



**ÇOKLU BAĞLANTI DURUMUNDA TUTARLI DEĞİŞEN VARYANS
KOVARYANS MATRİSİNE DAYALI TAHMİN EDİCİLER**

Tuba GÜZEY

**YÜKSEK LİSANS TEZİ
İSTATİSTİK ANA BİLİM DALI**

**GAZİ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

TEMMUZ 2023

ETİK BEYAN

Gazi Üniversitesi Fen Bilimleri Enstitüsü Tez Yazım Kurallarına uygun olarak hazırladığım bu tez çalışmada;

- Tez içinde sunduğum verileri, bilgileri ve dokümanları akademik ve etik kurallar çerçevesinde elde ettiğimi,
- Tüm bilgi, belge, değerlendirme ve sonuçları bilimsel etik ve ahlak kurallarına uygun olarak sunduğumu,
- Tez çalışmada yararlandığım eserlerin tümüne uygun atıfta bulunarak kaynak gösterdiğimi,
- Kullanılan verilerde herhangi bir değişiklik yapmadığımı,
- Bu tezde sunduğum çalışmanın özgün olduğunu,

bildirir, aksi bir durumda aleyhime doğabilecek tüm hak kayıplarını kabullendiğimi beyan ederim.

Tuba GÜZEY

26/07/2023

ÇOKLU BAĞLANTI DURUMUNDA TUTARLI DEĞİŞEN VARYANS KOVARYANS MATRİSİNE DAYALI TAHMİN EDİCİLER

(Yüksek Lisans Tezi)

Tuba GÜZEY

GAZİ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

Temmuz 2023

ÖZET

Regresyon analizinde, regresyon katsayılarının tahmininde kullanılan en yaygın yöntem en küçük kareler (*EKK*) yöntemidir. Bu yöntem, sabit varyans varsayımı ve bağımsız değişkenler arasında ilişki olmaması gibi önemli varsayımlara dayanır. Hata teriminin varyansının farklı olması durumunda değişen varyans söz konusudur. Çoklu regresyon modellerinde bağımsız değişkenler arasında lineer bir ilişki bulunması ise çoklu bağlantı olarak ifade edilir. Çoklu bağlantı varlığında *EKK* tahmin edicisi sapmasızdır. Ridge tahmin edicisi ise sapmalı olmasına rağmen *EKK* tahmin edicisinden daha küçük hata kareler ortalamasına sahiptir. Varyansı *EKK*'nın varyansına göre daha küçük olmaktadır. Bu tez çalışmasında, çoklu regresyon modellerinde değişen varyans ve çoklu bağlantı problemi olması durumunda tutarlı değişen varyans kovaryans matrisine (*TDVKM*) dayalı *EKK* ve ridge regresyona ilişkin tahmin ediciler üzerinden elde edilen varyans tahminleri ve hata kareler ortalaması (*HKO*) sonuçları farklı varyans desenleri altında incelenmiştir. Farklı örnek çapları için *TDVKM*'ye dayalı, *EKK* yöntemiyle elde edilen *TDV1*, *TDV2* ve *TDV3* varyans tahminleri ile ridge regresyona ilişkin *RTDV1*, *RTDV2* ve *RTDV3* varyans tahminleri üzerine Matlab programı kullanılarak simülasyon çalışması yapılmıştır. *TDVKM*'ye dayalı ridge regresyon üzerinden elde edilen varyans tahmin edicisi *RTDV3* sonuçlarının, *EKK* yöntemi üzerinden elde edilen varyans tahmin edicisi *TDV3* sonuçlarına göre daha iyi olduğu gözlenmiştir. Ayrıca, sabit varyans ve değişen varyans durumlarında örnek çapı arttıkça tahmin edicilere ilişkin *HKO* değerlerinin düştüğü ve ridge tahmin edicisine ilişkin *HKO* sonuçlarının *EKK* tahmin edicisine ilişkin *HKO* sonuçlarına göre daha düşük olduğu sonucuna varılmıştır.

Bilim Kodu : 20513

Anahtar Kelimeler : Çoklu bağlantı, en küçük kareler, ridge regresyon, tutarlı değişen varyans kovaryans matrisi

Sayfa Adedi : 49

Danışman : Prof. Dr. Meltem EKİZ

ESTIMATORS BASED ON VARYING VARIANCE COVARIANCE MATRIX
IN CASE OF MULTICOLLINEARITY

(M. Sc. Thesis)

Tuba GÜZEY

GAZİ UNIVERSITY

GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

July 2023

ABSTRACT

The most commonly used method in estimating the regression coefficients in regression analysis is the least squares (*LS*) method, which relies on assumptions such as constant variance and independence among independent variables. In the case of different variances of the error term, heteroscedasticity is present. The presence of a linear relationship among independent variables in multivariate regression models is defined as multicollinearity. The presence of multicollinearity, *LS* estimator is unbiased. The ridge estimator has a smaller mean squared error than the *LS* estimator, although it is biased. Its variance is smaller than the variance of *LS*. This thesis study aims to analyze variance estimations and mean square error (*MSE*) results obtained from estimators of *LS* and ridge regression based on the consistent varying variance-covariance matrix (*CVCM*) in the case of varying variance and multicollinearity problem in multiple regression models, under different variance patterns. The simulation was run on Matlab, which is based on the *CVCM* estimators for different sample sizes, the variance estimators *CV1*, *CV2*, and *CV3* related to *LS* method, and the variance estimators *RCV1*, *RCV2*, and *RCV3* related to ridge regression. *RCV3* results observed from the ridge regression based on *CVCM* were lower than the *CV3* results obtained from the *LS* method. In addition, it was concluded that in cases of constant and varying variance, the *MSE* results for estimators were lower as the sample size increased, and the *MSE* results for ridge estimators were lower than the *MSE* results for *LS* estimators.

Science Code : 20513

Key Words : Multicollinearity, least squares, ridge regression, consistent varying variance covariance matrix

Page Number : 49

Supervisor : Prof. Dr. Meltem EKİZ

TEŞEKKÜR

Bu tezin hazırlanmasının her aşamasında değerli fikirleri ile beni yönlendiren ve destek olan tez danışmanım değerli hocam Sayın Prof. Dr. Meltem EKİZ'e, manevi desteklerinden dolayı sevgili aileme, her konuda yanımda olan ve aldığım kararları her zaman destekleyen eşim Ümit Müfit GÜZEY'e sonsuz teşekkürlerimi sunarım.



İÇİNDEKİLER

	Sayfa
ÖZET	iv
ABSTRACT.....	v
TEŞEKKÜR.....	vi
İÇİNDEKİLER	vii
ÇİZELGELERİN LİSTESİ.....	ix
ŞEKİLLERİN LİSTESİ.....	x
SİMGELER VE KISALTMALAR.....	xi
1. GİRİŞ.....	1
2. REGRESYON ANALİZİ.....	5
3. ÇOKLU BAĞLANTI (MULTICOLLINEARTY)	7
3.1. Çoklu Doğrusal Bağlantı Varlığının Tespiti	7
3.2. Çoklu Bağlantıyı Gidermede Kullanılan Yöntemler.....	8
4. RIDGE REGRESYON ANALİZİ.....	11
4.1. Ridge Tahmin Edicisi ve Özellikleri	11
4.2. Ridge Tahmin Edicisinin En Küçük Kareler Tahmin Edicisi ile İlişkisi.....	12
4.3. Ridge Tahmin Edicisinin Beklenen Değeri ve Varyansı	14
4.4. Ridge Parametresinin Tahmini	17
4.5. Ridge Tahmin Edicisine İlişkin Hata Kareler Ortalaması ve Varyans Tahminleri.....	18
5. TUTARLI DEĞİŞEN VARYANS KOVARYANS MATRİSİNE DAYALI TAHMİN EDİCİLER	21
6. UYGULAMA	25
7. SONUÇ VE ÖNERİLER.....	45

Sayfa

KAYNAKLAR	47
ÖZGEÇMİŞ	49



ÇİZELGELERİN LİSTESİ

Çizelge	Sayfa
Çizelge 6.1. Değişken sayısı 4 ve $n = 20, 30, 50, 100, 200$ iken <i>EKK</i> ve ridge tahmin edicilerine ait <i>HKO</i> 'ları	28
Çizelge 6.2. Değişken sayısı 4 iken <i>VD1</i> , <i>VD2</i> ve <i>VD3</i> durumlarında <i>TDVKM</i> dayalı <i>EKK</i> ve ridge regresyona ilişkin <i>TDV1</i> ve <i>RTDV1</i> varyans tahminleri	30
Çizelge 6.3. Değişken sayısı 4 iken <i>VD1</i> , <i>VD2</i> ve <i>VD3</i> durumlarında <i>TDVKM</i> dayalı <i>EKK</i> ve ridge regresyona ilişkin <i>TDV2</i> ve <i>RTDV2</i> varyans tahminleri	31
Çizelge 6.4. Değişken sayısı 4 iken <i>VD1</i> , <i>VD2</i> ve <i>VD3</i> durumlarında <i>TDVKM</i> dayalı <i>EKK</i> ve ridge regresyona ilişkin <i>TDV3</i> ve <i>RTDV3</i> varyans tahminleri	33
Çizelge 6.5. Değişken sayısı 4 ve $n = 20, 30, 50, 100, 200$ iken <i>EKK</i> ve ridge tahmin edicilerine ilişkin $\hat{\beta}$ tahminleri	35
Çizelge 6.6. Değişken sayısı 6 ve $n = 20, 30, 50, 100, 200$ iken <i>EKK</i> ve ridge tahmin edicilerine ait <i>HKO</i> 'ları.....	36
Çizelge 6.7. Değişken sayısı 6 iken <i>VD1</i> , <i>VD2</i> ve <i>VD3</i> durumlarında <i>TDVKM</i> dayalı <i>EKK</i> ve ridge regresyona ilişkin <i>TDV1</i> ve <i>RTDV1</i> varyans tahminleri	37
Çizelge 6.8. Değişken sayısı 6 iken <i>VD1</i> , <i>VD2</i> ve <i>VD3</i> durumlarında <i>TDVKM</i> dayalı <i>EKK</i> ve ridge regresyona ilişkin <i>TDV2</i> ve <i>RTDV2</i> varyans tahminleri	39
Çizelge 6.9. Değişken sayısı 6 iken <i>VD1</i> , <i>VD2</i> ve <i>VD3</i> durumlarında <i>TDVKM</i> dayalı <i>EKK</i> ve ridge regresyona ilişkin <i>TDV3</i> ve <i>RTDV3</i> varyans tahminleri	41
Çizelge 6.10. Değişken sayısı 6 ve $n = 20, 30, 50, 100, 200$ iken <i>EKK</i> ve ridge tahmin edicilerine ilişkin $\hat{\beta}$ tahminleri	43

ŞEKİLLERİN LİSTESİ

Şekil	Sayfa
Şekil 6.1. Değişken sayısı 4 ve farklı örnek çapları için $VD3$ değişen varyans durumunda EKK için HKO	29
Şekil 6.2. Değişken sayısı 4 ve farklı örnek çapları için $VD3$ değişen varyans durumunda ridge için HKO	29
Şekil 6.3. Değişken sayısı 4 iken $VD1$ sabit varyans durumunda $TDVKM$ dayalı EKK ve ridge regresyona ilişkin $TDV3$ ve $RTDV3$ varyans tahminleri	34



SİMGELER VE KISALTMALAR

Bu çalışmada kullanılmış simgeler ve kısaltmalar, açıklamaları ile birlikte aşağıda sunulmuştur.

Simgeler

Açıklamalar

β	Regresyon katsayısı
ε	Hata terimi
X	Bağımsız değişken
Y	Bağımlı değişken
σ^2	Varyans
k	Ridge yan parametresi
n	Örnek çapı

Kısaltmalar

Açıklamalar

EKK	En küçük kareler
HKO	Hata kareler ortalaması
RTDV	Ridge tutarlı değişen varyans
TDV	Tutarlı değişen varyans
TDVKM	Tutarlı değişen varyans kovaryans matrisi
VD	Varyans deseni
VŞD	Varyans şişirme değeri

1. GİRİŞ

Regresyon analizi, aralarında neden-sonuç ilişkisi bulunan iki veya daha fazla nicel değişken arasındaki ilişkiyi ölçmek için kullanılan bir yöntemdir. Doğrusal regresyon modeli bir bağımlı değişken ve bir bağımsız değişkenden oluşuyorsa basit doğrusal regresyon olarak tanımlanır. Regresyon modelinde birden fazla bağımsız değişken olduğunda ise çoklu regresyon olarak ifade edilir.

Regresyon analizinde, regresyon katsayılarının tahmininde kullanılan en yaygın yöntem en küçük kareler (*EKK*) yöntemidir. Çoklu regresyon modelinde bağımsız değişkenler arasında ilişki olması durumu çoklu doğrusal bağlantı olarak tanımlanır. Regresyon katsayıları tahmin edilirken *EKK* yönteminin kullanılabilmesi için bağımsız değişkenler arasında ilişki olmaması varsayımını sağlaması gerekmektedir. Bu varsayım sağlanmadığında, yani bağımsız değişkenler arasında çoklu doğrusal bağlantı olması durumunda yanlış tahmin yöntemlerinin kullanılması önerilir. Ridge regresyon analizi, çoklu doğrusal bağlantı olması durumunda kullanılan yanlış bir analiz yöntemidir. Ridge regresyon analizinde amaç, *EKK* tahmin edicilerine göre daha küçük varyanslı tahmin ediciler elde etmektir. Ridge regresyon analizinin diğer bir amacı ise yan teriminin karesiyle varyansı küçülterek hata kareler ortalamasını (*HKO*) azaltmaktır [1, 2].

EKK yönteminde, sabit varyans varsayımı önemli varsayımlar arasındadır. Değişen varyans varlığı durumunda, tutarlı değişen varyans kovaryans matrisine (*TDVKM*) dayalı varyans tahmin edicileri daha doğru sonuçlar elde edilmesini sağlamaktadır. Çoklu bağlantı ve değişen varyans durumu, *EKK* yönteminin önemli varsayımlarının sağlanmadığını gösterir [3].

Bağımsız değişkenler arasında çoklu bağlantı olması durumunda *EKK* yöntemine alternatif olarak ridge regresyon yöntemi önerilmiştir. Ridge regresyon yönteminde, çoklu regresyon modeline küçük pozitif bir k yan parametresinin eklenmesiyle hata kareler ortalamasını küçültmek amaçlanmaktadır [4]. *TDVKM*'ye dayalı tahmin edicilerin performansları üzerine yapılan çalışmada, değişen varyans varlığında örnek çapının çok küçük olduğu durumlarda iyi sonuçlar veren ve *TDV3* olarak tanımlanan tahmin edicinin kullanılması önerilmiştir [5]. *EKK* yönteminin en önemli varsayımlarından biri sabit varyans varsayımı

olup bu varsayım sağlanmadığında yansızlık ilkesinin bozulduğu belirtilmiştir. Değişen varyans durumunda parametre tahminlerinin tutarsız olabileceği ve bu durumda *TDVKM*'ye dayalı tahmin edicilerin kullanılması önerilmektedir [6]. Çoklu bağlantı varlığında hem istatistiksel olarak anlamlı hem de teorik olarak tutarlı tahminler sağlayan bir yaklaşım sunulmuştur. Bu yaklaşımın, temel istatistiksel yöntemlere ek olarak bir ön bilgi içeren sağlam varsayım yaklaşımı olduğu ifade edilmektedir [7]. Bağımlı ve bağımsız değişkenler arasındaki korelasyona dayalı olarak ridge parametresini tahmin etmek için bazı yeni tahmin edicilerin performansı üzerine yapılan bir çalışmada, hatalara ilişkin değişen varyans durumunda farklı k ridge parametresine ilişkin önerilen tahmin edicilerin daha iyi sonuçlar verdiği belirtilmiştir [8]. Çoklu doğrusal bağlantı problemi varlığında farklı ridge tahmin edicileri performansları üzerine yapılan bir çalışmada, *EKK* tahmin edicisine göre en iyi performans gösteren ridge tahmin edicisi belirlenmiştir. Simülasyon sonuçlarına göre, ridge tahmin edicilerinin *EKK* tahmin edicisinden daha küçük *HKO*'ya sahip olduğu ifade edilmiştir [9]. Bağımsız değişkenler arasında çoklu bağlantı olması durumunda *EKK* yöntemi ile elde edilen β katsayılarının olduğundan daha büyük ve bazı katsayıların ters işaretli olduğu belirtilmektedir. Ridge regresyon yöntemi ile elde edilen β katsayılarının büyük ve ters işaretli olma durumu düzeltilmektedir. Bu durumda ridge regresyon analizinin kullanılması önerilmektedir [10]. Bağımsız değişkenler arasındaki korelasyonu arttırmanın *HKO* üzerinde olumsuz etki ettiği belirtilmektedir. Ancak, bağımsız değişken sayısının artmasının *HKO* üzerinde olumlu etkisi olduğunu ve örnek çapı arttığında bağımsız değişkenler arasındaki korelasyon büyük olsa bile *HKO*'sının azaldığı ifade edilmektedir [11]. Çoklu bağlantı sorununun bazı önemli değişkenleri, istatistiksel olarak anlamsız kıldığı belirtilmiş olup çoklu bağlantıyı tespit etmek için korelasyon katsayıları, varyans şişirme faktörü ve özdeğer yöntemi üzerine çalışma yapılmıştır. Ayrıca çoklu bağlantı varlığında temel bileşenler regresyonu, ağırlıklı regresyon ve ridge regresyon yönteminin kullanılabileceği belirtilmiştir [12].

