

**DOĐRUSAL REGRESYONDA RIDGE, LIU VE LASSO  
TAHMİN EDİCİLERİ ÜZERİNE BİR ÇALIŞMA**

**A STUDY ON RIDGE, LIU AND LASSO ESTIMATOR IN  
LINEAR REGRESSION**

**Ayşe KÜÇÜK**

**Prof. Dr. Meral ÇETİN**

**Tez Danışmanı**

Hacettepe Üniversitesi

Lisansüstü Eğitim-Öğretim ve Sınav Yönetmeliğinin

İstatistik Anabilim Dalı için Öngördüğü

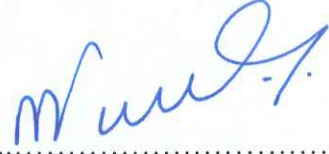
YÜKSEK LİSANS TEZİ olarak hazırlanmıştır.

**2019**

AYŞE KÜÇÜK'ün hazırladığı “Doğrusal Regresyonda Ridge, Liu ve LASSO Tahmin Edicileri Üzerine Bir Çalışma” adlı bu çalışma aşağıdaki jüri tarafından İSTATİSTİK ANABİLİM DALI'nda YÜKSEK LİSANS TEZİ olarak kabul edilmiştir.

Doç. Dr. Meral EBEGİL

Başkan



Prof. Dr. Meral ÇETİN

Danışman



Prof. Dr. Serpil AKTAŞ ALTUNAY

Üye



Prof. Dr. Duru KARASOY

Üye



Dr. Öğr. Üyesi Onur TOKA

Üye



Bu tez Hacettepe Üniversitesi Fen Bilimleri Enstitüsü tarafından YÜKSEK LİSANS TEZİ olarak ..... / ..... /..... tarihinde onaylanmıştır.

Prof. Dr. Menemşe GÜMÜŞDERELİOĞLU

Fen Bilimleri Enstitüsü Müdürü

## ETİK

Hacettepe Üniversitesi Fen Bilimleri Enstitüsü, tez yazım kurallarına uygun olarak hazırladığım bu tez çalışmada,

- tez içindeki bütün bilgi ve belgeleri akademik kurallar çerçevesinde elde ettiğimi,
- görsel, işitsel ve yazılı tüm bilgi ve sonuçları bilimsel ahlak kurallarına uygun olarak sunduğumu,
- başkalarının eserlerinden yararlanılması durumunda ilgili eserlere bilimsel normlara uygun olarak atıfta bulunduğumu,
- atıfta bulunduğum eserlerin tümünü kaynak olarak gösterdiğimi,
- kullanılan verilerde herhangi bir tahrifat yapmadığımı,
- ve bu tezin herhangi bir bölümünü bu üniversitede veya başka bir üniversitede başka bir tez çalışması olarak sunmadığımı

beyan ederim.

22 / 07 / 2019

İmza  
Ayşe KÜÇÜK



## YAYINLAMA VE FİKRİ MÜLKİYET HAKLARI

Enstitü tarafından onaylanan lisansüstü tezimin tamamını veya herhangi bir kısmını, basılı (kâğıt) ve elektronik formatta arşivleme ve aşağıda verilen koşullarla kullanıma açma iznini Hacettepe üniversitesine verdiğimi bildiririm. Bu izinle Üniversiteye verilen kullanım hakları dışındaki tüm fikri mülkiyet haklarım bende kalacak, tezimin tamamının ya da bir bölümünün gelecekteki çalışmalarda (makale, kitap, lisans ve patent vb.) kullanım hakları bana ait olacaktır.

Tezin kendi orijinal çalışmam olduğunu, başkalarının haklarını ihlal etmediğimi ve tezimin tek yetkili sahibi olduğumu beyan ve taahhüt ederim. Tezimde yer alan telif hakkı bulunan ve sahiplerinden yazılı izin alınarak kullanması zorunlu metinlerin yazılı izin alarak kullandığımı ve istenildiğinde suretlerini Üniversiteye teslim etmeyi taahhüt ederim.

Yükseköğretim Kurulu tarafından yayınlanan “ *Lisansüstü Tezlerin Elektronik Ortamda Toplanması, Düzenlenmesi ve Erişime Açılmasına İlişkin Yönerge* ” kapsamında tezim aşağıda belirtilen koşullar haricince YÖK Ulusal Tez Merkezi / H. Ü. Kütüphaneleri Açık Erişim Sisteminde erişime açılır.

Enstitü / Fakülte yönetim kurulu kararı ile tezimin erişime açılması mezuniyet tarihimden itibaren iki yıl ertelenmiştir.

Enstitü / Fakülte yönetim kurulu gerekçeli kararı ile tezimin erişime açılması mezuniyet tarihimden itibaren... ay ertelenmiştir.

Tezim ile ilgili gizlilik kararı verilmiştir.

22.10.2019

(imza)  
Ayşe KÜÇÜK

## ÖZET

### DOĞRUSAL REGRESYONDA RIDGE, LIU VE LASSO TAHMİN EDİCİLERİ ÜZERİNE BİR ÇALIŞMA

Ayşe KÜÇÜK

Yüksek Lisans, İstatistik Bölümü

Tez Danışmanı: Prof. Dr. Meral ÇETİN

Haziran 2019, 45 sayfa

Çoklu doğrusal regresyon analizinde, çoklu bağlantı sorununun çözümü için Ridge ve Liu regresyon kestiricisi sıklıkla kullanılır. Bu tez çalışmasında doğrusal regresyonda çoklu bağlantı durumunda literatürde en çok kullanılan yanlı tahmin edicilerden Ridge, Liu yöntemlerinin yanı sıra LASSO yöntemide ayrıntılı biçimde ele alınmıştır. Ayrıca bu yöntemlerde sağlam(robust) parametre kestirim değerleri kullanılarak klasik yöntemlerle karşılaştırılmaları da yapılmıştır. Çoklu bağlantı ve aykırıdeğer sorununun birlikte görüldüğü sayısal örnekler üzerinden yöntemlerin HKO'ları karşılaştırılmış ve sonuçlar yorumlanmıştır.

**Anahtar Kelimeler:** Regresyon Analizi, Çoklu Bağlantı, Ridge, Liu, LASSO,

## **ABSTRACT**

### **A STUDY ON RIDGE, LIU AND LASSO ESTIMATOR IN LINEAR REGRESSION**

**Ayşe KÜÇÜK**

**Master of Science, Department of Statistics**

**Supervisor: Prof. Dr. Meral ÇETİN**

**June 2019, 45 pages**

In multiple linear regression analysis, the Ridge and Liu regression estimators are often used to solve the multicollinearity problem. In this thesis, Ridge, Liu methods and LASSO method, which is one of the most widely used biased estimators in the literature in case of multiple connections in linear regression, are discussed in detail. In addition, robust parameter estimation values were compared with classical methods in these methods. The MSE of the methods were compared and the results were interpreted through numerical examples where multicollinearity and outlier problems were seen together.

**Keywords:** Regression Analysis, Multicollinearity, Ridge, Liu , LASSO,

## **TEŐEKKÜR**

Bu tezi hazırlarken ihtiyacım olduĐu her anda bana yol gĐsteren deĐerli danıŐmanım Sayın Prof. Dr. Meral ETİN hocama, sevgilerini ve inanlarını benden eksik etmeyen ve bugünlere ulaŐmamda büyük pay sahibi olan aileme, öĐretmenlerime, arkadaşlarıma ve Paku' ya teŐekkürü bir bor bilirim.

**AyŐe KÜÜK**

**Haziran 2019, Ankara**

## İÇİNDEKİLER

ÖZET .....	i
ABSTRACT .....	ii
TEŞEKKÜR .....	iii
İÇİNDEKİLER.....	iv
ÇİZELGELER DİZİNİ.....	vi
SİMGELER VE KISALTMALAR DİZİNİ.....	vii
1. GİRİŞ .....	1
2. GENEL BİLGİLER.....	4
2.1. Çoklu Bağlantı Sorunu.....	5
3. RIDGE TAHMİN EDİCİSİ .....	6
3.1. Ridge Tahmin Edicisinin Hata Kare Ortalaması ve k Değerleri.....	7
3.2. Ridge Tahmin Edicisinde k'nın Seçimi .....	9
3.2.1. Ridge İzi Yöntemi .....	9
3.2.2. 'k'nın Seçimi için Kullanılan Analitik Yöntemler .....	10
3.3. Ridge Tahmin Edicisinin Risk Fonksiyonu .....	14
3.4. Ridge Tahmin Edicisinin Kullanım Alanları .....	15
4. LIU TAHMİN EDİCİSİ .....	17
4.1. Liu Tahmin Edicisinde Yanlılık, Beklenen Değer, Varyans ve Hata Kare Ortalaması.....	18
4.2. Liu Tahmin Edicisinin Risk Fonksiyonu .....	19
4.2. Liu Tahmin Edicisinde Yanlılık Parametresi.....	19
4.2. Sağlam Liu Tahmin Edicisi .....	22
5. LASSO TAHMİN EDİCİSİ .....	24
5.1. Modified-Newton-Raphson (M-N-R) Algoritması ( $\gamma > 1$ ).....	25
5.2. Atış Algoritması.....	25
5.3. IRLS Yöntemiyle Algoritma .....	26
5.4. LASSO'nun Standart Hatası .....	26
5.5. LASSO'nun Risk Fonksiyonu .....	27
5.6. LASSO'nun Kullanım Alanları .....	27
6. UYGULAMA.....	29



6.1.	Hald'ın Verileri ile Uygulama .....	29
6.1.1.	Çoklu Bağlantı İncelemesi .....	29
6.1.2.	Ridge Tahmin Edicisinden Elde Edilen Sonuçlar .....	31
6.1.3.	Liu Tahmin Edicisinden Elde Edilen Sonuçlar .....	32
6.1.4.	LASSO Tahmin Edicisinden Elde Edilen Sonuçlar .....	34
6.2.	İstihdam Verileri ile Uygulama .....	36
6.2.1	Çoklu Bağlantı İncelemesi .....	36
6.2.2	Ridge Tahmin Edicisinden Elde Edilen Sonuçlar .....	38
6.2.3	Liu Tahmin Edicisinden Elde Edilen Sonuçlar .....	39
6.2.4	LASSO Tahmin Edicisinden Elde Edilen Sonuçlar .....	40
7.	SONUÇ .....	42
8.	KAYNAKLAR.....	43
	ÖZGEÇMİŞ.....	45

## ÇİZELGELER DİZİNİ

	<b>Sayfa No</b>
<b>Çizelge 1.</b> Hald veri kümesi .....	29
<b>Çizelge 2.</b> İlişki matrisi .....	30
<b>Çizelge 3.</b> Hald'ın verisi için özdeğerler ve koşul sayısı .....	30
<b>Çizelge 4.</b> Hald'ın verisi için özvektörler matrisi .....	30
<b>Çizelge 5.</b> Ridge tahmin edicisi parametre kestirim değerleri .....	31
<b>Çizelge 6.</b> Ridge tahmin edicisinden elde edilen sonuçlar .....	32
<b>Çizelge 7.</b> Aykırı değer varlığında Ridge tahmin edicisi .....	32
<b>Çizelge 8.</b> Liu parametre kestirim değerleri .....	33
<b>Çizelge 9.</b> Liu tahmin edicisine ilişkin parametre kestirimi ve HKO'ları.....	33
<b>Çizelge 10.</b> Aykırı değer varlığında Liu tahmin edicisi .....	34
<b>Çizelge 11.</b> LASSO parametre kestirim değerleri ve HKO'ları.....	34
<b>Çizelge 12.</b> Genel Sonuçlar .....	35
<b>Çizelge 13.</b> Aykırı değer varlığındaki sonuçlar.....	35
<b>Çizelge 14.</b> İstihdam veri kümesi .....	36
<b>Çizelge 15.</b> İlişki matrisi .....	37
<b>Çizelge 16.</b> İstihdam verisi için özdeğerler ve koşul sayısı .....	37
<b>Çizelge 17.</b> İstihdam verisi için özvektörler matrisi.....	38
<b>Çizelge 18.</b> Ridge parametre kestirim değerleri .....	38
<b>Çizelge 19.</b> Ridge tahmin edicisinden elde edilen sonuçlar .....	39
<b>Çizelge 20.</b> Liu parametre kestirim değerleri .....	39
<b>Çizelge 21.</b> Liu tahmin edicisine ilişkin parametre kestirimi ve HKO'ları.....	40
<b>Çizelge 22.</b> LASSO parametre kestirim değerleri ve HKO'sı .....	40
<b>Çizelge 23 .</b> Genel Sonuçlar .....	41

## SİMGELER VE KISALTMALAR DİZİNİ

EKK	En küçük kareler
HKO	Hata kareler ortalaması
HKT	Hata kareler toplamı
$\hat{\beta}$	En küçük kareler tahmin edicisi
$\hat{\beta}_k$	Ridge tahmin edicisi
$\hat{\beta}_d$	Liu tahmin edicisi
LASSO	Least absolute shrinkage and selection operator
k	Ridge tahmin edicisinin yan parametresi
$\alpha_{(p \times 1)}$	Tahmin edilecek parametre vektörü
$y_{(n \times 1)}$	Bağımlı değişken vektörü
$\varepsilon_{(n \times 1)}$	Hata vektörü
p	Bağımsız değişken sayısı
$\lambda_j$	X'X korelasyon matrisinin öz değerleri (j=1,2, ..., p)
d	Liu tahmin edicisinin yan parametresi

# 1. GİRİŞ

Doğrusal modellerin temelinde regresyon modelleri yer almaktadır. Regresyon sözcüğü "geriye doğru gidiş" ya da "geriye dönüş" olarak tanımlanmaktadır. Bu sözcük ilk kez 1885 yılında Sir Francis Galton (1822-1911) tarafından kullanılmıştır. U. Yule ve K. Pearson bu tekniği geliştirmişler ve böylece literatür de regresyon kullanılmaya başlanmıştır. Regresyon denklemlerinde bağımlı değişken ve bağımsız değişkenle birlikte bilinmeyen regresyon katsayıları da yer almaktadır. Regresyon analizinin amacı regresyon katsayılarının değerlerini tahmin etmek ve bir modele uydurmaktır. Regresyon analizinin birçok kullanım alanı vardır. Bunlardan bazıları; psikoloji, sosyoloji, meteoroloji, tıp, mühendislik, eğitim, spor, fizik, kimya, iktisat, işletme ve biyolojidir. Regresyon denkleminin derecesinin bir olduğu duruma doğrusal regresyon denklemi denilmektedir. Doğrusal modellerde bir ya da daha çok sayıda bağımlı değişken bağımsız değişken ile doğrusal olarak ilişkilendirilir. Bağımlı ve bağımsız değişkenin tek olduğu doğrusal denklemlere basit doğrusal regresyon modeli denilmektedir. Genellikle birden fazla bağımsız değişkenle açıklanır ve bu duruma çoklu doğrusal regresyon modeli denir. Çoklu doğrusal regresyon modelleri genel olarak Eşitlik (1.1)'de verildiği gibidir (Erar, 2017).

$$Y = X\beta + \varepsilon, \quad (1.1)$$

Burada  $Y$ ,  $(n \times 1)$  boyutlu rastlantı değişkenlerinin gözlenen vektörü,  $X$ ,  $(n \times p)$  boyutlu bağımsız değişkenlerin gözlem değerini içeren matrisi göstermektedir.  $\beta$   $(p \times 1)$  boyutlu bilinmeyen katsayılar vektörü ve  $\varepsilon$ ,  $(n \times 1)$  boyutlu hata terimi vektörüdür. Hataların dağılımı  $E(\varepsilon) = 0$  ortalama ve  $\sigma^2 I_n$  varyans ile normaldir.  $Y$ ' nin doğrusal olduğu durumda  $\beta$ ' nin en küçük kareler (EKK) tahmin edicisi aşağıdaki gibi gösterilmektedir;

$$\hat{\beta} = (X'X)^{-1}X'Y \quad (1.2)$$

EKK hata kareler toplamını en küçükmeyi hedefleyen yöntemdir. Bu yöntemin bazı varsayımları sağladığı durumlarda güvenilir tahminler elde edilir. Bu varsayımlar; gözlemlenen verilerin normal olması, sabit varyanslılık durumunu sağlamasıdır. Varsayımların sağlamaması durumunda elde edilen tahminlerin hata olasılığı yüksek olur (Aydın, 2014).

