

**UKUROVA NİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

DOKTORA TEZİ

Gölin TABAKAN

YARI PARAMETRİK REGRESYONDA TAHMİN METODLARI

İSTATİSTİK ANABİLİM DALI

ADANA, 2009

ÇUKUROVA ÜNİVERSİTESİ

FEN BİLİMLERİ ENSTİTÜSÜ

YARI PARAMETRİK REGRESYONDA TAHMİN METODLARI

Gülin TABAKAN

DOKTORA TEZİ

İSTATİSTİK ANABİLİM DALI

**Bu tez/...../2009 Tarihinde Aşağıdaki Jüri Üyeleri Tarafından
Oybirliği/Oyçokluğu İle Kabul Edilmiştir.**

İmza

Prof. Dr. Fikri AKDENİZ
DANIŞMAN

İmza

Prof. Dr. Hamza EROL
ÜYE

İmza

Prof. Dr. Altan ÇABUK
ÜYE

İmza

Prof. Dr. Olcay ARSLAN
ÜYE

İmza

Prof.Dr. Hülya ÇINGİ
ÜYE

Bu tez Enstitümüz İstatistik Anabilim Dalında hazırlanmıştır.

Kod No

Prof. Dr. Aziz ERTUNÇ

Enstitü Müdürü

İmza ve Mühür

Bu çalışma

**• Ç.Ü. Bilimsel Araştırma Projeleri Birimi Proje No: FEF2007D4
tarafından desteklenmiştir.**

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.

ÖZ

DOKTORA TEZİ

YARI PARAMETRİK REGRESYONDA TAHMİN METODLARI

Gülin TABAKAN

ÇUKUROVA ÜNİVERSİTESİ

FEN BİLİMLERİ ENSTİTÜSÜ

İSTATİSTİK ANABİLİM DALI

Danışman: Prof. Dr. Fikri AKDENİZ

Yıl: 2009, Sayfa: 131

Jüri: Prof. Dr. Fikri AKDENİZ

Prof. Dr. Hamza EROL

Prof. Dr. Altan ÇABUK

Prof. Dr. Olcay ARSLAN

Prof. Dr. Hülya ÇİNGİ

Bu tezde parametrik model varsayımlarının sağlanmaması durumunda parametrik ve parametrik olmayan yaklaşımların en iyi yönlerini alan yarı parametrik regresyon modeli tanıtılmıştır. Yarı parametrik regresyon modelinde parametre tahmini için üç farklı yaklaşım ele alınmıştır. İlk olarak, cezalı en küçük kareler temeline dayanan eğrisel çizgi düzeltme yöntemi tanıtılmıştır ve bu yöntemle dayalı olarak yarı parametrik modelin değerlendirilmesi için düzeltme parametresi seçim problemi ele alınmıştır. İkinci olarak, yarı parametrik regresyon modelinde parametre tahmini için dalgacık tabanlı tahmin yöntemi tanıtılmıştır. Son olarak yarı parametrik modelde parametre tahmini için fark alma fikri ele alınmıştır ve yarı parametrik modelde çoklu iç ilişkinin varlığı durumunda farka dayalı ridge tahmin edici olarak adlandırılan yeni bir tahmin edici önerilmiştir. Farka dayalı ridge tahmin edici ile farka dayalı tahmin edici hata kareler ölçütüne göre karşılaştırılmıştır.

Anahtar Kelimeler: Eğrisel çizgi düzeltme, Düzeltme parametresi, Fark alma, Yarı parametrik regresyon, Dalgacık tahmin.

ABSTRACT

PhD THESIS

ESTIMATION METHODS IN SEMIPARAMETRIC REGRESSION

Glin TABAKAN

DEPARTMENT OF STATISTICS

INSTITUTE OF NATURAL AND APPLIED SCIENCES

UNIVERSITY OF UKUROVA

Supervisor: Prof. Dr. Fikri AKDENİZ

Year: 2009, Pages: 131

Jury: Prof. Dr. Fikri AKDENİZ

Prof. Dr. Hamza EROL

Prof. Dr. Altan ABUK

Prof. Dr. Olcay ARSLAN

Prof. Dr. Hlya INGI

In this thesis, the semiparametric regression model that combines the best features of the parametric and the nonparametric approaches are introduced when the parametric model assumptions are violated. Three different approaches for the semiparametric regression model estimation are considered. Firstly, the smoothing spline estimation procedure based on penalized least squares is introduced and for the evaluation of a semiparametric model based on this procedure the smoothing parameter selection criteria are considered. Secondly, a wavelet based approach is introduced for estimating a semiparametric regression model. Finally the idea of differencing to the parameter estimation in semiparametric regression model is considered and a new difference-based estimator which is called difference-based ridge estimator when the presence of multicollinearity in the semiparametric regression model is suggested. The differencing estimator and difference-based ridge estimator are analyzed and compared in the sense of mean-squared error criterion.

Key Words: Differencing, Semiparametric regression, Smoothing parameter, Spline smoothing, Wavelet estimation.

TEŞEKKÜR

Doktora çalışmam sırasında vermiş olduğu her türlü destek ve katkı için danışman hocam, Sayın Prof. Dr. Fikri AKDENİZ'e ve çalışmalarında beni destekleyen bölüm hocalarıma teşekkürlerimi sunarım. Ayrıca, doktora süresi boyunca bana destek olarak, her an yanımda olan sevgili arkadaşım Gülsen KIRAL'a teşekkür ederim.

Çalışmalarım sırasında her an yanımda olan, çalışmam boyunca beni daima yüreklendiren ve büyük özveride bulunan biricik annem Günseli ve babam Necati TABAKAN'a teşekkürlerimi bir borç bilirim.

İÇİNDEKİLER	SAYFA
ÖZ.....	I
ABSTRACT.....	II
TEŞEKKÜR.....	III
İÇİNDEKİLER.....	IV
ÇİZELGELER DİZİNİ.....	VII
ŞEKİLLER DİZİNİ.....	VIII
SİMGELER VE KISALTMALAR.....	IX
TANIMLAR.....	XI
1.GİRİŞ.....	1
2. REGRESYON MODELLERİ, REGRESYONDA DÜZELTME KAVRAMI VE PARAMETRİK OLMAYAN REGRESYONDA DÜZELTME YÖNTEMLERİ.....	6
2.1. Parametrik Regresyon.....	6
2.2. Parametrik Olmayan Regresyon.....	8
2.3. Yarı Parametrik Regresyon.....	10
2.4. Parametrik Olmayan Regresyonda Düzeltme Kavramı.....	11
2.5. Pürüzlülük Cezaları.....	13
2.5.1. Bir Eğrinin Pürüzlülüğünü Ölçme.....	13
2.5.2. Cezalı En Küçük Kareler Regresyonu.....	14
2.6. Eğrisel Çizgiler (Splines).....	15
2.6.1. Kübik Eğrisel Çizgiler.....	15
2.6.2. İkinci Türev-Değer Gösterimi.....	17
2.7. Parametrik Olmayan Regresyonda Düzeltme Yöntemleri (Doğrusal Düzelticiler).....	18

2.7.1. Kernel (Çekirdek) Düzeltici.....	20
2.7.2. Yerel (Local) Regresyon Düzeltici.....	25
2.7.3. k -En Yakın Komşu Tahmin Edici	31
2.7.4. Eğrisel Çizgi Düzeltme (Spline Smoothing) Yöntemi.....	35
2.7.4.1. Eğrisel Çizgi Düzeltme Tahmin Edicisinin Elde Edilmesi.....	37
2.8. Doğrusal Düzelticilerin İstatistiksel Özellikleri	38
2.8.1. Hata Kareler Ortalaması (MSE).....	38
2.8.2. Serbestlik Derecesi.....	40
2.9. Hata Varyansının Tahmini	42
2.9.1. Hata Kareler Yaklaşımı.....	43
2.9.2. Yerel Fark Alma Yaklaşımı.....	43
 3. YARI PARAMETRİK MODELDE EĞRİSEL ÇİZGİ DÜZELTME YÖNTEMİNE DAYALI PARAMETRE TAHMİNLERİ VE PARAMETRELER İLE İLGİLİ SONUÇLAR.....	46
3.1. Yarı Parametrik Modeller için Cezalı En Küçük Kareler Yaklaşımı	48
3.1.1. Backfitting Yöntemi.....	50
3.1.2. Kısmi Eğrisel Çizgi (Partial Spline) Yaklaşımı.....	52
3.2. Speckman Yaklaşımı.....	53
3.3. Varyans-Kovaryans Tahmini.....	56
3.4. Yarı Parametrik Modele ait Çıkarımlar.....	57
3.4.1. Parametrik Bileşen için Çıkarım.....	57
3.4.2. Parametrik Olmayan Bileşen için Çıkarım.....	60
 4. DÜZELTME PARAMETRESİ (BANT GENİŞLİĞİ) SEÇİM YÖNTEMLERİ.....	62
4.1. Çapraz Geçerlilik Ölçütü (Cross Validation- CV).....	63
4.2. Genelleştirilmiş Çapraz Geçerlilik Ölçütü (Generalized Cross Validation- GCV).....	64
4.3. Mallows'un C_p Ölçütü (Mallows' C_p Criterion).....	66
4.3.1. GCV Ölçütü ve Mallows'un C_p Ölçütü Arasındaki İlişki.....	68

4.4. Akaike Bilgi Ölçütü (Improved Akaike Information Criterion- <i>AIC</i>).....	70
4.5. Klasik Pilotları Kullanan Risk Tahmini (Risk Estimation using Classical Pilots- <i>RECP</i>).....	71
4.6. Tam Kat Düzeltme (Exact Double Smoothing- <i>EDS</i>).....	72
4.7. Yerleştirme (Plug-in) Yöntemleri.....	73
5. YARI PARAMETRİK MODELLERİN DALGACIK TAHMİNİ.....	79
5.1. Parametrik Olmayan Regresyonda Dalgacık Tahmin.....	80
5.2. Yarı Parametrik Modelde Dalgacık Tahmin.....	82
6. YARI PARAMETRİK MODELDE FARK ALMA YÖNTEMİNE DAYALI TAHMİNLER.....	93
6.1. Fark Matrisi ve Özellikleri.....	94
6.2. Yarı Parametrik Modelin Farka Dayalı Tahmini.....	96
6.3. Düzeltme ve Fark Alma Arasındaki İlişki.....	102
6.4. Deneysel Uygulama.....	104
6.4.1. Kanada Hane Benzin Talebi.....	104
6.5. Yarı Parametrik Modelde Parametrelerin Farka Dayalı Ridge Tahmin Edicisi.....	107
6.5.1. Önerilen Farka Dayalı Ridge Tahmin Edici.....	108
6.5.2. Farka Dayalı Tahmin Edici ve Farka Dayalı Ridge Tahmin Edicinin Karşılaştırılması.....	109
6.5.3. Farka Dayalı Ridge Tahmin Edici $\hat{\beta}_{Rdiff}(k)$ 'nin Üstünlüğü.....	112
6.5.4. Deneysel Uygulama.....	114
7. SONUÇLAR VE ÖNERİLER.....	119
KAYNAKLAR.....	121
ÖZGEÇMİŞ.....	125
EK	126

ÇİZELGELER DİZİNİ

SAYFA

Çizelge 2.1. Kernel fonksiyonları.....	22
Çizelge 6.1. En uygun fark alma ağırlıkları.....	100
Çizelge 6.2. Benzin talebi için parametre tahminleri.....	105
Çizelge 6.3. Farka dayalı tahmin edici ve k 'nın çeşitli değerlerine karşılık gelen farka dayalı ridge tahmin edicinin tahmin edilen varyans ve mse değerleri.....	116

ŞEKİLLER DİZİNİ

SAYFA

Şekil 2.1. Doğrusal regresyon ve interpolasyon.....	12
Şekil 2.2. Ethanol verisine ait saçılım grafiği.....	19
Şekil 2.3. Bazı kernel fonksiyonları.....	23
Şekil 2.4. h düzeltme parametresinin farklı değerleri için ethanol veri setine ait kernel regresyon düzeltme.....	24
Şekil 2.5. Ethanol veri kümesi için normal (Gaussian) ve üçgen (triangle) kernel fonksiyonlarının kullanımı.....	24
Şekil 2.6. Tricube fonksiyonu.....	29
Şekil 2.7. $h = 3$ için yerel regresyon düzeltme.....	31
Şekil 2.8. Gözlem değerlerinin saçılım grafiği.....	33
Şekil 2.9. $k = 1$ için k -NN tahmin edici.....	34
Şekil 2.10. $k = 2$ için k -NN tahmin edici.....	34
Şekil 2.11. $k = 20$ için k -NN tahmin edici.....	35
Şekil 2.12. $k = 109$ için k -NN tahmin edici.....	35
Şekil 6.1. Benzin için hane talebi.....	106
Şekil 6.2. Benzin için hane talebi: Aylık etkiler.....	106
Şekil 6.3. Yarı parametrik model: Elektrik dağıtımında parametrik olmayan bileşen için uyum eğrisi.....	115
Şekil 6.4. k 'nın çeşitli değerleri için $\hat{\beta}_{rdiff}(k)$ ve $\hat{\beta}_{diff}$ tahmin edicilerinin tahmin edilen mse değerleri.....	117
Şekil 6.5. k 'nın çeşitli değerleri için $\hat{\beta}_{rdiff}(k)$ ve $\hat{\beta}_{diff}$ tahmin edicilerinin tahmin edilen örneklem varyans değerleri.....	117

SİMGELER VE KISALTMALAR DİZİNİ

ADD	: Ayrık dalgacık dönüşüm
AIC_C	: Geliştirilmiş Akaike bilgi ölçütü (Improved Akaike Information Criterion)
$CEKK$	: Cezalı en küçük kareler
CV	: Çapraz geçerlilik (Cross Validation)
C_p	: Mallow' un C_p ölçütü (Mallows' C_p Criterion)
D	: Fark alma matrisi
DD	: Dalgacık dönüşümü
df	: Serbestlik derecesi
EDF	: Eşdeğer serbestlik derecesi
EDS	: Tam kat düzeltme (Exact Double Smoothing)
$E(RSS)$:	Hata kareler toplamının beklenen değeri
EKK	: Alışılmış en küçük kareler (OLS)
f	: Düzgün fonksiyon
$f \in C^2[a, b]$	: İlk 2 türevi $[a, b]$ kapalı aralığında sürekli olan tüm f fonksiyonlarının kümesi
f''	: f fonksiyonunun 2. türevi
$\dot{f}(x)$	: x_i tasarım noktalarının yoğunluğu
GCV	: Genelleştirilmiş çapraz geçerlilik (Generalized Cross Validation)
h_{opt}	: En uygun (optimum) düzeltme parametresi
H	: Şapka matrisi
H_p	: Kısmi eğrisel çizgi tahmin edicisi için düzeltme matrisi
H_s	: Speckman tahmin edicisi için düzeltme matrisi
K	: Kernel fonksiyonu
k	: Ridge regresyon tahmin edicisi için yanlılık parametresi
k -NN	: k -En yakın komşu tahmin edici
L	: Gecikme matrisi

$\log$	: Logaritma
m	: Fark alma derecesi
mse	: Skaler değerli hata kareler ortalaması
MSE	: Hata kareler ortalaması
$MSSE$	: Ortalama hata kareler toplamı
N	: Tekrarlanma matrisi
NCS	: Doğal kübik eğrisel çizgi
ORR	: Alışılmış ridge regresyon tahmin edici
$PRESS$	: Ön tahmin hata kareler toplamı
PS	: Pürüzlülük cezası
R	: Risk fonksiyonu
$RECP$	: Klasik pilotları kullanan risk tahmini (Risk Estimation using Classical Pilots)
RSS	: Hata kareler toplamı
SE	: Standart hata
sgn	: İşaret (signum) fonksiyonu
S_{λ}	: Parametrik olmayan regresyon için şapka matrisi (düzeltme matrisi)
$tr(A)$	: A matrisinin izi
λ	: Düzeltme parametresi (bant genişliği), eşik değeri
λ_p	: Düzeltme parametresinin pilot tahmini
V_p	: Kısmi eğrisel çizgi yaklaşımı için varyans-kovaryans matrisi
V_s	: Speckman yaklaşımı için varyans-kovaryans matrisi
$\dot{=}$	: Son etkiler hariç eşittir
$\cong$	: Yaklaşık olarak eşittir

TANIMLAR

Bant matris: Sadece sıfır olmayan elemanları köşegen etrafındaki bir bantta olan bir kare matrise bant matris denir. Böylece eğer $A = (a_{ij})$ kare matrisi $|i - j| \leq 1$ olduğunda $a_{ij} = 0$ şartını sağlarsa, o zaman A bir bant matristir. 4×4 tipinde bir A bant matrisine örnek olarak aşağıdaki matrisi verebiliriz:

$$A = \begin{bmatrix} a_{11} & a_{12} & 0 & 0 \\ a_{21} & a_{22} & a_{23} & 0 \\ 0 & a_{32} & a_{33} & a_{34} \\ 0 & 0 & a_{43} & a_{44} \end{bmatrix}$$

İnterpolasyon: İlk kez Uygulamalı Matematik biliminin bir alt kategorisi olan sayısal analiz yöntemlerinde tanımlanan ve elde var olan değer noktalarından yola çıkarak, farklı bir yerde ve değeri bilinmeyen bir diğer noktadaki olası değeri bulmaya/tahmin etmeye yarayan yöntemlerin tümüne verilen genel isimdir. En basit tanımı ile ‘var olan sayısal değerleri kullanarak, boş noktalardaki değerlerin tahmin edilmesi’ olarak açıklanmaktadır. ‘interpolasyon’ sözcüğü yerine bazen ‘tahmin’ sözcüğü kullanılmaktadır.

İnterpolasyon genelde mühendislik ve deneylere/ölçümlere dayalı benzeri bilim dallarında, toplanan verilerin bir fonksiyon eğrisine uydurulması amacı ile kullanılmaktadır. Elde toplanan verinin dağınık ve özellikle aşırı heterojen olduğu durumlarda interpolasyon ile boş noktalardaki değerlerin bulunması önem kazanmaktadır.

Matrisin iz’i: Bir $A = (a_{ij})_{n \times n}$ kare matrisinin iz’i A matrisinin esas köşegen elemanlarının toplamı olarak tanımlanır ve $iz(A)$ ile gösterilir. Yani,

$$iz(A) = \sum_{i=1}^n a_{ii}$$

olarak yazılır.

Pozitif tanımlı matris: Simetrik A matrisi için $x'Ax$ karesel formu ele alınsın. $\forall x \neq 0$ için $x'Ax$ karesel formu 0'dan büyük ise A matrisine pozitif tanımlı (pd) matris denir (Graybill, 1983).

Pozitif yarı tanımlı matris: Simetrik A matrisi için $x'Ax$ karesel formu ele alınsın. En az bir $x \neq 0$ için $x'Ax$ karesel formu 0'dan büyük ya da 0'a eşit ise A matrisine pozitif yarı tanımlı (psd) matris denir (Graybill, 1983).

Negatif tanımlı olmayan matris: Bir A matrisi pozitif tanımlı ya da pozitif yarı tanımlı ise A matrisine negatif tanımlı olmayan (nnd) matris denir (Graybill, 1983).

o -küçük: $\varphi(x)$ ve $g(x)$ fonksiyonları, x_0 noktasının herhangi Ω_0 civarında tanımlanmış ve $g(x) \neq 0$ koşulunu sağlayan fonksiyonlar olsun (x_0 noktasında $g(x) = 0$ olabilir). Bu nokta sonlu veya sonsuz olabilir.

$$\lim_{x \rightarrow x_0} \frac{\varphi(x)}{g(x)} = 0$$

oluşu,

$$\varphi(x) = o(g(x)), \quad x \rightarrow x_0$$

şeklinde yazılmakta ve $x \rightarrow x_0$ 'da $\varphi(x)$ fonksiyonu $g(x)$ 'e göre o -küçüktür diye okunmaktadır (Halilov ve ark., 1999).

Tek yanlı yönlü türev (One-sided directional dervative): L , $\mathbb{R}^m$ 'de bir fonksiyon ve x sonlu $L(x)$ 'de bir nokta olsun. Bir h yönü ile ilgili L 'nin x 'deki L' tek yanlı yönlü türevi varsa,

$$L'(x;h) = \lim_{\alpha \rightarrow 0^+} \frac{L(x+\alpha h) - L(x)}{\alpha}$$

biçimindedir. Burada,

$$-L'(x;-h) = \lim_{\alpha \rightarrow 0^-} \frac{L(x+\alpha h) - L(x)}{\alpha}$$

olduğuna dikkat ediniz. Böylece tek yanlı yönlü türev $L'(x;h)$, ancak ve ancak $L'(x;-h)$ varsa ve $L'(x;-h) = -L'(x;h)$ ise iki yanlı yönlü türevdir.

Yönlü türev: Belirli bir birim vektör doğrultusunda elde edilen türev.

1. GİRİŞ

İstatistik biliminin en önemli konularından birini regresyon analizi oluşturmaktadır. Regresyon analizi matematik, finans, ekonomi, tıp, ziraat, mühendislik gibi bilim dallarında çok yaygın olarak kullanılmaktadır. Gözlenen bir olayın değerlendirilirken hangi olayların etkisi altında olduğunun araştırılması regresyon analizinin temelini oluşturur. Regresyon analizi ile bağımlı ve bağımsız (açıklayıcı) değişkenler arasında bir ilişki var mıdır? Eğer bir ilişki varsa bu ilişkinin gücü nedir? Değişkenler arasında ne tür bir ilişki vardır? Belirli koşulların kontrol edilmesi durumunda özel bir değişken veya değişkenler grubunun diğer değişken veya değişkenler üzerindeki etkisi nedir ve nasıl değişir? gibi sorulara cevap aranmaya çalışılır. Regresyon analizi yapılırken gözlem değerlerinin ve etkilenilen olayların bir matematiksel gösterimle yani bir fonksiyon yardımı ile ifade edilmesi gerekir. Regresyon analizi bağımlı ve bağımsız değişkenler arasındaki ortalama ilişkinin matematiksel bir fonksiyonla ifade edilmesinde, bağımsız değişkenlerle bağımlı değişkenin doğrusal bir ilişki içerisinde olduğunu varsayar. Bir tek bağımsız değişkenin kullanıldığı regresyon analizi tek değişkenli regresyon analizi, birden fazla bağımsız değişkenin kullanıldığı regresyon analizi de çok değişkenli regresyon analizi olarak adlandırılır.

Regresyon analizi bazı varsayımlara dayanır. Bu varsayımların en önemlisi, bağımlı ve bağımsız değişkenler arasındaki ilişkinin şeklinin biliniyor olmasıdır. Varsayımların sağlanmadığı durumlarda yapılan tahminler iyi bir tahmin olma niteliğine sahip olamazlar. Bu durumda daha iyi tahmin yapabilmek amacıyla parametrik regresyondaki doğrusallık varsayımının esnetilmesine olanak sağlayan regresyon yöntemlerine ihtiyaç duyulur. Bu yöntemler parametrik olmayan (nonparametric) ve yarı parametrik (semiparametric) regresyon yöntemleri olarak bilinen regresyon modelleridir.

Yarı parametrik regresyon modeli için var olan yaklaşımların tümü farklı parametrik olmayan regresyon yöntemlerine bağlıdır. Yarı parametrik regresyon modelleri karmaşık veri kümelerini bizim anlayabileceğimiz biçimde özetleyip, uygulamada verilerin önemsiz detaylarını göz ardı ederken önemli olan özelliklerini

muhafaza eder ve böylece sağlam kararlar verilmesini sağlarlar (Ruppert, Wand ve Carroll, 2003).

Zamana bağlı olarak elde edilen verilerin analizinde yarı parametrik regresyon yöntemi yaygın olarak kullanılmaktadır. Genellikle tarım, tıp ve biometri alanlarındaki denemelerden elde edilen boylamsal (longitudinal) veriler zamana bağlı olarak sürekli bir ölçekle ölçülür ve aynı deneme ünitesinden (bireyden) farklı zamanda elde edilen ölçümler farklı değerler alırlar. Ancak veriler birbirleriyle ilişkilidir. Bu veriler birden fazla davranışın birbirini izleyecek şekilde aynı deneme ünitelerine uygulanması sonucu elde edilen ölçümlerdir (Laird ve Ware, 1982). Boylamsal çalışmaların çoğunda, ilgilenilen sonuç değişkeni üzerine zamanın ve sürekli bağımsız değişkenin etkileri modelde yer alır. Aynı birey üzerinde yer ve zamana bağlı olarak birden fazla gözlem yapıldığı zaman hata değişkenleri arasında korelasyon (otokorelasyon) söz konusu olmaktadır. Bu gibi durumlarda bazı varsayımlar geçerli olmaz. Bundan dolayı zamanla ilgili değerlendirmeler yapmak parametrik yöntemler için genel bir sorun olmaktadır. Bu gibi durumlarda parametrik olmayan yöntemler kullanılabilir. Ancak parametrik olmayan yöntemlerle bağımsız değişken sayısı ikiden fazla olduğu zaman analizlerin yapılması ve grafiklerin yorumlanması zor olmaktadır. Alternatif bir yöntem olarak yarı parametrik modeller kullanılabilir. Yarı parametrik modellerde şansa bağlı etkiler ve zaman etkisi parametrik olmayan yöntemlerle, sürekli bağımsız değişken etkileri ise parametrik olan yöntemlerle modele dahil edilir.

Yarı parametrik regresyon modeli parametrik ve parametrik olmayan regresyon fonksiyonunun birleşiminden oluşması nedeni ile “*kısmi doğrusal model*” olarak da adlandırılmaktadır. Bu çalışmada kısmi doğrusal model yerine yarı parametrik model ifadesi kullanılacaktır. Yarı parametrik regresyon modeli, parametrik değişkenlerin etkilerinin sıfır olması ya da bu tür değişkenlerin analizde yer almadığı durumlarda parametrik olmayan regresyon modeli olarak ele alınır. Ayrıca yarı parametrik modeller en az esneklikten (düşük dereceli bir polinom) en çok esnekliğe (yüksek dereceli bir polinom-interpolasyon) doğru bir genişletmesi olan, bir düzeltme parametresine (λ) sahip az sayıda parametre ile özetlenemeyen

bir parametrik olmayan regresyon fonksiyonu içerir. Bu nedenle parametrik regresyon modellerinden çok daha esnektir.

Bu çalışmada parametrik regresyon yöntemleri ile çözümlenemeyen regresyon problemlerinin çözümünde oldukça kapsamlı uygulama alanına sahip olan ve temelde parametrik olmayan bir yöntem olan yarı parametrik regresyon modeli ele alınmış, yarı parametrik regresyon modelinde parametre tahmin yöntemleri incelenmiş ve bu modelle ilgili çıkarımlar yapılmıştır. Çalışmada yarı parametrik regresyon modelinde parametrelerin tahmini için üç farklı yaklaşım ele alınmıştır. Bu yaklaşımlardan ilki rezidü tabanlı tahmin ediciler, ikinci yaklaşım dalgacık tabanlı tahmin yaklaşımı, üçüncü yaklaşım ise f fonksiyonundan kaynaklanan verideki eğilimi elimine etmeyi amaçlayan fark alma yaklaşımıdır. Bahsedilen bu yaklaşımlar temelde parametrik olmayan regresyon yöntemleri için kullanılan yaklaşımlardır.

İlk yaklaşım olan rezidü tabanlı tahmin edicilerde, parametrik olmayan ve yarıparametrik regresyon modellerinde parametre tahmini için çoğunlukla düzeltme (smoothing) yöntemleri kullanılır. Düzeltme yöntemlerinden biri eğrisel çizgi düzeltme (spline smoothing) yöntemidir. Bu yöntemin temeli, cezalı en küçük kareler regresyonuna dayanır. Parametrik olmayan ve yarı parametrik olan regresyon modellerinin kestiriminde kullanılan cezalı en küçük kareler yönteminde alışılmış en küçük kareler yönteminden farklı olarak hata kareler toplamına bir düzeltme parametresine (λ) sahip olan bir ceza fonksiyonu eklenir. Ceza fonksiyonunun eklenmesindeki amaç esnek eğimli uyumlar ile sabit eğimli uyumlar arasında bir uzlaşma sağlamaktır. Rezidü tabanlı tahmin ediciler düzeltme miktarına bağlıdır ve düzeltme parametresinin seçimi pratikte zor bir problemdir (Ruppert ve ark. 1997).

Bu çalışmada parametrik olmayan ve yarı parametrik regresyon modellerinde düzeltme parametresinin seçimi için, çapraz geçerlilik ölçütü (CV), genelleştirilmiş çapraz geçerlilik ölçütü (GCV), geliştirilmiş Akaike bilgi ölçütü (AIC_c) ve Mallows' un (C_p) ölçütü olarak bilinen *klasik yöntemler*, pilot tahminlerin seçimini gerektiren *risk tahmin yöntemleri* EDS ve $RECP$, yerel doğrusal düzeltmede (local linear smoothing) kullanılan alışılmış *yerleştirme* (plug-in) yöntemleri ele alınmıştır.

Ele alınan ikinci yaklaşımda yarı parametrik modelin parametre tahmini; parametrik olmayan bileşenin dalgacık katsayılarının l_1 normunun cezalandırılıp, rezidü vektörünün l_2 normunun karesinin minimize edilmesi ile elde edilmiştir.

Parametre tahmini için ele alınan üçüncü yaklaşımda ise $f(\cdot)$ 'nin türevlenebilir ve x ordinatlarının birbirine yakın olması ile parametrik olmayan ve yarı parametrik regresyon modellerindeki f fonksiyonunun elimine edilmesi sağlanır. Parametrik olmayan ve yarı parametrik regresyon modellerindeki parametrik olmayan etkinin elimine edilmesini sağlayan fark alma fikri yarı parametrik modelde Ahn ve Powell (1993), Yatchew (1997) tarafından ele alınmıştır.

Sonuç olarak yukarıda bahsedilen tahmin yöntemleri aracılığı ile yarı parametrik regresyon modeli ile çok daha tutarlı tahminler yapılır.

Son yıllarda klasik regresyon yöntemlerinin yetersiz kaldığı göz önüne alınarak parametrik olmayan ve yarı parametrik regresyonla ilgili bir çok çalışma yapılmıştır; Engle, Granger, Rice, Weiss (1986) Hava durumu ve elektrik satışları arasındaki ilişkinin yarı parametrik tahminleri, Heckman (1986) Kısmi doğrusal modelde eğrisel çizgi düzeltme (spline smoothing), Robinson (1988) Yarı parametrik regresyon modelinin tahmini, Speckman (1988) Kısmi doğrusal modellerde kernel düzeltmesi, Wahba (1990) Gözleme dayalı veriler için eğrisel çizgi modelleri, Cuzick (1992) Yarı parametrik toplamsal regresyon, Carroll ve ark. (1997) Genelleştirilmiş kısmi doğrusal single index model, Yatchew (1997) Kısmi doğrusal modelin bir temel tahmin edicisi, Eubank, Kambour, Kim, Klipple, Reese (1998) Kısmi doğrusal modellerde tahmin, Härdle ve ark. (1998) Genelleştirilmiş kısmi doğrusal modelde parametre tahmini ve tahmin edilen modeli değerlendiren bir test istatistiği, Liang ve ark. (1999) Yarı parametrik kısmi doğrusal hata ölçümlü modelde tahmin, Schimek (2000) Eğrisel çizgi düzeltme ile kısmi doğrusal modellerde tahmin ve çıkarımlar, Yatchew (2000) elektrik dağıtımında ölçek ekonomileri, Lee (2003) Eğrisel çizgi düzeltmede düzeltme parametresinin seçimi ile ilgili bir simülasyon çalışması, Liang ve Wang (2005) Kısmi doğrusal single index ölçüm hatalı modeller, Klipple ve Eubank (2007) Kısmi doğrusal modeller için farka dayalı varyans tahmin edicileri ve bunun gibi konu ile ilgili bir çok çalışma yapılmıştır.

Bu alıřma 7 blmden oluřmaktadır. Birinci blm olan giriř blmnde tezin konusu ve nemi, bu konuda yapılmıř alıřmalar, tezin ierięi hakkında bilgiler verilmiřtir.

İkinci blmde parametrik, parametrik olmayan ve yarı parametrik regresyon modelleri tanıtılmıř, regresyonda dzeltme kavramı, przllk ceza yaklařımı ve parametrik olmayan regresyonda dzeltme yntemleri incelenmiřtir.

nc blmde bu alıřmanın temel konusu olan yarı parametrik regresyon modeli ele alınmıř, bu modelin parametrelerinin tahmini iin, przllk ceza yaklařımı ve Speckman (1988) tarafından nerilen Speckman yaklařımı ele alınmıřtır.

Drdnc blmde parametrik olmayan ve yarı parametrik regresyon uyumunun iyi bir řekilde yapılabilmesi iin gerekli olan ve bant geniřlięi olarak da adlandırılan dzeltme parametresi seim yntemleri, klasik yntemler, risk tahmin yntemleri ve yerel doęrusal dzeltme ynteminde kullanılan yerleřtirme (plug-in) yntemleri olarak incelenmiřtir.

Beřinci blmde, Chang ve Qu (2004) tarafından yarı parametrik modelde parametre tahmini iin geliřtirilen bir dalgacık tabanlı tahmin yaklařımı ele alınmıřtır. Bu yaklařım eęrisel izgi dzeltme, kernel ve paralı polinom yaklařımları gibi yarı parametrik modeller iin var olan geleneksel dzeltme yntemlerindeki parametrik olmayan fonksiyonun kısıtlayıcı przszlęnden kaınan, parametrik olmayan regresyon yntemleri iin dalgacık yaklařımının bir uzamasıdır.

Altıncı blmde parametrik olmayan ve yarı parametrik regresyon modellerinde yaygın olarak kullanılan fark alma yntemi tanıtılıp bu ynteme dayalı olarak parametre tahmini elde edilmiřtir. Ayrıca regresyon problemlerinde oklu i iliřkinin varlıęı durumunda nerilen ridge regresyon kavramı ele alınmıř, yarı parametrik modelde oklu i iliřkinin varlıęı durumunda farka dayalı ridge tahmin edici olarak adlandırılan yeni bir tahmin edici nerilmiřtir ve farka dayalı ridge tahmin edici ile farka dayalı tahmin edici hata kareler ltne gre karřılařtırılmıřtır.

Yedinci blmde ise sonular ve neriler verilmiřtir.

2. REGRESYON MODELLERİ, REGRESYONDA DÜZELTME KAVRAMI VE PARAMETRİK OLMAYAN REGRESYONDA DÜZELTME YÖNTEMLERİ

Regresyon, ele alınan değişkenlerden birinin bağımlı ve diğerlerinin bağımsız (açıklayıcı) olması durumunda bağımlı değişkenin (y) bağımsız değişkenlerin ($z_1, z_2, \dots, z_p$) bir fonksiyonu olarak ifade edilmesidir. Bir fonksiyonel bağıntı olan regresyon analizi; bağımsız değişkenlerdeki değişmelerin bağımlı değişkeni hangi yönde ve hangi oranda etkilediğini belirtir. Regresyon analizinin asıl amacı; değişkenler arasındaki fonksiyonel bağıntıyı en iyi şekilde ifade edecek matematiksel denklemi kurmak ve bu denklemi bağımlı değişkenin değerlerinin tahmin edilmesinde ve istatistiksel analizde kullanmaktır. Regresyon analizi ($z_1, z_2, \dots, z_p$) bağımsız değişkenlerine göre (y) bağımlı değişkenin koşullu ortalamasının fonksiyonel bağımlılığını ortaya koyar.

Bu bölümde bağımsız değişkenlere göre bağımlı değişkenin fonksiyonel bağımlılığını belirleyen parametrik regresyon, parametrik olmayan regresyon ve yarı parametrik regresyon modelleri tanıtılmış, parametrik olmayan regresyonda düzeltme (smoothing) kavramı ele alınmış ve düzeltme yöntemleri incelenmiştir.

2.1. Parametrik Regresyon

Parametrik regresyon bağımlı ve bağımsız değişkenler ile bu değişkenler arasındaki ortalama ilişkinin matematiksel bir fonksiyonla ifade edilmesi ve bu fonksiyondaki parametre vektörlerinin açık bir şekilde gösterilmesidir. Parametrik regresyon, regresyon fonksiyonunun $z_1, z_2, \dots, z_p$ bağımsız değişkenlerinin bir doğrusal fonksiyonu olarak yazılabildiğini varsayar. $E(y|Z)$ koşullu beklenen değeri Z biliniyorken y 'nin ortalama dağılımının Z ile fonksiyonel ilişkisini gösterir. Başka bir ifade ile $Z = (z_1, z_2, \dots, z_p)$ bağımsız değişkenlerindeki değişime karşılık y bağımlı değişkeninin ortalama tepkisini ifade eder ve

$$E(y|Z) = Z\beta \quad (2.1)$$

biçiminde veya,

$$y = Z\beta + \varepsilon \quad (2.2)$$

biçiminde yazılır. Burada $\varepsilon = y - E(y|Z)$ ifadesi $E(y|Z)$ koşullu beklenen değerinin y ’den sapması olarak tanımlanır (Härdle ve ark., 2004). Model (2.2)’de gözlem sayısı n , bağımsız değişken sayısı p olmak üzere; y $(n \times 1)$ boyutlu bağımlı değişken vektörü, Z $(n \times p)$ boyutlu ve p ranklı bağımsız değişkenler matrisi, β $(p \times 1)$ boyutlu bilinmeyen regresyon katsayıları vektörü, ε ise gözlenemeyen 0 ortalamalı ve sabit varyanslı rasgele hataların $(n \times 1)$ boyutlu vektörüdür. Parametrik regresyon modelinde amaç modelin uydurulması ve model uygunluğunun araştırılmasıdır. Doğrusal regresyon modelini belirlemek için bilinmeyen β parametrelerini tahmin etmek gerekir. Eşitlik (2.2)’de (y_i, z_i) gözlem değerlerine karşılık gelen nokta ile bu noktanın en küçük kareler yöntemi ile elde edilmiş olan doğru üzerindeki izdüşümleri toplamı yani $\varepsilon_i = y_i - \hat{y}_i$ sıfır olmalıdır. Ayrıca bu farkın kareler toplamı minimum olmalıdır. Bu durumda (2.2) ile verilen modelde parametre tahminleri,

$$\sum_{i=1}^n \varepsilon_i^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n (y_i - z_i' \hat{\beta})^2$$

eşitliğinden elde edilir.

Parametrik yaklaşım tümüyle varsayımlara dayalıdır. Eşitlik (2.2)’de Z bağımsız değişkenler arasındaki fonksiyonel yapının $Z\beta$ biçiminde doğrusal olduğu ve β parametrelerinin sonlu olduğu varsayılmaktadır. Ancak burada değişkenler

arasında doğrusal olmayan ilişkiler varsa parametrik yöntemler yerine parametrik olmayan yöntemlerin kullanılması gerekir.

Günümüzde parametrik olan istatistiksel yöntemlere karşılık olarak parametrik olmayan yöntemler geliştirilmiştir. Parametrik olmayan yaklaşımda β parametre vektöründen bahsedilmemekte ve bağımsız değişkenler arasındaki fonksiyonel yapı da bilinmemektedir. Hiçbir varsayım gerektirmediğinden dolayı bu yaklaşımın uygulamalarda daha fazla tercih edilmesi gerekirken, birçok problem ile karşılaşılmasından dolayı nadiren kullanılmaktadır. Bu problemler içinde en fazla dikkat çeken, bağımsız değişken sayısının fazla olması durumunda tahmin ve yorumlamada güçlük çekilmesidir (Härdle ve ark., 2004).

Bağımlı değişkeninin bağımsız değişkenlerden bazıları ile doğrusal ilişki içinde fakat bazıları ile de ilişkisinin kolayca parametreleştirilemediği durumlar vardır. Bu gibi durumlarda parametrik ve parametrik olmayan regresyon yöntemlerinin yeterli olamayacağı açıktır. Bu nedenle, hem parametrik hem de parametrik olmayan regresyon modellerini içeren ve bu modellerin özel bir durumu olan '*kısmi doğrusal model*' olarak da adlandırılan '*yarı parametrik regresyon modeli*' parametre tahmininde daha uygun bir sonuç verecektir.

