

**ÇUKUROVA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

YÜKSEK LİSANS TEZİ

HÜLYA TOSUN

**MATEMATİKSEL PROGRAMLAMANIN İSTATİSTİKTE
UYGULAMALARI**

**T.C. YÜKSEKÖĞRETİM KURULU
DOKÜMANTASYON MERKEZİ**

MATEMATİK ANABİLİM DALI

ADANA, 20002

119824

ÇUKUROVA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

MATEMATİKSEL PROGRAMLAMANIN İSTATİSTİKTE
UYGULAMALARI

HÜLYA TOSUN
YÜKSEK LİSANS TEZİ
MATEMATİK ANABİLİM DALI

Bu tez 18/09/2002 Tarihinde Aşağıdaki Jüri Üyeleri Tarafından
Oybirliği/Oyçokluğu İle Kabul Edilmiştir.

İmza.....

İmza.....

İmza.....

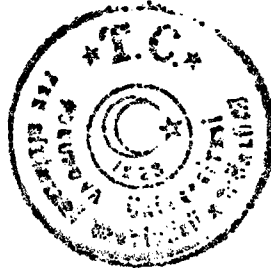
Doç.Dr. Selahattin KAÇIRANLAR
DANIŞMAN

Doç.Dr. Rızvan EROL
ÜYE

Doç.Dr. Hamza EROL
ÜYE

Bu tez Enstitümüz Matematik Anabilim Dalında hazırlanmıştır.

Kod No: 2054



Prof. Dr. Fikri AKDENİZ

Enstitü Müdürü

İmza ve Mühür

Not: Bu tezde kullanılan özgün ve başka kaynaktan yapılan bildirişlerin, çizelge, şekil ve fotoğrafların kaynak gösterilmeden kullanımı, 5846 sayılı Fikir ve Sanat Eserleri Kanunundaki hükümlere tabidir.

İÇİNDEKİLER

Sayfa no

ÖZ.....	I
ABSTRACT.....	II
TEŞEKKÜR.....	III
TABLolar DİZİNİ.....	IV
ŞEKİLLER DİZİNİ.....	V
1. GİRİŞ.....	1
1.1. Matematiksel Programlama Problemi.....	3
2. LİNEER REGRESYON ANALİZİ.....	5
2.1. Mutlak Sapmaların Ortalamasını Minimum Yapma (MINMAD)	
Regresyon.....	6
2.2. Mutlak Sapmaların Maksimumunu Minimum Yapma	
(MINMAXAD) Regresyon.....	32
2.3. Sapmalar Arasındaki Farkın Mutlak Değerinin Toplamını Minimum	
Yapma (MINSADBED) Regresyon.....	33
2.4 Mutlak Sapmalar Arasındaki Mutlak Farkların Toplamını Minimum	
Yapma (MINSADBAD) Regresyon Analizi.....	37
3. MATEMATİKSEL PROGRAMLAMA VE REGRESYON ANALİZİ.....	38
3.1. MINMAD Regresyon.....	38
3.2. MINMAD Regresyon İçin Simpleks Yöntemi.....	46
3.3 Genel Lineer Programlama Problemi.....	65

3.4. Minmad Regresyonda Dualite.....	68
3.5. Lineer Programlamada Dualite.....	71
3.6. MINMAD İçin Alternatif Bir Formülasyon.....	79
3.7. MINMAD Kullanarak Yansız Tahminler Bulma.....	81
3.8 Maksimum-Likelihood ve MINMAD Tahminleri.....	85
3.9 Mutlak Sapmaların Maksimumunu Minimum Yapma (MINMAXAD).....	85
3.10. MINMAXAD Kullanarak Yansız Tahminler Elde Etme.....	93
3.11. En Küçük Kareler ve MINMAD Regresyonun Konveks Kombinasyonu	94
KAYNAKLAR.....	117
ÖZGEÇMİŞ.....	118

ÖZ

YÜKSEK LİSANS TEZİ

**MATEMATİKSEL PROGRAMLAMANIN İSTATİSTİKTE
UYGULAMALARI**

HÜLYA TOSUN

ÇUKUROVA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
MATEMATİK ANABİLİM DALI

Danışman: Doç. Dr. Selahattin KAÇIRANLAR
Yıl: 2002, Sayfa: 118
Jüri : Doç. Dr. Selahattin KAÇIRANLAR
Doç. Dr. Rızvan EROL
Doç. Dr. Hamza EROL

Bu çalışmada en küçük karelere alternatif olarak önerilen lineer regresyon tahmin edicileri incelenmiştir. Matematiksel programlama yöntemleri kullanılarak bu tahmin edicilerin hesaplanması ve böylece istatistikte matematiksel programlamanın uygulamaları ele alınmıştır.

Anahtar kelimeler: Matematiksel programlama, lineer regresyon, MINMAD regresyon, MINSADBED regresyon, simpleks yöntemi.

ABSTRACT

MSc THESIS

**THE APPLICATIONS OF MATHEMATICAL PROGRAMMING
IN STATISTICS**

HÜLYA TOSUN

DEPARTMENT OF MATHEMATICS
INSTITUTE OF NATUREL AND APPLIED SCIENCES
UNIVERSITY OF ÇUKUROVA

Supervisor: Assoc. Prof. Dr. Selahattin KAÇIRANLAR
Year: 2002, Pages: 118
Jury : Assoc. Prof. Dr. Selahattin KAÇIRANLAR
Assoc. Prof. Dr. Rızvan EROL
Assoc. Prof. Dr. Hamza EROL

In this study, linear regression estimators which can be proposed as alternative to the least squares estimator are examined. Computing these estimators and therefore application of mathematical programming in statistics are considered by using mathematical programming procedures.

Key Words: Mathematical programming, linear regression. MINMAD regression, MINSADBED regression, simplex procedure.

TEŞEKKÜR

Bilim yolunda böyle bir adım atarken bana büyük destek veren, sonsuz sevgi ve saygı duyduğum ve her zaman örnek aldığım değerli hocam Doç. Dr. Nedret BİLLOR 'a; beni her zaman iyiyi ve doğruyu bulmaya yönlendiren, bu tezi hazırlama aşamasında ilgisini ve yardımlarını esirgemeyen değerli danışmanım Doç. Dr. Selahattin KAÇIRANLAR 'a sonsuz teşekkürlerimi sunarım

Ayrıca çalışmalarım sırasında her zaman yanımda olan aileme ve destek ve yardımlarından dolayı İbrahim Karabatak 'a teşekkür ederim.



ÇİZELGELER DİZİNİ

Çizelge 2.1 Kuzey ve Merkez Amerika'daki 14 ülkeye ait doğum oranı ve şehre ait yüzdelik.....	16
Çizelge 2.2 Yangın kazaları ile ilgili veri.....	24
Çizelge 3.1 α_{ij} leri ve bazdaki vektörleri gösteren ilk çizelge.....	53
Çizelge 3.2 Örnek 3.2.1 için oluşturulan ilk çizelge.....	56
Çizelge 3.3 a_2 ile a_8 in yer değiştirmesi ile elde edilen çizelge.....	57
Çizelge 3.4 a_1 ile a_5 in yer değiştirmesi ile elde edilen çizelge.....	57
Çizelge 3.5 a_{10} ile a_6 nın yer değiştirmesi ile elde edilen çizelge.....	58
Çizelge 3.6 a_4 ün a_3 ile yer değiştirmesinden elde edilen çizelge.....	58
Çizelge 3.7 Örnek 3.2.2 için oluşturulan ilk çizelge.....	62
Çizelge 3.8 a_2 ile a_5 in yer değiştirmesi ile elde edilen çizelge.....	64
Çizelge 3.9 a_3 ile a_1 in yer değiştirmesi ile elde edilen çizelge.....	64

ŞEKİLLER DİZİNİ

Şekil 2.1 $f_1(\beta)= 3-\beta $, $f_2(\beta)= 1-\beta $, $f_3(\beta)= 4-2\beta $ ve $\sum f_i(\beta)$ fonksiyonlarını gösteren grafik.....	8
Şekil 2.2 LS, LS (T/T hariç) ve LAD regresyon doğrularının karşılaştırılması.....	17
Şekil 2.3 $f_1(\beta)= 3-\beta $, $f_2(\beta)= 1-\beta $, $f_3(\beta)= 4-2\beta $, $\max f_i(\beta)=g(\beta)$ yı gösteren grafik.....	32
Şekil 2.4 Şekil 2.1.3 den farklı bir lineer konveks fonksiyon grafiği.....	33
Şekil 3.1 Yarı konveks olan fakat konveks olmayan bir fonksiyon.....	102
Şekil 3.2 Yarı konveks olan fakat aşıkarak yarı konveks olmayan bir fonksiyon.....	104

1. GİRİŞ

İstatistik, verileri anlamlandırma bilimidir ve insan aktivitelerinin olduğu hemen her yerde kaçınılmaz bir rol oynar. Verileri anlamlandırmak için farklı veri tiplerine uygun farklı istatistiksel yöntemler geliştirilmiştir. Regresyon analizi, tahmin problemi, hipotez testleri, örnekleme teorisi ve verilerin sınıflandırılması ve gruplandırılması çeşitli alanlarda birçok uygulamaya sahip temel istatistik yöntemlerinin bazılarıdır. Bununla birlikte tüm istatistiksel işlemler uygun formülize edilmiş optimizasyon problemlerinin çözümleridir.

Birçok birimin her biri için değişkenlerin değerleri ölçüldüğünde verilerin genel yapısı meydana gelir. Örneğin; endüstriyel bir deneyde bir kimyasalın 20 birimi için reaksiyon zamanı, reaksiyon sıcaklığı ve çıkan ürün miktarı kaydedilir. Bu veriler kullanılarak bir değişkenin (örneğin çıkan ürün miktarının) diğer değişkenlerle (reaksiyon zamanı ve reaksiyon sıcaklığı ile) arasında nasıl bir ilişki olduğu ile ilgilenilir. Lineer regresyon analizi genelde böyle verileri incelemek için kullanılan bir istatistiksel yöntemdir. Regresyon, veri analizi için mevcut istatistiksel yöntemler arasında en kapsamlı ve ayrıntılı yöntemlerden biridir. Regresyon teorisi *bağımlı değişken* olarak adlandırılan bir ya da daha fazla değişkenin *bağımsız değişkenler* olarak adlandırılan diğer değişkenler bazında tahminine dayanır. Bu tür problemlerle deneysel bilimin ve teknolojinin hemen her alanında karşılaşılabilir.

Regresyon analizi yapılması, gözlenen veriyi tanımlamak için stokastik bir model seçilmesi, kullanışlı verilerden sonuç çıkarılması veya hipotez testi yapılması problemlerinde bilinmeyen parametreler üzerinde verilen kısıtlamalara göre minimum ya da maksimum yapılmak istenen bir amaç fonksiyonu seçmeliyiz. Klasik optimizasyon metodları diferansiyel hesaplamaya dayanır ve çok kısıtlıdır. Klasik optimizasyon metodlarını istatistiksel çalışmalarda ortaya çıkan birçok durumda uygulamak zordur ya da olanaksızdır. Son otuz yıl boyunca geniş uygulamaya sahip diğer optimizasyon teknikleri bulunmaya çalışılmıştır. İstatistikte etkili uygulamaya sahip tekniklerden biri matematiksel programlamadır. İstatistiğin çeşitli alanlarında matematiksel programlamanın uygulamaları ile ilgili birçok çalışma yapılmış ve halen de yapılmaktadır.

Bu konudaki temel çalışma Charhes,A ; Cooper,W.W. ve Ferguson,R.O. (1955) tarafından yapılan istatistikte matematiksel programlamanın uygulamalarıdır. Bunlar lineer regresyonda en küçük kareler (EKK)'e alternatif olarak sapmaların mutlak değerinin toplamını minimize etmeyi (MINMAD) seçtiler ve MINMAD problemi ve lineer programlama problemi arasındaki denklığı gösterdiler. İstatistikte matematiksel programlamanın diğer uygulama alanları aynı anda gelişti. Lee, T.C. ; Judge, G.G. ve Zallner, A.(1968) maksimum olabilirlik tahminde, Rao, J.N.K.(1979) örneklemede, Liittschwager,J.M. ve Wang, C.(1978) küme analizinde, Meeks, H.D. ve Francis, R.L. (1973) ve Pukelsheim, F. (1978) hipotez testinde matematiksel programlamayı kullandılar.

Bu çalışmanın amacı da istatistikte matematiksel programlamanın uygulamaları üzerindeki elde edilebilir sonuçların çoğunu bir araya getirmek, yeni uygulama alanları bulmaya çalışmak ve böylece gerek teorik gerekse uygulamalı istatistik çalışan araştırmacılara ışık tutmaktır.

Bu çalışma üç ana bölümden oluşmaktadır. Çalışmanın birinci bölümü olan giriş bölümünde regresyon ve matematiksel programlama problemleri hakkında kısaca bilgi verilmiştir ve tezde anlatılan konular tanıtılmıştır. İkinci bölümde regresyon analizi ele alınmış ve basit tek bağımsız değişkenli durumda farklı optimizasyon kriterleri tanıtılmıştır. Mutlak sapmaların ortalamasının minimum yapılması (MINMAD) (bir başka adıyla LAD) regresyon hakkında bilgi verilmiş ve ardından çoklu regresyonda MINMAD bir örnekle açıklanmıştır. MINMAXAD ve MINSADBED regresyonlara değinilmiştir. Daha sonra MINSADBAD regresyon hakkında kısaca bilgi verilmiştir.

Üçüncü bölümde regresyon analizinde matematiksel programlamanın kullanımı ele alınmıştır. İlk olarak MINMAD regresyonun bir matematiksel programlama problemi olarak nasıl ele alındığı anlatılmıştır. Ayrıca, “en uygun sonuca ardışık yaklaşım yoluyla ulaşmak” şeklinde tanımlanan simpleks yöntemi MINMAD regresyon için geliştirilmiştir. MINMAD problemi için özel bir algoritma tanımlanmıştır. Bir başlangıç çözümü ile başlayıp ardışık iterasyonlarla optimal çözüm elde edilmiştir. Daha sonra genel lineer programlama problemi hakkında bilgi

verilmiştir. MINMAD regresyonda ve lineer programlamada modelin çözümüne kolaylık sağlayan dualiteden bahsedilmiştir.

Hatalar simetrik dağılıma sahip ve $E(\varepsilon)=0$ olsa bile MINMAD probleminin bir çözümü β nın yansız tahminini vermeyebilir. Bu yüzden eklenen kısıtlamalar ile birlikte MINMAD için alternatif bir formülasyondan bahsedilmiştir. ε un simetrik dağılıma sahip ve $E(\varepsilon)=0$ olduğunun kabulü ile MINMAD kullanarak yansız tahminler elde edilmiştir. Ardından maksimum likelihood ve MINMAD tahminlerinden bahsedilmiştir. Ayrıca farklı lineer programlama formülasyonları kullanılarak MINMAXAD kriteri ve problemin geometrik özellikleri ele alınmıştır. Son olarak da MINMAD ve en küçük kareler arasında bir uyuşma olarak bu ikisinin bir konveks kombinasyonu ele alınmıştır.

1.1 Matematiksel Programlama Problemi

Bir matematiksel programlama problemi

$$f(x^*) \leq f(x) \quad (\text{tüm } x \in F \text{ için}) \quad (1.1.1)$$

olacak şekilde bir x^* bulma şeklinde ifade edilebilir. Burada F, uygun küme olarak adlandırılır. F, R^n in bir alt kümesi olması gerekmeyen herhangi bir kümedir. f, F üzerinde tanımlı reel değerli bir kümedir. Bu şekilde bir x^* in bulunuyor olması genelde garanti değildir. Bununla birlikte, bu problemin çözümü, böyle bir x^* bulmak veya olmadığını göstermek şeklinde anlaşılabilir.

F, R^n in bir alt kümesi olarak verildiğinde aşağıdaki gibi ifade edilen bir lineer olmayan programlama problemine sahip olunur.

$$F = \{x \mid g_i(x) \leq b_i, i = 1, \dots, m; x \geq 0\} \text{ iken}$$

$$f(x^*) \leq f(x) \text{ olacak şekilde bir } x^* \in F \text{ nin} \quad (1.1.2)$$

bulunması , (tüm $x \in F$ için)

Burada f , amaç fonksiyonudur. $g_i(x) \leq b_i$ ler, g_i F de tanımlı reel değerli fonksiyon iken problemin kısıtlamalarıdır. $x \geq 0$ kısıtlaması da negatif olmayan kısıtlama olarak bilinir. Bu problem genellikle aşağıdaki gibi ifade edilir:

$$\text{Min } f(x)$$

$$g_i(x) \leq b_i \quad , \quad i = 1, \dots, m$$

$$x \geq 0$$



2. LİNEER REGRESYON ANALİZİ

İstatistik, verileri anlamlandırma bilimidir. Verilerin ne anlama geldiğini belirtmek amacıyla, verilerden uygun bilgiler çıkarmak için farklı veri tiplerine uygun farklı istatistiksel teknikler geliştirilmiştir. Birçok birimin her biri için değişkenlerin değeri ölçüldüğünde verilerin genel şekli meydana gelir. Örneğin; tıbbi bir çalışmada 100 birimlik bir grubun ağırlığı, boyu, ve kan basıncı kaydedilebilir. Burada dört değişkenin değeri 100 birimin her biri için ölçülür. Böyle veriler bir değişkenin diğerine nasıl bağlı olduğunu görme amacıyla toplanır. Yani kan basıncı, ile kişinin yaşı, boyu ve ağırlığı arasında nasıl bir ilişki olduğu ile ilgilenilebilir. Lineer regresyon analizi genelde böyle veriler için kullanılan bir istatistiksel tekniktir.

Veri analizi için, mevcut istatistiksel metotlar arasında en kapsamlı ve ayrıntılı metotlardan biri de regresyondur. Regresyon teorisi *bağımlı değişken* olarak adlandırılan bir ya da daha fazla değişkenin *bağımsız değişkenler* olarak adlandırılan diğer değişkenler bazında tahminine dayanır. Bu tarz problemlerle deneysel bilimin ve teknolojinin hemen her dalında karşılaşılabilir.

Model, bağımsız değişkenler ile bağımlı değişken arasında lineer bir ilişki olduğu varsayımı ile açıklandığında bir lineer regresyon modeline sahip olunur. Aksi halde lineer olmayan bir regresyon modeli oluşturulur.

Regresyon probleminin klasik yaklaşımında amaç; bağımlı değişkenin tahmin edilmiş değerinin gözlenen değerinden sapmalarının karelerinin toplamını minimum yapmaktır. Bu yöntem *en küçük kareler yöntemi* olarak bilinir.

Kullanılan bir başka yöntem; bağımlı değişkenin gözlenen değerinin tahmin edilen değerden sapmalarının mutlak değerinin ortalamasını minimum yapmaktır. Bu problem *L_1 -norm minimizasyon* olarak bilinir.

Literatürde ele alınan üçüncü bir yöntem de bağımlı değişkenin tahmin edilmiş değeri ile gözlenen değeri arasındaki mutlak sapmaların maksimumunu minimum yapan Chebyshev kriteridir.

Lineer regresyon doğrularının bulunması için ikiden fazla kriter önerilir. Sonunda ,bağımsız değişkenlerin mevcut kümesinden uygun şekilde seçilen bir alt küme için problem ele alınır.

2.1. Mutlak Sapmaların Ortalamasını Minimum Yapma (MINMAD) Regresyon

MINMAD metodu, bir diğer adıyla mutlak sapmaların en küçüğünü bulma (LAD) metodu EKK metodundan hemen hemen 50 yıl önce, 1757 de Roger Joseph Bascovich tarafından tanımlandı.

Metod, dünyanın şeklini tahmin etmeyi amaçlayan tutarlı olmayan ölçümleri uzlaştırmak için tasarlandı. Pierre Simon Laplace metodu 30 yıl sonra kabul etti. Bu metod ara sıra kullanıldı, fakat en küçük kareler metodu ile kullanımı azaldı.

En küçük karelerin yaygın olarak kullanılması kısmen hesaplama kolaylığına ve teorisinin Gauss ve Laplace tarafından geliştirilmesine bağlıdır. Bugün ise hesaplamalar sınırlı değil ve teorik gelişmeler MINMAD ı içeren çeşitli alternatif metodlara götürmektedir.

$$Y = \beta_0 + \beta_1 X + \varepsilon$$

basit lineer regresyon modelini ele alalım.

$$\frac{1}{n} \sum_{i=1}^n |Y_i - \beta_0 - \beta_1 X_i|$$

ifadesi minimum olacak şekilde β_0 ve β_1 bulmak istiyoruz. Bu ifade; gözlemlerin, bağımlı değişkenin tahmin edilmiş değerlerinden mutlak sapmalarının ortalaması olarak bilinir. Bunu minimum yapmak demek

$$\sum_{i=1}^n |Y_i - \beta_0 - \beta_1 X_i|$$

ifadesini minimum yapmak demektir. Bu ise mutlak sapmaların toplamıdır.

Önce, mutlak sapmaların toplamını minimum eden β_0 ve β_1 'i bulmak için, (X_0, Y_0) verilen bir değer olmak üzere, $Y_0 = \beta_0 + \beta_1 X_0$ kısıtlaması ile problemi göz önüne alalım. Bu kısıtlama analizi basitleştirir ve yaklaşımı anlamamızı sağlar. Sonra bu kısıtlama olmaksızın problemi göz önüne alalım.

(X_0, Y_0) verilsin. Verilen veriye aşağıdaki dönüşümü uygulayalım:

$$x_i = X_i - X_0$$

$$y_i = Y_i - Y_0$$

Şimdi problem

$$\sum_{i=1}^n |y_i - \beta x_i|$$

ifadesini minimum yapan β bulma problemi olarak ifade edilebilir. Bu problemi incelemeden önce herhangi bir i için $|y_i - \beta x_i|$ nin minimumunu göz önüne alalım.

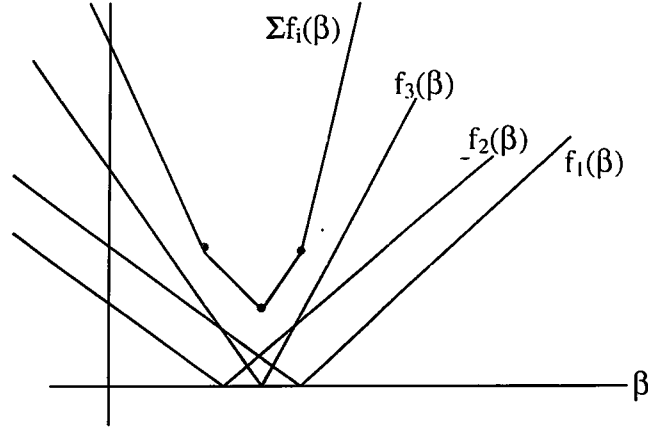
$f(\beta) = |y_i - \beta x_i|$, $\left(\frac{y_i}{x_i}, 0\right)$ da minimum olan ve $-|x_i|$ ve $|x_i|$ eğimlerine sahip olan iki doğrudan oluşur.

i	x_i	y_i
1	1	3
2	1	1
3	2	4

verilerini göz önüne alalım. $|y_i - \beta x_i|$ nin ve $\sum |y_i - \beta x_i|$ nin grafiğini çizelim.

$$f_1(\beta) = |3 - \beta|, \quad f_2(\beta) = |1 - \beta|, \quad f_3(\beta) = |4 - 2\beta|$$

$\sum f_i(\beta) = |3 - \beta| + |1 - \beta| + |4 - 2\beta|$ olup bir parçalı lineer konveks fonksiyondur.



Şekil 2.1 $f_1(\beta) = |3 - \beta|$, $f_2(\beta) = |1 - \beta|$, $f_3(\beta) = |4 - 2\beta|$ ve $\sum f_i(\beta)$ fonksiyonlarını gösteren grafik

Sonuç 2.1.1 :

$f(\beta) = \sum_{i=1}^n |y_i - \beta x_i|$ fonksiyonu verilen (x_i, y_i) için $(i = 1, 2, \dots, n)$ parçalı

lineer konveks fonksiyondur.

İspat: $\beta' < \beta''$, $0 \leq \lambda \leq 1$ ve $\beta = \lambda\beta' + (1 - \lambda)\beta''$ için $f(\beta) \leq \lambda f(\beta') + (1 - \lambda)f(\beta'')$ olduğunu göstermeliyiz.

$$f_i(\beta) = |y_i - \beta x_i| \text{ olsun.}$$

$$f_i(\beta) = f_i(\lambda\beta' + (1 - \lambda)\beta'') = |y_i - \lambda\beta'x_i - (1 - \lambda)\beta''x_i|$$

$$= |\lambda(y_i - \beta'x_i) + (1 - \lambda)(y_i - \beta''x_i)|$$

$$\leq \lambda|y_i - \beta'x_i| + (1 - \lambda)|y_i - \beta''x_i|$$

$$= \lambda f_i(\beta') + (1 - \lambda)f_i(\beta'').$$

$f_i(\beta)$ nın konveks olduğu gösterildi. İki konveks fonksiyonun toplamının konveks olması kullanılırsa $f(\beta)$ nın da konveks olduğunu söyleyebiliriz.

Sonuç 2.1.2 : $f(\beta)$ fonksiyonu aşağıdaki özelliklere sahiptir.

1. Sol uçtaki lineer parçanın eğimi $-\sum |x_i|$ ve benzer şekilde sağ uçtaki lineer parçanın eğimi $\sum |x_i|$ dir.

2. $\left(\frac{y_i}{x_i}\right)$, $f_i(\beta)$ nin minimum noktası olduğu zaman $f(\beta)$ nin grid (ardışık parçaların kesim noktası) noktaları $\left(\frac{y_i}{x_i}\right)$ formundadır.

$(i_1, i_2, \dots, i_n)$, $\frac{y_{i_1}}{x_{i_1}} \leq \dots \leq \frac{y_{i_n}}{x_{i_n}}$ olacak şekilde ise $f(\beta)$ nin eğimi her $\beta_{(k)}$ grid noktasında $2|x_{i_k}|$ kadar artar. Burada , $\beta_{(k)} = \frac{y_{i_k}}{x_{i_k}}$ dır.

Böylece yukarıdaki sonuçtan, $f(\beta)$ nin minimumunu bulmak için bir algoritmaya sahip oluruz :

$\beta_{(r)}$ için minimum

$$-\sum_{i=1}^n |x_i| + 2 \sum_{k=1}^{r-1} |x_{i_k}|$$

negatif fakat

$$-\sum_{i=1}^n |x_i| + 2 \sum_{k=1}^r |x_{i_k}|$$

negatif olmayan bir değer ise

$$\hat{\beta}_1 = \frac{y_r}{x_r} \text{ ve } \hat{\beta}_0 = Y_0 - \frac{y_r}{x_r} X_0$$

olarak elde edilir.

$$-\sum_{i=1}^n |x_i| + 2\sum_{k=1}^r |x_{i_k}| = 0 \text{ ise } \beta_{(r)} \leq \beta \leq \beta_{(r+1)} \text{ aralığındaki tüm } \beta \text{ lar için}$$

$f(\beta)$ optimaldir.

Bu durumda eşit olasılık ile $\beta_{(r)}$ ya da $\beta_{(r+1)}$ seçebiliriz.

Örnek 2.1.1 :

i	Y_i	X_i	y_i	x_i	y_i / x_i	rank
1	22	50	-8.4	-12.5	0.672	8
2	25	54	-5.4	-8.5	0.635	6
3	34	56	3.6	-6.5	-0.554	2
4	28	59	-2.4	-3.5	0.686	9
5	26	60	-4.4	-2.5	1.760	12
6	32	61	1.6	-1.5	-1.067	1
7	30	62	-0.4	-0.5	0.800	10
8	30	65	-0.4	2.5	-0.160	4
9	28	67	-2.4	4.5	-0.534	3
10	34	71	3.6	8.5	0.423	5
11	36	71	5.6	8.5	0.659	7
12	40	74	9.6	11.5	0.835	11
Σ	365	750				

Verilen veriyi göz önüne alalım. Y_0, X_0 ; başlangıç değerleri $\bar{Y}$ ve $\bar{X}$ olsun.

$\bar{Y} = 30,4$ ve $\bar{X} = 62,5$ dir. Böylece

$$\sum_{i=1}^{12} |x_i| = 71 \text{ olur.}$$

Soldan sağa farklı lineer parçaların eğimleri S_r negatif olmayan oluncaya kadar

S_{k-1} 'e $2|x_{i_k}|$ eklenmesiyle bulunur. Burada

$$S_{k-1} = -\sum_{i=1}^{12} |x_i| + 2\sum_{i=1}^{k-1} |x_{i_l}|, k = 1, 2, 3, \dots, r \text{ dir. Böylece,}$$

l	0	1	2	3	4	5	6	7
S_l	-71	-68	-55	-46	-41	-24	-7	10
$2 x_{il} $	-	3	13	9	5	17	17	17

olur. O halde,

$\frac{y_{i7}}{x_{i7}}$, β 'nın optimal değeridir. Yani $\frac{y_{i7}}{x_{i7}} = 0.659$ dir. Regresyon doğrusu;

$$\hat{Y} = -10.8 + 0.659X$$

ile verilir.

Şimdi kısıtlama olmaksızın $Y = \alpha + \beta X + e$ lineer regresyon modelini ele alarak MINMAD metodu ile α ve β katsayılarını tahmin etme problemini ele alalım.

En küçük kareler (EKK) metodunda α ve β tahminleri rezidülerin kareleri toplamı, $\sum \hat{e}_i^2$, minimum olacak şekilde seçilir. MINMAD metodunda, α ve β rezidülerin mutlak değerlerinin toplamı, $\sum |\hat{e}_i|$, minimum olacak şekilde seçilir.

Yani, MINMAD $\hat{\alpha}$ ve $\hat{\beta}$

$$\sum |y_i - (a + bx_i)| \quad (2.1.1)$$

toplamını minimum eden a ve b değerleridir.

$y_i - (a + bx_i)$ farkı $\hat{Y} = a + bX$ doğrusundan (x_i, y_i) noktasının sapması olarak adlandırılır.

LAD tahmin kavramı EKK tahmin kavramından daha zor değildir. Aslında, $|\hat{e}|$, $\hat{e}^2$ den daha doğru ve basit bir ölçümdür.

