

**T.C.
MİMAR SİNAN GÜZEL SANATLAR ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

**DOĞRUSAL OLMAYAN REGRESYON MODELLERİNDE
İÇ İLİŞKİ OLMASI DURUMUNDA
PARAMETRE KESTİRİMİ**

DOKTORA TEZİ

ALİ ERKOÇ

İstatistik Anabilim Dalı

İstatistik Programı

Tez Danışmanı: Prof. Dr. M. Aydın ERAR

HAZİRAN 2017

**T.C.
MİMAR SİNAN GÜZEL SANATLAR ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

**DOĞRUSAL OLMAYAN REGRESYON MODELLERİNDE
İÇ İLİŞKİ OLMASI DURUMUNDA
PARAMETRE KESTİRİMİ**

DOKTORA TEZİ

**ALİ ERKOÇ
(20127209)**

İstatistik Anabilim Dalı

İstatistik Programı

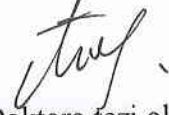
Tez Danışmanı: Prof. Dr. M. Aydın ERAR

HAZİRAN 2017

Ali ERKOÇ tarafından hazırlanan DOĞRUSAL OLMAYAN REGRESYON MODELLERİNDE İÇ İLİŞKİ OLMASI DURUMUNDA PARAMETRE KESTİRİMİ adlı bu tezin Doktora tezi olarak uygun olduğunu onaylarım.

Prof. Dr. M. Aydın ERAR

Tez Yöneticisi



Bu çalışma, jürimiz tarafından İstatistik Anabilim Dalında Doktora tezi olarak kabul edilmiştir.

Başkan: : Prof. Dr. M. Aydın ERAR



Üye : Prof. Dr. Gülay BAŞARIR



Üye : Prof. Dr. Müjgan TEZ



Üye : Prof. Dr. Ali Hakan BÜYÜKLÜ



Üye : Doç. Dr. Barış AŞIKGİL



Bu tez, Mimar Sinan Güzel Sanatlar Üniversitesi Fen Bilimleri Enstitüsü tez yazım kurallarına uygundur.

ÖNSÖZ

Bu çalışmada, bilgisi ve tecrübesi ile her konuda bana destek olan, hiçbir sorumu cevapsız bırakmayan, tez danışmanım Sayın Prof. Dr. M. Aydın ERAR' a katkıları ve emeği için çok teşekkür ederim. Her konuda bana yardımcı olan, bilgisini paylaşmaktan çekinmeyen Sayın Doç. Dr. Kadri Ulaş AKAY' a desteğinden dolayı teşekkür ederim. Konuyu çalışmamda bana yol gösteren, anlamamda büyük katkısı olan Sayın Prof. Dr. Müjgan TEZ' e teşekkürlerimi bir borç bilirim. Ayrıca, çalışmamda emeği olan tüm hocalarım ve arkadaşlarıma teşekkür ederim.

Hayatım boyunca bana her türlü desteği veren ve yanımda olan, annem, babam ve kardeşime, manevi desteğini her zaman hissettiğim hayat arkadaşşıma ayrıca çok teşekkür ederim.

Haziran, 2017

Ali ERKOÇ

İÇİNDEKİLER

ÖNSÖZ.....	i
İÇİNDEKİLER.....	ii
ÖZET.....	iv
SUMMARY.....	v
ÇİZELGE LİSTESİ.....	vi
ŞEKİL LİSTESİ.....	vii
SEMBOL LİSTESİ.....	ix
KISALTMALAR LİSTESİ.....	xi
1. GİRİŞ.....	1
2. DOĞRUSAL OLMAYAN REGRESYON MODELLERİNDE PARAMETRE KESTİRİMİ.....	6
2.1. Genel Yapı.....	6
2.2. Doğrusal Olmayan En Küçük Kareler Yöntemi.....	7
2.3. Doğrusal Olmayan En Küçük Kareler Yönteminin Geometrisi.....	7
2.4. Modelin Doğru Belirlenmesi.....	13
2.5. Gauss – Newton Yöntemi.....	16
2.6. Levenberg – Marquardt Yöntemi.....	18
3. DOĞRUSAL OLMAYAN REGRESYON MODELLERİNDE ÇOKLU İÇ İLİŞKİ PROBLEMİ ve YENİ BİR YAKLAŞIM.....	20
3.1. Çoklu İç İlişki Teşhisi.....	20
3.2. Kötü Koşulluluk.....	23
3.3. Çoklu İç İlişkinin Etkileri.....	24

3.4. X Matrisindeki İç İlişkinin J Jakobiyen Matrisindeki İç İlişki İle Bağlantısı	27
3.5. Optimizasyon Problemi.....	30
3.6. İç İlişki Problemine Yeni Yaklaşım	31
3.6.1. Ön Bilgiye Dayalı Yeni Kestirici	35
3.6.2. Hata Kareler Ortalamaları Karşılaştırılması.....	41
4. BENZETİM ÇALIŞMASI.....	44
4.1. Yöntem.....	45
4.2. Sonuçlar	46
4.3. Sonuçların Yorumlanması.....	67
5. UYGULAMA	69
6. SONUÇ VE TARTIŞMA.....	72
KAYNAKLAR.....	74
EKLER.....	76
EK A 1. Bazı Teorem Ve İspatları	77
EK B 1. Diğer Çizelge Ve Grafikler	81
ÖZGEÇMİŞ.....	103

ÖZET

Bilim ve teknolojinin gelişimi ile verilerin bilgisayar ortamında modellenmesi ve geleceğe dönük tahmin yöntemlerinin geliştirilmesi popüler hale gelmiştir. Fizik, kimya, mühendislik, tıp, uzay bilimleri gibi uygulamalı temel bilimlerde, elde edilen verilerin modellenmesi ile bir sonraki adımın tahmini önem kazanmıştır.

Doğrusal modeller ile modellenebilecek veriler olmasına rağmen, oluşturulan modeller, çoğunlukla, diferansiyel denklem sistemlerinin çözümünden elde edildiği için genellikle doğrusal olmayan fonksiyonlar ile belirtilir. Örneğin, bir uzay aracının ya da bir gök cisminin yörüngesinin belirlenmesi için doğrusal olmayan regresyon modelleri kullanılır. Dolayısıyla yörüngenin doğru tahmini için parametrelerin tutarlı kestirimi önem kazanmaktadır.

Regresyon analizinde, iç ilişki problemi, parametrelerin tutarlı ve güvenilir kestirimini engelleyen bir problemdir. Doğrusal olmayan regresyonda, parametrelerin güvenilir ve tutarlı olacak şekilde kestirilmesi, modelin veriyi en iyi şekilde temsili ve tutarlı kestirimlerin yapılması bakımından çok önemlidir.

Bu amaçla, bu tezde, doğrusal olmayan regresyon modellerinde iç ilişki olması durumunda parametre kestirimine yeni bir yaklaşım sunulmuştur. Yapılan benzetim çalışması ile, önerilen yaklaşımın geçerliliği test edilmiştir. Son olarak, teorisi ve benzetim çalışması ile anlatılan yöntem, bir örnek ile uygulamaya geçirilmiştir. Böylelikle birçok bilim dalında ihtiyaç olan doğrusal olmayan regresyon modellerinin, daha tutarlı ve güvenilir parametre kestirimlerine sahip olmaları konusunda bir kestirim yöntemi literatüre kazandırılmıştır.

SUMMARY

With the advancement of science and technology, the computer modeling of data and the development of future predictive methods have become popular. By modelling of obtained data, the estimation of the next step is gained importance, specifically in applied basic sciences such as physics, chemistry, engineering, medicine, space sciences.

Although these data sets can be modelled by using linear models, the generated models are often specified by nonlinear functions, since they are derived from solving systems of differential equations. For instance, the orbit of a spacecraft or a celestial body is generally determined by nonlinear regression models. Therefore, consistent estimation of the parameters is important for the accurate estimation of the orbit.

In regression analysis, the multicollinearity is a problem that prevents consistent and reliable estimation of parameters. In nonlinear regression, the estimation of reliable and consistent parameters is crucial to make consistent predictions of the model and to represent data as good as possible.

For this purpose, in this thesis, a new approach to parameter estimation is presented in case of multicollinearity in nonlinear regression models. The validity of the proposed approach was tested with the simulation study. Finally, the method explained both by a theoretical and methodological aspect is put into practice with an example. Thus, a predictive method to have more consistent and reliable parameter estimates in nonlinear regression models that are used in many fields of science is gained to the literature.

ÇİZELGE LİSTESİ

Sayfa No

Çizelge 4.1 $n=10$ ve $\sigma=0.1$ için benzetim sonuçları	46
Çizelge 4.2 $n=10$ ve $\sigma=1$ için benzetim sonuçları.....	49
Çizelge 4.3 $n=10$ ve $\sigma=3$ için benzetim sonuçları	52
Çizelge 4.4 $n=20$ ve $\sigma=0.1$ için benzetim sonuçları	55
Çizelge 4.5 $n=20$ ve $\sigma=1$ için benzetim sonuçları	57
Çizelge 4.6 $n=20$ ve $\sigma=3$ için benzetim sonuçları	59
Çizelge 4.7 $n=50$ ve $\sigma=0.1$ için benzetim sonuçları	61
Çizelge 4.8 $n=50$ ve $\sigma=1$ için benzetim sonuçları	63
Çizelge 4.9 $n=50$ ve $\sigma=3$ için benzetim sonuçları	65
Çizelge 5.1 Örnek verileri	69
Çizelge 5.2 Verilen örnek için parametre kestirim değerleri	70
Çizelge 5.3 Kestirilen modellerin press değerleri	71
Çizelge 6.1 Benzetim çalışması için özet tablo	73
Çizelge B.1 $n=10$ ve $\sigma=0.5$ için benzetim sonuçları.....	81
Çizelge B.2 $n=10$ ve $\sigma=2$ için benzetim sonuçları	83
Çizelge B.3 $n=20$ ve $\sigma=0.5$ için benzetim sonuçları	85
Çizelge B.4 $n=20$ ve $\sigma=2$ için benzetim sonuçları	87
Çizelge B.5 $n=50$ ve $\sigma=0.5$ için benzetim sonuçları	89
Çizelge B.6 $n=50$ ve $\sigma=2$ için benzetim sonuçları	91
Çizelge B.7 $n=100$ ve $\sigma=0.1$ için benzetim sonuçları	93
Çizelge B.8 $n=100$ ve $\sigma=0.5$ için benzetim sonuçları.....	95
Çizelge B.9 $n=100$ ve $\sigma=1$ için benzetim sonuçları	97
Çizelge B.10 $n=100$ ve $\sigma=2$ için benzetim sonuçları.....	99
Çizelge B.11 $n=100$ ve $\sigma=3$ için benzetim sonuçları	101

ŞEKİL LİSTESİ

Sayfa No

Şekil 1.1 Doğrusallaştırmanın geometrik görünümü	3
Şekil 1.2 Model Kestirim Eğrileri	5
Şekil 2.1 Hata Kareler Toplamı Fonksiyonu	9
Şekil 2.2 Doğrusal olmayan bir fonksiyon grafiği.....	10
Şekil 2.3 Doğrusal olmayan fonksiyona ait kontur grafiği	10
Şekil 2.4 Doğrusal en küçük kareler yönteminin geometrisi	11
Şekil 2.5 Yüzeğe çizilen tanjant düzlemleri.....	12
Şekil 2.6 y' nin hiperdüzleme izdüşümü.....	12
Şekil 2.7 Yüksek dereceli doğrusal olmayan hiperyüzey	13
Şekil 2.8 $S(\theta)$ fonksiyonunun doğru tanımlanmış minimum değeri	14
Şekil 2.9 $S(\theta)$ fonksiyonunun doğru tanımlanmamış minimum değeri	15
Şekil 3.1 $n = 30$ iken X ile J arasındaki iç ilişki incelemesi.....	28
Şekil 3.2 $n = 50$ iken X ile J arasındaki iç ilişki incelemesi.....	29
Şekil 3.3 En küçük kareler ve ridge kestiricilerinin parametre uzayındaki yeri	30
Şekil 3.4 Farklı kısıt altında kestiricilerin geometrik yeri.....	31
Şekil 3.5 İki boyutlu uzayda $\hat{\beta}_{0,\phi}$ kestiricisinin konumu.....	35
Şekil 3.6 Ön bilgiye dayalı kestiricinin parametre uzayındaki konumu	39
Şekil 4.1 $n = 10$ ve $\sigma = 0.1$ için parametre dağılım grafikleri	47
Şekil 4.2 $n = 10$, $\sigma = 0.1$ ve $\rho = 0.85$ için parametre dağılım grafikleri	48
Şekil 4.3 $n = 10$ ve $\sigma = 1$ için parametre dağılım grafikleri	50
Şekil 4.4 $n = 10$, $\sigma = 1$ ve $\rho = 0.999$ için parametre dağılım grafikleri	51
Şekil 4.5 $n = 10$ ve $\sigma = 3$ için parametre dağılım grafikleri.....	53
Şekil 4.6 $n = 10$, $\sigma = 3$ ve $\rho = 0.999$ için parametre dağılım grafikleri.....	54
Şekil 4.7 $n = 20$ ve $\sigma = 0.1$ için parametre dağılım grafikleri.....	56
Şekil 4.8 $n = 20$ ve $\sigma = 1$ için parametre dağılım grafikleri	58
Şekil 4.9 $n = 20$ ve $\sigma = 3$ için parametre dağılım grafikleri.....	60
Şekil 4.10 $n = 50$ ve $\sigma = 0.1$ için parametre dağılım grafikleri.....	62
Şekil 4.11 $n = 50$ ve $\sigma = 1$ için parametre dağılım grafikleri	64
Şekil 4.12 $n = 50$ ve $\sigma = 3$ için parametre dağılım grafikleri.....	66

Şekil 4.13 Farklı iterasyonlardan sonuçlar	67
Şekil 4.14 Jakobiyen matrisinden hesaplanan koşul sayılarının değişim grafiği	68
Şekil 5.1 Ardıl artık kareler toplamı grafikleri.....	70
Şekil 5.2 Kestirim modellerinin farklı açılardan grafikleri.....	71
Şekil B.1 $n = 10, \sigma = 0.5$ ve $\rho = 0.99$ için parametre dağılım grafikleri.....	82
Şekil B.2 $n = 10, \sigma = 2$ ve $\rho = 0.85$ için parametre dağılım grafikleri	84
Şekil B.3 $n = 20, \sigma = 0.5$ ve $\rho = 0.80$ için parametre dağılım grafikleri	86
Şekil B.4 $n = 20, \sigma = 2$ ve $\rho = 0.99$ için parametre dağılım grafikleri	88
Şekil B.5 $n = 50, \sigma = 0.5$ ve $\rho = 0.85$ için parametre dağılım grafikleri.....	90
Şekil B.6 $n = 50, \sigma = 2$ ve $\rho = 0.999$ için parametre dağılım grafikleri	92
Şekil B.7 $n = 100, \sigma = 0.1$ ve $\rho = 0.90$ için parametre dağılım grafikleri	94
Şekil B.8 $n = 100, \sigma = 0.5$ ve $\rho = 0.99$ için parametre dağılım grafikleri.....	96
Şekil B.9 $n = 100, \sigma = 1$ ve $\rho = 0.99$ için parametre dağılım grafikleri	98
Şekil B.10 $n = 100, \sigma = 2$ ve $\rho = 0.85$ için parametre dağılım grafikleri	100
Şekil B.11 $n = 100, \sigma = 3$ ve $\rho = 0.85$ için parametre dağılım grafikleri.....	102

SEMBOL LİSTESİ

x	: Tasarım matrisi
X	: Doğrusal modellerdeki tasarım matrisi
y	: Yanıt vektörü
θ	: Doğrusal olmayan modellerdeki parametre vektörü
ε	: Hata vektörü
$f(\cdot)$: Doğrusal olmayan model
$S(\theta)$: Amaç fonksiyonu
$\hat{\theta}$: Doğrusal olmayan parametre kestirici vektörü
J	: Jakobiyen matrisi
j_i	: Jakobiyen matrisinin i . sütunu
σ	: Hatanın standart sapması
$\hat{\beta}$: Doğrusal modeldeki parametre kestirici vektörü
β	: Doğrusal modellerdeki parametre vektörü
$L(\cdot)$: Lagrange fonksiyonu
H_{ij}	: Hesyen matrisi
Λ	: Özdeğerler matrisi
I	: Birim matris
λ_i	: i . özdeğer
λ_{\max}	: En büyük özdeğer
λ_{\min}	: En küçük özdeğer
r_{ij}	: ij . korelasyon
p	: Parametre sayısı
n	: Gözlem sayısı
$\text{İz}(\cdot)$: Bir matrisin izi
ρ	: Korelasyon katsayısı
ϕ	: Lagrange çarpanı
$k(\beta)$: Kısıt fonksiyonu
k	: Ridge parametresi

b	: Stokastik olmayan ön bilgi vektörü
B	: Stokastik ön bilgi vektörü
M	: Önerilen kestiricinin ön bilgi vektörü
$\hat{\sigma}$: σ 'nın kestiricisi
Σ	: Varyans - kovaryans matrisi
$\tilde{\beta}$: β 'nın herhangi bir kestiricisi



KISALTMALAR LİSTESİ

HKO	: Hata kareler toplamı
KS	: Koşul sayısı
VIF	: Varyans şişirici faktör
MHKO	: Matris hata kareler ortalaması
SHKO	: Skaler hata kareler ortalaması
HKB	: Hoerl-Kennard-Baldwin



1. GİRİŞ

Çıkarsamalı istatistikte, rastlantı değişkenleri arasındaki bağıntıyı bulmak ya da modellemek önemli bir konudur. Regresyon analizi, bu anlamda, bağımlı ve bağımsız değişkenler arasında var olan bağıntıyı açıklamak ya da modellemek için kullanılan istatistiksel bir yöntemdir. Bu bağıntı bir f fonksiyonu yardımıyla,

$$y \approx f(x_1, x_2, \dots, x_k; \theta) \quad (1.1)$$

biçiminde açıklanabilir. Burada f fonksiyonu θ bilinmeyen parametreler vektörü hariç bilinen bir fonksiyondur. Bununla birlikte, uygulamada açıklanmaya çalışılan ilişki, çoğunlukla, karmaşık yapıda olup, modeli açıklayacak değişkenlerin tamamının bilinmesi imkansızdır. Dolayısıyla amaç, minimum hata ile, bulunabilecek en yakın (en doğru) modeli kestirmektir. ε hata terimi olmak üzere, bağıntı,

$$y = f(x_1, x_2, \dots, x_k; \theta) + \varepsilon \quad (1.2)$$

daha genel olarak, matris formunda,

$$y = f(x, \theta) + \varepsilon \quad (1.3)$$

şeklinde yazılabilir. Yukarıda yazılan model parametrik bir aile olup, amaç bağımlı değişkeni (y) en iyi açıklayacak modelin parametrelerini kestirmek, ara değer bulmak ve önkestirim yapmaktır.

Değişkenler arasındaki bağıntının modellenmesi için doğrusal regresyon analizi güçlü ve sık kullanılan bir yöntem olmasına rağmen, verinin yapısını açıklayacak model, çoğunlukla, doğrusal olmayan regresyon modelleri ile ifade edilir. Burada modelin doğrusal olmaması, fonksiyonun değişkenlere göre değil, parametrelere göre doğrusal olmamasından kaynaklanmaktadır. Yani, doğrusal regresyonda modelin parametrelere göre türevleri parametrelerden bağımsız iken; modelin en az bir parametreye göre

türevi, en az bir parametreye bağlı ise bu model doğrusal olmayan model olarak tanımlanır.

Doğrusal regresyon modelleri ile bağımlı ve bağımsız değişkenler arasındaki ilişkinin açıklandığı örnekler olmakla birlikte, fizik, biyoloji, kimya, ekonomi, mühendislik bilimleri vb. gibi temel bilimlerde yetersiz kalmıştır. Doğrusal olmayan modeller amaca özel modellerdir. Bu modellerin altında yatan bilimsel kuram, fiziksel olayları tanımlamak için matematiksel formülleri kullanır. Birçok durumda, fiziksel ilişkiyi tanımlayan bu formüller, diferansiyel denklemler kümesinin çözümünden elde edilir (Montgomery, et al., 2012). Örneğin, kimyada Puromycin verileri ile madde konsantrasyonu ve enzimatik reaksiyonun başlangıç hızı arasındaki ilişki, Michaelis-Menten modeli olarak bilinen,

$$y = \frac{\theta_1 x}{x + \theta_2} + \varepsilon$$

doğrusal olmayan modeli ile açıklanmıştır (Montgomery, et al., 2012). Bir başka örnek olarak, biyokimyasal oksijen ihtiyacı verileri,

$$f(x, \theta) = \theta_1(1 - e^{-\theta_2 x}) + \varepsilon$$

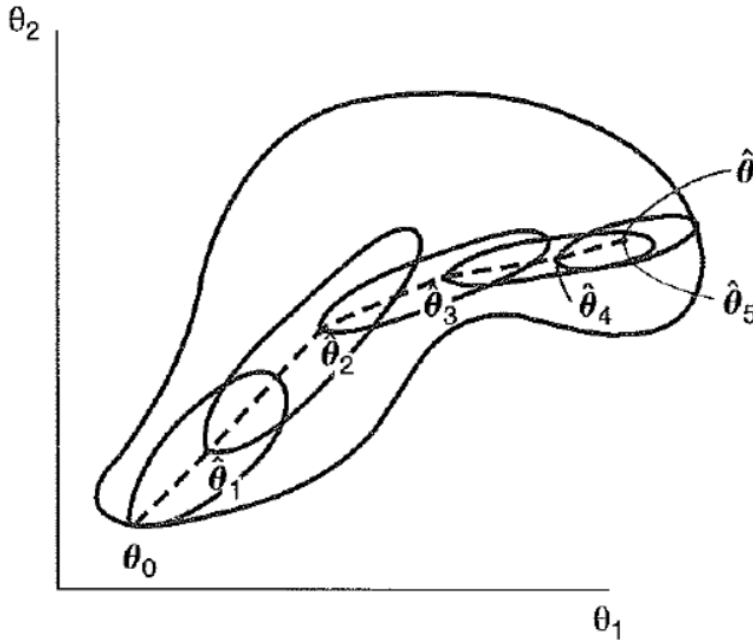
modeli ile açıklanmıştır (Bates & Watts, 1988). Her bilim dalından sayısız örnek vermek mümkündür. Durum böyle olunca, doğrusal olmayan regresyon modellerine olan ihtiyaç fazlasıyla artmıştır. Ancak, asıl amacın model kestirimi, dolayısıyla parametre kestirimi olmasından dolayı, parametre kestirimi doğrusal regresyon modellerinde olduğu kadar kolay değildir. Parametre kestirim yöntemi olarak, doğrusal regresyonda olduğu gibi, doğrusal olmayan regresyonda da en küçük kareler yöntemi, $\mathbf{x}_i^T = (x_{i1}, x_{i2}, \dots, x_{ip})$ ve $\boldsymbol{\theta}^T = (\theta_1, \theta_2, \dots, \theta_p)$ iken,

$$S(\boldsymbol{\theta}) = \sum_{i=1}^n [y_i - f(\mathbf{x}_i, \boldsymbol{\theta})]^2 \quad (1.4)$$

amaç fonksiyonunun minimize edilmesiyle çoğu zaman en iyi sonucu verebilir. Ancak doğrusal olmayan modellerin karmaşık yapısından dolayı, $j = 1, 2, \dots, p$ için,

$$\sum_{i=1}^n [y_i - f(x_i, \theta)] \left[\frac{\partial f(x_i, \theta)}{\partial \theta_j} \right]_{\theta=\hat{\theta}} \quad (1.5)$$

normal denklemlerinin çözümü, çoğunlukla, analitik olarak mümkün olmamaktadır. Dolayısıyla doğrusal olmayan regresyon modellerinde parametre kestiriminde iteratif süreç tercih edilmektedir (Seber & Wild, 2003). Gauss-Newton yöntemi doğrusal olmayan regresyon modellerinin kestiriminde en sık kullanılan iteratif ve sayısal bir yöntemdir. Bu yöntem, doğrusal olmayan regresyon modelinin Taylor serisi yardımıyla doğrusallaştırılması temeline dayanır. Doğrusallaştırma yöntemi, seçilen bir başlangıç noktasından başlanarak hesaplanan $S(\theta)$ amaç fonksiyonunun (artık kareler fonksiyonu) iteratif olarak minimize edilmesine dayanır. Bu sürecin geometrik görünümü Şekil 1.1’ de olduğu gibidir.



Şekil 1.1 Doğrusallaştırmanın geometrik görünümü

θ_0 başlangıç parametre vektörü olmak üzere iteratif olarak çizilen eliptik şekiller $S(\theta)$ amaç fonksiyonunun kontur çizimleridir. İteratif olarak amaç fonksiyonunun azalarak dengeye ulaştığı anda yakınsama sağlanır ve kestirilen parametre vektörü Şekil 1.1’ de olduğu gibi $\hat{\theta}$ olarak elde edilir.

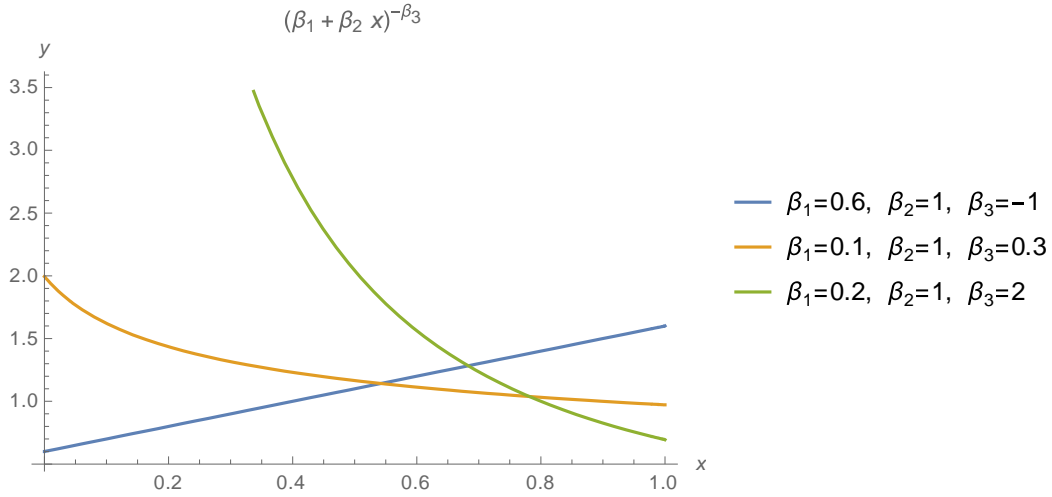
Parametre kestirim sürecinde karşılaşılan en büyük problemlerden bir tanesi, doğrusal regresyonda olduğu gibi, açıklayıcı değişkenler arasındaki iç ilişki (collinearity,

multicollinearity) problemidir. İç ilişki, doğrusal regresyonda, bağımsız değişkenler arasındaki doğrusal bağımlılık anlamına gelmektedir. Daha genel olarak, modeldeki açıklayıcı değişkenlerden birisi, diğer açıklayıcı değişkenlerin herhangi bir alt kümesinin tam doğrusal kombinasyonu olarak yazılıyorsa modelde tam çoklu iç ilişki (multicollinearity) var demektir (Belsley, 1991). Bu, açıklayıcı değişkenler matrisinin (tasarım matrisi) tam ranklı olmadığı anlamına gelir ve tek bir çözüm elde edilemez. Açıklayıcı değişkenlerden birisi, diğer açıklayıcı değişkenlerin herhangi bir alt kümesinin yaklaşık doğrusal kombinasyonu olarak yazılıyorsa, tasarım matrisinin rankı düşmez (Gruber, 2010). Ancak tasarım matrisi ile elde edilen korelasyon matrisi “kötü koşullu” olarak tanımlanır. Kötü koşulluluk parametre kestirim değerlerini olumsuz etkilemektedir. Bu durumda, parametre kestirim sürecinde izlenmesi gereken yol, söz konusu problemin çözümüne yönelik olmalıdır. Doğrusal olmayan modellerde de, bu problem, modelin yapısından kaynaklı olarak karşımıza çıkmaktadır. Fakat doğrusal modellerden farklı olarak, burada, çoklu iç ilişki tasarım matrisindeki değişkenler arasında değil, birinci türevlerden oluşan,

$$\mathbf{J} = \left[\frac{\partial f(\mathbf{x}, \boldsymbol{\theta})}{\partial \theta_j} \right] \quad j = 1, 2, \dots, p \quad (1.6)$$

Jakobiyen matrisinin sütunları arasındadır. Bir başka deyişle doğrusal olmayan modeldeki \mathbf{J} Jakobiyen matrisi, doğrusal modeldeki açıklayıcı değişkenler matrisinin rolünü üstlenmektedir (Magel & Hertsgaard, 1987). \mathbf{J} Jakobiyen matrisindeki iç ilişki yüzünden meydana gelen kötü koşulluluk parametre kestirimini olumsuz etkilemektedir.

Yukarıda da değindiğimiz gibi, veriyi açıklayacak en iyi modeli bulmak regresyon analizinin en temel amacıdır. Bu amaç doğrultusunda, en iyi modeli kestirmek demek, bilinmeyen parametrelerin en doğru ve en düşük hata kareler toplamına sahip olacak şekilde kestirilmesi demektir. Öyle ki, parametre kestirim sürecinde, iç ilişki problemi göz ardı edilirse elde edilen model veriyi en iyi açıklayan model olmayacaktır. Örnek olarak, farklı parametre değerlerine karşı elde edilen model eğrileri Şekil 1.2’ de verilmiştir.



Şekil 1.2 Model Kestirim Eğrileri

Yukarıdaki Şekil 1.2 incelenecek olursa, aynı veriler ve aynı model ile elde edilen farklı parametre kestirimleriyle oluşan kestirim modelinin eğrilerinin birbirinden ne kadar farklı olduğu görülebilir. Dikkat edilecek olursa ikinci parametre değerleri aynı, birinci parametre değerleri çok yakın olmasına rağmen, üçüncü parametredeki farklılıklar kestirim modelini fazlasıyla etkilemektedir. İç ilişkili modelde, bu problem göz ardı edilerek elde edilen parametre kestirim değerleri yanlış işarete sahip olabilirler (Montgomery, et al., 2012). Bu durum, yukarıdaki şekilden de izleneceği gibi, farklı kestirim değerlerinin, veriyi temsil etmeyen modellerin elde edilmesine sebep olur. Doğrusal olmayan modellerde parametre kestirim sürecinde iç ilişki problemi doğrusal modellerde olduğu kadar incelenmemiştir. Bu tezde, yukarıda bahsedilen iç ilişki problemine, doğrusal olmayan regresyon modellerinde karşılaşılabildiği durumunda, parametre kestirimine yönelik yeni bir yaklaşım geliştirilmiştir. Geliştirilen yaklaşım, doğrusal modellerde kullanılan ridge regresyonuna benzer yapıdadır. Doğrusal modellerde iteratif kestirim süreci olmadığından, ridge regresyon veriyi bir kez uygulanır. Burada önerilen yeni yaklaşım, veriyi iteratif olarak uygulanır ve iç ilişkili modelde, Gauss – Newton yöntemine göre daha küçük hata kareler ortalamasına sahip ve daha kararlı (tutarlı) kestirim değerleri elde edilir. Böylece veriyi en iyi temsil eden model kestirilmiş olur. Yeni yaklaşımın performansı yapılan benzetim çalışması ile test edilmiştir. Farklı parametre değerleri ile üretilen veriler yardımıyla önerilen yeni yaklaşım ile elde edilen kestiricilerin Gauss – Newton yöntemine göre büyük ölçüde üstünlük sağladığı görülmüştür. Ayrıca yöntem, gerçek verilere uygulanmış, performansının, yine, daha iyi olduğu gösterilmiştir.

2. DOĞRUSAL OLMAYAN REGRESYON MODELLERİNDE PARAMETRE KESTİRİMİ

Bu bölümde doğrusal olmayan regresyon modellerinde en çok kullanılan parametre kestirim yöntemleri ve doğrusal olmayan en küçük kareler yönteminin geometrisi ile modelin doğru belirlenmesi anlatılacaktır. Bir sonraki bölümde, çoklu iç ilişki probleminin varlığında, bu bölümde anlatılacak olan parametre kestirim yöntemlerine alternatif, yeni bir yaklaşım sunulacaktır.

2.1. GENEL YAPI

(x_i, y_i) $i = 1, 2, \dots, n$ n sayıda gözlem olmak üzere, bilinen bir f fonksiyonu yardımıyla, doğrusal olmayan regresyonun genel yapısı,

$$y_i = f(x_i, \theta) + \varepsilon_i, \quad i = 1, 2, \dots, n \quad (2.1)$$

şeklinde olup, matris formunda,

$$y = f(x, \theta) + \varepsilon \quad (2.2)$$

biçiminde yazılabilir. Burada x_i $k \times 1$ boyutunda açıklayıcı değişkenler vektörü; ε_i , $E(\varepsilon_i) = 0$ ve $Var(\varepsilon_i) = \sigma^2$ özelliklerine sahip hata terimidir. Genellikle hataların normal dağıldığı ve ilişkisiz olduğu varsayılır. Normallik varsayımı, parametre kestirimi için gerekli bir varsayım olmasa da, model kestirimi sonrası, parametre ve model çıkarımları, hipotez testleri, güven aralıkları gibi işlemlerde gerekli olmaktadır.

2.2. DOĞRUSAL OLMAYAN EN KÜÇÜK KARELER YÖNTEMİ

Doğrusal olmayan regresyon modellerinde parametreleri kestirmek için, doğrusal regresyonda olduğu gibi, hata kareler toplamı minimize edilebilir:

$$S(\boldsymbol{\theta}) = \sum_{i=1}^n \varepsilon_i^2 = \sum_{i=1}^n [y_i - f(\mathbf{x}_i, \boldsymbol{\theta})]^2 \quad (2.3)$$

Burada, amaç fonksiyonu, $\boldsymbol{\theta}$ bilinmeyen parametreler vektörünün fonksiyonudur. Bu fonksiyona hata kareler toplamı (HKT) fonksiyonu da denebilir. En küçük kareler kestiricisi olan $\hat{\boldsymbol{\theta}}$ değerini elde etmek için (2.3) ile verilen amaç fonksiyonunun $\boldsymbol{\theta}$ parametre vektörüne göre türevi alınıp, sıfıra eşitlenir ve aşağıdaki normal denklemler elde edilir:

$$\sum_{i=1}^n [y_i - f(\mathbf{x}_i, \boldsymbol{\theta})] \left[\frac{\partial f(\mathbf{x}_i, \boldsymbol{\theta})}{\partial \theta_j} \right]_{\theta_j = \hat{\theta}_j} = 0, \quad j = 1, 2, \dots, p \quad (2.4)$$

$\hat{\boldsymbol{\theta}}$ kestirim vektörü yukarıdaki (2.4) denklemi ile verilen normal denklemleri sağlayan çözüm vektörüdür. (2.4) ile ifade edilen normal denklemlerin sağ tarafındaki

$\left[\frac{\partial f(\mathbf{x}_i, \boldsymbol{\theta})}{\partial \theta_j} \right]_{\theta_j = \hat{\theta}_j}$ kısmi türevler ifadesi, f fonksiyonun doğrusal olmamasından dolayı,

parametrelere bağlıdır. Doğrusal regresyonda, bu ifade parametrelerden bağımsızdır. Dolayısıyla çözümü zor değildir. Ancak, (2.4) normal denklemlerinin çözümü, çoğunlukla, analitik olarak mümkün değildir (Seber & Wild, 2003). Fonksiyonun yapısından kaynaklanan bu problem, parametre kestirim sürecinde, iteratif yöntemler kullanılarak çözülmeye çalışılır.

2.3. DOĞRUSAL OLMAYAN EN KÜÇÜK KARELER YÖNTEMİNİN GEOMETRİSİ

Doğrusal olmayan regresyonda parametre kestirimi için en küçük kareler yöntemi kullanılabilir. (2.3) ile verilen hata kareler toplamı,

$$S(\boldsymbol{\theta}) = (\mathbf{y} - f(\mathbf{x}, \boldsymbol{\theta}))^T (\mathbf{y} - f(\mathbf{x}, \boldsymbol{\theta})) \quad (2.5)$$

matris formunda gösterilsin. Hata kareler toplamının geometrik olarak ne anlama geldiğini görebilmek için,

$$S(\boldsymbol{\theta}) = \|\mathbf{y} - f(\mathbf{x}, \boldsymbol{\theta})\|^2 \quad (2.6)$$

şeklindeki norm ile ifade edilmesi daha uygundur. $\|\mathbf{y} - f(\mathbf{x}, \boldsymbol{\theta})\|^2$ ifadesi Öklid uzayında $\mathbf{y} - f(\mathbf{x}, \boldsymbol{\theta})$ vektörünün uzunluğunun karesini göstermektedir. $S(\boldsymbol{\theta})$ fonksiyonunun (2.6) gösterimini minimize ettiğimizde, amaç geometrik olarak anlam kazanmaktadır. Çünkü bu ifadenin minimize edilmesi, \mathbf{y} ile $f(\mathbf{x}, \boldsymbol{\theta})$ fonksiyonunun arasındaki Öklid uzaklığının minimize edilmesi anlamına gelmektedir. Bir başka deyişle, kestirim uzayında, i . gözleme en yakın kestirim değerini elde etmek anlamındadır. (2.5) denklemini daha açık yazacak olursak, $S(\boldsymbol{\theta})$ aşağıdaki gibi elde edilir:

$$S(\boldsymbol{\theta}) = \mathbf{y}^T \mathbf{y} - 2\mathbf{y}^T f(\mathbf{x}, \boldsymbol{\theta}) + f(\mathbf{x}, \boldsymbol{\theta})^T f(\mathbf{x}, \boldsymbol{\theta})$$

Bu ifadenin $\boldsymbol{\theta}$ parametre vektörüne göre kısmi türevleri sıfıra eşitlenirse,

$$\left. \frac{\partial S(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} = 0$$

ya da,

$$\mathbf{J} = \left. \frac{\partial f(\mathbf{x}, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}$$

kısaltması ile normal denklemler,

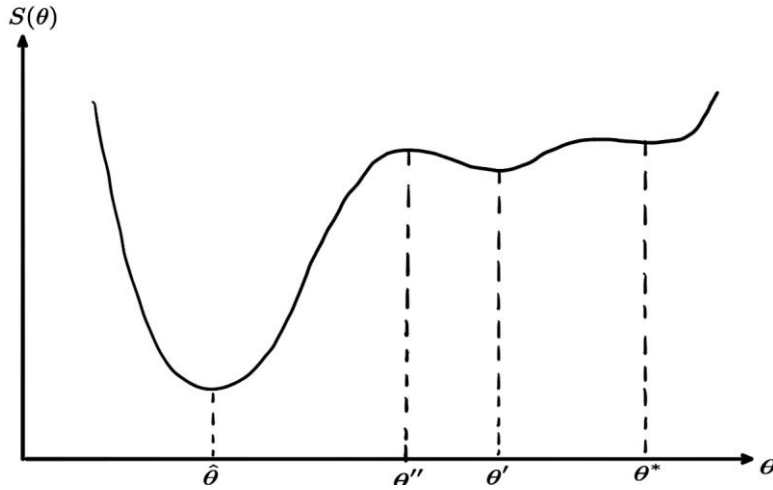
$$-2\mathbf{J}(\mathbf{x}, \hat{\boldsymbol{\theta}})^T \mathbf{y} + 2\mathbf{J}(\mathbf{x}, \hat{\boldsymbol{\theta}})^T f(\mathbf{x}, \hat{\boldsymbol{\theta}}) = 0 \quad (2.7)$$

şeklinde elde edilmiş olur. \mathbf{J} ile adlandırdığımız matris, sütunları türevlerden oluşan Jakobiyen (Jacobian) matristir. Bu matris, doğrusal olmayan regresyonda parametre kestirimi yaparken sıkça kullanılan bir matristir. Kısmi türevlerden oluştuğu için karmaşık yapıdadır. Bu yüzden (2.7) normal denklemlerinin analitik olarak çözülmesi genellikle mümkün değildir (Rawlings, et al., 1998). Normal denklemler daha sade olarak,

$$\mathbf{J}(\mathbf{x}, \hat{\boldsymbol{\theta}})^T [\mathbf{y} - f(\mathbf{x}, \hat{\boldsymbol{\theta}})] = 0 \quad (2.8)$$

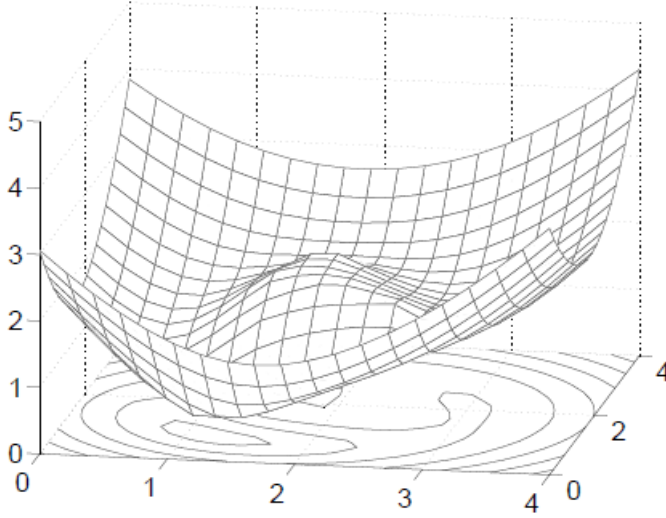
denklemini ile ifade edilebilir. (2.8) normal denklemlerine göre, normal koşullarda, $\mathbf{y} - f(\mathbf{x}, \hat{\boldsymbol{\theta}})$ artıklarının $\mathbf{J}(\mathbf{x}, \hat{\boldsymbol{\theta}})^T$ Jakobiyen matrisine dik olması gerektiği söylenebilir. Bu durumun doğrusal modellerdeki karşılığı, \mathbf{X} matrisinin $\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}$ artıklarına dik olmasıdır.

