

Çok Değişkenli Çoklu Regresyon Modelinin Minmad Problemi Olarak Modellenmesi ve Global Kriter Yöntemi ile Çözümü

Nimet YAPICI PEHLİVAN *

Ayşen APAYDIN **

ÖZET

İstatistik teorisinin gelişim sürecinde optimizasyon problemleri ile oldukça sık karşılaşılmaktadır. Bu tür problemlerin çözümü için kullanılan teknikler: klasik yöntemler, sayısal yöntemler, değişimsel yöntemler ve matematiksel programlama olarak sınıflandırılmaktadır. Çeşitli alanlarda uygulaması olan istatistik yöntemlerinden bazıları; regresyon analizi, tahmin, istatistiksel hipotez testleri, deney düzenleme ve analizi, veri sınıflandırma ve gruplandırma, zaman serisi analizidir (Arthanari ve Dodge 1981).

Regresyon Analizi 'nde etkin olarak kullanılan En Küçük Kareler Yöntemi (EKK) modelin bilinmeyen parametrelerinin tahmin edicileri için, hatalar bağımsız olduğunda optimal sonuçlar vermektedir. Bu tahmin ediciler, sıfır ortalamalı ve σ^2 varyanslı normal dağılıma uymaktadır. Özellikle normal dağılıma uymayan durumlarda ve uç değerler olduğunda bu yöntem optimallikten çok uzaklaşmaktadır (Cade ve Richards, 1996, Narula ve Wellington 1982).

Charnes, Cooper ve Ferguson (1955)'un birlikte yaptıkları makalede, istatistik için matematiksel programlamanın bir uygulaması ele alınmış ve MİN MAD (Ortalama Mutlak Sapmaların En Küçüklenmesi) Problemi, doğrusal regresyon modelinin çözümü için EKK yöntemine bir alternatif olarak seçilmiştir. Bu makalede, MİN MAD problemi, Doğrusal Programlama (DP) modeli olarak formüle edilmiş ve çözülmüştür (Arthanari ve Dodge 1981, Narula 1987).

*Bu çalışmada, çok değişkenli çoklu regresyon modeli MİN MAD problemi olarak düşünülmüş ve Çok Amaçlı MİN MAD *Problemi elde edilmiştir. Elde edilen bu problem aslında bir çok amaçlı programlama modelidir. Bu nedenle problemin çok amaçlı programlama yöntemlerinden biri olan Global Kriter yöntemi ile çözülmesi amaçlanmıştır.*

Anahtar Kelimeler: MİN MAD Problemi, Çok amaçlı Programlama, Global Kriter Yöntemi.

*Araş. Gör. Selçuk Üniversitesi Fen-Edebiyat Fakültesi İstatistik Bölümü Kampüsü KONYA
nimet@selcuk.edu.tr

**Prof.Dr. Ankara Üniversitesi Fen Fakültesi İstatistik Bölümü Tandoğan ANKARA

1. REGRESYON ÇÖZÜMLEMESİ

Regresyon çözümlemesi, iki ya da daha çok değişken arasındaki ilişkinin yapısını incelemektedir. Regresyon çözümlemesinde, ilgilenilen olayı tanımlayan rasgele değişken “bağımlı (açıklanan) değişken” ve bu olayla ilgili ya da olayı etkileyen değişken ise “bağımsız (açıklayıcı) değişken” olarak tanımlanır (Apaydın, Kutsal, Atakan, 1997).

Bir bağımlı değişken (Y) ve birden fazla bağımsız değişken X arasındaki ilişkiyi inceleyen çoklu doğrusal regresyon modeli,

$$Y_j = \beta_0 + \beta_1 X_{j1} + \beta_2 X_{j2} + \dots + \beta_p X_{jp} + \varepsilon_j, j=1,2,\dots,n \quad (1)$$

eşitliği ile verilmektedir. Eşitlik (2.1) ile verilen modelin matris gösterimi

$$Y = X\beta + \varepsilon \quad (2)$$

biçimindedir. Burada,

Y : $n \times 1$ boyutlu bağımlı değişken vektörünü

X : $n \times (p+1)$ boyutlu tasarım matrisini

β : $(p+1) \times 1$ regresyon katsayıları vektörünü

ε : $n \times 1$ boyutlu hata vektörünü

göstermektedir.

