KLASİK DOĞRUSAL REGRESYON ANALİZİNE ALTERNATİF BİR YÖNTEM OLARAK BULANIK DOĞRUSAL REGRESYON ANALİZİ VE BİR UYGULAMA

Leyla İŞBİLEN YÜCEL*

Özet

Klasik regresyon analizi kesin bilgilerle çalışan sayısal bilimlerin ihtiyaçlarına somut cevaplar verebilmektedir. Ancak sosyal bilimler söz konusu olduğunda tutarlı ve doğru sonuçlara ulaşmak sanıldığı kadar kolay değildir. Çünkü gerçek yaşamda her şey bir derecelendirme sorunudur. Bulanık mantık, insan zekası gibi etraflı düşünerek, yaklaşık karar verebilmeyi sağlamaktadır. Bulanık regresyon yöntemi, bulanık mantığın klasik regresyon yöntemlerine uygulanması sonucunda bulunmuştur. Bulanık regresyon, klasik regresyon analizi için gerekli olan varsayımların sağlanmasına ihtiyaç duymaz. Artık veriler normal dağılmak zorunda değildir, durağanlık testleri ve büyük örneklemler (sonsuz denemeler) olmasa da olur. Fakat tüm bu anlatılanlar, bulanık regresyonun sadece durum tespiti yapabilmesine olanak tanımaktadır. Klasik regresyonda olduğu gibi dönem dışı tahmin yapılamaması, bulanık regresyon analizinin en büyük eksikliğidir. Bunun yanında bulanık regresyon analizi klasik regresyona göre olayların çok daha ayrıntılı bir resmini sunabilmektedir. Bunu da yeterli veriye ulaşılamaması veya bağımlı değişken ve bağımsız değişkenler arasında net olarak ortaya konamayan bir takım ilişkiler gibi imkansızlıklar nedeniyle klasik modele dahil edilemeyen açıklayıcı değişkenleri modele dahil ederek başarmaktadır.

Anahtar Kelimeler: *Bulanık doğrusal regresyon, bulanık parametreler, h terimi, üyelik fonksiyonu.*

* İstanbul Üniversitesi, Ekonometri Bölümü.

Abstract

Classical regression analysis is able to reply to requirements of numerical sciences that study with definite information. Hardly, when social sciences are in question, it is not easy to estimate coherent and consistent estimations as it is supposed. Because everything in real life is a gradation problem. Fuzzy logic supplies approximate reasoning by thinking as exhaustive as human intelligence. Fuzzy regression method was invented by applying fuzzy logic techniqurs to the classical regression method. Assumptions of classical regression need not be proved for fuzzy regression. Any longer, fuzzy regression does not require normal distributed data, stability tests and big samples (infinite trials). Yet, all these properties of fuzzy regression that are mentioned, only enables case stabilization. It can not estimate external terms like classical regression. But it exhibits more detailed photos of the events than classical method. Fuzzy regression overcomes this, by including some explanatory variables that classical regression can not include them to the classical model for some reasons such as insufficient data, intuitive but not certain relations between dependent variable and these explanatory variables.

Key Words: *Fuzzy linear regression, fuzzy parameters, h term, membership function.*

1. Giriş

Bulanık doğrusal regresyon analizi, bulanık (sınırları kesin tanımlanamayan) ortamda, bağımlı ve bağımsız değişkenler arasında üyelik fonksiyonlarını temel alarak, fonksiyonel ilişkiler kurar. Üyelik fonksiyonunun istatistik teorisindeki karşılığı, rastlantı değişkeninin olasılık fonksiyonudur. Üyelik fonksiyonu [0,1] aralığında değer alır ve μ (x) ile gösterilir. Örneğin $\mu_{\tilde{A}}(x) = 0.7$ olduğunda, x elemanının $\tilde{A}$ bulanık kümesine ait olma derecesi 0.7'dir yani %70 olasılıkla $\tilde{A}$'nın elemanı iken, aynı anda %30 olasılıkla başka bir kümenin veya kümelerin elemanı olabilir demektir. Görüldüğü gibi bulanık mantıkta klasik küme teorisindeki gibi ikili mantık yoktur, bir elaman bir kümeye kesinlikle ait olmak veya olmamak zorunda değildir. Bulanık küme teorisinde kısmi üyelik kavramı söz konusudur. Kısmi üyelik ise, bir elemanın tamamıyla bir kümeye ait olmak veya olmamak gibi bir kısıtlamasının olmaması, birden fazla kümeye aynı anda ait (elemanı) olabilmesi durumudur. Bulanık mantık sadece {0,1} ile ifade edilen ikili mantığın daha genişletilmiş halidir ve bunun sonucu olarak [0,1] kapalı aralığında sonsuz değer alabilmektedir. Diğer bir deyişle, her bir elemanın kendine ait bir üyelik derecesi(değeri) söz konusudur.

Bulanık doğrusal regresyon yöntemi, klasik doğrusal regresyon yönteminin esnetilmesiyle bulunmuştur. Klasik yaklaşımı inkar etmez, sadece klasik yollardan çözümü olmayan problemleri daha gevşetilmiş olarak ele alır.

Klasik model rasgeleliği temel alır. Rasgelelik ise bir olayın meydana gelişindeki belirsizliğin ölçümüdür. Bir olayın olması veya olmaması rasgeleliğin konusudur.[1] Deneme sayısı arttıkça rasgelelik önemini yitirir ve belirli bir değere yakınsar. Bir olayın gerçekleşebilme derecesi ise bulanıklığın konusudur. Bulanıklık; kesin sınırları olmayan, açık seçik bir biçimde tanımlanamayan, durumdan duruma farklı anlamlar alabilen demektir. Gerçek yaşamdaki problemler genellikle bulanıktırlar. Fakat bu bulanıklık (belirsizlik), iki değerli klasik mantığın sınırları arasında kalmıştır. Bilim ve tekniğin gelişmesiyle yeni yaklaşımlar üretilmiş ve daha kapsamlı, gerçekleri olduğu gibi ele alabilen çözüm yolları araştırılmıştır. Bilginin artması ile kesinliğin azalması sonucunda klasik teori genişletilerek bulanık teoriye ulaşılmıştır.

Bulanık regresyon yöntemini dünyaya tanıtan Hideo Tanaka olmuştur. Yöntemin temelinde Zadeh'in bulanık küme teorisi ve genişletme prensibi vardır. Tanaka'nın modeli ilk kurulan model olması sebebiyle, kabul görmesine rağmen bazı eleştirilere de maruz kalmıştır. Fakat bu eleştiriler yeni yaklaşımlar doğurarak yöntemin daha da gelişmesini sağlamıştır.[2]

2. Bulanık Doğrusal Regresyon Analizinin Ortaya Çıkış Nedenleri

Klasik doğrusal regresyonun pek çok uygulaması olmasına rağmen bazı durumlarda sorunlara yol açmaktadır. Bu sorunlar şöyle sıralanır:

- Küçük örnek sorunu (yeterli veriye ulaşılamaması),
- Dağılımlara ilişkin varsayımları yerine getirme zorluğu,
- Değişkenler arasındaki ilişkileri doğrusallaştırmak adına ham verilerin dönüştürülmesi,
- Hata terimlerinin normal dağılmaması,

[1] Baykal, Beyan, a.g.e., s. 310.

[2] (Çevrimiçi) http://www.libertasmedia.com/alan/dusveaks/nispidus.html, 11 Aralık 2004.

- Değişkenlerin sayısal kodlama yapılamayacak derecede bulanık olmaları,
- Değişkenler arasındaki ilişkilerin kesin bir fonksiyonla gösterilememesi.

Bulanık doğrusal regresyonda veri yapısına ilişkin varsayımlar göz ardı edilir. Değişkenlere ilişkin gözlem değerlerinin (verilerin) hangi dağılıma uyduğu önemsizdir. Klasik doğrusal regresyonda hata, genellikle ölçmedeki dikkatsizliklerden, modelin doğru belirlenmemiş olmasından ya da ulaşılamayan veri kümesinden kaynaklanır. Bulanık doğrusal regresyonda ise hata, sistem parametrelerinin (model katsayılarının) kesin olarak belirli olmamasından yani bulanıklığından kaynaklanır. Bu nedenle bulanık regresyonda hata miktarı, modeldeki bulanık parametrelerin yayılımlarının toplamına eşittir.

