KARADENİZ TEKNİK ÜNİVERSİTESİ * FEN BİLİMLERİ ENSTİTÜSÜ

MATEMATİK ANABİLİM DALI

MATEMATİK PROGRAMI

EŞİTSİZLİK KISITLAMALI LİNEER MODELLER

Selahattin MADEN

Karadeniz Teknik Üniversitesi Fen Bilimleri Enstitüsünce

" Yüksek Lisans (Matematik) "

Ünvanının Verilmesi İçin Kabul Edilen Tezdir

Tezin Enstitüye Verildiği Tarih: 12 Ocak 1994

Tezin Sözlü Savunma tarihi : 04 Şubat 1994

Tez Danışmanı: Doç. Dr. Cemil YAPAR

Jüri Üyesi : Yrd. Doç. Dr. Abdullah ÇAVUŞ

Jüri Üyesi : Yrd. Doç. Dr. İhsan ÜNVER

Enstitü Müdürü : Prof. Dr. Temel SAVAŞKAN

OCAK-1994

TRABZON

ÖNSÖZ

Bu çalışmanın konusu "Eşitsizlik Kısıtlamalı Lineer Modeller" dir.

Tez konusunun seçiminde ve çalışmanın gerçekleşmesinde yardımlarını esirgemeyen hocam Sayın Doç. Dr. Cemil YAPAR'a teşekkür eder saygılarımı sunarım.

Ayrıca çalışmalarım sırasında bana daima yardımcı olan Sayın Arş. Gör. A. Yaşar ÖZBAN'a teşekkür ederim.

Ocak,1994 Selahattin MADEN

İÇİNDEKİLER

Sayfa No

ÖZET

Bu çalışmada regresyon parametreleri üzerinde eşitlik veya eşitsizlik kısıtlamaları altında lineer modellerde tahmin problemi değişik metodlar kullanarak incelenmiştir. Lineer modellerde varyans-kovaryans matrisinin çeşitli durumlarına göre farklı formlarda tahminler verilip farklı formlara karşılık gelen farklı teoremler ifade ve ispat edildi. Lineer eşitsizlik kısıtlamalı en küçük kareler regresyon problemi için kapalı form çözümler verilmeden önce model dönüştürüldü ve dönüştürmeden önceki ve sonraki kısıtlamalı tahminler verildi.Daha sonra eşitsizlik kısıtlamalı en küçük kareler tahmininin bazı özellikleri verildi. Bunlara ilaveten regresyon katsayıları üzerinde eşitsizlik kısıtlamalı lineer modellerde hipotez testi problemi çalışıldı. Bu nedenle son kısımda Likelihood Oran, Wald ve Kuhn-Tucker çarpanlar testini modelin sapmalarının varyans-kovaryans matrisinin bilinip bilinmemesi durumlarında ele aldık ve bu istatistikler arasında nasıl bir ilişkinin bulunduğu problemi ele alındı.

Çalışmanın temel yapısını [5],[6],[7] ve [10] çalışmaların oluşturmasına rağmen bu çalışmayı tamamlayıcı diğer bazı çalışmalarda incelendi ve tartışıldı.

SUMMARY

In this study, the estimation problem in linear models with equality or inequality constraints on the regression parameters has been investigated by using several different methods. Estimation in different forms with respect to various situations of variance-covariance matrix in linear models is given and different theorems corresponding to different situations have been stated and proved. Before considering the closed-form solutions for the least-squares regression problem with linear inequality constraints, the model is transformed and a relation between pre- and post- transformation constrained estimates is established. Then some properties concerning inequality constrained least-squares estimation are given. In addition to theses the problem of testing statistical hypotheses in linear regression models with inequality constraints on the regression coefficients is studied. Therefore in the last section we consider the Likelihood Ratio test, the Wald test and the Kuhn-Tucker multiplier test when the variance-covariance matrix of the disturbances of the model is known or unknown and examine the relations between these statistics.

Although the references [5],[6],[7] ve [10] constitute the framework of this study, some others have also been studied and discussed as the comlement to the references mentioned above.

GİRİŞ

Bu çalışma da iki veya daha çok değişkenin ortak davranışını inceleyip bilinen değişkenlerden hareket ederek bilinmeyen değişkenler hakkında tahminlerde bulunacağız. Genel lineer modelde bilinmeyen parametre üzerine konulan kısıtlamalar altında tahminler verilip modelin tam ranklı, keyfi ranklı ve varyans-kovaryans matrisinin yapı ve rank durumlarına göre bu tahminlerin durumları incelendi. Bu nedenle çalışmanın birinci bölümünde daha sonraki bölümler için malzeme niteliği taşıyan bazı tanım ve teoremler verilecek ancak fazla detaya girilmeyecektir.

İkinci bölümde lineer modeller tanıtılıp tahmin etme problemi ele alınacaktır. Ancak burada modelin kısıtlamasız model veya eşitlik kısıtlamalı model olması durumları gözönüne alınacaktır. Bunlara ait " en küçük kareler" tahmin yöntemi ve bu tahminin bazı özellikleri incelenecektir.

Son bölümde ise lineer modellerde eşitsizlik kısıtlamaları ve bu kısıtlamalar altında parametre tahminleri verilecektir. Bu bölümde Firoozi[7], Werner[10] ve Escobar ve Skarpness[6] de verilen çalışmalar incelendi. Ayrıca Gouriéroux, Holly ve Monfort (GHM) [5] de verilen test istatistikleri ele alınıp bunlar arasında nasıl bir ilişki bulunduğu problemi verilecektir. Yine bu bölümde varyans-kovaryans matrisinin bilinen veya sonlu sayıda bilinmeyen parametreye bağlı olması durumlarında bu test istatistiklerinin ifadeleri verildi.

Ayrıca çalışmanın akışı içerisinde verilen bilgileri kullanacak nitelikte bazı örnekler ele alınmıştır.

BÖLÜM I

ÖN BİLGİLER

1.1. GEREKLİ TANIMLAR VE TEOREMLER

Tanım 1.1.1.[20]: a) Bir E denemesinin örnek uzayı S olsun. S nin her $s\in S$ elemanına $X(s)=x$ gibi bir sayı karşılık getiren X fonksiyonuna bir rastgele değişken denir.

b) X bir rastgele değişken olsun. X in mümkün değerlerinin sayısı sonlu ya da sayılabilir sonsuz ise bu takdirde X rastgele değişkenine kesikli rastgele değişken adı verilir.

c) X rastgele değişkeninin bir (a,b) ,$a<b$, aralığındaki tüm değerleri aldığını farzedelim. Bu aralık $(-\infty,\infty)$ olabilir. Bu durumda X'e sürekli rastgele değişken denir.

Bir (a,b) aralığındaki noktaların sayılamadığını biliyoruz. Bu bakımdan sürekli bir rastgele değişkenin aldığı değerler X'in ihtimal dağılımını vermek üzere $f(x)$ fonksiyonunu tanımlayabiliriz. Yani $f(x)$ yoğunluk fonksiyonu aşağıdaki iki şartı gerçekleyen bir fonksiyondur:

1) $\forall\, x\in(a,b)$ için $f(x)\geq 0$

ve

2) $$\int_a^b f(x)dx=1$$

dir.

Rastgele değişkenin değer almadığı yerlerde $f(x)=0$ olduğunu kabul edersek, yukarıdaki iki şart şu şekilde de ifade edilebilir:

1') $f(x)\geq 0$, $-\infty\leq x\leq\infty$

2') $\int_{-\infty}^{+\infty} f(x)dx=1$

dır. Ayrıca şunu da kabul edeceğiz.

$$P(c<X<d) = \int_{c}^{d} f(x)dx$$

dir.

Teorem 1.1.1.[1]: A ve B nxn boyutlu kare matrisler olmak üzere

$iz(A+B)=iz(A)+iz(B)$ ve $iz(AB)=iz(BA)$

dir , burada iz(C) ile C matrisinin köşegen elemanlarının toplamı gösterilmektedir.

Tanım 1.1.2.[19]: a) A nxn boyutlu bir kare matris ve I da aynı boyutlu bir birim matris olsun. Bu takdirde $|A-uI|=0$ bağıntısını sağlayan $u_1,u_2,\ldots,u_n$ sayılarına A matrisinin öz değerleri veya karekteristik değerleri denir. $|A-uI|=0$ deklemine de A matrisinin karekteristik denklemi denir.

b) A nxn boyutlu bir matris ve u da A'nın bir öz değeri olsun. Eğer $X\neq 0$ vektörü $AX = uX$ koşulunu sağlayan nx1 boyutlu bir vektör ise bu takdirde X vektörüne A matrisinin u öz değerine karşılık gelen öz vektörü denir.

c) A nxn boyutlu matrisinin karekteristik polinomu,P_A,

$$P_A(t)=|tI-A|$$

ile tanımlanır.

Teorem 1.1.2.[19]: A bir K cismi üzerindeki nxn boyutlu bir matris olsun. Bir $u\in K$ nın A matrisinin öz değeri olması için gerek ve yeter şart u nun A nın karekteristik polinomunun bir kökü olmasıdır.

Teorem 1.1.3.[19]: V kompleks sayılar üzerinde sonlu (≥ 1) boyutlu bir vektör uzayı ve $A:V \longrightarrow V$ bir lineer dönüşüm olsun. P de A nın bir karekteristik polinomu olsun. Bu takdirde P(A)=0 dır.

Bu teorem Hamilton–Cayley Teoremi olarak bilinir.

Sonuç 1.1.1.[19]: A nxn boyutlu bir matris ve P de A matrisinin karekteristik polinomu olsun. Bu takdirde P(A)=0 dır.

Tanım 1.1.3[4]: Y, elemanları y_i ler olan nx1 boyutlu bir vektör ve A da elemanları a_{ij} ler olan nxn boyutlu bir simetrik matris olsun. Bu takdirde

$$Q(Y) = Y'A\,Y = \sum_{i=1}^{n} \sum_{j=1}^{n} y_i y_j a_{ij}$$

ifadesine y_i elemanlarının bir kuadratik formu ve A matrisine de kuadratik form matrisi denir.

Tanım 1.1.4.[4]: Eğer sıfırdan farklı her Y vektörü için $Y'A\,Y>0$ ise $Y'A\,Y$ kuadratik formuna pozitif definit kuadratik form ve A matrisine de bir (simetrik) pozitif definit matris denir. Şayet sıfırdan farklı her Y vektörü için $Y'A\,Y\geq 0$ ise $Y'A\,Y$ kuadratik formuna pozitif semi–definit (non–negatif definit) kuadratik form ve A ya da pozitif semi–definit (non–negatif definit) matris adı verilir.

Teorem 1.1.4.[2]:(Genişletilmiş Cauchy–Schwarz Eşitsizliği)
b ve d nx1 boyutlu herhangi iki vektör ve B nxn boyutlu pozitif definit bir matris olmak üzere

$$(b'd)^2 \leq (b'Bb)(d'B^{-1}d)$$

dir. Eşitlik durumu c bir sabit olmak üzere, ancak $b=cB^{-1}d$ olması ile mümkündür.

Tanım 1.1.5.[2]: nxm boyutlu bir A matrisi $r(A)=r$ rankına sahip olsun. Aşağıdaki dört şartı sağlayan bir A matrisi mevcut ise $\bar{A}$ ya A nın moore–penrose tipi genelleştirilmiş inversi denir.

1) $A\bar{A}A=A$

2) $\bar{A}A\bar{A}=\bar{A}$

3) $(A\bar{A})'=A\bar{A}$

4) $(\bar{A}A)'=\bar{A}A$

A nın singüler olmaması durumunda normal inversi olan A^{-1} matrisi özel bir genelleştirilmiş inverstir.

Teorem 1.1.5.[2]: a) nxm boyutlu her matrisin bir tek genelleştirilmiş inversi vardır.

b) Herhangi bir A matrisi için $(\bar{A})'=\overline{(A')}$ dir.

c) B ve C sırasıyla nxr ve rxn boyutlarına sahip ve $r(B)=r(C)=r$ $(r>0)$ olsun. Bu takdirde $\overline{BC}=\bar{C}\bar{B}$ dir.

Lemma 1.1.1.[13]: $Ax=g$ lineer denklem sistemini gözönüne alalım, burada g bilinenlerin nx1 boyutlu bir vektörü, A nxm boyutlu bilinen bir matris ve x mx1 boyutlu bilinmeyenlerin bir vektörüdür. Bu durumda denklem sisteminin tutarlı olması için gerek ve yeter şart $A\bar{A}g=g$ olmasıdır.

Lemma 1.1.2.[13]: $Ax = g$ denklem sistemi tutarlı ise genel çözüm, w uygun boyutlu rastgele bir vektör olmak üzere

$$x=\bar{A}g+(I-\bar{A}A)w$$

dır.

Lemma 1.1.3.[13]: A nxn boyutlu simetrik ve idempotent bir matris ise A nın genelleştirilmiş inversi kendisidir.

Lemma 1.1.4.[1]: A simetrik idempotent bir matris ise $r(A)=r$ olması için gerek ve yeter şart A nın r tane bir ve n-r tane sıfır özdeğerlerine sahip olmasıdır. Simetrik idempotent bir matris projeksiyon (izdüşüm) matrisi olarak adlandırılır.

Sonuç 1.1.2.[1]: P bir izdüşüm matrisi ise $r(P)=iz(P)$ dir. Ayrıca izdüşüm matrisleri pozitif semi-definittir.

Lemma 1.1.5.[1]: A pozitif definittir $\Longleftrightarrow$ $A=RR'$ olacak şekilde nonsingüler bir R matrisi mevcuttur.

Teorem 1.1.6.[1]: (Cholesky Ayrışımı) A pozitif definit bir matris ise $A=U'U$ olacak şekilde pozitif diagonal (köşegen) elemanlara sahip bir tek üst üçgensel U matrisi mevcuttur.

Lemma 1.1.6.[1]: A nxn boyutlu r ranklı pozitif semi-definit bir matris ise bu takdirde $S'A\,S=I_r$ olacak şekilde nxr boyutlu ve r ranklı bir S matrisi mevcuttur.

Tanım 1.1.6.[1]: A ve D simetrik ve tüm inversler mevcut olmak üzere

$$\begin{bmatrix} A & B \\ B' & D \end{bmatrix}^{-1} = \begin{bmatrix} A^{-1}+FE^{-1}F' & -FE^{-1} \\ -E^{-1}F' & E^{-1} \end{bmatrix}$$

dır. Burada $E=D-B'A^{-1}B$ ve $F=A^{-1}B$ dir.

Tanım 1.1.7.[1]: Z_{ij} $(i=1,2,\ldots,m,\ j=1,2,\ldots,n)$ $E[Z_{ij}]$ beklenen değerlerine sahip bir rastgele değişkenler kümesi olsun. Rastgele değişkenleri ve onların beklenen değerlerini matris formunda ifade ederek $Z=[(Z_{ij})]$ nin beklenen değeri $E[Z]=[(E[Z_{ij}])]$ ile tanımlanır. özellikle m=n=1 olduğunda $E[Z]=E[Z_{11}]$ dir.

Teorem 1.1.7.[1]: $A=[(a_{ij})]$, $B=[(b_{ij})]$ ve $C=[(c_{ij})]$ sırasıyla lxm,nxp ve lxp boyutlu matrisler olmak üzere

$$E[AZB+C]=AE[Z]B+C$$

dir.

Sonuç 1.1.3.[1]: A ve B sabitlerden oluşan mxn boyutlu matrisler X ve Y ise nx1 boyutlu rastgele vektörler olmak üzere

$$E[AX+BY]=AE[X]+BE[Y]$$

dir.

Benzer şekilde vektörler için kovaryans ve varyans notasyonları genelleştirilebilir. X ve Y sırasıyla mx1 ve nx1 boyutlu rastgele vektörler ise bu takdirde genelleştirilmiş kovaryans ve varyans operatöleri C ve D aşağıdaki gibi tanımlanır.

Tanım 1.1.8.[1]: X ve Y rastgele vektörleri için,

$$C[X,Y]=[(cov[X_i,Y_j])],$$

$$D[Y]=[(cov[Y_i,Y_j])]$$

dir.

Teorem 1.1.8.[1]: X ve Y sırasıyla mx1 ve nx1 boyutlu rastgele vektörler A ve B ise sırasıyla lxm ve pxn boyutlu bilinen matrisler olsun. Bu takdirde

i) $C[X,Y]=E[(X-E[X])(Y-E[Y])']$

ii) $C[AX,BY]=AC[X,Y]B'$

iii) $C[AX,Y]=AC[X,Y]$ ve $C[X,BY]=C[X,Y]B'$

iv) $D[X]=E[XX']-(E[X])(E[X])'$

v) $D[AX]=C[AX,AX]=AC[X,X]A'=AD[X]A'$

dir.

Teorem 1.1.9.[1]: X, Y, U ve V nx1 boyutlu (farklı olmaları gerekmez.) rastgele vektörler olsun. Her a,b,c ve d (sıfır dahil) reel sayıları için

i) $C[aX+bY,cU+dV]=acC[X,U]+adC[X,V]+bcC[Y,U]+bdC[Y,V]$

ii) $D[aX+bY]=C[aX+bY,aX+bY]$

$$=a^2D[X]+2abC[X,Y]+b^2D[Y]$$

bağıntıları gerçeklenir.

Teorem 1.1.10.[1]: X nx1 boyutlu bir rastgele vektör ve A nxn boyutlu bir simetrik matris olsun. Eğer $E[X]=\theta$ ve $D[X]=\Sigma=[(\sigma_{ij})]$ ise bu takdirde

$$E[X'A\,X]=iz[A\Sigma]+\theta'A\,\theta$$

dır.

Sonuç 1.1.4.[1]: Y=X-b konularak

$$E[(X-b)'A\,(X-b)]=iz[A\Sigma]+(\theta-b)'A\,(\theta-b)$$

elde edilir.

Tanım 1.1.9.[1]: X ve Y rastgele değişkenlerine (istatistiksel) bağımsızdır denir, şayet f(x,y) birleşik yoğunluk fonksiyonu

$$f(x,y)=f_1(x)f_2(y)$$

formunda parçalanabilirse, burada f_1 ve f_2 sırasıyla X ve Y nin marjinal yoğunluk fonksiyonlarıdır.

Teorem 1.1.11.[1]: Eğer X ve Y bağımsız ise $\alpha(X)$ ve $\beta(Y)$ de bağımsızdır. Burada $\alpha(X)$ ve $\beta(Y)$ fonksiyonları rastgele değişkenlerdir, yani α ve β ölçülebilir fonksiyonlardır.

Teorem 1.1.12.[1]: $X_1, X_2, \ldots, X_n$ $\theta_1, \theta_2, \ldots, \theta_n$ ortalamalı, μ_2 ortak varyanslı, üçüncü ve dördüncü ortak momentleri sırasıyla μ_3 ve μ_4 (yani $\mu_r = E[(X_i-\theta_i)^r]$) olan n bağımsız değişken olsun. Eğer A nxn boyutlu herhangi bir simetrik matris ve a da A nın diagonal elemanlarının sütun vektörü ise bu takdirde

$$Var[X'A\,X] = (\mu_4 - 3\mu_2^2)a'a + 2\mu_2^2 iz(A^2) + 4\mu_2\theta'A^2\theta + 4\mu_3\theta'Aa$$

dir.

Sonuç1.1.5.[1]: $X_1, X_2, \ldots, X_n$ normal dağılmış ise bu takdirde $\mu_3=0$, $\mu_4=3\mu_2^2$ olup

$$Var[X'A\,X] = 2\mu_2^2 iz(A^2) + 4\mu_2\theta'A^2\theta$$

dır. Ayrıca $\theta=0$ ve $\mu_2=\sigma^2$ ise bu takdirde

$$Var[X'A\,X] = 2\sigma^4 iz(A^2)$$

dır.

Teorem 1.1.13.[12]: A nxn boyutlu simetrik bir matris olsun. Bu takdirde A matrisinin spektral (tayfi) ayrışımı olarak bilinen

$$A = u_1e_1e_1' + u_2e_2e_2' + \ldots + u_ne_ne_n'$$

özelliği gerçeklenir. Burada $u_1, u_2, \ldots, u_n$ A nın özdeğerleri ve $e_1, e_2, \ldots, e_n$ de A nın bu özdeğerlerine ilişkin normalleştirilmiş özvektörleridir.

Uyarı: Bir X vektörü $X'/(X'X)^{\frac{1}{2}}$ şeklinde yazılarak birim uzunluklu yapılabilir. Böyle vektörlere normalleştirilmiş vektörler denir.

Tanım 1.1.10.[12]: A spektral ayrışımı $\sum_{i=1}^{n} u_ie_ie_i'$ olan nxn boyutlu pozitif definit bir matris olsun. Bu takdirde

$$P = [e_1, e_2, \ldots, e_n]; \quad u_i > 0 \text{ ve}$$

$$T = \begin{bmatrix} u_1 & 0 & \dots & 0 \\ 0 & u_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & u_n \end{bmatrix}$$

olmak üzere $\sum_{i=1}^{n}(u_i)^{\frac{1}{2}}e_ie_i' = PT^{\frac{1}{2}}P'$ matrisine A matrisinin karekökü denir ve $A^{\frac{1}{2}}$ ile gösterilir.

Sonuç 1.1.6.[12]: $A^{\frac{1}{2}}$ karekök matrisi aşağıdaki özelliklere sahiptir.

i) $(A^{\frac{1}{2}})' = A^{\frac{1}{2}}$

ii) $(A^{\frac{1}{2}}A^{\frac{1}{2}}) = A$

iii) $(A^{\frac{1}{2}})^{-1} = \sum_{i=1}^{n}(1/\sqrt{u_i})e_ie'_i = PT^{-\frac{1}{2}}P'$

dir, burada $T^{-\frac{1}{2}}$ i-yinci köşegen elemanı $1/\sqrt{u_i}$ olan bir köşegen matrisi göstermektedir.

iv) $A^{\frac{1}{2}}A^{-\frac{1}{2}} = A^{-\frac{1}{2}}A^{\frac{1}{2}} = I$, $A^{-\frac{1}{2}}A^{-\frac{1}{2}} = A^{-1}$

dir, burada $A^{-\frac{1}{2}} = (A^{\frac{1}{2}})^{-1}$ dir.

Lemma 1.1.7.[17]: A matrisi

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

formunda parçalanmış olsun.

i) A ve A_{11} non-singüler ise bu takdirde

$$M=(A_{22}-A_{21}A_{11}^{-1}A_{12})^{-1}$$

olmak üzere

$$A^{-1} = \begin{bmatrix} A_{11}^{-1}+A_{11}^{-1}A_{12}MA_{21}A_{11}^{-1} & -A_{11}^{-1}A_{12}M \\ -MA_{21}A_{11}^{-1} & M \end{bmatrix}$$

dir.

ii)Eğer A ve A_{22} non-singüler ise bu takdirde

$$N=(A_{11}-A_{12}A_{22}^{-1}A_{21})^{-1}$$

olmak üzere

$$A^{-1} = \begin{bmatrix} N & -NA_{12}A_{22}^{-1} \\ -A_{22}^{-1}N & A_{22}^{-1}+A_{22}^{-1}A_{21}NA_{12}A_{22}^{-1} \end{bmatrix}$$

dir.

Tanım 1.1.10.[18]: $x,y\in R^n$ olmak üzere $[x,y]=\{\ ty+(1-t)x:\ 0\leq t\leq 1\ \}$ uç noktaları x ve y olan doğrudur. $K\subset R^n$ olsun. K nın herhangi iki noktasını birleştiren doğru parçası K ya ait ise, yani $x,y\in K$ herhangi iki nokta olmak üzere $[x,y]\subset K$ ise K ya "konveks küme" denir.

Örnek 1.[18]: R nin boştan farklı konveks altkümeleri aralıklardır.R^n deki her açık top konvekstir. Herhangi bir kapalı yarı-uzay konvekstir. Ayrıca

$$K=\{x\in R^n \mid \|x\|\leq 1\}$$

kümesi konvekstir.

Tanım 1.1.12.[18]: f reel değerli bir foksiyon ve K da f nin tanım kümesinin bir konveks altkümesi olsun. Eğer her $x,y\in K$ ve $t\in[0,1]$ için

$$f(tx+(1-t)y)\leq tf(x)+(1-t)f(y)$$

ise " f fonksiyonu K üzerinde konvekstir " denir, x y den farklı ve $0<t<1$ olmak üzere yukarıdaki kesin eşitsizlik sağlanıyorsa bu takdide f ye K üzerinde kesin olarak konvekstir denir.

Tanım 1.1.13.[1]: R^n n-boyutlu Euklidean uzayı olmak üzere $\Omega \subset R^n$ olsun. Bu takdirde her $y \in R^n$ vektörü $u \in \Omega$ ve $v \in \Omega^{\perp}$ olmak üzere $y=u+v$ formunda tektürlü olarak ifade edilebilir, buna y vektörünün ortogonal (dik) ayrışımı denir.

Sonuç 1.1.7.[1]: a) Eğer $u=P_\Omega y$ ise P_Ω tektir.

b) P_Ω matrisi $P_\Omega=TT'$ formunda ifade edilebilir, burada T nin sütunları Ω için bir ortonormal tabandır.

c) P_Ω simetrik ve idempotenttir.

Sonuç 1.1.8.[1]: a) $\Omega = R[X]$ ise $P_\Omega = X(X'X)^{-}X'$ dir. Şayet X 'in sütunları lineer bağımsız ise $P_\Omega = X(X'X)^{-1}X'$ dir.

b) $R[P_\Omega]=\Omega$ dir.

c) I_n-P_Ω $\Omega^{\perp}$ üzerindeki dik izdüşümü göstermektedir.

Lemma 1.1.8.[1]: $\mathcal{N}[C]$ C matrisinin sıfır uzayı (çekirdeği) ise bu takdirde $\mathcal{N}[C] =\{R[C']\}^{\perp}$ dir. Ayrıca

$$(\Omega_1 \cap \Omega_2)^{\perp} = \Omega_1^{\perp}+\Omega_2^{\perp}$$

dir.

Teorem 1.1.14.[1]: $W \subset \Omega$ ise bu takdirde

$$P_\Omega P_W = P_W P_\Omega = P_W$$

ve

$$P_\Omega - P_W = P_{W^{\perp} \cap \Omega}$$

dir.

Sonuç 1.1.9.[1]: Eğer A_1 $W = \mathcal{N}[A_1]\cap\Omega$ olacak şekilde bir matris ise bu takdirde

$$W^{\perp}\cap\Omega = R[P_{\Omega}A_1']$$

dir.

Sonuç 1.1.10.[1]: A_1 qxn boyutlu q ranklı bir matris ise $r(P_{\Omega}A_1')=q$ olması için gerek ve yeter şart $R[A_1']\cap\Omega^{\perp} = 0$ olmasıdır.

1.2. ÖNEMLİ BAZI DAĞILIMLAR VE BAZI ÖZELLİKLERİ

Tanım 1.2.1.[1]:

$$f(y)=(2\pi\sigma^2)^{-\frac{1}{2}}e^{-\frac{1}{2}(y-\theta)^2/\sigma^2} \qquad -\infty<y<\infty$$

$$=(2\pi\alpha)^{-\frac{1}{2}}e^{-\frac{1}{2}(y-\theta)'\alpha^{-1}(y-\theta)} \qquad \alpha=\sigma^2>0$$

tekli normal yoğunluk fonksiyonuna benzer olarak $-\infty<y_i<\infty,(i=1,2,\ldots,n)$ ve Σ nxn boyutlu pozitif bir matris olmak üzere

$$f(y_1,y_2,\ldots,y_n)=k^{-1}e^{-\frac{1}{2}(y-\theta)'\Sigma^{-1}(y-\theta)} \qquad (1.2.1)$$

çoklu yoğunluk fonksiyonunu tanımlarız.