En küçük kareler (EKK) yöntemi kullanılarak,

$$\hat{\beta} = (X^T X)^{-1} X^T Y \quad (3)$$

tahmin edicileri ve

$$\hat{Y} = X\hat{\beta} \quad (4)$$

tahmini regresyon denklemi elde edilir (Apaydın, Kutsal, Atakan 1997).

m bağımlı değişken (Y) ve p bağımsız değişken (X) arasındaki ilişkiyi inceleyen çok değişkenli çoklu regresyon modeli,

$$\begin{aligned} Y_{j1} &= \beta_{01} + \beta_{11} X_{j1} + \beta_{21} X_{j2} + \dots + \beta_{p1} X_{jp} + \varepsilon_{j1} \\ Y_{j2} &= \beta_{02} + \beta_{12} X_{j1} + \beta_{22} X_{j2} + \dots + \beta_{p2} X_{jp} + \varepsilon_{j2} \\ &\vdots \\ Y_{jm} &= \beta_{0m} + \beta_{1m} X_{j1} + \beta_{2m} X_{j2} + \dots + \beta_{pm} X_{jp} + \varepsilon_{jm} \end{aligned} \quad (5)$$

eşitliği ile ifade edilir $j=1,2,\dots,n$ (Johnson ve Wichern 1988, Rencher 1995).

Eşitlik (5)'de verilen model,

$$Y = X\beta + \varepsilon \quad (6)$$

biçiminde matris formunda yazılabilir. Burada,

Y : $n \times m$ boyutlu bağımlı değişken matrisini

X : $n \times (p+1)$ boyutlu tasarım matrisini

β : $(p+1) \times m$ boyutlu regresyon katsayıları matrisini

ε : $n \times m$ boyutlu hata matrisini

göstermektedir. ε hata matrisinin, $E(\varepsilon)=0$ ve $\text{Var}(\varepsilon)=\Sigma$ ile normal dağıldığı varsayılmaktadır.

Bağımlı değişkenler üzerinden alınan gözlemlerden elde edilen EKK tahmin edicileri,

$$\hat{\beta} = (X^T X)^{-1} X^T Y \quad (7)$$

tahmini regresyon denklemi

$$\hat{Y} = X\hat{\beta} \quad (8)$$

ve hatalar

$$\varepsilon = Y - \hat{Y} \quad (9)$$

eşitliklerinden elde edilir (Johnson ve Wichern 1988, Tatlıdil 1992, Apaydın ve ark 1997).

2. ÇOK DEĞİŞKENLİ ÇOKLU REGRESYON MODELİ İÇİN MİN MAD PROBLEMİ

MİN MAD problemi, ilk olarak Charnes, Cooper ve Ferguson (1955)'in yaptığı çalışma ile doğrusal regresyon modelinin çözümü için EKK yöntemine alternatif olarak seçilmiş ve bir doğrusal programlama problemi gibi formüle edilmiştir.

MİN MAD yöntemi, mutlak hataların minimum toplamı (MSAE, LSAE), mutlak sapmaların minimumu (MAD, LAD), mutlak hatalar ya da değerlerin minimumu (MAE, LAV), L_1 -normu olarak da bilinir (Hawley ve Gallagher 1994, Narula 1987, Şanlı 1999).

