

T.C.

MARMARA ÜNİVERSİTESİ

FEN BİLİMLERİ ENSTİTÜSÜ

**LİNEER OLMAYAN REGRESYON MODELLERİNDE
ÇOKLU İÇ İLİŞKİ PROBLEMİ**

ALİ ERKOÇ

YÜKSEK LİSANS TEZİ

MATEMATİK ANABİLİM DALI

UYGULAMALI MATEMATİK PROGRAMI

DANIŞMAN

Prof. Dr. Müjgân TEZ

EŞ DANIŞMAN

Yrd. Doç. Dr. Kadri Ulaş AKAY

İSTANBUL 2011

T.C.

MARMARA ÜNİVERSİTESİ

FEN BİLİMLERİ ENSTİTÜSÜ

**LİNEER OLMAYAN REGRESYON MODELLERİNDE
ÇOKLU İÇ İLİŞKİ PROBLEMİ**

ALİ ERKOÇ

(141103120070213)

YÜKSEK LİSANS TEZİ

MATEMATİK ANABİLİM DALI

UYGULAMALI MATEMATİK PROGRAMI

DANIŞMAN

Prof. Dr. Müjgân TEZ

EŞ DANIŞMAN

Yrd. Doç. Dr. Kadri Ulaş AKAY

İSTANBUL 2011

ÖNSÖZ (TEŞEKKÜR)

Çalışmam boyunca yardımlarını esirgemeyen, beni her konuda destekleyen ve fikirleriyle aydınlatan saygıdeğer hocam Sayın Prof. Dr. Müjgân TEZ'e, bu tezde emeği ve katkısı büyük olan, bilgilerini paylaşmaktan çekinmeyen değerli hocam Sayın Yrd. Doç. Dr. Kadri Ulaş AKAY'a, yardımlarından dolayı tüm değerli hocalarıma ve her zaman yanımda olan sevgili ailem ve arkadaşlarıma teşekkürlerimi sunarım.

Şubat, 2011

Ali ERKOÇ

İÇİNDEKİLER

	SAYFA
ÖNSÖZ.....	i
İÇİNDEKİLER.	ii
ÖZET.....	iv
ABSTRACT	v
SEMBOLLER	vi
KISALTMALAR.....	viii
ŞEKİLLER	ix
TABLOLAR	x
BÖLÜM I.GİRİŞ ve AMAÇ	1
BÖLÜM II. LİNEER MODELLERDE İÇ İLİŞKİ PROBLEMİNİN İNCELENMESİ	3
II.1. LİNEER MODELLERİN GENEL YAPISI ve MODEL VARSAYIMLARI..	3
II.2. MODEL PARAMETRELERİNİN KESTİRİLMESİ.....	4
II.2.1. EN KÜÇÜK KARELER KESTİRİCİSİNİN ÖZELLİKLERİ.....	4
II.3. İÇ İLİŞKİ PROBLEMİ.....	5
II.3.1. İÇ İLİŞKİNİN KAYNAKLARI	5
II.3.2. İÇ İLİŞKİNİN ETKİLERİ	6
II.3.3. İÇ İLİŞKİ TEŞHİSLERİ.....	8
II.3.3.1 KORELASYON MATRİSİNİN İNCELENMESİ.....	8
II.3.3.2 KORELASYON MATRİSİNİN TERSİNİN İNCELENMESİ: VARYANS ŞİŞİRİCİ FAKTÖRLER	9
II.3.3.3 ÖZDEĞER – ÖZVEKTÖR ANALİZİ.....	11
II.3.4. İÇ İLİŞKİNİN GİDERİLME YÖNTEMLERİ	14
II.3.4.1 İLAVE VERİ TOPLAMAK.....	14
II.3.4.2 MODELİ YENİDEN DÜZENLEMEK.....	15
II.3.4.3 RİDGE REGRESYON	15

BÖLÜM III. LİNEER OLMAYAN MODELLERDE ÇOKLU İÇ İLİŞKİ PROBLEMİ	22
III.1. GENEL YAPI.....	23
III.2. EN KÜÇÜK KARELER KESTİRİCİSİNİN BELİRLENMESİ	23
III.2.1. GAUSS-NEWTON YÖNTEMİ	23
III.2.2. EN KÜÇÜK KARELER KESTİRİCİSİ İÇİN ÇIKARSAMALAR	25
III.3. ÇOKLU İÇ İLİŞKİ PROBLEMİ	26
III.3.1 İÇ İLİŞKİ TEŞHİSİ.....	27
III.3.1.1 J MATRİSİNİN TEKİL-DEĞER AYRIŞIMI	27
III.3.1.2 VARYANS AYRIŞIMI	27
III.3.2 TEŞHİS YÖNTEMİNİN UYGULANMASI	31
III.3.3 İÇ İLİŞKİ DURUMUNDA PARAMETRE KESTİRİMİNE ALTERNATİF BİR YAKLAŞIM	32
III.3.3.1 İTERATİF RIDGE REGRESYON	32
BÖLÜM IV. SİMÜLASYON ÇALIŞMASI	35
IV.1. YÖNTEM.....	35
IV.2. ALGORİTMA.....	36
IV.3. SİMÜLASYON ANALİZİ	36
BÖLÜM V. SONUÇ VE TARTIŞMA	41
KAYNAKLAR	43
ÖZGEÇMİŞ	45

ÖZET

LİNEER OLMAYAN REGRESYON MODELLERİNDE ÇOKLU İÇ İLİŞKİ PROBLEMİ

Regresyon analizi, bir veya birden fazla bağımsız değişken ile bağımlı bir değişken arasındaki ilişkinin analizi ve modellenmesi için pek çok teknik içermektedir. Lineer ve lineer olmayan regresyon modeller uygulamalı bilimlerin birçoğunda geniş bir kullanım alanına sahiptir. Regresyon analizinde sıklıkla karşılan problemlerden biri açıklayıcı değişkenler arasındaki iç ilişki problemidir. Bu çalışmada, lineer olmayan modellerde iç ilişki olması durumunda, iç ilişkinin teşhisi ve model parametrelerinin kestiriminde en küçük kareler yöntemine alternatif olan bir yaklaşım incelenecektir.

“Lineer Olmayan Regresyon Modellerinde Çoklu İç İlişki Problemi” adlı bu tez beş bölümden oluşmaktadır. Birinci bölümde, tezin amacına yönelik ön bilgiler verilmiştir. İkinci bölümde, lineer modellerde iç ilişki probleminin, tanısı ve giderilmesi için temel bilgiler verilmiştir. Üçüncü bölümde, lineer olmayan modellerde iç ilişki probleminin tanısı ve giderilmesine yardımcı olabilecek alternatif yaklaşımlar önerilmiştir. Bu yaklaşımlar doğrultusunda, Bölüm IV te bir simülasyon çalışması düzenlenerek önerilen bu yeni yaklaşımın tutarlılığı gösterilmiştir. Bölüm V de elde edilen sonuçlar yorumlanmıştır.

Şubat, 2011

Ali ERKOÇ

ABSTRACT

ON MULTICOLLINEARITY IN NONLINEAR REGRESSION MODELS

Regression analysis includes many techniques for modeling and analyzing the relationship between a dependent variable and one or more independent variables. Linear and nonlinear regression models have widely used in many fields of applied science. One of the frequency problems in regression analysis is multicollinearity problem between the explanatory variables. In this study, diagnosis of multicollinearity and a new approach, which is an alternative to the least squares method for the estimation of the model parameters, in case multicollinearity exists on nonlinear models, will be examined.

This thesis named “on multicollinearity in nonlinear regression models” consists five chapters. In the first chapter, the preliminary information for the purpose of the thesis is given. In the second chapter, for diagnosing and remedying to multicollinearity on linear models, basis information is given. In the third chapter, alternative approaches that may help diagnosis and remedy to multicollinearity on nonlinear models are suggested. According to these approaches, the consistence of this approach which is proposed by a simulation study is shown in chapter IV. In the chapter five, the results obtained here are taken into consideration.

Şubat, 2011

Ali ERKOÇ

SEMBOLLER

y	: Yanıt değişkeni
X	: Açıklayıcı değişkenler matrisi
β	: Lineer modellerde bilinmeyen parametreler vektörü
$\hat{\beta}$	: Lineer modeller için en küçük kareler kestiricisi
ε	: Hata vektörü
X_j	: j . açıklayıcı değişken
$f(x, \theta)$	: Lineer olmayan fonksiyon
y_i	: i . gözlem değeri
θ	: Lineer olmayan modellerde bilinmeyen parametreler vektörü
$\hat{\theta}$	: Lineer olmayan modellerde en küçük kareler kestiricisi
σ^2	: Hata varyansı
$S(\theta)$	: Amaç fonksiyonu
r_{ij}	: i . açıklayıcı değişken ile j . açıklayıcı değişken arasındaki korelasyon katsayısı
R^2	: Belirleyicilik katsayısı
C	: Korelasyon matrisinin tersi
L_1^2	: Kesticinin gerçek parametreye olan uzaklığının karesi
λ_j	: Korelasyon matrisinin j . özdeğeri
k_j	: Koşul indeksi

Λ	: Köşegen elemanları korelasyon matrisinin özdeğerleri olan köşegen matris
T	: Elemanları korelasyon matrisinin özvektörleri olan ortogonal matris
μ_j	: X matrisinin j . tekil değeri
η_j	: X matrisinin tekil değerleri ile ifade edilen j .koşul indeksi
π_{ij}	: Varyans ayrışım oranları
$\hat{\beta}^*$	: β parametresinin herhangi bir yanlı kestiricisi
J	: f fonksiyonunun Jacobian matrisi
Π	: Varyans ayrışım oranları matrisi
$\hat{\theta}_{iR}$	: i . iterasyondaki ridge kestirici
$\hat{\theta}_R$	: Ridge kestirici

KISALTMALAR

HKO	: Hata kareler ortalaması
Cov	: Kovaryans
VIF	: Varyans şişirici faktör
HKT	: Hata kareler toplamı

ŞEKİLLER

SAYFA NO

Şekil II.1 EKK ve ridge kestiricilerinin k değerine bağlı olarak hata kareler ortalaması değişim grafiği.....	18
Şekil II.2 Ridge izi grafiği.....	20

TABLÖLAR

SAYFA NO

Tablo II.1 Webster, Gunst ve Mason Verisi için Özvektörler	12
Tablo III.1 Varyans Ayrışım Oranlarının Π Matrisi	30
Tablo IV.1 $n = 100$ ve $\rho = 0.75$ için Hata Kareler Ortalama Değerleri	39
Tablo IV.2 $n = 50$ ve $\rho = 0.85$ için Hata Kareler Ortalama Değerleri	40

BÖLÜM I

GİRİŞ ve AMAÇ

Regresyon analizi, bir veya birden fazla bağımsız değişken ile bağımlı bir değişken arasındaki ilişkinin analizi ve modellenmesi için pek çok teknik içermektedir. Regresyon analizi birçok alanda (veri analizi, veri madenciliği vb.) çok yaygın olarak kullanılmaktadır. Genel olarak, mühendisler ve bilim adamları veri kümelerini analiz etmek ve yorumlamak için sıklıkla regresyon analizinden yararlanırlar. Yapılan regresyon analizinde kullanılan modeller lineer ya da lineer olmayan yapıya sahip olabilirler. Lineer regresyon analizi birçok durumda değişkenlerin modellenmesinde kullanılabilir olmasına rağmen, her duruma uygun olmayabilir. Fen ve mühendislikte bilinen ve lineer olmayan modellere dayanan birçok problem vardır. Örneğin, kimya mühendisleri, kimyasal bir tepkimede, x , substrat konsantrasyonunun, y , enzim reaksiyonunun başlangıç hızı, üzerindeki etkisini araştırmak için Michaelis-Menten modeli olarak adlandırılan $y = \beta_1 x / (x + \beta_2) + \varepsilon$ lineer olmayan modeli kullanırlar [4,7].

Regresyon modelleri ile teorik olarak analiz yaparken, genellikle, açıklayıcı değişkenlerin bağımsız olduğu kabul edilir. Ancak uygulamada, değişkenlerin tam anlamıyla bağımsız ya da ilişkisiz olması çok düşük olasılıklıdır. Dolayısıyla, modeldeki açıklayıcı değişkenlerin bağımsız olması varsayımı sağlanmaz ve bu da iç ilişki problemine (collinearity) neden olur [4].

Lineer regresyon modellerinde, açıklayıcı değişkenler arasında tam iç ilişki söz konusu ise, $X'X$ matrisinin rankı düşer. Dolayısıyla, tam ranka sahip olmayan $X'X$ matrisinin tersi alınamaz. Ayrıca modelde güçlü bir iç ilişki var ise, bu durumda, X matrisindeki yakın lineer ilişki, $X'X$ matrisinin kötü koşullu olmasına neden olur. $X'X$ matrisinin kötü koşulluluğu, model parametrelerinin kestirimini olumsuz şekilde etkiler.

İç ilişki olması durumunda, parametre kestirimi için kullanılan en küçük kareler yöntemi iyi sonuçlar vermez [4]. Bu problemi gidermek için lineer modellerde birçok analitik çözüm üretilmiştir. Belsley (1991), Belsley, Kuh, Welsch (1980) bu problemin teşhisine yönelik çalışmalarda bulunurken, Hoerl ve Kennard (1970a) bu problemin çözümüne yönelik çalışmalarda bulunmuş ve ridge regresyonu öne sürmüştür. Ridge regresyon, lineer modellerde, bu problemi gidermek için sıklıkla kullanılan yöntemlerden biridir [3]. Bu yöntemle, model parametrelerinin kestiriminde en küçük kareler kestiricisine göre daha verimli sonuçlar elde edilebilir.

