




İstatistikçiler Dergisi: İstatistik & Aktüerya
Journal of Statisticians: Statistics and Actuarial Sciences
IDIA 15, 2022, 1, 1-18
Geliş/Received:06.06.2022, Kabul/Accepted: 25.06.2022

Araştırma Makalesi / Research Article

Ridge regresyon parametre seçimi: Türkiye'nin doğrudan yabancı yatırım örneği


Bahadır YÜZBAŞI

*İnönü Üniversitesi
İktisadi ve İdari Bilimler Fakültesi
Ekonometri Bölümü
Malatya, Türkiye
b.yzb@hotmail.com*

 0000-0002-6196-3201

Mustafa PALA

*İnönü Üniversitesi
İktisadi ve İdari Bilimler Fakültesi
Ekonometri Bölümü
Malatya, Türkiye
mustafapalaa@gmail.com*

 0000 0002 5390 1190

Öz

Bu çalışmada çoklu doğrusal regresyon modelin bağımsız değişkenler arasında çoklu doğrusal bağlantı olduğu durumlarda En Küçük Kareler (EKK) yöntemine alternatif olarak kullanılan Ridge regresyon metodu için ayar parametresi seçimine yardımcı olacak bazı kriterler, Akaike Bilgi Kriteri (AIC), Bayes Bilgi Kriteri (BIC), Mallow's Cp, Çapraz Geçerlilik (CV) ve Genelleştirilmiş Çapraz Geçerlilik Ölçütü (GCV) karşılaştırılmıştır. Kullanılan model seçim kriterlerinin performansları Monte Carlo simülasyon çalışması ve ekonometrik bir veri kullanılarak hata kareler ortalaması (HKO) ve tahmin hatası (TH) kriterleri yardımıyla karşılaştırılmıştır. Nümerik çalışmalar sonucunda, çoklu doğrusal bağlantının olduğu durumlarda önerilen kriterler ile ayar parametresi seçilen Ridge regresyon yöntemlerinin daha düşük HKO ve TH değerleri ile daha üstün performans gösterdiği bulunmuştur.

Anahtar sözcükler: Çoklu Doğrusal Bağlantı, Ridge Regresyon, LASSO.

Abstract

Ridge regression parameter selection: Turkey's example of foreign direct investment

In this study, some criteria such as Akaike Information Criteria (AIC), Bayes Information Criteria (BIC), Mallow's Cp, Cross Validity (CV) and Generalized Cross Validity Measure (GCV) that will help the selection parameter for the Ridge regression method, which is used as an alternative to the Least Squares (Least Squares) method in cases where the multiple linear regression model has multiple linear connections between the independent variables, are compared. The performances of the model selection criteria used were compared using the Monte Carlo simulation study and econometric data, with the help of mean squares error (MSE) and prediction error (PE) criteria. As a result of the numerical studies, it was found that the Ridge regression methods, whose adjustment parameter was selected with the suggested criteria in cases of multicollinearity, showed superior performance with lower MSE and PE values.

Keywords: Multicollinearity, Ridge Regression, LASSO.

1. Giriş

Aşağıda verilen çoklu doğrusal regresyon modelini göz önüne alalım;

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_p X_{ip} + \varepsilon_i, \quad i = 1, 2, \dots, n, \quad (1)$$

Burada Y_i 'ler bağımlı rassal değişkenler, $X_{i1}, X_{i2}, \dots, X_{ip}$ 'ler bağımsız açıklayıcı değişkenler, β_0 sabit regresyon katsayısı, $\beta_1, \beta_2, \dots, \beta_p$ kısmi regresyon katsayıları ve $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ aynı dağılımlı özdeş rassal hata terimidir. i alt indisleri belirlenmiş olan mümkün gözlemi, n gözlem sayısını ve p ise parametre sayısını ifade eder. (1) eşitliğini matris formatında yazacak olursak:

$$Y = X\beta + \varepsilon$$

Burada $Y = (Y_1, Y_2, \dots, Y_n)'$, $X = (X_1, X_2, \dots, X_n)'$, $X_i = (X_{i1}, X_{i2}, \dots, X_{ip})'$ $\beta = (\beta_1, \beta_2, \dots, \beta_p)'$ ve $\varepsilon = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)'$ şeklinde ifade edilmektedir ve (\cdot) kullanılan üst simge bir vektörün veya matrisin devriğini almayı gösterir.

Çoklu doğrusal regresyon analizinde her bir gözlem için bağımsız değişkenler arasında tam veya tama yakın doğrusal bir korelasyon olursa ortaya çıkan probleme çoklu doğrusal bağlantı (ÇDB) adı verilir. Bu durumda doğrusal ve yansız tahmin edicileri içinde küçük varyansa sahip olan EKK tahmin edicisi önemli oranda olumsuz etkilenmektedir [1]. Modeldeki bağımsız değişkenlerin aynı eğilime sahip olmaları, modelin değişken sayısının gözlem sayısından çok olması, kullanılan veri toplama yöntemlerinin uygun olmaması çoklu doğrusal bağlantının nedenleri arasında gösterilebilir. Bağlantı sorunun çözümü için panel veri, soruna neden olan değişkenleri modelden çıkarmak ve yanlış tahmin edicileri kullanarak bağlantı sorununu çözmek için başvurulacak yollardandır.

Çoklu doğrusallık probleminin tarihi Frisch [2] çalışmasına dayanır. İlk olarak Hoerl ve Kennard [3, 4] tarafından tanımlanan Ridge regresyon bu probleme çözüm yöntemlerinden biridir. Ridge regresyonun ayar parametresinin bulmak araştırmacıların temel ilgisidir. Pek çok araştırmacı bu ayar parametresinin tahminini elde etmek farklı yöntemler önermişlerdir. Bunlardan bazıları, McDonald ve Galarneau [5], Monte Carlo simülasyonları ile k ayar Ridge parametresini belirlemek için iki analitik yöntem önermiş ve hata kareler ortalaması açısından değerlendirmişlerdir. Lawless ve Wang [6], Ridge ve diğer regresyon parametre tahmin edicileriyle yaptıkları simülasyon çalışması sonucunda, yaygın olarak kullanılan iki hata kareler ortalaması kriterine göre, iki sıradan Ridge tahmin edicisinin hem en küçük karelerden hem de diğer tahmin edicilerden önemli ölçüde daha iyi performans gösterdiğini belirtmişlerdir. Golub, Heath ve Wahba [7], iyi bir Ridge parametresi seçme yöntemi olarak genelleştirilmiş çapraz doğrulama yöntemini incelemişlerdir. Khalaf ve Shukur [8], tasarım matrisinin sütunları arasında çoklu bağlantı olduğunda, Ridge ayar parametresinin k seçimi için yeni bir yaklaşım önermiş ve hata kareler ortalaması açısından simülasyon teknikleri ile değerlendirmişlerdir. Alkhamisi, Khalaf ve Shukur [9], tasarım matrisinin sütunları arasında çoklu bağlantı olduğunda k Ridge ayar parametresini seçmek için Khalaf ve Shukur [8] tarafından önerilen tahmin edicilerin geliştirilmiş versiyonu olan tahmin edicilere dört değişiklik önermişlerdir. Alkhamisi ve Shukur [10], k Ridge ayar parametresini elde etmek için yeni bir yaklaşım önermiş ve ardından Monte Carlo simülasyonları ile değerlendirmişlerdir. Lukman ve Olatunji [11], regresyon katsayılarından bağımsız ve standart regresyon hatasının bir fonksiyonu olan bir Ridge parametresi önermişlerdir. Owolabi, Ayinde ve Alabi [12], iki parametrelili bir Ridge tipi tahmin edici önermişlerdir ve istatistiksel özelliklerini teorik olarak ve Monte Carlo simülasyon çalışmaları yoluyla ortaya koymuşlardır.

Bu çalışma aşağıdaki şekilde organize edilmiştir: Ridge regresyonun ayar parametre seçim kriterleri çalışmanın ikinci bölümde ayrıntıları ile birlikte incelenecektir. Çalışmanın üçüncü bölümünde, Monte Carlo simülasyon modeli ile seçim kriterlerinin performansları karşılaştırılmıştır. Bu karşılaştırmalara, literatürde en popüler cezalı tahmin edicilerinden biri olan En Küçük Mutlak Büzülme ve Seçim Operatörü (LASSO) tahmin edicisi de eklenmiştir. Çalışmanın dördüncü bölümünde ekonometrik bir veri seti kullanılarak, doğrudan yabancı yatırımı etkileyen faktörlerle kurulan çoklu doğrusal regresyon modeliyle

tahminler yapılacaktır. Yapılan ön incelemeler sonucunda, ilgili verinin bağımsız değişkenler arasında korelasyon olması sebebiyle EKK yönteminden elde edilen tahminin varyansının olması gerektiğinden daha büyük çıkmasına neden olmaktadır. Söz konusu bu durumda çoklu doğrusal regresyon modelinin tahmininde kullanılan EKK yöntemi doğru sonuçlar verememektedir. Bağlantı sorununun çözümü için önerilen yöntemlerden yanlı regresyon tahmin edicisi olan Ridge regresyon ve modele dahil edilen yanlılık parametresi ile analiz tekrardan yapılip sorun giderilmeye çalışılmaktadır. Çalışmanın asıl amacı da modele dahil edilen Ridge regresyon ayar parametresi olan k 'nın seçimi için kullanılan kriterler de analize dahil edilerek en iyi sonucu veren kriter seçimi de yapılmıştır.

2. Ridge regresyon

ÇDB problemi durumunda EKK yönteminin kullanılması tahminlerin büyük varyansa sahip olmasına neden olmaktadır. 1970 yılında Hoerl ve Kennard çoklu doğrusal bağlantı problemi olduğunda bu problemi gidermek için Ridge tahmin ediciyi önermişlerdir. Böylelikle daha küçük varyanslı tahmin ediciler elde edilmektedir. ÇDB problemi olduğunda $X'X$ matrisi tekil değildir. Hoerl ve Kennard ilk kez 1962 yılında $X'X$ matrisine k negatif olmayan bir sayı olmak üzere, kI_p sabitini modele ekleyerek, Ridge tahmin edicisini elde etmişlerdir [3].

$X'X$ matrisinde ÇDB' dan dolayı bir veya daha fazla öz değer küçük olacağını ve bu nedenle β ile onun EKK tahmin edicisi $\hat{\beta} = (X'X)^{-1}X'Y$ ile değerleri arasındaki uzaklığın yüksek olacağını açıklamışlardır. Bu sorunun çözümü için de en uygun tahmin edicinin Ridge tahmin edicisi olduğunu söylemişlerdir [3].

Açıklayıcı değişkenler arasında ÇDB olması durumunda Ridge regresyon yöntemi ile tahmin edilen β regresyon katsayılarının EKK yöntemiyle yapılan tahminlerden daha küçük HKO'ya sahip olduğu Hoerl ve Kennard [3] tarafından gösterilmiştir. Hoerl ve Kennard [3] tarafından önerilen Ridge regresyon tahmin edicisi aşağıdaki kayıp fonksiyondan elde edilmiştir;

$$L(\beta) = \sum_{i=1}^n (Y_i - \beta_0 - X'_i\beta)^2 + k \sum_{j=1}^p \beta_j^2 .$$

β 'ya türevini alıp, sifira eşitlenirse;

$$\hat{\beta}_k = (X'X + kI_p)^{-1}X'Y, \quad (2)$$

elde edilir. Burada I_p ise $p \times p$ boyutunda birim matrisidir ve $k \geq 0$ ayar parametresi olup, $k = 0$ ise $\hat{\beta}_k = \hat{\beta}$, $k = \infty$ ise $\hat{\beta}_k = 0$ 'dır.

2.1. Ridge tahmin edicisi ile EKK tahmin edicisi arasındaki ilişki

EKK tahmin edicisinin her iki tarafı da $X'X$ ile çarpılırsa;

$$X'X\beta = X'Y$$

elde edilir.

Denklem 2'de Ridge tahmin edicisi olan denklemde $X'Y$ yerine eşiti yazılırsa;

$$\hat{\beta}_k = (X'X + kI_p)^{-1}X'X\hat{\beta}$$

elde edilir.

$X'X$ matrisinin tersinin tersi kendisi olduğu için;

$$\hat{\beta}_k = (X'X + kI_p)^{-1} [(X'X)^{-1}]^{-1} \hat{\beta}$$

şeklinde yazılabilir. Her iki matriste tekil olmadıkları için;

$$\hat{\beta}_k = [(X'X)^{-1} (X'X + kI_p)]^{-1} \hat{\beta}$$

şeklinde yazılabilir. Buradan da,

$$\hat{\beta}_k = [(X'X)^{-1} X'X + k(X'X)^{-1}]^{-1} \hat{\beta}$$

şeklinde ifade edilir. Yapılan bu işlemlerden sonra

$$\hat{\beta}_k = [I_p + k(X'X)^{-1}]^{-1} \hat{\beta}$$

olur. $Z = [I_p + k(X'X)^{-1}]^{-1}$ olarak tanımlanırsa;

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

şeklinde ifade edilir. Bu eşitlik Ridge tahmin edicisinin EKK tahmin edicisinin bir dönüşümü olduğunu göstermektedir [13].

2.2. Ridge regresyon parametre seçim kriterleri

Ayar parametresinin seçimi yanlı regresyon tahmin edicilerinin performanslarını etkileyen en önemli etkidir. Ayar parametre tahmini için birçok kriter bulunmaktadır. Bu kriterlerden AIC, BIC, CV, GCV ve CP alt başlıklar halinde verilecektir. Bu kriterlerde kullanılan değerler aşağıdaki şekilde özetlenmiştir.

$$\hat{\varepsilon} = Y - X\hat{\beta}_k$$

$$H = X(X'X + kI_p)^{-1}X'$$

$$df = tr(H)$$

$$HKT(\text{Hata Kareler Toplamı}) = \hat{\varepsilon}'\varepsilon$$

Burada, H izdüşüm matrisi olarak bilinmektedir, df serbestlik derecesi ve $tr(\cdot)$ ifadesi bir matrisin izini göstermektedir.

2.2.1. Akaike Bilgi Kriteri

Akaike'nin bilgi kriterleri, Hirotugu Akaike [14] tarafından geliştirilmiştir. AIC, tahmin edilen herhangi bir istatistiksel modelin uyum iyiliğinin bir göstergesi olarak adlandırılabilir. AIC genellikle yüksek boyutlu gerçekliğe sahip bilinmeyen bir model bulmaya çalışır. Bu modellerin AIC'deki gerçek modeller olmadığı anlamına gelir. AIC kriterleri asimptotik olarak çapraz geçerliliğe eşdeğerdir. AIC aşağıdaki şekilde tanımlanmaktadır.

$$AIC = n \times \log(HKT) + 2 \times df ,$$

Burada, n gözlem sayısını ve $\log(\cdot)$ doğal logaritmayı göstermektedir.

2.2.2. Bayes Bilgi Kriteri

Bayesian bilgi kriterleri [15] Gideon E. Schwarz Bayesçi bilgi ölçütünü geliştirmiştir. BIC, farklı sayıda parametreye sahip bir grup parametrik model arasında bir model seçimidir. BIC serbest parametreleri daha güçlü cezalandırmaktadır. Bayes Bilgi kriterleri sadece gerçek modellerle karşılaşmaktadır. Bayes bilgi kriterlerinin tutarlı olduğu söylenebilir.

$$BIC = n \times \log(HKT) + df \times \log(n).$$

2.2.3. Çapraz Geçerlilik Model Seçim Kriteri

Eğer kare hatalar kullanılacaksa iyi bir tahmin edici olarak birini dışarıda bırakarak çapraz geçerlilik tahmin edicisi önerilebilir. Bu metot modellerin öngörü yeteneğine göre model seçimi yapılmasını benimsemektedir. Çapraz geçerlilik modeli;

$$CV = \sum_{i=1}^n (Y_i - \hat{Y}_i^{-1})^2 = \sum_{i=1}^n \left(\frac{Y_i - \hat{Y}_i}{1 - h_{ii}} \right)^2$$

olarak elde edilir. Burada h_{ii} , H şapka matrisinin i . köşegen elemanıdır. Ridge ve LASSO regresyon parametre seçiminde K kat çapraz geçerlilik kullanılmaktadır. n sayıda gözlemi olan veri seti K sayıda eşit parçaya bölünür. K parçaya bölünen veri için $K - 1$ parça eğitim ve geriye kalan ise test seti olarak seçilir. İşlem K adım tekrarlanarak her adımda farklı test seti seçilir. Söz konusu her adım için hata kareler ortalaması elde edilir ve tüm değerlerin ortalaması ele alınarak çapraz doğrulama hata eğrisi elde edilir ve bu eğride minimum değeri veren parametre seçilir. Yapılan bu çalışmada $K = 10$ alınmıştır.

