

İÇİNDEKİLER

	Sayfa
SİMGE LİSTESİ.....	iv
KISALTMA LİSTESİ.....	v
ŞEKİL LİSTESİ.....	vi
ÇİZELGE LİSTESİ.....	vii
ÖNSÖZ.....	viii
ÖZET.....	ix
ABSTRACT.....	x
1. GİRİŞ	1
2. ÇOKLU DOĞRUSAL REGRESYON MODELİ	3
2.1 Çoklu Doğrusal Regresyon Modelinin Varsayımları	4
2.2 Çoklu Doğrusal Regresyonun Matris Notasyonu ile Anlatım	7
2.3 Varyans Analizi Sonuçları	9
3. EN KÜÇÜK KARELER TEKNİĞİ (EKK).....	11
3.1 En Küçük Kareler Tekniğinin Varsayımları	13
4. ÇOKLU BAĞLANTI PROBLEMİ.....	17
4.1 Çoklu Bağlantının Saptanması	17
4.2 Çoklu Bağlantı Probleminin Çözümü.....	18
4.2.1 Varyans Büyütme Faktörü VIF (Variance Inflation Factor).....	19
5. RIDGE REGRESYON	21
5.1 Yanlı Tahminler.....	22
5.1.1 Yanlı k Sabitinin Seçimi	25
5.2 Genelleştirilmiş Ridge Regresyon Denklemi.....	27
5.3 Standartlaştırma	27
5.4 Ridge Tahminleri	28
5.5 Ridge Varyans	28
5.6 Ridge Regresyonun Ortalama Kare Hatası	28
5.6.1 5.14 denkleminin gerekçesi.....	29
5.7 Ridge Parametresinin Optimal Değeri.....	29
5.8 Varyans, Yan ve k (ridge parametresi) Arasındaki İlişki	31
5.9 Ridge İzi.....	32
5.10 Ridge Parametresinin Seçimi	33
5.11 Ridge Regresyonun Yapılan Diğer Çalışmalarla İlişkisi	35
5.12 Sıralı Ridge Tahmin Edicisi	35

6.	DOĞRUSAL OLMAYAN REGRESYON MODELİ.....	36
6.1	Doğrusal Olmayan Regresyonda Parametre Tahmini	37
6.2	Doğrusal Olmayan Ridge Regresyon Modeli	38
6.3	Kanonik Yöntem ya da İlişkisiz Bileşen.....	39
6.4	Ridge Regresyonun Kanonik Formu	41
6.5	Bayesgil İstatistiksel Çıkarsama.....	41
6.6	Bayesgil Çıkarsama ve Ridge Regresyon	45
7.	ZAMAN SERİLERİ ve TEMEL BİLEŞENLER REGRESYON.....	48
7.1	Zaman Serileri Analizi	48
7.1.1	Otoregresif (AR) Süreç	48
7.1.2	Hareketli Ortalama (MA) Süreci	49
7.1.3	ARMA ve ARIMA Süreci	49
7.2	Zaman Serilerinde Durağanlık ve Trend.....	50
7.3	Durağanlığın Tespiti	52
7.3.1	Görsel Saptama -Korelogram.....	52
7.3.2	Dickey-Fuller Testi.....	52
7.4	Granger Causality Nedensellik Testi	54
7.5	Eşbütünleşme (Koentegrasyon).....	54
7.6	Var Analizi	55
7.7	Temel Bileşenler Regresyonu	56
8.	UYGULAMA.....	58
8.1	Türkiye İthalatının Ridge Regresyon ile Analizi.....	58
8.2	Veri ile İlgili Bilgi	58
8.3	Çoklu Bağlantının Ölçülmesi ve Betimsel İstatistikler.....	62
8.4	Ridge Regresyon (RR) ve En küçük kareler (LS) Yöntemi ile Elde Edilen Sonuçlar	64
8.5	Zaman Serisi Uygulaması	71
8.6	Temel Bileşenler Regresyon Analizi Uygulaması	75
	SONUÇ.....	78
	KAYNAKLAR.....	79
	EKLER.....	82
	Ek 1 Etki – Tepki Grafikleri.....	83
	ÖZGEÇMİŞ.....	84

SİMGELİSTESİ

β_i	Çoklu doğrusal regresyon katsayısı
e_i	Artık
ε_i	Distürbans
Var	Varyans
I	Birim matris
$N()$	Normal dağılım
Kov / Cov	Kovaryans
$\hat{\beta}$	En küçük kareler tahmin edicisi
N	Ana kütle
R^2	Belirlilik katsayısı
β^R	Ridge regresyon tahmin edicisi
R	Korelasyon katsayısı
G	Ortogonal matris
λ_i	Özdeğer
S_j	x_j 'nin standart sapması

KISALTMA LİSTESİ

MLR	Multivariate Linear Regression (Çoklu doğrusal regresyon)
LS	Least Squares (En küçük kareler)
RR	Ridge regression (Ridge regresyon)
RSS	Residual Sum of Squares (Artık kareler toplamı)
SYT	β_0 artık kareler toplamı
SS_{REG}	Y artık kareler toplamı
EKK	En Küçük Kareler
VIF	Variance Inflation Factor (Varyans Büyütme Faktörü)
SD	Standart Deviation (Standart Sapma)
SSE	Sum of Square Due to Error (Hata Kareler Toplamı)
SSTO	Y'nin Kareler Toplamı
MSE	Mean Square Error (Ortalama Kare Hata)
VAR	Vector Autoregressive
ÇDDRM	Çok değişkenli doğrusal regresyon metodu
PLS	Penalized Least Squares
SVD	The single value decomposition

ŞEKİL LİSTESİ

	Sayfa
Şekil 2.1 Çoklu doğrusal regresyon modelinin geometrik açıklaması	4
Şekil 3.1 Hata terimi ve hata terimi dağılımının grafiği	14
Şekil 5.1 Yanlı ve yansız tahmin edici.....	23
Şekil 5.2 Ridge parametreleri grafiği.....	31
Şekil 5.3 Ridge yan-varyans grafiği	32
Şekil 5.4 Ridge izi grafiği	33
Şekil 8.1 Logaritmik ölçekli ridge regresyon grafikleri (ridge izi)	69
Şekil 8.2 Reel Gayri Safi Yurt İçi Hasılaya (RGDP) bir standart hatalık şok karşısında gösterilen tepkiler.....	73
Şekil 8.3 Nispi Fiyata (RP) bir standart hatalık şok karşısında gösterilen tepkiler	74

ÇİZELGE LİSTESİ

	Sayfa
Çizelge 2.1 Çoklu bağlantı çözüm istatistikleri.....	6
Çizelge 2.2 Dönüşüm şekilleri	7
Çizelge 2.3 Varyans analizi tablosu.....	10
Çizelge 8.1 Değişkenler tablosu	58
Çizelge 8.2 Türkiye ekonomisinin yıllık verileri; ithalat verileri (\$)	59
Çizelge 8.3 Betimsel istatistikler	63
Çizelge 8.4 Korelasyon tablosu	63
Çizelge 8.5 Model özeti (LS)	66
Çizelge 8.6 ANOVA (LS)	66
Çizelge 8.7 Tahmin edilen regresyon katsayıları ve güven aralıkları.....	66
Çizelge 8.8 EKK çoklu doğrusal bağlantının saptanması.....	67
Çizelge 8.9 Korelasyonların özdeğerleri.....	67
Çizelge 8.10 $k=2$ Yanlılık sabitiyle ridge regresyon sonuçları	67
Çizelge 8.11 ANOVA tablosu	67
Çizelge 8.12 Modeldeki değişkenler.....	68
Çizelge 8.13 k , R-Kare ve standartlaştırılmış ridge regresyon katsayıları.....	68
Çizelge 8.14 EKK- RR sonuç karşılaştırma tablosu.....	69
Çizelge 8.15 Geliştirilmiş Dickey – Fuller (ADF) birim kök test sonuçları.....	71
Çizelge 8.16 Granger nedensellik testi	72
Çizelge 8.17 Eşbütünleşme (Koentegrasyon) testi	72
Çizelge 8.18 İthalat için ortalama varyans ayrıştırması	74
Çizelge 8.19 Regresyon katsayıları sonuçları	76
Çizelge 8.20 Varyans (ANOVA) analizi	76
Çizelge 8.21 LS, RR ve PC teknikleriyle elde edilen sonuçların karşılaştırılması	76

ÖNSÖZ

Bu tez çalışmasında, yansız tahmin yöntemlerinden biri olan ridge regresyon incelenmiştir. Öncelikle, doğrusal regresyon modeli ve en küçük kareler yöntemi ele alınmıştır. Ridge regresyon ile birlikte, doğrusal olmayan regresyon ve doğrusal olmayan ridge regresyon modelleri incelenmiştir. Ridge ve en küçük kareler yönteminden elde edilen analiz sonuçları karşılaştırılmıştır. Bu çalışma için ridge regresyona alternatif olabileceği düşünülen yöntemlerden zaman serileri analizi ve temel bileşenler regresyon analizinden elde edilen sonuçlar ışığında modeller elde edilmiştir. Elde edilen analizlerin sonucunda tüm modeller karşılaştırılmıştır.

Bazı kavramların, Türkçe’de tam karşılıkları olmadığı için ya da birden fazla karşılıkları olduğundan, bu sözcüklerin yanına parantez içerisinde İngilizce karşılıkları da yazılmıştır.

Bu tezin hazırlanmasında yardımları için değerli hocam ve danışmanım Yrd. Doç. Dr. Atıf EVREN’e teşekkürü borç bilirim.

Tüm bunların yanında, tez çalışmam sırasında, maddi ve manevi hiçbir yardımı esirgemeyen sevgili eşim Adem SONUVAR’a ve aileme teşekkür ederim.

Esra TOKUR SONUVAR

ÖZET

Esra TOKUR SONUVAR
İstatistik, Yüksek Lisans Tezi

Yanlı tahmin yöntemlerinden Ridge Regresyon ve Temel Bileşenler Regresyon; yansız tahmin yöntemlerinden En Küçük Kareler metodu ve zaman serileri analizi yöntemleri elimizdeki ekonometrik modele uygulanmıştır. Elde edilen sonuçların güvenilirliğini anlamak için sonuçlar birbirleriyle karşılaştırılmıştır. Veri kümesinden, öncelikle En küçük kareler yöntemi ile bir model elde edilmiştir. Uygulama sırasında, çoklu bağlantı problemi olduğu saptanmıştır. Veri kümesinin çoklu bağlantı problemi olduğundan, ridge regresyon analizi uygulanmıştır. Elde edilen model ile en küçük kareler yöntemiyle elde edilen modeller karşılaştırılmıştır. Daha sonra aynı veri kümesi için zaman serisi analizi ve temel bileşenler regresyon analizi yapılmıştır. Ridge regresyon analizi sonucunda elde edilen sonuçlar ile temel bileşenler regresyon analizi sonuçlarının birbirleriyle paralellik gösterdiği görülmüştür. Ridge regresyon analizi sonuçları ve diğer analizlerden elde edilen sonuçlar birbirleriyle karşılaştırılmıştır. Ridge regresyon analizi sonuçlarının iktisat teorisini desteklediği görülmüştür. Sonuç olarak elde edilen analizler sonucunda nispi fiyatın, ithalatı önemli miktarda etkilemediği görülmüştür. Bu uygulamalar için, Excel 2003, SPSS 11.5, Eviews 5 ve NCSS 2007 kullanılmıştır.

Eldeki veri seti için uygulanan ridge regresyon analizinin sonucunda ithalatın, nispi fiyattan çok etkilenmediği görülmüştür. Aynı zamanda temel bileşenler regresyon analizi sonuçlarıyla da benzerlik gösterdiği ve dolayısıyla sonuçlarına güvenilebileceği şeklinde yorumlanmıştır.

Anahtar kelimeler: Ridge regresyon, çoklu bağlantı, en küçük kareler yöntemi, doğrusal regresyon, doğrusal olmayan regresyon, zaman serileri, temel bileşenler regresyon

ABSTRACT

Esra TOKUR SONUVAR
Statistics, M.S. Thesis

Ridge regression and principal component regression as biased estimation methods have been applied to data at hand as well as time series methodology. To confirm the reliability of all the results of obtained by all methodologies mentioned above, a comparison between all application results have been made. First of all, a model has been reached from the data with least of square. During the application the multicollinearity problem has been figured out. Since the data has a multicollinearity problem, the ridge regression analysis has been applied. Its results and least square model has been compared to each other. After that, for the same data, time series analysis and principal component regression analysis have been executed. Both the results of ridge regression and principal component regression analysis have had an parallelity. Then the results of ridge regression and other analysis have been compared to each other. Finally it has seen that the results of ridge regression supports the economical theory. At the end of the all analysis and comprasions; it has been revealed that the relative price doesn't affect the import too much. Excel 2003, SPSS 11.5, EvIEWS 5 and NCSS 2007 have been implied for these applications.

As a results of ridge regression analysis which has been applied for handled data set, it has been seen that import isn't affected by the relative price. At the same time has been interpreted that these results can be reliable.

Key words: Ridge regression, multicollinearity, least of square (LS), linear regression, nonlinear regression , time series, principal component regression

1. GİRİŞ

Çok değişkenli regresyon analizi ekonometrik verilerin analizleri için çok geniş bir alanda kullanılan ekonometrik araçlardan birisidir. Regresyon analizi birden fazla değişken ve bunlar arasındaki bağıntıların incelenmesinde kullanılan bir yöntemdir. (Anderson, T. W. , 2003) Regresyon analizindeki standart yaklaşım; bir veri setindeki ilişkileri özetlemek, tanımlamak, katsayı tahmininde bulunmak, bir sistemin kontrol ve gözetimini yapmak, değişken seçmek gibi amaçlarla bir regresyon eşitliği araştırmak şeklindedir. Bu amaçla regresyon analizinin kullanımı için bazı varsayımların gerçekleşmesi gereklidir. Tahmin edicilerin sonuçlarının tatmin edici olup olmadığı t, F testleri gibi testlerle sınanır. Belirlilik katsayısı R^2 elde edilir. (Draper, N.R., H. Smith, 1981)

Bu tezin amacı, yanlış tahmin yöntemlerinden Ridge Regresyon (RR), Temel Bileşenler Regresyon (PC); yansız tahmin yöntemi olan En Küçük Kareler (EKK) ve zaman serileri analizi yöntemlerinin elimizdeki ekonometrik veriye uygulanmasıdır.

Problem olarak Türkiye ithalatı fonksiyonunun tahmini seçilmiştir. Çözümüne geçilmeden önce problemle ilgili istatistik yöntemleri anlatılmış ve bazı özel durumlar açıklanmıştır. Bu istatistik yöntemlerinden çoklu doğrusal regresyon, EKK, ridge regresyon, doğrusal olmayan regresyon, kanonik yöntem, Bayesgil istatistik çıkarsama, zaman serileri analizi ve temel bileşenler regresyon analizleri ele alınmıştır.

İkinci bölüm çoklu doğrusal regresyon modelinin genel yapısı hakkında bilgi vermektedir. Bu bölümde çoklu doğrusal regresyon modelinin varsayımları, matris notasyonu ve varyans analizi incelenmiştir.

Üçüncü bölümde EKK yöntemi genel olarak ele alınmıştır. Varsayımları incelenmiştir.

Dördüncü bölümde yukarıda bahsedilen özel durumlardan biri olan çoklu bağlantı problemi incelenmiştir. Bu bölümde çoklu bağlantının saptanması ve probleminin çözümü hakkında bilgi verilmiştir.

Beşinci bölümde ridge regresyon modelinin genel yapısı incelenmiştir. Ridge tahminleri ridge izi ayrıntılı olarak ele alınmıştır.

Altıncı bölümde doğrusal olmayan regresyon modeli genel olarak ele alınmıştır. Bu bölümde doğrusal olmayan ridge regresyon, ridge regresyonun kanonik ve Bayesgil formları incelenmiştir.

Yedinci bölümde zaman serileri ve temel bileşenler regresyon modelinin genel yapıları ele alınmıştır. Zaman serileri ve temel bileşenler regresyonun varsayımları incelenmiştir.

Sekizinci bölümde uygulama için seçilen veri setine ilk önce EKK uygulanmış ve bu analiz sonucu model hakkında bilgiler elde edilmiştir. Daha sonra aynı veri setine ridge regresyon analizi uygulanmıştır. Bu iki analizin sonucu karşılaştırılmıştır. Aynı şekilde veri setine önce zaman serileri analizi, sonra temel bileşenler regresyon analizi yapılmıştır. Tüm analizlerden elde edilen sonuçlar birbirleri ile karşılaştırılmıştır.

Bu uygulamalar sırasında, Excel 2003, SPSS 11.5, Eviews 5 ve NCSS 2007 paket programları kullanılmıştır.

2. ÇOKLU DOĞRUSAL REGRESYON MODELİ

Çoklu doğrusal regresyon (MLR) bağımlı bir değişkenle bir ya da daha fazla bağımsız değişkenler arasındaki doğrusal ilişkiyi modellemek için kullanılan bir yöntemdir. Bağımlı değişken tahmin edilen, bağımsız değişkenler ise tahmin edici değişkenler olarak da adlandırılır. k sayıda bağımsız değişkeni ve bir bağımlı değişkeni olan modeli;

$$Y_i = P_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + \varepsilon \quad (2.1)$$

şeklinde ifade edilebilir.

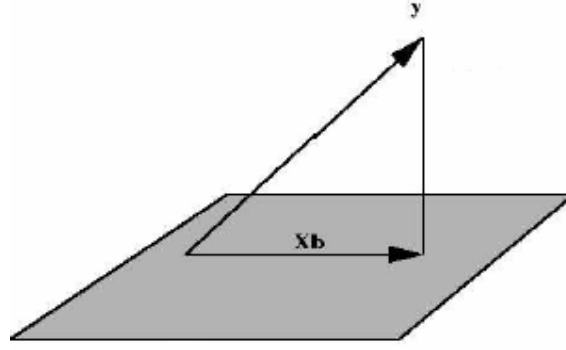
Modelde Y bağımlı değişkeni ve $X_1, X_2, \dots, X_k$ bağımsız değişkenleri gösterir. $\beta_0, \beta_1, \dots, \beta_k$ terimleri çoklu doğrusal regresyonun katsayıları olup alacakları değerler sabittir. ε , sıfır etrafında normal dağıldığı kabul edilen bir rassal değişkendir. “Burada bağımsız değişken sayısı k, parametre sayısı ise k+1” dir. (Draper, N.R., H. Smith, 1981)

Çoklu regresyonun temel amaçlarından biri bağımsız değişkenler yardımı ile bağımlı değişkenin tahminidir. Bu tahminlerde çoklu regresyonun iki hedefi vardır. Birinci hedef rastlantı değişkeni gibi varsayılan bağımsız değişkenlerin, toplam tahmin edici gücünü maksimize etmektir. Tahmin edicinin doğruluğunun ölçüsü, tahmin edicinin gücünü öğrenmeye ilişkin önemlilik istatistik testleri yapılarak oluşturulur. Bütün örneklerde tahmin edicinin ilk odak noktası olsun olmasın, regresyon analizi uygulamalarında, tahmin edicinin doğruluğu kabul edilebilir düzeyde başarılı olmalıdır. Analizi yapan kişi, önemliliği hem istatistiksel olarak hem de pratik olarak tetkik ederek sağlamalıdır. (David Birkes, Y. Dodge, 1993)

Çoklu regresyonda ikinci bir hedef iki ya da daha fazla bağımsız değişkenler dizisinin, rastlantı değişkenlerinin tahmin edici gücünün araştırılmasıdır. Bu analiz tipinin öncelikli odak noktası, herhangi bir durumda seçilen modelin tahmininde hem istatistiksel hem de pratik de önemliliğin gösterilmesine rağmen, modeller arasında göreceli tahmin edici gücünün olmasıdır. (Anderson, T. W. , 2003)

Sonuç olarak, çoklu regresyon bağımlı değişkenin tahmininde, bağımsız değişkenler arasındaki ilişkilerin içine nüfuz etmesini sağlar. (Stein, C. M., 1960)

Çoklu regresyon modelinin geometrik olarak açıklaması aşağıdaki gibidir.



Şekil 2.1 Çoklu doğrusal regresyon modelinin geometrik açıklaması

Bağımsız değişkenlerde meydana gelen bir birimlik değişim, bağımlı değişkende sabit bir değişiklik meydana getiriyorsa; bu değişkenlerin etkisi çoklu doğrusal regresyonla değerlendirilir. (Stein, C. M., 1960)

2.1 Çoklu Doğrusal Regresyon Modelinin Varsayımları

ÇDDRM (çok değişkenli regresyon modeli) çeşitli varsayımlara dayanır. Regresyon tahminleri varsayımları sağlandığında yansızlık, etkinlik (efficiency) ve tutarlılık gibi tahminin optimal olmasını sağlayacak kriterlere de uyması gerekmektedir. (Anderson, T. W. , 2003)

Yansızlık, tahmin edicinin beklenen değerinin, parametrenin doğru değerine eşit olması anlamına gelmektedir. Etkinlik, tahmin edicinin diğer tahmin edicilerin varyansından daha küçük bir varyansa sahip olması demektir . Tutarlılığın anlamı, tahmin edicinin yanlı olması ve tahmin edicinin varyansının sıfıra yaklaşması, örnek boyutunun sonsuza yaklaşması demektir. Tutarlılık asimptotik bir özelliktir. n arttıkça sistematik hata azalıyorsa tahminci tutarlıdır. (Seber, G. A. F. , 2003)

Ostrom (1990), regresyon modeli için 6 temel varsayım listelemiştir:

1. Doğrusallık: Tahmin edici ve tahmin edilen arasındaki ilişki doğrusaldır. ÇDDRM doğrusal ilişkiler için varsayılr. Eğer ilişkiler doğrusal değilse, iki yönteme başvurulur:

- a. Aradaki ilişkiyi doğrusal yapmak için verileri dönüştürmek,
- b. Alternatif bir istatistiksel yöntem kullanmak.

Çizilen saçılım grafiği incelenerek açıklayıcı değişkenin doğrusallığı kontrol edilmelidir.

2. Çoklu Doğrusal Bağlantı (Multicollinearity): Çoklu doğrusal bağlantı terimi, bağımsız değişkenler gözlem matrisinin (X) sütunlarının doğrusal bağımlılığı şeklinde tanımlanabilir. Bağımsız değişkenler arasındaki çoklu bağlantı problemi;

- a. Veri toplama yönteminin yanlış olması,
- b. Örnekleme veya modelde yapılan kısıtlamalar,
- c. Modelin tanımlanmasındaki hatalar,
- d. Model seçiminde yapılan hatalar,

gibi nedenlerden kaynaklanabilir. Verideki bu problem, regresyon katsayılarına ait en küçük kareler tahmincilerinin varyans ve kovaryans değerlerinin büyük olmasına, buna bağlı olarak da regresyon modeline dayanan yorumların hatalı olmasına neden olabilir. (Ostrom, C.W., Jr., 1990)

Bağımsız değişkenler arasında çoklu bağlantı problemi olup olmadığının belirlenmesi için birçok yöntem önerilmektedir. Bu problemin tespiti için önerilen yöntemlerin çoğunluğu X bağımsız değişken matrisi ile ilgilidir. Bunlar arasında en çok kullanılan istatistiklerle ilgili özet bilgiler Çizelge 2.1’de verilmiştir. (David Birkes, Y. Dodge, 1993)

Bağımsız değişkenler arasında tespit edilen birlikte değişim probleminin çözümü için modeli yeniden tanımlamak, konu ile ilgili ek veriler toplamak ve en küçük kareler yöntemi yerine ridge regresyon yöntemi önerilmektedir. (Seber, G. A. F. , 2003)

Çizelge 2.1 Çoklu bağlantı çözüm istatistikleri

Istatistik (Used Statistic)	Formülü (Formüla)	Deger (Critical Value)	Istatistiklerin Önemli Özellikleri (Some Important Properties of Statistics)
F ve t istatistikleri	-	-	• Her zaman geçerli bir yol degil.
Korelasyon Katsayısı(r_{ij})	-	-	• Sadece ikili bagimliliği belirler, çoklu bagimlilikta yetersiz kalir.
Varyans Sisme Faktörü (VIF _j)	$(1-R_j^2)^{-1}$	10	• Birlikte degisim nedeniyle j'nci regresyon katsayisinin kötü bir tahminin yapıldigini ifade eder • Bagimsiz degiskenin birlikte degisime katkisinin derecesini ölçer
$\det(X'X)$	$ X'X $	-	• Birlikte degisimin hangi degiskenler arasında olduğunu belirlemeyip sadece bagimliligin varligi ile ilgili genel sonuç verir.
Durum Sayisi (κ_j)	$\lambda_{\max} / \lambda_j$	100	• Hesaplanan özdegerlerin en büyüğünü diger özdegerlere bölmek suretiyle hesaplandigindan belli bir degisken ile diger degiskenler arasındaki bagimliliği belirler.
Durum İndeksi (η_j)	$\mu_{\max} / \mu_j$	30	• iki degisken arasındaki ikili bagimliliği belirledigi gibi, çoklu bagimliliği da belirler.
Varyans Ayrısım Oranı (π_j)	$\frac{(v_{jj}^2 / \mu_j^2)}{VIF_j}$	0.5	• iki degisken arasındaki ikili bagimliliği belirledigi gibi, çoklu bagimliliği da belirler.. • İlgili j'nci bagimsiz degiskenler arasında birlikte degisim olduğunu ve birlikte degisim ile ilgili olan j'nci tekil degerin (μ_j), $\text{Var}(\beta_j)$ 'yi önemli derecede etkiledigini ifade eder. • En fazla önerilen istatistiktir.