Lineer modellerde bu problemin çözümüne yönelik birçok çalışma yapılmış olmasına rağmen, lineer olmayan modellerde bu alanda henüz yeterli bir çalışma yapılmamıştır. Belsley, Kuh, Welsch (1980) iç ilişki probleminin teşhisini lineer olmayan modellere genişletmiştir. Magel ve Hertsgaard (1987) yine lineer olmayan modellerde iç ilişki teşhisine yönelik çalışmalarda bulunmuştur. Hill ve Adkins (2003) bu problemi öncelikle lineer modeller için ele almış ve daha sonra bu problemin lineer olmayan modellerde de olabileceği göz önünde bulundurularak, teşhisi ve parametre kestirimleri üzerindeki etkilerini incelemişlerdir. Yapılan literatür taramasına göre, lineer olmayan modellerde iç ilişki probleminin teşhisine yönelik çalışmalarda, genellikle, Belsley, Kuh, Welsch (1980)' in lineer modellerdeki çalışması temel alınmıştır. Ancak bu çalışmalarda, lineer olmayan modellerde iç ilişki problemini gidermek için herhangi bir analitik yaklaşım öne sürülmemiştir.

Bu tezin amacı, lineer modellerde sözü edilen çoklu iç ilişki probleminin (multicollinearity), lineer olmayan modellerde de olabileceği göz önünde bulundurularak, bu probleme bir çözüm üretmektir. Bu amaç doğrultusunda, Bölüm II de, lineer modeller temel alınarak, problemin kaynağı, teşhisi, etkileri ve giderilme yöntemleri verilmiştir. Bölüm III te, iç ilişki probleminin lineer olmayan modellerdeki genişlemesi incelenmiştir. Ayrıca, iç ilişkinin giderilebilmesi için lineer modellerde yaygın olarak kullanılan ridge regresyon lineer olmayan modellere uyarlanmıştır. Bölüm IV te de, bir simülasyon çalışması ile ele alınan bu problem örneklendirilmiştir. Son olarak, bölüm V de elde edilen sonuçlar yorumlanmıştır.

BÖLÜM II

LİNEER MODELLERDE İÇ İLİŞKİ PROBLEMİNİN İNCELENMESİ

Çalışmalarında deneysel olmayan veriler kullanarak lineer regresyon ile parametre kestirmek isteyen birçok araştırmacı, iç ilişki problemini göz ardı ederler. Dolayısıyla, göz ardı edilen iç ilişki, yüksek standart hataya sahip, anlamsız veya yüksek duyarlı parametrelerin kestirilmesine neden olur [1].

Lineer modellerde iç ilişkinin teşhisi ve giderilmesi açısından birçok çalışma yapılmıştır. Bu bölümde, lineer olmayan modellerde iç ilişki incelemesine zemin hazırlamak amacıyla lineer modellerde iç ilişkinin teşhisi, etkileri ve giderilmesi incelenecektir.

II.1. LİNEER MODELLERİN GENEL YAPISI VE MODEL VARSAYIMLARI

Çoklu lineer regresyon modeller, y $n \times 1$ boyutlu yanıt vektörü, X $n \times p$ boyutlu açıklayıcı değişkenler matrisi (tasarım matrisi), β $p \times 1$ boyutlu bilinmeyen parametreler vektörü ve ε $n \times 1$ boyutlu hata vektörü olmak üzere;

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \varepsilon \quad (\text{II.1})$$

veya matris formunda

$$y = X\beta + \varepsilon \quad (\text{II.2})$$

şeklinde yazılır. Lineer regresyon model kullanılarak yapılan analizlerde, tüm çıkarsamaların tamamıyla doğru ya da güvenilir olması için, aşağıdaki varsayımların sağlanması gerekir.

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

$$E(y_i) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik}$$

(ii) $Var(\varepsilon_i) = \sigma^2$, $i = 1, 2, \dots, n$ veya özdeş olarak $var(y_i) = \sigma^2$

(iii) $Cov(\varepsilon_i, \varepsilon_j) = 0$, $\forall i \neq j$

Bu üç varsayım sağlandığında β parametresinin en küçük kareler kestiricisinin birtakım özellikleri vardır. Bir veya daha fazla varsayımın sağlanmadığı durumda parametre kestirimine dayanan sonuç çıkarımları zayıf olabilir [5].

II.2. MODEL PARAMETRELERİNİN KESTİRİLMESİ

En küçük kareler kestiricisi regresyon parametrelerini kestirmek için kullanılan en yaygın kestiricidir. (II.2) modelindeki parametrelerin kestirimi için en küçük kareler yaklaşımı ile hata kareler toplamı minimize edilir.

$$\sum_{i=1}^n \hat{\varepsilon}_i^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (II.3)$$

ifadesinin minimize edilmesiyle β parametresinin en küçük kareler kestiricisi

$$\hat{\beta} = (X'X)^{-1} X'y \quad (II.4)$$

şeklinde elde edilir.

II.2.1. EN KÜÇÜK KARELER KESTİRİCİSİNİN ÖZELLİKLERİ

(II.4) ile verilen $\hat{\beta}$ en küçük kareler kestiricisi $E(y) = X\beta$ ($\sim E(\varepsilon) = 0$) ve $Cov(y) = \sigma^2 I$ varsayımları kullanılmadan elde edilebilir. Eğer bu varsayımlar sağlanırsa $\hat{\beta}$ kestiricisi aşağıdaki özellikleri sağlar:

- i. Eğer $E(y) = X\beta$ ise $\hat{\beta}$, β parametresinin yansız kestiricisidir. Yani $E(\hat{\beta}) = \beta$ sağlanır.
- ii. Eğer $Cov(y) = \sigma^2 I$ ise $\hat{\beta}$ kestiricisinin kovaryans matrisi $Cov(\hat{\beta}) = \sigma^2 (X'X)^{-1}$ ile verilir.

- iii. Eğer $E(\mathbf{y}) = \mathbf{X}\boldsymbol{\beta}$, $Cov(\mathbf{y}) = \sigma^2 \mathbf{I}$ ise Gauss-Markov teoremine göre, $\hat{\boldsymbol{\beta}}$, $\boldsymbol{\beta}$ parametresinin tüm yansız lineer kestiricileri arasında en küçük varyansa sahip ve en iyi kestiricidir.

En küçük kareler kestirim yöntemi, ortogonal olmayan veriye uygulandığında, regresyon parametrelerinin çok zayıf kestirimleri elde edilir. Regresyon parametrelerinin en küçük kareler kestirimlerinin varyansları gözle görülür bir şekilde büyür. Bu ise en küçük kareler kestirimlerinin, mutlak değerinin çok büyük ve tutarsız olduğunu, yani verilen değişik bir örnekleme işaretlerinin ve büyüklüklerinin değişebileceğini gösterir.

II.3. İÇ İLİŞKİ PROBLEMİ

Regresyon modelleri ile teorik olarak analiz yaparken, genellikle, açıklayıcı değişkenlerin bağımsız olduğu kabul edilir. Fakat uygulamada, değişkenlerin tam anlamıyla ilişkisiz olması çok düşük olasılıklıdır. Dolayısıyla, modeldeki açıklayıcı değişkenlerin bağımsız olması varsayımı sağlanmaz ve bu da iç ilişki problemine neden olur.

II.3.1. İÇ İLİŞKİNİN KAYNAKLARI

(II.2) modelindeki $\mathbf{X}$ matrisinin sütunları X_j olmak üzere; tasarım matrisi $\mathbf{X} = [X_1, X_2, \dots, X_p]$ şeklinde yazılabilir. İç ilişkiyi tasarım matrisindeki sütunların lineer ilişkisi olarak tanımlayabiliriz [4]:

$$\sum_{j=1}^p t_j X_j = 0 \quad (II.5)$$

denklemi hepsi birden sıfır olmayan $t_1, t_2, \dots, t_p$ sabitleri için gerçekleşirse, $X_1, X_2, \dots, X_p$ vektörleri lineer bağımlıdır denir ve iç ilişki problemi mevcuttur. Dolayısıyla $\mathbf{X}'\mathbf{X}$ matrisinin rankı p değerinden küçük olup $(\mathbf{X}'\mathbf{X})^{-1}$ hesaplanamaz. Bundan dolayı, (II.5) denkleminin yaklaşık olarak sağlanması durumunda, $\mathbf{X}$ matrisindeki iç ilişkiyi, yaklaşık lineer ilişki olarak adlandıracaktır.

İç ilişkinin kaynaklarından temel olan dört tanesi aşağıdaki gibidir:

Veri toplama yöntemi iç ilişki problemine yol açabilir. Analiz yapan araştırmacı, örneklemini sadece (II.5) denklemini sağlayan değişkenlerden yararlanarak oluşturursa muhtemel iç ilişkiyi yaratmış olur. Bundan kaynaklanan iç ilişki problemi, veri kümesi mümkün olduğu kadar geniş tutularak giderilebilir.

Model üzerindeki veya kitledeki zorunluluklar iç ilişki problemi yaratabilir. Kitledeki zorunluluklar, açıklayıcı değişkenlerin kimyasal veya üretim süreçlerinde ortaya çıkar.

Model seçimi yine iç ilişki problemini yaratan kaynaklardan biridir. Örneğin modele, bir açıklayıcı değişken varken, bu değişkenin polinomal kuvvetlerinin de alınması iç ilişkiye neden olur.

Modelin aşırı tanımlanması modelde gözlem sayısından çok açıklayıcı değişkenin olmasıdır. Bu da iç temel ilişki probleminin kaynaklarından biridir. Bu tür modeller, bazen tıbbi veya davranış araştırmalarında ortaya çıkar [4].

II.3.2. İÇ İLİŞKİNİN ETKİLERİ

İç ilişkinin varlığı, regresyon parametrelerinin en küçük kareler kestirimini ciddi şekilde etkiler. Bu etkilerden bazılarını göstermek amacıyla, iki parametrelili bir lineer model düşünölsün. Model x_1 ve x_2 açıklayıcı değişkenler ve y yanıt değişkeni olmak üzere

$$y = \beta_1 x_1 + \beta_2 x_2 + \varepsilon \quad (\text{II.6})$$

şeklinde olsun. Açıklayıcı ve yanıt değişkenlerinin birim uzunluk ölçeklemesine göre standartlaştırıldığını kabul edelim. (II.6) modeli için normal denklemler

$$(X'X)\hat{\beta} = X'y$$

$$\begin{bmatrix} 1 & r_{12} \\ r_{12} & 1 \end{bmatrix} \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} = \begin{bmatrix} r_{1y} \\ r_{2y} \end{bmatrix} \quad (\text{II.7})$$

şeklindeir. Burada r_{12} , x_1 ve x_2 değişkenleri arasındaki basit korelasyon, r_{jy} ise x_j ve y arasındaki basit korelasyondur ($j = 1, 2$). $(X'X)$ matrisinin tersi

$$C = (\mathbf{X}'\mathbf{X})^{-1} = \begin{bmatrix} \frac{1}{(1-r_{12}^2)} & \frac{-r_{12}}{(1-r_{12}^2)} \\ \frac{-r_{12}}{(1-r_{12}^2)} & \frac{1}{(1-r_{12}^2)} \end{bmatrix} \quad (\text{II.8})$$

olmak üzere regresyon parametrelerinin kestirimleri

$$\hat{\beta}_1 = \frac{r_{1y} - r_{12}r_{2y}}{(1-r_{12}^2)} \quad , \quad \hat{\beta}_2 = \frac{r_{2y} - r_{12}r_{1y}}{(1-r_{12}^2)} \quad (\text{II.9})$$

şeklindedir. Eğer x_1 ile x_2 arasında güçlü bir iç ilişki varsa r_{12} korelasyon katsayısı büyük olacaktır. (II.8) denkleminde de görüleceği gibi $|r_{12}| \rightarrow 1$ olması durumunda $Var(\hat{\beta}_j) = C_{jj}\sigma^2 \rightarrow \infty$ ve $Cov(\hat{\beta}_1, \hat{\beta}_2) = C_{12}\sigma^2 \rightarrow \mp\infty$ olacaktır. Bundan dolayı, güçlü iç ilişki, varyans ve kovaryansı artırarak, regresyon parametrelerinin en küçük kareler kestirimlerini kötü şekilde etkiler. Bu durum, aynı x değişkeni için alınan farklı örneklerden çok farklı model parametrelerinin kestirilmesine sebep olur.

Modelde iki açıklayıcı değişkenden daha fazla değişken olduğunda, çoklu iç ilişki, yine aynı etkileri gösterir. Bu durumda R_j^2 , x_j ile geriye kalan $p-1$ tane açıklayıcı değişkenin regresyonundan elde edilen çoklu belirleyicilik katsayısı olmak üzere, $C = (\mathbf{X}'\mathbf{X})^{-1}$ matrisinin köşegen elemanı

$$C_{jj} = \frac{1}{1-R_j^2} \quad , \quad j=1,2,\dots,p \quad (\text{II.10})$$

şeklinde yazılır. Eğer x_j ile diğer $p-1$ değişken arasında güçlü bir iç ilişki varsa bu durumda R_j^2 1 değerine yaklaşacaktır. $\hat{\beta}_j$ kestiricisinin varyansı $Var(\hat{\beta}_j) = C_{jj}\sigma^2 = (1-R_j^2)^{-1}\sigma^2$ olduğundan, güçlü iç ilişki parametrelerin en küçük kareler kestirimlerinin varyanslarında artışa sebep olacaktır.

