

**ANKARA ÜNİVERSİTESİ  
FEN BİLİMLERİ ENSTİTÜSÜ**

**YÜKSEK LİSANS TEZİ**

**DOĞRUSAL OLMAYAN REGRESYON MODEL PARAMETRELERİNİN  
NOKTA VE ARALIK TAHMİNİ İÇİN BİR YAKLAŞIM**

**Fikret AKGÜN**

**İSTATİSTİK ANABİLİM DALI**

**ANKARA  
2018**

**Her hakkı saklıdır**

## TEZ ONAYI

Fikret AKGÜN tarafından hazırlanan “Doğrusal Olmayan Regresyon Model Parametrelerinin Nokta ve Aralık Tahmini İçin Bir Yaklaşım” adlı tez çalışması 06 / 04 / 2018 tarihinde aşağıdaki jüri tarafından oybirliği ile Ankara Üniversitesi Fen Bilimleri Enstitüsü İstatistik Anabilim Dalı'nda **YÜKSEK LİSANS TEZİ** olarak kabul edilmiştir.

**Danışman** : Doç. Dr. Özlem TÜRKŞEN  
Ankara Üniversitesi İstatistik Anabilim Dalı

### Jüri Üyeleri:

**Başkan** : Prof. Dr. Müjgan TEZ  
Marmara Üniversitesi İstatistik Anabilim Dalı

**Üye** : Prof. Dr. Ayşen APAYDIN  
Ankara Üniversitesi Uygulamalı Bilimler Fakültesi  
Sigortacılık ve Aktüerya Bilimleri Anabilim Dalı

**Üye** : Doç. Dr. Özlem TÜRKŞEN  
Ankara Üniversitesi İstatistik Anabilim Dalı

**Yukarıdaki sonucu onaylarım.**

**Prof. Dr. Atila YETİŞEMİYEN**  
Enstitü Müdürü

## ETİK

Ankara Üniversitesi Fen Bilimleri Enstitüsü tez yazım kurallarına uygun olarak hazırladığım bu tez içindeki bütün bilgilerin doğru ve tam olduğunu, bilgilerin üretilmesi aşamasında bilimsel etiğe uygun davrandığımı, yararlandığım bütün kaynakları atıf yaparak belirttiğimi beyan ederim.

06 / 04 / 2018



Fikret AKGÜN

## ÖZET

Yüksek Lisans Tezi

### DOĞRUSAL OLMAYAN REGRESYON MODEL PARAMETRELERİNİN NOKTA VE ARALIK TAHMİNİ İÇİN BİR YAKLAŞIM

Fikret AKGÜN

Ankara Üniversitesi  
Fen Bilimleri Enstitüsü  
İstatistik Anabilim Dalı

Danışman: Doç. Dr. Özlem TÜRKŞEN

Doğrusal olmayan regresyon model parametrelerinin nokta tahminlerinin elde edilmesinde en çok izlenen yaklaşım, türe ve dayalı iteratif algoritmalar ile modele En Küçük Kareler (EKK) yaklaşımının uygulanmasıdır. Gauss-Newton, En Hızlı İniş ve Levenberg-Marquardt (L-M) algoritmaları literatürde tanımlı yaygın olarak kullanılan türe ve dayalı algoritmalarlardır. Fakat, bu algoritmalar kullanılarak yapılan aramalarda, başlangıç noktasının iyi tanımlanamadığı durumlarda yerel tuzaklara takılma, global çözüme yakınsayamama gibi sorunlar ortaya çıkabilmektedir. Bu çalışmada, doğrusal olmayan regresyon model parametrelerinin nokta tahminlerinin elde edilmesi amacıyla türe ve dayalı algoritmalara alternatif olarak, Nelder-Mead Simpleks (NMS) algoritması ve Genetik Algoritma (GA) gibi türevden bağımsız algoritmalar önerilmektedir. Çalışmada, NMS algoritması ve GA hakkında detaylı bilgiler verilmiştir. GA kullanılarak parametrelerin nokta tahminlerinin elde edilmesinde, GA ayarlanabilir parametrelerinin probleme uygun olarak belirlenmesi global sonuca ulaşmayı etkileyeceğinden, Taguchi deney tasarımı ile optimal GA ayarlanabilir parametre değerleri belirlenmiştir. Ayrıca çalışmada, NMS algoritması ve GA'nın avantajlı yönlerinin birlikte kullanılması ile oluşturulan ve GANMS olarak adlandırılan bir hibrit algoritma ile parametrelerin nokta tahminlerinin elde edilmesine de yer verilmiştir. Model parametrelerinin aralık tahminleri için Bootstrap yönteminden yararlanılmıştır. Çalışmada ifade edilen nokta ve aralık tahmini yaklaşımları, literatürde tanımlı bir veri setine uygulanarak, bir negatif-üstel regresyon modeline ait parametreler için nokta ve aralık tahminleri elde edilmiştir. Elde edilen sonuçlar literatürde bulunan sonuçlar ile karşılaştırıldığında, ayarlanabilir parametreleri uygun tanımlanan GA'nın ya da GANMS hibrit algoritmasının, parametrelerine göre doğrusal olmayan regresyon modellerinde bir optimizasyon aracı olarak kullanılabileceği sonucuna ulaşılmıştır.

**Nisan 2018, 72 sayfa**

**Anahtar Kelimeler:** Doğrusal Olmayan Regresyon, Türevden Bağımsız Optimizasyon Algoritmaları, Nelder-Mead Simpleks (NMS) Algoritması, Genetik Algoritma (GA), Genetik Algoritma ve Nelder-Mead Simpleks Algoritması Hibrit Algoritması (GANMS), Taguchi Deney Tasarımı, Bootstrap Örnekleme

## ABSTRACT

Master Thesis

### AN APPROXIMATION FOR POINT AND INTERVAL ESTIMATIONS OF NONLINEAR REGRESSION MODEL PARAMETERS

Fikret AKGÜN

Ankara University  
Graduate School of Nature Applied Science  
Department of Statistics

Supervisor: Assoc. Prof. Dr. Özlem TÜRKŞEN

A common used approach for obtaining point estimates of nonlinear regression model parameters is applying Least Squares (LS) approach by using derivative-based iterative algorithms. The Gauss-Newton, the Steepest Descent and Levenberg-Marquardt (L-M) algorithms are widely used derivative-based algorithms which are defined in the literature. However, if the initial values for these algorithms are not well-defined, some of the problems occur during the search, e.g. trapping to local solutions and non-convergence to global solution. In this study, derivative-free optimization algorithms, Nelder-Mead Simplex (NMS) algorithm and Genetic Algorithm (GA), are used as alternative approaches to derivative-based algorithms to obtain point estimations of nonlinear regression model parameters. The detailed informations are given about NMS algorithm and GA with defining the advantages and disadvantages. It is known that defining the proper values of GA tuning parameters effect global solution. Therefore, optimal GA tuning parameters are defined by using Taguchi experimental design. In addition, in the study, a hybrid algorithm, called GANMS and obtained by combining the advantageous properties of NMS and GA, is used for point estimation of parameters. The Bootstrap approach is used for interval estimation of model parameters. The point and interval estimation approaches, defined in the study, are applied on a data set from the literature and estimation of the negative-exponential regression model parameters are obtained. The obtained solutions are compared with the previous results. It is seen that the GA with proper tuning parameters or GANMS hybrid algorithm can be used as an optimization tool for point estimation of nonlinear regression models.

**April 2018, 72 pages**

**Key Words:** Nonlinear Regression, Derivative-Free Optimization Algorithms, Nelder-Mead Simplex (NMS) Algorithm, Genetic Algorithm (GA), Hybrid Algorithm of Genetic Algorithm with Nelder-Mead Simplex Algorithm (GANMS), Taguchi Experimental Design, Bootstrap Resampling

## TEŞEKKÜR

Tez çalışmalarım süresince, önerileri ile beni yönlendiren, karşılaştığım problemlerin çözümüne zaman ayırarak fikirlerini benimle paylaşan, yapıcı eleştirileri ile çalışmalarımda bana ivme katan, bilgisi ve anlayışıyla üzerimde büyük emeği olan danışman hocam Sayın Doç. Dr. Özlem TÜRKŞEN'e (Ankara Üniversitesi İstatistik Anabilim Dalı) sonsuz teşekkürlerimi sunarım.

İş yoğunluğuna rağmen vakit ayırıp güzel Türkçesiyle tezime son biçimini vermemde bana yardım eden edebiyat öğretmeni ve geleceğin ünlü yazarı Ayşenur KILIÇER'e teşekkürlerimi sunarım.

Bilgisayar programı kullanımı konusunda tecrübelerini benimle paylaşan Gözde ÖZÇIRPAN'a teşekkürlerimi sunarım.

Çok uzakta olsalar bile her zaman benimle olan aileme hayatım boyunca bana verdikleri koşulsuz destek ve sevgi için minnetimi ve teşekkürlerimi sunarım.

Fikret AKGÜN  
Ankara, Nisan 2018

# İÇİNDEKİLER

<b>TEZ ONAYI</b>	
<b>ETİK</b> .....	<b>i</b>
<b>ÖZET</b> .....	<b>ii</b>
<b>ABSTRACT</b> .....	<b>iii</b>
<b>TEŞEKKÜR</b> .....	<b>iv</b>
<b>KISALTMALAR DİZİNİ</b> .....	<b>vi</b>
<b>ŞEKİLLER DİZİNİ</b> .....	<b>vii</b>
<b>ÇİZELGELER DİZİNİ</b> .....	<b>viii</b>
<b>1. GİRİŞ VE ÖNCEKİ ÇALIŞMALAR</b> .....	<b>1</b>
<b>2. PARAMETRELERİNE GÖRE DOĞRUSAL OLMAYAN REGRESYON MODELİ</b> .....	<b>7</b>
<b>2.1 Model Parametrelerinin Nokta Tahmini</b> .....	<b>8</b>
<b>2.2 Model Parametrelerinin Aralık Tahmini</b> .....	<b>10</b>
<b>3. DOĞRUSAL OLMAYAN MODEL PARAMETRELERİNİN NOKTA TAHMİNLERİNİN ELDE EDİLMESİNDE KULLANILAN TÜREVDEN BAĞIMSIZ OPTİMİZASYON ALGORİTMALARI</b> .....	<b>13</b>
<b>3.1 Nelder-Mead Simpleks Algoritması</b> .....	<b>14</b>
<b>3.2 Genetik Algoritma</b> .....	<b>21</b>
<b>3.3 Nelder-Mead Simpleks Algoritması ile Genetik Algoritmanın Hibrit Edilmesi</b> .....	<b>35</b>
<b>3.4 Genetik Algoritma Ayarlanabilir Parametrelerinin Belirlenmesinde Taguchi Deney Tablolarının Kullanımı</b> .....	<b>40</b>
<b>4. DOĞRUSAL OLMAYAN MODEL PARAMETRELERİNİN ARALIK TAHMİNİNDE BOOTSTRAP YAKLAŞIMI</b> .....	<b>46</b>
<b>4.1 Bootstrap Güven Aralığı Yaklaşımları</b> .....	<b>48</b>
<b>5. MİKROORGANİZMALARIN BİYOKİMYASAL OKSİJEN İHTİYACINA İLİŞKİN DOĞRUSAL OLMAYAN REGRESYON MODELİNDE PARAMETRELERİN NOKTA VE ARALIK TAHMİNLERİNİN ELDE EDİLMESİ</b> .....	<b>53</b>
<b>6. TARTIŞMA VE SONUÇ</b> .....	<b>64</b>
<b>KAYNAKLAR</b> .....	<b>68</b>
<b>ÖZGEÇMİŞ</b> .....	<b>72</b>

## KISALTMALAR DİZİNİ

BC	Yanlılıđı Düzeltilmiş (Bias Corrected)
BCa	Yanlılıđı Düzeltilmiş ve Hızlandırılmış (Bias Corrected and Accelerated)
BOD	Biyokimyasal oksijen ihtiyacı (Biochemical Oxygen Demand)
EKK	En Küçük Kareler
GA	Genetik Algoritma
GANMS	Genetik Algoritma ve Nelder-Mead Simpleks Algoritması hibriti
L-M	Levenberg-Marquardt
NMS	Nelder-Mead Simpleks





## ŞEKİLLER DİZİNİ

Şekil 3.1 İki boyutlu Öklid uzayında yansıtma işlemiyle $R$ parametre tahmin vektörünün elde edilmesi.....	16
Şekil 3.2 İki boyutlu Öklid uzayında genişleme işlemiyle $E$ parametre tahmin vektörünün elde edilmesi.....	17
Şekil 3.3 İki boyutlu Öklid uzayında büzülme işlemiyle $P_1$ ve $P_2$ parametre tahmin vektörlerinin elde edilmesi .....	18
Şekil 3.4 İki boyutlu Öklid uzayında küçülme işlemiyle yeni parametre tahmin vektörlerinin elde edilmesi .....	19
Şekil 3.5 Nelder Mead Simpleks algoritmasının akış diyagramı .....	20
Şekil 3.6 Genetik Algoritma adımlarının akış diyagramı .....	25
Şekil 3.7 Gerçek değer kodlamalı tek nokta çaprazlama .....	31
Şekil 3.8 Gerçek değer kodlamalı iki nokta çaprazlama.....	31
Şekil 3.9 Scattered çaprazlama .....	32
Şekil 3.10 Aritmetik çaprazlama.....	32
Şekil 3.11 Heuristic çaprazlama.....	33
Şekil 3.12 Genetik Algoritma Nelder Mead Simpleks hibrit algoritması akış diyagramı .....	38
Şekil 5.1 Deneme-15’te bir yineleme için 500 kuşakta hesaplanan amaç fonksiyonu değerleri .....	58
Şekil 5.2 Deneme-7’de bir yineleme için 75 kuşakta hesaplanan amaç fonksiyonu değerleri .....	59
Şekil 5.3 Deneme-15 için uygulanan 200 yinelemeye göre GA ve GANMS ile bulunan hata kareler toplamalarının karşılaştırılması .....	62

## ÇİZELGELER DİZİNİ

Çizelge 3.1 Oluşturulma biçimlerine göre Taguchi tasarımları .....	42
Çizelge 3.2 $L_8(2^7)$ için üreteçlerin listesi (Kacker vd. 1991).....	43
Çizelge 3.3 $L_8(2^7)$ tablosu .....	44
Çizelge 3.4 $L_{25}(5^6)$ için üreteç tablosu (Kacker vd. 1991).....	44
Çizelge 3.5 $L_{25}(5^6)$ tablosu .....	45
Çizelge 5.1 Mikroorganizmaların zamana göre biyokimyasal oksijen ihtiyacı (Bates ve Watts 1988) .....	54
Çizelge 5.2 Bates ve Watts (1988) tarafından yapılan çalışmada bulunan modelin hata kareler toplamı ve parametre tahmin değerleri.....	54
Çizelge 5.3 Uygulamada kullanılacak GA ayarlanabilir parametreleri ve seviyeleri .....	55
Çizelge 5.4 GA ile 25 deneme için bulunan amaç fonksiyonu değerleri ve istatistikler .....	56
Çizelge 5.5 Ayarlanabilir GA parametrelerinin uygulama için optimal seviyeleri .....	57
Çizelge 5.6 Parametre başlangıç değerleri ve GANMS ile 25 deneme için bulunan amaç fonksiyonu değerlerinin istatistiği.....	60
Çizelge 5.7 GANMS ile 25 deneme için bulunan parametre tahminleri .....	61
Çizelge 5.8 BOD veri seti için Gauss-Newton, GA ve GANMS yöntemleriyle bulunan hata kareler toplamı ve parametre tahmin değerleri.....	62
Çizelge 5.9 BOD verisi için parametrelerin Bootstrap BCa güven aralıkları .....	63

## 1. GİRİŞ VE ÖNCEKİ ÇALIŞMALAR

Gerçek dünyada karşılaşılan bir problemin çözümünün elde edilmesindeki ilk aşama, probleme uygun bir matematiksel model tanımlanması olarak düşünülebilir. Fen, mühendislik ve sosyal alanlardaki bir çok problem, matematik terimleri ile ifade edildiği zaman, bu problemlerin matematiksel modelleri, bilinmeyen fonksiyonun bir veya birden daha yüksek dereceden türevlerini içeren bir denklemi sağlayan fonksiyonun bulunması problemine dönüşür. Bu yaklaşımla oluşturulmuş denklemlere diferansiyel denklemler denir. Bu amaçla, ilgilenilen sisteme ya da sürece ilişkin gerçekçi matematiksel modellerin oluşturulmasında diferansiyel denklemlerden yararlanır. Bir çok diferansiyel denklem sistemi doğrusal olmayan yapıdadır. Bu doğrusal olmama durumu modelin parametrelerine göre doğrusal olmamasından kaynaklanmaktadır. Matematiksel olarak bir modelin parametrelerine göre doğrusal olmaması, o modelin parametrelerine göre birinci türevleri (kısmi türevleri) alındığında elde edilen sonuçların model parametrelerine bağlı olması durumu olarak tanımlanabilir.

Rasgelelik ve rasgele hata kavramlarını göz ardı etmeden gözlenen veri kümesine en uygun doğrusal olmayan modelin seçilmesi, model parametrelerinin en düşük hata ile tahmin edilmesi istatistiksel analizin temel zorluklarından biridir. Mevcut veri kümesinde açıklanan ve açıklayıcı değişkenler arasındaki ilişki yapısının oluşturulmasında doğrusal olmayan regresyon analizinden yararlanır. Doğrusal regresyon modellemesine yönelik analizlerin kullanımı yaygın olmasına rağmen doğrusal olmayan regresyon modelleri ile ilgili çalışmalar daha yavaş ilerlemiştir. Çünkü, doğrusal olmayan regresyon modellerinde model parametrelerinin nokta ve aralık tahminlerine ilişkin istatistiksel sonuç çıkarımı yapmak, hesaplamalarda karşılaşılan güçlüklerin olması ve normal dağılım teorisinin tam olarak uygulanamaması gibi nedenlerden dolayı oldukça zor ve zaman alıcı olmuştur.

Marquardt (1963), Hartley ve Booker (1963), Gallant (1975), Seber ve Wild (1989) doğrusal olmayan regresyon modellerinde istatistiksel sonuç çıkarımı konusunda çalışmalar yapmışlardır.

Doğrusal olmayan regresyon modellerinde parametre tahmin edicilerinin analitik ifadesi elde edilemediğinden, Jennrich (1969), Malinvaud (1970), Box (1971), Gallant ve Fuller (1973) çalışmalarında yer verdikleri tahmin ve hipotez testi konuları için asimptotik ifadelerden ve doğrusal yaklaşımlardan faydalanmışlardır.

Bates ve Watts (1988), çalışmalarında doğrusal olmayan regresyon modelleri konusunda açıklayıcı örneklere yer vermişlerdir.

Şahinbaşoğlu (2005), doğrusal olmayan regresyonda artık analizi uygulamış ve sonuçları test edebilmek adına modelin eğrisellik ölçülerini araştırmıştır. Doğrusal olmayan regresyon modellerinde, model parametrelerinin nokta tahminlerini yapmak amacıyla izlenen ilk çözüm yaklaşımı dönüşüm uygulayarak modeli doğrusal hâle getirmek ve modeli doğrusal regresyon analizi ile çözmek olmuştur. Fakat bu durumda model ile birlikte hata terimleri de dönüşüme uğradığından, hata değişkenlerinin sabit varyanslı olması varsayımının bozulduğu görülmüştür.

Doğrusal olmayan modellerde parametrelerin nokta tahminine bir diğer yaklaşım türeve dayalı iteratif algoritmaları kullanarak hata kareler toplamının en küçüklenmesine dayalı olarak oluşturulan En Küçük Kareler (EKK) yaklaşımının uygulanmasıdır. Gauss-Newton, En Hızlı İniş ve Levenberg-Marquardt algoritmaları yaygın olarak kullanılan türeve dayalı optimizasyon algoritmalarıdır. Fakat, bu algoritmalarla en iyileme çalışmalarında problem için global değere yakınsayamama, yerel çözüm tuzaklarına takılma gibi sorunlar ortaya çıkabilmektedir. Ayrıca, bu yöntemler başlangıç çözüm değeri bilgisine ihtiyaç duyar. Başlangıç çözüm değerinin iyi belirlenememesi durumunda problem için optimal çözüme ulaşmak oldukça zaman alıcı ve bazen mümkün olamamaktadır. Bu nedenle, iyi tanımlanmamış başlangıç değeri ile arama yapmak hatalı sonuçlar elde edilmesine sebep olabilmektedir.

Türevden bağımsız optimizasyon algoritmaları doğrusal olmayan regresyon modellerinde parametrelerin nokta tahmininde türeve dayalı algoritmalara alternatif olarak kullanılmaktadır. Türevden bağımsız algoritmalar, belirlenen bir amaç fonksiyonu için genellikle stokastik aramaya dayalı çözüm üretir. Bu durum kesin sonuç

yerine yaklaşık sonuçlar sunmakla beraber bu yöntemlerin esnek yapıları, yerel çözüm tuzaklarına takılma riskini azaltmaktadır. Bu çalışmada türevden bağımsız arama algoritmaları olan Nelder-Mead Simpleks (NMS) algoritması ve Genetik Algoritma (GA) ile doğrusal olmayan regresyon modellerinde parametrelerin nokta tahminlerinin elde edilmesine yer verilmiştir.

Spendley vd. (1962), ilk defa simpleks esaslı bir arama algoritmasını kullanarak evrimsel programlama içerisinde yer alan ve üretim süreçlerinin iyileştirilmesi için kullanılan bir algoritma geliştirmiştir. Nelder ve Mead (1965), Spendley vd. (1962) tarafından ortaya atılan simpleks esaslı arama algoritmasını daha da geliştirmişler ve doğrusal olmayan modellerin optimizasyonu için kullanılan bir algoritma olarak uyarlamışlardır. NMS algoritması olarak adlandırılan bu algoritma, bölgesel olarak tanımlanmış bir alanda tek nokta aramaya dayalı, pratik ve etkili sonuçlar veren bir optimizasyon algoritmasıdır.

GA ise popülasyon tabanlı aramaya dayalı başlangıç çözüm değeri bilgisine ihtiyaç duymayan sezgisel bir optimizasyon algoritmasıdır. Holland (1975), evrim mekanizmasını bilgisayar sistemlerine aktararak GA'nın temellerini oluşturmuştur. De Jong (1975), GA'nın optimizasyon problemlerinin çözümünde kullanılabileceğini göstermiştir. Goldberg (1989), çalışmasında problemlere GA'nın uygulanmasını sağlayacak bilgisayar tekniklerini, matematik araçlarını ve araştırma sonuçlarını bir araya getirmeyi amaçlamıştır. Koza (1992, 1994) çalışmalarında, genetik programlamayı kullanıp zor ve karmaşık bilgisayar programlarının hiyerarşik bir yaklaşımla basit ve küçük alt gruplara indirgenerek üretilebileceğini iddia etmiştir. Yeniay (1999), çalışmasında girdi değişkenlerinin seviyelerinin belirlenmesinde Taguchi deney tasarımı yaklaşımıyla GA yaklaşımını karşılaştırmıştır. Altunkaynak ve Esin (2004), çalışmasında doğrusal olmayan regresyon modellerinde parametre tahmini için GA'yı kullanmışlardır.

NMS algoritması ve GA ilgilenilen probleme özgü model parametreleri dışında kendi içinde de parametreler bulundurur. Algoritmalara özgü bu parametrelere "ayarlanabilir (tunable/tuning) parametreler" adı verilir. Özellikle GA ile elde edilecek model

parametrelerinin nokta tahmin deęerleri, GA'nın ayarlanabilir parametrelerinden etkilenebilmektedir. Çünkü, GA'nın ayarlanabilir parametrelerinin uygun olabilecek farklı kombinasyonları mevcuttur. Bu kombinasyonlar içinden seçim yapmak ciddi emek ve zaman kayıplarına yol açacağından, GA için etkili ve güçlü bir parametre deęeri belirleme yöntemine ihtiyaç vardır. Taguchi tasarımı, en az sayıda deney ile en güçlü parametre kombinasyonuna ulaşmak için etkili bir deneysel tasarım yöntemidir. Bu nedenle, çalışmada Taguchi deney tasarımından yararlanarak, doğrusal olmayan model parametrelerinin nokta tahminlerinin minimum hata ile elde edilmesi için, GA ayarlanabilir parametre kombinasyonunun belirlenmesi konusunda çalışılmıştır. Türkşen ve Tez (2016) çalışmalarında, NMS ve GA'nın avantajlı özelliklerinin bir arada kullanılması ile oluşturulan hibrit bir algoritma önermişler ve bu algoritmayı GANMS olarak adlandırmışlardır. Bu çalışmada, GANMS hibrit algoritması, doğrusal olmayan regresyon modellerinin parametrelerinin nokta tahmininde kullanılmıştır ve GA ile elde edilen parametre tahmini sonuçlarından daha az hatalı nokta tahmin deęerlerine ulaşıldığı gözlenmiştir.