2.2. Parametrik Olmayan Regresyon

Bir y bağımlı değişkeni ve bu değişkenle ne tür bir ilişki içerisinde olduğu bilinmeyen bir x bağımsız değişkeninin yer aldığı basit '*parametrik olmayan regresyon modeli*',

$$y_i = f(x_i) + \varepsilon_i, \quad i = 1, \dots, n \quad (2.3)$$

biçimindedir. Burada $f \in C^2[a, b]$ olan bir düzgün fonksiyon, x_i parametrik olmayan bağımsız değişkenlere ait gözlem değerleri, y_i bağımlı değişkene ait gözlem değerleri, ε_i 0 ortalamalı ve σ^2 sabit varyanslı, bağımsız özdeş olarak dağılan rasgele hata terimleridir.

Parametrik olmayan regresyon analizinin amacı parametreleri tahmin etmekten çok bilinmeyen yanıt fonksiyonu (ortalama fonksiyon) olan $f(x)$ 'i tahmin etmektir. $x = (x_1, x_2, \dots, x_n)$ bağımsız değişkenlerinin ve y bağımlı değişkeninin ölçümleri arasındaki ilişkiyi açıklayan en yaygın yöntem, $E(y|x) = f(x)$ koşullu beklenen fonksiyonunu tahmin etmektir. Bu ilişkiyi açıklayan genel parametrik olmayan regresyon modeli,

$$\begin{aligned} y &= E(y|x) + \varepsilon \\ &= f(x) + \varepsilon \end{aligned} \quad (2.4)$$

biçiminde ifade edilir. (2.4) eşitliğinde $y = (y_1, \dots, y_n)'$, $f = (f_1, \dots, f_n)' = (f(x_1), \dots, f(x_n))'$ ve $\varepsilon = (\varepsilon_1, \dots, \varepsilon_n)'$ olup f belirgin bir şekle sahip olmayan x açıklayıcı değişkenlerinin bir fonksiyonudur. Ayrıca f ikinci mertebeden sürekli türeve sahip olan fonksiyonlar uzayının bir elemanıdır ve $f \in C^2[a, b]$ modelin düzeltme kısmıdır. f fonksiyonunu tahmin etmek için alt bölüm 2.7'de verilen düzeltme yöntemleri kullanılır. Parametrik olmayan tahmin edicinin yakınsama oranı genellikle parametrik tahmin edicilerden daha yavaştır. Bu nedenle parametrik tahmin ile kıyaslamada parametrik olmayan yöntemler çok büyük örneklem hacimleri gerektirir (Yatchew, 1998). Ancak parametrik modellerde olduğu gibi bu tür modellerde de istatistiksel doğruluk örneklem hacmine değil tahmin edicilerin varyans ve kovaryanslarına bağlıdır (Heerde ve ark., 2001; Yatchew, 1998).

Parametrik olmayan regresyon yönteminin tahmin yaparken kısıtlayıcı varsayımları olmamasına rağmen bazı sakıncaları vardır. Bağımsız değişken sayısı fazla olduğu zaman tahmin yapmak zor olmakta ve elde edilen grafikler karmaşık bir yapıda olmaktadır. Bu durum '*boyutluluk sorunu* (curse of dimensionality)' olarak adlandırılır. Ayrıca parametrik olmayan yöntemle kesikli bağımsız değişkenleri dikkate almak ve bağımsız değişken sayısındaki artışa bağlı olarak y değişkenine ait bireysel etkileri yorumlamak zor olmaktadır. Parametrik olmayan bu yöntemin sakıncaları yarı parametrik regresyon modeli kullanılarak giderilmektedir. Yarı

parametrik regresyon modeli hem parametrik hem de parametrik olmayan regresyon modelinin her ikisini birlikte kullanır. Bu nedenle yarı parametrik regresyon modeli parametrik modellerin kısıtlayıcı varsayımlarından etkilenmemekle birlikte parametrik olmayan yöntemlerin cazip özelliklerini bir araya getirmektedir.

Parametrik olmayan regresyon aykırı gözlemlerin (outliers) bulunduğu veri setleri için önemli bir analiz yöntemidir. İstatistiksel çalışmalarda aykırı gözlemlerin etkilerini farklı biçimlerde ele alan güçlü (robust) parametrik yöntemler bulunmaktadır. Bununla birlikte, aykırı gözlemlerden dolayı parametreler bozulduğu için bu güçlü yöntemler bile uygun çözümler üretemeyebilir ve verinin gerçek yapısı modele yansıtılamaz. Bu durumda parametrik olmayan regresyon, ön bilgi sağlamaktadır (Härdle, 1994). Parametrik yaklaşımda çok fazla varsayım yapıldığından dolayı sonuçların güvenilirliği giderek azalmaktadır. Parametrik olmayan yaklaşımda ise hiçbir varsayım yapılmamakta fakat bağımsız değişken sayısının fazla olması durumunda model tahmininin elde edilmesi zor olmaktadır. Yarı parametrik yaklaşım, parametrik ve parametrik olmayan yaklaşım arasında bir orta yol bulmayı amaçlamaktadır.

2.3. Yarı Parametrik Regresyon

Yarı parametrik regresyon modelleri bağımlı değişkenin bağımsız değişkenlerden bazıları ile ilişkisinin parametreleştirilebildiği fakat diğer bağımsız değişken veya değişkenlerle ilişkisinin kolayca parametreleştirilemediği modellerdir. Değişkenlerin skaler olduğu ve f fonksiyonunun parametrik bir aile içerisinde bulunmadığı '*yarı parametrik regresyon modeli*',

$$\begin{aligned} y &= E(y|Z, x) + \varepsilon \\ y &= Z\beta + f(x) + \varepsilon \end{aligned} \tag{2.5}$$

biçimindedir. (2.5) eşitliğindeki y , Z , f ve ε önceden tanımlandığı gibidir. (2.5) modelinde z parametrik değişkenler vektörünün parametrik olmayan

$x = (x_1, x_2, \dots, x_n)$ değişkeni ile düzgün bir regresyon ilişkisine sahip olduğu varsayılır (Yatchew, 2003) ve bu modelle parametre tahmini yapılırken normallik varsayımına gerek duyulmaz (Zeger ve Diggle, 1994).

Parametrik olmayan model tahmininde yorumlanabilir sonuçlar elde etmek için en fazla iki açıklayıcı değişken ile çalışmak mümkün iken, yarı parametrik yöntemde k tane açıklayıcı değişkenin bağımlı etkisini incelemek mümkündür. Ayrıca parametrik modeldeki kadar varsayım yapılmaması nedeni ile bu yaklaşımın uygulamalı çalışmalarda kullanılması önerilmektedir (Horowitz, 1993).

2.4. Parametrik Olmayan Regresyonda Düzeltme Kavramı

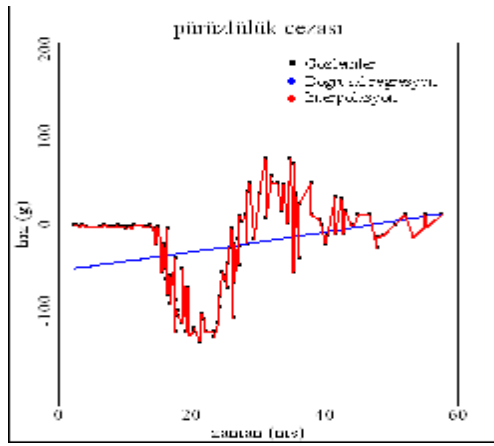
Parametrik olmayan regresyon modelinde y_i 'ye karşı x_i 'nin dağılım grafiği incelendiğinde açıklaması mümkün bir regresyon ilişkisi her zaman kurulamaz. Çünkü veri kümesindeki aykırı değerler uygun olmayan yorumlamalara neden olabilir. Parametrik olmayan regresyon analizinin amacı bilinmeyen yanıt fonksiyonu olan $f(x)$ 'in uygun analizini elde etmektir. Gözlenen hataların azaltılması ile y 'nin x 'e göre ortalama bağımlılığının önemli ayrıntılarını vermek yorumu kolaylaştırır. Bu eğri yaklaştırma işlemi genel olarak '*düzeltme (smoothing)*' olarak adlandırılır. Düzeltme fikrinin temelinde verileri bir eğriye uydurmak ve daha basit fonksiyonların birleşimi olabilen esnek fonksiyonları kullanmak yatar. '*Düzeltilici*

(smoother)' $\hat{f}(x) = \sum_{i=1}^n w_{xi} y_i$ ise $x_1, x_2, \dots, x_n$ olmak üzere bir ya da birden fazla

bağımsız değişkenin fonksiyonu olan y bağımlı değişkenin sahip olduğu eğilimi ifade etmek için kullanılan bir araçtır ve bağımlı değişkenin kendisinden daha az değişken olan bir eğriyi tahmin etmeyi amaçlamaktadır. Bir düzelticinin en önemli özelliği, değişkenler arasındaki ilişkinin şeklini kesin bir biçimde belirlememesidir (bir doğru gibi) ve bu özelliğinden dolayı parametrik olmayan regresyonda sık kullanılan bir araçtır (Hastie ve Tibshirani, 1990). Parametrik regresyonda elde edilen doğru kesin parametrik bir biçime sahip olduğu için düzeltici değildir.

Düzeltilme yönteminin sonuç değişkeni yani bir düzeltici tarafından oluşturulan tahmin ‘*düzgün (smooth)*’ olarak adlandırılır (Härdle, 1994).

Eşitlik (2.4) ve (2.5) deki f fonksiyonu yaklaşık olarak doğrusal ise f fonksiyonunu tahmin etmek için kullanılan klasik yöntemlerden biri doğrusal regresyon yaklaşımıdır. Bu yaklaşımda hata kareler toplamı minimum yapılarak f fonksiyonu tahmin edilmeye çalışılır. Fakat f fonksiyonu doğrusal değil ise bu yaklaşım başarısız olabilir ve bu durumda doğrusal regresyon uyumu sadece değişkenler arasındaki bir ilişkinin varlığını ifade eder (Bakınız Şekil 2.1).



Şekil 2.1. Doğrusal regresyon ve interpolasyon

Şekil 2.1’de görüldüğü gibi f regresyon fonksiyonunun diğer bir tahmini (x_i, y_i) , $i = 1, \dots, n$ verilerinin interpolasyonu ile elde edilebilir. ‘*İnterpolasyon*’ elde varolan (bilinen) değer noktalarından yola çıkarak, farklı bir yerde ve değeri bilinmeyen bir diğer noktadaki olası değeri tahmin etmeye yarayan yöntemlerin tümüne verilen genel isimdir. En basit tanımı ile var olan sayısal değerleri kullanarak, boş noktalardaki değerlerin tahmin edilmesi olarak açıklanır. Bu tahmin edici için doğrularla birleştirilen gözlemlerde bireysel eğimler ile sabit eğimlilik sağlanır. Bu tahmin edici için hata kareler toplamı $RSS(f) = 0$ olur. Doğrusal regresyon uyumu verideki bilginin çok azını kullanırken, bu uyumda verideki bilgi daha fazla içerilir ancak verilerin yararlı bir özetini sağlamada başarısız olur ve daha

önemlisi modelin rasgele hata kavramından kaynaklanan ve (2.4) denklemindeki regresyon fonksiyonuna bağlı olabilen verilerdeki özelliklerin veya esas eğilimin doyurucu bir şekilde açıklanmasını gerçekleştirmez. Bu durumda parametre tahmininde doğrusal regresyon ve interpolasyon uyumlarının esnetilmesini sağlayan daha esnek yaklaşımların kullanılması zorunlu hale gelir.

Alt bölüm 2.5 ve 2.6’da parametrik olmayan fonksiyonun tahmini için istatistiksel çalışmalarda çok sık olarak kullanılan ve düzeltme yöntemleri başlığı altında alt bölüm 2.7.4’de ele alınan eğrisel çizgi düzeltme yöntemi için temel kavramlar verilmiştir.

2.5. Pürüzlülük Cezaları

En basit şekliyle pürüzlülük ceza yaklaşımı, regresyon doğrusu boyunca klasik doğrusal regresyondaki model varsayımlarını esneten bir yöntemdir. Basit doğrusal regresyon Şekil 2.1’de görülen veriler için uygun değildir. Şekil 2.1’deki gözlem değerlerinden tamamıyla esnek eğimli uyumlar (interpolasyon) ile sabit eğimli uyumlar (bir doğru) arasında bir uzlaşmaya gereksinim olduğu görülmektedir. Bunu sağlamanın bir yolu, modele regresyon fonksiyonun eğimi ile bağlantılı olan ceza fonksiyonunun eklenmesidir. Pürüzlülük ceza yaklaşımının esas amacı hızlı olarak dalgalanan bir eğrinin eğilimini ölçmek ve daha sonra eğri tahmininde sabit eğimli uyumlar ile esnek eğimli uyumlar arasında gerekli uzlaşmayı sağlayacak şekilde tahmin problemini ortaya koymaktır (Green ve Silverman, 1994).

2.5.1. Bir Eğrinin Pürüzlülüğünü Ölçme

Bir $[a, b]$ aralığında tanımlı f eğrisinin ne kadar pürüzlü ve dalgalı olduğunu ölçmenin birçok farklı yolu vardır. İki kez sürekli türevlenebilir f eğrisinin pürüzlülüğünü ölçmenin bir sezgisel cazip yolu f fonksiyonunun ikinci türevinin karesinin integralini $\int_a^b \{f''(x)\}^2 dx$ almaktır. Bu ölçüme göre sadece doğrusal $f(x)$

fonksiyonları sıfır pürüzlülüğe sahipken diğer $C^2[a,b]$ sınıfındaki tüm fonksiyonlar pozitif bir pürüzlülüğe sahiptir. Pürüzlülüğü hesaplamak için çeşitli yöntemler bulunmaktadır. Özellikle regresyon konusunda, modele bir sabitin veya doğrusal bir fonksiyonun eklenmesi pürüzlülüğün ölçümünü etkilemez. Öyle ki eğer iki fonksiyon sadece bir sabit veya bir doğrusal fonksiyona göre farklı ise onların pürüzlülüğü benzer olmalıdır. Bu doğal olarak göz önüne alınan eğrinin ikinci türevine bağlı olan bir fonksiyonel pürüzlülük fikrine yol açar (Green ve Silverman, 1994).

2.5.2. Cezalı En Küçük Kareler Regresyonu

Parametrik olmayan regresyonda büyük bir öneme sahip olan ve 2.7.4 alt bölümünde ele alınan eğrisel çizgi düzeltme (spline smoothing) yönteminin temeli cezalı en küçük kareler regresyonuna dayanır. $[a,b]$ aralığında iki kez türevlenebilen herhangi bir f fonksiyonu ve bir $\lambda > 0$ düzeltme parametresi ile cezalı kareler toplamı,

$$S(f) = \sum_{i=1}^n \{y_i - f(x_i)\}^2 + \lambda \int_a^b \{f''(x)\}^2 dx \quad (2.6)$$

olarak tanımlanır. (2.6) eşitliğinden elde edilen cezalı en küçük kareler (CEKK) tahmin edici $\hat{f}$ iki kez türevlenebilen tüm f fonksiyonları sınıfında $S(f)$ fonksiyonunu minimum yapan değer olarak tanımlanır. Buradaki λ düzeltme parametresi belirlenmiş bir sabit olup, eğrinin pürüzlülüğünü ölçen $\int_a^b \{f''(x)\}^2 dx$ ile

verinin uyumunu ölçen $\sum_{i=1}^n \{y_i - f(x_i)\}^2$ ifadeleri arasında bir uzlaşma sağlar ve esneklik üzerine konulan önemin bir ölçüsü olarak görülebilir. $S(f)$ 'nin minimize edilmesi uyum iyiliği ve düzeltme arasındaki en iyi uzlaşmayı verecektir. Eğer λ büyükse $S(f)$ 'deki ana bileşen $\int_a^b \{f''(x)\}^2 dx$ pürüzlülük ceza terimi olacak ve

bundan dolayı minimum $\hat{f}$ çok az eğrilik gösterecektir. Limit durumunda ise λ sonsuza gitme eğiliminde ise $\int f''^2$ terimi sıfıra yaklaşacaktır ve $\hat{f}$ eğrisi doğrusal regresyon uyumu üretecektir. Eğer λ oldukça küçükse $S(f)$ 'ye ana katkı, hata kareler toplamı (RSS) olacaktır ve $\hat{f}$ tahmin eğrisi veriyi yakından izleyecektir. Limit durumunda λ sıfıra yaklaşma eğiliminde ise, $\hat{f}$ Şekil 2.1 de gösterilen esnek eğimli uyum (interpolasyon) eğrisine yaklaşacaktır. Burada önemli olan nokta verileri en iyi temsil eden bir eğri tahmini elde etmek için en uygun λ değerinin seçimidir (Green ve Silverman, 1994). λ değerinin seçimi ile ilgili olarak düzeltme parametresi seçim yöntemleri 4. Bölümde ayrıntılı olarak ele alınmıştır.

2.6. Eğrisel Çizgiler (Splines)

Bu kısımda (2.6) cezalı kareler toplamını minimum yapan f eğrisinin belirlenmesinde önemli kavramlar olan eğrisel çizgiler (spline), kübik eğrisel çizgiler ve doğal kübik eğrisel çizgiler ele alınmıştır.

'Eğrisel çizgi (spline)' bir dizi veri noktalarına polinom bir eğri uydurma veya bu noktalar arasından geçen ve bir çok parçadan oluşan esnek bir eğridir. İstatistik alanında eğrisel çizgilerin istatistiksel özelliklerinden çok sayısal çözümlemeleri hakkında çok geniş çalışmalar yapılmıştır (De Boor, 1978; Schumaker, 1993).

2.6.1. Kübik Eğrisel Çizgiler

Bir $[a, b]$ aralığında $a < x_1 < x_2 \dots < x_n < b$ koşulunu sağlayan $x_1, x_2, \dots, x_n$ reel sayılarının verildiğini varsayalım. $[a, b]$ aralığında tanımlanan f fonksiyonu aşağıdaki koşullar sağlanıyorsa bir 'kübik eğrisel çizgi' olarak adlandırılır:

- $(a, x_1), (x_1, x_2), (x_2, x_3), \dots, (x_n, b)$ alt aralıklarının her birinde f bir kübik polinom,

- f 'nin kendisi, birinci ve ikinci türevleri her bir x_i noktasında, böylece $[a, b]$ aralığının tümü üzerinde süreklidir.

Burada x_i noktaları '*düğüm noktaları*' olarak adlandırılır. Bir kübik eğrisel çizgiyi belirlemenin çeşitli yolları vardır. Bunlardan en açık şekilde olanı her bir kübik parçanın dört polinom katsayısı ile; Örneğin verilen $a_i, b_i, c_i, d_i, i = 0, \dots, n$ sabitleri için,

$$f(x) = f_i(x) = d_i(x - x_i)^3 + c_i(x - x_i)^2 + b_i(x - x_i) + a_i, \quad x_i \leq x \leq x_{i+1} \quad (2.7)$$

biçiminde belirlenmesidir. Burada $x_0 = a$ ve $x_{n+1} = b$ olarak tanımlanır. f ve ilk iki türevi üzerindeki süreklilik koşulları katsayılar arasındaki değişik bağıntıları ifade eder. Örneğin, $i = 0, \dots, n-1$ için x_{i+1} düğüm noktasındaki f 'nin sürekliliği,

$$f_{i+1}(x_{i+1}) = d_{i+1}(x_{i+1} - x_{i+1})^3 + c_{i+1}(x_{i+1} - x_{i+1})^2 + b_{i+1}(x_{i+1} - x_{i+1}) + a_{i+1} = a_{i+1}$$

olup, $f_i(x_{i+1}) = f_{i+1}(x_{i+1})$ 'e eşit olduğundan aşağıdaki eşitlik elde edilir:

$$f_i(x_{i+1}) = d_i(x_{i+1} - x_i)^3 + c_i(x_{i+1} - x_i)^2 + b_i(x_{i+1} - x_i) + a_i = a_{i+1}$$

$[a, b]$ aralığında bir kübik eğrisel çizgi a ve b noktalarında ikinci ve üçüncü türevleri sıfır oluyorsa '*doğal kübik eğrisel çizgi* (natural cubic spline: NCS)' olarak adlandırılır. Bu koşullar doğal sınır koşulları olarak adlandırılır. $d_0 = c_0 = d_n = c_n = 0$ ifadesi iki uç aralık olan $[a, x_1]$ ve $[x_n, b]$ aralıklarında f fonksiyonunun doğrusal olmasını ifade eder (Green ve Silverman, 1994).

2.6.2. İkinci Türev-Değer Gösterimi

Gerçekte (2.7) doğal kübik eğrisel çizgiyi hesaplamak için uygun bir gösterim değildir. Bir doğal kübik eğrisel çizgi her bir x_i düğüm noktasındaki değeri ve 2. türevinin verilmesi ile belirlenebilir. Bu gösterim ikinci türev-değer gösterimi olarak adlandırılır. f 'nin $x_1, x_2, \dots, x_n$ düğüm noktaları ile bir doğal kübik eğrisel çizgi olduğunu varsayalım. $f_i = f(x_i)$ ve $\gamma_i = f''(x_i)$, $i = 1, \dots, n$ olsun. Doğal kübik eğrisel çizginin tanımı gereği f 'nin x_1 ve x_n 'deki 2. türevi $\gamma_1 = \gamma_n = 0$ biçimindedir. $f = (f_1, \dots, f_n)'$ ve $\gamma = (\gamma_2, \dots, \gamma_{n-1})'$ vektörleri verilsin. Burada γ_i 'nin $(n-2)$ elemanı vardır. f ve γ vektörleri f eğrisini tamamen belirler ve herhangi bir x noktasındaki f 'nin türevleri ve değeri için f ve γ 'nin terimlerinde açık bir formül vermek mümkündür.

Verilen düğüm noktalarında vektörlerin bir doğal kübik eğrisel çizgi göstermesi için gerek ve yeter koşul Q ve R matrisleri olarak tanımlayacağımız iki matrise bağlıdır. Burada Q ve R $x_1 < \dots < x_n$ düğüm noktalarından elde edilen bant matrisleri olup $h_i = x_{(i+1)} - x_{(i)}$, $i = 1, \dots, n-1$ 'nin fonksiyonlarıdır. Daha genel olarak Q , $i = 1, \dots, n$; $j = 2, \dots, n-1$ için,

$$q_{j-1,j} = h_{j-1}^{-1}, \quad q_{jj} = -h_{j-1}^{-1} - h_j^{-1}, \quad q_{j+1,j} = h_j^{-1} \quad \text{ve} \quad |i-j| \geq 2 \quad \text{için} \quad q_{ij} = 0$$

elemanlarına sahip $n \times (n-2)$ tipinde bir matris olup Q 'nun sütunları γ 'nin girişlerine benzer biçimde $j = 2$ ile başlayarak numaralandırılmıştır ve Q 'nun ilk elemanı q_{12} 'dir. R elemanları $j = 2, \dots, n-1$ için,

$$r_{ii} = \frac{1}{3}(h_{i-1} + h_i), \quad i = 2, \dots, n-1$$

$$r_{i,i+1} = r_{i+1,i} = \frac{1}{6}h_i, \quad i = 2, \dots, n-2 \quad \text{ve} \quad |i-j| \geq 2 \quad \text{için} \quad r_{ij} = 0$$

olan $(n-2) \times (n-2)$ tipinde simetrik bir matristir (bkz. Ek). R matrisi her bir i için $|r_{ii}| > \sum_{i \neq j} |r_{ij}|$ olması anlamında kesin köşegen dominanttır. Nümerik doğrusal cebirdeki standart incelemeler R matrisinin pozitif tanımlı ($p.d$) olduğunu gösterir. Böylece,

$$K = QR^{-1}Q'$$

olacak şekilde bir K matrisi tanımlanabilir (Green ve Silverman, 1994).

Teorem 2.1. f ve γ vektörlerinin doğal kübik eğrisel çizgi belirtmesi için gerek ve yeter koşul aşağıdaki gibidir:

$$Q'f = R\gamma \quad (2.8)$$

(2.8) sağlanırsa bu durumda pürüzlülük cezası,

$$\int_a^b f''(x)^2 dx = \gamma'R\gamma = f'Kf \quad (2.9)$$

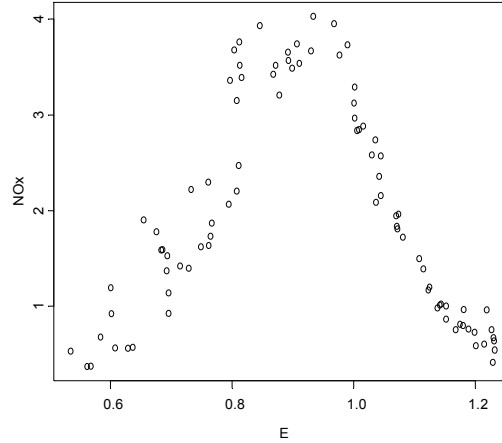
biçiminde yazılabilir (İspat için bkz. Green ve Silverman, 1994).

Teorem 2.1 pürüzlülük cezasının (2.9) biçiminde yazılmasından dolayı eğrisel çizgi düzeltme yöntemi ile tahmin yapmada gerekli bir teoremdir.

2.7. Parametrik Olmayan Regresyonda Düzeltme Yöntemleri (Doğrusal Düzelticiler)

Parametrik regresyon konusunda veri kümelerine ait grafikler çizildiğinde bağımlı değişken ile bağımsız değişkenler arasında doğrusal bir ilişkinin olduğu açık olarak görülmektedir. Böyle veriler için doğrusal model verileri kolay ve basit olarak analiz eden bir modeldir. İlgilenilen veri kümesinin doğrusal bir bağımlılık sergilemediği durumlarda bu modelin kullanımı uygun değildir. Örneğin, Şekil

2.2’de verilen saçılım grafiğini ele alalım. Şekilden bağımsız değişken E (ethanol) ve bağımlı değişken NOx (nitrik oksit) arasında bir ilişki olduğu fakat bu ilişkinin doğrusal bir uyum olmadığı açık olarak görülmektedir.



Şekil 2.2. Ethanol verisine ait saçılım grafiği

Şekil 2.2’deki saçılım grafiğine uygun bir model nasıl olmalıdır? Klasik doğrusal modele polinom terimlerini ekleyip en küçük kareler yönteminin uygulanması yani polinom regresyonun uygulanması bir yaklaşım olabilir. Polinom regresyon bir çok durumlarda kullanışlı olabilir. Bununla birlikte terimlerin seçimi her zaman açık değildir ve küçük etkiler önemli derecede büyük olabilir veya terimlerin yanlış seçimi ile tamamen kaybolabilir. Doğrusal olmayan verilerin analizi için bir başka yaklaşım, modelin şeklini belirlemek için verilere dayalıdır. Bu yaklaşım verilere uygun bir eğri için yerel noktaları kullanan cazip bir yaklaşımdır. Bu yöntemde, herhangi bir noktadaki eğri sadece bu noktadaki gözlemlere ve bazı belirlenmiş komşuluktaki noktalara bağlıdır. Böyle bir uyum orijinal gözlenen yanıtın çok farklı olmayan yanıtın tahminini elde ettiğinden dolayı elde edilen sonuç alt bölüm 2.4’de belirtildiği gibi ‘*düzgün (smooth)*’ olarak adlandırılır. Bu şekildeki uyumları elde etmek için kullanılan yöntemler ‘*saçılım grafiğini düzeltme*’ (scatterplot smoothers) olarak adlandırılır.

Düzeltilme yöntemleri farklı ölçümler arasındaki fonksiyonel ilişkileri bulmaya yarayan yöntemlerdir. Parametrik regresyon çerçevesinde olduğu gibi verilerin bir ya da daha fazla bağımsız değişken ve bir bağımlı değişken ölçümlerinden oluştuğu varsayılır. Parametrik regresyon yöntemleri bağımsız ve bağımlı değişkenler arasındaki ilişkiyi tanımlamak için fonksiyonel bir şekil (bir doğru gibi) belirlerler. Düzeltilme yöntemleri ise uydurulmuş eğrilerin şeklini belirlemek için kendileri veri noktaları sağlayan daha esnek bir yaklaşım sağlarlar (Loader, 2004).

Bu bölümde en yaygın düzeltilme yöntemlerinden bazıları ele alınacaktır. Ele alacağımız düzelticilerin hepsi,

$$\hat{f}(x) = \sum_{i=1}^n w_{xi} y_i \quad (2.10)$$

yazılışından dolayı doğrusal düzelticilerdir. Burada $\{w_{xi}\}_{i=1}^n$ vektörü $\{x_i\}_{i=1}^n$ vektörüne bağlı olarak hesaplanan ağırlık vektörüdür. Eşitlik (2.10)'daki $\hat{f}(x)$ regresyon tahmin edicisi 'düzeltici' olarak adlandırılır. Düzeltilme yöntemlerin amacı $f(x)$ ortalama fonksiyonu için parametrik bir model belirlemek değil verileri sağlayacak uygun bir fonksiyonel şekil belirlemektir (Loader, 2004).

2.7.1. Kernel (Çekirdek) Düzeltici

Düzeltilme yöntemlerinin en basiti kernel düzelticidir. Bu yöntemde bir x noktası ortalama fonksiyon $f(x)$ 'in tanım bölgesinde yer almıştır ve bir düzeltilme penceresi bu noktanın etrafında belirlenmiştir. En sık kullanılan düzeltilme penceresi basitçe $(x - h, x + h)$ aralığı olarak gösterilir. Kernel tahmini düzeltilme penceresi içindeki gözlemlerin bir ağırlıklandırılmış ortalamasıdır ve aşağıdaki şekilde gösterilir:

$$\hat{f}_h(x) = \frac{\sum_{i=1}^n K\left(\frac{x_i - x}{h}\right) y_i}{\sum_{j=1}^n K\left(\frac{x_j - x}{h}\right)} \quad (2.11)$$

Burada h düzeltme penceresinin yarıçapına karşılık gelen ve bant genişliği olarak da bilinen bir düzeltme parametresi, n gözlem sayısıdır. (2.11) ile verilen kernel düzeltici aşağıdaki biçimde de gösterilebilir (Loader, 2004):

$$\hat{f}_h(x) = \sum_{i=1}^n \left(\frac{K\left(\frac{x_i - x}{h}\right)}{\sum_{j=1}^n K\left(\frac{x_j - x}{h}\right)} \right) y_i = \sum_{i=1}^n w_{xi} y_i \quad (2.12)$$

(2.12) eşitliğinde w_{xi} $x - x_i$ uzaklığına bağlı i . gözlem y_i 'ye atanan ağırlık olarak yani i . gözlemin w ağırlığı $x_i - x$ uzaklığının bir fonksiyonu olarak tanımlanır. Bu yöntemde x 'e bağlı w_{xi} ağırlıkları yardımı ile y_i bağımlı değişkenlerinin bir ağırlıklı ortalaması olarak x noktasında regresyon fonksiyonu tahmin edilir ve w_{xi} katsayıları,

$$w_i(x) = w_{xi} = \frac{K\left(\frac{x_i - x}{h}\right)}{\sum_{j=1}^n K\left(\frac{x_j - x}{h}\right)} \quad (2.13)$$

biçiminde olup $\sum_{i=1}^n w_{xi} = 1$ dir. Bir doğrusal düzeltici uygun olarak tanımlanmış w_{xi} ağırlıkları için (2.12)'deki gibi gösterilen bir düzelticidir. Bir kernel düzeltici her bir hedef noktada tahmini üretmek için kernel olarak tanımlanan yerel ağırlıkların açıkça tanımlanmış bir kümesini kullanır. Bir kernel düzeltici hedef noktadan uzağa

taşındığından dolayı düzgün bir biçimde azalan ağırlıkları kullanır. Genellikle uzaklık küçükse ağırlık yüksek ve uzaklık büyükse ağırlık düşük olur. Ağırlıklar K tarafından belirtilir ve düzeltme parametresi (bant genişliği) olarak bilinen h tarafından kontrol edilir. (2.12) kernel tahmini bazen ‘Nadaraya-Watson’ tahmini olarak da adlandırılır (Härdle ve ark., 2004). Bu yaklaşım ağırlıkları açıkça belirtmek için K kernel fonksiyonunu kullandığından dolayı bir ‘kernel düzeltici’ olarak adlandırılır. Kernel fonksiyonları aşağıdaki özelliklere sahip fonksiyonlardır:

- 1) $K(u) \geq 0, \forall u$ için
- 2) $\int_{-\infty}^{+\infty} K(u) du = 1$
- 3) $K(-u) = K(u)$ simetriktir.

Bu özellikler aynı zamanda simetrik bir olasılık yoğunluk fonksiyonunun özellikleridir (Montgomery ve Peck 1992).

Uygulamada kullanılan bazı kernel fonksiyonları Çizelge 2.1’de verilmiştir;

Çizelge 2.1. Kernel fonksiyonları

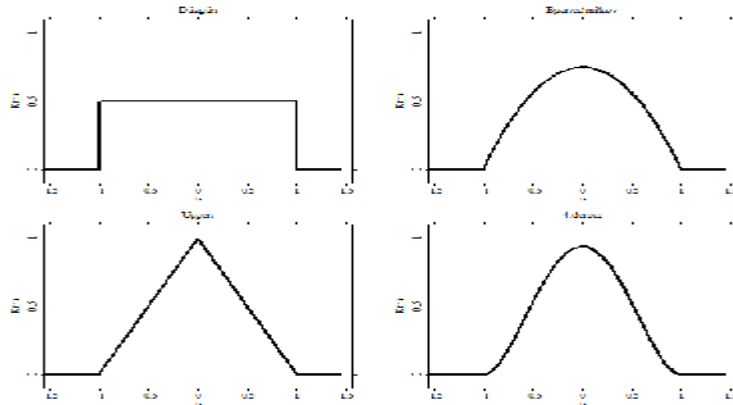
Kernel	$K(u)$	
Düzgün (Uniform)	$\frac{1}{2}I(u \leq 1)$	$\frac{1}{2}, u \in [-1, 1]$
Üçgen (Triangle)	$(1- u)I(u \leq 1)$	$(1- u), u \in [-1, 1]$
Epanechnikov	$\frac{3}{4}(1-u^2)I(u \leq 1)$	$\frac{3}{4}(1-u^2), u \in [-1, 1]$
4. dereceden (Quartic, Biweight)	$\frac{15}{16}(1-u^2)^2I(u \leq 1)$	$\frac{15}{16}(1-u^2)^2, u \in [-1, 1]$
6. dereceden (Triweight)	$\frac{35}{32}(1-u^2)^3I(u \leq 1)$	$\frac{35}{32}(1-u^2)^3, u \in [-1, 1]$

Normal (Gaussian)	$\frac{1}{\sqrt{2\Pi}} \exp\left(-\frac{1}{2}u^2\right) \quad u \in [-\infty, \infty]$
Cosinüs	$\frac{\Pi}{4} \cos\left(\frac{\Pi}{2}u\right) I(u \leq 1) \quad \frac{\Pi}{4} \cos\left(\frac{\Pi}{2}u\right), \quad u \in [-1, 1]$

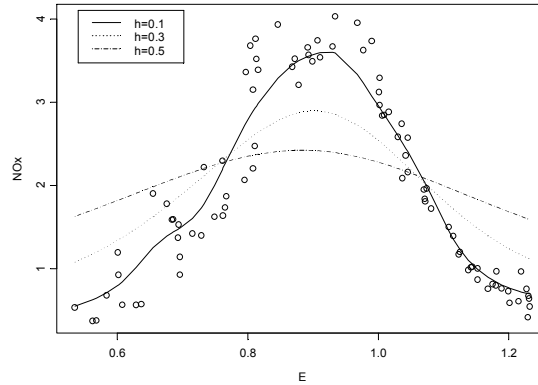
Çizelge 2.1’de verilen tüm fonksiyonlar yukarıda verilen kernel fonksiyonlarının özelliklerini sağlar. Örneğin 4.dereceden kernel fonksiyonu $K(u) = \frac{15}{16}(1-u^2)^2$, nun yukarıda verilen özelliklerine sahip olduğunu gösterelim:

- 1) $K(u) = \frac{15}{16}(1-u^2)^2 \geq 0, \quad u \in [-1, 1],$
- 2) $\int_{-1}^1 \frac{15}{16}(1-u^2)^2 du = \frac{15}{16} \int_{-1}^1 (1-2u^2+u^4) du = 1,$
- 3) $K(-u) = \frac{15}{16}(1-(-u)^2)^2 = \frac{15}{16}(1-u^2)^2 = K(u)$

Şekil 2.3’de bazı kernel fonksiyonları gösterilmiştir. Şekil 2.4’de ethanol veri setine ait kernel regresyon tahmini farklı düzeltme parametresi değerlerine bağlı biçimde grafiksel olarak gösterilmiştir:

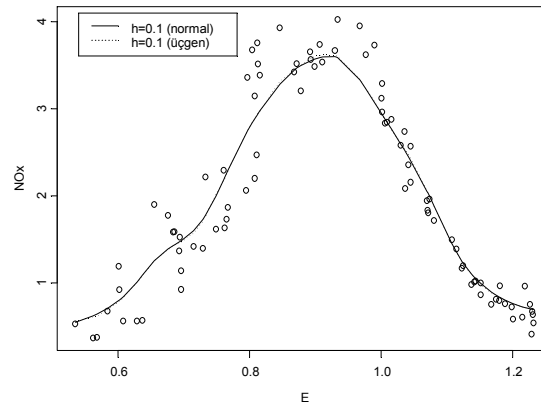


Şekil 2.3. Bazı kernel fonksiyonları: Düzgün (sol üst), Epanechnikov (sağ üst), Üçgen (sol alt), 4. dereceden (sağ alt) (Härdle ve ark., 2004)



Şekil 2.4. h düzeltme parametresinin farklı değerleri için ethanol veri setine ait kernel regresyon düzeltme

Çizelge 2.1’de verilen kernel fonksiyonlarından, üçgen (triangle) ve normal (Gaussian) kernel fonksiyonları bir çok veri kümesi için benzer düzeltme sağlamakta olup bu benzerlik grafiksel olarak Şekil 2.5’deki gibi gösterilmiştir:



Şekil 2.5. Ethanol veri kümesi için normal (Gaussian) ve üçgen (triangle) kernel fonksiyonlarının kullanımı

Hastie ve Tibshirani (1990) kernel fonksiyonunun seçiminin düzeltme parametresinin seçimine göre daha önemsiz olduğunu ortaya koymuştur. (2.11) kernel tahmininde h düzeltme parametresinin rolü önemlidir. Şöyle ki: Büyük h değerleri için eğri çok yavaş değişir ve düzeltme önemlidir. Bu durumda tahminin varyansı sınırlı fakat tahmin oldukça sapmalıdır. h değeri küçük olduğunda ise eğri

oldukça düzensizdir ve sapmalar sınırlı fakat tahminin varyansı büyüktür. Bu nedenle h parametresi sapmalar ve tahminin doğruluğu arasında bir uzlaşma sağlar (Härdle ve ark., 2004).

