T.C.

ORDU ÜNİVERSİTESİ

FEN BİLİMLERİ ENSTİTÜSÜ

RASGELE VEKTÖRLER, RASGELE MATRİSLER, ŞAPKA MATRİSİ VE MERKEZLEŞTİRİLMİŞ AÇIKLAYICILAR MATRİSİ, ONLARIN LİNEER MODELLERDE KULLANILMASI

BİLGE BAHAR KIRIŞOĞLU

YÜKSEK LİSANS TEZİ

MATEMATİK ANABİLİM DALI

ORDU 2020

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.

BİLGE BAHAR KIRIŞOĞLU

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

ÖZET

RASGELE VEKTÖRLER, RASGELE MATRİSLER, ŞAPKA MATRİSİ VE MERKEZLEŞTİRİLMİŞ AÇIKLAYICILAR MATRİSİ, ONLARIN LİNEER MODELLERDE KULLANILMASI

BİLGE BAHAR KIRIŞOĞLU

ORDU ÜNİVERSİTESİ FEN BİLİMLERİ ENSTİTÜSÜ

MATEMATİK ANABİLİM DALI

YÜKSEK LİSANS TEZİ, 100 SAYFA

TEZ DANIŞMANI: DOÇ. DR. MEHMET KORKMAZ

Bu çalışmada yaygın olarak kullanılan rasgele vektörler ve rasgele matrislerin lineer modellerde kullanılması sunulmuştur. Aynı zamanda şapka matrisi ve merkezileştirilmiş açıklayıcılar matrisi de lineer modellerde kullanılarak hata terimleri bulunmuştur. Burada en küçük kareler analizinde bireysel ve ortak olarak etkili gözlemleri saptamak için şapka matrisi ve açıklayıcılar matrisi kullanılmıştır. En küçük kareler yaklaşımında, herhangi bir duyarlılık (hassasiyet) analizi aslında noktaların nasıl gözlendiği bu nedenle şapka matrisinin elemanları üzerine nasıl yansıtıldığı ile ilgilidir. Regresyon tanılarında en yaygın kullanılan kavramlar olarak etkili gözlemler ve aykırı değerler bu niceliklerin büyüklüğü vasıtasıyla teşhis edilir.Bu tezde ayrıca gözlenen değerlerdeki matematiksel ilişkiler ayrıntılı olarak incelenmektedir ve bu inceleme bir akış şeması vasıtasıyla verilmektedir.

.

Anahtar Kelimeler: Lineer regresyon model, Şapka matrisi, Vektör; Rank, Merkezileştirilmiş açıklayıcılar matrisi, Lineer kombinasyon, Varyans, En küçük kareler analizi.

ABSTRACT

USE OF RANDOM VECTORS, RANDOM MATRICES, HAT MATRIX AND CENTRALIZED EXPLANTORIES MATRIX IN LINEAR MODELS

BİLGE BAHAR KIRISOGLU

ORDU UNIVERSITY INSTITUTE OF NATURAL SCIENCES

MATHEMATİCS

MSC THESİS, 100 PAGES

SUPERVISOR: ASSOC. PROF. DR. MEHMET KORKMAZ

In this study, the use of commonly used random vectors and random matrices in linear models is presented. Error terms were also found by using the hat matrix and the centralized explanatory matrix in linear models. Here, in the least-squares analysis, the hat matrix and explanatory matrix are used to identify effective observations individually and collectively. In the least squares approach, any sensitivity (tenderness) analysis actually concerns how points are observed, and therefore how they are reflected on the elements of the hat matrix. Effective observations and outliers are identified by the magnitude of these quantities as the most commonly used concepts in regression diagnoses. In this thesis the mathematical relationships in these observations are also examined in detail and given in a flow chart.

Keywords: Linear Regression Models, Hat Matrix, Vector, Rank, Centralized explanatory matrix, Linear combination, Variance, Least squares analysis.

TEŞEKKÜR

Tüm çalışmalarım boyunca her zaman bilgi ve deneyimleriyle yolumu açan, desteğini esirgemeyen değerli danışman hocam, sayın Doç. Dr. Mehmet KORKMAZ’a ve Ordu Üniversitesi Fen Edebiyat Fakültesi Matematik Bölümü öğretim üyelerine ve öğretim elemanlarına sonsuz teşekkür ve şükranlarımı sunarım.

Ayrıca tezin hazırlanma aşamasında yardımlarını esirgemeyen sayın Prof. Dr. Cemil YAPAR’a engin sabırlarından dolayı teşekkür ederim.

Öğrenim hayatım boyunca gösterdikleri maddi, manevi destekleri ve fedakarlıkları ile her zaman benim yanımda olan annem, babam, kardeşlerim ve eşim Volkan KIRIŞOĞLU ‘na teşekkürü bir borç bilirim.

İÇİNDEKİLER

ŞEKİL LİSTESİ

ÇİZELGE LİSTESİ

Sayfa

SİMGELER ve KISALTMALAR LİSTESİ

ANOVA	**:** Varyans Analizi
BLUE	**:** En İyi Lineer Yansız Tahmin Edici
PRESS	**:** Tahmin Edilen Ölçüm Hatası Kareler Toplamı
RSS	**:** Hata Kareler Toplamı
TSS	**:** Genel Kareler Toplamı
$\boldsymbol{SS_{reg}}$	**:** Tahmin Edilen Kareler Toplamı
SSE	**:** Tahmin Edilen Hata Kareler Toplamı
SSR	**:** Regresyon Kareler Toplamı
SST	**:** Genel Kareler Toplamı

1. GİRİŞ

1.1 Rasgele Vektörler ve Matrisler

1.1.1 Rasgele Vektör

Bir rasgele vektör elemanları rasgele değişkenler olan bir vektördür.Örneğin, x_1 , ... , x_k herbiri rasgele değişkenler olmak üzere

$$\mathbf{X}_{kx1} = (x_1, x_2, \ldots, x_k)^T \tag{1}$$

bir rasgele vektördür.

1.1.2 Rasgele Matris

Bir rasgele matris elemanları rasgele değişkenler olan bir matristir. Örneğin $x_1, x_2, \ldots, x_{nk}$ ların herbiri rasgele değişkenler olmak üzere $x_{nxk} = (x_{ij})$ bir rasgele matristir.

1.1.3 Beklenen Değer

Bir rasgele matrisin (veya vektörün) beklenen değeri (kitle ortalaması) beklenen değerlerin matrisi veya vektörüdür. $x_{n\,xk}$ için

$$E(\mathbf{X}) = \begin{pmatrix} E_{(x_{11})} & E_{(x_{12})} \cdots & E_{(x_{1k})} \\ \vdots & \vdots & \vdots \\ E_{(x_{n1})} & E_{(x_{n2})} \cdots & E_{(x_{nk})} \end{pmatrix} \tag{2}$$

dır.

- $E(X)$ çoğu kez μ_x veya μ nün ortalama olduğu rasgele matris (vektör) içerikten, açık olduğunda sadece μ ile gösterilecek.
- Birtek x rasgele değişkeni için,

$$E(\mathbf{X}) = \begin{cases} \int\limits_{-\infty}^{\infty} x\, f(x)dx & X\ sürekli\, ise, \\ \sum_{\forall x} x\, f_x(x) & X\ kesikli\, ise, \end{cases} \tag{3}$$

Burada $f_x(x)$ sürekli durumunda X in olasılık yoğunluk fonksiyonudur, $f_x(x)$ kesikli durumunda X' in olasılık fonksiyonudur.

1.2 (Kitle) Varyans-Kovaryans Matrisi

Bir $\mathbf{X}_{k\,x1} = (x_1, x_2, ..., x_k)$ rasgele vektörü için,

$$\begin{pmatrix} \mathrm{var}_{(x_1)} & kov_{(x_1,x_2)} & \cdots & kov_{(x_1,x_k)} \\ kov_{(x_2,x_1)} & \mathrm{var}_{(x_2)} & \cdots & kov_{(x_2,xk)} \\ \vdots & \vdots & & \vdots \\ kov_{(x_k,x_1)} & kov_{(x_k,x_2)} & & \mathrm{var}_{(x_k)} \end{pmatrix} \tag{4}$$

$$
= \begin{pmatrix} \sigma_{11} & \sigma_{12} \cdots & \sigma_{1k} \\ \sigma_{21} & \sigma_{22} \cdots & \sigma_{2k} \\ \vdots & \vdots & \\ \sigma_{k1} & \sigma_{k2} & \sigma_{kk} \end{pmatrix}
$$

matrisine X in varyans – kovaryans matrisi denir ve var(x) veya $\sum_x$ veya başvurulan rasgele vektör açık olduğunda $\sum$ ile gösterilir.

- var(.) fonksiyonunun vektör veya skaler olan bir tek argümanı olduğuna dikkat ediniz.

Bu tez X in var-kov matrisi için kov(**x**) rotasyonunu kullanır. Ancak bir argument var olduğunda var(**.**)' yı ve iki argument olduğunda $kov(.,.)$' yı kullanmayı uygun buluyoruz.

- μ beklenen değerli bir tek x rasgele değişkeni için,

$$\sigma_{ii} = \mathrm{var}(xi) = E\left[(x_i - \mu_i)^2\right] \tag{5}$$

olduğunu hatırlayınız.

- $x_i \; ve \; x_j$ rasgele değişkenleri için

$$\sigma_{ij} = k\,\mathrm{ov}(x_i, x_j) = E\left[(x_i - \mu_i)(x_j - \mu_j)\right] \tag{6}$$

olduğunu hatırlayınız.

- $\sigma_{ij} = \sigma_{ji}$ *olduğundan*, $\mathrm{var}(x)$ *simetriktir.*
- Vektör/matris cebirine göre, $\mathrm{var}(x)$:

$$\mathrm{var}(x) = E[(x - \mu_x)(x - \mu_x)^T] \tag{7}$$

formulüne sahiptir.

- Eğer X deki $x_1,....,x_k$ rasgele değişkenleri karşılıklı olarak bağımsız iseler, bu taktirde $i \neq j$ olduğunda $kov(x_i, x_j) = 0$ ve $\text{var}(x)$ köşegeni boyunca $(\sigma_{11}, \sigma_{22}, ...\sigma_{kk})^T$ ve başka yerlerde sıfırlara sahip köşegen matrisidir.

(Kitle) Kovaryans Matrisi: $\mathbf{X}_{k\,x1} = (x_1,....,x_k)^T$ ve $\mathbf{Y}_{n\,x1} = (y_1,...,y_n)^T$ rasgele vektörleri için

$\sigma_{ij} = kov(x_i, y_j)\,, i = 1,....,k\,, j = 1,....,n$

olsun

$$\begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2n} \\ \vdots & \vdots & & \vdots \\ \sigma_{k1} & \sigma_{k2} & \cdots & \sigma_{kn} \end{pmatrix} = \begin{pmatrix} kov_{(x_1,y_1)} & \ldots & kov_{(x_1,y_n)} \\ \vdots & \ddots & \vdots \\ kov_{(x_k,y_1)} & \cdots & kov_{(x_k,y_n)} \end{pmatrix} \tag{8}$$

matrisine x ve y nin kovaryans matrisi denir ve $kov(x, y)$ veya bazen $\sum_{x,y}$ ile gösterilir.

- $kov(.,.)$ fonksiyonunun, herbiri bir skaler veya bir vektör olabilen iki argument olduğuna dikkat ediniz.
- Vektör matris cebirine göre $kov(x, y)$;

$$kov(x, y) = E[(x - \mu_x)(y - \mu_y)]^T$$

formulüne sahiptir.

- $Var(x) = kov(x, x)$ olduğuna dikkat ediniz.

(Kitle) Korelasyon Matrisi: Bir $x_{k\,x1}$ rasgele değişkeni için, kitle korelasyon matrisi X in elemanları arasındaki korelasyonların matrisidir ve $korr(x)$ ile gösterilir.

$\rho_{ij} = korr(x_i, x_j)$ olmak üzere

$$korr(x)=\begin{pmatrix} 1 & \rho_{12} & \cdots & \rho_{1k} \\ \rho_{21} & 1 & \cdots & \rho_{2k} \\ \vdots & \vdots & & \vdots \\ \rho_{k1} & \rho_{k2} & & 1 \end{pmatrix} \tag{9}$$

dır.

- $x_i \; ve \; x_j$ rasgele değişkenleri için

$$\rho_{ij}=korr(x_i,x_j)=\frac{\sigma_{ij}}{\sqrt{\sigma_{ii}}\sqrt{\sigma_{jj}}} \text{ nin}$$

$x_i \; ve \; x_j$ arasındaki lineer ilişkinin miktarını ölçtüğünü hatırlayınız.

- Herhangi bir x için korr(x) simetriktir.
- Bazen iki argumentli $korr(x_{k\,x1},y_{n\,x1})$ korr fonksiyonunun x ve y nin elemanları

arasındaki korelasyonları $k\,x\,n$ matrisini ifade etmek için kullanacağız.

$$korr(x,y)=\begin{pmatrix} korr_{(x_1,y_1)} & \cdots & korr_{(x_1,y_n)} \\ \vdots & \ddots & \vdots \\ korr_{(x_k,y_1)} & \cdots & korr_{(x_k,y_n)} \end{pmatrix} \tag{10}$$

- $korr(x)=korr(x,x)$ olduğuna dikkat ediniz.
- $x_{k\,x1} \; ve \; y_{n\,x1}$ rasgele vektörleri için

$$\rho_x=korr(x), \quad \sum\nolimits_x=\mathrm{var}(x), \quad \rho_{x,y}=korr(x,y), \quad \sum\nolimits_{x,y}=kov(x,y),$$

$$v_x=k\ddot{o}\c{s}eg(\mathrm{var}(x_1),...,\mathrm{var}(x_k)) \;\; ve \;\; v_y=k\ddot{o}\c{s}eg(\mathrm{var}(y_1),...,\mathrm{var}(y_n)) \text{ olsun.}$$

$\rho_x \; ve \sum_x$ arasındaki ilişki

$$\sum\nolimits_x = v_x^{1/2}\rho_x^{1/2} \tag{11}$$

$$\rho_x=(v_x^{1/2})^{-1}\sum\nolimits_x (v_x^{1/2})^{-1}$$

ve x ve y nin varyans ve korelasyon matrisleri arasındaki ilişki

$$\sum_{x,y} = v_x^{1/2} \rho_{x,y} v_y^{1/2}$$

$$\rho_{x,y} = v_x^{-1/2} \sum_{x,y} v_y^{-1/2}$$

dir. Özellikleri X,Y aynı boyutlu rasgele matrisler olsun A,B ve AxB tanımlı olacak şekilde sabitlerin matrisi olsun.

1) E(X + Y) = E(X) + E(Y) dir.

2) E(A × B) = A E(X)B

özellikle $E(AX) = A\mu x$ dir.

Şimdi $x_{k\,x1}$ $y_{n\,x1}$ rasgele vektörleri olsun ve $c_{k\,x1}\, ve\ d_{n\,x1}$ sabitlerin vektörleri olsun . A,B; Ax**,** By çarpımları ile ilgili çarpılabilir matrisler olsun.

3) $kov(x.y) = kov(y,x)^T$ dur.

4) $kov(x+c, y+d) = \mathrm{cov}(x,y)$ dir.

5) $kov(Ax, By) = A\,kov(x,y)\,B^T$ dir.

$x_1, x_2\ iki\ k\,x1$ rasgele vektör ve $y_1, y_2\ iki\ n\,x1$ rasgele vektör olsun. Bu taktirde ,

6) $kov(x_1+x_2, y_1) = kov(x_1,y_1) + kov(x_2,y_1)\ ve\ kov(x_1, y_1+y_2) = kov(x_1,y_1) + kov(x_1,y_2)$ dır.Birlikte ele alındığında 5. ve 6. özellikleri $kov(gg)$ nın her iki argumente göre lineer olduğunu söyler.(yani, o bilineerdir).

$\mathrm{var}(g)$ nın çeşitli özellikleri $kov(gg)$ un özelliklerinden direkt olarak anlaşılır. Çünkü $\mathrm{var}(x) = kov(x,x)$

7) $\mathrm{var}(x_1+c) = kov(x_1+c, x_1+c) = kov(x_1,x_1) = \mathrm{var}(x_1)$ dir.

8) $\mathrm{var}(Ax) = A\,\mathrm{var}(x)\,A^T$ dir.

9) $\operatorname{var}(x_1+x_2)=kov(x_1+x_2,x_1+x_2)=\operatorname{var}(x_1)+kov(x_1,x_2)+kov(x_2,x_1)+\operatorname{var}(x_2)$

Eğer $x_1\, ve\, x_2$ bağımsız ise bu taktirde $kov(x_1,x_2)=0$ dır. Bu nedenle 9. özelliğin $\operatorname{var}(x_1+x_2)=\operatorname{var}(x_1)+\operatorname{var}(x_2)$ olduğunu gösterir. Bu sonuç n tane bağımsız x_i lerin bir toplamına kolayca genişler , bu nedenle

eğer $x_1,.....,x_n$ *bağımsız ise,*

$$\operatorname{var}(\sum_{i=1}^{n} x_i)=\sum_{i=1}^{n}\operatorname{var}(x_i)$$

dir. Ayrıca, eğer $\operatorname{var}(x_1)=...=\operatorname{var}(x_n)=\sum_x$ ise bu takdirde

$$var(\sum_{i=1}^{n} x_i)=n\sum_x$$

dir. Bir örnek ortalaması vektörünün varyansı için formülü

$$\operatorname{var}(\bar{x})=\operatorname{var}(\frac{1}{n}\sum_{i=1}^{n} x_i)=(\frac{1}{n})^2\operatorname{var}(\sum_{i=1}^{n} x_i)=(\frac{1}{n^2})n\sum_x=\frac{1}{n}\sum_x$$

olduğu kolayca anlaşılır.

- Bunun bir değişkenli durumda bilinen formüle genelleştiğine dikkat ediniz.

Lineer modellerde, her hangi bir x rasgele vektörü için $x^T Ax$ kuadratik formları ve A simetrik matrisi sık sık ortaya çıkar ve böyle niceliklerin beklenen değerinin nasıl alınacağı hakkında genel bir sonuca sahip olmak faydalıdır.

- Simetrik olmayan A için $x^T Ax$ yine bir kuadratik formdur. Çünkü simetrik $B=\frac{1}{2}(A+A^T)$ için $x^T Ax$ kuadratik formunu $x^T Bx$ olarak yapmak mümkündür. Yani $Q(x)=x^T Ax$:

$$Q(x)=x^T A_x=\frac{1}{2}(x^T Ax+\underbrace{x^T Ax}_{=(x^T A^T x)})=\frac{1}{2}(x^T Ax+x^T A^T x) \tag{13}$$

$$= x^T \left\{ \underbrace{\frac{1}{2}(A + A^T)}_{=B \text{ simetrik}} \right\} x$$

olarak yazılabilir. Kuadratik formların lineer modellerde yaygın ve önemli olduğunu herhangi bir kuadratik formun ağırlıklı kareler toplamı olarak yazılabildiğini fark ettik.

A ; $\sum_{i=1}^{n} \lambda_i u_i u_i^T$ tayfsal ayrışımlı bir $n x m$ simetrik matris olsun. Bu taktirde

$$Q(x) = x^T Ax = \sum_{i=1}^{n} \lambda_i x^T u_i u_i^T x = \sum_{i=1}^{n} \lambda_i (\underbrace{u_i x}_{\equiv w_i})(u_i^T x) = \sum_{i=1}^{n} \lambda_i w_i^2 \quad . \tag{14}$$

dır. Bir $x^T Ax$ kuadratik formunun beklenen değeri bir $Q(x, y) = (x_{kx1})^T A_{kxn} y_{nx1}$ bilinen formunun beklenen değeriyle ilgili daha güzel bir sonuçtan hemen görülür. (Hall 2000)

Teorem 1.2.1 $E(x) = \mu_x \; ve \; E(y) = \mu_y \, , kov(x, y) = \sum_{x,y} = (\sigma_{ij}) \; ve \; A = (a_{ij})$ olsun.

Bu taktirde

$$E(x^T Ay) = \sum_i \sum_j a_{ij}\sigma_{ij} + \mu_x^T A\mu_y = iz(A\sum_{x,y}^T) + \mu_x^T A\mu_y \tag{15}$$

dir.

İspat: Bilineer formu toplama rotasyonuna göre yazarak,

$x^T Ay = \sum_i \sum_j a_{ij} x_i y_i$ elde ederiz. Ayrıca

$$E(x_i y_i) = kov(x_i, y_i) + \mu_{x,i}\mu_{y,j} = \sigma_{ij} + \mu_{x,i}\mu_{y,j}$$

dır. Bu nedenle

$$E(x^T A_y) = \sum_{i=1}^{k} \underbrace{\sum_{j=1}^{n} a_{ij}\sigma_{ij}}_{=(A\sum_{x,y}^T nın\,(i,i) terimi} + \sum_{i=1}^{k}\sum_{j=1}^{n} a_{ij}\mu_{x,i}\mu_{y,j}$$

$$= iz(A\textstyle\sum_{x,y}^{T}) + \mu_x^{\,T} A \mu_y$$

dır. Böylece ispat tamamlanır. $y = x$ koyarak aşağıdaki teoremi elde ederiz.

Teorem 1.2.2 (Bir kuadratik formun beklenen değeri)

$Q(x) = x^T Ax, \operatorname{var}(x) = \sum \; , E(x) = \mu \, ve \, A = (a_{ij})$ olsun .Bu takdirde

$$E\{Q(x)\} = \sum_i \sum_j a_{ij} kov(x_i y_j) + \mu^T A \mu = iz(A\textstyle\sum) + Q(\mu) \tag{16}$$

dır.

Örnek 1.2.1 $x_1,....,x_n$ her biri μ ortalamalı ve σ^2 varyanslı bağımsız rasgele değişkenler olsun. Bu taktirde

$$x = (x_1,...,x_n)^T, E(x) = \mu \, j_n, \operatorname{var}(x) = \sigma^2 I_n$$

dır.

$$V = 1\,(j_n)\, ve \, P_v = (1/n) J_{n,n}$$

olmak üzere

$$Q(x) = \sum_{i=1}^{n} (x_i - \bar{x})^2 = \left\| P_v \perp x \right\|^2 = x^T (I - P_v) x \tag{17}$$

Kuadratik formunu göz önüne alınız . $E\{Q(x)\}$ elde etmek için , kuadratik formdaki matrisin $A = I_n - P_v$ olduğuna ve $\sum_x \sigma^2 I_n$ olduğuna dikkat ederiz.

Bu taktirde $A\sum_x = (I_n - P_v) \; ve \; iz(A\sum) = \sigma^2 (n-1)$ dır. (Bir izdüşüm matrisinin izi onun üzerine izdüşürdüğü uzayın boyutuna eşittir). Bu nedenle

$$E\{Q(x)\} = \sigma^2 (n-1) + \underbrace{Q(\mu j_n)}_{=0} \tag{18}$$

dır. Bunun hemen bir sonucu, örneklem varyansının yani

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

yansızlığıdır. Bu sonucu elde etmenin bir başka yöntemi

$$y = (x_1 - \bar{x}, x_2 - \bar{x}, \ldots.., x_n - \bar{x})^T$$

yi tanımlamak ve yukarıdaki teoremi

$$Q(x) = \sum_i (x_i - \bar{x})^2 = y^T I_n y$$

uygulamaktır.

$y = P_v \perp x$ olduğundan, $y; P_v \perp (\sigma^2 I_n)$ $P_v \perp = \sigma^2 P_v \perp$ varyansına sahip ($P_v \perp$ simetrik ve idempotent olduğundan) ve 0 ortalamaya sahiptir. Bu nedenle,

$$E\{Q(x)\} = iz\{I_n(\sigma^2 P_v \perp)\} = iz(\sigma^2(I_n - P_v)\} = \sigma^2(n-1)$$

dır . Yani önceki sonucun aynısıdır.

1.3 Çok Değişkenli Normal Dağılım

Bir $\mathbf{Y}_{nx1}$ rasgele değişkenine bir çok değişkenli normal dağılıma sahiptir denecektir, eğer herhangi bir p için z; N(0,1) bağımsız değişkenlerin bir vektörü A sabitlerin bir matrisi ve μ sabitlerin bir vektörü olmak üzere, y:

$$A_{nxp} z_{px1} + \mu_{nx1} \equiv x$$

ile aynı dağılımına sahip ise,

- Yukarıda z'den x'ye geçerken kullanılan dönüşüm tipine bir afin dönüşüm (Dik ve eğik koordinatlar arasındaki bir dönüşüm bağıntısı) denir. x'in biçimini ve z'nin elemanlarının bağımsız standart normal değişkenler olduğu gerçeğini kullanarak x'in bu nedenle y'nin dağılım fonksiyonunu belirleyebiliriz.

- Bu sadece n=p ve $rank(A) = p$ olduğu durumda mümkün olabilir. Dikkatimizi bu duruma vermeliyiz.

$g(z) = Az + \mu$ ‘yü z’den x’ e dönüşüm olarak tanımlayınız.

Bir $A\ p x p$ tam ranklı matrisi için, $g(z); R^p\ 'den\ R^p$ ye bir 1-1 fonksiyondur, bu nedenle x’in yoğunluğu için aşağıdaki değişken değişimini kullanabiliriz.

$$f_x(x) = f_z\{g^{-1}(x)\}abs(\left|\frac{\partial_g{}^{-1}(x)}{\partial x^T}\right|) = f_z\{A^{-1}(x-\mu\}abs(\left|A^{-1}\right|) \tag{19}$$

(Burada $abs(g)$ mutlak değeri ve $|g|$ determinantı gösterir).

z’nin elemanları bağımsız standart normaller olduğundan z’nin yoğunluğu

$$f_z(z) = \prod_{i=1}^{p} \frac{1}{\sqrt{2\pi}} e^{-z_i^2/2} = (2\pi)^{-p/2} \exp(-\frac{1}{2} z^T z) \tag{20}$$

dır. Değişken dönüşümü formülünde yerine koyarak

$$f_x(x) = (2\pi)^{-p/2} \underbrace{abs|A|}_{=|A|}{}^{-1} \exp\{-\frac{1}{2}(x-\mu)(AA^T)^{-1}(x-\mu)\} \tag{21}$$

elde ederiz.

$$\sum = \mathrm{var}(x) \quad in \ \mathrm{var}(Az+\mu) = \mathrm{var}(Az) = AIA^T = AA^T \tag{22}$$

eşit olduğuna bu nedenle

$$\left|\sum\right| = \left|AA^T\right| = |A|^2 \Rightarrow |A| = \left|\sum\right|^{1/2}$$

olduğuna dikkat ediniz. Ayrıca $E(x) = E(Az+\mu) = \mu$ dür.

