

## THE GENETIC ALGORITHM METHOD FOR PARAMETER ESTIMATION IN NONLINEAR REGRESSION

Bülent ALTUNKAYNAK\*, Alptekin ESİN  
Gazi Üniversitesi, Fen-Edebiyat Fakültesi, İstatistik Bölümü, 06500, Ankara, TÜRKİYE  
e-mail: bulenta@gazi.edu.tr

### ABSTRACT

In this study, an alternative method has been proposed for the parameter estimation in non-linear regression. This method is the genetic algorithms technique which is widely used in recent years. Unlike other parameter estimation methods, genetic algorithms do not require supplementary information and thus are functional for practical purposes. Genetic algorithms are shown to achieve effective solutions compared to other methods in the *S*-type growth models. The results have established that genetic algorithms with no need for auxiliary information can be used for parameter estimation in non-linear regression.

*Key Words: Nonlinear Regression, S-Type Growth Model, Genetic Algorithms*

## DOĞRUSAL OLMAYAN REGRESYONDA PARAMETRE TAHMİNİ İÇİN GENETİK ALGORİTMA YÖNTEMİ

### ÖZET

Bu çalışmada doğrusal olmayan regresyonda parametre tahmini için alternatif bir yöntem önerilmiştir. Bu yöntem son yıllarda oldukça geniş bir kullanım alanına sahip olan genetik algoritmalar yöntemidir. Genetik algoritmanın kullanımı diğer parametre tahmin yöntemleri gibi yardımcı bilgiler gerektirmez ve bu nedenle pratik uygulamalar için kullanışlıdır. Bu çalışmada genetik algoritmaların diğer yöntemler kadar etkin çözümlere ulaşabileceği *S*-biçimli büyüme modellerinde parametre tahminleri yapılarak gösterilmiştir. Sonuçlar genetik algoritmaların yardımcı bilgiler gerektirmeksizin doğrusal olmayan regresyonda parametre tahmini için kullanılabilir olduğunu göstermektedir.

*Anahtar Kelimeler: Doğrusal Olmayan Regresyon, S-Biçimli Büyüme Modeli, Genetik Algoritmalar*

### 1. GİRİŞ

Regresyon analizi, yığının ortalama değerini açıklayıcı değişkenlerin bilinen ya da değişmeyen değerleri cinsinden tahmin etme ya da kestirme amacıyla bir bağımlı değişkenin açıklayıcı değişkenlere olan bağımlılığını inceler (1). Regresyon analizi doğru değerlendirmeler için önemli bir araç olarak pek çok uygulamalı alanda kullanılan istatistiksel yöntemlerin en yaygın olanlarından biridir. Özellikle son yıllarda doğrusal olmayan regresyon analizi, pratik uygulamalar için alternatif bir araç haline gelmiştir. Çünkü, fizik, kimya, biyoloji ve mühendislik gibi pek çok alanda elde edilen modellerin hemen hemen hepsi doğrusal değildir. Doğrusal olmayan regresyon modellerindeki model parametreleri doğrusal olmayan yapıdadır ve sıradan küçük kareler yöntemiyle tahmin edilemezler. Bu nedenle doğrusal olmayan regresyon modelinde parametre tahminlerine ilişkin bir çok algoritma geliştirilmiştir. Bu algoritmalarından bazıları Gauss-Newton, Direkt Arama

### 1. INTRODUCTION

Regression analysis examines the dependency of a dependent variable on descriptive variables with the aim of estimating the mean value of a population in the type of a constant or known value of descriptive variables (1). Regression analysis is one of the most commonly used statistical methods employed in various applied fields for the purpose of sound evaluation. Specifically, in its new form developed in the recent years, non-linear regression analysis provides an alternative means for practical applications due to the fact that almost all of the models formulated in many fields like physics, chemistry, biology and engineering are of non-linear character. Model parameters in non-linear regression models are of non-linear nature and they cannot be estimated through ordinary least squares method. Hence, in non-linear regression model various algorithms as to parameter estimation have been developed. Some of these algorithms are Gauss-Newton, Direct Search Method, Hooke-Jeeves

Yöntemi, Hooke-Jeeves Yöntemi, Nelder-Mead Yöntemi, Gradyant Yöntemi, Değişken Ölçüm Yöntemidir (2). Bu algoritmalar genellikle belirli bir problem için uygundur ve süreklilik, türevlerin varlığı, tek modluluk ve diğer konulara ilişkin kısıtlayıcı varsayımlar gerektirebilir.