Çoklu bağlantı olması durumunda; ridge regresyon analizi ile 1. tip hata, testin gücü ve değişen varyans sorunu üzerine yapılan bir çalışmada doğrusal regresyon modeli $Y = X\beta + \lambda(X)\varepsilon$ olarak belirtilmiştir. Burada $\lambda(X)$, X matrisini n boyutlu bir köşegen matrisine iz düşüren bilinmeyen bir fonksiyondur. $\lambda(X)$ için üç varyans desenini *VD1*, *VD2* ve *VD3* olarak tanımlanmıştır. Burada *VD1* sabit varyans deseni, *VD2* ve *VD3* değişen varyans desenleri olarak ele alınmıştır. Korelasyon katsayısı yüksek olduğunda,

yani bağımsız değişkenler arasında güçlü ilişki olduğunda ridge tahmin edicisine dayalı 1. tip hata olasılıklarının *EKK* tahmin edicisine dayalı 1. tip hata olasılıklarından daha düşük olduğu gösterilmiştir. Bu durumda, ridge tahmin edicisi kullanıldığında hipotez doğru iken reddedilme olasılığı daha düşüktür. Ridge tahmin edicisine dayalı bir dizi hipotez testi yapılırken en az bir kez yanlış karar verme olasılığının *EKK* tahmin edicisine göre daha düşük olduğu gösterilmiştir. Çalışmada, korelasyon katsayısı arttıkça ridge tahmin edicisinin güç seviyesinin *EKK* tahmin edicisine göre daha iyi sonuçlar verdiğini ortaya konulmuştur [13].

Değişen varyans durumunda *EKK* tahmin edicileri yanlış çıkarımlara yol açabilir. *TDVKM*'ye dayalı tahmin ediciler, değişen varyans durumunda daha iyi sonuçlar verir. *TDVKM*'nin küçük örneklerdeki ($n \leq 250$) performanslarına ilişkin *TDV1*, *TDV2* ve *TDV3* olarak tanımlanan üç alternatif tahmin edici geliştirilmiştir. *TDVKM*'ye dayalı örnek çapının 25'e kadar küçük olduğu durumlarda bile doğru çıkarımlar veren tahmin edici *TDV3*'dür. Değişen varyans olması durumunda ve n küçük iken *TDV3* olarak tanımlanan tahmin edicinin kullanılması önerilmektedir [3].

Tezin önemi

Bu çalışmada çoklu bağlantı varlığında, sabit ve değişen varyans durumlarında *EKK* tahmin edicisi ile ridge tahmin edicine ait *HKO* sonuçları incelenmiştir. Ayrıca, *TDVKM*'ye dayalı *EKK* ve ridge regresyona ait varyans tahminleri farklı varyans desenleri altında karşılaştırılmıştır. Çoklu bağlantı sorunu olması durumunda, yanlış tahmin yöntemlerinden olan ridge regresyon yöntemi ve *EKK* yöntemiyle elde edilen $\hat{\beta}$ tahminleri incelenmiştir.

Simülasyon çalışmasıyla, çoklu bağlantı varlığında sabit ve değişen varyans durumlarında *TDVKM*'ye dayalı, *EKK* için *TDV1*, *TDV2*, *TDV3* varyans tahmin edicileri ile ridge için *RTDV1*, *RTDV2*, *RTDV3* varyans tahmin edicileri incelenmiştir. Ayrıca, bağımsız değişkenler arasında çoklu bağlantı olması durumunda farklı örnek çapları ile elde edilen *HKO* sonuçları değerlendirilmiştir. Sabit ve değişen varyans durumlarında örnek çapı arttıkça tahmin edicilere ilişkin *HKO* değerlerinin düştüğü ve ridge tahmin edicisine ilişkin *HKO* sonuçlarının *EKK* tahmin edicisine ilişkin *HKO* sonuçlarına göre daha düşük olduğu

sonucuna varılmıştır. *TDVKM*'ye dayalı olarak farklı örnek çapları için *EKK* ve ridge regresyona ait varyans tahminleri karşılaştırılmış olup ridge regresyona ait *RTDV3* sonuçlarının daha iyi olduğu tespit edilmiştir. Çoklu bağlantı durumunda *EKK* ve ridge yöntemi ile elde edilen $\hat{\beta}$ tahminleri karşılaştırılmıştır. *EKK* yöntemi ile elde edilen $\hat{\beta}$ katsayılarının, ridge yöntemi ile elde edilen katsayılardan daha büyük ve ters işaretli olduğu durumlar gözlenmiştir.

Tezin içeriği

Bu çalışmada, farklı varyans desenleri altında çoklu bağlantı olması durumunda ridge tahmin edicisi ile *EKK* tahmin edicisi *HKO*'ları karşılaştırılmış ve *TDVKM*'ye dayalı varyans tahminleri incelenmiştir. Bu çalışmanın ikinci bölümünde regresyon analizi ve bu analizin varsayımları açıklanmıştır. Üçüncü bölümde; çoklu bağlantı sorunu, çoklu doğrusal bağlantı varlığının tespit edilmesi ve çoklu bağlantıyı gidermede kullanılan yöntemler belirtilmiştir. Dördüncü bölümde ridge regresyon analizi ve özellikleri incelenmiştir. Beşinci bölümde, *TDVKM*'ye dayalı tahmin ediciler ele alınmıştır. Altıncı bölümde ise değişen varyans durumunda ve çoklu bağlantı varlığında, *TDVKM*'ye dayalı *EKK* yöntemi üzerinden *TDV1*, *TDV2*, *TDV3* varyans tahminleri ile ridge yöntemi üzerinden *RTDV1*, *RTDV2*, *RTDV3* varyans tahminleri elde edilmiştir. Matlab programı kullanılarak *TDVKM*'ye dayalı varyans tahminleri ve *HKO*'ları için simülasyon çalışması yapılmış ve sonuçlar özetlenmiştir.

2. REGRESYON ANALİZİ

Bir bağımlı değişken (Y) ve birden fazla bağımsız değişken ($X_1, X_2, \dots, X_n$) arasındaki fonksiyonel ilişki çoklu regresyon analizi ile incelenir. Regresyon analizinde amaç bağımsız değişkenlerden bağımlı değişkeni tahminleyebilen en iyi modelin oluşturulmasıdır. Diğer bir ifadeyle bağımlı değişkenin hangi bağımsız değişkenlerden daha çok etkilendiğini saptamaktır. Bunun için oluşturulacak modeldeki parametrelerin tahmini için en çok kullanılan yöntemlerden birisi en küçük kareler (*EKK*) tahmin yöntemidir. *EKK* yöntemi, hatalara ait kareler toplamını en küçük yapan parametre tahmincilerinin seçildiği yöntemdir. Çoklu regresyon analizinde parametrelerin tahmininde yaygın olarak kullanılan *EKK* yönteminin bazı varsayımları sağlanması gerekir.

Bu varsayımlar;

- i. Bağımsız değişkenler arasında çoklu bağlantının olmaması,
 - ii. Hata teriminin beklenen değerinin sıfır olması,
 - iii. Hata teriminin varyansının sabit olması,
 - iv. Hata terimi değerleri arasında bir ilişkinin (otokorelasyonun) olmaması,
 - v. Hata terimi (ε_i) ile bağımsız değişkenler (X_i) arasındaki kovaryansın sıfıra eşit olması yani ε_i ile bağımsız değişkenler arasında bir ilişki olmaması,
- şeklinde ifade edilebilir [14].

Çok değişkenli doğrusal regresyon analizinde en çok karşılaşılan problemlerden biri bağımsız değişkenler arasında doğrusal ilişkinin bulunması yani çoklu doğrusal bağlantı durumunun olmasıdır. Bağımsız değişkenlerden biri diğer bağımsız değişken veya değişkenlerin bir doğrusal fonksiyonu olarak ifade ediliyorsa bağımsız değişkenler arasında doğrusal ilişkinin varlığından söz edilir ve bu durum çoklu bağlantı sorununu ortaya çıkarır. Çoklu doğrusal bağlantı olması durumunda *EKK* yöntemi kullanılamadığından yanlış tahmin yöntemlerinden ridge regresyon analizinin kullanılması önerilmiştir [2].



3. ÇOKLU BAĞLANTI (MULTICOLLINEARITY)

Regresyon modeli, $Y = \beta_0 + \beta_1 X_1 + \dots + \beta_p X_p + \lambda(X)\varepsilon$ şeklinde verilmiş olsun. Açıklayıcı değişkenlerden en az bir tanesi diğer açıklayıcı değişkenlerin bir lineer birleşimi ise, çoklu bağlantı problemi söz konusudur. Böyle bir durumda normal denklemlerin çözümü ya yoktur ya da birden fazla çözümü vardır.

Regresyon modeli,

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

biçiminde ifade edilebilir. Burada Y , $n \times 1$ boyutlu bağımlı değişken vektörü, X , $n \times p$ boyutlu bağımsız değişkenler matrisi, β , $p \times 1$ boyutlu bilinmeyen katsayılar vektörü ve ε , $n \times 1$ boyutlu hata terimi vektörüdür. $\lambda(X)$, X matrisini n boyutlu bir köşegen matrisine iz düşüren bilinmeyen bir fonksiyondur. Hataların dağılımı, ortalaması $0_{n \times 1}$ ve varyansı $\sigma^2 I_n$ olmak üzere normal dağılımdır [13].

X matrisinin sütunlarından en az bir tanesi diğer sütunların bir lineer birleşimi olarak yazılabiliyorsa, $X'X$ matrisi singülerdir. Yani, $X'X$ matrisinin determinanı sıfırdır. Bazen, $X'X$ matrisinin determinanı sıfıra çok yakın olabilir.

Açıklayıcı değişkenlerin sayısı az olduğunda, açıklayıcı değişkenler arasında bir ilişki varsa bunun ortaya çıkarılması kolaydır. Ancak, regresyonda açıklayıcı değişkenlerin sayısı fazla ise bunun görülmesi kolay olmayabilir. Çoklu bağlantının varlığını ortaya çıkaran birçok gösterge vardır [15].

3.1. Çoklu Doğrusal Bağlantı Varlığının Tespiti

- i. Bağımsız değişkenler arasındaki korelasyon matrisinden faydalanılır. Eğer değişkenler arasındaki korelasyon katsayısının mutlak değeri 1'e yakınsa o bağımsız değişkenler arasında çoklu doğrusal bağlantının olduğu ifade edilir.
- ii. Varyans şişirme değerleri incelenir ($V\hat{S}D$); genelde $V\hat{S}D$ değeri 10'un üzerinde

- olduğunda bağımsız değişkenler arasında çoklu bağlantının varlığı kabul edilir.
- iii. Bağımsız değişkenlere ait korelasyon matrisinin determinanı 0 ile 1 arasında değişir ve çoklu doğrusal bağlantının belirlenmesinde kullanılabilir. Determinant değeri küçüldükçe çoklu bağlantının derecesi artar.
 - iv. Korelasyon matrisinin bir veya daha fazla karakteristik kökü (özdeğerleri) sıfır veya sıfıra yakın ise çoklu bağlantının olduğu söylenir.
 - v. $X'X$ matrisinin özdeğerlerinin en büyüğünün en küçüğüne oranı da $(\lambda_{\max}/\lambda_{\min})$ çoklu doğrusal bağlantının bir ölçüsü olarak kullanılır ve bu oran büyüdükçe çoklu doğrusal bağlantının derecesi de artar. Genelde $(\lambda_{\max}/\lambda_{\min})$ oranı 10'dan küçük olduğunda veri setinde önemli boyutta çoklu bağlantının olmadığı ifade edilir.
 - vi. Regresyon modelinin anlamlılığı konusunda bilgi veren F istatistiği ile regresyon katsayılarına ilişkin t -istatistikleri uyumlu olmayabilir. F testi sonucuna göre modelin anlamlı olduğu durumlarda katsayılara ilişkin t testleri anlamsız sonuçlar verebilir. Böyle bir durum çoklu bağlantıya bir işarettir [16].

3.2. Çoklu Bağlantıyı Gidermede Kullanılan Yöntemler

- i. Örnek genişliğinin artırılması; parametre tahminlerine ait varyansların küçülmesini sağlayabilir. Örnek genişliğinin artırılmasıyla parametre tahminlerine ait varyansların küçülebileceği bilinmektedir. Dolayısıyla, varyansların örneklem büyüklüğüyle ters orantılı olduğu söylenebilir [1].
- ii. Bağımsız değişkenlerden bazılarının modelden çıkarılması; aşamalı regresyon yöntemi kullanılarak bağımlı değişken üzerinde daha az etkili olan bağımsız değişkenler regresyon modelinden çıkarılabilir. Ancak bu yol çoklu bağlantının örneklemeden geldiği durumlarda sakıncalıdır. Çünkü gerçekte önemli bir değişken, başka değişkenlerle yapay olarak ilişkili gözüktüğünden modelden çıkarılırsa yanlış model üzerinden çıkarsamalara yol açmış olur. Ayrıca, değişkenlerin modelden çıkarılmasıyla modele gerçekten katkı yapan bir değişken çıkarılmış olabilir ki bu da bilgi kaybına sebep olur. Modeldeki değişkenleri çıkarmadan regresyon katsayılarını tahmin etmek için genellikle çoklu doğrusal bağlantıya en küçük kareler yönteminden daha az duyarlı olan yöntemler kullanılır. Bu yöntemlerin başında da yanlış tahminler veren ridge regresyon ve temel bileşenler regresyon yöntemleri gelir.
- iii. Bağımsız değişkenlerin birleştirilmesi; sorunun gerçek bir ilişkiden kaynaklandığı

durumlarda birbiriyle bağlantılı olan bağımsız değişkenlerden bazılarının birleştirilerek tek bir değişken olarak modelde yer alması düşünülebilir. Ancak tercih edilmesi genellikle uygun olmayan bir seçenektir. Çoklu bağlantı yaratan değişkenler tek bir değişken olarak kombine edilebilir. Örneğin; Fen, Matematik sınav sonuçları aralarında çoklu bağlantı olduğu belirlenen iki değişken olsun. Sınav sonuçları birleştirilerek iki değişkenin ortalaması hesaplanır ve bu yeni hesaplanan değer ilgili regresyonda tek bir değişken olarak kullanılırsa problem çözülebilir. Başka bir örnek ise ağırlık ve boy değişkenleri yerine vücut kitle indeksinin kullanılmasıdır.