Teorem 1.2.1.[1]: Eğer $Y=(Y_1,Y_2,\ldots,Y_n)'$ (1.2.1) ihtimal yoğunluk fonksiyonuna sahip rastgele değişkenlerin bir vektörü ise

i) $k=(2\pi)^{\frac{1}{2}n}|\Sigma|^{\frac{1}{2}}$

ii) $E[Y]=\theta$ ve $D[Y]=\Sigma$

iii) $Q=(Y-\theta)'\Sigma^{-1}(Y-\theta)\sim\chi^2_n$

dir.

Sonuç 1.2.1.[1]: a) Eğer $Y \sim N_n(\theta,\Sigma)$ ise $Y-\theta \sim N_n(0,\Sigma)$ dir.

b) C pxn boyutlu ve p ranklı bir matris ise $CY \sim N_p(C\theta, C\Sigma C')$ dır.

Teorem 1.2.2.[1]: Y nin bir çoklu normal dağılıma sahip olması için gerek ve yeter şart sıfırdan farklı her a reel değerli vektörü için a'Y nin tekli normal dağılmış olmasıdır.

Teorem 1.2.3.[1]: $Y \sim N_n(\theta,\Sigma)$ ve $Y^{(1)}$ p elemanlı (p<n) olmak üzere Y nin

$$Y = \begin{bmatrix} Y^{(1)} \\ \\ Y^{(2)} \end{bmatrix}$$

formunda parçalandığını kabul edelim. Bu takdirde $Y^{(1)}$ ve $Y^{(2)}$ nin istatistiksel bağımsız olması için gerek ve yeter şart

$$C[\ Y^{(1)}, Y^{(2)}\ ]=0$$

olmasıdır.

Sonuç 1.2.2.[1]: Yukarıdaki teorem aşağıdaki gibi genelleştirilebilir; Eğer $Y=(Y^{(1)},Y^{(2)},\dots,Y^{(k)})'$ ve her i,j $(i\neq j)$ için

$$C[\ Y^{(i)}, Y^{(j)}\ ]=0$$

ise bu takdirde $Y^{(i)}$, $(i=1,2,\dots,k)$ ler karşılıklı bağımsızdır.

Teorem 1.2.4.[1]: $Y \sim N_n(\theta,\sigma^2 I_n)$ ve U=AY, V=BY olsun. A_1 A nın lineer bağımsız satırlarını göstersin ve $U_1=A_1Y$ olsun. Bu takdirde eğer C[U,V]=0 ise

i) U_1 ile V'V bağımsızdır.

ii) U'U ile V'V bağımsızdır.

Sonuç 1.2.3.[1]: $Y \sim N(\theta, \sigma^2 I_n)$ ise bu takdirde a'Y ile b'Y nin bağımsız olması için gerek ve yeter şart a'b=0 olmasıdır.

Teorem 1.2.5.[1]: $Y \sim N_n(\theta, \sigma^2 I_n)$, P nxn boyutlu ve r ranklı simetrik bir matris olsun. Bu takdirde $Q=(Y-\theta)'P(Y-\theta)/\sigma^2$ nin χ^2 dağılımına sahip olması için gerek ve yeter şart $P^2=P$ olmasıdır.(yani P nin idempotent olmasıdır.)

Sonuç 1.2.4.[1]: $Y \sim N_n(0, I_n)$ ve $Q_i = Y'P_i Y$ (i=1,2) olmak üzere Q_1 ve Q_2 nin her ikiside ki-kare dağılımına sahip olsun. Bu takdirde Q_1 ile Q_2 nin bağımsız olmaları için gerek ve yeter şart $P_1P_2=0$ eşitliğinin sağlanmasıdır.

Tanım 1.2.2.[1]: $X_1, X_2, \ldots, X_n$ rastgele vektörleri karşılıklı bağımsız ve her biri $N_p(\mu, \Sigma)$ dağılımına sahip olsunlar. Bu takdirde

$$L(\mu,\Sigma)=\prod_{j=1}^{n}\left\{\frac{1}{(2\pi)^{\frac{1}{2}p}\,|\Sigma|^{\frac{1}{2}}}\, e^{-\frac{1}{2}(X_j-\mu)'\Sigma^{-1}(X_j-\mu)}\right\}$$

$$=\frac{1}{(2\pi)^{\frac{1}{2}np}\,|\Sigma|^{\frac{1}{2}n}}\, e^{-\frac{1}{2}\sum_{j=1}^{n}(X_j-\mu)'\Sigma^{-1}(X_j-\mu)}$$

fonksiyonuna Likelihood fonksiyonu denir. Bu fonksiyonu maksimumlaştıran paremetreleri bulma tekniğine maksimum likelihood tahmini ve maksimumlaştırmayı sağlayan parametrelere de maksimum likelihood tahminleri denir.

Teorem 1.2.6.[12]: $X_1, X_2, \ldots, X_n$ μ ortalamalı ve Σ kovaryans matrisli bir normal kitleden alınan bir rastgele örneklem olsun. Bu takdirde

$$\hat{\mu} = \bar{X} \text{ ve } \hat{\Sigma} = (1/n) \sum_{j=1}^{n} (X_j - \bar{X})(X_j - \bar{X})' = [(n-1)/n]\, S$$

sırasıyla μ ve Σ nın maksimum likelihood tahmin edicileridir.

Tanım 1.2.3.[16]: Y nxl boyutlu rastgele vektörü $N(\mu, I)$ dağılımına sahip olsun. Bu takdirde ihtimal yoğunluk fonksiyonu

$$\chi^2(U:n,t) = \begin{cases} \displaystyle\sum_{j=1}^{\infty} \frac{e^{-t} t^j}{j!} \; \frac{U^{\frac{1}{2}(n+2j-2)}\, e^{-\frac{1}{2}U}}{\Gamma((n+2j)/2)\, 2^{j+n/2}} & U>0 \\ 0 & U \le 0 \end{cases}$$

olan $U = Y'Y$ rasgele değişkenine ki-kare dağılımına sahiptir denir ve $t = \frac{1}{2}\mu'\mu$ değerine de merkezi olmama paremetresi adı verilir. $\mu = 0$ olduğunda bu dağılım merkezi ki-kare olarak adlandırılır ve kısaca χ^2_n ile gösterilir.

Lemma 1.2.1.[11]: A nxn boyutlu simetrik bir matris olmak üzere Y nxl rasgele vektörü $N(\mu, \Sigma)$ dağılımına sahip olsun. Bu takdirde $t = \mu' A\, \mu$ olmak üzere $U = Y'A\,Y$ nin $\chi^2(U:r,t)$ dağılımına sahip olması için gerek ve yeter şart $iz(A\Sigma) = r$ ve $\Sigma A \Sigma A \Sigma = \Sigma A \Sigma$ olmasıdır. Burada t merkezi olmama paremetresidir.

Sonuç 1.2.4.[11]: U rasgele değişkeni $\chi^2(U:n,t)$ dağılımına sahip olsun. Bu takdirde $E[U]=n+2t$ ve $var[U]=2(n+4t)$ dir.

Sonuç 1.2.5.[11]: U ve V rasgele değişkenleri sırasıyla $\chi^2(U:n,t)$ ve $\chi^2(V:m,t')$ dağılımına sahip olsun. Bu takdirde $W = U+V$ $\chi^2(W:m+n,t+t')$ dağılımına sahiptir.

Sonuç 1.2.6.[11]: $Y \sim N_n(\mu,\sigma^2 I)$ nx1 boyutlu bir rastgele vektör olsun. Bu takdirde $t=\mu'\mu/2\sigma^2$ olmak üzere $U=Y'Y/\sigma^2$ $\chi^2(U:n,t)$ dağılımına sahiptir.

Sonuç 1.2.7.[11]: $Y \sim N_n(\mu,\Sigma)$ nx1 boyutlu bir rastgele vektör olsun, burada Σ n ranklıdır. Bu takdirde $t = \frac{1}{2}\mu'\Sigma^{-1}\mu$ olmak üzere $U=Y'\Sigma^{-1}Y$ $\chi^2(U:n,t)$ dağılımına sahiptir.

Tanım 1.2 4.[6]: a) U_1 ve U_2 rasgele değişkenleri sırasıyla χ^2_n ve χ^2_m dağılımına sahip olsunlar. Bu takdirde $(U_1/n)/(U_2/m)$ rastgele değişkeninin dağılımına n, m serbestlik dereceli (merkezi) F-dağılımı denir ve $F_{n,m}$ ile gösterilir.

b) X rasgele vektörü ve U rastgele değişkeni sırasıyla $N_n(0,I)$ ve χ^2_n dağılımına sahip olsunlar. Bu takdirde $X/(U/n)^{\frac{1}{2}}$ rastgele değişkeninin dağılımına n serbestlik dereceli Student-t dağılımı adı verilmektedir.

Teorem 1.2.7.[2]; $Y \sim N_n(0,I)$ nx1 boyutlu bir rastgele vektör olsun. $Y'AY$ kuadratik formunun k serbestlik dereceli merkezi bir ki-kare dağılımına sahip olması için gerek ve yeter şart A nın k ranklı simetrik idempotent bir matris olmasıdır.

Teorem 1.2.8.[11]: $Y \sim N_n(\mu, I)$ nx1 boyutlu bir rastgele vektör olsun. $U=Y'A\,Y$ rastgele değişkeninin $t = \frac{1}{2}\mu' A\,\mu$ olmak üzere $\chi^2(U:k,t)$ dağılımına sahip olması için gerek ve yeter şart A nın k ranklı idempotent bir matris olmasıdır.

Teorem 1.2.9.[2]: Σ n ranklı olmak üzere $Y \sim N_n(\mu, \Sigma)$ nx1 boyutlu bir rastgele vektör olsun. Bu takdirde $U=Y'A\,Y$ kuadratik formunun $t=\frac{1}{2}\mu' A\,\mu$ olmak üzere $\chi^2(U:p,t)$dağılımına sahip olması için gerek ve yeter şart aşağıdaki üç şarttan herhangi birinin sağlanmasıdır:

i) $A\Sigma$ p ranklı idempotent bir matristir.

ii) ΣA p ranklı idempotent bir matristir.

iii) Σ A nın bir c-inversidir ve A matrisi p ranklıdır.

Tanım 1.2.5.[2]: A mxn boyutlu bir matris olsun. Bir A^c matrisine A nın bir şartlı inversi (c-inversi) denir:$\Longleftrightarrow$ $A\,A^c\,A=A$ dır.

Teorem 1.2.10.[2]: $Y \sim N_n(\mu, I)$ nx1 boyutlu bir rastgele vektör ve B de pxn boyutlu bir matris olsun. Bu takdirde $BA=0$ ise BY vektörü ile $Y'A\,Y$ kuadratik formu bağımsızdır.

Sonuç 1.2.7.[2] Σ n ranklı olmak üzere $Y \sim N_n(\mu, \Sigma)$ nx1 boyutlu bir rastgele vektör olsun. B pxn boyutlu bir matris olmak üzere eğer $B\Sigma A=0$ ise $Y'A\,Y$ kuadratik formu ve BY lineer formu bağımsızdır.

Sonuç 1.2.8.[2]: Σ n ranklı olmak üzere $Y \sim N_n(\mu, \Sigma)$ nx1 boyutlu bir rastgele vektör olsun. Eğer $A\Sigma B = 0$ ise bu takdirde $Y'A\,Y$ ve $Y'B\,Y$ kuadratik formları bağımsızdır.

Tanım 1.2.6.[3]: X $E[X]=\mu$ ortalamalı ve $cov[X]=\Sigma$ varyans-kovaryans matrisli nx1 boyutlu bir rastgele vektör olsun. Yani $\sigma_{ii}=\sigma_i^2$ ve

$$\sigma_{ij}=E[(X_i-\mu_i)(X_j-\mu_j)], \quad (i\neq j)$$

olmak üzere

$$\Sigma=\mathrm{cov}[X]=E[(X-\mu)(X-\mu)'] = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2n} \\ . & . & \cdots & . \\ . & . & \cdots & . \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_{nn} \end{bmatrix}$$

olsun. $\sigma_{ii}>0$ ve $\sigma_{jj}>0$ için

$$\rho_{ij} = \frac{\mathrm{cov}[X_i,X_j]}{(\mathrm{var}[X_i]\mathrm{var}[X_j])^{\frac{1}{2}}} = \frac{\sigma_{ij}}{(\sigma_{ii}\sigma_{jj})^{\frac{1}{2}}}$$

ye X_i ve X_j arasındaki korelasyon katsayısı adı verilir.

Eğer $\sigma_{ii}=0$ veya $\sigma_{jj}=0$ ise ρ_{ij} yi tanımlayamayız. Bunun sonucunda

$$\rho = \begin{bmatrix} 1 & \rho_{12} & \cdots & \rho_{1n} \\ \rho_{21} & 1 & \cdots & \rho_{2n} \\ . & . & \cdots & . \\ . & . & \cdots & . \\ \rho_{n1} & \rho_{n2} & \cdots & 1 \end{bmatrix}$$

simetrik matrisi kitle korelasyon katsayılar matrisi olarak ve

$$V^{\frac{1}{2}} = \begin{bmatrix} \sqrt{\sigma_{11}} & 0 & \cdots & 0 \\ 0 & \sqrt{\sigma_{22}} & \cdots & 0 \\ . & . & \cdots & . \\ . & . & \cdots & . \\ 0 & 0 & \cdots & \sqrt{\sigma_{nn}} \end{bmatrix}$$

matrisi de standart sapma matrisi olarak tanımlanır.

Sonuç 1.2.9.[3]: ρ , Σ ve $V^{\frac{1}{2}}$ yukarıdaki gibi tanımlı ise

$$V^{\frac{1}{2}}\rho V^{\frac{1}{2}} = \Sigma \quad \text{ve} \quad \rho = (V^{\frac{1}{2}})^{-1}\Sigma(V^{\frac{1}{2}})^{-1}$$

özellikleri sağlanır. Ayrıca ρ korelasyon matrisi non-negatif olup

$$r(\rho) = r(\Sigma)$$

eşitliği gerçeklenir.

Sonuç 1.2.10.[3]: X_i ve X_j rastgele değişkenleri için

$$-1 \leq \rho_{ij} \leq 1$$

eşitsizliği sağlanır.

BÖLÜM II

LİNEER MODELLER

2.1. EN KÜÇÜK KARELER TAHMİNİ

Bir değişkenin diğeri ile olan ilişkisinin nasıl olduğunun belirtilmesi problemi bir " model kurma " sorunudur. Y bilinmeyen bir η parametresine göre değişen bir rastgele değişken yani e hata terimi olmak üzere $Y=\eta+e$ olsun. örneğin e deneyde η ya sebep olan bir doğal değişim olabilir veya η nın ölçülmesindeki hatayı gösterebilir. Şimdi η parametresinin

$$\eta=\beta_o+\beta_1 x_1+\ldots+\beta_{m-1}x_{m-1}$$

şeklinde ifade edilebildiğini kabul edelim, burada x_i'ler bilinen parametreler (örneğin deneyci tarafından kontrol edimiş ve ihmal edilebilir bir hatayla ölçülmüş deneysel değişkenler) ve $\beta_j(j=o,1,\ldots,m-1)$ bilinmeyen parametrelerdir. Eğer x_j değiştiğinde Y nin $Y_1,Y_2,\ldots,Y_n$ gibi değerleri elde ediliyorsa bu takdirde x_{ij} x_j nin i-yinci değeri olmak üzere

$$Y_i=\beta_o+\beta_1 x_{i1}+\ldots+\beta_{m-1}x_{i,m-1}+e_i$$

dir. Bu n denklem matris formunda $x_{1o}=x_{2o}=\ldots=x_{no}=1$ olmak üzere

$$\begin{bmatrix} Y_1 \\ Y_2 \\ . \\ . \\ Y_n \end{bmatrix} = \begin{bmatrix} x_{1o} & x_{11} & x_{12} & \cdots & x_{1,m-1} \\ x_{2o} & x_{21} & x_{22} & \cdots & x_{2,m-1} \\ . & . & . & \cdots & . \\ . & . & . & \cdots & . \\ x_{no} & x_{n1} & x_{n2} & \cdots & x_{n,m-1} \end{bmatrix} \begin{bmatrix} \beta_o \\ \beta_1 \\ . \\ . \\ \beta_{m-1} \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ . \\ . \\ e_n \end{bmatrix}$$

veya kısaca

$$Y = X\beta + e \qquad (2.1.1)$$

olarak yazılır,burada Y rastgele değişkenlerin nx1 boyutlu bir gözlem vektörü, X bilinen değerlerin nxm boyutlu ve q($q \leq m<n$) ranklı bir matrisi, β bilinmeyen parametrelerin mx1 boyutlu bir vektörü ve e de gözlenemeyen hataların nx1 boyutlu bir vektörüdür. X matrisine " tasarım matrisi " veya "regresyon matrisi " denir. (2.1.1) modelinin oldukça genel bir model olduğunu belirtelim. Örneğin, $x_{ij}=x_i^{\,j}$ ve k=m-1 alarak

$$Y_i = \beta_o + \beta_1 x_i + \beta_2 x_i^2 + \ldots + \beta_k x_i^k + e_i$$

polinom modelini elde ederiz. (2.1.1) bağıntısının önemliliği β_j bilinmeyen parametrelerine göre lineer olmasıdır, bu nedenle buna "Lineer Model " adı verilir.

$$Y_i = \beta_o + \beta_1 e^{-\beta_2 x} + e_i$$

β_2 ye göre non-lineer olduğundan bir non-lineer modeldir.

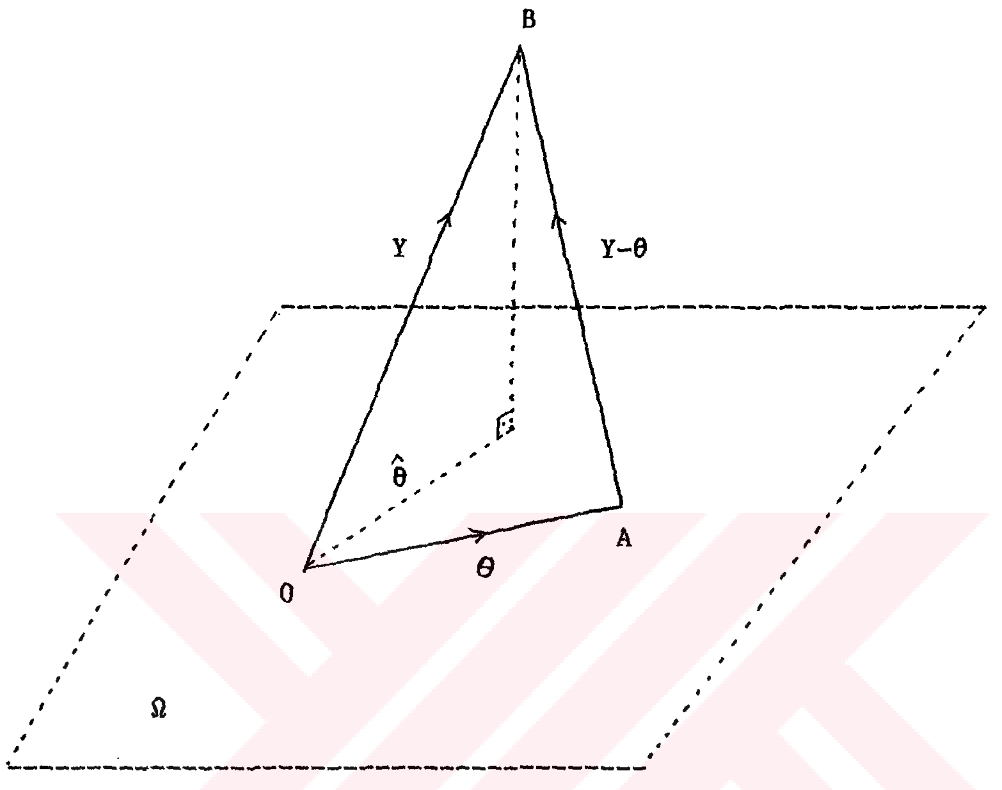
β nın tahmini problemini ele almadan önce buradaki ve daha sonraki kısımlardaki tüm teorilerin (2.1.1) modeli için geliştirildiğini belirtelim, burada x_{io} ler 1 olarak alınmayabilir. β nın bir tahminini elde etmek için bir metod " en küçük kareler " metodu olarak adlandırılır. Bu metod, $\sum_i e_i^2$ yi β ya göre minimumlaştırmaktan ibarettir. Yani $\theta = X\beta$ konularak Ω X in ranj (aralık) uzayı ($=\{y|$ herhangi bir x için $y=Xx\}$) olmak üzere $\theta \in R[X]=\Omega$ ya bağlı olarak $e'e= \|Y-\theta\|^2$ yi minimumlaştırırız. θ yı Ω üzerinde değiştirirsek $\|Y-\theta\|^2$ (Y-θ nın uzunluğunun karesi) $(Y-\hat{\theta}) \perp \Omega$ iken $\hat{\theta} = \theta$ için minimum olacaktır. (bkz.şekil 2.1.) Böylece

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

veya

$$X'\hat{\theta}=X'Y \qquad (2.1.2)$$

dır, burada $\hat{\theta}$ tek türlü olup Y nin Ω üzerindeki yegâne ortagonal projeksiyonu(izdüşümü) olur.



Sekil 2.1. AB minimum olacak şekilde A nın bulunmasından ibaret E.K.K Yöntemi

X in sütunları lineer bağımsız ve $\hat{\theta}=X\hat{\beta}$ olacak şekilde bir tek $\hat{\beta}$ nın mevcut olduğunu gözönüne alalım. Böylece (2.1.2) bağıntısından normal denklem(ler) denilen

$$X'X\,\hat{\beta}=X'Y \tag{2.1.3}$$

elde edilir. X m ranklı olduğundan X'X pozitif definit olup non-singülerdir. Bu nedenle (2.1.3) bir tek çözüme sahiptir, dolayısıyla

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

dir. Burada $\hat{\beta}$ ya β nın en küçük kareler tahmini denir.

X matrisinin m ranklı (yani tam sütun ranklı) olmaması durumunda böyle bir tek çözümden bahsetmek mümkün değildir. Çünkü bu durumda X'X matrisi singüler olabilir. Bu nedenle çözüm için X'X in genelleştirilmiş inversi kullanılır. Böylece G X'X in genelleştirilmiş inversi ve w uygun boyutlu rastgele bir vektör olmak üzere normal denklemlerin çözümü

$$\hat{\beta} = G X'Y + (I - G X'X)w \tag{2.1.4}$$

dır. $X\hat{\beta}$ uydurulmuş regresyonu $\hat{Y}(=[(\hat{Y}_i)])$ ile gösterilir.

$$\begin{aligned} Y-\hat{Y} = Y-X\hat{\beta} &= (I_n-X(X'X)X')Y \\ &= (I_n-P)Y \end{aligned} \tag{2.1.5}$$

nin elemanlarına rezidüler adı verilir ve

$$\begin{aligned} RSS = (Y-X\hat{\beta})'(Y-X\hat{\beta}) &= Y'Y-2\hat{\beta}'X'Y + \hat{\beta}'X'X\hat{\beta} \\ &= Y'Y-\hat{\beta}'X'Y + \hat{\beta}'(X'X\hat{\beta}-X'Y) \\ &= Y'Y-\hat{\beta}'X'Y \\ &= Y'Y-\hat{\beta}'X'X\hat{\beta} \end{aligned} \tag{2.1.6}$$

ye de rezidü kareler toplamı denir. $\hat{\theta}=X\hat{\beta}=X(X'X)^{-1}X'Y=PY$ olduğundan P n-boyutlu öklid uzayı R^n nın Ω ya olan izdüşümünü gösteren bir lineer transformasyondur. Benzer şekilde I_n-P de R^n nin Ω nın ortogonal bütünleyeni $\Omega^{\perp}$ e olan dik izdüşümünü gösterir. Böylece $Y=PY+(I_n-P)Y$, Y nin birinci bileşeni Ω ve ikinci bileşeni $\Omega^{\perp}$ üzerinde bulunan iki bileşenli üzerine dik izdüşümünü gösterir. P ve I_n-P nin bazı özellikleri aşağıdaki teoremde verilmektedir.

Teorem 2.1.1.[1]: $P=X(X'X)^{-1}X'$ olmak üzere

i) P ve I_n-P simetrik ve idempotenttir.

ii) $r(I_n-P)=iz(I_n-P)=n-m$

iii) $(I_n-P)X=0$

dır.

Ispat: i) P nin simetrik olduğu açıktır. Ayrıca

$$(I_n-P)'=I_n-P'=I_n-P$$

olup I_n-P de simetriktir. Öte yandan

$$P^2= X(X'X)^{-1}X'X(X'X)^{-1}X'= XI_m(X'X)^{-1}X'=P$$

ve

$$(I_n-P)^2= I_n-2P+P^2= I_n-2P+P = I_n-P$$

dir.

ii) I_n-P simetrik idempotent olduğundan rankı izine eşittir. Dolayısıyla

$$r(I_n-P) = iz(I_n-P) = n-iz(P)$$

elde edilir, burada

$$iz(P)=iz[X(X'X)^{-1}X']=iz[X'X(X'X)^{-1}]=iz(I_m)=m$$

dir.

iii) $(I_n-P)X = X-X(X'X)^{-1}X'X = X-X = 0$

dır. Böylece ispat biter.

Eğer $r(X)=q<m$ ise bu takdirde $\hat{\theta}=PY$ yine bir tek P izdüşüm matrisi tanımlar. X_1 X in q tane lineer bağımsız sütunundan oluşan nxq boyutlu bir matris olsun. Bu takdirde (2.1.2) den $X_1'(Y-\hat{\theta})=0$ ve bir $\hat{\alpha}$ için $\hat{\theta}=X_1\hat{\alpha}$ olup $P=X_1(X_1'X_1)^{-1}X_1'$dir. Bu durumda yukarıdaki teorem yine geçerlidir, ancak m ile q yerdeğiştirir. P nin $X(X'X)^{-}X'$ formunda da ifade edilebileceğini belirtelim.

Şimdi en küçük kareler tahmininin bazı özelliklerini verelim.

Teorem 2.1.2.[12]: $E[e]=0$ ve $E[ee']=\sigma^2 I_n$ ise (2.1.1) lineer modeli altında $\hat{\beta}=(X'X)^{-1}X'Y$ en küçük kareler tahmini için

$$E[\hat{\beta}] = \beta \text{ ve } D[\hat{\beta}] = \sigma^2(X'X)^{-1}$$

dir. Ayrıca $\hat{e} = Y-X\hat{\beta}$ olmak üzere

$$E[\hat{e}] = 0 \text{ ve } D[\hat{e}] = \sigma^2(I_n-X(X'X)^{-1}X')$$

dir, $E[\hat{e}'\hat{e}] = (n-m)\sigma^2$ olup

$$S^2 = \frac{\hat{e}'\hat{e}}{n-m} = \frac{Y'[I_n-X(X'X)^{-1}X']\,Y}{n-m}$$

için $E[S^2] = \sigma^2$ dir. Buna ilaveten $\hat{\beta}$ ve $\hat{e}$ bağımsızdırlar.

