

**ÇUKUROVA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

DOKTORA TEZİ

Mahmude Revan ÖZKALE

ÇOKLU İÇ İLİŞKİ İLE İLGİLİ PROBLEMLER

İSTATİSTİK ANABİLİM DALI

ADANA, 2007

ÇUKUROVA ÜNİVERSİTESİ

FEN BİLİMLERİ ENSTİTÜSÜ

ÇOKLU İÇ İLİŞKİ İLE İLGİLİ PROBLEMLER

Mahmude Revan ÖZKALE

DOKTORA TEZİ

İSTATİSTİK ANABİLİM DALI

Bu tez 24/09/2007 Tarihinde Aşağıdaki Jüri Üyeleri Tarafından Oybirliği/
Oyçokluğu İle Kabul Edilmiştir.

İmza.....
Prof.Dr. Selahattin KAÇIRANLAR
DANIŞMAN

İmza.....
Prof.Dr. Sadullah SAKALLIOĞLU
ÜYE

İmza.....
Doç. Dr. Rızvan EROL
ÜYE

İmza.....
Prof.Dr. Hamza EROL
ÜYE

İmza.....
Doç.Dr. Aşır GENÇ
ÜYE

Bu tez Enstitümüz İstatistik Anabilim Dalında hazırlanmıştır.

Kod No:

Prof. Dr. Aziz ERTUNÇ

Enstitü Müdürü

İmza ve Mühür

Bu çalışma

- Ç.Ü. Bilimsel Araştırma Projeleri Birimi Proje No: FEF2006D7
- Türkiye Bilimsel ve Teknoloji Araştırma Kurumu (TÜBİTAK)-Bilim İnsanı Destekleme Daire Başkanlığı (BİDEB)

tarafından desteklenmiştir.

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

ÖZ

DOKTORA TEZİ

ÇOKLU İÇ İLİŞKİ İLE İLGİLİ PROBLEMLER

Mahmude Revan ÖZKALE

ÇUKUROVA ÜNİVERSİTESİ

FEN BİLİMLERİ ENSTİTÜSÜ

İSTATİSTİK ANABİLİM DALI

Danışman : Pof. Dr. Selahattin KAÇIRANLAR

Yıl : 2007, Sayfa: 96

Jüri : Prof. Dr. Selahattin KAÇIRANLAR

Prof. Dr. Sadullah SAKALLIOĞLU

Doç. Dr. Rızvan EROL

Prof. Dr. Hamza EROL

Doç. Dr. Aşır GENÇ

Çoklu lineer regresyon modelinde açıklayıcı değişkenler arasındaki lineer ilişki olarak tanımlanan çoklu iç ilişkinin varlığı tahmin prosedüründe en küçük karelerin kötü bir tahmin edici olmasına neden olmaktadır. Bu nedenle ileri sürülen $r-d$ sınıf tahmin edici en küçük kareler, ridge ve temel bileşenler ve yansız ridge tahmin edici ridge, en küçük kareler ve $r-k$ sınıf tahmin edicilerle matris hata kareler ortalaması kriterine göre karşılaştırılmıştır. Çoklu iç ilişkinin kalite kontrol çalışmalarındaki etkilerinden dolayı Liu tahmin edici için PRESS istatistiği tanımlanmış ve bunu minimum yapan d nin tahmin edicisi verilmiştir.

Anahtar Kelimeler: Çoklu İç İlişki, Yanlı Tahmin Ediciler, Ön Tahmin Hata Kareler Toplamı, Kalite Kontrol

ABSTRACT

PhD. THESIS

**PROBLEMS RELATED WITH
MULTICOLLINEARITY**

Mahmude Revan ÖZKALE

DEPARTMENT OF STATISTICS
INSTITUTE OF NATURAL AND APPLIED SCIENCES
UNIVERSITY OF CUKUROVA

Supervisor : Prof. Dr. Selahattin KAÇIRANLAR
Year : 2007, Pages:96
Jury : Prof. Dr. Selahattin KAÇIRANLAR
Prof. Dr. Sadullah SAKALLIOĞLU
Assoc. Prof. Dr. Rızvan EROL
Prof. Dr. Hamza EROL
Assoc. Prof. Dr. Aşır GENÇ

Multicollinearity which is defined as the relationship between the explanatory variables in the multiple linear regression model causes the least squares estimator to be a poor estimator. The $r-d$ class estimator which is proposed for this reason is compared to the least squares, ridge and principal components regression and the unbiased ridge estimator is compared to the ridge, least squares and $r-k$ class estimators under the matrix mean square error criterion. Due to the effects of multicollinearity on quality control PRESS statistic for the Liu estimator is defined and the estimator of d which minimizes this statistic is given.

KeyWords: Multicollinearity, Biased Estimators, Prediction Error Sum of Squares, Quality Control

TEŞEKKÜR

Bu bilim yolculuğuna 2002 yılında başladım. Bu yoldaki zorluklarla mücadele etmek için azim ve çalışmak gerektiğini öğrendim. Her zaman için başaracağına inanmak aslında başarının ta kendisidir. Azim ve çalışmak kişinin kendi elinde olmasına rağmen elinde olmayan ve başarıyı tetikleyen bence en önemli faktörlerden birisi de belli bir potansiyele sahip olabilmektir. Belli bir potansiyele sahip olduğunu kişinin kendisi fark edemez ama çevresinde bunu fark eden ve kendisine bunu hissettiren insanların olması o kişi için en büyük şanstır. Bu anlamda bana inanan ve beni destekleyen tüm bölüm hocalarıma danışmanım sayın Prof. Dr. Selahattin KAÇIRANLAR nezninde teşekkürlerimi sunuyorum.

Çalışmalarım sırasında hep yanımda olan ve aldığım kararları hiçbir zaman irdelemeksizin beni destekleyen aileme de teşekkürlerimi sunuyorum.

İÇİNDEKİLER

SAYFA

ÖZ.....	I
ABSTRACT.....	II
TEŞEKKÜR.....	III
İÇİNDEKİLER.....	IV
TABLolar DİZİNİ.....	VII
ŞEKİLLER DİZİNİ.....	VIII
SİMGELER VE KISALTMALAR.....	IX
1. GİRİŞ.....	1
1.1. Çoklu İç İlişkinin Tanımlanması.....	1
1.2. Çoklu İç İlişkinin Belirlenmesi.....	2
1.3. Çoklu İç İlişkinin İstatistiksel Sonuçları.....	4
1.4. Çözüm Yöntemleri.....	7
2. BAZI YANLI TAHMİN EDİCİLER.....	12
2.1. Ridge Tahmin Edici.....	12
2.1.1. Ridge Tahmin Edicinin Bazı Özellikleri.....	13
2.1.2. k Yanlılık Parametresinin Seçimi.....	16
2.2. Genelleştirilmiş Ridge Tahmin Edici.....	19
2.3. Temel Bileşenler Tahmin Edici.....	20
2.4. Kısıtlı En Küçük Kareler Tahmin Edici.....	23
2.5. Modified Ridge Tahmin Edici.....	25
2.6. Yansız Ridge Tahmin Edici.....	27
2.7. Kısıtlı Ridge Tahmin Edici.....	27
2.8. r-k Sınıf Tahmin Edici.....	29
2.9. Liu Tahmin Edici.....	30
2.10. r-d Sınıf Tahmin Edici.....	32
2.11. Liu-Tipi Tahmin Edici.....	32
3. BAZI YANLI TAHMİN EDİCİLERİN KARŞILAŞTIRILMASI.....	34

3.1. r-d Sınıf Tahmin Edicinin Matris HKO Kriterine Göre Bazı Tahmin Ediciler Üzerindeki Üstünlüğü.....	35
3.1.1. r-d Sınıf Tahmin Edici ile EKK Tahmin Edicinin Matris HKO Kriterine Göre Karşılaştırılması.....	36
3.1.2. r-d Sınıf Tahmin Edicinin Temel Bileşenler Tahmin Edicisi ile Matris HKO Kriterine Göre Karşılaştırılması.....	38
3.1.3. r-d Sınıf Tahmin Edici ile Liu Tahmin Edicinin Matris HKO Kriterine Göre Karşılaştırılması.....	40
3.1.4. r-d Sınıf Tahmin Edicinin Temel Bileşenler ve Liu Tahmin Edicilerinin Her İkisinden Matris HKO Anlamında İyi Olması İçin Kriter.....	42
3.1.5. Koşullar İçin Hipotez Testleri.....	43
3.1.6. r-d ve r-k Sınıf Tahmin Edicilerin Matris HKO Kriterine Göre Karşılaştırılması.....	46
3.2. Yansız Ridge Tahmin Edicinin Diğer Tahmin Edicilerle Karşılaştırılması....	48
3.2.1. Yansız Ridge Tahmin Edicinin Ridge Tahmin Edici ile Matris HKO Kriterine Göre Karşılaştırılması.....	49
3.2.2. Yansız Ridge ve EKK Tahmin Edicilerinin Matris HKO Kriterine Göre Karşılaştırılması.....	52
3.2.3. Yansız Ridge Tahmin Edici ile $r - k$ Sınıf Tahmin Edicinin Matris HKO Kriterine Göre Karşılaştırılması.....	52
3.2.4. Rezidülerin İncelenmesi ve Analizi.....	54
3.2.5. Yansız Ridge ve Ridge Rezidülerinin Kovaryans Yapılarının Karşılaştırılması.....	55
3.2.6. $\hat{e}_{UR}$ ve $\hat{e}_R$ nin Matris HKO Kriterine Göre Karşılaştırılması.....	58
3.2.7. Yansız Ridge ve EKK Rezidülerinin Kovaryans Yapılarının Karşılaştırılması.....	59
3.2.8. $\hat{e}_{UR}$ ve $\hat{e}$ nin Matris HKO Kriterine Göre Karşılaştırılması.....	60
3.2.9. Yansız Ridge ve $r - k$ Sınıf Rezidülerin Kovaryans Yapılarının Karşılaştırılması.....	60
3.2.10. $\hat{e}_{UR}$ ve $\hat{e}_{r-k}$ nin Matris HKO Kriterine Göre Karşılaştırılması.....	61

3.2.11. Sayısal Örnek.....	62
4. LİU TAHMİNDE YANLILIK PARAMETRESİNİN SEÇİMİ İÇİN ÖN TAHMİN AMAÇLI BİR KRİTER.....	67
4.1. Kalite Kontrolde Çoklu İç İlişki Problemleri.....	67
4.2. Ön Tahmin Amaçlı Kriterler.....	71
4.2.1. Liu Tahmin Edicide d nin Seçimi İçin PRESS İstatistiği.....	73
4.2.2. Sayısal Örnek.....	81
5. SONUÇ VE ÖNERİLER.....	87
KAYNAKLAR.....	89
ÖZGEÇMİŞ.....	95

TABLolar DİZİNİ**SAYFA**

Tablo 3.1. Veri kümesi.....	62
Tablo 3.2. $SMSE_1$ ve $SMSE_2$ sırasıyla tahmin ediciler ve rezidüler için $SMSE$ yi göstermek üzere $\hat{k}_{hkb} = 0.0133$ ve $\hat{k}_{cjh} = 0.0436$ için $SMSE$ ve tahminlerin değerleri.....	63
Tablo 3.3. $SMSE_1$ ve $SMSE_2$ sırasıyla tahmin ediciler ve rezidüler için $SMSE$ yi göstermek üzere $k_1 = 0.0027$ için $SMSE$ ve tahminlerin değerleri.....	65
Tablo 4.1. k nın çeşitli değerleri için $PRESS_R(k)$ değerleri.....	82
Tablo 4.2. d nin çeşitli değerleri için $(1 - d)PRESS_R(1) + dPRESS$ değerleri.....	84
Tablo 4.3. k ve d nin çeşitli değerleri için $PRESS_R(k)$ ve $PRESS_d$ değerleri.....	85
Tablo 4.4. Tahmin özetleri.....	86

ŞEKİLLER DİZİNİ

SAYFA

Şekil 2.1. $\hat{\beta}(k)$ ve $\hat{\beta}$ nın SMSE sinin k ya bağlı grafiği (Hoerl ve Kennard, 1970).....	15
Şekil 2.2. Negatif ilişkili regresyon katsayılarının ridge trace-i (Vinod ve Ullah, 1981).....	16
Şekil 2.3. $p = 2$ için temel bileşenler (Myers, 1990).....	21
Şekil 3.1. k ya karşı EKK, ridge ve yansız ridge tahmin edicilerin SMSE grafikleri.....	64
Şekil 3.2. k ya karşı EKK, yansız ridge ve $r - k$ sınıf tahmin edicilerin SMSE grafikleri.....	64
Şekil 3.3. k ya karşı $r - k$ sınıf tahmin edicinin SMSE grafiği.....	65
Şekil 3.4. k ya karşı EKK, yansız ve ridge rezidü için SMSE grafikleri.....	66
Şekil 3.5. k ya karşı EKK, yansız ve $r - k$ sınıf rezidülerin SMSE grafikleri.....	66
Şekil 4.1. k ya karşı $PRESS_R(k)$ nın grafiği.....	83
Şekil 4.2. k ve d ye karşı sırasıyla $PRESS_R(k)$ ve $PRESS_d$ grafikleri.....	84
Şekil 4.3. d ye karşı $PRESS_d$ ve $PRESS$ grafikleri.....	85

SİMGELER VE KISALTMALAR

EKK : En Küçük Kareler

HKO : Hata Kareler Ortalaması

CUSUM: Kümülatif Toplam (Cumulative Sum)

MRRC : Çok Değişkenli Ridge Rezidü Grafikleri (Multivariate Ridge Residual Charts)

EWMA : Üstel Ağırlıklı Hareketli Ortalama (Exponentially Weighted Moving Average)

PRESS : Ön Tahmin Hata Kareler Toplamı (Prediction Error Sum of Squares)

GCV : Genelleştirilmiş Çapraz Geçerlilik (Generalized Cross-Validation)

1. GİRİŞ

1.1. Çoklu İç İlişkinin Tanımlanması

$$y = X\beta + e \quad (1.1)$$

şeklinde bir çoklu lineer modelimizin olduğunu varsayalım. Burada y ($n \times 1$) yanıt değişkenlerin vektörü, X ($n \times p$) stokastik olmayan açıklayıcı değişkenlerin gözlenen matrisi, β ($p \times 1$) bilinmeyen parametreler vektörü ve e sıfır ortalamalı, σ^2 varyanslı rasgele hataların ($n \times 1$) vektörüdür. Bir çoklu lineer model yeni gözlemlerin ön tahmini (prediction), regresyon katsayılarının tahmini (estimation), öngörü (forecasting) gibi amaçlar nedeniyle pek çok alanda kullanılmaktadır.

Çoklu lineer regresyon modelinde, genellikle açıklayıcı değişkenlerin bağımsız olduğu varsayılır. Fakat uygulamada, lineer ilişki olabilir. Bu durumda açıklayıcı değişkenler arasında bağımsızlık varsayımı geçerli olmaz ve bu, çoklu iç ilişki problemine neden olur.

X_j , X matrisinin j -inci kolon vektörü olsun: $X = [X_1, \dots, X_p]$.

$$\sum_{j=1}^p t_j X_j = 0 \quad (1.2)$$

olacak şekilde hepsi sıfır olmayan $t_1, t_2, \dots, t_p$ sabitler kümesi varsa $X_1, X_2, \dots, X_p$ vektörleri lineer bağımlıdır. X in kolonlarının bir kümesi için (1.2) tam olarak sağlanıyorsa “tam çoklu iç ilişki vardır” (Silvey, 1969). Bu durumda X matrisinin rankı p den küçük olur ve $X'X$ matrisi tekildir. Fakat (1.2) X in kolonlarının bazı alt kümeleri için yaklaşık olarak doğru ise “yaklaşık çoklu iç ilişki” vardır. Eğer açıklayıcı değişkenler arasında bir lineer ilişki yoksa, açıklayıcı değişkenler ortogonaldır denir. $X'X = I$ iken X matrisi ortogonal açıklayıcı değişkenleri içerir.

Çoklu iç ilişkinin pek çok nedeni olabilir (Judge ve ark., 1985; Montgomery ve Peck, 1992). Bunlardan bazıları:

1. Uygulanan veri toplama metodu: Analizci (1.2) ile tanımlı bir bölgenin alt uzayından örneklem almış ise çoklu iç ilişki problemi ile karşılaşacaktır.

2. Modeldeki veya kitledeki zorunluluklar: Kitledeki zorunluluklar daha çok, açıklayıcı değişkenlerin kimyasal veya üretim proseslerinde ortaya çıkar. Örneğin, bir kimyasal reaksiyonun gerçekleşmesi için belli içeriklerin sabit oranlarda olması gibi.

3. Modelin belirlenmesi: X değişkenlerinin değişim aralığı küçük iken bir regresyon modeline polinom terimin eklenmesi çoklu iç ilişkiye neden olabilir.

4. Modelin aşırı tanımlanması (overdefined): Gözlemlerden çok açıklayıcı değişkenlerin olduğu modellere aşırı tanımlanmış model denir. Daha çok tıbbi araştırmalarda ve ekonometrik modellerde ortaya çıkar. Bu gibi durumlarda açıklayıcı değişkenlerden bazılarını atmak gerekmektedir.

Çoklu iç ilişkiyi üç ana başlık altında inceleyeceğiz. Bunlar çoklu iç ilişkinin belirlenmesi, çoklu iç ilişkinin istatistiksel sonuçları ve çözüm yöntemleri şeklindedir.

1.2. Çoklu İç İlişkinin Belirlenmesi

Çoklu iç ilişkinin belirlenmesi için çeşitli yöntemler vardır. Bunlardan bazılarını şu şekilde verebiliriz.

1. Korelasyon matrisinin incelenmesi: $X'X$ in köşegen dışı elemanlarının (r_{ij}) incelenmesi çoklu iç ilişkinin belirlenmesini sağlar. Geometrik olarak r_{ij} , x_i ve x_j arasındaki açının kosünüsüdür (Farrar ve Glauber, 1967). x_i ve x_j açıklayıcı değişkenleri lineer bağımlı ise $|r_{ij}|$ bire yakın olur.

2. Varyans Şişirme Faktörü (VIF): $(X'X)^{-1}$ matrisinin j -inci köşegen elemanına j -inci varyans şişirme faktörü (VIF_j) denir. İlk olarak Farrar ve Glauber (1967) tarafından çoklu iç ilişkiyi belirlemek için kullanılmış fakat Marquardt (1970) tarafından VIF olarak adlandırılmıştır. R_j^2 , x_j nin diğer $p-1$ açıklayıcı değişken üzerine olan modelinden çoklu belirleyicilik katsayısı olmak üzere $VIF_j = (1 - R_j^2)^{-1}$

olarak ifade edilir (Theil, 1971). x_j diğer açıklayıcı değişkenlerin bazılarına yaklaşık olarak lineer bağımlı iken R_j^2 bire yakınlaşır ve VIF_j büyür. Herhangi bir VIF değeri 10 dan büyükse çoklu iç ilişki vardır. VIF lerin bir diğer ilginç yorumu ise şu şekildedir. $\hat{\beta} = (X'X)^{-1} X'y$, β ortalama vektörlü ve $\sigma^2(X'X)^{-1}$ varyanslı dağılıma sahip olduğundan j -inci regresyon katsayısı üzerindeki güven aralığının uzunluğu $L_j = 2(VIF_j \hat{\sigma}^2)^{1/2} t_{\alpha/2, n-p-1}$ olarak yazılabilir. Dolayısıyla, VIF_j nin karekökü çoklu iç ilişkiden dolayı j -inci regresyon katsayısı için güven aralığının ne kadar büyük olduğunu gösterir.

3. $X'X$ matrisinin özdeğerlerinin analizi: Çoklu iç ilişki problemini ciddi anlamda çalışan ilk araştırmacı Ragnar Frisch'dır. Frisch (1934) çoklu iç ilişkiyi özdeğerlerle ilişkilendirmiştir. Fakat bilgisayar kullanımının kısıtlı olması nedeniyle $X'X$ in özdeğerlerinin nümerik analizi desteklenememiştir (Vinod ve Ullah, 1981). Eğer veride bir veya daha fazla lineer bağımlılık varsa, bir veya daha fazla özdeğer küçük olacaktır. Küçük özdeğer, X in kolonları arasındaki yaklaşık lineer bağımlılığı gösterir. $\lambda_{\min}$ ve $\lambda_{\max}$, $X'X$ matrisinin minimum ve maksimum özdeğerleri olmak üzere $K = \frac{\lambda_{\max}}{\lambda_{\min}}$ olarak tanımlanan koşul sayısı çoklu iç ilişkiyi belirlemek için kullanılır. K , 100 den büyükse şiddetli bir çoklu iç ilişki vardır.

4. Diğer tanılama (diagnostics) metotları: Çoklu iç ilişki pek çok teknikle belirlenebilir. Bunlardan birisi $X'X$ in determinantıdır. Standartlaştırılmış formda $0 \leq |X'X| \leq 1$ dir. $|X'X| = 1$ ise açıklayıcı değişkenler ortogonaldır aksine $|X'X| = 0$ iken tam çoklu iç ilişki vardır. $|X'X|$ sifıra yaklaştıkça çoklu iç ilişkinin şiddeti artacaktır (Farrar ve Glauber, 1967).

Regresyon katsayılarının işaretlerinin veya büyüklüklerinin beklenenden farklı olması, açıklayıcı değişkenlerin eklenmesi veya çıkarılması ile regresyon katsayılarının tahminlerinde büyük değişikliklere neden olması çoklu iç ilişkinin göstergesi olabilir.

1.3. Çoklu İç İlişkinin İstatistiksel Sonuçları

(1.2) şeklindeki bir lineer ilişkinin varlığı, $c'\beta = c_1\beta_1 + \dots + c_p\beta_p$ gibi lineer fonksiyonların lineer yansız tahmin edicilerinin olamayacağı anlamına gelir (Silvey, 1969). Ancak c' , X in satırlarının bir lineer kombinasyonu olarak yazılabilirse $c'\beta$ nın bir lineer yansız tahmin edicisi bulunabilir. Fakat X tam ranklı değilken c' bu şekilde ifade edilemez. $c'\beta$ nın tahmin edilebilir olması için bir diğer koşul c vektörünün $X'X$ in kolonlarının bir lineer kombinasyonu olarak yazılabilmesidir. Silvey (1969) $c'\beta$ nın tahmin edilebilir olması için gerek ve yeter koşulun $X'X$ in sıfırdan farklı özdeğerlerine karşılık gelen özvektörlerin bir lineer kombinasyonu olarak yazılması olduğunu belirtmiştir.

Farrar ve Glauber (1967) çoklu iç ilişkiyi X in kolonlarının ortogonallikten uzaklaşması olarak tanımlamıştır. Bu şekilde tanımlamalarının temel sebebi istatistiksel sonuçların daha kolay elde edebilmesidir. Silvey (1969) bu tanım doğrultusunda çoklu iç ilişkinin istatistiksel sonuçlarını şu şekilde incelemiştir. (1.1) modelini yeniden parametreleme yaparak ortogonal formda yazabiliriz: V matrisi, XV nin kolonları ortogonal olacak şekilde bir tekil olmayan matris olsun. Bu durumda (1.1) modelini

$$y = X\beta + e = (XV)(V^{-1}\beta) + e \quad (1.3)$$

formunda yazabiliriz. V matrisinin pek çok seçimi vardır. Özel olarak kolonları $X'X$ matrisinin ortonormal özvektörleri olan T ortogonal matrisi olsun. Böylece $Z = XT$ ve $\alpha = T'\beta$ olmak üzere (1.3) modelini

$$y = Z\alpha + e \quad (1.4)$$

olarak yazabiliriz. $\lambda_1, \dots, \lambda_p$ $X'X$ matrisinin özdeğerleri olmak üzere

$$Z'Z = T'X'XT = \Lambda = \text{diag}(\lambda_1, \dots, \lambda_p) \quad (1.5)$$

olduğundan Z nin kolonları ortogonaldır. $X'X$ in J tane özdeğeri sıfır ise X in rankı $p - J$ olacaktır ve XT nin J tane kolonu sıfır olacaktır. Bu ise α nın son J bileşeni $\alpha_{p-J+1}, \dots, \alpha_p$ nin tahmin edilemeyeceği anlamına gelir. Diğer yandan ilk $p - J$ bileşen, $\alpha_1, \dots, \alpha_{p-J}$, ve bunların herhangi bir lineer kombinasyonu tahmin edilebilir. Bu nedenle $c'\beta$, $\alpha_1, \dots, \alpha_{p-J}$ nin lineer kombinasyonu olarak yazılabilirse tahmin edilebilirdir. $c'\beta = c'TT'\beta = (T'c)'\alpha$ olduğundan $c'\beta$ nın tahmin edilebilir olması için gerek ve yeter koşul $T'c$ nin son J bileşeninin sıfır olmasıdır. t sıfırdan farklı λ özdeğerine karşılık gelen $X'X$ in ortonormal bir özvektörü olsun. Bu durumda $t'\beta$ tahmin edilebilirdir ve b^* , $X'Xb^* = X'y$ normal denklemlerinin herhangi bir çözümü olmak üzere $t'\beta$ nın minimum varyans lineer tahmin edicisi $t'b^*$ dir. $t'X'X = \lambda t'$ kullanarak $t'X'y = t'X'Xb^* = \lambda t'b^*$ elde edilir ve $\lambda^2 \text{var}(t'b^*) = t'X'\text{var}(y)Xt = \sigma^2 t'X'Xt = \lambda \sigma^2$ olduğundan $t'b^*$ ın varyansı σ^2/λ dır. t_1 ve t_2 $X'X$ in sıfırdan farklı iki ortogonal özvektörü ise $\text{cov}(t_1'b^*, t_2'b^*) = 0$ dır. Böylece $c'\beta$ tahmin edilebilir ise yani k_i ler sabitler, t_i ler sıfırdan farklı $\lambda_1, \dots, \lambda_{p-J}$ özdeğerlerine karşılık gelen özvektörler olmak üzere $c = k_1 t_1 + \dots + k_{p-J} t_{p-J}$ iken $c'b^*$ tahmin edicisinin varyansı

$$\text{var}(c'b^*) = \sigma^2 \left(\frac{k_1^2}{\lambda_1} + \dots + \frac{k_{p-J}^2}{\lambda_{p-J}} \right) \quad (1.6)$$

dır. (1.6) ifadesi tahmin edilebilen ve edilemeyen fonksiyonlar arasındaki farkı açık bir şekilde göstermektedir, tahmin edilebilir olanlar $X'X$ in büyük özdeğerlerine karşılık gelen özvektörler yönünde olacaktır ve tahmini mümkün olmayanlar ise küçük özvektörlere karşılık gelen özvektörler yönünde olacaktır. Yani, Judge ve ark.'nın (1985) belirttiği gibi, eğer c küçük özdeğerlere karşılık gelen büyük k_i

ağırlıklarına sahipse tahmin edicinin varyansı büyük olacaktır ve $c' \beta$ güvenilebilir bir tahmine sahip olmayacaktır. Bu yaklaşım X in kolonları arasında tam lineer kombinasyon olmadığı durumda da geçerlidir. Bu nedenle, $c' \beta = (k_1 t_1 + \dots + k_p t_p)' \beta$ formundaki tüm fonksiyonlar tahmin edilebilirdir ve $\hat{\beta} = (X' X)^{-1} X' y$ en küçük kareler (EKK) tahmin edicisi olmak üzere minimum varyans tahmin edicisi $c' \hat{\beta}$ olup varyansı

$$\text{var}(c' \hat{\beta}) = \sigma^2 c' (X' X)^{-1} c = \sigma^2 c' \left[\sum_{i=1}^p \lambda_i^{-1} t_i t_i' \right] c = \sigma^2 \sum_{i=1}^p \frac{k_i^2}{\lambda_i} \quad (1.7)$$

dır. (1.7) ifadesi (1.6) ya benzer şekilde yorumlanabilir yani, çoklu iç ilişki olması durumunda neden bir modelin kötü bir tahmin edici olduğunu göstermektedir.

Regresyon katsayılarının varyansının büyük olması genellikle $H_0 : \beta_i = 0$ hipotezinin kabul edilmesinin uygun olacağı anlamına gelir.

Çoklu iç ilişki EKK tahminlerinin mutlak değerce büyük olmasına neden olacaktır (Montgomery ve Peck, 1992). Bunu görmek için $\hat{\beta}$ ile β arasındaki uzaklığın karesini, örneğin $L^2 = (\hat{\beta} - \beta)'(\hat{\beta} - \beta)$ yı inceleyelim.

$$\begin{aligned} E(L^2) &= E(\hat{\beta} - \beta)'(\hat{\beta} - \beta) \\ &= \sum_{j=1}^p E(\hat{\beta}_j - \beta_j)^2 \\ &= \sum_{j=1}^p \text{var}(\hat{\beta}_j) \\ &= \sigma^2 \text{tr}(X' X)^{-1} = \sigma^2 \sum_{j=1}^p \frac{1}{\lambda_j} \end{aligned} \quad (1.8)$$

olduğundan eğer $X' X$ çoklu iç ilişkiden dolayı kötü koşullu ise en az bir λ_j küçük olacaktır ve (1.8) $\hat{\beta}$ dan β ya olan uzaklığın büyük olacağı anlamına gelir. Denk olarak

$$\begin{aligned} E(L^2) &= E(\hat{\beta} - \beta)'(\hat{\beta} - \beta) \\ &= E(\hat{\beta}'\hat{\beta} - 2\hat{\beta}'\beta + \beta'\beta) \end{aligned}$$

veya

$$E(\hat{\beta}'\hat{\beta}) = \beta'\beta + \sigma^2 \text{tr}(X'X)^{-1}$$

olur. Yani, $\hat{\beta}$ vektörü genellikle β vektöründen büyüktür. Bu ise EKK metodunun regresyon katsayılarının mutlak değerce büyük tahminlerini ürettiği anlamına gelir.

Tam çoklu iç ilişki varken, EKK metodunun model parametrelerinin kötü bir tahminini vermesi uydurulan (fitted) modelin kötü bir tahmin edici olduğu anlamına gelmez. Eğer ön tahminler çoklu iç ilişkinin yaklaşık olarak geçerli olduğu x -uzayının bölgesinde sınırlı ise uydurulan model çoğu kez uygun ön tahminler üretecektir. Bunun sebebi, β_j tahminleri kötü olsa bile $\sum_{j=1}^p \beta_j x_{ij}$ lineer

kombinasyonunun gerçekten iyi tahmin edilebilmesidir. Yani, eğer orijinal veri (1.2) ile tanımlı hiperdüzlemde tanımlı ise model parametrelerinin yetersiz tahminlerine rağmen sonraki (future) gözlemler, hiperdüzlemin yakınına düşse bile, çoğu kez tam tahmin edilebilir.