Küçük örnek ortalaması durumunda, klasik doğrusal regresyon tahminleri etkinliğini yitirirler. Bulanık doğrusal regresyonda ise sadece bulanıklık derecesi bir miktar artar, fakat tahminler gerçeği yansıtma özelliklerini yitirmezler.

Bulanık regresyon, klasik regresyona göre daha karmaşık hesaplar gerektirir. Sadece üyelik fonksiyonlarının belirlenmesi bile başlı başına zaman alan bir süreçtir. Çünkü üyelik fonksiyonları uzman görüşüne başvurularak ve deneme yanılma yoluyla belirlenirler. Bu nedenle eğer klasik yollardan çözüme ulaşmak mümkünse, yani bulanık olmayan ortamda klasik regresyonun varsayımları sağlanıyor iken yeterli veriye ulaşılabiliyor ise, bulanık regresyon yöntemi tercih edilmemelidir.

Ancak uygulamada bulanık yöntemin kullanılmasının kaçınılmaz olduğu bazı durumlar vardır ve bunlar üç başlık altında özetlenebilir:

- Bağımlı ve bağımsız değişkenlerin bulanık olması,
- Değişkenler arasındaki ilişkinin bulanık olması,
- Hem değişkenlerin hem de aralarındaki ilişkinin bulanık olması.

Söz konusu durumlarda artık keskin (reel) değerlerle çalışılamaz, bunun yerine bulanık sayılar (bulanık kümeler) kullanılmalıdır. Bulanık ilişkiler daha az kesin olmasına rağmen, bazı durumlarda sezgisel olarak daha gerçekçi görünmektedir.

Klasik yaklaşımda nitel veriler belli gruplara ayrılarak sayısal kodlama yapılır ve böylece modele dahil edilirler. Örneğin "iyi" kelimesi için 5 üzerinden bir değerleme sistemi söz konusu olduğunda, "4" rakamı kodlanabilir. Fakat gruplandırılabilmesi mümkün olmayan nitel değişkenler de vardır. Böyle değişkenlere "özünde bulanık değişkenler" denir. Anlamları kişiden kişiye, durumdan duruma farklılık gösterir. Örneğin "sevimli, dürüst" gibi kelimeler özünde bulanık kavramlardır ve gerçekleşme dereceleri vardır. Bu dereceler üyelik fonksiyonuyla belirlenir. Bir insanın ne kadar dürüst olduğuna ilişkin bir soruya 0.8 gibi bir cevap alınıyorsa, bu kişinin %80 güvenilir olduğu anlaşılır.

Bulanık yaklaşım, önermeleri iki uçlu değer almaktan, gruplara ayrılıp kodlanmaktan kurtarır ve sorunların olduğu gibi ele alınmasını sağlar. Zira gerçek yaşamda her şeyin bir meydana geliş derecesi vardır. Çok kesin gibi görünen olaylar dahi daha yakından incelendikçe bulanıklaşabilecektir.

3. Bulanık Doğrusal Regresyon Modelinin Bileşenleri

Bulanık doğrusal regresyon modelinin genel yapısı aşağıdaki eşitlikteki gibidir:[3]

$$\tilde{Y}_i = \tilde{A}_0 + \tilde{A}_1 X_{i1} + ...+ \tilde{A}_k X_{kn} \quad i=1,2,....n \quad j=0,1,2....k \tag{3.1}$$

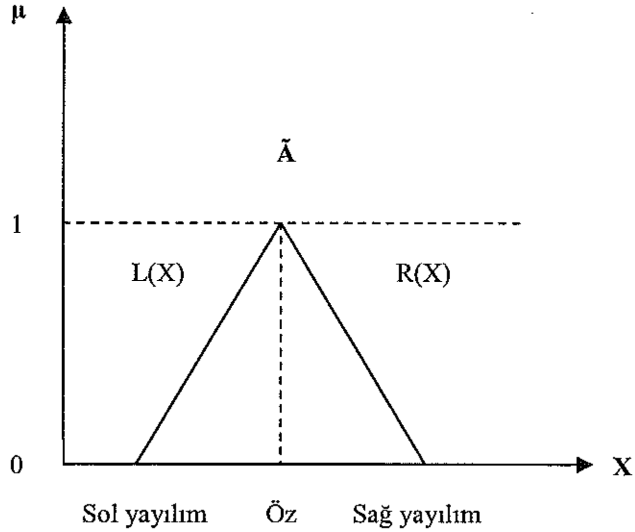
n, gözlem sayısıdır. k ise açıklayıcı değişken sayısıdır. X_j girdi değişkenleri (açıklayıcı değişkenler) keskin (reel) değerlidir. $\tilde{A}_j$'ler bulanık katsayılardır (sistem parametreleridir).

$$\tilde{A}_j = (a_j, c_j)_L \tag{3.2}$$

$$\tilde{A}_j = (a_j, c_j)_R \tag{3.3}$$

3 S. Ghoshray, "Fuzzy Linear Regression Analysis by Symmetric Triangular Fuzzy Number Coefficients", IEEE, 1997, pp. 307-313.

Yukarıdaki gösterimlerde a_j değeri $\tilde{A}_j$'nin merkezi, c_j ise yarı yayılımıdır (yarıçaptır). L ve R, $\tilde{A}_j$ üçgensel bulanık sayısının referans fonksiyonlarıdır. L(X) ve R(X) bir araya gelerek $\tilde{A}_j$'nin üyelik fonksiyonunu oluştururlar. Bu anlatım, Şekil 3.1.'de gösterilmektedir.



Şekil 3.1: Bulanık parametrenin geometrik gösterimi

Bulanık sayının özü; merkez değeri demektir, yayılımı ise yarıçap gibi düşünülmelidir. Söz konusu yarıçap, bulanık sayının tek bir sayı olarak değil, bir aralık (bir küme) olarak algılanmasını gerektirmektedir. Yayılım, değişkenlik olarak da tanımlanabilir. Yani belirtilen aralıkta bulanık sayı (bulanık küme) sonsuz sayıda değeri kapsamaktadır. Bulanık sayının matematiksel ifadesi aşağıdaki gibidir:

Sol yayılım $= a_j - c_j$

Sağ yayılım $= a_j + c_j$

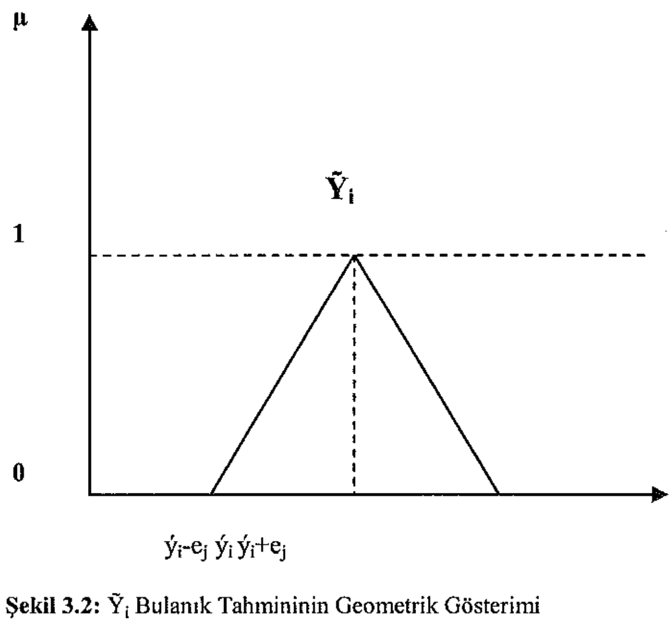
Öz (merkez) $= a_j$

Bulanık regresyonda genellikle $\tilde{A}_j$ bulanık parametreleri simetrik olduğundan, bundan sonraki gösterimlerde "L" ve "R" gösterimleri kullanılmayacaktır. $\tilde{A}_j$ bulanık parametreleri; $\tilde{A}_j=(a_j,\ c_j)$ olarak ifade edildiğinde, Tanaka modeli aşağıdaki gibi elde edilir:

$$\tilde{Y}_i = (a_0, c_0) + (a_1, c_1)X_{11} + (a_1, c_1)X_{12} + (a_1, c_1)X_{13} + \ldots + (a_j, c_j)X_{kn}$$

$$i = 1,2....n \qquad j=0,1,2......k \qquad (3.4)$$

$\tilde{Y}_i = (\acute{y}_i,\ e_i)$ şeklinde tanımlanır.[4] $\acute{y}_i$, $\tilde{Y}_i$ bulanık tahmininin merkezini ($\tilde{A}_j$ bulanık parametrelerinin merkez değerleri a_j'ler gibi anlaşılmalıdır), e_i ise yarı yayılımını ($\tilde{A}_j$ bulanık parametrelerinin yayılım değerleri c_j'ler gibi anlaşılmalıdır) göstermektedir.



