem tanımladığını yani üç vektörün farklı olduklarını farzedersek, bu taktirde üçüncü vektör diğer ikisinin bir lineer kombinasyonu olmalıdır.



Şekil 5.7. F-testi olasılığının bir geometrik yorumu

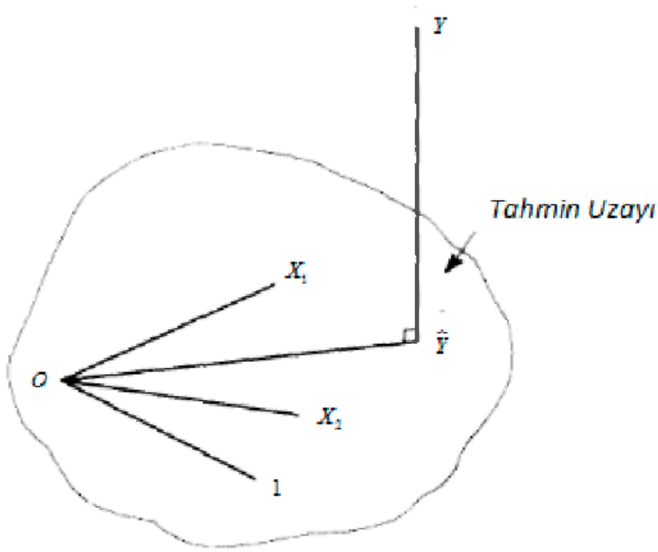
Şimdi $(\mathbf{X'X})^{-1}$ mevcut olmadığından, bazen yanlışlıkla bu şartlar altında en küçük kareler probleminin çözüme sahip olmadığı düşünülür.



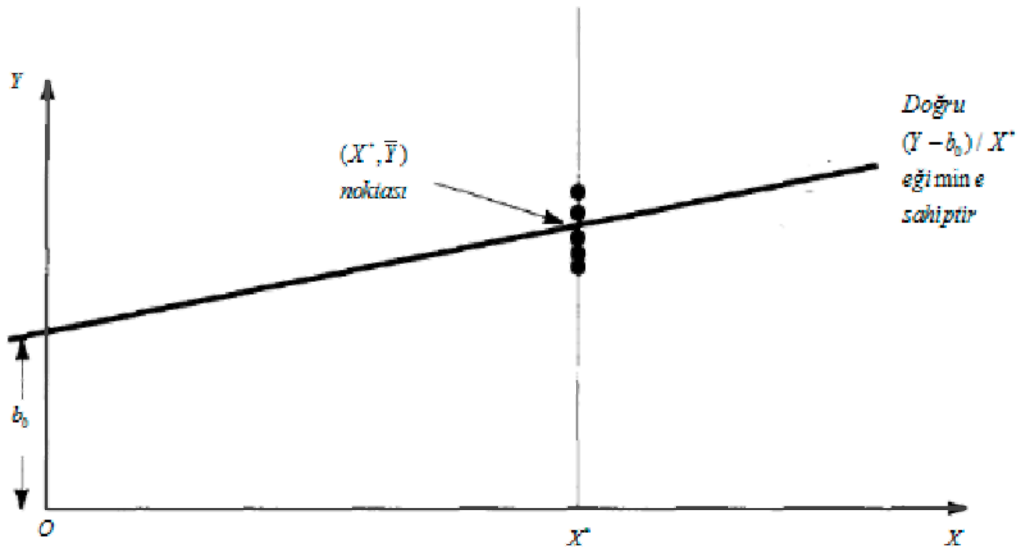
Şekil 5.8. $(\beta_0, \beta_1) = (\beta_{00}, \beta_{10})$ eşitliğinin test edilmesi

Her zaman tek olduğu halde, şimdi bir tek en küçük kareler çözümü olmadığı Şekil 5.9 dan görülebilir. Yani tahmin uzayı üzerine, $\hat{\mathbf{Y}}$ tek olmak üzere, $\hat{\mathbf{Y}}$ ayaklı bir dikmeyi inebiliriz ve böylece $\mathbf{X}$' in tüm sütunlarına ortogonal olan bir tek $\mathbf{Y} - \hat{\mathbf{Y}}$ vektörüne sahip oluruz. $\hat{\mathbf{Y}}$ nın $\mathbf{1}, \mathbf{x}_1, \mathbf{x}_2$' ye göre tanımı neden tek değildir? Oldukça fazla taban vektörlere sahip olduğumuzdan, $\hat{\mathbf{Y}}$' yı sonsuz sayıda ifade edebilme yolu vardır. Normal denklemler mevcuttur ve çözülebilirdir. Fakat parametre tahminleri için çözüm tek değildir. Bunu, mümkün olabilir en küçük örnek olan iki parametre ve her biri diğerinin katları olan iki $\mathbf{1}$ ve $\mathbf{x}$ vektörleri örneğiyle açıklayalım.

Örnek 5.4.1: $\mathbf{X} = \mathbf{X}^{*}$ da n tane veri değerine sahip olduğumuzu farz edelim (bkz. Şekil 5.10 herhangi bir n kullanılabilmekle beraber, orada n=5 tir). En küçük kareler vasıtasıyla $\mathbf{Y} = \beta_0 + \beta_1 \mathbf{X} + \varepsilon$ doğrusunu uydurmayı düşünelim. Aşikar olarak, sonsuz sayıda çözüm vardır. $(\mathbf{X}, \mathbf{Y}) = (\mathbf{X}^{*}, \overline{\mathbf{Y}})$ noktasından geçen herhangi bir doğru bir en küçük kareler uyumunu sağlayacaktır. Geometrik olarak problem, tahmin uzayını tanımlayan, biri diğerinin katları olan iki $\mathbf{1}$ ve $\mathbf{x} = \mathbf{X}^{*}\mathbf{1}$ vektörleridir.



Şekil 5.9. Üç vektör ya da daha birçok vektör tarafından tanımlanan iki boyutlu bir vektör uzayı



Şekil 5.10. Bir tek $\mathbf{Y}$ için çoklu tanımlara sahip singüler regresyon problemi

Böylece $\mathbf{Y}$' den tahmin uzayına inilen dikmenin ayağı, $\mathbf{1}$ ve $\mathbf{X}^*\mathbf{1}$ vektörlerinin sonsuz sayıdaki lineer kombinasyonlarıyla belirlenebilir. Bununla beraber $\hat{\mathbf{Y}} = \overline{\mathbf{Y}}\mathbf{1}$' nin bir tek olduğuna dikkat edelim. Şimdi iki $\mathbf{X'Xb} = \mathbf{X'Y}$ normal denklemini aşağıda

$$\begin{aligned} nb_0 + n\mathbf{X}^* b_1 &= n\overline{\mathbf{Y}} \\ n\mathbf{X}^* b_0 + n\mathbf{X}^{*2} b_1 &= n\mathbf{X}^* \overline{\mathbf{Y}} \end{aligned} \qquad (5.4.1)$$

olarak yazarız ve derhal onların bağımsız olmadıklarını fakat herhangi bir b_0 için

($b_0, b_1 = (\overline{\mathbf{Y}} - b_0) / \mathbf{X}^*$) çözümüne sahip olan bir denklem olduğunu görürüz.

Böylece uygun doğru, tümü $(\mathbf{X},\mathbf{Y}) = (\mathbf{X}^*, \overline{\mathbf{Y}})$ noktasından geçen, farklı eğilimli ve sabit terimli sonsuz sayıda doğruyu gösteren

$$\hat{\mathbf{Y}} = b_0 + b_1\mathbf{X} = b_0 + \mathbf{X}(\overline{\mathbf{Y}} - b_0) / \mathbf{X}^* \tag{5.4.2}$$

doğrusudur b_0' a verilen değer ne olursa olsun, bir en küçük kareler çözümüne sahibiz. b_0 hangi değeri alırsa alsın $\mathbf{X} = \mathbf{X}^*$ da $\hat{\mathbf{Y}} = \overline{Y}\mathbf{1}$ dir ve tektir.

5.5. GENEL REGRESYON DURUMUNDA ORTOGONALLEŞTİRME

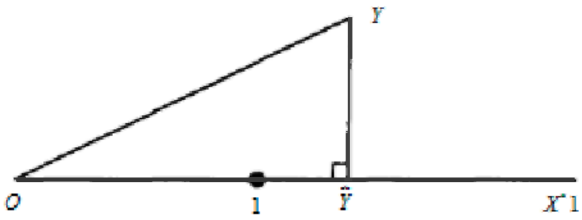
Bir $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ genel lineer modelinin uyumunu ele alacağız ve **X**' in genelde ortogonal olmayan iki kısıma herhangi bir parçalanışını, $\mathbf{X} = (\mathbf{Z}_1, \mathbf{Z}_2)$ ile göstereceğiz. $\boldsymbol{\beta}$ ' yı **X**' in parçalanışına uygun gelecek bir şekilde $\boldsymbol{\beta}' = (\boldsymbol{\theta}_1', \boldsymbol{\theta}_2')$ olarak parçalayacağız. Böylece, **Y** $n \times 1$, $\mathbf{Z}_1$ nxp_1, $\boldsymbol{\theta}_1$ p_1x1, $\mathbf{Z}_2$ nxp_2 ve $\boldsymbol{\theta}_2$ p_2x1 olmak üzere,

$$\mathbf{Y} = \mathbf{Z}_1\boldsymbol{\theta}_1 + \mathbf{Z}_2\boldsymbol{\theta}_2 + \boldsymbol{\varepsilon} \tag{5.5.1}$$

yazarız. Eğer sadece modelin $\mathbf{Z}_1\boldsymbol{\theta}_1$ kısmına uyarsak, p_1 izdüşüm matrisi, yani **Y**'yi **Z**'nin sütunları tarafından tanımlanan tahmin uzayına izdüşüren matris, olmak üzere,

$$\hat{\mathbf{Y}} = \mathbf{Z}_1(\mathbf{Z}_1'\mathbf{Z}_1)^{-1}\mathbf{Z}_1'\mathbf{Y} = \mathbf{P}_1\mathbf{Y} \tag{5.5.2}$$

elde ederiz. Hata vektörünün tahmini $\mathbf{e} = \mathbf{Y} - \hat{\mathbf{Y}} = (\mathbf{I} - \mathbf{P}_1)\mathbf{Y}$ dır. Yukarıdaki müzakerede, **Y**' yi $\hat{\mathbf{Y}}$ ' ya çevirdiğinden dolayı, $\mathbf{P}_1$'re şapka-tahmin matrisi diyeceğiz. Onun geometriksel özelliklerini vurgulamak için şimdi ona yeni ismiyle hitap edelim.



Şekil 5.11. Tamamen 1 ve $\mathbf{X}^{..}1$ vektörleri tarafından tanımlanan tahmin uzayı

$\mathbf{P}_1$ simetrik (yani $\mathbf{P}_1' = \mathbf{P}_1$) ve idempotent (yani $\mathbf{P}_1^2 = \mathbf{P}_1$) olduğundan,

$$\begin{aligned}\hat{\mathbf{Y}}'\mathbf{e} &= \mathbf{Y}'\mathbf{P}_1'(\mathbf{I}-\mathbf{P}_1)\mathbf{Y} \\ &= \mathbf{Y}'(\mathbf{P}_1' - \mathbf{P}_1'\mathbf{P}_1)\mathbf{Y} \\ &= 0\end{aligned} \tag{5.5.3}$$

olduğu açıktır. Bu hesaplama farklı notasyonlarla (5.1.4) ün bir tekrarıdır. $\mathbf{Z}_2, \mathbf{Z}_1$ ve ortogonal olduğunda, $\mathbf{Y}-\hat{\mathbf{Y}}$' ya benzeşmeyle ,

$$\begin{aligned}\mathbf{Z}_{2.1} &= \mathbf{Z}_2 - \hat{\mathbf{Z}}_2 \\ &= (\mathbf{I}-\mathbf{P}_1)\mathbf{Z}_2 \\ &= \mathbf{Z}_2 - \mathbf{Z}_1(\mathbf{Z}_1'\mathbf{Z}_1)^{-1}\mathbf{Z}_1'\mathbf{Z}_2 \\ &= \mathbf{Z}_2 - \mathbf{Z}_1\mathbf{A}\end{aligned} \tag{5.5.4}$$

olur. A ya genellikle alias veya bias (yan) matris denir.

Örnek 5.5.1: $\mathbf{Z}_{21}$ ve $\mathbf{Z}_1$' rın ortogonal matrisler olduklarını gösteriniz. Şimdi tam modeli ortogonalleştirilmiş (dik duruma getirilmiş) biçimde yazabiliriz. Böylece modelin ortogonalleştirilmiş biçimi

$$\begin{aligned}\mathbf{Y} &= \mathbf{Z}_1\theta_1 + \mathbf{Z}_2\theta_2 + \varepsilon \\ &= \mathbf{Z}_1(\theta_1 + \mathbf{A}\theta_2) + (\mathbf{Z}_2 - \mathbf{Z}_1\mathbf{A})\theta_2 + \varepsilon \\ &= \mathbf{Z}_1\theta_1 + \mathbf{Z}_{2.1}\theta_2 + \varepsilon\end{aligned} \tag{5.5.5}$$

olacaktır. $\hat{\boldsymbol{\theta}} = (\mathbf{Z}_1'\mathbf{Z}_1)^{-1}\mathbf{Z}_1'\mathbf{Y}$' nin θ_1' ri tam tahmin etmediğini ancak $\boldsymbol{\theta}_1 + \mathbf{A}\boldsymbol{\theta}_2$' yi tahmin ettiğini hemen görürüz. Bu nedenle eğer yalnız $\mathbf{Y} = \boldsymbol{\theta}_1\mathbf{Z}_1 + \boldsymbol{\varepsilon}$ modeli uydurulursa $\boldsymbol{\theta}_1$' in tahminlerinde $\mathbf{A}\boldsymbol{\theta}_2$ yanlılıkları ortaya çıkar.

Örnek 5.5.2: $E(\mathbf{Y})=\theta_1\mathbf{Z}_1+\theta_2\mathbf{Z}_2$ olmak üzere bunu, $\mathbf{E(\hat{Y})=E(Z_1\theta)}$ değerlendirmesiyle gösteririz. Tam regresyon için, $\boldsymbol{\eta}=\mathbf{Z}_1\theta_1+\mathbf{Z}_2\theta_2$ olmak üzere, çizelge 5.2 varyans analizi tablosuna sahip oluruz.

Çizelge 5.2. Ortogonalleştirilmiş genel regresyon için varyans analizi tablosu

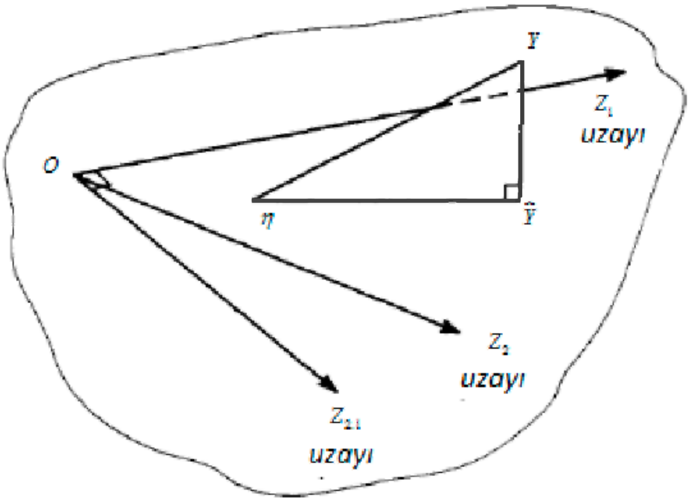
Kaynak	*Sd*	*KT*
Yalnız $\mathbf{Z}_1$ *için tepkime*	p_1	$(\boldsymbol{\theta}_1-\hat{\boldsymbol{\theta}}_1)'\mathbf{Z}_1'\mathbf{Z}_1(\boldsymbol{\theta}_1-\hat{\boldsymbol{\theta}}_1)$
Ekstra $\mathbf{Z}_2$ *için*	p_2	$(\boldsymbol{\theta}_2-\hat{\boldsymbol{\theta}}_2)'\mathbf{Z}_{2.1}'\mathbf{Z}_{2.1}(\boldsymbol{\theta}_2-\hat{\boldsymbol{\theta}}_2)$
Hata tahmini	$n-p_1-p_2$	*Çıkarmayla*
Toplam	n	$(\mathbf{Y}-\boldsymbol{\eta})'(\mathbf{Y}-\boldsymbol{\eta})$

Aşağıdaki noktalara dikkat edelim:

1. $\mathbf{Y}=\mathbf{Z}_1\theta_1+\mathbf{Z}_{2.1}\theta_2+\boldsymbol{\varepsilon}$ modelini uydurarak elde edilen $\hat{\theta}$ ile $\mathbf{Y}=\mathbf{Z}_1\theta_1+\boldsymbol{\varepsilon}$ modelini uydurmak suretiyle elde edilen $\hat{\theta}_1$ aynıdır.

2. $\mathbf{Y}=\mathbf{Z}_1\theta_1+\mathbf{Z}_2\theta_2+\boldsymbol{\varepsilon}$ modelini uydurarak elde edilen $\hat{\theta}_2$ ile $\mathbf{Y}=\mathbf{Z}_{2.1}\theta_2+\boldsymbol{\varepsilon}$ modelini uydurarak elde edilen $\hat{\theta}_2$ aynıdır.

3. Varyans analizi tablosundaki θ_1 ve θ_2 değerleri "test değerleri" olarak düşünülebilir ve geometri, orijinden $\boldsymbol{\eta}=\mathbf{Z}_1\theta_1+\mathbf{Z}_2\theta_2$ ye hareketle biraz değiştirilir (bkz Şekil 5.12).



Şekil 5.12. $\hat{\mathbf{Y}}$ ya ortogonal olmayan $\mathbf{Z}_1$ ve $\mathbf{Z}_2$ uzayları tarafından gerilen uzayların bir lineer kombinasyonu olarak ifade edilebilir ya da ortogonal olan $\mathbf{Z}_1$ ve $\mathbf{Z}_{2\,1}$ uzayları tarafından gerilen uzayların bir lineer kombinasyonu olarak ifade edilebilir.

4. Modelde $\mathbf{Z}_2\boldsymbol{\theta}_2$ ye olan ihtiyaç "$\mathbf{Z}_2$ için fazladan" büyük bir kareler toplamıyla gösterilecek.
5. $\mathbf{Z}_2$ için fazladan kareler toplamı $\mathbf{Y} = \mathbf{Z}_1\boldsymbol{\theta}_1 + \mathbf{Z}_2\boldsymbol{\theta}_2 + \boldsymbol{\varepsilon}$ modelini sonra da $\mathbf{Y} = \mathbf{Z}_1\boldsymbol{\theta}_1 + \boldsymbol{\varepsilon}$

Modelini uydurarak ve

a) İki regresyon kareler toplamı,
b) Her ikisi de $n\overline{\mathbf{Y}}^2$ ile düzeltilmiş iki regresyon kareler toplamı,
c) Ters sırada (pozitif işaretli bir sonuç vermek için) iki hata kareler toplamı arasındaki farkı bulmak suretiyle de elde edilebilir. Bunun geometrisi hakkında daha fazlası ilerideki bölümde verilecektir.

