

**TÜRKİYE’DE İHRACATIN İTHALATI KARŞILAMA
ORANLARININ SPLİNE REGRESYON MODELLERİ
YARDIMIYLA ARAŞTIRILMASI**

Ömer ALKAN

**Doktora Tezi
İşletme Anabilim Dalı
Prof. Dr. Erkan OKTAY**

**2013
Her Hakkı Saklıdır**

**ATATÜRK ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ
İŞLETME ANABİLİM DALI**

Ömer ALKAN

**TÜRKİYE’DE İHRACATIN İTHALATI KARŞILAMA
ORANLARININ SPLİNE REGRESYON MODELLERİ
YARDIMIYLA ARAŞTIRILMASI**

DOKTORA TEZİ

**TEZ YÖNETİCİSİ
Prof. Dr. Erkan OKTAY**

ERZURUM – 2013



T.C.
ATATÜRK ÜNİVERSİTESİ
SOSYAL BİLİMLERİ ENSTİTÜSÜ



TEZ BEYAN FORMU

17.01.2013

SOSYAL BİLİMLERİ ENSTİTÜSÜ MÜDÜRLÜĞÜNE

BİLDİRİM

Atatürk Üniversitesi Lisansüstü Eğitim-Öğretim ve Sınav Yönetmeliğine göre hazırlamış olduğum "Türkiye’de İhracatın İthalatı Karşılama Oranlarının Spline Regresyon Modelleri Yardımıyla Araştırılması " adlı tezin/raporun tamamen kendi çalışmam olduğunu ve her alıntıya kaynak gösterdiğimi taahhüt eder, tezimin/raporumun kağıt ve elektronik kopyalarının Atatürk Üniversitesi Sosyal Bilimler Enstitüsü arşivlerinde aşağıda belirttiğim koşullarda saklanması için verdiğimi onaylarım:

Lisansüstü Eğitim-Öğretim yönetmeliğinin ilgili maddeleri uyarınca gereğinin yapılmasını arz ederim.

- ☒ Tezimin/Raporumun tamamı her yerden erişime açılabilir.
☐ Tezim/Raporum sadece Atatürk Üniversitesi yerleşkelerinden erişime açılabilir.
☐ Tezimin/Raporumun yıl süreyle erişime açılmasını istemiyorum. Bu sürenin sonunda uzatma için başvuruda bulunmadığım takdirde, tezimin/raporumun tamamı her yerden erişime açılabilir.

17.01.2013

[Tarih ve İmza]

[Ömer ALKAN]



T.C.
ATATÜRK ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ



TEZ KABUL TUTANAĞI

SOSYAL BİLİMLER ENSTİTÜSÜ MÜDÜRLÜĞÜNE

Prof. Dr. Erkan OKTAY danışmanlığında, Ömer ALKAN tarafından hazırlanan bu çalışma 16 / 01 / 2013 tarihinde aşağıdaki jüri tarafından İşletme Anabilim Dalı'nda Doktora Tezi olarak kabul edilmiştir.

Başkan : Prof. Dr. Aşır GENÇ

İmza:

Jüri Üyesi : Prof. Dr. Erkan OKTAY

İmza:

Jüri Üyesi : Prof. Dr. Alaattin KIZILTAN

İmza:

Jüri Üyesi : Doç. Dr. Mehmet Suphi ÖZÇOMAK

İmza:

Jüri Üyesi : Yrd. Doç. Dr. Emrah TALAŞ

İmza:

Yukarıdaki imzalar adı geçen öğretim üyelerine aittir. / /

Prof. Dr. Mustafa YILDIRIM

Enstitü Müdürü

İÇİNDEKİLER

ÖZET.....	VI
ABSTRACT.....	VII
KISALTMALAR DİZİNİ	VIII
ŞEKİLLER DİZİNİ.....	IX
TABLolar DİZİNİ.....	XI
ÖNSÖZ.....	XIII
GİRİŞ	1

BİRİNCİ BÖLÜM

SPLİNE KAVRAMI

1.1. LİTERATÜR ARAŞTIRMASI	4
1.1.1. Dünyada Yapılan Çalışmalar	4
1.1.2. Türkiye’de Yapılan Çalışmalar	13
1.2. İNTERPOLASYON KAVRAMI.....	15
1.2.1. Global İnterpolasyon.....	16
1.2.2. Parçalı İnterpolasyon.....	16
1.3. SPLİNE TANIMI VE MAHİYETİ	16
1.4. SPLİNELARIN ÖNEMİ	19
1.5. SPLİNE FONKSİYONLARI.....	20
1.5.1. Doğrusal Spline Fonksiyonu	21
1.5.2. Kareli (Kuadratik) Spline Fonksiyon	23
1.5.3. Kübik Spline Fonksiyon.....	25
1.5.3.1. Kübik Spline Katsayılarının Bulunmasında Birinci Yaklaşım	27
1.5.3.2. Kübik Spline Katsayılarının Bulunmasında İkinci Yaklaşım	31
1.5.3.3. Kübik Spline Fonksiyonda Sınır Değerleri Sorunu	35
1.6. SPLİNE GÖSTERİMİ.....	36
1.6.1. Kesikli Üstel Fonksiyon Gösterimi (+ Fonksiyon)	36
1.6.2. B-Spline	36
1.6.3. İnterpolasyon Gösterimi.....	38