Şekil 3.2: $\tilde{Y}_i$ Bulanık Tahmininin Geometrik Gösterimi

[4] Y.H.O. Chang, B.M.Ayyub, "Fuzzy regression methods-a comparative assessment", Fuzzy Sets and Systems, 2001, pp.187-203.

Tanaka'nın merkez ve yayılım değerleri ile ifade edilen modeline tekrar bakılırsa, $\acute{y}_i = \sum a_j X_{ij}$ olduğu anlaşılır. Aynı şekilde $e_i = \sum c_j \left| X_{ij} \right|$ 'dır. Burada mutlak değer kullanılmasının sebebi, bir uzunluğun negatif olamamasındandır.

$\tilde{Y}_i$ bulanık tahmininin ($\tilde{Y}_i$ bulanık aralığının) alt sınırı;

$$\acute{y}_i - e_j = \sum a_j X_{ij} - \sum c_j \left| X_{ij} \right| \quad (3.5)$$

$\tilde{Y}_i$ bulanık tahmininin ($\tilde{Y}_i$ bulanık aralığının) üst sınırı;

$$\acute{y}_i + e_j = \sum a_j X_{ij} + \sum c_j \left| X_{ij} \right| \quad (3.6)$$

şeklinde elde edilir.[5]

$e_i = \sum c_j \left| X_{ij} \right|$ değeri Tanaka modelinde toplam yayılımı yani toplam hatayı temsil etmektedir. Bulanık doğrusal regresyon modelinde Klasik regresyon modelinde olduğu gibi ayrıca bir hata terimi yoktur. Modelin toplam hatası, bulanık parametrelerin toplam yayılımları (yarıçapları, değişkenliği) ile elde edilir. $e_i = \sum c_j \left| X_{ij} \right|$ değerinin klasik regresyondaki karşılığı $\sum e_i^2$'dir yani kalıntı kareler toplamıdır. Bu nedenle toplam yayılım (hata) minimize edilmek istenir.[6]

4. Bulanık Doğrusal Regresyon Analizinde "h terimi"nin anlamı

Klasik doğrusal regresyon analizinde model kurarken; bağımlı değişkeni açıklama ihtimali olan fakat modele dahil edilmemiş bir takım değişkenlerin yerini hata teriminin tuttuğu varsayılır. Böylece model kurarken yapılması muhtemel hatalar, tek bir hata teriminde toplanmış olur. Bulanık modelde ise hata; tüm parametrelere dağıtılmış vaziyettedir. Bu durumda her bir parametre

---

[5] Tanaka, H., Guo, P., **Possibilistic Data Analysis for Operations Research,** Physica - Verlag Heidelberg, New York, 1999, p. 96.

[6] Wang, H. F., Tsaur, R.C., "Insight of a Fuzzy Regression Model", Fuzzy Sets and Systems, 2000, p. 356.

belli bir bulanıklık seviyesinde tahmin edilir. Söz konusu bulanıklık seviyesi "h terimi" olarak adlandırılır ve [0,1] aralığında değer almaktadır. Bu değer, çalışmanın başında veri kümesinin eksik, yarım veya tam olma durumuna bakılarak analist tarafından belirlenir ve hesaplamalara sabit bir girdi (input) olarak dahil edilir.[7] h'ın 0 (sıfır) değerini alması, tahmin edilen model ile gözlem değerlerinin son derece uyumsuz olduğunu, 1 değerini alması ise son derece uyumlu olduğu anlamına gelmektedir. h'ın değerinin sıfıra yaklaşması, bulanık parametrelerin yayılımlarının artmasına sebep olur, 1'e yaklaşması ise bulanık parametrelerin yayılımlarının azalmasına sebep olur. Hatta h=1 alınması durumunda, bulanık parametreler sıfır yayılımla elde edilirler ve bulunan sonuçlar klasik regresyonun en küçük kareler tahmin değerlerine oldukça yakındır. Fakat h'ın 1 alınması yani sıfır bulanıklık düzeyinde çalışılması durumunda, bulanık regresyon yöntemini tercih etmek son derece gereksizdir, zira klasik regresyon tahmin değerlerine çok yakın sonuçlar elde etmek için bir sürü kısıt denklemi yazmak ve optimizasyon problemini çözmek gerekmektedir. Bulanık doğrusal regresyonda değişkenler arasında var olduğundan şüphelenilen ve tam olarak tanımlanamayan bazı belirsiz ilişkileri regresyon analizine dahil edebilmek amacıyla h terimi kullanılmaktadır. Klasik yollardan modele alınamayan veya Y bağımlı değişkeni üzerinde doğrudan bir etkisi olduğu saptanamayan, fakat sezgisel olarak etkisi olabileceği düşünülen herhangi bir açıklayıcı değişken, bulanık analiz uygulandığında, modele girmeyi hak edebilmektedir. h teriminin "veri kümesinin modele uygunluğu" şeklindeki tanımıyla anlatılmak istenen, veri kümesiyle ilgili şüphemizi en başından ortaya koymak ve bu bulanıklık seviyesinde çalışmaktır. Örneğin h=0.6 olarak belirlendiğinde, veri kümesinin %60 oranında doğru (veya tam) olduğu kabul edilmiş olunur. Veri kümesinin doğruluğu, veri kümesinin tam olup olmadığı veya ölçmedeki muhtemel hatalar göz önüne alındığında, ancak h seviyesinde güvenilebilir olduğudur. Klasik doğrusal regresyonda ise h=1 olduğu, yani verilerin tam doğru olduğu (hiç bulanık olmadığı) varsayılır. Zira yapılan herhangi bir ölçüm hatası ya da ulaşılamayan veri olduğunda, bu gibi hatalar zaten modeldeki hata terimi tarafından temsil edilmektedir. h seviyesinin en ideal değerinin ne olması gerektiği hala bir tartışma konusudur. Konuyla ilgili çalışmış olan bazı bilim adamları, h'ın optimum değerlerini belirleyerek bu değerler üzerinden bulanık tahmin yapılırsa başarılı olunacağını savunmuşlardır.

---

7 H.Moskowitz, K.J. Kim, "On assessing the h value in fuzzy linear regression", Fuzzy Sets and Systems, Volume 58, 1993, p. 304.

Örneğin Tanaka, Uejima ve Asai h=0.5 alınmasını, Gharpuray, Fan ve Lai ise h=0.9 alınmasını önermişlerdir. h'ın değerinin 1'e yaklaşması bulanıklığı azaltır ve parametrelerin tahmin aralığı azalarak toplam hata miktarı azalır, 0'a yaklaştığında ise bulanık aralıklar genişler ve toplam hata miktarı artar. Zira h'ın değeri araştırmacının elindeki verilere verdiği not olduğuna göre, sıfıra yaklaştıkça verilere verilen notun düştüğünü yani başarısız bir veri kümesiyle çalıştığımızı ve böylece bulanık aralık tahminlerinin genişleyerek toplam hatanın artacağını anlarız. Bire yaklaştığında ise bulanık aralıklar daralmakta ve toplam hata miktarı azalmaktadır. Dolayısıyla yapılan tahminler bulanıklıktan kesinliğe doğru gitmektedir. Bir benzetme yaparsak, h=1 (siyah veya beyaz), h=0 (beyaz veya siyah), h=0.4 (açık gri), h=0.8 (koyu gri) şeklindedir. Yani ikili mantıktan çok değerli mantığa yaptığımız bu geçiş, siyah ve beyaz arasındaki tonların da farkında olmak gibi algılanmalıdır.