Fakat tahminlerin gerçek hesaplamalarında LAD metodu daha karmaşıktır. LAD metodu için formüller yoktur, bunların hesaplaması için bir algoritma verilir.

Algoritma, verilen verilerin tek olmama ya da bozulma durumlarına sahip olmadığını kabul eder.

Tek olmama : Bir veri noktasından geçen birden fazla en iyi doğru olması.

Bozulma : Bir veri noktasından geçen en iyi doğrunun iki veya daha fazla noktadan geçmesi.

Algoritma:

Amaçımız mutlak sapmaların toplamını minimum yapan, verilere en iyi uyan doğruyu bulmaktır.

Algoritmanın temel kısmı, verilen herhangi (x_0, y_0) için buradan geçen doğrular arasında en iyisini bulma işlemidir.

Bu işlem LAD regresyon doğrusunun verilen iki noktadan geçmesi ile birlikte kullanılır. Böylece algoritma verilen noktalardan biri ile başlar. Diyelim ki (x_1, y_1) olsun. Buradan geçen en iyi doğru bulunur. Bu doğru aynı zamanda verilen noktalar içinde başka birinden de geçer. Buna da (x_2, y_2) diyelim. (x_2, y_2) den geçen en iyi doğruyu bulalım. Bu doğru başka noktadan da geçer. Bu nokta da (x_3, y_3) olsun. (x_3, y_3) den geçen en iyi doğruyu bulalım ve böyle devam edelim. En son doğru bir öncekiyle aynı oluncaya kadar algoritma devam eder. Bu doğru diğerleri arasında en iyisidir. Yani LAD regresyon doğrusudur.

Şimdi verilen bir (x_0, y_0) noktasından geçen tüm doğrular arasında en iyisinin bulunması işlemini tanımlayalım.

Verilen her (x_i, y_i) için (x_0, y_0) ve (x_i, y_i) den geçen doğruların eğimleri hesaplanır. Bazı i ler için $x_i = x_0$ ise eğim tanımlı değildir, bu noktalar göz ardı edilebilir.

Verilen noktaları

$$\frac{y_1 - y_0}{x_1 - x_0} \leq \frac{y_2 - y_0}{x_2 - x_0} \leq \dots \leq \frac{y_n - y_0}{x_n - x_0}$$

olacak şekilde yeniden sıralayalım. $T = \sum |x_i - x_0|$ olsun.

$$|x_1 - x_0| + \dots + |x_{k-1} - x_0| < \frac{1}{2}T$$

$$|x_1 - x_0| + \dots + |x_{k-1} - x_0| + |x_k - x_0| > \frac{1}{2}T \quad (2.1.2)$$

koşullarını sağlayan k indisini bulalım.

(x_0, y_0) dan geçen en iyi doğru $\hat{Y} = \alpha^* + \beta^* X$ dir.

Burada,

$$\left. \begin{aligned} \beta^* &= \frac{y_k - y_0}{x_k - x_0} \\ \alpha^* &= y_0 - \beta^* x_0 \end{aligned} \right\} \quad (2.1.3)$$

dır.

İki Noktadan Geçen En İyi doğrunun Açıklanması :

Algoritmanın bulunmasında kullanılan gerçek, LAD regresyon doğrusunun verilen verilerin iki noktasından geçmesidir. Niçin böyle olduğunu görmek için, $y = a + bx$ doğrusunu alalım ve verilen noktaları işaretleyelim. Doğrudan bir noktanın mutlak sapması, noktadan doğruya indirilen dik doğru parçasının uzunluğudur.

$\sum |y_i - (a + bx_i)|$, mutlak sapmaların toplamıdır. Doğrunun herhangi bir noktadan geçmediğini varsayalım. Eğer doğru küçük bir miktar, örneğin ϵ kadar, yukarıya kaydırılırsa her mutlak sapma verinin, yukarıda ya da aşağıda olmasına göre, ya ϵ kadar azalır ya da ϵ kadar artar. (2.1.1) in değeri doğrunun aşağı ya da yukarı hareket ettirilmesiyle azalabilir (veya en azından artmaz). Doğru bir noktadan geçinceye kadar hareket ettirilir.

Doğru bir noktadan geçiyorsa, saatin dönme yönünde ya da tersine ikinci bir noktadan geçinceye kadar hareket ettirilir. Bir nokta için mutlak sapma sıfır olur ve diğer mutlak sapmaların her biri azalır ya da artar. Mutlak sapmaların azalma ya da artmasına göre döndürme yönleri farklı etkilere sahip olur. Böylece, mutlak

sapmaların toplamını azaltabiliriz (ya da en azından arttırmayabiliriz). Bu, mutlak sapmaların toplamını minimum etmek için sadece verilen noktaların en az ikisinden geçen doğrulara bakmaya ihtiyaç duyduğumuzu gösterir.

Algoritmanın Açıklanması :

(x_0, y_0) dan geçen en iyi doğrunu bulunması işlemi aşağıdaki gibi açıklanabilir:

(x_0, y_0) dan geçen tüm doğrular arasından (2.1.1)'i minimum yapan doğruyu bulmak istiyoruz.

$\hat{Y} = a + bX$ doğrusunun (x_0, y_0) dan geçmesinin anlamı $y_0 = a + bx_0$ demektir.

Böylece $a = y_0 - bx_0$ olur. Bundan faydalananarak (2.1.3)'ün ikinci kısmı bulunur.

$y_i - (a + bx_i)$ sapması $(y_i - y_0) - b(x_i - x_0)$ şeklinde yazılabilir. Böylece ,

$$\sum |(y_i - y_0) - b(x_i - x_0)| \quad (2.1.4)$$

ifadesini minimum yapacak b değerini bulmak isteriz.

(2.1.4), b nin bir fonksiyonu olarak göz önüne alınır. Bu fonksiyonun minimumunu bulmak için fonksiyonun türevini alırız. Mutlak değer fonksiyonu $|t|$, $t = 0$ da türevlenemez, fakat her $t \neq 0$ için türevlenebilir. Bu, $(y_i - y_0) - b(x_i - x_0) = 0$ ı sağlayan b hariç tüm b lerde (2.1.4) ün türevinin olduğunu gösterir.

(2.1.2) eşitsizlikleri

$b < \beta^*$ için (2.1.4) ün türevinin negatif olması koşulunu

$b > \beta^*$ için (2.1.4) ün türevinin pozitif olması koşulunu

ifade eder. Bununla birlikte (2.1.4) ün türevlenemediği noktalarda bile, sürekli bir fonksiyon olması gerçeği

$b < \beta^*$ için (2.1.4) ün azalan

$b > \beta^*$ için (2.1.4) ün artan ve

böylece β^* ın b nin minimum değeri olmasını gerektirir. Bu da (2.1.3) ün ilk kısmını verir.

(x_0, y_0) dan geçen en iyi doğru verilen noktaların birinden de geçer, örneğin (x_k, y_k) . Bunu kontrol etmek için, (2.1.3) deki α^* ve β^* in tanımlarını kullanarak $y_k = \alpha^* + \beta^* x_k$ olduğu gösterilir.

$\frac{y_i - y_0}{x_i - x_0}$ eğiminin $x_i = x_0$ olduğu zaman tanımlı olmadığını hatırlayalım.

Böyle verilerin göz ardı edilmesine izin verilir. Çünkü (x_0, y_0) ile aynı x-değerine sahip bir noktanın en iyi doğrunun belirlenmesinde hiçbir katkısı yoktur. Bu, (2.1.4) toplamını oluşturan $|(y_i - y_0) - b(x_i - x_0)|$ ifadesinin her b için aynı olan $|y_i - y_0|$ değerine sahip olmasının bir sonucudur.

(2.1.2) Koşullarının Açıklanması:

(2.1.2) koşulları (2.1.4) fonksiyonunun türevinin alınmasıyla bulunabilir. Bu fonksiyon, $b_i = \frac{y_i - y_0}{x_i - x_0}$ oranları hariç her b için türevlenebilir.

$b_{i-1} < b < b_i$ aralığında (2.1.4) ün türevi

$|x_1 - x_0| + \dots + |x_{i-1} - x_0| - |x_i - x_0| - \dots - |x_n - x_0|$ dir. Bu aşağıdaki gibi görülebilir:

$d_i, (y_i - y_0) - b(x_i - x_0)$ olsun. O zaman (2.1.4) , $\sum |d_i|$ dir.

$$|d_i| = |(y_i - y_0) - b(x_i - x_0)| = |x_i - x_0| \left| \frac{y_i - y_0}{x_i - x_0} - b \right| = |x_i - x_0| |b_i - b|.$$

$$b < b_i \leq \dots \leq b_n \text{ ise } |d_i| = |x_i - x_0| (b_i - b), \dots, |d_n| = |x_n - x_0| (b_n - b).$$

$$b_1 \leq \dots \leq b_{i-1} < b \text{ ise } |d_1| = |x_1 - x_0| (b - b_1), \dots, |d_{i-1}| = |x_{i-1} - x_0| (b - b_{i-1}) \text{ dir.}$$

Bu yüzden, $b_{i-1} < b < b_i$ ise (2.1.4),

$b(|x_1 - x_0| + \dots + |x_{i-1} - x_0| - |x_i - x_0| - \dots - |x_n - x_0|) + c$ ye eşittir. Burada c , b içermeyen sabitlerdir. b nin lineer fonksiyonunun türevi b nin katsayısıdır. Böylece türev $R_i - S_i$ dir. Burada $R_i = |x_1 - x_0| + \dots + |x_{i-1} - x_0|$,

$S_i = |x_i - x_0| + \dots + |x_n - x_0|$ dır. $R_i + S_i = T$ olduğundan türev

$R_i - (T - R_i) = 2R_i - T$ şeklinde yazılabilir. Türev negatiftir $\Leftrightarrow R_i < \frac{1}{2}T$.

Böylece (2.1.2) deki ilk koşul (2.1.4) ün türevinin $b < b_k$ için negatif olduğunu ve ikinci koşul $b > b_k$ için türevin pozitif olduğunu söyler.

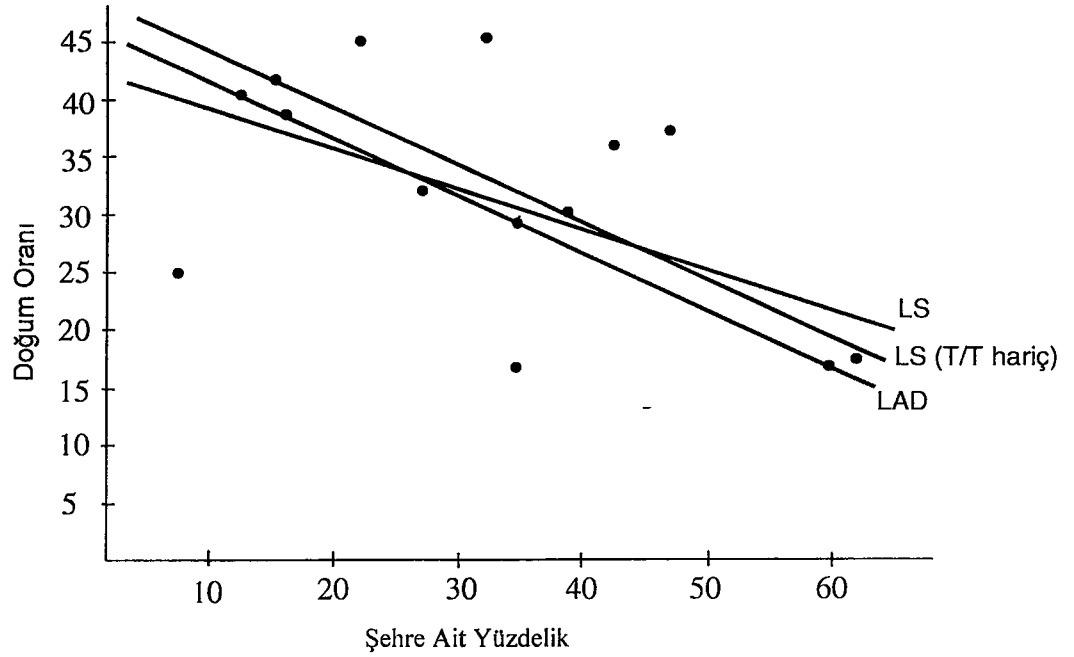
Doğum Oranı Örneği :

Tabloda Kuzey ve Merkez Amerika'da 1985 de nüfusu 1 milyonun üzerinde olan 14 ülke verildi. Her ülke için tabloda 1980-1985 için doğum oranı (1000 insanın yıldan yıla doğumlarının sayısı) ve 1980 de şehre ait yüzdelik (100.000 in üzerindeki insanın yaşadığı şehirlerdeki nüfus yüzdesi) verilmiştir.

Bu veriler için EKK metodu uygulanırsa $\hat{Y} = 43 - 0.40X$ olur. Sapan değer olan T/T çıkarılırsa $\hat{Y} = 49 - 0.55X$ olur. Görüldüğü gibi bir tek ülkenin çıkarılması EKK regresyon doğrusunu oldukça etkilenmektedir. Sapan değerlerin etkisini sınırlamak için kullanılan metodlardan biri LAD yöntemidir. Bu yöneme göre regresyon doğrusu $\hat{Y} = 46 - 0.54X$ olup T/T çıkarıldığında regresyon doğrusu aynıdır.

Çizelge 2.1 Kuzey ve Merkez Amerika'daki 14 ülkeye ait doğum oranı ve şehre ait yüzdelik

	Doğum Oranı (Y)	Şehre Ait Yüzdelik (X)
Canada	16,2	55
Costa Rica	30,5	27,3
Cuba	16,9	33,3
Dominican Republic	33,1	37,1
El Salvador	40,2	11,5
Guatemala	38,4	14,2
Haiti	41,3	13,9
Honduras	43,9	19
Jamaica	28,3	33,1
Mexico	33,9	43,2
Nicaragua	44,2	28,5
Panama	28	37,7
Trinidad/Tobago	24,6	6,8
United States	16	56,5



Şekil 2.2 LS, LS (T/T hariç) ve LAD regresyon doğrularının karşılaştırılması

Örnek için LAD regresyon doğrusunun bulunması:

İlk adım olarak Canada'ya ait (55,16.2) noktasından geçen en iyi doğruyu bulalım. Bunun için $\frac{y_i - 16.2}{x_i - 55}$ eğimlerini oluştururuz. 13 ülkenin eğimleri artan sırada aşağıdaki gibidir:

	$\frac{y_i - 16.2}{x_i - 55}$	$ x_i - 55 $	$ x_i - 55 $ in kümülatif toplamı
Mexico	-1,5	11,8	11,8
Nicaragua	-1,0566	26,5	38,3
Dominican Republic	-0,9441	17,9	56,2
Honduras	-0,7694	36	92,2
Panama	-0,6821	17,3	109,5
Haiti	-0,6107	41,1	150,6
Jamaica	-0,5525	21,9	172,5
El Salvador	-0,5517	43,5	216
Guatemala	-0,5447	40,8	256,8
Costa Rica	-0,5162	27,7	284,5
T/T	-0,1743	48,2	332,7
U.S.	-0,1333	1,5	334,2
Cuba	-0,0323	21,7	335,9

(2.1.2) yi uygulamak için, $\sum |x_i - 55| = 335.9$ bulunur. İkiye bölünürse 177.95 olur. $172.5 < 177.95$ ve $216 > 177.95$ olduğunu görürüz. Böylece $\beta^* = -0.5517$ ve $\alpha^* = 46.54$ bulunur. Bu aşamada aslında α^* ve β^* hesaplamaya gerek yoktur. Sadece Canada noktasından geçen en iyi doğrunun El Salvador noktasından geçtiğini bilmeye ihtiyaç duyarız.

Şimdi El Salvador'dan geçen en iyi doğruyu buluruz:

Eğimleri $\frac{y_i - 40.2}{x_i - 11.5}$ şeklinde oluşturur, artan sırada sıralar ve $|x_i - 11.5|$ ve $\sum |x_i - 11.5|$ için tablo oluştururuz.

Bu tablo oluşturulursa, $\sum |x_i - 11.5| = 265.5$ bulunur. Bu değer yarısı 132.75 tir. Kümülatif toplamı 132.75 i geçen ilk ülke U.S. dir. Bu yüzden El Salvador noktasından geçen en iyi doğru U.S. den de geçer.

Daha sonra U.S. den geçen en iyi doğru bulunur. Bu doğrunun El Salvador'dan geçtiği görülür. Fakat bu bir önceki adımda bulduğumuz doğru ile aynıdır. Böylece algoritma durur. LAD regresyon doğrusu El Salvador ve U.S. den geçen doğrudur.

$$\text{Eğimi } \hat{\beta} = \frac{40.2 - 16}{11.5 - 56.5} = -0.5378 \text{ ve } \hat{\alpha} = 40.2 - (-0.5378)(11.5) = 46.38 \text{ dir.}$$

Yani LAD regresyon doğrusu $\hat{Y} = 46.38 - 0.5378X$ dir.

LAD regresyon bir çok veri için bulunabilir. Fakat ara sıra tek olmama ve bozulma problemleri ortaya çıkar.

Tek olmamanın anlamı: Bir noktadan geçen en iyi doğrunun birden fazla olması demektir.

Bozulmanın anlamı: Bir noktadan geçen en iyi doğrunun aynı zamanda diğer iki veya daha fazla noktadan geçmesidir.

Bu durumda gelecek adımda kullanılacak çok nokta var demektir ve bunlardan iyi bir seçim yapılmamışsa algoritma döngüye girer ve LAD regresyon doğrusu bulunamaz.

Karşılaşılabilecek diğer bir problem ise (2.1.2) koşulunda eşitlik ortaya çıkmasıdır. Ya da

$\beta^* = \frac{y_k - y_0}{x_k - x_0}$ in $\frac{y_{k-1} - y_0}{x_{k-1} - x_0}$ ya da $\frac{y_{k+1} - y_0}{x_{k+1} - x_0}$ 'a eşit olmasıdır. Bu durumda aşağıdaki gibi algoritmayı uyguluyoruz.

Basit Algoritma: Bu algoritma kuramsal olarak basit olma avantajına ve çok hesaplama gerektirme dezavantajına sahiptir.

LAD regresyon doğrusunun en az iki noktadan geçtiği bilinmektedir böylece olabilecek tüm ikililerden geçen LAD regresyon doğrusu bulunabilir. Bunların bazıları aynıdır. Bu doğruların her biri için mutlak sapmaların toplamı (2.1.1) hesaplanır ve en küçük toplama sahip olan seçilir. Bu algoritmanın kullanılması örneklem hacmi olan n -e bağlıdır.

Bu algoritma bozulmadan etkilenmez. Tek olmama durumunda keyfi olarak birini seçeriz ya da onların ortalamalarını alabiliriz. Bu ortalamalı doğru hem de bir LAD regresyon doğrusudur.

$\beta = 0$ in test edilmesi :

Örnek için $\hat{\beta} = -0.5378$ bulmuştuk. Tahminin gerçek değere tam olarak eşit olmasını bekleyemeyiz. $\hat{\beta}$ negatif olsa bile, β nın gerçek değerinin 0 olması mümkündür. β nın 0 olup olmadığını test etmek isteyelim.

Testin tanımı: Önce LAD regresyon tahminleri olan $\hat{\alpha}$ ve $\hat{\beta}$ ve $\hat{e}_i = y_i - (\hat{\alpha} + \hat{\beta}x_i)$ rezidüleri hesaplanır.

$m = n - 2$ sıfırdan farklı rezidülerin sayısı olsun. Sıfırdan farklı rezidüleri artan sırada düzenleyelim. $\hat{e}_1$ en küçük ve $\hat{e}_m$ en büyük rezidüyü gösterecek. Örnek için bunlar,

i	1	2	3	4	5	6	7	8	9	10	11	12
$\hat{e}_i$	-18.128	-11.576	-1.203	-0.607	-0.348	-0.248	1.89	2.391	6.667	7.733	10.748	13.142

dir. $k_1, \frac{m+1}{2} - \sqrt{m}$ 'e en yakın tamsayı ve $k_2, \frac{m+1}{2} + \sqrt{m}$ 'e en yakın tamsayı olsun.

$$\hat{\tau} = \frac{\sqrt{m}}{4} [\hat{e}_{k_2} - \hat{e}_{k_1}] \quad (2.1.5)$$

hesaplanır. Sonra tahmin edicinin standart sapması

$$S(\hat{\beta}) = \frac{\hat{\tau}}{\sqrt{\sum (x_i - \bar{x})^2}} \quad (2.1.6)$$

ile bulunur. Test istatistiği

$$|t| = \frac{|\hat{\beta}|}{S(\hat{\beta})} \quad (2.1.7)$$

dir. Testin p-değeri $P[|T| \geq |t|]$ olarak hesaplanır. Burada T , $n-2$ serbestlik dereceli t-dağılımına sahip bir rastgele değişkeni gösterir.

(2.1.6) ve (2.1.7) nin açıklaması:

(2.1.7) istatistiği $\beta = 0$ olup olmadığını test etmek için uygundur. $\hat{\beta}$ tahmininin β 'ya mümkün olduğunca yakın olması beklenir. $\hat{\beta}$ ve β arasındaki fark bir veya iki standart sapma $S(\hat{\beta})$ dan fazla olmamalıdır. $|t|$ büyük ise bunun anlamı $\hat{\beta}$ ve 0 arasındaki fark $S(\hat{\beta})$ dan daha büyüktür. Böylece $\beta \neq 0$ deriz.

(2.1.6) dan x değerleri fazla yayılmaya sahipse $S(\hat{\beta})$ nın daha küçük tahmin edilebileceğini söyleyebiliriz. x değerlerinin geniş sınırı regresyon doğrusunun eğimi hakkında daha iyi bilgi verir. Hem de τ küçük ise tahmin daha değerlidir.

(2.1.5) deki τ değeri EKK regresyonda σ nın oynadığı role benzer rolü LAD regresyonda oynar.

$\hat{\beta}_{EKK}$ nın standart sapması

$$\frac{\sigma}{\sqrt{\sum (x_i - \bar{x})^2}},$$

$\hat{\beta}_{LAD}$ ın standart sapması yaklaşık olarak

$$\frac{\tau}{\sqrt{\sum (x_i - \bar{x})^2}}$$

dir. Örneklem hacmi büyükse yaklaşım daha iyidir. Böylece $\frac{\tau}{\sigma}$ oranı iki regresyon metodundan hangisinin regresyon doğrusunun eğimini tahmin etmede daha iyi olduğunu gösterir. τ ve σ nın her ikisi de rastgele hataların hacminin ölçümleridir. τ , $1/2\theta$ ya eşittir. Burada θ , τ nun medyanında hataların dağılımının olasılık yoğunluğudur.

σ büyük ise hataların dağılımı geniştir. Böylece medyan civarında olasılık yoğunluğu düşüktür, böylece θ küçülür bu da τ nun büyük olmasını gerektirir.

Yani kabaca söylemek gerekirse; τ , σ büyük iken büyük ve σ küçük iken küçüktür. Fakat τ / σ oranı hataların kitlesinin dağılımının şekline bağlıdır. Hatalar normal dağılıma sahipse, $\tau / \sigma = 1.235 > 1$, yani geniş hacimli örneklem için, LAD regresyon tahminleri EKK regresyon tahminlerinden daha az doğrudur. Hatalar Laplace dağılımına sahipse, $\tau / \sigma = 0.707 < 1$ dir.

(2.1.5) in açıklaması:

Örneklem hacmi n 'nin büyük olduğunu varsayalım. O zaman m , yaklaşık olarak n 'e eşittir ve $\hat{e}_i$ daha çok gerçek hatalar e_i gibi davranır. τ 'nın $\frac{\sqrt{n}}{4} [\hat{e}_{k_2} - \hat{e}_{k_1}]$ ifadesine yaklaşabileceğini tartışacağız. Burada k_1 ve k_2 $\frac{n+1}{2} \pm \sqrt{n}$ 'e en yakın tamsayılardır. EKK in standart sapması σ_c formundadır. LAD regresyon tahmine karşılık gelen standart sapma da τ_c formundadır. (Burada c aynıdır.) Regresyonun özel bir durumunu göz önüne alalım: Bağımsız. değişkenler olmasın ve bir kitleden seçilmiş $e_1, e_2, \dots, e_n$ gibi n bağımsız gözlemden oluşan bir örnekleme sahip olalım.

EKK regresyon tahmini örneklem ortalaması $\bar{e}$ dir. Çünkü $a = \bar{e}$, $\sum (e_i - a)^2$ toplamını minimum yapan a değeridir, standart sapması $\frac{\sigma}{\sqrt{n}}$ dir. LAD regresyon tahmini örneklem medyanı $\tilde{e}$ dir. Çünkü $a = \tilde{e}$, $\sum |e_i - a|$ toplamını minimum yapan a değeridir. Standart sapması yaklaşık olarak $\frac{\tau}{\sqrt{n}}$ dir.

Merkezi limit teoreminden, geniş hacimli örneklem için, örneklem ortalamasının dağılımının yaklaşık olarak normal dağıldığını biliyoruz.

v , hatalar kitlesinin medyanını gösterebilir. Büyük n için $\tilde{e}$ nin beklenen değeri v ve standart sapması $\frac{\tau}{\sqrt{n}}$ ile yaklaşık olarak normal dağıldığını ifade edebiliriz.

Böylece, eğer $\hat{\tau}$ τ nun bir tahmini ise v için %95 lik bir güven aralığını $\tilde{e} \pm \frac{2\hat{\tau}}{\sqrt{n}}$ şeklinde oluşturabiliriz.

Bu aralığın uzunluğu $\frac{4\hat{\tau}}{\sqrt{n}}$ dir. e_{k_1} den e_{k_2} ye olan aralık, v için bulunan %95 lik güven aralığına yaklaşıp. En azından büyük n ler için iki aralık benzerdir. Uzunlukları aynıdır.

$$\frac{4\hat{\tau}}{\sqrt{n}} \approx e_{k_1} - e_{k_2}$$

Bunu tam eşitlik yapmak için e yerine $\hat{e}$ ve n yerine $m = n-2$ yazılmasıyla (2.1.5)'i elde ederiz.

t dağılımı :

LAD test işlemi, $\beta = 0$ iken t rastgele değişkeninin $n-2$ serbestlik dereceli (s.d.li) t dağılımına sahip olduğunu kabul eder. Bu tam olarak doğru değil fakat n büyük iken t yaklaşık olarak standart normal dağılıma sahiptir. Bu n büyük iken LAD test işlemini sağlar, çünkü herhangi bir s.d.sinde t -dağılımı standart normal dağılıma çok benzer. n küçük iken standart normal dağılım yerine t -dağılımı kullanırız.

Örnek için test: $n=14$ küçük bir örneklem hacmi olduğu için bu bölümde anlatılan LAD testin geçerliliği şüphelidir. Fakat bir örnek olması için bu veriye uygulayalım.

Sıralanmış rezidüler tablo olarak verilmişti.

$$m=n-2=12, \frac{m+1}{2} - \sqrt{m} = 3.04, k_1 = 3 \text{ ve benzer şekilde } k_2 = 10 \text{ bulunur.}$$

$$\hat{\tau} = \frac{\sqrt{12}}{4} [\hat{e}_{10} - \hat{e}_3] = \frac{\sqrt{12}}{4} [7.733 + 1.203] = 7.739, \sum (x_i - \bar{x})^2 = 3151 \text{ bulunur.}$$

$$\text{Böylece } S(\hat{\beta}) = \frac{7.739}{\sqrt{3151}} = 0.1379 \text{ ve } |t| = |-0.5378| / 0.1379 = 3.9 \text{ bulunur. } p-$$

değerini hesaplamak için $n-2 = 12$ s.d.li t -dağılımını kullanırız. t -tablosundan p -değerinin 0.001 ve 0.01 arasında olduğunu buluruz. Böylece regresyon doğrusunun eğiminin negatif olduğunu sonuçlandırırız.

Çoklu Regresyona Örnek :

LAD metodunu aşağıdaki tabloda verilen yangın kazaları ile ilgili veriyi analiz etmek için uygulayalım. Yangın kazalarının bölgenin üç karakteristiği olan evin yaşı, hırsızlık kazası ve ailenin geliri ile ilgisini görmek isteyelim. B=Bölge, S=Seçilen her 1000 ev için yangın sayısı, Y=(Evin Yaşı) 1940 dan önce inşa edilen

evlerin oranı, H=Her 1000 ev için hırsızlık sayısı, G=Bölgedeki ailelerin gelirlerinin medyanı(1000\$) olsun.