1.4. Çözüm Yöntemleri

Çoklu iç ilişkiyi ortadan kaldırmak için kullanılan metotlardan bazıları şu şekilde özetlenebilir:

1. Ek verinin toplanması: Farrar ve Glauber (1967) ve Silvey (1969) çoklu iç ilişkiyi ortadan kaldıracak şekilde ek verinin toplanmasını önermişlerdir. Fakat bu, modeldeki veya kitledeki kısıtlamalardan dolayı her zaman mümkün olmayabilir. (1.1) ile tanımladığımız modelde çoklu iç ilişki varsa, $X'Xq = 0$ olacak şekilde bir q vektörü vardır ve $q'\beta$ tahmin edilemez. $X'Xq = 0$ olması için gerek ve yeter koşul $Xq = 0$ dır, X in her bir satırı q ye ortogonal ise çoklu iç ilişki problemi vardır. Bu nedenle q ye ortogonal olmayan yeni bir gözlem almamız gerekir. Uygun

seçimlerden birisi, l bir sabit olmak üzere $x_{q+1} = lq$ dır. Benzer bir yaklaşım $X'X$ in özdeğeri sıfır değilken fakat çok küçükken de geçerlidir. $X'X$ in bir küçük özdeğeri λ ve karşılık gelen özvektörü q olsun. Bu durumda $x_{q+1} = lq$ de y_{q+1} gözlemini alabiliriz. Böylece yeni model

$$\begin{bmatrix} y \\ y_{q+1} \end{bmatrix} = \begin{bmatrix} X \\ x'_{q+1} \end{bmatrix} \beta + \begin{bmatrix} e \\ e_{q+1} \end{bmatrix}$$

veya $y_* = X_* \beta + e_*$ olur. Buradan $X_*' X_* = X'X + x_{q+1} x'_{q+1} = X'X + l^2 q q'$ ve $X_*' X_* q = X'X q + l^2 q q' q = \lambda q + l^2 q = (\lambda + l^2) q$ dır. Yani q , $X_*' X_*$ ın $l^2 + \lambda$ özdeğerine karşılık gelen özvektördür. Böylece q yönünde yeni bir gözlem seçerek belirsiz yöndeki tahmin gelişmiş olur. Birden fazla çok küçük özdeğer olması durumunda da bu durum geçerlidir.

2. Modelin yeniden belirlenmesi: Çoklu iç ilişkinin nedeni model seçiminden kaynaklanabilir, örneğin ilişkili açıklayıcı değişkenlerin kullanılması gibi. Bu durumda ya açıklayıcı değişkenler yeniden tanımlanır ya da ilişkili açıklayıcı değişkenlerin biri çıkartılır. Fakat açıklayıcı değişkenlerin çıkartılması modelin etkinliğini azaltabilir. Çünkü Lipovetsky ve Conklin'in (2001) belirttiği gibi çoklu iç ilişki olsa bile değişkenler birbirini tam olarak temsil etmeyebilirler. Açıklayıcı değişkenlerin her biri uyumda (fitting) ve bağımlı değişkenin yapısını açıklamada özel bir role sahip olabilir.

3. Alternatif regresyon analizinin kullanılması: Regresyon analizinde asıl olarak regresyon katsayıları ile ilgileniyorsak EKK en çok kullanılan yöntemdir. Bilindiği gibi (1.1) modeli için EKK tahmin edici,

$$\hat{\beta} = (X'X)^{-1} X'y \quad (1.9)$$

dır. $\hat{\beta}$ tahmin edicisi rezidü kareler toplamı minimum olacak şekilde bulunur:

$$S(\beta) = (y - X\beta)'(y - X\beta).$$

$S(\beta)$ nın β ya göre türevini aldığımız da

$$X'X\beta = X'y \quad (1.10)$$

normal denklemlerini buluruz. X tam kolon ranklı, $rank(X) = p$, ise $X'X$ tekil değildir ve (1.10) un çözümü (1.9) ile verilen EKK tahmin edicidir.

$\hat{\beta}$ nın bazı özelliklerini şu şekilde verebiliriz:

1. Yansız bir tahmin edicidir. $E(e) = 0$ olduğundan

$$\begin{aligned} E(\hat{\beta}) &= E[(X'X)^{-1}X'y] = E[(X'X)^{-1}X'(X\beta + e)] \\ &= \beta + (X'X)^{-1}X'E(e) = \beta. \end{aligned}$$

2. Gauss-Markov teoremi gereğince $\hat{\beta}$, β nın tüm lineer yansız tahmin edicileri arasında minimum varyansa sahiptir (BLUE):

$$Var(\hat{\beta}) = Var[(X'X)^{-1}X'y] = \sigma^2(X'X)^{-1}. \quad (1.11)$$

Gauss-Markov teoreminin en önemli özelliği dağılımsal genelleştirilebilirliğidir. (1.1) modeli için hataların normal dağıldığı varsayımı da yapılırsa β nın maksimum olabilirlik (likelihood) tahmin edicisi yine (1.9) ile verilen $\hat{\beta}$ dır. Maksimum olabilirlik tahmin edicisi β nın yansız tahmin edicileri içerisinde minimum varyansa sahiptir.

σ^2 nin EKK tahmini ise şöyledir.

$$\hat{e} = y - X\hat{\beta} = [I_n - X(X'X)^{-1}X']y \quad (1.12)$$

olup (1.1) modeli göz önünde bulundurulursa (1.12) yeniden

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

şeklinde yazılabilir. $M = I_n - X(X'X)^{-1}X'$ simetrik, idempotent bir matris olduğundan

$$\hat{e}'\hat{e} = e'Me \quad (1.13)$$

dır. İz operatörü kullanılırsa ve (1.13) in beklenen değeri alınırsa

$$\begin{aligned} E(\hat{e}'\hat{e}) &= E(e'Me) = E(\text{tr}(e'Me)) = E(\text{tr}(Mee')) = \text{tr}(ME(ee')) \\ &= \sigma^2 \text{tr}[I_n - X(X'X)^{-1}X'] = \sigma^2 [\text{tr}(I_n) - \text{tr}(I_p)] = \sigma^2 (n - p) \end{aligned}$$

bulunur. Sonuç olarak $E\left(\frac{\hat{e}'\hat{e}}{n-p}\right) = \sigma^2$ olduğundan

$$\hat{\sigma}^2 = \frac{\hat{e}'\hat{e}}{n-p} = \frac{y[I_n - X(X'X)^{-1}X']y'}{n-p} \quad (1.14)$$

dır ve $\hat{\sigma}^2$, σ^2 nin yansız tahmin edicisidir.

İç ilişki olması durumunda en küçük kareler hala yansızdır fakat varyansı çok büyüktür. Bu nedenle, gerçek değerinden uzaklaşır. Bu problemi ortadan kaldırmak için önerilen tahmin prosedürleri yanlış tahmin edicilerin ortaya çıkmasına ve bunların incelenmesine neden olmuştur. Burada bu tahmin edicilerden bazılarını kısa bir giriş yapabiliriz.

Pek çok tahmin edici olması nedeniyle bunların karşılaştırması ile ilgileniriz. Gauss 1809'da tahmin edicilerin seçimi için en uygun kriteri hata kareler ortalaması (HKO) olarak önermiştir. β nin herhangi bir $\tilde{\beta}$ tahmin edicisi için HKO matrisi

$$MMSE(\tilde{\beta}, \beta) = E(\tilde{\beta} - \beta)(\tilde{\beta} - \beta)' \quad (1.15)$$

olarak tanımlanır. $\tilde{\beta}$ nin kovaryans matrisini $Var(\tilde{\beta})$ ile ve β ile $E(\tilde{\beta})$ arasındaki farkı $Bias(\tilde{\beta})$ ile gösterirsek (1.15)

$$MMSE(\tilde{\beta}, \beta) = Var(\tilde{\beta}) + Bias(\tilde{\beta})Bias(\tilde{\beta})' \quad (1.16)$$

olur. $\tilde{\beta}$ nin skaler HKO su, MMSE nin izi olarak tanımlanır:

$$SMSE(\tilde{\beta}, \beta) = E(\tilde{\beta} - \beta)'(\tilde{\beta} - \beta) = trVar(\tilde{\beta}) + Bias(\tilde{\beta})' Bias(\tilde{\beta}).$$

$SMSE(\tilde{\beta}, \beta)$ nin temel kriter olarak alınması $MMSE(\tilde{\beta}, \beta)$ nin köşegen dışı elemanlarını göz ardı etmemiz ve $(\tilde{\beta}_i - \beta_i)^2$ lerin tamamını eşit ağırlıklandırmamız anlamına gelir. Bu nedenle $MMSE$ daha iyi bir kriterdir fakat hesaplanması daha zordur.

2. BAZI YANLI TAHMİN EDİCİLER

2.1. Ridge Tahmin Edici

Ridge regresyon $X'X$ matrisinin köşegen elemanlarına pozitif küçük k değerleri ekleyerek koşul sayısını küçültmeyi amaçlar. Hoerl ve Kennard'ın (1970) önerdiği ridge tahmin edici,

$$\hat{\beta}(k) = (X'X + kI)^{-1} X'y, \quad k \geq 0 \quad (2.1)$$

dır. Ridge tahmin edici k yanlılık parametresine bağlıdır ve k nın seçimi tahmin edicinin performansını etkiler.

Hoerl ve Kennard (1970) ridge tahmin ediciyi şu şekilde elde etmişlerdir. $\tilde{\beta}$ herhangi bir tahmin edici olmak üzere $\tilde{\beta}$ nın hata kareler toplamı, $\hat{\beta}$ nın optimallik özelliğine göre,

$$\begin{aligned} S(\tilde{\beta}) &= (y - X\tilde{\beta})'(y - X\tilde{\beta}) \\ &= (y - X\hat{\beta})'(y - X\hat{\beta}) + (\hat{\beta} - \tilde{\beta})' X' X (\hat{\beta} - \tilde{\beta}) \\ &= S(\hat{\beta}) + \Phi(\tilde{\beta}) \end{aligned}$$

olarak yazılabilir. Burada $(\hat{\beta} - \tilde{\beta})' X' X (\hat{\beta} - \tilde{\beta})$, $\hat{\beta}$ yerine $\tilde{\beta}$ nın kullanılmasından dolayı kaynaklanan yanlılığın karesidir. $X'X$ kötü koşullu iken $\hat{\beta}$ ile β arasındaki uzaklık artmaktadır. Bu nedenle $\tilde{\beta}'\tilde{\beta}$ tahmin edicisinin uzaklığının karesini minimum yapmak isteriz. Fakat minimumlaştırma keyfi değildir. Rezidü kareler toplamının verilen belli bir seviyesi için uzaklığın karesini minimum yapmak isteriz (aynı rezidü kareler toplamına sahip pek çok tahmin edici olabilir, fakat bunlar içerisinde en küçük uzaklıklı olanı seçmek istiyoruz). $\Phi_0 > 0$ hata kareler toplamı için verilen bir sabit olsun. Bu durumda $S(\tilde{\beta}) = S(\hat{\beta}) + \Phi_0$ koşulunu sağlayan

tahmin edicilerin bir $\{\tilde{\beta}\}$ kümesi vardır. Bu küme içerisinde en küçük uzunluklu $\tilde{\beta}$ tahminini bulmak isteriz. Yani $1/k$ Lagrange çarpanı olmak üzere

$$\text{Min}_{\tilde{\beta}} \left\{ \tilde{\beta}' \tilde{\beta} + \frac{1}{k} \left[(\hat{\beta} - \tilde{\beta})' X' X (\hat{\beta} - \tilde{\beta}) - \Phi_0 \right] \right\} \quad (2.2)$$

ifadesini araştırırız. (2.2) nin $\tilde{\beta}$ ve $1/k$ ya göre türevleri

$$\tilde{\beta} + \frac{1}{k} (X' X) (\tilde{\beta} - \hat{\beta}) = 0 \quad (2.3)$$

ve

$$\Phi_0 = (\hat{\beta} - \tilde{\beta})' X' X (\hat{\beta} - \tilde{\beta}) \quad (2.4)$$

normal denklemlerini verir. Böylece (2.2) nin çözümü, (2.1) ile verilen ridge tahmin edicidir. k yarıllık parametresi ise (2.4) eşitliğini sağlayacak şekilde seçilmelidir (Rao ve Toutenburg, 1995).

Benzer şekilde r^2 sabit olmak üzere

$$\text{Min}_{\tilde{\beta}} \left\{ (y - X\tilde{\beta})' (y - X\tilde{\beta}) + \frac{1}{k} [\tilde{\beta}' \tilde{\beta} - r^2] \right\}$$

ifadesi ile de (2.1) bulunabilir.

2.1.1. Ridge Tahmin Edicinin Bazı Özellikleri

1. $\hat{\beta}(k)$, yarıçapı $\hat{\beta}(k)$ nın uzunluğu olan orijin merkezli küre üzerindeki rezidü kareler toplamını minimum yapar.
2. Rezidü kareler toplamı k nın artan bir fonksiyonudur.
3. $\hat{\beta}(k)' \hat{\beta}(k) < \hat{\beta}' \hat{\beta}$ ve $k \rightarrow \infty$ iken $\hat{\beta}(k)' \hat{\beta}(k) \rightarrow 0$ dır.

4. $\lambda_1 \geq \dots \geq \lambda_p$, $X'X$ in özdeğerleri olmak üzere $X'X + kI$ nın en büyük özdeğerinin en küçük özdeğere oranı $(\lambda_1 + k)/(\lambda_p + k)$ olup k nın azalan bir fonksiyonudur. Bu oranın kareköküne X in koşul sayısı denir (Judge ve ark., 1985).
5. Ridge tahmin edici $k = 0$ için EKK tahmin edicisini vermektedir. Ayrıca ridge tahmin edici EKK tahmin edicinin bir lineer dönüşümü olarak da yazılabilmektedir:

$$\hat{\beta}(k) = (X'X + kI)^{-1} X'X\hat{\beta}.$$

6. Hoerl ve Kennard (1970) ridge tahmin edicinin toplam varyansının k nın sürekli, monoton azalan bir fonksiyonu ve yanlılığın karesinin k nın sürekli, monoton artan bir fonksiyonu olduğunu göstermiştir. Bu nedenle varyanstaki azalma yanlılığın karesindeki artıştan fazla olduğu sürece ridge tahmin edicinin iyi bir teknik olduğu söylenebilir (Şekil 2.1). $S_k = X'X + kI$ olsun. Bu durumda

$$E(\hat{\beta}(k)) = S_k^{-1} X'X\beta$$

$$\text{Bias}(\hat{\beta}(k), \beta) = -kS_k^{-1} \beta$$

$$\text{Var}(\hat{\beta}(k)) = \sigma^2 S_k^{-1} X'X S_k^{-1}$$

olduğundan ridge tahmin edicinin matris hata kareler ortalaması

$$\text{MMSE}(\hat{\beta}(k), \beta) = S_k^{-1} (\sigma^2 X'X + k^2 \beta\beta') S_k^{-1}$$

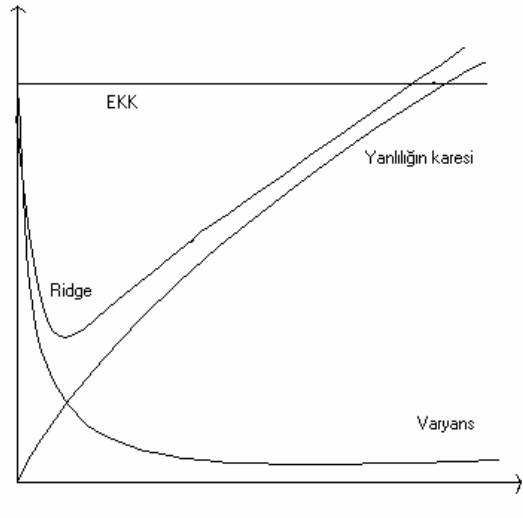
olur. $\lambda_1, \dots, \lambda_p$, $X'X$ in özdeğerleri olmak üzere skaler hata kareler ortalaması

$$\begin{aligned} \text{SMSE}(\hat{\beta}(k), \beta) &= \text{tr}(\text{MMSE}(\hat{\beta}(k), \beta)) \\ &= \sum_{i=1}^p \frac{\sigma^2 \lambda_i}{(\lambda_i + k)^2} + k^2 \beta' (X'X + kI)^{-2} \beta \end{aligned} \quad (2.5)$$

veya

$$SMSE(\hat{\beta}(k), \beta) = tr(MMSE(\hat{\beta}(k), \beta)) = \sum_{i=1}^p \frac{\sigma^2 \lambda_i + k^2 \beta_i^2}{(\lambda_i + k)^2}$$

olarak elde edilir. (2.5) teki ilk terim toplam varyans, ikinci terim yanlılığın karesi toplamıdır.



Şekil 2.1. $\hat{\beta}(k)$ ve $\hat{\beta}$ nın SMSE sinin k ya bağlı grafiği (Hoerl ve Kennard, 1970)

7. Sabit k için $\lim_{\beta' \beta \rightarrow \infty} MMSE(\hat{\beta}(k), \beta) = \infty$ olduğundan $\hat{\beta}(k)$ sınırsız riske sahip değildir.

8. $\hat{\beta}(k)$ nın HKO sı $\hat{\beta}$ dan daha iyi olacak şekilde her zaman bir $k > 0$ vardır. Hoerl ve Kennard (1970) karesel kayıp fonksiyonu altında 8. özelliğin gerçekleşmesi için bir yeter koşulu $k < \sigma^2 / \alpha_{maks}^2$ olarak vermişlerdir. Burada, α_{maks}^2 (1.4) modeli ile verilen $\alpha = T' \beta$ nın en büyük elemanıdır. Theobald (1974) bu sonucu ağırlıklı karesel kayıp fonksiyonuna genişletmiş ve yeter koşulu $k < 2\sigma^2 / \beta' \beta$ olarak bulmuştur.

(8) özelliği bazı ridge tahmin edicilerin EKK'dan daha iyi olduğunun bir göstergesidir. Fakat $\hat{\beta}(k)$ nın $\hat{\beta}$ üzerindeki üstünlüğü sadece parametre uzayının sınırlı bir bölgesinde geçerlidir ve bu bölge bilinmeyen β ve σ^2 tahminlerine

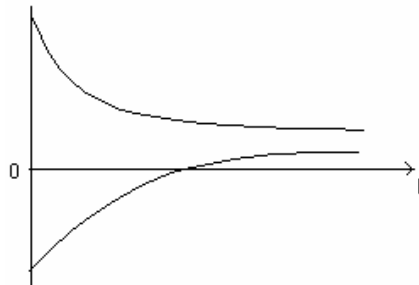
bağlıdır. Sonuç olarak pratik amaçla (8) özelliği kullanışsızdır. Fakat EKK'dan daha iyi olacak şekilde bir k yı veriye bağlı olarak seçmek mümkündür. Bu gibi bir tahmin ediciyi yani farklı verilerde farklı k nın uygun olduğu tahmin ediciyi Judge ve ark. (1985) “uyarlamacı (adaptive) ridge tahmin edici” olarak adlandırmışlardır.

2.1.2. k Yanlılık Parametresinin Seçimi

Ridge regresyonda k nın seçimi için pek çok prosedür regresyon katsayılarının tahmininde önemli bir gelişme olacak şekilde önerilir. k stokastik iken bunlardan bazıları şu şekildedir:

1. *Ridge Trace*: Hoerl ve Kennard (1970), k nın seçimi için *ridge trace* metodunu önermiştir. Ridge trace k ya karşı $\hat{\beta}(k)$ nın elemanlarının grafiğidir. k arttıkça tahminler küçülür ve belli bir k değerinden sonra durağan (stable) olurlar. $\hat{\beta}(k)$ nın durağan olduğu en küçük k değeri uygun bir değer olarak alınır. Hoerl ve Kennard (1970) $0 \leq k \leq 1$ aralığını önermişlerdir. Fakat her veri için bu aralık geçerli olmayabilir (Vinod ve Ullah, 1981).

Ridge trace-in değişkenliği (instability) çoklu iç ilişkiden kaynaklanan açıklayıcı değişkenler arası iç ilişkiyi gösterir. Örneğin, Şekil 2.2 de zıt yönde hareket eden iki açıklayıcı değişken görülmektedir. Bunun sebebi bu iki açıklayıcı değişken arasındaki negatif ilişkidir. Buradan da görüldüğü gibi ridge trace-in değişken olduğu bölge çoklu iç ilişkiden etkilenmektedir.



Şekil 2.2. Negatif ilişkili regresyon katsayılarının ridge trace-i (Vinod ve Ullah, 1981)

Hoerl ve Kennard (1970) farklı birimlerdeki regresyon katsayılarını karşılaştırmak amacıyla veriyi standartlaştırdıktan sonra ridge trace-i grafiklendirmeyi önermiştir.

2. Hoerl ve Kennard (1970) ridge tahmin edicinin EKK'dan daha küçük HKO ya sahip yanlılık parametresinin tahminini $\hat{k}_{HK} = \frac{\hat{\sigma}^2}{\hat{\alpha}_{maks}^2}$ olarak vermiştir. Burada $\hat{\alpha}_{maks}$,

kanonik formda regresyon katsayısının tahmininin maksimum değeridir.

3. Theobald (1974) her W için $E[(\hat{\beta}(k) - \beta)'W(\hat{\beta}(k) - \beta)]$, $E[(\hat{\beta} - \beta)'W(\hat{\beta} - \beta)]$

dan daha küçük olacak şekilde yanlılık parametresinin tahminini $\hat{k}_T = \frac{2\hat{\sigma}^2}{\hat{\beta}'\hat{\beta}}$ olarak

vermiştir.

4. Hoerl, Kennard ve Baldwin (1975) ridge tahmin edicinin HKO değerini minimum yapan $k_i = \frac{\sigma^2}{\alpha_i^2}$ değerlerinin harmonik ortalaması $\hat{k}_{HKB} = \frac{p\hat{\sigma}^2}{\hat{\beta}'\hat{\beta}}$ y1 önermiştir.

Simülasyon çalışması yaparak $\hat{k}_{HKB}$ için ridge tahmin edicinin EKK dan daha iyi HKO ya sahip olduğunu göstermişlerdir.

5. Lawless ve Wang (1976) bayesian yaklaşımla $\hat{k}_{LW} = \frac{p\hat{\sigma}^2}{\hat{\beta}'X'X\hat{\beta}}$ tahminini

vermişlerdir. Monte Carlo simülasyon çalışması yaparak $\hat{k}_{LW}$ için ridge tahmin edicinin $\hat{k}_{HKB}$ den ve EKK den daha iyi bir performans sergilediğini göstermişlerdir.

6. McDonald ve Galarneau (1975) $\hat{\beta}'(k)\hat{\beta}(k) = \hat{\beta}'\hat{\beta} - \hat{\sigma}^2 tr(X'X)^{-1}$ olacak şekilde k y1 seçmeyi önermişlerdir.

Hatanın normal dağıldığı varsayımı altında, k stokastik değil iken Rubio ve Firinguetti (2002) $\hat{k}_{HK}$, $\hat{k}_{HKB}$, $\hat{k}_{LW}$ için olasılık yoğunluk ve dağılım fonksiyonlarını belirlemişler ve bu stokastik yanlılık parametrelerinin olasılık yoğunluk fonksiyonlarını bir örnek üzerinde karşılaştırmışlardır.

$$\hat{k}_i = \frac{\hat{\sigma}^2}{\hat{\alpha}_i^2} \text{ için olasılık yoğunluk fonksiyonu } \theta_i = \frac{\alpha_i^2 \lambda_i}{\sigma^2} \text{ olmak üzere}$$

$$f(\hat{k}_i) = \frac{e^{-\theta_i/2} (n-p)^{(n-p)/2}}{B(1/2, (n-p)/2)} \frac{(1/\lambda_i)^{(n-p+1)/2} (1/\hat{k}_i)^{3/2}}{((n-p)/\lambda_i + 1/\hat{k}_i)^{(n-p+1)/2}}$$

$$\times \sum_{j=0}^{\infty} \left[\frac{\theta_i/2\hat{k}_i}{(n-p)/\lambda_i + 1/\hat{k}_i} \right]^j \frac{\Gamma((n-p+1)/2 + j)\Gamma(1/2)}{\Gamma(j+1)\Gamma((n-p+1)/2)\Gamma((j+1)/2)}$$

şeklindedir. Dağılım fonksiyonu ise $B_r(a, b) = \int_0^r t^{a-1} (1-t)^{b-1} dt$ eksik (incomplete)

beta fonksiyonu ve $r_i = \frac{\lambda_i}{x_i(n-p) + \lambda_i}$ olmak üzere

$$F(x_i) = 1 - \sum_{j=0}^{\infty} \frac{(\theta_i/2)^j e^{-\theta_i/2}}{j! B(1/2 + j, (n-p)/2)} B_{r_i}((j+1)/2, (n-p)/2)$$

dır.

$\hat{k}_{HKB}$ için olasılık yoğunluk ve dağılım fonksiyonları ise sırasıyla aşağıdaki gibidir. Δ , $|c_i| = |1 - \Delta/h_i| < 1$ $i=1, \dots, p$ olacak şekilde bir sayı ve h_i , $M_1 = X(X'X)^{-2}X'$ in köşegen elemanı olmak üzere $\hat{k}_{HKB} > 0$ için

$$f(\hat{k}_{HKB}) = \sum_{j=0}^{\infty} \frac{\Gamma(j+n/2) a_j [(n-p)/p]^{(n-p)/2} \hat{k}_{HKB}^{((n-p)/2)-1}}{\Gamma((n-p)/2) \Gamma(p/2 + j) \Delta^{(p/2)+j} [(n-p)\hat{k}_{HKB}/p + 1/\Delta]^{j+(n/2)}}$$

ve $\Delta_* = \Delta^{j+(n/2)} ((p/(n-p))\Delta)^{(n-p)/2}$ ve $r = [(n-p)/p]\Delta x$ olmak üzere

$$F(x) = \sum_{j=0}^{\infty} \frac{a_j \Gamma(j + (n/2)) B_r((n-p)/2, j + (p/2))}{\Gamma((n-p)/2) \Gamma((p/2) + j)}, \quad x > 0$$

dır.

$\hat{k}_{LW}$ için ise olasılık yoğunluk ve dağılım fonksiyonları ise sırasıyla aşağıdaki gibidir. $\theta = \hat{\beta}' X' X \hat{\beta} / \sigma^2$ olmak üzere

$$f(\hat{k}_{LW}) = \frac{e^{-\theta/2} p^{p/2} (n-p)^{(n-p)/2}}{B(p/2, (n-p)/2)} \frac{\hat{k}_{LW}^{((n-p)/2)-1}}{(\hat{k}_{LW} (n-p) + p)^{n/2}}$$

$$\times \sum_{j=0}^{\infty} \left(\frac{p \theta / 2}{\hat{k}_{LW} (n-p) + p} \right)^j \frac{\Gamma(n/2 + j) \Gamma(p/2)}{\Gamma(j+1) \Gamma(n/2) \Gamma(p/2 + j)}$$

ve $\gamma = p / [(n-p)x + p]$ ile $B_{\gamma}(p/2 + j, (n-p)/2)$ eksik beta fonksiyonu olmak üzere

$$F(x) = 1 - \sum_{j=0}^{\infty} \frac{(\theta/2)^j e^{-\theta/2}}{\Gamma(j+1)} \frac{\Gamma(n/2 + j)}{\Gamma(p/2 + j) \Gamma(p/2 + j) \Gamma((n-p)/2)} B_{\gamma}(p/2 + j, (n-p)/2)$$

şeklindedir.

2.2. Genelleştirilmiş Ridge Tahmin Edici

Hoerl ve Kennard (1970) genelleştirilmiş ridge tahmin ediciyi $K = \text{diag}(k_i)$, $k_i \geq 0$ olmak üzere

$$\hat{\beta}(K) = (X'X + K)^{-1} X'y$$

olarak tanımlamıştır. Tüm k_i değerleri eşit iken (2.1) ile verilen ridge tahmin edici elde edilir. α_i kanonik formda regresyon katsayılarının tahmini olmak üzere

$k_i = \frac{\sigma^2}{\alpha_i^2}$ $i = 1, \dots, p$ genelleştirilmiş ridge tahmin edicinin HKO değerini minimum

yapar. Hoerl ve Kennard (1970), k_i değerlerinin ardışık seçim yöntemini vermiştir. k_i nin ilk tahmini için EKK ile başlanır:

$$k_i^0 = \frac{\hat{\sigma}^2}{\hat{\alpha}_i^2}, i = 1, \dots, p.$$

k_i^0 kullanılarak $K = \text{diag}(k_1^0, \dots, k_p^0)$ olmak üzere başlangıç genelleştirilmiş ridge tahmin edici ($\hat{\alpha}_i^0$) hesaplanır ve bunun yardımıyla bir sonraki k_i değeri bulunur:

$$k_i^1 = \frac{\hat{\sigma}^2}{(\hat{\alpha}_i^0)^2}, i = 1, \dots, p.$$

Bu şekilde durağan parametre tahminleri elde edilene kadar iterasyona devam edilir. Durağanlığın bir ölçüsü olarak regresyon katsayılarının tahminlerinin uzunluklarının karesi kullanılır. Eğer $i - 1$ den i adımına geçerken parametre tahminlerinin boylarının karesinde herhangi bir değişiklik olmuyorsa iterasyon sonlandırılır. Aksi halde devam edilir.

Genelleştirilmiş ridge tahmin edicide farklı k_i değerleri olduğundan k_i lerin seçimi için ridge trace kullanılamaz.

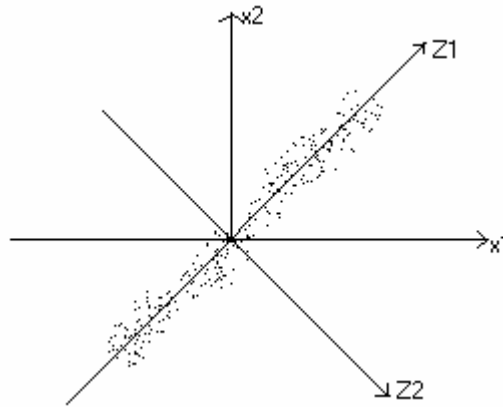
2.3. Temel Bileşenler Tahmin Edici

Çoklu iç ilişkiyi gidermek için ileri sürülen yanlı tahmin edicilerden birisi de temel bileşenler tahmin edicidir. İlk olarak Hotelling (1933) tarafından ele alınmıştır. Bu metotta, korelasyon matrisinin temel bileşenleri olarak adlandırılan yapay (artificial) değişkenlerin bir kümesi üzerinde EKK tahmini uygulanır. Varyansta önemli bir azalma olacak şekilde temel bileşenlerin belli bir sayısı atılır.