Bu nedenle μ ortalamalı ve $\sum$ pozitif tanımlı var-kovaryans matrisli p boyutlu bir çok değişkenli normal rasgele vektör:

Her x için,

$$f_x(x) = (2\pi)^{-p/2} \left|\sum\right|^{-1/2} \exp\{-\frac{1}{2}(x-\mu)^T (\sum)^{-1}(x-\mu)\} \tag{23}$$

yoğunluğuna sahiptir.

- A nın $rank(A) \neq n$ $ranklı\, n\, x\, p$ olduğu durumda, x yine çok değişkenli normaldir, fakat onun yoğunluğu mevcut değildir. Böyle durumlar nadiren ortaya çıkan ve bu çalışmada göz önüne almayacağımız pozitif tanımlı olmayan var-kov matrisli çok değişkenli normal dağılımlara karşılık gelir. Böyle dağılımlar olasılık yoğunluk fonksiyonlarına sahip olmaz ancak daima mevcut olan karakteristik fonksiyonu kullanarak tanımlanabilir.

Olasılık yoğunluk fonksiyonunun bir rasgele değişkenin (veya vektörün) dağılımını karakterize etmek (tam olarak tanımlamak) için doğru bir yöntem olduğunu hatırlayınız. Bu amaç için kullanılabilen başka bir fonksiyon moment çıkaran fonksiyonudur.

Bir x rasgele vektörünün moment çıkaran fonksiyonu $m_x(t) = E(e^{t^T x})$ dır. Bu nedenle $x = Az + \mu$: $N_n(\mu, AA^T = \sum)$ için x'in moment çıkaran fonksiyonu bu

$$m_x(t) = E[\exp\{t^T(Az+\mu)\}] = e^{t^T\mu} E(e^{t^T Az}) = e^{t^T\mu} m_z(A^T t) \tag{24}$$

dır. Bir z_i standart normal rasgele değişkenin moment çıkaran fonksiyonu:

$$m_{z_i} = (e)^{u_i^2/2}$$

$$m_z(u) = \prod_{i=1}^{p} \exp(u_i^2/2) = e^{u^T u/2}$$

dır. (24) denkleminde yerine koyarak.

$$m_x(t) = e^{t^T\mu} \exp\{\frac{1}{2}(A^T t)^T (A^T t)\} = e^{t^T\mu} \exp(\frac{1}{2} t^T \sum t) \tag{25}$$

elde ederiz. Bu nedenle, eğer $m_x(t)$ x in dağılımını tam olarak karakterize ederse ve $m_x(t)$ sadece x'in ortalamasına ve varyansına, yani μ ve $\sum$ ya bağlı ise, bu taktirde bu bir çok değişkenli normal dağılımın bu iki parametre ile tam olarak belirlendiğini söyler.

Yani $x_1 : N_n(\mu_1, \sum_1)$ *ve* $x_2 : N_n(\mu_2 \sum_2)$ için x_1 ve x_2 nin aynı dağılıma sahip olması için gerek ve yeter şart $\mu_1 = \mu_2$ ve $\sum_1 = \sum_2$ olmasıdır.

Teorem 1.3.1 μ *ve* R^n bir elemanı ve $\sum$ *bir* $n x n$ simetrik pozitif tanımlı matris olsun. Bu taktirde μ ortalamalı ve $\sum$ var-kov matrisli bir çok değişkenli normal dağılım mevcuttur.

İspat: $\sum$ simetrik ve pozitif tanımlı olduğundan $\sum = BB^T$ olacak şekilde bir B mevcuttur. (örneğin Cholesky ayrışımı) z bağımsız standart normallerin bir $n x 1$ vektörü olsun. Bu taktirde $x = Bz + \mu : N_n(\mu, \sum)$ dır. Böylece ispat tamamlanır.

- Bu sonuç bağımsız standart normallerin bir z vektörünü üreterek ve sonrada z'yi B alt-üçgen Cholesky çarpanıyla önden çarparak ve bundan sonra μ ortalama vektörünü ekleyerek, μ ortalamalı ve $\sum$var$-kov$ matrisi ile bir çok değişkenli normal vektörü daima üretebileceğimizi ortaya koyar.

Teorem 1.3.2 $\sum$ pozitif tanımlı olmak üzere $x : N_n(\mu, \sum)$ olsun. Sabitleri içeren c ve d için $y_{rx1} = c_{rxn}x + d$ olsun.Bu taktirde;

$$y : N_r(C\mu + d, C\sum C^T) \tag{26}$$

dır.

İspat: Tanımdan $AA^T = \sum$ ve $z : N_p(0, I_p)$ olacak şekilde herhangi A için $x = Az + \mu$

dür. Bu taktirde

$$y = Cx + d = C(Az + \mu) + d = CAz + C\mu + d$$

$$= (CA)z + (C\mu + d) \tag{27}$$

dır. Bu nedenle tanıma göre, y ; $C\mu + d$ ortalamalı ve $(CA)(CA)^T = C\sum C^T$ var-kov matrisli bir çok değişkenli normal dağılımına sahiptir. Bu teoremin basit sonuçları eğer $x : N_n(\mu, \sum)$ ise bu taktirde

i. x'in herhangi bir alt vektöründe μ 'nün karşılık gelen alt vektörüyle ve $\sum nın$ karşılık gelen alt matrisiyle verilen ortalama ve varyansa sahip bir çok değişkenli normal dağılıma sahip olduğu ve

ii. sabitlerin bir vektörü a için herhangi bir $a^T x$ lineer kombinasyonunu

$a^T x : N(a^T \mu, a^T \sum a)$ olduğudur.

Teorem 1.3.3 y_{nx1} bir çok değişkenli normal dağılıma sahip olsun ve y 'yi

$$y = \begin{pmatrix} y_{1px1} \\ \text{L} \\ y_{2(n-p)x1} \end{pmatrix}. \tag{28}$$

olarak parçalayınız.

Bu taktirde y_1 *ve* y_2 nin bağımsız olması için gerek ve yeter şart $kov(y_1, y_2) = 0$ olmasıdır.

İspat: İlki bağımsızlık 0 kovaryansı belirtir. y_1, y_2 *nin* μ_1, u_2 ortalamaları ile bağımsız olduklarını farz ediniz. Bu taktirde

$$kov(y_1, y_2) = E\{(y_1 - \mu_1)(y_2 - \mu_2)^T\} = E\{(y_1 - \mu_1)\}E\{(y_2 - \mu_2)^T\} = 0(0^T) \tag{29}$$

İkincisi 0 kovaryans ve normallik bağımsızlığı gösterir. Bunu yapmak için iki rasgele vektörün bağımsız olmaları için gerek ve yeter şartın onların ortak moment çıkaran fonksiyonlarının çarpımı olması gerektiği gerçeğini kullanırız. $kov(y_1, y_2) = 0$ olduğunu farz ediniz.

t_{nx1}, t_1 $p\,x1$ olmak üzere $t = (t_1^T, t_2^T)^T$ olarak parçalanmış olsun. Bu taktirde

$$\sum = \mathrm{var}(y) = \begin{pmatrix} \sum_{11} & 0 \\ 0 & \sum_{22} \end{pmatrix} = \begin{pmatrix} \mathrm{var}_{(y_1)} & 0 \\ 0 & \mathrm{var}_{(y_2)} \end{pmatrix} \tag{30}$$

olmak üzere, y;

$$m_y(t) = \exp\ (\underbrace{t^T\mu}_{=t_1^T\mu_1+t_2^T\mu_2})\ \exp(\frac{1}{2}t^T\sum t) \tag{31}$$

moment çıkaran fonksiyonuna sahiptir. $\sum$ nın biçimden dolayı

$$t^T\sum t = t_1^T\sum_{11}t_1 + t_2^T\sum_{22}t_2$$

dır . Bu nedenle

$$m_y(t) = \exp(t_1^T\mu_1 + \frac{1}{2}t_1^T\sum_{11}t_1 + t_2^T\mu_2 + \frac{1}{2}t_2^T\sum_{22}t_2) \tag{32}$$

$$= \exp(t_1^T\mu_1 + \frac{1}{2}t_1^T\sum_{11}t_1)\exp(t_2^T\mu_2 + \frac{1}{2}t_2^T\sum_{22}t_2) = m_{y_1}(t_1)m_{y_2}(t_2) \tag{33}$$

dir.

Lemma 1.3.1 $\sum_{21} = \sum_{12}{}^{T}$ olmak üzere

$$y = \begin{pmatrix} y_{1px1} \\ \mathrm{L} \\ y_{2(n-p)x1} \end{pmatrix}, \ \mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \ \begin{pmatrix} \sum_{11} & \sum_{12} \\ \sum_{21} & \sum_{22} \end{pmatrix} \tag{34}$$

parçalanmasıyla

$y : N_n(\mu, \sum)$ olsun $y_{2|1} = y_2 - \sum_{21}\sum_{11}^{-1}y_1$

olsun. Bu taktirde

$$y_1 : N_p(\mu_1, \sum_{11}) \ , \ y_{2|1} : N_{n-p}(\mu_{2|1}, \sum_{22|1})$$

olmak üzere y_1 ve $y_{2|1}$ bağımsızdır. Burada

$$\mu_{2|1} = \mu_2 - \sum_{21} \sum_{11}^{-1} \mu_1 \quad \text{ve} \quad \sum_{22|1} = \sum_{22} - \sum_{21} \sum_{11}^{-1} \sum_{12}$$

dir.

İspat: $C_1 = (I, 0)$ olmak üzere $y_1 = C_{1y}$ yazabiliriz ve $C_2 = (-\sum_{21} \sum_{11}^{-1}, I)$ olmak üzere

$y_{2|1} = C_{2y}$ yazabiliriz. y_1 *ve* $y_{2|1}$ normaldir. Onların ortalama ve varyans-kovaryansları y_1 için

$$C_1\mu = \mu \quad \text{ve} \quad C_1 \sum C_1^T = \sum_{11}$$

ve

$y_{2|1}$ için

$$C_2\mu = \mu_{2|1} \quad \text{ve} \quad C_2 \sum C_2^T = \sum_{22|1} \tag{35}$$

dır.Bağımsızlık bu iki rasgele vektörün

$kov(y_1, y_{2|1}) = kov(C_{1y}, C_{2y}) = C_1 \sum C_2^T = 0$ kovaryans matrisine sahip olması gerçeğinden anlaşılır.

Teorem 1.3.4 Önceki teoremdeki gibi tanımlanan y için y_1 verildiğinde y_2 nin şartlı dağılımı

$$y_2 | y_1 : N_{n-p}(\mu_2 + \sum_{21} \sum_{11}^{-1} (y_1 - \mu_1), \sum_{22|1}) \tag{36}$$

dır.

İspat: $y_{2|1}$; y_1 den bağımsız olduğundan y_1 ‘in verilen bir a değeri için onun şartlı dağılımı, onun marjinal dağılımı, yani $y_{2|1} : N_{n-p}(\mu_{2|1}, \sum_{22|1})$ ile aynıdır. $y_2 = y_{2|1} + \sum_{22} \sum_{11}^{-1} y_1$ olduğuna dikkat ediniz.

y_1 ‘in değeri üzerindeki şart yani ; $\sum_{21} \sum_{11}^{-1} y_1$ sabittir, bu nedenle y_2 nin şartlı dağılımı $y_{2|1}$ nin ki artı bir sabit veya

$\mu_{2|1} + \sum_{21} \sum_{11}^{-1} y_1 = \mu_2 - \sum_{21} \sum_{11}^{-1} \mu_1 + \sum_{21} \sum_{11}^{-1} y_1 = \mu_2 + \sum_{21} \sum_{11}^{-1} (y_1 - \mu_1)$ ortalamalı ve $\sum_{22|1}$ var-kovaryansı (n-p) değişkenli normaldir. Böylece ispat tamamlanmıştır.

1.4 Kısmi ve Çoklu Korelasyon

Örnek-Boy ve Okuma Yeteneği

Bir eğitim psikoloğun standartlaşan bir test üzerindeki puanlara dayanarak çocukların y_1 boyu ve y_2 okuma yeteneği arasındaki ilişkiyi incelediğini varsayınız. O; 3. , 4. ve 5. Sınıflardaki 200 çocuk için y_1 ve y_2 ölçüldü ve bu değişkenler arasındaki örnekli korelasyonu 0,56 bulundu.

Boy ve okuma yeteneği arasında bir lineer ilişki vardır.

Pekala, evet, ancak sadece bir veya daha fazla “karışıklılığa” neden olan değişkeni yok saydığımızda boyun okuma yeteneği üzerinde muhtemelen direkt etkisi yoktur.Bunun yerine daha fazla eğitim sürelerine sahip daha yaşlı çocuklar daha iyi okuma ve daha uzun boylu olma eğilimi gösterir. Hem y_1 ve hem de y_2 üzerinde yaşın etkisi sadece y_1 ve y_2 arasındaki basit korelasyonu (ilişkiyi) inceleyerek yok sayılmıştır.

Kısmi korelasyon katsayısı bir veya daha fazla başka değişkenin lineer etkisiyle iki değişken arasındaki lineer ilişkinin bir ölçüsüdür.

1.4.1 Kısmi Korelasyon

$v : N_{p+q}(\mu, \sum)$ olduğunu farz ediniz. v , μ ve $\sum$

$$v=\begin{pmatrix} x \\ y \end{pmatrix}, \mu=\begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix}, \sum=\begin{pmatrix} \sum_{xx} & \sum_{xy} \\ \sum_{yx} & \sum_{yy} \end{pmatrix}$$

olarak parçalanmış olsun. Burada $x=(v_1,...,v_p)^T$ $p x1$ dır ve $y=(v_{p+1},....,v_{p+q})^T$ $q x1$ dır. x verildiğinde y nin şartlı var-kov matrisinin

$\mathrm{var}(y|x)=\sum_{yy}-\sum_{yx}\sum_{xx}^{-1}\sum_{xy}\equiv\sum_{y|x}$ olduğunu hatırlayınız. $\sigma_{ij|1,...,p}$; $\sum_{y|x}$ in $(i,j)^{th}$ elemanını göstersin. Bu taktirde $x=c$ verildiğinde y_i ve y_j nin kısmi korelasyon katsayısı

$$\rho_{ij|1,...,p}=\frac{\sigma_{ij|1,...,p}}{[\sigma_{ii|1,...,p}\sigma_{jj|1,...p}]^{1/2}} \tag{37}$$

ile tanımlanır. (Paydanın sıfırdan farklı olması şartıyla, sıfır olması durumunda, $\rho_{ij|1,...,p}$ tanımlı değildir).

- Alışıla gelen korelasyon katsayısı gibi, kısmi korelasyon

  $-1\leq\rho_{ij|1,...,p}\leq 1$ eşitsizliğini sağlar.

- Yorum: $\rho_{ij|1,...,p}$ kısmi korelasyonu y_i ve x arasındaki ve y_j ve x arasındaki lineer ilişkiye açıklama getirdikten (veya, kaldırdıktan) sonra y_i ve y_j arasındaki lineer ilişkiyi ölçer.

Örneğin, eğer v_1 = yaş , v_2 =boy v_3 =okuma yeteneği $x=v_1$ ve $y=(v_2,v_3)^T$ ise, bu taktirde $\rho_{23|1}=$ bu değişkenlerin her biri üzerinde yaşın etkilerini uzaklaştırdıktan sonra boy ve okuma yeteneği arasındaki korelasyon olur. $\rho_{23|1}\approx 0$ olmasını beklemeliyiz. Kısmi korelasyon diğer bir kaç değişkenin etkilerini dışladıktan sonra iki değişken arasındaki lineer ilişkiyi ölçer. Kısmi korelasyon katsayısı bir değişken ve diğer birkaç değişkene sahip bir grup arasındaki lineer ilişkiyi ölçer.

1.4.2 Çoklu Korelasyon

$v: \ N_{p+1}(\mu,\sum)$ olduğunu farz ediniz ve v , μ , $\sum$

$$v=\begin{pmatrix} x \\ y \end{pmatrix}, \mu=\begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix}, \sum=\begin{pmatrix} \sum_{xx} & \sigma_y \\ \sigma_{yx} & \sigma_{yy} \end{pmatrix}$$

olarak parçalanmış olsun.

Burada $x=(v_1,.....,v_p)^T \ \ p\,x1$ dır ve $y=v_{p+1}$ bir skaler rasgele değişkendir.

x verildiğinde y'nin şartlı ortalamasının

$\mu_y+\sigma_{yx}\sum_{xx}^{-1}(x-\mu_x)\equiv\mu_{y|x}$ olduğunu hatırlayınız. Bu taktirde y ve x arasındaki çoklu korelasyon katsayısının karesi

$$\rho_{y,x}^2=\frac{k\,\mathrm{ov}(\mu_{y|x},y)}{[\mathrm{var}(\mu_{y|x})\,\mathrm{var}(y)]^{1/2}}$$

olarak tanımlanır.

- Sayısal olarak basit bir formül

$$\rho_{y,x}^2=\{\frac{\sigma_{yx}\sum_{xx}^{-1}\sigma_{xy}}{\sigma_{yy}}\}^{1/2}$$

ile verilir.

- Bu nicelik sayımız için rotasyonumuz $\rho_{y,x}$ den ziyade $\rho_{y,x}^2$ dır, çünkü o bir korelasyon katsayısının karesi gibi davranır. O sıfır ile bir arasındadır yani,

$0\leq\rho_{y,x}^2\leq 1$ dır.

Ve lineer ilişkinin gücünü ölçer, ancak doğrultu ve yönünü ölçmez.

- Örneklerin çoklu korelasyon katsayısının karesine determinasyon katsayısı denir ve genellikle R^2 ile gösterilir.

- Lineer modellerde kısmi ve çoklu korelasyon katsayılarının örnek versiyonları konusuna daha sonra değineceğiz.

1.5 Kuadratik Formların Dağılımları

x^2, F ve t Dağılımları

- Bu dağılımların her üçü normal olarak dağılırlar, rasgele değişkenlerin belirli fonksiyonlarının dağılımları olarak ortaya çıkar.
- Onların merkezi (alışıla gelen) versiyonları muhtemelen alışıla gelen sıfır hipotezleri altından ve güven aralıkları için bir temel olarak normal-teori istatistiklerinin dağılımları olarak bilinir.
- Bu dağılımların merkezi olmayan versiyonları (uyarlamaları) alışılagelen sıfır hipotezlerine karşıtlar(karşıt hipotezleri) altında normal-teori test istatistiklerinin dağılımları olarak ortaya çıkar. Bu nedenle, onlar örneğin testlerin gücünü belirlemede önemlidir.

1.5.1 Ki Kare Dağılımı

$x_1,.....,x_n$; $\mu_1,.....,\mu_n$ ortalamalı ve 1 ortak varyanslı bağımsız normal rasgele değişkenler olsun. Bu taktirde $x=(x_1,.....,x_n)^T$ olmak üzere

$y = x_1^2 + ... + x_n^2 = x^T x = \|x\|^2$ ye n serbestlik dereceli ve $\lambda = \frac{1}{2}\sum_{i=1}^{n} \mu_i^2$ merkezi olmama parametreli bir merkezi olmayan ki-kare dağılımına sahiptir denir. Bunu $y:\ x^2(n,\lambda)$ olarak gösteririz.

- T ve F dağılımlarında olduğu gibi λ sıfıra eşit olduğunda, merkezi ki-kare dağılımı
- n serbestlik dereceli merkezi X^2 genellikle

  $X^2(n) \equiv X^2(n,0)$ olarak gösterilecek.

Özellikle $z_1,...,z_n :\ N(0,1), z_1^2 + ... + z_n^2 :\ X^2(n)$

- Merkezi olmayan $X^2(n,\lambda)$ dağılımı sadece onun n ve λ parametrelerine bağlıdır.

Bir $Y: \; X^2(n)$ n serbestlik dereceli bir merkezli X^2 değişkeni

$$f_Y(y;n) = \frac{y^{n/2-1}e^{-y/2}}{\Gamma(n/2)2^{n/2}} \;, \; y > 0$$

için olasılık yoğunluk fonksiyonuna sahiptir.

- Bu $n/2$ kuvvet(güç) parametreli ve $1/2$ ölçek parametreli bir gamma yoğunluğunun bir ögeli durumudur.

Merkezi olmayan X^2 yoğunluğu merkezi X^2 lerin bir poison karmasıdır. Eğer $Z: \; X^2(n,\lambda)$ ise, bu taktirde Z

$f_z(z;n,\lambda) = \sum_{k=0}^{\infty} p(k;\lambda)\, f_Y(z;n+2k)$, $z > 0$ için olasılık yoğunluk fonksiyonuna sahiptir.

- Yani merkezi olmayan X^2 poison ağırlıklı merkezi X^2 lerin bir ağırlıklı toplamıdır.

Teorem 1.5.1.1 $Y: \; X^2(n,\lambda)$ olsun.Bu taktirde

$i. E(Y) = n + 2\lambda$

$ii. \operatorname{var}(Y) = 2n + 8\lambda$ ve

iii. y'nin moment olarak fonksiyonu

$$m_Y(t) = \frac{\exp[-\lambda\{1-1/(1-2t)\}]}{(1-2t)^{n/2}}$$

dır.

İspat : (i) ve (ii) nin ispatı kolay olduğundan ispatın yapılmasına gerek duyulmamıştır. (iii) nin ispatı, $x; \mu$ ortalamalı ve 1 sabit varyanslı bağımsız normal değişkenlerin bir $n\,x1$ vektörü olmak üzere, sadece

$$m_Y(t) = E(e^{tY}) = E\{e^{t(x^T x)}\} = \int_{-\infty}^{\infty} \mathrm{L} \int_{-\infty}^{\infty} e^{t(x^T x)} \underbrace{f_x(x)}_{N(\mu,1)} dx_1 ... dx_n$$

bir $N(\mu,1)$ yoğunluğunu değerlendirmek için beklenen değerin tanımını kullanmayı içerir. Ayrıntılar için bkz: Graybill (1976, s.126)

X^2 dağılımı bağımsız X^2 lerin toplamında

X^2 olduğu kullanışlı özelliğe sahiptir.

Teorem 1.5.1.2 Eğer $v_1,....,v_k$ sırasıyla $X^2(n_i,\lambda_i)$, $i=1,....,k$ olarak dağılan bağımsız rasgele değişkenler ise, bu taktirde

$$\sum_{i=1}^{k} v_i : X^2(\sum_{i=1}^{k} n_i, \sum_{i=1}^{k} \lambda_i)$$

dır.

İspat: Tanımdan kolayca anlaşılır.

1.6 Bir Kuadratik Fonksiyonun Dağılımı

Bir merkezi ki-karenin tanımından, eğer $y : N_n(\mu, I_n)$ ise, bu taktirde $\|y-\mu\|^2 = (y-\mu)^T(y-\mu) : X^2(n)$ olduğu hemen görülür. $\sum$ pozitif tanımlı olmak üzere $y : N_n(\mu,\sum)$ için, $\sum^{-1/2} = (\sum^{1/2})^{-1}$ *ve* $\sum^{1/2}$; $\sum$ nın bir tek simetrik karekökünün olmak üzere,

$$\begin{aligned} &= (y-\mu)^T \textstyle\sum^{-1}(y-\mu) = (y-\mu)^T \textstyle\sum^{-1/2}\textstyle\sum^{-1/2}(y-\mu) \\ &= \{\textstyle\sum^{-1/2}(y-\mu)\}^T \underbrace{\{\textstyle\sum^{-1/2}(y-\mu)\}}_{\equiv z} \\ &= z^T z \end{aligned}$$

olduğunu dikkate alarak, bunu

$(y-\mu)^T \sum^{-1}(y-\mu) : X^2(n)$ ye genişletebiliriz. $z = \sum^{-1/2}(y-\mu) : N_n(0, I_n)$ olduğundan

$$z^T z = (y-\mu)^T \textstyle\sum^{-1}(y-\mu) : X^2(n)$$

olduğunu görürüz.

Sabitlerin bir A matrisi için $y : N_n(\mu,\sum)$ olduğunda kuadratik formların dağılımı üzerindeki bu, sonuçları $y^T Ay$ nin dağılımını elde etmek için genelleştirmek istiyoruz.Bunu yapabiliriz, ancak ilk önce bir çift sonuca ihtiyacımız var.

Teorem 1.6.1 Eğer $y: N_n(\mu, \sum)$ ise bu taktirde $y^T Ay$ nin moment çıkaran fonksiyonu;

$$m_{y^T Ay}(t) = \left| I_n - 2tA\sum \right|^{-1/2} \exp[-\frac{1}{2}\mu^T \{I_n - (I_n - 2tA\sum)^{-1}\}\sum^{-1}\mu]$$

dır.

İspat: Yine, bu sonucun ispatı sadece

$$E(e^{ty^T Ay}) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{ty^T Ay} f_y(y) dy_1 \ldots dy_n$$

Bir çoklu normal yoğunluk y_i değerlendirmeyi içerir.

Ayrıntılar için bkz Searle 1971, sy. 55. Böylece ispat tamamlanır. Belki daha evvel ifade etmemiz gereken özdeğer sonuçlarının bir çiftine de ihtiyacımız var .

Sonuç 1.6.1 Eğer λ; A' nın bir özdeğeri ve x; A'nın bu özdeğerine karşılık gelen bir özdeğer vektörü ise, bu taktirde c *ve* k skalerleri için $(c\lambda + k, x)$; $cA + kI$ matrisinin bir özdeğer-özvektör çiftidir.

İspat: (λ, x); A için bir öz çift olduğundan, onlar

$$cAx = c\lambda x$$

bağıntısını gösteren, $Ax = \lambda x$ bağıntısını sağlar. Bu denklemin her iki yanına kx ekleyerek,

$$cAx + kx = c\lambda x + kx$$
$$\Rightarrow (cA + kI)x = (c\lambda + k)x$$

bağıntısını elde ederiz, bu nedenle $(c\lambda + k, x)$; $cA + kI$ matrisinin bir özdeğer-özvektör çiftidir.

Sonuç 1.6.2 A nın özdeğerlerinin tümü $-1 < \lambda < 1$ eşitsizliğini sağlarsa bu taktirde

$$(I - A)^{-1} = I + A + A^2 + A^3 + \ldots = I + \sum_{k=1}^{\infty} A^k$$

dır.

İspat: Bu birim matrisi elde etmek için $(I-A)(I+\sum_{k=1}^{\infty}A^k)$ çarpımıyla doğrulanabilir. A nın tüm özdeğerleri için $-1<\lambda<1$ eşitsizliğinin $\lim_{k\to\infty}A^k \to 0$ limitinin garantilendiğine, bu nedenle, $\sum_{k=1}^{\infty}A^k$ nın yakınsak olduğuna dikkat ediniz.

Böylece ispat tamamlanır.

Daha evvel, $boy(V)=k$ olmak üzere bir $V\in R^n$ altuzayı üzerine bir P izdüşüm matrisinin k sayıda 1 e eşit ve $n-k$ sayıda sıfıra eşit özdeğere sahip olduğunu saptadık. Bir izdüşüm matrisinin simetrik ve idempotent olduğunu hatırlayınız. Daha genel olarak bu sonuç tüm idempotent matrislere genişletilebilir.