Doğrusal regresyon analizinde varsayımlardan birisi de bağımsız değişkenlerin kendi arasında ilişkili olmamasıdır. Bu ilişkinin olması sonucu çoklu bağlantı sorunu (multicollinearity) ortaya çıkar. Bu sorunla birlikte tahminler kararsız olabilir, model parametreleri için gerçekte olması gerekenden farklı işaret ve tahminler bulunabilir, parametre varyansları olduğundan büyük çıkabilir, verideki çok küçük değişiklikler parametrelerdeki işaretleri olduğundan farklı gösterebilir. Bu sorunu gidermek için modelden birden çok bağımsız değişken modele alınmayabilir. Bazı durumlarda yeni gözlemlerin veri setine eklenmesiyle çoklu bağlantı sorunu giderilmiş olur. Ya da aralarında ilişki olan değişkenlerin yerine bu değişkenlerin toplamı modele alınabilir, değişkenler dönüştürülebilir, ya da EKK tahmin edicilerine göre daha küçük hata kareler ortalamasına (HKO) sahip olan tahmin ediciler; Ridge, Liu ve LASSO tahmin edicileri kullanılabilir.

Hoerl ve Kennard (1970) çoklu bağlantı durumunun üstesinden gelmek amacıyla yanlı tahmin edici yöntemlerinden ilki olan Ridge tahmin edicisini önermişlerdir. Ridge tahmin edicisi aşağıdaki gibi tanımlanmıştır (Akdeniz ve Kaçıranlar, 1995).

$$\hat{\beta}_k = (X'X + kI_p)^{-1}X'y, k \geq 0 \quad (1.3)$$

Burada  $k$  değeri EKK' ya göre daha küçük HKO verir ve Ridge parametresi ya da yan parametresi olarak adlandırılır (Arthur ve Kennard, 1970). Ridge tahmin edicisinin yorumlanması neticesinde çoklu bağlantı problemine yeni bir yanlı tahmin edici 1993 yılında Liu tarafından önerilmiştir. Liu tahmin edicisi aşağıdaki gibi tanımlanmıştır:

$$\hat{\beta}_d = (X'X + I_p)^{-1}(X'y + d\hat{\beta}), 0 < d < 1 \quad (1.4)$$

Belirlenen aralıkta  $d$  değerinin aldığı herhangi bir değer EKK yöntemiyle kıyasla daha küçük HKO ortalaması vermektedir.  $d$  parametresi büzülme parametresi olarak adlandırılmaktadır. EKK, Ridge tahmin edicisi ve Liu tahmin edicisi varyans büyüklüğü ve tahmin edici güvenilirliğinin azlığı gibi nedenlerle tam anlamıyla yeterli olmamıştır. Diğer yandan Ridge ve Liu tahmin edicileri, katsayıları azaltan ve böylece daha kararlı olan sürekli süreçlerdir; bununla birlikte, modelin yorumlanması sorunu, hataların hiçbirini 0 olarak belirlemediğinden hala çözümsüz kalmaktadır. Bu durum daha güvenilir sonuçlar elde etmek adına farklı ve yeni tahmin edicilerin doğmasını ihtiyaç kılmıştır. Tibshirani (1996) 'least absolute shrinkage and selection operator'

(LASSO) adında yeni bir yöntem önermiştir. Bu teknik artık kareleri minimum yapan bir operatördür. Bu tez çalışmasında, Ridge, Liu ve LASSO yöntemleri ayrıntılı biçimde verilmiş, ayrıca yöntemlerin karşılaştırmaları yapılmıştır. Bu amaçla tezin ikinci bölümünde genel bilgiler verilmiş, sırasıyla üçüncü, dördüncü ve beşinci bölümlerde Ridge, Liu ve LASSO tahmin edicileri ve risk fonksiyonları ayrıntılı biçimde anlatılmıştır. Altıncı bölümde ise değişik veri kümeleri üzerinden uygulamalar yapılmış ve sonuçları tartışılmıştır.

## 2. GENEL BİLGİLER

Bir çoklu doğrusal regresyon modelinin iyi bir tahmininin elde edilebilmesi için sağlanması gereken varsayımlar vardır. Bu varsayımlar; modelin şekli hakkındaki varsayımlar, hatalar hakkındaki varsayımlar, bağımsız değişkenler hakkındaki varsayımlar ve gözlemler hakkındaki varsayımlar olarak sınıflandırılabilir. Bu bölümde çoklu doğrusal regresyon, çoklu bağlantı sorunu ve bu sorunun giderilmesi için gerekli olan çözümlerden bahsedilecektir (Erar, 2017).

Çoklu doğrusal regresyon birden fazla bağımsız değişkenin bir bağımsız değişkeni açıkladığı durumlarda kullanılan istatistiksel yöntemlerden bir tanesidir. Değişken sayısının fazla olması durumu bağımlı değişkenin daha iyi açıklanmasını sağladığı için çoklu doğrusal regresyonun yaygın kullanımı söz konusudur Eşitlik (1.1)'de verilen doğrusal regresyon modeli daha genel olarak,

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_p X_{ip} + \varepsilon_i, i = 1, \dots, n \quad (2.1)$$

biçiminde gösterilir (Shah, 2015; Aydın, 2014).

Regresyon denkleminde amaç  $Y$ 'ye en yakın değerleri veren kestirim değerini elde edebilmek için model parametrelerinin en iyi tahminlerinin bulmaktır. Bu nedenle regresyon katsayılarının tahmininde genelde EKK yöntemi kullanılmaktadır. Örneğin örneklem regresyon eğrisi  $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i$  değerlerini verdiği için  $Y$  değerleri ile bu eğri arasındaki uzaklık hata terimlerini ( $\hat{e}_i$ ) vermektedir. EKK yöntemi bu HKT'nı ( $\sum e_i^2$ ) minimum yapacak tahminlerini hesaplar. Burada hata terimlerinin karelerinin modele girmemesini nedeni toplam alındığı zaman negatif ve pozitif değerlerin birbirini götürmesine engel olmaktır. EKK yönteminde temel olarak;

$$f(\hat{\beta}) = \sum_{i=1}^n e_i^2 = e'e = (Y - X\hat{\beta})'(Y - X\hat{\beta})$$

$$\hat{\beta} \in R^p$$

parametre kümesi üzerinden minimize eden değerleri bulmaktır. Burada,

$$\frac{\partial f(\hat{\beta})}{\partial \hat{\beta}} = -2X'Y + 2X'X\hat{\beta}$$

olmak üzere, ifadesini sağlayan  $\hat{\beta}$  vektörü  $\beta'$  nın EKK tahmin edicisidir.

$$\hat{\beta} = (X'X)^{-1}X'Y \quad (2.2)$$

şeklinde parametre tahmini elde edilir (Aydın, 2014).

### 2.1. Çoklu Bağlantı Sorunu

Çoklu doğrusal regresyon modelinin sağlanması gereken varsayımlardan biri, modelde yer alan bağımsız değişkenler arasında ilişki olmaması varsayımdır. Varsayımın sağlanmadığı durumlarda çoklu bağlantı sorunu meydana gelir. Bu değişkenler arasında doğrusal bir ilişki varsa gerçekte olandan oldukça farklı kestirimler meydana gelebilir. Bununla birlikte çoklu bağlantı, regresyon katsayılarının standart hata kestirimleri ve bununla birlikte de elde edilen  $t$  istatistiklerinin farklı çıkmasıyla sonuçlanabilir ve hesaplanan  $R^2$  çok büyük çıkabilir (Erar, 2017).

Çoklu bağlantıyı belirleyebilmek için birçok yöntem vardır. Bunlardan bazıları korelasyon matrisinin incelenmesi, varyans şişirme faktörleri (variance inflation factor), öz değerler ve öz vektör analizidir. Bu yöntemler sayesinde sorunun boyutunu ve hangi değişkenden kaynaklandığını görebiliriz (Aydın, 2014).

Çoklu bağlantının varlığında, regresyon katsayıları hem sayısal hem de işaretçe olduğundan farklıdır, katsayılarının varyansları olduğundan daha büyüktür ayrıca  $t$ -istatistiği, güven aralığı ve  $R^2$  de oldukça etkilenir (Erar, 2017).

Katsayı kestirimlerinin standart hatasını azaltmak ve tutarlı sonuçlar elde etmek için çoklu bağlantının ortadan kaldırılması gerekmektedir. Bu durumda önerilen yöntemlerden birkaçı yukarıda anlatılmış ve bir başka yöntem olan yanlı regresyon yöntemlerinin kullanılması ve bu yöntemlerin en çok kullanılanlardan Ridge, Liu ve LASSO yöntemleri bu tezde ele alınmıştır (Aydın, 2014).



### 3. RIDGE TAHMİN EDİCİSİ

EKK'nın yansız  $\hat{\beta}$  tahmin ediciler vermesi ve tüm doğrusal yansız  $\hat{\beta}$  tahmin edicilerinin en küçük varyanslı olması istenen bir özellik olmasına rağmen, bu tahmin edicilerin varyansının bir üst sınırı olmamaktadır ve çoklu bağlantı olması durumunda olması gerekenden daha büyük varyanslar ortaya çıkabilmektedir (Aydın, 2014). Büyük varyans oluşumu sorununu azaltmanın bir yolu,  $\hat{\beta}$  tahmin edicisinin yansız olma gereksinimini ortadan kaldıran yöntemlerden olan Ridge tahmin edicisidir.

Hoerl ve Kennard (1970) önerdikleri Ridge tahmin edicisi ile EKK yöntemiyle yapılan tahminlerden daha küçük HKO'sına sahip regresyon katsayılarını hesaplamışlardır. Ridge tahmin edicisi normal denklemlerin hafifçe değişiklikler yapılmış şekli çözülerek bulunmaktadır. Çoklu doğrusal regresyon modeli matris formunda aşağıdaki gibi verilmiş olsun:

$$Y = X\beta + \varepsilon \quad (3.1)$$

Burada  $Y$ ,  $(n \times 1)$  boyutlu bağımlı değişken,  $X$ ,  $(n \times p)$  boyutlu bağımsız değişkenler matrisi,  $\beta$  bilinmeyen katsayılar vektörü ve  $\varepsilon$   $(n \times 1)$  boyutlu hata terimi vektörüdür. Hataların dağılımı 0 ortalama ve  $\sigma^2 I_n$  varyans ile normaldir.

Regresyon modelinde tüm değişkenler standartlaştırıldığında  $X'X$  matrisi korelasyon matrisi formundadır. Bu modelde  $\beta$ 'nin EKK tahmin edicisi aşağıdaki gibidir:

$$\hat{\beta} = (X'X)^{-1}X'y \quad (3.2)$$

Bağımsız değişkenler arasında doğrusal ilişki olduğu durumda  $\beta$ 'nin EKK tahminlerinin güvenilirliği azalır. Bu durumda aşağıda Ridge tahmin edicisi önerilir:

$$\hat{\beta}_k = (X'X + kI_p)^{-1}X'y, \quad k \geq 0 \quad (3.3)$$

Burada  $k$  Ridge parametresi,  $I_p$  ise birim matrisidir.  $k = 0$  olduğunda Ridge tahmin edicisi EKK'ye dönüşür. Burada verilen  $k$  yan miktarı olarak adlandırılır.  $k > 0$  olduğunda ise  $X'$  in köşegen elemanlarına negatif olmayan küçük  $k$  değeri eklemiş olur ve dolayısıyla çoklu bağlantı etkisi kaldırılmış olur (Neter, 1996; Fox, 1997). Ridge

tahmin edicisinin HKO' sının EKK'nın HKO'sından daha küçük olması için küçük bir  $k$  seçilmelidir ve bunun için önerilen denklem aşağıda verilmiştir:

$$k = p\sigma^2/\beta'\beta, \quad (3.4)$$

Burada  $k$ , 1' den küçük 0'dan büyük bir değerdir. Bu tahminin beklenen değeri aşağıdaki gibidir;

$$E(\hat{\beta} - \beta) = [(X'X + kI)^{-1}X'X - I]\beta \quad (3.5)$$

ve kovaryans matrisi de aşağıdaki gibidir:

$$V(\hat{\beta}) = (X'X + kI)^{-1} X'X(X'X + kI)^{-1} \quad (3.6)$$

En uygun  $k$  değerini elde etmek için yanlı standartlaştırılan regresyon katsayıları ve  $k$  arasında hesaplanan ve Ridge Trace olarak adlandırılan grafiklerden yararlanılmaktadır (Arthur ve Kennard, 1970). Optimal  $k$  değeri, yanlı standartlaştırılmış regresyon katsayılarının durgun hale geldiği bölgelerden alınmaktadır (Orhunbilge ve Albayrak, 2000).

### 3.1. Ridge Tahmin Edicisinin Hata Kare Ortalaması ve $k$ Değerleri

HKO'nı,  $\hat{\beta}$ 'nin  $\beta$ 'ya olan uzaklığının bir fonksiyonu olarak tanımlayalım. Uzaklığı  $HKO(\hat{\beta})$  olarak gösterirsek, Ridge tahmin edicisinin HKO'sı,

$$\begin{aligned} HKO(\hat{\beta}) &= E((\hat{\beta} - \beta) (\hat{\beta} - \beta)') \\ &= \sigma^2 \sum_{j=1}^p \frac{\lambda_j}{(\lambda_j + k)^2} + \sum_{j=1}^p \frac{\alpha_j^2 k_j^2}{(\lambda_j + k)^2} \end{aligned} \quad (3.7)$$

olarak yazılır. Burada  $\lambda_1, \lambda_2, \dots, \lambda_p$ ,  $X'X$ 'in özdeğerleridir. Bu denklemin sağ tarafında yer alan ilk terim  $\hat{\beta}$ 'deki parametrelerin varyansları toplamıdır ve ikinci terim de yanın karesidir. Eğer  $k > 0$  ise  $\hat{\beta}$ 'deki yan  $k$  ile artmaktadır. Ancak  $k$  azaldıkça varyans da azalmaktadır.

Ridge tahmin edicisinin kullanımında  $k$ 'nın değeri, varyans terimindeki azalmanın kareli yandaki artıştan daha az olacak biçimde seçilmeye çalışılır (Montgomery, Peck ve Vining, 2012).

$k_j$  değerlerinin;

$$k_j = \frac{\sigma^2}{\alpha_j^2}, j = 1, 2, \dots, p \quad (3.8)$$

olarak seçilmesi HKO( $\hat{\beta}$ ) değerini minimize etmektedir. Ancak bu değer  $\sigma^2$  ve  $\alpha_j$  parametrelerine bağlı olarak değişmektedir.  $k_j$  değerlerini optimal olarak seçmek için Hoerl ve Kennard (1970) tarafından iteratif yöntemler önerilmiştir. Bu yöntemde öncelikle;

$$k_j^0 = \frac{\hat{\sigma}^2}{\hat{\alpha}_j^2}, j = 1, 2, \dots, p \quad (3.9)$$

başlangıç değeri alınır.  $k_j^0$  başlangıç değeri ile elde edilen ilk geliştirilmiş Ridge tahmin edicisi,

$$\hat{\alpha}_r^0 = (\Lambda + K^0)^{-1}Z \quad (3.10)$$

şeklinde iken,  $k^0 = \text{diag}(k_1^0, k_2^0, \dots, k_p^0)$  şeklinde köşegen bir matristir. Başlangıç tahmincisi aracılığıyla, optimal  $k_j$  değerleri,

$$k_j^1 = \frac{\hat{\sigma}^2}{(\hat{\alpha}_j^0)^2}, j = 1, 2, \dots, p \quad (3.11)$$

$$\begin{array}{ccc} : & & : \\ \cdot & & \cdot \\ : & & : \end{array}$$

şeklinde bulunur. Bu iteratif işleme  $\hat{\alpha}_r$  vektör uzunluğunun karesi,  $\hat{\alpha}_r' \hat{\alpha}_r$  değerinin i. ve (i-1). adımda aldığı değerler arasında anlamlı fark kalmayınca kadar devam edilmelidir (Fidanoğlu, 2009).

