T.C.
SELÇUK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

LİNEER OLMAYAN BULANIK REGRESYONDA TAHMİN

ÜMRAN MÜNİRE TEKŞEN

YÜKSEK LİSANS TEZİ
İSTATİSTİK ANA BİLİM DALI

Konya, 2008

ÖZET

Yüksek Lisans Tezi

LİNEER OLMAYAN BULANIK REGRESYONDA TAHMİN

Ümran Münire TEKŞEN

Selçuk Üniversitesi
Fen Bilimleri Enstitüsü
İstatistik Ana Bilim Dalı

Danışman: Doç. Dr. Aşır GENÇ

2008, 121 sayfa

Gözlemlerin bulanık sayı olması durumunda lineer olmayan regresyon analizi için bazı yöntemler vardır. Yapay sinir ağları, en küçük kareler destek vektör mekanizmasına dayalı bulanık regresyon ve bulanık en küçük mutlak sapmalar regresyonu bunlardan bazılarıdır. Bu çalışmada, belirtilen yöntemler ele alınacak ve sayısal uygulamaların sonuçları karşılaştırılacaktır.

Anahtar Kelimeler: Lineer Regresyon, Lineer Olmayan Regresyon, Bulanık Küme Kuramı, Yapay Sinir Ağları, En Küçük Kareler Destek Vektör Mekanizması, Bulanık En Küçük Mutlak Sapmalar Regresyonu, Doğrusal Programlama.

ABSTRACT

Ms Thesis

ESTIMATION IN FUZZY NONLINEAR REGRESSION

Ümran Münire TEKŞEN

Selçuk University
Graduate School of Natural and Applied Sciences
Department of Statistics

Supervisor: Assoc. Prof. Dr. Aşır GENÇ

2008, 121 page

There are some methods for fuzzy nonlinear regression analysis when the observations are fuzzy numbers. Some of them are artificial neural networks, least square-support vector machines and fuzzy least absolute deviations regression. In this study, explained methods will be handled and the results of numeric applications will be compared.

Key Words: Linear Regression, Nonlinear Regression, Fuzzy Sets, Artificial Neural Networks, Least Squares-Support Vector Machines, Fuzzy Least Absolute Deviations Regression, Linear Programming.

ÖNSÖZ

Lineer olmayan bulanık regresyonda tahmin adlı tezimin seçimi ve gerçekleşmesinde yardımını esirgemeyen, çalışmalarımda yol gösteren ve bilgileriyle katkıda bulunan değerli hocam Doç.Dr. Aşır GENÇ'e, yüksek lisans öğrenimim sırasında maddi olarak verdiği destekten dolayı TÜBİTAK'a ve olumlu yaklaşımlarıyla her zaman yanımda olan aileme teşekkür ederim.

İÇİNDEKİLER

ÖZET……..…...i
ABSTRACT …….....ii
ÖNSÖZ …….....iii
İÇİNDEKİLER……...iv
ŞEKİLLER DİZİNİ.....ix
ÇİZELGELER DİZİNİ.....x
KISALTMALAR.....xi

1. GİRİŞ VE ÖNCEKİ ÇALIŞMALAR.....1

1.1. Giriş…….……..……1
1.2. Önceki Çalışmalar……...……....2

2. BULANIK KÜME KURAMI……...…….....……...4

2.1. Bulanık Kümelerde Temel Kavramlar ..…….....……..5
2.2. Bulanık Kümelerde Temel İşlemler……..……..8
2.2.1. Kesişim ……...8
2.2.2. Birleşim …….....…8
2.2.3. Tümleme …….....9
2.2.4. Kapsama …….....9
2.2.5. Eşitlik …….....9
2.2.6. Cebirsel çarpım …….....10
2.2.7. Cebirsel toplam …….....…10
2.2.8. Fark …….....…10
2.3. Bazı İleri Bulanık Küme İşlemleri…….....11
2.3.1. Bazı t-conorm operatörleri …….....…11
2.3.2. Bazı t-norm operatörleri …….....12

2.4. Üyelik Fonksiyonları…….....…12
2.4.1. Üyelik fonksiyonları biçimleri …….....13
2.5. Bulanık Sayılar…….....…16
2.6. Bulanık Sayılarda İşlemler…….....…16
2.6.1. Toplama işlemi…….....……16
2.6.2. Çıkarma işlemi…….....…17
2.6.3. Çarpma işlemi…….....…17
2.6.4. Bölme işlemi…….....…18
2.7. Üçgensel ve Yamuksal Bulanık Sayılar…….....……18
2.7.1. Üçgensel ve yamuksal bulanık sayılar üzerinde bulanık aritmetik…...…19

3. LİNEER REGRESYON ANALİZİ 22

3.1. Parametre Tahmini.....24
3.1.1. EKK (En küçük kareler kestiricisi).....25
3.1.2. En çok olabilirlik yöntemi.....27
3.1.3. MLE tahmin edicilerinin bazı istatistiksel özelikleri.....28
3.2. Artıklar.....30
3.3. Çoklu Belirleyicilik Katsayısı 30
3.4. Bireysel Parametreler İçin Güven Aralığı.....31
3.5. Bireysel Parametreler İçin Hipotez testi ….....31
3.6. $\mu_{Y/\underline{x}'}$ 'nin Tahmini.....32
3.7. Sonradan Gözlemin Tahmini.....32

4. LİNEER OLMAYAN REGRESYON ANALİZİ….....34

4.1. Modelin Tanıtımı ve Bazı Gösterimler.....34
4.2. Parametre Tahmini.....35

4.2.1 Başlangıç noktasının seçimi.....40
4.2.2. Durdurma kuralı.....41
4.2.3. Gauss- Newton Yöntemi….....41

5. YAPAY SİNİR AĞLARI.....45
5.1. Giriş.....45
5.1.1. Biyolojik nöron yapısı 45
5.1.2. Aktivasyon fonksiyonları.....47
5.1.3. İşlemci elemanları (Yapay nöron).....48
5.2. YSA ile Hesaplamanın Özellikleri.....49
5.2.1. Doğrusal olmama 50
5.2.2. Öğrenme..... 50
5.2.3. Genelleme.....50
5.2.4. Uyarlanabilirlik 50
5.2.5. Dağıtılmış birleşik hafıza 51
5.2.6. Hata toleransı.....51
5.2.7. Paralel işlem yapma 51
5.3. Yapay Sinir Ağlarının Sınıflandırılması.....52
5.3.1. YSA'ların yapılarına göre sınıflandırılması 52
5.3.1.1. İleri beslemeli ağlar..... 52
5.3.1.2. Geri beslemeli ağlar 53
5.3.2. YSA'ların öğrenme algoritmalarına göre sınıflandırılması 53
5.3.2.1. Danışmanlı öğrenme (Supervised learning) 54
5.3.2.2. Danışmansız öğrenme (Unsupervised learning) 54
5.3.2.3. Takviyeli öğrenme (Reinforcement learning) 55
5.4. Çok Katmanlı Perseptronlar ve Öğrenme Algoritmaları.....56
5.4.1. Çok katmanlı perseptronlar 56
5.4.2. Geri yayılım algoritması 58

5.5. Diğer Yapay Sinir Ağları 61
5.5.1. LVQ (Learning vector quantization)…….....……61

6. YSA YARDIMIYLA LİNEER OLMAYAN BULANIK REGRESYON TAHMİNİ.....63

6.1. Bulanık Regresyon.....63
6.2. Lineer Olmayan Bulanık Regresyon.....63
6.3. Modelin Test Edilmesi…….....68

7. LS-SVM'YE DAYALI BULANIK LİNEER OLMAYAN REGRESYON.....70

7.1. LS-SVM.....70
7.1.1. Support vector machine (SVM) sınıflayıcısı.....70
7.1.2. En küçük kareler (Least Square) SVM.....72
7.2. Lagrange Yöntemi.....74
7.3. Bulanık Lineer LS-SVM Regresyon.....75
7.4. Bulanık Lineer Olmayan LS-SVM Regresyon…….....…79

8. DOĞRUSAL PROGRAMLAMA İLE BULANIK EN KÜÇÜK MUTLAK SAPMA REGRESYONU.....82

8.1. Simpleks Yöntem.....82
8.1.1. Doğrusal programlama probleminin standart biçimi.....84
8.1.2. Temel çözümler.....85
8.1.2.1. Temel uygun çözümü iyileştirme.....87
8.1.3. Başlangıç temel uygun çözümün bulunması.....95
8.1.3.1. Temel uygun çözümün bulunmasında bazı aykırı durumlar.....96
8.2. Bulanık En Küçük Mutlak Sapma Regresyonu 97
8.2.1. Genel bulanık lineer model.....97
8.2.2. Bulanık En Küçük Mutlak Sapma Regresyonu….....98

9. UYGULAMALAR.....102

9.1. Uygulama 1.....102
9.2. Uygulama 2.....108
9.3. Uygulama 3.....112

10. SONUÇ.....115

KAYNAKLAR.....118

ŞEKİLLER DİZİNİ

Sayfa

Şekil 2.1. Bulanık kümelerde zayıf α -seviye kesme kümesi 8

Şekil 2.2. Üçgensel bulanık sayı gösterimi 15

Şekil 4.1. $\underline{Y}$ vektörünün $F(\underline{\theta})$ sütun uzayına dik izdüşümü 39

Şekil 5.1. Biyolojik sinir sisteminin blok gösterimi 46

Şekil 5.2. YSA'larda en çok tercih edilen aktivasyon fonksiyonları a) sigmoid, b) doğrusal c) hiperbolik tanjant ve d) keskin sınırlayıcı 48

Şekil 5.3. Bir işlemci elemanı (yapay nöron) 49

Şekil 5.4. İleri beslemeli ağ için blok diyagram 52

Şekil 5.5. Geri beslemeli ağ için blok diyagram 53

Şekil 5.6. Danışmanlı öğrenme yapısı 54

Şekil 5.7. Danışmansız öğrenme yapısı 55

Şekil 5.8. Takviyeli öğrenme yapısı 55

Şekil 5.9. Geri yayılım MLP yapısı 57

Şekil 5.10. Çok katmanlı bir perseptron geri yayılım akış şeması 60

Şekil 5.11. LVQ ağı 61

Şekil 9.1. YSA eğitim verisi tahmin değerleri 105

Şekil 9.2. YSA test verisi tahmin değerleri 105

Şekil 9.3. YSA alt ve üst sınır değerleri 106

Şekil 9.4. LS-SVM tahmin değerleri 109

Şekil 9.5. LS-SVM alt ve üst sınır değerleri 111

Şekil 9.6. FLAD tahmin, alt ve üst sınır değerleri 113

ÇİZELGELER DİZİNİ

Sayfa

Çizelge 3.1. Artıkların hesaplanması 31

Çizelge 5.1. Sinir sistemi ile YSA'nın benzerlikleri 46

Çizelge 9.1. Uygulama verisi 103

Çizelge 9.2. YSA'da model yeterliliği değerleri 106

Çizelge 9.3. YSA uygulama sonuçları 107

Çizelge 9.4. LS-SVM uygulama verisi 108

Çizelge 9.5. LS-SVM uygulama sonuçları 110

Çizelge 9.6. FLAD uygulama verisi 112

Çizelge 10.1. Yöntemlerin hata kareler ortalaması ve R^2 değerleri 115

Çizelge 10.2. Yöntemlerin avantaj ve dezavantajları 116

KISALTMALAR

EKK: En küçük kareler

EOTE: En çok olabilirlik

YSA: Yapay sinir ağları

İE: İşlemci eleman

MLP: Çok katmanlı perseptron

BP: Geri yayılma

GA: Genetik algoritma

LVQ: Öğrenen vektör ölçümü

FNR: Lineer olmayan bulanık regresyon

AAE: Ortalama mutlak hata

LS-SVM: En küçük kareler destek vektör mekanizması

FLAD: Bulanık en küçük mutlak sapmalar regresyonu

Dpp: Doğrusal programlama problemi

TFN: Üçgensel bulanık sayı

QFN: Dörtgensel bulanık sayı

HKO: Hata kareler ortalaması

1. GİRİŞ VE ÖNCEKİ ÇALIŞMALAR

1.1. Giriş

Lineer regresyon analizinde, açıklanan değişken, açıklayıcı değişken veya değişkenlerin lineer formda bir fonksiyonuyla ifade edilir. Lineer olmayan regresyonda bu fonksiyon parametreye göre lineer olmayan yapıdadır ve fonksiyonun biçimsel olarak bilindiği varsayılır. Regresyon parametrelerinin tahmini sıradan en küçük kareler (EKK) yöntemiyle elde edilebilir. Fonksiyonel ilişkinin bilinmemesi durumunda ise bu yöntem kullanılamaz (Genç 1997).

Verilerin her zaman kesin sayılarla ifade edilemediği durumlarda, özellikle veri için niteliksel tanımlamalar getirildiğinde bulanık sayı ifadesi ortaya çıkmakta ve bulanık küme teorisine ihtiyaç duyulmaktadır.

Bu çalışmada verilerin kesin sayı olmaması durumunda bir bağımlı ve bir bağımsız değişken arasındaki ilişkinin ortaya konulması ve tahmin yapılması problemi ele alınmıştır. Bu bağlamda yapay sinir ağları ile bulanık regresyon, LS-SVM'ye dayalı bulanık regresyon ve bulanık en küçük mutlak sapma regresyonu konularında bilgi verilecektir.

Yapay sinir ağları (YSA), beynin fizyolojisinden faydalanılarak oluşturulan bilgi işleme modelleridir. YSA'lar sahip oldukları öğrenme özelliği yardımıyla, girdi ve çıktı arasındaki ilişkiyi, doğrusal olsun veya olmasın, daha önce hiç görülmemiş olaylara uygulayabilmektedir (Yazıcı ve ark. 2006). Böylece, girdi ve çıktı değişkenleri arasındaki ilişkinin biçimi bilinmese de tahmin yapılabilmektedir.

Girdi ve çıktısı bulanık veri setleri için LS-SVM'ye dayalı bulanık lineer regresyon yönteminde hataların en küçük hale getirilmesi amaçlandığından bir optimizasyon problemi olarak düşünülmüştür (Hong ve Hwang 2006).

Bulanık en küçük mutlak sapma regresyonu (FLAD), en küçük kareler yöntemine alternatif olarak sunulan L_1 normuna dayalı bir yöntemdir. Bu yöntem de sonuçta bir optimizasyon problemi haline getirilerek çözüm elde edilmektedir (Chang ve Lee 1994).

Çalışmanın ikinci bölümünde bulanık küme teorisine ilişkin bazı temel kavramlarla, bulanık kümelerde ve bulanık sayılarda işlemler ele alınacaktır.

Üçüncü bölümde doğrusal regresyon analizine giriş yapılacak, modele ilişkin gösterimler, parametre tahmin yöntemleri, çoklu belirleyicilik katsayısı, parametre tahminlerinin güven aralıkları ve bu parametreler için hipotez testleri verilecektir. Dördüncü bölümde de lineer olmayan regresyon analizinde gösterimler ve parametre tahmini için bir algoritma açıklanacaktır.

Beşinci bölümde yapay sinir ağları tanıtılacak ve altıncı bölümde yapay sinir ağları yardımıyla bulanık regresyon tahmini yapmak için bir algoritma verilecektir.

Yedinci bölümde LS-SVM'ye dayalı bulanık regresyon analizi için gerekli tanımlamalarla optimizasyon algoritması açıklanacak ve bulanık lineer ve lineer olmayan regresyon modeli için tahmin elde etme adımları verilecektir.

Sekizinci bölüm doğrusal programlama ve bulanık en küçük mutlak sapma regresyonunu içermektedir.

Dokuzuncu bölümde verilen yöntemler için uygulamalar yapılacak ve elde edilen tahminler grafikler yardımıyla açıklanmaya çalışılacaktır. Sonuç bölümünde ise uygulamalar yardımıyla yukarıda değinilen yöntemler kıyaslanacak, avantaj ve dezavantajları açıklanmaya çalışılacaktır.

1.2. Önceki Çalışmalar

Bulanık lineer regresyon analizi alanında ilk çalışma Tanaka ve ark. (1982) tarafından yapılmıştır. Girdinin kesin sayı çıktının ise bulanık sayı olduğu modelde bulanık katsayıların tahmini için model geliştirilmiştir.

Ishibuchi ve Tanaka (1992), bulanık regresyon için sinir ağlarını kullanarak basit fakat güçlü bir yöntem önermiştir. Bulanık regresyon analizini aralık regresyon analizine indirgeyip, çıktı aralığının alt ve üst sınırlarını iki bağımsız sinir ağının eğitilmesi ile tanımlamışlardır.

Xu ve ark. (2000), bulanık lineer olmayan regresyon analizi için aralık tahmini yapmayı hedeflemişlerdir.

Hong ve Hwang (2006), LS-SVM'ye dayalı bulanık lineer regresyon analizi için bir algoritma önermişler ve lineer olmayan regresyon için de uygulamışlardır.

Sun ve Sun (2003), bir sınıflayıcı ve regresyon tahmin edicisi olarak SVM'yi ele almıştır. Çalışmada bulanık mantıktan yararlanarak genel SVM modelinden çok katmanlı SVM yapısına geçiş açıklanmıştır.

Chang ve Lee (1994), bulanık regresyon analizi için genel bulanık lineer regresyon modeline dayalı bulanık fark sıralama yöntemi geliştirmiştir. Böylece bulanık en küçük mutlak sapma regresyonu oluşturmuşlardır.

Kim ve Kim (2005), bulanık girdi ve çıktıya sahip veride bulanık lineer regresyon modeli yerine bulanık en küçük mutlak sapma regresyonu yöntemini uygulamışlardır.

Dasgupta ve Mishra (2004), ekonometrik modellerde en küçük mutlak sapma regresyonu üzerine inceleme yapmışlardır.

2. BULANIK KÜME KURAMI

Bulanık küme kuramı ilk olarak 1965'te L.A. Zadeh tarafından ele alınmıştır. Sözel ifadelerdeki belirsizliği göstermek için matematiksel bir yöntem olarak tanıtılan bulanık kümeler, klasik küme teorisinin genişletilmesi olarak düşünülebilir (Pedrycz ve Gomido 1998).

Kesin veya klasik küme, elemanları belirli sınırlar içinde yer alan kümedir. Elemanlar kümenin tam olarak üyesidir.

Bulanık küme ve kesin küme arasındaki bağlantıyı görmek için öncelikle kesin kümelerdeki üyelik durumuna bakalım. X, ilgilendiğimiz elemanları içeren evrensel küme olsun. A da, X'in bir alt kümesini göstersin. A'nın, X üzerindeki üyelik fonksiyonu (veya karakteristik fonksiyonu) 0 veya 1 değerini alır. Yani, X'te bir x elemanı A'nın elemanı ise üyelik fonksiyonu değeri 1, diğer durumlarda 0'dır (Xu ve Khosgoftaar 2000).

Bir $A \subseteq X$ kümesinin üyelik derecesi [0,1] aralığı kabul edilirse A kümesi bulanık küme olarak adlandırılır ve $\widetilde{A}$ ile gösterilir. $\widetilde{A}$ bulanık kümesi

$$\widetilde{A} = \{(x, \mu_{\widetilde{A}}(x)) \mid x \in X\} \tag{2.1}$$

şeklinde tanımlanmıştır ve üyelik derecesi

$$\mu_{\widetilde{A}}(x) : X \to [0,1] \tag{2.2}$$

fonksiyonu ile tanımlanır. $\widetilde{A}$ bulanık kümesi, $X = \{x_1, x_2, ..., x_n\}$ olmak üzere,

$$\widetilde{A} = \mu_{\widetilde{A}}(x_1)/x_1 + \mu_{\widetilde{A}}(x_2)/x_2 + ... + \mu_{\widetilde{A}}(x_n)/x_n \tag{2.3}$$

olarak tanımlanır. Burada "/" işareti x elemanının $\widetilde{A}$ bulanık kümesine ait olma (üyelik) derecesini gösteren ayrım işaretidir. "+" işareti Boole cebri anlamında birleşme işaretidir.

2.1. Bulanık Kümelerde Temel Kavramlar

Bulanık kümelerde işlem yapabilmek için bazı temel kavramların açıklanmasına ihtiyaç duyulmaktadır. Burada $X \subset \Re$ 'dir.

- Destek

$\widetilde{A}$ bulanık kümesinin desteği,

$$Supp(\widetilde{A}) = \{x \in X \mid \mu_{\widetilde{A}}(x) > 0\} \tag{2.4}$$

olarak tanımlanır. Üyelik derecesi sıfırdan büyük olan elemanların oluşturduğu kümedir.

- α - kesme (α -seviye) kümesi

$\widetilde{A}$ bulanık kümesinin zayıf α -kesme kümesi,

$$\widetilde{A}_{\alpha} = \{x \in X \mid \mu_{\widetilde{A}}(x) \geq \alpha\} \tag{2.5}$$

ve $\widetilde{A}$ bulanık kümesinin güçlü α -kesme kümesi,

$$\widetilde{A}'_{\alpha} = \{x \in X \mid \mu_{\widetilde{A}}(x) > \alpha\} \tag{2.6}$$

biçiminde ifade edilir ve burada $\alpha \in (0,1]$ 'dir.

- Dışbükeylik

$\widetilde{A}$ bulanık kümesi için,

$$\mu_{\widetilde{A}}(\lambda x_1 + (1-\lambda)x_2) \geq \lambda\mu_{\widetilde{A}}(x_1) + (1-\lambda)\mu_{\widetilde{A}}(x_2) \quad \forall x_1, x_2 \in X \text{ ve } \lambda \in [0,1] \tag{2.7}$$

ifadesi sağlanırsa $\widetilde{A}$ bulanık kümesi dışbükeydir.

- Normallik

$\widetilde{A}$ bulanık kümesi ancak ve ancak

$$\sup_x \mu_{\widetilde{A}}(x) = 1 \tag{2.8}$$

eşitliği sağlandığında normaldir. Yani üyelik fonksiyonunun en büyük değeri 1 olduğunda bulanık küme normaldir.

- Yükseklik

$\widetilde{A}$ bulanık kümesinin yüksekliği,

$$h(\widetilde{A}) = \sup_{x \in X} \mu_{\widetilde{A}}(x) \tag{2.9}$$

ile tanımlanır (Yapıcı 2000).

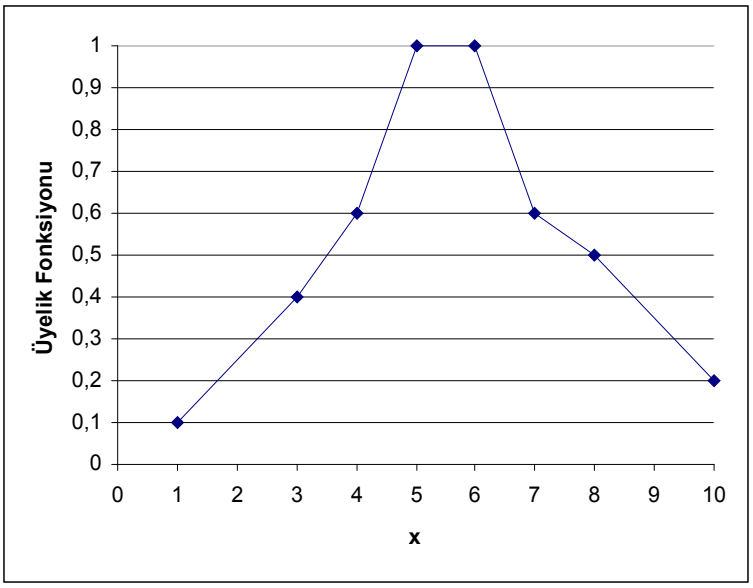
Örnek 2.1. $\widetilde{A} = \{0.1/1 + 0.4/3 + 0.6/4 + 1.0/5 + 1.0/6 + 0.6/7 + 0.5/8 + 0.2/10\}$ bulanık kümesi $X = \{1,2,3,4,5,6,7,8,9,10\}$ evrensel kümesi üzerinde tanımlıdır. Bulanık kümelerde temel kavramları $\widetilde{A}$ kümesi üzerinde aşağıdaki biçimde gösterilebilir

$Supp(\widetilde{A}) = \{1,3,4,5,6,7,8,10\}$

$\alpha = 0.4$ için;

Zayıf α -seviye kesme kümesi $\widetilde{A}_\alpha = \{3,4,5,6,7,8\}$

Güçlü α -seviye kesme kümesi $\widetilde{A}_\alpha = \{4,5,6,7,8\}$ olarak bulunur.



Şekil 2.1.

$\lambda = 0.5$

$\mu_{\widetilde{A}}(0.5 \times 4 + 0.5 \times 10) \geq 0.5 \times 0.6 + 0.5 \times 0.2$

$\mu_{\widetilde{A}}(7) \geq 0.4$

$0.6 \geq 0.4$ olduğundan $\widetilde{A}$ kümesi dışbükeydir (Şekil 2.1).

$\sup_x \mu_{\tilde{A}}(x) = 1$ olduğundan $\tilde{A}$ kümesi normaldir.

$$|\tilde{A}| = \sum_{x \in X} \mu_{\tilde{A}}(x) = 0.1 + 0.4 + 0.6 + 1 + 1 + 0.6 + 0.5 + 0.2$$
$$= 4.4$$

$$h(\tilde{A}) = \sup_{x \in X} \mu_{\tilde{A}}(x) = 1$$

dır.

2.2. Bulanık Kümelerde Temel İşlemler

Bulanık kümelerde işlemler üyelik fonksiyonları yardımı ile tanımlanır. $\tilde{A}$ ve $\tilde{B}$ iki bulanık küme olmak üzere Zadeh (1965) tarafından tanımlanan cebirsel işlemler aşağıda verilmiştir.

2.2.1. Kesişim

$\tilde{A}$ ve $\tilde{B}$ bulanık kümelerinin kesişimi $\tilde{A} \cap \tilde{B}$ nin üyelik fonksiyonu,

$$\mu_{\tilde{A} \cap \tilde{B}}(x) = \mu_{\tilde{A}}(x) \wedge \mu_{\tilde{B}}(x) = \min\{\mu_{\tilde{A}}(x), \mu_{\tilde{B}}(x) \quad , x \in X\} \tag{2.10}$$

ile tanımlanır. Burada "$\wedge$" işareti bulanık kümelerde kesişim işlemini göstermektedir.

2.2.2. Birleşim

$\tilde{A}$ ve $\tilde{B}$ bulanık kümelerinin birleşimi $\tilde{A} \cup \tilde{B}$ nin üyelik fonksiyonu,

$$\mu_{\widetilde{A}\cup\widetilde{B}}(x) = \mu_{\widetilde{A}}(x) \vee \mu_{\widetilde{B}}(x) = \max\{\mu_{\widetilde{A}}(x), \mu_{\widetilde{B}}(x) \quad , x \in X\} \tag{2.11}$$

ile tanımlanır. Burada "$\vee$" işareti bulanık kümelerde birleşme işlemini göstermektedir.

2.2.3. Tümleme

$\widetilde{A}$ bulanık kümesinin tümleyeni $\widetilde{A}^c$ ile gösterilsin. $\widetilde{A}^c$ nin üyelik fonksiyonu

$$\mu_{\widetilde{A}^c}(x) = 1 - \mu_{\widetilde{A}}(x) \tag{2.12}$$

şeklinde ifade edilir.

2.2.4. Kapsama

$\widetilde{A}$ ve $\widetilde{B}$ bulanık kümelerinde $\widetilde{A} \subset \widetilde{B}$ ise bunun üyelik fonksiyonu,

$$\mu_{\widetilde{A}}(x) \leq \mu_{\widetilde{B}}(x) \ , \ x \in X \tag{2.13}$$

eşitsizliği ile ifade edilir.

2.2.5. Eşitlik

$\widetilde{A}$ ve $\widetilde{B}$ bulanık kümeleri eşit ise $\widetilde{A} = \widetilde{B}$ biçiminde gösterilir ve üyelik fonksiyonu

$$\mu_{\widetilde{A}}(x) = \mu_{\widetilde{B}}(x) \ , \ x \in X \tag{2.14}$$

biçiminde ifade edilir.

2.2.6. Cebirsel çarpım

$\widetilde{A}$ ve $\widetilde{B}$ bulanık kümelerinin cebirsel çarpımı $\widetilde{A} \cdot \widetilde{B}$ nin üyelik fonksiyonu,

$$\mu_{\widetilde{A}\cdot\widetilde{B}}(x) = \mu_{\widetilde{A}}(x) \cdot \mu_{\widetilde{B}}(x) \quad , x \in X \tag{2.15}$$

eşitliği ile ifade edilir.

2.2.7. Cebirsel toplam

$\widetilde{A}$ ve $\widetilde{B}$ bulanık kümelerinin cebirsel toplamı $\widetilde{A} \oplus \widetilde{B}$ nin üyelik fonksiyonu,

$$\mu_{\widetilde{A}\oplus\widetilde{B}}(x) = \mu_{\widetilde{A}}(x) + \mu_{\widetilde{B}}(x) - \mu_{\widetilde{A}}(x) \cdot \mu_{\widetilde{B}}(x) \quad , x \in X \tag{2.16}$$

eşitliği ile ifade edilir.

2.2.8. Fark

$\widetilde{A}$ ve $\widetilde{B}$ bulanık kümelerinin farkı $\widetilde{A} - \widetilde{B} = \widetilde{A} \cap \widetilde{B}^c$, ($\widetilde{B}^c$, $\widetilde{B}$ nin tümleyenini göstermek üzere) biçimindedir ve üyelik fonksiyonu,

$$\mu_{\widetilde{A}\cap\widetilde{B}^c}(x) = \min\{\mu_{\widetilde{A}}(x), \mu_{\widetilde{B}^c}(x)\} = \min\{\mu_{\widetilde{A}}(x), 1 - \mu_{\widetilde{B}}(x)\} \tag{2.17}$$

eşitliği ile ifade edilir.

2.3. Bazı İleri Bulanık Küme İşlemleri

Bulanık birleşme operatörlerine t-conorm ($\oplus$)(s-norm) operatörü, bulanık kesişme operatörlerine t- norm (*) operatörü denir.

2.3.1. Bazı t-conorm operatörleri

- Sınırlı toplam

$\widetilde{A}$ ve $\widetilde{B}$ bulanık kümelerinin cebirsel toplamı $\widetilde{A} \oplus \widetilde{B}$ nin üyelik fonksiyonu,

$$\mu_{\widetilde{A}\oplus\widetilde{B}}(x) = \min\{1, \mu_{\widetilde{A}}(x) + \mu_{\widetilde{B}}(x)\} \ , \ x \in X \tag{2.18}$$

şeklinde ifade edilir.