Örneğin, yaygın olarak kullanılan algoritmalarından birisi olan Gauss-Newton yönteminin uygun olarak kullanılabilmesi için çok fazla yardımcı bilgi gerekir. Dahası bu algoritmanın kullanımı büyük bir dikkat ister çünkü algoritma için başlangıç noktası iyi seçilmezse yerel optimumlara takılabilir ve sadece bölgesel optimumları sağlar. Bu durumda uygun olmayan öncül bir alan, araştırmanın yönünü değiştirebilir ve optimal sonucun bulunmasını geciktirebilir.

Bu çalışmada ise, yardımcı bilgilere gerek duymayan ve optimizasyon problemleri için önemli bir araç olmaya başlayan Genetik Algoritmalar (GA) doğrusal olmayan regresyon modelinde parametreleri tahmin etmek için önerilmiştir. Çalışmanın ikinci bölümünde parametre tahmini için gerekli olan bir fonksiyon optimizasyon probleminin amaç fonksiyonu olarak tanımlanmıştır. Üçüncü bölümde ise bu çalışmada önerilen genetik algoritma yöntemi tanıtılmıştır. Dördüncü bölüm Gauss-Newton yöntemi ile Genetik Algoritma yöntemine ilişkin deneysel sonuçları içermektedir. Bu bölümde karşılaştırma için sadece Gauss-Newton yönteminin ele alınmasının nedeni en yaygın kullanılan yöntem olması ve yukarıda sayılan diğer yöntemlerle (deterministik yöntemler) aynı türden olmasıdır. Son bölümde ise elde edilen sayısal sonuçlar tartışılmıştır.

## 2. PARAMETRE TAHMİNİ

Tipik bir doğrusal olmayan regresyon modeli aşağıdaki gibidir.

$$y = f(x, \theta) + e \quad [1]$$

burada  $e \sim N(0, \sigma^2)$ ,  $x \in R^m$ ,  $y \in R^1$  ve  $f(x, \theta)$ ,  $\theta \in R^p$ 'nin doğrusal olmayan bir fonksiyonudur. Doğrusal olmayan regresyona örnek olarak aşağıdaki S-biçimli büyüme modelleri verilebilir (3).

**Table 1.** S-type growth models

**Çizelge 1.** S-biçimli büyüme modelleri

The functions, the parameter number of which are <b>three</b> / Bilinmeyen yığın parametresi sayısı <b>üç</b> olan fonksiyonlar		The functions, the parameter number of which are <b>four</b> / Bilinmeyen yığın parametresi sayısı <b>dört</b> olan fonksiyonlar	
<i>Gompertz</i>	$f(x) = \alpha \exp(-\exp(\beta - \gamma x))$	<b>Richards</b>	$f(x) = \alpha / (1 + \exp(\beta - \gamma x))^{1/\delta}$
<i>Logistic</i>	$f(x) = \alpha / (1 + \exp(\beta - \gamma x))$	<i>MMF</i>	$f(x) = (\beta \gamma + \alpha x^\delta) / (\gamma + x^\delta)$
		<i>Weibull Type</i>	$f(x) = \alpha - \beta \exp(-\gamma x^\delta)$

Method, Nelder-Mead Method, Gradyant Method and Variable Measurement Method (2). These algorithms are usually appropriate for a certain problem and may require restrictive assumptions regarding continuity, existence of derivatives, existence of a single mode and other issues.

For instance, much more supplementary information is necessary for the Gauss-Newton Method, one of the commonly used algorithms, to be used as it should be. Moreover, using this algorithm involves a lot of attention as unless the starting point for the algorithm is chosen correctly, the algorithm may incorrectly pinpoint local optimums and thus yields only local optimums. In this case, an inappropriate starting area may augment the orientation of the search and delay the catch of the optimal solution.

In this study, Genetic Algorithms (GA), not necessitating any auxiliary information and starting to be an important tool for optimization problems, has been suggested with the intention of estimating parameters in non-linear regression models. A function necessary for parameter estimation has been defined as the objective function of the optimization problem in the second part of the study. In the third part, the genetic algorithm method has been portrayed. The fourth part covers the experimental results concerning Gauss-Newton Method and genetic algorithm. The reason why this part deals with only Gauss-Newton Method for comparative purposes is that it is the most commonly used method and that it is of the same type as the above mentioned methods (deterministic methods). The final part presents a discussion of the results obtained.