2.7.2. Yerel (Local) Regresyon Düzeltici

Yerel regresyon tahmini istatistiksel çalışmalarda 1970'lerin sonlarında önem kazanmış bir yöntem olup bu yönetime göre bir düzgün fonksiyon herhangi bir x noktasının komşuluğunda düşük dereceli bir polinom tarafından çok iyi tahmin edilebilir. Örneğin bir yerel doğrusal tahmin $x - h \leq x_i \leq x + h$ için,

$$f(x_i) \approx a_0 + a_1(x_i - x) \quad (2.14)$$

biçimindedir. Burada a_0 ve a_1 iki yerel parametre ve h yerel komşulukların büyüklüğüdür. Bir yerel ikinci dereceden tahmin,

$$f(x_i) \approx a_0 + a_1(x_i - x) + \frac{a_2}{2}(x_i - x)^2 \quad (2.15)$$

biçiminde tanımlanır (Loader, 2004).

p . dereceden genel bir polinom,

$$f(x_i) = \sum_{k=0}^p a_k \frac{(x_i - x)^k}{k!} \quad (2.16)$$

biçimindedir. (2.16) ile verilen fonksiyonun bir x noktasındaki yerel regresyon değerinin tahmin edilebilmesi için a_k ile gösterilen katsayıların bulunması gerekir. Bunun için aşağıdaki (2.17) ile verilen eşitlik kullanılır.

$$\begin{pmatrix} \hat{a}_0 \\ . \\ . \\ . \\ . \\ \hat{a}_p \end{pmatrix} = (X'WX)^{-1} X'Wy \quad (2.17)$$

Burada X tasarım matrisi, W köşegen elemanları $K\left(\frac{x_i - x}{h}\right)$ olan bir köşegen matris ve $y = (y_1, \dots, y_n)'$ yanıt değişkenlerden oluşan bir matris olup aşağıdaki biçimde gösterilirler:

$$X = \begin{pmatrix} 1 & . & . & . & \frac{(x_1 - x)^n}{n!} \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ 1 & . & . & . & \frac{(x_n - x)^n}{n!} \end{pmatrix}, W = \begin{pmatrix} K\left(\frac{x_1 - x}{h}\right) & . & . & . & 0 \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ 0 & . & . & . & K\left(\frac{x_n - x}{h}\right) \end{pmatrix}, y = \begin{bmatrix} y_1 \\ . \\ . \\ . \\ y_n \end{bmatrix}$$

Yerel tahmin yerel olarak ağırlıklandırılmış en küçük kareler ile uydurulabilir. Yani daha uzak gözlemlere daha düşük ve daha yakın gözlemlere daha yüksek ağırlıklar atayan ağırlıklı regresyon yapılabilir. Ağırlık fonksiyonu ve düzeltme parametresi kernel regresyonda olduğu gibi tanımlandığından, bu işlem bazen kernel regresyonu olarak da ifade edilir.

3. dereceden yerel bir polinom (2.16) eşitliğinde $p = 3$ alınırsa,

$$f(x_i) \approx a_0 + a_1(x_i - x) + \frac{a_2}{2}(x_i - x)^2 + \frac{a_3}{6}(x_i - x)^3 \quad (2.18)$$

biçiminde yazılıp katsayı tahminleri,

$$\sum_{i=1}^n K\left(\frac{x_i - x}{h}\right) \left(y_i - \left(a_0 + a_1(x_i - x) + \frac{a_2}{2}(x_i - x)^2 + \frac{a_3}{3!}(x_i - x)^3 \right) \right)^2 \quad (2.19)$$

ifadesini minimum yapacak biçimde seçilir. Her bir yerel *EKK* problemi bir x noktasında $\hat{f}(x)$ tanımlar; Eğer x değiştirilirse düzeltme ağırlıkları $K\left(\frac{x_i - x}{h}\right)$ değişecek ve böylece $\hat{a}_0, \hat{a}_1, \hat{a}_2, \hat{a}_3$ tahminleri değişecektir. (2.19) ağırlıklandırılmış *EKK* olduğundan katsayı tahminleri,

$$X'W \left(y - X \begin{pmatrix} \hat{a}_0 \\ \hat{a}_1 \\ \hat{a}_2 \\ \hat{a}_3 \end{pmatrix} \right) = 0$$

normal denklemlerinin çözülmesi ile elde edilebilir. Burada yerel 3. dereceden regresyon için X tasarım matrisi aşağıdaki biçimdedir:

$$X = \begin{pmatrix} 1 & x_1 - x & \frac{(x_1 - x)^2}{2} & \frac{(x_1 - x)^3}{6} \\ 1 & x_2 - x & \frac{(x_2 - x)^2}{2} & \frac{(x_2 - x)^3}{6} \\ . & . & . & . \\ . & . & . & . \\ . & . & . & . \\ 1 & x_n - x & \frac{(x_n - x)^2}{2} & \frac{(x_n - x)^3}{6} \end{pmatrix}$$

Örnek: Bant genişliğini $h = 3$ alarak, aşağıda verilen A matrisi yardımı ile 3. dereceden yerel bir polinomun $x = 2$ noktasındaki tahminini elde edelim.

2. REGRESYON MODELLERİ, REGRESYONDA DÜZELTME
KAVRAMI VE PARAMETRİK OLMAYAN REGRESYONDA
DÜZELTME YÖNTEMLERİ

Gülin TABAKAN

$$A = \begin{bmatrix} 0 & 28 \\ 1 & 19 \\ 2 & 36 \\ 3 & 34 \\ 4 & 40 \end{bmatrix}$$

Çözüm için ilk önce 3. dereceden yerel polinomun $\hat{a}_0, \hat{a}_1, \hat{a}_2, \hat{a}_3$ ile gösterilen katsayılarını elde etmemiz gerekmektedir. Bunun için yukarıda tanımlanan X tasarım matrisini, W köşegen matrisini ve $y = (y_1, \dots, y_n)'$ 'yi aşağıdaki biçimde elde ederiz:

$$X = \begin{bmatrix} 1 & 0-2 & \frac{(0-2)^2}{2} & \frac{(0-2)^3}{6} \\ 1 & 1-2 & \frac{(1-2)^2}{2} & \frac{(1-2)^3}{6} \\ 1 & 2-2 & \frac{(2-2)^2}{2} & \frac{(2-2)^3}{6} \\ 1 & 3-2 & \frac{(3-2)^2}{2} & \frac{(3-2)^3}{6} \\ 1 & 4-2 & \frac{(4-2)^2}{2} & \frac{(4-2)^3}{6} \end{bmatrix} = \begin{bmatrix} 1 & -2 & 2 & -1.33 \\ 1 & -1 & 0.5 & -0.167 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 0.5 & 0.167 \\ 1 & 2 & 2 & 1.33 \end{bmatrix}, \quad y = \begin{bmatrix} 28 \\ 19 \\ 36 \\ 34 \\ 40 \end{bmatrix}$$

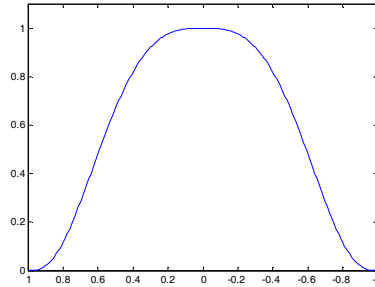
$$W = \begin{bmatrix} K\left(\frac{0-2}{3}\right) & 0 & 0 & 0 & 0 \\ 0 & K\left(\frac{1-2}{3}\right) & 0 & 0 & 0 \\ 0 & 0 & K\left(\frac{2-2}{3}\right) & 0 & 0 \\ 0 & 0 & 0 & K\left(\frac{3-2}{3}\right) & 0 \\ 0 & 0 & 0 & 0 & K\left(\frac{4-2}{3}\right) \end{bmatrix} = \begin{bmatrix} K\left(\frac{-2}{3}\right) & 0 & 0 & 0 & 0 \\ 0 & K\left(\frac{-1}{3}\right) & 0 & 0 & 0 \\ 0 & 0 & K(0) & 0 & 0 \\ 0 & 0 & 0 & K\left(\frac{1}{3}\right) & 0 \\ 0 & 0 & 0 & 0 & K\left(\frac{2}{3}\right) \end{bmatrix}$$

$$= \begin{bmatrix} \left(1 - \left|\frac{-2}{3}\right|\right)^3 & 0 & 0 & 0 & 0 \\ 0 & \left(1 - \left|\frac{-1}{3}\right|\right)^3 & 0 & 0 & 0 \\ 0 & 0 & \left(1 - \left|\frac{0}{3}\right|\right)^3 & 0 & 0 \\ 0 & 0 & 0 & \left(1 - \left|\frac{1}{3}\right|\right)^3 & 0 \\ 0 & 0 & 0 & 0 & \left(1 - \left|\frac{2}{3}\right|\right)^3 \end{bmatrix} = \begin{bmatrix} 0.3484 & 0 & 0 & 0 & 0 \\ 0 & 0.8929 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0.8929 & 0 \\ 0 & 0 & 0 & 0 & 0.3484 \end{bmatrix}$$

Burada ağırlıklardan oluşan W köşegen matrisini oluşturmada Şekil 2.6 ile gösterilen,

$$K(u) = \begin{cases} (1 - |u|^3)^3 & -1 \leq u \leq 1 \\ 0 & u > 1 \text{ veya } u < -1 \end{cases}$$

tricube fonksiyonu kullanılmıştır.



Şekil 2.6 Tricube fonksiyonu

Yukarıda verilen W matrisinden $x = 2$ noktasından daha uzak gözlemlere daha düşük ve $x = 2$ noktasından daha yakın gözlemlere daha yüksek ağırlıklar atandığı görülmektedir. Örneğin $x_1 = 0$ ve $x_5 = 4$ için atanan ağırlık 0.3484 iken $x_2 = 1$ ve $x_4 = 3$ için atanan ağırlık 0.8929'dur. Elde edilen X ve W matrislerinden hareketle,

$$(X^T W X)^{-1} = \begin{bmatrix} 0.5360 & 0.0000 & -0.3790 & -0.0000 \\ 0.0000 & 0.3722 & -0.0000 & 0.4514 \\ -0.3790 & 0.0000 & 0.5772 & 0.0000 \\ 0.0000 & 0.4514 & 0.0000 & 1.3273 \end{bmatrix}, \quad X^T W y = \begin{bmatrix} 107.0149 \\ -5.0319 \\ 71.0443 \\ -7.7972 \end{bmatrix}$$

elde edilip buradan,

$$(X'WX)^{-1}X'Wy = \begin{bmatrix} 30.4315 \\ -5.3924 \\ 0.4523 \\ -12.6202 \end{bmatrix} = \begin{bmatrix} \hat{a}_0 \\ \hat{a}_1 \\ \hat{a}_2 \\ \hat{a}_3 \end{bmatrix}$$

elde edilir. O halde 3. dereceden yerel polinom, (2.18) eşitliğinde katsayıların değerlerinin yerleştirilmesi ile,

$$f(x_i) \approx 30.4315 - 5.3924(x_i - x) + \frac{0.4523}{2}(x_i - x)^2 - \frac{12.6202}{6}(x_i - x)^3$$

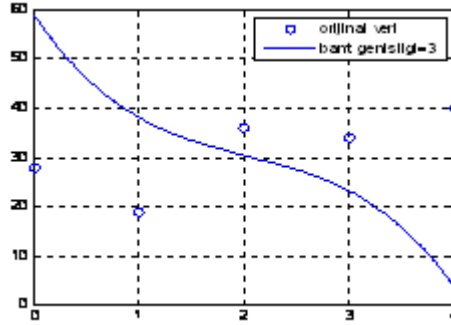
olarak elde edilir. Böylece $x = 2$ için x_3 noktası etrafındaki yerel regresyon değeri, $\hat{f}(x_3 = 2) = 30.4315$ olarak elde edilir. Eğer $x = 2.1$ noktasındaki değeri elde edersek bu sonuç $x = 2$ için bulunan sonuca yakın bir değer olacaktır. $x_i = x$ olduğunda $f(x_i)$, a_0 'ı hesaplamak için minimize edilir.

(2.17) eşitliği yerel regresyon tahmininin $\hat{f}_h(x) = \sum_{i=1}^n w_{xi}(x)y_i$ ile tanımlanmış

olan bir doğrusal tahmin olduğunu ifade eder ve w_{xi} katsayıları,

$$w(x) = \begin{pmatrix} w_{x1} \\ \cdot \\ \cdot \\ \cdot \\ w_{xn} \end{pmatrix} = e'_1(X'WX)^{-1}X'W \quad (2.20)$$

biçiminde açık olarak verilir (Loader, 2004). Burada $e'_1 = (1 \ 0 \dots 0)$ p -sütunlu birim vektördür. Yukarıdaki örnekte $h = 3$ için $e'_1(X'WX)^{-1}X'W$ değerleri -0.0773 0.3094 0.5360 0.3094 -0.0773 olarak elde edilmiştir. Şekil 2.7'de $h = 3$ için yerel regresyon düzeltme gösterilmiştir.



Şekil 2.7. Yerel regresyon düzeltme ($h=3$)

2.7.3. k -En Yakın Komşu Tahmin Edici (k -Nearest Neighbor Estimator: k -NN)

En yakın komşu tahmin edicinin yapısı kernel tahmin ediciden farklıdır. Şimdiye kadar gördüğümüz gibi, kernel regresyon tahmini x noktası civarındaki belirlenmiş bir komşulukta bağımlı değişkenlerin ağırlıklı ortalamalarının hesaplanması yöntemi olarak düşünülebilir. Bu komşuluğun genişliği düzeltme parametresi h ve K kernel fonksiyonu tarafından yönetilir.

$$y_i = f(x_i) + \varepsilon_i, \quad i = 1, \dots, n$$

modelinde $(x_1, y_1), \dots, (x_n, y_n)$ gözlem değerlerini göz önüne alalım. Burada $\{x_i\}_{i=1}^n$, x değişkeninin $[a, b]$ 'de aldığı farklı değerleri gösterir. Ayrıca x_i noktasındaki f 'in tahmin edicisi x_i 'ye en yakın k tane komşuya ait olan y_i değerlerinin ortalamasıdır. k -en yakın komşu tahmini değişen komşuluklarda ağırlıklı bir ortalama değildir. Bu komşuluk öklid (Euclidean) uzaklığındaki x 'in k . en yakın komşular arasında bulunan x değişkenleri aracılığı ile tanımlanır. Biçimsel olarak, k -en yakın komşu tahmini,

$$\hat{f}_k(x) = \sum_{i=1}^n w_{ki}(x) y_i \quad (2.21)$$

olarak yazılabilir. Burada $\{w_{ki}(x)\}_{i=1}^n$ ağırlık dizisi,

$$J_x = \{i: x_i, x' \text{ e en yakın } k \text{ tane gözlemden biri} \}$$

indeks dizisi ile,

$$w_{ki}(x) = \begin{cases} 1/k, & i \in J_x \text{ ise} \\ 0, & \text{aksi durumda ise} \end{cases}$$

olarak tanımlanır. Burada k bu tahmin edicinin düzeltme parametresidir ve tahmin edilen eğrinin pürüzlülüğünün derecesini düzenler ve kernel düzelticilerin düzeltme parametresine (bant genişliğine) benzer bir rol oynar. Ağırlıkların oluşturulması hakkında fikir sahibi olabilmek için aşağıdaki örnekleri ele alalım.

$$\{(x_i, y_i)\}_{i=1}^5 \text{ gözlemleri } \{(1,5), (7,12), (3,1), (2,0), (5,4)\} \text{ olsun. } x = 4 \text{ ve}$$

$k = 3$ için $\hat{f}_k(x)$ 'in k -NN tahminini hesaplayalım.

$$x' \text{ e en yakın } k \text{ gözlem son üç veri noktasıdır, böylece } J_x = J_4 = \{3, 4, 5\}$$

$$\text{ve bundan dolayı } w_{k1}(4) = 0, w_{k2}(4) = 0, w_{k3}(4) = 1/3, w_{k4}(4) = 1/3, w_{k5}(4) = 1/3$$

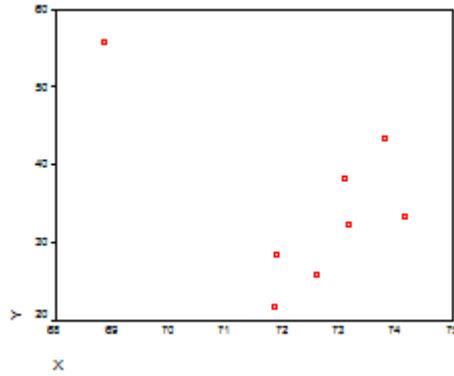
olup $\hat{f}_k(x)$ 'in k -NN tahmini $k = 3$ ve $x = 4$ için (2.21) ile verilen eşitlikten,

$$\hat{f}_3(x = 4) = (1 + 0 + 4)/3 = 5/3$$

olarak elde edilir (Härdle, 1994).

Kardiyovasküler hastalıklar için risk faktörleri ile ilgili olarak yapılan klinik bir çalışmada, yaşları 18 ve 42 arasında değişen 109 hastaya ait gözlem değerleri verilmiş olsun. Bu gözlem değerlerinden 8 gözlem değerini (x_i, y_i) göz önüne alalım. Burada x_i bağımsız değişkeni hastanın bel çevresi, y_i bağımlı değişkeni ise hastanın iç karın yağ dokusudur. Şekil 2.8'de 8 gözlem değerine ait saçılım grafiği verilmiştir.

- | | |
|-------------------|-------------------|
| 1. (68.85, 55.78) | 5. (73.10, 38.21) |
| 2. (71.85, 21.68) | 6. (73.20, 32.22) |
| 3. (71.90, 28.32) | 7. (73.80, 43.35) |
| 4. (72.60, 25.89) | 8. (74.15, 33.41) |



Şekil 2.8. Gözlem değerlerinin saçılım grafiği

$k = 1$ ile $x = 73.00$ 'ün komşuluğu, $J(x = 73.00) = \{5\}$ olup,

$$\hat{f}_2(x = 73.00) = \frac{38.21}{1} = 38.21$$

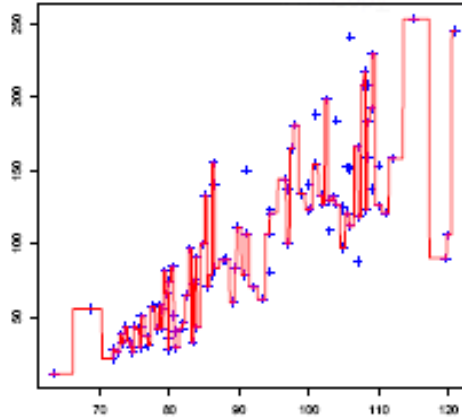
elde edilir. $k = 2$ ile $x = 73.00$ 'ün komşuluğu, $J(x = 73.00) = \{5, 6\}$ olup,

$$\hat{f}_2(x = 73.00) = \frac{38.21 + 32.22}{2} = 35.215$$

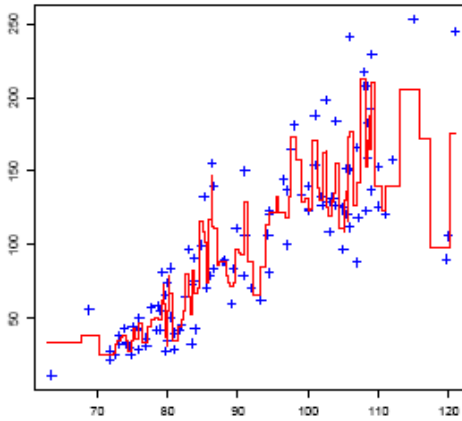
elde edilir. $k = 5$ ile $x = 73.00$ 'ün komşuluğu, $J(x = 73.00) = \{3, 4, 5, 6, 7\}$ olup

$$\hat{f}_5(x = 73.00) = 33.598$$

elde edilir. Şekil 2.9 ve 2.10'da sırasıyla $k = 1$ ve $k = 2$ için k -en yakın komşu fonksiyonunun uyum eğrileri verilerden elde edilmiştir.

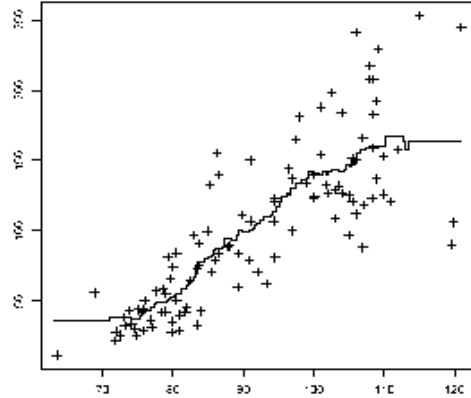


Şekil 2.9. $k = 1$ için k -NN tahmin edici

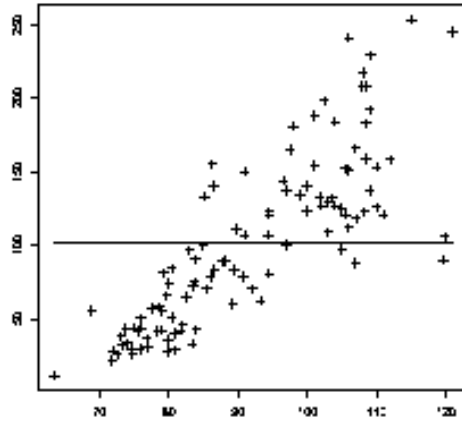


Şekil 2.10. $k = 2$ için k -NN tahmin edici

Benzer şekilde $k = 20$ ve $k = 109$ için verilerden elde edilen k -en yakın komşu fonksiyonunun uyum eğrileri Şekil 2.11 ve Şekil 2.12'deki gibidir.



Şekil 2.11. $k = 20$ için k -NN tahmin edici



Şekil 2.12. $k = 109$ için k -NN tahmin edici

Yukarıdaki şekillerden komşulukların sayısı k küçük ise sonuç fonksiyonunun oldukça değişken, büyük ise sonuç fonksiyonunun daha düzgün olduğu görülmektedir. Belirlenmiş bir k sabiti için genel olarak; Eğer veriler yoğun bir şekilde dağılmış ise komşuluklar arasındaki uzaklıklar küçük, tam tersi olarak veriler dağınık bir şekilde dağılmış ise komşuluklar arasındaki uzaklıklar büyük olur.

2.7.4. Eğrisel Çizgi Düzeltme (Spline Smoothing) Yöntemi

Parametrik olmayan regresyonda f fonksiyonunun biçimsel olarak değerlendirilmesinde kullanılan eğrisel çizgi düzeltme yönteminde f fonksiyonunun $\hat{f}_\lambda$ ile gösterilen tahmin edicisi uyum iyiliği için bir ölçüm olan RSS 'nin göz önüne

alınması ile elde edilir. Parametrik olmayan regresyon modeli kullanılarak hata kareler toplamı,

$$RSS = \sum_{i=1}^n \{y_i - f(x_i)\}^2 \quad (2.22)$$

ile hesaplanır. Pürüzlülük cezası olarak bilinen,

$$\int \{f''(x)\}^2 dx$$

fonksiyonu ile eğrisel çizgi düzeltme tahmin edicisi (2.6) ile verilen,

$$S_\lambda(f) = \sum_{i=1}^n \{y_i - f(x_i)\}^2 + \lambda \int \{f''(x)\}^2 dx \quad (2.23)$$

cezalı kareler toplamının minimum yapılması ile elde edilir (Härdle ve ark., 2004). Eşitlik (2.23)'de ilk terim hata kareler toplamını (RSS) ifade etmektedir ve parametrik regresyonda olduğu gibi uyumun verilere yakınlığını ölçer, ikinci terim ise pürüzlülük cezasını (roughness penalty) gösterip pürüzlüğe bir ceza yükler yani fonksiyondaki eğriliği cezalandırır. İkinci terimde yer alan λ alt bölüm 2.5.2'de tanımlanan düzeltme parametresi olup f eğrisinin pürüzlülüğü ve rezidü hata arasındaki değişimin oranını gösterir. Cezalı kareler toplamını minimum yapan f eğrisine '*eğrisel çizgi düzeltme tahmin edicisi*' denir ve $x_1, x_2, \dots, x_n$ düğüm noktaları ile bir doğal kübik eğrisel çizgi olarak bilinir. $\hat{f}$ 'nın doğal kübik eğrisel çizgi özelliklerinin pürüzlülük cezasının özel seçiminden kaynaklanan özellikler olduğuna dikkat edilmelidir. Burada bahsedilen eğrisel çizgi düzeltme tahmin edicisi,

$$\hat{f}_\lambda = (I + \lambda K)^{-1} y = S_\lambda y$$

biçiminde gösterilir. Buradan eğrisel çizgi düzelticinin y_i 'de bir doğrusal düzeltici olduğunu görebiliriz. Yani $\hat{f}_\lambda(x) = \sum_{i=1}^n w_{\lambda i}(x) y_i$ olacak şekilde $w_{\lambda i}(x)$ ağırlıkları vardır (Härdle ve ark., 2004). Burada yer alan S_λ matrisine $x = (x_1, x_2, \dots, x_n)$ düğüm noktaları vektörü ile belirlenen bir '*düzeltilme matrisi*' denir ve doğrusal regresyondaki şapka matrisinin rolünü oynar.

2.7.4.1. Eğrisel Çizgi Düzeltme Tahmin Edicisinin Elde Edilmesi

(2.6) ile verilen cezalı kareler toplamının minimum yapılması ile elde edilen eğrisel çizgi düzeltme tahmin edicisi $\hat{f}$ 'nın doğal kübik eğrisel çizgi olarak bilinmesi çok önemlidir. Çünkü bu durumda sonsuz boyutlu $C^2[a, b]$ düzgün fonksiyonlar kümesinde minimum problemi, (x_i, y_i) gözlem noktaları ve λ düzeltme parametresine bağlı olarak hesaplanan doğal kübik eğrisel çizgiler dikkate alınarak $S(f)$ 'nin minimum yapılması problemi ile yer değiştirmiş olur. Bu durumda $\hat{f}$ tahmin fonksiyonu eğrisel çizgi düzeltme tahmin edicisi olarak belirtilebilir. Doğrusal eşitliklerin bir kümesinin çözülmesi ile eğrisel çizgi eğrisinin minimum yapılması aşağıdaki gibidir:

f 'nin alt bölüm 2.6.2'deki gibi f ve y vektörleri, Q ve R matrisleri ile doğal kübik eğrisel çizgi olduğu varsayılın. Bu vektör ve matrislere göre (2.6) cezalı kareler toplamı yeniden yazılabilir.

$y = (y_1, \dots, y_n)'$ vektörü verilsin. x_i düğüm noktalarında $f(x_i)$ değerler vektörü $f = (f_1, \dots, f_n)' = (f(x_1), \dots, f(x_n))'$ olup (2.22) ile verilen rezidü kareler toplamını,

$$RSS = \sum_{i=1}^n \{y_i - f(x_i)\}^2 = (y - f)'(y - f)$$

biçiminde yazarız. Eşitlik (2.9)'dan dolayı $\int f''^2$ pürüzlülük ceza teriminin $f'Kf$ olarak ifade edilebileceğini biliyoruz. Böylece (2.6) ile verilen cezalı kareler toplamı

$$\begin{aligned} S(f) &= (y - f)'(y - f) + \lambda f'Kf \\ &= f'(I + \lambda K)f - 2y'f + y'y \end{aligned} \quad (2.24)$$

biçiminde yeniden yazılır. λK negatif tanımlı olmayan (non-negative) bir matris olduğundan $(I + \lambda K)$ pozitif tanımlı bir matristir. Bu yüzden (2.24) eşitliğinin f 'e göre türevinin alınması ile,

$$f = (I + \lambda K)^{-1} y \quad (2.25)$$

biçiminde bir tek minimum değeri elde edilir (Green ve Silverman, 1994).

2.8. Doğrusal Düzelticilerin İstatistiksel Özellikleri

Şimdiye kadar ele alınan düzeltme yöntemlerinin her biri, yapılan düzeltmenin miktarını kontrol eden düzeltme parametrelerine sahipti. Buna örnek olarak kernel düzeltici veya yerel regresyon yöntemlerindeki h düzeltme parametresi ve cezalı en küçük kareler ölçütünde λ parametresi verilebilir. Düzeltme parametresi seçimi ve diğer istatistiksel çıkarım yöntemlerini inceleyebilmek için başlangıçta ilk olarak doğrusal düzelticilerin özelliklerinin bilinmesi gerekir. Bu kısımda doğrusal düzelticilerin özelliklerinden bazılarını ele alacağız.

2.8.1. Hata Kareler Ortalaması (MSE)

Hata kareler ortalaması (MSE), regresyon modellerinin üstünlüğünün karşılaştırmasında çok sık kullanılan ölçütlerden biridir. MSE'yi hesaplamak için

$y_i = f(x_i) + \varepsilon_i$, $\varepsilon \sim N(0, \sigma^2)$, $i = 1, \dots, n$ ile verilen parametrik olmayan regresyon modelini tekrar ele alalım. Bu model için MSE ,

$$\begin{aligned} MSE(\hat{f}(x)) &= E\left[\left(\hat{f}(x) - f(x)\right)^2\right] = E\left[\left(\hat{f}(x) - F(x) + F(x) - f(x)\right)^2\right] \\ &= E\left[\left(\hat{f}(x) - F(x)\right)^2\right] + E\left[\left(F(x) - f(x)\right)^2\right] + 2E\left[\left(\hat{f}(x) - F(x)\right)\left(F(x) - f(x)\right)\right] \\ &= Var(\hat{f}(x)) + \left(Bias\hat{f}(x)\right)^2 + 2E\left[\left(\hat{f}(x)F(x) - \hat{f}(x)f(x) - (F(x))^2 + F(x)f(x)\right)\right] \\ &= Var(\hat{f}(x)) + \left(Bias\hat{f}(x)\right)^2 + 2\left[F(x)^2 - F(x)f(x) - (F(x))^2 + F(x)f(x)\right] \\ &= Var(\hat{f}(x)) + \left(Bias\hat{f}(x)\right)^2 \end{aligned}$$

biçiminde yazılır ve burada $F(x) = E[\hat{f}(x)]$ 'dir. $\hat{f}_\lambda = S_\lambda y$ ile gösterilen bir doğrusal düzelticiyi göz önüne alalım, $Bias(\hat{f}(x)) = E(S_\lambda y) - f = S_\lambda f - f = (I - S_\lambda)f$ yanlışlık terimi ile hata kareler ortalaması (MSE),

$$\begin{aligned} MSE(\hat{f}(x)) &= Var(\hat{f}(x)) + \left(Bias\hat{f}(x)\right)^2 \\ &= \sigma^2 S_\lambda' S_\lambda + \|(I - S_\lambda)f\|^2 \end{aligned} \tag{2.26}$$

ile ifade edilir.

MSE ölçümünde birinci terim varyansı ölçerken ikinci terim yanlışlığın karesini ölçmektedir (Hastie ve Tibshirani, 1990). Burada düzeltme parametresinin değerindeki artmalarda $\hat{f}(x)$ tahminini oluşturmak için daha fazla veri kullanıldığı önsezi ile söylenebilir ve bu durumda $var(\hat{f}(x))$ azalır. Diğer taraftan yerel polinom tahmini küçük aralıklar üzerinde en iyidir. Bundan dolayı λ değeri artar iken yanlışlık artışı beklenebilir. Sezgisel olarak büyük komşuluklar düşük varyanslı bir tahmin fakat potansiyel olarak yüksek yanlışlık üretecektir ve tam tersi küçük komşuluklar için de söylenebilir. Böylece seçilen düzeltme parametresi (bant genişliği) λ ile yönetilen yanlışlık ve varyans arasında bir değiş-tokuş (trade-off) olduğunu söyleriz.

2.8.2. Serbestlik Derecesi

Bir S_λ doğrusal düzeltici operatörü verildiğinde, eğrisel çizgi düzeltme konusunda yaygın olarak kullanılan en basit serbestlik derecesi tanımı $df = tr(S_\lambda)$ biçimindedir.

Bir doğrusal düzeltici için, serbestlik derecesinin seçiminde $tr(S_\lambda)$ 'dan başka iki yaygın tanım daha vardır. Bunlar $tr(S_\lambda)$ yerine $n - tr(2S_\lambda - S_\lambda S_\lambda')$ ve $tr(S_\lambda S_\lambda')$ ölçümlerinin kullanılmasıdır. Bu tanımlar, $tr(S_\lambda)$ 'da olduğu gibi, doğrusal regresyon modeline benzer biçimde harekete geçirilebilir ve farklı amaçlar için kullanışlıdır. Ayrıca, bu iki tanım doğrusal olmayan düzelticilere genişletilebilir (Hastie ve Tibshirani, 1990).

(2.2) ile verilen parametrik regresyon modelinde,

$$E(RSS) = (n - p)\sigma^2$$

ile ifade edilir. Bu modelde H şapka matrisinin izi modelin serbestlik derecesi p 'ye eşit olup, $(n - p)$ hataların serbestlik derecesidir. (2.3) ile verilen parametrik olmayan regresyon modelinde hata serbestlik derecesini (residual degrees of freedom),

$$df_{res}(\lambda) = n - tr(2S_\lambda - S_\lambda' S_\lambda) \quad (2.27)$$

olarak tanımlarız. Parametrik tahminde olduğu gibi, hata serbestlik derecesi σ^2 'nin tahmininde kullanılır. (2.27) eşitliğini göstermek için,

$$y = f + \varepsilon, \quad \text{cov}(\varepsilon) = \sigma^2 I$$

modelinin $\hat{f}_\lambda = S_\lambda y$ ile gösterilen eğrisel çizgi düzeltme tahmin edicisini ele alalım. Bu tahmin edici için hata kareler toplamı (RSS),

$$\begin{aligned}
 RSS &= \sum_{i=1}^n (y_i - \hat{f}(x_i))^2 = (y - \hat{f})'(y - \hat{f}) \\
 &= (y - S_\lambda y)'(y - S_\lambda y) \\
 &= y'(I - S_\lambda)'(I - S_\lambda)y \\
 &= y'(I - S_\lambda' - S_\lambda + S_\lambda' S_\lambda)y
 \end{aligned} \tag{2.28}$$

biçiminde elde edilir. Bir genel v rasgele vektörü için,

$$E(v'Av) = E(v)'AE(v) + tr\{ACov(v)\} \tag{2.29}$$

ile verilen beklenen değer tanımını kullanarak,

$$\begin{aligned}
 E(RSS) &= E\|\hat{f} - y\|^2 \\
 &= E\|(S_\lambda - I)y\|^2 \\
 &= E\{y'(S_\lambda - I)'(S_\lambda - I)y\} \\
 &= f'(S_\lambda - I)'(S_\lambda - I)f + \sigma^2 tr\{(S_\lambda - I)'(S_\lambda - I)\} \\
 &= \|(S_\lambda - I)f\|^2 + \sigma^2 \{tr(S_\lambda' S_\lambda) - 2tr(S_\lambda) + n\} \\
 &= \|(S_\lambda - I)f\|^2 + \sigma^2 df_{res}
 \end{aligned} \tag{2.30}$$

olarak elde edilir. Buradan hareketle RSS için beklenen değer, yanlışlık teriminin ihmal edilmesi ile,

$$E(RSS) = \sigma^2 df_{res} \tag{2.31}$$

olarak bulunur. Burada S_λ $n \times n$ tipinde kare matris olduğundan $tr(S_\lambda) = tr(S_\lambda')$ dir.

Böylece RSS/df_{res} , hata varyansı σ^2 'nin yansız bir tahmin edicisidir (Ruppert ve ark. 2003).

(2.3) ile verilen parametrik olmayan regresyon modeli için parametrik regresyona benzer biçimde, hata için eş değer serbestlik derecesi EDF ,

$$EDF = tr(I - S_\lambda) \quad (2.32)$$

tanımlanmıştır. (2.32) eşitliğindeki S_λ , λ düzeltme parametresi ile eğrisel çizgi düzeltme ile bağlantılı olan parametrik regresyondaki şapka matrisine karşılık gelen bir düzeltici matristir ve $df_{fit} = tr(S_\lambda)$ modelin serbestlik derecesidir. Hastie ve Tibshirani (1990) eş değer serbestlik derecesi kavramını daha ayrıntılı olarak incelemişler ve modelleri karşılaştırma konusunda EDF yerine (2.27) eşitliğinde verilen $EDF^* = tr\{(I - S'_\lambda)(I - S_\lambda)\} = n - tr(2S_\lambda - S'_\lambda S_\lambda)$ ifadesinin kullanılması gerektiğini savunmuşlardır. Eğer S_λ bir izdüşüm matrisi ise, klasik doğrusal regresyonda olduğu gibi (2.27) ve (2.32) serbestlik dereceleri birbirleriyle uyusurlar (Green ve Silverman, 1994).

Bu alt bölümde ele aldığımız serbestlik derecesi kavramı, ilerleyen bölümlerde düzeltme parametresinin değerinin belirlenmesinde kullanılacak olan varyansın tahmininde, yarı parametrik regresyon modelinde parametrik ve parametrik olmayan bileşenler hakkında çıkarımlar yapılmasında kullanılacaktır.

2.9. Hata Varyansının Tahmini

Bir doğrusal düzelticinin varyansını elde etmek için (2.3) modelindeki ε_i rasgele hataları hakkında varsayımlar yapmamız gerekir. En yaygın varsayım $var(\varepsilon_i) = \sigma^2$ varyansı ile hataların bağımsız ve özdeş olarak dağılmasıdır.

Bu çalışmada parametrik olmayan regresyon konusunda birçok istatistiksel yöntemler için ihtiyaç duyulan hata varyansı σ^2 'nin tahmini için, istatistiksel çalışmalarda yaygın olarak kullanılan iki yöntem ele alınmıştır. Bu yöntemlerden ilki parametrik regresyon konusunda olduğu gibi, hata kareler toplamının serbestlik

derecesine oranı ile elde edilir. σ^2 'nin tahmini için yaygın olarak kullanılan diğer bir yöntem ise fark almaya dayalı varyans tahminidir.

2.9.1. Hata Kareler Yaklaşımı

Parametrik regresyonda varyansın yansız bir tahmin edicisi hata kareler toplamının hata serbestlik derecesine bölünmesi ile elde edilir. Benzer olarak eğrisel çizgi düzeltme konusunda, hata kareler toplamı eşdeğer serbestlik derecesi olarak tanımlanan $tr(I - S_\lambda)$ ifadesine bölünerek varyansın bir tahmin edicisi elde edilir. λ düzeltme parametresi ile bu tahmin edici aşağıdaki biçimde verilir:

$$\begin{aligned}\hat{\sigma}_\lambda^2 &= \frac{\sum_i \{y_i - \hat{f}(x_i)\}^2}{tr(I - S_\lambda)} = \frac{\|(S_\lambda - I)y\|^2}{tr(I - S_\lambda)} \\ &= \frac{RSS}{n - tr(S_\lambda)}\end{aligned}\quad (2.33)$$

Burada $\hat{f}$, λ düzeltme parametresi ile hesaplanan eğrisel çizgi düzeltme tahmin edicisidir. Gerçek f regresyon eğrisinin bir düz doğru olduğu özel durumda, tüm λ düzeltme parametreleri için $\hat{\sigma}_\lambda^2$ tahmin edicisi, σ^2 'nin yansız bir tahmin edicisidir (Green ve Silverman, 1994).

2.9.2. Yerel Fark Alma Yaklaşımı

Model (2.3)'de $f(\cdot)$ fonksiyonunun türevlenebilir ve x ordinatlarının yeterince yakın olması durumunda fark almanın verilere uygulanması ile, f ortalama fonksiyonunun etkisinin modelden çıkarılması yaklaşık olarak mümkündür.