3. Sıfır Ortalama: $E[\varepsilon_i] = 0$. Distürbansların beklenen değeri sıfırdır. Bu bir problem değildir, çünkü en küçük kareler metoduyla tahmin edilen regresyon denklemlerinin ortalaması sıfırdır. (David Birkes, Y. Dodge, 1993)

4. Sabit Varyans: $E(\varepsilon_i^2) = \sigma^2$. Distürbansların varyansı sabittir. Sabit varyans varsayımı, regresyon analizinde temel varsayımlardan birisidir. Bu nedenle, distürbansların sabit varyansa sahip olup olmadıklarının belirlenmesi ve doğrulanması önemlidir. Eğer bu problem çözülmezse, en küçük kareler tahminçileri yine sapmasızdırlar, fakat minimum varyans özelliğine sahip olmayabilirler. Diğer bir ifade ile katsayılar gerekenden büyük standart hatalara sahip olacaklardır. (David Birkes, Y. Dodge, 1993)

Distürbans gözlemlenmeyebilir ama artıklar elde edilebilir. Problemin varlığının incelenmesi için en çok kullanılan yöntem, artıkların, tahmin değerlerine karşı grafiklerinin oluşturulmasıdır. Bu problemin çözümü için önerilen en etkili yöntem, Y değişkenine dönüşüm uygulamaktır. Yapılan dönüşüm şekli σ^2 ile $E(y)$ arasındaki ilişkiye bağlıdır. Buna göre dönüşümler aşağıdaki Çizelge 2.2.'de verilmiştir. (Anderson, T. W. , 2003)

Çizelge 2.2 Dönüşüm şekilleri

σ^2 ile $E(y)$ arasındaki ilişki (Relationship between σ^2 and $E(y)$)	Önerilen Transformasyon (Suggested Transformation)
$\sigma^2 = \text{sabit bir sayı}$	yok
$\sigma^2 \propto E(y)$	$y' = \sqrt{y}$
$\sigma^2 \propto E(y)(1-E(y))$	$y' = \sin^{-1} \sqrt{y}, 0 \leq y \leq 1$
$\sigma^2 \propto (E(y))^2$	$y' = \ln(y)$
$\sigma^2 \propto (E(y))^3$	$y' = y^{-1/2}$
$\sigma^2 \propto (E(y))^4$	$y' = y^{-1}$

5. Otoregresyonun Olmaması Varsayımı (Nonautoregression): $E(\varepsilon_i, \varepsilon_{i-m}) = 0, m \neq 0$. Distürlanslar rasgeledir ya da ilişkisizdir. Bu varsayım zaman serileri uygulamalarında bozulması en muhtemel varsayımlardan birisidir. (David Birkes, Y. Dodge, 1993)

6. Normallik Varsayımı: Distürlans normal dağılır. $\varepsilon_i \sim N(0, \sigma^2); i=1,2,\dots,n$. Bu varsayım, regresyon denkleminin geçerli olması için, katsayıların önemliliğini test eden geleneksel testlerden birisidir. (Scott M. Kowalski, 2001)

2.2 Çoklu Doğrusal Regresyonun Matris Notasyonu ile Anlatım

Çoklu doğrusal regresyon modelleri için istatistiksel sonuçlar, parametre tahmini, istatistiksel testler gibi bağımsız değişkenler iki ya da daha fazla olduğunda çok karmaşık olabilmektedir. Matris notasyonunda kullanılan karmaşık ifadeler sadeleştirilebilir. Genel olarak bir vektör ya da bir matris koyu renklerle gösterilir $\mathbf{X}$, $\boldsymbol{\varepsilon}$ gibi. Vektörlerin ve matrislerin elemanları x_{ij} , ε_i , formunda gösterilirler. (Seber, G. A. F. , 2003)

$\mathbf{Y}$ ve $\mathbf{e}$ elemanları y_i ve ε_i 'lerden oluşan $n \times 1$ 'lik bir vektördürler.

$$\mathbf{Y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \quad \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

$\boldsymbol{\beta}$, $(p+1) \times 1$ boyutlu $\boldsymbol{\beta}_0$ kesim noktasını da içeren vektör parametresi olarak tanımlanır.

$$\beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{bmatrix}$$

Ve $\mathbf{X}$ $n \times (p+1)$ boyutlu bir matristir;

$$\mathbf{X} = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1p} \\ 1 & x_{21} & x_{22} & \dots & x_{2p} \\ \vdots & \vdots & & & \vdots \\ 1 & x_{n1} & x_{n2} & & x_{np} \end{bmatrix}$$

$\mathbf{X}$ matrisi gözlenen tahmin edici değişkenlerin gözlemlerini içermektedir ve en soldaki sütun birlerden oluşur. Birlerden oluşan bu sütun sabit terimi hesaba katacak şekilde eklenmiş $n \times (k+1)$ boyutlu bir matris; $\mathbf{Y}$, $\mathbf{Y}'$ 'nin gözlenen değerlerinden oluşan $n \times 1$ boyutlu matris (ya da sütun vektörü); β , da, regresyon katsayılarının en küçük kareler tahminlerinden oluşan $(k+1) \times 1$ boyutlu bir matristir (ya da sütun vektörüdür). (Seber, G. A. F. , 2003)

Yukarıdaki matrisleri kullanarak çoklu doğrusal regresyon denklemini matris terimi olarak aşağıdaki gibi yazılabilir.

$$\mathbf{Y} = \mathbf{X} \beta + \varepsilon \quad (2.2)$$

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1p} \\ 1 & x_{21} & x_{22} & \dots & x_{2p} \\ \vdots & \vdots & & & \vdots \\ 1 & x_{n1} & x_{n2} & & x_{np} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

Modelin formülasyonunu yaklaşık olarak şu şekilde yapılabilir;

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki} + \varepsilon_i \quad (i = 1, 2, \dots, n; n \geq k + 1) \quad (2.3)$$

$$\varepsilon_i \sim N(0, \sigma^2)$$

β 'nın en küçük kareler tahmin edicisi olan $\hat{\beta}$ 'nin seçimi artık kareler toplamının minimize edilmesi ile seçilir. Varsayalım ki x_i^T , $\mathbf{X}$ vektörünün i. satırını oluştursun, $i=1,2,\dots,n$; x_i devrik yapmamızın sebebi; bütün vektörlerin sütun vektörleri olmasıdır.

$\hat{\beta}$ 'yi minimize eden fonksiyon;

$$RSS(\beta) = \sum (y_i - x_i^T \beta)^2 = (Y - X\beta)^T (Y - X\beta) \quad (2.4)$$

Buradan da β 'nın en küçük kareler tahmin edicileri

$$\hat{\beta} = (X^T X)^{-1} X^T Y \quad (2.5)$$

şeklinde bulunur. $\hat{\beta}$ tahmin edicisi sadece $(X^T X)$ ve $(X^T Y)$ istatistiklerine bağlıdır. Bunlar düzeltilmemiş kareler toplamalarının matrisleri ve çarpımlarıdır.

$$X^T X = \begin{bmatrix} \sum (x_{i1} - \bar{x}_1)^2 & \dots & \sum (x_{i1} - \bar{x}_1)(x_{ip} - \bar{x}_p) \\ \vdots & & \vdots \\ \sum (x_{ip} - \bar{x}_p)(x_{i1} - \bar{x}_1) & \dots & \sum (x_{ip} - \bar{x}_p)^2 \end{bmatrix}$$

$$X^T Y = \begin{bmatrix} \sum (x_{i1} - \bar{x}_1)(y_i - \bar{y}) \\ \vdots \\ \sum (x_{ip} - \bar{x}_p)(y_i - \bar{y}) \end{bmatrix}$$

$E(\varepsilon)=0$ ve $\text{Var}(\varepsilon)=\sigma^2 I_n$ ise o zaman $\hat{\beta}$ yansızdır, $E(\hat{\beta})=\beta$, ve

$$\text{Var}(\hat{\beta})=\sigma^2 (X^T X)^{-1} \text{ olur.} \quad (2.6)$$

σ^2 'nin tahmini

$$\sigma^2 = \frac{RSS}{n-p}, \quad p=p+1$$

olarak elde edilir. (John Fox, 1991)

2.3 Varyans Analizi Sonuçları

Varyans analizi değişkenleri bölmek için kullanılacağı gibi, farklı değişken gruplarını içeren modellerin karşılaştırılmasında da kullanılır. Varyans analizinin genel olarak tam modeli;

$$Y=X\beta +e \quad (2.7)$$

olarak gösterilir. Eğer modelde X değişkeni yoksa

$$Y=\beta_0 I +e \quad (2.8)$$

olarak gösterilir. Burada I $n \times 1$ 'lik bir vektördür.

Model için $\beta_0 = \bar{y}$ ve $SY\bar{Y}$ β_0 için artık kareler toplamıdır. Diğer yandan β 'nın artık kareler toplamı ise RSS dir. Açıkça $RSS < SY\bar{Y}$ olması gerekir ve aralarındaki fark ise;

$SS_{REG} = SY\bar{Y} - RSS$ 'dir. Bu Y 'nin kareler toplamına uymaktadır ancak büyük modeller için açıklanabilirken, küçük modeller için açıklanamamaktadır. (Sanford Weisberg, 1985)

Çizelge 2.3 Varyans analizi tablosu

Varyans Analizi Tablosu			
V.K (Değişim Kaynağı)	S.D. (D.F)	K.T (SS)	K.O (MS)
Regresyon	p	SS_{reg}	SS_{reg}/p
Hata	$\frac{n-p'}{n-1}$	$\frac{RSS}{SY\bar{Y}}$	$RSS/(n-p') = \hat{\sigma}^2$

Eğer SS_{REG} , $F(p, n-p')$ dağılımlı σ^2 'nin regresyonu için, ortalama karenin oranını karşılaştırabilecek kadar büyükse, X 'in regresyonunun önemliliği değerlendirilebilir. Eğer hesaplanan F değeri uygun bir kritik değeri aşarsa o zaman, X 'in önemlilik değerine bakarak modelin iyi bir model olduğu söylenir. Eğer hatalar $N(0, \sigma^2)$ dağılırsa, hesaplanan oran tam bir F dağılımına sahip olacaktır. F -testinin hipotezleri:

$$H_0 = \beta_1 = \beta_2 = \dots = \beta_{p-1} = 0$$

$$H_1 = \beta_k \text{ 'lardan en az biri } \neq 0$$

olarak gösterilir. (Sanford Weisberg, 1985)

3. EN KÜÇÜK KARELER TEKNİĞİ (EKK)

En küçük kareler tekniği günümüzde yaygın olarak kullanılan regresyon tekniklerinden bir tanesidir. Bu teknik 1795 yılında Alman Carl Frederich Gauss ve 1805 yılında Fransız Adrien Marie Legendre tarafından birbirinden habersiz olarak geliştirdikleri bir tekniktir. Tekniğe ilişkin ilk uygulamalar astronomi ve yerbilim verilerinin analizinde kullanılmıştır. (Sanford Weisberg, 1985) En küçük kareler tekniği ilk olarak 1805 yılında Legendre tarafından yayınlanan bir kitapta kuyruklu yıldızların yörüngelerinin hesaplanmasında kullanılmıştır. (Vijay Mahajan, Arun K. Jain, M. B. , 1986).

$$Y = \beta_0 + \beta_1 X + \varepsilon \quad (3.1)$$

doğrusal ilişkisi X ve Y değişkenlerinin ana kütleleri için geçerlidir. İstatistiksel çalışmaların çoğunda olduğu gibi regresyon analizinde de ana kütleye ait verilerin tümüne ulaşamadığından bu ana kütlede seçilen örneklem verileriyle analiz yapılır. Örneklem verileri kullanılarak ana kütle parametreleri olan β_0 ve β_1 'in kestirimlerini elde edilebilmek için en küçük kareler tekniğinden yararlanılır. X ve Y değerleri için (X_1, Y_1) , (X_2, Y_2) , ..., (X_k, Y_k) ikili gözlemleri bir örneklem oluşturur. Bunlar yardımıyla $Y = \beta_0 + \beta_1 X_i$ denkleminin katsayılarını bulma olanağı yoktur. Örnekten hareketle bunlar tahmin edilebilir. (Montgomery, D.C. , E. A. Peck, 1992)

Regresyon modelin de yer alan β_0 ve β_1 terimlerinin regresyon denklemindeki karşılıkları $\hat{\beta}_0$ ve $\hat{\beta}_1$ olarak ele alındığında, regresyon doğrusunun denklemi; $\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X$ şeklinde olacaktır. Denkleminde yer alan $\hat{\beta}_0$ ve $\hat{\beta}_1$ terimlerinin değerlerini bulmak için en küçük kareler tekniği kullanılır. Bu tekniğin temelinde, toplam sapmaların karelerinin toplamını en küçük yapacak değerlerin bulunması yatmaktadır. Burada $\hat{Y}$, X değişkeninin bir X değerine karşılık gelen Y değişkeninin hesaplanan değerini; $\hat{\beta}_0$ ve $\hat{\beta}_1$ ise β_0 ve β_1 için hesaplanan kestirim değerlerini göstermektedir. $\hat{Y}$ kestirim değeriyle Y gözlem değerleri arasındaki farklar hata terimlerini oluşturur. (Marquardt, D.W. 1963)

$$e = Y - \hat{Y} \quad (3.2)$$

şeklinde hesaplanan hata terimleri ‘pozitif’, ‘negatif’ veya ‘sıfır’ değerine sahip olurken; bu farkların toplamı sıfıra eşittir.

$$\sum_{i=1}^k e_i = \sum_{i=1}^k (Y_i - \hat{Y}_i) = 0 \quad (3.3)$$

En küçük kareler yönteminin esası; β_0 ve β_1 parametre tahmincilerinin kestirimleri olan $\hat{\beta}_0$ ve $\hat{\beta}_1$ 'yi söz konusu farkı en küçük yapacak şekilde belirlemektir. Yani,

$$\sum_{i=1}^k e_i^2 = \sum_{i=1}^k (Y_i - \hat{Y}_i)^2 = \text{en küçük} \quad (3.4)$$

olacak şekilde belirlemektir. Burada regresyon katsayılarının en küçük kareler tahminlerini elde edebilmek için,

$$\sum_{i=1}^k (Y_i - (\hat{\beta}_0 + \hat{\beta}_1 X_i))^2 = q \quad (3.5)$$

eşitliğinde $\hat{\beta}_0$ ve $\hat{\beta}_1$ 'ya göre kısmi türevler alınıp sıfıra eşitlendiğinde;

$$\frac{\partial q}{\partial \hat{\beta}_0} = \sum_{i=1}^k (-2)[Y_i - (\hat{\beta}_0 + \hat{\beta}_1 X_i)] = 0 \quad (3.6)$$

ve

$$\frac{\partial q}{\partial \hat{\beta}_1} = \sum_{i=1}^k (-2)X_i[Y_i - (\hat{\beta}_0 + \hat{\beta}_1 X_i)] = 0 \quad (3.7)$$

eşitlikleri ve buradan da;

$$\sum_{i=1}^k Y_i = \hat{\beta}_0 n + \hat{\beta}_1 \sum_{i=1}^k X_i \quad (3.8)$$

ve

$$\sum_{i=1}^k X_i Y_i = \hat{\beta}_0 \sum_{i=1}^k X_i + \hat{\beta}_1 \sum_{i=1}^k X_i^2 \quad (3.9)$$

normal denklemler olarak isimlendirilen eşitliklere ulaşılır. (Marquardt, D.W. 1963)

$\hat{\beta}_1$ formülü ise;

$$\hat{\beta}_1 = \frac{n \left[\sum_{i=1}^k X_i Y_i \right] - \left(\sum_{i=1}^k X_i \right) \left(\sum_{i=1}^k Y_i \right)}{n \left(\sum_{i=1}^k X_i^2 \right) - \left(\sum_{i=1}^k X_i \right)^2} = \frac{\sum_{i=1}^k (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^k (X_i - \bar{X})^2} \quad (3.10)$$

$$\hat{\beta}_1 = \frac{\sum_{i=1}^k Y_i + \hat{\beta}_1 \sum_{i=1}^n X_i}{n} = \bar{Y} - \hat{\beta}_1 \bar{X} \quad (3.11)$$

olarak bulunur. Burada $\bar{X}$ ve $\bar{Y}$ sırasıyla X_i ve Y_i değerlerinin ortalamalarını göstermektedir.

Genelleştirilmiş doğrusal regresyon modeli için en küçük kareler normal denklemi:

$$X'Xb = X'Y \quad (3.12)$$

ve

en küçük kareler tahmin edicisi denklemi

$$b = (X'X)^{-1}(X'Y) \quad (3.13)$$

$\begin{matrix} p \times 1 & & p \times p & & p \times 1 \end{matrix}$

şeklinde gösterilir. (Marquardt, D.W. 1963)

3.1 En Küçük Kareler Tekniğinin Varsayımları

E.K.K. tekniğinin gerek değişkenler gerekse hata terimine ait bir takım varsayımları bulunmaktadır. Bu temel varsayımlar şu şekilde özetlenebilir.

1. Distürbans stokastik (tesadüfi) bir değişkendir.

Hangi değer veya sonucu olacağı önceden bilinmeyen, aldığı değerler şansa bağlı olarak ortaya çıkan rassal bir değişkendir. ε_i 'ler sürekli aynı yönde bir sapma göstermeyip, düzensiz bir dağılım gösterirler. (Raymond H. Myers, 1990)

2. Distürbansın ortalaması (beklenen değeri) sıfırdır.

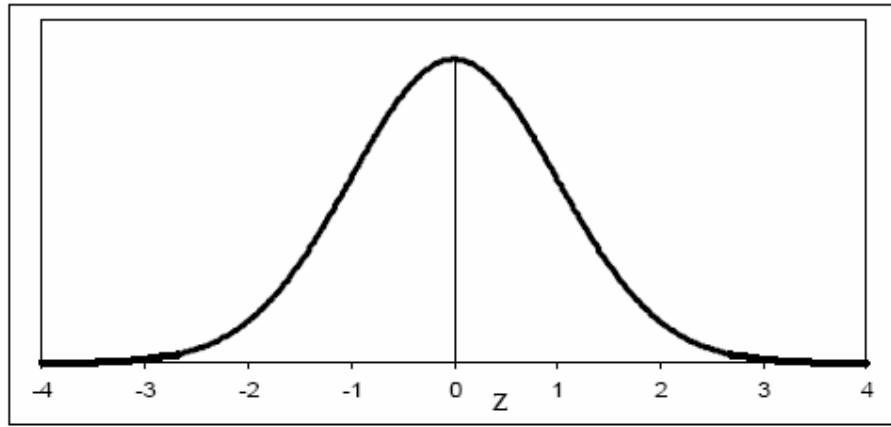
X' 'in aldığı farklı değerler için ε_i 'de artı, eksi ve sıfır gibi çeşitli değerler almaktadır. Bu değerlerin ortalaması sıfıra eşittir. (Tatlıldil Hüseyin, 1996)

$$E(\varepsilon_i) = 0, i = 1, 2, \dots, n$$

3. Distürbans normal dağılımla dağılır.

Distürbans normal dağılımla dağılır. “Yani bağımsız değişkene ait her bir distürbans değerleri kendi ortalamaları etrafında normal dağılım eğrisi şeklinde simetrik bir dağılım gösterir.” (Raymond H. Myers, 1990)

$$E(e_i) = 0$$



Şekil 3.1 Hata terimi ve hata terimi dağılımının grafiği

4. Distürbans sabit varyanslıdır. (Tatlıldil Hüseyin, 1996)

Distürbansların varyansı sabittir. Yani değişen varyanslılığın olmaması varsayımı olarak da ifade edilebilir. Bu varsayıma göre, distürbans varyansları, bağımsız değişkendeki değişmelere göre değişmeyip sabit kalmaktadır. Yani, σ^2 , X_i 'nin bir fonksiyonu değildir.

$$\sigma^2 \pm f(X_i)$$

$$\text{Var}(\varepsilon_i) = \sigma^2$$

5. Distürbanslar arasında bir ilişki (otokorelasyon) yoktur. (Sanford Weisberg, 1985)

Bu varsayıma göre, distürbanslar birbirinden bağımsız olup, birbirlerini etkilememektedirler. Bu durum, ε_i ve ε_j 'nin ortak varyanslarının (kovaryanslarının) sifıra eşit olmasını gerektirir.

$$\text{Kov}(\varepsilon_i, \varepsilon_j) = E\{[\varepsilon_i - E(\varepsilon_i)][\varepsilon_j - E(\varepsilon_j)]\} \quad (3.14)$$

$E(\varepsilon_i) = E(\varepsilon_j) = 0$ olduğundan,

$$\text{Kov}(\varepsilon_i, \varepsilon_j) = E(\varepsilon_i, \varepsilon_j) = 0, i \neq j$$

olmaktadır.

6. Bağımsız değişken ile hata terimi arasında bir ilişki yoktur. (Raymond H. Myers, 1990)

Bağımsız değişkene ait değerler daha önceden bilinen değerlerdir ve ana kütleden çekilecek her bir örneklem için bu değerler değişmez.

7. Bağımsız değişkenler arasında bir ilişki (çoklu doğrusal bağıntı) yoktur.

Bölüm 4’de ayrıntılı olarak açıklanmıştır.

8. Bağımlı değişken (Y) ile distürbans (ε_i) arasında korelasyon yoktur. (Raymond H. Myers, 1990)

$$\text{Kov}(\varepsilon_i, Y_i) = 0.$$

Yukarıda sayılan varsayımların dışında aşağıdaki varsayımlarda mevcuttur:

- Doğrusal regresyon analizinde bağımsız değişkenler sabit olmasına rağmen, bağımlı değişken tesadüfidir.
- Değişkenler hatasız ölçülmüştür.
- Bağımlı ve bağımsız değişkenler arasındaki ilişki doğrusaldır. Fakat bu doğrusallık koşulu kuşkusuz modelin parametreleri için gereklidir (değişkenler doğrusal olmayabilir).
- Birim değerleri sayısı (n), değişken sayısından (p) büyük olmalıdır: $n > p$.
- Bağımsız değişkenlerin (X_i) varyansı sıfırdan büyük olmalıdır. Değişkenlerin tüm gözlem değerleri birbirine eşit ise, $X_i = \bar{X}$ olacağından, regresyon doğrusunun eğimi tanımsız olmaktadır.
- Model doğru tanımlanmış olmalıdır. İlişkiye uygun fonksiyonun ve dahil edilmesi gereken tüm değişkenlerin dikkate alınması gerekmektedir. (Stein, C. M., 1960)

Yukarıdaki varsayımlardan birisinin sağlanmaması durumunda EKK tahmincileri yanlı (biased), tutarsız veya etkisiz olmaktadır. Söz konusu tahminciler aşağıdaki ilk üç koşulu sağlaması durumunda en iyi doğrusal tahminciler olarak kabul edilmektedir. $\hat{y}$ tahmin edilen değeri gösterirse; (Stein, C. M., 1960)

- Tahmin edilen istatistiğin beklenen değeri bilinmeyen anakütle parametresine eşitse, buna yansız (unbiased) tahmin denilmektedir:

$$E(\hat{\theta}) = \theta$$

- Diğer tekniklerle elde edilen sonuçlarla kıyaslandığında minimum varyansa sahip tahmine etkili tahmin denilmektedir:

$$\text{Var}(\hat{\theta}_1) < \text{Var}(\hat{\theta}_2) < \text{Var}(\hat{\theta}_3) \dots \text{Var}(\hat{\theta}_k) \text{ ise}$$

$\hat{\theta}_1$, θ 'nın en etkin tahmincisi olur.

- Tahmin, örnek terimlerinin doğrusal bir fonksiyonu ise bu tahmine doğrusal tahmin denilmektedir.

$$\bar{Y} = \frac{\sum Y}{N} = \left(\frac{1}{N} \right) (Y_1 + Y_2 + \dots) = \frac{Y_1}{N} + \frac{Y_2}{N} + \dots \quad (3.15)$$

Burada $\bar{Y}$, sırasıyla $Y_1, Y_2, \dots, Y_n$ değerlerinin doğrusal bir fonksiyonu olmaktadır.

