İstatistik Araştırma Dergisi Cilt 03 No 02 Özel Sayı Sayfa 13-21 Ağustos 2004

REGRESYON ANALİZİNDE GİRDİLERİN BULANIK OLMASI DURUMUNDA UYARLAMALI AĞ YAKLAŞIMI İLE PARAMETRE TAHMİNİ

Türkan ERBAY DALKILIÇ * Ayşen APAYDIN**

ÖZET

Regresyon analizinde, bağımsız değişkene ait gözlem değerlerinin tek bir sınıftan gelmemesi ve bazı gözlemlerin hangi sınıfa ait olduğunun kesin olmaması başka bir değişle bulanık olması durumlarına aynı anda rastlanabilir. Böyle bir durumda bilinmeyen parametrelerinin tahmininde karmaşık problemlerin çözümünde etkin olan uyarlamalı ağlardan yararlanılabilir. Bu çalışmada, girdilerin bulanık olması durumunda regresyon modelinin bilinmeyen parametrelerinin tahmini için uyarlamalı ağı içeren bir algoritma önerilmiştir.

Anahtar Kelimeler *: Bulanık Regresyon, Parametre Tahmini, Uyarlamalı Ağ.*

1.GİRİŞ

Regresyon analizinde, bir veri kümesi için gözlemlerin tek bir sınıftan geldiği düşünülür ve bağımlı değişken ile bağımsız değişkenler arasındaki basit fonksiyonel ilişki $Y = f(X) + \varepsilon$ şeklindeki genel model ile ifade edilebilir. Veri kümesi birbirinden farklı dağılımlara sahip birden fazla sınıftan elde edilen gözlemlerin bir araya getirilmesiyle meydana gelmiş olabilir. c sınıf sayısını göstermek üzere, her sınıf bir f_i fonksiyonuyla ifade edildiğinde oluşturulacak regresyon modeli, switching regresyon modeli olarak da adlandırılmakta ve;

$$Y_i = f_i(X) + \varepsilon_i \qquad 1 \le i \le c \qquad (1)$$

biçiminde ifade edilmektedir (Lung. 1984, Michel. 2001, Richard. 1972).

Bu çalışmada verilerin birden fazla sınıftan gelmesi ve her bir verinin sınıflara ait olma durumlarının kesin olmaması, bir başka değişle bulanık olması durumunda veri kümesi için oluşturulan ve (1) eşitliği ile ifade edilen modellerin bir araya getirilerek

* Ankara Üniversitesi, Fen Fakültesi, İstatistik Bölümü, Ankara, Türkiye, terbay@science.ankara.edu.tr
**Ankara Üniversitesi, Fen Fakültesi, İstatistik Bölümü, Ankara, Türkiye,apaydin@science .ankara.edu.tr

farklı sınıflardan gelen verilere ait tek bir modelin oluşturulması aşamasında uyarlamalı ağlardan faydalanılacaktır.

2. BULANIK ÇIKARSAMA SİSTEMİ VE BULANIK UYARLAMALI AĞ

Bulanık çıkarsama sistemi, bulanık küme teorisi, bulanık Eğer-İse kuralları ve bulanık muhakemeye dayalı kullanışlı bir hesaplama yapısı oluşturur. Regresyon modelinin bilinmeyen parametrelerinin tahmininde kullanılabilen bulanık uyarlamalı ağlar, bulanık Eğer-İse kuralları ve bulanık çıkarsama sistemine dayanır. Problem, farklı dağılımlardan gelen bulanık girdilere bir regresyon doğrusu tahmin etmek olduğunda, Sugeno Bulanık Çıkarsama Sistemi çözüm için kullanılabilecek uygun sistemlerden biridir ve bu durumda önerilen bulanık kural,

$$R^l = \text{Eğer; } (x_1 = F_1^l \text{ ise ve } x_2 = F_2^l \text{ ise ve} \ldots\ x_p = F_p^l \text{ ise})$$

$$Y = Y^l = c_0^l + c_1^l x_1 + \ldots + c_p^l x_p$$

biçiminde ifade edilir. Burada; F_i^l bulanık kümeyi ve Y^l de R^l kuralına göre sistem çıktısını ifade etmektedir.

