ÖZET

Doktora Tezi

ÇOK DEĞİŞKENLİ LİNEER OLMAYAN MODELLERDE GENETİK ALGORİTMA

Aydın KARAKOCA

Selçuk Üniversitesi

Fen Bilimleri Enstitüsü

Matematik Anabilim Dalı

Danışman: Doç.Dr. Aşır GENÇ

2009, 86 Sayfa

Jüri: Doç. Dr. Aşır GENÇ

Prof. Dr. Fahrettin ARSLAN

Doç. Dr. Galip OTURANÇ

Doç. Dr. Coşkun KUŞ

Yrd. Doç. Dr. Necati TAŞKARA

Çok değişkenli lineer olmayan modeller birçok uygulamada bağımlı değişken(ler) ile bağımsız değişken(ler) arasındaki ilişkiyi modellemek amacıyla kullanılmaktadır. Lineer olmayan modellerin parametreleri En Küçük Kareler (EKK) yöntemiyle tahmin edilebilmektedir. EKK yönteminde parametre tahmini için en çok Gauss-Newton, Marquardt ve En Hızlı İniş algoritmaları kullanılmaktadır. Bu algoritmaların kullanılabilmesi için bağımsız değişken(ler)in tepki fonksiyonunun en az iki kez türevlenebilmesi şartı gerekmektedir. Ayrıca bu algoritmaların seçilecek başlangıç noktasına göre çözüme ulaşamama riski vardır. Bu çalışmada çok değişkenli lineer olmayan modellerde parametre tahmini için belirtilen algoritmalara alternatif olarak bir genetik algoritma önerilmiştir. Çok değişkenli lineer olmayan modelde önerilen genetik algoritma ve EKK yöntemiyle elde edilen parametre tahmin sonuçları karşılaştırılmıştır.

Anahtar Kelimeler: Çok Değişkenli Lineer Olmayan Model, Genetik Algoritma, Parametre Tahmini

ABSTRACT

PhD Thesis

GENETIC ALGORITHM IN MULTIVARIATE NONLINEAR MODELS

Aydın KARAKOCA

Selcuk University

Graduate School of Natural and Applied Sciences

Department of Mathematics

Supervisor: Assoc. Prof. Dr. Asır GENC

2009, 86 Page

Jury: Assoc.Prof.Dr. Aşır GENÇ

Assoc.Prof.Dr. Galip OTURANÇ

Prof. Dr. Fahrettin ARSLAN

Assoc.Prof.Dr. Coşkun KUŞ

Assist.Prof. Dr. Necati TAŞKARA

Multivariate non-linear models have been used for modelling functional relationship between dependent and independent variable(s) in most of applications. Parameters of multivariate non-linear models can be estimated by least squares (LS) method. Gauss-Newton, Marquardt and Steepest Descent are most widely used algorithms in LS method. These algorithms requires the condition that the function of independent variables can be differentiable at least two times. Also these algorithms have a risk of unreachable solution which depends according to the chosen starting point. In this study as an alternative to these algorithms, genetic algorithms have recommended for parameter estimation in multivariate non-linear models. And then the parameter estimation results of multivariate nonlinear models that obtained by least squares method and genetic algorithm were compared.

Key Words: Multivariate Non-Linear Model, Genetic Algorithm, Parameter Estimation

İÇİNDEKİLER

ÖZET i
ABSTRACT ii
İÇİNDEKİLER iii
ŞEKİL VE ÇİZELGELER DİZİNİ v
SİMGELER VE KISALTMALAR vi
TEŞEKKÜR vii
1. GİRİŞ 1
2. GENETİK ALGORİTMALAR 9
2.1. Genetik Algoritma Terminolojisi 10
2.2. Genetik Algoritmaların Geleneksel Arama Metotlarından Farkı 12
2.3. Genetik Algoritmaların Genel Yapısı 13
2.4. Genetik Kodlama ve Uygunluk Fonksiyonu 13
2.4.1. İkili Kodlama 15
2.4.2. Gray Kodlama ve Hamming Uzaklığı 16
2.4.3. Gerçek Kodlama 18
2.5. Başlangıç Populasyonu 18
2.6. Kopyalama 19
2.6.1. Rulet Çemberi Yöntemi 20
2.6.2. Turnuva Seçim Yöntemi 20
2.6.3. Sıralı Seçim Yöntemi 21
2.6.4. Elitist Seçim Yöntemi 21
2.7. Çaprazlama 22
2.7.1. Tek Nokta Çaprazlama 22
2.7.2. İki Nokta Çaprazlama 24
2.7.3. Düzgün Çaprazlama 24
2.8. Mutasyon 25
2.9. Schema Teorisi ve Seçim Baskısı 26

3.ÇOK DEĞİŞKENLİ LİNEER OLMAYAN MODELLER.....28
3.1. Tek Değişkenli Lineer Olmayan Modelin Tanıtımı ve Bazı Gösterimler.....29
3.2. Çok Değişkenli Lineer Olmayan Modelin Tanıtımı ve Bazı Gösterimler 30
3.2.1. Denklemlere Göre Gruplanmış Gösterim.....32
3.2.2. Gözlemlere Göre Gruplanmış Gösterim.....35
3.3. Tek Değişkenli Lineer Olmayan Modelde Parametre Tahmini 37
3.3.1. Başlangıç noktasının seçimi 40
3.3.2. Durdurma kuralı.....41
3.3.3. Gauss- Newton Yöntemi.....41
3.3.4. En hızlı iniş yöntemi.....43
3.3.5. Marquardt yöntemi 45
3.4. Çok Değişkenli Lineer Olmayan Modelde Parametre Tahmini.....48
3.4.1. En Küçük Kareler Tahmin Edicisi.....48
4. ÇOK DEĞİŞKENLİ LİNEER OLMAYAN MODELDE GENETİK ALGORİTMA İLE PARAMETRE TAHMİNİ.....53
4.1. Tek Değişkenli Lineer Olmayan Genetik Algoritma (NLGA1).....53
4.2. Çok Değişkenli Lineer Olmayan Genetik Algoritma (NLGA2).....59
4.2.1. Jackknife Tahmini 62
4.2.2. Bootstrap Tahmini 63
5. SONUÇ.....69
6. KAYNAKLAR.....71
Ek-1 NLGA1 program kodları.....74
Ek-2 NLGA2 program kodları.....80

ŞEKİL VE ÇİZELGELER DİZİNİ

Şekil 2.2. Tek Nokta Çaprazlama
Şekil 2.3. Gerçek Kodlu GA'da Tek Nokta Çaprazlama
Şekil 2.4. İkili Kodlu GA'da Tek Nokta Çaprazlama
Şekil 2.5. Iki Nokta Çaprazlama
Şekil 2.6. Düzgün Çaprazlama
Şekil 2.7. Gerçek Kodlu GA'da mutasyon
Şekil 2.8. İkili Kodlu GA'da mutasyon
Şekil 3.1. Y vektörünün sütun uzayına dik izdüşümü
Çizelge 4.1. Örnek1 için üretilen değerler
Çizelge 4.2. Örnek1 verisi için parametre tahmin değerleri
Çizelge 4.3. Örnek 2 için parametre tahmin değerleri
Çizelge 4.4. Çok değişkenli model parametre tahmin sonuçları

SİMGELER VE KISALTMALAR

∂ : Kısmi türev operatörü
$E(\cdot)$: Beklenen değer
$\sum$: Varyans-kovaryans matrisi
I : Birim matris
$\otimes$: Kroneker Çarpım
$|\cdot|$: Mutlak değer
$\|\cdot\|$: Norm
$[\![\cdot]\!]$: Tamdeğer
$A_{m\times n}$: *mxn* boyutlu matris
$A'_{m\times n}$: A matrisinin transpozu (devriği)
$A^{-1}{}_{m\times n}$: A matrisinin tersi
GA : Genetik Algoritma
EKK : En Küçük Kareler

TEŞEKKÜR

Bu çalışmanın konusunun seçimini ve gerçekleşmesini sağlayan değerli hocam Sayın Doç.Dr. Aşır GENÇ'e çalışmamda bana yardımcı olan Doç. Dr. Coşkun KUŞ, Yrd.Doç.Dr. İsmail KINACI ve Tarık YILMAZ'a teşekkür eder, benden hiçbir şekilde desteklerini esirgemeyen aileme ayrıca teşekkürü bir borç bilirim.

1.GİRİŞ

Lineer olmayan modelde Y bağımlı değişken, $\underline{X}_1,\underline{X}_2,\ldots,\underline{X}_k$ açıklayıcı (bağımsız) değişkenler $\underline{X}=(\underline{X}_1,\underline{X}_2,\ldots,\underline{X}_k)'$ açıklayıcı değişkenlerin vektörü $\underline{\theta}=(\theta_1,\theta_2,\ldots,\theta_p)'$ $(\underline{\theta}\in\Theta\subset R^p)$, bilinmeyen parametre vektörü ve e_t gözlenemeyen veya deneysel hata terimi olmak üzere e_t'lerin bağımsız, aynı dağılıma sahip, beklenen değerleri sıfır ve varyanslarının eşit olduğu varsayılır. f, $\underline{\theta}$ bilinmeyen parametre vektörünün bileşenlerinin en az birine göre lineer olmayan bir fonksiyon olmak üzere lineer olmayan model,

$$Y_t = f(\underline{X};\underline{\theta}) + e_t, \quad t = 1,2,\ldots,n \tag{1.1}$$

biçiminde ifade edilir. Burada f biçimsel olarak bilindiğinden,

$$Q(\underline{\theta}) = \sum_{t=1}^{n}\left(Y_t - f(\underline{X};\underline{\theta})\right)^2 \tag{1.2}$$

hata kareler toplamı minimum olacak şekilde $\underline{\theta}$ parametre değerini belirlemek en küçük kareler (EKK) yöntemine göre en iyi tepki fonksiyonunu bulmak demektir. $Q(\underline{\theta})$'yı minimum yapmak için kullanılan algoritmanın ismine bağlı olarak tahmin yöntemi de aynı ismi almaktadır. Gauss-Newton algoritması(Gallant 1987), Marquardt algoritması(Marquardt 1963) ve En hızlı iniş algoritması(Milliken 1988) EKK yönteminde kullanılan başlıca algoritmalardır.

Eşitlik (1.2) ile verilen hata kareler toplamını $\underline{\theta}$'ya göre minimum yapacak değeri bulmak amacıyla kullanılabilecek teknikler 1)Analitik, 2)Sayımlama ve 3)Rassal teknikler olarak üç ana sınıfta ayrılabilir.

Analitik teknikler dolaylı ve doğrudan analitik teknikler olarak iki alt sınıfa ayrılabilir. Dolaylı analitik teknikler (1.2) eşitliğinin $\underline{\theta}$'nın bileşenlerine göre kısmi türevlerinin (gradyentinin) sıfıra eşitlendikten sonra elde edilen doğrusal olmayan denklem sisteminin çözülmesiyle en iyi çözümü bulmayı amaçlar. Doğrusal modeller söz konusu olduğunda dolaylı analitik teknikler kullanılarak en iyi çözümler bulunabilir. Ancak doğrusal olmayan modeller söz konusu olduğunda (1.2)

eşitliğinin birinci kısmi türevi sonucunda elde edilen eşitlikler sisteminin çözülmesi problem oluşturabilir. Bu nedenle doğrusal olmayan modeller söz konusu olduğunda çeşitli iteratif teknikler kullanılır. Bu amaçla kullanılan iteratif teknikler doğrudan analitik teknikler sınıfındandır(Goldberg 1989).

Yönlendirilmiş arama teknikleri olarak isimlendirilen bu teknikler belirli bir noktadan başlayarak her bir iterasyonda bu nokta üzerinde gelişme sağlayarak başka bir noktaya geçmeye dayalıdır. Kullanılacak başlangıç noktası araştırıcı tarafından tanımlanır(Koç 2001). Ancak başlangıç noktasının seçimi çözümü doğrudan etkilemektedir. Bu sebeple başlangıç noktasına göre çözümün duyarlılığı fazladır. Bu teknikler gerçek çözümün uzak bir komşuluğunda seçilecek bir başlangıç noktasına göre yerel bir en iyi noktasına takılma riskini bünyesinde barındırırlar.

Sayımlama Teknikleri çözüm uzayının tamamında veya kısıtlanmış bir kısmındaki bütün noktaların amaç fonksiyonundaki değerlerinin test edilmesine dayalıdır. Bu yöntemde sonlu çözüm uzayında veya kesikli sonsuz çözüm uzaylarında iyi çözümler elde edilebilmektedir. Çözüm uzayı küçük olduğunda sayımlama teknikleri ile iyi çözüm elde edilebilirken çözüm uzayının genişlemesi bu yönteminin uygulanabilirliğini kısıtlamaktadır.

Analitik ve sayımlama tekniklerinin belirtilen eksiklikleri rasgele teknikleri popüler hale getirmiştir. Rasgele teknikler de tümüyle rasgele ve rasgeleleştirilmiş teknikler olarak iki alt sınıfa ayrılır. Tümüyle rasgele teknikler bir başlangıç noktasından başlayarak her bir iterasyonda başka bir noktaya rasgele olarak geçerler ve o ana kadar ki en iyi çözümü saklarlar. Çözüm uzayının geniş olduğu durumlarda bu yöntemin performansının düşük olduğu bilinmektedir. Rasgeleleştirilmiş teknikler, rasgele ve yönleştirilmiş tekniklerin birleştirilmesiyle oluşur. Genetik algoritmalar rasgeleleştirilmiş teknikler sınıfında yer alır(Goldberg 1989).

Lineer olmayan modeller üzerindeki çalışmalar 1960'larda Marquardt(1963), Hartley ve Booker(1963)'ın en küçük kareler yöntemi ile lineer olmayan modellerde parametre tahmini için önerdikleri algoritmalar öncülüğünde gerçekleşmiştir. De Bruin(1971), Gallant(1975) ile Seber ve Wild(1989) lineer olmayan modellerde parametre tahmini ve hipotez testi üzerinde çalışmışlardır. Bard(1964),

Gallant(1975), Bates ve Watts(1988), Seber ve Wild (1989) çok değişkenli lineer olmayan modellerde parametre tahmini ve sonuç çıkarımı üzerindeki çalışmalar yapmışlardır(Genç 1997).

Bruin(1971)lineer olmayan modeller için hipotez testi ve test istatistikleri üzerinde çalışmış, parametre tahminleri için modifiye edilmiş Gauss-Newton algoritmasını kullanmıştır.

Goldfeld ve Quandt(1972) hesaplama zorlukları içerdiği için en çok olabilirlik ve genelleştirilmiş en küçük kareler tahmin yöntemlerini kullanmaktan kaçınmıştır. En hızlı iniş, Newton ve gradyent metodlarının karşılaştırılması sonucunda fonksiyonların birkaç kez farklı başlangıç noktalarına göre çözülmesiyle optimum çözüme ulaşılabileceğini belirtmiştir. Bu farklı çözümler eğer aynı optimum noktayı belirttiği takdirde bulunan noktanın global nokta olarak değerlendirilebileceğini belirtmişlerdir.

Bird(1974) lineer olmayan modellerde tahmin sorununa ilişkin modifiye edilmiş Gauss-Newton yöntemini ve yeniden parametrelendirme ve tahmin sorununu incelemiştir.

Günay(1978) lineer olmayan regresyon modellerinde en küçük kareler yöntemiyle parametre tahmini için özdeğer ve öz vektörlere dayalı tahmin yöntemi ile hartley yönteminin etkinliğini incelemiştir.

Tez(Aysever)(1982) lineer olmayan modelde parametrelerin kestirimi için pratikte yaygın olarak kullanılan Gauss-Newton, En hızlı iniş ve Marquart yöntemlerini ele almıştır.

Seber ve Wild(1989) lineer ve lineer olmayan regresyon modellerini tanıtmış, doğrusal ve doğrusal olmayan regresyon modelleri için çeşitli tahmin yöntemlerini vermiştir. Eğrilik ve doğrusal olmamanın ölçüsü ve istatistiksel sonuç çıkarımı ve güven aralıklarını incelemiştir. Çok değişkenli lineer olmayan modeller için lineer olmayan en küçük kareler yöntemi ve Gauss-Newton algoritmasını vermiştir.

Siepman ve Yang(1994) eksik gözlemli verilerde çok değişkenli lineer olmayan modeller için genelleştirilmiş en küçük kareler tahmini üzerinde çalışmıştır.

Genç(1997) çalışmasında tek değişkenli doğrusal olmayan regresyon modelleri ve tek değişkenli doğrusal olmayan regresyon modellerinde parametre tahmin yöntemlerini tanıtmış, tahmin edicilerin özellikleri ve hipotez testlerini inceledikten sonra çok değişkenli doğrusal olmayan regresyon modelleri için parametre tahmin yöntemlerinden en küçük kareler yöntemi, en çok olabilirlik yöntemi ve hipotez testleri konuları incelemiştir.

Genetik algoritmalar üzerindeki çalışmalar Holland(1975), De Jong(1975) ve Goldberg(1989) öncülüğünde başlamıştır. Literatürde optimizasyon alanında birçok uygulamaya sahip olan genetik algoritmaların lineer olmayan modeller üzerindeki uygulamaları diğer alanlardaki uygulamalarla karşılaştırıldığında oldukça azdır. Çok değişkenli lineer olmayan modeller üzerindeki bir uygulamaya ise yapılan literatür çalışmasında rastlanmamıştır.

Eter ve ark.(1982) lineer bir otoregresif modelin parametrelerini tahmin etmek amacıyla genetik algoritmayı kullanmıştır.

Karr ve ark.(1991) üç adet regresyon modeli için parametre tahminlerini, hata kareler toplamını uygunluk fonksiyonu olarak kullanıp, genetik algoritma ile elde etmiştir.

Buckley ve Hayashi(1994) bulanık lineer regresyon modelinin parametrelerini tahmin etmek amacıyla bulanık genetik algoritma kullanmıştır.

Leehter ve Sethares(1994)lineer ve lineer olmayan filtrelerde parametre tahmin problemi için modifiye edilmiş genetik algoritmayı kullanmış ve elde edilen tahmin hatasının olasılıkta sıfıra yakınsadığını göstermiştir.

Gülşen ve ark.(1995) 3 adet regresyon modelinin parametrelerini genetik algoritma ile tahmin etmiştir. Çalışmada kullanılan modellerden biri tek değişkenli lineer regresyon modeli olup bu model Karr ve ark. Çalışmasında da kullanılmış ve bu çalışmada bulunan sonuçların Karr ve ark.'dan daha iyi olduğu gözlenmiştir.

Hwang ve Woo(1995) Bir bulanık modelin parametrelerini tahmin etmek amacıyla genetik algoritmayı kullanmıştır.

Karr ve ark.(1995) aykırı değerlere duyarlılıkları açısından uygunluk fonksiyonu olarak hata kareler toplamı ve hata kareler medyanını karşılaştırmıştır.

Moros ve ark.(1996) lineer olmayan bir modelin parametrelerini genetik algoritma kullanarak tahmin etmiştir.

Gülşen, M.(1997) veri setine en uygun fonksiyon seçilmesi problemine, iki aşamalı bir genetik algoritmayı uygulamıştır.

Ahmad ve ark.(1998) genetik algoritma kullanarak bir ısıtma sisteminin parametrelerini tahmin etmiş ve elde edilen sonuçlar tekrarlamalı en küçük kareler yöntemiyle elde edilen sonuçlarla karşılaştırılmıştır.

Petridis ve ark.(1998) lineer olmayan dinamik bir sistemin parametrelerini tahmin etmek amacıyla genetik algoritmayı kullanmıştır.

Gülşen ve Smith(1999) modellerin fonksiyonel formlarının genetik algoritma ile belirlenmesi üzerine bir çalışma yapmıştır. Farklı formlardaki fonksiyonların parçalarının genetik algoritma ile seçilerek veri kümesine uygun modelin seçiminin amaçlandığı çalışmada algoritmanın çalışma süresi uzun olduğu için paralel bilgisayarlar kullanılmıştır.

Chatterjee ve ark.(1996) değerlendirme fonksiyonu olarak hata kareler toplamını kullanarak doğrusal ve lojistik regresyon modelin parametrelerini tahmin etmiştir.

Hasançebi ve Erbatur(2000) ikili kodlamada kullanılan GA'nın yerel en iyi noktaya takılma riskini azaltmak amacıyla iki çaprazlama modeli önermiştir. Bunlardan birincisi karma çaprazlama, ikincisi karar değişken değişim çaprazlamasıdır. Bu çaprazlama tekniklerinin tek nokta ve iki nokta çaprazlamaya göre daha iyi sonuçlar verdiği gözlenmiştir. Ancak gerçek kodlu GA'lar için bu çaprazlama tekniklerinin uygun olmadığı belirtilmiştir.

Koç(2001) zorluk derecelerine göre sınıflandırılmış lineer olmayan 20 regresyon modelinin genetik algoritma ile çözümüne yer vermiştir. İkili kodlamanın kullanıldığı genetik algoritma ile elde edilen parametre tahmin sonuçları Levenberg-

Marquardt algoritmasıyla elde edilen parametre tahmin sonuçlarıyla karşılaştırılmıştır.

Yang ve ark.(2002) lineer olmayan tahmin yöntemlerinin başlangıç noktalarına olan duyarlılıklarına bir çözüm getirmek amacıyla ağırlıklandırılmış mutasyon operatörünü kullanarak elektrokimyasal özdirenç verilerinde parametre tahmini için başlangıç noktası tahmininde genetik algoritma ve Gauss Newton metodunu beraber kullanmıştır.

Altunkaynak ve Esin(2004) GA kullanarak S biçimli büyüme eğrilerinde parametre tahminlerini 5 farklı model için 3 farklı veri setiyle gerçekleştirmiştir. İkili kodlamanın kullanıldığı çalışmada sıralı seçim, tek nokta çaprazlama ve mutasyon operatörü kullanılmış. GA ve Gauss Newton metoduyla birbirine çok yakın tahminler elde edilmiştir.

Olinsky ve ark.(2004) lineer olmayan tek değişkenli bir modelin parametrelerini gauss Newton algoritması ve ikili kodlu GA ile tahmin etmiş ve tahmin sonuçlarını binde bir hassasiyetle birbirlerine yakın bulunmuştur.

Hwang ve He(2006) çalışmalarında GA'ların kodlama ve deşifre işlemlerindeki zaman kaybı ve yerel en iyi noktaya takılma riskine karşı gerçek kodlama ile tavlamalı benzetimin özelliklerini taşıyan ARSAGA isimli bir algoritma önermiştir. 16 fonksiyon üzerinde ve iki optimizasyon problemi üzerindeki uygulamalarda farklı populasyon büyüklükleri için arama uzayını tarama ve yakınsama hızı açısından elde edilen sonuçlar karşılaştırılmıştır.

Chang(2006) çoklu çaprazlamalı gerçek kodlu GA kullanarak lineer olmayan iki bileşenli sistemin parametre tahmini üzerinde çalışmış. Elde edilen sonuçlar standart GA sonuçları ile karşılaştırılmış ve çoklu çaprazlamanın daha iyi sonuç verdiği gözlenmiştir.

Eşitlik (1.2) ile verilen $Q(\underline{\theta})$, $\underline{\theta}$'nın lineer olmayan bir fonksiyonudur. Lineer olmayan bu fonksiyonu $\underline{\theta}$'nın değerlerine göre minimum yapmak için Gauss-Newton, Marquardt ve En hızlı iniş algoritmalarının sıklıkla kullanılmaktadır. Ancak bu algoritmaların kullanılabilmesi için bazı şartların sağlanması gerekmektedir. Çok

değişkenli lineer olmayan modelde Gauss-Newton algoritması ile parametre tahmininin yapılabilmesi için bağımsız değişken(ler)in oluşturduğu tepki fonksiyonu $f(\underline{X};\underline{\theta})$ 'nın $\underline{\theta}$'nın bileşenlerine göre en az iki kez türevlenebilir olması, dolayısıyla $f(\underline{X};\underline{\theta})$ 'nın sürekli olması gerekmektedir. Gauss-Newton algoritması başlangıç noktası adı verilen ve genellikle araştırıcı tarafından belirlenen $\underline{\theta_0}$ noktasından başlayarak noktasal bazda arama yapmaktadır. İteratif olarak işleyen süreç her adımdaki nokta üzerinde gelişme sağlayarak $Q(\underline{\theta})$ 'yı minimum yapan $\underline{\theta^0}$ gerçek noktasını bulmayı amaçlamaktadır. Ancak Goldfeld ve Quandt(1972)'ın çalışmalarında da belirtildiği gibi gerçek çözümün uzağında seçilecek bir başlangıç noktasına göre çözümün geç bulunması veya bulunamaması riski vardır.

