



Araştırma Makalesi / Research Article

ÇOKLU DOĞRUSAL BAĞLANTI PROBLEMİ ALTINDA TAHMİN EDİCİLERİN KARŞILAŞTIRILMASI: GENELLEŞTİRİLMİŞ MAKSİMUM ENTROPİ, RIDGE, LIU

Sibel ÖRK ÖZEL¹

Fulya GEZER²

Öz

Genel lineer regresyon modellerinde, çoklu doğrusal bağlantının varlığı durumunda, en küçük kareler (EKK) yöntemi kullanıldığında parametre tahminleri kararsız ve yüksek varyanslı olarak elde edilirler. Bu durumda, çoklu doğrusal bağlantının olumsuz etkilerini azaltmak için EKK yerine alternatif yanlı tahmin edicilerin kullanılması önerilir. Çalışmada, çoklu doğrusal bağlantının tespit edildiği bir genel lineer regresyon modelinde daha kararlı parametre tahminleri elde etmek amacıyla yanlı tahmin edicilerden Ridge, Liu ve Genelleştirilmiş maksimum entropi (GME) tahmin edicileri kullanılmıştır. Tahmin edicilerin performansları hata kareleri ortalamaları (HKO) bazında karşılaştırılmıştır. Tahmin edicilerin HKO değerleri hem bootstrap yöntemi ile hem de Monte Carlo simülasyon çalışması ile elde edilmiştir. Sonuç olarak, çoklu doğrusal bağlantı durumu söz konusu iken en etkin tahmin edicinin GME olduğuna karar verilmiştir.

Anahtar Kelimeler: Çoklu iç ilişki, Genelleştirilmiş Maksimum Entropi, Liu, Ridge

JEL Kodları: C15, C49, C53

COMPARISON OF ESTIMATORS UNDER MULTICOLLINEARITY PROBLEM: GENERALIZED MAXIMUM ENTROPY, RIDGE, LIU

Abstract

In general linear regression models, in the presence of multicollinearity, the parameter estimates are obtained as unstable and with inflated variance when the least squares (OLS) method is used. In such cases, the usage of alternative biased estimators is recommended instead of OLS for reducing the negative effects of multicollinearity. In this study, biased estimators which are Ridge, Liu and Generalized Maximum Entropy (GME) have been used to obtain more stable parameter estimates in a general linear regression model with multicollinearity. Performances of these estimators have been compared according to the mean square error (MSE) criteria. MSE values of these estimators have been obtained by both the bootstrap method using observed data and Monte Carlo simulation study. As a result, in the presence of multicollinearity, it has been decided that GME estimator is the most efficient estimator.

Keywords: Multicollinearity, Ridge, Liu, Generalized Maximum Entropy

JEL Codes: C15, C49, C53

¹ Arş. Gör. Dr., sork@cu.edu.tr, Çukurova Üniversitesi, İ.İ.B.F., Ekonometri Bölümü, ORCID 0000-0002-7030-3512

Sorumlu yazar (Corresponding Author): sork@cu.edu.tr

² Arş. Gör., Ankara Hacı Bayram Veli Üniversitesi, İ.İ.B.F., Ekonometri Bölümü, ORCID 0000-0002-4885-1213

Başvuru Tarihi (Received): 27.02.2020 **Kabul Tarihi** (Accepted): 16.10.2020

Giriş

Regresyon analizinde, en küçük kareler (EKK) yönteminin uygulanması için bazı varsayımların sağlanması gerekmektedir. Ancak uygun modeli kurma, örneklem seçme ve veri oluşturma süreci aşamalarında bazı hatalar söz konusu olabilir. Bunun sonucu olarak analizlerde çoğu zaman bu varsayımlardan sapmalar görülmektedir. Genel lineer regresyon modelinin açıklayıcı değişkenleri arasında ilişki olmaması varsayımından sapma olarak bilinen çoklu doğrusal bağlantı sorunu en sık rastlanan sorunlardan biridir. Çoklu doğrusal bağlantı problemi zaman serisi veya kesit verilerinde, sağlık verilerinde, kimya verilerinde ya da farklı bilim dallarına ait verilerde sıkça karşımıza çıkabilmektedir.

Çoklu doğrusal bağlantı sorunu olan bir modele EKK yöntemi uygulanarak elde edilen parametre tahminlerinde olumsuz birtakım sonuçlarla karşılaşmaktadır. Örneğin, parametre tahminlerinin kararsız olması, işaretlerinin hatalı elde edilmesi, varyansının yüksek çıkması bununla birlikte geniş güven aralıklarının elde edilmesi, reddedilebilecek bir hipotezin kabulü gibi istenmeyen durumlar gözlemlenebilir. Tüm bu nedenlerden dolayı, modelde çoklu doğrusal bağlantı bulunması durumunda EKK yöntemi yerine çoklu doğrusal bağlantının olumsuz etkilerini azaltan, yanlı fakat daha kararlı parametre tahminleri veren alternatif tahmin edicilere başvurulmaktadır.

Çoklu doğrusal bağlantı durumunda, 20. yüzyıldan başlayarak kullanılan ve hala günümüzde sıklıkla kullanılmaya devam eden yanlı tahmin ediciler tercih edilmektedir. Literatür araştırmasında bu tahmin ediciler ile elde edilen parametre tahminlerinin yanlı fakat daha kararlı olarak elde edildiği görülmektedir. Kararlı parametre tahminleri elde etmek araştırmacılar için önem arz etmektedir.

Bir araştırmacı, çoklu doğrusal bağlantı sorunu olan bir genel lineer modelde kararlı tahminler elde etmek için kullanması gereken en etkin tahmin ediciyi, tahmin edicilerin hata kareleri ortalamalarını (HKO) karşılaştırarak belirleyebilir. Çalışmamızda bu tahmin edicilerden Hoerl ve Kennard (1970) tarafından önerilen Ridge tahmin edici, Liu (1993) tarafından önerilen Liu tahmin edici ve Golan, Judge ve Miller (1996) tarafından önerilen, son yıllarda uygulama sıklığı artan Genelleştirilmiş Maksimum Entropi (GME) tahmin edicisi kullanılmıştır.

GME, çoklu doğrusal bağlantı durumunda, Ridge ve Liu tahmin edicilerine alternatif olarak son yıllarda sıklıkla kullanılmaya başlanmıştır. Golan, Judge ve Miller tarafından 1996 yılında önerilen GME, genel lineer modelin bilinmeyen parametrelerini olasılık formunda yeniden parametreleştirerek, parametre ve hataların olasılık dağılımlarını tahmin eder. İktisatta artan sayıda uygulama alanı avantajı sağlayan GME, hem eksik-sunumlu hem de kötü-koşullu evrik problemlerin çözümü için kullanılmaktadır (Akdeniz, Çabuk ve Güler, 2014).