Çoklu iç ilişki, ayrıca, $\hat{\beta}_j$ parametre kestirimlerinin gerçek değerlerinden fazlaca sapmasına neden olur. Bunu göstermek için, $\hat{\beta}$ kestiriminin gerçek parametreden uzaklığının karesi olan

$$L_1^2 = (\hat{\beta} - \beta)'(\hat{\beta} - \beta) \quad (\text{II.11})$$

ifadesini düşünelim. Bu ifadenin beklenen değeri alınırsa

$$\begin{aligned} E(L_1^2) &= E[(\hat{\beta} - \beta)'(\hat{\beta} - \beta)] \\ &= \sum_{j=1}^p E[(\hat{\beta}_j - \beta_j)^2] \\ &= \sum_{j=1}^p \text{Var}(\hat{\beta}_j) \\ &= \sigma^2 \text{Tr}(\mathbf{X}'\mathbf{X})^{-1} \end{aligned} \quad (\text{II.12})$$

elde edilir. Matrisin izi aynı zamanda özdeğerlerin toplamına eşit olduğundan (II.12) denklemi,

$$E(L_1^2) = \sigma^2 \sum_{j=1}^p \frac{1}{\lambda_j} \quad (\text{II.13})$$

halini alır. Burada, $\lambda_j > 0, j = 1, 2, \dots, p$ değerleri $\mathbf{X}'\mathbf{X}$ matrisinin özdeğerleridir. (II.13) denkleminde, eğer $\mathbf{X}'\mathbf{X}$ matrisi, $\mathbf{X}$ matrisindeki iç ilişki yüzünden kötü koşullu ise, en az bir λ_j özdeğeri küçük olacak ve $\hat{\beta}$ en küçük kareler kestiriminin β gerçek parametreye olan uzaklığı oldukça büyük olacaktır.

II.3.3. İÇ İLİŞKİ TEŞHİSLERİ

Regresyon modellerde, iç ilişki teşhisi için uzun çalışmalar yapılmış, birçok teknik geliştirilmiştir. Bu tekniklerden bazıları; korelasyon matrisinin incelenmesi, korelasyon matrisinin tersinin incelenmesi, $\mathbf{X}'\mathbf{X}$ matrisinin determinantının incelenmesi ve özdeğer – özvektör analizidir. Bunlar arasında özdeğer – özvektör analizi en çok kullanılan yöntemlerden bir tanesidir. Bu tezde, iç ilişki teşhisi için, korelasyon matrisi, korelasyon matrisinin tersi ve özdeğer – özvektör analizi incelenecektir.

II.3.3.1. KORELASYON MATRİSİNİN İNCELENMESİ

Korelasyon matrisinin incelenmesi, iç ilişki teşhisi için, sıklıkla kullanılan bir yöntemdir. Değişkenler arasındaki yüksek korelasyon, iç ilişkinin göstergesidir. Eğer veri matrisinin sütunlarının (değişkenlerin), birim uzunluk ölçeklemesine göre

standartlaştırıldığını varsayarsak, $\mathbf{X}'\mathbf{X}$ matrisi korelasyon matrisi haline gelir ve r_{ij} korelasyon matrisinin köşegen dışı elemanları (ikili korelasyon) olmak üzere, bu elemanlar, basit bir iç ilişki ölçüsü olarak kullanılabilir. Eğer x_i ile x_j yakın lineer ilişkili iki açıklayıcı değişken ise $|r_{ij}|$ ifadesi 1 değerine yaklaşacaktır [4].

İki açıklayıcı değişken arasındaki yüksek korelasyon katsayısı, açıkça, bir iç ilişki problemine işaret olmasına rağmen, iç ilişkinin varlığı, yüksek korelasyon katsayısının varlığına kanıt olarak gösterilemez. Bunun nedeni olarak, korelasyon ile iç ilişkinin aynı şey olmadığı söylenebilir. Yüksek korelasyon iç ilişkiye işaretler, fakat bunun tersi doğru değildir [1]. Açıklayıcı değişkenler arasındaki basit korelasyon katsayısını incelemek, sadece iki açıklayıcı değişken arasındaki yakın lineer ilişkiyi belirlemek için yararlıdır. Fakat ikiden fazla değişken yakın lineer ilişki içinde ise, kısmi korelasyon katsayısının büyük olma garantisi yoktur [4]. Ayrıca, iki açıklayıcı değişken kendi aralarında lineer ilişkili değil iken, ikiden fazla değişken bir yakın lineer ilişki gösterebilir.

İç ilişki teşhisi için, korelasyon incelemesinin yukarıda sözü edilen zayıflığının yanı sıra bir diğer zayıflığı da korelasyon değerinin, alt sınırının kesin bir şekilde belli olmamasıdır. Yani ne kadar büyüklükte korelasyon, bir iç ilişki göstergesidir? Bu değer, araştırmacıdan araştırmacıya değişmekle birlikte, bazı alt sınır değerleri verilebilir. Ancak en iyi alt sınırın, araştırmacının kendisinin kazandığı deneysel tecrübelerden elde ettiği değer olduğu söylenebilir.

II.3.3.2. KORELASYON MATRİSİNİN TERSİNİN İNCELENMESİ: VARYANS ŞİŞİRİCİ FAKTÖRLER

$C = (\mathbf{X}'\mathbf{X})^{-1}$ matrisinin köşegen elemanlarının incelenmesi, iç ilişki teşhisi için çok kullanışlı bir yöntemdir. C_{jj} , C matrisinin j . köşegen elemanı olmak üzere,

$$C_{jj} = (1 - R_j^2)^{-1} \quad (\text{II.14})$$

şeklinde yazılabilir. Eğer x_j , kalan diğer açıklayıcı değişkenlere yaklaşık ortogonal ise R_j^2 sıfıra yakın küçük bir değer olacak ve C_{jj} 1 değerine yaklaşacaktır. Fakat

tersine, x_j kalan diğer açıklayıcı değişkenlerin herhangi bir alt kümesi ile yaklaşık lineer ilişkili ise, R_j^2 1 değerine yaklaşacak ve dolayısıyla, C_{jj} büyüyecektir. Regresyon modelinin j . parametre kestiricisinin varyansı, $Var(\hat{\beta}_j) = C_{jj}\sigma^2$ olduğundan, C_{jj} katsayısına, açıklayıcı değişkenler arasındaki yakın lineer ilişkiden kaynaklı olarak, $\hat{\beta}_j$ parametre kestiriminin varyansını artırıcı bir faktör olarak bakılabilir. Daha genel bir ifade ile

$$VIF_j = C_{jj} = (1 - R_j^2)^{-1} \quad (II.15)$$

denklemleri ile verilen ifadeye varyans şişirici faktör denir.

İç ilişki teşhisi için, korelasyon matrisinin incelenmesinde olduğu gibi korelasyon matrisinin tersinin incelenmesinde de (varyans şişirici faktörlerin incelenmesinde) bazı zayıf yönler vardır. İlk olarak, korelasyon incelemesinde olduğu gibi, burada da yüksek VIF değerleri, iç ilişki teşhisi için yeterli ancak gerekli bir koşul değil. Başka bir deyişle, yüksek VIF değerleri iç ilişki olduğu anlamına gelirken, bunun tersi doğru değildir. Yani iç ilişki olduğu durumda, yüksek VIF olacağı anlamına gelmez. İkinci olarak, VIF değerleri, X tasarım matrisindeki yakın lineer ilişkilerin sayılarının teşhisi için kullanılamaz. Yani, yüksek VIF değeri, tasarım matrisindeki iç ilişkiyi doğrular, fakat bu ilişkilerin sayısını bildirmez. Örneğin; X matrisinde iki tane yüksek iç ilişki olduğu varsayalım. Bunlardan ilk olanı X_1, X_2, X_3 ve X_4 sütunlarını içersin, ikinci ilişki ise X_3 ve X_4 sütunlarını içersin. Bu değişkenler için hesaplanan dört VIF değeri büyük olmasına rağmen, bu ilişkilerden sadece ilki tek başına sağlansaydı bile, VIF değeri yine büyük olurdu. Yani, bu iki ayrı ilişkinin varlığı, yüksek VIF değerine bakarak anlaşılmaz. Üçüncü olarak, iç ilişki teşhisi için, kesin bir VIF kesme değerinin olmadığı söylenebilir. Yani, “yüksek VIF değeri”, ifadesi bir kesinlik belirtmez. Ne kadar yüksek bir VIF değeri, bir iç ilişki belirtir? Ne kadar yüksek VIF değeri, var olan iç ilişkinin derecesini hakkında bilgi verir? Temel olarak önerilen VIF kesme değeri, 7 veya 10 olduğu söylenebilir [1].

II.3.3.3. ÖZDEĞER – ÖZVEKTÖR ANALİZİ

$X'X$ matrisinin $\lambda_1, \lambda_2, \dots, \lambda_p$ özdeğerleri (karakteristik kökleri) verideki iç ilişkinin derecesinin ölçülmesinde kullanılabilir. Eğer veride bir veya daha fazla yakın lineer ilişki var ise, bu durumda bir veya daha fazla özdeğer küçük olacaktır. Bir veya daha fazla küçük özdeğer, X matrisinin sütunları arasındaki yakın lineer ilişkinin göstergesidir. Bu yakın lineer ilişkinin teşhisi için **koşul sayısı**,

$$k = \frac{\lambda_{\max}}{\lambda_{\min}} \quad (\text{II.16})$$

incelemesi yaygın olarak kullanılır. Genellikle, koşul sayısı 100 değerinden küçük ise ciddi bir iç ilişki problemi yoktur denir. Eğer bu değer 100 ile 1000 arasında ise, güçlü bir iç ilişki söz konusudur. 1000 değerini aşan koşul sayısı, verideki çok şiddetli bir yakın lineer ilişkiye işaretler. Koşul sayısının yanı sıra, **koşul indeksi** $X'X$ matrisindeki yakın lineer ilişkinin sayısının tespiti için kullanılabilir. $X'X$ matrisinin koşul indeksi,

$$k_j = \frac{\lambda_{\max}}{\lambda_j} \quad j = 1, 2, \dots, p \quad (\text{II.17})$$

şeklindeir. (II.17) ifadesinden, en büyük koşul indeksi koşul sayısına eşittir. Büyük koşul indekslerinin (örneğin $k_j \geq 1000$) sayısı $X'X$ matrisindeki yakın lineer ilişkilerin sayısının belirlenmesinde kullanışlı bir ölçüdür.

Özdeğer – özvektör analizi, ayrıca, verideki yakın lineer ilişkinin doğasının belirlenmesinde de kullanılabilir. Bunun için $X'X$ matrisinin

$$X'X = T\Lambda T' \quad (\text{II.18})$$

şeklindeki ayrışımı düşünölsün. Λ , $p \times p$ boyutlu ve elemanları $X'X$ matrisinin λ_j ($j = 1, 2, \dots, p$) özdeğerleri olan köşegen matris, T $p \times p$ boyutlu ve sütunları $X'X$ matrisinin özvektörleri olan ortogonal matristir. T matrisinin sütunları $t_1, t_2, \dots, t_p$ ile gösterilsin. Eğer λ_j özdeğeri sıfıra yakın ise, bu yakın lineer ilişkiye işaretler, bu özdeğere ilişkin t_j özvektörünün elemanları yakın lineer ilişkinin doğasını belirler. t_j özvektörünün elemanları (II.5) denkleminde tanımlanan

$t_1, t_2, \dots, t_p$ katsayılarıdır. Tablo II.1, Webster, Gunst ve Mason verileri için özvektörleri gösterir [12].

Tablo II.1 Webster, Gunst ve Mason Verisi için Özvektörler

t_1	t_2	t_3	t_4	t_5	t_6
-0.39072	-0.33968	0.67980	0.7990	-0.25104	-0.44768
-0.45560	-0.05392	-0.70013	0.05769	-0.34447	-0.42114
0.48264	-0.45333	-0.16078	0.19103	0.45364	-0.54169
0.18766	0.73547	0.13587	-0.27645	0.01521	-0.57337
-0.49773	-0.09714	-0.03185	-0.56356	0.65128	-0.00605
0.35195	-0.35476	-0.04864	-0.74818	-0.43375	-0.00217

Bu veriler için en küçük özdeğer $\lambda_6 = 0,0011$ ve bu özdeğere karşılık gelen t_6 özvektörünün elemanları (II.5) denklemindeki açıklayıcı değişkenlerin katsayılarıdır. Yani

$$\begin{aligned} & -0.44768x_1 - 0.42114x_2 - 0.54169x_3 - 0.57337x_4 \\ & -0.00605x_5 - 0.00217x_6 = 0 \end{aligned} \quad (\text{II.19})$$

denklemini sağlanır. -0.00605 ve -0.00217 değerlerinin yaklaşık olarak sıfır olduğu varsayımı yapılırsa, (II.19) denklemini

$$x_1 \simeq -0.941x_2 - 1.210x_3 - 1.281x_4 \quad (\text{II.20})$$

halini alır. Bu ise ilk dört açıklayıcı değişkenin bir yakın lineer ilişki içinde olduğunu gösterir. t_6 özvektörünün elemanları ise, bu dört açıklayıcı değişken arasındaki ilişkinin doğasını ifade eder.

Belsley, Kuh ve Welsch (1980), iç ilişki teşhisi için, yukarıda anlatılan yöntemle benzer bir yaklaşım önermiştir. Önerilen yöntem, X tasarım matrisinin tekil değer ayrışımına dayanır. $n \times p$ boyutlu X matrisi,

$$X = UDT' \quad (\text{II.21})$$

şeklinde ayrıştırılabilir. Burada, U , $n \times p$, T , $p \times p$ boyutlu ve $U'U = I$ ve $T'T = I$ koşullarını sağlayan matrisler ve D , $p \times p$ boyutlu, köşegen elemanları negatif olmayan μ_j $j = 1, 2, \dots, p$ değerlerinden oluşan köşegen matristir. μ_j değerlerine, X matrisinin **tekil değerleri** ve $X = UDT'$ ayrışımına ise, X matrisinin **tekil değer ayrışımı** denir.

$$\begin{aligned}
 X'X &= (UDT')'(UDT') \\
 &= TD'U'UDT' \\
 &= TD'DT' \quad (U'U = I) \\
 &= TD^2T' \\
 &= T\Lambda T' \quad (D^2 = \Lambda)
 \end{aligned} \tag{II.22}$$

eşitliği sağlandığından dolayı, tekil değer ayrışımı özdeğer – özvektör kavramıyla yakından ilişkilidir. Yani, X matrisinin tekil değerinin karesi, $X'X$ matrisinin özdeğeridir. T , $X'X$ matrisinin önceden tanımlanmış özvektörlerinin matrisi ve U , sütunları $X'X$ matrisinin p tane sıfır olmayan özdeğerlerine ilişkin özvektörlerden oluşan matristir.