Bulanık kurallara göre elde edilen modellerin ağırlıklandırılmış ortalaması, Sugeno Bulanık Çıkarsama Sisteminin çıktısıdır ve farklı sınıflardan gelen veriler için ortak regresyon modeli bu ağırlıklandırılmış ortalama ile ifade edilir.

Bulanık regresyon analizi için, bulanık çıkarsama sisteminin işleyişine imkan veren sinir ağları uyarlamalı ağ olarak bilinmektedir. Regresyon fonksiyonuna iyi bir yaklaşım elde etmek için kullanılan, sinirlerden ve bağlantılardan meydana gelen uyarlamalı ağ beş tabakadan meydana gelmektedir (Hisao. 1992- 2001, Horia.1996).

Ağı meydana getiren sinirlere ait fonksiyonlar, parametre fonksiyonları ile karakterize edilir. Beş tabakadan oluşan uyarlamalı ağın işleyişi aşağıda verilmiştir. Burada her biri, iki farklı sınıftan elde edilen verilerden oluşan iki bağımsız değişken ve bir bağımlı değişken vardır. Uyarlamalı ağın işleyişinde bu değişkenler arasındaki fonksiyonel bağıntı modellenir:

Tabaka 1: Sistem iki düzeyli iki değişkenden meydana geldiği için dört bulanık kurala sahiptir. Bu kurallara ilişkin bulanık kümeler $F_1, F_2, F_3\ ve\ F_4$ ile gösterildiğinde tabakadaki h. sinirin çıktısı,

$$f_{1,h} = \mu_{F_h}(x_1) \qquad \text{h=1,2 için}$$

$$f_{1,h} = \mu_{F_h}(x_2) \qquad \text{h=3,4 için}$$

biçiminde tanımlanır. Burada μ_{F_h}, F_h'a ilişkin üyelik fonksiyonudur. F_h için farklı üyelik fonksiyonları tanımlanabilir. Burada verilerin, parametre kümesi $\{v_h, \sigma_h\}$ olan Normal Dağılım' dan geldiği düşünüldüğünde, üyelik fonksiyonları,

$$\mu_{F_h}(x_1)=\exp\left[-\left(\frac{x_1-v_h}{\sigma_h}\right)^2\right] \qquad \text{h=1,2 için,}$$

$$\mu_{F_h}(x_2)=\exp\left[-\left(\frac{x_2-v_h}{\sigma_h}\right)^2\right] \qquad \text{h=3,4 için}$$

biçimininde tanımlanır. Bu tabakadaki $\{v_h,\sigma_h\}$ parametre kümesi *önsel parametreleri* gösterir.

Tabaka 2: Bu tabakadaki her sinir sabitlenmiş sinirdir ve Λ_l ile etiketlenmiştir (l=1,...,4). Birinci tabakadan gelen sinyaller bu tabakanın girdi sinyalleridir ve Λ_l; bu girdi sinyallerinin çarpımı şeklinde tanımlanır. Bu tabakaya ilişkin sinir fonksiyonları,

$$f_{2,1}=w^1=\mu_{F_1}(x_1).\mu_{F_3}(x_2)$$
$$f_{2,2}=w^2=\mu_{F_1}(x_1).\mu_{F_4}(x_2)$$
$$f_{2,3}=w^3=\mu_{F_2}(x_1).\mu_{F_3}(x_2)$$
$$f_{2,4}=w^4=\mu_{F_2}(x_1).\mu_{F_4}(x_2)$$

ile ifade edilir.

Tabaka 3: Bu tabakadaki sinirler N_l ile etiketlenmiş ve ikinci tabakada olduğu gibi sabit sinirlerdir. Bu tabakanın çıktısı ikinci tabakanın çıktılarının bir normalizasyonudur ve sinir fonksiyonu;

$$f_{3,l}=\overline{w}^l=\frac{w^l}{\sum_{t=1}^{m} w^t} \qquad l=1,...,4 \qquad (2)$$

olarak tanımlanır.