Çalışmamızda çoklu doğrusal bağlantının tespit edildiği Marquart ve Snee (1975) çalışmasında kullanılan veri seti kullanılmıştır. Bu veri setinin çoklu doğrusal bağlantı sorununa değinen birçok uygulamalı çalışmada kullanıldığı görülmüştür ve bu çalışma için de çoklu doğrusal bağlantı sorununa sahip olduğu için tercih edilmiştir. Örneğin; Smith ve Campbell (1980) bu veri seti ile ilgili yapılan analizleri tekrar gözden geçirmiş, Oman (1982) ise bu veri setini Bayes yaklaşımı ile analiz etmiştir. Najarian, Arashi ve Kibria (2013) kısıtlı Ridge tahmin edicisi kullanarak bu veri seti ile bir simülasyon çalışması yapmıştır.

Çalışma, giriş bölümü hariç dört bölümden oluşmaktadır. Birinci bölümde literatür taramasına yer verilmiştir. İkinci bölümde veri seti ve uygulanan yöntemlerden bahsedilmiştir. Üçüncü bölüm, analiz sonucu elde edilen bulguları içermektedir. Çalışmanın son bölümü olan dördüncü bölüm ise analizin sonuç kısmını oluşturmaktadır.

1. Literatür Taraması

Hem ulusal hem de uluslararası çalışmalar incelendiğinde genel lineer regresyon modelleri başta olmak üzere çoklu doğrusal bağlantı probleminin tespit edildiği modellerde yanlı tahmin edicilerin kullanıldığı çalışmalar ile sıklıkla karşılaşılmaktadır. Modellerde kullanılan veri setlerinin çok farklı konularda olduğu göze çarpmaktadır. Çalışmanın bu bölümünde güncel çalışmalardan bazılarına yer verilmektedir. Ancak çalışmamızda kullanılan Ridge, Liu ve GME tahmin edicilerinin birlikte kıyaslandığı bir çalışmaya literatürde rastlanmamıştır. Bu açıdan çalışmamızın literatüre katkı sağlayacağı düşünülmektedir.

Aktaş ve Yılmaz (2003), çoklu regresyon analizlerinde sıklıkla karşılaşılan bir problem olan çoklu doğrusal bağlantı durumunda EKK tahmin edicisi kullanmak yerine daha küçük HKO değerine sahip yanlı tahmin edicilerin kullanılması gerektiğini belirtmiştir. Bundan hareketle yanlı tahmin edicilerden Ridge ve Liu tahmin edicilerini İMKB bileşik endeksi fonksiyonuna uygulamışlardır. Çalışmanın sonucunda Liu tahmin edicinin Ridge tahmin ediciye göre daha duyarlı olduğu bulgusuna ulaşılmıştır.

Mishra (2004) çalışmasında, Portland çimento verisini Maksimum Entropi Leuven (MEL) tahmin edicisi kullanarak analiz etmiş ve tahmin edilen regresyon katsayılarını Kısıtlı Liu tahmin edicisi ile elde edilen değerlerle karşılaştırmıştır.

Karakaş (2008) çalışmasında, çoklu doğrusal bağlantı durumunda kullanılan yanlı tahmin edicilerden Ridge ve Liu tip tahmin edicilerini bir istihdam modeli kullanarak tahmin etmiş ve bu tahmin edicilerin standart hatalarını bulmak için bootstrap yöntemini kullanmıştır. EKK, Ridge ve Liu tip tahmin edicilerini kıyaslayan çalışmada, hem tahmin edicinin işareti bakımından hem de parametrelerin anlamlılıkları açısından Liu tip tahmin edicinin, diğer iki tahmin ediciden daha üstün olduğu ifade edilmiştir.

Pires, Dionisio ve Coelho (2010) çalışmalarında, von Neumann ve Morgenstern fayda fonksiyonlarını EKK ve GME tahmin edicilerini karşılaştırarak tahmin etmişlerdir. Böylece iki tahmin edicinin performansını gerçek veri kümesi kullanarak göstermişler ve Monte Carlo simülasyonu kullanarak elde edilen bulguları onaylamışlardır. İki tahmin edici arasındaki farkın fayda fonksiyonları tahmini için küçük olduğu ve parametre destek vektörünün genişliği arttıkça aralarındaki farkın azaldığı çalışmanın bulgularındandır. Ancak GME tahmin edicisinin EKK tahmin edicisinden daha hassas olduğunu ifade etmişlerdir. Sonuç olarak GME tahmin edicisi, fayda fonksiyonlarının tahmini için EKK tahmin edicisinin bir alternatifi olarak kullanılabilirliği ifade edilmiştir.

Akdeniz, Çabuk ve Güler (2011), Portland çimento veri setini kullandıkları genel lineer regresyon modelinde çoklu doğrusal bağlantı problemini tespit ederek parametre tahmininde EKK, kısıtlı EKK ve GME tahmin edicilerini kullanmışlardır. Çalışmada, tahmin edicilerin HKO değerlerinin hesaplanmasında bootstrap yöntemi kullanılmış olup tahmin ediciler HKO ölçütüne göre karşılaştırılmıştır. Çalışmada kullanılan veri seti için, GME tahmin edicisinin bahsedilen diğer tahmin edicilere göre daha etkin olduğu sonucuna ulaşılmıştır.

Karakaya (2011) çalışmasında, çoklu regresyon analizlerinde karşılaşılan problemlerden biri olan çoklu doğrusal bağlantı probleminin olumsuz etkilerini azaltmak amacı ile yanlı tahmin edicilerden Ridge ve Liu tahmin edicilerini kullanmışlardır. Çalışmada, Türkiye turizm geliri fonksiyonu 1985:2006 dönem aralığında Liu ve Ridge tahmin edicileri ile tahmin edilmiştir. Tahmin edicilerin HKO değerlerine bakılarak yapılan karşılaştırmada, Liu tahmin edicinin Ridge tahmin ediciye göre daha tercih edilebilir olduğu ifade edilmiştir.

Gruber (2012) çalışmasında, Liu tahmin edicinin farklı varyasyonlarını ifade ederek çalışmanın uygulama kısmında Liu, Ridge ve EKK tahmin edicilerini HKO bazında karşılaştırmıştır. Liu tahmin edicinin EKK tahmin edicisinden daha küçük HKO değerine sahip olduğunu

göstermişlerdir. Bazı özel durumlar haricinde de Liu tahmin edicisinin HKO değerinin Ridge tahmin edicisinin HKO değerinden daha küçük olduğu çalışmanın sonuçlarındandır.