2.2.4. Genelleştirilmiş Çapraz Geçerlilik Ölçütü

$$CV = \sum_{i=1}^n (Y_i - \hat{Y}_i^{-1})^2 = \sum_{i=1}^n \left(\frac{Y_i - \hat{Y}_i}{1 - h_{ii}} \right)^2$$

Yukarıdaki CV modeli üzerinde yapılan değişikliklerle elde edilen GCV ölçütü CV modelindeki h_{ii} yerine \hat{H} matrisinin ortalaması alınarak elde edilen bir ölçüttür. GCV modeli ise;

$$GCV = \frac{n \times HKT}{(n - df)^2}$$

şeklindedir.

2.2.5. Mallow'un C_p Kriteri

Mallow'un model seçim kriteri aşağıdaki gibidir.

$$C_p = HKT + 2 \times df \times \hat{\sigma}^2$$

Burada; $\hat{\sigma}^2$, σ^2 'nin tahmin edicisidir. Bulunan alt küme modelleri arasında C_p değerini minimum olan model uygun model olarak seçilir.

2.4. En Küçük Mutlak Büzülme ve Seçim Operatörü

Ridge tahmin edicisinin tahmin gücü oldukça yüksektir. Ancak tahmin sürecinde birtakım problemler ortaya çıkmaktadır. Ridge regresyon katsayıları daraltan bir tahmin edici olması sebebiyle daha kararlı bir yapıya sahiptir. Fakat regresyon katsayılarını sıfıra daraltmadığı için modelin yorumlanması zorlaşır. Bu sebeple modeli yorumlamak için ve tahmin doğruluğunu elde edebilmek için bazı katsayıları sıfıra daraltmak söz konusu olabilir. Bu işlemden sonra varyans küçülmekte fakat yanlış açıdan sapma meydana gelmektedir. Çıkan sonuçta tahmin doğruluğu daha güvenilir olduğundan bu tahmin edicisinin yansızlık özelliği göz ardı edilebilir [16]. 1996 yılında Tibshirani, LASSO yöntemi ile bazı katsayıları sıfıra daraltarak parametre tahminini ve model seçimini eş zamanlı yapmıştır [17]. LASSO yöntemi genel olarak aşağıdaki gibidir;

$$\hat{\beta}_{LASSO} = \underset{\beta}{\operatorname{argmin}} \left\{ \sum_{i=1}^n (Y_i - X_i' \beta)^2 + k \sum_{j=1}^p |\beta_j| \right\},$$

Burada k ayar parametresidir.

3. Simülasyon Çalışması

Bu bölümde, Ridge ayar parametresinin seçilmesinde kullanılacak beş yöntemin (AIC, BIC, CV, GCV, Cp) ve LASSO'nun karşılaştırılması Monte Carlo simülasyon için çalışması ile yapılmıştır. McDonald ve Galarneau çalışması [5] takip edilerek farklı derecelerde çoklu bağlantıya sahip veri matrisi üretebilmek için, bağımsız değişkenler aşağıdaki şekilde üretilmiştir;

$$X_{ij} = (1 - \rho^2)^{1/2} z_{ij} + \rho z_{i(p+1)} \quad i = 1, 2, \dots, n \quad j = 1, 2, \dots, p,$$

Burada p , bağımsız değişken sayısı, n , örneklem hacmi olmak üzere ρ^2 herhangi iki açıklayıcı değişken arasındaki korelasyon katsayısı, z_{ij} ise $N(0,1)$ dağılımından olan rassal sayılardır. Amacımız, çoklu doğrusallığın derecesinin yüksek ve düşük olduğu durumlarda Ridge, LASSO ve EKK tahmin edicilerinin performanslarını karşılaştırmak ve ÇDB durumunda en iyi tahmini veren yöntemi seçmek amaçlanır. Bağımsız değişkenler arasındaki farklı korelasyon değerleri ile beş örnek için simülasyon çalışması yapılarak sonuç elde etmek amaçlanmıştır [18]. Simülasyonda kullanılacak doğrusal regresyon modeli aşağıdaki gibidir;

$$Y = X\beta + \sigma\varepsilon \quad \varepsilon \sim N(0, I)$$

Modellerin değerlendirilmesi için simülasyonda eğitim verisi ve test verisi için n birimlik örneklem kullanılmıştır. Simülasyon için R Studio'da *glmnet* paketi kullanılmıştır. Eğitim setinde en iyi ceza parametresinin değeri model seçim kriterleri ile belirlenip en iyi model oluşturmak amaçlanmaktadır. Ardından test seti ile modelin iyiliği test edilmektedir. Simülasyonlarda tüm değişkenleri eğitim setine dayalı olarak ortalarız. $\bar{X}_{eğitim} = (\bar{X}_{1,eğitim}, \dots, \bar{X}_{p,eğitim})$ eğitim verilerini vektörlerini belirtir. Test veri setindeki gözlem sayısı n_{test} ve $\bar{Y}_{eğitim} = (Y_{eğitim} - \bar{Y}_{eğitim})$ eğitim verilerinde ortalama yanıtları gösterir. Son olarak iki performans ölçüsü hesaplanmıştır. Bunlar, test hatası $HKO_Y = \frac{1}{n_{test}} r'_{sim} r_{sim}$ ise $r_{i,sim} = X_i \beta - (\bar{Y}_{eğitim} + (X_i + \bar{X}_{eğitim})' \hat{\beta})$ ve tahminin hata kareler ortalaması ise $HKO_{\beta} = |\hat{\beta} - \beta|^2$ olarak gösterilir [19].

Sırasıyla eğitim ve test setini göstermek için ./ gösterimi kullanılacaktır. Simülasyon çalışması için incelenen örnekler aşağıdaki gibidir.

1. Her veri seti 20/200 gözlemden oluşmaktadır. $\beta = (3, 1.5, 0, 0, 2, 0, 0, 0)'$ olarak $\sigma = 1, 3$ ve $X \sim N(0, \Sigma)$ ise $\Sigma_{ij} = \rho^{|i-j|}$, $\rho = 0.3, 0.6, 0.9$

2. Her veri seti 20/200 gözlemden oluşmaktadır. $\beta = \left(\frac{0.85, 0.85, \dots, 0.85}{8} \right)'$ olarak $\sigma = 1,3$ ve $X \sim N(0, \Sigma)$ ise $\Sigma_{ij} = \rho^{|i-j|}$, $\rho = 0.3, 0.6, 0.9$
3. Her veri seti 20/200 gözlemden oluşmaktadır. $\beta = (3, 1.5, 0, 0, 0, -1, -1)'$ olarak $\sigma = 1,3$ ve $X \sim N(0, \Sigma)$ ise $\Sigma_{ij} = \rho^{|i-j|}$, $\rho = 0.3, 0.6, 0.9$
4. Her veri seti 100/200 gözlemden ve 30 bağımsız değişkenlerden oluşmaktadır. $\beta = \left(\frac{2, \dots, 2}{8}, \frac{0, \dots, 0}{22} \right)'$ olarak $\sigma = 1,3$ ve $X \sim N(0, \Sigma)$ ise $\Sigma_{ij} = \rho^{|i-j|}$, $\rho = 0.3, 0.6, 0.9$
5. Her veri seti 100/200 gözlemden ve 40 bağımsız değişkenlerden oluşmaktadır. $\beta = \left(\frac{0, \dots, 0}{10}, \frac{2, \dots, 2}{10}, \frac{0, \dots, 0}{10}, \frac{2, \dots, 2}{10} \right)'$ olarak $\sigma = 1,3$ ve $X \sim N(0, \Sigma)$ ise $\Sigma_{ij} = \rho^{|i-j|}$, $\rho = 0.3, 0.6, 0.9$

Yukarıda verilen örnek modellerini 1000 veri seti ile simülasyon yaparak araştırıyoruz. Bu örnek simülasyon çalışmaları aşağıdaki çizelgelerde verilmiştir. Bu örneklerle yapılan simülasyon çalışmasına Koşul İndisi (CI) eklenerek çoklu doğrusal bağlantının varlığı durumunda seçim parametrelerinin sonuçları karşılaştırılmıştır [20]. Sonuçları daha kolay karşılaştırmak için, Rölatif HKO ($RHKO$) değeri şu şekilde tanımlanmıştır: $RHKO(\hat{\beta}^*) = \frac{MSE(\hat{\beta})}{MSE(\hat{\beta}^*)}$, burada $\hat{\beta}^*$ yukarıda belirttiğimiz tahmin edicilerden bir tanesidir. Eğer, $RHKO(\hat{\beta}^*)$ değerinin 1'den büyük ise, o zaman $\hat{\beta}^*$ tahmin edicisi EKK tahmin edicisinden daha üstün performans gösterdiğini belirtir. Simülasyon sonuçlarına bakıldığında tüm örneklerde EKK, $RHKO_\beta$ ve $RHKO_Y$ göre de en kötü sonuçları vermiştir. Simülasyon çıktılarının yorumları aşağıda verilmiştir.

Çizelge 1: Örnek 1 için Simülasyon Çıktısı

ρ	CI	σ	YÖNTEM	HKO_β	$RHKO_\beta$	HKO_Y	$RHKO_Y$
0.3	1.886	1	EKK	0.894	1.000	0.887	1.000
			RIDGE(GCV)	0.827	1.081	0.839	1.057
			RIDGE(AIC)	0.837	1.068	0.842	1.053
			RIDGE(BIC)	0.830	1.078	0.840	1.056
			RIDGE(CP)	0.827	1.081	0.840	1.056
			RIDGE(CV)	0.855	1.046	0.892	0.995
			LASSO(CV)	0.591	1.514	0.625	1.419
		3	EKK	8.049	1.000	7.985	1.000
			RIDGE(GCV)	5.104	1.577	5.595	1.427
			RIDGE(AIC)	5.546	1.451	5.882	1.357
			RIDGE(BIC)	5.448	1.478	6.014	1.328
			RIDGE(CP)	5.040	1.597	5.559	1.436
			RIDGE(CV)	5.013	1.606	5.529	1.444
			LASSO(CV)	5.116	1.573	5.587	1.429
0.6	3.605	1	EKK	1.494	1.000	0.887	1.000
			RIDGE(GCV)	1.249	1.195	0.802	1.106
			RIDGE(AIC)	1.288	1.160	0.807	1.100
			RIDGE(BIC)	1.257	1.189	0.805	1.102
			RIDGE(CP)	1.246	1.198	0.805	1.102
			RIDGE(CV)	1.250	1.195	0.872	1.017
			LASSO(CV)	0.863	1.730	0.606	1.464
		3	EKK	13.443	1.000	7.958	1.000
			RIDGE(GCV)	6.397	2.102	4.954	1.612
			RIDGE(AIC)	7.596	1.770	5.405	1.477
			RIDGE(BIC)	6.466	2.079	5.143	1.553
			RIDGE(CP)	6.246	2.152	4.961	1.610

			RIDGE(CV)	5.705	2.356	4.621	1.728
			LASSO(CV)	6.609	2.034	4.971	1.606
			EKK	6.322	1.000	0.887	1.000
			RIDGE(GCV)	3.541	1.785	0.633	1.401
			RIDGE(AIC)	4.052	1.560	0.673	1.318
			RIDGE(BIC)	3.613	1.750	0.651	1.363
		1	RIDGE(CP)	3.486	1.814	0.633	1.401
			RIDGE(CV)	3.046	2.076	0.639	1.390
			LASSO(CV)	3.087	2.048	0.561	1.581
			EKK	56.895	1.000	7.985	1.000
0.9	10.880		RIDGE(GCV)	14.562	3.907	3.558	2.244
			RIDGE(AIC)	21.676	2.625	4.341	1.840
			RIDGE(BIC)	12.978	4.384	3.528	2.263
		3	RIDGE(CP)	13.739	4.141	3.518	2.270
			RIDGE(CV)	8.250	6.896	2.860	2.792
			LASSO(CV)	18.000	3.161	3.884	2.056

Örnek 1 için simülasyon sonuçlarına Çizelge 1'den bakıldığında, çoklu doğrusallığın düşük, orta ve yüksek olduğu aralıklarda $\sigma = 1$ 'de $RHKO_{\beta}$ ve $RHKO_{\gamma}$ değerlerinde EKK'ya göre en iyi sonucu veren kriter LASSO(CV)'dir. $\sigma = 3$ ise en iyi sonucu veren kriter RIDGE (CV)'dir. Çoklu doğrusallığın yüksek ve çoklu doğrusal bağlantının arttığı (CI değeri yüksek olduğu değerler) durumda ise EKK'ya göre en kötü sonucu veren kriter RIDGE(AIC) en iyi sonucu veren kriter ise RIDGE(CV) olduğu görülmüştür.

Çizelge 2: Örnek 2 için Simülasyon Çıktısı

ρ	CI	σ	YÖNTEM	HKO_{β}	$RHKO_{\beta}$	HKO_{γ}	$RHKO_{\gamma}$
			EKK	0.894	1.000	0.887	1.000
			RIDGE(GCV)	0.649	1.358	0.740	1.199
			RIDGE(AIC)	0.707	1.266	0.756	1.173
			RIDGE(BIC)	0.665	1.346	0.747	1.187
		1	RIDGE(CP)	0.652	1.373	0.742	1.195
			RIDGE(CV)	0.618	1.447	0.708	1.253
			LASSO(CV)	0.940	0.952	0.964	0.920
			EKK	8.049	1.000	7.985	1.000
0.3	1.886		RIDGE(GCV)	3.193	2.521	4.227	1.889
			RIDGE(AIC)	4.003	2.011	4.780	1.671
			RIDGE(BIC)	3.243	2.482	4.603	1.735
		3	RIDGE(CP)	3.103	2.594	4.246	1.881
			RIDGE(CV)	3.198	2.517	4.286	1.863
			LASSO(CV)	5.098	1.579	6.174	1.293
			EKK	1.494	1.000	0.887	1.000
			RIDGE(GCV)	0.686	2.176	0.587	1.511
			RIDGE(AIC)	0.843	1.771	0.632	1.405
			RIDGE(BIC)	0.683	2.188	0.595	1.492
		1	RIDGE(CP)	0.657	2.273	0.587	1.511
			RIDGE(CV)	0.602	2.483	0.543	1.633
			LASSO(CV)	1.429	1.045	0.907	0.978
			EKK	13.443	1.000	7.958	1.000
0.6	3.605		RIDGE(GCV)	3.409	3.943	3.533	2.260
			RIDGE(AIC)	4.996	2.691	4.190	1.906
			RIDGE(BIC)	3.041	4.420	3.597	2.220
		3	RIDGE(CP)	3.171	4.240	3.537	2.258
			RIDGE(CV)	3.094	4.344	3.399	2.349
			LASSO(CV)	6.427	2.092	5.367	1.488
			EKK	6.322	1.000	0.887	1.000
			RIDGE(GCV)	1.297	4.876	0.369	2.406

			RIDGE(AIC)	2.090	3.024	0.450	1.971
			RIDGE(BIC)	1.140	5.544	0.364	2.436
		1	RIDGE(CP)	1.197	5.280	0.367	2.419
			RIDGE(CV)	0.574	11.011	0.287	3.092
			LASSO(CV)	3.674	1.720	0.654	1.356
			EKK	56.895	1.000	7.985	1.000
0.9	10.880		RIDGE(GCV)	8.990	6.328	2.681	2.979
			RIDGE(AIC)	16.477	3.453	3.549	2.250
			RIDGE(BIC)	6.601	8.619	2.531	3.155
		3	RIDGE(CP)	8.577	6.633	2.669	2.992
			RIDGE(CV)	3.449	16.498	2.049	3.897
			LASSO(CV)	17.448	3.261	3.948	2.023

Örnek 2 için simülasyon sonuçlarına Çizelge 2'den bakıldığında, çoklu doğrusallığın düşük, orta ve yüksek olduğu aralıklarda $\sigma = 1$ 'de $RHKO_{\beta}$ ve $RHKO_{\gamma}$ değerlerinde EKK'ya göre en iyi sonucu veren kriter RIDGE(CV)'dir. $\sigma = 3$ ise en iyi sonucu veren kriter LASSO(CV)'dir. Fakat çoklu doğrusallığın yüksek ve çoklu doğrusal bağlantının arttığı (CI değeri yüksek olduğu değerler) durumda ise EKK'ya göre en kötü sonucu veren kriter RIDGE(AIC) en iyi sonucu veren kriter ise RIDGE(CV) olduğu görülmektedir.