İstatistiksel sonuç çıkarımında model parametrelerinin nokta tahmini dışında bir dięer tahmin yaklaşımı model parametrelerinin aralık tahminidir. Parametrelerin nokta tahmini, parametrelere özgü tek bir aralık tanımlarken aralık tahmini belirli bir güven düzeyinde parametrelerin olabilecek deęerlerini içeren bir aralık tanımlar. Bu aralığa güven aralığı denir. Bir güven aralığı, parametreler için hesaplanan alt ve üst sınırların rasgele deęişken olduğu bir rasgele aralıktır.

Klasik güven aralığı hesaplamalarında kitle dağılımının normal dağılıma uyduğu varsayımı yapılır. Ancak, uygulamada karşılaşılan bir kitlenin normal dağılıma uygunluk göstereceğinin bir garantisi yoktur. Normal dağılıma uymayan kitle dağılımlarında merkezi limit teoreminden faydalanılabilir. Ancak örneklem hacminin küçük olduğu çalışmalarda kitle dağılımı hakkında yeterli bilgiye sahip olunamaması, kitle dağılımının çarpık olması gibi durumlardan dolayı hesaplanan güven aralıkları yanıltıcı olabilmektedir.

Örneklem hacminin küçük olduğu ve dağılım bilgisinin sınırlı olduğu problemlerde kitle parametresi için güven aralığı hesaplamasında Bootstrap yaklaşımı, klasik yaklaşıma alternatif olarak kullanılmaktadır. Efron (1979) tarafından yapılan çalışmada ilk defa Bootstrap yönteminin genel yapısından bahsedilmiş ve yöntemin temelleri oluşturulmuştur. Freedman (1981), Efron (1979) tarafından önerilen Bootstrap yönteminin regresyon modellerine uygulanmasına yönelik asimptotik teoriler geliştirmiştir. Efron ve Tibrishani (1986), DiCiccio ve Tibrishani (1987), DiCiccio ve Efron (1996) çalışmalarında, standart hata ve güven aralıkları hesaplamalarında Bootstrap yaklaşımını kullanmışlardır.

Bootstrap yaklaşımı dağılım varsayımına ihtiyaç duymaması, ortalama dışında medyan gibi ölçü birimlerinin de güven aralıklarının bulunmasında etkili sonuçlar vermesi gibi sebeplerle aralık tahmininde tercih edilmektedir. Bu amaçla literatürde tanımlanmış çeşitli Bootstrap yöntemleri mevcuttur. Bunlardan bazıları, yüzdeler yöntemi (percentile method), yanlılığı düzeltilmiş Bootstrap yöntemi (bias corrected method - BC method), yanlılığı düzeltilmiş ve hızlandırılmış Bootstrap yöntemi (bias corrected and accelerated method – BCa method) olarak tanımlanabilir. Bu çalışmada, model parametrelerinin aralık tahminlerinin elde edilmesinde BCa yöntemi kullanılmıştır.

Çalışmanın ikinci bölümünde, doğrusal olmayan regresyon modelinin genel yapısı tanımlanarak, model parametrelerinin nokta ve aralık tahminlerinin hesaplanmasında kullanılan yöntemler hakkında bilgiler sunulmuştur.

Çalışmanın üçüncü bölümünde, doğrusal olmayan regresyon modellerinin nokta tahminlerinin bulunmasında kullanılan türevden bağımsız optimizasyon algoritmaları NMS ve GA hakkında detaylı bilgiler verilmiştir. Bu iki algoritmanın adımları ile birlikte, NMS algoritması ve GA'nın hibriti ile elde edilen GANMS algoritmasının adımları tanımlanmıştır. GANMS'nin avantajlarından bahsedilmiştir. Ayrıca, GA'nın kendi ayarlanabilir parametrelerinin düzeylerinin ve seviyelerinin belirlenmesinde kullanılan Taguchi deney tasarım yöntemine yer verilmiştir.

Çalışmanın dördüncü bölümünde, Bootstrap yöntemi ile ilgili bilgilere yer verilerek Bootstrap yöntemi ile doğrusal olmayan model parametrelerinin güven aralıklarının belirlenmesine ilişkin kullanılacak BCa yaklaşımdan söz edilmiştir.

Çalışmanın beşinci bölümünde, literatürde tanımlı bir doğrusal olmayan regresyon modeli incelenmiştir. Uygulamaya ait veri seti Bates ve Watts (1988) tarafından yapılan çalışmadan alınmış olup, "mikroorganizmaların biyokimyasal oksijen ihtiyacı (Biochemical Oxygen Demand-BOD)" olarak tanımlanmıştır. BOD veri seti, negatif-üstel doğrusal olmayan regresyon yapısına göre modellenmiştir. Uygulamada kullanılan doğrusal olmayan modelin parametrelerinin nokta tahmini GA ve GANMS ile elde edilerek, türeve dayalı algoritmalar ile önceden bulunan sonuçlarla karşılaştırılması yapılmıştır. Model parametrelerinin aralık tahminleri ise Bootstrap BCa güven aralığı yöntemi ile hesaplanmıştır.

Çalışmanın altıncı bölümünde, yapılan analizlerin değerlendirilmesi yapılmış ve sonraki çalışmalar için öneriler sunulmuştur.



## 2. PARAMETRELERİNE GÖRE DOĞRUSAL OLMAYAN REGRESYON MODELİ

Parametrelerine göre doğrusal olmayan regresyon modelinde değişkenler arasındaki fonksiyonel ilişki, parametrelerin en az birinin doğrusal olmayan fonksiyonu biçiminde olabilir.  $n$  gözlemlili doğrusal olmayan regresyon modellerinin genel yapısı

$$Y_i = f(X_i; \theta) + \varepsilon_i, \quad i = 1, 2, \dots, n \quad (2.1)$$

biçiminde ifade edilebilir. Burada,  $\theta = (\theta_1, \theta_2, \dots, \theta_p)'$ ,  $\theta \in \Theta \subset R^p$  bilinmeyen parametre vektörü,  $Y_i$ ,  $i = 1, 2, \dots, n$  açıklanan değişken vektörü,  $X_i = (X_{i1}, X_{i2}, \dots, X_{ik})'$   $k$  adet girdi değişkenine sahip bağımsız değişken vektörü ve  $\varepsilon = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)$  rasgele hata vektörü olarak tanımlıdır.

Eşitlik (2.1)'de  $f$  fonksiyonunun biçimsel olarak bilindiği, sürekli olduğu, parametre vektörüne göre en az iki defa türevinin alınabildiği varsayılır. Ayrıca, modelin rasgele hata değişkenlerinin

$$\left. \begin{array}{l} E(\varepsilon_i) = 0 \\ Cov(\varepsilon_i) = \sigma^2 I \\ \varepsilon_i \sim (0, \sigma^2) \end{array} \right\} \quad i = 1, 2, \dots, n$$

varsayımlarını sağladığı kabul edilir.

Eşitlik (2.1) ile tanımlı açıklayıcı değişkenlerin gözlem değerleri bilindiğinde model parametrelerine bağlı olarak

$$Y = f(X, \theta) + \varepsilon \quad (2.2)$$

biçiminde vektörel formda yazılabilir. Burada,  $\mathbf{Y} = [Y_1 Y_2 \dots Y_n]'$ ,  $\boldsymbol{\varepsilon} = [\varepsilon_1 \varepsilon_2 \dots \varepsilon_n]'$ ,  $\boldsymbol{\theta} = [\theta_1 \theta_2 \dots \theta_p]'$  ve  $f(\boldsymbol{\theta}) = [f(x_1, \boldsymbol{\theta}) f(x_2, \boldsymbol{\theta}) \dots f(x_n, \boldsymbol{\theta})]'$  dir. Eşitlik (2.2)'de,  $\mathbf{X}$  ile tanımlı açıklayıcı değişken matrisine ilişkin gözlem değerleri genellikle çalışmalarda bilinen sabit değerler olarak kabul edilir. Bu nedenle (2.2) ile tanımlı eşitlik  $Y = f(\boldsymbol{\theta}) + \varepsilon$  biçiminde de yazılabilir.

## 2.1 Model Parametrelerinin Nokta Tahmini

Doğrusal olmayan modellerde parametre tahmini, doğrusal modellere göre daha zordur ve hesaplamaları daha uğraştırıcıdır. Doğrusal olmayan modellerde parametre tahmini yapılırken izlenen ilk çözüm yaklaşımının dönüşüm yapılarak doğrusal hâle getirmek olduğu düşünülmüştür. Fakat doğrusal olmayan bir fonksiyonun doğrusal hâle getirilmesi, dönüşüm ile elde edilen doğrusal modelin uygun olduğu anlamına gelmez. Çünkü doğrusallaştırma için uygulanan dönüşüm, modeldeki hata terimlerinin de dönüşümüne neden olur. Bu durumda, hata terimlerinin değerleri yanıltıcı olabilir. Ayrıca değişkenler arasındaki ilişkilerin fiziksel anlamda doğrusal olmaması nedeniyle doğrusallaştırma hâlinde parametrelerin fiziksel gerçekliğine uygun olarak yorumlanması mümkün olmayabilir.

Eşitlik (2.1) ile tanımlı doğrusal olmayan regresyon modelindeki  $f$  fonksiyonunun oluşturduğu yüzey doğrusal değildir ve sonlu büyüklüktedir.  $f$  fonksiyonunun biçimsel olarak bilinmesi varsayımına rağmen, parametre vektörünü tahmin etmek zordur. Bilinmeyen model parametrelerini tahmin etmek için bir amaç fonksiyonu belirlenerek, bilinmeyen katsayıların optimal değerleri birer tahmin olarak hesaplanır. Bu amaçla, doğrusal olmayan model için tanımlanan varsayımları dikkate alınan genel çözüm yaklaşımı

$$\phi(\boldsymbol{\theta}) = \sum_{i=1}^n [Y_i - f(\mathbf{X}_i, \boldsymbol{\theta})]^2 \quad (2.3)$$

biçiminde tanımlı hata kareler toplamını amaç fonksiyonu olarak ele alıp, bu amaç fonksiyonunu minimum yapan parametre vektörünü elde etmek biçiminde tanımlanabilir. Eşitlik (2.3)'te tanımlı fonksiyonunun minimize edilme işlemi, En Küçük Kareler (EKK) yöntemi olarak bilinir. Bu yöntem kullanılarak elde edilen parametre değerleri, parametrelerin EKK tahmin değerleri olarak adlandırılır. EKK yöntemi, eşitlik (2.2) ile tanımlı  $\varepsilon$  rasgele hata vektörünün dağılım bilgisi mevcut olmadığında tercih edilen bir tahmin edici bulma yöntemidir. Eşitlik (2.3) ile tanımlı  $\phi$  amaç fonksiyonu kısıtsız bir optimizasyon problemi olarak düşünüldüğünde, ilgilenilen  $f$  model fonksiyonunun sürekli ve iki kez türevlenebilir olması varsayımı altında,  $\phi$  fonksiyonun minimizasyonunda diskriminant yönteminden yararlanılabilir (Apaydın 2005).

Parametrelerin EKK tahminleri için, Eşitlik (2.3)'te tanımlı  $\phi$  amaç fonksiyonunun minimum noktasında parametre vektöründeki her bir parametreye göre birinci dereceden kısmi türevleri sıfır, ikinci dereceli kısmi türevlerinden oluşan Hessian matrisi pozitif tanımlı olmalıdır. Bu durumda EKK tahminlerini bulmak için,  $\phi$  amaç fonksiyonu sürekli ve türevlenebilir olduğundan,  $\phi$  fonksiyonunun model parametrelerine göre türevi alınarak normal denklemler elde edilir. Burada normal denklem sistemi

$$\frac{\partial \phi}{\partial \theta_j} = \sum_{i=1}^n [Y_i - f(\mathbf{X}, \boldsymbol{\theta})] \left( \frac{\partial f(\mathbf{X}, \boldsymbol{\theta})}{\partial \theta_j} \right) = 0, \quad j = 1, 2, \dots, p \quad (2.4)$$

biçiminde tanımlı türev fonksiyonları ile oluşturulur. Eşitlik (2.4)'te tanımlı kısmi türev fonksiyonları, bilinmeyen parametrelerin fonksiyonları olacaktır.  $f$  fonksiyonunun da parametrelerine göre doğrusal olmayan bir fonksiyon olduğu göz önünde bulundurulduğunda, normal denklem sisteminin analitik çözümlerinin elde edilmesinin zorlayıcı olacağı açıktır. Bu gibi zorlukların üstesinden gelebilmek için, EKK tahminlerini belirlemede Gauss-Newton algoritması, En Hızlı İniş algoritması, Levenberg-Marquardt (L-M) algoritması gibi türeve dayalı yinelemeli algoritmalar kullanılır. Şahinbaşoğlu (2005), Ünlü (2006) ve Serin (2010) bu algoritmalar hakkında çalışmalarında detaylı bilgiler vermişlerdir.

Bu yöntemler arasında en yaygın kullanılanı Gauss-Newton algoritmasıdır. Bu algoritma, doğrusallaştırma yaklaşımı olarak da bilinir. Doğrusal olmayan regresyon modeli Taylor serisi kullanılarak doğrusal modele yaklaştırılır ve EKK yöntemi ile parametre tahmini yapılır. Gauss-Newton algoritması ile optimal çözüm değerine yakınsama yavaş olabilir, geniş bir aralıkta olabilir veya hiç olmayabilir. İşlemlerin yakınsama hızını arttırabilmek ve yineleme işlem sayısını azaltabilmek için minimum hata ile parametre tahmin değerini belirlemede iyi bir başlangıç çözümünden başlamak önemlidir. En Hızlı İniş algoritmasında, Gauss-Newton algoritmasındaki yön vektörlerinin yerine gradyant vektörlerinin ters yönünde olan yön vektörü kullanılır.

Gauss-Newton algoritmasında, parametrelerin başlangıç tahmin değerleri, en iyi çözüm değerlerinden çok uzak seçilmişse, algoritma yavaş ilerler ve en iyi parametre tahmini değerine yaklaştıkça arama hızlanır. En Hızlı İniş algoritmasında ise parametrelerin başlangıç tahmin değerleri optimalden uzaktayken yakınsama hızlı olur, optimale yaklaştıkça yakınsama yavaşlar. L-M algoritması, Gauss-Newton ve En Hızlı İniş algoritmalarının avantajlarını bir araya getiren bir algoritmadır. L-M algoritmasında başlangıçta En Hızlı İniş algoritması kullanılmaktayken optimal değere yaklaşıldığında Gauss-Newton algoritması kullanılır. Böylece daha az sayıda yineleme ile en iyi çözüme yakınsama sağlatılmaya çalışılır. Ancak uygun olmayan başlangıç değerleri, adım aralıklarının gereğinden büyük veya küçük olması durumlarında parametrelerin minimum hataya sahip tahminleri bulunamayabilir veya bir yakınsama gerçekleşmeyebilir. Bu gibi dezavantajlarından dolayı türeve dayalı oluşturulmuş yinelemeli yöntemler yerine, türevden bağımsız fonksiyon değerleri ile arama yapan optimizasyon yöntemlerini kullanmak tercih edilir.

## **2.2 Model Parametrelerinin Aralık Tahmini**

Doğrusal olmayan regresyon modellerinin nokta tahminlerinin hesaplanmasında gerçek parametre değerlerinin tahmin ile aynı olması beklenemeyeceğinden parametreye en yakın tahminin bulunması amaçlanmaktadır. Ancak yapılan tahminin parametreye hangi olasılıkla ne kadar yakın olduğu bilgisi nokta tahmini ile açıklanamaz. İstatistiksel sonuç çıkarımında model parametrelerinin nokta tahmini dışında bir diğer tahmin

yaklaşımı model parametrelerinin aralık tahminidir. Mevcut veri setini kullanarak parametre değerini belli bir olasılıkla bir aralık içinde hesaplama işlemine aralık tahmini adı verilmektedir.

Aralık tahmininde, parametre için alt ve üst sınırlar belirlenmesi ve parametrenin ne kadar olasılıkla bulunan alt ve üst sınırlar içinde yer alacağını hesaplanması ile ilgilenilir. Burada hesaplanan alt ve üst sınırların birbirlerine yakın değerler çıkması istenilen durumdur. Sınır değerleri birbirine yaklaştıkça tahminin güvenilirliği artmaktadır. Bir parametrenin aralık tahmini, parametreyi tahmin etmek için kullanılan değerleri içeren bir aralıktır. Bir parametrenin aralık tahmininin güven düzeyi, seçilen aralığın parametreyi kapsayan aralıklardan biri olma olasılığıdır ve  $1-\alpha$  ile gösterilir. Burada,  $\alpha$  anlamlılık düzeyi adını alır. Tahminin güven düzeyini kullanarak bir parametre için belirlenen aralığa güven aralığı denir (Oral-Erbaş 2016). Örneğin, bir parametre için %95 güven düzeyinde belirlenen bir güven aralığı, ilgilenilen parametre için 100 defa güven aralığı hesaplaması yapıldığında bu aralıklardan 95 tanesinin parametreyi kapsaması demektir.

Klasik yaklaşım ile yapılan güven aralığı hesaplama işleminde, veri setinin normal dağılım gösteren bir kitleden rasgele olarak seçildiği varsayımı yapılmaktadır. Kitle, normal dağılımlı olmasa bile Merkezi Limit Teoremi (MLT) gereği, örnek sayısı arttıkça kitlenin dağılımı normal dağılıma yakınsamaktadır.  $\theta$ , nokta ve aralık tahmini elde edilmek istenilen parametre olmak üzere

$$\frac{\hat{\theta}-\theta}{S_{\hat{\theta}}} \sim N(0,1) \quad (2.5)$$

biçiminde standart normal dağılıma dönüşüm yapılır. Burada,  $S_{\hat{\theta}}$ ,  $\hat{\theta}$  tahmin edicisinin standart sapmasıdır. Standart normal dağılımın kümülatif fonksiyonunda güven sınırlarına karşılık gelen değerleri kullanılarak alt ve üst sınır belirlenir. Buna göre Eşitlik (2.5)'deki yakınsama ile oluşturulan  $(1-\alpha)$  güven düzeyindeki güven aralığı

$$P(Z_{\alpha/2} < \frac{\hat{\theta} - \theta}{S_{\hat{\theta}}} < Z_{1-\alpha/2}) = 1 - \alpha \quad (2.6)$$

şeklindedir. Eşitlik (2.6)'daki  $Z_{\alpha/2}$  ve  $Z_{1-\alpha/2}$ , standart normal dağılım fonksiyonunun  $(1-\alpha)$  güven düzeyine karşılık gelen değerleridir. Burada  $Z_{\alpha/2} < 0$ ,  $Z_{1-\alpha/2} > 0$  'dır ve standart normal dağılım simetrik olduğundan bu iki değer mutlak değerleri birbirine eşittir. Bu durumda parametrenin  $(1-\alpha)$  güven düzeyinde güven aralığı

$$P(\hat{\theta} + Z_{\alpha/2} \times S_{\hat{\theta}} < \theta < \hat{\theta} + Z_{1-\alpha/2} \times S_{\hat{\theta}}) = 1 - \alpha \quad (2.7)$$

biçimindedir. Eşitlik (2.7)'ye göre parametrenin  $(1-\alpha)$  güven düzeyindeki güven sınırları  $[\alpha_{alt}; \alpha_{üst}] = [\hat{\theta} + Z_{\alpha/2} \times S_{\hat{\theta}}; \hat{\theta} + Z_{1-\alpha/2} \times S_{\hat{\theta}}]$  biçiminde tanımlanır.

Normal dağılıma uymayan kitle dağılımlarında MLT'den faydalanılabilir. Ancak, örneklem hacminin küçük olduğu çalışmalarda kitle dağılımı hakkında yeterli bilgiye sahip olunamaması, kitle dağılımının çarpık olması gibi durumlardan ötürü hesaplanan güven aralıkları yanıltıcı olabilmektedir.

Örneklem hacminin küçük olduğu ve dağılım bilgisinin sınırlı olduğu problemlerde kitle parametresi için güven aralığı hesaplamasında Bootstrap yaklaşımı, klasik yaklaşıma alternatif olarak kullanılmaktadır. Bootstrap yaklaşımı dağılım varsayımına ihtiyaç duymaması, ortalama dışında medyan gibi ölçü birimlerinin de güven aralıklarının bulunmasında etkili sonuçlar vermesi gibi sebeplerle parametrelerin aralık tahminlerinin elde edilmesinde tercih edilmektedir.

### **3. DOĞRUSAL OLMAYAN MODEL PARAMETRELERİNİN NOKTA TAHMİNLERİNİN ELDE EDİLMESİNDE KULLANILAN TÜREVDEN BAĞIMSIZ OPTİMİZASYON ALGORİTMALARI**

Türevden bağımsız optimizasyon algoritmaları türev bilgisinin elde edilemediği, elde edilse bile türev bilgisi kullanımının pratik olmadığı problemlerde, problemin çözümü için belirlenen kurallar çerçevesinde yinelemeli olarak ilerleyen, her yinelemede elde edilen sonuçlar kullanılarak optimal sonuca ulaşmaya çalışan algoritmalar (Rios ve Sahinidis 2013).

Türevden bağımsız optimizasyon algoritmaları deterministik veya stokastik aramaya dayalı olabilir. Stokastik aramaya dayalı algoritmalar, aynı problem için kullanılan aynı yöntemden farklı sonuçların elde edilebildiği algoritmalar. Bu aşamada, başlangıç çözümü değerlerinin ve algoritmayı ilerletecek adımsal kuralların seçimi önemlidir. Stokastik algoritmalar kesin sonuç yerine yaklaşık sonuçlar sunmakla beraber bu algoritmaların esnek yapıları nedeniyle yerel çözüm tuzaklarına yakalanma olasılıkları Gauss-Newton, L-M gibi türeve dayalı yinelemeli algoritmalara göre daha azdır (Maringer, 2005).

Türevden bağımsız stokastik optimizasyon algoritmalarından her biri belirlenen bir amaç fonksiyonunu optimize etmek için kendine özgü çözüm stratejilerini kullanan algoritmalar. Algoritmaların başarısı, kullanılan çözüm stratejileri ve ilgilenilen problemin niteliğine göre farklılık gösterir.

Bu çalışmada, tek nokta aramalara dayalı türevden bağımsız bir optimizasyon algoritması olan NMS algoritması, çok nokta aramaya dayalı popülasyon tabanlı global arama algoritması olan GA ve bu iki algoritmanın avantajlı yönlerinin hibriti ile oluşturulan GANMS algoritması ile doğrusal olmayan model parametrelerinin nokta tahminleri elde edilecektir.

### 3.1 Nelder-Mead Simpleks Algoritması

Nelder-Mead Simpleks (NMS) algoritması, doğrusal olmayan modellerin optimizasyonunda çabuk sonuç veren, karmaşık olmayan, etkili bir yöntemdir. Stokastik aramaya dayalı bir optimizasyon algoritması olup çok boyutlu doğrusal olmayan problemlerin çözümüne türev bilgisine ihtiyaç duymadan, uygulanması kolay bir yaklaşım getirdiğinden NMS en çok kullanılan optimizasyon yöntemlerinden biridir. İstatistik, fizik, mühendislik ve tıp bilimlerinde yaygın olarak kullanılmaktadır.

NMS algoritması, Spendley vd. (1962) tarafından tanıtılan simpleks yönteminin genişletilmiş bir versiyonudur. Spendley vd. (1962) tarafından ortaya atılan simpleks esaslı arama algoritmasında,  $q$  boyutlu Öklid uzayında arama yaparken  $q+1$  noktaya ihtiyaç duyulur. Bu noktalara karşılık gelen amaç fonksiyonu değerlerine göre noktalar sıralanır. Yansıtma ve küçülme adı verilen operasyonlar kullanılarak adım adım optimum noktaya ulaşılır. Nelder ve Mead (1965) simpleks algoritmasında yansıtma ve küçülme operasyonlarına genişleme ve büzülme operasyonları da ekleyerek doğrusal olmayan modellerin optimizasyonu için NMS algoritmasını geliştirmişlerdir.