- Güçlü toplam

$\widetilde{A}$ ve $\widetilde{B}$ bulanık kümelerinin üyelik fonksiyonu ile bulanık birleşimi, üyelik fonksiyonu aşağıda gösterilen bulanık kümedir.

$$\mu_{\widetilde{A}\oplus\widetilde{B}}(x) = \begin{cases} \mu_{\widetilde{A}}(x), & \mu_{\widetilde{B}}(x) = 0 \\ \mu_{\widetilde{B}}(x), & \mu_{\widetilde{A}}(x) = 0 \\ 1, & \mu_{\widetilde{A}}(x), \mu_{\widetilde{B}}(x) > 0 \end{cases} \tag{2.19}$$

2.3.2. Bazı t-norm operatörleri

- Sınırlı çarpım

$\widetilde{A}$ ve $\widetilde{B}$ bulanık kümelerinin cebirsel çarpımı $\widetilde{A} \cdot \widetilde{B}$ nin üyelik fonksiyonu,

$$\mu_{\widetilde{A}\cdot\widetilde{B}}(x) = \max\{0, \mu_{\widetilde{A}}(x) + \mu_{\widetilde{B}}(x) - 1\} \ , \ x \in X \tag{2.20}$$

şeklinde ifade edilir.

- Güçlü çarpım

$\widetilde{A}$ ve $\widetilde{B}$ bulanık kümelerinin üyelik fonksiyonu ile bulanık kesişimi, üyelik fonksiyonu aşağıda gösterilen bulanık kümedir.

$$\mu_{\widetilde{A}\cdot\widetilde{B}}(x) = \begin{cases} \mu_{\widetilde{A}}(x), & \mu_{\widetilde{B}}(x) = 1 \\ \mu_{\widetilde{B}}(x), & \mu_{\widetilde{A}}(x) = 1 \\ 0, & \mu_{\widetilde{A}}(x), \mu_{\widetilde{B}}(x) < 1 \end{cases} \tag{2.21}$$

2.4. Üyelik Fonksiyonları

Üyelik fonksiyonu, X evrensel kümesine ait bir x elemanının $\widetilde{A}$ bulanık kümesine ait olma derecesini gösteren fonksiyondur. Üyelik fonksiyonu, karakteristik fonksiyon olarak da adlandırılır.

$$\mu_{\widetilde{A}}(x) : X \to [0,1]$$

eşitliği ile ifade edilir.

2.4.1. Üyelik fonksiyonları biçimleri

Üyelik fonksiyonu biçimleri dört ana grupta toplanmaktadır.

1. Deneysel karar vermeye dayalı üyelik fonksiyonları

Bu tip üyelik fonksiyonlarına örnek olarak Zadeh'in unimodal fonksiyonları verilebilir.

$$\mu(x) = \begin{cases} 1/\{1+[(x-25)/25]^2\}, & x > 25 \\ 1 \qquad\qquad\qquad\quad , & x \le 25 \end{cases} \tag{2.22}$$

ve

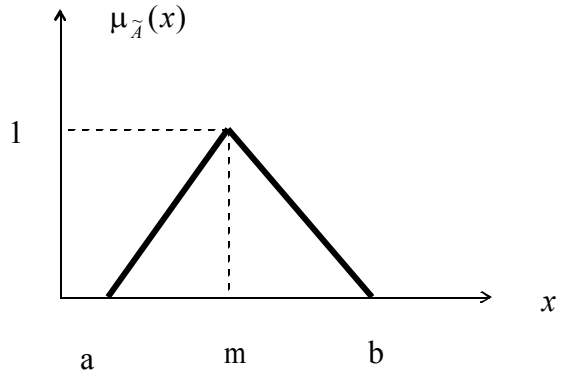
$$\mu(x) = \begin{cases} 1/\{1+[(x-50)/50]^{-2}\}, & x > 50 \\ 1 \qquad\qquad\qquad\quad , & x \le 50 \end{cases} \tag{2.23}$$

2. Güvenilirlik kavramına dayalı üyelik fonksiyonları

Üçgensel üyelik fonksiyonu bu tiptedir.

$$\mu(x) = \begin{cases} 0 & , x < a \\ \dfrac{x-a}{m-a} & , a \le x \le m \\ \dfrac{b-x}{b-m} & , m \le x \le b \\ 0 & , x > b \end{cases} \tag{2.24}$$

Burada üçgensel bulanık sayının gösterimi $\widetilde{A} = (a, m, b)$ şeklindedir.



Şekil 2.1.

Bu grupta örnek olarak Dubois ve Prade'in (1980) LR-tip bulanık sayısı da verilir.

$$\mu(x) = \begin{cases} L((a-x)/\alpha)\ , & x < a \\ R((x-b)/\beta)\ , & x > b \\ 1 & , a \le x \le b \end{cases} \tag{2.25}$$

Burada $L(\cdot)$ ve $R(\cdot)$ referans fonksiyonlarıdır. Üçgensel bulanık sayı

$$L(x) = R(x) = \begin{cases} 1-x & ,0 \le x < 1 \\ 0 & ,d.y. \end{cases} \tag{2.26}$$

olan özel bir LR-tip bulanık sayıdır.

3. Teorik isteklere dayalı üyelik fonksiyonları

Civanlar ve Trussel'in (Yapıcı 2000) fonksiyonu örnek verilebilir.

$$\mu(x) = \begin{cases} ap(x) \ , \ ap(x) \le 1 \\ 0 \quad\quad , \ d.y. \end{cases} \tag{2.27}$$

ile ifade edilmek üzere $a \in [0,1]$ bir parametre ve $p(x)$ olasılık yoğunluk fonksiyonudur.

4. Kişi kavramlarının bir modeli olarak üyelik fonksiyonları

Bu kategoride Dombi'nin (1990) fonksiyonu

$$\mu(x) = (1-s)x^2 / [(1-s)x^2 + s(1-x)^2] \tag{2.28}$$

şeklinde verilir. Burada s, durumun karakteristik değerini belirtir ve $y = \mu(x)$ ve $y = x$ değerlerinin kesişimidir.

Örnek 2.2. Simetrik $\tilde{5}$ üçgensel bulanık sayısının üyelik fonksiyonu,

$$\mu_{\tilde{X}}(x) = \begin{cases} 0 & x < 3 \ ve \ x > 7 \ için \\ (x-3)/2 & 3 \le x \le 5 \ için \\ (7-x)/2 & 5 \le x \le 7 \ için \end{cases}$$

olarak yazılabilir. Aynı sayı için LR-tip üyelik fonksiyonu aşağıdaki gibi verilir.

$$\mu(x) = \begin{cases} L((5-x)/2) \ , \ x < 5 \\ R((x-5)/2) \ , \ x > 5 \\ 1 \quad\quad\quad\quad , \ x = 5 \end{cases}$$

2.5. Bulanık Sayılar

Normal ve dışbükey bir bulanık küme için eğer zayıf α -kesme kümesi kapalı bir aralık ise "bulanık sayı" olarak adlandırılır. Bulanık sayılar, üyelik fonksiyonuna sahiptir. Eğer bulanık bir sayının üyelik fonksiyonu sürekli ise zayıf α -kesme kümesi kapalı bir aralık olacaktır.

Başka bir tanıma göre $\widetilde{A}$ bulanık sayısı , dışbükey, normal bir gerçek düzlem $\Re$ 'de tanımlı $\widetilde{A}$ bulanık kümesidir. Aşağıdaki özellikleri sağlar.

i) $\mu_{\widetilde{A}}(x_0) = 1$ ise (x_0 , $\widetilde{A}$ bulanık sayısının ortalama değeridir) kesinlikle $x_0 \in \Re$ mevcuttur.

ii) $\mu_{\widetilde{A}}(x_0)$ parçalı süreklidir.

$\widetilde{A}$ bulanık sayısının pozitif (negatif) olarak adlandırılabilmesi için $\mu_{\widetilde{A}}(x) = 0, \forall x < 0$ ($\forall x > 0$) olmalıdır (Yapıcı 2000).

2.6. Bulanık Sayılarda İşlemler

2.6.1. Toplama işlemi

$\widetilde{X}$ ve $\widetilde{Y}$ bulanık sayılar olmak üzere toplamları aşağıda verilen yöntemle elde edilir.

$\widetilde{X}$ ve $\widetilde{Y}$ nin α -seviye kesme kümeleri $\widetilde{X}_\alpha = [x_\alpha^L, x_\alpha^U]$ ve $\widetilde{Y}_\alpha = [y_\alpha^L, y_\alpha^U]$ dır. Burada $x_\alpha^L, x_\alpha^U, y_\alpha^L, y_\alpha^U \in \Re$ dir.

$\widetilde{X}$ ve $\widetilde{Y}$ bulanık sayılarının toplamı $\widetilde{Z}$,

$$\widetilde{Z}_\alpha = \widetilde{X}_\alpha (+)\widetilde{Y}_\alpha = [x_\alpha^L + y_\alpha^L, x_\alpha^U + y_\alpha^U], \quad \alpha \in [0,1] \tag{2.29}$$

şeklinde elde edilir.

2.6.2. Çıkarma işlemi

Bulanık sayılarda çıkarma işlemi toplama işlemine benzer.

$\widetilde{X}$ ve $\widetilde{Y}$ nin α -seviye kesme kümeleri $\widetilde{X}_\alpha = [x_\alpha^L, x_\alpha^U]$ ve $\widetilde{Y}_\alpha = [y_\alpha^L, y_\alpha^U]$ dır. Burada $x_\alpha^L, x_\alpha^U, y_\alpha^L, y_\alpha^U \in \Re$ dir.

$\widetilde{X}$ ve $\widetilde{Y}$ bulanık sayılarının farkı $\widetilde{Z}$,

$$\widetilde{Z}_\alpha = \widetilde{X}_\alpha (-)\widetilde{Y}_\alpha = [x_\alpha^L - y_\alpha^U, x_\alpha^U - y_\alpha^L], \quad \alpha \in [0,1] \tag{2.30}$$

şeklinde elde edilir.

2.6.3. Çarpma işlemi

$\widetilde{X}$ ve $\widetilde{Y}$ bulanık sayıları $\widetilde{X} = [x^L, x^U]$ ve $\widetilde{Y} = [y^L, y^U]$ dır. Burada $x^L, x^U, y^L, y^U \in \Re$ dir.

$\widetilde{X}$ ve $\widetilde{Y}$ bulanık sayılarının çarpımı $\widetilde{Z}$,

$$\widetilde{Z} = \widetilde{X}(\cdot)\widetilde{Y} = [\min(x^L y^L, x^L y^U, x^U y^L, x^U y^U), \max(x^L y^L, x^L y^U, x^U y^L, x^U y^U)] \tag{2.31}$$

eşitliği ile ifade edilir.

2.6.4. Bölme işlemi

$\widetilde{X}$ ve $\widetilde{Y}$ bulanık sayıları $\widetilde{X}=[x^L,x^U]$ ve $\widetilde{Y}=[y^L,y^U]$ dır. Burada $x^L,x^U,y^L,y^U\in\Re$ dir.

$\widetilde{X}$ ve $\widetilde{Y}$ bulanık sayılarının oranı $\widetilde{Z}$,

$$\widetilde{Z}=\widetilde{X}(/)\widetilde{Y}=[\min(x^L/y^L,x^L/y^U,x^U/y^L,x^U/y^U),\max(x^L/y^L,x^L/y^U,x^U/y^L,x^U/y^U)] \quad (2.32)$$

eşitliği ile ifade edilir.

2.7. Üçgensel ve Yamuksal Bulanık Sayılar

Bulanık sayılar içinde üçgensel ve yamuksal sayılar sıklıkla kullanılmaktadır. Üçgensel bulanık sayılar

$$\widetilde{X}=(x,\alpha,\beta) \text{ ve } \widetilde{Y}=(y,r,\delta)$$

biçiminde gösterilir. Burada x ve y üyelik fonksiyonu değerleri 1 olan merkezi değerleri, α ve r sol yayılma miktarını, β ve δ ise sağ yayılma miktarını göstermektedir. Sayının simetrik olması durumunda sağ ve sol yayılmalar birbirine eşittir.

Yamuksal bulanık sayılar

$$\widetilde{X}=(x_1,x_2,\alpha,\beta) \text{ ve } \widetilde{Y}=(y_1,y_2,r,\delta)$$

şeklinde gösterilir. Burada da x_1 ve x_2 , $\widetilde{X}$ bulanık sayısının, y_1 ve y_2 , $\widetilde{Y}$ bulanık sayısının üyelik fonksiyonu değerleri 1 olan merkez değerlerinin göstermektedir.

Üçgensel bulanık sayılar bir başka şekilde,

$\widetilde{X} = (x^c, x^l, x^r)$ ve $\widetilde{Y} = (y^c, y^l, y^r)$ biçiminde de gösterilebilir. Burada x^c ve y^c merkez değerlerini, x^l ve y^l ile x^r ve y^r değerleri de sırasıyla bulanık sayının sol ve sağ ayaklarını göstermektedir.

Aynı şekilde yamuksal bulanık sayılar da $\widetilde{X} = (x_1^c, x_2^c, x^l, x^r)$ ve $\widetilde{Y} = (y_1^c, y_2^c, y^l, y^r)$ şeklinde yazılabilir.

$\widetilde{A}$ bulanık sayısının üyelik fonksiyonu simetrik üçgensel ise, aşağıdaki eşitlik $\widetilde{A}$ bulanık sayısının α -seviye kümesi $\left[\widetilde{A}\right]_\alpha$ aralığını hesaplamada kullanılır.

$$\left[A\right]_\alpha = \left[a^C - (1-\alpha)a^W, a^C + (1-\alpha)a^W\right] \tag{2.33}$$

Burada a^C, sayının merkezini, a^W ise yayılma miktarını göstermektedir.

2.7.1. Üçgensel ve yamuksal bulanık sayılar üzerinde bulanık aritmetik

$\widetilde{X} = (x^c, x^l, x^r)$ ve $\widetilde{Y} = (y^c, y^l, y^r)$ olan üçgensel bulanık sayılar üzerinde bulanık aritmetik işlemleri şöyle uygulanır.

1. $\widetilde{Y}$ nin toplamaya göre tersi $-\widetilde{Y} = (-y^c, -y^r, -y^l)$ (2.34)

2. $\widetilde{Y}$ nin çarpmaya göre tersi $\widetilde{Y}^{-1} = (1/y^c, 1/y^r, 1/y^l)$ (2.35)

3. Toplama $\widetilde{X}+\widetilde{Y}=(x^c+y^c, x^l+y^l, x^r+y^r)$ (2.36)

4. Çıkarma $\widetilde{X}$ - $\widetilde{Y}=(x^c-y^c, x^l-y^r, x^r-y^l)$ (2.37)

5. Çarpma $\widetilde{X}>0$ ve $\widetilde{Y}>0$ ise $\widetilde{X}\ (\cdot)\ \widetilde{Y}=(x^c y^c, x^l y^l, x^r y^r)$

$\widetilde{X}<0$ ve $\widetilde{Y}>0$ ise $\widetilde{X}\ (\cdot)\ \widetilde{Y}=(x^c y^c, x^l y^r, x^r y^l)$ (2.38)

$\widetilde{X}<0$ ve $\widetilde{Y}<0$ ise $\widetilde{X}\ (\cdot)\ \widetilde{Y}=(x^c y^c, x^r y^r, x^l y^l)$

6. Bölme $\widetilde{X}>0$ ve $\widetilde{Y}>0$ ise $\widetilde{X}\ (:)\ \widetilde{Y}=(x^c / y^c, x^l / y^r, x^r / y^l)$

$\widetilde{X}<0$ ve $\widetilde{Y}>0$ ise $\widetilde{X}\ (:)\ \widetilde{Y}=(x^c / y^c, x^l / y^l, x^r / y^r)$ (2.39)

$\widetilde{X}<0$ ve $\widetilde{Y}<0$ ise $\widetilde{X}\ (:)\ \widetilde{Y}=(x^c / y^c, x^r / y^l, x^l / y^r)$

7. Skaler Çarpım $a>0$, $a\in\Re$ $\quad a(\cdot)\widetilde{X}=(ax^c, ax^l, ax^r)$

$a<0$, $a\in\Re$ $\quad a(\cdot)\widetilde{X}=(ax^c, ax^r, ax^l)$

$\widetilde{X}=(x_1^c, x_2^c, x^l, x^r)$ ve $\widetilde{Y}=(y_1^c, y_2^c, y^l, y^r)$ şeklinde verilen yamuksal bulanık sayılar üzerine bulanık aritmetik işlemler şu şekilde uygulanır.

1. $\widetilde{Y}$ nin toplamaya göre tersi $-\widetilde{Y}=(-y_2^c, -y_1^c, -y^r, -y^l)$ (2.40)

2. $\widetilde{Y}$ nin çarpmaya göre tersi $\widetilde{Y}^{-1}=(1/y_2^c, 1/y_1^c, 1/y^r, 1/y^l)$ (2.41)

3. Toplama $\widetilde{X}+\widetilde{Y}=(x_1^c+y_1^c, x_2^c+y_2^c, x^l+y^l, x^r+y^r)$ (2.42)

4. Çıkarma $\widetilde{X}$ - $\widetilde{Y}=(x_1^c-y_2^c, x_2^c-y_1^c, x^l-y^r, x^r-y^l)$ (2.43)

5. Çarpma $\widetilde{X} > 0$ ve $\widetilde{Y} > 0$ ise $\widetilde{X}\ (\cdot)\ \widetilde{Y} = (x_1^c y_1^c, x_2^c y_2^c, x^l y^l, x^r y^r)$

$\widetilde{X} < 0$ ve $\widetilde{Y} > 0$ ise $\widetilde{X}\ (\cdot)\ \widetilde{Y} = (x_2^c y_1^c, x_1^c y_2^c, x^l y^r, x^r y^l)$ (2.44)

$\widetilde{X} < 0$ ve $\widetilde{Y} < 0$ ise $\widetilde{X}\ (\cdot)\ \widetilde{Y} = (x_2^c y_2^c, x_1^c y_1^c, x^r y^r, x^l y^l)$

6. Bölme $\widetilde{X} > 0$ ve $\widetilde{Y} > 0$ ise $\widetilde{X}\ (:)\ \widetilde{Y} = (x_1^c / y_2^c, x_2^c / y_1^c, x^l / y^r, x^r / y^l)$

$\widetilde{X} < 0$ ve $\widetilde{Y} > 0$ ise $\widetilde{X}\ (:)\ \widetilde{Y} = (x_2^c / y_2^c, x_1^c / y_1^c, x^l / y^l, x^r / y^r)$ (2.45)

$\widetilde{X} < 0$ ve $\widetilde{Y} < 0$ ise $\widetilde{X}\ (:)\ \widetilde{Y} = (x_2^c / y_1^c, x_1^c / y_2^c, x^r / y^l, x^l / y^r)$

7. Skaler Çarpım $a > 0$, $a \in \Re$ $a(\cdot)\widetilde{X} = (ax_1^c, ax_2^c, ax^l, ax^r)$

$a < 0$, $a \in \Re$ $a(\cdot)\widetilde{X} = (ax_2^c, ax_1^c, ax^r, ax^l)$ (2.46)

3. LİNEER REGRESYON ANALİZİ

Y ile x arasında $\mu_{Y/x} = \beta_0 + \beta_1 x$ gibi bir bağıntının olduğunda gözlemler ile,

$$Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \qquad ,i = 1,2,...,n \tag{3.1}$$

şeklinde ifade edilir.ε_i'lerin sıfır ortalamalı, aynı varyanslı (σ^2), bağımsız gözlenemeyen rasgele değişkenler olduğu varsayıldığında x ile gösterilen değişkene açıklayıcı (bağımsız) değişken, Y ile gösterilen değişkene de bağımlı (açıklanan, tepki) değişkeni denir. Açıklayıcı değişkenin verilmiş x değeri için Y'nin dağılımının beklenen değeri $\mu_{Y/x}$ ile x arasındaki

$$\mu_{Y/x} = g(x;\beta_0,\beta_1) = \beta_0 + \beta_1 x$$

bağıntısına regresyon fonksiyonu denir (Genç 1997). Gözlemler ile ilgili yazılan

$$Y_i = g(x;\beta_0,\beta_1) + \varepsilon_i = \beta_0 + \beta_1 x_i + \varepsilon_i, \qquad i = 1,2,...,n$$

denklemine regresyon denklemi ve bu ε_i'ler üzerinde yapılan varsayımlarla birlikte bu regresyon denklemine regresyon modeli denir. $\beta_0, \beta_1, \sigma^2$ parametrelerine regresyon parametresi denir.

Birden çok açıklayıcı değişken $x_1, x_2, ..., x_p$ ve açıklanan değişken Y olmak üzere çoklu lineer regresyon modeli;

$$Y_i = \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_p x_p + \varepsilon_i, \quad i = 1, 2, \ldots, n \tag{3.2}$$

veya

$$Y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_p x_p + \varepsilon_i, \quad i = 1, 2, \ldots, n \tag{3.3}$$

şeklinde gösterilsin. Matris gösterimiyle;

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

olup

$$\underline{Y} = X\underline{\beta} + \underline{\varepsilon} \tag{3.4}$$

biçiminde yazılabilir. X tasarım matrisi

$$X = \left[\underline{x}_1, \quad \underline{x}_2, \quad \cdots, \quad \underline{x}_p\right] = \begin{bmatrix} \underline{x}^1 \\ \underline{x}^2 \\ \vdots \\ \underline{x}^n \end{bmatrix}$$

olup burada $\underline{x}^i = \left(\underline{x}_{i1}, \underline{x}_{i2}, \cdots, \underline{x}_{ip}\right) \in \Re^p$ bir satır vektörünü göstermektedir (Genç 1997).

Hata Terimi Üzerindeki Varsayımlar **:**

Regresyon modelinin rasgele terimini oluşturan hata terimi üzerinde bazı varsayımlar vardır. Bunlar:

1. $E(\varepsilon_i) = 0, \qquad i = 1, 2, \cdots, n$

2. $Var(\varepsilon_i) = \sigma^2$, $i = 1, 2, \cdots, n$ (hata terimleri sabit varyanslıdır)

3. $Cov(\varepsilon_i, \varepsilon_j) = 0$, $i \neq j$ için (hata terimleri birbiriyle ilişkili değildir)

4. $\underline{\varepsilon}$ vektörü X matrisinin her bir sütun vektörüne diktir.

5. İleride verilecek hipotez testleri için $\underline{\varepsilon} \sim N(\underline{0}, \sigma^2 I)$ 'dır. Hata terimleri normal olasılık dağılımına sahip rasgele değişkenlerdir.

Hata terimi üzerindeki varsayımlardan başka regresyon modeli ile ilgili olarak şu varsayımlar vardır.

1. $rank(X) = p$ 'dir.

2. $X'X$ matrisi tekil olmayan bir matristir.

3. Model doğru olarak belirlenmiştir (model spesifikasyonu).

3.1. Parametre Tahmini

Çoklu lineer regresyon modeli,

$$Y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k + \varepsilon_i \ , \ i = 1, 2, \ldots, n$$

ya da

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ 1 & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

dır. X matrisinde birinci sütunda 1'ler gelmiştir. $p=k+1$ dir. Burada parametre tahmin yöntemlerinden en küçük kareler yöntemi ve en çok olabilirlik yöntemi ele alınacaktır.

3.1.1. EKK (En küçük kareler tahmin edicisi):

Lineer regresyon modelinin matris gösterimiyle,

$$\underline{Y}_{n\times 1} = X_{n\times p}\underline{\beta}_{n\times 1} + \underline{\varepsilon}$$

modelinde $\underline{\beta} \in \Re^p$ parametresinin tahmin edilmesi problemi göz önüne alınsın. En küçük kareler yöntemi ele alındığında,

$$\min_{\underline{\beta}}(\underline{Y} - X\underline{\beta})'(\underline{Y} - X\underline{\beta}) = \min_{\underline{\beta}}(\underline{Y}'\underline{Y} - 2\underline{\beta}'X'\underline{Y} + \underline{\beta}'X'X\underline{\beta}) \qquad (3.5)$$

optimizasyon problemini çözmek için (3.5) eşitliğinin $\underline{\beta}$ 'ya göre türev alınıp sıfıra eşitlenir ve

$$X'X\underline{\hat{\beta}} = X'\underline{Y} \qquad (3.6)$$

normal denklemlerine ulaşılır. X tam sütun ranklı yani

$$rank(X) = rank(X'X) = p$$

olduğundan

$$\underline{\hat{\beta}} = (X'X)^{-1} X'\underline{Y} \tag{3.7}$$

elde edilir (Genç 1997). $\underline{Y}$ örneklem olmak üzere $\underline{Y}$ nin bir fonksiyonu olan $\underline{\hat{\beta}}$, $\underline{\beta}$ için bir tahmin edici olmak üzere bu tahmin ediciye "en küçük kareler tahmin edicisi" denir.

σ^2 ' nin tahmin edicisi; $\hat{\beta}$ 'nın EKK tahmin edicisine göre

$$\hat{\sigma}^2 = \frac{1}{n}(\underline{Y} - X\underline{\hat{\beta}})'(\underline{Y} - X\underline{\hat{\beta}}) \tag{3.8}$$

ya da

$$\begin{aligned} S^2 &= \frac{\underline{Y}'(\mathrm{I} - X(X'X)^{-1}X')\underline{Y}}{n-p} \\ &= \frac{Y'Y - Y'X(X'X)^{-1}X'Y}{n-p} \\ &= \frac{Y'Y - \hat{\beta}'\mathrm{X}'\mathrm{Y}}{n-p} \end{aligned} \tag{3.9}$$

dır. $\underline{\hat{\beta}}$ ve S^2 bağımsızdır.

Hipotez modellerinde

$$\underline{\varepsilon} \sim N(0, \sigma^2 \mathrm{I}) \ , \ \underline{\hat{\beta}} \sim N(\beta, \sigma^2 (X'X)^{-1})$$

$$\frac{(n-p)S^2}{\sigma^2} \sim \chi^2_{(n-p)}$$

dir (Genç 1997).

3.1.2. En çok olabilirlik yöntemi

Hataların normallik varsayımı altında regresyon modelinin matris gösterimi,

$$\underline{Y}_{n\times 1} = X_{n\times p}\underline{\beta}_{n\times 1} + \underline{\varepsilon}_{n\times 1}, \qquad \underline{\varepsilon} \sim N(\underline{0}, \sigma^2 I)$$

olmak üzere $\underline{\beta} \in \Re^p$ ve σ^2 parametrelerini tahmin etmek için en çok olabilirlik yöntemi şu şekilde verilebilir. Olabilirlik fonksiyonu,

$$\ell(\underline{\beta}, \sigma^2; \underline{Y}) = \frac{1}{(\sqrt{2\pi})^{n/2}(\sigma^2)^{n/2}} e^{\frac{-1}{2\sigma^2}(\underline{Y}-X\underline{\beta})'(\underline{Y}-X\underline{\beta})}$$

dır. Olabilirlik fonksiyonunun logaritması,

$$\ln \ell(\underline{\beta}, \sigma^2; \underline{Y}) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma^2) - \frac{1}{2\sigma^2}(\underline{Y} - X\underline{\beta})'(\underline{Y} - X\underline{\beta})$$

$$= -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma^2) - \frac{1}{2\sigma^2}(\underline{Y}'\underline{Y} - 2\underline{\beta}'X'\underline{Y} + \underline{\beta}'X'X\underline{\beta}) \qquad (3.10)$$

ve

$$\frac{\partial \ln \ell}{\partial \underline{\beta}} = -\frac{1}{2\sigma^2}(-2X'\underline{Y} + 2X'X\underline{\beta})$$

$$\frac{\partial \ln \ell}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}((\underline{Y} - X\underline{\beta})'(\underline{Y} - X\underline{\beta}))$$

olmak üzere türevlerin sıfıra eşitlenmesiyle $\underline{\beta} \in \Re^p$ ve σ^2 parametrelerinin En Çok Olabilirlik Tahmin Edicileri (EOTE-MLE),

$$\underline{\hat{\beta}} = (X'X)^{-1}X'\underline{Y} \tag{3.11}$$

ve

$$\hat{\sigma}^2 = \frac{1}{n}(\underline{Y} - X\underline{\hat{\beta}})'(\underline{Y} - X\underline{\hat{\beta}})$$

elde edilir (Genç 1997).

3.1.3. EOTE'nin bazı istatistiksel özelikleri

$\underline{\varepsilon} \sim N(\underline{0}, \sigma^2 I)$ ve $\underline{Y} \sim N(X\underline{\beta}, \sigma^2 I)$ olduğundan $\underline{Y}$ nin bir lineer dönüşümü olan $\underline{\hat{\beta}}$ da normal dağılımlıdır. Böylece,

$$\begin{aligned} E(\underline{\hat{\beta}}) &= E((X'X)^{-1}X'\underline{Y}) \\ &= (X'X)^{-1}X'E(\underline{Y}) \\ &= (X'X)^{-1}X'X\underline{\beta} = \underline{\beta} \end{aligned} \tag{3.12}$$

ve

$$\begin{aligned} \text{cov}(\underline{\hat{\beta}}) &= \text{cov}((X'X)^{-1}X'\underline{Y}) \\ &= (X'X)^{-1}X'\,\text{cov}(\underline{Y})((X'X)^{-1}X')' \\ &= (X'X)^{-1}X'\sigma^2 I(X'X)^{-1}X' \end{aligned}$$

$$= \sigma^2 (X'X)^{-1} \tag{3.13}$$

olmak üzere $\underline{\hat{\beta}} \sim N(\underline{\beta}, \sigma^2 (X'X)^{-1})$ dır. $\underline{\hat{\beta}}$, $\underline{\beta}$ nın yansız bir tahmin edicisidir (Genç 1997).

$\hat{\sigma}^2$ tahmin edicisine gelince,

$$E(\hat{\sigma}^2) = E(\frac{1}{n}(\underline{Y} - X\underline{\hat{\beta}})'(\underline{Y} - X\underline{\hat{\beta}}))$$

$$= \mathrm{E}(\frac{1}{n}(\underline{Y} - X(X'X)^{-1}X'\underline{Y})'(\underline{Y} - X(X'X)^{-1}X'\underline{Y})')$$

$$= \frac{1}{\mathrm{n}} E(\underline{Y}'(I - X(X'X)^{-1}X')\underline{Y}) \tag{3.14}$$

$$= \frac{1}{\mathrm{n}} \left\{ tr\left[(I - X(X'X)^{-1}X')\sigma^2 I\right] + (X\underline{\beta})'(I - X(X'X)^{-1}X')(X\underline{\beta}) \right\}$$

$$= \frac{1}{\mathrm{n}} (n - p)\sigma^2 \tag{3.15}$$

olmak üzere $\hat{\sigma}^2$ yansız değildir. σ^2 nin yansız bir tahmin edicisi

$$S^2 = \frac{(\underline{Y} - X\underline{\hat{\beta}})'(\underline{Y} - X\underline{\hat{\beta}})}{n - p} = \frac{\underline{Y}'(I - X(X'X)^{-1}X')\underline{Y}}{n - p} \tag{3.16}$$

dır ve dağılımı

$$\frac{\underline{Y}'(I - X(X'X)^{-1}X')\underline{Y}}{\sigma^2} = \frac{(n - p)S^2}{\sigma^2} \sim \chi^2_{(n-p)} \tag{3.17}$$

dir (Genç 1997).