Bu çalışmada, çoklu doğrusal regresyon modeli için düşünülen MİN MAD problemi temel alınarak, Eşitlik (5)'te verilen çok değişkenli çoklu regresyon modeli için MİN MAD problemi tasarlanmıştır ve

$$\text{Minimize } Z_1 = \sum_{j=1}^n d_{1j} + \sum_{j=1}^n d_{2j}$$

$$\text{Minimize } Z_2 = \sum_{j=1}^n d_{3j} + \sum_{j=1}^n d_{4j}$$

⋮

$$\text{Minimize } Z_m = \sum_{j=1}^n d_{(2m-1)j} + \sum_{j=1}^n d_{(2m)j}$$

$$X\beta_1 + d_1 - d_2 = Y_1 \quad (10)$$

$$X\beta_2 + d_3 - d_4 = Y_2$$

⋮

$$X\beta_m + d_{(2m-1)} - d_{2m} = Y_m$$

$$d_1, d_2, \dots, d_{(2m-1)}, d_{2m} \geq 0$$

β isareti belirtilmemis

çok amaçlı programlama (ÇAP) problemi elde edilmiştir (Apaydın ve Yapıcı 2001).

Eşitlik(10)'da, β değişkenleri β' ve β'' değişkenleri ≥ 0 alınarak $(\beta' - \beta'')$ biçiminde işareti belirtilmiş duruma dönüştürülür. Çözüm sonucunda β' ve β'' değişkenlerinin optimal çözümü elde edilecektir. $d_1, d_2, \dots, d_{(2m-1)}, d_{(2m)}$ sapma değişkenleri, k tek olduğunda pozitif sapmalar,

$$d_k = \begin{cases} Y_k - \hat{Y}_k, & Y_k - \hat{Y}_k \geq 0 \\ 0, & \text{diger durumlarda} \end{cases}$$

k çift olduğunda negatif sapmalar,

$$d_k = \begin{cases} -(Y_k - \hat{Y}_k), & Y_k - \hat{Y}_k < 0 \\ 0 & \text{diger durumlarda} \end{cases}$$

biçiminde ifade edilmektedir (Narula 1987).

3. ÇOK AMAÇLI MİNİMAD PROBLEMİ İÇİN GLOBAL KRİTER YÖNTEMİ ve BİR UYGULAMA

Global kriter yöntemi, problemle ilgili kısıtlar ve amaçlar tanımlandıktan sonra, karar vericinin tercihleriyle ilgili bilgisine ihtiyaç duymayan yöntemlerden birisidir. Karar verici, klasik optimizasyon yöntemlerinde olduğu gibi yöntemin bulunduğu çözümün kabul edilebilir olduğunu varsayar (Evren ve Ülengin 1992).

Global kriter yöntemi algoritması:

1.adım:

Çok değişkenli çoklu regresyon modelinin çok amaçlı MİN MAD problemi olarak tasarlandığı Eşitlik (10)'da verilen çok amaçlı programlama modelinin çözümü için; Z_1, Z_2, \dots, Z_m problemlerinin ayrı ayrı çözülmesiyle $Z_1(x^*), Z_2(x^*), \dots, Z_m(x^*)$ "ideal çözümler"i elde edilir.

2.adım:

Bulunan ideal çözümler yardımıyla $a=1$ alınarak,

$$\begin{aligned} \text{Minimize } & \sum_{i=1}^m \left[\frac{Z_i(x^*) - Z_i(x)}{Z_i(x^*)} \right]^a \\ & X(\beta'_1 - \beta''_1) + d_{1j} - d_{2j} = Y_1 \\ & X(\beta'_2 - \beta''_2) + d_{3j} - d_{4j} = Y_2 \\ & (d, \beta, \beta'') \geq 0 \end{aligned} \quad (11)$$

problemi çözülür. Eşitlik(11)'de, işlem kolaylığı sağlaması ve bir doğrusal programlama problemi elde etmek istenmesi nedeniyle $a=1$ alınmıştır. $a=2$ alınması durumunda, problem karesel programlama problemine dönüşmektedir. $a=3$ alınması durumunda ise problem daha karmaşık olacaktır. Çözüm sonucunda elde edilen (d, β, β'') değişkenleri optimal çözümü vermektedir.