- Tahmin, örnek büyüklüğü artarken gerçek değerine yaklaşıyorsa tutarlıdır denilmektedir. (Stein, C. M., 1960)

$$\lim_{n \rightarrow \infty} P\{|\theta - \hat{\theta}| = 0\} = 1 \quad (3.14)$$

Bu varsayımın birden çok bağımsız değişkeni olan regresyon modelleri için geçerlidir. Değişkenler arasındaki ilişkinin derecesi korelasyon katsayısı ile ifade edilir. Korelasyon katsayısı 0 ile 1 arasında değerler alır. Korelasyon katsayısının 0 değerini alması değişkenler arasında bir ilişkinin olmadığını, 1 değerini alması ise değişkenler arasında güçlü bir ilişkinin varlığını gösterir. (Tatlidil Hüseyin, 1996)

4. ÇOKLU BAĞLANTI PROBLEMİ

Bir regresyon modelindeki tahmin ediciler bağımsız değişkenler olarak adlandırılırlar. Fakat bu terim, elimizdeki tahmin edici değişkenlerin istatistiki manada birbirinden bağımsız olduğu anlamına gelmemektedir. Normalde tahmin ediciler birbirleriyle yüksek ilişkili olabilirler. Bir regresyon modelinin amacı veri setinden en uygun ve etkin tahminleri seçerek bir model oluşturmaktır. Eğer değişkenler arasında ilişki yoksa bu değişkenlere ortogonal denir. Regresyonda bazı değişkenler arasında yüksek ilişki olduğu zaman ise bu duruma çoklu bağlantı sorunu denir. “Çoklu bağlantı” terimi tahmin ediciler arasında yüksek ilişki olduğu durumlarda kullanılır. Çoklu bağlantı regresyon modelini geçersiz kılmaz. Ama regresyon modelinin faydasına etkisi düşünüldüğünde ciddi bir problemdir. (David Birkes, Y. Dodge, 1993) Aynı çok değişkenli uzayda denklemin kullanımını ayarlamak için tahmin edici kombinasyonlarına dayanması şartıyla denklemin tahmin değeri iyi olabilir. Fakat çoklu bağlantının çeşitli negatif etkileri vardır. Birincisi, regresyon katsayısının varyansı büyütülmüş olabilir. Aynı zamanda bu regresyon katsayıları istatistiksel olarak anlamlı değildirler. İkincisi, göreceli/nispi önem ve katsayıların işaretleriyle beraber yorumlanamayabilir. Üçüncüsü, regresyon katsayılarının değeri tahmin edici denklemlerin denkleme eklenmesi ya da çıkartılması ile kökten değiştirilebilir. Aslında regresyon katsayısının önemini bağlayıcıdır. (Marquardt, D.W. , Snee, Ronald D. 1975)

Regresyon analizinde çoklu bağlantı aşağıdaki problemlere yol açmaktadır:

1. Tam çoklu doğrusal bağlantı durumunda regresyon katsayıları belirsiz ve bu katsayıların standart hataları büyük olmaktadır.
2. Çoklu bağlantı halinde regresyon katsayılarının varyans ve kovaryansları artmaktadır.
3. Modelin R^2 açıklama değeri yüksek, ancak bağımsız değişkenlerden hiç biri veya çok azı kısmi t testine göre anlamlıdır.
4. İlgili bağımsız değişkenlerin bağımlı değişkenle olan ilişkilerinin yönü kuramsal ve ampirik beklentilerle çelişmektedir. (Vijay Mahajan, Arun K. Jain, M. B. , 1986)

4.1 Çoklu Bağlantının Saptanması

Çoklu bağlantının olup olmadığını aşağıdaki belirleme yöntemlerinde görüldüğü gibi 10 başlığa ayrılabilir:

1. İkişer ikişer tahmin edici değişkenler arasındaki yüksek korelasyon olduğunda,
2. Regresyon katsayılarının önemi ya da miktarının fiziksel anlamda iyi olmadığı durumda,
3. Önemli tahmin ediciler üzerinde, istatistiksel olarak önemli olmayan regresyon katsayıları olduğunda,
4. Regresyon katsayılarının önemini ve miktarının değişmesi, tahmin edici değişkenlerin silinmesi ya da eklenmesi gerektiği durumlarda ,
5. VIFs değerinin en büyüğü 10'un üstünde olduğu durumlarda,
6. Hoerl ve Kennard'ın (1970) önerdiği ridge izinde katsayıların grafiğinde dalgalanmalar olduğunda,
7. En büyük özdeğerin en küçük özdeğere bölümü 30'un üzerinde olduğunda,
8. Standartlaştırılmış matrisinin determinantı sıfır veya sıfıra çok yakın olduğunda,
9. j'nci bağımsız değişkenin çoklu belirlilik katsayısı değeri 1'e yakın olduğu durumlarda,
10. Özdeğerlerinin terslerinin toplamı bağımsız değişken sayısından çok büyük,

olduğunda çoklu bağlantının önemli olduğunu söylenebilir. (Neter, J. , W. Wesserman, M. H. , Kutner Terry Sincich, 1996)

4.2 Çoklu Bağlantı Probleminin Çözümü

Çoklu bağlantı probleminin çözüm yolları aşağıda verilmektedir:

- Bir ya da daha çok bağımsız değişken modelden çıkartılmaktadır. Fakat hangi değişkenler çıkartılacaktır? Böyle bir yaklaşım modeli yanlış tanımlamaya götürebilmektedir. (Vijay Mahajan, Arun K. Jain, M. B. , 1986)
- Farklar alınarak değişkenler dönüştürülebilir. Fakat böyle bir dönüşüm hatalar arasında otokorelasyon problemine yol açabilir. Ayrıca böyle bir yöntem sadece zaman serilerine uygulanırken, kesit verilerine uygulanmamaktadır. (Obeuchain, R. L., 1975)
- Bazen yeni gözlem değerlerinin elde edilmesi ile çoklu bağlantı problemi ortadan

kaldırılabilir. Fakat her zaman örneği büyütme mümkün olmamaktadır. (Vijay Mahajan, Arun K. Jain, M. B. , 1986)

- Birbiriyle ilişkili iki değişken yerine bu iki değişkenin toplamı (tek bir değişken olarak) alınmaktadır. (Hongehang Hu, 2005)
- Birbirinden bağımsız bileşenler türeten “Temel Bileşenler Regresyon Analizi” veya EKK tekniğinin düzeltilmiş şekli olan ve yanlış standartlaştırılmış regresyon katsayılarını tahmin eden “Ridge Regresyon” teknikleri kullanılmaktadır. (Obeuchain, R. L. , 1975)

Özetle çoklu bağlantı sorunu olduğunda elimizdeki tahmin edici yansız ama varyansı büyük bir değer olacaktır. Bunu önlemek için küçük bir yansız ama küçük varyanslı yanlış tahmin edicilerden “Ridge Regresyon” ve “Temel Bileşenler Regresyon” kullanılabilir. (Hongehang Hu, 2005)

4.2.1 Varyans Büyütme Faktörü VIF (Variance Inflation Factor)

VIF değeri istatistiki olarak bir matristeki tahmin edici değişkenlerdeki çoklu bağlantının tanımlanmasında kullanılır. VIF değeri, tahmin edilen regresyon katsayısının varyansındaki çoklu bağlantının etkisini gösterir. Çoklu bağlantı, sadece tahmin ediciler arasındaki ikili korelasyonlara değil, aynı zamanda bir tahmin ediciden bir diğer tahmin ediciye çok değişkenli tahmin edilebilirliğine de bağlıdır. Dolayısıyla VIF, çok değişkenli doğrusal regresyonda her bir tahmin edicinin belirlilik katsayısına (R^2) dayanır. (Jürgen Grob, 2003)

$$VIF_i = \frac{1}{1 - R_i^2}$$

Burada R_i^2 i. tahmin edicinin diğer bütün tahmin ediciler üzerindeki belirlilik katsayısıdır. Eğer i. tahmin edici diğer tahmin edicilerden bağımsız ise VIF_i bir tanedir. Eğer i. tahmin edici diğer tahmin ediciler tarafından mükemmel yakın bir şekilde tahmin edilebilirse, bu durumda VIF_i yaklaşımı sonsuzdur. Bu durumda belirlilik katsayısı sınırsızdır. (Neter, J. , W. Wesserman, M. H. Kutner Terry Sincich, 1996)

Bir ya da daha fazla VIF_i değeri büyük değerlere sahip olduğu zaman, bu durumda çoklu bağlantı sorunu olduğu söylenir. Özel bir karar nasıl büyük olarak görülür. Bazı

araştırmacılar kritik değer olarak VIF_i 'nin %5'lik ya da %10'luk kritik başlangıç değerini kullanırlar. VIF_i değerleri her biri ayrı ayrı olarak R_i^2 0,80 ve 0,90 değerleri için uygun bir değerdir. Bazı hesaplamalarda, bütün tahmin ediciler için ortalama VIF_i değeri önemli miktarda büyükse bir tane belirtilmiş çoklu bağlantıyı ifade eder. Durum ne olursa olsun önemli olan çoklu bağlantı için tahmin ediciler arasında güçlü ilişki olmalıdır. Sadece aralarında güçlü ilişki varsa çoklu bağlantı vardır anlamına gelmez. (Vijay Mahajan, Arun K. Jain, M. B. , 1986)

VIF_i yine bir istatistik terimi olan tolerance (dayanıklılık) ile yakından ilişkilidir ve tolerance $1 / VIF_i$ olarak gösterilir. İstatistik paket programlarının bazıları VIF_i 'yi, bazıları ise tolerance'yi raporlarında verirler. (Anderson, T. W. , 2003)

Eğer X_j diğer değişkenlerle doğrusal bileşen ise o zaman $R_i^2=1$ ve $VIF_i \rightarrow \infty$ 'dur. Eğer X_j diğer değişkenlerle tam olarak ilişkisizse o zaman $R_i^2=0$ ve $VIF_i=1$ 'dir.

X_j 'nin VIF_i değeri ile X_j 'nin regresyon katsayısının en küçük kareler tahmin edicisinin standart sapması arasında doğrudan bir bağlantı vardır.:

$$SD(\hat{\beta}_j) = \sqrt{VIF_j} \left(\frac{\sigma}{s_j} \right) / \sqrt{n-1} \quad (4.2)$$

Burada $\hat{\beta}_j = \hat{\beta}_{LS_j}$ ve s_j , X_j 'nin gözlenen değerinin standart sapmasıdır. (Anderson, T. W. , 2003)

5. RIDGE REGRESYON

Çoklu doğrusal regresyon probleminde bağımsız değişkenler, deney düzenlemesi yardımı olmadan verilerin toplanması veya sadece deney düzenleme yoluyla veriler toplansa dahi bu kez de olayın yapısındaki fiziksel ve matematiksel kısıtlar nedeniyle birbirleriyle ilişki hâlinde olabilirler. Regresyonda çoklu bağlantının ortaya çıkması ile problemdeki sebep-sonuç ilişkisini ortaya koyan parametrelerin tahmininden duyarlı sonuçlar elde edilemez. (Hongehang Hu, 2005)

Korelasyon matrisinin birim matrise yakın olması durumunda EKK yöntemi güvenilir sonuçlar vermektedir. Ancak korelasyon matrisinin birim matris olmaktan uzaklaşması, EKK tahminlerinin VIF değerlerinin de büyümesine neden olmakta ve dolayısıyla parametre tahminlerinin hatalarını artırmaktadır. (Hoerl, A.E., KENNARD, R.W. , 1970)

Bağımsız değişkenler arasındaki ilişkiler, EKK katsayı tahminlerin varyanslarını azaltmaktadır. Modeldeki her bir değişken için VIF değerleri; regresyon katsayılarının varyansları üzerindeki basit korelasyonların toplam etkisini gösterir. Çoklu bağlantı durumunda, korelasyon matrisinin tersinin köşegen elemanları VIF değerleri; her bir tahminin diğer tahminlerle olan çoklu korelasyonunda sonsuz hale gelir. Bu durumda EKK tahminleri yansız tahmin ediciler sınıfında en küçük varyanslı tahminler olma özelliklerini kaybederler. Çünkü çoklu bağlantı gerçek değerler arasında sapmaya neden olur. (Ryan, Thomas P. , 1997)

Çoklu bağlantının olması durumunda EKK yöntemiyle çözüm yapıldığında, parametre tahminlerinin Hata Kareler Toplamı (SSE), özdeğerlerden yararlanılarak hesaplanırsa, bir ya da daha fazla özdeğerin sıfır veya sıfıra yakın olması sapmaların beklenen değerini arttıracaktır. (Ferreira. Eurico J.,Sirmans. G. Stacy 1988)

Bu nedenle regresyon katsayılarının tahmini için, bağımsız değişkenlerin birbirleri üzerindeki etkilerini minimum yapmak ve kararlı katsayı tahminleri elde edebilmek için yanlı regresyon yöntemlerinden birisi olan Ridge Regresyon yöntemi kullanılmalıdır. (Ferreira. Eurico J.,Sirmans. G. Stacy 1988)

Ridge regresyon 1962 yılında bir kimya mühendisliği dergisine Arthur Hoerl tarafından yazılan bir makale ile girmiştir. (Ryan, Thomas P. , 1997)

Ridge tahmini çoklu bağlantı problemi ortaya çıktığında uygulanabilir. Eğer değişkenlerden biri diğer değişkenlerin doğrusal fonksiyonu ise bu değişken seti tam olarak bağımlıdır. Eğer değişkenlerden biri hemen hemen diğer değişkenlerin doğrusal fonksiyonu ise veri seti de

hemen hemen bağımlıdır. Eğer iki değişken yüksek derecede ilişkili ise aynı zaman da collinear (bağımlı) dir. (S. H. Ngo , S. Kemeny, A. Deak. 2003)

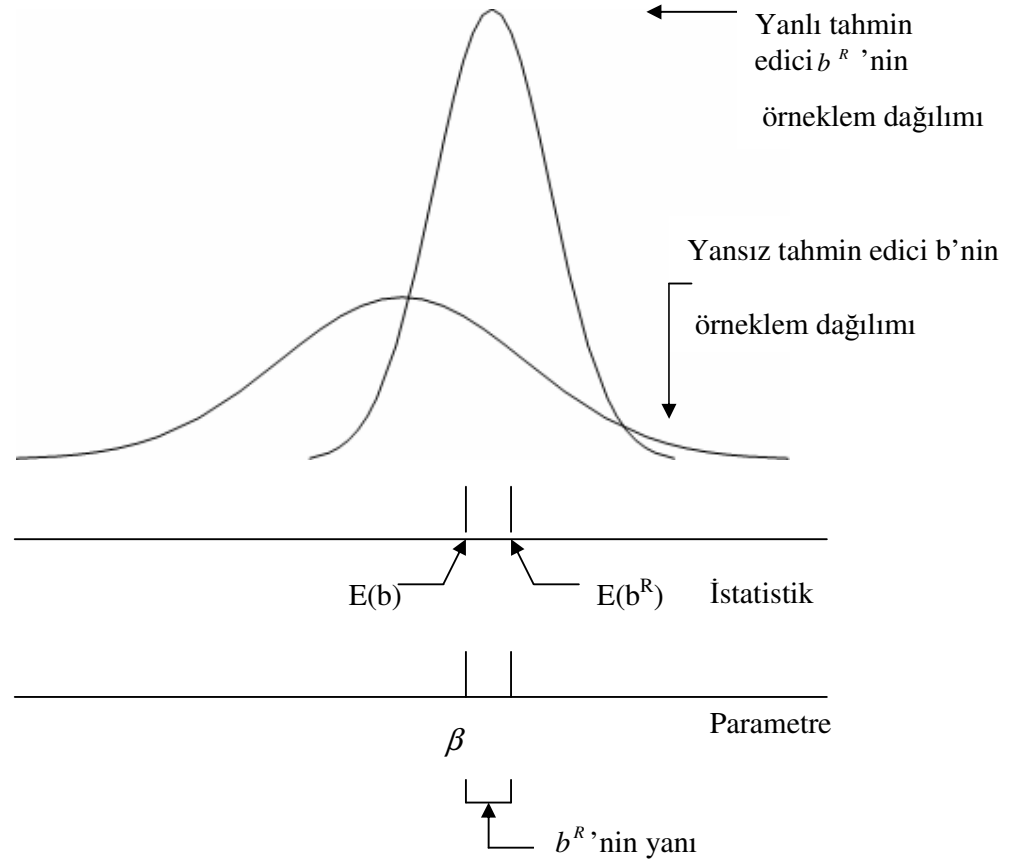
Yanlı tahmin, regresyon katsayısındaki kararlılığın artışı ile birlikte varyansta önemli bir azalmaya sebep olarak kullanılır. Katsayılar yansızdır. Varyanstaki azalma tahmin edicinin yanındaki azalmadan daha önemli ise o zaman başarılı bir tahmin yapılmış olur. (John Fox, 1991)

Ridge regresyonu etkili olarak kullanmanın çeşitli yolları vardır. Pek çok yaklaşımda regresyon sonuçlarında $X'X$ matrisinin özdeğerlerinin etkilerinde azalma meydana gelmektedir. Bilindiği gibi küçük özdeğer üretmek demek büyük varyans demektir. Bu da açıkça görüldüğü gibi $X'X$ matrisinin tersinin köşegen elemanları olan VIFs gibi değerlerle regresyon katsayısı tahmini çok etkisizleştiren bir yöntemdir. (Hoerl, A.E., KENNARD, R.W. , 1970)

Diğer yandan çoklu bağlantı durumunda köşegen elemanları durum üzerinde hakim değildirler. Köşegen elemanlarının duruma hakim olmadığı zamanlarda en az bir tane özdeğer küçüktür. (Ryan, Thomas P. , 1997)

5.1 Yanlı Tahminler

Ridge regresyon, çoklu bağlantı probleminin çözülmesi için en küçük kareler metodu ile bulunan regresyon katsayısının yanlı tahminini değiştirmek için önerilen çeşitli yöntemlerden birisidir. Bir tahmin edici sadece küçük bir yan, yansız bir tahmin ediciden daha doğru bir sonuç verir. Bu şekilde daha büyük bir olasılıkla doğru parametre değerine yaklaşılr. (S. H. Ngo , S. Kemeny, A. Deak. 2003)



Şekil 5.1 Yanlı ve yansız tahmin edici

Şekil 5.1’de de görüldüğü gibi b tahmin edicisi yansızdır. Fakat kesin olarak doğru bir tahmin değildir. Oysa ki b^R tahmin edicisinin doğruluğu daha kesindir ama küçük bir yana sahiptir. (Neter, J. , W. Wessarman, M. H. Kutner Terry Sincich, 1996)

Yanın (bias) ve örneklem değişikliğinin birleşik etkisinin ölçüsü, ortalama kare hatadır. Burada ortalama kare hata, β parametresinin doğru tahmin edicisi olan b^R yanlı tahmin edicisinin, standart sapmasının beklenen değeridir. Bu beklenen değer, tahmin edicinin varyansının ve yanın karesinin toplamıdır:

$$E\{b^R - \beta\}^2 = \sigma^2\{b^R\} + (E\{b^R\} - \beta)^2 \quad (5.1)$$

Sıralı en küçük kareler için verilen normal denklem:

$$(X'X)\beta = X'Y \quad (5.2)$$

dir. (Hoerl, A.E., KENNARD, R.W. , 1970)

Korelasyon dönüşümü ile bütün değişkenler dönüştürüldüğü zaman dönüştürülmüş regresyon denklemi:

$$Y_i' = \beta_1' X_{i1}' + \beta_2' X_{i2}' + \dots + \beta_{p-1}' X_{i,p-1}' + \varepsilon_i' \quad (5.3)$$

olarak verilir.

Ve en küçük kareler normal denklemi

$$r_{XX} b = r_{YX}$$

burada r_{XX} X değişkeninin korelasyon matrisi ve r_{YX} ise Y ile her bir X değişkeni arasındaki basit korelasyonun katsayılarının vektörüdür. (Hoerl, A.E., KENNARD, R.W. , 1976)

Standartlaştırılmış ridge regresyon tahmin edicileri sabit bir $k \geq 0$ ile en küçük kareler normal denkleminde (5.2)'den elde edilir.

$$(r_{XX} + kI) b^R = r_{YX} \quad (5.4)$$

Burada b^R , b_k^R standartlaştırılmış ridge regresyon katsayısının vektörüdür.

$$b_{(p-1) \times 1}^R = \begin{bmatrix} b_1^R \\ b_2^R \\ \vdots \\ b_{p-1}^R \end{bmatrix}$$

I (p-1)×(p-1)'lik birim matristir. (5.4)'nin çözümünden standartlaştırılmış regresyon katsayısı:

$$b^R = (r_{XX} + kI)^{-1} r_{YX} \quad (5.5)$$

olarak elde edilir. (Hoerl, A.E., KENNARD, R.W. , 1976)

Tahmin edicilerde yanın miktarı k sabitine yansımaktadır. $k=0$ olduğunda (5.4) denkleminde sıralı en küçük kareler tahmin edicisi azalır. $k>0$ olduğunda ridge regresyon katsayısı yanlı olur. Fakat sıralı en küçük kareler tahmincisiinden daha kararlı (stabil) bir eğilime sahiptir. (Ferreira. Eurico J., Sirmans. G. Stacy 1988)

5.1.1 Yanlı k Sabitinin Seçimi

Ridge regresyon tahmin edicisi b^R 'nin toplam ortalama kare hatasının yan bileşenleri k gittikçe artarken artmaktadır ve varyans bileşenleri küçülmektedir. Ridge regresyon tahmin edicisi b^R , sıralı en küçük kareler ortalaması tahmin edicisinden daha küçük bir toplam ortalama kare hataya sahiptir. Bu koşulu sağlayan daima bir k sabiti vardır. İşin zor kısmı ise optimal k değerini bulmaktır. (Hoerl, A.E., KENNARD, R.W. , 1976)

Yanlı k sabitinin bulunması için yaygın olarak kullanılan yöntemler ridge izi ve VIFs'dir. Ridge izi eş zamanlı olarak k'nın farklı değerleri için, genellikle 0 ile 1 arasındadır, p-1 tane tahmin edilmiş standartlaştırılmış ridge regresyon katsayısının değerleridir. Yapılan çalışmalar gösteriyor ki; k'nın hafifçe değişmesi için tahmin edilmiş regresyon katsayısı b_k^R genellikle değişir. Ridge izi ve VIFs değerleri için en küçük k değeri seçilir. Regresyon katsayısı seçiminde ridge izi değişmez ve VIFs değeri yeteri kadar küçük olmalıdır.

$$\begin{aligned}
 (1+k)b_1^R + r_{12}b_2^R + \dots + r_{1,p-1}b_{p-1}^R &= r_{Y1} \\
 r_{21}b_1^R + (1+k)b_2^R + \dots + r_{2,p-1}b_{p-1}^R &= r_{Y2} \\
 &\vdots \\
 r_{p-1,1}b_1^R + r_{p-1,2}b_2^R + \dots + (1+k)b_{p-1}^R &= r_{Y,p-1}
 \end{aligned} \tag{5.6}$$

Burada r_{ij} i. ve j. X değişkenleri arasındaki basit korelasyon katsayısıdır. r_{Yj} ise Y değişkeni ile j. X değişkeni arasındaki basit korelasyon katsayısıdır. (Hoerl, A.E., KENNARD, R.W. , 2000)

b_k^R ridge regresyon katsayısının VIF değeri sıralı en küçük kareler regresyon katsayısınınkinden benzerlik göstermektedir. Yani b_k^R 'nin VIF değerinin ölçüsü eğer tahmin edici değişkenlerle ilişkisiz ise, b_k^R 'nin varyansının ne kadar büyük olduğu ile ilişkilidir. Bu da gösteriyor ki VIF değerleri (p-1)×(p-1) boyutlu matrisin köşegen elemanlarıdır:

$$(r_{XX} + kI)^{-1} r_{XX} (r_{XX} + kI)^{-1} \tag{5.7}$$

En küçük kareler için verilen R^2 çoklu belirlilik katsayısı:

$R^2 = 1 - \frac{SSE}{SSTO}$ şeklindedir.

Ridge regresyon için de benzer şekilde tanımlanır. Olaylar basitleştirilirse, Y bağımlı değişkeninin korelasyon dönüşümü için kareler toplamının toplamı :

$$SSTO_R = \sum (Y_i' - \bar{Y}')^2 = 1 \quad (5.8)$$

dır. (Ryan, Thomas P. , 1997)

Ridge regresyon ile uygun değerler

$$\hat{Y}_i' = b_1^R X_{i1}' + \dots + b_{p-1}^R X_{i,p-1}' \quad (5.9)$$

burada X_{ik}' X değişkenlerinin korelasyon dönüşümüne göre dönüştürülmüş halleridir. Hata kareler toplamı ise:

$$SSE_R = \sum (Y_i' - \hat{Y}_i')^2 \quad (5.10)$$

bu toplamda $\hat{Y}_i'$ (5.7)'de verilen denklemdir. Buradan ise ridge regresyonun R^2 'si

$$R_R^2 = 1 - SSE_R \quad (5.11)$$

olur. (John Fox, 1991)

Ridge regresyon tahmini, veriler üzerindeki ufak değişikliklerde kararlı olma eğilimindedir. En küçük kareler tahmincisi ise bu durumun tersine tahmin ediciler arasında çoklu bağlantı sorunu yüksek olduğu zaman yüksek düzeyde değişiklik göstermektedir. (Ryan, Thomas P. , 1997)

Ridge regresyonun büyük kısıtlamaları vardır, sıralı çıkarsama usulleri uygulanabilir değildir ve tam dağılım özellikleri bilinmiyordur. Ridge regresyonun diğer bir kısıtlaması ise yanlı k sabit değerinin tartışılabilir olmasıdır. (John Fox, 1991)

Ridge regresyon ridge izi analizi ile gözleme dayalı araştırma çalışmalarında potansiyel tahmin edici değişkenlerin sayılarını azaltmak için yardımcı olabilir. İki durum vardır. Değişkenlerin ridge izi sabit değildir, katsayıların eğilimi sıfır çevresindedir. Değişkenlerin ridge izi sabittir fakat çok küçük bir değerde düşer. Sonuç olarak değişkenlik göstermeyen ridge izi ile değişkenler sıfır çevresinde bir eğilime sahip değildirler. (David Birkes, Y. Dodge, 1993)

5.2 Genelleştirilmiş Ridge Regresyon Denklemi

$X'X = G'\Lambda G$ burada G bir ortogonal matris ve Λ ise $\lambda_1, \dots, \lambda_p$ öz değerlerini içeren bir köşegen matristir. Ridge parametresinin $K = \text{diag}(k_1, \dots, k_p)$ bir köşegen matrisidir. (Hongehang Hu, 2005)

$$\hat{\beta}^R = (X'X + GKG')^{-1} X'y \quad (5.12)$$

5.3 Standartlaştırma

Açıklayıcı değişkenler ridge tahmini yapılmadan önce genellikle standartlaştırılırlar. Ridge tahmininin hesaplanmasında, bütün açıklayıcı değişkenler, aynı şekilde işlenirler. Bu nedenle değişkenlerin standardize edilmiş değişkenler olabilmesi için, ölçümün onların birimlerinde karşılaştırılabilir olması gerekmektedir. (Ferreira. Eurico J., Sirmans. G. Stacy 1988)

Bazı istatistikçiler en küçük kareler regresyonu yapmadan önce de standartlaştırmayı tavsiye etmektedirler. Standartlaştırmanın 3 yolu vardır. Birincisi, $X'X$ matrisinin tersini alarak hatayı azaltır. İkincisi, regresyon katsayılarını ve değişkenlerin değerlerinin artırır. Üçüncüsü standardizasyon, direk olarak farklı değişkenlerin regresyon katsayıları ile karşılaştırılabilmesini sağlamaktadır. Örnek olarak $\hat{Y} = -171 + 1,92X_1 + 0,286X_2$ $Y =$ Santilitre olarak akciğerin hacmi $X_1 =$ santilitrede yükseklik ve $X_2 =$ santilitrede ağırlık olarak regresyon denkleminde bir denklem alınır. 1,92 ve 0,286 olan katsayıları karşılaştırılmaz. Açıklayıcı değişkenler standartlaştırılır. Standartlaştırılan değişkenlere Z_1 ve Z_2 denilirse ve regresyon denklemindeki tahminleri $\hat{Y} = 193 + 12,9Z_1 + 3,28Z_2$ olur. Şimdi 12,9 ve 3,28'i karşılaştırılır ve sonuçlandırıldığında akciğer kapasitesindeki farklılıklarda, yükseklikteki farklılığın ağırlıktaki farklılıktan daha fazla etkili olduğundan kaynakladığını görülmektedir. (Draper, N.R., H. Smith, 1981)

X değişkeni standardize edildiğinde, $z_i = (x_i - \bar{x})/s_x$ olur burada $\bar{x}, x_i$ 'lerin ortalaması, s_x ise x_i 'lerin standart sapmasıdır.