Bu tez çalışmasında Gauss-Newton algoritması ve benzer mantıkla çalışan Marquardt ve En hızlı iniş algoritmalarına alternatif bir yöntem olarak genetik algoritma (GA) ile çok değişkenli lineer olmayan modelin parametre tahmini önerilmiştir. Yapılan literatür taramasında "multivariate nonlinear model" ve Türkçe "Çok değişkenli lineer olmayan model" konu başlıklarında yapılan taramalar sonucunda, genetik algoritma ile çok değişkenli lineer olmayan modelin parametre tahminin yapıldığı bir çalışmaya rastlanmaması çalışmanın önemini artırmaktadır. GA'nın önerilmesindeki en önemli etken $f(\underline{X};\underline{\theta})$ 'nın $\underline{\theta}$ parametre vektörünün bileşenlerine göre türevlenebilir olma şartını gerektirmemesidir. GA'nın ikinci bir avantajı ise çözümü noktasal bazda değil potansiyel çözümlerin oluşturduğu noktalar kümesi üzerinde aramasından dolayı Gauss-Newton algoritmasındaki başlangıç noktasına göre duyarlılığının GA'da bir sorun teşkil etmemesidir. Yapılan uygulamalar için NLGA1 ve NLGA2 isimli iki algoritma sırasıyla tek değişkenli ve çok değişkenli lineer olmayan modellerde parametre tahmini için önerilmiştir. Önerilen algoritmada elitist stratejiye yeni bir yaklaşım getirilmiş, tek nokta çaprazlama ve mutasyon operatörü önceki kullanımlarından farklılık gösterecek şekilde tanımlanıp kullanılmıştır. Yapılan uygulamalarda NLGA1 ve NLGA2 ile elde edilen parametre tahmin değerlerinin Gauss-Newton algoritması ile elde edilen parametre değerlerine göre daha küçük hata kareler toplamına sahip olduğu görülmüştür.

Bu tez çalışması beş bölümden oluşmaktadır. Tezin birinci bölümü giriş bölümü lineer olmayan modellerde parametre tahmini için kullanılan karesel formun optimizasyonu için kullanılan yöntemler ve sınıflandırılmasıyla ilgilidir. Ayrıca konu ile yapılan çalışmalara ilişkin literatür taramasını içermektedir.

Tezin ikinci bölümde genetik algoritmalar hakkında bilgi verilmiştir. Bu bölümde genetik algoritmalarda kullanılan başlıca terimler, kodlama yöntemleri ve genetik algoritmaların operatörleri tanıtılmıştır.

Tezin üçüncü bölümünde lineer olmayan model tanıtılmış ve lineer olmayan modelin parametrelerinin tahmini için kullanılan yöntemler verilmiştir. Bu bölümde tezin esas konusunu oluşturan çok değişkenli lineer olmayan model tanıtılmış ve çok değişkenli lineer olmayan modelin parametre tahmini için yöntemler belirtilmiştir.

Tezin esasını dördüncü bölüm oluşturmaktadır. Bu bölümde tek değişkenli lineer olmayan modellerde parametre tahmini için Delphi 5 görsel programlama dilinde kodlanmış NLGA1 isimli genetik algoritma verilmiştir. Ayrıca çok değişkenli lineer olmayan modelin parametre tahmini için yine Delphi 5 görsel programlama dilinde kodlanmış NLGA2 isimli genetik algoritma verilmiştir.

Tezin altıncı bölümde NLGA1, NLGA2 ve Gauss-Newton algoritmaları kullanılarak simülasyonla üretilen veri seti üzerinde 3 uygulama yapılmıştır ve elde edilen sonuçlar karşılaştırılmıştır.

2. GENETİK ALGORİTMALAR

Genetik algoritmalar doğal seçim ve genetik kavramlarına dayalı rasgeleleştirilmiş arama metotlarıdır. Genetik algoritmaların temeli Charles Darwin'in görüşlerine dayandırılan doğayı taklit etme ve bunun sonucu olarak güçlü olan bireylerin yaşaması (survival of fittest) prensibine dayanır (Michalewicz 1996). Bu prensibe göre bütün canlılar, yaşadıkları doğal ortamın koşullarına aynı ölçüde dayanıklı değildir, bu durumda dayanıklı olmayanlar ortam tarafından ayıklanır, dayanıklı olanlar ise varlıklarını sürdürürler. Buna "doğal seçilim" veya "doğal ayıklanma" mekanizması denir.

Genetik Algoritma(GA) kavramı literatüre 1960 yılında Michigan Üniversitesinden Prof. John Holland tarafından kazandırılmış ve 1970'lerde kendisi ve öğrencileri tarafından geliştirilmeye başlanmıştır. Holland'ın (1975) "Adaptation in Natural and Artificial Systems" adlı eserinde iki temel başlık olarak doğal sistemlerin uyum sürecini anlayıp özetlemek ve doğal sistemlerin önemli mekanizmalarını içeren yapay sistemlerin yazılım olarak tasarlanması amaçlanmıştır (Goldberg 1989).

John Holland'ın çalışması günümüzde Evrim stratejileri, Evrimsel programlama, yapay zeka, sınıflama sistemleri ve Genetik programlama gibi çalışma alanlarının açılmasına öncülük etmiştir(Reeves ve Rowe 2002).

John Holland'ın çalışmasına paralel olarak De Jong(1975) genetik algoritmaların en iyileme amaçlı kullanılabileceğini göstermiştir. Önceleri konu üzerinde benzer çalışmalar yapılsa da Holland'ın öğrencisi David Goldberg'in (1980) "Adaptive control of gas pipeline systems" isimli tezinde gaz boru hattının kontrolüne ilişkin bir problem genetik algoritma kullanılarak çözülmüştür. Goldberg'in optimizasyon problemlerinde kullanımı ile ilgili "Genetik Algorithms in Search, Optimization and Machine Learning" isimli eseri genetik algoritmaların uygulama alanlarının gelişmesini sağlamıştır(Michalewicz 1996).

2.1. Genetik Algoritma Terminolojisi

Genetik Algoritma terminolojisinin anlaşılmasında genetik bilimiyle ilgili temel kavramları bilmekte yarar vardır. Canlılar hücrelerden ve her bir hücre de bir ya da daha fazla sayıda kromozomdan oluşur. Yaşayan her organizma türü belirli sayıda kromozomdan oluşur. Örneğin İnsan 46 kromozomdan oluşurken alabalıkta 94 adet kromozom vardır(Bäck 1996).

Gen : Her kromozom gen adı verilen temel birimlerden oluşur. Genler belirli özelliklerin kalıtımını kontrol eder. Belirli bir özelliğe ait genler kromozomların belirli kısımlarında bulunur.

Lokus : Genlerin kromozomlar üzerinde bulunduğu yerin pozisyonu lokus olarak adlandırılır.

Allel : Genler her biri farklı karakteristiği temsil eden iki fonksiyonel formda meydana gelir. Bu formların her biri "allel" olarak bilinir. Allellerden biri baskın olurken diğeri çekinik kalır. Bir genin alleli o genin aldığı değerin göstergesidir.

Genotip : Kromozomlardaki allellerin kombinasyonları, fertlerin kişisel özelliklerini belirler. Örneğin allel değerine göre insanın saç renginin siyah veya sarı olması belirlenir. Allellerin oluşturdukları kombinasyonlara genotip denir.

Fenotip : Genotip sonucunda oluşan görünüm, örneğin insanın saç renginin siyah veya sarı olması fenotip olarak adlandırılır.

Birey(Kromozom) : Genetik algoritma kullanılarak elde edilecek her bir çözüm birey ya da kromozom olarak adlandırılır.

Populasyon : Bireylerin (kromozomların) temsil ettiği çözümlerin bir kümesine populasyon denir.

Kodlama : GA'da populasyonu oluşturan her birey için fenotip-genotip geçişinin simgelerle ifade edilmesidir.

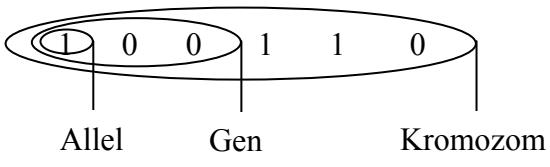
Genetik algoritmalar, biyolojik süreci yapay bir sistemde modelleyerek fonksiyonları optimize eden algoritmalardır. GA'da oluşturulan bir populasyondaki her birey eşit sayıda gen içeren bir kromozomdan oluşur. Bireyler populasyonda

kromozomları ile temsil edilir. Her bir birey ilgilenilen probleme ilişkin bir çözümü ifade eder. Her bir kromozom probleme ilişkin çözümün kodlanmış halini belirtir. Populasyonun uygunluğu, belirli kurallar dahilinde değerlendirilir (Koç 2001).

GA'ların evrim süreci potansiyel çözüm uzayı boyunca kromozom populasyonlarının GA'nın operatörlerinin işletilmesiyle gerçekleştirilir. GA'nın kullandığı operatörler Kopyalama, Çaprazlama ve Mutasyon operatörleridir. Bu operatörler yardımıyla GA, ilgilenilen problemin çözümü için çözüm uzayının taranması ve en iyi çözümün bulunmasını amaçlar(Goldberg 2004).

Evrim sürecinde her bir kromozomun uygunluk derecesi değerlendirme fonksiyonu (amaç fonksiyonu) ile belirlenir. Değerlendirme fonksiyonu problemin yapısına göre maksimum ya da minimum yapısında bulunur. Değerlendirme fonksiyonundaki değerine göre her bir kromozoma bir uygunluk değeri karşılık gelir.

Populasyon için hesaplanan uygunluk değerlerine göre kromozomlar iyi veya kötü olarak değerlendirilebilir. İyi bireylerin yaşamlarını devam ettirme olasılıkları ve kötü olan bireylerin ölme olasılıkları artırılarak iyi bireylerin çoğalmaları ve kötü bireylerin yok olmaları beklenir. Evrim sürecinde GA'nın evrim operatörleri yardımıyla kromozomlar farklı bir populasyon oluşturur ve bir iterasyon tamamlanır. Belirli iterasyon sayısına ulaşınca veya durdurma koşulu sağlandığında son populasyondaki en iyi uygunluk değerine sahip kromozomun kodladığı çözüm, problemin çözümü olarak belirlenir(Mitchell 1999).



Şekil 2.1. GA'da Kromozomun genel yapısı

Kromozomlar problemin çözümü için gerekli bilgilerin saklandığı yapılardır. Bir kromozomun genel yapısı Şekil 2.1.'de verilmiştir. Bir kromozomu oluşturan toplam gen sayısı kromozomun uzunluğunu belirtir. Kromozomun uzunluğu problemdeki parametre sayısına bağlıdır. Problemde değeri araştırılan tüm parametrelerin yan yana kodlanması bir kromozomu, dolayısıyla problem için bir çözümün kodlanmış halini temsil eder(Reeves 2002).

Genetik algoritmaların çalışma prensibine uygun işlem adımları aşağıda verilmiştir.

Adım1 : Başlangıç değerlerini belirle (Parametre Değişim Aralıkları, Amaç Fonksiyonu, Çaprazlama olasılığı, Mutasyon olasılığı, Populasyon büyüklüğü)

Adım 2 : Başlangıç populasyonunu oluştur.

Adım 3 : Populasyon uygunluk değerlerini hesapla

Adım 4 : Sonlandırma koşulu sağlandıysa Adım 8'e, sağlanmadıysa Adım 5'e git.

Adım 5 : Kopyalama operatörünü uygula

Adım 6 : Çaprazlama operatörünü uygula

Adım 7 : Mutasyon operatörünü uygula

Adım 8 : En iyi çözümü yaz

Adım 9 : Programı sonlandır.

2.2. Genetik Algoritmaların Geleneksel Arama Metotlarından Farkı

Goldberg 1989'e göre GA'ların geleneksel arama metotlarından farkı aşağıdaki şekilde verilebilir.

- Genetik algoritmalar parametrelerin kendisi ile değil parametreler kümesinin kodlanması ile çalışır.
- Genetik algoritmalar arama aracı olarak tek bir noktayı değil, noktalar kümesini kullanır. Bu sebepten yerel en iyiye takılma riskleri düşüktür.

- Genetik algoritmalar arama sırasında probleme ilişkin türevlenebilme koşulu veya diğer ek bilgilere ihtiyaç duymaz, amaç fonksiyonunun bilinmesi genetik algoritmalar için yeterlidir.
- Genetik algoritmalar belirli geçiş kurallarını değil stokastik geçiş kuralları ile çalışır.

2.3. Genetik Algoritmaların Genel Yapısı

Genetik algoritmaların yapısı son derecede basittir. Dizilerin kopyalanması ve kısmen değiştirilmesinden daha karışık değildir. Basitlik ve etkisinin gücü genetik algoritmaların en önemli özelliklerindendir. Genetik algoritmaların bir problemin çözümünde kullanılabilmesi için aşağıdaki beş bileşenin sağlanması gerekir. (Goldberg 1989)

- Problem için potansiyel çözümlerin genetik gösterimi (Çözümlerin kromozomlarla ifade edilmesi)
- Potansiyel çözümlerin (başlangıç populasyonunun) oluşturulması için bir yol
- Çözümlerin uygunluğunu değerlendiren değerlendirme fonksiyonunu
- Genetik operatörlerin (Kopyalama, Çaprazlama ve Mutasyon) belirlenmesi
- Genetik algoritmanın kullanacağı parametre değerlerinin belirlenmesi (Popülâsyon büyüklüğü ve genetik operatörlerin değerleri)

2.4. Genetik Kodlama ve Uygunluk Fonksiyonu

Genetik algoritmanın bir problemin çözümünde kullanılabilmesi için öncelikle problemin çözümünün kromozomlarla ifade edilebilmesi gerekir. Kromozomlar populasyonu oluşturan bireyleri temsil ederler ve genlerden oluşurlar. Kromozom gösterimi parametrelerin değerlerinin kodlanmasından oluşur. Kodlanmış kromozomların deşifre edilmesiyle parametrelerin gerçek değerlerine ulaşılır.

Kodlama ve deşifre işlemi biyolojik olarak genotip ve fenotip arasında geçişlerin sağlanmasına karşılık gelir. Genotip-fenotip arasındaki geçişi f_g ve fenotipten uygunluk değerine geçişi f_p göstermek üzere kodlama için iki aşamalı bir yol izlenebilir.

Φ_g, çaprazlama ve mutasyon gibi genetik operatörlerin uygulanacağı genotip arama uzayını göstermek üzere kesikli yada sürekli olarak tanımlanabilir.

$$f(x) : \Phi_g \rightarrow \Re$$
$$x \rightarrow f(x)$$

ile verilen fonksiyon Φ_g arama uzayındaki her elemanın reel sayılardaki karşılığını belirtir.

$$\hat{x} = \min_{x \in \Phi_g} f(x)$$

ile verilen bir optimizasyon problemi için *x*, karar değişkenler vektörü (alleller) ve *f*(*x*) ise amaç fonksiyonu ya da uygunluk fonksiyonu olarak adlandırılır, $\hat{x}$ ise yerel minimum noktayı ifade eder. Kodlama işleminde en önemli nokta problem için arama uzayını en iyi temsil eden kodlamanın seçilmesidir.

Uygunluk fonksiyonu *f* genotip arama uzayından fenotip arama uzayına geçiş (kodlama) ve fenotip uzaydan uygunluk uzayına geçiş (deşifre) olmak üzere iki kısma ayrılabilir.

Φ_p, fenotip uzay olmak üzere,

$$f_g : \Phi_g \rightarrow \Phi_p \qquad f_p : \Phi_p \rightarrow \Re$$

$$f = f_p \circ f_g = f_p(f_g)$$

dir. Genotip-fenotip geçişi f_g, kullanılan kodlama tipi tarafından belirlendiği için kodlama olarak adlandırılırken, f_p uygunluk fonksiyonu çözümün deşifre edilmiş halini belirtir.

Genetik operatörler genotiplere uygulanırken, bireylerin uygunluk değerleri fenotiplerden yararlanılarak hesaplanır(Reeves 2002).

2.4.1. İkili kodlama

İkili kodlama GA'ların ilk çıkışından beri kullanılan en yaygın kodlama biçimidir. İkili kodlama yöntemine göre her bir kromozom 0 veya 1 değerlerinden oluşan sayı dizilerinden oluşur. GA'da bir dizinin uzunluğu dizideki toplam karakter sayısı ile ifade edilir. Örneğin 10 karakterden oluşan bir kromozom için "10 bitle kodlanmış" veya "10 bitlik "ifadesi kullanılır.

Parametrelerin değişim aralıklarıyla belirlenen sınırlı parametre uzayında değerlendirme fonksiyonunun sürekli parametre değerlerine ihtiyaç vardır. Değerlendirme fonksiyonunun hesaplanması için her seferinde kromozomun onluk sayı sistemine çevrilmesi gerekir. Bu gereklilik ikili kodlamanın dezavantajlarından biridir ve işlem yükünün artmasına sebep olmaktadır. İkili kodlamanın diğer bir dezavantajı parametreleri kodlamak için kullanılacak kromozomların uzunluğudur.

Arama uzayı $\Phi_g = \{0,1\}^L$ olarak ifade edilmek üzere L uzunluğundaki ikili kodlanmış vektör $x^g = \left(x_1^g x_2^g \cdots x_L^g\right) \in \{0,1\}^L$ ile gösterilir(Goldberg 1989).

İkili kodlama sisteminde genotip fenotip geçişi amaç fonksiyonu tarafından belirlenir. Bir çok optimizasyon probleminde genotip fenotip geçişi doğrudan sağlanmaktadır(Haupt 1998).

Değişim aralıkları alt_i ve $üst_i$ olarak sırasıyla i. parametrenin alt ve üst sınırlarını ifade etmek üzere (h) ondalık hassasiyet değeri ile i. parametreyi kodlamak için gerekli gen sayısı,

$$2^{L-1} \le (üst_i\text{-}alt_i) \times 10^h \le 2^L \tag{2.1}$$

koşulunu sağlayan en küçük L değeri parametreyi kodlamak için gerekli toplam gen sayısını ifade eder. Belirlenen uzunlukta ikili sistemde kodlanan parametrenin onluk sistemde belirlenen hassasiyet değerine göre karşılığının bulunması dizinin deşifre edilmesidir. L uzunluğundaki ikili sistemdeki dizi aşağıdaki şekilde deşifre edilir.

1. $Parametre_{ikili} = (g_0\ g_1\ g_2 \cdots g_{L-2}\ g_{L-1})$ dizisinin onluk karşılığı,

$$Parametre_{onluk} = \sum_{i=0}^{L-1} g_i \times 2^i$$ ile hesaplanır.

2. Parametrenin değişim aralığındaki gerçek değeri,

$$Parametre_{gerçek} = alt_i + Parametre_{onluk} \times \frac{üst_i - alt_i}{2^L - 1}$$

eşitliği ile belirlenir(Rothlauf 2006).

Örnek : X_{ikili}= 10011101101011 şeklinde kodlanan L=14 uzunluğundaki dizinin [1 , 12] aralığındaki karşılığı,

$$\begin{aligned} X_{onluk} &= 1\times 2^0 + 0\times 2^1 + 0\times 2^2 + 1\times 2^3 + 1\times 2^4 + 1\times 2^5 + 0\times 2^6 + 1\times 2^7 + \\ &\quad 1\times 2^8 + 0\times 2^9 + 1\times 2^{10} + 0\times 2^{11} + 1\times 2^{12} + 1\times 2^{13} \\ &= 13753 \end{aligned}$$

ve

$$\begin{aligned} X_{gerçek} &= -1 + 13753 \times \frac{11}{2^{14} - 1} \\ &= 8.23 \end{aligned}$$

olarak bulunur.

2.4.2. Gray kodlama ve hamming uzaklığı

İkili kodlamada hamming uzaklığı problemi fenotipleri birbirine yakın kromozomların genotipleri arasındaki farkın büyük olmasıdır. Hamming uzaklığı iki kromozomun farklılık gösteren bit sayıları olarak tanımlanır (Taub ve Shilling 1986).

İkili kodlamanın hamming uzaklığından kaynaklanan eksikliğini gidermek amacıyla Gray kodlama önerilmiştir(Haupt, 1998).

[0-7] aralığındaki tamsayıların ikili ve gray kodlama değerleri çizelge 2.1.'de, hamming uzaklıkları ise Şekil 2.1'de verilmiştir.

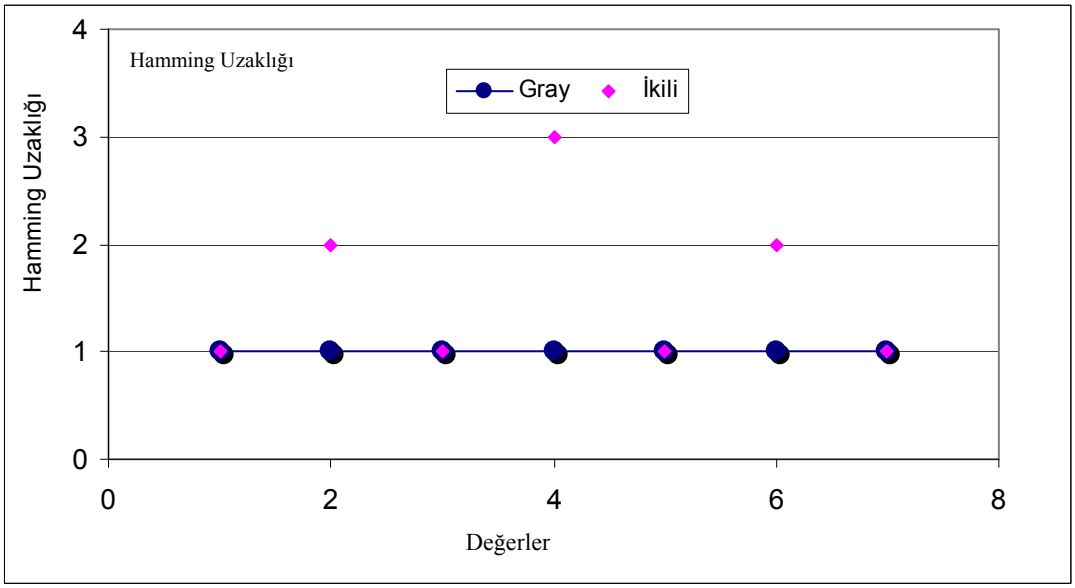
$$g_1 = b_1$$

$$g_k = b_{k-1}\bar{b}_k + b_k\bar{b}_{k-1} \qquad k = 1,2,\ldots,L \qquad (2.2)$$

İkili kodlama ile gray kodlama arasındaki geçiş için (2.2) eşitlikleri kullanılmaktadır. Burada, b_k ifadesindeki k indisi ikili kodlanan kromozomdaki bitin sırasını, b_1 en solda kodlanan biti, $\bar{b}_k$ da b_k'nın değilini (1 ise 0, 0 ise 1'i) ifade eder. (+) işlemi mantıksal işlemlerden VEYA' ya karşılık gelirken, çarpımlar mantıksal VE işleminin karşılığıdır.

Çizelge 2.1. İkili ve Gray Kodlama Gösterimi

Değer	İkili kodlama	Gray Kodlama
0	000	000
1	001	001
2	010	011
3	011	010
4	100	110
5	101	111
6	110	101
7	111	100



Şekil 2.1. İkili ve Gray Kodlama için Hamming uzaklıkları

Hamming uzaklığının önemi özellikle çaprazlama operatörünün uygulanışında ortaya çıkmaktadır. Uygunluk değeri birbirine yakın ancak hamming uzaklıkları açısından çok farklı bireylerin çaprazlanması sonucu oluşacak bireyler, her iki ebeveynden oldukça farklı olacağından optimum çözüme yakınsamayı geciktirecektir.

2.4.3. Gerçek kodlama

Gerçek kodlama GA'ların son yıllardaki uygulamalarında sıkça kullanılmaktadır. İkili kodlamanın fenotip – genotip arasındaki geçişlerdeki zaman kaybı gerçek kodlama ile giderilmeye çalışılmıştır. Gerçek değerli kodlamada kromozomlar parametrelerin kodlanmış hali yerine doğrudan parametrelerin gerçek değerlerinden oluşur.

Gerçek değerli kodlamada arama uzayı $\Phi_g = \Re^L$ 'dir. Burada L gerçek değerle kodlanan kromozom uzunluğudur. Örneğin iki parametreli bir problemin çözümünü temsil eden bir kromozom için kromozom uzunluğu L=2 ve arama uzayı da $\Re^2$' dir(Rothlauf 2006).

İkili kodlu GA'da çaprazlama operatörü esas operatör olarak kabul edilirken, gerçek kodlu GA'da esas operatör olarak mutasyon operatörü kabul edilir(Ahmad 1998).

