



Gölgelerin Oyunu: İzdüşümlerin İstatistiksel Çıkarımları ve Türkiye’de Döviz Kurlarını Etkileyen Makroekonomik Göstergeler Üzerine Bir Uygulama

Yunus Güral^{1*}, Arzu Ekinci², Mehmet Gürcan³

^{1*} Fırat Üniversitesi Fen Fakültesi İstatistik Bölümü, Elazığ, Türkiye, (ORCID: 0000-0002-0572-453X), ygural@firat.edu.tr

² Batman Üniversitesi İktisadi ve İdari Bilimler Fakültesi, Batman, Türkiye (ORCID: 0000-0001-9229-8198), arzu.demirelli@batman.edu.tr

³ Fırat Üniversitesi Fen Fakültesi İstatistik Bölümü, Elazığ, Türkiye (ORCID: 0000-0002-3641-8113), mgurcan@firat.edu.tr

(İlk Geliş Tarihi 22 Aralık 2021 ve Kabul Tarihi 21 Ağustos 2022)

(DOI: 10.31590/ejosat.1039913)

ATIF/REFERENCE: Güral, Y., Demirelli E. A. & Gürcan, M. (2022). Gölgelerin Oyunu: İzdüşümlerin İstatistiksel Çıkarımları ve Türkiye’de Döviz Kurlarını Etkileyen Makroekonomik Göstergeler Üzerine Bir Uygulama. *Avrupa Bilim ve Teknoloji Dergisi*, (38), 341-351.

Öz

Pratikte kullanılan istatistiksel yöntemlerin matematiksel arka planları çoğunlukla uygulayanlar için pek önem arz etmeyen detaylar olarak görülebilir. Ancak bilimsel araştırmaların farklı detayları için oldukça önem arz eden tekniklerdir. Örneğin en küçük kareler yöntemi “EKK” üç farklı şekilde detaylandırılabilir. Bir uygulamacı yöntemi paket programın analiz butonu olarak görünürken bir teorisyen EKK yöntemini değişken sayısı gözlem sayısından az olan bir denklem sisteminin çözüm tekniği olarak görmektedir. Bu aynı zamanda görüntü uzayının tanım uzayından büyük olduğu bir durumda görüntü uzayından alınan bir elemanın ters görüntüsünün elde edilmesi tekniğidir. Bu bakımdan bu çalışma araştırmalarda kullanılan temel bileşenler analizi, Karhunen-Loeve dönüşümü, kısmi en küçük kareler regresyonu ve Nipals algoritması gibi tekniklerin teorik arka planlarını vermeyi amaçlamaktadır. Yöntemleri anlaşılır kılmak için çalışmada USD ve EURO kurunun cevap değişkeni olarak seçildiği uygulamaya yer verilmiştir. Açıklayıcı değişkenler ise aralarında iç ilişkili olduğu düşünülen, ihracat, ithalat, borsa endeksi, işsizlik oranı, enflasyon, cari açık ve döviz rezervi değişkenleridir. 2014:01 / 2021:06 dönem aralığındaki aylık veriler alınarak oluşturulan modele çoklu doğrusal regresyon, NIPALS algoritması ve Karhunen-Loeve dönüşümü yapıldıktan sonra NIPALS algoritması uygulanmıştır.

Anahtar Kelimeler: Temel bileşenler analizi “PCA”, kısmi en küçük kareler yöntemi “PLS”, Karhunen-Loeve dönüşümü, Nipals algoritması, Online-Nipals.

Game of Shadows: Statistical Inferences of Projections and an Application on Macroeconomic Indicators Affecting Exchange Rates in Turkey

Abstract

The mathematical backgrounds of statistical methods used in practice can often be seen as insignificant details for those who apply them. However, they are very important techniques for different details of scientific research. For example, the least-squares method can be elaborated in three different ways. While a practitioner sees the method as the analysis button of the package program, a theorist sees the EKK method as a solution technique for a system of equations with a variable number less than the number of observations. This is also the technique of obtaining an inverted image of an element taken from the image space in a case where the image space is larger than the domain. In this respect, this study aims to give the theoretical background of techniques such as principal component

* Sorumlu Yazar: ygural@firat.edu.tr

analysis, Karhunen-Loeve transform, partial least squares regression, and Nipals algorithm used in research. In order to make methods understandable, the application in which USD and Euro rates are chosen as the answer variable is included in the study. Explanatory variables, on the other hand, are export, import, stock market index, unemployment rate, inflation, current account deficit, and foreign Exchange reserve variables, which are thought to be interrelated. The NIPALS algorithm was applied after multiple linear regression, NIPALS algorithm, and Karhunen-Loeve transform were applied to the model, which was created by taking monthly data in the period of 2014:01/ 2021:06.

Keywords: Principal component analysis “PCA”, partial least squares method “PLS”, Karhunen-Loeve transform, Nipals algorithm, Online-Nipals.

1. Giriş

Günümüzde çok değişkenli veri analizinde birçok farklı yöntem kullanılmaktadır. Bu yöntemler içerisinde ilklere giren yöntemlerin başında temel bileşenler analizi gelmektedir. Temel bileşenler analizi köken itibarıyla Cauchy, Pearson ve Hotelling'e kadar dayanmaktadır. Bu çalışmadaki amacımız kısaca bu yöntemi tanıtarak buradan hareketle kısmi en küçük kareler yöntemine kadar dayandırmaktır. Kısmi en küçük kareler yöntemi, bir veya daha fazla bağımlı değişkeni iki veya daha fazla bağımsız değişkenle ilişkilendirmek için geliştirilen bir algoritmadır. Bir regresyon yöntemi olarak, bir matrisin özvektörlerini ve özdeğerlerini belirlemeye yönelik güç yöntemine benzer olan NIPALS algoritmasını kullanan temel bileşenler regresyonu yönteminden uyarlanmıştır (Avraham Lorber ve ark. 1987). Kısmi en küçük kareler yönteminde uygulanan dokuz algoritma mevcuttur (Andersson, M. 2009). Bu algoritmaları kısaca sıralayalım; 1. Wold tarafından geliştirilen NIPALS (Wold, H. 1966), 2. Martens'in ortogonalize edilmiş puanlar algoritması (Martens H, Naes T. 1989), 3. Golub ve Kahan'ın Bidiag2'si (Golub, G. & Kahan, W. 1965), 4. Joung'un SIMPLS'si (De Jong, S. 1993), 5. Dayal tarafından geliştirilen çekirdek PLS (Dayal, B. S. & MacGregor, J. F. 1997), 6. Manne'nin PLSF'si (Wu, W. & Manne, R. 2000), bu algoritmada kendi içerisinde sonradan yapılan eklemelerle üç yeni algoritma daha üretmiştir, 7. Direkt olarak açıklayıcı ve cevap değişkenlerinden puanlar veren temel vektörlerin hesaplanması için yeni bir yinelenen formüle dayalı doğrudan puanlar PLS1 algoritması, 8. Sadece orijinal açıklayıcı ve cevap değişkenleri kullanılarak tanımlanmış Krylov PLS1, 9. Açıklayıcı değişkenden yinelemeli olarak tanımlanan regresyon vektörü ve önceki yinelemeli regresyon vektörleri ile PLSPLS1 algoritması. Bu algoritmalarından 1, 2, 5 ve 7'inci sırada yazılanlar en kararlı algoritmalar. Aynı zamanda 5 ve 7'inci sırada olanlar 1 ve 2'inci sırada olan algoritmalarından çok daha hızlıdır. En hızlı sayılabilecek algoritma olan 4 ve orta hızlı olan algoritma 9 ise faktör sayısı yükseldikçe kararsız hale gelme eğilimi göstermektedir. 3, 6 ve 8'inci sırada bulunan algoritmalar bazı uygulamalarda PLS1 regresyon vektörüne uymayan regresyon vektörleri oluşturmuş bu bakımdan yüklem ve puan vektörlerinin de yinelemeli denklemlerde saptığı sonucuna varılmıştır. Elde edilen artık toplamlarının ise en küçük kareler anlamında optimalin altında olduğu gösterilmiştir. Bu dokuz algoritma içerisinde, NIPALS gibi, veri matrislerinin indirgenmesine dayalı algoritmaların diğerlerinden sayısal anlamda daha kararlı olduğu da bilinmektedir (Ergon, R. 2009). Hesaplamalarda sayısal anlamda kararlılık ve tutarlılık önemlidir. Tüm PLS1 algoritmaları aslında aynı sayısal regresyon sonuçlarını vermek zorunda olmalarına rağmen bu durum uygulamalarda mümkün olmamaktadır. Bu durumda aynı veriler farklı algoritmalarla modellendiğinde farklı modeller elde edilecektir (Andersson, M. 2009).

Veri incelemede kullanılan istatistiksel teknikler araştırmacılara veri ile ilgili ilk bakışta göremedikleri birçok önemli bilgiyi ve detayı sunmaktadır. Bir araştırmada toplanan veriler görselleştirilmek istendiğinde ilk adımda verinin dağılımı, tablo ve grafiklerin kullanımı tercih edilmektedir. Ancak incelenmek istenen veri yüksek boyutlu olduğunda bu görsellik araştırmacılar için pekte önemli bilgiler içermez. Bundan dolayı verinin birtakım istatistiksel işlemlerden geçirilmesi gereklidir. Yüksek boyutlu verilerde ilk adımda tercih edilen yöntem temel bileşenler analizidir. Temel bileşenler analizi (PCA) kullanım kolaylığı bakımından araştırmacılar tarafından sıklıkla tercih edilen bir veri inceleme tekniğidir. Kullanım kolaylığı düzenli bir algoritmasının olmasından kaynaklanmaktadır. Araştırmacının bu analizi yapabilmek için yoğun bir teorik bilgiye ihtiyacı yoktur. Kullanım avantajlarının çok olmasının yanı sıra bazı dezavantajları da bulunmaktadır. Ancak bunlar araştırmacılar tarafından göz ardı edilebilecek konumdadır. Bu dezavantajlar kısaca iki kategoride toplanabilir: Birincisi kovaryans matrisi tekil değer ayrışımı (SVD) ile parçalandığından matrisin değerlendirilmesi zordur. İkincisi ise verinin bazı önemli görselleri PCA tarafından açıkça ortaya konulamamasıdır. Sağladığı avantajlar ise kısaca şu şekildedir: ortogonal bileşenler sayesinde veri düşük boyuta indirgenir, kovaryans karmaşıklığı azaltılır, tüm verileri sınıflara ayırabilir, büyük hesaplamalar gerektirmez, en geniş varyanslara sahip yönleri inceler (Karamizadeh vd., 2013:173-175). PCA analizi ilk olarak SVD ayrışımı ile başlar. Bu ayrışımın elde edilen öz vektörler ortogonal olan temel bileşenleri belirler. Bu sayede değişkenlerin doğrusal kombinasyonları birbirleri ile ilişkisiz olarak elde edilir.