Yukarıda ifade edilen (2.8) normal denklemlerini sağlayan fakat istenilen çözümü vermeyen noktalar olabilir. Bir başka deyişle, en küçük kareler kestiricisi $\hat{\boldsymbol{\theta}}$ nın, global minimum noktası olması için, (2.8) normal denklemleri sağlaması gerekli fakat yeterli değildir. (2.8) normal denklemlerini sağlayan fakat, yerel minimum, yerel maksimum gibi noktalar olabilir; Şekil 2.1 de bu duruma örnek gösterilmiştir.



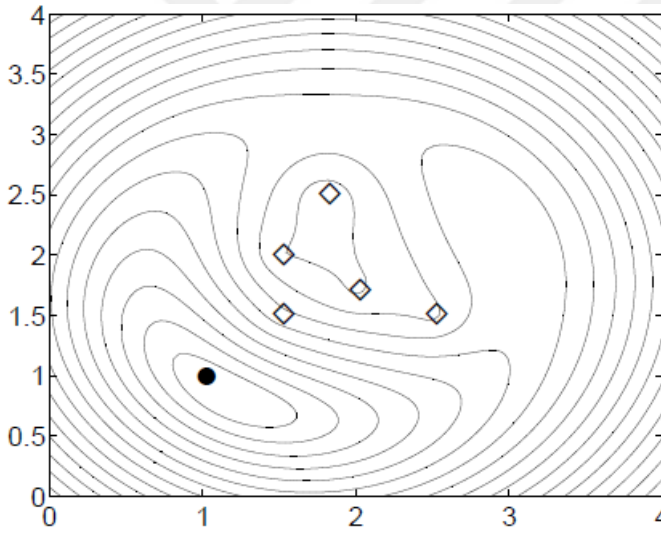
Şekil 2.1 Hata Kareler Toplamı Fonksiyonu

Şekil 2.1' de görüldüğü gibi hata kareler toplamı fonksiyonunun dört farklı noktada türevleri sıfır yani dört farklı noktada ekstremum noktası vardır. Fakat bunlardan bir tanesi, $\hat{\boldsymbol{\theta}}$ ile gösterilen değer, aranan kestirim değeridir. Bunun dışında θ' ile gösterilen nokta yerel minimum, θ'' ile gösterilen nokta yerel maksimum, θ^* ile gösterilen nokta normal denklemleri sağlayan fakat istenilen çözüm olmayan bir noktadır. Bu durum modelin karmaşık yapıda olması ve yüksek dereceli doğrusal olmamasından kaynaklanabilmektedir. Şekil 2.2'de de görüleceği gibi modelin farklı noktalarda minimum noktaları vardır. Ancak bunlardan bir tanesi global minimumdur. Global minimum (1.18 , 0.82), lokal minimum (2.99 , 2.12), lokal maksimum ise (1.94 , 1.87) dir. Bu üç nokta da normal denklemleri sağlayan noktalaradır.



Şekil 2.2 Doğrusal olmayan bir fonksiyon grafiği

Yine Şekil 2.2’ de grafiği verilen modele ait kontur grafiğini gösteren Şekil 2.3’ de de modelin farklı noktalarda minimum ve maksimum değerlerinin olduğu görülebilir.



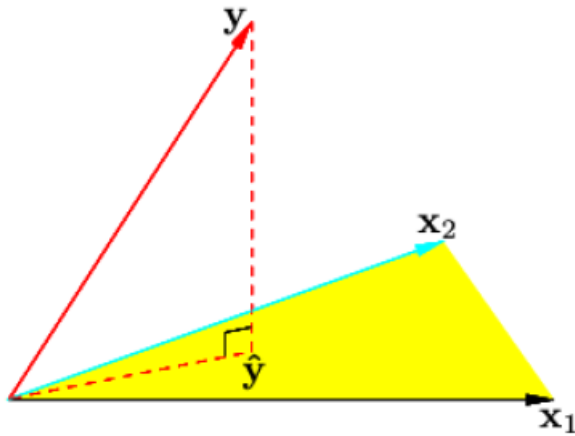
Şekil 2.3 Doğrusal olmayan fonksiyona ait kontur grafiği

Doğrusal olmayan regresyonda parametre kestiriminin zorluğu kontur çizimlerine bakılarak anlaşılabilir. Doğrusal regresyondaki kontur çizimi incelenecek olursa, düzgün iç içe geçmiş eliptik şekiller ve ortada tek minimum nokta olan en küçük kareler kestiricisi olduğu görülebilir. Burada ise birden çok ekstramum nokta vardır. Siyah nokta ile gösterilen nokta aranan global minimum değeridir.

Normal denklemleri sağlayan noktalardan hangilerinin hata kareler toplamını minimum yaptığını bulmak için 2. türev yöntemi kullanılabilir (Davidson &

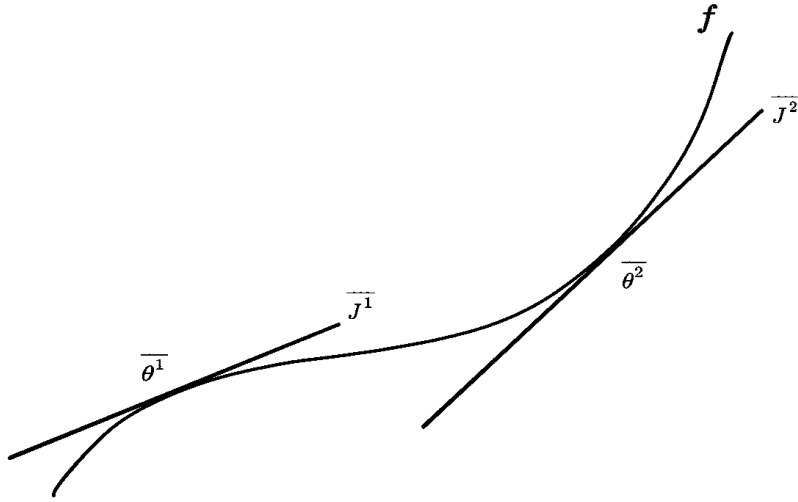
Mackinnon, 1993). Ancak yine de bu noktanın yerel minimum mu yoksa global minimum mu olduğu tespit edilemez. Analitik olarak bulunamayan global minimum noktasını bulabilmek için hata kareler toplamı fonksiyonu farklı başlangıç noktalarından başlanarak, birkaç kez minimize edilmeye çalışılabilir. (Davidson & Mackinnon, 1993). Örneğin Şekil 2.1 deki durum için, başlangıç noktası olarak θ' noktasının solunda herhangi bir nokta tercih edilirse doğru noktaya yakınsama, hem daha hızlı hem de daha olanaklıdır. Bir boyutlu durumlar için grafiksel yöntemler ile başlangıç noktasının seçimi ve global minimum noktasının kontrolü sağlanabilir. Ancak grafiksel yöntemlerin kullanılmadığı çok boyutlu durumlarda yakınsamanın global minimum noktası olmasının garantisi yoktur.

Doğrusal regresyonda, parametrelerin en küçük kareler kestiricisini geometrik olarak bulmak için, X tasarım matrisinin sütunları tarafından gerilen alt uzaya, Şekil 2.4' de gösterildiği gibi, gözlem vektörünün izdüşüm vektörü bulunur.



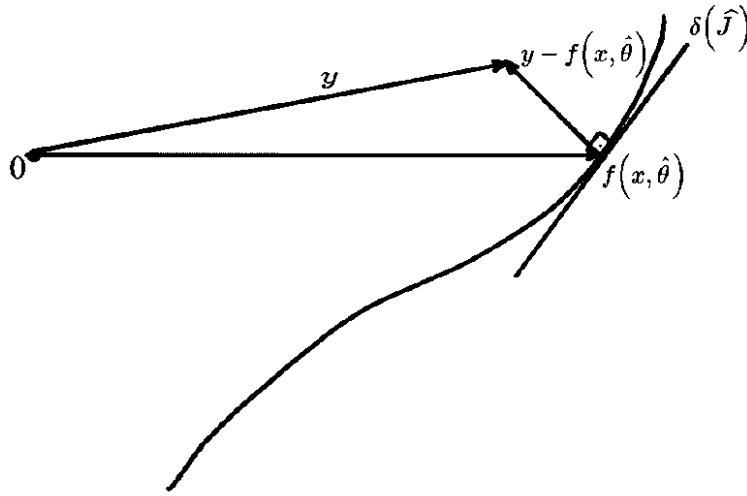
Şekil 2.4 Doğrusal en küçük kareler yönteminin geometrisi

Doğrusal olmayan regresyonda ise, en küçük kareler kestiricisinin geometrik gösterimi biraz farklıdır. $f(x, \theta)$ her yerde türevlenebilir doğrusal olmayan bir fonksiyon olsun. Bu durumda $f(x, \theta)$, k – boyutlu bir hiper yüzey (manifold) belirtir. Bu yüzeye herhangi sabit $\bar{\theta}$ noktasında oluşturulan $\bar{J} = J(x, \bar{\theta})$ matrisi modelin bu noktadaki tanjant uzayını (düzlemini) belirler. Bu uzayı $\delta(\bar{J})$ ile gösterelim. Şekil 2.5' de iki farklı $\bar{\theta}^1$ ve $\bar{\theta}^2$ noktalarında oluşturulmuş tanjant düzlemlerinin $k = 1$ için gösterimi verilmiştir.



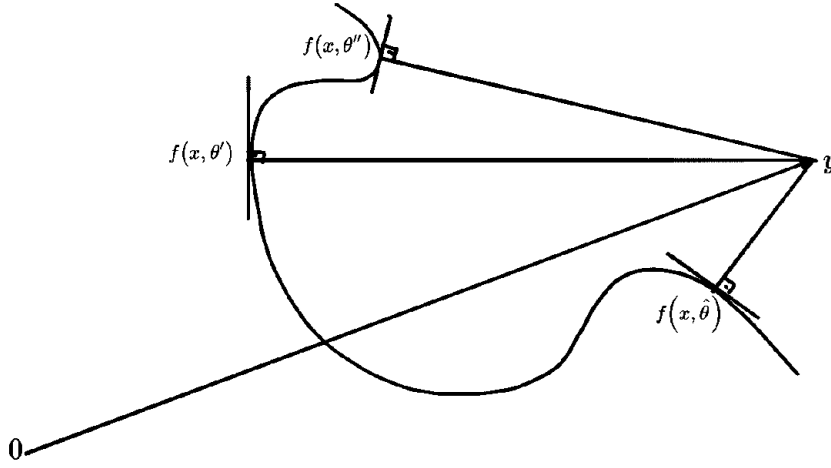
Şekil 2.5 Yüzeğe çizilen tanjant düzlemleri

$\bar{\theta}^1$ ve $\bar{\theta}^2$ normal denklemleri sağlayan iki nokta ve \bar{J}^1 ve \bar{J}^2 sırasıyla bu noktalarda hiperyüzeğe çizilen tanjant düzlemlerini belirtir. Oluşturulan tanjant düzlemleri J Jakobiyen matrisinin sütunları tarafından gerilen bir alt uzay olarak tanımlanabilir.



Şekil 2.6 y 'nin hiperdüzleme izdüşümü

Şekil 2.6' da, bağımlı değişkenin doğrusal olmayan yüzeğe, normal denklemlerin sağlandığı noktada çizilen tanjant düzlemine izdüşümü gösterilmiştir (Davidson & Mackinnon, 1993). Şekilden de görüleceği üzere $f(x, \theta)$ hiperyüzeyi orta derecede doğrusal olmayan bir yüzeydir. Dolayısıyla bu yüzey üzerinde global minimum bulmak zor değildir. Ancak Şekil 2.7 incelendiğinde, yüksek dereceli doğrusal olmayan bir hiperyüzey mevcut ise burada global minimum bulmak kolay değildir.



Şekil 2.7 Yüksek dereceli doğrusal olmayan hiperyüzey

Şekil 2.7' den de anlaşılacağı gibi, çalışılan yüzeylerin doğrusal olmama dereceleri parametre kestirimi açısından oldukça önemlidir. Bu modelde normal denklemleri sağlayan 3 farklı nokta vardır. Ancak önemli olan bu noktalardan global minimum olanı bulmaktır. Bu grafikten de geometrik olarak yine anlaşılıyor ki normal denklemleri sağlamak, global minimum değerini, yani aranan doğru kestiriciyi bulmak için gerekli fakat yeterli değildir. Söz konusu durumda parametre kestiriminde problemler ortaya çıkmaktadır.

2.4. MODELİN DOĞRU BELİRLENMESİ

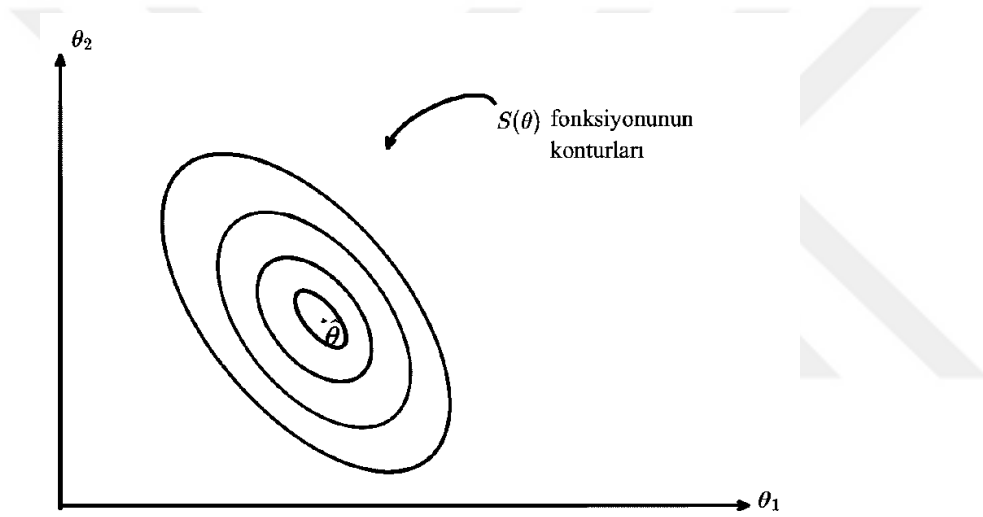
$S(\theta)$ amaç fonksiyonunu başarılı bir biçimde minimize etmek modelin doğru belirlenmesi için gerekli fakat yeterli değildir (Davidson & Mackinnon, 1993). Modelin doğru belirlenmesi kısmında oluşabilecek problemler vardır. Doğrusal olmayan regresyonda, modelin doğru belirlenebilmesi (tanımlanabilmesi) için, verilen bir veri kümesinde $S(\theta)$ amaç fonksiyonunu minimize eden sadece bir $\hat{\theta}$ değeri olması gerekir. Yani doğrusal olmayan model doğru tanımlanamıyorsa birden fazla $\hat{\theta}$ değeri var demektir (Davidson & Mackinnon, 1993).

Yerel ve mutlak olmak üzere iki tip doğru tanımlama vardır. En küçük kareler kestiricisi olan $\hat{\theta}$ değerinin ufak değişiminde $S(\theta)$ değeri artıyorsa, $\hat{\theta}$ yerel doğru tanımlı olacaktır. Bu durum, $S(\theta)$ fonksiyonunun $\hat{\theta}$ noktasında kesin dışbükey (konveks) olması gerektiği anlamına gelir. Bir başka ifade ile, bu durum, τ değerinin

ufak deęişimleri için $S(\hat{\theta}) < S(\hat{\theta} + \tau)$ eşitsizlięinin saęlanmasıdır. Kesin konvekslięin bir dięer garantisi ise Hesyen (Hessian) matrisi olarak tanımlanan,

$$H_{ij}(\theta) = \frac{\partial^2 S(\theta)}{\partial \theta_i \partial \theta_j}$$

ikinci türev matrisinin $\hat{\theta}$ noktasında pozitif tanımlı olmasıdır. Kesin konvekslik, $S(\theta)$ amaç fonksiyonunun $\hat{\theta}$ civarında düzlemsel bir yapıda deęil, eğrisel bir yapıda olması anlamına gelir. Bu durumu gösteren Şekil 2.8 aşıęıda resmedilmiştir:

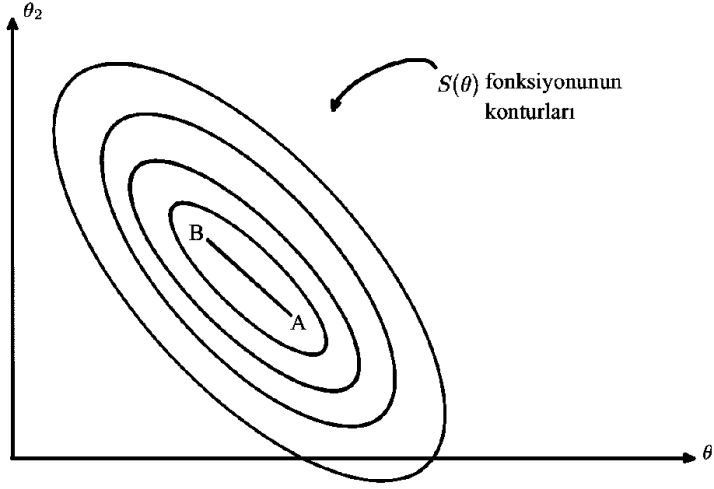


Şekil 2.8 $S(\theta)$ fonksiyonunun doğru tanımlanmış minimum deęeri

Yukarıdaki Şekil 2.8' de de görüldüęü gibi $S(\theta)$ fonksiyonunun konturlarını temsil eden eliptik şekiller küçölerek merkezde bir nokta elde edilmiştir. Bu nokta en küçük kareler kestirimini temsil etmektedir ve yerel olarak doğru tanımlıdır.

$S(\theta)$ fonksiyonunun kesin konveks olması her yöne doğru eğrilięinin olduęu yani düzlemsel bir bölgenin olmadıęı anlamına gelir. Eęer $S(\theta)$ fonksiyonu $\hat{\theta}$ civarında bir bölgede düzlemsel bir yapıya sahip ise yani düzlemsel bölgedeki her θ_i için $S(\hat{\theta}) = S(\theta_i)$ saęlanıyorsa bu durumda $S(\theta)$ fonksiyonunun deęiőeceęi bir yöne doğru ilerlemek gerekir. Dolayısıyla bu durum bize söz konusu bölgede $S(\theta)$

fonksiyonunu minimize eden sonsuz sayıda $\hat{\theta}$ kestirim değeri olduğunu gösterir ki bu da doğru tanımlanmamış (unidentified) en küçük kareler kestiricisi demektir.



Şekil 2.9 $S(\theta)$ fonksiyonunun doğru tanımlanmamış minimum değeri

Şekil 2.9' da, sözü edilen duruma örnek gösterilmiştir. AB doğru parçası boyunca tüm noktalar, $S(\theta)$ fonksiyonunu minimize eden noktalardır. Sadece bir $\hat{\theta}$ elde etmek için yerel doğru tanımlama gerekli fakat yeterli değildir. Bunun için mutlak doğru tanımlama gereklidir. Yani,

$$S(\hat{\theta}) < S(\theta^*) , \quad \forall \theta^* \neq \hat{\theta} \text{ için}$$

eşitsizliği sağlanmalıdır. Eğer bir model yerel doğru tanımlı ise, yüksek ihtimalle, iki ya da daha fazla parametre kestirim değeri eşit hata kareler ortalamasına sahip olacaktır. Bununla birlikte mutlak doğru tanımlamanın sağlandığı fakat yerel doğru tanımlamanın sağlanmadığı yani, ikinci türevler matrisi olan hesyen matrisinin pozitif tanımlı olmadığı durumlar söz konusu olabilir. Bu durum, aslında, doğrusal olmayan regresyon modellerinde parametre kestiriminde sıkça karşılaşılan bir problemdir. Bu duruma örnek olarak,

$$f(x_i, \theta) = \theta_1 + \theta_2 x_i^{\theta_3}$$

fonksiyonu verilebilir. Bu fonksiyonda dikkat edilecek olursa, $\hat{\theta}_2 = 0$ olduğu durumda, θ_3 parametresinin fonksiyon ve $S(\theta)$ üzerinde hiçbir etkisi yoktur. Buna benzer fonksiyonlar uygulamada sıkça karşımıza çıkmaktadır. Bu problem literatürde zayıf

tanımlı model olarak tanımlanmaktadır. Bu genel bir tabirdir. Değişik kaynaklarda bu problem çoklu iç ilişki (multicollinearity) olarak tanımlanır. Bu problemin varlığında Hesyen ya da Jakobiyen matrisi tekil olmayan ancak yaklaşık tekil olan matrislerdir. Dolayısıyla doğrusal olmayan regresyonda, parametre kestiriminde kullanılan J Jakobiyen matrisi ile elde edilen $J^T J$ matrisi yaklaşık tekil matristir ve bu matrisin tersinin alınması aşamasında sıkıntılar ortaya çıkmaktadır.

2.5. GAUSS – NEWTON YÖNTEMİ

Doğrusal olmayan en küçük kareler yönteminin Eşitlik (2.4)' de verilen normal denklemlerin çözümünde yetersiz kalmasından dolayı, Gauss – Newton yöntemi olarak bilinen doğrusallaştırma (linearization) yöntemi öne sürülmüştür. Amaç, doğrusal olmayan fonksiyonun Taylor serisi yardımıyla doğrusallaştırılması ve doğrusal regresyon analizinde olduğu gibi en küçük kareler yönteminin uygulanmasıdır. Bu süreç, iteratif olarak devam eder ve yakınsama sağlandığında kestirim elde edilmiş olur. Bu amaç doğrultusunda, $\theta_0 = (\theta_{10}, \theta_{20}, \dots, \theta_{p0})^T$ başlangıç noktası olmak üzere, $f(\mathbf{x}_i, \theta)$ fonksiyonunun θ_0 civarındaki Taylor serisine açılımı yaklaşık olarak,

$$f(\mathbf{x}_i, \theta) = f(\mathbf{x}_i, \theta_0) + \sum_{j=1}^p \left[\frac{\partial f(\mathbf{x}_i, \theta)}{\partial \theta_j} \right]_{\theta=\theta_0} (\theta_j - \theta_{j0}) \quad (2.9)$$

şeklindedir. Burada,

$$\begin{aligned} f_i^0 &= f(\mathbf{x}_i, \theta_0) \\ \beta_j^0 &= \theta_j - \theta_{j0} \\ J_{ij}^0 &= \left[\frac{\partial f(\mathbf{x}_i, \theta)}{\partial \theta_j} \right]_{\theta=\theta_0} \end{aligned} \quad (2.10)$$

kısaltmaları ile (2.9) eşitliği,

$$f(\mathbf{x}_i, \theta) = f_i^0 + \sum_{j=1}^p \beta_j^0 J_{ij}^0 \quad (2.11)$$

biçimini alır. Buradan da doğrusal olmayan (2.1) modelini aşağıdaki gibi yazabiliriz:

$$y_i - f_i^0 = \sum_{j=1}^p \beta_j^0 J_{ij}^0 + \varepsilon_i \quad i = 1, 2, \dots, n \quad (2.12)$$

Yukarıda elde ettiğimiz (2.12) doğrusallaştırılmış modelinin, doğrusal regresyon modelinden bir farkı yoktur. Öyle ki,

$$\mathbf{J}_0 = \begin{pmatrix} J_{11}^0 & \cdots & J_{p1}^0 \\ \vdots & \ddots & \vdots \\ J_{1n}^0 & \cdots & J_{pn}^0 \end{pmatrix}_{n \times p}$$

yeni tasarım matrisi, doğrusal regresyondaki \mathbf{X} tasarım matrisinin görevini görmektedir.

$$\hat{\boldsymbol{\beta}}_0 = \begin{pmatrix} \hat{\beta}_1^0 \\ \hat{\beta}_2^0 \\ \vdots \\ \hat{\beta}_p^0 \end{pmatrix}$$

kestirim vektörü olmak üzere,

$$\mathbf{y}_0 = \mathbf{y} - \mathbf{f}^0 = \begin{pmatrix} y_1 - f_1^0 \\ y_2 - f_2^0 \\ \vdots \\ y_n - f_n^0 \end{pmatrix}$$

yanıt vektörü, yine doğrusal modeldeki \mathbf{y} yanıt vektörünün görevini üstlenmektedir.

$\mathbf{y}_0 = y_i - f_i^0$ kısaltması ile (2.12) modelini, matris formunda,

$$\mathbf{y}_0 = \mathbf{J}_0 \boldsymbol{\beta}_0 + \boldsymbol{\varepsilon} \quad (2.13)$$

şeklinde yazacak olursak, $\boldsymbol{\beta}_0$ parametresinin kestirimi en küçük kareler yöntemi kullanılarak,

$$\hat{\boldsymbol{\beta}}_0 = (\mathbf{J}_0^T \mathbf{J}_0)^{-1} \mathbf{J}_0^T \mathbf{y}_0 = (\mathbf{J}_0^T \mathbf{J}_0)^{-1} \mathbf{J}_0^T (\mathbf{y} - \mathbf{f}_0) \quad (2.14)$$

olarak bulunur. Bu değer,

$$S(\boldsymbol{\theta}) = \sum_{i=1}^n \left[y_i - f_i^0 - \sum_{j=1}^p \beta_j^0 J_{ij}^0 \right]^2 \quad (2.15)$$

amaç fonksiyonunu minimize eden değerdir. Bir başka deyişle, (2.15) amaç fonksiyonun $\beta_j^0 = \theta_j - \theta_{j0}$ parametrelerine göre kısmi türevlerinin sıfıra eşitlenmesi ile elde edilen normal denklemlerin çözümü (2.14) ile verilen parametre kestirim değerlerini verecektir. Yukarıda (2.14) ile verilen ifadeye $\beta_0 = \boldsymbol{\theta} - \boldsymbol{\theta}_0$ dönüşümü yapıldığından, iteratif olarak elde edilen ilk parametre kestirimi,

$$\hat{\boldsymbol{\theta}}_1 = \boldsymbol{\theta}_0 + \hat{\boldsymbol{\beta}}_0 = \boldsymbol{\theta}_0 + (\mathbf{J}_0^T \mathbf{J}_0)^{-1} \mathbf{J}_0^T (\mathbf{y} - \mathbf{f}_0) \quad (2.16)$$

şeklinde bulunmuş olur. Bu iteratif kestirim süreci yakınsaklık sağlanıncaya kadar devam eder. Yakınsaklık sağlandığında elde edilen son kestirim, genel olarak,

$$\hat{\boldsymbol{\theta}}_{k+1} = \hat{\boldsymbol{\theta}}_k + \hat{\boldsymbol{\beta}}_k = \hat{\boldsymbol{\theta}}_k + (\mathbf{J}_k^T \mathbf{J}_k)^{-1} \mathbf{J}_k^T (\mathbf{y} - \mathbf{f}_k) \quad (2.17)$$

şeklinde elde edilir. Yakınsaklık kriteri için, iteratif olarak parametrelerin kararlı olduğu ya da amaç fonksiyonunun kararlı olduğu kriterler kullanılabilir. Örneğin,

$$\frac{\hat{\theta}_{j,k+1} - \hat{\theta}_{j,k}}{\hat{\theta}_{j,k}} < \tau \quad , \quad j = 1, 2, \dots, p \quad (2.18)$$

yakınsaması sağlanıncaya kadar, iteratif süreç devam edebilir. Burada τ , 10^{-6} gibi, sıfıra çok yakın pozitif bir reel sayı olarak seçilebilir. Duyarlılık arttırılmak istenirse, τ daha da küçük seçilebilir. Ancak τ değerinin çok küçük seçilmesi yakınsamanın hızını büyük ölçüde etkileyecektir. Ayrıca birikimli hesaplama yanlışları artacaktır. Bununla birlikte, iterasyon sürerken, her adımda $S(\hat{\boldsymbol{\theta}}_k)$ artık (residual) kareler toplamında bir azalma sağlanıp sağlanmadığı kontrol edilmelidir (Montgomery, et al., 2012).

2.6. LEVENBERG – MARQUARDT YÖNTEMİ

1944 yılında Levenberg, ardından 1963 yılında Marquardt, Levenberg – Marquardt yöntemi olarak bilinen bu yöntemi, Gauss – Newton yöntemine alternatif olarak önermişlerdir. Gauss – Newton ile yaklaşım olarak benzerlik gösteren bu yöntem,

$$S(\boldsymbol{\theta}) = \sum_{i=1}^n \varepsilon_i^2 = \sum_{i=1}^n [y_i - f(\mathbf{x}_i, \boldsymbol{\theta})]^2 = \|\mathbf{y} - f(\mathbf{x}, \boldsymbol{\theta})\|^2$$

amaç fonksiyonunu, $\boldsymbol{\theta}^*$ parametre uzayında bir nokta olmak üzere,

$$\|\boldsymbol{\theta}\|^2 = \|\boldsymbol{\theta}^*\|^2 \quad (2.19)$$

kısıtı altında, minimize etmeyi amaçlar. Bu problemin çözümü Lagrange fonksiyonu yardımıyla yapılır. Bu amaçla,

$$L(\boldsymbol{\theta}, \phi) = \|\mathbf{y} - f(\mathbf{x}, \boldsymbol{\theta})\|^2 + \phi(\|\boldsymbol{\theta}\|^2 - \|\boldsymbol{\theta}^*\|^2) \quad (2.20)$$

Lagrange fonksiyonu minimize edilir. Yöntem Gauss – Newton yönteminde olduğu gibi Taylor serisi yardımıyla doğrusal olmayan $f(\mathbf{x}, \boldsymbol{\theta})$ fonksiyonunun doğrusallaştırılması ile başlar ve kısıtlı en küçük kareler yöntemi uygulanır. Burada, $\phi > 0$ Lagrange çarpanı olarak tanımlanır. (2.20) Lagrange fonksiyonunun minimize edilmesiyle iteratif olarak elde edilen parametre kestiricisi,

$$\hat{\boldsymbol{\theta}}_{k+1} = \hat{\boldsymbol{\theta}}_k + \hat{\boldsymbol{\beta}}_k = \hat{\boldsymbol{\theta}}_k + (\mathbf{J}_k^T \mathbf{J}_k + \phi_k \mathbf{I})^{-1} \mathbf{J}_k^T (\mathbf{y} - \mathbf{f}_k) \quad (2.21)$$

olarak elde edilir. ϕ_k değerinin farklı hesaplanma teknikleri vardır. Genellikle, başlangıç olarak $\phi = 10^{-8}$ olarak başlar. Bir dizi deneme yanılma hesapları, her bir iterasyon için,

$$S(\hat{\boldsymbol{\theta}}_{k+1}) < S(\hat{\boldsymbol{\theta}}_k) \quad (2.22)$$

sağlanıncaya kadar, ϕ ' nin 10 ile çarpılmasıyla uygulanır. Süreç, aynı zamanda, (2.22) eşitsizliğinin sağlanması koşulu ile her bir iterasyonda ϕ ' yi 10 kat küçültür. Strateji, her bir iterasyondaki artık kareler toplamının azaldığını garanti ederken, ϕ ' yi olabildiğince küçük tutmaktır. Bu genel sürece Marquardt'ın yaklaşımı denir (Montgomery, et al., 2012).

3. DOĞRUSAL OLMAYAN REGRESYON MODELLERİNDE ÇOKLU İÇ İLİŞKİ PROBLEMİ ve YENİ BİR YAKLAŞIM

Doğrusal regresyon modellerinde çoklu iç ilişki problemi her yönüyle incelenmiş ve farklı çözüm yolları geliştirilmiştir. Fakat doğrusal olmayan regresyon modellerinde bu problem, modelin yapısı gereği karşılaşılan zorluklardan dolayı yeteri kadar incelenmemiştir. Var olan sınırlı sayıda çalışmalar da, doğrusal modellerde olduğu kadar ayrıntılı incelenmemiştir. Bu bölümde, doğrusal regresyon analizinde çoklu iç ilişkinin teşhisi ve etkilerini doğrusal olmayan regresyon analizine uyarlayacağız. Ardından bu probleme çözüm olarak geliştirdiğimiz yeni yaklaşımı sunacağız.

3.1. ÇOKLU İÇ İLİŞKİ TEŞHİSİ

Doğrusal regresyon modellerinde çoklu iç ilişki teşhisi için yapılan çalışmalarda çoğunlukla X tasarım matrisi incelenmiştir. Belsley (2004) çoklu iç ilişki teşhisini doğrusal olmayan regresyon modellerine genişletilebileceğini ve Gauss – Newton algoritmasında kullanılan kısmi türevler matrisi olan J Jakobiyen matrisinin sütunları arasındaki doğrusal ilişkinin teşhis edilebileceğini belirtmiştir. Magel ve Hertsgaard (1987) çalışmalarında doğrusal olmayan regresyondaki J Jakobiyen matrisinin doğrusal modellerdeki X tasarım matrisinin rolünü üstlendiğini söylemişlerdir. Bir başka çalışmada, doğrusal olmayan regresyonda iç ilişki teşhisi için Jakobiyen matrisinin incelenebileceği belirtilmiştir (Hill & Adkins, 2003).

Doğrusal olmayan regresyonda parametre kestirim süreci iteratif bir süreç olduğundan, yakınsaklık sağlanıncaya kadar, tüm iterasyonlarda elde edilen J Jakobiyen matrisi çoklu iç ilişkinin teşhisi için incelenmek zorundadır.

$y = f(X, \theta) + \varepsilon$ doğrusal olmayan modelini Gauss – Newton yöntemi ile iteratif olarak doğrusallaştırdığımızı ve herhangi bir iterasyonda $y^* = J\beta + \varepsilon$ doğrusal modeline sahip olduğumuzu düşünelim. $n \times p$ boyutlu J “yeni tasarım matrisi” ni,

$$J = H\Lambda^{1/2}G^T \quad (3.1)$$

olacak şekilde yazmak her zaman mümkündür (Vinod & Ullah, 1981). Burada \mathbf{H} $n \times p$ boyutlu, $\mathbf{H}^T \mathbf{H} = \mathbf{I}$ koşulunu sağlayan özel bir matristir. Bu matrisin sütunları dik (ortogonal) olup \mathbf{J} matrisinin sütun uzayı için bir dik taban oluşturur (Erkoç, 2011). $\mathbf{\Lambda}^{1/2}$ $p \times p$ boyutlu ve \mathbf{J} matrisinin sıralı $\lambda_i^{1/2}$ tekil değerlerini içeren ($\lambda_1^{1/2} \geq \lambda_2^{1/2} \geq \dots \geq \lambda_p^{1/2}$) bir köşegen matristir. $p \times p$ boyutlu \mathbf{G} matrisinin sütunları (\mathbf{g}_i) dik olup, \mathbf{J} matrisinin satır uzayı için dik bir taban oluşturur. $\mathbf{J}^T \mathbf{J} = (\mathbf{H} \mathbf{\Lambda}^{1/2} \mathbf{G}^T)^T (\mathbf{H} \mathbf{\Lambda}^{1/2} \mathbf{G}^T) = \mathbf{G} \mathbf{\Lambda}^{1/2} \mathbf{H}^T \mathbf{H} \mathbf{\Lambda}^{1/2} \mathbf{G}^T = \mathbf{G} \mathbf{\Lambda} \mathbf{G}^T$ matrisinde $\mathbf{\Lambda}$ matrisinin λ_i elemanları \mathbf{J} matrisinin tekil değerlerinin kareleri olup, $\mathbf{J}^T \mathbf{J}$ matrisinin özdeğerleridir.

Yukarıda yapılan ayrışım tekil değer ayrışımı (singular value decomposition, SVD) olarak bilinir. $\mathbf{y}^* = \mathbf{J} \boldsymbol{\beta} + \boldsymbol{\varepsilon}$ modelini, daha sonra çoklu iç ilişki teşhisi yapmak için, \mathbf{J} matrisinin tekil değer ayrışımından yararlanarak,

$$\begin{aligned} \mathbf{y}^* &= \mathbf{J} \boldsymbol{\beta} + \boldsymbol{\varepsilon} = \mathbf{H} \mathbf{\Lambda}^{1/2} \mathbf{G}^T \boldsymbol{\beta} + \boldsymbol{\varepsilon} \\ &= \mathbf{J}^* \boldsymbol{\gamma} + \boldsymbol{\varepsilon} \end{aligned} \quad (3.2)$$

şeklinde yazabiliriz. Burada, $\mathbf{J}^* = \mathbf{H} \mathbf{\Lambda}^{1/2}$ ve $\boldsymbol{\gamma} = \mathbf{G}^T \boldsymbol{\beta}$ kısaltmaları yapılmıştır.

Regresyon analizinde temel varsayımların dışında açıklayıcı değişkenleri temsil eden vektörlerin ya da daha genel olarak, parametre kestirim uzayında y_i gözlemlerinin dik izdüşümlerinin yapıldığı uzayı geren sütun vektörlerinin dik olması varsayımı da önemlidir. Bu varsayım özdeş olarak \mathbf{J} matrisinin sütunlarının dik olması yani kısaca \mathbf{J} matrisinin dik olması anlamına gelmektedir. Eğer $\mathbf{J}^T \mathbf{J} = \mathbf{I}$ ise \mathbf{J} matrisine dik matris denir. $\mathbf{J}^T \mathbf{J}$ matrisinin tüm özdeğerleri 1 ise yani $\mathbf{\Lambda} = \mathbf{I}$ ise \mathbf{J} dik sütunlardan oluşmuştur. Aksine, $\mathbf{\Lambda} \neq \mathbf{I}$ olduğu durumda \mathbf{J} matrisi dik bir matrisi değildir (nonorthogonality).

Çoklu iç ilişki \mathbf{J} matrisinin sütunları arasındaki doğrusal bağımlılık anlamına gelmektedir. Burada karıştırılmaya müsait iki farklı durum söz konusudur: Tam çoklu iç ilişki ve yakın (yaklaşık) çoklu iç ilişki.

Tam çoklu iç ilişki \mathbf{J} matrisinin rankının p ' den küçük olduğu yani rankının düştüğü durumdur. $\mathbf{j}_1, \mathbf{j}_2, \dots, \mathbf{j}_p$, \mathbf{J} matrisinin sütun vektörleri; a_i ($i = 1, 2, \dots, p$) hepsi birden sıfır olmayan sabitler olmak üzere,

$$\sum_{i=1}^p a_i \mathbf{j}_i = 0 \quad (3.3)$$

eşitliği sağlanıyorsa, \mathbf{J} matrisinin sütun vektörlerinin herhangi bir alt kümesi birbirleri ile doğrusal bağımlıdır denir. Bu durumda \mathbf{J} matrisinin rankı düşmüş olur.

Yakın çoklu iç ilişki ise, yukarıdaki (3.3) ile verilen denklemdeki eşitlik yerine yaklaşık eşitliği gösteren \approx sembolü ile,

$$\sum_{i=1}^p a_i \mathbf{j}_i \approx 0 \quad (3.4)$$

sağlanıyorsa geçerlidir.

$\mathbf{J}^T \mathbf{J} = \mathbf{G} \mathbf{\Lambda} \mathbf{G}^T$ matrisi tekil olduğunda tersi alınmaz. Tam çoklu iç ilişki $\mathbf{J}^T \mathbf{J}$ matrisinin tekil olmasına yol açar. Bu durumda, $\mathbf{J}^T \mathbf{J}$ matrisinin en küçük öz değeri $\lambda_p = 0$ dır.

Yaklaşık çoklu iç ilişki ise $\mathbf{J}^T \mathbf{J}$ matrisinin yaklaşık tekil olmasına neden olur. Bu durumda ise en küçük özdeğer sıfıra çok yakındır. Özdeğer analizi çoklu iç ilişkinin teşhisi için sıkça kullanılan, yararlı bir tekniktir. Özdeğerler 1 değerinden farklı ise bu durumda diklik yani değişkenler (sütun vektörleri) arasındaki tam bağımsızlık yoktur (nonorthogonality). Bu durumu şu şekilde özetleyebiliriz:

1 değerinden farklı en küçük özdeğer sıfır ya da sıfıra yakın demek değildir. Yani dikliğin sağlanmadığı her durum tekilliğe ya da çoklu iç ilişkiye işaret değildir. Ancak, çoklu iç ilişki dikliğin sağlanmadığı durumu yaratır (Vinod & Ullah, 1981).