Hoerl ve Kennard (1970 a, b), Hoerl, Kennard, Baldwin (1975) ve Lawless ve Wang (1976)  $k$  için sırasıyla aşağıdaki formülasyonları önermişlerdir (Sakallıoğlu, Kaçıranlar ve Akdeniz, 2001).

$$\hat{k}_{HK} = \frac{\hat{\sigma}^2}{\sum_{i=1}^p \hat{\alpha}_i^2}$$

$$\hat{k}_{HKB} = \frac{p\hat{\sigma}^2}{\sum_{i=1}^p \hat{\alpha}_i^2}$$

$$\hat{k}_{LW} = \frac{p\hat{\sigma}^2}{\sum_{i=1}^p \lambda_i \hat{\alpha}_i^2}$$

Troskie ve Chalton (1996) tarafından geliştirilmiş Ridge regresyonda optimal  $k_j$ 'lerin seçimi için; n gözlem sayısı, p değişken sayısı,  $\hat{\sigma}^2$  ve  $\hat{\beta}$ ,  $\sigma^2$  ve  $\beta$ 'nin EKK tahmincileri olmak üzere önerilen  $k$  değeri;

$$\hat{k}_j = \frac{\lambda_j \hat{\sigma}^2}{(n-p)\beta_j^2 + \sigma^2} \quad (3.12)$$

olarak tanımlamıştır. Firinqwtti (1999) ise,

$$\hat{k}_j = \frac{\lambda_j \hat{\sigma}^2}{(n-p)\hat{\sigma}^2 + \lambda_j \hat{\beta}_j^2} \quad (3.13)$$

formülünü önermiştir (Karakaş, 2008).

### 3.2. Ridge Tahmin Edicisinde $k$ 'nın Seçimi

Ridge tahmin edicisinde yanlılık parametresi olan  $k$ ,  $0 < k < 1$  arasında değer almaktadır. Yanlılık parametresi olan  $k$ 'nın seçimi büyük önem taşımaktadır ve bunun için birçok yöntem geliştirilmiştir. Bu yöntemler Ridge izi ve analitik yöntemler olarak ikiye ayrılmaktadır.

#### 3.2.1. Ridge İzi Yöntemi

Ridge izi yöntemi, Ridge tahmin edicisinin yanlılık parametresi  $k$ 'nın seçimi için kullanılan grafiksel bir yöntemdir. Ridge tahmin edicisinin en büyük avantajı parametrelerden hangisinin verilerdeki değişimler karşısında hassasiyet gösterdiğini anlamayı amaçlayan “Ridge İzi” yöntemini sağlamasıdır. Ridge izi, hangi parametrenin, verideki küçük değişimlerden etkilendiğini belirttiği için hassasiyet analizi olarak da tanımlanmıştır (Akdeniz ve Kaçiranlar, 2001).

Ridge izi yönteminde Ridge katsayıları  $\hat{\beta}_R$ 'ler düşey ekseninde,  $k$  değerleri  $k \geq 0$  şeklinde yatay ekseninde alınır ve iki boyutlu kartezyen düzlemde Ridge izi grafiği elde edilir. Çizim yapılırken karışıklık olmaması için 10'dan fazla  $k$  değeri gösterilmemelidir.

Grafikte gösterilen  $k$  değerleri arttıkça  $\hat{\beta}_R$ 'lerin HKO'sı azalır. Amaç, EKK yöntemiyle elde edilen regresyon katsayılarının HKO'dan daha küçük değerler veren kararlı Ridge tahmin edicisini bulmak olduğundan, katsayıların dengeye geldiği,  $k$ 'daki artışlara  $\hat{\beta}_R$ 'lerdeki değişimin çok yavaşladığı ve yatay eksene paralel bir konuma ulaştığı noktaya karşılık gelen  $k$  değeri Ridge parametresinin değeri olarak seçilir (Arthur ve Kennard, 1970).

### 3.2.2. 'k'nın Seçimi için Kullanılan Analitik Yöntemler

Ridge tahmin edicisinin yanlılık parametresi  $k$ 'nın seçimi için grafiksel yöntemlerin yanı sıra HKO'ları minimum yapma esasına dayalı analitik yöntemler önerilmiştir. Bu yöntemler aşağıda verilmiştir.

- **Hoerl, Kennard ve Baldwin'in Tahmin Edicisi**

Hoerl, Kennard ve Baldwin 1975 yılında Ridge tahmin edicisinin yanlılık parametresini;

$$\hat{k}_{HKB} = \frac{p \hat{\sigma}^2}{\hat{\beta}'\hat{\beta}} \quad (3.14)$$

biçiminde önermişlerdir.  $\hat{\sigma}^2$  ve  $\hat{\beta}$ , EKK tahmin değerleri;  $p$  bağımsız değişken sayısıdır.  $p$ 'nin sabit değeri için  $X'X$ 'in öz değerlerinde değişkenlik artarken,  $k$  parametresinin kullanımı sonucu daha küçük HKO elde edilebilir.

- **Mallows Tahmin Edicisi**

Mallows (1973) ,  $k$ 'nın seçiminde, değişken seçimi kriteri olarak önerdiği  $C_p$  istatistiğinde birtakım değişiklikler yaparak  $C_k$  istatistiğini geliştirmiştir:

$$C_k = \frac{HKT(k)}{\hat{\sigma}^2} - n + 2 + 2iz( X (X'X + kI)^{-1}X' ) \quad (3.15)$$

Eşitlik (3.15)'in minimize edilmesi sonucu elde edilen optimal  $k$  değerinin kullanımını önermiştir.

- **Lawless ve Wang Tahmin Edicisi**

Lawless ve Wang (1976)  $k$ 'nın hesaplanması için, Ridge tahmin edicisinin kanonik biçimini kullanmıştır. Regresyon modeli;

$$y = Z\alpha + \varepsilon \quad (3.16)$$

olarak verilir. Burada  $Z = XQ$  ve  $\alpha = Q'\beta$  olmak üzere;  $Q$  ortogonal matris,  $\Lambda$  öz değerleri 0 olmayan köşegen matristir ve  $\Lambda = Q'X'XQ$  dönüşümü sağlanmaktadır.  $Z_{(n \times p)}$  matrisi  $Z'Z = \Lambda$  olarak tanımlanmakta,  $p$  bağımsız değişken sayısını göstermek üzere,  $\alpha_{(p \times 1)}$ , tahmin edilecek parametre vektörü,  $y_{(n \times 1)}$ , bağımlı değişken vektörü ve  $\varepsilon_{(n \times 1)}$ , hata vektörüdür. Kanonik formdaki regresyon modelinde  $\alpha$ 'nın Ridge tahmin edicisi

$$\hat{\alpha}_{(k)} = (\Lambda + kI)^{-1}Z'y \quad (3.17)$$

$$\hat{\alpha}_{j(k)} = \frac{\lambda_j}{\lambda_j + k} \hat{\alpha}_j, \quad j = 1, 2, \dots, p \quad (3.18)$$

olarak hesaplanır.  $\lambda_j$ ,  $X'X$  korelasyon matrisinin öz değerleridir. Lawless ve Wang (1976) Ridge tahmin edicisi ile optimal yanlılık parametresi  $k$ 'nin seçimi için aşağıdaki eşitliği önermişlerdir:

$$\hat{k}_{LW} = \frac{p\hat{\sigma}^2}{\sum_{j=1}^p \lambda_j \hat{\alpha}_j^2} \quad (3.19)$$

Burada  $\hat{\sigma}^2$ ,  $\sigma^2$ 'nin EKK tahmin edicisidir.

#### ▪ Hocking, Speed ve Lynn Tahmin Edicisi

Hocking, Speed ve Lynn 1976 yılında ridge regresyonun yanlılık parametresi olan  $k$ 'yı;

$$k_{HSL} = \sum_{j=1}^p \frac{\lambda_j \tau_j^2}{\sum_{j=1}^p \tau_j^4} \quad (3.20)$$

olarak tanımlamıştır. Burada  $\tau_j^2$ ,

$$\tau_j^2 = \frac{\alpha_j^2 \lambda_j}{\sigma^2} \quad j = 1, 2, \dots, p \quad (3.21)$$

biçiminde tanımlanır. Eşitlik (3.20), (3.19) da yerine konulup,  $\sigma^2$  yerine EKK tahmin edicisi  $\hat{\sigma}^2$  ve  $\alpha$  değeri için (3.16) eşitliğinde verilen kanonik biçimli değeri kullanıldığında optimal Ridge parametresi aşağıdaki gibi yeniden elde edilir.

$$\hat{k}_{HSL} = \hat{\sigma}^2 \sum_{j=1}^p \frac{\lambda_j \hat{\alpha}_j^2}{\sum_{j=1}^p \lambda_j^2 \hat{\alpha}_j^4} \quad (3.22)$$

#### ▪ Demster, Schatzoff ve Wermuth Tahmin Edicisi

Demster, Wermuth ve Schatzoff,  $k$  parametresinin seçimi için doğrusal olmayan bir form önermişlerdir.  $\hat{k}_{RIDGM}$  olarak isimlendirdikleri optimal  $k$  değeri,

$$\sum_{j=1}^p \left[ \frac{\hat{\alpha}_j^2}{\frac{1}{k} + \lambda_j} \right] = p\hat{\sigma}^2 \quad (3.23)$$

eşitliğinin çözümüdür.  $\hat{\sigma}^2$ ,  $\sigma^2$ 'nin EKK tahmin edicisi ve  $\hat{\alpha}$ ;  $\alpha$ 'nın kanonik parametreler cinsinden tahmin edicisidir. Gosling ve Puterman (1985), Hoerl (1975)'in önerdiği  $\hat{k}_{HKB}$  ile  $\hat{k}_{RIDGM}$  tahmin edicilerini karşılaştırmışlar ve  $\hat{k}_{RIDGM} > \hat{k}_{HKB}$  sonucunu elde etmişlerdir.

#### ▪ Kibria'nın Tahmin Edicisi

Kibria (2003), genelleştirilmiş Ridge tahmin edicisi yaklaşımına bağlı olarak, yeni bir tahmin edici önermiş, yeni tahmin edici ve mevcut Ridge tahmin edicisinin performansını ölçmek için simülasyon çalışmaları yapmıştır.

Daha önceki çalışmalarda bahsedilen, Hoerl, Kennard (1970), Hoerl (1975), Lawless, Wang (1976), Hocking (1976)'in önerdiği  $k$ 'larına ilaveten;  $k$ 'nın seçimi için, aritmetik ortalama, geometrik ortalama ve ortancayı esas alan yeni tahmin ediciler önermiştir.

Eşitlik ( 3.16 )'da verilen modele dayanarak  $Z = XQ$  ve  $\alpha = Q'\beta$  olmak üzere;  $Q$  ortogonal matris ve  $\Lambda$ , özdeğerleri 0 olmayan köşegen matristir ve  $\Lambda = Q'X'XQ$  dönüşümünü sağlasın.  $K$ , asal köşegen matrisinin elemanları  $k_1, k_2, \dots, k_p$  olmak üzere  $k_i > 0$  için,

$$\hat{\alpha}_{(k)} = (\Lambda + KI)^{-1}Z'y \quad (3.24)$$

şeklinde formüle edilmiş olup,  $\alpha$ 'nın EKK tahmin edicisi,  $\lambda_j$ ,  $X'X$  korelasyon matrisinin öz değerleri olmak üzere;

$$\hat{\alpha} = \Lambda^{-1}Z'y$$

ya da,

$$\hat{\alpha}_{j(k)} = \frac{\lambda_j}{\lambda_j + k} \hat{\alpha}_j, \quad j = 1, 2, \dots, p \quad (3.25)$$

biçiminde tanımlanır. Literatürde önerilen üç tahmin edici;

$$\hat{k}_{AO} = \frac{1}{p} \sum_{j=1}^p \frac{\hat{\sigma}^2}{\hat{\alpha}_j^2} \quad (3.26)$$

$$\hat{k}_{GO} = \frac{\hat{\sigma}^2}{(\prod_{j=1}^p \hat{\alpha}_j^2)^{1/p}} \quad (3.27)$$

$$\hat{k}_{MED} = \text{Medyan} \left( \frac{\hat{\sigma}^2}{\hat{\alpha}_j^2} \right), \quad j = 1, 2, \dots, p \quad (3.28)$$

biçimindedir. Kibria'nın çalışmasında, performans kriteri olarak HKO'sını minimize etmeye odaklanılmıştır.  $\hat{\sigma}^2$  ve sinyal gürültü oranı  $\rho^2$ , arasındaki ilişkiyi belirten;

$$\rho^2 = \frac{\beta'\beta}{\sigma^2} \quad (3.29)$$

formülasyonun farklı değerleri için tahmin ediciler hesaplanıp karşılaştırılmıştır. Gözlem sayısı(n), değişken sayısı(p), korelasyon değerleri ( $\gamma$ ) ve  $\frac{\lambda_1}{\lambda_p}$  oranları için, işaret gürültü oranı  $\rho^2$  büyük iken EKK tahmin edicisi her çoklu bağlantı derecesinde diğer tüm tahmin edicilerden daha iyi sonuç vermekte ama diğer tüm tahmincilerin HKO'su, en küçük kareler tahmin edicisinin HKO'dan daha küçüktür.  $\hat{k}_{GO}$  ve  $\hat{k}_{LW}$  tahmin edicileri eşit derecede sonuçlar verirken,  $\hat{k}_{HKB}$ 'den daha iyi sonuçlar vermektedir.

Küçük veya orta seviyedeki korelasyonlarda,  $\hat{k}_{GM}$ 'nin performansı  $\hat{k}_{LW}$ 'den daha iyi performans gösterdiği görülmüş ancak güçlü korelasyonlarda  $\hat{k}_{LW}$ ,  $\hat{k}_{GM}$  ve  $\hat{k}_{HKB}$ 'den daha iyi olduğu düşünülmüştür. Son olarak, daha önce çalışılan simülasyon

çalışmalarıyla birlikte  $\hat{k}_{GM}$  tahmin edicisinin  $\hat{k}_{HKB}$  tahmin edicisiyle eşit ve diğer tahmin edicilere oranla biraz daha yüksek performanslı olduğu görülmüş ve  $k$ 'nin seçimi için iyi bir tahmin edici olabileceği sonucuna varılmıştır.

▪ **Khalaf ve Shukur Tahmin Edicisi**

Khalaf ve Shukur (2005) çalışmasında, Hoerl ve Kennard (1975)'in çalışmasına alternatif bir tahmin edici geliştirmişlerdir ve simülasyon çalışması yapmışlardır. Yeni elde ettikleri tahmin ediciye KS adını vermişlerdir.  $n$  gözlem sayısı,  $p$  bağımsız değişken sayısı ve  $\lambda_{max}$   $X'X$  korelasyon matrisinin öz değerlerini göstermek üzere önerilen tahmin edici aşağıdaki gibidir:

$$k_{KS}^{**} = \frac{\lambda_{max}\hat{\sigma}^2}{(n-p-1)\hat{\sigma}^2 + \lambda_{max}\hat{\beta}_{max}^2} \quad (3.30)$$

Burada,  $\hat{\sigma}^2$  ve  $\hat{\beta}$ , EKK tahmin edicileridir. Hoerl ve Kennard (1975)'in önerdiği tahmin ediciye HK adını vermişlerdir. KS tahmin edicisi HK tahmin edicisi ile karşılaştırdıklarında küçük varyans durumunda ve küçük örneklerde HK ve KS tahmin edicilerinin EKK'den daha iyi sonuç verdiği görülmüş olup, KS'nin de EKK'dan daha iyi sonuçlar verdiği sonucuna varmışlardır.