$\underline{\varepsilon} \sim N(\underline{0},\sigma^2 I)$ olan basit lineer hipotez modeli için $\underline{\beta}$ parametresinin en çok olabilirlik tahmin edicisi yukarıdaki en küçük kareler tahmin olmak üzere

$$\underline{\hat{\beta}} \sim N(\underline{\beta},\sigma^2 (X'X)^{-1}) \tag{3.18}$$

ve

$$\frac{(n-2)S^2}{\sigma^2} \sim \chi^2_{(n-2)} \tag{3.19}$$

dir (Genç 1997).

3.2. Artıklar

Artık terimi $e_i = Y_i - \hat{Y}_i$ şeklindedir. Buna göre artıkları tablo halinde aşağıdaki biçimde yazabiliriz.

Çizelge 3.1.

Y_i	$\hat{Y}_i$	e_i
Y_1	$\hat{Y}_1$	$Y_1 - \hat{Y}_1$
Y_2	$\hat{Y}_2$	$Y_2 - \hat{Y}_2$
$\vdots$	$\vdots$	$\vdots$
Y_n	$\hat{Y}_n$	$Y_n - \hat{Y}_n$

$\hat{Y}_j = \hat{\mu}_{Y/\underline{x}^j} = x_{j1}\hat{\beta}_1 + x_{j2}\hat{\beta}_2 + \ldots + x_{jp}\hat{\beta}_p = \underline{x}^j\,\underline{\hat{\beta}}$ şeklindedir.

3.3. Çoklu Belirleyicilik Katsayısı

Belirleyicilik katsayıları modelin gözlemlere uyumunun iyiliği veya bağımlı değişkendeki toplam değişimin regresyon modeli tarafından açıklanan kısmı için bir ölçüt olmak üzere, düzeltilmiş belirleyicilik katsayısı regresyondaki açıklayıcı değişken sayısını dikkate almaktadır.

Çoklu belirleyicilik katsayısı

$$R^2 = \frac{SSR}{SST} = \frac{\underline{\hat{\beta}} X' \underline{Y} - n\overline{Y}^2}{Y'Y - n\overline{Y}^2} = \frac{\sum_{i=1}^{n} (\hat{Y}_i - \overline{Y})^2}{\sum_{i=1}^{n} (Y_i - \overline{Y})^2} \tag{3.20}$$

şeklinde hesaplanmaktadır. R^2 değeri 0 ile 1 arasında değer alır. 1'e yakın olması bağımsız değişkenlerin bağımlı değişkeni açıkladığını gösterir. 0'a yakın olması bağımsız değişkenlerin bağımlı değişkeni açıklamadığını ifade eder. R^2 değeri, bağımsız değişkenlerin bağımlı değişkeni açıklama oranıdır.

3.4. Bireysel Parametreler için Güven Aralığı

i . bireysel parametre β_i için $1-\alpha$ ' lık güven aralığı

$$\left(\hat{\beta}_i - t_{1-\alpha/2} s\sqrt{c_{ii}} \, , \hat{\beta}_i + t_{1-\alpha/2} s\sqrt{c_{ii}} \right) \tag{3.21}$$

dır . Burada c_{ii} , $(X'X)^{-1}$ matrisinin $i.$ köşegen elemanıdır.

3.5. Bireysel Parametreler İçin Hipotez Testi

$H_0 : \beta_i = a$ hipotezini

$H_1 : \beta_i \neq a$ hipotezine karşı α anlam düzeyinde test etmek için test istatistiği

$$t_h = \frac{\hat{\beta}_j - a}{s\sqrt{c_{ii}}} \sim t_{n-p}$$

olup t_h , $n-p$ serbestlik dereceli t dağılımına sahiptir.

$$t_T = t_{1-\alpha/2;n-p}$$

şeklinde hesaplanıp $|t_h| > t_T$ ise H_0 hipotezi reddedilir (Genç 1997).

3.6. $\mu_{Y/\underline{x}^j}$ ' nin Tahmini

Açıklayıcı değişkenler üzerinde j. gözlem X matrisinin j. satırı olsun. Açıklayıcı değişkenlerin aldığı bu değerler Y ' nin koşullu dağılımının beklenen değeri olan $\mu_{Y/\underline{x}^j}$ için bir tahmin edici

$$\mu_{Y/\underline{x}^j} = \underline{x}^j \underline{\hat{\beta}} \qquad (3.22)$$

olmak üzere

$$E(\hat{\mu}_{y/\underline{x}^j}) = \underline{x}^j \underline{\beta}$$

$$Var(\hat{\mu}_{y/\underline{x}^j}) = (\underline{x}^j)Cov(\underline{\hat{\beta}})(\underline{x}^j)'$$

$$= \sigma^2(\underline{x}^j)(X'X)^{-1}(\underline{x}^j)'$$

dir. Hipotez modellerinde

$$\hat{\mu}_{y/\underline{x}^j} \sim N(\ \underline{x}^j\ \underline{\beta}\ \ ,\sigma^2(\underline{x}^j)(X'X)^{-1}(\underline{x}^j)')$$

dır (Genç 1997).

3.7. Sonradan Gözlemin Tahmini

$\underline{x}^*$, $p \times 1$ boyutlu bir satır vektörü olmak üzere, açıklayıcı değişkenlerin aldığı değer $\underline{x}^*$ olduğunda $\underline{Y}$ ' nin koşullu beklenen değeri $\mu_{Y/\underline{x}^*}$ için bir tahmin edici yerine $\mu_{Y/\underline{x}^*} = \underline{x}^*\ \underline{\hat{\beta}}$ olacaktır.

Açıklayıcı değişkenlerin $\underline{x}^*$ değerinde alınacak sonradan bir $Y_{\underline{x}^*}$ gözlemi için tahmin olarak $\mu_{Y/\underline{x}^*}$ değeri alınır, ancak bu durumda

$$Var(Y_{\underline{x}^*}) = \sigma^2(1 + \underline{x}^*(X'X)^{-1}(\underline{x}^*)')$$

dır (Genç 1997).

4. LİNEER OLMAYAN REGRESYON ANALİZİ

Lineer olmayan modelde değişkenler arasındaki bağıntı, parametrelerin en az birinin lineer olmayan fonksiyonu biçimindedir. Lineer olmayan modellerin teorisi ve metotları, lineer modeller teorisi ve metotları ile yakından ilişkilidir. Bu bölümde bir bağımsız değişkenli lineer olmayan modelin tanıtımı ve bu modellerde parametre tahmini verilecektir.

4.1. Modelin Tanıtımı ve Bazı Gösterimler

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

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

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

$$Y_t = f(\underline{x}_t, \underline{\theta}) + e_t, \quad t = 1,2,...,n \tag{4.1}$$

biçiminde yazılabilsin. f fonksiyonu bilinmeyen $\underline{\theta} = (\theta_1, \theta_2, ..., \theta_p)'$ $(\underline{\theta} \in \Theta \subset \Re^p)$, parametre vektörünün bileşenlerine göre lineer olmayan bir ifade olduğunda bu modele lineer olmayan model denir (Genç 1997). f fonksiyonuna bazen tepki fonksiyonu da denir. e_t gözlenemeyen hata terimi olmak üzere modelleme düşüncesi çerçevesinde belirtilen özelliklere sahip olduğu varsayılmaktadır. Açıklayıcı değişkenlerin vektörü ile ilgili $\underline{x}_1, \underline{x}_2, ..., \underline{x}_n$ gözlemleri rasgele bir vektörün gözlem değerleri gibi yorumlanmamaktadır. Örneğin, bunlar önceden belirlenen deney noktaları olarak tasarlanan değerler olabilir. Parametre kümesi Θ $(\Theta \subset \Re^p)$ olmak üzere bilinmeyen gerçek parametre değeri $\underline{\theta}^0$ $(\underline{\theta}^0 \in \Theta)$ ile gösterilecektir. Buna göre model esasında,

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

biçiminde olmak üzere, istatistik teorisi açısından matematiksel işlemleri yürütmek için $\underline{\theta}^0$ yerine $\underline{\theta}$ yazılmaktadır. (4.1) modeline atıf yapıldığında $\underline{\theta}$ yerine $\underline{\theta}^0$ gelebilir. Kısaca $f(\underline{x}_t, \underline{\theta})$ 'yı $f_t(\underline{\theta})$ olarak gösterip,

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

olmak üzere (4.1) modeli

$$\underline{Y} = \underline{f}(\underline{\theta}) + \underline{e} \tag{4.2}$$

şeklinde yazılabilir (Genç 1997).

4.2. Parametre Tahmini

Bu kısımda (4.1) ile verilen,

$$Y_t = f(\underline{x}_t, \underline{\theta}) + e_t, \quad t = 1,2,\ldots,n \tag{4.3}$$

lineer olmayan tek değişkenli modelin bilinmeyen $\underline{\theta}^0$ parametresinin tahmini için bazı algoritmalar sunulacaktır. Parametre tahmininde en küçük kareler yöntemi ve en çok olabilirlik yöntemi dikkate alınacaktır. Hatalar için

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

$$\text{Cov}(e_i, e_j) = \begin{cases} 0 & , \ i \neq j \\ \sigma^2 & , \ i = j \end{cases} \quad , \quad i, j = 1,2,\ldots,n \tag{4.4}$$

veya vektör gösterimi ile

$$E(\underline{e}) = \underline{0}$$

$$Cov(\underline{e}) = \sigma^2 I \tag{4.5}$$

olduğu varsayılacaktır. Normallik varsayımı altında

$$\underline{e} \sim N(\underline{0}, \sigma^2 I) \tag{4.6}$$

dağılımlı olacaktır.

İlk önce en küçük kareler yöntemini ele alalım. f biçimsel olarak bilindiğinden

$$Q(\underline{\theta}) = \sum_{t=1}^{n} \left[Y_t - f(\underline{x}_t, \underline{\theta}) \right]^2 \tag{4.7}$$

hata kareler toplamı minimum olacak şekilde $\underline{\theta}$ parametre değerini belirlemek en küçük kareler yöntemine göre en iyi tepki fonksiyonunu bulmak demektir (Genç 1997). Vektör gösterimi ile (4.7) eşitliği

$$Q(\underline{\theta}) = \left[\underline{Y} - \underline{f}(\underline{\theta}) \right]' \left[\underline{Y} - \underline{f}(\underline{\theta}) \right] \tag{4.8}$$

şeklinde yazılabilir.

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

$$F(\underline{\theta}) = \left(\frac{\partial}{\partial \theta_j} f(\underline{x}_t, \underline{\theta}) \right) = \begin{bmatrix} \frac{\partial}{\partial \theta_1} f(\underline{x}_1, \underline{\theta}) & \frac{\partial}{\partial \theta_2} f(\underline{x}_1, \underline{\theta}) & \cdots & \frac{\partial}{\partial \theta_p} f(\underline{x}_1, \underline{\theta}) \\ \frac{\partial}{\partial \theta_1} f(\underline{x}_2, \underline{\theta}) & \frac{\partial}{\partial \theta_2} f(\underline{x}_2, \underline{\theta}) & \cdots & \frac{\partial}{\partial \theta_p} f(\underline{x}_2, \underline{\theta}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial}{\partial \theta_1} f(\underline{x}_n, \underline{\theta}) & \frac{\partial}{\partial \theta_1} f(\underline{x}_1, \underline{\theta}) & \cdots & \frac{\partial}{\partial \theta_p} f(\underline{x}_n, \underline{\theta}) \end{bmatrix} \tag{4.9}$$

matrisi $\underline{f}(\underline{\theta})$ fonksiyonunun türev (jakobiyen) matrisi olmak üzere (4.8) ile verilen hata kareler toplamının bilinmeyen $\underline{\theta}$ parametre vektörüne göre türevi

$$\begin{aligned} \frac{\partial}{\partial \underline{\theta}'} Q(\underline{\theta}) &= -2\left[\underline{Y} - \underline{f}(\underline{\theta})\right]' F(\underline{\theta}) \\ &= -2F'(\underline{\theta})\left[\underline{Y} - f(\underline{\theta})\right] \end{aligned} \tag{4.10}$$

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

$$\min_{\underline{\theta} \in \Theta} Q(\underline{\theta}) = Q(\hat{\underline{\theta}}) \tag{4.11}$$

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

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

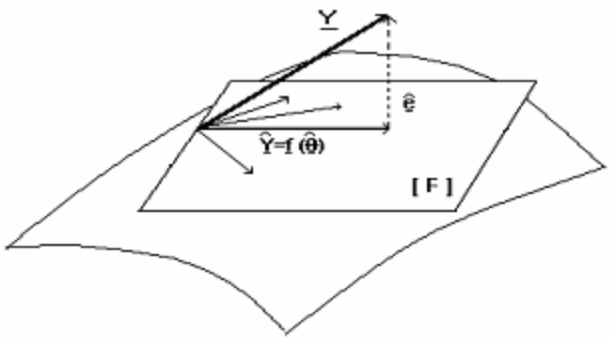
dır. Bu durumda (4.10) ifadesinden

$$F'(\hat{\underline{\theta}})\left[\underline{Y} - \underline{f}(\hat{\underline{\theta}})\right] = 0 \qquad (4.12)$$

olacağından

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

artık vektörü, $\underline{\theta} = \hat{\underline{\theta}}$ noktasında $\underline{f}(\underline{\theta})$ 'nın jakobiyen matrisi $F(\underline{\theta})$ 'nın sütun uzayına diktir (Şekil 4.1) (Genç 1997):



Şekil 4.1. $\underline{Y}$ vektörünün $F(\underline{\theta})$ sütun uzayına dik izdüşümü

Parametre tahmini yapmak için en çok olabilirlik yöntemi şöyle verilir. (4.1)'deki modelde e_t, $t = 1,2,...,n$ hata terimlerinin bağımsız ve her birinin sıfır ortalamalı bilinmeyen σ^2 varyanslı normal dağılıma sahip olduğu varsayılsın. Bu taktirde,

$$Y_t \sim N\left(f\left(\underline{x}_t, \underline{\theta}\right), \sigma^2\right) \quad , \quad t=1,2,...,n$$

dağılımına sahip olup, olasılık yoğunluk fonksiyonu

$$f\left(y_t; \underline{\theta}, \sigma^2\right) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2\sigma^2}\left[Y_t - f\left(\underline{x}_t, \underline{\theta}\right)\right]^2}, -\infty < y_t < \infty$$

olacaktır. $Y_1, Y_2, ..., Y_n$ rasgele değişkenlerinin olabilirlik fonksiyonu, $\underline{Y} = \left(Y_1, Y_2, ..., Y_n\right)'$ olmak üzere,

$$L\left(\underline{\theta}, \sigma^2; \underline{Y}\right) = \prod_{t=1}^{n} n\left(Y_t; \underline{\theta}, \sigma^2\right)$$

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

şeklindedir. $L\left(\underline{\theta}, \sigma^2; \underline{Y}\right)$ fonksiyonunu maksimum yapan bilinmeyen $\underline{\theta}$ ve σ^2 parametrelerinin tahmin edilmesi istenmektedir. $L\left(\underline{\theta}, \sigma^2; \underline{Y}\right)$ nin maksimize edilmesi yerine bunun logaritması alınarak elde edilen $log\, L\left(\underline{\theta}, \sigma^2; \underline{Y}\right)$ nin maksimize edilmesiyle

$$\sum_{t=1}^{n}\left[Y_t - f\left(\underline{x}_t, \underline{\theta}\right)\right]\frac{\partial f\left(\underline{x}_t, \underline{\theta}\right)}{\partial \underline{\theta}'} = 0$$

$$-n\sigma^2 + \sum_{t=1}^{n}\left[Y_t - f\left(\underline{x}_t, \underline{\theta}\right)\right]^2 = 0$$

denklem sistemine ulaşılır. Bu denklem sistemindeki birinci denklemin sol tarafı (4.10) daki ifadenin aynısıdır. Dolayısıyla $\underline{\theta}$ nın en küçük kareler yöntemi ile elde edilen tahmini en çok olabilirlik yöntemi ile elde edilen tahminine eşittir. $\underline{\theta}$ nın en çok olabilirlik yöntemi ile elde edilen tahmini $\underline{\hat{\theta}}$ olmak üzere ikinci denklemden σ^2 nin tahmini,

$$\hat{\sigma}^2 = \frac{\sum_{t=1}^{n}\left[Y_t - f\left(\underline{x}_t, \underline{\hat{\theta}}\right)\right]^2}{n} \qquad (4.13)$$

dir.

(4.11) deki minimizasyon probleminde Q, $\underline{\theta}$ 'nın lineer olmayan bir fonksiyonudur. Lineer olmayan bu fonksiyonun minimizasyonu ile ilgili kullanılan optimizasyon algoritmasının ismine bağlı olarak tahmin yöntemi de bu ismi almaktadır. Burada parametre tahmin yöntemi olarak Gauss-Newton algoritmasına değinilecektir.

Algoritmalar $f(\underline{x}_t, \underline{\theta})$ tepki fonksiyonunun Taylor serisine açılımının yapılması ve gradientinin hesaplanması civarında toplanmıştır. Nümerik hesaplamalar açısından amaç, işlemlerin yakınsama hızını artırmak ve ardışık işlem sayısını azaltmaktır. Bunun için ardışık işlem adımlarına iyi bir noktadan başlamak gerekir.

4.2.1 Başlangıç noktasının seçimi

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

Başlangıç noktasının seçimi için aşağıdaki yöntemlerden birisi uygulanabilir:

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

ii- Gözlem sonucu bilinen Y_t ve $\underline{x}_t$ değerlerini modelde yerine yazarak elde edilen denklem sisteminden parametre sayısı kadar bir alt sistem alınır. Alınan bu denklem sistemi çözülerek $\underline{\theta}_0$ başlangıç noktası bulunur.

t = 1,2,...,n için n tane Y_t , $\underline{x}_t$ gözlemi p gruba ayrılır, herbir grup için bulunan Y 'lerin $\overline{Y}_j$ (j = 1,2,...,p) ortalaması ile f değerlerinin $\bar{f}_j(\underline{\theta})$ (j = 1,2,...,p) ortalaması yardımıyla,

$$\overline{Y}_j = \bar{f}_j(\underline{\theta}) \quad , \; j = 1,2,...,p$$

denklem sistemi oluşturulur. Denklem sisteminden bulunan $\underline{\theta}$ başlangıç değer olarak alınır.

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

Bir tahmin yöntemi yardımıyla lineer olmayan bir model için parametre tahmini yapılmak istendiğinde ardışık işlem adımına $\underline{\theta}_0$ başlangıç noktası ile başlanır. Yakınsama sağlandığında ardışık işlem adımlarının durdurulması gerekir (Genç 1997).

4.2.2. Durdurma kuralı

$\varepsilon > 0$ ve $\tau > 0$ keyfi sabitler olsun. Genellikle $\tau = 10^{-3}$ ve $\varepsilon = 10^{-5}$ önerilir. Aşağıdaki kuralın her ikisi de aynı anda sağlandığında ardışık işlem adımlarına son verilir,

$$\left\| \underline{\theta}_i - \underline{\theta}_{i-1} \right\| < \varepsilon \left(\left\| \underline{\theta}_{i-1} \right\| + \tau \right) \tag{4.14}$$

ve

$$\left| Q\left(\underline{\theta}_i\right) - Q\left(\underline{\theta}_{i-1}\right) \right| < \varepsilon \left(Q\left(\underline{\theta}_{i-1}\right) + \tau \right) \tag{4.15}$$

eşitsizlikleri sağlandığında $\underline{\theta}_0 \to \underline{\theta}_1 \to \underline{\theta}_2 \to \ldots \to \underline{\theta}_i \to \ldots$ ardışık işlemi durdurulur (Genç 1997).

4.2.3. Gauss- Newton Yöntemi

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

$$f\left(.,\underline{\theta}\right) = f\left(.,\underline{\theta}^0\right) + \frac{\partial f\left(.,\underline{\theta}^*\right)}{\partial \underline{\theta}'}\left(\underline{\theta} - \underline{\theta}^0\right) \tag{4.16}$$

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

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

yazılır. Vektör gösterimi ile

$$\underline{f}(\underline{\theta}) \cong \underline{f}(\underline{\theta}^0) + \left. F(\underline{\theta})\right|_{\underline{\theta}=\underline{\theta}^0} (\underline{\theta} - \underline{\theta}^0) \tag{4.17}$$

dır. Hataların kareleri toplamının minimizasyonunda geçen (4.12) ifadesinde $\underline{\theta} = \underline{\theta}^0$ için f yerine (4.17) nin göz önüne alınmasıyla,

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

olur. O halde

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

olup

$$F'\left(\underline{\theta}^0\right)F\left(\underline{\theta}^0\right)\left(\underline{\theta} - \underline{\theta}^0\right) = F'\left(\underline{\theta}^0\right)\left[\underline{Y} - \underline{f}\left(\underline{\theta}^0\right)\right] \tag{4.18}$$

elde edilir. Elde edilen (4.18) eşitliğine normal denklem denir. Normal denklemlerden,

$$\underline{\theta} - \underline{\theta}^{0} = \left[F'\left(\underline{\theta}^{0}\right) F\left(\underline{\theta}^{0}\right) \right]^{-1} F'\left(\underline{\theta}^{0}\right) \left[\underline{Y} - \underline{f}\left(\underline{\theta}^{0}\right) \right] \tag{4.19}$$

olmak üzere,

$$\underline{d}_{0} = \left[F'\left(\underline{\theta}^{0}\right) F\left(\underline{\theta}^{0}\right) \right]^{-1} F'\left(\underline{\theta}^{0}\right) \left[\underline{Y} - \underline{f}\left(\underline{\theta}^{0}\right) \right] \tag{4.20}$$

gösterimi ile

$$\underline{\theta} = \underline{\theta}^{0} + \underline{d}_{0} \tag{4.21}$$

yazılır.

$\underline{\theta}^{0}$ bilinmediğinden böyle elde edilen $\underline{\theta}$ 'nın bir anlamı yoktur. Ancak, $\underline{\theta}^{0}$ yerine bir $\underline{\theta}_{0}$ başlangıç değerinden başlayıp $\underline{d}_{0}$ vektörü yönünde $\lambda_{0}\underline{d}_{0}$ $(0 \le \lambda_{0} \le 1)$ adımına dayalı $\underline{\theta}_{1} = \underline{\theta}_{0} + \lambda_{0}\underline{d}_{0}$ değeri ve benzer düşünce ile

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

yani $\underline{\theta}_{0} \to \underline{\theta}_{1} \to \underline{\theta}_{2} \to \ldots \to \underline{\theta}_{i} \to \ldots$ değerleri hesaplanabilir. Her yinelemedeki λ_{i} $(0 \le \lambda_{i} \le 1 \ , \ i = 0,1,2,\ldots)$ değerleri $Q(\underline{\theta}_{i}) \le Q(\underline{\theta}_{i-1})$ olacak şekilde seçilir. Bu şekilde oluşturulan ardışık yineleme yöntemi Gauss-Newton algoritması olarak bilinir (Genç 1997). Özet olarak bu algoritma bir $\underline{\theta}_{0}$ başlangıç değerinden başlayıp belirlenen bir durdurma kuralına göre,

$$\underline{\theta}_{i} = \underline{\theta}_{i-1} + \lambda_{i-1}\underline{d}_{i-1} \quad , \quad i = 1,2,\ldots. \tag{4.22}$$

değerlerini ardışık olarak hesaplar. Burada,

$$\underline{d}_i = \left[F'(\underline{\theta}_i)F(\underline{\theta}_i)\right]^{-1} F'(\underline{\theta}_i)\left[\underline{Y} - \underline{f}(\underline{\theta}_i)\right] \quad , \qquad \text{i} = 0,1,2,... \tag{4.23}$$

dır ve λ_i $(0 \le \lambda_i \le 1 \ , \ i = 0,1,2,\ldots)$ değeri $Q(\underline{\theta}_{i+1}) \le Q(\underline{\theta}_i)$ olacak şekilde ayrıca belirlenen bir kurala göre seçilir. Bu çalışmada,

$$\lambda = 1, \ \frac{1}{2}, \ \frac{1}{4}, \frac{1}{8}, \ldots$$

değerlerinden soldan sağa doğru

$$Q(\underline{\theta}_i + \lambda \underline{d}_i) \le Q(\underline{\theta}_i), \qquad i = 0,1,2,...$$

eşitsizliğini sağlayan ilk sayı λ_i olarak seçilmiştir.

(4.22) deki λ değerlerine adım uzunluğu, $\underline{d}$ vektörlerine de yön vektörü denir.

5. YAPAY SİNİR AĞLARI

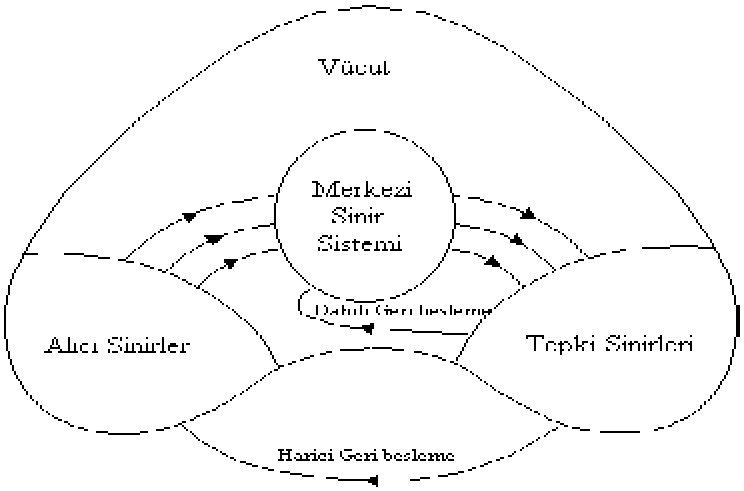
5.1. Giriş

Yapay Sinir Ağları (YSA), beynin fizyolojisinden yararlanılarak oluşturulan bilgi işleme modelleridir. Literatürde 100'den fazla yapay sinir ağı modeli vardır . Bazı bilim adamları, beynimizin güçlü düşünme, hatırlama ve problem çözme yeteneklerini bilgisayara aktarmaya çalışmışlardır. Bazı araştırmacılar ise, beynin fonksiyonlarını kısmen yerine getiren bir çok modeli oluşturmaya çalışmışlardır.

YSA'ların öğrenme özelliği, araştırmacıların dikkatini çeken en önemli özelliklerden birisidir. Çünkü herhangi bir olay hakkında girdi ve çıktılar arasındaki ilişkiyi, doğrusal olsun veya olmasın, elde bulunan mevcut örneklerden öğrenerek daha önce hiç görülmemiş olaylara uygulayabilmektedir. Önceki örneklerden çağrışım yaparak ilgili olaya çözümler üretebilme özelliği YSA'lardaki zeki davranışın da temelini teşkil eder.

5.1.1. Biyolojik nöron yapısı

Biyolojik sinir sistemi, merkezinde sürekli olarak bilgiyi alan, yorumlayan ve uygun bir karar üreten beynin (merkezi sinir ağı) bulunduğu 3 katmanlı bir sistem olarak açıklanmaktadır. Alıcı sinirler (receptor), organizma içerisinden ya da dış ortamlardan algıladıkları uyarıları, elektriksel sinyallere dönüştürerek beyne iletirler. Tepki sinirleri (effector) ise, beyinin ürettiği elektriksel sinyalleri organizma çıktısı olarak uygun tepkilere dönüştürür. Şekil 5.1'de biyolojik sinir sisteminin blok şeması görülmektedir.



Şekil 5.1. Biyolojik sinir sisteminin blok gösterimi

Sinir hücreleri nöron olarak bilinir. Nöron, özellikle beyin olmak üzere sinir sisteminin temel birimidir ve başlıca üç kısımdan oluşur. Bunlar:

- Gövde (cell body)
- Gövdeye gire sinyal alıcı lifler (dendrit),
- Gövdeden çıkan sinyal iletici lifler (axon).

YSA'lar, insan beyninin çalışma prensibi örnek alınarak geliştirilmeye çalışılmıştır ve aralarında yapısal olarak bazı benzerlikler vardır. Bu benzerlikler Çizelge 5.1'de verilmiştir.

Çizelge 5.1. Sinir sistemi ile YSA'nın benzerlikleri

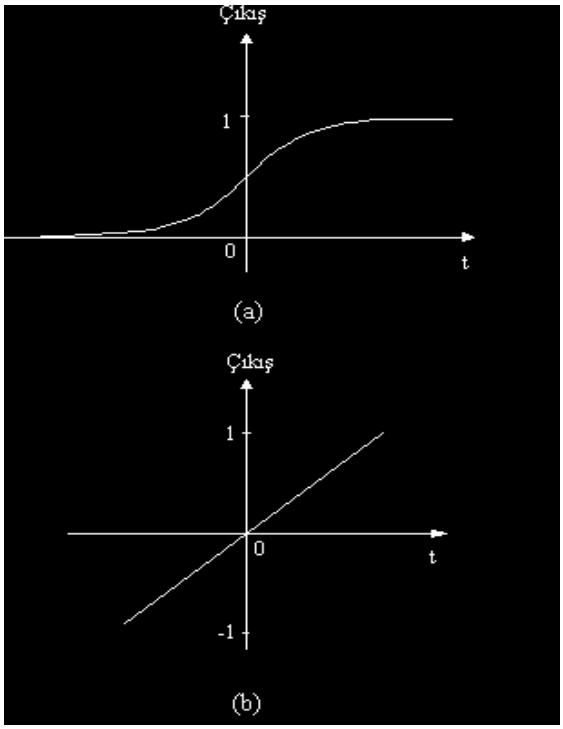
SİNİR SİSTEMİ	YSA SİSTEMİ
Neuron	İşlem elemanı
Dendrit	Toplama fonksiyonu
Hücre gövdesi	Transfer fonksiyonu
Aksonlar	Eleman çıkışı
Sinapslar	Ağırlıklar

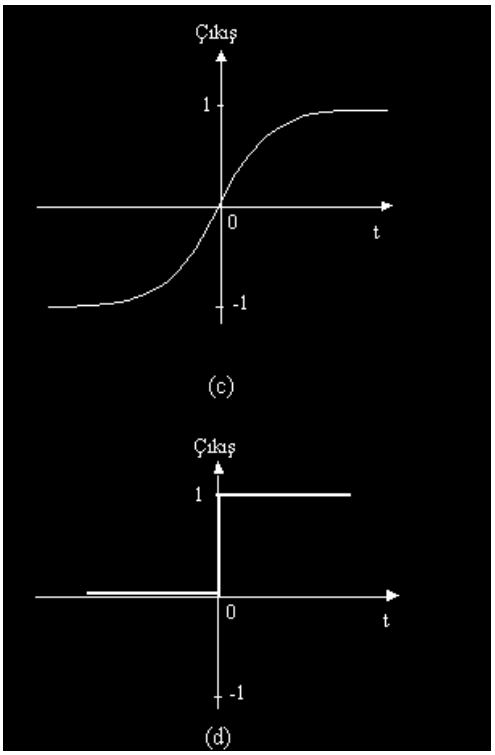
5.1.2. Aktivasyon fonksiyonları

Şekil 5.2'de, bir hücrenin girdisine göre hücre çıkışını tanımlayan aktivasyon fonksiyonlarından 5 tanesi verilmiştir.