Çizelge 3: Örnek 3 için Simülasyon Çıktısı

ρ	CI	σ	YÖNTEM	HKO_{β}	$RHKO_{\beta}$	HKO_{γ}	$RHKO_{\gamma}$
			EKK	0.894	1.000	0.887	1.000
			RIDGE(GCV)	0.805	1.111	0.819	1.083
			RIDGE(AIC)	0.821	1.089	0.827	1.073
			RIDGE(BIC)	0.808	1.107	0.821	1.081
		1	RIDGE(CP)	0.804	1.113	0.820	1.083
			RIDGE(CV)	0.775	1.154	0.816	1.087
			LASSO(CV)	0.660	1.355	0.691	1.285
			EKK	8.049	1.000	7.985	1.000
0.3	1.886		RIDGE(GCV)	4.773	1.686	5.321	1.501
			RIDGE(AIC)	5.321	1.513	5.704	1.400
			RIDGE(BIC)	5.161	1.560	5.860	1.363
		3	RIDGE(CP)	4.695	1.714	5.288	1.510
			RIDGE(CV)	4.686	1.718	5.265	1.517
			LASSO(CV)	4.989	1.613	5.541	1.441
			EKK	1.494	1.000	0.887	1.000
			RIDGE(GCV)	1.130	1.322	0.770	1.152
			RIDGE(AIC)	1.213	1.231	0.786	1.129
			RIDGE(BIC)	1.142	1.308	0.775	1.145
		1	RIDGE(CP)	1.116	1.338	0.771	1.150
			RIDGE(CV)	0.991	1.507	0.781	1.135
			LASSO(CV)	0.975	1.532	0.692	1.282
			EKK	13.443	1.000	7.958	1.000
0.6	3.605		RIDGE(GCV)	5.783	2.324	4.831	1.653
			RIDGE(AIC)	7.077	1.900	5.274	1.514
			RIDGE(BIC)	5.865	2.292	5.097	1.567
		3	RIDGE(CP)	5.603	2.399	4.840	1.650
			RIDGE(CV)	5.222	2.574	4.548	1.756
			LASSO(CV)	6.358	2.114	5.129	1.557
			EKK	6.322	1.000	0.887	1.000
			RIDGE(GCV)	2.838	2.227	0.578	1.534
			RIDGE(AIC)	3.491	1.811	0.629	1.411
			RIDGE(BIC)	2.843	2.224	0.589	1.507
		1	RIDGE(CP)	2.730	2.315	0.577	1.537
			RIDGE(CV)	1.815	3.484	0.540	1.644

0.9	10.880	3	LASSO(CV)	2.938	2.152	0.561	1.581
			EKK	56.895	1.000	7.985	1.000
			RIDGE(GCV)	13.620	4.177	3.581	2.230
			RIDGE(AIC)	20.929	2.719	4.345	1.838
			RIDGE(BIC)	12.436	4.575	3.734	2.138
			RIDGE(CP)	12.846	4.429	3.586	2.227
			RIDGE(CV)	8.227	6.916	3.049	2.619
			LASSO(CV)	16.571	3.433	4.022	1.986

Örnek 3 için simülasyon sonuçlarına Çizelge 3'den bakıldığında, çoklu doğrusallığın düşük olduğu aralıkta $\sigma = 1$ de $RHKO_{\beta}$ ve $RHKO_{\gamma}$ değerlerinde EKK'ya göre en iyi sonucu veren kriter LASSO(CV), $\sigma = 3$ ise en iyi sonucu RIDGE(CV) vermiştir. Çoklu doğrusallığın orta olduğu aralıkta $\sigma = 1$ de $RHKO_{\beta}$ ve $RHKO_{\gamma}$ değerlerinde EKK'ya göre en iyi sonucu veren kriter LASSO(CV), $\sigma = 3$ ise en iyi sonucu RIDGE(CV) vermiştir. Çoklu doğrusallığın yüksek ve ÇDB artışı (CI değeri yüksek olduğu değerler) durumunda ise EKK'ya göre en kötü sonucu veren kriter RIDGE(AIC) en iyi sonucu veren kriter ise RIDGE(CV) olduğu görülmektedir.

Çizelge 4: Örnek 4 için Simülasyon Çıktısı

ρ	CI	σ	YÖNTEM	HKO_{β}	$RHKO_{\beta}$	HKO_{γ}	$RHKO_{\gamma}$
0.3	2.551	1	EKK	0.530	1.000	0.457	1.000
			RIDGE(GCV)	0.498	1.065	0.439	1.042
			RIDGE(AIC)	0.504	1.052	0.440	1.038
			RIDGE(BIC)	0.493	1.075	0.444	1.029
			RIDGE(CP)	0.498	1.065	0.439	1.042
			RIDGE(CV)	0.536	0.989	0.532	0.860
			LASSO(CV)	0.220	2.409	0.224	2.037
		3	EKK	4.772	1.000	4.115	1.000
			RIDGE(GCV)	3.335	1.431	3.222	1.277
			RIDGE(AIC)	3.532	1.351	3.283	1.254
			RIDGE(BIC)	3.247	1.470	3.598	1.144
			RIDGE(CP)	3.328	1.434	3.224	1.276
			RIDGE(CV)	3.318	1.438	3.235	1.272
			LASSO(CV)	1.985	2.403	2.023	2.035
0.6	4.838	1	EKK	0.930	1.000	0.457	1.000
			RIDGE(GCV)	0.729	1.275	0.402	1.139
			RIDGE(AIC)	0.766	1.215	0.406	1.127
			RIDGE(BIC)	0.680	1.368	0.418	1.094
			RIDGE(CP)	0.728	1.277	0.402	1.138
			RIDGE(CV)	0.657	1.416	0.438	1.044
			LASSO(CV)	0.283	3.285	0.190	2.410
		3	EKK	8.369	1.000	4.115	1.000
			RIDGE(GCV)	3.556	2.353	2.538	1.621
			RIDGE(AIC)	4.051	2.066	2.622	1.569
			RIDGE(BIC)	2.838	2.949	2.882	1.428
			RIDGE(CP)	3.518	2.379	2.544	1.618
			RIDGE(CV)	3.531	2.370	2.547	1.616
			LASSO(CV)	2.546	3.287	1.708	2.409
1	EKK	4.119	1.000	4.115	1.000		
	RIDGE(GCV)	1.430	2.880	0.260	1.756		
	RIDGE(AIC)	1.670	2.467	0.270	1.694		
	RIDGE(BIC)	1.110	3.712	0.296	1.547		
	RIDGE(CP)	1.409	2.923	0.261	1.753		
	RIDGE(CV)	1.072	3.842	0.294	1.556		
	LASSO(CV)	0.933	4.415	0.155	2.952		
EKK	37.067	1.000	4.115	1.000			

0.9	18.898	3	RIDGE(GCV)	4.594	8.068	1.416	2.906
			RIDGE(AIC)	6.254	5.927	1.526	2.697
			RIDGE(BIC)	2.614	14.183	1.687	2.439
			RIDGE(CP)	4.511	8.217	1.427	2.883
			RIDGE(CV)	4.155	8.922	1.382	2.978
			LASSO(CV)	8.117	4.567	1.376	2.992

Örnek 4 için simülasyon sonuçlarına Çizelge 4'den bakıldığında, çoklu doğrusallığın düşük ve orta olduğu aralıkta $\sigma = 1$ ve $\sigma = 3$ de $RHKO_{\beta}$ ve $RHKO_{\gamma}$ değerlerinde EKK'ya göre en iyi sonucu veren kriter LASSO(CV) olduğu görülmüştür. Çoklu doğrusallığın yüksek ve ÇDB'nin yüksek oranda arttığı (CI değeri yüksek olduğu değerler) durumunda ise EKK'ya göre en kötü sonucu veren kriter $\sigma = 1$ 'de $RHKO_{\beta}$ 'ya göre RIDGE(AIC), $RHKO_{\gamma}$ 'ye göre RIDGE(BIC) en iyi sonucu ise LASSO(CV) vermiştir. $\sigma = 3$ ise en kötü sonucu $RHKO_{\beta}$ göre RIDGE(AIC), $RHKO_{\gamma}$ göre RIDGE(BIC)'dir, en iyi sonucu da $RHKO_{\beta}$ göre RIDGE(BIC), $RHKO_{\gamma}$ göre LASSO(CV) olduğu görülmüştür.

Çizelge 5: Örnek 5 için Simülasyon Çıktısı

ρ	CI	σ	YÖNTEM	HKO_{β}	$RHKO_{\beta}$	HKO_{γ}	$RHKO_{\gamma}$
0.3	2.724	1	EKK	0.825	1.000	0.710	1.000
			RIDGE(GCV)	0.787	1.048	0.688	1.033
			RIDGE(AIC)	0.797	1.035	0.691	1.028
			RIDGE(BIC)	0.783	1.053	0.691	1.027
			RIDGE(CP)	0.787	1.048	0.688	1.033
			RIDGE(CV)	0.931	0.886	0.938	0.757
			LASSO(CV)	0.514	1.604	0.488	1.454
		3	EKK	7.424	1.000	6.392	1.000
			RIDGE(GCV)	5.510	1.347	5.185	1.233
			RIDGE(AIC)	5.911	1.256	5.328	1.200
			RIDGE(BIC)	5.351	1.387	5.476	1.167
			RIDGE(CP)	5.496	1.351	5.185	1.233
			RIDGE(CV)	5.464	1.359	5.193	1.231
			LASSO(CV)	4.633	1.602	4.402	1.452
0.6	5.124	1	EKK	1.443	1.000	0.710	1.000
			RIDGE(GCV)	1.159	1.246	0.630	1.127
			RIDGE(AIC)	1.231	1.173	0.640	1.110
			RIDGE(BIC)	1.107	1.304	0.644	1.102
			RIDGE(CP)	1.156	1.249	0.630	1.127
			RIDGE(CV)	1.078	1.339	0.742	0.958
			LASSO(CV)	0.724	1.995	0.423	1.679
		3	EKK	12.991	1.000	6.392	1.000
			RIDGE(GCV)	5.736	2.265	3.980	1.606
			RIDGE(AIC)	6.879	1.889	4.180	1.529
			RIDGE(BIC)	4.804	2.704	4.390	1.456
			RIDGE(CP)	5.670	2.291	3.988	1.603
			RIDGE(CV)	5.698	2.280	3.988	1.603
			LASSO(CV)	6.543	1.985	3.825	1.671
0.9	20.497	1	EKK	6.400	1.000	0.710	1.000
			RIDGE(GCV)	2.567	2.493	0.417	1.701
			RIDGE(AIC)	3.120	2.051	0.444	1.599
			RIDGE(BIC)	2.210	2.896	0.464	1.530
			RIDGE(CP)	2.540	2.519	0.418	1.698
			RIDGE(CV)	2.222	2.880	0.567	1.252
			LASSO(CV)	2.927	2.187	0.435	1.633
		3	EKK	57.596	1.000	6.392	1.000
			RIDGE(GCV)	7.998	7.202	2.139	2.988
			RIDGE(AIC)	11.856	4.858	2.394	2.670

3	RIDGE(BIC)	5.423	10.621	2.565	2.492
	RIDGE(CP)	7.805	7.380	2.152	2.970
	RIDGE(CV)	7.341	7.845	2.080	3.073
	LASSO(CV)	22.891	2.516	3.135	2.039

Örnek 5 için simülasyon sonuçlarına Çizelge 5’den bakıldığında, çoklu doğrusallığın düşük olduğu aralıkta $\sigma = 1$ ve $\sigma = 3$ de $RHKO_{\beta}$ ve $RHKO_{\gamma}$ değerlerinde EKK’ya göre en iyi sonucu veren kriter LASSO(CV) olduğu görülmüştür. Çoklu doğrusallığın orta olduğu aralıkta $\sigma = 1$ en iyi sonucu veren kriter LASSO(CV), $\sigma = 3$ de $RMSE_{\beta}$ göre en iyi sonuç RIDGE(CV) ve $RMSE_{\gamma}$ değerlerinde ise LASSO(CV) en iyi sonucu vermiştir. Çoklu doğrusallığın ve ÇDB’ nin yüksek olduğu değerde $\sigma = 1$ için en iyi sonuç $RHKO_{\beta}$ göre RIDGE(BIC), $RHKO_{\gamma}$ göre ise RIDGE(GCV)’dir. $\sigma = 3$ için en kötü sonucu $RHKO_{\beta}$ ve $RHKO_{\gamma}$ değerlerinde LASSO(CV) vermiştir. En iyi sonucu ise $RHKO_{\beta}$ göre RIDGE(BIC), $RHKO_{\gamma}$ göre de RIDGE(CV) vermiştir.

Simülasyon sonuçlarına bakıldığında ÇDB’nin yüksek olduğu (CI yüksek olduğu değer) durumunda Ridge regresyon ayar parametresi seçim kriterleri EKK tahmininden daha iyi sonuçlar verdiği görülmüştür. ÇDB’nin en yüksek olduğu durumda EKK’dan sonra en kötü sonucu LASSO(CV) vermiştir. En iyi sonuçları ise RIDGE(BIC) ve RIDGE(CV) vermiştir.

4. Uygulamanın amacı ve modeli

Bu çalışmanın amacı, Doğrudan Yabancı Yatırımı (DYY) etkileyen faktörlerle kurulan modelle yapılan analiz sonucunda, meydana gelen ÇDB problemini çözmek için yanlış tahmin edici Ridge regresyon analizi yapmaktır. Yapılan Ridge regresyon analizi ile birlikte Ridge ayar parametre tahmin edicilerinin hangisinin daha iyi sonuç vereceği incelenmiştir.