▪ **Alkhamisi, Khalaf ve Shukur Tahmin Edicisi**

Alkhamisi, Khalaf ve Shukur (2006), ve Kibria (2003)'nin çalışmalarını, Khalaf ve Shukur (2005) tarafından önerilen KS tahmin edicisine uyarlanarak,  $k$ 'nin seçiminde dört tahmin edici önermişlerdir.  $K_{arithmetik}^{**}$ ,  $K_{medyan}^{**}$ ,  $K_{max}^{**}$ ,  $K_{max}^{HK}$  şeklinde isimlendirdikleri bu tahmin ediciler için,  $n$  gözlem sayısı,  $p$  bağımsız değişken sayısı,  $\lambda_j$ ,  $X'X$  korelasyon matrisinin öz değerleri ve  $\hat{\sigma}^2$ ,  $\hat{\beta}$  değerleri EKK tahmin edicisinden elde edilen değerler olmak üzere aşağıdaki gibi tanımlanmıştır:

$$K_{arithmetik}^{**} = \frac{1}{p} \sum_{j=1}^p \left( \frac{\lambda_j \hat{\sigma}^2}{\lambda_j \hat{\beta}_j^2 + (n-p)\hat{\sigma}^2} \right), \quad j = 1, 2, \dots, p \quad (3.31)$$

$$K_{medyan}^{**} = \text{medyan} \left( \frac{\lambda_j \hat{\sigma}^2}{\lambda_j \hat{\beta}_j^2 + (n-p)\hat{\sigma}^2} \right), \quad j = 1, 2, \dots, p \quad (3.32)$$

$$K_{max}^{**} = \max \left( \frac{\lambda_j \hat{\sigma}^2}{\lambda_j \hat{\beta}_j^2 + (n-p)\hat{\sigma}^2} \right), \quad j = 1, 2, \dots, p \quad (3.33)$$

$$K_{max}^{HK} = \max_j \left( \frac{\hat{\sigma}^2}{\hat{\beta}_j} \right), \quad j = 1, 2, \dots, p \quad (3.34)$$



$K_{HK}$ ,  $K_{KS}^{**}$  tahmin edicileri ile  $K_{arithmetik}^{**}$ ,  $K_{medyan}^{**}$ ,  $K_{max}^{**}$ ,  $K_{max}^{HK}$  tahmin edicileri HKO'na göre karşılaştıran geniş çaplı simülasyon çalışması yapmışlar ve  $K_{max}^{**}$ 'in iyi özellikler taşıdığını ortaya koymuşlardır.

- **Alkhamisi ve Shukur Tahmin Edicisi**

Alkhamisi ve Shukur (2007) daha önceki çalışmasının devamı niteliğinde  $k$ 'nın seçimi için yeni birkaç yöntem daha önerip, HKO'na karşılaştırılabilir monte carlo simülasyon çalışmaları yapmışlardır. Simülasyon çalışmaları sonucunda  $K_{nas}$  adını verdikleri yeni bir yöntem önermişlerdir (Karakaş, 2008).

$$K_{nas} = \max \left( \frac{\hat{\sigma}^2}{\hat{\beta}_j^2} + \frac{1}{\lambda_j} \right), \quad j = 1, 2, \dots, p \quad (3.35)$$

### 3.3. Ridge Tahmin Edicisinin Risk Fonksiyonu

$(\hat{\theta} - \theta)^2$  karesel kayıp fonksiyonu veya hata karenin kayıp fonksiyonu olarak adlandırılırsa,  $E(\hat{\theta} - \theta)^2$ , tahmin edicinin risk fonksiyonu olarak adlandırılır. (aslında  $\theta$  parametresinin  $\hat{\theta}$  tahmincisinin hata kare ortalamasıdır) Öncelikle EKK'nın HKO'sını tanımlayalım:

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

Aynı şekilde EKK parametresinin risk fonksiyonu:

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

olarak gösterilir. Bu şekilde  $X'X = I_p$  olarak tanımlandığında risk fonksiyonu tekrar aşağıdaki gibi yazılır:

$$Risk(\hat{\beta}) = \sigma^2 tr(I_p) = \sigma^2 p \quad (3.36)$$

Ridge tahmin edicisinin risk fonksiyonunu Shah, S.(2015) tarafından aşağıdaki gibi tanımlanmıştır (Shah, 2015).

$$\begin{aligned} \hat{\beta}_k &= (X'X + kI_p)^{-1}X'y, \\ &= (I_p + k(X'X)^{-1})^{-1}(X'X)^{-1}X'y \end{aligned}$$

Burada

$$W = (I_p + k(X'X)^{-1})^{-1}$$

eşitliği tanımlanırsa Ridge parametresini tekrar,

$$\hat{\beta}_k = W\hat{\beta}$$

biçiminde yazılabilir. HKO'sı da aşağıdaki gibidir.

$$\begin{aligned}
HKO(\hat{\beta}_k) &= E((\hat{\beta}_k - \beta) (\hat{\beta}_k - \beta)') \\
\hat{\beta}_k - \beta &= W\hat{\beta} - W\beta + W\beta - \beta \quad (W\beta \text{ eklenip çıkarılırsa}) \\
&= W(\hat{\beta} - \beta) + (W - I_p) \beta
\end{aligned}$$

W'lar türünden HKO'sı;

$$\begin{aligned}
HKO(\hat{\beta}_k) &= E\{(W(\hat{\beta} - \beta) + (W - I_p) \beta) (W(\hat{\beta} - \beta) + (W - I_p) \beta)'\} \\
&= W(E(\hat{\beta}_k - \beta) (\hat{\beta}_k - \beta)') W' + E(W - I_p) \beta \beta' (W - I_p)' \\
&= \sigma^2 W(X'X)^{-1} W' + (W - I_p) \beta \beta' (W - I_p)'
\end{aligned}$$

olarak elde edilir ve risk fonksiyonunda,

$$\begin{aligned}
Risk(\hat{\beta}) &= \sigma^2 tr(W(X'X)^{-1} W') + \beta' (W - I_p)^2 \beta \\
X'X = I_p &\Rightarrow W = (I_p + k I_p)^{-1} = \frac{1}{(k+1)} I_p \\
&= \sigma^2 tr\left(\frac{1}{(k+1)} I_p I_p I_p \frac{1}{(k+1)}\right) + \beta' \left(\frac{1}{(k+1)} I_p - I_p\right)^2 \beta \\
&= \frac{\sigma^2 p}{(1+k)^2} + \frac{k^2 \beta' \beta}{(1+k)^2}
\end{aligned}$$

$$\begin{aligned}
\Delta^2 &= \beta' \beta / \sigma^2 \\
&= \frac{\sigma^2}{(1+k)^2} [p + k^2 \Delta^2], \quad k > 0, \quad \Delta^2 \geq 0 \quad (3.37)
\end{aligned}$$

olarak elde edilir. Burada  $\Delta^2$  ıraksama (divergence) parametresi olarak tanımlanır. Normalize edilmiş katsayıların kareleri toplamıdır (Shah, 2015).

### 3.4. Ridge Tahmin Edicisinin Kullanım Alanları

Ridge yöntemi kimya mühendisliği, jeoloji, finans, sosyal bilimler, meteoroloji ve diğer birçok alanda kullanılmaktadır. Bu alanlarda yapılan çalışmalardan bazıları aşağıda verilmiştir:

- Miller (1972) “aggregate data” ile seçimlerde değişkenlik katsayısını incelerken datadaki çoklu doğrusallığı gidermek için Ridge tahmin edicisini ele almışlardır.

- Motor yađı sıcaklıđı, motor alıřtıran etmenler ve tasarımın deđiřkenlerinin arasında oklu bađlantının olmasından dolayı Gallopoulos (1974) alıřmasında Ridge regresyonu kullanmıřtır.
- Brown ve Beattie (1975) üretim fonksiyonlarının uygulama alanlarında ekonomik parametreleri tahmin ederken Ridge regresyonu kullanmıřlardır.
- Ekonomik modellerde oklu dođrusal bađlantı olabileceđi kuřkusuyula Watson ve White (1976) para talebinin tahmininde Ridge regresyonu kullanmıřlardır.
- Meisner (1979) kısa dnemde elde edilen verilerden uzun zaman aralıđında olabilecek yađıř miktarlarını tahmin edebilmek amacıyla Ridge regresyon analizini kullanmıřlardır.
- Maclaren (1980) İngiltere’de et toptan satıř fiyatını Ridge regresyon ile tahmin etmiřlerdir (Akdeniz ve Kaıranlar, 1995).

#### 4. LIU TAHMİN EDİCİSİ

Çoklu bağlantı durumunda önerilen yanlı tahmin edicilerden biri de Liu tahmin edicisidir. Liu (1993), Ridge regresyonun farklı bir kombinasyonu ile Stein tipi tahmin ediciyi ele alarak bir kombinasyon oluşturmuş ve bu tahmin edicilerin avantajlarını da bir araya getirerek yeni bir yanlı tahmin edici tanımlamıştır. Liu tanımladığı bu tahmin edici ile EKK'yı karşılaştırmış ve önerdiği tahmin edicinin EKK tahmin edicisinden az hata kareler ortalaması verdiğini göstermiştir (Kejian, 1993; Kaçiranlar ve Sakallıoğlu, 2000). Doğrusal Regresyon modeli,

$$Y = X\beta + e$$

biçiminde tanımlansın.  $X_{n \times (p+1)}$  bağımsız değişkenler vektörü iken, kanonik formda  $X'X = \Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_p)$  biçiminde yazılır ve  $\lambda_i$ ,  $X'X$ ' in özdeğerleridir.  $E(\varepsilon) = 0$  ve  $\text{cov}(\varepsilon) = \sigma^2 I$  hata vektörü olmak üzere  $\beta$ 'nin Liu tahmin edicisi  $\hat{\beta}(d)$ ,

$$\hat{\beta}_d = (X'X + I)^{-1}(X'y + d\hat{\beta}) \quad (4.1)$$

biçiminde tanımlanır. Burada  $d$ , Liu tahmin edicisinin yan parametresidir ve  $0 < d < 1$  aralığında tüm  $\sigma^2$  ve  $X$  değerleri için EKK kestricisinden daha küçük HKO vermektedir. Ridge tahmin edicisindeki  $k$  parametresi gibi  $d$  parametresinin kestirimi için birçok yöntem önerilmiştir (Kejian, 1993; Akdeniz ve Kaçiranlar, 1995).

Liu tahmin edicisi EKK tahmin edicisinin doğrusal bir fonksiyonu olarak gösterilebilir. Eşitlik (4.1)'de,  $X'y = (X'X)\hat{\beta}$  olarak alınırsa eşitlik tekrardan aşağıdaki gibi elde edilir:

$$\begin{aligned} \hat{\beta}_d &= (X'X + I)^{-1}((X'X)\hat{\beta} + d\hat{\beta}) \\ \hat{\beta}_d &= (X'X + I)^{-1}(X'X + dI)\hat{\beta} \end{aligned} \quad (4.2)$$

Bu eşitlikte  $A = (X'X + I)^{-1}(X'X + dI)$  olarak alınırsa, Liu tahmin edicisi  $\hat{\beta}_d = A\hat{\beta}$  biçiminde EKK tahmin edicisinin fonksiyonu olarak yazılabilir.  $d = 1$  olduğunda Liu tahmin edicisi EKK tahmin edicisine eşitlenir. Liu tahmin edicisinin genelleştirilmiş biçimi;  $0 < d_j < 1$  ve  $j = 1, 2, \dots, p+1$  olmak üzere aşağıdaki gibidir:

$$\hat{\beta}_d = (X'X + I)^{-1} \left( X'y + \begin{pmatrix} d_1 & \cdots & \\ \vdots & \ddots & \vdots \\ & \cdots & d_{p+1} \end{pmatrix} \right) \hat{\beta} \quad ,$$

Bu durumda yanlılık parametresi matris şeklinde ifade edilir (Akdeniz ve Kaçıranlar, 1995).

#### 4.1. Liu Tahmin Edicisinde Yanlılık, Beklenen Değer, Varyans ve Hata Kare Ortalaması

Eşitlik (4.2) de,  $A = (X'X + I)^{-1}(X'X + dI)$  olacak şekilde tanımlanır ve  $d=1$  alınırsa Liu tahmin edicisi EKK tahmin edicisine dönüşür:

$$\hat{\beta}_d = A \hat{\beta}$$

Bu eşitliğin beklenen değeri alınıp A'nın değeri yerine yazıldığında Eşitlik (4.3)'de  $\hat{\beta}_d$ 'nin beklenen değeri elde edilir:

$$E(\hat{\beta}_d) = E(A \hat{\beta}) = A E(\hat{\beta}) = A \beta$$

$$E(\hat{\beta}_d) = E[(X'X + I)^{-1}(X'X\hat{\beta} + d\hat{\beta})] = E[(X'X + I)^{-1}(X'X + dI)\hat{\beta}]$$

$$E(\hat{\beta}_d) = (X'X + I)^{-1}(X'X + dI)\beta \quad (4.3)$$

Liu tahmin edicisinin yanlılık miktarı ise,

$$yan(\hat{\beta}_d) = E(\hat{\beta}_d) - \beta$$

$$yan(\hat{\beta}_d) = (X'X + I)^{-1}(X'X + dI)\beta - \beta$$

$$yan(\hat{\beta}_d) = [(X'X + I)^{-1}(X'X + dI) - I]\beta$$

olarak hesaplanır. Liu tahmin edicisinin varyansı ise,

$$Var(\hat{\beta}_d) = \sigma^2 (X'X + I)^{-1}(X'X + dI)(X'X)^{-1}(X'X + dI)(X'X + I)^{-1},$$

biçimindedir. HKO'sı da aşağıdaki gibidir:

$$HKO(\hat{\beta}_d) = \sigma^2 \sum_{i=1}^p \frac{(\lambda_i + d)^2}{\lambda_i(\lambda_i + 1)^2} + (d - 1)^2 \sum_{i=1}^p \frac{\alpha_i^2}{(\lambda_i + 1)^2} \quad (4.4)$$

#### 4.2.Liu Tahmin Edicisinin Risk Fonksiyonu

Eşitlik (4.2),  $\hat{\beta}_d = F\hat{\beta}$  biçiminde yazılabilir. Burada  $F = (X'X + I_p)^{-1}(X'X + dI_p)$ 'dir.

$\hat{\beta}_d$  parametresinin HKO ve Risk fonksiyonları da sırasıyla aşağıdaki gibi tanımlanabilir:

$$HKO(\hat{\beta}_d) = E(\hat{\beta}_d - \beta)(\hat{\beta}_d - \beta)'$$

$$HKO(\hat{\beta}_d) = E\{(F\hat{\beta} - \beta) + (F - I_p)\beta\}(F\hat{\beta} - \beta) + (F - I_p)\beta\}'\}$$

$$= F(E(\hat{\beta} - \beta)(\hat{\beta} - \beta)')F' + E(F - I_p)\beta\beta'(F - I_p)'$$

$$= \sigma^2 F(X'X)^{-1}F' + (F - I_p)\beta\beta'(F - I_p)'$$

$$Risk(\hat{\beta}_d) = \sigma^2 \text{tr}(F(X'X)^{-1}F') + \beta'(F - I_p)^2\beta$$

$$X'X = I_p \Rightarrow F = (I_p + I_p)^{-1}(I_p + dI_p)$$

$$= \sigma^2 \text{tr}\left(\frac{1+d}{2}\right) I_p I_p I_p \frac{1+d}{2} + \beta' \left(\frac{1+d}{2} I_p - I_p\right)^2$$

$$= \frac{\sigma^2 (1+d)^2 p}{4} + \frac{(d-1)^2 \beta'\beta}{4}$$

$\Delta^2 = \beta'\beta/\sigma^2$  olarak ifade ettiğimizde;

$$Risk(\hat{\beta}_d) = \frac{\sigma^2}{4} [(1+d)^2 p + (d-1)^2 \Delta^2], \quad d > 0, \Delta^2 \geq 0$$

$\Delta^2$  iraksama (divergence) parametresi olarak tanımlanır (Shah, 2015).