Standardize edilmemiş X basit doğrusal regresyonda $Y = \alpha + \beta X + \varepsilon$ olarak gösterilirken; standardize edildikten sonra $Y = \mu + \gamma Z + \varepsilon$ olarak gösterilir. (Hoerl, A.E. , KENNARD, R.W. , 2000)

5.4 Ridge Tahminleri

μ ve γ 'nın en küçük kareler tahminleri formülünden elde edilebilir. z_i standardize edildiğinde, $\bar{z} = 0$ olur ve bu nedenle formüller basitleştirildiğinde:

$$\hat{\gamma}_{LS_I} = \frac{\sum z_i y_i}{\sum z_i^2} \quad (5.13)$$

$$\hat{\mu}_{LS_I} = \bar{y}$$

Ridge tahminlerinde $\hat{\mu} = \bar{y}$ 'dir. Ve

$$\hat{\gamma} = \frac{\sum z_i y_i}{\sum z_i^2 + k} \quad (5.14)$$

burada $k = \hat{\sigma}_{LS}^2 / \hat{\gamma}_{LS}^2$ 'dir ve $\hat{\sigma}_{LS}^2$ σ 'nın en küçük kareler tahminidir. (S. H. Ngo , S. Kemeny, A. Deak. , 2003)

5.5 Ridge Varyans

$$\text{Yan}(\hat{\beta}^R) = -k(X'X + kI)^{-1} \beta \quad (5.15)$$

$$\text{Var}(\hat{\beta}^R) = \sigma^2 (X'X + kI)^{-1} X'X (X'X + kI)^{-1} \quad (5.16)$$

$$\text{Var}(\beta) = \sigma^2 (X'X)^{-1} \text{ En küçük kareler tahmin edicisinin varyansı} \quad (5.17)$$

(John Fox, 1991)

5.6 Ridge Regresyonun Ortalama Kare Hatası

Eğer $\hat{\gamma}$ ile γ arasındaki mesafe küçük ise $\hat{\gamma}$ 'nın doğru bir tahmin edici olduğu söylenebilir. Bu fikrin sayısal hale getirilmesindeki genel yol $\hat{\gamma}$ tahmin edicisinin ortalama kare hatasının miktarını bularak $\text{MSE}(\hat{\gamma}) = E[(\hat{\gamma} - \gamma)^2]$ $\hat{\gamma}$ ile γ arasındaki mesafenin karesinin beklenen değerini almaktır. Ve $\text{MSE}(\hat{\gamma})$ en küçük ise o zaman $\hat{\gamma}$ en doğru tahmindir denilir. Bu ifadenin daha pratik bir hali

$$\text{MSE}(\hat{\gamma}) = \text{Var}(\hat{\gamma}) + [E(\hat{\gamma}) - \gamma]^2 \quad (5.18)$$

ya da

$$MSE(z_R) = \text{Variance}(z_R) + \text{Bias}(z_R) \quad (5.19)$$

Eğer $k < k_{\max} = 2\sigma^2 / \beta' \beta$ ise $MSE(\hat{\beta}) - MSE(\beta) < 0$

negatif tanımlıdır. . (Vinod, H.D., 1976)

Bu ifade $\hat{\gamma}$ 'nın ortalama kare hatası varyansı artı yanının karesidir. Bu nedenle ortalama kare hatası küçük bir tahmin için, aynı anda küçük varyansa ve küçük yana sahip olmalıdır. En küçük kareler tahmini $\hat{\gamma}_{LS}$ yansızdır $E(\hat{\gamma}_{LS}) = \gamma$. Ridge tahmini yanlıdır. Ancak daha küçük varyansa sahip olma eğilimindedir. Varyansın azalışından daha az bir şekilde yanın karesinin artması umulur, bu nedenle ridge tahmini daha küçük ortalama kare hataya sahiptir. (Vinod, H.D., 1976)

5.6.1 5.14 denkleminin gerekçesi

Ridge regresyon hakkında düşünmenin tek yolu $\hat{\gamma}_{LS}$ değerinin doğruluğunu ve azalan bir değer olduğunu göstermektir. Azalan $\hat{\gamma}_{LS}$ değeri 0 ile 1 arasında değişen c gibi bir değer ile çarpımına eşittir $\hat{\gamma}_{LS} = c \cdot \hat{\gamma}_{LS}$. Buna “shrinking” yani daralma denir. Çünkü $|c \hat{\gamma}_{LS}| = |\hat{\gamma}_{LS}|$ ’dir.

$\hat{\gamma}_{LS}$ ’nin ortalama kare hatası yani sıfır olduğu için kendisinin varyansına eşittir $v = \sigma^2 / \sum z_i^2$.

$c \hat{\gamma}_{LS}$ ’nin ortalama kare hatası

$$MSE(c \hat{\gamma}_{LS}) = c^2 v + (c - 1)^2 \gamma^2 \text{’dir.} \quad (5.20)$$

c ’nin türetilmesi ile ilgili 0’a eşit olması için c çözülür ve en küçük ortalama kare hata elde edilir $c = \gamma^2 / (\gamma^2 + v)$. Bu ifade c için, bilinmeyen γ ve σ parametrelerinin $\hat{\gamma}_{LS}$ ve $\hat{\sigma}_{LS}$ olarak değiştirilerek $\hat{c}$ ’nin elde edilir. Küçük bir cebir uygulaması yapıldığında $\hat{c} = \sum z_i^2 / (\sum z_i^2 + \hat{\sigma}_{LS}^2 / \hat{\gamma}_{LS}^2)$ ifadesini bulunur ve ridge tahminiyle $c \hat{\gamma}_{LS}$ elde edilir. (Ferreira, Eurico J., Sirmans. G. Stacy 1988)

5.7 Ridge Parametresinin Optimal Değeri

En küçük kareler tahmin edicisinin optimal değerine doğrusal yansız tahmin ediciler sınıfındaki minimum varyanstan ulaşılır. Çünkü tahmin edici yansızdır ve varyansın minimum olması ortalama kare hatanın da minimum olmasını sağlar. Bu durumda varyans

$E[\hat{\beta}_p - E(\hat{\beta}_p)]^2$ ve aynı şekilde ortalama kare hata ise $E[\hat{\beta}_p - \beta_p]^2$ şeklindedir. (Obeuchain, R. L. , 1975)

Fakat ridge tahmin edicisi yanlıdır. Bu durum en küçük kareler tahmin edicisinin doğrusal bir dönüşümü olan ridge tahmin edicisinde kolayca görülmektedir. Aşağıda görüldüğü gibi:

$$\begin{aligned}\beta^R &= (X'X + kI)^{-1} x'y \\ &= (X'X + kI)^{-1} x'x\hat{\beta}\end{aligned}\tag{5.21}$$

denklemdaki $\hat{\beta}$ en küçük kareler tahmin edicisi ile elde edilir.

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

$X'X$ terimine kI terimini bir ekleyip bir çıkarırsak ;

$$\begin{aligned}\beta^R &= (X'X + kI)^{-1} [(X'X + kI) - kI] \hat{\beta} \\ &= [I - k(X'X + kI)^{-1}] \hat{\beta}\end{aligned}\tag{5.22}$$

ve β^R ' in beklenen değeri;

$$E(\beta^R) = [I - k(X'X + kI)^{-1}] \beta\tag{5.23}$$

olur. (Obeuchain, R. L. , 1975)

Bu nedenle yan;

$$E(\beta^R) - \beta = -k(X'X + kI)^{-1} \beta\tag{5.24}$$

yanlı bir tahmin edici olduğu zaman örneğin β^R gibi, en uygun optimal değer ortalama kare hatadır. β^R ' in ortalama kare hatasını incelemek için regresyon modelinin kanonik formunun yeniden parametre edilmesi gerekir. (Jürgen Grob, 2003)

$$y = X\beta + u = Z\Theta + u\tag{5.25}$$

burada

$$Z = PX$$

$$\Theta = P' \beta$$

ve

$$P' P = PP' = I$$

$$p' = p$$

olarak tanımlanır.

P matrisi $X'X$ vektörünün normalize edilmiş karakteristik matrisidir. Örneğin;

$$Z'Z = P'X'XP = D = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_p)$$

Burada $\lambda_1, \lambda_2, \dots, \lambda_p$ $X'X$ 'nin P tane karakteristik köküdür.

Kanonik formda modelin en küçük kareler tahmin edicisi;

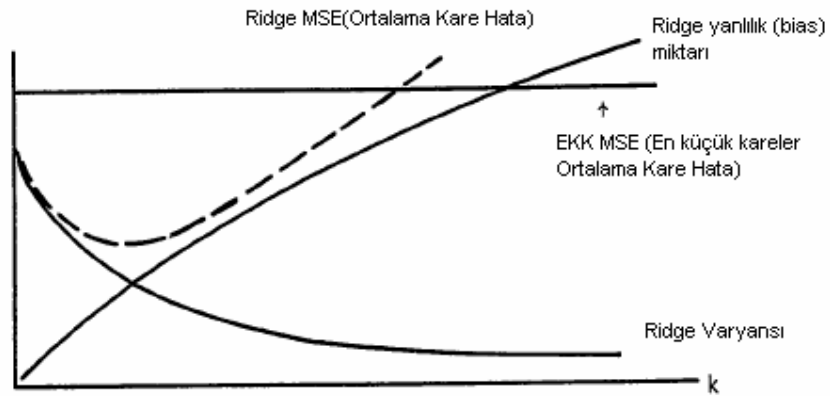
$$\hat{\Theta} = (Z'Z)^{-1}$$

ve $\hat{\beta}$ 'nin doğrusal dönüşümü

$$\begin{aligned} \hat{\Theta} &= (Z'Z)^{-1} Z' y = (P' X' X P)^{-1} P' X' y \\ &= (P^{-1} X' X P)^{-1} P' X' y = P^{-1} (X' X)^{-1} P P' X' y \\ &= P' \hat{\beta} \end{aligned} \quad (5.26)$$

şeklindedir. (Obeuchain, R. L. , 1975)

5.8 Varyans, Yan ve k (ridge parametresi) Arasındaki İlişki

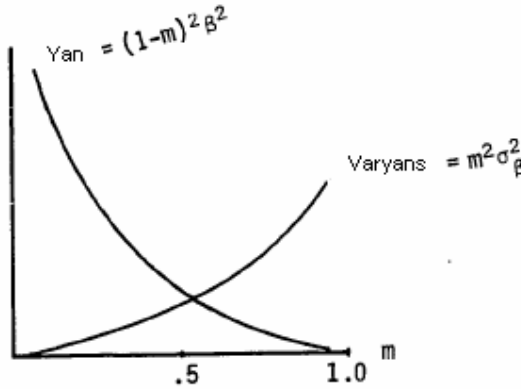


Şekil 5.2 Ridge parametreleri grafiği

Grafikte görüldüğü gibi en küçük kareler ortalama kare hatası k 'yı etkilemeyen yatay bir çizgi olarak görülmektedir. Diğer iki kesiksiz çizgi ise ridge sürecinin yanlılığı ve varyansını göstermektedir. $k > 0$ olarak kabul edildiğinde küçük bir yan ve azalan varyansa sahip tahmin ediciler elde edilir. (Marquardt, D.W. , 1970)

β 'nın m katsayısının bir fonksiyonu olarak $\hat{\beta}$ 'nın ortalama kare hatası, için yan ve varyansa katkısı aşağıdaki şekilde görülmektedir. (Marquardt, D.W. , 1970)

m =sabit terim



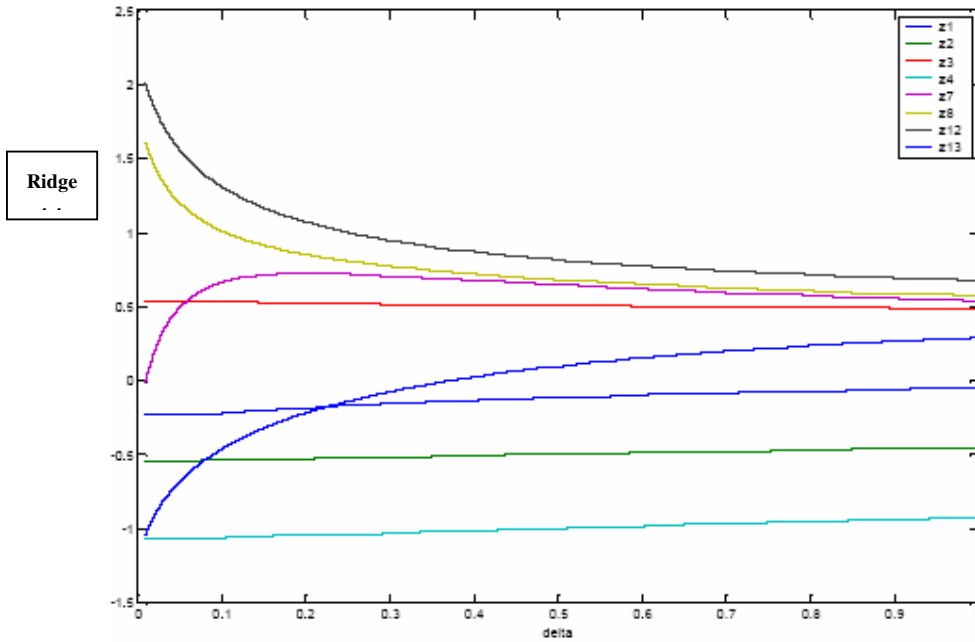
Şekil 5.3 Ridge yan-varyans grafiği

5.9 Ridge İzi

Ridge regresyonun diğer yöntemlerden farkı k değerinin seçimidir ve varyans terimi yanın karesinin artışından daha az bir şekilde artmaktadır. Hoerl ve Kennard k 'nın değerinin sıfırdan farklı olmasının en küçük kareler tahmincisinin varyansından daha az bir MSE'ye sahip olacağını savunmuşlardır. (Marquardt, D.W. , 1970) k 'nın değerine karar verme ile ilgili farklı söylemler vardır. Ridge izi Hoerl ve Kennard tarafından k değerinin uygun bir şekilde karar verilmesi için kontrol yöntemi olarak belirlenmiştir. Ridge izindeki k değeri $[0,1]$ arasında değişmektedir. Eğer bir veri setinde çoklu bağlantı varsa, regresyon katsayıları izde kararlı olmayacaklardır. k 'nın bazı değerlerinde ridge tahmini kararlıdır. k 'nın en küçük olarak seçilmesinin gayesi yanlı parametrenin kararlı olmasıdır. (Pierre Druilhet, Alain Mom , 2006)

Çoklu doğrusal regresyon problemlerinde çoklu bağlantı olduğunda katsayı tahminleri duyarlıdır. Yani veri kümesine birkaç gözlemin ilave edilmesiyle bu tahmin edicilerde değişikliklerin olduğu görülür. Böyle durumlarda regresyon katsayıları genellikle kararsız

katsayılar olarak bulunur. Bu kararsızlıkları izleyebilmek ve çoklu bağlantının etkisini açıkça görebilmek için grafiksel anlatım olan Ridge izinden yararlanılır. (Raymond H. Myers, 1990) Ridge regresyonun grafiksel bir gösterimi olan Ridge izi, regresyon katsayıları, katsayılar düşey ekseninde, k^* değerleri yatay ekseninde olacak şekilde iki boyutlu uzayda grafik elde edilir. (Marquardt ve Snee 1975) Ridge izi araştırmacıya hangi katsayıların verilere duyarlı olduğu konusunda yardımcı olur. Ridge izi karşı gelen her bir katsayı değerinin grafiğidir. Her bir katsayı için bir eğri ya da iz oluşur. Ridge izinde amaç EKK'dan daha küçük MSE'yi veren k^* değerini bulmak ve kararlı katsayılar kümesini oluşturmaktır. Kararlı katsayılar kümesinin anlamı; verideki küçük değişikliklere karşı katsayıların duyarlı olmamasıdır. Eğer bağımsız değişkenler yüksek ilişkili iseler k^* 'nın küçük değerleri için katsayılar çok hızlı değişecek ve k^* 'nın daha büyük değerlerinde derece derece kararlı olacaktır. Katsayıların kararlı olduğu k^* değeri katsayıların istenen kümesini verecektir. (Obeuchain, R. L. , (1975).)



Şekil 5.4 Ridge izi grafiği

5.10 Ridge Parametresinin Seçimi

Hoerl ve Kennard (1970) k^* 'nın tek bir değer olmadığını ancak k^* 'dan daha iyi olan k^* 'in her

zaman bulunabileceğini ifade etmişlerdir. Ridge parametresi k^* 'nın seçimine ilişkin çok sayıda teknik olmasına karşın, burada birkaç tanesi verilmiştir.

1. Hoerl ve Kennard (1976); k^* 'nın seçimini ridge izinden elde etmişlerdir.
2. Marquardt ve Snee (1975), VIF'in 1 ile 10 değerleri arasında olduğunda k^* 'nın seçilebileceğini belirtmişlerdir.
3. Hoerl, Kennard (1976); k^* 'nın seçimi için aşağıdaki eşitliği önermişlerdir.

$$k = p\hat{\sigma}^2 / \hat{\beta}'\hat{\beta} \quad (5.27)$$

4. Vinod (1976) ridge izinin dikey skalasını değiştirmeyi önermiştir. Eğer regresyon katsayılarına karşılık m çizilirse;

$$m = p - \sum_{i=1}^p \lambda_i / (\lambda_i + k) \quad (5.28)$$

daha kararlı bir yapı olacağını öne sürmüştür. Aynı zamanda Vinod m -ölçeği olarak bilinen, “Index of Stability of Relative Magnitudes” (ISRM)’yi geliştirmiştir. Bu katsayı $m < p$ için tanımlanır ve

$$\text{ISRM} = \sum_{i=1}^p [(p\lambda_i / (\lambda_i + k)^2 \bar{S}) - 1]^2 \text{ şeklindedir.} \quad (5.27)$$

Burada

$$\bar{S} = \sum_{i=1}^p \lambda_i / (\lambda_i + k)^2 \quad (5.28)$$

olarak gösterilir.

Bu nedenle her bir m (ve bu nedenle k) için ISRM hesaplanır ve m değeri için hesaplanan en küçük ISRM değeri, en kararlı değer olarak bulunur. (Vinod, 1976)

5. Obenchain (1975) k^* 'nın değerine karar vermenin iki yolu olduğunu belirtmiştir. Birincisi y 'nin tahmin edicisi için Xb^R 'nin doğruluğunun sorgulanması ve diğer ölçüte geçebilmek için ilgili verilerin minimumlarından b^R 'nin hesaplanmasıdır. Bu ölçüt prediction oriented criterion (yöneltmiş tahmin kriteri) olarak da adlandırılır ve k 'nın en küçük değerine eğilimli sonuçlar elde edilir. İkinci ölçüt ise control oriented criterion (yöneltmiş kontrol kriteri) olarak adlandırılır ve k 'nın en büyük değerine eğilimli

sonuçlar verir.

5.11 Ridge Regresyonun Yapılan Diğer Çalışmalarla İlişkisi

Ridge regresyonun diğer regresyon analizleri yaklaşımı ile aynı hedefleri paylaşan temas noktası / noktaları vardır. Üç yaklaşımdan söz edilebilir.

- Stein ve James (1960,1961) $0 \leq C < 1$ $C\hat{\beta}$ formunda $\hat{\beta}$ 'ya bir dönüşüm uygulayarak, ortalama kare hatadaki gelişmeyi incelemişlerdir. Bu durum $\hat{\beta}$ vektörünün kısaltılmış halidir. Onlar $C > 0$ olduğunda daima bir sonuç bulabildiklerini ve hesaplayabildiklerini göstermişlerdir.
- Schlaifer ve Raiffa (1961) tarafından bulunan Bayesgil yaklaşım bunlardan birisidir. Bu yaklaşım ilerleyen konularda ayrıntılı olarak ele alınmıştır.
- Balakrishnan (2006) regresyon da yapay tahmin ediciyi bulmuştur. Model için β 'yı ele aldığımızda; β , yapay, C konveks setinde sınırlı, örtülü ve C'de bir hiper kürenin R yarıçapına sahiptir. Tahmin edici ölçütü minimum artık kareler ortalamasıdır. $\phi = (Y - XB)'(Y - XB)$ burada B minimum olarak verilen değerdir. $\hat{\beta}'\hat{\beta} \leq R^2$ kısıtı altında B, $\hat{\beta}$ olarak seçilir aksi halde B, $\hat{\beta}^R$ olarak ve k ise $(\hat{\beta}^R)'\hat{\beta}^R = R^2$ 'den seçilir.

5.12 Sıralı Ridge Tahmin Edicisi

Hoerl ve Kennard (2000) bütün k_i 'ler içerisinde tek bir k seçilmesini sağlayan başka bir yöntem önermişlerdir. Bu yöntem "sıralı ridge tahmin edicisi" olarak adlandırılmaktadır. Genelleştirilmiş ridge metodunu temel alır ve minimum ortalama kare hata değeri $k_i = \sigma^2 / \alpha_i^2$ olduğu zaman elde edilir. Bu yolla tek bir k değeri elde edilirse, sıralı bir ortalama bir tane kullanılması istenmez. Çünkü çok küçük bir α_i tahmin gücünü etkileyeceği gibi büyük bir k ve büyük bir yan demektir.

6. DOĞRUSAL OLMAYAN REGRESYON MODELİ

Bir çok doğrusal olmayan fonksiyonun logaritmik dönüşümü yapılarak doğrusal hale dönüştürülebilir. Bu durumda dönüştürülmüş fonksiyonun parametreleri, EKK tekniği ile tahmin edilebilmektedir. Doğrusal olmayan regresyon modelinin yapısı doğrusal regresyon modelinin yapısına benzemektedir. Doğrusal regresyon modeli aşağıdaki şekilde yazılabilir:

$$y_i = x_i \beta + \varepsilon_i . \quad (6.1)$$

Burada x_i , n tane gözlemin i. tahmin ediciler satır vektörüdür. β tahmin edilecek regresyon parametrelerinin vektörüdür. ε_i beklenen değeri 0 ve sabit varyansa sahip $\varepsilon_i \sim N(0, \sigma^2)$ hata terimidir. (Bates, Douglas. M. , 1988)

Doğrusal olmayan regresyon modeli ise aşağıdaki gibi yazılabilir:

$$y_i = f(\beta, x_i) + \varepsilon_i \quad i=1, \dots, n \quad (6.2)$$

Yukarıdaki denklemde;

$$X_i = \begin{bmatrix} X_{i1} \\ X_{i2} \\ \vdots \\ X_{iq} \end{bmatrix} \quad \text{ve} \quad \gamma = \begin{bmatrix} \gamma_0 \\ \gamma_1 \\ \vdots \\ \gamma_p \end{bmatrix}$$

ve Y_i gözlemi, doğrusal olmayan $f(X_i, \gamma)$ fonksiyonu ve ε_i hata terimlerinin toplamıdır. Hata terimlerinin doğrusal regresyon modelindeki gibi, beklenen değerlerinin sıfır, sabit varyanslı ve korelasyonsuz olduğu varsayılır. $f(X_i, \gamma)$ fonksiyonu ise, Y bağımlı değişkeninin, X bağımsız değişkenlerinin bir fonksiyonu olduğunu gösterir. (Goron K. Smyth, 2002)

3 bağımsız değişkenli bir denklem örneği aşağıdaki gibidir.

$$y_i = \beta_1 + \beta_2 [1 - \exp(\beta_3 x_i)] + \varepsilon_i \quad (6.3)$$

Bu modelde β üç boyutludur. Dikkat edilirse bu model β_3 parametresinde doğrusal değildir. Eğer $\beta_3 < 0$ için, x tahmin edicisi gittikçe büyürken, y ise $\beta_1 + \beta_2$ 'ye yaklaşır. Eğer $x=0$ için, y β_1 'e yaklaşır. Bu nedenle model tasvirindeki bir ilişki, x git gide büyürken β_1 ve $\beta_1 + \beta_2$

değerlerinin bu duruma tepkisi başlamaktadır. Analizlerin bu durumu doğrusal olmayan regresyonda oldukça ortaktır ve sık sık $x=0$ veya x gittikçe büyürken fonksiyonların ve parametrelerin nasıl etkilendiği tartışılabilir. (Anderson, T. W. , 2003)

Doğrusal olmayan regresyonda denklem tahmini doğrusal olmayan bir ya da birden fazla bilinmeyen parametrelere bağlı olarak tanımlanmaktadır. Doğrusal olmayan regresyonda genellikle tahmin edilen ile tahmin edici değişkenler arasındaki ilişki fiziksel bir sonucu ortaya koymaktadır.(Raymond H. Myers, 1990)

6.1 Doğrusal Olmayan Regresyonda Parametre Tahmini

Uygun bir regresyon modeli formül halinde yazılmadan önce verilerden parametre tahmini ve bazı çıkarsamalar yapılır. Parametre tahmini, doğrusal olmayan regresyon için, doğrusal regresyondan daha zor ve hassastır. İstatistik analizleri yapılırken bazı varsayımlarda bulunulur. Toplam kare hatanın küçülmesi ya da olabilirlik fonksiyonunun maksimum olması gibi başlangıç varsayımları sağlandığı zaman, parametre değerinin veya β 'nın bulunması için bazı istatistik programları kullanılır. Toplam kare hata minimum seçilecektir. Bunun anlamı β 'nın potansiyel değerlerinin bilgisayarda hesaplanarak $\sum_{i=1}^n (y_i - f(x_i, \theta))^2$ formülünün minimize edilmesidir. (S. H. Ngo , S. Kemeny, A. Deak. , 2003)

β değerinin minimize edilmiş kareler toplamı p boyutlu bir vektördür ve $\hat{\beta}$ olarak adlandırılır. Tahmin edilen varyans-kovaryans matrisi:

$$Var(\hat{\beta}) = s^2 (W'W)^{-1} \quad (6.4)$$

şeklinde yazılır. Burada W , $\hat{\beta}$ 'dan hesaplanan f 'nin birinci türevlerinin oluşturduğu $n \times p$ 'lik bir matris ve $s^2 = RSS/(n-p)$, artık kareler toplamının $n-p$ 'ye bölünmesi ile elde edilir. Bu matrisin köşegen elementleri $\hat{\beta}$ 'nın bileşenlerinin varyans tahminleridir. Aynı zamanda köşegen olmayan (off-diagonal) elementler tahmincilerin kovaryans tahmin edicileridir.

Artık kareler toplamının minimize edilmiş olan parametre vektöründe, $\hat{\beta}$ varyans-kovaryans matrisi $\sigma^2 (W'W)^{-1}$ olan yaklaşık olarak bir normal dağılımla dağılmaktadır. (S. H. Ngo , S. Kemeny, A. Deak. , 2003)

Doğrusal olmayan bir regresyon modelinin parametre tahmini genellikle doğrusal regresyonda da olduğu gibi, en küçük kareler metodu ya da en çok olabilirlik metodu ile yapılabilir.