2.5. Başlangıç Populasyonu

Geleneksel optimizasyon yöntemlerinde çözümün aranmasına bir noktadan başlanır ve bu noktadan gelişme sağlanarak başka bir noktaya geçiş yapılır. Başlangıç noktasının seçimine göre çözüme ulaşmama veya çözüm süresinin uzaması problemiyle karşılaşılabilir. GA' da ise arama işlemi tek noktada değil noktaların (potansiyel çözümlerin) bir kümesi üzerinde gerçekleştirilir ve tatmin edici bir çözüme ulaşıncaya kadar çözümler değerlendirilir.

Başlangıç popülasyonun oluşturulması populasyon büyüklüğünün belirlenmesi problemiyle başlar, populasyon büyüklüğünün çok küçük seçilmesinin arama uzayının etkin bir biçimde aranmamasına yol açabileceği gibi çok büyük bir popülasyon büyüklüğünün de işlem yükünü artırma riski vardır.

Goldberg(1989)'e göre ikili kodlamanın kullanıldığı populasyonlarda en iyi çözüme ulaşmak için gerekli populasyon büyüklüğünün kromozom uzunluğuna göre üstel olarak arttığı belirtilmiştir.

Populasyon büyüklüğünün 30 olarak seçilmesinin birçok uygulamada yeterli olacağı belirtilmiştir(Reeves 2002).

Populasyon büyüklüğünün alt sınırının belirlenmesinde kullanılabilecek bir ölçütte $N_{pop} \geq (1.65) \times 2^{0.21 \times L}$ eşitliğidir. Burada *L,* bir bireyi kodlamak için gerekli toplam bit sayısıdır(Saruhan ve ark. 2003).

Populasyon büyüklüğü belirlendikten sonra bireyleri temsil eden kromozomlardan oluşan bir başlangıç populasyonu oluşturulur. Literatürde başlangıç populasyonunun rasgele oluşturulması benimsenmişse de bazı durumlarda araştırıcı tarafından belirlenecek aday çözümlerin bir kümesi başlangıç populasyonu olarak kullanılabilir.

2.6. Kopyalama

GA'da kopyalama operatörü, uygunluk değerlerine bağlı olarak populasyondaki bireylerden sonraki nesile aktarılacak bireylerin belirlenmesi işlemidir. Kopyalama işleminde populasyondaki bireyler arasında çeşitliliği sağlamak ve yüksek uygunluk değerlerine sahip bireyleri sonraki nesle aktarmak amaçlanır. Bu nedenle seçim operatörü olarak da isimlendirilir. Kopyalama operatörünün kullandığı yöntemler rulet çemberi yöntemi, turnuva seçim yöntemi ve sıralı seçim yöntemi olarak verilebilir.

2.6.1. Rulet çemberi yöntemi

GA'da en çok kullanılan seçim yöntemlerden biri rulet çemberi yöntemidir. Bu yöntemde her bir birey uygunluk değerleriyle orantılı olarak seçilme şansına sahiptir. Bu yöntemde çember, populasyon büyüklüğü kadar aralığa bölünür. Çemberdeki i. aralık populasyondaki i. bireyi temsil eder. Her bir aralık ilgili bireyin seçilme olasılığına ve çemberdeki aralıkların toplamı da 1'e eşittir. Bu yöntemde seçim işlemi yapılırken çember populasyon büyüklüğü kadar çevrilir. Her defasında [0, 1] aralığından rasgele sayı üretilerek, bu sayının dahil olduğu ilk aralıktaki birey sonraki nesil için aday olarak seçilir(Goldberg 1989).

Rulet çemberi yöntemi basit olmasına karşın bazı bireyler beklenende fazla sayıda sonraki nesil için seçilebilmektedir. Böylelikle teorik olarak tahmin edilenden farklı yönlerde arama yapılarak algoritmanın zamansız yakınsadığı gözlenmektedir(Booker 1987).

De Jong(1975) tarafından önerilen elitist, beklenen değer ve crowding metodları rulet çemberi yönteminin belirtilen eksikliğine çözüm olarak önerilmiştir.

N bireyden oluşan populasyondaki i. bireyin uygunluk değeri $f(i)$ ve populasyondaki bireylerin uygunluk değerleri toplamı $F=\sum_{i=1}^{N} f(i)$ olmak üzere i. bireyin sonraki populasyona seçilme olasılığı, $p(i) = \frac{1}{f(i)} \times \sum_{i=1}^{N} \frac{1}{f(i)}$ ile hesaplanır. $p(i)$ değerleri rulet çemberinde i. bireye ait dilimin büyüklüğüdür.

2.6.2. Turnuva seçim yöntemi

Bu yöntemde populasyon rasgele belirlenen k gruba ayrılır. Uygulamalarda genellikle k=2 alınmasına karşın bu değerin artırılabileceği belirtilmiştir(Goldberg 2004).

Turnuva seçim yönteminde k gruba ayrılan bireyler daha sonra bir turnuvaya katılır ve k grup içerisinden en yüksek uygunluk değerine sahip bireyler seçilir.

Turnuva seçiminin avantajı yüksek uygunluk değerine sahip bireylerin sonraki nesile aktarılması ve gruplardan seçilemeyen bireylerin başka bir gruba dahil olarak seçilme şanslarının ortaya çıkmasıdır.

2.6.3. Sıralı seçim yöntemi

Sıralı seçim yöntemi populasyon çeşitliliğini destekleyen bir yöntemdir. Populasyonda çok sayıda (populasyonun %90'ı civarında) yüksek uygunluk değerine sahip bireyin bulunması, düşük uygunluk değerine sahip bireylerin seçilme şansını düşürecektir. Sıralı seçim yöntemi bireyleri uygunluk değerlerinin büyüklük derecelerine göre değil büyüklük sırasına göre değerlendirerek populasyondaki çeşitliliğin sürdürülmesini sağlamaktadır.

Sıralı Seçim yönteminde N bireyden oluşan populasyon içerisinde bireyler uygunluk değerlerine göre büyükten küçüğe sıralanır. *f(i), i.* bireyin uygunluk değeri ve *n(i), f(i)*'ler büyükten küçüğe sıralandıktan sonra *i*.bireyin sıra sayısını göstermek üzere *i*.bireyin sonraki nesil yer alma olasılığı,

$$p(i) = \frac{2 \times n(i)}{N(N+1)}$$

ile belirlenir(Haupt 1998).

2.6.4. Elitist seçim yöntemi

Elitist seçim yöntemi başlı başına bir seçim yöntemi olmayıp daha önce belirtilen yöntemlerle beraber uygulanabilen bir yöntemdir. Bu yöntemde populasyon içerisindeki en iyi uygunluk değerine sahip birey korunarak doğrudan bir sonraki nesle aktarılır. Böylece seçim yöntemlerinin en iyi bireyi sonraki nesil için seçmeme riski ortadan kaldırılır. De Jong(1975) tarafından önerilen elitist seçim yöntemi her bir nesildeki en iyi çözümün korunarak, algoritmanın çözümden uzak noktalardaki aramalarına rağmen iyi çözümlerin saklanmasını sağlar.

2.7. Çaprazlama

Çaprazlama operatörü uygulama açısından oldukça basit bir yöntemdir. Rasgele sayı üretme, dizilerin kopyalanması ve kısmen değiştirilmesinden oluşur. Çaprazlama işlemi basit olmasının yanında rasgeleliği ve bilgi çeşitliliğini sağlaması açısından genetik algoritmaların en önemli operatörüdür(Goldberg 1989).

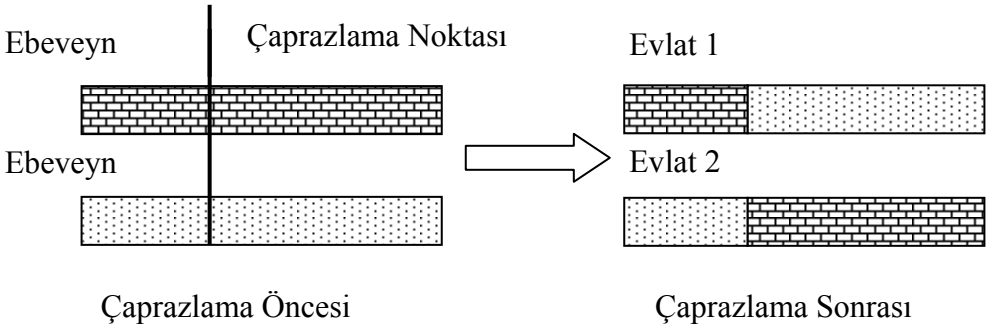
Çaprazlama operatörü farklı çözümler arasında bilgi değişimini sağlayarak arama uzayının benzer, ancak araştırılmamış bölgelerine ulaşmayı sağlayan bir arama operatörüdür(Booker 1987).

Çaprazlama operatörü populasyondaki her bireye uygulanmak zorunda değildir. Çalışmanın başında çaprazlama olasılığı (p_c), $0 < p_c < 1$ olacak şekilde tanımlanır. N bireyden oluşan bir populasyonda çaprazlanacak toplam kromozom sayısı ($N \times p_c$) tanedir. Bir kromozomun çaprazlanması için [0,1] aralığından rasgele bir sayı (k) üretilir, $k < p_c$ ise ilgili kromozom çaprazlama işlemine dahil olur.

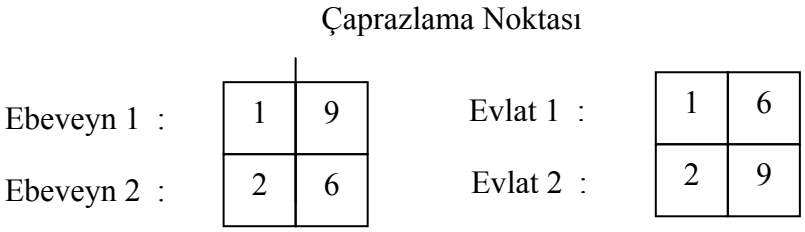
2.7.1. Tek nokta çaprazlama

Belirlenen seçim yöntemine göre sonraki nesli oluşturmak için aday kromozomlardan rasgele iki tanesi seçilir. Seçilen bireyler Ebeveynler olarak adlandırılırken bu ebeveynlerin çaprazlanması sonucu oluşacak yeni bireyler evlatlar olarak adlandırılır. Çaprazlama sonucunda oluşturulan evlatlar ebeveynlerin karakterlerini taşırlar.

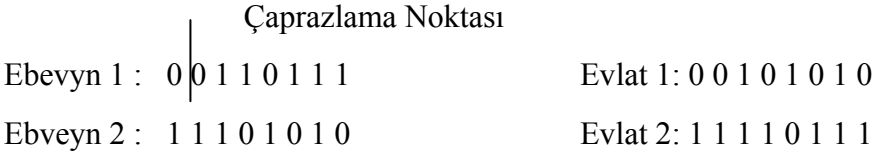
Seçilen ebeveynlerin kromozom uzunlukları L olmak üzere, [1 , L-1] aralığından rasgele bir tamsayı (k) seçilerek çaprazlama noktası oluşturulur. Seçilen iki kromozomdan oluşturulacak iki yeni evladın [1, k] arasındaki parçaları aynen kalırken, [k+1, L] arasındaki parçaları Şekil 2.2., Şekil 2.3 ve Şekil 2.4.'de gösterildiği gibi diğer ebeveynle değiştirilir(Koza 1994).



Şekil 2.2. Tek Nokta Çaprazlama



Şekil 2.3. Gerçek Kodlu GA'da Tek Nokta Çaprazlama



Şekil 2.4. İkili Kodlu GA'da Tek Nokta Çaprazlama

2.7.2. İki nokta çaprazlama

İki nokta çaprazlama operatörü tek nokta çaprazlama operatörünün gelişmiş biçimidir. Çaprazlama işlemi için seçilen ebeveynler için [1,*L*-1] aralığından rasgele iki nokta üretilir. Eşleşen ebeveynlerin rasgele üretilen iki nokta arasında kalan kısımları karşılıklı yer değiştirerek yeni evlatlar oluşur(Şekil 2.5.).

Çaprazlama Noktaları

Ebevyn 1 : 0 0 | 1 1 0 1 1 1 Evlat 1: 0 0 1 0 0 1 1 1

Ebveyn 2 : 1 1 | 1 0 1 0 1 0 Evlat 2: 1 1 1 1 1 0 1 0

Şekil 2.5. İki Nokta Çaprazlama

2.7.3. Düzgün çaprazlama

Düzgün çaprazlama ilk olarak Syswerda(1989) tarafından önerilmiştir. Bu operatörün uygulanabilmesi için öncelikle çaprazlama maskı adı verilen yeni bir kromozom oluşturulur. İkili kodlama ile oluşturulan kromozom üzerindeki 1 ile kodlanan değerler ilgili değerin birinci bireyden, 0 ile kodlanan değerler ilgili değerin ikinci bireyden kopyalanacağını ifade eder. Kopyalama maskındaki bit değerlerine göre ebeveynlerden ilgili kromozomlar kopyalanarak yeni evlat oluşturulur(Şekil 2.6.).

Çaprazlama Maskı 1 : 1 1 0 1 0 0 1 0 1 1

Ebeveyn 1 : 0 1 1 0 0 1 1 1 0 1

Ebeveyn 2 : 1 0 1 1 1 0 0 0 1 0

Evlat : 0 1 1 0 1 0 1 0 0 1

Şekil 2.6. Düzgün Çaprazlama

Syswerda(1989) yaptığı çalışmada 6 farklı problem için çaprazlama operatörlerini karşılaştırmış ve düzgün çaprazlamanın diğer yöntemlerden daha etkin olduğunu göstermiştir.

Spears ve De Jong(1991) çalışmalarında iki nokta çaprazlama operatörünün tek nokta çaprazlama operatöründen daha etkin olduğu sonucuna varmışlardır.

Yapılan çalışmalara göre düzgün çaprazlama küçük populasyonlar için, iki nokta çaprazlama ise büyük populasyonlar için daha iyi sonuçlar vermiştir. Büyük populasyonlar için çeşitlilik sağlanmakta ve küçük populasyonlara göre çeşitlilik ihtiyacı daha az olduğundan iki nokta çaprazlama daha iyi sonuçlar vermektedir(Srinivas ve Patnaik 1994).

2.8. Mutasyon

Mutasyon operatörü GA'nın çalışmasında önemli bir rol oynar. Kromozomların küçük bir yüzdesini değiştirerek arama uzayının farklı bölgelerindeki çözümlerin de araştırılmasını sağlayarak en iyi çözümün bulunmasına yardımcı olur(Goldberg,1989).

Mutasyona uğrayacak kromozomun belirlenmesi için çalışmanın başında belirlenen mutasyon olasılığı (p_m) $0 < p_m < 1$ kullanılır. Her kromozom için [0,1] aralığından rasgele bir (k) sayısı üretilir $k < p_m$ ise kromozom mutasyona uğratılır. Uygulamalarda mutasyon olasılığı genellikle 0.01 alınmaktadır.

Gerçek kodlu GA'nın mutasyon işlemi normal dağılımdan üretilen bir sayı dizisinin mutasyona uğrayacak genlere eklenmesi ya da çıkartılmasıyla gerçekleştirilir(Şekil 2.7.)(Haupt1998).

Ebeveyn 1 :	17	9	Evlat 1 :	17.1	9.7
Ebeveyn 2 :	13	6	Evlat 2 :	13.6	6.4

Şekil 2.7. Gerçek Kodlu GA'da mutasyon

İkili kodlamanın kullanıldığı GA'da mutasyon işlemi 0 olan bitlerin 1, 1 olan bitlerin de 0 olarak değiştirilmesiyle gerçekleştirilir(Şekil 2.8).

Ebevyn 1 : 0 0 1 1 0 1 1 1	Evlat 1: 1 1 0 0 1 0 0 0
Ebveyn 2 : 1 1 1 0 1 0 1 0	Evlat 2: 0 0 0 1 0 1 0 1

Şekil 2.8. İkili Kodlu GA'da mutasyon

2.9. Schema teorisi ve seçim baskısı

GA'ların teorik altyapısı Holland(1975) tarafından ikili kodlamanın kullanılması ve benzerlik profili (Schema) kavramlarına dayanır. Benzerlik profili ikili kodlama kullanılarak oluşturulan kromozomların daha genel olarak sınıflandırılması esasına dayanır. Benzerlik profili oluşturulurken ikili kodlamada kullanılan {0,1} elamanlarına {*} sembolü eklenir. * sembolü, bulunduğu konumdaki değerin {0} veya {1} olmasının önemsiz olduğunu gösterir.

(0 * 0 1 0 1 1 0) ile belirtilen benzerlik profili {(0 1 0 1 0 1 1 0) , (0 0 0 1 0 1 1 0)}

kromozomlarını belirtir. Benzer mantıkla (* * 0 0 0 1 1 0) benzerlik profili ise aşağıdaki 4 kromozomu belirtir.

(0 0 0 0 0 1 1 0)

(0 1 0 0 0 1 1 0)

(1 0 0 0 0 1 1 0)

(1 1 0 0 0 1 1 0)

"*" sembolü ile farklı sayılardaki kromozomlar gösterilebileceği gibi * sembolü kullanılmadan gösterilebilecek kromozom sayısı bir tanedir. (1 1 1 0 0 1 1 0) benzerlik profili yalnızca bir kromozomu belirtir. Benzerlik profilindeki * sembolünün sayısı t olmak üzere bir benzerlik profili en fazla 2^t tane kromozom belirtir. Bununla beraber k uzunluğundaki her bir kromozom 2^k tane benzerlik profili tarafından belirtilebilir.

Bir benzerlik profilinin üç önemli özelliği derecesi, tanım uzunluğu ve uygunluk değeridir.

S, bir benzerlik profilini göstermek üzere o(S), benzerlik profilinin derecesini ifade eder ve kromozom üzerinde * sembolü ile gösterilmeyen bit sayısıdır.

S benzerlik profilinin tanım uzunluğu $\delta(S)$ ile gösterilir ve * sembolü ile gösterilmeyen son ve ilk pozisyonlar arasındaki farkın değeridir.

S benzerlik profilinin t anındaki uygunluk değeri $f^t(s)$ ile gösterilir ve benzerlik profili tarafından belirtilen kromozomların uygunluk değerlerinin ortalamasıdır. Uygunluk değeri populasyon uygunluk değeri ortalamasının üstünde olan benzerlik profillerine ortalama üstü benzerlik profilleri denir.

Benzerlik profili teoremine göre kısa tanım uzunluklu, düşük dereceli ve ortalama üstü benzerlik profilleri sonraki jenerasyonlarda üstel artan sayıda kromozom belirtirler. Bu teoreme göre GA'lar tarafından yapılan aramalar tamamen rasgele bir arama olmayıp, daha iyi çözümlerin sonraki jenerasyonda üstel artan sayıda evlat verecek yapıda bir arama yöntemi olduğunu gösterir(Goldberg 1989).

3.ÇOK DEĞİŞKENLİ LİNEER OLMAYAN MODELLER

Birçok uygulamada elde edilen deneysel veya gözlemsel veriler arasındaki matematiksel ilişki incelenmektedir. Bu tip araştırmalarda gözlenen değişkenlerden biri bağımlı(Y), diğeri(leri) bağımsız değişken(ler) (X) olarak adlandırılmaktadır. Bağımlı değişken ile bağımsız değişken(ler) arasındaki ilişkinin lineer olmaması lineer olmayan modeli oluşturmaktadır. Bu bölümde kullanılacak bazı ifadelerin tanımı aşağıda verilmiştir.

Tanım 3.1 : $f(\underline{x},\underline{\theta})$, $\underline{\theta}=(\theta_1,\theta_2,\ldots,\theta_p)'$ $(\underline{\theta}\in\Theta\subset\Re^p)$ parametre vektörünün bileşenlerine göre türevlenebilir fonksiyon olmak üzere, $\dfrac{\partial f(\underline{x},\underline{\theta})}{\partial\theta_i}$ ifadesi en az bir θ_i'ye $(i=1,2,\ldots,p)$ bağlı ise $f(\underline{x},\underline{\theta})$'ya $\underline{\theta}$ parametre vektörüne göre lineer olmayan fonksiyon denir(Bates ve Watts 1988).

Tanım 3.2 : $\underline{X}$ elemanları rasgele değişkenlerden oluşan bir rasgele vektör olmak üzere, $\underline{X}$ rasgele vektörünün beklenen değeri $E(\underline{X})$ ile kovaryansı $Cov(\underline{X})=E(\underline{X}-E(\underline{X}))(\underline{X}-E(\underline{X}))'$ ile ifade edilir(Johnson ve Wichern 1992).

Tanım 3.3 : $A_{a\times b}$ ve $B_{c\times d}$ boyutlu matrisler olmak üzere A ve B matrislerinin kroneker çarpımları $A\otimes B$ şeklinde gösterilir. $A=\left[\left(a_{ij}\right)\right]_{a\times b}$ olmak üzere A ile B matrislerinin kroneker çarpımları,

$$A\otimes B=\begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1b}B \\ a_{21}B & a_{22}B & \cdots & a_{2b}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{a1}B & a_{a2}B & \cdots & a_{ab}B \end{bmatrix}_{(a\times c)\times(b\times d)}, \quad \left(A\otimes B\right)'=A'\otimes B'$$

dır. A ve B matrislerinin tersi mevcut olmak üzere $\left(\left(A\otimes B\right)^{-1}=A^{-1}\otimes B^{-1}\right)$ özelliği mevcuttur(Gallant 1987).

Tanım 3.4 : $\underline{x}'=\begin{bmatrix}x_1 & x_2 & \cdots & x_k\end{bmatrix}$ ve $A_{k\times k}=\left[\left(a_{ij}\right)\right]$ simetrik matris ($a_{ij}=a_{ji}$ $i,j=1,2,\ldots,k$) olmak üzere her $\underline{x}\neq\underline{0}$ için $\underline{x}'A\underline{x}>0$ oluyorsa A matrisine pozitif tanımlı matris denir(Johnson ve Wichern 1992).

Tanım 3.5 : $A_{k\times k} = \left[\left(a_{ij}\right)\right]$ pozitif tanımlı matris olmak üzere, $A = P'P$ eşitliğini sağlayan,

$$P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1k} \\ 0 & p_{22} & \cdots & p_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & p_{kk} \end{bmatrix}_{k\times k}, \; p_{ii} > 0 \quad i = 1,2,\cdots,k$$

şeklinde bir P üst üçgen matrisi vardır(Johnson ve Wichern 1992).

Tanım 3.6 : $\underline{A} = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix}_{1\times n}$, $(1\times n)$ boyutlu satır vektörü olmak üzere $\underline{A}$ 'nın normu ya da uzunluğu $\left\|\underline{A}\right\|$ ile gösterilir ve $\left\|\underline{A}\right\| = \sqrt{\left(a_1{}^2 + a_2{}^2 + \cdots + a_n{}^2\right)}$ şeklinde hesaplanır(Seber ve Wild 1989).

3.1. Tek Değişkenli Lineer Olmayan Modelin Tanıtımı ve Bazı Gösterimler

Y bağımlı değişken, $X_1, X_2, \ldots, X_k$ açıklayıcı(bağımsız) değişkenler ve $\underline{X} = (X_1, X_2, \ldots, X_k)'$ açıklayıcı değişkenlerin vektörü olmak üzere,

$$Y_t, \underline{X}_t = (X_{1t}, X_{2t}, \ldots, X_{kt})', \quad t=1,2,...,n$$

gözlemleri için $\underline{X_t} = \underline{x_t}$ $(\underline{x}_t \in \mathcal{X} \subset \Re^k)$ olarak gözlendiğinde,

$$Y_t = f(\underline{x}_t, \underline{\theta}) + e_t, \quad t = 1,2,...,n \tag{3.1}$$

biçiminde yazılabilsin. f fonksiyonu bilinmeyen $\underline{\theta} = (\theta_1, \theta_2, ..., \theta_p)'$ $(\underline{\theta} \in \Theta \subset \Re^p)$ parametre vektörünün bileşenlerine göre lineer olmayan bir ifade olduğunda bu modele lineer olmayan model denir(Bates veWatts 1988).

e_t gözlenemeyen veya deneysel hata terimi olmak üzere e_t 'lerin bağımsız, aynı dağılıma sahip, beklenen değerleri sıfır ve varyanslarının eşit olduğu varsayılır. Açıklayıcı değişkenlerin vektörü ile ilgili $\underline{x}_1, \underline{x}_2, \ldots \underline{x}_n$ gözlemleri rasgele bir vektörün gözlem değerleri gibi yorumlanmamaktadır. Örneğin bunlar önceden belirlenen deney noktaları olabilir. Parametre kümesi Θ $(\Theta \subset \Re^p)$ olmak üzere

bilinmeyen gerçek parametre değeri $\underline{\theta}^0$ ($\underline{\theta}^0 \in \Theta$) ile gösterilecektir. Buna göre model esasında,

$$Y_t = f(\underline{x}_t, \underline{\theta}^0) + e_t,\ t=1,2,...,n$$

biçiminde olmak üzere, istatistik teorisi açısından matematiksel işlemleri yürütmek için $\underline{\theta}^0$ yerine $\underline{\theta}$ yazılmaktadır. Eşitlik (3.1)'e atıf yapıldığında $\underline{\theta}$ yerine $\underline{\theta}^0$'ın gelebileceğini hatırlatalım. Kısaca $f(\underline{x}_t, \underline{\theta})$ 'yı $f_t(\underline{\theta})$ olarak gösterip,

$$\underline{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}, \qquad \underline{f}(\underline{\theta}) = \begin{bmatrix} f(\underline{x}_1, \underline{\theta}) \\ f(\underline{x}_2, \underline{\theta}) \\ \vdots \\ f(\underline{x}_n, \underline{\theta}) \end{bmatrix} = \begin{bmatrix} f_1(\underline{\theta}) \\ f_2(\underline{\theta}) \\ \vdots \\ f_n(\underline{\theta}) \end{bmatrix}, \qquad \underline{e} = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix}, \qquad \underline{\theta} = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_p \end{bmatrix}$$

olmak üzere (3.1) modeli

$$\underline{Y} = \underline{f}(\underline{\theta}) + \underline{e} \tag{3.2}$$

şeklinde yazılabilir. Bu şekilde (3.1) ve (3.2) eşitlikleri ile verilen model tek değişkenli lineer olmayan model olarak isimlendirilir.