Rice (1984), σ^2 'nin tahmini için (2.33) eşitliğindeki varyans tahmin edicisinden farklı olarak fark almaya dayalı bir tahmin edici önermiştir. Bu tahmin edici ardışık üç noktaya en küçük kareler doğrusunu uydurmak ile oluşturulur. Yerel olarak uydurulan bu doğruların her biri ile ilgili hata kareler toplamı σ^2 'nin bir

tahminini verir ve bu tahminlerin ortalaması varyansın tümünün tahminini verir. Bu tahmin edici tamamen Gasser ve ark. (1986) tarafından önerilen tahmine eşdeğerdir ve aşağıdaki gibi verilir (Green ve Silverman, 1994):

$$\hat{\sigma}_{GSJS}^2 = \frac{1}{n-2} \sum_{i=2}^{n-1} \tilde{\varepsilon}_i^2 \quad (2.34)$$

Burada $\tilde{\varepsilon}_i$,

$$\tilde{\varepsilon}_i = (y_i - a_i y_{i-1} - b_i y_{i+1}) / (1 + a_i^2 + b_i^2)^{1/2}, i = 2, \dots, n-1$$

olup, (x_{i-1}, y_{i-1}) ve (x_{i+1}, y_{i+1}) noktalarını birleştiren doğru ile (x_i, y_i) noktası arasındaki uzaklıktır ve burada a_i ve b_i aşağıdaki gibidir:

$$a_i = (x_{i+1} - x_i) / (x_{i+1} - x_{i-1})$$

$$b_i = (x_i - x_{i-1}) / (x_{i+1} - x_{i-1})$$

Gasser ve ark. (1986) tarafından önerilen varyans tahmin edicisinin yarı parametrik model konusuna uyarlaması şöyledir: A matrisi girişlerin tümü sıfır fakat i . girişi $a_i c_i$, $(i+1)$. girişi $-c_i$ ve $(i+2)$. girişi $b_i c_i$ olan $(n-2) \times n$ boyutlu bir matris olsun. Burada a_i ve b_i yukarıdaki gibi olup,

$$c_i = (a_i^2 + b_i^2 + 1)^{-1/2}, i = 1, \dots, n-1$$

biçimindedir. Böylece $\hat{\sigma}_{GSJS}^2$,

$$\hat{\sigma}^2 = \frac{y' A' A y}{tr A' A} \quad (2.35)$$

olarak yazılabilir ve $tr A'A = n - 2$ 'dir. Model (2.5)'ten $Ay = AZ\beta + A\varepsilon$ yazabiliriz. Burada $\beta = 0$ ise, eşitlik (2.35)'deki varyans tahmin edicisi σ^2 için etkili olacaktır, $\beta \neq 0$ ise (2.35) eşitliği parametrik regresyon modelleri göz önüne alınarak aşağıdaki gibi elde edilir:

$$\begin{aligned}
 \hat{\sigma}^2 &= \frac{(Ay - AZ\beta)'(Ay - AZ\beta)}{tr(A'(I - P)A)} \\
 &= \frac{y'A'Ay - y'A'AZ\beta}{tr(A'(I - P)A)} \\
 &= \frac{y'A'Ay - y'A'AZ \left[(AZ)'(AZ) \right]^{-1} (AZ)'(Ay)}{tr(A'(I - P)A)} \\
 &= \frac{y'A' \left\{ I - AZ \left[(AZ)'(AZ) \right]^{-1} (AZ)' \right\} Ay}{tr(A'(I - P)A)} \\
 &= \frac{y'A'(I - P)Ay}{tr(A'(I - P)A)}
 \end{aligned} \tag{2.36}$$

Burada $P = AZ(Z'A'AZ)^{-1}Z'A'$ biçimindedir (Eubank ve ark. 1998, Schimek, 2000). Farka dayalı varyans tahmin edicisi, düzeltme parametresinin herhangi bir seçimini gerektirmeyen avantaja sahiptir (Green ve Silverman, 1994).

3.YARI PARAMETRİK MODELDE EĞRİSEL ÇİZGİ DÜZELTME
YÖNTEMİNE DAYALI PARAMETRE TAHMİNLERİ VE
PARAMETRELER İLE İLGİLİ SONUÇLAR

Gülin TABAKAN

**3. YARI PARAMETRİK MODELDE EĞRİSEL ÇİZGİ DÜZELTME
YÖNTEMİNE DAYALI PARAMETRE TAHMİNLERİ VE
PARAMETRELER İLE İLGİLİ SONUÇLAR**

Yarı parametrik regresyon modelleri bağımlı değişkenin bazı açıklayıcı değişkenlerle ilişkisinin parametreleştirilebildiği fakat diğer bazı açıklayıcı değişken veya değişkenlerle ilişkisinin kolayca parametreleştirilemediği modellerdir. Yarı parametrik modelde doğrusal parametrik bileşenler modelin parametrik kısmını oluştururken, hem parametrik hemde doğrusal olmayan bileşenler ise modelin parametrik olmayan kısmını oluşturur. Bu model her bir değişkenin etkisinin daha kolay yorumlanmasına olanak sağlayan ve standart regresyon yöntemlerini genelleştiren toplamsal regresyon modellerin özel bir durumu şeklindedir (Härdle, Liang, Gao, 2000). Ayrıca yarı parametrik modeller bağımlı değişkenin bazı açıklayıcı değişkenlerle doğrusal fakat diğer özel bağımsız (açıklayıcı) değişkenlerle doğrusal olmayan ilişki içerisinde olduğu modellerdir. Yarı parametrik regresyon modeli,

$$y_i = z_i' \beta + f(x_i) + \varepsilon_i, \quad 1 \leq i \leq n \quad (3.1)$$

biçiminde ifade edilir. Burada,

y_i : y bağımlı değişkeninin i . gözlem değeri;

z_i : Parametrik kısma karşı gelen bağımsız değişkenlerin p boyutlu i . gözlemler vektörü;

β : p boyutlu regresyon katsayıları vektörü;

x_i : Parametrik olmayan açıklayıcı değişkenin i . gözlem değeri;

$f \in C^2[a,b]$: Modelin parametrik olmayan kısmına karşılık gelen ve $[a,b]$ aralığında ikinci mertebeden sürekli türeve sahip olan bilinmeyen bir düzgün fonksiyon;

3.YARI PARAMETRİK MODELDE EĞRİSEL ÇİZGİ DÜZELTME YÖNTEMİNE DAYALI PARAMETRE TAHMİNLERİ VE PARAMETRELER İLE İLGİLİ SONUÇLAR

Gülin TABAKAN

$\varepsilon_i : i = 1, 2, \dots, n$, bağımsız sıfır ortalamalı ve σ^2 ortak varyanslı rasgele hata terimleridir.

(3.1) modeli matris-vektör gösterimi ile,

$$y = Z\beta + f + \varepsilon \quad (3.2)$$

biçiminde de ifade edilebilir. Burada değişkenler aşağıdaki şekilde tanımlanır:

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \quad Z = \begin{bmatrix} z_{11} & \cdot & \cdot & \cdot & z_{1p} \\ z_{21} & & & & z_{2p} \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ z_{n1} & \cdot & \cdot & \cdot & z_{np} \end{bmatrix}, \quad \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \cdot \\ \cdot \\ \beta_p \end{bmatrix}, \quad f = \begin{bmatrix} f(x_1) \\ f(x_2) \\ \cdot \\ \cdot \\ f(x_n) \end{bmatrix} \quad \text{ve} \quad \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \cdot \\ \cdot \\ \varepsilon_n \end{bmatrix}$$

$y = (y_1, \dots, y_n)'$ gözlemlerin bir $(n \times 1)$ boyutlu vektörü, $Z' = [z_1, \dots, z_n]$ parametrik kısma karşı gelen bağımsız değişkenlerin $(n \times p)$ boyutlu gözlem matrisi, β regresyon katsayılarının $(p \times 1)$ boyutlu vektörü, $f = (f_1, \dots, f_n)' = (f(x_1), \dots, f(x_n))'$ eğrisel çizgi düzeltme tahmin edicisine karşı gelen $(n \times 1)$ boyutlu vektörü ve $\varepsilon = (\varepsilon_1, \dots, \varepsilon_n)'$ normal dağılan rasgele hataların $(n \times 1)$ boyutlu vektörüdür. (3.1) yarı parametrik regresyon modeli için farklı düzeltme yöntemlerine dayalı bir çok yaklaşım önerilmiştir: Green vd. (1985), Engle vd. (1986), Wahba (1990), Green ve Silverman (1994), Eubank vd. (1998) ve Schimek (2000) tarafından yapılan çalışmalarda (3.1) modeline eğrisel çizgi düzeltme yöntemi uygulanmıştır. Robinson (1988) ve Speckman (1988) kernel düzeltmesini, Chen (1988) ise bilinmeyen f fonksiyonunun bir parçalı polinom tahmin edicisine dayalı olarak parametrik bileşenin tahmini için bir yaklaşım önermişlerdir.

Bu bölümde amacımız (3.2) modelindeki f fonksiyonunu, β parametre vektörünü ve $\mu = Z\beta + f$ ortalama vektörünü etkili olarak tahmin etmektir. Bu

3.YARI PARAMETRİK MODELDE EĞRİSEL ÇİZGİ DÜZELTME YÖNTEMİNE DAYALI PARAMETRE TAHMİNLERİ VE PARAMETRELER İLE İLGİLİ SONUÇLAR

Gülin TABAKAN

nedenle çalışmada, (3.2) yarı parametrik regresyon modelinde parametrik ve parametrik olmayan bileşenlerin tahmini için, pürüzlülük ceza yaklaşımı ve Speckman (1988) tarafından önerilen Speckman yaklaşımı ele alınmıştır.

3.1. Yarı Parametrik Modeller için Cezalı En Küçük Kareler Yaklaşımı

$\{x_i, y_i, z_i\}_{i=1}^n$ gözlem değerleri eşitlik (3.1) deki modele uygulandığında β parametre vektörü ve f fonksiyonunun tahmini,

$$\sum_{i=1}^n \{y_i - z_i' \beta - f(x_i)\}^2 \quad (3.3)$$

olarak verilen hata kareler toplamının minimum yapılması ile elde edilebilir. Bununla birlikte, f üzerindeki kısıtlama azlığı bu yaklaşımı başarısız yapacaktır. Burada bir an için x_i 'lerin farklı olduğunu düşünelim. β 'nın herhangi bir değeri için $f(x_i) = y_i - z_i' \beta$ interpolasyonu yolu ile elde edilebilir ancak β bilinmiyorsa bu yaklaşımla belirlenemez. Eşitlik (3.1) de β parametre vektörünü ve f fonksiyonunu tahmin etmek için istatistiksel çalışmalarda yaygın olarak kullanılan birkaç yöntem vardır. Bu yöntemler içinde en çok kullanılanı eğrisel çizgi düzeltme yöntemidir. Giriş kısmında ve alt bölüm 2.7.4'te belirtildiği gibi söz konusu yöntemin temelini cezalı en küçük kareler toplamının minimum yapılması problemi oluşturur. Eğrisel çizgi düzeltme yöntemine göre (3.1) modelinin çözümü,

$$S(\beta, f) = \sum_{i=1}^n \{y_i - z_i' \beta - f(x_i)\}^2 + \lambda \int_a^b \{f''(x)\}^2 dx \quad (3.4)$$

ile verilen cezalı en küçük kareler toplamını minimum yapan f fonksiyonu ve β parametre vektörü olarak tanımlanır. Bu minimumu veren f fonksiyonu $x_1, \dots, x_n$ düğümlü bir doğal kübik eğrisel çizgidir (Green ve Silverman 1994, Wahba, 1990).

3.YARI PARAMETRİK MODELDE EĞRİSEL ÇİZGİ DÜZELTME YÖNTEMİNE DAYALI PARAMETRE TAHMİNLERİ VE PARAMETRELER İLE İLGİLİ SONUÇLAR

Gülin TABAKAN

Eşitlik (3.4)'de birinci terim hata kareler toplamını (RSS) gösterir ve bu ifade uyumun verilere yakınlığını ölçer. İkinci terim ise parametrik olmayan regresyonda olduğu gibi pürüzlülük cezasını gösterir. Ayrıca λ bir skaler olup $\{x_i, y_i, z_i\}_{i=1}^n$ gözlem değerlerine dayalı olarak belirlenir ve bilinmeyen regresyon fonksiyonu ile sabit bir düzeltme parametresidir. Parametrik olmayan regresyon modelinde olduğu gibi yarı parametrik regresyon analizinde de istenen çözüm ne doğrusal bir çözüm ne de bir interpolasyondur. Bu nedenle λ 'nın uygun bir değerinin seçilmesi gerekir. Bu da ancak 4. Bölümde ele alacağımız farklı seçim yöntemleri kullanılarak yapılabilir. Eşitlik (3.4)'de verilen cezalı en küçük kareleri minimum yapan $f(\cdot)$ fonksiyonu gerçekte $[x_i, x_{i+1}]$ aralığında bir parçalı kübik polinom ise o zaman $\int \{f''(x)\}^2 dx$ pürüzlülük ceza terimi, $K = QR^{-1}Q'$ ile ayrıştırılabilen bir K matrisi ile bir karesel form olarak yani (2.9) eşitliği ile $\int \{f''(x)\}^2 dx = f'Kf$ biçiminde ifade edilebilir. Burada Q ve R alt bölüm 2.6.2'de tanımlanan matrislerdir. Parametre tahmini için söz konusu düğüm noktalarının $x_1 < \dots < x_n$ koşulunu sağlaması, diğer bir ifade ile farklı ve sıralı olması gerekir. x_i 'ler farklı ve sıralı değilse, belirli bir ' N tekrarlanma matrisi (incidence matrix)' yardımı ile farklı ve sıralı hale getirilir. $x_1, \dots, x_n$ düğüm noktalarının farklı ve sıralı değerleri $s_1, \dots, s_q$ ile gösterilsin. Bu durumda N tekrarlanma matrisinin elemanları $x_i = s_j$ ise $N_{ij} = 1$ aksi durumda $N_{ij} = 0$ biçimindedir. Böylece N matrisi yardımı ile (3.4) eşitliği ile verilen $S(\beta, f)$ cezalı kareler toplamı,

$$S(\beta, f) = (y - Z\beta - Nf)'(y - Z\beta - Nf) + \lambda f'Kf \quad (3.5)$$

biçiminde yazılabilir. Pürüzlülük ceza yaklaşımı olarak adlandırılan (3.5) eşitliğindeki cezalı en küçük karelerin amacı (3.4) eşitliğini minimum yapan β ve f vektörlerinin tahminlerini elde etmektir. (3.5) denkleminin β ve f 'e göre türevlerinin sıfıra eşitlenmesi ile elde edilen denklemler aşağıdaki blok matris sistemini verir:

3.YARI PARAMETRİK MODELDE EĞRİSEL ÇİZGİ DÜZELTME
YÖNTEMİNE DAYALI PARAMETRE TAHMİNLERİ VE
PARAMETRELER İLE İLGİLİ SONUÇLAR

Gülin TABAKAN

$$\begin{bmatrix} Z'Z & Z'N \\ N'Z & N'N + \lambda K \end{bmatrix} \begin{pmatrix} \beta \\ f \end{pmatrix} = \begin{bmatrix} Z' \\ N' \end{bmatrix} y \quad (3.6)$$

(3.6) blok matris sistemini sağlayan β ve f (3.5) eşitliğini minimize yapan değerlerdir. Z ve β 'yı içeren (3.2) modelinin parametrik kısmı göz ardı edildiğinde yani $\beta = 0$ olduğunda (3.6),

$$(N'N + \lambda K)f = N'y \quad (3.7)$$

olarak indirgenir ve böylece Nf uyum değerleri vektörünü elde etmek için y vektörüne uygulanan ve $\lambda > 0$ sabitine bağlı bir düzeltme matrisi,

$$S_\lambda = N(N'N + \lambda K)^{-1} N' \quad (3.8)$$

biçiminde elde edilir. (3.1) modelinde x_i düğüm noktaları farklı ve önceden sıralı ise, $N=I$ olması nedeniyle, S_λ düzeltme matrisi aşağıdaki şekle indirgenir:

$$S_\lambda = (I + \lambda K)^{-1} \quad (3.9)$$

Burada I bir birim matristir (Green ve Silverman, 1994).

3.1.1. Backfitting Yöntemi

(3.6) ile verilen matris sistemi genellikle çok büyük olduğundan bu sistemin direk çözümü elverişli değildir. Bunun için bir yaklaşım (3.6) matris denklemini eş zamanlı matris eşitliklerinin çifti olarak aşağıdaki gibi yeniden yazmaktır:

$$Z'Z\beta = Z'(y - Nf) \quad (3.10)$$

$$(N'N + \lambda K)f = N'(y - Z\beta) \quad (3.11)$$

3.YARI PARAMETRİK MODELDE EĞRİSEL ÇİZGİ DÜZELTME
YÖNTEMİNE DAYALI PARAMETRE TAHMİNLERİ VE
PARAMETRELER İLE İLGİLİ SONUÇLAR

Gülin TABAKAN

Bu denklemler sezgisel olarak yorumlanabilir: (3.10) denklemi f biliniyorsa y_i 'den $(Nf)_i = f(x_i)$ çıkarılarak $y - Nf$ farklarının en küçük kareler regresyonu ile β 'nın

$$\hat{\beta} = (Z'Z)^{-1} Z'(y - Nf) \quad (3.12)$$

biçiminde tahmin edileceğini ifade ederken, (3.11) denklemi β biliniyorsa $y_i - z_i'\beta$ farkları için (3.7)'deki biçimde bir kübik eğrisel çizgi düzeltme uyumunu ifade eder:

$$\hat{f} = Nf = N(N'N + \lambda K)^{-1} N'(y - Z\beta) = S_\lambda \tilde{y} \quad (3.13)$$

(3.10) ve (3.11) denklem sistemini çözmek için backfitting algoritması önerilmiştir (Green ve ark., 1985). Backfitting algoritması orijinalinde bir toplamsal modeli değerlendirmek için yinelemeli bir algoritma olarak önerilmiştir. Hastie ve ark. (1989) backfitting algoritmasının normal denklem sistemlerinin uygun bir çözümünü bulan Gauss-Seidel yöntemi olduğunu göstermişlerdir. Backfitting algoritması aşağıdaki biçimdedir:

$$\hat{f}^{(yeni)} = S_\lambda (y - Z\beta^{(eski)}) \quad (3.14)$$

$$\beta^{(yeni)} = (Z'Z)^{-1} Z'(y - \hat{f}^{(yeni)}) \quad (3.15)$$

Burada $\hat{f} = Nf$, (3.14) eşitliği ile verilen $(n \times 1)$ boyutlu vektördür ve S_λ (3.8)'de verilen düzeltme matrisidir. (3.15)'deki algoritma cezalı en küçük kareler tahminine yakınsak bir $\hat{\beta}^{(0)}$ başlangıç değeri buluncaya kadar devam eder (Green ve Silverman, 1994).

3.YARI PARAMETRİK MODELDE EĞRİSEL ÇİZGİ DÜZELTME
YÖNTEMİNE DAYALI PARAMETRE TAHMİNLERİ VE
PARAMETRELER İLE İLGİLİ SONUÇLAR

Gülin TABAKAN

3.1.2. Kısmi Eğrisel Çizgi (Partial Spline) Yaklaşımı

(3.6) ile verilen matris denklemini çözmek için yinelemeli olmayan alternatif bir yöntem kısmi eğrisel çizgi yaklaşımıdır. Kısmi eğrisel çizgi yaklaşımına göre λ 'nın önceden belirlenmiş değeri için (3.2) modelindeki f ve β vektörlerinin tahmini bir S_λ düzeltme matrisi yardımı ile elde edilir. Buradan elde edilen tahmin edici bir kısmi eğrisel çizgi olarak adlandırılır. λ 'nın önceden belirlenmiş değeri için eşitlik (3.2)'ye dayalı f , β ve μ için karşılık gelen tahmin ediciler p alt indisi ile gösterilerek aşağıdaki biçimde elde edilir:

$$\hat{f}_p = Nf = S_\lambda(y - Z\beta) \quad (3.16)$$

$$\begin{aligned} \hat{\beta}_p &= [Z'(I - S_\lambda)Z]^{-1} Z'(I - S_\lambda)y \\ &= (\tilde{Z}'\tilde{Z})^{-1} \tilde{Z}'y \end{aligned} \quad (3.17)$$

Burada S_λ , (3.8)'de verilen düzeltme matrisidir. Bu yaklaşıma göre (3.2) ile verilen yarı parametrik regresyon modelinin ortalama vektörü,

$$\begin{aligned} \mu_p &= Z\hat{\beta}_p + \hat{f}_p = Z\hat{\beta}_p + S_\lambda(y - Z\hat{\beta}_p) \\ &= S_\lambda y + (I - S_\lambda)Z\hat{\beta}_p \\ &= S_\lambda y + (I - S_\lambda)Z[Z'(I - S_\lambda)Z]^{-1} Z'(I - S_\lambda)y \\ &= [S_\lambda + \tilde{Z}(\tilde{Z}'\tilde{Z})^{-1} \tilde{Z}']y \\ &= H_p y \end{aligned}$$

olarak elde edilir. Buradan elde edilen

$$H_p = S_\lambda + \tilde{Z}(\tilde{Z}'\tilde{Z})^{-1} \tilde{Z}' \quad (3.18)$$

3.YARI PARAMETRİK MODELDE EĞRİSEL ÇİZGİ DÜZELTME
YÖNTEMİNE DAYALI PARAMETRE TAHMİNLERİ VE
PARAMETRELER İLE İLGİLİ SONUÇLAR

Gülin TABAKAN

matrisi $\tilde{Z} = (I - S_\lambda)Z$ ile doğrusal regresyondaki şapka matrisinin rolünü oynayan bir matristir. $\tilde{Z}$ matrisi tam ranklı değilse (3.18) eşitliğindeki $(\tilde{Z}'Z)^{-1}$ genelleştirilmiş ters olarak yorumlanacaktır (Schimek, 2000).

3.2. Speckman Yaklaşımı

Bu alt bölümde kısmi eğrisel çizgi yaklaşımına bir alternatif olarak verilen, Speckman (1988) tarafından önerilen bir yaklaşım incelenecektir. Bu yaklaşımda (3.1) modelindeki z_i açıklayıcı değişkeni ile parametrik olmayan x_i bağımsız değişkeninin düzgün bir regresyon ilişkisine sahip olduğu varsayılır:

$$z_i = \zeta(x_i) + \eta_i, \quad i = 1, \dots, n \quad (3.19)$$

(3.19) eşitliğinde $\zeta(x_i)$, x_i 'nin düzgün vektör-fonksiyonu ve η_i , hata terimleri vektörüdür. (3.19) modeli (3.1)'de yerine yazılarak,

$$\begin{aligned} y_i &= (\zeta(x_i) + \eta_i)' \beta + f(x_i) + \varepsilon_i \\ &= \zeta(x_i)' \beta + f(x_i) + (\eta_i' \beta + \varepsilon_i) \end{aligned} \quad (3.20)$$

elde edilir. (3.20) eşitliğindeki $(\eta_i' \beta + \varepsilon_i) = hata$ olarak ifade edilsin. Bu durumda (3.20) ile verilen eşitliğin

$$y_i = f_0(x_i) + hata \quad (3.21)$$

biçiminde olabilmesi için aşağıdaki gibi bir f_0 fonksiyonu tanımlanabilir:

$$f_0(x_i) = \zeta(x_i)' \beta + f(x_i) \quad (3.22)$$

3.YARI PARAMETRİK MODELDE EĞRİSEL ÇİZGİ DÜZELTME
YÖNTEMİNE DAYALI PARAMETRE TAHMİNLERİ VE
PARAMETRELER İLE İLGİLİ SONUÇLAR

Gülin TABAKAN

(3.1) ve (3.22) eşitliklerinin farkı alınarak aşağıdaki ifade elde edilir:

$$y_i - f_0(x_i) = \{z_i - \zeta(x_i)\}' \beta + \varepsilon_i \quad (3.23)$$

S_λ verilen herhangi bir λ düzeltme parametresi için x_i düğüm noktaları ile belirlenen (3.8) eşitliğindeki düzeltme matrisi olsun. Ayrıca y i . gözlem değeri y_i ile gösterilen bağımlı değişken vektörü, Z i . satırları z_i' ile belirtilen bağımsız değişkenlerin gözlem değerlerinin matrisi ve Ξ i .satırı $\zeta(x_i)'$ ile belirtilen matris olarak gösterilirse, (3.23) eşitliğine göre aşağıdaki adımlar gerçekleştirilir (Green ve Silverman, 1994).

Adım 1 Sırası ile $S_\lambda Z$ ve $S_\lambda y$ tahminlerini veren Ξ ve $\{f_0(x_i)\}$ 'yi tahmin için eğrisel çizgi düzeltme yöntemi kullanılır:

$$S_\lambda Z = \Xi = (\zeta(x_1), \dots, \zeta(x_n))'$$

matrisi ve

$$S_\lambda y = f_0 = (f_0(x_1), \dots, f_0(x_n))'$$

vektörü hesaplanır.

Adım 2 (3.23) eşitliğinden

$$y - f_0 = y - S_\lambda y = (I - S_\lambda)y = \tilde{y}$$

ve

$$Z - \Xi = Z - S_\lambda Z = (I - S_\lambda)Z = \tilde{Z}$$

3.YARI PARAMETRİK MODELDE EĞRİSEL ÇİZGİ DÜZELTME
YÖNTEMİNE DAYALI PARAMETRE TAHMİNLERİ VE
PARAMETRELER İLE İLGİLİ SONUÇLAR

Gülin TABAKAN

olarak yazılabilir. Buradan, $\tilde{y} = (I - S_\lambda)y$ ve $\tilde{Z} = (I - S_\lambda)Z$ dönüştürülmüş değişkenleri elde edilir.

Adım 3 (3.23) denkleminin karşılık gelen regresyon denklemi,

$$\tilde{y} = \tilde{Z}\beta + \varepsilon \quad (3.24)$$

biçiminde ifade edilir. Bu modele en küçük kareler (EKK) yönteminin uygulanması ile model (3.1)'deki β parametre vektörünün tahmini aşağıdaki gibi bulunur:

$$\hat{\beta}_s = \{\tilde{Z}'\tilde{Z}\}^{-1} \tilde{Z}'\tilde{y}. \quad (3.25)$$

Adım 4 (3.25) eşitliğindeki $\hat{\beta}_s$ tahmini (3.1) modelinde yerine yazılarak (3.1) yarı parametrik regresyon modeli,

$$y_i^* = f(x_i) + \varepsilon_i \quad (3.26)$$

biçimindeki parametrik olmayan bir modele dönüşür. Burada $y_i^* = y_i - z_i'\hat{\beta}_s$ 'dır. Bağımlı y_i^* değişkenine uygulanan eğrisel çizgi düzeltme yönteminde, (3.1) modelinin parametrik olmayan bileşeni için aşağıdaki gibi bir $\hat{f}$ tahmini elde edilir (Green ve Silverman, 1994).

$$\hat{f}_s = (I + \lambda K)^{-1} y_i^* = S_\lambda y_i^* \quad (3.27)$$

Kısmi eğrisel çizgi yaklaşımına benzer biçimde Speckman yaklaşımına göre (3.2) ile verilen yarı parametrik regresyon modelinin ortalama vektörü,

$$\mu_s = Z\hat{\beta}_s + \hat{f}_s = H_s y$$

3.YARI PARAMETRİK MODELDE EĞRİSEL ÇİZGİ DÜZELTME YÖNTEMİNE DAYALI PARAMETRE TAHMİNLERİ VE PARAMETRELER İLE İLGİLİ SONUÇLAR

Gülin TABAKAN

olarak elde edilir. Buradaki H_s düzeltme matrisi S_λ düzeltme matrisi yardımı ile hesaplanan doğrusal regresyondaki şapka matrisinin rolünü oynayan ve aşağıdaki biçimde verilen bir matristir .

$$H_s = S_\lambda + \tilde{Z} \{ \tilde{Z}' \tilde{Z} \}^{-1} \tilde{Z}' (I - S_\lambda) \quad (3.28)$$

Görüldüğü gibi tahminler kısmi rezidülerde regresyon ile elde edilmiştir. Speckman yaklaşımına göre β 'nın bir tahmini, hem z_i hem de y 'den x 'nin (parametrik olmayan tahmin edici) etkisi çıkarıldıktan sonra yapılırken, kısmi eğrisel çizgi yaklaşımı x -bilgisini sadece z_i 'den çıkarır (Schimek, 2000).

3.3. Varyans- Kovaryans Tahmini

Yarı parametrik regresyon modelinin varyans ve kovaryanslarının tahminleri aşağıda belirtilen amaçlar için gereklidir:

- 1) Düzeltme parametresi seçim ölçütlerinin hesaplanmasında.
- 2) Modelin parametrik bileşeni hakkındaki çıkarımlar.
- 3) Modelin parametrik olmayan fonksiyonu hakkındaki çıkarımlar.

Yarı parametrik modelin parametrik bileşeni için geliştirilen varyans tahmin edicisi (2.36) ile verilen,

$$\hat{\sigma}^2 = \frac{y'A'(I-P)Ay}{tr(A'(I-P)A)}$$

tahmin edicidir ve $P = AZ(Z'AAZ)^{-1}Z'A'$ biçimindedir (Eubank ve ark. 1998, Schimek, 2000). Burada alt bölüm 2.9.2'de belirtildiği gibi $(n-2) \times n$ boyutlu A matrisinin herhangi bir i . satırının i . elemanı $a_i c_i$, $(i+1)$. elemanı $-c_i$ ve $(i+2)$. elemanı $b_i c_i$ olmak üzere, bu satırın diğer tüm elemanları sıfırdır ve a_i, b_i, c_i elemanları alt bölüm 2.9.2'de tanımlandığı gibidir.

3.YARI PARAMETRİK MODELDE EĞRİSEL ÇİZGİ DÜZELTME YÖNTEMİNE DAYALI PARAMETRE TAHMİNLERİ VE PARAMETRELER İLE İLGİLİ SONUÇLAR

Gülin TABAKAN

Schimek (2000) tarafından yapılan çalışmada yarı parametrik regresyon modelinin parametrik katsayılarının varyans-kovaryans tahmin edicileri hem kısmi eğrisel çizgi (β_p) hem de Speckman yaklaşımı (β_s) için hesaplanmıştır. Bu tahmin ediciler için sırasıyla varyans-kovaryans matrisleri aşağıdaki gibidir:

$$V_p = \sigma^2 (\tilde{Z}'Z)^{-1} \tilde{Z}'\tilde{Z}(Z'\tilde{Z})^{-1} \quad (3.29)$$

$$V_s = \sigma^2 (\tilde{Z}'\tilde{Z})^{-1} \tilde{Z}'(I - S_\lambda)^2 \tilde{Z}(\tilde{Z}'\tilde{Z})^{-1} \quad (3.30)$$

(3.29) ve (3.30)'da $\tilde{Z} = (I - S_\lambda)Z$ olup, S_λ (3.8) ile tanımlanan düzeltme matrisidir (Eubank ve ark. 1998, Schimek, 2000).

3.4. Yarı Parametrik Modele ait Çıkarımlar

(3.2) eşitliği ile verilen modeli değerlendirmek için hem parametrik hem de parametrik olmayan bileşenler üzerinde testler yapmak gerekir. Ele aldığımız yarı parametrik regresyon modeli hakkındaki çıkarımlar aşağıdaki varsayımlara dayanır:

- Bağımsız değişkenler ve bağımlı değişken sürekli ölçekle ölçülür.
- Hata varyansı σ^2 'nin tutarlı tahminidir.
- Bağımsız değişkenler arasında iç ilişki (korelasyon) yoktur.
- Parametrik olmayan regresyon tahmin edicisi f , bağımlı değişken y 'ye göre doğrusaldır.
- y bağımlı değişkeni bağımsız ve özdeş dağılıma sahiptir (Schimek, 2000).

3.4.1. Parametrik Bileşen için Çıkarım

Geleneksel regresyon analizine benzer olarak, yarı parametrik regresyon analizi de örneklem verileri üzerinden yapıldığından, elde edilen $\hat{\beta}$ tahmin vektörü,

3.YARI PARAMETRİK MODELDE EĞRİSEL ÇİZGİ DÜZELTME YÖNTEMİNE DAYALI PARAMETRE TAHMİNLERİ VE PARAMETRELER İLE İLGİLİ SONUÇLAR

Gülin TABAKAN

β parametrelerinin anlamlılığının test edilmesinde kullanılır. Böylece parametrik katsayıların testi, modelin anlamlılığını da test eder.

(3.2) ile verilen yarı parametrik regresyon modelinde asimptotik normallik verildiğinde parametrik kısmın katsayıları için güven aralıkları ve test istatistiklerini hesaplamak için β_p (kısmi eğrisel çizgi yaklaşımı) ve β_s 'nin (Speckman yaklaşımı) varyans-kovaryans matrisleri kullanılabilir. V_p ve V_s matrislerinin her biri tahmin edilen $\hat{\beta}$ katsayılarının standart hatalarını (standart errors: SE) verir. Bundan dolayı, parametrik katsayıların istatistiksel açıdan anlamlı olup olmadığını test etmek için kurulan hipotez,

$$H_0 : \hat{\beta}_j = 0$$

$$H_1 : \hat{\beta}_j \neq 0$$

biçimindedir. $\beta_j = 0$ hipotezi z_j bağımsız değişkeninin bağımlı değişken üzerinde etkili olmadığını gösterirken $\beta_j \neq 0$ hipotezi z_j bağımsız değişkeninin bağımlı değişken üzerinde etkili olduğunu göstermektedir. Her bir parametrik katsayıyı test etmek için aşağıdaki t-test istatistiği kullanılır:

$$t_{df} = \frac{\hat{\beta}_j}{SE(\hat{\beta}_j)}, \quad j = 1, \dots, p \quad (3.31)$$

Burada $SE(\hat{\beta}_j)$, $\hat{\beta}_j$ katsayılarının standart hataları olup (3.29) veya (3.30) varyans kovaryans matrisinin köşegen elemanları yardımı ile hesaplanır. Eşitlik (3.31)'de verilen t istatistiği $df = n - tr(S_\lambda) - p$ serbestlik derecesine göre t-tablosundan bulunan değerden büyükse sıfır hipotezi reddedilir, diğer bir ifade ile, β_j katsayısının anlamlı olduğu sonucuna varılır. Speckman (1988), σ^2 'nin bir tahmin edicisi olarak,

3.YARI PARAMETRİK MODELDE EĞRİSEL ÇİZGİ DÜZELTME
YÖNTEMİNE DAYALI PARAMETRE TAHMİNLERİ VE
PARAMETRELER İLE İLGİLİ SONUÇLAR

Gülin TABAKAN

$$\hat{\sigma}^2 = MSE = \frac{\text{Hata Kareler Toplamı}(RSS)}{\text{Serbestlik Derecesi}} = \frac{\|(I - H_s)y\|^2}{tr(I - H_s)'(I - H_s)} \quad (3.32)$$

biçiminde ifade edilen hata kareler ortalamasını kullanmayı önermiştir. Eşitlik (3.32)'nin payında yer alan hata kareler toplamı,

$$RSS = \|(I - H_s)y\|^2$$

olarak tanımlanır. (3.32) ile verilen σ^2 'nin tahmin edicisi pozitif ama asimptotik olarak ihmal edilebilir bir yanlılığa (sapmaya) sahiptir.

Parametrik katsayıların toplu olarak istatistiksel açıdan anlamlı olup olmadığını test edebilmek için hipotezler:

$$H_0 : \beta_1 = \dots = \beta_p = 0$$

$$H_1 : \beta_1 \neq \dots \neq \beta_p \neq 0 \text{ (en az bir } \beta_j \neq 0)$$

biçimindedir. Söz konusu bu hipotezler, $df_1 = p$ ve $df_2 = tr(I - H_s)'(I - H_s)$ serbestlik derecelerine sahip bir F istatistiği yardımıyla test edilir (Schimek, 2000):

$$F_{df_1, df_2} \approx \frac{MSS_{par}}{MSE}. \quad (3.33)$$

Eşitlik (3.33)'de parametrik bileşenin kareler toplamının ortalaması $MSS_{par} = n^{-1}SS_{par}$, parametrik bileşenin kareler toplamı $SS_{par} = \hat{\beta}'_s(Q'Q)^{-1}\hat{\beta}_s/n$ ve $Q' = (\tilde{Z}'\tilde{Z})^{-1}\tilde{Z}'(I - S_\lambda)'$ 'dir. Ayrıca Speckman tarafından önerilen (3.32)'deki varyans tahmin edicisi yerine değiştirilmiş farka dayalı varyans tahmin edicisi (2.36)'da kullanılabilir. Yaklaşık F testi, parametrik kısmın tahmin edicisi hakkında karara varmamızı sağlar (Schimek, 2000).

3.4.2. Parametrik Olmayan Bileşen için Çıkarım

Bu alt bölümde amacımız f eğrisinin biçimsel olarak şeklini değerlendirmektir. Test edilmek istenen sıfır ve alternatif hipotezler aşağıdaki şekilde ifade edilir:

$$H_0 : E(y_i) = \mu \quad (\text{doğrusal fonksiyon})$$

$$H_1 : E(y_i) = f(x_i) \quad (\text{düzgün fonksiyon})$$

Böyle bir test bir parametrik bir modelle karşılaştırılan yarı parametrik modelin bir anlam ifade edip etmeyeceğine karar verilmesine olanak sağladığı için önemlidir. Hastie ve Tibshirani (1990), yarı parametrik duruma uygulanabilen, bir parametrik olmayan uyum $\hat{f}_1$ 'e karşı bir doğru denklem uyumu $\hat{f}_0$ için yaklaşık bir F testi önermiştir. Bu test için gerekli varsayımlar aşağıdaki gibidir:

- Alışılmış en küçük karelerden (EKK) elde edilen $\hat{f}_0$ yansız bir tahmin edicidir.
- Speckman yaklaşımından elde edilen $\hat{f}_1$, yansız bir tahmin edicidir.
- En uygun (optimum) düzeltme parametresi seçilir.

Böylece, söz konusu F -test istatistiği,

$$F_{df_1-df_0, n-df_1} = \frac{\left(\sum_{i=1}^n \hat{\varepsilon}_i^2 - \sum_{i=1}^n \hat{v}_i^2 \right) / (df_1 - df_0)}{\sum_{i=1}^n \hat{v}_i^2 / (n - df_1)} \quad (3.34)$$

formülü ile verilir. Burada $\hat{\varepsilon}_i = (y_i - z_i' \hat{\beta}_{OLS})$ ve $\hat{v}_i = z_i' \hat{\beta}_s + \hat{f}(x_i) - z_i' \hat{\beta}_{OLS}$ dir. Serbestlik dereceleri df_0 , sıradan en küçük karelerde olduğu gibi, modeldeki parametre sayısına (p tane) ve $df_1 = \text{tr}(2H_s - H_s H_s')$ ifadesine eşittir. Eşitlik (3.34)'de verilen bu F -test istatistiği σ^2 'nin tahminini gerektirmez.

3.YARI PARAMETRİK MODELDE EĞRİSEL ÇİZGİ DÜZELTME YÖNTEMİNE DAYALI PARAMETRE TAHMİNLERİ VE PARAMETRELER İLE İLGİLİ SONUÇLAR

Gülin TABAKAN

Hastie and Tibshirani (1990) yaklaşık F testinin başarısının yeterli olduğunu söylemişlerdir. Fakat bazı araştırmacılar küçük örneklem sonuçlarına şüphe ile bakmaktadır (Bowman ve Azzalini (1997) normal ve bağımsız dağılan veriler için sözde-olabilirlik oran testi (pseudo-likelihood ratio test) olarak adlandırılan başka bir yaklaşım önermişlerdir). Raz (1990) tarafından yerel regresyon veya kernel tahmin gibi bir parametrik olmayan yöntem tarafından elde edilen bir düzgün regresyon fonksiyonu durumunda kullanılmak üzere yeni bir test tanımlanmıştır. Bu test farklı alanlarda önerilen bir kesin (exact) permütasyon teste bir yaklaşım olarak oluşturulmuş olup, sınırlı örneklem bilgisine dayalı bir karar için çok daha uygun olarak görülür ve bu testte hataların dağılımı ile ilgili herhangi bir özel varsayım yapılmamaktadır.

Bu kısımda son olarak parametrik olmayan bileşenin güven aralığı hakkında yorum yapacağız. Amacımız, doğrusal olup olmamasından başka, f eğrisinin şekli hakkında kesin karar vermektir. Örneklem değişkenliğinden dolayı parametrik olmayan tahminlerde bazı eğriliklerin her zaman beklenmesi gerekir. Güven aralıkları bazı ana hatlara sahip olmak için değerlendirilebilir. Bununla birlikte yanlılık ($bias(x)$) bilinmeksizin böyle bir aralığı oluşturmak mümkün değildir. Bu yüzden uygulamada yanlılık (sapma) düzeltilmesinden kaçınılır ve sözde (so-called) değişkenlik bantları (sınırları) geniş ölçüde kullanılır. Böylece $f(x)$ 'den daha ziyade $E(f(\hat{x}))$ için güven aralığı $\hat{f}(x) \pm 2SE(\hat{f}(x))$ biçimindedir. Burada $SE(\hat{f}(x))$, f eğrisinin tahmin edilen standart hatasıdır (Schimek, 2000).