PCA incelemesinde elde edilen temel bileşen sayısı araştırmacıya ait değildir. Ancak araştırmacı elde edilen temel bileşenlerden açıklama yüzdesi fazla olanları seçme imkanına sahiptir. Bu imkan yüksek boyutlu verilere temel bileşenler regresyonu uygulamada oldukça faydalıdır. PCA gibi SVD ayrışımına dayanan ikinci bir analiz yöntemi de kısmi en küçük kareler (PLS) yöntemidir. PLS yöntemini PCA yönteminden ayıran en karakteristik özellik: SVD ayrışımının PCA yönteminde $X^t X$ varyans kovaryans matrisine uygulanması, PLS yönteminde ise $X^t Y$ kovaryans matrisine uygulanmasıdır. Burada X bağımsız değişkenlerin, Y ise bağımlı değişkenlerin standartlaştırılmış veri matrisleridir. PLS, PCA yöntemi gibi uygulaması tanımlanan birtakım algoritmalar sayesinde kolaylaştırılmış önemli bir yöntemdir. Bu özelliğinden dolayı araştırmacılar tarafından tercih edilen ve literatürde geniş uygulama sahası bulmuş bir yöntemdir. PLS yöntemi için geliştirilen bazı algoritmalar temelde lineer olmayan kısmi en küçük kareler (Nipals) algoritmasına dayanmaktadır. Nipals algoritması için bakılması gereken önemli kaynaklar (Wold, 1966:391-420) ve (Oin vd., 2014:1073-1077)'de verilen kaynaklardır. Literatürde bulunan birçok kaynaktan Nipals algoritmasının adımları bulunmaktadır. Ancak bu adımların izahları yeterince açık değildir. Sonraki bölümde Nipals algoritması detaylı olarak anlatılacaktır. Vurgulanması gereken önemli bir nokta da PCA ve PLS yöntemlerinin yüksek boyutlu verilerin yanı sıra iç ilişki probleminde de Ridge regresyonuna

alternatif olarak kullanılabilmesidir (Rosipal ve Krämer, 2006:34-51; Wold vd., 1984). Regresyon incelemesinde bağımsız değişkenler arasında ilişki bulunduğu bu durum veri matrisinin herhangi bir sütununun ilişkili olduğu diğer bir sütun cinsinden yazılabilmesini sağlamakta ve X^tX matrisinin tersi alınamamaktadır. Böyle bir durumda regresyon katsayılar vektörü regresyonun normal denklemlerinden direkt olarak elde edilemez. Bu problemi ortadan kaldırmak için varyans kovaryans matrisinin asal köşegen elemanlarına eklemeler yapılarak singülerlik ortadan kaldırılmaya çalışılır. Yapılan eklemeler değişkenlerin bulunduğu eksen değerlerini artırdığından değişkenlerin arasındaki açılırları artırmaktadır. Bu sayede kovaryans değerleri azaltılarak bir çözüme gidilebilir. Ancak artan varyanslar direkt olarak tahminlere yansıtacağından bu durum tahminlerin hatasını artırmaktadır. Bunun yerine temel bileşenler veya kısmi en küçük kareler regresyonlarının kullanılması latent değişkenlerin eksenlerini ortagonalleştirdiği için daha uygundur. Literatürde yapılan birçok çalışmada bu durum vurgulanmıştır. Denklem sayısı fazla olan lineer denklem sistemlerinin çözümü en küçük kareler yöntemiyle yapılabilir. Bu yöntemde katsayılar matrisinin genelleştirilmiş tersi $(X^tX)^{-1}X^t$ matrisidir. $(X^tX)^{-1}$ ters matrisi hesaplanmadığında, matrisin sütunları lineer bağımlı olduğunda veya değişken sayısı denklem sayısından büyük olduğunda, lineer denklem sistemi, regresyon modelinin normal denklemleri, birinci mertebeden daha düşük bir türev yardımıyla oluşturulduğunda elde edilen çözüm Ridge çözümüdür. Bu durum detaylı olarak kaynak (Marquardt ve Snee, 1975:3-20)'da yer almaktadır.

Çalışmamızda USD ve EURO kuru ile iç ilişkili olduğu düşünülen makro-ekonomik değişkenler seçilerek kurulan modele, çoklu doğrusal regresyon, NIPALS algoritması ve Karhunen-Loeve dönüşümü yapıldıktan sonra NIPALS algoritması uygulanmıştır. Teorik olarak, döviz kuru ile ithalat, ihracat, Bist 100 endeksi, işsizlik, enflasyon, cari açık ve Merkez bankası döviz rezervleri arasındaki ilişkiler şu şekilde açıklanabilir. Döviz kuru, bir ülkenin ulusal para biriminin değerini yabancı bir ülkenin para birimleri cinsinden ifade etmektedir. Ülkelerin ekonomilerine bağlı olarak döviz kurlarında bazı dönemlerde azalış gözlenirken, bazı dönemlerde ise ani yükselişlerle karşılaşmaktadır. Bu durum piyasalarda belirsizlik ve tedirginliğe neden olurken, döviz ile iş yapan kesimleri olumsuz etkilemektedir. Döviz kurlarındaki ani yükselişler sonucunda üretimde kullanılan ithal malların fiyatları artmakta ve buna bağlı olarak üretilen ticari malların maliyetleri yükseltmektedir. Bunun sonucunda maliyetleri yükselen ürünlerin ihracatının yapılabilmesi konusunda üreticiler sıkıntı yaşamakta uluslararası pazarlardaki rekabet güçleri azalmaktadır ve bu şekilde dövizle iş yapan ticari kesimler döviz kurundaki değişimlerden olumsuz etkilenmektedir. Döviz kurlarındaki bu ani değişimler, ekonomi üzerinde yarattığı tahribat ve ekonomik göstergeleri olumsuz etkilemektedir (Kartal vd., 2018: 211).

Döviz kuru, sadece ithalat ve ihracatla ilişkili kalmamakta, aynı zamanda emek piyasası üzerinde de çeşitli etkilere sebep olmaktadır. Döviz kurundaki dalgalanmaların, kısa vadede ekonomik kalkınmayı ve dolayısıyla emek piyasasını olumsuz etkilediği varsayılmaktadır. Döviz kurlarındaki dalgalanmalar belirsizlik ortamında maliyetleri artırır, özellikle mikro düzeyde işletmelerin uzun vadede plan yapmalarını zorlaştırır. Döviz kurundaki belirsizliğin ortadan kalkmasının etkisi ekonomiye olumlu biçimde yansıtılarak ihracat, üretim ve istihdamı olumlu yönde etkiler. İhracat artışının yeni yatırımlarla desteklenmesi durumunda, istihdamı artırıcı etkinin daha fazla olacağı da açıktır.

Sonuç olarak, istikrarlı seyreden döviz kuruyla, dış ticaret işsizlik arasında doğrusal bir ilişkinin olduğundan bahsedilebilir (Buscher ve Mueller, 1997). Döviz kurunun emek piyasası üzerine olan etkisi, çoğunlukla ihracat ve ithalat yoluyla olmaktadır. İthalatın artması bununla birlikte ithal mallarla yerli mallar arasında rekabet oluşması, işverenlerle işçiler arasındaki ilişkiyi büyük ölçüde etkilemektedir. Reel döviz kurundaki yükselmenin, özellikle ithalatın artmasına yol açarak işsizliğe sebep olduğu söylenebilir (Bilgin, 2004).

Döviz kuru ile BİST 100 endeksinde işlem gören ulusal hisse senetlerinin fiyatları arasında ekonomik ilişkide ise Branson (1983)'un portföy dengesi yaklaşımına göre BİST 100 endeksinde meydana gelen bir artışın döviz kurlarında düşüşe sebep olacağı düşünülmektedir. Hisse senetlerinde gözlenen artış, Türk Lirasına olan talebin artmasına ve böylece TL'nin değerlenmesine yol açmaktadır, bu sebeple yabancı hisse senetlerini elden çıkaran yatırımcılar ulusal hisse senetlerine yönelerek, döviz kurlarında düşüşe sebep olmaktadır (Yurdakul, 2016: 24).

Ülke ekonomisinin en önemli sorunlarından biri olan enflasyon; fiyatlar genel seviyesindeki artışı ifade etmekle birlikte, yaşandığı ülkenin para biriminin diğer ülke para birimleri karşısındaki değerinin azalması olarak ta açıklanır (Gedik, 2020:324). Enflasyonun döviz kuru ile ilişkisi ise, literatürde kısır döngü hipotezi ile açıklanmaya çalışılmıştır. Bu hipoteze göre; döviz kurunda gözlenen artış ile üretim için temin edilen ithal hammaddelerin ve ara malların fiyatları artacaktır. Bunun sonucunda ise maliyeti artan ürünlerin üretimi azalacak, fiyatlar genel seviyesinde artışa neden olacaktır. Bu döngüde piyasayı canlandırmak için para arzının artırılması, enflasyon ve döviz kurlarının yükselmesine sebep olacaktır (Terzi ve Kurt, 2007: 3). Cari açık bir ülkenin ürettiğinden daha fazla harcama yaparak tüketmesi olarak tanımlanmaktadır, yabancı paralara talebi artırdığından, döviz kuru üzerinde yükselişe sebep olacak baskı oluşturmaktadır, bu sebeple döviz kuru ile ilişkisinin pozitif yönde olması beklenmektedir. Merkez bankası döviz rezervleri ile döviz kuru arasındaki ekonomik ilişkide ise Merkez bankası piyasadan döviz toplayarak döviz rezervlerini artırma yoluna gittiğinde piyasada döviz hacmi azalacaktır. Böyle bir müdahale sonrası kurların yükselişe geçmesi beklenmektedir. Bunun tersi bir müdahalede Merkez bankası piyasaya döviz sürdüğünde, rezervlerdeki döviz miktarı azalacak ve piyasadaki döviz miktarı artacaktır. Bunun sonucunda döviz kurlarında düşme beklenmektedir. (Şit ve Karadağ, 2019:153).