Ispat: $E[\hat{\beta}] = E[(X'X)^{-1}X'Y] = (X'X)^{-1}X'X\,\beta = \beta$

olup $\hat{\beta}$ β nın bir yansız tahminidir.

$$D[e] = \sigma^2 I_n \text{ ve } D[Y] = D[Y-X\beta] = D[e]$$

olduğundan

$$D[\hat{\beta}] = D[(X'X)^{-1}X'Y] = (X'X)^{-1}X'D[Y]\;X(X'X)^{-1}$$
$$= \sigma^2(X'X)^{-1}X'X(X'X)^{-1} = \sigma^2(X'X)^{-1} \qquad (2.1.7.)$$

dir. $\hat{e} = Y-X\hat{\beta}$ olmak üzere

$$E[\hat{e}] = E[Y-X\hat{\beta}] = E[Y]-XE[\hat{\beta}] = X\beta-X\beta = 0$$

ve

$$\hat{e} = Y-X(X'X)^{-1}X'Y = (I_n-X(X'X)^{-1}X')Y = (I_n-X(X'X)^{-1}X')(X\beta+e)$$
$$=(I_n-X(X'X)^{-1}X')e$$

olup $I_n-X(X'X)^{-1}X'$ idempotent olduğundan

$$D[\hat{e}] = E[(\hat{e}-E[\hat{e}])(\hat{e}-E[\hat{e}])'] = E[\hat{e}\hat{e}']$$
$$= E[(I_n-X(X'X)^{-1}X')ee'(I_n-X(X'X)^{-1}X')]$$
$$= (I_n-X(X'X)^{-1}X')E[ee'](I_n-X(X'X)^{-1}X')$$
$$= \sigma^2(I_n-X(X'X)^{-1}X')$$

elde edilir.

$$\hat{e}'\hat{e} = e'(I_n-X(X'X)^{-1}X')(I_n-X(X'X)^{-1}X')\;e$$
$$= e'(I_n-X(X'X)^{-1}X')\;e$$
$$= iz[\;e'(I_n-X(X'X)^{-1}X')\;e\;]$$
$$= iz[\;(I_n-X(X'X)^{-1}X')ee']$$

dir. Böylece

$$E[\hat{e}'\hat{e}] = iz[(I_n-X(X'X)^{-1}X')E[ee']]$$
$$= \sigma^2 iz[I_n-X(X'X)^{-1}X']$$

$$= \sigma^2 iz[I_n] - \sigma^2 iz[(X'X)^{-1}X'X] = \sigma^2 n - \sigma^2 m = \sigma^2(n-m)$$

elde edilir. Dolayısıyle

$$E[S^2] = \frac{1}{(n-m)} E[\hat{e}'\hat{e}] = \frac{1}{(n-m)} (n-m)\,\sigma^2 = \sigma^2$$

dir. Şimdi de $\hat{\beta}$ ile $\hat{e}$ nin bağımsız olduğunu gösterelim.

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

ve $E[\hat{e}] = 0$ olduğundan

$$\begin{aligned} cov[\hat{\beta},\hat{e}] &= E[\ (\hat{\beta}-E[\hat{\beta}])(\hat{e}-E[\hat{e}])'\ ] = E[(\hat{\beta}-\beta)\hat{e}'] \\ &= E[(X'X)^{-1}X'ee'(I_n-X(X'X)^{-1}X')] \\ &= (X'X)^{-1}X'E[ee'](I_n-X(X'X)^{-1}X') \\ &= \sigma^2(X'X)^{-1}X'(I_n-X(X'X)^{-1}X') \\ &= 0 \end{aligned}$$

dır. Çünkü $X'(I_n-X(X'X)^{-1}X')=X'-X'X(X'X)^{-1}X' = X'-X'=0$ dır. Böylece $\hat{\beta}$ ile $\hat{e}$ bağımsızdır.

Teorem 2.1.3.[1]: X nxm boyutlu m ranklı olmak üzere eğer Y $N_n(X\beta\ ,\ \sigma^2 I_n)$ dağılımına sahip ise bu takdirde

i) $\hat{\beta} \sim N_m(\beta\ ,\sigma^2(X'X)^{-1})$

ii) $(\hat{\beta}-\beta)'X'X(\hat{\beta}-\beta)/\sigma^2 \sim \chi^2_m$

iii) $RSS/\sigma^2 = (n-m)S^2/\sigma^2 \sim \chi^2_{n-m}$

dir.

Ispat:i) $\hat{\beta} = (X'X)^{-1}X'Y$ olduğundan $C = (X'X)^{-1}X'$ olarak tanımlanırsa $\hat{\beta} = CY$ olur, burada C mxn boyutlu olup $r(C) = r(X') = r(X) = m$ dir. $Y \sim N_n(X\beta\ ,\ \sigma^2 I_n)$ olduğundan $\hat{\beta} \sim N_m(\beta\ ,\ \sigma^2(X'X)^{-1})$ elde edilir.

ii) $(\hat{\beta}-\beta)'X'X(\hat{\beta}-\beta)/\sigma^2 = (\hat{\beta}-\beta)'(D[\hat{\beta}])^{-1}(\hat{\beta}-\beta)$

olduğundan χ^2_m dağılımına sahiptir.

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

$$= Y'(I_n-X(X'X)^{-1}X')Y = Y'(I_n-P)Y \qquad (2.1.8.)$$

dır. I_n-P simetrik idempotent olduğundan $RSS/\sigma^2 \sim \chi^2_{n-m}$ dağılımlıdır.

Teorem 2.1.4.[1]: $\hat{\theta}$ θ nın en küçük kareler tahmini olsun. Bu takdirde $c'\theta$ nın lineer yansiz tahminlerinin sınıfı içinde $c'\hat{\theta}$ minimum varyanslı yegane tahmindir. $c'\hat{\theta}$ ya $c'\theta$ nın en iyi lineer yansız tahmini denir.

Ispat: Yukarıda verilen teoremlerin sonucu olarak $\hat{\theta} = PY$ dir, burada $PX=X$ sağlanır. Böylece her $\theta \in \Omega = R[X]$ icin $E[c'\hat{\theta}] = c'P\theta = c'\theta$ ve $c'\hat{\theta}[=(Pc)'Y]$ $c'\theta$ nın bir lineer yansız tahminidir. $d'Y$ $c'\theta$ nın diğer herhangi bir lineer yansız tahmini olsun. Bu takdirde $c'\theta = E[d'Y] = d'\theta$ olup $(c-d)'\theta=0$ dır,dolayısıyla $(c-d)$ vektörü Ω ya diktir.Buradan $P(c-d) = 0$ ve $Pc = Pd$ elde edilir.

$$var[(Pd)'Y] = D[(Pd)'Y] = \sigma^2 d'P'Pd = \sigma^2 d'Pd$$

olup

$$\begin{aligned} var[d'Y]-var[c'\hat{\theta}] &= var[d'Y]-var[(Pd)'Y] \\ &= \sigma^2(d'd-d'Pd) \\ &= \sigma^2 d'(I_n-P)d \\ &= \sigma^2 d'(I_n-P)'(I_n-P)d \\ &= \sigma^2 f'f \geq 0 \text{ ,diyelim,} \end{aligned}$$

dır. Eşitlik $(I_n-P)d = 0$ veya $d = Pd = Pc$ ise sağlanır. Böylece $c'\hat{\theta}$ minimum varyansa sahip olup tektir.

Şimdi V nxn boyutlu pozitif definit bir matris olmak üzere $D[e] =\sigma^2V$ olsun. V pozitif definit olduğundan $V = KK'$ olacak şekilde nxn boyutlu nonsingüler bir K matrisi mevcuttur. $Z=K^{-1}Y$, $B=K^{-1}X$ ve $e^*=K^{-1}e$ konularak B nxm boyutlu m ranklı olmak uzere $Z=B\beta+e^*$ modeli elde edilir. Burada $E[e^*]=0$ ve $D[e^*]=\sigma^2 I_n$ dir. $e^*e^{*'}$ ifadesini β ya göre minimumlaştırarak dönüştürülmüş model için β nın E.K.K. tahmini

$$\beta^* = (B'B)^{-1}B'Z = (X'(KK')^{-1}X)^{-1}X'(KK')^{-1}Y$$
$$= (X'V^{-1}X)^{-1}X'V^{-1}Y \qquad (2.1.9)$$

elde edilir, burada β^* nın ortalaması

$$E[\beta^*] = (X'V^{-1}X)^{-1}X'V^{-1}X\beta = \beta$$

ve dağılım (varyans-kovaryans) matrisi

$$D[\beta^*] = \sigma^2(B'B)^{-1} = \sigma^2(X'V^{-1}X)^{-1}$$

olup rezidü kareler toplamı

$$(Z-B\beta^*)'(Z-B\beta^*) = (Y-X\beta^*)'(KK')^{-1}(Y-X\beta^*)$$
$$=(Y-X\beta^*)'V^{-1}(Y-X\beta^*)$$

dır. Buna alternatif olarak

$$e^{*'}e^* = e'V^{-1}e = (Y-X\beta)'V^{-1}(Y-X\beta)$$
$$= Y'V^{-1}Y-2\beta'X'V^{-1}Y + \beta'X'V^{-1}X\beta$$

dan β ya göre türev almak suretiyle de β^* elde edilebilir. Böylece

$$\partial e^{*'}e^* / \partial\beta = -2X'V^{-1}Y + 2X'V^{-1}X\beta$$

bağıntısını sıfıra eşitleyip buradan β^* çekilebilir: $X'V^{-1}X$ pozitif definit olduğundan bir inverse sahiptir.

β^* elde edildikten sonra β^* ile $\hat{\beta} = (X'X)^{-1}X'Y$ nın aynı olup olmadığını yani ne zaman $D[e]$ nin σ^2V olabileceği ve σ^2I_n olamayacağı sorulabilir. Bunun cevabı aşağıdaki teoremle verilir.

Teorem 2.1.5.[1]: β^* ile $\hat{\beta}$ nın özdeş olması için gerek ve yeter şart $R[V^{-1}X] = R[X]$ olmasıdır.

Ispat:

β^* ile $\hat{\beta}$ özdeştir $\Longleftrightarrow$ $\forall$ Y için $X'V^{-1}Y=(X'V^{-1}X)(X'X)^{-1}X'Y$ (*)

dır. $Y_1 \in R[X]$ ve $Y_2 \perp R[X]$ olmak üzere Y bir tek $Y = Y_1 + Y_2$ parçalanışına sahip olsun. Bir a için $Y_1 = Xa$ olduğundan her V^{-1} için (*) bağıntısı Y_1 ile sağlanır. $X'Y_2 = 0$ olduğundan $X'V^{-1}Y_2 = 0$ olan her Y_2 için (*) bağıntısı yine sağlanır. Böylece her Y için $X'V^{-1}Y = (X'V^{-1}X)(X'X)^{-1}X'Y$ olması için gerek ve yeter şart $X'x=0$ dan $X'V^{-1}x=0$

bağıntısının sağlanmasıdır, yani $R[V^{-1}X] \subset R[X]$ olmasıdır. Bu iki uzay aynı boyutlu olduğundan bu $R[V^{-1}X] = R[X]$ olmasını sağlar.

Sonuç 2.1.1.[1]: $\hat{\beta} = \beta^* \Longleftrightarrow R[VX] = R[X]$ (2.1.10)

Ispat: $R[V^{-1}X] = R[X]$ verilsin ve $z \in R[VX]$ olsun. Bu takdirde $z = VXb = V(V^{-1}Xc) = Xc \in R[X]$ olacak şekilde b ve c mevcuttur. O halde $R[VX] \subset R[X]$ olup yukarıda verilen boyut özelliğinden $R[VX] = R[X]$ elde edilir. Tersine olarak $R[VX]=R[X]$ ve $z \in R[V^{-1}X]$ olsun. Bu takdirde $z = V^{-1}Xd = V^{-1}(VXf) = Xf$ olacak şekilde f ve d meccuttur. Yani $R[V^{-1}X] \subset R[X]$ dir. Bu nedenle $R[V^{-1}X] = R[X]$ elde edilir. Böylece $R[VX] = R[X]$ olması için gerek ve yeter şart $R[V^{-1}X] = R[X]$ olmasıdır.

Sonuç 2.1.2.[1]: $Y = x\beta + e$ ise her x için $\beta^* = \hat{\beta} \Longleftrightarrow V = cI_n$ dir.

2.2. LİNEER KISITLAMALI TAHMİN

2.2.1. Lagrange Çarpanı Metodu

X nxm boyutlu m ranklı bir matris olmak üzere $Y=X\beta+e$ olsun. R pxm boyutlu (p<m) bilinen bir matris ve r de px1 boyutlu bilinen bir vektör olmak üzere $R\beta=r$ uygun lineer kısıtlaması altında $e'e$ nin minimumunu bulmak istediğimizi varsayalım. Bu problemin çözümü için bir metod Lagrange çarpanlarını kullanmaktır. Yani a_i' R nin i-yinci satırı olmak üzere $a_i'\beta = r_i$, $i=1,2,\ldots,p$, lineer sınırlaması için bir Lagrange çarpanı bulmaktır. Böylece

$$\sum_{i=1}^{p} \lambda_i(a_i'\beta - r_i) = \lambda'(R\beta - r)$$

ile ligileneceğiz. Lagrange çarpanları metodunu uygulamak için

$$L(\beta,\lambda) = e'e + \lambda'(R\beta - r) = (Y-X\beta)'(Y-X\beta) + \lambda'(R\beta - r)$$

ifadesini gözönüne alalım ve

$$\frac{\partial L}{\partial \beta} = -2X'Y + 2X'X\beta + R'\lambda = 0 \qquad (2.2.1)$$

ve

$$R\beta = r$$

denklemlerini çözelim. Bu iki denklemin çözümünü $\hat{\beta}_H$ ve $\hat{\lambda}_H$ ile gösterelim. (2.2.1) bağıntısından

$$\hat{\beta}_H = (X'X)^{-1}X'Y-\tfrac{1}{2}(X'X)^{-1}R'\hat{\lambda}_H = \hat{\beta}-\tfrac{1}{2}(X'X)^{-1}R'\hat{\lambda}_H \qquad (2.2.2)$$

elde edilir. $R\beta=r$ bağıntısından da

$$r = R\hat{\beta}_H = R\hat{\beta}- \tfrac{1}{2}R(X'X)^{-1}R'\hat{\lambda}_H \qquad (2.2.3)$$

elde edilir. $(X'X)^{-1}$ pozitif definit bir matrisin inversi olarak pozitif definittir ve dolayısıyla $R(X'X)^{-1}R'$ de pozitif definit olup nonsingülerdir. Böylece

$$-\tfrac{1}{2}\hat{\lambda}_H = [R(X'X)^{-1}R']^{-1}(r-R\hat{\beta})$$

dır. Bu ifadeyi (2.2.2) de yerine koyarak

$$\hat{\beta}_H = \hat{\beta} + (X'X)^{-1}R'[R(X'X)^{-1}R']^{-1}(r-R\hat{\beta}) \qquad (2.2.4)$$

elde edilir. $\hat{\beta}_H$ nın $e'e$ ifadesini $R\beta=r$ şartı altında minimumlaştırdığını görmek için ilk olarak

$$\begin{aligned}\| X(\hat{\beta}-\beta)\|^2 &= (\hat{\beta}-\beta)'X'X(\hat{\beta}-\beta)\\ &= (\hat{\beta}-\hat{\beta}_H+\hat{\beta}_H-\beta)'X'X(\hat{\beta}-\hat{\beta}_H+\hat{\beta}_H-\beta)\\ &= (\hat{\beta}-\hat{\beta}_H)'X'X(\hat{\beta}-\hat{\beta}_H)+(\hat{\beta}_H-\beta)'X'X(\hat{\beta}_H-\beta)\\ &=\| X(\hat{\beta}-\hat{\beta}_H) \|^2+\| X(\hat{\beta}_H-\beta)\|^2 \end{aligned} \qquad (2.2.5)$$

olduğunu belirtelim, çünkü (2.2.2) bağıntısından

$$2(\hat{\beta}-\hat{\beta}_H)X'X(\hat{\beta}_H-\beta) = \hat{\lambda}'_H R(\hat{\beta}_H-\beta) = \hat{\lambda}'_H(r-r) = 0 \qquad (2.2.6)$$

dır. Böylece

$$e'e = \| Y-X\hat{\beta} \|^2+\| X(\hat{\beta}-\beta) \|^2$$

$$= \| Y-X\hat{\beta} \|^2+\| X(\hat{\beta}-\hat{\beta}_H)\|^2+\| X(\hat{\beta}_H-\beta)\|^2$$

ifadesi $\| X(\hat{\beta}_H-\beta) \|^2=0$ olduğunda yani $X(\hat{\beta}_H-\beta)=0$ veya $\beta=\hat{\beta}_H$ (X in sütunları lineer bağımsız olduğundan) olduğunda bir minimumdur.

$\beta=\hat{\beta}_H$ konularak

$$\| Y-X\hat{\beta}_H \|^2 = \| Y-X\hat{\beta} \|^2 + \| X(\hat{\beta}-\hat{\beta}_H) \|^2 \qquad (2.2.7)$$

veya $\hat{Y} = X\hat{\beta}$ ve $\hat{Y}_H = X\hat{\beta}_H$ yazarak

$$\| Y-\hat{Y}_H \|^2 - \| Y-\hat{Y} \|^2 = \| \hat{Y}-\hat{Y}_H \|^2 \qquad (2.2.8)$$

elde edilir. Bu özdeşlik direkt olarakta bulunabilir.

2.2. Ortogonal İzdüşümler Metodu

Ortogonal izdüşümleri kullanarak (2.2.4) bağıntısını elde etmek mümkündür. Bunun için önce r yi yokedelim. β_0 $R\beta=r$ nin herhangi bir çüzümü olsun. Bu takdirde

$$Y-X\beta_0 = X(\beta-\beta_0)+e$$

veya $\tilde{Y}=X\delta+e$ ve $R\delta=R\beta-R\beta_0= 0$ dır. Böylece $\theta\in R[X](=\Omega)$ olmak üzere $\tilde{Y}=\theta+e$ modelini elde ederiz. X tam ranklı olduğundan $R(X'X)^{-1}X'\theta = R\delta = 0$ dır. $R_1 = R(X'X)^{-1}X'$ ve $W = \mathcal{N}[R_1]\cap\Omega$ konularak $W^{\perp}\cap\Omega = R[P_\Omega R_1']$ elde edilir, burada

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

nxp boyutlu olup p ranklıdır. Böylece

$$\begin{aligned} P_\Omega-P_W = P_{W^{\perp}\cap\Omega} &= (P_\Omega R_1')[R_1P^2{}_\Omega R_1']^{-1}(P_\Omega R_1')' \\ &= X(X'X)^{-1}R'[R(X'X)^{-1}R']^{-1}R(X'X)^{-1}X' \end{aligned}$$

dir. Bu nedenle $P_\Omega X\beta_0 = X\beta_0$ ve $R\beta_0 = r$ olduğundan

$$\begin{aligned} X\hat{\beta}_H-X\beta_0 &= X\hat{\delta}_H = P_W\tilde{Y} = P_\Omega\tilde{Y}-P_{W^{\perp}\cap\Omega}\tilde{Y} \\ &= P_\Omega Y-X\beta_0-X(X'X)^{-1}R'[R(X'X)^{-1}R']^{-1}(R\hat{\beta}-r) \qquad (2.2.9) \end{aligned}$$

dır. O halde her iki tarafı $(X'X)^{-1}X'$ ile çarparak (2.2.4) bağıntısı elde edilir.

Bu yaklaşımın bir avantajı X in q (q<m) ranklı olması durumunda da kolayca uygulanabilmesidir. Yalnız tahmin edilebilir fonksiyonları tahmin edebildiğimizden $a_i'\beta$ (i=1,2,...) nin tahmin edilebilir yani $a_i'=m_i'X$ olduğunu ve M pxn boyutlu bir matris olmak üzere R = MX

olduğunu kabul ederiz. Çünkü R pxm boyutlu p ranklı olduğundan $p \le q$ almalıyız. $r[R] \le r[M]$ olduğundan M p ranklıdır. Böylece $\theta \in R[X]$ ve $M\theta = MX\delta = R\delta = 0$ olmak üzere modelimizi $\tilde{Y} = \theta + e$ ye indirgeriz. Bu nedenle $P_\Omega M'$ $(= X(X'X)^- X'M' = X(X'X)^- R')$ nxp boyutlu ve p ranklı olmak üzere $W = \mathcal{N}[M] \cap \Omega$ ve $\Omega \cap W^\perp = R[P_\Omega M']$ dir. Buradan

$$\begin{aligned} P_\Omega - P_W &= (P_\Omega M')[MP_\Omega M']^{-1}(P_\Omega M') \\ &= X(X'X)^- R'[R(X'X)^- R']^{-1} R(X'X)^- X' \end{aligned} \qquad (2.2.10)$$

elde edilir. Sonuç olarak aynı özelliği kullanarak (2.2.9) elde edilir. O halde

$$\begin{aligned} X'X\hat{\beta}_H &= X'P_\Omega Y - X'P_\Omega M'[MP_\Omega M']^{-1} MP_\Omega (Y - X\beta_0) \\ &= X'Y - X'M'[MP_\Omega M']^{-1}(R(X'X)^- X'Y - MX\beta_0) \\ &= X'Y - R'[R(X'X)^- R']^{-1}(R\hat{\beta} - r) \end{aligned}$$

olup $\hat{\beta}_H$ için herhangi bir çözüm $\hat{\beta} = (X'X)^- X'Y$ olmak üzere

$$\hat{\beta}_H = \hat{\beta} - (X'X)^- R'[R(X'X)^- R']^{-1}(R\hat{\beta} - r)$$

formundadır.

2.3.KISITLAMALI VE KISITLAMASIZ TAHMİNLERİN KARŞILAŞTIRILMASI

Y nx1 boyutlu $E[Y] = X\beta$ ve $D[Y] = V$ olan bir rastgele vektör X nxm boyutlu bilinen bir matris, β mx1 boyutlu bilinmeyen parametre vektörü ve V de tamamen bilinen veya bir pozitif skaler çarpan hariç bilinen nxn boyutlu non-negatif definit bir matris olmak üzere $(Y, X\beta, V)$ genel lineer modelini gözönüne alalım. β üzerinde ilave bir $R\beta = r$ kısıtlamasının konulduğunu varsayalım, burada R pxm (p<m) boyutlu bilinen bir matris ve r de px1 boyutlu bilinen bir vektördür. $R\beta = r$ kısıtlaması altında $(Y, X\beta, V)$ modelini $(Y, X\beta, V | R\beta = r)$ ile gösterelim. $R\beta = r$ için bir genel çözümün $U = I - R^- R$ ve t keyfi bir vektör olmak üzere

$$\beta = R^- r + Ut \qquad (2.3.1)$$

ile verildiğini biliyoruz. O halde

$$E[Y] = X\beta = X(R^{-}r+Ut)$$

veya buna denk olarak

$$E[Y-XR^{-}r] = XUt$$

dir. Böylece $(Y,X\beta,V|R\beta=r)$ modeli denk olarak $(Y-XR^{-}r,XUt,V)$ formunda yazılabilir, burada t bir parametre vektörü ve $U=I-R^{-}R$ dir. $X\beta$ nın $(Y,X\beta,V)$ ve $(Y,X\beta,V|R\beta=r)$ altındaki en iyi lineer yansız tahmin edicilerini $X\hat{\beta}$ ve $X\hat{\beta}_H$ ile göstermiştik. Bu durumda $W=V+XUU'X'$ olmak üzere

$$X\hat{\beta}_H = XR^{-}r+XU(U'X'W^{-}XU)^{-}U'X'W^{-}(Y-XR^{-}r)$$

bağıntısı sağlanır. Şimdi hangi şartlar altında $X\hat{\beta}=X\hat{\beta}_H$ olduğunu veren aşağıdaki teoremi ifade edelim.

Teorem 2.3.1.[14]: $(Y,X\beta,V)$ ve $(Y,X\beta,V|R\beta=r)$ lineer modellerini gözönüne alalım. $X\beta$ nın $(Y,X\beta,V)$ altındaki en iyi lineer yansız tahmin edicisinin aynı zamanda $(Y,X\beta,V|R\beta=r)$ altında da en iyi lineer yansız tahmin edicisi olması için gerek ve yeter şart $R[V]\cap R[X]\subset R[X(I-R^{-}R)]$ olmasıdır.

Ispat: $X\beta$ nın $(Y,X\beta,V)$ altındaki en iyi lineer yansız tahmin edicisi $G = (V+XX')^{-}$ olmak üzere

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

dir. Açık olarak bu $(Y,X\beta,V|R\beta=r)$ altında da $X\beta$ için yansızdır. Bu nedenle $(Y,X\beta,V|R\beta=r)$ ve $(Y-XR^{-}r,XUt,V)$ modelleri denk olduğundan $X\hat{\beta}$ nın $X\beta$ nın $(Y,X\beta,V|R\beta=r)$ altında da en iyi lineer yansız tahmini olması için gerek ve yeter şart

$$X(X'G\,X)^{-}X'GV(XU)^{\perp} = 0$$

$$\Longleftrightarrow R[VGX]\subset R[XU] = R[X(I-R^{-}R)]$$

olmasıdır, burada herhangi bir A matrisi için $A^{\perp}$ $A^{\perp'}A=0$ şartını sağlayan maksimum ranklı matrisi göstermektedir. Öte yandan

$$R[VGX] = R[VGXX'] = R[V] \cap R[X]$$

olduğundan istenilen sonuç elde edilir.

Sonuç 2.3.1.[14]: Eğer $R[X] \subset R[V]$ ise bu takdirde $X\beta$ nın $(Y,X\beta,V)$ altındaki en iyi lineer yansiz tahmin edicisinin $(Y,X\beta,V|R\beta=r)$ altında da en iyi lineer yansız tahmin edicisi olması için gerek ve yeter şart $R[X'] \cap R[R'] = \{0\}$ olmasıdır.

Ispat: Eğer $R[X] \subset R[V]$ ise yukarıdaki teoremde verilen şart

$$R[X] = R[X(I-R^{-}R)] \Longleftrightarrow R[X'] \cap R[R'] = \{0\}$$

bağıntısına dönüşür, bu ise sonucun ispatını tamamlar.

Böylece $R[X'] \cap R[R'] = \{0\}$ şartı $X\beta$ nın $(Y,X\beta,V)$ altındaki en iyi lineer yansız tahmininin $(Y,X\beta,V|R\beta=r)$ altında da en iyi lineer yansız tahmini olması için yeter şart olmasına rağmen $R[X] \subset R[V]$ olmadıkça gerekmeyebilir. Örneğin,

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \end{bmatrix}, \quad X = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 1 & 0 \end{bmatrix}, \quad V = \begin{bmatrix} 4 & 2 & 2 \\ 2 & 1 & 1 \\ 2 & 1 & 1 \end{bmatrix}$$

$R = (1,-1)$ ve $r = 0$ alalım. Gerekli hesaplamalar yapılırsa

$$X\hat{\beta} = \begin{bmatrix} Y_1 \\ \frac{1}{2}(Y_1+Y_3) \\ \frac{1}{2}(Y_1+Y_2) \end{bmatrix} \quad \text{ve} \quad X\hat{\beta}_H = \begin{bmatrix} 1/3\ (2Y_1+Y_2+Y_3) \\ 1/6\ (2Y_1+Y_2+Y_3) \\ 1/6\ (2Y_1+Y_2+Y_3) \end{bmatrix}$$

olduğu görülür. Öte yandan $(Y,X\beta,V|R\beta=0)$ altında bir olasılıkla $X\hat{\beta} = X\hat{\beta}_H$ olmasına rağmen $R[X'] \cap R[R'] = R[R'] \neq \{0\}$ olduğu da gösterilebilir.

2.4. LİNEER MODELLERDE HİPOTEZ TESTİ

2.4.1. F-Testi

X nxm boyutlu ve m ranklı bir matris ve $e \sim N_n(0, \sigma^2 I_n)$ olmak üzere $Y = X\beta + e$ lineer modelini gözönüne alalım. R pxm boyutlu p ranklı bir matris ve r px1 boyutlu bir vektör olmak üzere $H: R\beta = r$ hipotezini test etmek istediğimizi varsayalım.

$$RSS = (Y - X\hat{\beta})'(Y - X\hat{\beta}) \quad [=(n-m)S^2]$$

ve

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

olsun, burada $\hat{\beta}_H$ (2.2.4) de verildiği gibidir ve RSS_H $R\beta = r$ altında e'e nin minimum değeridir. H hipotezinin test edilmesi için bir F-istatistiği aşağıdaki teoremle verilmektedir.