## 2. PARAMETER ESTIMATION

A typical non-linear regression model is like the following:

where  $e \sim N(0, \sigma^2)$ ,  $x \in R^m$ ,  $y \in R^1$  and  $f(x, \theta)$  is a non-linear function of  $\theta \in R^p$ . The following S-type growth models may set as example of non-linear regression (3):

Bu çizelgede verilen üç bilinmeyen parametrelili modeller için  $\theta' = (\alpha, \beta, \gamma)$  ve dört bilinmeyen parametrelili modeller için  $\theta' = (\alpha, \beta, \gamma, \delta)$  olduğu açıktır.

Eğer  $\{(x_i, y_i); i=1, 2, \dots, n\}$  eldeki veri çiftleri ise, doğrusal olmayan parametre tahmin problemi  $\theta$ 'nın en küçük kare tahminini bulmaktır ya da örneğin,  $\hat{\theta} \in R^P$  'yi belirlemektir, öyle ki;

$$S(\theta) = \sum_{i=1}^n [y_i - f(x_i, \theta)]^2 \quad [2]$$

Yani, doğrusal olmayan bir parametre tahmin problemi amaç fonksiyonu  $S(\theta)$  'nın en aza indirgenmesi için yapılan bir optimizasyon problemidir. Görüldüğü gibi  $S(\theta)$  doğrusal değildir ve  $f(x_i, \theta)$  fonksiyonuna bağlı olarak bir çok yerel minimuma sahip olabilir. Bu nedenle Giriş bölümünde anlatılan yöntemlerin sınırlılıklarından dolayı sezgisel (heuristic) yaklaşımları kullanmak analitik yöntemlere iyi bir alternatif olabilir. Sezgisel yöntemler kesin çözüme çok yakın çözümler veren ve hızla çözüme ulaşan yöntemlerdir. GA son yıllarda oldukça yaygın bir şekilde kullanılan sezgisel yöntemlerden birisidir ve özellikle doğrusal olmayan fonksiyonların optimizasyonunda sıkça kullanılmaktadır.

### 3. GENETİK ALGORİTMAYA GENEL BAKIŞ

GA hakkındaki çalışmalar ilk olarak 1975 yılında J.H. Holland tarafından yapılmıştır (4). GA 'en iyi' olanın hayatta kalma ilkesine ve doğal seçim mekanizmasını temel alan ve bu ilkelerin bilgisayarda simülasyonunun yapılmasıyla sonuca ulaşan bir yöntemdir.

Bu yöntemde bireylerin popülasyonları ile işleme başlanır. İkili dizinin temsil ettiği her birey parametre uzayı  $R^P$  'de bir noktayı temsil eder. Her nesilde her bir birey için amaç fonksiyonunun değeri onun uygunluğu olarak değerlendirilir ve daha uygun olan bireyler seçilerek yeni bir popülasyon elde edilir. Böylece, bireylerin uygunluğuna dayalı yeni çözümler oluşturulur. Yüksek uygunluk değerine sahip bireyler daha sıklıkla seçildiği için daha uygun olan bireylerin popülasyona katılması yönünde baskı vardır. Birkaç nesil sonra en iyi bireyin optimal çözümü temsil etmesi umulur (5). GA'nın işleyişi Şekil 1'de verilmiştir.

#### 3.1. Temsil ve Hesaplama

GA'da, bir popülasyondaki her bir birey genellikle sabit uzunluktaki bir ikili dizi olarak kodlanır. Dizinin uzunluğu alan parametrelerine ve istenen kesinliğe bağlıdır (6). Örneğin alan parametresi  $\theta$ 'nın aralığı  $[-2,5]$  ise ve gerekli kesinlik ondalık noktadan sonra altı basamak ise  $[-2,5]$  aralığı 7 milyon eşit aralığa bölünmelidir. Bu da dizi uzunluğunun 23 bit olması anlamına gelir. Bunun nedeni

It is obvious from Table 1 that for the three-unknown-parameter models  $\theta' = (\alpha, \beta, \gamma)$ , and for the four-unknown-parameter models  $\theta' = (\alpha, \beta, \gamma, \delta)$ .

If  $\{(x_i, y_i); i=1, 2, \dots, n\}$  is the data pairs at hand, then the non-linear parameter estimation problem is to find out the least square estimation of  $\theta$ , or, for instance, to identify  $\hat{\theta} \in R^P$ , so that

That is, a non-linear parameter estimation problem is an optimization problem whereby the objective function  $S(\theta)$  is minimized. As is seen,  $S(\theta)$  is non-linear and may have various local minimums, depending on  $f(x_i, \theta)$ .