Tekil değerlerin sayısı, X matrisindeki iç ilişkiyi yansıtır. Her bir yakın lineer ilişki için, bir küçük tekil değer olacaktır. Belsley, Kuh ve Welsch (1980), X matrisi için koşul indeksi,

$$\eta_j = \frac{\mu_{\max}}{\mu_j}, \quad j = 1, 2, \dots, p \tag{II.23}$$

şeklinde tanımlamıştır. η_j için en büyük değer X matrisinin koşul sayısıdır. Bu yaklaşım, iç ilişki teşhisinde, $X'X$ matrisi ile değil de, doğrudan asıl ilgilenilmesi gereken matris olan X matrisi ile ilgilidir.

$\hat{\beta}$ en küçük kareler kestiricisinin kovaryans matrisi,

$$Var(\hat{\beta}) = \sigma^2 (X'X)^{-1} = \sigma^2 T\Lambda^{-1}T' \tag{II.24}$$

ve $\hat{\beta}_j$ j .parametre kestiriminin varyansı (II.24) denklemi ile verilen matrisin j . köşegen elemanı olduğundan, varyans

$$Var(\hat{\beta}_j) = \sigma^2 \sum_{i=1}^p \frac{t_{ji}^2}{\mu_i^2} = \sigma^2 \sum_{i=1}^p \frac{t_{ji}^2}{\lambda_i} \quad (II.25)$$

şeklinde yazılabilir. Ayrıca (II.25) ifadesi σ^2 olmadan yazıldığında, $T\Lambda^{-1}T'$ matrisinin j . köşegen elmanı olan

$$VIF_j = \sum_{i=1}^p \frac{t_{ji}^2}{\mu_i^2} = \sum_{i=1}^p \frac{t_{ji}^2}{\lambda_i} \quad (II.26)$$

j . varyans şişirici faktörü verir. Dolayısıyla bir veya daha küçük tekil değer (veya küçük özdeğer) $\hat{\beta}_j$ parametre kestiriminin varyansını aşırı derecede büyütür. Belsley, Kuh ve Welsch (1980), (II.23) denklemi ile tanımladıkları, koşul indeksi kullanarak, iç ilişki teşhisine bir ölçü olarak

$$\pi_{ij} = \frac{t_{ji}^2 / \mu_i^2}{VIF_j}, \quad j = 1, 2, \dots, p \quad (II.27)$$

şeklinde tanımladıkları varyans ayrışım oranlarını önermişlerdir. Eğer π_{ij} değerleri $p \times p$ boyutlu π matrisine dizilirse, π matrisinin her bir sütununun elemanları, her bir $\hat{\beta}_j$ kestiriminin varyans oranlarıdır. İki ya da daha fazla parametre için, bir küçük tekil değere ilişkin yüksek varyans oranı iç ilişkinin göstergesidir. Örneğin π_{32} ve π_{34} değerleri büyük ise, 3. tekil değer $\hat{\beta}_2$ ve $\hat{\beta}_4$ kestirimlerinin varyanslarını şişirici iç ilişkiye neden olur. İç ilişki teşhisinde, referans değeri (eşik değeri) olarak, koşul indisinin 30 değerinden büyük, varyans ayrışım oranının 0,5 değerinden büyük olması alınabilir.

II.3.4. İÇ İLİŞKİNİN GİDERİLME YÖNTEMLERİ

II.3.4.1. İLAVE VERİ TOPLAMAK

İlave veri toplamak mevcut verideki iç ilişkinin derecesini düşürmek için önerilen en iyi yöntemdir. Ek veri toplarken, verideki iç ilişkiyi yok edecek şekilde veriler toplamak mantıklı olacaktır. Fakat ek veri toplamak ekonomik sebeplerden dolayı veya üzerinde çalışılan projenin artık veri toplamaya uygun olamamasından dolayı, her zaman mümkün olmayabilir. Bazen veri toplanmaya müsait olsa bile, toplanacak yeni veri araştırmacının ilgilendiği alandaki açıklayıcı değişkenleri temsil

etmezse, yeni veri kullanıma uygun olmayabilir. Sonuç olarak, modelde var olan iç ilişki eğer, model seçiminden veya yığından kaynaklanmış ise, iç ilişkiyi yok etmek için ilave veri toplamak geçerli bir çözüm olmayacaktır.

II.3.4.2. MODELİ YENİDEN DÜZENLEMEK

İç ilişki sıklıkla modeli oluşturan değişkenlerin seçiminden kaynaklanır. Örneğin yüksek ilişkili iki değişkeni aynı modelde kullanmak iç ilişki yaratır. Bu durumda, modeli yeniden gözden geçirmek ve modelde bazı değişiklikler yapmak iç ilişkinin derecesini düşürebilir. Modeli yeniden düzenlemek için bir yaklaşım açıklayıcı değişkenleri yeniden tanımlamak olabilir. Örneğin x_1 , x_2 ve x_3 yakın lineer ilişkili ise, bu değişkenler arasında $x = (x_1 + x_2) / x_3$ veya $x = x_1 x_2 x_3$ v.s. şeklinde orijinal değişkenlerin içeriklerini (bilgilerini) kaybetmeden kötü koşulluluğu azaltan yeni değişkenler bulunabilir.

Modeli yeniden düzenlemek için bir diğer yaklaşım ise, değişken eleme yöntemidir. Eğer x_1 , x_2 ve x_3 yakın lineer ilişkili ise, bir değişkeni (örneğin x_3) elemek, iç ilişkiyi azaltıcı bir yaklaşım olabilir. Değişken eleme yöntemi, çoğunlukla, etkili bir yöntem olmasına rağmen, modelden atılan değişkenin yanıt değişkeni üzerinde önemli bir etkisi varsa, geçerli bir çözüm sağlamaz. Başka bir deyişle, iç ilişkinin azalması için modele katkısı yüksek bir değişkeni modelden atmak, modelin kestirim gücünün azalmasına sebep olur.

II.3.4.3. RIDGE REGRESYON

İç ilişki problemine karşı önceki bölümlerde anlatılan çözümlerin yanı sıra, birçok analitik yöntem de geliştirilmiştir. Bunlardan en yaygın olanı ridge regresyon yöntemidir.

En küçük kareler yöntemi ortogonal olmayan bir veri grubuna (iç ilişkili veri grubuna) uygulanırsa, çok zayıf parametre kestirimleri elde edilebilir. Zayıf kestirim ifadesi parametre kestiriminin dengeli ya da tutarlı bir kestirim olmadığı anlamına gelir. Yani veri grubundaki küçük değişikliklerde, parametre kestirimi yüksek duyarlılık gösterir. Ayrıca, iç ilişkili bir veride, en küçük kareler kestirimlerinin varyansları aşırı büyük ve parametre kestirim vektörünün boyu ortalamanın üzerinde

olacaktır. Bu ise, en küçük kareler kestiriminin mutlak değerce büyük ve dengesiz bir yapıda olacağına işaretler. Böyle kestirimlerin işaretleri ve büyüklükleri, verilen farklı örneklerde farklılık gösterir.

İç ilişkili bir modelde, parametre kestirimi için kullanılan en küçük kareler yöntemindeki problem, aslında, $\hat{\beta}$ kestiricisinin β parametresinin yansız kestiricisi olması gerekliliğidir. Gauss-Markov teoreminden, en küçük kareler kestiricisinin tüm yansız lineer kestiriciler arasında minimum varyansa sahip olduğu bilinir. Ancak, bu teorem ile verilen “en küçük kareler kestiricisi, tüm yansız lineer kestiriciler arasında minimum varyansa sahiptir.” ifadesi varyansın küçük olacağının garantisini vermez [4].

İç ilişki olması durumunda, β parametresinin yansız kestiricisi olan $\hat{\beta}$ kestiricisinin örneklem dağılımı incelenirse, örneklemin varyansı geniş bir alana yayılmaktadır. Geniş bir alana yayılan $\hat{\beta}$ yansız kestiricisinin varyansı oldukça büyük, dolayısıyla, β parametresi için kurulacak güven aralığı geniş ve nokta kestirimi oldukça kararsız olacaktır [4].

Bu problemi gidermek için bir çözüm, β parametresinin kestiricisinin yansız olma koşulunun ortadan kaldırılmasıdır. Parametre kestiricisine ufak bir yanlılık katarak, daha küçük varyanslı ve dengeli kestirimler elde etmek mümkündür. β parametresinin $\hat{\beta}^*$ ile gösterilen, yanlı fakat $\hat{\beta}$ yansız kestiricisinden daha küçük varyansa sahip bir kestirici olduğu varsayalım. $\hat{\beta}^*$ kestiricisinin hata kareler ortalaması,

$$\begin{aligned} HKO(\hat{\beta}^*) &= E[(\hat{\beta}^* - \beta)^2] \\ &= Var(\hat{\beta}^*) + [E(\hat{\beta}^*) - \beta]^2 \end{aligned} \quad (II.28)$$

veya

$$HKO(\hat{\beta}^*) = Var(\hat{\beta}^*) + [yan(\hat{\beta}^*)]^2 \quad (II.29)$$

şeklinde tanımlanır. Burada HKO $\hat{\beta}^*$ kestiricisinin β parametresine olan uzaklığının karesinin beklenen değeridir. $\hat{\beta}^*$ kestiricisinin varyansı, kestiriciye ufak bir yanlılık ekleyerek, $HKO(\hat{\beta}^*)$ değeri, $Var(\hat{\beta})$ değerinden daha küçük kalacak şekilde küçültülebilir.

Regresyon parametrelerinin yanlı kestiricilerini elde etmek için birçok yöntem geliştirilmiştir. Bu yöntemlerden bir tanesi **ridge regresyondur**. İlk olarak 1962 yılında ridge analizini Hoerl geliştirmiş ve 1970 yılında Hoerl ve Kennard ridge regresyonu ortaya atmıştır. Ridge kestirici, $k \geq 0$ araştırmacı tarafından seçilen uygun bir sabit olmak üzere,

$$(X'X + kI)\hat{\beta}_R = X'y \quad (II.30)$$

denkleminde

$$\hat{\beta}_R = (X'X + kI)^{-1} X'y \quad (II.31)$$

şeklinde elde edilir. Dikkat edilecek olursa, $k = 0$ olduğunda, ridge kestirici en küçük kareler kestiricisine eşit olur.

Ridge kestirici,

$$\begin{aligned} \hat{\beta}_R &= (X'X + kI)^{-1} X'y \\ &= (X'X + kI)^{-1} (X'X)\hat{\beta} \\ &= Z_k \hat{\beta} \end{aligned} \quad (II.32)$$

olduğundan, en küçük kareler kestiricisinin bir lineer kombinasyonudur. Dolayısıyla $E(\hat{\beta}_R) = E(Z_k \hat{\beta}) = Z_k \beta$ eşitliği sağlandığından, $\hat{\beta}_R$ β parametresinin yanlı bir kestiricisidir. k parametresine yanlılık parametresi denir. $\hat{\beta}_R$ kestiricisinin kovaryans matrisi,

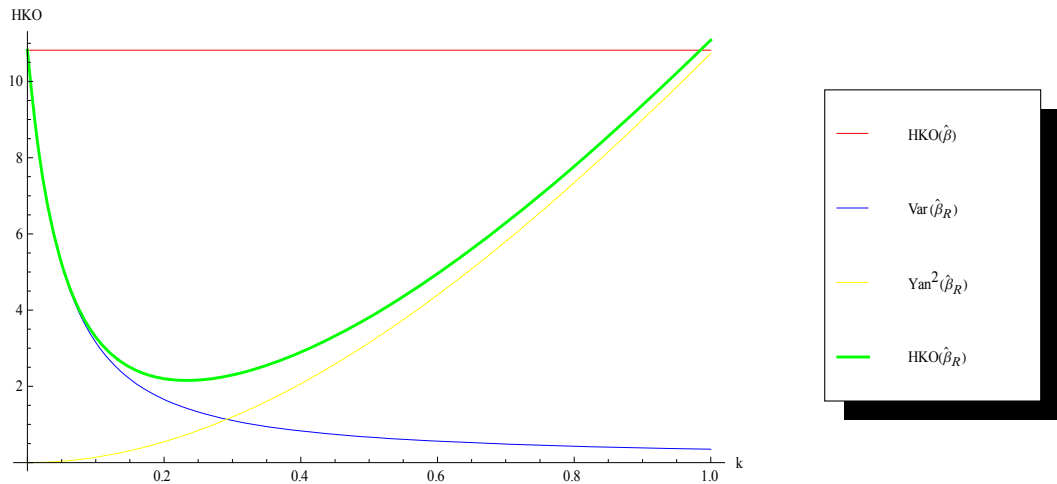
$$Var(\hat{\beta}_R) = \sigma^2 (X'X + kI)^{-1} X'X (X'X + kI)^{-1} \quad (II.33)$$

denklemleriyle ifade edilir. Ridge kestiricinin hata kareler ortalaması, $\lambda_1, \lambda_2, \dots, \lambda_p$ değerleri $\mathbf{X}'\mathbf{X}$ matrisinin özdeğerleri olmak üzere,

$$\begin{aligned} HKO(\hat{\beta}_R) &= Var(\hat{\beta}_R) + [yan(\hat{\beta}_R)]^2 \\ &= \sigma^2 Iz[(\mathbf{X}'\mathbf{X} + k\mathbf{I})^{-1} \mathbf{X}'\mathbf{X} (\mathbf{X}'\mathbf{X} + k\mathbf{I})^{-1}] + k^2 \beta' (\mathbf{X}'\mathbf{X} + k\mathbf{I})^{-2} \beta \quad (II.34) \\ &= \sigma^2 \sum_{j=1}^p \frac{\lambda_j}{(\lambda_j + k)^2} + k^2 \beta' (\mathbf{X}'\mathbf{X} + k\mathbf{I})^{-2} \beta \end{aligned}$$

şeklindedir. (II.33) denklemindeki eşitliğin sağ tarafındaki ilk terim $\hat{\beta}_R$ kestirim vektöründeki parametrelerin varyanslarının toplamıdır. İkinci terim ise yanlılığın karesidir. $k > 0$ değeri için, k değeri arttığında, $\hat{\beta}_R$ kestirimindeki yanlılık artarken, varyansta azalma görülür.