- iv. Çoklu bağlantı durumunda çoklu bağlantıdan etkilenmeyen yanlı tahmin yöntemleri kullanılabilir. Bağımsız değişkenler arasında ilişki olduğunda *EKK* tahmin edicisi, en küçük varyansa sahip tahmin edici özelliğini kaybetmektedir. Böyle durumlarda β , yansız tahmin edici $\hat{\beta}$ 'dan daha küçük varyansa sahip olan $\hat{\beta}^*$ yanlı tahmin edicisi ile tahmin edilir. Yanlı tahmin yöntemlerinde amaç, yanlılık terimi kullanarak varyansı daha küçük bir tahmin edici bulmaktır. Yanlı tahmin yöntemleri; temel bileşenler regresyonu, özdeğerler regresyonu ve ridge regresyondur [17].

Çoklu bağlantı, bağımsız değişkenler arasında yüksek korelasyon olduğunda ortaya çıkar. Bu durum, *EKK* ile elde edilen tahminleri olumsuz etkiler. Ridge regresyon, çoklu bağlantı sorununun olduğu durumlarda önerilen bir regresyon yöntemidir. Bağımsız değişkenler arasında doğrusal ilişki olması durumu yanlış çıkarımlara neden olabilir. Bu durumda *EKK* yöntemine göre daha doğru sonuçlar veren ridge regresyon yöntemi kullanılabilir.



4. RIDGE REGRESYON ANALİZİ

EKK'nın yansız $\hat{\beta}$ tahmin ediciler vermesi ve tüm doğrusal yansız $\hat{\beta}$ tahmin edicilerinin en küçük varyanslı olması istenen bir özelliktir. Ancak, çoklu bağlantı durumunda olması gerekenden daha büyük varyanslar ortaya çıkabilmektedir. Büyük varyans oluşumu sorununu azaltmanın bir yolu, $\hat{\beta}$ tahmin edicisinin yansız olma gereksinimini ortadan kaldıran yöntemlerden olan ridge tahmin edicisidir [18].

Ridge regresyon analizi, çoklu doğrusal bağlantı olması durumunda en küçük varyansla parametre tahminini amaçlayan yanlı bir tahmin yöntemidir. Ridge regresyonun özelliklerinin özeti aşağıda verilmiştir;

- i. Kuvvetli çoklu bağlantı olması durumunda katsayılardaki farklılıklar katsayı grafiği üzerinde görsel olarak izlenebilir,
- ii. Çoklu regresyon modelinde bağımsız değişkenler birbirleriyle ilişkili olduğunda *EKK* tahmininden daha küçük varyanslı tahmin ediciler elde edilir,
- iii. Bağımsız değişkenler arasında oluşan çoklu bağlantı sorunundan etkilenmeyen tahmin ediciler sunar,
- iv. Sapma miktarı artmasına karşın; varyansın küçük olmasından dolayı ridge regresyona ilişkin *HKO*, *EKK* yöntemine ait *HKO*'dan daha düşüktür [1,2].

Hoerl ve Kennard (1970) önerdikleri ridge tahmin edicisi ile *EKK* yöntemiyle yapılan tahminlerden daha küçük *HKO*'sına sahip regresyon katsayılarını hesaplamışlardır. *HKO*, parametreyi tahmin etmek için kullanılabilen tahmin edici ile parametre arasındaki farkların karelerinin ortalaması şeklinde ifade edilebilir.

4.1. Ridge Tahmin Edicisi ve Özellikleri

Ridge tahmin edicisi, klasik *EKK* tahmin edicisine benzerdir. Ancak, ridge regresyon yönteminde, $X'X$ matrisinin köşegen elemanlarına pozitif ve küçük bir k sabitinin eklenmesiyle ridge tahmin edicileri elde edilmesi amaçlanmaktadır. Buna göre ridge regresyon çözümü aşağıdaki gibidir [1].

Çoklu doğrusal regresyon modeli Eş. 3.1'deki gibi tanımlansın. Regresyon modelinde tüm değişkenler standartlaştırıldığında yani her bir değişkenin aldığı değer ortalama değerden çıkarılıp standart sapmaya bölüldüğünde, $X'X$ matrisi korelasyon matrisi formunu alır.

Bu modelde β 'nın EKK tahmin edicisi $\hat{\beta}_{EKK}$ aşağıdaki gibidir:

$$\hat{\beta}_{EKK} = (X'X)^{-1} X'Y \quad (4.1)$$

Bağımsız değişkenler arasında yüksek doğrusal ilişki olması durumunda β 'nın EKK tahminlerinin güvenilirliği azalır. Ridge regresyon yönteminde $X'X$ matrisi korelasyon matrisi formunda olmak üzere $X'X$ matrisi köşegen elemanlarına küçük bir k değerinin ($k \geq 0$) eklenmesiyle ridge tahmin edicisi elde edilir [19]. Çoklu bağlantıdan etkilenmeyen ridge tahmin edicisi,

$$\hat{\beta}_R = (X'X + kI)^{-1} X'Y \quad (4.2)$$

ile elde edilir. Burada k ridge parametresi, I ise pxp boyutlu birim matrisidir. $k=0$ iken ridge tahmin edicisi EKK tahmin edicisine dönüşür. Ridge parametresi k , yan miktarı olarak da isimlendirilebilir. $X'X$ matrisinin köşegen elemanlarına pozitif ve küçük bir k değeri eklendiğinde ($k > 0$) çoklu bağlantıdan etkilenmeyen tahmin edicilerin elde edilmesi sağlanır [4].

4.2. Ridge Tahmin Edicisinin En Küçük Kareler Tahmin Edicisi ile İlişkisi

EKK tahmin edicisi Eş. 4.1'de verilmiştir. Burada eşitliğin her iki tarafı $X'X$ ile çarpıldığında aşağıdaki denklem elde edilir;

$$X'Y = X'X \hat{\beta}_{EKK} \quad (4.3)$$

Ridge tahmin edicisi Eş. 4.2'de verilmiştir. Eş. 4.2'de $X'Y$ yerine $X'X \hat{\beta}_{EKK}$ yazıldığında;

$$\hat{\beta}_R = (X'X + kI)^{-1} X'X \hat{\beta}_{EKK} \quad (4.4)$$

ile tahmin edilir.

$X'X$ matrisinin tersinin tersi kendisi olduğundan,

$$\hat{\beta}_R = (X'X + kI)^{-1} \left[(X'X)^{-1} \right]^{-1} \hat{\beta}_{EKK} \quad (4.5)$$

şeklinde ifade edilebilir. Buradan da,

$$\hat{\beta}_R = \left[(X'X)^{-1} (X'X + kI) \right]^{-1} \hat{\beta}_{EKK} \quad (4.6)$$

yazılabilir. Eş. 4.6'daki $\hat{\beta}_R$ formülünden hareketle katsayı tahminleri Eş. 4.7 veya Eş. 4.8'deki gibi ifade edilir.

$$\hat{\beta}_R = \left[(X'X)^{-1} X'X + k(X'X)^{-1} \right]^{-1} \hat{\beta}_{EKK} \quad (4.7)$$

$$\hat{\beta}_R = \left[I + k(X'X)^{-1} \right]^{-1} \hat{\beta}_{EKK} \quad (4.8)$$

Eş. 4.8'deki formülde Z matrisi,

$$Z = \left[I + k(X'X)^{-1} \right]^{-1} \quad (4.9)$$

olarak tanımlanırsa;

$$\hat{\beta}_R = Z \hat{\beta}_{EKK} \quad (4.10)$$

şeklinde elde edilir. Bu eşitlik ridge tahmin edicisinin en küçük kareler tahmin edicisinin bir dönüşümü olarak yazılabildiğini göstermektedir [19].

Eş. 4.2'deki ridge regresyon tahmin edicisinde;

$$W = (X'X + kI)^{-1} \quad (4.11)$$

olarak tanımlandığında,

$$\hat{\beta}_R = WX'Y \quad (4.12)$$

şeklinde yazılabilir. W matrisinin özdeğerleri, $i = 1, \dots, p$ olmak üzere;

$$\tau_i(w) = \frac{1}{\lambda_i + k} \quad (4.13)$$

ile elde edilir. λ_i , $X'X$ ($p \times p$ boyutlu) matrisinin özdeğeridir.

Z matrisinin özdeğerleri ise $\tau_i(z)$ olmak üzere,

$$\tau_i(z) = \frac{\lambda_i}{\lambda_i + k} \quad (4.14)$$

şeklindedir [19].

4.3. Ridge Tahmin Edicisinin Beklenen Değeri ve Varyansı

Ridge tahmin edicisinin EKK tahmin edicisinin bir dönüşümü olarak yazılabildiği

Eş. 4.10'da verilmiştir. $\hat{\beta}_R$ 'in beklenen değeri;

$$E(\hat{\beta}_R) = E\left[(Z\hat{\beta}_{EKK})\right] \quad (4.15)$$

yazılabilir. Eş. 4.9'da verilen Z matrisi,

$$Z = (X'X + kI)^{-1} X'X \quad (4.16)$$

şeklinde de ifade edilebilir. Burada Z matrisinin tersi alındığında;

$$\begin{aligned}
Z^{-1} &= (X'X)^{-1} (X'X + kI) \\
&= \left((X'X)^{-1} X'X + k(X'X)^{-1} \right) \\
&= \left(I + k(X'X)^{-1} \right)
\end{aligned} \tag{4.17}$$

olur. Dolayısıyla da

$$(Z^{-1})^{-1} = Z = \left(I + k(X'X)^{-1} \right)^{-1} \tag{4.18}$$

dır. Eş. 4.16'da $W = (X'X + kI)^{-1}$ olmak üzere,

$$Z = WX'X \tag{4.19}$$

formunda da yazılabilir [19]. Böylece $\hat{\beta}_R$ 'in beklenen değeri,

$$E(\hat{\beta}_R) = E[Z\hat{\beta}_{EKK}] \tag{4.20}$$

$$E(\hat{\beta}_R) = (X'X + kI)^{-1} X'X \hat{\beta}_{EKK} \tag{4.21}$$

şeklindedir. Burada Eş. 4.21'ün her iki tarafı soldan $W^{-1} = \left[(X'X + kI)^{-1} \right]^{-1}$ ile çarpıldığında,

$$\left[(X'X + kI)^{-1} \right]^{-1} E(\hat{\beta}_R) = X'X \hat{\beta}_{EKK} \tag{4.22}$$

elde edilir. Eş. 4.22'in sağına $kI\hat{\beta}_{EKK}$ eklenip çıkartılarak;

$$\left[(X'X + kI)^{-1} \right]^{-1} E(\hat{\beta}_R) = X'X \hat{\beta}_{EKK} + kI\hat{\beta}_{EKK} - kI\hat{\beta}_{EKK} \tag{4.23}$$

$$\left[(X'X + kI)^{-1} \right]^{-1} E(\hat{\beta}_R) = (X'X + kI) \hat{\beta}_{EKK} - kI \hat{\beta}_{EKK} \quad (4.24)$$

bulunur. Burada Eş. 4.24 soldan $(X'X + kI)^{-1}$ ifadesi ile çarpılırsa ridge tahmin edicisinin beklenen değeri $E(\hat{\beta}_R)$;

$$(X'X + kI)^{-1} \left[(X'X + kI)^{-1} \right]^{-1} E(\hat{\beta}_R) = (X'X + kI)^{-1} \left[(X'X + kI) \hat{\beta}_{EKK} - kI \hat{\beta}_{EKK} \right] \quad (4.25)$$

$$E(\hat{\beta}_R) = \left[\hat{\beta}_{EKK} - k(X'X + kI)^{-1} \hat{\beta}_{EKK} \right] \quad (4.26)$$

şeklinde de ifade edilir [20].

Ayrıca v_i , $X'X$ matrisin özdeğerlerine karşılık gelen özvektörler olmak üzere;

$$E(\hat{\beta}_R) = \beta - k \sum_{i=1}^p (\lambda_i + k)^{-1} v_i v_i' \beta \quad (4.27)$$

olarak yazılabilir.

Eş. 4.1'de verilen $\hat{\beta}_{EKK}$ 'nin varyansı Eş. 4.28'deki gibidir.

$$\text{Var}(\hat{\beta}_{EKK}) = \sigma^2 (X'X)^{-1} \quad (4.28)$$

Eş. 4.2'de belirtilen ridge tahmin edicisinin varyansı; $\text{Var}(Y) = \sigma^2$ olmak üzere [21],

$$V(\hat{\beta}_R) = V \left[(X'X + kI)^{-1} X'Y \right] \quad (4.29)$$

$$V(\hat{\beta}_R) = (X'X + kI)^{-1} X' \text{Var}(Y) X (X'X + kI)^{-1} \quad (4.30)$$

$$V(\hat{\beta}_R) = \sigma^2 (X'X + kI)^{-1} X'X (X'X + kI)^{-1} \quad (4.31)$$

şeklinde elde edilir. Özdeğerler ve özvektörler cinsinden $V(\hat{\beta}_R)$,

$$V(\hat{\beta}_R) = \sigma^2 \sum_{i=1}^p \lambda_i (\lambda_i + k)^{-2} v_i v_i' \quad (4.32)$$

olarak yazılabilir [19].

4.4. Ridge Parametresinin Tahmini

Ridge tahmin edicisinin *HKO*'sının, *EKK*'nin *HKO*'sından daha küçük olması için küçük bir k parametresi belirlenmelidir. Bu amaçla k parametresinin belirlenmesinde bazı yöntemler kullanılmaktadır ve bu yöntemler aşağıda incelenmiştir.

- i. Ridge izi yöntemi; ridge regresyon modeline ait k parametre değerinin tahmini amacıyla kullanılan grafiksel bir yöntemdir. Bu amaçla regresyon katsayıları $\hat{\beta}_R$ 'lar düşey eksenle k değerleri yatay eksenle yer almak üzere iki boyutlu grafik çizilir. k 'nın farklı değerlerine göre $\hat{\beta}_R$ katsayıları hesaplanarak çizilen grafik üzerinden hangi katsayıların daha çok değiştiği belirlenebilir. Bağımsız değişkenler arasında yüksek ilişki olması durumunda $\hat{\beta}_R$ katsayıları, k 'da meydana gelen küçük artışlarda çok hızlı değişecektir. k 'nın büyük değerler alması durumunda $\hat{\beta}_R$ katsayıları kararlılık gösterme eğiliminde olacaktır.
- ii. k 'nın nokta kestirimi yöntemi; ridge regresyon modeline ait k parametre değerinin belirlenmesi için özdeğere dayalı bir formül önerilmiştir. Bu formülde $k \neq 0$ olmak üzere k sabitinin belirlenmesinde en büyük özdeğer ile en küçük özdeğere dayalı koşul indeksi önerilir. Koşul indeksinde Eş. 4.11'de verilen W matrisinin özdeğerleri kullanılmak üzere;

$$k \leq \frac{\lambda_{\max} - 100\lambda_{\min}}{99} \quad (4.33)$$

denkleminde göre elde edilir [2].

- iii. Basit ridge tahmin yöntemi, bütün k_j 'ler yerine tek bir k değerini seçmek için önerilen bir yöntemdir. Bu yöntemde $j=1,...,p$ olmak üzere k_j 'ler uygun bir şekilde birleştirilerek tüm katsayılar için tek bir k sapma parametresi,

$$k = \frac{p\hat{\sigma}^2}{\hat{\beta}'_{EKK}\hat{\beta}_{EKK}} \quad (4.34)$$

ile tahmin edilir [22].