Teorem 2.4.1.[1]:

i) $RSS_H - RSS = (R\hat{\beta} - r)'[R(X'X)^{-1}R']^{-1}(R\hat{\beta} - r)$

ii) $E[RSS_H - RSS] = \sigma^2 p + (R\beta - r)'[R(X'X)^{-1}R']^{-1}(R\beta - r)$

iii) H doğru olduğunda

$$F = \frac{(RSS_H - RSS)/p}{RSS/(n-m)} = \frac{(R\hat{\beta} - r)'[R(X'X)^{-1}R']^{-1}(R\hat{\beta} - r)}{p \cdot S^2}$$

oranı $F_{p,n-m}$ dağılımına (sırasıyla p ve n-m serbestlik dereceli F-dağılımına) sahiptir.

iv) r = 0 olduğunda F,

$$F = \frac{n-m}{p} \quad \frac{Y'(P - P_H)\, Y}{Y'(I_n - P)\, Y}$$

formundadır, burada P_H simetrik idempotent olup $P_H P = P P_H = P_H$ dır.

Ispat: i) $RSS_H - RSS = (Y-X\hat{\beta}_H)'(Y-X\hat{\beta}_H) - (Y-X\hat{\beta})'(Y-X\hat{\beta})$

$$= \| Y-X\hat{\beta}_H \|^2 - \| Y-X\hat{\beta} \|^2 \quad [\ (2.2.7.)\ \text{den}\ ]$$

$$= \| X(\hat{\beta}-\hat{\beta}_H) \|^2$$

$$= (\hat{\beta}-\hat{\beta}_H)'X'X(\hat{\beta}-\hat{\beta}_H) \quad [\ (2.2.4.)\ \text{den}\ ]$$

$$= (R\hat{\beta}-r)'[R(X'X)^{-1}R']^{-1}R(X'X)^{-1}X'X(X'X)^{-1}R' \times [R(X'X)^{-1}R']^{-1}(R\hat{\beta}-r)$$

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

elde edilir.

ii) R nin satırları lineer bağımsız ve $\hat{\beta} \sim N_m(\beta, \sigma^2(X'X)^{-1})$ olduğundan $R\hat{\beta} \sim N_p(R\beta, \sigma^2 R(X'X)^{-1}R')$ dir. $Z=R\hat{\beta}-r$ ve $B=R(X'X)^{-1}R'$ olsun. Bu takdirde $E[Z] = R\beta-r$ ve $D[Z] = \sigma^2 B$ dir. Böylece

$$E[RSS_H - RSS] = E[Z'B^{-1}Z]$$

$$= iz[\sigma^2 B^{-1}B] + (R\beta-r)'B^{-1}(R\beta-r)$$

$$= iz[\sigma^2 I_p] + (r\beta-r)'B^{-1}(R\beta-r)$$

$$= \sigma^2 p + (R\beta-r)'[R(X'X)^{-1}R']^{-1}(R\beta-r) \qquad (2.4.1)$$

dir.

iii) $RSS_H - RSS$ $\hat{\beta}$ nın bir sürekli fonksiyonu olup RSS den bağımsızdır. H doğru olduğunda $R\hat{\beta} \sim N_p(r, \sigma^2 R(X'X)^{-1}R')$ olup

$$\frac{RSS_H - RSS}{\sigma^2} = (R\hat{\beta}-r)'(D[R\hat{\beta}])^{-1}(R\hat{\beta}-r)$$

χ^2_p dağılımına sahiptir. Öte yandan RSS/σ^2 χ^2_{n-m} dağılımına sahip idi. O halde H doğru olduğunda

$$F = \frac{(RSS_H - RSS)/\sigma^2 p}{RSS/\sigma^2(n-m)}$$

$[\chi^2_p/p]/[\chi^2_{n-m}/n-m]$ formundadır. Böylece H doğru olduğunda $F \sim F_{p,n-m}$ dir.

iv) $r=0$ olmak üzere (2.2.4) bağıntısını kullanarak

$$\hat{Y}_H = X\hat{\beta}_H = \{X(X'X)^{-1}X'-X(X'X)^{-1}R'[R(X'X)^{-1}R']^{-1}R(X'X)^{-1}X'\}Y$$

$$= (P-P_1)Y = P_HY, \text{diyelim}, \qquad (2.4.2)$$

dir, burada P_H simetriktir. Ayrıca P_1 in simetrik idempotent olduğu ve $P_1P = PP_1 = P_1$ olduğu kolayca görülebilir. Böylece

$$P_H^2 = P^2-P_1P-PP_1 + P_1^2$$

$$= P-2P_1+P_1 = P-P_1 = P_H \, , \qquad (2.4.3.)$$

$$P_HP = (P-P_1)P = P^2-P_1P = P-P_1 = P_H \qquad (2.4.4.)$$

ve transpoze alarak $PP_H = P_H$ elde edilir. Ispatı tamamlamak için $RSS = Y'(I_n-P)\,Y$ olduğunu biliyoruz. Benzer şekilde

$$RSS_H = (Y-X\hat{\beta}_H)'(Y-X\hat{\beta}_H) = Y'(I_n-P_H)^2Y = Y'(I_n-P_H)\,Y \qquad (2.4.5.)$$

elde edilir. Buradan da $RSS_H - RSS = Y'(P-P_H)\,Y$ bulunur. Bu ise ispatı bitirir.

Eğer H doğru ise bu takdirde $R\hat{\beta}$ ($R\beta$ nin en iyi lineer yansız tahmini) r ye yaklaşır ve RSS_H-RSS "küçük" olur. Bununla beraber $R\beta$ r den çok uzak ise RSS_H-RSS büyük bir değere ulaşır. Böylece F-testimiz bir tek yanlı test olup eğer F anlamlı olarak büyürse H reddedilir.

2.4.2. F - Testinin Uyumu

Her şeyden önce F yi nasıl seçeceğimiz sorusu ortaya çıkar. Testimizin uyumu için iki metod vereceğiz. RSS_H-RSS/p yi S_H^2 olarak tanımlayalim. Bu takdirde

$E[S_H^2] = \sigma^2+(R\beta-r)'[R(X'X)^{-1}R']^{-1}(R\beta-r)/p = \sigma^2+\delta$, diyelim, burada $\delta \geq 0$ dır. (Çünkü $[R(X'X)^{-1}R']^{-1} = D[R\hat{\beta}]/\sigma^2$ pozitif definittir.) Ayrıca $E[S^2]=\sigma^2$ idi. O halde H doğru olduğunda S_H^2 ve S^2 nin her ikisi de σ^2 nin yansız tahminleridir, yani $F = S_H^2/S^2 \approx 1$ dir. H yanlış olduğunda $\delta>0$ ve $E[S_H^2]>E[S^2]$ olup $E[F]=E[S_H^2]E[1/S^2]>E[S_H^2]/E[S^2]>1$ dir. Böylece eğer F anlamlı olarak büyürse H hipotezi reddedilir.

F-istatistiği H ın likelihood oran testini kullanarak da uyumlu yapılabilir. Söz konusu model için likelihood fonksiyonu, $L(\beta,\sigma^2)$ diyelim, Y nin ihtimal yoğunluk fonksiyonudur, yani

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

dir, $\partial\log L/\partial\beta = 0$ ve $\partial\log L/\partial\sigma^2 = 0$ dan

$$\hat{\beta} = (X'X)^{-1}X'Y \quad \text{ve} \quad \hat{\sigma}^2 = RSS/n \qquad (2.4.6)$$

maksimum likelihood tahminleri elde edilir ve L nin maksimumu, $L(\hat{\beta},\hat{\sigma}^2)$, $(2\pi\hat{\sigma}^2)^{-\frac{1}{2}n}e^{-\frac{1}{2}n}$ dir. Lagrange metoduna benzer bir metod kullanarak $R\beta=r$ kısıtlamalarına göre maksimum likelihood tahminleri $\hat{\beta}_H$ ve $\hat{\sigma}_H^2=RSS_H/n$ olur. Böylece L nin maksimum değeri

$$L(\hat{\beta}_H,\hat{\sigma}_H^2) = (2\pi\hat{\sigma}_H^2)^{-\frac{1}{2}n} e^{-\frac{1}{2}n}$$

dir. Likelihood oran istatistiği

$$\ell = L(\hat{\beta}_H,\hat{\sigma}_H^2)/L(\hat{\beta},\hat{\sigma}^2) = [\hat{\sigma}^2/\hat{\sigma}_H^2]^{\frac{1}{2}n}$$

olup likelihood oran prensibine göre ℓ çok küçük ise H reddedilir. Öte yandan

$$F = (n-m)/p\ (\ell^{-\frac{1}{2}n}-1) \qquad (2.4.7)$$

l nin monoton artan bir fonksiyonu olduğu için eğer F çok büyük ise H hipotezi reddedilir.

Örnek 2.4.1.[1]:

$$Y_1 = \alpha_1 + \alpha_2 + e_1$$
$$Y_2 = 2\alpha_2 + e_2$$
$$Y_3 = -\alpha_1 + \alpha_2 + e_3$$

olsun, burada $e_i \sim N(0,\sigma^2)$ olarak bağımsız dağılmışlardır. $H: \alpha_1 = 2\alpha_2$ hipotezini testetmek için bir F-istatistiği türetiniz?

Çözüm:

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 2 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \\ \alpha_2 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ e_3 \end{bmatrix} = X\beta + e$$

$$H:[1 \ -2]\begin{bmatrix}\alpha_1\\ \\ \alpha_2\end{bmatrix} = R\beta = 0$$

dir.

$$(X'X)^{-1} = \begin{bmatrix}1/2 & 0\\ \\ 0 & 1/6\end{bmatrix} \text{ ve } X'Y = \begin{bmatrix}1 & 0 & -1\\ \\ 1 & 2 & 1\end{bmatrix}\begin{bmatrix}Y_1\\ Y_2\\ Y_3\end{bmatrix} = \begin{bmatrix}Y_1-Y_3\\ \\ Y_1+2Y_2+Y_3\end{bmatrix}$$

$$\hat{\beta} = \begin{bmatrix}\alpha_1\\ \\ \alpha_2\end{bmatrix} = (X'X)^{-1}X'Y = \begin{bmatrix}1/2 & 0\\ \\ 0 & 1/6\end{bmatrix}\begin{bmatrix}Y_1-Y_3\\ \\ Y_1+2Y_2+Y_3\end{bmatrix} = \begin{bmatrix}1/2(Y_1-Y_3)\\ \\ 1/6(Y_1+2Y_2+Y_3)\end{bmatrix}$$

dır.

$$R(X'X)^{-1}R' = [1 \ -2]\begin{bmatrix}1/2 & 0\\ \\ 0 & 1/6\end{bmatrix}\begin{bmatrix}1\\ \\ -2\end{bmatrix} = [1/2 \ -1/3]\begin{bmatrix}1\\ \\ -2\end{bmatrix} = 7/6$$

olduğundan $[R(X'X)^{-1}R']^{-1} = 6/7$ dir. Ayrıca $r(R) = 1$ ve $r(X) = 2$ dir.

$$R\hat{\beta} = [\ 1 \ -2\ ]\begin{vmatrix}\hat{\alpha}_1\\ \\ \hat{\alpha}_2\end{vmatrix} = (\ \hat{\alpha}_1 - 2\hat{\alpha}_2\ )$$

olup

$$F = \frac{(R\hat{\beta})'[R(X'X)^{-1}R']^{-1}(R\hat{\beta})}{S^2} = \frac{7}{6}\ \frac{(\hat{\alpha}_1-2\hat{\alpha}_2)^2}{S^2}$$

dir, burada

$$S^2 = Y'Y - \hat{\beta}'X'X\,\hat{\beta} = Y_1^2 + Y_2^2 + Y_3^2 - (2\hat{\alpha}_1 + 6\hat{\alpha}_2)$$

dir.

Şimdi H için bir kanonik form vereceğiz. $Y=X\beta+e$ tam ranklı modeli için R pxm boyutlu p ranklı olmak üzere $H:R\beta=0$ hipotezini test etmek istediğimizi varsayalım. R p tane lineer bağımsız sütuna sahip olduğundan genelliği bozmaksızın bunların son p sütun olduğunu kabul edebiliriz; böylece R_2 pxp boyutlu nonsingüler bir matris olmak üzere $R=[R_1,R_2]$ dir. β parametre vektörünü de aynı yolda parcalayarak

$$0 = R\beta = R_1\beta_1 + R_2\beta_2$$

ve bu bağıntıyı R_2^{-1} ile çarparak

$$\beta_2 = -R_2^{-1}R_1\beta_1 \tag{2.4.8}$$

elde edilir. Bu ise H hipotezi altında regresyon modelinin

$$X\beta = (X_1,X_2)\beta = X_1\beta_1 + X_2\beta_2$$

$$(X_1-X_2R_2^{-1}R_1)\beta_1 = X_Rh, \text{diyelim}, \tag{2.4.9}$$

kanonik formunu aldığını gösterir, burada X_R nx(m-p) boyutlu m-p ranklı bir matris ve $h=\beta_1$ dır.

$$X_R\beta_1 = 0 \Longleftrightarrow X\beta = 0 \Longleftrightarrow \beta = 0 \Longleftrightarrow \beta_1 = 0$$

olduğundan X_R lineer bağımsız sütunlara sahiptir. X_R nın kolayca ve kusursuz olarak hesaplanabilmesi şartıyla $H:E[Y] = X_Rh$ modelini $e[Y]= X\beta$ orjinal modeli ile aynı formda ifade ederek aynı paket bilgisayar programı RSS ve RSS_H nin her ikisini de hesaplamada kullanılabilir.

X nxm boyutlu ve m ranklı olmak üzere $H:R\beta=r$ yi testetmek isteyelim, burada R pxm boyutlu ve p ranklıdır. Bu takdirde test istatistiğimiz

$$F = \frac{n-m}{p}\,\frac{RSS_H - RSS}{RSS}$$

dir, burada

$$RSS_H - RSS = (R\hat{\beta}-r)'[R(X'X)^{-1}R']^{-1}(R\hat{\beta}-r)$$

dir.

2.5. TAM RANKLI OLMAYAN TASARIM MATRİSİ

2.5.1. F-testi ve İzdüşüm Matrisleri

X nxm boyutlu matrisinin q (q<m) ranklı olduğu durumu gözönüne almadan önce F-testinin daha önce verilenden çok daha genel bir teorisini verelim. $\theta \in \Omega$ (R^n nin q boyutlu bir altuzayı) olmak üzere $Y=\theta+e$ modeline sahip olduğumuzu ve W Ω nın q-p boyutlu bir altuzayı olmak üzere $H:\theta \in W$ yı testetmek istediğimizi varsayalım. Bu takdirde aşağıdaki teoremi verebiliriz.

Teorem 2.5.1.[1]: H doğru ve $e \sim N_n(0,\sigma^2 I_n)$ olduğunda

$$F = \frac{(RSS_H-RSS)/p}{RSS/(n-q)} = \frac{e'(P_\Omega-P_W)e/p}{e'(I_n-P_\Omega)e/(n-q)} \sim F_{p,n-q}$$

dir. P_Ω ve P_W R^n yi sırasıyla Ω ve W ya izdüşüren simetrik idempotent matrislerdir.

Ispat: $\hat{\theta} = P_\Omega Y$ ve $\hat{\theta}_H = P_W Y$ θ nın respektif (ayrı ayrı) en küçük kareler tahminleri olsun. Bu nedenle

$$RSS = \| Y-\theta \|^2 = Y'(I_n-P_\Omega)\, Y$$

ve

$$RSS_H = Y'(I_n-P_W)\, Y$$

dir. $\theta \in \Omega$ olduğundan $(I_n-P_\Omega)\theta = 0$ olup dolayısıyla

$$RSS = (Y-\theta)'(I_n-P_\Omega)\,(Y-\theta) = e'(I_n-P_\Omega)\, e$$

dir. Benzer şekilde H doğru olduğunda $\theta \in W$ ve

$$RSS_H = e'(I_n-P_W)\, e$$

dir. I_n-P_Ω ve $P_\Omega-P_W$ $\Omega^\perp$ ve $W^\perp \cap \Omega$ üzerine izdüşümler olup bu matrisler simetrik ,idempotenttirler ve rankları da sırasıyla n-q ve q-(q-p)=p dir. Bu nedenle $e'(P_\Omega-P_W)\, e/\sigma^2$ ve $e'(I_n-P_\Omega)\, e/\sigma^2$ bağımsız olarak $X^2{}_p$ ve $X^2{}_{n-q}$ dağılımlarına sahiptirler. Böylece $F\sim F_{p,n-q}$ dir.

2.5.2. Testedilebilir Hipotezler

$Y = \theta + e$ olsun, burada $\theta = X\beta$ ve X nxm boyutlu ve q (q<m) ranklıdır. R pxm boyutlu ve p ranklı olmak üzere $H{:}R\beta = 0$ ı test etmek isteyelim. X tam ranklı olmadığından H ın "testedilebilir" olup olmaması problemi ortaya çıkar. Örneğin, eğer R nin satırları X in satırlarından lineer bağımsız (ve dolayısıyle $p \leq m-q$) ise bu takdirde her $\theta \in R[X]$ için $\theta = X\beta$ ve $R\beta = 0$ olacak şekilde en az bir β mevcuttur; eğer $p = m-q$ ise bu β tektir. Bu durumda $R\beta = 0$ denklemleri β için basitçe belirlenebilen kısıtlamalardır ve dolayısıyle θ R[x] in bir uygun altkümesi ile kısıtlı değildir. Bu nedenle eğer a_i' R nin i-yinci satırı ise herhangi bir $a_i'\beta = 0$ denklemini önemsemeyebiliriz. a_i' X in satırlarından lineer bağımsız olup testedilebilir hipotezimiz geri kalan denklemlerden ibaret olacaktır. Şimdi aşağıdaki tanımı verelim.

Tanım 2.5.1.[1]: Eğer R nin satırları X in satırlarından lineer bağımsız ise yani

$$R = MX \tag{2.5.1}$$

olacak şekilde pxn boyutlu bir M matrisi mevcut ise bu takdirde $H{:}R\beta = 0$ hipotezine "testedilebilir" denir.

Bu tanım R nin satırları lineer bağımsız olmadığında uygulanır. Bununla beraber eğer R pxm boyutlu p ranklı ise bu takdirde $r[R] \leq r[M]$ olacağından M p ranklı olmalıdır. Bu tanım $r \neq 0$ ve $r \in R[R]$ olmak üzere $R\beta = r$ daha genel durumunada uygulanabilir. Bunu göstermek için orijini değiştirip modeli ve hipotezi

$$\tilde{Y} = Xu + e \text{ ve } Ru = 0 \tag{2.5.2}$$

ye indirgeriz, burada $u = \beta - \beta_0$, β_0 $R\beta = r$ nin bir çözümü ve $\tilde{Y} = Y - X\beta_0$ dir.

Açık olarak orijinal hipotezin testedilebilir olması için gerek ve yeter şart dönüştürülmüş hipotezin testedilebilir olmasıdır.

Teorem 2.5.2.[1]: R kxm boyutlu p ($p \leq q$) ranklı olmak üzere H: $R\beta = r$ testedilebilir olsun. RSS ve RSS_H $e'e = \| Y-\theta \|^2$ nin sırasıyle kısıtlamasız ve kısıtlamalı minimum değerleri olsun. Bu takdirde

i) H doğru olduğunda

$$F = \frac{(RSSS_H-RSS)/p}{RSS/(n-q)} \sim F_{p,n-q}$$

dir.

ii) $RSS_H-RSS = (R\hat{\beta}-r)'[R(X'X)^-R']^-(R\hat{\beta}-r)$

dir, burada $\hat{\beta}$ $X'X\beta = X'Y$ nin herhangi bir çözümüdür.

Ispat: i) önce modeli ve hipotezi (2.5.2) ye dönüştürelim. Böylece modelimiz $\theta \in R[X]$ $(=\Omega)$ olmak üzere $\tilde{Y} = \theta+e$ dir ve $M\theta = MXu = Ru = 0$ olduğundan H hipotezi $\theta \in W = \mathcal{N}[M] \cap \Omega$ olur. Kısım 2.2.2. den $W^{\perp} \cap \Omega = R[P_{\Omega}M']$ p boyutludur. Ayrıca $(I_n-P_{\Omega})X = 0$ olduğundan

$$(I_n-P_{\Omega})\tilde{Y} = (I_n-P_{\Omega})(Y-X\beta + X(\beta-\beta_0)) = (I_n-P_{\Omega})e$$

dir. H doğru olduğunda $(P_{\Omega}M')'Xu = MXu = Ru = 0$ olduğundan

$$(P_{\Omega}-P_W)\tilde{Y} = P_{W^{\perp}\cap\Omega}(Y-X\beta+Xu) = P_{W^{\perp}\cap\Omega}\, e$$

dir. Istenilen sonuç Teorem 2.5.1. i hatırlayarak elde edilir.

ii)
$$\begin{aligned} R(X'X)^-X'\tilde{Y} &= MP_{\Omega}(Y-X\beta_0) \\ &= MP_{\Omega}Y-MX\beta_0 \\ &= MX\hat{\beta}-R\beta_0 = R\hat{\beta}-r \end{aligned}$$

olduğundan Kısım 2.2.2. deki (2.2.10) bağıntısına göre

$$\begin{aligned} RSS_H-RSS &= \tilde{Y}'P_{W^{\perp}\cap\Omega}\tilde{Y} \\ &= \tilde{Y}'X(X'X)^-R'[R(X'X)^-R']^-R(X'X)^-X'\tilde{Y} \\ &= (R\hat{\beta}-r)'[R(X'X)^-R']^-(R\hat{\beta}-r) \end{aligned}$$

elde edilir.

Yukarıdaki sonuçlar g-inversin özelliklerini kullanarak da daha uzun bir yolda ispatlanabilirdi. Pratik olarak teoremin i) kısmı RSS ve RSS_H direkt olarak türev almak suretiyle hatta çoğu varyans analizi ifadelerinde olduğu gibi tetkik (yoklama) yapmak suretiyle elde edilebildiğinden oldukça önemlidir.

2.6. ÖN KISITLAMALI HİPOTEZ TESTİ

İlk adımda $Y = X\beta + e$ tam ranklı modelini gözönüne alacağız, burada X nxm boyutlu m ranklıdır fakat C kxm boyutlu k ranklı olmak üzere bir $C\beta = 0$ ilave kısıtlaması konuluyor. R pxm boyutlu p ranklı ve R nin satırları C nin satırlarından lineer bağımsız (böylece $p + k \leq m$) olmak üzere $H:R\beta = 0$ hipotezini test etmek isteyelim. Bu takdirde Kısım 2.5.1. deki notasyonu kullanarak $C\beta = C(X'X)^{-1}X'\theta$ olduğundan $\Omega = R[X] \cap \mathcal{N}[C(X'X)^{-1}X']$ ve $W = \mathcal{N}[R(X'X)^{-1}X'] \cap \Omega$ elde edilir. W ve R olarak sırasıyla Ω ve $(R',C')'$ alınarak Ω ve W nin sırasıyla m-k ve p boyutlu olduğu elde edilir. Bu nedenle H doğru olduğunda

$$F = \frac{(RSS_H - RSS)/p}{RSS/(n-(m-k))} \sim F_{p,n-m+k}$$

elde edilir. Bu sonuç $H:R\beta = c$ $(c \neq 0)$ durumu için yine sağlanacaktır.

$$\begin{bmatrix} R \\ C \end{bmatrix} \beta_0 = \begin{bmatrix} c \\ 0 \end{bmatrix}$$

olacak şekilde β_0 bulunur ve daha önceki kısımda olduğu gibi $\tilde{Y} = Y - X\beta_0$ yazılır.

BÖLÜM III

EŞİTSİZLİK KISITLAMALI LİNEER MODELLER

İstatistiksel pratikte tahmin ve ön kestirim için pek çok model ve ilgili metodlar geliştirilmiş ve uygulanmış olmasına rağmen günümüzde lineer regresyon modelleri hâlâ büyük bir güvene sahiptir. Bunun bir sebebi, bunların ekonomi ve diğer bilimlerdeki değişik yöntemlerin yeterli bir yaklaşımını mümkün kılmalarıdır. Eğer yalnızca örneklem bilgisi mevcut ise, bu takdirde Gauss ve Aitken'e uygun tahminler en kullanışlılarıdır. Bununla birlikte çoğu pratiksel problemlerde örneklem bilgisine ilaveten bilinmeyen parametre vektörünün bileşenleri hakkında bazı ön bilgiler vardır. Parametreler üzerinde eşitsizlik kısıtlamalı lineer modeller pratikte sıksık karşılaşılabilen modellerdir. Y nxl boyutlu bir gözlemler vektörü, X bilinen değerlerin nxm boyutlu bir matrisi, β bilinmeyen fakat tahmin edilebilir parametrelerin mxl boyutlu bir vektörü e ise $E[e] = 0$ ve $E[ee'] = \Sigma$ kovaryans matrisli gözlenemeyen hataların bir nxl boyutlu vektörü olmak üzere

$$Y = X\beta + e \qquad (3.1.1)$$

genel lineer modelini gözönüne alalım. β parametre vektörü üzerinde

$$R\beta \geq r \qquad (3.1.2)$$

gibi bir kısıtlama konulursa (3.1.1) modeline " eşitsizlik kısıtlamalı genel lineer model" denir. Burada R pxm ($p \leq m$) boyutlu bilinen bir matris ve r de pxl boyutlu bilinen bir vektördür. (3.1.2) kısıtlaması altında (3.1.1) lineer modelini

$$(Y, X\beta, \Sigma \mid R\beta \geq r) \qquad (3.1.3)$$

ile gösterelim. R matrisinin tam satır ranklı olduğu kabul edilecektir. S kısıtlanmış kümesini

$$S = \{\beta \in R^m : R\beta \geq r\} \qquad (3.1.4)$$

olarak tanımlarız. Σ kovaryans matrisi, I_n birim matris ve σ^2 pozitif bir skaler olmak üzere, $\sigma^2 I_n$ skaler matrisi olarak kabul edildiğinde (3.1.1) modeli bir " standart " lineer modeldir. Bu takdirde (3.1.1) modelindeki β nın $R\beta \geq r$ eşitsizlik kısıtlaması altındaki kısıtlamalı en küçük kareler tahmini

$$\min_{\beta} \{ \| Y-X\beta \|^2 : R\beta \geq r \} \qquad (3.1.5)$$

minimum - norm problemine dönüşür, burada $\| \cdot \|$ standart Euclidian uzaklığıdır. " Genelleştirilmiş " lineer model durumunda β nın $R\beta \geq r$ kısıtlaması altındaki en küçük kareler tahmini ise

$$\min_{\beta} \{ \| Y-X\beta \|^2_{\Sigma} : R\beta \geq r\} \qquad (3.1.6)$$

olarak verilir. Burada $\| w \|^2_{\Sigma} = w'\Sigma^{-1}w$ dir. (3.1.5) ve (3.1.6) bağıntılarının her ikisi de alışılmış kuadratik programlama problemi olup β nın kısıtlamalı tahmini $\tilde{\beta}$ için değişik algoritmalar uygulayarak çözülebilir. (3.1.4) bağıntısını kullanarak (3.1.6) bağıntısını

$$\min_{\beta \in S} \{ \| Y-X\beta \|^2_{\Sigma} : \} \qquad (3.1.7)$$

olarak yazabiliriz. β nın kısıtlamasız en küçük kareler tahmini $\hat{\beta}$

$$\min_{\beta} \{ \| Y-X\beta \|^2_{\Sigma} \} \qquad (3.1.8)$$

için bir çözüm olup X in tam sütun ranklı olması durumunda (3.1.8) için yegâne çözüm

$$\hat{\beta} = (X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}Y \qquad (3.1.9)$$

dır. Böylece $\tilde{\beta}$ $\hat{\beta}$ nın S üzerindeki bir izdüşümüdür. Bu (3.1.1) modelini $(X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}$ ile önden çarpmak suretiyle gösterilebilir. Bu durumda (3.1.1) modeli denk olarak $\delta=(X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}e$ olmak üzere

$$\hat{\beta} = \beta + \delta \qquad (3.1.10)$$

biçiminde yazılır, burada δ 0 ortalaması ve $\Gamma = (X'\Sigma^{-1}X)^{-1}$ kovaryans matrisine sahiptir. (3.1.10) modelini kullanarak $\tilde{\beta}$ nın

$$\min_{\beta \in S}\{ \| \hat{\beta}-\beta \|^2_\Gamma\} \tag{3.1.11}$$

için bir çözüm olduğu elde edilir. Böylece $\tilde{\beta}$, $\hat{\beta}$ nın Γ ya göre S üzerindeki bir izdüşümüdür. Ayrıca X tam sütun ranklı ve S konveks ise bu $\tilde{\beta}$ tek türlü belirlidir.