Teorem 1.6.2 Eğer A ranklı bir $n x n$ idempotent matris ise, bu taktirde A bire eşit n tane ve sıfıra eşit $n-r$ tane eşdeğere sahiptir.

İspat: Genel olarak eğer λ; A için bir özdeğer ise bu taktirde

$$A^2x = A(Ax) = A\lambda x = \lambda Ax = \lambda\lambda x = \lambda^2 x$$

olduğundan λ^2; A^2 için bir özdeğerdir. $A^2 = A$ olduğundan $A^2x = Ax = \lambda x$ elde ederiz. $A^2x = \lambda^2 x$ *ve* $A^2x = \lambda x$ in sağ yanlarını eşitleyerek, $\lambda x = \lambda^2 x$ *veya* $(\lambda-\lambda^2)x = 0$ elde ederiz. Ancak $x \neq 0$ dır, bu nedenle $\lambda-\lambda^2 = 0$ dır ki buradan λ ya 0 yada 1 olmalıdır. A idempotent olduğundan 0 pozitif yarı tanımlı olmalıdır. Bu nedenle, sıfırdan farklı özdeğerlerin sayısı $rank(A)=r$ ye eşittir ve böylece r sayıda özdeğer 1 ve $n-r$ sayıda özdeğer 0 dır. Bu ise ispatı tamamlar.

Şimdi asıl sonucumuzu ifade etmeye hazırız;

Teorem 1.6.3 $y: N_n(\mu,\sum)$ ve A; $rank(A)=r$ ranklı bir $n x n$ simetrik matrisi olsun. $\lambda = \frac{1}{2}\mu^T A\mu$ olsun. Bu taktirde

$y^T Ay : X^2(r,\lambda)$ olması için gerek ve yeter şart

$A\sum$ *idempotent* olmasıdır.

İspat: Teorem 1.6.1' den $y^T Ay$ nin

$m_{y^T Ay}(t) = \left|I_n - 2tA\sum\right|^{-1/2} \exp[-\frac{1}{2}\mu^T\{I_n - (I_n - 2tA\sum)^{-1}\}\sum^{-1}\mu]$ dur. Sonuç 1.6.1' e göre λ; $A\sum$ ' nın bir özdeğeri olmak üzere $I_n - 2tA\sum$ ' nın özdeğerleri $I - 2t\lambda_i$, $i = 1,...,n$ dir. Determinant özdeğerlerin çarpımına eşit olduğundan $\left|I_n - 2tA\sum\right| = \prod_{i=1}^{n}(1-2t\lambda_i)$

elde ederiz. Ayrıca teorem 1.6.2' ye göre her i için

$-1 < 2t\lambda_i < 1$ olması şartıyla $(I_n - 2tA\sum)^{-1} = I_n + \sum_{k=1}^{\infty}(2t)^k (A\sum)^k$ elde ederiz. Bu nedenle $m_{y^T Ay}$

$$m_{y^T Ay}(t) = (\prod_{i=1}^{n}(1-2t\lambda_i)^{-1/2})\exp[-\frac{1}{2}\mu^T\{-\sum_{k=1}^{\infty}(2t)^k (A\sum)^k\}\sum^{-1}\mu]$$

Şimdi $A\sum$ nın $r = rank(A)$ ranklı bir idempotent matris olduğunu farz ediniz. Bu taktirde $(A\sum)^k = A\sum$ dır ve λ_i lerin r tanesi 1' e ve λ_i lerin $n-r$ tanesi 0 eşittir. Bu nedenle, $-1 < 2t < 1$ veya $-\frac{1}{2} < t < \frac{1}{2}$ olması şartıyla

$$m_{y^T Ay}(t) = \left(\prod_{i=1}^{n}(1-2t\lambda_i)^{-1/2}\right)\exp[-\frac{1}{2}\mu^T\{-\sum_{k=1}^{\infty}(2t)^k\}A\sum\sum^{-1}\mu]$$

$$= (1-2t)^{-r/2}\exp[-\frac{1}{2}\mu^T\{1-(1-2t)^{-1}\}A\mu]$$

dır ki bu 0 ın komşuluğundaki t için momet çıkaran fonksiyonunun mevcut olduğu gereksinimi ile uyumludur. Burada $-1 < x < 1$ için $1/(1-x) = 1 + \sum_{k=1}^{\infty}x^k$ seri açılımını kullandık. Bu nedenle

$$m_{y^T Ay}(t) = \frac{\exp[-\frac{1}{2}\mu^T A\mu\{1-1/(1-2t)\}]}{(1-2t)^{r/2}}$$

dır ki bu bir $X^2(r, \frac{1}{2}\mu^T A\mu)$ rasgele değişkenin moment çıkaran fonksiyonudur.

Tersinin ispatı için (yani $y^T Ay : X^2(r,\lambda)$' nın $A\sum$ nın idempotent olduğunu gösterdiğinin ispatı için)

Bkz. Searle (1971, ss. 57-58). Birkaç faydalı sonuç önceki teoremden sonuçlar olarak kolayca görülür.

Sonuç 1.6.3 Eğer $y : N_n(0,\sigma^2 I)$ ise bu taktirde $\frac{1}{\sigma^2} y^T Ay : X^2(r)$ olması için gerek ve yeter şart A nın idempotent ve r ranklı olmasıdır.

Sonuç 1.6.4 $y : N_n(0,\sigma^2 I)$ ise bu taktirde $\frac{1}{\sigma^2} y^T Ay : X^2(r, \frac{1}{2\sigma^2}\mu^T A\mu)$ olması için gerek ve yeter şart A nın r ranklı idempotent olmasıdır.

Sonuç 1.6.5 $y : N_n(\mu,\sigma^2 I_n)$ olduğunu farz ediniz ve p_v $r \le n$ boyutlu bir $V \in R^n$ altuzayı üzerine izdüşüm matrisi olsun. Bu takdirde

$$\frac{1}{\sigma^2} y^T P_v y = \frac{1}{\sigma^2}\left\| p(y|V) \right\|^2 : X^2(r, \frac{1}{2\sigma^2}\mu^T P_v \mu) = X^2(r, \frac{1}{2\sigma^2}\left\| p(\mu|V) \right\|^2)$$

dır.

Sonuç 1.6.6 $y : N(\mu,\sum)$ olduğunuz farz ediniz ve c sabitlerin bir $n x 1$ vektörü olsun bu takdirde ;

$$\lambda = \frac{1}{2}(\mu - c)^T \sum^{-1}(\mu - c) \text{ için } (y-c)^T \sum^{-1}(y-c) : X^2(n,\lambda)$$

dır. Klasik lineer model $y : N_n(\mu,\sigma^2 I_n)$ biçimine sahiptir.

Burada, μ' nün bir $V = L(x_1,...,x_n) = C(x)$ altuzayında uzandığı kabul edilir. Yani herhangi bir β için $\mu = X\beta$' dır.

Bu nedenle $\hat{y}_v = p(y|V)$ yani y' nin V üzerine izdüşümünün

- $\hat{y}_V$ ' nin fonksiyonlarının istatistiksel özellikleri ve $y-\hat{y}_V = p(y|V^{\perp})$ hata vektörü ile ilgilenceğiz.
- y ' nin fonksiyonlarının dağılımsal biçimi (normal ki-kare vs.) (örneğin $\hat{y}_V$) y nin dağılımsal biçim (genellikle normal kabul edilen) vasıtasıyla belirlenir.
- y ' nin lineer fonksiyonlarının beklenen değeri sadece $E(y)=\mu$ vasıtasıyla belirlenir.
- y ' nin kuadratik fonksiyonlarının (Örneğin $\|\hat{y}_V\|^2$) beklenen değeri μ *ve* $\mathrm{var}(y)$ ile belirlenir.

Bunun bir sonucu olarak aşağıdaki Teorem 1.6.4' ü elde ederiz.

Teorem 1.6.4 V; R^n nin bir boyutlu bir alt uzayı olsun ve y; R^n de $E(y)=\mu$ ortalamalı bir rasgele vektör olsun

Bu taktirde

1. $E\{p(y|V)\} = p(\mu|V)$ dir.
2. Eğer $\mathrm{var}(y)=\sigma^2 I_n$ ise, bu taktirde

$$\mathrm{var}\left\{p(y|V)\right\} = \sigma^2 P_v \ ve\ E\left\{\| P(y|V) \|^2\right\} = \sigma^2 k + \| P(\mu|V) \|^2$$
$$dır, ve$$

3. Eğer ayrıca y nin m değişkenli normal olduğunu, yani, $y \sim N_n(\mu,\sigma^2 I_n)$ ise, bu taktirde

$$p(y|V) \sim N_n(p(\mu|V),\sigma^2 P_V)$$

ve

$$\frac{1}{\sigma^2}\|p(y|V\|^2 = \frac{1}{\sigma^2} y^T P_V y \sim X^2(k, \frac{1}{2\sigma^2}\underbrace{\mu^T P_V \mu}_{=\|p(\mu|V)\|^2})$$

Dır.

İspat:

1. İzdüşüm işlemi lineer olduğundan $E\{p(y|V)\} = p(E(y)|V) = p(\mu|V)$ dır.

2. $p(y|V) = P_V y$ dır, bu nedenle $\mathrm{var}\{p(y|V)\} = \mathrm{var}(P_V\, y) = P_V \sigma^2 I_n P_V^T = \sigma^2 P_V$ dır.

Ayrıca, $\|p(y|V\|^2 = p(y|V)^T p(y|V) = (P_V y)^T P_V y = y^T P_V y$

dır. Bu nedenle

$$E(\|p(y|V)\|^2) = E(y^T P_V y)$$

bir kuadratik formun beklenen değeridir ve bundan dolayı

$$E(\|p(y|V\|^2) = iz(\sigma^2 P_V) + \mu^T P_V \mu = \sigma^2 is(P_V) + \mu^T P_V^T P_V \mu$$

$$= \sigma^2 k + \|p(\mu|V\|^2$$

ye eşittir.

3. Teorem 1.6.4' ten kolayca anlaşılmaktadır.

Böylece, izdüşüm ve bu izdüşüm uzunluğunun karesi için dağılımsal sonuçları elde ettik. Lineer modellerde, genel olarak $\hat{y}_V = p(y|V)$ yi oluşturmak için örneklem uzayını V model uzayı üzerine ve $e = y - \hat{y}_V$ hatasını (hata tahminini) oluşturmak için V model uzayının ortogonal tümleyeni olan $V^{\perp}$ üzerine izdüşürerek ayrıştırırız. Bazı durumlarda daha öte gider ve model uzayını model uzayı içindeki altyazılara izdüşürerek ayrıştırırız.Böyle izdüşümlerin ortak dağılımı nedir ?

Pekala , eğer altuzaylar ortogonal iseler ve eğer klasik lineer modelin şartları (normallik bağımsızlık ve eşit varyanslılık) geçerli olur ise, bu taktirde bu altuzaylar üzerine izdüşümler, bağımsız normal rasgele vektörler ve onların uzunluk karelerinin, bir ANOVA (Analysis Of Variance) (Varyans Analizi)' daki karelerin toplamlarının bağımsız ki-kare rasgele değişkenleri oldukları ortaya çıkar.

Bu nedenle, eğer lineer modelin altında yatan geometriyi anlarsak, bu taktirde Teorem 1.6.5' i kullanabiliriz.

Teorem 1.6.5 $V_1,....,V_k$; R^n nin sırasıyla $d_1,....,d_k$ boyutlu karşılıklı olarak ortagonal alt uzayları oluşur ve y ; R^n de $E(y) = \mu$ ortalamasına sahip olan değerleri olan bir rasgele vektör olsun . P_i ; $\hat{y}_i = p(y|V_i) = P_i y$ olacak şekilde V_i üzerine izdüşüm matrisi olsun ve $\mu_i = P_i \mu = 1,...,n$ olsun . Bu taktirde

1. Eğer $\mathrm{var}(y) = \sigma^2 I_n$ ise , bu taktirde $i \neq j$ için $kov(\hat{y}_i, \hat{y}_j) = 0$ dır.

2. Eğer $y : N_n(\mu,\sigma^2 I_n)$ ise , bu taktirde $\hat{y}_i : N(\mu_i,\sigma^2 P_i)$ olmak üzere $\hat{y}_1,....,\hat{y}_k$ lar bağımsızdırlar ve

3. Eğer, $y : N_n(\mu,\sigma^2 I_n)$ ise , bu taktirde

$$\frac{1}{\sigma^2}\|\hat{y}_i\|^2 : X^2(d_i,\frac{1}{2\sigma^2}\|\mu_i\|^2)$$

olmak üzere,

$$\|\hat{y}_1\|^2,....,\|\hat{y}_k\|^2$$

ler bağımsızdırlar.

İspat :

1. Kısmı: $i \neq j$ i için

$$kov(\hat{y}_i,\hat{y}_j) = kov(P_i y, P_j y) = P_i kov(y,y)P_j = P_i\sigma^2 IP_j = \sigma^2 P_i P_j = 0$$

dır.

(Herhangi bir $z \in R^n$ için $P_i P_j z = 0 \Rightarrow P_i P_j = 0$)

2.Kısmı: Eğer, y m değişkenli normal dağılımına sahip ise bu taktirde

$\hat{y}_i = P_i y_i$, $i = 1,...,k$ lar birlikte çok değişkenli normal dağılıma sahiptirler ve bu nedenle bağımsız olmaları için gerek ve yeter şart $kov(\hat{y}_i,\hat{y}_j) = 0$, $i \neq j$ olmasıdır. $\hat{y}_i$ ların ortalama ve varyans-kovaryansları

$$E(\hat{y}_i) = E(P_i y) = P_i\mu = \mu_i \text{ ve } \mathrm{var}(\hat{y}_i) = P_i\sigma IP_i^T = \sigma^2 P_i$$

dır.

3.Kısmı : Eğer $\hat{y}_i = P_i y$, $i = 1,....,k$ lar karşılıklı olarak bağımsız iseler, bu taktirde herhangi (ölçülebilir)* $f_i(\hat{y}_i)$, $i = 1,....,k$ fonksiyonları karşılıklı olarak bağımsızdır. Bu nedenle $\|\hat{y}_i\|^2$, $i = 1,...,k$ ler karşılıklı olarak bağımsızdır.

$$\sigma^{-2}\|\hat{y}_i\|^2 : X^2(d_i,\frac{1}{2\sigma^2}\|\mu_i\|^2)$$

olduğu önceki teoremin kısım 3 ünden anlaşılır.

Bir başka türlü, iz düşümlerin ve onların uzunluk karelerinin bağımsız olup olmadığını belirlemek için bir cebirsel yaklaşımı ortaya koyabiliriz. Geometrik yaklaşım belki daha kolaydır, ancak geometriyi anlayabilirseniz. Fakat cebirsel yaklaşımıda ifade edeceğiz.

1.7 Lineer ve Kuadratik Formların Bağımsızlığı

Burada,

1.Bir lineer form ve bir kuadratik formun (örneğin bir örneklem probleminde $\bar{y}$ *ve* s^2 nin bağımsız olup olmadığını göz önüne alırız veya bir regresyon kurgulamasında $\hat{\beta}$ *ve* $\hat{\sigma}$ nin bağımsızlığını göz önüne alırız).

2. İki kuadratik formun (örneğin, bir regresyonda regresyona bağlı kareler toplamı ve hata kareler toplamı) ve

3. Birkaç kuadratik formun (örneğin, bir iki yönlü varyans analizi tasarımında SS_A , SS_B , SS_{AB} , SS_E nın ortak dağılımını göz önüne alınız)

istatistiksel bağımsızlığını gözönüne alırız.

* Tüm sürekli fonksiyonlar ve pek çok “iyi-huylu” fonksiyon ölçülebilirdir.

Devam etmeden önce, bir lemma ve onun sonucuna ihtiyacımız var.

Lemma 1.7.1 Eğer, $y : N_n(\mu, \sum)$ ise bu taktirde $kov(y, y^T Ay) = 2\sum A\mu$

dür.

İspat: Kuadratik formun kovaryans ve beklenen değerinin tanımını kullanarak,

$$kov(y, y^T Ay) = E[(y-\mu)\{y^T Ay - E(y^T Ay)\}]$$

$$= E[(y-\mu)\{y^T Ay - is(A\sum) - \mu^T A\mu\}]$$

elde ederiz. Şimdi

$$y^T Ay - \mu^T A\mu = (y-\mu)^T A(y-\mu) + 2(y-\mu)^T A\mu$$

cebirsel özdeşliği kullanarak

$$kov(y, y^T Ay) = E[(y-\mu)\{y-\mu)^T A(y-\mu) + 2(y-\mu)^T A\mu - iz(A\sum)\}]$$

$$= E\{(y-\mu)(y-\mu)^T A(y-\mu)\} + 2E\{(y-\mu)(y-\mu)^T A\mu\} - E\{(y-\mu)iz(A\sum)\}$$

$$= 0 + 2\sum A\mu - 0$$

elde ederiz. Burada çok değişkenli normal dağılımın üçüncü merkezi momentleri 0 olduklarından, ilk terim 0 a eşittir. Böylece ispat tamamlanır.

Sonuç 1.7.1 B sabitlerin bir $k\,x\,n$ matrisi olsun ve $y:\ N_n(\mu, \sum)$ olsun. Bu taktirde

$$kov(By, y^T Ay) = 2B\sum A\mu$$

dır. Şimdi Teorem 1.7.1' i gözönüne almak için hazırız.

Teorem 1.7.1 B nin sabitlerin $k\,x\,n$ matrisi **A** nın sabitlerin bir $n\,x\,n$ simetrik matrisi ve $y:\ N_n(\mu, \sum)$ olduğunu farz ediniz . Bu taktirde By ve $y^T Ay$ nın bağımsız olmaları için gerek ve yeter şart $B\sum A = 0_{k\,x\,n}$ olmasıdır.

İspat: Bu teoremi **A** nın bir izdüşüm matrisi olduğu ek varsayımı altında (yani, **A** nın simetrik olmasının yanı sıra idempotent olduğunu kabul ederek) ispatlarız ki bu çalışmada en çok ilgilendiğimiz durumdur, ispatı tamamlamak için bkz Searle (1971 , s.59) **A** nın simetrik ve idempotent olduğunu kabul ederek bu durumda

$$y^T Ay = y^T A^T = \|Ay\|^2$$

yazarız. Şimdi $B\sum A = 0$ olduğunu farz ediniz. Bu taktirde

$$kov(By, Ay) = B\sum A = 0$$

olmak üzere $By\ ve\ Ay$ nin herbiri normaldir. Bu nedenle, $By\ ve\ Ay$ biri diğerinden bağımsızdır. Ayrıca By ; Ay nin herhangi bir (ölçülebilir) fonksiyonundan bağımsızdır, bu nedenle By ; $\|Ay\|^2 = y^T Ay$ den bağımsızdır.

Şimdi, tersi için $By\ ve\ y^T Ay$ nın bağımsız olduğunu farz ediniz. Bu taktirde $kov(By, y^T Ay) = 0$ dır, bu nedenle yukarıdaki lemmayla ilgili sonuca göre,

$$0 = kov(By, y^T Ay) = 2B\sum A\mu$$

dür. Tümü mümkün μ için bu gerçekleştiğinden $B\sum A = 0$ olduğu anlaşılır. Böylece ispat tamamlanır.

Sonuç 1.7.1 Eğer $y :\ N_n(\mu, \sigma^2 I)$ ise bu taktirde $By\ ve\ Ay$ nın bağımsız olmaları için gerek ve yeter şart $BA = 0$ olmasıdır.

Örnek 1.7.1 Bir $y = (y_1, ..., .y_n)^T\ :\ N_n(\mu, \sigma^2 I_n)$ rasgele örneklemine sahip olduğumuzu farz ediniz ve

$$P_{L(jn)} \perp = I_n - \frac{1}{n} J_n J_n^T \text{ olmak üzere}$$

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i = \frac{1}{n} j_n^T y$$

ve

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2 = \frac{1}{n-1}\left\| P_{L(j_n)} \perp y \right\|^2 \text{ yi}$$

gözönüne alınız. Yukarıdaki sonuca göre $\bar{y}\ ve\ s^2$ bağımsızdır, çünkü

$$\frac{1}{n} j_n^T P_{L(j_n)} \perp = \frac{1}{n}(P_{L(j_n)}^T \perp j_n)^T = (P_{L(j_n)} \perp j_n)^T = \frac{1}{n}(0)^T = 0^T$$

dır.

Teorem 1.7.2 **A** ve **B** sabitlerin simetrik matrisleri olsun. Eğer $y :\ N_n(\mu, \sum)$ ise, bu taktirde $y^T Ay\ ve\ y^T By$ nın bağımsız olmaları için gerek ve yeter şart $A\sum B = 0$ olmasıdır.

İspat: Yine **A** ve **B** simetrik ve idempotent (izdüşüm matrisleri) olduğundaki özel durumunu gözönüne alırız ve ispatın sadece "gerek" kısmını veririz. (ispatı tamamlamak için bkz. Searle 1971 , s.59-60)

A ve **B** nin simetrik ve idempotent olduklarını ve $A\sum B=0$ olduğunu farz ediniz. Bu taktirde $y^T Ay=\|Ay\|^2$, $y^T By=\|By\|^2$ dır. Eğer $A\sum B=0$ ise bu taktirde

$kov(Ay,By)=A\sum B=0$ dır.

$Ay \; ve \; By$ nin herbiri çok değişkenli normaldir ve onlar 0 kovaryansına sahiptir, bu nedenle onlar bağımsızdır. Ayrıca $Ay \; ve \; By$ nin herhangi (ölçülebilir) fonksiyonları bağımsızdır, bu nedenle $\|Ay\|^2 \; ve \; \|By\|^2$ bağımsızdır. Böylece ispat tamamlanır.

Sonuç 1.7.2 Eğer $y : N_n(\mu,\sigma^2 I)$ ise bu taktirde $y^T Ay \; ve \; y^T By$ nin bağımsız olmaları için gerek ve yeter şart $AB=0$ olmasıdır.

Son olarak, normal rasgele vektörlere göre çeşitli kuadratik formların karşılıklı olarak bağımsızlığını ilgilendiren bir teorem ve sonucu elde ederiz

Teorem 1.7.3 $y : N_n(\mu,\sigma^2 I_n)$ olsun A_i ; r_i , $i=1,...,k$ ranklı simetrik matris olsun ve $y^T Ay=\sum_{i=1}^{k} y^T A_i y$ olacak şekilde $A=\sum_{i=1}^{k} A_i$ r rankına sahip olsun. Bu taktirde

1. $y^T A_i y/\sigma^2 : X^2(r_i,\mu^T A_i \mu/\{2\sigma^2\})\, i=1,....,k$; ve

2. $y^T A_i y \; ve \; y^T A_j y \;\; i\neq j$ için bağımsızdır ve

3. $y^T Ay/\sigma^2 : X^2(r,\mu^T A\mu/\{2\sigma^2\})$

olması için gerek ve yeter şart aşağıdaki ifadelerin herhangi ikisinin doğru olmasıdır.

a.Her biri A_i idempotentdir

b.Her $i\neq j$ için $A_i A_j=0$ dır

c.A idempotentdir

veya gerek ve yeter şart (d) aşağıdaki gibi olmak üzere (c) ve (d) nın doğru olmasıdır.

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

İspat: Bkz. Searle (1971 , ss.61-64)

- (a),(b) ve (c) den herhangi birinin üçüncüsünü ifade ettiğine dikkat ediniz.
- Önceki teorem bir $y^T Ay$ ağırlık kareler toplamının birkaç bileşene parçalanması

ile ilgilendi. Şimdi A=I olduğu $y^T y$ nin genel kareler toplamının (ağırlıklı olmayan) kuadratik formların bir toplamına ayrıştırıldığı özel durumu işleme sokan bir sonucu ifade edeceğiz.

Sonuç 1.7.3 $y:\ N_n(\mu,\sigma^2 I)$ olsun. A_i ; r_i , $i=1,...,k$, ranklı simetrik matris olsun ve $y^T y = \sum_{i=1}^{k} y^T A_i y$ (yani, $\sum_{i=1}^{k} A_i$ =I)

olduğunu farz ediniz.Bu taktirde

1.Herbir $y^T A_i y/\sigma^2 :\ X^2(r_i, \mu^T A_i \mu/\{2\sigma^2\})\ j$ ve

2.Herbir $y^T A_i y$ lerin karşılıklı olarak bağımsız olması için gerek ve yeter şart, aşağıdaki ifadelerin sağlanmasıdır:

a. Herbir A_i idempotenttir.

b.Her $i \neq j$ için $A_i A_j = 0$ dır.

c. n=$\sum_{i=1}^{k} r_i$ dir.

Bu teorem, bir varyans analizindeki genel kareler toplamının bir ayrışımında kareler toplamlarının bağımsızlığı için çoğu kez gerekçe olarak alıntılanan Cochran Teoremini kapsar. (Anonim 1)

2. Şapka (Hat) Matrisi

2.1 Gözlenen Değerlerden Uydurulan Değerlere

$$Y = x\beta + \in,\ X\ nxp,\ rak(x) = p,\ \in:\ N(\mu,\sigma^2 I)$$

Lineer modelinde β nın

$$\hat{\beta}_{px1} = (X^T X)^{-1} X^T y$$

OLS (Ordinary Least Squares) (Sıradan En Küçük Kareler) tahmin edicisi en küçük kareler yöntemiyle bulunur.

Tahmin edilen $\hat{y}$ değerleri bu durumda

$$\hat{y} = X\hat{\beta} = (X(X^T X)^{-1} X^T)y =: Hy$$

olarak yazılabilir. Burada y üzerine " şapka" koyan ve bu nedenle şapka matrisi olarak başvurulan $H := X(X^T X)^{-1} X^T$ bir nxn matristir. Aslında bu uydurulan değerlerin, herhangi bir y için $y' \in IR^n$ olacak şekilde gözlenen değerlerin bir lineer fonksiyonu olduğunu gösterir. $y' = f(y)$ koyarak

$$f(y + y') = H(y + y') = Hy + Hy' = f(y) + f(y')$$

ve herhangi bir $a \in R$ için, $f(ay) = af(y)$

elde ederiz. Benzer şekilde, hatalar (tahmin edilen hatalar) I nxn birim matris olmak üzere H nın bir fonksiyonu olarak da ifade edilebilir yani,

$$\hat{e} := y - \hat{y} = y - Hy = (I - H)y$$

dır ve burada yine hataların gözlenen değerlerin, yani y nin bir lineer fonksiyonu olduğu da görülebilir. Bu nedenle özet olarak

$$\hat{y} = Hy \; ve \; \hat{e} = (I - H)y$$

elde ederiz.Çok önemli olarak $H \; ve \; I - H$ ın her ikisinin de ortogonal izdüşümler (iz düşürücüler) oldukları gösterilebilir. Bu çoklu lineer regresyon üzerinde iki farklı bakış açısını karşılaştırarak görsel olarak gösterilebilir.

i. Buraya kadar aslında her bir deneğin bir nokta ve değişkenlerin ölçüler olduğu , değişken uzayı denilebilen R^p ile ilgilendik.

ii. Denek uzayın , yani R^n yi gözönüne alarak bir bakış açısı değişikliği yapmamız gerekir . Bu öklid uzayında, her bir denek bir ölçüdür(boyuttur) halbuki $y, \hat{y} \; ve \; \hat{e}$ vektörler olarak işleme sokulur.