5.6. BİR M MATRİSİNİN SÜTUN UZAYI VE SIFIR UZAYI

Bir **M** matrisinin, $R(\mathbf{M})$ ile gösterilen, sütun uzayı, **M**' nin sütunlarıyla tanımlanan tüm vektörlerin uzayıdır. Uzayın boyutu **M**' nin sütun rankıdır ve Cr(**M**) ile gösterilir. Cr(**M**) , **M'** nin lineer bağımsız sütunlarının sayısıdır. (Bu nedenle M deki sütunların sayısı onların tanımladığı uzayın boyutunu aşar.) **M'** nin, $N(\mathbf{M})$ ile gösterilen, sıfır uzayı, **Mv**=0 olacak şekilde tüm **v** vektörlerinin oluşturduğu uzaydan ibarettir. Yani, **M'** nin satırlarına ortogonal olan tüm vektörlerin sıfır uzayını tanımlamak istersek, $N(\mathbf{M}')$ yazmalıyız.

İzdüşüm Matrisleri

E_n n-boyutlu öklid uzayı (yani, alışılmış n- boyutlu uzay) olsun. Ω, E_n' nin bir p-boyutlu alt uzayı olsun. Ω , E_n' nin geriye kalanı yani E_n' nin Ω' ya ortogonal altuzayı olsun. $\mathbf{P}_\Omega$ genel bir n-boyutlu **Y** vektörünü tamamen Ω uzayına izdüşüren bir $n \times n$ izdüşüm matrisi olsun. (yalnız Ω uzayını hesaba katan ifadelerle basit olarak **P** yazacağız.) Bu taktirde bir takım ifadeler aşağıdaki gibi gösterilebilir.

1. $\hat{\mathbf{Y}}$ tamamen Ω da olmak üzere ($\hat{\mathbf{Y}} \in \Omega$ yazarız), her **Y** vektörü $\hat{\mathbf{Y}} + \mathbf{e}$ ve $\mathbf{e} \in \Omega^{-}$ biçiminde bir tek şekilde ifade edilebilir.

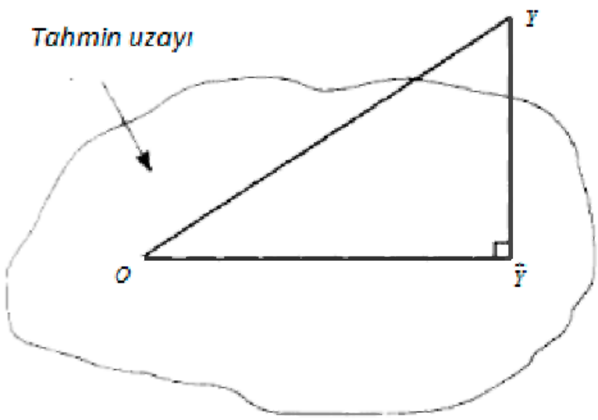
2. Eğer $\hat{\mathbf{Y}} = \mathbf{PY}$ ise, bu taktirde P tektir.
3. P tek olmakla beraber **T** $n \times p$ matrisinin p tane sütunu ortonormal bir taban (tek değildir) oluşturmak üzere, **P**, $\mathbf{P} = \mathbf{TT'}$ olarak yazılabilir.| Not: Ω' nın bir tabanı, $\boldsymbol{\eta}$' nın uzayını geren, yani Ω' nın her vektörünü, taban vektörlerin bir lineer kombinasyonu olarak ifade edebilmeye izin veren, vektörlerin bir kümesidir. Bir ortogonal taban, içindeki vektörlerinin tümünün birinin diğerine ortogonal olduğu bir tabandır. Bir ortonormal taban, taban vektörlerinin uzunluğu 1 olan bir ortogonal tabandır, yani v taban vektörleri için $v'v = 1$ dir. (uzunluğun karesi de 1 dir.)|
4. P simetrik $(\mathbf{P'} = \mathbf{P})$ ve idempotent $(\mathbf{P}^2 = \mathbf{P})$ dir.
5. **P**' nin vektörleri Ω uzayını gerer, yani $R(\mathbf{P}) = \Omega$ dır.
6. $\mathbf{I} - \mathbf{P}, \Omega^{\perp}$ için izdüşüm matrisidir. Burada $\Omega^{\perp}, E_n$' nin Ω da olmayan ortogonal kısmıdır.
7. Herhangi bir simetrik idempotent **P** $n \times n$ matrisi, **P**' nin sütunları tarafından gerilen uzaya yani $R(\mathbf{P})$' ye bir ortogonal izdüşüm matrisini gösterir.

1-7 ifadeleri, her ne kadar onları regresyon kavramına uygun bir notasyona göre vermiş olsak bile, genel olarak herhangi bir Ω için doğrudur. Şimdi 8 ve 9 ifadeleri regresyonla ilişkilendirilecektir. En küçük karelerle $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ modelini uydurmayı düşünelim. Şimdi $\Omega = R(\mathbf{X})$, **X**' in **P** tane sütunuyla tanımlanmış olacak ve bu nedenle regresyon problemimiz için tahmin uzayını oluşturacaktır.

8. Ω; **X** $n \times p$ matrisinin sütunları tarafından gerilen bir uzay olsun . $(\mathbf{X'X})^{-}$, $(\mathbf{X'X})$ matrisinin herhangi bir g-tersi olsun. Bu taktirde $\mathbf{P} = \mathbf{X}(\mathbf{X'X})^{-}\mathbf{X'}$, Ω için bir tek izdüşüm matrisidir. Ne **X**' in ne de $(\mathbf{X'X})^{-}$' nin tek olmadığını fakat

$\mathbf{P} = \mathbf{X}(\mathbf{X'X})^{-}\mathbf{X'}$ matrisinin tek ve bu nedenle de $\hat{\mathbf{Y}} = \mathbf{PY}$ ' nin tek olduğuna önemle dikkat edelim.

9. Eğer Cr (**X**) =*p* , yani **X**'in sütunları lineer bağımsız iseler ve $(\mathbf{X'X})^{-1}$ mevcutsa, bu taktirde $\mathbf{P} = \mathbf{X}(\mathbf{X'X})^{-1}\mathbf{X'}$ ve $R(\mathbf{X}) = R(\mathbf{P}) = \Omega$ dir.Bu tipik olarak sağlanacak olan durumdur. Bu ifadelerin ispatları için bu ifadelerin uygulanabilir sonuçları aşağıda verilmiştir (Seber 1977).Verilen bir Ω tahmin uzayı (ve böylece ister istemez, $\Omega^{\perp}$ hata uzayı için) herhangi bir **Y** vektörü bir tek $\mathbf{P} = \mathbf{X}(\mathbf{X'X})^{-1}\mathbf{X'}$ izdüşüm matrisi vasıtasıyla, Ω içine tek bir şekilde geri çekilebilir. **X** ve $(\mathbf{X'X})^{-}$ tek olmadıklarında bile, **P**' nin tek ve bu nedenle verilen bir Ω için $\hat{\mathbf{Y}} = \mathbf{PY}$ ' nin tek olduğu gerçekleri, geometrik olarak düşünüldüğünde tamamen açıktır. Verilen bir Ω ve **Y** için **Y'** nin Ω üzerine $\ddot{\mathbf{Y}}'$ yı Ω da tek olarak tanımlayan bir tek izdüşümü vardır. Ω'nın nasıl tanımlandığı[yani, **X**'in nasıl seçildiği ya da $\mathbf{X'X}$ singüler olduğunda $(\mathbf{X'X})^{-}$ seçiminin nasıl yapıldığı] veya $\hat{\mathbf{Y}}$'nın, b' lerin bir $\mathbf{Xb} = \mathbf{X}(\mathbf{X'X})^{-}\mathbf{X'Y}$ fonksiyonu olarak nasıl ifade edildiği mesele değildir. Şekil 5.13' ün $\mathbf{0}, \mathbf{Y}$ ve $\hat{\mathbf{Y}}$'yı birleştiren üçgeni, verilen Ω ve Y için daima sabit kalır.



Şekil 5.13. Verilen Ω ve Y için üçgen tektir.

5.7. EN KÜÇÜK KARELERİN GEOMETRİSİ

İlk bölümde çoklu lineer regresyon modelini, $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ matris denklemiyle sunduk. En küçük kareler tahmini yöntemini modelden sapmalara dayanarak tanımladık. Şimdi aşağıdaki örneği verelim.

Örnek 5.7.1 : Bir istatistik dersinde öğrencilerin ev ödevi yapması dönem ortasında yapılan sınav için, onları sınava hazırlamaya yardım etmediğini iddia edilmektedir. Sınıftaki 18 öğrenci için, sınav notu y ve ev ödevi notu x, (dönem ortasına kadarki ortalama notlar) aşağıdaki gibidir.

Çizelge 5.3: 18 öğrenci için sınav notu ve ev ödevi notu

y	x	y	x	y	x
95	96	72	89	35	0
80	77	66	47	50	30
0	0	98	90	72	59
0	0	90	93	55	77
79	78	0	18	75	74
77	64	95	86	66	67

Bu örnekten,

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} x_i y_i - n\bar{x}\bar{y}}{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2}$$

$$= \frac{81{,}195 - 18(58.056)(61.389)}{80{,}199 - 18(58.056)^2} = .8726,$$

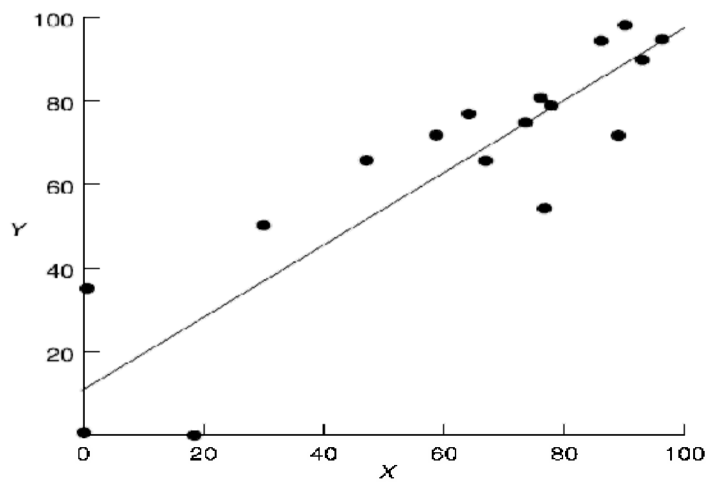
$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} = 61.389 - .8726(58.056) = 10.73.$$

elde ederiz.Bu nedenle tahmin (öngörü) denklemi

$$\hat{y} = 10.73 + .8726x \tag{5.7.1}$$

ile verilir.Bu denklemle verilen regresyon doğrusu ve 18 nokta Şekil 5.14 de grafiklenmiştir. Çizimde x değişirken $\hat{\beta}_1$ eğiminin $\hat{y}$ 'nın değişim oranı olduğu ve $\hat{\beta}_0$ sabitinin, $x = 0$ da $\hat{y}$ ' nın değeri olduğu hemen görülmektedir.

Şekil 5.14 de beliren lineer eğilim, ev ödevi ve test sonuçları arasındaki sebep ve sonucu belirlemez. Her $i = 1, 2, 3, \ldots, 18$ için, $\mathrm{var}(\varepsilon_i) = \sigma^2$ (sabit varyans) varsayımının akla uygun olacağı gözükmektedir.



Şekil 5.14. Ev ödevi ve test puanları (notları) için regresyon doğrusu ve veri

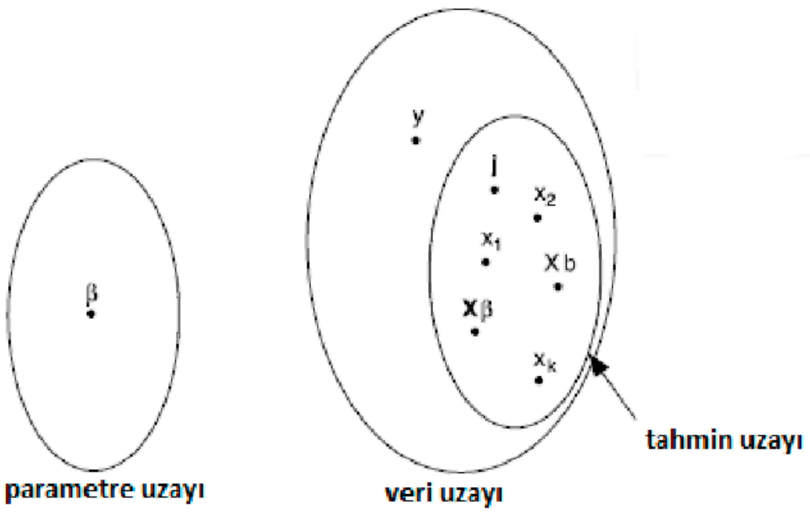
İlk olarak en küçük karelere neyin geometrik yaklaşım olmadığını açıklamak önemlidir. İki boyutta, n tane $(x_1 y_1)$, $(x_2, y_2),...,(x_n, y_n)$ noktasının iki boyutlu bir saçılım grafiğini (Şekil 5.14) oluşturarak en küçük kareler prensibini açıkladık. Sonra da veriye en uygun doğru olarak en küçük kareler doğrusunu görselleştirdik. Bu yaklaşım, $(k+1)$-boyutlu uzayda n- tane $(x_{11}, x_{12},..., x_{1k}, y_1)$, $(x_{21}, x_{22},..., x_{2k}, y_2),...,(x_{n1}, x_{n2},..., x_{nk}, y_n)$ noktasına en uygun hiperdüzlemin uydurulmasına dayalı çoklu lineer regresyonda, en küçük kareler tahminini sunmak için genelleştirilebilir.

Bu yaklaşım çoklu lineer regresyonu görselleştirmede bir dereceye kadar yararlı olmakla beraber çoklu lineer regresyonda en küçük kareler tahminine geometrik yaklaşım bu yüksek boyutlu genelleştirmeyi içermez.

Aşağıda ele alınacak geometrik yaklaşım onun matematiksel güzelliğinden dolayı sevimlidir. Örneğin; tahmin edici, matris analizini kullanmaksızın ortaya konur. Aynı zamanda, geometrik yaklaşım istatistiksel çıkarsamaya (sonuç elde etmeye) daha derin bir kavrayış getirir. Çekirdek yumuşatmasını içeren birkaç ileri istatistiksel yöntem (Eubank ve Eubank (1999)), Fourier analizi, (Bloomfield (2000)) ve dalgacık (küçük dalga) analizi (Ogden (1997)), bu geometrik yaklaşımın bir genelleştirmeleri olarak yorumlanabilir. Lineer modellere geometrik yaklaşım, ilk olarak Fisher Mahalanobis (1964) tarafından önerildi. Christensen (1996) ve Jammalamadaka ve Sengupta (2003) lineer istatistiksel modeli nerdeyse tamamen geometrik açıdan ele aldılar.

5.8. PARAMETRE UZAYI, VERİ UZAYI VE TAHMİN UZAYI

En küçük karelere geometrik yaklaşım, biri $(k+1)$-boyutlu ve diğeri n-boyutlu, iki yüksek boyutlu uzay ile başlar. Bilinmeyen $\boldsymbol{\beta}$ parametre vektörüne, $k+1$ tane $\beta_0, \beta_1, ..., \beta_k$ regresyon katsayılarına, karşılık gelen eksenler ile birlikte, $(k+1)$-boyutlu uzayda bir tek nokta olarak bakılabilir. Bu nedenle bu uzaya "parametre uzayı" (Şekil 5.15) diyeceğiz. Aynı şekilde, $\mathbf{y}$ veri vektörüne, n tane gözleme karşılık gelen eksenler ile birlikte n-boyutlu uzayda bir tek nokta olarak bakılabilir. Bu uzaya "veri uzayı" diyeceğiz.



Şekil 5.15. Temsil edici elemanlarla birlikte, parametre uzayı, veri uzayı ve tahmin uzayı

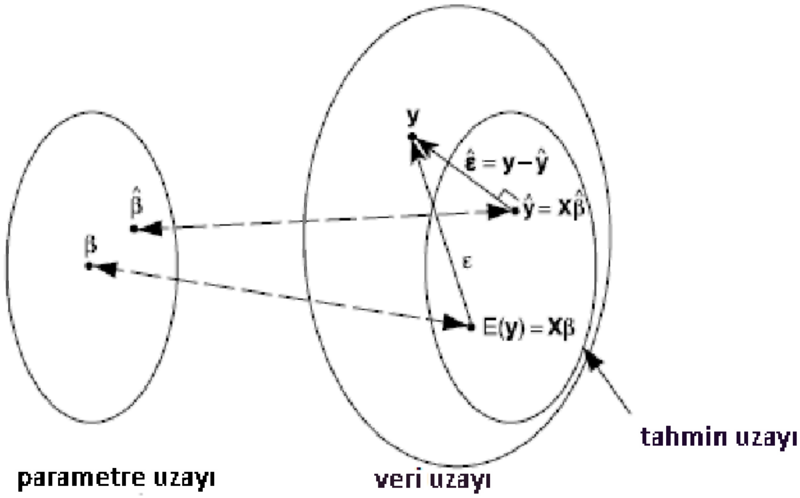
(5.1.1) çoklu regresyon modelinin $\mathbf{X}$ matrisi, $\mathbf{X} = \{\mathbf{1}, \mathbf{x}_1, ..., \mathbf{x}_k\}$ biçiminde onun $k+1$ tane sütununa dayanarak parçalanabilir. $\mathbf{1}$ dahil, $\mathbf{X}$'in, sütunlarının tümü n-boyutlu (bileşenli) vektörlerdir ve bu nedenle veri uzayındaki noktalardır. $\mathbf{X}$'in rankını $k+1$ kabul ettiğimizden, bu vektörlerin lineer bağımsız olduklarına dikkat edelim. $\mathbf{X}$'in sütunlarının tüm mümkün lineer kombinasyonları veri uzayının bir alt kümesini oluşturur. Bu kümenin elemanları, $\mathbf{b}$ herhangi bir $(k+1)$ vektör, yani, parametre uzayındaki herhangi bir vektör olmak üzere,

$$\mathbf{Xb} = b_0\mathbf{1} + b_1\mathbf{x}_1 + ... + b_k\mathbf{x}_k \tag{5.8.1}$$

olarak yazılabilir. Gerçekten bu alt küme toplama ve çarpma altında kapalı olduğundan, bir alt uzay statüsüne sahiptir (Harville 1997). Bu alt kümeye $\mathbf{X}$'in sütunları tarafından üretilen veya gerilen altuzay denecektir ve bu alt uzaya (öngörü) “tahmin uzayı” diyeceğiz. $\mathbf{X}$ ’in sütunları tahmin uzayı için bir taban kümesi oluştururlar.