(3.1.1) modelinin diğer bir dönüşümü ise (3.1.10) bağıntısını R ile çarpıp r yi çıkarmak suretiyle türetilebilir:

$$(R\hat{\beta}-r) = (R\beta-r) +\tau \tag{3.1.12}$$

olup burada $\tau=R\delta$ sıfır ortalamasına ve $\Omega = R\Gamma R'$ kovaryans matrisine sahiptir. $\hat{\mu} = R\hat{\beta}-r$ ve $\mu = R\beta-r$ dersek (3.1.1) modeli

$$\hat{\mu} = \mu + \tau \tag{3.1.13}$$

olarak yazılabilir. Test yapmak amacıyle (3.1.1) modelindeki $R\beta-r \geq 0$ kısıtlaması (3.1.13.) modelinde $\mu \geq 0$ kısıtlamasına denktir. (3.1.13) modeli için kısıtlanmış altküme R^p deki C pozitif orthant'dır:

$$C = \{\mu \in R^p \mid \mu \geq 0\} \tag{3.1.14}$$

$\tilde{\mu}$ (3.1.13) modelinde μ nün $\mu \geq 0$ kısıtlaması altında kısıtlamalı en küçük kareler tahmini olsun. Bu durumda $\tilde{\mu}$

$$\min_{\mu \in C} \| \hat{\mu} - \mu \|^2_\Omega \tag{3.1.15}$$

için bir çözümdür, yani $\tilde{\mu}$ $R\hat{\beta}-r$ nin Ω ya göre C üzerindeki bir izdüşümüdür. X tam sütun ranklı olduğundan $\tilde{\mu}$ tektir.

Tanım 3.1.1.[7]: Γ mxm boyutlu pozitif definit bir matris olsun. R^m üzerinde Γ ya göre bir iç çarpımı her $\alpha,\beta \in R^m$ için $(\alpha|\beta)_\Gamma=\alpha'\Gamma^{-1}\beta$ ile tanımlayalım. Bu iç çarpıma karşılık gelen norm her $\beta \in R^m$ için $\|\beta\|^2_\Gamma = \beta'\Gamma^{-1}\beta$ nin pozitif karekökü olarak tanımlanır. $T(\beta) = R\beta - r$ ile $T: R^m \longrightarrow R^p$ dönüşümünü ve R^m nin bir altuzayını

$$\ker R = \{\beta \in R^m : R\beta = 0\} \tag{3.1.16}$$

tanımlayalım. kerR nin R^m deki ortogonal komplementi

$$(\ker R)^{\perp} = \{\alpha \in R^m : (\alpha|\beta)_\Gamma=0 \ , \ \forall\beta\in\ker R\} \tag{3.1.17}$$

altuzayı ile tanımlanır. R^m nin bir direkt-toplam parçalanışı

$$R^m = (\ker R) \oplus (\ker R)^{\perp} \tag{3.1.18}$$

ile verilir. Her $\beta\in R^m$ vektörü $\tau\in\ker R$ ve $\delta\in(\ker R)^{\perp}$ vektörlerinin toplamı olarak tektürlü parçalanabilir. Yani $\beta = \tau + \delta$ dir.

Lemma 3.1.1.[7]: $T(S) = C$ dir.

Ispat: $\beta\in S$ olsun. Bu takdirde $T(\beta)=R\beta-r \geq 0$ olduğundan $T(\beta)\in C$ dir. Dolayısıyla $T(S)\subset C$ elde edilir. Tersine olarak $\mu\in C$ olsun. R tam satır ranklı olduğundan R^m den R^p ye bir lineer dönüşüm olarak T örtendir. Bu nedenle $R\beta=\mu+r$ olacak şekilde bir $\beta\in R^m$ mevcuttur. Bu takdirde $T(\beta)=R\beta-r=\mu\geq 0$ olup $\beta\in S$ ve $\mu=T(\beta)$ dır. Böylece $C\subset T(S)$ sağlanır.

Lemma 3.1.2.[7]: $\beta\in R^m = (\ker R) \oplus (\ker R)^{\perp}$ olsun. R^+ ve δ yı

$$R^+ = \Gamma R'(R\Gamma R')^{-1} \quad \text{ve} \quad \delta = R^+R\beta \tag{3.1.19}$$

olarak tanımlayalım. Bu takdirde δ β nın kerR ve $(\ker R)^{\perp}$ deki tek ayrışımı için β nın $(\ker R)^{\perp}$ deki bileşenidir.

Ispat: $RR^+ = I$ olduğunu belirtelim. Bu takdirde $R\delta = RR^+R\beta = R\beta$ olup $R(\beta-\delta) = 0$ elde edilir. Böylece $\beta-\delta\in\ker R$ dir. β yı $\beta=(\beta-\delta)+\delta$ olarak yazalım ve $\tau\in\ker R$ olduğundan $(\tau|\delta)_\Gamma = 0$ olduğunu görmek yeter:

$$(\tau|\delta)_\Gamma = \tau'\Gamma^{-1}\delta = \tau'\Gamma^{-1}R^+R\beta = \tau'\Gamma^{-1}\Gamma R'(R\Gamma R')^{-1}R\beta$$
$$= (R\tau)'(R\Gamma R')^{-1}(R\beta) = 0$$

olup buradaki son eşitlik $R\tau = 0$ a uygundur.

Lemma 3.1.3.[7]: $\delta_1,\delta_2\in(\ker R)^{\perp}$ ise bu takdirde $\Omega=R\Gamma R'$ olmak üzere

$$(R\delta_1|R\delta_2)_\Omega = (\delta_1|\delta_2)_\Gamma \tag{3.1.20}$$

Ispat: R tam satır ranklı ve Γ pozitif definit olduğundan Ω

pozitif definittir. Lemma 3.1.2. den R^{+} (3.1.19.) da tanımlandığı gibi olmak üzere $i=1,2$ için $\delta_i = R^{+}R\delta_i$ dir. Bu takdirde

$$\begin{aligned}(\delta_1|\delta_2)_\Gamma = \delta_1'\Gamma^{-1}\delta_2 &= (R^{+}R\delta_1)'\Gamma^{-1}(R^{+}R\delta_2)\\ &= (R\delta_1)'(R^{+})'\Gamma^{-1}(R^{+})(R\delta_2)\\ &= (R\delta_1)'(R\Gamma R')^{-1}(R\delta_2)\\ &= (R\delta_1|R\delta_2)_\Omega\end{aligned}$$

dir.

Sonuç 3.1.1.[7]: Eğer $\delta\in(\ker R)^{\perp}$ ise bu takdirde

$$\| R\delta \|_\Omega = \| \delta \|_\Gamma \qquad (3.1.21)$$

dir.

Ispat: Lemma 3.1.3. den direkt olarak elde edilir.

Lemma 3.1.4.[7]: $\beta\in S$ olsun. β, $\hat\beta$ ve $\tilde\beta$ yı $\ker R$ ve $(\ker R)^{\perp}$ teki tek ayrışımlarıyla tanımlayalım: $\beta = \tau + \delta$, $\hat\beta = \hat\tau + \hat\delta$ ve $\tilde\beta = \tilde\tau + \tilde\delta$ olsun, burada $\tau,\hat\tau,\tilde\tau\in\ker R$ ve $\delta,\hat\delta,\tilde\delta\in(\ker R)^{\perp}$ dir. Bu takdirde her $\beta\in S$ için

$$\| \hat\beta-\tilde\beta \|_\Gamma \le \| \hat\delta-\delta \|_\Gamma \qquad (3.1.22)$$

dir ve ayrıca

$$\| \hat\beta-\tilde\beta \|_\Gamma = \| \hat\delta-\tilde\delta \|_\Gamma \qquad (3.1.23)$$

eşitliği gerçeklenir.

Ispat: $\tilde\beta$ nın tanımına göre her $\beta\in S$ için $\| \hat\beta-\tilde\beta \|_\Gamma \le \| \hat\beta-\beta \|_\Gamma$ dir. $\beta\in S$ yi sabit tutup $\tau\in\ker R$ ve $\delta\in(\ker R)^{\perp}$ olmak üzere $\beta = \tau + \delta$ tanımlayalım. Bu takdirde $r \le R\beta = R\tau + R\delta = R\delta$ olup $\delta\in S$ dir. Bu nedenle $R(\hat\tau+\delta) = R\delta \ge r$ olup $\hat\tau + \delta \in S$ dir. Böylece

$$\| \hat\beta-\tilde\beta \|_\Gamma \le \| \hat\beta-(\hat\tau+\delta) \|_\Gamma = \| (\hat\tau+\hat\delta)-(\hat\tau+\delta) \|_\Gamma = \| \hat\delta-\delta \|_\Gamma$$

elde edilir. Öte yandan

$$\begin{aligned}\| \hat\delta-\tilde\delta \|^2_\Gamma \ge \| \hat\beta-\tilde\beta \|^2_\Gamma &= \| (\hat\tau-\tilde\tau)+(\hat\delta-\tilde\delta) \|^2_\Gamma\\ &= \| \hat\tau-\tilde\tau \|^2_\Gamma + \| \hat\delta-\tilde\delta \|^2_\Gamma \qquad (3.1.24)\end{aligned}$$

dir. Buradan $\| \hat\tau-\tilde\tau \|_\Gamma = 0$ ve $\hat\tau = \tilde\tau$ elde edilir. Bunu (3.1.24) de yerine koyarak $\| \hat\beta-\tilde\beta \|_\Gamma = \| \hat\delta-\tilde\delta \|_\Gamma$ olduğu görülür.

Lemma 3.1.5.[7]: Yukarıdaki tanımlar verildiğinde $R\tilde{\beta}-r$, $R\hat{\beta}-r$ nin Ω ya göre C üzerindeki izdüşümüdür. Burada $\Omega = R\Gamma R'$ dir.

Ispat: $\| T\hat{\beta}-T\tilde{\beta} \|_{\Omega} = \min\{ \| T\hat{\beta}-\mu \|_{\Omega} : \mu \in C\}$ olduğunu göstermek yeter. $\mu \in C$ keyfi olsun. Bu takdirde Lemma3.1.1. e göre $\mu = T(\beta)$ olacak şekilde bir $\beta \in S$ mevcuttur. O halde

$$\| T\hat{\beta}-\mu \|_{\Omega} = \| T\hat{\beta}-T\beta \|_{\Omega} = \| R\hat{\beta}-R\beta \|_{\Omega} = \| R\hat{\delta}-R\beta \|_{\Omega}$$

$$= \| \hat{\delta}-\delta \|_{\Gamma} \geq \| \hat{\beta}-\tilde{\beta} \|_{\Gamma} = \| \hat{\delta}-\tilde{\delta} \|_{\Gamma} = \| R\hat{\delta}-R\tilde{\delta} \|_{\Omega}$$

$$= \| T\hat{\beta}-T\tilde{\beta} \|_{\Omega}$$

elde edilir.

3.2. EŞİTSİZLİK KISITLAMALI TAHMİN EDİCİNİN ÖZELLİKLERİ

$$Y = X\beta + e \ , \ e \sim N_n(0, \sigma^2 I_n) \qquad (3.2.1)$$

standard lineer modelini gözönüne alalım, burada Y,X ve β (3.1.1) modelinde tanımlandığı gibidir. Ayrıca X in tam sütun ranklı olduğunu kabul edelim. β parametre vektörü üzerinde R tam satır ranklı olmak üzere (3.1.2) lineer kısıtlaması verilsin. Bu takdirde eşitsizlik kısıtlamalı en küçük kareler (ICLS) tahmin edicisi $\tilde{\beta}$

$$\min (Y-X\beta)'(Y-X\beta) \ : \ R\beta \geq r \qquad (3.2.2)$$

kısıtlamalı kuadratik programlama problemi için çözümdür. β nın kısıt lamasız en küçük kareler tahmin edicisi $\hat{\beta} = (X'X)^{-1}X'Y$ olup buna β nın alışılmış en küçük kareler (OLS) tahmin edicisi denir. (3.2.2) minimumlaştırma problemindeki objektif fonksiyon bir konveks fonksiyon olduğundan aşağıda verilen "Kuhn-Tucker " şartlarını uygulayarak $\tilde{\beta}$ nın (3.2.2) için bir çözüm olması için gerek ve yeter şart

$$\partial L/\partial\lambda \geq 0 \text{ veya } R\tilde{\beta}-r \geq 0 \tag{3.2.3}$$

$$\partial L/\partial\beta = 0 \text{ veya } X'X\tilde{\beta}-X'Y-R'\lambda_0 = 0 \tag{3.2.4}$$

$$\lambda_0'\,\partial L/\partial\lambda = 0 \text{ veya } \lambda_0'(R\tilde{\beta}-r) = 0 \tag{3.2.5}$$

olacak şekilde bir $\lambda_0 \geq 0$ vektörünün bulunmasıdır, burada

$$L(\beta,\lambda) = (Y-X\beta)'(Y-X\beta)-2\lambda'(R\beta-r) \tag{3.2.6}$$

olup türevler $(\tilde{\beta},\lambda_0)$ de hesaplanmıştır. (3.2.6) bağıntısı ile tanımlanan fonksiyona "Lagrangian fonksiyonu " denir.

Eğer $\hat{\beta} = (X'X)^{-1}X'Y$ alışılmış en küçük kareler tahmin edicisi (3.2.2) deki kısıtlamaları sağlarsa yani $R\hat{\beta}\geq r$ ise bu takdirde λ_0 vektörü olarak sıfır vektörünü almak yeterlidir ve bu durumda (3.2.2) için çözüm $\tilde{\beta} = \hat{\beta}$ dır.

Şimdi eşitsizlik kısıtlamalı en küçük kareler tahmin edicisi için bir kapalı form çözüm vereceğiz. Ancak daha önce bazı notastonları tanımlayalım. $\Delta = \{1,2,\ldots,p\}$ nin bir altkümesini s ile gösterelim. s nin Δ ya göre bütünleyenini $\bar{s}$ ile gösterelim. m_i ler sütun vektörleri olmak üzere her hangi bir $M = (m_1,m_2,\ldots,m_p)$ matrisi için $j\notin s$ ise M nin j-yinci sütununu silmek suretiyle elde edilen matrisi M_s ile gösterelim. örneğin $s = \Delta$ ise $M_s = M$ ve $s = \emptyset$, boş küme , ise M_s sıfır matrisidir. M_s nin transpozunu M_s' ile gösterelim. U matrisini

$$U'U = R(X'X)^{-1}R' \tag{3.2.7}$$

ile tanımlayalım, burada U pozitif köşegen elemanlarına sahip bir reel üst üçgensel matristir. Ayrıca $V = -(U')^{-1}$ olsun.Buradan $U'V = -I$ elde edilir. Son olarak Δ nın her bir s altkümesi için C_s konisini

$$C_s = \{ a \mid a\in R^p,\ a = V_s n_0 + U_{\bar{s}} m_0 \ ,\ n_0\leq 0 \ ,\ m_0\leq 0 \}$$

olarak tanımlayalım. C_s kümelerinin birleşiminin R^p olduğu aşikârdır.

Aşağıdaki teorem $\tilde{\beta}$ eşitsizlik kısıtlamalı en küçük kareler tahmin edicisi için bir kapalı form verir.

Teorem 3.2.1.[6]: $\hat{\beta} = (X'X)^{-1}X'Y$ β nın alışılmış en küçük kareler tahmin edicisi olsun. Eğer $\hat{a} = V(r-R\hat{\beta})\epsilon C_S$ ise bu takdirde (3.2.2) minimumlaştırma probleminin çözümü olan $\tilde{\beta}$ eşitsizlik kısıtlamalı en küçük kareler tahmin edicisi

$$\tilde{\beta} = \hat{\beta}-(X'X)^{-1}R_{\not{s}}'[R_{\not{s}}(X'X)^{-1}R_{\not{s}}']^{-1}(R_{\not{s}}\hat{\beta}-I_{\not{s}}'r) \qquad (3.2.8)$$

dir, burada I pxp boyutlu bir birim matristir.

Ispat: (3.2.3) - (3.2.5) gerek ve yeter şartlarını sağlayan bir $\lambda_0 \geq 0$ vektörünün mevcut olduğunu göstereceğiz. önce aşağıdaki bağıntıları verelim. $\hat{a}\epsilon C_S$ olduğundan

$$R\hat{\beta}-r = U'\hat{a} = U'V_S n_0 + U'U_{\not{s}}m_0 = -I_S n_0 + U'U_{\not{s}}m_0 \qquad (3.2.9)$$

olacak şekilde $n_0 \leq 0$ ve $m_0 \leq 0$ vektörleri mevcuttur. (3.2.7) bağıntısına göre

$$R_{\not{s}}(X'X)^{-1}R_{\not{s}}' = U_{\not{s}}'U_{\not{s}} \qquad (3.2.10)$$

ve

$$R(X'X)^{-1}R_{\not{s}}' = U'U_{\not{s}} \qquad (3.2.11)$$

dir. (3.2.9) ve (3.2.10) bağıntılarını kullanarak

$$\begin{aligned}[R_{\not{s}}(X'X)^{-1}R_{\not{s}}']^{-1}(R_{\not{s}}\hat{\beta}-I_{\not{s}}'r) &= (U_{\not{s}}'U_{\not{s}})^{-1}I_{\not{s}}'(-I_S n_0 + U'U_{\not{s}}m_0)\\ &= (U_{\not{s}}'U_{\not{s}})^{-1}(U_{\not{s}}'U_{\not{s}}m_0-I_{\not{s}}'I_S n_0)\\ &= m_0 \qquad (3.2.12)\end{aligned}$$

dır, çünkü $I_{\not{s}}'I_S n_0 = 0$ dır. Şimdi teoremin ispatına geçelim.

a) önce $r-R\tilde{\beta} \leq 0$ olduğunu gösterelim. Bunun için $\tilde{\beta}$ nın (3.2.8) deki ifadesini ve (3.2.9)-(3.2.11) bağıntılarını kullanarak

$$\begin{aligned}R\tilde{\beta}-r &= (R\hat{\beta}-r)-R(X'X)^{-1}R_{\not{s}}'[R_{\not{s}}(X'X)^{-1}R_{\not{s}}']^{-1}(R_{\not{s}}\hat{\beta}-I_{\not{s}}'r)\\ &= -I_S n_0 + U'U_{\not{s}}m_0-U'U_{\not{s}}m_0\\ &= -I_S n_0 \geq 0\end{aligned}$$

elde edilir. Böylece (3.2.3) şartı sağlanır.

b) $X'X\tilde{\beta}-X'Y-R'\lambda_0 = 0$ olduğunu göstermek için λ_0 ı

$$\lambda_0 = -I_S m_0 \qquad (3.2.13)$$

olarak tanımlayalım. $m_0 \leq 0$ olduğundan $\lambda_0 \geq 0$ dır. Böylece

$$X'X\tilde{\beta}-R'\lambda_0-X'Y = X'X\tilde{\beta} + R'I_{\not{s}}m_0-X'Y$$
$$= X'X[\hat{\beta}-(X'X)^{-1}R_{\not{s}}'[R_{\not{s}}(X'X)^{-1}R_{\not{s}}']^{-1}(R_{\not{s}}\hat{\beta}-I_{\not{s}}'r)]$$
$$+ R'I_{\not{s}}m_0-X'Y$$
$$= X'Y-R_{\not{s}}'[R_{\not{s}}(X'X)^{-1}R_{\not{s}}']^{-1}(R_{\not{s}}'\hat{\beta}-I_{\not{s}}'r)+R_{\not{s}}'m_0-X'Y$$
$$=-R_{\not{s}}'m_0 + R_{\not{s}}'m_0 = 0$$

elde edilir. Böylece (3.2.4) şartı gerçeklenmiş olur.

c) Şimdi $\lambda_0'(R\tilde{\beta}-r) = 0$ olduğunu gösterelim. (3.3.13) bağıntısını ve a) yı kullanarak

$$\lambda_0'(R\tilde{\beta}-r) = (-I_{\not{s}}m_0)'(-I_s n_0) = m_0'I_{\not{s}}'I_s n_0 = 0$$

olup (3.2.5) sağlanır. Böylece teoremin ispatı tamamlanır.

örnek 3.2.1.[6]: j=1,2,3 için $\beta_j \geq 0$ kısıtlamaları altında

$$Y_i = \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + e_i \quad (i=1,2,3)$$

lineer regresyon uydurma örneğini ele alalım, burada R 3x3 boyutlu bir birim matris r 3x1 boyutlu bir sıfır vektörü ve X ve Y gözlem vektörleri de

$$X = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 3 \\ 1 & 2 & 0 \end{bmatrix} \quad \text{ve} \quad Y = \begin{bmatrix} 2 \\ -3 \\ 6 \end{bmatrix}$$

dır. O halde (X'X) in inversi

$$U'U = (X'X)^{-1} = 1/49 \begin{bmatrix} 41 & -17 & 1 \\ -17 & 19 & -4 \\ 1 & -4 & 6 \end{bmatrix}$$

dir. Buradan gerekli hesaplamalar yapılarak

$$U = \begin{bmatrix} .915 & -.379 & .022 \\ 0 & .494 & -.148 \\ 0 & 0 & .316 \end{bmatrix}, \quad V = -\begin{bmatrix} 1.093 & 0 & 0 \\ .840 & 2.025 & 0 \\ .316 & .949 & 3.162 \end{bmatrix}$$

bulunur. öte yandan

$$X'Y = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 3 \\ 1 & 2 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ -3 \\ 6 \end{bmatrix} = \begin{bmatrix} 8 \\ 9 \\ -7 \end{bmatrix},$$

$$\hat{\beta} = (X'X)^{-1}X'Y = 1/49 \begin{bmatrix} 41 & -17 & 1 \\ -17 & 19 & -4 \\ 1 & -4 & 6 \end{bmatrix} \begin{bmatrix} 8 \\ 9 \\ -7 \end{bmatrix} = \begin{bmatrix} 3.429 \\ 1.286 \\ -1.429 \end{bmatrix}$$

ve

$$\hat{a} = V(r-R\hat{\beta}) = -V\hat{\beta} = \begin{bmatrix} 1.093 & 0 & 0 \\ .840 & 2.025 & 0 \\ .316 & .949 & 3.162 \end{bmatrix} \begin{bmatrix} 3.429 \\ 1.286 \\ -1.429 \end{bmatrix} = \begin{bmatrix} 3.748 \\ 5.484 \\ -2.214 \end{bmatrix}$$

elde edilir. v_i ve u_i ler sırasıyla U ve V nin sütun vektörleri olmak üzere

$$\hat{a} = -3.666\, v_1 - .333\, v_2 - 11.666\, u_3$$

olduğu görülür. Böylece $\hat{a} \in C_s = C_{\{1,2\}}$ olup $R_{\bar{s}} = R_{\{3\}} = (0,0,1)'$ dir. Bu nedenle (3.2.8) bağıntısını uygulamak suretiyle

$$\tilde{\beta} = \begin{bmatrix} 3.666 \\ .333 \\ 0 \end{bmatrix}$$

elde edilir.

Şimdi de V pozitif definit ve simetrik olmak üzere

$$(Y, X\beta, \sigma^2 V \mid R\beta \geq r) \qquad (3.1.3)'$$

genel lineer modelini gözönüne alalım, burada Y,X,β,R ve r daha önce verildiği şekildedir. β nın eşitsizlik kısıtlamalı genelleştirilmiş en küçük kareler tahmin edicisi $\tilde{\beta}$

$$\min_{\beta}\{\, \| Y-X\beta \|_V^2 : R\beta \geq r \} \qquad (3.1.6)'$$

ile verilir. Burada $\| Y-X\beta \|_V^2 = (Y-X\beta)'V^{-1}(Y-X\beta)$ dir. Şimdi (3.1.3)' modelinde β nın eşitsizlik kısıtlamalı genelleştirilmiş en küçük kareler tahmin edicisi için kapalı form ifadeler geliştireceğiz.Bunun için önce aşağıdaki lemmayı verelim.

Lemma 3.2.1.[10]: $R\in R^{p,m}$ p ranklı matrisi ve $r\in R^{p}=R(R)$ vektörü için $N\in R^{m,m-p}$ $R(N) = N(R)$ olacak şekilde bir matris olsun. Bu takdirde β nın $R\beta\geq r$ için bir çözüm olması için gerek ve yeter şart bir $v\in R_{+}^{p}$ (yani $v\geq 0$) ve bir $w\in R^{m-p}$ için $\beta = R^{+}r+R^{+}v+Nw$ olmasıdır.

Ispat:"$\Longrightarrow$" β $R\beta\geq r$ için bir çözüm olsun. Bu takdirde bir $v\in R_{+}^{p}$ için $R\beta=r+v$ dir. Bu son denklem bir β için çözülebilir; $R^{+}(r+v)$ bir özel çözümdür. Bu nedenle $R\beta=r+v$ için genel çözüm $w\in R^{m-p}$ keyfi olmak üzere $\beta = R^{+}(r+v)+Nw$ olarak yazılabilir.

"$\Longleftarrow$" $RR^{+} = I$ ve $RN = 0$ olduğundan ispat açıktır.

Yukarıda verilen lemma R^{+} yerine R nin bir R^{-} g-inversi alındığında da gerçeklenir. Bununla beraber bu lemmanın sonucunun R nin tam satır ranklı olması gerçeğine bağlı olduğuna dikkat edilmelidir. Aksi takdirde yani R bir satır kusuruna sahipse bu sonuç genelde doğru değildir.

Not 3.2.1.[10]: Lemma 3.2.1. in şartları altında (3.1.6)' kısıtlamalı kuadratik programlama probleminin yerine

$$\min(\tilde{Y}-\tilde{X}c)'(\tilde{Y}-\tilde{X}c)\ , \qquad (3.2.14)$$

$$c=(c_1',c_2')'\in R_{+}^{p}xR^{m-p}$$

yardımcı probleminin çözülebileceği kolayca görülebilir. Burada

$$\tilde{Y}:=V^{-1/2}(Y-XR^{+}r)\ ,\ \tilde{X}:=V^{-1/2}X(R^{+},N) \qquad (3.2.15a-b)$$

ve N de Lemma3.2.1. deki gibi seçilmiştir. Eğer $\hat{a}=(\hat{a}_1',\hat{a}_2')'$ (3.2.14) için bir çözüm ise bu takdirde $\tilde{\beta}:=R^{+}r+R^{+}\hat{a}_1+N\hat{a}_2$ (3.1.6)' orijinal problemi için bir çözümdür.