#### 4.2.Liu Tahmin Edicisinde Yanlılık Parametresi

Bir regresyon modelinde,  $X$  standartlaştırılmış gözlem değerleri olmak üzere,

$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$  olarak ifade edilen  $\lambda_j$  değerleri  $X'X$ 'in özdeğerleridir.

Özdeğerlere karşılık yazılan özvektörler ise  $Q = (q_1, q_2, \dots, q_p)$  şeklinde

yazılırsa,  $\Lambda = \begin{pmatrix} \lambda_1 & \cdots & \\ \vdots & \ddots & \vdots \\ & \cdots & \lambda_p \end{pmatrix}$  olarak tanımlanır ve Liu tahmin edicisi kanonik formda tekrar yazılır.

$Z = XQ$  ve  $\alpha = Q'\beta$  olsun.  $Z'Z = Q'(X'X)Q = \Lambda$  olmak üzere  $\alpha$ 'nın EKK, Ridge, Liu ve genelleştirilmiş Liu tahmin edicileri sırasıyla aşağıdaki gibi yazılabilir:

$$\hat{\alpha} = \Lambda^{-1}Z'y,$$

$$\hat{\alpha}_r = (\Lambda + kI)^{-1}Z'y,$$

$$\hat{\alpha}_d = (\Lambda + kI)^{-1}(Z'y + d\hat{\alpha}),$$

$$\hat{\alpha}_d = (\Lambda + kI)^{-1}\left(Z'y + \begin{pmatrix} d_1 & \cdots & \\ \vdots & \ddots & \vdots \\ & \cdots & d_p \end{pmatrix} \hat{\alpha}\right),$$

Parametreler kanonik formda yazıldığında ise

$$\hat{\beta} = Q\hat{\alpha},$$

$$\hat{\beta}_r = Q\hat{\alpha}_r,$$

$$\hat{\beta}_d = Q\hat{\alpha}_d,$$

$$\hat{\beta}_{Gd} = Q\hat{\alpha}_{Gd},$$

şeklinde yazılır ve

$$\hat{\alpha}_{rj}(k) = \frac{\lambda_j}{\lambda_j + k} \hat{\alpha}_j \text{ ve}$$

$\hat{\alpha}_{Gdj}(d) = \frac{\lambda_j + d_j}{\lambda_j + 1} \hat{\alpha}_j, j = 1, \dots, p$ 'dir. Liu (1993) bu fonksiyonların grafiklerini çizdirerek yanlılık parametresi seçimi için grafiksel yöntem önermiştir.

Teorik sonuçlara göre katsayılarının kestiriminde Liu tahmin edicisinin diğer tahmin edicilerden daha iyi olup olmaması bilinmeyen  $\beta$  ve  $\sigma^2$  parametrelerine ve  $d$ 'nin seçimine göre değişmektedir. Bunun için uygulanabilirlik bakımından, bilinmeyen parametrelerin yerine bu parametrelerin kestirim değerleri alınır. Liu, Ridgedeki  $k$ 'nın kestiriminden esinlenerek  $d$ 'nin kestirimi için Eşitlik (4.5) 'i önermiştir:

$$\hat{d}_{MME} = 1 - \hat{\sigma}^2 \left[ \sum_{i=1}^p \frac{1}{\lambda_i(\lambda_i+1)} / \sum_{i=1}^p \frac{\hat{\alpha}_i^2}{(\lambda_i+1)^2} \right] \quad (4.5)$$

Liu (1993), daha sonra formülü genelleştirerek,  $h > 0$  için,

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

biçiminde tanımlanmıştır. Daha sonra Liu, Ridge yönteminde optimal  $k$  seçimi için önerilen Mallows (1973)'un  $C_p$  istatistiğini de kullanarak aşağıdaki eşitlikleri tanımlanmıştır.

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

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

şeklinde tanımlanır. Burada  $h > 0$  ve  $\hat{\alpha}$  ve  $\hat{\sigma}^2$  EKK kestirim değerleridir (Akdeniz ve Kaçıranlar,1995).

Akdeniz ve Kaçıranlar (1995), Liu için HKO'sını ve  $d_{opt}$  değerini tekrardan aşağıdaki gibi tanımlamışlardır:

$$HKO(\hat{\beta}_{Liu}) = \sum_{i=1}^p \frac{\sigma^2(\lambda_i + d_i)^2 + \lambda_i(1-d_i)^2 \beta_i^2}{\lambda_i(\lambda_i+1)^2} \quad (4.6)$$

$$d_{i(opt)} = \frac{\lambda_i(\beta_i^2 - \sigma^2)}{\lambda_i \beta_i^2 + \sigma^2} \quad i = 1, \dots, p \quad (4.7)$$

Genelleştirilmiş Liu tahmin edicisinin HKO'sı ve  $d$  değeri de,

$$HKO(\hat{\alpha}_{Gd}) = \hat{\sigma}^2 \sum_{j=1}^p \frac{(\lambda_j + d)^2}{\lambda_j(\lambda_j+1)^2} + \sum_{j=1}^p \frac{(d-1)^2 \alpha_j^2}{(\lambda_j+1)^2} \quad (4.8)$$

$$d_j = \frac{\alpha_j^2 - \sigma^2}{\alpha_j^2 + (\sigma^2/\lambda_j)}, j = 1, \dots, p \quad (4.9)$$

şeklinindedir. Liu (1993)'nun Ridge regresyon tahmin edicisini kullanarak önerdiği tahmin edicilere ek olarak Akdeniz vd. (2006) optimal  $d$  parametresi önermişlerdir.  $X'X = I$  olması durumunda Eşitlik (4.2)'den elde edilen Liu tahmin edicisi,

$\hat{\beta}_d = \frac{1}{2}(I + d) X'y$  biçiminde elde edilir ve burada  $d$  parametresi,

$$d_{opt} = 1 - \frac{2p\sigma^2}{p\sigma^2 + \beta' \beta}$$

olarak elde edilir. Bilinmeyen parametre tahminleri yerine yazıldığında  $d$  parametresi,



$$\hat{d}_{opt} = 1 - 2p \frac{\hat{\sigma}^2}{\hat{\beta}'\hat{\beta}}$$

şeklinde yazılabilir.  $d_{opt}$  parametresi için verilen üst sınır ise,

$$d^* = 1 - \left(\frac{\lambda_{min}}{\lambda_{max}}\right)^2 p \frac{\hat{\sigma}^2}{\hat{\beta}'\hat{\beta}} \text{ olmak üzere,}$$

$$\hat{d}_{opt} < 1 - \left(\frac{\lambda_{min}}{\lambda_{max}}\right)^2 p \frac{\hat{\sigma}^2}{\hat{\beta}'\hat{\beta}} = d^* \text{ biçimindedir.}$$

Bu eşitsizlik,  $p > 2$  için,  $E(\hat{d}_{opt}) < E(d^*) = 1 - \left(\frac{\lambda_{min}}{\lambda_{max}}\right)^2 \frac{p}{p-2}$  biçimindedir (Neter, 1996; Fox, 1997).

## 4.2. Sağlam Liu Tahmin Edicisi

Doğrusal regresyonda çoklu bağlantı ve aykırı değer önemli sorunlardır. Bu sorunların birlikte görülmesi de daha sıkıntılı bir durumdur ve araştırmacılar için sorun yaratmaktadır. Çoklu bağlantı için önerilen Ridge tahmin edicisi  $\hat{\beta}_k$  aykırı değere duyarlı olduğu için Silvapulle (1991) tarafından alternatif olarak sağlam (robust) ridge-M tahmin edicisi ( $\hat{\beta}_{RM}$ ) önerilmiştir. Bununla birlikte,  $\hat{\beta}_{RM}$   $k$ 'nin karmaşık bir fonksiyonudur; bu nedenle,  $k$ 'nin seçimi için önerilen yöntemlerden bazılarının kullanılması,  $k$  için daha karmaşık denklemler verir. Bu zorlukların üstesinden gelmek için Liu tahmin edicisi önerilmiştir (Kejian, 1993).

$$\hat{\beta}_d = (X'X + dI)^{-1} X'X \hat{\beta}_{EKK}$$

Burada  $0 < d < 1$  aralığında bir parametredir.  $\hat{\beta}_d$ 'nin  $\hat{\beta}_k$ 'ya avantajı  $d$ 'nin  $\hat{\beta}_d$ 'nin doğrusal fonksiyonu olmasıdır. Böylece  $d$ 'yi seçmek  $k$ 'yi seçmekten daha kolay olur. Bununla beraber  $\hat{\beta}_d$ 'nin EKK'yi küçültme esasına dayanmasından dolayı aykırı değerler  $\hat{\beta}_d$ 'yi etkilemiştir. Bu problemlerin üstesinden gelmek için yani hem çoklu bağlantı hem de aykırı değer durumunda Liu-türü M kestiricisi, Arslan ve Billor tarafından önerilmiştir (Arslan ve Billor, 2000):

$$\hat{\beta}_{LM}(d) = (X'X + dI)^{-1} X'X \hat{\beta}_M$$

Burada,  $\hat{\beta}_M$ , Huber-M tahmin edicisidir ve amaç fonksiyonu aşağıdaki gibidir (Monsteller et al., 1982).

$$\rho(x) = \begin{cases} \frac{1}{2}u^2 & , |u| \leq k \\ k|u| - \frac{1}{2}k^2, & |u| > k \end{cases}, k = 1.345$$

$\lambda_1 \geq \lambda_2, \geq \dots \geq \lambda_p$   $X'X$ 'in özdeğerleri ve  $q_1, q_2, \dots, q_p$ , özvektörler olmak üzere.  $\Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_p)$  ve  $P = (q_1, q_2, \dots, q_p)$  olduğunda  $X'X = P \Lambda P'$  olarak yazılır. Regresyon eşitliğinin kanonik formu aşağıdaki gibi yazılır:

$$y = \beta_0 I + c\alpha + \varepsilon$$

Bu eşitlikte  $c = XP$  ve  $\alpha = P'\beta$  olarak alındığında Liu tahmin edicisi aşağıdaki formda tanımlanabilir:

$$\hat{\alpha}_L(d) = (\Lambda + I)^{-1}(\Lambda + dI) \hat{\alpha}_{OLS}$$

Liu-M tahmin edicisi de,

$$\hat{\alpha}_{LM}(d) = (\Lambda + I)^{-1}(\Lambda + dI) \hat{\alpha}_M \quad (4.10)$$

biçiminde yazılabilir (Arslan ve Billor, 2000). Liu türü M kestiricisinin HKO'sı aşağıdaki gibi elde edilir:

$$MSE(\hat{\alpha}_M) = \sum_{i=1}^p \Omega_{ii}$$

$$MSE(\hat{\alpha}_L(d)) = \sigma^2 \sum_{i=1}^p \frac{(\lambda_i + d)^2}{\lambda_i(\lambda_i + d)^2} + (d - 1)^2 \sum_{i=1}^p \frac{\alpha_i^2}{(\lambda_i + 1)^2}$$

$$MSE(\hat{\alpha}_{LM}(d)) = \sigma^2 \sum_{i=1}^p \frac{(\lambda_i + d)^2}{\lambda_i(\lambda_i + d)^2} \Omega_{ii} + 2(d - 1)^2 \sum_{i=1}^p \frac{\alpha_i^2}{(\lambda_i + 1)^2}$$

Eşitlik (4.10)'daki d parametresi M tahmini,

$$\hat{d}_M = 1 - A^2 \left[ \sum_{i=1}^p \frac{1}{(\lambda_i + 1)^2} \right] / \sum_{i=1}^p \frac{\alpha_{Mi}^2}{(\lambda_i + 1)^2} \text{ biçimindedir.}$$

## 5. LASSO TAHMİN EDİCİSİ

EKK, Ridge ve Liu tahmin edicilerinin tahmin kesinliği yüksek olmasına rağmen bu tahmin edicilerin varyansının bir üst sınırı olmamaktadır. Bu yöntemler çoklu bağlantı durumunda büyük varyansa sahip olmaları nedeniyle yeterli olmadıkları durumlar ortaya çıkmaktadır. EKK tahminleri genellikle düşük yanlı fakat büyük bir varyansa sahiptirler. Kestirim doğruluğunu elde etmek için bazen daralma (shrinkage) yapılabilir ya da bazı katsayılar 0'a ayarlanarak geliştirilebilir. Bunu yaparak, öngörülen değerlerin varyansını azaltmak için küçük bir sapmayı ihmal edebilir ve dolayısıyla genel tahmin doğruluğu geliştirebilir. Ridge ve Liu tahmin edicileri, katsayıları daraltan sürekli bir süreçtir ve bu nedenle daha kararlıdır; ancak model yorumlama sorunu katsayıların hiçbirini 0'a ayarlamadıkları için hala çözülememiştir. Bu gibi sorunlara çözüm olarak Tibshirani (1996), LASSO (Least Absolute Shrinkage and Selection Operator) adı altında yeni bir yöntem önermiştir. Bu yöntem katsayıların mutlak değerinin toplamının sabitten az olması durumunda artık karelerin toplamını en aza indirme mantığına göre kurulmuştur. LASSO fikri oldukça geneldir ve çeşitli istatistiksel modellerde uygulanabilir (Tibshirani, 1996).

i.gözleme ilişkin doğrusal regresyon modeli,

$$y_i = \alpha + \sum_{j=1}^p x_{ij} \beta_j + \varepsilon_i, \quad i = 1, 2, \dots, n \quad (5.1)$$

biçiminde tanımlansın. i. gözleme ilişkin için bileşen  $(x^i, y_i)$  olsun.

Burada  $x^i = (x_{i1}, x_{i2}, \dots, x_{ip})^T$  standartlaştırılmış bağımsız değişkenlerdir.

Yani  $\sum_i \frac{x_{ij}}{n} = 0$  ve  $\sum_i \frac{x_{ij}^2}{n} = 1$ 'dir.  $y_i$ 'ler bağımlı değişkenlerdir. Tibshirani tarafından LASSO,

$$\sum |\beta_j| \leq t \text{ kısıltması ile}$$

$$(\hat{\alpha}, \hat{\beta}) = \arg \min(\{\sum_{i=1}^n (y_i - \alpha - \sum_{j=1}^p x_{ij} \beta_j)\}^2) \quad (5.2)$$

olarak tanımlamıştır. Denkleminin çözümü doğrusal eşitsizlik kısıtlamalı bir karesel programlama problemidir. LASSO regresyon katsayıları  $y$ 'nin bir fonksiyonu olarak yazılamaz ancak karesel programlama ile optimize edilerek çözümü bulunabilir.

$t \geq 0$  parametresi, tahminlere uygulanan bzlmenin miktarını kontrol eder. Yukarıdaki  $t \geq 0$ , ayar (tuning) parametresidir. Tm  $t$ 'ler iin  $\alpha$ 'nın zm  $\hat{\alpha} = \bar{y}$ 'dir. Genelliđi bozmaksızın  $\bar{y} = 0$  kabul edilir. Bu nedenle  $\alpha$  ihmal edilir. LASSO regresyonun ama fonksiyonu ařađıdaki gibidir:

$$L(\beta) = \sum_{i=1}^n (y_i - \alpha - \sum_{j=1}^p x_{ij} \beta_j)^2 + \lambda_L \sum_{j=1}^p |\beta_j|^\gamma \rightarrow \min \text{ olur} \quad (5.3)$$

Burada ama Eřitlik (5.3)' minimize etmektir ya da eřdeđer olarak  $\lambda > 0$  ile  $\sum_{j=1}^p |\beta_j|^\gamma \leq t$  kısıtı ile  $\min (y - X\beta)^T (y - X\beta)$ 'yi minimum yapmaktır (Tibshirani,1996).