Tabaka 4: Bu tabakanın çıktı sinyalleri de bir fonksiyona bağlıdır ve bu fonksiyon;

$$f_{4,l}=\overline{w}^l Y^l \qquad l=1,...,4$$

ile ifade edilir. Burada Y^l, bulanık Eğer-İse kuralının sonuç kısmıdır ve

$$Y^l=c_0^l+c_1^l x_1+c_2^l x_2$$

ile verilir. c_i^l ise bulanık sayılardır ve *sonsal parametreleri* gösterirler.

Tabaka 5: Bu tabakadaki tek sinir sabitlenmiş sinirdir ve gelen sinyallerin tümünün toplamı olarak,

$$f_{5,1} = \hat{Y} = \sum_{l=1}^{4} \overline{w}^l Y^l$$

biçiminde hesaplanır (Chi-Bin. 1999, Chi-Bin 2001, Hisao.1993).

Verilen girdi-çıktı veri çiftleri arasındaki ilişkinin modelini elde etmeyi amaçlayan uyarlamalı ağın eğitimi hata ölçüsüne dayanmaktadır. Ağdan elde edilen tahmin ile hedeflenen çıktı arasındaki fark hata olarak tanımlandığında, ağ bu hata ölçüsünü en küçük yapacak şekilde model oluşturabilmek üzere eğitilmelidir. Eğitim, hata ölçütü önceden belirlenen bir değerden küçük olduğunda sona erer.

3. UYARLAMALI AĞ İLE PARAMETRE TAHMİNİ İÇİN BİR ALGORİTMA

Uyarlamalı ağ ile parametre tahmini, hata ölçütünün en küçüklenmesi prensibine dayanır. Chi-Bin C. (1999) tarafından, farklı sınıflardan gelen verilere ilişkin regresyon modellerinin oluşturulması ve bu regresyon modellerine dayanan ortak bir tahminin elde edilmesi süreci için bir algoritma önerilmiştir. Tahmin sürecinin iki önemli adımı, verilerin geldiği sınıfı karakterize eden önsel parametre setinin belirlenmesi ve bu parametrelerin süreç içinde güncellenmesidir. Chi-Bin tarafından önerilen algoritmada sonsal parametreler $c_i^l = (a_i^l, b_i^l)$, Tanaka tarafından önerilen ve

$$\begin{aligned}
&\min \sum_{k=1}^{N}\sum_{l=1}^{m}\sum_{i=0}^{p} \overline{w}^l b_i^l x_{ik} \quad (= \min \sum_{k=1}^{N} \hat{e}_k) \\
&b_i^l \geq 0 \quad i = 1,\ldots,p \quad l = 1,\ldots,m \\
&\sum_{l=1}^{m}\sum_{i=0}^{p} \overline{w}^l a_i^l x_{ik} + (1-\alpha)\sum_{l=1}^{m}\sum_{i=0}^{p} \overline{w}^l b_i^l x_{ik} \geq y_k + (1-\alpha)e_k \\
&-\sum_{l=1}^{m}\sum_{i=0}^{p} \overline{w}^l a_i^l x_{ik} + (1-\alpha)\sum_{l=1}^{m}\sum_{i=0}^{p} \overline{w}^l b_i^l x_{ik} \geq -y_k + (1-\alpha)e_k
\end{aligned} \tag{3}$$

biçiminde modellenen, doğrusal programlama probleminin çözülmesiyle elde edilir. (James. 2000, Jyh-Shing. 1993).

Elde edilen parametre setinin süreç içerisinde güncellenmesi ise,

$$\in_{5,1} = \frac{\partial \varepsilon_k^2}{\partial y_k}, \qquad \in_{r,l} = \sum_{h=1}^{M_{r+1}} \in_{r+1,h} \frac{\partial F_{r+1,h}}{\partial f_{r,l}}$$

biçiminde tanımlanan geri yayılım hatalarına bağlıdır ve

$$\Delta\rho = -\eta\frac{\partial(y_k - \hat{y}_k)^2}{\partial\rho} \tag{4}$$

eşitliğinden elde edilir. Burada ρ önsel parametreler ve η, öğrenme oranıdır. (0 1] aralığında değer alan öğrenme oranı karar verici tarafından belirlenir.