Doğrusal olmayan regresyon model parametrelerinin nokta tahminlerinin NMS algoritması ile elde edilmesi amacıyla öncelikli olarak amaç fonksiyonunun belirlenmesi ve seçilen  $q+1$  tane parametre tahmin vektörünün amaç fonksiyon değerlerinin bulunmasıyla işleme başlanır. Buna göre,  $(\hat{\theta}_1 \phi(\hat{\theta}_1)), (\hat{\theta}_2 \phi(\hat{\theta}_2)), \dots, (\hat{\theta}_{q+1} \phi(\hat{\theta}_{q+1}))$  hesaplanır.

Parametre tahminleri,  $\phi$  amaç fonksiyonundaki değerlerine göre sıralanır.

$$\phi(\hat{\theta}_1^*) < \phi(\hat{\theta}_2^*) < \dots < \phi(\hat{\theta}_{q+1}^*)$$

Bu sıralamada, dikkat edilmesi gereken bazı önemli parametre tahmini vektörü adlandırmaları mevcuttur. Bunlar;



$L$ : En düşük amaç fonksiyonu değerine sahip parametre tahmin vektörü,  
 $S$ : En düşük ikinci amaç fonksiyonu değerine sahip parametre tahmin vektörü,  
 $H$ : En yüksek amaç fonksiyonu değerine sahip parametre tahmin vektörü,  
 $N$ : En yüksek ikinci amaç fonksiyonu değerine sahip parametre tahmin vektörü,  
olarak adlandırılır.

Sıralama işleminden sonra NMS algoritmasında yansıtma, genişleme, büzülme ve küçülme olmak üzere dört operasyon uygulanır. Bu operasyonların her birinde modelin daha iyi çalışmasını sağlayacak NMS algoritmasına özgü ayarlanabilir parametreler kullanılır. Bu ayarlanabilir algoritma parametreleri

$\alpha$  : Yansıtma parametresi ( $\alpha > 0$ )  
 $\beta$  : Genişleme parametresi ( $\beta > 1$ )  
 $\gamma$  : Büzülme parametresi ( $0 < \gamma < 1$ )  
 $\sigma$  : Küçülme parametresi ( $0 < \sigma < 1$ )

biçiminde tanımlanmıştır. NMS algoritmasının ayarlanabilir parametrelerinin her biri için tanımlı aralıklarda farklı reel sayılar belirlenebilir. Çalışmalarda genellikle bu parametreler  $\alpha = 1$ ,  $\beta = 2$ ,  $\gamma = 1/2$ ,  $\sigma = 1/2$  olacak biçimde seçilir (Gurson 2000).

NMS algoritması, bu operasyonlar ile adım adım ilerleyerek iyi bir başlangıç çözüm değeri ile amaç fonksiyonu değerini en küçük yapacak parametre tahmin değerine yakınsar.

NMS algoritmasında tanımlı operasyonlar aşağıdaki gibi tanımlanabilir:

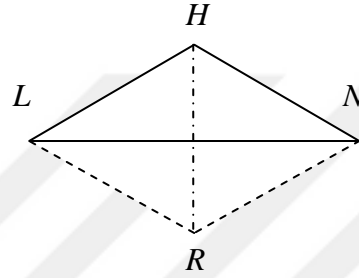
**i) Yansıtma operasyonu:** Öncelikle  $H$  parametre tahmin vektörü haricindeki bütün parametre tahmin vektörlerinin ağırlık merkezi bulunur.

$$\bar{\theta} = \frac{\sum_{i=1}^q \hat{\theta}_i}{q} \quad (3.1)$$

$H$  ile  $\bar{\hat{\theta}}$  arasındaki mesafe hesaplanır ve Öklid uzayında  $\bar{\hat{\theta}}$  'dan bu mesafe kadar daha ileriye gidilir. Yansıtma operasyonu ile elde edilen  $R$  parametre tahmin vektörü  $R = \bar{\hat{\theta}} + \alpha(\bar{\hat{\theta}} - H)$ ,  $\alpha > 0$  biçiminde tanımlıdır. Burada,  $\alpha = 1$  durumunda

$$R = 2\bar{\hat{\theta}} - H \quad (3.2)$$

olacaktır.



Şekil 3.1 İki boyutlu Öklid uzayında yansıtma işlemiyle  $R$  parametre tahmin vektörünün elde edilmesi

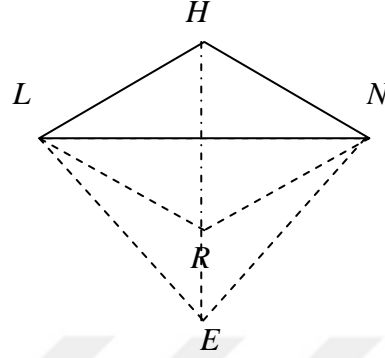
Şekil 3.1'de yansıtma operasyonu ile elde edilen  $R$  parametre tahmin vektörü gösterilmiştir. Optimizasyon işleminde yansıtma ile elde edilen  $R$  parametre tahmin vektörü, her adımda diğer operasyon işlemleri ile elde edilen parametre tahmini vektörleri ile karşılaştırılarak referans tahmin değeri olarak işlem görür.

**ii) Genişleme operasyonu:**  $R$  parametre tahmininin  $\phi$  fonksiyonu değeri  $\phi(R)$ ,  $H$  parametre tahmininin  $\phi$  fonksiyonu değeri  $\phi(H)$  ile karşılaştırılır. Eğer,  $\phi(R) < \phi(H)$  ise yansıtma işleminde  $\bar{\hat{\theta}}$  'den  $R$ 'ye gidilen mesafe kadar daha  $R$ 'den ileriye gidilerek  $E$  parametre tahmin değerine ulaşılır. Bu durumda  $E$  parametre tahmininin  $\phi$  fonksiyonu değeri  $\phi(E)$ ,  $R$  parametre tahmin değerinin  $\phi$  fonksiyonu değeri  $\phi(R)$  ile karşılaştırılır. Eğer,  $\phi(E) < \phi(R)$  ise bir sonraki adımda algoritmaya  $H$  yerine  $E$  dâhil edilir.

Genişleme operasyonu ile elde edilen  $E$  parametre tahmin vektörü  $E = \bar{\hat{\theta}} + \beta(R - \bar{\hat{\theta}})$ ,  $\beta > 1$  biçiminde tanımlıdır.  $\beta = 2$  için,

$$E = 2R - \bar{\underline{\theta}} \quad (3.3)$$

olur.



Şekil 3.2 İki boyutlu Öklid uzayında genişleme işlemiyle E parametre tahmin vektörünün elde edilmesi

Şekil 3.2’de genişleme operasyonu ile elde edilen R parametre vektörü tahmini gösterilmiştir.

**iii) Büzülme operasyonu:** Büzülme operasyonu sonucunda iki farklı parametre tahmin vektörü elde edilir. Eğer  $\phi(R) > \phi(H)$  ise,  $\bar{\underline{\theta}}$  ve  $H$  parametre tahmin vektörlerinin ortasındaki

$$P_1 = \bar{\underline{\theta}} + \gamma(H - \bar{\underline{\theta}}), \quad 0 < \gamma < 1 \quad (3.4)$$

eşitliği ile tanımlı  $P_1$  parametre tahmin vektörü bulunur. Eşitlik (3.4)’te  $\gamma = 1/2$  için

$$P_1 = (H + \bar{\underline{\theta}}) / 2 \quad (3.5)$$

olur. Eğer,  $\phi(P_1) < \phi(H)$  ise, sonraki adımda  $P_1$  parametre tahmin vektörü  $H$  parametre tahmin vektörünün yerine geçer.

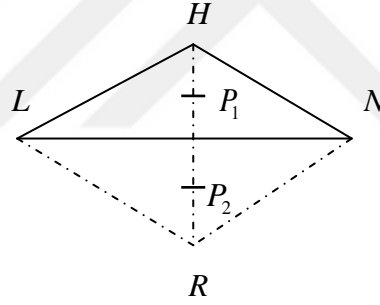
Eğer,  $\phi(N) < \phi(R) < \phi(H)$  ise ikinci tip büzülme gerçekleşir.

$$P_2 = \bar{\theta} + \gamma(R - \bar{\theta}), \quad 0 < \gamma < 1 \quad (3.6)$$

eşitliği ile tanımlı  $P_2$  parametre tahmin vektörü bulunur. Eşitlik (3.6)'da  $\gamma = 1/2$  için

$$P_2 = (R + \bar{\theta})/2 \quad (3.7)$$

olur. Eğer,  $\phi(P_2) < \phi(R)$  ise sonraki adımda  $P_2$  parametre tahmin vektörü,  $H$  parametre tahmin vektörünün yerine geçer. Şekil 3.3'te, büzülme operasyonu sonucunda bulunan  $P_1$  ve  $P_2$  parametre tahmin vektörlerinin konumları gösterilmiştir.



Şekil 3.3 İki boyutlu Öklid uzayında büzülme işlemiyle  $P_1$  ve  $P_2$  parametre tahmin vektörlerinin elde edilmesi

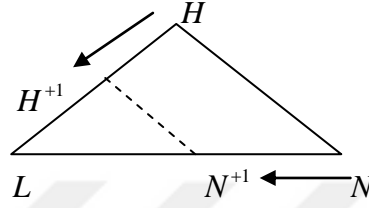
**iv) Küçülme Operasyonu:** Eğer  $\phi(P_1) > \phi(H)$  ise küçülme işlemi gerçekleşir. Bu durumda algorithmada yer alan bütün parametre tahmin değerleri L parametre tahmin vektörüne doğru çekilir. Buna göre, parametre tahmin değerleri

$$\bar{\theta}_{j+1} = \bar{\theta}_j + \sigma(L - \bar{\theta}_j), \quad 0 < \sigma < 1, \quad j = 2, 3, \dots, q+1 \quad (3.8)$$

$\sigma=1/2$  için

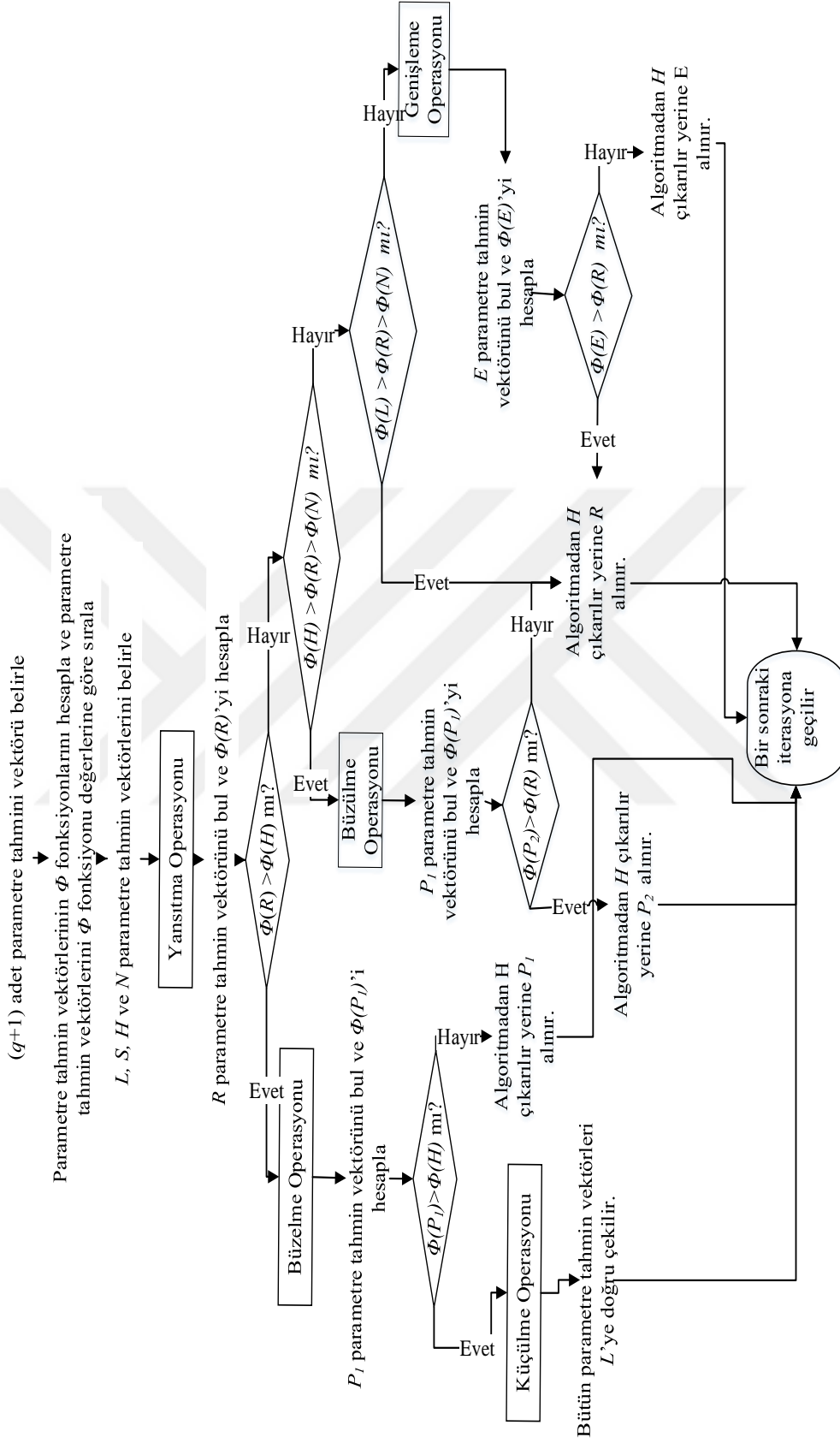
$$\bar{\theta}_{j+1} = (L + \bar{\theta}_j) / 2, \quad j = 2, 3, \dots, q+1 \quad (3.9)$$

olur. Şekil 3.4'te küçülme operasyonu ile elde edilen yeni parametre tahmin vektörleri gösterilmiştir. Burada  $H^{+1}$  ve  $N^{+1}$ ,  $H$  ve  $N$ 'nin bir sonraki iterasyonda alacağı değeri göstermektedir.



Şekil 3.4 İki boyutlu Öklid uzayında küçülme işlemiyle yeni parametre tahmin vektörlerinin elde edilmesi

NMS algoritmasında seçilen  $q+1$  adet parametre tahmin vektörüne dört NMS operasyonunun uygulanması ve bu uygulama sonucunda  $\phi$  fonksiyon değeri en büyük parametre tahmin vektörü olan  $H$ 'nin çıkarılıp yerine en uygun parametre tahmin vektörünün seçilmesi ile bir sonraki yinelemeye geçilir. Operasyonların bir veya birkaçı kullanılarak  $H$ 'nin yerine geçecek yeni parametre tahmin vektörü bulunmuş ise diğer operasyonları kullanmak gereksizdir. Her yineleme sonucunda algoritmada bulunan  $q+1$  parametre tahmin vektörü  $\phi$  fonksiyonu değerine göre yeniden sıralanır ve  $L$ ,  $H$ ,  $S$  ve  $N$  parametre tahmin vektörleri yeniden belirlenir. Şekil 3.5'teki akış diyagramında gösterilen işlemler her yinelemede uygulanır ve yeni parametre tahmin vektörleri belirlenerek  $H$  yerine bir sonraki yinelemede algoritmaya dâhil edilir. Algoritma, amaç fonksiyon değerlerinin birbirine çok yaklaşması veya belirli bir yineleme sayısına ulaşılması durumlarından birinin gerçekleşmesi yoluyla sonlandırılır. NMS algoritmasının bir yinelemesinde gerçekleşen işlemler şekil 3.5'te gösterilmektedir.



Şekil 3.5 Nelder Mead Simpleks algoritmasının akış diyagramı

### 3.2 Genetik Algoritma

Evrım, canlı türlerinin içinde yaşadıkları çevre koşullarında avantaj sağlayabilmeleri için kalıtsal deęişikliklere uğrayarak nesiller ilerledikçe farklı özellikler kazanılması sürecidir. Bu teoriye göre bitkiler, hayvanlar ve dünyadaki dięer tüm canlıların kökeni kendilerinden önce yaşamış türlere dayanır. Önceki atalardan farklı olarak kazanılmış yeni özellikler, başarılı nesillerde meydana gelmiş genetik deęişikliklerin bir sonucudur. İçinde yaşadığı ortama uyum sağlayarak hayatta kalmış ve üreme şansına sahip olmuş bireyler nesillerini devam ettirme fırsatı yakalarken, uyum sağlayamayıp üreme şansı elde edemeyen bireyler kuşaklar geçtikçe toplumdan elenirler. Doğal seçim sonucunda çevre koşullarına göre "en iyi" özelliklere sahip bireyler kuşaklar ilerledikçe topluluk içerisinde çoğunluğu oluştururlar (Ayala 2017). GA da evrim sürecini taklit ederek nesiller sonucunda "en iyi" çözümlerin topluluk içinde çoğunluğu oluşturması ve belli bir kuşak sayısı sonunda bu çözümler içinden de en iyisinin seçilmesi temeline dayanan bir optimizasyon algoritmasıdır.

GA, doğal seçim mekanizmasını esas alan ve evrim ilkelerinin bilgisayarda simülasyonunun yapılmasıyla sonuca ulaşan bir yöntemdir (Altunkaynak ve Esin 2004). Canlı türlerine ait özellikler kromozomlar içerisinde bulunan genlerde kodlanmış hâlde bulunur ve bu özellikler gen aktarımıyla gelecek kuşaklara geçer. GA'da da parametre değerlerinin olası çözümleri genler hâlinde kodlanır ve başarılı sonuç veren genler gelecek nesillere aktarılma fırsatı bulur. Algoritmanın amacı en iyi sonucu veren kromozomu oluşturmaktır. En iyi kromozoma ulaşma ise genlerin kromozom içindeki dizilişini deęiştirerek yani yeni nesiller yaratarak gerçekleştirilmektedir (Özçakar 1998). Nesiller ilerledikçe aktarılan ve deęişen gen özellikleri, kötü olanın elenmesi ve iyi olanın seçilmesi ile en iyi noktaya yaklaşmaktadır. Bütün bu işlemler bilgisayar simülasyonu yardımı ile gerçekleştirilmektedir.

1950'lerden itibaren birçok bilgisayar bilimcisi, evrimin mühendislik problemleri için bir optimizasyon aracı olarak kullanılabileceği fikriyle ilgilenmişlerdir. Bütün bu çalışmalarda ana fikir, genetik çeşitlilik ve doğal seleksiyondan esinlenerek elde edilen bazı operatörleri kullanarak belirli bir problemdeki aday çözümler popülasyonunu

geliştirmektir. Rechenberg (1965, 1973), gerçek değerli parametreleri optimize etmek amacıyla kullandığı bir yöntem olan "evrim stratejileri"ni ortaya koymuştur. Fogel vd. (1966) evrimsel programlamanın tanıtıldığı kitabı yayınlamışlardır.

Evrimin bilgisayar bilimlerine entegre edilmesinde 1950'li yıllardan itibaren pek çok bilim insanının katkısı olmasına rağmen GA kavramı John Holland ve öğrencileri tarafından geliştirilmiştir. Evrim stratejileri ve evrimsel programlama gibi Holland'ın yaptığı çalışmaların öncesinde geliştirilen çalışmalarda asıl amaç belirli problemleri çözmek için algoritmalar tasarlamak iken, Holland (1975)'in amacı doğada gerçekleştiği şekliyle adaptasyon olgusunu incelemek ve evrim mekanizmalarının bilgisayar sistemlerine aktarıldığı yollar geliştirmek olmuştur (Mitchell 1999).

Holland (1975), genetik algoritma üzerine çalışmalarını "Adaptation in Natural and Artificial Systems" isimli kitabında toparlayıp yayınlamıştır. De Jong (1975) ise "An Analysis of the Behavior of a Class of Genetic Adaptive Systems" isimli çalışmasında GA'nın optimizasyon problemlerinin çözümünde kullanılabileceğini göstermiştir. Yine bu çalışmasında GA'nın en hassas noktalarından biri olan optimum çözümü verecek GA parametrelerinin ayarlanması konusuna ilk defa değinmiştir. Holland'ın öğrencilerinden Goldberg (1989) ise yaptığı çalışmalarla GA'nın tanınmasına büyük katkı sağlamıştır.

GA, ilk çıktığı yıllardan günümüze kadar istatistik ve yöneylem araştırmalarında, çeşitli mühendislik dallarında, bilgi sistemlerinde, lojistik, pazarlama ve finans gibi birbirinden farklı bilim dallarında kullanılmıştır ve kullanılmasına devam edilmektedir.

GA, bir stokastik optimizasyon algoritması olup türev bilgisine ihtiyaç duymaz. GA, diğer stokastik yöntemlerin çoğundan daha etkili ve esnek bir yapıya sahiptir. Ayrıca diğer sezgisel yöntemlerden farklı olarak başlangıç değeri bilgisine gerek duymadan işlem yapar.



GA'yı diğer optimizasyon algoritmalarından ayıran özellikler,

- Parametrelerin kendileriyle değil de onların kodlanmış şekilleriyle çalışmasına olanak sağlaması,
- Popülasyon üzerinde tek bir noktadan değil, pek çok noktadan arama yapması,
- Türevsel veya diğer yardımcı bilgileri değil de doğrudan amaç fonksiyonunun kendisini kullanması,
- Deterministik kuralları değil, olasılıksal geçiş kurallarını kullanması

biçiminde sıralanabilir. Ayrıca hem sürekli ve hem kesikli değişkenlerde çalışabilmeleri, tek bir noktadan değil de pek çok noktadan aynı anda arama yapmaları gibi özellikleri GA'yı avantajlı kılmaktadır (Goldberg 1989).

GA, özellikle tahmin edilecek parametrenin sayısının fazla olduğu karmaşık problemlerin çözümünde diğer optimizasyon yöntemlerine göre daha etkilidir. Ayrıca matematiksel analiz ile çözüm elde edilemeyen geleneksel araştırma yöntemlerinin kendi karakteristik yapılarından ötürü, başarılı sonuç veremedikleri problemlerde GA'dan faydalanılmaktadır.

GA, yukarıda belirtildiği gibi evrim ilkelerinin bilgisayar ortamına uyarlanması fikrine dayanmaktadır. Bu sebeple genetik biliminde kullanılan pek çok terim GA için uyarlanmıştır. GA'da kullanılan bazı temel kavramlar aşağıda tanımlanmıştır.

**Gen:** Biyolojide canlı türlerinin kendi tür özelliklerinin şifreli hâlde bulunduğu ve anlamlı genetik bilgi içeren en küçük diziye verilen isimdir. GA'da da bu terim anlamlı bilgi taşıyan en küçük birimi ifade eder. Tıpkı biyolojide olduğu gibi probleme ait bilgiler genler içerisine kodlanır. Bütün GA işlemleri bu kodlanmış bilgi üzerinden gerçekleştirilir.

**Kromozom:** Genlerin bir araya gelerek oluşturduğu, bütünlük arz eden, kopyalanma ve yeni nesil oluşturma için tüm bilgileri içeren diziye verilen isimdir. GA'da kromozomlar problem için aday çözümlerden birini ifade eder. Örneğin, bir problemin çözümüne ilişkin dört parametre değerinin tahmin edilmesi gerektiği varsayıldığında, her bir parametre için olası çözümler kodlanarak genler oluşturulur. Kromozomlar da genlerden oluştuğuna göre, dört parametre tahmininin bir arada bulunduğu olası çözüme kromozom aracılığıyla ulaşılabilir. GA'da her bir kromozom aynı sayıda gen içermelidir ve algoritmaya kaç adet kromozomla başlanacağına önceden karar verilmelidir. Az sayıda kromozom ile başlamak kötü sonuç elde edilmesine, gereğinden fazla kromozom ile başlamak ise zaman kaybına sebep olabilmektedir.

**Popülasyon:** Kromozomlardan oluşan topluluğa verilen isimdir. Kromozomların her biri, bir çözüm içeren bireyler olarak düşünüldüğünde popülasyon, içinde pek çok muhtemel çözümü barındıran bir toplum olarak düşünülebilir. GA işlemlerinde ilk popülasyon kromozomların rasgele elde edilmesiyle oluşmaktadır.

**Kuşak:** Kuşaklar ilerledikçe popülasyon içindeki kromozomlar değişir ve gelişir. Bunun sonucunda popülasyon, optimum noktaya daha çok yaklaşan bir çözüm kümesi hâlini alır. Kuşak, işlemlerin tekrarlandığı yinleme sayısıdır. İşlemlere başlamadan önce kuşak sayısı belirlenir ve kuşak o sayıya ulaştığında GA sonlandırılır.