3.2. KÖTÜ KOŞULLULUK

Bir matrisin kötü koşulluluğunu incelemek ve tespit etmek için literatürde “koşul sayısı” olarak adlandırılan,

$$KS = \lambda_1^{1/2} \lambda_p^{-1/2} \quad (3.5)$$

ifadesi kullanılır. Burada $\lambda_1^{1/2}$ ve $\lambda_p^{1/2}$ sırasıyla \mathbf{J} matrisinin en büyük ve en küçük tekil değerleridir. Kısaca koşul sayısı, \mathbf{J} matrisinin tekil değer oranı olarak da tanımlanabilir. Burada “tekil değer” ifadesi ile bir matrisin tekilliği karıştırılmamalıdır. \mathbf{J} matrisinin tekil değerlerinin kareleri $\mathbf{J}^T \mathbf{J}$ matrisinin özdeğerlerine karşılık geldiğinden, farklı kaynaklarda koşul sayısı,

$$KS = \frac{\lambda_{\max}}{\lambda_{\min}} \quad (3.6)$$

olarak görmek mümkündür.

$\mathbf{J}^T \mathbf{J}$ matrisinin tekil matris olması $\lambda_p = 0$ olması anlamına gelir. Bu durumda tam çoklu iç ilişki vardır ve koşul sayısı $KS \rightarrow \infty$ olur. Tam tersi durumda, yani, dik $\mathbf{J}^T \mathbf{J}$ matrisinde $\mathbf{J}^T \mathbf{J} = \mathbf{I}$ sağlandığından ve $\lambda_1 = \lambda_2 = \dots = \lambda_p = 1$ tüm özdeğerler 1’e eşit olduğundan $KS = 1$ olur.

Koşul sayısı $[1, \infty)$ yarı açık aralığında değer alabilir. Koşul sayısı, KS , 1’den uzaklaştıkça kötü koşulluluğun arttığı söylenir.

Kötü koşulluluğun derecesinin tespiti için farklı kaynaklarda ölçütler verilmiştir. Genel olarak kabul gören kritik değer 30’ dur. Koşul sayısı 30 değerinden küçük ise, kötü koşulluluğun derecesi nispeten daha azdır. Burada önemli olan nokta bir matrisin kötü koşulluluğunun yapılan analizi ne ölçüde etkilediğidir. Bu anlamda, koşul sayısı 30 değerinden küçük ise yapılan analizlerin (nispeten) daha güvenli olduğu söylenebilir.

“Kötü koşulluluk” terimi bazı kaynaklarda çoklu iç ilişki ya da dik olmama durumu ile eş anlamlı olarak kullanılabilir. Aslında birbirlerinin nedeni olan bu kavramlar, anlam kargaşası yaratmamak için aynı anlamı taşıyormuş gibi kullanılabilir (Vinod & Ullah, 1981).

3.3. ÇOKLU İÇ İLİŞKİNİN ETKİLERİ

Çoklu iç ilişkinin parametre kestirimi üzerindeki etkisini incelemek için doğrusal olmayan modelde iki parametre olduğunu düşünelim. Böylelikle \mathbf{J} Jakobiyen matrisinde doğrusallaştırma sonucunda j_1 ve j_2 ile gösterilen, açıklayıcı değişkenleri temsil edecek iki sütun olacaktır. r_{12} bu açıklayıcı değişkenler arasındaki korelasyon katsayısı, r_{yi} j_i ile \mathbf{y} arasındaki korelasyon katsayısı olmak üzere, $\mathbf{y}^* = \mathbf{J}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ modeli için, en küçük kareler kestiricisinin,

$$(\mathbf{J}^T \mathbf{J})^{-1} = \begin{bmatrix} 1 & -r_{12} \\ -r_{12} & 1 \end{bmatrix} \cdot [\det(\mathbf{J}^T \mathbf{J})]^{-1} \quad (3.7)$$

matrisini kullanarak,

$$\hat{\boldsymbol{\beta}} = (\mathbf{J}^T \mathbf{J})^{-1} \mathbf{J}^T \mathbf{y}^* \quad (3.8)$$

şeklinde elde edildiğini önceki bölümlerde göstermiştik. Burada $\mathbf{y}^* = \mathbf{y} - \mathbf{f}$ olduğu önceki bölümlerde Gauss – Newton yöntemi anlatırken gösterilmişti. (3.7) ile verilen eşitlikte, $\mathbf{J}^T \mathbf{J}$ matrisinin determinant değeri, $\det(\mathbf{J}^T \mathbf{J}) = 1 - r_{12}^2$ şeklinde hesaplanır. $\hat{\boldsymbol{\beta}}$ nın bir başka gösterilişi,

$$\hat{\boldsymbol{\beta}} = \begin{bmatrix} r_{y1} - r_{12}r_{y2} \\ r_{y2} - r_{12}r_{y1} \end{bmatrix} [\det(\mathbf{J}^T \mathbf{J})]^{-1} \text{ şeklindedir. Eğer } r_{12} = 1 \text{ ise } \det(\mathbf{J}^T \mathbf{J}) = 0 \text{ olacaktır. Bu}$$

durum, modelde tam çoklu iç ilişki olması anlamına gelir ve $\hat{\boldsymbol{\beta}}$ kestirilemez. Bununla birlikte,

$$\begin{aligned} \hat{\beta}_1 + \hat{\beta}_2 &= [r_{y1} - r_{12}r_{y2}] [\det(\mathbf{J}^T \mathbf{J})]^{-1} + [r_{y2} - r_{12}r_{y1}] [\det(\mathbf{J}^T \mathbf{J})]^{-1} \\ &= [r_{y1} - r_{12}r_{y2} + r_{y2} - r_{12}r_{y1}] [\det(\mathbf{J}^T \mathbf{J})]^{-1} \\ &= [r_{y1}(1 - r_{12}) + r_{y2}(1 - r_{12})] [\det(\mathbf{J}^T \mathbf{J})]^{-1} \\ &= [(r_{y1} + r_{y2})(1 - r_{12})] [\det(\mathbf{J}^T \mathbf{J})]^{-1} \\ &= \frac{[(r_{y1} + r_{y2})(1 - r_{12})]}{1 - r_{12}^2} \\ &= \frac{(r_{y1} + r_{y2})}{1 + r_{12}} \end{aligned} \quad (3.9)$$

parametre kestirimlerinin toplamları yukarıdaki gibi elde edilebilir ve $r_{12} = 1$ olması durumunda toplam belirlenebilir. Ancak,

$$\begin{aligned}
\hat{\beta}_2 - \hat{\beta}_1 &= [r_{y_2} - r_{12}r_{y_1}] [\det(\mathbf{J}^T \mathbf{J})]^{-1} - [r_{y_1} - r_{12}r_{y_2}] [\det(\mathbf{J}^T \mathbf{J})]^{-1} \\
&= [r_{y_2} - r_{12}r_{y_1} - r_{y_1} + r_{12}r_{y_2}] [\det(\mathbf{J}^T \mathbf{J})]^{-1} \\
&= [r_{y_2}(1 + r_{12}) - r_{y_1}(1 + r_{12})] [\det(\mathbf{J}^T \mathbf{J})]^{-1} \\
&= [(r_{y_2} - r_{y_1})(1 + r_{12})] [\det(\mathbf{J}^T \mathbf{J})]^{-1} \\
&= \frac{[(r_{y_2} - r_{y_1})(1 + r_{12})]}{1 - r_{12}^2} \\
&= \frac{(r_{y_2} - r_{y_1})}{1 - r_{12}}
\end{aligned} \tag{3.10}$$

parametre kestirim farkları yukarıdaki (3.10) eşitliği ile elde edilir ve $r_{12} = 1$ olması durumunda fark belirlenemez. Bir başka ifade ile, modelde tam çoklu iç ilişki var ise yani $r_{12} = 1$ ise, parametre kestirimleri $\hat{\beta}_1$ ve $\hat{\beta}_2$ belirlenemez, $\mp(\hat{\beta}_1 - \hat{\beta}_2)$ farkları belirlenemez, ancak, $\mp(\hat{\beta}_1 + \hat{\beta}_2)$ toplamları belirlenir. Bu durum $p = 2$ için incelendiğinden şu şekilde de özetlenebilir:

β parametre vektörünün doğrusal bir kombinasyonu $\mathbf{U}^T \beta$ olsun. Bu durumda $\mathbf{U}^T \beta$, $\mathbf{U}^T = (1,1)$ olması durumunda kestirilebilir. Ancak $\mathbf{U}^T = (-1,1)$ olması durumunda kestirilemez.

Çoklu iç ilişki parametre kestirimlerinin varyansları üzerinde etkilidir. Gauss – Newton yöntemi ile elde edilen en küçük kareler kestiricisi $\hat{\beta}$ asimptotik olarak,

$$\hat{\beta} \sim N(\beta, \sigma^2 (\mathbf{J}^T \mathbf{J})^{-1}) \tag{3.11}$$

dağılımına sahiptir (Hill & Adkins, 2003). Buradan hareketle, varyans – kovaryans matrisi $Cov(\hat{\beta}) = \sigma^2 (\mathbf{J}^T \mathbf{J})^{-1}$ nin köşegen öğeleri toplamı,

$$HKO(\hat{\beta}) = \sigma^2 \mathbf{1}_z (\mathbf{J}^T \mathbf{J})^{-1} \tag{3.12}$$

hata kareler ortalamasını verir. Bu ifade, $\mathbf{J}^T \mathbf{J}$ matrisinin özdeğerleri cinsinden,

$$HKO(\hat{\beta}) = \sigma^2 \sum_{j=1}^p \lambda_j^{-1} \quad (3.13)$$

eşitliği ile ifade edilebilir.

Yaklaşık çoklu iç ilişki olması durumunda, $\lambda_p \rightarrow 0$ sağlanacak ve $HKO(\hat{\beta}) \rightarrow \infty$ olacaktır. Bu durumda çoklu iç ilişki, parametre kestiricisinin varyansını şişirecek, bu varyansı olduğundan çok fazla gösterecektir.

Standartlaştırılmış veriler için $J^T J$ matrisi korelasyon formundadır. $(J^T J)^{-1}$ matrisinin (ij) . elemanı r^{ij} olarak gösterilsin. Doğrusal modellerde, ilk olarak, Farrar ve Glauber (1967) çoklu iç ilişkinin teşhisi için r^{ii} değerini incelemeyi önermişlerdir. Marquardt (1970) “Varyans Şişirici Faktörler” (Variance Inflation Factors) olarak tanımlanan, VIF değerlerinin,

$$VIF(i) = r^{ii} > 5 \quad (3.14)$$

olması durumunda, güçlü çoklu iç ilişkinin olduğunu işaret etmiştir (Vinod & Ullah, 1981). Theil (1971) r^{ii} değerini,

$$r^{ii} = \frac{1}{(1 - R_i^2) \|j_i\|^2} \quad (3.15)$$

ifadesi ile tanımlamıştır. Burada, $\|j_i\|^2 = j_i^T j_i$ ve R_i^2 değeri j_i ile geriye kalan $p-1$ açıklayıcı değişken ile doğrusal regresyonundan elde edilen, çoklu belirtme katsayısıdır. Modelde çoklu iç ilişki var ise, açıklayıcı değişkenler arasında doğrusal ilişki mevcut demektir. Bu ise, R_i^2 değerinin 1 değerine yaklaşacağı anlamına gelir. Dolayısıyla r^{ii} ifadesinin paydası 0 (sıfır) değerine yaklaşır ve r^{ii} büyük değerler alır. Bir başka deyişle, çoklu iç ilişki (3.15) ile verilen eşitliğe bakarak da parametre kestirim değerlerinin varyanslarının çok büyük olmasına neden olur. Ayrıca parametre varyanslarının gereğinden büyük olması, parametrelerin modele olan anlamlı katkılarını ölçen,

$$H_0 : \beta_j = 0 \quad j = 1, 2, \dots, p \quad (3.16)$$

hipotezinin büyük olasılıkla kabul edilmesine neden olur (Vinod & Ullah, 1981). Yani çoklu iç ilişkinin neden olduğu yüksek varyans, açıklayıcı değişkenlerin modele anlamlı katkılarının olup olmadığı konusunda araştırmacıyı yanıltabilir.

3.4. X MATRİSİNDEKİ İÇ İLİŞKİNİN J JAKOBİYEN MATRİSİNDEKİ İÇ İLİŞKİ İLE BAĞLANTISI

Doğrusal olmayan modellerde modelin yapısı gereği iç ilişki problemi ortaya çıkmaktadır. Doğrusal modellerde olduğu gibi, sadece açıklayıcı değişkenlere bakılarak iç ilişki teşhisi yapılamaz. Parametre kestirim sürecinde Jakobiyeen matrisi ile çalışıldığından, iç ilişki teşhisi yapılması gereken yer burasıdır. Ancak J

Jakobiyeen matrisi $J(\mathbf{x}, \boldsymbol{\theta}) = \frac{\partial f(\mathbf{x}, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}}$ şeklinde hesaplandığından, bu matriste iç

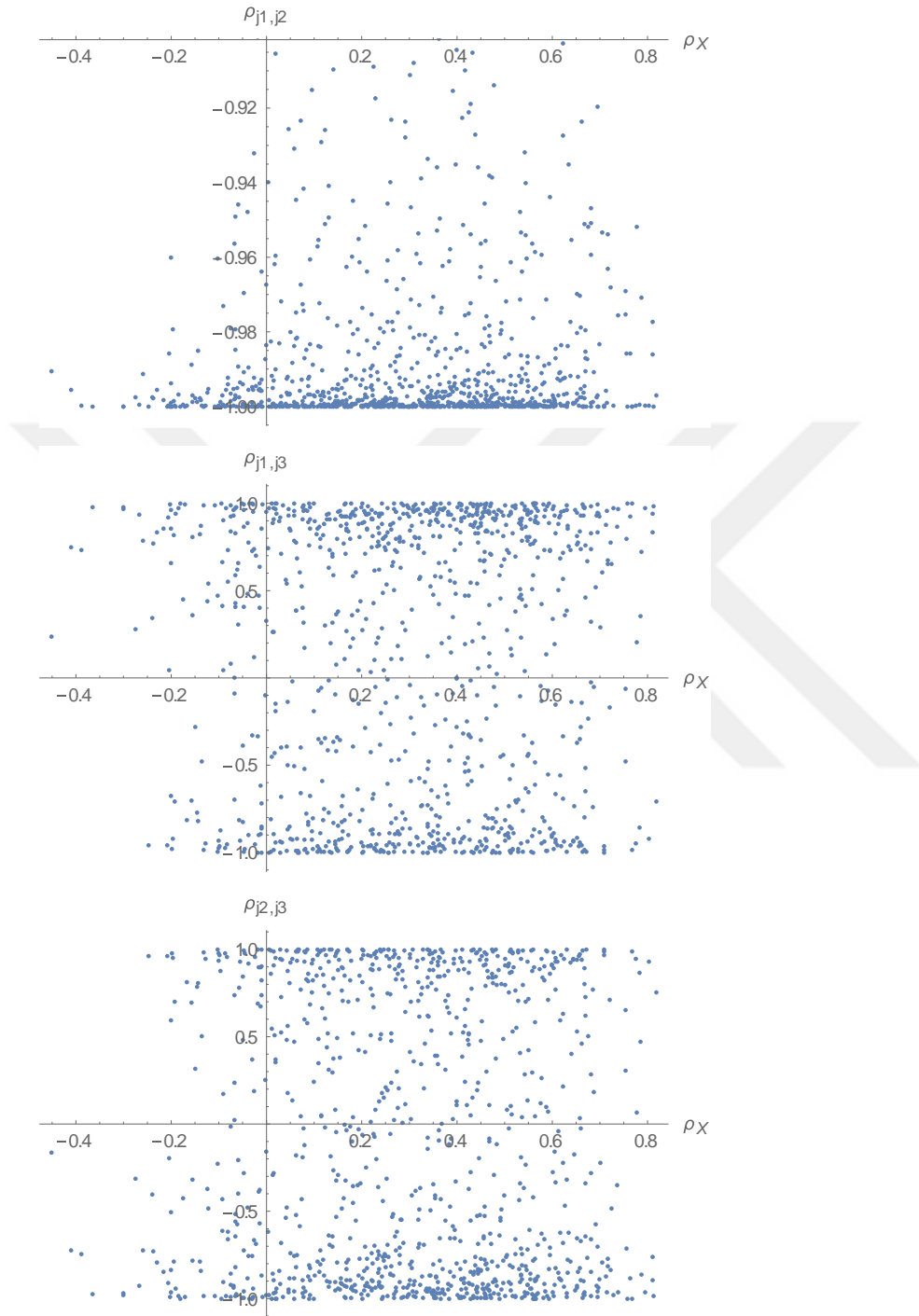
ilişkinin kaynaklarını bulmak zordur. Bunun nedeni olarak, Jakobiyeen matrisinin sütunlarının türevlerden oluştuğu söylenebilir. Ayrıca, J Jakobiyeen matrisi sadece X matrisinden değil, aynı zamanda parametre kestirim değerlerinden de oluştuğu için iç ilişkinin yapısını belirlemek analitik olarak imkansız hale gelir. Bu durumda, doğrusal olmayan modelin yapısının da büyük rolü vardır.

X matrisindeki iç ilişkinin $J(\mathbf{x}, \boldsymbol{\theta})$ matrisindeki karşılığı analitik olarak hesaplanamadığı için, burada, bilgisayar programı yardımıyla oluşturulan bir simülasyon çalışması ile bu probleme çözüm üretilmiştir. Amaç X matrisinin sütunları arasındaki ikili doğrusal ilişki olan Pearson korelasyon katsayısı ile, $J(\mathbf{x}, \boldsymbol{\theta})$ matrisinin sütunları arasındaki Pearson korelasyon katsayısının ilişkisini tespit etmektir. Bu amaçla, örnek model olarak, Meyer ve Roth (1972)' un kullandıkları

$y = \frac{\theta_1 \theta_3 x_1}{1 + \theta_1 x_1 + \theta_2 x_2}$ modeli seçilmiştir. Rasgele olarak atanan ρ korelasyon katsayısı

ile üretilen X matrisinden, bir başlangıç parametresi $\boldsymbol{\theta}_0$ ile J Jakobiyeen matrisi oluşturulmuştur. Farklı n gözlem sayısı ile oluşturulan grafikler ve yorumları aşağıda verilmiştir. Aşağıda Şekil 3.1' de gözlem sayısı olarak $n = 30$ seçilmiş ve 1000 tekrar ile buna bağlı olarak J matrisinin iç ilişkisinin bağıntısını veren grafikler verilmiştir. Modelde 2 açıklayıcı değişken, 3 parametre olduğundan, X matrisinin 2 sütunu, J matrisinin 3 sütunu bulunmaktadır. Dolayısıyla X matrisindeki tek korelasyon

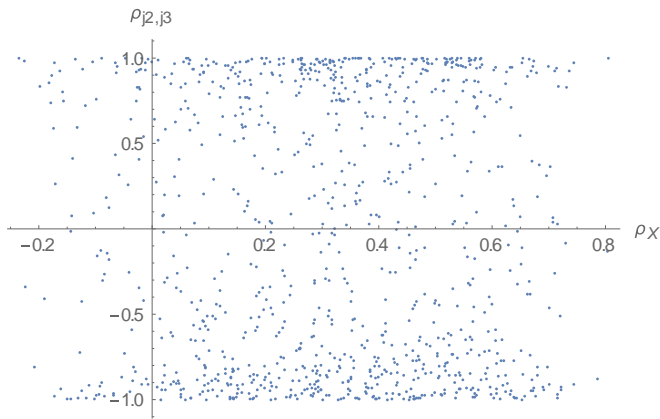
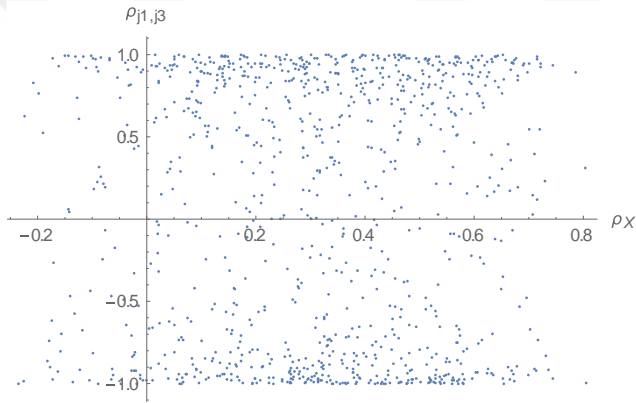
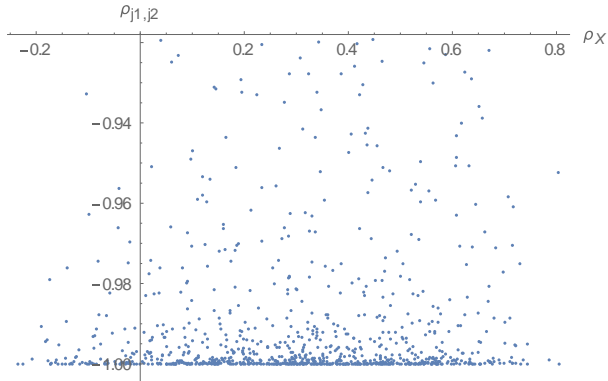
katsayısını ρ_x J matrisindeki üç korelasyon katsayılarını da sırası ile ρ_{j_1,j_2} , ρ_{j_1,j_3} , ρ_{j_2,j_3} temsil etmektedir.



Şekil 3.1 $n = 30$ iken X ile J arasındaki iç ilişki incelemesi

3 grafikten de anlaşılacağı gibi, X matrisindeki iç ilişki ile J matrisindeki iç ilişki arasında belirgin bir ilişki yoktur. Gözlem sayısını $n = 50$ olarak arttırdığımızda, yine,

aşağıdaki grafiklerde olduğu gibi X matrisindeki iç ilişkinin, J matrisindeki iç ilişkiye etkisinin olmadığı görülmüştür.

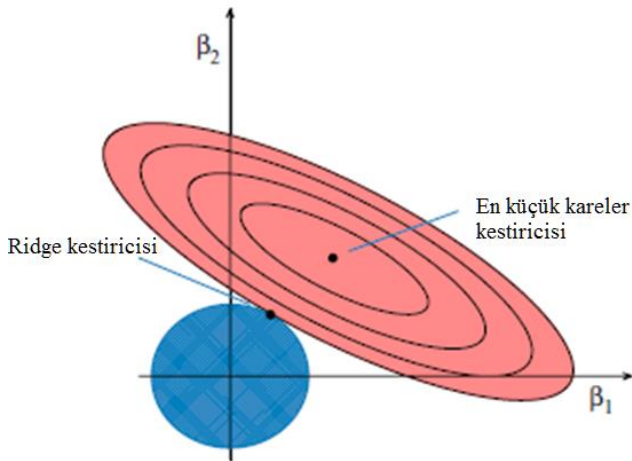


Şekil 3.2 $n = 50$ iken X ile J arasındaki iç ilişki incelemesi

Farklı modeller ve farklı gözlem sayıları ile tekrarlanan similasyonda benzer grafikler bulunmuştur. Sonuç olarak, analitik olarak gösterilemeyen, X matrisindeki iç ilişkinin J matrisine geçişi, similasyon yolu ile geometrik olarak gösterilmeye çalışılmıştır. Böyle bir geçişin olmayabileceği kanısına varılmıştır.

3.5. OPTİMİZASYON PROBLEMİ

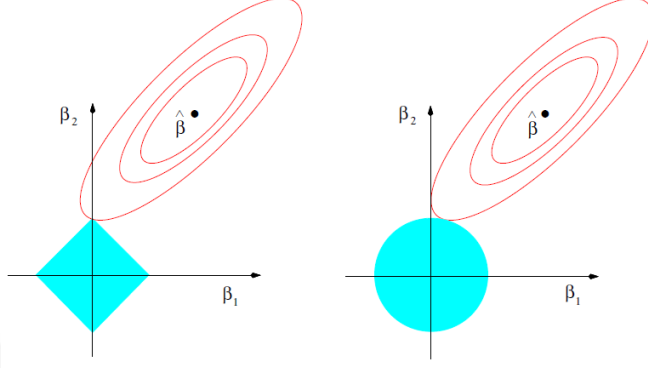
Modeldeki açıklayıcı değişkenler arasında iç ilişki olması durumunda, parametre tahminleri farklı veri kümelerinde çok farklı değerler alabilir. Yani aynı değişkenler ile 1000 farklı veri kümesi üzerinde parametre tahmini yapıldığında birbirinden çok farklı değerler aldığı görülür. Bu da, parametre uzayında, parametre tahminlerinin dağılımına bakıldığında büyük yayılıma (varyans) sahip olduğunu gösterir. Bilindiği gibi en küçük kareler kestiricisi tüm yansız kestiriciler arasında en küçük varyansa sahiptir (BLUE). Bununla birlikte iç ilişkinin parametre tahminleri üzerindeki bu etkisini azaltmak için parametre uzayında parametreleri sınırlayıcı bir kısıt ile tahminlerin ortalamadan sapmalarının azaltılması amaçlanmıştır. Bu amaç doğrultusunda yansızlık koşulu göz ardı edilerek yanlı fakat daha tutarlı kestiriciler elde edilmiştir.



Şekil 3.3 En küçük kareler ve ridge kestiricilerinin parametre uzayındaki yeri

Yukarıdaki grafikte de görüldüğü gibi en küçük kareler kestiricisi iç içe olan elipslerin ortasındaki noktadır. Bu nokta amaç fonksiyonunu en küçükleyen ve yansız olan kestiricidir. Tam merkezde bulunan çember (daire de olabilir) parametre tahmini üzerindeki kısıtı göstermektedir. Kestirmek istenen tahmin değeri, bu çember ile amaç fonksiyonunun değerlerini gösteren kontur grafiğinin kesişim noktasıdır. Amaç,

yansızlık kriterini kaldırarak yanlı fakat daha küçük varyansa sahip tahminler elde etmektir. Burada en önemli nokta, varyanstaki düşüşün büyüklüğünün yanlılıktaki artışın büyüklüğünden fazla olmasıdır. Dolayısıyla hata kareler ortalaması daha küçük kestirim elde edilmiş olur.



Şekil 3.4 Farklı kısıt altında kestiricilerin geometrik yeri

Şekil 3.4’ de ise farklı kısıtlar ile farklı kestiricilerin bulunuşu gösterilmektedir. Soldaki grafik Lasso kestiricisi olarak bilinir.

3.6. İÇ İLİŞKİ PROBLEMİNE YENİ YAKLAŞIM

$y = f(\mathbf{x}, \boldsymbol{\theta}) + \boldsymbol{\varepsilon}$ doğrusal olmayan regresyon modelinin parametrelerini iteratif olarak kestirmek isteyelim. $f(\mathbf{x}, \boldsymbol{\theta})$ fonksiyonunun bir $\boldsymbol{\theta}_0$ başlangıç noktasında Taylor serisine açılımı, yaklaşık olarak,

$$f(\mathbf{x}, \boldsymbol{\theta}) = f(\mathbf{x}, \boldsymbol{\theta}_0) + \sum_{j=1}^p \left[\frac{\partial f(\mathbf{x}, \boldsymbol{\theta})}{\partial \theta_j} \right]_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} (\theta_j - \theta_{j0}) \quad (3.17)$$

şeklindedir. Kesim 2.5’ de yapılan kısaltmalar ile $\mathbf{y} = f(\mathbf{x}, \boldsymbol{\theta}) + \boldsymbol{\varepsilon}$ modelini matris formunda, $\mathbf{y}_0 = \mathbf{J}_0 \boldsymbol{\beta}_0 + \boldsymbol{\varepsilon}$ şeklinde yazabiliriz. Birinci iterasyonda, elde ettiğimiz bu doğrusallaştırılmış modeli,

$$k(\boldsymbol{\beta}_0) \equiv 0 \quad (3.18)$$

özdeş olarak sifıra eşit olan kısıtı altında, kısıtlı en küçük kareler yöntemi ile kestirelim. Bu durumda, ϕ Lagrange çarpanı olmak üzere, yeni amaç fonksiyonu, Lagrange fonksiyonu olarak da ifade edebileceğimiz,

$$L(\phi, \boldsymbol{\beta}_0) = (\mathbf{y}_0 - \mathbf{J}_0 \boldsymbol{\beta}_0)^T (\mathbf{y}_0 - \mathbf{J}_0 \boldsymbol{\beta}_0) + 2\phi k(\boldsymbol{\beta}_0) \quad (3.19)$$

şeklinde elde edilir. Amaç, doğrusal regresyonda olduğu gibi, bu fonksiyonu minimize eden $\hat{\boldsymbol{\beta}}_{0,\phi}$ kestiricisini,

$$\arg \min_{\boldsymbol{\beta}_0} (\mathbf{y}_0 - \mathbf{J}_0 \boldsymbol{\beta}_0)^T (\mathbf{y}_0 - \mathbf{J}_0 \boldsymbol{\beta}_0) + 2\phi k(\boldsymbol{\beta}_0) \quad (3.20)$$

ile bulmaktır. (3.20) ile verilen denklem Levenberg – Marquardt yönteminde verilen (2.20) denklemi ile benzerlik göstermektedir. Bu amaçla,

$$\left. \frac{\partial L(\phi, \boldsymbol{\beta}_0)}{\partial \boldsymbol{\beta}_0} \right|_{\boldsymbol{\beta}_0 = \hat{\boldsymbol{\beta}}_{0,\phi}} = -2\mathbf{J}_0^T (\mathbf{y}_0 - \mathbf{J}_0 \hat{\boldsymbol{\beta}}_{0,\phi}) + 2\phi k'(\hat{\boldsymbol{\beta}}_{0,\phi}) = 0 \quad (3.21)$$

denklemini daha sade olarak,

$$\mathbf{J}_0^T \mathbf{J}_0 \hat{\boldsymbol{\beta}}_{0,\phi} + \phi k'(\hat{\boldsymbol{\beta}}_{0,\phi}) = \mathbf{J}_0^T \mathbf{y}_0 \quad (3.22)$$

denklemi ile verebiliriz. (3.22) ile ifade edilen denklem, doğrusal olmayan regresyonda parametre kestirim sürecinde önerdiğimiz, yeni parametre kestiricilerinin çıkış noktasıdır. Bu denklemi, “**kestirici üreten denklem**” olarak görebiliriz. Kısıtı ortadan kaldırdığımızda (3.22) denklemi,

$$\mathbf{J}_0^T \mathbf{J}_0 \hat{\boldsymbol{\beta}}_{0,\phi} = \mathbf{J}_0^T \mathbf{y}_0$$

doğrusal regresyondaki normal denklemlere dönüşür. Bu denklemlerden elde edilen kestirici,

$$\hat{\boldsymbol{\beta}}_{0,\phi} = \hat{\boldsymbol{\beta}}_0 = (\mathbf{J}_0^T \mathbf{J}_0)^{-1} \mathbf{J}_0^T \mathbf{y}_0$$

en küçük kareler kestiricisidir. Kısıt değıştikçe elde edilecek kestirici değışecektir. Örneğin, (3.22) ile verilen kestirici üreten denklemde, $\mathbf{1} = (1, 1, \dots, 1)^T$ $p \times 1$ boyutlu vektör olmak üzere, $k'(\hat{\boldsymbol{\beta}}_{0,\phi})$ yerine,

$$k'(\hat{\boldsymbol{\beta}}_{0,\phi}) = \hat{\boldsymbol{\beta}}_{0,\phi} + \mathbf{1} \quad (3.23)$$

olacak şekilde yazıldığında, elde edeceğimiz yeni kestirici,

$$\begin{aligned} \mathbf{J}_0^T \mathbf{J}_0 \hat{\boldsymbol{\beta}}_{0,\phi} + \phi(\hat{\boldsymbol{\beta}}_{0,\phi} + \mathbf{1}) &= \mathbf{J}_0^T \mathbf{y}_0 \\ \mathbf{J}_0^T \mathbf{J}_0 \hat{\boldsymbol{\beta}}_{0,\phi} + \phi \hat{\boldsymbol{\beta}}_{0,\phi} + \phi \mathbf{1} &= \mathbf{J}_0^T \mathbf{y}_0 \\ (\mathbf{J}_0^T \mathbf{J}_0 + \phi \mathbf{I}) \hat{\boldsymbol{\beta}}_{0,\phi} &= \mathbf{J}_0^T \mathbf{y}_0 - \phi \mathbf{1} \end{aligned}$$

normal denklemlerinden,

$$\hat{\boldsymbol{\beta}}_{0,\phi} = (\mathbf{J}_0^T \mathbf{J}_0 + \phi \mathbf{I})^{-1} (\mathbf{J}_0^T \mathbf{y}_0 - \phi \mathbf{1}) \quad (3.24)$$

şeklinde bulunur. Daha önce $\boldsymbol{\beta}_0 = \boldsymbol{\theta} - \boldsymbol{\theta}_0$ dönüşümü yapıldığından, ötelenmiş kestirici,

$$\hat{\boldsymbol{\theta}}_{1,\phi} = \boldsymbol{\theta}_0 + \hat{\boldsymbol{\beta}}_{0,\phi} = \boldsymbol{\theta}_0 + (\mathbf{J}_0^T \mathbf{J}_0 + \phi \mathbf{I})^{-1} (\mathbf{J}_0^T \mathbf{y}_0 - \phi \mathbf{1}) \quad \text{şeklinde elde edilmiş olur.}$$

$k'(\boldsymbol{\beta}_0) = \boldsymbol{\beta}_0 + \mathbf{1}$ eşitliğinde, kısıtı bulmak için, eşitliğin her iki tarafını integre edelim:

\mathbf{F} parametre uzayında $p \times 1$ boyutlu sabit bir vektör olmak üzere,

$$\begin{aligned} \int k'(\boldsymbol{\beta}_0) d\boldsymbol{\beta}_0 &= \int (\boldsymbol{\beta}_0 + \mathbf{1}) d\boldsymbol{\beta}_0 \\ k(\boldsymbol{\beta}_0) &= \frac{1}{2} \boldsymbol{\beta}_0^T \boldsymbol{\beta}_0 + \boldsymbol{\beta}_0^T \mathbf{1} + \mathbf{F}^T \mathbf{1} \end{aligned} \quad (3.25)$$

kısıtını elde etmiş oluruz. Kısıt özdeş olarak sıfıra eşit olduğundan, (3.25) ifadesini biraz daha düzenleyecek olursak,

$$\begin{aligned} \frac{1}{2} \boldsymbol{\beta}_0^T \boldsymbol{\beta}_0 + \boldsymbol{\beta}_0^T \mathbf{1} + \mathbf{F}^T \mathbf{1} &= 0 \\ \boldsymbol{\beta}_0^T \boldsymbol{\beta}_0 + 2\boldsymbol{\beta}_0^T \mathbf{1} + 2\mathbf{F}^T \mathbf{1} &= 0 \\ (\boldsymbol{\beta}_0 + \mathbf{1})^T (\boldsymbol{\beta}_0 + \mathbf{1}) &= \mathbf{1}^T \mathbf{1} - 2\mathbf{F}^T \mathbf{1} \end{aligned} \quad (3.26)$$

ifadesi elde edilir. $\mathbf{1}^T \mathbf{1} - 2\mathbf{F}^T \mathbf{1} = \mathbf{D}^T \mathbf{D}$ kısaltması ile kısıt,

$$(\boldsymbol{\beta}_0 + \mathbf{I})^T (\boldsymbol{\beta}_0 + \mathbf{I}) = \mathbf{D}^T \mathbf{D} \quad (3.27)$$

şeklini alır. Kısıtın geometrik olarak ne anlama geldiğini göstermek için Öklid uzayında (3.27) ifadesini normlu gösterebiliriz:

$$\|\boldsymbol{\beta}_0 + \mathbf{I}\|^2 = \|\mathbf{D}\|^2 \quad (3.28)$$

Bu kısıt ile kestirilmek istenen parametreyi, merkezi, parametre uzayında tüm bileşenleri -1 değerinden oluşan nokta olan çember üzerine kısıtlamış oluruz.

İteratif kestirim süreci yakınsaklık sağlanıncaya kadar devam eder. Bu süreçte, tüm iterasyonlarda yukarıda anlatılan kısıtlı en küçük kareler yöntem uygulanır.

Yakınsaklık sağlandığında elde edilen son kestirim, genel olarak, $\hat{\boldsymbol{\theta}}_{m+1,\phi} = \hat{\boldsymbol{\theta}}_{m,\phi} + \hat{\boldsymbol{\beta}}_{m,\phi} = \hat{\boldsymbol{\theta}}_{m,\phi} + (\mathbf{J}_m^T \mathbf{J}_m + \phi \mathbf{I})^{-1} (\mathbf{J}_m^T \mathbf{y}_m - \phi \mathbf{I})$ şeklinde bulunur.

Bu yaklaşım ile doğrusal regresyon modellerindeki ridge regresyon yaklaşımı benzerlik göstermektedir. Ridge regresyonda, parametre kestirimleri merkezi orijinde olan çember üzerinde kısıtlanır. Böylelikle iç ilişkiden kaynaklanan yüksek varyans problemine, yansızlık koşulu gözardı edilerek, çözüm üretilmiş olur. Yukarıda elde ettiğimiz kestirici, ridge regresyondaki merkezi orijinde olan çemberin ötelenmesi ile elde edilmiştir.

Bir başka örnek olarak, (3.22) denkleminde $k'(\hat{\boldsymbol{\beta}}_{0,\phi}) = \hat{\boldsymbol{\beta}}_{0,\phi} + \mathbf{I}$ yerine $k'(\hat{\boldsymbol{\beta}}_{0,\phi}) = \hat{\boldsymbol{\beta}}_{0,\phi} - \mathbf{I}$ yazacak olursak, kısıt, $\|\boldsymbol{\beta}_0 - \mathbf{I}\|^2 = \|\mathbf{D}\|^2$ şeklinde olur. Bu kısıt altında kestirilen bir diğer yeni kestirici,

$$\hat{\boldsymbol{\beta}}_{0,\phi} = (\mathbf{J}_0^T \mathbf{J}_0 + \phi \mathbf{I})^{-1} (\mathbf{J}_0^T \mathbf{y}_0 + \phi \mathbf{I}) \quad (3.29)$$

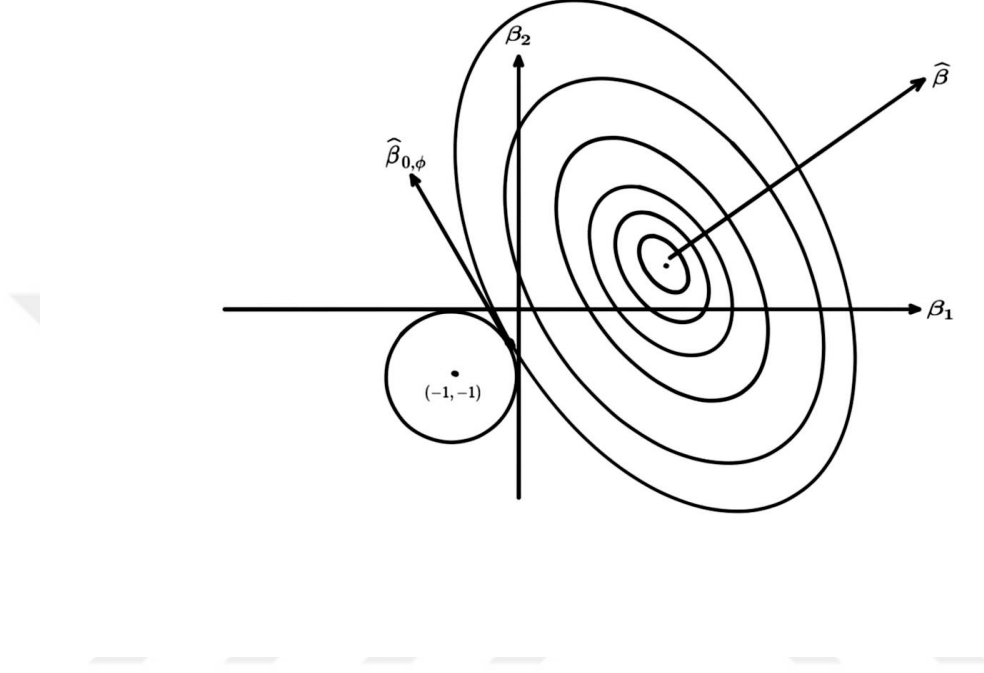
elde edilmiş olur. Ötelenmiş kestirici ise,

$$\hat{\boldsymbol{\theta}}_{1,\phi} = \boldsymbol{\theta}_0 + \hat{\boldsymbol{\beta}}_{0,\phi} = \boldsymbol{\theta}_0 + (\mathbf{J}_0^T \mathbf{J}_0 + \phi \mathbf{I})^{-1} (\mathbf{J}_0^T \mathbf{y}_0 + \phi \mathbf{I}) \quad \text{şeklinde elde edilmiş olur.}$$

Yakınsaklık sağlandığında elde edilen son kestirim,

$$\hat{\boldsymbol{\theta}}_{m+1,\phi} = \hat{\boldsymbol{\theta}}_{m,\phi} + \hat{\boldsymbol{\beta}}_{m,\phi} = \hat{\boldsymbol{\theta}}_{m,\phi} + (\mathbf{J}_m^T \mathbf{J}_m + \phi \mathbf{I})^{-1} (\mathbf{J}_m^T \mathbf{y}_m + \phi \mathbf{I}) \quad \text{şeklinde bulunur.}$$

Aşağıda verilen Şekil 3.5’ de görüldüğü üzere, $\hat{\beta}_{0,\phi}$ kestiricisinin kısıtlandığı çember bölgesi $(-1,-1)$ merkezlidir. Doğrusal regresyondaki ridge kestiricisinin kısıtlandığı çember merkezi orijindedir.