2.2 Ortogonal İzdüşüm Matrisi Olarak Şapka Matrisi

Bir izdüşümün aynı zamanda simetrik olan bir matrisi bir ortogonal izdüşümdür. $H \; ve \; I - H$ ın her ikisininde ortogonal izdüşümler olduklarını gösterebiliriz. Bu iki şart aşağıdaki gibi yeniden ifade edilebilir.

1. Bir A kare matrisi bir izdüşümdür, eğer o idempotent ise.

2. Bir A izdüşümü ortogonaldir, eğer o aynı zamanda simetrik ise.

Ortogonal olmayan bir izdüşüme "eğik izdüşüm" denir . Özellikle

H y yi X in sütun uzayı üzerine izdüşürür, halbuki

$I-H$; y yi H nın görüntüsünün ortogonal tümleyeni yerine izdüşürür. Matrisin sütun uzayı karşılık gelen lineer dönüşümün erimi veya görüntüsü olarak tanımlanır.

Formül olarak H bir nxn matristir. Onun sütun uzayı, $H.,j$, H nin J-yinci sütununu göstermek üzere,

$$s\ddot{u}t(H) := \{\sum_{i=1}^{n} c_j H._j : c_j \in R, \forall\ j = 1,....,n\}$$

olarak tanımlanır. Bu H nın lineer bağımsız sütunların gerenidir.

2.3 Şapka Matrisinin İdempotentliği

H bir nxn kare matristir ve ayrıca o idempotentdir. Bu aşağıdaki gibi doğrulanabilir:

$$\begin{aligned}
HH &= (X(X^T X)^{-1} X^T)(X(X^T X)^{-1} X^T) \\
&= X(X^T X)^{-1}(X^T X)(X^T X)^{-1} X^T \\
&= X(X^T X)^{-1} X^T \\
&= H
\end{aligned}$$

Benzer şekilde $I-H$ nın da idempotent olduğu, yani

$$(I-H)(I-H) = I - 2H + HH = (I-H)$$

olduğu gösterilebilir. Her kare ve idempotent matris bir izdüşüm matrisidir. Bir izdüşümün anlamı herhangi bir 2 boyutlu X vektörünü bir boyutlu bir altuzaya gönderen aşağıdaki 2x2 boyutlu bir P izdüşüm örneğiyle anlaşılabilir.

$$Px = \begin{pmatrix} 1 & \alpha \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 + \alpha x_2 \\ 0 \end{pmatrix}$$

P nin idempotentliği, bir vektörel bir kez bir altuzaya izdüşürüldüğünde aynı izdüşümü uygulasak bile, onun orada kaldığını ifade eder.

2.4 Şapka Matrisinin Simetrikliği

Herhangi kare ve tersinin matrisler için, ters ve transpoze işlemi yer değiştirir, yani,

$$(X^T)^{-1} = (X^{-1})^T$$

dır.Ayrıca $(X^T)^T = X$ olduğundan, transpoze birli (unary) işlemi bir involusyondur. Bu nedenle

$$[(X^T X)^{-1}]^T = [(X^T X)^T]^{-1} = (X^T X)^{-1}$$

olduğundan, $(X^T X)^{-1}$ in kendinin transpozesi olduğu(yani simetrik olduğu) görülür. H nin kendisinin transpozesi veya simetrik olduğunu göstermek için biri bu iki kuralı uygulayabilir. Buna göre,

$$HT = [X(X^T X)^{-1} X^T]^T = X[(X^T X)^{-1}]^T X^T = H$$

dır, yani H simetriktir. Bu simetri $I - H$ ile aşağıdaki tarzdadır

$$(I - H)^T = I^T - H^T = (I - H)$$

Bir matrisinin ortogonalliği bir örnekle en iyi şekilde anlaşılır.

Herhangi bir $x \in R^2$ vektörüne uygulandığında aşağıdaki p_{2x2} matrisi ve onun $I - p$ tümleyenini gözönüne alınız.

$$Px = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ 0 \end{bmatrix}$$

ve

$$(I - P)x = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ x_2 \end{bmatrix}$$

aşikar olarak, biri sonuçta ortaya çıkan iki vektörün dik olduklarını, yani $< px, (I - P)x >= 0$ olduğunu ve bu nedenle, onların R^2 nin iki ortogonal lineer uzayını gördüğünü hemen görür.

2.5 Ortogonal İzdüşümler ve Ortogonal Matrisler

Biri ortogonal matris ile bir ortogonal izdüşümün matrisini karıştırmamak için dikkatli olunmalıdır.

Bir A ortogonal matrisinin

$$A^T A = AA^T = I$$

eşitliğini sağladığını hatırlayınız.

Özellikle bir ortogonal matris ortogonal olan sütunlara sahip iken, bir H ortogonal izdüşümün vektörleri onun sütun uzayındaki vektörlere ortogonal izdüşürdüğünü gözlemleyebilir. Genel olarak, bir ortogonal matris bir ortogonal izdüşüme neden olmaz. Aslında hem bir ortogonal izdüşüm ve hem de bir ortogonal matris olan yegane matris birim matristir.

2.6 Ortogonal Ayrışım

2.6.1 Şapka Matrisinin Erimi (Range) ve Çekirdeği

Uydurulan değerler ve tahmin edilen hatalar(rezidüler) tahminimizi birleştirerek

$$\hat{y} = Hy \ \ ve \ \ \hat{e} = (I - H)y$$

elde ederiz. Bu denklemler

$$y = \hat{y} + \hat{e} = Hy + (I - H)y$$

olacak şekilde, gözlenen değerlerin ortogonal ayrışımına karşılık gelir. $süt(H)$ ile gösterilen, H'ın sütun uzayı ve görüntüsünün (eriminin), X in sütun uzayı ile aynı olduğuna dikkat ediniz. X in reel elemanlı bir matris olduğunu ve bu nedenle, X in rankının,

$$rank(x) = rank(x^T x) = P$$

olacak şekilde, $X^T X$ olarak tanımlanan, Gram matrisinin rankına eşit olduğunun bilindiğini hatırlayınız.

Ayrıca, kxk mertebeli herhangi bir A kare matrisi için, eğer B k ranklı bir nxk matris ise, bu taktirde

$$rank(BA) = rank(A) \ \ ve \ \ rank(AB^T) = rank(A)$$

olacak şekilde, matris rankları üzerinde bazı temel işlemeri kullanabiliriz.

Varsayımdan, $rank(X) = p$ elde ederiz. Bu nedenle

$$rank(H) = rank(X(X^T X)^{-1} X^T) = rank((X^T X)^{-1}) = rank(X^T X) = P$$

yi elde etmek üzere ve sadece tam ranklı matrislerin tersinir olduğuna, ki bunun matrisin tersinir rankı koruduğunu ifade ettiğine dikkat etmek için, bu kuralları uygulamak yeter. Bu nedenle, H ın sütun uzayı:

$$süt(H) = süt(X)$$

olacak şekilde, X in sütun uzayına eşittir. Burada X in sütun uzayı, X in sütunlarının lineer kombinasyonları olarak elde edilebilen tüm vektörlerin kümesidir. Buna umumiyetle X in sütunlarının gereni olarak başvurulur.

Bu vektör altuzayının ortogonal tümleyeni H ın (çekH) ile gösterilen çekirdeği veya sıfır uzayıdır.Özet olarak, H in n tane sütununun uzayı aşağıdaki iki ortogonal vektör altuzayına bölünebilir.

i. $y \in R^n$

ii. $\hat{y} \in X := süt(H) = geren(x_1, ..., x_p)$

iii. $\hat{e} \in x^{\perp} := çek(H) = \{v \in R^n \ : < v, v_j >= 0, \ \forall_j = 1, ..., p\}$

Burada R^n üzerindeki vektör yapısı $X \ ve \ X^{\perp}$ üzerindeki vektör yapılarının direkt toplamını alarak yeniden ele geçirilebilir. Burada $X \ ve \ X^{\perp}$ye R^n nin vektör alt uzayları veya lineer alt uzayları olarak başvurulur.

Daha özel olarak, $X^{\perp}$ lineer alt uzayı, X in ortogonal tümleyenidir. Doğal olarak $\hat{y} \in X \ ve \ \hat{e} \in X^{\perp}$ nin bundan sonraki kısımda gösterildiği gibi, cebirsel olarak kontrol edilebilen $\hat{y} \perp \hat{e}$ olduğunu ifade ettiği gerçektir.

2.6.2 Tahmin Edilen Hataların ve Uydurulan Değerlerin Ortogonalliği

İlk olarak, $(I-H)$ nın H ile soldan çarpımı bir nxn sıfır matrisi verir, yani,

$$H(I-H)=H-HH=H-H=0$$

dır. Geometrik olarak, tahmin edilen hatalar ve uydurulan değerler iki ortogonal vektördür. Bu nedenle, biri R^n deki bu iki vektör arasındaki öklid iç çarpım ve nokta çarpımın sıfır olduğunu doğrulayabilir. Yani ,

$$<\hat{y},\hat{e}>=\hat{y}^T\hat{e}=(Hy)^T(I-H)y=y^TH(I-H)y=y^T0y=0$$

dır. Bu nedenle $\hat{y}$ ve e ; R^n de ortogonaldir.

2.6.3 Tahmin Edilen Hataların ve Uydurulan Değerlerin İlişkisizliği

Ayrıca biri $\hat{y}$ ve $\hat{e}$ nın ilişkisiz olduklarını da gösterebilir. Herhangi iki $x,y\in R^n$ rasgele vektörü için

$$kov[Ax,By]=A\,kov[x,y]B^T$$

gerçeğini kullanmamız gerekecek . Bu nedenle,

$$\begin{aligned}
&kov[\hat{y},\hat{e}\,|\,x]=kov[Hy,(I-H)y\,|\,X]\\
&=H\,kov[y,y\,|\,X](I-H)^T\\
&=\sigma^2H(I-H)=0
\end{aligned}$$

dır. Önceki kısımda bu rasgele süreçlerden kavrayışların sadece iki vektörünü göz önüne almamıza rağmen, burada $\hat{y}$ ve $\hat{e}$ rasgele vektörlerini işlemlerden geçiriyoruz.

2.7 Regresyonun Geometrik Yorumu

2.7.1 $\hat{y}$ ve $\hat{e}$'nin Serbestlik Dereceleri

Geometrik olarak $\hat{y}$ ve $\hat{e}$ ile ilgili serbestlik derecelerinin basitce bu iki vektörün izdüşürüldüğü ilgili vektör alt uzaylarının boyutları alacağı görülebilir. Özellikle, bir V vektör uzayının rank-sıfırlık teoreminin basit uygulamasıyla kendi erim(görüntü)ve çekirdeğine ayrışımı

$$\begin{aligned}
&boy(R^n)=boy(X)+boy(X^{\perp})\\
&=rank(H)+s\imath f(H)
\end{aligned}$$

dır ve burada sıf(H), H ın sıfır uzayının boyutu olarak tanımlanan, H ın sıfırlığını gösterir. p sayıda farklı parametreye sahip bir ortalama fonksiyonu gözönüne alındığında, serbestlik derecelerinin aşağıdaki parçalanmasına sahip oluruz.

$$\begin{aligned} sd(y) &= sd(\hat{y}) + sd(\hat{e}) \\ &= p + (n-p) \end{aligned} \tag{38}$$

H nin bir ortogonal izdüşüm olduğu belirtildiğinden, biri H nin özdeğerlerinin bu vektör uzaylarının boyutlarını da belirlediğini gözlemleyebilir. Aslında bir ortogonal izdüşümün matrisinin özdeğerleri sadece 0 veya 1 olabilir. Bu nedenle, H nin spektrumundaki sıfırların sayısı H nın sıfırlığına eşittir. Halbuki onun spektrumundaki birlerin sayısı onun rankina eşittir.

2.7.2 Pisagor Teoremi Vasıtasıyla Varyans Parçalanması

$y, \hat{y}$ *ve* $\hat{e}$ vektörleri R^n de bir üçgen oluşturan üç noktayı belirler. $<\hat{y}, \hat{e}>=0$ olduğundan bu üçgenin bir dik üçgen veya (dik-açılıüçgen) olduğu anlaşılır. Temel geometriyi kullanarak regrasyon analizinin iki görünümü idrak edilebilir. İlk olarak, genel kareler toplamının (TSS veya SYY) tahmin edilen kareler toplamı (SS_{reg} veya ESS) ve hata (tahmin edilen hata) kareler toplamına (RSS) ayrışımı pisagor teoreminin özel bir durumunu oluşturmak için gösterilebilir. Uygunluk için, y_i leri $\overline{y}$ de merkezleştiririz. Bu taktirde genel kareler toplamı $\|\cdot\|$ öklid normu olmak üzere

$$SYY := \sum_{i=1}^{n} y_i^2 = \| y \|^2$$

ile verilir. Her $i:1,....,n$ için $y_i = \hat{y}_i + \hat{e}_i$ gerçeğini kullanarak,

$$\begin{aligned} \sum_{i=1}^{n} (\hat{y}_i + \hat{e}_i)^2 &= \sum_{i=1}^{n} \hat{y}_i^2 + 2\sum_{i=1}^{n} \hat{y}_i \hat{e}_i + \sum_{i=1}^{n} \hat{e}_i^2 \\ &= \| \hat{y} \|^2 + 2 < \hat{y}, \hat{e} > + \| \hat{e} \|^2 \end{aligned}$$

elde ederiz. Ayrıca, $\hat{y} \perp \hat{e}$ olduğunu önceden göstermiştik. Bu nedenle, $\| y \|$ karesi

$$\| y \|^2 = \| \hat{y} \|^2 + \| \hat{e} \|^2$$

olacak şekilde kenarların uygunluklarının karelerinin toplamına eşit olan, hipotenüsün uzunluğudur.

Bundan sonraki çalışmada bu nicelikler için serbestik derecelerinin nasıl hesaplanacağını göreceğiz. Bu serbestlik dereceleri merkezleştirilmiş veriye sahip olup olmamamıza bağlı olarak (38) bağıntısındaki rasgele vektörlerden farklı olabilir.

2.7.3 Trigonometriyi Kullanarak Determinasyon Katsayısı

İkinci olarak, $y \; ve \; \hat{y}$ arasındaki lineer bağımlılığın derecesini ölçmek için Pisagor Teoreminin başka bir sonucunu da uygulayabiliriz. θ radyanlar cinsinden ölçülmek üzere, herhangi bir dik üçgen için

$$\cos\theta = \frac{|\, yankenar \,|}{|\, hipotenüs \,|}$$

olduğunu hatırlayınız. Regresyanda , $y \; ve \; \hat{y}$ vektörleriyle verilen kenarlar ile R^n de 0 daki açı ile ilgileniriz. Bu nedenle benzer şekilde

$$(\cos\theta)^2 = \frac{\|\hat{y}\|^2}{\|y\|^2} = \frac{ESS}{TSS} = \frac{SYY - RSS}{SYY} = R^2$$

elde ederiz. Böylece, bir regresyon modelinin uyumunu maksimumlaştırma R^n deki $y \; ve \; \hat{y}$ arasındaki açıyı azaltmaya eşdeğerdir. Cos(0)=1 olacak şekilde bu açı sıfır olduğunda mükemmel bir korelasyon (ilişki) ortaya çıkar halbuki tam istatistiksel bağımsızlık (yani korelasyonun yokluğu) bir dik açıya eşdeğerdir, yani $\cos(\pi/2) = 0$ eşdeğerdir. Son olarak, $y \; ve \; \hat{y}$ arasındaki açı doğru açı olduğunda bir negatif korelasyon elde edilir ki bu durumda

$$\cos(\pi) = -1$$

dır.

2.7.4 Doğrudaş (Kolineer) Açıklayıcılar

Buna ek olarak çoklu regrasyon üzerinde bu geometrik bakış istatistikte “kolineer” teriminin kullanımını da aydınlığa kavuşturur. Çeşitli noktalara doğrudaş denir eğer onlar bir tek doğru üzerinde bulunuyor ise çoklu regresyon bağlamında X in herbiri sütunu R^n de bir vektördür. Eğer böyle sütun vektörleri diyelim ki $X_{\cdot,j}$ ve $X_{\cdot,k}$ arasındaki korelasyon mükemmel (tam) ise bu taktirde onlar aynı doğru üzerinde uzanırlar.

3. Çoklu Regresyon, Temel Teori

$$y = x\beta + u$$

dur. Burada $y = (y_1,....., y_n)'$, tepkime değişkeni üzerindeki n gözlemden oluşan veri vektörüdür.

X; açıklayıcı değişkenlerden oluşan bir $nx(p+1)$ açıklayıcılar matrisidir, ki x in birinci sütunu sadece 1 lerden oluşan bir sütun vektördür. $\beta = (\beta_0,...., \beta_p)'$ rasgele olmadığı kabul edilen, regresyon parametrelerinin bir $(p+1)x1$ vektörüdür ve $u = (u_1,...., u_n)'$ rasgele hataların bir $nx1$ vektörüdür.

X in sütunlarını $x_0,....., x_p$ ile gösteriniz

Herbir x_j; y deki n tane gözleme karşılık gelen j-yinci açıklayıcı değişkenin n tane değerini verir.

$u_1,...., u_n$ nin, $E[u_i] = 0$ ve $\text{var}[u_i] = \sigma_{ii}{}^2$ olmak üzere, bağımsız özdeş dağılımlı olduklarını kabul ederiz. Aynı zamanda hata teriminin X in elemanlarından bağımsız olduklarını da kabul ederiz. (Bu açık ve seçik olarak gerçekleşir, eğer açıklayıcı değişkenler rasgele değiller ise). Eğer X matrisi rasgele ise, bu taktirde ortaya koyacağımız dağılımsal ifadeler $x_0,....., x_p$ nin gözlenen değerleri üzerinde şartlı olacak.

β_j katsayısal regresyon fonksiyonundaki , yani

$$E[y \mid x] = x\beta = \sum_{k=0}^{p} x_k \beta_k$$

da ki değişmeyi, j-yinci açıklayıcı değişkende birim değişime karşılık gelen değişmeyi ölçer, eğer model doğru ise ve diğer açıklayıcı değişkenler sabit tutulur ise.

3.1 En Küçük Kareler Tahmin Edicisi

Gözlemlenmiş y verisinden β nın tahmin edilmek istendiğini farz ediniz. En küçük kareler tahmin edicisi $(y - x\beta^x)'(y - \beta^x)$ ölçüt fonksiyonunu minimum yapan β^x in değeridir.Jobson,s.224 çözümün

$$b = (X'X)^{-1}X'y$$

ile verildiğini göstermek için analizi (matematiksel analiz) kullanır. J-yinci b_j elemanı $\hat{y} = xb = \sum_{k=0}^{p} x_k b_k$ uydurulmuş modelindeki x_j nin katsayısıdır.

Bu nedenle, b_j yi y tepkime değişkeninin beklenen değerinde, x_j açıklayıcı değişkenindeki bir birim değişime karşılık gelen değişmenin bir tahmini olarak düşünebiliriz, eğer modelin doğru olduğunu kabul ederek, diğer tüm değişkenler sabit tutulur ise b nin mevcut olması için, $(X'X)^{-1}$ mevcut olmalıdır. b nin mevcut olması için gerek ve yeter şart X in sütunlarının lineer bağımsız olmasıdır. Bu şunu ifade eder, X in sütunlarının herhangi birini X in geriye kalan sütunlarının bir lineer kombinasyonları olarak ifade etmek münkün değildir.

Eğer X in bir sütununu diğer sütunların bir lineer kombinasyonu olarak ifade etmek mümkün ise, şarta çoklu doğrudaşlık denir.

Örneğin, $x_2 = x_0 + x_1 \;\; veya \;\; x_8 = 2x_4 - 3x_6$ ise bu olmalıdır.

3.2 En Küçük Karelerdeki Şapka Matrisi

En küçük kareler tahmin edicisinin gerçekten

$$b = (X'X)^{-1}X'y$$

denklemi ile verildiğinin bir ispatını vereceğiz. Analizi kullanmak yerine,

$$H = X(X'X)^{-1}X'$$

ile tanımlanan nxn şapka matrisine dayanan değer öğretici ve hoş bir ispatı vereceğiz.

$\hat{y} = Hy$ olduğundan, yani y üzerine "şapka" koyduğundan H ya "şapka matrisi" denir.

$$\begin{aligned} H' &= (X')'((X'X)^{-1})'X' \\ &= X((X'X)')^{-1}X' = X(X'X)^{-1}X' = H \end{aligned}$$

olduğundan, H simetriktir. H ın başka bir özelliği onun idempotent, yani $H^2 = H$ olmasıdır. Bunu görmek için

$$H^2 = HH = X\underbrace{(X'X)^{-1}X'X(X'X)^{-1}}_{(X'X)^{-1}}X' = X(X'X)^{-1}X'$$

olduğunu göstermek yeter.

$$H(I-H) = 0,\; HX = X \;\; ve \;\; (I-H)^2 = (I-H)$$

olduğunu göstermek kolaydır.

Herhangi bir $w = (w_1, ..., w_n)'$ n boyutlu vektörü için, w nun uygunluğunu

$$\| w \| = \left[\sum_{i=1}^{n} w_i^2 \right]^{1/2}$$

olarak tanımlarız. $\| w \|^2 = w'w$ olduğuna dikkat ediniz.

Tahmin edilen hata vektörü

$$\hat{e} = y - \hat{y} = (I-H)y$$

ve tahmin edilen hataların kareleri toplamı

$$\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 = \| y - \hat{y} \|^2 = \| (I-H)y \|^2 = y'(I-H)y$$

dir

Teorem 3.2 En küçük kareler tahminleri

$$b = (X'X)^{-1} X'y$$

ile verilir.

İspat: $\| y - x\beta^* \|^2$ nin $\beta^* = b$ seçerek minimumlaştırıldığını göstermememiz gerekir.

$$\| y - x\beta^* \|^2 = \| y - Hx\beta^* \|^2 = \| [Hy + (I-H)y] - Hx\beta^*] \|^2$$
$$= \| H(y - x\beta^*) + (I-H)y \|^2 = \| H(y - x\beta^*) \|^2 + \| (I-H)y \|^2$$

elde ederiz.

$$H(I-H) = 0$$

olduğundan, çapraz çarpıma terimlerinin sıfır olduklarına dikkat ediniz.

$\| H(y - x\beta^*) \|^2 + \| (I-H)y \|^2$ denklemindeki ikinci terim β^* a bağlı değildir.

Bir vektörün uzunluğu olduğundan birinci terim negatif değildir.

$$H = (y - x\beta^*) = 0$$

olacak şekilde

β^* i seçerek, dolayısıyla $\| y - x\beta^* \|^2$ yi minimumlaştırarak, bu terimi sıfır yapabiliriz.

Bu taktirde,

$$H(y - x\beta^*)$$

$$Hy - Hxb = Hy - xb = Hy - X(X'X)^{-1} X'y$$
$$= Hy - Hy = 0$$

a indirgendiğinden, eğer $\beta^* = b$ seçersek, bu gerçekten vuku bulacak. Bu nedenle, b, $\| y - x\beta^* \|^2$ yi minimum yapar.

3.3 En Küçük Kareler Tahmin Edicisinin Ortalama ve Kovaryansı

b en küçük kareler tahmin edicisi β için yansızdı: Gerçekten,

$$E[b] = E[(X'X)^{-1}X'(X\beta) + u)]$$
$$= \beta + (X'X)^{-1}X'E[u] = \beta$$

dır. b nın $(p+1)x(p+1)$ kovaryans matrisi

$$kov(b) = E[(b-\beta)(b-\beta)']$$

dür. Kov(b) nin (i,j) yinci elemanı $i, j = 0...p$ için, b_i b_j arasındaki kovaryanstır.

$b = \beta + (X'X)^{-1}X'u$ olduğundan,

$$kov(b) = E[(X'X)^{-1}X'uu'X(X'X)^{-1}]$$
$$= (X'X)^{-1}X'E[uu']X(X'X)^{-1}$$

elde ederiz.

u_i ler sıfır ortalama ve σ_u^2 varyans ile bağımsız özdeş dağıldıklarından,

$E[uu'] = kov(u) = \sigma_u^2 I$ ve bu nedenle,

$kov(b) = \sigma_u^2 (X'X)^{-1}$ elde ederiz.

3.4 Hata Varyansını Tahmin Etme

σ_u^2 hata varyansının bir tahmin edicisi β ve diğer amaçlar için güven bölgelerini oluşturmak için yararlı olmalı.

$$s^2 = \frac{1}{n-p-1} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = \frac{1}{n-p-1} y'(I-H)y$$

tahmin edicisini tanımlayınız. $n-p-1$ bölenini kullanma nedeni s^2 nin $\sigma_u{}^2$ için yansız olduğunu garanti etmektir.

Teorem 3.4 $E[S^2] = \sigma_u^2$

İspat: Herhangi bir A kere matrisi için

$iz(A) = \sum_i A_{ii}$ yani köşegen elemanlarının toplamı olarak tanımlayınız. Eğer AB ve BA her ikiside iyi tanımlı matrisler ise bu taktirde iz(AB)=iz(BA) dır. Bu nedenle

$$iz(H) = iz(X(X'X)^{-1}X') = iz((X'X)^{-1}X'X) = iz(IpH) = p+1$$

dir. Şimdi

$$\begin{aligned}
E[s^2] &= \frac{1}{n-p-1} E[y'(I-H)y] \\
&= \frac{1}{n-p-1} E[(X\beta+u)'(I-H)(X\beta+u)] \\
&= \frac{1}{n-p-1} [(X\beta)'(I-H)X\beta] + E[u'(I-H)u] \\
&= \frac{1}{n-p-1} [(X\beta)'X\beta - (X\beta)'HX\beta + \sum_{i=1}^{n}\sum_{j=1}^{n} E[u_i u_j](I-H)_{ij}] \\
&= \frac{1}{n-p-1} \sum_{i=1}^{n} \sigma_{ii}{}^2 (I-H)_{ii} = \frac{\sigma_u^2}{n-p-1} iz(I-H) \\
&= \frac{\sigma_u^2}{n-p-1}[n - iz(H)] = \sigma_u^2 \frac{n-p-1}{n-p-1} = \sigma_u^2
\end{aligned}$$

dır.