Çalışmada kullanılan veriler Dünya Bankasından (World Bank) <https://databank.worldbank.org/> (18.10.2021) elde edilmiştir. Elde edilen veri seti, Türkiye’nin 1974-2019 yıllarına ait doğrudan yabancı yatırımlarını etkileyen faktörlerin verileridir. 46 gözlem ve 8 değişkenle kurulan çoklu doğrusal regresyon modeli aşağıdaki gibidir [21];

$$\log(DYY)_i = \beta_0 + \beta_1 KBGB_i + \beta_2 EGD_i + \beta_3 MH\dot{I}H_i + \beta_4 MH\dot{I}T_i + \beta_5 GHNTH_i + \beta_6 AKH_i + \beta_7 CHD_i + \varepsilon_i, \\ i = 1, 2, \dots, 46.$$

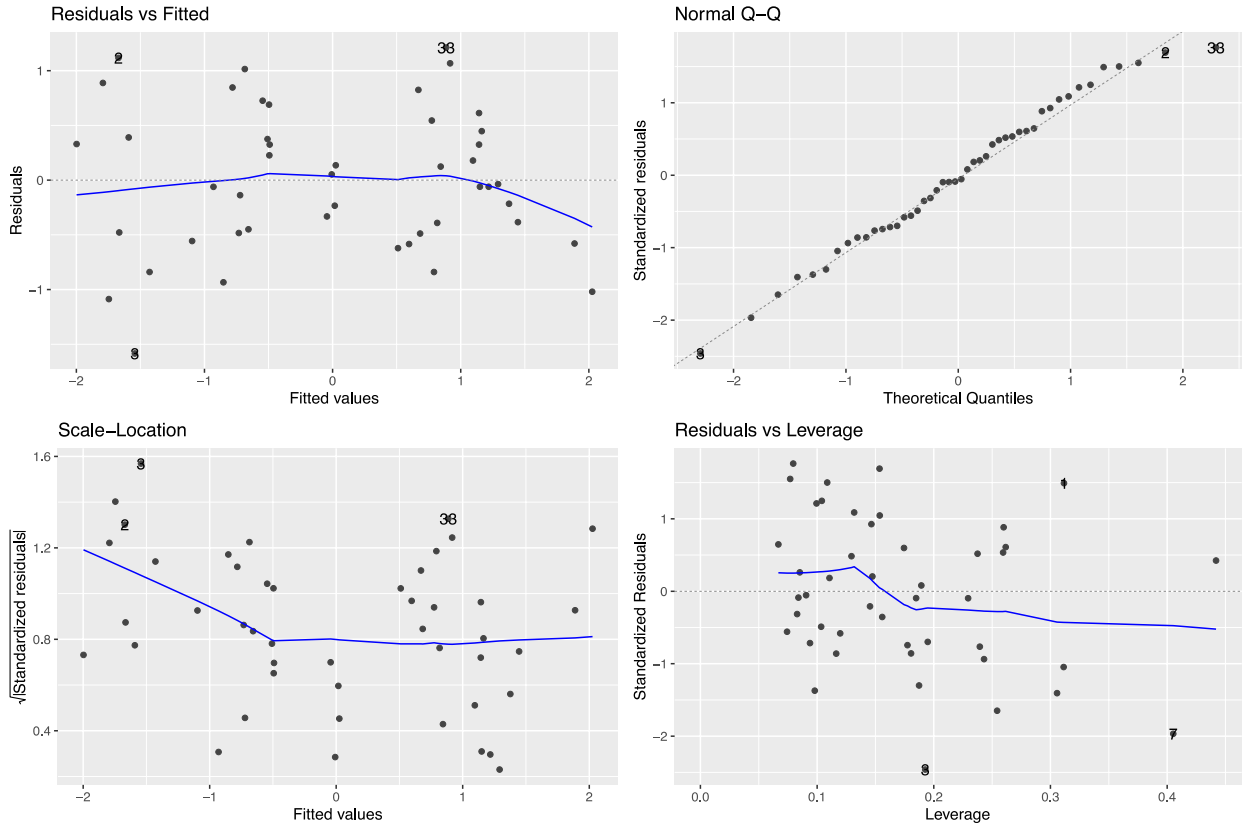
Kurulan regresyon modelindeki bağımlı ve bağımsız değişkenlerin açıklamaları Çizelge 6’da detaylı bir şekilde gösterilmiştir;

Çizelge 6: Değişkenler ve Açıklamaları

DEĞİŞKENLER		AÇIKLAMALAR
Bağımlı Değişkenler		
DYY	Doğrudan yabancı yatırımlar, net girişler (GSYİH’nin yüzdesi)	
Bağımsız Değişkenler		
KBGB	Kişi başına GSYİH büyümesi (Yıllık %)	
EGD	Enflasyon, GSYİH deflatörü (Yıllık %)	
MHİH	Mal ve hizmet ihracatı (GSYİH’nin yüzdesi)	
MHİT	Mal ve hizmet ithalatı (GSYİH’nin yüzdesi)	
GHNTH	Genel hükümet nihai tüketim harcaması (GSYİH’nin yüzdesi)	
AKH	Alınan kişisel havaleler (GSYİH’nin yüzdesi)	
CHD	Cari hesap dengesi (GSYİH’nin yüzdesi)	

4.1. Model varsayımların kontrolü

Doğrusal regresyon modelinde veriler hakkında bazı çeşitli varsayımlar bulunur. Bunları Şekil 1’de özetleyebiliriz. Şekil 1’in üst ve sol tarafında verilen grafikte tahmin edilen bağımlı değişken (x-ekseninde) ve kalıntılar (y-ekseninde) olmak üzere doğrusallık varsayımının sınanması yapılmaktadır. Burada, mavi renkteki doğru yatay eksene yakın olması doğrusallık varsayımının sağlandığını göstermektedir. Şekil 1’in üst ve sağ tarafında verilen grafikte kalıntıların normal dağılıp dağılmadığını incelemek için kullanılır. Burada, kalıntı noktaların düz kesikli çizgiyi takip etmesi normallik varsayımının dağıldığını gösterir ki, varsayım sağlanmış olur. Şekil 1’in alt ve sol tarafında verilen grafikte kalıntıların varyansının homojenliğini kontrol etmek için kullanılır. Mavi renkteki düz çizginin yatay ekseni izlemesi değişen varyans problemimizin olmadığını gösterir ki, burada varsayım sağlanmış olur. Son olarak, Şekil 1’in alt ve sağ tarafında verilen grafikte kalıntılara karşı kaldıraç grafiği olup, regresyon modelinde etkili gözlemleri tanımlamamıza yardımcı olur. Burada herhangi bir nokta Cook’un mesafesinin (normalde mavi kesikli çizgi olarak görünmeli) dışında kalmadığı için, varsayım sağlanmış olup aykırı bir değer yoktur. Ayrıca, Kolmogorov-Smirnov ve Breusch-Pagan testleri ile sırasıyla kalıntıların normallik ve sabit varyans varsayımlarının sağlandığı kontrol edilmiştir.



Şekil 1. Model varsayım grafikleri

ÇDB’yi belirleyebilmek için birçok yöntem vardır. Bunlardan bazıları korelasyon matrisinin incelenmesi, varyans artış faktörleri (VIF), öz değerler ve öz vektör analizidir. Bu yöntemler sayesinde sorunun boyutunu ve hangi değişkenden kaynaklandığı görülmektedir. İki bağımsız değişken arasındaki korelasyon katsayısının 1’e yakın olması doğrusal bağlantının olduğuna güçlü bir işarettir. MHİT ve MHİH bağımsız değişkenlerinin VIF değerlerini 10’dan büyük olduğu Çizelge 7’de belirtilen analiz sonucunda tespit edilmiş ve çoklu doğrusal bağlantının varlığı ispat edilmiştir. Ayrıca $CI = 19.86$ değerine bakıldığında 10’dan büyük olduğu görülmektedir [22]. İhracat ve ithalat verileri arasındaki kuvvetli korelasyon çoklu doğrusal bağlantı problemine neden olduğu görülmektedir.

Çizelge 7: VIF değerleri

KBGB	EDG	MHİT	GHNTH	KBGB	MHİH	CHD
7.88	2.2	45.18	3.57	2.32	54.78	7.05

Çizelge 8’de bağımsız değişkenler arasındaki korelasyon değerleri verilmiştir. Korelasyon değerleri -1 ile $+1$ arasında değerler alır. Korelasyon değerleri 0 ile 1 arasında ise pozitif yönde, -1 ile 0 arasında ise negatif yönde bir korelasyon söz konusudur. Genel olarak korelasyon katsayısının 0.6 ile 0.8 arasında olması yüksek korelasyon, 0.8 ’in üzerinde olması ise çok yüksek korelasyon olduğunu göstermektedir. Grafik incelendiğinde, bağımsız değişkenlerden MHİT ve MHİH arasında 0.938 oranında bir korelasyon olduğu görülmektedir. Buna göre ithalat ve ihracat arasındaki korelasyonun pozitif yönde oldukça yüksek olduğu görülmektedir. Diğer değişkenler arasında korelasyon değerleri ve saçılım grafiklerine bakıldığında, düşük ve orta düzeyde zayıf bir korelasyon olduğu gözlemlenmiştir. Değişkenlere bakıldığında, MHİT ile AKH arasında -0.718 , GHNTH ile AKH arasında -0.773 ve MHİH ile AKH arasında -0.717 oranında çıkan sonuçlara göre yüksek oranda ters yönde negatif bir korelasyon olduğu görülmektedir. EDG ile AKH arasında 0.566 , CHD ile EDG arasında 0.606 , GHNTH ile MHİT arasında 0.552 ve MHİH ile GHNTH arasında 0.515 oranında çıkan sonuçlara göre pozitif yönde orta düzeyde bir korelasyon olduğu görülmektedir.

Çizelge 8. Açıklayıcı değişkenler arasındaki korelasyon değerleri

	AKH	EDG	MHİT	GHNTH	KBGB	MHİH
EDG	0.566					
MHİT	-0.717	-0.339				
GHNTH	-0.773	-0.443	0.552			
KBGB	-0.167	-0.342	0.278	-0.063		
MHİH	-0.718	-0.251	0.938	0.515	0.069	
CHD	0.401	0.606	-0.246	-0.365	-0.513	0.024

Analize başlamadan önce veri seti rastgele olacak şekilde, 23 gözlemler eğitim ve 23 gözlemler test seti olarak iki parçaya bölünmüştür. Bağımlı ve bağımsız değişkenleri, eğitim veri setine dayalı olarak ortalarız, yani $\bar{X}_{eğitim} = (\bar{X}_{1,eğitim}, \dots, \bar{X}_{p,eğitim})$ ve $\bar{Y}_{eğitim} = (Y_{eğitim} - \bar{Y}_{eğitim})$. Model eğitim verisi kullanılarak kurulmuştur. Test veri setini kullanarak, herhangi bir $\hat{\beta}^*$ tahmin edicisi için TH değeri şu şekilde hesaplanmıştır: $TH(\hat{\beta}^*) = \frac{1}{n_{test}} r'_{test} r_{test}$ ise $r_{i,test} = Y_{test} - (\bar{Y}_{eğitim} + (X_i + \bar{X}_{eğitim})' \hat{\beta}^*)$. Veri setini ikiye bölme işlemi rastgele bir süreç olduğu için, rastgeleliği en az düzeye indirmek için, bu süreç 1000 kez tekrar edilmiştir ve 1000 adet TH’nin ortalaması alınarak sonuç raporlanmıştır. Çizelge 9’da yapılan analiz sonuçlarının TH ve Rölatif TH (RTH) değerleri verilmiştir. Burada $RTH(\hat{\beta}^*) = \frac{TH(\hat{\beta})}{TH(\hat{\beta}^*)}$ şeklinde tanımlanmıştır. Eğer RTH değeri 1 ’den büyük ise, bu $\hat{\beta}^*$ tahmin edicisinin EKK tahmin edicisinden daha iyi bir performans gösterdiğini gösterir.

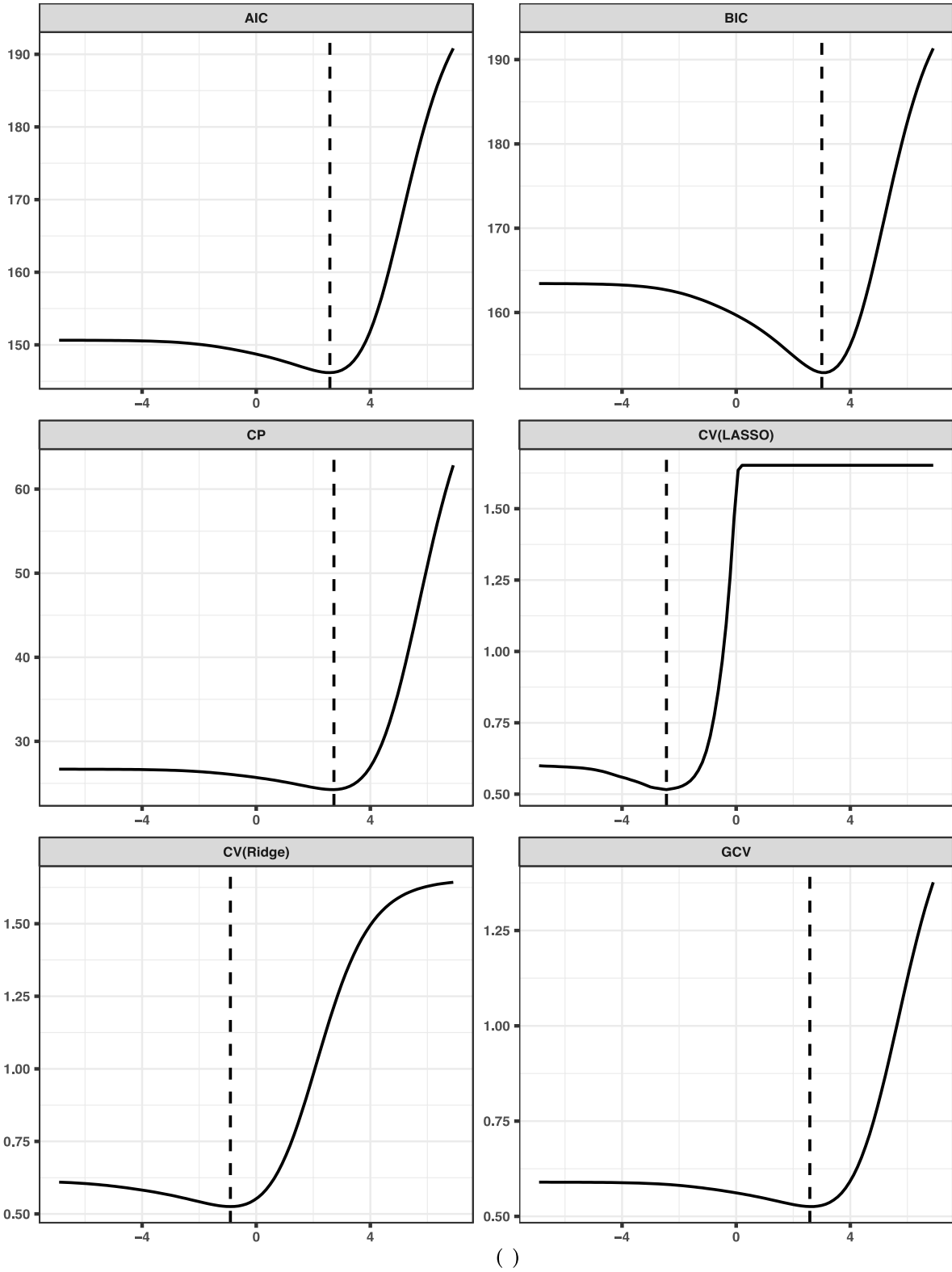
Çizelge 9: Parametre tahminleri ve tahmin edicilerin performansları

	EKK	RIDGE(GCV)	RIDGE(AIC)	RIDGE(BIC)	RIDGE(CP)	RIDGE(CV)	LASSO(CV)
AKH	-0.239	-0.209	-0.209	-0.209	-0.209	-0.209	0.197
EDG	-0.141	-0.096	-0.096	-0.091	-0.94	-0.095	0.023
MHİT	0.705	0.370	0.370	0.342	0.361	0.364	0.390
GHNTH	0.084	0.140	0.140	0.144	0.142	0.141	0.087
KBGB	-0.088	-0.028	-0.028	-0.012	-0.017	-0.018	0.000
MHİH	0.161	0.369	0.369	0.347	0.363	0.365	0.423
CHD	0.104	0.025	0.025	0.019	0.023	0.024	0.000
k	-	13.219	13.219	20.092	15.199	0.404	0.087
TH	0.973	0.676	0.785	0.666	0.676	0.656	0.685
RTH	1.000	1.440	1.240	1.462	1.440	1.483	1.421

Çizelge 9'a göre, tüm değerler için *RTH* değerleri üzerinden yorum yapılacak olursa en iyi tahmini *RIDGE(CV)* 1.483 değeriyle verdiği görülmektedir. Kurulan regresyon modelindeki bağımsız değişkenlerin bağımlı değişkene etkilerini analiz sonuçlarına göre yorumlamak da mümkündür. En iyi tahmini veren *RIDGE(CV)* kriterine bakıldığında, değişkenlerdeki %1 oranındaki artışta bağımlı değişken olan *DYY* etkileri açık olarak görülmektedir. *EKK*'ya göre en iyi sonucu veren bilgi kriteri *RIDGE(CV)*'ye göre analizi yorumlanırsa, bağımsız değişkenlerin %1'lik artışta bağımlı değişken olan *DYY*'yi etkilerini görmek mümkündür. %1'lik artışta bağımsız değişkenler *AKH* %0,20 oranında, *EDG* %0,09 oranında, *KBGB* %0,01 oranında *DYY*'yi azaltmaktadır. *MHİT* %0,36 oranında, *GHNTH* %0,14 oranında, *MHİH* %0,36 oranında ve *CHD* %0,02 oranında *DYY*'yi arttırdığı görülmektedir. Analiz yorumlarına bakıldığında, pozitif yönde *DYY*'yi etkileyen en iyi bağımsız değişkenler *MHİT* ve *MHİH* olduğu görülmüştür. Ekonomik olarak bakıldığında, doğrudan yabancı yatırımı etkileyen en iyi değişkenler ithalat ve ihracat olduğu söylenebilir. Çizelge 10'da tahmin edilen modeller verilmiştir;

Çizelge 10: Tahmin edilen modeller

<i>EKK</i>	$DYY = -0.239xAKH - 0.141xEDG + 0.705xMHİT + 0.084xGHNTH - 0.088xKBGB + 0.161xMHİH + 0.104xCHD$
<i>RIDGE(GCV)</i>	$DYY = -0.209xAKH - 0.096xEDG + 0.370xMHİT + 0.140xGHNTH - 0.020xKBGB + 0.369xMHİH + 0.025xCHD$
<i>RIDGE(AIC)</i>	$DYY = -0.209xAKH - 0.096xEDG + 0.370xMHİT + 0.140xGHNTH - 0.020xKBGB + 0.369xMHİH + 0.025xCHD$
<i>RIDGE(BIC)</i>	$DYY = -0.209xAKH - 0.091xEDG + 0.342xMHİT + 0.144xGHNTH - 0.012xKBGB + 0.347xMHİH + 0.019xCHD$
<i>RIDGE(CP)</i>	$DYY = -0.209xAKH - 0.094xEDG + 0.361xMHİT + 0.142xGHNTH - 0.017xKBGB + 0.363xMHİH + 0.023xCHD$
<i>RIDGE(CV)</i>	$DYY = -0.209xAKH - 0.095xEDG + 0.364xMHİT + 0.141xGHNTH - 0.018xKBGB + 0.365xMHİH + 0.024xCHD$
<i>LASSO(CV)</i>	$DYY = 0.197xAKH + 0.023xEDG + 0.390xMHİT + 0.087xGHNTH + 0.000xKBGB + 0.423xMHİH + 0.000xCHD$



Şekil 2. Seçim kriterleri için $\log k$ ayar parametresi grafikleri

Şekil 2’de k ayar parametresi seçimi için kullanılan kriterlerin (AIC, BIC, CP, CV(LASSO), CV(RIDGE), GCV) optimum noktaları verilmiştir.