Şekil 3.5 İki boyutlu uzayda $\hat{\beta}_{0,\phi}$ kestiricisinin konumu

Şekil 3.5 ile birlikte, düşünülen ilk soru, kestirilmek istenen parametreyi kısıtlayan çemberin (ya da daire) uzaydaki konumunun ne olması gerektiğidir. Çünkü, kısıt bölgesi gerçek parametreye ne kadar yakınsa, yakınsamanın hızı ve hata kareler ortalamasının minimum olması o kadar olasıdır. Bu durumda, parametre uzayında, kestirilmek istenen parametreye ilişkin ön bilgi önem kazanmaktadır.

3.6.1. ÖN BİLGİYE DAYALI YENİ KESTİRİCİ

Doğrusal regresyonda Hoerl ve Kennard (1970), iç ilişki probleminde çözüm olarak $\hat{\beta}_{RR} = (\mathbf{X}^T \mathbf{X} + k\mathbf{I})^{-1} \mathbf{X}^T \mathbf{y}$ ridge kestiricisini önermiştir. Önerdikleri kestiricinin en küçük kareler kestiricisinden daha küçük hata kareler ortalamasına sahip olacak şekilde $k > 0$ ridge parametresinin varlığını göstermişlerdir. Bununla birlikte, Swindel (1976), modifiye ridge regresyon kestiricisini (modified ridge regression estimator) önermiştir. Önerilen yaklaşımda, kestirilmek istenen parametreye ilişkin ön bilgi ile

ridge regresyon birleştirilmiştir. \mathbf{b} , $\boldsymbol{\beta}$ 'nın stokastik olmayan bir ön bilgisi olmak üzere Swindel' in önerdiği kestirici,

$$\hat{\boldsymbol{\beta}}(k, \mathbf{b}) = (\mathbf{X}^T \mathbf{X} + k\mathbf{I})^{-1} (\mathbf{X}^T \mathbf{y} + k\mathbf{b}) \quad , \quad k \geq 0 \quad (3.30)$$

şeklindedir. Swindel (1976), ayrıca, önerdiği kestircinin hata kareler ortalamasının en küçük kareler kestircisinin hata kareler ortalamasından daha küçük olmasını sağlayan k değerlerinin olduğunu ispatlamıştır. $\boldsymbol{\beta}$ 'nın ön bilgisi olan \mathbf{b} vektörünün sabit değil, stokastik yani bir dağılıma sahip olduğu durumda da Swindel (1976), önerdiği kestircinin hata kareler ortalamasının en küçük kareler kestircisinin hata kareler ortalamasından daha küçük olmasını sağlayan bir k değerinin olduğunu göstermiştir (Crouse, et al., 1995). Ancak \mathbf{b} ön bilgi vektörünün $\hat{\boldsymbol{\beta}}$ en küçük kareler kestircisinden bağımsız olmadığı durumda, Swindel'in önerdiği kestircinin hata kareler ortalaması kriterine göre daha iyi olduğunun garantisi yoktur (Crouse, et al., 1995). Örneğin, eğer $\mathbf{b} = \hat{\boldsymbol{\beta}}$ olarak seçilirse, $\hat{\boldsymbol{\beta}}(k, \mathbf{b}) = \hat{\boldsymbol{\beta}}$ sağlanır. Swindel (1976), k ridge parametresinin kestirim yönteminden bahsetmemiştir.

Crouse ve arkadaşları (1995), Swindel (1976)' in kestircisine benzer bir kestirici önermişlerdir. Yansız Ridge Regresyon Kestircisi (Unbiased Ridge Regression Estimator) olarak tanımladıkları kestirici,

$$\hat{\boldsymbol{\beta}}(k, \mathbf{B}) = (\mathbf{X}^T \mathbf{X} + k\mathbf{I})^{-1} (\mathbf{X}^T \mathbf{y} + k\mathbf{B}) \quad , \quad k \geq 0$$

şeklindedir. Burada $\mathbf{B} \sim N(\boldsymbol{\beta}, \frac{\sigma^2}{k} \mathbf{I})$ şeklinde normal dağılıma sahip $\boldsymbol{\beta}$ 'nın ön bilgisi olarak kabul edilir. Ayrıca ridge parametresi olan k değerini,

$$\hat{k}_{CJK} = \begin{cases} \frac{p\hat{\sigma}^2}{(\hat{\boldsymbol{\beta}} - \mathbf{B})^T (\hat{\boldsymbol{\beta}} - \mathbf{B}) - \hat{\sigma}^2 \text{tr}(\mathbf{X}^T \mathbf{X})^{-1}} \quad , \quad (\hat{\boldsymbol{\beta}} - \mathbf{B})^T (\hat{\boldsymbol{\beta}} - \mathbf{B}) > \hat{\sigma}^2 \text{tr}(\mathbf{X}^T \mathbf{X})^{-1} \\ \frac{p\hat{\sigma}^2}{(\hat{\boldsymbol{\beta}} - \mathbf{B})^T (\hat{\boldsymbol{\beta}} - \mathbf{B})} \quad , \quad \text{diğer durumlarda} \end{cases}$$

şeklinde önermişlerdir. Burada, $\hat{\sigma}^2$ değeri, σ^2 nin yansız kestircisi olup, aşağıdaki gibi tanımlamışlardır:

$$\hat{\sigma}^2 = \frac{(\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}})^T (\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}})}{n - p}$$

Crouse ve arkadaşları (1995), önerdikleri k parametresinin Hoerl ve arkadaşları

(1975) tarafından önerilen $\hat{k}_{HKB} = \frac{p\hat{\sigma}^2}{\hat{\boldsymbol{\beta}}^T \hat{\boldsymbol{\beta}}}$ ridge parametresinin genelleştirilmiş hali

olduğunu söylemişlerdir. Ekte sunulan teorem ile önerdikleri kestiricinin hangi koşul altında hata kareler ortalamasının en küçük olduğunu kanıtlamışlardır. Buna göre

önerilen $\beta(C, B) = C\hat{\boldsymbol{\beta}} + (I - C)B$ şeklindeki konveks kestiricinin hata kareler ortalaması C matrisinin $C = V(V + \Sigma)^{-1}$ seçimi ile en küçük karelere göre daha küçük bulunmuştur.

$\beta(C, B) = C\hat{\boldsymbol{\beta}} + (I - C)B$ şeklinde önerilen konveks kestiricide C matrisini yazdığımızda, hata kareler ortalaması minimum olan kestiriciyi buluruz. Bunun için öncelikle (Ek A.5) denklemini düzenleyelim:

$$\begin{aligned} C &= V(V + \Sigma)^{-1} \\ &= V[V(I + V^{-1}\Sigma)]^{-1} \\ &= VV^{-1}[I + V^{-1}\Sigma]^{-1} \\ &= [I + V^{-1}\sigma^2(X^T X)^{-1}]^{-1} \\ &= [I + k(X^T X)^{-1}]^{-1}, \quad k = V^{-1}\sigma^2 \end{aligned} \tag{3.37}$$

C matrisinin son denklemdaki eşitini konveks kestiricide yerine yazdığımızda,

$$\begin{aligned} \beta(C, B) &= C\hat{\boldsymbol{\beta}} + (I - C)B \\ &= C(X^T X)^{-1} X^T Y + (I - C)B \\ &= [I + k(X^T X)^{-1}]^{-1} (X^T X)^{-1} X^T Y + B - [I + k(X^T X)^{-1}]^{-1} B \\ &= (X^T X + kI)^{-1} (X^T X) (X^T X)^{-1} X^T Y + B - [I + k(X^T X)^{-1}]^{-1} B \\ &= (X^T X + kI)^{-1} X^T Y + B - [I + k(X^T X)^{-1}]^{-1} (X^T X)^{-1} (X^T X) B \\ &= (X^T X + kI)^{-1} X^T Y + B - (X^T X + kI)^{-1} (X^T X) (X^T X)^{-1} (X^T X) B \\ &= (X^T X + kI)^{-1} X^T Y + B - (X^T X + kI)^{-1} (X^T X + kI - kI) B \\ &= (X^T X + kI)^{-1} X^T Y + B - (X^T X + kI)^{-1} (X^T X + kI) B \\ &\quad + (X^T X + kI)^{-1} kB \\ &= (X^T X + kI)^{-1} X^T Y + (X^T X + kI)^{-1} kB \\ &= (X^T X + kI)^{-1} (X^T Y + kB) \end{aligned} \tag{3.38}$$

Crouse ve arkadaşlarının (1995) önerdikleri kestirici, minimum hata kareler ortalamasına sahip olarak bulunmuş olur.

Öte yandan, (3.22) ile verilen $\mathbf{J}_0^T \mathbf{J}_0 \hat{\boldsymbol{\beta}}_{0,\phi} + \phi k'(\hat{\boldsymbol{\beta}}_{0,\phi}) = \mathbf{J}_0^T \mathbf{y}_0$ kestirici üreten denklemde, kısıtın türevi olan,

$$k'(\hat{\boldsymbol{\beta}}_{0,\phi}) = \hat{\boldsymbol{\beta}}_{0,\phi} - \mathbf{M} \quad (3.39)$$

ifadesini yazdığımızda,

$$\mathbf{J}_0^T \mathbf{J}_0 \hat{\boldsymbol{\beta}}_{0,\phi} + \phi(\hat{\boldsymbol{\beta}}_{0,\phi} - \mathbf{M}) = \mathbf{J}_0^T \mathbf{y}_0 \quad (3.40)$$

denklemini elde ederiz. Bu denklemi düzenlediğimizde,

$$\begin{aligned} \mathbf{J}_0^T \mathbf{J}_0 \hat{\boldsymbol{\beta}}_{0,\phi} + \phi \hat{\boldsymbol{\beta}}_{0,\phi} - \phi \mathbf{M} &= \mathbf{J}_0^T \mathbf{y}_0 \\ (\mathbf{J}_0^T \mathbf{J}_0 + \phi \mathbf{I}) \hat{\boldsymbol{\beta}}_{0,\phi} &= \mathbf{J}_0^T \mathbf{y}_0 + \phi \mathbf{M} \end{aligned} \quad (3.41)$$

normal denklemlerinden,

$$\hat{\boldsymbol{\beta}}_{0,\phi} = (\mathbf{J}_0^T \mathbf{J}_0 + \phi \mathbf{I})^{-1} (\mathbf{J}_0^T \mathbf{y}_0 + \phi \mathbf{M}) \quad (3.42)$$

kestiricisi elde edilir. (3.42) ile verilen kestiriciyi elde etmek için kullanılan kısıtın türevi olan $k'(\boldsymbol{\beta}_{0,\phi}) = \boldsymbol{\beta}_{0,\phi} - \mathbf{M}$ ifadesini integre ettiğimizde,

$$\begin{aligned} \int k'(\boldsymbol{\beta}_0) d\boldsymbol{\beta}_0 &= \int (\boldsymbol{\beta}_0 - \mathbf{M}) d\boldsymbol{\beta}_0 \\ k(\boldsymbol{\beta}_0) &= \frac{1}{2} \boldsymbol{\beta}_0^T \boldsymbol{\beta}_0 - \boldsymbol{\beta}_0^T \mathbf{M} + \mathbf{F}^T \mathbf{1} \end{aligned} \quad (3.43)$$

kısıtını elde etmiş oluruz. Kısıt özdeş olarak sıfıra eşit olduğundan, (3.43) ifadesini biraz daha düzenleyecek olursak,

$$\begin{aligned} \frac{1}{2} \boldsymbol{\beta}_0^T \boldsymbol{\beta}_0 - \boldsymbol{\beta}_0^T \mathbf{M} + \mathbf{F}^T \mathbf{1} &= \mathbf{0} \\ \boldsymbol{\beta}_0^T \boldsymbol{\beta}_0 - 2\boldsymbol{\beta}_0^T \mathbf{M} + 2\mathbf{F}^T \mathbf{1} &= \mathbf{0} \\ (\boldsymbol{\beta}_0 - \mathbf{M})^T (\boldsymbol{\beta}_0 - \mathbf{M}) &= \mathbf{M}^T \mathbf{M} - 2\mathbf{F}^T \mathbf{1} \end{aligned} \quad (3.44)$$

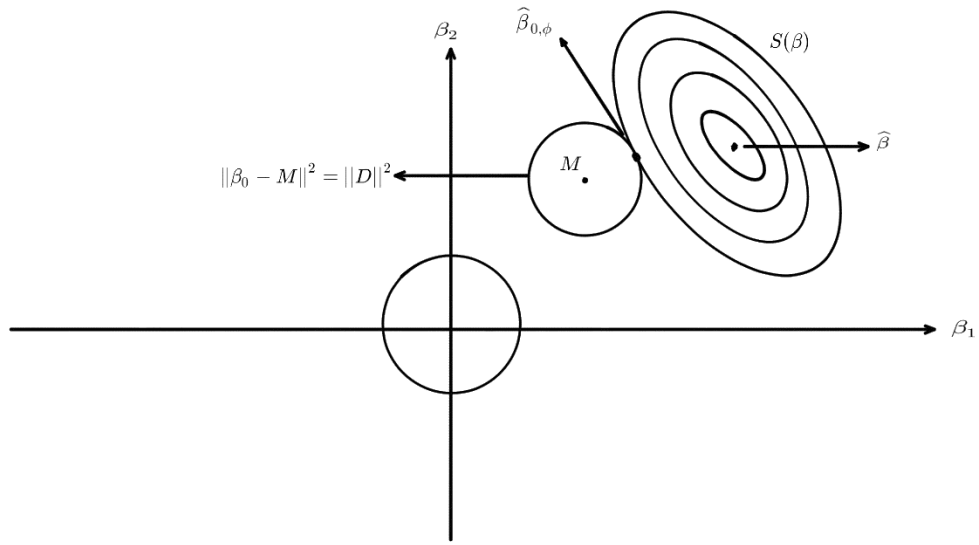
ifadesini elde ederiz. $\mathbf{D}^T \mathbf{D} = \mathbf{M}^T \mathbf{M} - 2\mathbf{F}^T \mathbf{1}$ kısaltması ile kısıt,

$$(\boldsymbol{\beta}_0 - \mathbf{M})^T (\boldsymbol{\beta}_0 - \mathbf{M}) = \mathbf{D}^T \mathbf{D} \quad (3.45)$$

şeklinde elde edilir. Normlu uzayda kısıt,

$$\|\boldsymbol{\beta}_0 - \mathbf{M}\|^2 = \|\mathbf{D}\|^2 \quad (3.46)$$

şeklindedir. Bu kısıt ile kestirilmek istenen parametreyi, merkezi parametre uzayında \mathbf{M} noktası, yarıçapı $\|\mathbf{D}\|$ olan çember üzerine, Şekil 3.6 'daki gibi, kısıtlamış oluyoruz.



Şekil 3.6 Ön bilgiye dayalı kestiricinin parametre uzayındaki konumu

Şekil 3.6 incelenecek olursa, $S(\boldsymbol{\beta})$ konturlarının merkezi \mathbf{M} noktasında olan çemberi kestiği ilk nokta önerdiğimiz kestiriciyi vermektedir. Merkezi orijinde olan çember, doğrusal modellerdeki ridge kestiricisinin elde edildiği kısıt çemberidir.

Daha önce $\boldsymbol{\beta}_0 = \boldsymbol{\theta} - \boldsymbol{\theta}_0$ dönüşümü yapıldığından, ötelenmiş kestirici,

$\hat{\boldsymbol{\theta}}_{1,\phi} = \boldsymbol{\theta}_0 + \hat{\boldsymbol{\beta}}_{0,\phi} = \boldsymbol{\theta}_0 + (\mathbf{J}_0^T \mathbf{J}_0 + \phi \mathbf{I})^{-1} (\mathbf{J}_0^T \mathbf{y}_0 + \phi \mathbf{M})$ şeklinde elde edilmiş olur. Bu iteratif kestirim süreci yakınsaklık sağlanıncaya kadar devam eder. Bu süreçte, tüm iterasyonlarda aynı yöntem uygulanır. Yakınsaklık sağlandığında elde edilen son kestirim, genel olarak,

$$\hat{\theta}_{m+1,\phi} = \hat{\theta}_{m,\phi} + \hat{\beta}_{m,\phi} = \hat{\theta}_{m,\phi} + (\mathbf{J}_m^T \mathbf{J}_m + \phi \mathbf{I})^{-1} (\mathbf{J}_m^T \mathbf{y}_m + \phi \mathbf{M})$$

şeklinde bulunur.

Geometrik olarak anlatılan bu yeni yaklaşım, doğrusal modellerde, Swindel (1976) ve Crouse ve arkadaşlarının (1995) önerdikleri, ön bilgiye dayalı kestiricilere benzer bir yaklaşımdır. Swindel' in (1976) önerdiği kestiricide, β parametresi için kullanılan ön bilgi vektörü, \mathbf{b} , ya da Crouse ve arkadaşlarının (1995) önerdikleri kestiricide ön bilgi vektörü \mathbf{B} , bizim yaklaşımımızda \mathbf{M} vektörüne karşılık gelmektedir. Önerdiğimiz yaklaşım ile, kestirilmek istenen β parametresini, uzayda, merkezi ön bilgi vektörü, \mathbf{M} , olan çember üzerine kısıtlamış oluyoruz. Bu şekilde Swindel (1976) ile Crouse ve arkadaşlarının (1995) kestiricilerine benzer bir kestirici ile doğrusal olmayan modellerde iç ilişki problemine çözüm üretmiş oluyoruz. Parametre uzayında, kestirilmek istenen parametreyi herhangi bir nokta ile belirlenen çember yerine, o parametreye ilişkin ön bilgi noktasının belirttiği çember üzerine kısıtlamak daha mantıklıdır. Çünkü, ön bilgi vektörü gerçek parametre vektörüne ne kadar yakınsa, kestirici de o ölçüde gerçek parametreye daha yakın olacaktır. Böylelikle, iç ilişkiden kaynaklanan, parametrelerin uzaydaki büyük yayılımları, ön bilgi civarında kısıtlanarak, daha tutarlı kestiriciler elde edilecektir.

Özkale ve Kaçıranlar (2007) çalışmalarında, Crouse ve arkadaşlarının (1995) ve Swindel' in (1976) önerdikleri kestiricilerin, doğrusal modellerde en küçük kareler kestircisinden, hata kareler ortalaması kriterine göre, daha iyi olduğunu göstermişlerdir. Bizim çalışmamızda ise iteratif olarak elde edilen \mathbf{J} Jakobiyeen matrisi, \mathbf{X} tasarım matrisinin rolünü üstlendiğinden, doğrusal modellerde en küçük kareler kestircisine üstünlük gösteren kestiricileri, iteratif olarak doğrusal olmayan modellere genişletebiliriz. Bununla birlikte, Gauss – Newton yöntemi ile $\hat{\theta}$ kestirilirken, \mathbf{J} Jakobiyeen matrisi her iterasyonda hesaplanıp kullanıldığından, doğrusal olmayan en küçük kareler için yapılacak çıkarsamalarda da bu matrisi kullanmak doğaldır (Bates & Watts, 1988). Dolayısıyla parametre kestiriminde önerdiğimiz yeni kestirici de doğrusallaştırma yaklaşımı ile elde edildiğinden, kestirici ile ilgili yapılacak olan çıkarsamalar da \mathbf{J} matrisinin kullanılması doğaldır.

3.6.2. HATA KARELER ORTALAMALARI KARŞILAŞTIRILMASI

$y = f(\mathbf{x}, \boldsymbol{\theta}) + \boldsymbol{\varepsilon}$ doğrusal olmayan regresyon modelinin parametrelerini kestirmek için $f(\mathbf{x}, \boldsymbol{\theta})$ fonksiyonunun yakınsaklık sağlanıncaya kadar iteratif olarak doğrusallaştırdığımızı düşünelim. Kesim 2.5' de $f(\mathbf{x}, \boldsymbol{\theta})$ fonksiyonunun bir $\boldsymbol{\theta}_0$ başlangıç noktasında matris formunda, $\mathbf{y}_0 = \mathbf{J}_0 \boldsymbol{\beta}_0 + \boldsymbol{\varepsilon}$ şeklinde yazabileceğimizi göstermiştik. Birinci iterasyonda, elde ettiğimiz bu doğrusallaştırılmış modeli, önceki bölümde anlatıldığı gibi, iteratif olarak, kısıtlı en küçük kareler yöntemi ile kestirelim. Yakınsaklığın sağlandığı son doğrusallaştırılmış modeli, daha anlaşılır olması açısından, indis kullanmadan,

$$\mathbf{y} = \mathbf{J} \boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (3.47)$$

olarak gösterelim. $\tilde{\boldsymbol{\beta}}$, (3.47) modelindeki $\boldsymbol{\beta}$ için herhangi bir kestirici olsun. Matris hata kareler ortalaması (*MHKO*),

$$MHKO(\tilde{\boldsymbol{\beta}}) = E[(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta})(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta})^T] \quad (3.48)$$

olarak tanımlanır. Aynı kestirici için varyans kovaryans matrisi,

$$Cov(\tilde{\boldsymbol{\beta}}) = E[(\tilde{\boldsymbol{\beta}} - E(\tilde{\boldsymbol{\beta}}))(\tilde{\boldsymbol{\beta}} - E(\tilde{\boldsymbol{\beta}}))^T] \quad (3.49)$$

şeklinde tanımlanır. $\tilde{\boldsymbol{\beta}}$ kestiricisinin yanlılığı ise,

$$yan(\tilde{\boldsymbol{\beta}}) = E(\tilde{\boldsymbol{\beta}}) - \boldsymbol{\beta} \quad (3.50)$$

ifadesi ile bulunur. Dolayısıyla (3.48) eşitliği, özdeş olarak,

$$MHKO(\tilde{\boldsymbol{\beta}}) = Cov(\tilde{\boldsymbol{\beta}}) + [yan(\tilde{\boldsymbol{\beta}})][yan(\tilde{\boldsymbol{\beta}})]^T \quad (3.51)$$

şeklinde ifade edilebilir (Özkale & Kaçırınlar, 2007). *MHKO* kestirici ile ilgili tüm bilgileri içerirken, Skaler Hata Kareler Ortalaması (*SHKO*) olarak tanımlanan,

$$SHKO(\tilde{\boldsymbol{\beta}}) = E[(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta})^T (\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta})] \quad (3.52)$$

ifadesi, kestiricinin tüm bileşenlerinin toplam varyansını vermektedir. Dolayısıyla (3.52) ifadesi,

$$SHKO(\tilde{\beta}) = \dot{I}z(MHKO(\tilde{\beta})) \quad (3.53)$$

olarak da hesaplanabilir. Theobald (1974) çalışmasında *SHKO* ile *MHKO* arasındaki ilişkiyi şöyle vermiştir: $\tilde{\beta}_1$ ve $\tilde{\beta}_2$ β ' nin iki kestiricisi olsun. Eğer $MHKO(\tilde{\beta}_1) - MHKO(\tilde{\beta}_2)$ matrisi negatif olmayan tanımlı (non negative definite, n.n.d.) matris ise, $SHKO(\tilde{\beta}_1) - SHKO(\tilde{\beta}_2) \geq 0$ eşitsizliği sağlanır. Ancak, bunun tersi doğru değildir.

Ön bilgiye dayalı önerdiğimiz $\hat{\beta}_{0,\phi} = (J_0^T J_0 + \phi I)^{-1} (J_0^T y_0 + \phi M)$ kestirici, $S_k = J^T J + kI$, $S = J^T J$ ve $\hat{\beta}$ en küçük kareler kestiricisi olmak üzere,

$$\hat{\beta}_{0,\phi} = S_k^{-1} S \hat{\beta} + k S_k^{-1} M \quad (3.54)$$

şeklinde de yazılabilir. Crouse ve arkadaşlarının (1995) önerdikleri $\hat{\beta}(k, B) = (X^T X + kI)^{-1} (X^T y + kB)$ kestiricinin yansız kestirici olduğunu aşağıdaki Teorem 3.1 ile vermişlerdir.

Teorem 3.1

$\hat{\beta}$, β ortalamalı, $\sigma^2 (X^T X)^{-1}$ varyans kovaryans matrisine sahip, herhangi bir dağılıma sahip olan, doğrusal modellerdeki en küçük kareler kestiricisi olsun. B , β ortalamalı, $\frac{\sigma^2}{k} I$ varyans kovaryans matrisine sahip olsun ($k > 0$).

$\beta(C, B) = C \hat{\beta} + (I - C) B$ kestiricisini tanımlayalım. En küçük HKO değerine sahip olacak şekilde seçilen optimal C için, $\beta(C, B) = \hat{\beta}(kI, B) = (X^T X + kI)^{-1} (X^T y + kB)$ kestiricisi β ' nin yansız bir kestiricisidir.

Önerdiğimiz $\hat{\beta}_{0,\phi}$ kestiricinin, iteratif olarak, her bir doğrusallaştırılmış düzlem (model) üzerinde, asimptotik olarak yansız olduğunu Teorem 3.1 'e göre

söyleyebiliriz. Bu durumda, $MHKO(\hat{\beta}_{0,\phi}) = Cov(\hat{\beta}_{0,\phi})$ eşitliği sağlanır. (3.49)
eşitliğinden yararlanarak,

$$MHKO(\hat{\beta}_{0,\phi}) = \sigma^2 \mathbf{S}_k^{-1} \mathbf{S} \mathbf{S}_k^{-1} + k \sigma^2 \mathbf{S}_k^{-2} \quad (3.55)$$

eşitliği elde edilir. $\hat{\beta}$ en küçük kareler kestiricisinin hata kareler ortalaması, $MHKO(\hat{\beta}) = \sigma^2 \mathbf{S}^{-1}$ şeklindedir. Önerdiğimiz kestiricinin, en küçük kareler kestiricisinden daha küçük hata kareler ortalamasına sahip olduğunu göstermek için fark matrisini bulmamız gerekir.

$$MHKO(\hat{\beta}) - MHKO(\hat{\beta}_{0,\phi}) = \sigma^2 \mathbf{S}^{-1} - \sigma^2 \mathbf{S}_k^{-1} \mathbf{S} \mathbf{S}_k^{-1} - k \sigma^2 \mathbf{S}_k^{-2} \quad (3.56)$$

ifadesini negatif olmayan tanımlı yapacak olan k değerini bulmak için biraz düzenlediğimizde,

$$\begin{aligned} MHKO(\hat{\beta}) - MHKO(\hat{\beta}_{0,\phi}) &= \sigma^2 \mathbf{S}^{-1} - \sigma^2 \mathbf{S}_k^{-1} \mathbf{S} \mathbf{S}_k^{-1} - k \sigma^2 \mathbf{S}_k^{-2} \\ &= \sigma^2 \mathbf{S}^{-1} - (\sigma^2 \mathbf{S}_k^{-1} \mathbf{S} \mathbf{S}_k^{-1} + k \sigma^2 \mathbf{S}_k^{-2}) \\ &= \sigma^2 \mathbf{S}^{-1} - [\sigma^2 \mathbf{S}_k^{-2} (\mathbf{S} + k \mathbf{I})] \quad (\mathbf{S} + k \mathbf{I} = \mathbf{S}_k) \\ &= \sigma^2 \mathbf{S}^{-1} - [\sigma^2 \mathbf{S}_k^{-2} \mathbf{S}_k] \\ &= \sigma^2 \mathbf{S}^{-1} - \sigma^2 \mathbf{S}_k^{-1} \\ &= \sigma^2 \mathbf{S}_k^{-1} \mathbf{S}^{-1} \mathbf{S}_k \mathbf{S} (\mathbf{S}^{-1} - \mathbf{S}_k^{-1}) \\ &= \sigma^2 \mathbf{S}_k^{-1} \mathbf{S}^{-1} (\mathbf{S}_k \mathbf{S} \mathbf{S}^{-1} - \mathbf{S}_k \mathbf{S} \mathbf{S}_k^{-1}) \\ &= \sigma^2 \mathbf{S}_k^{-1} \mathbf{S}^{-1} (\mathbf{S}_k - \mathbf{S}) \quad (\mathbf{S}_k - \mathbf{S} = k \mathbf{I}) \\ &= k \sigma^2 \mathbf{S}_k^{-1} \mathbf{S}^{-1} \end{aligned} \quad (3.57)$$

elde edilir. (3.57) ile verilen matris $k > 0$ için pozitif tanımlı (positive definite, p.d.) olduğundan, pozitif tüm k değerleri için önerdiğimiz kestirici, en küçük kareler kestiricisinden $MHKO$ kriterine göre daha iyi bir kestiricidir.

Doğrusal olmayan regresyon modellerinde, herhangi bir iterasyonda kestirilen parametre vektörü, bir sonraki kestirim için öteleme vektörü olarak alınır. Dolayısıyla her bir adımda elde edilen kestirim vektörüne öteleme vektörünün eklenmesi, beklenen değer dönüşümünün doğrusal bir dönüşüm olmasından dolayı, parametre kestiriminin asimptotik yansızlığını etkilememektedir.

4. BENZETİM ÇALIŞMASI

Önceki bölümde önerdiğimiz yaklaşımların performansını, yaptığımız benzetim çalışması ile bu bölümde inceleyeceğiz. Yaptığımız benzetim çalışması, Wolfram Mathematica 10 kodlu paket programı kullanılarak hazırlanmıştır. Yapılan çalışmada, Gauss – Newton kestiricisi,

$$\hat{\theta}_{GN} = (\mathbf{J}^T \mathbf{J})^{-1} (\mathbf{J}^T \mathbf{y})$$

ile bizim önerdiğimiz,

$$\hat{\theta}_{yeni1} = \hat{\theta}_{IRR} = (\mathbf{J}^T \mathbf{J} + \phi \mathbf{I})^{-1} (\mathbf{J}^T \mathbf{y})$$

$$\hat{\theta}_{yeni2} = (\mathbf{J}^T \mathbf{J} + \phi \mathbf{I})^{-1} (\mathbf{J}^T \mathbf{y} - \phi \mathbf{I})$$

$$\hat{\theta}_{yeni3} = (\mathbf{J}^T \mathbf{J} + \phi \mathbf{I})^{-1} (\mathbf{J}^T \mathbf{y} + \phi \mathbf{M})$$

üç yeni kestirici karşılaştırılmıştır. Önerdiğimiz ilk iki kestirici üçüncü kestiriciden elde edilmiştir. Üçüncü kestiricide ön bilgi vektörü $\mathbf{M} = (-1, -1, \dots, -1)$ olarak seçildiğinde, yani, kısıt bölgesi, merkezi $(-1, -1, \dots, -1)$ olan çember olarak alındığında ikinci kestirici elde edilmiş olur. Aynı şekilde ön bilgi vektörü $\mathbf{M} = (0, 0, \dots, 0)$ olarak seçildiğinde ilk kestirici elde edilmiş olur. Kısıt çemberini merkeze taşıyarak elde ettiğimiz ilk kestiricimiz doğrusal modeller için kullanılan ridge regresyona benzemektedir. Bundan dolayı bu kestiriciye “iteratif ridge kestirici” diyebiliriz.

4.1. YÖNTEM

Çalışmada Meyer ve Roth (1972)'un makalelerinde yer alan,

$$f(\mathbf{x}, \boldsymbol{\theta}) = \frac{\theta_1 \theta_3 x_1}{1 + \theta_1 x_1 + \theta_2 x_2} + \varepsilon$$

doğrusal olmayan model kullanıldı. İki değişkenli \mathbf{x} matrisi çok değişkenli normal dağılımdan üretildi. \mathbf{x} matrisinin sütunları arasındaki ilişki katsayısı, sırasıyla, $\rho = 0.80 - 0.85 - 0.90 - 0.99 - 0.999$ olarak seçildi. Hatalar, $\varepsilon \sim N(0, \sigma)$

dağılımına uygun bir şekilde üretildi. Hata varyansı, sırasıyla, $\sigma = 0.1 - 0.5 - 1 - 2 - 3$ olacak şekilde seçildi. Bağımlı değişken, $y = f(\mathbf{x}, \boldsymbol{\theta}) + \varepsilon$ eşitliğinden üretildi. Gözlem sayısı $n = 10 - 20 - 50 - 100$ olarak alındı. Her bir benzetim çalışması 2000 tekrar ile yapıldı. 2000 tekrarın sonunda her bir kestiricinin kestirim değerleri, hata kareler ortalaması ve iterasyon sonucu yakınsama oranları hesaplandı. Farklı kombinasyonlar ile elde edilen sonuçlar, aşağıda, tablolar halinde verilmiştir. Tabloları destekleyen parametre dağılım grafikleri de tabloların altında sunulmuştur. Gözlem sayıları sırasıyla 10, 20 ve 50 için, hatanın standart sapması sırasıyla, 0.1, 1 ve 3 için sonuçlar bu bölüm içerisinde, kalanları ise EKLER bölümünde verilmiştir.

Böylece, benzetim çalışmasında yer alacak etkenler,

- İlişki katsayısı $\rho = 0.80, 0.85, 0.90, 0.99, 0.999$
- Hata standart sapması, $\sigma = 0.1, 0.5, 1, 2, 3$
- Gözlem sayısı $n = 10, 20, 50, 100$

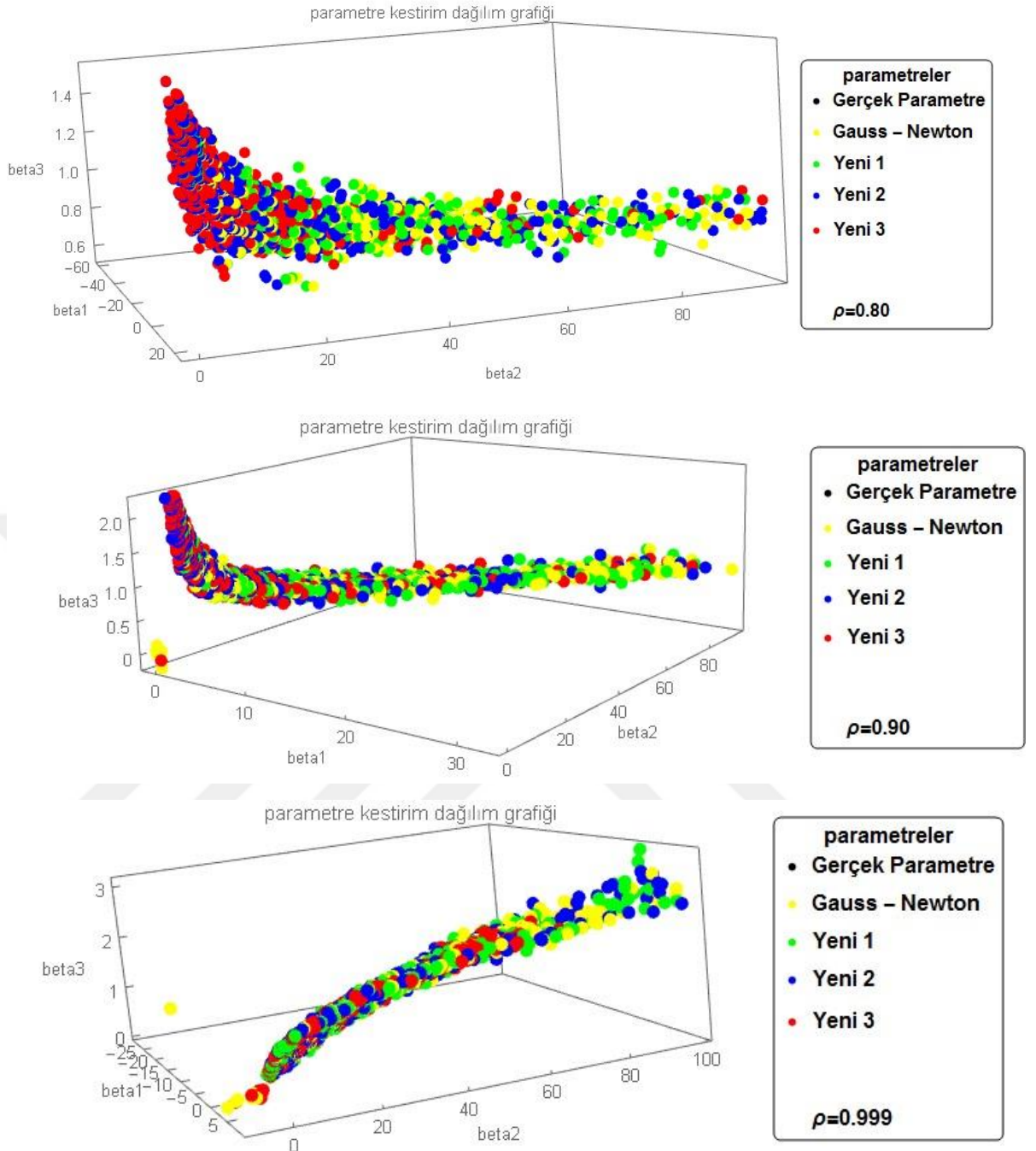
olacak şekilde belirlendi. Bu üç etkenin düzeyleri dikkate alındığında, çalışmada toplam $5 \times 5 \times 4 = 100$ model kullanıldı. Gerçek parametre değerinin $\boldsymbol{\theta} = (3, 12, 1)$ olarak bilinmesinden dolayı parametre kestirimlerinin yanları ile beraber hata kareler ortalamaları hesaplanarak karşılaştırma yapılmıştır. Ayrıca Jakobiyen matrisindeki iç ilişki araştırılmış ve koşul sayıları hesaplanmıştır. Koşul sayılarındaki değişimin parametre kestirimlerine etkisini gösteren grafikler çizilmiştir. Bu grafiklerden bazıları verilmiştir.