(1.4) ile verilen modeli ele alalım. Bu model için $Z = XT$ ve $\alpha = T'\beta$ olarak verilmişti. X_j , X in j -inci kolonu ve t_{ji} , T nin i -inci kolonunun elemanları olmak üzere $Z = XT$ den dolayı Z nin i -inci kolonu

$$Z_i = \sum_{j=1}^p t_{ji} X_j \quad (2.6)$$

dır. Açıklayıcı değişkenlerin ortogonal kümesi olan Z nin kolonlarına, $Z = [Z_1, \dots, Z_p]$, temel bileşenler denir. $Z_i' Z_i = t_i' (X' X) t_i = \lambda_i$ olduğundan $X' X$ in λ_i özdeğeri Z_i temel bileşenindeki değişimi ölçer. $\lambda_i \approx 0$ ise Z_i temel bileşeni model üzerinde küçük bir etkiye sahiptir. λ_i lerin tamamı bir ise (1.5) ten dolayı açıklayıcı değişkenler ortogonal olur fakat X matrisi kötü koşullu iken özdeğerlerden bazıları çok küçük olacaktır. Genelliği kaybetmeksizin $\lambda_{r+1} \approx 0, \dots, \lambda_p \approx 0$ olsun. Bu durumda (2.6) dan dolayı $Z_{r+1}, \dots, Z_p$ ler yaklaşık olarak sabittir. Dolayısıyla modelden atılabilirler. Bu, grafiksel olarak $p=2$ için Şekil 2.3 ile açıklanabilir.



Şekil 2.3. $p=2$ için temel bileşenler (Myers, 1990)

x_1 ve x_2 nin bağımlı olduğu bir veriyi ele alalım. Şekil 2.3 ten dolayı λ_1 büyük, λ_2 küçük özdeğerdir. $\hat{\alpha}_2$ nin varyansı, kanonik form için (1.11) den dolayı, $\sigma^2(1/\lambda_2)$

dır. Dolayısıyla Z_2 yönündeki küçük değişimin nedeni $\hat{\alpha}_2$ nin büyük varyansa sahip olmasıdır. Verideki değişkenliğin çoğu Z_1 eksenı yönündedir. Bu nedenle büyük λ_1 , $Var(\hat{\alpha}_1)$ in x_1 ve x_2 deki bağımlılıktan etkilenmemesine neden olmuştur. Şekil 2.3 ten de görüldüğü gibi veri Z_1 yönündedir ve Z_2 temel bileşeni model üzerinde bir etkiye sahip değildir. Bu nedenle modelden atılabilir.

(1.4) modeli için

$$\Lambda = \begin{pmatrix} \Lambda_r & 0 \\ 0' & \Lambda_{p-r} \end{pmatrix}, \Lambda_r : rxr, \alpha = \begin{pmatrix} \alpha_r \\ \alpha_{p-r} \end{pmatrix}, \alpha_r : rx1, \\ Z = (Z_r, Z_{p-r}), Z_r : nxr, T = (T_r, T_{p-r}), T_r : pxr$$

parçalanışlarını yapalım. Bu durumda (1.4) modelini

$$y = Z_r \alpha_r + Z_{p-r} \alpha_{p-r} + e$$

formunda yazabiliriz. Çoklu iç ilişkinin olması durumunda $Z_{p-r} \alpha_{p-r}$ modelden atılmalıdır. Bunun anlamı α_{p-r} tahmini için $\hat{\alpha}_{p-r} = 0$ ın kullanılması demektir. İlk r katsayı ise EKK ile tahmin edilir, $\hat{\alpha}_r = \Lambda_r^{-1} Z_r' y$. $\beta = T\alpha$ olduğundan β nın tahmin edicisi $\hat{\beta}_{PCR} = (T_r, T_{p-r}) \begin{pmatrix} \hat{\alpha}_r \\ 0 \end{pmatrix} = T_r \hat{\alpha}_r = T_r \Lambda_r^{-1} Z_r' y$ olur. $Z = TX$ olduğundan $Z_r' = T_r' X'$ olur, $X'X = T\Lambda T' = T_r \Lambda_r T_r' + T_{p-r} \Lambda_{p-r} T_{p-r}'$ ve T ortonormal özvektörleri içerdiğinden $\hat{\beta}$ (1.9) ile verilen EKK tahmin edici olmak üzere temel bileşenler tahmin edicisi

$$\hat{\beta}_{PCR} = T_r \Lambda_r^{-1} T_r' X' y = T_r T_r' \hat{\beta} \quad (2.7)$$

dır.

Temel bileşenlerin atılması regresyon katsayılarının varyansında bir azalmaya neden olacaktır.

$$\text{Var}(\hat{\beta}_{PCR}) = \sigma^2 T_r \Lambda_r^{-1} T_r'$$

ve

$$\text{Var}(\hat{\beta}) = \sigma^2 (X'X)^{-1} = \sigma^2 T_r \Lambda_r^{-1} T_r' + \sigma^2 T_{p-r} \Lambda_{p-r}^{-1} T_{p-r}'$$

olduğundan temel bileşenler ile EKK'nın varyans-kovaryans matrisleri arasındaki fark $T_{p-r} \Lambda_{p-r}^{-1} T_{p-r}'$ dir. Bu matrisin köşegen elemanları atılan temel bileşenlere karşılık gelen özdeğerlerin terslerinin ağırlıklı toplamıdır. Sonuç olarak, küçük özdeğerlere karşılık gelen temel bileşenler atılırsa varyansta önemli bir azalma olur.

Temel bileşenler tahmin edicisi $\hat{\alpha}_{p-r} = 0$ alınarak elde edildi. Dolayısıyla $T_{p-r}' \beta = 0$ kısıtlaması altında (1.1) modelinin kısıtlı EKK tahmin edicisinin bulunması ile de (2.7) elde edilebilir.

2.4. Kısıtlı En Küçük Kareler Tahmin Edicisi

X matrisinin tam kolon ranklı olmadığı yani modelin çoklu iç ilişkili olduğu (1.1) modeli için β nın tahmin problemini ele alalım. $\text{rank}(X) = p - J$ olsun. β üzerinde J lineer kısıtlamanın olduğunu varsayalım:

$$R\beta = r \tag{2.8}$$

burada r $J \times 1$ tipinde verilen bir vektör, R $J \times p$ tipinde J ranklı bilinen bir matristir. Ayrıca $\text{rank}\begin{pmatrix} X \\ R \end{pmatrix} = p$ olduğunu varsayalım. Bu durumda (2.8) kısıtlaması verilmişken (1.1) modelindeki β nın tahmini ile ilgileniriz.

EKK tahmin edicinin elde edilmesine benzer mantıkla $R\beta = r$ kısıtlaması altında hata kareler toplamını minimum yapmak isteriz. Yani λ Lagrange çarpanı olmak üzere

$$S(\beta, \lambda) = (y - X\beta)'(y - X\beta) - 2\lambda'(R\beta - r)$$

yi minimum yapmak isteriz. $S(\beta, \lambda)$ nın β ve λ ya göre türevlerini aldığımız da

$$X'X\beta - X'y - R'\lambda = 0 \quad (2.9)$$

$$R\beta - r = 0 \quad (2.10)$$

normal denklemlerini elde ederiz. (2.9) ve (2.10) denklemlerinin çözülmesi ile $\hat{\beta}_r$ ile verilen kısıtlı en küçük kareler tahmin edicisi elde edilir:

$$\hat{\beta}_r = \hat{\beta} - (X'X)^{-1}R'[R(X'X)^{-1}R']^{-1}(R\hat{\beta} - r). \quad (2.11)$$

(2.11) den de görüldüğü gibi kısıtlı EKK tahmin edicisi, EKK tahmin edicisi ile (2.8) kısıtlamasını tam olarak sağlayacak şekilde bir düzeltme teriminin toplamı olarak yazılmaktadır:

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

$R\beta = r$ kısıtlaması doğru iken $\hat{\beta}_r$ tahmin edicisi yansızdır aksi halde yanlıdır:

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

Yanlı olmasına rağmen kısıtlama koyulmasının sebebi varyansı küçültmektir.

$$Var(\hat{\beta}_r) = \sigma^2 (X'X)^{-1} - \sigma^2 (X'X)^{-1} R' [R(X'X)^{-1} R']^{-1} R (X'X)^{-1}$$

olup (1.11) den dolayı

$$Var(\hat{\beta}) - Var(\hat{\beta}_r) = \sigma^2 (X'X)^{-1} R' [R(X'X)^{-1} R']^{-1} R (X'X)^{-1} \quad (2.12)$$

olmaktadır. (2.12) eşitliğinin sağındaki ifade negatif olmayan tanımlı (non-negative definite) bir matris olduğundan lineer kısıtlamaların kullanılması varyansta azalmaya neden olmuştur diyebiliriz.

$\hat{\beta}_r$ nin yansızlığını $H_0 : R\beta = r$ hipotezini test ederek inceleyebiliriz. H_0 hipotezi doğru ise $\hat{\beta}_r$ yansızdır. H_0 hipotezi için test istatistiği ise (1.1) modeli için $e \sim N(0, \sigma^2 I)$ varsayımı altında

$$W = \frac{n-p}{J} \frac{(R\hat{\beta} - r)' [R(X'X)^{-1} R']^{-1} (R\hat{\beta} - r)}{y'(I - X(X'X)^{-1} X')y}$$

dır.

2.5. Modified Ridge Tahmin Edici

Swindel (1976) ridge tahmin edici ile ön (prior) bilgiyi birleştirerek modified ridge tahmin ediciyi önermiştir:

$$\hat{\beta}(k, b) = (X'X + kI)^{-1} (X'y + kb), \quad k \geq 0 \quad (2.13)$$

burada b , p boyutlu stokastik olmayan bir vektör olup β üzerindeki ön bilgiyi temsil edecek şekilde seçilmelidir. $\hat{\beta}(k, b)$ nin bazı özelliklerini ise şu şekilde vermiştir:

i) $\hat{\beta}(k, b)$, aynı rezidü kareler toplamına sahip β nin tahmin ediciler sınıfı içerisinde b ye en yakın olandır. Yani $(y - X\tilde{\beta})'(y - X\tilde{\beta}) = c$ olacak şekilde $(\tilde{\beta} - b)'(\tilde{\beta} - b)$ yi minimum yapmak isteriz. $1/k$ Lagrange çarpanı olmak üzere

$$\Psi_1 = (\tilde{\beta} - b)'(\tilde{\beta} - b) + \frac{1}{k} [(y - X\tilde{\beta})'(y - X\tilde{\beta}) - c]$$

nın $\tilde{\beta}$ ya göre türevini alıp sıfıra eşitlersek (2.13) ile verilen $\hat{\beta}(k, b)$ yi elde ederiz.

ii) $\hat{\beta}(k, b)$, b ye eşit uzaklıktaki β nin tahmin ediciler sınıfı içerisinde en küçük rezidü kareler toplamına sahiptir. Yani $\hat{\beta}(k, b)$, $(\tilde{\beta} - b)'(\tilde{\beta} - b) = c$ olacak şekilde $(y - X\tilde{\beta})'(y - X\tilde{\beta})$ yi minimum yapan $\tilde{\beta}$ tahmin edicisidir.

$$\Psi_2 = (y - X\tilde{\beta})'(y - X\tilde{\beta}) + k[(\tilde{\beta} - b)'(\tilde{\beta} - b) - c]$$

nın $\tilde{\beta}$ ya göre türevini alıp sıfıra eşitlersek (2.13) ile verilen $\hat{\beta}(k, b)$ yi elde ederiz.

iii) $\hat{\beta}(0, b) = \hat{\beta}$ dır ve $\lim_{k \rightarrow \infty} \hat{\beta}(k, b) = b$ olduğundan k arttıkça $\hat{\beta}(k, b)$, $\hat{\beta}$ dan b ye bir eğri çizer. Bu eğri rezidü kareler toplamında küçük bir atış olduğunda $\hat{\beta}(k, b)$ nin b ye olan uzaklığında büyük bir azalma olacak şekilde tek türlü belirlenir.

Swindel (1976) $\hat{\beta}(k, b)$ nin EKK dan daha küçük HKO ya sahip olacak şekilde bir k nin varlığını göstermiştir. Ayrıca b , $\hat{\beta}$ dan bağımsız bir rasgele vektör iken de böyle bir k nin var olduğunu ispatlamıştır. b , $\hat{\beta}$ dan bağımsız değilse böyle bir k nin varlığı ise garanti değildir.

2.6. Yansız Ridge Tahmin Edici

Crouse, Jin ve Hanumara (1995) EKK ile ön bilginin konveks kombinasyonu olarak yansız ridge tahmin ediciyi $k > 0$ için J , $J \sim N(\beta, \frac{\sigma^2}{k} I)$ dağılımına sahip bir rasgele vektör olmak üzere

$$\hat{\beta}(kI, J) = (X'X + kI)^{-1}(X'y + kJ) \quad (2.14)$$

olarak önermiştir. k yanlılık parametresinin bilinmemesi durumunda (2.14) pratikte kullanışsızdır. Crouse, Jin ve Hanumara (1995) k yanlılık parametresinin bir tahminini ise $\hat{\sigma}^2$, (1.14) ile verilen σ^2 nin yansız tahmin edicisi olmak üzere

$$\hat{k}_{cjh} = \begin{cases} \frac{p\hat{\sigma}^2}{(\hat{\beta} - J)'(\hat{\beta} - J) - \hat{\sigma}^2 \text{tr}(X'X)^{-1}}, & (\hat{\beta} - J)'(\hat{\beta} - J) - \hat{\sigma}^2 \text{tr}(X'X)^{-1} > 0 \\ \frac{p\hat{\sigma}^2}{(\hat{\beta} - J)'(\hat{\beta} - J)}, & \text{aksi halde} \end{cases} \quad (2.15)$$

olarak vermişlerdir ve $\hat{k}_{cjh}$ nin Hoerl, Kennard ve Baldwin'in (1975) önerdiği

$$\hat{k}_{HKB} = \frac{p\hat{\sigma}^2}{\hat{\beta}'\hat{\beta}} \text{ nin genel hali olduğunu belirtmişlerdir. Hoerl ve Kennard'ın (1981)}$$

ürettiği veriyi kullanarak ve simülasyon yaparak yansız ridge tahmin edicinin ridge ve EKK tahmin edicilerden daha küçük HKO ya sahip olduğunu göstermişler.

2.7. Kısıtlı Ridge Tahmin Edici

Sarkar (1992) ridge tahmin edici ile kısıtlı EKK tahmin edicilerini birleştirerek

$$\beta^*(k) = W\hat{\beta}_r, \quad W = (I_p + k(X'X)^{-1})^{-1}, \quad k \geq 0 \quad (2.16)$$

ile verilen kısıtlı ridge tahmin ediciyi tanımlamıştır. Burada $\hat{\beta}_r$ (2.11) ile verilen kısıtlı EKK tahmin edicidir ve (2.8) ile verilen kısıtlamalar geçerlidir. $\beta^*(k)$ nin bazı özellikleri ise $\beta^*(0) = \hat{\beta}_r$, $\lim_{k \rightarrow \infty} \beta^*(k) = 0$ ve

$$E(\beta^*(k)) = W\beta + W(X'X)^{-1}R'[R(X'X)^{-1}R']^{-1}(r - R\beta) \quad (2.17)$$

şeklindedir. (2.17) dikkate alınırsa $k=0$ ve $R\beta=r$ olmadığı sürece kısıtlı ridge tahmin edici yanlı bir tahmin edicidir. Sarkar (1992) kısıtlamalar hem tam olarak sağlanıyor iken hem de sağlanmıyorken HKO anlamında kısıtlı ridge tahmin edicinin ridge ve kısıtlı EKK dan iyi olması için gerekli koşulları vermiştir.

Gross (2003) $\beta^*(k)$ nin y nin her sonucu için $R\beta^*(k)=r$ kısıtlamasını sağlamadığını ve bu nedenle $R\beta=r$ kısıtlamasına göre bir kısıtlı ridge tahmin edici olmadığını belirtmiştir. Dolayısıyla $\beta^*(k)$, $R\beta=r$ eşitliği ile üretilen alt uzayın dışında değerler alabilir. Bu düşünceler Gross'un (2003) β için yeni bir ridge-tipi tahmin edici vermesine neden olmuştur. Bu tahmin ediciyi $S_k = X'X + kI_p$, $\beta_0 = R'(RR')^{-1}r$ ve $\hat{\beta}(k, \beta_0) = S_k^{-1}(X'y + kR'(RR')^{-1}r)$ olmak üzere

$$\hat{\beta}_r(k) = \hat{\beta}(k, \beta_0) - S_k^{-1}R'[RS_k^{-1}R']^{-1}(R\hat{\beta}(k, \beta_0) - r), \quad k \geq 0 \quad (2.18)$$

olarak vermiş ve kısıtlı ridge tahmin edici olarak adlandırmıştır. Burada $\hat{\beta}(k, \beta_0)$, (2.13) ile verilen modified ridge tahmin edicide $b = \beta_0$ alınarak elde edilmiştir. $\lim_{k \rightarrow \infty} \hat{\beta}_r(k) = \beta_0$ ve $\lim_{k \rightarrow 0} \hat{\beta}_r(k) = \hat{\beta}_r$ olduğundan k arttıkça $\hat{\beta}_r(k)$, $\hat{\beta}_r$ den β_0 a bir eğri çizmektedir. Gross (2003) önerdiği kısıtlı ridge tahmin edicinin kısıtlı EKK tahmin ediciden matris HKO anlamında üstün olması için gerek ve yeter koşulları vermiştir. Ayrıca $Q = I_p - R'(RR')^{-1}R$ olmak üzere $QX'XQ = Q$ iken $\hat{\beta}_r(k)$ yi

minimum yapan k değerini $k = \frac{(p-J)\sigma^2}{\beta' Q \beta}$ olarak bulmuştur. $QX'XQ = Q$

sağlanmasa bile k nın bir tahmin edicisini $\hat{\sigma}_r^2 = \frac{(y - X\hat{\beta}_r)'(y - X\hat{\beta}_r)}{n - (p - J)}$ olmak üzere

$$\hat{k} = \frac{(p - J)\hat{\sigma}_r^2}{\hat{\beta}_r' Q \hat{\beta}_r} \quad (2.19)$$

olarak önermiştir. Herhangi bir kısıtlama yoksa (2.19) Hoerl, Kennard ve Baldwin'in

(1975) önerdiği $\hat{k}_{HKB} = \frac{p\hat{\sigma}^2}{\hat{\beta}'\hat{\beta}}$ ye eşit olmaktadır. Gross (2003), (2.19) da $\hat{\beta}_r' Q \hat{\beta}_r$

yerine $\hat{\beta}_r(k)'Q\hat{\beta}_r(k)$ alarak $\hat{k}$ nın iteratif yöntemle belirlenmesini de önermiştir.

2.8. $r - k$ Sınıf Tahmin Edici

Lineer regresyon modelinde çoklu iç ilişki problemi ve onun istatistiksel sonuçlar üzerindeki etkisi çok iyi bilinmektedir. Bu problemi ortadan kaldırmak için farklı çalışmalar önerilmiştir. Baye ve Parker (1984) ridge ve temel bileşenler tahmin edicilerini birleştirerek $r - k$ sınıf tahmin ediciyi önermiştir. $\lambda_i, i = 1, \dots, p$ $X'X$ matrisinin özdeğerleri olmak üzere $\Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_p)$ ve $T = [t_1, t_2, \dots, t_p]$ $T'X'XT = \Lambda$ olacak şekilde ortogonal bir matris olsun. Ayrıca $T_r = [t_1, t_2, \dots, t_r]$, $r \leq p$ ve $T_{p-r} = [t_{r+1}, t_{r+2}, \dots, t_p]$ sırasıyla $T_r'X'XT_r = \Lambda_r = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_r)$ ve $T_{p-r}'X'XT_{p-r} = \Lambda_{p-r} = \text{diag}(\lambda_{r+1}, \lambda_{r+2}, \dots, \lambda_p)$ olacak şekilde matrisler olsun. Bu durumda $r - k$ sınıf tahmin edici

$$\hat{\beta}_r(k) = T_r(T_r'X'XT_r + kI_r)^{-1}T_r'X'y, \quad k \geq 0 \quad (2.20)$$

olarak tanımlanmıştır. Baye ve Parker (1984) bu tahmin edicinin EKK, temel bileşenler ve ridge tahmin edicilerin genel hali olduğunu belirtmişlerdir:

$$\hat{\beta}_p(0) = \hat{\beta} = (X'X)^{-1}X'y \quad \text{EKK tahmin edicisi,}$$

$$\hat{\beta}_r(0) = \hat{\beta}_r = T_r(T_r'X'XT_r)^{-1}T_r'X'y \quad \text{temel bileşenler tahmin edicisi,}$$

$$\hat{\beta}_p(k) = \hat{\beta}(k) = (X'X + kI_p)^{-1}X'y \quad \text{ridge tahmin edici.}$$

Baye ve Parker (1984) skaler HKO anlamında $r - k$ sınıf tahmin edicinin temel bileşenler tahmin ediciden iyi olduğunu göstermiştir. Nomura ve Ohkubo (1985) $r - k$ sınıf tahmin ediciyi EKK ve ridge tahmin ediciler ile skaler HKO anlamında karşılaştırmıştır. Sarkar (1996) ise $r - k$ sınıf tahmin ediciyi EKK, ridge ve temel bileşenler tahmin ediciler ile matris HKO anlamında karşılaştırmıştır.

2.9. Liu Tahmin Edici

Çoklu iç ilişki olması durumunda EKK'nin olumsuz yanları bilinmektedir. Bu olumsuzlukları gidermek için pek çok tahmin edici önerilmiştir. Ridge ve Stein ($\hat{\beta}_s = c\hat{\beta}$, $0 < c < 1$) tahmin ediciler bunlardan ikisidir. Fakat her iki tahmin edicinin de olumsuz yanları vardır.

$\hat{\beta}(k)$, k nın karmaşık bir fonksiyonudur. Bu nedenle k nın seçimi için önerilen metotlarda karmaşık denklemlerle karşılaşabiliriz. Ridge tahmin metodu $X'X$ matrisinin aksine $X'X + kI$, $k \geq 0$ matrisine bağlı olduğundan EKK'deki zorluklar önlenmiş olunur. $X'X + kI$ nın koşul sayısı k nın azalan bir fonksiyonu olduğundan yeteri kadar büyük k için $X'X + kI$ nın koşul sayısı küçük seviyeye düşürülebilir. Fakat uygulamada k nın küçük olması önerilir. Bu nedenle $X'X + kI$ nın koşul sayısını küçültebilecek bir k bulunamayabilir. Bu durumda, ridge regresyon hala durağan değildir. k yanlılık parametresinin farklı seçim yöntemleri vardır ve seçimi analizciye bağlıdır. k nın nasıl seçileceği ile ilgili bir uzlaşmanın olmamasından dolayı çoklu iç ilişki olması durumunda yeni metotlar araştırılmıştır.

Stein tahmin edicinin avantajı c nin bir lineer fonksiyonu olmasıdır. Fakat $\hat{\beta}_s$ nin her elemanının büzülmesi aynıdır. Bu da uygulamada iyi sonuç vermez.

İki tahmin edicinin kombinasyonunun bu iki tahmin edicinin avantajlarını birleştireceği düşüncesi Liu'yu (1993) yeni bir tahmin edici önermeye sevk etmiştir.

Çoklu iç ilişki problemini ortadan kaldırmak için Liu (1993), Stein (1956) tahmin edici ile ridge tahmin ediciyi birleştirmiş ve

$$\hat{\beta}_d = (X'X + I)^{-1}(X'y + d\hat{\beta}), \quad 0 < d < 1 \quad (2.21)$$

tahmin edicisini önermiştir. Bu tahmin edici Akdeniz ve Kaçıranlar (1995) ve Gruber (1998) tarafından Liu tahmin edici olarak adlandırılmış. $\hat{\beta}_d$ nın $\hat{\beta}(k)$ üzerindeki avantajı d nin bir lineer fonksiyonu olmasıdır. Dolayısıyla d nin seçimi k nın seçiminden daha kolaydır.

(2.1) ile verilen ridge tahmin edici (1.1) modeline $0 = k^{1/2}\beta + e'$ eşitliğinin ilave edilip EKK yönteminin kullanılması ile elde edilebilir. k büyüdükçe $k^{1/2}$ ile 0 arasındaki uzaklık artar, böylece ridge tahmin edicinin yanlılığı da artar. Bu problemi ortadan kaldırmak için (1.1) modeline $d\hat{\beta} = \beta + e'$ eşitliğinin ilave edilip EKK yönteminin kullanılması ile (2.21) ile verilen Liu tahmin edici elde edilir.

Liu tahmin edicinin bazı özellikleri ise şu şekildedir:

1. $d=1$ iken $\hat{\beta}_d = \hat{\beta}$ dir. Ayrıca Liu tahmin edici EKK tahmin edicinin bir lineer dönüşümü olarak da yazılabilmektedir: $\hat{\beta}_d = (X'X + I)^{-1}(X'X + dI)\hat{\beta}$.
2. $\|\hat{\beta}_d\| < \|\hat{\beta}\|$ dir.
3. $E(\hat{\beta}_d) = (X'X + I)^{-1}(X'X + dI)\beta$ olup yanlı bir tahmin edicidir.
4. $\hat{\beta}_d$ nın HKO sı $\hat{\beta}$ dan daha iyi olacak şekilde bir $0 < d < 1$ vardır.
5. (1.4) modeli için

$$SMSE(\hat{\beta}_d, \beta) = \sigma^2 \sum_{i=1}^p \frac{(\lambda_i + d)^2}{\lambda_i(\lambda_i + 1)} + (d-1)^2 \sum_{i=1}^p \frac{\alpha_i^2}{(\lambda_i + 1)^2}$$

dır ve Liu (1993) $SMSE(\hat{\beta}_d, \beta)$ yi minimum yapan d yanlılık parametresinin tahminini

$$\hat{d}_{mm} = 1 - \hat{\sigma}^2 \left(\sum_{i=1}^p \frac{1}{\lambda_i(\lambda_i + 1)} \right) / \left(\sum_{i=1}^p \frac{\hat{\alpha}_i^2}{(\lambda_i + 1)^2} \right) \quad (2.22)$$

olarak vermiştir.

2.10. $r - d$ Sınıf Tahmin Edici

$r - k$ sınıf tahmin ediciye alternatif olarak, Kaçıranlar ve Sakallıoğlu (2001) Liu ve temel bileşenler tahmin edicileri birleştirerek $r - d$ sınıf tahmin ediciyi önermişlerdir:

$$\hat{\beta}_r(d) = T_r(T_r'X'XT_r + I_r)^{-1}(T_r'X'y + dT_r'\hat{\beta}_r), \quad 0 < d < 1. \quad (2.23)$$

$r - d$ sınıf tahmin edici de EKK, temel bileşenler ve Liu tahmin edicilerin genel halidir:

$$\hat{\beta}_p(1) = \hat{\beta} = (X'X)^{-1}X'y \quad \text{EKK tahmin edicisi,}$$

$$\hat{\beta}_r(1) = \hat{\beta}_r = T_r(T_r'X'XT_r)^{-1}T_r'X'y \quad \text{temel bileşenler tahmin edicisi,}$$

$$\hat{\beta}_p(d) = \hat{\beta}_d = (X'X + dI_p)^{-1}(X'X + dI_p)\hat{\beta} \quad \text{Liu tahmin edicidir.}$$

Kaçıranlar ve Sakallıoğlu (2001) bu yeni tahmin ediciyi EKK, temel bileşenler ve Liu tahmin edicilerle skaler HKO anlamında karşılaştırmışlardır.

2.11. Liu-Tipi Tahmin Edici

Koşul sayısı çoklu iç ilişkiyi belirleme yöntemlerinden birisidir. $X'X$ kötü koşullu iken $X'X$ in koşul sayısı büyük olacaktır. Ridge regresyonda $X'X$ in köşegen elemanlarına pozitif küçük k sabitleri eklenerek $X'X$ in koşul sayısı küçültülmeye çalışılır. Fakat $X'X + kI$ nın koşul sayısı k nın azalan bir fonksiyonudur, dolayısıyla uygulamada $X'X + kI$ nın koşul sayısını küçültecek şekilde yeterince büyük bir k seçilmelidir. Fakat uygulamada k nın küçük olması

önerilir, bu durumda ise $X'X + kI$ kötü koşullu olup ridge regresyon durağan değildir.

$0 = k^{1/2} \beta + e'$ eşitliğinin kullanılması ile ridge tahmin edicinin elde edildiğini Bölüm 2.9 da belirtmiştik. Liu (2003), kötü koşulluluk problemini ortadan kaldırmak için (1.1) modeline $(-d/k^{1/2})\hat{\beta} = k^{1/2} \beta + e'$ eşitliğini ilave edip EKK tahmin yöntemini kullanarak $\tilde{\beta}$, β nın herhangi bir tahmin edicisi olmak üzere

$$\hat{\beta}_{k,d} = (X'X + kI)^{-1}(X'y - d\tilde{\beta}), \quad k > 0, \quad -\infty < d < \infty \quad (2.24)$$

tahmin edicisini önermiştir. k yanlılık parametresi $X'X + kI$ nın koşul sayısını uygun bir seviyeye düşürmek için kullanılır, d yanlılık parametresi ise tahmin edicinin istatistiksel özelliklerini geliştirmek için kullanılır.

Liu (2003) $\tilde{\beta}$ yerine EKK ve ridge tahmin edicilerini almış ve her iki durumda da $\hat{\beta}_{k,d}$ nın ridge tahmin ediciden daha iyi HKO ya sahip olacak şekilde k ve d tahminlerini vermiştir.