5. Sonuç ve öneriler

Çoklu doğrusal regresyon analizinde parametrelerin tahmini için genellikle kullanılan EKK yöntemi ile ancak belirli varsayımların sağlanması durumunda güvenilir sonuçlar elde edilir. Bu temel varsayımlardan biri olan bağımsız değişkenler arasında ÇDB'nin olmamasıdır. Değişkenler arasında ÇDB'nin varlığı durumunda EKK yöntemi ile elde edilen sonuçlar güvenilir olamamaktadır. ÇDB'nin EKK tahmin edicileri üzerindeki birçok olumsuz etkisi bulunduğu için, bu problemin saptanmasına yarayan yöntemler ve bu problemin düzeltilmesi veya en aza indirgenmesi amacıyla önerilen Ridge regresyon ayrıntılarıyla incelenmiştir.

Bu çalışmada hem Monte Carlo simülasyon çalışması hem de gerçek veri uygulaması yapılmış ve sonuçlar R programı yardımı ile elde edilmiştir. Yapılan simülasyon çalışmasının sonucunda, en iyi değeri veren RIDGE(CV) kriteri olduğu gözlemlenmiştir. Gerçek veri setinde açıklayıcı değişkenler arasındaki yüksek derecede korelasyon değerlerinin olmasından dolayı EKK ile uygun bir çözüm elde edilememiştir. Ayrıca açıklayıcı değişkenler için koşul indeksi, VIF değerleri ÇDB sorununu göstermiştir. Dolayısıyla çoklu doğrusal regresyon modeline alternatif Ridge regresyon tahmin edicisi önerilmiştir. Ridge regresyon EKK yöntemine göre daha iyi sonuçlar vermiştir. Gerçek veri seti ile yapılan çalışmada tüm kriterler için RTH değerleri üzerinden yapılan yorumlar sonucunda RIDGE(CV)'nin en iyi sonucu verdiği gözlemlenmiştir.

Bundan sonraki çalışmalar için, araştırmacılara genelleştirilmiş lineer modeller için Ridge regresyon tahmin edicisi için ayar parametresinin seçimini açık problem olarak bırakıyoruz.

Teşekkür

Yazarlar, dergi Editörü ve iki hakeme, makaleyi ayrıntılı olarak okudukları ve makalenin önemli ölçüde iyileştirilmesine yol açan değerli yorum ve önerileri için teşekkür eder. Bu çalışma, İnönü Üniversitesi Sosyal Bilimler Enstitüsü Ekonometri Anabilim Dalı'nda Doç. Dr. Bahadır Yüzbaşı danışmanlığında Mustafa Pala tarafından "Ridge Regresyon Parametre Seçimi: Türkiye'nin Doğrudan Yabancı Yatırım Örneği" başlığı ile tamamlanarak 17.01.2022 tarihinde savunulan Yüksek Lisans tezinden türetilmiştir.

Kaynaklar

- [1] P. Kennedy, 2003, *A guide to econometrics*. 5th. Edition. Cambridge.
- [2] R. Frisch, 1934, Statistical confluence analysis by means of complete regression systems, *Norway, Institute of Economics Oslo*.
- [3] A. E. Hoerl, R. W. Kennard, 1970a, Ridge regression: Applications to non-orthogonal problems, *Technometrics*, 12, 1, 69-82.
- [4] A. E. Hoerl, R. W. Kennard, 1970b, Ridge regression: Biased estimation for non-orthogonal problems, *Technometrics*, 12, 1, 55-67.
- [5] G. C. McDonald, D. I. Galarneau, 1975, A Monte Carlo evaluation of some Ridge-type estimators. *Journal of the American Statistical Association*, 70, 350, 407-416.
- [6] J. F. Lawless, P. Wang, 1976, A simulation study of Ridge and other regression estimators, *Communications in Statistics*, 5, 307-323.
- [7] G. H. Golub, M. Heath, G. Wahba, 1979, Generalized cross-validation as a method for choosing a good Ridge parameter. *Technometrics*, 21,2, 215-223.
- [8] G. Khalaf, G. Shukur, 2005, Choosing Ridge parameter for regression problems. *Communications in Statistics-Theory and Methods*, 34(5), 1177-1182.
- [9] M. Alkhamisi, G. Khalaf, G. Shukur, 2006, Some modifications for choosing Ridge parameters. *Communications in Statistics-Theory and Methods*, 35, 11, 2005-2020.
- [10] M. A. Alkhamisi, G. Shukur, 2007, A Monte Carlo study of recent Ridge parameters. *Communications in Statistics-Simulation and Computation*, 36, 3, 535-547.
- [11] A. F. Lukman, A. Olatunji, 2018, Newly proposed estimator for Ridge parameter: an application to the Nigerian economy. *Pakistan Journal of Statistics*, 34,2, 91-98.

- [12] A. T. Owolabi, K. Ayinde, O. O. Alabi, 2022. A new Ridge-type estimator for the linear regression model with correlated regressors. *Concurrency and Computation: Practice and Experience*, 34,15, 6933.
- [13] S. Karakaş, 2008, *Çoklu doğrusal bağlantı problemi ve yanlış regresyon tahmincileri*, İstanbul Üniversitesi, İstanbul.
- [14] H. Akaike, 1998, Information theory and an extension of the maximum likelihood principle. In *Selected Papers of Hirotugu Akaike*, Springer, New York, NY.
- [15] G. Schwarz, 1978, Estimating the Dimension of a Model. *Annals of Statistics*, 6, 461-464. <http://dx.doi.org/10.1214/aos/1176344136>.
- [16] A. Küçük, 2019, *Doğrusal regresyonda Ridge, Liu ve LASSO tahmin edicileri üzerine bir çalışma*, Yüksek Lisans Tezi, Hacettepe Üniversitesi, Ankara.
- [17] R. Tibshirani, 1996, Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58, 267-288. <https://doi.org/10.1111/j.2517-6161.1996.tb02080>.
- [18] H. Zou, T. Hastie, 2005, Regularization and variable selection via the elastic net. *Journal of the royal statistical society: series B (statistical methodology)*, 67,2, 301-320.
- [19] G. Tutz, J. Ulbricht, 2009, Penalized regression with correlation-based penalty. *Statistics and Computing*, 19,3, 239-253.
- [20] M. Arashi, Y. Asar, B. Yüzbaşı, 2021, SLASSO: SLASSO: A scaled LASSO for multicollinear situations. *J. Stat. Comput. Simul*, 91,15, 3170-3183, <https://doi.org/10.1080/00949655.2021.1924174>.
- [21] B. Yüzbaşı, S. E. Ahmed, 2020, Ridge type shrinkage estimation of seemingly unrelated regressions and analytics of economic and financial data from “fragile five” countries. *Journal of Risk and Financial Management*, 13,6, 131, <https://doi.org/10.3390/jrfm13060131>.
- [22] E. Karaoğlan, 2019, Regresyon analizinde çoklu doğrusal bağlantı probleminin incelenmesi: Temel bileşenler, *Ondokuz Mayıs Üniversitesi, Samsun*.



İstatistikçiler Dergisi: İstatistik & Aktüerya
Journal of Statisticians: Statistics and Actuarial Sciences

İDIA 15, 2022, 1, 19-33

Geliş/Received:13.03.2022, Kabul/Accepted: 30.06.2022


Araştırma Makalesi / Research Article

Evaluation and comparison of metaheuristic methods for Markowitz's mean-variance portfolio optimization model


Nimet Yapıcı PEHLİVAN

Berat YILDIZ

Selçuk Üniversitesi Fen Fakültesi
İstatistik Bölümü 42031-Kampüs
Konya, Türkiye
nimet@selcuk.edu.tr

 0000-0002-7094-8097

beratyildiz3535@gmail.com

 0000-0002-9556-8269

Abstract

Portfolio selection is the process of selecting a combination of assets among portfolios containing multiple assets to achieve a satisfactory return on investment. Mean-variance model proposed by Markowitz (1952) has been extensively used for portfolio selection problem. It is a quadratic programming model based on the minimum risk and maximum return by choosing assets in the portfolio. Generally, classical optimization algorithms have been used for solving the quadratic programming problem. Recently, metaheuristic optimization algorithms have been used in addition to classical optimization techniques for solving portfolio selection problems. Metaheuristic methods are designed to solve complex optimization problems that cannot be solved in a reasonable time with the definitive solution methods. Various metaheuristic methods have been developed for different areas. In this study, BIST30 index data set obtained from daily closing prices of 30 stocks between December 2016 - December 2017 was used. Markowitz's mean-variance model is considered to constitute an optimal portfolio. Particle Swarm Optimization, Differential Evolution, and Artificial Bee Colony which are mostly used metaheuristic methods, are applied to determine an optimal portfolio. Performances of these methods are compared by considering risk values, i.e. portfolio variances.

Keywords: Artificial Bee Colony, Differential Evolution, Metaheuristics, Particle Swarm Optimization, Portfolio Optimization, Quadratic programming.

Öz

Markowitz'in ortalama-varyans portföy optimizasyon modeli için metasezgisel yöntemlerin değerlendirilmesi ve karşılaştırılması

Portföy seçimi, tatmin edici bir yatırım getirisi elde etmek için birden fazla varlık içeren portföyler arasından bir varlık kombinasyonu seçme sürecidir. Markowitz (1952) tarafından önerilen ortalama varyans modeli, portföy seçim problemi için yaygın olarak kullanılmaktadır. Portföydeki varlıkları minimum risk ve maksimum getiriye dayalı olarak seçen bir karesel programlama modelidir. Karesel programlama probleminin çözümü için genellikle klasik optimizasyon algoritmaları kullanılmaktadır. Son yıllarda, portföy

seçim problemlerinin çözümü için klasik optimizasyon tekniklerine ek olarak metasezgisel optimizasyon algoritmaları kullanılmaktadır. Metasezgisel yöntemler, kesin çözüm yöntemleri ile makul bir sürede çözülemeyen karmaşık optimizasyon problemlerini çözmek için tasarlanmış algoritmalar. Farklı alanlar için çeşitli metasezgisel algoritmalar geliştirilmektedir. Bu çalışmada Aralık 2016 - Aralık 2017 tarihleri arasında BİST30 endeksinde işlem gören 30 hisse senedinin günlük kapanış fiyatlarından elde edilen veri seti kullanılmıştır. Optimal portföy oluşturmak için Markowitz'in ortalama varyans modeli ele alınmıştır. Optimal portföyü belirlemek için çoğunlukla metasezgisel yöntemlerden, Parçacık Sürüsü Optimizasyonu ve Diferansiyel Evrim ve Yapay Arı Kolonisi uygulanmıştır. Bu yöntemlerin performansları, risk değerleri yani portföy varyansları dikkate alınarak karşılaştırılmıştır.

Anahtar sözcükler: Yapay arı kolonisi, Diferansiyel evrim, Metasezgiseller, Parçacık Sürü Optimizasyonu, Portföy optimizasyonu, Karesel programlama.

1. Giriş

Portfolio is the general name of all financial assets such as cash, currency, gold, stocks, etc. Portfolio selection is the process of selecting a combination of assets among portfolios containing multiple assets to achieve a satisfactory return on investment. Portfolio management is a process of creating a portfolio and deciding which investment instrument to remove from that portfolio and when to replace it. Portfolio optimization, which is the allocation of wealth among several assets, is an essential problem in modern risk management. Expected returns and risks are the most important parameters in portfolio optimization problems. Investors generally prefer to maximize returns and minimize risk. However, high returns generally involve increased risk. It is very important to find ratios of the assets in a portfolio for decision makers. Mean-variance model proposed by Markowitz (1952) that can be described in terms of the mean return of the assets and the variance of return of these assets, is extensively used for portfolio selection problem. This model is a quadratic programming model based on the minimum risk and maximum return by choosing assets in the portfolio [1]. In general, classical optimization algorithms have been used to obtain optimal ratio of the assets. Recently, in addition to classical optimization algorithms, metaheuristic optimization algorithms have been suggested for solving portfolio selection problems.

Metaheuristic optimization methods are algorithms designed to solve complex optimization problems that cannot be solved in a reasonable time with the definitive solution methods. In practice, we are usually satisfied with “good” solutions, which are obtained by heuristic/metaheuristic methods. Metaheuristic methods provide “acceptable” solutions in a reasonable time for solving hard and complex problems in science and engineering. Unlike exact (classical) optimization algorithms, metaheuristic methods do not guarantee the optimality of the obtained solutions. The metaheuristic methods have been used for NP(Non-deterministic polynomial time)-hard problems, non-linear optimization problems, large-scale examples of P (Polynomial time) problems, real time limited problems, problems where calculation of the objective function or constraint takes a long time.

Metaheuristic methods allow much faster and more efficient solutions than classical optimization methods. The metaheuristic methods inspired by nature aim to bring appropriate solutions to combinatorial optimization problems by modeling an event that takes place in nature. Just like in nature, metaheuristic methods aim for the best performance, not the best solution. Metaheuristic methods have been applied in various fields in recent years. There are so many metaheuristic optimization algorithms in the literature. For example, Simulated Annealing (SA), Tabu Search (TS), Genetic Algorithm (GA), Ant Colony Optimization (ACO), Differential Evolution (DE), Particle Swarm Optimization (PSO), Artifice Bee Colony (ABC), etc.

In this study, different metaheuristic methods, Particle Swarm Optimization, Differential Evolution, and Artificial Bee Colony Algorithm are proposed to obtain an optimal portfolio using Markowitz's mean-variance model.

The rest of the paper is organized as follows. Section 2 provides a brief review of various studies conducted on the use of metaheuristic optimization algorithms for portfolio optimization models. In Section 3, the Markowitz's mean-variance model is presented. In Section 4, algorithms of three metaheuristic methods, Particle Swarm Optimization, Differential Evolution, and Artificial Bee Colony, are explained. In Section 5, Markowitz's mean-variance model is constituted for BIST30 index data set obtained from daily closing prices of 30 stocks between December 2016 - December 2017. Then, preferred metaheuristic methods are applied to obtain an optimal portfolio and performances of these methods are compared by considering risk value. Finally, some concluding remarks are made in Section 6.