3.1.Uygulama

Dokuz çocuğun; Y_1 - göğüs çevresi, Y_2 - dirsek üstü kol çevresi, X_1 - boy uzunluğu (cm cinsinden), X_2 - yaş (ay olarak) ları aşağıda verilmiştir (Tatlıdil, 1992).

| Gözlem no | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|-----------|------|------|------|------|------|------|------|------|------|
| Y_1 | 58.4 | 59.2 | 60.3 | 57.4 | 59.5 | 58.1 | 58.0 | 55.5 | 59.2 |
| Y_2 | 14.0 | 15.0 | 15.0 | 13.0 | 14.0 | 14.5 | 12.5 | 11.0 | 12.5 |
| X_1 | 80 | 75 | 78 | 75 | 79 | 78 | 75 | 64 | 80 |
| X_2 | 21 | 27 | 27 | 22 | 26 | 26 | 23 | 22 | 22 |

Y bağımlı değişkenleri ve X bağımsız değişkenleri göstermek üzere, çok değişkenli çoklu regresyon modeli hem EKK hem MİN MAD yöntemleri kullanılarak tahmin edilecektir.

EKK yöntemi ile çözüm yapıldığında; Tahmini çok değişkenli çoklu regresyon modeli

$$\hat{Y}_1 = 36.777 + 0.204X_1 + 0.254X_2$$

$$\hat{Y}_2 = -5.526 + 0.104X_1 + 0.347X_2$$

olarak elde edilmiştir. Burada ε hata terimlerine ilişkin değerler Çizelge 1'de verilmiştir (Özdamar 1999).

Çizelge 1. EKK yöntemiyle elde edilen ε hatalarına ait değerler

| | | | |
|--------------------|--------|--------------------|-------|
| ε_{11} | -0.031 | ε_{21} | 3.919 |
| ε_{12} | 0.265 | ε_{21} | 3.357 |
| ε_{13} | 0.753 | ε_{23} | 3.045 |
| ε_{14} | -0.265 | ε_{24} | 3.092 |
| ε_{15} | 0.003 | ε_{25} | 2.288 |
| ε_{16} | -1.193 | ε_{26} | 2.892 |
| ε_{17} | 0.081 | ε_{27} | 2.245 |
| ε_{18} | 0.079 | ε_{28} | 2.236 |
| ε_{19} | 0.515 | ε_{29} | 2.072 |

Y_1 ve Y_2 bağımlı değişkenlerine ait hata kareler toplamları, $\sum_{j=1}^m \varepsilon_{j1}^2 = 2.409$ ve

$\sum_{j=1}^m \varepsilon_{j2}^2 = 73.392$ olarak elde edilmiştir.

Çok amaçlı MİN MAD problemi Global Kriter Yöntemi ile çözüldüğünde:

Eşitlik (10) kullanılarak, verilere göre çok değişkenli çoklu regresyon modeli aşağıdaki çok amaçlı MİN MAD problemine dönüştürülmüştür.

$$\text{Minimize } \sum_{j=1}^9 d_{1j} + \sum_{j=1}^9 d_{2j}$$

$$\text{Minimize } \sum_{j=1}^9 d_{3j} + \sum_{j=1}^9 d_{4j}$$

$$(\beta_{01}' - \beta_{01}'') + 80(\beta_{11}' - \beta_{11}'') + 21(\beta_{21}' - \beta_{21}'') + d_{11} - d_{21} = 58.4$$

$$(\beta_{01}' - \beta_{01}'') + 75(\beta_{11}' - \beta_{11}'') + 27(\beta_{21}' - \beta_{21}'') + d_{12} - d_{22} = 59.2$$