Regresyon modellerini oluşturacak katsayıların belirlenmesi için verilen bu algoritmada, hatası en küçük olan tahmine ulaşılmaya çalışılmaktadır. Hatası en küçük olan tahmine ulaşmak, önsel parametrelerin doğru belirlenmesine ve güncellenmesine ve ayrıca sonsal parametrelere bağlıdır. Chi-Bin (1999) tarafından önerilen algoritmada sonsal parametreler (3) eşitliği ile verilen Tanaka modelinin çözümü ile elde edilir. Ancak çözüm aşamasında bu doğrusal programlama modeli bazı durumlarda kullanıcıyı çözümsüzlüğe taşıyabilmektedir ve her problem için etkin çözüm verememektedir. Önsel parametrelerin güncellenmesi de geri yayılım hatalarına bağlı olduğu için bir dizi karmaşık işlemi beraberinde getirmektedir. Tüm bu sakıncaları gidermek için önsel parametrelerin farklı belirlendiği ve güncellendiği bir algoritma önerildi. Burada merkezler ve yayılımlardan oluşan önsel parametre setinde oldukça önemli olan merkez parametrelerinin belirlenmesi amaçlandı. Önerilen algoritma adımları aşağıdaki gibi tanımlandı.

Adım 1: Önsel parametreler belirlenir: Yayılımlar, girdi değişkenlerinin değer aldığı aralığa ve değişkenlerin düzey sayılarına göre sezgisel olarak belirlenir. Bu parametredeki değişim tahmini çok fazla etkilememektedir. Merkez parametreleri de değişkenlerin değer aldığı aralığa ve düzey sayısına bağlıdır. Örneğin iki düzeyli X_i değişkeni için merkeze ilişkin ilk önsel parametre değerleri,

$$V_1 = \text{Enk}\,(X_i)\ ,\quad V_2 = \text{Enb}\,(X_i)$$

biçiminde belirlenir.

Adım 2: Bağımsız değişkenlerin bulanık, bağımlı değişkenin kesin sayılardan oluştuğu durumda, sonsal parametreleri ifade eden c_i^l de kesin sayılar olarak elde edilir. Bu durumda sonsal parametrelerin saptanması için,

$$Z = (B^T B)^{-1} B^T Y$$

eşitliği kullanılır (Chi-Bin. 2001). Burada,

$$B = \begin{bmatrix} \overline{w}_1^1, & \Lambda, & \overline{w}_1^m, & \overline{w}_1^1 x_{11}, & \Lambda, & \overline{w}_1^m x_{11}, & \Lambda, & \overline{w}_1^1 x_{p1}, & \Lambda, & \overline{w}_1^m x_{p1} \\ M & & & & & \overline{w}_k^l x_{jk} & & & & M \\ \overline{w}_n^1 & \Lambda, & \overline{w}_n^m, & \overline{w}_n^1 x_{1n}, & \Lambda, & \overline{w}_n^m x_{1n}, & \Lambda, & \overline{w}_n^1 x_{pn}, & \Lambda, & \overline{w}_n^m x_{pn} \end{bmatrix},$$

$$Y = [y_1, y_2, ..., y_n]^T \text{ ve}$$

$$Z = [a_0^1, ..., a_0^m, a_1^1, ..., a_1^m, a_p^1, ..., a_p^m]^T$$

biçiminde tanımlanır. B matrisinin oluşturulmasında kullanılan $\overline{w}_i^l$, (2) eşitliğinde de tanımlandığı gibi, uyarlamalı ağın üçüncü tabakasındaki sinirlerin çıktısıdır. Bu yöntem bağımlı değişken bulanık olduğu durumda da kullanılabilen bir yöntemdir.