5.9. ÇOKLU LİNEER REGRESYON MODELİNİN GEOMETRİK YORUMU

(5.1.1) Çoklu lineer regresyon modeli, $\mathbf{Y}$ ’nin, tahmin uzayındaki bir vektör ile yani $\mathbf{E(Y) = X\beta}$ ile rasgele hataların bir vektörünün, yani $\boldsymbol{\varepsilon}$ ’nun, toplamına, eşit olduğunu, ifade eder. (Şekil 5.16)



Şekil 5.16. Çoklu Lineer Regresyon Modeline İlişkin Vektörlerin Geometrik İlişkileri

Problem, ne $\boldsymbol{\beta}$ ’nın ne de $\boldsymbol{\varepsilon}$ ’nun bilinmemesidir. Bununla beraber, tahmin uzayında olmayan, $\mathbf{Y}$ veri vektörü bilinir. Ayrıca $\mathbf{E(Y)}$ ’nin tahmin uzayında olduğu da bilinir.

Çoklu lineer regresyon, geometrik olarak, tahmin uzayındaki $\mathbf{E(Y)}$ ’nin akla uygun bir tahminini bulma ve sonra da parametre uzayında bu tahmin ile ilgili vektörü belirleme işlemi olarak anlaşılabilir. (Şekil 5.16) $\mathbf{E(Y)}$ ’nin tahmini $\hat{\mathbf{Y}}$ olarak gösterilir ve parametre uzayındaki ilgili vektör $\hat{\boldsymbol{\beta}}$ olarak gösterilir. Akla uygun (mantıklı) bir geometrik düşünce tahmin uzayında $\mathbf{Y}$ ’ye en yakın olan noktayı kullanarak $\mathbf{E(Y)}$ ’yi tahmin etmektir. Bu düşünce, tahmin uzayında $\mathbf{Y}$ ’ ye en yakın $\hat{\mathbf{Y}}$ ’nın, $\hat{\boldsymbol{\varepsilon}} = \mathbf{Y} - \hat{\mathbf{Y}}$ fark vektörünün tahmin uzayına ortogonal (dik) olması gerektiğine dikkat etmek suretiyle bulunabildiğini ortaya koyar (Harville 1997). Ayrıca, tahmin uzayı, $\mathbf{X}$ ’in sütunları tarafından gerildiğinden, $\hat{\mathbf{Y}}$

noktası; $\hat{\boldsymbol{\varepsilon}}$, $\mathbf{X}$' in sütunlarına ortogonal olacak şekilde olmalıdır. Böylece $\mathbf{a'b}=0$' ın bir genişlemesini kullanarak,

$$\mathbf{X'\hat{\boldsymbol{\varepsilon}}=0}$$

veya

$$\mathbf{X'(Y-\hat{Y})=X'(Y-X\hat{\boldsymbol{\beta}})=X'Y-X'X\hat{\boldsymbol{\beta}}=0} \tag{5.9.1}$$

olacak şekilde $\hat{\mathbf{Y}}$' yı araştırırız. Bu ise

$$\mathbf{X'X\hat{\boldsymbol{\beta}}=X'Y} \tag{5.9.2}$$

normal denklemlerini ortaya koyar. Bu nedenle, sadece geometrik düşünceleri kullanarak, normal denklemleri ve sonuç olarak $\hat{\boldsymbol{\beta}}$ en küçük kareler tahmin edicisini elde ederiz. Sonra da, $\hat{\mathbf{Y}}$ yı, $\mathbf{X\hat{\boldsymbol{\beta}}=X(X'X)^{-1}X'Y=HY}$ olarak hesaplarız. $\hat{\boldsymbol{\varepsilon}}=\mathbf{Y-X\hat{\boldsymbol{\beta}}=(I-H)Y}$, $\boldsymbol{\varepsilon}$' nun bir tahmini olarak alınabilir. $\hat{\boldsymbol{\varepsilon}}$, $(n-k-1)$-boyutlu uzayda bir vektör olduğundan, σ^2' yi, $\hat{\boldsymbol{\varepsilon}}$' nun karesi alınmış uzunluğunu $(n-k-1)$ değerine bölerek, tahmin etmek akla uygun görünmektedir. Başka bir deyişle, σ^2' nin akla uygun bir tahmin edicisi, $E(s^2)=\sigma^2$ dekine eşit olan,

$$s^2=\mathbf{Y'(I-H)Y}/(n-k-1) \tag{5.9.3}$$

dır.

5.10. MERKEZİLEŞTİRİLMİŞ BİÇİMDEKİ MODEL

Her bir y_i için $y_i=\beta_0+\beta_1x_{i1}+\ldots+\beta_kx_{ik}+\varepsilon_i \quad i=1,2,\ldots,n$ modeli merkezileştirilen $\mathbf{X}$ değişkenlerine göre,

$$y_i=\beta_0+\beta_1x_{i1}+\ldots+\beta_kx_{ik}+\varepsilon_i \quad i=1,2,\ldots,n$$

$$=\alpha+\beta_1(x_{i1}-\bar{x}_1)+\beta_2(x_{i2}-\bar{x}_2)+\ldots+\beta_k(x_{ik}-\bar{x}_k)+\varepsilon_i \tag{5.10.1}$$

olarak yazılabilir. Burada

$$\alpha = \beta_0 + \beta_1 \bar{x}_1 + \dots + \beta_k \bar{x}_k \qquad (5.10.2)$$

ve $\bar{x}_j = \sum_{i=1}^{n} x_{ij} / n$, j=1,2,...,k dır. Modelin merkezileştirilmiş biçimi, etkili gözlemler için bir araştırmada ve diğer sezişleri temin etmede, belirli hipotezlerin ifade edilmesinde yararlıdır. $y_1, y_2, \dots, y_n$ için merkezileştirilmiş model matris formunda,

$$\mathbf{y} = (\mathbf{1}, \mathbf{X}_c)\begin{pmatrix} \boldsymbol{\alpha} \\ \boldsymbol{\beta}_1 \end{pmatrix} + \boldsymbol{\varepsilon} \qquad (5.10.3)$$

olur. Burada, $\boldsymbol{\beta}_1 = (\beta_1, \dots, \beta_k)'$,

$$\mathbf{X}_c = (\mathbf{I} - \frac{1}{n}\mathbf{J})\mathbf{X}_1 = \begin{pmatrix} x_{11} - \bar{x}_1 & \cdots & x_{1k} - \bar{x}_k \\ x_{21} - \bar{x}_1 & \cdots & x_{2k} - \bar{x}_k \\ \vdots & \vdots & \vdots \\ x_{n1} - \bar{x}_1 & \cdots & x_{nk} - \bar{x}_k \end{pmatrix}, \qquad (5.10.4)$$

ve $\mathbf{X}_1$ dır. $\mathbf{I} - (1/n)\mathbf{J}$ matrisine bazen "merkezileştirme matrisi" de denir.

$\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{X}'\mathbf{Y}$ de olduğu gibi, (5.10.3) deki model için normal denklemler

$$(\mathbf{1}, \mathbf{X}_c)'(\mathbf{1}, \mathbf{X}_c)\begin{pmatrix} \hat{\boldsymbol{\alpha}} \\ \hat{\boldsymbol{\beta}}_1 \end{pmatrix} = (\mathbf{1}, \mathbf{X}_c)'\mathbf{Y} \qquad (5.10.5)$$

dir. (5.10.5) in sol yanındaki $(\mathbf{1}, \mathbf{X}_c)'\mathbf{1}, \mathbf{X}_c)$ çarpımı

$$\begin{aligned} (\mathbf{1}, \mathbf{X}_c)'(\mathbf{1}, \mathbf{X}_c) &= \begin{pmatrix} \mathbf{1}' \\ \mathbf{X}_c' \end{pmatrix}(\mathbf{1}, \mathbf{X}_c) = \begin{pmatrix} \mathbf{1}'\mathbf{1}_1 & \mathbf{1}'\mathbf{X}_c \\ \mathbf{X}_c'\mathbf{1} & \mathbf{X}_c'\mathbf{X}_c \end{pmatrix} \\ &= \begin{pmatrix} n & \mathbf{0}' \\ \mathbf{0} & \mathbf{X}_c'\mathbf{X}_c \end{pmatrix} \end{aligned} \qquad (5.10.6)$$

olur. Burada, $\mathbf{X}_c$ ' nın sütunları toplamı sıfır olduğundan $\mathbf{1}'\mathbf{X}_c = \mathbf{0}'$ dır. (5.10.5) in sağ yanı $(\mathbf{1}, \mathbf{X}_c)'\mathbf{Y} = \begin{pmatrix} \mathbf{1}' \\ \mathbf{X}_c' \end{pmatrix}\mathbf{Y} = \begin{pmatrix} \mathbf{n}\bar{y} \\ \mathbf{X}_c'\mathbf{Y} \end{pmatrix}$ olarak yazılabilir. Bu durumda en küçük kareler tahmin edicileri

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

$$=\begin{pmatrix}1/n & \mathbf{0}'\\ \mathbf{0} & (\mathbf{X}_c'\mathbf{X}_c)^{-1}\end{pmatrix}\begin{pmatrix}n\bar{y}\\ \mathbf{X}_c'\mathbf{Y}\end{pmatrix}=\begin{pmatrix}\bar{y}\\ (\mathbf{X}_c'\mathbf{X}_c)^{-1}\mathbf{X}_c'\mathbf{Y}\end{pmatrix}$$

veya

$$\hat{\alpha} = \bar{y} \tag{5.10.7}$$

$$\hat{\boldsymbol\beta}_1 = (\mathbf{X}_c'\mathbf{X}_c)^{-1}\mathbf{X}'_c\mathbf{Y} \tag{5.10.8}$$

ile verilir. Bu tahmin ediciler, (5.10.2) deki α nın bir tahmin edicisinden elde edilen

$$\hat{\beta}_0 = \hat{\alpha} - \hat{\beta}_1\bar{x}_1 - \hat{\beta}_2\bar{x}_2 - \ldots - \hat{\beta}_k\bar{x}_k = \bar{y} - \hat{\beta}_1\bar{x} = \bar{y} - \hat{\boldsymbol\beta}_1'\bar{\mathbf{x}} \tag{5.10.9}$$

ayarlaması ile birlikte, $\hat{\boldsymbol\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$ alışılmış en küçük kareler tahmin edicisiyle aynıdır.

$\hat{\mathbf{Y}}$ ' yı

$$\hat{\mathbf{Y}} = \hat{\alpha} + \hat{\beta}_1(x_1 - \bar{x}_1) + \ldots + \hat{\beta}_k(x_k - \bar{x}_k)$$

merkezileşmiş formunda ifade ettiğimizde, uygun regresyon düzleminin $(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_k, \bar{y})$ noktasından geçtiği açıktır. HKT : Hata Kareler Toplamını göstermek üzere, HKT için, merkezileştirilmiş $\hat{\mathbf{Y}}$ lar ile sahip olunan merkezileştirilmiş modele uyarlayarak, $HKT = \mathbf{Y}'\mathbf{Y} - \hat{\boldsymbol\beta}'\mathbf{X}'\mathbf{Y}$ ye eşitliğini ortaya koyacak olan

$$HKT = \sum_{i=1}^{n}(y_i - \bar{y})^2 - \hat{\boldsymbol\beta}_1'\mathbf{X}_c'\mathbf{Y} \tag{5.10.10}$$

yi elde ederiz.

$\hat{\boldsymbol\beta}_1$ ve $\hat{\boldsymbol\beta}_0$ ' yı, örneklem varyansları ve örneklem kovaryanslarına bağlı olarak ifade etmek için (5.10.7)-(5.10.9) bağıntılarını kullanabiliriz. Bu rasgele- x durumu için elde edilecek olan tahmin edicilerle bu tahmin edicileri karşılaştırmada yararlı olacaktır. İlk olarak x

değişkenleri için bir örneklem kovaryans matrisini ve y ve x ler arasındaki örneklem kovaryanslarının bir vektörünü

$$\mathbf{S}_{xx} = \begin{pmatrix} s_1^2 & s_{12} \ldots & s_{1k} \\ \vdots & \vdots & \vdots \\ s_{k1} & s_{k2} & s_k^2 \end{pmatrix}, \quad \mathbf{s}_{yx} = \begin{pmatrix} s_{y1} \\ s_{y2} \\ \vdots \\ s_{yk} \end{pmatrix} \tag{5.10.11}$$

olarak tanımlarız.

$\mathbf{Y}$ $n \times 1$ örneklem vektörü için örneklem varyansı

$s^2 = \dfrac{\sum_{i=1}^{n}(y_i - \bar{y})^2}{n-1}$ ve $\mathbf{X}$ ve $\mathbf{Y}$ $n \times 1$ örneklem vektörleri için örneklem kovaryansı

$s_{xy} = \dfrac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{n-1}$ olarak tanımlanır. Burada s_i^2, s_{ij} ve s_{yi} ler, yukarıdaki s^2 ve s_{xy} ye

benzerdirler; örneğin, $\bar{x}_2 = \sum_{i=1}^{n} x_{i2} / n$ olmak üzere,

$$s_2^2 = \frac{\sum_{i=1}^{n}(x_{i2} - \bar{x}_2)}{n-1} \tag{5.10.12}$$

$$s_{12} = \frac{\sum_{i=1}^{n}(x_{i1} - \bar{x}_1)(x_{i2} - \bar{x}_2)}{n-1} \tag{5.10.13}$$

$$s_{y2} = \frac{\sum_{i=1}^{n}(x_{i2} - \bar{x}_2)(y_i - \bar{y})}{n-1} \tag{5.10.14}$$

dır. Bununla beraber, x ler sabit olduklarından bu örneklem varyansları ve kovaryansları kitle varyanslarını ve kovaryanslarını tahmin etmezler. Eğer , x ler rasgele değişkenler olsa idi, s_i^2, s_{ij}, s_{yi} değerleri kitle parametrelerini tahmin edeceklerdi. $\hat{\boldsymbol{\beta}}_1$ ve $\hat{\boldsymbol{\beta}}_0$ yı, $\mathbf{S}_{xx}$ ve $\mathbf{s}_{yx}$

değerlerine dayanarak tahmin etmek için, ilk olarak, $\mathbf{S}_{xx}$ ve $\mathbf{s}_{yx}$ değerlerini merkezileşmiş $\mathbf{X}_c$ matrisine dayanarak aşağıdaki gibi yazarız:

$$\mathbf{S}_{xx} = \frac{\mathbf{X}_c'\mathbf{X}_c}{n-1} \tag{5.10.15}$$

$$\mathbf{s}_{yx} = \frac{\mathbf{X}_c'\mathbf{Y}}{n-1} \tag{5.10.16}$$

(5.10.16) da ki $\mathbf{X}_c'\mathbf{Y}$ ' nin, (5.10.14) deki gibi $\sum_{i=1}^{n}(x_{ij} - \bar{x}_j)(y_i - \bar{y})$ den ziyade, $\sum_{i=1}^{n}(x_{ij} - \bar{x}_j)y_i$ biçimindeki terimleri ihtiva ettiğine dikkat edelim. $\sum_{i=1}(x_{ij} - \bar{x}_j)(y_i - \bar{y})$ $\sum_{i=1}(x_{ij} - \bar{x}_j)y_i$ olduğunda hemen gösterilebilir.

(5.10.8), (5.10.15) ve (5.10.16) dan

$$\hat{\boldsymbol{\beta}}_1 = (n-1)(\mathbf{X}_c'\mathbf{X}_c)^{-1}\frac{\mathbf{X}_c'\mathbf{Y}}{n-1} = \left(\frac{\mathbf{X}_c'\mathbf{X}_c}{n-1}\right)^{-1}\frac{\mathbf{X}_c'\mathbf{Y}}{n-1} = \mathbf{S}_{xx}^{-1}\mathbf{s}_{yx} \tag{5.10.17}$$

ve (5.10.9) ve (5.10.17) dan,

$$\hat{\beta}_0 = \hat{\alpha} - \hat{\boldsymbol{\beta}}_1'\bar{x} = \bar{y} - \mathbf{s}_{yx}'\mathbf{S}_{xx}^{-1}\bar{x} \tag{5.10.18}$$

elde ederiz.

Örnek 5.10.1: Aşağıdaki verileri göz önüne alalım. (5.10.17) ve (5.10.18) bağıntılarını kullanarak,

gözlem numaraları	Y	x_1	x_2
1	2	0	2
2	3	2	6
3	2	2	7
4	7	2	5
5	6	4	9
6	8	4	8
7	10	4	7
8	7	6	10
9	8	6	11
10	12	6	9
11	11	8	15
12	14	8	13

tam modelde yani merkezileştirilmiş modelde

$$\hat{\boldsymbol{\beta}}_1 = \mathbf{S}_{xx}^{-1}\mathbf{s}_{yx} = \begin{pmatrix} 6.4242 & 8.5455 \\ 8.5455 & 12.4545 \end{pmatrix}^{-1} \begin{pmatrix} 8.3636 \\ 9.7273 \end{pmatrix}$$

$$= \begin{pmatrix} 3.0118 \\ -1.2855 \end{pmatrix},$$

$$\hat{\beta}_0 = \bar{y} - \mathbf{s}_{yx}'\mathbf{S}_{xx}^{-1}\bar{x}$$

$$=7.5000-(3.0118,-1.2855)\begin{pmatrix} 4.3333 \\ 8.5000 \end{pmatrix}$$

$$=7.5000-2.1246=5.3754$$

elde ederiz.Bu değerler yukarıdaki değerlerin aynısıdır.

5.11. NORMAL MODEL

Varsayımlar

Buraya kadar $y_1, y_2, \ldots, y_n$ rasgele değişkenleri hakkında normallik varsayımı yapmadık. Şimdi

$\mathbf{Y} \sim N_n(\mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I})$ dır veya $\boldsymbol{\varepsilon} \sim N_n(\mathbf{0}, \sigma^2\mathbf{I})$ varsayımını ekleyeceğiz. Normallik varsayımı altında, $\sigma_{ij} = 0$, ilişkisizliğin yanı sıra, $\mathbf{y}$ (veya $\boldsymbol{\varepsilon}$,) değişkenlerinin bağımsız olduklarını da ifade eder.

$\boldsymbol{\beta}$ ve σ^2 İçin En Çok Olabilirlik Tahmin Edicileri

Normallik varsayımı ile, en çok olabilirlik tahmin edicilerini elde edebiliriz. Olabilirlik fonksiyonu, $\mathbf{L}(\boldsymbol{\beta},\sigma^2)$ile göstereceğimiz, $\mathbf{Y}$ lerin ortak yoğunluğudur. Örneklemde verilen $\mathbf{Y}$ ve x için, $\mathbf{L}(\boldsymbol{\beta},\sigma^2)$yi maksimum yapan (en büyük kılan) bilinmeyen $\boldsymbol{\beta}$ ve σ^2nin değerlerini araştıracağız.