4.5. Ridge Tahmin Edicisine İlişkin Hata Kareler Ortalaması ve Varyans Tahminleri

Doğrusal regresyon modeli Eş. 3.1'de verilmiştir. Homojen varyans durumu $\lambda(X)=1$, heterojen varyans durumları ise $\lambda(X)=|X_1|+1$ ve $\lambda(X)=1/(|X_1|+1)$ olarak tanımlanır. β 'nın EKK tahmin edicisi $\hat{\beta}_{EKK}$ olmak üzere,

$$\hat{\beta}_{EKK} = C^{-1}XY \quad (4.35)$$

ile tahmin edilir. Eş. 4.35'de verilen formülde $C = X'X$ 'dir. Çoklu bağlantı olması durumu regresyon katsayıları içerisinde büyük varyanslar ortaya çıkmasına neden olabilir. Hoerl ve Kennard, çoklu bağlantının özellikle mühendislik alanında yaygın bir sorun olduğunu ortaya çıkarmışlardır ve çözümü için negatif olmayan k yan parametresine bağlı olarak ridge tahmin edicisini önermişlerdir [13]. k , yan parametresi olmak üzere ridge tahmin edicisi Eş. 4.36'deki gibi elde edilir.

$$\hat{\beta}_R(k) = (C + kI_p)^{-1}XY \quad (4.36)$$

Ridge tahmin edicisinin EKK tahmin edicisi formunda beklenen değeri $k > 0$ olmak üzere; $E(Y) = X\beta$ olduğundan $E(\hat{\beta}_R(k)) = (C + kI_p)^{-1}C\beta_{EKK}$ 'dır. Burada k yan parametresini, HKO'yu en aza indirecek şekilde belirlemek gerekir.

ε 'nun normal dağıldığı ve $\lambda(X) \equiv I_n$ olduğu varsayıldığında, ridge tahmin edicisinin varyansı Eş. 4.33 ile ifade edilir [13].

$$\text{Var}(\hat{\beta}(k)) = \sigma^2 (C + kI_p)^{-1} X'X (C + kI_p)^{-1} \quad (4.37)$$

σ^2 , $\hat{\sigma}^2 = \sum r_i^2 / (n - p)$ ile tahmin edilir, burada $r_1, \dots, r_n$ artık terimlerdir.

Ridge regresyon analizinde *HKO* 'nun minimize edilmesi amaçlanmaktadır. Ridge tahmin edicisinin hata kareler ortalaması Eş. 4.38 ve Eş. 4.39,

$$HKO(\hat{\beta}(k)) = \text{Var}(\hat{\beta}(k)) + [E(\hat{\beta}(k)) - \beta]^2 \quad (4.38)$$

$$HKO(\hat{\beta}(k)) = \text{Var}(\hat{\beta}(k)) + (\text{sapma miktarı})^2 \quad (4.39)$$

şeklinde ifade edilir [23]. Ridge tahmin edicisinin *HKO*'yu en aza indirmeyi hedefleyen k 'nın tahmin edilmesi amaçlanmaktadır. Kibria, k ridge parametre tahmini için birkaç yöntemi karşılaştırmış ve en iyi performans gösteren tahmin ediciyi belirlemiştir [24]. k ridge parametre tahmini için bu yöntem aşağıda gösterilmiştir.

$C = X'X$ matrisinin özdeğerleri $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_p)$ olmak üzere $D'CD = \Lambda$ olacak şekilde ortogonal D matrisi belirlenir. D ve C matrisine bağlı olarak $\gamma = \Lambda^{-1}(XD)'Y$ ile elde edilir. Burada, γ_j 'nin i . elemanı $\hat{\alpha}_{ji}^2$ olmak üzere k ridge yan parametresi,

$$\hat{k} = \frac{\hat{\sigma}^2}{\left(\prod_{j=1}^p \hat{\alpha}_{ji}^2\right)^{1/p}} \quad (4.40)$$

ile tahmin edilir [24].

Bu tez çalışmasında Eş. 4.40 ile verilen $\hat{k}$ kullanılarak ridge tahmin edicisine ilişkin *HKO*'ları ve varyans tahminleri farklı varyans desenleri altında incelenmiştir.



5. TUTARLI DEĞİŞEN VARYANS KOVARYANS MATRİSİNE DAYALI TAHMİN EDİCİLER

EKK yönteminin en önemli varsayımlarından biri sabit varyans varsayımdır. Heteroskedastisite, hata teriminin varyansının farklı olması durumunda ortaya çıkmaktadır. Değişen varyans varlığında *EKK* tahminleri yansızdır ancak varyans kovaryans tahminleri minimum varyanslı olmadığından yanlış çıkarımlara yol açabilir. *TDVKM*'ye dayalı tahmin ediciler, değişen varyans varlığı durumunda daha doğru sonuçlar elde edilmesini sağlamaktadır. Tutarlı tahmin edici, örnek çapı arttıkça ölçüm hatası sıfıra yaklaşan tahmin edici olarak ifade edilebilir. Literatürde, *EKK* tahmin edicisinin standart hatasını tahmin ederken varyansların heterojenliği durumu için birkaç yöntem karşılaştırılmıştır [3]. *TDVKM*'nin *TDV0* olarak bilinen asimptotik formunun küçük örnek çaplarındaki performanslarına ilişkin *TDV1*, *TDV2* ve *TDV3* olarak tanımlanan üç alternatif tahmin edici geliştirilmiştir. *TDV1*, *TDV2* ve *TDV3* olarak tanımlanan tahmin ediciler, hata terimine bağlı olarak geliştirilen formlardır. Değişen varyans varlığında örnek çapının 25'e kadar küçük olduğu durumlarda bile doğru çıkarımlar veren *TDVKM*'ye dayalı tahmin edici *TDV3* olarak adlandırılmaktadır. Değişen varyans varlığında ve örnek çapının en fazla 250 olması durumunda *TDV3* olarak bilinen *TDVKM*'ye dayalı tahmin edicinin kullanılması önerilmektedir. *TDVKM* dayalı tahmin ediciler, değişen varyans varlığında regresyon katsayılarının kovaryans matrisinin tutarlı bir tahmin edicisini sağlar [3]. Değişen varyans durumunda daha iyi sonuçların elde edilebileceği *TDVKM*'ye dayalı tahmin edicilerin kullanılmasının uygun olduğu belirtilmiştir [3].

Varyans tahmini matrisi S 'nin genel formu aşağıdaki gibidir.

$$S = C^{-1} X' V X C^{-1} \quad (5.1)$$

$j = 0, \dots, p$ olmak üzere $\hat{\beta}_j(0)$ 'nin varyans tahmini, S matrisinin j . köşegen elemanı $s_{jj}^2, (j = 0, \dots, p)$ 'dir.

TDV0, *TDVKM*'ye dayalı tahmin edicilerden en sık kullanılan biçimdir. *TDV0*, heteroskedastisite varlığında $\text{Var}(\hat{\beta}_{EKK})$ için tutarlı bir tahmin edicidir. *TDVKM*'nin *TDV0* olarak tanımlanan asimptotik formu Eş. 5.2 ile gösterilmiştir [3].

$$TDV0 = C^{-1}X'diag[r_i^2]XC^{-1} \quad (5.2)$$

$TDV0$ 'ın küçük örneklem özelliklerini iyileştirmek için hata terimine bağlı olarak üç alternatif tahmin edici geliştirilmiştir. $V_1 = diag(r_i^2)$ olmak üzere; $n/(n-p)$ düzeltme faktörü ile geliştirilen varyans tahmin edicisi $TDV1$ olarak tanımlanmış olup Eş. 5.3'de verilmiştir [5].

$$TDV1 = (n/n-p)C^{-1}X'V_1XC^{-1} \quad (5.3)$$

Hata terimine bağlı olarak değişen üç alternatif tahmin ediciden $TDV2$ tahmin edicisini elde edebilmek için; $i = 1, \dots, n$ olmak üzere $V_2 = diag\left(\frac{r_i^2}{(1-h_{ii})}\right)$ olarak belirlenmektedir.

X matrisinin i . satırı x_i olmak üzere; $h_{ii} = x_i(X'X)^{-1}x_i'$ 'dir. Bu durumda $TDV2$ varyans tahmin edicisi Eş. 5.4 ile tahmin edilir.

$$TDV2 = C^{-1}X'V_2XC^{-1} \quad (5.4)$$

$TDVKM$ 'ye dayalı $TDV3$ tahmin edicisini elde edebilmek için; $i = 1, \dots, n$ olmak üzere $V_3 = diag\left(\frac{r_i^2}{(1-h_{ii})^2}\right)$ olarak belirlenir. Üçüncü alternatif yöntem olan $TDV3$ varyans tahmin edicisi Eş. 5.5 ile elde edilir.

$$TDV3 = C^{-1}X'V_3XC^{-1} \quad (5.5)$$

$k > 0$ olduğunda $TDVKM$ 'ye dayalı $TDV1$, $TDV2$ ve $TDV3$ tahmin edicileri genişletilerek ridge regresyon tahmin edicileri $RTDV1$, $RTDV2$ ve $RTDV3$ sırasıyla aşağıdaki gibi elde edilir [13].

$$RTDV1(k) = (n/n-p)(C + kI_p)^{-1}X'V_1X(C + kI_p)^{-1} \quad (5.6)$$

$$RTDV2(k) = (C + kI_p)^{-1} X' V_2 X (C + kI_p)^{-1} \quad (5.7)$$

$$RTDV3(k) = (C + kI_p)^{-1} X' V_3 X (C + kI_p)^{-1} \quad (5.8)$$

Çoklu bağlantı ve değişen varyans varlığında *EKK* ve ridge regresyona ait *HKO*'ları, *TDVKM*'ye dayalı ridge ve *EKK* varyans tahminleri ve β katsayı tahminleri üzerine simülasyon çalışması yapılmıştır. Simülasyon çalışması sonuçları uygulama kısmında değerlendirilmiştir.





6. UYGULAMA

Doğrusal regresyon modeli $Y = X\beta + \lambda(X)\varepsilon$ biçimindedir. Burada Y , $n \times 1$ boyutlu bağımlı değişken vektörü ve X , $n \times p$ boyutlu bağımsız değişkenlere ait bilinmeyen dizayn matrisidir. $\beta = (\beta_0, \beta_1, \dots, \beta_p)$, $p \times 1$ boyutlu bilinmeyen regresyon katsayıları vektörü ve ε $n \times 1$ boyutlu hata vektörüdür. $\lambda(X)$, X matrisini n boyutlu bir köşegen matrisine iz düşüren bilinmeyen bir fonksiyondur.

Regresyonun amaçlarından biri, bağımlı değişkenle bağımsız değişkenler arasındaki ilişkiyi ortaya çıkarmaktır. Bu amaçla, $\beta = (\beta_0, \beta_1, \dots, \beta_p)$ bilinmeyen regresyon katsayılarının tahmin edilmesi gerekmektedir. Modelin bilinmeyen regresyon katsayıları tahmin edildiğinde, bağımsız değişkenlerin farklı değerleri için bağımlı değişkenin alacağı değer tahmin edilebilir [25].

EKK yönteminin önemli varsayımlarından biri ε 'nin varyansının sabit olmasıdır ($Var(\varepsilon) = \sigma^2 I_n$). Bu varsayım sağlanmadığında değişen varyans durumu ortaya çıkmaktadır. Hata terimi varyansının farklılık göstermesi değişen varyans olarak ifade edilmektedir. Değişen varyans durumunun göz ardı edilmesi, *EKK* standart hatalarının yanlış tahminlerine ve dolayısıyla katsayıların yanlış istatistiksel testlerine yol açmaktadır [26].

EKK yönteminin önemli varsayımlarından diğeri bağımsız değişkenler arasında doğrusal ilişkinin olmamasıdır. Bağımsız değişkenlerin birbirleriyle ilişkili olması durumunda çoklu bağlantı sorunu ortaya çıkmaktadır. Bu durumda güvenilir sonuçlar elde edilememektedir [27]. İki veya daha fazla değişken arasındaki doğrusal ilişkinin derecesinin ve yönünün belirlenmesinde korelasyon matrisi kullanılmaktadır. Korelasyon matrisinde yer alan katsayılar incelenerek değişkenler arasındaki bağımlılık durumu değerlendirilebilmektedir. Değişkenlerden biri büyük değerler alırken diğer değişken de büyük değerler alıyorsa kovaryans pozitif olur. Ters durumda kovaryans negatif olur. Değişkenler arasındaki ilişki derecesi korelasyon katsayısı ile belirlenir. Korelasyon katsayısı +1 ya da -1'e yakın olduğunda ilişki kuvvetlidir. Korelasyon katsayısı sıfıra yakın olduğunda ilişki zayıftır [25].

Değişen varyans ve çoklu bağlantı durumu yanlış çıkarımlara neden olacağından ridge regresyon analizinin kullanılması önerilmektedir. Bu bölümde, değişen varyans ve çoklu bağlantı olması durumunda ridge ve *EKK* tahmin edicilerine ilişkin *HKO*'ları farklı varyans desenleri altında karşılaştırılmış ve sonuçları değerlendirilmiştir. Ayrıca *TDVKM*'ye dayalı *EKK* için *TDV1*, *TDV2* ve *TDV3* varyans tahminleri ile ridge için *RTDV1*, *RTDV2* ve *RTDV3* varyans tahminleri üzerine simülasyon uygulaması Matlab programında çalıştırılmış ve sonuçlar raporlanmıştır.

Simülasyon çalışmasında bağımsız değişkenler arasındaki çoklu bağlantı durumu için korelasyon matrisi oluşturulmuştur. Korelasyon matrisi oluşturulurken, iki bağımsız değişken arasındaki korelasyon katsayısı oldukça anlamlı olacak şekilde belirlenmiştir. Değişken sayısı $p=4$ için varyans-kovaryans matrisi (Σ) bağımsız değişkenler arasında şiddetli çoklu bağlantı olacak şekilde Eş. 6.1'deki gibi alınmıştır.

$$\Sigma = \begin{bmatrix} 1 & 0,9 & 0,9 & 0,9 \\ 0,9 & 1 & 0,999 & 0,9 \\ 0,9 & 0,999 & 1 & 0,9 \\ 0,9 & 0,9 & 0,9 & 1 \end{bmatrix} \quad (6.1)$$

Eş. 3.1'deki doğrusal regresyon model denkleminde yer alan bağımlı ve bağımsız değişkenler ile hata terimi Matlab programı kullanılarak simülasyon çalışmasıyla türetilmiştir. Matlab programında simülasyon çalışması 10 000 tekrara göre incelenmiştir. Varyans deseni $\lambda(X)$ için varyansların homojenliği durumu *VD1*, heterojenliği durumları için *VD2* ve *VD3* olarak tanımlanmış olup, bu durumlar Eş. 6.2, Eş. 6.3 ve Eş. 6.4'da verilmiştir.

$$\lambda(X) = 1 (\text{sabit varyans}) \quad (6.2)$$

$$\lambda(X) = |X_1| + 1 \quad (6.3)$$

$$\lambda(X) = 1/(|X_1| + 1) \quad (6.4)$$

Birinci durum sabit varyans durumudur. VD2 ve VD3 olarak tanımlanan varyansların heterojenliği durumu iki şekildedir. Bağımlı değişkenin, 1. bağımsız değişkene bağlı uç değerlerdeki değişimden kaynaklanan varyans durumu ve tersi olan durumdur [13].

Simülasyonda değişken sayısı $p=4, 6$ olarak belirlenmiştir. X değerleri çok değişkenli normal dağılımdan türetilmiş olup beklenen değer vektörü μ ve varyans değeri Eş. 6.5'deki gibi tanımlanmıştır.

$$\mu' = [0 \ 0 \ 0 \ 0], \sigma^2 = 1 \quad (6.5)$$

$\hat{\beta}_{EKK}$ ve $\hat{\beta}_R(k)$ vektörleri başlangıç değeri $p \times 1$ boyutlu 0 vektörü olarak belirlenmiştir. Örnek çapı $n=20, 30, 50, 100, 200$ için HKO 'ları, EKK ve ridge varyans tahminleri ile regresyon katsayıları incelenmiştir.

$\hat{\beta}_{EKK}$ ve $\hat{\beta}_R$ sırasıyla EKK ve ridge regresyon katsayıları olmak üzere Eş. 4.1 ve Eş. 4.2'ye göre hesaplanmıştır.

Bölüm 4.5.'de bahsedildiği gibi γ_j 'nin i . elemanı $\hat{\alpha}_{ji}^2$ olmak üzere k ridge parametresi,

$$\hat{k} = \frac{\hat{\sigma}^2}{\left(\prod_{j=1}^p \hat{\alpha}_{ji}^2\right)^{1/p}} \quad (6.6)$$

ile tahmin edilmiştir. Simülasyon çalışmasında kullanılan ridge tahmin edicisinin HKO 'su Eş. 4.39'de verilmiştir. Farklı varyans durumları $\lambda(X)=1$, $\lambda(X)=|X_1|+1$ ve $\lambda(X)=1/(|X_1|+1)$ için örnek çapı $n=20, 30, 50, 100, 200$ iken HKO sonuçları incelenmiştir.