2. Bölümde ele aldığımız tahmin edicileri homojen lineer tahmin ediciler ve heterojen lineer tahmin ediciler olarak iki farklı sınıfa koyabiliriz. C_1 , $p \times n$ tipinde bir matris olmak üzere $\tilde{\beta} = C_1 y$ formundaki tahmin edicilere homojen lineer tahmin ediciler denir. Ridge, temel bileşenler, $r - k$ sınıf, Liu ve $r - d$ sınıf tahmin ediciler bu sınıfa girmektedir. C_1 , $p \times n$ tipinde bir matris ve c_2 , $p \times 1$ tipinde bir vektör olmak üzere $\tilde{\beta} = C_1 y + c_2$ formundaki tahmin edicilere heterojen lineer tahmin ediciler denir. Kısıtlı EKK, modified ridge, yansız ridge ve kısıtlı ridge tahmin ediciler ise bu sınıfa girmektedir. (2.24) ile verilen Liu-tipi tahmin edici ise $\tilde{\beta}$ nın durumuna göre her iki sınıfa da girebilir.

3. BAZI YANLI TAHMİN EDİCİLERİN KARŞILAŞTIRILMASI

İstatistikçiler EKK ile yanlı tahmin edici veya iki yanlı tahmin edicinin seçimi ile ilgilenirler. Gauss ise tahmin edicilerin seçimi için en iyi kriteri HKO olarak belirlemiştir. β nın herhangi bir $\tilde{\beta}$ tahmin edicisi için HKO matrisi (1.15) ile verilmiştir. $Var(\tilde{\beta})$ ve $Bias(\tilde{\beta})$ terimleri cinsinden ise (1.16) ile ifade edilmiştir.

$\tilde{\beta}_1$ ve $\tilde{\beta}_2$, β nın herhangi iki tahmin edicisi olsun. $MMSE(\tilde{\beta}_1, \beta) - MMSE(\tilde{\beta}_2, \beta)$ farkı negatif olmayan tanımlı bir matris ise $\tilde{\beta}_2$ tahmin edicisi $\tilde{\beta}_1$ tahmin edicisinden matris HKO anlamında iyidir denir.

Bir tahmin edicinin iyiliği ile ilgili bir diğer kriter ise genelleştirilmiş HKO dır. B negatif olmayan tanımlı bir matris olmak üzere $\tilde{\beta}$ tahmin edicisinin genelleştirilmiş HKO su

$$GMSE(\tilde{\beta}, \beta) = E(\tilde{\beta} - \beta)' B (\tilde{\beta} - \beta)$$

olarak tanımlanır. $B = I$ olması durumunda $GMSE(\tilde{\beta}, \beta)$, $SMSE(\tilde{\beta}, \beta)$ ye dönüşür:

$$SMSE(\tilde{\beta}, \beta) = E(\tilde{\beta} - \beta)' (\tilde{\beta} - \beta). \quad (3.1)$$

İz operatörünün uygulanması ile $SMSE(\tilde{\beta}, \beta) = tr(MMSE(\tilde{\beta}, \beta))$ elde ederiz (Bölüm 1). $MMSE(\tilde{\beta}_1, \beta) - MMSE(\tilde{\beta}_2, \beta)$ negatif olmayan tanımlı bir matris ise $SMSE(\tilde{\beta}_1, \beta) \geq SMSE(\tilde{\beta}_2, \beta)$ dır. Fakat tersi genellikle doğru değildir. Ayrıca Theobald'a (1974) göre $GMSE(\tilde{\beta}_1, \beta) > GMSE(\tilde{\beta}_2, \beta)$ olması için gerek ve yeter koşul $MMSE(\tilde{\beta}_1, \beta) - MMSE(\tilde{\beta}_2, \beta)$ matrisinin negatif olmayan tanımlı bir matris olmasıdır. Bu nedenle, matris HKO kriteri skaler HKO ve genelleştirilmiş HKO kriterlerinden daha iyi bir kriterdir (Toutenburg, 1982; Rao ve Toutenburg, 1995).

Ayrıca, skaler HKO nun temel kriter olarak seçilmesi demek matris HKO nun köşegen dışı elemanlarının göz ardı edilmesi demek anlamına gelir. Sonuç olarak, iki tahmin edicinin matris HKO anlamında karşılaştırılması daha uygun bir kriterdir. Fakat matris HKO nun uygulanması zordur bu nedenle bazı durumlarda diğer iki kriter uygulanır.

Bölüm 2 de ele aldığımız tahmin ediciler arasındaki bazı karşılaştırmalar pek çok araştırmacı tarafından incelenmiştir. Bunlardan bazılarını burada vermemiz yararlı olacaktır. Hoerl ve Kennard (1970) ridge ile EKK tahmin edicilerini skaler HKO anlamında karşılaştırmış, ridge tahmin edicinin EKK dan daha iyi olacak şekilde bir $k > 0$ ın var olduğunu görmüştür. Matris HKO kriterine göre ridge tahmin edicinin EKK dan iyi olacak şekildeki gerek ve yeter koşulun $k \leq 2\sigma^2/\beta' \beta$ olduğu pek çok kaynaktan görülebilir (örneğin, Rao ve Toutenburg, 1995; Theobald, 1974). Swindel (1976) modified ridge ile EKK tahmin edicilerini matris HKO kriterine göre, Pliskin (1987) modified ridge ile ridge tahmin edicilerini matris HKO kriterine göre, Baye ve Parker (1984) $r - k$ sınıf ile temel bileşenler tahmin edicilerini skaler HKO kriterine göre, Nomura ve Ohkubo (1985) $r - k$ sınıf ile ridge ve EKK tahmin edicilerini skaler HKO kriterine göre, Sarkar (1996) $r - k$ sınıf ile EKK, ridge ve temel bileşenleri matris HKO kriterine göre, Sarkar (1992) kısıtlı ridge ile kısıtlı EKK ve ridge tahmin edicileri skaler HKO kriterine göre karşılaştırmıştır. Yapılan bu çalışmalara paralel olarak bazı yeni karşılaştırmalar yapılacaktır.

3.1. $r - d$ Sınıf Tahmin Edicinin Matris HKO Kriterine Göre Bazı Tahmin

Ediciler Üzerindeki Üstünlüğü

Kaçıranlar ve Sakallıoğlu (2001) $r - d$ sınıf tahmin ediciyi EKK, temel bileşenler ve Liu tahmin edicilerle skaler HKO anlamında karşılaştırmışlardır. Fakat, bu karşılaştırmaları matris HKO kriterine göre yapmamışlardır. Bu nedenle burada $r - d$ sınıf tahmin edici EKK, Liu ve temel bileşenler ile matris HKO kriterine göre karşılaştırılacaktır.

(1.1) modeli ve (2.23) ile verilen $r - d$ sınıf tahmin ediciyi ele alalım. Bu durumda $S_r(d) = \Lambda_r + dI_r$ olmak üzere

$$\text{Var}(\hat{\beta}_r(d)) = \sigma^2 T_r S_r(1)^{-1} \Lambda_r^{-1} S_r(d) \Lambda_r S_r(d) \Lambda_r^{-1} S_r(1)^{-1} T_r'$$

dır. $E(\hat{\beta}_r(d)) = T_r S_r(1)^{-1} \Lambda_r^{-1} S_r(d) \Lambda_r T_r' \beta$ ve

$$T_r S_r(1)^{-1} \Lambda_r T_r' - I_p = T_r' T_r - T_r S_r(1)^{-1} T_r' - I_p = -(T_{p-r} T_{p-r}' + T_r S_r(1)^{-1} T_r')$$

eşitliklerini kullanırsak

$$\text{Bias}(\hat{\beta}_r(d)) = [(d-1)T_r S_r(1)^{-1} T_r' - T_{p-r} T_{p-r}'] \beta$$

elde ederiz. Böylece $\hat{\beta}_r(d)$ için matris HKO

$$\begin{aligned} \text{MMSE}(\hat{\beta}_r(d), \beta) &= \sigma^2 T_r S_r(1)^{-1} \Lambda_r^{-1} S_r(d) \Lambda_r S_r(d) \Lambda_r^{-1} S_r(1)^{-1} T_r' \\ &\quad + [(d-1)T_r S_r(1)^{-1} T_r' - T_{p-r} T_{p-r}'] \beta \\ &\quad \times \beta' [(d-1)T_r S_r(1)^{-1} T_r' - T_{p-r} T_{p-r}'] \end{aligned} \quad (3.2)$$

şeklinde olur.

3.1.1. $r - d$ Sınıf Tahmin Edici ile EKK Tahmin Edicinin Matris HKO Kriterine Göre Karşılaştırılması

Teorem 3.1 : $r - d$ sınıf tahmin edicinin EKK tahmin ediciden üstün olması için gerek ve yeter koşul

$$\beta' T_r (1-d)^2 [2(1-d)I_r + (1-d^2)\Lambda_r^{-1}]^{-1} T_r' \beta + \beta' T_{p-r} \Lambda_{p-r} T_{p-r}' \beta \leq \sigma^2$$

olmasıdır.

İspat : EKK için $S = X'X$ olmak üzere $MMSE(\hat{\beta}, \beta) = \sigma^2 S^{-1}$ dır. Eğer

$$\Lambda = \begin{pmatrix} \Lambda_r & 0 \\ 0' & \Lambda_{p-r} \end{pmatrix} \text{ ve } T = (T_r, T_{p-r}) \text{ yazarsak}$$

$$S^{-1} = T\Lambda^{-1}T' = T_r\Lambda_r^{-1}T_r' + T_{p-r}\Lambda_{p-r}^{-1}T_{p-r}'. \quad (3.3)$$

elde ederiz. (3.3) ü kullanırsak $MMSE(\hat{\beta}, \beta) - MMSE(\hat{\beta}_r(d), \beta)$ farkını

$$\begin{aligned} MMSE(\hat{\beta}, \beta) - MMSE(\hat{\beta}_r(d), \beta) &= \sigma^2 T_r S_r(1)^{-1} [2(1-d)I_r + (1-d^2)\Lambda_r^{-1}] \\ &\quad \times S_r(1)^{-1} T_r' \\ &\quad + T_{p-r} [\sigma^2 \Lambda_{p-r}^{-1} - T_{p-r}' \beta \beta' T_{p-r}] T_{p-r}' \\ &\quad - (d-1)^2 T_r S_r(1)^{-1} T_r' \beta \beta' T_r S_r(1)^{-1} T_r' \\ &\quad + (d-1) T_r S_r(1)^{-1} T_r' \beta \beta' T_{p-r} T_{p-r}' \\ &\quad + (d-1) T_{p-r} T_{p-r}' \beta \beta' T_r S_r(1)^{-1} T_r' \end{aligned} \quad (3.4)$$

olarak yazabiliriz. $S^*(d)$ ve $\Lambda^*(d)^{-1}$ yı

$$S^*(d) = \begin{pmatrix} \frac{S_r(1)}{(1-d)} & 0 \\ 0' & I_{p-r} \end{pmatrix}$$

ve

$$\Lambda^*(d)^{-1} = \begin{pmatrix} \frac{1}{(1-d)^2} [2(1-d)I_r + (1-d^2)\Lambda_r^{-1}] & 0 \\ 0' & \Lambda_{p-r}^{-1} \end{pmatrix}$$

olarak tanımlarsak (3.4)

$$MMSE(\hat{\beta}, \beta) - MMSE(\hat{\beta}_r(d), \beta) = TS^*(d)^{-1} [\sigma^2 \Lambda^*(d)^{-1} - T' \beta \beta' T] S^*(d)^{-1} T'$$

ifadesine eşit olur. $MMSE(\hat{\beta}, \beta) - MMSE(\hat{\beta}_r(d), \beta)$ nin negatif olmayan tanımlı olması için gerek ve yeter koşul $\sigma^2 \Lambda^*(d)^{-1} - T' \beta \beta' T$ nin negatif olmayan tanımlı bir matris olmasıdır. $\sigma^2 \Lambda^*(d)^{-1} - T' \beta \beta' T$ negatif tanımlı değildir ancak ve ancak $\beta' T \Lambda^*(d) T' \beta \leq \sigma^2$ ise (Rao ve Toutenburg, 1995 syf. 308). $\Lambda^*(d)$ ifadesini yerine yazar ve gerekli kısaltmaları yaparsak ispatı tamamlamış oluruz.

3.1.2. $r - d$ Sınıf Tahmin Edicinin Temel Bileşenler Tahmin Edicisi ile Matris HKO Kriterine Göre Karşılaştırılması

$MMSE(\hat{\beta}_r, \beta) - MMSE(\hat{\beta}_r(d), \beta)$ farkını incelemek için Trenkler ve Toutenburg'deki (1990), Rao ve Toutenburg'deki (1995) gibi farklı teoremler kullanılabilir. Biz burada Baksalary ve Trenkler (1991) tarafından verilen ve Sarkar (1996) tarafından önerilen aşağıdaki teoremi kullanacağız.

Teorem : $\zeta_{n \times p}$ $n \times p$ tipinde kompleks matrisler kümesi ve H_n Hermityen matrisleri içeren $\zeta_{n \times n}$ nin alt kümesi olsun. Ayrıca, $L \in \zeta_{n \times p}$ verilsin, L^* , $\Re(L)$, $\zeta(L)$ gösterimleri sırasıyla L nin eşlenik transpozu, kolon uzayı ve genelleştirilmiş terslerin kümesi olsun. Şimdi, $A \in H_n$, a_1 ve $a_2 \in \zeta_{n \times 1}$ lineer bağımsız, $A^- \in \zeta(A)$ için $s = [a_1^* (I_n - AA^-)^* (I_n - AA^-) a_2] / [a_1^* (I_n - AA^-)^* (I_n - AA^-) a_1]$ ve $f_{ij} = a_i^* A^- a_j$, $i, j = 1, 2$ olsun ve $a_1 \notin \Re(A)$ sağlansın. Bu durumda, $A + a_1 a_1^* - a_2 a_2^*$ negatif tanımlı değildir ancak ve ancak aşağıdaki koşullardan biri sağlanırsa:

a) A negatif tanımlı olmasın, $a_1 \in \Re(A)$, $a_2 \in \Re(A)$ ve

$$(f_{11} + 1)(f_{22} - 1) \leq |f_{12}|^2;$$

b) A negatif tanımlı olmasın, $a_1 \notin \Re(A)$, $a_2 \in \Re(A: a_1)$ ve

$$(a_2 - s a_1)^* A^- (a_2 - s a_1) \leq 1 - |s|^2;$$

c) $A = U\Delta U^* - \delta vv^*$, $a_1 \in \Re(A)$, $a_2 \in \Re(A)$, $v^* a_1 \neq 0$ ve

$$(f_{11} + 1) \leq 0, (f_{22} - 1) \leq 0, (f_{11} + 1)(f_{22} - 1) \geq |f_{12}|^2$$

burada $(U : v)$ altuniter matris, Δ pozitif köşegen bir matris (U varken ortaya çıkan) ve δ pozitif bir skalerdir. Üstelik, (a)-(c) koşullarının tamamı $A^- \in \zeta(A)$ seçiminden bağımsızdır.

Teorem 3.2 : $MMSE(\hat{\beta}_r, \beta) - MMSE(\hat{\beta}_r(d), \beta)$ nın negatif olmayan tanımlı bir matris olması için gerek ve yeter koşul $N(W_1)$, $W_1 = \left(\frac{1-d}{\sigma} \right) [2(1-d)I_r + (1-d^2)\Lambda_r^{-1}]^{1/2} T'_r$ nın sıfır uzayı olmak üzere $\beta \in N(W_1)$ olmasıdır.

İspat : $d=1$ iken $\hat{\beta}_r$, $\hat{\beta}_r(d)$ nin özel durumu olduğundan (3.2) de $d=1$ alarak $MMSE(\hat{\beta}_r, \beta)$ yı elde ederiz:

$$MMSE(\hat{\beta}_r, \beta) = \sigma^2 T_r \Lambda_r^{-1} T'_r + (T_r T'_r - I_p) \beta \beta' (T_r T'_r - I_p). \quad (3.5)$$

Böylece

$$\begin{aligned} MMSE(\hat{\beta}_r, \beta) - MMSE(\hat{\beta}_r(d), \beta) &= \sigma^2 T_r S_r(1)^{-1} [2(1-d)I_r + (1-d^2)\Lambda_r^{-1}] \\ &\quad \times S_r(1)^{-1} T'_r \\ &\quad + (T_r T'_r - I_p) \beta \beta' (T_r T'_r - I_p) \\ &\quad - [(d-1)T_r S_r(1)^{-1} T'_r - T_{p-r} T'_{p-r}] \beta \\ &\quad \times \beta' [(d-1)T_r S_r(1)^{-1} T'_r - T_{p-r} T'_{p-r}]. \end{aligned} \quad (3.6)$$

dır. Baksalary ve Trenkler'in (1991) teoremini uygulayabilmek için

$$B = S_r(1)^{-1} [2(1-d)I_r + (1-d^2)\Lambda_r^{-1}] S_r(1)^{-1},$$

olmak üzere $A = \sigma^2 T_r B T_r'$, $a_2 = [(d-1)T_r S_r (1)^{-1} T_r' - T_{p-r} T_{p-r}'] \beta$ ve $a_1 = (T_r T_r' - I_p) \beta$ alalım. A nin Moore-Penrose tersi $A^+ = \sigma^{-2} T_r B^{-1} T_r'$ dir ve ayrıca $AA^+ = T_r T_r'$ olup $a_1 \in \mathfrak{R}(A)$ olması için gerek ve yeter koşul $a_1 = 0$ olmasıdır. Bu nedenle Teoremin (a) ve (c) kısımlarını uygulayamayız. Fakat (b) kısmını uygulayabiliriz. s nin tanımından $s = 1$ elde ederiz. Üstelik,

$$\eta_1 = \left(\frac{d-1}{\sigma^2} \right) T_r S_r (1) [2(1-d)I_r + (1-d^2)\Lambda_r^{-1}]^{-1} T_r' \beta$$

olmak üzere $a_2 - a_1 = A\eta_1$ dir ki bu, $a_2 \in \mathfrak{R}(A: a_1)$ olması anlamına gelir. Ayrıca $(a_2 - a_1)' A^- (a_2 - a_1) = \eta_1' A \eta_1$ dir. Böylece, $r-d$ sınıf tahmin edicinin temel bileşenler tahmin ediciden daha iyi olması için gerek ve yeter koşul

$$\eta_1' A \eta_1 = \frac{(d-1)^2}{\sigma^2} \beta' T_r [2(1-d)I_r + (1-d^2)\Lambda_r^{-1}]^{-1} T_r' \beta \leq 0 \quad (3.7)$$

olmasıdır. Fakat, her zaman $(a_2 - a_1)' A^- (a_2 - a_1) \geq 0$ olduğu açıktır. $0 < d < 1$ olduğundan $W_1 = \left(\frac{1-d}{\sigma} \right) [2(1-d)I_r + (1-d^2)\Lambda_r^{-1}]^{-1/2} T_r'$ olmak üzere $Q_1 = W_1 \beta$ ise $MMSE(\hat{\beta}_r, \beta) - MMSE(\hat{\beta}_r(d), \beta)$ nin negatif olmayan tanımlı olması için gerek ve yeter koşul $Q_1 = 0$ olmasıdır. $Q_1 = W_1 \beta = 0$ olması $N(W_1)$, W_1 nin sıfır uzayı olmak üzere $\beta \in N(W_1)$ anlamına gelir.

3.1.3. $r-d$ Sınıf Tahmin Edici ile Liu Tahmin Edicinin Matris HKO Kriterine Göre Karşılaştırılması

Teorem 3.3 : $r-d$ sınıf tahmin edicinin Liu tahmin ediciden daha iyi olması için gerek ve yeter koşul $N(W_2)$, $W_2 = \sigma^{-1} \Lambda_{p-r}^{1/2} T_{p-r}'$ nin sıfır uzayı olmak üzere $\beta \in N(W_2)$ olmasıdır.

İspat : (3.2) de $r = p$ alırsak $MMSE(\hat{\beta}_d, \beta)$ yi elde ederiz:

$$MMSE(\hat{\beta}_d, \beta) = \sigma^2 TS(1)^{-1} S(d) \Lambda^{-1} S(d) S(1)^{-1} T' + (d-1)^2 TS(1)^{-1} T' \beta \beta' TS(1)^{-1} T' \quad (3.8)$$

burada $S(d) = \Lambda + dI_p$ dir. (3.2) ve (3.8) in bir sonucu olarak

$$\begin{aligned} MMSE(\hat{\beta}_d, \beta) - MMSE(\hat{\beta}_r(d), \beta) &= \sigma^2 TS(1)^{-1} S(d) \Lambda^{-1} S(d) S(1)^{-1} T' \\ &\quad - \sigma^2 T_r S_r(1)^{-1} S_r(d) \Lambda_r^{-1} S_r(d) S_r(1)^{-1} T'_r \\ &\quad + (d-1)^2 TS(1)^{-1} T' \beta \beta' TS(1)^{-1} T' \\ &\quad - [(d-1) T_r S_r(1)^{-1} T'_r - T_{p-r} T'_{p-r}] \beta \\ &\quad \times \beta' [(d-1) T_r S_r(1)^{-1} T'_r - T_{p-r} T'_{p-r}]. \end{aligned} \quad (3.9)$$

elde ederiz. Bölüm 3.1.2 de olduğu gibi Baksalary ve Trenkler'in (1991) teoremini kullanalım. $a_1 = (d-1)TS(1)^{-1}T'\beta$, $a_2 = [(d-1)T_r S_r(1)^{-1}T'_r - T_{p-r} T'_{p-r}] \beta$ ve $B = S_{p-r}(1)^{-1} S_{p-r}(d) \Lambda_{p-r}^{-1} S_{p-r}(d) S_{p-r}(1)^{-1}$ olmak üzere $A = \sigma^2 T_{p-r} B T'_{p-r}$ alırsak teoremin (b) kısmını uygulayabiliriz. A nın Moore-Penrose tersi $A^+ = \sigma^{-2} T_{p-r} B^{-1} T'_{p-r}$ olup $AA^+ = T_{p-r} T'_{p-r}$ dir. Böylece $a_1 \notin \mathfrak{R}(A)$ dir ve $s=1$ olduğundan $a_2 \in \mathfrak{R}(A: a_1)$ dir.

$$\eta_2 = \frac{-1}{\sigma^2} T_{p-r} S_{p-r}(1) S_{p-r}(d)^{-1} \Lambda_{p-r} T'_{p-r} \beta$$

olmak üzere $a_2 - a_1 = A\eta_2$ dir. Ayrıca, $(a_2 - a_1)' A^-(a_2 - a_1) = \eta_2' A\eta_2$ dir. Böylece $r-d$ sınıf tahmin edicinin Liu tahmin ediciden üstün olması için gerek ve yeter koşul

$$\eta_2' A\eta_2 = \frac{1}{\sigma^2} \beta' T_{p-r} \Lambda_{p-r} T'_{p-r} \beta \leq 0$$

olmasıdır. Fakat her zaman $\eta'_2 A \eta_2 \geq 0$ olduğundan gerek ve yeter koşul $\eta'_2 A \eta_2 = 0$ olur. $W_2 = \sigma^{-1} \Lambda_{p-r}^{1/2} T'_{p-r}$ olmak üzere $Q_2 = W_2 \beta$ alırsak $\eta'_2 A \eta_2 = Q'_2 Q_2$ elde ederiz. Böylece $MMSE(\hat{\beta}_d, \beta) - MMSE(\hat{\beta}_r(d), \beta)$ nın negatif olmayan tanımlı olması için gerek ve yeter koşul $Q_2 = 0$ olmasıdır. $Q_2 = W_2 \beta = 0$ olması $N(W_2)$, W_2 nin sıfır uzayı olmak üzere $\beta \in N(W_2)$ anlamına gelmektedir.

3.1.4. $r - d$ Sınıf Tahmin Edicinin Temel Bileşenler ve Liu Tahmin Edicilerinin Her İkisinden Matris HKO Anlamında İyi Olması İçin Kriter

Teorem 3.4 : $r - d$ sınıf tahmin edicinin temel bileşenler ve Liu tahmin edicilerinin her ikisinden iyi olması için gerek ve yeter koşul $N(W_3)$, $W_3 = \sigma^{-1} \Lambda^*(d)^{1/2} T'$ nın sıfır uzayı olmak üzere $\beta \in N(W_3)$ olmasıdır.

İspat : $W_3 = \begin{pmatrix} W_1 \\ W_2 \end{pmatrix}_{p \times p}$ tanımlarsak $W_3 = \sigma^{-1} \Lambda^*(d)^{1/2} T'$ olur. W_3 ü β ile sağdan

çarparsak $W_3 \beta = 0$ elde ederiz. Böylece ispat tamamlanır.

Bölüm 3.1.1-3.1.4 ü özetleyecek olursak $r - d$ sınıf tahmin edicinin EKK, temel bileşenler ve Liu tahmin edicilerden iyi olması için gerek ve yeter koşullar verilmiştir. Bu koşullar β , σ^2 bilinmeyen parametrelerine ve d yanlılık parametresinin seçimine bağlıdır. Bunların bilinmediğini dikkate alırsak teoremlerdeki koşullar geçerli olmayabilir. Yani bazı β , σ^2 , d değerleri için $r - d$ sınıf tahmin edici diğer tahmin edicilerden üstünken β , σ^2 , d nin diğer değerleri için daha kötü bir tahmin edici olabilir. Bu nedenle, β ve σ^2 nin yerine yansız tahmin edicilerini ve d nin yerine uygun bir tahminini kullanarak veya parametreler hakkındaki ön bilgileri kullanarak uygulanabilir sonuçlar elde edebiliriz.

3.1.5. Koşullar İçin Hipotez Testleri

Teorem 3.1-3.4 de elde edilen gerek ve yeter koşullar β ve σ^2 bilinmeyen parametrelerine bağlıdır. Bir araştırmacının bu koşullara inanması için nedenlerinin olması gerekir. Bu anlamda hipotez testleri ile ilgilenilebilir. Yani bir test prosedürü verilen bir durumda koşulların sağlanıp sağlanmadığını incelemek için iyi bir metottur.

e nın $N(0, \sigma^2 I_n)$ dağılımına sahip olması varsayımı altında (1.1) modeli için testler verilebilir:

$$H_0 : W_i \beta = 0, \quad i = 1, 2, 3. \quad (3.10)$$

Genelleştirilmiş olabilirlik oranı

$$v(y) = \frac{\max_{\omega \text{ daki } (\beta, \sigma^2)} [L(\beta, \sigma^2 : y)]}{\max_{\Omega \text{ daki } (\beta, \sigma^2)} [L(\beta, \sigma^2 : y)]}$$

ile verilir. Burada $L(\beta, \sigma^2 : y)$ olabilirlik fonksiyonu, $\Omega = \{(\beta, \sigma^2) : \beta \in E_p, \sigma^2 > 0\}$ ve $\omega = \{(\beta, \sigma^2) : \beta \in E_p, W_i \beta = 0, \sigma^2 > 0, i = 1, 2, 3\}$ dir. ω nın W_i , $i = 1, 2, 3$ nin sıfır uzayı olduğunu da görüyoruz.

W_i $q \times p$ tipinde q ranklı bir matris olsun. (3.10) u test etmek için genelleştirilmiş olabilirlik oran test istatistiği

$$W = \left(\frac{n-p}{q} \right) \frac{(W_i \hat{\beta})' [W_i (X' X)^{-1} W_i']^{-1} (W_i \hat{\beta})}{y' [I_n - X X^{-1}] y} \quad (3.11)$$

dır. W , q ve $n-p$ serbestlik dereceli merkezi olmayan F dağılımına sahiptir. Merkezi olmama parametresi ise

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

ile verilir. H_0 hipotezi altında W , $F(q, n - p)$ dağılımına sahiptir (Graybill, 1976).

(3.10) daki hipotez $W_i \beta = 0$, $i = 1, 2, 3$ ilişkisinin sağlanıp sağlanmadığını test etmek için kullanılır. Bir başka deyişle, (3.10) daki test $r - d$ sınıf tahmin edicinin temel bileşenler ve Liu tahmin edicilerinden ve her ikisinden sırasıyla $i = 1, 2, 3$ için matris HKO anlamında iyi olduğunu ifade eder. Theorem 3.2 deki gerek ve yeter koşul

$$\left(\frac{1-d}{\sigma} \right) [2(1-d)I_r + (1-d^2)\Lambda_r^{-1}]^{-1/2} T'_r \beta = 0 \quad (3.12)$$

anlamına gelen $\beta \in N(W_1)$ olarak ifade edilmiştir. $0 < d < 1$ için

$\left(\frac{1-d}{\sigma} \right) [2(1-d)I_r + (1-d^2)\Lambda_r^{-1}]^{-1/2}$ matrisi pozitif tanımlı (positive definite)

olduğundan (3.12) koşulu $T'_r \beta = 0$ a denktir. $T'_r \beta = 0$ ise $\hat{\beta}_r(d)$ nin veya $\hat{\beta}_r$ nin kullanılmasına karar vermek için bir temel teşkil eder. Benzer şekilde Teorem 3.3 de verilen gerek ve yeter koşul

$$\sigma^{-1} \Lambda_{p-r}^{1/2} T'_{p-r} \beta = 0 \quad (3.13)$$

anlamına gelen $\beta \in N(W_2)$ olarak ifade edilmiştir. $\sigma^{-1} \Lambda_{p-r}^{1/2}$ pozitif tanımlı bir

matris olduğundan (3.13) koşulu $T'_{p-r} \beta = 0$ a denktir. $T'_{p-r} \beta = 0$ ise $\hat{\beta}_r(d)$ nin

veya $\hat{\beta}_d$ nin kullanılmasına karar vermek için bir temel teşkil eder. Böylece Teorem

3.2 ve 3.3 de elde edilen gerek ve yeter koşullar sırasıyla $T'_r \beta = 0$ ve $T'_{p-r} \beta = 0$

olarak da ifade edilebilir. Bu durumda

$$H_0 : T'_r \beta = 0 \quad (3.14a)$$

ve

$$H_0 : T'_{p-r} \beta = 0 \quad (3.14b)$$

hipotez testleri (3.10) un özel hali olur. Normallik varsayımı altında $T'_r \hat{\beta}_r(d)$ $S_r(1)^{-1} \Lambda_r^{-1} S_r(d) \Lambda_r T'_r \beta$ beklenen değerli ve $\sigma^2 S_r(1)^{-1} \Lambda_r^{-1} S_r(d) \Lambda_r S_r(d) \Lambda_r^{-1} S_r(1)^{-1}$ kovaryans matrisli normal dağılıma sahiptir. Benzer şekilde $T'_{p-r} \hat{\beta}_{p-r}(d)$, $S_{p-r}(1)^{-1} \Lambda_{p-r}^{-1} S_{p-r}(d) \Lambda_{p-r} T'_{p-r} \beta$ beklenen değerli ve $\sigma^2 S_{p-r}(1)^{-1} \Lambda_{p-r}^{-1} S_{p-r}(d) \Lambda_{p-r} S_{p-r}(d) \Lambda_{p-r}^{-1} S_{p-r}(1)^{-1}$ kovaryans matrisli normal dağılıma sahiptir. Böylece $e'e = y'[I_n - XX^{-}]y$ olmak üzere (3.14a) ve (3.14b) için test istatistikleri sırasıyla

$$F_1 = \left(\frac{n-p}{r} \right) \frac{\hat{\beta}'_r(d) T'_r \Lambda_r T'_r \hat{\beta}_r(d)}{e'e} \sim F \left(r, n-p, \frac{\beta' T'_r \Lambda_r T'_r \beta}{2\sigma^2} \right)$$

ve

$$F_2 = \left(\frac{n-p}{p-r} \right) \frac{\hat{\beta}'_{p-r}(d) T'_{p-r} \Lambda_{p-r} T'_{p-r} \hat{\beta}_{p-r}(d)}{e'e} \sim F \left(p-r, n-p, \frac{\beta' T'_{p-r} \Lambda_{p-r} T'_{p-r} \beta}{2\sigma^2} \right)$$

ile verilir ki bunlar (3.11) in özel halleridir.