Doğrusal olmayan regresyon modelinde hata terimleri sabit varyansla bağımsız normal dağıldığı zaman, bu yöntemlerin ikisi de aynı parametre tahminini yapabilmektedirler. (Marquardt, D.W. 1970)

Doğrusal olmayan regresyon modelinde en çok olabilirlik ve en küçük kareler yöntemleri için analitik bir ifade genellikle yoktur. Bunun yerine sayısal araştırma süreçleri, yoğun hesaplama gerektiren tahmin edici prosedürler kullanılır. Doğrusal olmayan regresyon modellerinin analizinde bu yüzden genellikle standart bilgisayar programları kullanılır. (SAS, STATISTICA, vs.) (Stein, C. M., 1960)

6.2 Doğrusal Olmayan Ridge Regresyon Modeli

Doğrusal olmayan bir regresyon modeli doğrusal regresyon modeline benzer,

$$y = f(\theta) + \varepsilon. \quad (6.5)$$

PLS fonksiyonu ele alındığında,

$$PLS(\theta) = \|y - f(\theta)\|^2 + \theta' \sum \theta, \quad (6.6)$$

burada $\sum$ $p \times p$ 'lik pozitif yarı belirli (positive semi definite) bir matristir. Bir ridge tahmin edicisi olan $\hat{\theta}$, $PLS(\theta)$ 'nin minimumuna karşılık gelen değerdir. (Seber, G. A. F. And Wild, C. J. , 1989)

Doğrusal olmayan bir regresyon modelinde parametre tahmini için ridge regresyon metodu uygulanmasının iki amacı vardır: Daha kararlı nümerik prosedürler yapmak ve aynı zamanda ortalama kare hatası daha küçük tahmin ediciler bulmaktır. (S. H. Ngo , S. Kemeny, A. Deak. 2003) Diğer dikkat edilecek nokta ise bu yöntem Levenberg-Marquardt yöntemi ile ilişkili olmasıdır. Doğrusal regresyon problemlerinin bir serisinin çözümü için Levenberg-Marquardt yöntemi eşdeğerdir. Bunun anlamı bu yöntem en küçük kareler yöntemi ile birbirlerine yakınsamaktadır. Levenberg-Marquardt yöntemi doğrusal olmayan fonksiyonların kareler toplamını minimum yapan popüler alternatif yöntemlerden birisidir. (Marquardt, D.W. 1963)

Ceza (penalty) parametresinin (PLS) amacı istatistiksel özelliği daha iyi olan parametrelerin tahmin edicilerini bulmaktır. (Seber, G. A. F. And Wild, C. J. , 1989)

θ 'nın küçük bir komşusu olan ξ doğrusal bir Taylor serisinin genişletilmesidir:

$$\hat{y} = F\theta + \varepsilon, \quad \varepsilon \sim N(0, \sigma^2)I, \quad (6.7)$$

burada $\bar{y} = y - f(\xi) + F\xi$ 'dir. $\hat{\theta}$ ve $\tilde{\theta}$ 'nın θ 'ya yakınsadığını varsayalım,

$$MSE(\hat{\theta}) \approx \sigma^2 (F' F)^{-1} \quad (6.8)$$

ve

$$MSE(\tilde{\theta}) \approx (F' F + \Sigma)^{-1} (\Sigma \theta \theta' \Sigma + \sigma^2 (F' F)) (F' F + \Sigma)^{-1}. \quad (6.9)$$

F matrisi doğrusal olmayan regresyonda, X matrisinin doğrusal regresyondaki rolüne benzer bir tasarıma sahiptir. (6.8) ve (6.9) denklemleri doğrusal durumda da aynı sonuca götürürler; bu denklemlerde: Σ daima vardır ve ridge regresyon tahmin edicisi $\tilde{\theta}$ en küçük kareler tahmin edicisinden $\hat{\theta}$ daha küçük bir toplam ortalama kare hataya sahiptir. Doğrusal olmayan ridge regresyonda bulunan Σ , doğrusal ridge regresyonda bulunan k parametresinin doğrusal olmayan durumdaki karşılığını oluşturmaktadır. (Seber, G. A. F. And Wild, C. J. , 1989)

6.3 Kanonik Yöntem ya da İlişkisiz Bileşen

Regresyon tahmin edicileri, matematiksel karmaşıklığı kolaylaştırmak için, kanonik forma ya da ilişkisiz bileşenlere dönüştürülürler. Kanonik formun diğer bir faydası ise bileşenlerin ilişkisiz olması ve her hangi bir bileşen için regresyon ağırlığının bağımlı değişkenle diğer bileşenlerinin aralarındaki ilişkiyi etkilememesidir. Bu nedenle her bileşen, regresyon modelinde sadece tahmin ediciler için incelenebilir. (Marquardt, D.W. , 1970)

$$X = H \Lambda^{1/2} T' \quad (6.10)$$

Yukarıdaki X matrisindeki, H X'in standartlaştırılmış formunun basit temel bileşenleridir. $H'H=I$ ve $\Lambda^{1/2}$ ise $(p \times p)$ 'lik, $\lambda_1^{1/2}, \lambda_2^{1/2}, \dots, \lambda_p^{1/2}$ köşegen elementlerinden oluşan bir köşegen matristir. $(p \times p)$ 'lik T matrisinin sütunları $v_1, v_2, \dots, v_p$ özvektörlerinden oluşmaktadır. Bu ifade X'in "tek değer ifadesi" (the single value decomposition SVD) olarak adlandırılır. (S. H. Ngo , S. Kemeny, A. Deak. 2003)

SVD'nin temeli en küçük karelere dayanmaktadır ve aşağıdaki şekilde yazılır:

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

$$= T\Lambda^{1/2} H'Y$$

ve

$$\text{Var} = \sigma^2 (X'X)^{-1} \quad (6.11)$$

$$= \sigma^2 T\Lambda^{-1} T'$$

$$= \sigma^2 \sum_{i=1}^p \lambda_i^{-1} v_i v_i'$$

olur.

β parametresi kanonik forma dönüştürüldüğünde:

$$\alpha = T' \beta$$

ve

$$Y = X^* \alpha + E$$

$$\text{burada } X^* = XT$$

ve

$$X^{*'} X = \Lambda \text{ dir}$$

Bu yüzden OLS tahmin edicisi α dönüştürülmüş parametresi için verilen denklem

$$\hat{\alpha} = T' \beta \quad (6.12)$$

ve

$$\text{Var}(\hat{\alpha}) = \sigma^2 (X^{*'} X^*)^{-1}$$

$$= \sigma^2 \Lambda^{-1} \quad (6.13)$$

şeklinde köşegen bir yapıya sahiptir. Bu durum birbiri ile ilişkisiz $\hat{\alpha}$ 'nın $\hat{\alpha}_1, \hat{\alpha}_2, \dots, \hat{\alpha}_p$ elemanları için bu parametrelerin ne kadar önemli olduğunu göstermektedir.

(6.12) denklemi

$$\beta = T\alpha$$

şeklinde gösterilebilir.

Bu nedenle, $\alpha_i \beta$ 'nın ilişkisiz bileşenlerinde biri olabilir. j. element olan β_j $\hat{\alpha}_1, \hat{\alpha}_2, \dots, \hat{\alpha}_p$ 'ların ağırlığı ile T'nin j. satırlarının ağırlıklarının toplamıdır. (Bates, Douglas. M. , 1988))

$$\beta_j = \sum_{i=1}^p v_{ji} \alpha_i \quad (6.14)$$

6.4 Ridge Regresyonun Kanonik Formu

Ridge regresyon tahmin edicileri $\hat{\beta}^R$, $\hat{\alpha}^R$ ilişkisiz bileşenlerine dönüştürülebilir.

$$\begin{aligned} X^* &= XT \\ \alpha &= T' \beta \\ \text{ve} \\ \Lambda &= T' X' XT \end{aligned} \quad (6.15)$$

α 'nın en küçük kareler tahmin edicisi ise:

$$\begin{aligned} (X^{*'} X) \hat{\alpha} &= X^{*'} Y \\ \text{yada} \\ \hat{\alpha} &= \Lambda^{-1} X^{*'} Y \end{aligned} \quad (6.16)$$

ve en küçük karelerin orijinal regresyon katsayıları

$$\hat{\beta} = T \hat{\alpha} \quad (6.17)$$

şeklinde gösterilir. (Marquardt, D.W. , 1970)

Benzer şekilde ridge tahmin edicisi

$$\begin{aligned} (\Lambda + K) \hat{\alpha}^R &= X^{*'} Y \\ \text{ya da} \\ \hat{\alpha}^R &= (\Lambda + K)^{-1} X^{*'} Y \end{aligned} \quad (6.18)$$

burada K , $k_1, k_2, \dots, k_p$ köşegen elementleri ve ridge tahmin edicisinin orijinal formunun kanonik formuna dönüştürülmesi ile bir köşegen matristir. (S. H. Ngo , S. Kemeny, A. Deak. 2003)

$$\hat{\beta}^R = T \hat{\alpha}^R \quad (6.19)$$

6.5 Bayesgil İstatistiksel Çıkarsama

Çok genel bir yaklaşım olan regresyona prior bilginin eklenmesi ile Bayesgil istatistiksel yöntemlere geçilebilir. Bayesgil çıkarsamanın temeli Bayes teoremine dayanmaktadır. Posterior (artçıl) olasılık yoğunluk fonksiyonu ve prior (öncül) olasılık yoğunluk fonksiyonu

en çok olabilirlik (likelihood) fonksiyonu yardımı ile hesaplanabilir. (Zellner Arnold, 1971)

$$\text{Posterior (oyf)} \propto \text{prior (oyf)} \times \text{en çok olabilirlik fonksiyonu} \quad (6.20)$$

En çok olabilirlik fonksiyonu bütün örneklem bilgisini içerirken; prior olasılık yoğunluk fonksiyonu da bütün prior bilgisini içermektedir. Bu yüzden Bayesgil yaklaşım, posterior dağılım ile prior dağılımın örneklem bilgisi ile oluşturdıkları bir kombinasyondur. β parametre vektörünün posterior dağılımın ortalaması, $\bar{\beta}$ Bayesgil tahmin edicisi olarak verilmektedir (David G. T., Denison, 2002)

Doğrusal regresyon modelini tahmin probleminde Bayesgil yaklaşım uygulandığında:

$$y = X\beta + u$$

β için posterior dağılım, β için prior dağılım ve en çok olabilirlik fonksiyonu elde edilir. Doğrusal regresyon modeli için örneklem bilgisi açıklayıcı değişkenler ve bağımlı veriler yolu ile elde edilir. Veri D matrisi tarafından elde edilir. (Zellner Arnold, 1971)

$$D = \begin{pmatrix} y & X \end{pmatrix} \\ T \times (P \times 1) \quad T \times 1 \quad T \times P$$

Eğer x ortalaması μ ve varyansı σ^2 ile normal dağılıyorsa ya da $x \sim N(\mu, \sigma^2)$ ise o zaman x'in olasılık yoğunluk fonksiyonu aşağıdaki şekilde tanımlanır.

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2\sigma^2}(x - \mu)^2\right] \quad (6.21)$$

En çok olabilirlik fonksiyonu ise:

$$L(\mu, \sigma^2) = \prod_{i=1}^T \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2\sigma^2}(x_i - \mu)^2\right] \\ = \frac{1}{(2\pi\sigma^2)^{T/2}} \exp\left[-\frac{1}{2\sigma^2} \sum_{i=1}^T (x_i - \mu)^2\right] \quad (6.22)$$

β 'nın Bayesgil tahmin edicisini bulmak için; doğrusal regresyon modelinde $u \sim N(0, \sigma^2 I)$ ve u_i bağımsız rasgele değişkenlerden oluşmakta varsayımları altında, örneklem değeri için en çok olabilirlik fonksiyonu:

$$L(\beta, \sigma/y) = \frac{1}{(2\mu\sigma^2)^{T/2}} \exp\left[-\frac{1}{2\sigma^2} u'u\right]$$

ya da

$$L(\beta, \sigma/y) \propto \frac{1}{\sigma^T} \exp\left[-\frac{1}{2\sigma^2} (y - x\beta)'(y - x\beta)\right] \quad (6.23)$$

şeklinde gösterilir. (David G. T., Denison, 2002)

En çok olabilirlik fonksiyonundaki üstel terim, $\hat{\beta}$ sıralı en küçük kareler tahmin edicisindeki terim gibi yazılabilir. Bazı işlemler yapıldığında:

$$(y - x\beta)'(y - x\beta) = y'y - (\beta'x'y) - \beta'x'y + \beta'x'\beta$$

elde edilir. $X'y$ 'nin yerine $x'x\hat{\beta}$ konulduğunda :

$$(y - x\beta)'(y - x\beta) = y'y - \hat{\beta}'x'x\beta - \hat{\beta}'x'x\hat{\beta} + \beta'x'x\beta$$

ve $\hat{\beta}'x'x\hat{\beta} - \hat{\beta}'x'y$ ile $\hat{\beta}'x'x\hat{\beta} - y'x\hat{\beta}$ terimleri eklendiğinde elde edilen denklem

$$(y - x\beta)'(y - x\beta) = (y - x\hat{\beta})'(y - x\hat{\beta}) + (\beta - \hat{\beta})'x'x(\beta - \hat{\beta})$$

şeklindedir. En küçük karelerden hareketle

$$S^2 = (y - x\hat{\beta})'(y - x\hat{\beta})/n \quad (6.24)$$

burada $n=T-P$ dir. Bu nedenle

$$(y - x\beta)'(y - x\beta) = ns^2 + (\beta - \hat{\beta})'x'x(\beta - \hat{\beta}) \quad (6.25)$$

ve en çok olabilirlik fonksiyonu ise

$$L(\beta, \sigma/y) \propto \frac{1}{\sigma^T} \exp\left\{-\frac{1}{2\sigma^2} [ns^2 - (\beta - \hat{\beta})'X'X(\beta - \hat{\beta})]\right\} \quad (6.26)$$

şeklindedir. (James, W. and Stein, C. 1961)

Posterior (artçıl) dağılımı elde edilebilmesi için prior yoğunluk fonksiyonu belirtilmelidir.

Eğer β 'nın prior bilgisi β_0 ortalama vektörü ile normal dağılmış olarak tanımlanırsa ve kovaryans matrisi $v^2 A^{-1}$ ya da $\beta \sim N(\beta_0, v^2 A^{-1})$ ise o zaman prior yoğunluk fonksiyonu:

$$g(\beta, \sigma) \propto \frac{1}{v^k} \exp\left[-\frac{1}{2}(\beta - \beta_0)' A / v^2 (\beta - \beta_0)\right] \quad (6.27)$$

şeklinde yazılır. (Zellner Arnold, 1971)

Prior yoğunluk fonksiyonu ve en çok olabilirlik fonksiyonu etkileşiminden

$$p(\beta / \sigma) \propto v^{-k} \sigma^{-T} \exp\left\{-\frac{1}{2}\left[\frac{ns^2 + (\beta - \hat{\beta})' X' X (\beta - \hat{\beta})}{\sigma^2} + (\beta - \beta_0)' \frac{A}{v^2} (\beta - \beta_0)\right]\right\} \quad (6.28)$$

elde edilir ve ikinci dereceden (kuadratik) formu genişletildiğinde:

$$-\frac{1}{2}\left[ns^2 + \beta' \left(\frac{X' X}{\sigma^2} + \frac{A}{v^2}\right) \beta + \text{diğer terimler}\right] \quad (6.29)$$

(6.29) denklemi aşağıdaki şekilde de yazılabilir: (Zellner Arnold, 1971)

$$-\frac{1}{2}\left[ns^2 + (\beta - \bar{\beta})' \left(\frac{X' X}{\sigma^2} + \frac{A}{v^2}\right) (\beta - \bar{\beta})\right]$$

burada

$$\bar{\beta} = \left(\frac{X' X}{\sigma^2} + \frac{A}{v^2}\right)^{-1} \left(\frac{X' X}{\sigma^2} \hat{\beta} + \frac{A}{v^2} \beta_0\right) \quad (6.30)$$

Bu yüzden, β 'nın posterior yoğunluk fonksiyonunun analizine göre;

$$p(\beta / \sigma) \propto v^{-k} \sigma^{-T} \exp\left\{-\frac{1}{2}\left[ns^2 + (\beta - \bar{\beta})' \left(\frac{X' X}{\sigma^2} + \frac{A}{v^2}\right) (\beta - \bar{\beta}) + \beta \text{ içermeyen terimler}\right]\right\}$$

Bu ortalaması $\bar{\beta}$ ve varyans-kovaryans matrisi $[X' X / \sigma^2 + A / v^2]^{-1}$ olan bir normal dağılıma

sahiptir. Bu nedenle $\bar{\beta}$, β 'nın Bayesgil tahmin edicisidir. (Zellner Arnold, 1971)

(6.30) denkleminde β_0 sağdan bir eklenip bir çıkarıldığında, aşağıdaki sonuca ulaşılır.

$$-(\frac{X'X}{\sigma^2} + \frac{A}{v^2})^{-1} (\frac{X'X}{\sigma^2} + \frac{A}{v^2})\beta_0 \quad (6.31)$$

Bu durum (6.30) denklemi için ikinci matris açıklamasına imkan vermektedir ve A/v^2 terimi çıkarılır. (David G. T., Denision, 2002) Sonuçta,

$$\bar{\bar{\beta}} = \beta_0 + (\frac{X'X}{\sigma^2} + \frac{A}{v^2})^{-1} \frac{X'X}{\sigma^2} (\hat{\beta} - \beta_0) \equiv \beta_0 + W(\hat{\beta} - \beta_0) \quad (6.32)$$

burada $W \equiv (\frac{X'X}{\sigma^2} + \frac{A}{v^2})^{-1} \frac{X'X}{\sigma^2}$, dir.

Sonuç olarak

$$\bar{\bar{\beta}} = (I - W)\beta_0 + W\hat{\beta} \quad (6.33)$$

elde edilir.

Bundan dolayı $\bar{\bar{\beta}}$, β_0 ve $\hat{\beta}$ 'nın ağırlıklı ortalamasıdır. (David G. T., Denision, 2002)

6.6 Bayesgil Çıkarsama ve Ridge Regresyon

Bir önceki bölümde β 'nın Bayesgil tahmin edicisi gösterilmiştir.

$$\bar{\bar{\beta}} = (\frac{X'X}{\sigma^2} + \frac{A}{v^2})^{-1} (\frac{X'X}{\sigma^2} \hat{\beta} + \frac{A}{v^2} \beta_0) \quad (6.34)$$

Eğer $\beta_0=0$ ve $A=I$ olarak seçilseydi bu durumda,

$$\begin{aligned} \bar{\bar{\beta}} &= (\frac{X'X}{\sigma^2} + \frac{1}{v^2} I)^{-1} (\frac{X'y}{\sigma^2}) \\ &= (X'X + \frac{\sigma^2}{v^2} I)^{-1} X'y \end{aligned} \quad (6.35)$$

şeklinde bir denklem elde edilir.

Bilindiği gibi β 'nın prior dağılımının ortak varyansı v^2 'dir. $k = \sigma^2/v^2$ olduğu zaman

$$\bar{\beta} = \beta^R = (X'X + KI)^{-1} X'y \quad (6.36)$$

ulaşılır. (David G. T., Denision, 2002)

Elde edilen $\bar{\beta}$ ya da β^R ridge tahmin edicisidir. Bundan dolayı $k = \sigma^2/v^2$ için ridge tahmin edicisi, Bayesgil bir tahmin edicidir, denir. Eğer $\beta_0=0$ varsayımı yapılırsa o zaman bütün regresyon katsayılarının ortak bir varyansa sahip olduğu söylenebilir. (Zellner Arnold, 1971)

Bayesgil tahmin edici ile ridge tahmin edicisi arasındaki yakın ilişki, ridge tahmin edicisinin bilinmeyen β parametreleri hakkındaki, prior bilgisini birleştirmeye yönelik bir teşebbüs olduğunu gösterir. (David G. T., Denision, 2002)

Çoklu bağlantı durumunda, araştırmacılar genellikle örneklem bilgisinin doğruluğundan şüphelenirler. Araştırmacılar elde edilen örneklem sonuçlarını bazı prior bilgilerle karşılaştırırlar. Özellikle araştırmacılar doğru parametrelerin örneklem tahmin edicisinden sıfıra daha yakın olduğuna inanmaktadırlar. (James, W. and Stein, C. 1961)

v^2 ve β_0 'ın özellikleri çok az sınırlandırılabilir. Eğer her bir katsayı için $(v_1^2 + v_2^2 + \dots + v_p^2)$ farklı bir prior varyans seçilirse,

$$k_1 = \sigma^2/v_1^2$$

$$k_2 = \sigma^2/v_2^2$$

.

.

.

elde edilir. (Zellner Arnold, 1971)

Ridge tahmin edicisinin genelleştirilmiş hali

$$\hat{\beta}^R = B_G^* = [X'X + \Omega]^{-1} X'y \quad (6.37)$$

dir. Burada

$$\Omega = \begin{bmatrix} k_1 & 0 & \cdots & 0 \\ 0 & k_2 & \cdots & 0 \\ 0 & 0 & \cdots & k_p \end{bmatrix}$$

olarak gösterilir. (David G. T., Denison, 2002)

Buna ilaveten β_0 sıfırdan farklı olarak seçilirse Bayes tahmin edicisi olarak kullanılan (6.34) numaralı formül elde edilir. (David G. T., Denison, 2002)

7. ZAMAN SERİLERİ ve TEMEL BİLEŞENLER REGRESYON

7.1 Zaman Serileri Analizi

Bir zaman serisi, ilgilenilen bir büyüklüğün zaman içerisinde sıralanmış ölçümlerinin bir kümesidir. (MADDALA, G.S. and Inn-Moo Kim, 2002) Zaman serisi ile ilgili bu analizin yapılma amacı ise, gözlem kümesince temsil edilen gerçeğin anlaşılması ve zaman serisindeki değişkenlerin gelecekteki değerlerinin doğru bir şekilde tahmin (forecast) edilmesidir. Ekonometrik yöntemlerin zaman serileri için uygulanmasına geçmeden önce zaman serilerini oluşturan farklı unsurları incelemek gerekmektedir.

Zaman serileri dört bileşenden oluşur;

1. Trend (Genel Eğilim) bileşeni; Zaman serilerinin uzun sürede gösterdiği düşme ve yükselme süreçlerinden sonra oluşan kararlı durumdur. Zaman serileri uzun dönem açısından kararlı alçalma ya da yükselme şeklinde bir eğilime sahiptir.
2. Mevsim Bileşeni; Zaman serilerinde mevsimlere göre değişmeyi ifade eder. Zaman serileri açısından kullanılan verilerin kimi dönemleri diğer dönemlere göre farklılık gösterir.
3. Çevrimsel Bileşen; Ekonomide, mevsimsel değişimler ile ilgili olmayan dönemsel değişimlerdir. Örneğin, ekonomide genel eğilimden bağımsız kısa süreli genişleme ya da daralma durumu çevrimsel süreci tarif eder.
4. Düzensiz Bileşen; Diğer unsurlar gibi belirli olmayan, hata terimi ile ifade edilebilecek değişimlerdir.

Zaman serileri tüm bu kendilerini oluşturan bileşenlere ayrıştırıldıktan sonra, bileşenlerin toplamı şeklinde, $Y_t = T_t + S_t + C_t + I_t$ ya da çarpma yöntemi ile $Y_t = T_t S_t C_t I_t$ şeklinde belli bir t döneminde Y zaman serisi ifade edilebilir. (KUTLAR Azia, 2000)

7.1.1 Otoregresif (AR) Süreç

Bir AR modelinde, bağımlı değişken geçmişteki değerinin bir fonksiyonudur. Bir çok zaman serisi verisi de bu süreci içermektedir. Bu durum aşağıdaki gibi bir denklemle ifade edilebilir.

$$x_t = a + a_1 * x_{(t-1)} + a_2 * x_{(t-2)} + a_3 * x_{(t-3)} + \dots + \varepsilon \quad (7.1)$$

Burada, a, sabit terimi temsil ederken $a_1, \dots$ gibi katsayılarda gecikmeli değerlerin şimdiki değerle olan ilişkisini temsil eder. Ayrıca ε hata terimi de rassal şokları tanımlamaktadır. Genel bir şekilde AR(p) şeklinde ifade edilir. (MADDALA, G.S. and Inn-Moo Kim, 2002)

AR süreci için aşağıdaki durum aktarılabilir:

“Bir limonata satıcısı olduğunuzu ve her saat beş bardak limonata sattığınızı düşünürseniz; eğer siz limonata sattığınız yeri kapatmak ve limonata bittiği için satmaktan vazgeçmek istemiyorsanız, her saat başına tükenen limonata yerine yeni limonata doldurmanız gerekir. Böylece her saat beş bardak limonata satılsa da siz her zaman yerine yenisini ilave ettiğinizden siz bir kaza geçirmediğiniz sürece asla limonata satışınızda bir aksama olmaz. Bu bir otoregresif süreci tarif eder. Çünkü daha az ya da daha fazla limonata satmanız şeklinde bir şok belli bir saatteki limonata seviyesini etkiler. “(Kutlar Azia, 2000)

7.1.2 Hareketli Ortalama (MA) Süreci

Eğer serinin gecikmeli hata terimi, şimdiki hata terimini etkiliyorsa hareketli ortalama süreci tanımlanır. Bir hareketli ortalama sürecinde değişkenin tahmin değeri hata terimlerinin tahmin değeri ile ilgilidir.

Bir hareketli ortalama süreci,

$$x_t = e_t - a_1 e_t - \dots, t = 1, 2, \dots, n \quad (7.2)$$

şeklinde ifade edilebilir. Genel şekil olarak MA(q) şeklinde ifade edilir.

Hareketli ortalama sürecinde, her bir gecikmeli hata terimi onun şimdiki değerini etkilemektedir. (Kutlar Azia, 2000)

Hareketli ortalama süreci için aşağıdaki durum aktarılabilir:

”yolda kalan kamyonları çekmek üzerine uzmanlaşmış olan bir şirkete sahip olduğunuzu düşünürseniz, her bir yolda kalan aracın çekilmesi bir bağımsız olay olacaktır. Deneyimleriniz aracın bozulduğu yere ve araca sahip olan şirketin bir tamir şirketinin tamirhanesinin olduğu yere bağlı olarak, bir aracın çekilmesi ve onun tamirhaneye götürülmesi için üç günün gerekli olduğunu göstermiştir. Eğer siz yeterli çekiciye sahip olmazsanız aracın sahipleri bu işi başkasına verecektir. Bir gündeki tamir edilmek için çekilmesi gerekli araç sayısı size gerekli olan çekici için bilgi vermektedir. Üç gün ötesinde, bu günkü tercihler size gelecekte olanlar hakkında bir şey söylemez. Bu süreç bir hareketli ortalama sürecidir.“ (Muammer ŞİMŞEK, 2003)

7.1.3 ARMA ve ARIMA Süreci

Çoğu zaman serisi gerek AR ve gerekse de MA sürecini içermektedir. Ayrıca I durağan

(integrated) ise serisinin içerdği trendi ifade etmektedir. ARIMA olarak ifade edilen süreç değerlendirilecek olursa, ARIMA(p,d,q), zaman serisinin p dereceden kendisinin gecikmesi ile ilişkisini ifade eden AR(p) sürecini içerdığını, q ile ifade edilen ve hata terimlerinin geçmiş değerleri ile ilişkisini ifade eden ve rassal süreci yansıtmaması açısından hata terimlerinin düzleştirilme yöntemlerinden olan MA(q) sürecini ifade etmektedir. Zaman serilerinde genelde olduğu gibi bir durağan olmama durumunun olması ya da durağan bir zaman serisi ile karşılaşmamız d ile ifade edilir. Zaman serisi hangi düzeyde durağan (integrated) ise belirtilir. (Kutlar Azia, 2000)

Söz konusu denklem,

$$Y_t = m + a_1 y_{t-1} + \dots + a_p y_{t-p} + u_t - b_1 u_{t-1} - \dots - b_q u_{t-q} \quad (7.3)$$

ARIMA(p,d,q) sürecini ifade eder.