Thus, due to the constraints of the methods presented in the introduction, it might be a better alternative to analytical methods to use heuristic methods. Heuristic methods are the methods that can produce solutions very close to the exact solution and that can arrive at the solution quicker. GA is one of the heuristic methods that have been widely used in recent years and especially genetic algorithms has a frequent use in the optimization of the non-linear functions.

### 3. AN OVERVIEW OF GENETIC ALGORITHMS

The pioneering study on GA was carried out by Holland in 1975 (4). GA is a method based upon the principle of the survival of "the fittest" and the natural selection mechanism and yielding the result through the simulation of these principles.

This method starts to operate with the population of individuals. Each individual represented by the binary string signifies a point in  $R^P$  the parameter space. The value of the objective function for each individual in each generation is considered to be the fitness of the individual and a new population is obtained through selecting the fitter individuals. In this way, new solutions based upon the fitness of the individuals are formulated. Since the individuals with a high fitness value are more frequently selected, there is pressure for the fitter individuals to be incorporated into the population. After a few generations, it is desired that the fittest individual represent the optimal solution (5). The way GA operates is illustrated in Figure 1.

#### 3.1. Representation and Evaluation

In GA, each individual in a population is generally encoded as a binary string with fixed length. The string length depends on the field parameters and the desired certainty (6). For example, if the range of the field parameter  $\theta$  is  $[-2,5]$  and the required certainty is a six-point in decimal, then the range  $[-2,5]$  should be divided into 7,000,000 equal interval, which means the length of the range equals to 23 bits. The reason for this is the following equation:

$$4194304 = 2^{22} < 7000000 < 2^{23} = 8388608$$

olmasıdır.  $\langle b_{22}b_{21}...b_0 \rangle$  şeklindeki ikili dizinin gerçek sayıya dönüştürülmesi iki adımda yapılır.

- İkili dizi  $\langle b_{22}b_{21}...b_0 \rangle$  2 tabanından 10 tabanına dönüştürülür

The binary string  $\langle b_{22}b_{21}...b_0 \rangle$  is transformed into real numbers in two steps.

- $\langle b_{22}b_{21}...b_0 \rangle$  is transformed from binary base into ten-digit base.

$$\theta' = \sum_{i=0}^{22} b_i 2^i \quad [3]$$

- $\theta$ 'ya karşılık gelen gerçek değer hesaplanır

- the real value equal to  $\theta$  is calculated.

$$\theta = -2.0 + \theta' \cdot \frac{7}{2^{23} - 1} \quad [4]$$

Parametre uzayı  $R^p$ 'nin boyutu  $p$ , 1'den büyükse her bir parametrenin dizilerini tek bir dizi halinde birleştirebilir ve bunu ayrıca başlı başına bir parametre vektörü olarak çözebiliriz.

Uygunluk değerleri  $S(\theta)$ 'nin değerleri ile ilişkilidir. Değerlendirme süreci, uygunluk değerleri arasından seçilen potansiyel çözümlere değer vererek çevre rolü üstlenir. Populasyonun değerlendirilmesine ilişkin genetik değerlendirmelerin yapılandırılması değerlendirme sürecinde olur.