5. Bulanık Modelin Tahmin Edilmesi (Tanaka Modeli)

Bulanık model aşağıdaki doğrusal programlama probleminin çözülmesiyle elde edilir.[8]

$$\min\,(c_0 + c_1 + \ldots c_k)$$

kısıtı altında (subject to);

$$\sum[a_j + (1\text{-}h)c_j]X_{ij} \geq \acute{y}_i + (1\text{-}h)e_i$$

$$\sum[a_j - (1\text{-}h)c_j]X_{ij} \leq \acute{y}_i - (1\text{-}h)e_i$$

$$i=1,2\ldots n \quad j=0,1,2\ldots k \tag{5.1}$$

Amaç fonksiyonu toplam yayılımı (hatayı, bulanıklığı) minimize eder. Kısıtların olduğu bölümde ise, verilerin belirlenen h düzeyinde tahmin edilen bulanık aralıklar tarafından içerilmesi gerekliliği gösterilmektedir. Doğrusal

[8] M. Hojati, C.R. Bector, K. Smimou, "A simple method for computation of fuzzy linear regression", European Journal of Operational Research, 2004, p. 3.

programlama problemi çözülerek bulanık parametreler yani $\tilde{A}_j$'lerin merkez (a_j) ve yayılım(c_j) değerleri elde edilir.

Tanaka'nın modeli hesaplamalarda son derece etkin ve kolay olmasına rağmen önemli bir eksikliği vardır. Aykırı gözlemleri de içerebilmek amacıyla bulanık aralıklar geniş tahmin edilmek zorunda kalınacağı için toplam hata miktarı artacaktır. Bunu çözmenin kolay ve akılcı bir yolu, amaç fonksiyonuna açıklayıcı değişkenleri de ekleyerek minimizasyon işlemini yapmaktır. Bu durumda yeni amaç fonksiyonu aşağıdaki gibi olacaktır:

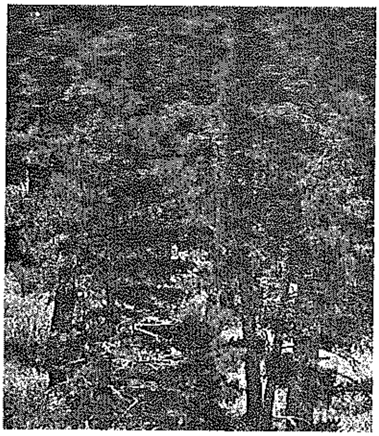
$$\min \sum\sum c_j X_{ij} \tag{5.2}$$

Amaç fonksiyonu bu şekilde belirlenirse, aykırı gözlemler de yayılımlarla beraber minimize edilerek aykırı gözlemlerden kaynaklanan bulanıklık artışının önüne geçilmiş olacaktır.

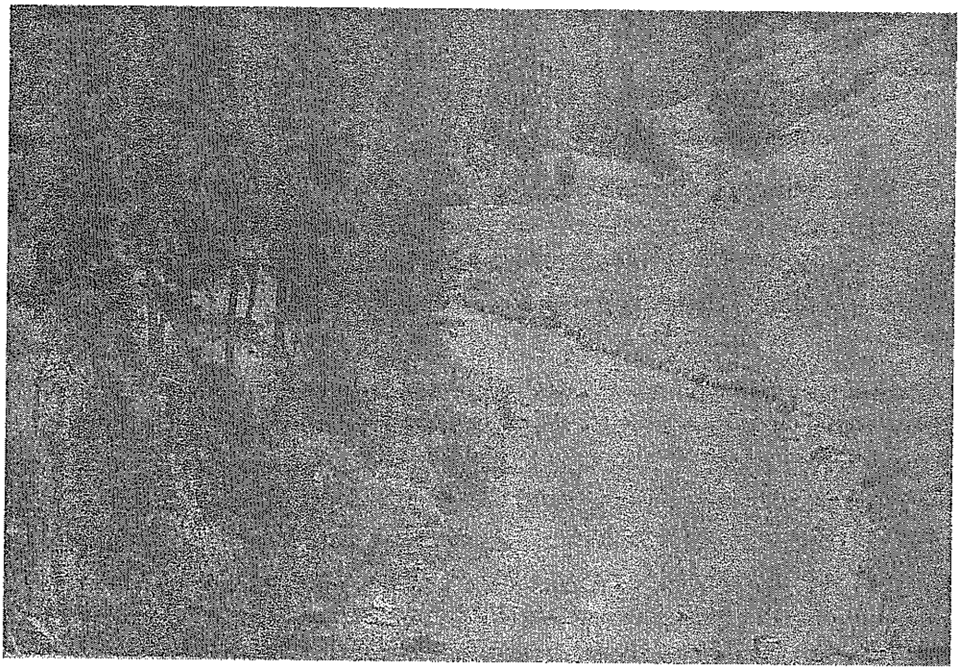
6. Klasik Doğrusal Regresyon ve Bulanık Doğrusal Regresyon Yöntemleri Arasındaki Ayrım

Bulanık regresyon analizi, klasik regresyon analizinin uygulanması için gerekli şartlar oluşturulamadığında; özellikle gözlem sayısının azlığı ve (veya) bağımlı değişkeni etkilemekte olduğu bilindiği halde, kurulan modelde anlamsız bulunan açıklayıcı değişkenlerin varlığı durumunda, klasik regresyon modelinin bulguları yetersiz ve tutarsız olduğunda ortaya çıkan alternatif bir yöntemdir.

Bulanık regresyon analizi aslında bir doğrusal programlama yöntemidir. Tahmin edilen Y değerleri klasik modeldeki gibi tek bir değer olarak değil de, bir aralık olarak sunulur. Yani her bir Y gözlemi için bir alt ($\tilde{Y}_{alt}$) ve bir de üst sınır ($\tilde{Y}^{üst}$) olmak üzere aralık tahmini yapılır. Klasik modelde de elbette güven aralıkları oluşturulabilmektedir, fakat burada esas vurgulanması gereken nokta şudur: Önemli olduğu halde modele dahil edilemeyen açıklayıcı değişkenlerin varlığı durumunda, elde edilen tahmin denklemi ve buradan yola çıkarak elde edilen güven aralıkları ne kadar güvenilir ve sağlıklı olacaktır? İşte bulanık regresyon analizinin en önemli başarısı, tüm açıklayıcı değişkenleri modele dahil etmesinden kaynaklanmaktadır. Yani klasik modelin olayın sadece bir bölümünü resmedebildiği durumlarda, bulanık model; bulanık da olsa tüm manzarayı gözler önüne serebilmektedir.



Şekil 6.1: Klasik modelin yeterli şartlar oluşmadığında sunduğu tahminlerin fotoğraf örneği ile sembolik bir gösterimi



Şekil 6.2: Bulanık modelin tahmin gücü.

Şekil 6.1 ve Şekil 6.2'den anlaşılacağı üzere, klasik model için yeterli şartlar oluşturulamadığında (yeterli veriye ulaşılamaması veya teoride bağımlı değişkeni etkilediği öngörülmesine rağmen çeşitli nedenlerden dolayı modele dahil edilemeyen bir takım açıklayıcı değişkenlerin varlığında); bulanık model, klasik modele göre çok daha nitelikli bulgular sunmaktadır. Yukarıdaki fotoğraf benzetmesinde klasik ve bulanık yöntemlerin tahmin güçleri sembolik bir şekilde kıyaslanmıştır. Şekil 6.1'de ormandaki ağaçların sadece bir kısmı görülebilirken, bulanık modelin tahmin gücünü temsilen gösterilmiş olan Şekil 6.2'de bulanık da olsa ormanın daha büyük bir kısmı görülebilmektedir.