Eřitlik (5.3)' te  $\gamma=2$  iken Ridge tahmin edici,  $\gamma=1$  iken LASSO tahmin edici adını alır. Burada optimal  $\gamma$  deđeri seim ynteminin en iyi sonucu elde etmek iin nemlidir. LASSO tahmin edicisinin zm iin algoritmalar tasarlanmıřtır. Bu alıřmada Modified-Newton-Raphson, Atıř ve IRLS (İteratif Ađırlıklandırılmıř EKK) algoritmaları ele alınmıřtır.

### 5.1. Modified-Newton-Raphson (M-N-R) Algoritması ( $\gamma > 1$ )

Bu algoritmanın adımları ařađıda verildiđi gibidir;

- (1)  $\hat{\beta}_0 = \hat{\beta}_{EKK} = (\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_p)^T$  ile bařlar.
- (2)  $m$ . adımda,  $j=1, \dots, p$  iin  $S_0 = S_j (0, \hat{\beta}^{-j}, X, y)$  olur .  $S_0=0$  iken  $\hat{\beta}_j = 0$   
 $\gamma > 2$  ise  $\hat{\beta}_j$ 'nin zm iin Newton-Raphson metodu uygulanır.  $\gamma < 2$  iken fonksiyon dzeltilir. Tm  $\hat{\beta}_j$ 'ler bulunduktan sonra  $\hat{\beta}_m = (\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_p)^T$  bulunur.
- (3)  $\hat{\beta}_m$  bir noktaya yakınsayana kadar 2. adımı tekrarlanır (Wenjang, 1998).

### 5.2. Atıř Algoritması

Bu algoritmanın adımları ařađıda verildiđi gibidir;

- (1)  $\hat{\beta}_0 = \hat{\beta}_{EKK} = (\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_p)^T$  ile bařlar.

(2) *m. adımda*,  $j = 1, \dots, p$  için  $S_0 = S_j(0, \hat{\beta}^{-j}, X, y)$  olur ve  $x_j$ ,  $x$ 'in  $j$ . kolonu olmak üzere;

$$\hat{\beta}_j = \begin{cases} \frac{\lambda - S_0}{2x_j^T x_j}, & S_0 > \lambda \\ \frac{-\lambda - S_0}{2x_j^T x_j}, & S_0 < -\lambda \\ 0, & |S_0| \leq \lambda \end{cases}$$

olarak hesaplanır. Tüm  $\hat{\beta}_j$ 'ler bulunduktan sonra  $\hat{\beta}_m = (\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_p)^T$  bulunur.

(3)  $\hat{\beta}_m$  bir noktaya yakınsayana kadar 2. adımı tekrarlanır.

### 5.3. IRLS Yöntemiyle Algoritma

Algoritmanın adımları aşağıda verilmiştir:

- (1)  $\hat{\beta}_0$  başlangıç değeri ile başlar.
- (2) Genel tahmin olarak  $\hat{\beta}$ 'ya bağlı düzeltilmiş değişken  $z$ 'yi

$$z = X\hat{\beta} + [V(\mu)]^{-1}(y - \mu) \text{ olarak tanımlar.}$$

- (3)  $W = V^{-1}$  olarak alınır ve  $WX$  üzerinde  $Wz$ 'nin lineer regresyonuna MNR

(Modified Newton Raphson) metodunu uygulanır ve  $\hat{\beta}$ 'yi günceller.

- (4)  $\hat{\beta}$ 'nin yaklaşımı elde edilinceye kadar 2. ve 3. adımları tekrarlanır.

Bu algoritmada ortak olabilirlik fonksiyonu yoksa Jacobian koşulu sağlandığı sürece LASSO tahmin edici elde etmek için MNR (Modified-Newton-Raphson) metodu ya da Atış metodu uygulanabileceğine dikkat çekilmektedir.  $\lambda > 0$  ve  $\frac{\partial S}{\partial \beta}$  pozitif tanımlı ise, Atış algoritması;  $\gamma=1$  için LASSO tahmin edicisine yakınsar. Atış algoritmaları çok hızlı yakınsar ve IRLS yöntemi ile birleştirilebilir (Fidanoğlu, 2009).

### 5.4. LASSO'nun Standart Hatası

LASSO için standart hatalarının doğru bir tahminini elde etmek zordur. Bu yaklaşım, bootstrap yolu ile yapılır: ya  $t$  sabitlenebilir ya da her bir bootstrap örneği için  $t$ 'ye göre optimize edebiliriz.  $t$ 'yi sabitleme, en iyi alt kümenin seçilmesine ve ardından bu alt

kümenin en küçük kareler standart hatasını kullanmaya benzemektedir. LASSO kestirim değeri aşağıda verilmektedir:

$$\beta^* = (X^T X + \lambda W^-)^{-1} X^T y, W \text{ köşegen matris, } \sum |\beta_j|^* = t \text{ dir.} \quad (5.4)$$

$\hat{\sigma}^2$ , hata varyansının bir tahminidir. Bu eşitlikle ilgili bir zorluk,  $\beta_j=0$  için tahmin edilen değerlerin 0'ı vermesidir. Bu yaklaşım aynı zamanda, LASSO tahmininin hesaplanması için yinelenen bir Ridge regresyonu algoritması önerilmiştir. Ancak bunun, oldukça verimsiz olduğu ortaya çıkmakla beraber LASSO parametresinin seçiminde yararlı olduğu görülmüştür (Tibshirani, 1996).

Geleneksel tahmin yöntemleri ile karşılaştırıldığında LASSO'nun önemli avantajı parametre tahmini ve değişken seçiminin her ikisini birden eşanlı olarak yerine getirmesi, sayısal olarak uygulanabilir olması olarak belirtilebilir.

### 5.5. LASSO'nun Risk Fonksiyonu

Lasso regresyon tahmin edicisinin risk fonksiyonu,

$$R(\hat{\beta}^L) = \sigma^2 (1 + 2 \ln(p)) \left( \frac{1+q}{p} \right) \quad (5.5)$$

biçimindedir. Burada  $p$  ve  $q$  sabit değer olarak alındığında farklı değerler için risk fonksiyonları hesaplanmaktadır (Shah, 2015).

### 5.6. LASSO'nun Kullanım Alanları

LASSO regresyon tahmin edicisi aktüerya, biyoloji, veterinerlik, genetik bilimi, makine öğrenmesi ve tıp gibi diğer birçok alanda kullanılmaktadır. Bu alanda yapılan çalışmalardan bazıları,

- Billor N. ve Godwin J. (2015) fMRI veri analizi adlı çalışmasında, daha iyi bir sınıflandırma yönteminin geliştirilmesine ihtiyaç duymuşlar ve LASSO L1 adı altında bir değişken seçim yolu geliştirmişlerdir.

- Alçin Ö. F., Şengür A., İnce M. İleri-Geri Takip Algoritması Tabanlı Seyrek Aşırı Öğrenme Makinesi çalışmasında, LASSO ve Elastic Net metodun karşılaştırılmasını yapmıştır.
- Çiftsüren M. ve Akkol S. tarafından yapılan çalışmada yumurta kalite özelliklerinin ve değişken seçim yöntemlerini kullanarak değişken seçiminin tahmin edilmesi için Ridge, LASSO ve elastik ağ yöntemleri çalışılmıştır. Bu çalışma, yumurtaların dış kabuk kalite özelliklerini kullanarak iç kalite özelliklerini belirlemek amacıyla yapılmıştır. Değişkenler Ridge, LASSO ve elastik net regülasyon yöntemleri kullanılarak en basit modeli elde etmek için seçilmiştir. Hem yumurta sarısı ağırlığı hem de yumurta ağırlığı için iki belirteç içeren LASSO'nun, yüksek tahmin doğruluğu açısından en iyi model olduğu ortaya çıkmıştır.
- Akdeniz F. ve Fidanoğlu I.(2010) LASSO ve Ridge tahmin edicileri ile çoklu bağlantı problemi incelenmiş, büzülme ve ayar parametreleri incelenmiş ve tahmin ediciler karşılaştırılmıştır. Sonuç olarak Ridge tahmin ediciye kıyasla ve EKK tahmin ediciye göre daha iyi, sıfır olmayan fakat küçük parametreler içerdiğinde ise Ridge regresyonun LASSO'dan daha iyi performans verdiği sonucuna varmışlardır.

## 6. UYGULAMA

Bu bölümde Ridge, Liu ve LASSO tahmin edicilerini karşılaştırmak için bir veri üzerinde uygulama yapılmıştır. Yöntemlerin karşılaştırılması için HKO kullanılmıştır. Uygulamada hesaplanan HKO'ları aşağıda verilmiştir. Verilerin değerlendirilmesi R Studio programında yapılmıştır.

### 6.1.Hald'ın Verileri ile Uygulama

Hald'ın veri kümesi 13 gözlem ve 4 bağımsız değişkenden oluşmaktadır. Y sıcaklık (cals/gm),  $X_1$  karışımdaki  $3CaO.Al_2O_3$  oranı,  $X_2$  karışımdaki  $3CaO.SiO_2$  oranı,  $X_3$  karışımdaki  $4CaO.Al_2O_3.Fe_2O_3$  oranı,  $X_4$  ise karışımdaki  $2CaO.SiO_2$  oranıdır. Hald'ın verilerinde değişken seçim ölçütleri Montgomery ve Peck (1982, p.256-266) ve Ronchetti ve Staudte (1994) tarafından da incelenmiştir.

**Çizelge 1.** Hald veri kümesi

	Y	$X_1$	$X_2$	$X_3$	$X_4$
1	78.5	7	26	6	60
2	74.3	1	29	15	52
3	104.3	11	56	8	20
4	87.6	11	31	8	47
5	95.9	7	52	6	33
6	109.2	11	55	9	22
7	102.7	3	71	17	6
8	72.5	1	31	22	44
9	93.1	2	54	18	22
10	115.9	21	47	4	26
11	83.8	1	40	23	34
12	113.3	11	66	9	12
13	109.4	10	68	8	12

#### 6.1.1. Çoklu Bağlantı İncelemesi

Veri kümesinde doğrusal bağımlılık olup olmadığını anlamak için çoklu bağlantı incelemesi yapılmıştır. Bunun için öncelikle değişkenlere ilişkin ilişki matrisi elde edilmiş ve Çizelge 2' de verilmiştir.



**Çizelge 2.** İlişki matrisi

		$X_1$	$X_2$	$X_3$	$X_4$
$X_1$	Pearson Correlation	1	.229	-.824**	-.245
	Sig. (2-tailed)		.453	.001	.419
	N	13	13	13	13
$X_2$	Pearson Correlation	.229	1	-.139	-.973**
	Sig. (2-tailed)	.453		.650	.000
	N	13	13	13	13
$X_3$	Pearson Correlation	-.824**	-.139	1	.030
	Sig. (2-tailed)	.001	.650		.924
	N	13	13	13	13
$X_4$	Pearson Correlation	-.245	-.973**	.030	1
	Sig. (2-tailed)	.419	.000	.924	
	N	13	13	13	13

İlişki matrisine göre  $X_1$  ve  $X_3$  değişkenleri arasında 0.824 büyüklüğünde,  $X_2$  ve  $X_4$  arasında 0.973 büyüklüğünde ters yönde ilişki vardır. Ancak ilişki matrisi iki değişken arasındaki ilişkiyi göstermektedir, daha fazla değişken arasında doğrusal bağımlılık varsa ilişki matrisi yeterli olmayacaktır. Şimdi de  $X'X$  matrisine ilişkin özdeğer ve özvektör matrisini inceleyelim.

**Çizelge 3.** Hald'ın verisi için özdeğerler ve koşul sayısı

Özdeğerler( $\lambda$ )	2.235704	1.576066	0.186606	0.001623
Koşul Sayısı(KS <sub>j</sub> )	1	1.191022	3.461341	37.11487

Çizelge 3'e göre  $\lambda_j$  değeri 0'a yakın olan bir tane özdeğer vardır ve buna karşılık gelen koşul sayısı 30'dan büyük olduğu için bir tane güçlü çoklu bağlantı olduğunu söyleyebiliriz.

**Çizelge 4.** Hald'ın verisi için özvektörler matrisi

	[,1]	[,2]	[,3]	[,4]
[1,]	-0.475955	0.508979	0.675500	0.241052
[2,]	-0.563870	-0.413931	-0.314420	0.641756
[3,]	0.394066	-0.604969	0.637691	0.268466
[4,]	0.547931	0.451235	-0.195421	0.676734

Özvektörler matrisi incelendiğinde  $X_1$  ve  $X_3$  ;  $X_2$  ve  $X_4$  değişkenlerinin doğrusal bağımlı olduğu görülmektedir. Değişkenler arasındaki doğrusal bağımlılık yapısı da aşağıdaki gibidir:

$$0.6755002 x_1 + 0.6376911 x_3 \cong 0$$

$$0.6417561 x_2 + 0.6376911 x_4 \cong 0$$

Bu veri kümesinin özelliği  $X_1$  ve  $X_3$  ;  $X_2$  ve  $X_4$  bağımsız değişkenlerinin ilişkili olmasıdır. Ayrıca veride hem çoklu bağlantı hem de aykırı değer olması durumunda kestirim yöntemlerinin karşılaştırılması amaçlanmıştır ve bağımlı değişkende tek aykırı değer oluşturulmuştur. Aykırı değer oluşturmak için bağımlı değişkenin 13. gözlem değeri 113.2 iken 200 olarak alınmıştır.

Tezde ele aldığımız yöntemlere ilişkin HKO'ları aşağıdaki çizelgede verilmiştir.

### 6.1.2. Ridge Tahmin Edicisinden Elde Edilen Sonuçlar

Hald'ın veri kümesi için Ridge tahmin edicisi

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

eşitliğinden elde edilmiştir.

Hald veri kümesine ilişkin Ridge parametre kestirim değerleri Çizelge 5 'te verilmiştir.

**Çizelge 5.** Ridge tahmin edicisi parametre kestirim değerleri

$\hat{\beta}_{x_1}$	$\hat{\beta}_{x_2}$	$\hat{\beta}_{x_3}$	$\hat{\beta}_{x_4}$
0.606512	0.527705	0.043389	-0.160287

Ridge'e ilişkin EKK ve sağlam parametre kestirimleri ve HKO'ları aşağıdaki eşitlikler kullanılarak elde edilmiş ve sonuçları Çizelge 6'da verilmiştir.

$$\hat{\beta}_{Ridge.EKK} = (X'X + k_{EKK}I_p)^{-1}\hat{\beta}_{EKK},$$

$$\hat{\beta}_{Ridge.sağlam} = (X'X + k_{sağlam}I_p)^{-1}\hat{\beta}_{EKK},$$

$$HKO\hat{\beta}_{Ridge.EKK} = \sigma^2 \sum_{j=1}^p \frac{\lambda_j \sigma^2 + k_{EKK}\beta_i^2}{(\lambda_j + k_{EKK})^2}$$

$$HKO\hat{\beta}_{Ridge.sağlam} = \sigma^2 \sum_{j=1}^p \frac{\lambda_j \sigma^2 + k_{sağlam}\beta_i^2}{(\lambda_j + k_{sağlam})^2}$$

Ridge yönteminde kullanılan k yanlılık parametresi  $k_{EKK}$ ,  $k_{sağlam}$ ,  $k_{medyan}$  olmak üzere üç farklı şekilde hesaplanmıştır. Eşitlik (3.14) ve (3.32)'den elde edilen k değerleri ve bunlara karşılık gelen HKO'ları Çizelge 6'da verilmiştir.