Güler, Akdeniz, Çabuk ve Örk Özel (2015), Türkiye'nin birinci düzeydeki 12 istatistikî bölgesi için yoksulluk oranı ve bu oranın belirleyicilerini bir genel lineer regresyon modeli ile analiz etmişlerdir. Modelde çoklu doğrusal bağlantı problemi olduğundan EKK tahmin edicisinin tutarsız sonuçlar verdiği belirlenmiştir. Tutarlı tahminler elde edebilmek için yanlı tahmin edicilerden Ridge ve GME tahmin edicileri kullanılarak model tahmin edilmiştir. GME tahmin edicisi kullanılırken iktisat teorisinin gerektirdiği eşitsizlik ve işaret kısıtlarına da yer verilmiştir. Tahmin edicilerin etkinlik sıralaması, bootstrap yöntemiyle elde edilen HKO değerlerine göre belirlenmiştir. Ele alınan veri seti için sözü edilen tahmin ediciler arasında en etkin tahmin edicinin GME tahmin edicisi olduğu sonucuna ulaşılmıştır.

Adegoke, Adewuyi, Ayinde ve Lukman (2016), çoklu doğrusal regresyon analizinde çoklu doğrusal bağlantı problemi olduğunda EKK tahmin edicinin değil Ridge ve Liu tahmin edicilerinin kullanılması gerektiğini belirtmişlerdir. Ayrıca verilerde aykırı değerler bulunduğu ise M, MM, LTS ve S gibi sağlam tahmin edicilerin tercih edildiğini ifade etmişlerdir. Çalışmada, bu iki problemi birlikte ele almak için Ridge ve Liu tahmin edicileri sırasıyla Robust Ridge ve Robust Liu tahmin edicilerini sağlamak için Robust tahmin edicilerle birleştirilmiştir. Tahmin edicilerin performanslarını karşılaştırmak için HKO ölçütü kullanılmıştır. Çoklu doğrusal bağlantı ve aykırı değer problemlerine sahip üç gerçek veri setine yapılan uygulama sonucu M-Liu ve LTS-Liu tahmin edicilerinin genellikle en etkin olduğu sonucu elde edilmiştir.

Gezer (2016), Klein (1950) çalışmasında yer alan ve çoklu doğrusal bağlantı probleminin tespit edildiği eşanlı denklem modelini kullanmıştır. İki aşamalı en küçük kareler (2AEKK), üç aşamalı en küçük kareler (3AEKK), Ridge ve GME tahmin edicileri ile tahmin edilen modelde tahmin edicilerin performansları bootstrap yöntemi ile elde edilen HKO değerleri kullanılarak karşılaştırılmıştır. Çalışmanın sonucunda, kullanılan model için en etkin tahmin edici GME tahmin edicisi olarak bulunmuştur.

Çabuk ve Örk Özel (2017) çalışmalarında, bir genel lineer regresyon modeli ele alıp Kaliforniya'ya ait yoksulluk veri kümesi (Ramanathan, 2002) kullanarak bir uygulama yapmışlardır. Çoklu doğrusal bağlantının tespit edildiği model için EKK ile tahminin hatalı sonuçlara yol açacağı ifade edilen çalışmada yanlı tahmin edicilerden Ridge ve GME tahmin edicileri kullanılmıştır. Bu tahmin ediciler HKO ölçütüne göre karşılaştırılmıştır. Çalışmada bu veri kümesi için en iyi tahmin edicinin GME tahmin edicisi olduğu sonucuna ulaşılmıştır.

Macedo (2017) çalışmasında, Ridge ve GME tahmin edicilerini birleştiren Ridge-GME parametre tahmin edicisini, Ridge izinin analizindeki öznellikleri ortadan kaldırmak için geliştirilmiştir. Çalışmasında, bir simülasyon uygulaması ve iki uygulama kullanılmıştır.

Şamkar ve Güner (2018), OECD ülkelerindeki beş yaş altı çocuk ölüm sayıları ve ölüm sayısını en çok etkilediğini düşündükleri değişkenlerle bir model oluşturmuşlardır. Oluşturdukları modelde çoklu doğrusal bağlantı probleminin tespit edilmesi üzerine EKK yerine yanlı tahmin edicilerin kullanılması gerektiğini belirterek çalışmada Ridge ve Liu tahmin edicilerini kullanmışlardır. Uygulama kısmında EKK, Ridge ve Liu tahmin edicileri HKO değerlerine göre karşılaştırılmıştır. Yanlı tahmin edicilerin HKO değerleri EKK tahmin edicisinin HKO değerinden küçük olmakla birlikte Ridge tahmin edicisinin HKO değeri en küçük bulunmuştur. Ayrıca çalışmada kullanılan veri seti için tahmin ediciler arasında bir seçim yapmak gerekirse Ridge tahmin edicisinin tercih edilmesi gerektiği ifade edilmiştir.

Derman (2019), tavuk yumurtası kalitesinin incelendiği modelde çoklu doğrusal bağlantı tespit edildiği için modelin tahmin edilmesinde yanlı tahmin edicilerin kullanılması gerektiğini ifade

etmiştir. Çalışmada bu amaçla Ridge, temel bileşenler ve kısmi EKK tahmin edicileri kullanılarak bu tahmin ediciler arasında karşılaştırma yapılmıştır. Çalışmanın sonucunda, yanlış tahmin edicilerin kullanılması durumunda standart hataların normalleştiği ifade edilerek böylelikle daha güvenilir sonuçlar elde edileceği belirtilmiştir.

Küçük (2019) çalışmasında, çoklu bağlantı probleminin yer aldığı bir kimya verisinin bir de gerçek istihdam verisinin kullanıldığı iki ayrı modeli ele alarak çoklu doğrusal bağlantı problemine sahip bu modelleri Ridge, Liu ve LASSO tahmin edicileri ile tahmin etmiştir. Çalışmada, kullanılan iki ayrı veri kümesi için de Liu ve LASSO tahmin edicilerine ait HKO değerleri Ridge tahmin edicisinin HKO değerinden daha küçük olarak elde edilmiştir.

Örk Özel (2019), Transcendental logaritmik (translog) bir modelde çoklu doğrusal bağlantı problemi olacağından bu modellerin tahmin edilmesinde yanlış tahmin edicilerin kullanılması gerektiğini belirtmiştir. Bu amaçla EKK, kısıtlı EKK, Ridge, kısıtlı Ridge (KRidge), GME, kısıtlı GME (KGME) tahmin edicilerini kullanarak bu tahmin edicileri HKO ölçütüne göre karşılaştırmıştır. Uygulamada HKO değerlerinin elde edilmesinde Monte Carlo simülasyonundan faydalanılmıştır. Sonuç olarak GME ve kısıtlı GME tahmin edicilerinin HKO değerleri genel olarak diğer tahmin edicilerin HKO değerlerinden daha küçük olarak elde edilmiştir. Uygun destek matrisinin seçilmesi durumunda Kısıtlı GME tahmin edicisinin daha etkin tahmin edici olduğu çalışmanın sonuçlarındandır.