4.2. SONUÇLAR

Çizelge 4.1 $n=10$ ve $\sigma=0.1$ için benzetim sonuçları

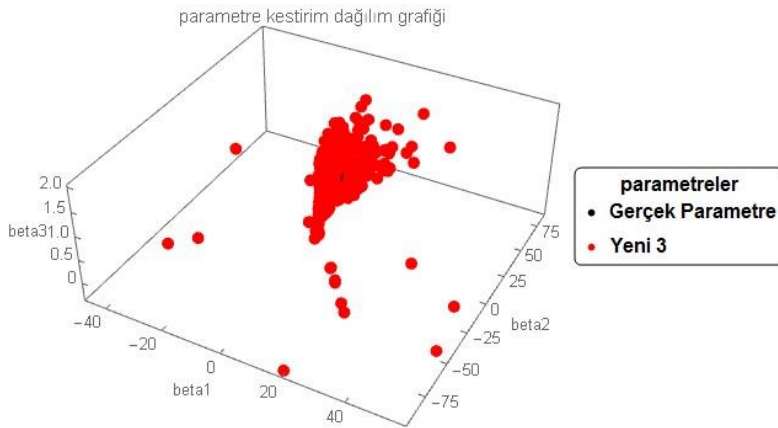
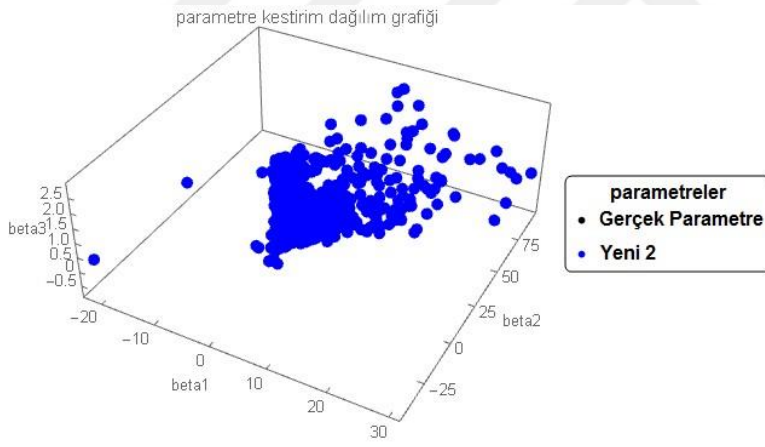
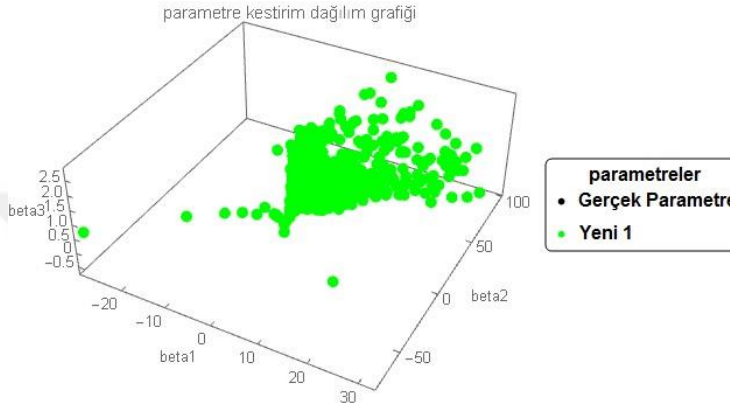
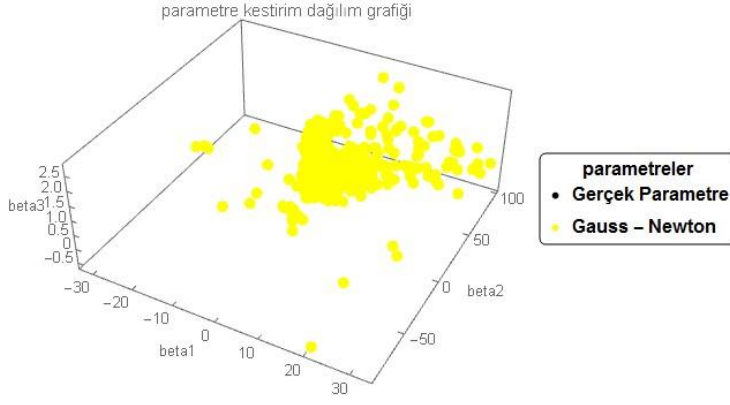
$n=10$	$\sigma=0.1$	$\theta^T=(3,12,1)$				
			$\hat{\theta}_{GN}$	$\hat{\theta}_{yeni1}$	$\hat{\theta}_{yeni2}$	$\hat{\theta}_{yeni3}$
			4.23	3.78	3.45	2.70
			17.57	15.32	13.91	10.69
			0.97	1.03	1.04	1.08
$\rho=0.80$						
	HKO	224.547	206.977	187.5	92.2402	
	Yakınsama oranı	% 69.45	% 86.65	% 88.1	% 84.75	
			5.53	4.79	3.36	3.04
			19.63	17.44	13.13	12.18
			0.96	1.05	1.27	1.29
$\rho=0.90$						
	HKO	303.549	295.891	174.746	144.712	
	Yakınsama oranı	% 54.25	% 68.95	% 78.3	% 74.1	
			3.31	3.20	3.11	3.11
			20.05	16.75	14.80	15.10
			1.18	1.08	1.01	1.03
$\rho=0.999$						
	HKO	285.08	275.725	251.013	200.607	
	Yakınsama oranı	% 60.6	% 80.6	% 84,85	% 75.6	

Yukarıda Çizelge 4.1 ile verilen tabloda, önerdiğimiz kestiriciler, tüm durumlarda en iyi sonucu vermişlerdir. Kurulan 3 farklı modelde de özellikle son sütundaki ön bilgiye dayalı olan kestirici her durumda en küçük hata kareler ortalamasına (HKO) sahiptir. Aşağıdaki Şekil 4.1 de verilen ortak dağılım grafiklerinden önerdiğimiz kestiricilerin yayılımlarının Gauss – Newton kestiricisine göre daha sınırlı olduğu görülebilir.



Şekil 4.1 $n=10$ ve $\sigma=0.1$ için parametre dağılım grafikleri

Parametre yayılımlarını daha açık görebilmek için aşağıdaki Şekil 4.2 de $n=10$, $\sigma=0.1$ ve $\rho=0.85$ için ayrı ayrı dağılımlar verilmiştir. Grafiklerdeki ölçeklere bakıldığında, önerdiğimiz kestiricilerin daha dar ölçekli bölgede yayılım gösterdikleri görülebilir.

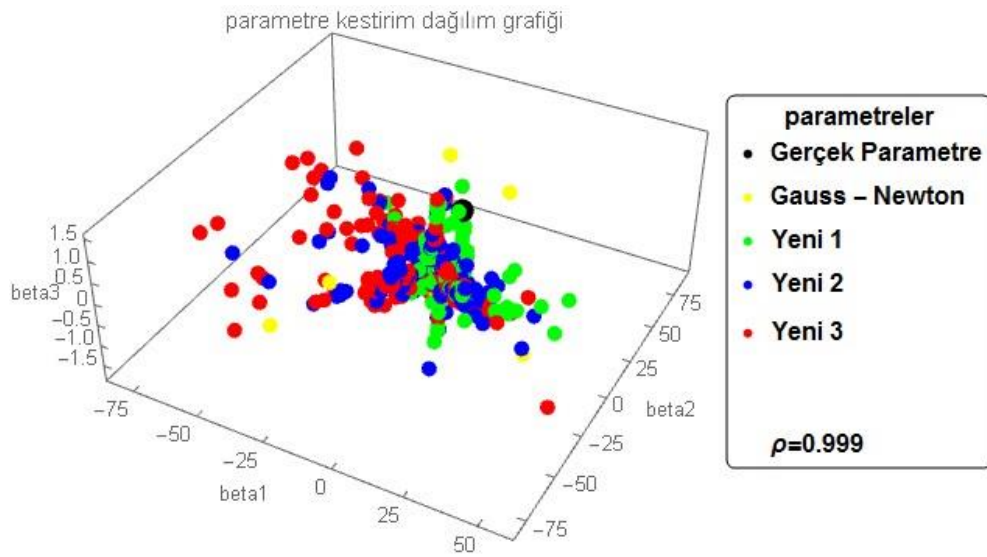
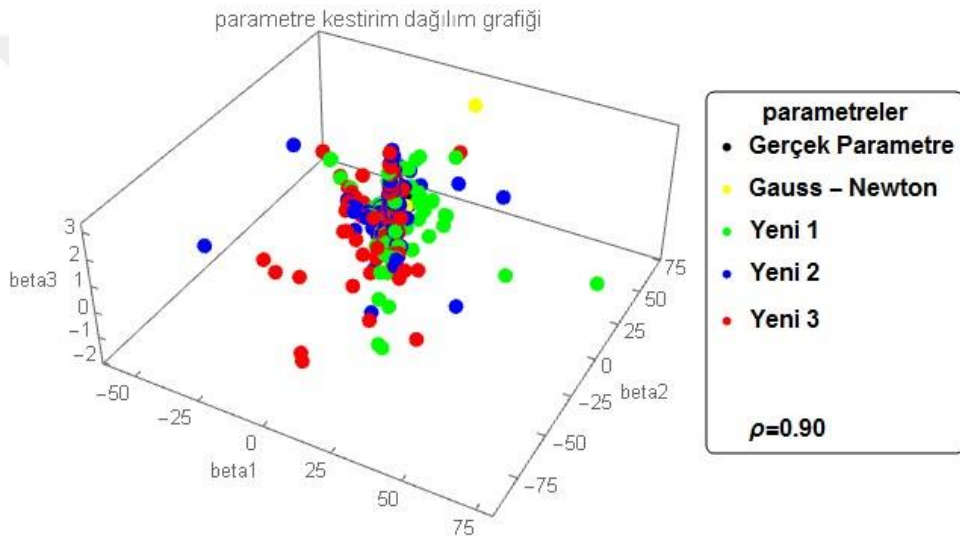
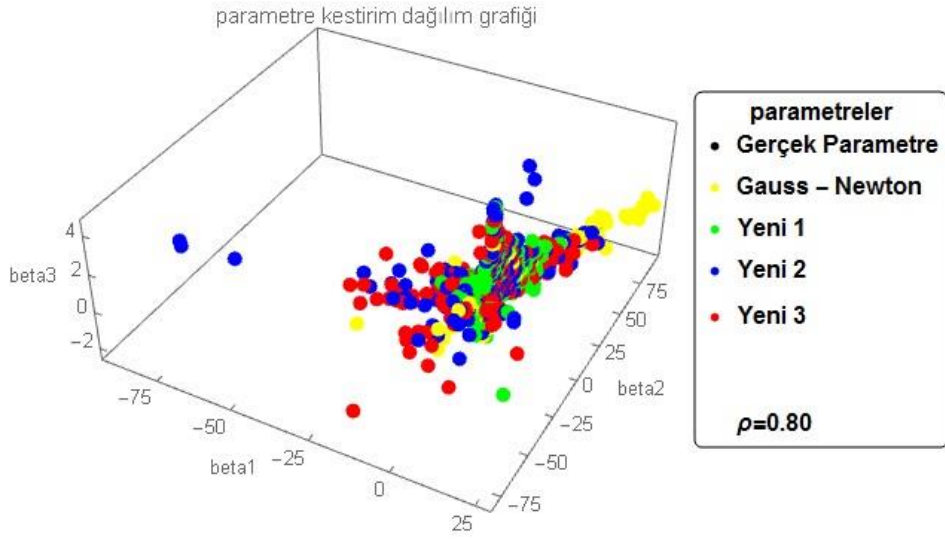


Şekil 4.2 $n = 10$, $\sigma = 0.1$ ve $\rho = 0.85$ için parametre dağılım grafikleri

Çizelge 4.2 $n=10$ ve $\sigma=1$ için benzetim sonuçları

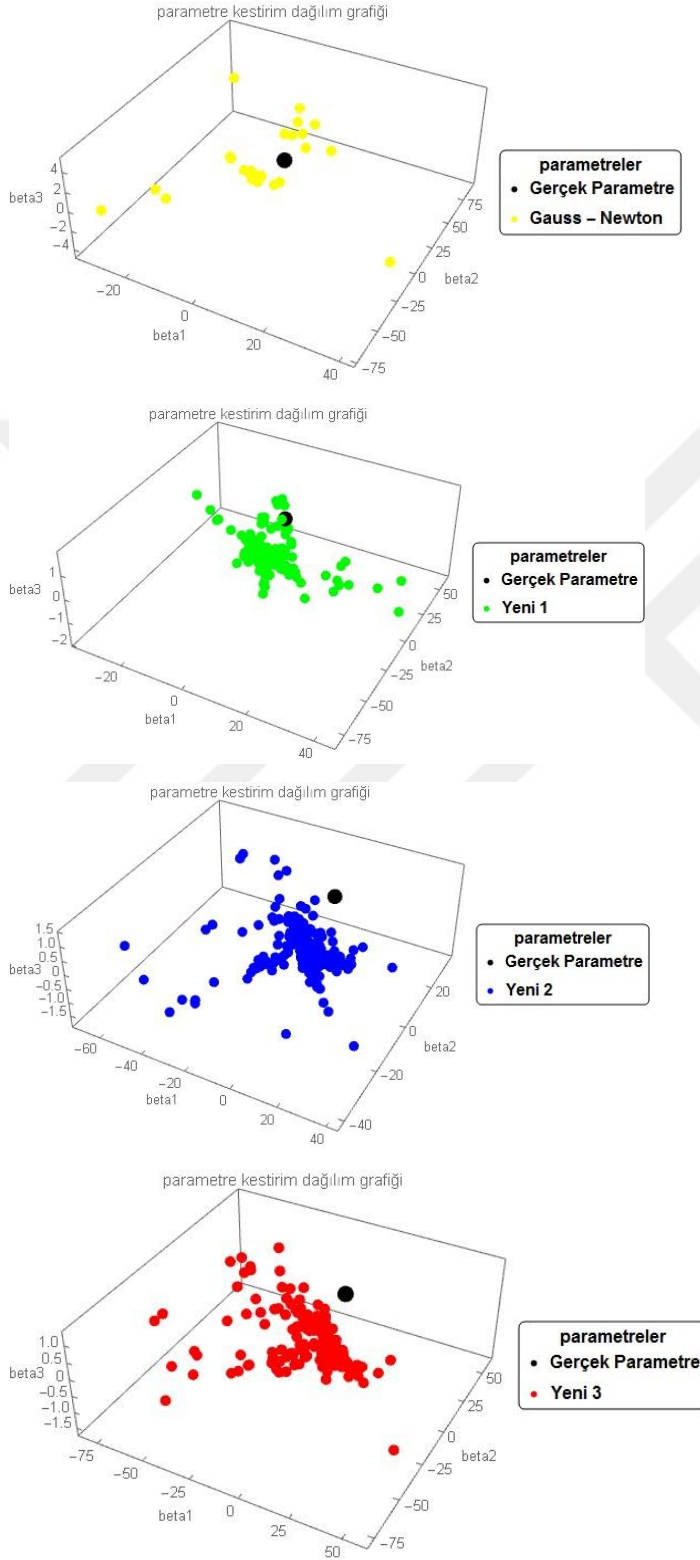
$n=10$	$\sigma=1$	$\theta^T=(3,12,1)$				
			$\hat{\theta}_{GN}$	$\hat{\theta}_{yeni1}$	$\hat{\theta}_{yeni2}$	$\hat{\theta}_{yeni3}$
			3.90	0.94	0.21	0.16
			14.81	5.20	4.38	3.63
			0.88	1.31	0.48	1.05
$\rho=0.80$						
	HKO	505.202	180.382	289.623	241.101	
	Yakınsama oranı	% 13.55	% 33.2	% 40.9	% 34.75	
			1.46	0.41	-0.03	-0.39
			8.73	4.11	2.24	0.37
			0.47	0.50	-1.39	0.24
$\rho=0.90$						
	HKO	461.453	429.066	379.967	369.366	
	Yakınsama oranı	% 0.5	% 9.3	% 17.15	% 16.1	
			-1.77	0.96	-2.12	-6.43
			3.17	1.57	0.03	0.23
			-0.63	0.33	-1.58	-1.54
$\rho=0.999$						
	HKO	1160.82	309.338	508.205	841.343	
	Yakınsama oranı	% 1.45	% 11	% 18.7	% 15.3	

Yukarıdaki Çizelge 4.2 ile verilen tabloda, hatanın standart sapmasını bir önceki tabloya göre 10 katına çıkardığımızda yakınsama oranlarında gözle görülür şekilde düşüş olduğu farkedilebilir. 3 modelde de önerdiğimiz kestiricilerin yakınsama oranları Gauss – Newton kestiricisine göre yaklaşık 10-15 kat daha fazladır. Yine tüm modellerde önerdiğimiz kestiricilerin hata kareler ortalamaları daha küçüktür. 3 model için parametre yayılım grafikleri aşağıdaki Şekil 4.3 ‘de olduğu gibidir.



Şekil 4.3 $n=10$ ve $\sigma=1$ için parametre dağılım grafikleri

Hata varyansının artmasıyla, yukarıda bahsedilen yakınsama oranlarındaki düşüşü aşağıdaki $n = 10$, $\sigma = 1$ ve $\rho = 0.999$ modeli için elde ettiğimiz kestiricilerin ayrı ayrı yayılımlarını Şekil 4.4 'de görebiliriz.

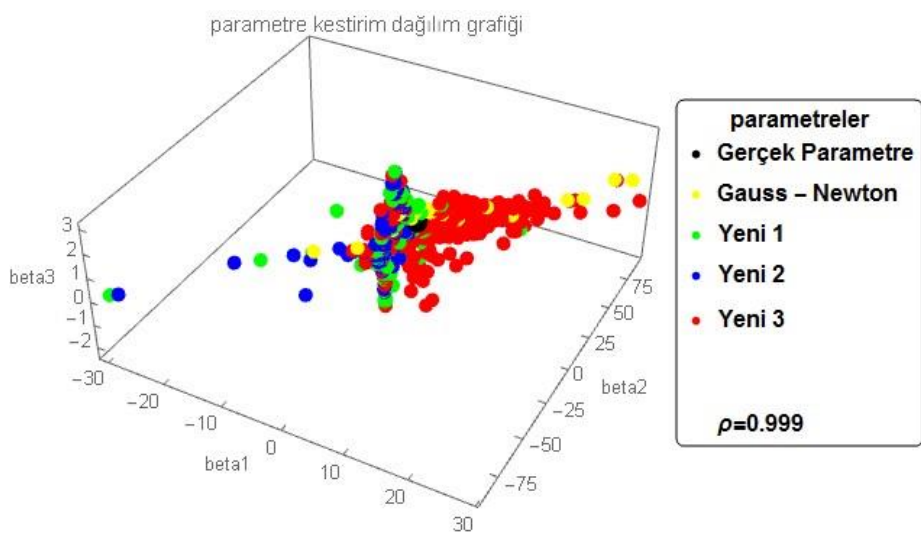
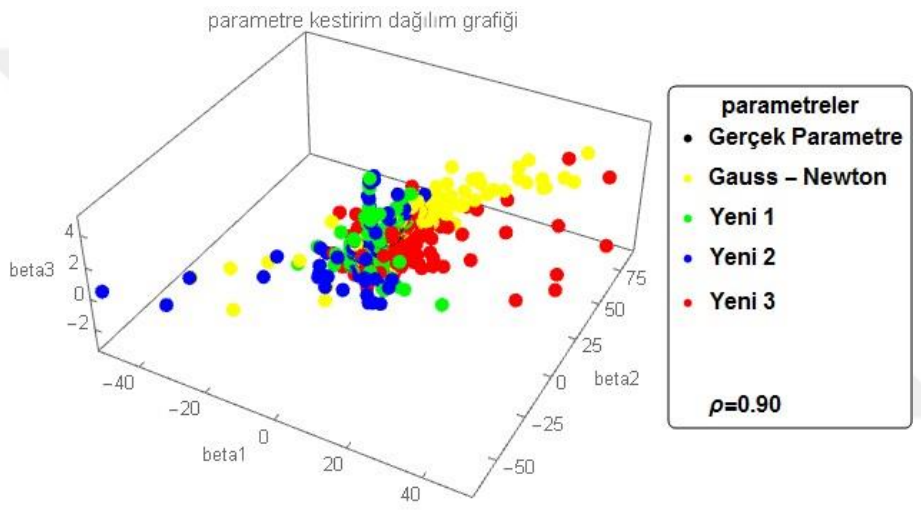
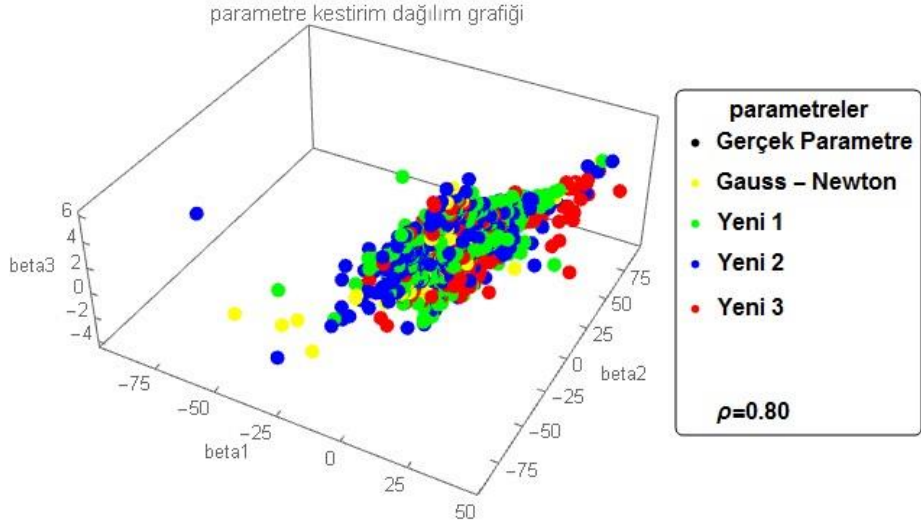


Şekil 4.4 $n = 10$, $\sigma = 1$ ve $\rho = 0.999$ için parametre dağılım grafikleri

Çizelge 4.3 $n=10$ ve $\sigma=3$ için benzetim sonuçları

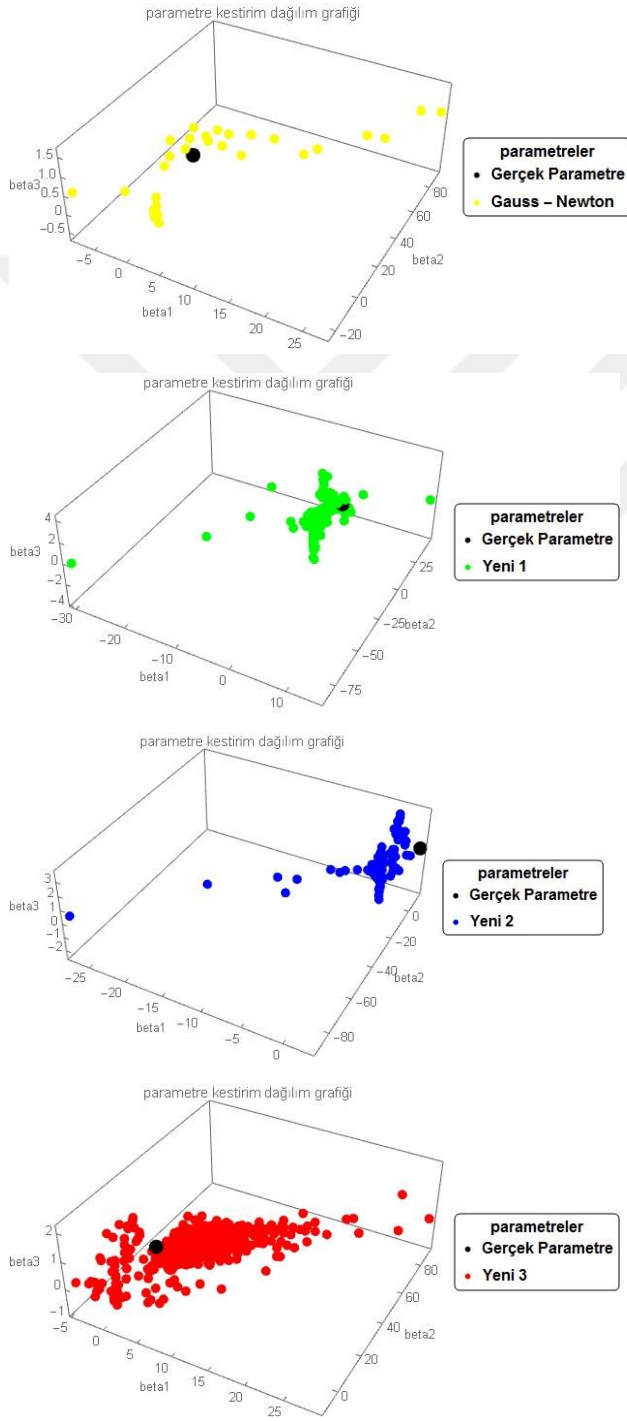
$n=10$	$\sigma=3$	$\theta^T=(3,12,1)$				
			$\hat{\theta}_{GN}$	$\hat{\theta}_{yeni1}$	$\hat{\theta}_{yeni2}$	$\hat{\theta}_{yeni3}$
$\rho=0.80$			1.61	0.18	0.39	3.29
			3.61	0.87	1.66	14.18
			1.24	1.14	1.17	0.36
	HKO	943.482	499.716	546.899	560.686	
	Yakınsama oranı	% 4.75	% 32.15	% 31.45	% 30.3	
$\rho=0.90$			6.46	-0.30	-1.94	4.70
			20.37	5.47	1.21	15.77
			1.41	1.41	1.25	0.67
	HKO	520.644	129.084	302.617	287.797	
	Yakınsama oranı	% 6.85	% 8.85	% 8.8	% 13.05	
$\rho=0.999$			5.55	-0.25	-0.99	5.87
			20.50	2.19	-0.73	21.49
			0.69	0.17	0.68	0.72
	HKO	832.052	230.952	277.983	320.736	
	Yakınsama oranı	% 1.55	% 6.85	% 8.6	% 33.1	

Çizelge 4.3 ile verilen tabloda hatanın standart sapmasının arttırılması ile yakınsama oranlarındaki düşüş göze çarpmaktadır. Bununla birlikte, önerdiğimiz kestiriciler yine en iyi hata kareler ortalamasına sahiptir. Aşağıda verilen Şekil 4.5 ile kestiricilerin parametre uzayındaki yayılımları incelenebilir.



Şekil 4.5 $n = 10$ ve $\sigma = 3$ için parametre dağılım grafikleri

Yukarıdaki Şekil 4.5 ile verilen ortak dağılım grafiklerinden $\rho = 0.999$ için ayrı ayrı kestiricilerin yayılımları aşağıdaki Şekil 4.6 ' da incelenebilir. Şekilden de görüleceği gibi önerdiğimiz kestiricilerin yakınsama oranları daha yüksek ve gerçek parametre civarındaki yayılımları daha fazladır. Şekil 4.6, yukarıdaki Çizelge 4.3 ile verilen tabloda önerdiğimiz kestiricilerin hata kareler ortalamasının daha küçük olduğunu görsel ispatıdır.

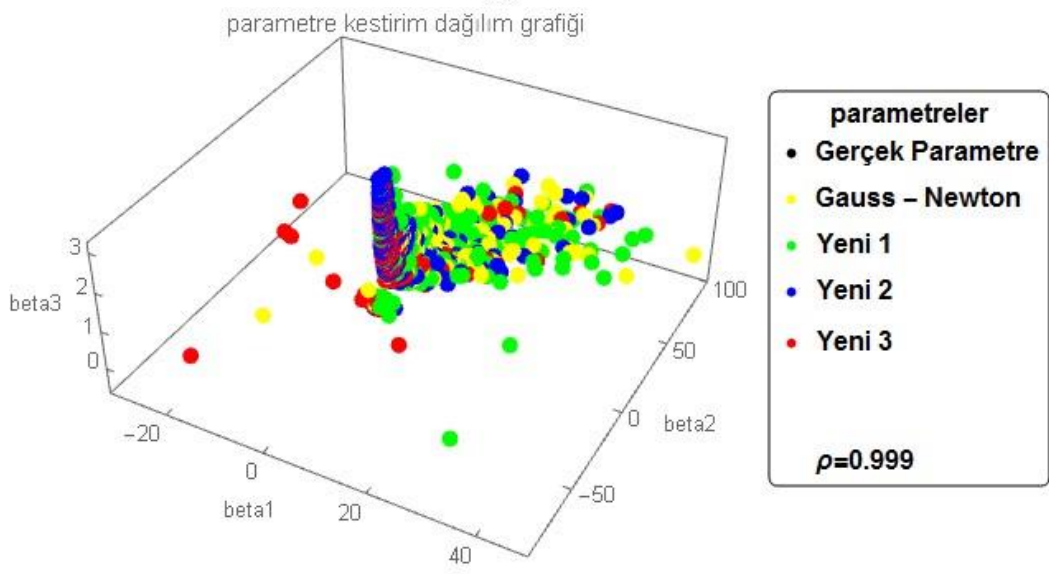
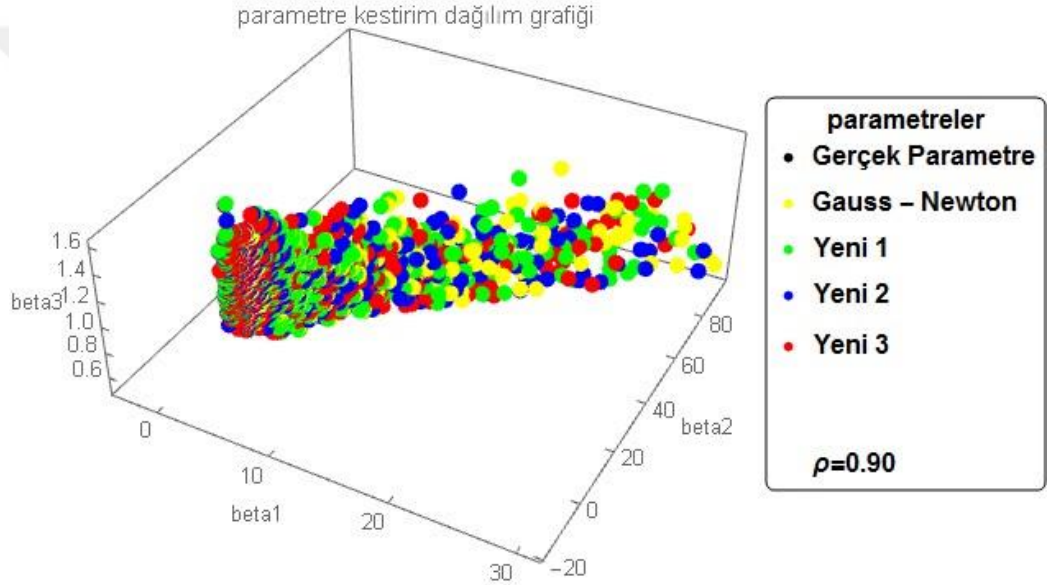
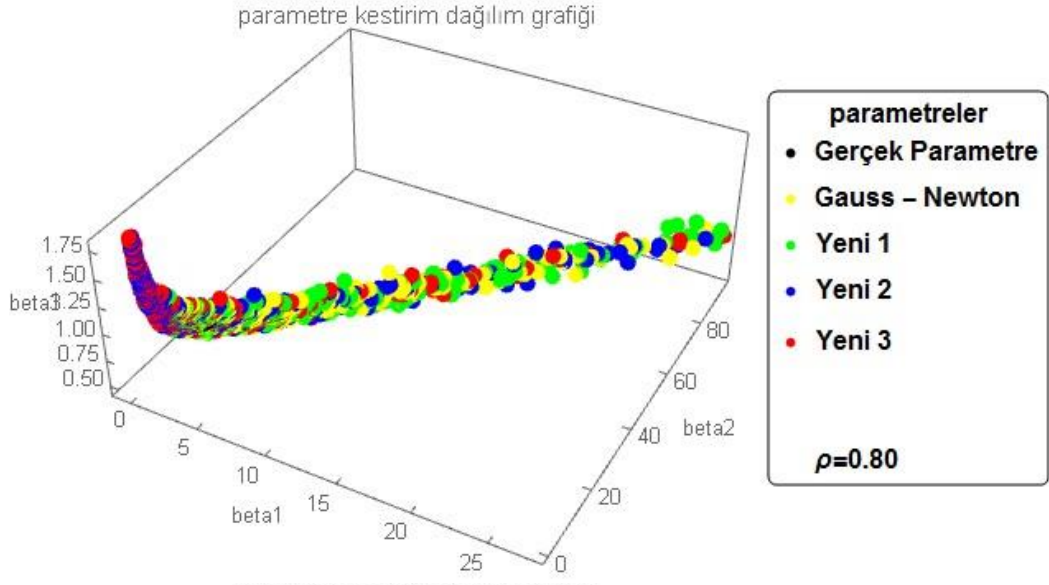


Şekil 4.6 $n = 10$, $\sigma = 3$ ve $\rho = 0.999$ için parametre dağılım grafikleri

Yukarıda $n = 10$ için bazı sonuç tabloları ve grafikler verilmiştir. $n = 20$ ve $n = 50$ için bazı tablo ve grafikler aşağıda verilmiştir. Sonuçlar benzer şekildedir. Her durum için önerdiğimiz kestiricilerin daha küçük hata kareler ortalamasına sahip olduğu aşağıdaki tablo ve grafiklerden incelenebilir.

Çizelge 4.4 $n = 20$ ve $\sigma = 0.1$ için benzetim sonuçları

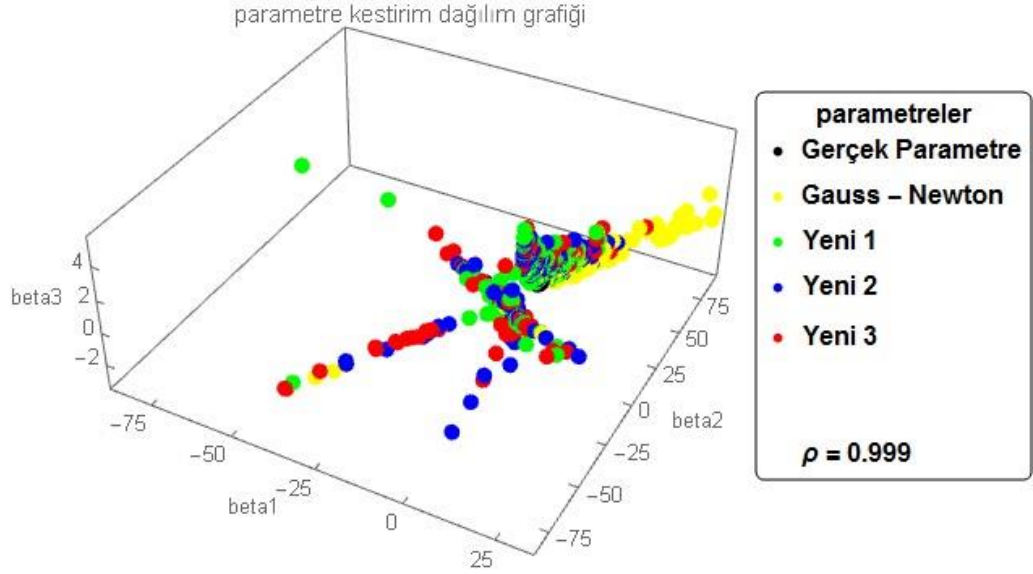
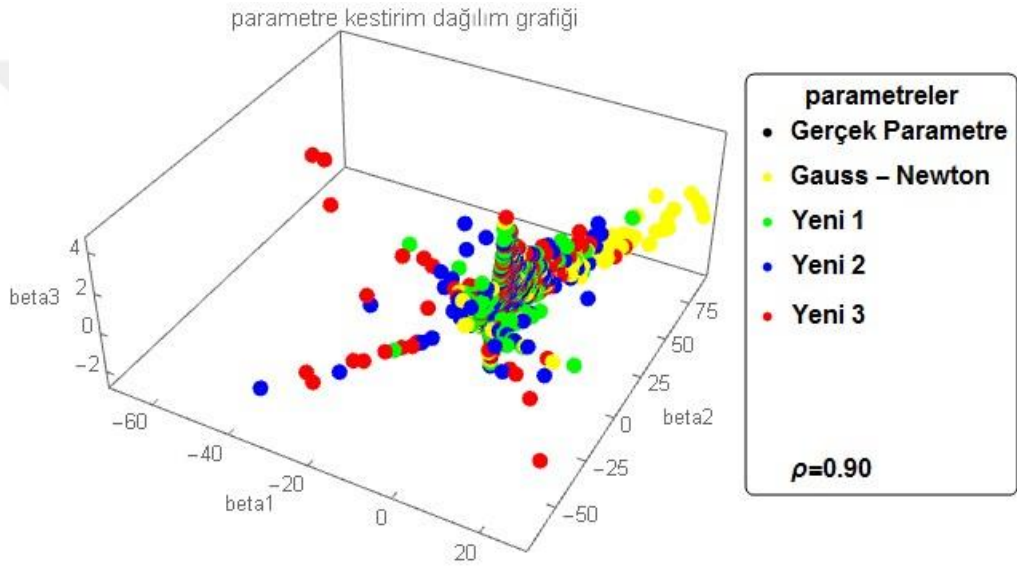
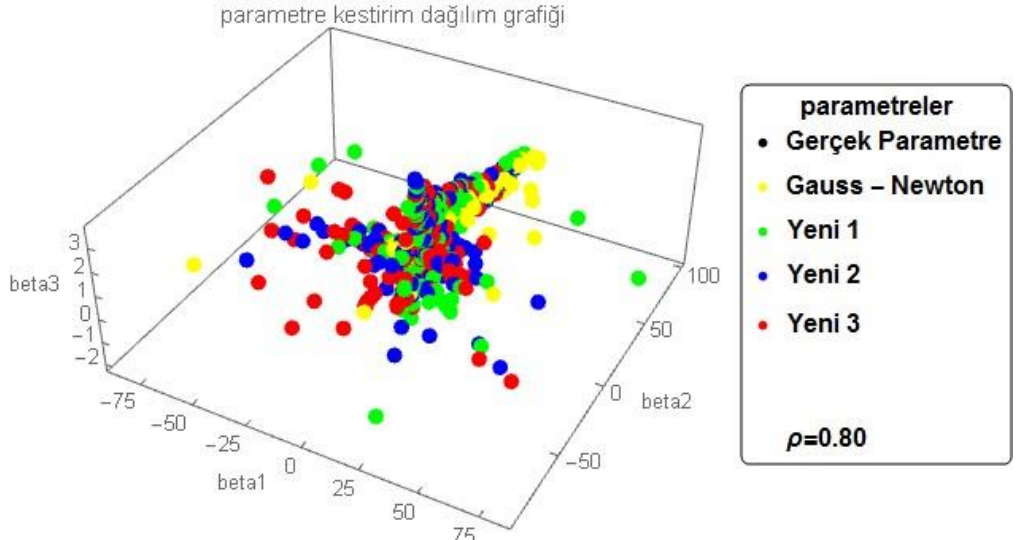
$n = 20$		$\sigma = 0.1$	$\theta^T = (3, 12, 1)$			
			$\hat{\theta}_{GN}$	$\hat{\theta}_{yeni1}$	$\hat{\theta}_{yeni2}$	$\hat{\theta}_{yeni3}$
$\rho = 0.80$			5.53	4.73	2.45	2.07
			20.70	17.99	10.13	8.81
			0.97	1.00	1.20	1.21
	HKO Yakınsama oranı		264.798 % 31.3	293.487 % 48	140.319 % 67.95	110.197 % 67.05
$\rho = 0.90$			4.69	4.16	3.76	3.57
			17.84	16.00	14.61	13.96
			0.99	1.00	1.01	1.01
	HKO Yakınsama oranı		193.885 % 73.75	192.01 % 91.75	150.607 % 88.7	136.308 % 88.2
$\rho = 0.999$			6.43	4.78	1.90	1.53
			22.27	16.36	8.34	7.05
			0.99	1.02	0.19	-0.16
	HKO Yakınsama oranı		384.274 % 15.5	252.609 % 36.8	201.12 % 59.5	229.553 % 50.05