Normal dağılım fonksiyonunun durumunda, diferansiyelle $\hat{\boldsymbol{\beta}}$ ve $\hat{\sigma}^2$maksimum (en çok) olabilirlik tahmin edicilerini bulmak mümkündür. Normal dağılım bir çarpım ve bir üsteli ihtiva ettiğinden, $\mathbf{L}(\boldsymbol{\beta},\sigma^2)$ile çalışmak daha basittir. Çünkü $\mathbf{L}_n\mathbf{L}(\boldsymbol{\beta},\sigma^2)$'yi maksimum yapan aynı $\boldsymbol{\beta}$ ve σ^2değerleri için $\mathbf{L}(\boldsymbol{\beta},\sigma^2)$de maksimum olur.

Teorem 5.11.1: Eğer $\mathbf{Y} \sim N_n(\mathbf{X}\boldsymbol{\beta},\sigma^2\mathbf{I})$ise, bu taktirde $\hat{\boldsymbol{\beta}}$ ve $\hat{\sigma}^2$, $\boldsymbol{\beta}$ ve σ^2için müştereken (birlikte) yeterlidir.

İspat: $\mathbf{f}(\mathbf{Y},\boldsymbol{\beta},\sigma^2)$ yoğunluğu verildi. Üstelde

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

$$=\left[(\mathbf{Y}-\mathbf{X}\hat{\boldsymbol{\beta}})+\mathbf{X}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})\right]'\left[(\mathbf{Y}-\mathbf{X}\hat{\boldsymbol{\beta}})+\mathbf{X}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})\right]$$

elde etmek için, $\mathbf{X}\hat{\boldsymbol{\beta}}$ ' yı bir ekler, birde çıkarırız. Bunu $\mathbf{Y}-\mathbf{X}\hat{\boldsymbol{\beta}}$ ve $\mathbf{X}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})$ya göre açarak, $\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{X}'\mathbf{Y}$ normal denklemlerinden dolayı ikisi sıfır olan, dört terim elde ederiz. Sonuç;

$$(\mathbf{Y}-\mathbf{X}\boldsymbol{\beta})'(\mathbf{Y}-\mathbf{X}\boldsymbol{\beta}) = (\mathbf{Y}-\mathbf{X}\hat{\boldsymbol{\beta}})'(\mathbf{Y}-\mathbf{X}\hat{\boldsymbol{\beta}})+(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})'\mathbf{X}'\mathbf{X}(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}}) \qquad (5.11.1)$$

$$= n\hat{\sigma}^2+(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})'\mathbf{X}'\mathbf{X}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})$$

dır. Şimdi, $h(y)=1$olmak üzere, $\mathbf{f}(\mathbf{Y},\boldsymbol{\beta},\sigma^2)=\mathbf{g}(\hat{\boldsymbol{\beta}},\hat{\sigma}^2,\boldsymbol{\beta},\sigma^2)\mathbf{h}(\mathbf{Y})$biçiminde olan,

$$\mathbf{f}(\mathbf{Y},\boldsymbol{\beta},\sigma^2)=\frac{1}{(2\pi\sigma^2)^{n/2}}\mathbf{e}^{-}\left[\mathbf{n}\hat{\sigma}^2+(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})'\mathbf{X}'\mathbf{X}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})\right]/2\sigma^2$$

olarak yazabiliriz. Bu nedenle, Neyman Çarpansallıma teoremine göre, $\hat{\boldsymbol{\beta}}$ ve $\hat{\sigma}^2$, $\boldsymbol{\beta}$ ve σ^2 için müştereken yeterlidirler. $\hat{\boldsymbol{\beta}}$ ve $\hat{\sigma}^2$'nin $\boldsymbol{\beta}$ ve σ^2 için müştereken yeterli, bağımsız olarak yeterli olmadıklarına, yani, $\mathbf{f}(\mathbf{Y},\boldsymbol{\beta},\sigma^2)$ nin $g_1(\hat{\boldsymbol{\beta}},\boldsymbol{\beta})g_2(\hat{\sigma}^2,\sigma^2)\mathbf{h}(\mathbf{Y})$ biçiminde çarpanlarına ayrılamadığına dikkat edelim. $\mathbf{s}^2 = n\hat{\sigma}^2/(n-k-1)$ olduğundan, $\hat{\boldsymbol{\beta}}$ ve $\mathbf{s}^2$'nin $\boldsymbol{\beta}$ ve σ^2 için müştereken yeterli olduklarını göstermek için, Teorem 5.11.1' rin kolaylıkla değiştirebildiğine de dikkat edelim. $\hat{\boldsymbol{\beta}}$ ve $\hat{\sigma}^2$ yeterli olduklarından, $\boldsymbol{\beta}$ ve σ^2'yi tahmin etmek için onları örneklemden çıkardıkları bilgiden ileriye gidebilen başka tahmin ediciler yoktur. Bu nedenle $\hat{\boldsymbol{\beta}}$ ve s^2'nin minimum varyans yansız tahmin ediciler ($\hat{\boldsymbol{\beta}}$ daki her bir $\hat{\boldsymbol{\beta}}_j$ minimum varyansa sahiptir) oldukları şaşırtıcı değildir. Bu sonuç aşağıdaki teoremde verilir.

Teorem 5.11.2:Eğer $\mathbf{Y} \sim N_n(\mathbf{X}\boldsymbol{\beta},\sigma^2\mathbf{I})$ ise, bu taktirde $\hat{\boldsymbol{\beta}}$ ve s^2, diğer tüm yansız tahmin ediciler arasında, minimum varyansa sahiptirler.

İspat: bkz. Graybill (1976)veya Christensen (1996)

Teorem 5.11.3: (Gauss-Markov Teoremi) Eğer $\mathbf{E(Y)} = \mathbf{X}\boldsymbol{\beta}$ ve $kov(\mathbf{Y}) = \sigma^2 I$ ise $\hat{\boldsymbol{\beta}}_j$, $j = 0,1,2,...k$ en küçük kareler tahmin edicileri diğer tüm lineer yansız tahmin ediciler arasında minimum varyansa sahiptirler. Bu teoremin bir sonucu aşağıda verilmiştir.

Sonuç 1: Eğer $\mathbf{E(Y)} = \mathbf{X}\boldsymbol{\beta}$ ve $kov(\mathbf{Y}) = \sigma^2 I$ ise $\mathbf{a}'\boldsymbol{\beta}$ nın en iyi lineer yansız tahmin edicisi $\mathbf{a}'\hat{\boldsymbol{\beta}}$ dır. Burada $\hat{\boldsymbol{\beta}}$, $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$ en küçük kareler tahmin edicidir ve $\mathbf{a}' = (a_0, a_1, ..., a_k)$ dır.

Teorem 5.11.4: $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ lineer modeli için

$$s^2 = \frac{1}{n-k-1}(\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}})'(\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}})$$

$$= \frac{\mathbf{Y}'\mathbf{Y} - \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{Y}}{n-k-1} = \frac{HKT}{n-k-1}$$

dir.

Burada $HKT = (\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}})'(\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}) = \mathbf{Y}'\mathbf{Y} - \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{Y}$ dır. Eğer bu model için $\mathbf{E}(\boldsymbol{\varepsilon}) = \mathbf{0}$, $kov(\boldsymbol{\varepsilon}) = \sigma^2 I$ ve $\mathbf{E}(\boldsymbol{\varepsilon}_i^2) = 3\sigma^4$ ise bu taktirde; s^2, σ^2'nin kuadratik yansız tahmin edicisidir.

Teorem 5.11.3 de, $\boldsymbol{\beta}$'nın elemanlarının diğer tüm lineer yansız tahmin ediciler arasında minimum varyansa sahip olacağı açıklandı. Teorem 5.11.3 de eklenen normallik varsayımı ile, $\hat{\boldsymbol{\beta}}$'nın elemanları tüm lineer yansız tahmin ediciler arasında minimum varyansa sahiptirler. Teorem 5.11.4' e göre, s^2, tüm kuadratik yansız tahmin ediciler arasında minimum varyansa sahiptir.

Teorem 5.11.2 ye eklenen aşağıdaki sonuç Teorem 5.11.3 nin sonucunun benzeridir.

Sonuç 2 : Eğer $\mathbf{Y} \sim N_n(\mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I})$ ise, bu taktirde, $\hat{\boldsymbol{\beta}}$ en çok olabilirlik tahmin edicisi olmak üzere, $\boldsymbol{\alpha}'\boldsymbol{\beta}$'nın minimum varyans tahmin edicisi $\boldsymbol{\alpha}'\hat{\boldsymbol{\beta}}$ dır.

5.12. SABİT-x REGRESYONUNDA R^2

(5.12.10) da, $HKT = \sum_{i=1}^{n}(y_i - \bar{y})^2 - \hat{\boldsymbol{\beta}}_1'\mathbf{X}_c'\mathbf{Y}$ elde ettik. Bu nedenle $GKT = \sum_{i}(y_i - \bar{y})^2$ düzeltilmiş genel kareler toplamı

$$\sum_{i=1}^{n}(y_i - \bar{y})^2 = \hat{\boldsymbol{\beta}}_1'\mathbf{X}_c'\mathbf{Y} + HKT \qquad (5.12.1)$$
$$GKT \quad RKT \quad HKT$$

olarak parçalanabilir. Burada $RKT = \hat{\boldsymbol{\beta}}_1'\mathbf{X}_c'\mathbf{Y}$ regresyon kareler toplamıdır. (5.12.8) de, $\mathbf{X}_c'\mathbf{Y} = \mathbf{X}_c'\mathbf{X}_c\boldsymbol{\beta}_1$ elde ederiz ve bunun $\hat{\boldsymbol{\beta}}_1'$ ile çarpımı $\hat{\boldsymbol{\beta}}_1'\mathbf{X}_c'\mathbf{Y} = \hat{\boldsymbol{\beta}}_1'\mathbf{X}_c'\mathbf{X}_c\hat{\boldsymbol{\beta}}_1$ yı verir. Bu taktirde

$$\begin{aligned} RKT &= \hat{\boldsymbol{\beta}}_1'\mathbf{X}_c'\mathbf{Y} \\ RKT &= \hat{\boldsymbol{\beta}}_1'\mathbf{X}_c'\mathbf{X}_c\hat{\boldsymbol{\beta}}_1 = (\mathbf{X}_c\hat{\boldsymbol{\beta}}_1)'(\mathbf{X}_c\hat{\boldsymbol{\beta}}_1) \end{aligned} \qquad (5.12.2)$$

olarak yazılabilir. Bu biçimiyle, RKT nin $\hat{\boldsymbol{\beta}}_1 = (\hat{\beta}_1, \hat{\beta}_2, ..., \hat{\beta}_k)'$ye bağlandığı açıktır. Genel kareler toplamının regresyona yüklenebilir (bağlı) oranı, determinasyon (belirleme) katsayısı veya çoklu korelasyonun karesi olarak bilinen,

$$R^2 = \frac{\hat{\boldsymbol{\beta}}_1' \mathbf{X}_c' \mathbf{X}_c \hat{\boldsymbol{\beta}}_1}{\sum_{i=1}^{n} (y_i - \bar{y})^2} = \frac{RKT}{GKT} \tag{5.12.3}$$

dir. (5.12.3) deki oran model uyumunun bir ölçüsüdür ve $\mathbf{X}$ ' lerin $\mathbf{Y}$ yi ne kadar iyi tahmin ettiğinin bir belirtisidir.

(5.12.1) deki parçalanma R^2 için değişik bir

$$R^2 = \frac{\hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y} - n\bar{y}^2}{\mathbf{y}'\mathbf{y} - n\bar{y}^2} \tag{5.12.4}$$

ifadesine götüren,

$$\sum_{i=1}^{n} (y_i - \bar{y})^2 = \mathbf{y}'\mathbf{y} - n\bar{\mathbf{y}}^2 = (\hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y} - n\bar{\mathbf{y}}^2) + (\mathbf{y}'\mathbf{y} - \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y})$$
$$= RKT + HKT$$

özdeşliği olarak tekrar yazılabilir.

(5.12.3) veya (5.12.4) den elde edilen R pozitif kareköküne çoklu kolerasyon katsayısı denir. Eğer $\mathbf{x}$ değişkenleri rasgele olsa idi, R bir kitle çoklu kolerasyon tahmini olacaktır.

R^2 ve R 'nın bazı özelliklerini kaydedelim:

1. R^2 'nin değişim aralığı $0 \leq R^2 \leq 1$ dır. Eğer $\hat{\beta}_0$ hariç, $\hat{\beta}_j$ ların tümü sıfır olsa idi, $R^2 = 0$ olacaktı. (Sürekli veri için bu olay 0 olasılığa sahiptir.) Eğer tüm $\mathbf{y}$ değerleri uygun yüzeye düşseydi, yani $y_i = \hat{y}_i$, $i = 1,2,\ldots,n$ olsaydı, bu taktirde $R^2 = 1$ olacaktı.

2. $R = r_{y\hat{y}}$ dır; yani çoklu kolerasyon, gözlenen y_i ler ve uygun $\hat{y}_i$ lar arasındaki basit kolerasyonu eşittir.

3. Modele bir x değişkeninin eklenmesi R^2 'nin değerini artırır. (eksiltemez)

4. Eğer $\beta_1 = \beta_2 = \ldots = \beta_k = 0$ ise, bu taktirde

$$E(R^2) = \frac{k}{n-1} \tag{5.12.5}$$

dir. $\boldsymbol{\beta}_j$ ler 0 olduğunda, $\hat{\boldsymbol{\beta}}_j$ ların sıfır olamayacağına dikkat edelim.

5. R^2, x ler karşılıklı olarak ortogonal olmadıkça, yani $j \neq m$ için $\sum_{i=1}^{n}(x_{ij} - \bar{x}_j)(x_{im} - x_m') \quad 0$ olmadıkça, her biri bir tek şekilde bir X_j ye bağlanabilir olan, k tane bileşene parçalanamaz.

6. R^2, $\mathbf{X}$ ler üzerindeki tam rank lineer dönüşümlere göre ve $\mathbf{y}$ üzerindeki bir ölçek (ölçü) değişmesine göre değişmez (fakat $\mathbf{Y}$ ve $\mathbf{X}$ ' leri içeren ortak bir lineer dönüşüme göre değişir.) 3 ve 4 özelliklerinde, eğer k, n nin büyük bir kesri ise, R^2 'nin anlamlı olmayan büyük bir değerine sahip olmak mümkündür. Bu durumda $\mathbf{Y}$ 'yi tahmin etmeye katkıda bulunmayan $\mathbf{X}$ 'lerin özel bir örnekte öyle olacağı görülebilir ve tahmin edilen regresyon modeli kitle modelinin yararlı bir tahmin edicisi olamaz. Bu eğilimi düzeltmek için R_a^2 ile gösterilen bir ayarlanmış R^2 Ezekiel (1930) tarafından önerildi. R_a^2 ' yi elde etmek için, ilk olarak $\beta_1 = \beta_2 = \ldots = \beta_k = 0$ durumundaki yanlılığın düzeltilmesi için R^2 den (5.12.5) deki $\dfrac{k}{n-1}$ ri çıkarırız. Bununla beraber, bu düzeltme, $\boldsymbol{\beta}$ 'lar çok olduğunda R_a^2 'yı oldukça küçük yapacaktı, bu nedenle, $R^2 = 1$ olduğunda $R_a^2 = 1$ olacak şekilde başka bir değişiklik yapılır. Bundan dolayı R_a^2 ,

$$R_a^2 \quad \frac{(R^2 - \frac{k}{n-1})(n-1)}{n-k-1} \quad \frac{(n-1)R^2 - k}{n-k-1} \tag{5.12.6}$$

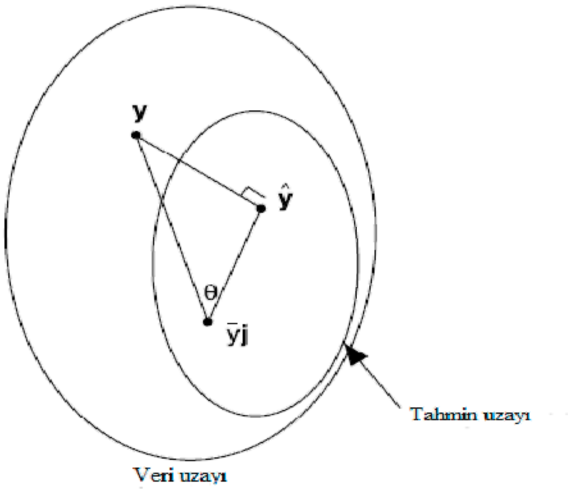
olarak tanımlanır.

Örnek 5.12.1: Çizelge 5.3. de verilen veriler için, R^2 'yi (5.12.4) bağıntısıyla ve R_a^2 'yi (5.12.6) ile elde ederiz. $\hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{Y}$ ve $\mathbf{Y}'\mathbf{Y}$ 'nin değerleri örneklerde verilirler.

$$\begin{aligned} R^2 &= \frac{\hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y} - n\bar{\mathbf{y}}^2}{y'y - n\bar{y}^2} = \frac{814.5410 - 12(7.5)^2}{840 - 12(7.5)^2} \\ &= \frac{139.5410}{165.0000} = .8457, \\ R_a^2 &= \frac{(n-1)R^2 - k}{n-k-1} = \frac{(11)(.8457) - 2}{9} = .8114 \end{aligned}$$

olur. (5.10.15) ve (5.10.17) yı kullanarak, (5.12.4) deki R^2 'yi örneklem varyansları ve örneklem kovaryanslarına dayanarak aşağıdaki gibi elde ederiz:

$$R^2 = \frac{\hat{\boldsymbol{\beta}}_1' \mathbf{X}_c' \mathbf{X}_c \hat{\boldsymbol{\beta}}_1}{\sum_{i=1}^{n} (y_i - \bar{y})^2} = \frac{\mathbf{s}_{yx}' \mathbf{S}_{xx}^{-1} (n-1) \mathbf{S}_{xx} \mathbf{S}_{xx}^{-1} \mathbf{s}_{yx}}{\sum_{i=1}^{n} (y_i - \bar{y})^2} = \frac{\mathbf{s}_{yx}' \mathbf{S}_{xx}^{-1} \mathbf{s}_{yx}}{s_y^2} \tag{5.12.7}$$



Şekil 5.17. $\boldsymbol{\theta}$**,** $\mathbf{y} - \bar{\mathbf{y}}\mathbf{1}$ ve $\hat{\mathbf{y}} - \bar{\mathbf{y}}\mathbf{1}$ arasındaki açı olmak üzere, $\boldsymbol{\theta}$'nın kosinüsü olarak R çoklu kolerasyonu.