Ridge regresyonda önemli olan yanlılık parametresi k değerinin optimal seçimidir. k parametresi öyle seçilmelidir ki, hata kareler ortalamasının azalması için, toplam varyans değerindeki düşüşün büyüklüğü, yanlığın karesindeki artışın büyüklüğünden daha fazla olsun. Eğer bu sağlanırsa, $\hat{\beta}_R$ ridge kestiricisinin hata kareler ortalaması $\hat{\beta}$ en küçük kareler kestiricisinin varyansından daha küçük olur.



Şekil II.1 EKK ve ridge kestiricilerinin k değerine bağlı olarak hata kareler ortalaması değişim grafiği.

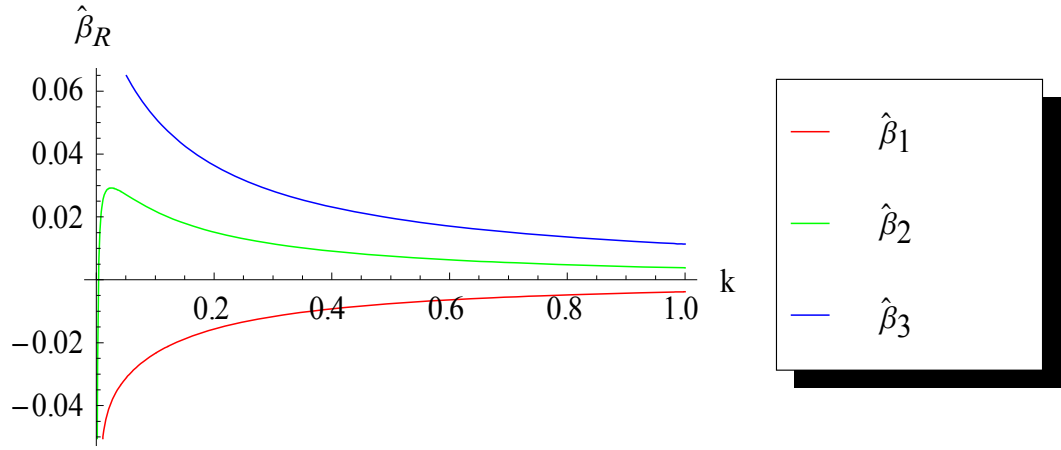
Şekil II.1 de k yanlılık parametresinin seçimi için bir aralık belirlenebilir. Ridge kestiricisinin en küçük kareler kestiricisinden daha iyi sonuç vermesi için, en geniş aralık, 0 ile hata kareler ortalaması eğrilerinin kesiştiği noktalar arası alınmalıdır. En iyi sonuç için ise, bu aralıkta, ridge kestiricisinin hata kareler ortalamasının minimum olduğu nokta alınabilir.

Hoerl ve Kennard (1970a), $\hat{\beta}_R$ ridge kestiricisinin hata kareler ortalamasının $\hat{\beta}$ en küçük kareler kestiricisinin varyansından daha küçük kalacak şekilde negatif olmayan bir k parametresinin olduğunu ispat etmiştir. Hata kareler toplamı,

$$\begin{aligned} HKT &= (\mathbf{y} - \mathbf{X}\hat{\beta}_R)'(\mathbf{y} - \mathbf{X}\hat{\beta}_R) \\ &= (\mathbf{y} - \mathbf{X}\hat{\beta})'(\mathbf{y} - \mathbf{X}\hat{\beta}) + (\hat{\beta}_R - \hat{\beta})' \mathbf{X}'\mathbf{X} (\hat{\beta}_R - \hat{\beta}) \end{aligned} \quad (\text{II.35})$$

şeklindedir. (II.35) denkleminde eşitliğin sağ tarafındaki ilk terim $\hat{\beta}$ en küçük kareler kestirimi için hata kareler toplamı olduğundan, k değeri arttığında hata kareler toplamı artar. Bundan dolayı, toplam hata kareler toplamı sabit olduğundan, k arttığında R^2 değeri düşer. Dolayısıyla k parametresinin seçimi, hem yanlılığı hem de belirleyicilik katsayısını etkilediği için çok önemlidir.

Hoerl ve Kennard, k parametresinin uygun seçimi için **ridge izi** yöntemini önermiştir. Ridge izi, k parametresinin 0 ile 1 arasındaki değerleri için, k değerlerine karşılık $\hat{\beta}_R$ kestirim elemanlarının bir grafiğidir. Modelde iç ilişki yüksek ise, oluşturulan ridge izi grafiğinde, regresyon parametrelerinin kestirimindeki kararsızlık açıkça görülecektir. k arttıkça bazı parametreler yüksek oranda değişecek ve belli değerlerde dengeye ulaşacaktır. Burada amaç, $\hat{\beta}_R$ kestirimlerinin dengeye ulaştığı noktada, yeteri kadar büyük (optimal) k değerinin seçilmesidir.



Şekil II.2 Ridge izi grafiği.

Şekil II.2 de, Bölüm IV de yapılan simülasyon çalışmasından bir örnek ridge izi verilmiştir. Bu grafikten, görsel olarak, analitik hesapla elde edilen k değerine yakın bir değer seçilebilir. Ancak, grafikte önemli olan, parametrelerin dengeye ulaştığı k değerini alırken, bu değeri yeteri kadar büyük seçmektir. Çünkü gereğinden büyük seçilen k değeri yanlılığı artırır. Dolayısıyla hata kareler ortalaması büyür.

k parametresinin seçimi için Hoerl ve Kennard (1970b) tarafından önerilen grafiksel bir yöntem olan ridge izinin yanı birçok analitik yöntem geliştirilmiştir. Hoerl, Kennard ve Baldwin (1975) k parametresinin seçimini

$$k = \frac{p\hat{\sigma}^2}{\hat{\beta}'\hat{\beta}} \quad (\text{II.36})$$

denklemleri vermiştir. Burada $\hat{\sigma}^2$ ve $\hat{\beta}$ kestirimleri en küçük karelerden elde edilir. Hoerl ve Kennard (1976), (II.36) denklemleri seçilen k parametresinden esinlenerek iteratif kestirim yöntemini önermiştir. Bu yöntem ile bir dizi k değeri ve buna bağlı olarak bir dizi ridge kestirimi elde edilir:

$$\begin{aligned}
\hat{\beta} & : k_0 = \frac{p\hat{\sigma}^2}{\hat{\beta}'\hat{\beta}} \\
\hat{\beta}_R(k_0) & : k_1 = \frac{p\hat{\sigma}^2}{\hat{\beta}_R'(k_0)\hat{\beta}_R(k_0)} \\
\hat{\beta}_R(k_1) & : k_2 = \frac{p\hat{\sigma}^2}{\hat{\beta}_R'(k_1)\hat{\beta}_R(k_1)} \\
& \vdots
\end{aligned} \tag{II.37}$$

$T = \hat{I}z(\mathbf{X}'\mathbf{X})^{-1} / p$ olmak üzere, eğer $\frac{k_{j+1} - k_j}{k_j} > 20T^{-1,3}$ ise algoritma devam eder, aksi halde algoritma durur ve son parametre kestirim değeri olarak alınır.

k parametresinin seçimi için analitik olarak hesaplanan bir diğer denklem

$$k = \frac{1}{\hat{\beta}'\hat{\beta}} \tag{II.38}$$

ifadesi ile verilir [11]. Bunların dışında, k yanlılık parametresinin seçimi için birçok analitik yöntem vardır. Ancak bu tezde k parametresinin seçimi için, IV. Bölümde yapılan simülasyon çalışmasında, (II.36) ve (II.38) ifadeleri ile verilen denklemler kullanılmıştır.

BÖLÜM III

LİNEER OLMAYAN MODELLERDE ÇOKLU İÇ İLİŞKİ PROBLEMİ

Lineer regresyon modeller, birçok durumda olayların modellenmesinde kullanılabilir olmasına rağmen, lineer olmayan yapıya sahip bir matematiksel ilişki ile açıklanabilecek birçok olay vardır. Bu tür lineer olmayan ilişkiler ancak lineer olmayan regresyon modelleri ile açıklanabilirler. Lineer olmayan regresyon modeller, kimya, biyoloji, ziraat, tıp, mühendislik vs. gibi uygulamalı bilimlerde geniş kullanım alanına sahiptir. Örneğin kimya mühendisleri, açıklayıcı değişkenlerin yanıt değişkeni üzerindeki etkisini araştırmak için Michaelis-Menten modeli olarak adlandırılan, lineer olmayan $y = \beta_1 x / (x + \beta_2) + \varepsilon$ modelini kullanırlar [4]. Biyologlar bir çiçeğin veya bir organizmanın büyüme oranını lineer olmayan modeller yardımıyla belirleyebilirler. Ziraat mühendisleri, toprağın gübre emilim oranı için yine lineer olmayan modelleri kullanabilirler. Bu nedenle regresyon analizinde modele iyi bir yaklaşım yapmak, optimal parametre kestirimine dayanır.

Lineer olmayan modellerde parametre kestirimi, iteratif bir süreç olduğu için, lineer modellere göre daha karmaşık bir yapıya sahiptir. Sözü edilen modellerde parametre kestirim sürecinde en büyük problem, fonksiyona iteratif olarak en iyi yaklaşım gerektiğinden, yakınsama problemidir. Bu problem modeldeki iç ilişki, başlangıç parametresinin seçimi vs. gibi nedenlerden dolayı meydana gelebilir. Parametre kestiriminde sıklıkla kullanılan en iyi yaklaşım algoritmaları, Newton-Raphson ve Gauss-Newton algoritmalarıdır [6].

III.1. GENEL YAPI

Genel olarak lineer olmayan regresyon modeller,

$$y = f(x, \theta) + \varepsilon \quad (III.1)$$

şeklinde yazılabilir. Burada f beklenen fonksiyon, θ , $p \times 1$ boyutlu bilinmeyen parametreler vektörü ve ε , $E(\varepsilon) = 0$, $Var(\varepsilon) = \sigma^2$ koşullarını sağlayan hata terimidir. Hata teriminin normal dağılıma sahip olduğu kabul edilir. Lineer olmayan modellerde, lineer modellerden farklı olarak, beklenen yanıtın parametrelere göre en az bir türevi parametrelerden en az birine bağlıdır.

III.2. EN KÜÇÜK KARELER KESTİRİCİSİNİN BELİRLENMESİ

Lineer olmayan modellerde en küçük kareler kestiricisi, lineer modellerde olduğu gibi,

$$S(\theta) = \|y - f(x, \theta)\|^2 \quad (III.2)$$

amaç fonksiyonunu minimize eden değerdir. n gözlemlili bir örnekleme sahip olduğumuzu varsayalım. (III.1) modelinin en küçük kareler kestiricisini bulmak, (III.2) denklemini θ parametresine göre türetilip,

$$\sum_{i=1}^n [y_i - f(x_i, \theta)] \left[\frac{\partial f(x_i, \theta)}{\partial \theta_j} \right]_{\theta=\hat{\theta}} = 0 \quad j = 1, 2, \dots, p \quad (III.3)$$

p tane normal denklemi çözmeyi gerektirir. f lineer olmayan bir model olduğundan, (III.3) denkleminde büyük parantez içindeki ifade bilinmeyen parametrelerin fonksiyonu olacaktır. Bundan dolayı bu denklem sistemini çözmek oldukça güç bir işlem olacaktır.

III.2.1. GAUSS-NEWTON YÖNTEMİ

Gauss tarafından önerilen yöntem, beklenen fonksiyona iteratif olarak lineer yaklaşımlarda bulunarak θ parametresini kestirmeye dayanır. $f(x, \theta)$ beklenen fonksiyonunu bir θ_0 başlangıç noktasında birinci dereceden Taylor serisine açalım:

$$f(\mathbf{x}_i, \boldsymbol{\theta}) \approx f(\mathbf{x}_i, \boldsymbol{\theta}_0) + \sum_{j=1}^p \left[\frac{\partial f(\mathbf{x}_i, \boldsymbol{\theta})}{\partial \theta_j} \right]_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} (\theta_j - \theta_{j0}) \quad (\text{III.4})$$

$$\begin{aligned} f_i^0 &= f(\mathbf{x}_i, \boldsymbol{\theta}_0) \\ \beta_j^0 &= \theta_j - \theta_{j0} \\ J_{ij}^0 &= \left[\frac{\partial f(\mathbf{x}_i, \boldsymbol{\theta})}{\partial \theta_j} \right]_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} \end{aligned} \quad (\text{III.5})$$

kısaltmaları ile (III.1) lineer olmayan modeline

$$y_i - f_i^0 = \sum_{j=1}^p \beta_j^0 J_{ij}^0 + \varepsilon_i, \quad i = 1, 2, \dots, n \quad (\text{III.6})$$

olacak şekilde lineer bir yaklaşımda bulunabiliriz. Burada J, f fonksiyonun Jacobian matrisidir. (III.6) modeli matris formunda,

$$\mathbf{y}_0 = \mathbf{J}_0 \boldsymbol{\beta}_0 + \boldsymbol{\varepsilon} \quad (\text{III.7})$$

gibi ifade edilebilir. Bu durumda (III.7) modelindeki $\boldsymbol{\beta}_0$ parametresinin en küçük kareler kestiricisi,

$$\hat{\boldsymbol{\beta}}_0 = (\mathbf{J}_0' \mathbf{J}_0)^{-1} \mathbf{J}_0' \mathbf{y}_0 \quad (\text{III.8})$$

$$S(\boldsymbol{\theta}) = \sum_{i=1}^n \left\{ y_i - f(\mathbf{x}_i, \boldsymbol{\theta}_0) - \sum_{j=1}^p \beta_j^0 J_{ij}^0 \right\}^2 \quad (\text{III.9})$$