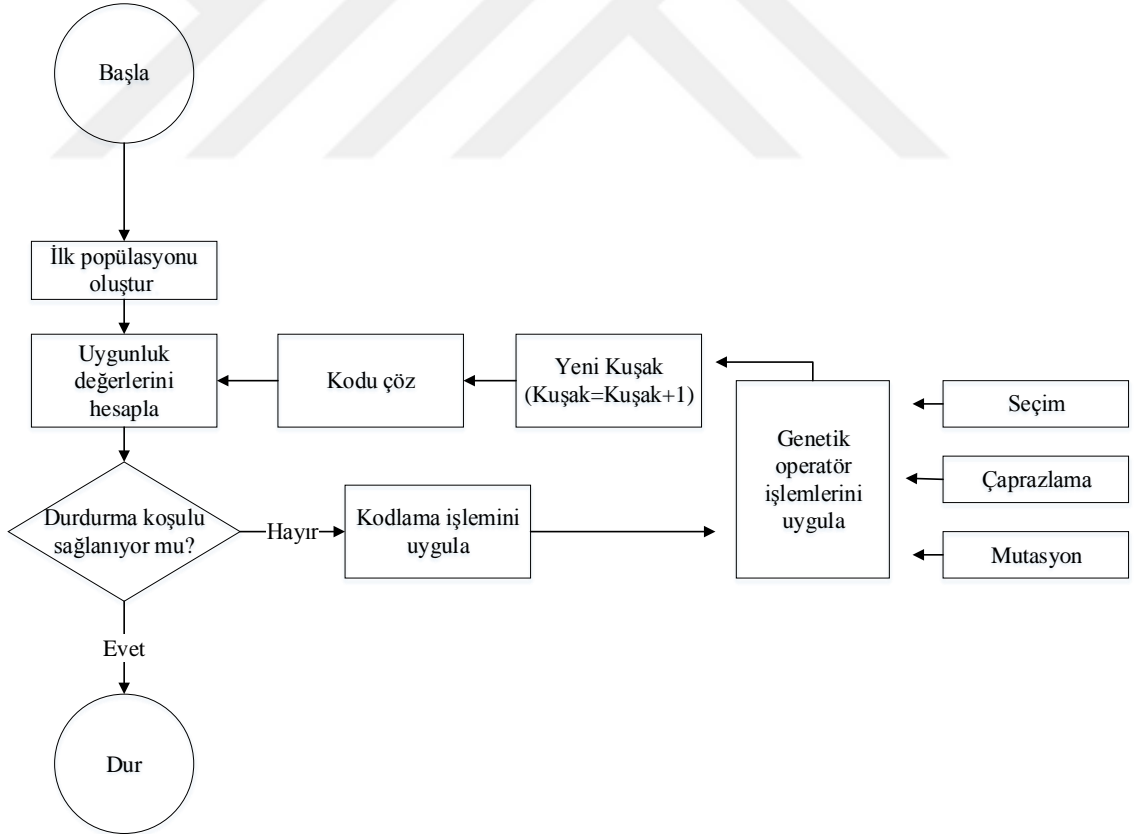
**Uygunluk Değeri:** Algoritmada her kuşak sonucunda oluşan kromozomların optimum sonuca yakınlıklarına göre aldıkları skorlardır. Bu skorlara göre kromozomların bir sonraki kuşağa dahil olma olasılıkları hesaplanır. Uygunluk değeri yüksek olan kromozomların, sonraki kuşaklara katılma olasılıkları daha fazla olduğundan bir süre sonra bu kromozomlar popülasyonda çoğunluğu oluşturur.

**Amaç Fonksiyonu:** Uygunluk değerinin hesaplanmasında kullanılan fonksiyondur. Bütün algoritma işlemleri genler içine kodlanmış bilgiler üzerinden yapılırken uygunluk değeri hesaplaması kodlanmamış veri üzerinden hesaplanmaktadır. Bu sebeple amaç fonksiyonunda kullanılacak kodlanmış çözüm kümelerinin (kromozomların) kodları çözülerek uygunluk değerleri hesaplanır.

Genetik Algoritma,

- i) Başlangıç popülasyonunun oluşturulması,
- ii) Uygunluk değerinin hesaplanması,
- iii) Model parametrelerinin kodlanması,
- iv) Genetik operatör işlemlerinin (seçim, çaprazlama, mutasyon) uygulanması,
- v) Yeni kuşak popülasyonlarının oluşturulması,
- vi) Durdurma koşulunun sağlanması,

adımlarından oluşmaktadır. Şekil 3.6'da GA adımlarının akış diyagramı gösterilmektedir. Bu adımlar aşağıdaki alt bölümlerde ayrıntılı olarak tanımlanmıştır.



Şekil 3.6 Genetik Algoritma adımlarının akış diyagramı

### **i) Başlangıç popülasyonunun oluşturulması**

Başlangıç popülasyonu oluşturulmasına popülasyon büyüklüğünün belirlenmesi (popülasyonda kaç kromozomun olacağına belirlenmesi) ile başlanır. Popülasyon büyüklüğü fazla alındığında seçilen popülasyonun kitleyi temsil düzeyi arttığından daha isabetli sonuçlar elde etme olasılığı artmaktadır. Ancak, bu durum popülasyon büyüklüğünü fazla almanın her zaman mantıklı olacağı anlamına gelmemektedir. Bazı problemlerde küçük hacimli popülasyonla çalışmak benzer kalitede sonuçları çok daha kısa sürede elde etmeyi sağlayabilmektedir. Fakat küçük popülasyonla çalışmak da parametre tahmin uzayını yeterli miktarda örnekleyememe veya zamansız yakınsama problemlerini beraberinde getirebilmektedir. (Bolat vd. 2004). Bazı problemlerde 20-30 kromozomlu bir popülasyon, parametre tahmini için yeterli iken bazı problemler için 50-100 yeterli olmaktadır. Bazı karmaşık problemlerde ise 200-500 arası büyüklüklerdeki popülasyonlar ile global çözüm sonucuna ulaşılmaktadır.

GA'da popülasyon büyüklüğü belirlendikten sonra başlangıç popülasyonu genellikle çözüm uzayından rasgele kromozomlar seçilerek elde edilir. Rasgele belirlemelerde çoğu zaman sayı üretici kullanılır.

### **ii) Uygunluk değerinin hesaplanması**

Başlangıç popülasyonunun oluşturulmasıyla birinci kuşak oluşturulmuş olur. Bundan sonraki amaç, çözüme uzak sonuç veren (kötü) kromozomların popülasyondan giderek elendiği, onların yerini ise optimal parametre tahminine daha yakın sonuçlar veren (iyi) yeni kuşakların aldığı yapıyı oluşturmaktır. Algoritmanın yeni kuşaklarının belirlenmesi için uygunluk değerinin hesaplanması gerekmektedir. Uygunluk değerleri, her bir parametre tahmininin (kromozomun) iyi veya kötü olduğunu tespit etmede kullanılan sayısal değerlerdir. Bu değerlere göre kromozomun sonraki nesillere aktarılma olasılıkları belirlenir. Bir kromozomun uygunluk değeri yüksekse, bir sonraki kuşağa katılma olasılığı da yüksektir. Bir optimizasyon probleminde her bir kromozomun uygunluk değeri, genellikle o noktadaki amaç fonksiyonunun (  $\phi$  ) değeridir. Kromozomlar, tahmin edilen parametrelerin aday çözümlerinin oluşturduğu bir küme

olduđuna gre  $\phi$  fonksiyonunda kromozomlar kullanılarak bařarı lt olarak uygunluk deęeri hesaplaması yapılmaktadır. GA'nın bařarılı sonu vermesi iin  $\phi$  fonksiyonunun optimal parametre tahminlerini bulacak biimde oluřturulması ve uygunluk deęerlerinin  $\phi$  fonksiyonuna baęlı olarak hesaplanması gerekmektedir.

### iii) Parametrelerin kodlanması

Genetikte cinsiyet, gz rengi, kan grubu gibi karakteristik zelliklerin kromozomlardaki DNA molekllerine nkleotidler řeklinde kodlanması gibi, GA'da da, kromozomlar iindeki genlere parametreler uygun bir řekilde kodlanır. Modelin etkili bir řekilde parametre tahminlerinin bulunması iin kodlama iřleminin doęru yapılması nemlidir. Aksi hlde algoritma alıřmayacak veya yanlıř alıřacaktır.

zellikle son yıllardaki alıřmalarla beraber farklı kodlama yntemleri geliřtirilmiřtir. İlk alıřmalarda, Holland (1975), ikili kodlamayı kullanmıřtır. İkili kodlamada ondalık tabandaki deęerler, ikili tabana evrilir. Buna gre, genler, 0 ve 1'lerden oluřur. rneęin (21 10 7) řeklinde oluřturulmuř kromozomun ikili sistemde kodlanması iřlemi ele alınırsa kodlama iřlemi ile beraber kromozom ( 10101 1010 111 ) hlini alacaktır.

İkili kodlama ile alıřmanın genetik operatr iřlemlerinde (aprazlama ve mutasyon iřlemlerinde) avantajları bulunmasına raęmen bazı durumlarda aslında birbirine ok yakın olan deęerlerin ikili koda yazılması sonucunda birbirlerine uzak dřmeleri sorunu ortaya ıkabilmektedir. Bu sorunun zm ise Gray kodlama yntemidir. Gray kodlama yntemiyle basamak fazlalıęından ortaya ıkan karıřıklıklar nlenebilmektedir (Akyol 2006).

Gezgin satıcı ve en kısa yol problemlerinin zmnde zellikle kullanılan bir dięer kodlama yntemi de permtasyon kodlamadır. Bu kodlama ynteminde sıralama esastır ve aynı deęer tekrar kullanılamaz. rneęin mahalleler arası uzaklıkların bilindięi ve altı mahallenin en kısa srede dolařılacaęı bir gezgin satıcı probleminin olası bir zm (5

3 4 1 2 6) şeklinde kodlanabilir. Bu kodlama yönteminde tekrarlama yapılamadığından genetik operatör işlemlerinden sonra yinelenen değerler düzeltme işlemine tabi olur.

Gerçek değerli kodlamada ise değerler kromozom içerisinde, dönüşüme uğramadan, genetik operatör işlemlerine tabi olmaktadır. GA'nın ortaya çıktığı ilk yıllarda ikili kodlama kullanılırken son yıllarda gerçek değerli kodlama yaygınlık kazanmıştır. Bu çalışmada hesaplama kolaylığı bakımından gerçek değerli kodlama kullanımı tercih edilmiştir.

#### **iv) Genetik operatörler**

Genetik operatörler, GA'da evrim süreçlerinin taklit edildiği bölümdür. Nesiller ilerledikçe çevre koşullarına uyum sağlayıp avantajlı konuma gelebilmek için türlerin genetik özelliklerinde bir takım değişimler ve gelişmeler gerçekleşir. Önceki atalardan farklı olarak kazanılmış yeni özellikler, başarılı nesillerde meydana gelmiş genetik değişikliklerin bir sonucudur. GA'da da önceki nesillere göre yeni nesillerin daha başarılı olması genetik operatörler aracılığıyla sağlanır. Üç tane genetik operatör mevcuttur. Bunlar, seçim, çaprazlama ve mutasyon operatörleridir.

**Seçim operatörü:** Tabiatта çevre şartlarına uyum sağlayabilen bireyler hayatlarını devam ettirip üreyebilme fırsatı yakalarken zayıf bireyler hayatta kalamaz ve üreyemezler. Bu mekanizmaya "doğal seçim" adı verilmektedir (Ayala 2017). GA'da seçim operatörü tabiatтаki doğal seçilimin karşılığıdır. Popülasyondaki uygunluk değeri yüksek kromozomlar daha fazla seçilme imkânı bulup genlerini sonraki kuşaklara taşıma imkânı yakalayabiliyorken uygunluk değeri düşük kromozomlar kuşaklar geçtikçe popülasyondan silinirler.

Seçim operatörünün işlevi, popülasyondaki kromozomların hangilerinin gelecek nesillerde çocuk kromozom yaratacağını belirlemektir. Bu seçim işlemi sırasında bazı kromozomlar bir sonraki kuşağa katılım sağlayamazken bazıları birden fazla seçilerek pek çok çocuk kromozomun oluşmasına sebep olur. Elbette bu seçim yapılırken denge

önemlidir. Çok güçlü seçim yapılması, uygun olmayan en iyi bireylerin popülasyonu ele geçirerek daha fazla değişim ve ilerleme için gerekli olan çeşitliliği azaltacağı anlamına gelir, çok zayıf seçim ise çok yavaş evrimle sonuçlanır (Mitchell 1999). Tıpkı kodlama işleminde olduğu gibi, GA literatüründe de çok sayıda seçim yöntemi önerilmiştir. Aşağıda, en yaygın yöntemlerden bazıları açıklanmıştır.

Rulet tekerleği yöntemi, ilk defa Holland tarafından ortaya atılmıştır. Kromozomların uygunluk değerleri bir tabloya yazılır. Tüm uygunluk değerleri toplanır ve her bir kromozomun uygunluk değeri elde edilen toplam değere oranlanır. Böylece her bir kromozomun seçilme şansı (0,1) aralığında bir değer alacak şekilde belirlenmiş olur. Her bir kromozomun seçilme olasılığı tabloya birikimli şekilde yazılır ve (0,1) aralığında rasgele popülasyon büyüklüğü kadar sayı (N) üretilir. Üretilen sayılar hangi kromozomun aralığına denk geliyorsa o kromozom, eşleştirme havuzu adı verilen diğer operatör işlemlerinin yapılacağı bölüme alınır.

Stokastik uniform yöntemi, rulet tekerleği yöntemine benzemesine rağmen rulet tekerleği ile seçilme şansı az olan ebeveynlerin pek çok kere seçilmesi ihtimali bu yöntem ile ortadan kaldırılmaktadır. Rulet tekerleği yönteminde popülasyon büyüklüğü (N) kadar sayı türetilirken bu yöntemde (0, 1/N) aralığında rasgele bir tane sayı (R) türetilir. Seçilen sayı rulet tekerleğinde hangi kromozomun aralığına denk geliyorsa o kromozom ilk ebeveyn olarak seçilir. İkinci ebeveyne ise  $(R+(1/N))$  değerine denk gelen kromozom atanır. Her ebeveyn bir önceki ebeveynin değerine (1/N) eklenerek atanacaktır. Bu şekilde N tane kromozom teker üzerinde eşit aralıkta belirlenmiş olur ve erken yakınsamanın önüne geçilir (Mitchell1999).

Remainder yönteminde, her bir kromozomun uygunluk değerinin popülasyonun ortalama uygunluk değerine bölünmesiyle elde edilen beklenen uygunluk değerlerinin tam sayı kısımları deterministik olarak eşleştirme havuzuna alınır. Geri kalan kontenjan ise beklenen uygunluk değerlerinin kesirli kısımları kullanılarak rulet tekerleği yöntemiyle seçilir.

Rank yöntemi, çok hızlı yaklaşmayı önlemek için olan alternatif bir yöntemdir. Popülasyondaki bireyler uygunluğa göre sıralanır ve bu sıralamaya göre bir seçim yapılır. Bu yöntemde popülasyondaki her bir kromozomun seçilme olasılığı uygunluk değerlerinden ziyade sıralamasına bağlıdır. Rulet tekerleği yönteminde uygunluk değerleri büyük olan bir kaç noktaya hızlı yaklaşma riskinin önüne rank seçimi ile geçilebilir. Bazı problemlerde de rank seçimi değişimi yavaşlattığından uygun bireylerin bulunmasını zorlaştırabilir (Mitchell 1999).

Turnuva yöntemi, popülasyondan iki kromozom rasgele seçilir. İkisi uygunluk değerlerine göre karşılaştırılır ve uygunluk değeri büyük olan turnuvayı kazanarak eşleştirme havuzuna alınır. Bu yöntemin sadece iki kromozomu değil pek çok kromozomu aynı anda yarıştırdığı çeşitleri de mevcuttur. Turnuva seçiminin avantajı yüksek uygunluk değerine sahip bireylerin sonraki kuşağa aktarılmasının yanı sıra seçilemeyen kromozomların başka gruplara dâhil edilip seçilme şanslarını devam ettirebilmeleridir (Karakoca 2009).

**Çaprazlama operatörü:** Çaprazlama operatörü ile iki ebeveyn kromozomun genleri takas edilerek yeni özelliklere sahip çocuk kromozomlar oluşturulur. Böylece popülasyonda çeşitlilik sağlanmış olur ve algoritmanın daha önce ulaşamadığı noktalara ulaşılır. Çaprazlama operatörü GA'nın önemli bir bölümüdür. Çünkü, iyi özelliklerin bir araya gelmesini kolaylaştırır. Bu nedenle, algoritmanın performansını büyük ölçüde etkiler.

Çaprazlama oranı önceden uygulayıcı tarafından belirlenen bir orandır ve yeni nesil elde edilirken eşleştirme havuzunda yer alan kromozomlardan kaç tanesinin çaprazlama işlemine tabi olacağını belirlemede kullanılır. Çaprazlama oranı (0,1) aralığında bir değer olmalıdır ve çalışmaların büyük çoğunluğunda 0.5 ile 1 arasında bir değer almaktadır. Çözüm uzayında daha kapsamlı araştırma yapmak için uygulamalarda genellikle 0.8 ve daha büyük çaprazlama oranları tercih edilmektedir. Literatürde en çok kullanılan çaprazlama yöntemlerinden bir kaçını aşağıda vermiştir.



Tek nokta çaprazlamada, eşleştirme havuzundan iki ebeveyn kromozom rasgele seçilir. Genlerin arasında rasgele bir nokta belirlenir. Belirlenen bu noktadan önce kalan bölüm birinci ebeveyn, kromozomdan sonra kalan bölüm diğer ebeveynden alınır ve birinci çocuk kromozom oluşturulur. Ebeveyn kromozomların arta kalan bölümlerinin birleştirilmesiyle de ikinci çocuk kromozom oluşturulur. Şekil 3.7’de gerçek değerli tek nokta çaprazlamaya bir örnek gösterilmektedir.

$$\begin{array}{l}
 \text{Ebeveyn 1} \quad [ \quad 3 \quad 0.5 \quad : \quad 5 \quad 1 \quad ] \\
 \text{Ebeveyn 2} \quad [ \quad 2 \quad 0.75 \quad : \quad 6 \quad 1.5 \quad ] \\
 \\
 \text{Çocuk 1} \quad [ \quad 3 \quad 0.5 \quad : \quad 6 \quad 1.5 \quad ] \\
 \text{Çocuk 2} \quad [ \quad 2 \quad 0.75 \quad : \quad 5 \quad 1 \quad ]
 \end{array}$$

Şekil 3.7 Gerçek değer kodlamalı tek nokta çaprazlama

İki nokta çaprazlama, tek nokta çaprazlamaya benzemektedir. Sadece değiş tokuş işlemi için seçilen çaprazlama noktalarının sayısı ikiye yükselmiştir. Bu yöntemde ebeveyn kromozomların parça alışverişleri belirlenen iki nokta arasındaki bölümü kapsar. Şekil 3.8’de gerçek değerli iki nokta çaprazlamaya bir örnek gösterilmektedir.

$$\begin{array}{l}
 \text{Ebeveyn 1} \quad [ \quad 3 \quad : \quad 0.5 \quad 5 \quad : \quad 1 \quad ] \\
 \text{Ebeveyn 2} \quad [ \quad 2 \quad : \quad 0.75 \quad 6 \quad : \quad 1.5 \quad ] \\
 \\
 \text{Çocuk 1} \quad [ \quad 3 \quad : \quad 0.75 \quad 6 \quad : \quad 1 \quad ] \\
 \text{Çocuk 2} \quad [ \quad 2 \quad : \quad 0.5 \quad 5 \quad : \quad 1.5 \quad ]
 \end{array}$$

Şekil 3.8 Gerçek değer kodlamalı iki nokta çaprazlama

Scattered çaprazlamada, çaprazlama işleminden önce, 0 ve 1’lerden oluşan ve kromozomdaki gen sayısı büyüklüğünde bir şablon dizisi hazırlanır. İki ebeveyn kromozom rasgele seçilir. Şablon dizisinde 1’e karşılık gelen gen birinci ebeveynden, 0’ a karşılık gelen gen ikinci ebeveynden seçilir ve birinci çocuk kromozom oluşturulur. Ebeveyn kromozomların arta kalan bölümlerinin birleştirilmesiyle de ikinci çocuk

kromozom oluşturulur. Şekil 3.9’da gerçek değerli scattered çaprazlamaya bir örnek gösterilmektedir.

$$\begin{array}{l}
 \text{Şablon dizi} \quad [ 1 \quad 1 \quad 0 \quad 1 ] \\
 \\ 
 \text{Ebeveyn 1} \quad [ 3 \quad 0.5 \quad 5 \quad 1 ] \\
 \text{Ebeveyn 2} \quad [ 2 \quad 0.75 \quad 6 \quad 1.5 ] \\
 \\ 
 \text{Çocuk 1} \quad [ 3 \quad 0.5 \quad 6 \quad 1 ] \\
 \text{Çocuk 2} \quad [ 2 \quad 0.75 \quad 5 \quad 1.5 ]
 \end{array}$$

Şekil 3.9 Scattered çaprazlama

Aritmetik çaprazlamada, (0,1) aralığında rasgele bir sayı ( $r$ ) seçilir. Bir geninin  $\alpha$  kadar ilk ebeveynden alınırken (1- $r$ ) kadar ikinci ebeveynден alınır (Herrera vd. 2003). Şekil 3.10’da gerçek değerli aritmetik çaprazlamaya bir örnek gösterilmektedir.

$$\begin{array}{l}
 \text{Ebeveyn 1} \quad [ \quad 3 \quad \quad \quad 0.5 \quad \quad \quad 5 \quad \quad \quad 1 \quad ] \\
 \text{Ebeveyn 2} \quad [ \quad 2 \quad \quad \quad 0.75 \quad \quad \quad 6 \quad \quad \quad 1.5 \quad ] \\
 (\alpha = 0.4) \\
 \text{Çocuk 1} \quad [ \quad 3 \times 0.4 + 2 \times 0.6 \quad 0.5 \times 0.4 + 0.75 \times 0.6 \quad 5 \times 0.4 + 6 \times 0.6 \quad 1 \times 0.4 + 1.5 \times 0.6 \quad ] \\
 \text{Çocuk 2} \quad [ \quad 3 \times 0.6 + 2 \times 0.4 \quad 0.5 \times 0.6 + 0.75 \times 0.4 \quad 5 \times 0.6 + 6 \times 0.4 \quad 1 \times 0.6 + 1.5 \times 0.4 \quad ] \\
 \quad \quad \quad \quad \quad \quad || \quad \quad \quad || \quad \quad \quad || \quad \quad \quad || \\
 \text{Çocuk 1} \quad [ \quad 2.4 \quad \quad \quad 0.65 \quad \quad \quad 5.6 \quad \quad \quad 1.3 \quad ] \\
 \text{Çocuk 2} \quad [ \quad 2.6 \quad \quad \quad 0.6 \quad \quad \quad 5.4 \quad \quad \quad 1.2 \quad ]
 \end{array}$$

Şekil 3.10 Aritmetik çaprazlama

Heuristic çaprazlamada, (0,1) aralığında rasgele bir sayı ( $r$ ) belirlenir. Gen havuzundan rasgele seçilen iki genden amaç fonksiyonu bakımından iyi olanın Ebeveyn 1, kötü olanın Ebeveyn 2 olduğu varsayıldığında çocuk kromozomlar

$$\begin{aligned}\text{Çocuk 1} &= \text{Ebeveyn 1} \\ \text{Çocuk 2} &= \text{Ebeveyn 1} + r \times (\text{Ebeveyn 1} - \text{Ebeveyn 2})\end{aligned}$$

eşitlikleriyle hesaplanır (Gopi 2007).  $r=0.03$  için Heuristic çaprazlamaya bir örnek, Şekil 3.11’de verilmektedir.

Ebeveyn 1	[	3	0.5	5	1	]
Ebeveyn2	[	2	0.75	6	1.5	]
Çocuk 1	[	3	0.5	5	1	]
Çocuk 2	[	3.3	0.425	4.7	0.85	]

Şekil 3.11 Heuristic çaprazlama

**Mutasyon operatörü:** Mutasyon operatörü ile kromozomlardaki genlerin bir kısmı değiştirilir ve gen havuzunda bulunmayan yeni genlerin ortaya çıkması sağlanır. Böylece seçim ve çaprazlama operatörleri ile sağlanamayan farklılaşma sağlanmış olur. Farklı özellikteki yeni kromozomların elde edilmesi daha önceden inceleme imkânı bulunamayan bilgilerin incelenmesine olanak tanır. Sadece seçme ve çaprazlama operatörleri ile üretilen kuşaklar bir zaman sonra popülasyon içinde birbirine benzer özelliklere sahip çocuk kromozomların ortaya çıkmasına sebep olur ve eğer popülasyon seçimi düzgün yapılamamışsa algoritma yerel minimum noktalara takılabilir. Mutasyon operatörü yerel optimum tuzaklarından kurtulmak için çözüm sağlayabilir.