Kullanılacak olan diğer fonksiyonların türevi alınabilir ve süreklilik arz etmesi gerekmektedir. YSA'nın kullanım amacına uygun olarak tek veya çift yönlü aktivasyon fonksiyonları da kullanılabilir.

YSA'larda en çok kullanılan aktivasyon fonksiyonu Şekil 5.2a'da verilmiştir. Doğrusal YSA'larda ve katmanlı YSA'ların çıkış katmanında kullanılan doğrusal fonksiyon Şekil 5.2b'de verilmiştir ve işlemci elemanın girişini doğrudan işlemci elemanın çıkışı olarak verir. Bipolar özelliğe sahip olan tanjant hiperbolik fonksiyonu Şekil 5.2c'de gösterilmiştir. Giriş uzayının genişletilmesinde etkili bir aktivasyon fonksiyonudur. Şekil 5.2d'de görülen keskin sınırlayıcı aktivasyon fonksiyonu, mantıksal çıkış verir ve sınıflandırıcı ağlarda tercih edilir.





Şekil 5.2. YSA'larda en çok tercih edilen aktivasyon fonksiyonları a) sigmoid, b) doğrusal c) hiperbolik tanjant ve d) keskin sınırlayıcı

5.1.3. İşlemci eleman (Yapay nöron)

Bir YSA modelinin temel birimi, Şekil 5.3'te gösterilen işlem elemanıdır. Burada girişler dış kaynaklardan veya diğer işlem elemanlarından gelen işaretlerdir. Bu işaretler, kaynağına göre kuvvetli veya zayıf olabileceğinden ağırlıkları da farklıdır.

YSA'da girilen giriş değerlerine önce toplama fonksiyonları uygulanır ve her bir işlem elemanının çıkış (İEÇ) değeri,

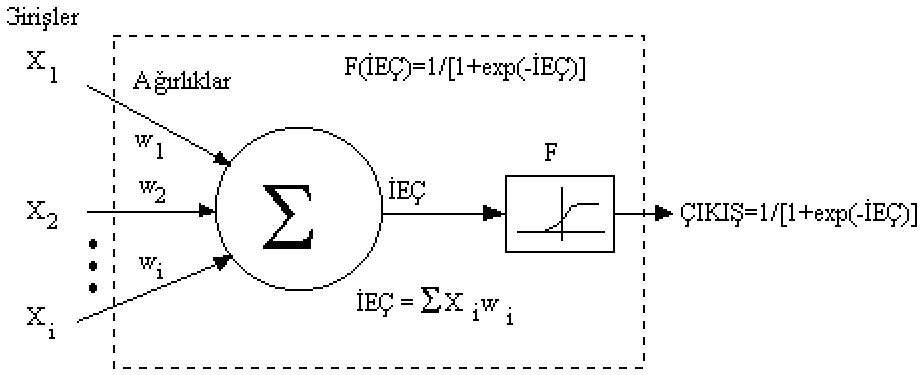
$$İEÇ = \sum_{i=1}^{N} X_i W_{ij} - \theta_i$$

olarak bulunur. Burada X_i i'inci girişi, W_{ij} j'inci elemandan i'inci elemana bağlantı ağırlığını ve θ_i eşik (threshold) değerini göstermektedir. Daha sonra bu çıkış değerleri sigmoidal aktivasyon fonksiyonuna yani öğrenme eğrisine uygulanır. Çıkış değeri,

$$ÇIKIŞ = \frac{1}{1+e^{-İEÇ}}$$

şeklinde bulunur.

Transfer fonksiyonları olarak çoğunlukla, Şekil 5.2'de verilen fonksiyonlardan birisi kullanılabilir. Uygulamalarda, en çok hiperbolik tanjant veya sigmoid fonksiyonu kullanılmaktadır. Şekil 5.3'te işlemci eleman çıkışında kullanılan sigmoid fonksiyona göre çıkış değerinin hesaplanması gösterilmiştir. Bu işlemci elemanın çıkış değeri diğer işlemci elemanlarına giriş veya ağın çıkış değeri olabilir.



Şekil 5.3. Bir işlemci elemanı (yapay nöron)

5.2. YSA ile Hesaplamanın Özellikleri

YSA'nın hesaplama özelliklerini, paralel dağılmış yapısından, öğrenebilme ve genelleme yapma yeteneğinden aldığı söylenebilir. Genelleme, eğitim ya da öğrenme süresince kullanılmayan girişler için de YSA'nın uygun tepkileri üretmesi olarak tanımlanır. Bu özellikleri ile YSA'lar karmaşık ve çözümlenmesi güç problemleri de çözebilme

yeteneğine sahiptir. Nesne tanıma , işaret işleme, sistem tanımlama ve denetimi gibi birçok mühendislik alanında YSA'lar, aşağıda belirtilen özellikleri nedeniyle başarılı olmuşlardır (Yazıcı ve ark. 2006).

5.2.1. Doğrusal olmama

YSA'nın temel işlem elemanı olan hücre doğrusal değildir. Dolayısıyla hücrelerin birleşmesinden meydana gelen YSA da doğrusal değildir ve bu özellik tüm ağa yayılmış durumdadır. Bu özelliği ile YSA, doğrusal olmayan karmaşık problemlere çözüm getirmektedir.

5.2.2. Öğrenme

Aslında öğrenmeden kasıt, ilgili problemdeki girdi-çıktı ilişkisini en güzel tanımlayacak optimum ağırlıkların bulunmasıdır. Problemden alınan ömeklerden faydalanılarak ilgili problemi kendisine uygulanan örneklerden öğrenmeye çalışır. Probleme farklı bir çözüm sağlar.

5.2.3. Genelleme

YSA, ilgilendiği problemi öğrendikten sonra eğitim sırasında karşılaşmadığı test örnekleri için de belirtilen tepkiyi üretme kabiliyetine sahiptir. Örneğin, karakter tanıma amacıyla eğitilmiş bir YSA, bozuk karakter girişlerinde de doğru karakteri verirler. Nöral hesaplamada hafızalar birleşiktir. Yani eğitilmiş ağa girişin sadece bir kısmı verilse bile, ağ hafızadan bu girişe en yakınını seçerek tam bir giriş verisi alıyormuş gibi kabul eder ve buna uygun bir çıkış değeri üretir. Veri YSA'ya, eksik , bozuk veya daha önce hiç karşılaşmadığı şekilde verilse bile, ağ kabul edilebilir en uygun çıkışı üretecektir. Bu özellik ağın genelleştirme özelliğidir.

5.2.4. Uyarlanabilirlik

YSA ağırlıkları, uygulanan probleme göre değiştirilir. Yani, belirli bir problemi çözmek amacıyla eğitilen YSA, problemdeki değişimlere göre tekrar eğitilebilir. Değişimler devamlı ise gerçek zamanda da eğitime devam edilebilir. Bu özelliği ile YSA, uyarlamalı örnek tanıma, işaret işleme, sistem tanımlama ve denetim gibi alanlarda etkin olarak kullanılır.

5.2.5. Dağıtılmış birleşik hafıza

YSA'ların en önemli özelliklerinden biri de bilgiyi depolamalarıdır. Nöral hesaplamalarda bilgi ağırlıklar üzerine dağıtılmıştır. Bağlantıların ağırlıkları nöral ağın hafıza birimidir. Bu ağırlıklar ağın o andaki sahip olduğu bilgiyi veya uygulanan örneklerden öğrenmiş olduğu davranışı verir. Bu bilgiler, ağdaki birçok ağırlıklar üzerine (hafıza birimine) dağıtılır.

Eğitilmiş ağa, eğitimde kullanılmamış herhangi farklı bir giriş uygulanırsa ağ daha önceki girişlerden öğrenmiş olduğu davranış doğrultusunda beklenen çıkışa uygun bir çıkış değeri üretebilecektir. YSA'ya uygulanan veri eksik, gürültülü veya daha önce hiç karşılaşmamış olsa bile, ağ kabul edilebilir en uygun çıkışı üretecektir. Bu özelliğe, genelleştirme özelliği denir.

5.2.6. Hata toleransı

YSA'lar, çok sayıda işlemci elemanların bağlantısı paralel dağılmış bir yapıya sahiptir ve ağın sahip olduğu bilgi, ağdaki tüm bağlantılara dağılmıştır. Giriş data setinde bulunabilecek herhangi bir gürültü, bütün ağırlıklar üzerine dağıtıldığından dolayı, gürültü etkisi tolere edilebilir. Geleneksel yöntemlere göre hatayı tolere etme yetenekleri daha fazladır.

5.2.7. Paralel işlem yapma

YSA'lar, paralel yapısı nedeniyle büyük ölçekli entegre devre teknolojisi ile gerçeklenebilir. Bu özellik, YSA'nın hızlı bilgi işleme yeteneğini ve örnek tanıma, işaret işleme, sistem kimliklendirme ve denetim gibi gerçek zaman uygulamalarında kullanımı artırır.

5.3. Yapay Sinir Ağlarının Sınıflandırılması

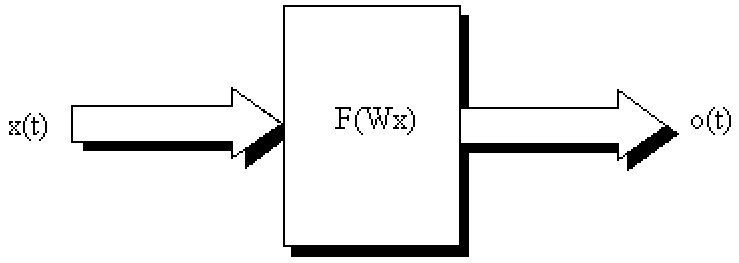
YSA'lar, genel olarak birbirleri ile bağlantılı işlemci birimlerden veya diğer bir ifade ile işlemci elemanlardan (neurons) oluşurlar. Her bir sinir hücresi arasındaki bağlantıların yapısı ağın yapısını belirler. İstenilen hedefe ulaşmak için bağlantıların nasıl değiştirileceği öğrenme algoritması tarafından belirlenir. Kullanılan bir öğrenme kuralına göre, hatayı sıfıra indirecek şekilde, ağın ağırlıkları değiştirilir. YSA'lar yapılarına ve öğrenme algoritmalarına göre sınıflandırılırlar.

5.3.1. YSA'ların yapılarına göre sınıflandırılması

Yapay sinir ağları, yapılarına göre, ileri beslemeli (feedforward) ve geri beslemeli (feedback) ağlar olmak üzere iki şekilde sınıflandırılırlar.

5.3.1.1. İleri beslemeli ağlar

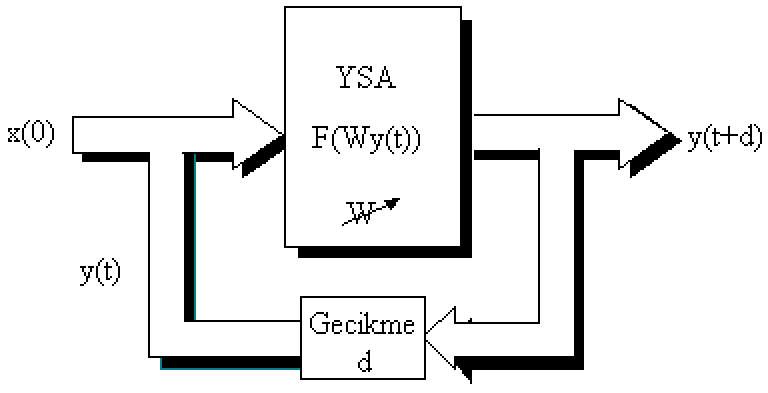
İleri beslemeli bir ağda işlemci elemanlar (İE) genellikle katmanlara ayrılmışlardır. İşaretler, giriş katmanından çıkış katmanına doğru tek yönlü bağlantılarla iletilir. İE'ler bir katmandan diğer bir katmana bağlantı kurarlarken, aynı katman içerisinde bağlantıları bulunmaz. Şekil 5.4'te ileri beslemeli ağ için blok diyagram gösterilmiştir. İleri beslemeli ağlara örnek olarak çok katmanlı perseptron (Multi Layer Perceptron-MLP) ve Learning Vector Quantization ağları verilebilir.



Şekil 5.4. İleri beslemeli ağ için blok diyagram

5.3.1.2. Geri beslemeli ağlar

Bir geri beslemeli sinir ağı, çıkış ve ara katlardaki çıkışların, giriş birimlerine veya önceki ara katmanlara geri beslendiği bir ağ yapısıdır. Böylece, girişler hem ileri yönde hem de geri yönde aktarılmış olur. Şekil 5.5'de bir geri beslemeli ağ görülmektedir. Bu çeşit sinir ağlarının dinamik hafızaları vardır ve bir andaki çıkış hem o andaki hem de önceki girişleri yansıtır. Bundan dolayı, özellikle önceden tahmin uygulamaları için uygundurlar. Bu ağlar çeşitli tipteki zaman-serilerinin tahmininde oldukça başarı sağlamışlardır. Bu ağlara örnek olarak Hopfield, Self Organizing Map, Elman ve Jordan ağları verilebilir.



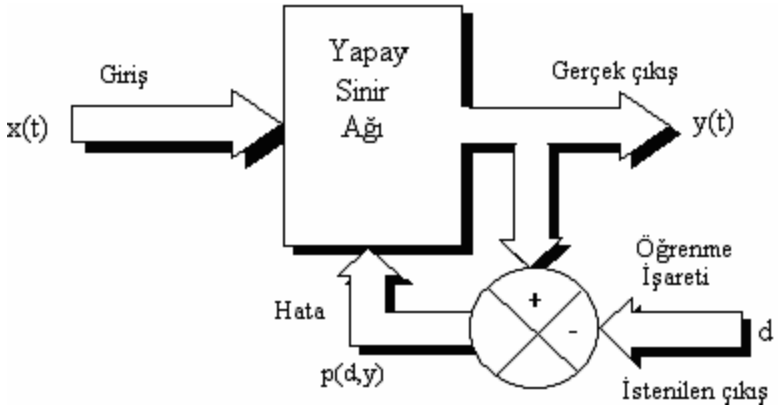
Şekil 5.5. Geri beslemeli ağ için blok diyagram

5.3.2. YSA'ların öğrenme algoritmalarına göre sınıflandırılması

Öğrenme; gözlem, eğitim ve hareketin doğal yapıda meydana getirdiği davranış değişikliği olarak tanımlanmaktadır. O halde, birtakım metot ve kurallar, gözlem ve eğitime göre ağdaki ağırlıkların değiştirilmesi sağlanmalıdır. Bunun için genel olarak üç öğrenme metodundan ve bunların uygulandığı değişik öğrenme kurallarından söz edilebilir. Bu öğrenme kuralları aşağıda açıklanmaktadır.

5.3.2.1. Danışmanlı öğrenme (Supervised learning)

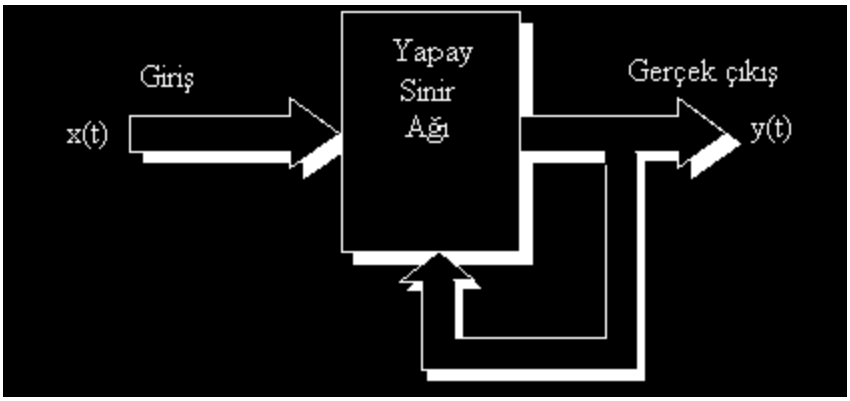
Bu tip öğrenmede, YSA'ya örnek olarak bir doğru çıkış verilir. İstenilen ve gerçek çıktı arasındaki farka (hataya) göre İE'ler arası bağlantıların ağırlığını en uygun çıkışı elde etmek için sonradan düzenlenebilir. Bu sebeple danışmanlı öğrenme algoritmasının bir "öğretmene" veya "danışmana" ihtiyacı vardır. Şekil 5.6'da danışmanlı öğrenme yapısı gösterilmiştir.



Şekil 5.6. Danışmanlı öğrenme yapısı

5.3.2.2. Danışmansız öğrenme (Unsupervised learning)

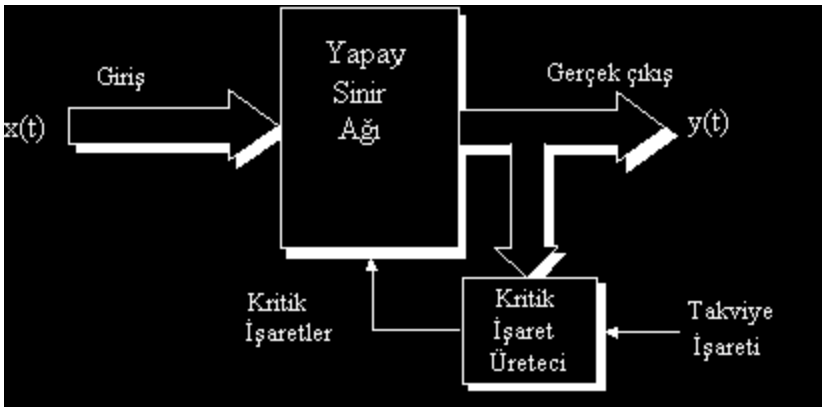
Girişe verilen örnekten elde edilen çıkış bilgisine göre ağ sınıflandırma kurallarını kendi kendine geliştirmektedir. Bu öğrenme algoritmalarında, istenilen çıkış değerinin bilinmesine gerek yoktur. Öğrenme süresince sadece giriş bilgileri verilir. Ağ daha sonra bağlantı ağırlıklarını aynı özellikleri gösteren desenler (patterns) oluşturmak üzere ayarlar. Şekil 5.7'de danışmansız öğrenme yapısı gösterilmiştir.



Şekil 5.7. Danışmansız öğrenme yapısı

5.3.2.3. Takviyeli öğrenme (Reinforcement learning)

Bu öğrenme kuralı danışmanlı öğrenmeye yakın bir metoddur. Denetimsiz öğrenme algoritması, istenilen çıkışın bilinmesine gerek duymaz. Hedef çıktıyı vermek için bir "öğretmen" yerine, burada YSA'ya bir çıkış verilmemekte fakat elde edilen çıkışın verilen girişe karşılık iyiliğini değerlendiren bir kriter kullanılmaktadır. Şekil 5.8'de takviyeli öğrenme yapısı gösterilmiştir.



Şekil 5.8. Takviyeli öğrenme yapısı

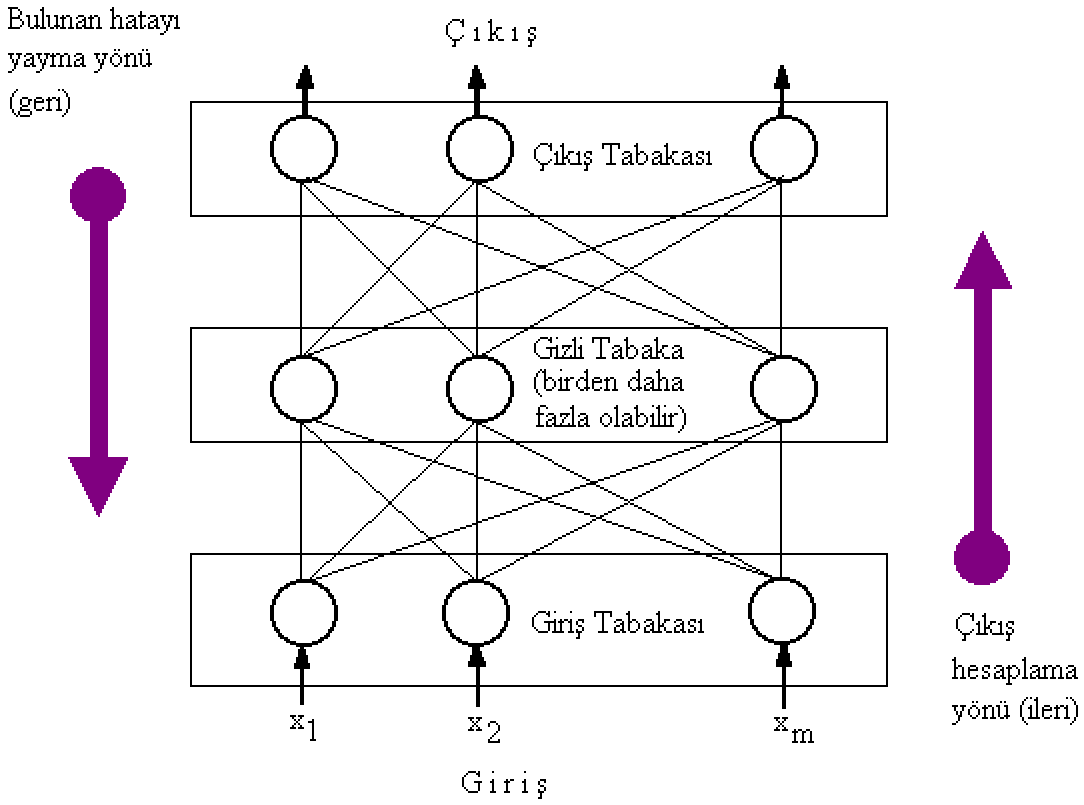
5.4. Çok Katmanlı Perseptronlar ve Öğrenme Algoritmaları

MLP'ler bir çok öğretme algoritması kullanılarak eğitilebilirler. Bu tez çalışmasında, danışmanlı öğrenme algoritmaları olan geri-yayılım (BP:back-propagation), hızlı yayılım, delta-bar-delta, genişletilmiş delta-bar-delta ve takviyeli öğrenme metodu olan genetik algoritma (GA) mikroşerit antenlerin parametrelerinin hesaplanmasında ve tasarımında kullanılan nöral modelleri eğitmede kullanılmıştır. Bu algoritmalar, bu bölümde detaylı olarak açıklanacaktır.

5.4.1. Çok katmanlı perseptronlar

Bir çok katmanlı perseptron sinir ağı modeli, Şekil 5.9'da gösterilmiştir. Bu ağ modeli özellikle mühendislik uygulamalarında en çok kullanılan sinir ağı modeli olmuştur. Bir çok öğretme algoritmasının bu ağı eğitmede kullanılabilir olması, bu modelin yaygın kullanılmasının sebebidir. Bir MLP modeli, bir giriş, bir veya daha fazla ara ve bir de çıkış katmanından oluşur. Bir katmandaki bütün işlem elemanları bir üst katmandaki bütün işlem elemanlarına bağlıdır. Bilgi akışı ileri doğru olup geri besleme yoktur. Bunun için ileri beslemeli sinir ağı modeli olarak adlandırılır. Giriş katmanında herhangi bir bilgi işleme

yapılmaz. Buradaki işlem elemanı sayısı tamamen uygulanan problemin giriş sayısına bağlıdır. Ara katman sayısı ve ara katmanlardaki işlem elemanı sayısı ise, deneme-yanılma yolu ile bulunur. Genel olarak bir veya iki tanedir. Çıkış katmanındaki eleman sayısı ise yine uygulanan probleme dayanılarak belirlenir.



Şekil 5.9. Geri yayılım MLP yapısı

MLP ağlarında, ağa bir örnek gösterilir ve örnek neticesinde nasıl bir sonuç üreteceği de bildirilir (danışmanlı öğrenme). Örnekler giriş katmanına uygulanır, ara katmanlarda işlenir ve çıkış katmanından da çıkışlar elde edilir. Kullanılan eğitme algoritmasına göre, ağın çıkışı ile arzu edilen çıkış arasındaki hata tekrar geriye doğru yayılarak hata minimuma düşünceye kadar ağın ağırlıkları değiştirilir.

İleri beslemeli ağlar, en genel anlamıyla giriş uzayıyla çıkış uzayı arasında statik haritalama yapar. Bir andaki çıkış, sadece o andaki girişin bir fonksiyonudur.

5.4.2. Geri yayılım algoritması

Bu algoritmanın akış şeması Şekil 5.10'da verilmiştir. Algoritmanın adımları aşağıdaki gibi sıralanabilir.

Adım 1: Girdi ve gizli birimleri arasındaki ağırlık W_{ij}'ler ile gizli ve çıktı birimleri arasındaki ağırlık V_j'ler küçük rasgele sayılar seçilerek başlatılır.

Adım 2: Bir p örneği seçilerek, ileriye doğru yayılma işlemi yapılır. Bunun için, önce gizli birim değerleri,

$$h_{pj} = f(Net_{pj}) = \frac{1}{1+e^{-Net_{pj}}} \tag{5.1}$$

eşitliği kullanılarak hesaplanır. Burada,

$$Net_{pj} = \sum_{i=1}^{n_I} W_{ij} X_{pi}$$

biçiminde tanımlanır. W_{ij}, i. girdi biriminden j. gizli birime olan ağırlığı, X_{pi}, p örneği için i.girdinin değerini, n_I, girdi düğümlerinin sayısını göstermektedir.

Daha sonra çıktı birimi değerleri,

$$\hat{Y}_p = f(Net_p) = \frac{1}{1+e^{-Net_p}} \tag{5.2}$$

eşitliği kullanılarak hesaplanır. Burada,

$$Net_p = \sum_{j=1}^{n_H} V_j h_{pj}$$

biçiminde tanımlanır. n_H, gizli birim sayısını, V_j, j. gizli birimden çıktı birimine olan ağırlığı göstermektedir.

Adım 3: Çıktı değerleri hesaplandıktan sonra, geriye doğru yayılma işlemi yapılır. Öncelikle çıktı hataları,

$$\delta_p=(Y_p-\hat{Y}_p)f'(\hat{Y}_p)$$

eşitliğinden hesaplanır.

Daha sonra gizli tabaka hataları,

$$\psi_{pj} = \sum_{k=1}^{l} \delta_p V_j h_{pj}(1 - h_{pj})$$

eşitliğinden hesaplanır.

Adım 4: Ağırlıkların güncellenmesinde kullanmak üzere,

$$\Delta V_j(t) = \eta \delta_p h_{pj} \tag{5.3}$$

$$\Delta W_{ij}(t) = \eta \psi_{pj} X_{pi} \tag{5.4}$$

eşitlikleri elde edilir. Burada η denemeden bağımsız öğrenme oranıdır $(0 < \eta < 1)$.

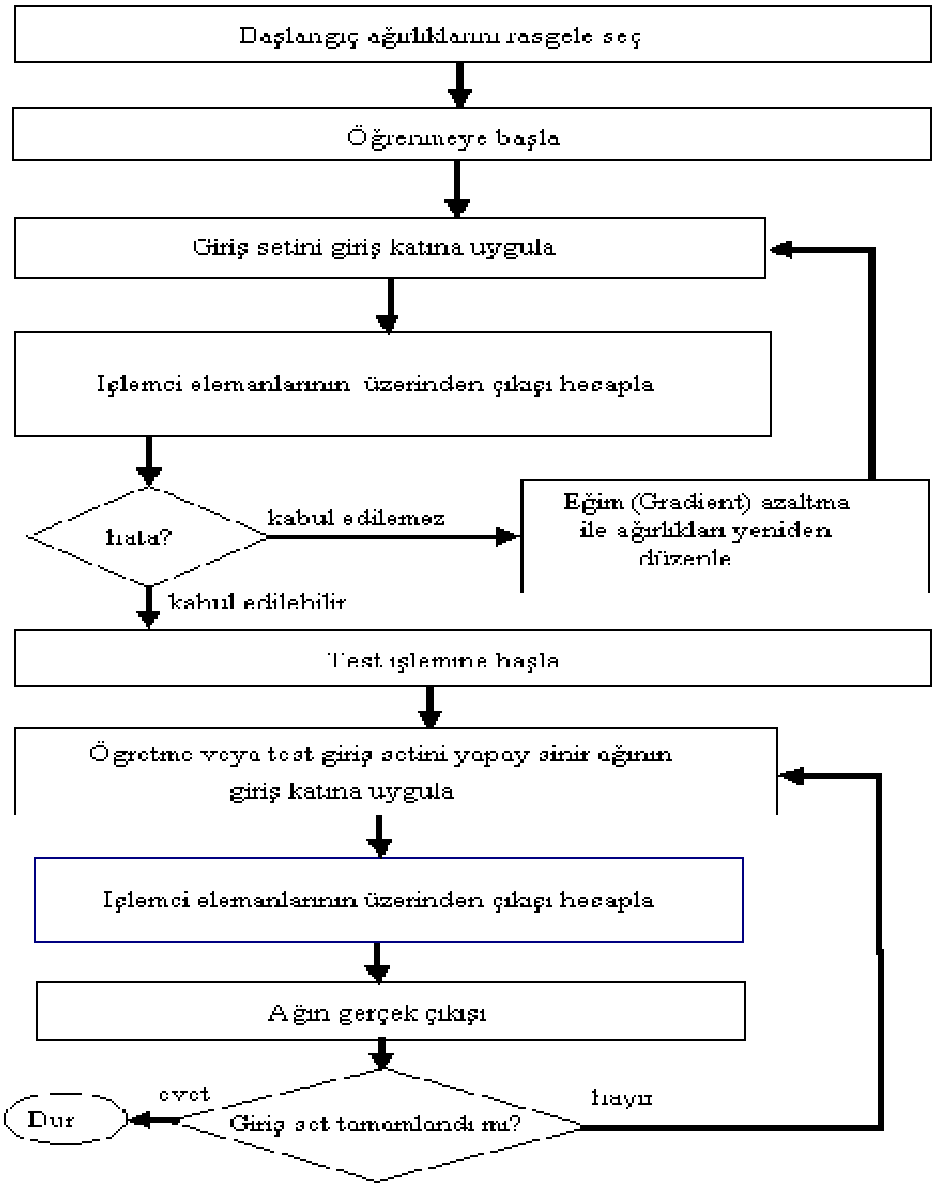
Eşitlik (5.3) ve (5.4)'ten yararlanarak V_j ve W_{ij}'nin yeni değerleri,

$$V_j(t+1) = V_j(t) + \Delta V_j(t) \tag{5.5}$$

$$W_{ij}(t+1) = W_{ij}(t) + \Delta W_{ij}(t) \tag{5.6}$$

biçiminde elde edilir.

Adım 5: Her bir örnek için yukarıdaki adımlar tekrarlanır. $E = \frac{1}{2} \sum_{p=1}^{n} \sum_{k=1}^{o} (y_{pk} - \hat{y}_{pk})^2$ ile ifade edilen hata kriterine göre yakınsama olana kadar iterasyonlar tekrarlanır (Yapıcı Pehlivan ve Apaydın 2005).