amaç fonksiyonunu minimize eden değerdir. $\boldsymbol{\beta}_0 = \boldsymbol{\theta} - \boldsymbol{\theta}_0$ olduğundan, (III.1) modelinin 1. iterasyon sonucunda elde edilen parametre kestirimi,

$$\hat{\boldsymbol{\theta}}_1 = \hat{\boldsymbol{\beta}}_0 + \boldsymbol{\theta}_0 \quad (\text{III.10})$$

elde edilmiş olur. Bir sonraki adımda (III.10) denklemindeki $\hat{\boldsymbol{\theta}}_1$, bir önceki adımda başlangıç parametresi olarak kullanılan, $\boldsymbol{\theta}_0$ parametresinin rolünü oynar ve $\hat{\boldsymbol{\theta}}_2$ parametresi kestirilir. Bu işlemle beraber k . iterasyonda,

$$\begin{aligned}\hat{\theta}_{k+1} &= \hat{\theta}_k + \hat{\beta}_k \\ &= \hat{\theta}_k + (J'_k J_k)^{-1} J'_k y_k\end{aligned}\quad (\text{III.11})$$

parametre kestirim değeri elde edilir. Bu iteratif süreç yakınsaklık sağlanıncaya kadar devam eder. Bir yakınsaklık ölçütü olarak

$$\left| (\hat{\theta}_{j,k+1} - \hat{\theta}_{j,k}) / \hat{\theta}_{j,k} \right| < \delta, \quad j=1,2,\dots,p \quad (\text{III.12})$$

eşitsizliği verilebilir. Burada amaç, parametre kestirim değerlerinin birbirlerine yeteri kadar çok yaklaşımları halinde iterasyonu durdurmak olduğu için $\delta = 10^{-6}$ seçilebilir [4]. Yakınsaklığın sağlanması için, her iterasyonda elde edilen parametre kestirim değeri ile hesaplanan hata kareler toplamında, iteratif olarak düşüşün gerçekleşmesi gerekmektedir. Ayrıca bu iteratif sürece bir başlangıç noktası seçilerek başlandığı için, başlangıç parametresinin seçimi yakınsaklık açısından çok önemlidir. Çünkü iyi seçilmeyen bir başlangıç parametre değeri, yakınsamamaya veya geç yakınsamaya yol açabilir.

III.2.2. EN KÜÇÜK KARELER KESTİRİCİSİ İÇİN ÇIKARSAMALAR

θ parametresi yakınsama ölçütünü sağladığında, hata varyansı σ^2 değerinin kestirimini, hata kareler ortalamasından elde edebiliriz.

$$HKO = \hat{\sigma}^2 = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n-p} = \frac{\sum_{i=1}^n [y_i - f(x_i, \hat{\theta})]^2}{n-p} = \frac{S(\hat{\theta})}{n-p} \quad (\text{III.13})$$

Lineer olmayan regresyon modellerinde oluşturulan J Jacobian matrisi, lineer modellerdeki X tasarım matrisinin rolünü oynar [9]. Bu nedenle lineer en küçük karelerin bilinen özellikleri, lineer olmayan en küçük karelerde **asimptotik** (yaklaşık) olarak sağlanır. Gallant (1987), eğer örneklem büyüklüğü yeteri kadar büyük ve $\varepsilon \sim N(0, \sigma^2)$ varsayımı sağlanıyorsa, $\hat{\theta}$ kestiricisinin,

$$\hat{\theta} \sim N[\theta, (J'J)^{-1} \sigma^2] \quad (\text{III.14})$$

asimptotik olarak özelliğine sahip olduğunu göstermiştir [7].

III.3. ÇOKLU İÇ İLİŞKİ PROBLEMİ

Lineer olmayan modellerde, iç ilişkinin derecesinin (şiddetinin) ve sonuçlarının (etkilerinin) teşhisi, lineer modellere göre daha karmaşıktır. Fakat (III.7) ve (III.14) denklemlerinden J Jacobian matrisi, lineer modellerdeki X matrisinin rolünü oynamaktadır. Bu sebeple, lineer modellerdeki iç ilişki teşhis yöntemlerini lineer olmayan modellere genişletmek mümkündür.

J Jacobian matrisi iç ilişki için incelendiğinde, bu matrisin yapısından kaynaklanan bir problem ortaya çıkar. J matrisi sadece X tasarım matrisine değil, ayrıca θ parametresine de bağlıdır. Bu da parametre uzayında, iç ilişkinin noktadan noktaya değişiklik göstermesine sebep olur. Ayrıca X tasarım matrisinin sütunları arasındaki ilişki J matrisinin sütunları arasındaki ilişkiye karşılık gelebilir veya gelmeyebilir [9].

İç ilişki problemi lineer olmayan regresyonu iki şekilde etkiler. İlk etki, Gauss-Newton algoritmasının ıraksaması ya da çok geç yakınsamasıdır. Bunun nedeni, her adımda hesaplanan $J'J$ matrisinin, J matrisindeki iç ilişki nedeniyle kötü koşullu olması ve dolayısıyla tersinin alınmasında zorlukların meydana gelmesidir. İç ilişkinin yüksek olduğu durumda (gerçek ilişki) Jacobian matrisi tam ranklı olmayacak ve $(J'J)^{-1}$ hesaplanamayacaktır. Bunun sonucu olarak yakınsaklık sağlanamayacaktır. İkinci etki ise, lineer olmayan en küçük kareler kestiricisinin kestirilen asimptotik varyans-kovaryans matrisi,

$$Cov(\hat{\theta}) = \sigma^2 (J'J)^{-1} \quad (III.15)$$

$J'J$ matrisini içerdiğinden, gerçek değerinden sapmasıdır. Buna bağlı olarak kestirilen parametre değerleri için kurulan güven aralıkları, lineer modellerde olduğu gibi, genişleme gösterebilir.

Jacobian matrisindeki iç ilişki, verinin dönüştürülmesinden etkilenebilir. Bu anlamda, bir yaygın problem de $J(\theta_k)$ matrisinin sütunlarının oldukça farklı değerlere sahip olmasıdır. $J(\theta_k)$, $f(x, \theta)$ fonksiyonunun türevlerini içerdiğinden dolayı, bu matrisin elemanları fonksiyonun eğimleridir. Eğer bu eğimler birbirlerinden çok farklı ise, bu durumda fonksiyon bazı yönlerde dik, bazı yönlerde

ise daha yüzeyssel (derin olmayan, sıg) bir yapıdadır. Bu yapıda olan bir yüzey ile çalışmak oldukça güçtür. Bazen orijinal veri ölçeklenerek, Jacobian matrisinin sütunları arasındaki dengeyi sağlamak mümkündür. Bu yüzeyden kaynaklanan bir sorun olduğu için yapılan ölçekleme yüzeyin biraz daha düzleştirilmesi anlamına gelmektedir [9].

III.3.1 İÇ İLİŞKİ TEŞHİSİ

Linear olmayan regresyon modellerde iç ilişki teşhisi için, önerilecek olan yöntem, Belsley (1991)' in linear modellerde önerdiği iç ilişki teşhis yöntemidir. Belsley' in iç ilişki teşhisinde bahsettiği linear modeldeki X matrisinin yerini burada J Jacobian matrisi alacaktır. Ayrıca iç ilişki teşhis yöntemini vermeden önce teşhis için bilinmesi gereken bazı teknik gereksinimler verilecektir.

III.3.1.1 J MATRİSİNİN TEKİL-DEĞER AYRIŞIMI

Herhangi bir $n \times p$ boyutlu J matrisi, $U^T U = V^T V = I$ ve D , J matrisinin tekil değerleri olan $\mu_1, \mu_2, \dots, \mu_p$ ifadelerinden oluşan negatif olmayan köşegen matris olmak üzere,

$$J = UDV^T \quad (III.16)$$

şeklinde ayrıştırılabilir. Bu çok etkili bir ayrışımıdır. Çünkü ayrışımın üç bileşeni de çok özel matrislerdir: U , sütunları ortogonal bir matris, V , hem satırları hem de sütunları ortogonal bir matris ve D , negatif olmayan köşegen bir matristir. Burada U , $n \times p$, V , $p \times p$ ve D , $p \times p$ boyutlu olarak kabul edilir. Fakat ayrıştırmaya uygun farklı boyut seçimi yapılabilir ve ayrıca farklı boyut seçimleri yapılan çalışmaya göre daha kullanışlı olabilir. (III.16) ayrışımı incelenecek olursa, J matrisinin sütunları, U matrisinin sütunlarının lineer birleşimi ve satırları, V matrisinin sütunlarının birleşimidir. Bu anlamda, U matrisinin sütunları J matrisinin sütun uzayı için bir ortogonal taban, V matrisinin sütunları (ve satırları) J matrisinin satır uzayı için bir ortogonal taban oluşturur.

III.3.1.2 VARYANS AYRIŞIMI

Varyans ayrışımında ilk adım olarak J matrisi ölçeklenir. Ölçekleme işlemi birim uzunluk ölçeklemesine göre, J matrisinin j . sütunu x_j olmak üzere, j . sütunun

her elemanı $\sqrt{x'_j x_j}$ ile bölünür. Bu işlem p sütun için gerçekleştirilir. Böylece matrisinin tüm sütunları normalize edilmiş ve sütunlar arasındaki denge sağlanmış olur. Bir sonraki adımda $\mathbf{J}$ matrisinin p tane koşul indeksi olan,

$$\eta_j = \frac{\sqrt{\lambda_{\max}}}{\sqrt{\lambda_j}} \quad j = 1, 2, \dots, p \quad (\text{III.17})$$

değerleri hesaplanır. Burada λ_j değeri $\mathbf{J}'\mathbf{J}$ matrisinin j . özdeğeri, $\lambda_{\max}$ ise $\mathbf{J}'\mathbf{J}$ matrisinin en büyük özdeğeridir.

(III.17) denklemi ile verilen koşul indeksi, lineer modellerde olduğu gibi lineer olmayan modeller için iç ilişki teşhisinde bir ölçüt olarak kullanılabilir.

Genel regresyon varsayımları altında (III.7) denklemi ile verilen lineer modelin en küçük kareler kestiricisinin varyans-kovaryans matrisi $\sigma^2(\mathbf{J}'\mathbf{J})^{-1}$ ile verilmişti. $\mathbf{J}$ matrisinin tekil değer ayrışımını kullanarak, $\mathbf{J} = \mathbf{U}\mathbf{D}\mathbf{V}^T$ olmak üzere, varyans-kovaryans matrisi,

$$\begin{aligned} \text{Cov}(\hat{\boldsymbol{\theta}}) &= \sigma^2(\mathbf{J}'\mathbf{J})^{-1} \\ &= \sigma^2(\mathbf{V}\mathbf{D}^T\mathbf{U}^T\mathbf{U}\mathbf{D}\mathbf{V}^T)^{-1} \\ &= \sigma^2(\mathbf{V}\mathbf{D}^2\mathbf{V}^T)^{-1} \\ &= \sigma^2\mathbf{V}\mathbf{D}^{-2}\mathbf{V}^T \end{aligned} \quad (\text{III.18})$$

şeklinde ayrıştırılabilir. Buradan k . parametrenin varyansı,

$$\text{Var}(\hat{\theta}_k) = \sigma^2 \sum_j \frac{v_{kj}^2}{\mu_j^2} \quad (\text{III.19})$$

μ_j değerleri $\mathbf{J}$ matrisinin tekil değerleri ve $\mathbf{V} \equiv (v_{ij})$ olmak üzere elde edilir. (III.19) denkleminde, $\mu_j^2 = \lambda_j$ eşitliğinin sağlandığı görülür. Bu ise $\mathbf{J}$ matrisinin tekil noktalarının, $\mathbf{J}'\mathbf{J}$ matrisinin özdeğerlerine karşılık geldiğini gösterir.

(III.19) denklemi, $Var(\hat{\theta}_k)$ değerini $n \times p$ boyutlu J matrisinin her biri p tane tekil değerden bir ve yalnız biri ile ilişkili bileşenlerin toplamına ayırıştırır. μ_j^2 paydada olduğundan, yakın bağıllık gösteren yani küçük μ_j tekil değerine sahip bileşenler diğer bileşenler ile büyük ilişkili olacaktır. Dolayısıyla, aynı küçük tekil değere ilişkin bileşenlere karşılık gelen iki ya da daha fazla parametrenin aşırı derecede yüksek varyans oranı, J matrisinin bu parametrelere ilişkin sütunlarının yakın ilişki içerdiğinin kanıtıdır [1]. Her bir tekil değere, dolayısıyla koşul sayısına karşılık gelen **varyans-ayrışım oranı**,

$$\phi_{kj} \equiv \frac{v_{kj}^2}{\mu_j^2} \quad , \quad \phi_k \equiv \sum_{j=1}^p \phi_{kj} \quad k = 1, 2, \dots, p \quad (III.20)$$

tanımlamaları ile

$$\pi_{jk} = \frac{\phi_{kj}}{\phi_k} \quad k, j = 1, 2, \dots, p \quad (III.21)$$

şeklindedir. Varyans-ayrışım oranları Tablo 3.1 ile Π matrisinde verilmiştir.