3.5 En Küçük Karelerin Geometrisi

Bir geometrik bakış açısından H şapka matrisi bir izdüşüm matrisi olarak düşünülebilir. Fazla teknik olmadan bunun ne ifade ettiğini açıklamaya çalışacağız iki n boyutlu w ve v vektörlerine ortogonal diyeceğiz, eğer $w'v = 0$ oluyor ise. Bu tanımın n=2 ve n=3 için sezgimizle uyuştuğuna dikkat ediniz.

Örneğin $(0,1)'$ ve ($(2,0)'$ vektörleri ortogonaldir. $(1,0,2)'$ ve $(0,1,0)'$ vektörleri de ortogonaldir. Herhangi bir n boyutlu w vektörü iki kısma ayrılabilir:

$$w = Hw + (I - H)w$$

veya

$$w = \hat{w} + e$$

birinci bileşen, yani

$$\hat{w} = Hw = X[(X'X)^{-1}X'w]$$

X in sütunlarının bir lineer kombinasyonudur. İkinci bileşen, yani

$$e = w - \hat{w} = (I - H)w$$

X in sütunlarının tüm lineer kombinasyonlarına ortogonaldir. Bunu göstermek için herhangibir

$$X_0\beta_0^* + ... + X_p\beta_p^* = X\beta^*$$

lineer kombinasyonunu gözönüne alınız. Bu taktirde

$$(I - H)X = X - X = 0$$

olduğundan,

$$e'X\beta^* = w'(I - H)X\beta^* = 0$$

dır.

Bu nedenle, w yi Hw ye götüren lineer dönüşümünün, W nun X in sütun uzayına bir ortogonal izdüşümünü oluşturduğunu söyleriz. $\hat{w} = Hw$ olmak üzere, $w = \hat{w} + e$ nin w nun X in sütun uzayına (yani X in sütunlarının tam oluş kombinasyonlarına) izdüşümü olduğuna ve $e = w - \hat{w}$ nın w nun X in X in sütun uzayına ortogonal olan kısmı olduğunu tam anlamıyla göstermiş olduk.

$b = (X'X)^{-1} X'y$ denkleminin ispatında verilenle hemen hemen aynı bir düşünceyle $\hat{w}$ nın, X in sütun uzayında w ya en yakın olan vektör olduğu gösterilebilir.

Yani, $\| w - X\beta^* \|^2$ β^* ı $X\beta^* = \hat{w}$ gibi olarak tüm $\beta^* = (\beta_0^*, \ldots, \beta_p^*)$ üzerinden minimumlaştırılır.

Sonuçta oluşan minimum değeri

$$\| w - \hat{w} \|^2$$

dir.

Şimdi veri vektörümüzü w=y koyarsak bu taktirde $\hat{y} = Hy = Xb$ nin X in sütun uzayında y ye mümkün olduğu kadar yakın olan, vektör olduğunu, yani kriter fonksiyonunun kareler toplamı olan $\| y - X\beta^* \|^2$ yi minimum yapan vektör olduğunu görürüz.

$e = y - \hat{y}$ hata tahmini vektörü X in sütunlarının tüm lineer kombinasyonlarına ortogonaldir.

Özellikle, e ; $\hat{y}$ ya ortogonaldir, bu nedenle

$$\| y \|^2 = \| \hat{y} \|^2 + \| e \|^2$$

dik üçgen ilişkisini elde ederiz.

$$\frac{1}{n}\sum_{i=1}^{n}\hat{y}_i = \frac{1}{n}\hat{y}'X_0 = \frac{1}{n}(Hy)'X_0$$
$$= \frac{1}{n}y'HX_0 = \frac{1}{n}y'X_0 = \bar{y}$$

olduğundan, uydurulan değerlerin ortalaması $\bar{y}$ dır. Bu nedenle $y \; ve \; \hat{y}$ nın ortalama düzeltilmiş versiyonları

$$y - \bar{y} = \hat{y} - \bar{y} + e$$

ile bağlı olur.

Sonuncu nicelik X in sütün uzayında olduğundan e nin $\hat{y} - \bar{y}$ ye ortogonal olduğuna dikkat ediniz . Bu çoğu kez SST=SSR+SSE olarak yapılan çok önemli

$\| y - \bar{y} \|^2 = \| \hat{y} - \bar{y} \|^2 + \| y - \hat{y} \|^2$ ilişkisini verir. Burada

$SST = \sum (y_i - \bar{y})^2$ genel kareler toplamıdır.

$SSR = \sum (\hat{y}_i - \bar{y})^2$ regresyon kareler toplamıdır

$SSE = \sum (y_i - \hat{y}_i)^2$ tahmin edilen hata kareler toplamıdır.

Çoklu determinasyon katsayısı

$$R^2 = {SSR}/{SST}$$

dir.Bu niceliklerin yorumları aslında basit lineer regresyondaki ile aynıdır.

3.6 Normal Dağılan Hatalar için b , s^2 nin Dağılımı

Lineer regresyon için standard dağılım teorisini elde etmek için u_i lerin normal dağılımlı olduklarını kabul etmemiz gerekir. Bu varsayım daha evvel gerekli değildir.

$$V'V = VV' = I$$ (V bir ortogonal matris ve)

$$HV = [V_0, ..., V_p, 0, ..., 0]$$

olacak şekilde bir $V = [V_0, ..., V_{n-1}]$ nxn matrisini oluşturmak mümkündür. Bunu yapmak için bir çok yol(yöntem) vardır. Herşeyden önce $V_0, ..., V_p$ ler X in sütun uzayında olacak şekilde karşılıklı olarak ortgonal $V_0, ..., V_{n-1}$ vektörlerinin bir kümesini bulmamız gerekir.

Lineer cebirden Gram-Schmidt algoritmasını kullanarak, bu yapılabilir. H ile soldan çarpımın V nin ilk p+1 tane sütununu değiştirmediğine, ancak geri kalan sütunları sıfıra götürdüğüne dikkat ediniz.

Şimdi $z = V'y$ ile n boyutlu normal dağılan z rasgele vektörünü tanımlarız.

V ile her iki yanı soldan çarparak

$$y = Vz = z_0 V_0 + ... + z_{n-1} V_{n-1}$$

elde ederiz. Bunun her iki yanını H ile soldan çarparak

$$\hat{y} = z_0 V_0 + ... + z_p V_p$$

elde ederiz.

$$e = y - \hat{y} = \sum_{j=p+1}^{n-1} z_j V_j$$

olduğu anlaşılır. z normal olduğundan (normal dağıldığından) onun özellikleri onun beklenen değeri ve kovaryans matrisiyle tam olarak belirlenir. z' nin beklenen değeri

$$\eta = E[z] = V'X\beta$$ dır

$$\eta' = \beta' X'V = \beta'(HX)'V$$
$$= \beta' X'HV = \beta' X'[V_0, ..., V_p, 0, ..., 0]$$

olduğuna dikkat ediniz. n boyutlu η' satır vektörünün son n-p-1 tane elemanının bileşeninin tümünün sıfır olduğu, yani

$$\eta_{p+1} = ... = \eta_{n-1} = 0$$

olduğu görülür ve z için kovaryans matrisi

$$kov(z) = E[(z-\eta)(z-\eta)']$$

dir. Şimdi

$$z - \eta = V'y - E[V'y] = V'(y - X\beta) = V'\mu$$

dur . Bu nedenle,

$$kov(z) = E[V'uu'V] = V'E[uu']V = \sigma_u^2 V'V = \sigma_u^2 I$$

dır. Bundan dolayı $z_0, ..., z_{n-1}$ $E[z_i] = \eta_j$ ve $\text{var}[z_j] = \sigma_u^2$ olmak üzere ilişkisiz ve normal dağılımlıdır (bu nedenle bağımsızdır).

Teorem 3.6 u_i lerin normal dağıldıklarını farz ediniz. Bu taktirde $b \; ve \; s^2$ bağımsız olarak normal dağılan rasgele değişkenlerdir b; β ortalamalı ve $\sigma_u^2 (X'X)^{-1}$ kovaryans matrisli çok değişkenli normal dağılıma sahiptir ve

$$(n-p-1)s^2 / \sigma_u^2 \quad \chi^2_{n-p-1}$$

olarak dağılır.

İspat: b; normalların bir lineer kombinasyonu olduğundan o bir çok değişkenli normal dağılıma sahip olmalı b nin ortalaması ve kovaryansı daha evvel ortaya kondu $b \; ve \; s^2$ nin bağımsızlığını saptamak için

$$b = (X'X)^{-1} X'y = (X'X)^{-1} (HX)'Vz$$
$$= (X'X)^{-1} X'H (V_0 z_0 + \ldots + V_{n-1} z_{n-1})$$
$$= (X'X)^{-1} X' (V_0 z_0 + \ldots + V_p z_p)$$

olduğuna, bunun $z_0, \ldots., z_p$ nin bir fonksiyonu olup $z_{p+1}, \ldots, z_{n-1}$' e bağlı olmadığına dikkat ediniz. Öte yandan, $y - \hat{y} = \sum_{j=p+1}^{n-1} z_j V_j$ tahmin edilen hata vektörü, $z_{p+1}, \ldots, z_{n-1}$ in bir fonksiyonudur.

Ancak $z_0, \ldots z_p$ ye bağlı değildir. z_j ler bağımsız olduğundan $b \ ve \ s^2 = \dfrac{1}{n-p-1} \| y - \hat{y} \|^2$ nin bağımsız olduğu anlaşılır. $\| v_j \|^2 = 1$ olmak üzere, v_j vektörleri karşılıklı olarak ortogonal olduklarından,

$$(n-p-1) s^2 / \sigma_u^2 = \| z_{p+1} V_{p+1} + \ldots. + z_{n-1} V_{n-1} \|^2 / \sigma_u^2$$
$$= (z_{p+1}^2 + \ldots.. + z_{n-1}^2) / \sigma_u^2$$

olduğunu görürüz. $\{z_j\}_{j=p+1}^{n-1}$ ler $N(0, \sigma_u^2)$ bağımsız özdeş dağılımına sahip olduğundan

$(n-p-1) s^2 / \sigma_u^2$ nın X_{n-p-1}^2

olarak dağıldığı görülür.

4. Lineer Regresyon Modelinde Şapka Matrisinin Köşegen ve Köşegen-Dışı Elemanları için Sınırlar Üzerine

En küçük kareler yaklaşımında, herhangi bir duyarlılık(hassasiyet) analizi aslında noktaların nasıl gözlendiği, bu nedenle şapka matrisinin elemanları üzerinde nasıl yansıtıldığı ile ilgilidir. Regresyon tanılarında en yaygın kullanılan kavramlar olarak, etkili gözlemler ve aykırı değerler bu niceliklerin büyüklüğü vasıtasıyla elde edilir.

y_i i-yinci gözlenen yanıt, x_i *bir* $px1$ deterministik (rasgele olmayan) vektör, $\beta \in ¡^p$ parametrelerinin bilinmeyen bir $px1$ vektörü ve $\in_i$ ler sıfır ortalamalı ve σ^2 varyanslı ilişkisiz hatalar olmak üzere

$$y_i = x_i'\beta + \in_i \ , \ i = 1,2,....,n \tag{39}$$

genel lineer regresyon modelini gözönüne alınız.

$$y = (y_1,..., y_n)', \in = (\in_1,....,\in_n)' \ ve \ x = (x_1,....,x_n)'$$

yazarak (39) modeli

$$y = X\beta + \in \tag{40}$$

olarak yazılabilir

İnterceptli (sabit terimli) modelde 1 ler sütununu içeren, X matrisine tasarım matrisi denilir. Başından sonuna kadar X in tam ranklı (tam sütun ranklı) matris olduğunu, bu nedenle X'X in tekil olmadığını kabul ederiz . Bu durumda β nın alışılmış en küçük kareler tahmin edicisi

$$\hat{\beta} = (X'X)^{-1}X'y \tag{41}$$

dır.Tepkime değişkeninin alışılmış tahmin edilen değerlerinin nx1 vektörü $\hat{y} = Hy$ dir. Burada nxn tahmin veya şapka matrisi,

$$H = X(X'X)^{-1}X' \tag{42}$$

ile verilir I_n n mertebeli birim matris olmak üzere $V = (I_n - H)\sigma^2$ varyans-kovaryans matrisli, residual (tahmin edilen hata) vektörü , $e = (I_n - H)y$ ile verilir. H matrisi lineer regresyon modelinde önemli bir rol oynar. h_{ij}, H nin (i,j)-yinci elemanını göstersin.

Buradan,

$$hij = x_i'(X'X)^{-1}x_j \quad (i, j) = 1, 2, ..., n \tag{43}$$

dir h_{ij} köşegen elemanına i-yinci veri noktasının temayülü(kaldıraç gücü) de denir ve x_i gözleminin X-uzayındaki noktaların geri kalanında ne kadar uzakta olduğunu ölçer . h_{ii} nın büyük değerlerine sahip herhangi bir nokta etkili (etkileyici) bir gözlem olmaya meyleder. Böyle bir noktaya yüksek-temayül denir. Cook ve Weisberg(1982,s13) h_{ii} nin büyük olması için aşağıdaki şartları işaret eder.

$x_i'x_i$, x_j *vektörünün* $x_j'x_j$ normunun karesine göre büyüktür, yani x_i , veri kümesindeki diğer noktaların ekseriyetinden çok uzaktır, veya

$x_i'x_i$ aslında X'X in küçük bir özdeğerine karşılık gelen bir özvektörün doğrultusunda ve yönündedir.

x_i nin yüksek-temayül olmasına göre h_{ii} nın büyüklüğü için çeşitli ölçüler (kriterler) önerilir.

Öte yandan şapka matrisinin köşegen dışı elemanlarına regresyon analizinde başka bir kriter(ölçüt) olarak bakılabilir. σ^2 sabitini yok sayarak bu elemanlar tahmin edilen hataların herhangi çiftinin kovaryanslarıdır, bu nedenle bağımsızlık varsayımını kontrol etmek için faydalı olabilir. Teorik nokta gözlemlerin birlikte ancak bireysel olarak değil, etkili oldukları durumlar mevcut olabilir.

Hube (1975) genel olarak h_{ij} nin büyük değerlerinin tasarım noktalarının merkez dışı olmasına karşılık geldiğini zikreder.

Hadi(1990) gözlemlerin potansiyel olarak etkili altkümelerini gösterimini teklif etti.

Burada biz interceptli (sabit terimli) ve sabit terimsiz lineer regresyon modellerinde şapka matris elemanlarının bazı çok büyük (uç – aşırı) değerlerine sahip olmak için tasarım matrisi için gerek ve yeter şartları ele alırız.

Sabit terimli lineer modelde şapka matrisinin köşegen dışı elemanları için daha mükemmel bir alt sınır elde ederiz ki bu sabit terimsiz model için olandan 1/n kadar daha kısadır.

Aşağıdaki birinci lemmanın tekrarlanan uygulaması yapılır. Bu lemmanın (a) şıkkı Chipman(1964) ‘ e aittir.

Lemma 4.1 $A\ p-m_1\ (m_1>0)$ ranklı bir nxp matris olsun

(a) m_1xp mertebeli ve tam satır ranklı B,A nın satırlarından lineer bağımsız satırlara sahipse bu taktirde

$$A(A'A+B'B)^{-1}B'=0_{nxm_1}\ ve\ B(A'A+B'B)^{-1}B'=I_m$$

dır

(b) Eğer $m_2xp\ (m_2\leq m_1)$ mertebeli ve 1 ranklı R, ilk r’ satırı $\delta=(1,\delta_2,...,\delta_n)'$ olmak üzere , $R=\delta r'$ biçimde ve r; A nın satırlarından lineer bağımsız ise bu taktirde

$$R(A'A+R'R)^{-1}R'=\frac{\delta\delta'}{\|\delta\|^2}$$

dir.

Lemma 4.2 A ve B nxp matrisler olsun. Bu taktirde rank(A-B)=rank(A)-rank(B) olması için gerek ve yeter şart $\bar{A}$; $A nın\ A\bar{A}A=A$ bağıntısını sağlayan bir genelleştirilmiş tersi olmak üzere $A\bar{A}B=B\bar{A}A=B\bar{A}B=B$ olmasıdır.

Bundan böyle bu çalışmada, i-yinci gözlemin ihmal edilmesini göstermek için bir niceliğe bir altindis olarak yazılan(i) notasyonunu kullanırız. Örneğin $X_{(i)}\ ve\ X_{(ij)}$ sırasıyla bilinen i-yinci ve j-yinci satırlara sahip X matrisidir. $\bar{x}$ vektörü X’in satırlarının ortalamasını gösterir ve J_p 1 lerin bir pxp matrisidir.

4.1 Şapka Matrisinin Köşegen Elemanları için Sınırlar

Bu kısım X gözlem matrisinin alacağı değerler için gerek ve yeter şartlar ile birlikte h_{ii} nin alt ve üst sınırların belirlemeye tahsis edilir. Bu şartlar esas olarak x_i ve $x_{(i)}$ nin bazı özel biçimlerine dayalı olur.

İnterceptsiz ve interceptli iki geneksel tam rank lineer regresyon modelini gözönüne alırız.

Lemma 4.1.1 X nxk 1 ler sütunu olmayan tüm sütun ranklı matris olsun . Bu taktirde

(i) $0 \le h_{ii} \le 1 dir$

(ii) $h_{ii} = 0$ olması için gerek ve yeter şart $x_i = 0$ olmasıdır.

(iii) $h_{ii} = 1$ olması için gerek ve yeter şart $rank(x_{(i)}) = k-1$ olmasıdır.

İspat: H ve $I_n - H$ pozitif-yarı tanımlı (p.y.t) matrisler olduğundan, (i) kısmı hemen ispatlanır.

Benzer şekilde $(X'X)^{-1}$ bir pozitif tanımlı (p.t) matris olduğundan, (ii) kısmı elde edilir. (iii) kısmını doğrulamak için, genelliği yitirmeksizin x_i nin, X in son satırı olduğunu yani $X' = [X'_{(i)} x_i]$ olduğunu farz ediniz. Eğer $h_{ii} = 1$ ise

$$H = \begin{bmatrix} x_{(i)}(X'X)^{-1}X'_{(i)} & 0_{(n-1)x1} \\ 0,_{x(n-1)} & 1 \end{bmatrix} \text{dir. H bir idempotent matris olduğundan} \qquad (44)$$

$x_{(i)}(X'X)^{-1}X'_{(i)}$ de idempotentdir. Bu nedenle

$rank(X_{(i)}) = rank(X_{(i)}(X'X)^{-1}X'_{(i)}) = iz(X_{(i)}(X'X)^{-1}X'_{(i)} = k-1$ dir. Tersine olarak $rank(X_{(i)}) = k-1$ olsun. $rank[X'_{(i)} x_i] = rank(x') = k$ olduğundan x_i ; $x'_{(i)}$ nin satırlarından lineer bağımsızdır. Lemma 4.1 in (a) kısmını kullanarak.

$$x'_i(X'_{(i)}X_{(i)} + x_i x'_i)^{-1} x_i (= h_{ii}) = 1$$

dir ve ispat tamamlanır.

Lemma 4.1.2 Eğer $X_{nx(k+1)}$ tam sütün ranklı matrisi 1 lerden oluşan sütununu ihtiva ederse bu taktirde

(i) $\frac{1}{h} \leq h_{ii} \leq 1 dir.$

(ii) $h_{ii} = \frac{1}{n}$ olması için gerek ve yeter şart $x_i = \overline{x}$ olmasıdır.

(iii) $h_{ii} = 1$ olması için gerek ve yeter şart $rank(x_{(i)}) = k$ olmasıdır.

İspat: Bu durumda $H - \frac{1}{n} J_n$ ve $I_n - H$ nın her ikiside pozitif yarı tanımlı matrislerdir , bu nedenle (i) kısmı gerçeklenir. (ii) yi doğrulamak için interceptli (sabit terimli) modelde

$$(X'X)^{-1}\overline{x} = \frac{1}{n}\begin{bmatrix} 1 \\ 0_{kx1} \end{bmatrix} \tag{45}$$

elde ettiğimize dikkat ediniz yeter şart

$$x_i'(X'X)^{-1}\overline{x} = \overline{x}'(X'X)^{-1}\overline{x} = \frac{1}{n} \tag{46}$$

olduğuna dikkat ederek saptanır.

Tersine olarak eğer $h_{ii} = 1/n$ ise

$$(x_i - \overline{x})'(X'X)^{-1}(x_i - \overline{x}) = 0$$

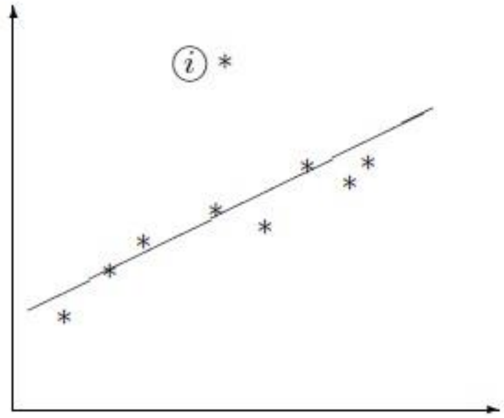
dır ki bu $x_i = \overline{x}$ yi sağlar. Lemma 4.1.1 in (iii) kısmına göre , (iii) kısmı benzer şekilde gerçeklenir.

Örnek 4.1.1 Alışılagelen varsayımla, $y_i = \beta_0 + \beta_1 x_i + \in_i$ basit lineer regresyon modelini gözönüne alınız.

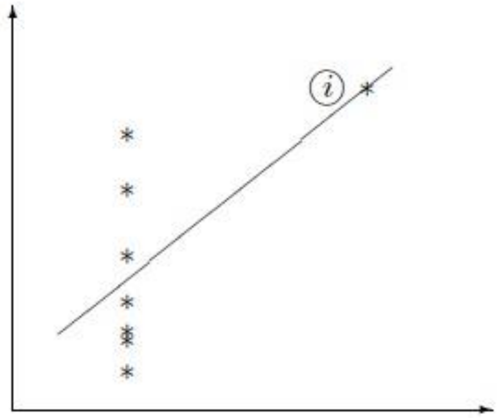
Bu durumda,

$$h_{ii} = \frac{1}{h} + \frac{(x_i - \overline{x})^2}{\sum_{k=1}^{n}(x_k - \overline{x})^2}$$

dir. $x_i = \overline{x}$ nin $h_{ii} = 1/n$ eşitliğini sağladığı açıktır. Şekil 1 ve Şekil 2 durumların iki örneğini gösterir. Şekil 1 de i-yinci gözlem minimum olası h_{ii} değerini verir ve uygun eğim bu gözlemle etkilenmez. Tersine olarak, Şekil 2 h_{ii} için maksimum olası değerli bir örneği gösterir. Bu durumda uygun doğrunun eğimi y_i ile belirlenir ve böyle gözlemi silme (yok sayma) $X'_{(i)}\ X_{(i)}$ yi tekil bir matrise çevirir.



Şekil 1



Şekil 2

Şekil 1: $h_{ii} = \frac{1}{h}$ olduğu interceptli bir basit lineer regresyon modeli

Şekil 2: $h_{ii} = 1$ olduğu interceptli bir basit lineer regresyon modeli

4.2 Şapka Matrisinin Köşegen-Dışı Elemanları için Sınırlar

Bu durumda lineer regresyon modelinde sabit terimli iki durumu kabul ederiz. Aşağıdaki lemmanın (i) kısmı Chatterjee ve Hadi (1988,s.18) de gösterilir.

(Onlar bu lemmanın (i) kısmını belirtme için Professor J. Brain Gray, takdir etti)

Lemma 4.2.1 X_{nxk} 1 ler sütunu olmayan tüm sütun ranklı matris olsun. Bu taktirde,

(i) $-\frac{1}{2} \le h_{ij} \le \frac{1}{2}$ dir

(ii) $h_{ij=-\frac{1}{2}}$ olması için gerek ve yeter şart $x_i = -x_j$ ve $rank(X_{(ij)}) = k-1$ olmasıdır.

(iii) $h_{ij} = \frac{1}{2}$ olması için gerek ve yeter şart $x_i = x_j$ ve $rank(X_{(ij)}) = k-1$ olmasıdır.

İspat: H idempotent olduğundan

$$h_{ii} = \sum_{i=1}^{n} h_{ik}^2 = h_{ii}^2 + h_{ij}^2 + \sum_{k \neq (i,j)} h_{ik}^2$$

elde ederiz ki bu

$h_{ij}^2 = h_{ii}(1-h_{ii}) + \sum_{k \neq (i,j)} h_{ik}^2$ olduğunu gösterir. $0 \leq h_{ii} \leq 1$ olduğundan (i) kısmı

$h_{ii} = h_{jj} = 1/2$ ve her $k(\neq i, j) = 1, 2, ..., n$ için $h_{ik} = h_{jk} = 0$ şartlarıyla elde edilir.

(ii) kısmının yeter şartını doğrulamak için $h_{ij} = -1/2$ olsun. (46) dan $h_{ii} = h_{jj} = 1/2$ elde ederiz. Bu nedenle

$$(x_i + x_j)'(X'X)^{-1}(x_i + x_j) = 0$$

dır ki bunun gerçekleşmesi için yeter şart $x_i = -x_j$ olmasıdır. Yine eğer $X' = [X'_{(i,j)} x_i x_j]$ ise bu taktirde

$$H = \begin{bmatrix} X_{(ij)} \quad (X'X)^{-1} X'_{(ij)} & 0_{(n-2)x2} \\ 0_{2x(n-2)} & \frac{1}{2}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \end{bmatrix} \tag{47}$$

dir. H idempotent olduğundan, (47) bağıntısından $X(ij)(X'X)^{-1}X'_{(ij)}$ nin de idempotent olduğu görülür. Bu nedenle

$$rank(X_{(ij)}) = rank(X_{(ij)}(X'X)^{-1}X'_{(ij)}) = iz(X_{(ij)}(X'X)^{-1}X'_{(ij)}) = k-1 \text{ dir.}$$

Tersine olarak eğer $x_i = x_j$ ve $rank(X_{(ij)}) = k-1$ ise $rank[X'_{(ij)} x_i x_j] = rank(X) = k$ olduğundan x_i nin $x_{(ij)}$ nin satırlarından lineer bağımsız olduğu görülür $y = (1.-1)$ olmak üzere A ve R yerine $X_{(ij)}$ ve $[x_i - x_i]$ koyarak Lemma 4.1 in (b) kısmını uygulama

$$R(A'A + C'C)^{-1}R^1 = \begin{bmatrix} h_{ii} & h_{ij} \\ h_{ij} & h_{jj} \end{bmatrix} = \begin{bmatrix} 1/2 & -1/2 \\ -1/2 & 1/2 \end{bmatrix}$$

olduğunu ortaya koyar. x_j y_i -1 ile çarparak (ii) kısmı benzer şekilde ispatlanır. Böylece ispat tamamlanır.