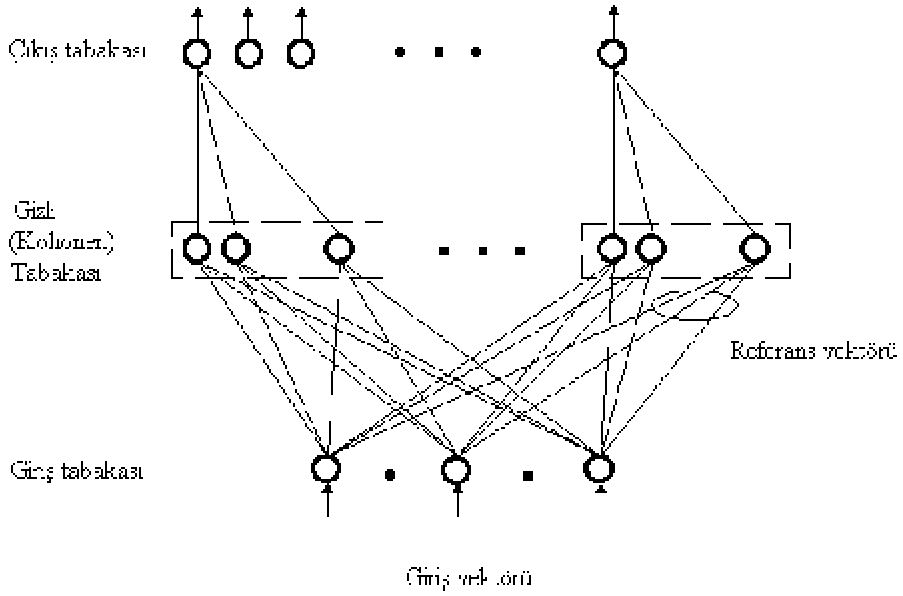
Şekil 5.10. Çok katmanlı bir perseptron geri yayılım akış şeması

5.5. Diğer Yapay Sinir Ağları

Bu kısımda, literatürde adı en çok geçen YSA yapıları kısaca açıklanacaktır.

5.5.1. LVQ (Learning vector quantization)

Şekil 5.11, üç kattan oluşan LVQ ağını göstermektedir. Bu tabakalar, giriş tabakası, gizli tabaka ve çıkış tabakasıdır. LVQ ağı, giriş ve gizli tabaka arasında tamamen, gizli ve çıkış tabakası arasında da kısmen bağlıdır. Her çıkış işlemci elemanı farklı bir gizli işlemci eleman kümesine bağlıdır. Gizli ve çıkış işlemci elemanları arasındaki ağırlıklar 1'e sabitlenmiştir. Giriş-gizli işlemci eleman bağlantılarının ağırlıkları "referans" vektörlerinin elemanlarını oluşturur (her gizli işlemci elemana bir referans vektör atanmıştır). Ağın öğretilmesi esnasında bunlar yeniden değerler alırlar. Hem gizli işlemci elemanlar (bunlar Kohonen işlemci elemanları olarak da bilinir) hem de çıkış işlemci elemanları ikili (binary) çıkışa sahiptir. Ağa bir giriş deseni verildiğinde referans vektörü giriş desenine en yakın olan gizli işlemci eleman kümesi "1", diğerleri "0" üretir. "1" üreten çıkış işlemci elemanı giriş işaretini sınıflar ve her işlemci eleman ayrı bir sınıfa atanmıştır.



Şekil 5.11. LVQ ağı

En basit LVQ öğretme prosedürü sırasıyla şöyle verilebilir.

i. Referans vektörlerinin ağırlıklarının başlangıç değerlerinin belirlenmesi,

ii. Ağa giriş deseni uygulanması,

iii. Giriş deseni ile her referans vektörü arasındaki "Öklit (Euclidian) uzaklığı"nın hesaplanması,

iv. Giriş desenine en yakın referans vektörünün (bu vektör kazanan gizli işlemci elemanının referans vektörüdür) ağırlıklarının yeniden düzenlenmesi. Eğer bu gizli işlemci eleman, o çıkış işlemci elemanın bağlı olduğu gizli işlemci eleman kümesine bağlı ise referans vektörü giriş desenine daha yakın hale getirilir. Aksi takdirde referans vektörü uzaklaştırılır, (ii) numaralı adıma yeni bir giriş deseni ile dönülür ve işlemlere bütün giriş desenleri sınıflandırılıncaya kadar devam edilir.

6. YSA YARDIMIYLA LİNEER OLMAYAN BULANIK REGRESYON TAHMİNİ

6.1. Bulanık Regresyon

Standart regresyon modellerinde, veri ile modelden elde edilen çıktı arasındaki fark gözlemlenmiş hata adını alır. Bulanık regresyon modellerinde ise gözlenen veri ile model çıktısı arasındaki fark, girdi-çıktı arasında olan sistem yapısındaki bir belirsizlik olarak varsayılır.

Bulanık regresyon modeli çıktılarının bulanık sayı olması nedeniyle klasik lineer regresyondan farklıdır. Tanaka ve ark. (1982) tarafından verilen bulanık regresyon modeli;

$$\underline{\widetilde{Y}} = \widetilde{A}_0 + \widetilde{A}_1 x_1 + \cdots + \widetilde{A}_p x_p$$

şeklinde tanımlanmıştır. Modelde $\underline{\widetilde{Y}}$ ile bulanık çıktılar ve $\underline{\widetilde{A}}$ ile bulanık katsayılar ifade edilmektedir. Girdi değerleri x_j 'ler ise bulanık olmayan kesin sayılardır.

6.2. Lineer Olmayan Bulanık Regresyon

Lineer olmayan bulanık regresyon (Fuzzy Nonlinear Regression [FNR]) modeli ile, aralarında lineer olmayan bir ilişkinin olduğu girdisi kesin çıktısı bulanık veri seti için öngörü yapılabilir. Ayrıca FNR modeli, kesin bir olasılıkla çıktı sonuçlarının içinde yer alacağı bir aralık da öngörebilir. Modelde bulanıklık, modelin çıktısının bulanık sayı olmasını, lineer olmama durumu, sinir ağlarının kullanılmış olmasını ve regresyon da regresyon analizine benzer olarak öngörü yapılmasını göstermektedir.

Geri beslemeli sinir ağlarına dayalı aralık regresyon analizi için bir model hazırlanmıştır. Bir ağ, veri aralıklarının en üst yanı olarak tanımlanır ve diğeri de veri aralıklarının en alt yanı ile tanımlanır. Bu yöntemin amacı lineer olmayan bir regresyon modeli meydana getirmektir.

Ancak bu yaklaşım iki sinir ağı (en üst yan ve en alt yan) için adımsal bir maliyet fonksiyonu seçimine odaklanmıştı. Bu algoritma, verinin yalnızca kutuplarına odaklandığından uç noktalara aşırı duyarlıdır. Xu ve ark. (2000), uç noktalara daha az duyarlı olan bağımlı değişkenin ortalamasına dayalı yeni bir algoritma geliştirdiler. Bu algoritmada farklı maliyet fonksiyonları araştırmak yerine çalışmadan önce makul bir aralık hesaplamak için bağımsız değişkenlerin ortalamasına güvenilmiştir. Aralığın alt ve üst sınırı iki bağımsız sinir ağının amaçları olacaktır. Bu iki sinir ağı her çalışma algoritmasına uygulanabilir ve çalışma düzgün ve çok kesin olabilir.

$\{x_i \mid i = 1,2,...,n\}$ ve $\{y_i \mid i = 1,2,...,n\}$ sırasıyla sinir ağları çalışma verisinde bağımsız ve bağımlı değişkenler olsun. Çalışma verisi n tane girdi-çıktı çifti (x_i, y_i) olarak gösterilebilir ve burada $i = 1,2,...,n$ 'dir. Bağımlı değişken y_i 'nin genişliği 1 iken, bağımsız değişken vektörü $x_i = (x_{i1}, x_{i2},..., x_{im})$ 'nin genişliği m'dir. $\underline{y} + (x_i)$ ve $\underline{y} - (x_i)$ bağımsız değişken vektörü x_i 'ye karşı gelen iki sinir ağından elde edilen çıktılar olsun. Burada her bir ağ m girdi düğümü ve tek çıktı düğümüne sahiptir. Amacımız $\underline{y} + (x_i)$ ve $\underline{y} - (x_i)$ çıktılarının yaklaşık olarak aşağıdaki durumu sağlamasıdır.

$$\underline{y} - (x_i) \le y_i \le \underline{y} + (x_i) \tag{6.1}$$

Burada anahtar nokta, çalışma sürecinden önce aralığın makul bir en üst ve en alt sınırının nasıl belirleneceğidir. Diğer bir deyişle y_i 'nin $y_i \in (y_i - y_i^W, y_i + y_i^W)$ olacak biçimde y_i^W yayılması ortaya konulmak zorundadır, böylece aynı bağımsız değişken vektörü ile üst taraf sinir ağı $y + (x_i)$ 'ye yaklaşırken, alt taraf sinir ağı $y - (x_i)$ 'ye yaklaşmaya çalışacaktır. y_i^W 'nin genişliği şunları gerektirir.

- pozitif olması
- sinir ağının doğruluğuna bağlı olarak olabildiğince küçük olması

- y_i 'nin dağılımını yansıtması

y_i^W 'nin hesaplanmasında y_i 'nin ortalama ve standart sapmasının gözlenmesi gerekir.

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i \tag{6.2}$$

$$s_y = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{6.3}$$

Burada ortalamanın yayılmayı belirlemede özel bir durum olarak kabul edilmesinin nedeni, bağımlı değişken değerleri y_i 'lerin, sinir ağının salınımının meydana geldiği $\bar{y}$ etrafında kümelendiğinin varsayılmasıdır. Ancak bu her zaman doğru değildir. Bazı veri setleri için, bağımlı değişken değeri kesin bir nokta etrafında toplanmaz. Bu durumda, yaklaşım güvenli olmaz. Bu nokta açıklamaya ihtiyaç duyan açık bir noktadır. Böylece, sinir ağının yayılmasını azaltabilmek için $\bar{y}$ etrafında nispeten büyük yayılma gösteren y_i 'ler verilmelidir. Buna göre y_i^W şöyle tanımlanabilir:

$$y_i^W = \frac{\phi}{\bar{y}} y_i + f(y_i) \tag{6.4}$$

Burada ϕ bir sabittir ve y_i 'lerin ortalama noktasındaki yayılması olarak yorumlanır. Deneysel çalışmalar $0{,}6\,\bar{y} \le \phi \le 0{,}7\,\bar{y}$ olarak elde edilen ϕ 'nin iyi sonuç verdiğini göstermiştir. Küçük ϕ 'ler (6.1)'i garanti etmez. Diğer taraftan aşırı büyük ϕ değerleri, aralığın duyarsız olmasına yol açabilen geniş bir yayılma sonucu verir. Örneğin bir öğrencinin mezuniyet not ortalamasının 0'la 4.0 arasında olduğunu söylemek çok geniş bir aralık tanımlar. Telafi fonksiyonu $f(y_i)$ sinir ağının çıktısının $\bar{y}$ etrafındaki yayılımını azaltır. Şöyle tanımlanır:

$$f(y_i)=\begin{cases}\dfrac{\psi\ y_i}{1+e^{-\theta_1(y_i-\omega_1)}} & y_i \le \omega_1 \\ \psi\ y_i & \omega_1 \le y_i \le \omega_2 \\ \dfrac{\psi\ y_i}{1+e^{-\theta_2(y_i-\omega_2)}} & y_i \ge \omega_2\end{cases} \tag{6.5}$$

Burada ψ bir ölçü sayısıdır ve bu çalışmada ψ =1 olarak alınacaktır. ψ parametresi, yayılma alanının ne kadar daraltılmak istendiğini gösterir. ω_1 ve ω_2 parametreleri deneysel olarak meydana gelen yayılım alanının alt ve üst sınırlarıdır. θ_1 ve θ_2 parametreleri fonksiyonunun iki tarafının ne yükseklikte aşağı ve yukarı gideceğini gösterir. θ_1=θ_2 ise fonksiyon simetriktir. θ_1 ve θ_2 şöyle tanımlanır.

$$\theta_1=\frac{\bar{y}-\omega_1}{s_y} \tag{6.6}$$

$$\theta_2=\frac{\omega_2-\bar{y}}{s_y} \tag{6.7}$$

İki sinir ağı bağımsız olarak çalışıldığından $y-(x_i) > y+(x_i)$ durumu ortaya çıkmaktadır. O zaman aralık modeli i. modül için anlamsız olacaktır. Bu normal olmama durumunu elemek için $z-(x_i)$ ve $z+(x_i)$ fonksiyonlarını kullanırız.

$$\underline{z}-(x_i)=\begin{cases}\underline{y}-(x_i) & \underline{y}-(x_i) \le \underline{y}+(x_i) \\ \dfrac{1}{2}[\underline{y}-(x_i)+\underline{y}+(x_i)] & \underline{y}-(x_i) > \underline{y}+(x_i) \\ 0 & \underline{y}-(x_i) < 0\end{cases} \tag{6.8}$$

$$\underline{z}+(x_i)=\begin{cases}\underline{y}+(x_i) & \underline{y}-(x_i) \le \underline{y}+(x_i) \\ \dfrac{1}{2}[\underline{y}-(x_i)+\underline{y}+(x_i)] & \underline{y}-(x_i) > \underline{y}+(x_i)\end{cases} \tag{6.9}$$

$\underline{z}-(x_i) \le \underline{z}+(x_i)$ tüm x_i 'leri kapsar ve $[\underline{z}-(x_i), \underline{z}+(x_i)]$ lineer olmayan aralığı daima anlamlıdır. Bu aralıktan bulanık lineer olmayan bir model oluşturabiliriz. Her bir modüle ait hataların sayısının genişliğini öngöreceksek ki bu sayı negatif olamaz, FNR'yi oluşturmadan önce negatif $\underline{z}-(x_i)$ sıfıra kadar kesilmelidir. $\hat{Y}(x_i)$, y_i bağımlı değişkeninden öngörülen bir bulanık sayı olsun. Eşitlik (6.11)'den simetrik üçgensel üyelik fonksiyonunu kullanılarak FNR yazılabilir.

$$\hat{Y}(x_i) = \left(f^C(x_i), f^W(x_i)\right) \tag{6.10}$$

Eşitlik (6.13)'ten $\hat{Y}(x_i)$ 'nin α -seviye kümesi aşağıdaki aralıkla hesaplanır.

$$[\hat{Y}(x_i)]_\alpha = \left[f^C(x_i) - (1-\alpha) f^W(x_i), f^C(x_i) + (1-\alpha) f^W(x_i)\right] \tag{6.11}$$

$f^C(x_i)$ ve $f^W(x_i)$ sırasıyla $\hat{Y}(x_i)$ 'nin merkezi ve yayılmasıdır. $\hat{Y}(x_i)$ 'nin bulanık modelini oluşturabilmek için $\hat{Y}(x_i)$ 'nin α -seviye kümesinin verilen bir $\alpha \in (0,1]$ için aşağıdaki durumu sağlaması gerekir.

$$[\hat{Y}(x_i)]_\alpha = \left[\underline{z}-(x_i), \underline{z}+(x_i)\right], \text{ tüm } x_i \text{'ler için} \tag{6.12}$$

$[\hat{Y}(x_i)]_\alpha^+$, $[\hat{Y}(x_i)]_\alpha$ 'in üst sınırı, $[\hat{Y}(x_i)]_\alpha^-$ alt sınırı ve $[\hat{Y}(x_i)]_\alpha^C = f^C(x_i)$ de merkezi olsun. (6.13) ve (6.14)'ten verilen bir $\alpha \in (0,1]$ için $f^C(x_i)$ ve $f^W(x_i)$ yazılabilir.

$$f^C(x_i) = \frac{1}{2}\left(\underline{z}-(x_i) + \underline{z}+(x_i)\right) \tag{6.13}$$

$$f^W(x_i) = \frac{1}{2h}\left(\underline{z}+(x_i) - \underline{z}-(x_i)\right) \tag{6.14}$$

Böylece, eşitlik (6.10)'dan x_i'ye bağlı *i.* modül için y_i'yi öngören $\hat{Y}(x_i)$ bulanık sayısı elde edilir.

Özet olarak bir lineer olmayan bulanık regresyon modelinin adımları şöyle sıralanabilir:

Adım 1 Bağımlı değişken y_i'nin $\bar{y}$, ortalama ve s_y, standart sapma değerleri hesaplanır.

Adım 2 $\bar{y}$'nin %60-%70'i şeklinde bir ϕ seçilir.

Adım 3 Telafi fonksiyonu $f(y_i)$ olmadan (6.4)'ü kullanarak bir y_i^W yayılması belirlenir.

Adım 4 Aralığın alt ve üst sınırı olarak $y_i - y_i^w$ ve $y_i + y_i^w$ hesaplanır.

Adım 5 (5.1)-(5.6)'yı kullanarak alt ve üst sınıra uygun iki sinir ağı oluşturulur.

Adım 6 y_i'nin yayılma alanını daraltmak için;

- $y_i \in [\omega_1, \omega_2]$ olmak üzere yayılma alanı tanımlanır.
- (6.5)-(6.7)'yi kullanarak $f(y_i)$ telafi fonksiyonu hesaplanır ($\psi = 1$).
- y_i^W'yi hesaplamak ve yeni alt ve üst sınır belirlenir.
- Yeni alt ve üst sınıra karşı gelen sinir ağlarını oluşturulur.

Adım 7 (6.8) ve (6.9)'u kullanarak $\underline{z}-(x_i)$ ve $\underline{z}+(x_i)$ hesaplanır.

Adım 8 Değişik α değerleri seçerek ve eşitlik (6.13) ile (6.14)'den $f^C(x_i)$ ve $f^W(x_i)$ hesaplanır.

Adım 9 (6.11)'den $[\hat{Y}(x_i)]_\alpha$ hesaplanır (Xu ve Khosgoftaar 2000).

6.3. Modelin Test Edilmesi

Modelin doğruluğunu test etmek için ortalama mutlak hata (AAE) değeri kullanılmıştır. Modelimizde AAE^+, AAE^- ve AAE^C olmak üzere üç tip AAE vardır. Bunlar sırasıyla $[\hat{Y}(x_i)]_\alpha$'ın üst sınırı, alt sınırı ve merkezi ile ilgilidir.

$$AAE^{*} = \sum_{i=1}^{n} \left| y_i - [\hat{Y}(x_i)]_{\alpha}^{*} \right| \tag{6.15}$$

Burada *, +, - ve C yerine geçmektedir.

AAE^{+} , $[\hat{Y}(x_i)]_{\alpha}$ 'ın üst sınırının, AAE^{-} , $[\hat{Y}(x_i)]_{\alpha}$ 'ın alt sınırının ve AAE^{C} , $[\hat{Y}(x_i)]_{\alpha}$ 'ın merkezinin y_i 'ye ne kadar yaklaştığını açıklar.

AA^{W} , $[\hat{Y}(x_i)]_{\alpha}$ 'ın ortalama mutlak genişliği olsun.

$$AA^{W} = \frac{2}{n} \sum_{i=1}^{n} \left| f^{W}(x_i) \right| \tag{6.16}$$

AA^{W} , verilen bir α -seviye için ne genişlikte aralık olduğunu açıklar (Xu ve Khosgoftaar 2000).

7. LS-SVM'YE DAYALI BULANIK LİNEER OLMAYAN REGRESYON

7.1. En Küçük Kareler Destek Vektör Mekanizması

7.1.1. Support vector machine (SVM) sınıflayıcısı

$x_i \in \Re^p$ $i.$ girdi verisi ve $y_i \in \{-1,+1\}$ şeklinde veri etiketi olmak üzere N gözlemli $\{x_i, y_i\}, i = 1,..., N$ seti verilsin. SVM yaklaşımının amacı aşağıdaki biçimde verilen bir sınıflayıcı bulmaktır (Suykens ve ark. 2002).

$$y(x) = \operatorname{sgn}\left[\sum_{i=1}^{N} \alpha_i y_i K(x_i, x) + b\right] \tag{7.1}$$

Burada a pozitif reel bir sabit sayı, b ise bir reel sayıdır. $K(x_i, x) = \langle \phi(x_i), \phi(x) \rangle$, $\langle \cdot, \cdot \rangle$ iç çarpım ve $\phi(x$ de orijinal uzaydan çok boyutlu uzaya eşleştirme yapan lineer olmayan bir fonksiyondur. sgn fonksiyonu ileride tanımlanacaktır. Çok boyutlu uzayda verinin bir yüzey tarafından bölündüğü varsayılsın. Bu durumda

$$\begin{cases} w^T \phi(x_i) + b \geq 1, & y_i = +1 \\ w^T \phi(x_i) + b \leq -1, & y_i = -1 \end{cases} \tag{7.2}$$

olur. Bu iki eşitlik

$$y_i \left[w^T \phi(x_i) + b \right] \geq 1, \quad i = 1,..., N, \tag{7.3}$$

olarak bir eşitlikte yazılabilir. Bölen yüzeyin var olmadığı durumda, ξ_i gibi bir gevşek değişken eklenir.

$$\begin{cases} y_i \left[w^T \phi(x_i) + b \right] \geq 1 - \xi_i, & i = 1,..., N, \\ \xi_i \geq 0, & i = 1,..., N, \end{cases} \tag{7.4}$$

Yapısal risk minimizasyon prensibine göre, risk sınırı aşağıdaki minimizasyon problemi ile en küçük yapılır.

$$\min_{w,\xi} J(w,\xi) = \frac{1}{2} w^T w + C \sum_{i=1}^{N} \xi_i \tag{7.5}$$

Burada (7.4) kısıtı göstermektedir ve Lagrange fonksiyonu

$$L_1(w,\xi,b,\alpha,\beta) = J(w) - \sum_{i=1}^{N} \alpha_i \left\{ y_i \left[w^T x(\phi_i) + b \right] - 1 + \xi_i \right\} - \sum_{i=1}^{N} \beta_i \xi_i \tag{7.6}$$

biçiminde oluşturulur. Burada $\alpha_i \geq 0, \beta_i \geq 0 \quad (i = 1, \ldots, N)$, Lagrange çarpanlarıdır. Lagrange fonksiyonunda optimal nokta

$$\max_{\alpha,\beta} \min_{w,\xi} L_1(w,\xi,\alpha,\beta) \tag{7.7}$$

dır.

$$\begin{cases} \dfrac{\partial L_1}{\partial w} = 0 \rightarrow w = \sum_{i=1}^{N} \alpha_i y_i \phi(x_i) \\ \dfrac{\partial L_1}{\partial b} = 0 \rightarrow \sum_{i=1}^{N} \alpha_i y_i = 0 \\ \dfrac{\partial L_1}{\partial \xi_i} = 0 \rightarrow 0 \leq \alpha_i \leq C \quad i = 1, \ldots, N \end{cases} \tag{7.8}$$

türevleri oluşturulur. (7.8) eşitliği, (7.6) eşitliğinde yerine konulursa (7.9)'daki karesel programlama problemi elde edilir.

$$\max_{\alpha} Q_1(\alpha) = \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{N} y_i y_j K(x_i, x_j) \alpha_i \alpha_j \tag{7.9}$$

Burada $K(x_i, x_j) = \langle \phi(x_i), \phi(x_j) \rangle$ kernel fonksiyonu olarak adlandırılır. (7.8)'deki kısıtları göz önüne alarak optimizasyon problemi çözüldüğünde çok boyutlu uzayda bir yüzey elde edilir ve asıl uzayda (7.1)'deki sınıflayıcı oluşturulmuş olur (Suykens ve ark. 2002).

7.1.2. En küçük kareler (Least Square) SVM

(7.5)'teki minimizasyon problemi yeniden formüle edilerek SVM sınıflayıcısına bir en küçük kareler yaklaşımı getirilebilir.

$$\min_{w,b,e} J_2(w,b,e) = \frac{\mu}{2} e^T e + \frac{\zeta}{2} w^T w \quad \text{s.t.} \quad \sum_{i=1}^{N} e_c^2 \tag{7.10}$$

Eşitlik kısıtları,

$$y_i [w^T \phi(x_i) + b] = 1 - e_{c,i}, \quad i = 1, \ldots, N \tag{7.11}$$

şeklindedir. LS-SVM sınıflayıcısı formülü $y_i = \pm 1$ şeklinde ikili hedefi olan bir regresyon olarak düşünülebilir. $y_i^2 = 1$ kullanılarak,

$$\sum_{i=1}^{N} e_{c,i}^2 = \sum_{i=1}^{N} (y_i e_{c,i})^2 = \sum_{i=1}^{N} e_{c,i}^2 = \sum_{i=1}^{N} \left(1 - y_i (w^T \phi(x_i) + b)\right)^2 \tag{7.12}$$

elde edilir ve

$$e_i = y_i - (w^T \phi(x_i) + b) \tag{7.13}$$

dir. Böylece LS-SVM sınıflayıcısı,

$$J_2(w, b, e) = \mu E_D + \zeta E_W \tag{7.14}$$

ifadesine eşit olur. Burada,

$$E_W = \frac{1}{2} w^T w \tag{7.15}$$

ve

$$E_D = \frac{1}{2}\sum_{i=1}^{N} e_i^2 = \frac{1}{2}\sum_{i=1}^{N} \left(y_i - (w^T \phi(x_i) + b)\right)^2 \tag{7.16}$$

dır. μ ve ζ , hata kareler toplamına karşı düzeltme miktarını ayarlayan parametreler olarak düşünülmelidir. (7.14) çözümü $\gamma = \mu / \zeta$ oranına bağlı değildir, bu nedenle, yalnızca γ parametresi düzeltme parametresi olarak kullanılabilir.

LS-SVM çözümü Lagrange fonksiyonu oluşturulduktan sonra elde edilir.

$$\begin{aligned} L_2(w, b, e, \alpha) &= J(w, e) - \sum_{i=1}^{N} \alpha_i \left\{ w^T \phi(x_i) + b + e_i - y_i \right\} \\ &= \frac{1}{2} w^T w + \frac{\gamma}{2}\sum_{i=1}^{N} e_i^2 - \sum_{i=1}^{N} \alpha_i \left\{ w^T \phi(x_i) + b + e_i - y_i \right\} \end{aligned} \tag{7.17}$$

Burada $\alpha_i \in R$ Lagrange çarpanlarıdır. Optimallik için koşullar şöyle verilir.

$$\begin{cases} \dfrac{\partial L_2}{\partial w} = 0 \quad \rightarrow \quad w = \sum_{i=1}^{N} \alpha_i \phi(x_i) \\ \dfrac{\partial L_2}{\partial b} = 0 \quad \rightarrow \quad 0 = \sum_{i=1}^{N} \alpha_i \\ \dfrac{\partial L_2}{\partial e_i} = 0 \quad \rightarrow \quad \alpha_i = \gamma e_i, \quad i = 1, \ldots, N \\ \dfrac{\partial L_2}{\partial \alpha_i} = 0 \quad \rightarrow \quad y_i = w^T \phi(x_i) + b + e_i, \quad i = 1, \ldots, N \end{cases} \tag{7.18}$$

Burada e ve w çıkarılarak karesel programlama yerine $Y = [y_1, \ldots, y_N]^T$, $1_N = [1, \ldots, 1]^T$, $\alpha = [\alpha_1, \ldots, \alpha_N]^T$ ve I_N'nin $N \times N$'lik birim matris olduğu (7.19)'daki doğrusal sistem elde edilebilir. $\Omega \in \Re^{N \times N}$, $\Omega_{ij} = \varphi(x_i)^T \varphi(x_j) = K(x_i, x_j)$ olan kernel matrisidir. Bu LS-SVM sınıflayıcısı (7.1) ile aynı biçime sahiptir (Suykens ve ark. 2002).

$$\begin{bmatrix} 0 & 1_N^T \\ 1_N & \Omega + \gamma^{-1} I \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ Y \end{bmatrix} \tag{7.19}$$

7.2. Lagrange Yöntemi:

$$\begin{aligned} &Min\ f(x) \\ &g_i(x) = 0 \qquad i = 1, \ldots, n \end{aligned}$$

problemi verilsin. Lagrange yönteminde,

$$L(x_1, \ldots, x_n, \lambda_1, \ldots, \lambda_n) = f(x_1, \ldots, x_n) - \sum_{i=1}^{m} \lambda_i [g(x_1, \ldots, x_n)]$$

şeklinde bir fonksiyon oluşturulur. Bu fonksiyona Lagrange fonksiyonu denir ve kısaca $L(\underline{x}\lambda) = f(\underline{x}) - \lambda g(\underline{x})$ şeklinde gösterilebilir. λ_i parametreleri Lagrange çarpanları olarak adlandırılır. $f(\underline{x})$'in $g(\underline{x})$ kısıtları altında optimizasyonu, $L(\underline{x}\lambda)$ Lagrange fonksiyonunun optimizasyonuna eşdeğerdir. Problemde $f(\underline{x})$ ve $g(\underline{x})$ fonksiyonlarının sürekli ve türevlenebilir olduğu varsayılmaktadır. $L(\underline{x}\lambda)$ fonksiyonunun $(n+m)$ tane değişkene göre kısmi türevi alınıp sıfıra eşitlenirse,

$$\frac{\partial L}{\partial x_j} = \frac{\partial f}{\partial x_j} - \sum_{i=1}^{m} \lambda_i \frac{\partial g_i}{\partial x_j} = 0 \quad (j = 1, \ldots, n)$$

$$\frac{\partial L}{\partial \lambda_i} = -g_i(x_1,...,x_n) + b_i = 0 \quad (i = 1,...,m)$$

bulunur. Bu sistemin çözümü aranılan optimal noktayı verir. $(x_1^*,...,x_n^*,\lambda_1^*,...,\lambda_n^*)$, $L(\underline{x}\lambda)$ için bir optimal noktayı gösterirken $(x_1^*,...,x_n^*)$ kümesi de asıl fonksiyonun optimali olur (Oral 1989).

7.3. Bulanık Lineer LS-SVM Regresyon

Bulanık girdi ve çıktıdan oluşan bulanık çok değişkenli lineer regresyon modelleri konveks optimizasyon problemi olarak LS-SVM yaklaşımı ile çözümlenmektedir. LS-SVM nin amacı bulanık regresyon modellerinin çözümünde hesaplama kolaylığı ve yeterlilik sağlamaktır. Burada üçgensel bulanık sayılara dayalı bulanık regresyon modellerine odaklanılmıştır, çünkü bu tip bulanık sayılar yaygın olarak kullanılmaktadır. Yamuksal ve Gaussian sayılara dayalı bulanık regresyon modelleri benzer davranışlar içinde yapılandırılabilir (Hong ve Hwang 2006).