2. Veri Seti ve Yöntem

Çalışmanın uygulama kısmında Marquart ve Snee (1975) çalışmasında ele alınan ve çoklu doğrusal bağlantının tespit edildiği veri seti kullanılmıştır. Bu veri seti, çoklu doğrusal bağlantının incelendiği birçok uygulamalı çalışmada kullanılmış olup 16 gözlem değeri içermektedir. Çalışmada, söz konusu veri seti kullanılarak EKK, Ridge, Liu ve GME tahmin edicileri ile parametre tahminleri elde edilmiştir. Bu tahmin edicilerin performansları HKO bazında karşılaştırılmıştır. HKO değerinin hesaplanmasında kullanılan gerçek veri seti ile elde edilen sonucun tesadüfi olabileceği ihtimaline karşı çalışma Monte Carlo simülasyon çalışması ile de desteklenmiştir.

2.1. Ridge Tahmin Edicisi

Çoklu doğrusal bağlantının tespit edildiği modellerde çoklu doğrusal bağlantının olumsuz etkilerini azaltmak amacıyla modeldeki değişkenleri çıkarmadan, regresyon katsayılarının yanlış olarak tahmininin gerçekleştirilmesi önerilmektedir. Bu amaçla Hoerl ve Kennard (1970) tarafından geliştirilen Ridge tahmin edicisi,

$$\hat{\beta}_k = (X'X + kI)^{-1} X'Y, \quad (k > 0) \quad (1)$$

şeklinde ifade edilir.

Burada k yanlılık parametresi olup çalışmada iki farklı k yanlılık parametresi kullanılarak Ridge tahmin edicileri elde edilmiştir. İlk olarak (2) nolu eşitlikte verilen Hoerl, Kennard ve Baldwin (1975) tarafından önerilen,

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

ile hesaplanan yanlılık parametresi kullanılmıştır. Burada p parametre sayısı, $\hat{\sigma}^2$ ise genel lineer modelin EKK ile elde edilen regresyon varyansının tahminidir.

Çalışmada kullanılan diğer yanlılık parametresi ise Lawless ve Wang (1976) tarafından önerilen,

$$k_{LW} = \frac{p\hat{\sigma}^2}{\sum_{i=1}^p \lambda_i \hat{\alpha}_i^2} \quad (3)$$

ile hesaplanan yanlılık parametresidir. Burada, $\hat{\alpha}_i$ kanonik model parametrelerinin EKK tahminleri olup, λ_i ise XX matrisine ait özdeğerlerdir.

2.2. Liu Tahmin Edicisi

Çoklu doğrusal bağlantı durumunda uygulanan bir diğer tahmin edici Liu tahmin edicisidir. Liu (1993) tarafından önerilen Liu tahmin edicisi (4) nolu eşitlikte verilmiştir.

$$\hat{\beta}_d = (\mathbf{S} + \mathbf{I})^{-1} (\mathbf{X}'\mathbf{Y} + d\hat{\beta}) \quad , \quad d \in (0,1) \quad (4)$$

Burada, $\hat{\beta} = \mathbf{S}^{-1}\mathbf{X}'\mathbf{Y}$ ve $\mathbf{S} = \mathbf{X}'\mathbf{X}$ olarak ifade edilmektedir.

Çalışmada, iki farklı d yanlılık parametresi kullanılarak Liu tahmin edicileri elde edilmiştir. Liu (1993) tarafından önerilen yanlılık parametresi (5) nolu eşitlikte, Akdeniz ve Kaçıranlar (1995) tarafından önerilen en uygun d yanlılık parametresi ise (6) nolu eşitlikte verilmiştir:

$$\hat{d}_1 = 1 - \hat{\sigma}^2 \left[\frac{\sum_{i=1}^p \frac{1}{\lambda_i(\lambda_i + 1)}}{\sum_{i=1}^p \frac{\hat{\alpha}_i^2}{\lambda_i(\lambda_i + 1)}} \right] \quad (5)$$

$$\hat{d}_2 = \frac{\lambda_i(\hat{\beta}_i^2 - \hat{\sigma}^2)}{\lambda_i\hat{\beta}_i^2 + \hat{\sigma}^2} \quad , \quad i = 1, \dots, p \quad (6)$$

2.3. Genelleştirilmiş Maksimum Entropi (GME) Tahmin Edicisi

Golan, Judge ve Miller tarafından 1996 yılında önerilen GME, Shannon'un (1948) entropinin ölçüsü teoreminden faydalanılarak Jaynes (1957) tarafından sunulan Maksimum Entropi (ME) prensibinin lineer regresyon çerçevesinde genelleştirilmiş halidir. GME yönteminin ME yöntemine göre üstünlüğü sadece eksik-sunumlu evrik problemlerin çözümü için değil aynı zamanda kötü-koşullu evrik problemlerin çözümü için de kullanılıyor olmasıdır. Buna göre,

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u} \quad (7)$$

şeklindeki genel lineer modeli düşünelim.

Burada \mathbf{y} : $T \times 1$ boyutlu açıklanan değişken vektörü, \mathbf{X} : $T \times K$ boyutlu açıklayıcı değişkenler matrisi, $\boldsymbol{\beta}$: $K \times 1$ boyutlu bilinmeyen parametreler vektörü, \mathbf{u} : $T \times 1$ boyutlu hata vektörüdür.

Golan, Judge ve Miller (1996) (7) ile belirtilen modeldeki $\boldsymbol{\beta}$ parametrelerini M ($2 \leq M \leq \infty$) sayıda tıkHz desteklerle birlikte kesikli rassal olasılık değişkenleri kullanarak yeniden parametrelendirmişlerdir. Hatalar için ise $J \geq 2$ olmak üzere hata destekleri tanımlamışlardır. Böylelikle (7) numaralı model

$$\mathbf{y} = \mathbf{XZp} + \mathbf{Vw} \quad (8)$$

ifadesine dönüşmüştür. Burada \mathbf{Z} : $K \times KM$ boyutlu tıkHz destek matrisi, \mathbf{p} : $KM \times 1$ boyutlu ağırlıklar vektörüdür, \mathbf{V} : $T \times TJ$ boyutlu destek noktaları matrisi, \mathbf{w} : $TJ \times 1$ boyutlu bilinmeyen ağırlıklar vektörüdür.

Pukelsheim (1994) tarafından önerilen üç sigma kuralını, Golan, Judge ve Miller (1996) hata bileşenlerinin sınırlarının belirlenmesi aşamasında kullanılması gerektiğini ifade etmişlerdir. Bu çalışmada da hata sınırlarının belirlenmesinde bu kural uygulanmıştır.