7. Uygulama

Çalışmanın bu bölümünde, Tanzi'nin nakit para talebi denklemi, Türkiye için hem klasik hem de bulanık regresyon analizi yöntemleriyle tahmin edilmiştir. Elde edilen bulgular karşılaştırmalı olarak sunulmuştur. Burada amaç kayıt dışı ekonominin büyüklüğünü ölçmek değildir. Yapılmak istenen sadece klasik ve bulanık yöntemi kıyaslamaktır. Çalışmada yaralanılan örnek büyüklüğü on üç (13) gözlemden ibarettir. Daha fazla bilgiye (veriye) ulaşmak olanaklıysa da, *özellikle yetersiz örnek büyüklüğü olması durumunda*, klasik regresyonun yeterince iyi tahminler yapamadığını göstererek, bulanık regresyon yöntemini ön plana çıkarmak amaçlanmıştır.

Tanzi'nin nakit para talebi denklemi, kayıt dışı ekonominin ekonometrik olarak ölçülmesini sağlamaktadır. Kayıt dışı ekonominin tahmininde nakit para denkleminin kullanılmasının nedeni, kayıt dışı ekonomi büyüdüğünde, buna bağlı olarak, nakit para talebinin de büyümesidir. Para talebindeki artış ise ekonometrik bir nakit para modelinin tahmin edilmesiyle tespit edilebilmektedir.

Tanzi'nin nakit para talebi denklemi;

$$Ln(C/M2) = \beta_0 + \beta_1 Ln(1+TW)_t + \beta_2 Ln(WS/Y)_t + \beta_3 LnR_t + \beta_4 Ln(Y/N)_t + \varepsilon_t \tag{7.1}$$

Modeldeki açıklayıcı değişkenler aşağıdaki gibidir:

Y : Milli gelir

Y/N : Kişi başına reel milli gelir

R : Mevduat faiz oranı

TW : Ortalama vergi oranı

WS : Maaş ve ücretlilere yapılan harcamalar

C : Dolaşımdaki nakit para

M2 : Geniş anlamlı para arzı

Nakit para denkleminde β_1 (vergi oranı), β_2 (maaş ve ücretlilerin gelirlerinin milli gelire oranı), β_4 (kişi başına milli gelir)'in işaretleri pozitif (kayıt dışı ekonomiyi büyüten), β_3 (faiz oranı)'ün işareti ise negatif (kayıt dışı ekonomiyi azaltan) olması beklenir. Çünkü yüksek faiz gelirleri nakit para talebini azaltmaktadır. Bu bekleyişler matematiksel olarak $\beta_1>0$, $\beta_2>0$, $\beta_4>0$ ve $\beta_3< 0$ şeklinde ifade edilebilir.

Türkiye için yapılan bu çalışmada değişkenler uyarlanarak kullanılmıştır.[9] Örneğin M2 yerine M1 tercih edilmiştir. M2'nin ödemelerde kullanım olanağı kısıtlı olan vadeli mevduatları da kapsaması nedeniyle, M2 yerine M1 (dar anlamlı para arzı) kullanılmıştır. Bunun başlıca nedeni, Türkiye'de çalışmaya konu olan dönemde yüksek enflasyon oranlarının olmasıdır. Tahmin denklemindeki açıklayıcı değişkenler aşağıdaki gibi ifade edilmektedir:

X_1 : Kişi başına reel GSMH

X_2 : Yıllık ortalama mevduat faiz oranı

X_3 : Tefe genel endeks

X_4 : Ortalama vergi oranı + 1 (toplam vergi gelirlerinin GSMH'a oranının 1 ile toplamı)

X_5 : Kamu personel harcamalarının toplam kamu harcamalarına oranı

Y : Nakit para stoğu / M1

Tanzi denkleminin öngördüğü şekilde değişkenlerin logaritmik değerleri alınarak 1. farkları alındığında tahmin edilen model aşağıdaki gibidir:

$$\nabla LnY = -2{,}92 + 0{,}192 * \nabla LnX1 - 0{,}178 * \nabla LnX2 - 2.35 * \nabla LnX4 + 0{,}194 * \nabla LnX5 \quad (7.2)$$

[9] E.Aktürk, "Kayıt dışı Ekonomi ve Türkiye Üzerine Bir Uygulama", Atatürk Üniversitesi Sosyal Bilimler Enstitüsü İktisat Anabilim Dalı Doktora Tezi, Erzurum, 2003, s.25.

Tablo 7.1: Nakit para talebi denkleminin en küçük kareler yöntemi ile tahmini

Dependent Variable: LNY
Method: Least Squares
Date: 04/11/08 Time: 14:27
Sample: 1994 2006
Included observations: 13

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-2.925810	2.963995	-0.987117	0.3565
LNX1	0.192667	0.211585	0.910591	0.3928
LNX2	-0.178284	0.051802	-3.441617	0.0108
LNX3	0.002975	0.036882	0.080657	0.9380
LNX4	-2.350742	1.037050	-2.266758	0.0578
LNX5	0.194547	0.276261	0.704215	0.5040

R-kare	0.810144	Bağımlı değişkenin ort.	-0.723142
Düzeltilmiş R-kare	0.674533	Bağımlı değişkenin st. Sapması	0.071205
Regresyon standart hatası	0.040622	Akaike bilgi kriteri (AIC)	-3.264957
Kalıntı kareler toplamı	0.011551	Schwarz bilgi kriteri	-3.004211
Log likelihood	27.22222	F-istatistiği	5.974011
Durbin-Watson	2.299405	F istatistiğinin P-değeri	0.018220

Tablo 7.1'e bakıldığında, açıklayıcı değişkenlerin işaretleri (X_4:vergi oranı hariç) teoriyle tutarlı bulunmuştur. Sadece mevduat faiz oranı değişkeni istatistiksel olarak anlamlı bulunmuştur. Hata payı %5 yerine %10 olarak düşünülürse bir anlamda vergi oranı da anlamlı sayılabilir. Fakat teorik olarak vergi oranı arttığında nakit para talebinin artması beklenir. Yani yükselen vergi oranları kayıt dışını körüklemektedir. Halbuki burada yapılan tahmin sonucunda vergi oranının işareti negatif bulunmuştur.

Vergi oranını anlamlı olarak kabul edip, anlamsız bulunan değişkenler modelden dışlandıktan sonra tahmin edilen ikinci model aşağıdaki gibidir:

$$LNY = -0.336311897 - 0.2039834976*LNX2 - 2.724324603*LNX4 \quad (7.3)$$

Tablo 7.2: Nakit para talebi denkleminin anlamlı bulunan açıklayıcı değişkenler üzerinden en küçük kareler yöntemi ile tahmini

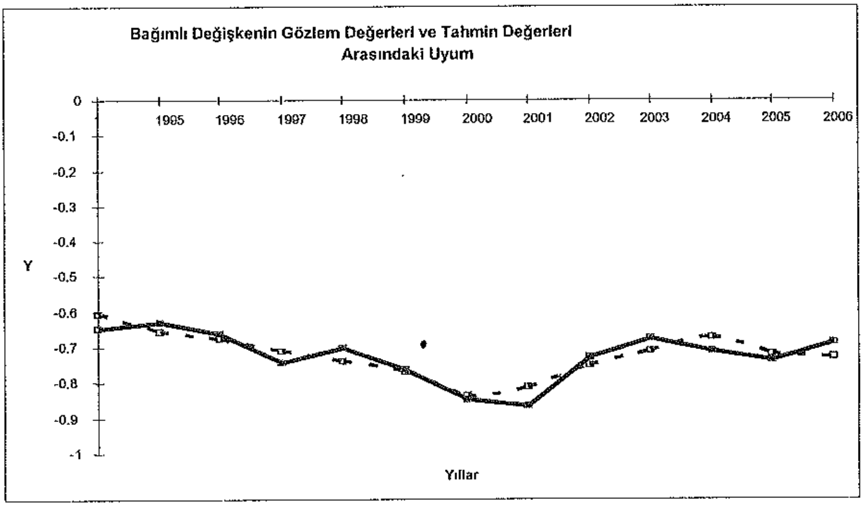
Dependent Variable: LNY
Method: Least Squares
Date: 04/02/08 Time: 16:22
Sample: 1994 2006
Included observations: 13

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.336312	0.087950	-3.823898	0.0034
LNX2	-0.203983	0.046508	-4.385996	0.0014
LNX4	-2.724325	0.587090	-4.640385	0.0009

R-kare	0.687393	Bağımlı değişkenin ort.	-0.723142
Düzeltilmiş R-kare	0.624872	Bağımlı değişkenin st. Sapması	0.071205
Regresyon standart hatası	0.043612	Akaike bilgi kriteri (AIC)	-3.227814
Kalıntı kareler toplamı	0.019020	Schwarz bilgi kriteri	-3.097441
Log likelihood	23.98079	F-istatistiği	10.99452
Durbin-Watson	1.709786	F istatistiğinin P-değeri	0.002985

Tahmin edilen ikinci modelde tüm açıklayıcı değişkenler anlamlıdır. Kalıntıların normal dağılmakta olduğu Jarque-Bera testinin sonucundan anlaşılmaktadır (JB=0.9647, Probability=0.6173). Yokluk hipotezi; "H_0:Kalıntılar normal dağılmaktadır" reddedilemez. Durbin-Watson değeri, sabit terim hariç 2 açıklayıcı değişken ve 13 gözleme karşılık gelen 1.562 değerinden büyük olduğu için birinci dereceden bir otokorelasyon yoktur.