Deney verisi $\{X_i, Y_i\}_{i=1}^n \subset T(\Re)^p \times T(\Re)$ şeklinde verilsin. Burada,

$X_i = ((m_{X_{i1}}, \alpha_{X_{i1}}, \beta_{X_{i1}})...(m_{X_{id}}, \alpha_{X_{id}}, \beta_{X_{ip}}))$

$Y_i = (m_{Y_i}, \alpha_{Y_i}, \beta_{Y_i})$

$T(\Re)$: Üçgensel bulanık sayılar kümesi

$T(\Re)^p$: Üçgensel bulanık sayıların p -boyutlu vektör kümesi

$\underline{m}_{X_i} = (m_{X_{i1}},...,m_{X_{ip}})$

$\underline{\alpha}_{X_i} = (\alpha_{X_{i1}},...,\alpha_{X_{ip}})$

$\underline{\beta}_{X_i} = (\beta_{X_{i1}},...,\beta_{X_{ip}})$

$B = (m_B, \alpha_B, \beta_B)$

$\underline{w} = (w_1,...,w_p)$

Bulanık girdi ve bulanık çıktı için,

$$Y(\underline{X}) = \langle \underline{w}, \underline{X} \rangle + B$$
$$= \left(\langle \underline{w}, \underline{m}_{\underline{X}} \rangle + m_B, \langle |\underline{w}|, \underline{\alpha}_{\underline{X}} \rangle + \alpha_B, \langle |\underline{w}|, \underline{\beta}_{\underline{X}} \rangle + \beta_B \right), \quad B \in T(\Re), w \in \Re^p$$

modeli düşünülsün. Burada $|\underline{w}| = (|w_1|,...,|w_d|)$ dir. Bulanık olmayan çoklu lineer regresyon düşüncesinden hareketle yukarıdaki model için konveks bir optimizasyon problemi yazabiliriz.

$$\min \frac{1}{2}\|\underline{w}\|^2 + \frac{C}{2}\sum_{k=!}^{3}\sum_{i=1}^{n} e_{ki}^2 \tag{7.20}$$

Burada kısıtlar;

$$\begin{aligned} & m_{Y_i} - \langle \underline{w}, \underline{m}_{X_i} \rangle - m_B = e_{1i} \\ & \left(m_{Y_i} - \alpha_{Y_i}\right) - \left(\langle \underline{w}, \underline{m}_{X_i} \rangle + m_B - \langle |\underline{w}|, \underline{\alpha}_{X_i} \rangle - \alpha_B \right) = e_{2i} \\ & \left(m_{Y_i} + \beta_{Y_i}\right) - \left(\langle \underline{w}, \underline{m}_{X_i} \rangle + m_B - \langle |\underline{w}|, \underline{\beta}_{X_i} \rangle + \beta_B \right) = e_{3i} \end{aligned} \tag{7.21}$$

şeklindedir.

Burada C parametresi pozitif reel bir sabittir ve algoritmada düzenleyici bir parametre olarak düşünülebilir.

$B = (m_B, \alpha_B, \beta_B)$, Lagrange çarpanları $\lambda_{1i}, \lambda_{2i}$, λ_{3i} ve $\underline{w}$ nin en uygun çözümünü veren Lagrange fonksiyonu şu şekildedir:

$$L = \frac{1}{2}\|w\|^2 + \frac{C}{2}\sum_{k=1}^{3}\sum_{i=1}^{n} e_{ki}^2 + \sum_{i=1}^{n}\lambda_{1i}\left(e_{1i} - m_{Y_i} + \left\langle \underline{w}, \underline{m}_{X_i}\right\rangle + m_B\right)$$
$$-\sum_{i=1}^{n}\lambda_{2i}\left(e_{2i} - \left(m_{Y_i} - \alpha_{Y_i}\right) + \left(\left\langle \underline{w}, \underline{m}_{X_i}\right\rangle + m_B - \left\langle |\underline{w}|, \underline{\alpha}_{X_i}\right\rangle - \alpha_B\right)\right)$$
$$-\sum_{i=1}^{n}\lambda_{3i}\left(e_{3i} - \left(m_{Y_i} + \beta_{Y_i}\right) + \left(\left\langle \underline{w}, \underline{m}_{X_i}\right\rangle + m_B + \left\langle |\underline{w}|, \underline{\beta}_{X_i}\right\rangle + \beta_B\right)\right)$$

Optimallik için koşullar şöyledir.

$$\frac{\partial L}{\partial \underline{w}} = 0 \rightarrow \underline{w} = \sum_{i=1}^{n}\lambda_{1i}\underline{m}_{X_i} + \sum_{i=1}^{n}\lambda_{2i}\left(\underline{m}_{X_i} - \mathrm{sgn}(w)\cdot\underline{\alpha}_{X_i}\right) + \sum_{i=1}^{n}\lambda_{3i}\left(\underline{m}_{X_i} + \mathrm{sgn}(w)\cdot\underline{\beta}_{X_i}\right) \tag{7.22}$$

$$\frac{\partial L}{\partial m_B} = 0 \rightarrow \sum_{k=1}^{3}\sum_{i=!}^{n}\lambda_{ki} = 0 \tag{7.23}$$

$$\frac{\partial L}{\partial \alpha_B} = 0 \rightarrow \sum_{i=!}^{n}\lambda_{2i} = 0 \tag{7.24}$$

$$\frac{\partial L}{\partial \beta_B} = 0 \rightarrow \sum_{i=!}^{n}\lambda_{3i} = 0 \tag{7.25}$$

$$\frac{\partial L}{\partial e_{ki}} = 0 \rightarrow e_{ki} = \frac{\alpha_{ki}}{C}, \quad k = 1,2,3 \tag{7.26}$$

$$\frac{\partial L}{\partial \lambda_{1i}} = 0 \rightarrow m_{Y_i} - \left\langle \underline{w}, \underline{m}_{X_i}\right\rangle - m_B - e_{1i} = 0 \tag{7.27}$$

$$\frac{\partial L}{\partial \lambda_{2i}} = 0 \rightarrow m_{Y_i} - \alpha_{Y_i} - \left\langle \underline{w}, \underline{m}_{X_i}\right\rangle - m_B + \left\langle |\underline{w}|, \underline{\alpha}_{X_i}\right\rangle + \alpha_B - e_{2i} = 0 \tag{7.28}$$

$$\frac{\partial L}{\partial \lambda_{3i}} = 0 \rightarrow m_{Y_i} + \beta_{Y_i} - \left\langle \underline{w}, \underline{m}_{X_i} \right\rangle - m_B - \left\langle |\underline{w}|, \underline{\beta}_{X_i} \right\rangle - \beta_B - e_{3i} = 0 \tag{7.29}$$

Burada $\text{sgn}(w) = (\text{sgn}(w_1),...,\text{sgn}(w_p))$ dir. $t > 0$ veya $t < 0$ olmasına göre $\text{sgn}(t) = 1$ veya -1 değerini alır. $\frac{\partial}{\partial t} | t | = \text{sgn}(t)$ olduğuna dikkat edilmelidir

$B = (m_B, \alpha_B, \beta_B)$ nin en uygun değerleri ve $\lambda_{1i}, \lambda_{2i}$ ve λ_{3i} Lagrange çarpanları aşağıdaki lineer eşitlikten oluşturulabilir.

$$\begin{pmatrix} 0 & 0 & 0 & \underline{1}' & \underline{1}' & \underline{1}' \\ 0 & 0 & 0 & \underline{0}' & \underline{1}' & \underline{0}' \\ 0 & 0 & 0 & \underline{0}' & \underline{0}' & \underline{1}' \\ \underline{1} & \underline{0} & \underline{0} & S_{11} & S_{12} & S_{13} \\ \underline{1} & \underline{-1} & \underline{0} & S'_{12} & S_{22} & S_{23} \\ \underline{1} & \underline{0} & \underline{1} & S'_{13} & S'_{23} & S_{33} \end{pmatrix} \begin{pmatrix} m_B \\ \alpha_B \\ \beta_B \\ \underline{\lambda}_1 \\ \underline{\lambda}_2 \\ \underline{\lambda}_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \underline{m}_Y \\ \underline{m}_Y - \underline{\alpha}_Y \\ \underline{m}_Y + \underline{\beta}_Y \end{pmatrix} \tag{7.30}$$

$$S_{11} = [\left\langle \underline{m}_{X_i}, \underline{m}_{X_j} \right\rangle] + I / C$$

$$S_{12} = [\left\langle \underline{m}_{X_i}, \underline{m}_{X_j} - \text{sgn}(\underline{w}) \cdot \underline{\alpha}_{X_j} \right\rangle]$$

$$S_{13} = [\left\langle \underline{m}_{X_i}, \underline{m}_{X_j} + \text{sgn}(\underline{w}) \cdot \underline{\beta}_{X_j} \right\rangle]$$

$$S_{22} = [\left\langle \underline{m}_{X_i} - \text{sgn}(\underline{w}) \cdot \underline{\alpha}_{X_i}, \underline{m}_{X_j} - \text{sgn}(\underline{w}) \cdot \underline{\alpha}_{X_j} \right\rangle] + I / C$$

$$S_{23} = [\left\langle \underline{m}_{X_i} - \text{sgn}(\underline{w}) \cdot \underline{\alpha}_{X_i}, \underline{m}_{X_j} + \text{sgn}(\underline{w}) \cdot \underline{\beta}_{X_j} \right\rangle]$$

$$S_{22} = [\left\langle \underline{m}_{X_i} + \text{sgn}(\underline{w}) \cdot \underline{\beta}_{X_i}, \underline{m}_{X_j} + \text{sgn}(\underline{w}) \cdot \underline{\beta}_{X_j} \right\rangle] + I / C$$

Burada $\underline{\lambda}_1, \underline{\lambda}_2, \underline{\lambda}_3, \underline{m}_Y, \underline{\alpha}_Y$ ve $\underline{\beta}_Y$, sırasıyla, $\lambda_{1i}, \lambda_{2i}, \lambda_{3i}, m_{Y_i}, \alpha_{Y_i}$ ve β_{Y_i} nin $n \times 1$ lik vektörleridir ve $[a_{ij}]$, $n \times 1$ lik matrisin a_{ij} elemanlarını göstermektedir.

Böylece, LS-SVM ile verilen X_q verisi üzerine $Y(X_q)$ nin öngörü değeri

$$\hat{Y}(\underline{X}_q) = (\langle \underline{w}, \underline{m}_{X_q} \rangle + m_B, \langle |\underline{w}|, \underline{\alpha}_{X_q} \rangle + \alpha_B, \langle |\underline{w}|, \underline{\beta}_{X_q} \rangle + \beta_B) \tag{7.31}$$

olarak elde edilir (Hong ve Hwang 2006).

7.4. Bulanık Lineer Olmayan LS-SVM Regresyon

Bulanık girdi ve çıktı için bulanık çok değişkenli regresyon modeli tahmini LS-SVM ye dayalı bulanık lineer regresyon modeli oluşturarak elde edilir. Bunu yapmak için, bulanık olmayan nonlineer regresyon için LS-SVM de kullanılan düşünceye kısaca göz atmak gerekir. Bu da bir özellik uzayı F ye $\phi : \Re^d \to F$ ile eşleştirilen x_i girdi örnekleri üretme ön süreci ile kolaylıkla yapılabilir. Sonra standart LS-SVM regresyon algoritması uygulanır. Başlangıç olarak, algoritmada veri nokta biçiminde $(\langle x_i, x_j \rangle)$ gösterilecektir. Algoritma sadece F deki noktalar yoluyla veriye bağlı olacaktır, yani, $\langle \phi(x_i), \phi(x_j) \rangle$ biçimindeki fonksiyonlardır. Böylece, $\phi(\cdot)$ yerine $K(x_i, x_j) = \langle \phi(x_i), \phi(x_j) \rangle$ yı bilmek ve kullanmak yeterli olacaktır. Regresyon problemi için iyi tanımlı kernel fonksiyonları aşağıda verilmiştir.

$K(\underline{x}, \underline{y}) = \left(1 + \langle \underline{x}, \underline{y} \rangle\right)^p$: Polinomial kernel

$K(\underline{x}, \underline{y}) = e^{-\frac{\|x-y\|^2}{2\sigma^2}}$: Gaussian kernel

Örnek olarak iki girdiye sahip ve derecesi 2 olan polynomial kernel bir özellik uzayı düşünülsün.

$$\begin{aligned} K(\underline{x},\underline{y}) &= \left(1+\left\langle \underline{x},\underline{y}\right\rangle\right)^2 \\ &= \left(1+x_1y_1+x_2y_2\right)^2 \\ &= 1+2x_1y_1+2x_2y_2+(x_1y_1)^2+(x_2y_2)^2+2x_1y_1x_2y_2 \end{aligned}$$

Eğer, $\phi(x)=(1,\sqrt{2}x_1,\sqrt{2}x_2,x_1^2x_2^2,\sqrt{2}x_1x_2)$ alınırsa, o zaman $K(x,y)=\langle\phi(x),\phi(y)\rangle$ olur.

ϕ fonksiyonu bulanık girdiler için artan olmalıdır, çünkü, üçgensel bulanık sayının biçimi, ϕ ile fonksiyonunda korunmalıdır. $m_{X_{ij}}-\alpha_{X_{ij}}\geq 0$ olduğu zaman polynomial kernelle ilgili eşleştirme fonksiyonu Φ artandır ve bu yüzden üçgensel bulanık sayının biçimi korunur. Ancak, Gaussian kernelle ilgili eşleştirme fonksiyonu bu durumda yeterli değildir. Bu yüzden, bu çalışmada her durumda tüm veriye uygulanacak basit bir dönüştürme ile $m_{X_{ij}}-\alpha_{X_{ij}}\geq 0$ olduğu varsayıldı ve yalnızca polynomial kernelle ilgili eşleştirme fonksiyonu düşünüldü. Gerçekte, kernel fonksiyona tercihen eşleştirme fonksiyonu kullanıldı. $\phi^*: T(R)^d \rightarrow T(F)$ ile tanımlı ϕ^* fonksiyonu tanımlansın. Böylece,

$$\phi^*((m_X,\alpha_X,\beta_X))=(\phi(m_X),\phi(m_X)-\phi(m_X-\alpha_X),\phi(m_X+\beta_X)-\phi(m_X))$$

olur. $\alpha_{X_i}^{\phi}$ ve $\beta_{X_i}^{\phi}$ şöyle tanımlansın.

$$\underline{\alpha}_{X_i}^{\phi}=\phi\left(\underline{m}_{X_i}\right)-\phi\left(\underline{m}_{X_i}-\underline{\alpha}_{X_i}\right)$$
$$\underline{\beta}_{X_i}^{\phi}=\phi\left(\underline{m}_{X_i}+\underline{\beta}_{X_i}\right)-\phi\left(\underline{m}_{X_i}\right)$$

Lineer durumdakine benzer olarak şu eşitliği elde ederiz.

$$\underline{w}^{\phi} = \sum_{i=1}^{n} \lambda_{1i} \phi\left(\underline{m}_{X_i}\right) + \sum_{i=1}^{n} \lambda_{2i} \left(\phi\left(\underline{m}_{X_i}\right) - \operatorname{sgn}(\underline{w}^{\phi}) \cdot \underline{\alpha}^{\phi}_{X_i}\right) + \sum_{i=1}^{n} \lambda_{3i} \left(\phi\left(\underline{m}_{X_i}\right) + \operatorname{sgn}(\underline{w}^{\phi}) \cdot \underline{\beta}^{\phi}_{X_i}\right) \tag{7.32}$$

Bundan başka, $\underline{w}, \underline{m}_{X_i}, \underline{\alpha}_{X_i}, \underline{\beta}_{X_i}$ in $\underline{w}^{\phi}, \phi(\underline{m}_{X_i}), \underline{\alpha}^{\phi}_{X_i}, \underline{\beta}^{\phi}_{X_i}$ ile yer değiştirmesi sonucu lineer eşitlik ve çözümü, $\underline{X}_q$ verisi üzerine $\hat{Y}(X_q)$ lineer olmayan öngörü değeri şöyle verilebilir (Hong ve Hwang 2006):

$$\hat{Y}(\underline{X}_q) = (\langle \underline{w}^{\phi}, \phi(\underline{m}_{X_q}) \rangle + m_B, \langle | \underline{w}^{\phi} |, \underline{\alpha}^{\phi}_{X_q} \rangle + \alpha_B, \langle | \underline{w}^{\phi} |, \underline{\beta}^{\phi}_{X_q} \rangle + \beta_B) \tag{7.33}$$

LS-SVM'ye dayalı lineer bulanık regresyon modelinin adımları şu şekilde sıralanır:

Adım 1 C sabiti 1 alınarak (7.30)'daki lineer eşitlik oluşturulur.

Adım 2 Elde edilen denklem sistemi çözülerek $\underline{\lambda}_1, \underline{\lambda}_2, \underline{\lambda}_3$ değerleri elde edilir.

Adım 3 Bu değerler (7.22)'de yerine konularak $\underline{w}$ katsayıları bulunur.

Adım 4 Optimallik koşullarının sağlanıp sağlanmadığı kontrol edilir.

Adım 5 Regresyon denklemi eşitlik (7.31) kullanılarak oluşturulur.

Lineer olmayan bulanık regresyon modeli için de aynı işlemler uygulanır (Hong ve Hwang 2006).

8. DOĞRUSAL PROGRAMLAMA İLE BULANIK EN KÜÇÜK MUTLAK SAPMA REGRESYONU

8.1. Simpleks Yöntem

Simpleks yöntem, herhangi bir doğrusal programlama problemini (dpp) sınırlı adım sonunda tam olarak çözen ya da sınırsız çözüm olduğunu belirten yinelemeli bir işlemdir. Bir dpp'nin en iyi çözümü varsa uç noktalardan biri en iyi çözümdür ve sınırlı sayıda uç nokta vardır. Simpleks yöntem bir uç noktadan en iyi uç noktaya adım adım yaklaşma işlemidir (Apaydın 2005). Her adımda en iyi uç noktaya daha yakın olan uç noktaya geçilir. Her uç noktaya gelindiğinde, bu yöntem, uç noktanın en iyi olup olmadığını, değilse bir sonraki uç noktanın hangisi olacağını belirtir.

Tanım 8.1. n bilinmeyenli m denklemden oluşan bir eşanlı doğrusal denklem sistemi,

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{k1}x_1 + a_{k2}x_2 + \cdots + a_{kn}x_n &= b_k \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

olarak tanımlanır. Matris gösterimi ile,

$$\underline{A}x = \underline{b}$$

biçiminde verilir.

$(A:b)$, m satır ve $(n+1)$ kolondan oluşan eklemeli matris olsun. Bu sistemin çözümü için iki durum göz önüne alınacaktır.

Durum 1: $(A:b)$'nin rankı $\underline{A}$'nın rankından büyük ise, $\underline{b}$, $a_1, a_2, ..., a_n$'nin doğrusal bileşimi olarak yazılamaz ve $\underline{A}x = \underline{b}$ sisteminin çözümü yoktur.

Durum 2: $r(\underline{A}) = r(A:b) = k$ olduğunu varsayalım. Yani doğrusal bağımsız satır sayısı k olsun. Bu durumda $(A:b)$ eklemeli matrisi,

$$(A:b) = \begin{bmatrix} A_1 & b_1 \\ A_2 & b_2 \end{bmatrix}$$

şeklindedir. Burada $(A_1)_{k\times n}$ ve $(A_2)_{(m-k)\times n}$ matris ve $(b_1)_{k\times 1}$ ve $(b_2)_{(m-k)\times 1}$ vektördür. $r(A_1) = r(A_1 : \underline{b}_1) = k$ olur. Eğer $\underline{x}$ vektörü $A_1\underline{x} = \underline{b}_1$ sistemini sağlarsa, $A_2\underline{x} = \underline{b}_2$ 'yi de sağlar. Böylece $A_2\underline{x} = \underline{b}_2$ doğrusal bağımlı kısıtları gözardı edilip $A_1\underline{x} = \underline{b}_1$ bağımsız kısıtları ayrılabilir. $r(A_1) = k$ olduğundan A_1 'in k tane doğrusal bağımsız kolonunu seçilebilir. A_1 matrisi $A_1 = (BN)$ olarak yazılsın. Burada $(B)_{k\times k}$ boyutlu tekil olmayan kare matristir. $r(A_1) = k$ olduğundan böyle bir B matrisi yazılabilir ve temel matris olarak adlandırılır. $(N)_{k\times(n-k)}$ ise temel olmayan matris olarak adlandırılır.

$\underline{x}$ vektörü de,

$$\underline{x} = \begin{bmatrix} \underline{X}_B \\ \underline{X}_N \end{bmatrix}; \quad \begin{array}{l} \underline{X}_B = [x_1 \quad x_2 \quad \dots \quad x_k] \\ \underline{X}_N = [x_{k+1} \quad x_{k+2} \quad \dots \quad x_n] \end{array}$$

biçiminde ayrılabilir. Burada $\underline{X}_B$ temel değişken, $\underline{X}_N$ ise temel dışı değişken olarak tanımlanır ve buradan,

$$\begin{aligned} A_1\underline{x} &= \underline{b}_1 \\ (B,N)\begin{bmatrix} \underline{X}_B \\ \underline{X}_N \end{bmatrix} &= \underline{b}_1 \\ B\underline{X}_B + N\underline{X}_N &= \underline{b}_1 \end{aligned}$$

olur. B 'nin kolonları bağımsız olduğu için tersi alınabilir. Eşitliğin her iki tarafı B^{-1} ile çarpıldığında,

$$\underline{X}_B = B^{-1}\underline{b}_1 - B^{-1}N\underline{X}_N$$

bulunur. Bu n bilinmeyenli k eşitlikte aşağıdaki durumlarla karşılaşılır:

a) $k = n$ olduğunda $A_1\underline{x} = \underline{b}_1$ sisteminin,

$$\underline{X}_B = B^{-1}\underline{b}_1$$

den bulunan tek çözümü vardır.

b) $n > k$ olduğunda $\underline{X}_N$ vektörüne keyfi değerler verilir. Bu durumda $A_1\underline{x} = \underline{b}_1$ sisteminin sınırsız çözümü vardır (Apaydın 2005).

8.1.1. Doğrusal programlama probleminin standart biçimi

Genel bir doğrusal programlama probleminin biçimi matematiksel olarak;

$$Max(Min)z = \sum_{j=1}^{n} c_j x_j \tag{8.1}$$

$$\sum_{j=1}^{n} a_{ij} x_j \{\leq, =, \geq\} b_i, \quad i = 1,...,m \tag{8.2}$$

$$x_j \geq 0, \qquad i = 1,...,n \tag{8.3}$$

şeklindedir. Burada a_{ij}, b_i ve c_j 'ler bilinen değişmezler, x_j 'ler karar değişkeni olarak tanımlanmaktadır. (8.2) kısıtlarını sağlayan herhangi bir x_j çözümü, (8.33) eşitliği ile verilen negatif olmama kısıtlarını da sağlarsa uygun çözüm olur. z 'yi en küçükleyen veya en büyükleyen herhangi uygun çözüm en iyi çözüm adını alır.

Dpp'nin genel bir çözümle çözülebilmesi için standart biçime dönüştürülmesi gerekir. Genel biçimde verilen matematiksel modelde eşitsizlik kısıtları eşitlik haline getirilerek standart biçim,

$$Max(Min)z = \sum_{j=1}^{n} c_j x_j$$

$$\sum_{j=1}^{n} a_{ij} x_j = b_i, \quad i = 1,...,m$$

$$x_j \geq 0, \qquad i = 1,...,n$$

olarak elde edilir. Standart biçime döndürme işleminde şunlar yapılır:

1) Tüm kısıtlar eşitlik biçimindedir ve sağ taraftaki b_i'ler negatif değildir. Eşitsizlik kısıtlarının eşitlik haline getirilmesinde gevşek ve fazla değişkenler kullanılır. Sağ taraf değerlerinde (b_i) negatif olan varsa (-1) ile çarpılarak pozitif hale getirilir.

2) Tüm değişkenler negatif olmayandır.

3) Amaç fonksiyonu en küçükleme ya da en büyükleme şeklindedir.

Standart biçimde değişken sayısı, eklenen veya çıkarılan gevşek ve fazla değişkenlere bağlı olarak n ile $n+m$ arasında değişir. Ancak bu değişkenler eklendiğinde amaç fonksiyonunun değişmemesi istenir. Bu nedenle bu değişkenlere ilişkin c_j değerlerinin sıfır olduğu varsayılır (Apaydın 2005).

8.1.2. Temel çözümler

n bilinmeyenli m tane eşanlı doğrusal denklem sistemi,

$$A\underline{x} = \underline{b}$$

nin $(n > m)$ çözümlerini inceleyelim. $r(A) = k$ ise A'dan k tane doğrusal bağımsız kolon seçilir. Geri kalan $(n-k)$ değişken bu kolonlarla ilişkili değildir ve herhangi değerler

verilebilir. Böyle bir sistemde $(n-k)$ değişkene sıfır verilirse geri kalan k değişken eşitliklerin sağlanması için sıfırdan farklıdır.

$r(A) = r(A:b) = m$ olduğunu varsayalım.

Tanım 8.2. n bilinmeyenli m tane eşanlı doğrusal denklem sistemi $A\underline{x} = \underline{b}$ verilsin. A'dan herhangi tekil olmayan $m \times m$ boyutlu bir B matrisi seçilir ve bu kolonlarla ilgili olmayan geri kalan $n-m$ değişkene sıfır verilirse bu sistem eşitliklerinin sonucu temel çözüm adını alır. Sıfırdan farklı olan m değişken ise temel değişkenler olarak adlandırılır ve

$$\underline{X}_B = B^{-1}\underline{b}$$

ile bulunur. Geri kalan $n-m$ temel olmayan değişkenlerdir.

$\underline{X}_B > \underline{0}$ ise temel uygun çözümdür. En az bir ögesi sıfır ise bozulmuş temel uygun çözüm adını alır. Temel çözüm sayısı,

$$\binom{n}{m} = \frac{n!}{(n-m)!m!}$$

ile bulunur. Ancak bu etkin bir yol değildir. Çünkü değişken sayısı arttıkça temel çözüm sayısı hızla artacaktır. Ayrıca yalnız temel çözümlerin incelenmesi problemin sınırsız olup olmadığını belirlemeye yetmez. Uç noktalar arasında yinelemeli bir yöntemle en iyi uç nokta bulunmalıdır.

$A\underline{x} = \underline{b}$ sisteminin çözümünün olabilmesi için $[A:b]$ eklemeli matrisinden oluşturulan her m kolon kümesinin doğrusal bağımsız olması gerekir.

8.1.2.1. Temel uygun çözümü iyileştirme

Bir temel uygun çözüm verildiğinde amaç fonksiyonuna daha iyi değer veren yeni temel uygun çözümün nasıl bulunacağını araştıralım. Aşağıdaki dpp verilsin.

$$Min\, z = c\underline{x}$$
$$A\underline{x} = \underline{b}$$
$$\underline{x} \geq \underline{0}$$

Burada A, $(m \times n)$ boyutlu matristir ve $r(A) = m$ 'dir. Bu sistemin temel uygun çözümü,

$$\underline{x} = \begin{bmatrix} B^{-1}\underline{b} \\ \underline{0} \end{bmatrix} = \begin{bmatrix} \underline{X}_B \\ \underline{X}_N \end{bmatrix}$$

olsun. Bu çözümün amaç fonksiyonuna verdiği değer,

$$\begin{aligned} z_0 = \underline{c}\begin{bmatrix} B^{-1}\underline{b} \\ \underline{0} \end{bmatrix} &= \begin{bmatrix} \underline{C}_B & \underline{C}_N \end{bmatrix}\begin{bmatrix} B^{-1}\underline{b} \\ \underline{0} \end{bmatrix} \\ &= \underline{C}_B B^{-1}\underline{b} \\ &= \underline{C}_B \underline{X}_B \end{aligned} \qquad (8.4)$$

dir. Şimdi $\underline{\hat{x}} = \begin{bmatrix} \underline{\hat{X}}_B \\ \underline{\hat{X}}_N \end{bmatrix}$ gibi herhangi bir uygun çözüm alalım. Burada $\underline{\hat{X}}_B \geq 0$, $\underline{\hat{X}}_N \geq 0$ dır. Kısıt kümesinde bu uygun çözüm yerine konulursa,

$$\begin{aligned} A\underline{\hat{x}} &= \underline{b} \\ \begin{bmatrix} B & N \end{bmatrix}\begin{bmatrix} \underline{\hat{X}}_B \\ \underline{\hat{X}}_N \end{bmatrix} &= \underline{b} \\ B\underline{\hat{X}}_B + N\underline{\hat{X}}_N &= \underline{b} \qquad (8.5) \\ \underline{\hat{X}}_B &= B^{-1}\underline{b} - B^{-1}N\underline{\hat{X}}_N \\ \underline{\hat{X}}_B &= B^{-1}\underline{b} - \sum_{j \in R} B^{-1}a_j x_j \end{aligned}$$

bulunur. (8.5) eşitliğinde R , temel dışı vektörlere ilişkin indis kümesidir. $\hat{x}$, uygun çözümünün amaç fonksiyonuna verdiği değer,

$$z = \underline{c}\hat{\underline{x}} = \begin{bmatrix} \underline{C}_B & \underline{C}_N \end{bmatrix} \begin{bmatrix} \hat{\underline{X}}_B \\ \hat{\underline{X}}_N \end{bmatrix} = \underline{C}_B \hat{\underline{X}}_B + \underline{C}_N \hat{\underline{X}}_N \tag{8.6}$$

olarak bulunur. (8.6) eşitliğinde (8.5) eşitliği konulduğunda,

$$\begin{aligned} z &= \underline{c}B\left[B^{-1}\underline{b} - \sum_{j\in R} B^{-1}a_j x_j \right] + \underline{C}_N \hat{\underline{X}}_N \\ &= \underline{C}_B B^{-1}\underline{b} - \underline{C}_B \sum_{j\in R} B^{-1}a_j x_j + \sum_{j\in R} c_j x_j \\ &= z_0 - \sum_{j\in R} (C_B B^{-1}a_j - c_j)x_j \\ z &= z_0 - \sum_{j\in R} (z_j - c_j)x_j \end{aligned} \tag{8.7}$$

bulunur. (8.7) eşitliğindeki tüm temel dışı değişkenlere ilişkin,

$$z_j = \underline{C}_B B^{-1} a_j$$

değerinden

$$\begin{aligned} z_j &= y_{1j}C_{B1} + \cdots + y_{mj}C_{Bm} \\ &= \sum_{i=1}^{m} y_{ij}C_{Bi}, \qquad i = 1,...,n \\ z &= \underline{C}_B y_j \end{aligned}$$

olarak da tanımlanır.

(8.7) eşitliği en son temek uygun çözümü iyileştirme için yol gösterir. Burada amaç z 'yi en küçüklemektir. $x_j > 0$ olacağından $(z_j - c_j) > 0$ olacak biçimde a_j vektörü

seçilirse $z < z_0$ olur. Böylece amaç fonksiyonu küçülür. Uygulamada hesaplamaları mümkün olduğunca azaltmak için $(z_j - c_j) > 0$ değerleri içinden en büyüğü seçilir. Bu durumda temele girecek olan yeni vektör,

$$z_k - c_k = Max\{z_j - c_j ; z_j - c_j > 0\} \tag{8.8}$$

ölçütüne göre seçilir. Buradaki k indisi temele a_k vektörünün gireceğini ifade eder. Bu seçimden sonra yeni amaç fonksiyonunun değeri,

$$z = z_0 - (z_k - c_k)x_k \tag{8.9}$$

olur. Yeni temel değişkenin değeri (8.5)'ten,

$$\underline{\hat{X}}_B = B^{-1}\underline{b} - B^{-1}a_k x_k$$
$$\underline{\hat{X}}_B = \underline{X}_B - y_k x_k$$

olarak elde edilir.