AR ve MA için bir örnek olarak,

“Ulusal park yakınında bir otele sahip olduğunuzu düşünün, Otel defteri bazı rezervasyonları içermektedir. Yani, müşterilerin bazıları otelinizde bir günden daha fazla zaman harcamakta ve ayrıca müşterilerden bazıları da ulusal parkta bir haftalık tatil geçirirken aynı zamanda evlerine dönmeyen önce gece kalmak için sizin otelinize gelmektedirler. Belli bir günde meydana gelecek çok otelde sürekli kalan müşterileri artan-birden fazla dönem- bir şekilde etkileyecektir. Ancak bu şokun, ulusal parkta bir tatil geçirdikten sonra gece için kalmaya gelenler üstünde bir hafta sonra tek bir etkisi olacaktır” verilebilir. (Muammer ŞİMŞEK, 2003)

Box-Jenkins yönteminin temeli ARIMA(p,d,q) paradigmasına dayanır. Otoregresif ve hareketli ortalamaya sahip zaman serilerini incelemekte ve bu yönteme göre zaman serilerinin durağanlığı korelogram ile tespit edilmekte ve yine zaman serisinin ne tür bir süreç içerdği de korelasyon fonksiyonları ile analiz edilmektedir. Zaman serilerinde durağan olmama durumunda farkı alınarak durağanlaştırılmakta, ancak eğer yine durağanlık ile karşılaşılmaz ise bu sefer verilere yine geri dönülmektedir. Ancak eğer durağanlık sağlanmış ise bu durumda model ARMA ile tahmin edilip, öngörü işlemi yapılmaktadır. (Kutlar Azia, 2000)

7.2 Zaman Serilerinde Durağanlık ve Trend

Zaman serilerinin durağan olması olarak ifade edilen şey, zaman içinde varyansın ve ortalamanın sabit olması ve gecikmeli iki zaman periyodundaki değişkenlerin ko-varyansının

değişkenler arasındaki gecikmeye bağlı olup zamana bağlı olmamasıdır.

$$\text{Ortalama} = E(Y_t) = \mu$$

$$\text{Varyans} = \text{var}(Y_t - \mu)^2 = \sigma^2$$

$$\text{Kovaryans} = \gamma_k = E((Y_t - \mu)(Y_{t-k} - \mu))$$

Zaman serilerinin durağan olmaması durumunda, zaman serileri trend içerebilir. Bu durumda zaman serilerinin kullanılacağı öngörü ve regresyon denklemlerinde sahte regresyon benzeri durumlar ortaya çıkacaktır. Zaman serileri durağan değilse, stokastik ya da deterministik trend içermektedirler. Ancak seri üzerinde uzun dönemde deterministik bir trendin varlığı ile düzensiz modellerde zaman içinde ortaya çıkan ve bir müddet sonra kayıp olan trendler, birbirinden farklıdır. (MADDALA, G.S. and Inn-Moo Kim, 2002)

Deterministik trend, oldukça uzun bir dönemde ortaya çıkan ve yükseliş ve alçalış zikzakları arasında belli bir yöndeki uzun dönemdeki eğilimi ifade etmektedir. Zaman serisi içerisinde trendin bütünü ile kestirilebilir oluşu deterministik trende işaret etmektedir. (GUJARATI, Damodar, 1995)

$$Y_t = a + \beta t + \varepsilon \quad (7.4)$$

Söz konusu denklem durağan olmayan bir zaman serisi içerisindeki deterministik trendi tarif etmektedir. Eğer rassal yürüyüş (random walk) modeli,

$$Y_t = Y_{t-1} + \varepsilon \quad (7.5)$$

stokastik bir trendi ifade etmiş oluruz. Yine stokastik ve deterministik trendle birlikte gösterimi de mümkündür. Ayrıca söz konusu denklemde yer alan sabit terimde, zaman serileri belli bir başlangıç değerine sahip olması gerektiğinden ve bu sabit değerinin anlamsızda olsa denkleme dahil edilmemesi durumunda durağan olmama durumlarının ortaya çıkabileceğinden dahil edilmektedir. Durumu, trend durağan süreç (trend stationary process) ve fark durağan süreç (difference stationary process) açısından incelendiğinde zaman serisine yapılacak doğrusal trend ilavesi ile yani trendden arındırma sonucunda eğer zaman serisi durağan hale geliyorsa bu zaman serisinin deterministik bir trende sahip olduğunu, ancak eğer doğrusal bir trend ilavesi eğer zaman serisini durağan hale getirmiyorsa bu durumda fark alınarak zaman serisi durağan hale getirileceğinden stokastik bir trend geçerlidir. (MADDALA, G.S. and Inn-Moo Kim, 2002)

7.3 Durağanlığın Tespiti

7.3.1 Görsel Saptama -Korelogram

Bu durağanlık testi otokorelasyon fonksiyonuna (ACF) dayanır. Otokorelasyon fonksiyonu serinin bazı değerleri ve gecikmeli değerleri arasındaki ilişkinin boyutunu belirler. Değişik zaman aralıkları (k) için bulunacak ACF(k) katsayısı değerleri ilişkilendirildiğinde, korelogram elde edilir. ACF(k) değerleri 1 ve -1 arasında yer almaktadır. (GUJARATI,Damodar, 1995)

$$S(X_t - X_{bar})(X_{t-k} - X_{bar})$$

$$ACF(k) = \frac{S(X_t - X_{bar})(X_{t-k} - X_{bar})}{S(X_{t-k} - X_{bar})^2}$$

$$S(X_{t-k} - X_{bar})^2$$

Durağanlık tespiti için korelogramdan şu şekilde yararlanılır: ACF eğer çok yüksek bir değerden başlayıp çok yavaş küçülüyorsa, bu serinin durağan olmadığının bir göstergesidir. Söz konusu hipotez testi her bir ACF(k) değeri için $\pm 1.96(1/\sqrt{n})$ değeri bulunarak yapılır. Eğer ACF(k) değeri güven aralığı sınırları dışında kalıyorsa otokorelasyon vardır. Kısmi korelasyon fonksiyonu gecikmeli değişkenler arasındaki ilişkiyi ifade eder. Kısmi korelasyon fonksiyonu ile korelasyon Y ve Y_{t-k} değerleri arasındaki terimlerin etkisi çıkarılarak bulunur. (MADDALA, G.S. and Inn-Moo Kim, 2002)

Bütün bu ACF(k) değerlerinin eşanlı olarak sıfıra eşit olduğunun testi için diğer bir yöntemde, Box-PierceQ ve Ljung-Box istatistiğinin kullanılmasıdır. (GUJARATI,Damodar, 1995)

n, örneklem büyüklüğü, m gecikme sayısı iken Q ve LB istatistiği ki-kare dağılımından yararlanılarak test edilir.

H_0 =Bütün ACF(k) lar sıfıra eşittir.

H_a =Bütün Acf(k) lar sıfırdan farklıdır.

Hipotezleri geçerli iken eğer hesaplanan Q ve LB değeri ki-kare çizelgesindeki eşik değerinden büyükse H_0 red edilir. Yani seri durağan değildir denir. (KUTLAR Azia, 2000)

7.3.2 Dickey-Fuller Testi

Dickey-Fuller testi, gözlenen serilerde birim kökün var olup olmadığının belirlenmesinde kullanılan bir testtir. Ancak yapılan uygulamalarda serinin birim kök taşıyıp taşımadığının

saptanması için mutlak suretle DF (Dickey-Fuller) testinin yapılması şart niteliğinde bulunmaktadır. (GUJARATI,Damodar, 1995)

Testin kullanımını açıklamak için aşağıdaki veri üreten süreç kullanılır:

$$\text{Model: } Y_t = p Y_{t-1} + u_t$$

u_t = stokastik hata terimi

Eşitliği aşağıdaki gibi gösterilir.

$$Y_t - Y_{t-1} = (p-1) Y_{t-1} + u_t \quad (7.6)$$

Denklemin her iki tarafından Y_{t-1} çıkarıldığında, $(p-1) = \gamma$ olmak üzere denklem aşağıdaki şekle gelir:

$$\Delta Y_t = \gamma Y_{t-1} + u_t$$

$$H_0 : p=1, H_1 : p<1$$

$(p-1) = 0$ veya $\gamma = 0$ durumunda Y_t serisi bir birim kök içermektedir. Ancak $|p| < 1$ durumunda seri durağan olur. Burada Dickey ve Fuller'ın Monte Carlo uygulamasında ortaya çıkarılan "T" (tau) istatistiği kullanılmaktadır.

Hesaplanan "T" değerinin mutlak değeri Dickey-Fuller veya McKinnon Dickey-Fuller kritik değerlerinin mutlak değerini aşıyorsa, zaman serisinin durağan olduğu hipotezi reddedilmez. ' $H_0 : p=1$ ' reddedilirse zaman serisi durağandır.

Dickey-Fuller'in ortaya koyduğu üç denklem türü bulunmaktadır ;

$$\text{Sabitli trendsiz Dickey-Fuller denklemi : } \Delta Y_t = \gamma Y_{t-1} + u_t$$

$$\text{Sabitli trendsiz Dickey-Fuller denklemi : } \Delta Y_t = a + \gamma Y_{t-1} + u_t$$

$$\text{Sabitli trendli Dickey-Fuller denklemi : } \Delta Y_t = a + bt + \gamma Y_{t-1} + u_t$$

Birim kökün varlığının sınanması için kullanılan iki hipotez kullanılmaktadır. Bunlar;

$$H_1 : \gamma < 0 \text{ (} p < 1 \text{) (seride birim kök yoktur.) (seri durağandır.)}$$

$$H_0 : \gamma = 0 \text{ (} p = 1 \text{) (seride birim kök vardır.) (seri durağan değildir.)}$$

hipotezlerinin oluşturulduktan sonra mevcut model içinde sınanması şu şekilde olmaktadır;

Dickey-Fuller testinin uygulanmasında “ $\Delta Y_t = \gamma Y_{t-1} + u_t$ ” regresyonunda yer alan γ parametresinin sahip olduğu ‘t’ değerinin, Dickey-Fuller’a özel olarak hazırlanan “T” istatistik tablo değeri ile karşılaştırılarak, önceden hazırlanan H_0 ve H_1 hipotezlerine göre birim kökün varlığı tespit edilmektedir. (MADDALA, G.S. and Inn-Moo Kim, 2002)

7.4 Granger Causality Nedensellik Testi

Engle-Granger (E-G) yöntemi basit bir analize dayanmaktadır. Buna göre (7.4) numaralı denklem en küçük kareler (EKK) yöntemi kullanılarak tahmin edildikten sonra, bu regresyonun hata terimi çekilerek durağanlığının araştırılmasına dayanmaktadır. Engle-Granger yönteminde önemli olan, $Y_t = \alpha + \beta X_t + \varepsilon_t$ ilişkisi $\varepsilon_t = Y_t - \alpha - \beta X_t$ şeklinde ifade edildiğinde ε_t hata terimlerinin $I(0)$ ve dolayısıyla durağan olup olmama hususudur. Eğer ε_t terimleri durağan ise iki zaman serisi arasında eş-bütünleşme var demektir. Bu amaçla iki serinin uzun dönem ilişkisine sahip olup olmadıkları iki aşamalı E-G yöntemiyle şu şekilde analiz edilmektedir:

$Y_t = \alpha + \beta X_t + \varepsilon_t$ için $\hat{Y}_t = a + bX_t$ regresyonu ve bunun yardımı ile $e_t = Y_t - \hat{Y}_t$ hata terimleri bulunur. $\Delta e_t = \delta e_{t-1} + v_t$ ilişkisi ele alınır. Burada v_t bir hata terimidir. Bunun için regresyon τ -istatistiği(ADF test istatistiği) bulunur ve Mac-Kinnon kritik değerleri çerçevesinde ADF-testi yapılır. Yapılan ADF testi sonucunda elde edilen τ -istatistiği(ADF test istatistiği) mutlak değer olarak tablo değerinden büyükse, yani, e_t hata terimi $I(0)$ ise iki zaman serisi arasında uzun dönemde bir ilişki söz konusudur. Değişkenlerin her ikisi de aynı seviyede durağan ise, Engle-Granger eş bütünleşme regresyonu aşağıdaki denklemlerle ifade edilir. (KUTLAR Azia, 2000)

7.5 Eşbütünleşme (Koentegrasyon)

Birim kök testlerinde serilerin düzeyde durağan olmadığının ortaya çıkmasından sonra sahte regresyon probleminden kaçınmak için daha önce yapılan iş, regresyon analizinde serilerin durağan hale geldikleri düzeyde kullanılmalarıdır. Ancak Gujarati’nin deyimiyle bu, kaş yaparken göz çıkarmaya benzemektedir. Çünkü durağan hale getirmek için serilerin farklarının alınması, özgün düzeylerince belirlenen uzun dönem ilişkisinin yitirilmesine neden olabilmektedir. (GUJARATI,Damodar, 1995)

Bu sorunun çözümü Engle ve Granger tarafından geliştirilen eşbütünleşme (cointegration)

testi sayesinde bulunmuştur. Eşbütünleşme testi düzeyde durağan olmayan serilerin uzun dönemde birlikte hareket edip etmediklerini ortaya koymaktadır. Eğer seriler arasında bir eşbütünleşme ilişkisi varsa yani uzun dönemde birlikte hareket ediyorsa, düzey değerleriyle yapılacak analizde sahte regresyon problemiyle karşılaşılmayacaktır. (MADDALA, G.S. and Inn-Moo Kim, 2002)

7.6 Var Analizi

VAR yöntemi modellerde yer alan değişkenlerin bir sistem olarak birbirleriyle etkileşimlerini göstermesi açısından oldukça önemli bir ekonometri tekniğidir. VAR sisteminden elde edilen katsayıların doğrudan yorumlanması zor olmasına karşılık F testi sonuçları sağ tarafta yer alan değişkenlerin beklenen etkileri konusunda bilgi verdiği için oldukça yararlıdır. Beklenmeyen şokların makro değişkenler üzerindeki etkilerini incelemek için Sims (1980), VAR modelinin hareketli ortalamalar bölümünden elde edilen etki tepki analizlerinin ve varyans ayrıştırmalarının yapılmasını önermektedir.

VAR sistemi içinde yer alan değişkenlerin durağan olup olmaması gereği tartışma konusudur. Fark alımına karşı temel tartışma, verilerdeki birlikte hareketle ilgili bilgilerde bir kayba yol açmasıdır. Bu konuda yaygın görüş sistemdeki değişkenlerin gerçek veri türetme prosedüründen elde edilmesi şeklindedir. Son zamanlardaki görgül (ampirik) çalışmalar, model tahmininde veri setinin ilk farklarının veya seviye değerlerinin kullanılmasının daha uygun olup olmadığını test etmeye yönelmiştir. (MADDALA, G.S. and Inn-Moo Kim, 2002) İstatistiksel teknikler, zaman serisi değişkenlerinin seviye değerlerinin kullanılmasının daha uygun olup olmadığını, yani verilerin zaman süreci içinde değişip değişmediğini belirlemeye yöneliktir. Serilerin bu özelliğinin belirlenmesi oldukça önemlidir. Çünkü durağan olmayan serilerle yapılan tahminlerin analizi, durağan seriler kullanarak yapılan tahminlerden farklı bir şekilde değerlendirilmelidir. Birçok durumda, durağan olmayan seriler kullanılarak yapılan tahminler, istatistiksel açıdan doğru olmayan sonuçlar ortaya koyabilmektedir. Sims (1980) ve Doan (1992) birim kök içerseler bile fark alımına karşı görüş belirterek, VAR analizinin amacının parametre tahmini olmadığını, fakat değişkenler arası ilişkilerin belirlenmesi olduğunu ifade etmektedirler. (Sims, C. 1980, Doan, T., 1992) Aynı şekilde Cooley ve Roy (1985)'e göre de VAR analizinin amacı, parametre tahmininden ziyade değişkenler arasındaki ilişkileri bulmaktır. (Cooley, T.F. and Roy, S.F, 1985)

7.7 Temel Bileşenler Regresyonu

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

Y, bağımlı değişkeni; X, bağımsız değişkenleri; β , regresyon katsayılarını ve ε , hata terimini göstermektedir.

Temel bileşenler (PC) analizinde ilk olarak hem bağımlı hem de bağımsız değişkenler ortalamadan farkları alınıp standart sapmalarına bölünerek standartlaştırılmaktadır. Nihai regresyon katsayıları gösterildiğinde ise, değişkenler orijinal ölçeğine dönüştürülmektedir. Daha önce belirtildiği gibi EKK tekniğiyle regresyon katsayıları aşağıdaki eşitlikle hesaplanmaktadır: (Hawkins, D.M. 1973)

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

Değişkenler standartlaştırıldığından $X'X = R'$ dir. Burada R, bağımsız değişkenler için korelasyon matrisini göstermektedir. Bu analizi gerçekleştirmek için bağımsız değişkenler temel bileşenlere dönüştürülür. Bu matematik olarak aşağıdaki gibi yazılmaktadır:

$$X'X = PDP' = Z'Z \quad (7.7)$$

Eşitlikte PC modelini tanımlayan D, $X'X$ özdeğerlerinin köşegen matrisini; P, $X'X$ özvektör matrisini ve Z, veri matrisini (X yapısına benzer) göstermektedir. Temel bileşenler (P) ortogonal olduğundan $P'P = I$ dir. (GUJARATI,Damodar, 1995)

Böylece orijinal değişkenlerin (X) ağırlıklı ortalamalarını ifade eden yeni değişkenler (Z) türetilmektedir. Bu durum, regresyon analizine başlamadan önce değişkenlerin logaritmik veya karekök dönüşümlerinin alınmasından farklı bir şey değildir. Bu yeni değişkenler temel bileşen olduğu için, bu bileşenler arasındaki korelasyonlar sıfırdır. X1, X2, X3 gibi değişkenlerle analize başlanmışsa, Z1, Z2 ve Z3 gibi dönüştürülmüş değişkenler elde edilmektedir. (GUJARATI,Damodar, 1995)

Çok küçük özdeğerler hesaplanması durumunda çok güçlü çoklu doğrusal bağlantı (severe multicollinearity) sorunu söz konusu olmaktadır. Bu sorunun üstesinden gelebilmek için düşük özdeğerlerle özdeşleşen bileşenler analizden çıkartılmaktadır. Genelde göreceli olarak özdeğeri küçük bir veya iki (en çok) temel bileşen elde edilmektedir. Örneğin üç değişkenli bir modelde özdeğeri küçük sadece bir temel bileşen elde edilmişse, Z3 (üçüncü temel bileşen) analizden çıkartılmaktadır. (GUJARATI,Damodar, 1995)

Böylece Y değişkeni bağımlı ve Z1 ve Z2 bileşenleri ise bağımsız değişkenler olarak alınan

modelde artık çoklu doğrusal bağlantı söz konusu olmamaktadır. Daha sonra sonuçlar X ölçeğine geri dönüştürülerek B'nin tahminleri elde edilir. Bu tahminler yanlış olacaktır. Fakat bu yanlışlığın büyüklüğü varyansın azaltılmasıyla dengeleneceği umulmaktadır. (GUJARATI,Damodar, 1995) Diğer bir anlatımla PC tahminlerinin hata kareleri ortalamasının EKK tahminlerinden daha küçük olması beklenmektedir. (Hawkins, D.M. 1973)

Matematik olarak, asal bileşenlerin özel doğası gereği, regresyon katsayılarının tahmini aşağıdaki gibidir :

$$\hat{A}=(Z'Z)^{-1}Z'Y=D^{-1}Z'Y \quad (7.8)$$

Artık bu eşitlik, değişik bağımsız değişken seti üzerine uygulanan EKK regresyonudur. Böylece $\hat{A}$ ve $\hat{B}$ gibi iki regresyon katsayıları seti arasındaki ilişkiler ise aşağıdaki gibi yazılmaktadır:

$$\hat{A} = P' \hat{B}$$

ve

$$\hat{B} = P' \hat{A}$$

$\hat{A}$ 'nın ilgili elementi sıfıra eşitlenerek ilgili temel bileşen analizden çıkartılabilir. Böylece PC regresyonu aşağıdaki adımlarla özetlenebilir:

X matrisi için PC analizi uygulanarak temel bileşenler (Z) elde edilir.

$\hat{A}$ katsayılar matrisinin en küçük kareler tahminlerini elde etmek için Y ile Z arasında EKK regresyonu uygulanır.

$\hat{A}$ 'nın en son elementi sıfıra eşitlenir.

Son aşamada $\hat{B} = P\hat{A}$ eşitliği kullanılarak hesaplanan katsayılar orijinal ölçeğine dönüştürülmektedir. (Hawkins, D.M. 1973)

Son olarak, RR tekniğinde k yanlışlık sabitinin seçiminde yaşanan belirsizliğin aksine, PC analizinde ortadan kaldırılacak temel bileşenlerin sayısı göreceli olarak daha kesindir. Diğer bir anlatımla çoklu doğrusal bağlantı sorununu çözmek için özdeğeri göreceli olarak küçük olan bileşenlerin analizden çıkartılması gerektiği hususu daha kesindir. Görüldüğü gibi PC tekniğinde, RR tekniğindeki k yanlışlık sabitinin seçimi gibi bir belirsizlik söz konusu değildir. (David G. T., Denison, 2002) (David G. T., Denison, 2002)

8. UYGULAMA

8.1 Türkiye İthalatının Ridge Regresyon ile Analizi

Bu bölümün amacı gerçek bir veri analizinin ridge regresyon ile yapılmasını bir örnek olarak göstermektir.

Gerçek bir veri seti kullanılarak elde edilen deneysel sonuçlardan ekonomik teori için önemli bir sonuç elde edilmemiştir.

Bu bölüm, tahmin işleminde kullanılan veri ve modelin tartışılması, çoklu bağlantının ölçülmesi ve en küçük kareler metodunun geliştirilmesi, ridge regresyon tahminleri, ridge izi k 'nın seçim metodunun sunulması, temel bileşenler regresyon analizinin yapılması, zaman serileri analizinin yapılması ve modelin tahmin edilmesi şeklinde işlenmiştir.

8.2 Veri ile İlgili Bilgi

Bu uygulama Türkiye ekonomisindeki ithalat hareketi ile ilgili toplanılan verileri temel alan bir örnek üzerinde yapılmıştır. Bağımlı değişken ve iki açıklayıcı değişkenler arasında doğrusal bir ilişkinin bulunduğu varsayılmıştır. Hata terimi olan u 'nun rastlantısal bir sayısal değer ve ortalaması sıfır, varyansı σ^2 olan bir terim olduğu düşünülmektedir. Değişkenler aşağıdaki tabloda görülmektedir.