Tablo 2.1.2 Yangın kazaları ile ilgili veri

B	S	Log(S)	Y	H	G
1	6.2	1.825	0.604	29	11.744
2	9.5	2.251	0.765	44	9.323
3	10.5	2.351	0.375	36	9.948
4	7.7	2.041	0.669	37	10.656
5	8.6	2.152	0.814	53	9.730
6	34.1	3.529	0.526	68	8.231
7	11.0	2.398	0.426	75	21.480
8	6.9	1.932	0.785	18	11.104
9	7.3	1.988	0.901	31	10.694
10	15.1	2.715	0.898	25	9.631
11	29.1	3.371	0.827	34	7.995
12	2.2	0.788	0.402	14	13.722
13	5.7	1.740	0.279	11	16.250
14	2	0.693	0.077	11	13.686
15	2.5	0.916	0.638	22	12.405
16	3	1.099	0.512	17	12.198
17	5.4	1.686	0.851	27	11.600
18	2.2	0.788	0.444	9	12.765
19	7.2	1.974	0.842	29	11.084
20	15.1	2.715	0.898	30	10.510
21	16.5	2.803	0.727	40	9.784
22	18.4	2.912	0.729	32	7.342
23	36.2	3.589	0.631	41	6.565
24	39.7	3.681	0.830	147	7.459
25	18.5	2.918	0.783	22	8.014
26	23.3	3.148	0.790	29	8.177
27	12.2	2.501	0.480	46	8.212
28	5.6	1.723	0.715	23	11.230
29	21.8	3.082	0.731	4	8.330
30	21.6	3.073	0.650	31	5.583
31	9	2.197	0.754	39	8.564
32	3.6	1.281	0.208	15	12.102
33	5	1.609	0.618	32	11.876
34	28.6	3.353	0.781	27	9.742
35	17.4	2.856	0.686	32	7.52
36	11.3	2.425	0.734	34	7.388
37	3.4	1.224	0.020	17	13.842
38	11.9	2.477	0.57	46	11.040
39	10.5	2.351	0.559	42	10.332
40	10.7	2.37	0.675	43	10.908
41	10.8	2.380	0.580	34	11.156
42	4.8	1.569	0.152	19	13.323
43	10.4	2.342	0.408	25	12.96
44	15.6	2.747	0.578	28	11.26
45	7	1.946	0.114	3	10.080
46	7.1	1.96	0.492	23	11.428
47	4.9	1.589	0.466	27	13.731

$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + e$ modelini ele alalım. $Y = \log(S)$, $X_1 = Y$, $X_2 = H$, $X_3 = G$ olsun. LAD tahminleri $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3$, $\sum |\hat{e}_i|$ minimum olacak şekilde seçilir. Yani; $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3$

$$\sum |y_i - (b_0 + b_1 x_{i1} + b_2 x_{i2} + b_3 x_{i3})| \quad (2.1.8)$$

minimum olacak şekilde b_0, b_1, b_2, b_3 ün değerleridir. Şimdi bir algoritma tanımlayacağız. Bu algoritma tek olmama veya bozulma problemi dışında herhangi bir problem içermez.

olsun.

$$b = \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix} \quad \text{ve} \quad x_i = \begin{bmatrix} 1 \\ x_{i1} \\ x_{i2} \\ x_{i3} \end{bmatrix}$$

$$\sum |y_i - b'x_i| \quad (2.1.9)$$

şeklinde yazılabilir. (2.1.9)'u minimum yapacak b vektörünü bulmak istiyoruz. Basit LAD regresyonda olduğu gibi çoklu LAD regresyon algoritması da iteratiftir. b vektörü ile başlarız ve en iyi vektör $\hat{\beta}$ yı buluncaya kadar devam ederiz. Her adımda uygun d vektörü yönünde en iyi vektör b^* 'ı buluruz ve $b^* = b + td$ en iyi olacak şekilde t nin değeri bulunur.

d yönünde en iyi tahmin vektörünün bulunması :

$$\sum |y_i - (b + td)'x_i| \quad (2.1.10)$$

minimum olacak şekilde t nin en iyi değerini bulmak için bir işleme ihtiyaç duyarız.

$z_i = y_i - b'x_i$ ve $w_i = d'x_i$ alırsak işlem

$$\sum |z_i - tw_i| \quad (2.1.11)$$

minimum olacak şekilde t değeri bulmaya indirgenir. Bu (2.1.4)'ü minimum yapan

b bulma problemi ile aynıdır ve çözümü biliyoruz. $\frac{z_i}{w_i}$ oranlarını bulur ve artan

sıraya koyarız. z ve w ları bu sıraya göre yeniden indeksler ve

$$|w_1| + \dots + |w_{k-1}| < \frac{1}{2}T$$

$$|w_1| + |w_2| + \dots + |w_{k-1}| + |w_k| > \frac{1}{2}T$$

sağlanacak şekilde k indeksini buluruz. Burada $T = \sum |w_i|$ ve t nin minimum değeri $\frac{z_k}{w_k}$ dır.

En Uygun Yönün Bulunması :

Her adımda algoritma d_1, d_2, d_3, d_4 gibi dört yön vektörünü göz önüne alır. (Genelde p bağımsız değişken varsa $p+1$ vektör göz önüne alınır.) Bunlar pozitif ve negatif yönler olmak üzere 8 farklı yön temsil eder. Bu 8 farklı yön arasında en iyi yön $t=0$ civarında (2.1.10)'un değerini azaltandır. (2.1.10)'un nasıl azaldığını belirlemek için $t=0$ da sağ türevini hesaplarız. Bu türev, $W_- + W_0 - W_+$ dır. Burada

W_- , $\frac{z_i}{w_i}$ yi negatif yapan i için $|w_i|$ nin toplamı, W_0 , $z_i=0$ olan i için $|w_i|$ nin

toplamı ve W_+ , $\frac{z_i}{w_i}$ yi pozitif yapan i için $|w_i|$ nin toplamıdır. 8 yönün her biri için

bu türevi hesaplarız ve türevi negatif olanlar içinden en küçüğünü alarak en uygun yönü seçeriz. Tüm türevler pozitif ise başlangıç vektörü b en iyi vektör $\hat{\beta}$ dır ve algoritma durur.

Algoritmanın başlangıcı :

Basit regresyonda , LAD regresyon doğrusunun verilen noktaların ikisinden geçtiğini biliyoruz. Benzer şekilde, p bağımsız değişkenli çoklu regresyonda LAD regresyon denklemi p+1 noktayı sağlar. Bizim örnekte p=3 olduğundan 4 noktanın seçimi ile başlanır. b tahmin vektörü başlangıç aşamasında $y_i = b'x_i$, $i=1,2,3,4$ ile belirlenir. Matris notasyonunda bu 4 eşitlik $Ab=c$ olarak yazılabilir. Burada,

$$A = \begin{bmatrix} x'_1 \\ x'_2 \\ x'_3 \\ x'_4 \end{bmatrix} \text{ ve } c = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix}.$$

Böylece $b = A^{-1}c$ olur. d_1, d_2, d_3, d_4 yön vektörlerinin başlangıç kümesi A^{-1} in 4 kolonudur.

Algoritmanın İterasyonları :

Her adımda tahmin edilen regresyon denklemi verilerin 4 noktası tarafından sağlanacak şekilde b tahmininin bir geçerli vektörüne sahip oluruz. Noktalar i_1, i_2, i_3, i_4 ile indekslenir.

$$A = \begin{bmatrix} x'_{i_1} \\ x'_{i_2} \\ x'_{i_3} \\ x'_{i_4} \end{bmatrix}$$

olsun. Geçerli yön vektörleri A^{-1} in 4 kolonudur. Yukarıda tanımlandığı gibi, yönlerdeki 8 türevi hesaplarız ve türevi negatif olanlar içinden en küçüğünü alarak en uygun yönü seçeriz. Varsayalım ki bu d_3 olsun. b den daha iyi tahmin vektörü $b^* = b + t^* d_3$ ile bulunur. Burada t^* (2.1.10)'u minimum eden t nin değeridir.

$$t^* = \frac{y_k - b'x_k}{d_3'x_k} \text{ ile bulunur. } A \text{ nın 3 üncü satırı yerine } d_3 \text{ kullanıldığı için } x_k'$$

yerleştirilir. Bu yeni matris A^* ile gösterilir. Yeni yön vektörleri A^{*-1} in 4 kolonudur. İterasyon 8 yöndeki tüm türevler pozitif olana kadar devam eder.

Yangın Verileri İçin Algoritmanın Uygulanması :

$$A = \begin{bmatrix} 1 & 0.604 & 29 & 11.744 \\ 1 & 0.765 & 44 & 9.323 \\ 1 & 0.735 & 36 & 9.948 \\ 1 & 0.669 & 37 & 10.656 \end{bmatrix} \quad c = \begin{bmatrix} 1.825 \\ 2.251 \\ 2.351 \\ 2.041 \end{bmatrix}$$

tahminlerin başlangıç vektörü

$$b = A^{-1}c = \begin{bmatrix} 47.93 \\ -23.26 \\ -0.1161 \\ -2.443 \end{bmatrix} \text{ ve } A^{-1} = \begin{bmatrix} -284.9 & -308.3 & 157.8 & 436.4 \\ 164.5 & 176.6 & -79.71 & -261.3 \\ 0.5233 & 0.6744 & -0.4656 & -0.7321 \\ 14.59 & 15.51 & -8.185 & -21.91 \end{bmatrix}$$

dir. Şimdi, b den daha iyi tahmin vektörünü bulmak isteyelim. Bunu yapmak için A^{-1} in 4 kolonu ve bunların negatifleri ile temsil edilen 8 yön vektörü için (2.1.10)'un sağ türevlerini hesaplarız. A^{-1} in birinci kolonunu d_1 ile gösterelim. d_1 yönündeki türevlerin bulunması için hesaplamalar aşağıdaki tabloda gösterilmiştir. 7 ve 24 nolu bölgeler modelden çıkarılmıştır. $z_i = y_i - b'x_i$ ve $w_i = d_1'x_i$ olduğundan örneğin z_5

$$z_5 = 2.152 - [47.93 \quad -23.26 \quad -0.1161 \quad -2.443] \begin{bmatrix} 1 \\ 0.814 \\ 53 \\ 9.730 \end{bmatrix} = 3.08$$

$$w_5 = d_1' x_5 = [-284.9 \quad 164.5 \quad 0.5233 \quad 14.59] \begin{bmatrix} 1 \\ 0.814 \\ 53 \\ 9.730 \end{bmatrix} = 18.70$$

şeklinde bulunmuştur. $y_i = b'x_i$ yi $i = 1,2,3,4$ için sağlayan başlangıç tahmin vektörünü seçeriz. Bu $z_1 = z_2 = z_3 = z_4 = 0$ dır. A nın i inci satırı x_i' ve A^{-1} in birinci kolonunu d_1 dir. Böylece AA^{-1} çarpımının $(i,1)$ girişi $x_i' d_1$, yani w_i dir. $AA^{-1} = I$ olduğundan $w_1 = 1$ ve $w_2 = w_3 = w_4 = 0$ dır. z_i ve w_i hesaplandıktan sonra $\frac{z_i}{w_i}$ nin işaretini belirleriz. Eğer $w_i = 0$ ise bu oran tanımlı değildir. Fakat bu sorun değildir. $|w_i| = 0$ olduğundan böyle bir veri noktası türevi vermez. Şimdi, $\frac{z_i}{w_i}$ ya negatif ya da sıfır olan $|w_i|$ leri toplarız ve işareti pozitif olan $|w_i|$ yi çıkarırız. Bu d_1 yönünde türev olarak -1221 'i verir. d_1 için hesaplamalar yapıldıktan sonra, $-d_1$ yönünde türev hesaplanır. Tablo çok az değişir. z_i ve $|w_i|$ nin değerleri aynı kalır, fakat w_i nin işareti değişir ve böylece $z_i = 0$ olması durumu hariç $\frac{z_i}{w_i}$ nin işareti değişir. $z_i = 0$ olan 4 nokta vardır. Bunlardan biri için $w_i = 1$ ve diğer üçü için $w_i = 0$ dır. Bu $-d_1$ yönündeki türevin $-(-1221 - 1) + 1 = 1223$ olmasını gerektirir. Genelde d ve $-d$ yönündeki türevler 2 eklenmiş olmalıdır. A^{-1} in dört kolonu için türevler ve onların negatifleri $-1221, 1223, -1323, 1325, 654, -652, 1903, -1901$ dir. Bunların en küçüğü $-d_4$ yönündeki -1901 dir. Böylece, $b + t d_4$ formundaki en iyi vektörü

ararız.(Burada t negatif, fakat işaretini belirtmeye gerek yoktur. Çünkü algoritma otomatik olarak t nin işaretini göz ardı eder.)

i	z_i	w_i	$\frac{z_i}{w_i}$ nin işareti	$ w_i $
1	0	1	0	1
2	0	0	*	0
3	0	0	*	0
4	0	0	*	0
5	3.08	18.71	+	18.71
6	-4.16	-42.68	+	42.68
8	1.48	15.67	+	15.67
9	4.74	35.57	+	35.57
10	2.1	16.43	+	16.43
11	-1.84	-14.42	+	14.42
12	-2.65	-11.21	+	11.21
13	1.27	3.88	+	3.88
14	-10.74	-66.75	+	66.75
15	0.68	12.57	+	12.57
16	-3.15	-13.78	+	13.78
17	5.02	38.48	+	38.48
18	-4.59	-20.88	+	20.88
19	4.07	30.51	+	30.51
20	4.83	31.87	+	31.87
21	0.33	-1.62	-	1.62
22	-6.41	-41.11	+	41.11
23	-8.87	-63.85	+	63.85
25	-4.67	-27.65	+	27.65
26	-3.06	-20.46	+	20.46
27	-8.86	-62.04	+	62.04
28	0.53	8.62	+	8.62
29	-7.03	-41.01	+	41.01
30	-12.5	-80.29	+	80.29
31	-2.74	-15.50	+	15.50
32	-10.51	-66.23	+	66.23
33	0.78	6.8	+	6.8
34	0.52	-0.15	-	0.15
35	-7.03	-45.58	+	45.58
36	-6.44	-38.57	+	38.57
37	-10.46	-70.71	+	70.71
38	0.12	-5.97	-	5.97
39	-2.46	-20.2	+	20.2
40	1.78	7.8	-	7.8
41	-0.86	-8.91	+	8.91
42	-8.08	-55.53	+	55.53
43	-1.54	-15.58	+	15.58
44	-0.98	-10.86	+	10.86
45	-18.36	-117.48	+	117.48
46	-3.94	-25.17	+	25.17
47	1.18	6.25	+	6.25

$z_i = y_i - b'x_i$ ve $w_i = d'_4x_i$ olsun. T nin değeri (0,0) ve (w_i, z_i) , $i = 1, 2, \dots, 47$, $i \neq 7, 24$ verilerinden geçen en iyi doğrunun eğimi olan en iyi vektörü verir. Bunun için algoritma daha önce verilmişti. Bu algoritmanın uygulanmasıyla, (0,0) dan geçen en iyi doğrunun aynı zamanda (w_{14}, z_{14}) den geçtiğini buluruz. Bu yüzden A ve c nin 4 üncü satırları yerine 14 nolu bölgenin verilerini yerleştiririz. Yani,

$$A = \begin{bmatrix} 1 & 0.604 & 29 & 11.744 \\ 1 & 0.765 & 44 & 9.323 \\ 1 & 0.735 & 36 & 9.948 \\ 1 & 0.077 & 11 & 13.686 \end{bmatrix} \quad c = \begin{bmatrix} 1.825 \\ 2.251 \\ 2.351 \\ 0.693 \end{bmatrix}$$

olur. Her adımda 4 veri geçerli regresyon denklemini belirler. Bunlar bazen temel olarak adlandırılır. Her adımda temeldeki verilerden biri yerine temel dışından bir veri yerleştirilir. Biz başlangıç temel olarak 1,2,3,4 bölgelerinin verileri ile başladık. İlk adımda 4 nolu bölge yerine 14 nolu bölge yerleştirildi. Algoritma 3 yerine 29, 2 yerine 26, 1 yerine 23, 26 yerine 10, 29 yerine 39, 10 yerine 19, 14 yerine 45 ve 23 yerine 37 yerleştirilmesiyle devam eder. Bu son adımdan sonra temel 37,19,39 ve 45 noktalarından oluşur. 8 yön için türevler bu adımda hesaplanırsa hepsinin pozitif olduğunu görürüz. Böylece algoritma durur. LAD regresyon denklemi 37,19,39 ve 45 veri noktaları ile belirlenir.

$$\hat{\beta} = \begin{bmatrix} 1 & 0.020 & 17 & 13.842 \\ 1 & 0.842 & 29 & 11.084 \\ 1 & 0.559 & 42 & 10.332 \\ 1 & 0.114 & 3 & 10.080 \end{bmatrix}^{-1} \begin{bmatrix} 1.224 \\ 1.974 \\ 2.351 \\ 1.946 \end{bmatrix} = \begin{bmatrix} 4.362 \\ -0.09098 \\ 0.01299 \\ -0.2425 \end{bmatrix}$$

ve

$$\hat{Y} = 4.362 - 0.09098X_1 + 0.01299X_2 - 0.2425X_3$$

olur.

2.2 Mutlak Sapmaların Maksimumunu Minimum Yapma (MINMAXAD) Regresyon

Burada β_0 ve β_1 kestirimini mutlak sapmaların maksimumunu minimum yaparak bulmaya çalışacağız. Bu kriterler altında β_0 ve β_1

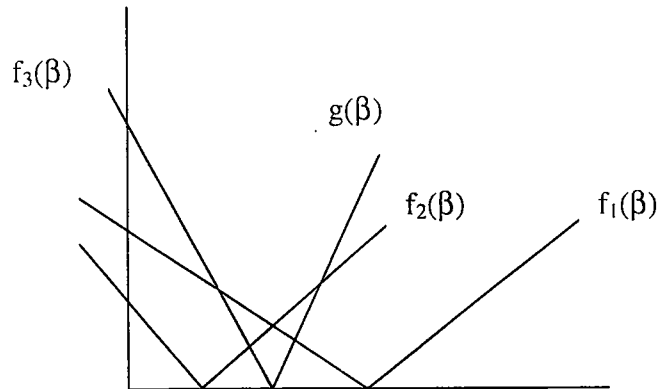
$$\min_{(\beta_0, \beta_1)} \left[\max_{1 \leq i \leq m} |Y_i - (\beta_0 + \beta_1 X_i)| \right]$$

nin bir çözümü olacaktır.

İlk olarak problemi β_0 olmaksızın yani $Y = \beta X + \varepsilon$ şeklinde ele alalım. Herhangi bir i için $f_i(\beta) = |Y_i - \beta X_i|$, $(Y_i/X_i, 0)$ da minimumu olan ve $-|X_i|, |X_i|$ eğimlerine sahip iki doğrudan oluşur.

Aşağıdaki verileri ele alarak $f_i(\beta)$ nın grafiğini çizelim.

i	Y_i	X_i
1	3	1
2	1	1
3	4	2



Şekil 2.3 $f_1(\beta) = |3 - \beta|$, $f_2(\beta) = |1 - \beta|$, $f_3(\beta) = |4 - 2\beta|$, $\max f_i(\beta) = g(\beta)$ yı gösteren grafik.

Görüldüğü gibi $\max_i f_i(\beta) = g(\beta)$ dir. $g(\beta)$ parçalı lineer konveks fonksiyondur.

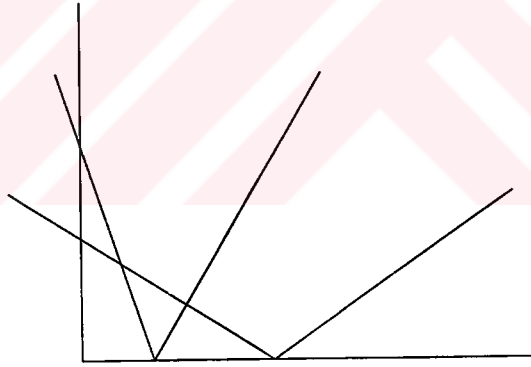
Sonuç 2.2.1 : $g(\beta) = \max_i |Y_i - X_i \beta|$ verilen (X_i, Y_i) için $(i = 1, 2, 3, \dots, n)$ parçalı lineer konveks fonksiyondur.

Aşağıdaki grafikten de görülebileceği gibi $g(\beta)$ nin grid noktalarının X_i/Y_i noktaları olması gerekmez. Bununla birlikte $g(\beta)$ nın grid noktaları

$$g(\beta) = Y_i + X_i \beta \text{ veya } g(\beta) = -(Y_i + X_i \beta) \quad (i \neq j \text{ için})$$

$$g(\beta) = Y_j + X_j \beta \text{ veya } g(\beta) = -(Y_j + X_j \beta)$$

gibi bazı doğruların kesim noktalarıdır. Bu nedenle bunun gibi kesim noktalarını bulup bunların minimumunu seçerek problemin çözümüne ulaşabiliriz.



Şekil 2.4 Şekil 2.1.3 den farklı bir lineer konveks fonksiyon grafiği

2.3. Sapmalar Arasındaki Farkın Mutlak Değerinin Toplamını Minimum Yapma (MINSADBED) Regresyon

$\sum_{i=1}^n d_i^2$ nin minimizasyonu için EKK regresyon kriterini görmüştük ve d_i 'nin i-inci gözleme karşılık gelen gözlenen değerin kestirilmiş değerden sapması

olduğunu biliyoruz. Şimdi $\sum_{i=1}^n d_i^2$ yi $\frac{1}{n} \sum_{i=1}^n (d_i - \bar{d})^2$ şeklinde sapmanın varyansı olarak düşünelim. Burada

$$\bar{d} = \frac{1}{n} \sum_{i=1}^n d_i = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)$$

dır. Fakat EKK doğrusunun $(\bar{X}, \bar{Y})$ dan geçtiği durum için $\bar{d} = 0$ dır. O halde

$$\sum_{i=1}^n (d_i - \bar{d})^2 \approx \frac{1}{n} \sum_{i < j} (d_i - d_j)^2 \text{ şeklinde yazabiliriz.}$$

Sonuç olarak $\sum_{i < j} (d_i - d_j)^2$ 'yi minimize etmek ,bir regresyon doğrusu bulmak için bir kriter olarak kullanılabilir. Aslında $\bar{d} = 0$ EKK regresyon doğrusuna eşittir. $(d_i - d_j)^2$ yerine $|d_i - d_j|$ kullanarak ‘sapmalar arasındaki farkın mutlak değerinin toplamı’ adında bir kriter elde edebiliriz.

Burada

$$\text{Minimize } \sum_{i < j} |d_i - d_j|$$

kullanarak β_0 ve β_1 parametrelerinin kestirimini ele alacağız.

d_i ve d_j 'nin tanımından

$$\sum_{i < j} |d_i - d_j| = \sum_{i < j} |(Y_i - \beta_0 - \beta_1 X_i) - (Y_j - \beta_0 - \beta_1 X_j)| = \sum_{i < j} |(Y_i - Y_j) - \beta_1 (X_i - X_j)|$$

olur.

$$Y_i - Y_j = Y_{ij} \quad , \quad X_i - X_j = X_{ij} \quad (i < j) \quad \text{dersek}$$

$$\sum_{i < j} |d_i - d_j| = \sum_{i < j} |Y_{ij} - \beta_1 X_{ij}|$$

elde edilir.

Burada bizim amaç fonksiyonumuz sadece β_1 'in kestirildiği MINMAD regresyondaki fonksiyon ile aynıdır. Bu problemi çözmek için $\frac{n(n-1)}{2}$ tane X_{ij}, Y_{ij} noktası ile MINMAD regresyonu uygulayabiliriz.

Bu yöntemle β_0 'ı kestiremeyiz. β_0 'ın kestirimini elde etmenin bir yolu; doğruyu $(\bar{X}, \bar{Y})$ 'dan geçirmek ve sabit terim β_0 ile kesişimini almaktır. Yani -

$$\hat{\beta}_0 = \bar{Y} - \bar{X}\hat{\beta}_1$$

olur.

β_0 'ın kestirimi için başka yollar da vardır. Bunlardan biri de

$$\hat{\beta}_0 = \text{medyan}_{i < j} \frac{1}{2}(Y_i + Y_j)$$

dir.

Örnek 2.3.1.: Aşağıdaki verileri ele alalım

Gözlem no	Y	X
1	2	5
2	5	4
3	4	6
4	8	9
5	3	7

i, j	Y_{ij}	X_{ij}	Y_{ij}/X_{ij}	Rank
1,2	-3	1	-3	1
1,3	-2	-1	2	9
1,4	-6	-4	1.5	8
1,5	-1	-2	0.5	5
2,3	1	-2	-0.5	4
2,4	-3	-5	0.6	6
2,5	2	-3	-0.67	3
3,4	-4	-3	1.33	7
3,5	1	-1	-1	2
3,5	5	2	2.5	10

$\sum |X_{ij}| = 24$ tür. $n=5$ olduğundan $5(5-1)/2=10$ tane X_{ij}, Y_{ij} noktası ile MINMAD regresyon yöntemini uygulayacağız. Buna göre aşağıdaki tablo oluşturulur.

l	0	1	2	3	4	5	6
S_l	-24	-22	-20	-14	-10	-6	4
$2 X_{(ij)_l} $	0	2	2	6	4	4	10

Tablodan anlaşıldığı gibi rankın 6 olduğu durum fonksiyonumuzun minimum olduğu noktadır. $i=2$ ve $j=4$ için $Y_{24}/X_{24} = 0.6$ bulunur. Yani $\hat{\beta}_1 = 0.6$ 'dır. Şimdi de β_0 'ın kestirimi için bir tablo oluşturalım.

i, j	1,2	1,3	1,4	1,5	2,3	2,4	2,5	3,4	3,5	4,5
$1/2(Y_i + Y_j)$	3.5	3	5	2.5	4.5	6.5	4	6	3.5	5.5
Rank	3	2	7	1	6	10	5	5	4	8

$\hat{\beta}_0 = \text{medyan}_{i < j} \left\{ \frac{1}{2}(Y_i + Y_j) \right\} = 4.25$ bulunur. Buna göre MINSADBED regresyon doğrusu

$$\hat{Y} = 4.25 + 0.6X$$

olarak bulunur.

2.4 Mutlak Sapmalar Arasındaki Mutlak Farkların Toplamını Minimum Yapma (MINSADBAD) Regresyon

$$d_i = Y_i - (\beta_0 + \beta_1 X_i) \quad \text{için} \quad \min_{\beta_0, \beta_1} \sum_{i < j} \|d_i - d_j\| \quad \text{nin kullanılması ile}$$

β_0 ve β_1 in kestiriminin bulunduğu bir yöntemdir.

Bu çeşit kestirimlerin elde edilmesini ileriki bölümlerde daha genel yapıda ele alacağız.

3. MATEMATİKSEL PROGRAMLAMA VE REGRESYON ANALİZİ

3.1. Mutlak Sapmaların Ortalamasının Minimum Yapılması (MINMAD) Regresyon

β ya bağlı olarak $\sum |d_i|$ nin minimizasyonunu düşünelim. d_i : i-inci gözlem için, Y_i nin gözlenen değeri ile tahmin edilmiş değeri arasındaki sapmadır. Bu metod L_1 - norm minimizasyon problemi olarak da bilinir.

Bu problemi aşağıdaki gibi ifade edebiliriz:

$$\begin{aligned} \text{Min} \quad & \sum |d_i| \\ & X\beta + d = Y \\ & d, \beta \text{ işaret kısıtlaması yok} \end{aligned} \quad (3.1.1)$$

$|d_i| = d_{1i} + d_{2i}$ şeklinde ifade edebiliriz. Burada d_{1i} ve d_{2i} nonnegatif ve $d_i = d_{1i} - d_{2i}$ dir ve bu problemi aşağıdaki şekilde tekrar formüle edebiliriz.

$$\begin{aligned} \text{Min} \quad & \sum d_{1i} + \sum d_{2i} \\ & X\beta + d_1 - d_2 = Y \\ & \beta \text{ işaret kısıtlaması yok} \\ & d_1 - d_2 \geq 0 \end{aligned} \quad (3.1.2)$$

Şimdi bu problemi çözmek için kullanılan belli sonuçların ispatı ile devam edelim.

Tanım 3.1.1 : $X\beta + Id_1 - Id_2 = Y$ yi sağlayan herhangi (β, d_1, d_2) vektörüne (3.1.2) nin bir çözümü denir.

A, $(X, I, -I)_{n \times p+2n}$ ve W, $(\beta, d_1, d_2)_{p+2n \times p}$ şeklinde gösterilsin, A nın kolonları a_j olarak belirlensin. $AW=Y$ yi sağlayan W (3.1.2) nin bir çözümüdür. C' ,

$(0, e', e')$ şeklinde bir vektör olsun. $0: 1 \times p$ vektör. $e' = (1, 1, \dots, 1)_{1 \times n}$ vektör olsun. $C'W$ (3.1.2) nin *amaç fonksiyonu* olarak adlandırılır.

Tanım 3.1.2 : (3.1.2) nin herhangi bir W çözümü

$$W_j \geq 0, j = p+1, \dots, p+2n$$

koşulunu sağlıyorsa W ya problem için *uygun çözüm* deriz.

Not 3.1.1 : Bir W uygun çözümü varsa aşağıdaki gibi bir W^* uygun çözümü bulabiliriz.

$$W_j^* = \begin{cases} W_j, & j = 1, \dots, p \\ W_{p+r} - W_{p+n+r} \text{ eger } W_{p+r} - W_{p+n+r} > 0 \text{ ise, } j = p+r, r = 1, \dots, n \\ 0, & \text{diger hallerde} \\ -W_{p+r} + W_{p+n+r} \text{ eger } W_{p+r} - W_{p+n+r} < 0 \text{ ise, } j = p+n+r, r = 1, \dots, n \\ 0, & \text{diger hallerde} \end{cases}$$

$W_j \geq 0, j = p+1, \dots, p+2n$ ve

$$AW^* = [X, I, -I]W^* = \sum_{l=1}^p X_l W_l + \sum_{r=1}^n e_r W_{p+r}^* - \sum_{r=1}^n e_r W_{p+n+r}^*$$

W^* in tanımından W_{p+r}^* veya W_{p+n+r}^* in sadece biri pozitif olabilir. Bu durumda

$$AW^* = AW = Y$$

yazarız veya W^* problemin uygun bir çözümüdür deriz.

(Başka türlü belirtilmedikçe ‘çözüm’ (3.1.2) için uygun çözüm anlamına gelecektir.)

Tanım 3.1.3 : W nun sıfırdan farklı bileşenlerine karşılık gelen A nın kolonları R^n de lineer bağımsız vektör kümeleri formunda ise W ya *temel uygun çözüm* denir.

Not 3.1.2 : Eğer $[a_1, \dots, a_p]$ nin rankı r ise $[a_1, \dots, a_p]$ de lineer bağımsız en fazla r kolon vardır. Bu yüzden p değişkenden en fazla r tanesi için (temel uygun çözümde) $-W_j$ sıfırdan farklı olacaktır.

Not 3.1.3 :
$$W_{p+r} = \begin{cases} Y_r & \text{eger } Y_r > 0 \text{ ise} \\ 0 & \text{diğer hallerde} \end{cases}$$

ve

$$W_{p+n+r} = \begin{cases} -Y_r & \text{eger } Y_r < 0 \text{ ise} \\ 0 & \text{diğer hallerde} \end{cases}$$

tanımlanır ve W_j lerin geri kalanları sıfırdır. Yukarıdaki gibi tanımlanan W bir temel uygun çözümdür.

Sonuç 3.1.1 : $F = \{W \mid AW = Y, W_j \geq 0, j = p+1, \dots, p+2n\}$ bir konveks kümedir.