3.2. Çok Değişkenli Lineer Olmayan Modelin Tanıtımı ve Bazı Gösterimler

Eşitlik (3.1) ile verilen *M* tane tek değişkenli lineer olmayan model göz önüne alınsın. Bunların her biri,

$$Y_{\alpha t} = f_\alpha\left(\underline{x}_t, \underline{\theta}_\alpha^{\,0}\right) + e_{\alpha t},\ \alpha = 1,2,\ldots,M \quad t = 1,2,\ldots,n \tag{3.3}$$

veya

$$\underline{Y}_\alpha = \underline{f}_\alpha\left(\underline{x}, \underline{\theta}_\alpha^{\,0}\right) + \underline{e}_\alpha,\ \alpha = 1,2,\ldots,M$$

ya da

$$
\begin{aligned}
Y_{1t} &= f_1\left(\underline{x}_t\,,\underline{\theta}_1^{\,0}\right)+e_{1t}\\
Y_{2t} &= f_2\left(\underline{x}_t\,,\underline{\theta}_2^{\,0}\right)+e_{2t}\\
&\vdots\\
Y_{Mt} &= f_M\left(\underline{x}_t\,,\underline{\theta}_M^{\,0}\right)+e_{Mt}\,,\ t=1,2,\ldots,n
\end{aligned}
$$

olarak yazılabilir. M tane denklemi bir arada,

$$
Y_{M\times n}=\begin{bmatrix} Y_{11} & Y_{12} & \cdots & Y_{1n}\\ Y_{21} & Y_{22} & \cdots & Y_{2n}\\ \vdots & \vdots & \ddots & \vdots\\ Y_{m1} & Y_{m2} & \cdots & Y_{mn}\end{bmatrix}=\begin{bmatrix} \underline{Y}'_1\\ \underline{Y}'_2\\ \vdots\\ \underline{Y}'_M\end{bmatrix}
$$

$$
f_{M\times n}\left(\underline{x},\underline{\theta}^0\right)=\begin{bmatrix} f_1\left(\underline{x}_1,\underline{\theta}_1^0\right) & f_1\left(\underline{x}_2,\underline{\theta}_1^0\right) & \cdots & f_1\left(\underline{x}_n,\underline{\theta}_1^0\right)\\ f_2\left(\underline{x}_1,\underline{\theta}_2^0\right) & f_2\left(\underline{x}_2,\underline{\theta}_2^0\right) & \cdots & f_2\left(\underline{x}_n,\underline{\theta}_2^0\right)\\ \vdots & \vdots & \ddots & \vdots\\ f_M\left(\underline{x}_1,\underline{\theta}_M^0\right) & f_M\left(\underline{x}_2,\underline{\theta}_M^0\right) & \cdots & f_M\left(\underline{x}_n,\underline{\theta}_M^0\right)\end{bmatrix}=\begin{bmatrix} \underline{f}_1{}'\left(\underline{x},\underline{\theta}_1^{\,0}\right)\\ \underline{f}_2{}'\left(\underline{x},\underline{\theta}_2^{\,0}\right)\\ \vdots\\ \underline{f}_M{}'\left(\underline{x},\underline{\theta}_M^{\,0}\right)\end{bmatrix}
$$

$$
e_{Mn}=\begin{bmatrix} e_{11} & e_{12} & \cdots & e_{1n}\\ e_{21} & e_{22} & \cdots & e_{2n}\\ \vdots & \vdots & \ddots & \vdots\\ e_{M1} & e_{M2} & \cdots & e_{Mn}\end{bmatrix}=\begin{bmatrix} \underline{e}'_1\\ \underline{e}'_2\\ \vdots\\ \underline{e}'_M\end{bmatrix}\qquad \theta_{\mathrm{M\times p}}=\begin{bmatrix} \theta_{11} & \theta_{12} & \cdots & \theta_{1p_1}\\ \theta_{21} & \theta_{22} & \cdots & \theta_{2p_2}\\ \vdots & \vdots & \ddots & \vdots\\ \theta_{M1} & \theta_{M2} & \cdots & \theta_{Mp_M}\end{bmatrix}=\begin{bmatrix} \underline{\theta}'_1\\ \underline{\theta}'_2\\ \vdots\\ \underline{\theta}'_M\end{bmatrix}
$$

gösterimleri altında $Y_{M\times n}=f_{M\times n}\left(\underline{x}\,,\,\underline{\theta}\right)+e_{M\times n}$ biçiminde göstermek mümkündür. Burada, $Y_{\alpha t}$, $\alpha=1,2,\ldots,M$, $t=1,2,\ldots,n$ tek değişkenli tepki değeri, $\underline{x}_t$, k -boyutlu girdi vektörü, $\underline{\theta}_{\alpha}^{\,0}$, p_α boyutlu bilinmeyen parametre vektörü, $p=\sum_{\alpha=1}^{M}p_\alpha$, $e_{\alpha t}$ deneysel hata, $\underline{\theta}_{\alpha}^{\,0}$ bilinmeyen $\underline{\theta}_{\alpha}$ nın gerçek değerini göstermektedir. Ayrıca $\underline{e}_t$ hata vektörü için,

$$\underline{e}_t = \begin{bmatrix} e_{1t} \\ e_{2t} \\ \vdots \\ e_{Mt} \end{bmatrix}_{\mathrm{M}\times 1} \quad , \; E(\underline{e}_{\mathrm{t}}) = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}_{\mathrm{M}\times 1}$$

ve

$$Cov(\underline{e}_t) = \Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1M} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{M1} & \sigma_{M2} & \cdots & \sigma_{MM} \end{bmatrix}_{\mathrm{M}\times\mathrm{M}} \quad , t = 1,2,\ldots,n$$

olmak üzere $\underline{e}_1, \underline{e}_2, \ldots, \underline{e}_M$ hata vektörlerinin bağımsız, aynı dağılımlı, sıfır ortalamalı ve bilinmeyen $\sum$ varyans-kovaryans matrisine sahip olduğu varsayılmaktadır.

Çok değişkenli lineer olmayan modelde, ya bağımlı değişkenler ile ilgili gözlemlerin aynı birey üzerinde (t. birey) alınmasıyla $t = 1,2,\ldots,n$ için elde edilen,

$$\mathrm{Y}_{\alpha\,\mathrm{t}} \quad , \; \alpha = 1,2,\ldots,M$$

gözlem değerlerinin ilişkili olması sebebiyle ya da bazı denklemler arasında bazı parametrelerin ortak kullanımı sebebiyle M tane denklem ayrı ayrı düşünülemez.

Çok değişkenli lineer olmayan model iki şekilde vektör formunda yazılmaktadır(Gallant 1987).

3.2.1. Denklemlere göre gruplanmış gösterim

Çok değişkenli lineer olmayan model M tane tek değişkenli lineer olmayan model şeklinde göz önüne alınır.

$$\underline{Y}_{\alpha} = \begin{bmatrix} Y_{\alpha 1} \\ Y_{\alpha 2} \\ \vdots \\ Y_{\alpha n} \end{bmatrix}_{n\times 1} \quad , \quad \underline{f}_{\alpha}(\underline{\theta}_{\alpha}) = \begin{bmatrix} f_{\alpha}(\underline{x}_1 \, , \, \underline{\theta}_{\alpha}) \\ f_{\alpha}(\underline{x}_2 \, , \, \underline{\theta}_{\alpha}) \\ \vdots \\ f_{\alpha}(\underline{x}_n \, , \, \underline{\theta}_{\alpha}) \end{bmatrix}_{n\times 1} \quad , \quad \underline{e}_{\alpha} = \begin{bmatrix} e_{\alpha 1} \\ e_{\alpha 2} \\ \vdots \\ e_{\alpha n} \end{bmatrix}_{n\times 1} \quad \alpha = 1,2,\ldots,M$$

ve

$$\underline{Y} = \begin{bmatrix} \underline{Y}_1 \\ \underline{Y}_2 \\ \vdots \\ \underline{Y}_M \end{bmatrix}_{(M\times n)\times 1} \quad , \quad \underline{f}(\underline{\theta}) = \begin{bmatrix} \underline{f}_1(\underline{\theta}_1) \\ \underline{f}_2(\underline{\theta}_2) \\ \vdots \\ \underline{f}_M(\underline{\theta}_M) \end{bmatrix}_{(M\times n)\times 1} \quad , \quad \underline{e} = \begin{bmatrix} \underline{e}_1 \\ \underline{e}_2 \\ \vdots \\ \underline{e}_M \end{bmatrix}_{(M\times n)\times 1}$$

olmak üzere,

$$\underline{Y}_\alpha = \underline{f}_\alpha\left(\underline{\theta}_\alpha^{\,0}\right) + \underline{e}_\alpha \ , \qquad \alpha = 1,2,...,M$$

modelleri vektör gösterimi ile,

$$\underline{Y} = \underline{f}(\underline{\theta}) + \underline{e} \tag{3.4}$$

biçiminde gösterilebilir. Buradaki $\underline{\theta}$ parametre vektörünün bileşenleri $\underline{\theta}_1, \underline{\theta}_2, \ldots, \underline{\theta}_M$ vektörlerindeki bileşenlerdir. Ancak $\underline{\theta}_1, \underline{\theta}_2, \ldots, \underline{\theta}_M$ vektörlerinin bazı bileşenleri ortak (aynı) olabilir. Ortak bileşenler $\underline{\theta}$ vektöründe bir kez yer almaktadır. Bu sebepten $\underline{\theta}$'nın boyutu $\underline{\theta}_1, \underline{\theta}_2, ..., \underline{\theta}_M$ vektörlerinin boyutları toplamından daha küçük olabilir(Genç 1997).

Hatalar bağımsız olduğundan $\alpha, \beta = 1,2,\ldots,M$ için

$$Cov\left(\underline{e}_\alpha, \underline{e}'_\beta\right) = \begin{bmatrix} Cov\left(e_{\alpha 1}, e_{\beta 1}\right) & \cdots & Cov\left(e_{\alpha 1}, e_{\beta n}\right) \\ \vdots & \ddots & \vdots \\ Cov\left(e_{\alpha n}, e_{\beta 1}\right) & \cdots & Cov\left(e_{\alpha n}, e_{\beta n}\right) \end{bmatrix}_{n\times n}$$

ve

$$= \begin{bmatrix} \sigma_{\alpha\beta} & 0 & \ldots & 0 \\ 0 & \sigma_{\alpha\beta} & \ldots & 0 \\ \vdots & & & \\ 0 & 0 & \ldots & \sigma_{\alpha\beta} \end{bmatrix}_{n\times n} = \sigma_{\alpha\beta} I_n$$

dir. Buna göre,

$$Cov(\underline{e}) = \begin{bmatrix} \sigma_{11}I_n & \sigma_{12}I_n & \cdots & \sigma_{1M}I_n \\ \sigma_{21}I_n & \sigma_{22}I_n & \cdots & \sigma_{2M}I_n \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{M1}I_n & \sigma_{M2}I_n & \cdots & \sigma_{MM}I_n \end{bmatrix}_{(M\times n)\times(M\times n)}$$

$$= \Sigma \otimes I$$

dır. $\Sigma^{-1} = P'P$ Cholesky ayrışımı ile yazılırsa,

$$(\Sigma \otimes I)^{-1} = \Sigma^{-1} \otimes I$$

$$= P'P \otimes I$$

$$= (P' \otimes I)(P \otimes I)$$

elde edilir. Eşitlik (3.4)'ün soldan $(P \otimes I)$ ile çarpılmasıyla,

$$(P \otimes I)\underline{Y} = (P \otimes I)\underline{f}(\underline{\theta}) + (P \otimes I)\underline{e}$$

ve

$$\underline{Y}^* = (P \otimes I)\underline{Y} \quad , \quad \underline{f}^*(\underline{\theta}) = (P \otimes I)\underline{f}(\underline{\theta}) \quad , \quad \underline{e}^* = (P \otimes I)\underline{e}$$

dönüşümü yardımıyla,

$$\underline{Y}^* = \underline{f}^*(\underline{\theta}) + \underline{e}^*$$

modeline ulaşılır. $\underline{e}^*$ hatalarının varyans-kovaryans matrisi,

$$\begin{aligned} Cov(\underline{e}^*) &= Cov\left((P \otimes I)\underline{e}, \underline{e}'(P \otimes I)'\right) \\ &= (P \otimes I)Cov(\underline{e})(P \otimes I)' \\ &= (P \otimes I)(\Sigma \otimes I)(P \otimes I)' \\ &= (P\Sigma P') \otimes I \\ &= \left[PP^{-1}(P')^{-1}(P')\right] \otimes I \\ &= I_{M\times M} \otimes I_{n\times n} \\ &= I_{nM\otimes nM} \end{aligned}$$

dır. Ayrıca,

$$S(\underline{\theta},\Sigma)=\left[\underline{Y}^{*}-\underline{f}^{*}(\underline{\theta})\right]'\left[\underline{Y}^{*}-\underline{f}^{*}(\underline{\theta})\right]$$

$$=\left[(P\times I)\underline{Y}-(P\times I)\underline{f}(\underline{\theta})\right]'\left[(P\times I)\underline{Y}-(P\times I)\underline{f}(\underline{\theta})\right]$$

$$=\left[\underline{Y}-\underline{f}(\underline{\theta})\right]'(P\otimes I)'(P\otimes I)\left[\underline{Y}-\underline{f}(\underline{\theta})\right]$$

$$=\left[\underline{Y}-\underline{f}(\underline{\theta})\right]'(\Sigma^{-1}\otimes I)\left[\underline{Y}-\underline{f}(\underline{\theta})\right]$$

dır. Bu durumda hata kareler toplamları,

$$S(\underline{\theta},\Sigma)=\left(\underline{Y}-\underline{f}(\underline{\theta})\right)'(\Sigma\otimes I_n)^{-1}\left(\underline{Y}-\underline{f}(\underline{\theta})\right)$$

şeklinde elde edilir(Genç 1997).

3.2.2. Gözlemlere göre gruplanmış gösterim

Gözlemlere göre yapılan gruplama,

$$\underline{Y}_t=\begin{bmatrix}\underline{Y}_{1t}\\ \underline{Y}_{2t}\\ \vdots\\ \underline{Y}_{Mt}\end{bmatrix}_{M\times 1},\quad \underline{f}(\underline{x}_t,\underline{\theta})=\begin{bmatrix}\underline{f}_1(\underline{x}_t,\underline{\theta}_1)\\ \underline{f}_2(\underline{x}_t,\underline{\theta}_2)\\ \vdots\\ \underline{f}_M(\underline{x}_t,\underline{\theta}_M)\end{bmatrix}_{M\times 1},\quad \underline{e}_t=\begin{bmatrix}\underline{e}_{1t}\\ \underline{e}_{2t}\\ \vdots\\ \underline{e}_{Mt}\end{bmatrix},\ \mathrm{t}=1,2,\ldots,\mathrm{n}$$

olmak üzere

$$\underline{Y}=\begin{bmatrix}\underline{Y}_1\\ \underline{Y}_2\\ \vdots\\ \underline{Y}_M\end{bmatrix},\quad \underline{f}(\underline{\theta})=\begin{bmatrix}\underline{f}(\underline{x}_1,\underline{\theta})\\ \underline{f}(\underline{x}_2,\underline{\theta})\\ \vdots\\ \underline{f}(\underline{x}_n,\underline{\theta})\end{bmatrix},\ \underline{e}=\begin{bmatrix}\underline{e}_1\\ \underline{e}_2\\ \vdots\\ \underline{e}_M\end{bmatrix}$$

gösterimleri altında M tane denklem $\underline{Y}=\underline{f}(\underline{\theta})+\underline{e}$ biçiminde olmak üzere Σ hataların varyans-kovaryans matrisi olup $\Sigma^{-1}=\left[\left(\sigma^{\alpha\beta}\right)\right]$ in elemanları $\sigma^{\alpha\beta}$ $\alpha,\beta=1,2,...,M$ ile gösterilsin. Bu durumda,

$$S(\underline{\theta},\Sigma)=\sum_{t=1}^{n}\left[\underline{Y}_t-\underline{f}(x_t,\underline{\theta})\right]'\Sigma^{-1}\left[\underline{Y}_t-\underline{f}(x_t,\underline{\theta})\right]$$

$$=\sum_{t=1}^{n}\sum_{\alpha=1}^{M}\sum_{\beta=1}^{M}\sigma^{\alpha\beta}\left[Y_{\alpha t}-f_\alpha(x_t,\theta_\alpha)\right]'\left[Y_{\beta t}-f_\beta(x_t,\theta_\beta)\right]$$

$$=\sum_{\alpha=1}^{M}\sum_{\beta=1}^{M}\sigma^{\alpha\beta}\left[\underline{Y}_\alpha-\underline{f}(\underline{\theta}_\alpha)\right]'\left[\underline{Y}_\beta-\underline{f}(\underline{\theta}_\beta)\right]$$

$$=\left[\underline{Y}-\underline{f}(\theta)\right]'\left(\Sigma^{-1}\otimes I\right)\left[\underline{Y}-\underline{f}(\theta)\right]$$

olur. $\Sigma^{-1}=P'P$ olduğu göz önünde tutulursa ve P matrisinin α. satırı $\underline{p'}(\alpha)$ ile gösterilirse,

$$S(\underline{\theta},\Sigma)=\sum_{t=1}^{n}\left[\underline{Y}_t-\underline{f}(x_t,\underline{\theta})\right]'P'P\left[\underline{Y}_t-\underline{f}(x_t,\underline{\theta})\right]$$

$$=\sum_{t=1}^{n}\left[P\underline{Y}_t-P\underline{f}(x_t,\underline{\theta})\right]'\left[P\underline{Y}_t-P\underline{f}(x_t,\underline{\theta})\right]$$

$$=\sum_{t=1}^{n}\sum_{\alpha=1}^{M}\left[\underline{p'}(\alpha)\underline{Y}_t-\underline{p'}(\alpha)\underline{f}(\underline{x}_t,\underline{\theta})\right]^2$$

elde edilir. $\alpha=1,2,...,M$ ve $t=1,2,...,n$ için,

$$\begin{aligned}
s &= M(t-1)+\alpha\\
\underline{Y}_s^* &= \underline{p'}(\alpha)\underline{Y}_t\\
\underline{x}_s^* &= \left(\underline{p'}(\alpha),\underline{x'}_t\right)'\\
\underline{f}_s^*(\theta) &= \underline{f}_s^*\left(\underline{x}_s^*,\underline{\theta}\right)=\underline{p'}(\alpha)\underline{f}(x_t,\underline{\theta})
\end{aligned}$$

dönüşümleri uygulanırsa,

$$S(\underline{\theta},\Sigma)=\sum_{s=1}^{nM}\left[\underline{Y}_s^*-\underline{f}_s^*(\underline{\theta})\right]^2$$

veya

$$S(\underline{\theta}, \Sigma) = \sum_{t=1}^{n} \left[\underline{Y}_t - \underline{f}(x_t, \underline{\theta})\right]' \Sigma^{-1} \left[\underline{Y}_t - \underline{f}(x_t, \underline{\theta})\right]$$

biçiminde hata kareler toplamına ulaşılır(Genç 1997).

3.3. Tek Değişkenli Lineer Olmayan Modelde Parametre Tahmini

Lineer olmayan tek değişkenli modelin bilinmeyen $\underline{\theta}^0$ parametresinin tahmini için en küçük kareler yöntemi dikkate alınacaktır. Eşitlik (3.1) ile verilen,

$$Y_t = f(\underline{x}_t, \underline{\theta}) + e_t, \quad t = 1,2,\ldots,n$$

tek değişkenli model ve bu modelin hataları için,

$$E(e_t) = 0, \qquad t = 1,2,\ldots,n$$

$$Cov(\mathrm{e_i}, e_j) = \begin{cases} 0 & , i \neq j \\ \sigma^2 & , i = j \end{cases} \quad , \quad i, j = 1,2,...,n$$

veya vektör gösterimi ile $E(\underline{e}) = \underline{0}$ ve $Cov(\underline{e}) = \sigma^2 I$ olduğu varsayılacaktır.

f biçimsel olarak bilindiğinden hata kareler toplamı ifadesi,

$$Q(\underline{\theta}) = \sum_{t=1}^{n} \left[Y_t - f(\underline{x}_t, \underline{\theta})\right]^2$$

veya vektörel formda,

$$Q(\underline{\theta}) = \left[\underline{Y} - \underline{f}(\underline{\theta})\right]' \left[\underline{Y} - \underline{f}(\underline{\theta})\right] \tag{3.5}$$

şeklinde yazılabilir. Verilen hata kareler toplamı minimum olacak şekilde $\underline{\theta}$ parametre değerini belirlemek en küçük kareler yöntemine göre en iyi tepki fonksiyonunu bulmak demektir.