**Çizelge 6.** Ridge tahmin edicisinden elde edilen sonuçlar

Yöntem	Parametre	HKO
$\hat{\beta}_{Ridge.EKK}$	$k_{EKK} = 0.011623$	$HKO_{k_{EKK}} = 2.000666$
$\hat{\beta}_{Ridge.sağlam}$	$k_{sağlam} = 0.010173$	$HKO_{k_{sağlam}} = 0.591642$
$\hat{\beta}_{Ridge.medyan}$	$k_{medyan} = 0.005061$	$HKO_{k_{medyan}} = 0.919880$

Çizelge 6 incelendiğinde, yanlılık parametresi olarak Huber-M kestiricisinin kullanıldığı  $\hat{\beta}_{Ridge.sağlam}$ 'a ilişkin HKO'sının EKK ve medyana dayalı olarak elde edilen HKO'larından daha küçük olduğu görülmektedir. EKK yönetime ilişkin HKO'sı diğer HKO'larına göre daha büyük çıkmıştır.

Aykırı değerlerin yöntemler üzerindeki etkisini görmek için k değerleri ve HKO'ları aykırı değer durumunda tekrar hesaplanmış ve Çizelge 7'de verilmiştir.

**Çizelge 7.** Aykırı değer varlığında Ridge tahmin edicisi

Yöntem	Parametre	HKO
$\hat{\beta}_{Ridge.EKK}$	$k_{EKK} = 0.006277$	$HKO_{k_{EKK}} = 1166.26$
$\hat{\beta}_{Ridge.sağlam}$	$k_{sağlam} = 0.064717$	$HKO_{k_{sağlam}} = 14.92$
$\hat{\beta}_{Ridge.medyan}$	$k_{medyan} = 0.005061$	$HKO_{k_{medyan}} = 38.82$

Aykırı değerlerin tüm HKO değerlerini etkilediği Çizelge 7 'den açıkça görülebilmektedir. Ancak kendi içinde değerlendirdiğimizde yine en küçük HKO'sının Huber-M kestiricili Ridge yöntemine ait olduğu görülmektedir. EKK yönteminin HKO'sı diğerleri ile kıyaslanamayacak kadar büyüktür.

### 6.1.3. Liu Tahmin Edicisinden Elde Edilen Sonuçlar

Hald'ın veri kümesi için Liu tahmin edicisi

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

eşitliğinden elde edilmiştir.

Hald veri kümesine ilişkin Liu parametre kestirim değerleri Çizelge 8’de verilmiştir:

**Çizelge 8.** Liu parametre kestirim değerleri

	$\hat{\beta}_{x_1}$	$\hat{\beta}_{x_2}$	$\hat{\beta}_{x_3}$	$\hat{\beta}_{x_4}$
Liu EKK	8.792565	9.031684	-37.484673	-22.764014
Liu Sağlam	0.636295	0.600194	-0.086365	-0.076172
Liu M	-0.302459	-0.248041	0.054882	-0.051336

Liu’ya ilişkin EKK ve sağlam parametre kestirimleri ve HKO’ları aşağıdaki eşitlikler kullanılarak elde edilmiş ve sonuçlar Çizelge 9’da verilmiştir.

$$\hat{\beta}_{Liu.EKK} = (X'X + I)^{-1}(X'y + d_{EKK})\hat{\beta}_{EKK}$$

$$\hat{\beta}_{Liu.sağlam} = (X'X + I)^{-1}(X'y + d_{sağlam})\hat{\beta}_{huber}$$

$$\hat{\beta}_{Liu.M} = (\wedge + I)^{-1}(\wedge + d_{sağlam})\hat{\beta}_{huber}$$

$$HKO(\hat{\beta}_{Liu.EKK}) = \sum_{i=1}^p \frac{\sigma^2(\lambda_i + d_{EKK})^2 + \lambda_i(1 - d_{EKK})^2 \beta_i^2}{\lambda_i(\lambda_i + 1)^2}$$

$$HKO(\hat{\beta}_{Liu.sağlam}) = \sum_{i=1}^p \frac{\sigma^2(\lambda_i + d_{sağlam})^2 + \lambda_i(1 - d_{sağlam})^2 \beta_i^2}{\lambda_i(\lambda_i + 1)^2}$$

$$HKO(\hat{\beta}_{Liu.M}) = \sum_{i=1}^p \frac{\sigma^2(\lambda_i + d_M)^2 + \lambda_i(1 - d_M)^2 \beta_i^2}{\lambda_i(\lambda_i + 1)^2}$$

Liu yönteminde kullanılan d yanlılık parametresi  $d_{EKK}$  ve  $d_{sağlam}$  olmak üzere iki farklı şekilde hesaplanmıştır. Elde edilen d değerleri de Çizelge 9’da verilmiştir.

**Çizelge 9.** Liu tahmin edicisine ilişkin parametre kestirimi ve HKO’ları

Yöntem	Parametre	HKO
$\hat{\beta}_{Liu.EKK}$	$d_{EKK} = 0.922090$	$HKO_{liu_{EKK}} = 1.450127$
$\hat{\beta}_{Liu.sağlam}$	$d_{sağlam} = 0.991421$	$HKO_{liu_{sağlam}} = 1.249062$
$\hat{\beta}_{Liu.M}$	$d_{sağlam} = 0.991421$	$HKO_{liu_M} = 1.235476e-06$

Çizelge 9 incelendiğinde, yanlılık parametresi olarak Huber-M kestiricisinin kullanıldığı

$\hat{\beta}_{Liu.M}$ 'e ilişkin HKO'sının klasiğe dayalı olarak elde edilen HKO'sından daha küçük olduğu görülmektedir. EKK yöntemine ilişkin elde edilen HKO diğer HKO'larına göre daha büyük çıkmıştır.

Aykırı değer yöntemleri üzerindeki etkisini görmek için d değerleri ve HKO'ları aykırı değer durumunda tekrar hesaplanmış ve Çizelge 10'da verilmiştir.

**Çizelge 10.** Aykırı değer varlığında Liu tahmin edicisi

Yöntem	Parametre	HKO
$\hat{\beta}_{Liu.EKK}$	$d_{EKK} = 0.923797$	$HKO_{Liu.EKK} = 28.93737$
$\hat{\beta}_{Liu.sağlam}$	$d_{sağlam} = 0.998994$	$HKO_{Liu.sağlam} = 33.6363$
$\hat{\beta}_{Liu.M}$	$d_{sağlam} = 0.998994$	$HKO_{Liu.M} = 7.752391e-05$

Aykırı değer tüm HKO değerlerini etkilediği Çizelge 10 'dan açıkça görülebilmektedir. Ancak kendi içinde değerlendirdiğimizde yine en küçük HKO'sının Huber-M kestiricili Liu.M yöntemine ait olduğu görülmektedir.

#### 6.1.4. LASSO Tahmin Edicisinden Elde Edilen Sonuçlar

Hald'in veri kümesine ilişkin LASSO parametre kestirim değerleri ve HKO'ları Çizelge 11' de verilmiştir.

**Çizelge 11.** LASSO parametre kestirim değerleri ve HKO'ları

	Aykırı Değersiz LASSO Parametreleri	Aykırı Değerli LASSO Parametreleri
$HKO$	1.081578	2.318041
$\hat{\beta}_{x_1}$	-0.651944	-0.850071
$\hat{\beta}_{x_2}$	-0.602580	-0.953266
$\hat{\beta}_{x_3}$	-0.470394	-0.828320
$\hat{\beta}_{x_4}$	0.268647	0.011836

Çizelge 11'e göre aykırı değer varlığından LASSO tahmin edicisinin çok fazla etkilenmediğini söyleyebiliriz.

Bu tez çalışmasındaki asıl amaç yukarıda verilen yöntemleri karşılaştırmak olduğundan sonuçların tamamı tekrardan Çizelge 12 ve Çizelge 13'de verilmiştir.

**Çizelge 12.** Genel Sonuçlar

Yöntem	Parametre	HKO
Ridge EKK	$k_{EKK} = 0.011623$	$HKO_{k_{EKK}} = 2.000666$
Ridge Sağlam	$k_{sağlam} = 0.010173$	$HKO_{k_{sağlam}} = 0.591642$
Ridge Medyan	$k_{medyan} = 0.005061$	$HKO_{k_{medyan}} = 0.919880$
Liu EKK	$d_{EKK} = 0.922095$	$HKO_{liu_{EKK}} = 1.450127$
Liu Sağlam	$d_{sağlam} = 0.991421$	$HKO_{liu_{sağlam}} = 1.249062$
Liu M	$d_{sağlam} = 0.991421$	$HKO_{liu_{M\beta}} = 1.23547e-06$
LASSO	-	$HKO_{LASSO} = 1.081578$

**Çizelge 13.** Aykırı değer varlığındaki sonuçlar

Yöntem	Parametre	HKO
Ridge EKK	$k_{EKK} = 0.006277$	$HKO_{k_{EKK}} = 1166.265$
Ridge Sağlam	$k_{sağlam} = 0.064717$	$HKO_{k_{sağlam}} = 14.92555$
Ridge Medyan	$k_{medyan} = 0.005061$	$HKO_{k_{medyan}} = 38.82282$
Liu EKK	$d_{EKK} = 0.923797$	$HKO_{liu_{EKK}} = 28.93737$
Liu Sağlam	$d_{sağlam} = 0.9989944$	$HKO_{liu_{sağlam}} = 33.6363$
Liu M	$d_{sağlam} = 0.998994$	$HKO_{liu_{M\beta}} = 7.752391e-05$
LASSO	-	$HKO_{LASSO} = 2.318041$

Çizelge 12' den en küçük HKO' sının, Liu-M kestiricisinin kullanıldığı Liu yöntemine ait olduğu görülmektedir. Bunun yanı sıra EKK Ridge tahmin edicisinin dışında diğer tahmin edicilerinde HKO değerleri küçük çıkmıştır. Aykırı değer varlığından etkilenmeyen tahmin edicilerin Liu-M'li Liu yöntemi ve LASSO olduğunu Çizelge 13'den söyleyebiliriz. Bunun yanı sıra Sağlam Ridge dışında diğer tüm yöntemlerin HKO'ları çok büyümüşür ve aykırı değer varlığından oldukça etkilenmişlerdir.

## 6.2. İstihdam Verileri ile Uygulama

İstihdam veri kümesi 19 gözlem ve 4 bağımsız değişkenden oluşmaktadır.  $Y$  bağımlı değişkendir ve 1998-2006 yılları arasında Türkiye’de istihdam edilen insan sayısını göstermektedir.  $X_1$  değişkeni 14 yaşından büyük sivil nüfusu,  $X_2$  değişkeni işsiz insan sayısını,  $X_3$  değişkeni yılları,  $X_4$  ise GSMH’yı temsil etmektedir (Kan, Alpu ve Yazıcı, 2013). Veri kümesi Çizelge 14’de verilmiştir.

**Çizelge 14.** İstihdam veri kümesi

	$Y$	$X_1$	$X_2$	$X_3$	$X_4$
1	22,328	47,391	2445	2006	5477
2	22,046	46,62	2519	2005	5008
3	21,79	45,813	2502	2004	4172
4	21,146	44,974	2492	2003	3383
5	21,35	44,224	2467	2002	2598
6	21,525	43,455	1969	2001	2123
7	21,582	42,612	1499	2000	2965
8	22,056	41,809	1836	1999	2879
9	21,785	40,915	1611	1998	3255
10	21,207	40,02	1557	1997	3079
11	21,198	39,071	1505	1996	2928
12	20,594	38,115	1704	1995	2759
13	20,011	37,114	1877	1994	2184
14	18,506	36,153	1821	1993	3004
15	19,467	35,277	1810	1992	2708
16	19,294	34,248	1725.5	1991	2621
17	18,541	33,066	1615.5	1990	2682
18	18,222	31,948	1712.5	1989	1959
19	17,754	31,461	1637	1988	1684

### 6.2.1 Çoklu Bağlantı İncelemesi

Veri kümesinde doğrusal bağımlılık olup olmadığını anlamak için çoklubağlantı incelemesi yapılmıştır. Bunun için öncelikle değişkenlere ilişkin ilişki matrisi elde edilmiş ve Çizelge 15’te verilmiştir.

**Çizelge 15.** İlişki matrisi

		$X_1$	$X_2$	$X_3$	$X_4$
$X_1$	Pearson Correlation	1	,696**	,998**	,732**
	Sig. (2-tailed)		,001	,000	,000
	N	19	19	19	19
$X_2$	Pearson Correlation	,696**	1	,731**	,616**
	Sig. (2-tailed)	,001		,000	,005
	N	19	19	19	19
$X_3$	Pearson Correlation	,998**	,731**	1	,745**
	Sig. (2-tailed)	,000	,000		,000
	N	19	19	19	19
$X_4$	Pearson Correlation	,732**	,616**	,745**	1
	Sig. (2-tailed)	,000	,005	,000	
	N	19	19	19	19

İlişki matrisine göre  $X_1$  ve  $X_3$  değişkenleri arasında 0.998 büyüklüğünde,  $X_3$  ve  $X_4$  arasında 0.745 büyüklüğünde ilişki vardır.

Veri kümesine ilişkin aykırı değer incelemesi yapılmış ve  $X$  yönünde bir aykırıdeğer olduğu görülmüştür.

Veri kümesine ilişkin özdeğerleri, özvektör matrisi ve koşul sayıları hesaplanmış ve Çizelgelerde verilmiştir.

**Çizelge 16.** İstihdam verisi için özdeğerler ve koşul sayısı

Özdeğerler ( $\lambda_j$ )	3.274776	0.387140	0.337572	0.000510
Koşul Sayısı ( $KS_j$ )	1	2,908417	3,114638	80,131950

Çizelge 16'ya göre, 0'a yakın olan bir tane özdeğer vardır ve buna karşılık gelen koşul sayısı 30'dan büyük olduğu için çok güçlü bir çoklu bağlantı olduğunu söyleyebiliriz.



**Çizelge 17.** İstihdam verisi için özvektörler matrisi

	[,1]	[,2]	[,3]	[,4]
[1,]	0.528616	0.130124	-0.481046	0.687187
[2,]	0.462939	-0.819968	0.334971	0.033640
[3,]	0.534561	0.099671	-0.421939	-0.725449
[4,]	0.469569	0.548437	0.691632	0.019092

Özvektörler matrisi incelendiğinde  $X_3$  ve  $X_4$  değişkenlerinin doğrusal bağımlı olduğu görülmektedir. Değişkenler arasındaki doğrusal bağımlılık yapısı da aşağıdaki gibidir:

$$-0.725449 x_3 + 0.691632 x_4 \cong 0$$

İstihdam veri kümesi için yöntemlere ilişkin HKO'ları aşağıdaki çizelgede verilmiştir.

### 6.2.2 Ridge Tahmin Edicisinden Elde Edilen Sonuçlar

İstihdam veri kümesi için Ridge tahmin edicisi

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

eşitliğinden elde edilmiştir.

İstihdam veri kümesine ilişkin Ridge parametre kestirim değerleri Çizelge 18 'de verilmiştir:

**Çizelge 18.** Ridge parametre kestirim değerleri

$\hat{\beta}_{x_1}$	$\hat{\beta}_{x_2}$	$\hat{\beta}_{x_3}$	$\hat{\beta}_{x_4}$
0.6065120	0.5277056	0.0433897	-0.1602874

Ridge'e ilişkin EKK ve sağlam parametre kestirimleri ve HKO'ları aşağıdaki eşitlikler kullanılarak elde edilmiş ve sonuçları Çizelge 19'da verilmiştir.

$$\hat{\beta}_{Ridge.EKK} = (X'X + k_{EKK}I_p)^{-1}\hat{\beta}_{EKK},$$

$$\hat{\beta}_{Ridge.saglam} = (X'X + k_{saglam}I_p)^{-1}\hat{\beta}_{EKK},$$

$$HKO\hat{\beta}_{Ridge.EKK} = \sigma^2 \sum_{j=1}^p \frac{\lambda_j \sigma^2 + k_{EKK}\beta_i^2}{(\lambda_j + k_{EKK})^2}$$

$$HKO\hat{\beta}_{Ridge.saglam} = \sigma^2 \sum_{j=1}^p \frac{\lambda_j \sigma^2 + k_{saglam}\beta_i^2}{(\lambda_j + k_{saglam})^2}$$

Ridge yönteminde kullanılan k yanlılık parametresi  $k_{EKK}$ ,  $k_{sağlam}$ ,  $k_{medyan}$  olmak üzere üç farklı şekilde hesaplanmıştır. Eşitlik (3.14) ve (3.32)'den elde edilen k değerleri ve bunlara karşılık gelen HKO'ları Çizelge 19'da verilmiştir.