2. Materyal ve Metot

2.1. Matematiksel Arka Plan: Kısa Matris Bilgisi

Temel bileşenler analizi, kısmi en küçük kareler yöntemi ve NIPALS algoritmasının anlaşılabilmesi için matris bilgilerinin kısa bir yorumu daima gereklidir. Öncelikle matris denildiğinde akla ilk gelen kullanım yerinin lineer dönüşümler olması gerekir. Bu sayede lineer regresyon modelini daha rahat anlamlandırabiliriz. Lineer bir regresyon modeli cevap değişkeninin katsayılar vektörünün lineer bir dönüşümüne eşlenmesiyle oluşturulur. Bu sayede regresyon katsayıları veri matrisinin tersinin soldan denklemle çarpılması sonucunda elde edilir. Bu işlem beraberinde birçok soruları getirmektedir. Örneğin hangi şartlarda veri matrisinin tersi mevcuttur. İlk akla gelen veri matrisinin tersinin alınabilmesi için karesel olması zorunluluğudur. Fakat lineer bir modelde veri matrisi hiçbir zaman karesel olmaz. Bu durum matrislerin genelleştirilmiş tersleri kavramını gündeme getirir ve literatürde oldukça yoğun

bir şekilde incelenmiştir. Bu konuda literatürün ilk kaynağı Penrose, R. (1955, July)'a aittir. Kısaca bu çalışmanın bizi ilgilendiren kısmı lineer bir model denkleminin her iki tarafını veri matrisinin transpozunu ile çarpılması ve katsayılar vektörüyle çarpılan veri matrisinin bu sayede karesel hale getirilmesidir. Ancak bazı veri matrisleri soldan transpozları ile çarpılarak karesel hale getirilselerde çarpım sonucu oluşan varyans kovaryans matrisinin singülerlikten dolayı tersi olmayabilir. Bu durum iç ilişki problemini ortaya çıkartır. Bundan dolayı veri matrisini oluşturan sütunların, değişkenlerin, birbirleri ile ilişkisiz olması gerekmektedir. Eğer herhangi bir ilişki mevcutsa veri matrisinin spektral ayrışımı sonucunda temel bileşenler dediğimiz yeni değişkenler ilişkisiz hale getirilir. Genelleştirilmiş matris tersleri sıklıkla ters dönüşüm problemlerinde kullanıldığından oldukça geniş bir uygulama alanına sahiptir. Bu yönüyle literatürde bu konuda değişik algoritmalar mevcuttur (Courrieu, P. 2008). Bu çalışmalarda farklı matris tersleri tanıtılarak hesaplamaların gerek hızları gerekse tahmin güçleri birbirleri ile kıyaslanmaktadır. Özellikle Drazin matris tersi hem uygulama açısından hem de elde edilen sonuçların duyarlılığı açısından genelleştirilmiş matris tersleri arasında önemli bir yere sahiptir (Koliha, J. J. 1996). Bu bölümde genelleştirilmiş ters matrislerin lineer regresyon modellerindeki uygulamasına ait kısa bir örnek sunulacaktır.

Matrislerle lineer dönüşümler özdeştir. Her matris lineer bir dönüşümdür. Bunun yanı sıra her lineer dönüşüme bir matris karşılık gelmektedir. A , $n \times p$ tipli bir matris ise tanım kümesi p ve görüntü kümesi n boyutludur. A matrisinin görüntüsünü $\mathcal{R}(A)$ ile gösterelim. Matrisin tanım kümesinin boyutu görüntü kümesinin boyutundan küçükse $\mathcal{R}(A) \subset R^p$ olacağından R^p uzayında tanım kümesiyle eşlenmeyen elemanlar bulunacaktır. Bu durum en küçük kareler çözümünün temelini oluşturur. Şöyle ki $Ab = y$ denklem sisteminde y cevap değişkeni $\mathcal{R}(A)$ içerisinde değilse $\mathcal{R}(A)$ içerisinde cevap değişkenine en yakın olan eleman bulunarak ters görüntü alınabilir. Bu sayede b çözüm vektörü tahmin edilmiş olur. Yeri gelmişken aşağıdaki kısa bir örnekte bu işlemi gösterelim. Basit doğrusal regresyon denklemi $Y = b_0 + b_1X$ şeklinde seçilen bir modelde bağımsız ve bağımlı değişken değerleri yazıldığında $Ab = y$ lineer denklem sistemi;

$$\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \end{bmatrix} = \begin{bmatrix} 8 \\ 6 \\ 14 \end{bmatrix} \quad (1)$$

şeklinde oluşturulur.

Burada dönüşüm matrisinin görüntüsü;

$$\mathcal{R}(A) = \{(b_0 + b_1, b_0 + 2b_1, b_0 + 3b_1) : b_0, b_1 \in R\} \quad (2)$$

olarak elde edilir.

Cevap değişkeni $Y \notin \mathcal{R}(A)$ görüntünün elemanı olmadığından çözüm için cevap değişkenine görüntüler içerisinde en yakın eleman seçilir. Bu eleman, $r \in \mathcal{R}(A)$ olmak üzere iki vektör arasındaki kare uzaklığı gösteren ve Eşitlik (3) ile verilen $d^2(Y, r)$ fonksiyonunun minimum değeri olup;

$$d^2(Y, r) = (8 - b_0 - b_1)^2 + (6 - b_0 - 2b_1)^2 + (14 - b_0 - 3b_1)^2 \quad (3)$$

şeklinde ifade edilir.

Yukarıdaki ifadenin b_0 ve b_1 parametrelerine göre türevleri alınarak sıfıra eşitlendiğinde $b_0 = 3,33$ ve $b_1 = 3$ çözümü bulunur.

Matrisin tanım kümesinin boyutu görüntü kümesinin boyutundan büyükse dönüşümün görüntüsü $\mathcal{R}(A) = R^p$ görüntü kümesini tamamen kaplar. Ancak tanım kümesinin boyutu büyük olduğundan birebir eşlemeyle görüntü kümesine sığamayacağından tanım kümesinin elemanlarından birçoğu görüntü kümesinde bir elemana eşlenir. Bu durum ise ters görüntünün seçilmesini bozacağından, lineer denklem sisteminin parametreleri tek türlü bulunamaz. Literatürde bu problem "ill posed" olarak adlandırılır. Günümüzde halen çözüm arayan problemlerin başında gelmektedir. Bu problem için kullanılan yöntemler genel olarak cevap değişkeninin varyansını en iyi şekilde temsil eden değişkenlerin seçilebilmesiyle yapılmaktadır. Literatürde bu gibi durumlarda sıklıkla PCA ve PLS yöntemleri kullanılmaktadır.

2.2. İç Çarpım Kavramı

İç çarpım kavramı ilk olarak diklik, ikinci olarak izdüşüm için sıklıkla kullanılan önemli bir kavramdır. Son yüz elli yılda gelişen modern matematiğin neredeyse en önemli işlemidir. İlk olarak sabit vektörler üzerinde tanımlanmasına rağmen sonradan fonksiyonel diziler üzerine genişletilmiştir. Bu sayede bir fonksiyonun birbirlerine dik olan sayılabilir çoklukta fonksiyonlar cinsinden ifadesi mümkün olmuştur. Bu özellik bilime önemli katkılar sunmaktadır. Günümüzde daha yeni sayılabilecek fonksiyonel veri analizi bizim için bu katkılarının önemli bir göstergesidir (Ramsay, J. O. & Dalzell, C. J. 1991).

İki vektörün iç çarpımı $\langle x, y \rangle = x^t y$ şeklinde tanımlanıp $x^t y = 0$ olduğunda bu iki vektörün dik olduğu anlaşılır ve bu durum $x \perp y$ şeklinde gösterilir. Şayet $\|y\| = 1$, yani y vektörü birim vektör ise $x^t y$ değeri x vektörünün y vektörü üzerindeki dik izdüşüm uzunluğunu veya kısaca gölgesini verir. x ve y tesadüfi değişkenlerinin değerleri standartlaştırılmış ise iç çarpım değeri $x^t y = \cos \alpha = \rho_{xy}$ korelasyon katsayısına eşit olacaktır. Burada α iki değişken arasındaki açıyı temsil eder. Doğal olarak korelasyon katsayısı iki değişken arasındaki açının kosinüs değeridir. Aynı şekilde X bağımsız değişkenlerin ve Y ise bağımlı değişkenlerin bir matrisi ise $X^t X$ bağımsız değişkenlerin varyans kovaryans matrisi, $X^t Y$ ise bağımsız değişkenler ile bağımlı değişkenlerin kovaryans matrisidir.

İç çarpım $x^t y$ değerinin x vektörünün y üzerindeki izdüşümünü vermesi önemli bir özelliktir. Bu sayede y yerine $\{u_1, \dots, u_p\}$ birim dik vektörler alındığında eşitlik (4) ile verilen toplam, x vektörünün u_j vektörleri cinsinden açılımı

$$\sum_{j=1}^p (x^t u_j) u_j \quad (4)$$

eşitliği ile verilir.

Başta Fourier ve Karhunen-Loeve açılımları olmak üzere bu tipli seri açılımları birçok araştırmada oldukça önemlidir. Ayrıca bağımsız değişkenler birbirlerine dik olduklarında bu açılım regresyon denklemini göstermektedir.