Değişen varyans olması durumunda ve örnek çapının en fazla 250 olması durumunda EKK yöntemi için $TDV1$, $TDV2$ ve $TDV3$ olarak tanımlanan varyans tahmin edicileri simülasyon çalışmasında Eş. 5.3, Eş. 5.4, Eş. 5.5'deki gibi alınmıştır.

$k > 0$ iken ridge regresyon yöntemi için $RTDV1$, $RTDV2$ ve $RTDV3$ varyans tahminleri sırasıyla Eş. 5.6, Eş. 5.7, Eş. 5.8'de verilen tahmin ediciler ile hesaplanmıştır. $TDVKM$ 'ye dayalı ridge ve EKK varyans tahminleri sonuçları elde edilmiştir.

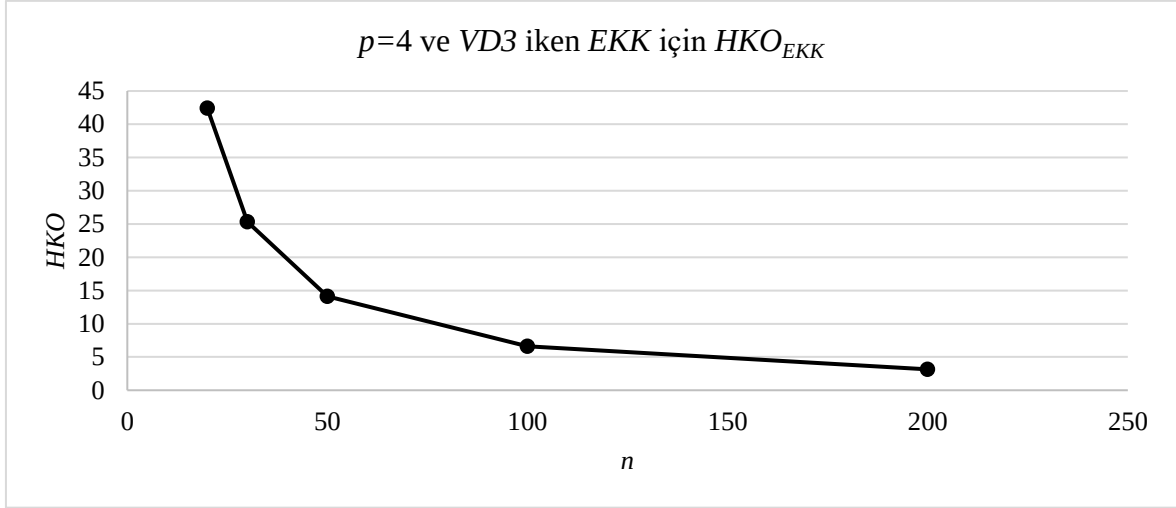
Uygulama çıktıları

Bu çalışmada, bağımsız değişkenler arasında çoklu bağlantı olması durumunda ve değişen varyans varlığında farklı örnek çapları ile elde edilen HKO 'ları ve $TDVKM$ 'ye dayalı varyans tahminleri üzerine çalışılmıştır.

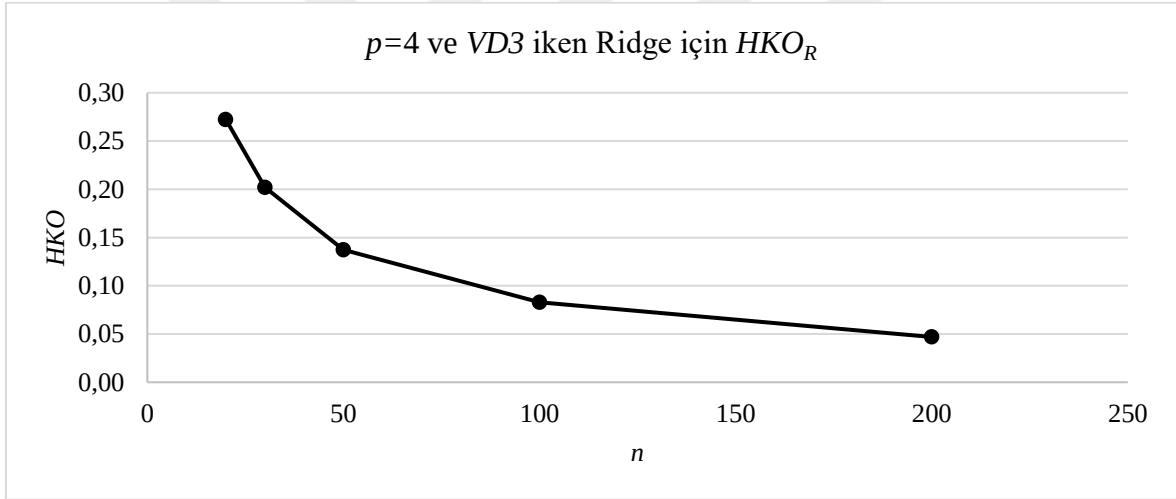
Değişken sayısı $p = 4$ ve 6 olduğunda $TDVKM$ dayalı varyans tahminleri, EKK ve ridge tahmin edicilerine ilişkin HKO sonuçları, regresyon katsayı tahminleri $n = 20, 30, 50, 100, 200$ örnek çapları için ele alınmıştır. EKK yöntemi için HKO 'su HKO_{EKK} ve ridge yöntemi için HKO 'su HKO_R olarak ifade edilmiştir. Simülasyon çalışmasında sabit varyans $\lambda(X) = 1$ ve değişen varyans durumları $\lambda(X) = |X_1| + 1$ ve $\lambda(X) = 1/(|X_1| + 1)$ alınarak elde edilen sonuçlar değerlendirilmiştir. Sabit varyans durumu $VD1$ olmak üzere $\lambda(X) = 1$, değişen varyans durumu $VD2$ için $\lambda(X) = |X_1| + 1$ ve $VD3$ için $\lambda(X) = 1/(|X_1| + 1)$ olarak ele alınmıştır. Simülasyonda tekrar sayısı 10 000 olarak çalışılmıştır. Simülasyon çalışması sonucunda elde edilen verilere çizelgelerde yer verilmiştir.

Çizelge 6.1. Değişken sayısı 4 ve $n = 20, 30, 50, 100, 200$ iken EKK ve ridge tahmin edicilerine ait HKO 'ları

n	$VD1$			$VD2$			$VD3$		
	HKO_{EKK}	k	HKO_R	HKO_{EKK}	k	HKO_R	HKO_{EKK}	k	HKO_R
20	69,017	1,007	0,358	115,58	1,715	0,623	42,405	0,231	0,272
30	41,302	0,440	0,258	70,356	1,757	0,467	25,312	1,175	0,202
50	22,704	1,426	0,180	39,576	2,393	0,308	14,130	3,228	0,137
100	10,311	5,812	0,103	18,963	10,921	0,184	6,619	1,405	0,083
200	5,200	2,156	0,062	9,349	3,280	0,111	3,145	2,830	0,047



Şekil 6.1. Değişken sayısı 4 ve farklı örnek çapları için $VD3$ değişen varyans durumunda EKK için HKO_{EKK}



Şekil 6.2. Değişken sayısı 4 ve farklı örnek çapları için $VD3$ değişen varyans durumunda ridge için HKO_R

Çizelge 6.1’de, değişken sayısı $p = 4$ ve $n = 20, 30, 50, 100, 200$ için $VD1$, $VD2$ ve $VD3$ varyans durumlarında simülasyon çalışmasıyla elde edilen EKK ve ridge tahmin edicilerine ait HKO sonuçları verilmiştir. Sabit varyans ($VD1$) ve değişen varyans ($VD2$, $VD3$) durumlarında örnek çapı arttıkça tahmin edicilere ilişkin HKO değerleri düşmektedir. Ridge tahmin edicisine ilişkin HKO_R sonuçlarının EKK tahmin edicisine ilişkin HKO_{EKK} sonuçlarına göre daha düşük olduğu görülmektedir. Ridge regresyon analizinde yan teriminin karesiyle varyansı küçültürerek HKO ’sını azaltmak amaçlanmaktadır. Elde edilen sonuçlara göre ridge regresyon HKO sonuçları EKK tahmin edicisine göre çok düşüktür. Ridge tahmin edicisi için yan parametresi k ’ya ilişkin elde edilen tahminlere Çizelge 6.1’de

yer verilmiştir. $n = 50$ ve $VD2$ olması durumunda EKK tahmin edicisine ilişkin HKO değeri 39,576, ridge regresyona ait HKO değeri $k = 2,393$ için 0,308 olarak elde edilmiştir [28]. Şekil 6.1 ve Şekil 6.2’de değişken sayısı $p = 4$ ve $n = 20, 30, 50, 100, 200$ için $VD3$ değişen varyans durumunda örnek çapı arttıkça ridge ve EKK ’ya ilişkin HKO ’ların düştüğü görülmektedir.

Çizelge 6.2. Değişken sayısı 4 iken $VD1$, $VD2$ ve $VD3$ durumlarında $TDVKM$ dayalı EKK ve ridge regresyona ilişkin $TDV1$ ve $RTDV1$ varyans tahminleri

	n	$VD1$			$VD2$			$VD3$		
		$TDV1$	k	$RTDV1$	$TDV1$	k	$RTDV1$	$TDV1$	k	$RTDV1$
$\hat{\beta}_0$	20	0,675	1,007	2,821	1,389	1,715	6,438	0,165	0,231	0,463
	30	0,048	0,440	0,597	0,553	1,757	9,573	0,081	1,175	1,233
	50	0,100	1,426	2,493	0,334	2,393	10,711	0,195	3,228	6,244
	100	0,030	5,812	8,530	0,155	10,921	49,120	0,048	1,405	2,250
	200	0,016	2,156	3,231	0,061	3,280	10,323	0,016	2,830	3,602
$\hat{\beta}_1$	20	30,424	1,007	32,570	34,182	1,715	49,402	10,663	0,231	11,388
	30	6,247	0,440	6,797	22,823	1,757	30,947	9,665	1,175	15,382
	50	6,170	1,426	8,563	8,985	2,393	19,889	13,152	3,228	29,873
	100	2,172	5,812	10,671	8,830	10,921	8,500	3,177	1,405	8,297
	200	1,186	2,156	4,400	4,643	3,280	14,875	1,428	2,830	8,993
$\hat{\beta}_2$	20	33,231	1,007	35,377	37,844	1,715	32,440	9,607	0,231	9,573
	30	6,227	0,440	6,777	23,766	1,757	21,550	9,156	1,175	7,172
	50	6,221	1,426	8,614	8,430	2,393	7,476	12,808	3,228	7,976
	100	2,242	5,812	10,741	8,886	10,921	46,979	3,177	1,405	3,250
	200	1,185	2,156	4,400	4,779	3,280	16,038	1,412	2,830	3,046
$\hat{\beta}_3$	20	1,356	1,007	3,502	0,575	1,715	5,791	0,141	0,231	0,537
	30	0,069	0,440	0,618	0,437	1,757	3,671	0,148	1,175	2,404
	50	0,092	1,426	2,486	0,295	2,393	9,214	0,162	3,228	5,596
	100	0,031	5,812	8,530	0,084	10,921	22,440	0,042	1,405	2,062
	200	0,016	2,156	3,231	0,049	3,280	8,373	0,017	2,830	4,359

Değişken sayısı $p=4$ iken $TDVKM$ dayalı EKK için $TDV1$, $TDV2$, $TDV3$ varyans tahminleri ile ridge yöntemi için hesaplanan $RTDV1$, $RTDV2$, $RTDV3$ varyans tahminleri $n = 20, 30, 50, 100, 200$ örnek çapları için incelenmiştir. Çizelge 6.2’de $\lambda(X)$ varyans deseni için sabit varyans durumu $VD1$, değişen varyans durumları $VD2$ ve $VD3$ olarak gösterilmiştir. Ridge yöntemi için $VD1$ sabit varyans durumunda $n=20$ ve $k=1,007$ iken; $j=0$ olmak üzere $\hat{\beta}_0(k)$ ’ın varyans tahmini $RTDV1$ matrisinin köşegen elemanına karşılık gelmek üzere 2,821, $j=1$ olmak üzere $\hat{\beta}_1(k)$ ’ın varyans tahmini $s_{11}^2 = 32,570$, $j=2$ olmak üzere $\hat{\beta}_2(k)$ ’ın varyans tahmini $s_{22}^2 = 35,377$ ve $j=3$ olmak üzere $\hat{\beta}_3(k)$ ’ın varyans tahmini $s_{33}^2 = 3,502$ olarak elde edilmiştir. EKK yöntemi için $VD1$ sabit varyans durumunda $n=20$ iken, $j=0$ olmak üzere $\hat{\beta}_0(0)$ ’ın varyans tahmini $TDV1$ matrisinin köşegen elemanı olmak üzere 0,675, $j=1$ olmak üzere $\hat{\beta}_1(0)$ ’ın varyans tahmini $s_{11}^2 = 30,424$, $j=2$ olmak üzere $\hat{\beta}_2(0)$ ’ın varyans tahmini $s_{22}^2 = 33,231$ ve $j=3$ olmak üzere $\hat{\beta}_3(0)$ ’ın varyans tahmini $s_{33}^2 = 1,356$ olarak elde edilmiştir. Çizelge 6.2’de homojen varyans olması durumunda ridge tahmin edicisine ilişkin $RTDV1$ varyans tahmini sonuçlarının EKK tahmin edicisine ilişkin $TDV1$ sonuçlarından daha büyük değerler olduğu görülmektedir. $VD2$ ve $VD3$ varyans desenleri için ise $RTDV1$ sonuçları genellikle $TDV1$ sonuçlarından daha büyüktür.