Mutasyon işlemi, ikili kodlamalı çalışmalarda genin alabileceği değer sadece 0 veya 1 olacağından, mutasyon oranı dikkate alınarak rasgele seçilmiş bitlerde, 0’ların 1 ve 1’lerin 0 yapılması ile gerçekleştirilir. Gerçek değer kodlamalı çalışmalarda ise mutasyon işlemi, yer değişimi veya mevcut değere mutasyon adımı denilen küçük değerler eklenerek veya çıkartılarak elde edilir (Öztürk 2002).

Mutasyon oranı teorik olarak (0,1) aralığında seçilmelidir ve küçük bir oran olarak belirlenmelidir. Yüksek belirlendiği durumlarda çocuk kromozomlar ebeveynlerinden

farklı özellikler gösterecek ve iyi özelliklere sahip kromozomlar popülasyondan kaybolmaya başlayacaktır. Mutasyon oranının çok küçük seçilmesi durumunda ise mevcut popülasyon dışındaki yeni noktaların keşfedilmesi zorlaşacaktır (Artaç 2003). Uygulamalarda mutasyon oranı için genellikle 0.01'den daha küçük bir değer seçilir.

#### **v) Yeni kuşak popülasyonun oluşturulması**

Genetik operatör işlemlerinden sonra, popülasyon büyüklüğü değişmeden, ebeveyn kromozomlar yerlerini, oluşan çocuk kromozomlara bırakır. Kuşaklar ilerledikçe çocuk kromozomlar bir sonraki kuşağın ebeveyn kromozomları olur. GA çok defa çalıştırılarak en iyi uygunluk değerine sahip kromozomlara ulaşılmaya çalışılır.

GA, kuşaklar boyunca elinde bulundurduğu olası çözümler kümesini değiştirir. Bu nedenle her bir kuşak sonunda elde edilen sonuçların önceki kuşaklarla karşılaştırılması gerekir (Özçakar 1998). Bu karşılaştırma, uygunluk dereceleri aracılığıyla yapılır. Burada amaç, en iyi kromozomun oluşturulması, dolayısıyla en iyi parametre tahmin kümesinin belirlenmesidir. En iyi parametre tahmin kümeleri, popülasyon içerisinde hep aynı kalmasalar da seçim operatörüyle (büyük olasılıkla) seçilir, çaprazlama ve mutasyon işlemleriyle geliştirilir. Seçilen kromozomlar üzerindeki işlemler algoritma sonlanıncaya kadar devam eder.

#### **vi) Durdurma koşulunun sağlanması**

GA işlemlerinin sonlandırılmasında kullanılan en yaygın yöntem algoritma başlamadan önce kuşak sayısını belirlemek ve bu kuşak sayısına ulaşıldığında algoritmayı durdurmaktır. Bundan başka, popülasyondaki en iyi uygunluk değerinin ortalama uygunluk değerleriyle karşılaştırılması da algoritmayı sonlandırmada kullanılabilir. Ayrıca kuşaklar boyunca artık en iyi kromozomun uygunluk skoru değişmiyorsa bu durum algoritmanın durdurulmasında ölçüt olarak değerlendirilebilir (Artaç 2003).

### 3.3 Nelder-Mead Simpleks Algoritması ile Genetik Algoritmanın Hibrit Edilmesi

NMS ve GA yöntemleri, yaygın olarak kullanılan türevden bağımsız optimizasyon algoritmalarıdır. Doğrusal olmayan regresyon modellerinde parametrelerin nokta tahminlerinin elde edilmesinde türeve dayalı algoritmalara alternatif olarak bu algoritmaların kullanımları yaygınlık kazanmaktadır. Bu durumun başlıca sebebi türev bilgisine ihtiyaç duymadan tahmin sonuçlarının elde edilebilmesidir. Ayrıca, yerel optimum tuzaklarına takılma olasılıkları Gauss-Newton ve L-M algoritmasına göre daha azdır.

NMS, bölgesel ve doğrudan aramaya dayalı bir yöntemdir. Bu, optimuma ulaşmak için bir sonraki adımda geçerli noktanın yakınındaki noktaları arayacağı anlamına gelmektedir. Sonuç bulmada pratik ve etkilidir. Fakat modelde tahmin edilecek parametre sayısı arttıkça yakınsama kabiliyeti azalır. Ayrıca NMS, işleme başlamak için başlangıç noktasına ihtiyaç duyar. Başlangıç noktasının arama uzayındaki konumu, algoritmanın optimum noktaya doğru ilerlemesi bakımından önemlidir.

GA, NMS'nin aksine global ve çoklu arama yapan bir yöntemdir. Bu, optimuma ulaşmak için bütün arama uzayında aynı anda pek çok noktadan arama işlemine başladığı anlamına gelmektedir. Bu sebeple başlangıç noktası bilgisine ihtiyaç duymaz. Ayrıca, parametrelerin kendisi yerine kodlanmış kümeleriyle çalıştığından sürekli ve kesikli değişkenlerde çalışabilir. NMS'nin aksine sonuç bulurken tahmin edilecek parametre sayısından etkilenmez. Ancak bu avantajlı özelliklerine rağmen dezavantajlı özellikleri de mevcuttur. Algoritmanın iyi sonuç vermesi GA ayarlanabilir parametrelerin iyi belirlenmesine bağlıdır. Ayarlanabilir GA parametrelerinin pek çok farklı seviyesi bulunur. Bu seviyelerin de oluşturduğu pek çok farklı kombinasyon da mevcuttur.

NMS algoritması ve GA'nın avantajlı yönlerinin bir arada kullanıldığı hibrit bir algoritma ile amaç fonksiyonunun en iyilenmesi sağlanabilir. GA ve NMS hibriti olarak (GANMS) adlandırılan bu algoritma ile öncelikle GA ile global arama yapılarak, geniş arama uzayı, global çözümün bulunduğu bölgeye indirgenir. Daha sonra, GA ile

bulunan model parametre tahminleri NMS için başlangıç değeri olarak kabul edilerek global parametre tahminlerine ulaşılması hedeflenir. Bu çalışmada, Türkşen ve Tez (2016)'in çalışmalarında önerdikleri GANMS algoritması doğrusal olmayan regresyon modellerinin parametrelerinin nokta tahminlerinin elde edilmesinde kullanılmıştır.

Bu çalışmada kullanılan GANMS yönteminin algoritma adımları aşağıda tanımlanmıştır.

**Adım 1:** GA ayarlanabilir parametreleri ve bu parametrelerin değerleri belirlenir.

**Adım 2:** Modelin hata kareler toplamı amaç fonksiyonu olarak belirlenir.

**Adım 3:** GA adımları işletilir.

i) GA başlangıç popülasyonu oluşturulur.

ii) Popülasyonundaki her kromozomun amaç fonksiyonu üzerinden uygunluk değerleri hesaplanır.

iii) Model parametreleri kromozomların içinde kodlanır.

iv) Genetik operatör işlemleri uygulanır.

- Seçim

- Çaprazlama

- Mutasyon

v) Yeni kuşak popülasyon oluşturulur.

vi) Uygunluk değeri hesaplaması yapılır.

vii) Durdurma koşulu kontrol edilir. Koşul sağlanıyorsa durulur ve parametrelerin nokta tahminleri belirlenir. Aksi halde, adım (iii)'ye dönlür.

**Adım 3:** GA ile bulunan parametre tahmin değerleri NMS için başlangıç noktası olarak seçilir.

**Adım 4:**  $q$  parametre sayısı olmak üzere, arama uzayında GA ile hesaplanan tahmin noktasını da içeren  $q+1$  nokta belirlenir.

**Adım 5:** NMS operatör işlemleri uygulanır.

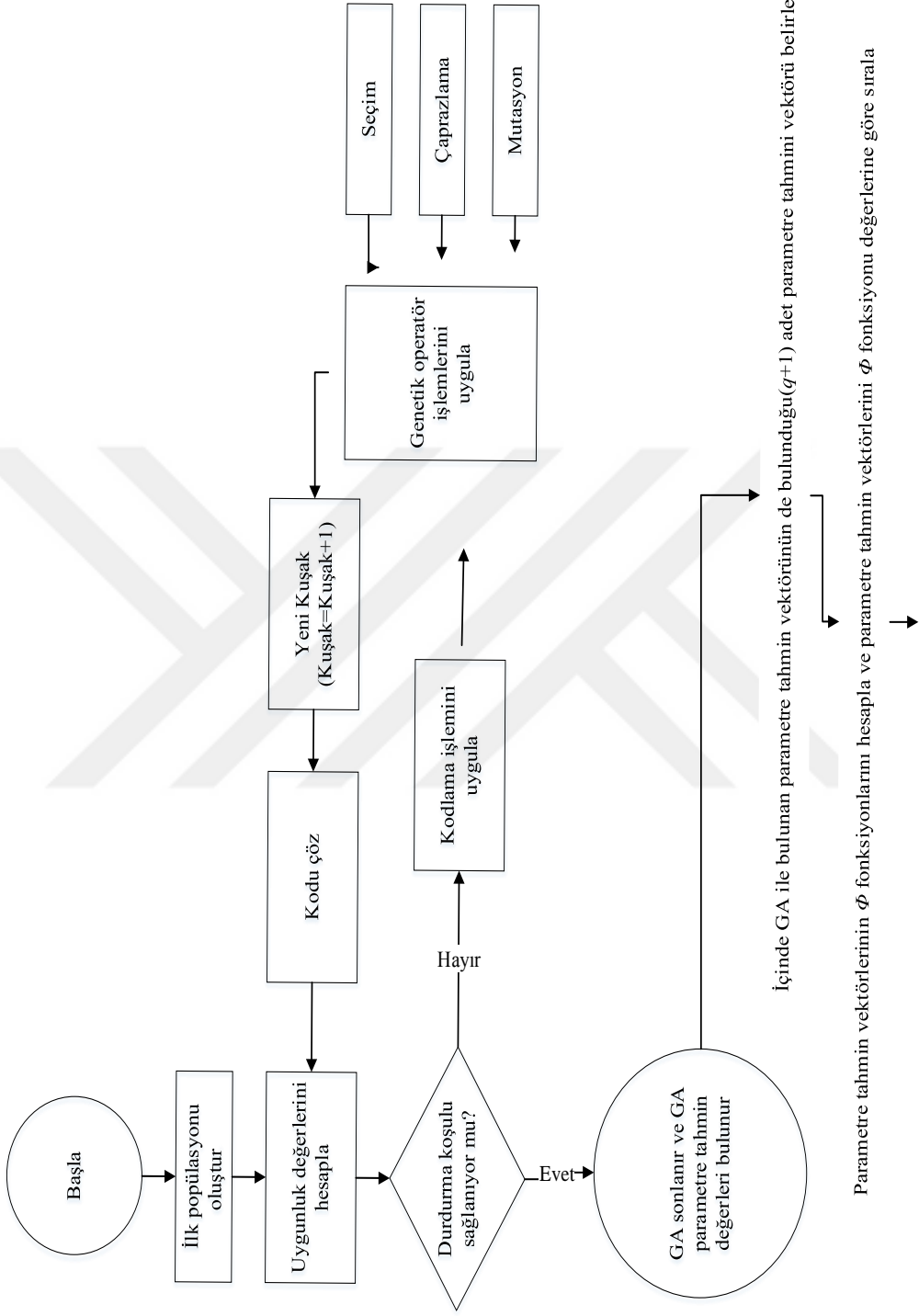
- i) Yansıtma
- ii) Büzülme
- iii) Genişleme
- iv) Küçülme

**Adım 6:** Yeni çözüm noktaları belirlenir.

**Adım 7:** Durdurma koşulu kontrol edilir. Koşul sağlanıyorsa durulur ve parametrelerin nokta tahminleri belirlenir. Aksi halde, adım 5'e dönülür.

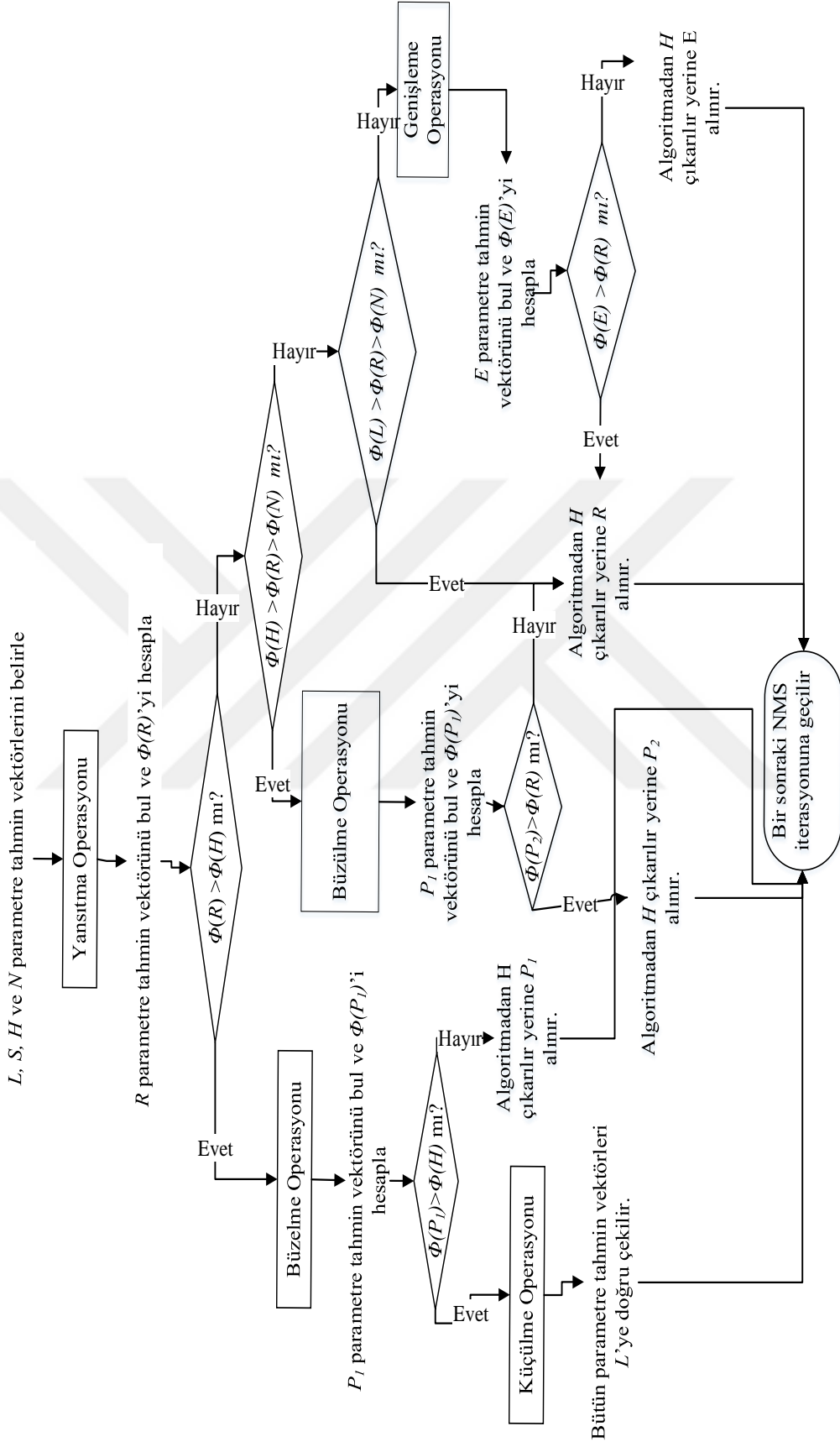
GANMS algoritmasının adımlarının akış diyagramı şekil 3.12'de gösterilmiştir.

GA ayarlanabilir parametrelerinin probleme uygun olarak tanımlanamaması durumunda global çözüm sonuçlarına ulaşmak konusunda başarılı olunamayacağı bilindiğinden, çalışmada, GA ayarlanabilir parametrelerinin uygun kombinasyonlarının belirlenmesi için Taguchi deney tasarımından yararlanılmıştır.



Şekil 3.12 Genetik Algoritma Nelder Mead Simpleks hibrit algoritması akış diyagramı





Şekil 3.12 Genetik Algoritma Nelder Mead Simpleks hibrit algoritması akış diyagramı (devam)

### 3.4 Genetik Algoritma Ayarlanabilir Parametrelerinin Belirlenmesinde Taguchi Deney Tablolarının Kullanımı

Doğrusal olmayan regresyon model parametrelerinin GA ile minimum hatalı tahmin edilmesinde, “Popülasyon büyüklüğü ne olacak?”, “Algoritma kaç kuşak sonra sonlanacak?”, “Seçim yöntemi ne olacak?”, “Çaprazlama ve mutasyon işlemleri nasıl yapılacak?” gibi soruların ilgilenecek probleme özgü en iyi sonucu verecek biçimde yanıtlanması gerekir. Bu nedenle, GA ayarlanabilir parametrelerinin sahip oldukları pek çok seviye arasından en uygun olanına karar verilmesi model parametrelerinin nokta tahmin sonuçlarını etkileyecektir.

GA ayarlanabilir parametrelerinin uygun olabilecek çok farklı kombinasyonları mevcuttur. Bu kombinasyonlar içinden seçim yapmak ciddi emek ve zaman kayıplarına yol açacağından, etkili ve güçlü bir GA araması için ayarlanabilir parametre değeri belirleme yöntemine ihtiyaç vardır. Taguchi tasarımı, en az sayıda deney ile en güçlü parametre kombinasyonuna ulaşmak için etkili bir deneysel tasarım yöntemidir. Bu nedenle çalışmada en etkin ayarlanabilir parametre kombinasyonuna ulaşabilmek amacıyla etkili bir deney tasarım yöntemi olan Taguchi tasarımı kullanılmıştır.

Taguchi'nin kullandığı deney tasarımı yöntemleri klasik çok etkenli deney tasarımı ve kesirli çok etkenli deney tasarımı geliştirilmiştir. Fakat, bu iki yöntemle göre pratiklik bakımından avantajları bulunmaktadır.

Çok etkenli bir deney tasarımı, belirli bir etken kümesi için olası tüm kombinasyonları deneyecektir. Bu da etken ve düzeylerin fazla olduğu deneylerde çok fazla denemenin yapılacağı anlamına gelmektedir. Örneğin, üç etken ve bu etkenlerin her birinin beş düzeyinin bulunduğu bir problem için 125 deneme yapılması gerekmektedir. Oysa Taguchi tasarımında aynı deneyin yapılması için 25 deneme yeterlidir. Taguchi yaklaşımı ile yapılan bir çalışmanın tam faktöriyel tasarım yaklaşımına göre yapılanın en az %90'ı kadar etkin olduğu gösterilmiştir. Çalışmadaki maliyetlerin büyük oranda azalması dikkate alındığında etkinlikteki bu kayıp kabul edilebilir düzeyde bulunmaktadır (Çelik 1993).

Kesirli çok etkenli deneylerde de tam tasarımda yapılması öngörülen bazı denemeler ihmal edilerek, zaman, çaba ve maliyet açısından avantaj elde edilir. Fazla etkenli ve yüksek düzeyli problemlerde bazı yüksek çoklu etkileşimlerin göz ardı edilebileceği varsayılarak bu etkileşimler hata olarak kabul edilir ve denemelerde belirlenen bir kesir oranında kesintiye gidilir. Fakat kesirli çok etkenli tasarımda etkileşimlerin tümü ihmal edilmedikçe ana etkilerin tahminleri hakkında bilgi elde edilmesinin çok zor olduğu söylenmektedir (Dervişoğlu ve Muluk 2007). Ayrıca denemelerde kesir sayısı arttıkça deney pratik olmaktan uzaklaşmaktadır (Muluk vd. 1998). Taguchi yöntemi kesirli denemelerden farklı olarak yapısında dikey dizin özelliğini bulundurmaktadır ve bu sebeple Taguchi tasarımında dikey dizinin görselliği, sapmanın hangi etkileşim bileşenlerinden kaynaklandığını gösterdiğinden bazı tasarımlar için daha iyi tahminler elde edilebilmektedir (Dervişoğlu ve Muluk 2007).

Taguchi deney dizilimlerinin hazırlanması dikey dizin adı verilen matris kullanılarak gerçekleştirilir. Dikey dizinlerden üretilen Taguchi tablolarının gösterimi  $L_N(s^k)$  formatındadır. Burada

$N$ : Dizindeki deneme sayısı (sıra sayısı),

$s$ : Düzey sayısı (eleman sayısı),

$k$ : Etken sayısıdır (sütun sayısı),

dır. Dikey dizinlerin temeli ise Bose ve Bush (1952) tarafından yürütülen deney tasarımı ve ayırım kümeleri teorisi çalışmalarına dayanır. Dikey dizinler dik latin karelerinin uzantıları olarak geliştirilmiştir (Kacker vd. 1991). Taguchi dizinlerin üç önemli özelliği vardır.

i) Her bir sütunda (etkende) düzeylerin görülme sıklığı aynıdır.

ii) Herhangi iki sütunda düzeylerin birlikte görülme sıklığı aynıdır. Başka deyişle herhangi iki sütun iki etkenli bir faktöriyel tasarım modeli oluşturur.

iii) Bazı sütunlar silinse veya sütunların ve satırların yerleri değiştirilse bile oluşan matris yine diklik koşullarını sağlar.

Taguchi tabloları, denemelerde seviyelerin en az değişkenlik gösterdiği sütunlar (etkenler) en soldakiler olacak şekilde tasarlanmıştır. Seviyelerin değişme sıklıkları sütunlar sağa doğru ilerledikçe artmaktadır. Bu sebeple etkenlerin maliyetlerine göre soldan sağa doğru sıralanmaları ve en maliyetli etkenlerin en sol sütunda yer alması doğru olacaktır.

Taguchi kataloğunda 20 adet deney tablosu mevcuttur. Bu tablolardan ikisi hariç hepsi dikey dizinlerden elde edilmiştir ( $L_9(2^{21})$  ve  $L_{27}(3^{22})$  tabloları dikey dizin değildir). Oluşturulma yöntemlerine göre 20 tabloyu 3 gruba ve 9 alt gruba bölmek mümkündür. Çizelge 3.1’de oluşturulma biçimlerine göre Taguchi tasarımları sınıflandırılmıştır.

Çizelge 3.1 Oluşturulma biçimlerine göre Taguchi tasarımları

Aynı Düzeyli Dikey Dizine Oluşturulanlar				Karışık Düzeyli Diken Dizine Oluşturulanlar	Dikey Dizine Kullanılmadan Oluşturulanlar
İki Düzeyli $2^t, t=2,3,4,5,6$	$L_4(2^3)$	Üç Düzeyli $3^t, t=2,3,4$	$L_9(3^4)$	$L_{18}(2^1.3^7)$ $L_{32}(2^1.4^9)$ $L_{50}(2^1.5^{11})$	$L_9(2^{21})$  $L_{27}(3^{22})$
	$L_8(2^7)$		$L_{27}(3^{13})$		
	$L_{16}(2^{15})$		$L_{81}(3^{40})$		
	$L_{32}(2^{31})$	Dört Düzeyli $4^t, t=2,3$	$L_{16}(4^5)$	$L_{36}(2^{11}.3^{12})$	
	$L_{64}(2^{61})$		$L_{64}(4^{21})$	$L_{36}(2^3.3^{13})$	
İki Düzeyli Diğer	$L_{12}(2^{11})$	Beş Düzeyli $5^t, t=2$	$L_{25}(5^6)$	$L_{54}(2^1.3^{25})$	

Aynı düzeyli oluşturulan Taguchi tablolarında tam bir dikey dizin,  $(s^t - 1) / (s - 1)$  sütuna ve  $s^t$  satıra sahiptir. Bu dizinin oluşturulması için öncelikle  $t$  sütunlu ve  $(s^t - 1) / (s - 1)$  satırlı üreteç matris oluşturulur. Bu üreteç matrisin  $0,1,2,\dots,s-1$  ile ifade edilen  $s$  tane

elemanı bulunmaktadır. Üreteç matrisinin  $t$  sütunu dikey dizinde yer alan temel sütunlarını belirtmektedir ve bu sütunlar  $x_1, x_2, x_3, \dots, x_t$  olarak adlandırılmaktadırlar. Geriye kalan sütunlar, temel sütunların üreteç matristeki uygun düşen katsayıyla çarpılması (  $a_1x_1 + a_2x_2 + a_3x_3 + \dots + a_t x_t$  ) sonucu bulunur. Çarpım ve toplama işlemlerinde modüler aritmetik veya galois cebri kullanılır. Örnek olarak  $L_8(2^7)$  tasarımı için,  $N=8=2^3$ ,  $t=3$  temel sütun mevcuttur. Üç temel sütundan elde edilen toplamda yedi sütun, üreteç matrisindeki katsayılar dikkate alınarak çizelge 3.2'deki gibi oluşturulur. Taguchi tablolarının oluşturulmasında kullanılan üreteç matrislerine ilişkin detaylı bilgi Kacker vd. (1991) çalışmasında yer almaktadır.