Tanım 3.1.4 : Eğer $0 < \lambda < 1$, $W^1, W^2 \in F$, $W^1 \neq W^2$ için

$$W = \lambda W^1 + (1 - \lambda) W^2$$

olacak şekilde bir λ yoksa $W \in F$ ye F nin bir uç noktası denir.

Biz bu uç noktaların temel uygun çözümler ile ilişkilendirilmesi ile ilgileneceğiz.

Sonuç 3.1.2 : Eğer $W \in F$, F nin uç noktası ise W nun sıfırdan farklı bileşenlerine karşılık gelen A nın kolonları lineer bağımsızdır.

İspat: W, F nin bir uç noktası olsun. genelliği bozmadan, ilk s bileşeninin sıfırdan farklı olduğunu farz edelim. Eğer s sıfır ise buna karşılık gelen kolonların kümesi boştur. Bu yüzden lineer bağımsızdır.

Eğer $(a_1, \dots, a_s)$ lineer bağımlı ise hepsi sıfır olmayan $\alpha_1, \dots, \alpha_s$ gerçel sayıları vardır öyle ki

$$\sum_{l=1}^s \alpha_l a_l = 0 \text{ dır.}$$

1.Durum: Eğer bu a_l lerin hiçbiri kısıtlı W_j lere karşılık gelmiyor ise

$$W_j \pm \alpha_j, \quad j = 1, \dots, s, \quad \text{ve} \quad W_j = 0 (\text{diğerleri için})$$

$AW=Y$ yi sağlar. Bu

$$\sum_{j=1}^s (W_j \pm \alpha_j) a_j = \sum_{j=1}^s W_j a_j \pm \sum_{j=1}^s \alpha_j a_j = Y + 0 = Y$$

dir.

$W_j \pm \alpha_j$ uygundur. Çünkü tüm kısıtlı W_j ler sıfırdır ve

$$W_j = \frac{1}{2}(W_j + \alpha_j) + \frac{1}{2}(W_j - \alpha_j)$$

dir. W, nun, F nin bir uç noktası idi. Dolayısıyla bu bir çelişkiye yol açar.

2.Durum: a_l lerin bir veya daha fazlası kısıtlı W_j lere karşılık geliyor ise yeteri kadar küçük bir $\delta > 0$ vardır öyle ki kısıtlı W_j ler için $W_j \pm \delta \alpha_j > 0$ dır.

Ayrıca

$$\sum_{j=1}^s (W_j \pm \delta \alpha_j) a_j = \sum_{j=1}^s W_j a_j \pm \delta \sum_{j=1}^s \alpha_j a_j$$

dir.

$$W_j^1 = \begin{cases} W_j + \delta\alpha_j & j = 1, 2, \dots, s \\ 0 & \text{diğer hallerde} \end{cases}$$

ve

$$W_j^2 = \begin{cases} W_j - \delta\alpha_j & j = 1, 2, \dots, s \\ 0 & \text{diğer hallerde} \end{cases}$$

olsun.

Burada W^1 ve W^2 uygun çözümlerdir ve

$$W = \frac{1}{2}W^1 + \frac{1}{2}W^2$$

dir. Bu da W nun F için bir uç noktası olması ile çelişir.

Böylece ispat tamamlanmış olur.

Örnek 3.1.1 :

$$X = \begin{bmatrix} 1 & 2 & 4 \\ 1 & 1 & 3 \\ 1 & 2 & 5 \end{bmatrix} \quad ; \quad Y = \begin{bmatrix} 4 \\ 2 \\ 4 \end{bmatrix}$$

olarak verilsin.

$W = (-1, 0, 1, 1, 0, 0, 0, 0, 0)$, $(X, I_3, -I_3)W = Y$ yi sağlasın ve (a_1, a_3, a_4) kolonları lineer bağımsız olsun. W , F nin bir uç noktası değildir. Çünkü

$$W = \frac{1}{2}W^1 + \frac{1}{2}W^2$$

dir.

$$W^1 = (-\frac{1}{2}, 1, \frac{1}{2}, \frac{1}{2}, 0, 0, 0, 0) \quad , \quad W^2 = (-\frac{3}{2}, -1, \frac{3}{2}, \frac{3}{2}, 0, 0, 0, 0)$$

W^1 ve W^2 problem için uygun çözümlerdir.

Eğer tüm değişkenler işaretlerinde kısıtlı ise tersi de doğrudur.

Sonuç olarak; eğer W , F nin bir uç noktası ise W nun en fazla n tane sıfırdan farklı bileşeni vardır. Geri kalanlar sıfırdır.

Sonuç 3.1.3 : F sadece sonlu sayıda temel uygun çözüm içerir.

İspat: Her bir temel uygun çözümde sıfırdan farklı bileşenler, A dan karşılık gelen kolonların kümesi ile tek bir şekilde hesaplanır. Fakat, büyüklüğü n veya daha az olan lineer bağımsız kolon vektörlerinin alt kümeleri sonlu sayıdadır. Sonuç olarak; F sonlu sayıda birçok uç noktaya sahiptir.

Sonuç 3.1.4 : Eğer F boş değil ise, F nin uç noktalarının kümesi F^0 da boş değildir.

İspat: Herhangi bir $W \in F$ için, W nun sıfırdan farklı bileşenlerinin sayısı ile bir $\rho(W)$ fonksiyonu tanımlayabiliriz.

$0 \leq \rho(W) \leq p + 2n$ dir. Eğer F boş değilse, fonksiyon F üzerinde ρ_0 minimum noktasına ulaşır. $\rho(\bar{W}) = \rho_0$ olduğunu varsayalım. $\bar{W}$ nin F nin bir uç noktası olduğunu göstereceğiz. Eğer $\rho_0 = 0$ ise $\bar{W} = 0$ dır ve tanımdan $\bar{W}$ bir uç noktadır. Çünkü sıfırdan farklı bileşenlere karşılık gelen kolonların kümesi boştur ve ayrıca lineer bağımsızdır.

Eğer $\rho_0 > 0$ ise, $\bar{W} = (\bar{W}_1, \dots, \bar{W}_{\rho_0}, 0, \dots, 0)$ kabul edebiliriz. $\bar{W}$ nin F için bir uç nokta olmadığını kabul edersek bir çelişki elde ederiz. $a_1, \dots, a_{\rho_0}$ kolonları lineer

bağımsızdır ve $\sum_{j=1}^{\rho_0} \alpha_j a_j = 0$ olacak şekilde hepsi sıfır olmayan $\alpha_1, \dots, \alpha_{\rho_0}$ lar vardır.

$\alpha_j \neq 0$ için $\bar{W}_j / |\alpha_j|$ yi ele alalım ve kısıtlı verilere karşılık gelen j lerin en

küçüğünü bulalım. λ bu sayı olsun ve $\lambda = \overline{W}_{j_0} / |\alpha_{j_0}|$ olsun. Genelliği bozmadan $\alpha_{j_0} > 0$ diyebiliriz. Bu noktada

$$\overline{W} = \begin{cases} \overline{W}_j - \lambda \alpha_j & j = 1, \dots, \rho_0 \\ 0 & \text{diger} \end{cases} \in F$$

dir. Çünkü

$$A\overline{W} = \overline{A} - \lambda \sum_{j=1}^{\rho_0} \alpha_j a_j = Y$$

dir ve kısıtlı verilere karşılık gelen $\overline{W}$ ler λ nın seçiminden negatif değildir. Ayrıca $\overline{W}$ nın ρ_0 dan daha az bileşeni vardır çünkü $\overline{W}_{j_0} - \lambda \alpha_{j_0} = 0$ dır, F üzerinde $\rho(W)$ nun minimumu ρ_0 olduğundan çelişki oluşturur.

Böylece $\overline{W}$ bir uç noktadır.

F sınırlı olduğunda ona *konveks polihedron* denir.

Sonuç 3.1.5 : Eğer F bir konveks polihedron ise F nin her W noktası, F nin uç noktalarının bir konveks fonksiyonu olarak yazılabilir.

Tanım 3.1.5 : $C'W$ yu minimum yapan herhangi $W \in F$ ye problemin optimal çözümü denir.

Sonuç 3.1.6 : Eğer F konveks polihedron ise, $C'W$ en azından F nin bir uç noktasında minimuma ulaşır.

İspat: F sınırlı konveks küme iken, $C'W$ minimuma bazı $W^0 \in F$ de ulaşsın. Eğer W^0 bir uç nokta ise ispat bitmiş demektir. F sadece $W^1, W^2, \dots, W^k$ gibi sonlu sayıda uç noktaya sahiptir

Sonuç 3.1.5 den

$$W^0 = \sum_{j=1}^k \lambda_j W^j, \sum_{j=1}^k \lambda_j = 0$$

olacak şekilde bir $\lambda_j \geq 0$ vardır.

$$C'W^0 = \sum_{j=1}^k \lambda_j C'W^j \text{ dir.}$$

$W^{j_0}, C'W^{j_0} = \min_{1 \leq j \leq k} C'W^j$ olacak şekilde olsun. Böylece

$$C'W^0 = \sum_{j=1}^k \lambda_j C'W^j \geq \sum_{j=1}^k \lambda_j C'W^{j_0} = C'W^{j_0} \quad \lambda_j \geq 0, \sum_{j=1}^k \lambda_j = 1$$

olur.

Bir başka deyişle $C'W^0 \geq C'W^{j_0}$ dır.

Diğer yandan $C'W^0 \leq C'W$ ($W \in F$ için) dur. Öyle ise $C'W \leq C'W^{j_0}$ olur.

Buradan,

$$C'W = C'W^{j_0}$$

dır.

F sınırlı olmadığı durumda Sonuç 3.1.5 e benzer bir sonuç doğru değildir.

Sonuç 3.1.7 : Eğer F sınırlı ve $C'W$ F de minimuma ulaşıyor ise en azından F nin bir uç noktası optimal çözümdür.

Buradan şu sonucu çıkarabiliriz: Bir optimal çözüm ararken, F nin uç noktalarını içermesinden dolayı temel uygun çözümler göz önüne alınabilir.

3.2. MINMAD Regresyon İçin Simpleks Yöntemi

Bu bölümde, bir çözüm var iken optimal temel uygun çözümü elde etmek için “simpleks yöntemi” denen bir metod geliştireceğiz.

Problemimiz, sıfır tarafından alttan sınırlı kapalı konveks bir küme üzerinde sürekli bir fonksiyonun minimizasyonu problemi olduğundan her zaman bir uygun çözümü vardır.

Tanım 3.2.1 : A matrisinde n lineer bağımsız kolonun oluşturduğu kümeye A'nın *bazı* denir.

Bu kolonlara karşılık gelen $B=[a_{i1}, \dots, a_{in}]$ de olmayan $n+p$ tane kolon vardır. Bu kolonlara *temel olmayan kolonlar* denir. Bu kolonlar için $W_j = 0$ da bir uygun çözüm bulabiliyorsak bu, temel uygun çözümdür. Bu çözümde temel kolonlara karşılık gelen W_j lerin bazıları sıfır olabilir. Bu gibi durumda bu çözüme bir *dejenere temel uygun çözüm* denir. Eğer bu W_j lerin n-tanesi de sıfırdan farklı ise buna *dejenere olmayan temel uygun çözüm* denir.

Bir temel uygun çözüme karşılık gelen bir B bazı ile başlayarak daha iyi bir temel uygun çözüm bulmayı deneyeceğiz. B, temel uygun çözüm W ya karşılık gelen bir baz olsun. W_B , B de temel kolonlara karşılık gelen W_j lerin vektörünü temsil etsin. C_B de C_j sabitlerine karşılık gelsin.

Bazda olmayan herhangi bir vektör bazdaki vektörlerin lineer kombinasyonları olarak ifade edilebilir. Bu sabitler α_{ij} ile gösterilsin. a_j , $(a_{B_1}, \dots, a_{B_n})$ in bir lineer kombinasyonu olarak ifade edilir. B_i , i-inci temel kolon olsun. Öyle ise,

$$a_j = \sum_{i=1}^n \alpha_{ij} a_{B_i} \quad \text{veya} \quad B^{-1} a_j = \alpha_j$$

olur.

Eğer temel olmayan a_j kolonlarını içeren B bazından yeni bir baz elde etmek istiyorsak B deki bazı B_r kolonlarını çıkarmamız gerekir. Fakat bunu keyfî olarak yapamayız. Bunun için bazı bağıntılar vereceğiz.

$$\sum_{i=1}^n W_{B_i} B_i = Y$$

dir ve işaretiyle kısıtlı veriler yerine geçen B_i ler için $W_{B_i} \geq 0$ dir. R_B , bazdaki bu kısıtlı verilerin indisini gösterebilir. B nin rankının n ve sadece p tane kısıtsız değişkenin modelde olduğuna dikkat edelim, $p < n$ dir. Burada R_B her zaman boş olmayacaktır. Böylece, genelleştirilmiş teoride R_B lerin boş olması olasılığı bir teori geliştirmeye yardım eder.

Şimdi, a_j , herhangi bir $\alpha_{ij} \neq 0$ için, B_r vektörünün yerine geçebilir ve yeni vektörlerin kümesi lineer bağımsız n vektörün kümesi formundadır. Biz, en azından bir $\alpha_{ij} \neq 0$ olacak şekilde bir a_j seçeriz ve bazın içine koyarız. Sonra

$$B_r = \frac{1}{\alpha_{ij}} a_j - \sum_{i=1, i \neq r}^n \frac{\alpha_{ij}}{\alpha_{rj}} B_i$$

olur. Y nin gösteriminde B_i lerle B_r leri yerine koyarsak

$$\sum_{i=1, i \neq r}^n (W_{B_i} - W_{B_r} \frac{\alpha_{ij}}{\alpha_{rj}}) B_i + \frac{W_{B_r}}{\alpha_{rj}} a_j = Y$$

elde ederiz.

Biz, $W_{B_i} - W_{B_r} \alpha_{ij} / \alpha_{rj} \geq 0$, $i \in R_B$, $i \neq r$ olmak üzere uygun olacak yeni bir çözüm istiyoruz. Aynı zamanda, eğer W_j işaretinde kısıtlı ise $W_{B_r} / \alpha_{rj} \geq 0$ olmasını isteyebiliriz.

Durum 1: W_j işaretinde kısıtlı değildir. W_{B_r} / α_{rj} nin negatif olmaması gerekir ve biz sadece

$$W_{Bi} - \frac{W_{Br}}{\alpha_{rj}} \alpha_{ij} \geq 0, \quad i \in R_B, \quad i \neq r \quad (3.2.1)$$

olmasını isteriz.

R_B boş değilken $W_{Br} \geq 0$ olacak şekilde işaretinde kısıtlı bir W_B nin bulunduğuna dikkat edelim. Bu yüzden $\alpha_{rj} > 0$ iken tüm $\alpha_{ij} \leq 0$ ve $\alpha_{rj} < 0$ iken tüm $\alpha_{ij} \geq 0$ için (3.2.1) ifadesi sağlanır. Bu yüzden eğer biz r- yi

$$\frac{W_{Br}}{\alpha_{rj}} = \min_{\substack{i \in R_B \\ \alpha_{ij} > 0}} \left[\frac{W_{Bi}}{\alpha_{ij}} \right]$$

veya

$$\frac{W_{Br}}{\alpha_{rj}} = \max_{\substack{i \in R_B \\ \alpha_{ij} < 0}} \left[\frac{W_{Bi}}{\alpha_{ij}} \right]$$

olacak şekilde seçersek (3.2.1) sağlanacaktır.

R_B boş iken (3.2.1) sağlanır. Böylece herhangi bir B_r bazdan çıkarılabilir.

Durum 2: W_j işaretinde kısıtlıdır. Öyle ise $W_{Br}/\alpha_{rj} \geq 0$ dır. Ayrıca biz (3.2.1) in sağlanmasını isteriz. Böylece

$$\begin{aligned} W_{Bi} - \frac{W_{Br}}{\alpha_{rj}} \alpha_{ij} &\geq 0, \quad i \in R_B, \quad i \neq r \\ W_{Br}/\alpha_{rj} &\geq 0 \end{aligned} \quad (3.2.2)$$

elde ederiz.

Eğer $r \in R_B$ ise $W_{Br} \geq 0$ dır. Böylece α_{ij} lerin pozitif olması gerekir. Durum 1 deki gibi r-yi

$$\frac{W_{Br}}{\alpha_{rj}} = \min_{\substack{i \in R_B \\ \alpha_{ij} > 0}} \left[\frac{W_{Bi}}{\alpha_{ij}} \right]$$

şeklinde seçeriz.

Diğer yandan eğer R_B boş ise biz r -yi $W_{Br}/\alpha_{rj} \geq 0$ olacak şekilde seçebiliriz. Sonra, amaç fonksiyonu geliştirmemizdeki durum gibi a_j kolonunu seçmek için denemeler yaparız.

$$Z_j = \sum_{i=1}^n C_{Bi} \alpha_{ij}, \quad Z(B) = \sum_{i=1}^n C_{Bi} W_{Bi} \quad \text{olsun} \quad \text{ve} \quad Z(\hat{B}) \quad \text{da}$$

$\hat{B} = (B_1, \dots, B_{i-1}, a_j, B_{r+1}, \dots, B_n)$ bazda iken V_B nın amaç fonksiyonunda karşılık gelen değerini gösterebiliriz. Biz $Z(\hat{B}) \leq Z(B)$ olmasını istiyoruz.

$$\begin{aligned} Z(\hat{B}) &= \sum_{\substack{i=1 \\ i \neq r}}^n C_{Bi} (W_{Bi} - \frac{W_{Br}}{\alpha_{rj}} \alpha_{ij}) + C_j \frac{W_{Br}}{\alpha_{rj}} \\ &= \sum_{i=1}^n C_{Bi} (W_{Bi} - \frac{W_{Br}}{\alpha_{rj}} \alpha_{ij}) + C_j \frac{W_{Br}}{\alpha_{rj}} \\ &= \sum_{i=1}^n C_{Bi} W_{Bi} + \frac{W_{Br}}{\alpha_{rj}} \left[- \sum_{i=1}^n C_{Bi} \alpha_{ij} + C_j \right] \end{aligned}$$

göz önüne alalım. Böylece $Z(\hat{B}) \leq Z(B)$ olur, eğer

$$\frac{W_{Br}}{\alpha_{rj}} [C_j - Z_j] \leq 0 \quad (3.2.3)$$

ise.

Durum 1: $C_j - Z_j < 0$

Eğer W_j işaretinde kısıtlı değilse W_{Br}/α_{rj} herhangi bir reel sayı olabilir. Fakat (3.2.3) ü sağlamak için biz $W_{Br}/\alpha_{rj} \geq 0$ olmasını isteriz. Bu nedenle eğer R_B boş değil ise $\alpha_{rj} > 0$ dır.

Eğer W_j işaretinde kısıtlı ise $W_{Br}/\alpha_{rj} \geq 0$ olmasını isteriz. Bundan dolayı eğer R_B boş değil ise, $W_{Br}/\alpha_{rj} = 0$ olmadıkça, $\alpha_{rj} > 0$ dır.

Durum 2: $C_j - Z_j > 0$

Eğer W_j nin işaretinde kısıtlama yok ise (3.2.3) ün sağlanması için $W_{Br}/\alpha_{rj} \leq 0$ olmasını isteriz. Bunun için eğer R_B boş değil ise $\alpha_{rj} < 0$ dır.

Eğer W_j nin işaretinde kısıtlama var ise (3.2.3) ün sağlanması için $W_{Br}/\alpha_{rj} \geq 0$ olmalıdır. Aksi halde $W_{Br}/\alpha_{rj} = 0$ olur. Yeni çözüm önceki kadar iyidir. Fakat biz R_B nin Problem (3.1.2) için boş olmadığını biliyoruz.

Verilen bir $W_B = B^{-1}Y$ temel uygun çözümü için Problem (3.1.2) nin $AW=Y$ kısıtlamalarının amaç fonksiyonu değeri $Z = C_B W_B$ herhangi bir a_j kolonu için A dadır fakat B de değildir. Böylece $C_j - Z_j < 0$ durumu sağlanır. En azından bir $\alpha_{ij} > 0 (i \in R_B)$ ise B nin bir kolonunun a ile yer değiştirmesinden yeni bir çözüm elde etmek mümkündür ve amaç fonksiyonunun yeni değeri $\hat{Z}, \hat{Z} \leq Z$ yi sağlar. Diğer taraftan, eğer $C_j - Z_j > 0$ sağlanırsa ve eğer a_j işaretinde kısıtlama olmayan bir W_j ye karşılık geliyorsa (en azından bir $\alpha_{ij} < 0, i \in R_B$ ise) B nin kolonlarından birinin a_j ile yer değiştirmesi ile yeni bir temel uygun çözüm elde edilmesi mümkündür ve amaç fonksiyonunun yeni değeri $\hat{Z}, \hat{Z} \leq Z$ yi sağlar.

Eğer bazdaki vektörlerin değiştirilmesi yöntemine devam edersek her bir adımda amaç fonksiyonu geliştiririz. Eğer

$$\frac{W_{Br}}{\alpha_{rj}} [C_j - Z_j] < 0$$

ise tam bir ilerlemeye sahip oluruz.

Not 3.2.1 : W_{Br}/α_{rj} sıfır değilse $C_j - Z_j$ sıfır değildir. Fakat $\alpha_{rj} \neq 0$ dır. Bu yüzden eğer $W_{Bi} \neq 0$, $i \in R_B$, ise amaç fonksiyonunda kuvvetli bir değişmeye sahip olabiliriz. Böylece, eğer bazdaki tüm kısıtlı değişkenler pozitif seviyede ise her bir adımda amaç fonksiyonu geliştiririz. Aşağıdaki durum ortaya çıktığında bazdaki değişim yönteminde bir sona varırız: tüm pozitif kısıtlı temel olmayan W_j ler için $C_j - Z_j \geq 0$ ve tüm temel olmayan kısıtsız W_j ler için $C_j - Z_j = 0$ dır. (3.1.2) için ne (1) bir temel olmayan kısıtsız değişken için $C_j - Z_j > 0 (<0)$ ve α_{ij} lerin ($i \in R_B$) hiçbirinin $<0 (>0)$ olması ne de (2) bir temel olmayan kısıtlı değişken için $C_j - Z_j < 0$ ve tüm $i \in R_B$ için $\alpha_{ij} \leq 0$ olması durumlarında bu yöntem son vermemiz mümkün değildir. Yukarıdaki (1) inci durumun olduğunu varsayalım. Burada kısıtsız W_j değişkenine karşılık gelen bir a_j için $C_j - Z_j > 0$ ve α_{ij} ($i \in R_B$) lerin hiçbirisi sıfırdan küçük değildir. $Z = C_B W_B$ olmak üzere $\sum_{i=1}^n W_{Bi} B_i = Y$ yi ele alalım. θ bir skaler olmak üzere θa_j yi ekleyip çıkaralım.

Böylece

$$\sum_{i=1}^n W_{Bi} B_i + \theta a_j - \theta a_j = Y$$

elde ederiz. Fakat $-\theta a_j = -\theta \sum_{i=1}^n \alpha_{ij} B_i$ dir. Önceki eşitlikte yerine koyarak

$$\sum_{i=1}^n (W_{Bi} - \theta \alpha_{ij}) B_i + \theta a_j = Y \quad (\theta < 0)$$

elde ederiz. $\alpha_{ij} \geq 0$ ($i \in R_b$) olduğundan $W_{Bi} - \theta \alpha_{ij} \geq 0$ dır. Böylece $W_{Bi} - \theta \alpha_{ij}$, ($i = 1, 2, \dots, n$) ve θ Problem (3.1.2) için bir temel uygun çözümdür. Bu çözümün amaç fonksiyonu değerine bakalım:

$$\hat{Z} = \sum_{i=1}^n C_{Bi} (W_{Bi} - \theta \alpha_{ij}) + C_j \theta = Z + \theta (C_j - Z_j)$$

Fakat $C_j - Z_j < 0$ ve $\theta < 0$ dır. Bunun için θ nın küçük seçilmesiyle $\hat{Z}$ küçük yapılabilir. $C'W$ alttan sıfır ile sınırlı olduğundan bu bir çelişkidir ve böyle bir durumun ortaya çıkamayacağı ispatlanmış olur. Aynı şekilde (2) nci durumun olanaksızlığı da ispat edilebilir. Bununla birlikte, bu tarz problemlerde amaç fonksiyonu alttan sınırlı olmayabilir.

Şimdi yukarıdaki düşünciyi temel alarak bir algoritma tanımlayacağız:

Algoritma:

I.Adım: Daha önce bahsedildiği gibi bir temel uygun çözüm ile başlayalım.

$$W_{p+r} = \begin{cases} Y_r & Y_r > 0 \\ 0 & - \end{cases}, \quad 1 \leq r \leq n$$

$$W_{p+n+r} = \begin{cases} -Y_r & Y_r < 0 \\ 0 & - \end{cases}, \quad 1 \leq r \leq n \quad \text{ve } W_j = 0, j = 1, \dots, p$$

olsun.

Bir tabloda α_{ij} lerin karşılıklarını sunalım. Dikkat edelim ki bu çözüme karşılık gelen B bazı tam olarak e_r ya da $-e_r$ kolonunu içermektedir (tüm $r = 1, \dots, n$ için) ve $B^{-1} = B$ dir. Burada eğer W_{p+r} bazda ise r-inci kolonun +1 ile, eğer W_{p+n+r} bazda ise r-inci kolonun -1 ile çarpılmasıyla $\alpha_j = B^{-1}a_j$ elde edilir ve $B^{-1}Y = W_B$ temel uygun çözümün karşılığını verir. Ayrıca vektörlerin ve C_j lerinin bazda olduğuna dikkat etmeliyiz. İlk tablo aşağıdaki şekilde çizelge 3.1 olarak oluşturulmuştur.

$$C_j - Z_j, \sum_{i=1}^n C_{Bi} \alpha_{ij} \text{ de her } j \text{ için } C_j \text{ nin yerine koyulmasıyla elde edilebilir.}$$

Bazdaki a_j ler için $C_j - Z_j = 0$ dir.

Çizelge 3. 1 α_{ij} leri ve bazdaki vektörleri gösteren ilk tablo

C_B	Bazdaki vektörler	W_B	α_1	α_2	$\dots$	α_j	$\dots$	α_{2n+p}
1	a_{p+1} ya da a_{p+n+1}	$ Y_1 $	α_{11}	α_{12}	$\dots$	α_{1j}	$\dots$	$\alpha_{1,2n+p}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
1	a_{p+n} ya da a_{p+2n}	$ Y_n $	α_{n1}	α_{n2}	$\dots$	α_{nj}	$\dots$	$\alpha_{n,2n+p}$
$C_k - Z_k$		$Z = \sum_{l=1}^n Y_l $	$C_j - Z_j$					

2.Adım: Aşağıdaki gibi bazdaki bir vektörle yer değiştirmek için bazda olmayan bir a_j seçelim.

$$(a) \quad C_{j1} - Z_{j1} = \max_{\substack{k \\ C_k - Z_k > 0}} C_k - Z_k \quad (W_k \text{ kısıtlı olmayan değişken})$$

$$(b) \quad |C_{j_2} - Z_{j_2}| = \max_{\substack{k \\ C_k - Z_k < 0}} |C_k - Z_k|$$

olsun. j yi $|C_j - Z_j| = \max[|C_{j_1} - Z_{j_1}|, |C_{j_2} - Z_{j_2}|]$ olacak şekilde seçelim.

Bu adım baza girecek olan vektörün seçimi için bir kriter verir. Eğer j_1 ve j_2 bulunamazsa 5. adıma, aksi taktirde 3. adıma geçilir.

3.Adım: Eğer $C_j - Z_j > 0$ ise r -yi aşağıdaki gibi seçeriz:

$$\frac{W_{Br}}{\alpha_{rj}} = \max_{i \in R_B} \left[\frac{W_{Bi}}{\alpha_{ij}}, \alpha_{ij} < 0 \right]$$

Eğer $C_j - Z_j < 0$ ise r -yi aşağıdaki gibi seçeriz:

$$\frac{W_{Br}}{\alpha_{rj}} = \min_{i \in R_B} \left[\frac{W_{Bi}}{\alpha_{ij}}, \alpha_{ij} > 0 \right]$$

Bu adım bazdan çıkarılacak vektörün seçimi için bir kriter verir. 4. adıma geçilir.

4.Adım: Yeni B bazına karşılık gelen tablo aşağıdaki gibidir.

$$\hat{W}_{Br} = \frac{W_{Br}}{\alpha_{rj}}, \quad \hat{\alpha}_{rl} = \frac{\alpha_{rl}}{\alpha_{rj}}$$

ve

$$\begin{aligned}\hat{W}_{Bi} &= W_{Bi} - \frac{W_{Br}}{\alpha_{rj}} \alpha_{ij} \quad , \quad i \neq r \quad , \quad i = 1, 2, \dots, n \\ \hat{\alpha}_{il} &= \alpha_{il} - \frac{\alpha_{ij} \alpha_{rl}}{\alpha_{rj}} \quad , \quad i \neq r \quad , \quad i = 1, \dots, n \quad , \quad l = 1, \dots, p + 2n \\ \hat{Z} &= Z + (C_j - Z_j) \frac{W_{Br}}{\alpha_{rj}} \\ \hat{C}_l - \hat{Z}_l &= (C_l - Z_l) - (C_j - Z_j) \frac{\alpha_{rl}}{\alpha_{rj}} \quad , \quad l = 1, \dots, p + 2n\end{aligned}$$

2. adıma geri dönülüp $B = \hat{B}$ ile devam edilir.

5.Adım: Durun. Şimdiki baz optimaldir. $W_1, \dots, W_p$, $\beta_0, \dots, \beta_p$ parametrelerini verir. $W_{p+r} + W_{p+n+r}$ r-inci gözlemdaki mutlak hatayı verir. Dikkat etmeliyiz ki, eğer amaç fonksiyonunun alttan sınırlı olmadığı bir problemin çözümü için bu algoritmayı uygularsak, sonlu minimuma sahip olmayan problemde gösterildiği gibi sona erdirebiliriz. Bu 3. adımda r bulmak için yetersizlik şeklinde belirtilecektir.

Bu algoritmayı aşağıdaki örnekte verilen verilerle bir MINMAD regresyon problemi çözerek örneklendirebiliriz.