R^2 nin bu biçimi rasgele –x durumu için R^2 ile bir karşılaştırma kolaylığı getirecektir. Geometrik olarak, R, ortalamaya göre düzeltilmiş $\mathbf{y}$ ve $\hat{\mathbf{y}}$ arasındaki $\boldsymbol{\theta}$ açısının kosinüsüdür. $\hat{\mathbf{y}}_1, \hat{\mathbf{y}}_2, \ldots, \hat{\mathbf{y}}_n$'nın ortalaması $\bar{\mathbf{y}}$,yani $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n$'nın ortalamasının aynıdır. Merkezileştirilmiş biçimleri sırasıyla $\mathbf{y} - \bar{\mathbf{y}}1$ ve $\hat{\mathbf{y}} - \bar{\mathbf{y}}1$ dır. Onların arasındaki açı Şekil 5.17 de gösterilmiştir. ($\mathbf{X}$' birinci sütununun bir katı olduğundan, $\bar{\mathbf{y}}1$nın tahmin uzayında olduğuna dikkat edelim.)

$\cos\theta$'mn, (5.12.4) deki gibi verilen R^2nın kareköküne eşit olduğunu göstermek için, iki vektör arasındaki açının kosinüsü için (5.12.8) bağıntısını kullanırız. Buna göre,

$$cos\theta = \frac{(\mathbf{y} - \bar{\mathbf{y}}\mathbf{1})'(\hat{\mathbf{y}} - \bar{\mathbf{y}}\mathbf{1})}{\sqrt{\left[(\mathbf{y} - \bar{\mathbf{y}}\mathbf{1})'(\mathbf{y} - \bar{\mathbf{y}}\mathbf{1})\right]\left[(\hat{\mathbf{y}} - \bar{\mathbf{y}}\mathbf{1})'(\hat{\mathbf{y}} - \bar{\mathbf{y}}\mathbf{1})\right]}} \tag{5.12.8}$$

dır.

(5.12.8) bağıntısını basitleştirmek için, Şekil 5.17 de geometrik olarak da görülebilen, $\mathbf{y} - \bar{\mathbf{y}}\mathbf{1} = (\hat{\mathbf{y}} - \bar{\mathbf{y}}\mathbf{1}) + (\mathbf{y} - \hat{\mathbf{y}})$ özdeşliğini kullanırız. $\hat{\mathbf{y}} - \bar{\mathbf{y}}\mathbf{1}$ tahmin uzayında olduğundan, bu özdeşliğin sağ yanındaki $\hat{\mathbf{y}}-\bar{\mathbf{y}}\mathbf{1}$ ve $\mathbf{y}-\hat{\mathbf{y}}$ vektörleri ortogonaldırlar. Bu nedenle (5.12.8) ın payı

$$\begin{aligned}(\mathbf{y}-\bar{\mathbf{y}}\mathbf{1})'(\hat{\mathbf{y}}-\bar{\mathbf{y}}\mathbf{1}) &= \left[(\hat{\mathbf{y}}-\bar{\mathbf{y}}\mathbf{1})+(\mathbf{y}-\hat{\mathbf{y}})\right]'(\hat{\mathbf{y}}-\bar{\mathbf{y}}\mathbf{1}) \\ &= (\hat{\mathbf{y}}-\bar{\mathbf{y}}\mathbf{1})'(\hat{\mathbf{y}}-\bar{\mathbf{y}}\mathbf{1})+(\mathbf{y}-\hat{\mathbf{y}})'(\hat{\mathbf{y}}-\bar{\mathbf{y}}\mathbf{1}) \\ &= (\hat{\mathbf{y}}-\bar{\mathbf{y}}\mathbf{1})'(\hat{\mathbf{y}}-\bar{\mathbf{y}}\mathbf{1})+\mathbf{0}\end{aligned}$$

olarak yazılabilir. Bu durumda;

$$\cos\theta = \frac{\sqrt{(\hat{\mathbf{y}}-\bar{\mathbf{y}}\mathbf{1})'(\hat{\mathbf{y}}-\bar{\mathbf{y}}\mathbf{1})}}{\sqrt{\left[(\mathbf{y}-\bar{\mathbf{y}}\mathbf{1})'(\mathbf{y}-\bar{\mathbf{y}}\mathbf{1})\right]}} = R \tag{5.12.9}$$

olur ki, bunun (5.12.4) da verilen R^2 nın karekökü olacağı kolayca gösterilir. Bu, (5.12.4) bağıntısını izleyen özellikle 2 ye eş değerdir, yani $R = \mathbf{r}_{xy}$ dır.

(5.12.10) bağıntısını

$$R^2 = \frac{\sum_{i=1}^{n}(\hat{\mathbf{y}}_i-\bar{\mathbf{y}})^2}{\sum_{i=1}^{n}(\mathbf{y}_i-\bar{\mathbf{y}})^2} = \frac{RKT}{GKT}$$

biçiminde yazabiliriz. Burada $RKT = \sum_{i=1}^{n}(\mathbf{y}_i-\bar{\mathbf{y}})^2$, $\mathbf{y}_i$ lar için bir kareler toplamıdır. Bu taktirde (5.12.1) in aşağısındaki $GKT = RKT + HKT$ parçalanması, basit lineer regresyon için başka bir bağıntıya benzer olan

$$\sum_{i=1}^{n}(\mathbf{y}_i-\bar{\mathbf{y}})^2 = \sum_{i=1}^{n}(\hat{\mathbf{y}}_i-\bar{\mathbf{y}})^2 + \sum_{i=1}^{n}(\mathbf{y}_i-\hat{\mathbf{y}}_i)^2$$

olarak yazılabilir.

5.13. LİNEER MODELDE EN İYİ LİNEER YANSIZ TAHMİN EDİCİ (EİLYTE) NİN BİR GEOMETRİK GÖRÜNÜM

$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta}+\boldsymbol{\varepsilon}$ $\quad E(\boldsymbol{\varepsilon})=0$ $\quad D(\boldsymbol{\varepsilon}) = \sigma^2\mathbf{V}$ model denklemini göz önüne alalım. $\boldsymbol{\varepsilon}$'na hata denirken $\mathbf{X}\boldsymbol{\beta}$ terimi $\mathbf{Y}$'nin sistematik kısmıdır. $\boldsymbol{\beta}$'ya bilinmeyen olarak bakmak suretiyle, model; sistematik kısmın, $\mathbf{X}$'in sütunlarının bir bilinmeyen lineer kombinasyonu ile verildiğini varsayar. Model denkleminin deneysel eşdeğeri

$$\mathbf{Y} = \mathbf{X}\hat{\boldsymbol{\beta}} + \mathbf{e} = \hat{\mathbf{Y}} + \mathbf{e}$$

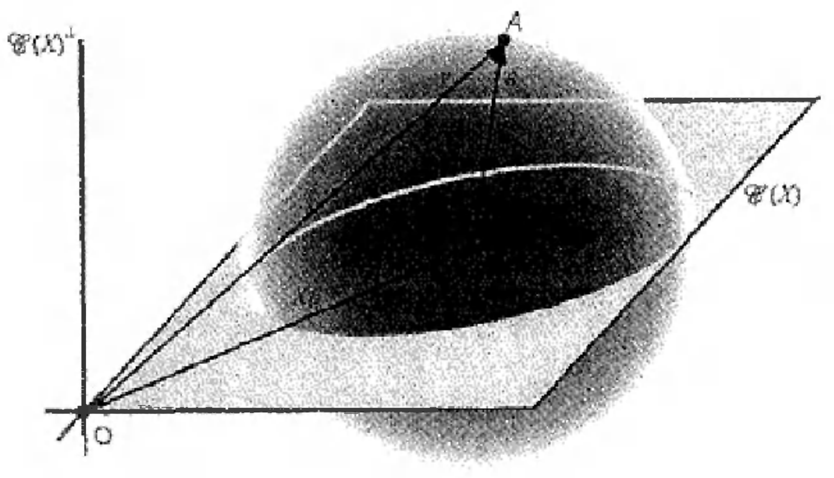
dir. Burada $\hat{\mathbf{Y}} = \mathbf{X}\hat{\boldsymbol{\beta}}$ sistematik kısmın tahmin edicisi olan uygun değerlerin vektörüdür. Tahmin edilen $\mathbf{e}$ açıklanamayan kısım iken, $\hat{\mathbf{Y}}$' ya $\mathbf{Y}$ nin açıklanan kısmı diyeceğiz. Modeli yorumlamanın başka bir şekli aşağıda verilmiştir.

Hata kısmı sıfır (0) ortalama ve $\sigma^2\mathbf{V}$ varyansına sahip iken model; sistematik kısmın $C(\mathbf{X})$ ($\mathbf{X}$'in sütun uzayı) sütun uzayında olduğunu varsayar. Bu yorumda vurgulama "sistematik" kısım olarak belirlenen vektör uzayı üzerindedir. Açıklayıcı değişkenlerinin ayrımsızlığı önemli değildir. Yeniden bir parametreleme vektör uzaylarını korur ve bu nedenle bu parametrelemenin $\hat{\mathbf{Y}}$ ve $\mathbf{e}$' nin aynı değerlerini üretmesi beklenir. Bu yeniden parametreleme, en iyi lineer yansız tahminin vektör uzayı yorumundan öğrenilecek çok şey olduğunu ortaya koyar. Bu kısmın amacı tamamiyle geometrik görüş açısından en iyi lineer yansız tahmin ediciyi değerlendirmektir.

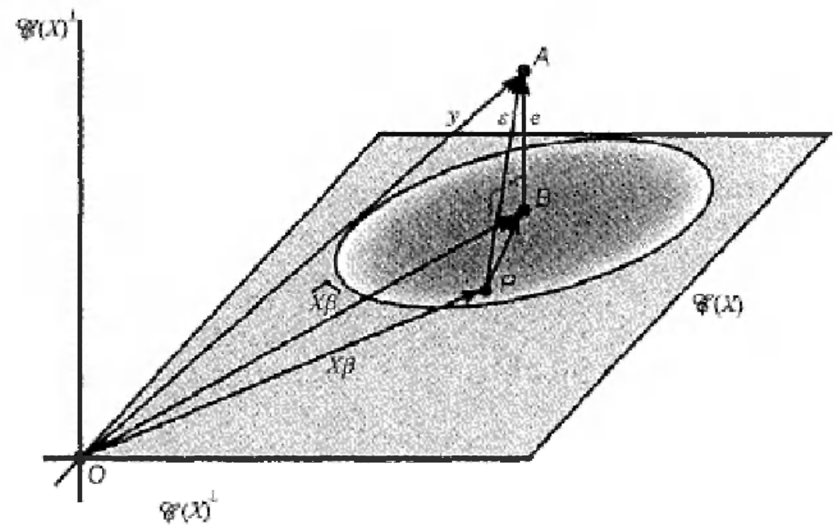
5.14. VARYANSLARIN EŞİT OLDUĞU DURUM

$\mathbf{V} = \mathbf{I}$ olsun. Şekil 5.18 $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ ayrışımını açıklar. $\mathbf{Y}$, $\mathbf{X}\boldsymbol{\beta}$ ve $\boldsymbol{\varepsilon}$ sırasıyla OA, OP ve PA doğru parçaları ile gösterilirler. Yatay düzlem $C(\mathbf{X})$'i gösterirken, düşey eksen $C(\mathbf{X})^{\perp}$ uzayını gösterir. Gerçekte bu uzayların her ikisi de ikiden daha büyük boyutlara sahiptirler ve özellikle $C(\mathbf{X})^{\perp}$, $C(\mathbf{X})$ den daha büyük bir boyuta sahip olmalıdır. Göz önünde canlandırmak için, $C(\mathbf{X})^{\perp}$ ve $C(\mathbf{X})$ için sırasıyla 1 ve 2 boyutlarını seçeriz. P noktası etrafındaki renkli bölge başka bir rasgele denemede A noktasının muhtemelen görülebildiği bölgeyi gösterir. Bölgenin çekirdeğindeki daha karanlık gölge orada olasılık kütlesinin daha yüksek konsantrasyonunu (yoğunluğunu) gösterir.

Model OP'nin yatay düzlem üzerindeki bir yerde uzandığını belirler. Tahmin problemi OP vektörünü belirlenmesinden ibarettir. Şekil 5.19 $C(\mathbf{X})$ düzleminde B nin tüm mümkün yerleri üzerinde BA doğrusunun uzunluğunu minimumlaştırmak suretiyle elde edilen OB doğrusu ile gösterilen, $\mathbf{X}\boldsymbol{\beta}$'nın en iyi lineer yansız tahmin edicisinin bireşimini açıklar.

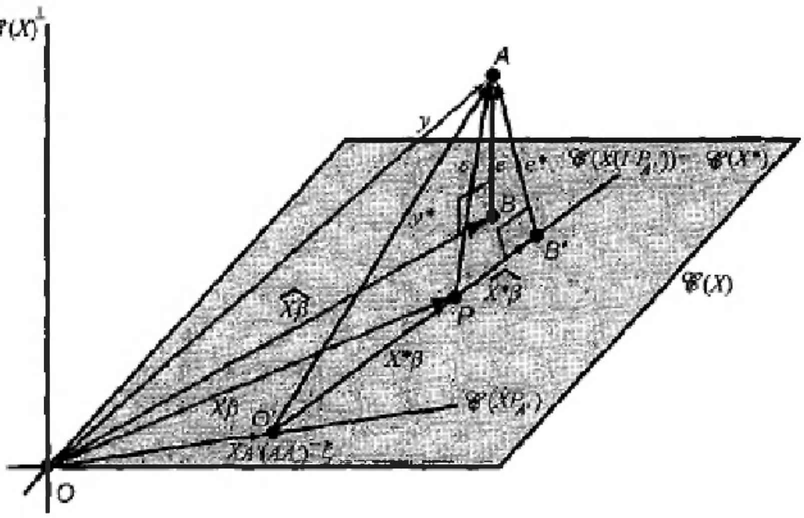


Şekil 5.18. Eşit Varyanslı Lineer Modelin Bir Geometrik Görünümü



Şekil 5.19. Eşit Varyanslı Lineer Modelde EİLYTE nin Bir Geometrik Görünümü

AB yatay düzleme dik (ortogonal) olduğunda minimum uzunluk ortaya çıkar. Bu nedenle $\mathbf{Y} = \hat{\mathbf{Y}} + \mathbf{e}$ ayrışımında $\hat{\mathbf{Y}}$ ve $\mathbf{e}$ 'ye karşılık gelen OB ve BA 'yı elde ederiz. Açık olarak, $\hat{\mathbf{Y}}$ $C(\mathbf{X})$ üzerinde $\mathbf{Y}$ 'nin ortogonal izdüşümüdür. Bu nedenle $\hat{\mathbf{Y}} \in C(\mathbf{X})$ ve $\mathbf{e} \in C(\mathbf{X})^{\perp}$ dir. B etrafındaki eliptik bölge $\mathbf{X\beta}$ için tipik bir güven bölgesidir. Bu bölge tamamiyle $C(\mathrm{D}(\hat{\mathbf{Y}})) = C(\mathbf{X})$ gerçeğine karşılık gelen $C(\mathbf{X})$ düzleminde yer alır.

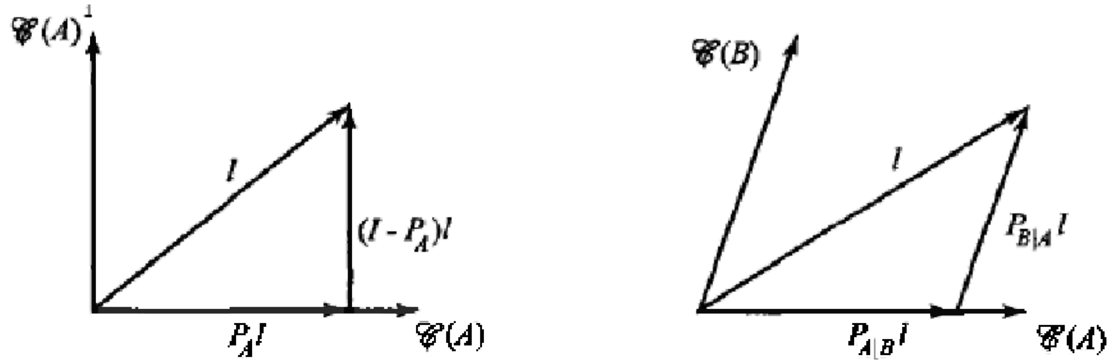


Şekil 5.20. Lineer Modelde Kısıtlanmış EİLYTE nin Bir Geometrik Görünümü

5.15. LİNEER KISITLAMALARIN ETKİSİ

Eşit varyanslı durumda $\mathbf{A\beta = \xi}$ kısıtlamasının etkisi Şekil 5.19 un bir genişlemesi olan, Şekil 5.20 de açıklanır. Bu diyagram (şekil) $\mathbf{(Y - XA'(AA')^{-}\xi, X(I - P_{A'}), \sigma^2 I)}$ "eşdeğer" modeli bağlamında anlaşılabilir. Gerçekten $\mathbf{Y = X\beta + \varepsilon}$ modelinde $\boldsymbol{\beta}$ parametre vektörü üzerinde $\mathbf{A\beta = \xi}$ kısıtlaması konulduğunda $\mathbf{\beta = A^{-}\xi + (I - A^{-}A)\theta}$ dır. $\boldsymbol{\beta}$ nın bu değeri $\mathbf{Y = X\beta + \varepsilon}$ modelinde yerine konulduğunda

$\mathbf{(Y - XA^{-}\xi, X\left[I - A^{-}A\right]\theta, \sigma^2 I)}$ kısıtlanmamış modeli elde edilir. Burada $\mathbf{A^{-} = A'(AA')^{-}}$ dır ve böylece $\mathbf{(Y - XA'(AA')^{-}\xi, X(I - P_{A'}), \sigma^2 I)}$ kısıtlanmış modeli elde edilir. Burada $\mathbf{P_{A'} = A'(AA')^{-}A}$ dır. $\mathbf{XA'(AA')^{-}\xi}$ $\mathbf{Y}$ 'nin, kısıtlamalardan dolayı, tam olarak bilinen bir parçasıdır. $\mathrm{OO'}$ doğru parçası ile gösterilen bu vektör $C(\mathbf{X})$ 'in düzlemi üzerinde koyu bir doğru ile gösterilen $C(\mathbf{XP_{A'}})$ düzleminde uzanır. Diğer koyu doğru $\mathbf{u} \in C(\mathbf{X(I - P_{A'})})$ olmak üzere $\mathbf{XA'(AA')^{-}\xi + u}$ olarak yazılabilen tüm vektörlerin uçlarını geometrik gösterir. P noktası (burada OP bilinmeyen $\mathbf{X\beta}$ dir) bu doğru üzerinde herhangi bir yerde yer alır. $\mathrm{OO'}$ tamamen bilindiğinden tahmin problemi orjini $\mathrm{O'}$ ye kaydırmak suretiyle basitleştirilebilir. Bu yer değiştirmeye bağlı $\mathrm{O'}$ ve P den geçen koyu doğru $C(\mathbf{X(I - P_{A'})})$ vektör uzayını gösterir. Tahmin problemi OP doğru parçasını bulmaya ya da bu doğru üzerindeki P noktasını bulmaya indirgenir. "Kısıtlanmış" en iyi lineer yansız tahmin edicinin B'A nın uzunluğu en küçük olacak şekilde bu doğru üzerindeki bir B' noktasının bulmak suretiyle elde edilir. B'A bu doğruya dik olduğunda ve özellikle O'B' ne dik olduğunda bu başarılır. Eğer $(\mathbf{Y}-\mathbf{XA'(AA')^{-}\xi}, \mathbf{X(I}-\mathbf{P_{A'}}), \sigma^2\mathbf{I})$ modeline kısaca $(\mathbf{Y}^*, \mathbf{X}^*\boldsymbol{\beta}, \sigma^2\mathbf{I})$ başvurursak bu taktirde $\mathbf{Y}^*, \mathbf{X}\boldsymbol{\beta}^*$ ve $\mathbf{e}^*$ sırasıyla O'A, O'B' ve B'A doğru parçaları ile gösterilirler.