2.3. Öz Vektör Kavramı

Bu bölümde özvektör kavramı sadece belirli özellikleri ile kısaca tanıtılacaktır. Dönüşümler genelde tanım kümesinin her bir önemli özelliğini görüntü kümesine taşımazlar. Taşındığı taktirde dönüşüm altında değişmeyen ve dönüşüme has karakteristik özelliklerdir. Lineer dönüşümler tanım kümesinin bazıını oluşturan birbirine dik vektörlerden oluşmuş bir yapıyı görüntü kümesine taşıırken yapının dikliğini koruyamazlar. Ancak her lineer dönüşüm dik vektörlerden oluşmuş en az bir yapıyı

dikliğini koruyarak görüntü kümesine taşır. İşte bu dikliği bozulmadan taşınan vektörler dönüşümün öz vektörleridir. Bu durumu şu şekilde açıklayalım. A , $n \times p$ tipli bir matris ve $\{u_1, \dots, u_p\}$ tanım kümesinde birbirlerine dik birim vektörlerden oluşmuş bir yapı olsun. $i \neq j$ için $u_i \perp u_j$ iken $\langle u_i, u_j \rangle = 0$ $Au_i \perp Au_j$ olabilmesi için $\langle \lambda_i u_i, \lambda_j u_j \rangle = \lambda_i \lambda_j \langle u_i, u_j \rangle = 0$ eşitliği dikkate alındığında $Au_i = \lambda_i u_i$ ve $Au_j = \lambda_j u_j$ seçilmesi yeterlidir. Gerçekten $(A - \lambda_i I)u_i = 0$ denkleminin çözümü $Au_i = \lambda_i u_i$ şeklinde elde edileceğinden bu durum böyle bir seçim yapma şansımızın olduğunu garanti eder.

u_i sıfırdan farklı olduğunda $(A - \lambda_i I)u_i = 0$ denklem sisteminin çözümünün olabilmesi için $\det(A - \lambda_i I) = 0$ olmalıdır. Bu ise λ_i sabitinin öz değer, u_i vektörünün de öz vektör olduğunu göstermektedir.

Burada kısaca şuna da değinelim; $\langle Au_i, Au_j \rangle = (Au_i)^t (Au_j) = u_i^t A^t A u_j$ ifadesi literatürde oldukça önemlidir. $u_i = u_j = w$ alındığında aşağıdaki eşitlik,

$$R_A(w) = \frac{w^t A w}{w^t w} \quad (5)$$

Rayleigh bölümü olarak adlandırılır (Stott ve ark, 2017).

Rayleigh bölümü online PLS (OL-PLS) yönteminin önemli bir fonksiyonudur. OL-PLS yöntemi "Gradient-Descent Total Least Squares" yönteminden adapte edilerek oluşturulmuş bir algoritmadır. Nipals algoritmasından sonra bu yöntem de kısaca değinilecektir.

2.3. Tekil Değer ve Tekil Değer Ayırışımı: SVD

Bu bölümde sadece bir matrisin SVD ayrışımının kısaca nasıl gerçekleştirildiğinden bahsedilecektir. Kare ve tam ranklı matrislerin tanım ve görüntü kümelerinin boyutları eşit olduğundan sorunsuz lineer dönüşümlerdir. Boyutu $(n \times n)$ olan bir kare matrisin rankı n sayısından küçük olduğunda matrisin görünüşte kare olmasının hiçbir önemi yoktur. Örneğin bu şekildeki bir matrisin rankı $n - 1$ olduğunda matrisin sütun vektörlerinden bir tanesi (veya satır vektörlerinden bir tanesi) diğerleri ile lineer bağımlı olacağından görüntü kümesinde faklı bir boyutu belirtmeyecektir. Dolayısıyla görüntü kümesinin boyutu tanım kümesinin boyutundan bir eksik olur. Bu ise kötü görüntü problemini ortaya çıkarır. Kare, tam ranklı ve $(n \times n)$ tipindeki bir kare matrisin öz vektörleri $\{u_1, \dots, u_n\}$ ise lineer dönüşümün tanım kümesi bu sistem tarafından rahatlıkla tanımlanabilir. Aynı şekilde aşağıdaki eşitliği dikkate aldığımızda,

$$\mathcal{R}(A) = Sp\{Au_1, \dots, Au_n\} = Sp\{\lambda_1 u_1, \dots, \lambda_n u_n\} \quad (6)$$

görüntü kümesinin de özdeğer ve öz vektörler yardımıyla tanımlanabileceğini görürüz. Bu önemli özellikten dolayı matrisin öz değer ve öz vektör ayrışımı,

$$A = U \Lambda U^t \quad (7)$$

şeklinde yapılabilir. Burada $U = (u_1, \dots, u_n)$ öz vektörlerin matrisi, $\Lambda = \text{diagonal}(\lambda_1, \dots, \lambda_n)$ öz değerlerin köşegen matrisidir. Ancak kare olmayan veya tam ranklı olmayan matrisler için $A = U \Lambda U^t$ şeklinde yazılım mümkün değildir. Ancak bu formdaki bir matris parçalanması mümkündür. O zaman bu şekildeki bir matrisin,

$$A = U \Lambda V^t \quad (8)$$

şeklinde parçalandığını varsayalım. Bu parçalanmada A matrisinin rankının s olduğunu kabul edersek $A: (n \times p)$, $U: (n \times$

$s)$, $\Lambda: (s \times s)$ ve $V^t: (s \times p)$ boyutlu matrislerdir. $AA^t = U \Lambda^2 U^t$ olduğundan U , AA^t matrisinin öz vektörlerinin matrisi, $A^t A = V \Lambda^2 V^t$ olduğundan V , $A^t A$ matrisinin öz vektörlerinin matrisidir. Her iki durumda da öz vektörlere karşılık gelen öz değerler aynıdır. $A^t A$ veya AA^t matrislerinin öz değerleri Λ matrisinin köşegen elemanlarıdır. A matrisinin singüler değerleri ise bu öz değerlerin kare kökleridir.

Burada karşımıza temelde iki durum çıkabilir. Birincisi $s = p$ olabilir. Bu durumda $A^t A$ matrisi $(p \times p)$ boyutlu ve tam ranklıdır. Eğer iç ilişki problemi yoksa öz değerlerin hepsi sıfırdan büyük ve birbirinden farklıdır. Problemin çözümü en küçük kareler yöntemine indirgenir. Öz değerlerin sıfıra yakın olması iç ilişki probleminin en önemli göstergesidir. İkincisi $s < p$ ise $A^t A$ matrisi de tam ranklı olmaz. Bu durumda iç ilişki problemi kaçınılmazdır. Böyle bir durumda yapılacak ilk iş var olan öz vektörlerle, $\{u_1, \dots, u_s\}$, görüntü kümesini tanımlayarak çözüme gitmektir. Görüntü kümesinin baz elemanları bağımsız değişkenlerin lineer birleşimleri olan temel bileşenlerdir.

2.5. Temel Bileşenler Analizi: PCA

Temel bileşenler analizi çok değişkenli istatistiksel teknikler arasında neredeyse en popüler olanıdır. Bu konuda literatürde dayanan en eski kaynaklar Cauchy, A. L. (1829), Pearson, K. (1901) ve Hotelling, H. (1933)'e aittir (Abdi, H., & Williams, L. J. 2010). Bu bölümde kısa bir örnek üzerinde işlemin nasıl gerçekleştirildiği gösterilecektir.

Temel bileşenler kısaca en temel manasıyla bağımsız değişkenlerin öz vektörlerin katsayıları yardımıyla lineer birleşimler şeklinde yazılabilmesidir. Bağımsız değişkenlerin matrisine $X = U \Lambda V^t$ şeklinde tekil değer ayrışımı yapılsın. Bu durumda V , $X^t X$ çarpımının öz vektörleri, XV ise temel bileşenlerdir. Şimdi bağımsız değişkenler arasında iç ilişkinin bulunduğu bir örneği göz önüne alalım. Bu örnekte regresyon denklemi

$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 3 & 6 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \\ 8 \end{bmatrix} \quad (9)$$

eşitliği ile verilsin. Lineer dönüşümün görüntü kümesi $\mathcal{R}(A) = \{(t, 2t, 3t) : t = b_0 + 2b_1, b_0, b_1 \in R\}$ tek boyutludur. Cevap değişkeni $(2, 5, 8) \notin \mathcal{R}(A)$ görüntü kümesinin elemanı değildir. Böyle bir durumda b_0 ve b_1 tek türlü elde edilemez. En küçük kareler denkleminin çözümünden $t = 18/7$ ve $b_0 + 2b_1 = 18/7$ olduğundan parametreler tek türlü elde edilemeyecektir. Eşitliği sağlayacak şekilde $b_0 = 2/7$ ve $b_1 = 8/7$ alınarak tahmin denklemi $\hat{y} = (2/7)x_1 + (8/7)x_2$ şeklinde oluşturulabilir. Veri matrisine SVD ayrışımı yaparsak bağımsız değişkenlerin lineer birleşimini yeni bağımsız değişken olarak kullanabiliriz. $X^t X$ matrisinin tek öz değeri $\lambda = 70$ ve öz vektörü $v^t = (1/\sqrt{5}, 2/\sqrt{5})$ olarak elde edilir. Bu durumda var olan tek temel bileşen $Xv = (1/\sqrt{5})X_1 + (2/\sqrt{5})X_2$ olarak bulunur. Sonuç olarak veri kümesi aşağıdaki şekilde değiştirilmiş olur,

Tablo 1. Temel bileşenden elde edilen değerler (Table 1. Values from the principal component)

X_1	X_2	$Z = Xv$	Y
1	2	$5/\sqrt{5}$	2
2	4	$10/\sqrt{5}$	5
3	6	$15/\sqrt{5}$	8

Yeni regresyon denklemi $\hat{y} = -1 + (3/\sqrt{5})z$ şeklinde tahmin edilir.