İkinci model temel bir takım testlerden başarıyla geçmiş olmasına rağmen, acaba bağımlı değişkene ilişkin gözlem değerlerini ne ölçüde tahmin edebilmiştir? Bu sorunun cevabı Şekil 7.1'de verilmektedir.



Şekil 7.1: Y Gözlem Değerleri ile Tahmin Değerleri Arasındaki Uyum

Şekil 7.1'de kesikli çizgilerle gösterilen seri $\hat{Y}$ tahmin değerlerini, sürekli çizgili olan ise Y gözlem değerlerini göstermektedir. İki seri adeta bir sarmal gibi uyumlu görünmesine rağmen, aslında gözlem değerlerinin düştüğü bazı noktalarda tahmin değerlerinin yükselmekte olduğu görülmektedir.

Bu durumda ikinci model de yetersiz bir model olarak karşımıza çıkmaktadır. Zira gözlem sayısı oldukça azdır (13 gözlem) ve teoride önemli olduğu bilinen bir takım değişkenler (Kişi başına reel GSMH, Tefe genel endeks, Kamu personel harcamalarının toplam kamu harcamalarına oranı) dışlanmak zorunda kalınmıştır. Bu aşamada devreye bulanık regresyon analizi girmelidir.

Öncelikle $\tilde{A}_j$ bulanık parametreleri; $\tilde{A}_j=(a_j,\ c_j)$ şeklinde doğrusal programlama yöntemiyle elde edilecek, ardından $\tilde{Y}_i$ bulanık aralıkları (alt ve üst regresyon eğrileri) hesaplanacaktır.

Bulandırma seviyesi h=0.9 olarak seçilmiştir (h'ın değeri 0 ile 1 arasında keyfi bir değerdir, 1'e yaklaştıkça bulanıklık miktarı azalacağından, 0.9 gibi yüksek bir değer seçilmiştir.)

A katsayılar matrisi; 2*n yani 26 adet satır (Y'ler için bir alt ve bir de üst sınır hesaplanacağı için gözlem sayısının 2 katı kadar sayıda kısıt denklemi yazılması gerekmektedir). Bunun yanında, 5 açıklayıcı değişken ve bir de sabit terim için olmak üzere, birer merkez ve birer de yayılım değeri hesaplamamız gerekecektir. Bu durumda A matrisi 12 sütuna sahip olacaktır ve boyutu 26*12 olacaktır.

Kısıt denklemleri aşağıdaki gibi yazılmaktadır:

Y_1 (1994 yılı için) İlk Kısıt Denklemi:

$[-a_0 - a_1(14,23) - a_2(-0,48) - a_3(13,27) - a_4\,(0,14) - a_5(-1,19)] - 0,9[\,c_0 + c_1(14,23) + c_2(-0,48) + c_3(13,27) + c_4\,(0,14) + c_5(-1,19)] \leq LnY_1$

Y_1 (1994 yılı için) İkinci Kısıt Denklemi:

$[a_0 + a_1(14,23) + a_2(-0,48) + a_3(13,27) + a_4\,(0,14) + a_5(-1,19)] - 0,9[\,c_0 + c_1(14,23) + c_2(-0,48) + c_3(13,27) + c_4\,(0,14) + c_5(-1,19)] \leq LnY_1$

Y_2 (1995 yılı için) İlk Kısıt Denklemi:

$[-a_0 - a_1(14,28) - a_2(-0,17) - a_3(13,91) - a_4\,(0,13) - a_5(-1,23)] - 0,9[\,c_0 + c_1(14,28) + c_2(-0,17) + c_3(13,91) + c_4\,(0,13) + c_5(-1,23)] \leq LnY_2$

Y_2 (1995 yılı için) İkinci Kısıt Denklemi:

$[a_0 + a_1(14,28) + a_2(-0,17) + a_3(13,91) + a_4\,(0,13) + a_5(-1,23)] - 0,9[\,c_0 + c_1(14,28) + c_2(-0,17) + c_3(13,91) + c_4\,(0,13) + c_5(-1,23)] \leq LnY_2$

Bu şekilde 26 adet (2*n sayıda) kısıt denklemi yazıldıktan sonra yapılan hesaplamalar sonucunda A katsayılar matrisi aşağıdaki gibi elde edilmiştir:

A =

-1	-14.23	0.48	-13.27	-0.14	1.19	-0.9	-12.807	0.432	-11.943	-0.126	1.071
1	14.23	-0.48	13.27	0.14	-1.19	-0.9	-12.807	0.432	-11.943	-0.126	1.071
-1	-14.28	0.17	-13.91	-0.13	1.23	-0.9	-12.852	0.153	-12.519	-0.117	1.107
1	14.28	-0.17	13.91	0.13	-1.23	-0.9	-12.852	0.153	-12.519	-0.117	1.107
-1	-14.34	0.27	-14.45	-0.14	1.4	-0.9	-12.906	0.243	-13.005	-0.126	1.107
1	14.34	-0.27	14.45	0.14	-1.4	-0.9	-12.906	0.243	-13.005	-0.126	1.107
-1	-14.42	0.24	-15.03	-0.15	1.35	-0.9	-12.978	0.216	-13.527	-0.135	1.215
1	14.42	-0.24	15.03	0.15	-1.35	-0.9	-12.978	0.216	-13.527	-0.135	1.215
-1	-14.44	0.24	-15.55	-0.16	1.39	-0.9	-12.966	0.216	-13.995	-0.144	1.251
1	14.44	-0.24	15.55	0.16	-1.39	-0.9	-12.966	0.216	-13.995	-0.144	1.251
-1	-14.37	0.32	-15.91	-0.17	1.4	-0.9	-12.933	0.288	-14.319	-0.153	1.26
1	14.37	-0.32	15.91	0.17	-1.4	-0.9	-12.933	0.288	-14.319	-0.153	1.26
-1	-14.38	0.2	-16.35	-0.2	1.54	-0.9	-12.942	0.18	-14.715	-0.18	1.386
1	14.38	-0.2	16.35	0.2	-1.54	-0.9	-12.942	0.18	-14.715	-0.18	1.386
-1	-14.26	0.51	-16.8	-0.2	1.66	-0.9	-13.4	0.459	-15.12	-0.18	1.494
1	14.26	-0.51	16.8	0.2	-1.66	-0.9	-13.4	0.459	-15.12	-0.18	1.494
-1	-14.32	0.79	-17.18	-0.2	1.61	-0.9	-12.888	0.711	-15.462	-0.18	1.449
1	14.32	-0.79	17.18	0.2	-1.61	-0.9	-12.888	0.711	-15.462	-0.18	1.449
-1	-14.37	1.27	-17.39	-0.21	1.53	-0.9	-12.933	1.143	-15.651	-0.189	-1.377
1	14.37	-1.27	17.39	0.21	-1.53	-0.9	-12.933	1.143	-15.651	-0.189	-1.377
-1	-14.44	1.52	-17.5	-0.23	1.58	-0.9	-12.966	1.368	-15.75	-0.207	1.422
1	14.44	-1.52	17.5	0.23	-1.58	-0.9	-12.966	1.368	-15.75	-0.207	1.422
-1	-14.51	1.59	-17.5	-0.25	1.52	-0.9	-13.059	1.431	-15.75	-0.225	1.368
1	14.51	-1.59	17.5	0.25	-1.52	-0.9	-13.059	1.431	-15.75	-0.225	1.368
-1	-14.56	1.46	-17.63	-0.27	1.53	-0.9	-13.104	1.314	-15.867	-0.243	1.377
1	14.56	-1.46	17.63	0.27	-1.53	-0.9	-13.104	1.314	-15.867	-0.243	1.377