$$(\beta_{01}' - \beta_{01}'') + 78(\beta_{11}' - \beta_{11}'') + 27(\beta_{21}' - \beta_{21}'') + d_{13} - d_{23} = 60.3$$

$$(\beta_{01}' - \beta_{01}'') + 75(\beta_{11}' - \beta_{11}'') + 22(\beta_{21}' - \beta_{21}'') + d_{14} - d_{24} = 57.4$$

$$(\beta_{01}' - \beta_{01}'') + 79(\beta_{11}' - \beta_{11}'') + 26(\beta_{21}' - \beta_{21}'') + d_{15} - d_{25} = 59.5$$

$$(\beta_{01}' - \beta_{01}'') + 78(\beta_{11}' - \beta_{11}'') + 26(\beta_{21}' - \beta_{21}'') + d_{16} - d_{26} = 58.1$$

$$(\beta_{01}' - \beta_{01}'') + 75(\beta_{11}' - \beta_{11}'') + 23(\beta_{21}' - \beta_{21}'') + d_{17} - d_{27} = 58.0$$

$$(\beta_{01}' - \beta_{01}'') + 64(\beta_{11}' - \beta_{11}'') + 22(\beta_{21}' - \beta_{21}'') + d_{18} - d_{28} = 55.5$$

$$(\beta_{01}' - \beta_{01}'') + 80(\beta_{11}' - \beta_{11}'') + 22(\beta_{21}' - \beta_{21}'') + d_{19} - d_{29} = 59.2$$

$$(\beta_{02}' - \beta_{02}'') + 80(\beta_{12}' - \beta_{12}'') + 21(\beta_{22}' - \beta_{22}'') + d_{31} - d_{41} = 14.0$$

$$(\beta_{02}' - \beta_{02}'') + 75(\beta_{12}' - \beta_{12}'') + 27(\beta_{22}' - \beta_{22}'') + d_{32} - d_{42} = 15.0$$

$$(\beta_{02}' - \beta_{02}'') + 78(\beta_{12}' - \beta_{12}'') + 27(\beta_{22}' - \beta_{22}'') + d_{33} - d_{43} = 15.0$$

$$(\beta_{02}' - \beta_{02}'') + 75(\beta_{12}' - \beta_{12}'') + 22(\beta_{22}' - \beta_{22}'') + d_{34} - d_{44} = 13.0$$

$$(\beta_{02}' - \beta_{02}'') + 79(\beta_{12}' - \beta_{12}'') + 26(\beta_{22}' - \beta_{22}'') + d_{35} - d_{45} = 14.0$$

$$(\beta_{02}' - \beta_{02}'') + 78(\beta_{12}' - \beta_{12}'') + 26(\beta_{22}' - \beta_{22}'') + d_{36} - d_{46} = 14.5$$

$$(\beta_{02}' - \beta_{02}'') + 75(\beta_{12}' - \beta_{12}'') + 23(\beta_{22}' - \beta_{22}'') + d_{37} - d_{47} = 12.5$$

$$(\beta_{02}' - \beta_{02}'') + 64(\beta_{12}' - \beta_{12}'') + 22(\beta_{22}' - \beta_{22}'') + d_{38} - d_{48} = 11.0$$

$$(\beta_{02}' - \beta_{02}'') + 80(\beta_{12}' - \beta_{12}'') + 22(\beta_{22}' - \beta_{22}'') + d_{39} - d_{49} = 12.5$$

$$d_{11}, d_{21}, \dots, d_{19}, d_{29} \geq 0$$

$$d_{31}, d_{41}, \dots, d_{39}, d_{49} \geq 0$$

$$\beta_{01}', \beta_{01}'', \beta_{11}', \beta_{11}'', \beta_{21}', \beta_{21}'' \geq 0$$

$$\beta_{02}', \beta_{02}'', \beta_{12}', \beta_{12}'', \beta_{22}', \beta_{22}'' \geq 0$$