2.6. Karhunen-Loeve Dönüşümü: KL

Bu dönüşüm temelinde Fourier serisi gibi bir stokastik sürecin seriyeye açılması ilkesine dayanmaktadır (Karhunen, K. 1947 ve Loeve, M. 1948). Burada sadece veriden elde edilen kovaryans matrisi için bu işlemin nasıl yapıldığı ile ilgileniyoruz. Bu dönüşüm sayesinde kovaryans matrisi birbirine dik sütunlardan oluşur.

Bir matrisin sabit bir sayıyla çarpılması öz değer ve öz vektörlerini değiştirmez. Bu özellikten dolayı $X^t X$ matrisini bazı durumlarda $CovX$ varyans kovaryans matrisi olarak da kullanabiliriz. Bu farklı bir durum gibi algılanmamalıdır. Kovaryans matrisi değişkenler arasında varyans ve kovaryanslar hesaplanarak oluşturulacağı gibi tek tek gözlemler için ilgili çarpımlar hesaplanıp toplandığı taktirde de bulunabilir. KL dönüşümü kovaryans matrisinin öz vektörleri yardımıyla oluşturulur. Kovaryans matrisinin öz vektörleri $U = \{u_1, \dots, u_s\}$ olmak üzere dönüşüm $\mathcal{K} = U^t(X - EX)$ şeklinde tanımlanır. Burada EX , X tesadüfi değişkeninin beklenen değerini göstermektedir. Aynı zamanda EX , X tesadüfi değişkeninin gözlem değerlerinin aritmetik ortalaması olarak hesaplanmaktadır. Dönüşüm uygulama bakımından PCA yönteminden daha kolaydır. Kovaryans matrisi;

$$nCovX = \ell_1 \ell_1^t + \dots + \ell_n \ell_n^t \quad (10)$$

eşitliğinden bulunabilir. Burada $\ell_j^t = (x_{j1}, \dots, x_{jp})$ j -inci gözlem vektörüdür. Ayrıca \mathcal{K} matrisinin satır elemanları birbirlerine diktir. $\mathcal{K}^t = (k_1, \dots, k_p)$ olmak üzere,

$$k_i = (u_i \ell_1^t, \dots, u_i \ell_n^t) \Rightarrow u_i \ell_1^t u_j \ell_1^t = \ell_1 u_i^t u_j \ell_1^t = 0 \Rightarrow k_i k_j^t = 0 \quad (11)$$

gerektirmeleri sağlar.

Sonuç olarak \mathcal{K} matrisinin satır elemanları x_1, \dots, x_p değişkenlerinin birbirlerine dik (ilişkisiz) olacak şekildeki lineer birleşimlerinden oluşur.

2.7. Kısmi En Küçük Kareler Yöntemi: PLS ve Nipals

Kısmi en küçük kareler yöntemi çok değişkenli veri kümelerinin modellenmesinde oldukça etkili bir yöntemdir. PLS modellemesi ve kullanılan algoritmalar Martens H, Naes T. (1989) tarafından kaleme alınan kitapta ayrıntılı olarak bulunur. PLS için ortogonal ayırıştırma, kısıtlama koşullarına tabi olan bir ekstremum problemidir. Bu nedenle Lagrange çarpanları yöntemi kullanılarak incelenebilir. Zhu, E. ve Barnes, R. M. (1995) bu yöntemle 'basit yineleme algoritması' denilen etkili bir yineleme algoritması elde etmişlerdir. Burada ana fikir, tahmin için regresyon modelinin uygun şekilde hesaplanmasıdır. Bunun için lineer regresyon modelinde veri matrisi parçalanarak katsayılar matrisi ile çarpıldığında iki parçaya ayrılır. İkinci parça gizli değişkenleri içermektedir. Gizli değişkenlerin iterasyonunda yinelemeli denklemler ters matris yardımıyla bulunur. Bu şekildeki basit yineleme algoritmaları, NIPALS algoritmasından daha hızlı, daha fazla bellek tasarrufu sağlar ve programlanması daha kolaydır. PLS regresyonu için basit yineleme algoritması ile NIPALS algoritması karşılaştırıldığında, yeni algoritmanın kesinliği klasik algoritmadan daha kötü değildir. Yeni yöntem, klasik algoritmanın amacı olan gizli değişkenler ile Y arasındaki

kovaryansın maksimum olduğu ekstremum problemlerini de çözebilir (Zhu, E. & Barnes, R. M. 1995). Ancak algoritmalar gerçekleştirilirken bilgisayarda yapılan yuvarlama hataları sonucunda birbirine dik çıkması beklenen vektörlerin dik olmadıkları sonucu ile de karşılaşılabilir (Seasholtz, M. B., Pell, R. J. & Gates, K. E. 1990). NIPALS algoritması ile ilgili çokça kaynak literatürde mevcuttur. Bunların herbirisi mevcut algoritmaya yaptıkları küçük dokunuşlarla gerek algoritmanın yakınsama durumunu gerekse tahminlerin güvenilirliğini artıtmayı amaçlamaktadır (Ergon, R. 2009). Burada gerçekleştirdiğimiz kısa çalışmada bunların hepsine ayrıntılı bir şekilde değinmek imkânımız maalesef mümkün değil. Bu bölümde sadece NIPALS algoritmasının Wold, H. (1966) ve Wold, S. ve ark. (1984) tarafından verilen ilk haline değineceğiz.

PLS yönteminde ilk dikkat edilmesi gereken unsur "kısmi" olma özelliğidir. Bu kısaca şu anlama gelir, PCA yönteminde SVD ayrışımı yapılırken tüm öz vektörler öncelikle elde edilmeli ve matris parçalanması gerçekleştirilmelidir. PLS yönteminde ise bu durum adımsal olarak işletilir. İlk önce en büyük öz değere karşılık öz vektör bulunarak işlem başlatılır. Daha sonra bu öz vektör ilgili matristen atılarak eksiltilmiş olan veri matrisinin en büyük öz değeri, veri matrisinin ikinci öz değeri ve bu öz değere karşılık gelen öz vektör bulunarak işlemler bir döngü içerisinde takip edilir. Eksiltilmiş matris kavramına kısaca açıklık getirelim. Determinantı sıfırdan farklı olan $(n \times n)$ tipli A matrisinin öz değerleri sırasıyla $\lambda_1 > \dots > \lambda_n$ ve öz vektörleri sırasıyla v_1, \dots, v_n olmak üzere bu matris,

$$A = \lambda_1 v_1 v_1^t + \dots + \lambda_n v_n v_n^t \quad (12)$$

şeklinde özdeğer ve özvektörler cinsinden yazılabilir. A matrisinin bu yazılışı spektral ayrışım olarak bilinmektedir.

Bu durumda $A - \lambda_1 v_1 v_1^t = \lambda_2 v_2 v_2^t + \dots + \lambda_n v_n v_n^t$ olup $A - \lambda_1 v_1 v_1^t$ eksiltilmiş matrisinin en büyük öz değeri λ_2 olur. Benzer işlem SVD ayrışımı için de geçerlidir. Veri matrisinin SVD ayrışımı $X = UAV^t$ olmak üzere öz değerleri $\lambda_1 > \dots > \lambda_p$ ve $U = [u_1, \dots, u_p]$, $V = [v_1, \dots, v_p]$ ise X matrisi,

$$X = \lambda_1 u_1 v_1^t + \dots + \lambda_p u_p v_p^t \quad (13)$$

şeklinde yazılabilir.

Kısmi en küçük kareler yöntemi genel olarak çok değişkenli regresyon analizine göre düzenlenmiştir. Ancak çoklu regresyon olduğu zamanda uygulanabilir. Regresyonda bağımsız değişkenlerin veri matrisinin sütunları ile bağımlı değişkenlerin veri matrisinin sütunlarının arada olan ilişkiyi belirtebilmesi için aynı uzayı tanımlamaları gerekmektedir. Basit olarak düşündüğümüzde çoklu regresyon denkleminde regresyon tahmin denklemi $Xb = \hat{y}$ şeklinde yazıldığında bağımsız değişken veri matrisinin sütunlarının lineer birleşimleri bağımlı değişkenin tanımladığı bölgeyi göstermektedir. Bu ise bağımlı değişkenin tanımladığı bölgenin aslında bağımsız değişkenlerin veri matrisinin sütun uzayı tarafından tanımlandığını açık bir şekilde gösterir. Sonuç olarak X ve Y matrislerinin SVD ayrışımalarının sırasıyla $X = T\xi^t$, $Y = T\eta^t$ şeklinde yazılması bize bu imkanı vermektedir.

Ancak X ve Y matrislerinin SVD ayrışımalarını ayrı ayrı yaptığımızda $X = U_X \Lambda_X V_X^t$ ve $Y = U_Y \Lambda_Y V_Y^t$ parçalanmalarından elde edilen $U_X \Lambda_X \neq U_Y \Lambda_Y$ olduğundan aranan T matrisi elde edilemez. Bu durumda $X^t X B = X^t Y$ denklem sisteminin sütun uzayları bu özelliği sağladığından $X^t Y$ matrisinin en büyük öz değerine karşılık gelen öz vektörü ile işe başlanmalıdır. Bütün öz

değerler bir anda bulunamaz. Gerçekte PLS yönteminin kısmi yapılmasının nedeni de budur. Buna göre $S = X^t Y$ olmak üzere SS^t matrisinin en büyük öz değeri λ_1 ve bu öz değere karşılık gelen öz vektör w_1 olsun. Bu öz vektörü ilk adımda ξ matrisinin ilk sütunu olarak değerlendiririz. $X\xi = T$ eşitliğinden T matrisinin ilk sütunu $t_1 = Xw_1$ çarpımından elde edilir. Bunun devamında T matrisinin ilk sütunu elimizde olduğundan ξ ve η matrislerinin ilk sütunlarını da,

$$X = T\xi^t \Rightarrow X^t = \xi T^t \Rightarrow X^t T = \xi T^t T \Rightarrow X^t t_1 = \xi_1 t_1^t t_1 \Rightarrow \xi_1 = X^t t_1 / t_1^t t_1 \quad (14)$$

$$Y = T\eta^t \Rightarrow Y^t = \eta T^t \Rightarrow Y^t T = \eta T^t T \Rightarrow Y^t t_1 = \eta_1 t_1^t t_1 \Rightarrow \eta_1 = Y^t t_1 / t_1^t t_1 \quad (15)$$

şeklinde elde edebiliriz.