Bu hipotez testlerine ilave olarak Teorem 3.1 için de hipotez testi verilebilir. $r-d$ sınıf tahmin edicinin EKK dan üstün olması için gerek ve yeter koşul $\sigma^{-2} \beta' T \Lambda^*(d) T' \beta \leq 1$ e denktir. Bu nedenle, $\hat{\beta}_r(d)$ nın $\hat{\beta}$ dan matris HKO anlamında iyi olup olmadığını test etmek istersek $H_0 : \sigma^{-2} \beta' T \Lambda^*(d) T' \beta \leq 1$ hipotezinin test edebiliriz. Normallik varsayımı altında $\sigma^{-1} \Lambda^*(d)^{1/2} T' \hat{\beta} \sim N(\sigma^{-1} \Lambda^*(d) T' \beta, \Lambda^*(d) \Lambda^{-1})$ dir. Böylece $H_0 : \sigma^{-2} \beta' T \Lambda^*(d) T' \beta \leq 1$ hipotezini

$$F_3 = \left(\frac{n-p}{p} \right) \frac{\hat{\beta}' T \Lambda T' \hat{\beta}}{e' e} \sim F \left(p, n-p, \frac{\beta' T \Lambda T' \beta}{2\sigma^2} \right)$$

test istatistiği ile test edebiliriz.

Sonuç olarak, H_0 hipotezleri kabul edilirse Teorem 3.1-3.4 deki koşullar geçerli olur, yani $r-d$ sınıf tahmin edici diğer tahmin edicilerden matris HKO anlamında daha iyidir.

3.1.6. $r-d$ ve $r-k$ Sınıf Tahmin Edicilerin Matris HKO Kriterine Göre Karşılaştırılması

Amacımız, iki tahmin edicinin üstünlüğünü matris HKO anlamında karşılaştırmaktır. Baye ve Parker (1984) tarafından ileri sürülen ve (2.20) ile verilen $r-k$ sınıf tahmin ediciyi ele alalım. Bu tahmin edici

$$MMSE(\hat{\beta}_r(k), \beta) = \sigma^2 T_r S_r(k)^{-1} \Lambda_r S_r(k)^{-1} T_r' + (T_r S_r(k) \Lambda_r T_r' - I_p) \beta \beta' (T_r S_r(k) \Lambda_r T_r' - I_p)$$

matris HKO ya ve $r-d$ sınıf tahmin edici, $\hat{\beta}_r(d)$, (3.2) ile verilen matris HKO ya sahiptir. $MMSE(\hat{\beta}_r(k), \beta) - MMSE(\hat{\beta}_r(d), \beta)$ farkının negatif olmayan tanımlı bir matris olması için gerek ve yeter koşullar Baksalary ve Trenkler'in (1991) verdiği teorem ile bulunabilir.

Teorem 3.5 : $MMSE(\hat{\beta}_r(k), \beta) - MMSE(\hat{\beta}_r(d), \beta)$ nin negatif olmayan tanımlı olması için gerek ve yeter koşul $(1-k-d)\lambda_i - kd \leq 0$ ve $(1+k+d)^2 - 8kd \leq 0$ eşitsizliklerinin aynı anda gerçekleşmesidir.

İspat : $a_1 = (T_r S_r(k)^{-1} \Lambda_r T_r' - I_p) \beta$, $a_2 = [(d-1)T_r S_r(1)^{-1} T_r' - T_{p-r} T_{p-r}'] \beta$ ve

$B = S_r(k)^{-1} \Lambda_r S_r(k)^{-1} - S_r(1)^{-1} \Lambda_r^{-1} S_r(d) \Lambda_r S_r(d) \Lambda_r^{-1} S_r(1)^{-1}$ olmak üzere

$A = \sigma^2 T_r B T_r'$ alırsak $MMSE(\hat{\beta}_r(k), \beta) - MMSE(\hat{\beta}_r(d), \beta) = A + a_1 a_1' - a_2 a_2'$

yazabiliriz. Bu kısaltmalarla birlikte, $\eta_3 = \sigma^{-2} T_r B^{-1} D T_r' \beta$ ve $D = [(d-1)S_r(1)^{-1} + kS_r(k)^{-1}]$ iken teoremin (b) kısmı için koşulların uygulanabilir olduğunu görüyoruz. $(a_2 - a_1)' A^{-1} (a_2 - a_1) \leq 0$ eşitsizliği

$$\sigma^{-2} \beta' T_r D' B^{-1} D T_r' \beta \leq 0 \quad (3.15)$$

a denktir ve bu, $MMSE(\hat{\beta}_r(k), \beta) - MMSE(\hat{\beta}_r(d), \beta)$ farkının negatif olmayan tanımlı olmasını sağlar. (3.15) in sıfıra eşit ve küçük olması için gerek ve yeter koşul $B^{-1} \leq 0$ olmasıdır.

$$C = S_r(k) \Lambda_r^{-1} S_r(k) S_r(1)^{-1} \Lambda_r^{-1} S_r(d) \Lambda_r S_r(d) \Lambda_r^{-1} S_r(1)^{-1}$$

yazarsak $B^{-1} = [I_r - C]^{-1} S_r(k) \Lambda_r^{-1} S_r(k)$ olur. $S_r(k) \Lambda_r^{-1} S_r(k)$ her zaman pozitif olduğundan, $B^{-1} \leq 0$ olması için gerek ve yeter koşul C nin özdeğerlerinin (γ_i olsun) bire eşit veya büyük olmasıdır:

$$\gamma_i = \left[\frac{(\lambda_i + k)(\lambda_i + d)}{(\lambda_i + 1)\lambda_i} \right]^2 \geq 1. \quad (3.16)$$

γ_i karesel bir fonksiyon olduğundan, (3.16) iki eşitsizlik gösterir:

$$\frac{2\lambda_i^2 + (k + d + 1)\lambda_i + kd}{\lambda_i(\lambda_i + 1)} \leq 0 \quad (3.17)$$

ve

$$\frac{(1 - k - d)\lambda_i - kd}{\lambda_i(\lambda_i + 1)} \leq 0. \quad (3.18)$$

(3.17) nin paydası daima pozitifdir ve payını $f(\lambda_i)$ ile gösterelim. $f(\lambda_i)$, λ_i nin negatif olan karesel bir fonksiyonu olduğundan $f(\lambda_i)$ nin diskriminantı sıfıra eşit veya küçüktür. Bu sebeple, (3.17) ve (3.18) sırasıyla $(1 + k + d)^2 - 8kd \leq 0$ ve $(1 - k - d)\lambda_i - kd \leq 0$ a eşittir. Böylece ispat tamamlanmış olur.

Teorem 3.5 teki gerek ve yeter koşullar ridge tahmin edicideki k yanlılık parametresinin ve Liu tahmin edicideki d yanlılık parametresinin seçimine bağlıdır. Bu koşulların uygulanabilir olması için k ve d nin yerine uygun tahminleri koyulmalıdır.

3.2. Yansız Ridge Tahmin Edicinin Diğer Tahmin Edicilerle Karşılaştırılması

Çoklu iç ilişki olması durumunda EKK nin yetersiz olduğunu ve buna alternatif bazı tahmin edicilerin ileri sürüldüğünü belirtmiştik. Crouse, Jin ve Hanumara (1995) EKK ile ön bilginin konveks kombinasyonu olarak (2.14) ile verdiğimiz yansız ridge tahmin ediciyi önermişlerdir ve Hoerl ve Kennard'ın (1981) ürettiği veriyi kullanarak ve simülasyon çalışması yaparak yansız ridge tahmin ediciyi EKK ve ridge tahmin edicilerle karşılaştırmışlar, her iki durumda da yansız ridge tahmin edicinin ridge ve EKK den daha küçük hata kareler ortalamasına sahip olduğunu görmüşlerdir. Fakat, yansız ridge tahmin edici ile ridge ve EKK tahmin edicilerini teorik olarak matris HKO anlamında karşılaştırmamışlardır.

Varsayılan model ile gözlenen veri arasındaki uyumu ölçmek için çoğu kez rezidüler kullanılır. Hataların durumu model varsayımlarını gösterir ve rezidüler hataların davranışlarını gözlemlemek için yararlıdır. Bu nedenle varsayımlardan sapmaları belirlemek için rezidüler incelenir. Karmaşık olan varyans kovaryans yapılarına rağmen, rezidüler yararlı olabilecek bilgiyi içerebilir.

Burada, yansız ridge tahmin edici Farebrother (1984) ve Troskie ve ark.'nın (1994) çalışmalarından iki farklı şekilde elde edilecek. Yansız ridge tahmin edicinin bazı özellikleri incelenecek, EKK, ridge ve $r - k$ sınıf tahmin edicilerle matris HKO ve rezidü anlamında karşılaştırılması yapılacaktır. Ayrıca sonuçlarımız Hoerl ve Kennard'ın (1981) ürettiği veri kullanılarak örneklendirilecektir.

**3.2.1. Yansız Ridge Tahmin Edicinin Ridge Tahmin Edici ile Matris HKO
Kriterine Göre Karşılaştırılması**

(2.14) ile verilen $\hat{\beta}(kI, J)$, $\hat{\beta}(kI, J) = S_k^{-1}S\hat{\beta} + kS_k^{-1}J$ olarak da yazılabilir, burada $\hat{\beta}$ EKK tahmin edici, $S_k = X'X + kI$ ve $S = X'X$ dır. Crouse, Jin ve Hanumara (1995), $\hat{\beta} \sim N(\beta, \sigma^2(X'X)^{-1})$ ve $J \sim N(\beta, \frac{\sigma^2}{k}I)$ olmak üzere J nin $\hat{\beta}$ dan bağımsız iken $\hat{\beta}(kI, J)$ nin yansız bir tahmin edici olduğunu göstermiştir.

Troskie ve ark. (1994) stokastik ön bilgi olması durumunda (1.1) modelindeki β nin tahminini vermişlerdir. h $l \times 1$ tipinde rasgele bir vektör, H $l \times p$ tipinde tasarım matrisi, u $l \times 1$ tipinde δ ortalamalı ve σ^2 ile R bilinmemek üzere $\sigma^2 R$ varyanslı normal dağılıma sahip rasgele vektör olmak üzere ön bilgi

$$h = H\beta + u \quad (3.19)$$

dır. (3.19) a bağlı (1.1) in tahmini,

$$\tilde{\beta} = (X'X + H'R^{-1}H)^{-1}(X'y + H'R^{-1}h) \quad (3.20)$$

dır. Troskie ve ark. (1994) ridge, genelleştirilmiş ridge, temel bileşenler ve Sclove-Stein shrunk tahmin ediciler gibi çok iyi bilinen tahmin edicileri de (3.20) nin özel durumu olarak üretmiştir. Benzer sonuçlar Belsley, Kuh ve Welsch'den (1980) de görülebilir. Troskie ve ark. (1994) ve Belsley, Kuh ve Welsch'de (1980) verilen benzer sonuçları kullanarak yansız ridge tahmin ediciyi elde edeceğiz.

h rasgele bir vektör olsun. $\tilde{\beta}$ nin beklenen değeri ve kovaryans matrisi

$$E(\tilde{\beta}) = \beta + (X'X + H'R^{-1}H)^{-1}H'R^{-1}\delta$$

ve

$$Var(\tilde{\beta}) = \sigma^2(X'X + H'R^{-1}H)^{-1}$$

dır. $h = J$, $H = I$, $R = \frac{1}{k}I$ ve $\delta = 0$ alırsak, (3.20) yansız ridge tahmin edici olur. Farebrother (1984), (1.1) ve (3.19) modellerini h rasgele, $\delta = 0$ ve $R = \frac{1}{k}I$ iken birleştirmiş böylece ridge tahmin edicinin bir genelleştirilmesini elde etmiştir:

$$\hat{\beta}_g(k) = (X'X + kH'H)^{-1} (X'X\hat{\beta} + kH'H\hat{\beta}_R) \quad (3.21)$$

burada $\hat{\beta}$ ve $\hat{\beta}_R$ sırasıyla (1.9) ile verilen EKK ve (3.19) modeli için

$$\hat{\beta}_R = \hat{\beta} - (X'X)^{-1}H'(H(X'X)^{-1}H')^{-1}(H\hat{\beta} - h)$$

ile vereceğimiz kısıtlı EKK'dir. Farebrother (1984) ayrıca $\hat{\beta}_g(0) = \hat{\beta}$ ve $\hat{\beta}_g(\infty) = \hat{\beta}_R$ olduğunu da göstermiştir. $H'H = I$ ise ve (3.19) her zaman doğru ise, (3.21) $J = \hat{\beta}_R$ özel durumu ile bir yansız ridge tahmin edicidir. (2.14) ten yaklaşımları incelersek $\lim_{k \rightarrow 0} \hat{\beta}(kI, J) = \hat{\beta}$ ve $\lim_{k \rightarrow \infty} \hat{\beta}(kI, J) = J$ elde ederiz ki bu sonuçlar Farebrother'ın (1984) sonuçlarına denktir.

$\hat{\beta}(kI, J)$ yansız olduğundan, $MMSE(\hat{\beta}(kI, J), \beta) = Var(\hat{\beta}(kI, J))$ olup

$$MMSE(\hat{\beta}(kI, J), \beta) = M_1 = \sigma^2 S_k^{-1} S S_k^{-1} + k \sigma^2 S_k^{-2} \quad (3.22)$$

dır ve (2.1) ile verilen ridge tahmin edici için $Var(\hat{\beta}(k)) = \sigma^2 S_k^{-1} S S_k^{-1}$ ve $Bias(\hat{\beta}(k)) = -k S_k^{-1} \beta$ olduğundan

$$MMSE(\hat{\beta}(k), \beta) = M_2 = \sigma^2 S_k^{-1} S S_k^{-1} + k^2 S_k^{-1} \beta \beta' S_k^{-1} \quad (3.23)$$

olur.

$\hat{\beta}(k)$ nın $\hat{\beta}(kI, J)$ den iyi olacak koşullar ile ilgilendiğimizden $\Delta = MMSE(\hat{\beta}(kI, J), \beta) - MMSE(\hat{\beta}(k), \beta)$ farkını incelememiz gerekmektedir. Δ

negatif olmayan tanımlı bir matris iken $\hat{\beta}(k)$, $\hat{\beta}(kI, J)$ ye tercih edilir.

(3.22) ve (3.23) u göz önünde bulundurursak, $\Delta = S_k^{-1}[k\sigma^2 I - k^2 \beta \beta'] S_k^{-1}$ yazabiliriz. Pozitif yarı tanımlı matrislerin özelliklerini (Rao and Toutenburg, 1995 incelenebilir), kullanırsak aşağıdaki teoremi elde ederiz.

Teorem 3.6 : β nın ridge tahmin edicisinin matris HKO kriterine göre yansız ridge tahmin ediciden üstün olması için gerek ve yeter koşul

$$k \leq \frac{\sigma^2}{\beta' \beta} \quad (3.24)$$

olmasıdır.

(3.24) deki k , σ^2 ve β bilinmeyen parametrelerine bağlıdır. $\hat{\beta}(k)$, $\hat{\beta}(kI, J)$ den daha küçük matris HKO ya sahip olacak şekilde uygun $\sigma^2 > 0$ ve β parametreleri için bir k vardır.

$\hat{\beta}(k)$ nın $\hat{\beta}$ dan matris HKO kriterine göre iyi olması için gerek ve yeter koşul

$$k \leq \frac{2\sigma^2}{\beta' \beta} \quad (3.25)$$

olmasıdır (Rao ve Toutenburg, 1995 incelenebilir). (3.24) ve (3.25) koşulları karşılaştırıldığında ridge tahmin edici yansız ridge tahmin ediciden matris HKO anlamında üstünken, ridge tahmin edicinin matris HKO kriterine göre EKK tahmin ediciden aynı k için üstün olduğunu görüyoruz. Fakat, tersi doğru olmayabilir. Bu noktada, hangi koşullar altında yansız ridge tahmin edicinin EKK tahmin edicisine üstün olduğunu merak edebiliriz.

3.2.2. Yansız Ridge ve EKK Tahmin Edicilerinin Matris HKO Kriterine Göre Karşılaştırılması

(1.1) den, (1.9) ile verilen EKK tahmin edici için

$$MMSE(\hat{\beta}, \beta) = M_3 = \sigma^2 S^{-1}$$

olduğunu biliyoruz. (3.22) yi dikkate alırsak $k > 0$ için daima pozitif tanımlı olan

$$MMSE(\hat{\beta}, \beta) - MMSE(\hat{\beta}(kI, J), \beta) = k\sigma^2 S_k^{-1} S^{-1}$$

farkını elde ederiz. Bu sonucu aşağıdaki teoremle ifade edebiliriz.

Teorem 3.7 : Yansız ridge tahmin edici EKK tahmin ediciden matris HKO anlamında $k > 0$ için her zaman daha iyidir.

3.2.3. Yansız Ridge Tahmin Edici ile $r - k$ Sınıf Tahmin Edicinin Matris HKO Kriterine Göre Karşılaştırılması

Sarkar (1996), (2.20) ile verilen $r - k$ sınıf tahmin edici için EKK, ridge ve temel bileşenler tahmin edicilerinden matris HKO kriterine göre daha iyi olacak şekildeki gerek ve yeter koşulları vermiştir. Sarkar (1996) $r - k$ sınıf tahmin edicinin matris hata kareler ortalamasını $S_r(k) = \Lambda_r + kI_r$ olmak üzere

$$MMSE(\hat{\beta}_r(k), \beta) = M_4 = \sigma^2 T_r S_r(k)^{-1} \Lambda_r S_r(k)^{-1} T_r' + [T_{p-r} T_{p-r}' + k T_r S_r(k)^{-1} T_r'] \beta \beta' [T_{p-r} T_{p-r}' + k T_r S_r(k)^{-1} T_r']$$

olarak vermiştir. $\Lambda = \begin{pmatrix} \Lambda_r & 0 \\ 0' & \Lambda_{p-r} \end{pmatrix}$ olduğundan

$$\Lambda + kI_p = \begin{pmatrix} \Lambda_r + kI_r & 0 \\ 0' & \Lambda_{p-r} + kI_{p-r} \end{pmatrix} = \begin{pmatrix} S_r(k) & 0 \\ 0' & S_{p-r}(k) \end{pmatrix}$$

dır. Böylece, $S_k = X'X + kI_p = T(\Lambda + kI_p)T'$ den

$$S_k^{-1} = T_r S_r(k)^{-1} T_r' + T_{p-r} S_{p-r}(k)^{-1} T_{p-r}' \quad (3.26)$$

yazabiliriz. (3.22) eşitliğini $MMSE(\hat{\beta}(kI, J), \beta) = \sigma^2 S_k^{-1}$ olarak sadeleştirir ve (3.26) yi bu yeni formda yerine yazarsak

$$MMSE(\hat{\beta}(kI, J), \beta) = \sigma^2 T_r S_r(k)^{-1} T_r' + \sigma^2 T_{p-r} S_{p-r}(k)^{-1} T_{p-r}'$$

elde ederiz. Bu nedenle, $MMSE(\hat{\beta}(kI, J), \beta) - MMSE(\hat{\beta}_r(k), \beta)$ farkı için

$$\begin{aligned} M_1 - M_4 &= T_r S_r(k)^{-1} [k\sigma^2 I_r - k^2 T_r' \beta \beta' T_r] S_r(k)^{-1} T_r' \\ &\quad + T_{p-r} [\sigma^2 S_{p-r}(k)^{-1} - T_{p-r}' \beta \beta' T_{p-r}] T_{p-r}' \\ &\quad - k T_{p-r} T_{p-r}' \beta \beta' T_r S_r(k)^{-1} T_r' - k T_r S_r(k)^{-1} T_r' \beta \beta' T_{p-r} T_{p-r}' \end{aligned} \quad (3.27)$$

yazabiliriz. Şimdi $S^*(k)$ ve $\Lambda^*(k)^{-1}$ i aşağıdaki gibi tanımlayalım:

$$S^*(k) = \begin{pmatrix} \frac{1}{k} S_r(k) & 0 \\ 0' & I_{p-r} \end{pmatrix} \quad \text{ve} \quad \Lambda^*(k)^{-1} = \begin{pmatrix} k^{-1} I_r & 0 \\ 0' & S_{p-r}(k)^{-1} \end{pmatrix}.$$

Böylece (3.27) deki ifade

$$M_1 - M_4 = T S^*(k)^{-1} [\sigma^2 \Lambda^*(k)^{-1} - T' \beta \beta' T] S^*(k)^{-1} T'$$

farkına eşit olur. Bu durumda, $MMSE(\hat{\beta}(kI, J), \beta) - MMSE(\hat{\beta}_r(k), \beta)$ farkının

negatif olmayan tanımlı olması için gerek ve yeter koşul $\beta' T \Lambda^*(k) T' \beta \leq \sigma^2$ olmasıdır. $\Lambda^*(k)$ için ifadeyi yerine yazarsak, bu koşul daha da basitleşir ve aşağıdaki teoremi elde ederiz.

Teorem 3.8 : β nın $r - k$ sınıf tahmin edicisinin yansız ridge tahmin ediciye matris hata kareler ortalamasına göre üstün olması için gerek ve yeter koşul

$$k \leq \frac{\sigma^2 - \beta' T_{p-r} \Lambda_{p-r} T_{p-r}' \beta}{\beta' \beta} \quad (3.28)$$

olmasıdır.

$r = p$ iken $r - k$ sınıf tahmin edici ridge tahmin ediciye eşit olur ve (3.24) ile (3.28) ün denk olduğunu da görürüz.

(3.28) ve (3.24) koşulları karşılaştırıldığında, $r - k$ sınıf tahmin edicinin matris HKO kriterine göre yansız ridge tahmin ediciden üstün olması için bir k var ise ridge tahmin edici matris HKO kriterine göre yansız ridge tahmin ediciden üstün olacak şekilde bir k her zaman vardır sonucuna varırız.

3.2.4. Rezidülerin İncelenmesi ve Analizi

Regresyon analizi ile ilgilendiğimizde hatalar hakkında bazı varsayımlar yaparız. Hatalar (e) gözlenemediğinden, model doğru ise rezidüler ($\hat{e}$) gözlenen hatalar olarak ele alınır. Bu nedenle, rezidüler varsayımların doğruluğu ile ilgili bilgileri taşıması anlamında önemlidir.

EKK için rezidü vektörü $\hat{e} = [I - X(X'X)^{-1}X']e$ dir, skaler formda ise $i = 1, 2, \dots, n$ için $H = (h_{ij}) = X(X'X)^{-1}X'$ şapka matrisi olmak üzere

$\hat{e}_i = e_i - \sum_{j=1}^n h_{ij} e_j$ olarak ifade edilir. Bu eşitlik $\hat{e}$ ve e arasındaki ilişkiyi gösterir.

$(I - H) \approx I$ ise yani, tüm h_{ij} ler yeteri derecede küçük ve eğer köşegen elemanları h_{ii} ler aynı büyüklükte ise $\hat{e}$, e nin iyi bir tahmin edicisidir.

Regresyon çalışmasında hataların ikiyeşerli olarak ilişkisiz olduğunu varsayalım. Fakat, hataların ilişkili olduğu pek çok durum vardır. Bu nedenle rezidülerin varyans kovaryans yapılarını incelemek yararlı olabilir.

İki tahmin ediciyi hata kareler ortalamasına göre karşılaştırmak istatistikte sıkça kullanılan bir yöntemdir. Rezidüleri hataların birer tahmin edicisi olarak aldığımızdan, rezidüler üzerinde HKO kriterini uygulayabiliriz. Akdeniz (2001), Liu (1993) tarafından önerilen Liu tahmin edici ve Akdeniz ve Kaçıranlar (1995) tarafından ileri sürülen hemen hemen yansız Liu tahmin edicilerin rezidüleri ile EKK tahmin edicisinin rezidüsünü hata kareler ortalaması kriterine göre karşılaştırmıştır.

3.2.5. Yansız Ridge ve Ridge Rezidülerinin Kovaryans Yapılarının Karşılaştırılması

(2.1) den ridge rezidü ve kovaryans matrisini sırasıyla

$$\hat{e}_R = (I - XS_k^{-1}X')y$$

ve

$$Var(\hat{e}_R) = \sigma^2(I - 2XS_k^{-1}X' + XS_k^{-1}SS_k^{-1}X') \quad (3.29)$$

olarak buluruz.

(2.14) den yansız ridge rezidüyü

$$\hat{e}_{UR} = (I - XS_k^{-1}X')y - kXS_k^{-1}J \quad (3.30)$$

olarak yazabiliriz. J , $\hat{\beta}$ dan bağımsız olduğundan $Cov(\hat{\beta}, J) = 0$ dır. $Cov(\hat{\beta}, J)$ nin tanımından yani $Cov(\hat{\beta}, J) = E[(\hat{\beta} - E(\hat{\beta}))(J - E(J))']$ dan

$$E(\hat{\beta}J') = \beta\beta' \quad (3.31)$$

olduğu görülür. (1.1), EKK tahmin edicide yazılırsa, EKK nin yeni formu $\hat{\beta} = \beta + (X'X)^{-1}X'e$ şeklinde olacaktır. Bu yeni formla, (3.31) i

$$E(\hat{\beta}J') = E(\beta J') + (X'X)^{-1}X'E(eJ') = \beta\beta'$$

olarak yeniden yazabiliriz ki bu $E(eJ') = 0$ sonucuna götürür.

(3.30) dikkate alınırsa $E(\hat{e}_{UR}) = 0$ olduğu görülür. Şimdi tanımdan yararlanarak $Var(\hat{e}_{UR})$ yi bulalım:

$$Var(\hat{e}_{UR}) = E[(\hat{e}_{UR} - E(\hat{e}_{UR}))(\hat{e}_{UR} - E(\hat{e}_{UR}))'] = E(\hat{e}_{UR}\hat{e}_{UR}'). \quad (3.32)$$

$\hat{e}_{UR}$ ve $\hat{e}_{UR}'$ yi çarparsak

$$\begin{aligned} \hat{e}_{UR}\hat{e}_{UR}' &= (I - XS_k^{-1}X')yy'(I - XS_k^{-1}X') - k(I - XS_k^{-1}X')yJ'S_k^{-1}X' \\ &\quad - kXS_k^{-1}Jy'(I - XS_k^{-1}X') + k^2XS_k^{-1}JJ'S_k^{-1}X' \end{aligned} \quad (3.33)$$

elde ederiz. yy' , yJ' ve JJ' için (1.1) modelini kullanırsak aşağıdaki denklemleri elde ederiz:

$$yy' = X\beta\beta'X' + X\beta e' + e\beta'X' + ee' \quad (3.34)$$

$$yJ' = X\beta J' + eJ' \quad (3.35)$$

ve

$$\begin{aligned} JJ' &= [(J - \beta) + \beta][(J - \beta) + \beta]' \\ &= (J - \beta)(J - \beta)' + J\beta' - \beta\beta' + \beta J'. \end{aligned} \quad (3.36)$$

(3.34), (3.35) ve (3.36) eşitliklerini (3.33) de yazarsak,

$$\begin{aligned}
 \hat{e}_{UR}\hat{e}_{UR}' &= (I - XS_k^{-1}X')X\beta\beta'X'(I - XS_k^{-1}X') \\
 &+ (I - XS_k^{-1}X')X\beta e'(I - XS_k^{-1}X') \\
 &+ (I - XS_k^{-1}X')e\beta'X'(I - XS_k^{-1}X') \\
 &+ (I - XS_k^{-1}X')ee'(I - XS_k^{-1}X') \\
 &- k(I - XS_k^{-1}X')X\beta J'S_k^{-1}X' - k(I - XS_k^{-1}X')eJ'S_k^{-1}X' \\
 &- kXS_k^{-1}Je'(I - XS_k^{-1}X') - kXS_k^{-1}J\beta'X'(I - XS_k^{-1}X') \\
 &+ k^2XS_k^{-1}(J - \beta)(J - \beta)'S_k^{-1}X' + k^2XS_k^{-1}J\beta'S_k^{-1}X' \\
 &- k^2XS_k^{-1}\beta\beta'S_k^{-1}X' + k^2XS_k^{-1}\beta J'S_k^{-1}X'
 \end{aligned}$$

olur. $E(e) = 0$, $E(ee') = \sigma^2 I$, $E(J) = \beta$, $E[(J - \beta)(J - \beta)'] = \frac{\sigma^2}{k} I$ ve $E(eJ') = 0$ olduğundan (3.32) den

$$\begin{aligned}
 Var(\hat{e}_{UR}) &= (I - XS_k^{-1}X')X\beta\beta'X'(I - XS_k^{-1}X') \\
 &+ \sigma^2(I - XS_k^{-1}X')(I - XS_k^{-1}X') \\
 &- k(I - XS_k^{-1}X')X\beta\beta'S_k^{-1}X' - kXS_k^{-1}\beta\beta'X'(I - XS_k^{-1}X') \\
 &+ k\sigma^2XS_k^{-2}X' + k^2XS_k^{-1}\beta\beta'S_k^{-1}X'
 \end{aligned} \tag{3.37}$$

yazabiliriz. Parantez içindeki ifadeleri çarparak (3.37) ü yeniden düzenlersek

$$\begin{aligned}
 Var(\hat{e}_{UR}) &= X\beta\beta'X' - X\beta\beta'SS_k^{-1}X' - XS_k^{-1}S\beta\beta'X' + XS_k^{-1}S\beta\beta'SS_k^{-1}X' \\
 &+ \sigma^2 I - 2\sigma^2XS_k^{-1}X' + \sigma^2XS_k^{-1}SS_k^{-1}X' - kX\beta\beta'S_k^{-1}X' \\
 &+ kXS_k^{-1}S\beta\beta'S_k^{-1}X' - kXS_k^{-1}\beta\beta'X' + kXS_k^{-1}\beta\beta'SS_k^{-1}X' \\
 &+ k\sigma^2XS_k^{-2}X' + k^2XS_k^{-1}\beta\beta'S_k^{-1}X'
 \end{aligned}$$

olur. Ayrıca $S = S_k - kI$ yazar ve sadeleştirmeleri yaparak bu eşitliği

$$Var(\hat{e}_{UR}) = \sigma^2 I - \sigma^2 XS_k^{-1}X' \tag{3.38}$$

şeklinde ifade edebiliriz. Bu durumda $Var(\hat{e}_{UR}) - Var(\hat{e}_R)$ farkı için

$$Var(\hat{e}_{UR}) - Var(\hat{e}_R) = k\sigma^2 XS_k^{-2}X' \quad (3.39)$$

elde ederiz. $k > 0$ için (3.39) pozitif tanımlı olduğundan aşağıdaki teoremi verebiliriz.