Çizelge 8.1 Değişkenler tablosu

Değişkenler		
Değişken	Sembol	Tanım
İthalat	RIMP	İthalat 1985 fiyatları ile Dolar
Gelir	RGDP	Gayri safi yurt içi hasıla (GSYİH)
Nisbi Fiyat (Relative Prices)	RP	İthalat Birim Değeri / GSYİH Deflatörü 1985 =100

Nisbi Fiyat (RP) değişkeninin formülü şu şekilde tanımlanmaktadır:

$$\text{Nisbi Fiyat} = \frac{\text{İthalat Birim Değeri}}{\text{GSYİH Deflatörü}}$$

Standartlaştırılmamış değişkenleri ele aldığımızda modeller:

$$\text{Model : } \log RIMP = \beta_0 + \beta_1 \log RGDP + \beta_2 \log RP + u$$

ve standartlaştırılmış değişkenler için ise:

$$\text{Model : } \log SRIMP = \beta_1 \log SRGDP + \beta_2 \log SRP + u$$

Standartlaştırılmış her bir z değişkenin i. gözlemi, standartlaştırılmamış x değişkeninden türetilmiştir ve aşağıdaki gibi gösterilir.

$$Z_i = \frac{x_i - \bar{x}}{\sqrt{\frac{1}{n-1} \sum (x_i - \bar{x})^2}} \quad (8.1)$$

İktisat teorisine göre, açıklayıcı değişken olan gayri safi yurt içi hasıla ithalatı pozitif olarak etkilemelidir. Bunun aksine, ithalat nispi fiyatın (RP) azalışını arttırmalıdır. Bunların matematiksel olarak ifade edilişi:

$$\frac{\partial RIMP}{\partial RGDP} > 0, \quad \frac{\partial RIMP}{\partial RGDP_{-1}} > 0, \quad \frac{\partial RIMP}{\partial RP} < 0, \quad \frac{\partial RIMP}{\partial RP_{-1}} < 0$$

şeklindedir. (William H. Greene, 2000)

Açıklayıcı değişkenlerin katsayıları olan β_i 'ler x_i 'lerin formülleri:

$$\tilde{\beta}_i = \beta_i \frac{S_{RIM}}{S_{x_i}} \quad (8.2)$$

Çizelge 8.2 Türkiye ekonomisinin yıllık verileri; ithalat verileri (\$)

			UVIMP		
	RGSYİH (RGDP)	İthalat (IMP)	(İthalat Birim Değeri)	GSYİH Deflatörü (GDP Deflator)	Nisbi Fiyat (RP)
1985-01	159,058	112,50	86,6	54,987	2,0509381
1985-02	160,579	113,10	87,3	55,068	2,0588244

1985-03	163,570	113,59	87,8	55,397	2,055472
1986-01	175,148	110,74	85,9	74,985	1,4818287
1986-02	174,298	110,93	86,3	75,069	1,4827072
1986-03	178,695	111,51	86,9	75,364	1,4846189
1987-01	198,254	139,68	89,6	99,014	1,4157096
1987-02	197,546	140,92	90,0	99,947	1,4149473
1987-03	198,731	141,96	90,3	101,014	1,4103795
1988-01	208,579	142,60	89,6	164,977	0,8693387
1988-02	209,346	143,10	90,4	165,164	0,8714116
1988-03	211,163	143,53	90,8	165,397	0,8727908
1989-01	217,687	156,98	91,6	287,345	0,5513119
1989-02	218,349	156,39	92,4	288,346	0,5473692
1989-03	218,157	157,98	92,6	288,294	0,5529823
1990-01	245,346	222,67	94,7	468,970	0,4798065
1990-02	247,564	223,03	95,0	469,641	0,4798947
1990-03	246,364	223,01	95,6	468,123	0,4813919
1991-01	256,479	209,57	84,9	768,476	0,2777086
1991-02	257,468	210,03	85,0	769,246	0,2780336
1991-03	258,987	210,98	85,9	769,648	0,2791253
1992-01	278,452	227,89	88,7	1269,547	0,184505
1992-02	277,157	228,06	89,1	1270,342	0,1845265
1992-03	279,845	228,97	89,8	1270,856	0,1851699
1993-01	308,493	293,82	91,6	2138,641	0,1423863
1993-02	309,012	293,94	92,6	2139,534	0,1423851
1993-03	309,678	294,59	92,8	2139,103	0,1427166
1994-01	299,657	231,90	86,7	4394,301	0,0577729

1994-02	300,579	231,89	87,6	4394,398	0,0577695
1994-03	300,999	232,97	87,7	4394,891	0,0580093
1995-01	327,571	356,84	115,1	8095,348	0,0490796
1995-02	328,627	356,92	116,9	8096,001	0,049086
1995-03	328,942	357,33	116,4	8096,649	0,0491331
1996-01	356,238	435,61	103,6	13666,314	0,0368747
1996-02	357,158	436,12	103,9	13677,697	0,0368855
1996-03	357,247	436,34	104,8	13677,035	0,0369031
1997-01	389,579	484,98	94,7	25498,971	0,0240196
1997-02	390,579	485,31	94,9	25499,001	0,0240325
1997-03	390,987	485,96	95,1	25499,961	0,0240573
1998-01	406,981	458,70	90,8	46094,015	0,0149513
1998-02	407,413	459,10	91,1	46095,008	0,0149599
1998-03	407,589	459,36	91,2	46095,697	0,0149654
1999-01	392,567	405,68	85,9	75192,698	0,0103952
1999-02	393,873	405,98	86,0	75193,069	0,0103992
1999-03	393,994	406,89	86,3	75193,290	0,0104113
2000-01	431,573	544,90	89,5	107622,971	0,010063
2000-02	432,015	544,23	89,9	107623,137	0,0100568
2000-03	432,195	545,30	90,4	107623,460	0,0100667
2001-01	408,579	412,58	88,1	171214,638	0,0074097
2001-02	408,983	412,97	88,9	171215,305	0,007412
2001-03	409,957	414,06	90,0	171215,148	0,0074184
2002-01	447,967	514,09	87,6	234701,569	0,0071904
2002-02	448,104	514,93	88,1	234702,978	0,007194
2002-03	448,998	515,98	88,7	234702,002	0,0071984

2003-01	482,900	692,83	99,9	285849,002	0,0074238
2003-02	483,103	693,30	99,2	285850,301	0,0074254
2003-03	483,981	693,89	100,9	285850,108	0,0074275
2004-01	528,348	972,51	114,9	314801,697	0,0080893
2004-02	528,512	973,40	115,1	314802,231	0,0080921
2004-03	529,891	973,81	116,0	314802,458	0,0080934
2005-01	569,540	1164,81	123,9	339985,147	0,0084261
2005-02	569,503	1164,90	122,7	339986,200	0,0084263
2005-03	570,997	1165,96	123,9	339986,648	0,0084294

Kaynak: International Monetary Fund, TCMB Dış Ticaret İstatistikleri, TÜİK Veritabanı

Az önce değinildiği gibi iktisat teorisine göre, tahmin edilecek modellerde katsayılardan; RGDP'nin katsayılarının artı işaretli ve RP'nin katsayılarının eksi işaretli olmaları beklenir.

Modeldeki β_1 ve β_2 katsayılarının, her biri ayrı ayrı olmak şartı ile gelir ve fiyat üzerinde elastikiyetleri vardır. Bu ifade matematiksel olarak:

$$E_{RIMP, RGDP} = \frac{\partial \log RIMP}{\partial \log RGDP}$$

$$E_{RIMP, RP} = \frac{\partial \log RIMP}{\partial \log RP}$$

şeklinde de gösterilir. (Sims, C. 1980)

International Monetary Fund, TCMB ve TÜİK web sitelerinden elde edilmiş olan, Çizelge 8.2'de verilen veriler kullanılarak analizler yapılmıştır. Yukarıdaki zamana endeksli veriler, bağımlı değişken olan İthalat (IMP), GSYİH, RGSYİH ve UVIMP açıklayıcı değişkenlerin 1985 ile 2005 aralığındaki 4 aylık dönemler halinde elde edilmiş olan değerlerinden oluşmaktadır.

8.3 Çoklu Bağlantının Ölçülmesi ve Betimsel İstatistikler

Bütün değişkenlerin betimsel istatistikleri aşağıdaki çizelgede görülmektedir.

Çizelge 8.3 Betimsel istatistikler

		IMP	RGDP	RP
N	Değer	63	63	63
Ortalama		395,6343	338,6851	0,237536
Ortanca		357,09	328,071	0,014429
Standart Sapma		277,6455	117,8039	0,443613
Minimum		111,05	161,069	0,00035
Maksimum		1165,37	570,748	1,58509
Yüzde	25	189,815	243,742	0,000761
	50	357,09	328,071	0,014429
	75	507,645	424,955	0,301919

Çoklu bağlantının varlığı araştırılmadan önce değişen varyanslılık varsayımı White Testi ile test edilmiştir. P anlamlılık derecesi 0,29 olarak elde edilir. %95 güvenirlilik ile yapılan bu çalışmada $p > \alpha$ olduğu için değişen varyanslılık olmadığı söylenir.

Bağımlı değişken ile açıklayıcı değişkenler arasındaki korelasyon tablosu ise:

Çizelge 8.4 Korelasyon tablosu

Korelasyon			IMP	RGDP	RP
Spearman's rho	IMP	Correlation Coefficient	1,0000	0,979	-0,920
		Sig. (2-tailed)	.	0,0000	0,0000
		N	63	63	63
	RGDP	Correlation Coefficient	0,979	1,0000	-0,964
		Sig. (2-tailed)	0,0000	.	0,0000
		N	63	63	63

	RP	Correlation Coefficient	-0,920	-0,964	1,0000
		Sig. (2-tailed)	0,0000	0,0000	.
		N	63	63	63

Değişkenler arasındaki en küçük ilişki miktarı %92 olarak tabloda görülmektedir. İthalat ile Reel Gayri Safi Yurtiçi Hasıla arasında %97,9'luk pozitif yönde bir ilişki bulunurken, Reel fiyatla arasında %92'lik negatif yönde bir ilişki bulunmaktadır. Reel fiyat ile Reel Gayri Safi Yurtiçi Hasıla arasında ise %96,4'lük pozitif yönde güçlü bir ilişki bulunmaktadır.

Korelasyonun özdeğerlerine baktığımızda:

$$\lambda_1 = 2,756 \quad \lambda_2 = 0,243 \quad \lambda_3 = 1,1178.10^{-4}$$

Bulduğumuz özdeğerler çoklu bağlantının seviyesini açıklamak için yardımcı olacaklardır.

Regresyon katsayılarının tahmininde çoklu bağlantının belirlenmek için tahmin ediciler arasındaki yüksek ilişkinin önemli olduğu bilinmektedir. Çoklu bağlantıyı en iyi belirleme yöntemlerinden birisi ise VIF (varyans büyütme etkeni)değerlerine bakmaktır.

$$VIF_{LGRGDP} = 133,971 \quad VIF_{LGRP} = 133,971$$

VIF değerleri 10'dan büyük ise çoklu bağlantının varlığından söz edilebilir. Yukarıdaki VIF değerlerine bakıldığında üzerinde çalışılan verilerde bir çoklu bağlantı sorunu olduğunu söylenebilir.

8.4 Ridge Regresyon (RR) ve En küçük kareler (LS) Yöntemi ile Elde Edilen Sonuçlar

Açıklayıcı değişkenler arasında güçlü doğrusal bir ilişki var ise, regresyon katsayılarının en küçük kareler ile tahminin değeri kararsız bir yapıya sahip olmaktadır. Çoklu bağlantının derecesi ve en küçük kareler tahmin metodundaki tahmin edicinin gücü model için regresyon sonuçlarını gösterir. En küçük kareler uygulandığında ithalat ile açıklayıcı değişkenler arasındaki doğrusal ilişkinin %99 olduğu ve ithalatla meydana gelen değişimlerin yaklaşık %98,1'inin bağımsız değişkenler tarafından açıklandığı anlaşılmaktadır [Çizelge 8.5].

Ayrıca EKK tekniğiyle elde edilen ANOVA tablosu incelendiğinde %5 veya %1 geleneksel anlamlılık düzeylerinde modelin anlamlı (güvenilir) olduğu ($F=1524,60$; $p=0,000$)

görülmektedir [Çizelge 8.6]. Modelin yüksek belirlilik ($R^2=\%98,1$) ve korelasyon katsayısının ($R=\%99$) aksine katsayılar tablosu incelendiğinde hiçbir değişkenin kısmi t testine göre anlamlı olmadığı görülmektedir [Çizelge 8.7]. Bu durum daha önce belirtildiği gibi, çoklu doğrusal bağlantı probleminin bir göstergesidir.

Aynı sonuca çoklu doğrusal bağlantı ile ilgili istatistiklerin incelenmesiyle de varılmaktadır [Çizelge 8.8]. burada bağımsız değişkenler için hesaplanan korelasyon matrisi ve VIF değerleri verilmektedir [Çizelge 8.7]. Korelasyon matrisi incelendiğinde X1 ve X2 değişkenleri arasında çok yüksek bir ilişki ($\%98,0$) görülmektedir. X1 ve X2 değişkenleri için hesaplanan VIF'ler sırasıyla yaklaşık olarak 133,971 ve 133,971 olarak elde edilmiş ve bu değerler kritik VIF değerinin (10) oldukça üstünde olduğu görülmektedir.

Hesaplanan korelasyonlara ait özdeğerlerden yararlanarak elde edilen koşul sayısı kritik değer olan 1000'den büyüktür [Çizelge 8.8]. Bu istatistiklerin tamamı verilerde çok güçlü bir doğrusal bağlantı problemi olduğunun bir kanıtıdır.

Ridge regresyon analizi sonuçlarının birinci kısmında belirli yanlılık katsayılarına göre R^2 , standartlaştırılmış ridge regresyon katsayıları verilmektedir [Çizelge 8.13]. Bu verilerden yararlanarak optimum k sabiti standartlaştırılmış ridge regresyon katsayılarını durağanlaşmaya başladığı ve bu katsayılara ait VIF değerlerinin birlikte 1'e yaklaştığı noktada ($k=\%2$) seçilir. Yanlılık sabiti (k) ile yanlı standartlaştırılmış regresyon katsayıları arasındaki grafiğe bakıldığında çok küçük ($k=0,02$) bir yanlılık sabitinden sonra regresyon katsayılarının durağan hale geldiği görülmektedir. [Şekil 8-1]. Yanlılık sabiti ile modelin R^2 değerleri arasındaki ilişki incelendiğinde k, 0 ile 1 aralığında değişim gösterirken; R^2 , $\%84,17$ ile $\%97,36$ aralığında değişmektedir [Çizelge 8.13]. Logaritmik ölçekli grafikte standartlaştırılmış ridge regresyon katsayılarının nasıl durağanlaştığı çok açık bir şekilde görülmektedir. [Şekil 8-1].

EKK tekniğinde olduğu gibi RR sonuçları da $\%5$ ve $\%1$ anlamlılık düzeylerinde anlamlı ($F=248,496$; $p=0,000$) olduğu görülmektedir [Çizelge 8.10]. Diğer bir anlatımla ithalat (Y)'daki değişim, açıklayıcı değişkenler tarafından anlamlı bir şekilde açıklanmaktadır. Beklendiği gibi tahmincilerin standart hataları da küçülmüştür. [Çizelge 8.10].

EKK tekniği ile yanlı tahmin tekniklerinin sonuçları arasındaki en önemli çelişki ise, X2 değişkenine ait regresyon katsayısının işaretinde görülmektedir. Diğer bir anlatımla, EKK tekniğinde kuramsal beklentilerin aksine ithalat (Y) ile nisbi fiyat (X2) değişkeni arasında pozitif yönlü bir ilişki söz konusu iken, yanlı tahmin tekniği olan ridge regresyon ile elde

edilen sonuçta kuramsal beklentilere uygun (negatif) yönlü bir ilişki elde edilmiştir. Görüldüğü gibi yanlış tahmin teknikleri kuramsal yaklaşımı doğrulamaktadır.

En küçük kareler (LS) regresyonu ve ridge regresyon (RR) teknikleriyle elde edilen sonuçlar Çizelge 8.14’te özetlenmektedir. Çizelge 8.14 dikkatle incelendiğinde yanlış tahmin tekniğiyle (RR) elde edilen sonucun en küçük kareler (LS) tekniğiyle elde edilen sonuçlar arasında bazı tutarsızlıkların ve katsayı büyüklüklerinin de önemli düzeyde farklılıkların olduğu açık bir şekilde görülmektedir.

Çizelge 8.5 Model özeti (LS)

R	R-Kare	Düzeltilmiş R Kare	Tahminin Standart Hatası	Durbin-Watson(d)
0,990	0,981	0,980	0,09350	0,658

Çizelge 8.6 ANOVA (LS)

Kaynak	SS	DF	MS	F	p
Regresyon	4,848	2	2,424	1524,600	0,000
Hata	0,078	60	0,0013		
Toplam	4,926	62			

Çizelge 8.7 Tahmin edilen regresyon katsayıları ve güven aralıkları

	B	SH	Beta	t	p	b için %95 Güven Aralığı	
						Alt	Üst
Sabit	-269,548	182,146		-1,665	0,179	-351,256	-1752,005
X1	15,719	11,349	1,511	1,253	0,205	2,351	128,537
X2	3,004	1,031	0,482	1,257	0,129	-3,007	15,269

Çizelge 8.8 EKK çoklu doğrusal bağlantının saptanması

	Korelasyon Matrisi		
	X1	X2	VIF
X1	1,000	0,983	133,971
X2	0,983	1,000	133,971

Çizelge 8.9 Korelasyonların özdeğerleri

No	Özdeğer	Koşul sayısı	Varyans Oranları		
			Sabit	X1	X2
1	2,756	1,000	0,000	0,000	0,001
2	0,243	3,365	0,000	0,000	0,06
3	0,000	1570,31	1,000	1,000	0,94

Çizelge 8.10 k=%2 Yanlılık sabitiyle ridge regresyon sonuçları

Çoklu R	0,983
R-Kare	0,978
Düzeltilmiş R-Kare	0,975
Standart Hata	0,069

Çizelge 8.11 ANOVA tablosu

Kaynak	DF	SS	MS
Regresyon	2	1,546	0,773
Hata	60	0,08	0,0013
F Değeri	265,386		
F-Anl.(p)	0,000		

Çizelge 8.12 Modeldeki değişkenler

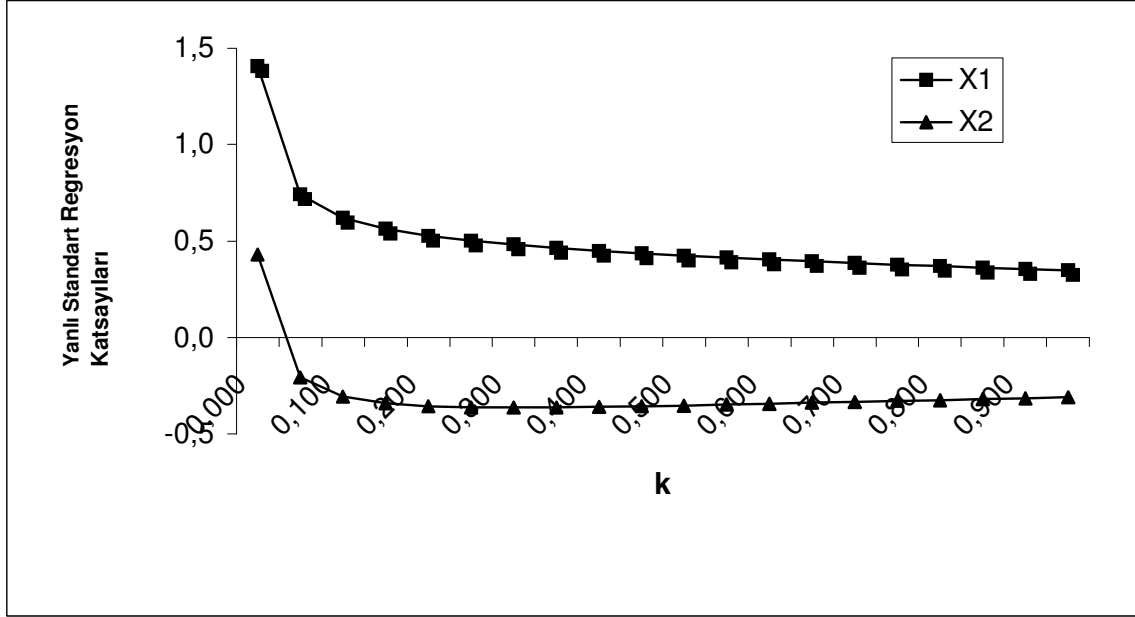
Değişken	B	SH	Beta	B/SH(B)	VIF
Sabit	-19,745	20,348	0,000	-0,970	
X1	4,518	1,567	0,971	2,883	1,478
X2	-0,003	0,069	-0,015	0,0434	1,478

Ridge Regresyon (RR) Analizi

Çizelge 8.13 k, R-Kare ve standartlaştırılmış ridge regresyon katsayıları

k	R- Kare	Standartlaştırılmış Ridge Regresyon Katsayıları	
		X1	X2
0	0,9807	1,491644	0,523053
0,05	0,95697	0,827729	-0,11692
0,1	0,94439	0,676695	-0,24517
0,15	0,93647	0,604976	-0,29517
0,2	0,93002	0,560538	-0,3189
0,25	0,92402	0,528858	-0,33079
0,3	0,91809	0,504271	-0,33647
0,35	0,91208	0,484102	-0,33854
0,4	0,90592	0,466917	-0,33839
0,45	0,8996	0,451875	-0,33681
0,5	0,89312	0,438445	-0,33429
0,55	0,8865	0,426277	-0,33115
0,6	0,87976	0,415125	-0,32758
0,65	0,87291	0,404814	-0,32373
0,7	0,86598	0,395212	-0,3197
0,75	0,85898	0,386219	-0,31557
0,8	0,85193	0,377756	-0,31137

0,85	0,84486	0,369761	-0,30716
0,9	0,83776	0,362183	-0,30296
0,95	0,83066	0,354978	-0,29879
1	0,82357	0,348113	-0,29466



Şekil 8.1 Logaritmik ölçekli ridge regresyon grafikleri (ridge izi)

Çizelge 8.14 EKK- RR sonuç karşılaştırma tablosu

	B		SH		Beta		B/SH(B)		VIF	
	EKK	RR	EKK	RR	EKK	RR	EKK	RR	EKK	RR
Sabit	-269,548	-19,745	182,146	20,348	0,00	0,000	-1,479	-0,970		
X1	15,719	4,518	11,349	1,567	1,511	0,971	1,385	2,883	133,971	1,478
X2	3,004	-0,003	1,031	0,069	0,482	-0,015	2,913	0,0434	133,971	1,478

$$\text{EKK : } y_i = -269,548 + 15,719x_1 + 3,004x_2 \quad (8.3)$$

$$R^2 = 0,981 \quad F=1524,600 \quad SH=0,09350$$

$$\text{RR : } y_i = -19,745 + 4,518x_1 - 0,003x_2 \quad (8.4)$$

$$R^2 = 0,978 \quad F=265,386 \quad SH=0,069$$

Bağımsız değişkenlerin çoklu bağlantı ile ilgili istatistikleri incelendiğinde verilerde çok güçlü doğrusal bağlantı probleminin olduğu anlaşılmaktadır. Bu nedenle EKK tekniğiyle elde edilen sonuçlar geçerliliğini kaybetmektedir. Diğer bir anlatımla, çoklu doğrusal bağlantılı verilerde regresyon katsayılarının standart hataları, büyüklükleri ve işaretleri uygun bir biçimde tahmin edilememektedir. Verilerde çoklu doğrusal bağlantı olması durumunda tahmin edilememektedir. Verilerde çoklu bağlantı olması durumunda RR tekniği EKK tekniğine göre daha durağan ve kuramsal (veya ampirik) beklentilere uygun sonuçlar sağlamaktadır. Her ne kadar EKK ve yanlı tahmin tekniklerinden (RR) birisinin seçimi, yanlı veya yansız tahmincilerden birisinin seçimi anlamına gelse de, gerçekte durum bu değildir. Bilindiği gibi, pratik anlamda, EKK tahmincileri sadece modelin hatasız tanımlanması durumunda yansızdırlar. Bu nedenle pratikte EKK tahmincilerinin genelde yanlı olacağı kabul edilmektedir. Kısaca yanlı tahmin teknikleriyle, çoklu doğrusal bağlantı sorununu azaltmak amacıyla, birbirleriyle anlamlı ilişki içinde olan açıklayıcı değişkenler birlikte analiz edilmektedir.

Araştırmamızda çoklu doğrusal bağlantı probleminin bir sonucu olarak tahmincilerin standart hatalarının yüksek ve X_2 değişkenine ait katsayının işareti kuramsal ve ampirik beklentilerle çeliştiği görülmektedir. Ayrıca regresyon katsayılarının büyüklükleri de çoklu doğrusal bağlantıdan olumsuz bir şekilde etkilenmektedir. Bu olumsuzlukları ortadan kaldırmak için uygulanan yanlı tahmin tekniklerinin birbirleriyle tutarlı ve kuramsal beklentilere uygun sonuçlar verdiği görülmektedir. Optimum yanlılık sabitini araştırmak amacıyla ridge izi grafiğinden yararlanarak, yanlı regresyon katsayılarının durağanlaştığı bölgede yaklaşık bir k değeri seçilerek iterasyonlara başlanmaktadır. İterasyonlar sonucunda RR için seçilen optimum 0,02 yanlılık sabiti ($k=2$) ile elde edilen sonuçlarda X_2 değişkenine ait katsayının işareti değişmekte ve tahmincilerin standart hataları küçülmektedir. Tahminler yanlı olduğundan RR tekniği için F testi geçerli değildir. Burada F istatistiği sadece bir endeks görevi gördüğünden ihtiyatlı bir şekilde yorumlanmalıdır.

Sonuç olarak, ithalatı açıklayan değişkenler arasında çoklu doğrusal bağlantı olduğundan,

Ridge Regresyon En küçük kareler tekniğine göre daha tutarlı, geçerli, durağan ve kuramsal beklentilere uygun tahminler sağladığı görülmektedir.

8.5 Zaman Serisi Uygulaması

Bu çalışmada kurulan VAR modelinde yer alan değişkenlerdeki durağanlığın tespiti Geliştirilmiş Dickey Fuller (ADF) birim kök testi ile gerçekleştirilmiştir. VAR analizindeki değişkenler durağan oldukları seviyede kullanılmıştır.

ADF testi sonuçları aşağıdaki tabloda verilmiştir:

Çizelge 8.15 Geliştirilmiş Dickey – Fuller (ADF) birim kök test sonuçları

Değişkenler	Sabit			Sabit ve Trendli		
	ADF	Gecikme Sayısı	Kritik Değer	ADF	Gecikme Sayısı	Kritik Değer
İTHALAT	-8,308	1	%1 -3,542	-8,778	1	%1 -4,115
RGSYİH (RGDP)	-9,361	1	%1 -3,542	-7,797	1	%1 -4,118
RP	-18,772	0	%1 -4,348	-9,538	0	%1 -5,743

RGSYİH:Reel Gayri Safi Yurt İçi Hasıla

RP:Nispi Fiyat

Çizelge 8.15’den de görüldüğü gibi ithalat ve reel gayri safi yurt içi hasıla (RGSYİH) değişkenlerinin durağan olmadığı, nispi fiyat (RP)’nin durağan olduğu belirlenmiştir. Bu durumda durağan olmayan değişkenlerin birinci farkları alınıp, durağanlık testi trendsiz modele göre yapılmıştır.

Değişkenlere ilişkin serilerde ithalat ve RGSYİH serileri aynı seviyede durağan oldukları için eşbütünleşme analizi bu iki seri arasında yapılmıştır. Ancak nispi fiyat (RP) serisi için eşbütünleşme analizinin gerçekleştirilememiş olması nispi fiyat (RP) ile değişkenler arasında uzun dönem ilişkinin olmadığını göstermektedir.

Seriler arasında k=2 gecikme için Granger nedenselliği ile ilgili değerler aşağıdaki tabloda verilmektedir. Değişkenlerin karşılıklı Granger nedenselliği taşımadıkları görülmektedir. Değişkenler arasında çift yönlü nedensellik ilişkisi bulunmamaktadır. Bundan dolayı

değişkenler arasında kısa dönem ilişki olmadığı söylenebilir.

Çizelge 8.16 Granger nedensellik testi

Sıfır Hipotezi	F	p
RGSYİH İthalatın Granger nedeni değildir.	1,5812	0,193
İthalat RGSYİH'in Granger nedeni değildir.	0,6257	0,646
RP İthalatın Granger nedeni değildir.	0,0406	0,996
İthalat RP'in Granger nedeni değildir.	0,0881	0,985

Çizelge 8.17 Eşbütünleşme (Koentegrasyon) testi

Öz değer (λ_i)	Ho: rank=p	$\lambda_{\max}$	p	95%	λ_{trace}	p	95%
0.969471**	$r = 0^{**}$	6.534.147**	0.0000**	20,51423	8.011.143**	0.0000**	28,96402
0.424136	$r \leq 1$	1.042.167	0.1785	13,92601	1.005.284	0.2130	14,99113
0.003145	$r \leq 2$	0.065704	0.7331	3,41717	0.02495	0.6947	3,51493

Koentegre vektör sayısını bulmak için yapılan koentegrasyon testinde koentegre vektörün olmadığı Çizelge 8.17’de görülmektedir. Bu durum değişkenler arasında uzun dönem ilişkinin olmadığını göstermektedir.