Global kriter yöntem ile çözüm yapıldığında çok değişkenli çoklu regresyon modeli,

$$\hat{Y}_1 = 36.100 + 0.200X_1 + 0.300X_2$$

$$\hat{Y}_2 = -6.8571 + 0.107X_1 + 0.500X_2$$

biçiminde elde edilmiştir. Burada, ε_{1j} ve ε_{2j} değerleri

$$\varepsilon_{1j} = d_{1j} - d_{2j}$$

ve

$$\varepsilon_{2j} = d_{3j} - d_{4j} \quad (j=1,2,\dots,9)$$

eşitliklerinden elde edilebilir. Bu değerler Çizelge2'de gösterilmiştir.

Çizelge 2. Çok amaçlı MİN MAD probleminin Global Kriter yöntemiyle elde edilen hatalarına ait değerler

| | | | |
|--------------------|------|--------------------|---------|
| ε_{11} | 0 | ε_{21} | 1.7857 |
| ε_{12} | 0 | ε_{21} | 0.3214 |
| ε_{13} | 0.5 | ε_{23} | 0 |
| ε_{14} | -0.3 | ε_{24} | 0.8214 |
| ε_{15} | -0.2 | ε_{25} | -0.6071 |
| ε_{16} | -1.4 | ε_{26} | 0 |
| ε_{17} | 0 | ε_{27} | -0.1786 |
| ε_{18} | 0 | ε_{28} | 0 |
| ε_{19} | 0.5 | ε_{29} | -0.2143 |

Y_1 ve Y_2 bağımlı değişkenlerine ait hata kareler toplamları $\sum_{j=1}^m \varepsilon_{j1}^2 = 2.59$

ve $\sum_{j=1}^m \varepsilon_{j2}^2 = 4.413$ olarak elde edilmiştir.

4.SONUÇ

Bütün regresyon modellerinde olduğu gibi, çok değişkenli çoklu regresyon modelinde de amaç, ε hatalarını minimum yapacak şekilde β parametrelerinin tahmin edilmesidir. Bu nedenle, çok değişkenli çoklu regresyon modelinde hataların minimum yapılması amaçlanmıştır ve çözüm için iki yöntem kullanılmıştır. Bunlar

- En küçük kareler yöntemi
- MİN MAD yöntemi

dir.

Her iki yöntemle elde edilen sonuçlar karşılaştırıldığında, genel olarak MİN MAD yönteminden elde edilen tahmini β parametrelerinin EKK yöntemine göre daha minimum olduğu görülmektedir.

Çizelge1 ve Çizelge2 karşılaştırıldığında, genel olarak MİN MAD yönteminden elde edilen ε hataları değerlerinin EKK yöntemine göre daha minimum olduğu görülmektedir. Bu da istenen sonuçtur.

EKK yönteminde hataların mutlaka normal dağıldığı varsayımları vardır. Bu varsayımlar bozulduğunda ve çoklu bağlantı durumunda MİN MAD yöntemi etkin çözümler vermektedir.

Bu çalışmada, çok değişkenli çoklu regresyon modeli MİN MAD problemi olarak oluşturulmuş ve bir ÇAP problemi elde edilmiştir. Ele alınan veriler için çok değişkenli çoklu regresyon modeli hem EKK hem de çok amaçlı MİN MAD yöntemi ile tahmin edilmiştir. Bunun sonucunda, EKK'dan elde edilen hata kareler toplamı Y_1 bağımlı değişkeni için 2.409 ve Y_2 bağımlı değişkeni için 73.392 dir. MİN MAD yönteminden elde edilen hata kareler toplamı ise Y_1 bağımlı değişkeni için 2.590 ve Y_2 bağımlı değişkeni için 4.413 tür. Dolayısıyla MİN MAD yönteminin hata kareler toplamının daha küçük olduğu söylenebilir.