Teorem 3.9 : $\hat{e}_{UR}$ ve $\hat{e}_R$ sırasıyla yansız ridge ve ridge rezidü vektörleri olsun. Bu durumda, $\hat{e}_{UR}$ $k > 0$ için her zaman $\hat{e}_R$ den daha büyük varyansa sahiptir.

Teorem 3.9 σ^2 nin tahmininde ridge rezidünün tercih edilmesi gerektiğini vurgulamaktadır.

3.2.6. $\hat{e}_{UR}$ ve $\hat{e}_R$ nin Matris HKO Kriterine Göre Karşılaştırılması

$Var(\hat{e}_{UR})$ nin hesaplanmasında gördük ki rezidülerin kovaryansını analiz etmek zordur. Biz aynı zamanda $\hat{e}_{UR}$ nin e ye olan uzaklığının, $\hat{e}_R$ nin e ye olan uzaklığından daha kısa olduğu durumlarla veya tam tersi durumlarla ilgileniriz. Bir tahmin edicinin performansını ölçmek için, rezidüleri hataların tahmin edicisi olarak aldığımızdan, matris HKO kriteri sıkça kullanılır.

$X\beta$ yı $X\hat{\beta}(k)$ ile tahmin ettiğimizden, $\hat{e}_R = y - X\hat{\beta}(k)$ rezidüsünü inceleyebiliriz. Böylece,

$$\hat{e}_R - e = (y - X\hat{\beta}(k)) - (y - X\beta) = -X(\hat{\beta}(k) - \beta)$$

elde ederiz. $MMSE(\hat{e}_R) = E[(\hat{e}_R - e)(\hat{e}_R - e)']$ olduğundan M_2 (3.23) de verilen şekilde olmak üzere $MMSE(\hat{e}_R) = XM_2X'$ dır. Benzer şekilde, $\hat{e}_{UR} - e = -X(\hat{\beta}(kI, J) - \beta)$ ve M_1 (3.22) ile verilmek üzere

$$MMSE(\hat{e}_{UR}) = XM_1X' \quad (3.40)$$

dır. Böylece,

$$MMSE(\hat{e}_{UR}) - MMSE(\hat{e}_R) = X[M_1 - M_2]X' = XS_k^{-1}[k\sigma^2 I - k^2\beta\beta']S_k^{-1}X'$$

elde ederiz. Bölüm 3.2.1 de verilen benzer düşünceleri kullanarak, karşılaştırmayı aşağıdaki teoremle sonuçlandırabiliriz.

Teorem 3.10 : $\hat{e}_{UR}$ ve $\hat{e}_R$ sırasıyla yansız ridge rezidü ve ridge rezidü vektörleri olsun. Bu durumda, $MMSE(\hat{e}_{UR}) - MMSE(\hat{e}_R)$ nin pozitif yarı tanımlı olması için gerek ve yeter koşul $k \leq \frac{\sigma^2}{\beta'\beta}$ olmasıdır.

3.2.7. Yansız Ridge ve EKK Rezidülerinin Kovaryans Yapılarının

Karşılaştırılması

(1.9) dan EKK rezidüyü ve kovaryans matrisini sırasıyla $\hat{e} = [I - X(X'X)^{-1}X']y$ ve

$$Var(\hat{e}) = \sigma^2(I - X(X'X)^{-1}X') \quad (3.41)$$

olarak buluruz. (3.38) ve (3.41) i kullanarak, $Var(\hat{e}_{UR}) - Var(\hat{e})$ farkını yazabiliriz:

$$Var(\hat{e}_{UR}) - Var(\hat{e}) = k\sigma^2 XS_k^{-1}S^{-1}X'. \quad (3.42)$$

(3.42) dikkate alınırsa aşağıdaki teoremi verebiliriz.

Teorem 3.11 : $\hat{e}_{UR}$ ve $\hat{e}$ sırasıyla yansız ridge ve EKK rezidü vektörleri olsun. Bu durumda, $k > 0$ için $Var(\hat{e}_{UR}) - Var(\hat{e})$ daima pozitif tanımlı bir matristir.

Teorem 3.11 in bir sonucu olarak $\hat{e}$ nin varyansı $\hat{e}_{UR}$ nin varyansından daha küçüktür diyebiliriz. Bu, σ^2 nın tahmininde EKK rezidünün tercih edilmesi gerektiğini gösterir. Bu sonuç Gauss-Markov teoreminin de bir sonucudur.

3.2.8. $\hat{e}_{UR}$ ve $\hat{e}$ nin Matris HKO Kriterine Göre Karşılaştırılması

$\hat{e} = y - X\hat{\beta}$ olduğundan $\hat{e} - e = -X(\hat{\beta} - \beta)$ ve $M_3 = MMSE(\hat{\beta}, \beta)$ olmak üzere $MMSE(\hat{e}) = XM_3X'$ dir. (3.40) ile verilen $MMSE(\hat{e}_{UR})$ den

$$MMSE(\hat{e}) - MMSE(\hat{e}_{UR}) = X[M_3 - M_1]X' = k\sigma^2 XS_k^{-1}S^{-1}X' \quad (3.43)$$

elde ederiz. $k > 0$ için (3.43) daima pozitif tanımlı olduğu için aşağıdaki teoremi verebiliriz.

Teorem 3.12 : $\hat{e}_{UR}$ ve $\hat{e}$ sırasıyla yansız ridge ve EKK rezidü vektörleri olsun. Bu durumda, $k > 0$ için e nin $\hat{e}_{UR}$ tahmin edicisi matris HKO anlamında $\hat{e}$ tahmin edicisinden daha iyidir.

Teorem 3.12, $\hat{e}_{UR}$ nin e ye olan uzaklığının $\hat{e}$ nin e ye olan uzaklığından daha yakın olduğunu söyler.

3.2.9. Yansız Ridge ve $r - k$ Sınıf Rezidülerin Kovaryans Yapılarının Karşılaştırılması

(2.20) den, $r - k$ sınıf rezidüyü ve kovaryans matrisini sırasıyla

$$\hat{e}_{r-k} = (I_p - XT_r(T_r'X'XT_r + kI_r)^{-1}T_r'X')y = (I_p - XT_rS_r(k)^{-1}T_r'X')y$$

ve

$$Var(\hat{e}_{r-k}) = \sigma^2 I - \sigma^2 XT_rS_r(k)^{-1}T_r'X' - k\sigma^2 XT_rS_r(k)^{-2}T_r'X' \quad (3.44)$$

olarak yazabiliriz. (3.26) yi (3.38) de yazarsak, $Var(\hat{e}_{UR})$

$$Var(\hat{e}_{UR}) = \sigma^2 I - \sigma^2 XT_rS_r(k)^{-1}T_r'X' - \sigma^2 XT_{p-r}S_{p-r}(k)^{-1}T_{p-r}'X' \quad (3.45)$$

olur. (3.44) ve (3.45) eşitliklerini kullanarak, $Var(\hat{e}_{r-k}) - Var(\hat{e}_{UR})$ farkını

$$Var(\hat{e}_{r-k}) - Var(\hat{e}_{UR}) = \sigma^2 X [T_{p-r} S_{p-r}(k)^{-1} T_{p-r}' - k T_r S_r(k)^{-2} T_r'] X' \quad (3.46)$$

olarak yazabiliriz. Şimdi $S^{**}(k)$ yı

$$S^{**}(k) = \begin{pmatrix} -k S_r(k)^{-2} & 0 \\ 0' & S_{p-r}(k)^{-1} \end{pmatrix}$$

olarak tanımlayalım. Bu durumda, (3.46) ile verilen ifade

$$Var(\hat{e}_{r-k}) - Var(\hat{e}_{UR}) = \sigma^2 X T S^{**}(k) T' X'$$

ye eşit olur. $\lambda_1 \geq \dots \geq \lambda_r \geq \lambda_{r+1} \geq \dots \geq \lambda_p \geq 0$ olduğundan ve ilk r özdeğer çok büyük, son $p-r$ özdeğer çok küçük ise $-k S_r(k)^{-2}$ çok küçük elemanları (yaklaşık sıfır) ve $S_{p-r}(k)^{-1}$ çok büyük pozitif elemanları içerir. Bu durumda $Var(\hat{e}_{r-k}) - Var(\hat{e}_{UR})$ pozitif yarı tanımlı olacaktır, böylece aşağıdaki teoremi verebiliriz.

Teorem 3.13 : $\hat{e}_{r-k}$ ve $\hat{e}_{UR}$ sırasıyla $r-k$ sınıf rezidü ve yansız ridge rezidü vektörler olsun ve λ_i , $i=1, \dots, p$ $X'X$ matrisinin özdeğerleri olsun. Eğer ilk r özdeğer çok büyük ve son $p-r$ özdeğer çok küçükse, $\hat{e}_{r-k}$ nin varyansı $\hat{e}_{UR}$ nin varyansından daha büyük olacaktır.

3.2.10. $\hat{e}_{UR}$ ve $\hat{e}_{r-k}$ nin Matris HKO Kriterine Göre Karşılaştırılması

$$\hat{e}_{r-k} = y - X \hat{\beta}_r(k) \quad \text{olduğundan,} \quad \hat{e}_{r-k} - e = -X (\hat{\beta}_r(k) - \beta) \quad \text{ve}$$

$M_4 = MMSE(\hat{\beta}_r(k), \beta)$ olmak üzere $MMSE(\hat{e}_{r-k}) = X M_4 X'$ dir. (3.40) ile verilen $MMSE(\hat{e}_{UR})$ den

$$MMSE(\hat{e}_{UR}) - MMSE(\hat{e}_{r-k}) = X[M_1 - M_4]X' \quad (3.47)$$

yazabiliriz. (3.47) nin negatif olmayan tanımlı olması için gerek ve yeter koşul $M_1 - M_4$ ün negatif olmayan tanımlı olmasıdır. Bölüm 3.2.3 deki sonuçlardan aşağıdaki teoremi elde ederiz.

Teorem 3.14 : $\hat{e}_{r-k}$ ve $\hat{e}_{UR}$ sırasıyla $r - k$ sınıf rezidü ve yansız ridge rezidü vektörler olsun. Bu durumda, e nin $\hat{e}_{r-k}$ sınıf tahmin edicisini $\hat{e}_{UR}$ tahmin edicisinden matris HKO anlamında daha iyi olması için gerek ve yeter koşul $k \leq \frac{\sigma^2 - \beta' T_{p-r} \Lambda_{p-r} T_{p-r}' \beta}{\beta' \beta}$ olmasıdır.

3.2.11. Sayısal Örnek

Bu bölümde Hoerl ve Kennard'ın (1981, Tablo1) ürettiği veri analiz edilecektir. Veri kümesi $\beta' \beta = 300$ kısıtlaması altında $\beta_1, \beta_2, \beta_3, \beta_4, \beta_5$ yı rasgele seçerek ve e yi $\sigma^2 = 1$ varyanslı sıfır ortalamalı normal dağılıma sahip varsayarak üretilmiştir. Veri Tablo 3.1 de verilmiştir. Model $y = X\beta + e$ olup $e, N(0, \sigma^2 I)$ dağılımına sahiptir.

Tablo 3.1. Veri kümesi

X_1	X_2	X_3	X_4	X_5	y
-0.4239	-0.2943	-0.1829	-0.2857	-0.3571	-12.3061
-0.3912	-0.2943	-0.2815	-0.2993	-0.3135	-9.8269
-0.1867	-0.1723	-0.3800	-0.2452	-0.1176	-6.5725
-0.1704	-0.1501	-0.2158	-0.2722	-0.2482	-6.8947
-0.1581	-0.1057	-0.3307	-0.1235	-0.1829	-6.1596
-0.1009	-0.1057	-0.2486	-0.1505	-0.2047	-2.6332
-0.0559	-0.2943	0.0306	-0.1776	-0.1611	-5.7948
-0.0232	-0.2277	0.0471	-0.1100	-0.1611	-5.9351
-0.0232	-0.0170	0.0142	0.0117	0.0131	-0.5752
0.0177	-0.0059	0.0799	-0.0288	0.2961	5.0592
0.0586	0.1826	0.0964	0.1875	0.1872	5.9096
0.2549	0.2714	0.2442	0.2957	0.4267	10.4006
0.3080	0.3157	0.2771	0.3498	0.2961	9.6367
0.4266	0.4044	0.3592	0.3633	0.1219	11.2918
0.4675	0.4932	0.4906	0.4850	0.4050	14.4004

Crouse, Jin ve Hanumara (1995) EKK, ridge ve yansız ridge tahmin edicilerin HKO performanslarını karşılaştırmak için bu veriyi analiz etmişlerdir. Bizim buradaki hesaplamalarımız Matlab 6.5 kullanılarak yapılmıştır. Bu nedenle Crouse, Jin ve Hanumara'nın (1995) hesaplamaları ile Matlab 6.5 hesaplamalar arasında makul farklılıklar olabilecektir.

$X'X$ matrisinin özdeğerleri 4.5792, 0.1940, 0.1549, 0.0584, 0.0138 dir. Koşul sayısı 331.1251 olup veri kümesinde çoklu iç ilişkinin olduğunu gösterir. Çoklu iç ilişki olması durumunda, EKK tahmin edicinin regresyon katsayıları Tablo 3.2 de görüldüğü gibi yanlış işarete sahip olacaktır. Bu nedenle, EKK tahmin edicisine güvenemeyiz. Alternatif olarak burada ridge, yansız ridge ve $r-k$ sınıf tahmin edicileri kullanacağız.

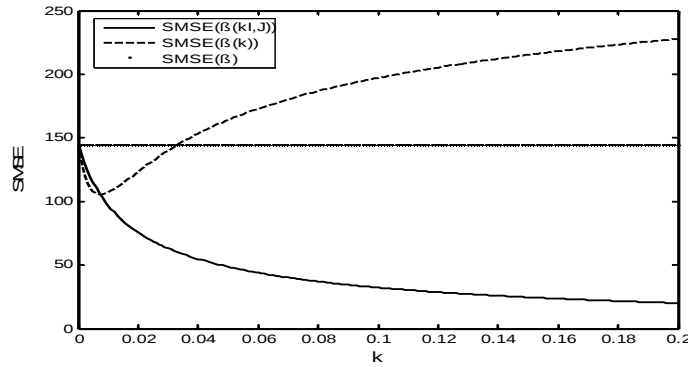
Tablo 3.2. $SMSE_1$ ve $SMSE_2$ sırasıyla tahmin ediciler ve rezidüler için $SMSE$ yi göstermek üzere $\hat{k}_{hkb} = 0.0133$ ve $\hat{k}_{cjh} = 0.0436$ için $SMSE$ ve tahminlerin değerleri

	β_1	β_2	β_3	β_4	β_5	$SMSE_1$	$SMSE_2$
Kitle	9.0269	8.3384	3.0903	3.3411	11.3258		
EKK	7.9567	16.6563	2.6446	-5.9090	12.3692	144.6341	7.1404
$\hat{\beta}(\hat{k}_{hkb})$	7.4966	12.5610	1.4810	0.0517	11.8267	111.7236	6.0336
$\hat{\beta}(\hat{k}_{cjh})$	6.9439	10.2121	1.4999	3.4230	11.0095	157.6149	6.2852
$\hat{\beta}(\hat{k}_{hkb}I, \bar{J})$	7.5058	12.6224	1.5332	-0.0062	11.8634	88.1927	5.9648
$\hat{\beta}(\hat{k}_{cjh}I, \bar{J})$	6.9812	10.3265	1.6252	3.3621	11.1187	52.3663	4.8554
$\hat{\beta}_r(\hat{k}_{hkb})$	6.6963	6.6401	6.4407	6.8108	6.3897	314.1622	20.7492
$\hat{\beta}_r(\hat{k}_{cjh})$	6.7405	6.6839	6.4832	6.8558	6.4318	314.1483	20.6856

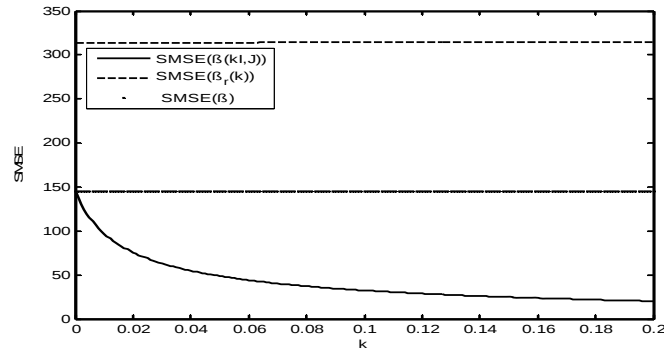
Tablo 3.2 de EKK, yansız ridge, ridge ve $r-k$ sınıf tahmin ediciler arasındaki karşılaştırmalar görülmektedir. $\hat{k}_{hkb}$ ile ridge yanlış işaret problemini ortadan kaldırmaktadır. Fakat, $SMSE(\hat{\beta}, \beta)$ ve $SMSE(\hat{\beta}(k), \beta)$ veya $SMSE(\hat{e})$ ve $SMSE(\hat{e}_R)$ arasındaki karşılaştırmalar büzülme (shrinkage) miktarının yeterli olmadığını göstermektedir (Şekil 3.1). Yanlı ve yansız tahmin ediciler varken yansız tahmin ediciyi tercih ederiz. Crouse, Jin ve Hanumara (1995) $1_{5 \times 1}$ birlerden oluşan

vektör olmak üzere $\bar{J} = \left[\sum_{i=1}^5 \hat{\beta}_i / 5 \right] 1_{5 \times 1}$ i gerçeğe uygun bir deneysel ön bilgi olarak önermişlerdir. $\hat{k}_{hkb}$ ile yansız ridge daha küçük hata kareler ortalamasına sahiptir fakat yanlış işaret probleminin ortadan kaldırmamaktadır. Yanlış işareti ortadan kaldırmak için $\hat{k}_{cjh}$ yeterli olmaktadır. Tablo 3.2 den görüyoruz ki, $\bar{J} = (6.7437 \ 6.7437 \ 6.7437 \ 6.7437 \ 6.7437)'$ ile yansız ridge tahmin edici diğer tahmin edicilerden daha küçük hata kareler ortalamasına sahip olması anlamında daha iyidir.

Tablo 3.2 özel k değerleri için tahmin edicilerin performanslarını özetlemektedir. Farklı k değerleri için yansız ridge tahmin edicinin diğer tahmin ediciler üzerindeki HKO anlamındaki üstünlüğünü inceleyebilmek için Şekil 3.1 ve Şekil 3.2 incelenebilir.



Şekil 3.1. k ya karşı EKK, ridge ve yansız ridge tahmin edicilerin SMSE grafikleri



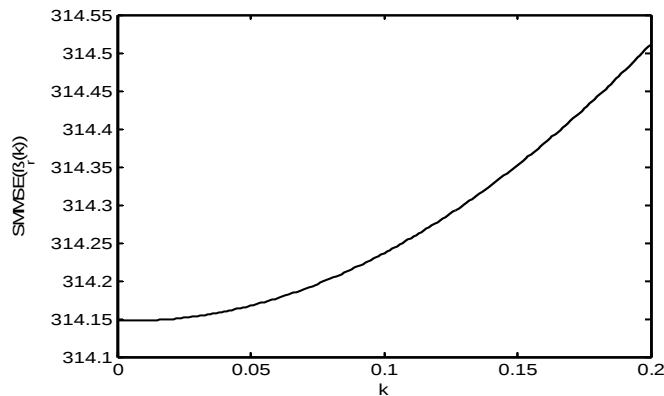
Şekil 3.2. k ya karşı EKK, yansız ridge ve $r - k$ sınıf tahmin edicilerin SMSE grafikleri

Teorem 3.6 da belirtildiği gibi, 0.0033 den küçük ($k < \frac{1}{300}$) k değerleri için ridge tahmin edicinin yansız ridge tahmin ediciden daha iyi olduğu Şekil 3.1 de görülmektedir. Örneğin, $k_1 = \hat{\sigma}^2 / \hat{\beta}'\hat{\beta} = 0.0027$ alırsak ridge tahmin edicinin yansız ridge ve EKK tahmin edicilerinin her ikisinden de HKO anlamında daha iyi olduğu görülür (Tablo 3.3).

Tablo 3.3. $SMSE_1$ ve $SMSE_2$ sırasıyla tahmin ediciler ve rezidüler için $SMSE$ yi göstermek üzere $k_1 = 0.0027$ için $SMSE$ ve tahminlerin değerleri

	β_1	β_2	β_3	β_4	β_5	$SMSE_1$	$SMSE_2$
EKK	7.9567	16.6563	2.6446	-5.9090	12.3692	144.6341	7.1404
$\hat{\beta}(k_1)$	7.8463	15.3118	2.1963	-3.9619	12.2385	117.1793	6.5907
$\hat{\beta}(k_1 I, J)$	7.8479	15.3307	2.2099	-3.9831	12.2463	126.6180	6.8029

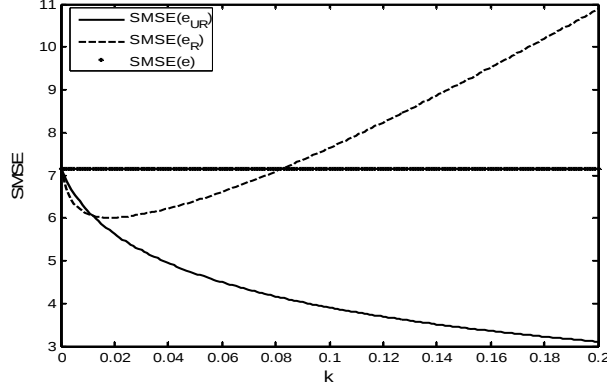
(3.28) eşitliğinin sayısal değeri k nın -0.0198 ($k < \frac{1-6.9512}{300}$) den küçük değerleri için $r-k$ sınıf tahmin edicinin yansız ridge tahmin ediciden daha iyi olduğunu belirtir. Fakat, $k > 0$ olduğundan -0.0198 den küçük bir k değeri yoktur. Ayrıca, Şekil 3.3 de görüldüğü gibi $\hat{\beta}_r(k)$ k nın artan bir fonksiyonudur. Bu nedenle, $\hat{\beta}(kI, J)$ nin her zaman $\hat{\beta}_r(k)$ dan iyi olduğunu Şekil 3.2 den görebiliriz.



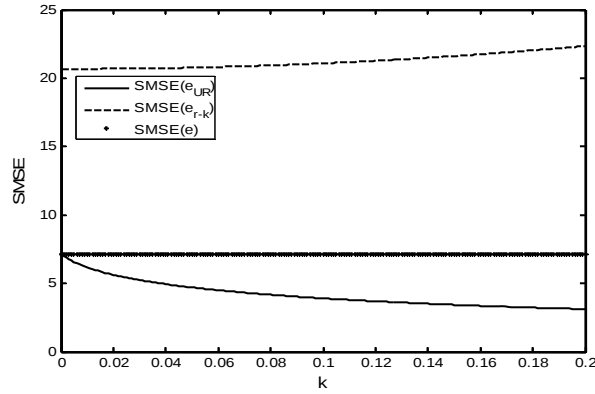
Şekil 3.3. k ya karşı $r-k$ sınıf tahmin edicinin $SMSE$ grafiği

Bölüm 3.3.4-.3.3.10 da rezidüleri de inceledik. Rezidülerden elde edilen sonuçların

tahmin edicilerden elde edilen sonuçlara benzer olduğu Şekil 3.1 ve Şekil 3.2 nin Şekil 3.4 ve Şekil 3.5 ile karşılaştırılmasından görülmektedir.



Şekil 3.4. k ya karşı EKK, yansız ve ridge rezidü için SMSE grafikleri



Şekil 3.5. k ya karşı EKK, yansız ve $r - k$ sınıf rezidülerin SMSE grafikleri

Sonuç olarak, teorik sonuçlar yansız ridge tahmin edicinin diğer tahmin edicilerden HKO anlamında her zaman iyi olmadığını göstermektedir. Hangi tahmin edicinin daha iyi olduğu bilinmeyen β ve σ^2 parametrelerine ve de k nın seçimine bağlıdır. Bu parametrelerin uygun seçimleri ile yansız ridge tahmin edici HKO anlamında iyi tahmin edicilerden biri olarak ele alınabilir. Ayrıca, yansız ridge rezidüsü daha küçük hata kareler ortalamasına sahip olması anlamında diğer rezidülerden daha iyi performans sergiler.

4. LİU TAHMİNDE YANLILIK PARAMETRESİNİN SEÇİMİ İÇİN ÖN TAHMİN AMAÇLI BİR KRİTER

4.1. Kalite Kontrolde Çoklu İç İlişki Problemleri

Çoklu iç ilişki problemi lineer regresyonda sıklıkla meydana gelen ciddi problemlerden biridir. Lineer regresyon analizinin mühendislerinde çok sık kullandığı istatistiksel tekniklerden biri olması nedeniyle bu bölümde mühendislik alanı olan kalite kontrolde çoklu iç ilişki problemine değineceğiz.

Bir kalite karakteristiği veya proses çıktı değişkeni ile ilgileniyorsak tek değişkenli metotlar kullanılır. Fakat uygulamada biribiri ile ilişkili pek çok değişken olabilir. Bu durumda tek değişkenli metotlar yerine çok değişkenli metotlar kullanılır. Kontrol grafikleri çok kullanışlı proses kontrol teknikleridir. Günümüzde otomatik denetleme prosedürlerin üretilen ürününü her bir birimi üzerinde pek çok parametreyi ölçmesi nedeniyle çok değişkenli kontrol grafiklerinin kullanımı önem kazanmaktadır. Bu kontrol grafikleri proses değişkenlerinin sayısı çok değilken (10 veya daha az) kullanışlıdır. Değişkenlerin sayısı arttıkça çok değişkenli kontrol grafikleri değişimi (shift) belirlemede etkinliklerini kaybederler. Bu gibi durumlarda en sık kullanılan yaklaşımlar temel bileşenler (principal components) analizi gibi boyut indirgeme teknikleridir.

En sık kullanılan çok değişkenli kontrol grafiği Hotelling T^2 'dir (Hotelling, 1947). Bu grafik çok değişkenli normallik varsayımı altında

$$T^2 = n(\bar{x} - \mu)' \Sigma^{-1} (\bar{x} - \mu) \quad (4.1)$$

istatistiğine dayanmaktadır. Burada μ her bir kalite karakteristiği için kontrol altındaki ortalama vektörü ve Σ kovaryans matrisidir. T^2 istatistiği kontrol sınırlarını aştığı zaman proses sinyal verecektir. T^2 kontrol grafiği için kontrol sınırları Alt (1985), Tracy, Young ve Mason (1992) ve Özkale (2004) gibi pek çok kaynaktan incelenebilir. Hotelling T^2 sık kullanılan bir yöntem olmasına rağmen

olumsuz yanları vardır. Bunlardan birincisi T^2 grafiği verinin ortalama ve varyansındaki değişimlerin her ikisine de duyarlıdır. Dolayısıyla T^2 grafiği bir değişim gösterdiğin de bunun ortalama ve varyansından mı kaynaklandığını bilemeyiz. İkinci olarak, T^2 grafiği prosesin kontrol dışına çıktığını belirlemesine rağmen bunun hangi değişkenden veya değişkenlerden kaynaklandığını belirleyemez. Bu olumsuzlukları gidermek için pek çok araştırmacı farklı yöntemler önermişlerdir.

Murphy (1987) diskriminant analizine dayalı bir yöntem önermiştir. Murphy (1987) değişkenlerin kümesini iki alt kümeye ayırmış ve hangi alt kümenin kontrol dışı sinyale neden olduğunu belirlemeye çalışmıştır. (4.1) de verilen T^2 istatistiğinin sinyal verdiğini varsayalım. Amacımız p değişkenden hangisinin veya hangi alt kümenin sinyal verdiğini belirlemektir. p_1 alt kümesinin sinyale neden olduğundan şüphelendiğimizi kabul edelim. Veri kümesini $\bar{x} = (\bar{x}^{(1)}, \bar{x}^{(2)})$ olarak parçalayalım öyle ki $\bar{x}^{(1)}$ sinyal verdiğinden şüphelendiğimiz p değişkenin alt kümesi ve $\bar{x}^{(2)}$ geri kalan p_2 değişkeni ($p = p_1 + p_2$) olsun.

$$T_p^2 = n(\bar{x} - \mu)' \Sigma^{-1} (\bar{x} - \mu)$$

ve

$$T_{p_1}^2 = n(\bar{x}^{(1)} - \mu^{(1)})' \Sigma_{11}^{-1} (\bar{x}^{(1)} - \mu^{(1)})$$

olmak üzere ($\mu^{(1)}$ ve Σ_{11}^{-1} , $\bar{x}$ nin parçalanışında p_1 değişkenine karşılık gelen ortalama ve kovaryans matrisidir) $D = T_p^2 - T_{p_1}^2$ ile ilgileniriz. Murphy (1987) D nin $\chi_{p_2}^2$ dağılımına sahip olduğunu göstermiştir ve D büyükse p_1 alt kümesinin sinyale neden olduğu hipotezinin red edileceğini, küçükse kabul edileceğini belirtmiştir.