4. DÜZELTME PARAMETRESİ (BANT GENİŞLİĞİ) SEÇİM YÖNTEMLERİ

Parametrik olmayan ve yarı parametrik regresyon uyumunun iyi bir şekilde yapılabilmesi için bant genişliği olarak da adlandırılan düzeltme parametresinin değerinin doğru olarak seçilmesi gerekir. Düzeltme parametresi λ 'nın $\hat{f}_\lambda$ 'nın varyansı ve yanlılığı arasında bir uzlaşma sağladığını alt bölüm 2.8.1'den biliyoruz. Bundan dolayı düzeltme parametresi $\lambda > 0$, $\hat{f}_\lambda$ 'nın kalitesi üzerinde çok önemli bir etkiye sahiptir. Böylece parametrik olmayan ve yarı parametrik regresyon konusunda, verilere dayalı olarak elde edilen düzeltme parametresinin seçiminin önemli bir problem olduğunu söyleyebiliriz. Aynı λ değerinin her veri seti ile aynı derecede iyi çalışması beklenemez. Bu bölümde ele alınan düzeltme parametresi seçim yöntemleri; Çapraz Geçerlilik (CV) ölçütü ve bunun değiştirilmesi ile elde edilen Genelleştirilmiş Çapraz Geçerlilik (GCV) ölçütü, klasik Akaike bilgi ölçütünün (AIC) gelişmiş bir biçimi olan AIC_c ölçütü, Mallows'un C_p ölçütü olarak bilinen 'klasik yöntemler', 'risk tahmin yöntemleri' olarak bilinen Tam Kat Düzeltme (EDS) ve Klasik Pilotları Kullanan Risk Tahmini (RECP) ve 'yerleştirme (plug-in) yöntemleri' olarak bilinen yöntemlerdir. Hem yerel doğrusal regresyon hem de kernel yoğunluk tahmin (kernel density estimation) konularında klasik seçicilerin birçoğunun değişkenlik göstermeye son derece eğilimli olması ve aynı zamanda daha az düzgün (undersmooth) eğilime sahip olması gibi sakıncaları vardır. Bu sakıncaları gidermek için bu iki konuda kullanılmak üzere risk tahmin veya yerleştirme (plug-in) yöntemleri önerilmiştir. Bununla birlikte eğrisel çizgi düzeltme yöntemi için bu yöntemlerin kullanımı yaygın değildir. Lee (2003) eğrisel çizgi düzeltme konusunda risk tahmin yöntemlerini incelemiş ve eğrisel çizgi düzeltme için klasik yöntemler ile risk tahmin yöntemlerinden hangilerinin daha iyi olduğunu belirlemek amacı ile bir simülasyon çalışması yapmıştır.

(3.2) ile verilen yarı parametrik regresyon modelinde düzeltme parametresinin seçimi için aşağıda bahsedilen düzeltme parametresi seçim

yöntemlerinde S_λ düzeltme matrisi yerine (3.19) ve (3.28)'de verilen H_s ve H_p matrisleri kullanılarak düzeltme parametresinin değeri elde edilir.

4.1. Çapraz Geçerlilik Ölçütü (Cross Validation-CV)

Çapraz Geçerlilik (CV) ölçütü düzeltme parametresinin seçiminde çok sık olarak kullanılan ölçütlerden biridir.

Çapraz Geçerlilik, (x_i, y_i) veri noktalarından birinin çıkarılması ve geri kalan $(n-1)$ veri noktasına bağlı olarak x_i noktasındaki pürüzsüzlüğün yani x_i noktasındaki uyum $\hat{f}_\lambda^{-i}(x_i)$ 'in tahmin edilmesidir. $\hat{f}_\lambda^{-i}(x_i)$ 'yi tanımlamak için n veri noktasına bağlı doğrusal düzeltmeler ile başlanırsa $n-1$ nokta için karşılık gelen düzeltmeler tanımlanmak zorundadır. y^{-i} , orijinal y bağımlı değişken vektöründen y_i gözleminin çıkarılması ile elde edilen $n-1$ boyutlu gözlemlerin vektörü olsun. Ayrıca, söz konusu y^{-i} vektörüne uygun düzeltme tahmini $\hat{f}_\lambda^{-i}(x_i)$ olsun. Kestirim hatasının çapraz geçerlilik (alışılmış çapraz geçerlilik) ölçütü,

$$CV(\lambda) = \sum_{i=1}^n \{y_i - \hat{f}_\lambda^{-i}(x_i)\}^2 \quad (4.1)$$

olarak tanımlanır. Burada $\hat{f}_\lambda^{-i}(x_i)$, x_i 'deki uyumu gösterir ve i . veri noktasının çıkarılması ile hesaplanır. Eğrisel çizgi düzeltme için ele alınan ve parametrik regresyonda *PRESS* olarak bilinen çapraz geçerliliğin asıl amacı $CV(\lambda)$ fonksiyonunu minimum yapan λ değerini seçmektir. Doğrusal düzelticiler için çapraz geçerlilik parametrik regresyona benzer biçimde elde edildiğinden dolayı ilk olarak parametrik regresyon açısından çapraz geçerlilik ölçütünü ele alalım: (2.2) ile verilen parametrik regresyon modeli için,

$$\hat{y} = Hy = Z(Z'Z)^{-1}Z'y$$

biçimindedir. Böylece *PRESS* ölçütünden hareketle parametrik regresyon için çapraz geçerlilik ölçütü,

$$y_i - \hat{y}_i^{-i} = \left(\frac{y_i - \hat{y}_i}{1 - h_{ii}} \right) \quad (4.2)$$

eşitliğinin her iki tarafının karesinin toplamından hareketle,

$$CV = \sum_{i=1}^n (y_i - \hat{y}_i^{-i})^2 = \sum_{i=1}^n \left(\frac{y_i - \hat{y}_i}{1 - h_{ii}} \right)^2 \quad (4.3)$$

olarak elde edilir. Burada h_{ii} H şapka matrisinin i . köşegen elemanıdır. Yukarıdaki adımlara benzer olarak $\hat{f} = S_\lambda y$ ile verilen doğrusal düzelticiler için çapraz geçerlilik ölçütü aşağıdaki biçimde elde edilir:

$$y_i - \hat{f}_i^{-i}(x_i) = \frac{y_i - \hat{f}(x_i)}{1 - S_{ii}}$$

$$CV(\lambda) = \sum_{i=1}^n \left\{ y_i - \hat{f}_i^{-i}(x_i) \right\}^2 = \sum_{i=1}^n \left\{ \frac{y_i - \hat{f}_\lambda(x_i)}{1 - S_{ii}} \right\}^2 \quad (4.4)$$

Böylece λ parametresinin CV tahmini $CV(\lambda)$ fonksiyonunu minimum yapan değerdir (Ruppert ve ark., 2003).

4.2. Genelleştirilmiş Çapraz Geçerlilik Ölçütü (Generalized Cross Validation-*GCV*)

Çapraz geçerlilik üzerindeki değişikliklerden elde edilen genelleştirilmiş çapraz geçerlilik ölçütü GCV , parametrik regresyon için (4.3) ile verilen $CV(\lambda)$

ölçütünde h_{ii} 'nin yerine H şapka matrisinin izinin (tr) ortalamasının alınması ile elde edilen bir ölçüttür ve aşağıdaki biçimde verilir:

$$GCV = \sum_{i=1}^n \left(\frac{y_i - \hat{y}_i}{1 - df/n} \right)^2 = \sum_{i=1}^n \left(\frac{y_i - \hat{y}_i}{1 - tr(H)/n} \right)^2 \quad (4.5)$$

Benzer biçimde doğrusal düzelticiler için GCV (4.4) eşitliğindeki S_{ii} 'nin ortalama değeri $tr(S_\lambda)/n$ ile yer değiştirmesi temeline dayanır ve (2.4) ile verilen parametrik olmayan regresyon modeli için genelleştirilmiş çapraz geçerlilik ölçütü (4.5)'e benzer biçimde,

$$\begin{aligned} GCV(\lambda) &= \sum_{i=1}^n \left\{ \frac{y_i - \hat{f}_\lambda(x_i)}{1 - tr(S_\lambda)/n} \right\}^2 \\ &= \frac{\|(I - S_\lambda)y\|^2}{\{1 - n^{-1}tr(S_\lambda)\}^2} \\ &= \frac{RSS(\lambda)}{\{1 - n^{-1}df_{fit}\}^2} \end{aligned} \quad (4.6)$$

olarak elde edilir. Burada $y = (y_1, y_2, \dots, y_n)'$, $(\hat{f}(x_1), \dots, \hat{f}(x_n))' = S_\lambda y$ ve $df_{fit} = tr(S_\lambda)$ olup modelin serbestlik derecesidir. λ parametresinin GCV tahmini $GCV(\lambda)$ fonksiyonunu minimum yapan değerdir (Ruppert ve ark., 2003).

(4.6) eşitliğine benzer biçimde (3.2) modeli için GCV ölçütü aşağıdaki biçimde tanımlanır:

$$\begin{aligned} GCV(\lambda) &= \frac{n(y - \mu_\lambda)'(y - \mu_\lambda)}{(n - tr(H_\lambda))^2} \\ &= \frac{n\|(I - H_\lambda)y\|^2}{(n - tr(H_\lambda))^2} \end{aligned}$$

$$= \frac{n^{-1}RSS(\lambda)}{\{1 - n^{-1}tr(H_\lambda)\}^2}$$

Burada $y = (y_1, y_2, \dots, y_n)'$ ve μ_λ , Speckman ve kısmi eğrisel çizgi yaklaşımları için sırası ile $H_s y$ ve $H_p y$ 'dir (Eubank ve ark., 1998).

4.3. Mallows'un C_p Ölçütü (Mallows' C_p Criterion)

Mallows'un C_p ölçütü eğrisel çizgi düzeltme konusunda yansız risk yöntemi olarak bilinir. σ^2 bilindiğinde λ için bir yansız risk tahmini mevcuttur. Regresyonda böyle bir tahmin Mallows tarafından önerilmiş ve Craven ve Wahba tarafından eğrisel çizgi düzeltmeye uygulanmıştır.

$$y = f + \varepsilon, \quad \text{cov}(\varepsilon) = \sigma^2 I$$

parametrik olmayan regresyon modelini göz önüne alalım. Bu model için $\hat{f} = S_\lambda y$ ve ortalama hata kareler toplamı (mean summed squared error) $MSSE$,

$$\begin{aligned} MSSE(\hat{f}) &= E \sum_{i=1}^n \{\hat{f}(x_i) - f(x_i)\}^2 \\ &= E \|\hat{f} - f\|^2 \end{aligned}$$

biçimindedir. Buradan hareketle,

$$\begin{aligned} MSSE(\hat{f}) &= E(S_\lambda y - f)'(S_\lambda y - f) \\ &= f' S_\lambda' S_\lambda f + tr S_\lambda' S_\lambda \text{cov}(y) - 2 f' S_\lambda f + f' f \\ &= f' S_\lambda' S_\lambda f + \sigma^2 tr(S_\lambda' S_\lambda) - 2 f' S_\lambda f + f' f \\ &= \|(S_\lambda - I)f\|^2 + \sigma^2 tr(S_\lambda' S_\lambda) \end{aligned} \tag{4.7}$$

olarak elde edilir.

Mallows (1973) tarafından tanımlanan ve kaynaklarda çeşitli biçimlerde gösterilen bir seçim ölçütü aşağıdaki gibidir:

$$C_p = RSS(p) + 2\hat{\sigma}^2 p \quad (4.8)$$

Burada p aday modeldeki terimlerin sayısıdır. $RSS(p)$ aynı aday model için hata kareler toplamıdır. $\hat{\sigma}^2$ göz önüne alınan en büyük modele bağlı hata varyansının tahminidir. (4.8)'in parametrik olmayan regresyona uygulaması şöyledir: Parametrik olmayan regresyon modeli için hata kareler toplamı,

$$RSS = \sum_{i=1}^n (y_i - \hat{f}(x_i))^2 = \|(I - S_\lambda)y\|^2$$

biçimindedir. (2.30) ile verilen eşitlikten,

$$E(RSS) = \|(S_\lambda - I)f\|^2 + \sigma^2 \text{tr}(S'_\lambda S_\lambda) + \sigma^2(n - 2\text{tr} S_\lambda) \quad (4.9)$$

biçiminde yazılabilir. (4.7) ve (4.9) eşitliklerinden hareketle,

$$E(RSS) = MSSE(\hat{f}) + \sigma^2(n - 2df_{fit})$$

biçiminde yazılır. $\hat{\sigma}^2, \sigma^2$ 'nin yansız tahmini ise o zaman, $E(RSS + 2\hat{\sigma}^2 df_{fit}) = MSSE(\hat{f}) + n\sigma^2$ olduğundan $RSS + 2\hat{\sigma}^2 df_{fit}$ ifadesi $MSSE(\hat{f}) + n\sigma^2$ 'nin yansız tahminidir. $n\sigma^2, S_\lambda$ 'ya bağlı olmadığından C_p 'nin minimumu yaklaşık olarak $MSSE(\hat{f})$ 'nın minimizasyonuna benzerdir. Yani parametrik olmayan regresyon modeli için C_p ölçütü (4.8)'den

$$C_p(\lambda) \equiv RSS(\lambda) + 2\hat{\sigma}^2 df_{fit} \quad (4.10)$$

olarak elde edilir. Bu yöntemi uygulamak için varyansın tahmininin gerekli olduğu görülmektedir. Parametrik regresyona benzer olarak eğrisel çizgi düzeltme yönteminde de hata kareler toplamı (2.32) ile verilen eş değer serbestlik derecesine bölünerek, (4.10) ifadesinde yer alan varyansın tahmin edicisi aşağıdaki gibi elde edilir:

$$\hat{\sigma}_{\lambda_p}^2 = \frac{RSS(\lambda_p)}{n - df_{fit}} = \frac{1}{tr(I - S_{\lambda_p})} \sum_{i=1}^n (y_i - \hat{f}_{\lambda_p}(x_i))^2 = \frac{\|(S_{\lambda_p} - I)y\|}{n - tr(S_{\lambda_p})} \quad (4.11)$$

(4.11) eşitliğinde λ_p , CV ölçütü tarafından seçilen λ 'nın bir ön tahminidir. Böylece model (3.2) için C_p ölçütünü aşağıdaki biçimde verebiliriz:

$$\begin{aligned} C_p(\lambda) &\equiv \left\{ (y - \mu_\lambda)' (y - \mu_\lambda) + 2\sigma^2 tr H_\lambda \right\} \\ &\equiv \left\{ \|(I - H_\lambda)y\|^2 + 2\sigma^2 tr H_\lambda \right\} \end{aligned}$$

Burada $y = (y_1, y_2, \dots, y_n)'$ ve μ_λ Speckman ve kısmi eğrisel çizgi yaklaşımları için sırası ile $H_s y$ ve $H_p y$ 'dir (Ruppert ve ark., 2003).

4.3.1. GCV Ölçütü ve Mallows'un C_p Ölçütü Arasındaki İlişki

Genelleştirilmiş çapraz geçerlilik ölçütünden hareketle,

$$\begin{aligned} \sum_{i=1}^n \left\{ \frac{y_i - \hat{f}_\lambda(x_i)}{1 - tr(S_\lambda)/n} \right\}^2 &= \frac{RSS(\lambda)}{\{1 - n^{-1} df_{fit}\}^2} \\ &= n^2 RSS(\lambda) \{n - df_{fit}\}^{-2} \end{aligned} \quad (4.12)$$

elde edilir.

$$\begin{aligned}\{n - df_{fit}\}^{-2} &= \left\{ n \left(1 - \frac{df_{fit}}{n} \right) \right\}^{-2} \\ &= n^{-2} \left(1 - \frac{df_{fit}}{n} \right)^{-2}\end{aligned}$$

ve

$$\frac{1}{(1-u)^2} = 1 + 2u + 3u^2 + 4u^3 + \dots$$

olup bir terimli Taylor seri yaklaşımı ile,

$$\{n - df_{fit}\}^{-2} \cong n^{-2} \{1 + 2df_{fit}/n\} \quad (4.13)$$

elde edilip (4.11) eşitliği,

$$\hat{\sigma}^2 = \frac{RSS(\lambda)}{n - df_{fit}} \cong \frac{RSS(\lambda)}{n} \quad (4.14)$$

yaklaşık olarak alınıp (4.12) eşitliği aşağıdaki biçimde yazılır:

$$\begin{aligned}n^2 RSS(\lambda) \{n - df_{fit}\}^{-2} &= n^2 RSS(\lambda) \{n^{-2} (1 + 2df_{fit}/n)\} \\ &= RSS(\lambda) + 2n^{-1} df_{fit} RSS(\lambda) \\ &= RSS(\lambda) + 2\hat{\sigma}^2 df_{fit}\end{aligned} \quad (4.15)$$

Buradan,

$$GCV(\lambda) = \sum_{i=1}^n \left\{ \frac{y_i - \hat{f}_\lambda(x_i)}{1 - \text{tr}(S_\lambda)/n} \right\}^2$$

ölçümünün $\{RSS(\lambda) + 2\hat{\sigma}^2 df_{fit}\}$ ifadesine yaklaşık olarak eşit olduğu görülür.

Böylece GCV yaklaşık olarak $C_p(\lambda)$ 'ya eşittir deriz (Ruppert ve ark., 2003).

4.4. Akaike Bilgi Ölçütü (Improved Akaike Information Criterion- AIC)

Çok yaygın olarak kullanılan düzeltme parametresi seçim ölçütlerinden biri Akaike (1973) tarafından önerilen,

$$AIC(\lambda) \equiv \log\{RSS(\lambda)\} + 2df_{fit}/n \quad (4.16)$$

eşitliğidir. Hurvich, Simonoff ve Tsai (1998) kernel regresyon için AIC ölçütünden çok daha düşük yanlılık içeren düzeltilmiş AIC , AIC_c olarak adlandırılan düzeltme parametresi seçim ölçütünü önermişlerdir:

$$\begin{aligned} AIC_c(\lambda) &\equiv \log\{RSS(\lambda)\} + \frac{2\{df_{fit} + 1\}}{n - df_{fit} - 2} \\ &= \log \sum_{i=1}^n \{y_i - \hat{f}_\lambda(x_i)\}^2 + \frac{2\{\text{tr}(S_\lambda) + 1\}}{n - \text{tr}(S_\lambda) - 2} \\ &= \log\|(I - S_\lambda)y\|^2 + \frac{2\{\text{tr}(S_\lambda) + 1\}}{n - \text{tr}(S_\lambda) - 2} \end{aligned} \quad (4.17)$$

(3.2) ile verilen model için (4.17) ile verilen eşitlik,

$$AIC_c(\lambda) \equiv \log\|(I - H_\lambda)y\|^2 + \frac{2\{\text{tr}(H_\lambda) + 1\}}{n - \text{tr}(H_\lambda) - 2}$$

biçimindedir. Burada $y = (y_1, y_2, \dots, y_n)'$ ve H_λ yerine Speckman ve kısmi eğrisel çizgi yaklaşımları için sırası ile H_s ve H_p matrisleri kullanılır. Diğer yöntemlerde olduğu gibi $AIC_c(\lambda)$ ölçütünü minimum yapan λ değeri düzeltme parametresi olarak seçilir (Ruppert ve ark., 2003).

Alt bölüm 4.5 ve 4.6'da ele alınan λ değerinin seçimi için pilot tahminlerin seçimini gerektiren iki risk tahmin yöntemi *RECP* ve *EDS*, yerel doğrusal düzeltmede düzeltme parametresinin seçimi için kullanılan alt bölüm 4.7'de verilen alışılmış yerleştirme (plug-in) yöntemlerine benzerdir. Bununla birlikte alışılmış yerleştirme (plug-in) yöntemlerinden farklı olarak bu iki risk tahmin yöntemi ne f fonksiyonunun 2. türevi f'' 'yi nede en uygun (optimal) düzeltme parametresi için analitik bir ifadenin varlığını gerektirmez (Lee, 2003).

4.5. Klasik Pilotları Kullanan Risk Tahmini (Risk Estimation using Classical Pilots-*RECP*)

Risk fonksiyonu R , tahmin ile gerçek regresyon fonksiyonu arasındaki uzaklığı ölçer. Risk fonksiyonu aşağıdaki gibi yanlılık ve varyansın toplamı olarak gösterilebilir:

$$R(f, \hat{f}_\lambda) = \frac{1}{n} E \|f - \hat{f}_\lambda\|^2 = \frac{1}{n} \left\{ \|(S_\lambda - I)f\|^2 + \sigma^2 \text{tr}(S_\lambda S'_\lambda) \right\} \quad (4.18)$$

(4.18) ile verilen eşitlikte σ^2 ve f 'nin uygun pilot tahminlerinin yerleştirilmesi ile $R(f, \hat{f}_\lambda)$ riskini tahmin etmek ve risk tahmin edicisini minimum yapan λ parametresini seçmek gerekir. Pilot tahminleri seçmek için önerilen bir yöntem; klasik bir yöntem (*CV*, *GCV*, *AIC_c*, *C_p*) kullanarak bir pilot λ_p değerini seçmek ve bu λ_p değeri ile f ve σ^2 'nin pilot tahminleri olan $\hat{f}_{\lambda_p}$ ve $\hat{\sigma}_{\lambda_p}^2$ 'nin hesaplanmasıdır. Pilotların seçimi için bu yol yerel doğrusal regresyon konusunda çok başarılı olarak

gösterilmiştir (Lee ve Solo, 1999). Böylece parametrik olmayan regresyon modeli için $R(f, \hat{f}_\lambda)$ 'nin tahmin edicisi aşağıdaki biçimde verilir:

$$R(\hat{f}_{\lambda_p}, \hat{f}_\lambda) = \frac{1}{n} E \|\hat{f}_{\lambda_p} - \hat{f}_\lambda\|^2 = \frac{1}{n} \left\{ \|(S_\lambda - I)\hat{f}_{\lambda_p}\|^2 + \sigma_{\lambda_p}^2 \text{tr}(S_\lambda S_\lambda') \right\} \quad (4.19)$$

(4.19)'a benzer biçimde (3.2) modeli için *RECP* ölçütü,

$$R(\hat{f}_{\lambda_p}, \hat{f}_\lambda) = \frac{1}{n} \left\{ \|(H_\lambda - I)\hat{f}_{\lambda_p}\|^2 + \sigma_{\lambda_p}^2 \text{tr}(H_\lambda H_\lambda') \right\}$$

olarak tanımlanır. Böylece *RECP* yöntemi diğer yöntemlerde olduğu gibi, $R(\hat{f}_{\lambda_p}, \hat{f}_\lambda)$ ifadesini minimum yapan λ düzeltme parametresini seçer (Lee, 2003).

4.6. Tam Kat Düzeltme (Exact Double Smoothing-EDS)

Wand ve Gutierrez (1997) (4.18)'deki pilot tahminlerin seçimi için tam kat düzeltme olarak adlandırılan başka bir yaklaşım önermişlerdir. Bu yaklaşım pilot tahminlerin iki seviyesinin seçimini içerir ve aşağıdaki gibi tanımlanır. λ_0 $R(f, \hat{f}_\lambda)$ 'yi minimize eden en uygun λ değeri olsun. Bu durumda amacımız $E\{(\lambda_0 - \lambda)^2\}$ 'ı minimize eden λ_{p_1} pilot düzeltme parametresini seçmektir. Bununla birlikte λ_0 bilinmediğinden $E\{(\lambda_0 - \lambda)^2\}$ 'nin pratik olarak minimizasyonu uygun değildir. Bu nedenle Wand ve Gutierrez böyle bir minimizasyonu uygulamak için aşağıdaki yöntemi önermişlerdir. İlk olarak $E\{(\lambda_0 - \lambda)^2\}$ için bir kapalı-form yaklaşımı $L(\lambda_0, \lambda)$ elde edilir. Daha sonra, bilinmeyen λ_0 ile ikinci bir pilot düzeltme parametresi λ_{p_2} yer değiştirilir. Yani, bilinmeyen $E\{(\lambda_0 - \lambda)^2\}$ $L(\lambda_{p_2}, \lambda)$ ile yaklaşık olarak hesaplanabilir. $L(\lambda_{p_2}, \lambda)$ için elde edilen ifade aşağıdaki gibidir:

$$\begin{aligned}
L(\lambda_{p_2}, \lambda) = & \left\{ \hat{f}'_{\lambda_{p_2}} D_{\lambda\lambda_{p_2}} \hat{f}_{\lambda_{p_2}} + \hat{\sigma}_{\lambda_{p_2}}^2 \text{tr}(D_{\lambda\lambda_{p_2}}) \right\}^2 + 4\hat{\sigma}_{\lambda_{p_2}}^2 \hat{f}'_{\lambda_{p_2}} D^2_{\lambda\lambda_{p_2}} \hat{f}_{\lambda_{p_2}} \\
& + 2\hat{\sigma}_{\lambda_{p_2}}^4 \text{tr}(D^2_{\lambda\lambda_{p_2}}) + 4\hat{\sigma}_{\lambda_{p_2}}^2 \text{tr}(\dot{S}_{\lambda_{p_2}} S'_{\lambda_{p_2}}) \left\{ \hat{f}'_{\lambda_{p_2}} D_{\lambda\lambda_{p_2}} \hat{f}_{\lambda_{p_2}} + \hat{\sigma}_{\lambda_{p_2}}^2 \text{tr}(D_{\lambda\lambda_{p_2}}) \right\} \\
& + 4\hat{\sigma}_{\lambda_{p_2}}^4 \text{tr}^2(\dot{S}_{\lambda_{p_2}} S'_{\lambda_{p_2}})
\end{aligned} \quad (4.20)$$

(4.20) eşitliğinde $\dot{S}_{\lambda_{p_2}} = \lambda_{p_2}^{-1} S_{\lambda_{p_2}} (S_{\lambda_{p_2}} - I)$ ve $D_{\lambda\lambda_{p_2}} = 2\lambda_{p_2}^{-1} S_{\lambda} S_{\lambda_{p_2}} (S_{\lambda_{p_2}} - I)^2 S_{\lambda}$ 'dir.

Sonuç olarak *EDS* yöntemi $R(\hat{f}_{\lambda_{p_1}}, \hat{f}_{\lambda})$ ifadesini minimize edecek şekilde λ değerini seçer. Burada λ_{p_1} , $L(\lambda_{p_2}, \lambda)$ 'yi minimize eder ve λ_{p_2} değeri *CV* yöntemi kullanılarak seçilebilir.

Lee (2003) eğrisel çizgi düzeltme için klasik yöntemler ile risk tahmin yöntemlerinden hangilerinin daha iyi olduğunu belirlemek amacı ile bir simülasyon çalışması yapmıştır ve simülasyon çalışması sonucunda aşağıdaki sonuçları elde etmiştir:

- Ele alınan hiçbir yöntem düzgün olarak en iyi performansı göstermemiştir.
- Üç klasik yöntem *CV*, *GCV* ve C_p çok benzer sonuçlar vermiştir.
- Yüksek gürültülü basit bir regresyon fonksiyonu için *RECP* daha üstündür.
- Heterokedastik hatalar altında AIC_c en iyi yöntemdir.

Böylece, yukarıda bahsedilen hiçbir yöntemin, eğrisel çizgi düzeltme konusunda düzgün olarak en iyi olduğu söylenemez. Ancak sabit hata varsayımı sağlanırsa *RECP* ölçütünün, diğer durumlarda ise AIC_c ölçütünün kullanımı önerilmiştir. Ayrıca yukarıdaki seçim yöntemlerinden *CV*, *GCV* ve *AIC* benzer biçimde hesaplama zamanı gerektirirken, C_p ve *RECP* iki nümerik minimizasyon, *EDS* ise üç nümerik minimizasyon gerektirir (Lee, 2003).

4.7. Yerleştirme (Plug-in) Yöntemleri

Düzeltilme parametresi seçim yöntemlerinin tamamen farklı bir sınıfı, yanlılık ve varyansın tahmin edilmesi ile bir risk ölçümünü tahmin etmeye çalışan

‘yerleştirme (*plug-in*) yöntemleri’ olarak ifade edilen yöntemlerdir. Yöntem daha çok kernel yoğunluk tahmini durumunda geliştirilmiştir fakat kernel regresyon ve yerel polinom regresyona uyarlaması da bulunmaktadır.

Gerçek regresyon fonksiyonu ile tahmini arasındaki uzaklığı ölçen risk fonksiyonuna örnek olarak,

$$R(f, \hat{f}) = \frac{1}{\sigma^2} \sum_{i=1}^n E\left(\left(\hat{f}(x_i) - f(x_i)\right)^2\right)$$

verilebilir (Loader, 2004). İdeal olarak iyi bir tahmin düşük riskli olmalıdır. Karesel-hata riski için varyans yanlılık ayrışımı,

$$\begin{aligned} \sigma^2 R(f, \hat{f}) &= \sum_{i=1}^n \left(bias(\hat{f}(x_i))\right)^2 + \sum_{i=1}^n \text{var}(\hat{f}(x_i)) \\ &= \sum_{i=1}^n \left(\sum_{j=1}^n w_j(x_i) f(x_j) - f(x_i)\right)^2 + \sigma^2 \sum_{i=1}^n \|w(x_i)\|^2 \end{aligned} \quad (4.21)$$

ile verilir. Yerleştirme tahmin ortalama fonksiyon $f(\cdot)$ ’nin bir başlangıç pilot tahminini oluşturma ile başlar. O zaman bu h düzeltme parametresi üzerinde minimize edilebilen (4.21) risk tahmininin yerini tutar.

İstatistiksel çalışmalarda yerleştirme fikrinin çok değişik biçimleri vardır. En basiti yanlılık ve varyans için,

$$E(\hat{f}(x)) - f(x) \approx \frac{f''(x)}{2} h^2 \frac{\int u^2 K(u) du}{\int K(u) du} \quad (4.22)$$

ve

$$\text{var}(\hat{f}(x)) \approx \frac{\sigma^2}{nh\dot{f}(x)} \frac{\int K(u)^2 du}{\left(\int K(u) du\right)^2} \quad (4.23)$$

asimptotik yaklaşımlarını kullanan risk fonksiyonudur. (4.22) ve (4.23) eşitlikleri aşağıdaki gibi elde edilir:

$$\hat{f}(x) = \sum_{i=1}^n w_i(x) y_i$$

ile gösterilen bir doğrusal düzelticinin yanlılığı,

$$\begin{aligned} E(\hat{f}(x)) - f(x) &= \sum_{i=1}^n w_i(x) E(y_i) - f(x) \\ &= \sum_{i=1}^n w_i(x) f(x_i) - f(x) \end{aligned} \quad (4.24)$$

ile verilir. Bu ifade bilinmeyen ortalama fonksiyon $f(x)$ 'e bağlıdır. Yanlılığı daha fazla kavrayabilmek için şu yaklaşımlar verilebilir:

- 1) $f(\cdot)$ 'nin düşük dereceli bir Taylor açılımı uydurulan x noktası civarındadır.
- 2) Toplamların tahmini integraller vasıtasıyla olur.

Örneğin, yerel doğrusal regresyonun yanlılığını göz önüne alalım. Bir üç terimli Taylor serisi,

$$f(x_i) = f(x) + (x_i - x)f'(x) + \frac{(x_i - x)^2}{2} f''(x) + o(h^2)$$

olup bu ifadeyi (4.24)'de yerine yazarsak,

$$\begin{aligned} E(\hat{f}(x)) - f(x) &= \sum_{i=1}^n w_i(x) f(x_i) - f(x) \\ &= \sum_{i=1}^n w_i(x) \left(f(x) + (x_i - x)f'(x) + \frac{(x_i - x)^2}{2} f''(x) + o(h^2) \right) - f(x) \\ &= f(x) \sum_{i=1}^n w_i(x) + f'(x) \sum_{i=1}^n (x_i - x) w_i(x) + \frac{f''(x)}{2} \sum_{i=1}^n (x_i - x)^2 w_i(x) - f(x) + o(h^2) \end{aligned} \quad (4.25)$$

elde edilir. Yerel doğrusal regresyon için,

$$\sum_{i=1}^n w_i(x) = 1$$

$$\sum_{i=1}^n (x_i - x) w_i(x) = 0$$

olduğu gösterilebilir. Bu basitleştirme ile yanlılık,

$$E(\hat{f}(x)) - f(x) = \frac{f''(x)}{2} \sum_{i=1}^n (x_i - x)^2 w_i(x) + o(h^2) \quad (4.26)$$

olur. Bu anlatım ortalama fonksiyondaki yanlılığın bağımlılığını karakterize eder. Yanlılığın baskın terimi ortalama fonksiyonun 2. türevi f'' ile orantılıdır. Şimdiki adım hem (4.26) hem de (2.20) matris eşitliğinde tanımlanan $w_i(x)$ 'e integraller vasıtasıyla yaklaşım olmalıdır. Buradan hareketle,

$$E(\hat{f}(x)) - f(x) = \frac{f''(x)}{2} \sum_{i=1}^n (x_i - x)^2 w_i(x) + o(h^2)$$

$$\approx \frac{f''(x)}{2} h^2 \frac{\int u^2 K(u) du}{\int K(u) du} \quad (4.27)$$

olur. Burada $f''(x)$ 'deki bağımlılığa ek olarak şimdi de h bant genişliğindeki bağımlılığı görüyoruz: h bant genişliği artarken yanlılık bant genişliği ile karesel olarak artmaktadır. (4.27)'ye benzer yanlılık genişlemeleri Ruppert ve Wand'ın, (1994) çalışmasında çok daha genel olarak elde edilmiştir; Elde edilen sonuçlar keyfi derecede yerel polinomlar ve çok boyutlu uyumları da kapsar.

Bir doğrusal düzelticinin (4.23) ile verilen varyansını elde etmek için (2.3)'teki ε_i rasgele hataları hakkında varsayımlar yapmamız gerekir. En yaygın

4. DÜZELTME PARAMETRESİ (BANT GENİŞLİĞİ)

SEÇİM YÖNTEMLERİ

Gülin TABAKAN

varsayım $\text{var}(\varepsilon_i) = \sigma^2$ varyansı ile hataların bağımsız ve özdeş dağılmasıdır. Bir

doğrusal düzeltici için $\hat{f}(x) = \sum_{i=1}^n w_i(x) y_i$ tahmininin varyansı,

$$\text{var}(\hat{f}(x)) = \sum_{i=1}^n w_i(x)^2 \text{var}(y_i) = \sigma^2 \|w(x)\|^2 \quad (4.28)$$

biçiminde yazılır. Varyans için tahminler integraller ile toplamaların yer değiştirmesi ile elde edilebilir. Yerel doğrusal regresyon için bu,

$$\text{var}(\hat{f}(x)) \approx \frac{\sigma^2}{nh\dot{f}(x)} \frac{\int K(u)^2 du}{\left(\int K(u) du\right)^2} \quad (4.29)$$

şeklinde yazılabilir. Burada $\dot{f}(x)$, x_i tasarım noktalarının yoğunluğudur.

(4.21) eşitliğinde (4.22) ve (4.23)'ü yerlerine yazarak,

$$\sigma^2 R(f, \hat{f}) \approx h^4 \left(\frac{\int u^2 K(u) du}{2 \int K(u) du} \right)^2 \sum_{i=1}^n f''(x_i)^2 + \frac{\sigma^2}{nh} \frac{\int K(u)^2 du}{\left(\int K(u) du\right)^2} \sum_{i=1}^n \frac{1}{\dot{f}(x_i)}$$

elde edilir. Eğer tasarım noktaları bir $[a, b]$ aralığında düzgün olarak yer alırsa, bu durumda belirli integral için ortalama değer teoreminden toplamlara integraller ile yaklaşımla,

$$\sigma^2 R(f, \hat{f}) \approx nh^4 \left(\frac{\int u^2 K(u) du}{2 \int K(u) du} \right)^2 \frac{1}{b-a} \int_a^b f''(x)^2 dx + \frac{(b-a)\sigma^2}{h} \frac{\int K(u)^2 du}{\left(\int K(u) du\right)^2}$$

elde edilir. Bu ifadeyi h üzerinde minimum yapmak asimptotik olarak en iyi düzeltme parametresi değerini verecektir. Yani,

$$\begin{aligned}
R(f, \hat{f}) &\approx \frac{nh^4}{\sigma^2} \left(\frac{\int u^2 K(u) du}{2 \int K(u) du} \right)^2 \frac{1}{b-a} \int_a^b f''(x)^2 dx + \frac{(b-a)\sigma^2}{\sigma^2 h} \frac{\int K(u)^2 du}{\left(\int K(u) du \right)^2} \\
\frac{4nh^3}{\sigma^2} \left(\frac{\int u^2 K(u) du}{2 \int K(u) du} \right)^2 \frac{1}{b-a} \int_a^b f''(x)^2 dx - \frac{(b-a)}{h} \frac{\int K(u)^2 du}{\left(\int K(u) du \right)^2} &= 0 \\
h_{opt}^5 &= \frac{\sigma^2 (b-a)^2 \int K(u)^2 du}{n \left(\int u^2 K(u) du \right)^2 \int_a^b f''(x)^2 dx} \quad (4.30)
\end{aligned}$$

olacaktır. h_{opt} ifadesi σ^2 'nin ve $\int_a^b f''(x)^2 dx$ 'in yerine tahminlerini kullanmayı gerektirir. σ^2 tahmini için,

$$\hat{\sigma}^2 = \frac{1}{n - 2tr(S_\lambda) + tr(S_\lambda S'_\lambda)} \sum (y_i - \hat{f}(x_i))^2 \quad (4.31)$$

eşitliği kullanılabilir fakat $\int_a^b f''(x)^2 dx$ tahmini daha problemlidir. Bilinen bir yöntem düzeltmenin bir pilot tahminini kullanarak ikinci türevi tahmin etmektir. Bu durumda $\int_a^b \hat{f}''(x)^2 dx$ tahmini kullanılır. Eğer yerel ikinci dereceden tahmin bir pilot tahmin durumunda kullanılırsa eğrilik katsayısı $\hat{a}_2$, $f''(x)$ 'in bir tahmini olarak kullanılabilir. Fakat ikinci türevin tahmini için bir pilot tahminin kullanımı problemlidir. Pilot tahmini için düzeltme parametresinin değeri seçilmek zorundadır ve tahmin edilen uygun düzeltme parametresi $\hat{h}_{opt}$, pilot düzeltme parametresinin seçimine oldukça duyarlıdır. Konu ile ilgili daha detaylı bilgi Loader'ın (1999) çalışmasında bulunabilir (Loader, 2004).

5. YARI PARAMETRİK MODELLERİN DALGACIK TAHMİNİ

Yarı parametrik modellerde parametre tahmini için çalışmalarda çeşitli yöntemler önerilmiştir. Bu modeldeki parametrik olmayan bileşenin tahmini için kullanılan yaklaşımlardan birinin eğrisel çizgi düzeltme yöntemi olduğunu ve kernel regresyon ile yerel polinom yöntemlerinin de yarı parametrik modellerde yaygın olarak kullanıldığını önceki bölümlerden biliyoruz. Bilinmeyen parametrik olmayan bileşen f 'nin tüm bu yöntemlerde yüksek düzgünlüğe sahip olduğu varsayılır. Fakat gerçekte bu varsayım sağlanmayabilir. İşaret ve görüntü işleme gibi bazı uygulamalı alanlarda nesneler çoğunlukla homojen değildir. Bundan dolayı bu modelde parametrik olmayan bileşenin daha az düzgün olması durumunda bir dalgacık tabanlı tahmin yöntemi geliştirilmiştir (Chang ve Qu, 2004). Bu bölümde dalgacık tahmin yönteminin parametrik olmayan ve yarı parametrik regresyon modellerine uygulanışı ele alınacaktır.

Dalgacık dönüşümü DD (wavelet transform) matematikçiler tarafından ortaya atılmış bir sinyal işleme yöntemi olup son yıllarda kullanımında artan bir eğilim gözlenmiştir. Kullanıldığı alanlar arasında görüntü işleme teknikleri, ses ve görüntü sıkıştırma, tıbbi uygulamalar, radar, akustik, gürültünün yok edilmesi gibi alanları sayabiliriz.

Dalgacıklar, veri veya fonksiyonların gösteriminde kullanılan belirli matematiksel gereksinimleri karşılayan fonksiyonlardır. Bir sinyaldeki özelliklerin belirlenmesi için gerekli olan zaman ve frekans verilerini içermesi dalgacık kullanımının en önemli sebebidir.