. t (t=1,2,...,n) satır ve j (j=1,2,...,p) sütun indisi olmak üzere,

$$F(\underline{\theta}) = \left(\frac{\partial}{\partial \theta_j} f(\underline{x}_t, \underline{\theta}) \right)$$

$$= \begin{bmatrix} \frac{\partial}{\partial \theta_1} f(\underline{x}_1, \underline{\theta}) & \frac{\partial}{\partial \theta_2} f(\underline{x}_1, \underline{\theta}) & \cdots & \frac{\partial}{\partial \theta_p} f(\underline{x}_1, \underline{\theta}) \\ \frac{\partial}{\partial \theta_1} f(\underline{x}_2, \underline{\theta}) & \frac{\partial}{\partial \theta_2} f(\underline{x}_2, \underline{\theta}) & \cdots & \frac{\partial}{\partial \theta_p} f(\underline{x}_2, \underline{\theta}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial}{\partial \theta_1} f(\underline{x}_n, \underline{\theta}) & \frac{\partial}{\partial \theta_1} f(\underline{x}_1, \underline{\theta}) & \cdots & \frac{\partial}{\partial \theta_p} f(\underline{x}_n, \underline{\theta}) \end{bmatrix}$$

matrisi $\underline{f}(\underline{\theta})$ fonksiyonunun türev (jakobiyen) matrisi olmak üzere (3.5) ile verilen hata kareler toplamının bilinmeyen $\underline{\theta}$ parametre vektörüne göre türevi,

$$\begin{aligned} \frac{\partial}{\partial \underline{\theta}'} Q(\underline{\theta}) &= \frac{\partial}{\partial \underline{\theta}'} \left(\left(\underline{Y} - \underline{f}(\underline{\theta}) \right)' \left(\underline{Y} - \underline{f}(\underline{\theta}) \right) \right) \\ &= \left(\underline{Y} - \underline{f}(\underline{\theta}) \right)' \frac{\partial}{\partial \underline{\theta}'} \left(\underline{Y} - \underline{f}(\underline{\theta}) \right) + \left(\underline{Y} - \underline{f}(\underline{\theta}) \right)' \frac{\partial}{\partial \underline{\theta}'} \left(\underline{Y} - \underline{f}(\underline{\theta}) \right) \\ &= 2\left(\underline{Y} - \underline{f}(\underline{\theta}) \right)' \left(-\frac{\partial}{\partial \underline{\theta}'} \underline{f}(\underline{\theta}) \right) \\ &= 2\left(\underline{Y} - \underline{f}(\underline{\theta}) \right)' F(\underline{\theta}) \\ &= -2F'(\underline{\theta}) \left(\underline{Y} - \underline{f}(\underline{\theta}) \right) \end{aligned} \tag{3.6}$$

şeklindedir. $\underline{\theta}$ 'nın en küçük kareler tahmin edicisi, Θ parametre uzayı üzerinde $Q(\underline{\theta})$ ' nın,

$$\min_{\underline{\theta} \in \Theta} Q(\underline{\theta}) = Q(\hat{\underline{\theta}})$$

minimize edilmesiyle elde edilen $\hat{\underline{\theta}}$ değeridir. Yani,

$$\frac{\partial}{\partial \underline{\theta}'} Q(\underline{\theta}) \Big|_{\underline{\theta} = \hat{\theta}} = 0$$

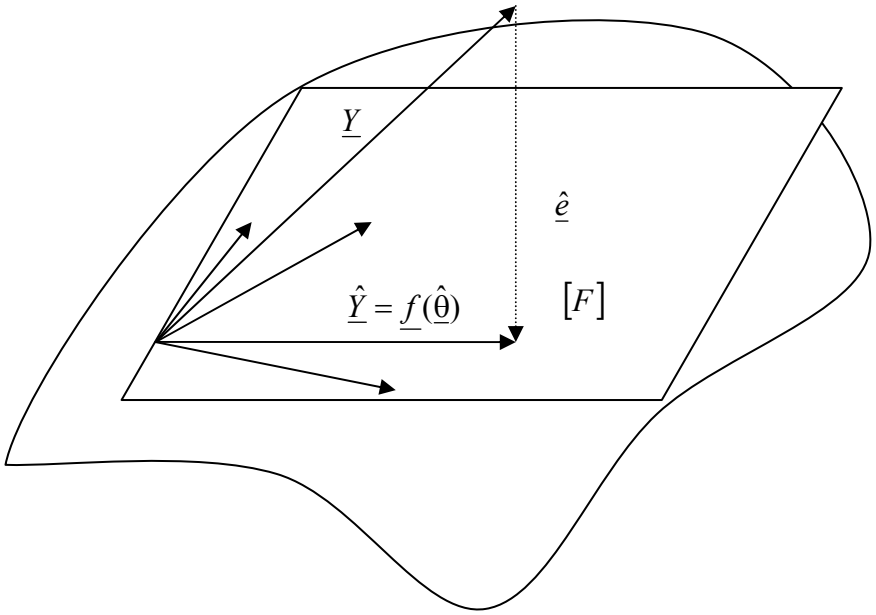
dır. Bu durumda (3.6) ifadesinden

$$F'(\hat{\underline{\theta}}) \left[\underline{Y} - \underline{f}(\hat{\underline{\theta}}) \right] = 0 \tag{3.7}$$

olacağından

$$\underline{\hat{e}} = \underline{Y} - \underline{f}(\underline{\hat{\theta}})$$

artık vektörü, $\underline{\theta} = \underline{\hat{\theta}}$ noktasında $\underline{f}(\underline{\theta})$ 'nın jakobiyen matrisi $F(\underline{\theta})$ 'nın sütun uzayına diktir (Şekil 3.1)(Seber ve Wild 1989).



Şekil 3.1. $\underline{Y}$ vektörünün $F(\underline{\theta})$ sütun uzayına dik izdüşümü

Eşitlik (1.2)'yi minimum yapan değeri araştırmakla ilgili algoritmalar $f(\underline{x}_i, \underline{\theta})$ tepki fonksiyonunun Taylor serisine açılımının yapılması ve gradyentinin hesaplanması civarında toplanmıştır. Nümerik hesaplamalar açısından amaç, işlemlerin yakınsama hızını artırmak ve ardışık işlem sayısını azaltmaktır. Bunun için ardışık işlem adımlarına iyi bir noktadan başlamak gerekir(Gallant 1987).

3.3.1. Başlangıç noktasının seçimi

Parametre tahmini yapmak için bir ardışık işlem adımına başlamadan önce iyi bir başlangıç noktasının saptanması gerekir. Saptanan nokta optimal noktadan çok uzakta ise ardışık işlem sayısı büyüyecektir ve belki de gidilmek istenen noktaya yakınsama gerçekleşmeyecektir.

Başlangıç noktasının seçimi için aşağıdaki yöntemlerden birisi uygulanabilir:

- Önce f fonksiyonunun varsa, lineer kısmı alınıp lineer olmayan kısmında lineerliği bozan parametrelere başlangıç değerleri verilip elde edilen lineer modelde parametreler tahmin edilip bu değerler $\underline{\theta}_0$ başlangıç noktası olarak alınır.

- Gözlem sonucu bilinen Y_t ve $\underline{x}_t$ değerlerini modelde yerine yazarak elde edilen denklem sisteminden parametre sayısı kadar bir alt sistem alınır. Alınan bu denklem sistemi çözülerek $\underline{\theta}_0$ başlangıç noktası bulunur. t = 1,2,...,n için n tane Y_t , $\underline{x}_t$ gözlemi p gruba ayrılır, her bir grup için bulunan Y 'lerin $\overline{Y}_j$ $(j = 1,2,...,\mathrm{p})$ ortalaması ile f değerlerinin $\bar{f}_j(\underline{\theta})$ (j = 1,2,...,p) ortalaması yardımıyla, $\overline{Y}_j = \bar{f}_j(\underline{\theta})$, $j = 1,2,...,\mathrm{p}$ denklem sistemi oluşturulur. Denklem sisteminden bulunan $\underline{\theta}$ başlangıç değeri olarak alınır.

- Konuya hakim olan uygulayıcının önerdiği bir nokta $\underline{\theta}_0$ başlangıç noktası olarak kabul edilir.

Bir tahmin yöntemi yardımıyla lineer olmayan bir model için parametre tahmini yapılmak istendiğinde ardışık işlem adımına $\underline{\theta}_0$ başlangıç noktası ile başlanır. Yakınsama sağlandığında ardışık işlem adımlarının durdurulması gerekir.

3.3.2. Durdurma kuralı

$\varepsilon > 0$ ve $\tau > 0$ sabitleri için,

$$\left\| \underline{\theta}_i - \underline{\theta}_{i-1} \right\| < \varepsilon\left(\left\| \underline{\theta}_{i-1} \right\| + \tau \right)$$

ve

$$\left| Q\left(\underline{\theta}_i\right) - Q\left(\underline{\theta}_{i-1}\right) \right| < \varepsilon\left(Q\left(\underline{\theta}_{i-1}\right) + \tau \right)$$

eşitsizlikleri sağlandığında $\underline{\theta}_0 \to \underline{\theta}_1 \to \underline{\theta}_2 \to \ldots \to \underline{\theta}_i \to \ldots$ ardışık işlemi durdurulur. Uygulamalarda genellikle $\tau = 10^{-3}$ ve $\varepsilon = 10^{-5}$ alınır(Hartley ve Booker 1963).

3.3.3. Gauss-newton yöntemi

Eşitlik (3.1) de verilen lineer olmayan tek değişkenli modelde açıklayıcı değişkenlerin gözlenen değerlerinde $\underline{\theta}$ 'nın bir fonksiyonu olan tepki fonksiyonunun $\underline{\theta}^0$ başlangıç noktası civarında Taylor serisine açılımı

$$f\left(\cdot, \underline{\theta}\right) = f\left(\cdot, \underline{\theta}^0\right) + \frac{\partial f\left(\cdot, \underline{\theta}^*\right)}{\partial \underline{\theta}'}\left(\underline{\theta} - \underline{\theta}^0\right)$$

olsun. Burada $\underline{\theta}^* = \underline{\theta} + \eta\left(\underline{\theta}^0 - \underline{\theta}\right)$, $0 \le \eta \le 1$ dir. $\underline{\theta}^* = \underline{\theta}^0$ alınırsa

$$f\left(\cdot, \underline{\theta}\right) \cong f\left(\cdot, \underline{\theta}^0\right) + \left.\frac{\partial f\left(\cdot, \underline{\theta}\right)}{\partial \underline{\theta}}\right|_{\underline{\theta}=\underline{\theta}^0}\left(\underline{\theta} - \underline{\theta}^0\right)$$

veya vektör gösterimi ile,

$$\underline{f}(\underline{\theta}) \cong \underline{f}(\underline{\theta}^0) + \left. F(\underline{\theta}) \right|_{\underline{\theta}=\underline{\theta}^0} (\underline{\theta} - \underline{\theta}^0) \qquad (3.8)$$

şeklindedir. Hataların kareleri toplamının minimize edilmesinde kullanılan (3.7) eşitliğinde $\underline{\theta} = \underline{\theta}^0$ için f yerine (3.8)'nin göz önüne alınmasıyla,

$$-2F'(\underline{\theta}^0)\left[\underline{Y}-\underline{f}(\underline{\theta}^0)-F(\underline{\theta}^0)(\underline{\theta}-\underline{\theta}^0)\right]=0$$

olur. O halde

$$-2F'(\underline{\theta}^0)\left[\underline{Y}-\underline{f}(\underline{\theta}^0)\right]+2F'(\underline{\theta}^0)F(\underline{\theta}^0)(\underline{\theta}-\underline{\theta}^0)=0$$

olup

$$F'(\underline{\theta}^0)F(\underline{\theta}^0)(\underline{\theta}-\underline{\theta}^0)=F'(\underline{\theta}^0)\left[\underline{Y}-\underline{f}(\underline{\theta}^0)\right] \tag{3.9}$$

elde edilir. Elde edilen (3.9) eşitliğine normal denklemler denir. Normal denklemlerden $\left[F'(\underline{\theta}^0)F(\underline{\theta}^0)\right]$ matrisinin tersinin varlığı koşuluyla,

$$\underline{\theta}-\underline{\theta}^0=\left[F'(\underline{\theta}^0)F(\underline{\theta}^0)\right]^{-1}F'(\underline{\theta}^0)\left[\underline{Y}-\underline{f}(\underline{\theta}^0)\right]$$

olmak üzere,

$$\underline{d}_0=\left[F'(\underline{\theta}^0)F(\underline{\theta}^0)\right]^{-1}F'(\underline{\theta}^0)\left[\underline{Y}-\underline{f}(\underline{\theta}^0)\right]$$

gösterimi ile

$$\underline{\theta}=\underline{\theta}^0+\underline{d}_0$$

yazılır. $\underline{\theta}^0$ bilinmediği için bu şekilde elde edilen $\underline{\theta}$ 'nın bir anlamı yoktur. Ancak, $\underline{\theta}^0$ yerine bir $\underline{\theta}_0$ başlangıç değerinden başlayıp $\underline{d}_0$ vektörü yönünde $\lambda_0\underline{d}_0$ $(0\le\lambda_0\le 1)$ adımına dayalı $\underline{\theta}_1=\underline{\theta}_0+\lambda_0\underline{d}_0$ değeri ve benzer düşünce ile,

$$\underline{\theta}_i=\underline{\theta}_{i-1}+\lambda_{i-1}\underline{d}_{i-1}\quad,\quad i=1,2,\ldots$$

alınarak $\underline{\theta}_0\to\underline{\theta}_1\to\underline{\theta}_2\to\ldots\to\underline{\theta}_i\to\ldots$ değerleri hesaplanabilir. Her adımdaki λ_i $(0\le\lambda_i\le 1\ ,\ i=0,1,2,\ldots)$ değerleri $Q(\underline{\theta}_i)\le Q(\underline{\theta}_{i-1})$ olacak şekilde seçilir. Bu şekilde oluşturulan ardışık yineleme yöntemi Gauss-Newton algoritması olarak bilinir. Özet olarak bu algoritma bir $\underline{\theta}_0$ başlangıç değerinden başlayıp belirlenen bir durdurma kuralına göre,

$$\underline{\theta}_i=\underline{\theta}_{i-1}+\lambda_{i-1}\underline{d}_{i-1}\quad,\quad i=1,2,\ldots \tag{3.10}$$

değerlerini ardışık olarak hesaplar. λ değerlerine adım uzunluğu, $\underline{d}$ vektörlerine de yön vektörü denir(Genç 1997).

Burada,

$$\underline{d}_i = \left[F'(\underline{\theta}_i)F(\underline{\theta}_i)\right]^{-1} F'(\underline{\theta}_i)\left[\underline{Y} - \underline{f}(\underline{\theta}_i)\right] \qquad i = 1,2,\ldots$$

alınır ve λ_i $(0 \le \lambda_i \le 1\ ,\ i = 0,1,2,\ldots)$ değeri $Q(\underline{\theta}_{i+1}) \le Q(\underline{\theta}_i)$ olacak şekilde ayrıca belirlenen bir kurala göre seçilir(Gallant 1987, Seber ve Wild 1989, Hartley ve Booker 1963, Milliken 1988).

3.3.4. En hızlı iniş yöntemi

$$Q(\underline{\theta}) = \sum_{t=1}^{n} \left[Y_t - f(x_t, \underline{\theta})\right]^2$$

ifadesinin $\underline{\theta}$ üzerinden minimizasyonu için Q fonksiyonunun $\underline{\theta} \in \Re^p$ noktasındaki gradiyenti,

$$\nabla Q(\underline{\theta}) = \begin{bmatrix} \dfrac{\partial Q(\underline{\theta})}{\partial \theta_1} \\ \dfrac{\partial Q(\underline{\theta})}{\partial \theta_2} \\ \vdots \\ \dfrac{\partial Q(\underline{\theta})}{\partial \theta_p} \end{bmatrix}$$

dir. En hızlı iniş yöntemi bir gradiyent yöntemidir. Bu yöntemde (3.10) deki $\underline{d}_{i-1}$ vektörünün yerini gradiyent vektörünün ters yönünde olan $-\nabla Q(\underline{\theta}_{i-1})$ vektörü almaktadır. Buna göre algoritma, $\underline{\theta}_0$ başlangıç değeri olmak üzere,

$$\underline{\theta}_i = \underline{\theta}_{i-1} - \lambda_{i-1} \nabla Q(\underline{\theta}_{i-1}) \qquad , \qquad i = 1,2,\ldots$$

biçimindedir. λ_{i-1} değerleri genel olarak $Q(\underline{\theta}_i) \le Q(\underline{\theta}_{i-1})$ olacak şekilde ayrıca belirlenen kurallara göre seçilebilir olmasına rağmen $Q(\underline{\theta}_{i-1} + \lambda_{i-1}\underline{d}_{i-1})$ 'yi minimum yapan değer λ_{i-1} değeri olarak seçilebilir. İndisleri bir tarafa bırakarak $Q(\underline{\theta} + \lambda \underline{d})$

değerini minimum yapan λ değerini bulmak için $Q(\underline{\theta}+\lambda\underline{d})$ fonksiyonu $\underline{\theta}$ noktası civarında Taylor serisine açılırsa,

$$Q(\underline{\theta}+\lambda\underline{d}) = Q(\underline{\theta}) + \lambda\frac{\partial Q(\underline{\theta})}{\partial\underline{\theta}}\underline{d} + \frac{1}{2}\lambda^2\,\underline{d}'\frac{\partial^2 Q(\underline{\theta})}{\partial\underline{\theta}\partial\underline{\theta}'}\underline{d} + \cdots$$

olur. λ değeri

$$Q(\underline{\theta}+\lambda\underline{d}) \le Q(\underline{\theta})$$

koşulunu sağlayacak biçimde seçilir. Bu durumda,

$$Q(\underline{\theta}+\lambda\underline{d}) - Q(\underline{\theta}) \le 0$$

olup

$$Q(\underline{\theta}+\lambda\underline{d}) - Q(\underline{\theta}) = \lambda\frac{\partial Q(\underline{\theta})}{\partial\underline{\theta}}\underline{d} + \frac{1}{2}\lambda^2\,\underline{d}'\frac{\partial^2 Q(\underline{\theta})}{\partial\underline{\theta}\partial\underline{\theta}'}\underline{d} \le 0$$

dır. λ değerini elde etmek için $Q(\underline{\theta}+\lambda\underline{d}) - Q(\underline{\theta})$ ifadesi λ üzerinden minimize edilirse,

$$\frac{\partial Q(\underline{\theta})}{\partial\underline{\theta}}\underline{d} + \lambda\underline{d}'\frac{\partial^2 Q(\underline{\theta})}{\partial\underline{\theta}\partial\underline{\theta}'}\underline{d} = 0$$

eşitliğinden

$$\lambda^* = -\frac{\dfrac{\partial Q(\underline{\theta})}{\partial\underline{\theta}}\underline{d}}{\underline{d}'\dfrac{\partial^2 Q(\underline{\theta})}{\partial\underline{\theta}\partial\underline{\theta}'}\underline{d}}$$

bulunur(Milliken 1988, Ross 1990).

3.3.5. Marquardt yöntemi

Verilen (1.1) modeli bir $\underline{\theta}_0$ başlangıç değerine bağlı olarak $\underline{\theta} = \underline{\theta}_0 + \underline{d}$ olarak yazılsın. Bu durumda model,

$$Y_t = f\left(\underline{x}_t, \underline{\theta}_0 + \underline{d}\right) + e_t \qquad , \quad t = 1,2,...n$$

olur. $f\left(\underline{x}_t, \underline{\theta}_0 + \underline{d}\right)$ fonksiyonu $\underline{\theta}_0$ noktası civarında birinci dereceden Taylor serisine açılırsa

$$Y_t = f\left(\underline{x}_t, \underline{\theta}_0\right) + \sum_{i=1}^{p}\left(\frac{\partial f\left(\underline{x}_t, \underline{\theta}_0\right)}{\partial \theta_i}\right) d_i + e_t, \quad t = 1,2,...,n$$

ve vektör gösterimi ile,

$$\underline{Y} = \underline{f}(\underline{\theta}_0) + F'(\underline{\theta}_0)\underline{d} + \underline{e}$$

elde edilir.

$$\underline{u}_0 = \underline{Y} - \underline{f}(\underline{\theta}_0)$$

değişken değiştirmesi ile,

$$\underline{u}_0 = F(\underline{\theta}_0)\underline{d} + \underline{e}$$

modeline ulaşılır. Amaç uygun bir $\underline{d}$ yön vektörü belirlemektir. Normal denklemler

$$F'(\underline{\theta}_0)F(\underline{\theta}_0)\underline{d} = F'(\underline{\theta}_0)\underline{u}$$

dır. $C_0 = F'(\underline{\theta}_0)F(\underline{\theta}_0)$ matrisinin köşegen elemanlarının dışındakilerin yerine sıfır ve c_{ii} köşegen elemanlarının yerine $1/\sqrt{c_{ii}}$ yazılmasıyla elde edilen köşegen matris,

$$D_0 = \begin{bmatrix} 1/\sqrt{c_{11}} & 0 & \cdots & 0 \\ 0 & 1/\sqrt{c_{22}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1/\sqrt{c_{pp}} \end{bmatrix}$$

olmak üzere,

$$C_0^* = D_0' C_0 D_0$$

$$\underline{g}_0^* = D_0 F'(\underline{\theta}_0)\underline{u}$$

ve

$$\underline{d}^* = D_0^{-1}\underline{d}_0$$

dönüşümleri altında normal denklemler

$$C_0^* \underline{d}^* = \underline{g}_0^*$$

şeklinde yazılır. Zayıf koşulluluğa bir tedbir olarak C_0 matrisinde köşegen elemanlarına λ $(\lambda > 0)$ gibi küçük bir sayının eklenmesiyle

$$(C_0^* + \lambda I)\underline{d}^* = \underline{g}_0^*$$

yazılır. Buradan

$$\underline{d}_0^* = (C_0^* + \lambda I)^{-1}\underline{g}_0^* \tag{3.11}$$

olmak üzere,

$$\underline{\theta} = \underline{\theta}_0 + \underline{d}_0 \tag{3.12}$$

iyileştirilmiş $\underline{\theta}$ değerine geçilebilir. $\underline{d}_0 = D_0\underline{d}_0^*$ yön vektörü yani λ $(\lambda > 0)$ değeri $Q(\underline{\theta}) < Q(\underline{\theta}_0)$ olacak şekilde seçilmelidir. λ 'nın seçimi ile ilgili bir kural aşağıdaki gibi olabilir(Marquardt 1963, Milliken 1988).

1. Adım. ϑ $(\vartheta > 1)$ sayısı seçilir. (Örneğin, $\vartheta = 10$ alınabilir.)

2. Adım. $\lambda_0 = 10^{-2}$ alınsın.

3. Adım. $a = \lambda_0/\vartheta$, $b = \lambda_0$ ve

$$Q_a = Q(\underline{\theta}_0 + D_0(C_0^* + \lambda_0/\vartheta I)^{-1}\underline{g}_0^*)$$

$$Q_b = Q(\underline{\theta}_0 + D_0(C_0^* + \lambda_0 I)^{-1}\underline{g}_0^*)$$

olmak üzere,

(i) eğer $Q_a < Q(\underline{\theta}_0)$ ise $\lambda = a$ alınsın.

(ii) eğer $Q_a > Q(\underline{\theta}_0)$ ve $Q_b < Q(\underline{\theta}_0)$ ise $\lambda = b$ alınsın.

(iii) eğer $Q_a > Q(\underline{\theta}_0)$ ve $Q_b > Q(\underline{\theta}_0)$ ise $\lambda = \lambda_0 \vartheta^{\omega}$ alınsın. Burada ω, $Q(\underline{\theta} + D_0(C_0^* + \lambda_0 \vartheta^{\omega} I)^{-1} \underline{g}_0^*) < Q(\underline{\theta}_0)$ eşitsizliğini sağlayan en küçük pozitif tamsayıdır.

Özetlersek, $\underline{\theta}_0$ gibi bir başlangıç değerinden başlayıp uygun bir λ seçildikten sonra (3.11) ve (3.12)' den,

$$\underline{\theta}_1 = \underline{\theta}_0 + \underline{d}_0$$

iyileştirilmiş değerine geçilebilir. Adımsal olarak işlemler aşağıda verilmiştir.

1. Adım. $\underline{\theta}_0$ başlangıç değeri seçilir, λ_0 yukarıdaki kurala göre belirlenir ve

$$\underline{\theta}_1 = \underline{\theta}_0 + \underline{d}_0$$

bulunur.

2. Adım. Belli bir durdurma kuralına göre, durdurma kuralı sağlanıyorsa adımlar durdurulur aksi halde bulunan $\underline{\theta}_1$ başlangıç değeri olarak alınır ve birinci adıma dönülür. Bu çalışmada aşağıdaki gibi bir durdurma kuralı kullanılmaktadır.

$$\left|Q(\underline{\theta}_1) - Q(\underline{\theta}_0)\right| < \varepsilon(Q(\underline{\theta}_0) + \tau)$$

ve

$$\left\|\underline{\theta}_1 - \underline{\theta}_0\right\| < \varepsilon(\left\|\underline{\theta}_0\right\| + \tau)$$

koşulları sağlanana kadar işlemler 1. Adım'a dönülerek tekrar edilir. Burada sabitler $(\varepsilon = 10^{-5}\ ,\ \tau = 10^{-3})$ alınır(Genç 1997).

3.4. Çok Değişkenli Lineer Olmayan Modelde Parametre Tahmini

Bu kesimde (3.3) ile verilen, lineer olmayan çok değişkenli modelin bilinmeyen parametreleri olan $\underline{\theta}$ ve Σ nın tahminleri hesaplanmak istensin. $\underline{e}$'nin bağımsız, aynı dağılımlı ve bilinmeyen Σ varyans - kovaryans matrisli olduğu varsayımları ışığı altında göz önüne alınan model ya 1. tip yada 2. tip olarak düşünülsün. Parametre tahmini yapılırken en küçük kareler tahmin yöntemi ele alınmıştır.

3.4.1. En küçük kareler tahmin edicisi

Ele alınan (3.1) ile verilen lineer olmayan çok değişkenli modelin 1. tipte olduğu kabul edilsin. Tek değişkenli lineer olmayan her bir model için Jakobiyen matrisi,

$$F_\alpha(\underline{\theta}_\alpha) = \frac{\partial}{\partial \theta'_\alpha} \underline{f}_\alpha(\underline{\theta}_\alpha) \ , \ \alpha = 1,2,...,M$$

$n \times p_\alpha$ boyutlu matrisi olmak üzere her model için

$$S_\alpha(\underline{\theta}_\alpha) = \left[\underline{Y}_\alpha - \underline{f}_\alpha(\underline{\theta}_\alpha)\right]' \left[\underline{Y}_\alpha - \underline{f}_\alpha(\underline{\theta}_\alpha)\right] \ , \ \alpha = 1,2,...,M \qquad (3.13)$$

karesel formu bilinmeyen $\underline{\theta}_\alpha$ parametresine göre minimize edilerek $\hat{\underline{\theta}}_\alpha^{\#}$ tahmin değeri elde edilir. Her bir model için hata teriminin tahmini,

$$\hat{\underline{e}}_\alpha = \underline{Y}_\alpha - \underline{f}_\alpha\left(\hat{\underline{\theta}}_\alpha^{\#}\right) \ , \ \alpha = 1,2,...,M$$

olup

$$\hat{E} = \left[\hat{\underline{e}}_1 \ \vdots \ \hat{\underline{e}}_2 \ \vdots \ \cdots \ \vdots \ \hat{\underline{e}}_M\right]$$

yazılırsa tüm modellerin hatalarının varyans-kovaryans matrisinin tahmini,

$$\hat{\Sigma} = \frac{1}{n} \hat{E}' \hat{E}$$

olarak hesaplanır(Gallant 1987).