Mason, Tracy ve Young (1995) T^2 istatistiğini bağımsız bileşenlere ayırmayı önermiştir. Böylece her bir bileşenin T^2 ye katkısı belirlenebilir. Bu şekildeki ayrıştırma ile değişken veya değişkenlerin sinyale katkısı belirlenmiş olacaktır. Mason, Tracy ve Young (1995) T^2 istatistiğini

$$T^2 = T_1^2 + \sum_{j=1}^{p-1} T_{j+1,1,\dots,j}^2 \quad (4.2)$$

olarak parçalamışlardır. Burada T_1^2 ilk değişken için T^2 istatistiği

$$T_1^2 = \frac{(x_{11} - \bar{x}_1)^2}{s_1^2}$$

ve

$$T_{j+1,1,\dots,j}^2 = \frac{(x_{i,j+1} - \bar{x}_{j+1,1,\dots,j})^2}{s_{j+1,1,\dots,j}^2}$$

istatistiği $x_1, \dots, x_j$ verilmişken x_{j+1} in dağılımının ortalama ve standart sapması ile düzeltilmiş x_i vektörünün $(j+1)$ -inci bileşenidir (ayrıntılı bilgi için Mason, Tracy ve Young, 1995; Mason ve Young, 2002; Özkale, 2004 incelenebilir).

T^2 nin (4.2) ile verilen parçası tek türlü olmayıp $p!$ şekilde parçalanabilir.

$T_j^2 > \frac{n+1}{n} F_{\alpha,1,n-1}$ ise kontrol dışı durumun x_j değişkeninden kaynaklandığını ve k koşullu terim sayısı olmak üzere $T_{j+1,1,\dots,j}^2 > \frac{(n+1)(n-1)}{n(n-k-1)} F_{\alpha,1,n-1}$ ise kontrol dışı durumun $x_1, \dots, x_j$ değişkenleri arasındaki ilişkiden kaynaklandığını söyleriz.

Mandel (1969) regresyon kontrol grafiğini önermiştir. Regresyon kontrol grafiğindeki temel fikir regresyon düzeltmesidir (regression adjustment). Verileri doğrudan grafiklendirmek yerine çoklu iç ilişki ortadan kaldırıldıktan sonra grafiklendirmeye regresyon düzeltmesi denir. Montgomery'ye (2005) göre regresyon kontrol grafiği lineer regresyonu ve kontrol grafiğini birleştiren bir teknik olup EKK tahmin yöntemine dayalıdır ve EKK yönteminin bir çok bilgisayar programında bulunması nedeniyle geniş bir kullanım alanına sahiptir.

Hawkins (1991) regresyon düzeltmesi fikrini çok değişkenli kalite kontrole uygulamıştır. Genel düşünce bir değişkenin diğer değişkenler üzerindeki regresyonundan elde edilen rezidülerin tek değişkenli Shewhart ve kümülatif toplam

(Cumulative Sum-CUSUM) kontrol grafiklerinin çizimi şeklindedir. Rezidüler birbiri ile ilişkisiz olduğundan çoklu iç ilişki problemi ortadan kalkmış olacak ve tek değişkenli kontrol grafikleri kullanılabilecektir. Hawkins (1993) yöntemini daha da geliştirmiş ve girdi değişkenlerini (input process variables veya upstream observations) bağımsız değişkenler, çıktı değişkenlerini (output variables) bağımlı değişkenler olarak şelale (cascade) proseslere uygulamıştır. Değişime maruz kalan bir ölçüm kendinden sonra gelen ölçümlerin bir kısmını veya tamamını etkiliyor fakat kendinden önceki ölçümleri etkilemiyorsa bu tip proseslere şelale proses denir.

Houshmand ve Javaheri (1998), çok değişkenli ridge rezidü grafikleri (multivariate ridge residual charts-MRRC) olarak adlandırdıkları yeni bir grafiklendirme yöntemi önermişlerdir. Bu yöntemde bir değişkenin diğer değişkenler üzerine olan regresyonundan elde edilen ridge rezidülerin Shewhart ve üstel ağırlıklı hareketli ortalama (exponentially weighted moving average-EWMA) grafiklerinin herhangi birinde kontrol dışı sinyalin ortaya çıkması karşılık gelen bağımlı değişkenin sinyale neden olduğu şeklinde yorumlanır. Houshmand ve Javaheri (1998) veriyi ilk olarak standartlaştırmayı önermişlerdir. Ridge yanlılık parametreleri k_i 'ler ($i = 1, \dots, p$) önceden belirlendikten sonra MRRC aşağıdaki dört adımla önerilmiştir;

1. adım: Girdi değişkenlerinin elde edilmesi,
2. adım: Önceden belirlenen k_i parametrelerinin kullanılması ile bir değişkenin diğer değişkenler üzerindeki regresyonundan ridge rezidülerin elde edilmesi,
3. adım: Rezidülerin Shewhart ve EWMA grafiklerinin oluşturulması,
4. adım: Kontrol dışı durum olması durumunda düzeltmelerin yapıp 1. adıma tekrar dönülmesi.

Değişken sayısı arttıkça yöntemin uygulanması zor olacaktır fakat değişkenlerdeki trendler belirlenebilir. Houshmand ve Javaheri'ye (1998) göre MRRC nin etkinliği yanlılık parametresinin seçimine bağlıdır.

Literatürde regresyon katsayılarının durağanlığını geliştirecek şekilde k yanlılık parametresinin seçimi için pek çok kaynak vardır. Alternatif bir diğer yaklaşım da modelin ön tahmin performansını geliştirecek şekilde k yı seçmektir.

Houshmand ve Javaheri (1998) ön tahmin hata kareler toplamını (prediction error sum of squares-PRESS) minimum yapacak şekilde k nın seçimini önermişlerdir.

4.2. Ön Tahmin Amaçlı Kriterler

Yanlı tahmin edicilerdeki yanlılık parametreleri, ya tahminler daha durağan olacak şekilde ya da ön tahminlerde bir ilerleme olacak şekildedir. Ridge regresyonda k daha çok katsayıların tahmininde bir gelişme olacak şekilde seçilir. Fakat model ön tahmin amaçlı kullanılıyorsa k nın seçimi için ön tahmin amaçlı kriterler kullanılır. Bu kriterler kavramsal ön tahmin (conceptual predictive- C_p), genelleştirilmiş çapraz geçerlilik (generalized cross-validation-GCV) ve ön tahmin hata kareler toplamı (PRESS) kriterleridir.

C_p istatistiği uydurulan modelin hata kareler ortalaması ile ilgilidir. Örneğin, x_i noktasındaki $\hat{y}(x_i)$ nin yani bir model için ön tahminin hata kareler ortalaması

$$MSE[\hat{y}(x_i)] = Var[\hat{y}(x_i)] + [Bias(\hat{y}(x_i))]^2$$

ile verilir. Eğer standartlaştırılmış hata kareler ortalamasını

$$\sum_{i=1}^n \frac{MSE[\hat{y}(x_i)]}{\sigma^2} = \sum_{i=1}^n \frac{var(\hat{y}(x_i)) + [Bias(\hat{y}(x_i))]^2}{\sigma^2} \quad (4.3)$$

olarak tanımlarsak ve eğer σ^2 biliniyorsa (4.3) deki ifade EKK tahmin edici için C_p istatistiği olarak adlandırılır (Montgomery ve Peck, 1992; Myers, 1990);

$$C_p = \frac{SS_{Res}}{\hat{\sigma}^2} - n + 2p .$$

Burada SS_{Res} , EKK kullanıldığında elde edilen rezidü kareler toplamıdır. Küçük C_p

değerleri küçük HKO anlamına geldiğinden küçük C_p değerleri tercih edilir. Mallows (1973) k yı belirlemek için C_p istatistiğini ridge tahmin edici için C_k istatistiğine

$$C_k = \frac{SS_{Res,k}}{\hat{\sigma}^2} - n + 2 + 2tr[H_k]$$

olarak uyarlamıştır. Burada $H_k = X(X'X + kI)^{-1}X'$ dir ve $SS_{Res,k}$ ridge regresyon kullanıldığında elde edilen rezidü kareler toplamıdır. C_k yı minimum yapacak şekilde k nın seçimi önerilir. Prosedür olarak k ya karşı C_k nın grafiği çizilerek C_k yı minimum yapan k değeri belirlenir.

Bir modelin ön tahmin performansını ölçmek için bir diğer kriter de PRESS istatistiğini minimum yapmaktır. Montgomery ve Friedman'ın (1993) belirttiği gibi bu prosedür Allen (1974) tarafından önerilmiş ve Golub, Heath ve Wahba (1979) tarafından geliştirilmiştir. PRESS istatistiğinin küçük değerleri, küçük değerli toplam hata kareler ortalamasına sahip modeli temsil eder yani regresyon modeli yeni gözlemlerde iyi ön tahminler sağlar.

Golub, Heath ve Wahba (1979) k nın seçimi için genelleştirilmiş çapraz geçerlilik (GCV) istatistiğini

$$GCV = \frac{\sum_{i=1}^n \hat{e}_{Ri}^2(k)}{\{n - [1 + tr(H_k)]\}^2}$$

olarak önermiştir. Burada $\hat{e}_{Ri}(k)$ belli bir k değeri için i -inci rezidüdür. GCV yi minimum yapacak şekilde k nın seçimi önerilir. Prosedür olarak k ya karşı GCV nın grafiği k yı belirlemede yararlı olacaktır.

Myers (1990) ve Montgomery ve Peck (1992) bu ön tahmin amaçlı kriterlerin benzer k değerlerini üretmeyebileceğini belirtmişlerdir.

Liu (1993) k nın tahminlerine paralel olarak d nin bazı tahminlerini

vermiştir. Skaler HKO'yu minimum yapan d nin tahminini (2.22) ile vermiştik. Liu (1993) aynı zamanda Liu tahmin edici için ön tahmin amaçlı kriterler olan C_p ve GCV istatistiklerini de

$$C_d = \frac{SS_{Res,d}}{\hat{\sigma}^2} + 2tr[\tilde{H}_d] - (n-2)$$

ve

$$GCV_d = \frac{SS_{Res,d}}{\{n - [1 + tr(\tilde{H}_d)]\}^2}$$

olarak vermiştir. Burada $SS_{Res,d}$ Liu tahmin edici için rezidü kareler toplamı ve $\tilde{H}_d = X(X'X + I)^{-1}(X'X + dI)(X'X)^{-1}X'$ dir. Bu istatistikleri minimum yapan d nin tahmin edicisini (her iki durumda da aynı tahmin ediciyi bulmuş)

$$\hat{d}_{CL} = 1 - \hat{\sigma}^2 \left(\sum_{i=1}^p \frac{1}{\lambda_i + 1} \middle/ \sum_{i=1}^p \frac{\lambda_i \hat{\alpha}_i^2}{(\lambda_i + 1)^2} \right)$$

olarak vermiştir. Fakat ön tahmin amaçlı bir kriter olan PRESS istatistiğini Liu tahmin edici için tanımlamamış ve bunu minimum yapan d nin tahminini vermemiştir.

4.2.1. Liu Tahmin Edicide d nin Seçimi İçin PRESS İstatistiği

Regresyon denklemleri yeni gözlemlerin ön tahmini için veya ortalama yanıtın (mean response) tahmini için sıklıkla kullanılırlar. Genel olarak, bir analizci ön tahminin hata kareler ortalaması minimum olacak şekildeki açıklayıcı değişkenleri seçmek ister. Bu ise küçük etkiye sahip açıklayıcı değişkenlerin modelden atılması anlamına gelmektedir. Bu anlamda PRESS istatistiği kullanılabilir (Montgomery ve Peck, 1992).

Allen (1971, 1974) tarafından önerilen PRESS istatistiği

$$\hat{e}_{(i)} = y_i - \hat{y}_{(i)}, \quad i = 1, 2, \dots, n \quad (4.4)$$

olarak tanımlanan PRESS rezidülere bağlıdır. Burada $\hat{y}_{(i)}$, i -inci gözlem atılarak diğer $n-1$ gözleme uydurulan regresyon modeli ile elde edilen y_i nin ön tahminidir. PRESS rezidü EKK rezidüden de hesaplanabilir. Şöyleki, $y_i - \hat{y}_i$ EKK rezidü ve h_{ii} $H = X(X'X)^{-1}X'$ şapka matrisinin i -inci köşegen elemanı olmak üzere

$$\hat{e}_{(i)} = \frac{y_i - \hat{y}_i}{1 - h_{ii}} \quad (4.5)$$

dır. Böylece PRESS istatistik n PRESS rezidünün kareleri toplamı olarak

$$PRESS = \sum_{i=1}^n \left(\frac{y_i - \hat{y}_i}{1 - h_{ii}} \right)^2 = \sum_{i=1}^n \left(\frac{y_i - \hat{y}_i}{1 - x'_i (X'X)^{-1} x_i} \right)^2 \quad (4.6)$$

şeklinde hesaplanır. Burada x'_i , X in i -inci satırıdır. PRESS $n-1$ gözlemin her alt kümesini tahmin kümesi olarak kullanır ve diğer gözlemi de ön tahmin kümesini oluşturmak için kullanır. PRESS istatistiği yeni gözlemlerin iyi bir ön tahmini olan modeli belirler. En iyi modelin seçimi için küçük değerli PRESS tercih edilir. PRESS istatistiği uydurulmuş çeşitli modellerin durağanlığı hakkında çok fazla bilgi sağlaması ve etkili (influential) gözlemlerin belirlenmesin de yararlı bir istatistiktir.

PRESS istatistiği ridge tahmin ediciye

$$PRESS_R(k) = \sum_{i=1}^n \left[\frac{\hat{e}_{Ri}(k)}{1 - h_{k,ii}} \right]^2 = \sum_{i=1}^n \left[\frac{y_i - x'_i \hat{\beta}_R(k)}{1 - x'_i (X'X + kI)^{-1} x_i} \right]^2 \quad (4.7)$$

olarak uyarlanmış olup $\hat{e}_{Ri}(k)$ belli bir k değeri için i -inci rezidü ve $h_{k,ii}$ H_k nın i -inci köşegen elemanıdır. PRESS modelin ön tahmin performansının bir ölçüsü

olduğundan $PRESS_R(k)$ yı minimum yapacak şekilde k yı seçmek önerilir. Fakat $PRESS_R(k)$, k nın bir konveks fonksiyonu olmadığından k yı seçmek zordur. Montgomery ve Friedman (1993), $PRESS_R(k)$ nın yaklaşık olarak $0 \leq k \leq 2$ aralığında minimum olduğunu ve k ya karşı $PRESS_R(k)$ nın grafiğinin bu minimum değeri bulmada yararlı olabileceğini belirtmişlerdir.

Şimdi PRESS istatistiğini Liu tahmin edici için tanımlayalım. Bu durumda $\hat{\beta}_d$, d nin lineer bir fonksiyonu olduğundan hesaplamalar daha kolay olacaktır.

$$\hat{e}_{Ri}(k) = y_i - x'_i \hat{\beta}_R(k) \quad (4.8)$$

ridge rezidünün i -inci elemanı,

$$h_{k-ii} = x'_i (X'X + kI)^{-1} x_i \quad (4.9)$$

ridge şapka matrisi H_k nın i -inci köşegen elemanı ve

$$\hat{e}_{d(i)} = y_i - x'_i \hat{\beta}_{d(i)} \quad (4.10)$$

i -inci gözlem silindikten sonraki Liu rezidü olsun. Bu durumda aşağıdaki teoremi verebiliriz.

Teorem 4.1 : Liu tahmin edici için PRESS istatistik

$$PRESS_d = \sum_{i=1}^n (\hat{e}_{d(i)})^2 = \sum_{i=1}^n \left[(1-d) \frac{\hat{e}_{Ri(1)}}{1-h_{1-ii}} + d \frac{\hat{e}_i}{1-h_{ii}} \right]^2 \quad (4.11)$$

dır ve

$$\hat{d}_{PR} = \frac{\sum_{i=1}^n \frac{\hat{e}_{Ri}(1)}{1-h_{1-ii}} \left(\frac{\hat{e}_{Ri}(1)}{1-h_{1-ii}} - \frac{\hat{e}_i}{1-h_{ii}} \right)}{\sum_{i=1}^n \left(\frac{\hat{e}_{Ri}(1)}{1-h_{1-ii}} - \frac{\hat{e}_i}{1-h_{ii}} \right)^2} \quad (4.12)$$

de minimum olur.

İspat : (4.11) de verilen $PRESS_d$ yi tanımlayalım. (2.21) ile verilen $\hat{\beta}_d$ yi $\hat{\beta}_d = (X'X + I)^{-1}(X'X + dI)\hat{\beta}$ olarak da yazabiliriz.

$$\hat{\beta}_{d(i)} = (X'_{(i)} X_{(i)} + I)^{-1}(X'_{(i)} X_{(i)} + dI)\hat{\beta}_{(i)} \quad (4.13)$$

i -inci gözlem silindikten sonraki Liu tahmin edici ve $\hat{\beta}_{(i)} = (X'_{(i)} X_{(i)})^{-1} X'_{(i)} y_{(i)}$, i -inci gözlem silindikten sonraki EKK tahmin edici olsun. $X_{(i)}$ ve $y_{(i)}$, i -inci gözlem silindikten sonraki X ve y dir. Sırasıyla $B = X'X$ ve $B = X'X + I$, $c = x_i$ (X in i -inci kolonu) alarak ve $(B - cc')^{-1} = B^{-1} + \frac{B^{-1}cc'B^{-1}}{1 - c'B^{-1}c}$ yı kullanarak

$$(X'_{(i)} X_{(i)})^{-1} = (X'X - x_i x'_i)^{-1} = (X'X)^{-1} + \frac{(X'X)^{-1} x_i x'_i (X'X)^{-1}}{1 - h_{ii}} \quad (4.14)$$

ve

$$\begin{aligned} (X'_{(i)} X_{(i)} + I)^{-1} &= (X'X + I - x_i x'_i)^{-1} \\ &= (X'X + I)^{-1} + \frac{(X'X + I)^{-1} x_i x'_i (X'X + I)^{-1}}{1 - h_{1-ii}} \end{aligned} \quad (4.15)$$

elde ederiz. Burada $h_{ii} = x'_i (X'X)^{-1} x_i$ ve $h_{1-ii} = x'_i (X'X + I)^{-1} x_i$ dir.

$X'_{(i)} y_{(i)} = X'y - x_i y_i$ olduğundan (4.14) den $\hat{\beta}_{(i)}$

$$\begin{aligned}\hat{\beta}_{(i)} &= \left[(X'X)^{-1} + \frac{(X'X)^{-1}x_i x_i' (X'X)^{-1}}{1-h_{ii}} \right] (X'y - x_i y_i) \\ &= \left[I + \frac{(X'X)^{-1}x_i x_i'}{1-h_{ii}} \right] (\hat{\beta} - (X'X)^{-1}x_i y_i)\end{aligned}$$

olarak yazılabilir. Parantez içerisindeki terimleri çarparak $\hat{\beta}_{(i)}$ yi tekrar yazarsak

$$\hat{\beta}_{(i)} = \hat{\beta} - (X'X)^{-1}x_i y_i + \frac{(X'X)^{-1}x_i x_i' \hat{\beta}}{1-h_{ii}} - \frac{(X'X)^{-1}x_i x_i' (X'X)^{-1}x_i y_i}{1-h_{ii}}$$

elde ederiz. $\hat{e}_i = y_i - x_i' \hat{\beta}$ ve $h_{ii} = x_i' (X'X)^{-1} x_i$ olduğundan son eşitlik

$$\hat{\beta}_{(i)} = \hat{\beta} - \frac{\hat{e}_i}{1-h_{ii}} (X'X)^{-1} x_i \quad (4.16)$$

şeklinde de ifade edilebilir (Montgomery ve Peck, 1992). Ayrıca

$$(X'_{(i)} X_{(i)} + dI) = (X'X - x_i x_i' + dI) = (X'X + dI) - x_i x_i' \quad (4.17)$$

dır. $\hat{y}_i = x_i' \hat{\beta}$, $\hat{y}_{di} = x_i' \hat{\beta}_d$ ve $\tilde{h}_{d-ii}$, $\tilde{H}_d$ nin i -inci köşegen elemanı olmak üzere

(4.15), (4.16), (4.17) yı kullanarak ve parantez içindeki ifadeleri sadeleştirirsek

(4.13) deki $\hat{\beta}_{d(i)}$ yi

$$\begin{aligned}\hat{\beta}_{d(i)} &= \hat{\beta}_d - \frac{(X'X + I)^{-1}x_i \hat{y}_i}{1-h_{1-ii}} + \frac{(X'X + I)^{-1}x_i \hat{y}_{di}}{1-h_{1-ii}} \\ &\quad - \frac{\hat{e}_i}{1-h_{ii}} (X'X + I)^{-1} (X'X + dI) (X'X)^{-1} x_i \\ &\quad + \frac{\hat{e}_i h_{ii}}{1-h_{ii}} \frac{(X'X + I)^{-1} x_i}{1-h_{1-ii}} - \frac{\hat{e}_i \tilde{h}_{d-ii}}{1-h_{ii}} \frac{(X'X + I)^{-1} x_i}{1-h_{1-ii}}\end{aligned}$$

olarak yazabiliriz. Son eşitliği soldan x'_i ile çarparsak sadeleştirmelerden sonra

$$x'_i \hat{\beta}_{d(i)} = \frac{1}{1-h_{1-ii}} \hat{y}_{di} - \frac{h_{1-ii}}{1-h_{1-ii}} y_i + \frac{\hat{e}_i}{(1-h_{1-ii})(1-h_{ii})} (h_{1-ii} - \tilde{h}_{d-ii})$$

elde edilir. Böylece $\hat{e}_{di} = y_i - \hat{y}_{di}$ olmak üzere $\hat{e}_{d(i)} = y_i - x'_i \hat{\beta}_{d(i)}$ ifadesinin

$$\hat{e}_{d(i)} = \frac{\hat{e}_{di}}{1-h_{1-ii}} - \frac{\hat{e}_i}{(1-h_{1-ii})(1-h_{ii})} (h_{1-ii} - \tilde{h}_{d-ii}) \quad (4.18)$$

ifadesine eşit olduğu görülebilir.

$$(X'X + I)^{-1}(X'X + I) = I \quad \text{olduğundan} \quad (X'X + I)^{-1} = I - (X'X + I)^{-1} X'X$$

olarak yazılabilir. Bu son eşitliği sağdan $(X'X)^{-1}$ ile çarparsak

$$(X'X + I)^{-1}(X'X)^{-1} = (X'X)^{-1} - (X'X + I)^{-1} \quad (4.19)$$

elde ederiz. (4.19) u kullanarak, $\tilde{h}_{d-ii}$ aşağıdaki formda yazılabilir

$$\begin{aligned} \tilde{h}_{d-ii} &= x'_i (X'X + I)^{-1} (X'X + dI) (X'X)^{-1} x_i \\ &= x'_i (X'X + I)^{-1} (I + d(X'X)^{-1}) x_i \\ &= x'_i (X'X + I)^{-1} x_i + dx'_i (X'X + I)^{-1} (X'X)^{-1} x_i \\ &= x'_i (X'X + I)^{-1} x_i + dx'_i (X'X)^{-1} x_i - dx'_i (X'X + I)^{-1} x_i \\ &= (1-d)h_{1-ii} + dh_{ii}. \end{aligned} \quad (4.20)$$

Ayrıca

$$\begin{aligned}
 \hat{e}_{di} - d\hat{e}_i &= (y_i - x'_i \hat{\beta}_d) - d(y_i - x'_i \hat{\beta}) \\
 &= (1-d)y_i + x'_i (d\hat{\beta} - \hat{\beta}_d) \\
 &= (1-d)y_i + x'_i [dI - (X'X + I)^{-1}(X'X + dI)]\hat{\beta} \\
 &= (1-d)y_i + (d-1)x'_i (X'X + I)^{-1}(X'X)\hat{\beta} \\
 &= (1-d)y_i + (d-1)x'_i (X'X + I)^{-1}X'y \\
 &= (1-d)(y_i - \hat{y}_{Ri}(1)) = (1-d)\hat{e}_{Ri}(1)
 \end{aligned}$$

olduğundan

$$\hat{e}_{di} = (1-d)\hat{e}_{Ri}(1) + d\hat{e}_i \quad (4.21)$$

dır. (4.20) ve (4.21) yi kullanarak (4.18), $a = \frac{\hat{e}_{Ri}(1)}{1-h_{1-ii}}$ ve $b = \frac{\hat{e}_i}{1-h_{ii}}$ olmak üzere

$$\hat{e}_{d(i)} = (1-d)\frac{\hat{e}_{Ri}(1)}{1-h_{1-ii}} + d\frac{\hat{e}_i}{1-h_{ii}} = (1-d)a + db \quad (4.22)$$

olarak da yazılabilir. Böylece (4.11) ile verilen $PRESS_d$ yi elde ederiz.

(4.22) yi dikkate alırsak $PRESS_d = \sum_{i=1}^n [(1-d)a + db]^2$ yazabiliriz. $PRESS_d$, d nin karesel bir fonksiyonu olduğundan $PRESS_d$ nin türevini alıp,

$$\begin{aligned}
 \frac{dPRESS_d}{dd} &= \sum_{i=1}^n 2(-a+b)[(1-d)a + db] \\
 &= 2 \sum_{i=1}^n [d(a^2 - 2ab + b^2) - a^2 + ab] \\
 &= 2 \sum_{i=1}^n [d(a-b)^2 - a(a-b)] \\
 &= 2 \sum_{i=1}^n d(a-b)^2 - 2 \sum_{i=1}^n a(a-b)
 \end{aligned}$$

sıfıra eşitleyerek minimum yapan değerini

$$d = \frac{\sum_{i=1}^n a(a-b)}{\sum_{i=1}^n (a-b)^2} \quad (4.23)$$

olarak buluruz. a ve b yi (4.23) de yazarsak ispat tamamlanmış olur.

$$\frac{\mathbf{d}^2 PRESS_d}{\mathbf{d}d^2} = 2 \sum_{i=1}^n \left(\frac{\hat{e}_{Ri}(1)}{1-h_{1-ii}} - \frac{\hat{e}_i}{1-h_{ii}} \right)^2 \geq 0$$

olduğundan $PRESS_d$ konveks bir fonksiyondur, bu nedenle konveks fonksiyonlar için geçerli olan aşağıdaki özelliklere sahip olur (Giorgi ve ark., 2004): $\varphi : T \rightarrow R$ konveks bir fonksiyon olsun, bu durumda aşağıdaki özellikler geçerli olur;

- i) $t_1 < t_2 < t_3$ olmak üzere $\forall t_1, t_2, t_3 \in T$ için $\varphi(t_2) \leq \frac{t_3 - t_2}{t_3 - t_1} \varphi(t_1) + \frac{t_2 - t_1}{t_3 - t_1} \varphi(t_3)$
- ii) $\forall t_1, t_2 \in T$ için $\varphi(t_1) - \varphi(t_2) \geq (t_1 - t_2) \frac{d\varphi(t_2)}{dt_2}$.

$$\varphi(t) = \sum_{i=1}^n \left[(1-t) \frac{\hat{e}_{Ri}(1)}{1-h_{1-ii}} + t \frac{\hat{e}_i}{1-h_{ii}} \right]^2, \quad t_1 = 0, \quad t_2 = d \quad \text{ve} \quad t_3 = 1 \quad \text{alırsak (i) özelliğinden}$$

$PRESS_d \leq (1-d)PRESS_R(1) + dPRESS$ elde ederiz. (ii) özelliği için $t_1 = 0, t_2 = d$ ve $t_3 = 1$ alırsak sırasıyla $\hat{d}_{PR}$ için $PRESS_d \leq PRESS_R(1)$ ve $PRESS_d \leq PRESS$ elde ederiz. Sonuç olarak $\hat{d}_{PR}$ için $PRESS_d$ nin $PRESS$ ve $k=1$ için $PRESS_R(k)$ nin her ikisinden ve bunların konveks kombinasyonundan daha iyi olduğunu söyleyebiliriz.

Analizci regresyonun amacına göre bir prosedür takip etmelidir. Önemli olan tahmin ise ridge trace, HKO gibi metotlar incelenmeli. Ön tahmin önemli ise C_p , $PRESS$ ve GCV metotları uygulanmalıdır. Belli bir veri kümesinde çoklu iç ilişki belirlendiğin de bu yöntemlerden her hangi birinin diğerlerine göre iyi sonuç vermesi

için kesin bir kriter yoktur.

$\hat{d}_{CL}$ yi hesaplayabilmek için σ^2 nin yansız tahminine gereksinim vardır, fakat $\hat{d}_{PR}$ için yoktur. Bu nedenle $\hat{d}_{PR}$, σ^2 nin tahminine gerek duyulmadığı durumlarda kullanılabilir. Ayrıca, (4.22) den görüldüğü gibi $\hat{e}_{d(i)}$, $\frac{\hat{e}_{Ri}(1)}{1-h_{1-ii}}$ ve $\frac{\hat{e}_i}{1-h_{ii}}$ nin kombinasyonudur. $\frac{\hat{e}_{Ri}(1)}{1-h_{1-ii}}$, $k=1$ iken ridge şapka matrisinin köşegen elemanları ile ridge rezidülerin ağırlıklandırılması ve $\frac{\hat{e}_i}{1-h_{ii}}$, şapka matrisinin köşegen elemanlarının EKK rezidü ile ağırlıklandırılması olduğundan $PRESS_d$, EKK ve $k=1$ iken ridge tahmin edicilerden hesaplanabilir.

4.2.2. Sayısal Örnek

Teorikte bulduğumuz sonuçları örneklendirmek için Woods, Steinour ve Starke'nin (1932) literatüre kazandırdığı ve daha sonra Hald (1952), Gorman ve Toman (1966), Kaçıranlar ve ark. (1999), Kaçıranlar ve Sakallıoğlu (2001) gibi çalışmalarda pek çok kez analiz edilen Portland çimento verisini inceleyeceğiz. Bu veri çimentonun oluşturulması ve sertleştirilmesinde açığa çıkan ısı miktarını göstermektedir. Çimentonun bileşenleri olan dört bağımsız değişken, tricalcium aluminate (X_1), tetracalcium sillicate (X_2), tetracalcium alumino ferrite (X_3) ve dicalcium silicate (X_4) dir. Bağımlı değişken (y) ise bir gram çimento başına düşen kalori cinsinden ısı miktarıdır.