Bu çalışmada aynı düzeyli dikey dizinlerle üreteç oluşturulan tablolardan  $L_8(2^7)$  ve  $L_{25}(5^6)$  örnek olarak incelenmiştir. Diğer tablolar ile ilgili detaylı bilgi Kacker vd. (1991) ve Dervişoğlu ve Muluk (2007) tarafından yapılan çalışmalarda bulunabilir.

Bu çalışmada örnek olarak  $L_8(2^7)$  tablosu için oluşturulan üreteç tablosu çizelge 3.2'de gösterilmektedir.

Çizelge 3.2  $L_8(2^7)$  için üreteçlerin listesi (Kacker vd. 1991)

Sütun Sayısı	Üreteç
1	$x_1$
2	$x_2$
3	$x_1 + x_2$
4	$x_3$
5	$x_1 + x_3$
6	$x_2 + x_3$
7	$x_1 + x_2 + x_3$

1., 2. ve 4. sütunlar temel sütunlardır. Temel sütunların elemanları en soldaki en az en sağdaki en fazla değişecek şekilde yazılır. Temel sütunlar dışında kalan diğer sütunların bulunmasında gerekli hesaplamalar yapılırken ikili mod kullanılır ( $0+0=0$ ,  $0+1=1$ ,  $1+1=0$ ). İşlemler sonucunda  $L_8(2^7)$  Taguchi tasarımı çizelge 3.3'teki gibi oluşur.

Çizelge 3.3  $L_8(2^7)$  tablosu

<b>Etkenler</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>
<b>Denemeler</b>							
<b>1</b>	0	0	0	0	0	0	0
<b>2</b>	0	0	0	1	1	1	1
<b>3</b>	0	1	1	0	0	1	1
<b>4</b>	0	1	1	1	1	0	0
<b>5</b>	1	0	1	0	1	0	1
<b>6</b>	1	0	1	1	0	1	0
<b>7</b>	1	1	0	0	1	1	0
<b>8</b>	1	1	0	1	0	0	1

Taguchi tablolarının oluşturulmasına ikinci örnek  $L_{25}(5^6)$  tablosu için verilmektedir. Beş düzeyli  $L_{25}(5^6)$  tablosunun oluşturulması sırasında gerçekleşen işlemler  $L_8(2^7)$  tablosunun oluşturulmasındaki gibidir. Üreteç tablosu aynı şekilde oluşturulur ve işlemler mod 5 kullanılarak yapılır.  $L_{25}(5^6)$  için oluşturulan üreteç tablosu çizelge 3.4'deki gibidir.

Çizelge 3.4  $L_{25}(5^6)$  için üreteç tablosu (Kacker vd. 1991)

Sütun Sayısı	Üreteç
1	$x_1$
2	$x_2$
3	$x_1 + x_2$
4	$2x_1 + x_2$
5	$3x_1 + x_2$
6	$4x_1 + x_2$

İşlemler Sonucunda  $L_{25}(5^6)$  Taguchi tasarımı çizelge 3.5'teki gibi oluşacaktır.

Çizelge 3.5  $L_{25}(5^6)$  tablosu

DENEME	A	B	C	D	E	F
1	0	0	0	0	0	0
2	0	1	1	1	1	1
3	0	2	2	2	2	2
4	0	3	3	3	3	3
5	0	4	4	4	4	4
6	1	0	1	2	3	4
7	1	1	2	3	4	0
8	1	2	3	4	0	1
9	1	3	4	0	1	2
10	1	4	0	1	2	3
11	2	0	2	4	1	3
12	2	1	3	0	2	4
13	2	2	4	1	3	0
14	2	3	0	2	4	1
15	2	4	1	3	0	2
16	3	0	3	1	4	2
17	3	1	4	2	0	3
18	3	2	0	3	1	4
19	3	3	1	4	2	0
20	3	4	2	0	3	1
21	4	0	4	3	2	1
22	4	1	0	4	3	2
23	4	2	1	0	4	3
24	4	3	2	1	0	4
25	4	4	3	2	1	0

#### 4. DOĞRUSAL OLMAYAN MODEL PARAMETRELERİNİN ARALIK TAHMİNİNDE BOOTSTRAP YAKLAŞIMI

Gerçek dünyada ilgilenilen bir probleme ilişkin verilerine ilişkin kitlenin bütünü hakkında bilgi sahibi olmak genellikle pek mümkün değildir. Kitle yerine kitleden seçilecek örneklem ile kitle parametrelerinin nokta tahminlerini yapmak emek, zaman, maliyet gibi konularda bazı avantajlar sağlıyor olsa da seçilen örneklemin kitleyi en iyi şekilde temsil edememesi durumunda hatalı sonuçlar elde edilmesine neden olacaktır.

Kitle dağılımı hakkında herhangi bir varsayım yapılmamışsa, örneklem hacminin küçük olduğu durumlarda, kitle parametrelerinin aralık tahminlerini yapmak olanaklı değildir. Bu gibi sorunların üstesinden gelebilmek amacıyla rasgele seçilen  $n$  hacimli bir örneklem kitle olarak kabul edilip, bu örneklemden çok sayıda yeniden  $n$  hacimli rasgele örnekler alınıp, ilgili tahmin edicinin değeri gözlenerek, tahmin edicinin yapay bir örnekleme dağılımı oluşturulabilir. İadeli olarak çok sayıda yeniden  $n$  birimlik rasgele örnek çekilmesine "Bootstrap Örnekleme Yöntemi" denir. Bilgisayar yardımıyla çok sayıda bootstrap örneği artık günümüzde yapılan çalışmalar ile kolaylıkla elde edilebilmektedir.

Efron (1979) geliştirdiği Bootstrap yöntemi ile parametrelerin nokta tahmininde uygun tahmin değerinin elde edilmesini, parametre tahminlerinin standart hatalarının ve güven aralıklarının belirlenmesini hedeflemiştir (Chernick 2008). Bootstrap yönteminin regresyon modellerinde kullanılması, Freedman (1981) ve Wu (1986) tarafından yapılan çalışmalarla başlamıştır. Bootstrap yöntemi ile ilgili daha sonra yapılan önemli çalışmalar, Efron ve Tibshirani (1986, 1993), DiCiccio ve Tibshirani (1987), Shao ve Tu (1995) ve Davison ve Hinkley (1997) çalışmaları olarak kabul edilebilir. Bootstrap yöntemi standart hata, güven aralığı ve regresyon analizi hesaplamaları dışında zaman serileri, kümeleme analizi ve diskriminant analizlerinde de kullanılmaktadır.

Bootstrap, örneklem büyüklüğünün normallik varsayımını sağlayamayacak kadar küçük olması, değişkenler arasında doğrusal olmayan bir kombinasyonun bulunması veya ortalamadan başka bir konum istatistiği ile ilgilenilmesi gibi istatistiksel sonuç



çıkarmasını zorlaştırıcı koşullarda başarılı sonuçlar verdiği için tercih edilen bir yöntemdir (Bello vd. 2015).

Niteliği bilinmeyen bir  $F$  dağılımından seçilen bağımsız  $n$  birimlik  $(X_1, X_2, \dots, X_n)$  örnekleme ile bu  $F$  dağılımına ait  $\theta$  parametresi ve bu parametrenin güven aralıkları tahmin edilmeye çalışılsın. Seçilen  $n$  adet gözlemin içinden değişkenlerin her birinin seçilme olasılığı eşit  $(1/n)$  olan kesikli dağılıma deneysel dağılım fonksiyonu denir. Deneysel dağılım fonksiyonu

$$\hat{F}(x) = \frac{1}{n} \sum_{i=1}^n 1_{\{x_i \leq x\}} \quad (4.1)$$

biçiminde ifade edilir. Deneysel dağılım fonksiyonu ( $\hat{F}$ ),  $F$  dağılım fonksiyonunun bir tahminidir.  $\hat{F}$  dağılımından elde edilecek parametre tahmini de  $\hat{\theta}$  ile ifade edilir. Deneysel dağılım fonksiyonunda  $\hat{\theta}$ 'yı bulmak için Bootstrap yöntemi kullanılması durumunda  $n$  birimli değişkenden, seçilen değişkeni tekrar yerine koyarak, eşit olasılıkla yeniden  $n$  birim değişken seçilir. Bu durum yeni oluşturulacak  $n$  hacimli örnekleme  $(X^* = (X_1^* X_2^* \dots X_n^* ))$  asıl örnekleme'deki değişkenlerin bazılarının birden fazla bulunacağı bazılarının ise hiç bulunmayacağı anlamına gelmektedir. Uygulanan işlem bağımsız olarak  $B$  defa tekrar edildiğinde bootstrap veri kümesi

$$\begin{aligned} X^{*1} &= (X_1^{*1}, X_2^{*1}, \dots, X_n^{*1}) \\ X^{*2} &= (X_1^{*2}, X_2^{*2}, \dots, X_n^{*2}) \\ &\cdot \\ &\cdot \\ &\cdot \\ X^{*B} &= (X_1^{*B}, X_2^{*B}, \dots, X_n^{*B}) \end{aligned}$$

biçiminde tanımlanabilir.  $X^{*1}, X^{*2}, \dots, X^{*B}$  bootstrap veri kümelerinden her biri için ayrı ayrı  $\hat{\theta}_1^*, \hat{\theta}_2^*, \dots, \hat{\theta}_B^*$  tahmin değerleri hesaplanabilir.

Örneğin; tahmin edilmek istenen istatistik, ortalama ise

$$\hat{\theta}_1^* = \bar{X}^{*1} = \sum_{i=1}^n X_i^{*1} / n,$$

$$\hat{\theta}_2^* = \bar{X}^{*2} = \sum_{i=1}^n X_i^{*2} / n,$$

$\hat{\theta}_B^* = \bar{X}^{*B} = \sum_{i=1}^n X_i^{*B} / n$  biçiminde hesaplanır. Bu hesaplamaların aritmetik ortalaması,  $\theta$  parametresi için

$$\hat{\theta}_{(.)}^* = \sum_{b=1}^B \hat{\theta}_b^* / B \quad (4.2)$$

biçiminde tanımlı bootstrap tahmini olacaktır.  $\theta$  parametresi için bootstrap tahmininin standart hatası ise

$$s\hat{e}^* = \left[ \sum_{b=1}^B (\hat{\theta}_b^* - \hat{\theta}_{(.)}^*)^2 / (B-1) \right]^{1/2} \quad (4.3)$$

biçiminde hesaplanır (Efron ve Tibshirani 1993).

#### 4.1 Bootstrap Güven Aralığı Yaklaşımları

Güven aralıklarını oluşturmanın tartışmasız en temel yolu, dağılımı bilinen tahmin edicinin bir fonksiyonunu aramak ve daha sonra parametre için güven aralığı oluşturmak adına bu bilinen dağılımın yüzdelik dilimlerini kullanmaktır (Carpenter ve Bithell 2000). Klasik yaklaşıma göre  $((\hat{\theta} - \theta) / S_{\hat{\theta}})$  dağılımının standart normal dağılıma asimptotik yaklaştığı varsayımı altında genel olarak standart normal dağılımın  $\alpha=0.05$  yüzdelik diliminde güven aralıkları belirlenmektedir. Ancak, küçük hacimli örneklerde örneklemin seçildiği dağılım hakkında yeterli bilgiye sahip olunamaması durumunda parametrenin ve tahmin edicinin hangi fonksiyonunun seçileceği belirsizleşmektedir ve asimptotik yaklaşım yanıltıcı olabilmektedir (Carpenter ve Bithell 2000).

Ayrıca, örneklemin seçildiği kitle çarpık bir dağılıma sahipse simetrik bir yapıya sahip olan standart normal dağılıma göre hesaplanan güven aralıkları gerçek güven aralıklarını tam olarak karşılayamayabilir. Ayrıca, ortalama gibi bir istatistik, analitik olarak hesaplanabiliyorken daha karmaşık istatistiklerin hesaplanmasında analitik formüller yetersiz kalabilmektedir (Wehrens vd. 2000).

Araştırmacılar, normallik varsayımı altında gerçekleştirilen klasik güven aralığı hesaplama yönteminin belirtilen zorlukları aşması adına yan düzeltmeleri ve parametre dönüşümleri gibi bir takım çözümlere başvurmuşlardır. Bootstrap yöntemi bu zorluklar karşısında herhangi bir müdahale olmadan başarıyla sonuçlar üretmektedir (Di Ciccio ve Efron 1996). Bootstrap güven aralıklarını hesaplamının pek çok yöntemi bulunmaktadır. Bu yöntemlerden en bilineni ve uygulaması en kolay olanı Yüzdalık yöntemidir.

Yüzdalık yöntemi ile yüzde aralığı  $\hat{\theta}_b^*$  değerlerinin sıralanması sonucu  $(1-\alpha)$  yüzde aralığında  $(\alpha/2)$  ve  $(1-\alpha/2)$ 'ye karşılık gelen değerlerin bulunması ile hesaplanır. Bu yöntemde göre bootstrap örneklem istatistiğinin dağılımı örneklem istatistiğinin doğrudan bir yaklaşımı olarak kullanılır. Yöntemin algoritmik adımları aşağıdaki gibi tanımlıdır.

**Adım 1:**  $\theta$  parametresini tahmin etmek için  $X_1, X_2, \dots, X_n$  örneklemini kitle olarak düşünülüp,  $X_1, X_2, \dots, X_n$  birimlerinden bağımsız  $B$  adet bootstrap veri kümesi oluşturulur ve her bir veri kümesi için  $\hat{\theta}_b^*$ ,  $b=1,2,\dots,B$  tahminleri hesaplanır.

**Adım 2:** Hesaplanan  $\hat{\theta}_b^*$ ,  $b=1,2,\dots,B$  tahminleri küçükten büyüğe sıralanır.

**Adım 3:**  $(1-\alpha)$  güven düzeyindeki yüzde aralığı sınırlarına denk gelen  $\hat{\theta}_b^*$  değerleri bulunur. (Örneğin,  $\alpha=0.05$  ve  $B=1000$  için 25. ve 975. sıradaki  $\hat{\theta}_{(\alpha/2)}^*$  ve  $\hat{\theta}_{(1-\alpha/2)}^*$  değerleri yüzde aralığı sınırları olarak belirlenir. Eğer  $\alpha.B$  bir tamsayı değilse  $\hat{\theta}_{(\alpha/2)}^*$  alt

sınırını  $\alpha.(B+1)$ 'den küçük olan en büyük tamsayı olarak tanımlanan bir  $k$  sayısı için  $k$ . sıradaki değer,  $\hat{\theta}_{(1-\alpha/2)}^*$  üst sınırını da  $(B-k)$ . sıradaki değer verecektir. )

Son durumda  $(1-\alpha)$  güven düzeyinde hesaplanan yüzde aralıkları

$$P\left[\left(\hat{\theta}_{(\alpha/2)}^*\right) < \theta < \left(\hat{\theta}_{((1-\alpha)/2)}^*\right)\right] = 1 - \alpha \text{ biçiminde olur.}$$

Bootstrap tahmini genellikle yanlı bir tahmindir. Yan miktarı, bootstrap örneklem dağılımından hesaplanan tahminin gerçek örneklem dağılımından hesaplanan tahminden farkı kadardır. Bootstrap örnek sayısı arttıkça yan miktarının azalıp bootstrap tahmini ile gerçek örneklem parametre tahminin birbirine yaklaşması beklenmektedir. Yüzdelik yöntemi, yan miktarını dikkate almadan hesaplama yapan bir yöntemdir. Yanlılığı düzeltilmiş (Bias Corrected - BC) güven aralığı yöntemi yüzdelik yönteminin bu eksikliğini gidermek için geliştirilmiştir. Yöntemin algoritmik adımları aşağıdaki gibidir.

**Adım 1:**  $\theta$  parametresini tahmin etmek için  $X_1, X_2, \dots, X_n$  örneklemini kitle olarak ele alınıp, gerçek örneklem parametre tahmini ( $\hat{\theta}$ ) hesaplanır.

**Adım 2:** Rasgele değişkenlerden bir birinden bağımsız  $B$  adet bootstrap veri kümesi oluşturulur ve her bir küme için  $\hat{\theta}_b^*$ ,  $b=1, 2, \dots, B$  tahmini hesaplanır. Hesaplanan  $\hat{\theta}_b^*$  tahminleri küçükten büyüğe sıralanır.

**Adım 3:**  $\hat{\theta}$  değerinden küçük olan  $\hat{\theta}_b^*$  tahminlerinin sayısı, ( $r$ ), bulunur. Yan düzeltme katsayısı  $\hat{b} = \phi^{-1}(r/B)$  formülü ile bulunur. Buradaki  $\phi^{-1}$ , kümülatif standart normal dağılım fonksiyonunun tersidir. (Örneğin,  $r=0.05 \times B$  olduğu durumda  $\hat{b} = \phi^{-1}(0.05) \approx -1.645$ ,  $r=B/2$  olduğu durumda  $\hat{b} = \phi^{-1}(0.5) \approx 0$ ,  $r=0.95 \times B$  olduğu durumda  $\hat{b} = \phi^{-1}(0.95) \approx 1.645$  olacaktır.)

**Adım 4:**  $\phi$  standart normal kümülatif dağılım fonksiyonu olmak üzere  $\alpha_{alt} = \phi(2\hat{b} + Z_{\alpha/2})$  ve  $\alpha_{üst} = \phi(2\hat{b} + Z_{1-\alpha/2})$  hesaplanır.

**Adım 5:** Küçükten büyüğe sıralanan bootstrap tahminlerinden  $\alpha_{alt} \times B$ . sıradaki değer  $\hat{\theta}_{alt}^*$ ,  $\alpha_{üst} \times B$ . sıradaki değer  $\hat{\theta}_{üst}^*$  olarak belirlenir.

Son durumda  $(1-\alpha)$  güven düzeyinde hesaplanan BC güven aralıkları  $P[(\hat{\theta}_{alt}^*) < \theta < (\hat{\theta}_{üst}^*)] = 1 - \alpha$  biçiminde olur.

BC güven aralığı yaklaşımı bootstrap örneklem dağılımının asimetri düzeltmesini yapar fakat çarpıklık düzeltmesini yapmaz. Bu yönteme çarpıklık düzeltmesi için bir katsayı eklenerek yanlılığı düzeltilmiş ve hızlandırılmış (Bias Corrceted and Accelerated - BCa) güven aralıklarını oluşturmak mümkündür (Efron 1992).

BCa güven aralığı yaklaşımı, BC yaklaşımının güven aralığı değerlerinin örneklem dağılımının çarpıklığına göre düzeltilmiş versiyonudur. BC yaklaşımına çarpıklık düzeltmesi için ivme ( $a$ ) olarak adlandırılan katsayı eklenerek hesaplanır. BCa yaklaşımı, gerçek örneklem dağılımına göre yanlı olan bootstrap örnekleme dağılımı üzerinde bir takım ayarlamalarla yüzdellik yöntem ve BC yaklaşımı ile bulunan güven aralıklarında iyileştirme yapar (Houkoos ve Lewis 2005). Yöntemin algoritmik adımları aşağıdaki gibidir.

**Adım 1:**  $\theta$  parametresini tahmin etmek için  $X_1, X_2, \dots, X_n$  örnekleme kitle olarak varsayıлып, parametre tahmini ( $\hat{\theta}$ ) hesaplanır.

**Adım 2:** Mevcut örnekten birbirinden bağımsız  $B$  adet bootstrap veri kümesi oluşturulur ve her bir küme için  $\hat{\theta}_b^*$ ,  $b = 1, 2, \dots, B$  tahmini hesaplanır. Hesaplanan  $\hat{\theta}_b^*$  tahminleri küçükten büyüğe sıralanır.

**Adım 3:** BC yönteminde olduğu gibi, yan düzeltme katsayısı  $\hat{b} = \phi^{-1}(r/B)$  formülü ile hesaplanır.

**Adım 4:** İvme sabiti ( $a$ ) hesaplanır.  $a$  hesaplamasında Jackknife parametre tahmini hesaplamasına ihtiyaç duyulur. Jackknife yöntemi tıpkı Bootstrap yöntemi gibi bir yeniden örnekleme yöntemidir. Jackknife yönteminde, örneklemden her seferinde  $X_1, X_2, \dots, X_n$  rasgele değişkenlerinin bir tanesi çıkartılarak  $n$  tane Jackknife örneği

oluşturulur. Örneğin,  $b$ . örnek için Jackknife örneği  $\hat{\theta}_{(b)}^j = \frac{\sum_{i=1}^{b-1} x_i + \sum_{i=b+1}^n x_i}{n-1}$  biçiminde

oluşur. Jackknife parametre nokta tahmini ise  $\hat{\theta}_{(.)}^j = \sum_{i=1}^n \frac{\hat{\theta}_{(i)}^j}{n}$  şeklinde bulunur. Jackknife

hakkında detaylı bilgi Efron (1982) tarafından yapılan çalışmada bulunabilir. Jackknife

örnek seti kullanılarak ivme sabiti  $a = \frac{\sum_{i=1}^n (\hat{\theta}_{(i)}^j - \hat{\theta}_{(.)}^j)^3}{6 \left[ \sum_{i=1}^n (\hat{\theta}_{(i)}^j - \hat{\theta}_{(.)}^j)^2 \right]^{3/2}}$  biçiminde hesaplanır.

**Adım 5:**  $\phi$  standart normal kümülatif dağılım fonksiyonu olmak üzere

$$\alpha_{alt} = \phi \left( \hat{b} + \frac{Z_{\alpha/2} + \hat{b}}{1 - a(Z_{\alpha/2} + \hat{b})} \right) \text{ ve } \alpha_{üst} = \phi \left( \hat{b} - \frac{Z_{1-\alpha/2} + \hat{b}}{1 - a(Z_{1-\alpha/2} + \hat{b})} \right) \text{ hesaplanır.}$$

**Adım 6:** Küçükten büyüğe sıralanan bootstrap tahminlerinden  $(\alpha_{alt} \times B)$ . sıradaki değer

$\hat{\theta}_{alt}^*$ ,  $(\alpha_{üst} \times B)$ . sıradaki değer  $\hat{\theta}_{üst}^*$  olarak belirlenir.

Son durumda  $(1-\alpha)$  düzeyinde hesaplanan BCa güven aralıkları

$$P \left[ (\hat{\theta}_{alt}^*) < \theta < (\hat{\theta}_{üst}^*) \right] = 1 - \alpha \text{ biçiminde olur.}$$

Bu çalışmada, doğrusal olmayan model parametrelerinin aralık tahmininde Bootstrap BCa güven aralığı yaklaşımı kullanılacaktır.

## **5. MİKROORGANİZMALARIN BİYOKİMYASAL OKSİJEN İHTİYACINA İLİŞKİN DOĞRUSAL OLMAYAN REGRESYON MODELİNDE PARAMETRELERİN NOKTA VE ARALIK TAHMİNLERİNİN ELDE EDİLMESİ**

Çalışmanın bu bölümünde, literatürde tanımlı, farklı bilimsel disiplinlerde yaygın olarak kullanılan parametrelerine göre doğrusal olmayan bir regresyon modelinde model parametrelerinin nokta ve aralık tahminleri elde edilmiştir. Çalışmadaki uygulama örneği, mikroorganizmaların biyokimyasal oksijen ihtiyacını (biochemical oxygen demand - BOD) belirlemek üzere nehir suyundan alınmış bir veri setine uygulanmış negatif-üstel modeldir. Model parametrelerin nokta tahmini, EKK yaklaşımına göre oluşturulmuş amaç fonksiyonunun GA ve GANMS algoritmaları ile optimizasyonu sonucunda belirlenmiştir. GA ayarlanabilir parametrelerinin belirlenmesinde Taguchi deney tasarımı tablolarından faydalanılmıştır.

GA ve GANMS algoritmaları kullanılarak, hata kareler toplamlarını en küçük yapan model parametrelerinin nokta tahminleri bulunduktan sonra bu tahminler, varyans büyüklükleri ve hata kareler toplamı değerleri bakımından karşılaştırılmıştır.