Şekil 5.21. Ortogonal ve Eğik İzdüşümler

$\mathbf{X}\boldsymbol{\beta}^*$ vektörü $\mathbf{Y}^*$'ın $C(\mathbf{X}^*)$ üzerinde ortogonal izdüşümüdür. OB' doğru parçası , $\mathbf{A}\boldsymbol{\beta}=\xi$ kısıtlaması altında $\mathbf{Y}$'nin uygun değerleri olan $\hat{\mathbf{Y}}_{\mathbf{kisit}}$ değerini gösterir. B'B doğru parçasına $\hat{\mathbf{Y}}_{\mathbf{kisit}}-\hat{\mathbf{Y}}$ veya $\mathbf{e}^*-\mathbf{e}$ olarak bakılabilir. OB' ve B'B vektörleri şekilde açık olarak gösterilmemişlerdir.

5.16. GENEL LİNEER MODEL

Genel lineer modelde $\mathbf{X}\boldsymbol{\beta}$'nın EİLYTE sine bir izdüşüm olarak bakmak için ortogonal izdüşümden ziyade eğik izdüşümleri kullanmamız gerekir. Eğer A ve B aynı satır sayıları olan matrisler ise ve $C(\mathbf{A})$ ve $C(\mathbf{B})$ karşılıklı olarak ayrık iseler bu taktirde $C(\mathbf{A}:\mathbf{B})$ deki herhangi bir l vektörünün, $l_1 \in C(\mathbf{A})$ ve $l_2 \in C(\mathbf{B})$ olmak üzere tek bir l_1+l_2 gösterimine sahip olduğuna dikkat edelim. (Eğer $l_{1^*}+l_{2^*}$ böyle başka bir gösterim ise bu taktirde $l_1-l_{1^*}=l_2-l_{2^*}$ $C(\mathbf{B})$'ye ait olduğu gibi $C(\mathbf{A})$'ya da aittir. Bu ise bu sütun uzaylarının karşılıklı olarak ayrık olmadıklarını ifade eder.) Bu ayrışımı mümkün kılacak olan bir izdüşüm matrisine ihtiyaç duyarız.

Tanım 5.16.1: Eğer her $l \in C(\mathbf{A}:\mathbf{B})$ için $P_{A/B}l \in C(\mathbf{A})$ ve $(I - P_{A/B})l \in C(\mathbf{B})$ ise bir $\mathbf{P}_{A/B}$ matrisine $C(\mathbf{B})$ boyunca veya $C(\mathbf{B})$'ye paralel olarak $C(\mathbf{A})$ üzerinde bir izdüşürücü denir. Burada $\mathbf{P}_{\mathbf{AB}}$ $\mathbf{A(B'A)^{-}B'}$ matrisidir. Bu matris idempotenttir.

Gerçekten $\mathbf{P}_{A/B}.\mathbf{P}_{A/B} = \mathbf{A(B'A)^{-}B'A(B'A)^{-}B'} = \mathbf{A(B'A)^{-}B'} = \mathbf{P}_{A/B}$, $\mathbf{P}_{B/A} = \mathbf{B(A'B)^{-}A'}$ matrisi de idemptenttir.

$rank(A_{mxn}) = m, \quad rank(B_{mxn}) = m$

$\mathbf{P}_{B/A}.\mathbf{P}_{B/A} = \mathbf{B(A'B)^{-}A'B(A'B)^{-}A'} = \mathbf{B(A'B)^{-}A'} = \mathbf{P}_{B/A}$ Moore Penrose g-ters 1 ve 2 kuralından bu sağlanır.

Şekil 5.21, her iki uzayda 1 boyutlu olmak üzere bir sütun uzayı üzerine ortogonal izdüşüm ve başka bir uzay boyunca sütun uzayı üzerine eğik izdüşüm arasındaki zıtlığı açıklar.

Asıl fark $\mathbf{P}_{A/B}l$ ve $\mathbf{P}_{B/A}l$'nin ortogonal olması gerekmezken, $\mathbf{P}_{A}l$ ve $(\mathbf{I}-\mathbf{P}_{A})l$'nin birinin diğerine ortogonal olmasıdır. $C(\mathbf{A}:\mathbf{B})$ tam satır ranklı olduğunda $\mathbf{P}_{A/B}$ tektir ve idempotenttir. Aksi halde $\mathbf{P}_{A/B}$ ne tek ne de idempotent olabilir. Her hangi bir durumda herhangi bir $l \in C(\mathbf{A}:\mathbf{B})$ için $\mathbf{P}_{A/B}l$ izdüşümü bir tanedir. $\mathbf{I}-\mathbf{P}_{A/B}$'nın $\mathbf{P}_{B/A}$'nın bir seçimi olduğu gerçeklenebilir. Aynı zamanda, $\mathbf{P}_{A}$ da $\mathbf{B}$, $(\mathbf{I}-\mathbf{P}_{A})$ veya $C(\mathbf{A})^{-}$'yı geren diğer herhangi bir matris olduğunda, $\mathbf{P}_{A/B}$'nın bir tek seçimidir. Şimdi $\mathbf{P}_{A/B}$'nin yararlı bir seçimini vereceğiz.

Yardımcı Teorem 5.16.1: **A** ve **B** aynı satır sayılarına sahip matrisler olsunlar.

(a) $C(\mathbf{A})$ ve $C(\mathbf{B})$'nin karşılıklı olarak ayrık olmaları için gerek ve yeter şart $C(\mathbf{A}') = C(\mathbf{A}'(\mathbf{I}-\mathbf{P}_{B}))$ olmasıdır.

(b) Eğer $C(\mathbf{A})$ ve $C(\mathbf{B})$ karşılıklı olarak ayrık iseler, bu taktirde $\mathbf{A}[(\mathbf{I}-\mathbf{P}_{B})\mathbf{A}]^{-}(\mathbf{I}-\mathbf{P}_{B})$ $C(\mathbf{B})$ boyunca $C(\mathbf{A})$ üzerine bir izdüşürücüdür.

İspat: $C(\mathbf{A})$ ve $C(\mathbf{B})$'nin karşılıklı olarak ayrık olduklarını varsayalım. $C(\mathbf{A}'(\mathbf{I}-\mathbf{P}_{B})) \subseteq C(\mathbf{A}')$. Eğer kapsama tam ise veya kesin ise $\mathbf{k}$ $C(\mathbf{A}')$de $C(\mathbf{A}'(\mathbf{I}-\mathbf{P}_{B}))$ ye dik olan bir vektör olsun. $\mathbf{Ak} \in C(\mathbf{B})$ olduğu görülür. $C(\mathbf{A})$ ve $C(\mathbf{B})$ karşılıklı olarak ayrık olduklarından **Ak** (sıfır) 0 olmak zorundadır. Çünkü $\mathbf{k} \in \mathcal{C}(\mathbf{A}')$ olduğundan $\mathbf{A}'\mathbf{t} = \mathbf{k}$, $\mathbf{A}\mathbf{A}'\mathbf{t} = \mathbf{A}\mathbf{k} = \mathbf{0}$ olmalıdır. $\mathbf{A}'\mathbf{m}$ formunda olan $\mathbf{k}$ nın kendisinin de sıfır olacağını ifade eder. Tersine olarak eğer $\mathcal{C}(\mathbf{A}') \subseteq \mathcal{C}(\mathbf{A}'(\mathbf{I} - \mathbf{P_B}))$ ise herhangi bir $\mathbf{T}$ matrisi için $\mathbf{A} = \mathbf{T}(\mathbf{I} - \mathbf{P_B})\mathbf{A}$ olsun. Eğer herhangi bir u_1 ve u_2 vektörleri için $\mathbf{Au_1} = \mathbf{Bu_2}$ ise $\mathbf{Au_1} = \mathbf{T}(\mathbf{I} - \mathbf{P_B})\mathbf{Au_1} = \mathbf{T}(\mathbf{I} - \mathbf{P_B})\mathbf{Bu_2} = \mathbf{0}$ bağıntısını elde ederiz. Bu nedenle $\mathcal{C}(\mathbf{A})$ ve $\mathcal{C}(\mathbf{B})$ karşılıklı olarak ayrıktırlar. Bu (a) kısmının ispatını tamamlar. (b) yi ispatlamak için $\mathbf{P_{A|B}} = \mathbf{A}[(\mathbf{I} - \mathbf{P_B})\mathbf{A}]^-(\mathbf{I} - \mathbf{P_B})$ olsun ve $\boldsymbol{l}$ $\mathcal{C}(\mathbf{A}:\mathbf{B})$ de bir vektör olsun. $\boldsymbol{l}_1 \in \mathcal{C}(\mathbf{A})$ ve $\boldsymbol{l}_2 \in \mathcal{C}(\mathbf{B})$ olmak üzere $\boldsymbol{l}$'nin tek bir $\boldsymbol{l}_1 + \boldsymbol{l}_2$ gösterimine sahip olduğunu varsayalım. $\mathbf{P_{A|B}}\boldsymbol{l} \in \mathcal{C}(\mathbf{A})$ olduğunu görmek kolaydır. Öte yandan;

$$
\begin{aligned}
(\mathbf{I} - \mathbf{P_{A|B}})\boldsymbol{l} &= \boldsymbol{l} - \mathbf{A}[(\mathbf{I} - \mathbf{P_B})\mathbf{A}]^-(\mathbf{I} - \mathbf{P_B})\boldsymbol{l} \\
&= \boldsymbol{l} - \mathbf{T}(\mathbf{I} - \mathbf{P_B})\mathbf{A}[(\mathbf{I} - \mathbf{P_B})\mathbf{A}]^-(\mathbf{I} - \mathbf{P_B})\boldsymbol{l}] \\
&= \boldsymbol{l} - \mathbf{T}(\mathbf{I} - \mathbf{P_B})\boldsymbol{l}_1 = \boldsymbol{l} - \boldsymbol{l}_1 = \boldsymbol{l}_2 \text{ dir.}
\end{aligned}
$$

Yukarıda $\mathbf{T}$, $\mathbf{A}' = \mathbf{A}'(\mathbf{I} - \mathbf{P_B})\mathbf{T}'$ olacak şekilde bir matristir. $\boldsymbol{l}_2$'nin $\mathcal{C}(\mathbf{B})$ de olduğuna dikkat edelim. Şimdi bir eğik izdüşürücü kavramını kullanarak genel lineer modelde $\mathbf{X}\boldsymbol{\beta}$'nın en iyi lineer yansız tahmin edicisini ele alacağız.

Yardımcı teorem 5.16.2: $\mathcal{M}$, parametreler üzerinde bilinen kısıtlamanın olmadığı $(\mathbf{Y}, \mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{V})$ lineer modeli olsun. Bu taktirde,

(a) $\boldsymbol{\varepsilon}_r = \mathcal{C}(\mathbf{X})^{\perp}$;

(b) $\boldsymbol{\varepsilon}_s = \mathcal{C}(\mathbf{V}(\mathbf{I} - \mathbf{P}_X))$;

(c) $\boldsymbol{\varepsilon}_r \cap \boldsymbol{\varepsilon}_s = \mathcal{C}(\mathbf{V}:\mathbf{X})^{\perp}$;

(d) $\boldsymbol{\varepsilon}_r + \boldsymbol{\varepsilon}_s = \mathbf{R}^n$ dır.

Parametreler üzerinde bilinen kısıtlama olmadığında (a),(c) ve (d) kısımlarının doğruluğu için (Sengupta ve Jammalamadaka 2003).

(c) Kısmını ispatlamak için, $l'\mathbf{Y}$ nin $\mathcal{M}$ de bir EİLYTE olması için gerek ve yeter şartın onun $(\mathbf{I}-\mathbf{P_X})\mathbf{Y}$ ile ilişkisiz olması gerektiğine dikkat edelim. Bu şart $l'\mathbf{V}(\mathbf{I}-\mathbf{P_X})=\mathbf{0}$, veya $l \in \mathcal{C}(\mathbf{V}(\mathbf{I}-\mathbf{P_X}))^{\perp}$ eşitliğine eşdeğerdir.

Uyarma 5.16.1: Eğer parametre uzayı üzerinde bilinen kısıtlama yoksa ve $\mathbf{Y}_i=\mathbf{L}_i\mathbf{Y}, i=1,2,3,4$ gibi iseler bu taktirde $\boldsymbol{\varepsilon}_r=\mathcal{C}(\mathbf{L_3}':\mathbf{L_4}')$, $\boldsymbol{\varepsilon}_s=\mathcal{C}(\mathbf{L_1}',\mathbf{L_2}',\mathbf{L_4}')$ olarak tanımlanır(Sengupta ve Jammalamadaka 2003).

Yardımcı teorem 5.16.3: $(\mathbf{Y},\mathbf{X}\boldsymbol{\beta},\sigma^2\mathbf{V})$ genel lineer modelini göz önüne alalım.

(a) $\mathbf{Y}$ cevap (tepki-gözlemler) vektörü hemen hemen kesinlikle (neredeyse kesin olarak) $\mathcal{C}(\mathbf{V}(\mathbf{I}-\mathbf{P_X}):\mathbf{X})$ uzayında uzanır.

(b) $\mathcal{C}(\mathbf{V}(\mathbf{I}-\mathbf{P_X}))$ ve $\mathcal{C}(\mathbf{X})$ karşılıklı olarak ayrıktırlar.

(c) $\mathbf{X}\boldsymbol{\beta}$ ' nın EİLYTE si hemen hemen kesin olarak $\mathbf{P}_{\mathbf{X}|\mathbf{V}(\mathbf{I}-\mathbf{P_X})}\mathbf{Y}$ ya eşittir.

İspat:

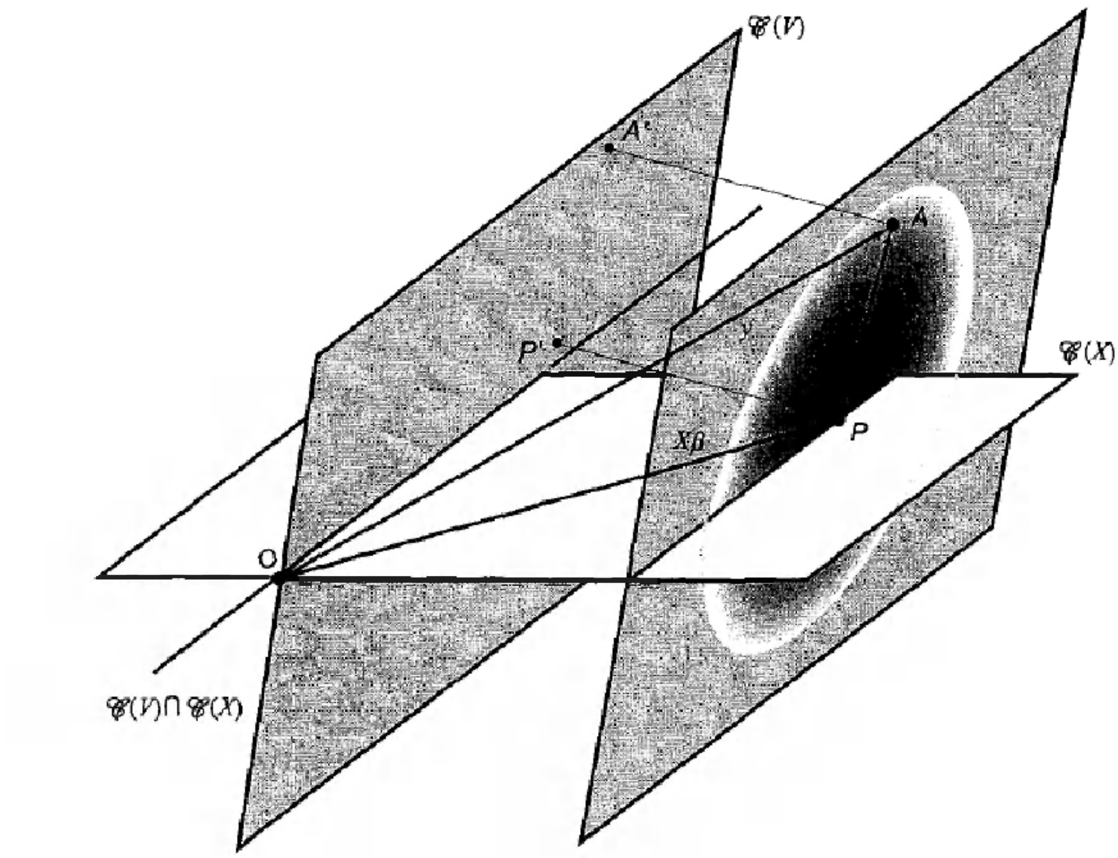
(a) Hemen hemen kesin olarak $\mathbf{Y}\in\mathcal{C}(\mathbf{V}:\mathbf{X})$ olduğundan, $\mathcal{C}(\mathbf{V}:\mathbf{X})=\mathcal{C}(\mathbf{V}(\mathbf{I}-\mathbf{P_X}):\mathbf{X})$ göstermek yeter. $\mathcal{C}(\mathbf{V}(\mathbf{I}-\mathbf{P_X}):\mathbf{X})$ in $\mathcal{C}(\mathbf{V}:\mathbf{X})$ de kapsandığı açıktır. Ters kapsamayı göstermek için, $l\in\mathcal{C}(\mathbf{V}(\mathbf{I}-\mathbf{P_X}):\mathbf{X})^{\perp}$ olsun. $\mathbf{X}'l=\mathbf{0}$ olduğundan, l 'yi $(\mathbf{I}-\mathbf{P_X})\mathbf{m}$ gibi yazabiliriz. $l'\mathbf{V}(\mathbf{I}-\mathbf{P_X})=\mathbf{0}$ olduğundan dolayı $\mathbf{m}'(\mathbf{I}-\mathbf{P_X})\mathbf{V}(\mathbf{I}-\mathbf{P_X})=\mathbf{0}$ elde ederiz yani $\mathbf{m}'(\mathbf{I}-\mathbf{P_X})\mathbf{V}=\mathbf{0}$ elde ederiz. Bu nedenle $l'(\mathbf{V}:\mathbf{X})=\mathbf{0}$ dır. $\mathcal{C}(\mathbf{V}(\mathbf{I}-\mathbf{P_X}):\mathbf{X})^{\perp}\subseteq\mathcal{C}(\mathbf{V}:\mathbf{X})$ olduğu görülür.