Örnek 3.2.1:

Verilen veriler için $\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2$ eşitliğindeki $\beta_0, \beta_1, \beta_2$ tahmin edicilerini yukarıdaki algoritmayı kullanarak bulalım.

$$X = \begin{bmatrix} 1 & 1 & 3 \\ 1 & 2 & 2 \\ 1 & 3 & 1 \\ 1 & 4 & 1 \\ 1 & 5 & 3 \end{bmatrix} \quad Y = \begin{bmatrix} 4 \\ 3 \\ 4 \\ 5 \\ 5 \end{bmatrix}$$

1.Adım:

$d_{11}, \dots, d_{15}$ şeklindeki ilk bazı bulalım. Elimizde Çizelge 3.2 de verilen ilk çizelge olsun.

Çizelge 3.2 Örnek 3.2.1 için α_{ij} leri ve bazdaki vektörleri gösteren ilk çizelge

C_B	Bazdaki vektörler	W_B	α_1	α_2	α_3	α_4	α_5	α_6	α_7	α_8	α_9	α_{10}	α_{11}	α_{12}	α_{13}
1	a_4	4	1	1	3	1	0	0	0	0	-1	0	0	0	0
1	a_5	3	1	2	2	0	1	0	0	0	0	-1	0	0	0
1	a_6	4	1	3	1	0	0	1	0	0	0	0	-1	0	0
1	a_7	5	1	4	2	0	0	0	1	0	0	0	0	-1	0
1	a_8	5	1	5	3	0	0	0	0	1	0	0	0	0	-1
$C_k - Z_k$		$Z=21$	-5	-15	-11	0	0	0	0	0	2	2	2	2	2

2.Adım:

$C_k - Z_k$ ($k = 1,2,3$) için negatiftir. Bu yüzden adım 2(a) hiçbir j_i için sağlanmaz. adım 2(b) ye gidelim;

$$C_2 - Z_2 = 15 = \max [|-5|, |-15|, |-11|] \quad \text{burada } j = 2 \quad \text{ve} \quad C_j - Z_j < 0 \quad \text{dır.}$$

3:Adım:

$$\frac{W_{B1}}{\alpha_{i2}}, \alpha_{i2} > 0 \text{ oranlarından } r = 5 \text{ seçeriz. } \frac{W_{B5}}{\alpha_{52}} = 5/5 = [4/1, 3/2, 4/3, 5/4, 5/5].$$

Böylece bazda a_2 ile a_8 yer değiştirir. Yeni tablo 4.adımdaki gibi hesaplanır ve Çizelge 3.3 elde edilir

Çizelge 3.3 a_2 ile a_8 vektörünün yer değiştirmesi ile elde edilen çizelge

C_B	Bazdaki vektörler	W_B	α_1	α_2	α_3	α_4	α_5	α_6	α_7	α_8	α_9	α_{10}	α_{11}	α_{12}	α_{13}
1	a_4	3	4/5	0	12/5	1	0	0	0	-1/5	-1	0	0	0	1/5
1	a_5	1	3/5	0	4/5	0	1	0	0	-2/5	0	-1	0	0	2/5
1	a_6	1	2/5	0	-4/5	0	0	1	0	-3/5	0	0	-1	0	3/5
1	a_7	1	1/5	0	-2/5	0	0	0	1	-4/5	0	0	0	-1	4/5
0	a_2	1	1/5	1	3/5	0	0	0	0	1/5	0	0	0	0	-1/5
$C_k - Z_k$		6	-2	0	-2	0	0	0	0	3	2	2	2	2	-1

Örneğin; $\hat{\alpha}_{23} = 2 - 2 \cdot 3/5 = 4/5$. 2. adıma dönersek

$$|C_1 - Z_1| = \max [|-2|, |-2|, |-1|]$$

bulunur. Bu yüzden $j = 1$ ve $C_j - Z_j < 0$ dir.

3.adımda $r = 2$ bulunur. Çünkü $W_{B2}/\alpha_{21} = 5/3 = \min [15/4, 5/3, 5/3, 5/1, 5/1]$.

Böylece bazda a_1 ile a_5 yer değiştirir. Yeni çizelge 4.adımdaki gibi elde edilir.

Çizelge 3.4 a_1 ile a_5 in yer değiştirmesi ile elde edilen çizelge

C_B	Bazdaki vektörler	W_B	α_1	α_2	α_3	α_4	α_5	α_6	α_7	α_8	α_9	α_{10}	α_{11}	α_{12}	α_{13}
1	a_4	5/3	0	0	4/3	1	-4/3	0	0	1/3	-1	4/3	0	0	-1/3
0	a_1	5/3	1	0	4/3	0	5/3	0	0	-2/3	0	-5/3	0	0	2/3
1	a_6	1/3	0	0	-4/3	0	-2/3	1	0	-1/3	0	2/3	-1	0	1/3
1	a_7	2/3	0	0	-2/3	0	-1/3	0	1	-2/3	0	1/3	0	-1	2/3
0	a_2	3/3	0	1	1/3	0	-1/3	0	0	1/3	0	1/3	0	0	-1/3
$C_k - Z_k$		8/3	0	0	2/3	0	10/3	0	0	5/3	2	-4/3	2	2	1/3

2. adıma tekrar dönelim. 2(a) dan $j_1 = 1$ bulunur. $C_3 - Z_3 = 2/3 = \max (C_k - Z_k)$ dır. $C_k - Z_k > 0$, W_k işaretinde kısıtlama olmayan veriler. 2(b) $j_2 = 10$ olarak sonuçlanır. $|C_{10} - Z_{10}| = 4/3$ tür. Böylece $j = 10$, $\max[2/3 , 4/3] = 4/3$ ve $C_{10} - Z_{10} < 0$ bulunur. 3.adımda $r = 3$ buluruz.

Çünkü $W_{B3}/\alpha_{310} = 1/2 = \min [5/4 , 1/2, 2/2]$ dir. Böylece a_{10} ile a_6 yer değiştirir. Yeni çizelge 4.adımdan elde edilir.

Çizelge 3.5 a_{10} ile a_6 nın yer değiştirmesi ile elde edilen çizelge

C_B	Bazdaki vektörler	W_B	α_1	α_2	α_3	α_4	α_5	α_6	α_7	α_8	α_9	α_{10}	α_{11}	α_{12}	α_{13}
1	a_4	1	0	0	4	1	0	-2	0	1	-1	0	2	0	-1
0	a_1	5/2	1	0	-2	0	0	5/2	0	-3/2	0	0	-5/2	0	3/2
1	a_{10}	1/2	0	0	-2	0	-1	3/2	0	-1/2	0	1	-3/2	0	1/2
1	a_7	1/2	0	0	0	0	0	-1/2	1	-1/2	0	0	1/2	-1	1/2
0	a_2	1/2	0	1	1	0	0	-1/2	0	1/2	0	0	1/2	0	-1/2
	$C_k - Z_k$	2	0	0	-2	0	2	2	0	1	2	0	0	2	1

Şimdi 2.adımda $j = 3$ buluruz. Çünkü a_3 baza girecek tek uygun vektördür. 3.adımda $r = 1$ bulunur. a_4 ün a_3 ile yer değiştirmesinden Çizelge 3.6 yı elde ederiz

Çizelge 3.6 a_4 ün a_3 ile yer değiştirmesinden elde edilen çizelge

C_B	Bazdaki vektörler	W_B	α_1	α_2	α_3	α_4	α_5	α_6	α_7	α_8	α_9	α_{10}	α_{11}	α_{12}	α_{13}
0	a_3	1/4	0	0	1	1/4	0	-1/2	0	1/4	-1/4	0	1/2	0	-1/4
0	a_1	3	1	0	0	1/2	0	3/2	0	-1	-1/2	0	-3/2	0	1
1	a_{10}	1	0	0	0	1/2	-1	1/2	0	0	-1/2	1	-1/2	0	0
1	a_7	1/2	0	0	0	0	0	-1/2	1	-1/2	0	0	1/2	-1	1/2
0	a_2	1/4	0	1	0	-1/4	0	0	0	1/4	1/4	0	0	0	-1/4
	$C_k - Z_k$	3/2	0	0	0	1/2	2	1	0	3/2	3/2	0	1	2	1/2

Çizelge 3.6. da girecek hiçbir uygun j bulamayız. Böylece 5.adıma gideriz ve dururuz.

$\beta_0, \beta_1, \beta_2$ tahminlerini 2, 5 ve 1 inci satırları sırasıyla yerine koyarak Çizelge 3.6 dan okuyabiliriz. Böylece

$$\hat{Y} = 3 + \frac{1}{4}X_1 + \frac{1}{4}X_2$$

elde ederiz.

Yukarıdaki çizelgelerden aşağıdakileri gözleyebiliriz:

1- $(C_{p+i} - Z_{p+i}) + (C_{p+n+i} - Z_{p+n+i})$ daima 2 dir. ($i = 1, 2, \dots, n$ için) Bu sonuç sadece bu örneğe özgü değil geneldir. Verilen herhangi bir B bazı için

$$B^{-1}a_{p+i} = -B^{-1}a_{p+n+i}$$

dir. Çünkü

$$a_{p+i} = e_i = -a_{p+n+i} \quad \alpha_{p+i} = -\alpha_{p+n+i}$$

dir. (Bu sonuç 1 den 5 e kadar olan tüm çizelgelerde görülebilir.)

Bu yüzden $C_{p+i} - Z_{p+i} = 1 - C_B \alpha_{p+i} = 1 + C_B \alpha_{p+n+i}$ dir.

Buradan $(C_{p+i} - Z_{p+i}) + (C_{p+n+i} - Z_{p+n+i}) = 1 + C_B \alpha_{p+n+i} + 1 - C_B \alpha_{p+n+i} = 2$ dir.

$C_{p+i} - Z_{p+i}$ yi veya $C_{p+n+i} - Z_{p+n+i}$ yi bilirse diğerini algoritmanın 4. adımındaki hesaplamaları yapmadan bulabiliriz.

2- Eğer a_{p+i} bazda ise $B^{-1}a_{p+i} = \alpha_{p+i} = e_i$ dir. Burada $\alpha_{p+n+i} = -e_i$ olur. Sonuç olarak bizim, herhangi bir tablodaki bu iki kolonu hesaplamamıza gerek yoktur. Çünkü $\alpha_{p+i}(\alpha_{p+n+i})$ nin $+e_i$ yada $-e_i$ olması $a_{p+i}(a_{p+n+i})$ nin bazda olup

olmamasına göre değişir. Lineer bağımlı olduklarından dolayı her ikisinin de bazda olmayacağı bir durum yoktur.

3- Eğer a_{p+i} ve a_{p+n+i} nin her ikisi de temel olmayan kolonlar ise bu α_{p+i} ve α_{p+n+i} nin hesaplanması için yeterlidir.

Eğer $a_{p+i}(a_{p+n+i})$ bazı ile a_j nin yerini değiştirirsek bir sonraki tabloda temel olmayan k lar için $\hat{\alpha}_k$ kolonlarına ihtiyacımız olur öyle ki k, ya β_k 'ya ya da (her ikisi de temel olmayan ise) d_{1i} ve d_{2i} ye karşılık gelir. Buradan her iterasyonda baza giren α_{p+r} ya da α_{p+n+r} leri bazdan çıkan d_{1i} veya d_{2r} lere göre belirleyebiliriz.

$$\alpha_{p+r} \text{ ya da } \alpha_{p+n+r} = \begin{bmatrix} -\alpha_{1j}/\alpha_{rj} \\ \vdots \\ 1/\alpha_{rj} \\ \vdots \\ -\alpha_{nj}/\alpha_{rj} \end{bmatrix}$$

elde ederiz. Bu çok büyük hesapsal gelişme anlamına gelir.

Üstelik biz ne zaman bir temel d_{1r} veya d_{2r} ile temel olmayan d_{2r} veya d_{1r} yi yer değiştirirsek bazda sadece, sırasıyla, d_{1r} veya d_{2r} ye karşılık gelen i-inci satır -1 ile çarpılır ve tablodaki $C_k - Z_k$ satırı

$$\hat{C}_k - \hat{Z}_k = (C_k - Z_k) + 2\alpha_{rk}$$

şeklinde değişir ve

$$\hat{Z} = Z - 2W_{Br}$$

olur.

Bir optimal çözüm elde etmek için Simplex Yöntemi kullanılırken iterasyon sayısının az olması istenir. Bu noktada şu iki durumu gerçekleştirebiliriz.

1- Algoritmanın 2. adımında, ilk birkaç iterasyonda sadece temel olmayan β_j leri ele alırız. Hiç temel olmayan β_j yok ise temel d_{1i} ya da d_{2i} lerle yer değiştirebiliriz ve diğer temel olmayan değişkenleri ele alırız.

2- Baza girdirmek için uygun bir β_j değişkeninin değerini $C_j - Z_j > 0$ veya $C_j - Z_j < 0$ a göre arttırabilir ya da azaltabiliriz.

Eğer (duruma göre) azaltılır ya da arttırılır ise W_{Br}/α_{rj} nin dışında işaretinde kısıtlı bazı temel değişkenler sonunda negatif olacaktır.

En son gözlemden, eğer d_{1r} veya d_{2r} negatif olursa uygun d_{2r} veya d_{1r} sırasıyla baza girdirilebilir ve uygun değişken pozitif yapılabilir. Bu, tablonun $C_k - Z_k$ satırını etkiler ve r-inci satırı -1 ile çarpabiliriz.

Buraya kadar bahsedilenlere göre aşağıdaki gibi bazdan vektör çıkararak değişiklik yapabiliriz: 3. adımda bir r seçilir. $\hat{C}_j - \hat{Z}_j = C_j - Z_j + 2\alpha_{rj}$ hesaplanır. Eğer $C_j - Z_j$ ve $\hat{C}_j - \hat{Z}_j$ zıt işaretli ise uygun değişkeni belirler ve B_r yi bazdan çıkartırız. Başka bir deyişle tüm $C_k - Z_k$ ları $(C_k - Z_k) + 2\alpha_{rk}$ ile ve Z yi $Z - 2W_{Br}$ ile değiştiririz. Bazdaki a_{p+r} veya a_{p+n+r} vektörü ile r-inci satırdaki a_{p+n+r} veya a_{p+r} lerle sırasıyla yer değiştiririz. Temel değişkenler d_{1r} veya d_{2r} sırasıyla d_{2r} veya d_{1r} ile değiştirilir. r-inci satır -1 ile çarpılır. B_r yer değiştirme için artık göz önünde tutulmaz. 3. adım da çıkarmak için olası başka bir vektör bulunur ve bu adım tekrarlanır. Bir sonraki tabloyu elde etmek için gerekli dönüşümler yapılır. Bu yöntemin bitimi algoritmadaki gibidir.

Bir örnekle algoritmadaki değişikliklere göz atalım.

Örnek3.2.2

$$X = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \\ 1 & 5 \end{bmatrix} \quad Y = \begin{bmatrix} 1 \\ 1 \\ 2 \\ 3 \\ 2 \end{bmatrix}$$

olsun. İlk çizelgemiz Çizelge 3.7 olarak oluşturulur.

Çizelge 3.7 Örnek 3.2.2 için oluşturulan ilk çizelge

C_B	Bazdaki vektörler	W_B	α_1	α_2	α_3	α_4	α_5	α_6	α_7	α_8	α_9	α_{10}	α_{11}	α_{12}
1	a_3	1	1	1	1	0	0	0	0	-1	0	0	0	0
1	a_4	1	1	2	0	1	0	0	0	0	-1	0	0	0
1	a_5	2	1	3	0	0	1	0	0	0	0	-1	0	0
1	a_6	3	1	4	0	0	0	1	0	0	0	0	-1	0
1	a_7	2	1	5	0	0	0	0	1	0	0	0	0	-1
$C_k - Z_k$		9	-5	-15	0	0	0	0	0	2	2	2	2	2

Algoritmanın 2. adımındaki gibi bir β_2 seçeriz. Buna göre $r = 5$ buluruz.

$W_{B5}/\alpha_{52} = \min[W_{B_i}/\alpha_{i2}, \alpha_{i2} > 0] = \frac{2}{5}$ olarak belirlenir. Bir sonraki çizelge

C_B	Bazdaki vektörler	W_B	α_1	α_2
1	a_3	1	1	1
1	a_4	1	1	2
1	a_5	2	1	3
1	a_6	3	1	4
1	a_7	2	1	5
$C_k - Z_k$		9	-5	-15

şeklinde olur.

Algoritmadaki gibi a_7 ile a_{12} nin yer değiştirdiğine dikkat edelim. 3. adımdaki değişiklikte olduğu gibi

$$\hat{C}_2 - \hat{Z}_2 = -15 + 2 \times 5 = -5$$

hesaplanır.

β_2 nin $\frac{2}{5}$ arttığını ele alalım. a_7 nin a_{12} ile yer değiştirmesi ile 5. satırı -1 ile çarpıp tüm $C_k - Z_k$ lara $2\alpha_{52}$ ekleriz ve Z den $2W_B$ yi çıkartırız. Bu bize en son satırın yeni halini aşağıdaki gibi verir:

$C_k - Z_k$	$Z = 5$	-3	-5
-------------	---------	----	----

Sonra 5. satır hariç yer değiştirmek için yeni bir r ele alırsak. 3. adımda olduğu gibi $r=2$ ve $W_{B2}/\alpha_{22} = \frac{1}{2}$ buluruz.

$$\hat{C}_2 - \hat{Z}_2 = -5 + 2 \times 2 = -1$$

hesaplarız.

β_2 nin $\frac{1}{2}$ kadar arttığını düşünelim. a_4 ile a_9 yer değiştirir. 2. satır -1 ile çarpılır ve son satır değişir. Yeni satırı şu şekilde elde ederiz:

$C_k - Z_k$	$Z = 3$	-1	-1
-------------	---------	----	----

Daha sonra $r = 3$ olur ve $W_{Br}/\alpha_{r2} = \frac{2}{3}$ olur. $\hat{C}_2 - \hat{Z}_2 = -1 + 2 \times 3 = 5$ hesaplarız.

Sonuç olarak $\hat{C}_2 - \hat{Z}_2$ yi pozitif olarak buluruz. Bunun anlamı baza a_2 girecek, a_5 çıkacak demektir. Yeni çizelge;

Çizelge 3.8 a_2 ile a_5 in yer değiştirmesi ile elde edilen tablo

C_B	Bazdaki vektörler	W_B	α_1	α_5
1	a_3	1/3	2/3	-1/3
1	a_9	1/3	-1/3	2/3
0	a_2	2/3	1/3	1/3
1	a_6	1/3	-1/3	-4/5
1	a_{12}	4/3	2/3	5/3
$C_k - Z_k$		$Z=7/3$	-2/3	1/3

şeklindedir. Yeni çizelgede α_2 nin yerine α_5 olduğuna dikkat edelim. Bir sonraki iterasyonda baza girmesi için β_1 i seçeriz. $C_1 - Z_1 = -\frac{2}{3} < 0$ olduğundan 3. adımda $r = 1$, $W_{B1}/\alpha_{11} = \frac{1}{2}$ buluruz.

$$\hat{C}_1 - \hat{Z}_1 = -\frac{2}{3} + 2 \times \frac{2}{3} = \frac{2}{3}$$

hesaplarız. Burada a_3 ü bazdan çıkarır ve a_1 i girdiririz. Bu dönüşümlerden sonra

Çizelge 3.9 a_3 ile a_1 in yer değiştirmesi ile elde edilen çizelge

C_B	Bazdaki vektörler	W_B	α_3	α_5
1	a_1	1/2	3/2	-1/2
1	a_9	1/2	1/2	1/2
0	a_2	1/2	-1/2	1/2
1	a_6	1/2	1/2	-3/2
1	a_{12}	1	-1	2
$C_k - Z_k$		$Z=2$	1	0

şeklinde tablo elde edilir.

Şimdi temel olmayan değişkenler için tüm $C_j - Z_j$ ler pozitifdir ve biri çıkacak olanı içerir.

Temel olmayan a_4 , a_1 ve a_7 lere karşılık gelen $C_k - Z_k$ lar 2 dir ve temel olmayan a_8 ve a_{10} da sırasıyla 1 ve 2 dir. Böylece yöntemi bitirmiş oluruz. MINMAD regresyon doğrusu;

$$\hat{Y} = \frac{1}{2} + \frac{1}{2} X$$

şeklinde elde edilir.

3.3 Genel Lineer Programlama Problemi

Genelde (3.1.2) ye benzer problemler

$$\begin{aligned} \text{Min} \quad & C'W \\ & A^1 W \leq b^1 \\ & A^2 W \geq b^2 \\ & A^3 W = b^3 \end{aligned} \quad (3.3.1)$$

şeklinde ifade edilebilir. Burada $W_1, W_2, \dots, W_p$ lerin işaretlerinde kısıtlama yoktur.

$W_{p+1} \geq 0, \dots, W_n \geq 0$ dir. $A^k : r_k \times n$ matris $k = 1, 2, 3$ ve

$$A = \begin{bmatrix} A^1 \\ A^2 \\ A^3 \end{bmatrix}$$

$A : m \times n$ dir. $m = r_1 + r_2 + r_3$. Aynı şekilde $b^k : r_k \times 1$ vektör ve $C' : n \times 1$ vektördür.

Genelliği kaybetmeden $b^k \geq 0$ varsayabiliriz. Bu gibi problemler lineer programlama problemi olarak bilinir.

Bu gibi genel problemler için herhangi bir uygun çözüme sahip olmayabiliriz ($F = \emptyset$). F boş kümeden farklı olsa bile, $C'W$ alttan sınırlı olmak zorunda olmadığından bir optimal çözüm olmayabilir. Bu nedenle, ilk olarak bu problemi, tüm kısıtlamaları eşitlik halinde olan bir probleme dönüştürerek F nin boş küme olup olmadığını bulmaya çalışırız.

Eşitliklere uygun, “yapay değişken” denen değişkenleri ekleriz. M büyük, pozitif bir sabit olmak üzere

$$\sum_{j=1}^n a_{ij} W_j + W_{ai} = b$$

$$W_{ai} \geq 0, \quad C_{ai} = M > 0$$

olur. $A^1 W \leq b^1$, şeklindeki kısıtlamalar her birinin sol tarafına negatif olmayan bir W_{si} değişkeninin eklenmesiyle eşitlik durumuna dönüştürülür. Aynı şekilde $A^2 W \geq b^2$ şeklindeki kısıtlamalar, sol taraflarından negatif olmayan bir W_{si} değişkeninin çıkarılmasıyla eşitlik haline dönüştürülür. Bu değişkenler sırayla “gevşek değişken” ve “fazlalık değişken” olarak adlandırılır. Bu değişkenler için amaç fonksiyonunun katsayıları sıfırdır. Bu değişikliklerden sonra problemimiz aşağıdaki gibi olur

$$\min C'W + M \sum_{i=1}^{r_3} W_{ai}$$

$$\begin{bmatrix} A^1 & I & 0 & 0 \\ A^2 & 0 & -I & 0 \\ A^3 & 0 & 0 & I \end{bmatrix} \begin{bmatrix} W \\ W_s \\ W_a \end{bmatrix} = b$$

W_i işaretinde kısıtlama yok $i = 1, \dots, p$

$$W_i \geq 0 \quad i = p+1, \dots, n$$

$$W_{si} \geq 0, \quad W_{ai} \geq 0 \quad (3.3.2)$$

Burada $W_s : (r_1 + r_2) \times 1$ vektör ve $W_a : r_3 \times 1$ vektördür.

Şimdi bu problem algoritma (3.2.1) kullanılarak çözülebilir. Bu problemde gevşek ve fazlalık ve yapay değişkenlere karşılık gelen temel formda bir matris olur. Simpleks yönteminin sonunda üç olasılık karşımıza çıkar:

1- Algoritmadaki optimallik şartları sağlanır ve optimal tabloda bazı $W_{ai} > 0$ dır.

2- Algoritmadaki optimallik şartları sağlanır ve optimal tabloda hiçbir $W_{ai} > 0$ değildir.

3- Üçüncü adımdaki durumları sağlayacak hiçbir r bulamayabiliriz.

1-inci durumda orijinal problem için uygun çözüm yoktur. Eğer F boş kümeden farklı ise bir $W \in F$ nin amaç fonksiyonu değeri $C'W$ ve $W_{ai} = 0$ olacaktır. Bu aynı zamanda genişletilmiş problem için de doğrudur.

2-nci durumda, problem için en son tabloda W_{Bi} lerle verilen bir optimal çözüme sahip oluruz. Eğer bazdaki yapay değişkenlerin bazıları orijinal ya da gevşek ve fazlalık değişkenlerle yer değiştiremiyorsa bu, A matrisinin bunlara karşılık gelen satırının gereksiz olduğu anlamına gelir. A matrisinin rankı m den küçüktür.

3-üncü durum problemin sonlu minimumunun olmadığını gösterir.

Bir lineer programlama probleminin çözümü olan bu metod Charnes M- metodu olarak adlandırılır. Bir lineer programlama problemi için diğer bir metod da problemi iki aşamada çözer. Birinci aşamada; (3.3.2) deki kısıtlamalar ile $\sum_{i=1}^{r_3} W_{ai}$ nin minimizasyonunu göz önüne alırız. Tüm $W_{ai} \geq 0$ olduğundan $\min \sum W_{ai} = 0$ dır. Eğer $\sum W_{ai} > 0$ ise orijinal problem için uygun çözüm yoktur.

Birinci aşamanın sonunda tüm yapay değişkenler için $W_{ai} = 0$ ise ve bunlardan bazıları bazda ise F boş kümeden farklıdır. Bununla birlikte A nın satırlarından gereksiz olanlar olabilir. Tüm yapay değişkenler için $W_{ai} = 0$ ise ve tüm yapay değişkenler temel olmayan değişken ise F boş kümeden farklıdır ve gereksizlik yoktur. $\text{rank} A = m$ dir.

Birinci aşamada amaç fonksiyonunun değeri sıfır olduğunda problem (3.3.1) deki $C'W$ orijinal amaç fonksiyonu ile ikinci aşamaya geçeriz ve simpleks dönüşümleri ile devam ederiz. Bunun sonunda ya problem (3.3.1) için bir optimal çözüm elde ederiz ya da problem için sonlu optimum olmadığını buluruz. Bu metod İki Aşamalı Metod olarak bilinir.

Metodun bir başka uyarlaması gözlemlerden ortaya çıkar. Öyle ki; eğer herhangi bir iterasyonda mevcut bazın tersi varsa baza girecek vektör için α_{ij} leri hesaplayabiliriz. Bu yol bizim hesap yükümüzü önemli ölçüde azaltabilir. Simpleks yöntemini gözden geçirerek bu görüşün kullanılabilirliğini görebiliriz.

3.4. Minmad Regresyonda Dualite

$$\text{Min } C'W$$

$$AW = Y$$

$$W_{p+r} \geq 0, \quad r = 1, \dots, 2n$$

$$W_1, \dots, W_p \quad \text{işaretlerinde kısıtlama yok}$$

problemi verilsin. Bu problemin duali aşağıdaki şekilde tanımlanır

$$\text{Max } Y'\lambda$$

$$A'\lambda = \begin{bmatrix} X' \\ I \\ -I \end{bmatrix} \lambda \leq \begin{bmatrix} 0 \\ e' \\ e' \end{bmatrix}$$

$$\lambda \quad \text{işaretinde kısıtlama yok} \quad (3.4.1)$$

Burada $0 : p \times 1$ vektör, $e : 1 \times n$ vektör $e = (1, \dots, 1)$ ve $\lambda : n \times 1$ vektördür. Şimdi n tane $\lambda_1, \dots, \lambda_n$ değişkenine ve $2n+p$ tane kısıtlamaya sahibiz. Bu kısıtlamaların p -tanisi eşitlik, geri kalanı eşitsizlik şeklindedir.

Örnek 3.4.1:

Örnek 3.2.1 de ele alınan problemin duali şu şekilde ortaya çıkar

$$\text{Max } 4\lambda_1 + 3\lambda_2 + 4\lambda_3 + 5\lambda_4 + 5\lambda_5$$

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 \\ 3 & 2 & 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \\ \lambda_5 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad (1)$$

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \\ \lambda_5 \end{bmatrix} \leq \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \quad (2)$$

$$\begin{bmatrix} -1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \\ \lambda_5 \end{bmatrix} \leq \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \quad (3)$$

$\lambda_1, \dots, \lambda_n$ işaretinde kısıtlama yok

İkinci ve üçüncü kısıtlama grupları $-1 \leq \lambda_i \leq 1$ iken eşitlik şeklinde ifade edilebilir.

İlk olarak bu problem ile Problem (3.1.2) arasındaki ilişkiyi saptarız ve dual problemin boyutunu nasıl küçülteceğimizi açıklarız.

Sonuç 3.4.1 : Problem (3.1.2) nin herhangi bir uygun çözümü ve dualinin herhangi bir çözümü için

$$Y'\lambda \leq C'W$$

dir.

İspat: λ dual için bir çözüm olsun. Böylece

$$A'\lambda \begin{cases} = \\ \leq \\ \leq \end{cases} C \quad (3.4.2)$$

(3.4.2) nin her iki tarafının W' ile çarpılması ile eşitsizlik yön değiştirmez. Çünkü $r = 1, \dots, 2n$ için $W_{p+r} \geq 0$ dır ve $W_1, \dots, W_p$ lere karşılık gelen eşitlikler vardır. W nun Problem (3.1.2) için bir uygun çözüm olmasından dolayı $W'A' = Y'$ dür. Buradan $Y'\lambda \leq C'W$ olur. Problem (3.1.2) de bir optimal çözüme sahip olduğumuzu biliyoruz. Şimdi Problem (3.1.2) ve dual problem için optimal olacak şekilde W^* ve λ^* ele alacağız.

Sonuç 3.4.2 : Eğer W^* Problem (3.1.2) için uygun, λ^* dual için bir çözüm ise ve $C'W^* = Y'\lambda^*$ ise λ^* dual problem için optimal, W^* Problem (3.1.2) için optimaldir.

İspat: Sonuç 3.4.1 den, herhangi bir λ çözümü için $C'W^* \geq Y'\lambda$ olduğunu biliyoruz. Fakat $C'W^* = Y'\lambda^*$ dır. Buradan, $Y'\lambda^* \geq Y'\lambda$ ya da λ^* dual problem için bir optimaldir sonucu çıkar. Aynı şekilde W^* in optimal olduğunu da gösterebiliriz. Bu sonucun tersi, aşağıdaki sonucun ispatında elde ediliyor. Bu sonuçta her iki problem de optimal çözüme sahip ve her iki problemin amaç fonksiyonlarının değerleri birbirine yakındır.