Şekil 4.7 $n = 20$ ve $\sigma = 0.1$ için parametre dağılım grafikleri

Çizelge 4.5 $n = 20$ ve $\sigma = 1$ için benzetim sonuçları

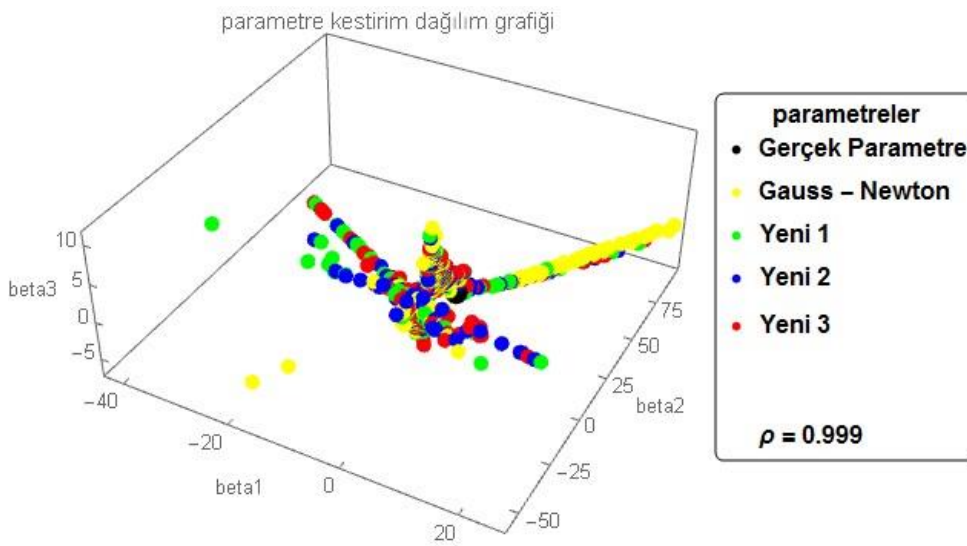
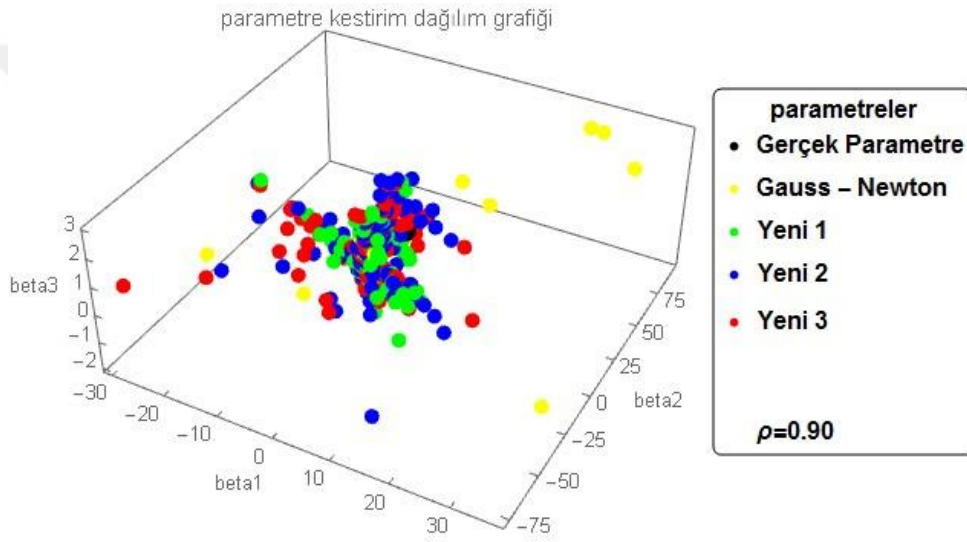
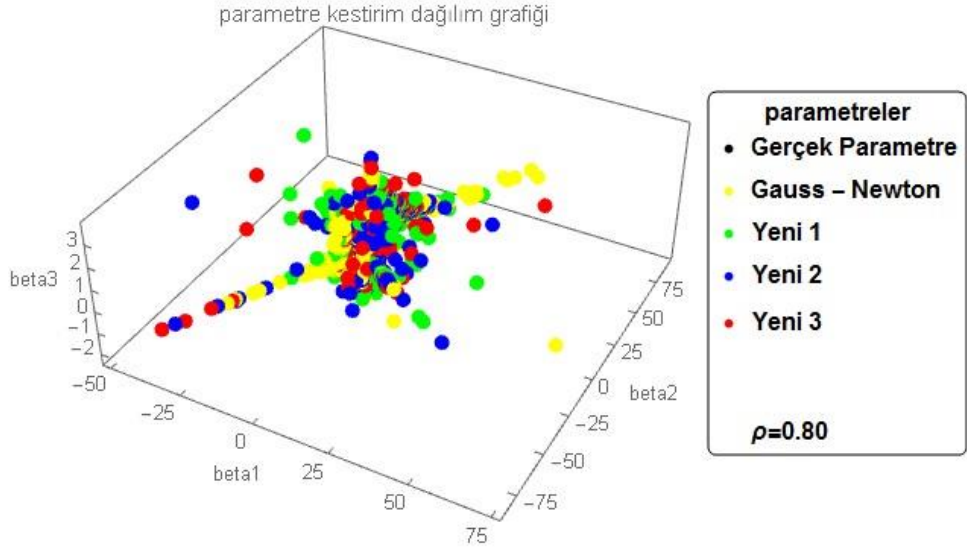
$n = 20$	$\sigma = 1$	$\theta^T = (3,12,1)$				
			$\hat{\theta}_{GN}$	$\hat{\theta}_{yeni1}$	$\hat{\theta}_{yeni2}$	$\hat{\theta}_{yeni3}$
			2.54	0.85	0.67	-0.28
			20.17	11.45	10.01	8.54
			0.69	0.86	0.76	0.66
$\rho=0.80$						
	HKO		770.822	436.468	407.617	412.756
	Yakınsama oranı		% 10.65	% 33.4	% 36.2	% 34.05
			3.99	1.41	1.43	0.94
			16.47	7.68	7.79	7.45
			0.90	0.91	1.04	0.84
$\rho=0.90$						
	HKO		384.663	97.9144	112.022	170.408
	Yakınsama oranı		% 17.8	% 40.5	% 46.95	% 46.5
			6.64	1.13	1.07	-0.09
			21.43	8.22	5.53	1.93
			1.12	1.10	1.01	0.56
$\rho=0.999$						
	HKO		445.586	170.701	244.586	444.028
	Yakınsama oranı		% 13.2	% 20.75	% 19.25	% 15.8



Şekil 4.8 $n = 20$ ve $\sigma = 1$ için parametre dağılım grafikleri

Çizelge 4.6 $n = 20$ ve $\sigma = 3$ için benzetim sonuçları

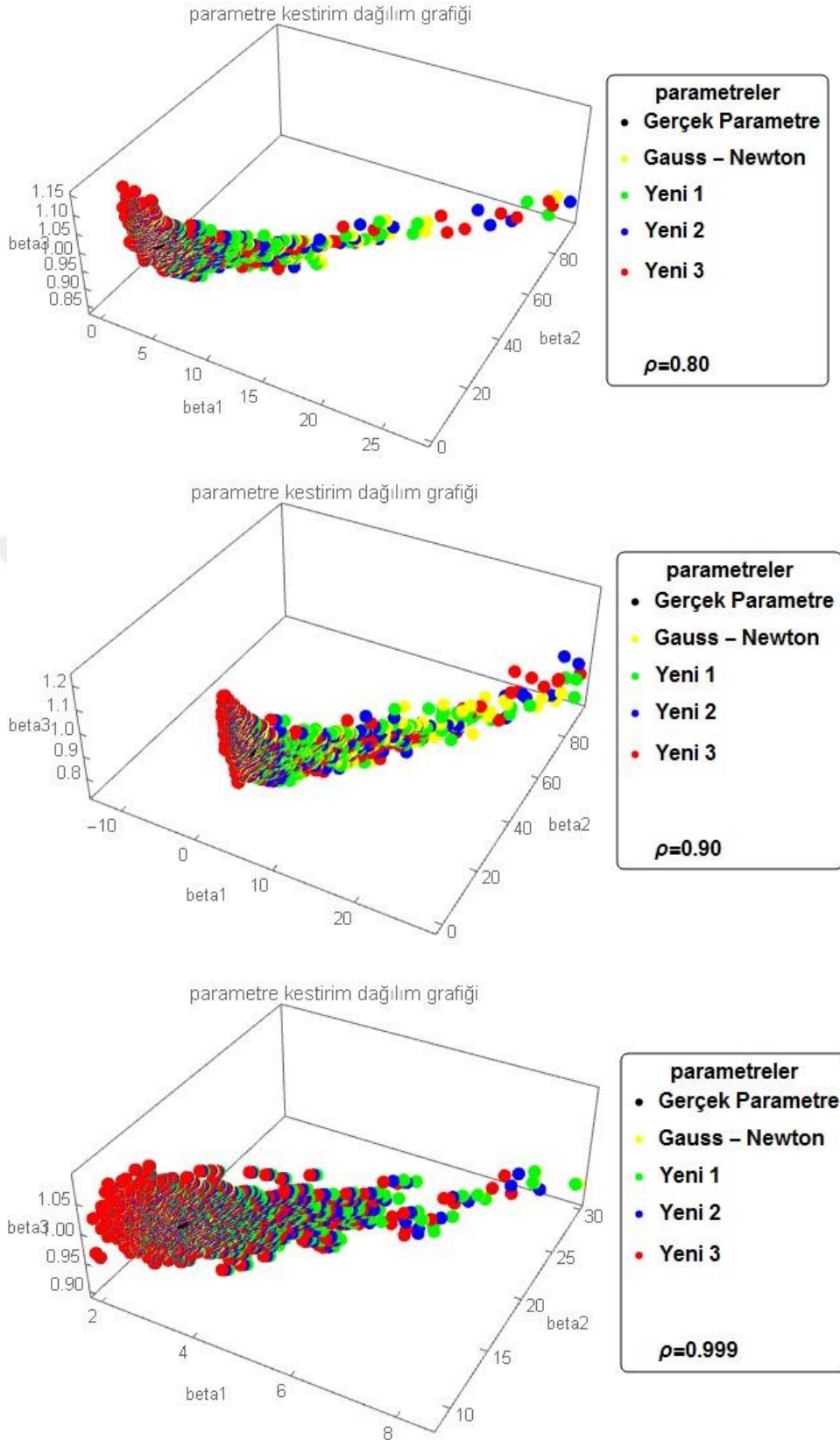
$n = 20$	$\sigma = 3$	$\theta^T = (3,12,1)$				
			$\hat{\theta}_{GN}$	$\hat{\theta}_{yeni1}$	$\hat{\theta}_{yeni2}$	$\hat{\theta}_{yeni3}$
			1.05	1.46	1.77	1.62
			5.86	7.94	7.59	7.36
			0.50	0.79	0.77	0.82
$\rho=0.80$						
	HKO	288.681		127.513	131.298	121.399
	Yakınsama oranı	% 27.25		% 44.2	% 50.85	% 50.3
			6.64	-0.46	-0.52	-1.37
			12.71	-0.05	0.63	0.08
			1.01	0.23	0.81	0.32
$\rho=0.90$						
	HKO	1835.36		222.624	265.415	314.564
	Yakınsama oranı	% 0.85		% 7.45	% 12.4	% 12.9
			2.86	-0.71	0.18	-0.01
			12.04	4.92	4.13	4.08
			1.50	1.53	1.66	2.49
$\rho=0.999$						
	HKO	558.209		249.797	217.953	283.5
	Yakınsama oranı	% 17.5		% 14	% 15.2	% 15.3



Şekil 4.9 $n = 20$ ve $\sigma = 3$ için parametre dağılım grafikleri

Çizelge 4.7 $n = 50$ ve $\sigma = 0.1$ için benzetim sonuçları

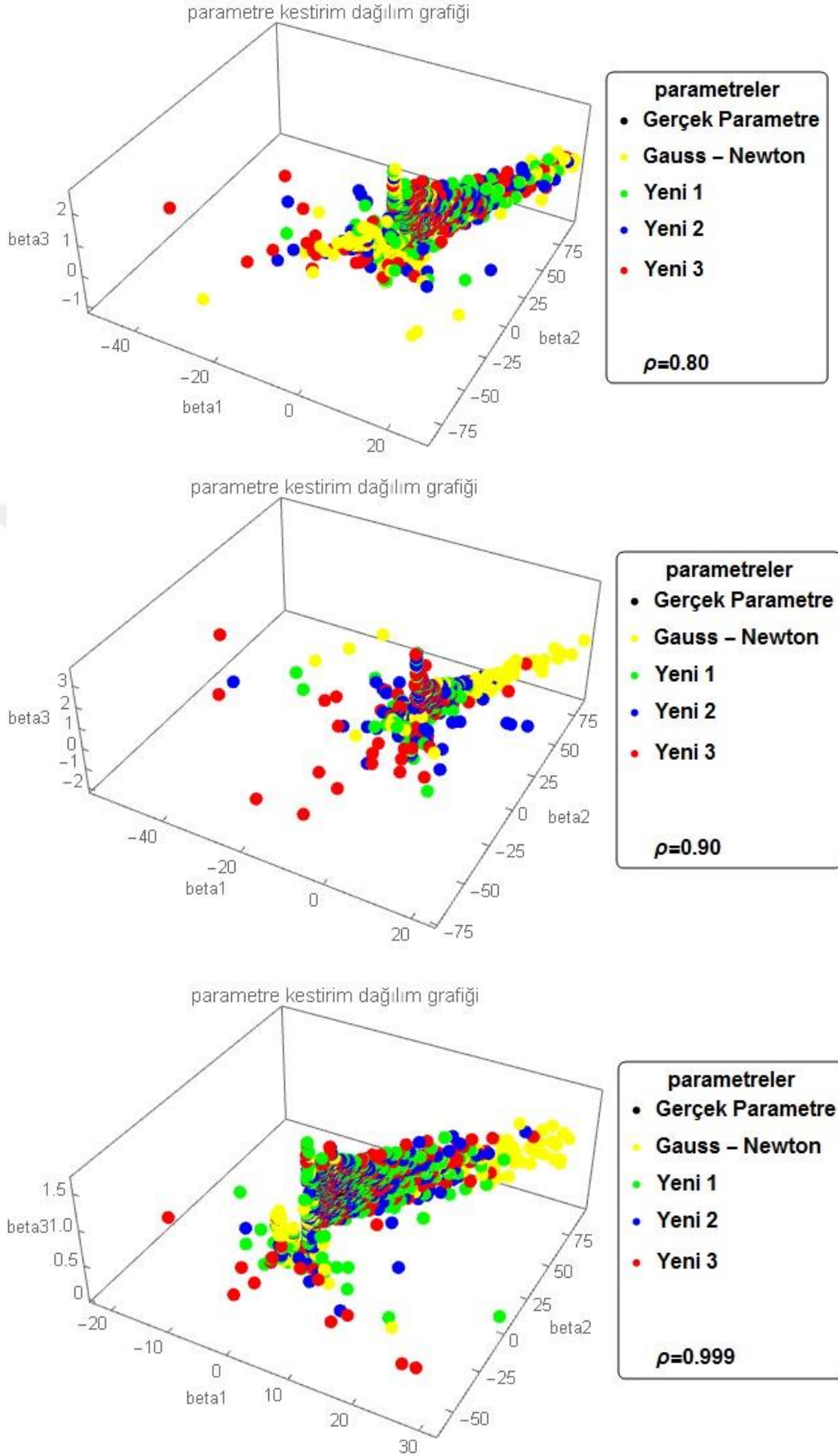
$n = 50$		$\sigma = 0.1$	$\theta^T = (3,12,1)$			
		$\hat{\theta}_{GN}$	$\hat{\theta}_{yeni1}$	$\hat{\theta}_{yeni2}$	$\hat{\theta}_{yeni3}$	
$\rho=0.80$		3.60	3.51	3.36	3.28	
		13.96	13.67	13.18	12.93	
		0.99	0.99	1.00	1.00	
	HKO	41.1575	45.033	38.3624	39.581	
	Yakınsama oranı	% 93.85	% 99.4	% 98.85	% 98.7	
$\rho=0.90$		3.90	3.72	3.46	3.36	
		14.94	14.33	13.50	13.16	
		0.98	1.00	1.00	1.01	
	HKO	91.4303	89.6685	67.9357	67.6305	
	Yakınsama oranı	% 88.1	% 97.45	% 96.05	% 95.75	
$\rho=0.999$		3.14	3.13	3.11	3.09	
		12.46	12.43	12.36	12.29	
		0.99	0.99	0.99	0.99	
	HKO	5.26434	5.52366	5.11245	4.76695	
	Yakınsama oranı	% 99.2	% 100	% 100	% 100	



Şekil 4.10 $n = 50$ ve $\sigma = 0.1$ için parametre dağılım grafikleri

Çizelge 4.8 $n = 50$ ve $\sigma = 1$ için benzetim sonuçları

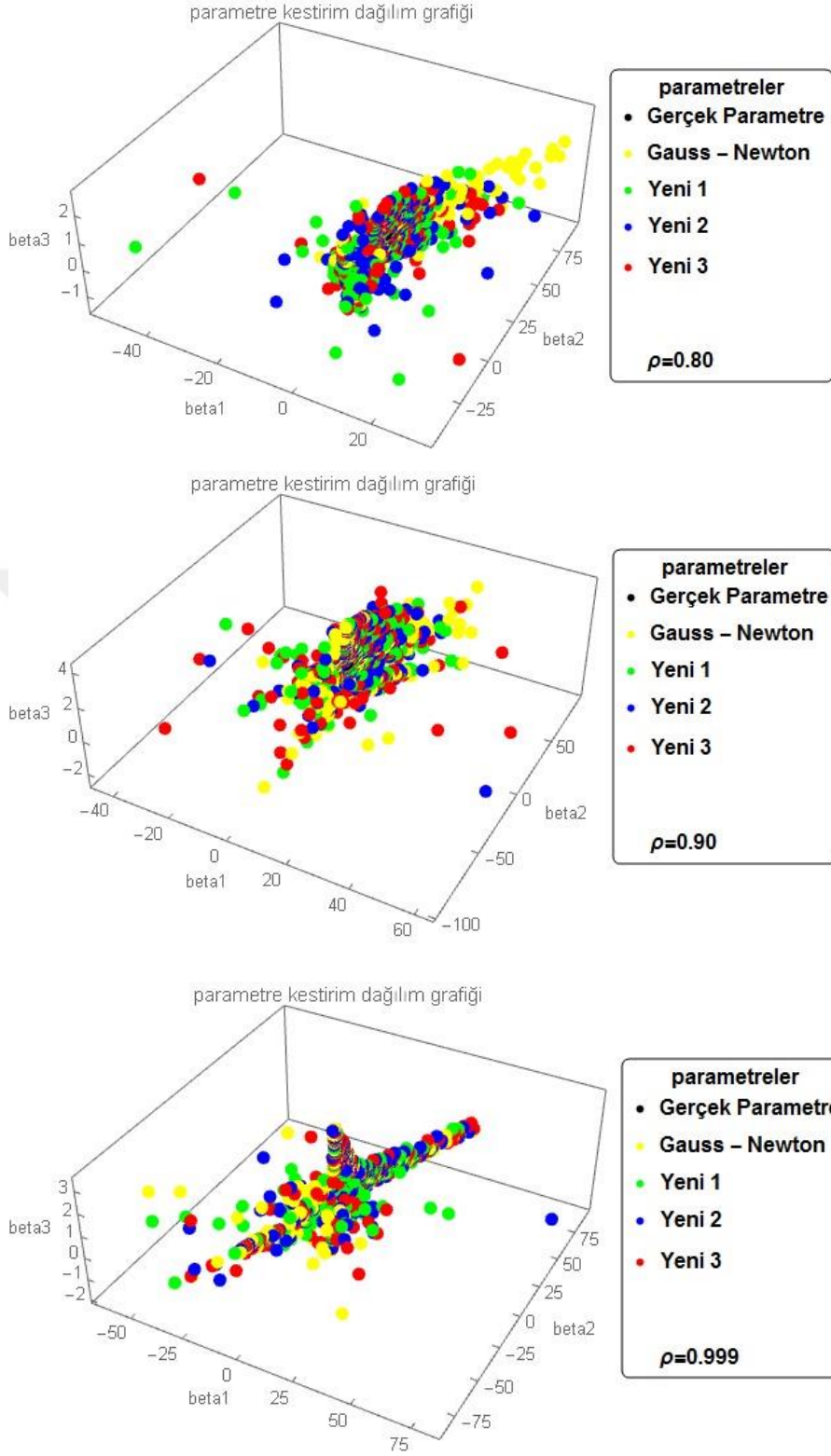
$n = 50$	$\sigma = 1$	$\theta^T = (3,12,1)$				
			$\hat{\theta}_{GN}$	$\hat{\theta}_{yeni1}$	$\hat{\theta}_{yeni2}$	$\hat{\theta}_{yeni3}$
			3.37	3.35	3.34	2.96
			12.63	13.06	13.18	12.22
			0.55	0.85	0.86	0.78
$\rho=0.80$						
	HKO		389.839	150.552	153.028	168.595
	Yakınsama					
	oranı		% 48.8	% 72.25	% 67.5	% 64.85
			3.91	0.90	1.22	0.92
			16.33	5.55	6.12	5.33
			0.69	0.95	1.02	0.94
$\rho=0.90$						
	HKO		386.034	94.3507	86.871	130.259
	Yakınsama					
	oranı		% 12.45	% 30	% 40.7	% 39.4
			5.03	3.48	3.13	3.04
			18.28	13.33	12.33	11.80
			0.92	0.83	0.99	1.02
$\rho=0.999$						
	HKO		334.818	131.017	110.667	136.414
	Yakınsama					
	oranı		% 51.4	% 67.4	% 67.55	% 68.6



Şekil 4.11 $n = 50$ ve $\sigma = 1$ için parametre dağılım grafikleri

Çizelge 4.9 $n = 50$ ve $\sigma = 3$ için benzetim sonuçları

$n = 50$	$\sigma = 3$	$\theta^T = (3,12,1)$				
			$\hat{\theta}_{GN}$	$\hat{\theta}_{yeni1}$	$\hat{\theta}_{yeni2}$	$\hat{\theta}_{yeni3}$
			4.30	1.90	2.12	2.11
			16.07	8.08	8.81	8.85
			0.80	0.53	0.60	0.62
$\rho=0.80$						
	HKO		320.791	127.201	99.4486	97.2231
	Yakınsama oranı		% 25.8	% 42.85	% 47.15	% 47.25
			3.53	3.06	2.97	2.74
			13.69	12.61	12.10	11.43
			0.98	0.89	0.88	0.86
$\rho=0.90$						
	HKO		225.213	145.374	130.342	181.105
	Yakınsama oranı		% 52.95	% 64.45	% 64.6	% 60.5
			1.36	1.67	1.74	1.61
			6.90	7.63	7.85	7.27
			0.69	0.71	0.77	0.80
$\rho=0.999$						
	HKO		353.576	292.116	289.854	313.883
	Yakınsama oranı		% 53.8	% 62.25	% 62.6	% 64.15



Şekil 4.12 $n = 50$ ve $\sigma = 3$ için parametre dağılım grafikleri

Yukarıda 9 tabloda herbirinde 3 model olmakla birlikte toplam 27 model sonuçları ve bu sonuçlara ilişkin grafikler verilmiştir. 27 model için önerdiğimiz kestiriciler en iyi sonucu vermiştir.

4.3. SONUÇLARIN YORUMLANMASI

Doğrusal olmayan modeldeki parametre kestirimi için önerdiğimiz yeni yaklaşımın tutarlılığını ve daha küçük hata kareler ortalamasına sahip olduğunu göstermek için yapılan benzetim çalışmasında 100 farklı model kurulmuştur. Kurulan bütün modellerde önerdiğimiz yaklaşımdan elde edilen kestiricilerimiz daha küçük hata kareler ortalamasına sahip çıkmıştır.

Doğrusal olmayan regresyonda parametre kestiriminde karşılaşılan en büyük problemlerden birisi yakınsama problemidir. Önerdiğimiz yaklaşım büyük ölçüde yakınsama problemine de çözüm olmuştur. Sonuçlar incelendiğinde, önerdiğimiz kestiricilerin yakınsama oranlarındaki artış görülmüştür.

Veri sayısının büyük, hata varyansının küçük olduğu durumda önerdiğimiz kestiriciler, Gauss – Newton yöntemine çok yakın sonuçlar vermesine rağmen yine de en iyi hata kareler ortalamasına sahip çıkmıştır. Hata varyansı arttıkça yakınsamalardaki düşüş gözlenmiş fakat kestiricilerimizin tutarlılığı hep iyi çıkmıştır.

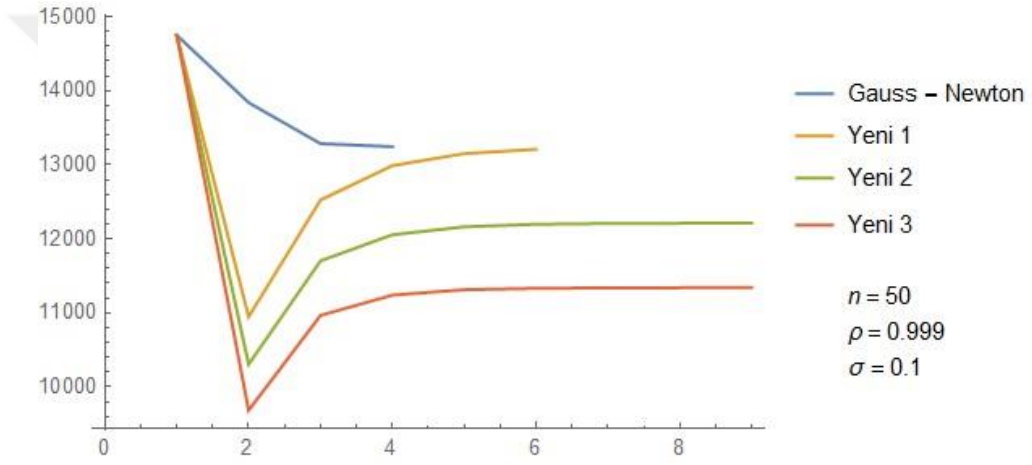
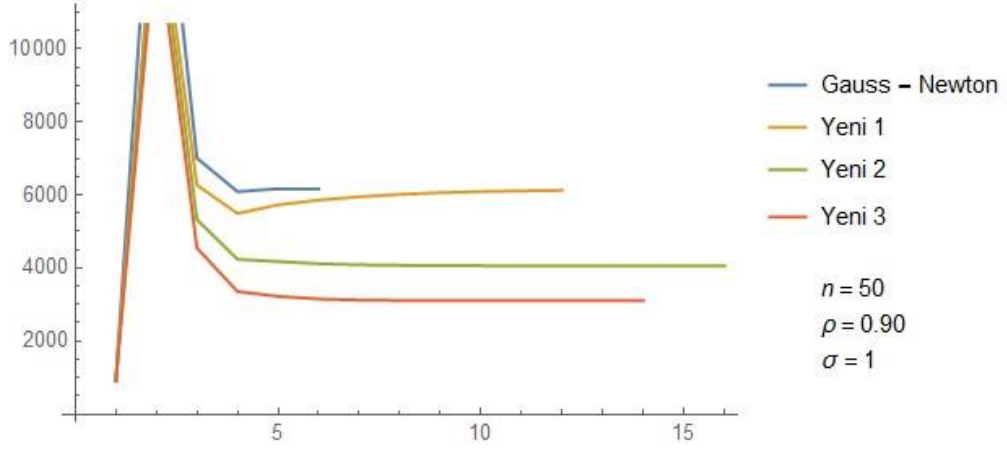
Aynı veriler ve aynı koşullarda, pek çok durumda, Gauss – Newton yöntemi yakınsama sağlamazken önerdiğimiz kestiriciler gerçek parametreye çok yakın noktalara yakınsamıştır. Bu duruma örnek olarak, aşağıda verdiğimiz, benzetim çalışmasının farklı iterasyonlarından elde edilen sonuçlar verilebilir:

	Gauss – Newton	Yeni 1	Yeni 2	Yeni 3
Parametre	$\begin{pmatrix} 9000.05 \\ 34238.5 \\ -0.843656 \end{pmatrix}$	$\begin{pmatrix} 3.26471 \\ 12.8348 \\ 0.925666 \end{pmatrix}$	$\begin{pmatrix} 3.26225 \\ 12.827 \\ 0.926276 \end{pmatrix}$	$\begin{pmatrix} 3.2598 \\ 12.8192 \\ 0.926887 \end{pmatrix}$
Varyans	35691.9	0.602527	0.611379	0.612147

	Gauss – Newton	Yeni 1	Yeni 2	Yeni 3
Parametre	$\begin{pmatrix} 14787.6 \\ 44103.8 \\ 0.710849 \end{pmatrix}$	$\begin{pmatrix} 10.6644 \\ 36.0848 \\ 0.756237 \end{pmatrix}$	$\begin{pmatrix} 3.22412 \\ 13.7378 \\ 0.818367 \end{pmatrix}$	$\begin{pmatrix} 2.37044 \\ 11.1316 \\ 0.834984 \end{pmatrix}$
Varyans	3.50317×10^6	16.9563	114.722	70.4343

Şekil 4.13 Farklı iterasyonlardan sonuçlar

Bununla birlikte Jakobiyen matrisindeki iç ilişkiyi gösteren koşul sayılarına ilişkin grafikler ve kestirim değerleri aşağıda verilmiştir.



Şekil 4.14 Jakobiyen matrisinden hesaplanan koşul sayılarının değişim grafiği

Yukarıdaki Şekil 4.14' de yatay eksen iterasyon sayısını dikey eksen ise koşul sayısını göstermektedir. İteratif olarak, Jakobiyen matrisinin özdeğerleri ile hesaplanan koşul sayılarındaki düşüş, önerdiğimiz kestiricilerde daha fazla olduğu gözlenebilir. Bu ise önerdiğimiz kestiricilerin iç ilişki problemine duyarlılıklarını göstermektedir.

5. UYGULAMA

Meyer ve Roth (1972)'nin makalelerinde kullandıkları veriler aşağıda Çizelge 5.1 ile verilmiştir.

Çizelge 5.1 Örnek verileri

	x_1	x_2	y		x_1	x_2	y
1	0	0	40	13	4.6	2.8	0.3
2	0.6	0.4	10	14	3.2	3	0.22
3	0.6	1	5	15	1.6	3.2	0.2
4	1.4	1.4	2.5	16	4.2	3.4	0.1
5	2.6	1.4	2.5	17	2	3.8	0.05
6	3.2	1.6	2	18	3.2	3.8	0.07
7	0.8	2	1	19	2.8	4.2	0.03
8	1.6	2.2	0.7	20	4.2	4.2	0.03
9	2.6	2.2	0.8	21	5.4	4.4	0.03
10	4	2.2	0.7	22	5.6	4.8	0.02
11	1.2	2.6	0.4	23	3.2	5	0.01
12	2	2.6	0.4				

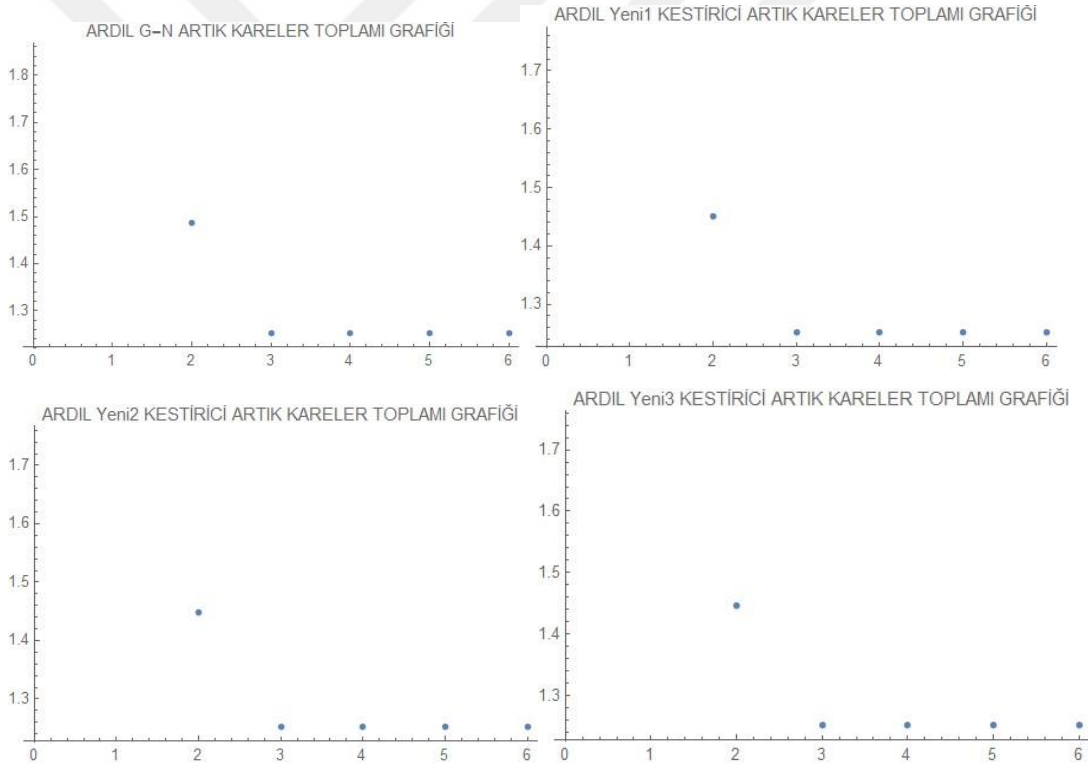
Bu veriler ile $y = \theta_3(e^{-x_1\theta_1} + e^{-x_2\theta_2})$ modelindeki parametreler kestirilmeye çalışılmıştır.

Çalışmada başlangıç parametresi olarak, $\theta^0 = (12,1,25)$ seçilmiş ve optimal parametre kestirim değerini $(31.5,1.51,19.9)$ olarak belirlemişlerdir. Yukarıdaki veriler ile verilen model parametreleri, Gauss – Newton ve önerdiğimiz yaklaşım ile aşağıdaki gibi kestirilmiştir:

Çizelge 5.2 Verilen örnek için parametre kestirim değerleri

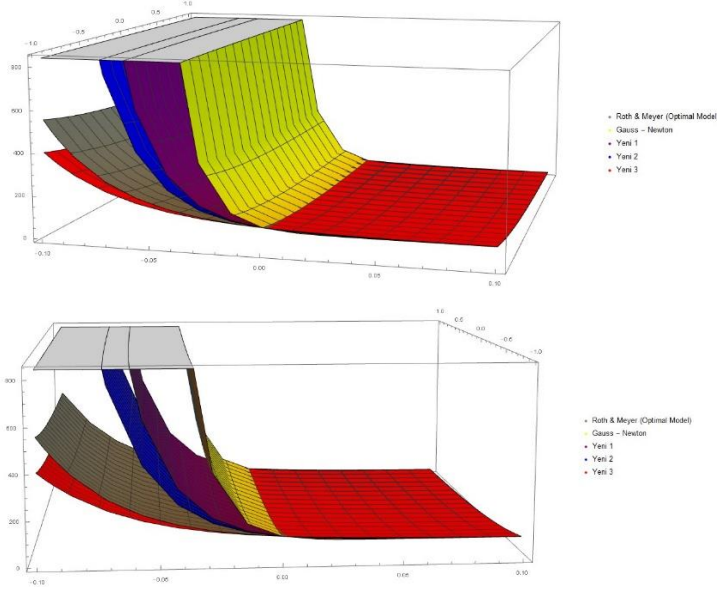
	Gauss –Newton	Yeni 1	Yeni 2	Yeni 3
Parametre	$\begin{pmatrix} 94.8355 \\ 1.50761 \\ 19.9203 \end{pmatrix}$	$\begin{pmatrix} 57.4441 \\ 1.50761 \\ 19.9203 \end{pmatrix}$	$\begin{pmatrix} 49.4315 \\ 1.50761 \\ 19.9203 \end{pmatrix}$	$\begin{pmatrix} 27.5012 \\ 1.50759 \\ 19.9201 \end{pmatrix}$
HKO	4011.4	673.115	1087.83	370.001

Hata kareler ortalaması yukarıda verilen Roth ve Meyer (1972) ‘in optimal parametre değeri baz alınarak hesaplanmıştır. Dikkat edilecek olursa ikinci ve üçüncü parametre kestirimleri aynıdır. Farklılık birinci parametre kestirimindedir. Gauss- Newton kestiricisinin varyansı biraz daha küçük olmasına rağmen birinci parametre kestiriminden dolayı yanı çok büyüktür. Yukarıdaki 4 kestiricinin de yakınsama hızları aynıdır. Bunun için aşağıda verilen amaç fonksiyonlarının grafikleri incelenebilir:



Şekil 5.1 Ardıl artık kareler toplamı grafikleri

Önerdiğimiz yaklaşımdan üretilen kestiriciler Gauss – Newton yöntemine göre varyansı biraz daha büyük fakat yanı çok daha küçüktür. Bununla birlikte yandaki düşüşün büyüklüğü varyanstaki artışın büyüklüğünden çok daha küçük olduğu görülmüştür. Dolayısıyla hata kareler ortalaması daha küçük ve daha tutarlı kestirim değerleri elde edilmiştir. Aşağıdaki şekilden kestirim modelleri incelenebilir.



Şekil 5.2 Kestirim modellerinin farklı açılardan grafikleri

Grafikten de görüleceği üzere gri renkte olan Roth ve Meyer (1972)' nin çalışmalarında belirledikleri optimal model kestirimine en yakın model kestirimleri önerdiğimiz yaklaşımdan elde edilen model kestirimleridir. En yakın olan kestirim ise kırmızı renkli olan Yeni 3 ile adlandırdığımız ön bilgi ile kestirilen modeldir. Roth ve Meyer (1972) 'in optimal modeli ile aralarındaki fark oldukça azdır.

Önerdiğimiz kestiricilerin Gauss – Newton ile karşılaştırılması hata kareler ortalamasının yanı sıra, çapraz geçerlilik kullanılarak yapılmıştır. Bunun için veri ikiye bölünmüş ve herbir kestiriciye ait model için Press değerleri hesaplanmıştır. Güvenilirliği arttırmak için, bu işlem 100 kez tekrarlanmış ve ortalama press değerleri hesaplanmıştır. Aşağıdaki tabloda press değerleri verilmiştir. Gauss – Newton yöntemi ile kestirilen modellerde press değerleri çok büyük çıktığı için farklı ortalama yöntemleri de hesaplanmış ve tabloda verilmiştir. Her durumda önerdiğimiz kestiriciler ile kurulan modellere ait press değerleri en küçük çıkmıştır.

Çizelge 5.3 Kestirilen modellerin press değerleri

	Gauss – Newton	Yeni 1	Yeni 2	Yeni 3
Aritmetik Ortalama (PRESS)	$>10^3$	13.5477	13.6031	13.7027
Medyan (PRESS)	57.7053	2.13266	2.1329	2.134

6. SONUÇ VE TARTIŞMA

Doğrusal olmayan regresyon modellerinde parametre kestirimine alternatif olarak sunulan yeni yaklaşımımız 3. Bölümde teorik olarak incelenmiş ve Gauss – Newton yöntemine göre hata kareler ortalamasının daha küçük olduğu gösterilmiştir. Önerilen yaklaşımın tutarlılığı 4. Bölümde yapılan benzetim çalışması ile gösterilmiştir. Yapılan benzetim çalışması ile oluşturulan 100 farklı durumda önerdiğimiz yaklaşımın yakınsaklık problemlerine büyük ölçüde çözüm olduğu görülmüştür. Her durumda veriyi en iyi modelleyen parametre kestirimleri elde edilmiştir. Elde edilen parametre kestirim değerlerinin hata kareler ortalaması her zaman daha küçük bulunmuştur. Tüm bu sonuçlar çizilen grafiklerle görsel olarak da desteklenmiştir. Özet olarak sunulan aşağıdaki Çizelge 6.1 ile verilen tabloda da, yine, önerdiğimiz yaklaşım ile elde edilen kestiriciler tüm durumlarda farklı bir şekilde daha tutarlı olduklarını göstermişlerdir. 5. Bölümde daha önce kullanılmış bir veri seti ile yeni yaklaşımımızın performansının iyiliği gösterilmiştir.

Benzetim çalışması sırasında, üretilen veri grupları dikkate alındığında, veriler dağılımı yaygın olduğu durumda belirgin bir şekilde önerdiğimiz yaklaşım başarılı olmuştur. Veriler hata varyansı daha küçük seçilerek üretildiğinde ise kestiriciler yaklaşık aynı sonucu verselerde yine de önerdiğimiz yaklaşım daha iyi çıkmıştır.

Çalışma boyunca gözlemlediğimiz kadarıyla, kestiricilerin başarısını en çok etkileyen faktör hata varyansının değişimi olmuştur. Yayılım arttıkça verileri minimum hata ile modellemek zor hale geldiğinden kestiricilerin başarısı burada ortaya çıkmıştır. Modelin genel yapısı, kestiricilerin başarısında en önemli bir diğer faktör olarak gözlemlenmiştir. Model ne kadar karmaşık yapıda ise çoklu iç ilişki o derece artacağından, yine, kestiricilerin iç ilişkiye olan duyarlılıklarının bu noktada devreye girdiği görülmüştür.

KAYNAKLAR

Bates, D. M. & Watts, D. G., 1988. *Nonlinear Regression Analysis and Its Applications*. New York: John Wiley & Sons.

Belsley, D. A., 1991. *Conditioning Diagnostics: Collinearity and Weak Data in Regression*. 1st dü. New York: Wiley.

Belsley, D. A., Kuh, E. & Welsch, R. E., 2004. *Regression Diagnostics Identifying Influential Data and Sources of Collinearity*. New Jersey: John Wiley & Sons.

Crouse, r. H., Jin, C. & Hanumara, R. C., 1995. Unbiased Ridge Estimation with Prior Information and Ridge Trace. *Communications in Statistics - Theory and Methods*, 24(9), pp. 2341-2354.

Davidson, R. & Mackinnon, J. G., 1993. *Estimation and Inference in Econometrics*. 1st dü. New York: Oxford University Press.

Erkoç, A., 2011. *Lineer Olmayan Regresyon Modellerinde Çoklu İç İlişki Problemi*. İstanbul: Marmara Üniversitesi Fen Bilimleri Enstitüsü.

Farrar, D. E. & Glauber, R. R., 1967. Multicollinearity in Regression Analysis: The Problem Revisited. *Review of Economics and Statistics*, Cilt 49, pp. 92-107.

Gruber, M. H. J., 2010. *Regression Estimator A Comparative Study*. 2nd dü. Maryland: The John Hopkins University Press.

Hill, R. C. & Adkins, L. C., 2003. *A Companion to Theoretical Econometrics*. Berlin: Blackwell Publishing.

Hoerl, A. E. & Kennard, R. W., 1970. Ridge Regression : Biased Estimation for Nonorthogonal Problems. *Technometrics*, Cilt 12, pp. 55-67.

Hoerl, A. E., Kennard, R. W. & Baldwin, K. F., 1975. Ridge Regression: Some Simulations. *Communications in Statistics*, Cilt 4, pp. 105-123.

Levenberg, K., 1944. A Method for the Solution of Certain Non-linear Problems in Least Squares. *Quart. Appl. Math.*, Cilt 2, pp. 164-168.