Hesaplamalarımızda Matlab 6.5 programı kullanılmıştır, bu nedenle diğer hesaplamalarla arasında farklılıklar olabilecektir. Hesaplamalara başlamadan önce değişkenler standartlaştırılmıştır. XX matrisinin özdeğerleri $\lambda_1 = 2.235704$, $\lambda_2 = 1.576066$, $\lambda_3 = 0.186606$, $\lambda_4 = 0.001623$ olup koşul sayısı $\frac{\lambda_1}{\lambda_4} \cong 1376.880621$ çoklu iç ilişkinin olduğunu gösterir. Bunun nedeni dört bağımsız değişken arasındaki yaklaşık lineer ilişkidir: $X_1 + X_2 + X_3 + X_4 \approx 100$. Bu nedenle EKK tahmin ediciyi

kullanamayız, alternatif olarak ridge ve Liu tahmin edicileri kullanabiliriz.

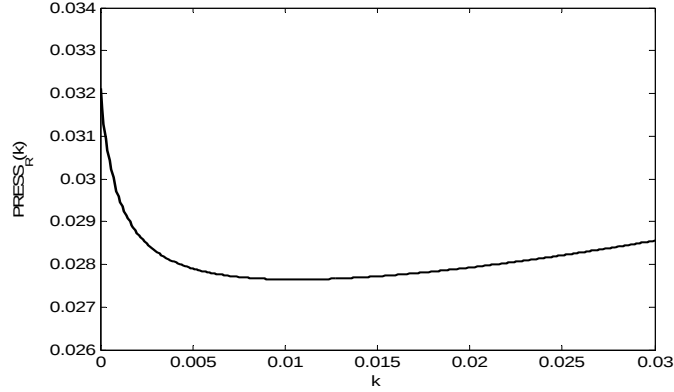
Montgomery ve Friedman (1993), $PRESS_R(k)$ nın minimum değerinin $0 \leq k \leq 2$ aralığı içerisinde elde edildiğini belirtmiştir. Bu nedenle k değerleri 0 dan 2 ye kadar 0.0001 lik artışlarla belirlenmiş ve bu değerlere karşılık gelen $PRESS_R(k)$ değerleri hesaplanmıştır. $PRESS_R(k)$ yı minimum yapan k değeri 0.0106 olarak elde edilmiştir. Bu değeri görebilmek için k nın 0.009 dan 0.0119 a kadar olan değerleri Tablo 4.1 de gösterilmiştir. Ayrıca Şekil 4.1 de modelin ön tahmin yeteneğinin azalan k ile birlikte arttığını ve yaklaşık olarak 0.0106 da en iyi düzeyde olduğunu göstermektedir.

Tablo 4.1. k çeşitli değerleri için $PRESS_R(k)$ değerleri

k	$PRESS_R(k)$	k	$PRESS_R(k)$
0.009	0.02765992619505	0.0105	0.02764580302603
0.0091	0.02765812489628	0.0106*	0.02764574381239
0.0092	0.02765645868829	0.0107	0.02764578252858
0.0093	0.02765492442657	0.0108	0.02764591730666
0.0094	0.02765351908193	0.0109	0.02764614633744
0.0095	0.02765223973501	0.0110	0.02764646786806
0.0096	0.02765108357125	0.0111	0.02764688019973
0.0097	0.02765004787603	0.0112	0.02764738168551
0.0098	0.02764913003015	0.0113	0.02764797072831
0.0099	0.027648327500550	0.0114	0.02764864577885
0.01	0.02764763786100	0.0115	0.02764940533384
0.0101	0.02764705873875	0.0116	0.02765024793416
0.0102	0.02764658786037	0.0117	0.02765117216318
0.0103	0.02764622302353	0.0118	0.02765217664511
0.0104	0.02764596209870	0.0119	0.02765326004346

Tablo 4.1 de $PRESS_R(k)$ nın minimum değerini yüz milyonuncu ondalık basamağında aldığı görülüyor. k için kullanılan ondalık sayı da $PRESS_R(k)$ nın performansını etkileyecektir. k nın 0 dan 2 ye kadar olan değerlerini hesaplamak için büyük ondalıklı sayılar kullandığımız da $PRESS_R(k)$ nın minimum değerinin değiştiğini gördük. Bu nedenle küçük ondalıklı sayılar kullanmak daha uygun

sonuçlar verecektir, fakat bu defa hesaplamalar zaman alıcı olabilmektedir. Sonuç olarak $0 \leq k \leq 2$ aralığı $PRESS_R(k)$ nın minimum değerini hesaplamak için geniş bir aralıktır.



Şekil 4.1. k ya karşı $PRESS_R(k)$ nın grafiği

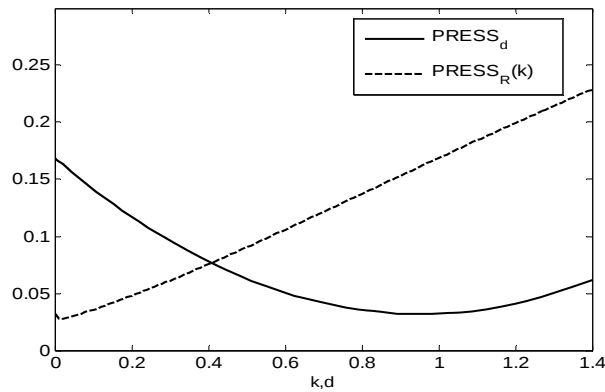
$k = 0.0106$ için ridge tahmin ediciyi kullandığımız da $XX + kI$ nın koşul sayısı 183.765605 e düştü. Fakat hala çoklu iç ilişkinin varlığını görüyoruz. Çoklu iç ilişkinin var olma nedeni koşul sayısını uygun seviyeye düşürecek kadar k nın yeterince büyük olmamasıdır. Hala çoklu iç ilişki olduğundan $k = 0.0106$ iken ridge tahmin ediciye güvenemeyiz. Alternatif olarak Liu tahmin ediciyi kullanabiliriz.

Şimdi Liu tahmin ediciyi ön tahmin amaçlı olarak kullanalım. (4.12) yi kullandığımız da $PRESS_d$ nin $\hat{d}_{PR} = 0.95403160232377$ de $PRESS_d = 0.03180660516500$ değeri ile minimum olduğunu görüyoruz. Ayrıca $PRESS_R(1) = 0.16875043046731$ ve $PRESS = 0.03212453820644$ olduğundan $\hat{d}_{PR}$ için $PRESS_d < PRESS_R(1)$ ve $PRESS_d < PRESS$ özellikleri sağlanır. $PRESS_d \leq (1 - d)PRESS_R(1) + dPRESS$ özelliği ise Tablo 4.2 den görülebilir.

Tablo 4.2. d nin çeşitli değerleri için $(1-d)PRESS_R(1)+dPRESS$ değerleri

d	$(1-d)PRESS_R(1)+dPRESS$	$PRESS_d$
0	0.16875043046731	0.1687504346731
0.1	0.15508784124122	0.14154656984012
0.2	0.14142525201513	0.11735188063541
0.3	0.12776266278905	0.09616636285316
0.4	0.11410007356296	0.07799001649337
0.5	0.10043748433687	0.06282284155605
0.6	0.08677489511079	0.05066483804120
0.7	0.07311230588470	0.04151600594881
0.8	0.05944971665861	0.03537634527889
0.9	0.04578712743253	0.03224585603143
0.94	0.04032209174209	0.03183622833074
0.95	0.03895583281948	0.03180905069112
0.96	0.0375895789687	0.03181196476574
1	0.03212453820644	0.03212453820644

Şekil 4.2 k ve d ye karşı sırasıyla $PRESS_R(k)$ ve $PRESS_d$ grafiklerini göstermektedir. Böylece aldıkları minimum değerler de görülmektedir. Şekil 4.2 aynı zamanda bir PRESS istatistiğinin diğerinden her zaman iyi olmadığını da göstermektedir. Tablo 4.3 bu PRESS istatistiklerin hangisinin diğerinden daha iyi olduğunu gösterir. $PRESS_R(k)$ nın 0.408 den küçük değerler için daha iyi olduğunu görüyoruz. $PRESS_R(k)$ ve $PRESS_d$ nin minimum değerlerini karşılaştırdığımız da $PRESS_R(k)$ nın küçük bir farkla daha küçük olduğu görülüyor. Fakat $PRESS_R(k)$ nın hesaplanması zordur.

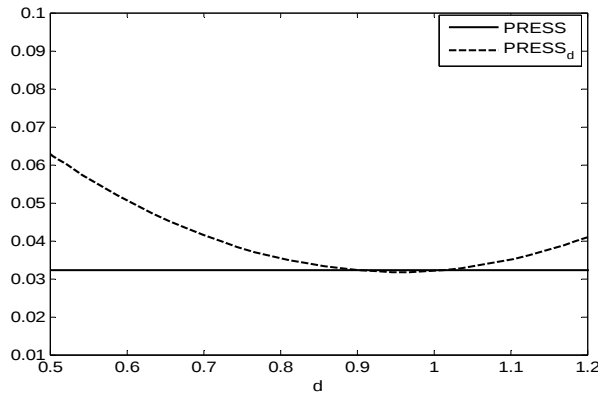


Şekil 4.2. k ve d ye karşı sırasıyla $PRESS_R(k)$ ve $PRESS_d$ grafikleri

Tablo 4.3. k ve d nin çeşitli değerleri için $PRESS_R(k)$ ve $PRESS_d$ değerleri

$k = d$	$PRESS_R(k)$	$PRESS_d$
0.4077	0.07662444032125	0.07671521161212
0.4078	0.07663912452134	0.07669877306225
0.4079	0.07665380939432	0.07668233752156
0.408*	0.07666849493995	0.07666590499004
0.4081	0.07668318115802	0.07664947546769
0.4082	0.07669786804830	0.07663304895451
0.4083	0.07671255561055	0.07661662545051

Şekil 4.3 de d ye karşı $PRESS_d$ ve $PRESS$ grafikleri görülmektedir. $PRESS_d$ nin her zaman $PRESS$ den iyi olmadığını görüyoruz. Fakat $\hat{d}_{PR}$ için $PRESS_d$, $PRESS$ den iyidir, daha küçük değere sahiptir. Bu ise $\hat{d}_{PR}$ için $PRESS_d < PRESS$ özelliğini doğrular.



Şekil 4.3. d ye karşı $PRESS_d$ ve $PRESS$ grafikleri

Karşılaştırma yapmak amacıyla regresyon katsayılarının $k = 0.0106$ için ridge, $\hat{d}_{PR} = 0.95403160232377$ için Liu ve EKK tahminlerini elde ettik. Tablo 4.4 tahmin prosedürlerini, rezidü kareler toplamını (SS_{Res}), HKO ve çoklu belirleyicilik katsayısını (R^2) özetlemektedir.

Tablo 4.4. Tahmin özetleri

	OLS	$PRESS_R(k = 0.0106)$	$PRESS_{\hat{d}_{PR}=0.95403160232377}$
β_1	0.60651195174781	0.51295957673165	0.58992944855781
β_2	0.52770563116372	0.31510630482537	0.51464075570808
β_3	0.04338969836570	-0.05613659092619	0.03471734831340
β_4	-0.16028741596540	-0.38309841449252	-0.16486607558979
SS	0.0176243795923199	0.01787914389276	0.01784534806967
MSE	1.21862916628714	0.03501491837107	1.32252582868391
R^2	0.98237562040765	0.97669010274632	0.96799158490726

Çoklu iç ilişki olması durumunda bir lineer regresyon modelinin ön tahmin performansı için iki yanlı tahmin edici (ridge ve Liu) kullandık. Ridge tahmin edici Montgomery ve Friedman (1993) tarafından da kullanılmış ve $0 \leq k \leq 2$ aralığında $PRESS_R(k)$ yı minimum yapan k değerini seçmek olarak önerilmiştir. Farklı veri kümeleri için k bir değerini alabilir. $PRESS_d$ nin Bölüm 4.2.1 de verilen özelliklerinden dolayı $PRESS_R(k)$ yı minimum yapan k değeri bir veya yaklaşık bir olması durumunda hesaplanması kolay olduğu için $PRESS_d$ yi kullanabiliriz.

5. SONUÇ VE ÖNERİLER

Bu çalışmada çoklu iç ilişki problemleri ve onun istatistiksel sonuçları ele alınmıştır. Çoklu iç ilişki olması durumunda modeldeki parametrelerin tamamı tam olarak tahmin edilemez. Verilen bir veri kümesinde EKK yöntemini kullanarak parametrelerin hangi lineer kombinasyonlarının tahmin edilebilir olduğunu belirlemek için $X'X$ matrisinin özdeğerlerini ve özvektörlerini incelemek gerekmektedir. İlgilendiğimiz parametrelerin lineer kombinasyonu çoklu iç ilişkiden etkilenmiyorsa klasik metotlar uygulanabilir. Fakat ilgilendiğimiz parametreler çoklu iç ilişkiden etkileniyorsa alternatif metotlar kullanılmalıdır.

Çoklu iç ilişki problemini ortadan kaldırmak için EKK yöntemine alternatif olarak ileri sürülen tahmin ediciler yanlı tahmin edicilere neden olmuştur. Burada bu tahmin edicilerden ridge, genelleştirilmiş ridge, temel bileşenler, kısıtlı EKK, modified ridge, yansız ridge, kısıtlı ridge, $r - k$ sınıf, Liu, $r - d$ sınıf, Liu-tipi tahmin ediciler ele alınmış ve bazı özellikleri incelenmiştir.

Çoklu iç ilişkinin istatistiksel sonuçlarını ortadan kaldırmak için pek çok tahmin edicinin olması, bunlardan hangisinin daha iyi bir performans sergilediğinin araştırılmasına neden olmuştur. Daha küçük HKO ya sahip tahmin edicinin daha iyi olması nedeniyle literatürde pek çok yanlı tahmin edicinin karşılaştırmaları yapılmıştır ve yapılmaya devam edilmektedir. Bazı araştırmacılar tarafından yapılan karşılaştırmalar verilmiş ve henüz yapılmamış olan bazı karşılaştırmalar yapılmıştır.

Matris HKO kriterinin skaler HKO ve genelleştirilmiş HKO kriterlerinden daha güçlü bir kriter olması nedeniyle $r - d$ sınıf tahmin edicinin EKK, ridge ve temel bileşenlerden matris HKO anlamında üstün olması için gerek ve yeter koşullar verilmiştir. Bu koşulların bilinmeyen parametrelere bağlı olduğu görülmüş ve bu koşullar için hipotez testleri oluşturulmuştur. $r - d$ ve $r - k$ sınıf tahmin ediciler matris HKO kriterine göre karşılaştırılmıştır.

Çoklu iç ilişki olması durumunda Crouse, Jin ve Hanumara'nın (1995) önerdiği yansız ridge tahmin edici EKK tahmin ediciye alternatif olarak ileri sürülen bir tahmin edicidir. Crouse, Jin ve Hanumara (1995) yansız ridge tahmin ediciyi Swindel'in (1976) önerdiği modified ridge tahmin edicideki ön bilginin stokastik

olması durumunda önermişlerdir. Bu çalışmada ise yansız ridge tahmin edici Farebrother (1984) ve Troskie ve ark.'nın (1994) çalışmaları yardımıyla iki farklı şekilde elde edilmiş ve bazı özelliklerini incelenmiştir. Crouse, Jin ve Hanumara (1995), Hoerl ve Kennard'ın (1981) ürettiği veriyi kullanarak önerdikleri tahmin ediciyi EKK ve ridge tahmin edici ile karşılaştırılmışlardır. Bu çalışmada ise yansız ridge tahmin edici ridge, EKK ve $r - k$ sınıf tahmin edicilerle matris HKO kriterine göre karşılaştırılmıştır. Ayrıca, yansız ridge rezidü ile ridge, EKK, $r - k$ sınıf rezidülerin varyans kovaryans yapıları incelenmiş ve matris HKO kriterine göre karşılaştırılmaları yapılmıştır. Bunlara paralel olarak Hoerl ve Kennard'ın (1981) ürettiği veri kullanılarak teorikte bulduğumuz sonuçlar örneklendirilmiştir.

Çoklu iç ilişki probleminin lineer regresyonda sıklıkla meydana gelen ciddi problemlerden biri olması ve lineer regresyon analizinin mühendislerin de çok sık kullandığı istatistiksel tekniklerden biri olması nedeniyle mühendislik alanı olan kalite kontrolde çoklu iç ilişki problemine değinilmiş. Çoklu iç ilişki olması durumunda Hotelling T^2 kontrol grafiği önerilmiştir. Fakat bu kontrol grafiğinin bazı olumsuz yanları vardır. Bu olumsuzlukları gidermek için ileri sürülen çözüm yöntemlerinden bahsedilmiştir.

Çoklu iç ilişki olması durumunda önerilen ridge ve Liu tahmin edicileri sırasıyla yanlılık parametreleri k ve d ye bağlıdır. Yanlılık parametreleri, ya tahminler daha durağan ya da ön tahminlerde bir ilerleme olacak şekilde seçilir. Ridge regresyonda k daha çok katsayıların tahmininde bir gelişme olacak şekilde seçilmiştir. Fakat model ön tahmin amaçlı kullanılıyorsa k nın seçimi için ön tahmin amaçlı kriterler kullanılır. Ön tahmin amaçlı olarak da k nın seçimi ile ilgili C_k , GCV, PRESS istatistikleri önerilmiştir. Liu (1993) ön tahmin amaçlı olarak d nin seçimi ile ilgili C_d ve GCV_d istatistiklerini önermiş ve bunları minimum yapan d nin tahmin edicisini vermiştir. Bu çalışmalara paralel olarak Liu tahmin edici için PRESS istatistiği tanımlanmış, bazı özellikleri incelenmiş ve bunu minimum yapan d nin tahmin edicisi verilmiştir. Teorikte bulduğumuz sonuçları Portland çimento verisi üzerinde örneklendirilmiştir.

Bu çalışmada elde edilen sonuçlar yanlış tahmin edicilerden birinin diğerine her zaman için üstün olmadığını göstermektedir. Fakat pek çok tahmin ediciyi içerek şekilde yeni bir tahmin edici verilebilir. Bu yeni tahmin edicinin kendini oluşturan tahmin edicilerden daha iyi bir tahmin yöntemi olması beklenir. Bu beklenti simülasyon çalışması yapılarak gözlemlenebilir. Ayrıca, Houshmand ve Javaheri'nin (1998) çalışmalarına paralel bir çalışma olarak Liu tahmin edici için PRESS istatistiği kullanılarak yeni bir çok değişkenli kontrol grafiği oluşturulabilir. Simülasyon çalışması yardımıyla bu yeni kontrol grafiğinin etkinliği çok değişkenli ridge rezidü kontrol grafiği ile karşılaştırılabilir.

KAYNAKLAR

- ALLEN, D. M., 1971. Mean Square Error of Prediction as a Criterion for Selection of Variables. *Technometrics*, 13, 469-475.
- _____, 1974. The Relationship Between Variable Selection and Data Augmentation and a Method for Prediction. *Technometrics*, 16, 125-127.
- ALT, F. B., 1985. Multivariate Quality Control, in *Encyclopedia of Statistical Sciences*, Vol. 6, N. L. Johnson and S. Kotz (eds) Wiley, New York.
- AKDENİZ, F., 2001. The Examination and Analysis of Residuals for Some Biased Estimators in Linear Regression. *Comm. Statist. Theory Methods*, 30, 1171-1183.
- AKDENİZ, F., and KAÇIRANLAR, S., 1995. On the Almost Unbiased Generalized Liu Estimator and Unbiased Estimation of the Bias and MSE. *Comm. Statist. Theory Methods*, 24, 1789-1797.
- BAKSALARY, J. K., and TRENKLER, G., 1991. Nonnegative and Positive Definiteness of Matrices Modified by Two Matrices of Rank One. *Linear Algebra Appl.*, 151, 169-184.
- BAYE, M. R., and PARKER, D. F., 1984. Combining Ridge and Principal Component Regression: A Money Demand Illustration. *Comm. Statist. Theory Methods*, 13, 2, 197-205.
- BELSLEY, D. A., KUH, E., and WELSCH, R. E., 1980. *Regression Diagnostics Identifying Influential Data and Sources of Collinearity*. John Wiley and Sons: New York.
- CROUSE, R. H., JIN, C., HANUMARA, R. C., 1995. Unbiased Ridge Estimation with Prior Information and Ridge Trace. *Comm. Statist. Theory Methods*, 24, 9, 2341-2354.
- FAREBROTHER, R. W., 1984. The Restricted Least Squares Estimator and Ridge Regression. *Comm. Statist. Theory Methods*, 13, 191-196.
- FARRAR, D. E. and GLAUBER, R. R., 1967. Multicollinearity in Regression Analysis: The Problem Revisited. *The Review of Economics and Statistics*, 49, 1, 92-107.

- FRISCH, R., 1934. Statistical Confluence Analysis by Means of Complete Regression Systems. Universitets Økonomiske Institut, Oslo.
- GIORGI, G., GUERRAGGIO, A., and THIERFELDER, J., 2004. Mathematics of Optimization: Smooth and Nonsmooth Case. Amsterdam: Elsevier.
- GOLUB, G. H., HEATH, C. G., and WAHBA, W., 1979. Generalized Cross-Validation as a Method for Choosing a Good Ridge Parameter. *Technometrics*, 21, 215-223.
- GORMAN, J. W., and TOMAN, R. J., 1966. Selection of Variables for Fitting Equations to Data. *Technometrics*, 8, 27-51.
- GRAYBILL, F. A., 1976. Theory and Application of the Linear Model. Duxbury Press, California, 704p.
- GROSS, J., 2003. Restricted Ridge Estimation. *Statistics and Probability Letters*, 65, 57-64.
- GRUBER, M. H. J., 1998. Improving Efficiency by Shrinkage: The James-Stein and Ridge Regression Estimators. Marcell Dekker, Inc. New York.
- HALD, A., 1952. Statistical Theory with Engineering Applications. New York: Wiley.
- HAWKINS, D. M., 1991. Multivariate Quality Control Based on Regression-Adjusted Variables. *Technometrics*, 33, 1, 61-75.
- _____, 1993. Regression Adjustment for Variables in Multivariate Quality Control. *Journal of Quality Technology*, 25, 3, 170-182.
- HOERL, A. E., and KENNARD, R. W., 1970. Ridge Regression: Biased Estimation for Nonorthogonal Problems. *Technometrics*, 12, 1, 55-67.
- HOERL, A. E., and KENNARD, R. W., 1981. Ridge Regression-1980 Advances, Algorithms, and Applications. *American Journal of Mathematical and Management Sciences*, 1, 5-83.
- HOERL, A.E., KENNARD, R.W., and BALDWIN, K. F., 1975. Ridge Regression: Some Simulation. *Communication in Statistics*, 4, 105-123.
- HOTELLING, H., 1933. Analysis of a Complex of Statistical Variables into Principal Components. *Journal of Educational Psychology*, 24, 417-441 and 489-520.

- _____, 1947. Multivariate Quality Control Techniques of Statistical Analysis. C. Eisenhart, M. W. Hastay, W. A. Wallis (eds). McGrawhill: New York.
- HOUSHMAND, A. A., and JAVAHERİ, A., 1998. Multivariate Ridge Residual Charts. *Quality Engineering*, 10, 617-624.
- JUDGE, G. G., GRIFFITHS, W. E., HILL, R. C., LÜTKEPOHL, H., and LEE, T-C., 1985. *The Theory and Practice of Econometrics*, John Wiley and Sons, New York, 1018p.
- KAÇIRANLAR, S., and SAKALLIOĞLU, S., 2001. Combining the Liu Estimator and the Principal Component Regression Estimator. *Comm. Statist. Theory Methods*, 30, 12, 2699-2705.
- KAÇIRANLAR, S., SAKALLIOĞLU, S., AKDENİZ, F., STYAN., G. P. H., and WERNER, H. J., 1999. A New Biased Estimator in Linear Regression and a Detailed Analysis of the Widely-Analysed Dataset on Portland Cement. *Sankhya, Series B, Indian Journal of Statistics*, 61, 443-459.
- LAWLESS, J. F. and WANG, P., 1976. A Simulation Study of Ridge and Other Regression Estimators. *Communication in Statistics*, 7, 139-164.
- LIPOVETSKY, S. and CONKLIN, W. M., 2001. Multiobjective Regression Modifications for Collinearity. *Computers and Operations Research*, 28, 1333-1345.
- LIU, K., 1993. A New Class of Biased Estimate in Linear Regression. *Comm. Statist. Theory Methods*, 22, 2, 393-402.
- _____, 2003. Using Liu-Type Estimator to Combat Collinearity. *Comm. Statist. Theory Methods*, 32, 5, 1009-1020.
- MALLOWS, C. L., 1973. Some Comments on C_p . *Technometrics*, 15, 661-675.
- MANDEL, B. J., 1969. The Regression Control Chart. *Journal of Quality Technology*, 1, 1-9.
- MARQUARDT, D. W., 1970. Generalized Inverses, Ridge Regression, Biased Linear Estimation, and Nonlinear Estimation. *Technometrics*, 12, 591-612.
- MASON, R. L., and YOUNG, J. C., 2002. *Multivariate Statistical Process Control with Industrial Applications*. ASA-SIAM Series on Statistics and Applied Probability: Philadelphia.

- MASON, R. L., TRACY, N. D., and YOUNG, J. C., 1995. Decomposition of T^2 for multivariate control chart interpretation. *Journal of Quality Technology*, 27, 2, 99-108.
- McDONALD, G. C., and GALARNEAU, D. I., 1975. A Monte Carlo Evaluation of Some Ridge-Type Estimators. *Journal of the American Statistical Association*, 70, 350, 407-416.
- MONTGOMERY, D. C., 2005. *Introduction to Statistical Quality Control*. 5th edition John Wiley and Sons: Hoboken, NJ.
- MONTGOMERY, D. C., and FRIEDMAN D. J., 1993. Prediction Using Regression Models with Multicollinear Predictor Variables. *IIE Transactions*, 25 (3), 73-85.
- MONTGOMERY, D. C. and PECK, E. A., 1992. *Introduction to Linear Regression Analysis*. John Wiley and Sons, New York, 526p.
- MURPHY, B. J., 1987. Selecting Out of Control Variables with the T^2 Multivariate Quality Control Procedure. *The Statistician*, 36, 571-583.
- MYERS, R. H., 1990. *Classical and Modern Regression with Applications*. Duxbury Press: California.
- NOMURA, M., and OHKUBO, T., 1985. A Note on Combining Ridge and Principal Component Regression. *Comm. Statist. Theory Methods*, 14, 10, 2489-2493.
- ÖZKALE, M. R., 2004. Yüksek Lisans Tezi, İstatistiksel Kalite Kontrol Yöntemleri ve Uygulamalar. (Yayınlanmamış)
- PLISKIN, J. L., 1987. A Ridge-Type Estimator and Good Prior Means. *Comm. Statist. Theory Methods*, 16, 12, 3429-3437.
- RAO, C. R. and TOUTENBURG, H., 1995. *Linear Models: Least Squares and Alternatives*. Springer, New York, 351p.
- RUBIO, H., FIRINGUETTI, L., 2002. The Distribution of Stochastic Shrinkage Parameters in Ridge Regression. *Comm. Statist. Theory Methods*, 31, 9, 1531-1547.
- SARKAR, N., 1992. A New Estimator Combining the Ridge Regression and the Restricted Least Squares Methods for Estimation. *Comm. Statist. Theory Methods*, 21, 7, 1987-2000.

- _____, 1996. Mean Square Error Comparison of Some Estimators in Linear Regressions with Multicollinearity. *Statistics and Probability Letters*, 30, 133-138.
- SILVEY, S. D., 1969. Multicollinearity and Imprecise Estimation. *Journal of the Royal Statistical Society, Series B (Methodological)*, 31, 3, 539-552.
- STEIN, C., 1956. Inadmissibility of Usual Estimator for the Mean of a Multivariate Normal Distribution. *Proceeding of the Third Berkeley Symposium on Mathematical Statistics and Probability*. Univeristy of California Press, Berkeley, 197-206.
- SWINDEL, F. F., 1976. Good Ridge Estimators Based on Prior Information. *Comm. Statist. Theory Methods*, A5 (11), 1065-1075.
- TOUTENBURG, H., 1982. *Prior Information in Linear Models*. John Wiley and Sons, New York, 212p.
- THEIL, H., 1971. *Principles of Econometrics*. John Wiley and Sons, New York.
- THEOBALD, C. M., 1974. Generalizations of Mean Square Error Applied to Ridge Regression. *Journal of the Royal Statistical Society. Series B (Methodological)*, 36, 1, 103-106.
- TRACY, N. D., YOUNG, J. C., and MASON, R. L., 1992. Multivariate Control Charts for Individual Observations. *Journal of Quality Technology*, 24, 88-95.
- TRENKLER, G., and TOUTENBURG, H., 1990. Mean Squared Error Comparisons of Biased Regression Estimators. *Comm. Statist. Theory Methods*, A9 (12), 1247-1259.
- TROSKIE, C. G., CHALTON, D. O., STEWART, T. J., and JACOBS, M., 1994. Detection of Outliers and Influential Observations in Regression Analysis Using Stochastic Prior Information. *Comm. Statist. Theory Methods*, 23, 3453-3476.
- VINOD, H. D. and ULLAH, A., 1981. *Recent Advances in Regression Methods*. Marcel Dekker, New York, 361p.
- WOODS, H., STEINOUR, H. H., and STARKE, H. R., 1932. Effect of Composition of Portland Cement on Heat Evolved During Hardening. *Industrial and*

Engineering Chemistry, 24, 1207-1214.

ÖZGEÇMİŞ

1979 yılında Kadirli’de doğdum. İlk ve orta öğrenimimi Kadirli’de, lise öğrenimimi ise Adana’da tamamladıktan sonra 1998 yılında Çukurova Üniversitesi Fen Edebiyat Fakültesi Matematik Bölümü’nde lisans öğrenimime başladım. 2002 yılında bu bölümden mezun olduktan sonra aynı bölümde yüksek lisansa başlayıp yine aynı yıl bölümde araştırma görevlisi olarak göreve başladım. 2004 yılında İstatistik Bölümü’nden yüksek lisansımı tamamlayarak aynı bölümde doktora öğrenimime başladım. Şu anda da İstatistik Bölümünde araştırma görevlisi olarak görevime devam etmekteyim.