Adım 3: Hedeflenen sonuç ile elde edilen çıktı arasındaki fark olarak tanımlanan hata

$$\varepsilon_k = Y_k \{-\} \hat{Y}_k$$

biçiminde hesaplanır. Burada {-} bağımlı değişkenin de bulanık olması durumunda fark operatörüdür.

Eğer hata ölçüsü, önceden belirlenmiş kabul edilebilir bir değerden küçükse durulur. Bu durumda Adım 2'de sonucun bulanık ya da kesin olması durumuna bağlı olarak izlenen yöntem sonucunda ulaşılan sonsal parametre, kurulacak olan regresyon modellerinin parametreleri olarak elde edilmiştir, sürece son verilir.

Eğer hata ölçüsü yeterince küçük değilse adım 4'e geçilir ve sürece devam edilir.

Adım 4: Enk(X_i) ve Enb(X_i) olarak belirlenen merkezi önsel parametreler, en küçük değerden en büyük değere doğru artacak, en büyük değerden en küçük değere doğru azalacak şekilde,

$$V_1' = \text{Enk}(X_i) + t, \quad V_2' = \text{Enb}(X_i) - t$$

ile yenilenir. Burada t bağımsız değişkenlerin aldığı değerlere bağlı olacak şekilde belirlenir. Değişim ile elde edilen her önsel parametre için tahminler ve bu tahminlere ilişkin hata ölçütleri hesaplanır.

Adım 5: Hesaplanan hata ölçütlerinden en küçük olanı belirlenir. Belirlenen en küçük hatayı veren önsel parametreler ve bu parametrelere ilişkin modellerden elde edilen tahmin çıktı olarak alınır.

Önerilen algoritma MATLAB'da yazılan bir program ile işletildi. Bu adımsal işletim aşamasında farklı veri setleri ele alındı ve daha önce sonuçlanmış model yapıları ile karşılaştırıldı. Yapılan karşılaştırmalardan, uyarlamalı ağdan elde edilen modellerin en küçük hataya ulaştıkları görüldü. Çalışmanın son kesiminde, önerilen algoritmanın işlerliği bir uygulama ile gösterildi.

4. UYGULAMA

Bağımsız değişkenlerin normal dağılımdan gelmesi durumunda önerilen yöntemi irdeleyebilmek ve klasik yöntemler ile karşılaştırabilmek için, simülasyon ile bir veri seti türetildi. Daha sonra MATLAB da hazırlanan program kullanılarak çözümlemeler yapıldı.

Üç bağımsız değişken ve bir bağımlı değişkenin yer aldığı veri setine için uyarlamalı ağ'dan elde edilen tahmin değerleri ($\hat{y}_{Ağ}$) ve bu tahminlere ilişkin hatalar ($e_{(Ağ)i}$ $(i=1,...,n)$) ve en küçük kareler yöntemi ile elde edilen tahminler ($\hat{y}_{EKK}$) ve bu tahminlere ilişkin hatalar ($e_{(EKK)i}$ $(i=1,...,n)$) Tablo 1'de verildi.

Tablo 1. Normal dağılımdan gelen üç bağımsız değişkene ilişkin veri setine ait tahminler ve hatalar