(8) numaralı ifade ile gösterilen yeniden parametrelendirilmiş model GME problemi olarak;

$$\max H(\mathbf{p}, \mathbf{w}) = -\sum_{k=1}^K \sum_{m=1}^M p_{km} \ln p_{km} - \sum_{t=1}^T \sum_{j=1}^J w_{tj} \ln w_{tj}$$

(9)

şeklinde ifade edilebilir. Problem için kısıtlar;

$$y_t = \sum_{k=1}^K \sum_{m=1}^M x_{tk} z_{km} p_{km} + \sum_{j=1}^J w_{tj} v_{tj} \quad , \quad t = 1, 2, \dots, T \quad (10)$$

$$\sum_{m=1}^M p_{km} = 1 \quad , \quad k = 1, 2, \dots, K \quad (11)$$

$$\sum_{j=1}^J w_{tj} = 1 \quad , \quad t = 1, 2, \dots, T \quad (12)$$

şeklindedir. Lagrange fonksiyonu çözüldüğünde,

$$\hat{\boldsymbol{\beta}} = \mathbf{Z}\hat{\mathbf{p}} \quad (13)$$

$$\hat{\mathbf{u}} = \mathbf{V}\hat{\mathbf{w}} \quad (14)$$

GME tahmin edicilerine ulaşılır. Burada,

$$\Omega_k^p(\hat{\boldsymbol{\lambda}}) = \sum_{m=1}^M \exp\left(-\sum_{t=1}^T \hat{\lambda}_t z_{km} x_{tk}\right) \quad (15)$$

ve

$$\Psi(\hat{\boldsymbol{\lambda}}) = \sum_{j=1}^J \exp(-\hat{\lambda}_t v_{tj}) \quad (16)$$

olmak üzere

$$\hat{p}_{km} = \frac{\exp\left(-\sum_{t=1}^T \hat{\lambda}_t z_{km} x_{tk}\right)}{\Omega_k^p(\hat{\boldsymbol{\lambda}})} \quad (17)$$

ve

$$\hat{w}_{tj} = \frac{\exp(-\hat{\lambda}_t v_{tj})}{\Psi(\hat{\boldsymbol{\lambda}})} \quad (18)$$

şeklindedir.

3. Bulgular

Veri seti kullanılarak EKK, Ridge, Liu ve GME tahmin edicileri ile parametre tahminleri elde edilmiştir. Tahmin edicilerin etkinlik karşılaştırması, bootstrap yöntemi ile elde edilen HKO değerlerine göre yapılmıştır. Tahmin edicilerin performanslarını karşılaştırmak amacıyla

uygulamada Monte Carlo simülasyon çalışmasına da yer verilmiştir. Analizler için SPSS, Gauss 10 ve MATLAB programları kullanılmıştır.

Çalışmada kullanılan değişkenlerin tanımları Tablo 1’de verilmiştir. Ayrıca çalışmada kullanılan değişkenlere ait normal dağılım testi sonuçları Tablo 2’de, veri kümesine ait özet istatistikler Tablo 3’te ve değişkenler arasındaki korelasyonlar Tablo 4’te yer almaktadır.

Tablo 1: Çalışmada Kullanılan Değişkenler Listesi

Değişken	Tanımı
Y	n-heptanın asetilene dönüşümü (%)
X_1	Reaktör sıcaklığı (°C)
X_2	H2 / n-heptan oranı (mol oranı)
X_3	Temas süresi (saniye)

Çalışmada kullanılan genel lineer regresyon modeli (19) nolu eşitlikte verilmiştir.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + u \quad (19)$$

(19) nolu modelde reaktör sıcaklığı (X_1) ve H2 / n-heptan oranı (X_2) arttıkça n-heptanın asetilene dönüşümünün artması beklenirken, temas süresi (X_3) arttıkça n-heptanın asetilene dönüşümünün azalması beklenmektedir. Buna göre (19) nolu eşitlikte verilen modelin parametrelerinin işaretlerine ait önsel bilgisinin $\beta_1, \beta_2 > 0$ ve $\beta_3 < 0$ şeklinde olduğu söylenebilir.

Tablo 2: Değişkenlere Ait Normal Dağılım Testi Sonuçları

Değişken	Jarque-Bera	Olasılık
Y	1,093261	0,578897
X_1	1,309270	0,519632
X_2	1,184030	0,553212
X_3	2,267017	0,321902

Tablo 2’de verilen çalışmada kullanılan değişkenlere ait Jarque-Bera normal dağılım testi sonuçları incelendiğinde bütün değişkenlerin normal dağılımlı olduğu görülmektedir. Buna göre tamamlayıcı istatistik bilgisi olarak değişkenlere ait ortalama ve standart sapmalar Tablo 3’te verilmiştir.

Tablo 3: Değişkenlere Ait Özet İstatistikler

Değişken	Ortalama	Standart Sapma
Y	36,11	11,899
X_1	1212,50	80,623
X_2	12,444	5,6620
X_3	0,040313	0,0316385

Tablo 4: Örneklem Korelasyon Değerleri

Değişkenler	X_1	X_2	X_3
X_1	1	0,224	-0,958**
X_2	0,224	1	-0,240
X_3	-0,958**	-0,240	1

Not: **: $p < 0,01$

Tablo 4'te verilen değişkenler arasındaki korelasyon değerleri incelendiğinde X_1 ile X_3 değişkenler arasındaki ikili korelasyonların oldukça yüksek olduğu görülmektedir. Bu durumda X_1 ile X_3 açıklayıcı değişkenleri arasında yüksek dereceli korelasyondan dolayı modelin çoklu doğrusal bağlantıdan etkilenebileceğini söylemek mümkündür.

(19) nolu modelin parametrelerinin EKK yöntemi ile elde edilen tahmin sonuçları Tablo 5'te verilmiştir.

Tablo 5: EKK Tahmin Sonuçları

Değişkenler	Tahmin	Standart Hata	t	p değeri	%95 lik Güven aralığı	
					Alt Sınır	Üst Sınır
Sabit	-121,270	55,436	-2,188	0,049	-242,054	-0,486
X_1	0,127	0,042	3,007	0,011	0,035	0,219
X_2	0,348	0,177	1,967	0,073	-0,038	0,734
X_3	-19,022	107,928	-0,176	0,863	-254,177	216,134

Ayrıca modelin belirleyicilik katsayısı $R^2 = 0,918$ olup, modelin F değeri %5 anlamlılık düzeyinde istatistiksel olarak anlamlıdır. Parametrelerin EKK tahmin sonuçları önsel beklentilere uymaktadır. X_1 değişkenine ait parametre değeri %5 anlamlılık düzeyinde istatistiksel olarak anlamlı, fakat X_2 ve X_3 değişkenlerine ait parametre değerleri %5 anlamlılık düzeyinde istatistiksel olarak anlamsız bulunmuştur. Bu durumun çoklu doğrusal bağlantıdan dolayı olduğu düşünülmektedir. Çoklu doğrusal bağlantının derecesini ölçmek için çeşitli yöntemler kullanılmaktadır. Çalışmada, Belsley, Kuh ve Welsch (1980) tarafından önerilen $(\mathbf{X}'\mathbf{X})$

matrisinin özdeğerleri kullanılarak hesaplanan $\kappa = \sqrt{\lambda_{\max} / \lambda_{\min}}$ ile koşul sayılarına bakılmıştır.