Böylelikle PLS yönteminin ilk adımı tamamlanmış olur. İkinci adıma geçerken X ve Y matrislerini t_1 , ξ_1 ve η_1 yardımıyla,

$$X = t_1 \xi_1^t + \dots + t_p \xi_p^t \Rightarrow X_1 = X - t_1 \xi_1^t \quad (16)$$

$$Y = t_1 \eta_1^t + \dots + t_p \eta_p^t \Rightarrow Y_1 = Y - t_1 \eta_1^t \quad (17)$$

şeklinde bir kademe eksilterek yeniden ifade edebiliriz.

Burada X_1 ve Y_1 ikinci döngüde kullanacağımız eksiltilmiş yeni matrislerimizdir. İşlem tamamlandığında T , ξ ve η matrislerinin tüm sütunları elde edilmiş olur. Kısaca $XB = Y$ denkleminde X matrisinin $X = T\xi^t$ şeklinde yazılabileceği dikkate alındığında B parametre matrisi;

$$XW = T\xi^t W \Rightarrow \hat{X} = XWW^t = T\xi^t W W^t \Rightarrow B = W(\xi^t W)^{-1} T^t Y \quad (18)$$

olarak elde edilir. Burada W matrisi $X^t Y$ matrisinin öz vektörlerinden oluşurken \hat{X} düşük ranklı yeni veri matrisidir (Stott ve ark, 2017).

2.8. Online Nipals Algoritması: OL-PLS

Kısmi en küçük kareler yöntemi kullanım kolaylığı ve iyi sonuç vermesi bakımından veri analizinde sıkça kullanılan önemli bir analizdir. Yöntemin uygulanmasında takip edilen algoritma, Nipals, SVD yönteminin adimsal uygulamaya dönüştürülmesiyle başarılı bir şekilde tanımlanmıştır. Uzunlamasına gözlenen verilerde araştırmacı bu yöntemi uygulayabilmek için gözlemi durdurarak hesaplama yapmalıdır. Ancak veri gözlem durduktan sonra da sürdürülebiliyorsa ilk yapılan hesaplamalara sonraki gözlemlerin eklenmesi doğrudan mümkün olmamaktadır. Bu aşamada geliştirilen OL-PLS yöntemi bu bakımdan oldukça kullanışlıdır (Stott ve ark,2017). Bu yöntem birinci veriden başlayarak Nipals algoritmasını işletebilmekte ve bu sayede gözlem ilerledikçe her adımda hesaplama yapılabilmektedir. Veri matrisi $X = (x^1, \dots, x^p)$ ve cevap değişkeni $Y = (y^1, \dots, y^r)$ olmak üzere $X^t Y$ matrisi,

$$S = X^t Y = g_{X1}^t g_{Y1} + \dots + g_{Xn}^t g_{Yn} \quad (19)$$

eşitliği ile ifade edilebilir. Burada $g_{Xj} = (x_j^1, \dots, x_j^p)$ ve $g_{Yj} = (y_j^1, \dots, y_j^r)$ j -inci gözlem vektörleridir. Diğer ilgili adımlar Nipals algoritmasında olduğu gibi takip edilir.

3. Pratik Bir Uygulama

Veri üzerinde NIPALS algoritmasının adım adım uygulanması oldukça güç bir süreçtir. Bu bakımdan işlemler için bir paket program kullanmak daha avantajlı olacağından Matlab paket programı ile çalışılmıştır. Bu aşamada paket programın verdiği *e-ISSN: 2148-2683*

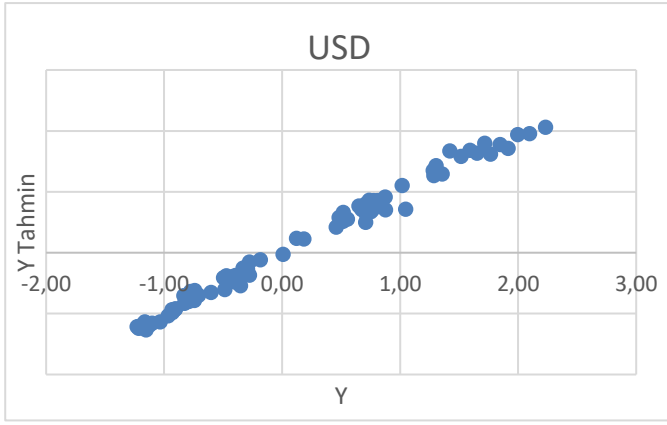
çıktıların yorumlanması ve anlamlandırılması bakımından veri üzerinde kısa bir uygulama yapacağız. Araştırmaya konu olan veriler TÜİK ve Türkiye Merkez Bankası Elektronik Veri Dağıtım Sisteminden elde edilmiştir. Çalışmada USD ve EURO kurları bağımlı değişken alınırken ihracat (X_1), ithalat (X_2), borsa endeksi (X_3), işsizlik oranı (X_4), enflasyon (X_5), cari açık (X_6) ve döviz rezervi (X_7) bağımsız değişken olarak alınmıştır. Bu verilere bağımlı değişken olarak önce USD sonra EURO alınmış ve çoklu lineer regresyon analizi uygulanmıştır.

İlk önce USD bağımlı değişken alınarak regresyon analizi yapıldığında R^2 , %99 olarak bulunurken model istatistiksel olarak anlamlı çıkmıştır. Yapılan analizin sonuçları Tablo 2’de sunulmuştur.

Tablo 2. Cevap değişkeni USD için yapılan çoklu regresyon analizi sonuçları (Table 2. Multiple regression analysis results for the response variable USD)

	Kareler toplamı	Serbestlik derecesi	Kareler ortalaması	F değeri	p değeri
Regresyon	88,227	7	12,604	1337,146	,000
Hata	,773	82	,009		
Toplam	89,000	89			
Modelin uyumu R^2 : 0,991					
Tahminin standart hatası: 0,09709					
Cevap değişkeni: USD					
Faktörler:	Katsayılar	Katsayıların standart hatası	p değeri		
Model sabiti	-2,54e-12	0,010	1,000		
X_1 : ihracat	-0,006	0,012	0,597		
X_2 : ithalat	-0,003	0,011	0,791		
X_3 : borsa endeksi	-0,080	0,025	0,002		
X_4 : işsizlik oranı	-0,072	0,021	0,001		
X_5 : enflasyon	0,821	0,052	0,000		
X_6 : cari açık	0,043	0,012	0,001		
X_7 : döviz rezervi	-0,296	0,034	0,000		

Tablo 2 incelendiğinde model sabiti dahil olmak üzere birinci ve ikinci değişkenlerin modele katkısının anlamsız olduğu görülmektedir. Bu durumda USD değerlerindeki değişimi ihracat ve ithalat etkilememektedir. Modele katkısı en fazla olan üç değişken ise sırasıyla enflasyon, döviz rezervi ve borsa endeksidir. Dikkat edilecek olursa USD ile enflasyon arasında beklendiği gibi pozitif bir ilişki, döviz rezervi ve borsa endeksi ile de zıt yönlü bir ilişki modelde yakalanmıştır. Cevap değişkeni USD için çoklu doğrusal regresyon modelinden elde edilen tahmin değerlerinin USD değerlerini ne ölçüde doğru tahmin ettiğinin görülebilmesi bakımından USD değerleri ile tahmin değerlerinin grafiği Şekil 1’de verilmektedir. Modelin uyumu grafiğinin birinci açıortay doğrusu üzerinde çıkmasına bağlıdır.



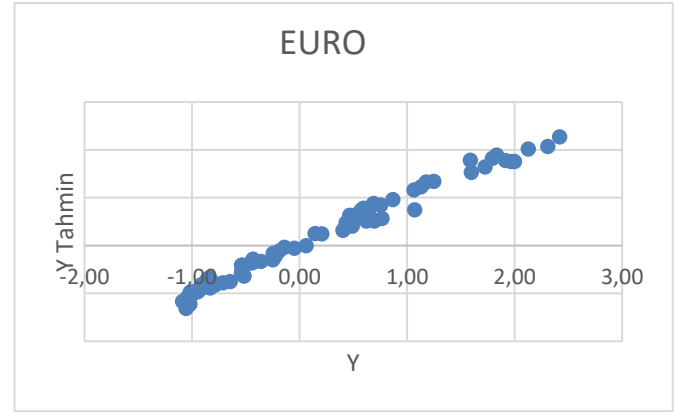
Şekil 1. Cevap değişkeni USD değerleri ile tahmin değerlerinin grafiği (Figure 1. Graph of response variable USD values and predicted values)

İkinci olarak cevap değişkeni EURO için çoklu regresyon analizi yapıldığında R^2 %98 olarak bulunurken model istatistiksel olarak anlamlı çıkmıştır. Yapılan analizin sonuçları Tablo 3'de özetlenmiştir.