### 3.2. Genetik Operatörler

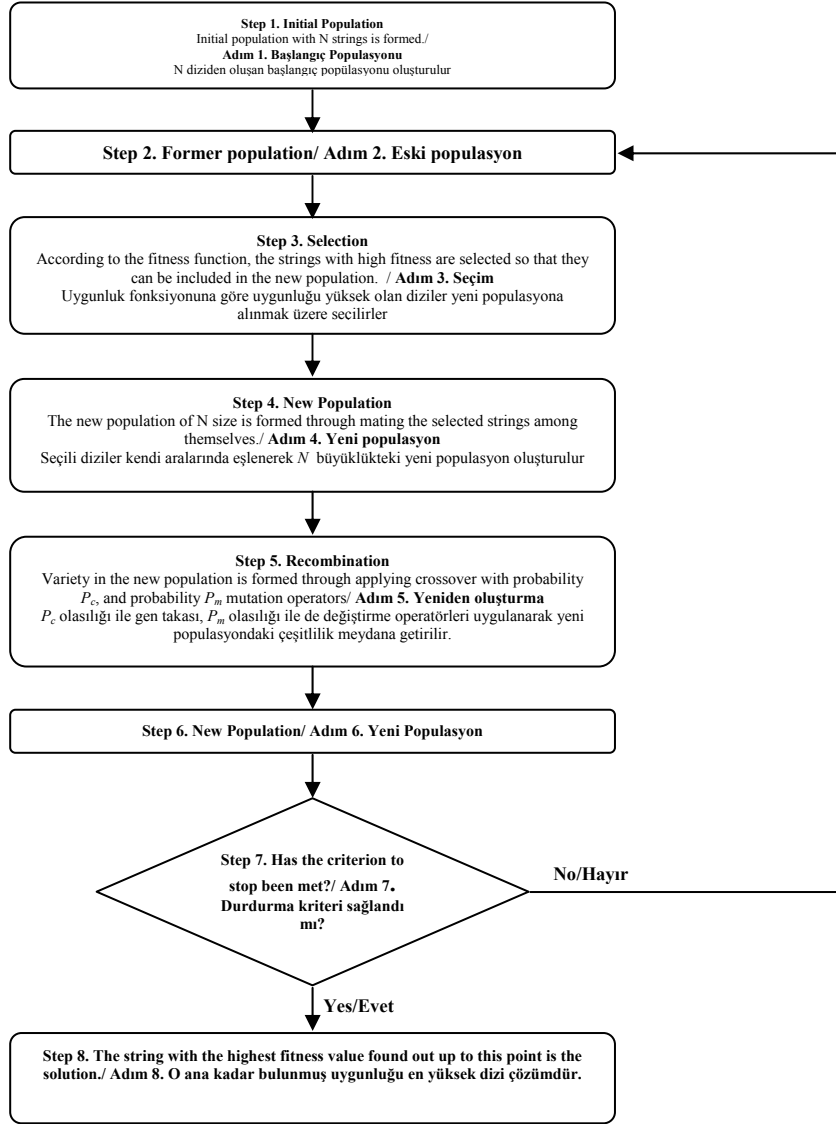
GA'daki bir populasyonun evrimsel süreci Darwin'in en iyi olanın hayatta kalması prensibine dayalıdır ve doğal olarak ortaya çıkan genetik işlemlerden sonra genetik operatöre yönlendirilir. En yaygın olarak kullanılan genetik operatörler, yeniden üretme, gen takası ve mutasyondur. Genetik işlemcileri uygulamak için üzerinde işlem yapılacak populasyondan bireylerinin seçilmesi gerekir. Seçim stratejisi temelde populasyonda gerçekten yer alan bireylerin uygunluk seviyesine bağlıdır. Uygunluğa bağlı farklı seçim stratejileri vardır ve bunlardan en çok kullanılanı uygunluk oranı seçimidir. Turnuva seçimi ve sıra seçimi de diğer iki alternatif stratejidir. Bu makalede, populasyondaki bireylerin amaç fonksiyonunun uygunluk değerlerinin sırasına bağlı olan sıra seçimi stratejisini kullanacağız (7). Bu stratejide,

If  $p$ , the dimension of the parameter space  $R^p$ , is bigger than 1, then the string of each parameter can be united into a single string and this can be solved as parameter vector on its own.

The fitness values are related to the values of  $S(\theta)$ . The evaluation process functions as the environment by attaching value to potential solutions selected among from fitness values. It is in the evaluation process that the genetic evaluations as to the evaluation of the population are structured.

### 3.2. Genetic Operators

The evolutionary process of a population in GA is based upon the Darwin's principle of "survival of the fittest" and directed to genetic operators after the genetic operations emerging naturally. The most widely used genetic operators are recombination, crossover and mutation. The individuals from the population to be worked on must be selected in order to apply the genetic operators. The selection strategy is largely dependent upon the fitness level of the individuals actually existing in the population. There are various selection strategies based on fitness, the most commonly used one of which is the fitness proportion selection. Tournament selection and ranking selection are the other two alternative strategies. In this study, the ranking selection strategy, based on the rank of the fitness values of the objective function of the individuals in the population, will be employed (7). In this strategy,



**Figure 1.** Operation of Genetic Algorithm  
**Şekil 1.** Genetik algoritmanın işleyişi

$S(\theta_i, t)$ :  $t$ 'inci popülasyondaki  $\theta_i$  bireyinin uygunluk değeri,

$n_i(t)$ :  $S(\theta_i, t)$ 'ler büyükten küçüğe doğru sıralandığında  $S(\theta_i, t)$ 'nin sıra sayısı