Tablo 6'da verilen koşul sayıları incelendiğinde koşul sayısı için 151,123 değeri oldukça yüksek bulunmuştur. Tüm bu yukarıda söz edilen nedenlerden dolayı (19) nolu modelde yüksek derecede çoklu doğrusal bağlantının varlığının söz konusu olduğu ve bu durumun tahmin sonuçlarını olumsuz yönde etkileyebileceği söylenebilir.

Tablo 6: Özdeğerler ve Koşul Sayıları

Boyut	Özdeğer	Koşul sayısı
1	3,539	1,000
2	0,367	3,107
3	0,094	6,130
4	0,0001	151,123

(19) nolu modelde var olan çoklu doğrusal bağlantı sorununun yol açacağı olumsuzlukları en aza indirmek, yanlı fakat daha kararlı tahminciler elde etmek için Hoerl ve Kennard (1970) tarafından önerilen (14) nolu eşitlikte verilen Ridge tahmin edicisi kullanılmıştır.

İki farklı yanlılık parametresi kullanılarak uygulanan Ridge tahmin edici ile elde edilen parametre tahmin sonuçları Tablo 7 ve Tablo 8’de verilmiştir.

Tablo 7: Ridge HKB Tahmin Sonuçları

Değişkenler	Tahmin	Standart Hata	t	p değeri	%95 lik Güven aralığı	
					Alt Sınır	Üst Sınır
Sabit	-103,3117	243,991	-0,42342	0,6795	-634,968	428,3443
X_1	0,1128	0,000159	709,43396	0,0001	0,1093	0,1131
X_2	0,35	0,027	12,96296	0,0001	0,2911	0,4088
X_3	-43,309	462,418	-0,09365	0,9269	-1050,91	964,299

Tablo 8: Ridge LW Tahmin Sonuçları

Değişkenler	Tahmin	Standart Hata	t	p değeri	%95 lik Güven aralığı	
					Alt Sınır	Üst Sınır
Sabit	-106,1015	299,084	-0,3547	0,729	-757,805	545,602
X_1	0,1151	0,000183	628,9617	0,0001	0,1111	0,1154
X_2	0,3485	0,029	12,0172	0,0001	0,2853	0,4116
X_3	-41,527	831,372	-0,0499	0,9610	-1853,086	1770,0325

Çoklu doğrusal bağlantı durumunda uygulanan bir diğer tahmin edici Liu tahmin edicisidir. İki farklı d yanlılık parametresi kullanılarak elde edilen Liu parametre tahmin sonuçları Tablo 9 ve Tablo 10’da verilmiştir.

Tablo 9: Liu (d_1) Tahmin Sonuçları

Değişkenler	Tahmin	Standart Hata	t	p değeri	%95 lik Güven aralığı	
					Alt Sınır	Üst Sınır
Sabit	6,1238	2886,268	0,0212	0,9834	6283,0541	6295,3007
X_1	0,0203	0,0018	11,2777	0,0001	-0,0189	0,0242
X_2	0,4783	0,2334	2,0492	0,0630	-0,0302	0,9868
X_3	-2,0759	7012,997	-0,0029	0,9977	-15283,39	15279,2441

Tablo 10: Liu (d_2) Tahmin Sonuçları

Değişkenler	Tahmin	Standart Hata	t	p değeri	%95 lik Güven aralığı	
					Alt Sınır	Üst Sınır
Sabit	-116,8243	2715,384	-0,043	0,9664	-6033,64	5799,997
X_1	0,1231	0,0017	72,4117	0,0001	0,0861	0,1268
X_2	0,3527	0,2155	1,6366	0,1277	-0,1168	0,8222
X_3	-18,43	6982,228	-0,0263	0,9795	-15232,7	15195,8448

Çoklu doğrusal bağlantının olumsuz etkisini azaltması beklentisiyle uygulanan, Ridge ve Liu tahmin edicileriyle elde edilen tahminlerin işaretleri beklentileri genel olarak karşılasa da standart hataları yüksek olarak elde edilmiştir. Bu nedenle, alternatif bir tahmin yöntemi olan GME tahmin edicisi ile parametreler tahmin edilmiştir, Önsel ortalamalar sıfır olacak şekilde oluşturulan parametre destek vektörleri Tablo 11’de, GME tahmin sonuçları ise Tablo 12’de verilmiştir.

Tablo 11: Parametre Destek Vektörleri

Değişkenler	Parametre Destek Vektörleri	Önsel Ortalama
Sabit	$\mathbf{z}_1 = \{-240, -120, 0, 120, 240\}$	0
X_1	$\mathbf{z}_2 = \{-2, -1, 0, 1, 2\}$	0
X_2	$\mathbf{z}_3 = \{-2, -1, 0, 1, 2\}$	0
X_3	$\mathbf{z}_4 = \{-40, -20, 0, 20, 40\}$	0

Tablo 12: GME Tahmin Sonuçları

Değişkenler	Tahmin	Standart Hata	%95 lik Güven aralığı	
			Alt Sınır	Üst Sınır
Sabit	-111,50159	196,474	-131,71842	-81,04388
X_1	0,11897	0,00013	0,09220	0,13594
X_2	0,28132	0,02189	-0,03268	0,53207
X_3	-1,79551	0,1535	-2,43683	-0,84243

Tablo 12’de verilen GME tahmin sonuçları incelendiğinde, parametrelerin işaretlerinin beklentilere uygun olduğu ve elde edilen güven aralıkları sıfır değerini içermediğinden tüm parametre tahminlerinin %95 güven düzeyinde istatistiksel olarak anlamlı olduğu görülmektedir, Ayrıca, GME ile elde edilen parametre tahminlerinin standart hataları, Ridge ve Liu ile elde edilen standart hatalara göre daha küçük olarak elde edilmiştir,

Tahmin edicileri karşılaştırmada sadece parametre tahminlerin işaretlerinin tutarlılığına ve standart hataların büyüklüğüne bakmak yetersizdir. Tahmin edicilerin karşılaştırılmasında HKO değerlerine göre karşılaştırma yapmak gerekmektedir.

GME tahmin edicisi için HKO değeri hesaplanırken kapalı formda ifade edilemediği için HKO değeri Efron (1979) tarafından önerilen bootstrap yaklaşımı ile tahmin edilmiştir, Bootstrap yaklaşımının adımları şu şekildedir:

1. Adım: $y = X\beta + e$ lineer regresyon modelinde, $\tilde{\beta}$ parametresi herhangi bir tahmini olmak üzere, $e = y - X\tilde{\beta}$ eşitliği ile kalıntılar hesaplanır.
2. Adım: j -inci bootstrap denemesinde hesaplanan e ’nin elemanlarından n tanesi rastgele çekilir.
3. Adım: Rastgele çekilen e ’lerin oluşturduğu vektör e_j^* olmak üzere, X ve $\tilde{\beta}$ kullanılarak j -inci bootstrap denemesi için y_j^* hesaplanır. Bağımlı değişken y için yeni bir örneklem elde edilmiş olur.