(b) Kısmı $\mathbf{A}=\mathbf{V}(\mathbf{I}-\mathbf{P_X})$ ve $\mathbf{B}=\mathbf{X}$ seçmek suretiyle yardımcı teorem 5.3.1. (a) kısmından doğrudan doğruya görülür. Yardımcı teorem 5.3.1.in (b) kısmını kullanarak ve sonuçta oluşan ifadeyi, model $(\mathbf{Y},\mathbf{X}\boldsymbol{\beta},\sigma^2\mathbf{V})$ ve $\mathbf{X}$ eksik ranklı olduğunda elde edilen

$$\mathbf{e}=\mathbf{Y}-\mathbf{Y}=\mathbf{V}(\mathbf{I}-\mathbf{P_X})\{(\mathbf{I}-\mathbf{P_X})\mathbf{V}(\mathbf{I}-\mathbf{P_X})\}^{-}(\mathbf{I}-\mathbf{P_X})\mathbf{Y}$$

Sengupta ve Jammalamadaka (2003) ifadesiyle karşılaştırarak $\mathbf{e}=\mathbf{P}_{\mathbf{V}(\mathbf{I}-\mathbf{P_X})}\mathbf{Y}$ elde ederiz. Bu taktirde (c) kısmı $\mathbf{I}-\mathbf{P}_{\mathbf{A}|\mathbf{B}}$ 'nin $\mathbf{P}_{\mathbf{B}|\mathbf{A}}$ 'nın bir seçimi olduğu gerçeğinden görülür.

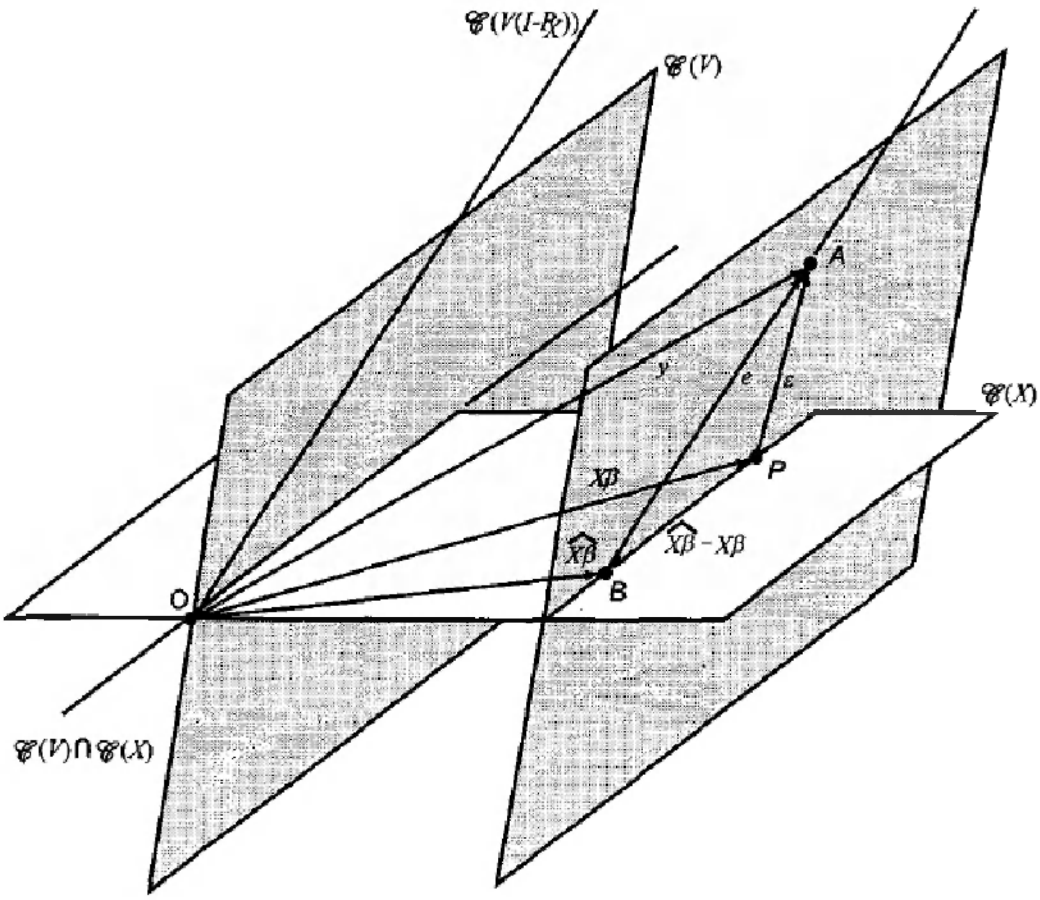
Yardımcı teoremin 5.16.2 uyarınca gözlem vektörünün ayrışımı ve tahmin ve hata uzayları arasında enteresan bir ilişki vardır. Bu yardımcı teoreme göre, $\mathbf{Y}$ tepkimesi $\mathbf{Y}_1 \in C(\mathbf{X})$ ve $\mathbf{Y}_2 \in C(\mathbf{V}(\mathbf{I}-\mathbf{P}_X))$ olmak üzere, $\mathbf{Y}_1 + \mathbf{Y}_2$ olarak bir tek şekilde ayrıştırılabilir. Bu, l vektörü tahmin uzayında bulunduğunda $l'\mathbf{Y}$ nin hemen hemen kesin olarak $l'\mathbf{Y}_1$ e eşit olduğu yardımcı teorem 5.16.2 den görülür. Aynı şekilde l , hata uzayında olduğunda, hemen hemen kesin olarak $l'\mathbf{Y} = l'\mathbf{Y}_2$ dir. Eğik izdüşümler hakkında daha fazla bilgi için Rao [1974] e bakınız. Şekil 5.22 muhtemelen $\mathbf{V}$ singüler matrisli $(\mathbf{Y}, \mathbf{X\beta}, \sigma^2\mathbf{V})$ genel lineer modelinde $\mathbf{Y}$ 'nin sentezini (bireşimini) gösterir. Sistematik kısım yani $\mathbf{X\beta}$ $C(\mathbf{X})$'i gösteren düzlemde uzanan OP doğrusu ile gösterilir. Eğer $\mathbf{V}$ singüler ise bu taktirde $\boldsymbol{\varepsilon}$, $C(\mathbf{V})$ uzayında uzanır. Bu düzlem üzerinde P noktası etrafındaki karartılmış bölge A noktasının başka bir rasgele örneklemde muhtemelen gözükebildiği bölgeyi gösterir.



Şekil 5.22. Tekil lineer modelin bir geometrik görünümü

$C(\mathbf{V})$'nin düzleminin O noktasından geçtiğini düşünelim. $\mathbf{PA(\varepsilon)}$ bu düzleme paralel olduğundan A ve P den $C(\mathbf{V})$ üzerine indirilen $\mathbf{A\,A}$ ve $\mathbf{P\,P}$ dikmeleri aynı uzunluğa sahip olmalıdır. Bu iki vektör sırasıyla $\mathbf{(I-PV)Y}$ ve $\mathbf{(I-PV)X\beta}$'yı gösterir. Bu nedenle, modelin singülerliğinden dolayı, kesin olarak $\mathbf{A\,A}$ ile bilinen $\mathbf{P\,P}$, OP nin parçasıdır.

Şekil 5.23 Genel lineer modelde EİLYTE' nin yapısını gösterir. $C(\mathbf{V(I-P_X)})$'nin, doğrusunun $C(\mathbf{X})$ düzlemine dik olmadığına dikkat edelim. Bu, $C(\mathbf{X})^{\perp}$ ekseninin $C(\mathbf{X})$'in düzlemine dik olduğu Şekil 5.22 ile çelişir.



Şekil 5.23. Genel lineer modelde EİLYTE'nin bir geometrik görünümü

$\mathbf{X\beta}$'nın EİLYTE si, OA'nın, $C(\mathbf{V(I-P_X)})$ boyunca $C(\mathbf{X})$ düzlemi üzerine eğik izdüşümünü düşürmek suretiyle elde edilir. Başka bir deyişle B noktası, $C(\mathbf{V(I-P_X)})$' e bir paralel doğru çizerek, bu doğrunun $C(\mathbf{X})$'in düzlemini kestiği yere konur. $C(\mathbf{V(I-P_X)})$,

$\mathcal{C}(\mathbf{V})$ de kapsandığından (içerildiğinden), BA (e hata tahmini vektörüne karşılık gelen doğru) doğrusu bu düzleme paraleldir. Bu nedenle, PB doğru parçası , P den geçen, $\mathcal{C}(\mathbf{V})\cap\mathcal{C}(\mathbf{X})$'e paralel olan doğruda yer alır. PB, $\mathbf{X\beta}$ nın tahminindeki hataya, yani $\mathbf{X\beta} - \mathbf{X\beta}$'ya karşılık gelir. Bu, D($\mathbf{X\beta}$) ile gerilen uzayın, niçin $\mathcal{C}(\mathbf{V})\cap\mathcal{C}(\mathbf{X})$ olduğunu ve $\mathbf{D(e)}$ ile gerilen uzayın, niçin $\mathcal{C}(\mathbf{V}(\mathbf{I} - \mathbf{P}_X))$ olduğunu açık olarak gösterir.

Bunun için aşağıdaki yardımcı teoremleri ifade ve ispat edelim.

Yardımcı teorem 5.16.4: a) Eğer $\mathbf{B}$ negatif definit (tanımlı) değilse, bu taktirde $\mathcal{C}(\mathbf{ABA'}) = \mathcal{C}(\mathbf{AB})$, ve $rank(\mathbf{ABA'}) = rank(\mathbf{AB}) = rank(\mathbf{BA'})$ dür.

b) Eğer $\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B'} & \mathbf{C} \end{pmatrix}$ negatif tanımlı olmayan bir matris ise, bu taktirde $\mathcal{C}(\mathbf{B}) \subseteq \mathcal{C}(\mathbf{A})$ ve $\mathcal{C}(\mathbf{B'}) \subseteq \mathcal{C}(\mathbf{C})$ dir.

İspat: $\mathbf{CC'}$ nün, $\mathbf{B}$'nin bir rank- çarpanlaması olduğunu farz edelim. Bu taktirde

$\mathcal{C}(\mathbf{ABA'}) = \mathcal{C}(\mathbf{ACC'A'}) \subseteq \mathcal{C}(\mathbf{ACC'}) \subseteq \mathcal{C}(\mathbf{AC})$ dir.

Bununla beraber, $\mathcal{C}(\mathbf{AC}) = \mathcal{C}((\mathbf{AC})(\mathbf{AC})') = \mathcal{C}(\mathbf{ABA'})$dür. Bu nedenle yukarıdaki tüm sütun uzayları aynıdırlar. Özellikle, $\mathcal{C}(\mathbf{AB}) = \mathcal{C}(\mathbf{ACC'}) = \mathcal{C}(\mathbf{ABA'})$dür. Sonuç olarak, $rank(\mathbf{ABA'}) = rank(\mathbf{AB}) = rank(\mathbf{BA'})$dür.

b) kısmını ispatlamak için, $\mathbf{TT'}$ verilen negatif tanımlı olmayan matrisin bir rank çarpanları olsun ve $\begin{pmatrix} \mathbf{T}_1 \\ \hline \mathbf{T}_2 \end{pmatrix}$, $\mathbf{T}_1$, $\mathbf{A}$ ile anı sayıda satıra sahip olacak şekilde $\mathbf{T}$'nin bir parçalanması olsun. Bu taktirde

$$\mathbf{TT'} = \begin{pmatrix} \mathbf{T}_1\mathbf{T}_1' & \mathbf{T}_1\mathbf{T}_2' \\ \mathbf{T}_2\mathbf{T}_1' & \mathbf{T}_2\mathbf{T}_2' \end{pmatrix} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B'} & \mathbf{C} \end{pmatrix}$$

elde ederiz. Blokları karşılaştırarak, $\mathbf{A} = \mathbf{T}_1\mathbf{T}_1'$ ve $\mathbf{B} = \mathbf{T}_1\mathbf{T}_2'$ elde ederiz. Ayrıca,

$\mathcal{C}(\mathbf{B}) = \mathcal{C}(\mathbf{T}_1\mathbf{T}_2') \subseteq \mathcal{C}(\mathbf{T}_1) = \mathcal{C}(\mathbf{TT}_1') = \mathcal{C}(\mathbf{A})$ dır. Bu muhakemeyi transpozesi alınan matris üzerinde tekrarlayarak, $\mathcal{C}(\mathbf{B'}) \subseteq \mathcal{C}(\mathbf{C})$ elde ederiz.

Yardımcı teorem 5.16.5: $\mathbf{A}_{m\times n}$ ve $\mathbf{B}_{m\times 1}$matrislerinin bilindiklerini farz edelim.

a) $\mathbf{Ax}=\mathbf{b}$ denklemlerinin bir çözüme sahip olması için gerek ve yeter şart $\mathbf{b}\in\mathcal{C}(\mathbf{A})$ olmasıdır.

b) $\mathbf{Ax}=\mathbf{b}$ denklemlerinin bir tek çözüme sahip olması için gerek ve yeter şart $\mathbf{b}\in\mathcal{C}(\mathbf{A})$ ve $rank(\mathbf{A})=n$ olmasıdır.

Eğer $\mathbf{b}\in\mathcal{C}(\mathbf{A})$ise, $\mathbf{Ax}=\mathbf{b}$ denklemlerine her çözüm $\overline{\mathbf{A}}\mathbf{b}+(\mathbf{I}-\overline{\mathbf{A}}\mathbf{A})\mathbf{c}$ biçimine sahiptir. Burada $\overline{\mathbf{A}}$, $\mathbf{A}$'nın herhangi sabit g-tersidir ve $\mathbf{c}$; keyfi bir vektördür (Sengupta ve Jammalamadaka 2003).

Yardımcı teorem 5.16.6:Eğer $\mathbf{u}$ ve $\mathbf{v}$ sonlu ortalama ve dağılım matrislerine (varyans-kovaryans matrislerine) sahip rasgele vektörler iseler, bu taktirde

a)1 olasılıkla kesin olarak $[\mathbf{u}-\mathbf{E}(\mathbf{u})]\in\mathcal{C}(\mathbf{D}(\mathbf{u}))$,

b) $\mathcal{C}(\mathbf{Cov}(\mathbf{u},\mathbf{v}))\subseteq\mathcal{C}(\mathbf{D}(\mathbf{u}))$dır.

İspat: $\mathbf{Y}=(\mathbf{I}-\mathbf{P}_{\mathbf{D}(\mathbf{u})})[\mathbf{u}-\mathbf{E}(\mathbf{u})]$ olsun. $\mathbf{E}(\mathbf{Y})=\mathbf{0}$ ve $\mathbf{D}(\mathbf{Y})=\mathbf{0}$ olduğunu görmek kolaydır. Bu nedenle $\|\mathbf{Y}\|^2$ rasgele değişkeni

$$\begin{aligned}\mathbf{E}(\|\mathbf{Y}\|^2)&=\mathbf{E}(\mathbf{Y}'\mathbf{Y})=\mathbf{E}(\mathbf{iz}\mathbf{Y}\mathbf{Y}')=\mathbf{iz}\mathbf{E}(\mathbf{Y}\mathbf{Y}')\\&=\mathbf{iz}\ \mathbf{E}\left[\{\mathbf{E}(\mathbf{Y})+(\mathbf{Y}-\mathbf{E}(\mathbf{Y})\}\{\mathbf{E}(\mathbf{Y})+(\mathbf{Y}-\mathbf{E}(\mathbf{Y})\}'\right]\\&=\mathbf{iz}\left[\mathbf{E}(\mathbf{Y})\mathbf{E}(\mathbf{Y}')\right]+\mathbf{iz}\mathbf{D}(\mathbf{Y})=\mathbf{0}\end{aligned}$$

şartını sağlar. Bu, olasılık teorisinin standart bir muhakemesiyle sıfır beklenen değerli herhangi bir negatif olmayan rasgele değişkenin $\mathbf{1}$ olasılıkla (kesin olarak) sıfır olması gerektiği görülür. Bu nedenle $\mathbf{Y}$ vektörü $\mathbf{1}$ olasılıkla sıfır vektör olmalıdır. Bu (a) kısmını ispatlar. (b) kısmı yardımcı teoremin (b) kısmının, $\mathbf{u}$ ve $\mathbf{v}$ nin birleştirilmiş (birleşik) dağılım matrisine uygulamak suretiyle elde edilir.(Yani, $\mathbf{cov}=\begin{bmatrix}\mathbf{D}(\mathbf{u}) & \mathbf{cov}(\mathbf{u},\mathbf{v})\\ \mathbf{cov}(\mathbf{u},\mathbf{v}) & \mathbf{D}(\mathbf{v})\end{bmatrix}$

matrisine uygulayarak, $\mathbf{cov}(\mathbf{u},\mathbf{v})\subseteq\mathbf{D}(\mathbf{u})$olduğu görülür.)

Yardımcı teorem 5.16.7: $\mathbf{u}$ ve $\mathbf{v}$, $\mathbf{E(v)=0}$ olacak şekilde, birinci ve ikinci mertebeden momentlere sahip olan rasgele vektörler olsunlar. Bu taktirde, $\mathbf{u-Bv}$ lineer birleşiminin $\mathbf{v}$ ile ilişkisiz olması için gerek ve yeter şart $\mathbf{1}$ olasılıkla, $\mathbf{Bv=cov(u,v)[D(v)]^{-}v}$ olmasıdır.

İspat: "Gerek" kısmı göstermek kolaydır. "yeter" kısmını ispatlamak için, $\mathbf{cov(u-Bv,v)=0}$ olsun. $\mathbf{BD(v)=cov(u,v)}$, veya $\mathbf{D(v)B'=cov(v,u)}$ olduğu görülür. Yardımcı teorem 5.16.2 nin c) kısmına göre $\mathbf{B'}$ $\mathbf{\overline{V}cov(v,u)-(I-\overline{V}V)G}$ formuna sahip olmalıdır. Burada $\mathbf{V=D(v)}$ dir ve $\mathbf{G}$ herhangi bir matristir. $\mathbf{E(v)=0}$ olduğundan, yardımcı teorem 5.16.9 un a) kısmı $\mathbf{1}$ olasılıkla, $\mathbf{v}\in\mathcal{C}\mathbf{(V)}$ olduğunu ifade eder. Genelliği yitirmeksizin, $\mathbf{\overline{V}}$ 'yi herhangi bir simetrik matris olarak seçebiliriz ve $\mathbf{cov(u,v)\overline{V}}$ 'nin değeri g-tersin seçimine bağlı olmaz. Sonuç olarak, $\mathbf{v'B'=v'\overline{V}cov(u,v)-0}$ yani $\mathbf{Bv=cov(u,v)\overline{V}}$ dır.