$N$ : Popülasyondaki birey sayısı

olmak üzere,  $\theta_i$ 'nin sonraki nesile ebeveyn olarak seçilme olasılığı,

$S(\theta_i, t)$ : the fitness value of the  $\theta_i$  individual in the  $t^{\text{th}}$  population,

$n_i(t)$ : rank number of  $S(\theta_i, t)$ s when they are ranked from high to low

$N$ : the number of individuals in the population

and the probability of  $\theta_i$  to be selected as parent for the following generation is

$$r(\theta_i, t) = \frac{2n_i(t)}{N(N+1)} \quad [5]$$

dir. Seçim süreci ile yeni bir popülasyon oluşturduktan sonra, genetik işlemciler aracılığıyla yeni çözümler üretmek üzere popülasyonun bazı üyeleri dönüştürme

After the formation of the new population through the selection process, in order to produce new solutions through genetic operators, some members of the new

işlemine girer (yeniden eşleştirme adımı). Sezgisel benzerliklerden dolayı GA'nın yeniden eşleştirme sürecinde sadece yeniden üretme, gen takası ve mutasyon işlemcileri kullanılır.

Üç genetik işlemci şöyle açıklanabilir; bir birey seçildiğinde yeniden üretme işlemcisi onu değişime uğratmadan mevcut popülasyondan yeni popülasyona sadece kopyalar. Gen takası işlemcisi seçilen iki bireyle başlar ve daha sonra gen takası noktası (1 ile  $L-1$  arasında bir tamsayı, burada  $L$  dizi uzunluğudur) tesadüfi olarak belirlenir. Ebeveyne ilişkin iki bireyin  $\theta_1$  ve  $\theta_2$  olduğunu, gen takası noktasının 5 ( $L=20$ ) olduğunu varsayarsak. Eğer

$$\theta_1 = (01001 | 1001100001000101)$$

$$\theta_2 = (11010 | 011100000010000)$$

ise, ortaya çıkan iki döl aşağıdaki gibidir.

$$\theta'_1 = (01001 | 011100000010000)$$

$$\theta'_2 = (11010 | 1001100001000101)$$

Üçüncü genetik işlem olan mutasyon, popülasyon içindeki yapılarda rastlantısal değişimler yol açar ve bazen bölgesel optimumdan kurtulma gibi yararlı sonuçlar ortaya koyabilir. Mutasyon dizilerin her birini olumsuzlamak için kullanılır. Örneğin  $p_m$  olasılığı ile 1'i 0'a, 0'ı 1'e çevirir.

Bu genetik işlemcilerinin mekaniği oldukça basittir ve sadece rastlantısal sayı üretimi, dizi kopyalama, kısmi dizi alışverişleri ve bit olumsuzlamaları gerektirir. GA'nın gerçek avantajı bunların ortak etkisinden kaynaklanır.

#### 4. DENEYSEL SONUÇLAR

Sayısal çözümlerimizde GA kontrol parametreleri, popülasyon boyutu  $N=60$ ; maksimum nesil sayısı 1000; yeniden üretme olasılığı  $p_r=0.1$ ; gen takası olasılığı  $p_c=0.9$ ; mutasyon olasılığı  $p_m=0.01$ , dizi uzunluğu her bir parametre 30 olmak üzere  $L=90$  (üç parametrelilik fonksiyonlarda) veya  $L=120$  (dört parametrelilik fonksiyonlarda) dir.

Tablo 1'de verilen büyüme modellerinin parametre tahminlerinde, Ratkowsky'dan alınmış olan Tablo 2'deki veri kümesi kullanılmıştır (2).

population start to be transformed (recombination step). Due to the heuristic similarities, only the recombination, crossover and mutation operators are used in GA mating process.

The three genetic operators can be explained as: when an individual is selected, the recombination operator just copies it, with no change whatsoever, from the current population to new population. The crossover operator starts with the selected two individuals and later crossover point (an integer between 1 and  $L-1$ ,  $L$  is the string length) is determined at random. Assuming that the two individuals regarding the parent are  $\theta_1$  and  $\theta_2$  and the crossover point is 5 ( $L=20$ ), if

then the emergent two offspring are like the following

Being the third genetic operation, mutation, brings about random changes in the structures in the population and can sometimes result in such useful outcomes as evading local optimums. Mutation uses each of the strings for negation purposes. For example, with the probability  $p_m$ , it changes 1 to 0 and 0 to 1.