4. Adım: Elde edilen y_j^* ve X kullanılarak j -inci deneme için $\tilde{\beta}_j$ hesaplanır.

Son adım yeterince tekrar edilerek $\tilde{\beta}$ için bir örnekleme dağılımının tahmini elde edilmiş olur (Çabuk ve Akdeniz, 2007 & Akdeniz, Çabuk ve Güler, 2011).

EKK, Ridge, Liu ve GME tahmin edicilerinin HKO değerlerini elde etmek için bu işlem 400 defa tekrarlanarak her bir tahmin edici için elde edilen HKO değerleri Tablo 13'te verilmiştir.

Tablo 13: Bootstrap Sonuçları

Tahmin Yöntemi	k Değeri	d Değeri	HKO Tahminleri
EKK	-	-	2676,5249
Ridge HKB	0,00376	-	1691,8112
Ridge LW	0,00248	-	1276,9812
Liu (d_1)	-	-0,12399	18763,156
Liu (d_2)	-	0,96077	13614,146
GME	-	-	614,3667

Tablo 13'te verilen HKO değerlerine bakılıp yanlış tahmin edicilerin HKO değerleri incelendiğinde; $HKO_{GME} < HKO_{RidgeLW} < HKO_{RidgeHKB} < HKO_{Liu(d_1)} < HKO_{Liu(d_2)}$ sonucuna ulaşılmıştır.

Daha sonra modelin parametrelerinin tahmininde kullanılan tahmin edicilerin performansını daha doğru bir şekilde kıyaslayabilmek adına Monte Carlo simülasyon çalışması yapılmıştır. Uygulanan Monte Carlo simülasyon adımları şu şekildedir:

1. Adım: Açıklayıcı değişkenler $X_1 \sim N(1,2)$, $X_2 \sim N(1,3)$, $X_3 \sim N(1,4)$ olarak $\mu=[1,1,1]$ ortalamalı ve sırasıyla 2, 3, 4 varyanslı olarak çok değişkenli normal dağılımdan üretilmiştir.

2. Adım: Çoklu doğrusal bağlantının etkisini görebilmek amacıyla tüm açıklayıcı değişkenler arasındaki korelasyon değeri ρ için sırası ile 0,5 ve 0,9 değerleri seçilmiştir.

3. Adım: Hata terimleri ise $\varepsilon \sim N(0; 0,25)$ olarak $\mu=0$ ortalamalı ve $\sigma^2 =0,25$ varyanslı normal dağılımdan üretilmiştir. 4. Adım: Çalışmada oluşturulan X matrisi tüm denemeler sırasında sabit

tutularak $\beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix}$ EKK ile elde edilen tahmin değerleri başlangıç değerleri olarak kullanılıp $Y^* =$

$\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \varepsilon$ modelinden Y^* üretilmiştir.

5. Adım: Son adım yeterince tekrar edilerek $\tilde{\beta}$ için bir örnekleme dağılımı elde edilmiş olur.

Simülasyon çalışmasında yapılan 1000 deneme sonucu elde edilen tahmin edicilerin HKO değerleri Tablo 14'te verilmiştir.

Tablo 14: Monte Carlo Simülasyon Çalışmasına Ait HKO Tahminleri

Tahmin Yöntemi	$\rho=0,5$ için HKO tahminleri	$\rho=0,9$ için HKO tahminleri
Ridge HKB	3,49024	1,06867
Ridge LW	1,08105	2,06535
Liu (d_1)	1,02651	1,74279
Liu (d_2)	5,92034	5,09523
GME	0,1539	0,6906

Tablo 14’te verilen HKO değerlerine bakıldığında göze çarpan ilk husus korelasyon değeri 0,9 seçildiğinde HKO değerlerinin genel olarak artıyor olmasıdır. Bu durum çoklu doğrusal bağlantının etkisinden kaynaklanmaktadır. Bununla birlikte, her iki korelasyon değeri için de GME tahmin edicisinin HKO değeri, diğer yanlı tahmin edicilerin HKO değerlerine göre daha küçük elde edilmiştir.

Tahmin edicileri etkinlik açısından değerlendirirken en düşük HKO değerine sahip tahmin ediciyi en etkin tahmin edici olarak belirleyebiliriz. Bu çalışmada elde edilen analiz sonuçlarına bakıldığında, HKO hesaplanmasında kullanılan her iki yöntem için de en etkin tahmin edicinin GME tahmin edicisi olduğu görülmektedir.

4. Sonuç

Bu çalışmada çoklu doğrusal bağlantının tespit edildiği bir genel lineer regresyon modelinin parametreleri yanlı tahmin ediciler olan Ridge, Liu ve GME ile tahmin edilerek GME tahmin edicisinin diğer tahmin edicilerden performans bazında üstünlüğünü göstermek amaçlanmıştır. Bu amaç doğrultusunda hem bootstrap hem de bir Monte Carlo simülasyon çalışması ile tahmin edicilerin HKO değerleri hesaplanmıştır. Hem gerçek veri seti için bootstrap yaklaşımı ile elde edilen HKO değerleri incelendiğinde hem de Monte Carlo simülasyon çalışması ile elde edilen HKO değerleri incelendiğinde GME tahmin edicinin Ridge ve Liu tahmin edicilerinden daha düşük HKO değerine sahip olduğu görülmektedir. Bu doğrultuda genel lineer bir modelde çoklu doğrusal bağlantı söz konusu iken en etkin tahmin edicinin GME tahmin edicisi olduğu sonucuna ulaşılmıştır.

Elde edilen sonuçların genel olarak literatür ile tutarlı olduğu görülmektedir. Akdeniz, Çabuk ve Güler (2011); Güler, Akdeniz, Çabuk ve Örk Özel (2015); Gezer (2016); Çabuk ve Örk Özel (2017) ve Örk Özel (2019) GME ve Ridge tahmin edicileri arasında kıyaslama yaparak GME tahmin edicisinin Ridge tahmin ediciye göre daha etkin olduğunu, Şamkar ve Güner (2018) ile Küçük (2019) ise Ridge tahmin edicinin Liu tahmin edicisinden daha etkin olduğunu ifade etmişlerdir.

Literatür taramasında belirtildiği üzere, çoklu doğrusal bağlantı durumunda en etkin tahmin edicinin belirlenmesi amaçlı yapılan çalışmalarda karşılaştırmalarda Ridge, Liu ve GME yanlı tahmin edicilerinin üçüne birden yer verilmemiştir. Bu çalışmanın bu konuda literatüre katkı sağlayacağı düşünülmektedir.

Ulaşılan sonuçlar çerçevesinde, araştırmacılar tarafından ele alınan modelde çoklu doğrusal bağlantı probleminin tespit edilmesi durumunda daha kararlı parametre tahminleri elde etmek amacıyla yanlı tahmin edicilerden GME tahmin edicisine başvurulması önerilebilir.