Son yıllarda, dalgacık düzeltme (wavelet smoothing) parametrik olmayan regresyon çalışmalarında büyük önem kazanmıştır. Bu tahmin ediciler alışılmış parametrik olmayan regresyon yöntemlerinin doğal bir uzamasıdır, diğer bir deyişle ortogonal seri yöntemleridir. Üstelik bu tahmin ediciler hızlı algoritmalar tarafından kolayca gerçekleştirildiğinden dolayı pratik durumlarda çok caziptirler (Antoniadis, 2007).

Bu bölümde Chang ve Qu (2004) tarafından yarı parametrik modelde parametre tahmini için geliştirilen bir dalgacık yaklaşımı ele alınmıştır. Bu

yaklaşımında yarı parametrik modelin bir tahmin edicisi; parametrik olmayan bileşenin dalgacık katsayılarının l_1 normunun cezalandırılıp, rezidü vektörünün l_2 normunun karesinin minimize edilmesi ile elde edilmiştir. Bu yaklaşım, eğrisel çizgi düzeltme, kernel ve parçalı polinom yaklaşımları gibi yarı parametrik modeller için var olan geleneksel düzeltme yöntemlerinde parametrik olmayan fonksiyonun kısıtlayıcı düzgünlüğünden kaçınan, parametrik olmayan regresyon yöntemleri için dalgacık yaklaşımının bir uzamasıdır.

Yarı parametrik modeller için dalgacık yaklaşımı uygulamanın temel nedeni f için var olan yaklaşımların tümünün yüksek düzgünlüğe sahip olmasıdır. Bu bölümde yarı parametrik modelin dalgacık tahmini için parametrik olmayan dalgacık regresyon hakkında kısa bir bilgi verilecek ve daha sonra bu yaklaşımın yarı parametrik modele genişlemesi verilip en uygun dalgacık tahmin edicisi için Chang ve Qu (2004) tarafından ele alınan gerek ve yeter koşullar verilecektir.

5.1. Parametrik Olmayan Regresyonda Dalgacık Tahmin

$$y_i = f(t_i) + \varepsilon_i, \quad i = 1, \dots, n \quad (5.1)$$

ile verilen parametrik olmayan regresyon modelini ele alalım. Bu modelde f fonksiyonu bilinmeyen bir fonksiyon ve $\{\varepsilon_i\}$ 'ler gürültüler olup genel olarak bağımsız özdeş olarak dağılan 0 ortalamalı ve σ^2 varyanslı dağılıma sahip olduğu varsayılır. Bu modelde ayrık dalgacık dönüşümünü (ADD) uygulamak için $\{t_i\}$ 'lerin eşit olarak yer aldığı ve n 'in 2'nin kuvveti olduğu varsayılacaktır. Vektör gösterimi ile (5.1) eşitliğindeki model,

$$y = f + \varepsilon \quad (5.2)$$

biçiminde yazılır. Bir ortogonal W matrisi yardımı ile ayrık dalgacık dönüşümü gösterilebilir. W 'nin gürültülü gözlem y 'ye uygulanması ile, y 'nin dalgacık

dönüşümü $y : w = Wy$ olarak elde edilir. $\theta = Wf$, f 'nin dalgacık dönüşümü olsun. Bu durumda ayırık dalgacık dönüşüm matrisinin transpozu W' ile,

$$\begin{aligned} W'\theta &= W'Wf \\ W'\theta &= f \end{aligned}$$

elde edilir. Böylece (5.2)'den gözlenen veri dalgacık bölgesinde,

$$y = W'\theta + \varepsilon \quad (5.3)$$

biçiminde bir doğrusal model olarak ifade edilebilir. θ 'nın alışılmış en küçük kareler tahmin edicisi $\hat{\theta}_{LS} = Wy$ deneysel dalgacık katsayılarıdır. $\hat{\theta}_{LS}$ θ 'nın bir yansız tahminidir ve kovaryans matrisi $\sigma^2 I$ 'dir. $\hat{f} = W'\hat{\theta}_{LS}$ tahmin edicisi gözlenen y verilerini basit olarak interpolate eder. Bu yüzden veriler gürültüden (sinyallerden) tamamen temizlenmezler.

Veriyi gürültüden temizlemek için genel olarak cezalı en küçük kareler yaklaşımı kullanılır. Verilen bir $\lambda > 0$ için,

$$\min_{\theta} \left\{ \frac{1}{2} \|y - W'\theta\|_2^2 + \lambda \|\theta\|_1 \right\} = \frac{1}{2} \|w - \theta\|_2^2 + \lambda \|\theta\|_1$$

cezalı en küçük kareler problemi için çözüm w 'nin yumuşak eşikle tutmasıdır:

$$\hat{\theta}_s = \text{sgn}(w) \left(|w| - \lambda e \right)_+ = \begin{cases} \text{sgn}(w) (|w| - \lambda e), & |w| > \lambda \\ 0, & |w| \leq \lambda \end{cases} \quad (5.4)$$

Burada, $e = (1, \dots, 1)'$ ve

$$\text{sgn}(w_i) \left(|w_i| - \lambda \right)_+ = \arg \min_{\theta_i} \frac{1}{2} (w_i - \theta_i)^2 + \lambda |\theta_i|, \quad i = 1, \dots, n$$

biçimindedir. λ ile gösterilen düzeltme parametresi veya eşik değerinin (treshold) seçimi için bir çok yöntem önerilmiştir. Eşik değerinin belirlenmesi, $\lambda_{UV} = \sigma \sqrt{2 \log(n)}$ ifadesine veya çapraz geçerlilik ölçütüne göre yapılır. Burada n işaretin uzunluğu ve σ gürültünün standart sapmasıdır.

5.2. Yarı Parametrik Modelde Dalgacık Tahmin

Parametrik olmayan regresyondaki dalgacık tahmin yönteminin yarı parametrik modellere uygulanabilmesi için,

$$y_i = x_i' \beta + f(t_i) + \varepsilon_i, \quad 1 \leq i \leq n$$

ile verilen yarı parametrik modelde alt bölüm 5.1’de olduğu gibi $\{t_i\}$ ’lerin eşit olarak yer aldığı ($t_i = i/n$) ve n ’in 2’nin kuvveti olduğu varsayılır. Bu modelde gözlem sayısı n , bağımsız değişken sayısı p olmak üzere; y_i ’ler gözlemler, x_i ’ler bilinen tasarım noktalarının p – boyutlu vektörü ($p \leq n$), β p – boyutlu bilinmeyen parametre vektörü, f bilinmeyen bir fonksiyon, ε ise gözlenemeyen bağımsız ve özdeş $N(0, \sigma^2)$ dağılan rasgele hatalar olarak varsayılır. Ayrıca bu modelde genelliği bozmaksızın $t \in [0,1]$ varsayılacaktır.

Dalgacık bölgesinde gözlemlenen veriler bir doğrusal model biçiminde,

$$y = X\beta + W'\theta + \varepsilon \quad (5.5)$$

olarak ifade edilebilir. Burada β biliniyorsa (5.5) ile verilen model (5.3) ile aynıdır ve θ ’yı tahmin için parametrik olmayan dalgacık dönüşümü uygulanır. Bu durumda (5.5) eşitliğinde temel amaç β ’yı tahmin etmek olmalıdır. Verilen bir λ için θ ’nın l_1 -normunun cezalandırılması ile,

$$l(\beta, \theta) = \frac{1}{2} \|y - X\beta - W'\theta\|_2^2 + \lambda \|\theta\|_1$$

ifadesini minimum yapan β ve θ 'yı buluruz. Bu durumda W matrisinin ortogonallığı ile,

$$\begin{aligned} l(\beta, \theta) &= \frac{1}{2} \|Wy - WX\beta - \theta\|_2^2 + \lambda \|\theta\|_1 \\ &= \frac{1}{2} \|w - U\beta - \theta\|_2^2 + \lambda \|\theta\|_1 \end{aligned} \quad (5.6)$$

bulunur. Burada $U = [u_1, \dots, u_n]' = WX$, X matrisinin her bir sütununun dönüşümü ile elde edilen, X matrisinin ayrıık dalgacık dönüşümüdür ve tam sütun ranklıdır. Eşitlik (5.6)'da $|\beta| \rightarrow \infty$ veya $|\theta| \rightarrow \infty$ iken $l(\beta, \theta)$ 'nın sonsuza gitme eğiliminde olduğu görülmektedir. Böylece $l(\beta, \theta)$ 'nın sonlu minimumları vardır. $l(\beta, \theta)$ her biri bir norm ile ifade edilen iki terimin toplamı biçiminde olduğundan β ve θ 'nın bir konveks fonksiyonudur. Üçgen eşitsizliğinden dolayı herhangi bir norm konvektir ve konveks bir fonksiyonun en az bir minimum noktaya sahip olduğu iyi bilinmektedir (Qu, 2001). Bazı i için $\theta_i = 0$ olduğunda, $l(\beta, \theta)$ θ 'ya göre türevlenemediğinden minimumların özellikleri için yönlü türevleri kullanırız.

Burada $l(\beta, \theta)$ bazı noktalarda, örneğin $\theta = 0$ 'da iki yanlı yönlü türeve sahip olmayabilir. Bu yüzden $L(x)$ 'in minimumları aşağıdaki lemmaya göre belirlenebilir.

Lemma 5.1. $L(x)$, $\Re^m$ 'de herhangi bir noktada tek yanlı yönlü türeve sahip olan konveks bir fonksiyon olsun. Bu durumda $\hat{x}$ 'nın $L(x)$ 'in bir minimumu olması için gerek ve yeter koşul,

$$L'(\hat{x}, h) \geq 0, \quad \forall h \in \Re^m \quad (5.7)$$

olmasıdır.

Böylece, yarı parametrik modelin dalgacık tahmin edicisini belirleyen aşağıdaki teoremi verebiliriz.

Teorem. 5.1.

$$\{\hat{\beta}, \hat{\theta}\} = \arg \min_{\{\beta, \theta\}} l(\beta, \theta) = \arg \min_{\{\beta, \theta\}} \frac{1}{2} \|w - U\beta - \theta\|_2^2 + \lambda \|\theta\|_1 \quad (5.8)$$

eşitliği ancak ve ancak aşağıdaki koşullar sağlanırsa doğrudur.

$$U'(w - U\hat{\beta} - \hat{\theta}) = 0, \quad (5.9)$$

$$\hat{\theta} = \text{sgn}(w - U\hat{\beta}) \left(|w - U\hat{\beta}| - \lambda e \right)_+ \quad (5.10)$$

İspat. $\theta \in \mathfrak{R}^n$ için aşağıda verilen indeks kümelerini tanımlayalım:

$$\zeta(\theta) = \{i : \theta_i = 0\}, \bar{\zeta}(\theta) = \{i : \theta_i \neq 0\}. \quad (5.11)$$

Lemma 5.1'den hareketle (5.9) ve (5.10) koşullarının,

$$\forall q \in \mathfrak{R}^p \text{ ve } t \in \mathfrak{R}^n \text{ için, } l'(\hat{\beta}, \hat{\theta}; q, t) \geq 0,$$

ifadesine denk olduğunu göstermek yeterlidir. Burada,

$$l'(\hat{\beta}, \hat{\theta}; q, t) = \lim_{\alpha \rightarrow 0^+} \frac{l(\hat{\beta} + \alpha q, \hat{\theta} + \alpha t) - l(\hat{\beta}, \hat{\theta})}{\alpha}$$

dır. $\alpha > 0$ için aşağıdaki eşitliği tanımlayalım:

$$\Delta l(\alpha) = \frac{1}{\alpha} [l(\hat{\beta} + \alpha q, \hat{\theta} + \alpha t) - l(\hat{\beta}, \hat{\theta})].$$

(5.6) ve (5.11)'den hareketle basit cebirsel işlemler sonucunda,

$$\begin{aligned}
\Delta l(\alpha) &= \frac{1}{\alpha} \left\{ \frac{1}{2} \|w - U(\hat{\beta} + \alpha q) - (\hat{\theta} + \alpha t)\|_2^2 + \lambda \|\hat{\theta} + \alpha t\|_1 - \frac{1}{2} \|w - U\hat{\beta} - \hat{\theta}\|_2^2 - \lambda \|\hat{\theta}\|_1 \right\} \\
&= \frac{1}{\alpha} \left\{ \frac{1}{2} \|w - U\hat{\beta} - U\alpha q - (\hat{\theta} + \alpha t)\|_2^2 + \lambda \|\hat{\theta} + \alpha t\|_1 - \frac{1}{2} \|w - U\hat{\beta} - \hat{\theta}\|_2^2 - \lambda \|\hat{\theta}\|_1 \right\} \\
&= \frac{1}{\alpha} \left\{ \frac{1}{2} \left[(w - U\hat{\beta} - \hat{\theta} - \alpha(Uq + t))' (w - U\hat{\beta} - \hat{\theta} - \alpha(Uq + t)) \right] + \lambda \|\hat{\theta} + \alpha t\|_1 \right. \\
&\quad \left. - \frac{1}{2} \left[(w - U\hat{\beta} - \hat{\theta})' (w - U\hat{\beta} - \hat{\theta}) \right] - \lambda \|\hat{\theta}\|_1 \right\} \\
&= \frac{1}{\alpha} \left\{ \frac{1}{2} \left[(w - U\hat{\beta} - \hat{\theta})' (w - U\hat{\beta} - \hat{\theta}) + \alpha^2 (Uq + t)' (Uq + t) \right. \right. \\
&\quad \left. \left. - \alpha (w - U\hat{\beta} - \hat{\theta})' (Uq + t) - \alpha (Uq + t)' (w - U\hat{\beta} - \hat{\theta}) \right] \right. \\
&\quad \left. + \lambda \|\hat{\theta} + \alpha t\|_1 - \frac{1}{2} \left[(w - U\hat{\beta} - \hat{\theta})' (w - U\hat{\beta} - \hat{\theta}) \right] - \lambda \|\hat{\theta}\|_1 \right\} \\
&= \frac{1}{2} \alpha (Uq + t)' (Uq + t) + (Uq + t)' (-w + U\hat{\beta} + \hat{\theta}) + \frac{1}{\alpha} (\lambda \|\hat{\theta} + \alpha t\|_1 - \lambda \|\hat{\theta}\|_1) \\
&= \frac{1}{2} \alpha (Uq + t)' (Uq + t) + (Uq + t)' (-w + U\hat{\beta} + \hat{\theta}) + \frac{1}{\alpha} \left(\alpha \lambda \sum_{i \in \zeta(\hat{\theta})} |t_i| + \lambda \sum_{i \in \bar{\zeta}(\hat{\theta})} |\hat{\theta}_i + \alpha t_i| - \sum_{i \in \bar{\zeta}(\hat{\theta})} |\hat{\theta}_i| \right) \\
&= \frac{1}{2} \alpha (Uq + t)' (Uq + t) + (Uq + t)' (-w + U\hat{\beta} + \hat{\theta}) + \lambda \left(\sum_{i \in \zeta(\hat{\theta})} |t_i| + \frac{1}{\alpha} \sum_{i \in \bar{\zeta}(\hat{\theta})} (|\hat{\theta}_i + \alpha t_i| - |\hat{\theta}_i|) \right)
\end{aligned} \tag{5.12}$$

elde edilir. Yeterince küçük $\alpha > 0$ için,

$$\text{sgn}(\hat{\theta}_i + \alpha t_i) = \text{sgn}(\hat{\theta}_i), \quad \forall i \in \bar{\zeta}(\hat{\theta}),$$

olduğundan ve sgn işaret fonksiyonunun $|x| = \text{sgn}(x)x$ özelliğinden dolayı (5.12) eşitliğinden,

$$\begin{aligned} \Delta l(\alpha) = & \frac{1}{2} \alpha (Uq + t)' (Uq + t) + (Uq + t)' (-w + U\hat{\beta} + \hat{\theta}) \\ & + \lambda \left(\sum_{i \in \zeta(\hat{\theta})} |t_i| + \sum_{i \in \bar{\zeta}(\hat{\theta})} \text{sgn}(\hat{\theta}_i) t_i \right) \end{aligned} \quad (5.13)$$

eşitliği elde edilir. (5.13) eşitliğinde $\alpha \rightarrow 0^+$ için limit alırsak,

$$\begin{aligned} l'(\hat{\beta}, \hat{\theta}; q, t) = & (Uq + t)' (-w + U\hat{\beta} + \hat{\theta}) + \lambda \sum_{i \in \zeta(\hat{\theta})} |t_i| + \lambda \sum_{i \in \bar{\zeta}(\hat{\theta})} \text{sgn}(\hat{\theta}_i) t_i \\ = & q' U' (-w + U\hat{\beta} + \hat{\theta}) + \sum_{i \in \zeta(\hat{\theta})} [-w_i + u_i' \hat{\beta} + \lambda \text{sgn}(t_i)] t_i \\ & + \sum_{i \in \bar{\zeta}(\hat{\theta})} [-w_i + u_i' \hat{\beta} + \hat{\theta}_i + \lambda \text{sgn}(\hat{\theta}_i)] t_i. \end{aligned} \quad (5.14)$$

elde edilir. Böylece Lemma 5.1'i kullanarak ancak ve ancak (5.9) ve (5.10) eşitlikleri sağlanırsa $\forall q \in \mathbb{R}^p, t \in \mathbb{R}^n$ için $l'(\hat{\beta}, \hat{\theta}; q, t) \geq 0$ eşitliğinin doğru olduğunu görüyoruz.

Yukarıda Lemma 5.1'i kullanarak $l(\beta, \theta)$ 'yı minimum yapan $\{\hat{\beta}, \hat{\theta}\}$ için gerekli koşulların (5.9) ve (5.10) eşitlikleri olduğu gösterilmiş oldu. Chang ve Qu (2004) çalışmalarında Lemma 5.1'i kullanmaksızın $l(\beta, \theta)$ 'yı minimum yapan $\{\hat{\beta}, \hat{\theta}\}$ için gerekli koşulların (5.9) ve (5.10) eşitlikleri olduğunun kolayca gösterilebileceğine değinmişlerdir. Gerçekte θ bilindiğinde doğrusal en küçük kareler tahmini için ancak ve ancak $U'(w - U\hat{\beta} - \theta) = 0$ ise $\hat{\beta} = (U'U)^{-1}U'(w - \theta)$, $l(\beta, \theta)$ 'nın minimumudur. Diğer yandan β biliniyorsa, parametrik olmayan tahmin yöntemi (5.4)'den dolayı ancak ve ancak $\hat{\theta} = \text{sgn}(w - U\beta)(|w - U\beta| - \lambda e)_+$ ise $\hat{\theta}$, $l(\beta, \theta)$ 'nın minimumudur.

(5.9) ve (5.10) eşitliklerinin yapısından dolayı çözümleri bulmak için iteratif bir algoritmanın kullanabileceği kolayca görülür. Chang ve Qu (2004) aşağıdaki algoritmayı önermişlerdir:

- θ için başlangıç değeri $\theta^0 = 0$ alınır.

- $k = 1, 2, \dots$ yakınsak oluncaya kadar,

$$\beta^k = (U'U)^{-1}U'(w - \theta^{k-1}) \quad (\theta^{k-1} \text{ verildiğinde alışılmış EKK tahmin edicisi}) \quad (5.15)$$

$$\theta^k = \text{sgn}(w - U\beta^k) \left(|w - U\beta^k| - \lambda e \right)_+ \quad (5.16)$$

Yukarıda verilen algoritma Hastie ve Tibshirani (1990) tarafından toplamsal modeli uydurmak için kullanılan yapıya benzediğinden dolayı backfitting algoritması olarak adlandırılır. Chang ve Qu (2004) çalışmalarında backfitting algoritmasına dayalı ve backfitting algoritmasından daha etkili, iteratif bir algoritma (exact line search algorithm) önermişlerdir. Bu algoritma için her bir iterasyon adımında algoritma tarafından kullanılacak olan asıl yön q 'yu (descent directional) bulmak için ilk olarak backfitting algoritmasını analiz etmişler ve daha sonra algoritmayı (exact line search algorithm) belirlemişler ve algoritmada kullanılacak olan basamak uzunluğunun (α) etkili olarak nasıl hesaplanacağını göstermişlerdir.

Aşağıda backfitting algoritmasına yakından bakalım:

Yarı parametrik modelde parametrik olmayan bileşen bulunmadığında ($f = 0$) β 'daki dönüştürülmüş rezidü,

$$r = w - U\beta$$

olarak tanımlanır. Buradan β^k 'daki dönüştürülmüş rezidüyü,

$$r^k = w - U\beta^k \quad (5.17)$$

biçiminde gösterebiliriz.

$$\xi(r) = \{i : |r_i| \leq \lambda\} \text{ ve } \bar{\xi}(r) = \{i : |r_i| > \lambda\}$$

indeks kümeleri tanımlansın. (5.16)'dan θ^k 'nın i . elemanı,

$$\begin{aligned}
\theta_i^k &= \text{sgn}(w - u_i' \beta^k) \left(|w - u_i' \beta^k| - \lambda \right)_+ \\
&= \text{sgn}(r_i^k) \left(|r_i^k| - \lambda \right)_+ \\
&= \begin{cases} 0, & i \in \xi(r^k) \\ r_i^k - \lambda \text{sgn}(r_i^k), & i \in \bar{\xi}(r^k) \end{cases} \quad i = 1, \dots, n
\end{aligned} \tag{5.18}$$

biçimindedir. Böylece,

$$\theta^k = r^k - z^k \quad \text{ve} \quad z_i^k = \begin{cases} r_i^k, & i \in \xi(r^k) \\ \lambda \text{sgn}(r_i^k), & i \in \bar{\xi}(r^k) \end{cases} \quad i = 1, \dots, n \tag{5.19}$$

yazılabilir. (5.15) iterasyon formülü, (5.17) ve (5.19) eşitliklerinden,

$$\begin{aligned}
\beta^{k+1} &= (U'U)^{-1} U'(w - r^k + z^k) \\
&= (U'U)^{-1} U'(w - r^k) + (U'U)^{-1} U'z^k \\
&= (U'U)^{-1} U'U\beta^k + (U'U)^{-1} U'z^k \\
&= \beta^k + (U'U)^{-1} U'z^k
\end{aligned} \tag{5.20}$$

elde edilir. Böylece,

$$q^k = (U'U)^{-1} U'z^k \tag{5.21}$$

k . adımda β ile ilgili bir arama yönü olarak tanımlanır. Chang ve Qu (2004), $q^k \neq 0$ ise q^k 'nin asıl yön olduğunu göstermişlerdir. (5.20)'deki birim basamak uzunluğu ($\alpha = I$) en uygun basamak uzunluğu olmayabilir veya muhtemelen $l(\beta^{k+1}, \theta^{k+1})l(\beta^k, \theta^k)$ 'dir. Böylece (5.16), (5.20) ve (5.21)'den,

$$\beta^{k+1} \equiv \beta^{k+1}(\alpha) = \beta^k + \alpha q^k, \tag{5.22}$$

$$\theta^{k+1} \equiv \theta^{k+1}(\alpha) = \text{sgn}(w - U\beta^{k+1}(\alpha)) \left(|w - U\beta^{k+1}(\alpha)| - \lambda e \right)_+ \tag{5.23}$$

elde edilir. Dikkat edersek, θ^{k+1} 'nin bu son tanımı (5.16) ile tutarlıdır. $\theta^{k+1}, \beta^{k+1}$ tarafından belirlendiğinden θ 'ya karşılık gelen yönü bilmemize gerek yoktur. Eğer (5.22) ve (5.23) denklemlerinde $q^k = 0$ ise bu durumda $\{\beta^k, \theta^k\}$, $l(\beta, \theta)$ 'nin minimumudur. Gerçekte her bir iterasyon adımında (5.17), (5.19) ve (5.21)'den,

$$\begin{aligned} q^k &= (U'U)^{-1}U'z^k = (U'U)^{-1}U'(r^k - \theta^k) \\ &= (U'U)^{-1}U'(w - U\beta^k - \theta^k) = 0 \end{aligned}$$

elde edilir. Böylece, $\{\beta^k, \theta^k\}$ (5.9)'daki şartı sağlar ve aynı zamanda (5.10)'daki koşulu (aynı zamanda (5.23)'ü) sağladığından $l(\beta, \theta)$ 'nin minimumu olmak zorundadır. $q^k \neq 0$ varsayımı altında q^k 'nin asıl yön olduğu aşağıdaki gibi gösterilir:

$$h^k = -Uq^k \quad (5.24)$$

olarak tanımlansın. Böylece (5.22)'deki β 'nin güncellenmesinden r 'nin güncellemesi,

$$r^{k+1}(\alpha) = w - U\beta^{k+1}(\alpha) = w - U(\beta^k + \alpha q^k) = w - U\beta^k - U\alpha q^k = r^k + \alpha h^k,$$

olarak elde edilir ve (5.23)'den,

$$\theta^{k+1}(\alpha) = \text{sgn}(r^k + \alpha h^k) \left(|r^k + \alpha h^k| - \lambda e \right)_+$$

bulunur. (5.19)'dan $z^{k+1}(\alpha) = r^{k+1}(\alpha) - \theta^{k+1}(\alpha)$ yazılabilir ve böylece,

$$\begin{aligned}
J(\alpha) &\equiv l(\beta^{k+1}(\alpha), \theta^{k+1}(\alpha)) = \frac{1}{2} \|z^{k+1}(\alpha)\|_2^2 + \lambda \|\theta^{k+1}(\alpha)\|_1 \\
&= \sum_{i=1}^n \frac{1}{2} (z_i^{k+1}(\alpha))^2 + \lambda |\theta_i^{k+1}(\alpha)|
\end{aligned}$$

elde edilir. Basit hesaplamalardan sonra aşağıdaki eşitlik yazılabilir:

$$J_i(\alpha) \equiv \frac{1}{2} (z_i^{k+1}(\alpha))^2 + \lambda |\theta_i^{k+1}(\alpha)| = \begin{cases} \frac{1}{2} (r_i^k + \alpha h_i^k)^2, & i \in \xi(r^k + \alpha h^k), \\ \frac{1}{2} (2\lambda |r_i^k + \alpha h_i^k| - \lambda^2), & i \in \bar{\xi}(r^k + \alpha h^k) \end{cases}$$

ve buradan $J_i(\alpha)$ 'nin türevi,

$$J'_i(\alpha) = \begin{cases} h_i^k (r_i^k + \alpha h_i^k), & i \in \xi(r^k + \alpha h^k), \\ h_i^k \lambda \operatorname{sgn}(r_i^k + \alpha h_i^k), & i \in \bar{\xi}(r^k + \alpha h^k) \end{cases} \quad (5.25)$$

biçiminde olup, (5.19), (5.21) ve (5.24)'den,

$$\begin{aligned}
J'(0) &= \sum_{i=1}^n J'_i(0) = \sum_{i \in \xi(r^k)} h_i^k r_i^k + \sum_{i \in \bar{\xi}(r^k)} h_i^k \lambda \operatorname{sgn}(r_i^k) \\
&= (h^k)' z^k = -(q^k)' U' z^k = -(z^k)' U (U' U)^{-1} U' z^k
\end{aligned}$$

elde edilir. Böylece,

$$q^k = (U' U)^{-1} U' z^k \neq 0$$

olduğundan $J'(0) \neq 0$ olmalıdır. Bu durumda q^k bir asıl yöndür. Chang ve Qu (2004) algoritmada kullanılacak olan ve $J(\alpha)$ 'yı minimum yapan α değerinin tahmini için aşağıdaki yolu izlemişlerdir:

$J(\alpha)$, α 'nın parçalı karesel, konveks, sürekli diferansiyellenebilir bir fonksiyonu olduğundan algoritmada kullanılacak olan minimum $\hat{\alpha}$ değeri $J'(\alpha) = 0$ eşitliğini sağlayan değerdir. $q^k \neq 0$ ve U tam sütun ranklı olduğundan $h^k = -Uq^k \neq 0$ olur. (5.25)'den $J'(\alpha)$ 'nın α 'nın azalmayan bir fonksiyonu olduğu ve $\lim_{n \rightarrow \infty} J(\alpha) = \infty$ olduğu görülmektedir. $J'(0) < 0$ olduğundan $\hat{\alpha}$ sonlu ve pozitif olmak zorundadır. En uygun $\hat{\alpha}$ değerini bulmak için,

$$A = \{\alpha : |r_i^k + \alpha h_i^k| = \lambda, \quad \alpha > 0, \quad h_i^k \neq 0, \quad i = 1, \dots, n\} \quad (5.26)$$

kümesi tanımlansın. A 'nın $0 < \alpha_1 < \alpha_2 < \dots < \alpha_m$ olacak biçimde elemanlara sahip olduğu varsayılın ve $\alpha_0 = 0$ olarak tanımlansın. Eğer j , $J'(\alpha_{j+1}) \geq 0$ olacak biçimde en küçük indeks ise bu durumda $\hat{\alpha} \in (\alpha_j, \alpha_{j+1}]$ aralığında veya $J'(\alpha_m) < 0$ ise $\hat{\alpha} \in (\alpha_m, \infty)$ aralığında olmalıdır. $j = 1$ 'den $J'(\alpha_j) \geq 0$ olacak biçimde ilk j 'ye kadar $J'(\alpha_j)$ etkili bir biçimde hesaplamak istensin. (5.25)'den,

$$J'(\alpha) = \sum_{i \in \xi(r^k + \alpha h^k)} h_i^k r_i^k + \alpha \sum_{i \in \xi(r^k + \alpha h^k)} (h_i^k)^2 + \lambda \sum_{i \in \bar{\xi}(r^k + \alpha h^k)} h_i^k \operatorname{sgn}(r_i^k + \alpha h_i^k) \quad (5.27)$$

elde edilir. $\alpha \in (\alpha_j, \alpha_{j+1})$ olduğunda $\xi(r^k + \alpha h^k)$ ve $\bar{\xi}(r^k + \alpha h^k)$ kümeleri yerine sırası ile ξ_j ve $\bar{\xi}_j$ gösterimleri kullanılabilir. (5.27)'den,

$$J'(\alpha) = c_j + \alpha d_j, \quad \alpha \in (\alpha_j, \alpha_{j+1})$$

biçiminde yazılabilir. Burada,

$$c_j = \sum_{i \in \xi_j(r^k + \alpha h^k)} h_i^k r_i^k + \lambda \sum_{i \in \xi_j(r^k + \alpha h^k)} (h_i^k)^2 \quad \text{ve} \quad d_j = \sum_{i \in \xi_j(r^k + \alpha h^k)} h_i^k \operatorname{sgn}(r_i^k + \alpha h_i^k) \quad (5.28)$$

biçimindedir ve $\text{sgn}(r_i^k + \alpha h_i^k) = \text{sgn}(r_i^k + \alpha_j h_i^k)$, $i \in \bar{\xi}_j$ eşitliği kullanılmıştır. c_j ve d_j sabit olduğundan $J'(\alpha)$, (α_j, α_{j+1}) aralığında doğrusal bir fonksiyondur ve $J'(\alpha)$ herhangi bir α 'da süreklidir. Böylece $J'(\alpha_{j+1}) = c_j + \alpha_{j+1} d_j$ ile $J'(\alpha_{j+1})$ hesaplanabilir. j , $J'(\alpha_{j+1}) \geq 0$ olacak şekilde en küçük indeks ise bu durumda en uygun $\hat{\alpha}$ değeri,

$$\hat{\alpha} = -c_j / d_j$$

eşitliğini sağlar. Aksi durumda $J'(\alpha_m) < 0$ ve bu durumda,

$$\hat{\alpha} = -c_m / d_m$$

olarak elde edilir. Böylece iteratif algoritma (exact line search algorithm) y , X ve tolerans δ değeri verildiğinde aşağıdaki biçimde verilir:

Adım 1 y 'nin ayrık dalgacık dönüşümü ADD , $X : w = Wy$ ve $U = WX$ hesaplanır.

Adım 2 $\beta^1 = (U'U)^{-1}U'w$ hesaplanır.

$k = 1, 2, \dots$ için,

Adım 3 $q^k = (U'U)^{-1}U'z^k$ hesaplanır. Burada z^k (5.19) eşitliği ile hesaplanır.

Adım 4 $\hat{\alpha} = \arg \min_{\alpha} J(\alpha)$ eşitliğini sağlayan en uygun $\hat{\alpha}$ hesaplanır.

Adım 5 $\beta^{k+1} = \beta^k + \hat{\alpha} q^k$ alınır,

$$\frac{\|\beta^{k+1} - \beta^k\|_2}{\|\beta^k\|_2} \leq \delta \text{ ise } \hat{\beta} = \beta^{k+1} \text{ alınır.}$$

Adım 6 $\hat{\theta} = \text{sgn}(w - U\hat{\beta}) \left(|w - U\hat{\beta}| - \lambda \right)_+$ ve ters ayrık dalgacık dönüşümü $\hat{f} = W'\hat{\theta}$ hesaplanır ve böylece algoritma bitirilir.

6. YARI PARAMETRİK MODELDE FARK ALMA YÖNTEMİNE DAYALI TAHMİNLER

Önceki bölümlerde parametrik olmayan ve yarı parametrik regresyon modellerinde parametre tahmini için düzeltme fikri ele alınmıştır. Bu bölümde ise (3.2) modelinin tahmini için düzeltme yaklaşımına alternatif olarak, f fonksiyonundan kaynaklanan verideki eğilimi elimine etmeyi amaçlayan fark alma yaklaşımı ele alınacaktır. Bu yaklaşımda $f(.)$ 'nin türevlenebilir ve x değerlerinin birbirine yakın olması ile (3.2) modelindeki f fonksiyonunun elimine edilmesi sağlanır. Parametrik olmayan regresyon modellerindeki parametrik olmayan etkinin elimine edilmesini sağlayan fark alma fikri kısmi doğrusal modelde Powell (1987), Ahn ve Powell (1993) tarafından ele alınmıştır. Hall ve ark.'larının (1990) sonuçlarının kullanılması ile Yatchew (1997), kısmi doğrusal modelin bir basit fark alma tahmin edicisini elde etmiştir. Bu yöntemin cazip bir özelliği, başlangıç parametrik olmayan tahmin yöntemleri için gerekenleri elimine etmesidir. Parametrik olmayan regresyon konusunda fark alma düşüncesi rezidü varyansının tahmin edicilerini elde etmek için kullanılan uzun bir geçmişe sahiptir.

Bu bölümde yarı parametrik modelde parametre tahmini için farklı bir yaklaşım olan farka dayalı tahmin ediciyi açıklamak için, ilk olarak temel kavramlar verilmiş ve daha sonra yarı parametrik modelde parametre tahmini için farka dayalı tahmin yöntemi ele alınmıştır. Ayrıca (3.2) yarı parametrik modelinde çoklu iç ilişkinin varlığı durumunda kullanılabilecek ve farka dayalı ridge tahmin edici olarak adlandırılan yeni bir tahmin edici önerilmiştir. Farka dayalı ridge tahmin edici $\hat{\beta}_{Rdiff}(k)$ ile farka dayalı tahmin edici $\hat{\beta}_{diff}$ hata kareler ortalaması ölçütüne göre karşılaştırılmıştır ve k ridge parametresinin küçük değerleri için tasarım noktalarına ve bilinmeyen regresyon parametrelerine bağlı olarak belirli bir aralık üzerinde $\hat{\beta}_{Rdiff}(k)$ 'nin $\hat{\beta}_{diff}$ üzerinde MSE anlamında üstün olduğu gösterilmiştir. Parametrik olmayan çalışmalarda yaygın olarak kullanılan bir veri seti alınarak bulunan sonuçların geçerliliği gösterilmiştir.

Giriş kısmında da belirtildiği gibi (3.2) modelinde parametre tahmini için genel olarak üç farklı yaklaşım söz konusudur. Bu yaklaşımlardan ilki rezidü tabanlı tahmin ediciler, ikincisi dalgacık tabanlı tahmin yaklaşımı ve üçüncü yaklaşım ise bu bölümde ele alınan ve f fonksiyonundan kaynaklanan verideki eğilimi elimine etmeyi amaçlayan fark alma yaklaşımıdır. Bu yaklaşımda $f(.)$ 'nin türevlenebilir ve x ordinatlarının birbirine yakın olması ile (3.2) modelindeki f fonksiyonunun elimine edilmesi sağlanır.

6.1. Fark Matrisi ve Özellikleri

m fark alma derecesi ve $d_0, d_1, \dots, d_m$ aşağıdaki koşulları sağlayan fark alma ağırlıkları olsun.

$$\sum_{j=0}^m d_j = 0, \quad \sum_{j=0}^m d_j^2 = 1 \quad (6.1)$$

Aşağıda (6.2) ile verilen matrise '*fark matrisi*' denir.

$$D_{n \times n} = \begin{pmatrix} d_0, d_1, d_2, \dots, d_m, 0, \dots, 0 \\ 0, d_0, d_1, d_2, \dots, d_m, 0, \dots, 0 \\ \vdots \\ 0, \dots, 0, d_0, d_1, d_2, \dots, d_m, 0 \\ 0, \dots, 0, \dots, 0, d_0, d_1, d_2, \dots, d_m \\ \vdots \\ 0, \dots, 0, \dots, 0 \end{pmatrix} \quad (6.2)$$

(6.2)'de verilen matriste son m satıra D matrisi kare bir matris olacak şekilde 0 yerleştirilmiştir. Bu matris gecikme (lag) matris L_i nin kullanımı için uygundur. $i > 0$ için L_i matrisi ana köşegen altındaki i . köşegende 1'lere diğer her yerde 0'lara sahip bir kare matristir. Eğer $i < 0$ ise L_i matrisi ana köşegen üstündeki i . köşegende 1'lere sahip bir kare matristir. L_0 matrisi bilinen birim matris olarak tanımlanır ve $L'_i = L_{-i}$

6. YARI PARAMETRİK MODELDE FARK ALMA YÖNTEMİNE

DAYALI TAHMİNLER VE İLGİLİ ÇIKARIMLAR

Gülin TABAKAN

ve $L_i L_j \doteq L_{i+j}$ dir. m . dereceden fark matrisinin gecikme matrislerin bir ağırlıklandırılmış toplamı olduğu (6.3)'den açıkça görülebilir. Yani,

$$D_{n \times n} \doteq d_0 L_0 + d_1 L'_1 + \dots + d_m L'_m \quad (6.3)$$

dir. Ana köşegeninde 1 elemanlarına ve simetrik bir bant yapısına sahip olan $D'D$ matrisini ele alalım. Bu durumda $tr(D'D) = n$ olur. Ana köşegeni çıkartırsak ardışık köşegenler, $\sum_{j=0}^{m-k} d_j d_{j+k}$, $k = 1, \dots, m$ biçiminde olup matrisin geri kalan elemanları 0'dır. Denk olarak (6.3)'ü ve gecikme matrislerin özelliklerini kullanarak,

$$\begin{aligned} D'D \doteq L_0 + (L_1 + L'_1) \sum_{j=0}^{m-1} d_j d_{j+1} + (L_2 + L'_2) \sum_{j=0}^{m-2} d_j d_{j+2} + \dots \\ + (L_{m-1} + L'_{m-1}) \sum_{j=0}^1 d_j d_{j+m-1} + (L_m + L'_m) d_0 d_m \end{aligned} \quad (6.4)$$

elde edebiliriz. Bant yapısı matris çarpımı ile korunduğundan, $D'DD'D$ matrisi de hem simetrik hem de bant yapısına sahip bir matristir. Bu matriste ana köşegen üzerindeki değer (6.4)'ü kendisi ile çarparak elde edilebilir. Yani,

$$[D'DD'D]_{ii} \doteq 1 + 2 \sum_{k=1}^m \left(\sum_{j=0}^{m-k} d_j d_{j+k} \right)^2 \quad (6.5)$$

olacaktır. (6.5) eşitliğinde

$$\delta = \sum_{k=1}^m \left(\sum_{j=0}^{m-k} d_j d_{j+k} \right)^2 \quad (6.6)$$

olarak alınırsa bu durumda,

$$[D'DD'D]_{ii} \doteq 1 + 2\delta \text{ ve } tr(D'DD'D) \doteq n(1 + 2\delta)$$

olarak yazılabilir (Yatchew, 2003).