The mechanics of these genetic operators is quite simple and they require only random number generation, string duplication, partial string exchange, and bit negations. The real advantage of GA originates from the combined effect of these.

#### 4. EXPERIMENTAL RESULTS

In the numerical analysis, the GA control parameters are population size  $N=60$ ; maximum generation number 1000; recombination probability  $p_r=0.1$ ; crossover probability  $p_c=0.9$ ; mutation probability  $p_m=0.01$ , string length for each parameter being 30  $L=90$  (for three-parameter functions) or  $L=120$  (for four-parameter functions).

In the parameter estimations of the growth models given in Table 1, the data sets in Table 2, taken from Ratkowsky, have been used.

Table 2. Data Sets

Tablo 2. Veri Kümeleri

Data sets/ Veri Kümeleri					
A		B		C	
Y	X	Y	X	Y	X
8.93	9	16.08	1	1.23	0
10.80	14	33.83	2	1.52	1
18.59	21	65.80	3	2.95	2
22.33	28	97.20	4	4.34	3
39.35	42	191.55	5	5.26	4
56.11	57	326.20	6	5.84	5
61.73	63	386.87	7	6.21	6
64.62	70	520.53	8	6.50	8
67.08	79	590.03	9	6.83	10
		651.92	10		
		724.93	11		
		699.56	12		
		689.96	13		
		637.56	14		
		717.41	15		

Daha sonra her bir veri kümesi için GA ile ayrı ayrı çözümler yapılmış ve sonuçlar Tablo 3’de literatürdeki sonuçlarla karşılaştırılmaları olarak verilmiştir.

Separate solutions are produced through GA for each data set and results are comparatively presented in Table 3.

Table 3. The results of Gauss-Newton and Genetic Algorithm for parameter estimation

Tablo 3. Parametre tahminleri için Gauss-Newton ve Genetik Algoritma sonuçları

Data Sets/ Veri Seti	Parameters/ Parametre	Gompertz		Logistic		Richards		Morgan-Mercer-Flodin (MMF)		Weibull Type	
		Gauss-Newton	GA	Gauss-Newton	GA	Gauss-Newton	GA	Gauss-Newton	GA	Gauss-Newton	GA
A	$\alpha$	82.830	82.730	72.46	72.534	69.62	69.170	80.69	81.000	69.96	70.986
	$\beta$	1.224	1.224	2.618	2.612	4.255	4.544	8.895	9.000	61.68	63.909
	$\gamma$	0.037	0.037	0.067	0.067	0.089	0.093	49577	47207	0.0001	0.0002
	$\delta$					1.724	1.875	2.828	3.000	2.378	2.2015
	$\sigma$	3.630	3.636	1.34	1.344	1.21	1.260	2.71	2.714	1.68	1.9928
B	$\alpha$	723.1	722.75	702.9	700.56	699.6	698.76	723.9	723.80	695.00	692.57
	$\beta$	2.500	2.503	4.443	4.444	5.277	5.422	33.35	33.600	673.50	673.98
	$\gamma$	0.450	0.451	0.689	0.689	0.760	0.775	6266	6418.0	0.0015	0.002
	$\delta$					1.279	1.321	4.641	4.700	3.262	3.197
	$\sigma$	1134	1133.9	744	744.17	799	799.36	1015	1015.1	712.00	724.48
C	$\alpha$	6.925	6.9213	6.687	6.691	6.684	6.659	6.986	6.993	6.656	6.691
	$\beta$	0.768	0.7696	1.745	1.764	1.780	2.089	1.181	1.182	5.549	5.577
	$\gamma$	0.493	0.4934	0.755	0.754	0.759	0.801	12.96	13.011	0.118	0.117
	$\delta$					1.017	1.174	2.475	2.480	1.763	1.757
	$\sigma$	0.0619	0.0619	0.0353	0.035	0.0424	0.043	0.0048	0.005	0.0268	0.0260

Tablo 3’deki  $\hat{\sigma}^2$  değerleri

The values of  $\hat{\sigma}^2$  in Table 3 are calculated from the following equation

$$\hat{\sigma}^2 = S(\hat{\theta}) / (n - p) \tag{6}$$

formülünden hesaplanmıştır. Burada  $n$  bir veri kümesindeki veri çiftlerinin sayısını  $p$  ise regresyon

where  $n$  represents the number of data pairs in a data set and  $p$  the number of parameters in regression equation.

denklemindeki parametre sayısını gösterir. Regresyon modeline bağlı olarak  $\hat{\theta}$ 'nin  $(\hat{\alpha}, \hat{\beta}, \hat{\gamma})'$  ya da  $(\hat{\alpha}, \hat{\beta}, \hat{\gamma}, \hat{\delta})'$  olduğuna dikkat edilmeli.