Çizelge 6.3. Değişken sayısı 4 iken $VD1$, $VD2$ ve $VD3$ durumlarında $TDVKM$ dayalı EKK ve ridge regresyona ilişkin $TDV2$ ve $RTDV2$ varyans tahminleri

	n	$VD1$			$VD2$			$VD3$		
		$TDV2$	k	$RTDV2$	$TDV2$	k	$RTDV2$	$TDV2$	k	$RTDV2$
$\hat{\beta}_0$	20	0,739	1,007	0,210	1,479	1,715	0,230	0,176	0,231	0,150
	30	0,052	0,440	0,041	0,593	1,757	0,231	0,086	1,175	0,066
	50	0,104	1,426	0,063	0,348	2,393	0,181	0,201	3,228	0,057
	100	0,031	5,812	0,013	0,156	10,921	0,041	0,049	1,405	0,038
	200	0,016	2,156	0,013	0,062	3,280	0,047	0,016	2,830	0,012

Çizelge 6.3. (devam) Değişken sayısı 4 iken $VD1$, $VD2$ ve $VD3$ durumlarında $TDVKM$ dayalı EKK ve ridge regresyona ilişkin $TDV2$ ve $RTDV2$ varyans tahminleri

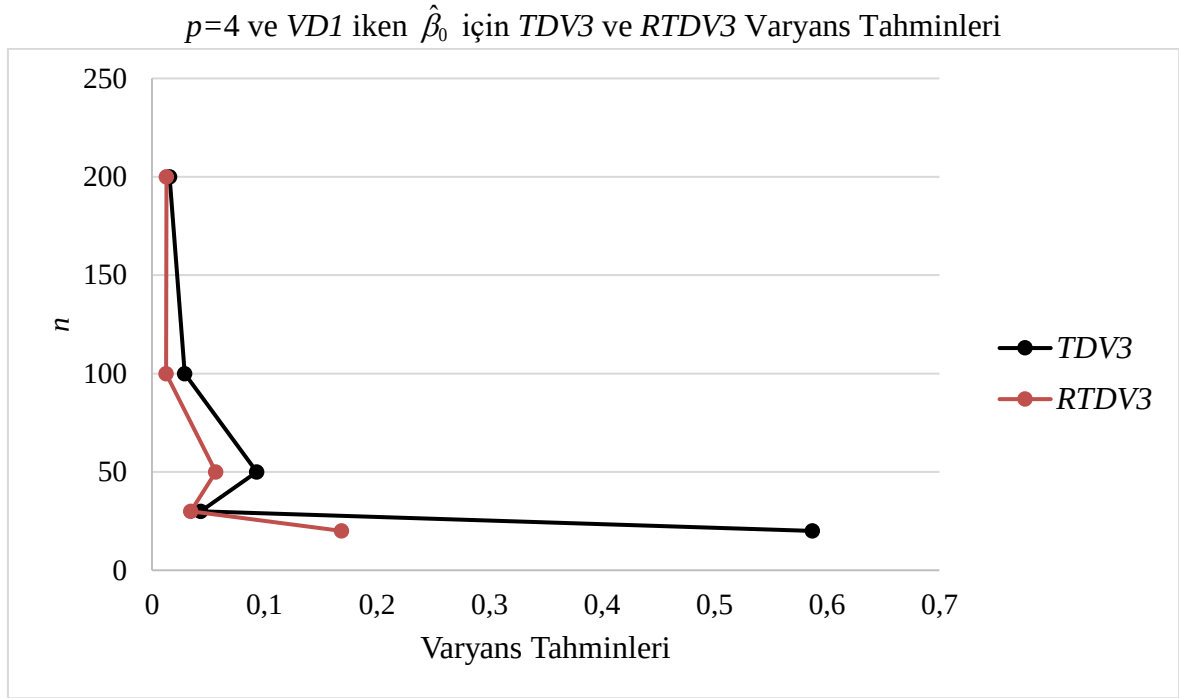
	n	$VD1$			$VD2$			$VD3$		
		$TDV2$	k	$RTDV2$	$TDV2$	k	$RTDV2$	$TDV2$	k	$RTDV2$
$\hat{\beta}_1$	20	35,600	1,007	0,086	38,905	1,715	0,117	11,330	0,231	0,106
	30	6,962	0,440	0,032	24,255	1,757	0,035	10,141	1,175	0,044
	50	6,324	1,426	0,025	9,106	2,393	0,045	13,383	3,228	0,020
	100	2,202	5,812	0,005	8,904	10,921	19,080	3,205	1,405	0,032
	200	1,197	2,156	0,013	4,685	3,280	0,027	1,432	2,830	0,010
$\hat{\beta}_2$	20	39,166	1,007	0,076	42,635	1,715	0,090	10,362	0,231	0,125
	30	6,968	0,440	0,033	25,243	1,757	0,048	9,511	1,175	0,018
	50	6,375	1,426	0,025	8,573	2,393	0,040	12,973	3,228	0,014
	100	2,272	5,812	0,005	8,951	10,921	0,011	3,206	1,405	0,024
	200	1,197	2,156	0,013	4,819	3,280	0,031	1,416	2,830	0,009
$\hat{\beta}_3$	20	1,488	1,007	0,211	0,688	1,715	0,212	0,153	0,231	0,122
	30	0,072	0,440	0,053	0,467	1,757	0,134	0,155	1,175	0,095
	50	0,094	1,426	0,059	0,310	2,393	0,159	0,170	3,228	0,050
	100	0,032	5,812	0,014	0,084	10,921	0,023	0,042	1,405	0,033
	200	0,016	2,156	0,013	0,049	3,280	0,039	0,018	2,830	0,014

Çizelge 6.3’de $VD2$ durumunda $n=100$ iken $\hat{\beta}_1(10,921)=19,080$ dışında ridge tahmin edicisine ilişkin varyans tahmini olan $RTDV2$ sonuçlarının EKK tahmin edicisine ilişkin $TDV2$ sonuçlarından daha düşük olduğu görülmektedir. Ridge yöntemi için $VD1$ sabit varyans durumunda $n=20$ ve $k=1,007$ iken; $j=0$ olmak üzere $\hat{\beta}_0(k)$ ’ın varyans tahmini, $RTDV2$ varyans tahmini matrisinin köşegen elemanına karşılık gelmek üzere 0,210’dır. $j=1$ olmak üzere $\hat{\beta}_1(k)$ ’ın varyans tahmini $s_{11}^2=0,086$, $j=2$ olmak üzere $\hat{\beta}_2(k)$ ’ın varyans tahmini $s_{22}^2=0,076$ ve $j=3$ olmak üzere $\hat{\beta}_3(k)$ ’ın varyans tahmini $s_{33}^2=0,211$ olarak elde edilmiştir. EKK yöntemi için $VD1$ sabit varyans durumunda $n=20$

iken, $j=0$ olmak üzere $\hat{\beta}_0(0)$ 'ın varyans tahmini $TDV2$ matrisinin köşegen elemanı olmak üzere 0,739, $j=1$ olmak üzere $\hat{\beta}_1(0)$ 'ın varyans tahmini $s_{11}^2 = 35,600$, $j=2$ olmak üzere $\hat{\beta}_2(0)$ 'ın varyans tahmini $s_{22}^2 = 39,166$ ve $j=3$ olmak üzere $\hat{\beta}_3(0)$ 'ın varyans tahmini $s_{33}^2 = 1,488$ olarak elde edilmiştir. Sabit varyans ($VD1$) ve değişen varyans ($VD2, VD3$) durumlarında EKK ve ridge tahmin edicileri üzerinden hesaplanan $TDV2$ ve $RTDV2$ varyans tahminleri incelendiğinde, $RTDV2$ varyans tahminlerinin genellikle daha düşük olduğu görülmektedir.

Çizelge 6.4. Değişken sayısı 4 iken $VD1, VD2$ ve $VD3$ durumlarında $TDVKM$ dayalı EKK ve ridge regresyona ilişkin $TDV3$ ve $RTDV3$ varyans tahminleri

	n	$VD1$			$VD2$			$VD3$		
		$TDV3$	k	$RTDV3$	$TDV3$	k	$RTDV3$	$TDV3$	k	$RTDV3$
$\hat{\beta}_0$	20	0,587	1,007	0,169	1,193	1,715	0,182	0,141	0,231	0,119
	30	0,043	0,440	0,034	0,498	1,757	0,196	0,073	1,175	0,054
	50	0,093	1,426	0,057	0,312	2,393	0,162	0,182	3,228	0,051
	100	0,029	5,812	0,013	0,149	10,921	0,039	0,046	1,405	0,036
	200	0,016	2,156	0,013	0,060	3,280	0,045	0,015	2,830	0,012
$\hat{\beta}_1$	20	27,309	1,007	0,071	30,227	1,715	0,090	9,134	0,231	0,087
	30	5,750	0,440	0,027	20,523	1,757	0,030	8,675	1,175	0,036
	50	5,747	1,426	0,023	8,350	2,393	0,041	12,225	3,228	0,017
	100	2,092	5,812	0,005	8,500	10,921	0,009	3,058	1,405	0,030
	200	1,163	2,156	0,012	4,555	3,280	0,026	1,400	2,830	0,010
$\hat{\beta}_2$	20	29,903	1,007	0,061	33,335	1,715	0,070	8,275	0,231	0,097
	30	5,739	0,440	0,028	21,367	1,757	0,042	8,189	1,175	0,015
	50	5,795	1,426	0,023	7,839	2,393	0,036	11,892	3,228	0,012
	100	2,159	5,812	0,004	8,554	10,921	0,010	3,058	1,405	0,023
	200	1,163	2,156	0,012	4,688	3,280	0,030	1,385	2,830	0,009
$\hat{\beta}_3$	20	1,179	1,007	0,169	0,519	1,715	0,161	0,122	0,231	0,096
	30	0,062	0,440	0,045	0,394	1,757	0,115	0,133	1,175	0,081
	50	0,086	1,426	0,054	0,276	2,393	0,141	0,152	3,228	0,045
	100	0,030	5,812	0,013	0,080	10,921	0,022	0,041	1,405	0,032
	200	0,016	2,156	0,013	0,048	3,280	0,037	0,017	2,830	0,013



Şekil 6.3. Değişken sayısı 4 iken $VD1$ sabit varyans durumunda $TDV3$ dayalı EKK ve ridge regresyona ilişkin $TDV3$ ve $RTDV3$ varyans tahminleri

Değişken sayısı $p = 4$ olduğunda, $TDV3$ dayalı EKK için $TDV3$ varyans tahminleri ile ridge yöntemi için hesaplanan $RTDV3$ varyans tahminleri $n = 20, 30, 50, 100, 200$ örnek çapları için Çizelge 6.4’de özetlenmiştir. Sabit varyans ($VD1$) ve değişen varyans ($VD2, VD3$) durumlarında her örnek çapı için ridge regresyon üzerinden hesaplanan $RTDV3$ varyans tahmini değerleri EKK tahmin edicisi üzerinden hesaplanan $TDV3$ sonuçlarına göre daha düşüktür. Örneğin, $VD1$ varyans deseninde $\hat{\beta}_0$ için $RTDV3$ varyans tahminlerinin $TDV3$ varyans tahminlerinden daha düşük olduğu Şekil 6.3’te görülmektedir. Ridge yöntemi için $VD2$ değişen varyans durumunda $n=100$ ve $k=10,921$ iken; $j=0$ olmak üzere $\hat{\beta}_0(k)$ ’ın varyans tahmini, $RTDV3$ varyans tahmini matrisinin köşegen elemanına karşılık gelmek üzere 0,039, $j=1$ olmak üzere $\hat{\beta}_1(k)$ ’ın varyans tahmini $s_{11}^2 = 0,009$, $j=2$ olmak üzere $\hat{\beta}_2(k)$ ’ın varyans tahmini $s_{22}^2 = 0,010$ ve $j=3$ olmak üzere $\hat{\beta}_3(k)$ ’ın varyans tahmini $s_{33}^2 = 0,022$ olarak elde edilmiştir. EKK yöntemi için $VD2$ değişen varyans durumunda $n=20$ iken, $j=0$ olmak üzere $\hat{\beta}_0(0)$ ’ın varyans tahmini $TDV3$ matrisinin köşegen elemanı olmak üzere 0,149, $j=1$ olmak üzere $\hat{\beta}_1(0)$ ’ın varyans tahmini $s_{11}^2 = 8,500$, $j=2$ olmak üzere $\hat{\beta}_2(0)$ ’ın varyans tahmini $s_{22}^2 = 8,554$ ve $j=3$

olmak üzere $\hat{\beta}_3(0)$ 'ın varyans tahmini $s_{33}^2 = 0,080$ olarak elde edilmiştir. Ridge regresyona ilişkin $RTDV3$ varyans tahmini değerlerinin EKK tahmin edicisi üzerinden hesaplanan $TDV3$ varyans tahmini değerlerinden daha düşük olması sebebiyle $RTDV3$ varyans tahminlerinin daha iyi olduğu sonucuna varılmıştır.

Çizelge 6.5. Değişken sayısı 4 ve $n = 20, 30, 50, 100, 200$ iken EKK ve ridge tahmin edicilerine ilişkin $\hat{\beta}$ tahminleri

$\hat{\beta}$	n	VD1			VD2			VD3		
		EKK	k	$Ridge$	EKK	k	$Ridge$	EKK	k	$Ridge$
$\hat{\beta}_0$	20	0,809	1,007	0,750	0,031	1,715	0,452	1,039	0,231	0,937
	30	0,717	0,440	0,730	1,030	1,757	0,926	1,227	1,175	1,287
	50	1,154	1,426	1,144	0,939	2,393	0,873	1,360	3,228	1,151
	100	0,976	5,812	0,994	1,045	10,921	0,979	1,093	1,405	1,095
	200	0,819	2,156	0,838	1,471	3,280	1,408	1,133	2,830	1,121
$\hat{\beta}_1$	20	10,363	1,007	1,173	3,244	1,715	1,048	-2,508	0,231	0,473
	30	2,944	0,440	1,182	6,072	1,757	1,106	-3,071	1,175	0,812
	50	1,748	1,426	0,843	-2,469	2,393	0,611	0,956	3,228	0,889
	100	0,371	5,812	0,988	0,872	10,921	1,078	-2,414	1,405	0,662
	200	-0,023	2,156	0,869	1,439	3,280	0,806	-0,588	2,830	0,873
$\hat{\beta}_2$	20	-8,701	1,007	0,843	-0,981	1,715	1,014	3,655	0,231	0,889
	30	-0,693	0,440	1,019	-3,768	1,757	0,955	4,987	1,175	1,023
	50	-0,214	1,426	0,766	3,770	2,393	0,793	0,681	3,228	0,897
	100	1,662	5,812	1,011	1,428	10,921	1,082	4,285	1,405	1,167
	200	1,929	2,156	1,038	0,066	3,280	0,743	2,547	2,830	1,085
$\hat{\beta}_3$	20	1,412	1,007	1,014	1,469	1,715	1,107	1,985	0,231	1,838
	30	0,969	0,440	0,994	0,575	1,757	0,806	0,819	1,175	0,801
	50	1,428	1,426	1,333	1,240	2,393	1,157	1,179	3,228	1,103
	100	1,056	5,812	1,020	0,769	10,921	0,859	1,003	1,405	1,024
	200	1,197	2,156	1,167	1,011	3,280	1,012	0,933	2,830	0,933

Değişken sayısı 4 iken $n = 20, 30, 50, 100, 200$ örnek çapları için *EKK* ve ridge tahmin edicilerine dayalı olarak elde edilen $\hat{\beta}$ regresyon katsayıları Çizelge 6.5'te özetlenmiştir. Çoklu bağlantı durumunda *EKK* ve ridge yöntemi ile elde edilen $\hat{\beta}$ katsayıları Çizelge 6.5'te karşılaştırılmış olup, *EKK* yöntemi ile elde edilen katsayılarda ters işaretli olma durumu gözlemlenmiştir. *VD2* değişen varyans durumunda $n=50$ için, *EKK* yöntemi ile $\hat{\beta}_1$ tahmini -2,469 hesaplanmış olup $k = 2,393$ için ridge yöntemi ile $\hat{\beta}_1$ tahmini 0,611 olarak elde edilmiştir. Ridge yöntemi ile $\hat{\beta}$ katsayılarının ters işaretli olma durumu gözlenmemiştir.

Çizelge 6.6. Değişken sayısı 6 ve $n = 20, 30, 50, 100, 200$ iken *EKK* ve ridge tahmin edicilerine ait *HKO*'ları

n	<i>VD1</i>			<i>VD2</i>			<i>VD3</i>		
	HKO_{EKK}	k	HKO_R	HKO_{EKK}	k	HKO_R	HKO_{EKK}	k	HKO_R
20	3,235	14,510	0,346	5,708	16,381	0,587	2,048	4,393	0,276
30	1,819	2,167	0,272	3,198	8,820	0,465	1,125	16,299	0,214
50	0,955	10,442	0,187	1,747	11,013	0,341	0,602	15,054	0,151
100	0,461	4,132	0,119	0,802	10,743	0,204	0,281	6,751	0,090
200	0,220	62,327	0,069	0,392	29,776	0,121	0,133	7,104	0,054

Çizelge 6.6'da, *VD1*, *VD2* ve *VD3* varyans desenleri olmak üzere değişken sayısı $p = 6$ iken *EKK* ve ridge tahmin edicilerine ilişkin *HKO* sonuçları $n = 20, 30, 50, 100, 200$ örnek çapları için özetlenmiştir. Simülasyonda tekrar sayısı 10 000 olarak çalışılmıştır. Sabit varyans (*VD1*) ve değişen varyans (*VD2*, *VD3*) durumlarında örnek çapı arttıkça tahmin edicilere ilişkin *HKO* değerleri düşmektedir. Ayrıca, belirlenen örnek çapları için ridge tahmin edicisine ilişkin HKO_R sonuçlarının *EKK* tahmin edicisine ilişkin HKO_{EKK} sonuçlarına göre daha düşük olduğu görülmektedir. $n = 200$ ve *VD2* değişen varyans durumunda, $k = 29,776$ için ridge tahmin edicisine ilişkin *HKO* 0,121 iken *EKK* tahmin edicisine ilişkin *HKO* 0,392 olarak elde edilmiştir.