6.2. Yarı Parametrik Modelin Farka Dayalı Tahmini

(3.2) modeli için bağımsız ve özdeş dağılan $(y_1, z_1, x_1), \dots, (y_n, z_n, x_n)$ gözlemlerinin verildiğini varsayalım. (3.2) modelinde $E(y|Z, x) = Z\beta + f(x)$ ve $\sigma_\varepsilon^2 = Var[y|Z, x]$ 'dir. Kolaylık olsun diye (3.2)'de bütün değişkenlerin skaler olduğunu varsayacağız. Ayrıca x 'ler birim aralık olarak bilinen bir kapalı aralıkta sınırlandırılmıştır ve gözlemler $x_1 \leq x_2 \leq \dots \leq x_n$ olacak şekilde yeniden düzenlenmiştir. n örneklemdaki gözlemlerin sayısı olup f hakkında bütün bilinen birinci türevinin sabit bir L sayısı ile sınırlı olduğudur $(|f(x_i) - f(x_{i-1})| \leq L|x_i - x_{i-1}|)$. y yanıt değişkeninin Z ile doğrusal fakat x ile doğrusal olmayan bir ilişki içerisinde bulunduğu varsayıldığında bu modellerin kullanımı uygundur. Burada Z 'nin koşullu ortalamasının x 'in düzgün bir fonksiyonu yani $E(Z|x) = f(x)$ olduğu varsayılacak ve x verildiğinde Z 'nin beklenen koşullu varyansı $var(Z|x) = \sigma_u^2$ ile gösterilecektir. Bu durumda $Z = g(x) + u$ yazabiliriz. Böylece model (3.1)'de verilere dayalı ilk farklar çıkarılırsa,

$$\begin{aligned} y_i - y_{i-1} &= (z_i - z_{i-1})\beta + (f(x_i) - f(x_{i-1})) + \varepsilon_i - \varepsilon_{i-1} \quad i = 2, \dots, n \\ &= (g(x_i) - g(x_{i-1}))\beta + (u_i - u_{i-1})\beta + (f(x_i) - f(x_{i-1})) + \varepsilon_i - \varepsilon_{i-1} \\ y_i - y_{i-1} &\cong (u_i - u_{i-1})\beta + \varepsilon_i - \varepsilon_{i-1} \end{aligned} \quad (6.7)$$

elde edilir. Böylece (6.7)'den de görüleceği gibi Z tarafından meydana gelen direk olmayan etki $g(x)$ ve parametrik olmayan x değişkeninin direk etkisi $f(x)$ çıkarılmış olur. Z bağımsız (açıklayıcı) değişkeni x ile tamamen ilişkili olmadığı

müddetçe, farkı alınmış verilere alışılmış en küçük kareler tahmin edicisi uygulanarak,

$$\hat{\beta}_{diff} = \frac{\sum_{i=2}^n (y_i - y_{i-1})(z_i - z_{i-1})}{\sum_{i=2}^n (z_i - z_{i-1})^2} \quad (6.8)$$

elde edilir. Böylece, $z_i - z_{i-1} \cong u_i - u_{i-1}$ ve $y_i - y_{i-1} \cong (u_i - u_{i-1})\beta + \varepsilon_i - \varepsilon_{i-1}$ yaklaşımları (6.8)'de yerine yazılırsa,

$$\begin{aligned} n^{1/2}(\hat{\beta}_{diff} - \beta) &= n^{1/2} \left[\frac{1}{n} \sum_{i=2}^n (f(x_i) - f(x_{i-1}) + \varepsilon_i - \varepsilon_{i-1}) \right. \\ &\quad \cdot (g(x_i) - g(x_{i-1}) + u_i - u_{i-1}) \Big] \\ &\quad / \frac{1}{n} \sum_{i=2}^n (g(x_i) - g(x_{i-1}) + u_i - u_{i-1})^2 \end{aligned} \quad (6.9)$$

elde edilir. Kolaylık olsun diye x_i 'lerin birim aralıkta bulunduğu durum göz önüne alınmış ve f ile g 'nin birinci türevlerinin sabit bir L sayısı ile sınırlı olduğu varsayılmıştır. Bu durumda,

$$\begin{aligned} n^{1/2} \left| \frac{1}{n} \sum_{i=2}^n (f(x_i) - f(x_{i-1}))(g(x_i) - g(x_{i-1})) \right| \\ \leq \frac{n^{1/2}}{n} \sum L^2 |x_i - x_{i-1}|^2 = L^2 \frac{n^{1/2}}{n} n \frac{1}{n^2} = \frac{L^2}{n^{3/2}} \end{aligned} \quad (6.10)$$

ve varyans,

$$Var \left[\frac{n^{1/2}}{n} \sum u_i (f(x_i) - f(x_{i-1})) \right] \quad (6.11)$$

$$\leq \frac{\sigma_u^2}{n} \sum (f(x_i) - f(x_{i-1}))^2 \leq \frac{L^2 \sigma_u^2}{n^2}$$

olarak elde edilir. Böylece (6.10) ve (6.11) sıfıra yakınsar. Benzer işlemler kullanılarak (6.9)'da $(f(x_i) - f(x_{i-1}))$ ve $(g(x_i) - g(x_{i-1}))$ 'i içeren diğer terimlerin yeterince çabuk sıfıra yakınsadığı gösterilebilir. Bu durumda (6.9) yaklaşık olarak,

$$n^{1/2}(\hat{\beta}_{diff} - \beta) \cong \frac{n^{1/2} \frac{1}{n} \sum_{i=2}^n (\varepsilon_i - \varepsilon_{i-1})(u_i - u_{i-1})}{\frac{1}{n} \sum_{i=2}^n (u_i - u_{i-1})^2} \quad (6.12)$$

biçiminde yazılabilir. (6.12) eşitliğinde payda $2\sigma_u^2$ 'ye yakınsar ve pay 0 ortalamalı ve

$6\sigma_\varepsilon^2 \sigma_u^2$ varyanslıdır. Böylece (6.12) eşitliği $\frac{6\sigma_\varepsilon^2 \sigma_u^2}{(2\sigma_u^2)^2} = 1,5\sigma_\varepsilon^2 / \sigma_u^2$ varyansına sahiptir:

$$n^{1/2}(\hat{\beta}_{diff} - \beta) \rightarrow N\left(0, \frac{1,5\sigma_\varepsilon^2}{\sigma_u^2}\right) \quad (6.13)$$

En etkin tahmin edici için (6.13)'teki varyansın karşılığı $\sigma_\varepsilon^2 / \sigma_u^2$ olacaktır, böylece ilk farklara dayalı önerilen tahmin edici $\frac{2}{3} = 1/1.5$ oranında etkinliğe sahiptir. Farka dayalı bu yöntem uygulanması basit olan bir yöntem olması ile birlikte, daha önce bahsedilen kernel tabanlı yöntemler ve parametrik olmayan EKK yöntemlerine göre daha az etkindir. Etkinlik temelde yüksek dereceden farkların kullanılması ile geliştirilebilir (Yatchew, 1997).

Yatchew (1997) çalışmasında daha etkili tahmin ediciler elde etmek için yüksek dereceden farka dayalı tahmin edicileri ele almıştır. Yatchew (1997) fark alma derecesi m 1'den 2'ye, 3'e artarken etkinliğin sırası ile %66.7, %80, %85.7'ye artacağını ve m fark alma derecesinin yeterince büyük seçilmesi ile, tahmin edicinin asimptotik etkinliğe ulaşacağını belirtmiştir.

m . dereceden fark alma için (6.7)'yi genelleştirerek aşağıdaki gibi yazabiliriz:

$$\sum_{j=0}^m d_j y_{i-j} = \left(\sum_{j=0}^m d_j z_{i-j} \right) \beta + \sum_{j=0}^m d_j f(x_{i-j}) + \sum_{j=0}^m d_j \varepsilon_{i-j}, \quad i = m+1, \dots, n \quad (6.14)$$

Ağırlıkların en uygun seçimi ile (6.14) EKK ile tahmin edilebilir ve böylece (6.8)'in genelleştirilmiş halini aşağıdaki biçimde yazabiliriz:

$$\hat{\beta}_{diff} = \frac{\sum_i \left(\sum_{j=0}^m d_j y_{i-j} \right) \left(\sum_{j=0}^m d_j z_{i-j} \right)}{\sum_i \left(\sum_{j=0}^m d_j z_{i-j} \right)^2}, \quad (6.15)$$

(6.15)'de m fark alma derecesi ve $d_0, \dots, d_m$ (6.1)'deki koşulları sağlayan ağırlıklardır. (6.1)'deki ilk koşul x 'ler yakın olduğunda (3.2) modelindeki parametrik olmayan etkinin modelden çıkarılmasını sağlar, ikinci koşul ise (3.2)'deki rezidülerin σ_ε^2 varyansına sahip olmasını ifade eder. $d_0, \dots, d_m$ ağırlıkları en uygun olarak seçilirse,

$$\hat{\beta}_{diff} \sim N \left(\beta, \frac{1}{n} \left(1 + \frac{1}{2m} \right) \frac{\sigma_\varepsilon^2}{\sigma_u^2} \right)$$

olduğu gösterilebilir.

Çizelge 6.1'de $m = 10$ 'a kadar olan en uygun (optimal) fark alma ağırlıkları listelenmiştir (Yatchew, 2003).

Çizelge 6.1. En uygun fark alma ağırlıkları

m	$(d_0, d_1, \dots, d_m)$
1	(0.7071, -0.7071)
2	(0.8090, -0.5000, -0.3090)
3	(0.8582, -0.3832, -0.2809, -0.1942)
4	(0.8873, -0.3099, -0.2464, -0.1901, -0.1409)
5	(0.9064, -0.2600, -0.2167, -0.1774, -0.1420, -0.1103)
6	(0.9200, -0.2238, -0.1925, -0.1635, -0.1369, -0.1126, -0.0906)
7	(0.9302, -0.1965, -0.1728, -0.1506, -0.1299, -0.1107, -0.0930, 0.0768)
8	(0.9380, -0.1751, -0.1565, -0.1389, -0.1224, -0.1069, -0.0925, -0.0791, -0.0666)
9	(0.9443, -0.1578, -0.1429, -0.1287, -0.1152, -0.1025, -0.0905, -0.0792, -0.0687, -0.0588)
10	(0.9494, -0.1437, -0.1314, -0.1197, -0.1085, -0.0978, -0.0877, -0.0782, -0.0691, -0.0606, -0.0527)

m fark alma derecesini belirleyip, $d_0, \dots, d_m$ fark alma ağırlıklarını en uygun olarak seçerek ve (3.2) yarı parametrik regresyon modeline D fark matrisini uygulayarak,

$$Dy = DZ\beta + Df(x) + D\varepsilon \cong DZ\beta + D\varepsilon \quad (6.16)$$

elde ederiz (Yatchew, 1998). Burada Dy elemanları $(Dy)_i = \sum_{j=0}^m d_j y_{i-j}$ olan

$(n-m) \times 1$ tipinde bir vektör, DZ elemanları $(DZ)_i = \sum_{j=0}^m d_j z_{i-j}$ olan $(n-m) \times p$

tipinde bir matristir. Böylece, fark alma fikri (3.2) modelinde parametrik olmayan bileşen f yokmuş gibi β üzerinde çıkarımlar yapmamıza izin verir (Yatchew, 2003). Bu durumda $\hat{\beta}_{diff}$ tahmin edilirse, β biliniyormuş gibi f tahmini için çeşitli parametrik olmayan yöntemler kullanılabilir.

Yukarıda verilen (6.16) eşitliğinde parametrik olmayan etkinin modelden çıkarılmasının mümkün olduğu görülmektedir. Bunu daha iyi kavrayabilmek için

6. YARI PARAMETRİK MODELDE FARK ALMA YÖNTEMİNE

DAYALI TAHMİNLER VE İLGİLİ ÇIKARIMLAR

Gülin TABAKAN

burada D fark matrisinin uygulanması ile (6.17) eşitliğindeki gibi yazılabilen parametrik olmayan regresyon modeli ele alınarak parametrik olmayan etkinin modelden yaklaşık olarak nasıl çıkarıldığı gösterilmiştir:

$$Dy = Df(x) + D\varepsilon \quad (6.17)$$

modelini daha açık olarak aşağıdaki biçimde yazabiliriz:

$$\begin{pmatrix} d_0, d_1, 0, \dots, 0 \\ 0, d_0, d_1, 0, \dots, 0 \\ \vdots \\ 0, \dots, 0, d_0, d_1, 0 \\ 0, \dots, 0, d_0, d_1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} d_0, d_1, 0, \dots, 0 \\ 0, d_0, d_1, 0, \dots, 0 \\ \vdots \\ 0, \dots, 0, d_0, d_1, 0 \\ 0, \dots, 0, d_0, d_1 \end{pmatrix} \begin{pmatrix} f(x_1) \\ f(x_2) \\ \vdots \\ f(x_n) \end{pmatrix} + \begin{pmatrix} d_0, d_1, 0, \dots, 0 \\ 0, d_0, d_1, 0, \dots, 0 \\ \vdots \\ 0, \dots, 0, d_0, d_1, 0 \\ 0, \dots, 0, d_0, d_1 \end{pmatrix} \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}$$

$$\begin{aligned} d_0 y_1 + d_1 y_2 &= d_0 f(x_1) + d_1 f(x_2) + d_0 \varepsilon_1 + d_1 \varepsilon_2 \\ d_0 y_2 + d_1 y_3 &= d_0 f(x_2) + d_1 f(x_3) + d_0 \varepsilon_2 + d_1 \varepsilon_3 \\ d_0 y_3 + d_1 y_4 &= d_0 f(x_3) + d_1 f(x_4) + d_0 \varepsilon_3 + d_1 \varepsilon_4 \\ \vdots & \\ \vdots & \end{aligned} \quad (6.18)$$

$m=1$ için $\sum_{j=0}^1 d_j = 0$ ve $d_0 + d_1 = 0.7071 - 0.7071 = 0$ 'dir ve ortalama değer

teoreminden birim aralıkta sınırlı n gözlem için, $f(x_i) - f(x_{i-1}) = f'(c)(x_i - x_{i-1}) \leq L \cdot \frac{1}{n} \rightarrow 0$ olur. Buradan hareketle (6.18) denklem sistemi,

$$\begin{aligned} d_0 f(x_1) + d_1 f(x_2) &= -d_1 f(x_1) + d_1 f(x_2) = d_1 (f(x_2) - f(x_1)) \rightarrow 0 \\ d_0 f(x_2) + d_1 f(x_3) &= -d_1 f(x_2) + d_1 f(x_3) = d_1 (f(x_3) - f(x_2)) \rightarrow 0 \end{aligned}$$

$$d_0 f(x_3) + d_1 f(x_4) = -d_1 f(x_3) + d_1 f(x_4) = d_1 (f(x_4) - f(x_3)) \rightarrow 0$$

$$\begin{array}{cccccc} \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{array}$$

biçiminde olur. Böylece (6.18) denklem sistemi,

$$d_0 y_1 + d_1 y_2 = d_0 \varepsilon_1 + d_1 \varepsilon_2$$

$$d_0 y_2 + d_1 y_3 = d_0 \varepsilon_2 + d_1 \varepsilon_3$$

$$d_0 y_3 + d_1 y_4 = d_0 \varepsilon_3 + d_1 \varepsilon_4$$

$$\begin{array}{cccc} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{array}$$

veya

$$Dy \cong D \varepsilon \quad (6.19)$$

olarak yazılabilir. Aynı şekilde (6.16) ile verilen modelden parametrik olmayan etkinin çıkarılması gösterilebilir. (6.16) eşitliğine *EKK*'nın uygulanması ile β 'nın farka dayalı tahmin edicisi,

$$\hat{\beta}_{diff} = \left[(DZ)' (DZ) \right]^{-1} (DZ)' Dy \quad (6.20)$$

biçiminde tanımlanır.

6.3. Düzeltme ve Fark Alma Arasındaki İlişki

2. ve 3. bölümlerde ele alınan parametrik olmayan ve yarı parametrik regresyon modellerinde temel amaç parametrik olmayan regresyon fonksiyonunun iyi tahminlerini elde etmektir. Örneğin,

$$y = f(x) + \varepsilon \quad (6.21)$$

ile verilen parametrik olmayan regresyon modelini yeniden ele alalım. Bir S düzeltici matrisinin verilere uygulanması ile f fonksiyonuna uygun bir yaklaşım elde edilebilir:

$$Sy = Sf(x) + S\varepsilon \cong Sf \cong f \quad (6.22)$$

burada $S\varepsilon \cong 0$ 'dır çünkü düzeltilmiş rasgele hata terimi '0' a eşittir. Bu bölümde ele alınan fark alma yöntemindeki amaç daha önceden de bahsedildiği gibi parametrik olmayan etkiyi modelden çıkarmak ve parametrik olmayan modeli,

$$Dy = Df(x) + D\varepsilon \cong D\varepsilon \quad (6.23)$$

biçiminde yazmaktır. Düzeltme ve fark alma yöntemi arasındaki ilişkinin temeli budur. Bir düzeltme yöntemi parametrik olmayan etkinin çıkarılması için her zaman kullanılabilir. Örneğin, parametrik olmayan regresyon modelini kullanarak,

$$(I - S)y = (I - S)f(x) + (I - S)\varepsilon \cong (I - S)\varepsilon \quad (6.24)$$

yazılabilir. Burada $D = I - S$ 'yi bir fark alma yöntemi olarak düşünebiliriz. Diğer yandan fark alma ile genel olarak parametrik olmayan etkinin tam kullanışlı bir tahmini elde edilemez. Bunun nedeni örneklem büyüklüğünün artışı ile fark alma derecesine ihtiyaç olmamasıdır.

Bu bölümde (6.16) eşitliği ile gösterildiği gibi fark alma düşüncesi yarı parametrik modelin bir tahmin edicisi ile sonuçlanır:

$$\begin{aligned} Dy &= Df(x) + DZ\beta + D\varepsilon \\ &\cong DZ\beta + D\varepsilon, \end{aligned}$$

burada D fark matrisi x değerlerinin birbirine yakın olacak biçimde yeniden sıralanması ile verilere uygulanmıştır. Farkı alınmış modeldeki rezidüler $D\varepsilon$ 'dir ve

ancak fark alma derecesi büyük ise ε 'a yaklaşık olarak eşittir. Benzer şekilde düzeltme yöntemi için,

$$\begin{aligned}(I - S)y &= (I - S)f(x) + (I - S)Z\beta + (I - S)\varepsilon \\ &\cong (I - S)Z\beta + (I - S)\varepsilon \\ &\cong (I - S)Z\beta + \varepsilon,\end{aligned}\tag{6.25}$$

olarak yazılabilir.

(6.25) eşitliği aynı zamanda (3.24) ile verilen eşitliktir. Buradan da görüldüğü gibi fark alma yönteminin uygulanmasında, D fark matrisi tüm veriye uygulanır. Düzeltme yöntemi uygulandığında ise parametrik olmayan değişken x 'de hem Z 'nin her bir sütunu hem de y için ayrı ayrı parametrik olmayan tahminleri elde etmek gerekir (Yatchew, 2003).

6.4. Deneysel Uygulama

6.4.1. Kanada Hane Benzin Talebi

Yatchew ve No (2001) Kanada da ev halkının benzin talebini bir yarı parametrik model ile tahmin etmişler ve değişkenleri aşağıdaki gibi tanımlamışlardır:

Gidilen yol: log(Ev halkı tarafından her ay gidilen yolun uzaklığı);

Gelir: log(Ev halkının yıllık geliri);

Fiyat: log(Bir litre benzinin fiyatı);

E.sürücü: log(Ev halkı içindeki ehliyetli sürücülerin sayısı);

Kişi sayısı: log(Hanede bulunan kişi sayısı);

Genç bekâr: Hane içindeki 35 yaşını geçmeyen genç bekârların sayısı için yapay değişken;

Yaş: log(Yaşın logaritması);

Emekli: Hane içindeki 65 yaşını geçen emeklilerin sayısı için yapay değişken;

Şehirde yaşayanlar: Şehirde ikamet edenler için yapay değişkendir.

6. YARI PARAMETRİK MODELDE FARK ALMA YÖNTEMİNE

DAYALI TAHMİNLER VE İLGİLİ ÇIKARIMLAR

Gülin TABAKAN

Tanımlanan değişkenlere ait yarı parametrik regresyon modeli,

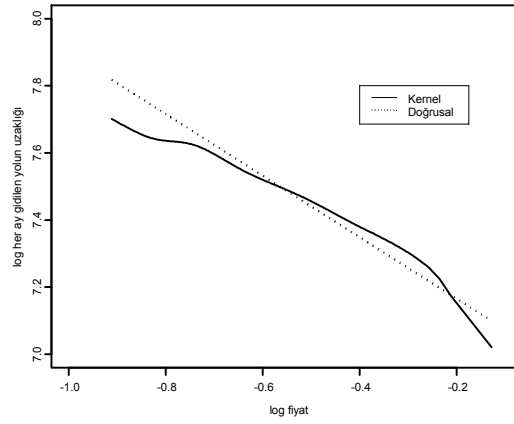
$$\begin{aligned} gidilenyol = & f(fiyat) + \beta_1(gelir) + \beta_2(e.sürücü) + \beta_3(kişisayı) + \beta_4(gençbekar) \\ & + \beta_5(yaş) + \beta_6(emekli) + \beta_7(şehirdeyaşayanlar) + \varepsilon \end{aligned} \quad (6.26)$$

biçimindedir.

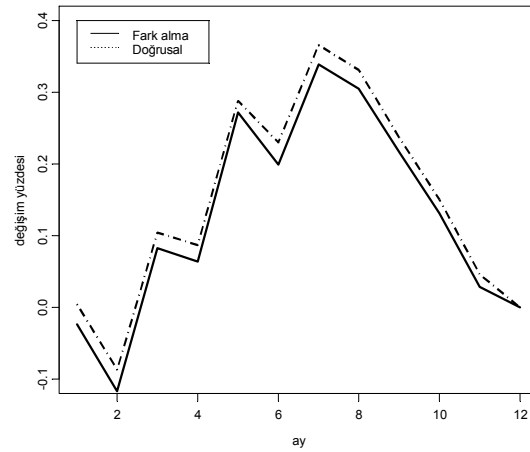
Çizelge 6.2 sonuçları özetlemektedir ve çizelgede verilen parametrik tahminler fiyatın log-doğrusal olarak girildiği bir modeli ifade etmektedir. Ayrıca Çizelge 6.2’de (3.2) yarı parametrik modelindeki parametre tahminleri fark alma yöntemine dayalı olarak elde edilmiştir.

Çizelge 6.2. Benzin talebi için parametre tahminleri

Değişkenler	Parametrik tahminler		3. dereceden farka dayalı tahminler		10. dereceden farka dayalı tahminler	
	Katsayı	SE	Katsayı	SE	Katsayı	SE
Fiyat	-0.9170	0.0960	-	-	-	-
Gelir	0.2890	0.0200	0.2802	0.0223	0.2816	0.0209
Sürücü sayısı	0.5610	0.0330	0.5593	0.0363	0.5686	0.0338
Kişi sayısı	0.1000	0.0260	0.1067	0.0297	0.0892	0.0274
Genç bekar	0.1930	0.0610	0.2219	0.0664	0.2099	0.0622
Yaş	-0.0780	0.0440	-0.1141	0.0781	-0.1171	0.0555
Emekli	-0.2070	0.0320	-0.2248	0.0536	-0.2113	0.0387
Şehirde yaşayanlar	-0.3310	0.0200	-0.3391	0.0217	-0.3331	0.0203
Aylık etkiler (Şekil 6.1)						
s_{ε}^2	.5003		.5067		.4997	
R^2	.2635		.2539		.2644	



Şekil 6.1. Benzin için hane talebi



Şekil 6.2. Benzin için hane talebi: Aylık etkiler

Bağımlı değişkenin varyansı 0.6794 ve gözlem sayısı $n = 6230$ 'dur. Rezidüler parametrik etkileri tahmin için kullanılmıştır. Farka dayalı tahminler için 3. ve 10. dereceden fark alma kullanılmıştır. Tahminlerin üçü de önemsiz derecede büyük olan standart hatalar dışında çok benzerdir. Fark alma tarafından tahmin edilen parametrik etkiler çıkarılmış ve fiyat etkisinin (Şekil 6.1 deki kesiksiz çizgi) parametrik olmayan bir tahminini elde etmek için kernel regresyon kullanılmıştır. Şekil 6.2 mevsimlik etkileri gösterir (Yatchew, 2003).

6.5. Yarı Parametrik Modelde Parametrelerin Farka Dayalı Ridge Tahmin Edicisi

İstatistikte çoklu iç ilişki problemi ve bir doğrusal regresyon modelindeki istatistiksel sonuçları çok iyi bilinir. Çoklu iç ilişki $y = Z\beta + \varepsilon$ doğrusal modelinde Z tasarım matrisinin sütun vektörleri arasında hemen hemen doğrusal bağımlılığın var olması olarak tanımlanır. Burada y , $n \times 1$ boyutlu açıklayıcı değişkenlerin bir vektörü, Z rankı p olan $n \times p$ boyutlu bağımsız değişkenlerin gözlem matrisi, β bilinmeyen bir parametre vektörü ve ε , $E(\varepsilon) = 0$ ve $E(\varepsilon\varepsilon') = \sigma^2 I$ olan hata vektörüdür. Çoklu iç ilişkinin varlığı regresyon katsayılarının en küçük kareler tahminlerinin büyük varyansa sahip olmalarına neden olur. Ridge regresyon modeli, tahminlerin varyanslarını azaltabilmek için regresyon denklemindeki katsayıları yanı olarak tahmin eder. Çoklu iç ilişkinin varlığını ölçmenin bir yolu koşul sayısına bakmaktır. Z matrisinin koşul sayısı çoklu iç ilişkinin varlığı hakkında bazı bilgiler verse bile $Z_1, Z_2, \dots, Z_p$ sütun vektörleri arasındaki doğrusal bağımlılığın yapısını anlatmaz. Çoklu iç ilişkinin varlığını ve yapısını anlatmanın en iyi yolu $Z'Z$ matrisinin öz değerlerine ve öz vektörlerine bakmaktır. $Z'Z$ büyük bir koşul sayısı ile kötü koşullu ise β parametresinin tahmini için ridge regresyon tahmin edici kullanılabilir. Çoklu iç ilişkinin varlığı bireysel parametreler (kararlı olmayan tahminler) için geniş güven aralıkları ile sonuçlanabilir, bir hipotez testinde kararımızda etkili olabilir ve bundan dolayı yanlış tahminler elde edilebilir. Şiddetli çoklu iç ilişki tahminleri çok kararsız yapabilir, bu durumda pratikte kullanışsız olurlar. Bu problemi aşmak için farklı çözüm yöntemleri önerilmiştir. Çoklu iç ilişki için çok yaygın olarak kullanılan sayısal bir yöntem ‘ridge regresyon’ yöntemidir (Hoerl ve Kennard (1970)). Alışılmış ridge regresyon (ORR) tahmin edici $\hat{\beta}_{or} = (Z'Z + kI)^{-1} Z'y$ yanlıdır fakat (ORR)’nin elemanlarının varyansı alışılmış en küçük karelerin (EKK) karşılık gelen elemanlarının varyansından daha küçüktür. Groß (2003)’te parametre vektöründe eklemeli doğrusal kısıtlamaların bulunduğu varsayıldığında bir doğrusal regresyon modelindeki parametre vektörü için bir ridge tahmin ediciyi tanıtmıştır. Yakın zamanda Hu (2005), yarı parametrik regresyon

modelinin ridge tahmin edicisini önermiş ve iki-adımlı tahmin yöntemine bağlı ridge tahmin yöntemini tanıtmıştır.

6.5.1. Önerilen Farka Dayalı Ridge Tahmin Edici

Farka dayalı ridge tahmin ediciyi açıklamak için (6.16) eşitliğini aşağıdaki biçimde yeniden yazabiliriz.

$$y^* = Z^* \beta + \varepsilon^*$$

Burada $y^* = Dy$, $Z^* = DZ$ ve $\varepsilon^* = D\varepsilon$ dir. (3.2) yarı parametrik modelinde farka dayalı tahmine bağlı olarak elde edilen ve farka dayalı ridge tahmin edici olarak adlandırdığımız tahmin edici için aşağıdaki adımlar izlenir:

Burada amaç k yanlılık parametresine sahip aşağıda verilen denklemi minimum yapan parametre değerini bulmaktır.

$$L^* = \left\{ (y^* - Z^* \beta)' (y^* - Z^* \beta) + k \beta' \beta \right\} \quad (6.27)$$

ile verilen denklemin β parametre vektörüne göre türevinin sıfıra eşitlenmesinden,

$$\frac{\partial L^*}{\partial \beta} = 0 \quad (6.28)$$

aşağıdaki eşitlik elde edilir:

$$-2Z^{*'} y^* + 2Z^{*'} Z^* \beta + 2k\beta = 0.$$

Böylece bu eşitliğin β 'ya göre çözümünden,

$$\hat{\beta}_{Rdiff}(k) = \left(Z^{*'} Z^* + kI \right)^{-1} Z^{*'} y^* \quad (k \geq 0)$$

$$\hat{\beta}_{Rdiff}(k) = S_k Z^{*'} y^* \quad (6.29)$$

elde edilir. Burada $S_k = \left(Z^{*'} Z^* + kI \right)^{-1}$, I $p \times p$ birim matris ve k araştırmacı tarafından seçilen sabit bir parametredir. Farklı araştırmacılar tarafından k parametresinin seçimi için farklı yöntemler önerilmiştir (Hoerl ve Kennard, 1970; Hoerl ve ark., 1975; Lawles, 1978; Kibria, 1996).

Doğrusal modeldeki ridge tahmin edici ve (6.29)'da verilen tahmin edici arasında biçimsel bir benzerlik olduğundan dolayı (6.29) denklemi ile verilen tahmin edici yarı parametrik modelin farka dayalı ridge tahmin edicisi olarak adlandırılır (Tabakan ve Akdeniz, 2008).

6.5.2. Farka Dayalı Tahmin Edici ve Farka Dayalı Ridge Tahmin Edicinin Karşılaştırılması

Bu alt bölümde $\hat{\beta}_{diff}$ ve $\hat{\beta}_{Rdiff}$ tahmin edicileri örneklem varyansı ölçütüne göre karşılaştırılmıştır. $S_k Z^{*'} Z^* = I - kS_k$ olduğundan $\hat{\beta}_{Rdiff}(k)$ 'nin beklenen değeri

$$E(\hat{\beta}_{Rdiff}(k)) = E(S_k Z^{*'} y^*) = \beta - kS_k \beta \quad (6.30)$$

ve yanlılığı,

$$bias(\hat{\beta}_{Rdiff}(k)) = -kS_k \beta. \quad (6.31)$$

biçimindedir. $\hat{\beta}_{Rdiff}(k)$ 'nin varyans-kovaryans matrisi aşağıdaki gibidir:

$$V(\hat{\beta}_{Rdiff}(k)) = V\left(S_k Z^{*'} y^*\right) = \left(S_k Z^{*'}\right) \text{var}(y^*) \left(S_k Z^{*'}\right)' \quad (6.32)$$

(6.32) denkleminde $S = (D'Z^*)'(D'Z^*)$ alınırsa,

$$M_1 = V(\hat{\beta}_{Rdiff}(k)) = \sigma_\varepsilon^2 S_k S_k' \quad (6.33)$$

olur. Benzer şekilde, $U = (Z^{*'} Z^*)^{-1}$ alınarak $V(\hat{\beta}_{diff})$ aşağıdaki gibi yazılabilir:

$$\begin{aligned} M_2 &= V(\hat{\beta}_{diff}) = V\left\{\left[(DZ)'(DZ)\right]^{-1} (DZ)' Dy\right\} \\ &= \sigma_\varepsilon^2 \left(Z^{*'} Z^*\right)^{-1} S \left(Z^{*'} Z^*\right)^{-1} \\ &= \sigma_\varepsilon^2 U S U \end{aligned}$$

burada $U = \left(Z^{*'} Z^*\right)^{-1}$ 'dir. Buradan,

$$\begin{aligned} M_2 &= \sigma^2 S_k \left(S_k^{-1} U S U S_k^{-1}\right) S_k' \\ M_2 &= \sigma^2 S_k \left\{ \left(Z^{*'} Z^* + kI\right) \left(Z^{*'} Z^*\right)^{-1} (D'Z^*)' (D'Z^*) \left(Z^{*'} Z^*\right)^{-1} \left(Z^{*'} Z^* + kI\right) \right\} S_k' \\ &= \sigma^2 S_k \left\{ \left[(D'Z^*)' (D'Z^*) \left(Z^{*'} Z^*\right)^{-1} + k \left(Z^{*'} Z^*\right)^{-1} (D'Z^*)' (D'Z^*) \left(Z^{*'} Z^*\right)^{-1}\right] \left(Z^{*'} Z^* + kI\right) \right\} S_k' \\ &= \sigma^2 S_k \left\{ \left[(DZ^*)' (DZ^*) + k \left(Z^{*'} Z^*\right)^{-1} (DZ^*)' (DZ^*) + k (DZ^*)' (DZ^*) \left(Z^{*'} Z^*\right)^{-1} + k^2 \left(Z^{*'} Z^*\right)^{-1} (DZ^*)' (DZ^*) \left(Z^{*'} Z^*\right)^{-1}\right] \right\} S_k' \\ &= \sigma^2 S_k \left\{ S + kUS + kSU + k^2 USU \right\} S_k' \quad (6.34) \end{aligned}$$

olur. Böylece, $V(\hat{\beta}_{diff}) - V(\hat{\beta}_{Rdiff}(k))$ farkı,

$$\begin{aligned}
&= \sigma^2 \left(USU - S_k S S_k' \right) \\
&= \sigma^2 S_k \left(S_k^{-1} USU S_k^{-1} - S \right) S_k' \\
&= k^2 \sigma^2 S_k \left(USU + \frac{1}{k} (SU + US) \right) S_k' \quad (6.35)
\end{aligned}$$

olarak ifade edilebilir. Burada S simetrik bir matris ve U singüler olmayan simetrik bir matristir.

Şimdi $t = \frac{1}{k} > 0$, $M = USU$, $N = US + SU$ ve $M = \left(Z'^* Z^* \right)^{-1} (DZ^*)' (DZ^*) \left(Z'^* Z^* \right)^{-1} = V'V$

olsun. $rank(V) = p \leq n - m$ olduğundan M , $p \times p$ boyutlu pozitif tanımlı bir matris ve N simetrik bir matristir. Böylece, (6.35) eşitliği aşağıdaki gibi yeniden yazılabilir:

$$\begin{aligned}
\Delta &= k^2 \sigma^2 S_k \left(T^{-1} \right)' T' [M + tN] T T^{-1} S_k, \\
\Delta &= k^2 \sigma^2 S_k \left(T^{-1} \right)' [T' M T + t T' N T] T^{-1} S_k, \\
\Delta &= k^2 \sigma^2 S_k \left(T^{-1} \right)' [I + tG] T^{-1} S_k, \quad (6.36)
\end{aligned}$$

burada $[I + tG] = diag(1 + tg_{11}, \dots, 1 + tg_{pp})$ 'dır.

M pozitif tanımlı bir matris ve N simetrik bir matris olduğundan, singüler olmayan bir T matrisi vardır öyle ki,

$$T' M T = I \text{ ve } T' N T = G$$

olur. Burada G köşegen elemanları $|N - \lambda M| = 0$ polinom eşitliğinin (λ) kökleri olan bir köşegen matristir (Graybill, 1983) ve $N = US + SU \neq 0$ olduğundan G 'nin köşegen elemanlarının en az biri sıfırdan farklıdır. $g_{ii} \neq 0$,

$$0 < t < \min_{g_{ii} \neq 0} \left| \frac{1}{g_{ii}} \right|$$

olsun. Bu durumda $1 + tg_{ii} > 0 \quad i = 1, \dots, p$ ve $I + tG$ pozitif tanımlı bir matris olur. Bundan dolayı, Δ 'nın pozitif tanımlı bir matris olduğunu söyleyebiliriz.

Böylece örneklem varyansı ölçütü ile $\hat{\beta}_{Rdiff}(k)$ 'nin $\hat{\beta}_{diff}$ 'den daha üstün olması için gerek ve yeter koşulu Teorem 6.1'deki gibi verebiliriz (Tabakan ve Akdeniz, 2008).

Teorem 6.1. $\hat{\beta}_{Rdiff}(k)$ tahmin edicisi ancak ve ancak aşağıdaki koşul sağlandığında $\hat{\beta}_{diff}$ tahmin edicisine göre varyans anlamında daha üstündür.

$$\lambda_{\min}(M^{-1}N) > -\frac{1}{t} = -k \quad \text{veya} \quad \lambda_{\min}[(USU)^{-1}(US + SU)] > -k. \quad (6.37)$$

6.5.3. Farka Dayalı Ridge Tahmin Edici $\hat{\beta}_{Rdiff}(k)$ 'nin Üstünlüğü

Bu alt bölümde amacımız iki tahmin edici $\hat{\beta}_{diff}$ ve $\hat{\beta}_{Rdiff}$ 'nin hata kareler ortalama matrislerinin farkını incelemektir. Bu bakımdan $Y = Z\beta + \varepsilon$ modelinde β 'nin tahmin edicisini b^* ile gösterelim. b^* 'nin MSE si aşağıdaki gibi tanımlanır.

$$MSE(b^*, \beta) = E[(b^* - \beta)(b^* - \beta)']. \quad (6.38)$$

Bir b^* tahmin edicisinin kovaryans matrisini $V(b^*)$ ile gösterirsek (6.38) eşitliği

$$MSE(b^*, \beta) = V(b^*) + (bias(b^*))(bias(b^*))' \quad (6.39)$$

olarak yazılır. Burada $bias(b^*) = E(b^*) - \beta$ dir. Skaler değerli hata kareler ortalaması mse ,

$$mse(b^*, \beta) = E \left[(b^* - \beta)' (b^* - \beta) \right] = tr \{ MSE(b^*, \beta) \} \quad (6.40)$$

ile verilir. Buradan ve (6.31) ile (6.33) eşitliklerinden hareketle farka dayalı ridge tahmin edicinin MSE 'sini,

$$\begin{aligned} MSE(\hat{\beta}_{Rdiff}(k), \beta) &= V(\hat{\beta}_{Rdiff}(k)) + (bias(\hat{\beta}_{Rdiff}(k)))(bias(\hat{\beta}_{Rdiff}(k)))' \\ &= S_k (\sigma_\varepsilon^2 S + k^2 \beta \beta') S_k'. \end{aligned} \quad (6.41)$$

olarak elde ederiz. $\hat{\beta}_{diff}$, β için yansız tahmin edici olduğundan benzer biçimde farka dayalı tahmin edicinin MSE 'sini,

$$MSE(\hat{\beta}_{diff}, \beta) = V(\hat{\beta}_{diff}) = \sigma_\varepsilon^2 USU \quad (6.42)$$

olarak elde ederiz. (6.41) ve (6.42) den, $MSE(\hat{\beta}_{diff}, \beta) - MSE(\hat{\beta}_{Rdiff}(k), \beta)$ farkını

$$\begin{aligned} \tilde{\Delta} &= MSE(\hat{\beta}_{diff}, \beta) - MSE(\hat{\beta}_{Rdiff}(k), \beta) \\ &= \sigma^2 S_k \{ S + kUS + kSU + k^2 USU \} S_k' - S_k (\sigma^2 S + k^2 \beta \beta') S_k' \\ &= S_k [\sigma^2 (k(SU + US) + k^2 USU) - k^2 \beta \beta'] S_k', \\ &= k^2 S_k [\sigma^2 (M + tN) - \beta \beta'] S_k', \\ \tilde{\Delta} &= k^2 S_k [\sigma^2 W - \beta \beta'] S_k' \end{aligned} \quad (6.43)$$

olarak yazabiliriz. Burada $W = M + tN$ dir. W simetrik, pozitif tanımlı bir matris ve $\sigma^2 > 0$ olduğundan lineer modellerdeki negatif tanımlı olmayan matrislerin

özelliklerini (Trenkler ve Toutenburg, 1990; Rao ve Toutenburg, 1995; Groß, 2003) veya Farebrother'ın (1976) sonuçlarını kullanarak ancak ve ancak $\beta'W^{-1}\beta < \sigma^2$ ise $\tilde{\Delta}$ nın bir pozitif tanımlı matris olduğunu görebiliriz.