Yardımcı teorem 5.16.8: $\mathbf{(y,X\beta,\sigma^2V)}$ modelinde tahmin edilebilir olan $\mathbf{LX\beta}$ *LPF* (lineer parametrik fonksiyonu) vektörünü göz önüne alalım. Bu *LPF* nin EİLYTE si

$\mathbf{LX\beta}$ = $\mathbf{L\left[I-V(I-P_X)\{(I-P_X)V(I-P_X)\}^{-}\right](I-P_X)Y}$ ile verilir. Ayrıca EİLYTE tektir.

İspat: $\mathbf{LY}$ nin, $\mathbf{LX\beta}$ 'nın bir lineer yansız tahmin edicisi, (*LYTE* si) olduğunu ve karşılık gelen EİLYTE nin $\mathbf{(I-P_X)Y}$ ile ilişkisiz olması gerektiğini bu nedenle onun tüm $\mathbf{LSF}$ (lineer sıfır fonksiyon) ler ile ilişkisiz olması gerektiğini belirtelim. Yardımcı teorem (5.3.7) de $\mathbf{u=LY}$ ve $\mathbf{v=(I-P_X)Y}$ koyarak, $\mathbf{u-Bv}$ lineer bileşiminin $\mathbf{v}$ ile ilişkisiz olması için gerek ve yeter şartın yardımcı teoremde verilen ifade olması gerektiğini buluruz. Bu niceliğin ortalaması $\mathbf{E(u)=LX\beta}$ olduğundan, $\mathbf{u=LY}$, $\mathbf{LX\beta}$ nin (tek) EİLYTE si olmalıdır.

Uyarma5.16.2 : $\mathcal{C}\mathbf{((I-P_X)V)}=\mathit{Cov}\mathbf{((I-P_X)Y,Y)}$

ve $\mathcal{C}\mathbf{((I-P_X)V(I-P_X))}=\mathcal{C}\mathbf{(D(I-P_X)Y)}$ olduğuna dikkat edelim. Yardımcı teorem (5.16.6) dan $\mathbf{(I-P_X)Y}$ nin $\mathcal{C}\mathbf{((I-P_X)V(I-P_X))}$ de ihtiva edildiği görülür, ve $\mathcal{C}\mathbf{((I-P_X)V)}$, $\mathcal{C}\mathbf{((I-P_X)V(I-P_X))}$ in bir alt kümesidir. Bu yardımcı teoremde verilen EİLYTE nin ifadesi g-tersin seçimine bağlı değildir (Sengupta ve Jammalamadaka 2003).

Uyarma 5.16.3: Herhangi bir tahmin edilebilir $\mathbf{A\beta}$ *LPF* vektörü $\mathbf{AX^{-}X\beta}$ gibi yazılabildiğinden, onun EİLYTE si

$\mathbf{A\beta = AX^{-}\left[I - V(I - P_X)\{(I - P_X)V(I - P_X)\}^{-}\right](I - P_X)Y}$ dır. Aşağıdaki uyarmaya göre yukarıdaki ifade, $\mathbf{X^{-}}$ nin seçimine bağlı değildir.

Uyarma 5.16.4: $\mathbf{E(\hat{Y}) = X\beta}$ olduğundan, yardımcı teorem (5.16.6) ve (5.16.9) un

a) kısmından dolayı **1** olasılıkla

$$\mathbf{\hat{Y} \in C(X)}$$

$$\mathbf{e} \in C(\mathbf{V(I - P_X)})$$

olduğu görülür. (5.16.6) ve (5.16.9) yardımcı teoremlerinden

$$\mathbf{(X\beta - X\beta) \in C(X) \cap C(V)}$$

olduğu da görülür. Bu sonuç hoş bir yorumdur. EİLYT'nin bir sonucu bizi $\mathbf{Y = X\beta + e}$ vasıtasiyle bilinmeyen $\mathbf{Y = X\beta + \varepsilon}$ ayrışımına yaklaştırır. Bu ayrışımdaki hata yani $\mathbf{(X\beta - X\beta)}$ hatanın parçası olmuş olmakla beraber sistematik kısımda, yani $\mathbf{X\beta}$ da yanlışlıkla konulmuş olan bir vektör olmalı. Böyle bir "karışıklık" lığın vuku bulması için gerek ve yeter şart bu vektör aynı anda sistematik kısmın sütun uzayına yani $C(\mathbf{X})$'e, ve $\boldsymbol{\varepsilon}$ 'nun sütun uzayına yani $C(\mathbf{V})$'ye ait olmalıdır. Eğer $C(\mathbf{X})$ ve $C(\mathbf{V})$ karşılıklı olarak ayrık iseler bu taktirde böyle bir hatanın ortaya çıkması mümkün değildir. Gerçekten, böyle bir durumda **1** olasılıkla $\mathbf{X\beta = X\beta}$ dır. **y** nin çeşitli alt uzaylara ait bileşenlere ayrılmasının bir geometrik yorumu yukarıda incelenmiştir (Sengupta ve Jammalamadaka 2003).

Uyarma 5.16.5: $\mathbf{V = I}$ olduğunda $\mathbf{A\beta}$ için verilen ifade Gauss-Markov Teoremindeki gibi

$\mathbf{AX^{-}P_XY}$ veya $\mathbf{A(X'X)^{-}X'Y}$ ye sadeleşir.

Bu yardımcı teoremde $\mathbf{L = I}$ koymak suretiyle elde edilen, uygun değerlerin vektörü:

$$\hat{\mathbf{Y}} = \left[\mathbf{I} - \mathbf{V}(\mathbf{I}-\mathbf{P_X})\{(\mathbf{I}-\mathbf{P_X})\mathbf{V}(\mathbf{I}-\mathbf{P_X})\}^{-}\right](\mathbf{I}-\mathbf{P_X})\mathbf{Y} \tag{5.16.1}$$

dır. Bu nedenle $\mathbf{A\beta} = \mathbf{AX^{-}\hat{Y}}$ dır. Böylece, $\boldsymbol{\beta}$ 'nın aşağıdaki tahmin edicisini aşağıdaki formülle tanımlayabiliriz.

$$\hat{\boldsymbol{\beta}} = \mathbf{X^{-}\hat{Y}} = \mathbf{X^{-}}\left[\mathbf{I} - \mathbf{V}(\mathbf{I}-\mathbf{P_X})\{(\mathbf{I}-\mathbf{P_X})\mathbf{V}(\mathbf{I}-\mathbf{P_X})\}^{-}\right](\mathbf{I}-\mathbf{P_X})\mathbf{Y} \tag{5.16.2}$$

Burada $\mathbf{X^{-}}$, $\mathbf{X}$ in bir keyfi g-tersidir. Binaenaleyh tahmin edici genelde tek bir şekilde tanımlanmaz. Tek olarak tanımlanabilmesi için gerek ve yeter şart $\mathbf{X}$ 'in tam sütun ranka sahip olmasıdır ki; bu durumda $\hat{\boldsymbol{\beta}}$, $\boldsymbol{\beta}$ 'nın EİLYTE sidir. Eğer $\mathbf{X}$ eksik ranklı olsa bile, $\hat{\boldsymbol{\beta}}$, tahmin edicide bir musluk gibi kullanılabilir. Özellikle, herhangi bir tahmin edilebilir $\mathbf{A\beta}$ lineer parametrik fonksiyonunun (tek bir) EİLYTE si $\mathbf{A\hat{\beta}}$ ile özdeştir.(5.16.1) de verilen $\hat{\mathbf{Y}}$ yı kullanarak, $\mathbf{e}$ hata tahmini vektörü:

$$\mathbf{e} = \mathbf{Y} - \hat{\mathbf{Y}} = \left[\mathbf{V}(\mathbf{I}-\mathbf{P_X})\{(\mathbf{I}-\mathbf{P_X})\mathbf{V}(\mathbf{I}-\mathbf{P_X})\}^{-}\right](\mathbf{I}-\mathbf{P_X})\mathbf{Y} \tag{5.16.3}$$

ile verilir. Olası singüler lineer modelde $\mathbf{X\beta}$ 'nın EİLYTE sinin bir geometrik görünümü yukarıda verilmniştir (Sengupta ve Jammalamadaka 2003).

Dağılımlar

(5.16.1) ve (5.16.3) den,

$$\mathbf{D}(\hat{\mathbf{Y}}) = \sigma^2\left[\mathbf{V} - \mathbf{V}(\mathbf{I}-\mathbf{P_X})\{(\mathbf{I}-\mathbf{P_X})\mathbf{V}(\mathbf{I}-\mathbf{P_X})\}^{-}(\mathbf{I}-\mathbf{P_X})\mathbf{V}\right], \tag{5.16.4}$$

$$\mathbf{D}(e) = \sigma^2 V\left[(\mathbf{I}-\mathbf{P_X})\{(\mathbf{I}-\mathbf{P_X})\mathbf{V}(\mathbf{I}-\mathbf{P_X})\}^{-}(\mathbf{I}-\mathbf{P_X})\mathbf{V}\right], \tag{5.16.5}$$

$$\begin{aligned}\mathbf{D}(\mathbf{A}\hat{\boldsymbol{\beta}}) &= \mathbf{D}(\mathbf{AX^{-}\hat{Y}}) \\ &= \sigma^2\mathbf{AX^{-}}\left[\mathbf{V} - \mathbf{V}(\mathbf{I}-\mathbf{P_X})\{(\mathbf{I}-\mathbf{P_X})\mathbf{V}(\mathbf{I}-\mathbf{P_X})\}^{-}(\mathbf{I}-\mathbf{P_X})\mathbf{V}\right](\mathbf{AX^{-}})'\end{aligned} \tag{5.16.6}$$

oldukları görülür.

Yardımcı teorem 5.16.9: $(\mathbf{Y}, \mathbf{X\beta}, \sigma^2\mathbf{V})$ lineer modeli için

(a) $\mathcal{C}(\mathbf{D}(\hat{\mathbf{Y}})) = \mathcal{C}(\mathbf{X}) \cap \mathcal{C}(\mathbf{V})$ dir.

(b) $\mathcal{C}(\mathbf{D(e)}) = \mathcal{C}(\mathbf{V(I - P_X)})$

İspat: (5.16.3) deki $\mathbf{D(\hat{Y})}$ 'nın ifadesi doğrudan doğruya $\mathcal{C}(\mathbf{D(\hat{Y})}) \subseteq \mathcal{C}(\mathbf{V})$ olduğunu ifade eder. Ayrıca, $(\mathbf{I - P_X})$. $\mathbf{D(\hat{Y})}$ 'nın sıfır olacağı kolayca görülür. Sonuç olarak, $\mathcal{C}(\mathbf{D(\hat{Y})}) \subseteq \mathcal{C}(\mathbf{X})$ ve gerçekten $\mathcal{C}(\mathbf{D(\hat{Y})}) \subseteq \mathcal{C}(\mathbf{V}) \cap \mathcal{C}(\mathbf{X})$ dır. Şimdi $\boldsymbol{l} \in \mathcal{C}(\mathbf{D(\hat{Y})})^{\perp}$ olsun, bu nedenle $\left[\mathbf{V - V(I - P_X)\{(I - P_X)V(I - P_X)\}^{-}(I - P_X)V}\right]\boldsymbol{l} = 0$ dır. $\boldsymbol{l} \in [\mathbf{C(X)} \cap \mathbf{C(V)}]^{\perp}$ olduğunu göstereceğiz. Bunu ispatlamak için $\mathcal{C}(\mathbf{X}) \cap \mathcal{C}(\mathbf{V})$ den bir $\mathbf{m}$ vektörü alırız. $\mathbf{m}$ 'yi $\mathbf{V}\boldsymbol{l}_1$ olarak ve sonrada $\mathbf{X}\boldsymbol{l}_2$ olarak yazarız. Böylece

$$\boldsymbol{l}'\mathbf{m} = \boldsymbol{l}'\mathbf{V}\boldsymbol{l}_1 = \boldsymbol{l}'\left[\mathbf{V(I - P_X)\{(I - P_X)V(I - P_X)\}^{-}(I - P_X)V}\right]V\boldsymbol{l}_1$$

$$= \boldsymbol{l}'\left[\mathbf{V(I - P_X)\{(I - P_X)V(I - P_X)\}^{-}(I - P_X)V}\right]X\boldsymbol{l}_2 = 0$$

dır.

Bu nedenle, $\mathcal{C}(\mathrm{D}(\hat{\mathbf{Y}})) \subseteq [\mathcal{C}(\mathbf{X}) \cap \mathcal{C}(\mathbf{V})]^{\perp}$ dır. Bu (a) kısmını ispatlar.

$\mathbf{D}(e)$ nin (5.23) de verilen ifadesi $\mathcal{C}(\mathbf{D(e)}) \subseteq \mathcal{C}(\mathbf{V(I - P_X)})$ olduğunu belirtir. Ters kapsama, yani $\mathcal{C}(\mathbf{V(I - P_X)}) \subseteq \mathcal{C}(\mathbf{D(e)})$, $\mathbf{D(e)\,(I - P_X)} = \sigma^2\mathbf{V(I - P_X)}$ olması gerçeğinden görülür. Bu (b) kısmını ispatlar.

Şekil 5.23 aynı zamanda $\mathbf{X\beta}$ 'nın EİLYTE sini elde etmek için $\mathbf{Y}$ 'nin "rasgele olmayan kısmının" ayrılmasına gerek olmadığını da açıklar. İzdüşümler, eşit varyanslı lineer model durumunda aynı şekilde çok daha fazla iş görür. $\mathbf{V}$ singüler olduğunda $\mathbf{X\beta}$ yı bundan ayrıca biri deterministik ve biri stokastik parçaya ayırabiliriz. Deterministik parça 0 dan bu izdüşümün tabanına çizilen dikme iken stokastik parça OB nin $\mathcal{C}(\mathbf{V}) \cap \mathcal{C}(\mathbf{X})$ doğrusu üzerindeki izdüşümüdür. Genel lineer modeldeki kısıtlamalar benzer tarzda gözden geçirilebilir. Singüler durumda bir diyagram (şekil) çizimi $\mathcal{C}(\mathbf{X})$ den $\mathcal{C}(\mathbf{X(I - P_{A'})})$ ya ve sonra da $\mathcal{C}(\mathbf{V}) \cap \mathcal{C}(\mathbf{X(I - P_{A'})})$ ya geçişi izler iken boyutların birinin yok olması nedeniyle zor yapılır.

6.SONUÇ VE ÖNERİLER

Bu tezde lineer modellerde kullanılan matris cebiri ve matrislerin genelleştirilmiş tersleri hakkında gereken bilgiler verilerek; lineer modellerde en küçük kareler yönteminin geometrisi açıklayıcı şekillerle birlikte sunuldu.

Bu tezde ele alınan konular temel istatistiksel kavramların bir takımı olup özellikle lineer modellerin geometrisi alanında çalışan araştırmacılara ve bu konuda doktora ve benzeri araştırmalar yapacaklara azda olsa bir katkı sağlayacaktır.

7.KAYNAKLAR

Bloomfield, P., 2000. Fourier Analysis of Time Series: An Introduction. New York: Wiley.

Bring, J., 1996, A geometric Approach to compare Variables in a Regression Model, The American statistician 50, 1, pp. 57-62

Christensen, R.,1996. Plane Answers to Complex Questions: The Theory of Linear Models (2nd ed.). New York: Springer Verlag. pp 25-27.

Davidson, R. ve Mackinon, R., 1993, Estimation and Inference in Econometrics, Oxford University Press, Oxford.

Draper, N.R. ve Smith, H .,1998. Applied Regression Analysis, Canada

Ekni , M., 1999.Lineer Modeller.Ankara

Eubank, R. L. ve R. L. Eubank, 1999. Nonparametric Regression and Spline Smoothing. New York: Marcel Dekker.

Ezekiel, M., 1930. Methods of Correlation Analysis. New York: Wiley.

Fisher, R. A., 1921. On the probable error of a coefficient of correlation deduced from a small sample. Metron 1, 1–32.

Frisch, R. Waugh, F. 1933, Partial time regressions as compared with individual trends, Econometrica, 45, 939-53

Graybill F.A., 1976. Theory and Application of the Linear Model. Duxbury Press, Boston, Massachusetts.

Grimmett, G. ve Stirzaker, D., 2004. Probability and random processes. (3 rd ed.) Oxford University Press. pp. 343-347

Harville, D. A., 1997. Matrix Algebra from a Statistician's Perspective. New York: Springer Verlag. p 170 and pp 28-29

Hsu, H.P., 1997. Probability, random variables and random processes. New York: McGraw-Hill.

Jammalamadaka, S. R. ve D. Sengupta 2003. Linear Models an Integrated Approach. Singapore: World Scientific Publications. p 484 uyarma 11.1.23, p 253, pp 47-48, pp 36-37 proposition 2.4.1.(f), p 257 remark 7.3.10, pp 252-253 proposition 7.3.1. John Wiley ve Sons inc. (Second Edition)

Lovell, M., 1963, Seasonal adjustment of econmic time series, Journal of the American Statistical Association, 58, 993-1010.

Magnus, J.R., 1990. Matrix Differential Calculus with Applications in Statistics and Econometrics. John Wiley & Sons Ltd.

Mahalanobis, P. C., 1964. Professor Ronald Aylmer Fisher. Biomtrics 20, pp 238-250.

Ogden, R. T. 1997. Essential Wavelets for Statistical Applications and Data Analysis. Birhauser.

Rao 1974. Projectors, generalized inverse ve BLUES J. Roy Statistic

Rao, C.R., Toutenburg, H., 1999. Linear Models: Least Squares and Alternatives. Second Edition, Wad Swarth, inc., California.

Rencher, A.C., ve Schaalje, B.G., 2007. Linear Models in Statistics Second Edition, A John Wiley &Sons. İnc., Publication, Hoboken, New Jersey

Seber, G.A.F., 1977. Linear Regression Analysis. John Wiley &Sons, Inc. New York

ÖZGEÇMİŞ

Adı Soyadı: Aysun KÖR

Doğum Yeri: ERBAA/TOKAT

Doğum Tarihi: 01.08.1988

Medeni Hali: Evli

Yabancı Dili: İngilizce

E-mail: zuhalim.33@hotmail.com

İletişim Bilgileri: 05417958945

Öğrenim Durumu:Yüksek Lisans

Derece	Bölüm /Program	Üniversite	Yıl
Lisans	Matematik	Ordu Üniv.	2010
Y. Lisans	Matematik	Ordu Üniv.	2013

İş Deneyimi:3 Yıl

Görev	Görev Yeri	Yıl
Uzman Öğretici	Ordu/Ünye	2010-2012
Uzman Öğretici	Ordu/Merkez	2012-2013