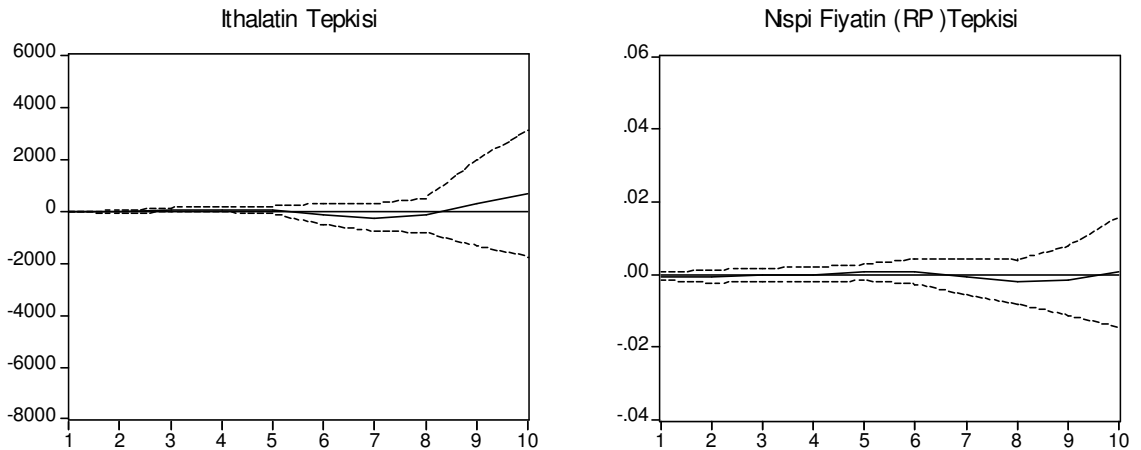
VAR analizinde ise, değişkenler durağan oldukları seviyelerde, yani, reel gayri safi yurt içi hasıla ile ithalat birinci farklarında, nispi fiyat ise seviye değerlerinde kullanılmıştır. VAR modelinde uygun gecikme uzunluğu belirlenirken LR (Log Likelihood), AIC (Akaike Information Criterion) ve SIC (Schwarz Information Criterion) test sonuçlarından faydalanılmıştır. Bu test sonuçlarına göre LR, AIC ve SIC dört gecikme uzunluğunu %5 anlam düzeyinde vermişlerdir.

VAR analizi sonucunda elde edilen denklem aşağıda verilmiştir:

$$\begin{aligned} \text{İthalat} = & 1,0147 \text{ İthalat}(-1) - 0,02302 \text{ İthalat}(-2) - 0,66759 \text{ RGDP}(-1) + 0,81939 \text{ RGDP}(-2) \\ & - 1,54705 \text{ RP}(-1) + 4,90266 \text{ RP}(-2) - 23,466 \end{aligned} \quad (8.5)$$

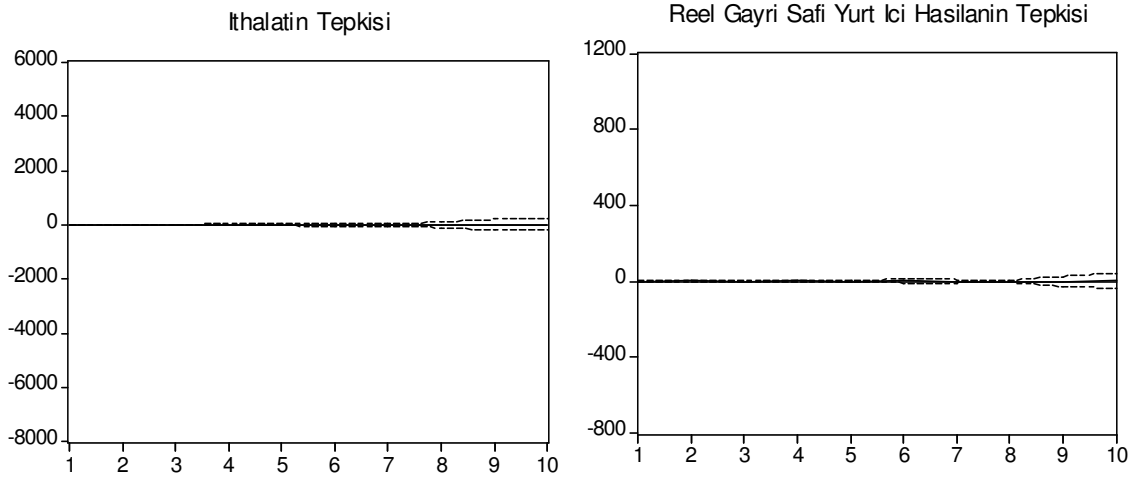
Etki-tepki fonksiyonları, sistemde yer alan değişkenlerden birine bir birimlik şok uygulandığında; diğer değişkenlerin bu değişime gösterdikleri tepkiyi verir. İktisadi değişkenlerin maruz kaldıkları şoklar genellikle iki kategoride toplanabilir. Bir kısım şoklar “kalıcı” (permanent) bir karaktere sahiptir ve iktisadi değişkenler üzerindeki etkileri uzun süre devam eder. Diğer bir deyişle, bu tür bir şok sisteme girdiği anda sadece o dönemi değil daha sonraki uzunca bir dönemi etkileyebilmektedir. Teknolojik nitelikteki şoklar bunlara iyi bir örnek oluşturur. Geçici (transitory) nitelikteki diğer bazı şoklar ise sisteme girdikleri anda önemli bir etkiye sahip olmakla birlikte bu etki kalıcı bir niteliğe sahip değildir ve takip eden birkaç dönem içinde etkileri yok olur.

Tahmin edilen etki-tepki katsayıları çerçevesinde güven sınırlarının da yer alması, etki-tepki fonksiyonlarının istatistiksel olarak anlamlı olup olmadığının analiz edilebilmesi için gereklidir. Diğer bir ifade ile değişkenlerin birbirlerindeki şoklara gösterdikleri tepkiler, güven sınırlarında anlamlı (anlamsız) ise, ithalatı azaltmayı amaçlayan reel gayri safi yurt içi hasıla ve/ veya nispi fiyat değişimleri etkin değildir (etkindir) denir. Runkle (1987) etki-tepki fonksiyonlarının istatistiksel olarak anlamlı olup olmadığının analiz edilmesi için güven sınırlarının oluşturulması gerektiğini vurgulamaktadır. İstatistiksel bakımdan anlamlı ilişkiler sunan ve amaç açısından da önem arz eden etki-tepki fonksiyonlarına ait grafiklerden bazıları aşağıda verilmiştir. Grafiklerdeki düz çizgi nokta tahminlerini, kesikli çizgiler ise bir standart hatalık güven sınırlarını ifade etmektedir.



Şekil 8.2 Reel Gayri Safi Yurt İçi Hasılaya (RGDP) bir standart hatalık şok karşısında gösterilen tepkiler

Şekil 8.2, reel gayri safi yurt içi hasıla serisindeki bir standart hatalık şok karşısında, ithalat ve nispi fiyat serilerinde meydana gelen değişimleri göstermektedir. Buna göre, reel gayri safi yurt içi hasıla serisindeki bir standart hata kadar şoka, her iki sektöründe yaklaşık ilk beş dönem boyunca tepkisi belirsiz kaldığı daha sonra negatif ve pozitif yönlerde ufak etkiler yaptığı, güven sınırları dikkate alındığında ise etkinin hiçbir dönemde anlam kazanmadığı görülmektedir.



Şekil 8.3 Nispi Fiyata (RP) bir standart hatalık şok karşısında gösterilen tepkiler

Şekil 8.3, nispi fiyat serisindeki bir standart hatalık şok karşısında reel gayri safi yurt içi hasıla ve ithalat serilerinin gösterdiği tepkiler görülmektedir. Buna göre güven sınırları dikkate alındığında sektörler tarafından uygulanan bir birimlik şoka reel gayri safi yurt içi hasıla ve ithalatın hiçbir dönemde anlamlı bir tepki vermediği görülmektedir.

Çizelge 8.18 İthalat için ortalama varyans ayrıştırması

Değişim Kaynakları			
Değişim	İthalat	Gayri Safi Yurt İçi Hasıla	Nispi Fiyat
İthalat	87,3	13,71	0,015
Gayri Safi Yurt İçi Hasıla	85,17	12,61	0,071
Nispi Fiyat	64,48	7,5	27,78

Varyans ayrıştırması, bir değişkende meydana gelen değişmelerin kaynaklarının araştırılmasında kullanılan bir yöntemdir. Yapılan varyans ayrıştırması analizinin sonucunda ise ithalat incelendiğinde, on dönem boyunca da %87,3 gibi oranlarda kendi kendini açıkladığını, diğer değişkenlerdeki değişimin de ithalat üzerinde önemli etkisinin olduğu görülmektedir.

Eş bütünleşme ve VAR analizi sonucunda elde edilen bulgular, ithalat ile Reel Gayri Safi Yurt İçi Hasıla ve Nispi Fiyat arasında bir nedensellik ilişkisi göstermemektedir. Yapılan durağanlık testi sonucunda sadece iki değişken aynı seviyede durağan olduğu için, eş bütünleşme analizi bu iki değişken arasında yapılmıştır ve değişkenler arasında uzun dönem ilişkinin olmadığı saptanmıştır. VAR analizinde ise değişkenler durağan oldukları seviyelerde, yani, nispi fiyat seviye değerinde, ithalat ve reel gayri safi yurt içi hasıla ise birinci farklarında kullanılmıştır. VAR modelinde uygun gecikme uzunluğu belirlenirken LR, AIC ve SIC test sonuçlarından faydalanılmıştır. Bu test sonuçlarına göre LR, AIC ve SIC dört gecikme uzunluğunu %5 anlam düzeyinde vermişlerdir. Granger nedensellik test sonucuna göre herhangi bir nedensellik ilişkisi saptanamamıştır.

VAR analizinden elde edilen etki-tepki fonksiyonları ve varsayım ayrıştırmaları söz konusu seriler arasında uzun dönemli bir etkileşimin söz konusu olmadığına işaret etmektedir. Tahmin edilen etki-tepki katsayıları çerçevesinde, etki-tepki fonksiyonlarının istatistiksel olarak anlamlı olması için güven sınırları da çizilmiş, ancak değişkenlerin birbirlerindeki bir standart hatalık şoklara gösterdikleri tepkilerin anlamlı olmadığı görülmüştür. Varyans ayrıştırması analizinden elde edilen sonuçlarda, söz konusu değişkenlerde meydana gelen değişimlerin nerede ise tamamının kendisindeki şoklardan kaynaklandığı görülmektedir. Bu analizlerin ışığında, ne reel gayri safi yurt içi hasılanın ne de nispi fiyatın, ithalatın uzun dönem dengesi üzerinde bir etkinliği olmadığı sonucuna ulaşılmıştır.

8.6 Temel Bileşenler Regresyon Analizi Uygulaması

Temel Bileşenler (PC) Regresyon Raporu:

Çizelge 8.19 Regresyon katsayıları sonuçları

Değişken	$\hat{b}$	SH	$\hat{\beta}$	VIF
Sabit	125,925			
X1	1,014	0,096	0,4302	0,2848
X2	-201,493	19,094	-0,4302	0,2848

Çizelge 8.20 Varyans (ANOVA) analizi

Kaynak	DF	SS	MS	F	p
Sabit	1	9844907	9844907		
Model	2	3006779	1503390	55,6776	0,000000
Hata	60	1620102	27001,7		
Toplam	62	4626881	74627,11		
Standart Hata (SH)	164,3219				
R-Kare	0,6499				

Çizelge 8.21 LS, RR ve PC teknikleriyle elde edilen sonuçların karşılaştırılması

	B			SH			Beta			B/SH(B)			VIF		
	EKK	RR	PC	EKK	RR	PC	EKK	RR	PC	EKK	RR	PC	EKK	RR	PC
Sabit	-269,548	-19,745	125,925												
X1	15,719	4,518	1,014	11,349	1,567	1,511	0,971	0,947	-0,4302	1,385	2,883	5,525	133,971	133,971	0,284
X2	3,004	-0,003	-201,493	1,031	0,069	0,482	-0,015	-0,018	0,4302	2,913	-0,0434	-5,514	133,971	133,971	0,284

Çizelge 8.21 'de görüldüğü gibi yanlı tahmin tekniklerinin (Ridge Regresyon ve Temel Bileşenler Regresyon (PC)) sonuçları birbirini doğrulamaktadır. X2 değişkenine bakıldığında

İthalat ile aralarında kuramsal olarak negatif bir ilişki beklenirken EKK’de bunun aksine pozitif bir ilişki elde edilmiştir. Oysa Ridge regresyon ve Temel Bileşenler regresyon analizleri sonuçları birbirleri ile örtüşmektedir.

SONUÇ

Türkiye'nin ithalat verileri ile yapılan bu çalışmada kullanılan yöntemler; en küçük kareler, ridge regresyon, zaman serileri ve temel bileşenler regresyon yöntemleridir. Bu yöntemlerden ridge regresyon ve temel bileşenler regresyon yöntemleri, yanlış tahmin yöntemleridir.

Ridge regresyon, regresyon analizinde karşılaşılan ve "çoklu bağlantı" diye adlandırılan, açıklayıcı değişkenlerin bağımsızlık varsayımının bozulması sorunundan kurtulabilmek amacıyla geliştirilen bir regresyon yöntemidir. Çoklu bağlantı varlığının, regresyon katsayılarının en küçük kareler tahminleri üzerinde önemli etkileri mevcuttur. Bu sorunun en önemli etkisi, regresyon katsayılarının en küçük kareler tahminlerinin büyük varyansa sahip olmalarına yol açmasıdır. Çoklu bağlantı sorununu giderme yollarından biri, β 'nin yansız olma gerektirimini göz ardı etmektir. Ridge modeli, tahminlerin varyanslarını azaltabilmek için regresyon denklemindeki katsayıları yanlış olarak tahmin eder. Bu çalışmanın amacı, çoklu bağlantı sorununu inceleyip, çözüm yollarından biri olan ridge regresyon modelini ithalat verisine uygulamak ve ridge metodunun özelliklerini ve performansını EKK, zaman serileri, temel bileşenler regresyon ile karşılaştırmaktır. Çalışma kapsamına ridge izi, basit, genelleştirilmiş, sıralı, Bayesgil, kanonik ridge olmak üzere altı farklı ridge yaklaşımı alınmıştır. Her bir tahmin değeri büyüklük, işaret ve istatistiksel anlamlılık açısından karşılaştırılırken, modeller ise doğruluk dereceleri açısından karşılaştırılmıştır. Gerçek veri üzerinde yapılan uygulama sonucu, ridge yaklaşımı ve temel bileşen yaklaşımının bu veri için EKK ve zaman serilerinden üstün olduğu görülmüştür.

Sonuç olarak, ithalatı açıklayan değişkenler arasında çoklu doğrusal bağlantı olduğundan, Ridge Regresyon ve Temel Bileşenler Regresyon analizlerinin, En küçük kareler tekniğine göre daha tutarlı, geçerli, durağan ve kuramsal beklentilere uygun tahminler sağladığı görülmektedir.

KAYNAKLAR

- Anderson, T. W. , (2003) “An Introduction to Multivariate Statistical Analysis”, 3th ed.. -- Hoboken, N.J. : Wiley-Interscience
- Balakrishnan N., (2006) “Precedence-Type Tests and Applications”. John Wiley-Sons
- Bates, Douglas. M. , (1988) “Nonlinear Regression Analysis and its Applications”, John Wiley & Sons, Ltd, Newyork
- Cooley, T.F. and Roy, S.F., (1985) “A theoretical Macro econometrics: A Critique” Journal of Monetary Economics, v. 16, .283-308
- David Birkes, Y. Dodge, (1993) “Alternative Methods of Regression”, John Wiley & Sons, Ltd, Newyork, sf 178-188
- David G. T., Denision, (2002) “Bayesian methods for nonlinear classification and regression”, John Wiley & Sons, Ltd, Chichester,
- Doan, T., (1992) “RATS User’s Manual”. Evanston, III.: Estima
- Draper, N.R., H. Smith, (1981) “Applied Regression Analysis”, John Wiley & Sons, Ltd, Chichester, Sayı 2
- Ferreira. Eurico J.,Sirmans. G. Stacy (1988) “Ridge Regression in Real Estate Analysis”, The Apraisal Journal 56,3 sf311
- Goron K. Smyth, (2002) “Nonlinear Regression”, Encyclopedia of Environmetrics, John Wiley & Sons, Ltd, Chichester, Sayı 3, sf 1405-1411, Newyork
- GUJARATI,Damodar, (1995) “Basic Econometrics”, third edit, MC-Graw-Hill, Inc.,USA.
- Hawkins, D.M. (1973). “On the investigation of alternative regressions by principal component analysis.” Technometrics, Sayı 12 (1), sf 55-67
- Hoerl, A.E., KENNARD, R.W. , (1970). “Ridge Regression:Biased Estimation For Nonorthogonal Problems” Technometrics, Sayı 12 (1), sf 55-67
- Hoerl, A.E., KENNARD, R.W. , (1976). “Ridge Regression: Iterative Estimation of the Biasing Parameter” Communications in Statistics-Theory Method. A5, sf 77-88
- Hoerl, A.E. , KENNARD, R.W. , (2000). “Ridge Regression: Biased Estimation for Nonorthogonal Problems” Technometrics; Sayı 42 sf 80
- Hongehang Hu, (2005). “Ridge Estimation of a Semiparametric Regression Model” Journal of Computational and Applied Mathematics Sayı 176, sf 215
- KUTLAR Azia, (2000). “Ekonometrik Zaman Serileri-Teori ve Uygulama”,Gazi K.Ankara.
- James, W. and Stein, C. (1961). Estimation with quadratic loss. Proc. Fourth Berkeley Symp. Math. Statist. Probab. 1 361--379. Univ. California Press, Berkeley.
- John Fox, (1991) “Regression Diagnostics”, Newbury Park, Calif. : Sage Publications sf 82
- Jürgen Grob, (2003) “Resricted Ridge Estimation”, Statistics Probability Letters Sayı 65, sf 57-64
- MADDALA, G.S. and Inn-Moo Kim, (2002) “Unit Roots,Cointegration and Structural

Changes”, Cambridge University Press,U.K.

Marquardt, D.W. (1963). An algorithm for least-squares estimation of nonlinear parameters. Journal of the Society for Industrial and Applied Mathematics, 11: 431-441.

Marquardt, D.W. , (1970). “Generalized Inverses, Ridge Regression, Biased Estimation and Nonlinear Estimation” Techometrics Sayı 12, sf 591-612

Marquardt, D.W. , Snee, Ronald D. (1975). “Ridge Regression in Practice” American Statistician Sayı 29, sf 3-19

Montgomery, D.C. , E. A. Peck, (1992). ”Introduction to Linear Regression Analysis”. John Wiley & Sons, Inc, Canada

Muammer ŞİMŞEK, (2003) “Kamu Harcamalarının Özel Yatırımlara Etkileri” . C.Ü. İktisadi ve İdari Bilimler Dergisi, Cilt 4, Sayı 2

Neter, J. , W. Wessarman, M. H. Kutner Terry Sincich, (1996) “Applied Linear Statistical Models”, 4th ed.. - Chicago : Irwin, sf 225-289

Obeuchain, R. L. , (1975). ”Good and Optimal Ridge Estimators”. Techometrics Sayı 6, sf 1111-1121

Ostrom, C.W., Jr., (1990). “Time Series Analysis, Regression Techniques”. Second Edition: Quantitative Applications in the Social Sciences, v. 07-009: Newbury Park, Sage Publications.

Pierre Druilhet, Alain Mom , (2006). ”Shrinkage Structure in Biased Regression”. IRMAR et Universite Rennes www.sciencedirect.com

Raiffa, H. and Schlaifer, R. (1961). “Applied Statistical Decision Theory” Harvard University Press.

Ryan, Thomas P. , (1997). ”Modern Regression Analysis”. John Wiley & Sons, Inc, Newyork

Raymond H. Myers, (1990) “Classical and Modern Regression with Applications”, 2th ed.. – Boston : PWS - KENT

Sanford Weisberg, (1985) “Applied Linear Regression”, John Wiley & Sons, Ltd, Newyork, 2th ed..

Scott M. Kowalski, (2001) “Introduction to Linear Regression Analysis”, John Wiley & Sons, Inc.

Seber, G. A. F. And Wild,C. J. , (1989). ”Nonlinear Regression ” Wiley-Newyork

Seber, G. A. F. , (2003). ”Linear Regression Analysis”. 2nd ed.. -- Hoboken, N.J. : Wiley-Interscience

Stein, C. M., (1960). “Multiplie Regression” . Standford, California: Standford University Press.

Sims, C. (1980) "Macroeconomics and Reality" Econometrica, 48, 1-48.

S. H. Ngo , S. Kemeny, A. Deak. , (2003). ”Performance of the Ridge Regression Method as Applied to Complex Linear and Nonlinear Models” Budapest University

Tatlıdil Hüseyin, (1996). ”Uygulamalı Çok Değişkenli İstatistiksel Analiz” Cem Ofset Ltd.

Şti. Ankara.

Vijay Mahajan, Arun K. Jain, M. B. , (1986). "Parameter Estimation in Marketing Models in the Presence of Multicollinearity". JMR, Journal of Marketing Research Sayı 14, sf 586

Vinod, H.D., (1976) "Application of Ridge Regression Methods to a study of Bell System scale economics". Journal of the American Statistical Association. 71, sf 835-841

William Mendenhall, Terry Sincich, (2003) "A Second Course in Statistics: Regression Analysis", 6th ed.. -- Upper Saddle River, N.J. : Prentice Hall, sf 456-459

William H. Greene, (2000) "Econometric Analysis", 4th ed.. -- Upper Saddle River, N.J. : Prentice Hall, 4th ed.. sf 456-459

Zellner Arnold, (1971) "An Introduction to Bayesian Inference in Econometrics", John Wiley & Sons, Newyork

INTERNET KAYNAKLARI

[1] www.fbe.yildiz.edu.tr

[2] www.sciencedirect.com

[3] turk.internet.com/haber/yazigoster.php3?yaziid=9310-43k-

[4] www.spss.com

[5] www.ncss.com

[6] <http://www.tcmb.gov.tr>

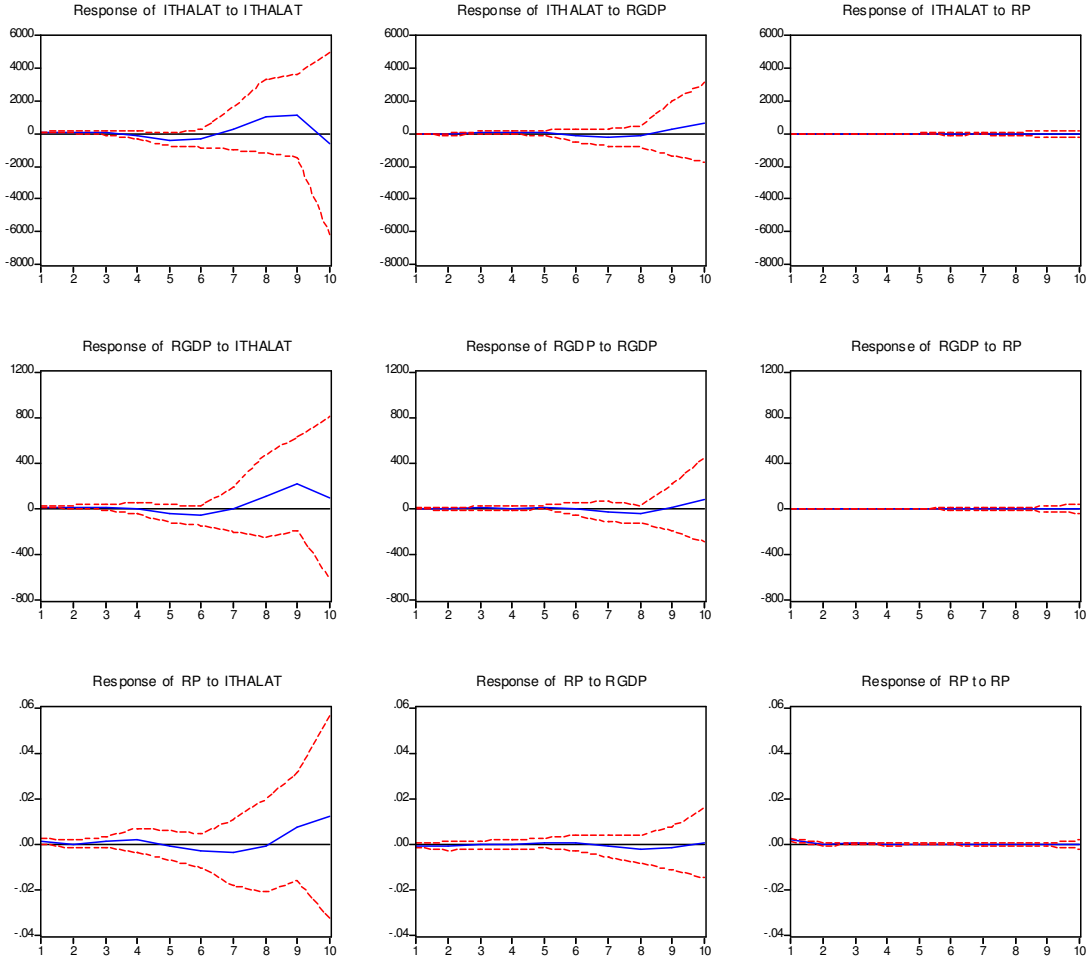
[7] <http://www.imf.org/external/pubs/ft/weo/2005/01/data/dbcsefm.cfm?G=2200>

EKLER

Ek 1 Etki Tepki Grafikleri

Ek 1 Etki – Tepki Grafikleri

Response to Cholesky One S.D. Innovations ± 2 S.E.



Şekil Ek 1.1 Etki Tepki Grafikleri

ÖZGEÇMİŞ

Doğum tarihi	25.12.1981	
Doğum yeri	Gümüşhane	
Lise	1995-1998	Eryaman Lisesi
Lisans	1999-2004	Hacettepe Üniversitesi Fen Fak. İstatistik Bölümü
Yüksek Lisans	2005-...	Yıldız Teknik Üniversitesi Fen Bilimleri Enstitüsü İstatistik Anabilim Dalı
Çalıştığı kurum	2004-Devam ediyor	Hava Kuvvetleri Komutanlığı