Sonuç 3.4.3: Problem (3.1.2) için bir W_B optimal temel uygun çözüm ele alalım. $\lambda' = C_B B^{-1}$ dual problem için bir optimal çözümdür ve amaç fonksiyonun değerleri birbirine yakındır.

İspat : W_B , Problem (3.1.2) için optimal olduğundan dolayı tüm kısıtsız değişkenler için $C_j - Z_j = 0$ ve tüm kısıtlı pozitif değişkenler için $C_j - Z_j \geq 0$ dır. Ayrıca tüm kısıtsız değişkenler için $Z_j = C_B B^{-1} a_j = C_j$ dir ve $j = 1, \dots, p$ için $C_j = 0$ dır.

Bununla birlikte

$$C_B B^{-1} a_j = C_B B^{-1} e_r \leq C_j = 1 \quad j = p+r, \quad r = 1, \dots, n$$

ve

$$C_B B^{-1} a_j = C_B B^{-1} (-e_r) \leq C_j = 1 \quad j = p+n+r, \quad r = 1, \dots, n$$

$\lambda' = C_B B^{-1}$ Problem (3.4.1) in tüm kısıtlamalarını sağlar.

Şimdi biz aşağıdaki dual amaç fonksiyonunu ele alacağız

$$Y'\lambda = \lambda'Y = C_B B^{-1}Y$$

Fakat $B^{-1}Y = W_B$ dir. Bu yüzden $Y'\lambda = C_B W_B = Z$ olur. Sonuç(3.4.2) den λ Problem (3.4.1) için optimaldir. Sonuç(3.4.3) den dolayı Problem (3.1.2.) nin dualini çözebilir ve bir optimal çözüm bulabiliriz.

3.5. Lineer Programlamada Dualite

Aşağıdaki gibi bir lineer programlama problemi verilsin

$$\text{Min } C'W$$

$$AW \geq b$$

$$W \geq 0$$

(3.5.1)

Bu probleme primal problem diyelim. Bu problemin duali aşağıdaki şekilde tanımlanır

$$\begin{aligned} \text{Max } b'U \\ A'U &\leq C \\ U &\geq 0 \end{aligned} \quad (3.5.2)$$

Not 3.5.1 : Her bir primal kısıtlama için dual problemde bir değişken tanımlandı ve primal problemdeki her değişken için bir dual kısıtlama oluşturuldu. Primal problemde eşitsizlikler $\geq$ şeklinde, dual problemde $\leq$ şeklindedir. Her iki problemde de negatif olmayan değişkenler vardır.

Not 3.5.2: Bir lineer programlama modelinde

$$a_{i1}W_1 + \dots + a_{in}W_n \leq b_i$$

gibi bazı kısıtlamalar

$$-a_{i1}W_1 - \dots - a_{in}W_n \geq -b_i$$

ifadesine eşittir. Aynı şekilde

$$a_{i1}W_1 + \dots + a_{in}W_n = b_i$$

gibi eşitlik şeklindeki bazı kısıtlamalar

$$a_{i1}W_1 + \dots + a_{in}W_n \geq b_i$$

$$-a_{i1}W_1 - \dots - a_{in}W_n \geq -b_i$$

şeklinde iki kısıtlamaya eşittir. Aynı zamanda eğer amaç fonksiyonu $C'W$ minimum yerine maksimum yapılacaksa $-C'W$ nun minimizasyonu ele alınır.

Eğer bazı W_j değişkenlerinin işaretinde kısıtlama yok ise, $W_j^1 \geq 0$, $W_j^2 \geq 0$ olmak üzere, W_j yi $W_j = W_j^1 - W_j^2$ şeklinde ifade edebiliriz.

Böylece herhangi bir lineer programlama problemi Problem (3.5.1) şekline dönüştürülebilir. Bununla birlikte aşağıdaki sonuca göre herhangi bir lineer programlama probleminin dualini, (3.5.1) formuna getirmeden yazmak mümkündür.

Sonuç 3.5.1 : Dual problemin duali primal modeldir.

İspat: Dual problem (3.5.2) yi aşağıdaki şekilde ifade edebiliriz

$$\begin{aligned} \text{Min } & -b'U \\ & -A'U \geq -C \\ & U \geq 0 \end{aligned}$$

Şimdi bu problemin dualini yazalım

$$\begin{aligned} \text{Max } & -C'W \\ & -AW \leq -b \\ & W \geq 0 \end{aligned}$$

Görüldüğü gibi bu ifade problem (3.5.1) e eşittir.

Sonuç 3.5.2 : Eğer primal problemdeki i-inci kısıtlama eşitlik şeklinde ise i-inci dual değişkenin işaretinde kısıtlama yoktur.

İspat: Problemi (3.5.1) formunda yazmak için eşitliğin yerine iki eşitsizlik yazacak olursak

$$a_{i1}W_1 + \dots + a_{in}W_n = b_i$$

eşitliğine karşılık

$$a_{i1}W_1 + \dots + a_{in}W_n \geq b_i$$

$$-a_{i1}W_1 - \dots - a_{in}W_n \geq -b_i$$

eşitsizliklerini elde ederiz.

Şimdi, dual problemde bu eşitsizliklere karşılık gelen iki değişkenimiz vardır. Bu değişkenler U_i^+ ve U_i^- şeklinde gösterilsin. U_i^+ ya karşılık gelen dual kolon primaldeki i-inci satırın transpozudur ve U_i^- ye karşılık gelen dual kolon ise primaldeki i-inci satırın transpozunun negatifidir. Bununla birlikte U_i^+ nin amaç fonksiyonu katsayısı b_i , U_i^- ninki de $-b_i$ dir. Burada U_i^+ ve U_i^- yerine U_i yazıldığında U_i nin işaretinde kısıtlama olmaz. U_i ye karşılık gelen kolon primal problemin i-inci satırının transpozu olur ve b_i amaç fonksiyonu katsayısı olur.

Buradan, primal problemde eşitlik şeklinde olan i-inci satıra karşılık gelen değişkenin, dual problemde işaretinde kısıtlama olmayan değişken olduğu görülür.

Sonuç 3.5.3 : Eğer primal problemdeki j-inci değişkenin işaretinde kısıtlama yok ise dual problemdeki j-inci kısıtlama eşitlik şeklindedir.

Şimdi Problem (3.1.2) nin duali olan (3.5.1) e dönelim. Dikkat edecek olursak bu dual bir lineer programlama probleminin dualinin genel tanımına uygundur.

Sonuç 3.5.4 : Eğer W Problem (3.5.1) için, U Problem (3.5.2) için uygun bir çözümse

$$C'W \geq b'U$$

dur.

Sonuç 3.5.5 : Eğer W^* (3.5.1) için, U^* (3.5.2) için uygun bir çözüm ise ve $C'W^* = b'U^*$ ise W^* ve U^* kendi problemleri için optimaldirler.

Sonuç 3.5.6 : Problem (3.5.1) ya da Problem(3.5.2) den biri optimal çözüme sahip ise diğeri de bir optimal çözüme sahiptir ve amaç fonksiyonlarının değerleri eşittir.

Not 3.5.3 : Sonuç 3.5.3 ün ispatı Sonuç 3.5.5 in ispatına benzer. Primal problem için optimal olan W_B nin bir temel uygun çözüme karşılık geldiği gösterilir ve $C_B B^{-1}$ dual problem için bir optimal çözüm verir.

Bu sonuç Problem (3.1.2) nin dualinin çözülmesi ile bir optimal çözüm elde edilmesinde çok kullanışlıdır. Genelde bir lineer programlama probleminin sonlu optimuma sahip olmak zorunda değildir. Bu gibi durumlarda problem sınırsız çözüme sahiptir deriz. Aşağıdaki sonuç böyle durumlar için verilir.

Sonuç 3.5.7 : Eğer primal(dual) problem sınırsız bir çözüme sahip ise dual(primal) problemin uygun çözümü yoktur.

Not 3.5.4 : Aşağıdaki problemde görüldüğü gibi her iki problem de uygun çözüme sahip olmayabilir.

$$\begin{aligned} \text{Min} \quad & -CX_1 - CX_2 \\ & aX_1 - aX_2 \geq b \\ & -aX_1 + aX_2 \geq b \\ & X_1, X_2 \geq 0 \end{aligned}$$

ve

$$\begin{aligned} \text{Max} \quad & bU_1 + bU_2 \\ & aU_1 - aU_2 \leq -C \\ & -aU_1 + aU_2 \leq -C \\ & U_1, U_2 \geq 0 \end{aligned}$$

Açıkça görülüyor ki her iki problem de uygun çözüme sahip değildir ve bunlar birbirlerinin dualidir.

Primal ve dual problemler için optimal çözümler sağlayan ve çok kullanışlı olan özellik, primal artan değişkenler ile dual değişkenler ve primal değişkenler ile dual gevşek değişkenler arasındaki tamamlayıcılık özeliğidir. Bu tamamlayıcılığa dayanarak dual ve primal problemlere gevşek ve artan değişkenler ekleriz. Daha sonra aşağıdaki problemi elde ederiz.

$$\text{Min } C'W$$

$$\begin{bmatrix} A & -I \end{bmatrix} \begin{bmatrix} W \\ W_s \end{bmatrix} = b$$

$$W, W_s \geq 0 \quad (3.5.3)$$

ve

$$\text{Max } b'U$$

$$\begin{bmatrix} A' & I \end{bmatrix} \begin{bmatrix} U \\ U_s \end{bmatrix} = C$$

$$U, U_s \geq 0 \quad (3.5.4)$$

Sonuç 3.5.8 : Eğer artan değişken W_{si} , herhangi bir optimal çözümde pozitif olan Problem (3.5.3) ün kısıtlamasına eklenirse Problem (3.5.4) de i-inci değişken U_i , (3.5.4) ün her optimal çözümü için sıfır olur. Diğer yandan, eğer U_i (3.5.4) ün herhangi bir optimal çözümünde pozitif ise (3.5.4) ün optimal çözümü için $W_{si} = 0$ dır.

İspat: $AW - IW_s = b$ olsun. Her iki tarafın $U' \geq 0$ ile çarpılması ile aşağıdaki ifade elde edilir

$$U'AW + U'W_s = U'b = b'U \quad (3.5.5)$$

Bununla birlikte Problem(3.5.1) ve (3.5.2) den $C'W \geq U'AW \geq U'b$ dir ve $W, U \geq 0$ dır. W ve U nun sırasıyla (3.5.3) ve (3.5.4) için optimal çözümler olduğunu varsayarsak $C'W = U'AW = U'b$ olur. Bu yüzden (3.5.5) $U'W_s = U'W_s = 0$ şeklinde ifade edilir. $W, U \geq 0$ olduğundan U_i ve W_{si} nin ikisi birlikte pozitif olamaz. Bu yüzden tüm i -ler için $W_{si} \cdot U_i = 0$ şeklinde ifade edebiliriz. W ve U bu problemler için optimal çözümlerdir. Aynı şekilde optimal W ve U ya karşılık gelen tüm j -ler için $W_j U_{sj} = 0$ olduğunu da gösterebiliriz.

Bu özellik dual linear programlama probleminin tamamlayıcı gevşeklik özeliği olarak bilinir.

Problem (3.1.2) ye dönecek olursak, dual problemin Problem (3.1.2) den daha fazla satıra sahip olduğunu buluruz. Bu nedenle duali çözerken daha büyük bir baz ele almamız gerekir. Ayrıca eşitsizlik şeklindeki kısıtlamaları dualde eşitlik yapmak için fazladan $2n$ tane işaretinde kısıtlı değişken eklememiz gerekir. Bununla birlikte, aşağıdaki gibi, problemi çözerken hesaplama sayısını azaltabiliriz.

$V = \lambda + e$, $Y'\lambda = Y'V - Y'e$ ve $Y'e$ bir sabit olsun. Böylece $Y'\lambda$ nin minimizasyonu $Y'V$ nin minimizasyonu ile aynı olur. Şimdi

$$A'\lambda = A'(V - e)$$

$$= \begin{bmatrix} X'V - X'e \\ V - e \\ -V + e \end{bmatrix}$$

olur. Böylece kısıtlamalar şu duruma gelir

$$X'V = X'e$$

$$V \leq e + e$$

$$-V \leq e - e$$

ya da

$$\begin{aligned} X'V &= X'e \\ 0 \leq V &\leq 2e \end{aligned}$$

Buradan da aşağıdaki problem elde edilir

$$\begin{aligned} \text{Max } Y'V \\ X'V &= X'e \\ 0 \leq V &\leq 2e \end{aligned} \quad (3.5.6)$$

Şimdi burada sadece p tane lineer eşitlik vardır. Çünkü X' : $p \times n$ matristir. Fakat şimdi, V alttan $2e$ ile sınırlı olacak şekilde bir sınırlama getirmeliyiz. Neyse ki biz simpleks yöntemini geliştirme yönünde bir değişiklik yapabiliriz öyle ki bu üstten sınırlı kısıtlamalar açıkça eklemekten de kısıtlama olarak kalırlar.

Örnek 3.5.1 :

Örnek 3.2.1 i ele alalım. Dönüşümlerden sonra $\lambda_1, \dots, \lambda_5$ ler $V_1, \dots, V_5$ şekline dönüşür. $V = \lambda + e$ dir.

$$\text{Max } 4V_1 + 3V_2 + 4V_3 + 5V_4 + 5V_5$$

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 \\ 3 & 2 & 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \\ V_3 \\ V_4 \\ V_5 \end{bmatrix} = \begin{bmatrix} 5 \\ 15 \\ 11 \end{bmatrix}$$

$$0 \leq V_i \leq 2$$

3.6. MINMAD İçin Alternatif Bir Formülasyon

Daha önce Bölüm 3.1 de ele alınan problem aşağıdaki gibi yeniden formüle edilebilir

$$\begin{aligned} &\text{Min } d'e \\ &-d \leq Y - X\beta \leq d \\ &d \geq 0 \\ &\beta \text{ işaretinde kısıtlama yok} \end{aligned} \quad (3.6.1)$$

Bu oluşum, (3.6.1) deki kısıtlamaları iki kısımda yazarak ve β_1 ve β_2 negatif olmayan vektörler olmak üzere $\beta = \beta_1 - \beta_2$ şeklinde tanımlayarak matris formunda yazılabilir. Böylece

$$\begin{aligned} &\text{Min } d'e \\ &\begin{pmatrix} X & -X & -I \\ -X & X & -I \end{pmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ d \end{bmatrix} \leq \begin{pmatrix} Y \\ -Y \end{pmatrix} \\ &\beta_1, \beta_2, d \geq 0 \end{aligned} \quad (3.6.2)$$

elde edilir. Bu problemde $2n$ -tane kısıtlama ve $2p+n$ tane değişken vardır. Biz yine β nın yansız tahmin edicisini elde etmek için bu formülü kullanacağız. Genelde, hatalar simetrik bir dağılıma sahip ve $E(\varepsilon)=0$ olsa bile MINMAD probleminin çözümü β nın yansız tahmin edicisini vermez.

Örnek 3.6.1:

$Y' = [Y_1, Y_2]$, $X' = [1, 1]$ ve $\varepsilon' = (\varepsilon_1, \varepsilon_2)$ alalım. Lineer model $Y = X\beta + \varepsilon$ şeklindedir. $\beta = 1$ ve ε_i aşağıdaki şekilde dağılıma sahip olsun.

$$\varepsilon_i = \begin{cases} -1 & 1/2 \text{ olasılık} \\ +1 & 1/2 \text{ olasılık} \end{cases} \quad i = 1, 2$$

Lineer programlama kullanılarak $[\varepsilon_1, \varepsilon_2]$ nin farklı olasılıklarına uygun şekilde (3.6.1) in çözülmesiyle β nin tahmin edicisinin bulunması istenir. Örneğin $\varepsilon_1 = \varepsilon_2 = -1$ iken problemimiz şu şekilde olur:

$$\begin{aligned} & \text{Min } d_1 + d_2 \\ & -\begin{pmatrix} d_1 \\ d_2 \end{pmatrix} \leq \begin{pmatrix} 0 \\ 0 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \end{pmatrix} \beta \leq \begin{pmatrix} d_1 \\ d_2 \end{pmatrix} \\ & d_1, d_2 \geq 0 \end{aligned}$$

Burada $d_1 = d_2 = \hat{\beta} = 0$ iken minimum tek şekilde bulunur. Aynı şekilde $\varepsilon_1 = \varepsilon_2 = 1$ iken problemimiz

$$\begin{aligned} & \text{Min } d_1 + d_2 \\ & -\begin{pmatrix} d_1 \\ d_2 \end{pmatrix} \leq \begin{pmatrix} 2 \\ 2 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \end{pmatrix} \beta \leq \begin{pmatrix} d_1 \\ d_2 \end{pmatrix} \\ & d_1, d_2 \geq 0 \end{aligned}$$

şeklindedir. Burada $\hat{\beta} = 0$ ve $d_1 = d_2 = 0$ tek şekildedir. Fakat $\varepsilon_1 = -1$ ve $\varepsilon_2 = +1$ veya $\varepsilon_1 = +1$ ve $\varepsilon_2 = -1$ e uygun oluşturduğumuz problem tek optimal çözüme sahip olmaz. Bu iki durumda $0 \leq \hat{\beta} \leq 2$ uygun problemin bir optimal çözümüdür. Örneğin $\varepsilon_1 = -1$ ve $\varepsilon_2 = +1$ iken

$$\begin{aligned} & \text{min } d_1 + d_2 \\ & -\begin{pmatrix} d_1 \\ d_2 \end{pmatrix} \leq \begin{pmatrix} 0 \\ 2 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \end{pmatrix} \beta \leq \begin{pmatrix} d_1 \\ d_2 \end{pmatrix} \\ & d_1, d_2 \geq 0 \end{aligned}$$

şeklinde problem elde ederiz. Bu iki durumda $\hat{\beta} = 0$ ı ele alalım. Aşağıdaki tabloyu oluştururuz

Durum	ε_1	ε_2	$\hat{\beta}$
1	-1	-1	0
2	1	1	2
3	-1	1	0
4	1	-1	0

Buradan $\beta_1 = 1$ yanlı olacak şekilde $E(\hat{\beta}) = \frac{1}{2}$ elde edilir. Gerçekte, eğer $\hat{\beta}_1 = 1$ ele alırsak 3. durumda ve $E(\hat{\beta}) = 1 = \beta$ elde ederiz. Böylece, lineer programlama probleminin çözümü tek olmadığında yanlı tahmin ediciler elde etme tehlikesi vardır, sonucuna ulaşırız.

3.7. MINMAD Kullanarak Yansız Tahminler Bulma

MINMAD regresyon kullanarak yansız tahminler elde etmek için ε un simetrik dağılıma sahip ve $E(\varepsilon) = 0$ olduğunu kabul edelim.

Tanım 3.7.1 : β nın β_0 ile gösterilen bir tahmin edicisini tanımlarız, öyle ki herhangi bir ε için $\beta - \beta_0(\varepsilon) = -(\beta - \beta_0(-\varepsilon))$ olur.

Örnek 3.7.1 :

$\beta_0 = (X'X)^{-1} X'Y$ en küçük kareler tahmin edicisidir.

$$\beta_0(\varepsilon) = (X'X)^{-1} X'(X\beta + \varepsilon)$$

$$= \beta + (X'X)^{-1} X'\varepsilon$$

$$\beta_0(-\varepsilon) = (X'X)^{-1} X'(X\beta - \varepsilon)$$

$$= \beta - (X'X)^{-1} X'\varepsilon$$

Burada $\beta - \beta_0(\varepsilon) = -(\beta - \beta_0(-\varepsilon))$ olur. β_0 antisimetriktr.

Örnek 3.7.2 :

$$\beta_0 = X_A^{-1} Y_A$$

olsun. Burada; $X_A : p \times p$, rankı p ve X in alt matrisi. $Y_A : p$ gözlemlerine karşılık gelen vektör.

$$\beta_0(\varepsilon) = X_A^{-1}(X_A \beta + \varepsilon_A) = \beta + X_A^{-1} \varepsilon_A$$

$$\beta_0(-\varepsilon) = X_A^{-1}(X_A \beta - \varepsilon_A) = \beta - X_A^{-1} \varepsilon_A$$

β_0 in antisimetrik olduğu görülebilir.

Şimdi de Problem (3.6.2) yi ele alalım. Bu problemi aşağıdaki şekilde ifade edebiliriz.

Min $d'e$

$$\begin{pmatrix} -X & X & -I \\ X & -X & -I \end{pmatrix} \begin{bmatrix} \beta_2 \\ \beta_1 \\ d \end{bmatrix} \leq \begin{pmatrix} Y \\ -Y \end{pmatrix} \quad (3.7.1)$$

Şimdi herhangi β_0 antisimetrik tahmin edicisini kullanarak yeni değişkenler tanımlayabiliriz.

$$\bar{\beta}_1 = \beta_1 + \beta_0^{(2)} \quad , \quad \bar{\beta}_2 = \beta_2 + \beta_0^{(1)}$$

Burada $\beta_0 = \beta_0^{(1)} - \beta_0^{(2)}$ öyle ki $\beta_0^{(1)}, \beta_0^{(2)} \geq 0$ dır. Problem (3.6.2) ve (3.7.1) aşağıda verilen (3.7.2) ve (3.7.3) e eşit şekilde ifade edilebilir.

Min $d'e$

$$\begin{pmatrix} -X & X & -I \\ X & -X & -I \end{pmatrix} \begin{bmatrix} \bar{\beta}_1 \\ \bar{\beta}_2 \\ d \end{bmatrix} \leq \begin{pmatrix} -Y & +X\beta_0 \\ Y & -X\beta_0 \end{pmatrix}$$

$$\bar{\beta}_1, \bar{\beta}_2, d \geq 0 \quad (3.7.2)$$

Min $d'e$

$$\begin{pmatrix} -X & X & -I \\ X & -X & -I \end{pmatrix} \begin{bmatrix} \bar{\beta}_2 \\ \bar{\beta}_1 \\ d \end{bmatrix} \leq \begin{pmatrix} Y & -X\beta_0 \\ -Y & +X\beta_0 \end{pmatrix}$$

$$\bar{\beta}_1, \bar{\beta}_2, d \geq 0 \quad (3.7.3)$$

Problem (3.7.2) nin satırlarını ve sütunlarını tekrar düzenleyerek Problem (3.7.3) ü elde edebiliriz.

Aşağıdaki gibi elde edilen yeni bir β_0^* tahmin edicisi tanımlarız:

Problem (3.7.2) ve (3.7.3) eşit olasılıkla seçilir. P_1 ve P_2 sırasıyla Problem (3.7.2) ve (3.7.3) ü gösterebilir. P_1 ve P_2 nin farklı temel optimal çözümlerinin olduğunu varsayalım. Bazı zamanlar bu problemler çözmek için seçilir, tüm farklı optimal çözümlerden verilen olasılıklarla sonlu bir optimal çözüm seçeriz. Örneğin seçilen problem için optimal çözüm $\bar{\beta}_1^*$ ve $\bar{\beta}_2^*$ ile verilsin. Daha sonra

$$\beta_0^* = \bar{\beta}_1^* - \bar{\beta}_2^* + \beta_0$$

tanımlanır.

Sonuç 3.7.1 : β_0^* , β_0 ın yansız tahmin edicisidir.

İspat:

$$\begin{aligned} E(\beta - \beta_0^*) &= E[\beta - (\bar{\beta}_1^* - \bar{\beta}_2^* + \beta)] \\ &= E[\beta - \beta_0] - E[\bar{\beta}_1^* - \bar{\beta}_2^*] \end{aligned}$$

Şimdi $E(\beta_0) = \beta$ dir. Çünkü β_0 antisimetriktir ve ε simetriktir. Bu yüzden $E(\beta_0 - \beta_0^*) = -E\{[\bar{\beta}_1^* - \bar{\beta}_2^*|\varepsilon] + [\bar{\beta}_1^* - \bar{\beta}_2^*|-\varepsilon]\}$ olur. Bu noktada ε un aldığı herhangi bir ε değeri için sağ-yan elemanı olduğu gözlenir.

$$\begin{aligned} Y(-\varepsilon) - X\beta_0(-\varepsilon) &= X\beta - \varepsilon - X\beta_0(-\varepsilon) \\ &= -\varepsilon + X[\beta - \beta_0(-\varepsilon)] \\ &= -\varepsilon - X[\beta - \beta_0(\varepsilon)] , \beta_0 \text{ antisimetrik} \\ &= -Y(\varepsilon) + X\beta_0(\varepsilon). \end{aligned}$$

Kısıtlamaların farklı sıralarda yazıldığı durumu saymazsak, ε , ε değerini aldığı anda Problem (3.7.2); ε , $-\varepsilon$ değerini aldığı anda Problem (3.7.3) aynıdır. Bunun için ε , ε değerini aldığı anda Problem P_1 den $E(\bar{\beta}_1^* - \bar{\beta}_2^*|\varepsilon, P_1)$ elde edilir, aynı şekilde ε , $-\varepsilon$ değerini aldığı anda Problem P_2 den $E(\bar{\beta}_2^* - \bar{\beta}_1^*|-\varepsilon, P_2)$ elde edilir.

Şimdi

$$\begin{aligned} E(\beta - \beta_0^*) &= -E\{[\bar{\beta}_1^* - \bar{\beta}_2^*|\varepsilon, P_1]_{\frac{1}{2}} + [\bar{\beta}_1^* - \bar{\beta}_2^*|\varepsilon, P_2]_{\frac{1}{2}} \\ &\quad + [\bar{\beta}_1^* - \bar{\beta}_2^*|-\varepsilon, P_1]_{\frac{1}{2}} + [\bar{\beta}_1^* - \bar{\beta}_2^*|-\varepsilon, P_2]_{\frac{1}{2}}\} \end{aligned}$$

Yukarıda yapılanları kullanarak aşağıdaki eşitliği elde ederiz.

$$E(\beta - \beta_0^*) = E\{0\} = 0$$

Bu yüzden β_0^* , β nın yansız tahmin edicisidir.

Böylece, eğer bir antisimetrik β_0 tahmini var ise iki lineer programlama modeli tanımlarız ve yansız tahmin ediciler elde etmek için eşit olasılıkla bir tanesini seçeriz.

3.8 Maksimum-Likelihood ve MINMAD Tahminleri

Hataların, σ parametrelili üstel dağılıma sahip olduğunu varsayalım.

$$f(\varepsilon_i) = (2\sigma)^{-1} e^{-|\varepsilon_i|/\sigma} \quad -\infty < \varepsilon_i < \infty$$

Maksimum-Likelihood metodunun β nın tahminindeki uygulamasının

$$\sum_{i=1}^n |\varepsilon_i|$$

nin minimizasyonu şeklinde ifade edildiğini biliyoruz. Bu yüzden MINMAD tahmin edicisini ele almada bir sakınca yoktur ve lineer programlama uygulaması bu gibi durumlarda kullanışlıdır. Birçok araştırmacı, hataların normallik varsayımları ihmal edildiği belli durumlarda en küçük kareler tahmininin daha üstün olduğunu bulmuşlardır.

3.9 Mutlak Sapmaların Maksimumunu Minimum Yapma (MINMAXAD)

Bu bölümde hataların maksimumunun minimizasyonunu ele alacağız ve bir lineer programlama problemi gibi lineer görünmeyen problem formüle edeceğiz. Ayrıca aynı problemi farklı yollardan ele alacağız.

Bu problem aşağıdaki şekilde ifade edilir:

$$\min_{\beta} \max_{1 \leq i \leq n} \left| Y_i - \sum_{j=1}^p X_{ij} \beta_j \right|$$

(3.9.1)

β işaretinde kısıtlama yok

(1) Eğer $Y, X_1, \dots, X_p$ nin bir lineer kombinasyonu ise $X\beta - Y = 0$ ın bir β^* çözümü vardır ve bu çözüme karşılık gelen ifade

$$\max_{1 \leq i \leq n} \left| Y_i - \sum_{j=1}^p X_{ij} \beta_j^* \right|.$$

Burada β^* (3.9.1.) için bir optimal çözümdür.

(2) Eğer herhangi bir j -için $X_j = 0$ ise bu kolonu ve uygun β_j yi problemten çıkarabiliriz. Buradan şu tahminleri yaparız:

(1) $Y, X_1, \dots, X_p$ nin lineer kombinasyonu değildir.

(2) Tüm $j = 1, \dots, p$ için $X_j \neq 0$ dir.

$$d = \max_{1 \leq i \leq n} \left| Y_i - \sum_{j=1}^p X_{ij} \beta_j \right|$$

olsun. d nin β ya bağlı olduğu görülüyor. Bu varsayımlar altında X_j lerin normal olduğunu varsayabiliriz. Böylece

$$\sum_{j=1}^p X_{ij}^2 = 1 \quad \text{tüm } i = 1, \dots, n \text{ için}$$

olur. Sonra

$$Y_i - \sum_{j=1}^p \beta_j X_{ij} = 0, \quad i = 1, \dots, n \quad (3.9.2)$$

ifadesi R^p de β noktasından meydana gelen bir n-hiperdüzlemin eşitliğini verir. Eğer β , R^p de herhangi bir nokta ise $\left| Y_i - \sum_{j=1}^p \beta_j X_{ij} \right|$, i-inci hiperdüzlemden noktaya olan uzaklıktır. d, bu uzaklıkların maksimumudur. Problem; maksimal uzaklığı minimum yapan $\beta \in R^p$ noktasını bulma problemidir. Bu gözlemler MINMAD regresyon için de doğrudur.