Tablo III.1 Varyans Ayrışım Oranlarının Π Matrisi

Koşul Sayısı	$Var(\hat{\theta}_1)$	$Var(\hat{\theta}_2)$	...	$Var(\hat{\theta}_p)$
	Oranları			
η_1	π_{11}	π_{12}	...	π_{1p}
η_2	π_{21}	π_{22}	...	π_{2p}
.	.	.		.
.	.	.		.
.	.	.		.
η_p	π_{p1}	π_{p2}	...	π_{pp}

Π matrisinde her satır, verilen μ_j tekil değerine ya da özdeş olarak $\eta_j = \mu_{\max} / \mu_j$ koşul indeksine karşılık gelir. Bu satırlar, koşul indeksinin artan (azalan) sırasına göre sıralanabilir. Eğer koşul indeksi artan sırasına göre sıralanırsa, en küçük koşul indeksi en üstte olmak üzere daima 1 değerine eşit olacak, en büyük koşul indeksi ise daima koşul sayısına eşit olacaktır. Ek olarak, her koşul sayısı tamsayı değerine yuvarlatılarak, virgülden sonraki değeri göz ardı edilebilir. Ayrıca Π matrisinde (III.22) denkleminde olduğu gibi, varyans-ayrışım oranlarına ilişkin her sütunun elemanları toplamı 1 değerine eşittir.

$$\begin{aligned}
\pi_{11} + \pi_{21} + \dots + \pi_{p1} &= \sum_{j=1}^p \pi_{j1} \\
&= \sum_{j=1}^p \frac{\phi_{1j}}{\phi_1} = \frac{1}{\phi_1} \sum_{j=1}^p \phi_{1j} \\
&= \frac{1}{\sum_{j=1}^p \phi_{1j}} \sum_{j=1}^p \phi_{1j} = 1
\end{aligned} \tag{III.22}$$

Π matrisindeki yüksek koşul indeksi için, bu koşul indeksine ilişkin satırda iki veya daha fazla yüksek varyans ayrışım oranı (örneğin, 0.5 den büyük) varsa, bu durumda, bu varyans ayrışım oranlarına ilişkin bağımsız değişkenler yakın lineer ilişki içindedir [10].

III.3.2 TEŞHİS YÖNTEMİNİN UYGULANMASI

Verilen herhangi bir veri matrisinde iç ilişki teşhisinde bulunurken, iki eşik değeri hesaplanmalıdır. Bunlardan biri, koşul indeksi için kesme değeri, η^* , diğeri ise varyans ayrışım oranı için kesme değeri, π^* , ifadeleridir [8]. Yapılan çalışmalar lineer olmayan modellerde olduğundan, uygulanacak adımlar lineerleştirilmiş model olan

$$y = J\beta + \varepsilon \quad (III.23)$$

denklemini üzerinden yapılacaktır.

Adım 1. J matrisinin sütunları birim uzunluk ölçeklemesine göre standartlaştırılır.

Adım 2. J matrisinin tekil değer ayrışımı elde edilir ve bundan yararlanarak (III.17) denkleminde olduğu gibi η_j koşul indeksi ve Tablo III.1 de olduğu gibi varyans-ayrışım oranlarının Π matrisi hesaplanır.

Adım 3. Seçilen η^* eşik değerini aşan koşul sayısına göre, yakın ilişkinin (bağılılığın) sayısı ve derecesi belirlenir. Bu eşik değeri 10, 15 veya 30 alınabilir [8].

Adım 4. 3. Adımda seçilen eşik değerine göre, ılımlı ilişkiye işaret eden (eşik değerine yakın koşul indeksleri) koşul indeksleri ve kuvvetli ilişkiye işaret eden (eşik değerini çok aşan koşul indeksleri) koşul indeksleri araştırılır.

Adım 5. Bu adımda, bir önceki adımda belirlenen potansiyel ilişkilerin dereceleri araştırılır. Bu ilişkilerden kaynaklanan regresyon kestirimlerindeki düşüş belirlenir. İlişki derecelerini araştırmak için varyans ayrışım oranı eşik değeri olarak $\pi^* = 0,5$ alınabilir.

III.3.3 İÇ İLİŞKİ DURUMUNDA PARAMETRE KESTİRİMİNE ALTERNATİF BİR YAKLAŞIM

Lineer olmayan modellerde iç ilişki problemi, teşhisi ve giderilmesi açısından lineer modellere göre daha karmaşık bir yapıya sahiptir. İç ilişki teşhisi için, lineer modelde X tasarım matrisinin sütunları arasındaki ilişki incelenirken, lineer olmayan modelde J Jacobian matrisinin sütunları incelenir. Jacobian matrisi sadece X matrisinin sütunlarına değil, θ parametre vektörüne de bağlı olduğundan, Jacobian matrisinin sütun vektörleri arasındaki ilişki, X matrisinin sütunları arasındaki ilişkiye bağlı olabilir ya da olmayabilir [9]. Bir diğer ifadeyle, lineer olmayan modelde tasarım matrisindeki iç ilişki, modeli lineerleştirdikten (düzlemleştirdikten) sonra oluşan yeni tasarım matrisine taşınmayabilir. Bir sonraki bölümde anlatılacak olan simülasyon çalışmasında, iç ilişkinin lineer olmayan modelden, lineer modele taşınması konusundan bahsedilecektir.

III.3.3.1 İTERATİF RIDGE REGRESYON

$J'J$, Jacobian matrisindeki iç ilişki yüzünden kötü koşullu ise Gauss-Newton yöntemi ile kestirilmek istenen (III.1) modelindeki parametreler için en küçük kareler yöntemi optimal bir yaklaşım vermeyecektir.

II. Bölümde de bahsedildiği gibi, lineer modelde iç ilişki olması durumunda kullanılan ridge kestirici, en küçük kareler kestiricisinden daha iyi sonuçlar verir. Bu anlamda, iç ilişki olması durumunda, (III.1) denklemini ile verilen lineer olmayan modelin parametre kestirimi için, Gauss-Newton yöntemi ile adımsal olarak uygulanan en küçük kareler kestiricisi yerine, ridge regresyon uygulanabilir. O halde, (III.1) lineer olmayan modelin ilk adımdaki lineerleştirilen (III.7) lineer modeli, β_0 yerine eşiti olan $\theta - \theta_0$ değeri yazılarak,

$$y_0 = J_0(\theta - \theta_0) + \varepsilon \quad (III.24)$$

şeklinde düzenlenirse, ilk adımda oluşturulan bu lineer model için parametre kestiriminde,

$$\widehat{\theta - \theta_0} = \hat{\theta}_1 - \theta_0 = (J_0' J_0)^{-1} J_0' y_0 \quad (III.25)$$

en küçük kareler kestiricisi kullanmak yerine, bu adımda,

$$\widehat{(\theta - \theta_0)_R} = \hat{\theta}_{1R} - \theta_0 = (J'_0 J_0 + k_0 I)^{-1} J'_0 y_0 \quad (\text{III.26})$$

ridge kestiricisi kullanmak yararlı olacaktır. O halde, (III.24) modeli için elde edilen ilk kestirim,

$$\hat{\theta}_1 = \theta_0 + (J'_0 J_0)^{-1} J'_0 y_0 \quad (\text{III.27})$$

yerine,

$$\hat{\theta}_{1R} = \theta_0 + (J'_0 J_0 + k_0 I)^{-1} J'_0 y_0 \quad (\text{III.28})$$

şeklinde elde edilmiş olur. Burada, ridge kestirici kullanmaktaki amaç, iç ilişki durumunda, en küçük kareler kestiricisinden daha iyi sonuçlar elde edilebilmektedir. Dolayısıyla (III.28) ile elde edilen kestiricinin (III.27) ile elde edilen kestiriciden daha iyi olduğunu gösterilmesi gerekir.

Hata kareler ortalaması karşılaştırması ile hangi parametre kestiriminin gerçek parametreye daha yakın olduğu gösterilebilir. (III.28) denklemiyle elde edilen, ridge kestirim için hata kareler ortalaması (III.27) denklemiyle elde edilen en küçük kareler kestiriminin hata kareler ortalamasıyla karşılaştırılırsa,

$$\begin{aligned} HKO(\hat{\theta}_R) &= E[(\hat{\theta}_R - \theta)'(\hat{\theta}_R - \theta)] \\ &= E[(\hat{\theta}_R - \theta + \theta_0 - \theta_0)'(\hat{\theta}_R - \theta + \theta_0 - \theta_0)] \\ &= E\{[(\hat{\theta}_R - \theta_0) - (\theta - \theta_0)]'[(\hat{\theta}_R - \theta_0) - (\theta - \theta_0)]\} \\ &= HKO(\hat{\theta}_R - \theta_0) \\ &< HKO(\hat{\theta} - \theta_0) \\ &= E\{[(\hat{\theta} - \theta_0) - (\theta - \theta_0)]'[(\hat{\theta} - \theta_0) - (\theta - \theta_0)]\} \\ &= E[(\hat{\theta} - \theta)'(\hat{\theta} - \theta)] \\ &= HKO(\hat{\theta}) \end{aligned} \quad (\text{III.29})$$

elde edilir. (III.29) eşitsizliğinde ridge kestiricisinin, en küçük kareler kestiricisinden daha küçük hata kareler ortalamaya sahip olduğu görülür. Burada, $\hat{\theta}_R - \theta_0$ ve $\hat{\theta} - \theta_0$ sırasıyla, (III.24) lineer modelindeki $\theta - \theta_0$ parametre vektörünün ridge ve en küçük

kareler kestiricileridir. $HKO(\hat{\theta}_R - \theta_0) < HKO(\hat{\theta} - \theta_0)$ eşitsizliğinin en az bir $k > 0$ yarıllık parametresi için sağlanacağı Hoerl ve Kennard (1970) tarafından ispatlanmıştır. Ridge kestirici gerçek parametreye daha yakın olduğundan, lineer olmayan modelde de parametre kestirimi için yeni bir yaklaşım olarak kullanılabilir.

Birinci iterasyonda modele yapılan bu yeni yaklaşım, parametre kestirim sürecinde yakınsaklık sağlanıncaya kadar, her iterasyonda iç ilişki incelemesi yapılarak, genelleştirilebilir. Bunun için, Bölüm IV de verilen simülasyon çalışması için bir algoritma geliştirilmiştir.

BÖLÜM IV

SİMÜLASYON ÇALIŞMASI

Lineer olmayan modelde iç ilişki olması durumunda, parametre kestirimine yeni bir yaklaşım olan iteratif ridge regresyon, üçüncü bölümde teorik olarak anlatılmıştır. Bu bölümde, teorisi anlatılan bu yeni yaklaşımın geçerliliğini göstermek için bir uygulaması verilecektir. Ancak, uygulamada, gerçek veriler üzerinden parametrelerin yanlılığı hesaplanamayacağından, gerçek veriler yerine bilgisayar ortamında üretilen yapay veriler kullanılmıştır.

IV.1. YÖNTEM

Veriler, Mathematica 7.0 programında, Roth ve Meyer (1972)' in makalelerinde kullandıkları veri ve $y = \frac{\theta_1 \theta_3 x_1}{1 + \theta_1 x_1 + \theta_2 x_2} + \varepsilon$ lineer olmayan model referans alınarak üretilmiştir. Çok değişkenli normal dağılımdan üretilen veriler, gerçek verilere yakınlık göstermesi açısından, gerçek verilerin ortalaması ve varyansı esas alınarak üretilmiştir. Hata terimi ise, $N(0,1)$ dağılımından üretilmiştir. Gerçeğe yakın olarak üretilen bu veriler kullanılarak, en küçük kareler kestiricisi ile bu tezde önerilen üç farklı kestirici karşılaştırılmıştır.

Önerilen ilk kestirici en küçük kareler kestiricisi,

$$\hat{\beta} = (X'X)^{-1} X'y \quad (IV.1)$$

ile ridge kestiricisi,

$$\hat{\beta}_R = (X'X + kI)^{-1} X'y \quad (IV.2)$$

birleşiminden oluşan ve iteratif olarak elde edilen bir kestiricidir. Bu kestirici elde edilirken, Bölüm IV.2 ile verilen algoritma izlenmiştir. Önerilen ikinci kestiricinin birinci kestiriciden tek farkı, iteratif olarak elde edilen ridge kestiricilerin elde

edilişinde kullanılan yanlılık parametresi k değerin farklı seçimidir. Son kestirici ise modelde iç ilişki olup olmadığına bakılmaksızın her adımda ridge regresyon uygulanarak elde edilen kestiricidir.

IV.2. ALGORİTMA

1. Lineer olmayan model Gauss-Newton yöntemiyle lineerleştirilir.
2. Oluşturulan lineer düzlemi geren, tasarım (Jacobian) matrisi elde edilir.
3. Tasarım matrisi birim uzunluk ölçeklemesi ile standartlaştırıldıktan sonra korelasyon formu elde edilir ($J'J$).
4. Oluşturulan bu korelasyon matrisi üzerinden iç ilişki teşhisi yapılır (örneğin, koşul sayısı incelenebilir).
5. İç ilişki teşhisine göre; eğer iç ilişki var ise bu adımda, birinci adımda elde edilen model üzerinden ridge regresyon yapılır ve ridge kestirim elde edilir, iç ilişki yok ise en küçük kareler yöntemi uygulanarak, en küçük kareler kestiricisi elde edilir.
6. Beşinci adımda elde edilen kestirici ile bir sonraki iteratif düzlem elde edilir ve ikinci adımdan itibaren işlemler devam ettirilir.
7. Yakınsaklık sağlanıncaya kadar devam eden bu iteratif süreç yakınsaklık sağlandığında sona erer ve elde edilen son parametre kestirim değeri olarak alınır.

Simülasyon çalışması için oluşturulan bu algoritma, Mathematica 7.0 programında kodlanarak yapılmıştır.

IV.3. SİMÜLASYON ANALİZİ

Lineer olmayan modellerde parametre kestirimi için iç ilişki olması durumunda, en küçük kareler yöntemine alternatif olarak önerilen iteratif ridge

regresyon, farklı n veri sayısı ve değişkenler arasındaki ρ korelasyon katsayısına göre incelenmiştir. Parametrelerin gerçek değerlerine yakınsaması ve hata kareler ortalamasındaki iyileşme, üretilen veri sayısı n ve değişkenler arasındaki ρ , korelasyon katsayısının farklı değerleri için değişim göstermiştir. Bununla birlikte, verilerin standartlaştırılması durumu da göz önünde bulundurularak analiz yapılmıştır. Ancak, standartlaştırılmamış veri ile yapılan analizlerin standartlaştırılmış verilere göre daha iyi sonuçlar verdiği görülmüştür.