- Magel, R. C. & Hertsgaard, D., 1987. A Collinearity Diagnostic for Nonlinear Regression. *Communications in Statistics - Simulation and Computation*, 16(1), pp. 85-97.
- Marquardt, D. W., 1963. An Algorithm for Least-Squares Estimation of Nonlinear Parameters. *Journal of the Society for Industrial and Applied Mathematics*, 11(2), pp. 431-441.
- Marquardt, D. W., 1970. Generalized Inverses, Ridge Regression, Biased Linear Estimation and Nonlinear Estimation. *Technometrics*, Cilt 12, pp. 591-612.
- Meyer, R. R. & Roth, P. M., 1972. Modified Damped Least Squares: An Algorithm for Non-linear Estimation. *J. Inst. Math. Appl.*, Cilt 9, pp. 218-233.
- Montgomery, D. C., Peck, E. A. & Vining, G. G., 2012. *Introduction to Linear Regression Analysis*. New Jersey: John Wiley & Sons.
- Özkale, M. R. & Kaçiranlar, S., 2007. Comparison of the Unbiased Ridge Estimation to the Other Estimations. *Communications in Statistics - Theory and Methods*, 36(4), pp. 707-723.
- Rawlings, J. O., Pantula, S. G. & Dickey, D. A., 1998. *Applied Regression Analysis: A Research Tool*. 2nd dü. New York: Springer.
- Seber, G. A. F. & Wild, C. J., 2003. *Nonlinear Regression*. New Jersey: John Wiley & Sons.
- Swindel, B. F., 1976. Good Ridge Estimators Based on Prior Information. *Communications in Statistics - Theory and Methods*, 5(11), pp. 1065-1075.
- Theil, H., 1971. *Principles of Econometrics*. New York: John Wiley & Sons.
- Theobald, C. M., 1974. Generalizations of Mean Square Error Applied to Ridge Regression. *Journal of the Royal Statistical Society. Series B*, 36(1), pp. 103-106.
- Vinod, H. D. & Ullah, A., 1981. *Recent Advances in Regression Methods*. New York: Marcel Dekker.

EKLER

EK A : BAZI TEOREM VE İSPATLARI

EK B : BENZETİM ÇALIŞMASI İÇİN DİĞER ÇİZELGE VE GRAFİKLER



EK A 1. BAZI TEOREM VE İSPATLARI

Teorem Ek A.1.

$Y = X\beta + \varepsilon$ doğrusal modelinde, hatalar $\varepsilon \sim N(0, \sigma^2 I)$ normal dağılıma sahip, $\hat{\beta} \sim N(\beta, \sigma^2 (X^T X)^{-1})$ en küçük kareler kestiricisi olsun. $\hat{\beta}$ dan bağımsız ön bilgi vektörü B , $B \sim N(\beta, V)$ dağılımına sahip olsun. V nin tam ranklı kovaryans matrisi olduğunu kabul edelim. I $p \times p$ boyutlu birim matris, C $p \times p$ boyutlu bir matris olmak üzere, konveks kestirici,

$$\beta(C, B) = C\hat{\beta} + (I - C)B$$

olarak tanımlandığında, bu kestiricinin en küçük hata kareler ortalamasına (HKO) sahip olması için en iyi C matrisi $C = V(\sigma^2 (X^T X)^{-1} + V)^{-1}$ olarak seçilmelidir.

İspat

En küçük kareler kestiricisini konveks kestiricide yerine yazalım.

$$\begin{aligned} \beta(C, B) &= C\hat{\beta} + (I - C)B \\ &= C(X^T X)^{-1} X^T Y + (I - C)B \end{aligned} \quad (\text{Ek A.1})$$

Bu kestiricinin hata kareler ortalaması,

$$\begin{aligned} HKO(\beta(C, B)) &= E[(\beta(C, B) - \beta)^T (\beta(C, B) - \beta)] \\ &= E[(C(X^T X)^{-1} X^T (X\beta + \varepsilon) + B - CB - \beta)^T (C(X^T X)^{-1} X^T (X\beta + \varepsilon) + B - CB - \beta)] \\ &= E[(C(X^T X)^{-1} X^T X\beta + C(X^T X)^{-1} X^T \varepsilon + B - CB - \beta)^T \\ &\quad (C(X^T X)^{-1} X^T X\beta + C(X^T X)^{-1} X^T \varepsilon + B - CB - \beta)] \\ &= E[(C\beta + C(X^T X)^{-1} X^T \varepsilon + B - CB - \beta)^T (C\beta + C(X^T X)^{-1} X^T \varepsilon + B - CB - \beta)] \\ &= E[\beta^T C^T C\beta + \beta^T C^T C(X^T X)^{-1} X^T \varepsilon + \beta^T C^T B - \beta^T C^T CB \\ &\quad - \beta^T C^T \beta + \varepsilon^T X(X^T X)^{-1} C^T C\beta + \varepsilon^T X(X^T X)^{-1} C^T C(X^T X)^{-1} X^T \varepsilon \\ &\quad + \varepsilon^T X(X^T X)^{-1} C^T B - \varepsilon^T X(X^T X)^{-1} C^T CB - \varepsilon^T X(X^T X)^{-1} C^T \beta \\ &\quad + B^T C\beta + B^T C(X^T X)^{-1} X^T \varepsilon + B^T B - B^T CB - B^T \beta \\ &\quad - B^T C^T C\beta - B^T C^T C(X^T X)^{-1} X^T \varepsilon - B^T C^T B + B^T C^T CB + B^T C^T \beta \\ &\quad - \beta^T C\beta - \beta^T C(X^T X)^{-1} X^T \varepsilon - \beta^T B + \beta^T CB + \beta^T \beta] \end{aligned}$$

gerekli düzeltmeleri yapalım:

$$\begin{aligned}
& E[-\beta^T C^T C(B - \beta) + \beta^T C^T (B - \beta) - B^T C(B - \beta) \\
& \quad + B^T C - B^T C^T C(B - \beta) - B^T C^T (B - \beta) + \beta^T C(B - \beta) \\
& \quad + B^T B - B^T \beta - \beta^T B + \beta^T \beta \\
& \quad + \beta^T C^T C(X^T X)^{-1} X^T \varepsilon + \varepsilon^T X(X^T X)^{-1} C^T C \beta \\
& \quad + \varepsilon^T X(X^T X)^{-1} C^T C(X^T X)^{-1} X^T \varepsilon \\
& \quad + \varepsilon^T X(X^T X)^{-1} C^T B - \varepsilon^T X(X^T X)^{-1} C^T C B \\
& \quad - \varepsilon^T X(X^T X)^{-1} C^T \beta + B^T C(X^T X)^{-1} X^T \varepsilon \\
& \quad - B^T C^T C(X^T X)^{-1} X^T \varepsilon - \beta^T C(X^T X)^{-1} X^T \varepsilon] \\
& = E[(-\beta^T C^T C + B^T C^T C + \beta^T C^T - B^T C^T + \beta^T C - B^T C)(B - \beta) \\
& \quad + (B - \beta)^T (B - \beta) \\
& \quad + \beta^T C^T C(X^T X)^{-1} X^T \varepsilon + \varepsilon^T X(X^T X)^{-1} C^T C \beta \\
& \quad + \varepsilon^T X(X^T X)^{-1} C^T C(X^T X)^{-1} X^T \varepsilon \\
& \quad + \varepsilon^T X(X^T X)^{-1} C^T B - \varepsilon^T X(X^T X)^{-1} C^T C B \\
& \quad - \varepsilon^T X(X^T X)^{-1} C^T \beta + B^T C(X^T X)^{-1} X^T \varepsilon \\
& \quad - B^T C^T C(X^T X)^{-1} X^T \varepsilon - \beta^T C(X^T X)^{-1} X^T \varepsilon] \\
& = E\{[(B^T - \beta^T)C^T C - (B^T - \beta^T)C^T - (B^T - \beta^T)C](B - \beta) \\
& \quad + (B - \beta)^T (B - \beta) \\
& \quad + \beta^T C^T C(X^T X)^{-1} X^T \varepsilon + \varepsilon^T X(X^T X)^{-1} C^T C \beta \\
& \quad + \varepsilon^T X(X^T X)^{-1} C^T C(X^T X)^{-1} X^T \varepsilon \\
& \quad + \varepsilon^T X(X^T X)^{-1} C^T B - \varepsilon^T X(X^T X)^{-1} C^T C B \\
& \quad - \varepsilon^T X(X^T X)^{-1} C^T \beta + B^T C(X^T X)^{-1} X^T \varepsilon \\
& \quad - B^T C^T C(X^T X)^{-1} X^T \varepsilon - \beta^T C(X^T X)^{-1} X^T \varepsilon\} \\
& = E[(B - \beta)^T (C^T C - C^T - C)(B - \beta) \\
& \quad + (B - \beta)^T (B - \beta) \\
& \quad + \beta^T C^T C(X^T X)^{-1} X^T \varepsilon + \varepsilon^T X(X^T X)^{-1} C^T C \beta \\
& \quad + \varepsilon^T X(X^T X)^{-1} C^T C(X^T X)^{-1} X^T \varepsilon \\
& \quad + \varepsilon^T X(X^T X)^{-1} C^T B - \varepsilon^T X(X^T X)^{-1} C^T C B \\
& \quad - \varepsilon^T X(X^T X)^{-1} C^T \beta + B^T C(X^T X)^{-1} X^T \varepsilon \\
& \quad - B^T C^T C(X^T X)^{-1} X^T \varepsilon - \beta^T C(X^T X)^{-1} X^T \varepsilon]
\end{aligned}$$

$$\begin{aligned}
&= E[(B - \beta)^T (C^T C - C^T - C + I)(B - \beta) \\
&\quad + \beta^T C^T C (X^T X)^{-1} X^T \varepsilon + \varepsilon^T X (X^T X)^{-1} C^T C \beta \\
&\quad + \varepsilon^T X (X^T X)^{-1} C^T C (X^T X)^{-1} X^T \varepsilon \\
&\quad + \varepsilon^T X (X^T X)^{-1} C^T B - \varepsilon^T X (X^T X)^{-1} C^T C B \\
&\quad - \varepsilon^T X (X^T X)^{-1} C^T \beta + B^T C (X^T X)^{-1} X^T \varepsilon \\
&\quad - B^T C^T C (X^T X)^{-1} X^T \varepsilon - \beta^T C (X^T X)^{-1} X^T \varepsilon]
\end{aligned}$$

Beklenen değer fonksiyonunun doğrusal dönüşüm özelliğini kullanarak,

$$\begin{aligned}
&= E[(B - \beta)^T (C^T C - C^T - C + I)(B - \beta)] \\
&+ E[\beta^T C^T C (X^T X)^{-1} X^T \varepsilon + \varepsilon^T X (X^T X)^{-1} C^T C \beta \\
&\quad + \varepsilon^T X (X^T X)^{-1} C^T C (X^T X)^{-1} X^T \varepsilon \\
&\quad + \varepsilon^T X (X^T X)^{-1} C^T B - \varepsilon^T X (X^T X)^{-1} C^T C B \\
&\quad - \varepsilon^T X (X^T X)^{-1} C^T \beta + B^T C (X^T X)^{-1} X^T \varepsilon \\
&\quad - B^T C^T C (X^T X)^{-1} X^T \varepsilon - \beta^T C (X^T X)^{-1} X^T \varepsilon]
\end{aligned}$$

B vektörünün kovaryans matrisi $Kov(B) = E[(B - \beta)^T (B - \beta)] = V$ olarak tanımlandığından, yukarıdaki eşitliği,

$$\begin{aligned}
&= \dot{I}z(CVC^T - CV - VC^T + V) \\
&+ E[\beta^T C^T C (X^T X)^{-1} X^T \varepsilon + \varepsilon^T X (X^T X)^{-1} C^T C \beta \\
&\quad + \varepsilon^T X (X^T X)^{-1} C^T C (X^T X)^{-1} X^T \varepsilon \\
&\quad + \varepsilon^T X (X^T X)^{-1} C^T B - \varepsilon^T X (X^T X)^{-1} C^T C B \\
&\quad - \varepsilon^T X (X^T X)^{-1} C^T \beta + B^T C (X^T X)^{-1} X^T \varepsilon \\
&\quad - B^T C^T C (X^T X)^{-1} X^T \varepsilon - \beta^T C (X^T X)^{-1} X^T \varepsilon]
\end{aligned}$$

şeklinde yazabiliriz. Yine yukarıdaki eşitlikte beklenen değerın lineer dönüşüm özelliğinden faydalanarak,

$$\begin{aligned}
&= \dot{I}z(CVC^T - CV - VC^T + V) \\
&+ E[2\beta^T C^T C (X^T X)^{-1} X^T \varepsilon] \\
&+ E[2B^T C (X^T X)^{-1} X^T \varepsilon] \\
&- E[2B^T C^T C (X^T X)^{-1} X^T \varepsilon] \\
&- E[2\beta^T C (X^T X)^{-1} X^T \varepsilon] \\
&+ E[\varepsilon^T X (X^T X)^{-1} C^T C (X^T X)^{-1} X^T \varepsilon]
\end{aligned}$$

$E(B) = \beta$ olduğundan, yukarıdaki eşitlikte sadeleşen terimler ile birlikte, ifade,

$$= \dot{I}z(CVC^T - CV - VC^T + V) \\ + E[\varepsilon^T X (X^T X)^{-1} C^T C (X^T X)^{-1} X^T \varepsilon]$$

şeklinde sadeleşmiş olur. Varyans tanımından,

$$= \dot{I}z(CVC^T - CV - VC^T + V) \\ + \dot{I}z(C(X^T X)^{-1} C^T) E[\varepsilon \varepsilon^T] \\ = \dot{I}z[\sigma^2 C (X^T X)^{-1} C^T + CVC^T - CV - VC^T + V]$$

elde edilmiş olur. Böylelikle, önerilen konveks kestiricinin C matrisine bağlı hata kareler ortalaması,

$$HKO(\beta(C, B)) = \dot{I}z[\sigma^2 C (X^T X)^{-1} C^T + CVC^T - CV - VC^T + V] \quad (\text{Ek A.2})$$

şeklinde elde edilir. Hata kareler ortalamasını minimum yapan C matrisini bulabilmek için, (Ek A.2) ile bulunan ifadenin C matrisine göre türevi alınıp sıfıra eşitlendiğinde,

$$\frac{\partial HKO(\beta(C, B))}{\partial C} = 2\sigma^2 C (X^T X)^{-1} + 2CV - V - V = 0 \quad (\text{Ek A.3})$$

ifadesi bulunur. Bu ifadede C yalnız bırakıldığında,

$$2C(\sigma^2 (X^T X)^{-1} + V) - 2V = 0 \\ C = V(\sigma^2 (X^T X)^{-1} + V)^{-1} \quad (\text{Ek A.4})$$

eşitliği bulunur. $Kov(\hat{\beta}) = \sigma^2 (X^T X)^{-1} = \Sigma$ kısaltması ile (Ek A.4) eşitliği,

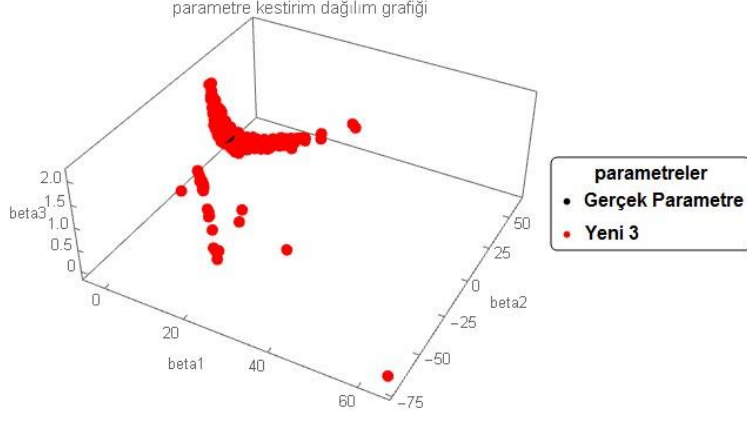
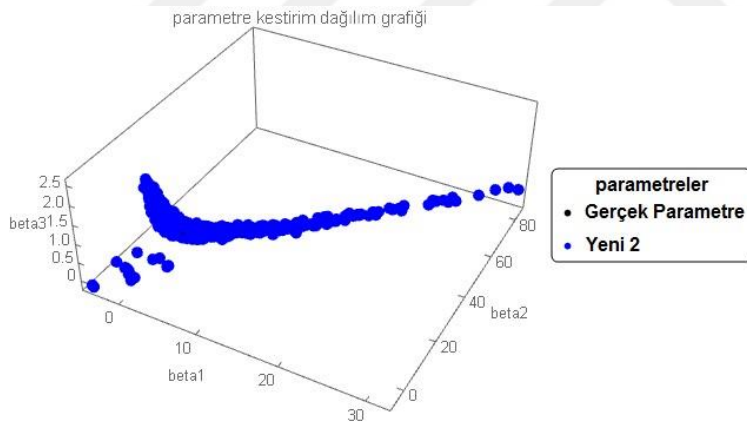
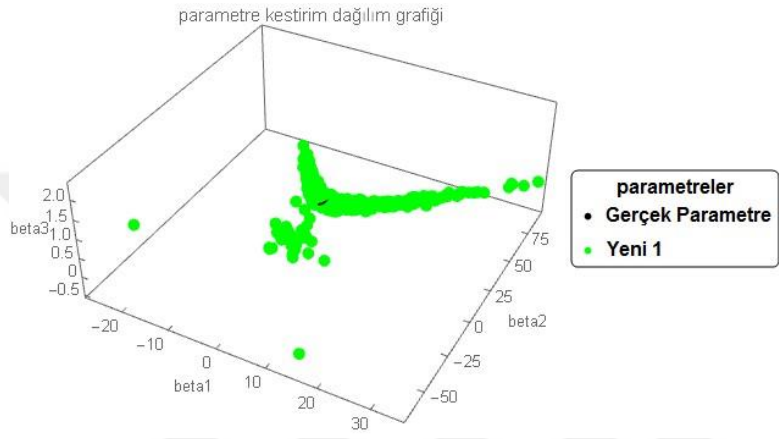
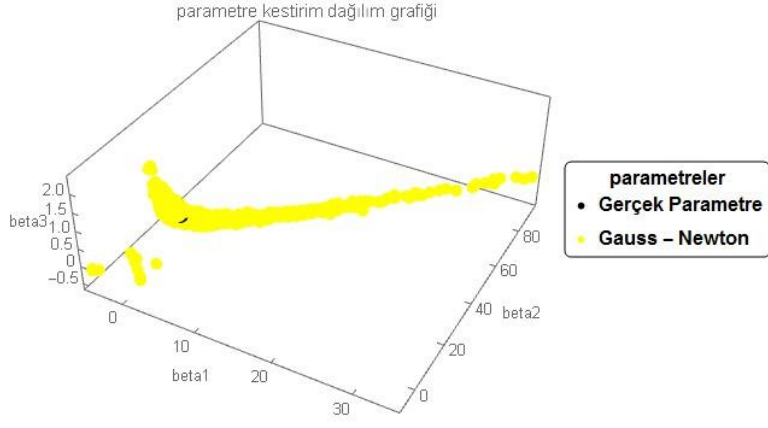
$$C = V(\Sigma + V)^{-1} \quad (\text{Ek A.5})$$

şeklinde bulunmuş olur.

EK B 1. DİĞER ÇİZELGE VE GRAFİKLER

Çizelge B.1 $n = 10$ ve $\sigma = 0.5$ için benzetim sonuçları

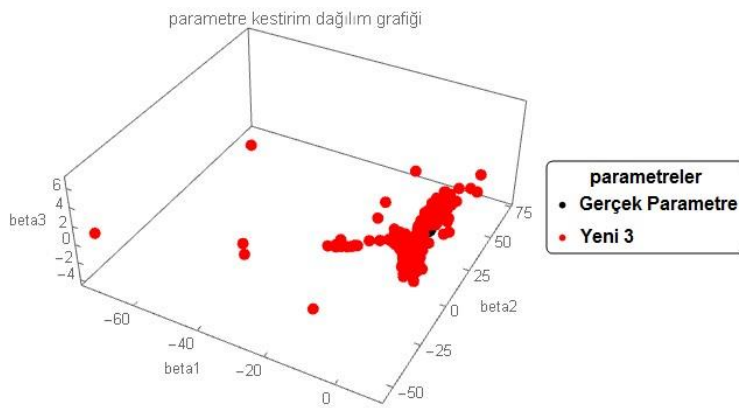
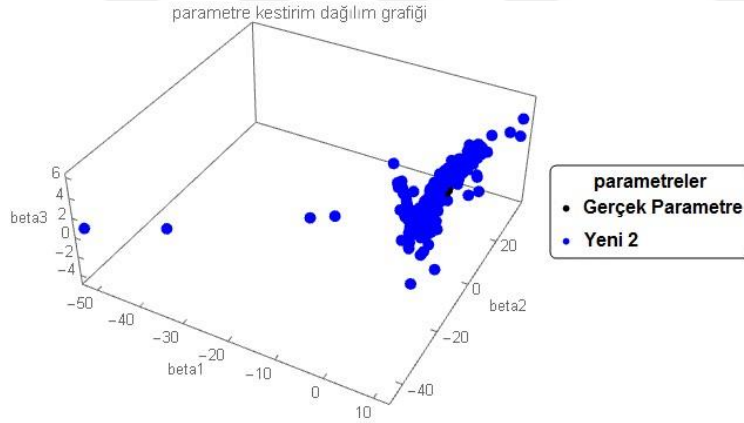
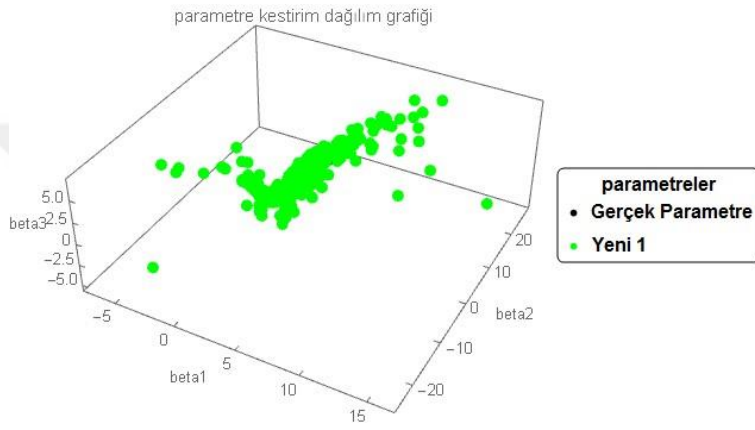
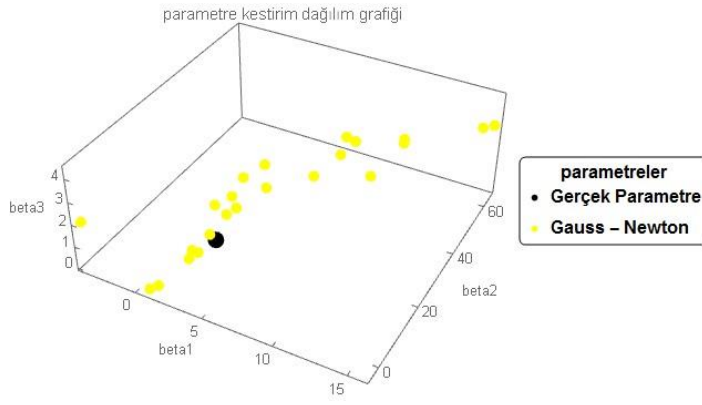
$n = 10$	$\sigma = 0.5$	$\theta^T = (3,12,1)$				
			$\hat{\theta}_{GN}$	$\hat{\theta}_{yeni1}$	$\hat{\theta}_{yeni2}$	$\hat{\theta}_{yeni3}$
			5.96	0.90	0.48	-0.05
			20.41	6.76	6.05	5.15
			0.96	1.01	0.54	0.68
$\rho=0.80$						
	HKO	443.545		78.2578	134.951	201.853
	Yakınsama oranı	% 5.8		% 22.3	% 31.55	% 27.1
			4.50	3.05	2.73	2.56
			18.95	13.50	12.29	11.62
			0.93	0.96	0.87	0.91
$\rho=0.85$						
	HKO	275.705		131.933	114.192	106.748
	Yakınsama oranı	% 47.8		% 68.4	% 70.8	% 67.85
			3.79	3.30	3.23	3.02
			18.61	14.51	13.89	10.91
			0.96	0.94	0.95	0.76
$\rho=0.90$						
	HKO	364.809		202.584	164.238	236.181
	Yakınsama oranı	% 46.9		% 77.95	% 80.1	% 51.2
			4.41	3.99	3.79	3.39
			15.58	14.34	13.98	12.16
			0.97	1.01	1.07	0.89
$\rho=0.99$						
	HKO	137.176		123.633	100.155	77.1064
	Yakınsama oranı	% 80.35		% 87.7	% 88.15	% 65.9
			4.42	4.68	4.51	4.03
			12.86	10.44	10.34	7.39
			0.20	0.80	0.77	0.40
$\rho=0.999$						
	HKO	484.729		195.124	226.423	323.596
	Yakınsama oranı	% 28.6		% 54.95	% 68.2	% 34



Şekil B.1 $n = 10$, $\sigma = 0.5$ ve $\rho = 0.99$ için parametre dağılım grafikleri

Çizelge B.2 $n = 10$ ve $\sigma = 2$ için benzetim sonuçları

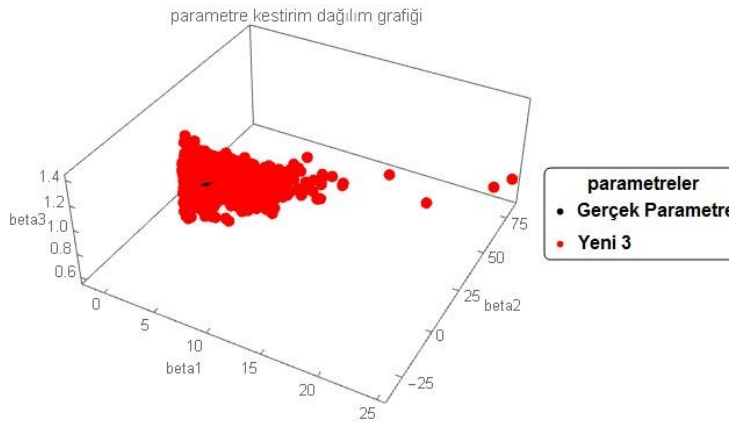
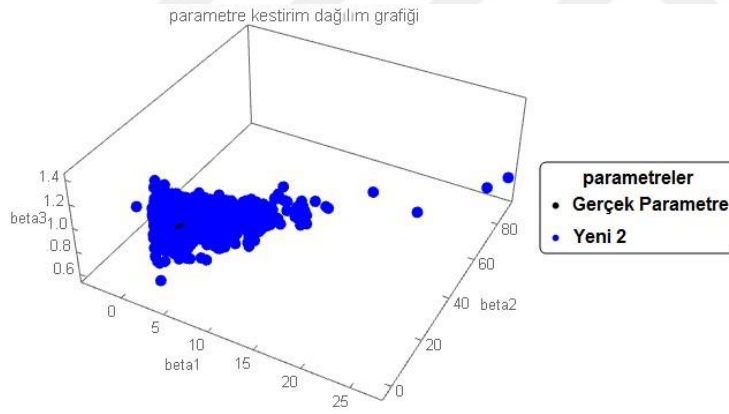
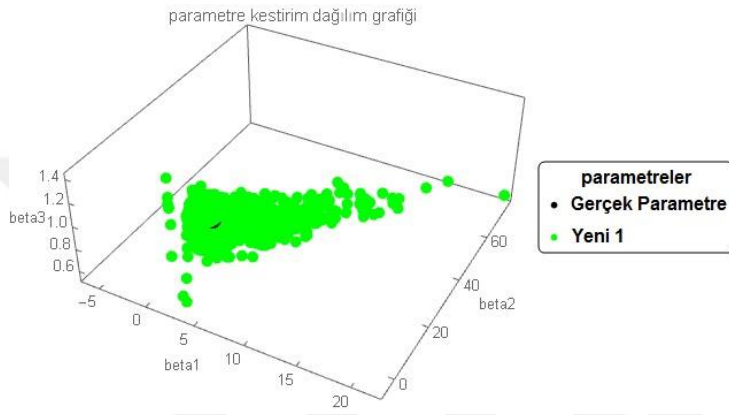
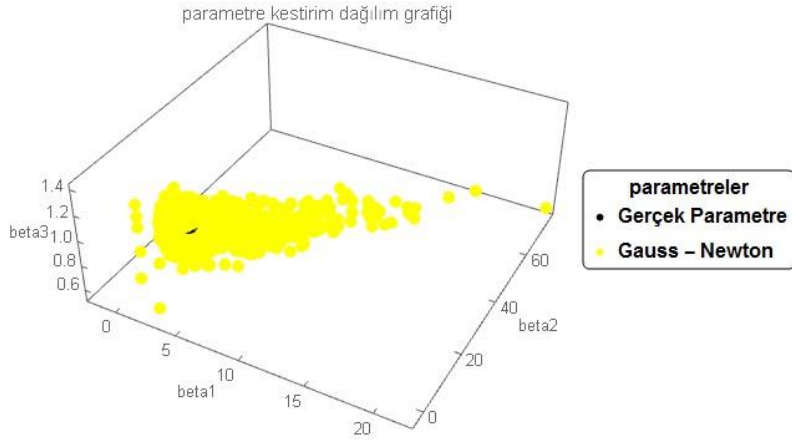
$n = 10$	$\sigma = 2$	$\theta^T = (3,12,1)$				
			$\hat{\theta}_{GN}$	$\hat{\theta}_{yeni1}$	$\hat{\theta}_{yeni2}$	$\hat{\theta}_{yeni3}$
$\rho=0.80$			-2.50	-1.70	-0.06	-1.02
			2.95	0.12	3.56	2.82
			0.44	0.27	-0.17	0.14
		HKO	560.704	301.082	384.585	460.874
		Yakınsama oranı	% 3.2	% 23.4	% 22.7	% 22.5
$\rho=0.85$			5.99	2.00	1.75	1.25
			22.56	5.11	4.57	4.62
			2.54	0.52	0.55	0.55
		HKO	403.533	129.89	114.737	167.78
		Yakınsama oranı	% 1.15	% 20.35	% 33.55	% 29.45
$\rho=0.90$			2.61	0.80	2.54	2.23
			8.74	4.29	5.62	4.98
			1.48	2.01	2.74	3.18
		HKO	1041.5	311.632	312.27	315.217
		Yakınsama oranı	% 9.35	% 26.65	% 18.7	% 19.05
$\rho=0.99$			4.80	1.82	0.77	6.73
			20.63	2.54	-0.44	18.52
			2.15	0.91	1.03	1.00
		HKO	1530.23	339.448	330.467	775.827
		Yakınsama oranı	% 1.25	% 4.05	% 7.2	% 6.7
$\rho=0.999$			6.10	-1.55	-2.46	5.27
			18.74	0.37	-1.96	17.71
			0.70	0.48	0.79	0.79
		HKO	953.753	240.945	317.31	312.381
		Yakınsama oranı	% 7.9	% 20.7	% 22.95	% 33.15



Şekil B.2 $n = 10$, $\sigma = 2$ ve $\rho = 0.85$ için parametre dağılım grafikleri

Çizelge B.3 $n = 20$ ve $\sigma = 0.5$ için benzetim sonuçları

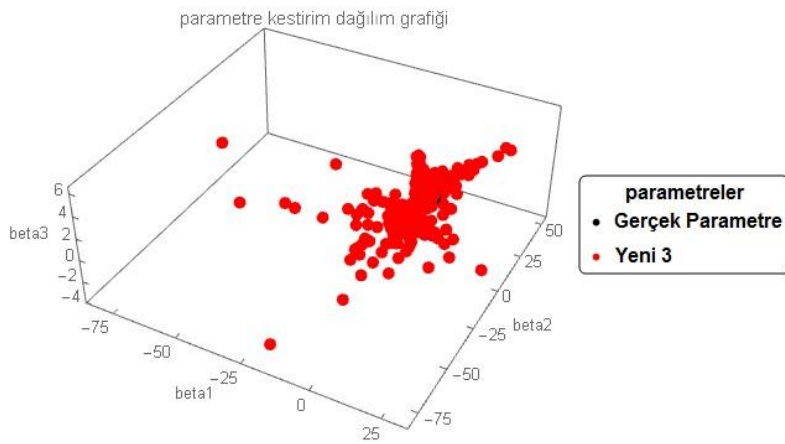
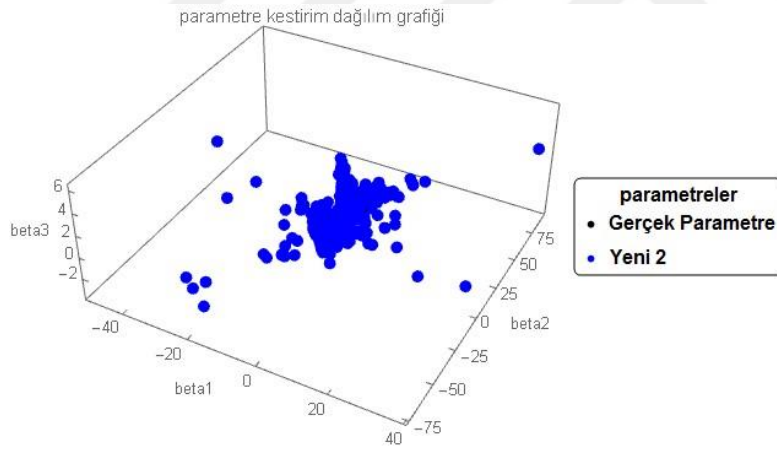
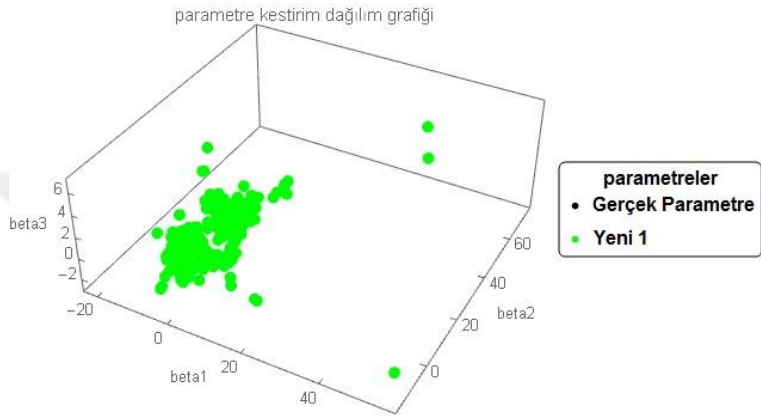
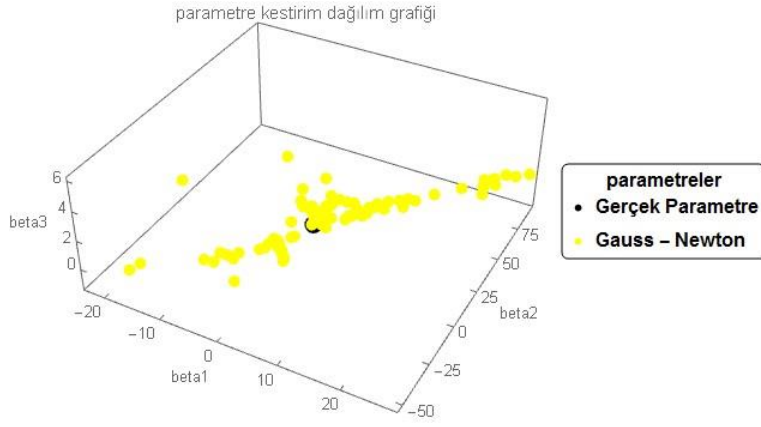
$n = 20$	$\sigma = 0.5$	$\theta^T = (3,12,1)$				
			$\hat{\theta}_{GN}$	$\hat{\theta}_{yeni1}$	$\hat{\theta}_{yeni2}$	$\hat{\theta}_{yeni3}$
			3.86	3.72	3.49	3.37
			14.74	14.34	13.62	13.16
			0.94	0.98	0.97	1.00
$\rho=0.80$						
	HKO		63.7493	59.8943	50.2405	42.3981
	Yakınsama oranı		% 85.15	% 91.05	% 94.35	% 93.8
			0.52	2.26	1.21	0.72
			5.27	9.57	6.58	5.56
			0.72	0.98	0.87	0.88
$\rho=0.85$						
	HKO		659.42	298.118	275.197	299.639
	Yakınsama oranı		% 10.85	% 37.65	% 50.6	% 45.95
			4.14	4.01	3.63	3.39
			15.27	14.94	13.77	13.13
			0.92	0.98	1.06	1.03
$\rho=0.90$						
	HKO		145.951	143.147	125.718	123.184
	Yakınsama oranı		% 81.5	% 89.8	% 84.25	% 79.05
			-0.75	2.18	0.26	-1.18
			6.45	8.29	6.61	5.91
			0.32	0.49	-1.28	-1.22
$\rho=0.99$						
	HKO		750.841	298.905	424.005	457.615
	Yakınsama oranı		% 7.4	% 28	% 35.6	% 32.15
			3.51	0.93	0.35	-0.31
			18.00	4.96	3.22	1.33
			0.76	0.87	0.005	-0.51
$\rho=0.999$						
	HKO		609.282	271.082	330.516	473.578
	Yakınsama oranı		% 6.3	% 27.45	% 34.95	% 32.55



Şekil B.3 $n = 20$, $\sigma = 0.5$ ve $\rho = 0.80$ için parametre dağılım grafikleri

Çizelge B.4 $n = 20$ ve $\sigma = 2$ için benzetim sonuçları

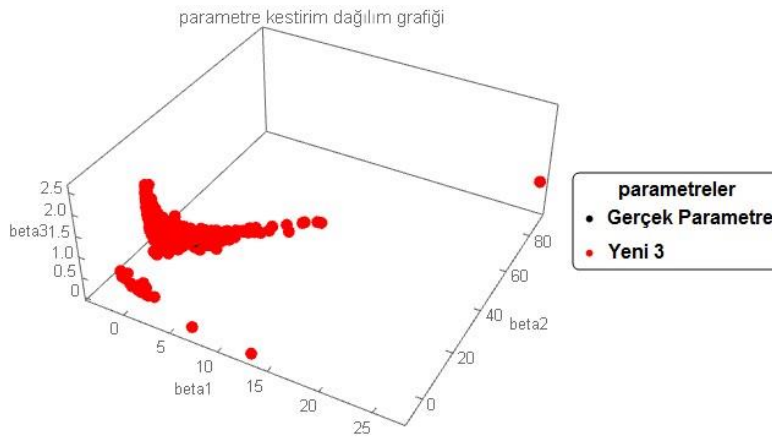
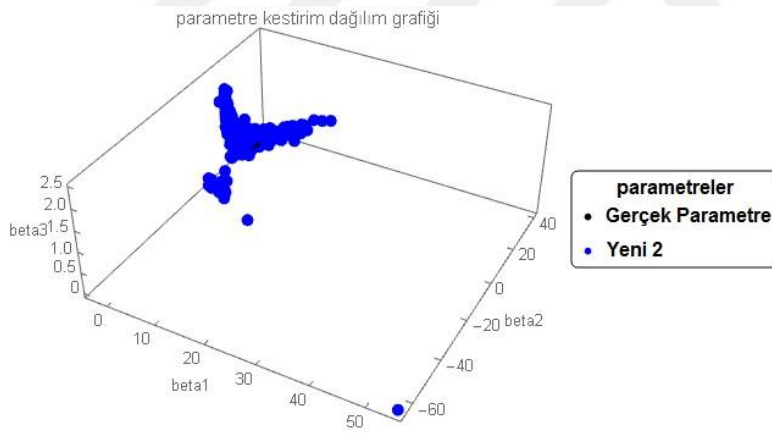
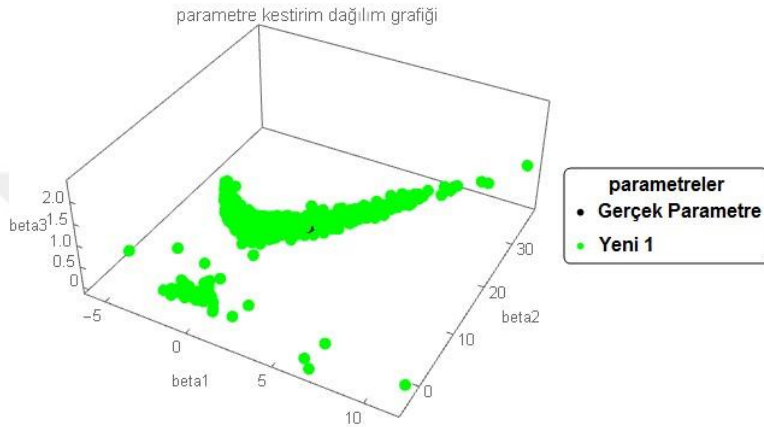
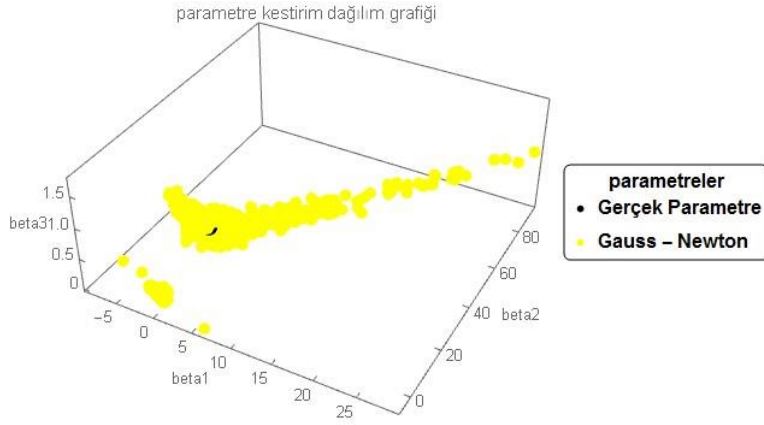
$n = 20$	$\sigma = 2$	$\theta^T = (3,12,1)$				
			$\hat{\theta}_{GN}$	$\hat{\theta}_{yeni1}$	$\hat{\theta}_{yeni2}$	$\hat{\theta}_{yeni3}$
			1.48	2.81	2.26	1.05
			5.91	10.79	8.78	5.75
			1.37	1.01	1.09	0.95
$\rho=0.80$						
	HKO	622.434		210.711	258.513	501.738
	Yakınsama oranı	% 32.3		% 54.4	% 50.55	% 48.85
			3.45	1.78	1.64	0.86
			11.61	5.46	6.15	4.40
			0.86	0.93	0.84	0.57
$\rho=0.85$						
	HKO	434.907		141.644	185.963	274.918
	Yakınsama oranı	% 12.2		% 33.3	% 35.15	% 36.45
			-4.58	-2.12	-1.15	-2.14
			-12.92	2.32	2.85	1.58
			0.54	0.59	0.03	0.29
$\rho=0.90$						
	HKO	599.617		278.569	208.338	329.181
	Yakınsama oranı	% 2.25		% 5.9	% 10.1	% 10.85
			5.41	-0.10	-0.15	-0.22
			19.03	0.75	0.67	0.39
			1.23	1.08	1.15	1.02
$\rho=0.99$						
	HKO	984.84		242.1	322.599	322.249
	Yakınsama oranı	% 4.1		% 24.15	% 29.05	% 29.8
			-0.51	3.35	2.96	0.44
			1.08	4.98	5.33	1.52
			0.57	0.78	0.46	0.67
$\rho=0.999$						
	HKO	1159.29		674.253	485.304	770.347
	Yakınsama oranı	% 12.8		% 36.35	% 37.35	% 35.25