Kaynakça

- Adegoke, A. S., Adewuyi, E., Ayinde, K. & Lukman, A. F. (2016). A comparative study of some robust Ridge and Liu estimators. *Science World Journal*, 11(4), 16-20.
- Akdeniz, F., Çabuk, A., & Güler, H. (2011). Generalized maximum entropy estimators: Applications to the Portland cement dataset. *The Open Statistics and Probability Journal*, 3, 13-20.
- Akdeniz, F., Çabuk, A., & Güler, H. (2014). Doğrusal eşanlı denklem modellerinin tahmininde kullanılan tahmincilerin kötü koşulluluk durumunda HKO bazında performanslarının özyetim ile karşılaştırılması. K. O. Oruç & H. Demirgil (Ed.), *15th International Symposium on Econometrics, Operations Research and Statistics Abstract Book* içinde (ss.76). Ankara: Nobel Akademik Yayıncılık

- Akdeniz, F., & Kaçıranlar, S. (1995). On the almost unbiased generalized Liu estimator and unbiased estimation of the bias and MSE. *Communications in Statistics-Theory and Methods*, 24(7), 1789-1797.
- Aktaş, C., & Yılmaz, V. (2003). Çoklu bağıntılı modellerde Liu ve Ridge regresyon kestiricilerinin karşılaştırılması. *Anadolu Üniversitesi Bilim ve Teknoloji Dergisi*, 4(2), 189-194.
- Belsley, D. A., Kuh, E., & Welsch, R. E. (1980). *Regression diagnostics*. New York: Wiley.
- Çabuk, A., & Akdeniz, F. (2007). İlişki ve genelleştirilmiş maksimum entropi tahmin edicileri. *Journal of Statistical Research*, 5(2), 1-19.
- Çabuk, H. A., & Örk Özel, S. (2017). Çoklu iç ilişki sorunu olan regresyon modelinin HKO ölçütü ile bir etkin tahmin edicisi. *Ç.Ü. Sosyal Bilimler Enstitüsü Dergisi*, 26(3), 13-25.
- Derman, D. A. (2019). *Çoklu bağlantı durumunda yanlı regresyon yöntemlerinin incelenmesi*. Yayımlanmamış yüksek lisans tezi, Ordu Üniversitesi, Ordu.
- Efron, B. (1979). Bootstrap methods: Another look at the Jackknife. *The Annals of Statistics*, 7(1), 1-26.
- Gezer, F. (2016). *Eşanlı denklem modellerinde çoklu iç ilişkinin etkileri ve alternatif tahmin edicilerin karşılaştırılması*. Yayımlanmamış yüksek lisans tezi, Çukurova Üniversitesi, Adana.
- Golan, A., Judge, G., & Miller, D. (1996). *Maximum entropy econometrics: Robust estimation with limited data*. New York, USA: John Wiley & Sons
- Gruber, M. H. J. (2012). Liu and ridge estimators-a comparison. *Communications in Statistics - Theory and Methods*, 41(20), 3739-3749.
- Güler, H., Akdeniz, F., Çabuk H. A., & Örk Özel, S. (2015). Poverty rate and its determinants for 12 statistical regions of Turkey: Generalized maximum entropy approach. *Ç.Ü. Sosyal Bilimler Enstitüsü Dergisi*, 24(2), 337-348.
- Hoerl, A. E., & Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, 12(1), 55-67.
- Hoerl, A. E., Kennard, R. W., & Baldwin, K. F. (1975). Ridge regression: Some simulation. *Communication in Statistics*, 4, 105-123.
- Jaynes, E. T. (1957). Information theory and statistical mechanics. II. *Physics Review*, 108(2), 171-190.
- Karakaş, S. (2008). *Çoklu doğrusal bağlantı problemi ve yanlı regresyon tahminçileri*. Yayımlanmamış yüksek lisans tezi, İstanbul Üniversitesi, İstanbul.
- Karakaya, E. (2011). *Ridge ve Liu tahminçilerinin etkinliklerinin ve yanlılıklarının karşılaştırılması*. Yayımlanmamış yüksek lisans tezi, Anadolu Üniversitesi, Eskişehir.
- Klein, L. R. (1950). *Economic fluctuations in the United States, 1921-1941*. John Wiley and Son, New York.
- Küçük, A. (2019). *Doğrusal regresyonda Ridge, Liu ve LASSO tahmin edicileri üzerine bir çalışma*. Yayımlanmamış yüksek lisans tezi, Hacettepe Üniversitesi, Ankara.
- Lawless, J. F., & Wang, P. (1976). A simulation study of Ridge and other regression estimators. *Communications in Statistics-Theory and Methods*, 14, 1589-1604.

- Liu, K. (1993). A new class of biased estimate in linear regression. *Communications in Statistics-Theory and Methods*, 22(2), 393-402.
- Macedo, P. (2017). Ridge regression and generalized maximum entropy: An improved version of the Ridge-GME parameter estimator. *Communications in Statistics-Simulation and Computation*, 46(5), 1532-4141.
- Marquardt, D. W., & Snee, R. D. (1975). Ridge regression in practice. *The American Statistician*, 29(1), 3-20.
- Mishra, S. K. (2004). *Estimation under multicollinearity: application of restricted Liu and maximum entropy estimators to the Portland cement dataset*. MPRA Paper 1809, University Library of Munich, Germany.
- Najarian, S., Arashi, M., & Kibria, B. M. G. (2013). A simulation study on some restricted ridge regression estimators. *Communications in Statistics - Simulation and Computation*, 42(4), 871-890.
- Oman, S. D. (1982). Shrinking towards subspaces in multiple linear regression. *Technometrics*, 24(4), 307-311.
- Örk Özel, S. (2019). *Transcendental logaritmik (translog) modelin etkin tahmini: Tahmin edicilerin monte carlo ile karşılaştırılması*. Yayımlanmamış doktora tezi, Çukurova Üniversitesi, Adana.
- Pires, C., Coelho, L. A., & Dionisio, A. (2010). *GME versus OLS – Which is the best to estimate utility functions?* (CEFAGE-UE Working Paper 2010/02). Retrieved from https://www.researchgate.net/profile/Andreia_Dionisio/publication/46448323_GME_vs_OLS_Which_is_the_best_to_estimate_utility_functions/links/0912f50f3e8b29627a000000/GME-versus-OLS-Which-is-the-best-to-estimate-utility-functions.pdf
- Pukelsheim, F. (1994). The three sigma rule. *The American Statistician*, 48(2), 88-91.
- Ramanathan, R. (2002). *Introductory econometrics with applications* (5th Edition). Harcourt College Publishers.
- Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, 27(3), 379-423.
- Smith, G., & Campbell, F. (1980). A critique of some ridge regression methods. *Journal of the American Statistical Association*, 75(369), 74-81.
- Şamkar, H., & Güner, D. (2018). OECD ülkelerindeki beş yaş altı çocuk ölüm sayılarının yanlış tahmin teknikleriyle modellenmesi. *UIİİD-IJEAS*, 21, 273-284.