X_1	X_2	X_3	Y	$\hat{y}_{Ağ}$	$e_{(Ağ)i}$	$\hat{y}_{EKK}$	$e_{(EKK)i}$
20.8841	17.3119	28.2459	17.4275	18.3357	0.9082	13.6840	3.7434
25.9724	20.4055	27.5454	14.9750	16.1843	1.2093	14.0011	0.9739
19.9736	19.5829	30.2403	13.6189	15.0042	1.3853	14.1848	-0.5659
23.6113	16.5098	26.3084	21.3823	22.0432	0.6609	13.3614	8.0209
21.2345	23.5512	22.3080	24.3170	23.5137	-0.8033	15.9467	8.3703
20.3018	19.9537	24.4027	12.3872	12.0337	-0.3535	14.8677	-2.4805
20.9154	21.6087	15.2039	15.5171	12.4232	-3.0939	16.2194	-0.7023
23.3933	17.8507	26.8931	10.9618	11.7609	0.7991	13.6833	-2.7216
25.6947	18.0333	25.4678	18.4022	19.0136	0.6114	13.6189	4.7833
24.5355	20.9431	22.6356	3.1771	2.8065	-0.3705	14.8331	-11.6561
29.2565	20.3204	22.3812	19.9506	20.1203	0.1698	14.1497	5.8008
24.0185	25.5446	20.6774	16.0945	15.1277	-0.9668	16.3335	-0.2390
18.9298	19.1747	24.8091	16.3083	15.9394	-0.3690	14.7738	1.5345
25.0389	26.6377	26.5174	21.0672	21.7754	0.7082	15.8870	5.1802
25.2855	24.5256	23.6786	13.6267	13.6220	-0.0046	15.5952	-1.9685
26.2661	14.1648	23.0004	14.3343	14.4642	0.1299	12.7797	1.5546
26.9993	14.9584	24.8561	8.6475	9.4438	0.7963	12.7103	-4.0628
30.1123	18.2794	24.3010	6.6820	7.6003	0.9183	13.3002	-6.6182
22.8087	19.4426	21.1709	11.4822	10.3834	-1.0988	14.7861	-3.3039
26.0432	20.0268	30.1702	16.4044	18.2827	1.8782	13.6121	2.7923
24.9473	22.5108	26.7636	12.2940	13.0556	0.7616	14.7662	-2.4722
22.6789	17.8332	30.1238	8.3323	9.9482	1.6159	13.4170	-5.0847
33.5452	17.8355	23.0091	20.3634	21.5100	1.1466	12.9234	7.4400
23.9695	19.3965	20.5251	11.4396	10.3599	-1.0797	14.7088	-3.2692
19.3619	19.9386	28.2306	14.9436	15.6531	0.7096	14.5645	0.3791
32.0804	20.8367	25.1648	14.9959	16.2294	1.2335	13.6664	1.3295
26.3022	23.1749	21.9752	13.7528	13.3411	-0.4117	15.2977	-1.5448
20.5238	21.8650	24.6435	16.9829	16.7977	-0.1852	15.3284	1.6545
27.4814	14.7482	16.9646	13.6799	12.1240	-1.5560	13.4384	0.2415
30.0791	22.0920	29.3357	6.6799	8.6132	1.9333	13.7891	-7.1092