Şekil B.4 $n = 20$, $\sigma = 2$ ve $\rho = 0.99$ için parametre dağılım grafikleri

Çizelge B.5 $n = 50$ ve $\sigma = 0.5$ için benzetim sonuçları

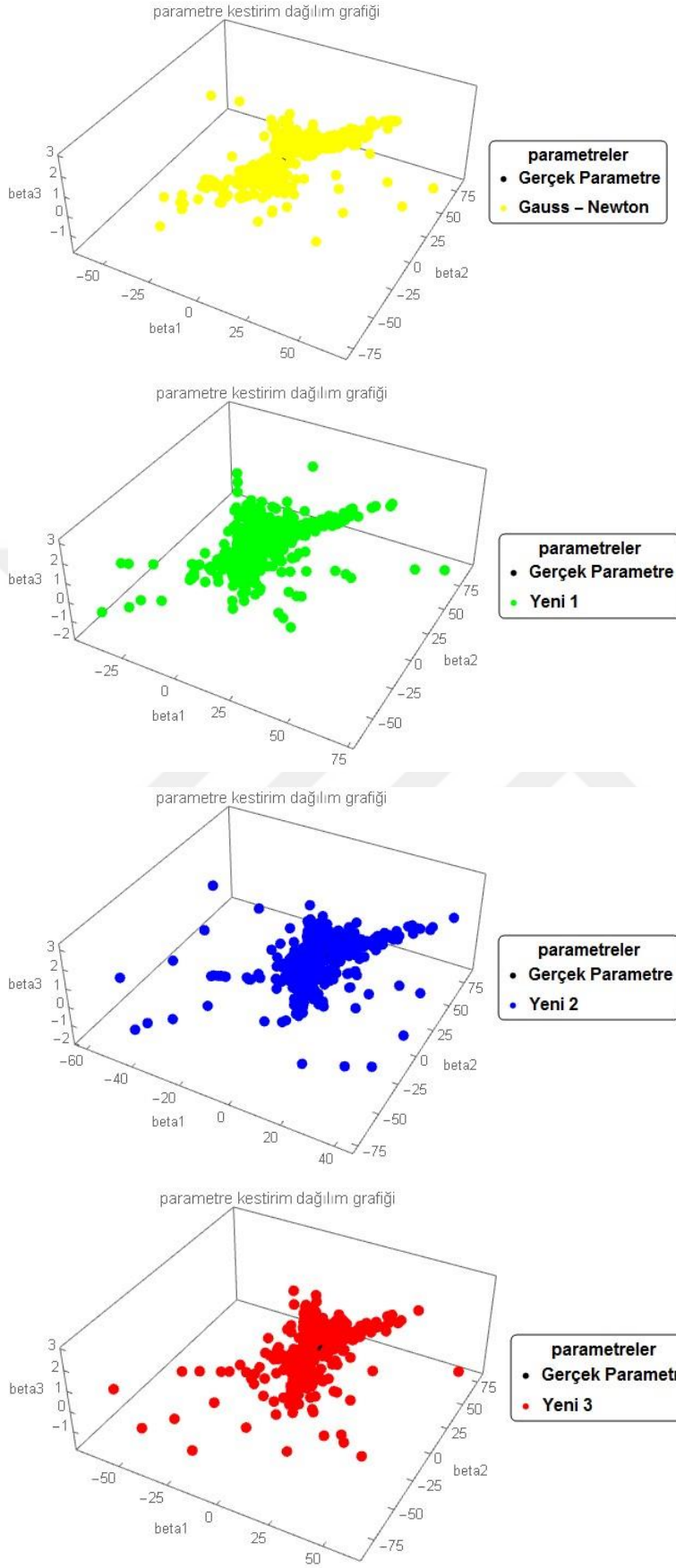
$n = 50$	$\sigma = 0.5$	$\theta^T = (3,12,1)$				
			$\hat{\theta}_{GN}$	$\hat{\theta}_{yeni1}$	$\hat{\theta}_{yeni2}$	$\hat{\theta}_{yeni3}$
			3.41	1.87	1.48	1.22
			13.59	8.06	6.99	6.34
			0.89	1.23	1.22	1.26
$\rho=0.80$						
	HKO	302.963		96.1235	97.322	92.4354
	Yakınsama oranı	% 13.65		% 43.65	% 53.3	% 54.55
			4.50	1.83	1.66	1.49
			16.91	8.15	7.44	7.11
			0.90	1.10	1.17	1.30
$\rho=0.85$						
	HKO	250.277		54.5132	61.9388	58.1765
	Yakınsama oranı	% 27.8		% 50.35	% 53.05	% 53.7
			3.89	1.96	1.69	1.46
			15.36	8.73	7.99	7.34
			0.92	1.19	1.22	1.22
$\rho=0.90$						
	HKO	458.468		90.6287	88.2861	91.374
	Yakınsama oranı	% 24.4		% 48.15	% 53.25	% 51.8
			2.84	2.87	2.72	2.64
			11.49	11.50	11.24	11.04
			0.99	0.99	1.02	1.03
$\rho=0.99$						
	HKO	42.5951		27.1803	27.0986	28.8649
	Yakınsama oranı	% 86.15		% 94.85	% 94.8	% 93.5
			4.31	3.95	3.33	3.08
			16.23	14.81	13.38	12.32
			0.81	1.03	0.87	0.97
$\rho=0.999$						
	HKO	350.31		230.374	485.304	167.593
	Yakınsama oranı	% 51.6		% 74.4	% 72.6	% 74.5



Şekil B.5 $n = 50$, $\sigma = 0.5$ ve $\rho = 0.85$ için parametre dağılım grafikleri

Çizelge B.6 $n = 50$ ve $\sigma = 2$ için benzetim sonuçları

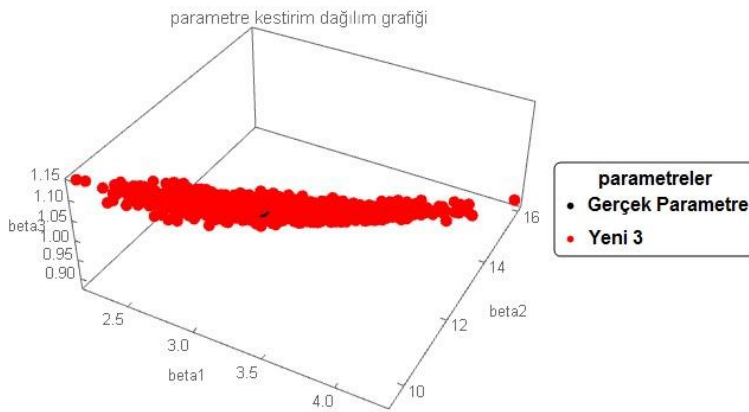
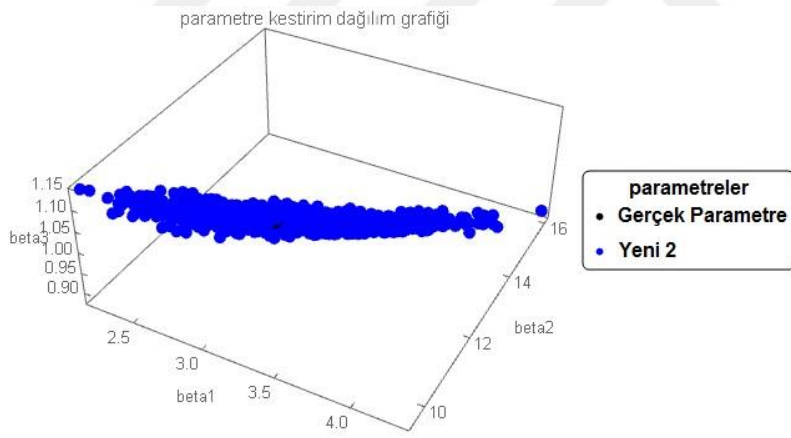
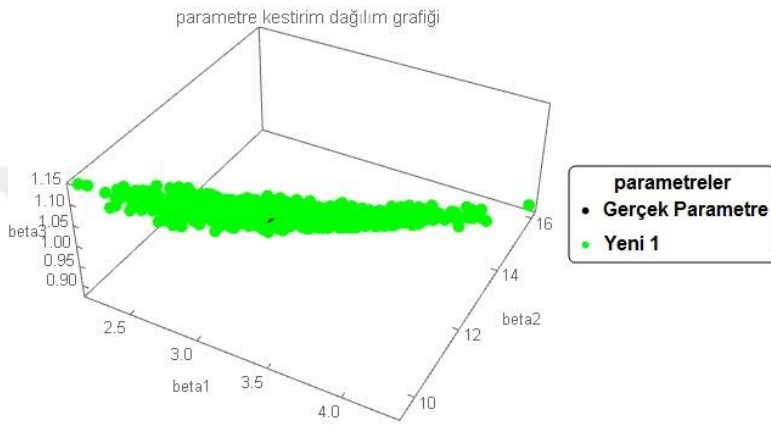
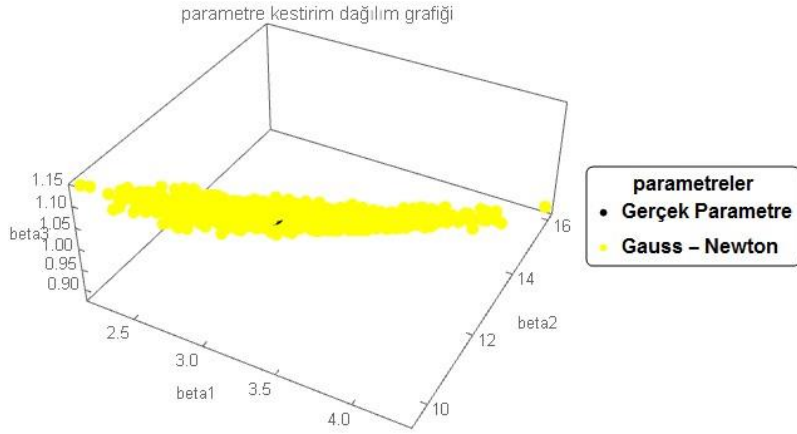
$n = 50$	$\sigma = 2$	$\theta^T = (3,12,1)$				
			$\hat{\theta}_{GN}$	$\hat{\theta}_{yeni1}$	$\hat{\theta}_{yeni2}$	$\hat{\theta}_{yeni3}$
			1.33	1.37	1.85	1.43
			6.01	6.95	7.97	7.06
			0.66	0.53	0.67	0.70
$\rho=0.80$						
	HKO	348.745		213.024	248.335	273.956
	Yakınsama oranı	% 28.05		% 46.5	% 50.85	% 51.75
			2.78	2.91	2.88	2.16
			7.35	11.76	11.63	10.02
			0.59	0.89	0.93	0.76
$\rho=0.85$						
	HKO	304.895		275.296	283.8	301.319
	Yakınsama oranı	% 42.4		% 61	% 59.6	% 57.2
			2.45	1.85	2.24	1.97
			9.87	8.66	9.51	9.23
			0.61	0.66	0.67	0.77
$\rho=0.90$						
	HKO	224.98		118.81	89.1331	107.328
	Yakınsama oranı	% 40.05		% 57.95	% 64.1	% 62.45
			0.01	0.82	1.82	1.23
			1.25	4.91	6.55	5.85
			0.76	0.68	0.81	0.72
$\rho=0.99$						
	HKO	1001.27		525.162	484.175	556.68
	Yakınsama oranı	% 25.2		% 50.4	% 51.55	% 52.85
			2.55	1.96	1.52	1.28
			8.87	10.09	8.62	7.39
			0.48	0.52	0.46	0.43
$\rho=0.999$						
	HKO	538.149		346.625	342.994	352.528
	Yakınsama oranı	% 31.75		% 51.85	% 46.45	% 45.95



Şekil B.6 $n = 50$, $\sigma = 2$ ve $\rho = 0.999$ için parametre dağılım grafikleri

Çizelge B.7 $n = 100$ ve $\sigma = 0.1$ için benzetim sonuçları

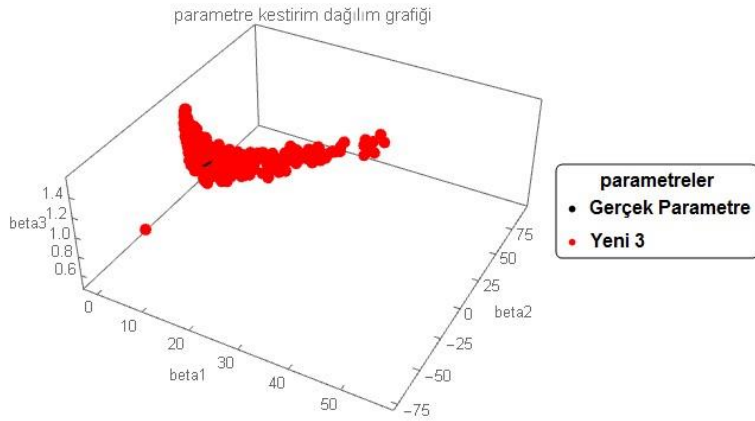
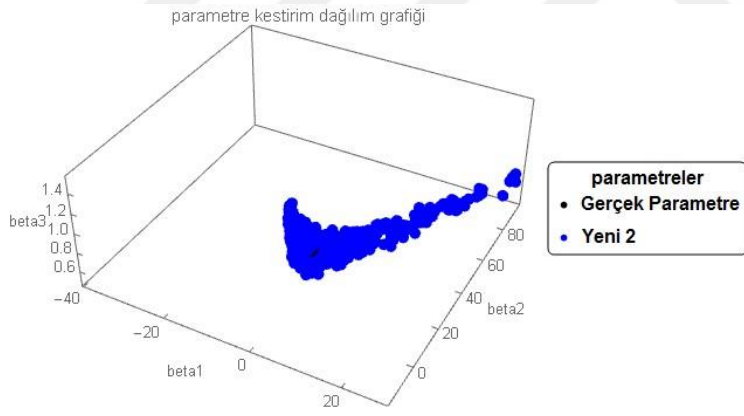
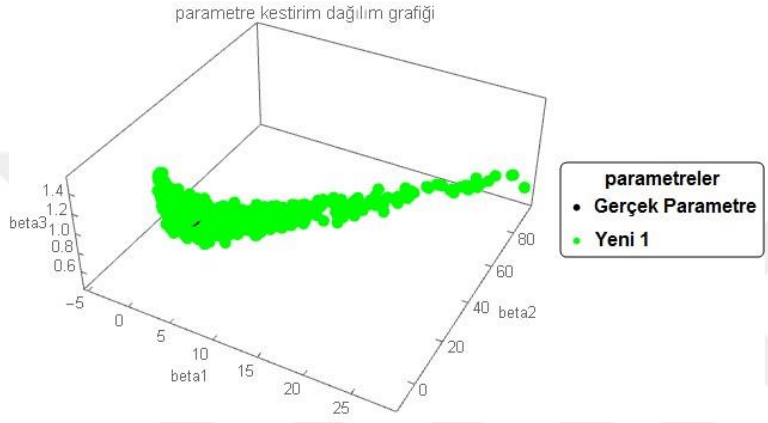
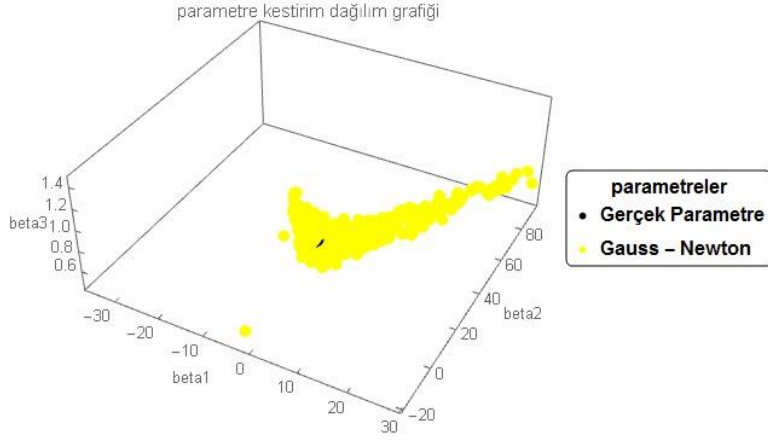
$n = 100$	$\sigma = 0.1$	$\theta^T = (3,12,1)$				
			$\hat{\theta}_{GN}$	$\hat{\theta}_{yeni1}$	$\hat{\theta}_{yeni2}$	$\hat{\theta}_{yeni3}$
$\rho=0.80$	HKO Yakınsama oranı		3.0006	3.0006	2.999	2.999
			12.002	12.002	12.0001	11.998
			1.0004	1.0004	1.0005	1.0006
			0.122968	0.122973	0.122916	0.122865
			% 100	% 100	% 100	% 100
$\rho=0.85$	HKO Yakınsama oranı		3.06	3.06	3.05	3.04
			12.22	12.22	12.18	12.14
			0.99	0.99	1.0005	1.001
			2.46764	2.52332	2.47426	2.42987
			% 100	% 100	% 100	% 100
$\rho=0.90$	HKO Yakınsama oranı		3.02	3.02	3.01	3.01
			12.07	12.07	12.05	12.04
			1.0002	0.96	1.0007	1.001
			0.962232	0.967982	0.960534	0.953591
			% 100	% 100	% 100	% 100
$\rho=0.99$	HKO Yakınsama oranı		3.004	3.004	3.0008	2.99
			12.01	12.01	11.99	11.98
			0.99	0.99	0.99	0.99
			1.1197	1.12	1.11387	1.10837
			% 100	% 100	% 100	% 100
$\rho=0.999$	HKO Yakınsama oranı		3.002	3.002	3.001	3.0009
			12.006	12.006	12.004	12.002
			1.0003	1.0003	1.0004	1.0005
			0.135335	0.135383	0.13523	0.135086
			% 100	% 100	% 100	% 100



Şekil B.7 $n = 100$, $\sigma = 0.1$ ve $\rho = 0.90$ için parametre dağılım grafikleri

Çizelge B.8 $n = 100$ ve $\sigma = 0.5$ için benzetim sonuçları

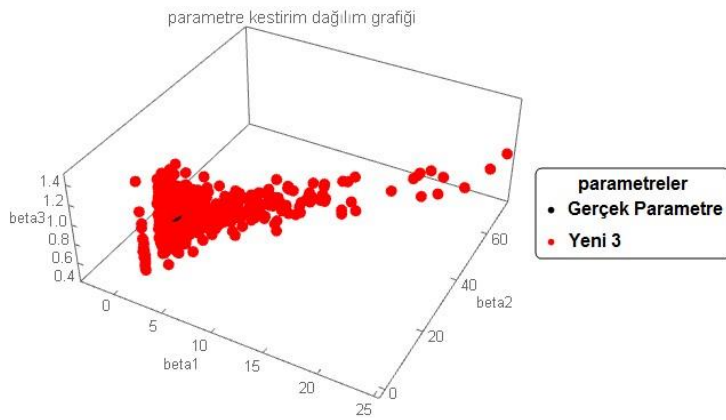
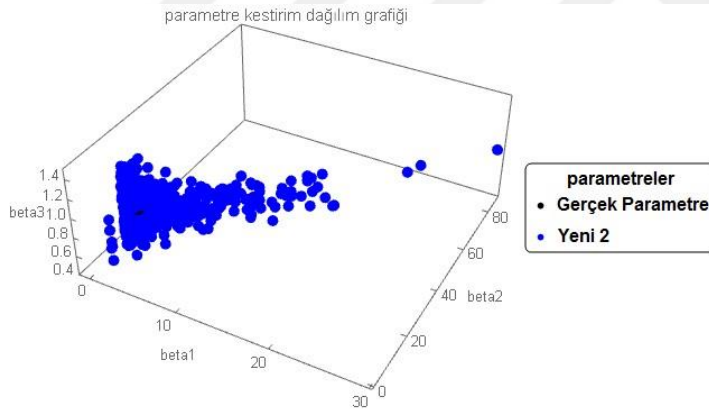
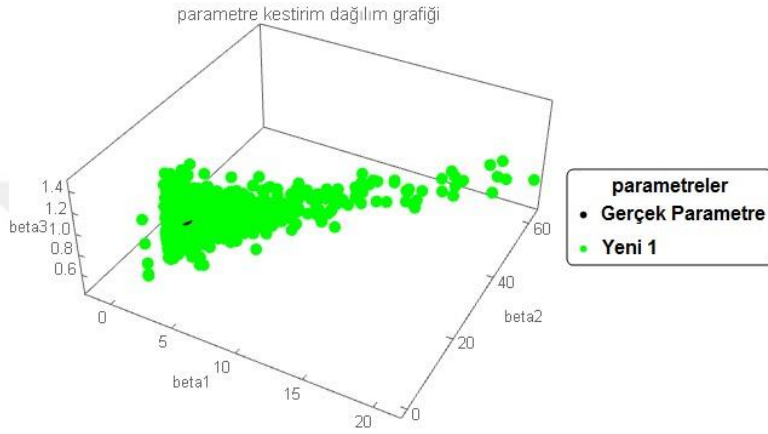
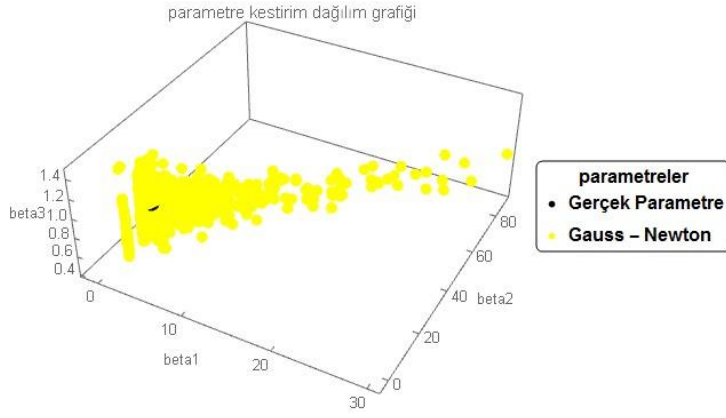
$n = 100$	$\sigma = 0.5$	$\theta^T = (3,12,1)$				
			$\hat{\theta}_{GN}$	$\hat{\theta}_{yeni1}$	$\hat{\theta}_{yeni2}$	$\hat{\theta}_{yeni3}$
$\rho=0.80$			3.03	3.03	3.009	2.98
			12.14	12.14	12.04	11.96
			1.003	1.003	1.007	1.01
		HKO	6.24506	6.16583	6.1407	6.08974
		Yakınsama oranı	% 99.65	% 99.75	% 99.95	% 99.95
$\rho=0.85$			5.08	4.05	3.59	2.90
			17.55	14.81	14.25	13.33
			0.78	0.87	0.78	0.59
		HKO	326.506	158.416	182.563	239.745
		Yakınsama oranı	% 48.35	% 73.4	% 66.05	% 58.4
$\rho=0.90$			3.15	3.12	3.09	3.06
			12.53	12.43	12.31	12.21
			0.99	0.99	0.99	1.002
		HKO	7.43607	7.75201	7.46989	7.20518
		Yakınsama oranı	% 96.7	% 99.35	% 99.7	% 99.7
$\rho=0.99$			4.52	4.28	3.73	3.54
			17.11	16.26	14.56	13.47
			0.94	1.007	1.01	0.98
		HKO	191.111	191.73	158.471	164.928
		Yakınsama oranı	% 81	% 90.35	% 84.75	% 81.65
$\rho=0.999$			3.06	3.06	3.05	3.03
			12.20	12.20	12.15	12.11
			0.99	0.99	0.99	0.99
		HKO	3.17839	3.22448	3.11353	3.01367
		Yakınsama oranı	% 100	% 100	% 100	% 100



Şekil B.8 $n = 100$, $\sigma = 0.5$ ve $\rho = 0.99$ için parametre dağılım grafikleri

Çizelge B.9 $n=100$ ve $\sigma=1$ için benzetim sonuçları

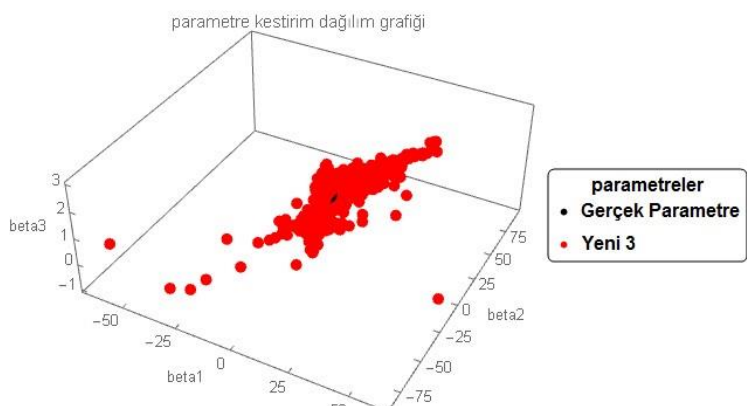
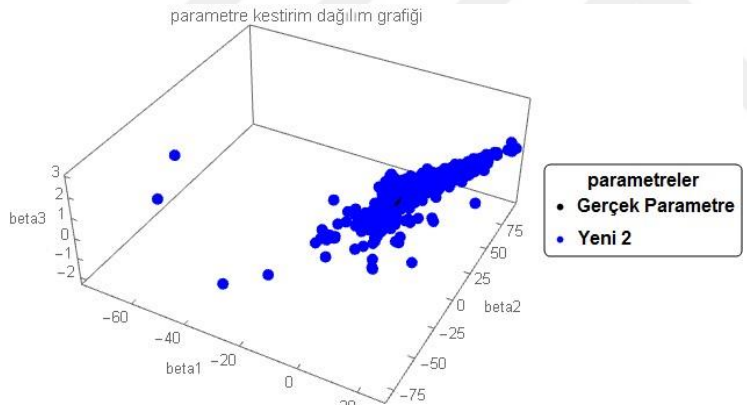
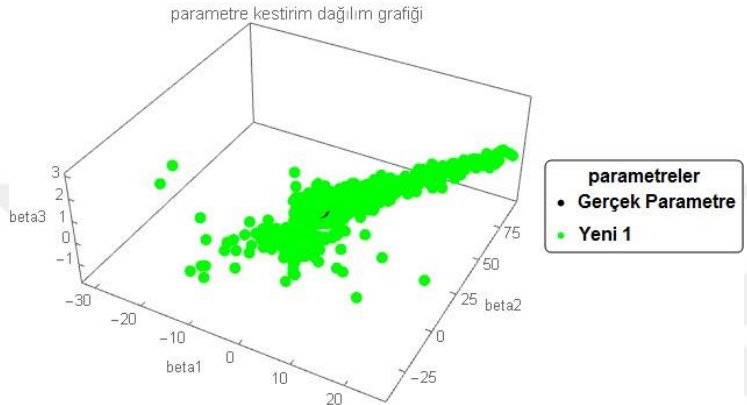
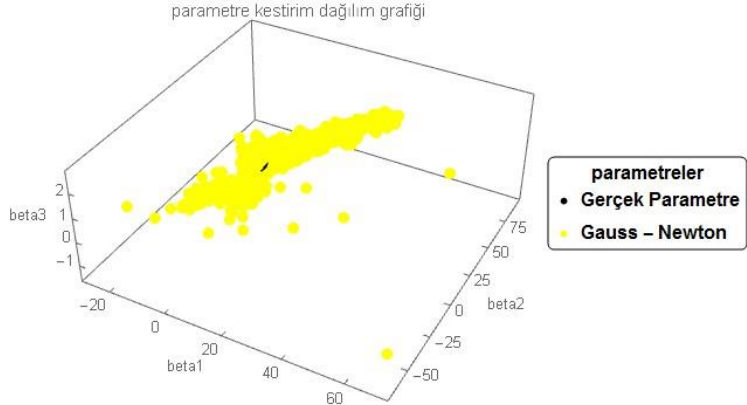
$n=100$	$\sigma=1$	$\theta^T = (3,12,1)$				
			$\hat{\theta}_{GN}$	$\hat{\theta}_{yeni1}$	$\hat{\theta}_{yeni2}$	$\hat{\theta}_{yeni3}$
$\rho=0.80$			3.01	3.19	3.05	2.93
			11.84	12.54	12.07	11.89
			0.89	0.97	0.93	0.89
		HKO	41.5114	17.0159	23.5997	35.4589
		Yakınsama oranı	% 87.65	% 97.85	% 94.65	% 91
$\rho=0.85$			4.87	4.17	3.50	3.27
			17.94	15.47	13.59	12.71
			0.82	0.89	0.87	0.89
		HKO	353.223	270.533	234.014	230.99
		Yakınsama oranı	% 51.25	% 69.2	% 67.65	% 66.8
$\rho=0.90$			5.30	3.40	3.14	3.13
			19.66	13.71	12.75	12.19
			0.79	0.79	0.79	0.96
		HKO	403.628	176.615	161.009	192.993
		Yakınsama oranı	% 38.8	% 52.95	% 58	% 58.35
$\rho=0.99$			3.96	3.61	3.45	3.44
			14.73	13.74	13.29	13.25
			0.97	0.98	0.98	0.98
		HKO	85.877	47.6136	41.4426	43.0813
		Yakınsama oranı	% 80.45	% 92.5	% 92.35	% 85.15
$\rho=0.999$			2.74	3.46	3.39	3.06
			12.10	14.96	13.93	12.94
			0.75	0.89	0.85	0.84
		HKO	375.479	254.514	208.293	201.964
		Yakınsama oranı	% 49.6	% 68	% 70.05	% 70.95



Şekil B.9 $n = 100$, $\sigma = 1$ ve $\rho = 0.99$ için parametre dağılım grafikleri

Çizelge B.10 $n = 100$ ve $\sigma = 2$ için benzetim sonuçları

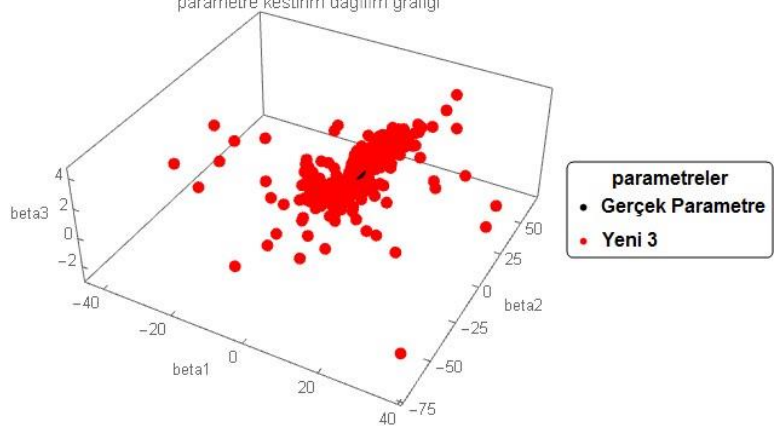
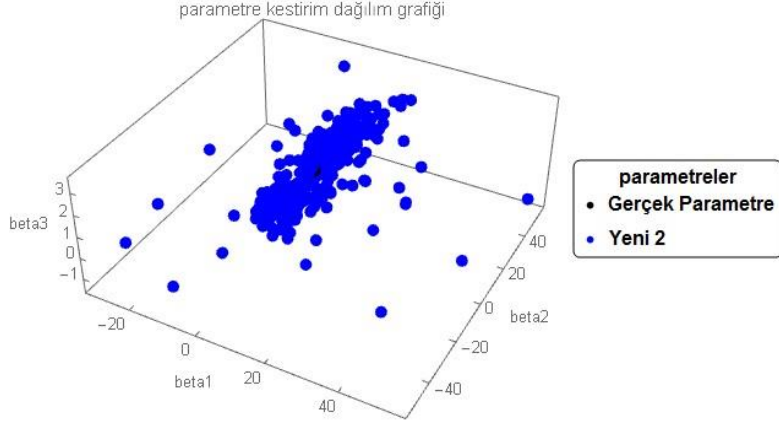
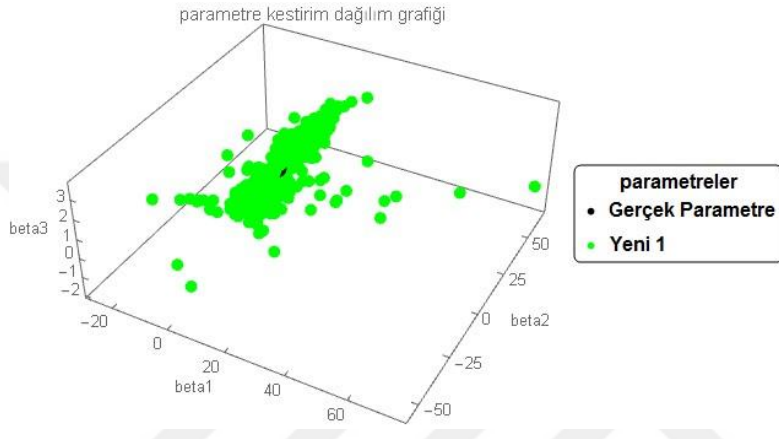
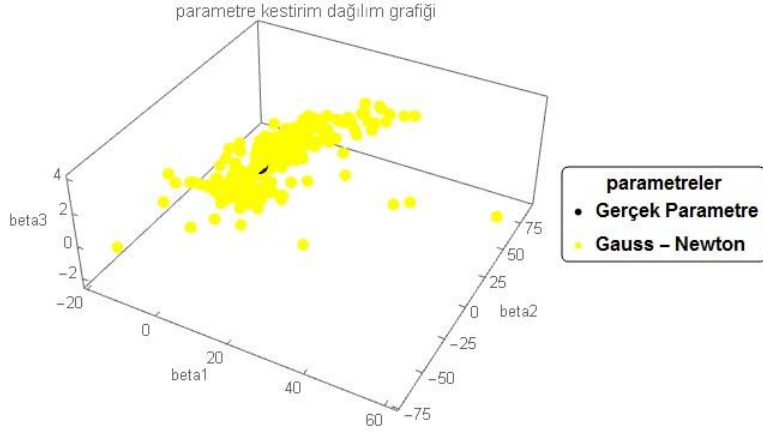
$n = 100$	$\sigma = 2$	$\theta^T = (3,12,1)$				
			$\hat{\theta}_{GN}$	$\hat{\theta}_{yeni1}$	$\hat{\theta}_{yeni2}$	$\hat{\theta}_{yeni3}$
$\rho=0.80$	HKO Yakınsama oranı		3.03	3.01	2.76	2.69
			11.83	11.88	11.27	10.89
			0.76	0.78	0.88	0.92
			133.959	88.9977	88.6242	82.1021
			% 78.1	% 87.95	% 86.65	% 83.25
$\rho=0.85$	HKO Yakınsama oranı		4.11	4.25	3.68	3.52
			15.54	16.69	14.87	13.99
			0.68	0.77	0.78	0.78
			611.464	387.707	326.738	320.899
			% 57.5	% 72.15	% 68.85	% 66.8
$\rho=0.90$	HKO Yakınsama oranı		2.54	2.28	2.25	2.00
			10.73	9.46	9.50	8.44
			0.75	0.69	0.81	0.65
			233.166	133.021	128.741	183.329
			% 45.85	% 60.5	% 58.65	% 57.9
$\rho=0.99$	HKO Yakınsama oranı		4.66	4.56	4.11	3.69
			16.72	16.54	15.18	14.16
			0.60	0.70	0.77	0.76
			468.039	372.012	271.042	276.084
			% 60.95	% 82.75	% 78.35	% 75.35
$\rho=0.999$	HKO Yakınsama oranı		2.94	2.94	2.93	2.94
			11.91	11.89	11.86	11.93
			0.94	0.95	0.95	0.94
			15.4069	13.1413	9.87438	6.71784
			% 90.4	% 93.6	% 92.85	% 91.1



Şekil B.10 $n = 100$, $\sigma = 2$ ve $\rho = 0.85$ için parametre dağılım grafikleri

Çizelge B.11 $n = 100$ ve $\sigma = 3$ için benzetim sonuçları

$n = 100$	$\sigma = 3$	$\theta^T = (3,12,1)$				
			$\hat{\theta}_{GN}$	$\hat{\theta}_{yeni1}$	$\hat{\theta}_{yeni2}$	$\hat{\theta}_{yeni3}$
$\rho=0.80$			2.92	2.56	2.46	2.43
			11.48	9.86	10.22	9.48
			0.73	0.73	0.75	0.75
		HKO	301.249	188.833	160.29	174.713
		Yakınsama oranı	% 56.8	% 67.55	% 68.9	% 68.3
$\rho=0.85$			3.04	2.21	2.20	1.70
			11.75	8.64	8.75	7.88
			0.77	0.73	0.67	0.56
		HKO	281.888	161.005	151.932	170.071
		Yakınsama oranı	% 40.8	% 51.3	% 55.8	% 55.35
$\rho=0.90$			2.80	2.88	2.84	2.68
			11.52	11.58	11.44	11.03
			0.89	0.91	0.91	0.89
		HKO	81.9804	72.2875	69.4487	64.7271
		Yakınsama oranı	% 72.3	% 87.4	% 81.2	% 77.45
$\rho=0.99$			3.15	2.81	2.76	2.45
			11.24	10.24	10.01	8.93
			0.73	0.66	0.67	0.67
		HKO	354.368	284.699	299.65	300.67
		Yakınsama oranı	% 58.25	% 76.15	% 69.85	% 66.6
$\rho=0.999$			2.51	2.75	2.53	2.58
			10.24	11.35	10.63	9.86
			0.77	0.81	0.82	0.80
		HKO	180.315	131.732	162.16	159.338
		Yakınsama oranı	% 54.9	% 75.4	% 71.7	% 67.45



Şekil B.11 $n = 100$, $\sigma = 3$ ve $\rho = 0.85$ için parametre dağılım grafikleri

ÖZGEÇMİŞ

Ali ERKOÇ 1985 yılında Yozgat' ın Sarınınören köyünde doğmuştur. İlk ve orta öğrenimini İstanbul' da tamamladıktan sonra 2003 yılında Trakya Üniversitesi Matematik Bölümünü kazanmıştır. Bir sene sonra Marmara Üniversitesi Matematik Bölümüne yatay geçiş yapmış ve 2007 yılında bölümü derece ile tamamlamıştır. Aynı yıl Marmara Üniversitesi Fen Bilimleri Enstitüsü Matematik Bölümünde yüksek lisansa başlamış, 2011 yılında istatistik alanında yazdığı tez ile buradan mezun olmuştur. 2012 yılında Mimar Sinan Güzel Sanatlar Üniversitesi Fen Bilimleri Enstitüsü İstatistik Programında Doktora başlamıştır. 2013 yılında MSGSÜ İstatistik bölümünde Araştırma Görevlisi olarak başlamıştır ve halen bu görevi sürdürmektedir. Doğrusal olmayan regresyon modelleri üzerine çalışmalar yapmış, yurtiçi ve yurtdışında makaleler yayınlamış, bildiriler sunmuştur.