$\hat{\underline{\theta}}_{\alpha}^{\#}$, $\alpha = 1,2,...,M$ ve ardından $\hat{\Sigma}$ 'yı elde etmek için, (3.13)'deki $S_{\alpha}\left(\underline{\theta}_{\alpha}\right)$ karesel formu bölüm 3.3.'de anlatılan yöntemlerden birisi kullanılarak minimize edilir. Başlangıç noktası olarak ${}_{0}\underline{\theta}_{\alpha}$ değeri ve yöntem olarak Gauss-Newton yöntemi seçilsin. ${}_{0}\underline{\theta}_{\alpha}$'dan atılan ilk Gauss - Newton adımı,

$${}_{1}\underline{\theta}_{\alpha} = {}_{0}\underline{\theta}_{\alpha} + {}_{0}\lambda_{\alpha}\left[F'_{\alpha}\left({}_{0}\underline{\theta}_{\alpha}\right)F_{\alpha}\left({}_{0}\underline{\theta}_{\alpha}\right)\right]^{-1}F'_{\alpha}\left({}_{0}\underline{\theta}_{\alpha}\right)\left[\underline{Y}_{\alpha} - \underline{f}_{\alpha}\,{}_{0}\underline{\theta}_{\alpha}\right],\ \alpha = 1,2,...,M$$

dir. Burada ${}_{0}\lambda_{\alpha}$ değeri 0 ile 1 arasında değer alır. Bu işlem belirlenen bir durdurma kuralı sağlanıncaya kadar sürer.

$\hat{\Sigma}$ hesaplandıktan sonra ikinci aşamada Cholesky ayrışımı ile $\hat{\Sigma}^{-1} = \hat{P}'\hat{P}$ olacak şekilde $\hat{P}'$ matrisi bulunur. Böylece

$$\left(\hat{P}\otimes I\right)\underline{Y} = \left(\hat{P}\otimes I\right)\underline{f}(\underline{\theta}) + \left(\hat{P}\otimes I\right)\underline{e}$$

şeklinde çok değişkenli lineer olmayan modelinden tek değişkenli lineer olmayan modeline varılır. $\underline{\theta}^{\,0}$ bilinmeyen parametresinin tahmin edicisi

$$S(\underline{\theta},\Sigma) = \left[\underline{Y}^{*} - \underline{f}^{*}(\underline{\theta})\right]'\left[\underline{Y}^{*} - \underline{f}^{*}(\underline{\theta})\right]$$

$$= \left[\underline{Y} - \underline{f}(\underline{\theta})\right]'\left(\Sigma^{-1}\otimes I\right)\left[\underline{Y} - \underline{f}(\underline{\theta})\right]$$

kareler toplamında Σ yerine önceden hesaplanan $\hat{\Sigma}$ değeri ve $\underline{\theta}_{0} = ({}_{0}\underline{\theta}_{1}, {}_{0}\underline{\theta}_{2}, ..., {}_{0}\underline{\theta}_{M})$ başlangıç değeri yerine $\left(\hat{\underline{\theta}}_{1}^{\#}, \hat{\underline{\theta}}_{2}^{\#}, ..., \hat{\underline{\theta}}_{M}^{\#}\right)$ tahmin değerleri alınıp,

$$\min_{\underline{\theta}\in\Theta} S\left(\underline{\theta},\hat{\Sigma}\right)$$

$\underline{\theta}$'ya göre minimize edilerek $\hat{\underline{\theta}}$ olarak hesaplanır. Örneğin Gauss-Newton yönteminde minimizasyon algoritmasındaki adımlar,

1. Adım. İlk önce her bir model için ${}_0\underline{\theta}_1, {}_0\underline{\theta}_2, ..., {}_0\underline{\theta}_M$ değerleri başlangıç noktası olarak alınıp M tane lineer olmayan modelin $\left(\underline{\hat{\theta}}_1^{\#}, \underline{\hat{\theta}}_2^{\#}, ..., \underline{\hat{\theta}}_M^{\#}\right)$ tahmin değerleri hesaplanır. Bu tahmin değerleri kullanılarak $\hat{\Sigma}$ değeri hesaplanır.

2. Adım. Bu aşamada $\underline{\theta}_0 = \left(\underline{\hat{\theta}}_1^{\#}, \underline{\hat{\theta}}_2^{\#}, ..., \underline{\hat{\theta}}_M^{\#}\right)$ başlangıç değeri olarak alınıp ve $\hat{\Sigma}$ kullanarak

$$\underline{\theta}_1 = \underline{\theta}_0 + \lambda_0 \left[F'(\underline{\theta}_0)(\hat{\Sigma} \otimes I_n)^{-1} F(\underline{\theta}_0) \right]^{-1} F'(\underline{\theta}_0)(\hat{\Sigma} \otimes I_n)^{-1} [\underline{Y} - \underline{f}(\underline{\theta}_0)]$$

olmak üzere

$$S(\underline{\theta}_1, \hat{\Sigma}) < S(\underline{\theta}_0, \hat{\Sigma})$$

olacak biçimde $0 < \lambda_0 < 1$ şeklinde bir λ_0 bulunur.

3. Adım. Bu işlemler belirlenen bir durdurma kuralı sağlanıncaya kadar yürütülür.

Bu yönteme 2 Aşamalı Gauss-Newton yöntemi denir(Genç 1997).

Şimdi ele alınan çok değişkenli lineer olmayan modelin 2. tipte olduğu kabul edilsin, yani çok değişkenli lineer olmayan model,

$$\underline{Y}_t = \underline{f}(\underline{x}_t, \underline{\theta}) + \underline{e}_t \qquad , \qquad t = 1,2,\ldots,n$$

biçiminde ve aynı zamanda $\Sigma^{-1} = P'P$ olduğu da göz önüne alınırsa ve P matrisinin α . satırı $\underline{p}'(\alpha)$ olmak üzere,

$$S(\underline{\theta}, \Sigma) = \sum_{t=1}^{n} \sum_{\alpha=1}^{M} \left[\underline{p}'(\alpha) \underline{Y}_t - \underline{p}'(\alpha) \underline{f}(\underline{x}_t, \underline{\theta}) \right]^2$$

şeklinde hata kareler toplamı yazılabilir. Buradan,

$$s = M(t-1) + \alpha$$

$$\underline{Y}_s^* = \underline{p}'(\alpha) \underline{Y}_t$$

$$\underline{x}_s^* = \left(\underline{p}'(\alpha)\ ,\ \underline{x}'_t\right)'$$

$$\underline{f}_s^*(\underline{\theta}) = \underline{p}'(\alpha)\underline{f}(\underline{x}_t\ ,\theta)\ ,\ \alpha = 1,2,\ldots,M \qquad t = 1,2,...,n$$

dönüşümü yardımıyla çok değişkenli lineer olmayan modelin

$$\underline{Y}_S^* = \underline{f}_S^*\left(\underline{x}_S^*\ ,\underline{\theta}\ \right) + \underline{e}_S^*\ ,\ s = 1,2,\ldots,nM$$

biçimindeki tek değişkenli lineer olmayan modeline ulaşılır. Dolayısıyla hata kareler toplamı

$$S(\underline{\theta}\ ,\Sigma) = \sum_{s=1}^{nM}\left[\underline{Y}_S^* - \underline{f}_S^*(\underline{\theta})\right]^2 \tag{3.14}$$

olarak yazılıp parametre tahmini yaparken $S(\underline{\theta}\ ,\Sigma)$ karesel formu doğrudan $\underline{\theta}$' ya göre minimize edilir. Tahmin yöntemi olarak Gauss-Newton yöntemi kullanıldığında Gauss-Newton yön vektörü,

$$\underline{D}(\underline{\theta}\ ,\Sigma) = \sum_{s=1}^{nM}\left(\frac{\partial}{\partial\,\underline{\theta}}\underline{f}^*\left(\underline{x}_s^*\ ,\underline{\theta}\right)\right)\left(\frac{\partial}{\partial\,\underline{\theta}}\underline{f}^*\left(\underline{x}_s^*\ ,\underline{\theta}\right)'\right)^{-1}$$

$$\times\ \sum_{s=1}^{nM}\left(\frac{\partial}{\partial\,\underline{\theta}}\underline{f}^*\left(\underline{x}_s^*\ ,\ \underline{\theta}\right)\right)\left[\underline{Y}_s^* - \underline{f}^*\left(\underline{x}_s^*\ ,\ \underline{\theta}\right)\right]$$

$$= \left[\sum_{t=1}^{n}\sum_{\alpha=1}^{M}\left(\frac{\partial}{\partial\,\underline{\theta}}\underline{f}'\left(\underline{x}_{\mathrm{t}}\ ,\underline{\theta}\right)\right)\underline{p}(\alpha)\underline{p}'(\alpha)\left(\frac{\partial}{\partial\,\underline{\theta}}\underline{f}(\underline{x}_{\mathrm{t}}\ ,\underline{\theta})\right)\right]^{-1}$$

$$\times\sum_{t=1}^{n}\sum_{\alpha=1}^{M}\left(\frac{\partial}{\partial\,\underline{\theta}}\underline{f}'\left(\underline{x}_{\mathrm{t}}\ ,\underline{\theta}\right)\right)\underline{p}(\alpha)\underline{p}'(\alpha)\left[\underline{Y}_t - \underline{f}(\underline{x}_{\mathrm{t}}\ ,\underline{\theta})\right]$$

$$= \left[\sum_{t=1}^{n}\left(\frac{\partial}{\partial\,\underline{\theta}}\underline{f}'\left(\underline{x}_{\mathrm{t}}\ ,\underline{\theta}\right)\right)\Sigma^{-1}\left(\frac{\partial}{\partial\,\underline{\theta}}\underline{f}'\left(\underline{x}_{\mathrm{t}}\ ,\underline{\theta}\right)\right)\right]^{-1}$$

$$\times\sum_{t=1}^{n}\left(\frac{\partial}{\partial\,\underline{\theta}}\underline{f}'\left(\underline{x}_{\mathrm{t}}\ ,\underline{\theta}\right)\right)\Sigma^{-1}\left[\underline{Y}_t - \underline{f}(\underline{x}_{\mathrm{t}}\ ,\underline{\theta})\right] \tag{3.15}$$

şeklinde olur.

Burada da ilk önce Σ nın bir tahmini elde edilir. Bu tahmin 1. Gösterimden elde edilen $\hat{\Sigma}$ olabilir. Aynı $\hat{\Sigma}$, (3.14) de Σ yerine I alınıp, örneğin (3.15) kullanılarak Gauss-Newton yöntemi ile hesaplanabilir. $\hat{\Sigma}$ nın elde edilişi sırasındaki $\hat{\underline{\theta}}^{\#}$ değeri başlangıç değer olarak alınıp belli bir yöntem ile

$$\min_{\underline{\theta} \in \Theta} S\left(\underline{\theta}, \hat{\Sigma}\right)$$

optimizasyon problemi çözülür.

Durdurma kuralı verilen $\varepsilon > 0$ ve $\tau > 0$ için

$$\frac{\left\|\underline{\theta}_{i+1} - \underline{\theta}_i\right\|}{\left\|\underline{\theta}_i\right\| + \tau} < \varepsilon$$

ve

$$\frac{\left|S\left(\underline{\theta}_i, \hat{\Sigma}\right) - S\left(\underline{\theta}_{i+1}, \hat{\Sigma}\right)\right|}{\left|S\left(\underline{\theta}_i, \hat{\Sigma}\right)\right| + \tau} < \varepsilon$$

koşullarının sağlanmasıdır($\varepsilon = 10^{-5}$ ve $\tau = 10^{-3}$)(Genç 1997).

4. ÇOK DEĞİŞKENLİ LİNEER OLMAYAN MODELDE GENETİK ALGORİTMA İLE PARAMETRE TAHMİNİ

Çok değişkenli lineer olmayan modelin en küçük kareler tahmin edicisi olarak Gauss Newton yöntemi bölüm 3'te verilmişti. Bu bölümde Çok değişkenli lineer olmayan modelin parametre tahmini genetik algoritma kullanılarak iki aşamada gerçekleştirilecektir. İlk aşamada tek değişkenli lineer olmayan modelin parametrelerinin tahmin edilmesi amacıyla lineer olmayan genetik algoritma NLGA1 tanıtılacaktır. İkinci aşamada ise çok değişkenli lineer olmayan modelin parametrelerinin tahmin edilmesi amacıyla çok değişkenli lineer olmayan genetik algoritma NLGA2 tanıtılacaktır.

4.1. Tek Değişkenli Lineer Olmayan Genetik Algoritma (NLGA1)

Bölüm 3'te eşitlik (3.1) ile verilen tek değişkenli lineer olmayan model,

$$Y_t = f(\underline{x}_t, \underline{\theta}) + e_t, \quad t = 1,2,...,n$$

ve modelin bilinmeyen parametreleri $\underline{\theta} = (\theta_1, \theta_2, ..., \theta_p)'$ olmak üzere $\underline{\theta}$ parametre vektörünün tahmin edilmesi için,

$$\min Q(\underline{\theta}) = \sum_{t=1}^{n} \left[Y_t - f(\underline{x}_t, \underline{\theta}) \right]^2 \tag{4.1}$$

hata kareler toplamı ifadesini minimum yapan $\underline{\theta}$ değeri araştırılacaktır. Bu amaçla oluşturulan genetik algoritma tek değişkenli lineer olmayan genetik algoritma NLGA1 olarak adlandırılmıştır. NLGA1'e ilişkin seçilen parametreler başlangıç değerleri ve kullanılan operatör bilgileri aşağıda verildikten sonra algoritmanın işleyiş adımları örnek üzerinde gösterilmiştir.

NLGA1'de kullanılan parametreler,

N : Gözlem sayısı

Y_i : Bağımlı değişken gözlem değerleri $(i = 1,2,\ldots,N)$

X_i : Bağımsız değişken gözlem değerleri $(i = 1,2,\ldots,N)$

p : Modeldeki toplam parametre sayısı

$f(\underline{x}_t, \underline{\theta})$: Tepki fonksiyonu

θ_i : Bilinmeyen *i*.parametre

p_c : Çaprazlama olasılığı

p_m : Mutasyon olasılığı

alt_i : *i*. parametre alt sınır değeri

$üst_i$: *i*. parametre üst sınır değeri

N_{pop} : Her bir nesildeki populasyon büyüklüğü

Max_{nesil} : Maksimum nesil büyüklüğü

hkt_i^t : *t*. nesildeki i.bireyin uygunluk değeri $(t = 1,2,...,Max_{nesil})$

f_i^t : *t*. nesildeki *i*. bireyin tepki fonksiyonu değeri $(t = 1,2,...,Max_{nesil})$

olmak üzere gerçek kodlama kullanılan NLGA1'de seçim yöntemi olarak rulet çemberi yöntemi, mutasyon operatörü için çok nokta mutasyon operatörü kullanılmıştır. NLGA1'de elitist seçim stratejisi kullanılarak optimum çözüme daha hızlı yakınsama amaçlanmıştır. Örnek1 üzerinde NLGA1'in uygulaması yapılarak algoritmanın işleyişi gösterilmiştir.

Örnek1 : $Y_1 = \theta_1 x_t + \theta_2 e^{\theta_3 x_t} + e_{1t}$, 3 parametreli modelin parametre değerleri $\theta_1 = 5$, $\theta_2 = 3$, $\theta_3 = 1$ ve e_{1t} hata değerleri normal dağılımdan üretilerek elde edilen X ve Y değerleri çizelge (4.1.) de verilmiştir(Genç 1997).

Çizelge 4.1. Örnek1 için üretilen değerler

X	-3	-2.5	-2	-1.5	-1	-0.5	0	0.5	1.5	2
Y_1	-16.29	-14.71	-10.81	-7.06	-7.03	-0.083	2.11	8.45	20.24	31.69
Y_2	6.94	-1.21	3.52	-3.65	-1.56	0.56	3.96	4.48	13.88	22.38

Örnek1 için kullanılan amaç fonksiyonu (4.1) eşitliği ile verilen hata kareler toplamı olmak üzere NLGA1 parametre değerleri,

N=10 (Gözlem sayısı)

$alt_i = 0$ $üst_i = 10$ i=1,2,3 (Parametre alt ve üst sınır değerleri)

$f(x) = \theta_1 x + \theta_2 e^{\theta_3 x}$, tepki fonksiyonu

N_{pop}=2000, her bir nesildeki populasyon büyüklüğü

$p_c = 0.90$ ve $p_m = 0.10$

şekilde belirlenmiştir. NLGA1'in işleyiş adımları aşağıda verilmiştir.

1. İterasyon indisi t=1 alınır. Her bir parametre için ($alt_i \le \theta_i \le üst_i$ i=1,2,..p) değerleri düzgün dağılımdan rasgele üretilir. Üretilen bu değerler GA'da bir bireyi dolayısıyla bir çözümü ifade eder. *Npop* parametre değeri aynı şekilde rasgele üretilerek başlangıç populasyonu oluşturulur.

2. Başlangıç populasyonundaki bireylerin ($Npop$ adet) $f(x) = \theta_1 x + \theta_2 e^{\theta_3 x}$ tepki fonksiyonundaki değerleri (f_i^t) hesaplanır.

3. Hesaplanan tepki değerleri (4.1) eşitliğinde yerine yazılarak her bir bireyin uygunluk değeri (Hata Kareler Toplamı) hesaplanarak hkt_i^t değerleri oluşturulur.

4. hkt_i^t değerlerine göre bireyler küçükten büyüğe sıralanır. Uygunluk değeri en kötü olan populasyon büyüklüğünün %75'i kadar birey yeniden üretilerek hkt_i^t değerleri hesaplanır ve bireyler küçükten büyüğe sıralanır. hkt_i^t değeri küçük olan bireyler çözüm için daha uygun bireyleri temsil ettikleri için güçlü olan yaşar ilkesine göre bu bireylerin sonraki nesillere aktarılma olasılıklarının yüksek olması gerekir. Populasyondaki bireylerin hkt_i^t değerleri ile ters orantılı olarak sonraki nesle seçilme olasılıkları,

$$p(i) = \frac{1}{hkt_i^t} \frac{1}{\sum_{i=1}^{N_{pop}} \frac{1}{hkt_i^t}} \tag{4.2}$$

ve $\sum_{i=1}^{N_{pop}} p(i) = 1$ olacak şekilde hesaplanır.

5. Adım 4'te hesaplanan *p(i)* değerleri rulet çemberi seçim yönteminde oluşturulan her bir çember diliminin büyüklüğünü temsil eder. Bir sonraki nesle aktarılacak birey sayısı N_{pop} kadardır. [0,1] aralığından N_{pop} adet rasgele sayı üretilir, üretilen sayıların rulet çemberindeki karşılığı olan birey bir sonraki nesle seçilerek N_{pop} adet birey sonraki nesil için aday populasyonu oluşturur.

6. Adım 5'te oluşturulan aday populasyondaki her bir birey için [0,1] aralığından rasgele bir r sayısı üretilir. Eğer $r < p_c$ oluyorsa ilgili birey çaprazlama işlemine tabi tutulur. Çaprazlama işlemi gerçekleştirilirken iki birey $birey1 = \{a_1, a_2, \ldots, a_p\}$ ve $birey2 = \{b_1, b_2, \ldots, b_p\}$ ebeveyn olarak

seçilir. a_i ve b_i ebeveyn bireylerin parametre değerlerini ifade etmek üzere p parametreli problem için [1,p] aralığından rasgele bir tamsayı değeri üretilir. Rasgele üretilen tamsayı değeri(i) ile ebeveyn bireylerin çaprazlama noktası oluşturulur. Çaprazlama noktası oluşturulduktan sonra $0 < \alpha_i < 1$ rasgele sayısı üretilir. $a_i^{'} = \alpha_i a_i + (1-\alpha_i) b_i$ ve $b_i^{'} = \alpha_i b_i + (1-\alpha_i) a_i$ olmak üzere yeni evlatlar,

$$evlat1 = \left\{ a_1, a_2, \ldots, a_i^{'}, b_{i+1}, \ldots, b_n \right\}$$

$$evlat2 = \left\{ b_1, b_2, \ldots, b_i^{'}, a_{i+1}, \ldots, a_n \right\}$$

şeklinde oluşturulur. Bu işlem N_{pop} kez tekrarlanarak çaprazlama işlemi tamamlanır.

7. Çaprazlama işleminden sonra populasyondaki her bir birey için [0,1] aralığından rasgele bir r sayısı üretilir. Eğer $r < p_m$ oluyorsa ilgili birey mutasyon işlemine tabi tutulur. Mutasyon işlemine tabi tutulacak birey ($birey1 = \left\{ a_1, a_2, \ldots, a_p \right\}$) seçilir. Seçilen bireyin mutasyona uğrayacak geni [1,p] aralığından rasgele bir tamsayı değeri (i) üretilerek belirlenir. Populasyondan rasgele bir birey $birey2 = \left\{ b_1, b_2, \ldots, b_p \right\}$ seçildikten sonra birinci bireyin mutasyona uğrayacak geni (a_i) için $0 < \alpha_i < 1$ rasgele sayısı üretilir. Mutasyona uğrayan yeni gen $a_i^{''} = \alpha_i a_i + (1-\alpha_i) b_i$ ve b_i ile değiştirilerek mutasyon işlemi gerçekleştirilir. Bu işlem N_{pop} birey için uygulanarak mutasyon işlemi tamamlanır.

8. Bu adımda çaprazlama ve mutasyon işlemi sonunda oluşturulan bireyler bir sonraki nesil için başlangıç populasyonu ile değiştirilir. Bu nesildeki en küçük hkt_i^t değerine sahip birey bir sonraki nesildeki başlangıç populasyonuna gönderilerek en iyi çözüm saklanmış olur böylece elitist seçim stratejisi gerçekleştirilerek sonraki nesillerde daha küçük hkt_i^t değerlerinin korunması sağlanır.

9. Max_{nesil} değerine ulaşıldıysa işlem durdurulur ve en küçük hkt_i^t değerine sahip bireyler çözüm olarak belirlenir. Max_{nesil} değerine ulaşılana kadar $t = t + 1$ alınarak 2. adıma dönülür.

Çizelge 4.2. Örnek1 verisi için parametre tahmin değerleri

Parametre	**Gauss Newton**	**NLGA1**
$\hat{\theta}_1$	5.659	5.702
$\hat{\theta}_2$	2.807	2.627
$\hat{\theta}_3$	0.985	1.019
HKT	9.386	9.159

Gauss-Newton yöntemi için parametre başlangıç noktaları $\theta_1 = 3.5$ $\theta_2 = 4$ ve $\theta_3 = 1.5$ alınarak elde edilen modellerin tahmin değerleri ve NLGA1 ile elde edilen tahmin değerleri çizelge (4.2.)de verilmiştir. Çizelge (4.2.)'deki sonuçlara göre Gauss Newton algoritması ile elde edilen parametre tahmin değerleri incelendiğinde NLGA1 ile elde edilen hata kareler toplamının daha düşük olduğu görülmektedir.

Örnek 2 : $Y_2 = \theta_4 + \theta_5 x_t + \theta_6 x_t^2 + e_{2t}$, 3 parametreli modelin parametre değerleri $\theta_4 = 2$, $\theta_5 = 5$, $\theta_6 = 2$ ve e_{1t} hata değerleri normal dağılımdan üretilerek elde edilen değerler çizelge (4.1.)'de verilmişti. Bu değerler kullanılarak NLGA1 ile elde edilen tek değişkenli modele ilişkin parametre tahmin değerleri çizelge (4.3.)'de verilmiştir.

Çizelge 4.3. Örnek2 verisi için parametre tahmin değerleri

Parametre	Gauss Newton	NLGA1
$\hat{\theta}_4$	2.041	2.083
$\hat{\theta}_5$	5.425	5.420
$\hat{\theta}_6$	2.193	2.183
HKT	39.526	39.535

Gauss-Newton yöntemi ve NLGA1 ile elde edilen modellerin tahmin değerleri çizelge (4.3.)'de verilmiştir. Çizelge (4.3.)'deki sonuçlara göre Gauss Newton algoritması ile elde edilen parametre tahmin değerleri ile NLGA1 ile elde edilen değerlerin birbirine yakın olduğu görülmektedir.