Örneğin Gompertz modeli için,

According to the regression model,  $\hat{\theta}$  is either  $(\hat{\alpha}, \hat{\beta}, \hat{\gamma})'$  or  $(\hat{\alpha}, \hat{\beta}, \hat{\gamma}, \hat{\delta})'$ .

For example, for Gompertz model, the following equation is calculated:

$$\hat{\sigma}^2 = \sum_{i=1}^n (y_i - \alpha \exp(-\exp(\beta - \gamma x)))^2 / (n - 3) \quad [7]$$

elde edilir.

## 5. SONUÇ

Tablo 3'deki sonuçlara bakıldığında GA sonuçları ile Gauss-Newton sonuçlarının birbirine çok yakın olduğu görülmektedir. Gauss-Newton gibi yöntemlerde parametre tahmininin başarılı olması aramanın başlangıç noktasına bağlıdır. Bu nedenle bu gibi yöntemlerde aramanın yapılacağı başlangıç noktasının belirlenmesi için önsel bir çalışma gerekir. Bunun yanında fonksiyonla ilgili türevlenebilme, değişkenlerin sürekliliği gibi bazı bilgilere de ihtiyaç duyulabilir. Oysa GA yöntemi, bu çalışmada da gösterildiği gibi hiçbir yardımcı bilgi ve önsel bir çalışma gerektirmez. Buna rağmen elde edilen sonuçlarında gösterdiği gibi GA yöntemi Gauss-Newton yöntemi kadar etkin olabilmektedir. Ayrıca GA'lar olasılığa ilişkin algoritmalar sınıfına dahildirler ancak doğrudan ve olasılıklı aramanın öğelerini birleştirdikleri için rastlantısal algoritmalarından çok farklıdır. Bir diğer önemli nitelikleri de bir potansiyel çözüm popülasyonunu sağlamasıdır. Ancak bütün diğer yöntemler araştırma evreninin tek bir noktasını inceler. Bu özelliklerinden dolayı GA yöntemi çok daha karmaşık fonksiyonların parametre tahminlerinde kullanılabilir. Bu nedenle genetik tabanlı yöntem diğer yöntemlere göre iyi bir alternatiftir.

## 5. CONCLUSION

The results in Table 3 suggest that it is clear that the results obtained through GA and those through Gauss-Newton are very close to one another. The success of the parameter estimation in such methods as Gauss-Newton depends on the starting point of the search. Thus, in such methods a preliminary study is necessary to identify the starting point where the search is to start. In addition, some information concerning derivativeness and continuous variables may be needed. However, as illustrated by this study, GA requires no supplementary information or preliminary study. All the same, GA may prove as effective as Gauss-Newton, as the results obtained have shown. Further, GA is classified under the probability algorithms class. Nonetheless, since GA combines direct search with probability search, it is largely different from random algorithms. Another important characteristic of GA is that GA produces a potential solution population. In contrast, all the other methods scan only a single point of the search space. Due to these features, GA can be used for the parameter estimation of the more complex functions. Consequently, genetic-based method is a good alternative for the other methods.

## REFERENCES/ KAYNAKLAR

1. Draper, N.R. and Smith, H., *Applied Regression Analysis, 2nd Edition*, Wiley, New York (1980).
2. Nash, J.C. and Walker-Smith, M., *Nonlinear Parameter Estimation*, Marcel Dekker, Inc., New York, Basel (1987).
3. Ratkowsky, D. A., *Nonlinear Regression Modeling*, Marcel Dekker, New York (1983).
4. Holland, J., *Adaptation in Natural and Artificial Systems*, University of Michigan Press, Ann Arbor (1975).
5. Goldberg, D.E., *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison Wesley, Reading, MA (1989).
6. Gen, M. and Cheng, R., *Genetic Algorithms and Engineering Design*, Wiley, New York (1996).
7. Whitley, D., "The Genitor Algorithm and Selection Pressure: Why Rank-based Allocation of Reproductive Trials is Best", in J.D.Schaffer (Ed.), *Proc. of the Third Int. Conf. on Genetic Algorithms*, 116-121 (1989).