Model parametrelerin aralık tahmininin bulunmasında Bootstrap yöntemi kullanılmıştır. GA ve GANMS algoritmaları ile elde edilen nokta tahminlerinin Bootstrap yöntemiyle bulunan güven sınırları içinde kalıp kalmadıkları incelenmiştir. Çalışmada, analiz işlemlerinin yapılmasında Matlab R2013a programından yararlanılmıştır.

Mikroorganizmaların biyokimyasal oksijen ihtiyacının modellenmesi amaçlanan, Bates ve Watts (1988) tarafından yayınlanan çalışmadan alınan veri setinde, BOD modelini belirlemek üzere nehir suyundan örnek alınmıştır. Suyu çözünmüş organik madde, inorganik besinler ve çözünmüş oksijen ilave edilmiştir. Su, şişelere bölüştürülmüştür. Her bir şişeye sabit sıcaklıkta ve eşit miktarda mikroorganizma kültürü enjekte edilmiş ve şişelerin ağzı kapatılmıştır. Daha sonra, şişeler günlere bağlı olarak sırasıyla açılmış ve şişelerdeki suyun içinde çözünmüş oksijen miktarları ölçülmüştür. Ölçümlere göre BOD verileri çizelge 5.1’de verilmiştir.

Çizelge 5.1 Mikroorganizmaların zamana göre biyokimyasal oksijen ihtiyacı (Bates ve Watts 1988)

$t$ (gün)	1	2	3	4	5	7
<b>BOD</b> (mg/l)	8.3	10.3	19.0	16.0	15.6	19.8

Bates ve Watts (1988) tarafından yapılan çalışmada biyokimyasal oksijen ihtiyacı verisine uygun model, negatif-üstel matematiksel yapısında

$$Y = \theta_1 (1 - e^{-\theta_2 t}) + \varepsilon \quad (5.1)$$

biçiminde modellenmiştir. Buna göre modelin hata kareler toplamını veren amaç fonksiyonu,

$$\phi(\theta) = \sum_{i=1}^6 (Y_i - \theta_1 (1 - e^{-\theta_2 t}))^2 \quad (5.2)$$

biçimindedir. Burada,  $\theta = [\theta_1 \ \theta_2]$  'dir.

Bates ve Watts (1988) çalışmasında, eşitlik (5.2) ile tanımlı amaç fonksiyonu Gauss-Newton algoritmasıyla minimize edilmiştir. Çalışmalarının sonucunda, parametre tahminlerini ( $\hat{\theta} = [\hat{\theta}_1 \ \hat{\theta}_2]$ ) ve amaç fonksiyonu değerini ( $\phi$ ) ve çizelge 5.2'deki gibi bulmuşlardır.

Çizelge 5.2 Bates ve Watts (1988) tarafından yapılan çalışmada bulunan modelin hata kareler toplamı ve parametre tahmin değerleri

<b>Hata kareler toplamı (<math>\phi</math>)</b>	<b>Parametre Tahminleri</b>	
	$\hat{\theta}_1$	$\hat{\theta}_2$
25.99027	19.1426	0.5311



Bu çalışmada,  $\phi$  amaç fonksiyonunu minimize eden parametre tahmin değerleri, GA ve GANMS hibrit algoritması ile bulunmuştur.

Öncelikli olarak GA ile yapılan model parametrelerinin nokta tahminlerinin elde edilmesi için GA'nın ayarlanabilir parametrelerinin belirlenmesinde Taguchi deney tasarımı tablolarından yararlanılmıştır. Bu çalışmada, GA'nın ayarlanabilir beş parametresi için beşer seviye belirlenmiştir. Çizelge 5.3'te, GA'nın ayarlanabilir parametreleri ve bu parametrelerin uygun görülen seviyeleri tanımlanmıştır.

Çizelge 5.3 Uygulamada kullanılacak GA ayarlanabilir parametreleri ve seviyeleri

ETKENLER		SEVİYELER				
		1	2	3	4	5
<b>A</b>	<b>Popülasyon Büyüklüğü</b>	30	50	100	150	200
<b>B</b>	<b>Kuşak Sayısı</b>	50	75	100	200	500
<b>C</b>	<b>Seçim Yöntemi</b>	Stokastik uniform	Remainder	Uniform	Rulet tekerleği	Turnuva
<b>D</b>	<b>Çaprazlama Yöntemi</b>	Scattered	İki nokta	Tek nokta	Heuristic	Aritmetik
<b>E</b>	<b>Mutasyon Yöntemi</b>	Uniform 0.2	Uniform 0.01	Uniform 0.05	Uniform 0.1	Adaptive feasible

Beş parametrelili ve beş düzeyli duruma göre Taguchi'nin  $L_{25}$  tablosunun kullanılması gerekir. Taguchi'nin  $L_{25}$  tablosuna göre ayarlanabilir parametre denemelerinin kombinasyonları çizelge 3.5'te tanımlanmıştı. Çizelge 3.5'teki ilk beş sütunun sıralaması dikkate alınarak denemeler oluşturulmuştur. Matlab R2013a programında 25 farklı denemenin her biri için algoritma 200 defa çalıştırılmıştır. Her bir denemede 200 yineleme için hata kareler toplamının en küçük, ortalama, medyan ve en büyük değerleri ile standart sapması hesaplanmıştır. Elde edilen sonuçlara göre hata kareler toplamının en küçük olduğu durumdaki parametre tahmini değerleri model parametrelerinin nokta tahmini olarak seçilmiştir. 25 deneme için elde edilen amaç fonksiyonu (hata kare

toplamı) sonuçları ve bu sonuçlara ilişkin istatistikler çizelge 5.4'te verilmiştir. Çizelge 5.4'teki sonuçlara göre, en küçük amaç fonksiyonu sonucunu veren denemenin 15 numaralı deneme olduğu gözlenmektedir. Deneme-15 için GA ayarlanabilir parametrelerinin seviyeleri Taguchi tasarımına göre çizelge 5.5'teki gibi belirlenmiştir.

Çizelge 5.4 GA ile 25 deneme için bulunan amaç fonksiyonu değerleri ve istatistikler

$\phi$					
Deneme	En küçük	Ortalama	Medyan	En büyük	Std. Sapma
1	1266.1069782	1268.1609067	1267.7243168	1275.3777614	1.6948250
2	1266.1809718	1271.8305209	1270.6387372	1293.6238881	4.8862309
3	1266.1666047	1270.3261400	1269.2471317	1288.8627747	3.6710961
4	25.9902673	26.9687439	26.0649720	44.1037351	2.6238390
5	26.0154535	28.4509584	27.3296316	45.0494851	2.9837203
6	1266.0196816	1267.5012760	1267.1452347	1273.4304262	1.3149505
7	25.9902674	26.0450894	25.9907407	36.6875144	0.7563178
8	1266.1611138	1267.5561497	1267.2924319	1272.1777694	1.0619193
9	1266.2624885	1270.1013345	1269.7053374	1281.3135309	2.6713834
10	1266.1006273	1268.4848709	1267.8787772	1280.2719414	2.2139618
11	1266.5785123	1273.5718813	1272.8794212	1286.0550176	4.2411381
12	1266.0437978	1267.3515693	1267.1156560	1273.5329477	1.1553424
13	1266.0015581	1266.8038999	1266.5714179	1270.5032173	0.7304958
14	25.9902744	57.6098187	26.0004321	359.7778313	86.2752688
<b>15</b>	<b>25.9902673</b>	<b>25.9904614</b>	<b>25.9902673</b>	<b>25.9983044</b>	<b>0.0010175</b>
16	25.9902694	63.1431873	26.0036731	425.3242484	108.9082730
17	1266.0071279	1266.3736030	1266.2989030	1267.7036856	0.2942195
18	25.9902673	27.4609854	26.1367039	37.1909464	2.3986946
19	1266.0686172	1267.4308291	1267.0441337	1272.2281456	1.2543129
20	1266.0460691	1266.8802603	1266.6852634	1268.9876794	0.6845241
21	1211.4667096	1232.8044643	1233.3594158	1252.2715201	8.3630561
22	1266.0348819	1266.7298825	1266.5747933	1269.9986701	0.6045408
23	25.9902675	58.6575752	25.9915420	310.2354825	86.7679697
24	1266.0026506	1266.4471295	1266.3649733	1268.2568688	0.3673478
25	1266.0278769	1267.0354283	1266.7937059	1270.8183973	0.8838733

Çizelge 5.4'e göre denemelerin büyük bir kısmında amaç fonksiyonu değerlerinin çok yüksek çıktığı gözlenmiştir. Ancak bazı denemelerde (4., 5., 7., 14., 15., 16., 18. ve 23. denemeler),  $\phi$  fonksiyon değerinin Bates ve Watts (1988) tarafından bulunan fonksiyon değerlerine yakın sonuçlar elde edilmiştir. Yineleme değerine ilişkin istatistikler incelendiğinde söz konusu 14., 16. ve 23. denemelerin standart sapma değerlerinin oldukça büyük olduğu söylenebilir. Bu durum, GA ayarlanabilir parametrelerinin model parametrelerinin nokta tahmini sonuçları üzerindeki etkisini ortaya çıkarmaktadır.

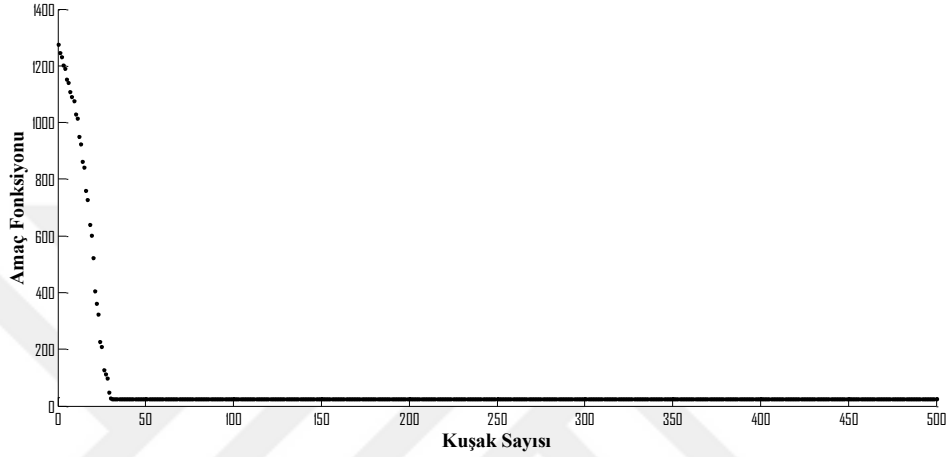
Çizelge 5.5 Ayarlanabilir GA parametrelerinin uygulama için optimal seviyeleri

GA ayarlanabilir parametreleri	Seviyeler
Popülasyon büyüklüğü	100
Kuşak sayısı	500
Seçim yöntemi	Remainder
Çaprazlama yöntemi	Heuristic
Mutasyon yöntemi	Uniform 0.2

Deneme-15 için yapılan yinelemelerde amaç fonksiyonunun en küçük, ortalama, medyan ve en büyük sonuçları diğer denemelere göre daha düşük çıkmıştır. Ayrıca, bu denemenin standart sapma değeri diğer denemelere göre dikkat çekici derecede düşük bulunmuştur. Bu durum algoritmanın, deneme-15 kombinasyonu için 200 yinelemede birbirine çok yakın sonuçlar bulduğu anlamına gelmektedir. Buna göre, deneme-15 için algoritma defalarca yinelense bile, bulunacak parametre tahminleri birbirine çok yakın değerler olarak elde edilecektir.

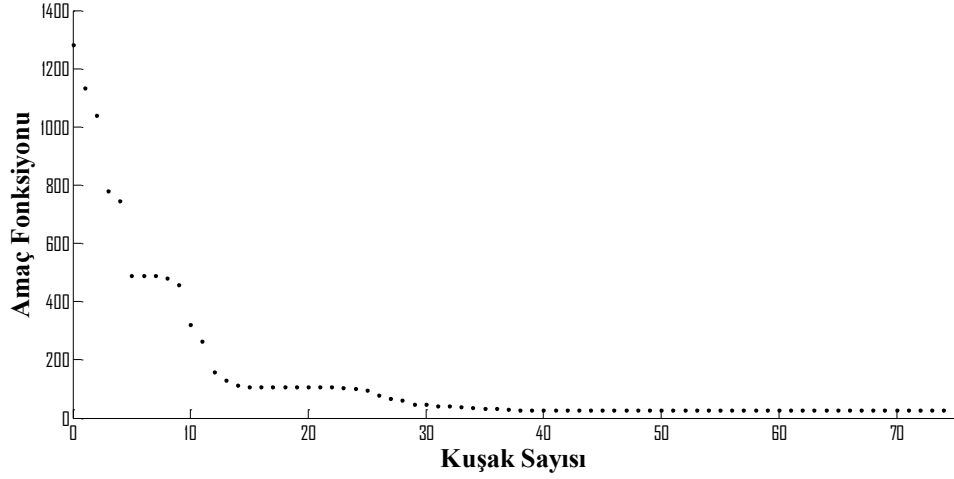
25 deneme içinde en iyi çözüm olarak bulunan deneme-15 için tanımlanan kuşak sayısı (500), Taguchi deney tasarımında belirlenen kuşak sayısı seviyeleri içinde en yüksek olanıdır. Bu nedenle, deneme-15'in, Matlab R2013a programında işlem süresi en uzun olan denemelerden biri olduğu görülmüştür. Kuşak sayısının gereğinden büyük seçilmesi hesaplama süresinin uzamasına neden olacaktır. GA'nın bir yinelemede kaçınıcı kuşaktan itibaren optimum değere yaklaştığı bilgisi, GA'nın optimal sonuca ne kadar çabuk ulaşabildiğinin bir yanıtıdır. En iyi çözüme yakınsama sağlandıktan sonraki

kuşaklar, denemelerin büyük çoğunluğunda, fazladan işlem yükü olabilmektedir. 25 deneme için bir yinelemede kaçınıcı kuşaktan itibaren optimum değere yakınsama sağlandığı çalışmada araştırılmıştır. Buna göre deneme-15 için bir yinelemede 500 kuşak boyunca her bir kuşakta bulunan amaç fonksiyonu değerleri şekil 5.1’de verilmiştir.



Şekil 5.1 Deneme-15’te bir yineleme için 500 kuşakta hesaplanan amaç fonksiyonu değerleri

Şekil 5.1’de görüldüğü üzere kuşak sayısı yaklaşık olarak 30’a ulaştığında amaç fonksiyonu en iyi değerine yakınsamaktadır. 25 deneme içinde deneme-15 dışında amaç fonksiyonuna göre deneme-15’e yakın sonuçlar bulunan (4., 5., 7., 14., 16., 18. ve 23.) denemeler de ayrı ayrı incelendiğinde genel olarak kuşak sayısının 30-50 arasında bir değere ulaştığında amaç fonksiyonunun optimum değere yaklaştığı gözlenmiştir. Bu denemeler arasında sadece 75 kuşak içeren deneme-7, zaman tasarrufu sağlaması ve deneme-15’e yakın sonuçlar bulması sebebiyle önerilebilecek bir diğer deneme olarak ön plana çıkmaktadır. Deneme-7 için bir yinelemede 75 kuşak boyunca her bir kuşakta bulunan amaç fonksiyonu değerleri şekil 5.2’de verilmiştir.



Şekil 5.2 Deneme-7’de bir yineleme için 75 kuşakta hesaplanan amaç fonksiyonu değerleri

Şekil 5.2 incelendiğinde, amaç fonksiyonu değerine göre 30.-50. kuşak aralığında deneme-7’nin en iyi değere (25.99) yaklaştığı görülmektedir. Deneme-7’ye özgü GA ayarlanabilir parametre değeri, 50 birimlik popülasyon, 75 kuşak sayısı, Uniform seçim, Heuristic çaprazlama ve Adaptive feasible mutasyon olarak tanımlıdır. Buna göre, GA kullanılarak yapılan aramalarda deneme-15’in en iyi amaç fonksiyon değerini veren deneme olduğu ancak hesaplama süresi dikkate alındığında, deneme-7’de belirlenen GA ayarlanabilir parametre kombinasyonunun, çizelge 5.5’te verilen 100 birimlik popülasyon ve 500 kuşak sayısına göre tercih edilebilen bir kombinasyon olduğu düşünülebilir.

Elde edilen tahmin değerlerinin Bates ve Watts (1988) tarafından bulunan sonuçlarla aynı olduğu görülmüştür. Çalışmada, GA ile elde edilen 25 farklı parametre tahmini, NMS algoritması için başlangıç değeri olarak kabul edilip, GANMS hibrit algoritması ile hesaplama yapılmıştır. Çizelge 5.6’da, GANMS hibrit algoritmasında, NMS kısmı için kullanılacak parametre tahmini başlangıç değerlerine göre elde edilen GANMS hibrit algoritma sonuçlarına yer verilmiştir. Çizelge 5.6’dan anlaşılacağı üzere GANMS ile elde edilen amaç fonksiyonu değerlerinin tamamı, GA yöntemiyle bulunan optimal noktaya ulaşmış veya bu noktaya çok yakın sonuç vermiştir. GANMS, amaç fonksiyonu değeri GA ile yüksek hesaplanan denemeleri bile optimal noktaya ulaştırmıştır.

GANMS yöntemiyle bulunan model parametre tahminleri ve bu tahminlere göre elde edilen  $\phi$  amaç fonksiyonu değerleri, çizelge 5.7’de verilmiştir.

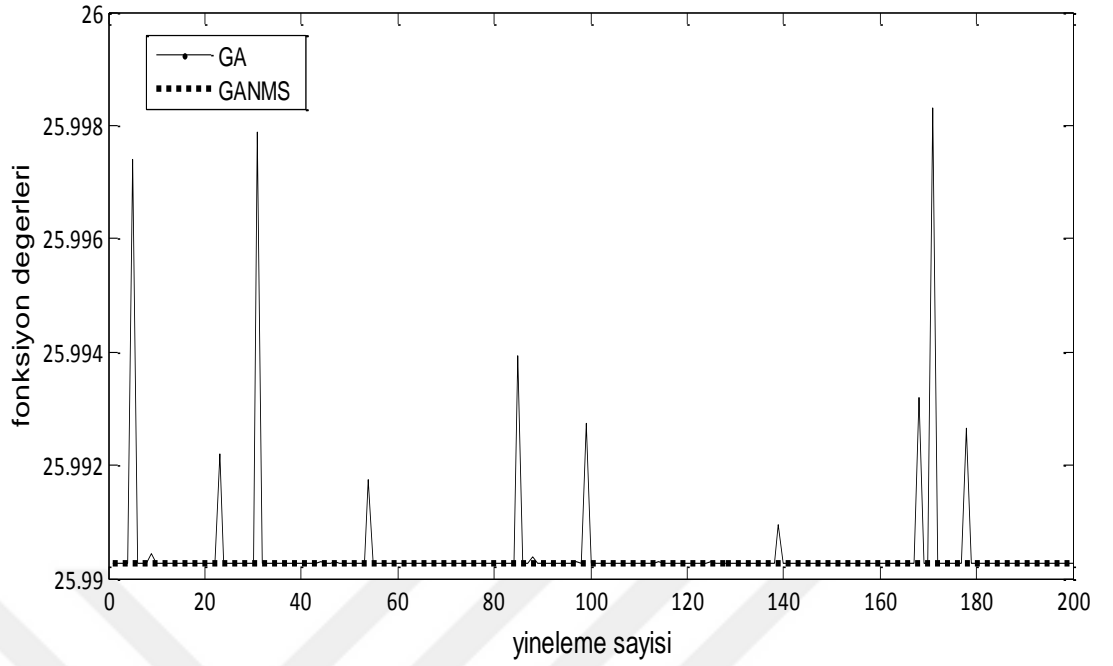
Çizelge 5.6 Parametre başlangıç değerleri ve GANMS ile 25 deneme için bulunan amaç fonksiyonu değerlerinin istatistiği

Deneme	Parametre başlangıç değerleri		$\phi$				
	$\hat{\theta}_1$	$\hat{\theta}_2$	Minimum	Ortalama	Medyan	Maksimum	Std. Sapma
1	0.999577	0.9977956	25.990267	25.990267	25.990267	25.99027	0.000000003
2	0.998979	0.9987884	25.990267	25.990267	25.99027	25.99027	0.000000002
3	0.999441	0.9958488	25.990267	25.990267	25.99027	25.99027	0.000000002
4	19.14263	0.531078	25.990267	25.990267	25.99027	25.99027	0.000000004
5	18.99722	0.5437043	25.990267	25.990267	25.99027	25.99027	0.000000002
6	0.999968	0.9991326	25.990267	25.990267	25.99027	25.99027	0.000000002
7	19.14291	0.5310688	25.990267	25.990267	25.99027	25.99027	0.000000002
8	0.999733	0.9938237	25.990267	25.990267	25.99027	25.99027	0.000000002
9	0.999462	0.9915422	25.990267	25.990267	25.99027	25.99027	0.000000002
10	0.999519	0.9985803	25.990267	25.990267	25.99027	25.99027	0.000000002
11	0.996438	0.9987802	25.990267	25.990267	25.99027	25.99027	0.000000002
12	0.999842	0.9989053	25.990267	25.990267	25.99027	25.99027	0.000000003
13	0.999979	0.9999672	25.990267	25.990267	25.99027	25.99027	0.000000002
14	19.14429	0.5310564	25.990267	39.256637	25.99027	136.7617	36.01567333
15	19.14258	0.5310914	25.990267	25.990267	25.99027	25.99027	0.000000002
16	19.14142	0.5311351	25.990267	37.607966	25.99027	136.9752	34.00365989
17	0.999944	0.9999637	25.990267	25.990267	25.99027	25.99027	0.000000002
18	19.14258	0.5310914	25.990267	25.990267	25.99027	25.99027	0.000000002
19	0.999558	0.9999052	25.990267	25.990267	25.99027	25.99027	0.000000002
20	0.999929	0.998097	25.990267	25.990267	25.99027	25.99027	0.000000002
21	1.32985	1.1665875	25.990267	25.990267	25.99027	25.99027	0.000000002
22	0.999987	0.9982023	25.990267	25.990267	25.99027	25.99027	0.000000002
23	19.14281	0.5310916	25.990267	39.808805	25.99027	136.8726	36.65218492
24	0.999982	0.9998889	25.990267	25.990267	25.99027	25.99027	0.000000004
25	0.999971	0.9986856	25.990267	25.990267	25.99027	25.99027	0.000000003

Çizelge 5.7 GANMS ile 25 deneme için bulunan parametre tahminleri

Deneme	$\hat{\theta}_1$	$\hat{\theta}_2$
1	19.14257068	0.53109158
2	19.14257762	0.53109068
3	19.14256286	0.53109214
4	19.14566667	0.53109158
5	19.14257612	0.53109064
6	19.14258065	0.53109128
7	19.14258700	0.53109057
8	19.14258043	0.53109109
9	19.14258821	0.53109053
10	19.14257435	0.53109209
11	19.14256763	0.53109231
12	19.14258885	0.53109048
13	19.14258499	0.53109071
14	19.14256289	0.53109245
15	19.14257521	0.53109138
16	19.14258394	0.53109058
17	19.14259095	0.53109000
18	19.14257698	0.53109142
19	19.14257651	0.53109075
20	19.14256221	0.53109212
21	19.14256472	0.53109240
22	19.14258272	0.53109107
23	19.14258561	0.53109027
24	19.14256771	0.53109166
25	19.14256848	0.53109127

GANMS ile 25 denemenin tamamında parametre tahminleri birbirine çok yakın değerler olarak hesaplanmıştır. Bu değerler, GA'nın deneme-15'te hesapladığı değerlerle benzer veya o değerlere çok yakındır. GA ile GANMS'nin deneme-15'te 200 yineleme için hesapladıkları amaç fonksiyonu değerleri şekil 5.3'te görülmektedir. Buna göre, GANMS ile her yinelemede GA'dan daha tutarlı sonuçlara ulaşıldığı söylenebilir.



Şekil 5.3 Deneme-15 için uygulanan 200 yinelemeye göre GA ve GANMS ile bulunan hata kareler toplamlarının karşılaştırılması

GA ve GANMS yöntemi sonuçlarıyla Bates ve Watts (1988) tarafından elde edilen sonuçlar çizelge 5.8’de birlikte gösterilmiştir.

Çizelge 5.8 BOD veri seti için Gauss-Newton, GA ve GANMS yöntemleriyle bulunan hata kareler toplamı ve parametre tahmin değerleri

Yöntem	Amaç fonksiyonu	Parametre tahminleri	
	Hata kareler toplamı ( $\phi$ )	$\hat{\theta}_1$	$\hat{\theta}_2$
Gauss-Newton Yöntemi (Bates ve Watts (1988))	25.99027	19.1426	0.5311
GA	25.990267	19.14257521	0.53109138
GANMS	25.990267	19.14257521	0.53109138

Çizelge 5.8 incelendiğinde, BOD veri setine uygun doğrusal olmayan modelin belirlenmesinde, GA’nın deneme-15 değeri ve GANMS ile bulunan bütün sonuçların Bates ve Watts (1988) tarafından bulunan sonuçlarla aynı olduğu görülmektedir.