Aşağıdaki Lemma interceptli modelde h_{ij} nin sınırını verir. Onun üst sınırını interceptsiz modelin durumuna benzer şekilde bulacağız, bununla beraber alt sınır 1/n sabitiyle netleştirilir.

Lemma 4.2.2 $x_{nx(k+1)}$ 1lerin sütununa sahip tam sütun ranklı matris ise bu taktirde,

(i) $\frac{1}{n}-\frac{1}{2}\leq h_{ij}\leq {}^{1}\!/_{2}$ dir.

(ii) $h_{ij}=\frac{1}{n}-\frac{1}{2}$ olması için gerek ve yeter şart

$x_i+x_j=2\overline{x}$ ve $rank(X_{(ij)})=k$ olmasıdır.

(iii) $h_{ij}=\frac{1}{2}$ olması için gerek ve yeter şart $x_i=x_j$ ve $rank(X_{(ij)})=k$ olmasıdır.

İspat: Bu durumda H idempotentdir ve bir geçiş olasılık matrisinin özelliğine sahiptir, yani H1=1 dir. Burada $1=(1,1,...,1)'$ dir Bu nedenle h_{ij} yi

$$\sum_{i=1}^{n}h_{ik}=1 \tag{48}$$

ile birlikte (46) kısıtlaması ile minimumlaştırmalıyız. λ yı bir Lagrange çarpanı olarak kullanarak

$$h_{ij}=1-h_{ii}-\sum_{k\neq(i,j)}h_{ik}+\lambda[h_{ii}(1-h_{ii})-h_{ij}^{2}-\sum_{k\neq i,j}h_{ik}^{2}] \tag{49}$$

yi λ ya ve $k\neq j=1,2,..,n$ elemanlarına göre minimumlaştırırız. $\partial h_{ij}/\partial\lambda=0$ ın (46) yı verdiği ve $\partial h_{ij}/\partial h_{ii}=0$ ın $h_{ii}=\frac{1}{2}-\frac{1}{2\lambda}$ yı verdiği açıktır. Öte yandan $\partial h_{ij}/\partial h_{ik}=0$ koyma $h_{ik}=-\frac{1}{2\lambda}$ sonucunu ortaya koyar (46) de yerine koyma

$$h_{ij}^{2}=\frac{1}{4}(1-\frac{n-1}{\lambda}) \tag{50}$$

yı verir ve böylece (48)

$$h_{ij}=\frac{1}{2}(1+\frac{n-1}{\lambda}) \tag{51}$$

yı verir. (50) ve (51) denklemlerini λ ya göre çözme h_{ij} nin sınırını

$$\frac{1}{n}-\frac{1}{2}\leq h_{ij}\leq 1/2$$

olarak verir. (ii) kısmını ispatlamak için,

$h_{ij}=1/n-1/2$ nin $h_{ii}=h_{jj}=1/n+\frac{1}{2}$ ye götüren $1/n$ ye eşit tüm $h_{ik}, (k\neq i, j)$ ürettiğine dikkat ediniz. Buradan, sadece $x_i+x_j=2\bar{x}$ ise sağlanan

$$(x_i+x_j-2\bar{x})'(\boldsymbol{X}'\boldsymbol{X})^{-1}(x_i+x_j-2\bar{x})=0$$

elde edilir . Ayrıca

$$\boldsymbol{H}-\frac{1}{n}\boldsymbol{J}_n=\begin{bmatrix}\boldsymbol{X}_{(ij)}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'_{(ij)}-\frac{1}{n}\boldsymbol{J}_{n-2} & \boldsymbol{0}_{(n-2)x2}\\ \boldsymbol{0}_{2x(n-2)} & \frac{1}{2}\begin{bmatrix}1 & -1\\ -1 & 1\end{bmatrix}\end{bmatrix} \tag{52}$$

elde ederiz. $\boldsymbol{H}-\frac{1}{h}\boldsymbol{J}_n$ idempotent

olduğundan (52) bağıntısı $\boldsymbol{X}_{(ij)}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'_{(ij)}-\frac{1}{n}\boldsymbol{J}_{n-2}$ nin de idempotent olduğunu ortaya koyar.

Bu nedenle

$$k-1=iz(\boldsymbol{X}_{(ij)}\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'_{(ij)}-\frac{1}{n}\boldsymbol{J}_{n-2})=rank(\boldsymbol{X}_{(ij)}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'_{(ij)}-\frac{1}{n}\boldsymbol{J}_{n-2})$$

dır. Şimdi fark matrisinin son rankının matrislerin karşılık gelen rankının farkına eşit olduğunu göstereceğiz.

$\boldsymbol{A} = \boldsymbol{X}_{(ij)}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'_{(ij)}$ *ve* $\boldsymbol{B} = \frac{1}{n}\boldsymbol{J}_{n-2}$ olsun. A simetrik olduğundan $A^2A' = A$ yı ortaya koyan $\mathbf{A\bar{A} = \bar{A}A}$ elde ederiz. (52) bağıntısını kullanarak ve $(\boldsymbol{H} - \frac{1}{n}\boldsymbol{J})1 = 0$ olduğuna dikkat ederek

$$\mathbf{AB = BA = B}^2 = \frac{(n-2)}{n}\boldsymbol{B} \ ve\ \boldsymbol{A}^2 = \boldsymbol{A} - \frac{2}{n}\boldsymbol{B}$$

elde ederiz. Bu nedenle

$$(\boldsymbol{A} - \frac{2}{n}\boldsymbol{B})\boldsymbol{A}^{-} = \boldsymbol{A} \tag{53}$$

dır. (53) bağıntısını soldan **A** ile çarparak $\mathbf{ABA^- = B}$ buluruz. Benzer şekilde , $\mathbf{A^-BA = B}$ eşitliği $\mathbf{ABA^- = B}$ buluruz. Benzer şekilde , $\mathbf{A^-BA = B}$ eşitliği gerçeklenir. Geriye $\mathbf{BA^-B = B}$ olduğunu göstermek kalır. (53) bağıntısını sağ yandan **B** ile çarparak ve $\mathbf{A^-B}$ nin simetrik olduğuna dikkat ederek.

$$\mathbf{BA^-B} = \frac{n}{2}(\mathbf{AA^-B - AB}) = \frac{n}{2}[\boldsymbol{B} - (\frac{n-2}{n})\boldsymbol{B}] = \boldsymbol{B}$$

elde ederiz. Lemma 4.2.2 yi kullanarak $rank(\boldsymbol{X}_{(ij)}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}_{(ij)} - rank(\boldsymbol{J}_{n-2} = k-1$ ve böylece $rank(\boldsymbol{X}_{(ij)}) = k$ elde ederiz .

Tersine olarak $(n-2)x(k+1)$ mertebeli $x_{(ij)}$ nin k rankına sahip olduğuna ve $x_i + x_j = 2\bar{x}$ olduğunu farz ediniz. Bu taktirde $\bar{x} = \bar{x}_{(ij)}$ *yani* $\boldsymbol{X}_{(ij)}$ nin satır ortalamalarıdır. Bu durumda $h_{ij} = 2/n - h_{ij}$ dir. Şimdi $\boldsymbol{X}_{(i)}$ tam sütun ranklı olduğundan x_j ; $\boldsymbol{X}_{(ij)}$ nin satırlarından lineer bağımsızdır. Lemma 4.1 (a) kısmını kullanarak.

$$
\begin{aligned}
x_i'(\boldsymbol{X}_{(i)}'\boldsymbol{X}_{(i)})^{-1}x_i &= (2\overline{x}-x_j)'(\boldsymbol{X}_{(ij)}'\boldsymbol{X}_{(ij)}+x_jx_j')^{-1}(2\overline{x}-x_j)\\
&= 4\overline{x}_{(ij)}\,(\boldsymbol{X}_{(ij)}'\boldsymbol{X}_{(ij)}+x_jx_j')^{-1}\,\overline{x}_{(ij)}\\
&\quad -\frac{4}{n-2}1'\boldsymbol{X}_{ij)}(\boldsymbol{X}_{ij)}'\boldsymbol{X}_{(ij)}+x_jx_j')^{-1}x_j\\
&\quad +x_j'(\boldsymbol{X}_{ij)}'\boldsymbol{X}_{(ij)}+x_jx_j')^{-1}x_j\\
&= 4\overline{x}_{(ij)}'(\boldsymbol{X}_{ij)}'\boldsymbol{X}_{(ij)}+x_jx_j)^{-1}\overline{x}_{(ij)}+1\\
&= 4\left(\frac{(n-1)\overline{x}_{(i)}x_i}{n}\right)'(\boldsymbol{X}_{(i)}'\boldsymbol{X}_{(i)}{}^{-1}\left(\frac{(n-1)\overline{x}_{(i)}+x_i}{n}\right)+1\\
&= 4[x_i'(\boldsymbol{X}_{(i)}'\boldsymbol{X}_{(i)})^{-1}n+1]+1
\end{aligned}
$$

elde ederiz. Buradan $x_i'(\boldsymbol{X}_{(i)}'\boldsymbol{X}_{(i)})^{-1}x_i=\dfrac{h_{ii}}{1-h_{ii}}=\dfrac{n+2}{n-2}$ dır ki bu $h_{ij}=1/n-1/2$ olduğunu ifade eder.

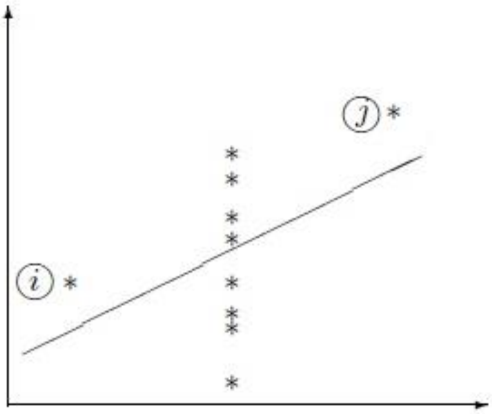
Örnek 4.2.1 Alışıla gelen varsayımlarla, $y_i=\beta_0+\beta_1x_i$ basit lineer regresyon modelini göz önüne alınız.

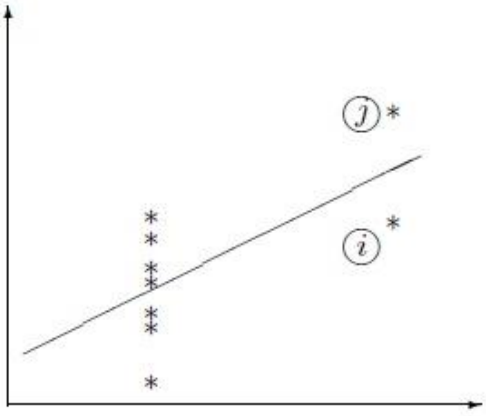
Bu modelde

$$
h_{ij}=\frac{1}{n}+\frac{(x_i-\overline{x})(x_j-\overline{x})}{\sum_{k=1}^{n}(x_k-\overline{x})^2}
$$

dir. Şimdi eğer $x_i=x_j=c\ ve\ x_k=d\neq c\ (her\,k\neq i,\,j\,için)$ ise, bu taktirde $\overline{x}=d+2(c-d)/n$ dir.

$h_{ij}=1/2$ olduğunu göstermek kolaydır. Öte yandan eğer $x_i\neq x_j\ x_k=d\neq x_i,x_j\ (her\,k\neq i,\,j\,için)\,ve\,x_i+x_j=2\overline{x}=2d$ ise bu taktirde $h_{ij}=1/n-1/2$ dir. Şekil 3 ve Şekil 4 h_{ij} onun maksimum ve minimum olası değerlerini verdiği durumda zikredilen durumların iki örneğini gösterir.





Şekil 3 Şekil 4

Şekil3: $h_{ij} = \frac{1}{n} - \frac{1}{2}$ olduğu interceptli bir basit lineer regresyon modeli

Şekil4: $h_{ij} = \frac{1}{n} + \frac{1}{2}$ olduğu olduğu interseptli bir basit lineer regresyon modeli

Örnek 4.2.2 X tasarım matrisinin,

$$x = \begin{bmatrix} 1 & 3 & 8 & 4 \\ 1 & 1 & 6 & 6 \\ 1 & 3 & 5 & 8 \\ 1 & 1 & 1 & 2 \\ 1 & 1 & 14 & 2 \\ 1 & 4 & 9 & 11 \\ 1 & 1 & 7 & 2 \\ 1 & 2 & 6 & 5 \end{bmatrix}$$

olarak verildiği çoklu lineer regresyon modelinin

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \varepsilon_i$$

olduğunu farz ediniz.

$$H=\begin{bmatrix} 0.625 & -0.375 & 0.125 & 0.125 & 0.125 & 0.125 & 0.125 & 0.125 \\ -0.375 & 0.625 & 0.125 & 0.125 & 0.125 & 0.125 & 0.125 & 0.125 \\ 0.125 & 0.125 & 0.298 & 0.133 & -0.161 & 0.335 & -0.003 & 0.148 \\ 0.125 & 0.125 & 0.133 & 0.618 & -0.196 & -0.235 & 0.242 & 0.188 \\ 0.125 & 0.125 & -0.161 & -0.196 & 0.791 & 0.008 & 0.296 & 0.049 \\ 0.125 & 0.125 & 0.335 & -0.235 & 0.008 & 0.658 & -0.123 & 0.106 \\ 0.125 & 0.125 & -0.003 & 0.242 & 0.260 & -0.123 & 0.250 & 0.124 \\ 0.125 & 0.125 & 0.148 & 0.188 & 0.049 & 0.106 & 0.124 & 0.126 \end{bmatrix}$$

dır

$h_{12}=-0.375=\frac{1}{n}-\frac{1}{2}$ olduğu gözlenir ve bu $x_i+x_j=2\bar{x}$ ve herhangi $i\geq 3$ için $x_{i3}=3x_{i1}-1$ olduğundan dolayıdır.

Mahalanobis uzaklığı, ağırlıklaştırılmış kare standartlaştırılmış uzaklık, PRESS (Predicted Residual Error Sum of Squares) (Tahmin Edilen Ölçüm Hatası Kareler Toplamı) v.s gibi çok sayıda istatistiksel ölçü genel olarak h_{ij} leri esas alan teşhisi koyan etkili gözlemlerin literatüründe önerildi. i-yinci noktayı ve (i,j)-yinci noktaları birlikte uzaklaştırma (yok sayma-silme) regresyon teşhislerindeki tesir gücünü (baskıyı) ortaya çıkarmak (sezmek) için yararlı olabilir.Lemma 4.3.1 ve Lemma 4.3.2 den elde edilir.

- $h_{ii}=0$ (veya regresyon sabitine sahip modelde $h_{ii}=1/n$)

Bu durumda i-yinci gözlem potansiyel olarak y_i ve $\bar{y}$ arasındaki büyük uzaklıkla tanınır.Sabit terimli (interceptli) modelde (bkz.Şekil 1) sabit terim hariç, bu nokta bilinmeyen β parametresinin tahmini üzerinde etkiye sahip değildir.

Bu durumda , y_i ; $\hat{y}$ yı belirlemek için minimum etkiye sahiptir.

- $h_{ii}=1$. Böyle noktanın varlığı x in bazı sütunlarının eş doğrultuluğunu giderir,

bu nedenle o büyük olasılıkla etkili bir gözlem olmalıdır. Bu nokta regresyon doğrusunun kendini sürdürmesi (uzatması) için önemlidir. Başka bir deyişle, uygun regresyon doğrusu i-yinci gözlemi yerleştirmek için diğer veri noktalarının arasından geçer. Bu durumda $e_i=0$ olduğunu görürüz .(bkz. Şekil 2)

- $h_{ij} = -\frac{1}{2}$ (veya sabit terimli modelde $h_{ij} = \frac{1}{n} - \frac{1}{2}$)

Bu durum i-yinci ve j-yinci gözlemler arasında bir rekabet olarak açıklanabilir.Lemma 4.2.1 ve lemma 4.2.2 yi kullanarak, eğer bu noktaların herhangi biri uzaklaştırılır

(atılır) ise, bu taktirde diğer noktanın geriye kalan n-1 tane gözlemi esas alan karşılık oluşturulmuş şapka matrisinin köşegen elemanının 1 maksimum değerine sahip olduğu, bu nedenle bir etkileyici gözlem olacağı gösterilebilir.

Bu durumda , $e_i = e_j$, bu nedenle p($\hat{y}_i$, $\hat{y}_j$)=-1 dir.

Bu durum (i,j)- yinci noktalar diğer noktaların büyük çoğunluğunun farklı taraflarında olduğunda gerçekleşir.(bkz.Şekil 3)

- $h_{ij} = \frac{1}{2}$ önceki kısıma zıt, bu durumda i-yinci ve j-yinci gözlemler,diğer

noktaların büyük çoğunluğunun (bulutunun) aynı tarafındadır.Bu gözlemlerin tahmin edilen değerlerinin aynı doğrultuda oldukları yani, p($\hat{y}_i$, $\hat{y}_j$)=1 olduğu gösterilebilir.

(bkz.Şekil 4)

5.Tahmin

5.1 Tam Ranklı Olmayan Modeller

$E(\varepsilon) = 0$ ve $\mathrm{var}(\varepsilon) = \sigma^2 I$ aynı zamanda X bir tam sütun ranka sahip olmamak yani $rank(x) = k < p \leq n$ olmak ve $X'X$ tekil olmak üzere tam ranklı olmayan modeller

$$y = X\beta + \varepsilon$$

biçimindedir (Anonim 2, 2020).

Bu modelde $\hat{B}$ de ki p sayı da parametreler bir tek değildirler.

Örnek 5.1

$$\begin{pmatrix} y_{11} \\ y_{12} \\ y_{13} \\ y_{21} \\ y_{22} \\ y_{23} \end{pmatrix} = \begin{pmatrix} 110 \\ 110 \\ 110 \\ 101 \\ 101 \\ 101 \end{pmatrix} \begin{pmatrix} \mu \\ \tau_1 \\ \tau_2 \end{pmatrix} + \begin{pmatrix} \varepsilon_{11} \\ \varepsilon_{12} \\ \varepsilon_{13} \\ \varepsilon_{21} \\ \varepsilon_{22} \\ \varepsilon_{23} \end{pmatrix}$$

5.2 β ‘nin Tahmini

En küçük kareler yaklaşımı

$$X'X\hat{B} = X'y$$

normal denklemlerini çözmeye götürür.

Teorem 5.2 Eğer X $k < p \leq n$ ranklı nxp matrisi ise bu taktirde $X'X\hat{B} = X'y$ denklemler sistemi tutarlıdır, yani çözüme sahiptir.

İspat: Sistemin tutarlı olması için gerek ve yeter şart ($Ax = c$ lineer denklemler sisteminin tutarlı olması için gerek ve yeter şart $A\bar{A}c = c$ olduğundan)

$$X'X(X'X)^{-}X'y = X'y$$

olmasıdır.

$$X'X(X'X)^{-}X' = X'$$

olduğundan sistem tutarlıdır. Böylece ispat tamamlanmış olur.

Normal denklemler tutarlı olduklarından, bir çözüm yani bir özel çözüm

$$\hat{B} = (X'X)^{-}X'y$$

dir. Bazı genel sonuçlar aşağıda verilmiştir.

1) $\hat{B}$ y nin bir lineer fonksiyonudur.

2) $E(\hat{B}) = (X'X)^{-}X'E(y) = (X'X)^{-}X'XB$ dir ki bu $(X'X)$’ ye bağlıdır ve bu durumda $\hat{B}$ yanlıdır.

3) β tahmin edilebilir değildir. Bir Ay lineer fonksiyonun β’ yı tahmin ettiğini yani $\beta = E(Ay) = E(AX\beta + A\varepsilon) = AX\beta$

$$B = E(Ay) = E(Ax\beta + A\varepsilon) = AxB$$

olduğunu farzediniz. Bu her β için sağlanması gerektiğinden $Ax = Ip$ elde etmeliyiz. Ancak

$rank(Ax) \leq rank(x) < p$ dir. Bu nedenle $Ax \neq Ip$ dir.

Ve böyle bir A yoktur.

Tanım 5.2 Eğer gözlemlerin beklenen değeri $\lambda'\beta$ ’ye eşit olan bir liner kombinasyonu var ise parametrelerin bir $\lambda'\beta$ lineer fonksiyonuna tahmin edilebilir denir, yani

$$E(a'y) = \lambda'\beta$$

olacak şekilde bir a vektörü mevcut ise $\lambda'\beta$ tahmin edilebilirdir.

Teorem 5.2.1 $E(y) = X\beta$ ve X $k < p \le n$ renkli bir nxp matris olmak üzere, $y = X\beta + \varepsilon$ modelinde $\lambda'\beta$ lineer fonksiyonunun tahmin edilebilir olması için gerek ve yeter şart aşağıdaki şartlarda herhangi birinin sağlanmasıdır.

(i) λ^i X' in satırlarının bir lineer kombinasyonudur.

$a'X = \lambda'$

olacak şekilde bir a vektörü mevcuttur.

(ii) λ' $X'X$ 'in satırlarının bir lineer kombinasyonudur veya λ; $X'X$ in sütunlarının bir lineer kombinasyonudur, yani

$r'X'X = \lambda'$ veya $X'Xr = \lambda$

olacak şekilde bir r vektörü mevcuttur.

(iii) $(X'X)^{-}$; $X'X$ in herhangi bir (simetrik) genelleştirilmiş tersi olmak üzere, λ veya λ' $X'X(X'X)^{-}\lambda = \lambda$ veya

$$\lambda'(X'X)^{-}X'X = \lambda'$$

olacak şekildedir.

İspat: Sadece " gerek" kısmını ispatlayacağız.

(i) Eğer $\lambda' = a'x$ olacak şekilde bir a vektörü mevcut ise bu taktirde

$$E(a'y) = a'E(y) = a'X\beta = \lambda'\beta$$

dır

(ii) Eğer $X'Xr = \lambda$ ye bir çözüm varsa bu taktirde tanımlamaya göre $a = Xr$ dir,

$$E(a'y) = E(r'X'y) = r'X'E(y) = r'X'X\beta = \lambda'\beta$$

elde ederiz.

(iii) Eğer $X'X(X'X)^{-}\lambda = \lambda$ ise bu taktirde (ii) kısmına göre $(X'X)\lambda$ $X'Xr = \lambda$ ya bir çözümdür.

Tersine olarak eğer $\lambda'\beta$ tahmin edilebilir ise, bu taktirde

$X'Xr = \lambda$ $r = (X'X)^{-}\lambda$

olarak bulunabilen bir çözüm vektörüne sahiptir. $X'Xr = \lambda$ ' da yerine koyma

$$X'X(X'X)^{-}\lambda = \lambda$$

olduğunu ortaya koyar. Bu ise ispatı tamamlar.

Bu nedenle parametrelerin hangi fonksiyonlarının tahmin edilebilir olduğunu görmek için X’in veya $X'X$ ’in satırlarının lineer kombinasyonlarını inceleyebiliriz.

Örnek 5.2.1

$$X = \begin{pmatrix} 11010 \\ 11001 \\ 10110 \\ 10101 \end{pmatrix} \qquad \beta = \begin{pmatrix} \mu \\ a_1 \\ a_2 \\ \beta_1 \\ \beta_2 \end{pmatrix}$$

yı göz önüne alınız. Örneğin 3.satır – 1. satır’a göre, $\lambda' = (0, -1, 1, 0, 0)$ elde ederiz. Bu nedenle

$$\lambda'\beta = -a_1 + a_2$$

tahmin edilebilirdir.

Üç lineer bağımsız satır, elde etmek için X ‘in satırlarının $a'X$ lineer kombinasyonlarını alırız. X’de ki art arda gelen her bir satırda birinci satırları çıkarma, sonrada matrisin dördüncü satırından ikinci ve üçüncü satırları çıkarma

$$\begin{pmatrix} 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 & 1 \\ 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

matrisini ortaya koyar.

Böylece, üç lineer bağımsız

$$\lambda_1'\beta = \mu + a_1 + \beta_1, \quad \lambda_2'\beta = \beta_2 - \beta_1, \quad \lambda_3'\beta = a_2 - a_1$$

tahmin edilebilir fonksiyonlarını elde ederiz.

5.3 $\lambda'\beta$ nın Tahmin Edicileri

Teorem 5.2.1 (i) ve (ii) den a' ve y' sırasıyla

$\lambda' = a'X$ ve $\lambda' = r'x'x$

bağıntılarını sağlamalı üzere,

$\lambda'\beta$ için $a'y$ ve $r'x'x$

tahmin edicilerini elde ederiz. $\lambda'\beta$ için bir üçüncü tahmin edici $\lambda'\beta$ dır. Burada $\hat{\beta}$; $x'x\hat{\beta} = x'y$ nin bir çözümüdür.

Teorem 5.3 $E(y) = X\beta$ ve X k ranklı $(k \le p \le n)$ nxp matris olmak üzere, $\lambda'\beta$

$y = X\beta + \in$ modelinde β nın bir tahmin edilebilir fonksiyonu olsun,

$\hat{\beta}$; $x'x\hat{\beta} = x'y$ denklemlerine herhangi bir çözüm olsun ve r ; $x'xr = \lambda$ ya herhangi bir çözüm olsun. Bu taktirde; iki $\lambda'\beta$ ve $r'x'y$ tahmin edicisi aşağıdaki özelliklere sahiptir.

(i) $E(\lambda'\hat{\beta}) = E(r'x'y) = \lambda'\beta$ dır.

(ii) Herhangi bir $\hat{\beta}$ veya herhangi bir r için,

$\lambda'\hat{\beta}$; $r'x'y$ ye eşittir.

(iii) $\lambda'\hat{\beta}$ ve $r'x'y$; $\hat{\beta}$ ve r seçimine göre değişmez

İspat (i)

$E(\lambda'\hat{\beta}) = \lambda' E(\hat{\beta}) = \lambda'(X'X)^{-} X'X\beta$ olduğundan, Teorem 5.2.1 (ii) ye göre

$\lambda'(X'X)^{-} X'X = X'$ dır, bu nedenle $E(\lambda'\hat{\beta})$;

$E(\lambda'\hat{\beta}) = \lambda'\beta$ olur. Teorem 5.2.1 (ii) den

$E(r'x'y) = r'x'E(y) = r'x'x\beta = \lambda'\beta$

dır. Teorem 5.2.2 (ii) ye göre, eğer $\lambda'\hat{\beta}$ tahmin edilebilir herhangi r için,

$\lambda' = r'x'x$

dır. $X'X\beta = X'y$ normal denklemlerini r ile önden çarpma

$$r'X'X\hat{\beta} = r'x'y$$

olduğunu ortaya koyar. $r'x'x = \lambda'$ olduğundan

$$\lambda'\hat{\beta} = r'x'y$$

elde ederiz.