Temele girecek a_k vektörü belirlendikten sonra, temelden çıkacak vektör belirlenir. $B = (b_1, b_2, ..., b_r, ..., b_m)$ temelinde r. ögenin yerine a_k vektörünün girdiğini varsayalım. Yeni temel, $\hat{B} = (b_1, b_2, ..., a_k, ..., b_m)$ olacaktır. Temel dışındaki bir vektör temel vektörlerin doğrusal bileşimi olarak yazılabiliyordu. O halde,

$$a_k = \sum_{i=1}^{m} y_{ik}\underline{b}_i = y_{1k}b_1 + \cdots + y_{rk}b_r + y_{mk}b_m \tag{8.10}$$

yazılabilir. Matris gösterimi ile,

$$a_k = By_k \Rightarrow y_k = B^{-1}a_k$$

dir. (8.10) eşitliğinde $y_{rk} \neq 0$ ise a_k ile b_r sütunları yer değiştirebilir ve

$$b_r = \frac{1}{y_{rk}} a_k - \sum_{i=1}^{m} \frac{y_{ik}}{y_{rk}} b_i , \qquad i \neq r \tag{8.11}$$

elde edilir. Bir dpp'nin temel uygun çözümü

$$\underline{X}_B = B^{-1}\underline{b} \;\Rightarrow\; B\underline{X}_B = \underline{b}$$

$$\sum_{i=1}^{m} X_{B_i} b_i = \underline{b}$$

$$X_{B_1} b_1 + \cdots + X_{B_r} b_r + \cdots + X_{B_m} b_m = \underline{b}$$

şeklinde yazılabilir. b_r yerine (8.10) eşitliği yerine yazılırsa,

$$X_{B_1} b_1 + \cdots + \left[\frac{1}{y_{rk}} a_k - \sum_{i=1}^{m} \frac{y_{ik}}{y_{rk}} b_i \right] + \cdots + X_{B_m} b_m = \underline{b}$$

elde edilebilir. Düzenlenirse

$$\sum_{\substack{i=1 \\ i \neq r}}^{m} X_{B_i} b_i + \frac{X_{B_r}}{Y_{rk}} a_k - X_{B_r} \sum_{\substack{i=1 \\ i \neq r}}^{m} \frac{y_{ik}}{y_{rk}} b_i = \underline{b}$$

$$\sum_{i=1}^{m} \left(X_{B_i} - X_{B_r} \frac{y_{ik}}{y_{rk}} \right) b_i + \frac{X_{B_r}}{y_{rk}} a_k = \underline{b}$$

bulunur. $\underline{X}_B$ değişkenlerinin yeni temel çözümdeki değerleri $\hat{X}_{B_i}$ ile gösterilirse,

$$\left.\begin{aligned} a)\; \hat{X}_{B_i} &= X_{B_i} - X_{B_r} \frac{Y_{ik}}{Y_{rk}} > 0 \qquad i \neq r \\ b)\; \hat{X}_{B_r} &= \frac{X_{B_r}}{Y_{rk}} \end{aligned}\right\} \tag{8.12}$$

elde edilir. temel çözümün uygun olması için (8.12) eşitliğindeki yeni değişkenlerin (≥ 0) olması gerekir. Bu nedenle y_{rk}'ların seçimi önemlidir. (8.12) eşitliğinin (b) şıkkında X_{B_r} ele aldığımız ilk temelin elemanlarıdır. O halde X_{B_r}>0'dır. bu durumda $X_{B_r} \geq 0$ olması için $y_{rk} > 0$ olmalıdır. (a) şıkkı için; X_B ve X_{B_r} ilk temelin elemanları olduğu için (≥ 0)'dır. (b) şıkkından $y_{rk} > 0$ olması gerektiği görülür. O halde $y_{ik} < 0$ olursa $\hat{X}_{B_i} \geq 0$ olması sağlanır.

Ancak $y_{ik} > 0$ olduğu durumda (a) şıkkındaki

$$X_{B_i} - X_{B_r} \frac{y_{ik}}{y_{rk}} \geq 0$$

eşitsizliği y_{ik}'ya bölünerek

$$\frac{X_{B_i}}{y_{ik}} - \frac{X_{B_r}}{y_{rk}} \geq 0$$

$$\frac{X_{B_i}}{y_{ik}} \geq \frac{X_{B_r}}{y_{rk}} \text{ ya da } \frac{X_{B_i}}{y_{ik}} \leq \frac{X_{B_r}}{y_{rk}} ; \qquad y_{ik} > 0$$

bulunur. Burada temelden çıkarılacak olan $r.$ vektör , X_{B_i} / y_{ik}'ların en küçüğü olacaktır. O halde B temelinin $r.$ sütunu,

$$\frac{X_{B_r}}{y_{rk}} = Min\left\{\frac{X_{B_i}}{y_{ik}},\ y_{ik} > 0\right\} \qquad (8.13)$$

ölçütüne göre seçilirse yeni temel uygun çözüm olur. Birden fazla küçük değer var ise herhangi biri keyfi olarak seçilebilir. Yine bir temel uygun çözüm elde edilir. Eşit değerli bu sütunlar için,

$$\frac{X_{B_i}}{y_{ik}} = \frac{X_{B_r}}{y_{rk}}$$

olacağından $\hat{X}_{B_i} = 0$ bulunur. Bu durumda yeni temel çözüm bozulmuştur. r. değişken ise,

$$\hat{X}_{B_r} = X_{B_r} - y_{rk}\frac{X_{B_r}}{y_{rk}} = 0 = \hat{X}_{B_i} \ , \qquad i = r$$

olarak bulunur. Burada ortaya çıkabilecek durumlar şunlardır:

1) Eğer temele giren a_k vektörüne ilişkin tüm $y_{ik} \le 0$ ise problemin en iyi çözümü sınırlanmamıştır. (8.7) eşitliğinde $(z_k - c_k) > 0$ olduğundan a_k 'ya ilişkin x_k değişkeni istenildiği kadar artırıldığında yine $z \le z_0$ bulunur.

$$\underline{\hat{X}}_B = \underline{X}_B - y_k x_k$$

eşitliğinde $y_k \le 0$ ise x_k değişkeni istenildiği kadar artırılabilir. $\underline{\hat{X}}_B$ 'lar yine pozitif olacaktır. Bu durumda problemin en yi çözümü sınırlanmamıştır.

2) Eğer tüm temel dışı değişkenler için $z_j - c_j < 0$, ancak tek $z_k - c_k = 0$ ise problemin seçenek en iyi çözümü vardır. Çünkü $z_k - c_k = 0$ olan x_k değişkeni temele girerse z amaç fonksiyonunun değeri değişmez. (8.7) eşitliğinden

$$z = z_0$$

bulunur.

3) Bozulma olmaması durumunda, yani $X_{B_r} > 0$ iken, amaç fonksiyonunun değeri her yinelemede kesinlikle azalır. Sınırlı sayıda temel çözüm olduğundan, yöntem sınırlı

sayıda adım sonucunda ya sonlu bir en iyi çözüm ile ya da en iyi çözümün sınırlanmadığı durum ile son bulur.

Bozulma durumu $X_{B_r} = 0$ olduğunda, temele giren x_k değişkeni için en büyük artış sıfır olur. Bu durumda amaç fonksiyonunun değeri bir önceki ile aynıdır. Burada aynı uç nokta ve amaç değerine sahip sınırsız sayıda temeller dizisi bulunabilir. Bu durum döngü olarak adlandırılır.

Buraya kadar verilen bilgilerden

$$Min\, z = \underline{c}\underline{x}$$
$$A\underline{x} = \underline{b}$$
$$\underline{x} \geq \underline{0}$$

dpp için Simpleks yöntem şöyle özetlenebilir.

Adım 1 Bir başlangıç uygun çözümü veren B temeli seçilir.

Adım 2 $\underline{X}_B = B^{-1}\underline{b}$, $\underline{X}_N = \underline{0}$ ve $z = \underline{C}_B \underline{X}_B$ hesaplanır.

Adım 3 Adım 2'de bulunan z değerinin en küçük olup olmadığının testi için temelde olmayan değişkenlere ilişkin $(z_j - c_j)$'ler hesaplanır.

$$z_k - c_k = Max\{z_j - c_j ; z_j - c_j > 0\}$$

ölçütünü sağlayan hiçbir değişken yoksa yani tüm $(z_j - c_j) \leq 0$ ise son bulunan çözüm en iyi çözümdür. Değilse dördüncü adıma geçilir.

Adım 4 $z_k - c_k$'ya ilişkin $k.$ vektör temele alınır ve

$$y_k = B^{-1}a_k$$

hesaplanır. Eğer bütün $y_k \leq 0$ ise en iyi çözüm, sınırlanmamış ya da sınırsızdır denir. Eğer $y_k > 0$ ise beşinci adıma geçilir.

Adım 5 Temelden çıkacak olan değişken (8.13) eşitliği ile verilen

$$\frac{X_{B_r}}{y_{rk}} = Min\left\{\frac{X_{B_i}}{y_{ik}},\ y_{ik} > 0\right\}$$

ölçütü ile belirlenir ve temelin $r.$ kolonu temelden çıkarılır.

Adım 6 Beşinci adımda bulunan temelin $r.$ vektörü çıkarılır, temel dışındaki a_k vektörü dolayısıyla x_k değişkeni temel alınır. Bu şekilde yeni temel oluşturulur. Bu temelin vereceği uygun çözüm ve z 'nin hesaplanması için yeniden ikinci adıma dönülür.

En büyükleme probleminin Simpleks yöntem ile çözümünde yukarıdaki adımlar uygulanır. Ancak temele girecek vektör seçiminde

$$z_k - c_k = Min\{z_j - c_j ; z_j - c_j < 0\} \qquad (8.14)$$

ölçütü kullanılır. Tüm temel olmayan değişkenlere ilişkin

$$(z_j - c_j) \geq 0$$

ise en iyi çözüme ulaşılır (Apaydın 2005).

8.1.3. Başlangıç temel uygun çözümün bulunması

$$A\underline{x} \leq \underline{b}$$
$$\underline{x} \geq \underline{0}$$

kısıt kümesi verilsin. Bu kısıtlar standart biçime dönüştürüldüğünde x_G gevşek değişkenlerinin eklenmesi ile

$$A\underline{x} + I\underline{x}_G = \underline{b}$$
$$\underline{x}, \underline{x}_G \geq \underline{0}$$

denklem sistemi oluşur ve $\begin{bmatrix} A & I \end{bmatrix}$ matrisi elde edilir. Burada A, $m \times (m+n)$ boyutlu katsayılar matrisi, I ise $m \times m$ boyutlu birim matristir. Tanımlamalardan

$$\begin{bmatrix} A & I \end{bmatrix} \begin{bmatrix} \underline{x} \\ \underline{x}_G \end{bmatrix} = \underline{b}$$

yazılabilir. $\underline{x}$ temel dışı vektör olduğunda $\underline{x} = \underline{0}$ olur. Buradan $\underline{x}_G = \underline{b}$ 'dir. $\underline{b} > 0$ olduğu için $\underline{x}_G > 0$ olacaktır. Dolayısıyla uygun çözümdür. $B = I = B^{-1}$ olduğundan birim matris, başlangıç temel uygun çözüm olarak alınabilir. Başlangıç temelin birim matris alınmasının birçok kolaylığı vardır. Tersi kendisine eşit olduğundan ters alma işlemi yapılmaz.

$$y_j = B^{-1}a_j = I = a_j \qquad j = 1,...,n$$

vektörlerinin hesaplanmasında da kolaylık sağlar. Gevşek değişkenlere ilişkin $C_B = 0$ olduğundan,

$$z_j - c_j = \underline{C}_B y_j - c_j = -c_j$$
$$z = \underline{C}_B \underline{X}_B = 0$$

olur. Birim matris problemin yapısından gelirse $\underline{C}_B \neq 0$ 'dır. Bu durumda

$$z = \underline{C}_B \underline{X}_B$$
$$z_j - c_j = \underline{C}_B y_j - c_j$$

olur (Apaydın 2005).

8.1.3.1. Temel uygun çözümün bulunmasında bazı aykırı durumlar

Standart biçime getirilmiş bazı dpp'nin katsayılar matrisinde birim matris oluşmaz. Bu durumda bir başlangıç temel uygun çözüm bulmak zordur. Sisteme uygun çözümü veren çeşitli temelleri denemek gerekir. Fakat kısıt ve değişken sayısı büyük olduğunda olası temel sayısı artacağı için her temeli denemek güçtür. Bu nedenle

$$A\underline{x}\begin{Bmatrix}=\\ \leq\\ \geq\end{Bmatrix}\underline{b}$$

biçiminde verilen dpp'nde birim matris oluşturabilmek için yapay değişkenler tanımlanır.

$$A\underline{x} = \underline{b}$$
$$\underline{x} \geq \underline{0}$$

şeklinde standart biçime dönüştürülmüş dpp'nin olduğunu düşünelim. İlk temel olarak bir birim matris seçilmesinin bize sağladığı yarar düşünülerek A matrisinin yanına $m \times m$'lik bir matris eklenir. Bu, her kısıt denklemine x_{ai} gibi bir yapay değişken eklenmesi demektir. Yapay değişkenin eklenmesiyle sistem,

$$A\underline{x} + I\underline{x}_G = \underline{b}$$
$$\underline{x}, \underline{x}_G \geq \underline{0}$$

olur. $B = I$ alınarak yeni sisteme temel uygun çözüm bulunur ve $\underline{X}_B = I\underline{q} = \underline{b}$ olur. Burada $\underline{q}$, yapay değişkenlere ilişkin vektördür.

Genişletilmiş sistem için bulunan çözüm orijinal problem için en iyi çözüm değildir. Çünkü temel olarak yapay vektörleri aldığımızdan sonuçta bulunan yapay değişkenlerdir.

Orijinal problemin en iyi çözümünün bulunabilmesi için, yapay değişkenlerin sıfır olması gerekir. Çünkü $A\underline{x} = \underline{b}$ sisteminin çözümü yalnız ve ancak $A\underline{x} + \underline{q} = \underline{b}$ sisteminde $\underline{q} = \underline{0}$ ise mümkündür. Bu sistemin çözümü için çeşitli yöntemler mevcuttur (Apaydın 2005).

8.2. Bulanık En Küçük Mutlak Sapma Regresyonu

8.2.1. Genel bulanık lineer regresyon

Genel bir bulanık lineer regresyon aşağıdaki biçimde verilir.

$$Y_i = A_0(+)A_1(\times)x_{i1}(+)\cdots(+)A_p(\times)x_{ip}\{+\}\varepsilon_i,\ i = 1,...,n \qquad (8.15)$$

Burada Y_i bulanık çıktıları ve $x_{i1},...,x_{ip}$ girdi birimlerini göstermektedir. $\underline{A} = (A_0,...,A_p)$ sistemin bulanık parametreleri veya regresyon katsayılarıdır. ε_i'ler gözlemsel hatalardır ve yalnızca sıfır ortalamalı rasgele hatalar olarak değil, sistemin bulanıklığından kaynaklanan hatalar olarak da düşünülebilir (Kim ve ark. 1994). Hem bulanıklık hem de rasgelelik içeren karma bir hata da olabilir. $(+)$ ve $(\times)$ bulanık toplama ve çarpma işlemlerini göstermektedir. $\{+\}$ ise daha sonra tanımlanacak bir operatördür.

A_j bulanık parametreleri LR-tip bulanık sayılar olarak tanımlanır ve $A_j = (a_j, c_j^L, c_j^R)_{LR}$ şeklinde gösterilir.

Y_i *, genel bulanık lineer regresyonun çıktısını göstersin. $x_{ij} \geq 0, i = 1,...,n$ ve $j = 1,...,p$ olmak üzere, Y_i * bulanık aritmetik kullanılarak LR-tip bulanık sayılarla şöyle ifade edilir.

$$
\begin{aligned}
Y_i^* &= A_0(+)A_1(\times)x_{i1}(+)\cdots(+)A_p(\times)x_{ip} \\
&= \left(a_0 + \sum_{j=1}^{p} a_j x_{ij}, c_0^L + \sum_{j=1}^{p} c_j^L x_{ij}, c_0^R + \sum_{j=1}^{p} c_j^R x_{ij} \right)_{LR} \\
&= \left(a^t x_i, (c^L)^t x_i, (c^R)^t x_i \right)
\end{aligned}
\tag{8.16}
$$

Burada $\underline{a} = (a_0, a_1, ..., a_p)^t$, $\underline{c}^L = (c_0^L, c_1^L, ..., c_p^L)^t$, $\underline{c}^R = (c_0^R, c_1^R, ..., c_p^R)^t$ ve $\underline{x}_i = (1, x_{i1}, ..., x_{ip})^t$ dir. Y_i^* çıktılarının üyelik fonksiyonu (2.29)'dan şöyle yazılabilir.

$$
\mu_{Y_i^*}(\underline{y}) = \begin{cases}
1 - \left(\dfrac{\underline{a}^t \underline{x}_i - \underline{y}}{(\underline{c}^L)^t \underline{x}_i} \right), & (\underline{c}^L)^t \underline{x}_i \geq (\underline{a}^t \underline{x}_i - \underline{y}) \geq 0, \;\; (\underline{c}^L)^t \underline{x}_i > 0 \\
1 - \left(\dfrac{\underline{y} - \underline{a}^t \underline{x}_i}{(\underline{c}^R)^t \underline{x}_i} \right), & (\underline{c}^R)^t \underline{x}_i \geq (\underline{y} - \underline{a}^t \underline{x}_i) \geq 0, \;\; (\underline{c}^R)^t \underline{x}_i > 0 \\
0 \qquad\qquad , & dy
\end{cases}
$$

8.2.2. Bulanık en küçük mutlak sapma regresyonu

Eşitlik (8.15)'teki genel bulanık lineer regresyonunu ele alalım. Y_i değerleri ile Y_i^* çıktı değerleri arasındaki farkı gösteren ε_i'ler bulanık sayılardır.

$$Y_i = Y_i^* \{+\} \varepsilon_i$$

ve

$$\varepsilon_i = Y_i \{-\} Y_i^* \tag{8.17}$$

dır. Burada ε_i 'ler gözlemlerin bulanık hatalarıdır ve Y_i değerleri ile Y_i* çıktıları arasındaki bulanık farkı gösterir. Bu problemde en küçük mutlak sapma çözümünü bulmak için, bu bulanık sapma en küçük hale getirilmelidir (Chang ve Lee 1994).

$$Min \sum_{i=1}^{n} |\varepsilon_i| = \sum_{i=1}^{n} |Y_i \{-\} Y_i^*| \tag{8.18}$$

Eşitlik (8.18)'de bulanık fark yaklaşımı kullanılırsa $\{-\}$ operatörü $(-)$ ile veya bulanık çıkarma işlemi ile yer değiştirebilir. Burada bulanık fark yaklaşımlarından "overall existence ranking index (OERI)" kullanılacaktır.

OERI, normal veya normal olmayan bulanık sayılar veya kümelerin tercih edilen sıralamaları için karşılaştırılabilir bir ölçü sağlayan bulanık sıralama fonksiyonudur. Bir A_j bulanık kümesinin indeksi şöyle tanımlanır.

$$OM(A_j) = \int_0^{w^*_{hgt}} g_j\left(\{\mu_{A_j}^{-1}(w)\}\right) dw \tag{8.19}$$

$$w^*_{hgt} = \min_j [hgt(A_j)]$$

$$\{\mu_{A_j}^{-1}(w)\} = \{a_j \mid \mu_{A_j}^{-1}(a_j) = w\}$$

Burada $a_j^n(w) = \sup\{a_j \mid \mu_{A_j}(a_j) = w\}$ ve $a_j^1(w) = \inf\{a_j \mid \mu_{A_j}(a_j) = w\}$ alınırsa $g_j(\{\mu_{A_j}^{-1}(w)\})$ şöyle tanımlanır.

$$g_j(\{\mu_{A_j}^{-1}(w)\}) = \omega(w) \times [x_1(w) a_j^1 + x_2(w) a_j^n] \tag{8.20}$$

Burada w üyelik derecesine uygun mevcut dereceyi göstermektedir. $\omega(w)$, w için karar verici belirlemede kişisel bir ağırlık fonksiyonudur ve $x_1(w)$ ve $x_2(w)$ w seviyesinde sırasıyla $a_j^1(w)$ ve $a_j^n(w)$ için kişisel ağırlık fonksiyonlarıdır. Burada aşağıdaki basit ağırlıklandırma kullanılmıştır.

$$\omega(w) = 1, \quad x_1(w) = x_2(w) = 1/2, \quad \forall w \in (0,1] \tag{8.21}$$

Özellikle üçgensel ve karesel LR-tip bulanık sayılar için eşitlik (8.19)'daki ağırlıklandırma şu hale gelir.

(i) Üçgensel LR-tip bulanık sayı için:

$$OM(A_j) = \frac{1}{2}\int_0^1 \left[(a_j - c_j^L(1-w)) + (a_j + c_j^R(1-w))\right]dw \\ = (4a_j - c_j^L + c_j^R)/4 \tag{8.22}$$

(ii) Dörtgensel LR-tip bulanık sayı için:

$$OM(A_j) = \frac{1}{2}\int_0^1 \left[(a_j - c_j^L\sqrt{1-w}) + (a_j + c_j^R\sqrt{1-w})\right]dw \\ = (3a_j - c_j^L + c_j^R)/3 \tag{8.23}$$

Bulanık sapmaların mutlak değerleri toplamını minimize etmek için bir doğrusal programlama modeli düşünülebilir.

$$Min\, J_1^T = \frac{1}{4}\sum_{i=1}^{n}\left|4y_i^m - y_i^L + y_i^R - 4a^t x_i + (c^L)^t x_i - (c^R)^t x_i\right| \tag{8.24}$$

Negatif olmayan d_i^{T1} ve d_i^{T2} değişkenlerinin eklenmesi ile (8.24)'teki optimizasyon problemi şöyle yazılır.

$$Min\left(1/4\right)\sum_{i=1}^{n}(d_i^{T1} + d_i^{T2}), \qquad c^L, c^R, x_i, d_i^{T1}, d_i^{T2} \geq 0$$

Burada,

$$d_i^{T1} - d_i^{T2} = 4y_i^m - y_i^L + y_i^R - 4a^t x_i + (c^L)^t x_i - (c^R)^t x_i, \quad i = 1,...,n \tag{8.25}$$

dır. Benzer şekilde J_1^Q için d_i^{Q1} ve d_i^{Q2} değişkenleri eklenebilir ve J_2 için de d_i^{L1}, d_i^{L2}, d_i^{R1} ve d_i^{R2} (8.25) deki gibi tanımlanabilir.

$$y_i^L - (c^L)^t x_i = d_i^{L1} - d_i^{L2} \tag{8.26}$$

$$y_i^R - (c^R)^t x_i = d_i^{R1} - d_i^{R2} \tag{8.27}$$

Böylece en küçük mutlak sapmalar için model,

Üçgensel Bulanık Sayılar için (TFN):

$$Min\, J_1^T + J_2 = \sum_{i=1}^{n} \left[\left(1/4\right) d_i^{T1} + \left(1/4\right) d_i^{T2} + d_i^{L1} + d_i^{L2} + d_i^{R1} + d_i^{R2} \right]$$

$$\textit{Kısıtlar:} \quad y_i^m - (1-\alpha) y_i^L \geq a^t x_i - (c^L)^t x_i \tag{8.28}$$

$$y_i^m + (1-\alpha) y_i^R \leq a^t x_i + (c^R)^t x_i\,,\ i = 1,...,n\,, \alpha \in (0,1] \tag{8.29}$$

$$x_i,\ d_i^{T1},\ d_i^{T2},\ d_i^{L1},\ d_i^{L2},\ d_i^{R1},\ d_i^{R2} \geq 0\,, \quad i = 1,...,n \text{ için } c^L, c^R \geq 0$$

Dörtgensel Bulanık Sayılar için (QFN):

$$Min\, J_1^T + J_2 = \sum_{i=1}^{n} \left[\left(1/3\right) d_i^{Q1} + \left(1/3\right) d_i^{Q2} + d_i^{L1} + d_i^{L2} + d_i^{R1} + d_i^{R2} \right]$$

$$d_i^{Q1} - d_i^{Q2} = 3y_i^m - y_i^L + y_i^R - 3a^t x_i + (c^L)^t x_i - (c^R)^t x_i\,, \quad i = 1,...,n$$

$$\textit{Kısıtlar:} \quad y_i^m - (1-\alpha)^{1/2} y_i^L \geq a^t x_i - (c^L)^t x_i \tag{8.30}$$

$$y_i^m + (1-\alpha)^{1/2} y_i^R \leq a^t x_i + (c^R)^t x_i\,,\ i = 1,...,n\,,\ \alpha \in (0,1] \tag{8.31}$$

$$x_i,\ d_i^{Q1},\ d_i^{Q2},\ d_i^{L1},\ d_i^{L2},\ d_i^{R1},\ d_i^{R2} \geq 0\,, \quad i = 1,...,n \text{ için } c^L, c^R \geq 0$$

biçiminde verilebilir (Chang ve Lee 1994).

9. UYGULAMALAR

9.1. Uygulama 1. YSA ile lineer olmayan bulanık regresyon uygulama verisi olarak bir değişkenli bir regresyon modeli göz önüne alınmıştır. Bağımsız değişken x_i değerleri $[0.0,10]$ aralığından düzgün dağılıma sahip olarak üretilmiştir. Bağımlı değişken,

$$y_i = 2.1 + \exp(0.2x_i) + \varepsilon_i$$

modeline uygun olarak elde edilmiştir. Burada $\varepsilon_i \sim N(0,0.01)$ ve $i = 1,...,25$ dir (Çizelge 9.1).

Veride 19 gözlem modeli oluşturmak için eğitim seti olarak düşünülmüş geri kalan 6 gözlem test verisi olarak alınmıştır. YSA uygulaması için .net C# ortamında bir Windows uygulaması yazılmıştır.

Çizelge 9.1.

i	y_i	x_i	ε_i
1	4.301	3.952	0.002
2	6.003	6.814	0.012
3	5.192	5.672	-0.006
4	6.000	6.815	0.009
5	9.212	9.826	0.018
6	6.073	6.901	0.013
7	9.286	9.891	0.001
8	3.265	0.785	-0.005
9	4.749	4.822	0.033
10	4.166	3.661	-0.009
11	3.580	1.921	0.014
12	6.839	7.827	-0.023
13	3.378	1.230	0.000
14	8.633	9.417	-0.005
15	5.060	5.446	-0.002
16	4.186	3.654	0.013
17	4.851	5.044	0.017
18	5.058	5.419	0.012
19	6.098	6.940	0.008
20	5.554	6.212	0.004
21	3.145	0.225	-0.001
22	5.815	6.572	0.007
23	3.365	1.235	-0.014
24	5.599	6.303	-0.015
25	4.173	3.685	-0.013

Lineer olmayan bulanık regresyon modeli için anahtar nokta y_i^w yayılmasının belirlenmesidir. y için ortalama ve standart sapma değerleri $\bar{y} = 5.34$ ve $s_y = 1.73$ olarak bulunmuştur. Bulanık lineer olmayan regresyon modelini oluşturmak için ilk olarak $\phi = 0.65\bar{y} = 3.49$ ve $\psi = 1$ olarak alınmıştır. Elde edilen y_i lere ait yayılmayı gösteren birinci ve üçüncü çeyreklik değerleri $\omega_1 = 4.17$ ve $\omega_2 = 6.04$ dir. Böylece beklenen yayılma değerleri y_i^w,

$$y_i^w = \frac{y_i}{5.34} + f(y_i)$$

olarak elde edilmektedir. Burada,

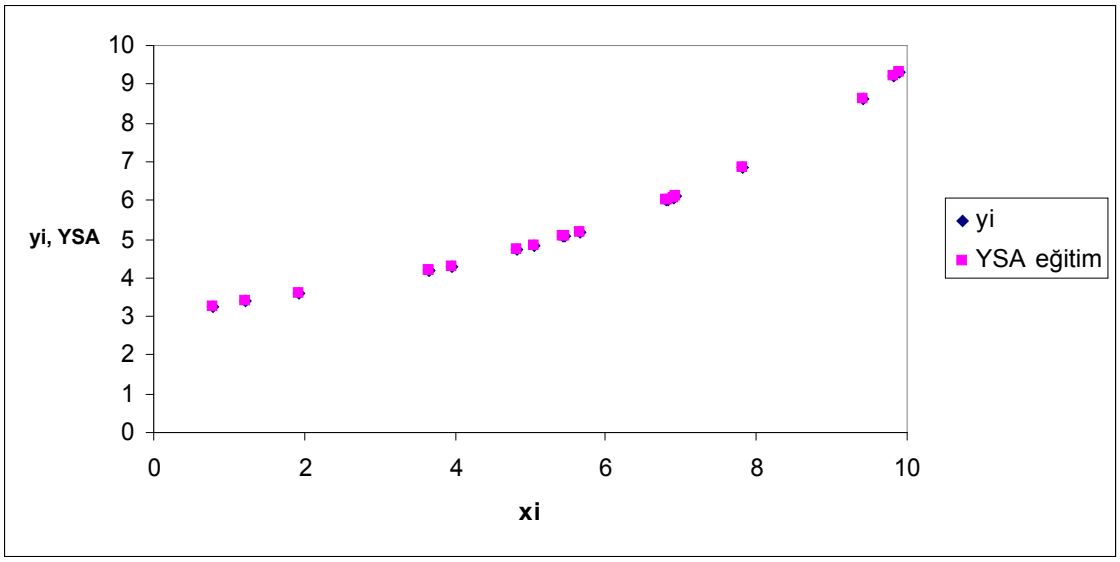
$$f(y_i) = \begin{cases} \dfrac{y_i}{1+e^{-0.68(y_i-4.17)}} & y_i \le 4.17 \\ y_i & 4.17 \le y_i \le 6.04 \\ \dfrac{y_i}{1+e^{-0.40(y_i-6.04)}} & y_i \ge 6.04 \end{cases}$$

olarak verilmektedir.

Uygulamada girdi birimleri x_i'lere karşılık çıktı birimleri $y+(x_i)$ ve $y-(x_i)$ olan iki tane sinir ağı ile çalışılmıştır. Sinir ağlarında sigmoid aktivasyon fonksiyonu kullanıldığı için çıktı verisi [0,1] aralığına getirilmiştir. Çalışma sürecinden sonra eski büyüklüklerine geri dönülmüştür. x_i'lere dönüşüm uygulanmamıştır.