Model parametrelerinin güven aralıklarının hesaplanmasında Bootstrap BCa yöntemi kullanılmıştır. Veri setinden 2000 adet bootstrap örnek kümesi oluşturulmuş ve her bir örnek kümesi için parametre tahmin değerleri hesaplanmıştır. Hesaplanan parametre tahmin değerleri küçükten büyüğe sıralanmıştır. %95 güven düzeyinde hesaplanan güven aralığı alt ve üst sınırlara çizelge 5.9’da yer verilmiştir.

Çizelge 5.9 BOD verisi için parametrelerin Bootstrap BCa güven aralıkları

Parametreler	Alt sınır	Üst Sınır
$\theta_1$	16.3105	23.573
$\theta_2$	0.223	116.5347

Buna göre, oldukça küçük örnek hacmine sahip ( $n=6$ ) BOD veri setine ilişkin doğrusal olmayan model parametrelerinin aralık tahmini için elde edilen değerlerin %95 güven düzeyinde parametre tahminlerini içerdiği söylenebilir.

## 6. TARTIŞMA VE SONUÇ

Doğrusal olmayan regresyon modellerinin parametre tahminlerinin hesaplanması doğrusal regresyon modellerinin parametre tahminlerinin hesaplanması kadar kolay değildir. Doğrusal regresyon modellerinde parametre tahminleri hesaplanırken EKK yaklaşımı kullanılmaktadır. Bu yaklaşıma göre modelin hata kareler toplamını minimum yapan parametre tahminleri parametrelere göre kısmi türev alma vasıtasıyla hesaplanabilir. Doğrusal olmayan regresyon modellerinde ise çoğu zaman hata kareler toplamının parametrelere göre kısmi türevi alınarak sonuç elde edilemez.

Doğrusal olmayan regresyon modellerinde parametre tahmini yapılırken karşılaşılan zorluklara karşı çeşitli çözüm yolları geliştirilmiştir. Bu çözüm yollarından ilki modele doğrusal dönüşüm uygulamaktır. Bu yöntemde eğer model üstel yapıdaysa modele logaritmik dönüşüm uygulanmakta ve model doğrusal yapıya dönüştürülmektedir. Bu sayede modelin parametreleri, doğrusal regresyon modellerinde olduğu gibi, kısmi türev vasıtasıyla tahmin edilebilmektedir. Ancak, bu yöntem ile modelin yanı sıra hataların da dönüşüme uğradığı için hataları en küçükleme mantığına dayanan EKK yaklaşımı sonucunda yanıltıcı sonuçlar elde edilebilmektedir.

Doğrusal olmayan regresyon modellerinin parametrelerinin nokta tahmininde yaygın olarak kullanılan bir diğer yaklaşım, türeve dayalı algoritmaları kullanmaktır. Gauss-Newton Algoritması, En Dik İniş Algoritması ve L-M Algoritması yaygın olarak kullanılan türeve dayalı algoritmalarıdır. Bu algoritmalar yaygın olarak kullanılmalarına rağmen türeve dayalı çözüm bulmaları sebebiyle türev almanın zor veya imkansız olduğu modellerde zorlanabilmektedir. Ayrıca bu algoritmalar, türevsel yaklaşım ile çözüm aradıklarından yakınsayamama veya lokal minimum tuzaklarına yakalanma gibi istenmeyen durumlarla karşılaşabilmektedir.

Çalışmada, doğrusal olmayan regresyon modellerinin çözümünde doğrusal dönüşüm uygulamalarına ve türeve dayalı algoritmalara alternatif olarak türevden bağımsız algoritmalar önerilmektedir. Bu algoritmalar genellikle stokastik çözümler üretmeleri ve türeve dayalı yöntemlere göre daha esnek yapıda olmaları sebebiyle tercih edilmektedir.

Çalışmada türevden bağımsız GA ve NMS algoritmaları hakkında bilgi verilmiştir. Bu algoritmaların avantajlı ve dezavantajlı yönleri üzerinde durulmuştur. Buna göre NMS başlangıç değeri bilgisine ihtiyaç duyan, bölgesel ve doğrudan arama yapan bir algoritmadır. Algoritma yapısının kolay olması sebebiyle pratik ve etkili sonuçlar bulmasına rağmen modelde tahmin edilecek parametre sayısı arttıkça yakınsama kabiliyeti azalır.

GA, NMS'nin aksine global ve çoklu arama yapan bir yöntemdir. Bu, optimuma ulaşmak için bütün arama uzayında aynı anda pek çok noktadan arama işlemine başladığı anlamına gelmektedir. Bu sebeple başlangıç noktası bilgisine ihtiyaç duymaz. Ayrıca, parametrelerin kendisi yerine kodlanmış kümeleriyle çalıştığından sürekli ve kesikli değişkenlerde çalışabilir. NMS'nin aksine sonuç bulurken tahmin edilecek parametre sayısından etkilenmez. Ancak bu avantajlı özelliklerine rağmen dezavantajlı özellikleri de mevcuttur. Algoritmanın iyi sonuç vermesi GA ayarlanabilir parametrelerin iyi belirlenmesine bağlıdır. Ayarlanabilir GA parametrelerinin pek çok farklı seviyesi bulunur. Bu seviyelerin de oluşturduğu pek çok farklı kombinasyon da mevcuttur.

GA ve NMS'nin diğer algoritmalara göre avantajlı olduğu yönlerinin öne çıkarılması ve dezavantajlı olduğu özelliklerinin etkisini minimuma indirilmesi için çalışmada, GA ve NMS algoritmalarının bir arada kullanıldığı hibrit bir algoritma, doğrusal olmayan regresyon parametrelerinin tahminlerinin hesaplanmasında önerilmektedir. GANMS adı verilen bu algoritmaya göre parametre tahminine başlangıç değeri bilgisine ihtiyaç duymadan global ve çoklu arama yaparak GA ile başlanmakta ve bu algoritma ile öncül tahminler elde edilmektedir. Elde edilen öncül sonuçlar, NMS algoritması için başlangıç değeri olarak kabul edilerek bu değerlere NMS operasyonları uygulanmaktadır. NMS operasyonları sonucunda parametre tahmin değerleri hesaplanmaktadır.

Bu çalışmada, GA ve GANMS algoritmaları kullanılarak doğrusal olmayan regresyon modellerinin EKK yaklaşımı ile parametre tahminlerinin bulunması, bulunan sonuçların birbirleriyle ve türeve dayalı bir algoritma olan Gauss-Newton algoritması sonuçlarıyla hata kareler toplamı üzerinden karşılaştırılması amaçlanmıştır. Ayrıca, bulunan nokta

tahminlerinin Bootstrap BCa yöntemiyle güven aralığı tahminleri yapılmıştır. Nokta tahmin değerlerinin bulunan bootstrap güven aralıkları içinde olup olmadığı incelenmiştir.

Uygulama bölümünde Bates ve Watts (1988) tarafından yapılan çalışmadan alınan doğrusal olmayan regresyon modelinin, GA ve GANMS algoritmaları kullanılarak hata kareler toplamı minimize edilmiştir. Hata kareler toplamını minimum yapan model parametre değerleri hesaplanmıştır. Uygulama bölümünde yapılan bütün işlemlerde Matlab R2013a paket programı kullanılmıştır.

Doğrusal olmayan model parametrelerinin nokta tahmini için GA ve GANMS yöntemleri kullanılırken, GA'nın en temel dezavantajlarından bir tanesi olan GA ayarlanabilir parametrelerinin ve seviyelerinin belirlenmesi işleminde  $L_{25}$  Taguchi deney tasarımı tablosundan yararlanılmıştır. Taguchi deney tasarımı tablosu ile farklı GA ayarlanabilir parametreleri kullanılarak yapılabilecek yüzlerce seviye kombinasyonu içinden etkili olabilecek kombinasyonlar belli bir sistematikte seçilmiştir. Böylece, GA ayarlanabilir parametrelerinin istenen seviye kombinasyonlarının belirlenmesinde zaman ve emek tasarrufu sağlanmıştır.

GA ile Taguchi deney tasarımı sistematığına göre yapılan 25 farklı deneme ile birbirinden farklı sonuçlara ulaşılmıştır. Bu sonuçlara göre sadece {4, 5, 7, 14, 15, 16, 18, 23} numaralı denemelerin sonuçları hata kareler toplamı bakımından makul seviyelerde bulunmuş diğer denemelerin hata kareler toplamı yüksek çıkmıştır. Bu durum, GA ile bulunan sonuçların başarısının modele uygun GA ayarlanabilir parametrelerinin seçimine bağlı olduğunu göstermektedir. Buna göre GA ile başarılı tahminler yapılabilmesi için GA ayarlanabilir parametrelerinin isabetli seçilmesi gerekmektedir. Taguchi deney tasarım tabloları kullanılarak, modele uygun GA ayarlanabilir parametre ve seviyelerinin belirlenebildiği ve bunun sonucunda model parametresine daha yakın tahminlerin yapılabildiği sonucuna varılmıştır.

GA ile bulunan parametre tahmin sonuçları başlangıç değeri olarak kabul edilerek optimizasyon işlemine NMS ile devam edilmesiyle GANMS sonuçlarına ulaşılmıştır. GANMS ile elde edilen hata kareler toplamı değerlerinin tamamı, GA ile bulunan en iyi hata kareler toplamı değerlerine ulaşmış veya bu değere çok yakın sonuç vermiştir. GANMS yöntemi ile yapılan denemelerin hepsinde, GA'nın aksine, birbirine yakın ve optimal sonuçlar bulunmuştur. Bunun sebebinin, GANMS'nin GA ve NMS yöntemlerinin avantajlı özelliklerini bir arada kullanarak GA'nın ayarlanabilir parametreler konusundaki dezavantajını ortadan kaldırması olduğu düşünülmektedir. Ayrıca, GANMS'nin GA ayarlanabilir parametrelerinden GA'ya oranla daha az etkilendiği gözlenmiştir. Böylece, GANMS ile yapılan nokta tahminlerinde, GA'nın aksine, deney tasarım tabloları kullanmadan da parametre tahmininde başarılı sonuçlar elde edilebileceği sonucuna varılmıştır.

Yapılan uygulama sonucunda GA ve GANMS algoritmalarıyla bulunan parametrelerin nokta tahmin değerleri hata kareler toplamı değerlerine bakılarak Bates ve Watts (1988) tarafından Gauss-Newton algoritmasıyla bulunan sonuçlarla karşılaştırılmıştır. Yapılan karşılaştırma çizelge 5.8'de verilmiştir. Buna göre, GA ve GANMS ile bulunan hata kareler toplamı değerleri ve hesaplanan parametre tahmin değerleri Gauss-Newton yöntemiyle bulunan sonuçlarla aynı çıkmıştır. Bu sebeple, doğrusal olmayan regresyon model parametrelerinin nokta tahminlerinin bulunmasında türevden bağımsız GA ve GANMS algoritmaları, Gauss-Newton gibi türeve dayalı algoritmalara alternatif olarak önerilmektedir.

Model parametrelerinin aralık tahminleri, Bootstrap BCa yöntemiyle yapılmıştır. Nokta tahmin değerlerinin bulunan bootstrap güven aralıkları içinde olup olmadığı incelenmiştir. Bu inceleme sonucunda, nokta tahmin değerlerinin güven aralığı alt ve üst sınırları arasında kaldığı gözlenmiştir. Ancak, hesaplanan alt ve üst sınırlar arasındaki uzunluk değeri büyük çıktığı gözlenmiştir. Bu durumun rasgele seçilen farklı Bootstrap örneklerinden kaynaklandığı düşünülmektedir. Sonraki çalışmalarda parametrelerin aralık tahmininin hesaplanmasında alt ve üst sınırlar arasındaki mesafeleri azaltacak alternatif yöntemlerin belirlenmesi üzerinde durulması düşünülmektedir.

## KAYNAKLAR

- Akyol, A.P. 2006. Doğrusal Olmayan Ekonometrik Modellerin Genetik Algoritma Yaklaşımı ile Parametre Tahmini. Yüksek Lisans Tezi. Gazi Üniversitesi, Sosyal Bilimler Enstitüsü, Ankara.
- Altunkaynak, B. ve Esin, A. 2004. Doğrusal Olmayan Regresyonda Parametre Tahmini İçin Genetik Algoritma Yöntemi. G.Ü. Fen Bilimleri Dergisi, 17(2), 43-51.
- Apaydın, A. 2005. Optimizasyon. Kılavuz Kitabevi, 388, Ankara.
- Artaç, T. 2003. Genetik Algoritma ile Dağıtım Şebekelerinin Optimum Tasarımı. Yüksek Lisans Tezi. İstanbul Teknik Üniversitesi, Fen Bilimleri Enstitüsü, İstanbul.
- Ayala, J. F. 2017. Web sitesi: <https://www.britannica.com/science/evolution-scientific-theory>. Erişim Tarihi: 11.02.2018
- Bates, D.M. and Watts, D.G. 1988. Nonlinear Regression Analysis and Its Applications, John Willey&Sons, Inc.,365, New York.
- Bello, O.A., Bamidura, T.A., Chuwkwu, U.A. and Osowole, O.I. 2015. Bootstrap Nonlinear Regression Application in a Design of an Experiment Data for Fewer Sample Size. International Journal of Research, 2(2), 428-441.
- Bolat, B., Erol, K.O. ve İmrak, C.E. 2004. Mühendislik Uygulamalarında Genetik Algoritmalar ve Operatörlerin İşlevleri. Sigma Mühendislik ve Fen Bilimleri Dergisi, 22(4), 264-271.
- Bose, R.C. and Bush, K.A. 1952. Orthogonal Arrays of Strength Two and Three. The Annals of Mathematical Statistics, 23(4), 508-524.
- Box, M. J. 1971. Bias in Nonlinear Estimation. Journal of Royal Statistic. Society Series B. 33(2), 171-201.
- Carpenter, J. and Bithell, J. 2000. Bootstrap Confidence Intervals: When, Which, What? A Practical Guide for Medical Statisticians. Statistics in Medicine, 19(9), 1141-1164.
- Chernick, M.R. 2008. Bootstrap Methods: A Guide for Practitioners and Researcers. John Willey&Sons, Inc., 369, New Jersey.
- Çelik, C. 1993. Kalite Geliştirmede Tasarım En İyileme Problemine Taguchi Yöntemlerinin Uygulanmasında Sistemik Bir Yaklaşım. Doktora Tezi. Anadolu Üniversitesi, Fen Bilimleri Enstitüsü, Eskişehir.
- Davison, A.C. and Hinkley, D.V. 1997. Bootstrap Methods and Their Application, Cambridge University Press, 582, New York.

- De Jong, K.A. 1975. An Analysis of the Behavior of a Glass of Genetic Adaptive System. Doctoral Dissertation. Michigan University, USA.
- Dervişoğlu, N. ve Muluk, Z. 2007. Taguchi Tasarımının Uygulanması ve Klasik Kesirli Çok Etkenli Tasarımla Karşılaştırılması. Anadolu Üniversitesi Bilim ve Teknoloji Dergisi, 8(1), 65-78.
- DiCiccio, T.J. and Efron, B. 1996. Bootstrap Confidence Intervals. *Statistical Science*, 11(3), 189-228.
- DiCiccio, T.J. and Tibshirani, R.J. 1987. Bootstrap Confidence Intervals and Bootstrap Approximations. *Journal of the American Statistical Association*, 82(1), 163-170.
- Efron, B. 1979. Bootstrap Methods: Another Look at the Jackknife. *The Annals of the Statistics*, 7(1), 1-26.
- Efron, B. 1982. *The Jackknife, the Bootstrap, and Other Resampling Plans (CBMS-NSF Regional Conference Series in Applied Mathematics)*, Capital City Press, 92, Philadelphia.
- Efron, B. and Tibshirani, R. J. 1986. Bootstrap Methods for Standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracy. *Statistical Science*, 1(1), 54-77.
- Efron, B. 1992. More Accurate Confidence Intervals in Exponentialfamily. *Biometrika*, 79(2), 231-245.
- Efron, B. and Tibshirani, R. J. 1993. *An Introduction to the Bootstrap*. Chapman&Hall, 436, New York.
- Fogel, L.J., Owens, A.J., and Walsh, M.J. 1966. *Artificial Intelligence through Simulated Evolution*. John Willey&Sons, Inc.,170,USA.
- Freedman, D.A. 1981. Bootstrapping Regression Models. *The Annals of the Statistics*, 9(6), 1218-1228.
- Gallant, A.R. 1975. Nonlinear Regression. *Journal of the American Statistical Association*, 29 (2), 73-81.
- Gallant, A.R. and Fuller, W.A. 1973. Fitting Segmented Polynomial Regression Models Whose Join Points Have to Be Estimated. *Journal of the American Statistical Association*, 68 (1), 144-147.
- Goldberg, D.E. 1989. *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley, 372, Boston.
- Gopi, E.S. 2007. *Algorithm Collections for Digital Signal Processing Applications Using Matlab*. Springer Netherland, Dordrecht.

- Gurson, A.P. 2000. Simplex Search Behaviour in Nonlinear Optimization. Underground Honors Thesis. College of William & Mary, USA.
- Herrera, F., Lozano, M. and Sanchez, A.M. 2003. A Taxonomy for the Crossover Operator for Real-Coded Genetic Algorithms: An Experimental Study. *International Journal of Intelligent Systems*, 18(3), 309-338.
- Holland, J.H. 1975. *Adaptation in Natural and Artificial Systems*. The University of Michigan Press, 211, Ann Arbor.
- Houkoos, J.S. and Lewis, R.J. 2005. Advanced Statistics: Bootstrapping Confidence Intervals for Statistics with “Difficult” Distributions. *Academic Emergency Medicine*, 12(4), 360-365.
- Hartley, H.O. and Booker, A. 1963. *Non-linear Least Squares Estimation*. Unpublished Report, Iowa State University.
- Jennrich, R.I. 1969. Asymptotic Properties of Non-linear Least Squares Estimators. *The Annals of Mathematical Statistics*, 40(2), 633-643.
- Kacker, R.N., Lagergren, E.S. and Filliben, J.J. 1991. Taguchi’s Orthogonal Arrays Are Classical Designs of Experiments. *Journal of Research of the National Institute of Standards and Technology*, 96(5), 577-591.
- Karakoca, A. 2009. Çok Değişkenli Lineer Olmayan Modellerde Genetik Algoritma. Doktora Tezi. Selçuk Üniversitesi, Fen Bilimleri Enstitüsü, Konya.
- Koza, J.R. 1992. *Genetic Programming: On the Programming of Computers by Means of Natural Selection*, The Mit Press, 840, London.
- Koza, J.R. 1994. *Genetic Programming II: Automatic Discovery of the Reusable Programs*, The Mit Press. 1994, 768, London.
- Malinvaud, E. 1970. Consistency of Nonlinear Regressions. *The Annals of Mathematical Statistics*, 41 (3), 956-969.
- Maringer, D. 2005. *Portfolio Management with Heuristic Optimization*. Springer, 223, Dordrecht.
- Marquardt, D.W. 1963. An Algorithm for Least-Squares Estimation of Nonlinear Parameters. *Journal of the Society of Industrial and Applied Mathematics*, 11(2), 431-441.
- Mitchell, M. 1999. *An Introduction to Genetic Algorithms*. The Mit press, 158, London.
- Muluk, Z., Saraçbaşı, T. ve Aktaş, S. 1998. Taguchi Üzerine Araştırma. DPT Proje No: 94K120340/2, Ankara.
- Nelder, J. A. and Mead, R. 1965. A Simplex Method for Function Minimization. *The Computer Journal*, 7(4), 308-313.



- Oral-Erbaş, S. 2016. Olasılık ve İstatistik Problemler ve Çözümleri İle. Gazi Kitabevi, 5. Baskı, 637, Ankara.
- Özçakar, N. 1998. Genetik Algoritmalar. İ. Ü. İşletme Fakültesi Dergisi, 27(1), 69-82.
- Öztürk, A. 2002. Gerçel Sayı Kodlamalı Genetik Algoritmaların Optimizasyonda Kullanımı. Yüksek Lisans Tezi. İstanbul Teknik Üniversitesi, Fen Bilimleri Enstitüsü, İstanbul.
- Rechenberg, I. 1965. Cybernetic Solution Path of an Experimental Problem. Ministry of Aviation, Royal Aircraft Establishment of U.K.
- Rechenberg, I. 1973. Evolutionsstrategie: Optimierung Technischer Systeme nach Prinzipien der Biologischen Evolution. Frommann-Holzboog, Stuttgart.
- Rios, L.M. and Sahinidis, N.V. 2013. Derivative-Free Optimization: a Review of Algorithms and Comparison of Software Implementations. Journal of Global Optimization, 56(3), 1247-1293.
- Seber, G.A.F. ve Wild, C.J. 1989. Nonlinear Regression. John Willey&Sons, Inc.,768, New York.
- Serin, T. 2010. Doğrusal Olmayan Regresyon Modellerinde Parametre Tahmin Yöntemleri, Öneriler ve Karşılaştırmaları. Doktora Tezi. Gazi Üniversitesi, Fen Bilimleri Enstitüsü, Ankara.
- Shao, J. ve Tu, D. (1995). The Jackknife and Bootstrap. Springer, 499, New York.
- Spendley, W., Hext, G.R. and Himsworth, F.R. 1962. Sequential Application of Simplex Designs in Optimization and Evolutionary Operation. Technometrics, 4(4), 441-461.
- Şahinbaşoğlu, Z.Z. 2005. Doğrusal Olmayan Regresyonda Bazı Eğrisellik Ölçüleri. Yüksek Lisans Tezi. Yıldız Teknik Üniversitesi, Fen Bilimleri Enstitüsü, İstanbul.
- Türkşen,Ö. and Tez, M. 2016. An Application of Nelder-Mead Heuristic-based Hybrid Algorithms: Estimation of Compartment Model Parameters. International Journal of Artificial Intelligence, 14(1), 112-129.
- Ünlü, A.R. 2006. Doğrusal Olmayan Regresyon Modelleri ve Bilgisayarlı Çözümleme. Yüksek Lisans Tezi. Marmara Üniversitesi, Fen Bilimleri Enstitüsü, İstanbul.
- Wehrens, R., Putter, H., and Buydens, L.M.C. 2000. The Bootstrap: a Tutorial. Chemometrics and Intelligent Laboratory Systems, 54(1), 35-52.
- Wu, C.F.T. 1986. Jackknife, Bootstrap and Other Resampling Methods in Regression Analysis. The Annals of the Statistics, 14(4), 1261-1295.
- Yeniay, Ö. 1999. Taguchi Deney Tasarımı Problemlerine Genetik Algoritma Yaklaşımı. Doktora Tezi. Hacettepe Üniversitesi, Fen Bilimleri Enstitüsü, Ankara.

## ÖZGEÇMİŞ

Adı Soyadı : Fikret AKGÜN

Doğum Yeri : Aydın

Doğum Tarihi : 1990

Medeni Hali : Bekar

Yabancı Dili : İngilizce

### Eğitim Durumu

Lise : Aydın Süleyman Demirel Anadolu Lisesi (2004-2008)

Lisans : Hacettepe Üniversitesi Fen Fakültesi İstatistik Bölümü (2008-2013)  
Eskişehir Anadolu Üniversitesi İşletme Fakültesi İşletme Bölümü (2011-2015)

Yüksek Lisans : Ankara Üniversitesi Fen Bilimleri Enstitüsü İstatistik Anabilim Dalı (Eylül 2013- Mayıs 2018)

### Çalıştığı Kurum/Kurumlar

T.C. Enerji Piyasaları Düzenleme Kurumu (2015- )

### Uluslararası Kongre Sunum

**Akgün, F.**, Türkşen, Ö. ve Tez, M. **2017**. Obtaining Point Estimates of Nonlinear Regression Model Parameters by Hybridization of Iterative and Direct Optimization Algorithms. 3rd International Researchcers, Statisticians and Young Statisticians Congress. Konya Selçuk Üniversitesi, Konya.

**Akgün, F.** ve Türkşen, Ö. **2017**. An Application of Parameter Estimation with Genetic Algorithm for Replicated Response Measured Nonlinear Data Set: Modified Michaelis-Menten Model. 10th International Statistics Congress. Ankara Üniversitesi, Ankara.