Çizelge 6.7. Değişken sayısı 6 iken $VD1$, $VD2$ ve $VD3$ durumlarında $TDVKM$ dayalı EKK ve ridge regresyona ilişkin $TDV1$ ve $RTDV1$ varyans tahminleri

	n	$VD1$			$VD2$			$VD3$		
		$TDV1$	k	$RTDV1$	$TDV1$	k	$RTDV1$	$TDV1$	k	$RTDV1$
$\hat{\beta}_0$	20	0,631	14,51	43,707	1,089	16,381	5,379	0,461	4,393	7,995
	30	0,028	2,167	2,332	0,786	8,820	25,652	0,055	16,299	22,411
	50	0,572	10,44	37,682	0,417	11,013	35,118	0,096	15,054	19,153
	100	0,073	4,132	7,132	0,222	10,743	59,322	0,057	6,751	9,413
	200	0,031	62,33	106,51	0,087	29,776	121,4	0,027	7,104	11,518
$\hat{\beta}_1$	20	1,192	14,51	44,269	0,796	16,381	87,778	0,379	4,393	9,356
	30	0,017	2,167	2,322	0,457	8,820	17,703	0,177	16,299	28,893
	50	0,498	10,44	37,607	0,310	11,013	33,024	0,120	15,054	20,591
	100	0,065	4,132	7,123	0,178	10,743	0,168	0,072	6,751	11,524
	200	0,029	62,33	106,51	0,096	29,776	66,672	0,020	7,104	10,746
$\hat{\beta}_2$	20	3,069	14,51	46,145	0,905	16,381	46,016	0,725	4,393	13,931
	30	0,022	2,167	2,326	0,381	8,820	27,340	0,206	16,299	34,925
	50	0,477	10,44	37,587	0,593	11,013	44,725	0,139	15,054	28,035
	100	0,076	4,132	7,135	0,137	10,743	27,065	0,038	6,751	11,075
	200	0,030	62,33	106,51	0,072	29,776	56,042	0,026	7,104	11,986
$\hat{\beta}_3$	20	1,244	14,51	44,321	0,345	16,381	16,909	0,389	4,393	6,088
	30	0,023	2,167	2,328	0,578	8,820	17,476	0,597	16,299	36,974
	50	0,507	10,44	37,616	0,265	11,013	20,407	0,128	15,054	28,264
	100	0,063	4,132	7,122	0,173	10,743	25,836	0,053	6,751	12,130
	200	0,027	62,33	106,51	0,065	29,776	69,623	0,023	7,104	10,001
$\hat{\beta}_4$	20	1,124	14,51	44,201	1,169	16,381	53,359	0,602	4,393	5,634
	30	0,024	2,167	2,329	0,349	8,820	31,679	0,154	16,299	21,001
	50	0,521	10,44	37,631	0,297	11,013	29,548	0,113	15,054	22,491
	100	0,060	4,132	7,118	0,148	10,743	36,002	0,090	6,751	14,090
	200	0,028	62,33	106,51	0,063	29,776	87,618	0,029	7,104	11,353

Çizelge 6.7. (devam) Değişken sayısı 6 iken $VD1$, $VD2$ ve $VD3$ durumlarında $TDVKM$ dayalı EKK ve ridge regresyona ilişkin $TDV1$ ve $RTDV1$ varyans tahminleri

	n	$VD1$			$VD2$			$VD3$		
		$TDV1$	k	$RTDV1$	$TDV1$	k	$RTDV1$	$TDV1$	k	$RTDV1$
$\hat{\beta}_5$	20	1,091	14,51	44,167	0,631	16,381	27,676	0,398	4,393	6,814
	30	0,018	2,167	2,323	0,243	8,820	21,578	0,446	16,299	32,262
	50	0,480	10,44	37,590	0,350	11,013	26,295	0,098	15,054	26,621
	100	0,054	4,132	7,113	0,192	10,743	21,056	0,051	6,751	9,913
	200	0,026	62,33	106,51	0,087	29,776	106,03	0,019	7,104	10,032

Değişken sayısı $p = 6$ iken, $n = 20, 30, 50, 100, 200$ örnek çapları için $TDVKM$ dayalı EKK için $TDV1$ varyans tahminleri ile ridge için $RTDV1$ varyans tahminleri Çizelge 6.7’de özetlenmiştir. $VD3$ değişen varyans durumunda $n=100$ ve $k=6,751$ iken $TDVKM$ dayalı $RTDV1$ varyans tahminleri; $j=0$ olmak üzere $\hat{\beta}_0(k)$ ’in varyans tahmini, $RTDV1$ matrisinin köşegen elemanına karşılık gelmek üzere 9,413, $j=1$ olmak üzere $\hat{\beta}_1(k)$ ’in varyans tahmini 11,524, $j=2$ olmak üzere $\hat{\beta}_2(k)$ ’in varyans tahmini 11,075, $j=3$ olmak üzere $\hat{\beta}_3(k)$ ’in varyans tahmini 12,130, $j=4$ olmak üzere $RTDV1$ varyans tahmini 14,090, $j=5$ olmak üzere $RTDV1$ varyans tahmini 9,913’tür. EKK yöntemi için $VD3$ değişen varyans durumunda $n=100$ iken $TDV1$ varyans tahminleri ise; $j=0$ olmak üzere $\hat{\beta}_0(0)$ ’in varyans tahmini $TDV1$ matrisinin köşegen elemanı olmak üzere 0,057, $j=1$ olmak üzere $\hat{\beta}_1(0)$ ’in varyans tahmini 0,072, $j=2$ olmak üzere $\hat{\beta}_2(0)$ ’in varyans tahmini 0,038, $j=3$ olmak üzere $\hat{\beta}_3(0)$ ’in varyans tahmini 0,053, $j=4$ olmak üzere $TDV1$ varyans tahmini 0,090, $j=5$ olmak üzere $TDV1$ varyans tahmini 0,051’dir. Sabit varyans ($VD1$) ve değişen varyans ($VD2, VD3$) durumlarında EKK varyans tahminleri $TDV1$ değerlerinin ridge regresyona ilişkin $RTDV1$ varyans tahminlerinden genellikle daha düşük olduğu Çizelge 6.7’de görülmektedir.

Çizelge 6.8. Değişken sayısı 6 iken $VD1$, $VD2$ ve $VD3$ durumlarında $TDVKM$ dayalı EKK ve ridge regresyona ilişkin $TDV2$ ve $RTDV2$ varyans tahminleri

	n	$VD1$			$VD2$			$VD3$		
		$TDV2$	k	$RTDV2$	$TDV2$	k	$RTDV2$	$TDV2$	k	$RTDV2$
$\hat{\beta}_0$	20	0,723	14,51	0,019	1,095	16,381	0,006	0,414	4,393	0,034
	30	0,028	2,167	0,006	0,809	8,820	0,037	0,056	16,299	0,002
	50	0,597	10,44	0,043	0,426	11,013	0,035	0,103	15,054	0,006
	100	0,074	4,132	0,032	0,227	10,743	0,058	0,058	6,751	0,023
	200	0,031	62,33	0,002	0,087	29,776	0,015	0,027	7,104	0,014
$\hat{\beta}_1$	20	1,259	14,51	0,016	0,862	16,381	0,009	0,385	4,393	0,041
	30	0,017	2,167	0,006	0,542	8,820	0,018	0,192	16,299	0,012
	50	0,519	10,44	0,047	0,324	11,013	0,034	0,121	15,054	0,005
	100	0,065	4,132	0,031	0,181	10,743	28,373	0,073	6,751	0,023
	200	0,029	62,33	0,002	0,097	29,776	0,011	0,020	7,104	0,012
$\hat{\beta}_2$	20	3,412	14,51	0,009	0,921	16,381	0,009	0,833	4,393	0,035
	30	0,024	2,167	0,007	0,391	8,820	0,031	0,213	16,299	0,003
	50	0,497	10,44	0,046	0,596	11,013	0,022	0,146	15,054	0,006
	100	0,080	4,132	0,034	0,141	10,743	0,033	0,038	6,751	0,016
	200	0,030	62,33	0,002	0,073	29,776	0,012	0,026	7,104	0,013
$\hat{\beta}_3$	20	1,296	14,51	0,016	0,401	16,381	0,005	0,387	4,393	0,032
	30	0,024	2,167	0,006	0,866	8,820	0,034	0,626	16,299	0,013
	50	0,538	10,44	0,046	0,272	11,013	0,031	0,138	15,054	0,011
	100	0,065	4,132	0,031	0,178	10,743	0,035	0,055	6,751	0,022
	200	0,028	62,33	0,002	0,065	29,776	0,012	0,023	7,104	0,012
$\hat{\beta}_4$	20	1,241	14,51	0,013	1,174	16,381	0,008	0,570	4,393	0,029
	30	0,025	2,167	0,006	0,392	8,820	0,045	0,159	16,299	0,003
	50	0,535	10,44	0,043	0,301	11,013	0,030	0,118	15,054	0,007
	100	0,061	4,132	0,029	0,150	10,743	0,030	0,093	6,751	0,027
	200	0,028	62,33	0,002	0,064	29,776	0,011	0,029	7,104	0,014

Çizelge 6.8. (devam) Değişken sayısı 6 iken $VD1$, $VD2$ ve $VD3$ durumlarında $TDVKM$ dayalı EKK ve ridge regresyona ilişkin $TDV2$ ve $RTDV2$ varyans tahminleri

	n	$VD1$			$VD2$			$VD3$		
		$TDV2$	k	$RTDV2$	$TDV2$	k	$RTDV2$	$TDV2$	k	$RTDV2$
$\hat{\beta}_5$	20	1,218	14,51	0,013	0,618	16,381	0,012	0,398	4,393	0,016
	30	0,020	2,167	0,007	0,356	8,820	0,050	0,460	16,299	0,005
	50	0,505	10,44	0,047	0,354	11,013	0,031	0,101	15,054	0,007
	100	0,055	4,132	0,029	0,198	10,743	0,043	0,051	6,751	0,018
	200	0,027	62,33	0,002	0,088	29,776	0,013	0,019	7,104	0,011

Değişken sayısı $p = 6$ olduğunda, $TDVKM$ dayalı EKK ve ridge tahmin edicileri için $TDV2$ ve $RTDV2$ sonuçları Çizelge 6.8’de gösterilmiştir. $VD2$ değişen varyans durumunda $n = 50$ ve $k = 11,013$ iken $TDVKM$ dayalı $RTDV2$ varyans tahminleri; $j = 0$ olmak üzere $\hat{\beta}_0(k)$ ’ın varyans tahmini $RTDV2$ matrisinin köşegen elemanına karşılık gelmek üzere 0,035, $j = 1$ olmak üzere $\hat{\beta}_1(k)$ ’ın varyans tahmini 0,034, $j = 2$ olmak üzere $\hat{\beta}_2(k)$ ’ın varyans tahmini 0,022, $j = 3$ olmak üzere $\hat{\beta}_3(k)$ ’ın varyans tahmini 0,031, $j = 4$ olmak üzere $RTDV2$ varyans tahmini 0,030, $j = 5$ olmak üzere $RTDV2$ varyans tahmini 0,031’dir. EKK yöntemi için $n = 50$ iken $VD2$ değişen varyans durumunda $TDV2$ varyans tahminleri ise; $j = 0$ olmak üzere $\hat{\beta}_0(0)$ ’ın varyans tahmini $TDV2$ matrisinin köşegen elemanı olmak üzere 0,426, $j = 1$ olmak üzere $\hat{\beta}_1(0)$ ’ın varyans tahmini 0,324, $j = 2$ olmak üzere $\hat{\beta}_2(0)$ ’ın varyans tahmini 0,596, $j = 3$ olmak üzere $\hat{\beta}_3(0)$ ’ın varyans tahmini 0,272, $j = 4$ olmak üzere $TDV2$ varyans tahmini 0,301, $j = 5$ olmak üzere $TDV2$ varyans tahmini 0,354’tür. Farklı varyans desenleri altında, ridge regresyona ilişkin $RTDV2$ varyans tahminlerinin EKK yöntemi üzerinden hesaplanan $TDV2$ varyans tahminlerinden genellikle daha düşük olduğu görülmektedir.

Çizelge 6.9. Değişken sayısı 6 iken $VD1$, $VD2$ ve $VD3$ durumlarında $TDVKM$ dayalı EKK ve ridge regresyona ilişkin $TDV3$ ve $RTDV3$ varyans tahminleri

	n	$VD1$			$VD2$			$VD3$		
		$TDV3$	k	$RTDV3$	$TDV3$	k	$RTDV3$	$TDV3$	k	$RTDV3$
$\hat{\beta}_0$	20	0,524	14,51	0,014	0,848	16,381	0,005	0,341	4,393	0,028
	30	0,023	2,167	0,005	0,666	8,820	0,028	0,046	16,299	0,002
	50	0,518	10,44	0,037	0,375	11,013	0,031	0,088	15,054	0,005
	100	0,069	4,132	0,030	0,211	10,743	0,054	0,054	6,751	0,021
	200	0,030	62,33	0,002	0,084	29,776	0,014	0,026	7,104	0,014
$\hat{\beta}_1$	20	0,949	14,51	0,012	0,639	16,381	0,006	0,295	4,393	0,031
	30	0,014	2,167	0,005	0,418	8,820	0,014	0,152	16,299	0,009
	50	0,450	10,44	0,040	0,281	11,013	0,030	0,107	15,054	0,004
	100	0,061	4,132	0,029	0,168	10,743	0,035	0,068	6,751	0,021
	200	0,028	62,33	0,002	0,093	29,776	0,011	0,020	7,104	0,011
$\hat{\beta}_2$	20	2,500	14,51	0,006	0,706	16,381	0,007	0,601	4,393	0,025
	30	0,019	2,167	0,005	0,324	8,820	0,025	0,175	16,299	0,003
	50	0,432	10,44	0,040	0,531	11,013	0,020	0,126	15,054	0,006
	100	0,073	4,132	0,031	0,130	10,743	0,031	0,036	6,751	0,015
	200	0,029	62,33	0,002	0,070	29,776	0,011	0,025	7,104	0,013
$\hat{\beta}_3$	20	0,981	14,51	0,012	0,289	16,381	0,004	0,302	4,393	0,025
	30	0,020	2,167	0,005	0,603	8,820	0,024	0,507	16,299	0,010
	50	0,461	10,44	0,040	0,239	11,013	0,027	0,118	15,054	0,010
	100	0,060	4,132	0,028	0,164	10,743	0,033	0,051	6,751	0,020
	200	0,027	62,33	0,002	0,063	29,776	0,011	0,022	7,104	0,012
$\hat{\beta}_4$	20	0,913	14,51	0,010	0,904	16,381	0,006	0,455	4,393	0,024
	30	0,020	2,167	0,005	0,308	8,820	0,036	0,131	16,299	0,002
	50	0,470	10,44	0,038	0,266	11,013	0,027	0,103	15,054	0,006
	100	0,056	4,132	0,027	0,140	10,743	0,028	0,085	6,751	0,025
	200	0,027	62,33	0,002	0,062	29,776	0,011	0,028	7,104	0,014

Çizelge 6.9. (devam) Değişken sayısı 6 iken $VD1$, $VD2$ ve $VD3$ durumlarında $TDVKM$ dayalı EKK ve ridge regresyona ilişkin $TDV3$ ve $RTDV3$ varyans tahminleri

	n	$VD1$			$VD2$			$VD3$		
		$TDV3$	k	$RTDV3$	$TDV3$	k	$RTDV3$	$TDV3$	k	$RTDV3$
$\hat{\beta}_5$	20	0,893	14,51	0,010	0,484	16,381	0,008	0,308	4,393	0,013
	30	0,016	2,167	0,005	0,249	8,820	0,034	0,377	16,299	0,004
	50	0,435	10,44	0,040	0,314	11,013	0,027	0,089	15,054	0,006
	100	0,051	4,132	0,027	0,182	10,743	0,039	0,048	6,751	0,016
	200	0,026	62,33	0,002	0,085	29,776	0,013	0,018	7,104	0,011

Değişken sayısı $p = 6$ ve $n = 20, 30, 50, 100, 200$ örnek çapları iken, EKK için $TDV3$ varyans tahmin edicileri ile ridge için $RTDV3$ varyans tahmin edicileri Çizelge 6.9'da özetlenmiştir. $VD3$ değişen varyans durumunda $n = 100$ ve $k = 6,751$ iken $TDVKM$ dayalı $RTDV3$ varyans tahminleri; $j = 0$ olmak üzere $\hat{\beta}_0(k)$ 'ın varyans tahmini $RTDV3$ matrisinin köşegen elemanına karşılık gelmek üzere 0,021, $j = 1$ olmak üzere $\hat{\beta}_1(k)$ 'ın varyans tahmini 0,021, $j = 2$ olmak üzere $\hat{\beta}_2(k)$ 'ın varyans tahmini 0,015, $j = 3$ olmak üzere $\hat{\beta}_3(k)$ 'ın varyans tahmini 0,020, $j = 4$ olmak üzere $\hat{\beta}_4(k)$ varyans tahmini 0,025, $j = 5$ olmak üzere $\hat{\beta}_5(k)$ 'ın varyans tahmini 0,016'dır. EKK yöntemi için $n = 100$ iken $TDV3$ varyans tahminleri ise; $j = 0$ olmak üzere $\hat{\beta}_0(0)$ 'ın varyans tahmini $TDV3$ matrisinin köşegen elemanı olmak üzere 0,054, $j = 1$ olmak üzere $\hat{\beta}_1(0)$ 'ın varyans tahmini 0,068, $j = 2$ olmak üzere $\hat{\beta}_2(0)$ 'ın varyans tahmini 0,036, $j = 3$ olmak üzere $\hat{\beta}_3(0)$ 'ın varyans tahmini 0,051, $j = 4$ olmak üzere $TDV3$ varyans tahmini 0,085, $j = 5$ olmak üzere $TDV3$ varyans tahmini 0,048'dir. Farklı varyans desenleri altında, ridge regresyona ilişkin $RTDV3$ varyans tahminlerinin EKK yöntemi üzerinden hesaplanan $TDV3$ varyans tahminlerinden daha düşük olduğu görülmektedir.