**Çizelge 19.** Ridge tahmin edicisinden elde edilen sonuçlar

Yöntem	Parametre	HKO
$\hat{\beta}_{Ridge.EKK}$	$k_{EKK} = 0.006974$	$HKO_{k_{EKK}} = 3.211669$
$\hat{\beta}_{Ridge.sağlam}$	$k_{sağlam} = 0.006104$	$HKO_{k_{sağlam}} = 0.824423$
$\hat{\beta}_{Ridge.medyan}$	$k_{medyan} = 0.002909$	$HKO_{k_{medyan}} = 1.450882$

Çizelge 19 incelendiğinde, yanlılık parametresi olarak Huber-M kestiricisinin kullanıldığı  $\hat{\beta}_{Ridge.sağlam}$ 'a ilişkin HKO'sının klasik ve medyana dayalı olarak elde edilen HKO'larından daha küçük olduğu görülmektedir. EKK yöntemine ilişkin HKO'sı diğer HKO'larına göre daha büyük çıkmıştır.

### 6.2.3 Liu Tahmin Edicisinden Elde Edilen Sonuçlar

İstihdam veri kümesi için Liu tahmin edicisi

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

eşitliğinden elde edilmiştir.

İstihdam veri kümesine ilişkin Liu parametre kestirim değerleri Çizelge 20'de verilmiştir:

**Çizelge 20.** Liu parametre kestirim değerleri

	$\hat{\beta}_{x_1}$	$\hat{\beta}_{x_2}$	$\hat{\beta}_{x_3}$	$\hat{\beta}_{x_4}$
Liu EKK	4.196758	4.745540	4.335119	5.552620
Liu Sağlam	5.436626	6.115406	5.600834	7.043015
Liu M	0.336216	-0.491522	-0.036393	-0.001451

Liu'ya ilişkin EKK ve sağlam parametre kestirimleri ve HKO'ları aşağıdaki eşitlikler kullanılarak elde edilmiş ve sonuçlar Çizelge 21'de verilmiştir.

$$\hat{\beta}_{Liu.EKK} = (X'X + I)^{-1}(X'y + d_{EKK})\hat{\beta}_{EKK}$$

$$\hat{\beta}_{Liu.sağlam} = (X'X + I)^{-1}(X'y + d_{sağlam})\hat{\beta}_{huber}$$

$$\hat{\beta}_{Liu.M} = (\Lambda + I)^{-1}(\Lambda + d_{sağlam})\hat{\beta}_{huber}$$

$$HKO(\hat{\beta}_{Liu.EKK}) = \sum_{i=1}^p \frac{\sigma^2(\lambda_i + d_{EKK})^2 + \lambda_i(1 - d_{EKK})^2 \beta_i^2}{\lambda_i(\lambda_i + 1)^2}$$

$$HKO(\hat{\beta}_{Liu.sağlam}) = \sum_{i=1}^p \frac{\sigma^2(\lambda_i + d_{sağlam})^2 + \lambda_i(1 - d_{sağlam})^2 \beta_i^2}{\lambda_i(\lambda_i + 1)^2}$$

$$HKO(\hat{\beta}_{Liu.M}) = \sum_{i=1}^p \frac{\sigma^2(\lambda_i + d_M)^2 + \lambda_i(1 - d_M)^2 \beta_i^2}{\lambda_i(\lambda_i + 1)^2}$$

Liu yönteminde kullanılan d yanlılık parametresi  $d_{EKK}$  ve  $d_{sağlam}$  olmak üzere iki farklı şekilde hesaplanmıştır. Elde edilen d değerleri ve bunlara karşılık gelen HKO'ları Çizelge 21'de verilmiştir.

**Çizelge 21.** Liu tahmin edicisine ilişkin parametre kestirimi ve HKO'ları

Yöntem	Parametre	HKO
$\hat{\beta}_{Liu.EKK}$	$d_{EKK} = 0.790042$	$HKO_{liu_{EKK}} = 1.453934$
$\hat{\beta}_{Liu.sağlam}$	$d_{sağlam} = 0.9982572$	$HKO_{liu_{sağlam}} = 2.326859$
$\hat{\beta}_{Liu.M}$	$d_{sağlam} = 0.9982572$	$HKO_{liu_M} = 8.555155e-07$

Çizelge 21 incelendiğinde, yanlılık parametresi olarak Huber-M kestiricisinin kullanıldığı  $\hat{\beta}_{Liu.M}$ 'e ilişkin HKO'sının EKK ve sağlam olarak elde edilen HKO'larından daha küçük olduğu görülmektedir. Yani en iyi sonucu Huber-M' li Liu yöntemi vermiştir.

#### 6.2.4 LASSO Tahmin Edicisinden Elde Edilen Sonuçlar

İstihdam veri kümesine ilişkin LASSO parametre kestirim değerleri ve HKO'sı Çizelge 22' de verilmiştir.

**Çizelge 22.** LASSO parametre kestirim değerleri ve HKO'sı

LASSO Parametreleri	

$HKO$	0.7701807
$\hat{\beta}_{x_1}$	0.000000
$\hat{\beta}_{x_2}$	8.282951e-01
$\hat{\beta}_{x_3}$	-7.804216e-06
$\hat{\beta}_{x_4}$	2.900135e-01

Tüm sonuçlar Çizelge 23’de tekrardan verilmiştir.

**Çizelge 23.** Genel Sonuçlar

Yöntem	Parametre	HKO
Ridge EKK	$k_{EKK} = 0.006974$	$HKO_{k_{EKK}} = 3.211669$
Ridge Sağlam	$k_{sağlam} = 0.006104$	$HKO_{k_{sağlam}} = 0.824423$
Ridge Medyan	$k_{medyan} = 0.002909$	$HKO_{k_{medyan}} = 1.450882$
Liu EKK	$d_{EKK} = 0.790042$	$HKO_{d_{EKK}} = 1.453934$
Liu Sağlam	$d_{sağlam} = 0.9982572$	$HKO_{d_{sağlam}} = 2.326859$
Liu M	$d_{sağlam} = 0.9982572$	$HKO_{d_{liu_M}} = 8.55515e-07$
LASSO	-	$HKO_{LASSO} = 0.7701807$

Çizelge 23’ den en küçük HKO’ sının, Liu-M kestiricisinin kullanıldığı Liu yöntemine ait olduğu görülmektedir. Liu-M kestiricisinin kullanıldığı Liu yöntemi de iyi bir alternatif olabilir.

## 7. SONUÇ

Bu bölüme değin, giriş bölümünde verilen tezin amacı doğrultusunda genel bilgiler derlenmiş, çoklu bağlantı durumunda kullanılan Ridge, Liu ve LASSO yöntemleri kullanım gerekçeleri, kullanım alanları ele alınarak ayrıntılı biçimde verilmiştir. Ayrıca aykırı değer varlığında daha iyi sonuçlar vereceği düşünülerek bu yöntemlerin sağlam versiyonları da incelenmiştir.

Bu amaçlar doğrultusunda iki veri kümesi üzerinden yanlı kestirim yöntemlerinden olan Ridge, Liu ve LASSO tahmin edicileri HKO'ları üzerinden karşılaştırılmıştır. Bu karşılaştırmalar klasik ve sağlam versiyonları üzerinden yapılmıştır.

Öncelikle çoklu bağlantılı Hald veri kümesi incelenmiş ve yöntemlerin karşılaştırması yapılmıştır. Bu veri için aykırı değerlerin yöntemler üzerindeki etkisini görmek adına  $y$ -yönünde aykırı değer oluşturulmuştur. Böylece hem çoklu bağlantı hem de  $y$ - yönünde aykırı değer olduğu durum için yöntemlerin HKO'ları karşılaştırılmıştır. En küçük HKO'lı tahmin edicinin Liu-M kestiricili Liu yöntemi ve LASSO olduğu görülmüştür. Bu yöntemlerin yanı sıra alternatif yöntem olarak Ridge-sağlam yöntemi de kullanılabileceği söylenebilir.

Diğer bir uygulama olarak gerçek bir veri kümesi kullanılmıştır. Bu veri kümesi içinde yöntemlerin HKO'ları karşılaştırılmıştır. Burada ele alınan İstihdam veri kümesinde çoklu bağlantının yanı sıra  $x$ -yönünde aykırı değer mevcuttur. Bu durumda da en küçük HKO'lı tahmin edicinin Liu-M kestiricili Liu yöntemi ve LASSO olduğu görülmüştür.

Yani bu çalışma sonucunda diyebiliriz ki, bir veride çoklu bağlantı varsa, ve  $x$  ya da  $y$  önünde aykırı değer varsa doğru bir parametre tahmini yapabilmek için, EKK yöntemine alternatif olarak Liu-M kestiricili Liu, LASSO ve sağlam-Ridge yöntemlerinin kullanımı önerilebilir. Ancak bu sonucu bu çalışmada ele aldığımız veri kümeleri için söyleyebiliyoruz. Genelleme yapabilmek için de bir benzetim çalışması yapmak gerekir.

## 8. KAYNAKLAR

- Akdeniz, F., Kaçiranlar, S., On The Almost Unbiased Generalized Liu Estimator And Unbiased Estimation Of The Bias And *MSE*. Communications in Statistics-Theory and Methods, 24(7), 1789-179,( **1995**).
- Akdeniz F., Kaçiranlar S., More on the New Biased Estimator in Linear Regression, The Indian Journal of Statistics, Series B (1960-2002), 63, 3, 321-325,( **2001**).
- Arslan O.,Billor N., Robust Liu estimator for regression based on an M-estimator, Journal of Applied Statistic, 27, 1,39-47,( **2000**).
- Arthur, E., Hoerl and Robert W. Kennard, Ridge regression: Biased Estimation for Nonorthogonsl Problems, Technometrics,12,1,55-67, (**1970**).
- Aydın D.,Uygulamalı Regresyon Analizi, Kavramlar ve R hesaplamaları,**2014**.
- Çetin Candan M., Robust Estimator in Linear Regression Analysisjournal of Applied Statistic, Master of Science, Department of Statistics, Hacettepe University, **1995**.
- Erar, A., Doğrusal Regresyon Analizine Giriş, 2013. Aydın, D., Uygulamalı Regresyon Analizi, **2017**.
- Fidanoğlu, I., İstatistiksel Daraltıcı ( Shrinkage ) Model ve Uygulamaları, Ç.Ü., **2009**.
- Kaçiranlar, S., Sakallıoğlu, S., Liu Ve Temel Bileşenler Regresyon Tahmin Edicilerinin Birleştirilmesi, İstatistik Araştırma Sempozyumu, **2000**.
- Karakaş, S., Çoklu Doğrusal Bağlantı Problemi ve Yanlı Regresyon Tahmincileri, **2008**.
- Kan, B., Alpu, Ö., Yazıcı, B., Robust Ridge and Robust Liu Estimator for Regression Based on the LTS Estimator, **2013**.

Kejian,L., A New Class of Blased Estimate in Linear Regression, Communications in Statistics-Theory and Metthods,(**1993**).

Monsteller,F., Hoaglin,D. C., Turkey, J. W., **1982**, Understanding Robust and Exploratory Data Analysis, John-Wiley and Sons, New York, 447p.

Montgomery, D., Peck, E., Vining, G., Lineer Regression Analysis, **2012**.

Neter vd. **1996**, Fox **1997**, Regresyon Analizinde Kullanılan En Küçük Kareler ve En küçük Medyan Kareler Yöntemlerinin Karşılaştırılması.

Orhunbilge, N., ALBAYRAK, S., Çoklu Doğrusal Bağlantı Halinde En Küçük Kareler Tekniğinin Alternatifi Yanlı Tahmin Teknikleri Ve Bir Uygulama, **2000**.

Sakallioğlu S. , Kaçiranlar S. & Akdeniz F.,Mean Squared Error Comparisons Of Some Biased Regression Estimators , Communications in Statistics - Theory and Methods, 30:2, 347-361, doi: 10.1081/sta-100002036, (**2001**).

Shah, S., Comparison of Some Improved Estimators for Linear Regression Model under Different Conditions, Florida İnternational University, **2015**.

Tibshiran, R. Regression Shrinkage and Selection via the Lasso, January (**1995**).

Wenjang J. Fu., Penalized Regressions: The Bridge Versus the Lasso, Journal of Computational and Graphical Statistics,7:3, 397-416, **1998**.



HACETTEPE ÜNİVERSİTESİ  
FEN BİLİMLERİ ENSTİTÜSÜ  
YÜKSEK LİSANS/~~DOKTORA~~-TEZ ÇALIŞMASI ORJİNALLİK RAPORU

HACETTEPE ÜNİVERSİTESİ  
FEN BİLİMLER ENSTİTÜSÜ  
İSTATİSTİK ANABİLİM DALI BAŞKANLIĞI'NA

Tarih: 08/07/2019

Tez Başlığı / Konusu: DOĞRUSAL REGRESYONDA RIDGE, LIU VE LASSO TAHMİN EDİCİLERİ ÜZERİNE BİR ÇALIŞMA

Yukarıda başlığı/konusu gösterilen tez çalışmamın a) Kapak sayfası, b) Giriş, c) Ana bölümler d) Sonuç kısımlarından oluşan toplam 42 sayfalık kısmına ilişkin, 08/07/2019 tarihinde tez danışmanım tarafından Turnitin adlı intihal tespit programından aşağıda belirtilen filtrelemeler uygulanarak alınmış olan orijinallik raporuna göre, tezimin benzerlik oranı % 4'tür.

Uygulanan filtrelemeler:

- 1- Kaynakça hariç
- 2- Alıntılar hariç/dâhil
- 3- 5 kelimedenden daha az örtüşme içeren metin kısımları hariç

Hacettepe Üniversitesi Fen Bilimleri Enstitüsü Tez Çalışması Orjinallik Raporu Alınması ve Kullanılması Uygulama Esasları'nı inceledim ve bu Uygulama Esasları'nda belirtilen azami benzerlik oranlarına göre tez çalışmamın herhangi bir intihal içermediğini; aksinin tespit edileceği muhtemel durumda doğabilecek her türlü hukuki sorumluluğu kabul ettiğimi ve yukarıda vermiş olduğum bilgilerin doğru olduğunu beyan ederim.

Gereğini saygılarımla arz ederim.

Tarih ve İmza

Adı Soyadı: Ayşe KÜÇÜK

Öğrenci No: N13127757

Anabilim Dalı: İstatistik Anabilim Dalı

Programı: İstatistik

Statüsü:  Y.Lisans  Doktora  Bütünleşik Dr.

22.07.19

**DANIŞMAN ONAYI**

UYGUNDUR.

(Prof. Dr. Meral Çetin)



## ÖZGEÇMİŞ

**Adı Soyadı** : Ayşe KÜÇÜK  
**Doğum yeri** : Adana  
**Doğum tarihi** : 16.02.1990  
**Medeni hali** : Bekâr  
**Yazışma adresi** : Aydınlikevler Mah. Uzayan Sokak 13/8 Altındağ/Ankara  
**Telefon** : 05320617557  
**Elektronik posta adresi** : aysekucuk22@gmail.com  
**Yabancı dili** : İngilizce

### EĞİTİM DURUMU

**Lisans** : Hacettepe Üniversitesi – İstatistik Bölümü  
**Yüksek Lisans** : Hacettepe Üniversitesi – İstatistik Bölümü  
**Doktora** :

### İş Tecrübesi

2016- ----- : Adalet Bakanlığı

### Diğer