2. Literature review

This section provides a brief review of various studies conducted on the use of metaheuristic optimization algorithms for portfolio optimization models.

Crama and Schyns [2] applied Simulated Annealing approach to solve a complex portfolio selection model which is Markowitz's classical mean-variance model enriched with additional constraints. Doerner et al. [3] proposed Pareto Ant Colony Optimization for solving portfolio selection problems and compared it with other metaheuristic approaches, Simulated Annealing and Genetic Algorithm. Ehrgott and Gandibleux [4] applied Simulated Annealing, Tabu Search and Genetic Algorithm to a portfolio optimization model that extends Markowitz's mean-variance model. Cura [5] presented the Particle Swarm Optimization algorithm to the cardinality constrained mean-variance portfolio optimization problem, taking into account the data set of weekly prices from the Hang Seng index, DAX 100 index, FTSE 100 index, S&P 100 index and Nikkei index. Golmakani and Fazel [6] introduced Particle Swarm Optimization based metaheuristic method for solving the extended Markowitz mean-variance portfolio selection model and compared it with the Genetic Algorithm. Zhu et al. [7] proposed a metaheuristic approach to the portfolio optimization problem using the Particle Swarm Optimization method and tested it on various constrained and unconstrained risk investment portfolios and compared it with Genetic Algorithms. Deng et al. [8] proposed Particle Swarm Optimization method to solve cardinality constrained Markowitz portfolio optimization problem. Test results showed that the proposed PSO is much more powerful and effective than existing PSO algorithms, especially for low-risk investment portfolios. Lwin and Qu [9] proposed a hybrid algorithm combining Population Based Incremental Learning and Differential Evolution algorithms for portfolio selection problems. Çelenli et al. [10] applied classical and guaranteed convergence Particle Swarm Optimization methods for portfolio optimization created from stocks included in the BIST 30 index. The results were compared with the results obtained from mathematical programming. Akyer et al. [11] developed the Particle Swarm Optimization algorithm to solve a portfolio optimization problem with cardinality constraints applied to the Istanbul Stock Exchange data. Doering et al. [12] systematically reviewed the scientific literature on the use of metaheuristic optimization algorithms to solve portfolio optimization and risk management problems. Kalaycı et al. [13] analyzed publications based on deterministic models and applications in the mean-variance portfolio optimization literature. The publications for solving this type of problems by exact solution techniques which are nonlinear programming, and inexact solution techniques which are metaheuristic methods and machine learning algorithms, were reviewed in a systematic way. Kalaycı et al. [14] introduced a hybrid metaheuristic algorithm combining continuous ant colony optimization, artificial bee colony optimization, and genetic algorithms to obtain optimal solutions for cardinality constrained portfolio optimization problem. Corazza et al. [15] proposed a hybrid metaheuristic based on Particle Swarm Optimization for a portfolio selection problem formulated as an unconstrained problem by means of penalty functions.

3. Markowitz's mean-variance portfolio optimization model

The mean -variance model proposed by Markowitz (1952) who is the founder of modern portfolio theory, has been widely used for the portfolio selection problem. Markowitz's mean variance model is a quadratic programming model based on selecting assets in the portfolio with minimum risk and maximum return [1]. In general, classical optimization algorithms have been applied to solve the quadratic programming problem. In addition to classical optimization techniques, metaheuristic optimization algorithms have been frequently used in the solution of portfolio selection problems in recent years.

Mathematical model of the mean-variance model is given as follows:

$$\begin{aligned}
 \text{Min } Z &= \sum_{i=1}^n \sum_{j=1}^n x_i x_j \sigma_{ij} \\
 \sum_{i=1}^n x_i E(r_i) &\geq R \\
 \sum_{i=1}^n x_i &= 1 \\
 0 \leq x_i &\leq 1, \quad i = 1, 2, \dots, n
 \end{aligned} \tag{1}$$

In Equation (1); n is the number of available assets, σ_{ij} is the covariance value between asset i and asset j , x_i is the ratio of the asset i in the portfolio, R is the expected return level, and $E(r_i)$ is the expected return of the asset i .

4. Metaheuristic Methods

Heuristics is a solution strategy by trial-and-error to produce acceptable solutions to a complex problem within an acceptable time. Metaheuristic methods are defined as further development over the heuristic algorithms. Most metaheuristic methods are inspired by nature. Some nature-inspired algorithms have been proposed based on Darwin's evolution theory. On account of this, the methods are also called biology-inspired or bio-inspired. Unlike classical optimization algorithms, metaheuristic methods ensure "acceptable" solutions within a reasonable time to solve hard and complex optimization problems. Any metaheuristic methods have two major components: selection of the best solutions and randomization. The selection of the best solutions allows that convergence to the optimality, whereas randomization prevents solutions from jamming in the local optimum as well as increase the variety of solutions. The good combination of these two components usually makes it possible to achieve the global optimality. There are various classification criteria used for metaheuristic methods: nature inspired vs nonnature inspired, memory usage versus memoryless methods, deterministic versus stochastic, population-based search versus single-solution based search, and iterative versus greedy. For instance, metaheuristic methods are divided into two groups in accordance with the solution numbers: single-solution based (eg, SA) that uses a single solution and population-based (e.g., GA, PSO, DE, and ABC) that uses a whole population of solutions, during the search [16, 17].

The three metaheuristic methods, PSO, DE, and ABC, used for obtaining optimal portfolio are briefly explained in the following subsections.

4.1. Particle swarm optimization

Particle swarm optimization (PSO), one of the stochastic population-based metaheuristic, is designed for continuous optimization problem. The first application of the PSO is carried out by Kennedy and Eberhart (1995, 2001) for optimization problems [16, 25, 26, 27]. On the basis of the behavior of a colony or swarm of insects, such as ants, termites, bees, and wasps; a flock of birds; or a school of fish, the PSO mimics the social behavior of these organisms [26, 28]. The particle means that a bee in a colony or a bird in a flock. If one particle finds out a good path to food, the rest of the swarm will also be able to follow the good path instantly even if their location is far away in the swarm. The swarm is assumed to be of specified or fixed size with each particle located initially at random locations in the multidimensional design space. It is suggested that each particle have two characteristics: a position and a velocity. Each particle circuits around in the design space and remembers the best position concerning the food source or objective function value. The particles communicate information or good positions to each other and regulate their individual positions and velocities on the basis of received information about the good positions [28].

The purpose of PSO is to find the best global among all the current best solutions until the objective function no longer improves or until a predetermined number of iterations [29]. The population is formed of a set of particles. Each particle records its personal best position ($pbest$) and swarm's best position found by all particles ($gbest$). Then, all particles update their velocity (v_i) and position (x_i) in each iteration [16, 17, 29].

Due to the ease of application and high convergence speed, the PSO is one of the most widely used metaheuristic methods [29, 30]. Also, PSO is much simpler than GA and ant algorithms as it does not use mutation/crossover operators or pheromones [17].

Pseudo code of Particle Swarm Optimization (PSO) Algorithm

Determine the objective function $f = f(x)$

Set particle numbers (n), max iteration number (G)

Set inertia weight (w) and acceleration coefficients (c_1, c_2)

Initialize position (x_i) and velocity (v_i) of n particles

Produce random numbers (r_1, r_2) from uniform distribution in $[0,1]$,

Find $gbest^g$ from $\min\{f(x_1), f(x_2), \dots, f(x_n)\}$ at $g = 0$

while $g < G$ or stopping criteria

for particles $i = 1, 2, \dots, n$

Evaluate $f(x_i^g)$

Determine the $pbest_i^g$ and $gbest^g$

if $f(x_i^g) < f(pbest_i^g)$ then $pbest_i = x_i^g$

if $f(x_i^g) < f(gbest^g)$ then $gbest^g = x_i^g$

Generate new velocity $v_i^{g+1} = wv_i^g + c_1r_1(pbest_i^g - x_i^g) + c_2r_2(gbest^g - x_i^g)$

Calculate new position $x_i^{g+1} = x_i^g + v_i^{g+1}$

end for

$g = g + 1$

end while

4.2. Differential evolution

Differential evolution (DE), one of the population-based evolutionary algorithm, is proposed by Storn (1996) and Storn and Price (1997) for continuous optimization problems. Like Genetic

Algorithm (GA), DE algorithm uses crossover, mutation, and selection operators. DE uses the mutation operator as a search mechanism and selection operator to navigate the candidate areas in the search space to obtain better solutions unlike GA uses crossover operator. The mutation operator in the DE is based on the difference between randomly selected two solutions from the population [16, 33, 34]. The notation $DE/x/y/z$ is generally used to classify the different variants of the DE. x specifies the vector to be used in the mutation, y is the number of difference vectors and z indicates the crossing stage. In this study, the standard variant, $DE/rand/1/bin$, has been used. “*rand*” indicates a randomly selected population vector for x , “*bin*” implies the cross based on independent binomial experiments for z [16, 35, 36].

Pseudo code of Differential Evolution (DE) Algorithm

Determine the objective function $f = f(x)$

Set max iteration number (G) and population size ($NP \geq 4$)

Set scaling factor ($F \in (0,1]$) and crossover factor (C_r).

Initialize the population $\mathbf{x}_i^g = [x_{1,i}^g, x_{2,i}^g, \dots, x_{D,i}^g]$

Calculate the $f(x)$ of each solution of the population

while $g < G$ or stopping criteria

for $i = 1, i \leq k, i = i + 1$

Mutate and Recombine

Generate a random index $j_{rand} \in [0, D]$

for $j = 1, j \leq D, j = j + 1$ and distinct vectors $r_1, r_2, r_3; r_1 \neq r_2 \neq r_3 \neq i \in (0, NP)$

if $rand_{i,j}(0,1) \leq C_r$ or $j = j_{rand}$

then generate new donor vector by $\mathbf{u}_{j,i}^{g+1} = \mathbf{v}_{j,i}^g = \mathbf{x}_{r_1}^g + F(\mathbf{x}_{r_2}^g - \mathbf{x}_{r_3}^g)$

else $\mathbf{u}_{j,i}^{g+1} = \mathbf{x}_{j,i}^g$

Replace

$$\mathbf{x}_i^{g+1} = \begin{cases} \mathbf{u}_i^g, & \log L(\mathbf{u}_i^g) \geq \log L(\mathbf{x}_i^g) \\ \mathbf{x}_i^g, & \log L(\mathbf{u}_i^g) < \log L(\mathbf{x}_i^g) \end{cases}$$

end for $g = g + 1$

end while

4.3. Artificial Bee Colony

The Artificial Bee Colony (ABC) algorithm, population based stochastic metaheuristic algorithm, is proposed by Karaboğa (2005) for numerical optimization problem inspiring the intelligent foraging behavior of honey bees. Later, it was adapted to solve optimization problem by Akay ve Karaboğa (2012). In the ABC algorithm, artificial bees are classified into three groups: employed bees, onlooker bees, and scout bees [37, 38]. The employed bees are pioneers of the swarm discovering food sources, gathering honey, and sharing information on the nectar amount the food sources within the hive. The onlooker bees select and exploit better food sources on the basis of this information. A bee with a bad food source randomly changes to be a scout bee in search of new food sources [39]. In ABC, position of a food source indicates a possible solution to the problem and the nectar amount of a food source corresponds to the quality (fitness) of the associated solution. Due to the fact that employed bee is associated with one and only one food source, the number of employed bees is equal to the number of food sources (i.e. solutions) [38].

ABC is a very simple and robust technique that perform better than PSO and DE in most studies [17, 40].

Pseudo code of Artificial Bee Colony (ABC) Algorithm

Determine the objective function $f = f(x)$

Set the population size (SN) and number of variables to be optimized (D)

Generate the initial population $x_i = [x_{i,1}, x_{i,2}, \dots, x_{i,n}]^T$ by

$$x_{ij} = x_j^{\min} + \text{rand}[0,1](x_j^{\max} - x_j^{\min}), i = 1, 2, \dots, SN; j = 1, 2, \dots, D$$

while $g < G$ or stopping criteria

for each employed bee

Produce new solution $v_{ij} = x_{ij} + \phi_{ij}(x_{ij} - x_{mj})$ with a random number ϕ_{ij} in $[-1,1]$

Calculate the fitness value of the solution by

$$\text{fit}_i(x_i) = \begin{cases} \frac{1}{1 + f_i(x_i)}, & \text{if } f_i(x_i) \geq 0 \\ 1 + \text{abs}(f_i(x_i)), & \text{if } f_i(x_i) < 0 \end{cases}$$

Apply greedy selection process between x_i and v_i , then select the better one

Calculate the probability values $p_i = \frac{\text{fit}_i(x_i)}{\sum_{i=1}^{SN} \text{fit}_i(x_i)}$ for the solution x_i

for each onlooker bee

Select a solution x_i depending on p_i

Produce new solution v_i

Calculate the fit_i value

Apply greedy selection process between x_i and v_i , then select the better one

if there is an abandoned solution x_i

then

Replace it using a scout bee with a new solution randomly produced by x_{ij}

Memorise the best solution so far

$g = g + 1$

end while

5. Numerical evaluation and comparison of metaheuristic methods

Borsa İstanbul (BIST), formerly known as the Istanbul Stock Exchange, was established in Turkey in 1985. BIST is the institution that provides custody and clearing services to Turkish banks, foreign banks and intermediary institutions operating in the capital market. It brings together all buyers, sellers and intermediary institutions by ensuring that the transactions are carried out within legal frameworks. The BIST 30 index measures the joint performance of the stocks of 30 companies with the highest market value and transaction volume, traded on the Borsa İstanbul. In order to determine the stocks to be included in the BIST 30 index, the market values and daily average trading volumes of the stocks in circulation are ranked in decreasing order. Those who are at the top of both rankings are included in the index.

In this study, daily closing prices of 30 stocks traded on the BIST30 index between December 1, 2016 and December 29, 2017 are considered. The stocks between the dates discussed are GARAN, AKBNK, EREGL, TCELL, TUPRS, BIMAS, THYAO, KCHOL, SAHOL, ISCTR, HALKB, EKGYO, PETKM, ASELS, VAKBN, TOASO, SISE, ARCLK, YKBNK, TAVHL, TKFEN, TTKOM, KRDM, KOZAL, PGSUS, KOZAA, DOHOL, OTKAR, ECILC, SKBNK. At first, average monthly closing prices of the stocks are calculated using the daily closing prices

for the periods between December 2016 and December 2017, as shown in Table 5.1. Then, returns and expected returns (means) of all stocks are calculated for the 13 periods discussed, and are given in Table 5.2. For instance, SKBNK's return for the period December 2017 is computed by using the average closing prices of that period and the previous period, with the following formula:

$$r_{\text{SKBNK_Dec2017}} = \left(\frac{P_{\text{Dec2017}} - P_{\text{Nov2017}}}{P_{\text{Nov2017}}} \right) = \left(\frac{1,5542 - 1,3645}{1,3645} \right) = 0,1391$$

By using the returns given in Table 5.2, deviation values related to all stocks are calculated by taking the difference between the stocks' return in that period and its average of all periods. For example, deviation value of SKBNK for the period December 2017 is computed as follows:

$$\text{SKBNK_Dec2017} = (r_{\text{Dec2017}} - \bar{r}_{2017}) = (0,1391 - 0,0280) = 0,1110$$

In case of two or more stocks, the risk can be measured with the covariance matrix. Covariance matrix of all stocks is given in Table 5.3. After calculating the returns and covariance matrix for the stocks using the discussed one-year data, Markowitz's mean variance model is created for portfolio selection. The value of expected return level R is considered as 0 in the constraint $\sum_{i=1}^n x_i E(r_i) \geq R$. The created model is solved with the classical optimization technique as well as metaheuristic optimization methods, Particle Swarm Optimization, Differential Evolution, and Artificial Bee Colony.