Problem (3.9.1) i bir lineer programlama problemi gibi tekrar ifade edebiliriz

Min d

$$d + \sum_{j=1}^p \beta_j X_{ij} \geq Y_i$$

$$d - \sum_{j=1}^p \beta_j X_{ij} \geq -Y_i$$

$$i = 1, \dots, n$$

$$d, \beta_j \text{ işaretinde kısıtlama yok} \quad (3.9.3)$$

Bu problemde d nin işaretinde kısıtlama olup olmaması açık değildir fakat d negatif olmayacaktır. Ayrıca d alttan sınırlıdır. Böylece varsayım (2) den $d > 0$, bu problem için bir optimal çözüme sahip oluruz. Değişken dönüşümleri aşağıdaki gibidir:

$$b_0 = \frac{1}{d}, \quad b_j = -\frac{\beta_j}{d}, \quad j = 1, \dots, p$$

Böylece denk bir problem elde etmiş oluruz.

$$\text{Max } b_0$$

$$b_0 Y_i + \sum_{j=1}^p b_j X_{ij} \leq 1$$

$$-b_0 Y_i - \sum_{j=1}^p b_j X_{ij} \leq 1$$

$$i = 1, \dots, n$$

$$b_0, b_j \text{ işaretinde kısıtlama yok} \quad (3.9.4)$$

Şimdi her bir kısıtlamaya gevşek değişkenler ekledikten ve $-b_0$ ı minimum yaptıktan sonra bu problemi çözmek için simpleks metodunu uygulayabiliriz.

Ayrıca, bazın büyüklüğünü azaltacak bir değişiklik de mümkündür. Problem (3.9.4) ü aşağıdaki gibi de yazabiliriz

$$\text{Max } b_0$$

$$b_0 Y_i + \sum_{j=1}^p b_j X_{ij} + b_{si} = 0$$

$$-1 \leq b_{si} \leq 1, \quad i = 1, \dots, n$$

$$b_0 \text{ işaretinde kısıtlama yok} \quad (3.9.5)$$

Dikkat edilecek olursa, bu problem, değişkenlerin $W_{si} = b_{si} + 1$, $0 \leq W_{si} \leq 2$ şeklinde dönüştürülmesinden sonra, sınırlı değişkenler metodu ile de çözülebilir. Böylece Problem(3.9.5) aşağıdaki gibi olur

$$\text{Max } b_0$$

$$b_0 Y_i + \sum_{j=1}^p b_j X_{ij} + W_{si} = 1$$

$$0 \leq W_{si} \leq 2$$

$$i = 1, \dots, n$$

$$(3.9.6)$$

Problem (3.9.3) ün dualine dönersek aşağıdaki matris formunu elde ederiz

$$\begin{aligned}
 & \text{Max } Y'U_1 - Y'U_2 \\
 & [X', -X'] \begin{pmatrix} U_1 \\ U_2 \end{pmatrix} = 0 \\
 & U_1'e + U_2' = 1 \\
 & U_1 \geq 0, U_2 \geq 0
 \end{aligned} \tag{3.9.7}$$

Problem (3.9.4) ün duali

$$\begin{aligned}
 & \text{Min } V_1'e + V_2'e \\
 & \begin{pmatrix} Y', -Y' \\ X', -X' \end{pmatrix} \begin{pmatrix} V_1 \\ V_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\
 & V_1 \geq 0, V_2 \geq 0
 \end{aligned} \tag{3.9.8}$$

şeklinde olur.

Problem (3.9.3) sonlu optimal çözüme sahip olduğundan Problem (3.9.7) ve Problem (3.9.8) oluşturuldu. Problem (3.9.8) in bir optimal çözümünde ya V_{1i} ya da V_{2i} sıfır değildir, çünkü bazı i -ler için $\delta = \min(V_{1i}, V_{2i})$ pozitiftir öyle ki $V_{1i} - \delta, V_{2i} - \delta$ problem için uygundur ve amaç fonksiyonunun değeri 2δ ile azaltılır. Bununla birlikte, eğer biz sadece temel uygun çözümleri ele alıyorsak V_{1i} ve V_{2i} nin her ikisi de pozitif olmayabilir. Çünkü kolonlar lineer bağımlıdır. Bundan dolayı $V_i = V_{1i} - V_{2i}$ ve $|V_i| = V_{1i} + V_{2i}$ yazabiliriz. Bundan sonra Problem (3.9.8) denk bir şekilde aşağıdaki gibi ifade edilebilir:

$$\begin{aligned}
 & \text{Min } \sum_{i=1}^n |V_i| \\
 & Y'V = 1 \\
 & X'V = 0 \\
 & V \text{ işaretinde kısıtlama yok}
 \end{aligned} \tag{3.9.9}$$

Problem (3.9.9) u Problem (3.9.7) ye benzeterek Problem (3.9.10) şekline dönüştürebiliriz. Bunun için

$$U_i = \frac{V_i}{\sum_{l=1}^n \|V_l\|}$$

olsun. $\sum_{i=1}^n \|U_i\| = 1$ dir ve aşağıdaki problem elde edilir:

$$\max Y'U$$

$$X'U = 0$$

$$\sum_{i=1}^n \|U_i\| = 1$$

$$U \text{ işaretinde kısıtlama yok} \quad (3.9.10)$$

Eğer şimdi U_i ile $U_{1i} - U_{2i}$ yi ve $|U_i|$ ile $U_{1i} + U_{2i}$ yi yer değiştirirsek Problem (3.9.7) yi elde ederiz. Sonuç olarak (1) ve (2) varsayımlarından aşağıdaki varsayımı yaparız: (3) X in her p-satırlı alt kare matrisi tekil olmayan matristir. Problem (3.9.7) ve (3.9.8) in uygun çözüm kümelerinin uç noktalarını bozmadan belli sonuçları ispatlayabiliriz.

Sonuç 3.9.1 : Eğer varsayım 1-3 sağlanırsa Problem (3.9.7) ve (3.9.8) in hiç dejenere uç noktası yoktur.

İspat: Problem (3.9.7) nin kısıtlamalarını ele alalım

$$[X', -X'] \begin{pmatrix} U_1 \\ U_2 \end{pmatrix} = 0$$

$$U_1'e + U_2'e = 1$$

ya da

$$\sum_{j=1}^n X_{ij} (U_{1i} - U_{2i}) = 0 \quad , \quad i = 1, \dots, p$$

$$\sum_{j=1}^n (U_{1i} + U_{2i}) = 1$$

Burada $U_{1i} \geq 0$, $U_{2i} \geq 0$ dır. $U_{1i} - U_{2i}$ hiçbir i -için sıfır olamaz. Çünkü

$$\sum_i (U_{1i} - U_{2i}) = 1$$

dir. X' matrisinin herhangi p kolonu lineer bağımsız olduğundan (varsayım 3), kısıtlamalar sadece, U_{1i} , U_{2i} değişkenlerinin en azından p tanesi sıfırdan farklı ise sağlanabilir. Bu problemin herhangi bir bazı $p+1$ büyüklüğünde olacaktır. Problem için uygun çözümlerin kümesinin bir (U_1, U_2) uç noktasının dejenere olduğunu varsayalım. (U_1, U_2) deki sıfırdan farklı bileşenlerin sayısı $p+1$ den az olacaktır. Bu da çözümün uygun olmasına ters düşer. Problem (3.9.7) için sonuç ispatlanmış olur. Aynı şekilde bu sonucu Problem (3.9.8) için de ispatlayabiliriz.

Bu sonucun bir sonucu olarak; lineer programladaki dualite teorisinin tamamlayıcı gevşeklik özelliğini kullanarak aşağıdaki sonucu elde ederiz:

Sonuç 3.9.2 : Eğer varsayım 1-3 sağlanırsa ve Problem (3.9.3) ün bir optimal çözümü bulunursa. kısıtlamaların en azından $p+1$ tanesi bu problem için sağlanır. Bu aynı zamanda , gözlemlerin en azından $p+1$ tanesi için elde edilecek maksimal değer, d , olarak da ifade edilebilir.

Sonuç (3.9.1) ve (3.9.2) den, yapılan varsayımlar altında, uç noktaların hiçbir dejenere değildir. Bu yüzden Problem (3.9.7) deki U_{1i} ve U_{2i} değişkenlerinin tam olarak $p+1$ tanesi bu şekildeki her bir uç noktada sıfırdan farklıdır. Ayrıca uygun kolonlar lineer bağımlı iken, herhangi bir i -için U_{1i} ve U_{2i} nin ikisi de pozitif olamaz. Böylece, $U_i = U_{1i} - U_{2i}$ şeklinde ifade edersek Problem (3.9.7) nin uygun

bölgesinin herhangi bir uç noktası için U_i nin sıfırdan farklı olduğu $p+1$ tane i -indisi vardır. Bu indislerin kümesi S ile gösterilsin. Diğer yandan bu S kümesi bir lineer eşitlik sistemi şeklinde tanımlanır:

$$\sum_{i \in S} X_{ij} U_i = 0 \quad , \quad j = 1, \dots, p \quad (3.9.11)$$

ve

$$\sum_{i \in S} \|U_i\| = 1 \quad (3.9.12)$$

Buradan, $U_i = 0 \quad i \in S$ (3.9.11) ve (3.9.12) nin bir çözümü değildir. Böyle bir durumda tüm U_j değişkenlerinin sıfırdan farklı olduğunu iddia ederiz. Eğer herhangi bir i -için $U_i = 0$ ise (3.9.11) in sağ tarafındaki kolonların geri kalanı sıfır olarak ifade edilir. Varsayım 3 den dolayı U_j lere karşılık gelenlerin hepsi sıfırdır. Dahası U_i lerin tek şekilde hesaplandığı ortak bir çarpan vardır. Bu çarpan, varsayım 3 ün tekrar kullanılması ile, doğrulanabilir. Ortak çarpanın λ olduğunu varsayalım. Böylelikle (3.9.12) sağlanır;

$$|\lambda| \sum_{i \in S} \|U_i\| = 1 \quad \text{veya} \quad \lambda = \pm 1$$

Bu yüzden (3.9.11) ve (3.9.12) sisteminin tam olarak iki çözümüne sahip oluruz.

Şimdi $p+1$ indisin verilen bir S kümesine karşılık gelen kısıtlı regresyon problemini ele alalım:

$$\min_{\beta} \max_{i \in S} \left| Y_i - \sum_{j=1}^p X_{ij} \beta_j \right|$$

Bu kısıtlı problemin β_s çözümü, R^p deki $p+1$ tane hiper düzlem tarafından sınırlandırılmış kürenin merkezidir.

$$Y_i - \sum_{j=1}^p X_{ij} \beta_j = 0 \quad , \quad i \in S \quad (3.9.13)$$

Minimum değer $-d_s$ kürenin yarıçapıdır. Bundan dolayı Problem (3.9.7), simpleksteki (3.9.13) ile verilen maksimum yarıçapa sahip küreye karşılık gelen $p+1$ indisi tüm S altkümelerinden bulmaya denktir.

Bazı varsayımlar altında MINMAXAD probleminin bu geometrik yorumu yukarıdaki eşitliği ispatlasa bile, C_{p+1}^n küçük ve $n \geq 2$ ya da 3 olmadıkça – bu gibi durumlarda problemi grafiksel olarak çözebiliriz – bu problemi çözmek için basit bir metod ortaya koyamaz.

3.10. MINMAXAD Kullanarak Yansız Tahminler Elde Etme

Bölüm 3.7. de olduğu gibi bu bölümde de bir antisimetrik β_0 tahmini kullanarak ve iki denk problem tanımlayarak yansız tahminler elde edebiliriz. Burada iki problem vardır:

Min d

$$\begin{pmatrix} -X & X & -e \\ X & -X & -e \end{pmatrix} \begin{bmatrix} \bar{\beta}_1 \\ \bar{\beta}_2 \\ d \end{bmatrix} \leq \begin{pmatrix} -Y & +X\beta_0 \\ Y & -X\beta_0 \end{pmatrix}$$

$$\bar{\beta}_1 \geq 0 \quad , \quad \bar{\beta}_2 \geq 0 \quad , \quad d \geq 0 \quad (3.10.1)$$

ve

Min d

$$\begin{pmatrix} -X & X & -e \\ X & -X & -e \end{pmatrix} \begin{bmatrix} \bar{\beta}_2 \\ \bar{\beta}_1 \\ d \end{bmatrix} \leq \begin{pmatrix} Y & -X\beta_0 \\ -Y & +X\beta_0 \end{pmatrix}$$

$$\bar{\beta}_1 \geq 0, \bar{\beta}_2 \geq 0, d \geq 0 \quad (3.10.2)$$

Bölüm 3.7 de tanımlandığı gibi burada da $\bar{\beta} = \bar{\beta}_1 - \bar{\beta}_2$ ve $\bar{\beta}_1 = \beta_1 + \beta_0^{(2)}$, $\bar{\beta}_2 = \beta_2 + \beta_0^{(1)}$ dir.

Problem (3.10.1) ve (3.10.2) sırasıyla P_1^* ve P_2^* ile gösterilsin. Böylelikle β nın yansız tahminini aşağıdaki gibi elde ederiz:

- 1- Antisimetrik bir β_0 seç.
- 2- Eşit olasılıkla P_1^* ve P_2^* problemlerini seç.
- 3- Bölüm 3.7. de tanımlandığı gibi seçilen problemin bir optimal çözümünü bul. Bu optimal çözümler $\bar{\beta}_1^*$ ve $\bar{\beta}_2^*$ şeklinde ifade edilsin. Burada $\beta^* = \bar{\beta}_1^* - \bar{\beta}_2^* + \beta_0$ dır.

Bu tahminin yansızlığının ispatı Bölüm 3.7 deki gibi yapılabilir.

3.11. En Küçük Kareler ve MINMAD Regresyonun Konveks Kombinasyonu

Farklı araştırmacılar regresyon problemlerinde tahminler elde etmek için MINMAD ve en küçük karelerin birleşiminin kullanılmasını öne sürmüşlerdir. Bu bölüm bu problemi bir quadratik problem olarak formüle etmeye ve böyle problemleri çözmek için teori geliştirmeye ayrılmıştır. En küçük kareler regresyon problemine denk olacak şekilde aşağıdaki ifadeyi yazabiliriz:

$$\text{Min } d'd$$

$$X\beta + d = Y$$

$$\beta, d \text{ işaretinde kısıtlama yok} \quad (3.11.1)$$

Problem (3.11.1) aynı zamanda şu şekilde de formüle edilebilir

$$\text{Min } d'd$$

$$-d \leq Y - X\beta \leq d$$

$$\beta \text{ işaretinde kısıtlama yok}$$

$$d \geq 0 \quad (3.11.2)$$

veya

$$\text{Min } d'd$$

$$X\beta - d \leq Y$$

$$-X\beta - d \leq -Y$$

$$\beta \text{ işaretinde kısıtlama yok, } d \geq 0$$

$$\beta = \beta_1 - \beta_2, \quad \beta_1, \beta_2 \geq 0$$

olsun. Ayrıca

$$\bar{Q} = \begin{bmatrix} 0_{p \times p} & 0_{p \times p} & 0_{p \times n} \\ 0_{p \times p} & 0_{p \times p} & 0_{p \times n} \\ 0_{n \times p} & 0_{n \times p} & I_{n \times n} \end{bmatrix}, \quad A = \begin{pmatrix} X & -X & -I \\ -X & X & -I \end{pmatrix}$$

$$V' = (\beta_1, \beta_2, d) \text{ ve } b = \begin{pmatrix} Y \\ -Y \end{pmatrix}$$

Burada $Q : (2p + n) \times (2p + n)$ ve $A : 2n \times (2p + n)$ matrislerdir. Böylece aşağıdaki ifade elde edilir:

$$\begin{aligned}
& \text{Min } V'QV \\
& AV \leq b \\
& V \geq 0
\end{aligned} \tag{3.11.3}$$

En küçük karelerin işleyişinin hatalar üzerindeki normallik varsayımları ihmal edildiğinde zayıf olduğu biliniyor. Böyle durumlarda β parametrelerini, alternatif-olarak, ortalama mutlak sapma ve sapmaların karelerinin ortalamasının konveks kombinasyonunu kullanarak tahmin edebiliriz.

$\gamma_1 + \gamma_2 = 1, \gamma_i \geq 0, i = 1, 2$ olsun. d_i ler negatif olmadığında

$$C' = (0_{1 \times 2p}, e'_{1 \times n}) \text{ için } C'V, \sum_{i=1}^n |d_i| = \sum_{i=1}^n d_i$$

ifadesine denktir. Bundan dolayı aşağıdaki problemi elde etmiş oluruz.

$$\begin{aligned}
& \text{Min } \gamma_1 d' d + \gamma_2 \sum_{i=1}^n d_i \\
& -d \leq Y - X\beta \leq d \\
& \beta \text{ işaretinde kısıtlama yok} \\
& d \geq 0
\end{aligned} \tag{3.11.4}$$

Veya

$$\begin{aligned}
& \text{Min } \gamma_1 [V'QV] + \gamma_2 C'V \\
& AV \leq b \\
& V \geq 0
\end{aligned} \tag{3.11.5}$$

Burada dikkat edecek olursak; $Q; \gamma_1 \geq 0$ ve $\gamma_2 C'V$ lineer olduğunda amaç fonksiyonu bir konveks fonksiyon şeklinde ifade edilen pozitif yarı tanımlı bir matristir. Şimdi, Problem (3.11.5) in optimal çözümü üzerindeki belli sonuçları

ispatlayarak Bölüm 3.7 de verilen simpleks yönteminin değişik bir şekli olan başka bir yöntem geliştiririz.

$$f(V) = \gamma_1 V' Q V + \gamma_2 C' V \quad \text{ve} \quad g_i(V) = A_i V - b_i, \quad i = 1, 2, \dots, 2n \quad \text{olsun.}$$

Böylelikle $g_i(V) \leq 0$, $i = 1, 2, \dots, 2n$ kısıtlaması ile $f(V)$ yi minimum yapmak isteriz.

Tanım 3.11.1 : $F(V, \lambda)$ Lagrange fonksiyonu aşağıdaki gibi tanımlanır:

$$F(V, \lambda) = f(V) + \lambda G(V)$$

Burada

$$G(V) = \begin{bmatrix} g_1(V) \\ \vdots \\ g_{2n}(V) \end{bmatrix}$$

ve $\lambda' = (\lambda_1, \dots, \lambda_{2n})$ R^{2n} de bir vektördür. λ nın bileşenleri Lagrange çarpanları olarak adlandırılır.

Tanım 3.11.2 : Eğer tüm $V \geq 0, \lambda \geq 0$ için

$$F(V^0, \lambda) \leq F(V^0, \lambda^0) \leq F(V, \lambda^0) \quad (3.11.6)$$

ise, $V^0 \geq 0$, $\lambda^0 \geq 0$ olmak üzere R^{2p+3n} de bir (V^0, λ^0) noktasına $F(V, \lambda)$ nın *saddle (eyer) noktası* denir.

Sonuç 3.11.1 : Eğer (V^0, λ^0) , $F(V, \lambda)$ nin eyer noktası ise V^0 Problem (3.11.5) için bir optimal çözümdür.

İspat: (3.11.6), $V \geq 0, \lambda \geq 0$ için şu şekilde ifade edilir:

$$f(V^0) + \lambda' G(V^0) \leq f(V^0) + \lambda^0' G(V^0) \leq f(V) + \lambda^0' G(V) \\ \Rightarrow \lambda' G(V^0) \leq \lambda^0' G(V), \quad \text{tüm } \lambda \geq 0 \text{ için.}$$

Bu dönüşüm $G(V^0) \leq 0$ şeklinde ifade edilir. Bu yüzden $\lambda^0' G(V^0) \leq 0$ dır. Diğer yandan $\lambda = 0$ için $\lambda' G(V^0) = 0 \leq \lambda^0' G(V^0)$ dır. Bundan dolayı $0 \leq \lambda^0' G(V^0) \leq 0$ veya $\lambda^0' G(V^0) = 0$ olur. Şimdi, eğer V uygun ise $G(V) \leq 0$,

$$f(V^0) + \lambda^0' G(V^0) \leq f(V) + \lambda^0' G(V) \leq f(V)$$

dir. Veya V^0 (3.11.5) için optimaldir. Böylece ispat tamamlanmış olur.

Bunun tersi sonuç 3.11.2. de gösterilmektedir.

Sonuç 3.11.2 : V^0 in (3.11.5) için optimal olduğunu varsayalım. Öyle ise $\lambda^0 \geq 0$ vardır öyle ki, tüm $V \geq 0, \lambda \geq 0$ için

$$F(V^0, \lambda) \leq F(V^0, \lambda^0) \leq F(V, \lambda^0)$$

dir.

İspat. V^0 Problem (3.11.5) için bir optimal çözüm olsun. $U = (U_0, U_1, \dots, U_{2n})$

vektörler olsun. Aşağıdaki gibi D_1 ve D_2 kümeleri tanımlarız:

$$D_1 = \{U \mid U_0 \geq f(V), U_i \geq g_i(V), i = 1, 2, \dots, 2n \exists V \geq 0 \text{ için}\} \\ D_2 = \{U \mid U_0 < f(V^0), U_i < 0, i = 1, 2, \dots, 2n\}$$

$f(V)$ konveks ve $g_i(V)$ lineer olduğundan dolayı D_1 ve D_2 konveks kümelerdir ve D_2 açıktır. V^0 optimal iken tüm uygun V ler için $f(V^0) \leq f(V)$ dir.

Bu yüzden $D_1 \cap D_2 = \emptyset$ dir. D_1 ve D_2 R^{2n+1} in uygun alt kümeleridir. Şimdi, aşağıdaki şekilde ifade edilen, konveks kümelerin ayırma teoremini kullanacağız:

D_1 ve D_2 , R^Z nin, hiç ortak noktaları olmayan iki uygun konveks alt kümesi olsun. D_2 açık olsun. D_1 ve D_2 yi ayıran bir $WU = \alpha$ hiper düzlemi vardır öyle ki, $W \neq 0$ vektörü ve tüm $U^1 \in D_1$ ve $U^2 \in D_2$ için $WU^2 \leq \alpha < WU^1$ olacak şekilde bir α reel sayısı vardır.

Bu teoremden ifade edilen durumlar sağlandığında bir $W = \{W_1, W_2, \dots, W_{2n}\}$ $W \neq 0$ vektörüne sahip oluruz öyle ki,

$$WU^1 > WU^2, U^1 \in D_1 \text{ ve } U^2 \in D_2 \text{ için} \quad (3.11.7)$$

olur.

$U^2 \in D_2$ nin bileşenlerinin büyük negatif sayılar olabilir olması $W \geq 0$ sonucunu ortaya koyar. D_2 nin bir sınır noktası ve bir D_1 noktası için hala (3.11.7) ye sahip oluruz. Fakat tam bir eşitsizlik yerini daha büyük bir eşitsizliğe veya bir eşitliğe bırakır. $U^2 = (f(V^0), 0, \dots, 0, 0, \dots, 0)'$ D_2 nin sınırındadır ve $V \geq 0$ olduğu sürece $U^1 = (f(V), g_1(V), \dots, g_{2n}(V))'$ D_1 e aittir. Bu yüzden (3.11.7) bu U^1 ve U^2 noktaları için aşağıdaki gibi olur:

$$WU^1 \geq WU^2$$

veya

$$W_0 f(V) + \sum_{i=1}^{2n} W_i g_i(V) \geq W_0 f(V^0) \quad (3.11.8)$$

tüm $V \geq 0$ için. Şimdi, ya $W_0 = 0$ ya da $W_0 > 0$ dır. Biz $W_0 > 0$ olduğunu göstereceğiz. Bu yüzden $W_0 = 0$ olduğunu varsayalım. Böylelikle

$$\sum_{i=1}^{2n} W_i g_i(V) \geq 0 \quad , \quad \text{tüm } V \geq 0 \text{ için} \quad (3.11.9)$$

olur. $W_0 \neq 0$ olduğundan en azından bir $W_{i_0} > 0$, $1 \leq i_0 \leq 2n$ dir. i-inci kısıtlama $g_{i_0}(V)$ yi ele alalım.

$$g_{i_0} = \begin{cases} (X\beta_1 - X\beta_2 - \bar{I}d - Y)_{i_0} \leq 0 & 1 \leq i_0 \leq n \\ (-X\beta_1 + X\beta_2 - Id + Y)_{i_0} \leq 0 & n+1 \leq i_0 \leq 2n \end{cases}$$

$d_{i_0} > 0$ keyfi büyük seçildiğinde daima $V \geq 0$ ve $G(V) \leq 0$ dır. Bu gibi bir seçim için $g_i(V) < 0$ olur. Burada $W_{i_0} g_{i_0}(V) < 0$ ve (3.11.9) doğru değildir. Buradan $W_0 > 0$ sonucuna varırız. Şimdi $\lambda^0 = (1/W_0)(W_1, \dots, W_{2n})'$ olsun. Böylece $\lambda^0 \geq 0$ dır ve (3.11.8)

$$f(V) + \sum_{i=1}^{2n} \lambda_i^0 g_i(V) \geq f(V^0) \quad , \quad \text{tüm } V \geq 0 \text{ için} \quad (3.11.10)$$

olur.

$V = V^0$ için $\lambda^0 G(V^0) \geq 0$ olur. Ayrıca V^0 Problem (3.11.5) için uygun çözüm iken $G(V^0) \leq 0$ dır. $\lambda^0 \geq 0$ olduğundan $\lambda^0 G(V^0) \leq 0$

$$\lambda^0 G(V^0) = 0 \quad (3.11.11)$$

ve

$$\lambda' G(V^0) \leq 0 \quad , \quad V \geq 0 \text{ için} \quad (3.11.12)$$

şeklinde ifade edilir. Böylece (3.11.10), (3.11.11) ve (3.11.12) den

$$f(V^0) + \lambda' G(V^0) \leq f(V^0) + \lambda'^0 G(V^0) \leq f(V) + \lambda'^0 G(V)$$

olur ($V \geq 0$, $\lambda \geq 0$ için). Veya (V^0, λ^0) $F(V, \lambda) = f(V) + \lambda' G(V)$ nin bir eyer noktasıdır. Böylece ispat tamamlanır.

Sonuç 3.11.2 genelde herhangi bir G için doğru değildir.

Örnek 3.11.1 :

$V^2 \leq 0$, $V \geq 0$ kısıtlamaları altında V nin minimizasyonunu ele alalım. Şu açıktır ki sadece $V = 0$ kısıtlamaları sağlar ve bu optimaldir. Lagrange fonksiyonu $F(V, \lambda) = -V + \lambda V^2$ dir. Eğer $V^0 = 0$ ve uygun, $\lambda^0 \geq 0$ ise (V^0, λ^0) , $F(V, \lambda)$ nin bir eyer noktasıdır, $0 \leq -V + \lambda^0 V^2$ tüm $V \geq 0$ için doğru değildir.

İspattaki can alıcı nokta $W_0 > 0$ ın gösterilmesidir. Bu $G(V)$ nin yapısından dolayı mümkündür, yani, en azından bir $V \geq 0$, her i -için $g_i(V) < 0$ dır. Genelde bu durum, kısıtlamalar tarafından sağlanan bir durum olarak aşağıdaki şekilde ifade edilir (Bu, Slater in şartlı kısıtlaması olarak bilinir.):

$g_i(V) \leq 0$ kısıtlamaları $g_i(V^i) < 0$ iken her i -için bir $V^i \geq 0$ olacak şekilde veya $g_i(V) \leq 0$ kısıtlamaları tüm i -ler için $g_i(V) < 0$ iken bir $V \geq 0$ olacak şekildedir.

Şimdi, Sonuç(3.11.1) ve (3.11.2) yi birlikte uygun bir şekilde, Problem (3.11.5) in bir çözümünün optimizasyonu şeklinde oluşturulan Sonuç (3.11.3) olarak ifade edebiliriz.

Sonuç 3.11.3 : $V^0 \geq 0$ Problem (3.11.5) için bir optimal çözümdür, ancak ve yalnız, (V^0, λ^0) $F(V, \lambda) = f(V) + \lambda' G(V)$ nin bir eyer noktası olacak şekilde bir $\lambda^0 \geq 0$ varsa.

(V^0, λ^0) ın $F(V, \lambda)$ nin bir eyer noktası olduğunu söylemek için $F(V^0, \lambda^0)$ ile $F(V, \lambda^0)$ ve $F(V^0, \lambda)$ yı karşılaştırmamız gerekmektedir. $F(V)$ ve $G(V)$

fonksiyonlarının diferansiyellenebilir olmasından kaçınılırız; klasik optimizasyon metodlarında kullanılan parçalı türevlerin veya gradyentlerin terimlerinde eşit durumlar bulmaya çalışılır. Fakat gradyentlerin terimlerindeki optimallik durumları sadece bölgesel optimallik durumlarını sağlar. Neyse ki, değindiğimiz problemin tipi herhangi bir bölgesel optimumu aynı zamanda global optimal olacak şekildedir. Bu sonucu genel bir şekilde gösterebilir ve burada beklenen sonuçla özelleştirebiliriz.

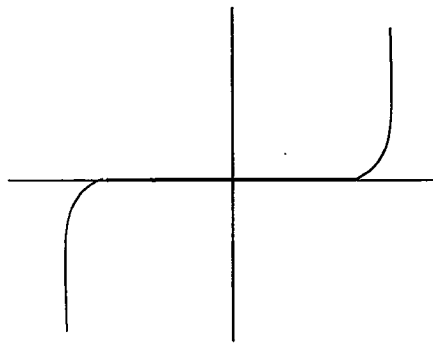
Tanım 3.11.3 : Eğer tüm $V^1, V^2 \in F$ için, $V^0 = \lambda V^1 + (1 - \lambda)V^2$, $1 > \lambda > 0$, $f(V^1) \leq f(V^2) \Rightarrow f(V^0) \leq f(V^2)$ veya buna denk olarak

$$f(V^0) \leq \max[f(V^1), f(V^2)] \quad (3.11.13)$$

ise $F \subset R^Z$ konveks bölgesi üzerinde tanımlanan $f(V)$ fonksiyonu *yarı konvektir*.

Not 3.11.1 : Herhangi bir konveks fonksiyon yarı konvektir, fakat bunun tersi doğru değildir. Şekil 3.1. de grafiği verilen fonksiyon

$$f(V) = \begin{cases} 0 & |V| \leq 1 \\ (V-1)^3 & V \geq 1 \\ (V+1)^3 & V \leq -1 \end{cases}$$



Şekil 3.1 Yarı konveks olan fakat konveks olmayan bir fonksiyon

Yukarıdaki $f(V)$ nin yarı konveks olduğunu fakat konveks olmadığını doğrulamak kolaydır.