KAYNAKLAR

1. APAYDIN, A., KUTSAL, A., ATAKAN, C.(1994), *Uygulamalı İstatistik*, Ankara: Kültür Kitap ve Yayınevi
2. APAYDIN, A., YAPICI, N. (2001), *Minmad Problemi Olarak Modellenen Çok Değişkenli Çoklu Regresyon Modelinin Çözümü İçin Çok Amaçlı Doğrusal Programlama Yaklaşımı*, 2. İstatistik Kongresi, Bildiriler Kitabı, 66-69, 2-6 Mayıs, Antalya.
3. ARTHANARI, T.S., DODGE, Y.(1981), *Mathematical Programming in Statistics*, John Wiley and Sons Inc.
4. CADE, B.S., RICHARDS, J.D. (1996), *Permutation Tests for Least Absolute Deviation Regression*, *Biometrics*, 52, 886-902.
5. EVREN, R., ÜLENGİN, F.(1992), *Yönetimde Çok Amaçlı Karar Verme*, İ.T.Ü. Matbaası.
6. HAWLEY, R.W., GALLAGHER, N.C. (1994), *On Edgeworth's Method for Minimum Absolute Error Linear Regression*, *IEEE Transactions on Signal Processing*, 42, 2045-2054.
7. JOHNSON, R.A., WICHERN, D.W.(1988), *Applied Multivariate Statistical Analysis*, Prentice-Hall Inc.
8. NARULA, S.C., (1987), *The Minimum Sum of Absolute Errors Regression*, *Journal of Quality Technology*, 19, 1, 37-44.
9. NARULA, S.C., WELLINGTON, J.F.(1982), *The Minimum Sum of Absolute Errors Regression: A State of The Art Survey*, *International Statistical Review*, 50, 317-326.
10. ÖZDAMAR, K.(1999), *İstatistiksel Programlar İle Veri Analizi-2*, Kaan Kitabevi, Eskişehir.

11. RENCHER, A.C.(1995), *Methods of Multivariate Analysis*, John Wiley and Sons Inc.
12. ŞANLI, K.(1999), *MİN MAD Yöntemiyle Rasgele Bloklar Model Denklemindeki Parametrelerin Tahmini*, Yüksek Lisans Tezi (basılmamış), Ankara Üniversitesi Fen Bilimleri Enstitüsü, Ankara.
13. TATLIDİL, H.(1992), *Uygulamalı Çok Değişkenli İstatistiksel Analiz*, Ankara.

Global Criteria Method for Solving Multivariate Multiple Regression Model Modelled as MINMAD Problem

ABSTRACT

In the development of the theory underlying statistical methods, one is often faced with an optimization problem. The techniques for solving such problems can be classified as classical, numerical, variational methods and mathematical programming. Regression analysis, estimation, testing of statistical hypotheses, design and analysis experiments, data classification and grouping, time series analysis are most of the major statistical methods that have found many applications in various fields (Arthanari and Dodge 1981).

The least squares method has dominated the statistical literature. This method is optimal and results in the estimators of the unknown parameters of the model if the errors are independent and follow a normal distribution with mean zero and a variance σ^2 . These estimators are very far from the optimal in many nonnormal situations and when the extreme values exist (Narula and Wellington, 1982).

The fundamental paper by Charnes, Cooper and Ferguson (1955) introduced the application of mathematical programming to statistics. As an alternative to the least squares method to linear regression, MINMAD (Minimizing Mean Absolute Deviations) problem is chose. In that paper, MINMAD problem is formulated and solved as Linear Programming (Arthanari and Dodge 1981, Narula and Wellington, 1982).

In this study, Multivariate Multiple Regression is considered as MINMAD Problem and obtained Multi-objective MINMAD Problem. Actually, this problem is Multi-objective Programming Problem and then solved by using the Global Criteria Method, that one of the Multi-objective Programming Problem Methods.

Key Words: *MİN MAD (Minimizing Mean Absolute Deviations) Problem, Multi-objective Programming Problem, Global Criteria Method.*