Table 5.1. Average monthly closing prices of all stocks for the periods between December 2016 and December 2017

Period	GARAN	AKBNK	EREGL	TCELL	TUPRS	BIMAS	THYAO	KCHOL	SAHOL	ISCTR	HALKB	EKGYO	PETKM	ASELS	VAKBN
Dec 2016	7,3377	7,5118	4,7018	8,3459	65,5686	47,6273	4,8991	12,7836	8,8764	4,8686	9,0432	2,9595	3,3255	11,5905	4,3009
Jan 2017	7,5132	7,6064	5,0255	9,4282	70,0577	50,4023	5,1332	13,7232	9,2286	5,2123	10,0014	3,0091	3,6355	13,1500	4,5009
Feb 2017	8,4805	8,5565	5,4830	10,3085	78,6760	51,5310	5,6640	14,9915	10,0345	5,9840	11,3630	3,1255	3,9615	13,8550	5,1935
Mar 2017	8,6300	8,6633	5,5792	10,9679	82,6379	54,4454	5,6513	15,0679	9,9167	6,3575	11,0267	3,0729	4,4029	16,8167	5,3717
Apr 2017	9,2180	9,0960	5,7815	10,9280	85,3740	55,2950	5,6400	15,9180	10,2705	6,7845	11,0160	2,8895	4,7660	17,7305	5,8150
May 2017	9,5171	9,4014	6,4538	10,5757	93,5357	60,3914	6,7257	16,5638	10,7129	7,0019	11,9757	2,9190	5,4300	21,1229	6,0890
Jun 2017	9,6780	9,7850	6,5150	10,8225	98,9650	63,7675	7,4880	16,4480	10,8715	7,1815	12,8530	2,9595	5,8835	21,9170	6,4780
Jul 2017	10,3005	10,2662	7,5600	11,3738	104,8786	66,3471	8,6543	16,5186	10,9743	7,6843	14,1381	3,0876	6,2414	22,6686	6,9452
Aug 2017	10,6932	10,3505	7,9945	12,0332	114,9045	71,1345	9,2436	17,2668	10,5682	7,5509	14,8077	3,0350	6,2600	26,0591	7,0823
Sep 2017	10,3021	9,9521	7,8168	12,0263	121,2684	74,8332	9,4937	17,0621	10,4189	7,1595	12,8053	2,7726	6,0942	26,4689	6,6358
Oct 2017	10,0332	9,6009	8,3718	13,1255	128,0455	74,8432	9,4750	16,4095	10,1877	6,9091	11,6982	2,6595	6,3068	28,8673	6,2495
Nov 2017	10,0627	9,5245	9,1473	13,8764	128,6409	76,8223	10,7718	16,9027	10,6591	6,6145	9,9091	2,5945	6,5677	35,4795	5,8591
Dec 2017	10,0795	9,3024	9,2986	14,8857	118,4286	75,1000	14,2600	17,5490	10,6119	6,6238	10,1443	2,6624	7,1605	32,3033	6,2114
Period	TOASO	SISE	ARCLK	YKBNK	TAVHL	TKFEN	TTKOM	KRDMD	KOZAL	PGSUS	KOZAA	DOHOL	OTKAR	ECILC	SKBNK
Dec 2016	22,7059	3,1695	20,0664	3,3700	13,0282	6,1732	5,2077	1,1045	16,2105	12,6786	2,2900	0,6855	122,0605	2,7023	1,1468
Jan 2017	24,0914	3,4509	20,6418	3,4805	14,2532	6,6809	5,4564	1,1795	17,1395	14,5895	2,0995	0,8155	132,1659	2,7936	1,1882
Feb 2017	25,8875	3,5205	21,2830	3,8915	15,0140	7,4635	5,6795	1,2590	20,1875	15,9445	2,1130	0,8190	136,5820	3,0865	1,2340
Mar 2017	26,6025	3,6363	21,3717	3,8717	14,5663	8,5554	5,8058	1,1608	19,5263	14,7317	2,0271	0,6988	125,7688	3,1079	1,2050
Apr 2017	28,3500	3,7845	22,8670	4,1080	14,7540	9,0010	5,9830	1,2290	20,2640	15,2530	2,0695	0,7185	125,2300	3,4150	1,1845
May 2017	29,0400	4,1157	23,8876	4,3571	16,1252	9,6129	6,2910	1,4681	18,2943	17,2500	2,1124	0,7314	131,6429	3,8490	1,2190
Jun 2017	29,4570	4,2760	25,9220	4,4400	18,2050	8,8880	6,2855	1,6540	19,2850	18,8930	2,5775	0,7400	116,9350	4,4545	1,2165
Jul 2017	30,1333	4,4005	26,2676	4,5905	20,5438	9,7886	6,5929	1,9295	26,7200	21,3576	3,9829	0,7762	121,1333	4,7152	1,3690
Aug 2017	31,6373	4,3936	24,6755	4,5218	21,0209	12,1714	7,1005	2,1559	32,8436	26,0100	6,3277	0,8577	117,3682	4,6177	1,3555
Sep 2017	31,3916	4,1716	23,0579	4,4568	19,1911	11,6747	6,7605	2,3753	31,1326	27,1453	7,2505	0,9358	109,0053	4,4653	1,5374
Oct 2017	30,3709	4,2327	22,0318	4,3964	17,8559	12,4464	6,6223	2,5232	32,2027	27,8882	6,7232	0,9232	105,8955	4,3641	1,5155
Nov 2017	32,6136	4,4977	20,1645	4,3705	19,0132	14,3073	5,9118	2,3309	30,7627	27,7427	5,3968	0,8818	117,5545	4,5050	1,3645
Dec 2017	32,0257	4,3786	20,5610	4,2171	21,2871	15,9695	6,1814	2,7248	33,0771	31,3724	5,7257	0,8438	119,6952	4,3076	1,5543

Table 5.2. Returns and expected returns (means) of all stocks for the periods between December 2016 and December 2017

Period	GARAN	AKBNK	EREGL	TCELL	TUPRS	BIMAS	THYAO	KCHOL	SAHOL	ISCTR	HALKB	EKGYO	PETKM	ASELS	VAKBN
Dec 2016															
Jan 2017	0,0239	0,0126	0,0688	0,1297	0,0685	0,0583	0,0478	0,0735	0,0397	0,0706	0,1060	0,0167	0,0932	0,1346	0,0465
Feb 2017	0,1287	0,1249	0,0910	0,0934	0,1230	0,0224	0,1034	0,0924	0,0873	0,1481	0,1361	0,0387	0,0897	0,0536	0,1539
Mar 2017	0,0176	0,0125	0,0175	0,0640	0,0504	0,0566	-0,0023	0,0051	-0,0117	0,0624	-0,0296	-0,0168	0,1114	0,2138	0,0343
Apr 2017	0,0681	0,0499	0,0363	-0,0036	0,0331	0,0156	-0,0020	0,0564	0,0357	0,0672	-0,0010	-0,0597	0,0825	0,0543	0,0825
May 2017	0,0325	0,0336	0,1163	-0,0322	0,0956	0,0922	0,1925	0,0406	0,0431	0,0320	0,0871	0,0102	0,1393	0,1913	0,0471
Jun 2017	0,0169	0,0408	0,0095	0,0233	0,0580	0,0559	0,1133	-0,0070	0,0148	0,0256	0,0733	0,0139	0,0835	0,0376	0,0639
Jul 2017	0,0643	0,0492	0,1604	0,0509	0,0598	0,0405	0,1558	0,0043	0,0095	0,0700	0,1000	0,0433	0,0608	0,0343	0,0721
Aug 2017	0,0381	0,0082	0,0575	0,0580	0,0956	0,0722	0,0681	0,0453	-0,0370	-0,0174	0,0474	-0,0170	0,0030	0,1496	0,0197
Sep 2017	-0,0366	-0,0385	-0,0222	-0,0006	0,0554	0,0520	0,0271	-0,0119	-0,0141	-0,0518	-0,1352	-0,0864	-0,0265	0,0157	-0,0630
Oct 2017	-0,0261	-0,0353	0,0710	0,0914	0,0559	0,0001	-0,0020	-0,0382	-0,0222	-0,0350	-0,0865	-0,0408	0,0349	0,0906	-0,0582
Nov 2017	0,0029	-0,0080	0,0926	0,0572	0,0047	0,0264	0,1369	0,0301	0,0463	-0,0426	-0,1529	-0,0244	0,0414	0,2291	-0,0625
Dec 2017	0,0017	-0,0233	0,0165	0,0727	-0,0794	-0,0224	0,3238	0,0382	-0,0044	0,0014	0,0237	0,0261	0,0903	-0,0895	0,0601
Mean	0,0277	0,0189	0,0596	0,0503	0,0517	0,0391	0,0969	0,0274	0,0156	0,0275	0,0140	-0,0080	0,0670	0,0929	0,0330
Period	TOASO	SISE	ARCLK	YKBNK	TAVHL	TKFEN	TTKOM	KRDMD	KOZAL	PGSUS	KOZAA	DOHOL	OTKAR	ECILC	SKBNK
Dec 2016															
Jan 2017	0,0610	0,0888	0,0287	0,0328	0,0940	0,0822	0,0477	0,0679	0,0573	0,1507	-0,0832	0,1897	0,0828	0,0338	0,0361
Feb 2017	0,0746	0,0202	0,0311	0,1181	0,0534	0,1171	0,0409	0,0674	0,1778	0,0929	0,0064	0,0043	0,0334	0,1048	0,0386
Mar 2017	0,0276	0,0329	0,0042	-0,0051	-0,0298	0,1463	0,0222	-0,0780	-0,0328	-0,0761	-0,0407	-0,1468	-0,0792	0,0069	-0,0235
Apr 2017	0,0657	0,0408	0,0700	0,0610	0,0129	0,0521	0,0305	0,0587	0,0378	0,0354	0,0209	0,0283	-0,0043	0,0988	-0,0170
May 2017	0,0243	0,0875	0,0446	0,0606	0,0929	0,0680	0,0515	0,1945	-0,0972	0,1309	0,0207	0,0180	0,0512	0,1271	0,0292
Jun 2017	0,0144	0,0389	0,0852	0,0190	0,1290	-0,0754	-0,0009	0,1266	0,0542	0,0952	0,2202	0,0117	-0,1117	0,1573	-0,0021
Jul 2017	0,0230	0,0291	0,0133	0,0339	0,1285	0,1013	0,0489	0,1666	0,3855	0,1305	0,5452	0,0489	0,0359	0,0585	0,1254
Aug 2017	0,0499	-0,0016	-0,0606	-0,0150	0,0232	0,2434	0,0770	0,1173	0,2292	0,2178	0,5887	0,1050	-0,0311	-0,0207	-0,0099
Sep 2017	-0,0078	-0,0505	-0,0656	-0,0144	-0,0870	-0,0408	-0,0479	0,1017	-0,0521	0,0436	0,1458	0,0910	-0,0713	-0,0330	0,1342
Oct 2017	-0,0325	0,0147	-0,0445	-0,0136	-0,0696	0,0661	-0,0205	0,0623	0,0344	0,0274	-0,0727	-0,0135	-0,0285	-0,0227	-0,0143
Nov 2017	0,0738	0,0626	-0,0848	-0,0059	0,0648	0,1495	-0,1073	-0,0762	-0,0447	-0,0052	-0,1973	-0,0448	0,1101	0,0323	-0,0996
Dec 2017	-0,0180	-0,0265	0,0197	-0,0351	0,1196	0,1162	0,0456	0,1690	0,0752	0,1308	0,0609	-0,0431	0,0182	-0,0438	0,1391
Mean	0,0297	0,0281	0,0034	0,0197	0,0443	0,0855	0,0157	0,0815	0,0687	0,0812	0,1013	0,0207	0,0005	0,0416	0,0280

Table 5.3. Covariance matrix related to all stocks for the periods between December 2016 and December 2017

	GARAN	AKBNK	EREGL	TCELL	TUPRS	BIMAS	THYAO	KCHOL	SAHOL	ISCTR	HALKB	EKGYO	PETKM	ASELS	VAKBN
GARAN	0,0020	0,0019	0,0011	0,0002	0,0010	0,0001	0,0003	0,0012	0,0010	0,0023	0,0031	0,0010	0,0009	-0,0001	0,0025
AKBNK	0,0019	0,0021	0,0010	0,0000	0,0013	0,0002	0,0001	0,0010	0,0012	0,0024	0,0032	0,0010	0,0010	0,0000	0,0026
EREGL	0,0011	0,0010	0,0026	0,0003	0,0009	0,0002	0,0012	0,0004	0,0007	0,0010	0,0020	0,0011	0,0006	0,0014	0,0008
TCELL	0,0002	0,0000	0,0003	0,0022	-0,0001	-0,0005	-0,0003	0,0004	0,0001	0,0006	0,0007	0,0007	0,0000	0,0000	0,0002
TUPRS	0,0010	0,0013	0,0009	-0,0001	0,0026	0,0011	-0,0025	0,0004	0,0004	0,0012	0,0020	0,0001	0,0000	0,0020	0,0008
BIMAS	0,0001	0,0002	0,0002	-0,0005	0,0011	0,0010	-0,0007	0,0001	0,0000	0,0001	0,0008	0,0000	0,0001	0,0017	0,0000
THYAO	0,0003	0,0001	0,0012	-0,0003	-0,0025	-0,0007	0,0094	0,0007	0,0006	-0,0001	0,0027	0,0024	0,0014	-0,0034	0,0016
KCHOL	0,0012	0,0010	0,0004	0,0004	0,0004	0,0001	0,0007	0,0014	0,0009	0,0014	0,0020	0,0006	0,0007	0,0002	0,0016
SAHOL	0,0010	0,0012	0,0007	0,0001	0,0004	0,0000	0,0006	0,0009	0,0013	0,0013	0,0013	0,0006	0,0009	0,0004	0,0012
ISCTR	0,0023	0,0024	0,0010	0,0006	0,0012	0,0001	-0,0001	0,0014	0,0013	0,0035	0,0044	0,0014	0,0018	-0,0003	0,0035
HALKB	0,0031	0,0032	0,0020	0,0007	0,0020	0,0008	0,0027	0,0020	0,0013	0,0044	0,0094	0,0031	0,0026	-0,0016	0,0055
EKGYO	0,0010	0,0010	0,0011	0,0007	0,0001	0,0000	0,0024	0,0006	0,0006	0,0014	0,0031	0,0016	0,0011	-0,0005	0,0018
PETKM	0,0009	0,0010	0,0006	0,0000	0,0000	0,0001	0,0014	0,0007	0,0009	0,0018	0,0026	0,0011	0,0022	0,0007	0,0019
ASELS	-0,0001	0,0000	0,0014	0,0000	0,0020	0,0017	-0,0034	0,0002	0,0004	-0,0003	-0,0016	-0,0005	0,0007	0,0088	-0,0018
VAKBN	0,0025	0,0026	0,0008	0,0002	0,0008	0,0000	0,0016	0,0016	0,0012	0,0035	0,0055	0,0018	0,0019	-0,0018	0,0043
	TOASO	SISE	ARCLK	YKBNK	TAVHL	TKFEN	TTKOM	KRDMD	KOZAL	PGSUS	KOZAA	DOHOL	OTKAR	ECILC	SKBNK
TOASO	0,0013	0,0008	0,0002	0,0009	0,0007	0,0012	0,0001	-0,0012	0,0007	0,0002	-0,0008	0,0006	0,0012	0,0011	-0,0013
SISE	0,0008	0,0017	0,0009	0,0008	0,0014	0,0002	0,0002	-0,0006	-0,0010	0,0000	-0,0035	0,0003	0,0014	0,0017	-0,0016
ARCLK	0,0002	0,0009	0,0030	0,0014	0,0022	-0,0019	0,0015	0,0018	0,0005	0,0007	-0,0002	0,0000	-0,0005	0,0027	0,0006
YKBNK	0,0009	0,0008	0,0014	0,0019	0,0009	-0,0004	0,0007	0,0005	0,0011	0,0004	-0,0010	0,0006	0,0009	0,0023	-0,0002
TAVHL	0,0007	0,0014	0,0022	0,0009	0,0057	0,0002	0,0014	0,0028	0,0037	0,0031	0,0037	0,0008	0,0021	0,0026	0,0008
TKFEN	0,0012	0,0002	-0,0019	-0,0004	0,0002	0,0071	0,0013	-0,0019	0,0039	0,0012	0,0032	-0,0010	0,0024	-0,0022	-0,0017
TTKOM	0,0001	0,0002	0,0015	0,0007	0,0014	0,0013	0,0027	0,0025	0,0035	0,0026	0,0062	0,0012	-0,0001	0,0005	0,0013
KRDMD	-0,0012	-0,0006	0,0018	0,0005	0,0028	-0,0019	0,0025	0,0075	0,0041	0,0053	0,0116	0,0033	-0,0003	0,0010	0,0043
KOZAL	0,0007	-0,0010	0,0005	0,0011	0,0037	0,0039	0,0035	0,0041	0,0187	0,0058	0,0248	0,0033	0,0007	0,0000	0,0033
PGSUS	0,0002	0,0000	0,0007	0,0004	0,0031	0,0012	0,0026	0,0053	0,0058	0,0063	0,0118	0,0047	0,0012	0,0004	0,0023
KOZAA	-0,0008	-0,0035	-0,0002	-0,0010	0,0037	0,0032	0,0062	0,0116	0,0248	0,0118	0,0589	0,0071	-0,0050	-0,0006	0,0068
DOHOL	0,0006	0,0003	0,0000	0,0006	0,0008	-0,0010	0,0012	0,0033	0,0033	0,0047	0,0071	0,0072	0,0012	0,0001	0,0018
OTKAR	0,0012	0,0014	-0,0005	0,0009	0,0021	0,0024	-0,0001	-0,0003	0,0007	0,0012	-0,0050	0,0012	0,0045	0,0002	-0,0005
ECILC	0,0011	0,0017	0,0027	0,0023	0,0026	-0,0022	0,0005	0,0010	0,0000	0,0004	-0,0006	0,0001	0,0002	0,0046	-0,0012
SKBNK	-0,0013	-0,0016	0,0006	-0,0002	0,0008	-0,0017	0,0013	0,0043	0,0033	0,0023	0,0068	0,0018	-0,0005	-0,0012	0,0053