4.2. Çok Değişkenli Lineer Olmayan Genetik Algoritma (NLGA2)

Çok değişkenli lineer olmayan model (3.1) eşitliği ile verilen *M* tane tek değişkenli lineer olmayan modelin (3.3) eşitliği ile belirtilmesi durumunda bağımlı değişkenin gözlem değerlerinin aynı bireyler üzerinde alınması veya *M* modelin ortak parametre içermesi sebebiyle bağımlı değişkenlerin ayrı ayrı düşünülmemesi gerektiği durumlarda kullanılmaktadır. *M* modelin ayrı ayrı tek değişkenli model olarak düşünülmesi gerek bağımlı değişkenler arasındaki ilişkilerin göz ardı edilmesi veya ortak parametre olması durumunda ortak parametrelerin farklı değerlerle tahmin edilmesi sorununu ortaya çıkarmaktadır.

Çok değişkenli lineer olmayan modelde EKK tahmini için Gauss Newton algoritması 3. bölümde verilmişti. Bu bölümde çok değişkenli lineer olmayan modelin parametre tahmini için Çok Değişkenli Lineer Olmayan Genetik Algoritma (NLGA2) önerilmiştir. NLGA2'nin işleyiş adımları verildikten sonra bir örnek üzerinde uygulama sonuçları verilmiştir.

NLGA2'de kullanılan parametreler,

N :Gözlem sayısı

$\underline{Y}_i$: i. model için bağımlı değişken gözlem değerleri i=1,2,…,M

X_i : Bağımsız değişken gözlem değerleri i=1,2,…,N

p : Modeldeki toplam parametre sayısı

$f_i(\underline{x}_t, \underline{\theta})$: i. model için tepki fonksiyonu

θ_i : Bilinmeyen i.parametre

p_c : Çaprazlama olasılığı

p_m : Mutasyon olasılığı

alt_i : i. parametre alt sınır değeri

$üst_i$: i. parametre üst sınır değeri

N_{pop} : Her bir nesildeki populasyon büyüklüğü

Max_{nesil} : Maksimum nesil büyüklüğü

hkt_i^t : t. nesildeki i.bireyin uygunluk değeri ($t = 1,2,...,Max_{nesil}$)

f_i^t : t. nesildeki i. bireyin tepki fonksiyonu değeri ($t = 1,2,...,Max_{nesil}$)

olmak üzere NLGA2'nin işleyiş adımları aşağıda verilmiştir.

1. $f_\alpha(\underline{x}_t, \underline{\theta})$, $\alpha = 1,\ldots,M$ tepki fonksiyonları olmak üzere, M model için NLGA1 kullanılarak her bir tepki fonksiyonu için parametre tahminleri ($\underline{\hat{\theta}}_\alpha$) elde edilir.

2. M modelin parametre tahmin değerleri kullanılarak her bir model için,

$$\underline{\hat{e}}_\alpha = \underline{Y}_\alpha - \underline{f}_\alpha\left(\underline{x}_t, \underline{\hat{\theta}}_\alpha\right),\ \alpha = 1,2,...,M$$

hata vektörleri elde edilir. M model için elde edilen hata vektörleri sütunlar halinde yazılarak,

$$\hat{E} = [\underline{\hat{e}}_1 \ \vdots \ \underline{\hat{e}}_2 \ \vdots \ \ldots \ \vdots \ \underline{\hat{e}}_M]$$

hata matrisi oluşturulduktan sonra tüm modellerin hatalarının varyans-kovaryans matrisinin tahmini $\hat{\Sigma} = \frac{1}{\mathrm{N}} \hat{E}'\hat{E}$ elde edilir.

3. M modelden her birinin bilinmeyen parametre sayısı p_α ve M modeldeki toplam parametre sayısı $p = \sum_{\alpha=1}^{\mathrm{M}} p_\alpha$ olmak üzere, M modele ait bilinmeyen parametre vektörü $\underline{\theta} = (\theta_1, \theta_2, \ldots, \theta_p)$ için $(alt_i < \theta_i < üst_i)$ olacak şekilde N_{pop} x p adet parametre rasgele üretilerek başlangıç populasyonu oluşturulur.

4. Her bir $f_\alpha(\underline{x}_t, \underline{\theta})$ tepki fonksiyonunun değeri başlangıç populasyonunda oluşturulan parametre değerleri için hesaplanır. Hesaplanan tepki fonksiyonlarının değerleri alt alta yazılarak,

$$\underline{f}(\underline{\theta}) = \begin{bmatrix} f_1(\underline{x}_t, \theta) \\ \vdots \\ f_{\mathrm{M}}(\underline{x}_t, \theta) \end{bmatrix}$$

oluşturulur ve bu değerlerden bağımlı değişken gözlem değerlerinin alt alta yazılmasıyla oluşturulan,

$$\underline{Y} = \begin{bmatrix} Y_1 \\ \vdots \\ Y_M \end{bmatrix}$$

vektörü kullanılarak t. populasyondaki i. bireyin uygunluk değeri $hkt_i = \left(\underline{Y} - \underline{f}(\underline{\theta})\right)' (\Sigma \otimes I_n)^{-1} \left(\underline{Y} - \underline{f}(\underline{\theta})\right)$ ile hesaplanır.

5. hkt_i^t değerlerine göre bireyler küçükten büyüğe sıralanır. Uygunluk değeri en kötü olan populasyon büyüklüğünün %75'i kadar birey yeniden üretilerek

hkt_i^t değerleri hesaplanır ve bireyler küçükten büyüğe sıralanır. Uygunluk değerlerine göre sıralanan bireylerin seçilme olasılıkları (4.2) eşitliğiyle hesaplanır.

6. Hesaplanan *p(i)* değerlerine göre rulet çemberi seçim yöntemi N_{pop} adet birey sonraki nesil için aday populasyonu oluşturur.

7. Oluşturulan aday populasyondaki her bir birey için p_c çaprazlama olasılığıyla çaprazlama NLGA1'de kullanılan çaprazlama operatörü uygulanır

8. Çaprazlama işleminden sonra populasyondaki her bir birey için p_m çaprazlama olasılığıyla NLGA1'de kullanılan mutasyon operatörü uygulanır.

9. Çaprazlama ve mutasyon işlemi sonunda oluşturulan bireyler bir sonraki nesil için başlangıç populasyonu ile değiştirilir. Bu nesildeki en küçük hkt_i^t değerine sahip birey bir sonraki nesildeki başlangıç populasyonuna gönderilerek elitist seçim stratejisi uygulanır. Max_{nesil} değerine ulaşıldıysa işlem durdurulur ve en küçük hkt_i^t değerine sahip bireyler çözüm olarak belirlenir. Max_{nesil} değerine ulaşılana kadar $t = t + 1$ alınarak 4. adıma dönülür.

4.2.1. Jackknife tahmini

Bu yöntem bir tahmin ediciye ait yan miktarını azaltmak amacıyla kullanılmaktadır. $X_1, X_2, \ldots, X_n$ gerçek değerli $\underline{\theta}$ parametresine sahip bir kitlenin n birimden oluşan örneklemi olsun. Tüm gözlemler için $\underline{\theta}$ parametresinin tahmin edicisi $\hat{\theta}_n = \hat{\theta}_n(X_1, X_2, \ldots, X_n)$ olarak gösterilsin. Örneklem $m = \left[\left|\frac{n}{k}\right|\right]$ büyüklüğünde *k* eşit gruba ayrılır. Her adımda bir grup çıkarılarak geriye kalan gözlemlerle $\underline{\theta}$ parametresi için $\underline{\hat{\theta}}_i$ tahmini elde edilir.

$$X_1, X_2, \cdots, X_{(i-1)m}, X_{(im-1)}, \cdots, X_n \quad i = 1,2,\ldots,k$$

Sırasıyla *i*. grup çıkartılarak geriye kalan *n-m* gözlemden $\underline{\theta}$ parametresi için elde edilen $\underline{\hat{\theta}}_i$ tahmin edicisine Jackknife tahmin edicisi denir ve

$$J_i = k\hat{\theta} - (k-1)\hat{\theta}_i, \qquad i = 1,2,\ldots k$$

olmak üzere,

$$J(\hat{\theta}) = \frac{1}{k}\sum_{i=1}^{k}(k\hat{\theta} - (k-1)\hat{\theta}_i) = \frac{1}{k}\sum_{i=1}^{k} J_i$$

ifadesi Jackknife tahmin edicisidir.

4.2.2. Bootstrap tahmini

Bu yöntemde *n* gözlemden oluşan bir veri setinde $\hat{\theta}_n = \hat{\theta}_n(X_1, X_2, \ldots, X_n)$ tahmin edicisinin bootstrap tahmini V_B aşağıdaki şekilde bulunur.

Adım1 :Gözlem değerlerinden iadeli olarak n kere çekiliş yapılır. Bu çekilişler, $X_1^*, X_2^*, \ldots, X_n^*$ ile gösterilir.

Adım2 :Her rasgele çekilen gruptan ele alınan tahmin edici hesaplanır. $\hat{\theta}^*_n = \hat{\theta}^*_n(X^*_1, X^*_2, \ldots, X^*_n)$

Adım 3: Adım1 ve Adım2 B kere tekrar edilsin. Elde edilen tahmin değerleri olarak gösterilirse, Bootstrap tahmini,

$$V_B = \frac{1}{B}\sum_{i=1}^{B}(\hat{\theta}_i^* - \frac{1}{B}\sum_{j=1}^{B}\hat{\theta}_j^*)^2$$

şeklinde elde edilir.

Örnek 3: $Y_1 = \theta_1 x_t + \theta_2 e^{\theta_3 x_t} + e_{1t}$, modelin parametre değerleri $\theta_1 = 5$, $\theta_2 = 3$, $\theta_3 = 1$ ve $Y_2 = \theta_4 + \theta_5 x_t + \theta_6 x_t^2 + e_{2t}$,modelin parametre değerleri $\theta_4 = 2$, $\theta_5 = 5$, $\theta_6 = 2$ olmak üzere çizelge 4.1.'deki veriler kullanılarak çok değişkenli lineer

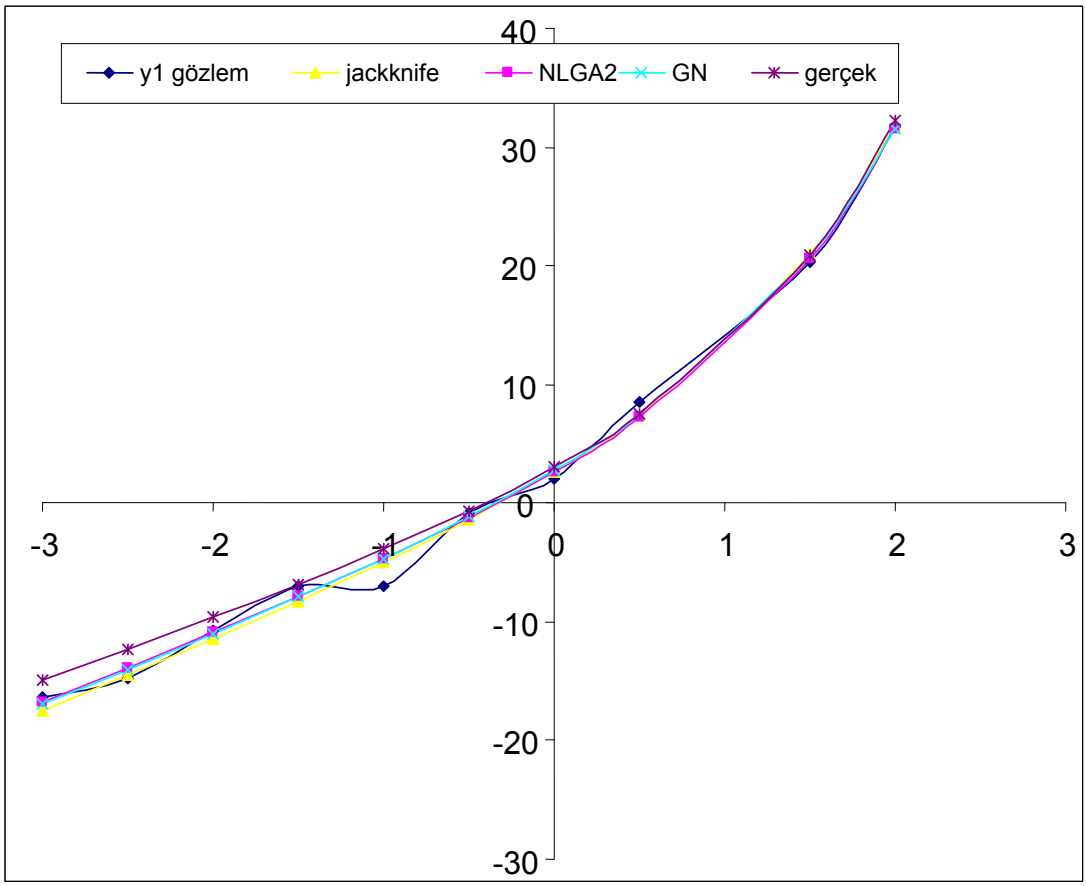
olmayan model için parametre tahmin değerleri Gauss Newton algoritması ve NLGA2 ile elde edilmiş ve sonuçları Çizelge 4.4.'de verilmiştir.

Çizelge 4.4. Çok değişkenli model parametre tahmin sonuçları

Parametre	**Gauss Newton**	**NLGA2**	**Jackknife**	**Bootstrap**
$\hat{\theta}_1$	4.991	5.642	5.894	0.810
$\hat{\theta}_2$	2.748	2.64	2.636	1.877
$\hat{\theta}_3$	1.022	1.018	1.017	0.054
$\hat{\theta}_4$	1.961	2.034	3.138	4.931
$\hat{\theta}_5$	4.991	5.642	5.894	0.810
$\hat{\theta}_6$	1.865	2.211	1.532	2.401
HKT	19.999	19.987	37.337	63.410

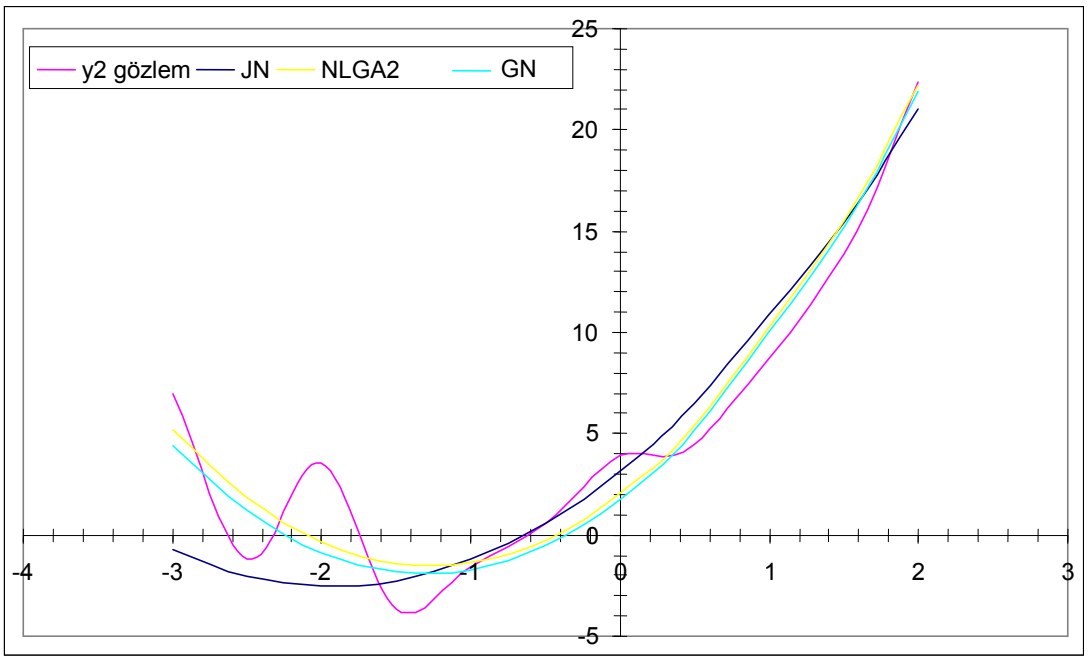
Çizelge 4.4.' de görüldüğü üzere Gauss Newton yöntemiyle elde edilen parametre tahmin değerleri ile NLGA2 ile elde edilen parametre tahmin değerlerinden elde edilen hata kareler toplamlarına göre NLGA2'nin daha düşük HKT değerine sahip olduğu görülmektedir.

Gerek Jackknife ve Bootstrap metodu ile parametre tahminlerindeki yan miktarını azaltmak amacıyla kullanılmıştır. Ancak elde edilen Çizelge 4.4.' deki sonuçlara göre Gauss Newton ve NLGA2 ile elde edilen tahmin sonuçlarının yan miktarlarının diğerlerinden daha az olduğu görülmektedir. Gerek yan miktarı ve gerekse hata kareler toplamının küçük çıkması NLGA2 ile elde edilen tahmin sonuçlarının Gauss Newton ile elde edilen sonuçlardan daha iyi olduğunu göstermektedir.



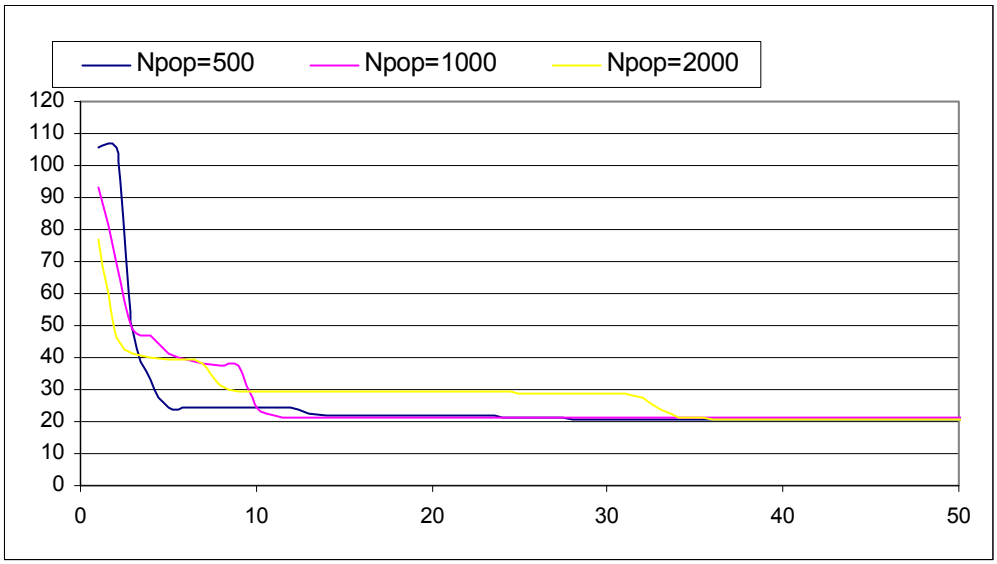
Şekil 4.1. Örnek1 verisi için tahmin değerleri ile gözlem değerlerinin saçılım grafiği

Çizelge 4.4. ile verilen parametre tahmin değerleri kullanılarak örnek1 verisi için elde edilen tahmin değerleri ile gözlem değerlerinin saçılım grafiği Şekil 4.1.'de verilmiştir. Şekil 4.1. incelendiğinde Gauss Newton ve NLGA2 ile elde edilen tahmin değerlerinin gözlem değerlerine yakın olduğu görülmektedir.



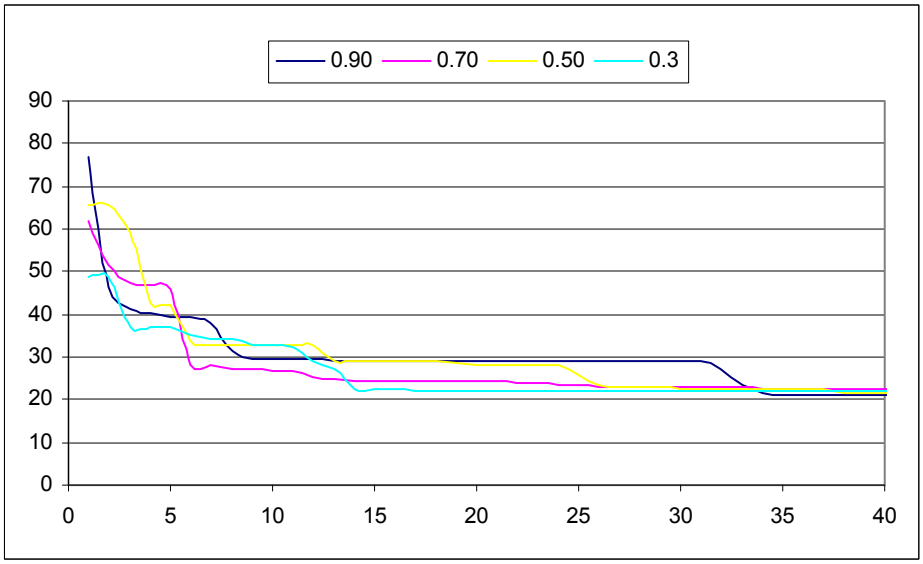
Şekil 4.2. Örnek2 verisi tahmin ve gözlem değerlerinin saçılım grafiği

Çizelge 4.4. ile verilen parametre tahmin değerleri kullanılarak örnek2 verisi için elde edilen tahmin değerleri ile gözlem değerlerinin saçılım grafiği Şekil 4.2.'de verilmiştir. Şekil 4.2. incelendiğinde Gauss Newton ve NLGA2 ile elde edilen tahmin değerlerinin gözlem değerlerine daha yakın olduğu görülmektedir.



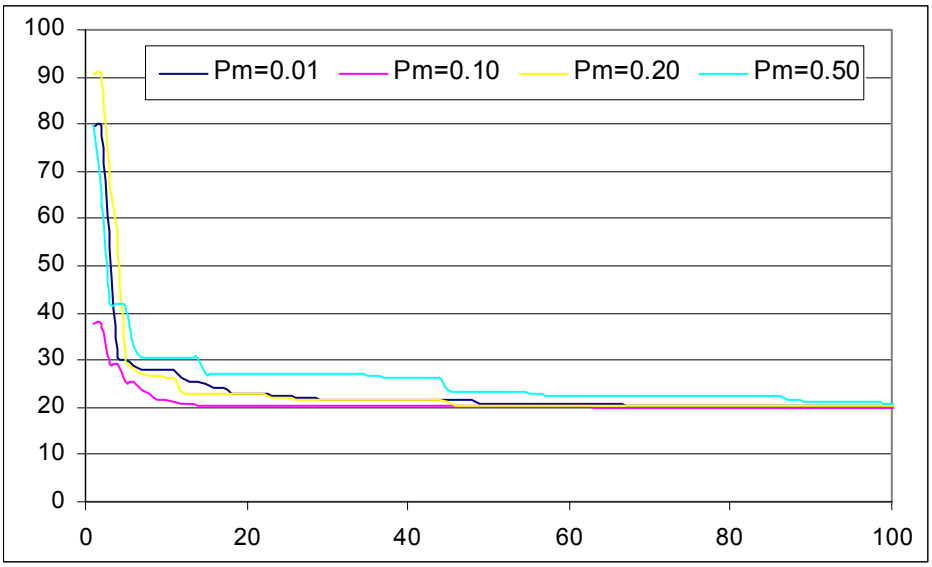
Şekil 4.3. Populasyon büyüklüğüne göre HKT değerlerinin değişimi

Şekil 4.3.'de NLGA2 için farklı populasyon büyüklüklerine göre her bir nesil için elde edilen HKT değerlerinin değişimi verilmiştir. Şekil 4.3.'e göre başlangıç populasyonunun büyüklüğü arttıkça arama uzayında daha çok noktaya ulaşılmakta ve daha küçük HKT değerine sahip bireyler bulunmaktadır.



Şekil 4.4. Çaprazlama olaılıklarına göre HKT değerlerinin değişimi

Şekil 4.4.'de NLGA2 için farklı çaprazlama olasılıkları ile her bir nesil için elde edilen HKT değerlerinin değişimi verilmiştir. Şekil 4.4.'e göre populasyon büyüklüğünün artması çaprazlama operatörünün etkinliğini azaltmakla beraber çözümü değiştirmemektedir.



Şekil 4.5. Mutasyon olasılıklarına göre HKT değerlerinin değişimi

Şekil 4.5.'de NLGA2 için farklı mutasyon olasılıkları ile her bir nesil için elde edilen HKT değerlerinin değişimi verilmiştir. Şekil 4.5.'e göre mutasyon olasılığı arttıkça gerçek çözümden daha uzak noktalarda arama yapılmakta ve HKT değerleri daha büyük çıkmaktadır. Şekil 4.5.'e göre mutasyon olasılığı p_m=0.10 alınarak elde edilen değerler diğer seçimlerden daha iyi sonuçlar vermektedir.