Detaya inmeden önce diğer bazı notasyonları vermek uygundur. $V^{-1/2}$ V^{-1} in pozitif karekökünü göstermek üzere $\bar{X}:=V^{-1/2}X$ olsun. $\triangle$

$P:=\{1,2,..,p\}$ nin bir altkümesi olsun. Δ nın P ye göre bütünleyeni $\not\Delta$ ile gösterelim. Herhangi bir Δ için

$$S_\Delta := [\ldots, \underset{\substack{\uparrow \\ i\in\not\Delta \text{ ise}}}{\bar{X}'\bar{X}R_i^+}, \ldots, \underset{\substack{\uparrow \\ j\in\Delta \text{ ise}}}{-R^j}, \ldots; X'XN] \qquad (3.2.16)$$

blok parçalanmış matrisini tanımlayalım, burada R_i^+ ve R^j sırasıyla R^+ nin i-yinci sütun vektörünü ve R' nün j-yinci sütun vektörünü gösterir. Diğer bir deyişle S_Δ nin ilk p sütun vektörü aşağıdaki gibi verilir; eğer $i\in\not\Delta = P\backslash\Delta$ ise S_Δ nin i-yinci sütunu, $S_{\Delta i}$ diyelim,

$$S_{\Delta i} = \bar{X}'\bar{X}R_i^+$$

ile verilir. Aksi takdirde yani $i\in\Delta$ ise

$$S_{\Delta i} = -R^i$$

dir. Herhangi bir $A\in R^{m,m}$ matrisi ve yukarıdaki gibi herhangi bir Δ için A nın i-yinci sütun vektörünü (satır vektörünü) içermesi için gerek ve yeter şart $i\in\Delta$ olan A nın bir alt matrisi için $A_{c\Delta}(A_{r\Delta})$ yazmak uygundur. Eğer $\Delta = \emptyset$ ise bu takdirde $A_{c\Delta}$ ve $A_{r\Delta}$ yı "mevcut değildir" biçiminde yorumlarız.

YÖNTEM 3.2.1.[10]: (3.2.14) ün çözümü için bir metod aşağıdaki adımlar üzerine kurulmuştur:

1.Adım: P nin bir Δ altkümesi seçilir.

2.Adım: $(S_\Delta)^{-1}\bar{X}'\tilde{Y}$ hesaplanır.

3.Adım: Eğer $[(S_\Delta)^{-1}\bar{X}'\tilde{Y}]_{rP} \geq 0$ ise 4.Adım' a geçilir, aksi takdirde Δ değiştirilip 2.Adım'a dönülür.

4.Adım: (3.2.14) için $\hat{a}=(\hat{a}_1',\hat{a}_2')'\in R_+^p x R^{m-p}$ optimal çözümünü bulmak için

$$(\hat{a}_1)_i = \begin{cases} [(S_\Delta)^{-1}\bar{X}'\tilde{Y}]_i & i\in\not\Delta \\ \\ 0 & i\in\Delta \end{cases} \qquad (3.2.17a)$$

ve $j=1,2,\ldots,m-p$ için

$$(\hat{a}_2)_j = [(S_\Delta)^{-1}\bar{X}'\tilde{Y}]_{p+j} \qquad (3.2.17b)$$

kullanılır. Sonlu sayıda farklı adımdan sonra bu metod (3.2.14) için bir optimal çözüm ile son bulacaktır.

Basitlik için yukarıdaki yöntemin bir ispatı sadece

$$(Y, X\beta, \sigma^2 I \mid \beta \geq 0) \qquad (3.2.18)$$

özel modeli için daha sonra verilecektir.

Teorem 3.2.2.[10]: $\hat{a} = (\hat{a}_1', \hat{a}_2')'$ (3.2.14) için bir çözüm olsun. Bu takdirde (3.1.6)' modelinin eşitsizlik kısıtlamalı genelleştirilmiş en küçük kareler tahmin edicisi aşağıdaki ifadelerden herhangi birisiyle gösterilebilir:

i) $\tilde{\beta} = R^+ r + R^+ \hat{a}_1 + N\hat{a}_2$ dir, burada N Lemma 3.2.1 de verilmiştir.

ii) $\tilde{\beta} = R^+ r + P_\Delta \bar{X}'\tilde{Y}$, $P_\Delta := P_{N(R) \oplus R((R^+)_{c\not\Delta}), R((\bar{X}'\bar{X})^{-1}(R')_{c\Delta})}$

iii) $\tilde{\beta} = R^+ r + P_\Delta(\hat{\beta} - R^+ r)$, burada $\hat{\beta} := (X'V^{-1}X)^{-1}X'V^{-1}Y$ $(Y, X\beta, \sigma^2 V)$ kısıtlamasız modelinin alışılmış en küçük kareler tahmin edicisidir.

iv) $\tilde{\beta} = \hat{\beta} + (X'V^{-1}X)^{-1}(R_{r\Delta})'[R_{r\Delta}(X'V^{-1}X)^{-1}(R_{r\Delta})']^{-1}(I_{r\Delta}r - R_{r\Delta}\hat{\beta})$

dır.

Ispat: $\hat{a} = (\hat{a}_1', \hat{a}_2')'$ (3.2.14) için bir çözüm olsun. Not 3.2.1 e göre (i) veya buna denk olarak $\tilde{\beta} = R^+ r + (R^+, N)\hat{a}$ elde edilir. Şimdi $\bar{X}' = \bar{X}'\bar{X}\bar{X}^+$ olduğu gözönüne alınırsa $(S_\Delta)^{-1}\bar{X}' = (S_\Delta)^{-1}\bar{X}'\bar{X}\bar{X}^+ = [(\bar{X}'\bar{X})^{-1}S_\Delta]^{-1}\bar{X}^+$ elde edilir, dolayısıyla (3.2.16) ya göre

$$(S_\Delta)^{-1}\bar{X}' = [\ldots, \underset{\substack{\uparrow \\ i \notin \Delta \text{ ise}}}{R_i^+}, \ldots, \underset{\substack{\uparrow \\ j \in \Delta \text{ ise}}}{-(\bar{X}'\bar{X})^{-1}R^j}, \ldots; N]^{-1}\bar{X}^+ = (A_\Delta)^{-1}\bar{X}^+$$

dir. Burada A_Δ açık olarak tanımlanmıştır. Böylece

$$\begin{aligned}
\tilde{\beta} &= R^+ r + (R^+, N)\hat{a} \\
&= R^+ r + [\ldots, \underset{\substack{\uparrow \\ i \notin \Delta}}{R_i^+}, \ldots, \underset{j \in \Delta}{0}, \ldots; N](A_\Delta)^{-1}\bar{X}^+\tilde{Y} \\
&= R^+ r + [(R^+)_{c\not\Delta}, N][(R^+)_{c\Delta}, N]^{\{2, R((\bar{X}'\bar{X})^{-1}(R')_{c\Delta})\}}\bar{X}^+\tilde{Y} \\
&= R^+ r + P_{R(N) + R((R^+)_{c\not\Delta}), R((\bar{X}'\bar{X})^{-1}(R')_{c\Delta})}\bar{X}^+\tilde{Y} \\
&= R^+ r + P_\Delta \bar{X}^+\tilde{Y}
\end{aligned}$$

elde edilir yani ii) sağlanır. iii) bağıntısı ii) , (3.2.15a-b) bağıntıları ve $\bar{X} := V^{-1/2}X$ için $\bar{X}^{+} = (\bar{X}'\bar{X})^{-1}\bar{X}' = (X'V^{-1}X)^{-1}X'V^{-1/2}$ gerçeğinden elde edilir. iv) bağıntısı ise $I-P_{\Delta}$ nın X , R ve V cinsinden aşağıdaki gibi elde edilebildiğini gözönüne almak suretiyle iii) bağıntısından elde edilir:

$$I-P_{\Delta} = (X'V^{-1}X)^{-1}(R_{r\Delta})'[R_{r\Delta}(X'V^{-1}X)^{-1}(R_{r\Delta})']^{-1}R_{r\Delta}$$

dır. Böylece ispat tamamlanır.

Şimdi $X \in R^{n,m}$ tam sütun ranklı olmak üzere (3.2.18) özel modelini gözönüne alalım. Bu takdirde β nın eşitsizlik kısıtlamalı en küçük kareler tahmin edicisi $\tilde{\beta}$

$$\min_{\beta}\{\ \| Y-X\beta \|^2 \ |\ \beta \geq 0\} \qquad (3.1.5)'$$

optimizasyon problemi için çözümdür, burada $\|Y-X\beta\|^2 = (Y-X\beta)'(Y-X\beta)$ dir. $S=(S_1,S_2,\ldots,S_m):=X'X$ olsun. Herhangi bir $\Delta \subseteq M = \{1,2,\ldots,m\}$ için $\not\Delta$ Δ nın M ye göre bütünleyeni olsun. Ayrıca $S_{\Delta}:=(S_{\Delta 1},\ldots,S_{\Delta m})$

$$S_{\Delta i} = \begin{cases} S_i & i \in \not\Delta \\ -e_i & i \in \Delta \end{cases} \qquad (3.2.19)$$

ya uygun olarak tanımlansın, burada $e_i = (0,0,\ldots,0,1,0,\ldots,0)'$ R^m nin i-yinci standard birim vektörünü göstermektedir.

(3.2.18) özel modeli için Yöntem 3.2.1 aşağıdaki formu alır:

YÖNTEM 3.2.2.[10]: (3.1.5)' yü çözmek için bir metod aşağıdaki adımlar üzerine kurulmuştur:

1.Adım: $M = \{1,2,\ldots,m\}$ nin bir Δ altkümesi seçilir.

2.Adım: $(S_{\Delta})^{-1}X'Y$ heaplanır.

3.Adım: Eğer $(S_{\Delta})^{-1}X'Y \geq 0$ ise bu takdirde $\tilde{\beta}$

$$\tilde{\beta} = \begin{cases} [(S_\Delta)^{-1}X'Y]_i & i \in \not\Delta \\ 0 & i \in \Delta \end{cases}$$

ile verilir. Aksi takdirde Δ değiştirilip 2.Adıma dönülür. Sonlu bir sayıda farklı adımdan sonra bu yöntem (3.1.5)' için bir optimal çözüm ile son bulacaktır.

Ispat: X tam sütun ranklı olduğundan $X^+X = I$ dır. Bu nedenle $X\beta \in XR_+^m$ olması için gerek ve yeter şart $\beta = X^+X\beta \in R_+^m$ olmasıdır. $W=(W_1, W_2,\ldots,W_m):=-X^{+'}$ olarak tanımlayalım. $\Delta \subset M$ için $X_\Delta = (X_{\Delta 1}, X_{\Delta 2},\ldots, X_{\Delta m})$ yi

$$X_{\Delta i} = \begin{cases} X_i & i \in \not\Delta \\ W_i & i \in \Delta \end{cases} \tag{3.2.20}$$

ile tanımlayalım. $X^+X = I$ olduğunda yukarıdaki inşaa gözönüne alınırsa $R(W_{c\Delta}) \subseteq [R(X_{c\not\Delta})]^{\perp}$ elde edilir. $R(X^{+'}) = R(X)$ bağıntısından

$$R(X_\Delta) = R(X_{c\not\Delta}) \overset{\perp}{\oplus} R(W_{c\Delta}) = R(X)$$

elde edilir. $\Delta \subset M$ için K_Δ ile X_Δ nın sütunlarıyla üretilen çok yüzlü koniyi gösterelim, yani $K_\Delta := X_\Delta R_+^m$ dir. Bu takdirde X in ranjı K_Δ konilerinin birleşimidir yani

$$R(X) = \bigcup_{\Delta \subseteq M} K_\Delta \tag{3.2.21}$$

dır. Sadece iki regressör durumunda tipik bir ifade Şekil 3.2.1.de gösterilmiştir, burada X_1 ve X_2 X in sütun vektörleri W_1 ve W_2 ise W nin sütun vektörleridir. Böylece eğer $x \in R(X)$ ise bir $\Delta \subseteq M$ için $x \in K_\Delta$ dır. Üstelik $x \in R(X)$ için $x \in K_\Delta$ olması için gerek ve yeter şart $X_\Delta^+ x \geq 0$ olmasıdır.(çünkü $X_\Delta^+ X_\Delta = I$ dır.) $W = -X^{+'} = -X(X'X)^{-1}$ i gözönüne alalım. $X'X^{+'} = X^+X$ ve $X^+X = I$ olduğundan (3.2.19) ve (3.2.20) bağıntılarından $X'X_\Delta = S_\Delta$ elde edilir. Bu nedenle $X'X_\Delta X_\Delta^+ = S_\Delta X_\Delta^+$ dır. Fakat $X_\Delta X_\Delta^+ = P_{R(X_\Delta)} = P_{R(X)} = XX^+$ ve $X'XX^+ = X'$olduğundan $X'X_\Delta X_\Delta^+ = X'P_{R(X_\Delta)} = X'$dır. Sonuç olarak $X' = S_\Delta S_\Delta^+$ veya buna denk olarak $X_\Delta^+ = (S_\Delta)^{-1}X'$dir.

$X_\Delta^+ = X_\Delta^+ X_\Delta X_\Delta^+ = X_\Delta^+ P_{R(X_\Delta)} = X_\Delta^+ P_{R(X)}$ olduğundan her $Y \in R^n$ için $X_\Delta^+ Y = X_\Delta^+ \hat{Y}$ elde edilir, burada $\hat{Y} = P_{R(X)} Y$ $X\beta$ nın $(Y, X\beta, \sigma^2 I)$ kısıtlamasız modeli altındaki en küçük kareler tahminidir. Bu sonuçları birleştirerek

$$(S_\Delta)^{-1} X'Y \geq 0 \Longleftrightarrow X_\Delta^+ Y \geq 0 \Longleftrightarrow X_\Delta^+ \hat{Y} \geq 0 \Longleftrightarrow \hat{Y} \in K_\Delta$$

elde edilir. $\beta \in R_+^m$ için

$$\| Y - X\beta \|^2 = \| Y - P_{R(X)} Y + P_{R(X)}(Y - X\beta) \|^2 = \| P_{N(X')} Y \|^2 + \| \hat{Y} - X\beta \|^2$$

olduğundan β nın (3.1.5)' problemini minimumlaştırması için gerek ve yeter şart

$$\min\{ \| \hat{Y} - X\beta \|^2 \mid \beta \geq 0 \text{ yani } X\beta \in XR_+^m \}$$

problemini minimumlaştırmasıdır.

örnek 3.2.2.[10]: örnek 3.2.1. i gözönüne alalım.

$$X = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 3 \\ 1 & 2 & 0 \end{bmatrix}, \quad Y = \begin{bmatrix} 2 \\ -3 \\ 6 \end{bmatrix}$$

olmak üzere

$$Y = X\beta + e \,, \quad \beta \geq 0$$

dır. Buradan

$$(X'X, X'Y) = \begin{bmatrix} 2 & 2 & 1 & 8 \\ 2 & 5 & 3 & 9 \\ 1 & 3 & 10 & -7 \end{bmatrix}$$

idi. $= \{3\}$ alınarak

$$S_{\{3\}} = \begin{bmatrix} 2 & 2 & 0 \\ 2 & 5 & 0 \\ 1 & 3 & -1 \end{bmatrix}, (S_{\{3\}})^{-1} = \begin{bmatrix} 5/6 & -1/3 & 0 \\ -1/3 & 1/3 & 0 \\ -1/6 & 2/3 & -1 \end{bmatrix}$$

olup

$$(S_{\{3\}})^{-1} X'Y = 1/6 \begin{bmatrix} 22 \\ 2 \\ 70 \end{bmatrix} = \begin{bmatrix} 3.666 \\ 0.333 \\ 11.666 \end{bmatrix} \geq 0$$

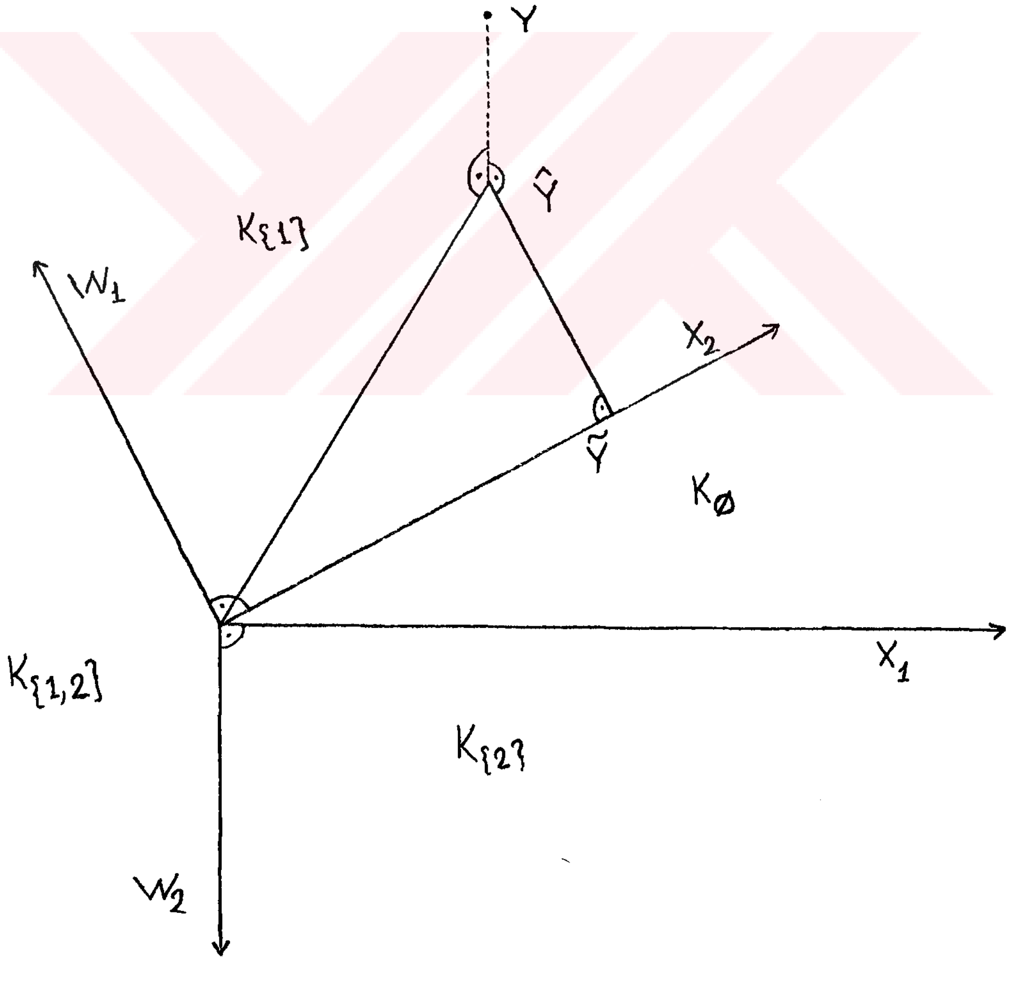
elde edilir. Böylece

$$\tilde{\beta} = \begin{bmatrix} 3.666 \\ 0.333 \\ 0 \end{bmatrix} \quad \text{ve} \quad \tilde{Y} = \begin{bmatrix} 3.666 \\ 0.333 \\ 4.333 \end{bmatrix}$$

bulunur. O halde $\tilde{\beta}$ daha önce örnek 3.2.1. deki ile aynıdır. Alışılmış en küçük kareler tahminleri için $\hat{\beta} = (S_{\emptyset})^{-1}X'Y = (X'X)^{-1}X'Y$ ve $\hat{Y} = X\hat{\beta}$ dan gerekli hesaplamalar yapılırsa

$$\hat{\beta} = \begin{bmatrix} 3.429 \\ 1.286 \\ -1.429 \end{bmatrix}, \quad \hat{Y} = \begin{bmatrix} 2 \\ -3 \\ 6 \end{bmatrix}$$

elde edilir.



Şekil 3.2.1.

İki-Adım Tahmin Edici

Şimdi özel lineer modeller durumunda Yöntem 3.2.1. in iki basit uyarlamasını geliştireceğiz. Önce $x_1 \in R^n$, $X_2 \in R^{n,m}$ ve $X:=(x_1,X_2)$ tam sütun ranklı olmak üzere

$$Y = x_1\beta_1 + X_2\beta_2 + e \ , \ \beta_1 \geq 0$$

regresyon modelini gözönüne alalım. $\beta:=(\beta_1,\beta_2')'$ diyelim. Bu model için β nın eşitsizlik kısıtlamalı en küçük kareler tahmini aşağıdaki yöntemi uygulayarak elde edilebilir.

Yöntem 3.2.3.[10]:(İki-adım Yöntemi)

1.Adım: β nın alışılmış en küçük kareler tahmini $\hat{\beta} = (X'X)^{-1}X'Y$ hesaplanır. Eğer $\hat{\beta}$ nın birinci bileşeni negatif değilse yani $(\hat{\beta})_1 \geq 0$ ise $\tilde{\beta} = \hat{\beta}$ dir. Aksi takdirde yani $(\hat{\beta})_1 < 0$ ise 2.adıma geçilir.

2.Adım: $(X_2'X_2)^{-1}X_2'Y$ hesaplanır, bu $Y = X_2\beta_2^* + e^*$ dan β_2^* ın en küçük kareler tahminidir. Bu takdirde

$$\tilde{\beta} = \begin{bmatrix} 0 \\ \\ (X_2'X_2)^{-1}X_2'Y \end{bmatrix}$$

dir. Bu metodla elde edilen tahmin edici sıksık iki-adım tahmin edici olarak adlandırılır. Yukarıdaki iki adım yöntemi Yöntem 3.2.1. de verilen daha genel yöntemden

$$(S_{\{1\}})^{-1}X' = \begin{bmatrix} -x_1'(I-X_2(X_2'X_2)^{-1}X_2') \\ \\ (X_2'X_2)^{-1}X_2' \end{bmatrix}$$

olduğu gözönüne alınarak kolayca elde edilebilir.

Şimdi de $X_1 \in R^{n,m_1}$, $m_1 \geq 2$ olmak üzere

$$Y = X_1\beta_1 + X_2\beta_2 + e \ , \ \beta_1 \geq 0$$

regresyon modelini gözönüne alalım. $X:=(X_1,X_2)$ olsun ve X in tam sütun ranklı olduğunu kabul edelim. $\beta = (\beta_1',\beta_2')'$olsun. Bu durumda aşağıdaki genişletilmiş iki adım yönteminin doğru olması için gerek ve yeter şart $X_1'X_1$ in diagonal ve $X_1'X_2 = 0$ olmasıdır. X in her $X=(Z_1, Z_2)$ blok parçalanışı için daima

$$(S_{\{1,\ldots,s\}})^{-1}X' = \begin{bmatrix} -Z_1'(I-Z_2(Z_2'Z_2)^{-1}Z_2') \\ \\ (Z_2'Z_2)^{-1}Z_2' \end{bmatrix}$$

elde edilir, burada s Z_1 in sütun sayısını gösterir. Ayrıca

$$(X'X)^{-1}X' = \begin{bmatrix} (Z_1'Z_1)^{-1}Z_1' \\ \\ (Z_2'Z_2)^{-1}Z_2' \end{bmatrix}$$

eşitliğinin doğru olmasi için gerek ve yeter şart $Z_1'Z_2 = 0$ olmasıdır.

Yöntem 3.2.4.[10]:(Genişletilmiş iki-adım yöntemi)

1.Adım: β nın alışılmış en küçük kareler tahmin edicisi $\hat{\beta}=(X'X)^{-1}X'Y$ hesaplanır. Eğer $(\hat{\beta})_{i=1,2,\ldots,m} \geq 0$ ise $\tilde{\beta}=\hat{\beta}$ dir. Aksi takdirde 2.adıma geçilir.

2.Adım: $\Delta := \{ i \leq m_1 \mid (\hat{\beta})_i < 0 \}$ olsun. Bu takdirde

$$\tilde{\beta} = \begin{cases} 0 & i \in \Delta \text{ ise} \\ \\ (\hat{\beta})_i & \text{aksi takdirde} \end{cases}$$

dır.

Şimdi parçalı genelleştirilmiş lineer regresyon modelinin bir özelliğini vereceğiz.

$$Y = X_1\beta_1+X_2\beta_2+e \qquad (*)$$

lineer modelini ele alalım, burada Y nx1 boyutlu bir rastgele vektör, $X=[X_1;X_2]$ nx(p+q) boyutlu tam sütun ranklı bilinen bir matris, E[Y]=

$X_1\beta_1 + X_2\beta_2$, $\mathrm{cov}[Y]=\Omega$, Ω pozitif definit nxn boyutlu bilinen bir matris ve β_1, β_2 de sırasıyla px1 ve qx1 boyutlu bilinmeyen parametre vektörleridir. (*) modelini $\beta = [\beta_1':\beta_2']'$ olmak üzere

$$\varphi = (Y, X\beta, \Omega)$$

ile gösterelim. I_n nxn boyutlu bir birim matris olmak üzere $M_1=I_n-P_1$ ve $P_1=X_1(X_1'X_1)^{-1}X_1'$ olarak tanımlayalım. Böylece (*) modelini soldan M_1 ile çarparak

$$M_1Y = M_1X_2\beta_2 + \ddot{e}$$

modelini elde ederiz, burada $E[\ddot{e}] = 0$ ve $\mathrm{cov}[\ddot{e}] = M_1\Omega M_1$ dir. Ayrıca

$$\phi = (M_1Y,\ M_1X_2\beta_2,\ M_1\Omega M_1)$$
$$\Psi = (M_1Y,\ M_1X_2\beta_2,\ \Omega)$$

olarak gösterilsin. φ modeli altında β nın alışılmış en küçük kareler tahmini ve en iyi liner yansız tahmini sırasıyla

$$\hat{\beta}(\varphi) = (X'X)^{-1}X'Y \quad \text{ve} \quad \beta^*(\varphi) = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}Y$$

dir. Bu iki tahminin eşit olması için bazı gerek ve yeter şartları daha önce vermiştik. Ayrıca $P\Omega=\Omega P$, $M\Omega=\Omega M$, ve $\Omega^{-1}M=M\Omega^{-1}M$ de bu tahminlerin eşit olması için denk şartlardır, burada $P=X(X'X)^{-1}X'$ ve $M=I_n-P$dir. Şimdi aşağıdaki lemmayı verelim.

Lemma 3.2.2: φ, ϕ ve Ψ yukarıdaki gibi tanımlanmak üzere

i) $\beta_2^*(\varphi) = \beta_2^*(\phi)$

ii) Eğer $(Y, X_1\beta_1, \Omega)$ altında $\beta_1 = \beta_1^*$ ise $\beta_2^*(\varphi) = \beta_2^*(\Psi)$ dir.

Ispat: i) $X'\Omega^{-1}X$, $X_1'\Omega^{-1}X_1$ ve $X_2'\Omega^{-1}X_2$ matrisleri nonsingüler olduğundan $X'\Omega^{-1}X$ in inversi $T = X_2'\Omega^{-1}X_2-X_2'\Omega^{-1}X_1(X_1'\Omega^{-1}X_1)^{-1}X_1'\Omega^{-1}X_2$ olmak üzere

$$\left[\begin{array}{c|c} (X_1\Omega^{-1}X_1)^{-1}[I+X_1\Omega^{-1}X_2T^{-1}X_2\Omega^{-1}X_1(X_1\Omega^{-1}X_1)^{-1}] & -(X_1\Omega^{-1}X_1)^{-1}X_1\Omega^{-1}X_2T^{-1} \\ \hline -T^{-1}X_2\Omega^{-1}X_1(X_1\Omega^{-1}X_1)^{-1} & T^{-1} \end{array}\right]$$

olduğundan $\beta^*(\varphi) = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}Y$ olup

$$\beta_2^*(\varphi) = -T^{-1}X_2'\Omega^{-1}X_1(X_1'\Omega^{-1}X_1)^{-1}X_1'\Omega^{-1}Y + T^{-1}X_2'\Omega^{-1}Y$$

dir. öte yandan $\dot{M}_1 = \Omega^{-1}-\Omega^{-1}X_1(X_1'\Omega^{-1}X_1)^{-1}X_1'\Omega^{-1}$ olarak tanımlanırsa $T = X_2'\dot{M}_1X_2$ olup

$$\begin{aligned}\beta_2^*(\varphi) &= (X_2'\dot{M}_1X_2)^{-1}X_2'[\Omega^{-1}-X_1(X_1'\Omega^{-1}X_1)^{-1}X_1'\Omega^{-1}]Y \\ &= (X_2'\dot{M}_1X_2)^{-1}X_2'\dot{M}_1Y\end{aligned}$$

elde edilir. Ω pozitif definit olduğunda $R[M_1\Omega M_1]=R[M_1\Omega]=R[M_1]$ olduğundan $R[M_1X_2]\subset R[M_1\Omega M_1]$ dir. Bu nedenle

$$\beta_2^*(\phi) = [X_2'M_1(M_1\Omega M_1)^-M_1X_2]^{-1}X_2'M_1(M_1\Omega M_1)^-M_1Y$$

dir. Ayrıca $\dot{M}_1 = \Omega^{-1}[I_n-X_1(X_1'\Omega^{-1}X_1)^{-1}X_1'\Omega^{-1}] = \Omega^{-1}[\Omega M_1(M_1\Omega M_1)^-M_1] = M_1(M_1\Omega M_1)^-M_1 = (M_1\Omega M_1)^+$ dir. M_1 simetrik idempotent ve Ω nonsingüler olduğundan $(M_1\Omega M_1)^+ = M_1\Omega^{-1}M_1$ olup

$$\beta_2^*(\phi) = (X_2'\dot{M}_1X_2)^{-1}X_2'\dot{M}_1Y = \beta_2^*(\varphi)$$

elde edilir.

ii) $\beta_2^*(\Psi) = (X_2'M_1\Omega^{-1}M_1X_2)^{-1}X_2'M_1\Omega^{-1}M_1Y$ olduğu kolayca görülebilir. öte yandan $(Y, X_1\beta_1, \Omega)$ altında $\hat{\beta}_1 = \beta_1^*$ ise bu takdirde $P_1 = X_1(X_1'\Omega^{-1}X_1)^{-1}X_1'\Omega^{-1}$ olduğundan $\dot{M}_1=\Omega^{-1}M_1$ dir. $\dot{M}_1$ simetrik olduğundan $M_1\Omega^{-1}=\Omega^{-1}M_1$ dir. Bu nedenle $\dot{M}_1=M_1\Omega^{-1}M_1$ olup $\beta_2^*(\Psi)=(X_2'\dot{M}_1X_2)^{-1}X_2'\dot{M}_1Y= \beta_2^*(\varphi)$ elde edilir. Böylece lemmanın ispatı tamamlanır.