Tablo 3. Cevap değişkeni EURO için yapılan çoklu regresyon analizi sonuçları (Table 3. Multiple regression analysis results for the response variable EURO)

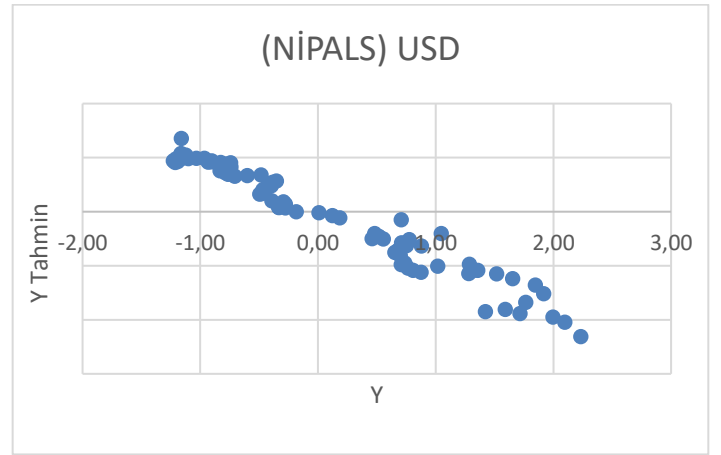
	Kareler toplamı	Serbestlik derecesi	Kareler ortalaması	F değeri	p değeri
Regresyon	87,782	7	12,540	844,422	0,000
Hata	1,218	82	0,015		
Toplam	89,000	89			
Modelin uyumu R^2 : 0,986					
Tahminin standart hatası: 0,12186					
Cevap değişkeni: EURO					
Faktörler:	Katsayılar	Katsayıların standart hatası	p değeri		
Model sabiti	2,698e-11	0,013	1,000		
X_1 : ihracat	0,063	0,015	0,000		
X_2 : ithalat	0,003	0,014	0,849		
X_3 : borsa endeksi	0,002	0,031	0,951		
X_4 : işsizlik oranı	-0,150	0,027	0,000		
X_5 : enflasyon	0,823	0,065	0,000		
X_6 : cari açık	0,056	0,016	0,001		
X_7 : döviz rezervi	-0,306	0,043	0,000		

Tablo 3'e dikkat edilecek olursak model sabiti ve kullanılan açıklayıcı değişkenlerden ithalat ve borsa endeksi değerlerinin modele katkısı anlamlı değildir. Bu durumda EURO değerlerindeki değişimi ithalat ve borsa endeksi değerleri etkilememektedir. Modele katkısı en fazla olan üç değişken ise sırasıyla enflasyon, döviz rezervi ve işsizlik oranıdır. Dikkat edilecek olursa EURO ile enflasyon arasında beklendiği gibi pozitif bir ilişki, döviz rezervi ve işsizlik oranı ile de zıt yönlü bir ilişki modelde yakalanmıştır. Cevap değişkeni EURO için çoklu doğrusal regresyon modelinden elde edilen tahmin değerlerinin EURO değerlerini ne ölçüde doğru tahmin ettiğinin görülebilmesi bakımından EURO değerleri ile tahmin değerlerinin grafiği Şekil 2'de verilmektedir. Grafik birinci açıortay doğrusu üzerinde ve etrafında yer almaktadır.



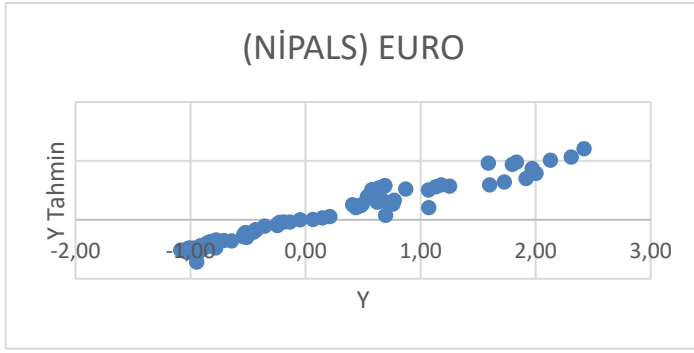
Şekil 2. Cevap değişkeni EURO değerleri ile tahmin değerlerinin grafiği (Figure 2. Graph of response variable EURO values and predicted values)

Şimdi NIPALS algoritmasını uygulayarak USD ve EURO değerlerini açıklayıcı değişkenlerin nasıl tahmin ettiğini görelim. Kullanılan model çok değişkenli doğrusal model olup açıklayıcı değişkenlerin USD ve EURO değerlerini birlikte nasıl açıkladığını göstermektedir. Elde edilen tahmin değerleri ile verinin grafiği, Şekil 3 ve Şekil 4'de sırasıyla verilmektedir.



Şekil 3. NIPALS algoritması kullanılarak USD değişkeninin değerleri ile tahmin değerlerinin grafiği (Figure 3. Graph of the values of the USD variable and the predicted values using the NIPALS algorithm)

Hata kareler ortalaması HKO= 0,9889 olarak hesaplanmıştır. Yukarıdaki şekle bakıldığında tahmin değerlerinin dar bir alanda toplandığı dikkati çekmektedir. Bu durumda USD kurunu etkilediği düşünülen açıklayıcı değişkenlerin USD kuru üzerinde etkisiz olduğu sonucuna varabiliriz. Ancak çoklu regresyon incelemesinde açıklayıcı değişkenlerin USD kuru üzerinde etkili olduğunu görmüştük. Burada karşımıza çıkan zıtlığın nedenleri kullanılan analiz yöntemlerinin birinin yetersizliğinden mi yoksa verinin yetersizliğinden mi kaynaklandığına öncelikli olarak karar vermemiz gerekmektedir. Bunun yerine kullanılan yöntemlerden çoklu doğrusal modelin veriye daha uygun olduğunu söylemek hata olur. Bu soruna yazının son bölümünde tekrar değinilecektir.



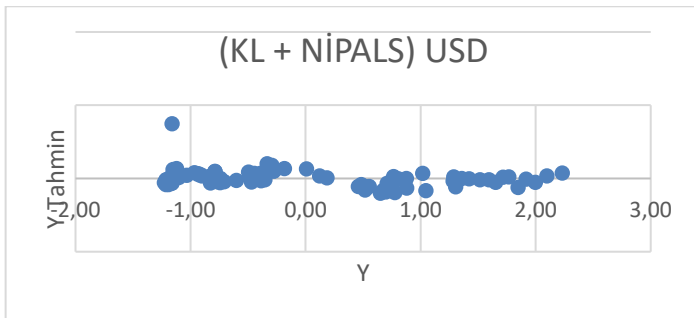
Şekil 4. NİPALS algoritması kullanılarak EURO değişkeninin değerleri ile tahmin değerlerinin grafiği (Figure 4. Graph of the values of the EURO variable and the predicted values using the NİPALS algorithm)

Hata kareler ortalaması HKO= 0,9889 olarak hesaplanmıştır. Şimdi KL dönüşümü yaptıktan sonra NİPALS algoritmasını veriye uygulayalım. Açıklayıcı değişkenlerin katsayıları Tablo 4'te verilmiştir.

Tablo 4. Regresyon katsayıları: değerler 1.0e-12 ile çarpılır (Table 4. Regression coefficients: values multiplied by 1.0e-12)

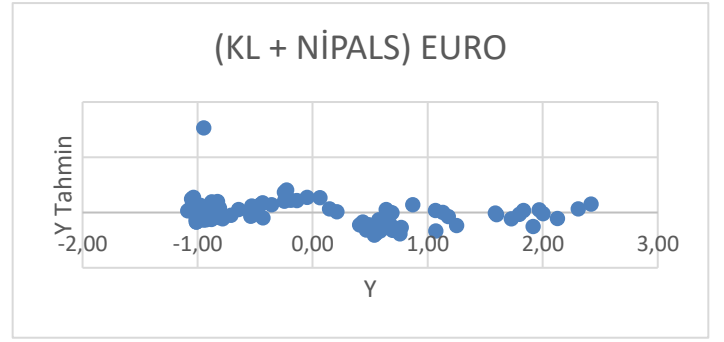
X_1 : ihracat	0.2764	-0.0064
X_2 : ithalat	-0.0032	0.0001
X_3 : borsa endeksi	-0.0145	0.0003
X_4 : işsizlik oranı	0.0958	-0.0022
X_5 : enflasyon	-0.0070	0.0002
X_6 : cari açık	-0.0195	0.0005
X_7 : döviz rezervi	-0.2828	0.0066

Tablo 4'te verilen katsayılar göre USD EURO kurlarını etkileyen en önemli açıklayıcı değişken döviz rezervidir. Döviz rezervi USD kurunu zıt yönlü etkilerken EURO kurunu aynı yönlü etkilemektedir. Şekil 5 ve Şekil 6 ile verilen grafikler sırasıyla USD ve EURO kur değerlerinin tahminlerine aittir.



Şekil 5. KL+NİPALS algoritması kullanılarak USD değişkeninin değerleri ile tahmin değerlerinin grafiği (Figure 5. Graph of the values of the USD variable and the predicted values using the KL+NİPALS algorithm)

Hata kareler ortalaması HKO= 0,0071 olarak hesaplanmıştır.



Şekil 6. KL+NİPALS algoritması kullanılarak EURO değişkeninin değerleri ile tahmin değerlerinin grafiği (Figure 6. Graph of the values of the EURO variable and the predicted values using the KL+NİPALS algorithm)

Hata kareler ortalaması HKO= 1,9707 olarak hesaplanmıştır.

4. Sonuç

Çalışmamızda sırasıyla çoklu regresyon, temel bileşenler, KL dönüşümü ve NİPALS algoritması tanıtılmıştır. Bu analizlerin neden gerekli olduğu matematiksel arka planları ile verilmiştir. Bunlara ek olarak son yapılan uygulama ile USD ve EURO kurlarını etkileyen faktörler KL dönüşümü ile birlikte NİPALS algoritması yardımıyla araştırılmıştır. Bu uygulamada USD ve EURO kurları ayrı ayrı iki çoklu regresyon modeli ile tahmin edildikten sonra NİPALS algoritması ile çok değişkenli regresyon analizi yapılarak cevap değişkeni USD ve EURO kurları birlikte kullanılmıştır. Cevap değişkenlerinin birbirlerini etkilemesi doğaldır. İki döviz kuru doğal olarak makroekonomik göstergelere bağlı birbirlerini etkileyebilir. Bu bakımdan tüm değişkenlere ilk önce KL dönüşümü uygulanarak bağımlılık ortadan kaldırılmaya çalışılmıştır.