Tanım 3.11.4 : Eğer $F(V)$, F de yarı konveks ve $V^1, V^2 \in F$ için $V^0 = \lambda V^1 + (1 - \lambda)V^2$, $1 > \lambda > 0$,

$$f(V^1) < f(V^2) \Rightarrow f(V^0) < f(V^2)$$

veya

$$f(V^1) \neq f(V^2) \Rightarrow f(V^0) < \max[f(V^1), f(V^2)] \quad (3.11.14)$$

ise $F \subset R^Z$ konveks bölgesi üzerinde tanımlı olan bir $f(V)$ fonksiyonu F de *aşıkarak yarı konvektir*.

Not 3.11.2 : Eğer $f(V)$, konveks $F \subset R^Z$ kümesinde konveks ise F de *aşıkarak yarı konvektir*.

İspat: $f(V)$ nin F de konveks ve $f(V^1) \leq f(V^2)$ olduğunu varsayalım

$$\begin{aligned} f(V^0) &= f(\lambda V^1 + (1 - \lambda)V^2) \\ &\leq f(V^1) + (1 - \lambda)f(V^2) \quad (f \text{ nin konveks oluşundan.}) \\ &= f(V^2) + \lambda [f(V^1) - f(V^2)] \\ &\leq f(V^2) \end{aligned} \quad \text{tüm } 1 > \lambda > 0 \text{ için.}$$

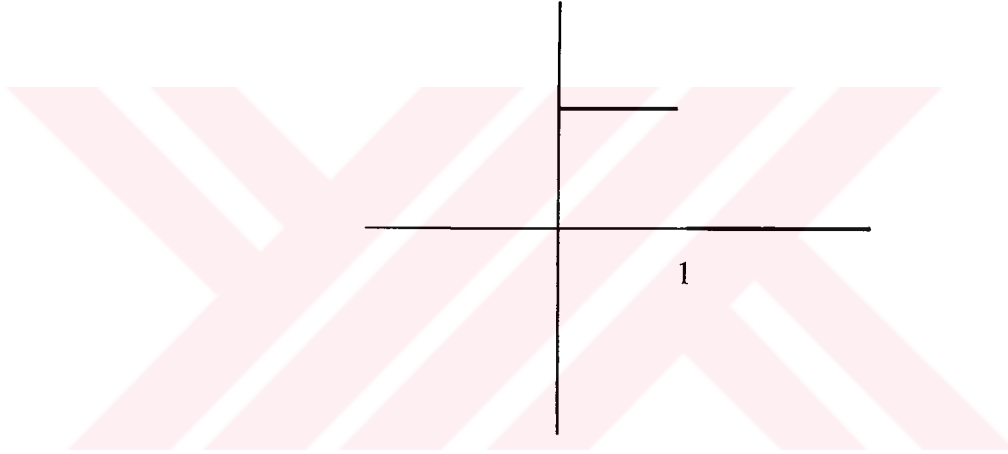
olur. Böylece $f(V)$, F de yarı konvektir. Dahası

$$f(V^0) \leq f(V^2) + \lambda [f(V^1) - f(V^2)]$$

$$< f(V^2) \quad \text{tüm } 1 > \lambda > 0 \text{ ve } f(V^1) < f(V^2)$$

dir. Bu yüzden $f(V)$, F de aşıkarak yarı konvektir.

Not 3.11.3 : Herhangi yarı konveks fonksiyonun aşıkarak yarı konveks olmadığına dikkat edelim



Şekil 3.2. Yarı konveks olan fakat aşıkarak yarı konveks olmayan bir fonksiyon

Şekil 3.2. ye ait fonksiyon aşağıda verilmiştir

$$f(V) = \begin{cases} 1 & 0 \leq V \leq 1 \\ 0 & V > 1 \end{cases}$$

Sonuç 3.11.4 : Eğer $f(V)$ $f \subset R^Z$ konveks kümesinde aşıkarak yarı konveks ise herhangi bölgesel minimum nokta $f(V)$ yi F de global minimum yapar.

İspat: $V^1 \in F$ global olmayan bölgesel bir minimum nokta olsun. Söylenen şey $f(V^2) < f(V^1)$ olacak şekilde $V^2 \in F$ olmasıdır. f nin aşikar olarak yarı konveks olmasından

$$f(V^0) < f(V^1) \text{ tüm } V^0 = \lambda V^1 + (1-\lambda)V^2, \quad 1 > \lambda > 0 \text{ için.}$$

Şimdi her $\varepsilon > 0$ için, V^1 etrafında ε yarıçaplı açık bir küre $B(V^1, \varepsilon)$ için $B(V^1, \varepsilon) \cap (V^1, V^2) \neq \emptyset$ ifadesine sahip oluruz. Bu yüzden V^1 bölgesel bir minimum noktası olamaz. Bu da bir çelişki yaratır. Buradan sonuç ispatlanmış olur.

Böylece, amaç fonksiyonunun konveks, bundan dolayı aşikar olarak yarı konveks olması problem (3.11.5) için gradyent terimlerinde bölgesel optimallik koşulları bulunması için yeterlidir.

F_V ve F_λ $F(V, \lambda)$ nın V ve λ yönlerinde gradyentlerini gösterebiliriz.

$$F_V = \nabla_V F = \left(\frac{\partial F}{\partial V_1}, \dots, \frac{\partial F}{\partial V_{2p+n}} \right)$$

$$F_\lambda = \nabla_\lambda F = \left(\frac{\partial F}{\partial \lambda_1}, \dots, \frac{\partial F}{\partial \lambda_{2n}} \right)$$

Ayrıca $F_V(\hat{V}, \hat{\lambda})$, $(\hat{V}, \hat{\lambda})$ değerlerinde $\nabla_V F$, aynı şekilde $F_\lambda(\hat{V}, \hat{\lambda})$, $(\hat{V}, \hat{\lambda})$ değerlerinde $\nabla_\lambda F$ olur.

Sonuç 3.11.5 : $V^0 \geq 0$ Problem (3.11.5) için bir optimal çözümdür ancak ve yalnız

$$F_V(V^0, \lambda^0) \geq 0, \quad V^0{}' F_V(V^0, \lambda^0) = 0, \quad (3.11.15)$$

$$F_\lambda(V^0, \lambda^0) \leq 0, \quad \lambda^0{}' F_\lambda(V^0, \lambda^0) = 0, \quad (3.11.16)$$

olacak şekilde bir $\lambda^0 \geq 0$ varsa.

İspat: Biz (3.11.15) ve (3.11.16) eşitliklerinin eyer noktası eşitliklerine denk olduğunu yani,

$$F(V^0, \lambda) \leq F(V^0, \lambda^0) \leq F(V, \lambda^0) \quad (3.11.17)$$

olduğunu göstereceğiz. $V \geq 0, \lambda \geq 0$ için ilk olarak (3.11.17)'nin (3.11.15) ve (3.11.16) yı ifade ettiğini gösteririz.

Bazı k-lar için $\frac{\partial F}{\partial V_k}$ nın (V^0, λ^0) daki değerinin negatif olmasının (3.11.15)

e ters düştüğünü varsayalım. $V_l = V_l^0$ ($l \neq k$ için) bir $V \geq 0$ vektörünün ve $F(V, \lambda^0) < F(V^0, \lambda^0)$ olacak şekilde $V_k \geq V_k^0$ bulunabilmesi (3.11.17) ye terstir. Bu yüzden (3.11.17) $F_V(V^0, \lambda^0) \geq 0$ ı ifade eder. Çünkü $V^0 \geq 0, V^0' F_V(V^0, \lambda^0)$ iç çarpımındaki tüm bileşenler negatif değildir. Bazı k-lar için $V_k^0 \cdot \partial F(V^0, \lambda^0) / \partial V_k > 0 \Rightarrow V_k^0 > 0, \partial F(V^0, \lambda^0) / \partial V_k > 0$ olduğunu varsayalım. Bu. $F(V, \lambda^0) < F(V^0, \lambda^0)$ için

$$V_l = \begin{cases} V_l^0, & l \neq k \\ V_l^0 > V_l \geq 0, & l = k \end{cases}$$

olacak şekilde bir V vektörünün olduğu anlamına gelir ve bu (3.11.17) için bir çelişkidir. Buradan, tüm k-lar için $V_k^0 \cdot \partial F(V^0, \lambda^0) / \partial V_k$ sifıra eşittir veya (3.11.17) $\Rightarrow$ (3.11.15) tir. Aynı şekilde (3.11.17) $\Rightarrow$ (3.11.16) yı da gösterebiliriz. Diğer taraftan (3.11.15) ve (3.11.16) $\Rightarrow$ (3.11.17) olduğunu da göstereceğiz. $F(V, \lambda^0)$ ın V nin bir konveks fonksiyonu olduğuna dikkat edelim. Diferansiyellenebilir konveks fonksiyon tanımından $V \geq 0$ için

$$F(V, \lambda^0) \geq F(V^0, \lambda^0) + (V - V^0)' F_V(V^0, \lambda^0) \quad (3.11.18)$$

yazarız. $F(V^0, \lambda)$, λ da lineer olduğundan $\lambda \geq 0$ için

$$F(V, \lambda^0) = F(V^0, \lambda^0) + (\lambda - \lambda^0)' F_\lambda(V^0, \lambda^0) \quad (3.11.19)$$

dir. Şimdi $(V - V^0)' F_V(V^0, \lambda^0) = V' F_V(V^0, \lambda^0) \geq 0$ ((3.11.15) ve $V \geq 0$ dan dolayı) dir.

Bu yüzden (3.11.18) $\Rightarrow$

$$F(V, \lambda^0) \geq F(V^0, \lambda^0) \quad (3.11.20)$$

dir. Ayrıca, $(\lambda - \lambda^0)' F_\lambda(V^0, \lambda^0) = \lambda' F_\lambda(V^0, \lambda^0) \leq 0$ dir [(3.11.16) ve $\lambda \geq 0$ dan dolayı.] Bu yüzden (3.11.19) $\Rightarrow$

$$F(V^0, \lambda^0) \geq F(V^0, \lambda) \quad (3.11.21)$$

(3.11.20) ve (3.11.21) den (3.11.17) sağlanır ve ispat tamamlanmış olur.

Şimdiye kadar $f(V)$ ve $g_i(V)$ nin konveksliğini ve diferansiyellenebilirliğini kullandık. Bu yüzden, bu sonuç, genelde $f(V)$ amaç fonksiyonu ve konveks ve diferansiyellenebilen $g_i(V)$ fonksiyonları için doğrudur ve daha önce sözü edilen kısıtlama şartlarını sağlar. Problem (3.11.5) e geri dönelim:

$$f(V) = \gamma_1 V' Q V + \gamma_2 C' V$$

ve

$$g_i(V) = (AV)_i - b_i$$

dir. Buradan 3.11.5 ve 3.11.16 eşitlikleri aşağıdaki şekilde ifade edilir:

$$F_V(V^0, \lambda^0) = \gamma_2 C + 2\gamma_1 QV^0 + A'\lambda^0 \geq 0; V^{0'} F_V(V^0, \lambda^0) = 0$$

ve

$$F_V(V^0, \lambda^0) = AV^0 - b \leq 0, \lambda^{0'} F_\lambda(V^0, \lambda^0) = 0$$

dır. $\mu^0 = F_V(V^0, \lambda^0)$ ve $\eta^0 = -F_\lambda(V^0, \lambda^0)$ olsun. Öyle ise $V^{0'} \mu^0 = 0$ ve $\lambda^{0'} \eta^0 = 0$ olur. Ya da $V^{0'} \mu^0 + \lambda^{0'} \eta^0 = 0$ yazabiliriz. V^0, λ^0, μ^0 ve η^0 in hepsi nonnegatif olduğundan dolayı Sonuç 3.11.5 den aşağıdakiler elde edilir.

Eğer

$$\begin{aligned} AV^0 + \eta^0 &= L \\ \mu^0 - 2\gamma_1 QV^0 - A'\lambda^0 &= \gamma_2 C \\ V^0 \geq 0, \mu^0 \geq 0, \eta^0 &\geq 0 \end{aligned} \quad (3.11.22)$$

ve

$$V^{0'} \mu^0 + \lambda^{0'} \eta^0 = 0 \quad (3.11.23)$$

olacak şekilde λ^0, μ^0 ve η^0 var ise $V^0 \geq 0$ Problem (3.11.5) için bir optimal çözümdür.

Şimdiye kadar geliştirilen, Problem (3.11.5) için bir optimal çözüm tanımlayan teoriden, (3.11.5) i çözmek için, Bölüm 3.1. de gösterilen simpleks yönteminin değişik bir şekli olan bir metod oluşturabiliriz. Simpleks yöntemi, problem için bir optimal çözüm ararken sadece temel uygun çözümler ele alır. Buradaki benzer sonuç da (3.11.22) nin lineer eşitlik sistemlerinin temel uygun çözümleri arasında bir optimal çözümü garanti eder.

Sonuç 3.11.6 : Aşağıdaki ifadeleri sağlayan herhangi V, λ, μ, η çözümlerini ele alalım

$$(1) \quad V, \lambda, \mu, \eta \geq 0$$

$$(2) \quad AV + \eta = b$$

$$(3) \quad \mu - 2\gamma_1 QV - A'\lambda = \gamma_2 C \quad (3.11.24)$$

ve

$$(4) \quad V'\mu + \eta'\lambda = 0 \quad (3.11.25)$$

Böyle bir çözümün (V, λ) parçası

$$\hat{F} = \{(V, \lambda) \mid AV \leq b, -2\gamma_1 QV - A'\lambda \leq \gamma_2 C, V \geq 0, \lambda \geq 0\}$$

nın bir uç noktasıdır.

İspat: $(V^*, \lambda^*, \mu^*, \eta^*)$, (3.11.24) ve (3.11.25) i sağlasın. $(V^*, \lambda^*) \in F$ dir. F nin tanımında negatif olmama koşullarını içeren $(2n + (2n+p) + (2p+n) + 2n) = 4p + 2n$ tane kısıtlama olduğuna dikkat edelim. $(V^*, \lambda^*, \mu^*, \eta^*)$, (3.11.25) i sağladığında (V^*, μ^*) in $2(2p+n)$ çarpanının en az $(2p+n)$ tanesi ve (λ^*, η^*) in $4n$ çarpanının en az $2n$ tanesi yok olur. Bu yüzden (V^*, λ^*) , $4p+6n$ eşitsizliğin en azından $2p+3n$ tanesini eşitlik olarak sağlar. Ya da (V^*, λ^*) bir temel uygun çözüm oluşturur. Buradan ispat tamamlanır.

Böylece, eklenen lineer olmayan (3.11.25) kısıtlamalarına rağmen (2) ve (3) te verilen sistemin sadece (4) ü sağlayan temel uygun çözümlerinin ele almak yeterlidir. Bundan sonra Problem (3.11.5) i çözmek için bir algoritma geliştiririz:

Algoritma 3.11.1 :

1.Adım: Bölüm 3.3. te tanımlanan yapay değişken tekniğini kullanarak bir $(A, I) \begin{bmatrix} V \\ \eta \end{bmatrix} = b$ uygun çözüm buluruz. (Eğer böyle bir çözüm yoksa dururuz. Problem (3.11.5) için böyle bir durum olmamalıdır.) Bu çözüm $\bar{V} \geq 0$, $\bar{\eta} \geq 0$ olsun ve temel kolonların indisleri sırasıyla $J_{\bar{V}}$ ve $J_{\bar{\eta}}$ ile gösterilsin.

2.Adım: $h = \gamma_2 C + 2\gamma_1 Q \bar{V}$ hesaplarız. Eğer $h = 0$, $V = \bar{V}$, $\mu = 0$, $\lambda = 0$ ve $\bar{\eta} = b - A \bar{V} \geq 0$ ise (3.11.24) ve (3.11.25) sağlanır ve dururuz. Aksi halde 3.adıma geçeriz.

3.Adım: Lineer programlama probleminde

min ζ

$$\begin{pmatrix} A & I & 0 & 0 & 0 \\ -2\gamma_1 Q & 0 & I & -A' & h \end{pmatrix} \begin{bmatrix} V \\ \eta \\ \mu \\ \lambda \\ \zeta \end{bmatrix} = \begin{pmatrix} b \\ \gamma_2 C \end{pmatrix}$$

$$V, \eta, \mu, \lambda \geq 0, \quad \zeta \geq 0 \quad (3.11.26)$$

aşağıdaki gibi seçilen $2p+3n$ kolon ile verilen bir başlangıç bazımız olur:

$$(1) \quad j \in J_{\bar{V}} \text{ için } \begin{pmatrix} a_j \\ -2\gamma_1 Q_j \end{pmatrix} \text{ ve } j \in J_{\bar{\eta}} \text{ için } \begin{pmatrix} e_j \\ 0 \end{pmatrix} \text{ şeklinde } 2n \text{ kolon}$$

$$(2) \quad j \notin J_{\bar{V}} \text{ için } \begin{bmatrix} 0 \\ I \end{bmatrix} \text{ dan ve } j \notin J_{\bar{\eta}} \text{ için } \begin{pmatrix} 0 \\ -A' \end{pmatrix} \text{ den } 2p+n \text{ kolon.}$$

(1) ve (2) deki tüm kolonlar lineer bağımsızdır.

(3) Şimdi ζ değişkenine karşılık gelen $\begin{bmatrix} 0 \\ h \end{bmatrix}$ kolonunu ele alalım. (1), (2) ve (3) ün hepsinde $2p+3n+1$ tane kolon vardır. (1) ve (2) deki kolonların birini (1), (2) ve $\begin{bmatrix} 0 \\ h \end{bmatrix}$ teki kolonlar lineer bağımsız olacak şekilde çıkaralım. Böylece $V = \bar{V}$, $\eta = \bar{\eta}$, $\mu = 0$, $\lambda = 0$ ve $\zeta = 1$ olan $2p+3n$ büyüklüğünde bir baz elde ederiz. Bu durumda $V'\mu = 0$ ve $\eta'\lambda = 0$ olduğuna dikkat edelim. Böylece Problem (3.11.26) için bir başlangıç bazı elde ederiz ve 4.adıma geçeriz.

4.Adım: Problem için simpleks uygulaması 3.adımda elde edilen baz ile başlar; eklenen kısıtlamalar

$$\text{Eğer } \begin{Bmatrix} V_i \\ \eta_i \\ \mu_i \\ \lambda_i \end{Bmatrix} \text{ temel ise } \begin{Bmatrix} \mu_i \\ \lambda_i \\ V_i \\ \eta_i \end{Bmatrix} \text{ baza girmemelidir.}$$

Eğer daha fazla olası değişiklik yapılamayacak bir baz elde etmişsek dururuz.

Sonuç 3.11.7 : Algoritma, (3.11.26) için şöyle bir çözüm bulur: $\hat{V}, \hat{\eta}, \hat{\mu}, \hat{\lambda} \geq 0$ ve $\hat{\zeta} = 0$. Ya da (3.11.24) ve (3.11.25) için bir çözüm elde ederiz.

İspat: $\zeta \geq 0$ olduğunu varsayalım. $\hat{V}, \hat{\eta}, \hat{\mu}, \hat{\lambda}$ ve $\hat{\zeta}$ aşağıdaki problem için çözüm olur.

Min ζ

$$\begin{pmatrix} A & I & 0 & 0 & 0 \\ -2\gamma_1 Q & 0 & I & -A' & h \end{pmatrix} \begin{bmatrix} V \\ \eta \\ \mu \\ \lambda \\ \zeta \end{bmatrix} = \begin{pmatrix} b \\ \gamma_2 C \end{pmatrix}$$

$$\hat{\mu}'V + \hat{V}'\mu = 0,$$

$$V \geq 0, \eta \geq 0, \mu \geq 0, \lambda \geq 0, \zeta \geq 0 \quad (3.11.27)$$

Fakat bu problemin duali

$$\text{Max } b'W^1 + \gamma_2 C'W^2$$

$$\begin{bmatrix} A' & -2\gamma_1 Q & \hat{\mu} \\ I & 0 & 0 \\ 0 & I & \hat{V} \\ 0 & -A & 0 \\ 0 & h' & 0 \end{bmatrix} \begin{bmatrix} W^1 \\ W^2 \\ W^3 \end{bmatrix} \leq \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

$$W^i, i=1,2,3 \text{ işaretinde kısıtlama yok} \quad (3.11.28)$$

Bu problem bir çözüme sahiptir.(Bölüm 3.4 te geliştirilen duallik teorisinden.) Bu çözüm $\hat{W}^1, \hat{W}^2$ ve $\hat{W}^3$ olsun. Öyle ki

$$b'\hat{W}^1 + \gamma_2 C'\hat{W}^2 = \hat{\xi} > 0.$$

Ayrıca, $\hat{\xi} > 0$ iken çözüm, tamamlayıcı gevşeklik özelliğinden çıkarılır ve dual problemdeki uygun kısıtlama eşitlik olara sağlanır:

$$h'\hat{W}^2 = 1$$

Üstelik çözüme uygun $\hat{V}$, $\hat{\eta}$, $\hat{\mu}$, $\hat{\lambda}$ için aşağıdaki durumların biri mutlaka sağlanır

$$(1) \quad \hat{V}_j > 0, \hat{\mu}_j = 0$$

$$(2) \quad \hat{V}_j = 0, \hat{\mu}_j > 0$$

$$(3) \quad \hat{V}_j = 0, \hat{\mu}_j = 0$$

$$1.\text{Durum} \Rightarrow (A'\hat{W}^1 - 2\gamma_1 Q\hat{W}^2)_j = 0$$

$$2.\text{Durum} \Rightarrow (\hat{W}^2)_j = 0$$

$$3.Durum \Rightarrow (A' \hat{W}^1 - 2\gamma_1 Q \hat{W}^2)_j \leq 0 \text{ ve } (\hat{W}^2)_j \leq 0$$

(Bu üç ifade tamamlayıcı gevşeklik özelliğinin bir sonucudur.)

Böylece

$$\hat{W}^2 (A' \hat{W}^1 - 2\gamma_1 Q \hat{W}^2)_j \geq 0$$

olur. Aynı şekilde $\hat{\eta}_i$ ve $\hat{\lambda}_i$ lerin pozitif veya sıfır olduğu üç durumun ele alınmasıyla

$$\hat{W}^1 A \hat{W}^2 \leq 0$$

ifadesini yazabiliriz. Buradan

$$\hat{W}^2 Q \hat{W}^2 \leq 0$$

olur.

$$Q = \begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix} \text{ ve } C = (0, e)$$

olduğunu hatırlatalım. Böylece V için

$$h = \gamma_2 C + 2\gamma_1 Q \hat{V} = \begin{pmatrix} 0 \\ \gamma_2 e + 2\gamma_1 \hat{V}_R \end{pmatrix}$$

dir. Burada $\hat{V}_R, (\hat{V}_{2p+1}, \dots, \hat{V}_{2p+n})'$ ile verilen vektördür. Burada $(h_1, \dots, h_{2p})$ şeklindeki h nin hepsi sıfırdır. Bu

$$h' \hat{W}^2 = 1 = \sum_{j=2p+1}^{2p+n} h_j \hat{W}_j^2$$

eşitliğini ifade eder. Şimdi, $\hat{W}^2 Q \hat{W}^2 \leq 0$ ifadesine tekrar bakarsak Q pozitif yarı tanımlı olduğunda $\hat{W}^2 Q \hat{W}^2 \geq 0$ dır. Buradan $\hat{W}^2 Q \hat{W}^2 = 0$ ortaya çıkar. Ya da

$$\sum_{j=2p+1}^{2p+n} (\hat{W}_j^2)^2 = 0 \quad (Q \text{ nun yapısından})$$

dır. Bu, ancak ve yalnız, $\hat{W}_j^2 = 0$ $j = 2p+1, \dots, 2p+n$ ise mümkündür. Bu da $h' \hat{W}^2 = 1$ olması ile çelişir. Böylece ispat tamamlanır.

3.11.1 Düşünceler ve Tartışmalar:

(1) Problem (3.11.2) de d nin kesin olarak işaretinde kısıtlama olmasına gerek yoktur. Ayrıca β , $\beta_1 - \beta_2$ şeklinde ifade edilmeden de kullanılabilir. Daha sonra Sonuç (3.11.1), (3.11.3) ve (3.11.5) in kısıtsız değişkenler için de ispatlanmasını isteriz. Bu mümkündür ve V daha büyük pozitif olduğunda $V' \mu = 0$ kısıtlaması da olmaz.

(2) Eğer en küçük kareler problemini çözmek isteseydik Problem (3.11.5) te sadece $\gamma_2 = 0$ olması yeterli olacaktı. Fakat bu durumda Problem (3.11.1) i ele alırken aşağıdaki sonuç uygulanır.

Sonuç 3.11.8 : V^0 aşağıdaki problem için bir optimal çözümdür

$$\text{Min } V' Q V + C' V$$

$$A V = b$$

$$V \text{ işaretinde kısıtlama yok} \quad (3.11.29)$$

ancak ve yalnız $AV^0 = b$ ve

$$-2QV^0 - A'\lambda^0 = C'$$

olacak şekilde bir λ^0 var ise.

Problem (3.11.1) de $C = 0$, $A = [X, I]$, $V = [\beta, d]$, $Q = \begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix}$ ve $b = Y$ dir.

Burada, yukarıda yazdığımız sonuç aşağıdakileri ifade eder

$$X\beta^0 + d^0 = Y \quad (3.11.30)$$

$$-X'\lambda^0 = 0 \quad (3.11.31)$$

ve

$$-2d^0 - \lambda^0 = 0 \quad (3.11.32)$$

veya

$$\lambda^0 = -2d^0 \quad (3.11.33)$$

ve

$$d^0 = Y - X\beta^0 \quad (3.11.34)$$

(3.11.31) de (3.11.33) ve (3.11.34) ü kullanarak ve daha basitleştirerek

$$X'X\beta^0 = X'Y \quad (3.11.35)$$

elde ederiz. Eğer $X'X$ in rankı p ise klasik en küçük kareler çözümü olan

$$\beta^0 = (X'X)^{-1} X'Y \quad (3.11.36)$$

yazabiliriz. Böylece, eğer sadece, genel en küçük kareler regresyonla uğraşırsak, matematiksel programlama yaklaşımına ihtiyaç olmadığını gözleriz. Tabi ki, bu yaklaşımda β üzerinde negatif olmama ve eşitsizlik şeklinde kısıtlamaların sağlanmasına ihtiyaç vardır.

(3) Algoritma (3.11.1), Q nun pozitif tanımlı olduğu genel quadratik programlama (QP) probleminde de kullanılabilir. Bu durumda Sonuç (3.11.7) nin ispatında $\hat{W}^2 Q \hat{W}^2 = 0 \Rightarrow \hat{W}^2 = 0$ olur fakat $h' \hat{W}^2 = 1$ şeklinde bir çelişki oluşur. Bu algoritma aynı zamanda Q pozitif yarı tanımlı olduğunda da uygulanabilir.

(4) Optimallik için gerekli ve yeterli durumları içeren sonuçlara dayalı QP yi çözmek için başka algoritmalar da vardır.

(5) Hesaplama bakımından daha üstün olan metodlar, lineer tamamlayıcılık probleminden çıkan sonuçlar kullanılarak elde edilebilir.

(6) Algoritma (3.11.1) deki bazın büyüklüğü $2p+3n$ dir. p ve n küçük olsa bile $2p+3n$ büyük olacaktır. Bu yüzden Problem (3.11.5) i çözerken bu algoritmayı yürütmek için bir bilgisayar programı kullanırız.

KAYNAKLAR

- ANDREWS, D.F., and A.M. HERZBERG** (1985) Data. Springer-Velag, New York
- ARTHANARI, T.S. and DODGE, Y.** (1993). "Mathematical Programming in Statistics." John Wiley and Sons.
- BLOOMFIELD, P., and W. STEIGER** (1983). Least Absolute Deviations: Theory, Applications and Algorithms. Birkhauser, Boston.
- CHARNES, A. ; COOPER, W.W. and FERGUSON, R.O.** (1955). "Optimal Estimation of Executive Compensation by Lineer Programming." Manage. Sci. 1, 138.
- KOESKER, R., and BASSET** (1982). Tests of linear hypotheses and L_1 estimation. Econometrica, vol.50, pp. 1577-1583
- KOLMAN, B. and BECK, R.E.** (1980). "Elementary Linear Programming in Statistics." John Wiley and Sons.
- LEE, T.C. ; JUDGE, G.G. and ZELLNER, A.** (1968). "Maximum Likelihood and Bayes Estimation of Transition Probabilities." J. Am. Stat. Assoc. 63, 1162.
- LIITTSCHWAGER, J.M. and WANG, C.** (1978). "Integer Programming Solution of a Classification Problem." Manage. Sci. 24, 1515.
- McKEAN, J.W., AND R.M. SCHRADER** (1987). Least absolute errors analysis of variance. In Y. Dodge (ed.) (1987a)
- PUKELSHEİM, F.** (1978). "A Quick Introduction to Mathematical Programming with Applications to Most Powerful Tests, Nonnegative Variance Estimation and Optimal Desing Theory. Technical Report 128, Stanford Univ. Press, Stanford, Calif.
- RAO, J.N.K.** (1979). "Optimization in the Desing of Sample Survets. In Optimizing Methods in Statistics. J.S. Rustagi, Ed. Academic. Preaa, New York, pp. 419-434.
- TAHA. H.** (2000). Operations Research an Introdicton. Prentice-Hall, Inc.
- VINCENT, A.S.** (1989). Linear Programming with statistical Applications. Iowa State University Press, Ames, Iowa.

ÖZGEÇMİŞ

1978 yılında Antakya’da doğdum. İlk orta ve lise öğrenimimi Hatay’ın Dört Yol ilçesinde tamamladım. 1995 yılında Çukurova Üniversitesi Fen Edebiyat Fakültesi Matematik Bölümü’ne girmeye hak kazandım. Dört yıllık lisans eğitiminden sonra 1999 yılında mezun oldum. Aynı yıl Çukurova Üniversitesi Fen Edebiyat Fakültesi Matematik Bölümü’nde yüksek lisans öğrenimime ve ayrıca matematik öğretmeni olarak göreve başladım. Bir yıl YADEM’de dil hazırlık programına devam ettim. Halen Antakya’da bir ilköğretim okulunda matematik öğretmeni olarak görev yapmaktayım.