Analyses are performed for two population sizes (50, 100), five iteration numbers (100, 250, 500, 750, 1000), and three run times (100, 250, 500) for the metaheuristic optimization algorithms. The minimum and maximum values for the parameter variables ($x_{\min} = 0.2$, $x_{\max} = 0.8$) and the crossover value ($CR = 0.2$) are determined for DE algorithm. The cooling rate ($\alpha = 0.99$), initial temperature value ($T = 1000$), and final temperature value ($T_s = 0.1$) are considered for SA algorithm. The number of onlooker bee (Pop) and dropping limit of the food source ($0,6 * nVar(30) * Pop$) are determined for ABC algorithm. The acceleration coefficients ($c_1 = 1.5$, $c_2 = 2$) and inertia weight ($w = 1$) are considered for PSO algorithm.

Investmentable stocks and their optimal ratios with the minimum portfolio variances obtained from the PSO, DE, ABC, and classical optimization method are shown in Table 5.4.

Table 5.4. Optimal solutions obtained from metaheuristic and classical optimization methods

	Methods			
	PSO	DE	ABC	Classical optimization
Pop size	100	50	50,100	-
Run time	500	100	100,250,500	-
Iteration number	100	750	1000	-
Variance	0,0002568557	0,0002568614	0,0002568557	0,0002570465
TCELL	0,2142	0,2146	0,2142	0,2144
BIMAS	0,3943	0,3935	0,3943	0,3934
THYAO	0,0162	0,0169	0,0162	0,0169
SAHOL	0,1734	0,1705	0,1735	0,1696
ASELS	0	0	0	0,0300
TOASO	0,0270	0,0305	0,0269	0,0324
ARCLK	0,0510	0,0511	0,0510	0,0505
TKFEN	0,0279	0,0267	0,0279	0,0264
SKBNK	0,0961	0,0962	0,0961	0,0963

According to Table 5.4., minimum portfolio variance (0.0002568557) is computed using the ABC and the optimal ratios of the stocks related to this variance are BIMAS (39.43%), TCELL (21.42%), SAHOL (17.35%), SKBNK (9.61%), ARCLK (5.10%), TKFEN (2.79%), TOASO (2.69%), and THYAO (1.62%). With the classical optimization method, the optimal ratios of the stocks regarding the minimum portfolio variance (0.0002570465) are computed as BIMAS (39.34%), TCELL (21.44%), SAHOL (16.96%), SKBNK (9.63%), ARCLK (5.05%), TOASO (3.24%), ASELS (3%), TKFEN (2.64%) and THYAO (1.69%). The minimum portfolio variances with ratios of the stocks obtained from the both classical optimization method and preferred metaheuristic optimization methods are very close to each other.

6. Conclusions

The most important decision in portfolio management is the choice of asset mix that refers to the division of the portfolio investment into different asset types such as stocks, bonds, warrants, treasury bills, financing bills, asset-backed securities, repo, gold, and foreign currency. In the traditional portfolio theory, diversification is made without considering the interaction of the assets in the portfolio with each other. The key element is the number of financial assets in the portfolio. Accordingly, it is thought that the risk can be reduced by simply increasing the number of assets in the portfolio, without considering the relationships between the returns of the assets in the portfolio. However, in the modern portfolio theory, it is predicted that the assets in the portfolio move in the same or the opposite direction, and therefore the risk cannot be reduced by simply diversifying the portfolio. The modern portfolio theory is based on a number of theories that explain how the information available in the market is analyzed, how investors behave, and how these behaviors affect prices. While traditional portfolio theory is based on qualitative variables, modern portfolio theory tries to quantify the relevant variables. The mean-variance model is provided to find out which assets will be in the portfolio and the ratios of the stocks in the portfolio.

In this study, it is aimed to determine optimal portfolio with minimum risk and maximum return by using the Markowitz' mean-variance model which is a quadratic programming problem. As an application, stocks traded in the BIST30 index between December 1, 2016 and December 29, 2017 are discussed. Optimal solutions for this problem are obtained with both classical optimization and metaheuristic optimization methods, PSO, DE, and ABC. In the application of metaheuristic optimization methods, various trials are carried out for different population sizes, iteration numbers and run times. It is found that there are very small differences between the objective function values, i.e. portfolio variances, obtained from the PSO, DE, and ABC.

When the results obtained from all metaheuristic methods are compared, it is seen that the minimum portfolio variance is calculated as 0.0002568557 from the ABC. The optimal ratios of the stocks to be included in the portfolio for this variance are ordered as BIMAS (39.43%), TCELL (21.42%), SAHOL (17.35%), SKBNK (9.61%), ARCLK (5.10%), TKFEN (2.79%), TOASO (2.69%), and THYAO (1.62%). On the other hand, the optimal ratios of the stocks obtained according to the minimum portfolio variance of 0.0002570465 calculated by the classical optimization method are BIMAS (39.34%), TCELL (21.44%), SAHOL (16.96%), SKBNK (9%, 63), ARCLK (5.05%), TOASO (3.24%), ASELS (3%), TKFEN (2.64%) and THYAO (1.69%). It is seen that the portfolio variances and ratios of the stocks obtained from the all methods are very close to each other.

The differences between the metaheuristic methods discussed are due to the structure of the algorithms and the different parameters they use. Although PSO, DE, and ABC methods give optimal solutions very close to each other, they have differences in terms of fastest convergence. PSO has the fastest convergence performance, followed by ABC and DE, respectively.

In future studies, it can be suggested to obtain solutions for portfolio selection problems with mean absolute deviation models proposed by Konno and Yamazaki [41], Feinstein and Thapa [42] using the metaheuristic and classical optimization methods. Also, hybrid algorithms which are composed with a derivative-free simple local search algorithm, e.g. Nelder-Mead Simplex and population based artificial intelligence algorithms, Genetic Algorithm and Particle Swarm Optimization, proposed by Türkşen and Tez [43] can be applied to portfolio selection problems.

References

- [1] H. Markowitz, 1952, Portfolio selection, *The Journal of Finance*, 7 (1), 77-91.
- [2] Y. Crama, M. Schyns, 2003, Simulated Annealing for complex portfolio selection problems, *European Journal of Operational Research*, 150 (3), 546-571.
- [3] K. Doerner, W. J. Gutjahr, R. F. Hartl, C. Strauss, C. Stummer, 2004, Pareto Ant Colony Optimization: A metaheuristic approach to multiobjective portfolio selection, *Annals of Operations Research*, 131 (1-4), 79-99.
- [4] M. Ehrgott, X. Gandibleux, 2004, Approximative solution methods for multiobjective combinatorial optimization, *Sociedad de Estadística e Investigaci3n Operativa*, 12 (1), 1-89.
- [5] T. Cura, 2009, Particle swarm optimization approach to portfolio optimization, *Nonlinear Analysis: Real World Applications*, 10 (4), 2396-2406
- [6] H. R. Golmakani, M. Fazel, 2011, Constrained portfolio selection using Particle Swarm Optimization, *Expert Systems with Applications*, 38 (7), 8327-8335.
- [7] H. Zhu, Y. Wang, K. Wang, Y. Chen, 2011, Particle Swarm Optimization for the constrained portfolio optimization problem, *Expert Systems with Applications*, 38 (8), 10161-10169.
- [8] G.-F. Deng, W.-T. Lin, C.-C. Lo, 2012, Markowitz-based portfolio selection with cardinality constraints using improved Particle Swarm Optimization, *Expert Systems with Applications*, 39 (4), 4558-4566.
- [9] K. Lwin, R. Qu, 2013, A hybrid algorithm for constrained portfolio selection problems, *Applied Intelligence*, 39 (2), 251-266.
- [10] A. Z. Çelenli, E. Eğri3glü, B. Ş. Çorba, 2015, İMKB 30 endeksini oluşturan hisse senetleri için Parçacık Sürü Optimizasyonu yöntemlerine dayalı portföy optimizasyonu, *Doğuş Üniversitesi Dergisi*, 16 (1), 25-33.
- [11] H. Akyer, C.B. Kalayci, H. Aygören, 2018, Ortalama-varyans portföy optimizasyonu için Parçacık Sürü Optimizasyonu algoritması: Bir Borsa İstanbul uygulaması, *Pamukkale University Journal of Engineering Sciences*, 24 (1), 124-129.
- [12] J. Doering, R. Kizys, A.A. Juan, À. Fitó, O. Polat, 2019, Metaheuristics for rich portfolio optimisation and risk management: Current state and future trends, *Operations Research Perspectives*, 6, 100121.
- [13] C.B. Kalayci, O. Ertenlice, M.A. Akbay, 2019, A comprehensive review of deterministic models and applications for mean-variance portfolio optimization, *Expert Systems with Applications*, 125, 345-368
- [14] C.B. Kalayci, O. Polat, M.A. Akbay, 2020, An efficient hybrid metaheuristic algorithm for cardinality constrained portfolio optimization, *Swarm and Evolutionary Computation*, 54, 100662, 1-16.
- [15] M. Corazza, G. di Tollo, G. Fasano, R. Pesenti, 2021, A novel hybrid PSO-based metaheuristic for costly portfolio selection problems, *Annals of Operations Research*, 304, 109-137
- [16] E-G. Talbi, 2009, Metaheuristics: From design to implementation, John Wiley & Sons Inc.
- [17] X.-S. Yang, 2010, Engineering optimization an introduction with metaheuristic applications, John Wiley & Sons Inc., New Jersey.
- [18] S. Kirkpatrick, C. D. Gelatt, M. P. Vecchi, 1983, Optimization by Simulated Annealing, *Science*, 220 (4598), 671-680.
- [19] B. Abbasi, A.H.E. Jahromi, J. Arkat, M. Hosseinkouchack, 2006, Estimating the parameters of Weibull distribution using Simulated Annealing Algorithm, *Applied Mathematics and Computation*, 183 (1), 85-93.
- [20] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, E. Teller, 1953 Equation of state calculations by fast computing machines, *The Journal of Chemical Physics*, 21, 1087-1092.
- [21] V. Cerny, 1985, Thermodynamical approach to the traveling salesman problem: an efficient simulation algorithm, *Journal of Optimization Theory and Applications*, 45, 41-51.
- [22] İ. Boussaïd, J. Lepagnot, P. Siarry, 2013, A survey on optimization metaheuristics, *Information Sciences*, 237, 82-117
- [23] D.S. Johnson, C. R. Aragon, L.A. McGeoch, C. Chevon, 1989, Optimization by Simulated Annealing: An experimental evaluation; Part 1, Graph Partitioning, *Operations Research*, 37 (6), 865-892.
- [24] C. C. Ribeiro, S.L. Martins, I. Rosetti, 2007, Metaheuristics for optimization problems in computer communication, *Computer Communications*, 30, 656-669.
- [25] J. Kennedy, R. C. Eberhart, 1995, Particle Swarm Optimization, *IEEE International Conference on Neural Networks*, Perth, Australia, 1942-1948.

- [26] J. Kennedy, R. C. Eberhart, 2001, *Swarm Intelligence*. Morgan Kaufmann, San Francisco, CA,
- [27] Y. Shi, R. Eberhart, 1998, A modified Particle Swarm Optimizer, *Proceedings of IEEE World congress on computational intelligence*. The 1998 I.E. international conference on evolutionary computation, 69–73.
- [28] S.S. Rao, 2009, *Engineering optimization: Theory and practice*, John Wiley&Sons.
- [29] H.H. Örkücü, V.S. Özsoy, E. Aksoy, M. İ. Doğan, 2015b, Estimating the parameters of 3-p Weibull distribution using Particle Swarm Optimization: A comprehensive experimental comparison, *Applied Mathematics and Computation*, 268, 201–226.
- [30] Ş. Acıtaş, Ç.H. Aladağ, B. Şenoğlu, 2019, A new approach for estimating the parameters of Weibull distribution via particle swarm optimization: An application to the strengths of glass fibre data, *Reliability Engineering&System Safety*, 183, 116–127.
- [31] R. Storn, 1996, On the usage of Differential Evolution for function optimization, *Fuzzy Inf. Process. Soc. 1996. NAFIPS., 1996 Bienn. Conf. North Am., IEEE*, 519–523.
- [32] R. Storn, K. Price, 1997, Differential Evolution—a simple and efficient heuristic for global optimization over continuous spaces, *Journal of Global Optimization*, 11, 341–359.
- [33] K. Price, R.M. Storn, J.A. Lampinen, 2006, *Differential Evolution: a practical approach to global optimization*, Springer Science & Business Media.
- [34] L. Gui, X. Xia, F. Yu, H. Wu, R. Wu, B. Wei, Y. Zhang, X. Li, G. He, 2019, A multi-role based Differential Evolution, *Swarm and Evolutionary Computation*, 50, 100508.
- [35] H.H. Örkücü, E. Aksoy, M. İ. Doğan, 2015a, Estimating the parameters of 3-p Weibull distribution through Differential Evolution, *Applied Mathematics and Computation*, 251, 211–224.
- [36] S. Das, S.S. Mullick, P.N. Suganthan, 2016, Recent advances in Differential Evolution—an updated survey, *Swarm and Evolutionary Computation*, 27, 1–30.
- [37] D. Karaboğa, 2005, An idea based on honey bee swarm for numerical optimization, Technical Report-tr06, Erciyes University, Engineering Faculty, Computer Engineering Department.
- [38] B. Akay, D. Karaboğa, , 2012, A modified Artificial Bee Colony Algorithm for real-parameter optimization, *Information Sciences*, 192, 120-142.
- [39] A. Rajasekhar, S. Lynn, S. Das, P.N. Suganthan, 2017, Computing with the collective intelligence of honey bees—a survey, *Swarm and Evolutionary Computation*, 32, 25–48.
- [40] D. Karaboğa, C. Öztürk, 2011, A novel clustering approach: Artificial Bee Colony (ABC) algorithm, *Applied Soft Computing*, 11, 652–657.
- [41] H. Konno, H. Yamazaki, 1991, Mean-absolute deviation portfolio optimization model and its applications to Tokyo Stock Market, *Management Science*, 37 (5), 519-531.
- [42] C. D. Feinstein, M. N. Thapa, 1993, A reformulation of a mean-absolute deviation portfolio optimization model, *Management Science*, 39 (12), 1552-1554.
- [43] Ö. Türkşen, M. Tez, (2016). An Application of Nelder-Mead Heuristic-based Hybrid Algorithms: Estimation of Compartment Model Parameters. *International Journal of Artificial Intelligence*, 14(1), 112-129.”