Bu analizin sonucunda her iki kurda birlikte etkileyen ilk üç değişken sırasıyla döviz rezervi, ihracat ve işsizlik oranı olarak elde edilmiştir. Bu durum Bilgin, M. H. (2004) ve Branson, W. H. (1983)'ün çalışmaları da bunu teyit etmektedir. Döviz kuru ile dış ticaret ve işsizlik arasında doğrusal bir ilişki mevcuttur. Ancak her iki kura aynı açıklayıcı değişkenler ile çoklu regresyon yapıldığında USD kurunu etkileyen değişkenler enflasyon, döviz rezervi ve borsa endeksi çıkarken EURO kurunu etkileyen önemli değişkenler enflasyon, döviz rezervi ve işsizlik oranı çıkmaktadır. İlk olarak burada dikkati çeken önemli bir nokta enflasyonun çoklu regresyonda önemli olduğuna karar verilirken çok değişkenli regresyonda önemli olmamasıdır. Doğal olarak enflasyon zaten diğer açıklayıcı değişkenlerden etkilenen bir değişken olması nedeniyle NİPALS algoritması bu değişken etkisini diğer değişkenlere atfederek önemini zayıflatmıştır. Ancak çoklu regresyon bunu yapabilme kabiliyetine sahip olmadığından bu değişkeni en önemli değişken olarak analiz sonucunda vermektedir. Burada dikkati çeken bir başka önemli nokta da her iki regresyon incelemesinde de döviz rezervinin anlamlı olarak bulunmasıdır. Tabii ki döviz rezervi de diğer açıklayıcı değişkenlerden etkilenir. Ancak diğer değişkenlerin sonucu olabildiği gibi aslında daha çok nedeni de olmaktadır. Bundan dolayı NİPALS algoritmasında önemli bir değişken olarak karşımıza çıkmaktadır. Burada önemli bir soru merkez bankasındaki döviz rezervinin kuru hangi yönde etkilediğidir. İhracatın artması piyasalarda dönen döviz miktarının ve merkez bankası döviz rezervinin de artmasını sağlar. Ancak sadece merkez bankası döviz rezervinin artması piyasalardaki dövizin

merkez bankası tarafından çekildiğinin de bir göstergesi olabilir (Gedik, A. 2020). Döviz rezervinin artmasıyla kurun artışa geçmesi ise tamda bunun bir göstergesidir. EURO kurunda bunun gerçekleştiğini görebilmekteyiz (Buscher, H. S. ve Claudia Mueller 1997). Burada dikkat edilmesi gereken bir başka önemli ayrıntı da dış ticaretin hangi oranda EURO ve USD bölgesinde yapıldığıdır. USD kuruna bakıldığında döviz rezervinin azalması kurun artışını beraberinde getirmektedir. Merkez bankası yeterli döviz rezervine sahip olmadığında piyasalara döviz sürememekte bunun sonucu olarak ta USD kuru artmaktadır (Branson, W. H. 1983). İhracatın artmasının USD kurunu artırması ise direkt olarak dış ticaretin EURO bölgesiyle yapılmasından kaynaklıdır. Analiz sonuçlarına bakıldığında ele alınan makroekonomik göstergelerin USD kurunu değil EURO kurunu daha iyi açıklamasının nedeni de budur. Son olarak burada vurgulamak isteriz ki her araştırmada ilk önce bir ön çalışma ile seçilen değişkenlerin nasıl kullanılacağına karar verilmelidir. İnceleme sonuçlarında da vurguladığımız gibi en azından dış ticaretin böyle bir araştırmada kullanılabilmesi için yeni bir değişken olarak ticaret bölgelerinin sınıflandırılması gerekmektedir. Ele alınan veri seti tek bir yöntemle değil en azından iki yöntemle yapılmalı ve yöntemlerin verdiği sonuçların karşılaştırılması yerine farklılıklarının değerlendirilmesi gerekmektedir.

Kaynakça

- Abdi, H., & Williams, L. J. (2010). Principal component analysis. Wiley interdisciplinary reviews: computational statistics, 2(4), 433-459.
- Alexander E. Stott, Sithan Kanna, Danilo P. Mandic, William T. Pike (2017), An Online Nipals Algorithm for Partial Least Squares, Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, pp. 4177-4181.
- Andersson, M. (2009). A comparison of nine PLS1 algorithms. Journal of Chemometrics: A Journal of the Chemometrics Society, 23(10), 518-529.
- Avraham Lorber, Lawrence E. Wangent, and Bruce R. Kowalski (1987) "A Theoretical Foundation for The PLS Algorithm" Journal of Chemometrics, Vol. 1, 19-31.
- Bilgin, M. H. (2004). Döviz Kuru İşsizlik İlişkisi: Türkiye Üzerine Bir İnceleme, Kocaeli Üniversitesi Sosyal Bilimler Dergisi, (8), 80-94.
- Branson, W. H. (1983), A Model of Exchange-rate Determination with Policy Reaction: Evidence from Monthly Data, NBER Working Paper, No: 1135.
- Buscher, H. S. ve Claudia Mueller (1997), Exchange Rate Volatility Effects On The German Labour Market: A Survey Of Recent Results And Extensions, Discussion Paper Series, No. 37, Institute For The Study Of Labor, March 1997.
- Cauchy, A. L. (1829). Sur l'équation l'aide de laquelle on détermine les inégalités séculaires des mouvements des planetes. Oeuvres Completes (Ileme Série), 9.
- Courrieu, P. (2008). Fast computation of Moore-Penrose inverse matrices. arXiv preprint arXiv:0804.4809.
- Dayal, B. S., & MacGregor, J. F. (1997). Improved PLS algorithms. Journal of Chemometrics: A Journal of the Chemometrics Society, 11(1), 73-85.
- De Jong, S. (1993). SIMPLS: an alternative approach to partial least squares regression. Chemometrics and intelligent laboratory systems, 18(3), 251-263.
- Ergon, R. (2009). Re-interpretation of NIPALS results solves PLSR inconsistency problem. Journal of Chemometrics: A Journal of the Chemometrics Society, 23(2), 72-75.
- Faber, M. N., & Ferré, J. (2008). On the numerical stability of two widely used PLS algorithms. Journal of Chemometrics: A Journal of the Chemometrics Society, 22(2), 101-105.
- Gedik, A. (2020), Türkiye'de İthalat İhracat İle Enflasyon Arasında Nedensellik Analizi, Avrupa Bilim ve Teknoloji Dergisi,(19), 323-333.
- Golub, G., & Kahan, W. (1965). Calculating the singular values and pseudo-inverse of a matrix. Journal of the Society for Industrial and Applied Mathematics, Series B: Numerical Analysis, 2(2), 205-224.
- Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. Journal of educational psychology, 24(6), 417.
- Karamizadeh, S., Abdullah, S. M., Manaf, A. A., Zamani, M., & Hooman, A. (2013). An overview of principal component analysis. Journal of Signal and Information Processing, 4(3B), 173.
- Karhunen, K. (1947). Under lineare methoden in der wahr scheinlichkeitsrechnung. Annales Academiae Scientiarum Fennicae Series A1: Mathematica Physica, 47.
- Kartal, M., Depren, S. K., & Depren, Ö. (2018). Türkiye'de Döviz Kurlarını Etkileyen Makroekonomik Göstergelerin Belirlenmesi: MARS Yöntemi ile Bir İnceleme, MANAS Sosyal Araştırmalar Dergisi, 7(1), 209-229.
- Koliha, J. J. (1996). A generalized Drazin inverse. Glasgow mathematical journal, 38(3), 367-381.
- Loeve, M. (1948). Fonctions aleatoires du second ordre. Processus stochastique et mouvement Brownien, 366-420.
- Marquardt, D. W., & Snee, R. D. (1975). Ridge regression in practice, The American Statistician, 29(1), 3-20.
- Martens H, Naes T. (1989) Multivariate Calibration 3rd edn. John Wiley: New York.
- Pearson, K. (1901). LIII. On lines and planes of closest fit to systems of points in space. The London, Edinburgh, and Dublin philosophical magazine and journal of science, 2(11), 559-572.
- Penrose, R. (1955, July). A generalized inverse for matrices. In Mathematical proceedings of the Cambridge philosophical society (Vol. 51, No. 3, pp. 406-413). Cambridge University Press.
- Qin, L., Snoussi, H., & Abdallah, F. (2014, September). Online learning partial least squares regression model for univariate response data. 22nd European Signal Processing Conference (EUSIPCO) (pp. 1073-1077). IEEE.
- Ramsay, J. O., & Dalzell, C. J. (1991). Some tools for functional data analysis. Journal of the Royal Statistical Society: Series B (Methodological), 53(3), 539-561.
- Rosipal, R., & Krämer, N. (2005, February). Overview and recent advances in partial least squares, In International Statistical and Optimization Perspectives Workshop" Subspace, Latent Structure and Feature Selection" (pp. 34-51). Springer, Berlin, Heidelberg.
- Seasholtz, M. B., Pell, R. J., & Gates, K. E. (1990). Comments on the power method. Journal of chemometrics, 4(4), 331-334.
- Şit, M., & Karadağ, H. (2019). Döviz Kurunu Belirleyen Ekonomik Faktörler: Türkiye Ekonomisi İçin ARDL Sınır Testi Uygulaması, International Journal of Economic & Administrative Studies, (23).

- Terzi, H., & Kurt, S. (2007). Türkiye'de Dolarizasyon Sürecinde Döviz Kuru ve Enflasyon İlişkisi, *Ekonomik Yaklaşım*, 18(64), 1-22.
- Wold, H. (1966). Estimation of principal components and related models by iterative least squares. *Multivariate analysis*, 391-420.
- Wold, S., Ruhe, A., Wold, H., & Dunn, Iii, W. J. (1984). The collinearity problem in linear regression. The partial least squares (PLS) approach to generalized inverses. *SIAM Journal on Scientific and Statistical Computing*, 5(3), 735-743.
- Wu, W., & Manne, R. (2000). Fast regression methods in a Lanczos (or PLS-1) basis. Theory and applications. *Chemometrics and intelligent laboratory systems*, 51(2), 145-161.
- Yurdakul, F. (2016). Döviz Kuru Modellemesi ve Türkiye Üzerine Bir Uygulama, Ed. (Yurdakul, F.) içinde: Döviz Kurunun Belirleyicileri: Kısmi ve Koşullu Granger Nedensellik, SETAR, LSTAR, TVAR Modelleri, Ss. 1-51. Ankara: Gazi Kitabevi.
- Zhu, E., & Barnes, R. M. (1995). A simple iteration algorithm for PLS regression. *Journal of chemometrics*, 9(5), 363-372.