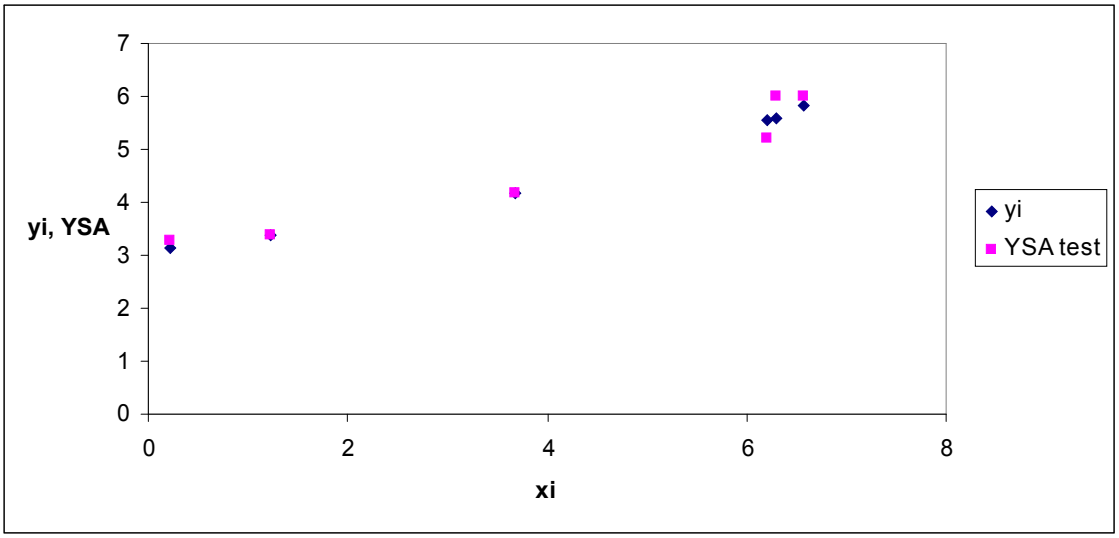
Sinir ağının oluşturulmasında başlangıç girdi ve çıktı ağırlıkları sıfır, öğrenme oranı $\eta = 0.70$ olarak alınmıştır. 10000 iterasyon sonucu hata kriteri $E = 0.011$ bulunmuştur. Değer oldukça küçüktür.

Çalışma sonucunda eğitim setindeki y_i'lere ait tahmin değerleri Şekil 9.1'de verilmiştir. YSA algoritması yapılan tahminler y_i'ler ile çakışmaktadır.



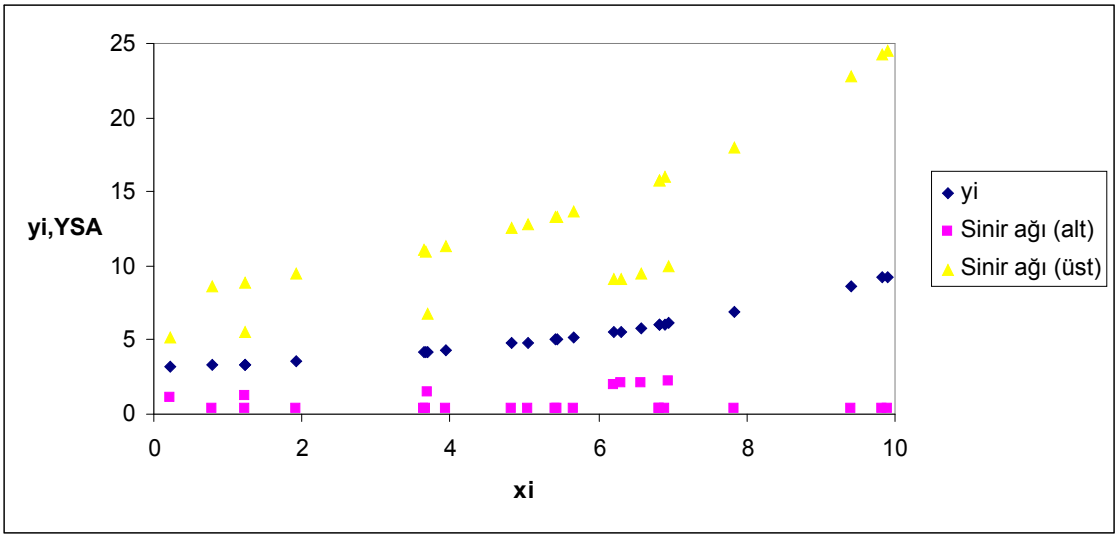
Şekil 9.1.

Test verisi ile yapılan tahmin değerleri ile y_i'ler ise Şekil 9.2'de görülmektedir.



Şekil 9.2.

x_i'lere karşılık çıktı birimleri $y+(x_i)$ ve $y-(x_i)$ alınarak yapılan YSA çalışması sonucunda alt ve üst sinir ağlarının y_i değerlerini kapsadığı Şekil 9.3'ten görülebilir. Çizelge 9.3'te bu değerler verilmiştir.



Şekil 9.3.

Modelin yeterliliğini görmek için $\hat{Y}(x_i)$ 'nin 2 farklı α -seviye kümesi kullanılmıştır.

Çizelge 9.2.

	α	AAE^-	AAE^c	AAE^+	AA^w	**Uyum Oranı**
Eğitim verisi	0.10	2.319	0.000	0.660	13.463	1
	0.05	2.605	0.000	0.286	14.211	1
Test verisi	0.10	6.946	0.182	7.065	7.784	1
	0.05	7.335	0.182	7.455	15.568	1

Çizelge 9.3.

i	y_i	$\hat{y}_i$	**Alt Sınır**	**Üst Sınır**	e_i
1	4.301	4.301	0.426	11.337	0.000
2	6.003	6.003	0.426	15.824	0.000
3	5.192	5.192	0.426	13.687	0.000
4	6.000	6.000	0.426	15.816	0.000
5	9.212	9.212	0.426	24.282	0.000
6	6.073	6.073	0.426	16.007	0.000
7	9.286	9.286	0.426	24.479	0.000
8	3.265	3.265	0.426	8.606	0.000
9	4.749	4.749	0.426	12.517	0.000
10	4.166	4.166	0.426	10.983	0.000
11	3.580	3.580	0.426	9.438	0.000
12	6.839	6.839	0.426	18.027	0.000
13	3.378	3.378	0.426	8.905	0.000
14	8.633	8.633	0.426	22.758	0.000
15	5.060	5.060	0.426	13.338	0.000
16	4.186	4.186	0.426	11.033	0.000
17	4.851	4.851	0.426	12.788	0.000
18	5.058	5.058	0.426	13.332	0.000
19	6.098	6.098	2.220	9.976	0.000
20	5.554	5.192	2.022	9.087	0.362
21	3.145	3.265	1.157	5.145	-0.120
22	5.815	6.003	2.117	9.513	-0.188
23	3.365	3.378	1.233	5.505	-0.013
24	5.599	6.003	2.038	9.159	-0.404
25	4.173	4.166	1.520	6.826	0.006

Modele ait hata kareler toplamı değeri 0.345 olarak bulunmuştur.

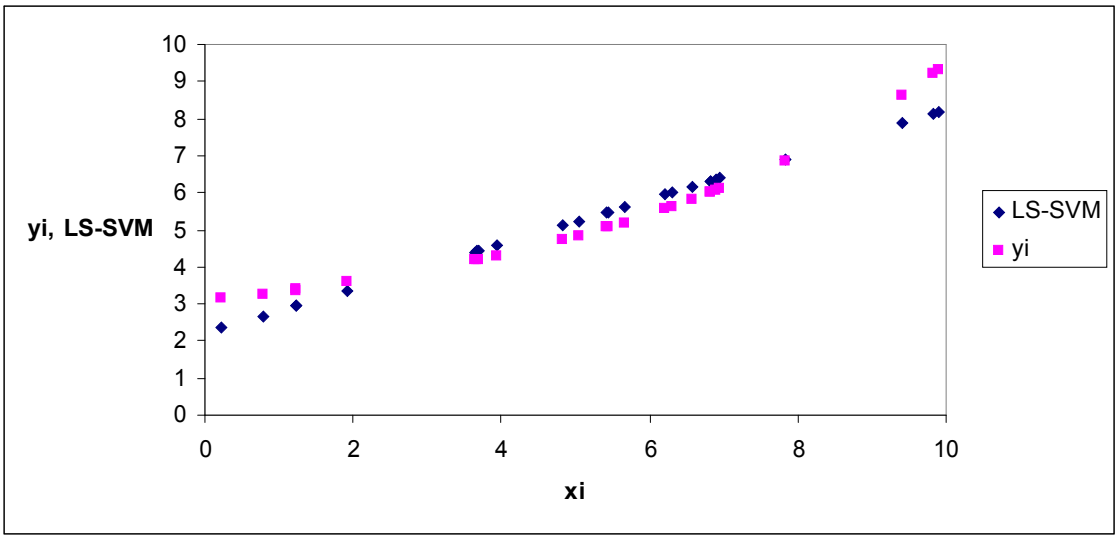
9.2. Uygulama 2. LS-SVM'ye dayalı lineer olmayan bulanık regresyon algoritmasının performansını görmek amacıyla Uygulama 1'deki veri seti göz önüne alınmıştır (Çizelge 9.4). x 'lere ait yayılma değerleri de [0,1] aralığından düzgün dağılıma sahip olarak üretilmiştir.

Çizelge 9.4.

i	m_{Y_i}	α_{Y_i}	β_{Y_i}	m_{X_i}	α_{X_i}	β_{X_i}	ε_i
1	4.301	3.723	4.879	3.952	3.480	4.424	0.002
2	6.003	5.394	6.612	6.814	5.900	7.728	0.012
3	5.192	4.867	5.518	5.672	5.254	6.091	-0.006
4	6.000	5.259	6.741	6.815	6.419	7.211	0.009
5	9.212	8.244	10.179	9.826	9.207	10.444	0.018
6	6.073	5.228	6.917	6.901	6.081	7.722	0.013
7	9.286	8.643	9.930	9.891	9.119	10.662	0.001
8	3.265	2.649	3.880	0.785	0.027	1.543	-0.005
9	4.749	3.810	5.687	4.822	3.826	5.818	0.033
10	4.166	3.655	4.678	3.661	2.815	4.507	-0.009
11	3.580	2.790	4.371	1.921	1.495	2.347	0.014
12	6.839	5.873	7.805	7.827	7.188	8.465	-0.023
13	3.378	2.530	4.227	1.230	0.812	1.647	0.000
14	8.633	8.194	9.073	9.417	8.945	9.889	-0.005
15	5.060	4.549	5.571	5.446	4.888	6.003	-0.002
16	4.186	3.816	4.555	3.654	2.779	4.529	0.013
17	4.851	4.539	5.163	5.044	4.497	5.592	0.017
18	5.058	4.735	5.380	5.419	5.077	5.761	0.012
19	6.098	5.681	6.515	6.940	5.987	7.893	0.008
20	5.554	4.575	6.533	6.212	5.382	7.042	0.004
21	3.145	2.319	3.970	0.225	0.085	0.366	-0.001
22	5.815	4.996	6.633	6.572	5.580	7.563	0.007
23	3.365	2.441	4.289	1.235	0.448	2.022	-0.014

24	5.599	5.045	6.152	6.303	5.883	6.724	-0.015
25	4.173	3.817	4.528	3.685	3.191	4.180	-0.013

LS-SVM uygulaması için Selçuk STAT paket programından yararlanılmıştır. Şekil 9.4'ten görüleceği gibi hazırlanan model lineer olmayan bulanık girdi ve çıktı verisi için yaklaşık sonuç vermektedir.

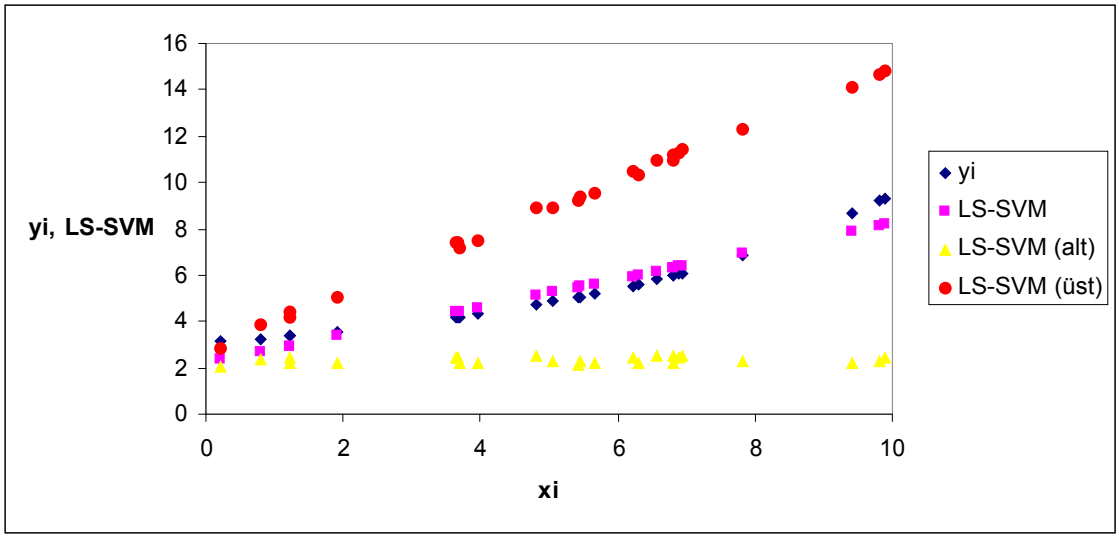


Şekil 9.4.

Çizelge 9.5'te modele ilişkin tahmin değerleri ile modelden elde edilen hatalar yer almaktadır. Tahmin değerlerine ait yayılmalar Şekil 9.5'te verilmiştir. Modelin hata kareler toplamı değeri 6.289 olarak bulunmuştur.

Çizelge 9.5.

i	m_{Y_i}	α_{Y_i}	β_{Y_i}	$m_{\hat{Y}_i}$	$\alpha_{\hat{Y}_i}$	$\beta_{\hat{Y}_i}$	e_i
1	4.301	3.723	4.879	4.585	2.361	2.930	-0.284
2	6.003	5.394	6.612	6.309	3.819	4.920	-0.306
3	5.192	4.867	5.518	5.621	3.430	3.934	-0.429
4	6.000	5.259	6.741	6.310	4.132	4.608	-0.309
5	9.212	8.244	10.179	8.123	5.811	6.556	1.089
6	6.073	5.228	6.917	6.362	3.928	4.916	-0.289
7	9.286	8.643	9.930	8.162	5.758	6.687	1.124
8	3.265	2.649	3.880	2.678	0.282	1.195	0.587
9	4.749	3.810	5.687	5.109	2.570	3.770	-0.361
10	4.166	3.655	4.678	4.410	1.961	2.980	-0.244
11	3.580	2.790	4.371	3.362	1.165	1.679	0.219
12	6.839	5.873	7.805	6.919	4.595	5.364	-0.080
13	3.378	2.530	4.227	2.946	0.754	1.257	0.433
14	8.633	8.194	9.073	7.877	5.653	6.221	0.757
15	5.060	4.549	5.571	5.485	3.209	3.881	-0.425
16	4.186	3.816	4.555	4.406	1.939	2.993	-0.220
17	4.851	4.539	5.163	5.243	2.974	3.633	-0.392
18	5.058	4.735	5.380	5.469	3.323	3.735	-0.411
19	6.098	5.681	6.515	6.385	3.871	5.019	-0.287
20	5.554	4.575	6.533	5.947	3.507	4.507	-0.392
21	3.145	2.319	3.970	2.341	0.316	0.486	0.804
22	5.815	4.996	6.633	6.163	3.626	4.821	-0.349
23	3.365	2.441	4.289	2.949	0.535	1.483	0.416
24	5.599	5.045	6.152	6.001	3.808	4.315	-0.403
25	4.173	3.817	4.528	4.425	2.187	2.783	-0.252


16
14
12
10
8
6
4
2
0
yi, LS-SVM
0
2
4
6
8
10
xi
yi
LS-SVM
LS-SVM (alt)
LS-SVM (üst)

Şekil 9.5.

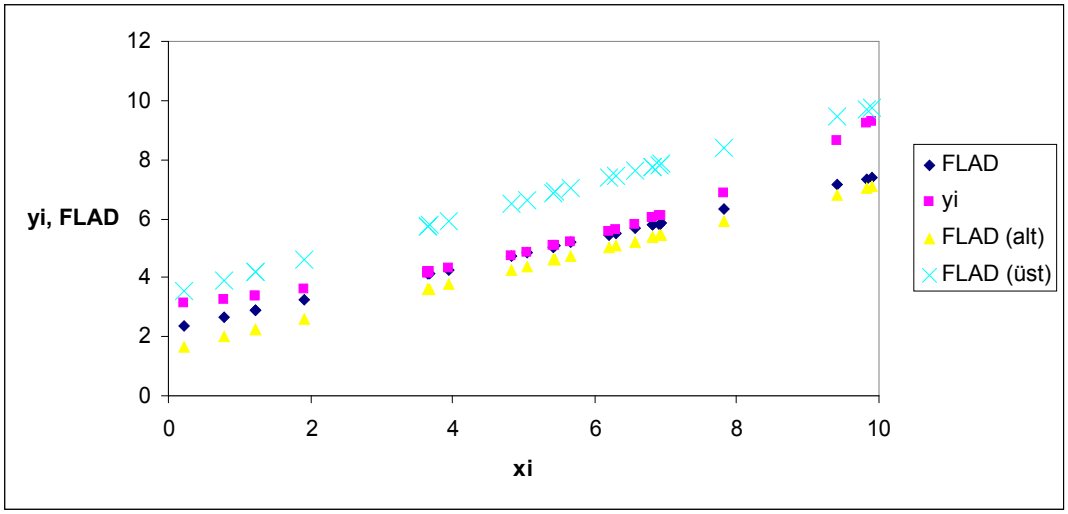
9.3. Uygulama 3: Bulanık en küçük mutlak sapma regresyonu uygulaması için önceki uygulamalardaki veri seti göz önüne alınmıştır (Çizelge 9.6).

Çizelge 9.6.

i	y_i^m	y_i^l	y_i^r	x_i
1	4.301	0.578	0.578	3.952
2	6.003	0.609	0.609	6.814
3	5.192	0.325	0.325	5.672
4	6.000	0.741	0.741	6.815
5	9.212	0.968	0.968	9.826
6	6.073	0.845	0.845	6.901
7	9.286	0.644	0.644	9.891
8	3.265	0.615	0.615	0.785
9	4.749	0.939	0.939	4.822
10	4.166	0.512	0.512	3.661
11	3.580	0.790	0.790	1.921
12	6.839	0.966	0.966	7.827
13	3.378	0.848	0.848	1.230
14	8.633	0.440	0.440	9.417
15	5.060	0.511	0.511	5.446
16	4.186	0.370	0.370	3.654
17	4.851	0.312	0.312	5.044
18	5.058	0.323	0.323	5.419
19	6.098	0.417	0.417	6.940
20	5.554	0.979	0.979	6.212
21	3.145	0.825	0.825	0.225
22	5.815	0.818	0.818	6.572
23	3.365	0.924	0.924	1.235
24	5.599	0.554	0.554	6.303
25	4.173	0.356	0.356	3.685

Veri üçgensel bulanık sayılardan oluştuğu için üçgensel bulanık sayılar için en küçük mutlak sapma regresyonu uygulanmıştır. Doğrusal programlama probleminin çözümünde WinQSB paket programından faydalanılmıştır.

$\alpha = 1$ alınarak tahminlerin merkez değerleri bulunmuş ve $\alpha = 0.5$ olarak alınıp tahmin değerleri için alt ve üst sınır belirlenmiştir. Şekil 9.6 en küçük mutlak sapma regresyonu yönteminin bulanık regresyon analizi için etkili bir yöntem olabileceğini göstermektedir.



Şekil 9.6.

Çizelge 9.7'de bulanık en küçük mutlak sapma regresyonuna ilişkin tahmin ve hata değerleri yer almaktadır. Modelin hata kareler toplamı 10.090 olarak hesaplanmıştır.

Çizelge 9.7.

$\hat{y}_i^m$	$\hat{y}_i^l$	$\hat{y}_i^r$	e_i
4.285	3.763	5.939	0.016
5.773	5.366	7.771	0.230
5.179	4.726	7.040	0.013
5.774	5.366	7.772	0.226
7.340	7.053	9.699	1.872
5.819	5.415	7.827	0.254
7.373	7.089	9.740	1.913
2.638	1.990	3.912	0.627
4.737	4.250	6.496	0.011
4.134	3.600	5.753	0.033
3.229	2.626	4.639	0.352
6.300	5.933	8.419	0.539
2.870	2.239	4.197	0.509
7.127	6.824	9.437	1.507
5.062	4.600	6.895	-0.002
4.130	3.596	5.749	0.056
4.853	4.375	6.638	-0.002
5.048	4.585	6.878	0.010
5.839	5.436	7.852	0.259
5.460	5.029	7.386	0.094
2.347	1.676	3.554	0.798
5.647	5.230	7.616	0.167
2.872	2.242	4.200	0.493
5.508	5.080	7.444	0.091
4.146	3.614	5.768	0.026

10. SONUÇ

Bulanık lineer olmayan regresyon modelinde çıktı değerleri için bir aralık meydana getirmek amaçlanmıştır. Bu yönde yapılan uygulamalar sonucunda hazırlanan üç yönteme ilişkin hata kareler ortalaması ve R^2 değerleri Çizelge 10.1'den görülmektedir.

Çizelge 10.1.

Model	**HKO**	R^2
YSA	0.014	0.99
LS-SVM	0.252	0.90
FLAD	0.404	0.73

Sonuçlara göre en küçük HKO değeri YSA yöntemine aittir. Yine en yüksek belirleyicilik katsayısını da YSA vermektedir. LS-SVM ve FLAD yöntemlerinden ilkinin HKO değeri daha düşük olduğundan yöntemleri aralarında bu şekilde bir sıralamaya tabi tutabiliriz. Verinin merkez değerlerinin EKK yöntem ile yapılan regresyon analizi sonucunda HKO değeri 0.004 ve belirleyicilik katsayısı da 0.99 olarak bulunmuştur. Sonuçlar Selçuk STAT paket programı ile elde edilmiştir. Klasik yöntemle elde edilen sonuçlar çok daha iyi olmasına rağmen burada modelin bilinmesi gerekliliği bir dezavantaj olarak ortaya çıkmaktadır. Ayrıca bu yöntem bulanık sayıları kapsamamakta yalnızca kesin sayılar için işlem yapmaktadır.

Çalışmada incelenen üç yöntemin yukarıda verilen kriterler dışında kendine göre bazı avantajları ve dezavantajları bulunmaktadır. Bunlar daha ayrıntılı bir şekilde Çizelge 10.2'de verilmiştir.

Çizelge 10.2.

	Avantajları	Dezavantajları
YSA	- Öğrenme yeteneği sayesinde bilinen örnekleri kullanarak daha önce karşılaşılmamış durumlarda genelleme yapabilmektedir. Eksik bilgileri de işleyerek sonuca varabilir. - Sınırsız sayıda değişken ve parametre ile kolaylıkla çalışılabilmektedir. - İşlemci elemanlar paralel bağlıdır ve ağın sahip olduğu bilgi tüm ağa dağıtılmış durumdadır. Bundan dolayı hatayı tolere etme yeteneği fazladır. - Değişken uzayı karmaşık ve verinin dağılımı bilinen istatistiksel dağılımlardan farklıysa, istatistiksel sınıflama yöntemlerinden daha güvenilir sonuçlar verir.	- Sinir ağlarının eğitilmesine ve test edilebilmesine yetecek genişlikte veri setine ihtiyaç duyulmaktadır. - Uygun bir ağ yapısı belirleme genellikle deneme yanılma yöntemiyle gerçekleştirilir. - Gizli tabakada işlemci sayısı az olduğunda bazı karmaşık ilişkiler belirlenememektedir. Ancak gizli tabakadaki işlemci sayısının çok artması da hesaplama yükünü artırmaktadır. - Parametre tahmini vermez ve model oluşturmaz. - Hatalı öğrenme meyili olabilir.

LS-SVM	- Girdi ve çıktısı bulanık lineer olmayan regresyon için modelden bağımsız bir yöntemdir. Bu nedenle aralarında lineer olmayan ilişkinin olduğu değişkenlere kolaylıkla uygulanabilir. - Matris işlemleri ile çözüm yapıldığından hesaplama kolaylığı sağlamaktadır. - Çok değişkenli veri için de uygulanması basittir.	- Kernel parametresi p ve kontrol parametresi C 'nin tespiti için önerilen bir yöntem yoktur. Deneme yoluyla elde edilmektedir.
FLAD	- Regresyon modelinin hataları normal dağılıma uygun olmadığında veya veride aykırı gözlemler olması durumunda EKK yöntemine göre daha sağlam sonuçlar vermektedir. - Üçgensel ve dörtgensel bulanık sayılar için işlem kolaylığı vardır.	- Bulanık regresyon modelinde LAD uygulandığında yayılmalar için her zaman doğru sonuç vermemektedir. - Büyük veri setlerinde kısıt sayısı çok fazla olduğundan işlem zorluğu ortaya çıkmaktadır.

KAYNAKLAR

1. Apaydın, A., (2005), Optimizasyon, AÜFF Döner Sermaye İşletmesi Yayınları, Ankara.

2. Bard, Y., (1974), Nonlinear Parameter Estimation, Academic Press, New York.

3. Bates, D.M., Watts, D.G., (1988), Nonlinear Regression Analysis and Its Application, John Wiley&Sons, Inc., USA.

4. Borowiak, D.S., (1989), Models Discrimination for Nonlinear Regression Models, Marcel Dekker, Inc., New York.

5. Buckley, J.J., Eslami, Esfandiar, (2002), An introduction to fuzzy logic and fuzzy sets, Physica-Verl., Heidelberg, New York.

6. Chang, P., Lee, S., (1994), "Fuzzy Least Absolute Deviations Regression Based on The Ranking of Fuzzy Numbers", IEEE World Congress on Computational Intelligence., Proceedings of the Third IEEE Conference on, vol.2, pp. 1365 – 1369.

7. Çetin, M.C., Orsoy, A., (2001), "Doğrusal Regresyonda Sağlam Tahmin Ediciler ve Bir Uygulama", Anadolu Üniversitesi Bilim ve Teknoloji Dergisi, vol. 2, no. 2, pp. 265-270.

8. Dasgupta, M., Mishra S., (2004), "Least absolute deviation estimation of linear economic models: A literature review", Munich Personal RePEc Archive.

9. Dombi, J., (1990), Membership function as an evaluation. Fuzzy Sets and Systems 35: 1-21.

10. Dubois, D., Prade, H., (1980), Fuzzy Sets and Systems: Theory and Applications, New York: Academic.

11. Dudewicz, E.J., Mishra, S.N., (1988), Modern Mathematical Statistics, John wiley&Sons, Canada.

12. Genç, A., (1997), "Çok Değişkenli Lineer Olmayan Modeller: Parametre Tahmini Ve Hipotez Testi", Ankara Üniversitesi Fen Bilimleri Enstitüsü Doktora Tezi, Ankara.

13. Hong, D.H., Hwang, C., (2006), "Fuzzy Nonlinear Regression Model Based on LS-SVM in Feature Space", FSKD 2006, LNAI 4223, pp. 208-216.

14. İnal, C., Günay, S., (1999), Olasılık ve Matematiksel İstatistik, Hacettepe Üniversitesi Fen Fakültesi Basımevi, Ankara.

15. Ishibuchi, H., Tanaka, H., (1992), "Fuzzy Regression Analysis Using Neural Networks: Fuzzy Sets and Systems", Vol.50, No.3, pp.257-266.

16. Kim, K., Kim, D., Choi, S., (2005), "Least Absolute Deviation Estimator in Fuzzy Regression", J. Appl. Math. &Computing, vol. 18, pp. 649-656.

17. Myers, R.H., (1986), Classical and Modern Regression With Applications, PWS-KENT Publishing Company, USA.

18. Mendenhall, W., Sincich, T., (1996), A Second Course in Statistics: Regression Analysis, Prentice Hall, Inc., USA.

19. Montgomery, D.C., Peck, E.A., Vining, G.G., (2001), Introduction to Linear Regression Analysis, John Wiley&Sons, Canada.

20. Oral, G., (1989), Doğrusal Olmayan Programlama, Akademi Matbaası, Ankara.

21. Pedrycz, W., Gomido, F., (1998), An introduction to fuzzy sets analysis and design, The MIT Press, USA.

22. Rawlings, J.O., Pantula, S.G., Dickey, D.A., (1998), Applied Regression Analysis, Sprinter-Verlag, New York.

23. Seber, G.A.F., Wild, C.J., (1989), Nonlinear Regression, John Wiley&Sons, Canada.

24. Sun, Z., Sun, Y., (2003), "Fuzzy Support Vector Machine for Regression Estimation", IEEE International Conference on, Vol. 4, pp. 3336 – 3341.

25. Suykens, J., Gestel, T., Brabanter, J., DeMoor, B., Vandewalle, J., (2002), Least Squares Support Vector Machines, World Scientific Publishing Co. Pte. Ltd., London.

26. Şanlı, K., Apaydın, A., (2005), "Verilerin Bulanık ve Veri Kümesinde Aykırı Değer Olması Durumunda Bulanık Robust Regresyon Çözümlemesi", Anadolu Üniversitesi Bilim ve Teknoloji Dergisi, vol. 6, no. 1, pp. 37-44.

27. Tanaka, H., Uejima, S., Asai, K., (1982). "Linear Regression Analysis with Fuzzy Model", IEEE Transactions on Systems, Man and Cybernetics, Vol. SMC-12, No.6.

28. Terano T., Asai K., Sugeno M., (1981), Fuzzy Systems Theory and Its Applications, Academic Press, Tokyo.

29. Xu, Z., Khosgoftaar, T.M., (2000), "Prediction of Software Faults Using Fuzzy Nonlinear Regression Modeling", Fifth IEEE International Symposim on. HASE 2000 Volume , Issue , pp. 281 – 290.

30. Yapıcı, N., (2000), "Bulanık Doğrusal Programlamaya Sinir Ağları Yaklaşımı", Ankara Üniversitesi Fen Bilimleri Enstitüsü Yayımlanmamış Yüksek Lisans Tezi, Ankara.

31. Yapıcı Pehlivan, N., Apaydın, A., (2005), "Bulanık k-En Yakın Komşuluk Tahmin Edicisi ve Bulanık Radyal Tabanlı Fonksiyon Ağları", Selçuk Üniversitesi Fen Edebiyat Fakültesi Fen Dergisi, sayı. 26, p. 19-32.

32. Yazıcı, A., Öğüş E., Ankaralı, S., Canan, S., Ankaralı, H., Akkuş, Z., (2006), "Yapay Sinir Ağlarına Genel Bakış", Türkiye Klinikleri J. Med Sci 2007, sayı. 27, p. 65-71.

33. Zadeh, L. A., 1965, Fuzzy Sets, Information and Control, 8, 338-353.

34. Zhang, D., Deng, L., Cai, K., So, A., (2005), "Fuzzy Nonlinear Regression With Fuzzified Radial Basis Function Network", IEEE Transactions on Fuzzy Systems, vol. 13, no. 6, p. 742-760.