Çizelge 6.10. Değişken sayısı 6 ve $n = 20, 30, 50, 100, 200$ iken *EKK* ve ridge tahmin edicilerine ilişkin $\hat{\beta}$ tahminleri

$\hat{\beta}$	n	VD1			VD2			VD3		
		<i>EKK</i>	k	<i>Ridge</i>	<i>EKK</i>	k	<i>Ridge</i>	<i>EKK</i>	k	<i>Ridge</i>
$\hat{\beta}_0$	20	0,775	14,51	0,903	1,368	16,381	0,877	1,633	4,393	1,227
	30	1,397	2,167	1,183	1,451	8,820	1,052	0,950	16,299	0,912
	50	1,542	10,44	1,058	1,577	11,013	1,037	0,974	15,054	1,001
	100	0,757	4,132	0,830	1,175	10,743	1,021	1,219	6,751	1,107
	200	0,963	62,33	0,944	0,915	29,776	0,909	1,059	7,104	1,052
$\hat{\beta}_1$	20	0,151	14,51	0,863	0,314	16,381	0,870	1,074	4,393	1,008
	30	1,069	2,167	1,020	1,018	8,820	0,990	0,877	16,299	0,905
	50	-0,203	10,44	0,543	1,263	11,013	0,966	1,198	15,054	0,948
	100	1,119	4,132	1,075	0,691	10,743	0,830	0,873	6,751	0,940
	200	0,821	62,33	0,868	0,922	29,776	0,929	1,211	7,104	1,140
$\hat{\beta}_2$	20	2,384	14,51	0,890	0,020	16,381	0,796	1,034	4,393	0,841
	30	0,841	2,167	0,951	0,535	8,820	0,800	1,462	16,299	1,004
	50	1,586	10,44	1,136	0,414	11,013	0,960	0,682	15,054	0,840
	100	1,459	4,132	1,269	1,308	10,743	1,192	1,154	6,751	1,113
	200	1,116	62,33	0,964	1,185	29,776	1,036	0,707	7,104	0,777
$\hat{\beta}_3$	20	1,105	14,51	1,072	1,595	16,381	0,988	1,291	4,393	1,045
	30	0,805	2,167	0,889	1,000	8,820	0,939	1,403	16,299	1,049
	50	1,173	10,44	1,016	1,204	11,013	0,945	0,603	15,054	0,838
	100	1,013	4,132	1,022	1,572	10,743	1,217	0,508	6,751	0,657
	200	1,310	62,33	1,025	0,999	29,776	0,915	0,607	7,104	0,691
$\hat{\beta}_4$	20	1,115	14,51	0,966	2,086	16,381	0,982	0,238	4,393	0,745
	30	0,896	2,167	0,928	1,018	8,820	1,017	0,598	16,299	0,919
	50	0,990	10,44	0,917	0,289	11,013	0,698	1,106	15,054	0,978
	100	0,523	4,132	0,683	0,885	10,743	0,988	1,203	6,751	1,073
	200	0,818	62,33	0,895	1,264	29,776	1,092	1,243	7,104	1,157
$\hat{\beta}_5$	20	0,980	14,51	0,891	0,536	16,381	0,822	0,709	4,393	0,949
	30	0,962	2,167	0,915	1,615	8,820	1,236	0,881	16,299	0,880
	50	0,598	10,44	0,825	1,141	11,013	0,916	1,344	15,054	1,006
	100	1,186	4,132	1,131	0,584	10,743	0,807	0,965	6,751	0,953
	200	0,961	62,33	0,915	0,520	29,776	0,772	1,118	7,104	1,082

Çizelge 6.10’da değişken sayısı 6 iken *EKK* ve ridge tahmin edicilerine ilişkin $\hat{\beta}$ regresyon katsayı tahminleri $n = 20, 30, 50, 100, 200$ örnek çapları için özetlenmiştir. *VD3* değişen varyans durumunda $n=50$ iken, *EKK* yöntemi ile $\hat{\beta}_2$ tahmini 0,682 hesaplanmış olup $k = 15,054$ için ridge yöntemi ile $\hat{\beta}_2$ tahmini 0,840’dır.



7. SONUÇ VE ÖNERİLER

Çoklu regresyon modelinde regresyon katsayılarının tahmini için *EKK* yöntemi yaygın olarak kullanılmaktadır. *EKK* tahmin yönteminin amaçlarından biri hatayı minimize etmektir. Bağımsız değişkenler arasında çoklu bağlantı olması durumu hata teriminde büyümelere neden olmaktadır. Hata teriminin varyansının farklı olması durumunda değişen varyans sorunu ortaya çıkmaktadır. Çoklu bağlantı ve değişen varyans varlığı yanlış çıkarımlara sebebiyet verdiğinden yanlış tahmin yöntemi olan ridge regresyon yöntemine başvurulmaktadır. Değişen varyans varlığında örnek çapının çok küçük olduğu durumlarda bile iyi sonuçlar veren tutarlı değişen varyans-kovaryans matrisine dayalı *TDV3* tahmin edicisinin kullanılması önerilmektedir [3,5,6]. Bağımsız değişkenler arasında çoklu bağlantı olması durumunda *HKO* bakımından ridge tahmin edicisi üzerine çalışmalar yapılmıştır. Ridge tahmin edicisiyle elde edilen sonuçların *EKK* tahmin edicisine göre daha iyi performans gösterdiği sonucuna varılmıştır [4,9,11].

Bu çalışmada, sabit ve değişen varyans durumlarında, çoklu bağlantı varlığında *EKK* ve ridge tahmin edicilerinin *HKO* sonuçları ile *TDVKM*'ye dayalı model parametre tahminleri üzerinden varyans tahminleri $n < 200$ iken simülasyonla incelenmiştir. Örnek çapı arttıkça ridge tahmin edicisine ilişkin *HKO* sonuçlarının *EKK* tahmin edicisine ilişkin *HKO* sonuçlarına göre daha düşük olduğu sonucuna varılmıştır. Aynı zamanda çoklu bağlantı varlığında sabit varyans ve değişen varyans durumlarında *TDVKM*'ye dayalı *EKK* ve ridge regresyona ait varyans tahminleri de irdelenmiştir. *TDVKM*'ye dayalı *EKK* yöntemi için *TDV1*, *TDV2*, *TDV3* varyans tahminleri sırasıyla ridge regresyon için elde edilen *RTDV1*, *RTDV2* ve *RTDV3* varyans tahminleri ile karşılaştırılmış ve *RTDV3* sonuçlarının diğerlerine göre daha iyi olduğu gözlenmiştir. Sabit varyans (*VD1*) ve değişen varyans (*VD2*, *VD3*) durumlarında örnek çapı arttıkça *EKK* ve ridge regresyona ilişkin *TDV3* ve *RTDV3* varyans tahmin sonuçları düşmüştür. Ayrıca, ridge tahmin edicisine ilişkin *RTDV3* sonuçlarının *EKK* tahmin edicisine ilişkin *TDV3* sonuçlarından daha düşük olduğu da görülmüştür. Belirlenen örnek çapları için ridge regresyona ait varyans tahmini (*RTDV3*) sonuçlarının daha iyi olduğu tespit edilmiştir. Değişen varyans ve çoklu bağlantı durumunda *EKK* ve ridge yöntemi ile elde edilen $\hat{\beta}$ katsayıları da tahmin edilmiştir. *EKK* yöntemi ile elde edilen katsayılarda ters işaretli olma durumu gözlenmiştir.

Üç farklı varyans deseni ve çoklu bağlantı durumunda $p=4, 6$ ve $n= 20, 30, 50, 100, 200$ iken yanlı tahmin yöntemi olan ridge regresyon tahmin edicisine ait varyans tahmini *RTDV3* ve *HKO* sonuçlarının *EKK* tahmin edicisine göre daha iyi sonuçlar verdiği görülmüştür. Bu durumda *TDVKM*'ye dayalı ridge regresyon üzerinden elde edilen *RTDV3* varyans tahminleri ile daha doğru çıkarımlar elde edilebilir.



KAYNAKLAR

1. Büyükuysal M. Ç. (2010). *Ridge Regresyon Analizi ve Bir Uygulama*, Yüksek Lisans Tezi, Uludağ Üniversitesi Sağlık Bilimleri Enstitüsü, Bursa, 18-19.
2. Eker S. (2011). *Regresyon Analizinde Çoklu Bağlantı Durumunda En Küçük Kareler, Ridge Regresyon ve Temel Bileşenler Yaklaşımlarının Karşılaştırılması*, Yüksek Lisans Tezi, Ondokuz Mayıs Üniversitesi Fen Bilimleri Enstitüsü, Samsun, 12-13.
3. Long, J.S., Ervin, L.H. (2000). Using heteroscedasticity consistent standart errors in the linear regression model. *American Statistical Association*, 54, 217-224.
4. Hoerl, A., Kennard, R. (1970). Ridge Regression: Biased Estimation for Nonorthogonal Problems. *Technometrics*, 12(1), 55-67.
5. MacKinnon, J., G., White, H. (1983). Some Heteroskedasticity Consistent Covariance Matrix Estimators with Improved Finite Sample Properties. *Queen's Economics Department Working Paper*, 1985(4), 1-22.
6. Hayes, A., F., Cai, L. (2007). Using heteroskedasticity-consistent standard error estimators in OLS regression: An introduction and software implementation. *Behavior Research Methods*, 39(4), 709-722.
7. Chen, G., J. (2012). A simple way to deal with multicollinearity. *Journal of Applied Statistics*, 39(9), 1893-1909.
8. Dorugade, A., V. (2015). Correlation Based Ridge Parameters in Ridge Regression with Heteroscedastic Errors and Outliers. *Journal of Statistical Theory and Applications*, 14(4), 413-424.
9. Kibria, B., M., G., Banik, S. (2016). Some Ridge Regression Estimators and Their Performances. *Journal of Modern Applied Statistical Methods*, 15(1), 206-238.
10. Büyükuysal, M., Ç., Öz, İ., İ. (2016). Çoklu Doğrusal Bağlantı Varlığında En Küçük Karelere Alternatif Yaklaşım: Ridge Regresyon. *Düzce Üniversitesi Sağlık Bilimleri Enstitüsü Dergisi*, 6(2), 110-114.
11. Månsson, K., Shukur, G. (2018). Performance of some ridge regression estimators for the multinomial logit model. *Communications in Statistics*, 47(12), 2795-2804.
12. Shrestha, N. (2020). Detecting Multicollinearity in Regression Analysis. *American Journal of Applied Mathematics and Statistics*, 8(2), 39-42.
13. Wilcox, R., R. (2019). Multicollinearity and Ridge Regression: Results on Type I Errors, Power and Heteroscedasticity. *Journal of Applied Statistics*, 46(5), 946-957.
14. Akkaya, Ş., Pazarlıoğlu, M. V. (1998). *Ekonometri I*. İzmir: Anadolu Matbaacılık, 88-127.
15. Topal, M., Eydurhan, E., Yağanoğlu, A. M., Sönmez, A. Y., Keskin, S. (2010). Çoklu Doğrusal Bağlantı Durumunda Ridge ve Temel Bileşenler Regresyon Analiz Yöntemlerinin Kullanımı. *Atatürk Üniversitesi Ziraat Fakültesi Dergisi*, 41(1), 53-57.

16. Hines, W. W. (1990). *Probability and Statistics in Engineering and Management Science*. New York: John Wiley & Sons, Inc, 50-63.
17. Kasalar, S. (2015). *Multicollinearity Çoklu Doğrusal Bağlantı*. İzmir: Yaşar Üniversitesi.
18. Aydın D. (2014). *Uygulamalı Regresyon Analizi, Kavramlar ve R Hesaplamaları*. Ankara: Nobel Akademik Yayıncılık, 71-125.
19. Kurtuluş M. (2001). *Ridge Regresyon Üzerine Bir Çalışma*, Yüksek Lisans Tezi, Gazi Üniversitesi Fen Bilimleri Enstitüsü, Ankara, 10-13.
20. Karakaş S. (2008). *Çoklu Doğrusal Bağlantı Problemi ve Yanlı Regresyon Tahmincileri*, Yüksek Lisans Tezi, İstanbul Üniversitesi Sosyal Bilimler Enstitüsü, İstanbul, 37-38.
21. Arashi M., Saleh E., Kibria B.M. (2019). *Theory of Ridge Regression Estimation with Applications*. USA: John Wiley & Sons, Inc., 1-18.
22. Uslu Y. R. (1991). *Ridge Regresyon ve Öğrenci Başarısı Üzerine Bir Uygulama*, Yüksek Lisans Tezi, Ondokuz Mayıs Üniversitesi Fen Bilimleri Enstitüsü, Samsun, 26-28.
23. Montgomery, D.C., Peck, E.A., Vining, G.G. (2021). *Introduction to Linear Regression Analysis*. (5th Edition). Newyork: Wiley, 69-83.
24. Kibria, B.M.G. (2003). Performance of some new ridge regression estimators. *Communications in Statistics Simulation and Computation*, 32, 419-435.
25. Gamgam, H., Altunkaynak, B., Ünver, Ö. (2011). *SPSS Uygulamalı Temel İstatistik Yöntemler*. Ankara: Seçkin, 239-303.
26. Kaufman, R., L. (2013). *Heteroskedasticity in Regression: Detection and Correction*. London: SAGE Publications, 3-4.
27. Fekedulegn, D., B.. (2002). *Coping with Multicollinearity: An Example on Application of Principal Components Regression in Dendroecology*. Delaware: USDA Forest Service, 1.
28. Güzey, T., Ekiz, M. (2023). *Çoklu Bağlantı Durumunda Tutarlı Değişen Varyans Kovaryans Matrisine Dayalı Tahmin Ediciler*. 8. Uluslararası Akademik Öğrenci Çalışmaları Kongresi, İstanbul, 156-163.



Gazili olmak ayrıcalıktır