5. SONUÇ

Çok değişkenli lineer olmayan modeller iki sebepten tek değişkenli modelden farklı şekilde düşünülmekte ve çözülmektedir. Bunlardan birincisi model hatalarının ilişkili olması, ikincisi ise model fonksiyonlarının bazı parametrelerinin aynı olmasıdır.

Çok değişkenli lineer olmayan modellerde tahmin ediciler analitik olarak elde edilemediği için iteratif teknikler kullanılmaktadır. Lineer olmayan ifadelerin optimizasyonu için kullanılan yöntemler bazen yetersiz olmaktadır. Bu yöntemler optimal noktaya her zaman yakınsayamama, yakınsasalar bile global optimum yerine yerel optimum noktaya yakınsayabilmektedirler.

Çok değişkenli lineer olmayan modellerin parametre tahmini için literatürde kullanılan iteratif tekniklerden Gauss-Newton metodu parametre tahmininde gerek başlangıç noktasına olan duyarlılıkları ve gerekse fonksiyonların türevlenebilir olması koşulunu gerektirmektedir. Bu koşulların her zaman sağlanamaması bazı problemlerin çözümü zorlaşmaktadır. Bu çalışmada mevcut yöntemlere bir alternatif olarak genetik algoritma önerilmiştir.

Genetik algoritmalar parametre tahmininde kullanım için fonksiyonların türevlerine ihtiyaç duymamakta ve başlangıç noktasına göre optimum çözüm değişmemektedir. Belirtilen sebeplerden alternatif olarak düşünülen yöntem için Delphi 5 görsel programlama dilinde iki program yazılmıştır. Bu amaçla Pentium M işlemcili, 2 GB ram konfigürasyonlu bilgisayar kullanılmıştır.

Delphi 5 görsel programlama dilinde NLGA1 ve NLGA2 isimli iki program kodlanmıştır. NLGA1 tek değişkenli lineer olmayan modelin parametre tahmini için, NLGA2 ise çok değişkenli lineer olmayan modelin parametre tahmini için kullanılmıştır.

Örnek veri setleri hataları bağımsız ve normal dağılıma sahip rasgele dağılımlardan gelecek şekilde belirli parametre değerleri ve fonksiyonel biçimi belirtilen fonksiyonlardan farklı gözlem sayılarından üretilmiştir.

NLGA1 ve NLGA2 program kodları sırasıyla Ek1, Ek2 de verilmiştir. Elde edilen sonuçlara bakıldığında tek değişkenli lineer olmayan modellerin parametre tahminlerinde NLGA1'in klasik yöntemle benzer sonuçlar verdiği hatta bazen daha iyi sonuçlar verdiği gösterilmiştir.

Çok değişkenli lineer olmayan modellerin parametre tahminlerinde NLGA2 ile elde edilen hata kareler toplamının Gauss-Newton yöntemiyle elde edilen sonuçlardan daha küçük olduğu dolayısıyla daha iyi sonuç verdiği gösterilmiştir.

Populasyon büyüklüğünün seçimi genetik algoritmaların çözüme yakınsama hızını doğrudan etkilemektedir. NLGA2 üç farklı populasyon büyüklüğü için (500,1000,2000) çalıştırılmış ve HKT değerlerinin değişimi Şekil 4.3'de verilmiştir. Populasyon büyüklüğünün 2000 olarak seçildiği durumda arama uzayının daha kapsamlı bir kısmı tarandığı için daha küçük HKT değerleri elde edilmiş ve çözüme daha kısa süre yaklaşıldığı gösterilmiştir.

NLGA2 ile elde edilen sonuçlar program üzerinde farklı seçim kriterlerine göre irdelendiğinde populasyon büyüklüğü arttıkça çaprazlama operatörünün etkisinin azaldığı, aynı şekilde populasyon büyüklüğü arttıkça mutasyon olasılığının düşük değerlerinin daha iyi sonuçlar verdiği Şekil 4.4 ve Şekil 4.5. 'de gösterilmiştir.

Sonuç olarak bu çalışmada önerilen her iki algoritma için elde edilen sonuçlar Gauss Newton ile elde edilen sonuçlarla benzerlik göstermekle beraber bazı durumlarda daha iyi sonuçlar vermektedir. Genetik algoritmanın sahip olduğu kullanım kolaylıkları açısından, Gauss Newton yönteminden daha avantajlı olduğu gösterilmiştir.

KAYNAKLAR

Ahmad, M.,Zhang,L.,Readle,J.C. 1998. Parameter Estimation of A Heating System Using A Genetic Algorithm, UKACC International Conference On Control'98 :195-200.

Altunkaynak, B.,Esin,A. 2004 The genetic algorithm method for parameter estimation in nonlinear regression.G.Ü. Fen Bilimleri Dergisi.17(2): 43-51.

Bäck, T. 1996. Evolutionary algorithms in theory and practice: evolution strategies, evolutionary programming,genetic algorithms, Oxford University Press, New York.

Bard, Y. 1964, Nonlinear Parameter Estimation, Academic Prens, Inc. New York.

Bates, D.M.,Watts,D.G. 1988. Nonliner Regression and Its Applications, John Wiley&Sons, Usa.

Bird, H.A. 1974. Estimable Functions in the Nonlinear Model, Ph. D. Thesis Kansas State University, Manhattan.

Booker, L. 1987, Improving search in genetic algorithms, Genetic algorithms and Simulated annealing,61-73.

Bruin, R.L. 1971. Testing Hypothesis For Nonlinear Models, Ph. D. Thesis Kansas State University, Manhattan.

Buckley, J.J., Hayashi,Y. 1994. Fuzzy Genetic Algortihm and Applications, Fuzzy Sets and Systems,61:129-136.

Chang, W.D. 2006 An improved real-coded genetic algorithm for parameters estimation of nonlinear systems. Mechanical Systems and Signal Processing. 20:236–246.

Chatterjee, S.,Laudato,H.,Lynch,L.A. 1996. Genetic Algorithms And Their Statistical Applications An Introduction, Computational Statistic And Data Analysis, 22:633-651.

De Jong, K.A. 1975. An Analysis of the Behavior of a Glass of Genetic Adaptive System, Ph.D. Thesis, Michigan University, Usa..

Eter, D.M.,Hicks,M.J.,Cho,K.H. 1982, Recursive Adaptive Filter Design Using an Adaptive Genetic Algorithm, Proceedings of IEEE International Conference On ASSP:635-638.

Gallant, A. R. 1975, Testing a Subset of the Parameters of a Nonlinear Regression Model, JASA, 70, 927-932.

Gallant, A.R. 1987. Nonlinear Statistical Models, John Wiley&Sons, Usa.

Genç, A. 1997. Çok Değişkenli Lineer Olmayan Regresyon Modelleri Parametre Tahmini ve Hipotez Testleri, Ankara Üniversitesi Fen Bilimleri Enstitüsü Doktora Tezi, Ankara.

Goldberg, D.E. 1989. Genetic Algorithms in Search, Optimization, and Machine Learning, Addison-Wesley, Massachusetts.

Goldberg, D.E. 2004. Frontiers of Evolutionary Computation, Kluwer Academic Publishers, New York,

Goldfeld, S.M.,Quandt,R.E.,1972 Nonlinear Methods in Econometrics, Amsterdam London :North Holland.

Gülşen, M. 1997. Curve Fitting Using A Hierarchical Genetic Algorithm Framework, Ph. D. Thesis University of Pittsburg, Usa.

Gülşen, M.,Smith,A.E. 1999. A Hierarchical Genetic Algorithm For System Identification and Curve Fitting With A Supercomputer Implement, IMA Volumes In Matematics and Its Applications,vol 111.

Gülşen, M.,Smith,A.E.,Tate,D.M. 1995. A Genetic Algorithm Approach To Curve Fitting, International Journal of Production Research vol 33,7:1911-1923.

Günay, S. 1978. Doğrusal Olmayan Regresyon Fonksiyonlarında En Küçük Kareler Tahmini ve Bu Tahminlerin Yorumu, Hacettepe Üniversitesi Fen Fakültesi İstatistik Bölümü , Ankara.

Hartley, H.O., Booker, A. 1963, Nonlinear Least Squares Estimation, Annals of Math. Stat., 36, 638-650.

Hasançebi, O., Erbatur, F. 2000 Evaluation of crossover techniques in genetic algorithm based optimization structural design. J Comput Struct .78:435–48.

Haupt, R.L.,Haupt,S.E. 1998. Practical Genetic Algorithms, Willey-Interscience Publication, Usa.

Holland, J.H. 1975, Adaptation in Natural and Artificial Systems, Ann Arbor: The University of Michigan Press.

Hwang, H.S.,Woo,K.B. 1995. Linguistic Fuzzy Model Identification, IEE Proceedings-Control Theory Applications, vol 42,6:537-544.

Hwang, S.F. He, R.S.2006 A hybrid real-parameter genetic algorithm for function optimization. Advanced Engineering Informatics. 20: 7–21.

Johnson, R.A., Wichern, D.W., 1992, Applied multivariate statistical analysis, Prentice hall, New Jersey.

Karr, C.L.,Stanley,D.A.,Scheiner,B.J. 1991. Genetic Algorithm Applied to Least Squares Curve Fitting, Bureau of Mines Report.

Karr, C.L.,Weck,B.,Massart,D.L.,Vankeerberghen,P. 1995 Least median squares curve fitting using a genetic algorithm. Engineering applications in artificial intelligence. Vol 8: No2:177-189.

Koç, İ.O. 2001. Doğrusal Olmayan Regresyon Modellerinin Parametrelerinin Genetik Algoritma İle Tahmin Edilmesi, Osman Gazi Üniversitesi Fen Bilimleri Enstitüsü Yüksek Lisans Tezi, Eskişehir.

Koza, J. R. 1994. Genetic Programming II, Mit Press, Cambridge.

Leehter, Y., Sethares,W. 1994. Nonlinear Parameter Estimation Via The Genetic Algorithm, IEEE Transactions On Signal Processing Vol 42, 4:927-935.

Marquart, D.W. 1963 An Algorithm for Least Squares Estimation of Nonlinear Parameters. Journal of The Society For Industrial and Applied Mathematics,11:431-441.

Michalewicz, Z. 1996. Genetic Algorithms + Data Structures = Evolution Programs 3.ed., Springer, Usa.

Milliken, G.A.1988, Applications of Nonlinear Statistical Models,Kansas State University.

Mitchell, M. 1999. An Introduction to Genetic Algorithms, 5th Ed. Mit pres,England.

Moros, R,Kalies,H.,Rex,H.G. 1996. A Genetic Algorithm For Generating Initial Parameter Estimations For Kinetic Models Of Catalytic Processes, Computers and Chemical engineering vol 20,10:1257-1270.

Olinsky, A.D.,Quinn,J.T.,Mangiameli,P.M.,Chen,S.K. 2004 A genetic algorithm approach to nonlinear least squares estimation. International journal of mathematical education in science and technology.vol 35:2:207-217.

Petridis, V.,Paterakis,E.,Kehagias,A. 1998. A Hybrid Neural Genetic Multimodel Parameter Estimation Algorithm, IEEE Transactions On Neural Networks vol 9,5: 862-876.

Reeves, C.R., Rowe, J.E. 2002. Genetic Algorithms Principles and Perspectives A Guide to GA Theory, Kluwer Academic Pres, Usa.

Ross, G. J. S. 1990, Nonlinear Estimation, Springer -Verlag, New York, Inc.

Rothlauf,F. 2006. Representations for Genetic and Evolutionary Algorithms, 2. ed.,Springer,Berlin.

Saruhan, H.,Uygur,İ. 2003. Design Optimization of Mechanical Systems Using Genetic Algorithms, S.Ü. Fen Bilimleri Enstitüsü Dergisi, 7.Cilt,2.Sayı.

Seber, G.A.F.,Wild, C.J. 1989. Nonlinear Regression, John Wiley&Sons, Usa.

Siepman, H.R.,Yang,S.S. 1994. Generalized Least Square Estimation of Multivariate Nonlinear Models With Missing Data, Communication Statistics Theory M,23: 1565-1579.

Spears, W. M. and DeJong, K. A. 1991, An analysis of multi-point crossover. In Rawlins, G. J. E., editor, Foundations of Genetic Algorithms-1, pages 301-315. Morgan Kauffman.

Srinivas, M.,Patnaik, L.M. 1994. Genetic Algorithms: A Survey, Computer, Vol:27, 6, 17-26.

Syswerda, G. 1989, Uniform crossover in genetic algorithms. In Proceedings of the Third International Conference on Genetic Algorithms, pages 2-8. Morgan Kauffman.

Taub, H.Schillind, D, 1986,Principles of communication systems Edition: 2nd ed. McGraw-Hill, New york.

Tez(Aysever), M. 1982. Lineer Olmayan Modellerde Parametrelerin Kestirimlerinin İncelenmesi, Ankara Üniversitesi Fen Fakültesi Yüksek Lisans Tezi, Ankara.

Yang, M., Zhang, X.,Li, X.,Wu,X. 2002 A hybrid genetic algorithm for the fitting of models to electrochemical impedance data. Journal of Electroanalytical Chemistry, 519:1–8.

Ek-1 NLGA1 Program Kodları

```
var

k8,p1,k10,p2,secim,k9,gecici,pm,pc,tbbhkt:double;

g,ts,gs,k,i1,kkk,k1,k2,k3,ps,populasyon,n,i,j,m,t:integer;

x:array[1..20] of double;
```

```
y:array[1..20] of double;
bp:array[1..3000,1..2]of double;
hkt:array[1..3000,1..9]of double;
eniyi:array[1..10000,1..2] of double;
p:array[0..3000]of double;
sira:array[1..3000]of integer;
aday:array[1..3000,1..2]of double;
f:array[1..20,1..3000]of double;
implementation
{$R *.DFM}
procedure TForm1.Button1Click(Sender: TObject);
begin
randomize;
gs:=12;   // gözlem sayısı
populasyon:=2000; // populasyon büyüklüğü
ts:=50; // iterasyon sayısı
pm:=0.10; //mutasyon olasılığı
pc:=0.90; // Çaprazlama olasılığı
FOR i:=1 TO gs do
begin
x[i]:=gr.NumberRC[i,1];
y[i]:=gr.NumberRC[i,2];
end;
for i:=1 to populasyon do
begin
bp[i,1]:=random*400;
bp[i,2]:=random*15;
end;
 for g:=1 to ts do
```

```
begin
     FOR m:=1 TO populasyon do
    begin
          hkt[m,4]:=0;
          For i:=1 to gs do
          begin
               f[i,m]:=bp[m,1]*(1-exp(-bp[m,2]*x[i]));
               hkt[m,4]:=hkt[m,4]+sqr(f[i,m]-y[i]);
          end;
    end;
    FOR i:=1 TO populasyon do
    begin
          hkt[i,1]:=i;
          hkt[i,2]:=bp[i,1];
          hkt[i,3]:=bp[i,2];
    end;
    for i:=1 to populasyon do
    for j:=1 to 5 do
     f1book1.numberrc[i,j]:=hkt[i,j];
     f1book1.Sort3(1,1,populasyon,5,true,4,1,1);
     for i:=1 to populasyon do
     for j:=1 to 5 do
     hkt[i,j]:=f1book1.numberrc[i,j];
    FOR i:=1 TO populasyon do
    begin
          hkt[i,5]:=populasyon+1-i;
    end;
      for k:=1 to 2 do
      begin
```

```
  eniyi[g,k]:=hkt[1,k+1];
  end;
  for i:=1 to populasyon do
  begin
  gr.numberrc[20+i,1]:=i;
  p[i]:=(2*hkt[i,5])/(populasyon*(populasyon+1));
  end;
  kkk:=populasyon-10;
  for k:=kkk to populasyon do
  begin
  bp[k,1]:=random*400;
  bp[k,2]:=random*10;
  end;
 //Rulet Çarkı Seçim aşaması
 p[0]:=0;
 FOR i:=1 TO populasyon do
 begin
      secim:=random;
      P1:= 0; P2:= 0;
      FOR j:= 1 to populasyon do
      begin
           P1:= P1+ p[j-1];
           P2:= P2+ p[j];
           if (secim >P1) and (secim<P2) then sira[i]:=j;
      end;
 end;
 button2.click;
//çaprazlama
 i:=0;
```

```
        while (i<=populasyon)do
        begin
         i:=i+1;
         k8:=random;
            IF k8< pc THEN
            begin
              k:=1+random(2);
              k9:=random;
              for t:=k to 2 do
              begin
              gecici:=aday[i,t];
              aday[i,t]:=k9*aday[i+1,t]+(1-k9)*aday[i,t];
              aday[i+1,t]:=(1-k9)*gecici+k9*aday[i+1,t];
              end;
              i:=i+1;
           end;//if
        end;
//mutasyon
        FOR i:= 1 TO populasyon do
        begin
          k10:=random(populasyon);
              IF k8 < pm THEN
              begin
                 k9:=random;
                  for k:=1 to 2 do
                  begin
                  gecici:=aday[i,k];
                  aday[i,k]:=k9*aday[i+1,k]+(1-k9)*aday[i,k];
                  aday[i+1,k]:=k9*gecici+(1-k9)*aday[i+1,k];
```

```
                end;
            end;
        end;
       FOR i:=1 TO populasyon do
        begin
            FOR j:= 1 TO 2 do
            begin
            bp[i,j]:=aday[i,j];
            end;
        end;
              bp[1,1]:=eniyi[g,1];
              bp[1,2]:=eniyi[g,2];
       gr.NumberRC[g,4]:=eniyi[g,1];
       gr.NumberRC[g,5]:=eniyi[g,2];
       gr.NumberRC[g,6]:=hkt[1,4];
  end;//ts
  end;
procedure TForm1.Button2Click(Sender: TObject);
var
vvv,kkk,sss:integer;
begin
 for vvv:=1 to populasyon do
        begin
            for kkk:= 1 to 2 do
            begin
             aday[vvv,kkk]:= bp[trunc(hkt[trunc(sira[vvv]),1]),kkk
            end;
        end;
end;
```

```
end.
```

Ek-2 NLGA2 Program Kodları

```
var
a,b,c,k8,p1,k10,p2,secim,k9,gecici,pm,pc,tbbhkt:double;
g,ts,gs,k,i1,kkk,k1,k2,k3,ps,populasyon,n,i,j,m,t:integer;
x:array[1..20] of double;
y:array[1..20] of double;
```

```
bp:array[1..3000,1..6]of double;  // Başlangıç populasyonu
hkt:array[1..3000,1..10]of double;
eniyi:array[1..1000,1..6] of double;
p:array[0..3000]of double;
sira:array[1..3000]of integer;
aday:array[1..3000,1..6]of double;
f:array[1..20,1..3000]of double;
e:array[1..20,1..3000]of double;  //hatalar için y-f
efkt:array[1..3,1..3000]of double;//  hata karelerin toplamı
implementation
{$R *.DFM}
procedure TForm1.Button1Click(Sender: TObject);
begin
randomize;
gs:=10;   // gözlem sayısı
populasyon:=1000; // populasyon büyüklüğü
ts:=100; // iterasyon sayısı
pm:=0.10; //mutasyon olasılığı
pc:=0.90; // Çaprazlama olasılığı
a:=1.07390;
b:=-0.046517;
c:=0.255007;
FOR i:=1 TO 20 do
begin
x[i]:=gr.NumberRC[i,1];
y[i]:=gr.NumberRC[i,2];
end;
for i:=1 to populasyon do
begin
```

```
bp[i,1]:=random*8;
bp[i,2]:=random*8;
bp[i,3]:=random*8;
bp[i,4]:=random*8;
bp[i,5]:=bp[i,1];
bp[i,6]:=random*8;
end;
 for g:=1 to ts do
 begin
      For k:=1 To populasyon do
      begin
           For j:=1 To gs do
           begin
           f[j,k]:=bp[k,1]*x[j]+bp[k,2]*exp(bp[k,3]*x[j]);
           f[j+10,k]:=bp[k,4]+bp[k,5]*x[j]+bp[k,6]*(x[j]*x[j]);
           end;
      end;
      For k:=1 To populasyon do
      begin
           For j:=1 To gs do
           begin
           e[j,k]:=y[j]-f[j,k];
           e[j+10,k]:=y[j+10]-f[j+10,k];    // Hata matrisi
           end;
      end;
      for k:=1 to populasyon do
      begin
           efkt[1,k]:=0;
           efkt[2,k]:=0;
```

```
        efkt[3,k]:=0;
        for j:=1 to gs do
        begin
        efkt[1,k]:=efkt[1,k]+ sqr(e[j,k]);
        efkt[2,k]:=efkt[2,k]+sqr(e[j+10,k]);
        efkt[3,k]:=efkt[3,k]+e[j,k]*e[j+10,k];
        end;
end;
FOR m:=1 TO populasyon do
begin
        hkt[m,4]:=0;
        hkt[m,4]:=a*efkt[1,m]+2*b*efkt[3,m]+c*efkt[2,m];
end;
FOR i:=1 TO populasyon do
begin
        hkt[i,1]:=i;
        hkt[i,2]:=bp[i,1];
        hkt[i,3]:=bp[i,2];
        hkt[i,6]:=bp[i,3];
        hkt[i,7]:=bp[i,4];
        hkt[i,8]:=bp[i,1];
        hkt[i,9]:=bp[i,6];
end;
for i:=1 to populasyon do
for j:=1 to 10 do
 f1book1.numberrc[i,j]:=hkt[i,j];
 f1book1.Sort3(1,1,populasyon,10,true,4,1,1);
 for i:=1 to populasyon do
 for j:=1 to 10 do
```

```
 hkt[i,j]:=f1book1.numberrc[i,j];
FOR i:=1 TO populasyon do
begin
      hkt[i,5]:=populasyon+1-i;
end;
  eniyi[g,1]:=hkt[1,2];
  eniyi[g,2]:=hkt[1,3];
  eniyi[g,3]:=hkt[1,6];
  eniyi[g,4]:=hkt[1,7];
  eniyi[g,5]:=hkt[1,8];
  eniyi[g,6]:=hkt[1,9];
  for i:=1 to populasyon do
  begin
  p[i]:=(2*hkt[i,5])/(populasyon*(populasyon+1));
  end;
  kkk:=populasyon-500;
  for k:=kkk to populasyon do
  begin
  bp[k,1]:=random*10;
  bp[k,2]:=random*10;
  bp[k,3]:=random*10;
  bp[k,4]:=random*10;
  bp[k,5]:=bp[k,1];
  bp[k,6]:=random*10;
  end;
p[0]:=0;
FOR i:=1 TO populasyon do
begin
      secim:=random;
```

```
            P1:= 0; P2:= 0;
            FOR j:= 1 to populasyon do
            begin
                  P1:= P1+ p[j-1];
                  P2:= P2+ p[j];
                  if (secim >P1) and (secim<P2) then sira[i]:=j;
            end;
      end;
      for i:=1 to populasyon do
      begin
            for j:= 1 to 6 do
            begin
             aday[i,j]:= bp[trunc(hkt[trunc(sira[i]),1]),j
            end;
      end;
//çaprazlama
      i:=0;
      while (i<=populasyon)do
      begin
       i:=i+1;
       k8:=random;
          IF k8< pc THEN
          begin
            k:=1+random(6);
            k9:=random;
            for t:=k to 6 do
            begin
            gecici:=aday[i,t];
            aday[i,t]:=k9*aday[i+1,t]+(1-k9)*aday[i,t];
```

```
            aday[i+1,t]:=(1-k9)*gecici+k9*aday[i+1,t];
            end;
            i:=i+1;
          end;//if
      end;
//mutasyon
      FOR i:= 1 TO populasyon do
      begin
        k10:=random(populasyon);
          IF k8 < pm THEN
          begin
            k9:=random;
             for k:=1 to 6 do
             begin
             gecici:=aday[i,k];
             aday[i,k]:=k9*aday[i+1,k]+(1-k9)*aday[i,k];
             aday[i+1,k]:=k9*gecici+(1-k9)*aday[i+1,k];
             end;
          end;
      end;
      FOR i:=1 TO populasyon do
      begin
          FOR j:= 1 TO 6 do
          begin
          bp[i,j]:=aday[i,j];
          end;
      end;
       for i:=1 to 6
         begin
```

```
            bp[1,i]:=eniyi[g,i];
            end;
        gr.NumberRC[g,3]:=hkt[1,4];
        gr.NumberRC[g,4]:=eniyi[g,1];
        gr.NumberRC[g,5]:=eniyi[g,2];
        gr.NumberRC[g,6]:=eniyi[g,3];
        gr.NumberRC[g,7]:=eniyi[g,4];
        gr.NumberRC[g,8]:=eniyi[g,5];
        gr.NumberRC[g,9]:=eniyi[g,6];
  end;//ts
  end;
end.
```