(ii) $r'x'y$ nin r nın seçimine göre değişmediğini göstermek için r_1 ve r_2

$$X'Xr_1 = X'Xr_2 = \lambda$$

olacak şekilde olsunlar. Bu taktirde

$$r_1'X'X\hat{\beta} = r_1'X'y \quad \text{ve} \quad r_2'X'X\hat{\beta} = r_2X'y$$

dır. $r_1'X'X = r_2'X'X$ olduğundan, $r_1'X'y = r_2'X'y$ elde ederiz. Herbirinin $\lambda'\hat{\beta}$ yu eşit olduğu açıktır. Böylece ispat tamamlanır (Rencher (2000),s.273-274).

Teorem 5.3.1 X k ranklı $(k \le p \le n)$ nxp ve $kov(y) = \sigma^2 \mathrm{I}$ olmak üzere $\lambda'\hat{\beta}$, $y = X\beta + \in$ modelinde tahmin edilebilir bir fonksiyon olsun. r ; $X'Xr = \lambda$ ye herhangi bir çözüm ve $\hat{\beta}$; $X'X\hat{\beta} = X'y$ ye herhangi bir çözüm olsun. Bu taktirde $\lambda'\hat{\beta}$ nın veya $r'x'y$ nin varyansı aşağıdaki özelliklere sahiptir.

i. $\mathrm{var}(r'X'y) = \sigma^2 r'X'Xr = \sigma^2 r'\lambda$ dır.

ii. $\mathrm{var}(\lambda'\hat{\beta}) = \sigma^2 \lambda'(X'X)^{-}\lambda$ dır.

iii. $\mathrm{var}(\lambda'\hat{\beta} =$ tektir, yani r veya $(X'X^{-})$ nin seçimine göre değişmedir.

İspat(i)

$$\mathrm{var}(r'x'y) = r'X'\,kov(y)xy = \sigma^2 r'x'xr = \sigma^2 r'\lambda$$

dır.

İspat(ii)

$$\mathrm{var}(\lambda'\hat{\beta}) = \lambda' kov(\hat{\beta})\lambda = \sigma^2 \lambda'(X'X)^{-} X'X\,(X'X)^{-}\lambda$$

dır.

$$\lambda'(X'X)^{-} X'X = \lambda'$$

olduğundan böylece (ii) ispatlanır.

İspat(iii)

$$\lambda'(X'X)^{-}\lambda = r'(X'X)(X'X)^{-}(X'X)r = r'x'xr = r'\lambda$$

olduğundan sadece $\lambda'(X'X)^{-}\lambda$ nın $(X'X)$ nın seçimine göre değişmez olduğunu göstermek gerekir G_1 ve G_2 $X'X$ in iki genelleştirilmiş tersi olsun. Bu taktirde

$XG_1X' = XG_2X'$ dür. $a'X = \lambda'$ olacak şekilde her iki yanı a ile çarparak,

$$a'XG_1X'a = a'XG_2X'a$$

veya

$$\lambda_1'G_1\lambda_2 = \lambda_1'G_1\lambda_2$$

elde ederiz. Bu ise ispatı tamamlar.

Teorem 5.3.2 X k ranklı $(k \leq p \leq n)$ nxp matris olmak üzere, eğer $\lambda'\hat{\beta}$ ve $r'X'y$

tahmin edicileri BLUE (Best Linear Unbiased Estimator) (En İyi Lineer Yansız Tahmin Edici)' dir.

İspat: Genelliği yitirmeksizin $a'y = r'bX'y + C'y$, yani $a' = r'X' + c'$ olmak

üzere $\lambda'\hat{\beta}$ nin bir lineer tahmin edicisi $a'y$ ile gösterilmiş olsun.

Burada r' ; $\lambda' = r'X'X$ e bir çözümdür. Yansızlık için

$$\lambda'\beta = E(a'y) = a'X\beta = r'X'X\beta + c'X\beta = (r'X'X + c'X)\beta$$

elde etmeliyiz. Bu nedenle her β için gerçeklenmeli ve bu nedenle

$$\lambda' = r'X'X + c'X$$

elde ederiz.

$\lambda' = r'X'X$ olduğundan , $c'X = a'$

$$\begin{aligned}\text{var}(a'y) &= a'kov(y)a = \sigma^2 a'a \\ &= \sigma^2 (r'X' + c')(rX + c) \\ &= \sigma^2 (r'X'Xr + r'X'c + c'Xr + c'c) \\ &= \sigma^2 (r'X'Xr + c'c)\end{aligned}$$

olduğu anlaşılır. Bu nedenle , $\text{var}(a'y)$ yi minimumlaştırmak için $c = 0$ elde ederiz ve $r'Xy$ BLUE dır. Böylece ispat tamamlanır.

5.4 σ^2 nin Tahmin Edicisi

σ^2 nin bir tahmin edicisi

$$s^2 = \frac{SSE}{n-k}$$

dır. Bu da

$$SSE = (y - X\hat{\beta})'(y - X\hat{\beta}),$$

$\hat{\beta}$; $X'X\hat{\beta} = X'y$ ye herhangi bir çözümdür ve $k = rank(X) dır$

Teorem 5.4.1 $E(y) = X\beta$ ve $kov(y) = \sigma^2 \mathrm{I}$ olmak üzere , $y = X\beta + \in$ tam ranklı olmayan modeli için

1) Denkleminde tanımlanan s^2 için aşağıdaki özelliklere sahibiz

i. $E(s^2) = \sigma^2$ dir

ii. s^2 ; β nın veya $(X'X)^{-}$ genelleştirilmiş tersinin seçimine göre değişmezdir.

İspat(i) $SSE = y'(\mathrm{I} - X(X'X)^{-}X')y$ olduğundan

$$\begin{aligned}E(SSE) &= is\{[\mathrm{I} - X(X'X)^{-}X'](\sigma^2\mathrm{I})\} + \beta'X'[\mathrm{I} - X(X'X)^{-}X']\} \\ &= is\{[\mathrm{I} - X(X'X)^{-}X'](\sigma^2\mathrm{I})]\} \\ &= \sigma^2\{is(\mathrm{I}) - is[X(X'X)^{-}X']\} \\ &= (n-k)\sigma^2\end{aligned}$$

Burada $k = rank(X'X) = rank(X)$

dır.

İspat(ii)

$X(X'X)^{-}X'$; $(X'X)^{-}$ nın seçimine göre değişmezdir.

Böylece ispat tamamlanmış olur.

5.5 Normal Model

Tam ranklı olmayan $y = X\beta + \in$ modeli için, Şimdi

$y : N_n(X\beta, \sigma^2 \mathrm{I})$ veya $\in : N_n(0, \sigma^2 \mathrm{I})$

olduğunu kabul ederiz. Normallik varsayımı ile en çok olabilirlik tahmin edicilerini elde ederiz.

Teorem 5.5.1 X k ranklı $(k \le p \le n)$ nxp matris olmak üzere, eğer y; $N_n(X\beta, \sigma^2 \mathrm{I})$ dağılımına sahip ise, bu taktirde $\hat{\beta}$ ve σ^2 (yanlılık için ayarlanan) en çok olabilirlik tahmin edicileri

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

$$\hat{\sigma}^2 = \frac{1}{n}(y - X\hat{\beta})'(y - X\hat{\beta})$$

dır.

Teorem 5.5.2 X k ranklı $(k \le p \le n)$ nxp matris olmak üzere eğer y; $N_n(X\beta, \sigma^2 \mathrm{I})$ dağılımına sahip ise bu taktirde $\hat{\beta}$ ve s^2 (yanlılık için ayarlanan) en çok olabilirlik tahmin edicileri aşağıdaki özelliklere sahiptir.

(i) $\hat{\beta}$; $Np((X'X)^{-}X'X\beta, \sigma^2(X'X)^{-}X'X(X'X)^{-})$

dağılımına sahiptir.

(ii) $(n-k)s^2/\sigma^2$; $\lambda^2(n-k)$

dağılımına sahiptir.

(iii) $\hat{\beta}$ ve s^2

bağımsızdır.

6. Hipotez Test Etme

Bir hipotez tahmin edilebilir fonksiyonlar cinsinde ifade edilemedikçe, onun test edilebilir olamadığı gösterilebilir (bkz. Searle 1971, ss.193-196)

Bu bizi aşağıdaki tanıma götürür:

Tanım 6.1 Bir hipoteze test edilebilir denecektir, eğer 0 tahmin edilebilir fonksiyonlar cinsinden ifade edilebilir ise

Genel olarak bir test edilebilir hipotez

$$C = c'_{(1)}, c'_{(2)}, \ldots\ldots, c'_{(m)})', c_{(i)}\beta = t_i, i = 1, \ldots., m$$

olmak üzere

$H_0 = C\beta = t$ olarak yazılabilir.

(1) C nin tam satır ranka, yani $rak(C) = m$

(2) $C_{(i)}\beta$ nin her i için tahmin edilebilir olduğunu kabul ederiz.

Teorem 6.1 X k ranklı $(k \le p \le n)$ nxp matris olmak üzere, eğer $y\ N_n(X\beta, \sigma^2 \mathrm{I})$ ise eğer C nin $C\beta$ m sayıda lineer bağımsız tahmin edilebilir fonksiyonun bir kümesi olacak şekilde, $m \le k$ ranklı mxp matris ise, ve eğer $\hat{\beta} = (X'X)^{-} X'y$ ise bu taktirde

(i) $C(X'X)^{-}C'$ tekil değildir ve $(X'X)^{-}$ ye göre değişmezdir.

(ii) $C\hat{\beta} :\ N_m(C\beta, \sigma^2((X'X)^{-}C')$

dir

(iii) $SSH / \sigma^2 = (C\hat{\beta})'[C(X'X)^{-}C']^{-1}(C\hat{\beta}) / \sigma^2 :\ X^2(m, \lambda)$

burada

$$\lambda = (C\beta)'[C(X'X)^{-}C']^{-1}(C\beta) / 2\sigma^2$$

dir

(iv) $SSE/\sigma^2 = y'[\mathrm{I} - X(X'X)^{-}X']y/\sigma^2 : X^2(n-k)$

dır.

(v) SSH ve SSE bağımsızdır.

İspat:

$$C\beta = \begin{pmatrix} c_{(1)}\beta \\ \mathrm{M} \\ c_{(m)}\beta \end{pmatrix}$$

m sayıda lineer bağımsız tahmin edilebilir fonksiyonların bir kümesi olduğundan, bu durumda Teorem 5.2.1 (ii) ye göre her $i = 1,...,m$ için,

$c_{(i)}\ (X'X)^{-}X'X = c_{(i)}$ elde ederiz. Buradan

$C(X'X)^{-}X'X = C$

dır . $rank(AB) \leq rank(A)$ olduğundan

$rank(C) \leq rank(C(X'X)^{-}X' \leq rank(C)$ yazarız. Yani $rank(C(X'X)^{-}X') = m$ dir

$Rank(A) = rank(AA')$ olduğundan

$$\begin{aligned} Rank(C) &= rank(C(X'X)^{-}X') \\ &= rank(C(X'X)^{-}X'X(X'X)^{-}C') \\ &= rank(C(X'X)^{-}C') \end{aligned}$$

elde ederiz. Son eşitlikte, $C(X'X)^{-}X'X = C$ eşitliğini kullanırız. Bu nedenle, $C(X'X)^{-}C'$ tekil değildir.

$C(X'X)^{-}C'$ nun değişmezliği $X(X'X)^{-}X'$ nin değişmezliğinden anlaşılır.

(ii)

$E(C\hat{\beta}) = CE(\hat{\beta}) = C(X'X)^{-}X'X\beta$ dır.

$C(X'X)^{-}X'X = C$

denklemine göre, $E(C\hat{\beta}) = C\beta$ elde ederiz.

$kov(C\hat{\beta}) = Ckov(\hat{\beta}C') = \sigma^2 C(X'X(X'X)^{-}C'$ dir.

$$s^2 = \frac{SSE}{n-k}$$

denklemine göre, $kov(C\hat{\beta}) = \sigma^2 C(X'X)^{-}C'$ elde ederiz. $C\hat{\beta}$; y nin bir lineer fonksiyonu olduğundan, (ii) ispatlanmış olur.

(iii) (ii) şıkkına göre , $kov(C\hat{\beta}) = \sigma^2 C(X'X)^{-}C'$

dır.

$$\sigma^2[C(X'X)^{-}C']^{-1}C(X'X)^{-}C'/\sigma^2 = \mathrm{I}$$

olduğundan sonuç anlaşılır. Böylece ispat tamamlanmış olur.

Teorem 6.2 X ; k ranklı $(k \le p \le n)$ nxp matris olmak üzere $y: N_n(X\beta, \sigma^2\mathrm{I})$ olsun ve $C, C\beta \, ve \, \hat{\beta}$ teorem 6.1 deki gibi tanımlanmış olsun. Bu takdirde, eğer $H_0: C\beta = 0$ doğru ise,

$$F = \frac{SSH/m}{SSE/(n-k)}$$
$$= \frac{(C\hat{\beta})'[C(X'X)^{-}C']^{-1}(C\hat{\beta})/m}{SSE/(n-k)}$$

$F(m, n-k)$ olarak dağılır.

6.1 Yeniden (Tekrar) Parametreleme

Yeniden parametrelemede, X k ranklı $(k \le p \le n)$ nxp matris olmak üzere, $y = X\beta + \in$ tüm ranklı olmayan modelini, Z; k ranklı nxp matris ve $\lambda = \mu\beta$ β nın k sayıdaki bağımsız tahmin edilebilir fonksiyonlarının bir kümesi olmak üzere, $y = \mathrm{Z}\gamma + \in$ tam ranklı modeline dönüştürürüz. Bu nedenle $\mathrm{Z}\gamma = X\beta$ dır ve $X = ZU$ olmak üzere,

$$\mathrm{Z}\gamma = \mathrm{Z}U\beta = X\beta$$

yazarız. UU' tekil olmadığından $(rank(UU') = k)$

$$ZUU' = XU'$$

ve

$$Z = XU'(UU')^{-1}$$

elde ederiz.

Şimdi, Z tam sütun ranklı $(rank(Z) = k)$ bir matristir ve bu durumda burada tam-model için sonuçlar uygulanabilir. Buradan

$$\hat{\gamma} = (Z'Z)^{-1}Z'y \ ,$$

$$s^2 = \frac{1}{n-k}(y - Z\hat{\gamma})'(y - Z\hat{\gamma}) = \frac{SSE}{n-k}$$

(en küçük kareler tahmin edicilerini)

elde ederiz.

$$Z\gamma = X\beta$$

olduğundan,

$$Z\hat{\gamma} = X\hat{\beta} \, ,$$

ve bu nedenle

$$SSE = (y - X\hat{\beta})'(y - X\hat{\beta}) = (y - Z\hat{\gamma})'(y - Z\gamma)$$

dır. Aynı zamanda, herhangi bir $\lambda'\beta$ tahmin edilebilir fonksiyonu için,

$$\lambda'\beta = a'X\beta = a'Z\gamma$$

buradan da

$$\lambda'\beta = a'Z\hat{\gamma}$$

elde ederiz yani, $\lambda'\beta$ nın tahmin edicisi yeniden parametrelemeye göre değişmezdir.

Örnek 6.1.1

$y_{ij} = \mu + \tau_i + \in_{ij}, i = 1,2,\ j = 1,2$ modeli için yeniden (tekrar) parametreleme tekniğini açıklayacağız. Matris formunda(yapısında) model:

$$y = X\beta + \in = \begin{bmatrix} 1\,1\,0 \\ 1\,1\,0 \\ 1\,0\,1 \\ 1\,0\,1 \end{bmatrix} \begin{pmatrix} \mu \\ \tau_1 \\ \tau_2 \end{pmatrix} + \begin{pmatrix} \in_{11} \\ \in_{12} \\ \in_{21} \\ \in_{22} \end{pmatrix}$$

dir. X 2 ranklı olduğundan, 2 tane lineer bağımsız tahmin edilebilir fonksiyon vardır. Bunları birçok şekilde seçebiliriz, bunlardan biri $\mu + \tau_1$ ve $\mu + \tau_2$ dir. Bu nedenle,

$$\gamma = \begin{pmatrix} \gamma_1 \\ \gamma_2 \end{pmatrix} = \begin{pmatrix} \mu + \tau_1 \\ \mu + \tau_2 \end{pmatrix} = \begin{pmatrix} 1\,1\,0 \\ 1\,0\,1 \end{pmatrix} = U\beta$$

$$U = \begin{pmatrix} 1\,0 \\ 1\,0 \\ 0\,1 \\ 0\,1 \end{pmatrix} \text{dır.}$$

$$Z\gamma = X\beta \text{ ve } ZU = X$$

olduğunu doğrulamak kolaydır.

6.2 Yan Şartlar

Yan şartlar koyma tekniği, parametreler tek ve ayrı ayrı tahmin edilebilir olacak şekilde bir tam-rank olmayan model üzerinde (lineer) kısıtlamaları temin eder. Yan şartlar için başka bir kullanım normal denklemleri sadeleştirmek (basitleştirmek) için kısıtlama koymaktır.

Yan şartların, β nın tahmin edilebilir olmayan fonksiyonları olması gerektiğine dikkat ediniz.

X matrisi $k < p$ ranklı nxp matristir. Bu nedenle, X in rankındaki noksanlık $p-k$ dır. Parametrelerin hepsinin tek olması için rankta bu noksanlığı oluşturan yan şartları tanımlamalıyız. Buna bağlı olarak, $T\beta = 0$ yan şartlarını tanımlarız. Burada T; $T\beta$ tahmin edilebilir olmayan fonksiyonların bir kümesi olacak şekilde $p-k$ ranklı bir $(p-k)xp$ matristir.

Teorem 6.2.1

X k ranklı $(k \le p \le n)$ nxp matris olmak üzere , eğer $y = X\beta + \in$ ise ve $T\beta$ tahmin edilebilir olmayan fonksiyonların bir kümesi olacak şekilde eğer T; $p-k$ ranklı bir $(p-k)xp$ ise bu taktirde hem $X'X\hat{\beta} = X'y$ ve $T\hat{\beta} = 0$ sağlayan tek bir $\hat{\beta}$ vektörü vardır.

İspat: İki denklemi birleştirerek

$$\begin{pmatrix} y \\ 0 \end{pmatrix} = \begin{pmatrix} X \\ T \end{pmatrix}\beta + \begin{pmatrix} \in \\ 0 \end{pmatrix}$$

elde ederiz. Bu nedenle $\begin{pmatrix} X \\ T \end{pmatrix}' \begin{pmatrix} X \\ T \end{pmatrix}$ p ranklı (tekil olmayan) pxp matristir, ve

$$\begin{aligned} \hat{\beta} &= [\begin{pmatrix} X \\ T \end{pmatrix}' \begin{pmatrix} X \\ T \end{pmatrix}]^{-1} \begin{pmatrix} X \\ T \end{pmatrix} \begin{pmatrix} y \\ 0 \end{pmatrix} \\ &= (X'X + T'T)^{-1} X'y \end{aligned}$$

elde ederiz.

Örnek 6.2.1 $y_{ij} = \mu + \tau_i + \in_{ij}, i = 1,2, j = 1,2$ modelini gözönüne alınız, $\tau_1 + \tau_2$ fonksiyonunun tahmin edilebilir olmadığı gösterilebilir. $\tau_1 + \tau_2$ yan şartı ; $(0,1,1)\beta = 0$ olarak ifade edilebilir ve $X'X + T'T$

$$X'X + T'T = \begin{pmatrix} 4 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 2 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} (0\ 1\ 1) = \begin{pmatrix} 4 & 2 & 2 \\ 2 & 3 & 1 \\ 2 & 1 & 3 \end{pmatrix}$$

olur. Bu takdirde

$$(X'X+T'T)^{-1}=\frac{1}{4}\begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & 0 \\ -1 & 0 & 2 \end{pmatrix}$$

dir.

$X'y=(y_{..},y_{1.},y_{2.})$ ile $y_{1.}+y_{2.}=y_{..}$ ve $\bar{y}_{i.}=y_{i.}/2$ olduğundan

$$\hat{\beta}=(X'X+T'T)^{-1}X'y$$
$$=\begin{pmatrix} \bar{y}_{..} \\ \bar{y}_{1.}-\bar{y}_{..} \\ \bar{y}_{2.}-\bar{y}_{..} \end{pmatrix}$$

elde ederiz.

7.Tam ve İndirgenmiş Model Testi

β $px1$ ve X le ranklı $(k \le p \le n)$ nxp matris olmak üzere, $y=X\beta+\in$ tam ranklı olmayan modelinde $H_0;\beta_1=...=\beta_q$ hipotezini test etmeyle ilgilendiğimizi farz ediniz. Eğer H_0 test edilebilir ise H_0,

$$H_0:\gamma_1=\begin{pmatrix} \lambda_1'\beta \\ \vdots \\ \lambda_t'\beta \end{pmatrix}=\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

a eşdeğer olacak şekilde $\lambda_1'\beta,...,\lambda_t'\beta$ lineer bağımsız tahmin edilebilir fonksiyonlarının bir kümesini bulabiliriz. k sayıda $\lambda_1'\beta,...,\lambda_k'\beta$ fonksiyonu lineer bağımsız ve tahmin edilebilir, $k=rank(x)$ olacak şekilde,

$$\gamma_2=\begin{pmatrix} \lambda_{t+1}'\beta \\ \vdots \\ \lambda_k'\beta \end{pmatrix}$$

yi bulmak da mümkündür.

$$\gamma = \begin{pmatrix} \gamma_1 \\ \gamma_2 \end{pmatrix}$$

olsun. Şimdi tam-ranklı olmayan modelden

$$y = z\gamma + \in = z\gamma_1 + z_2\gamma_2 + \in$$

tam ranklı modeline yeniden parametreleme yapabiliriz. Burada , $Z = (Z_1, Z_2)$

γ_1 *ve* γ_2 deki elemanların sayısı ile uyumlu olacak şekilde parçalanır.

$y = Z\gamma + \in$ bir tam ranklı model olduğundan,

$H_0 : \gamma_1 = 0$ tam-ranklı modeldeki gibi test edilebilir.

Test Tablo1 de ana hatlarıyla belirtilir. $SS(\gamma_1 | \gamma_2)$ için t, serbestlik derecesinin H_0 rı ifade etmek için gerek lineer bağımsız tahmin edilebilir fonksiyonların sayısı olduğuna dikkat ediniz.

Tablo 1. Tekrar Parametrelenmiş dengeli modellerde

$H_0 : \gamma_1 = 0$ hipotezini test etmek için ANOVA

Değişim Kaynağı	Sd	Kareler Toplamı	F-İstatistiği
γ_2 *için ayarlanan* γ_1 *nedeniyle*	t	$SS(\gamma_1 \| \gamma_2) = \hat{\gamma}'Z'y - \gamma_2{}'Z_2{}'y$	$\frac{SS(\gamma_1 \| \gamma_2)/t}{SSE/(n-k)}$
Hata	$n-k$	$SSE = y'y - \hat{\gamma}'Z'y$	
Toplam	$n-1$	$SST = y'y - n\bar{y}^2$	

Tablo 2. Yeniden parametrelenmiş dengeli modellerde $H_0 : \gamma_1 = 0$ hipotezini test etmek için, ANOVA

Değişim Kaynağı	Sd	Kareler Toplamı	F-İstatistiği
β_2 *için ayarlanmış* β_1 *den dolayı*	t	$SS(\beta_1 \mid \beta_2) = \hat{\beta}' X' y - \hat{\beta}_2' X_2' y$	$\dfrac{SS(\beta_1 \mid \beta_2) / t}{SSE / (n-k)}$
Hata	$n-k$	$SSE = y'y - \hat{\beta}' X' y$	
Toplam	$n-1$	$SST = y'y - n\bar{y}^2$	

Yeniden parametrelenmiş modelde $X\hat{\beta} = Z\hat{\gamma}$ elde ettiğimizden $\hat{\beta}$; $X'X\hat{\beta}\wedge = X'y$ normal denklemine göre herhangi bir çözüm olmak üzere,

$$\hat{\beta}' X' y = \hat{\gamma}' Z' y$$

elde ederiz. Benzer şekilde, $y = Z_2\gamma_2^* + \in^*$ modeline karşılık olarak, $\beta_1 = ... = \beta_q$ koyarak elde edilen bir $y = X_2\beta_2^* + \in^*$ indirgenmiş modelini elde ederiz. Bu taktirde,

$$\hat{\beta}_2^{*\prime} X_2' y = \hat{\gamma}_2^{*\prime} Z_2' y$$

dır.Bu durumda test Tablo 2 deki gibi ifade edilebilir, ki orada $\hat{\beta}' X' y$; $y = X\beta + \in$

tam modelinden elde edilir ve $\hat{\beta}_2' X_2' y$; $H_0 : \beta_1 = ... = \beta_q$ hipoteziyle indirgenmiş olan $y = X_2\beta_2 + \in$ modelinden elde edilir.

8. Bir Yönlü Varyans Analizinin Dengelenmiş Durumu

8.1 Bir – Yönlü Model

Bir yönlü dengelenmiş model aşağıdaki gibi ifade edilebilir.

$$y_{ij} = \mu + \alpha_i + \in_{ij} \quad , \quad i = 1, 2, .., k \quad , \quad j = 1, 2, .., n \tag{54}$$

Eğer $\alpha_1, \alpha_2, ..., \alpha_k$ her biri n deneysel birime uygulanan k sayıda işlemin etkilerini gösterir ise, bu takdirde y_{ij} , i-yince işleme maruz kalan (uğrayan) n birim arasında j-yinci gözlemin tepkimesidir. Model için varsayımlar;

1) Her i, j için , $E(\in_{ij}) = 0$

2) Her i, j için , $\mathrm{var}(\in_{ij}) = \sigma^2$

3) Her $(i, j) \neq (r, s)$ için , $kov(\in_{ij}, \in_{rs}) = 0$

Bazen

4) $\in_{ij}$ nin $N(0, \sigma^2)$ olarak dağıldığı varsayımına da sahibiz.

Bu modelde, i-yinci işlem için ortalamayı göstermek için ekseriyetle μ_i y_i kullanırız, yani (1) varsayımını kullanarak , $E(y_{ij}) = \mu_i$ dir. $\mu_i = \mu + \alpha_i$ dir. Bu nedenle modeli $y_{ij} = \mu_i + \in_{ij}$, $i = 1, 2, .., k$, $j = 1, 2, .., n$ biçiminde yazabiliriz. Modelin bu biçimde $H_0 : \mu_1 = \mu_2 = ... = \mu_k$ hipotezi ile ilgilidir.