Şimdi bu lemmanın eşitlik ve eşitsizlik kısıtlamalı durumlarda nasıl verileceğini görelim. $R=[R_1:R_2]$ β ya uygun parçalanmış $m\times(p+q)$ boyutlu ve tam satır ranklı bilinen bir matris olsun. $R\beta=r$ kısıtlaması altnda φ modelinde β nn eşitlik kısıtlamalıgenelleştirilmiş en küçük kareler tahmini

$$\beta_H^*(\varphi)=\beta^*(\varphi)-(X'\Omega^{-1}X)^{-1}R'[R(X'\Omega^{-1}X)^{-1}R']^{-1}(R\beta^*(\varphi)-r) \quad (**)$$

dır. Buradan

$$\begin{bmatrix}\beta_{1H}^* \\ \\ \beta_{2H}^*\end{bmatrix}(\varphi) = \begin{bmatrix}\beta_1^*(\varphi) \\ \\ \beta_2^*(\varphi)\end{bmatrix} - \begin{bmatrix}X_1'\Omega^{-1}X_1 & X_1'\Omega^{-1}X_2 \\ \\ X_2'\Omega^{-1}X_1 & X_2'\Omega^{-1}X_2\end{bmatrix}^{-1}\begin{bmatrix}R_1' \\ \\ R_2'\end{bmatrix}$$

$$.\left\{ [R_1 \vdots R_2] \begin{bmatrix} X_1'\Omega^{-1}X_1 & X_1'\Omega^{-1}X_2 \\ X_2'\Omega^{-1}X_1 & X_2'\Omega^{-1}X_2 \end{bmatrix}^{-1} \begin{bmatrix} R_1' \\ R_2' \end{bmatrix} \right\}^{-1} (R_1\beta_1^*(\varphi)+R_2\beta_2^*(\varphi)-r)$$

elde edilir. Burada $R_1=0$ olarak alınırsa $R_2\beta_2=r$ kısıtlaması altında modelinde β_2 nin eşitlik kısıtlamalı genelleştirilmiş en küçük kareler tahmini

$$\beta_{2H}^*(\varphi) = \beta_2^*(\varphi)-T^{-1}R_2'(R_2T^{-1}R_2')^{-1}(R_2\beta_2^*(\varphi)-r)$$

olarak elde edilir, burada 0 tüm elemanları sıfır olan uygun boyutlu bir sıfır matrisidir.

Lemma 3.2.3: φ, ϕ ve Υ daha önce tanımlanan şekilde olsun.

i) $R_2\beta_2=r$ kısıtlaması altında $\beta_{2H}^*(\varphi) = \beta_{2H}^*(\phi)$ dır.

ii) $(Y,\ X_1\beta_1,\ \Omega)$ altında $\hat{\beta}_1 = \beta_1^*$ ise $R_2\beta_2=r$ kısıtlaması altında $\beta_{2H}^*(\varphi) = \beta_{2H}^*(\Upsilon)$ dır.

ispat: i) (**) ın elde edilmesinde kullanılan yönteme göre

$$\beta_{2H}^*(\phi) = \beta_2^*(\phi)-(X_2'M_1(M_1\Omega M_1)^-M_1X_2)^{-1}R_2'$$
$$.[R_2(X_2'M_1(M_1\Omega M_1)^-M_1X_2)^{-1}R_2']^{-1}(R_2\beta_2^*(\phi)-r)$$

elde edilir, burada A^-, A matrisinin genelleştirilmiş inversidir.

$$\dot{M}_1 = \Omega^{-1}[I_n-X_1(X_1'\Omega^{-1}X_1)^{-1}X_1'\Omega^{-1}] = \Omega^{-1}[\Omega M_1(M_1\Omega M_1)^-M_1]$$
$$= M_1(M_1\Omega M_1)^-M_1 = (M_1\Omega M_1)^+$$

dir, burada A^+, A matrisinin Moore-Penrose inversini gösterir. Öte yandan M_1 simetrik ve idempotent bir matris ve Ω nonsingüler olduğundan $(M_1\Omega M_1) = M_1\Omega^{-1}M_1$ dir. Böylece

$$\beta_{2H}^*(\phi)=\beta_2^*(\phi)-(X_2'\dot{M}_1X_2)^{-1}R_2'[R_2(X_2'\dot{M}_1X_2)^{-1}R_2']^{-1}(R_2\beta_2^*(\phi)-r)$$

dir. Lemma 3.2.2. den $\beta_{2H}^*(\varphi) = \beta_{2H}^*(\phi)$ elde edilr.

ii) Direkt hesaplama yapılırsa

$$\beta_{2H}^*(\Upsilon) = \beta_2^*(\Upsilon)-(X_2'M_1\Omega^{-1}M_1X_2)^{-1}R_2'[R_2(X_2'M_1\Omega^{-1}M_1X_2)^{-1}R_2']^{-1}$$
$$.(R_2\beta_2^*(\Upsilon)-r)$$

dir, burada $\beta_2^*(\Upsilon) = (X_2'M_1\Omega^{-1}M_1X_2)^{-1}X_2'M_1\Omega^{-1}M_1Y$ dir. Eğer $(Y,X_1\beta_1,\Omega)$

altında $\beta_1=\beta_1^*$ ise $\dot{M}_1 = M_1\Omega^{-1}M_1$ olduğunu daha önce göstermiştik. Bunu yukarıda yerine koymak suretiyle

$$\beta_{2H}^*(\varphi)=(X_2'\dot{M}_1X_2)^{-1}X_2'\dot{M}_1Y-(X_2'\dot{M}_1X_2)^{-1}R_2'[R_2(X_2'\dot{M}_1X_2)^{-1}R_2']^{-1}$$
$$\cdot(R_2(X_2'\dot{M}_1X_2)^{-1}X_2'\dot{M}_1Y-r)$$

elde edilir. Böylece $\beta_{2H}^*(\varphi) = \beta_{2H}^*(\varphi)$ elde edilir ki bu ise lemmanın ispatını tamamlar.

Şimdi $R\beta\geq r$ kısıtlaması altında modelinde β nın $\tilde{\beta}(\varphi)$ eşitsizlik kısıtlamalı genelleştirilmiş en küçük kareler tahminini ele alalım. φ yi $D = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}$ ile soldan çarpmak suretiyle

$$\beta^*(\varphi) = \beta + \acute{e}$$

elde edilir, burada $E[\acute{e}] = 0$ ve $cov[\acute{e}] = D\Omega D'$ dir. Böylece

$$R\beta^*(\varphi) - r = R\beta - r + \tilde{e} \qquad (***)$$

olup burada $E[\tilde{e}] = 0$ ve $cov[\tilde{e}] = RD\Omega D'R'$ dir. $\mu^* = R\beta^*(\varphi)-r$ olarak tanımlanırsa $\tilde{\mu}$

$$\min(\mu^*-\mu)'[RD\Omega D'R']^{-1}(\mu^*-\mu)$$
$$\mu\geq 0$$

probleminin bir çözümü olmak üzere

$$\tilde{\beta}(\varphi) = R^+(\tilde{\mu} + r) + (I-R^+R)h$$

idi, burada h uygun boyutlu keyfi bir vektördür. Eğer h $\beta^*(\varphi)$ olarak seçilirse $\tilde{\beta}$ $R\beta\geq r$ kısıtlaması altında β nın φ modeline göre minimum varyanslı yansız tahmin edicisi olacaktır. $\tilde{\mu}$ daha önce verilen yöntemler uygulanarak bulunabilir. O halde $R_2\beta_2\geq r$ kısıtlaması altında φ modeline göre β_2 nin $\tilde{\beta}_2(\varphi)$ tahminini elde edebiliriz. $R_1=0$ olarak alınırsa (***) dan $E[\eta] = 0$ ve $cov[\eta] = R_2(X_2'\dot{M}_1X_2)^{-1}R_2'$ olmak üzere $R_2\beta_2^*(\varphi) - r=R_2\beta_2 - r + \eta$ elde edilir. Bunun sonucu olarakta μ_2

$$\min(\mu_2^*-\mu_2)'[R_2(X_2'\dot{M}_1X_2)^{-1}(\mu_2^*-\mu_2)$$
$$\mu_2\geq 0$$

probleminin bir çözümü olmak üzere

$$\tilde{\beta}_2(\varphi) = R_2^+(\tilde{\mu}_2+r)+(I-R_2^+R_2)\beta_2^*(\varphi)$$

elde edilir. Böylece aşağıdaki lemmayı verebiliriz.

Lemma 3.2.4: i) $\tilde{\beta}_2(\phi)$ $R_2\beta_2 \geq r$ kısıtlaması altında ϕ modelinde β_2 nin eşitsizlik kısıtlamalı en küçük kareler tahmini ise

$$\tilde{\beta}_2(\phi) = \tilde{\beta}_2(\varphi)$$

dir.

ii) $(Y, X_1\beta_1, \Omega)$ altında $\hat{\beta}_1 = \beta_1^*$ ise bu takdirde

$$\tilde{\beta}_2(\Psi) = \tilde{\beta}_2(\varphi)$$

dir.

Ispat:i) ϕ modelini

$$D_1 = [X_2'(M_1\Omega M_1)^+X_2]^{-1}X_2'(M_1\Omega M_1)^+ = (X_2'\dot{M}_1X_2)^{-1}X_2'\dot{M}_1$$

ile çarparsak $\beta_2^*(\phi) = \beta_2 + \dot{e}$ elde edilir, burada $E[\dot{e}] = 0$ ve $cov[\dot{e}] = D_1M_1\Omega M_1D_1' = (X_2'\dot{M}_1X_2)^{-1}$ dir. $R_2\beta_2^*(\phi) - r = \tau^*$ ve $R_2\beta_2 - r = \mu_2$ olarak tanımlanırsa $\tau^* = \mu_2 + \delta$ elde edilir, burada $E[\delta] = 0$ ve $cov[\delta] = R_2D_1M_1\Omega M_1D_1'R_2' = R_2(X_2'\dot{M}_1X_2)^{-1}R_2'$ dir. $\beta^*(\varphi) = \beta^*(\phi)$ olduğundan $\tau^* = \mu_2^*$ dır. Ayrıca $E[\eta] = E[\delta]$ ve $cov[\eta] = cov[\delta]$ dir. Bu nedenle $\mu_2^* = \mu_2 + \eta$ modeli ile $\tau^* = \mu_2 + \delta$ modeli denktir. O halde $\tilde{\mu}_2$ $\tau^* = \mu_2 + \delta$ modeli için de çözüm olur yani $\tilde{\beta}_2(\varphi) = \tilde{\beta}_2(\phi)$ dir.

ii) $\beta_2^*(\Psi) = \beta_2^*(\varphi)$ olduğunu Lemma3.2.2.de göstermiştik. Öte yandan $\dot{M}_1 = M_1\Omega^{-1}M_1$ olduğundan $\beta_2^*(\Psi) = (X_2'\dot{M}_1X_2)^{-1}X_2'\dot{M}_1Y$ dir. Ψ yi D_1 ile soldan çarparsak $\beta_2^*(\Psi) = \beta_2 + \delta_1$ elde edilir, burada $E[\delta_1] = 0$ ve $cov[\delta_1] = D_1\Omega D_1'$ dir. Böylece $R_2\beta_2^*(\Psi) - r = R_2\beta_2 - r + \eta_1$ olup $E[\eta_1] = 0$ ve $cov[\eta_1] = R_2D_1\Omega D_1'R_2'$dir. Ayrıca $cov[\eta_1] = R_2(X_2'\dot{M}_1X_2)^{-1}R_2'$dir. $\mu_3^* = R_2\beta_2^*(\Psi) - r$ ve $\mu_3 = R_2\beta_2 - r$ olarak tanımlanırsa $\mu_3^* = \mu_3 + \eta_1$ elde edilir. Bu durumda $\beta_2^*(\varphi) = \beta_2^*(\Psi)$ olduğundan $\mu_2^* = \mu_3^*$, $\mu_2 = \mu_3$, $E[\delta] = E[\eta_1]$ ve $cov[\delta] = cov[\eta_1]$ olup $\mu_2^* = \mu_2 + \delta$ ve $\mu_3^* = \mu_3 + \eta_1$ modelleri denktir. Böylece $\tilde{\mu}_2$, $\mu_3^* = \mu_3 + \eta_1$ modeli için de çözümdür. Dolayısıyla $\tilde{\beta}_2(\Psi) = \tilde{\beta}_2(\varphi)$ elde edilir. Bu ise lemmanın ispatını tamamlar.

EŞİTSİZLİK KISITLAMALI LİNEER MODELLERDE LİKELİHOOD ORAN, WALD ve KUHN-TUCKER TESTLERİ

Bu kısımda X nxk boyutlu k ranklı bir matris ve R mxk boyutlu m ranklı bir matris, β kx1 boyutlu bir vektör olmak üzere

$$Y = X\beta + e \quad , \; e \sim N(0,\Sigma) \qquad (3.3.1)$$

lineer modeli altında

$$R\beta \geq r$$

alternatif hipotezine karşı

$$R\beta = r$$

sıfır hipotezinin test edilmesi problemini ele alacağız, burada Σ bilinen non-singüler bir matris ve r de mx1 boyutlu bir vektördür. Önce eşitsizlik kısıtlamalı tahmin edici $\tilde{\beta}$ nın

$$\max -(Y-X\beta)'\Sigma^{-1}(Y-X\beta)$$

$$R\beta \geq r$$

kuadratik programlama probleminin çözümü olduğunu hatırlayalım. $R\beta \geq r$ kısıtlamasıyla ilgili Kuhn-Tucker (K.T.) çarpanlar vektörünü $\tilde{\lambda}$ ile gösterelim. Eşitlik kısıtlamalı tahmin ediciyi $\hat{\beta}_0$ ile ve $R\beta=r$ kısıtlamasıyla ilgili Lagrange çarpanlar (L.M.) vektörünü $\hat{\lambda}_0$ ile gösterelim. Kısıtlamasız tahmin ediciyi $\hat{\beta}$ ile ve bununla ilgili lagrange çarpanlar vektörünü de uygunluk için $\hat{\lambda} = 0$ ile gösterebiliriz. Bu şartlar altında β nın üç tahmin edicisininde

$$\frac{\partial}{\partial\beta}\left[-(Y-X\beta)'\Sigma^{-1}(Y-X\beta) + \lambda'(R\beta-r)\right] = 0$$

şartını sağladığını kolayca görebiliriz. Bu şarttan

$$\tilde{\beta} = \hat{\beta} + (X'\Sigma^{-1}X)^{-1}R'\tilde{\lambda}/2 \; ,$$

$$\hat{\beta}_0 = \hat{\beta} + (X'\Sigma^{-1}X)^{-1}R'\hat{\lambda}_0/2$$

olduğu sağlanır. Böylece

$$\tilde{\beta} - \hat{\beta}_0 = (X'\Sigma^{-1}X)^{-1}R'(\tilde{\lambda} - \hat{\lambda}_0)/2 \qquad (3.3.2)$$

bağıntısı elde edilir. $R\beta=r$ sıfır hipotezini test etmek için aşağıdaki test istatistiklerinden birisini kullanabiliriz.

Likelihood oran test istatistiği (LR) genel olarak

$$\Phi_{LR} = -2\log\wedge = 2(\tilde{L} - \hat{L}_0)$$

ile tanımlanır, burada $\tilde{L}$ ve $\hat{L}_0$ sırasıyla $H:\{R\beta \geq r\}$ ve $H_0:\{R\beta = r\}$ hipotezleri altında Likelihood'un logaritmasının maksimum değerleridir. Böylece

$$\Phi_{LR} = -(Y-X\tilde{\beta})'\Sigma^{-1}(Y-X\tilde{\beta}) + (Y-X\hat{\beta}_0)'\Sigma^{-1}(Y-X\hat{\beta}_0) \qquad (3.3.3)$$

olduğu kolayca görülür. O halde Φ_{LR} nin

$$\max_{R\beta \geq r} -(Y-X\beta)'\Sigma^{-1}(Y-X\beta) + (Y-X\hat{\beta}_0)'\Sigma^{-1}(Y-X\hat{\beta}_0) \qquad (P)$$

maksimumlaştırma probleminin objektif fonksiyonunun optimum değeri olduğunu belirtmek önemlidir. (P) β ya göre kuadratik olan bir primal problemdir. Eşitsizlik kısıtlamaları altında kuadratik optimizasyon hakkındaki genel sonuçlardan (P) nin dual probleminin

$$\min_{\lambda \leq 0} (\lambda - \hat{\lambda}_0)'R(X'\Sigma^{-1}X)^{-1}R'(\lambda - \hat{\lambda}_0)/4 \qquad (D)$$

olduğu görülebilir. Kuhn-Tucker test istatistiği Φ_{KT} yi (D) nin optimum değeri olarak, yani

$$\Phi_{KT} = (\tilde{\lambda} - \hat{\lambda}_0)'R(X'\Sigma^{-1}X)^{-1}R'(\tilde{\lambda} - \hat{\lambda}_0)/4 \qquad (3.3.4)$$

ile tanımlarız. (P) ve (D) nin optimumda hesaplanmış objektif fonksiyonları eşit olduğundan

$$\Phi_{LR} = \Phi_{KT} \qquad (3.3.5)$$

dir. Bu son eşitlik Φ_{KT} Kuhn-Tucker test istatistiğinin tanımlanması için yeterli olabilir. Son olarak Wald test istatistiği genel olarak

$$\Phi_W = (R\tilde{\beta}-r)'[R(X'\Sigma^{-1}X)^{-1}R']^{-1}(R\tilde{\beta}-r)$$

ile tanımlanır. $R\hat{\beta}_0 = r$ olduğundan (3.3.2) bağıntısına göre

$$R\tilde{\beta}-r = R(\tilde{\beta}-\hat{\beta}_0) = [R(X'\Sigma^{-1}X)^{-1}R'](\tilde{\lambda} - \hat{\lambda}_0)/2$$

olup dolayısıyla $\Phi_W = \Phi_{KT}$ elde edilir. Böylece sapmaların bilinen Σ kovaryans matrisli bir normal dağılıma sahip olması durumunda birbirine eşit olan üç test istatistiği tanımlanmıştır.

3.3.1. Σ Bilindiğinde Üç Test İstatistiğinin Dağılımı

Likelihood oran test istatistiğinin dağılımını tartışmadan önce yukarıda belirtilen regresyon modeli için likelihood oran test istatistiğinin v kx1 boyutlu bir $N(0,(X'\Sigma^{-1}X)^{-1})$ rastgele vektör ve $\hat{\beta}$ β nın kısıtlamasız maksimum likelihood tahmin edicisi yani G.E.K.K tahmin edicisi olmak üzere

$$\hat{\beta} = \beta + v \ , \ R\beta \geq r \tag{3.3.6}$$

olarak belirtilen model için likelihood oran test istatistiğine denk olduğunu belirtelim. Bu yeterli bilgileri kullanarak ispatlanabilir. Bununla beraber oldukça direkt bir yaklaşım kullanarak bunu yapmak daha da basittir,

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

yazarak ve normal denklemleri kullanarak

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

elde edilir. Bu sonuç Φ_{LR} nin

$$\max-(\beta-\hat{\beta})'(X'\Sigma^{-1}X)(\beta-\hat{\beta}) + (\hat{\beta}_0-\hat{\beta})'(X'\Sigma^{-1}X)(\hat{\beta}_0-\hat{\beta})$$

$$R\beta \geq r$$

maksimumlaştırma probleminin objektif fonksiyonunun optimum değeri olduğunu göstermek için kullanılabilir ki bu açık olarak (3.3.6) ile verilen belirleyici ile ilgili maksimumlaştırma problemidir. Gouriéroux, Holly ve Monfort (1982) $R\beta=r$ sıfır hipotezi altında likelihood oran, Wald ve Kuhn-Tucker test istatistiklerinin ortak dağılım fonksiyonunun

$$\Pr(\Phi_A<a) = \sum_{i=0}^{m} w(m,i)[\Pr\chi^2(i)<a] \tag{3.3.7}$$

ki-kare dağılımlarının bir ağırlıklı toplamı olduğunu göstermişlerdir, burada $w(m,i)$ bilinen ağırlıkların bir kümesi olup $\chi^2(0)$ orijindeki birim toptur. Genel teoriyi açıklamak için bir ve iki kısıtlamalı özel durumları tartışacağız. Ancak bu özel durumlara geçmeden önce bu test yöntemlerinin güçleri hakkında bir yorum yapalım. Gerçekten, sezgimiz bir ön kısıtlamanın konulduğu bir metod ile türetilen bir testetme yönteminin ön kısıtlamanın gözönünde tutulmadığı test yöntemlerinden daha iyi güç karekteristiklerine sahip olduğunu ortaya koyar. Özellikle bu ileride açıklanacağı gibi bir tek kısıtlama durumlarında açıktır. Prensipte üç yöntemin güç fonksiyonu sıfır hipotezi altında bunların dağılımını bulmak için kullanılan metodlara benzer metodlarla bulunabilir. Bununla beraber en önemli zorluk ağırlıkların alternatif hipotezlere de bağlı olacağı gerçeğidir.

Bir Eşitsizlik Kısıtlaması

Bu durumda R 1xk boyutlu satır vektörü ve Φ_W de

$$\Phi_W = \frac{(R\tilde{\beta}-r)^2}{R(X'\Sigma^{-1}X)^{-1}R'} = \frac{(R\tilde{\beta}-r)^2}{var[R\hat{\beta}]}$$

olacak şekildedir. Ayrıca

$$R\tilde{\beta} = \begin{cases} R\hat{\beta} & R\hat{\beta} > r \\ r & \text{aksi takdirde} \end{cases}$$

olup $R\hat{\beta}-r/(var[R\hat{\beta}])^{\frac{1}{2}}$ yi t ile gösterirsek bu takdirde

$$\Phi_W = \begin{cases} t^2 & t>0 \\ 0 & \text{aksi takdirde} \end{cases}$$

elde edilir. $H_0:(R\beta=r)$ sıfır hipotezi altında t istatistiği bir standart normal dağılıma sahiptir ve Φ_W nin dağılımı ise $\frac{1}{2}\chi^2(0) + \frac{1}{2}\chi^2(1)$ dir. α anlam düzeyindeki kritik bölge c $P(\Phi_W>c)=\alpha$ ile tanımlı olmak üzere $\{\Phi_W>c\}$ formundadır; c $P(Z>c)=2\alpha$ yazarak kolayca elde edilebilir. Burada Z $\chi^2(1)$ dağılımına sahiptir. Ayrıca kritik bölge $\{t>\sqrt{c}\}$

olarak da yazılabilir. Eğer standart teori kullanılırsa kritik bölge $\{\Phi_W>c'\}$ olur, burada c' $P(Z>c') = \alpha$ olacak şekilde verilmiştir. c' c den daha büyük olduğundan bu yöntem sıfır hipotezinin çoğu kez kabul edilmesine yol açar.