F-10

b çözüm vektörü ise, Y'nin bir negatif bir de pozitif değerlerinin [-lnY_1, lnY_1, -lnY_2, lnY_2,......,-lnY_{13}, lnY_{13}] şeklinde sıralanmasıyla elde edilir;

b Çözüm Vektörü:

b=[0.64 -0.64 0.62 -0.62 0.66 -0.66 0.74 -0.74 0.7 -0.7 0.76 -0.76 0.84 -0.84 0.86 -0.86 0.72 -0.72 0.67 -0.67 0.71 -0.71 0.73 -0.73 0.68 -0.68]

z Amaç Fonksiyonunun Katsayılar Vektörü:

Doğrusal programlama ile minimizasyon problemini çözebilmek için gerekli olan z amaç fonksiyonu, A katsayılar matrisi ile b çözüm vektörü elde edildikten sonra aşağıdaki gibi elde edilir;

$$[z]_{1*12} = [b^{-1}]_{1*26}*[A]_{26*12} \tag{7.4}$$

z=[0.076 14.38 -0.702 16.044 0.195 -1.46 1 186.98 -9.12 208.58 2.535 -18.9854]

Matlab 7.3.0'da komut satırında (z, A ve b tanıtıldıktan sonra) **>>[x,fval,exitflag,output]=linprog[z,A,b)** komutu yazılarak bulanık merkez ve yayılım değerleri aşağıdaki gibi elde edilmiştir:

x =[-0.0016 -0.0021 -0.0013 5.9070 0.0005 0.0001 -0.0007 -0.0025 -0.001 6.5635 0.0006 0]

X vektörünün ilk altı değeri merkez değerler, son altı değeri ise yayılımlardır.

a_0 = -0.0016

a_1 = -0.0021

a_2 = -0.0013

a_3 = 5.9070

a_4 = 0.0005

a_5 = 0.0001

$c_0 = -0.0007$

$c_1 = -0.0025$

$c_2 = -0.001$

$c_3 = 6.5635$

$c_4 = 0.0006$

$c_5 = 0$

Bulanık merkez ve yayılım değerleri bulunduktan sonra; alt ve üst regresyon eğrileri aşağıdaki gibi hesaplanmaktadır:

Tablo 7.2: Bulanık parametrelerin alt ve üst sınır değerleri

	A_0	A_1	A_2	A_3	A_4	A_5
ÜST (a_j + cj)	-0.0023	-0.0046	-0.0023	12.4705	0.0011	0.0001
ALT (a_j - cj)	-0.0009	0.0004	-0.0003	-0.6565	-1E-04	0.0001

Alt ve Üst Regresyon eğrileri aşağıdaki gibi elde edilir:

$$Y^{üst} = A_0^{üst} + A_1^{üst} * X_{j1} + \ldots + A_n^{üst} * X_{jn} \qquad (7.5)$$

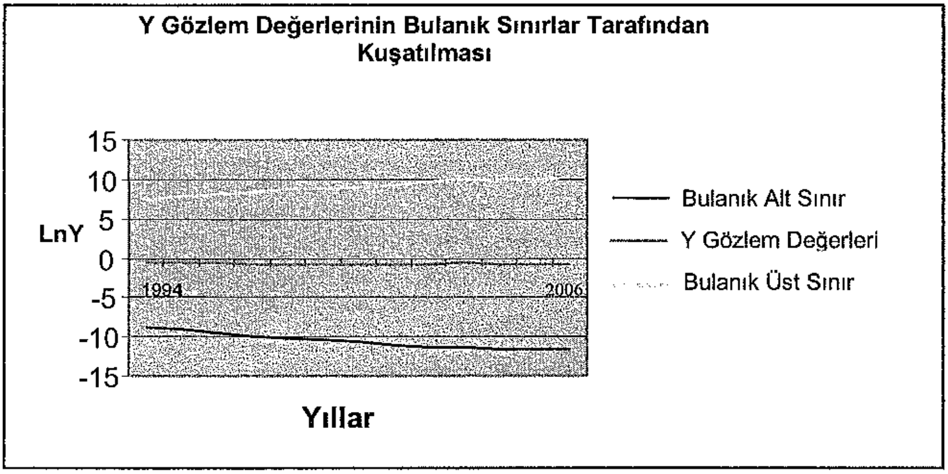
$$Y^{alt} = A_0^{alt} + A_1^{alt} * X_{j1} + \ldots + A_n^{alt} * X_{jn} \qquad (7.6)$$

Bu hesaplamalara göre elde edilen bulanık aralık değerleri ve gerçek Y gözlem değerleri aşağıdaki tabloda gösterilmektedir.

Tablo 7.3: Y gözlem değerleri ve bulanık tahmin değerleri

Yıllar	LnY_i	LnY^{alt}	$LnY^{üst}$
1994	-0.646452482	-8.711667152	7.418762188
1995	-0.628348501	-9.12725736	7.870560358
1996	-0.660853517	-9.482661055	8.160954021
1997	-0.742319206	-9.864295643	8.379657231
1998	-0.70207577	-10.20873653	8.804584994
1999	-0.762596261	-10.44656317	8.921370652
2000	-0.848326051	-10.72973069	9.033078588
2001	-0.865903731	-11.02686688	9.295059416
2002	-0.727692251	-11.27531903	9.819934533
2003	-0.677549991	-11.41505727	10.05995728
2004	-0.711112425	-11.48652299	10.06429814
2005	-0.738876063	-11.52333914	10.04558701
2006	-0.688740158	-11.57262921	10.1951489

Şekil 7.2: Y Gözlem Değerlerinin Bulanık Sınırlar Tarafından Kuşatılması



Şekil 7.2'den görüldüğü üzere, bulanık regresyon analizi kullanılarak yapılan tahmin sonucunda Y gözlem değerleri başarılı bir şekilde alttan ve üstten kuşatılmıştır. Halbuki Şekil 7.1'de gösterilen klasik regresyonun sunduğu tahmin değerleri Y gözlem değerlerini temsil etmekten uzaktır. Bunun başlıca

sebebi gözlem sayısının az olması (13 gözlem) ve teoride modele dahil edilmesi öngörülen bir takım değişkenlerin modele dahil edilememesidir.

Bulanık regresyon yöntemi istatistiksel bir tahmin değildir, sadece durum tespiti yapabilen, parametrelerin birer aralık şeklinde hesaplanabilmesini sağlayan bir doğrusal programlama yöntemidir. Dolayısıyla bulanık regresyon yönteminde parametrelerin anlamlılık sınaması yapılamaz. Seçilen bulanıklık seviyesinde hesaplanan (tahmin edilen denilemez, zira yapılmakta olan sadece bir doğrusal programla probleminin çözülmesidir) bulanık parametreler, h'ın değişen değerleri için genişleyebilir veya daraltılabilir. Örneğin çalışmada kabul edilen h=0.9 bulanıklık seviyesi, h=0.5 olarak alınmış olsaydı, bulanık parametreler daha geniş bir yayılımla (daha bulanık olarak) hesaplanacaktı. h'ın değerinin değişmesi, bulanık parametrelerin merkez değerlerini (a_j) etkilemez, sadece yayılımlarını değiştirir. Bu durum aşağıdaki teorem vasıtasıyla açıklanmaktadır:[10]

Teorem 7.1: Bulanık doğrusal regresyonun h_1 bulanıklık düzeyindeki optimal çözümü $\tilde{A}_j^*=(a_j^*, c_j^*)$ ise, h_2 bulanıklık düzeyindeki optimal çözümü;

$$\tilde{A}_j^{**} = \{a_j^*, [(1-h_2) / (1-h_1)]c_j^*\}\text{'dir.}$$

Örneğin uygulamada h=0,9 yerine h=0,5 almış olsaydık, $\tilde{A}_1$ bulanık parametresinin a_1 merkez değeri değişmeyecek, fakat c_1 yayılım değeri şu şekilde bulunacaktı:

$$c_1^{**} = [(1-0,5)/(1-0,9)]*c_1^* = 5*\,|-0.0025| = 0,0125 \text{ olarak elde edilir.}$$

Bulanıklık seviyesinin 0,9 dan 0,5'e yükselmesi sonucunda (h'ın değeri sıfıra yaklaştıkça bulanıklık artmaktadır) parametrelerin merkez değerleri aynı kalmış fakat yayılımları beş (5) kat artmıştır.