8.2 Parametrelerin Tahmini

(54) modelini genel bir k ve n ye genişleterek bir yönlü model matris formunda

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_k \end{pmatrix} = \begin{pmatrix} j & j & 0 & \cdots & 0 \\ j & 0 & j & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ j & 0 & 0 & & j \end{pmatrix} \begin{pmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_k \end{pmatrix} + \begin{pmatrix} \in_1 \\ \in_2 \\ \vdots \\ \in_k \end{pmatrix}$$

olarak yazılabilir. Burada J ve 0 ın herbiri $nx1$ boyutludur, ve y_i ve $\in_i$

$$y_i = \begin{pmatrix} y_{i1} \\ y_{i2} \\ \vdots \\ y_{in} \end{pmatrix}, \ \in = \begin{pmatrix} \in_{i1} \\ \in_{i2} \\ \vdots \\ \in_{in} \end{pmatrix}$$

olarak tanımlanır. Bu nedenle , $X'X\hat{\beta} = X'y$ normal denklemleri

$$\begin{pmatrix} kn & n & n & \cdots & n \\ n & n & 0 & \cdots & 0 \\ n & 0 & n & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ n & 0 & 0 & \cdots & n \end{pmatrix} \begin{pmatrix} \hat{\mu} \\ \hat{\alpha}_1 \\ \hat{\alpha}_2 \\ \vdots \\ \hat{\alpha}_k \end{pmatrix} = \begin{pmatrix} y_{..} \\ y_{1.} \\ y_{2.} \\ \vdots \\ y_{k.} \end{pmatrix}$$

biçimini alır. Burada $y_{..} = \sum_{ij} y_{ij}$ ve $y_{i.} = \sum_j y_{ij}$ dir.

$$(X'X) = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 1/n & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1/n \end{pmatrix} \tag{55}$$

ile verilir. Bu taktirde normal denkleme bir çözüm

$$\hat{\beta} = (X'X)^{-} X'y = \begin{pmatrix} 0 \\ \bar{y}_{1.} \\ \vdots \\ \bar{y}_{k.} \end{pmatrix} \tag{56}$$

olarak elde edilir. (56) daki tahmin ediciler farklı $(X'X)^{-}$ için farklıdırlar, ancak $\lambda'\beta$; $\hat{\beta}$ nın seçimine göre değişmediğinden, onlar tahmin edilebilir fonksiyonların aynı tahminlerini verirler. (56) daki $\hat{\beta}$ yı kullanarak, SSE Yİ aşağıdaki biçimde ifade edebiliriz:

$$\begin{aligned} SSE &= y'(\mathrm{I} - X(X'X)^{-}X')y \\ &= y'y - \hat{\beta}'X'y \\ &= \sum_{i=1}^{k} \sum_{j=1}^{n} y_{ij}^2 - \sum_{i=1}^{k} \bar{y}_{i.} y_{i.} \\ &= \sum_{ij} y_{ij}^2 - \sum_{i} \frac{y_{i.}^2}{n} \end{aligned}$$

Bu nedenle $s^2(E(s^2) = \sigma^2)$

$$s^2 = \frac{1}{k(n-1)}[\sum_{ij} y_{ij}^2 - \sum_i \frac{y_{i.}^2}{n}]$$

ile verilir.

8.3 $H_0 : \mu_1 = \mu_2 = ... = \mu_k$ Hipotezini Test Etme

$\mu_i = \mu + \alpha_i$ ilişkisini kullanarak hipotez

$H_0 : \alpha_1 = \alpha_2 = ... = \alpha_k$ olarak ifade edilebilir ki bu test edilebilir bir hipotezdir, çünkü bu $k-1$sayıda lineer bağımsız tahmin edilebilir kontrasta göre yazılabilir, örneğin, $H_0 : \alpha_1 - \alpha_2 = \alpha_1 - \alpha_3 = ... = \alpha_1 - \alpha_k = 0$

Basitlik için $k = 4$olmak üzere, test yöntemini açıklayalım. Bu durumda, $\beta = (\mu, \alpha_1, \alpha_2, \alpha_3, \alpha_4)'$ ve hipotez $H_0 : \alpha_1 = \alpha_2 = \alpha_3 = \alpha_4$ dür.

Üç lineer bağımsız tahmin edilebilir kontrastı kullanarak hipotez

$$H_0 : \begin{pmatrix} \alpha_1 - \alpha_2 \\ \alpha_1 - \alpha_3 \\ \alpha_1 - \alpha_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

biçiminde yazılabilir ki bu

$$C = \begin{pmatrix} 0 & 1 & -1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 & -1 \end{pmatrix} \tag{57}$$

olmak üzere , $H_0 : C\beta = 0$ olarak da yazılabilir. (57) deki C matrisi tek değildir. Örneğin,

$$C_2 = \begin{pmatrix} 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 & -1 \end{pmatrix}$$ da başka bir C matrisidir.(57)deki C yi ve (55)deki $(X'X)^-$ genelleştirilmiş tersini kullanarak,

$$C(X'X)^- C' = \frac{1}{n}\begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix} \tag{58}$$

elde ederiz.(58) ın tersini bulmak için onu

$$C(X'X)^- C' = \frac{1}{n}\left\{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}\right\} = \frac{1}{n}(\mathrm{I}_3 + J_3 J_3')$$

biçiminde yazarız. Formüle göre

$$(B+CC')^{-1}=B^{-1}-\frac{B^{-1}CC'B^{-1}}{1+C'B^{-1}C}$$ olduğundan J_3 $3x3$ olmak üzere

$$[C(X'X)^{-}C']^{-1}=n(\mathrm{I}_3-\frac{\mathrm{I}_3^{-1}J_3J_3{}'\mathrm{I}_3^{-1}}{1+J_3{}'\mathrm{I}_3^{-1}J_3}=n(\mathrm{I}_3-\frac{1}{4}J_3) \tag{59}$$

elde ederiz. Ayrıca $J_n{}'$ ve $0'$ $1xn$ olmak üzere

$$C(X'X)^{-}X'=\frac{1}{n}\begin{pmatrix} J_n{}' & -J_n{}' & 0' & 0' \\ J_n{}' & 0' & -J_n{}' & 0' \\ J_n{}' & 0' & 0' & -J_n{}' \end{pmatrix}=\frac{1}{n}A \tag{60}$$

yazarız.

(59) ve (60) bağıntılarını kullanarak

$$\begin{aligned}
SSH &= (C\hat{\beta})'[C(X'X)^{-}C']^{-1}C\hat{\beta} \\
&= y'X(X'X)^{-}C'[C(X'X)^{-}C']^{-1}C(X'X)^{-}X'y \\
&= y'[\frac{1}{n}A'n(\mathrm{I}_3-\frac{1}{4}J_3)\frac{1}{n}A]y \\
&= y'[\frac{1}{n}A'A-\frac{1}{4n}A'J_3A]y \\
&= y'[\frac{1}{4n}\begin{pmatrix} 3J_n & -J_n & -J_n & -J_n \\ -J_n & 3J_n & -J_n & -J_n \\ -J_n & -J_n & 3J_n & -J_n \\ -J_n & -J_n & -J_n & 3J_n \end{pmatrix}]y' \\
&= y'[\frac{1}{4n}B]y
\end{aligned}$$

elde ederiz.

$$\frac{1}{4n}B=\frac{1}{n}\begin{pmatrix} 4J_n & 0 & 0 & 0 \\ 0 & 4J_n & 0 & 0 \\ 0 & 0 & 4J_n & 0 \\ 0 & 0 & 0 & 4J_n \end{pmatrix}-\frac{1}{4n}\begin{pmatrix} J_n & J_n & J_n & J_n \\ J_n & J_n & J_n & J_n \\ J_n & J_n & J_n & J_n \\ J_n & J_n & J_n & J_n \end{pmatrix}$$

olduğuna dikkat ediniz. Buradan

$$
\begin{aligned}
SSH &= \frac{1}{n}\sum_{i=1}^{4} y_i{}' J_n y_i - \frac{1}{4n} y' J_{4n} y \\
&= \frac{1}{n}\sum_{i=1}^{4} y_i{}' J_n J_n{}' y_i - \frac{1}{4n} y' J_{4n} J_{4n}{}' y \\
&= \frac{1}{n}\sum_{i=}^{4} y_{i.}{}^2 - \frac{1}{4n} y_{..}^2
\end{aligned}
$$

elde ederiz.

Tablo 3. Üç paketleme yöntemi için askorbik asit (mg/100gr)

	A	B	C
	14.29	20.06	20.04
	19.10	20.64	26.23
	19.09	18.00	22.74
	16.25	19.56	24.04
	15.00	19.47	23.37
	16.61	19.07	25.02
	19.63	18.38	23.27
Toplamlar($y_{i.}$)	120.06	135.18	164.71
Ortalamalar ($\bar{y}_{i.}$)	17.15	19.31	23.53

Örnek 8.3.1 (Askorbik Asit) Dondurulmuş gıdaları paketlemenin üç yöntemi (A-C) Daniel(1974,s.196) tarafından karşılaştırıldı.

Cevap(tepkime) değişkeni askorbik asitdi. Veriler tabloda verilir.

$H_0 : \mu_1 = \mu_2 = \mu_3$ hipotezini test etmek için,

$$\frac{y_{..}^2}{k_n} = \frac{419.95^2}{(3)(7)} = 8398.0001,$$

$$\frac{1}{7}\sum_{i=1}^{3} y_{i.}^2 = \frac{1}{7}[120.06^2 + 135.18^2 + 164.71^2] = 8445.3457$$

$$\sum_{i=1}^{3}\sum_{J=1}^{7} y_{ij}^2 = 8600.3127$$

yi hesaplarız. İşlemler, hata ve genel kareler toplamları bu taktirde

$$\mathrm{SSH} = \frac{1}{7}\sum_{\mathrm{i}=1}^{3} \mathrm{y}_{\mathrm{i.}}^2 - \frac{\mathrm{y}_{..}^2}{21} = 8545.3457 - 8398.0001 = 147.3456,$$

$$SSE = \sum_{ij} y_{ij}^2 - \frac{1}{7}\sum_{i} y_{i.}^2 = 8600.3127 - 8545.3457 = 54.9670$$

$$SST = \sum_{ij} y_{ij}^2 - \frac{y_{..}^2}{21} = 8600.3127 - 8398.0001 = 202.3126 \text{ dır.}$$

Tablo 4. Askorbik Asit verisi için varyans analizi

Kaynak	sd	Kareler Toplamı	Kareler Ort	F
Yöntem	2	147.3456	73.6728	24.1256
Hata	18	54.9670	3.9537	
Toplam	20	202.312		

Bu kareler toplamları, Tabloda gösterildiği gibi, bir F testini elde etmek için kullanılabilir. F=24.1256 için p değeri 8.07×10^{-6} dır. Bu nedenle,

$H_0 : \mu_1 = \mu_2 = \mu_3$ sıfır hipotezini reddederiz.

8.4 Bir Kontrast için Hipotez Testi

Bir yönlü dengelenmiş model için, α_i lere dayanan kontrastların tahmin edilebilir oldukları ve $\sum_i c_i \alpha_i$ ların tahmin edilebilir olmaları için gerek ve yeter şartın $\sum_i c_i = 0$ olması olduğu gösterilir.

$$\sum_{i=1}^{k} c_i \mu_i = \sum_{i=1}^{k} c_i (\mu + \alpha_i) = \mu \sum_{i=1}^{k} c_i + \sum_{i=1}^{k} c_i \alpha_i = \sum_{i=1}^{k} c_i \alpha_i$$

olduğundan $\sum_i c_i \alpha_i$ kontrastı $\sum_i c_i \mu_i$ ye eşdeğerdir.

İlgili hipotez , eğer $\sum_i c_i = 0$ ise, ortalamaların bir karşılaştırmasını temsil eden,

$$H_0 : \sum_{i=1}^{k} c_i \alpha_i = 0 \text{ veya } H_0 : \sum_{i=1}^{k} c_i \mu_i = 0$$

dır. Örneğin,

$$H_0 : 3\mu_1 - \mu_2 - \mu_3 - \mu_4 = 0 ,$$

$\mu_2 , \mu_3 \, ve \, \mu_1$ in ortalaması ile μ_1 i karşılaştıran

$H_0 : \mu_1 = \frac{1}{3}(\mu_2 + \mu_3 + \mu_4)$ olarak yazılabilir.

$c' = (0, c_1, ..., c_k)$ ve $\beta = (\mu_1, \alpha_1, ..., \alpha_4)'$ olmak üzere hipotez $H_0 : C'\beta = 0$ olarak ifade edilebilir. y nin $N_{kn}(X\beta, \sigma^2 \mathrm{I})$ olarak dağıldığını kabul ederek, H_0 aşağıdaki F istatistiğini kullanarak test edilebilir.

$$F = \frac{(C'\hat{\beta})'[C'(X'X)^{-}C]^{-1}C'\hat{\beta}}{SSE/k(n-1)}$$

$$= \frac{(C'\hat{\beta})^2}{s^2 c'(X'X)^{-}C}$$

$$= \frac{(\sum_{i=1}^{k} c_i \bar{y}_{i.})^2}{s^2 \sum_{i=1}^{k} c_i / n}$$

Burada $s^2 = SSE/k(n-1)$ ve $(X'X)^{-}$ ve $\hat{\beta}$ sırasıyla (55) ve (56) bağıntılarıyla verilir.

9. İki – Yönlü Varyans Analizinin Dengelenmiş Durumu

$$y_{ij} = \mu + \alpha_i + \beta_j + \in_{ij} \ , \ i = 1,...,\alpha \ , \ j = 1,2,...,b$$

Toplamsal (etkileşimsiz) modeline sahip olduğumuzu farz ediniz. Bu model dengeli veri ile iki-faktör tasarımıdır.

A faktörü(α için) a sayıda düşeye sahiptir.

B faktörü (β için) b sayıda düşeye sahiptir.

Her bir (i, j) kutucuğunda sadece bir y_{ij} gözlemi vardır.

Matris formunda model

$$y = X\beta + \in$$

olarak yazılabilir. Burada

$$y = (y_{11},..., y_{1b}, y_{21},.., y_{2b},..., y_{a1},..., y_{ab})',$$

$$\beta = (\mu_1, \alpha_1,..., \alpha_a, \beta_1,..., \beta_b),$$

$\in = (\in_{11},..., \in_{1b}, \in_{21},..., \in_{2b},..., \in_{a1},..., \in_{ab})$ ve

$$X=\begin{pmatrix}
\mu & \alpha_1 & \alpha_2 & \cdots & \alpha_a & \beta_1 & \beta_2 & \cdots & \beta_b \\
1 & 1 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\
1 & 1 & 0 & \cdots & 0 & 0 & 1 & \cdots & 0 \\
\vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\
1 & 1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 1 \\
1 & 0 & 1 & \cdots & 0 & 1 & 0 & \cdots & 0 \\
1 & 0 & 1 & \cdots & 0 & 0 & 1 & \cdots & 0 \\
\vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\
1 & 0 & 1 & \cdots & 0 & 0 & 0 & \cdots & 1 \\
\vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\
1 & 0 & 0 & \cdots & 1 & 1 & 0 & & 0 \\
1 & 0 & 0 & \cdots & 1 & 0 & 1 & \cdots & 0 \\
\vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\
1 & 0 & 0 & \cdots & 1 & 0 & 0 & \cdots & 1
\end{pmatrix}$$

$$X'X=\begin{pmatrix}
ab & b & b & \cdots & b & a & a & \cdots & a \\
b & b & 0 & \cdots & 0 & 1 & 1 & \cdots & 1 \\
b & 0 & b & \cdots & 0 & 1 & 1 & \cdots & 1 \\
\vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\
b & 0 & 0 & \cdots & b & 1 & 1 & \cdots & 1 \\
a & 1 & 1 & \cdots & b & 1 & 1 & \cdots & 0 \\
a & 1 & 1 & \cdots & 1 & 0 & a & \cdots & 0 \\
\vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\
a & 1 & 1 & \cdots & 1 & 0 & 0 & \cdots & a
\end{pmatrix}$$

Bir $(X'X)^{-}$ genelleştirilmiş tersini bulmak kolay değildir.

Bunun yerine $X'X\hat{\beta} = X'y$ normal denklemlerini çözmek için iki $\sum_{i=1}^{a}\alpha_i = 0$ ve

$$\sum_{j=1}^{b}\beta_j = 0$$

yan şartını koyabiliriz.

$$X'X\hat{\beta} = \begin{bmatrix} ab\mu + b\sum_i \alpha_i + a \in_j \beta_j \\ b(\mu+\alpha_1) + \sum_j \beta_j \\ \vdots \\ b(\mu+\alpha_a) + \sum_j \beta_j \\ a(\mu+\beta_1) + \sum_i \alpha_i \\ \vdots \\ a(\mu+\beta_b) + \sum_i \alpha_i \end{bmatrix} = \begin{bmatrix} ab\mu \\ b(\mu+\alpha_1) \\ \vdots \\ b(\mu+\alpha_a) \\ a(\mu+\beta_1) \\ \vdots \\ a(\mu+\beta_b) \end{bmatrix}$$

$$X'y = (y_{..}, y_{1.}, \ldots, y_{a.}, y_{.1}, \ldots, y_{.b})'$$

dür. Bu nedenle, çözüm :

$$\hat{\beta} = (\hat{\mu}, \hat{\alpha}_1, \ldots, \hat{\alpha}_2, \hat{\beta}_1, \ldots, \hat{\beta}_b)'$$

dir. Burada

$$\hat{\mu} = y_{..} / (ab) = \bar{y}_{..},$$

$$\hat{\alpha}_i = y_{i.} / b - \hat{\mu} = \bar{y}_{i.} - \bar{y}_{..},\ i = 1,..,a,$$

$$\hat{\beta}_j = y_{ij} / a - \hat{\mu} = \bar{y}_{.j} - \bar{y}_{..}\ ,\ j = 1,\ldots,b$$

dır. Şimdi Tablo 2 deki ana hatları ile özetlemeyi izleyerek , $H_0 : \alpha_1 = \ldots = \alpha_a$ için testi elde etmek için ilerleriz (işlemi sürdürürüz) $H_0 : \alpha_1 = \alpha_2 = \alpha_3$ hipotezi $H_0 : \alpha_1 - \alpha_2 = 0$ ve $\alpha_1 - \alpha_3 = 0$

olarak ifade edilebilir. Bu nedenle, eğer $\alpha_1 - \alpha_2$ ve $\alpha_1 - \alpha_3$ test edilebilir ise, H_0 test edilebilirdir. Gözlemin her bir $E(y_{ij}) = \mu + \alpha_i + \beta_j$ beklenen değeri tahmin edilebilir olduğundan ve $(\mu + \alpha_i + \beta_j)$ nin herhangi bir lineer kombinasyonu tahmin edilebilir olduğundan $\alpha_1 - \alpha_2$ ve $\alpha_1 - \alpha_3$ in her ikisi de tahmin edilebilirdir. $(\alpha_1 - \alpha_2) = (\mu + \alpha_1 + \beta_1) - (\mu + \alpha_2 + \beta_1)$

dir. İlk olarak , $\sum_i y_{i.} = y_{..}$ ve $\sum_j y_{ij} = y_{..}$ olduğundan

$$SS(\mu,\alpha,\beta)=\hat{\beta}X'y=(\hat{\mu},\hat{\alpha}_1,...,\hat{\alpha}_a,\hat{\beta}_1,...,\hat{\beta}_b)\begin{pmatrix} y_{..} \\ y_{1.} \\ \vdots \\ y_{a.} \\ y_{.1} \\ \vdots \\ y_{.b} \end{pmatrix}$$

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

$$=\frac{y_{..}^2}{ab}+(\sum_{i=1}^{a}\frac{y_{i.}^2}{b}-\frac{y_{..}^2}{ab})+(\sum_{j=1}^{b}\frac{y_{.j}^2}{a}-\frac{y_{..}^2}{ab}),$$

yi hesaplarız.

9.1 SSE Hata Kareler Toplamı

$$y'y-\hat{\beta}'X'y=\sum_{ij}y_{ij}^2-\frac{y_{..}^2}{ab}-(\sum_{i=1}^{a}\frac{y_{i.}^2}{b}-\frac{y_{..}^2}{ab})-(\sum_{j=1}^{b}\frac{y_{ij}^2}{a}-\frac{y_{..}^2}{ab})$$

ile verilir. Tablo2 deki $\hat{\beta}_2X_2{}'y$ yi elde etmek için,

$\alpha_1=\alpha_2=\alpha_3=\alpha$ ve $\mu+\alpha$ yerine μ konulmak üzere, $y_{ij}=\mu+\alpha+\beta_j+\in_{ij}=\mu+\beta_j+\in_{ij}$ indirgenmiş modelini kullanırız. İndirgenmiş model için

$X_2{}'X_2\hat{\beta}_2=X_2{}'y$ normal denklemleri

$$ab\hat{\mu}+a\hat{\beta}_1+a\hat{\beta}_2=y_{..}$$
$$a\hat{\mu}+a\hat{\beta}_1=y_{.1}$$
$$a\hat{\mu}+a\hat{\beta}_2=y_{.2}$$

dır. $\hat{\beta}_1+\hat{\beta}_2=0$ yan şartını kullanarak, indirgenmiş normal denklemlere çözüm $\hat{\mu}=\bar{y}_{..},\hat{\beta}_1=\bar{y}_{.1}-\bar{y}_{..},...,\hat{\beta}_b=\bar{y}_{.b}-\bar{y}_{..}$ olarak kolayca elde edilebilir. Bu nedenle,

$$SS(\mu,\beta)=\hat{\beta}_2' X_2' y=\hat{\mu} y_{..}+\hat{\beta}_1 y_{.1}+...+\hat{\beta}_b y_{.b}=\frac{y_{..}^2}{ab}+(\sum_{j=1}^{b}\frac{y_{.j}^2}{a}-\frac{y_{..}^2}{ab}),$$

$SS(\alpha \mid \mu,\beta)=\hat{\beta}_2' X_2' y=\sum_{i=1}^{a}\frac{y_{i.}^2}{b}-\frac{y_{..}^2}{ab}$ elde ederiz. Test Tablo 5 de özetlenir. $H_0:\beta_1=...=\beta_b$ için test istatistiği benzer şekilde elde edilebilir.

Tablo 5 . İki – Yönlü Modeller için ANOVA

Değişim Kaynağı	Sd	Kareler Toplamı	F-İstastistiği
ve β için ayarlanan α nedeniyle	$a-1$	$SS(\alpha \mid \mu,\beta)=\sum_{i=1}^{a}\frac{y_{i.}^2}{b}-\frac{y_{..}^2}{ab}$	$F_1=\frac{SS(\alpha \mid \mu,\beta)/(a-1)}{SSE/(a-1)(b-1)}$
μ ve α için ayarlanan β nedeniyle	b-1	$SS(\beta \mid \mu,\alpha)=\sum_{j=1}^{b}\frac{y_{.j}^2}{a}-\frac{y_{..}^2}{ab}$	$F_2=\frac{SS(\beta \mid \mu,\alpha)/(b-1)}{SSE/(a-1)(b-1)}$
Hata	$(a-1)(b-1)$	$SSE=y'y-\hat{\beta}'X'y$	
Toplam	$ab-1$	$SST=\sum_{ij} y_{ij}^2-\frac{y_{..}^2}{ab}$	

10.SONUÇ VE ÖNERİLER

Günümüzde istatistik ve istatistiksel kavramlar hemen hemen her bilim dalında yaygın bir şekilde kullanılmaktadır. Özellikle istatistikte lineer modeller, istatistiksel tahmin teorisinde önemli bir kullanım alanıdır. Lineer modellerin incelenmesinde, lineer cebirde yer alan matris teorisine büyük miktarda ihtiyaç vardır. Özel yapılara sahip bir çok matris, tahmin teorisinin önemli bir aracıdır. Burada bu matrislerden bazılarının lineer modellerde kullanımına yer verilmiştir. Özellikle lineer modellerin özel bir şekli olan lineer regresyon modellerinde ortaya konan parametre tahminlerine, rasgele vektör ve matrislerin kullanımıyla yaklaşımlar verilmekte ve kontraslar tanıtılarak bir yönlü varyans analizindeki dengelenmiş durumda parametrelerin tahmini ve bazı hipotez tezleri ortaya konulmuştur.

Ayrıca burada en küçük kareler analizinde bireysel ve ortak olarak etkili gözlemleri saptamak için şapka matrisi ve açıklayıcılar matrisi kullanılmıştır.

Bu tezdeki çalışma, bu alanda çalışanlara önemli bir katkı sağlayacaktır.

Bu tezdeki çalışmaları anlamak için lineer cebirdeki matris ve vektör kavramlarının, olasılık teorisindeki bazı dağılımların ve temel istatistiksel kavramların incelenmesi önerilir.

11.KAYNAKLAR

Anonim 1 – Cochran’s Theorem https://en.wikipedia.org/wiki/Cochran's_theorem (Erişim Tarihi: 26.07.2020)

Anonim 2 – Stat 512-Semester (2017-2018)- Analysis of Variance - https://www.stat.purdue.edu/~fmliang/ANOVA.pdf (Erişim Tarihi :14.08.2020)

Chatterjee, S., Hadi A.S(1988). – Regression Analysis by Example, Wiley, *USA* (P.18) (P.100 - 101)

Chipman, John S. (1964) – *Econometrica, Journal of the Econometric Society* – The Treatment of Linear Restrictions in Regression Analysis, Evanston IL. England

Cook R.D, Weisberg S. (1982) - Residuals and Influence in Regression, Springer, Berlin Ginestet, C. E. Boston University - Department of Mathematics & Statistics – Hat Matrix: Properties and Interpretation (Lecture Notes), Boston

Hall D. B., (2000) University of Georgia – Department of Statistic - Theory of Linear Models (Lecture Notes), Georgia

Hurvich,C. Newyork University - Department of Mathematics - Multiple Regression, Key Theory (Lecture Notes), New York

Graybill, F. A. (1976) Theory and Application of the Linear Model, Duxbury Press,USA

Liang, F. Purdue University - Department of Statistics - Analysis of Varience Models (Lecture Notes), West Lafayette, Indiana

Mohammadi, M. Statistical Journal, Feb.2016, International Association for Official Statistics (*IAOS)*, Voorburg (P75-87)

Rencher, A.C. Linear Models in Statistics, Wiley, USA (P. 273-274)

Rencher, A.C., Schaalje, G.B., (2008) Linear Models in Statistics (Second Edition) (P.69-87), Wiley, USA

Searle, Shayle R. (1971) Linear Models, Wiley, USA (P. 55) (P. 61-64) (P.193-196)

Vasilic, A. Texas A&M University, Department of Mathematics, MATH304 Linear Algebra, The Gram-Schmidt Process (Lecure notes), USA

ÖZGEÇMİŞ

Kişisel Bilgiler	
Adı Soyadı	Bilge Bahar KIRIŞOĞLU
Doğum Yeri	Besni / Adıyaman
Doğum Tarihi	31.05.1990
E-Posta Adresi	bilgebahar02@hotmail.com
Eğitim Bilgileri	
Lisans	
Üniversite	Yüzüncü Yıl Üniversitesi
Fakülte	Fen Edebiyat Fakültesi
Bölümü	Matematik
Mezuniyet Yılı	01.01.2013
Yayınlar	
Ulusal Bildiriler	