Böylece aşağıdaki teoremi verebiliriz (Tabakan ve Akdeniz, 2008).

Teorem 6.2. $W = M + tN$ pozitif tanımlı bir matris olsun. O zaman $\hat{\beta}_{diff}(k)$ tahmin edicisi ancak ve ancak aşağıdaki koşul sağlandığında $\hat{\beta}_{diff}$ tahmin edicisine göre MSE-anlamında daha üstündür.

$$\beta'W^{-1}\beta \leq \sigma^2 \quad (6.44)$$

6.5.4. Deneysel Uygulama

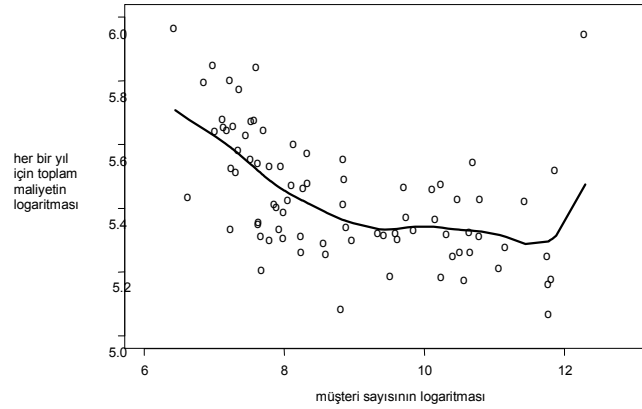
Uygulamada bu bölümden elde edilen sonuçların geçerliliği bir örnek üzerinde gösterilmiştir. Örnek için veriler Kanada'nın Ontario bölgesinde 1993 yılında belediyeye ait 81 elektrik dağıtıcısı ile ilgili bir çalışmadan elde edilmiştir (Yatchew 2000).

Yarı parametrik model Cobb-Douglas modelin basit bir yarı parametrik genelleştirmesine izin verir. Elektrik dağıtımının fiyatı için Cobb-Douglas modelin değişik bir biçimini göz önüne alalım.

$$tm = f(ms) + \beta_1 t\ddot{u} + \beta_2 sm + \beta_3 kk + \beta_4 kwh + \beta_5 gk\ddot{o} + \beta_6 lf + \beta_7 kkm + \varepsilon \quad (6.45)$$

burada tm müşteri başına düşen toplam maliyetin logaritması, ms müşteri sayısının logaritması, $t\ddot{u}$ teknisyen ücretinin logaritması, sm sermaye miktarının logaritması, kk kamu kuruluşu için yapay değişken (ekonomik açıdan fayda sağlayabilen ve ek servisler sunabilen), kwh müşteri başına düşen ortalama kilowatt saatin logaritması, $gk\ddot{o}$ dağıtım varlıklarının geri kalan ömrünün logaritması, lf load faktörünün logaritması (load factor: Bir elektrik santralinden alınan ortalama elektrik miktarının elde edilebilecek max. miktara oranı) ve kkm her bir müşteri için dağıtım kablusunun kilometresinin logaritmasıdır. Model (6.45) hem parametrik hem de parametrik

olmayan bileşenleri içerir. (6.45) ile verilen modelin parametrik olmayan bileşeni sayısal olarak özetlenemediğinden Şekil 6.3'deki gibi grafiksel olarak görüntülenmiştir (Yatchew, 2003).



Şekil 6.3. Yarı parametrik model: Elektrik dağıtımında parametrik olmayan bileşen için uyum eğrisi

Model (6.45)'de parametrik etki β , fark alma yöntemi ile tahmin edilir. Burada $m = 3$ durumundaki 3. dereceden fark alma katsayıları kullanılmıştır. $Z^* Z^* = Z' D' D Z$ matrisi $\lambda_1 = 0.1123423$, $\lambda_2 = 0.2382325$, $\lambda_3 = 1.54118196$, $\lambda_4 = 2.6134937$, $\lambda_5 = 3.4329283$, $\lambda_6 = 15.1239110$, $\lambda_7 = 19.4480101$ öz değerlerine sahiptir. En büyük öz değerin en küçük öz değere oranının karekökü veri kümesindeki çoklu iç ilişkinin varlığını ifade eden $K.S = \sqrt{\lambda_7 / \lambda_1} = \sqrt{173.11} = 13.5$ değeridir.

Farka dayalı ridge regresyon tahmin edici $\hat{\beta}_{Rdiff}(k)$ 'nin hata kareler ortalama matrisinin izi,

$$mse(\hat{\beta}_{Rdiff}(k), \beta) = tr(S_k (\sigma_\varepsilon^2 S + k^2 \beta \beta') S_k')$$
(6.46)

ve farka dayalı tahmin edici $\hat{\beta}_{diff}$ 'nin hata kareler ortalama matrisinin izi aşağıdaki gibi verilir:

6. YARI PARAMETRİK MODELDE FARK ALMA YÖNTEMİNE

DAYALI TAHMİNLER VE İLGİLİ ÇIKARIMLAR

Gülin TABAKAN

$$mse(\hat{\beta}_{diff}, \beta) = tr(\sigma_{\varepsilon}^2 USU) \quad (6.47)$$

(6.46)'da β ve σ^2 farka dayalı tahminleri $\hat{\beta}_{diff}$ ve $\hat{\sigma}^2$ ile yer değiştirirse, belirli bir aralıkta seçilen k değeri ile eşitlik (6.29) için bir tahmin edici ve tahmin edilen mse değerlerini elde ederiz. (6.46) ve (6.47) eşitliklerindeki σ^2 'nin bir tahmin edicisi (2.36)'da tanımlandığı gibidir:

$$\hat{\sigma}^2 = \frac{y'D'(I-P)Dy}{tr(D'(I-P)D)}$$

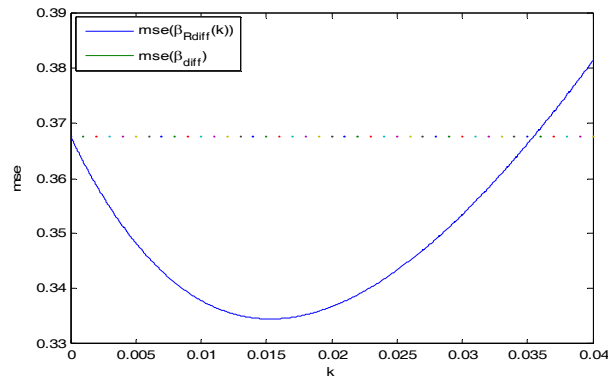
Bu bölümdeki hesaplamalar Matlab 7.0 kullanılarak yapılmıştır. Çizelge 6.3'de farka dayalı tahmin edici ve k 'nın çeşitli değerlerine karşılık gelen farka dayalı ridge tahmin edicinin tahmin edilen varyans ve mse değerleri gösterilmiştir.

Çizelge 6.3. Farka dayalı tahmin edici ve k 'nın çeşitli değerlerine karşılık gelen farka dayalı ridge tahmin edicinin tahmin edilen varyans ve mse değerleri

Katsayılar	k=0	k=0.001	k=0.005	k=0.007	k=0.01	k=0.05	k=0.1	k=0.5
$t\ddot{u}$	0.6113	0.6069	0.5898	0.5817	0.5699	0.4503	0.3582	0.1358
sm	0.5406	0.5399	0.5373	0.5361	0.5342	0.5142	0.4959	0.4165
kk	-0.0754	-0.0755	-0.0760	-0.0762	-0.0766	-0.0801	-0.0829	-0.0906
kwh	0.0134	0.0142	0.0172	0.0186	0.0207	0.0410	0.0558	0.0867
gkō	-0.6258	-0.6253	-0.6237	-0.6228	-0.6216	-0.6052	-0.5858	-0.4640
lf	1.3169	1.3050	1.2592	1.2375	1.2063	0.9014	0.6836	0.2306
kkm	0.4059	0.4057	0.4048	0.4043	0.4036	0.3940	0.3822	0.3051
var	0.3677	0.3628	0.3445	0.3359	0.3238	0.2155	0.1495	0.0437
$m\hat{s}e$	0.3677	0.3629	0.3483	0.3432	0.3379	0.4161	0.6206	1.5072
$\hat{\beta}_{Rdiff}(k) \hat{\beta}_{Rdiff}(k)$	2.9628	2.9246	2.7814	2.7148	2.6210	1.8095	1.3409	0.5693

k 'nın diğer değerleri için $\hat{\beta}_{Rdiff}(k)$ 'nın $\hat{\beta}_{diff}$ üzerinde MSE ölçütüne göre üstünlüğünü göstermek için Şekil 6.4 göz önüne alınabilir.

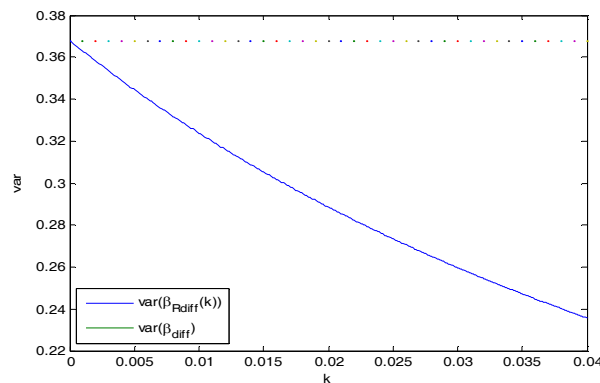
$k \cong 0.03545$ belirlenmiş bir değer olsun. Şekil 6.4'den görülebileceği gibi uygulamada 0.03545 değerinden daha küçük olan k değerleri için, farka dayalı ridge tahmin edici $\hat{\beta}_{Rdiff}(k)$ farka dayalı tahmin edici $\hat{\beta}_{diff}$ 'den daha iyidir. Örneğin, Çizelge 6.3'den de görüldüğü gibi $k = 0.007$ alındığında $\hat{\beta}_{Rdiff}(k)$, $\hat{\beta}_{diff}$ 'den daha küçük mse değerine sahiptir.



Şekil 6.4. k 'nın çeşitli değerleri için $\hat{\beta}_{Rdiff}(k)$ ve $\hat{\beta}_{diff}$ tahmin edicilerinin tahmin edilen mse değerleri

Teorem 6.2'den görüldüğü gibi k 'nın 0.017'den daha küçük değerleri için $\hat{\beta}'(0)W^{-1}\hat{\beta}(0) < \hat{\sigma}^2$ 'dır. Örneğin $k = 0.015$ ve $k = 0.007$ değerleri için sırası ile $\hat{\beta}'(0)W^{-1}\hat{\beta}(0) = 0.0196 < \hat{\sigma}^2 = 0.022$ ve $\hat{\beta}'(0)W^{-1}\hat{\beta}(0) = 0.0094 < \hat{\sigma}^2 = 0.022$ olduğu görülmektedir.

Şekil 6.5 tahmin edicilerin tahmin edilen örneklem varyans değerlerini gösterir.



Şekil 6.5. k 'nın çeşitli değerleri için $\hat{\beta}_{Rdiff}(k)$ ve $\hat{\beta}_{diff}$ tahmin edicilerinin tahmin edilen örneklem varyans değerleri

Teorem 6.2 ancak ve ancak $\beta'W^{-1}\beta \leq \sigma^2$ olduğunda $\hat{\beta}_{Rdiff}(k)$ 'nin $\hat{\beta}_{diff}$ üzerinde MSE ölçütüne göre daha üstün olacağını ifade etmektedir.

Teorik hesaplamalar $0 < k < 0.03545$ aralığında $\hat{\beta}_{Rdiff}(k)$ 'nin tahmin edilen mse değerlerinin $\hat{\beta}_{diff}$ 'nin tahmin edilen mse değerlerinden daha küçük olduğunu göstermektedir. Örneğin, $mse(\hat{\beta}_{Rdiff}(k=0.005)) = 0.3483$ $mse(\hat{\beta}_{diff}(k=0)) = 0.3677$ 'dir. Ayrıca, Çizelge 6.3'den farka dayalı ridge tahmin edicinin farka dayalı tahmin ediciden örneklem varyansı ölçütüne göre her zaman daha iyi olduğu görülmektedir.

Teorem 6.2'de farka dayalı ridge tahmin edicinin farka dayalı tahmin edici üzerinde üstünlüğü için, ridge parametresi k $0 < k < 0.017$ aralığında seçilebilir. Böylece, bu aralık Teorem 6.2 için gerek ve yeter koşulları sağlar (Tabakan ve Akdeniz, 2008).

7. SONUÇLAR VE ÖNERİLER

Regresyon analizi, bağımlı değişkenin bir veya birden fazla bağımsız değişkenle arasındaki ilişkinin matematiksel bir fonksiyon biçiminde yazılmasıdır. Parametrik regresyon analizinde biçimi önceden belirlenmiş modelin parametreleri tahmin edilirken, parametrik olmayan regresyon analizinde ise amaç regresyon fonksiyonu olan $f(x)$ 'i doğrudan tahmin etmektir. Parametrik olmayan regresyon analizi için model belirlenirken bilinmeyen regresyon eğrisini içeren uygun fonksiyon uzayı seçilir. Bu seçim yapılırken regresyon fonksiyonunun süreklilik ve diferansiyellenebilme gibi düzgünlük özelliklerine sahip olduğu varsayılır (Eubank, 1990).

Bu çalışmada parametrik regresyon yöntemleri ile çözümlenemeyen regresyon problemlerinin çözümünde oldukça kapsamlı uygulama alanına sahip olan ve temelde parametrik olmayan bir yöntem olan yarı parametrik regresyon modeli ele alınmış, yarı parametrik regresyon modelinde parametre tahmin yöntemleri incelenmiştir. Çalışmada yarı parametrik regresyon modelinde parametrelerin tahmini için üç farklı yaklaşım ele alınmıştır. Bu yaklaşımlardan ilki rezidü tabanlı tahmin edicilere dayalı düzeltme yaklaşımı, ikinci yaklaşım dalgacık tabanlı tahmin yaklaşımı, üçüncü yaklaşım ise f fonksiyonundan kaynaklanan verideki eğilimi elimine etmeyi amaçlayan fark alma yaklaşımıdır. Bahsedilen bu yaklaşımlar temelde parametrik olmayan regresyon yöntemleri için kullanılan yaklaşımlardır.

Yarı parametrik regresyon modelinde parametre tahmininde kullanılan fark alma yaklaşımından hareketle, yarı parametrik modelde çoklu iç ilişkinin varlığı durumunda farka dayalı ridge tahmin edici olarak adlandırılan yeni bir tahmin edici önerilmiştir. Çoklu iç ilişkinin istatistiksel sonuçlarını ortadan kaldırmak için önerilen farka dayalı ridge tahmin edici ile farka dayalı tahmin ediciden hangisinin daha iyi performans sergilediği araştırılmıştır. Daha küçük hata kareler ortalamasına sahip tahmin edicinin daha iyi olması nedeni ile farka dayalı ridge tahmin edicinin farka dayalı tahmin ediciden hata kareler ortalaması anlamında üstün olması için gerek ve yeter koşullar verilmiştir. Farka dayalı ridge tahmin edicinin farka dayalı tahmin ediciden örneklem varyansı ölçütüne göre daha üstün olması için gerek ve

yeter kořullar verilmiřtir. Parametrik olmayan alıřmalarda yaygın olarak kullanılan bir veri seti alınarak bulunan sonuların geerlilięi gsterilmiřtir.

Elde edilen sonulardan yanlı tahmin edici olan farka dayalı ridge tahmin edicinin farka dayalı tahmin ediciden hata kareler ortalaması ltne gre her zaman stn olmadıęı, sadece belirli kořullar altında stn olduęu grlmektedir. Bu alıřmada elde edilen farka dayalı ridge tahmin ediciye benzer biimde yeni tahmin ediciler elde edilebilir. Elde edilen tahmin edicilerin birbirlerine gre stnlę hata kareler ltne gre deęerlendirilebilir.

KAYNAKLAR

- AHN, H., POWELL, J., 1993. Semiparametric estimation of censored selection models with a nonparametric selection mechanism. *Journal of Econometrics*, 58:3-29.
- ANTONIADIS, A., 2007. Wavelet methods in statistics: Some recent developments and their applications. *Statistics Surveys*, 1:16-55.
- CARROLL, R.J., FAN, J., GIJBELS, I., WAND, M.P., 1997. Generalized partially linear single-index models. *J. Amer. Statist. Assoc.*, 92:477-489.
- CHANG, X., QU, L., 2004. Wavelet estimation of partially linear models. *Computational Statistics and Data Analysis*, 47:31-48.
- CHEN, H., 1988. Convergence rates for parametric components in a partly linear model. *Annals of Statistics*, 16:136-146.
- CUZICK, J., 1992. Semiparametric Additive Regression. *Journal of the Royal Statistical Society. Series B (Methodological)*, Vol. 54, No.3, 831-843.
- ENGLE, R.F., GRANGER, C.W.J., RICE, C.A., WEISS, A., 1986. Semiparametric estimates of the relation between weather and electricity sales. *Journal of the American Statistical Association*, 81:310-320.
- EUBANK, R.L., 1990. *Nonparametric regression and spline smoothing*. Second edition.
- EUBANK, R.L., KAMBOUR, E.L., KIM, J.T., KLIPPLE, K., REESE, C.S., SCHIMEK, M.G., 1998. Estimation in partially linear models. *Computational Statistics and Data Analysis*, 29:27-34.
- FAREBROTHER, R.W., 1976. Further results on the mean square error of ridge regression. *J.R Stat .Soc. B*, 38:248-250.
- GERALD, C.F., WHEATLEY, P.O., 2004. *Applied Numerical Analysis*, 7th ed., Addison-Wesley, Boston.
- GREEN, P., JENNISON, C., SEHEULT, A., 1985. Analysis of field experiments by least squares smoothing. *Journal of the Royal Statistical Society Series B (Methodological)*, 47(2):299-315.
- GREEN, P.J., SILVERMAN, B.W., 1994. *Nonparametric regression and generalized linear models*. Chapman & Hall London.

- GRAYBILL, F.A., 1983. Matrices with applications in statistics. Wadsworth, Belmont, CA.
- GROß, J., 2003. Linear regression. Springer Verlag, New York.
- HALILOV, H., HASANOĞLU, A., CAN, M., 1999. Yüksek Matematik 1-Tek değişkenli fonksiyonlar analizi. ISBN: 975-8431-08-0.
- HÄRDLE, W., 1994. Applied nonparametric regression. Cambridge University Press, Cambridge.
- HÄRDLE, W., MAMMEN, E., MULLER, M., 1998. Testing parametric versus semiparametric modelling in generalized linear models. Journal of the American Statistical Association, 93: 1461-1474.
- HÄRDLE, W., MULLER, M., SPERLICH, S., WERWATZ, A., 2004. Nonparametric and semiparametric models. Springer, New York.
- HASTIE, T., TIBSHIRANI, R.J., 1990. Generalized additive models. Chapman & Hall London.
- HECKMAN, N., 1986. Spline smoothing in partly linear models, J. Roy. Statist. Soc. B, 48: 244–248.
- HEERDE, H.J., LEEFLANG, P.S.H., WITTING, D.R., 2001. Semiparametric Analysis to estimate the deal effect curve. Journal of Marketing Research, 38:2, 197-215.
- HOERL, A.E., KENNARD, R.W., 1970. Ridge regression: biased estimation for nonorthogonal problems. Technometrics, 12:55-67.
- HOERL, A.E., KENNARD, R.W., BALDWIN, K.F., 1975. Ridge regression: some simulation. Communications in Statistics, 4:105-123.
- HOROWITZ, J.L., 1993. Semiparametric Estimation of a Work-Trip Mode Choice Model, Journal of Econometrics, 58, 49-70.
- HU, H., 2005., Ridge estimation of semiparametric regression model. J. Computat. Appl. Math., 176:215–222.
- HURVICH, C.M., SIMONOFF, J.S., TSAI, C., 1998. Smoothing parameter selection in nonparametric regression using an improved Akaike information criterion. Journal of the Royal Statistical Society, Series, B 60: 271-93.

- KIBRIA, B.M.G., 1996. On preliminary test ridge regression estimators for linear restriction in a regression model with non-normal disturbances. *Communications in Statistics A*, 25:2349-2369.
- KLIPPLE, K., EUBANK, R.L., 2007. Difference based variance estimators for partially linear models. *Festschrift in honor of Distinguished Professor Mir Masoom Ali on the occasion of his retirement*. May, 18-19, 313-323.
- LAWLESS, J.F., 1978. Ridge and related estimation procedure. *Communications in Statistics A*, 7:139-164.
- LEE, T.C.M., SOLO, V., 1999. Bandwidth selection for local linear regression: a simulation study. *Comput. Statist.*, 14, 515-532.
- LEE, C.M., 2003. Smoothing parameter selection for smoothing splines: A simulation study. *Comput. Statistics & Data Analysis*, 42, 139-148.
- LAIRD, N.M., WARE, J.H., 1982. Random Effects Models for Longitudinal Data. *Biometrics*, 38:963-974.
- LIANG, H., HÄRDLE, W., CARROLL, R., 1999. Estimation in a semiparametric partially linear errors-in-variables model. *Annals of Statistics*, 27:1519-1535.
- LIANG, H., WANG, N., 2005. Partially linear single-index measurement error models. *Statist. Sinica.*, 15:99–116.
- LOADER, C., 1999. *Local regression and likelihood*. Springer, New York.
- _____, 2004. Smoothing: Local regression principles. *Handbook of Computational Statistics* Ed., J Gentle, W Härdle, Y Mori, <http://www.herine.net/stat/papers.html>
- MALLOWS, C.L., 1973. Some comments on C_p . *Technometrics*, Vol.15, No.4.
- QU, L., 2001 Regularized wavelet estimation in partially linear models. *Referred paper, interface*, 2002.
- RAO, C.R., TOUTENBURG, H., 1995. *Linear models: least squares and alternatives*. Springer Verlag, New York.
- ROBINSON, M.P., 1988. Root-N-Consistent semiparametric regression. *Econometrica*, 56(4): 931-954.
- RUPPERT, D., WAND, M.P., 1994. Multivariate locally weighted least squares regression. *The Annals of Statistics*, 22:1346-1370.

- RUPPERT, D., WAND, M.P., HOLST, U., HOSSJER, O., 1997. Local polynomial variance-function estimation. *Technometrics*, 39:262-272.
- RUPPERT, D., WAND, M.P., CARROLL, R.J., 2003. *Semiparametric regression*. Cambridge University Press.
- SCHIMEK, G.M., 2000. Estimation and inference in partially linear models with smoothing splines. *Journal of Statistical Planning and Inference*, 91:525-540.
- SPECKMAN, P., 1988. Kernel Smoothing in Partially Linear Model. *J. Royal Statist. Soc., B*. 5:413-436.
- TABAKAN, G., AKDENIZ, F., 2008. Difference based ridge estimator of parameters in partial linear model. *Statistical Papers*, DOI: 10.1007/s00362-008-0192-6 (SCI Expanded).
- TRENKLER, G., TOUTENBURG, H., 1990. Mean squared error matrix comparisons between biased estimators-an overview of recent results. *Statistical Papers*, 31:165-179.
- WAHBA, G., 1990. *Spline models for observational data*. CBMS-NSF regional conference series in applied mathematics 59 Philadelphia, PA: SIAM.
- WAND, M.P., GUTIERREZ, R.G., 1997. Exact risk approaches to smoothing parameter selection. *J. Nonparametric Statist.*, 8, 337-354.
- YATCHEW, A., 1997. An elementary estimator of the partial linear model. *Economics Letters* 57:135-143. Additional examples contained in *Economics Letters*, (1998) 59:403-405.
- _____, 1998. Nonparametric regression techniques in economics. *Journal of Economic Literature*, Vol. 36, No.2. 669-721.
- _____, 2000. Scale economies in electricity distribution: A semiparametric analysis. *Journal of Applied Econometrics*, 15(2):187-210.
- _____, 2003. *Semiparametric regression for the applied econometrician*. Cambridge University press.
- YATCHEW, A., NO, A., 2001. Household gasoline demand in Canada. *Econometrica*, 69:1697-1709.

ÖZGEÇMİŞ

Adana doğumlu. İlk, orta ve lise eğitimimini Adana’da tamamladı. 1989 yılında Çukurova Üniversitesi Fen Edb. Fak. Matematik bölümünü kazandı ve 1993 yılında mezun oldu. 1994 yılında M.E.B Niğde Cumhuriyet ilköğretim okuluna Matematik öğretmeni olarak atandı ve 2000 yılına kadar M.E.B’na bağlı okullarda matematik öğretmeni olarak çalıştı. 2000 yılında Aksaray Üniversitesi Eğitim Fakültesine Öğretim Görevlisi olarak atandı ve N.Ü. Fen-Edb. Fak. Matematik bölümü matematik anabilim dalında 2004 yılında yüksek lisansını tamamladı. 2005 yılında Aksaray Üniversitesi Fen-Edb. Fak. Matematik bölümüne araştırma görevlisi olarak atandı ve aynı yıl doktora çalışmasını yapmak üzere Ç.Ü Fen-Edb. Fak. İstatistik Bölümüne 35. madde ile görevlendirildi. Halen bu üniversitede Prof.Dr. Fikri Akdeniz danışmanlığında doktora çalışmasına devam etmektedir.

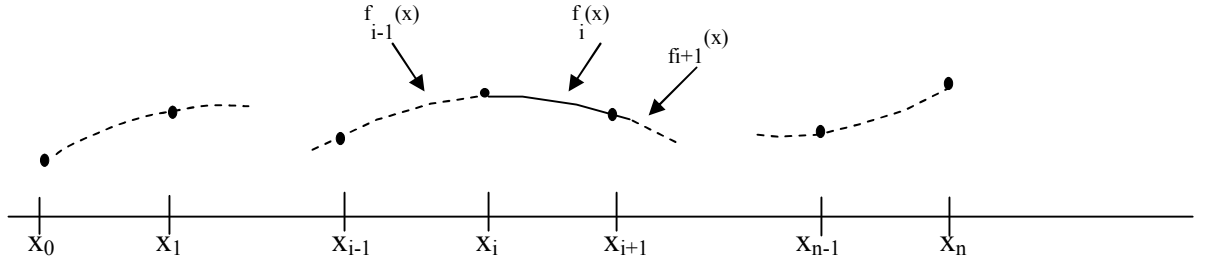
EK

Kübik eğrisel çizgi eğrileri ve Q ile R matrislerinin elde edilmesi

Kübik eğrisel çizgi yaklaşımında, verilen $[x_0, f_0], [x_1, f_1], \dots, [x_n, f_n]$ gözlem noktalarına birçok noktadan geçen polinomlar uydurmak yerine her bir $[x_i, x_{i+1}]$ alt aralığından 3. dereceden kübik bir polinom $f_i(x)$ geçirilir. i . aralıktaki (x_i, y_i) ve (x_{i+1}, y_{i+1}) noktaları arasındaki $f_i(x)$ kübik polinomunu aşağıdaki eşitlik ile verebiliriz:

$$f_i(x) = d_i + c_i(x - x_i) + b_i(x - x_i)^2 + a_i(x - x_i)^3, \quad x_i \leq x \leq x_{i+1}$$

Aşağıdaki şekilde kesiksiz olarak çizilen eğri $f_i(x)$ kübik polinomunu gösterirken kesikli eğriler diğer kübik polinomları göstermektedir.



Böylece $[x_i, x_{i+1}]$ aralığında kübik eğrisel çizgi fonksiyonu aşağıdaki biçimde tanımlanır.

$$f(x) = f_i(x), \quad i = 0, \dots, n-1 \quad (1)$$

Burada $n+1$ nokta varsa $f_i(x)$ 'lerin sayısı ve aralıkların sayısı n tane olup kübik eğrisel çizgi aşağıdaki özelliklere sahiptir:

1) $f_i(x_i) = y_i = f(x_i)$, $i = 0, \dots, n-1$ ve $f_{n-1}(x_n) = y_n$. Yani eğrisel çizgiler her bir x noktasında süreklidir.

2) Her bir $i = 0, \dots, n-2$ için $f_{i+1}(x_{i+1}) = f_i(x_{i+1})$ 'dir.

3) Her bir $i = 0, \dots, n-2$ için $f'_{i+1}(x_{i+1}) = f'_i(x_{i+1})$ 'dir. Yani 1. türev düğüm noktalarında süreklidir.

4) Her bir $i = 0, \dots, n-2$ için $f''_{i+1}(x_{i+1}) = f''_i(x_{i+1})$ 'dir. Yani 2. türev düğüm noktalarında süreklidir.

Yukarıda verilen 1. koşuldan,

$$f_i(x_i) = y_i = d_i, \quad i = 0, \dots, n-1$$

olduğu görülmektedir. 2. koşuldan

$$\begin{aligned} y_{i+1} = f_{i+1}(x_{i+1}) &= f_i(x_{i+1}) = y_i + c_i(x_{i+1} - x_i) + b_i(x_{i+1} - x_i)^2 + a_i(x_{i+1} - x_i)^3 \\ &= y_i + c_i h_i + b_i h_i^2 + a_i h_i^3, \quad i = 0, \dots, n-1 \end{aligned} \quad (2)$$

elde edilir. Burada $h_i = x_{i+1} - x_i$ $i = 0, \dots, n-1$ olduğuna dikkat edilmelidir.

(1) eşitliğinin türevlerinin alınması ile,

$$f'_i(x) = c_i + 2b_i(x - x_i) + 3a_i(x - x_i)^2 \quad (3)$$

$$f''_i(x) = 2b_i + 6a_i(x - x_i) \quad i = 0, \dots, n-1 \quad (4)$$

$$f'''_i(x) = 6a_i$$

elde edilir. (4) eşitliğinden bir kübiğin 2. türevinin lineer olduğunu görebiliriz.

Böylece $f''(x)$ $[x_i, x_{i+1}]$ aralığında doğrusaldır. 2. türevin x_i ve x_{i+1} uç noktalarındaki değerlerine sırası ile γ_i ve γ_{i+1} dersek,

$$f''_i(x_i) = 2b_i = \gamma_i$$

ve

$$f_i''(x_{i+1}) = 2b_i + 6a_i(x_{i+1} - x_i) = \gamma_{i+1}$$

olur. Buradan,

$$b_i = \frac{\gamma_i}{2} \text{ ve } a_i = \frac{\gamma_{i+1} - \gamma_i}{6h_i}$$

olduğu kolayca görülür. (2) eşitliğinde yukarıda bulunan a_i, b_i, d_i değerlerini yerine yazıp c_i 'ye göre çözersek,

$$\begin{aligned} c_i h_i &= y_{i+1} - y_i - b_i h_i^2 - a_i h_i^3, \\ c_i &= \frac{y_{i+1} - y_i}{h_i} - \left[\frac{\gamma_i}{2} h_i + \left(\frac{\gamma_{i+1} - \gamma_i}{6h_i} \right) h_i^2 \right], \\ &= \frac{y_{i+1} - y_i}{h_i} - \frac{2\gamma_i + \gamma_{i+1}}{6} h_i \end{aligned} \quad (5)$$

elde ederiz. Buradan kübik fonksiyonların a_i, b_i, c_i, d_i katsayılarının sadece γ_i parametrelerine bağlı olduğu görülmektedir. Yani γ_i parametreleri elde edilirse kübik fonksiyonların katsayıları elde edilmiş olacaktır. Kübik fonksiyonların 1. türevlerinin uç noktalarında sürekli olması gerektiğini belirten 3.koşul yani $f'_{i-1}(x_i) = f'_i(x_i)$ olduğu göz önüne alınırsa,

$$\begin{aligned} f'_{i-1}(x_i) &= c_{i-1} + 2b_{i-1}(x_i - x_{i-1}) + 3a_{i-1}(x_i - x_{i-1})^2 \\ f'_i(x_i) &= c_i + 2b_i(x_i - x_i) + 3a_i(x_i - x_i)^2 = c_i \end{aligned}$$

eşitliklerinden,

$$c_{i-1} + 2b_{i-1}(x_i - x_{i-1}) + 3a_{i-1}(x_i - x_{i-1})^2 = c_i$$

elde edilir. Elde edilen son eşitlikte a_i, b_i, c_i, d_i katsayılarının yukarıda bulunan değerleri yerlerine yazılırsa,

$$\frac{y_i - y_{i-1}}{h_{i-1}} - \frac{2\gamma_{i-1} + \gamma_i}{6} h_{i-1} + 2 \frac{\gamma_{i-1}}{2} h_{i-1} + 3 \frac{\gamma_i - \gamma_{i-1}}{6h_{i-1}} h_{i-1}^2 = \frac{y_{i+1} - y_i}{h_i} - \frac{2\gamma_i + \gamma_{i+1}}{6} h_i$$

elde edilir. Bulunan bu denklemin γ_i değerlerine göre düzenlenmesi ile ve $y_i = f(x_i)$ den hareketle,

$$h_{i-1}\gamma_{i-1} + 2(h_{i-1} + h_i)\gamma_i + h_i\gamma_{i+1} = 6\left(\frac{f_{i+1} - f_i}{h_i} - \frac{f_i - f_{i-1}}{h_{i-1}}\right), \quad i = 1, \dots, n-1 \quad (6)$$

denklemini elde edilir (Gerald ve Wheatley, 2004). (6) eşitliği $i = 1$ 'den $i = n-1$ 'e kadar her aralıkta sağlanmakta olup toplam $n+1$ nokta vardır. Bu γ_i 'nin $n+1$ değeri ile ilgili $n-1$ denklem verir. (6) eşitliği,

$$i=1 \text{ için} \quad \frac{1}{6}h_0\gamma_0 + \frac{1}{3}(h_0 + h_1)\gamma_1 + \frac{1}{6}h_1\gamma_2 = \frac{1}{h_0}f_0 - \left(\frac{1}{h_0} + \frac{1}{h_1}\right)f_1 + \frac{1}{h_1}f_2$$

$$\begin{array}{ccccccc} \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{array}$$

$$i = n-1 \text{ için} \quad \frac{1}{6}h_{n-2}\gamma_{n-2} + \frac{1}{3}(h_{n-2} + h_{n-1})\gamma_{n-1} + \frac{1}{6}h_{n-1}\gamma_n = \frac{1}{h_{n-2}}f_{n-2} - \left(\frac{1}{h_{n-2}} + \frac{1}{h_{n-1}}\right)f_{n-1} + \frac{1}{h_{n-1}}f_n$$

biçiminde olup matris gösterimi ile (6) eşitliği,

$$\begin{bmatrix} h_0 & 2(h_0+h_1) & h_1 & . & . & . & . & 0 \\ 0 & h_1 & 2(h_1+h_2) & h_2 & . & . & . & . \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ . & . & . & . & h_{n-2} & 2(h_{n-2}+h_{n-1}) & h_{n-1} & . \end{bmatrix} \begin{bmatrix} \gamma_0 \\ \gamma_1 \\ . \\ . \\ . \\ . \\ \gamma_{n-1} \\ \gamma_n \end{bmatrix} = 6 \begin{bmatrix} \frac{1}{h_0} & -\left(\frac{1}{h_0}+\frac{1}{h_1}\right) & \frac{1}{h_1} & . & . & . & . & 0 \\ 0 & \frac{1}{h_2} & -\left(\frac{1}{h_2}+\frac{1}{h_3}\right) & \frac{1}{h_3} & 0 & . & . & . \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ 0 & . & . & . & \frac{1}{h_{n-2}} & -\left(\frac{1}{h_{n-2}}+\frac{1}{h_{n-1}}\right) & \frac{1}{h_{n-1}} & . \end{bmatrix} \begin{bmatrix} f_0 \\ f_1 \\ . \\ . \\ . \\ . \\ f_{n-1} \\ f_n \end{bmatrix}$$

biçiminde yazılır. Yukarıda verilen kübik eğrisel çizgi koşullarına ilave olarak $[a, b]$ kapalı aralığının uç noktalarında f fonksiyonunun 2. ve 3. dereceden türevleri sıfır ise kübik eğrisel çizginin doğal kübik eğrisel çizgi olarak adlandırıldığını 2. Bölümden biliyoruz. Böylece uç noktalarında kübik eğrilerin birer doğruya dönüştükleri varsayılarak 2. ve 3. dereceden türevlerin sıfıra eşit olması ile $\gamma_0 = \gamma_n = 0$ elde edilir. Böylece yukarıdaki matris sistemini,

$$\begin{bmatrix} \frac{1}{3}(h_0+h_1) & \frac{1}{6}h_1 & 0 & . & . & 0 & 0 \\ \frac{1}{6}h_1 & \frac{1}{3}(h_1+h_2) & \frac{1}{6}h_2 & . & . & . & 0 \\ 0 & \frac{1}{6}h_2 & \frac{1}{3}(h_2+h_3) & \frac{1}{6}h_3 & 0 & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ 0 & . & . & . & 0 & \frac{1}{6}h_{n-2} & \frac{1}{3}(h_{n-2}+h_{n-1}) \end{bmatrix} \begin{bmatrix} \gamma_1 \\ \gamma_3 \\ . \\ . \\ . \\ \gamma_{n-1} \end{bmatrix} = \begin{bmatrix} \frac{1}{h_0} & -\left(\frac{1}{h_0}+\frac{1}{h_1}\right) & \frac{1}{h_1} & . & . & . & 0 \\ 0 & \frac{1}{h_2} & -\left(\frac{1}{h_2}+\frac{1}{h_3}\right) & \frac{1}{h_3} & 0 & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ 0 & . & . & . & \frac{1}{h_{n-2}} & -\left(\frac{1}{h_{n-2}}+\frac{1}{h_{n-1}}\right) & \frac{1}{h_{n-1}} \end{bmatrix} \begin{bmatrix} f_0 \\ f_1 \\ . \\ . \\ . \\ f_{n-1} \\ f_n \end{bmatrix}$$

biçiminde yazıp, bu ifadeleri $R\gamma = Q'f$ matris gösterimi ile yazabiliriz.

Örnek: Aşağıdaki çizelgede verilen değerlere uygun doğal kübik eğrisel çizgi eğrisini çizip bu eğrinin $f(0.76)$ ve $f(1.65)$ için değerlerini bulunuz ($f(x) = 2e^x - x^2$ dir).

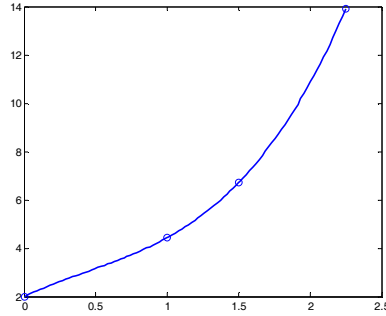
x	$f(x)$
0.0	2.0000
1.0	4.4366
1.5	6.7134
2.25	13.9130

Çizelgedeki değerlerden $h_0 = 1.0$, $h_1 = 0.5$ ve $h_2 = 0.75$ olduğu görülmektedir. Bir doğal kübik eğrisel çizgi için,

$$\begin{bmatrix} 3.0 & 0.5 \\ 0.5 & 2.5 \end{bmatrix} \begin{bmatrix} \gamma_1 \\ \gamma_2 \end{bmatrix} = \begin{bmatrix} 12.7020 \\ 30.2754 \end{bmatrix}$$

matris eşitliğinden $\gamma_1 = 2.2920$, $\gamma_2 = 11.6518$ ($\gamma_0 = \gamma_3 = 0$) elde ederiz. Bu değerleri kullanarak her bir aralıktaki kübik eğri çizgilerin katsayılarını aşağıdaki gibi hesaplayabiliriz.

i	Aralık	$f_i(x)$
0	$[0.0, 1.0]$	$0.3820(x-0)^3 + 0(x-0)^2 + 2.0546(x-0) + 2.0000$
1	$[1.0, 1.5]$	$3.1199(x-1)^3 + 1.146(x-1)^2 + 3.2005(x-1) + 4.4366$
2	$[1.5, 2.25]$	$-2.5893(x-1.5)^3 + 5.8259(x-1.5)^2 + 6.6866(x-1.5) + 6.7134$



Şekil 1. Kübik eğrisel çizgi

Şekil 1 kübik eğrisel çizgi eğrisini göstermektedir. Buradan $f(0.76) = 3.7519$ (Gerçek değeri=3.6989) ve $f(1.65) = 7.7218$ (Gerçek değeri=7.6914) değerlerinin elde edileceği kolayca görülmektedir.