[10] Tanaka, H., Guo, P., a.g.e., p. 103.

8. Sonuç

Bulanık regresyon metodu, klasik yöntemin ikili mantığa dayanan kısıtlı mantığından daha üstün olan bulanık mantık kavramını temel alması nedeniyle yukarıda bahsedilen bir takım üstünlüklere sahiptir. Sosyal bilimlerde kodlanamayacak derecede bulanık olan bazı nitel değişkenlerin modele alınabilmesini sağlar. Bunun yanı sıra, bağımlı değişkeni sezgisel olarak etkilediği bilinen fakat sayısal kodlama yapılamadığı ve aradaki ilişkinin mutlaka bir fonksiyonel kalıba uyum göstermesini zorunlu kılan klasik regresyon analizinin modele dahil edemediği açıklayıcı değişkenleri kolaylıkla modele katarak parametre tahminlerini elde edebilmektedir. Bulanık yöntem, sadece beş gözlemle dahi bir takım sonuçlar ortaya koyabilir, fakat klasik regresyon analizinde genellikle en az 15-20 gözlem gerekmekte, hatta gözlem sayısı arttıkça modelin anlamlılığı artmaktadır. Bulanık yöntemde değişkenler arasındaki ilişkileri bir fonksiyon kalıbına uydurmak gereksizdir, zira çözüm doğrusal programlama yöntemi ile elde edilmektedir. Bulanık yöntem bütün bu üstünlüklerinin yanında bir takım olumsuzluklara da sahiptir. Bunun başında bulanık kısıtların yazılması gelmektedir. Her bir gözlem için ikişer kısıt yazılmalıdır ki bu durumda 2*n tane kısıt anlamına gelir. Gözlem sayısının artması hesaplamaları daha da karmaşık hale getirir. Örneğin 50 tane gözlem ve 4 tane açıklayıcı değişkenin bulunduğu sabit terimli modelde 5 tane parametre tahmini için birer merkez ve birer yayılım değeri hesaplanacağından çözüm matrisi 100x10 büyüklüğünde olacaktır. Her ne kadar işlemler bilgisayarla yapılıyor olsa da, kısıtlar elde ve tek tek yazıldıktan sonra bilgisayara aktarılmaktadır. Ayrıca bulanık regresyon analizi tamamıyla eldeki verilere bağlı hesaplamalar yaptığından dolayı, klasik regresyon analizi gibi dönem dışı tahmin yapamaz. Klasik modeli bir kez tahmin ettikten sonra istenilen bir zaman için öngörü yapılabilmektedir. Fakat bulanık regresyon analizi sadece eldeki verilere ait döneme ilişkin yorumlar sunabilmektedir.

Her ne kadar bulanık yöntem klasik yöntemin yetersiz kaldığı durumlarda yardımcı bir yöntem olsa da, bir sorunun klasik yollardan çözümü var ise, bulanık yöntem tercih edilmemelidir. Çünkü klasik yöntem sonucunda tahmin edilen modelin türevlenebilir oluşu, özellikle iktisadi bakımdan esneklikleri ortaya koyabilmesi nedeniyle üstünlüğünü devam ettirmektedir. Bulanık regresyon yöntemi klasik regresyon için elverişli şartlar sağlanmadığında alternatif bir yöntem olarak tercih edilebilse de, mevcut haliyle bulanık regresyon analizinin, klasik regresyon analizinden daha üstün bir yöntem olduğu söylenemez.

KAYNAKÇA

Aktürk, E., **"Kayıt Dışı Ekonomi ve Türkiye Üzerine Bir Uygulama"**, Atatürk Üniversitesi Sosyal Bilimler Enstitüsü İktisat Anabilim Dalı Doktora Tezi, Erzurum, 2003.

Baykal, N., Beyan, T., **Bulanık Mantık İlke ve Temelleri, 2004, Bıçaklar Kitapevi,** Ankara.

Chang, Y. H. O., Ayyub, B.M., **"Fuzzy Regression Methods-A Comparative Assessment"**, Fuzzy Sets and Systems, 2001.

Elkan, C., **The Paradoxical Success of Fuzzy Logic**, MIT Press, San Diego.

Ghoshray, S., **"Fuzzy Linear Regression Analysis by Symmetric Triangular Fuzzy Number Coefficients"**, IEEE, 1997.

Kosko, B., **Fuzzy Thinking,** 1993, Harper Collins Publishers, London.

Moskowitz, H., Kim, K., **"On Assessing the h Value in Fuzzy Linear Regression"**, Fuzzy Sets and Systems, 1993.

Şen, Z., **Bulanık Mantık ve Modelleme İlkeleri**, 2001, Bilge Kültür Sanat Kitapevi, İstanbul.

Tanaka, H., Guo, P., **Possibilistic Data Analysis for Operations Research**, 1999, Physica - Verlag Heidelberg, NY.

Tanaka, H., Uejima, S., Asai, K., **"Linear Regression Analysis With Fuzzy Model"**, IEEE, 1982.

Tanaka, K., **An Introduction to Fuzzy Logic for Practical Applications**, 1997, Rassel Inc., NY.

Uysal, M., **Matlab ile Matematiksel Uygulamalar ve Mühendislik Uygulamaları**, 2004, Beta Basım A.Ş., İstanbul.

Wang, H.F., Tsaur, R. C.; **"Insight of a Fuzzy Regression Model"**, Fuzzy Sets and Systems, 2000, pp. 355 – 369.

Wang, H.F., Tsaur, R.C.; **"Resolution of Fuzzy Regression Model"**, European Journal of Operatinal Research, 2000, pp. 637-650.

Zadeh, L.A., **"Fuzzy Logic = Computing With Words"**, IEEE Transactions on Fuzzy Systems, 1996, pp. 103-111.

Ek-1: Veri Seti

Y^*	X_1^*	X_2^*	X_3^*	X_4^φ	X_5^*
0.523901032	1514346	0.6179	583722.7888	1.156680773	0.302577195
0.533472103	1606454	0.8366	1099223.007	1.144065407	0.2914991
0.516410381	1691943	0.7606	1888933.836	1.155528196	0.245915743
0.476008672	1838576	0.7827	3378298.83	1.169435245	0.257524858
0.495555576	1880016	0.7851	5708975.464	1.181152223	0.247911853
0.466453818	1741293	0.7211	8201291.263	1.208548727	0.246110184
0.428131002	1766124	0.812	12623396.6	1.220657596	0.213655455
0.420671207	1570770	0.5978	19850058.35	1.230431986	0.18878221
0.483022399	1670893	0.4518	28986543.19	1.221875136	0.199591243
0.50785973	1741783	0.2784	35871403.71	1.242760112	0.215083172
0.491097584	1884802	0.2181	40002765.03	1.267788159	0.20527456
0.477650463	2021100	0.2021	42313719.06	1.294610265	0.218045655
0.502208374	2115037	0.232	45611153.39	1.310191199	0.215241632

Y : Nakit para stoğu / M1
X_1 : Kişi başına reel GSMH
X2 : Yıllık ortalama mevduat faiz oranı
X3 : Tefe genel endeks
X4 : Ortalama vergi oranı + 1 (toplam vergi gelirlerinin GSMH'a oranının 1 ile toplamı)
X5 : Kamu personel harcamalarının toplam kamu harcamalarına oranı

* www.tcmb.gov.tr

φ http://www.gib.gov.tr/fileadmin/user_upload/VI/GBG/Tablo_3.xls.htm