Üretilen veriler, öncelikle, birim uzunluk ölçeklemesi ile standartlaştırılmıştır. Daha sonra, bulunan değerler, standartlaştırma işlemi yapmadan yapılan analizler ile karşılaştırılmıştır. Optimal parametre seçimi sürecinde, standartlaştırma ile ilgili en çok öne çıkan problemin, yakınsaklık problemi olduğu görülmüştür. Standartlaştırılmış veriler ile yapılan simülasyon çalışmasında, yapılan simülasyon çalışmasının yaklaşık %30'u yakınsaklık gösterirken, bu değer, standartlaştırılmamış veriler ile yaklaşık %50 değerine ulaştığı görülmüştür. Dolayısıyla, yapılan tüm simülasyon çalışmasında standartlaştırma yapılmadan parametre kestirimi yapılmıştır.

İşlemsel zorluklardan dolayı yapılan her simülasyon 100 tekrar ile yapılmıştır. Farklı parametre değerlerine göre yapılan her simülasyonda, Bölüm III de, kuramsal olarak gösterildiği gibi, varyansın azaldığı ve ortaya çıkan yanlılığın büyüklüğünün varyanstaki düşüşün büyüklüğünden fazla olmadığı görülmüştür. Dolayısıyla ridge kestiricisinin hata kareler ortalamasının, en küçük kareler kestiricisinin hata kareler ortalamasından daha düşük olduğu tespit edilmiştir. Yanlılık, tüm simülasyon sonucunda kestirilen parametrelerin ortalaması, beklenen değeri ile hesaplandığından, simülasyon sayısının fazla olması ile daha düşük değerler alabilir. Dolayısı ile yüksek boyutlu simülasyon çalışmalarının sonuçları daha makul sonuçlar verebilir.

Yapılan simülasyon çalışmasında, izlenilen bir diğer durum, Gauss-Newton yöntemi ile parametre kestirim sürecinde, başlangıç parametrelerinin uygun seçimidir. Bu durum yakınsaklığı büyük ölçüde etkilediğinden önemlidir. Farklı başlangıç parametreleri alınarak yapılan çalışmalarda, yakınsaklık oranlarının büyük

ölçüde değiştiği görülmüştür. Gerçek parametre değerinden uzak değerler başlangıç parametresi olarak seçildiğinde, yakınsaklığın zaman zaman sağlanmadığı ya da geç sağlandığı görülmüştür. Bununla birlikte, başlangıç parametresi, gerçek parametre değerine ne kadar yakın seçilirse, yakınsaklığın daha hızlı sağlandığı görülmüştür.

$n = 50, 100, 200, 300$ ve $\rho = 0.75, 0.80, 0.85$ değerleri göz önüne alınarak 11 farklı simülasyon çalışması yapılmıştır. Alınan sonuçlara göre, sunulan yeni yaklaşım, her simülasyonda, en küçük karelere göre daha iyi sonuçlar vermiştir. Tablo IV.1 de $n = 100$ ve $\rho = 0.75$ için ve Tablo IV.2 de $n = 50$ ve $\rho = 0.85$ için elde edilen sonuçlar verilmiştir.

Tablo IV.1 ve Tablo IV.2 de kestirilen parametrelerin hata kareler ortalama değerleri dört farklı kestirici için verilmiştir. Tablolarda “*” ile verilen hata kareler ortalaması dört kestirici arasında en iyi olanı göstermektedir. $HKO(\hat{\beta}_{R(k^*)})$ ile verilen hata kareler ortalamasındaki k^* değeri (II.38) denklemi kullanılarak hesaplanmıştır.

Tablo IV.1. $n = 100$ ve $\rho = 0.75$ için Hata Kareler Ortalama Değerleri

$HKO(\hat{\beta}_{EKK})$	$HKO(\hat{\beta}_{R(HKB)+EKK})$	$HKO(\hat{\beta}_{R(k^*)})$	$HKO(\hat{\beta}_{R(HKB)})$
17.6288	13.02687	10.92458*	13.02685
36.8682	19.86607	14.93999*	19.86605
20.8950	15.26487	12.10969*	15.26545
43.1319	19.04347	18.52799	15.29195*
11.4045	8.769070	7.52404*	8.76905
18.0447	12.28927	10.89067*	12.28925
14.4554	12.26377	11.63209*	12.26375
237.371	14.91427	9.515370*	14.91425
16.7145	8.19475	5.928390*	8.19473
59.9501	28.67407	22.59329*	28.67405
65.7715	10.79064	8.843980*	10.79062
112.848	32.47137	26.38399*	32.47135
19.3924	11.86057	10.01329*	11.86055
8.00672	7.432330	6.505100*	7.43231
11.4143	7.905230	7.218520*	7.90521
22.9103	14.12317	11.47419*	14.12315
8.90163	8.084840	7.425370*	8.08482
15.8127	1.942222	9.860140	1.942202*
527.970	27.60247	21.18769*	27.60245
15.5696	12.42897	10.66365*	12.42895
166.591	15.19917	8.125180*	15.19915
277.221	24.16167	14.93579*	24.16465

$$yan_{R(HKB)+Ekk}^2 = 1.61017, yan_{R(k^*)}^2 = 1.09449, yan_{R(HKB)}^2 = 1.61015$$

$n = 100, \rho = 0.75$

Tablo IV.2. $n = 50$ ve $\rho = 0.85$ için Hata Kareler Ortalama Değerleri

$HKO(\hat{\beta}_{EKK})$	$HKO(\hat{\beta}_{R(HKB)+EKK})$	$HKO(\hat{\beta}_{R(k^*)})$	$HKO(\hat{\beta}_{R(HKB)})$
35.8456	13.250281*	13.884777	13.250282
823.532	17.162281	10.284877*	17.162282
57.4161	11.204381*	12.858677	11.204382
252.342	9.7809510	5.7964970*	9.7809520
69.6839	10.465181	8.8189570*	10.4651820
38.0684	8.4627610	6.4887470*	8.4627620
15.0149	7.1733110	5.2566570*	7.1733120
178.318	33.625381	29.535177*	33.625382
20.7315	11.982981*	12.076477	11.982982
14.8563	6.9410210	6.0294470*	6.9410220
12.8275	8.1279110	6.2869370*	8.1279120
743.115	20.944581	9.9997870*	20.944582
4.19498	3.6202910	3.3786270*	3.6202920
4.60677	4.1089010	4.0100970*	4.1089020
54.7793	6.1441710	3.8076170*	6.1441720
$yan_{R(HKB)+Ekk}^2 = 0.258681, yan_{R(k^*)}^2 = 0.258577, yan_{R(HKB)}^2 = 0.258682$			
$n = 50, \rho = 0.85$			

SONUÇ VE TARTIŞMA

Regresyon modelleriyle yapılan çalışmalarda, yanıt değişkenini açıklayan değişkenlerin tam anlamıyla bağımsız ya da ilişkisiz olduğu beklenemez. Bu ilişkinin derecesine göre model parametrelerinin kestiriminde kullanılan en iyi yöntem araştırmacı tarafından seçilebilir. Örneğin, modeldeki açıklayıcı değişkenler arasındaki iç ilişkinin derecesi yok veya az ise en küçük kareler yöntemi bilinmeyen parametrelerin kestirilmesinde kullanışlıdır. Aksi takdirde, yani iç ilişkinin derecesi fazla ise, parametre kestirimi için sıklıkla kullanılan en küçük kareler yöntemi tutarlı sonuçlar vermez [4].

Çoklu iç ilişki problemini tamamen giderilmesi mümkün olmamasına rağmen, lineer modellerde, var olan ilişkinin derecesini azaltmak için, literatürde birçok yöntem ve analitik çözüm önerilmiştir. Hoerl ve Kennard (1970)'ın önerdiği ridge regresyon en çok kullanılan analitik çözümler arasındadır. Lineer modellerde kullanılan ridge regresyon ile yanlı fakat tutarlı ve daha küçük varyansa sahip kestirimlerin elde edilmesi amaçlanmaktadır. Bu anlamda, lineer modellerde iç ilişki olması durumunda, parametre kestirimi için kullanılan ridge regresyon, en küçük karelere göre daha güvenilir çıkarsamalar verir.

Lineer olmayan modellerde parametre kestirim süreci, lineer modellere göre daha karmaşık bir yapıya sahiptir. Model uzayda bir eğri veya yüzey oluşturduğundan, iç ilişkinin teşhis edilmesi ve giderilmesi lineer modellerden daha farklıdır. Lineer regresyon modellerinde, X tasarım matrisinin sütunları arasındaki lineer ilişki incelenirken, lineer olmayan modellerde Jacobian matrisinin sütunları incelenir. Jacobian matrisi sadece X matrisinin sütunlarına değil, θ parametre vektörüne de bağlı olduğundan Jacobian matrisinin sütun vektörleri arasındaki ilişki, X matrisinin sütunları arasındaki ilişkiye bağlı olabilir ya da olmayabilir [9].

Lineer olmayan modellerde, Jacobian matrisi ile elde edilen korelasyon matrisi, J , Jacobian matrisindeki iç ilişki yüzünden kötü koşullu ise, model parametrelerinin Gauss-Newton yöntemi ile en küçük kareler kullanarak kestirilmesi yanlış

çıkarsamalara neden olabilir. Bu aşamada, lineer modellerde iç ilişki olması durumunda, en küçük karelere göre daha iyi bir kestirici olan ridge kestiricisi bu tezde lineer olmayan modellere uyarlanmıştır. İç ilişki teşhisi X matrisi üzerinden değil J Jacobian matrisi üzerinden yapılmıştır. Lineer olmayan modellerde parametre kestirimi iteratif bir süreç olduğundan, yakınsaklık sağlanıncaya kadar her iterasyonda, iç ilişki teşhisi yapılmış ve buna bağlı olarak kestirici tercih edilmiştir. İç ilişki olması durumunda ridge kestirici, iç ilişki olmadığı durumda ise en küçük kareler kestirici tercih edilmiştir. Böylece, yakınsaklık sağlandığında elde edilen kestirim değeri, her iki kestiricinin kombinasyonundan elde edilmiştir. Bu yaklaşımla elde edilen sonuçlar doğrultusunda, kestirim değerlerinin daha tutarlı olduğu teorik olarak ve hata kare ortalamaları bir simülasyon çalışması ile karşılaştırılarak, gösterilmiştir.

Literatürde lineer regresyon modellerinde iç ilişki olması durumunda ridge kestiricisine alternatif olarak birçok yanlı kestirici önerilmiştir. Yapılan bu çalışma temel alınarak, lineer olmayan regresyon modellere de alternatif kestiricilerin uygulanabilirliği incelenebilir.

KAYNAKLAR

- [1] Belsley, D. A.: “Conditioning Diagnostics Collinearity and Weak Data in Regression”, John Willey & Sons, New York, USA, **(1991)**, 19-79, 342-377
- [2] Hoerl, A. E.; Kennard, R., W.: “Ridge Regression: Biased Estimation For Nonorthogonal Problems”, *Technometrics*, 12(1), **(1970a)**, 55-67
- [3] Hoerl, A. E.; Kennard, R., W.: “Ridge Regression: Application to Nonorthogonal Problems”, *Technometrics*, 12(1), **(1970b)**, 69-82
- [4] Montgomery, D.C.; Peck, A.E.; Vining, G.G.: “Introduction to the Linear Regression Analysis”, *John Willey & Sons*, Canada, **(2001)**.
- [5] Rencher, A., C.; Schaalje, G., B.: “Linear Models in Statistics”, *John Willey & Sons*, Hoboken, New Jersey, **(2008)**, 137-185.
- [6] Seber, G., A., F.; Wild, C., J.: “Nonlinear Regression”, *John Willey & Sons*, Hoboken, New Jersey, **(2003)**, 1-127
- [7] Rawlings, J., O.; Pantula, S., G.; Dickey, D., A.: “Applied Regression Analysis: A Research Tool”, Second Edition, Springer, New York, **(1998)**
- [8] Belsley, D., A.; Kuh, E.; Welsch, R., E.: “Regression Diagnostics Identifying Influential Data and Sources of Collinearity”, *John Willey & Sons*, Hoboken, New Jersey, **(1980)**.
- [9] Hill, R., C.; Adkins, L., C.: “Collinearity”, In Baltagi, B., eds.: “A Companion to Theoretical Econometrics”, *Blackwell Publishing*, Germany, **(2003)**, 256-277
- [10] Groß, J.: “Linear Regression”, *Springer*, Germany, **(2003)**
- [11] Schaefer, R., L.; Roi, L., D.; Wolfe, R., A.: “A Ridge Logistic Estimator”, *Commun. Statist. –Theor. Meth.*, 13(1), **(1984)**, 99-113.
- [12] Webster, J., T.; Gunst, R., F.; Mason, R., L.: “Latent Root Regression Analysis”, *Technometrics*, 16, **(1974)**, 513-522
- [13] Gallant, A., R.: “Nonlinear Statistical Models”, *John Willey & Sons*, New York, **(1987)**

- [14] Magel, R., C.; Hertsgaard, D.: "A Collinearity Diagnostic for Nonlinear Regression", *Communications in Statistics - Simulation and Computation*, 16(1), (1987), 85-97
- [15] Meyer, R., R.; Roth, P., M.: "Modified Damped Least Squares: An Algorithm for Non-linear Estimation", *J. Inst. Maths Applies*, 9, (1972), 218-233

ÖZGEÇMİŞ

Ali Erkoç, 1985 yılında Yozgat' ın Sarınınören Köyü'nde doğdu. İlk ve orta öğrenimini Prof. Muharrem Ergin İlköğretim Okulunda ve lise öğrenimini İstanbul Nuri Cıngıllıoğlu (Süper) Lisesi'nde tamamladı. 2003 yılında Trakya Üniversitesi Fen Edebiyat Matematik Bölümüne girdi. Bir sene sonra, Marmara Üniversitesi Fen-Edebiyat Fakültesi Matematik Bölümü'ne yatay geçiş yaptı. 2007 yılında bu bölümden derece ile mezun olduktan sonra aynı yıl Marmara Üniversitesi Fen Bilimleri Enstitüsü Matematik Ana Bilim Dalı Uygulamalı Matematik Programı'nda yüksek lisans yapmaya hak kazandı.