İki Eşitsizlik Kısıtlaması

İstatistiklerin dağılımını basit olarak ortaya çıkarmak için bunları bireysel modele dönüştürmek faydalıdır. Önce bunu (3.3.6) da verilen formda yazalım. Daha sonra uygun dönüştürülebilir bir afin dönüşümü kullanarak bireysel belirleyiciyi w kx1 boyutlu bir $N(0,I_k)$ rastgele vektör olmak üzere

$$\hat{u} = u + w \quad u_1 \geq 0 \quad u_2 - \theta u_1 \geq 0$$

olarak yazarız. Bu durumda sıfır hipotezi $H_0 = \{u_1=0, u_2-\theta u_1=0\}$ olur ki bu ise $\{u_1=0, u_2=0\}$ a denktir. Buna ilaveten Φ_{LR} istatistiği değişmez ve

$$\Phi_{LR} = \| \hat{u} - \hat{u}_0 \|^2 - \| \hat{u} - \tilde{u} \|^2$$

olarak yazılabilir, burada $\hat{u}_0$ ve $\tilde{u}$ sırasıyla u nun H_0 hipotezi ve iddia edilen hipotez altındaki maksimum likelihood tahminleridir. $\hat{u}_0$ $\tilde{u}$ nın H_0 a karşılık gelen altuzay üzerindeki ortogonal izdüşümü olup

$$\Phi_{LR} = \| \tilde{u} - \hat{u}_0 \|^2 = \tilde{u}_1^2 + \tilde{u}_2^2$$

dir, burada $\tilde{u}_1$ ve $\tilde{u}_2$ u nın ilk iki koordinatıdır. $(\tilde{u}_1, \tilde{u}_2)$ $(\hat{u}_1, \hat{u}_2)$ nin Şekil 3.3.1 de görüldüğü gibi $u_1 \geq 0, u_2 - \theta u_1 \geq 0$ ile tanımlı koni üzerindeki dik izdüşümüdür.

$$\begin{bmatrix} \tilde{u}_1 \\ \tilde{u}_2 \end{bmatrix} = \begin{cases} [\hat{u}_1 \;\; \hat{u}_2]' & \hat{u}_1 \geq 0, \; \hat{u}_2 - \theta\hat{u}_1 \geq 0 \\ [0 \;\; \hat{u}_2]' & \hat{u}_1 < 0, \; \hat{u}_2 \geq 0 \\ [0 \;\; 0\,]' & \hat{u}_2 \leq 0, \; \theta\hat{u}_2 + \hat{u}_1 \leq 0 \\ (\theta\hat{u}_2 + \hat{u}_1 / 1 + \theta^2)[1 \;\; \theta]' & \theta\hat{u}_2 + \hat{u}_1 \geq 0, \; \hat{u}_2 - \theta\hat{u}_1 \leq 0 \end{cases}$$

Φ_{LR} istatistiği dört farklı forma sahiptir:

$$\Phi_{LR} = \begin{cases} \hat{u}_1^2+\hat{u}_2^2 & \hat{u}_1 \geq 0,\ \hat{u}_2-\theta\hat{u}_1 \geq 0 \\ \hat{u}_2^2 & \hat{u}_1 \leq 0,\ \hat{u}_2 \geq 0 \\ 0 & \hat{u}_2 \leq 0, \theta\hat{u}_2+\hat{u}_1 \leq 0 \\ \dfrac{(\theta\hat{u}_2+\hat{u}_1)^2}{1+\theta^2} & \theta\hat{u}_2+\hat{u}_1 \geq 0, \hat{u}_2-\theta\hat{u}_1 \leq 0 \end{cases}$$

Φ_{LR} nin $H_0(u_1=u_2=0)$ hipotezi altındaki dağılımı

$$P[\Phi_{LR}=0] = P[\hat{u}_2 \leq 0 : \theta\hat{u}_2+\hat{u}_1 \leq 0] = q \quad (0 \leq q \leq \tfrac{1}{2})$$

ve

$$P[0<\Phi_{LR}<x] = P_1+P_2+P_3$$

ile verilir, burada

$$P_1=P[0<\hat{u}_1^2+\hat{u}_2^2<x; \hat{u}_1 \geq 0; \hat{u}_2-\theta\hat{u}_1 \geq 0],$$

$$P_2=P[0<\hat{u}_2^2<x; \hat{u}_1 \leq 0; \hat{u}_2 \geq 0],$$

$$P_3=P[0<(\theta\hat{u}_2+\hat{u}_1)^2/(1+\theta^2)<x; \theta\hat{u}_2+\hat{u}_1 \geq 0; \hat{u}_2-\theta\hat{u}_1 \leq 0]$$

dır. Şimdi P_1, P_2 ve P_3 ün değerlerini elde edelim: $\hat{u}_1^2+\hat{u}_2^2$ ile

$$\begin{bmatrix} \hat{u}_1 \\ \hat{u}_2 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

vektörleri arasındaki açı H_0 altında bağımsız olduğundan

$$P_1=P[0<\hat{u}_1^2+\hat{u}_2^2<x]P[\hat{u}_1 \geq 0; \hat{u}_2-\theta\hat{u}_1 \geq 0]$$

dır ve dolayısıyla F_2 $\chi^2(2)$ dağılımının c.d.f. si olmak üzere

$$P_1=(\tfrac{1}{2}-q)F_2(x)$$

dir. Öte yandan $\hat{u}_1$ ve $\hat{u}_2$ H_0 altında bağımsız olduğundan

$$P_2=P[0<\hat{u}_2^2<x : \hat{u}_2 \geq 0]P[\hat{u}_1 \leq 0]$$

veya H_0 altında $\hat{u}_1$ ve $\hat{u}_2$ simetrik dağılımlara sahip olduğundan

$$P_2=\tfrac{1}{2}P[0<\hat{u}_2^2<x; \hat{u}_2 \geq 0] = 1/4P[0<\hat{u}_2^2<x]$$

elde edilir. Bu nedenle F_1 $\chi^2(1)$ dağılımının c.d.f.si olmak üzere

$$P_2=1/4F_1(x)$$

dir. $\theta\hat{u}_2+\hat{u}_1$ ile $\hat{u}_2-\theta\hat{u}_1$ bağımsız olduğundan P_3 nün hesaplanması da P_2 nin hesaplanması ile aynıdır, yani

$$P_3=1/4F_1(x)$$

dir. Sonuç olarak H_0 altında Φ_{LR} nin dağılımı

$$w(2,0)\chi^2(0)+w(2,1)\chi^2(1)+w(2,2)\chi^2(2)$$

dir, burada $w(2,0)=q$; $w(2,1)=\frac{1}{2}$; $w(2,2)=\frac{1}{2}-q$ olup $q=(2\pi)^{-1}\cos^{-1}[-\theta/(1+\theta^2)^{\frac{1}{2}}]$ dır. q nun orijinal veriler cinsinden ifadesi ise $R_j(j=1,2)$ R nin j-yinci satırı olmak üzere

$$(2\pi)^{-1}\cos^{-1}\{R_1(X'\Sigma^{-1}X)^{-1}R_2'/[R_1(X'\Sigma^{-1}X)^{-1}R_1']^{\frac{1}{2}}[R_2(X'\Sigma^{-1}X)^{-1}R_2']^{\frac{1}{2}}\}$$

dir. Ayrıca $\rho(R_1\hat\beta,R_2\hat\beta)$ $R_1\hat\beta$ ve $R_2\hat\beta$ arasındaki korelasyon katsayısı olmak üzere $q=(2\pi)^{-1}\cos^{-1}\rho(R_1\hat\beta,R_2\hat\beta)$ dır. α anlam düzeyindeki kritik bölge c $P[\Phi_{LR}>c]=\alpha$ ile verilmek üzere $\{\Phi_{LR}>c\}$ formundadır. Bu nedenle c, $1-\alpha\geq q$ olmak üzere $q+\frac{1}{2}F_1(c)+(\frac{1}{2}-q)F_2(c)=1-\alpha$ nin bir çözümüdür.

Şimdi c nin nasıl belirleneceğini görelim. Öncelikle

$$q+(1-q)F_2(c)<q+\tfrac{1}{2}F_1(c)+(\tfrac{1}{2}-q)F_2(c)<q+(1-q)F_1(c)$$

olduğunu belirtelim. Bu nedenle $c_j=F_j^{-1}(1-\alpha/(1-q))$ $j=1,2$, olmak üzere $c\in[c_1,c_2]$ dir. Daha sonra verilen q ve α değerleri için c nin değeri $[c_1,c_2]$ aralığında bir araştırma yöntemiyle elde edilir.

3.3.2. Σ Sonlu Sayıda Bilinmeyen Parametreye Bağlı Olduğunda Üç Test İstatistiği

Tipik olarak Σ matrisi bilinmez ve Σ nın elemanları β ile ilişkisiz olan sonlu sayıda parametrenin sürekli fonksiyonlarıdır. Bu kısımda σ^2bilinmeyen olmak üzere $\Sigma=\sigma^2V$ olduğunu kabul edelim. Ayrıca H_0 altında ve H altında σ^2, V ve β nın maksimum likelihood tahmin edicilerinin mevcut ve tek olduklarını kabul edelim ve bunlar sırasıyla $\hat\sigma_0^2,\hat V_0,\hat\beta_0$ ve $\tilde\sigma^2,\tilde V,\tilde\beta$ ile gösterilsin. Log-likelihood

$$L=-(T/2)\log2\pi-(T/2)\log\sigma^2-(1/2)\log\det V-(1/2\sigma^2)(Y-X\beta)'V^{-1}(Y-X\beta)$$

olduğundan

$$\tilde L=-(T/2)\log2\pi-(T/2)\log\tilde\sigma^2-(1/2)\log\det\tilde V-T/2$$

ve

$$\hat L_0=-(T/2)\log2\pi-(T/2)\log\hat\sigma_0^2-(1/2)\log\det\hat V_0-T/2$$

dir. $2(\tilde{L}-\hat{L}_0)$ ye eşit olan Φ_{LR} likelihood oran test istatistiği,

$$S(0,0)=\hat{\sigma}_0{}^2(\det\hat{V}_0)^{1/T} \quad , \quad S(1,1)=\tilde{\sigma}^2(\det\tilde{V})^{1/T}$$

olmak üzere

$$\Phi_{LR} = T\log(S(0,0)/S(1,1)) \tag{3.3.8}$$

olarak yazılabilir. Wald test istatistiği Φ_W eşitsizlik kısıtlamalı tahmin ediciler üzerine kurulur ve

$$\Phi_W = (R\tilde{\beta}-r)'[R(X'\tilde{V}^{-1}X)^{-1}R']^{-1}(R\tilde{\beta}-r)/\ \tilde{\sigma}^2 \tag{3.3.9}$$

olup Σ yerine $\tilde{\Sigma} = \tilde{\sigma}^2\tilde{V}$ alınmak suretiyle (P) primal probleminin objektif fonksiyonunun optimum değeridir. Sonuç olarak Φ_W

$$\Phi_W = -\ (Y-X\tilde{\beta})'\tilde{V}^{-1}(Y-X\tilde{\beta})/\ \tilde{\sigma}^2 + (Y-X\tilde{\hat{\beta}}_0)'\tilde{V}^{-1}(Y-X\tilde{\hat{\beta}}_0)/\ \tilde{\sigma}^2$$

$$= -T + (Y-X\tilde{\hat{\beta}}_0)'\tilde{V}^{-1}(Y-X\tilde{\hat{\beta}}_0)/\ \tilde{\sigma}^2$$

olarak yazılabilir, burada $\tilde{\hat{\beta}}_0$ $\tilde{\Sigma} = \tilde{\sigma}^2\tilde{V}$ kovaryans matrisini kullanarak $R\beta=r$ kısıtlaması altında β nın eşitlik kısıtlamalı genelleştirilmiş en küçük kareler tahmin edicisidir.

$$S(0,1) = (1/T)(Y-X\tilde{\hat{\beta}}_0)'\tilde{V}^{-1}(Y-X\tilde{\hat{\beta}}_0)(\det\tilde{V})^{1/T}$$

olarak tanımlanırsa bu takdirde

$$\Phi_W = -\ T + T(S(o,1)/S(1,1)) \tag{3.3.10}$$

formunda yazılır.Şimdi Kuhn-Tucker çarpan istatistiğini tanımlayalım. $\tilde{\lambda}_0$ Σ yerine $\hat{\Sigma}_0 = \hat{\sigma}_0{}^2\hat{V}_0$ almak üzere (P) primal problemiyle ilgili k.T çarpanı olsun. Φ_{KT}, K-T çarpan istatistiği

$$\Phi_{KT} = (\tilde{\lambda}_0 - \hat{\lambda}_0)'R(X'\hat{V}_0{}^{-1}X)^{-1}R'(\tilde{\lambda}_0 - \hat{\lambda}_0)\hat{\sigma}_0{}^2/4 \tag{3.3.11}$$

olarak tanımlanır. Görüleceği gibi Φ_{KT} Σ yerine $\hat{\Sigma}_0$ almak üzere (P) nin objektif fonksiyonunun optimum değeri olup

$$\Phi_{KT} = -(Y-X\tilde{\beta}_0)'\hat{V}_0{}^{-1}(Y-X\tilde{\beta}_0)/\ \hat{\sigma}_0{}^2 + T$$

dir, burada $\tilde{\beta}_0$, sapmaların kovaryans matrisi $\hat{\Sigma}_0 = \hat{\sigma}_0{}^2\hat{V}_0$ olmak üzere eşitsizlik kısıtlamalı genelleştirilmiş en küçük kareler tahmin edicisidir. Eğer

$$S(1,0) = (1/T)(Y-X\tilde{\beta}_0)'\hat{V}_0{}^{-1}(Y-X\tilde{\beta}_0)(\det\hat{V}_0)^{1/T}$$

olarak tanımlanırsa bu takdirde

$$\Phi_{KT} = T\,[1 - S(1,0)/S(0,0)] \qquad (3.3.12)$$

formunda yazılır. Şimdi sıfır hipotezi altında bu üç test istatistiğinin asimptotik dağılımını gözönüne alalım. $\hat{\sigma}_0{}^2$ ve $\tilde{\sigma}^2$ σ^2 nin tutarlı tahmin edicileri, $\hat{V}_0$ ile $\tilde{V}$ V nin tutarlı tahmin edicileri olsun. Bu kısımda tanımlanan üç test istatistiğinin $\Sigma=\sigma^2 V$ doğru kovaryansı ile hesaplanmış aynı asimptotik dağılıma sahip olduğu sureklilik bilgileriyle sağlanabilir. önceki kısımda geliştirilen sonuçlara uygun olarak H_0 altında üç test istatistiği asimptotik olarak

$$\sum_{i=1}^{m} w(p,i)\,\chi^2(i)$$

dağılımına sahiptir, burada w(p,i) olasılık ağırlıklarıdır. Bu üç test istatistiği arasında

$$\Phi_W \geq \Phi_{LR} \geq \Phi_{KT} \qquad (3.3.13)$$

eşitsizliği geçerlidir. İlk olarak Φ_{LR} ve Φ_W arasındaki bağıntıyı ele alalım.

$$\Phi_{LR}=T\,\log[1 +(S(0,0)-S(1,1))/S(1,1)]$$

olup $z\geq 0$ olmak üzere $\log(1+z)\leq z$ eşitsizliğini kullanarak

$$\Phi_{LR} \leq T\,[(S(0,0)-S(1,1)/S(1,1)]$$

eşitsizliğini elde ederiz. $S(0,1)\geq S(0,0)$ olduğundan

$$\Phi_{LR} \leq T\,[(S(0,1)-S(1,1))\,/S(1,1)] = \Phi_W$$

elde edilir. Benzer şekilde $z\geq 0$ için $\log(1+z)\geq z/(1+z)$, $S(1,0)\geq S(1,1)$ eşitsizliklerini kullanarak $\Phi_{LR}\geq\Phi_{KT}$ elde edilir ve böylece (3.3.13) bağıntısı sağlanır. Şimdi Kuhn-Tucker çarpan istatistiğinin hesaplanması için bir alternatif verelim. Bu, bir determinasyon katsayısı olarak lagrange çarpan test istatistiğinin yorumu ile aynıdır. Gerçekten Φ_{KT} nin $\hat{\Sigma}_0=\hat{\sigma}_0{}^2\hat{V}_0$ ile ilgili (P) primal problemindeki objektif fonksiyonun optimum değeri olduğu gözlenebilir:

$$\max -(Y-X\beta)'\hat{\Sigma}_0^{-1}(Y-X\beta) + (Y-X\hat{\beta}_0)'\hat{\Sigma}_0^{-1}(Y-X\hat{\beta}_0)$$

$$R\beta \geq r$$

$\hat{e}_0$, modelin parametreleri $R\beta=r$ sıfır hipotezi altında maksimum likelihood ile tahmin edilmiş olduğunda tahmin edilmiş rezidüler olsun. Bu takdirde $\hat{e}_0 = Y-X\hat{\beta}_0$ olup yukarıdaki problem $b=\beta-\hat{\beta}_0$ olmak üzere

$$\max -(\hat{e}_0-Xb)'\hat{\Sigma}_0^{-1}(\hat{e}_0-Xb) + \hat{e}_0'\hat{\Sigma}_0^{-1}\hat{e}_0$$

$$Rb \geq 0$$

formuna dönüşür. $\tilde{b}$ b nin optimal değeri olsun.

$$\Phi_{KT}=-(\hat{e}_0-X\tilde{b})'\hat{\Sigma}_0^{-1}(\hat{e}_0-X\tilde{b}) + \hat{e}_0'\hat{\Sigma}_0^{-1}\hat{e}_0 \qquad (3.3.14)$$

bağıntısı sağlanır. Bu son eşitliğe göre K.T. istatistiğinin hesaplanması için aşağıdaki iki-adım yöntemini benimseyebiliriz.

i) Modeli $R\beta=r$ kısıtlaması altında maksimum likelihood metodu ile tahmin edip $\hat{e}_0$ tahmin edilmiş rezidüleri hesaplanır.

ii) v nin kovaryans matrisi $\hat{V}_0$ olmak üzere

$$\hat{e}_0 = Xb + v \ , Rb \geq 0$$

modeli tahmin edilir. Determinasyon katsayısı R^2

$$R^2 = 1- (\hat{e}_0-X\tilde{b})'\hat{V}_0^{-1}(\hat{e}_0-X\tilde{b}) \ / \ \hat{e}_0'\hat{V}_0^{-1}\hat{e}_0$$

ile tanımlanır. (3.3.14) bağıntısına uygun olarak

$$\Phi_{KT}=-(\hat{e}_0-X\tilde{b})'\hat{V}_0^{-1}(\hat{e}_0-X\tilde{b}) / \hat{\sigma}_0^2 + \hat{e}_0'\hat{V}_0^{-1}\hat{e}_0 \ / \ \hat{\sigma}_0^2$$

olup $\hat{\sigma}_0^2= \hat{e}_0'\hat{V}_0^{-1}\hat{e}_0 \ / \ T$ olduğundan $\Phi_{KT} = TR^2$ elde edilir.

Burada tanımlanan üç test istatistiğinin, regresyon parametreleri korunmuş hipotez altında kısıtlamasız olduğu durumda karşılık gelen test istatistiklerinin hemen hemen tüm özelliklerine sahip olduğu görülebilir. Bununla beraber genel durum ve eşitsizlik kısıtlamalı durum arasında önemli bir fark vardır. Birincisinde test istatistiklerinin asimptotik dağılımı p serbestlik dereceli (burada p kısıtlamaların sayısıdır) bir ki-kare dağılımıdır, ancak ikincisinde test istatistiklerinin asimptotik dağılımı serbestlik dereceleri sıfırdan p ye dolaşmak üzere ki-kare dağılımlarının bir karışımıdır.

Şimdi de X nxk boyutlu k ranklı bir matris , R mxk boyutlu p ranklı bir matris,β kx1 boyutlu bir parametre vektörü e nx1 boyutlu hata vektörü ve Y nx1 boyutlu bir vektör olmak üzere

$$Y=X\beta+e \quad , \quad e \sim N(0,\sigma^2 I_n)$$

modelini gözönüne alalım. r mx1 boyutlu bir vektör olmak üzere $R\beta \geq r$ alternatif hipotezine karşı $R\beta=r$ sıfır hipotezini test etmek isteyelim. Bunun için

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

$$\hat{\sigma}_0{}^2 = (Y-X\hat{\beta}_0)'(Y-X\hat{\beta}_0)/n$$

ve

$$\tilde{v} = n(\hat{\sigma}_0{}^2-\tilde{\sigma}^2)$$

eşitliklerini kullanırsak

$$\tilde{v} = (Y-X\hat{\beta}_0)'(Y-\hat{X}_0)-(Y-X\tilde{\beta})'(Y-X\tilde{\beta})$$

$$= (\tilde{\beta}-\hat{\beta}_0)'X'X(\tilde{\beta}-\hat{\beta}_0)-2(\tilde{\beta}-\hat{\beta}_0)'X'X(\tilde{\beta}-\hat{\beta}_0)$$

elde edilir. $V = (X'X)^{-1}$ olmak üzere $\tilde{\beta}-\hat{\beta}_0 = VR'\tilde{\lambda}$ ve $R\hat{\beta}_0 = r$ bağıntıları sağlandığından

$$(\tilde{\beta}-\hat{\beta}_0)'X'X(\tilde{\beta}-\hat{\beta}_0) = (\tilde{\beta}-\hat{\beta}_0)'R'\tilde{\lambda} = \tilde{\lambda}'(R\tilde{\beta}-r) = \tilde{\lambda}'\tilde{\delta} = 0$$

dır. Bu nedenle

$$\tilde{v} = (\tilde{\beta}-\hat{\beta}_0)'X'X(\tilde{\beta}-\hat{\beta}_0)$$

dir. öte yandan $\tilde{\beta}-\hat{\beta}_0 = VR'(\tilde{\lambda} - \hat{\lambda}_0)$ olduğundan $R\tilde{\beta}-R\hat{\beta}_0 = RVR'(\tilde{\lambda}- \hat{\lambda}_0)$ dır. Böylece

$$\tilde{v} = (\tilde{\lambda}-\hat{\lambda}_0)'RVR'(\tilde{\lambda}-\hat{\lambda}_0)$$

veya

$$\tilde{v} = (R\tilde{\beta}-r)'(RVR')^{-}(R\tilde{\beta}-r)$$

yazılır, burada $(RVR')^{-}$ matrisi $RVR'(RVR')^{-}RVR' = RVR'$ şatını sağlayan bir matristir. Bu durumda Likelihood oran, Wald ve Kuhn-Tucker istatistikleri

$$\Phi_{LR} = n\ \log(\hat{\sigma}_0{}^2/\ \tilde{\sigma}^2) \ ,$$

$$\Phi_W = \tilde{v}/\tilde{\sigma}^2$$

ile σ^2 bilinmediğinde

$$\Phi_{KT} = \tilde{v}/\hat{\sigma}_0{}^2$$

ve σ^2 bilindiğinde

$$\Phi_A = \tilde{v}/\sigma^2$$

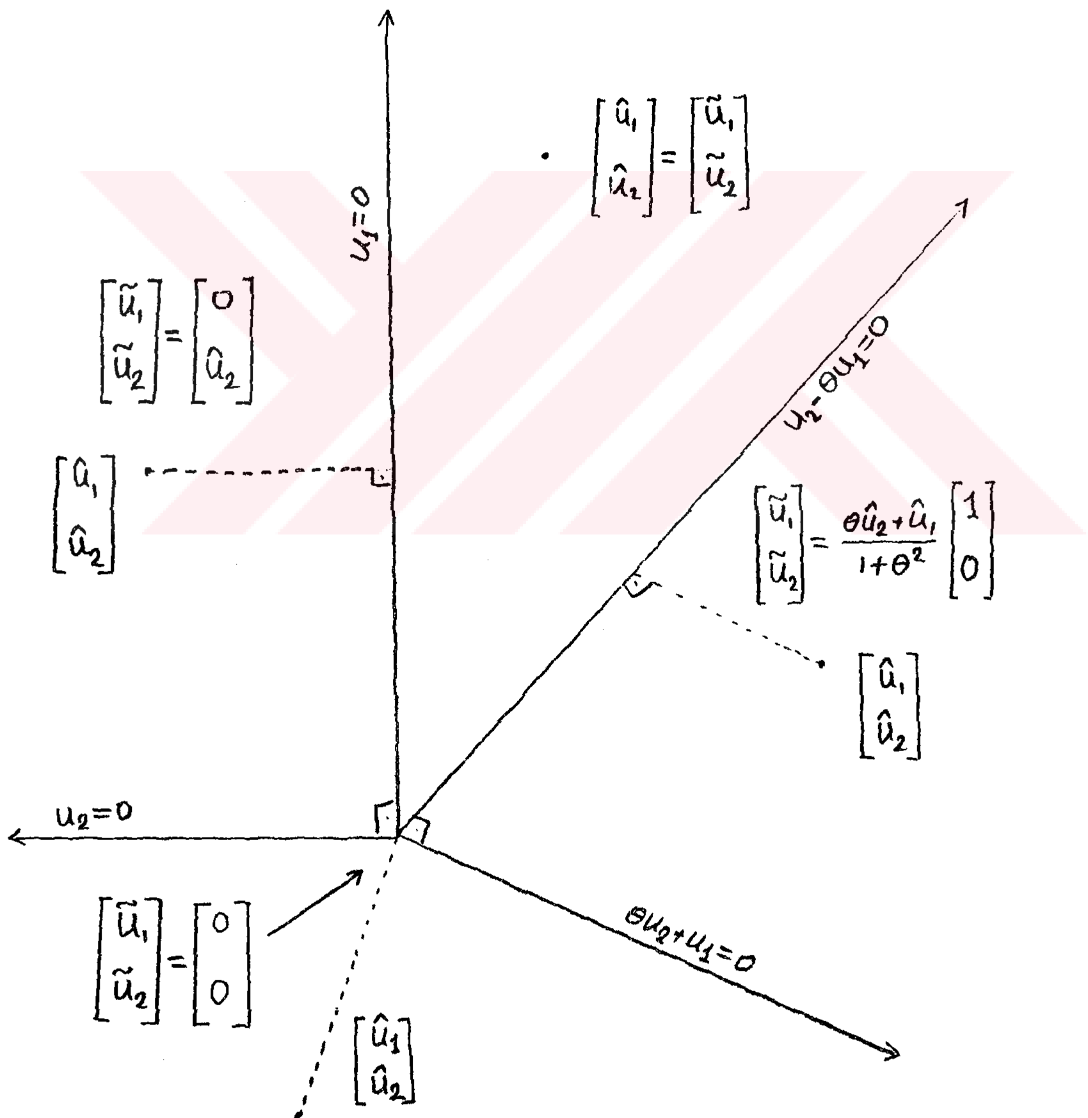
ile tanımlanır. Ayrıca

$$\Phi_{LR} = n \log(1 + \Phi_W/n)$$

ve

$$\Phi_{LR} = -n \log(1 - \Phi_{KT}/n)$$

bağıntıları gerçeklenir.



Şekil:3.3.1.

KAYNAKLAR

1. SEBER, G.S.F., Linear Regression Analysis, John Wiley and Sons Inc.,New York, (1977).
2. GRAYBIL, F.A., Theory and Application of the Linear Model, Wadsworth Publishing Company, Inc., Belmont, California, (1976).
3. KSHIRSAGAR, A.M., Multivariate Analysis, Marcel Dekker Inc., New York, (1978).
4. FAREBROTHER, R.W., Testing Linear İnequality Constraints in the Standart Linear Model, Comm. Stat, 15,(1986) 7-31.
5. GOURİEROUX,C.,HOLLY,A. and MONFORT,A., Likelihood ratio test, Wald test and Kuhn-Tucker test in Linear Models with Inequality Constraints on the Regresson Parameters, Econometrica 50, (1982) 63-80.
6. ESCOBAR,L.A. and SKARPNESS,B., A Closed Form Solution For the Least Squares Regression Problem with Linear Inequality Constraints, Comm. St. Theory and Meth, 13,(1984) 1127-1134.
7. FROOZİ, K.F., A Transformation of the İnequality Constrained Linear Model, Linear Algebra and its App. 133,(1990)153-163.
8. LIEW,C.K.,Inequality Constrainted Least-Squares Estimations. J, Amer. Stat.Assoc,71,(1976) 746-751.
9. JUDGE,G.G.and YANCEY,T., Sampling Properties of an Inequality Restricted Estimator,Economics Letters,7,(1981) 327-333.
10. WERNER,H.J., On Inequality Constrained Generalized Least-Squares Estimation,Lin. Algebra and its App,127,(1990) 379-392.
11. SEARLE,S.R., Linear Models, John Wiley and Sons Inc., New York, (1971).

12. BROWNLEE, K.A., Statistical Theory and Methodology In Science and Engineering, John Wiley and Sons Inc.,New York,(1965).

13. YAPAR,C., Lieer Modellerde Lineer Parametrik kısıtlamalar Altında Parametre Kestirimleri ve Hipotez Testleri,Doktora Tezi K.T.Ü. Trabzon, (1979).

14. THOMAS,M., A Note on Best Linear Unbiased Estimation in the Restricted General linear Model, Math. Oper. Stat. Ser. Stat. 10, (1984),3-6.

15. GRAYBILL,F.A., An Introduction to Linear Statistical Models Mc Graw-Hill Book Co.,Inc. New York,(1961).

16. JOHNSON,R.A.and WICHERN,D.W., Applied Multivariate Statistical Analysis, Prentice-Hall,Inc.Englewood Cliffs,(1982).

17. CHATTERJEE,S. and HADI,A.S., Sensivity Analysis in Linear Regresseon, John Wiley and Sons.Inc.,New York, (1988).

18. FLEMING,W.H., Functions of Several Variables, Addison-Wesley Publishing Co. Inc. New York, (1965).

19. LANG,S., Linear Algebra, Addison-Wesley Publishing Co. Inc. New York, (1970).

20. AKDENİZ,F., Olasılık ve İstatistik, A.Ü. Fen Fakültesi Yayınları,138, (1984)

21. PUNTANEN,S.and NURHONEN,M., A Property of Partitioned Generalized Regression, Comm. Stat. Theory and Methods, 21,(6) (1992) 1579-1583.

ÖZGEÇMİŞ

1965 yılında Samsun İlinin Vezirköprü İlçesine bağlı Gömlekhisar köyünde doğdu. İlk öğrenimini köyünde tamamladıktan sonra 1976-1977 öğretim yılında Vezirköprü İmam Hatip Lisesine kayıt yaptırdı. Buradan mezun olduktan sonra 1984-1985 öğretim yılında Karadeniz Teknik Üniversitesi Fen-Edebiyat Fakültesi Matematik Bölümüne girdi. 1989 yılında bu bölümden mezun oldu. Kasım 1989 tarihinde Karadeniz Teknik Üniversitesi Fen-Edebiyat Fakültesi Matematik Bölümünde Uygulamalı Matematik Anabilim Dalı'na Araştırma görevlisi olarak atandı ve halen bu görevine devam etmektedir.