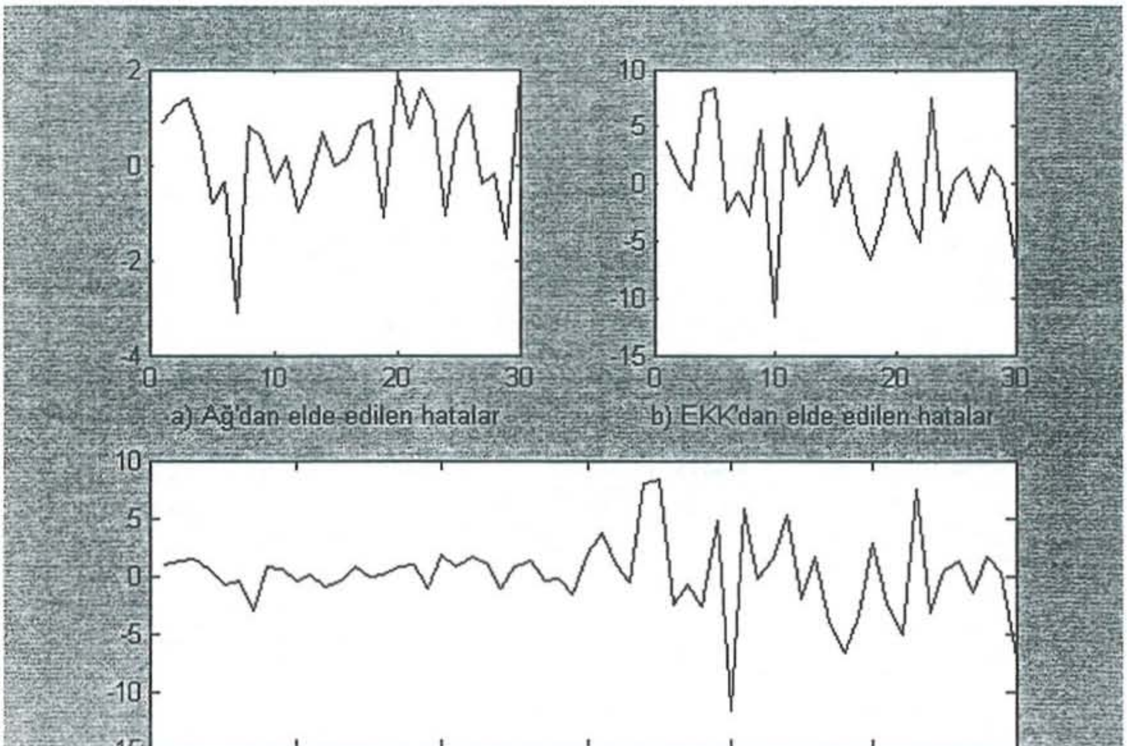
Ağ'dan elde edilen tahminler için hata

$$\varepsilon_{Ağ} = \frac{\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2}{n} = 1.2675$$

olarak ve En Küçük Kareler yönteminden elde edilen tahminler için hata ise,

$$\varepsilon_{EKK} = \frac{\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2}{n} = 20.8775$$

olarak elde edildi. Tahminlere ilişkin hataların grafikleri ise Şekil 1'de yer almaktadır.



Şekil 1. Tablo 1'de yer alan veri setine ilişkin hataların grafikleri

KAYNAKLAR

CHI-BIN C., LEE E. S. (2001), *Switching Regression Analaysis by Fuzzy Adaptive Network*, Europen Journal of Operational Research, 128. 647-663.

CHI-BIN C., LEE E. S. (1999), *Applying Fuzzy Adaptive Network to Fuzzy Regression Analysis,* An International Journal Computers & Mathematics With Applications. 38. 123-140.

HISAO I., HIDEO T. (1993) *An Architecture of Neural Networks with Interval Weights and Its Application to Fuzzy Regression Analysis*, Fuzzy Sets and Systems. 57. 27-39.

HISAO I., MANABU N. (2001), *Fuzzy Regression Usin Asymmetric Fuzzy Coefficients and Fuzzied Neural Networks,* Fuzzy Sets and Systems. 119. 273-290.

HISAO I., TANAKA H. (1992), *Fuzzy Regression Analaysis Using Neural Networks*, Fuzzy Sets and Systems. 50. 257-265.

HORIA F., COSTEL S. (1996), *A New Fuzzy Regression Algorithm,* Anal. Chem. 68. 771-778.

JAMES P.D., DONALT W. (2000), *Fuzzy Regression by Fuzzy Number Neural Networks,* Fuzzy Sets and Systems. 112. 371-380.

JYH-SHING ROGER JANG (1993), ANFIS: *Adaptive-Network-Based Fuzzy Inference System, IEEE Transaction on Systems*, Man and Cybernetics. 23. No:3. 665-685.

LUNG-FEI L., ROBERT H.P. (1984), *Switching Regression Models With Imperfect Sample Separation Information-With an Application on Cartel Stability,* Econometrica. 52. 391-418.

MICHEL M. (2001), *Fuzzy Clustering and Switching Regression Models Using Ambiguity and Distance Rejects,* Fuzzy Sets and Systems. 122. 363-399.

RICHARD E.Q. (1972), *A New Approach to Estimating Switching Regressions,* Journal of the American Statistical Association. 67. No:338. 306-310.

ADAPTIVE NETWORK APPROACH TO PARAMETER ESTIMATION IN FUZZY REGRESSION FOR FUZZY INPUT

ABSTRACT

In regression analysis, the condition that the observations do not come from one single class for a data set and some observations belong to which present classes are not clear can be seen at the same time. In this work, an algorithm is proposed for estimation of the unknown parameters in regression model by adaptive network when the input is fuzzy.

Key Words : *Adaptive Network, Fuzzy Regression, Parameter Estimation.*