1.7. SPLİNE TÜRLERİ	38
1.7.1. Regresyon Spline	38
1.7.2. Düzgünleştirici Spline (Smoothing Spline)	39
1.7.3. Çok Değişkenli Spline (Multivariate Spline)	42
1.7.3.1. İnce Tabakalı Spline (Thin Plate Spline)	42
1.7.3.2. Tensor Çarpım Spline (Tensor Product Spline)	43

İKİNCİ BÖLÜM

SPLİNE REGRESYONDA DÜĞÜM NOKTASI

2.1. DÜĞÜM NOKTASININ TANIMI	52
2.2. DÜĞÜM NOKTASININ ÖNCEDEN BİLİNMESİ (SABİT DÜĞÜM NOKTASI)	53
2.3. DÜĞÜM NOKTASI KONUMUNUN ÖNCEDEN BİLİNMEMESİ (DEĞİŞEN DÜĞÜM NOKTASI)	53
2.4. DÜĞÜM NOKTASININ BELİRLENMESİNDE KULLANILAN YÖNTEMLER	55
2.4.1. Noktasal Dağılımdan (Serpilme Diyagramı) Görsel Olarak Belirleme	55
2.4.2. Adımsal (Stepwise) Regresyon Yöntemi	56
2.4.3. Eşit Aralıklı Yöntem (Equally Spaced Method)	58
2.4.4. Eşit Aralıklı Örneklem Kantilleri Yöntemi (Equally Spaced Sample Quantiles Method)	58
2.4.5. Diğer Yöntemler	58

ÜÇÜNCÜ BÖLÜM

SPLİNE REGRESYON MODELLERİ

3.1. SPLİNE REGRESYONA GİRİŞ	60
3.2. GÖLGE DEĞİŞKENLİ REGRESYON	62
3.3. KESİKLİ PARÇALI REGRESYON	64
3.3.1. Parametreleri En Küçük Kareler Yöntemiyle Tahmin Edilen Kesikli Regresyon	65

3.3.2. Parametreleri En Çok (Maksimum) Olabilirlik Yöntemiyle Tahmin Edilen Kesikli Regresyon	67
3.4. SPLİNE REGRESYON ÇEŞİTLERİ	70
3.4.1. Doğrusal Spline Regresyon	72
3.4.2. Kareli (Kuadratik) Spline Regresyon	74
3.4.2.1. Kareli Spline Regresyonla Mevsimselliğin Hesaplanması	76
3.4.3. Kübik Spline Regresyon	79
3.4.3.1. Suits ve Ark. Tarafından Geliştirilen Kübik Spline Regresyon	82
3.4.4. Splineların Kısmi Sıralaması	88

DÖRDÜNCÜ BÖLÜM

SPLİNE REGRESYONDA MODEL SEÇİMİ

4.1. MODEL SEÇİM KRİTERLERİ	90
4.1.1. Belirlilik Katsayısı	90
4.1.2. Düzeltilmiş R^2	91
4.1.3. Bilgi Kriterleri	92
4.1.3.1. Akaike Bilgi Kriteri (AIC)	93
4.1.3.2. Schwarz Bilgi Kriteri (SIC)	94
4.1.4. PRESS İstatistiği	95
4.1.5. R^2_{pred} Değeri	96
4.1.6. Mallows'un C_p Kriteri	96
4.1.7. Modelin Hatası (S)	97
4.1.8. Hata Kareler Ortalaması (MSE)	97
4.1.9. Hata Kareler Ortalaması Karekökü (RMSE)	98
4.1.10. Ortalama Mutlak Hata (MAE) ve Ortalama Mutlak Hata Yüzdesi (MAPE) ..	99
4.1.11. Oransal Hata (Relative Error)	99
4.1.12. Hata Kümesinin Katsayısı (CRM)	100
4.1.13. Etkinlik İndeksi (Kapa İstatistiği)	100

BEŞİNCİ BÖLÜM

İHRACATIN İTHALATI KARŞILAMA ORANI DAĞILIMINDA DÜĞÜM NOKTASININ BELİRLENMESİ VE SPLİNE REGRESYON MODELLERİNİN KURULMASI

5.1. İHRACATIN İTHALATI KARŞILAMA ORANI DAĞILIMINDA ÖNERİLEN YÖNTEMLE DÜĞÜM NOKTALARININ BELİRLENMESİ.....	102
5.1.1. 1923-2010 Yılları Türk Dış Ticaretinin Kısaca Değerlendirilmesi	102
5.1.2. Yöntem ve Uygulaması.....	104
5.2. BELİRLENEN DÜĞÜM NOKTALARINA GÖRE SPLİNE REGRESYON MODELLERİNİN KURULMASI.....	110
5.2.1. Doğrusal Spline Regresyon Modelleri	112
5.2.1.1. Kısıtlandırılmamış Doğrusal Spline Regresyon Modeli	112
5.2.1.2. Kısıtlandırılmış Doğrusal Spline Regresyon Modeli	114
5.2.2. Kareli (Kuadratik) Spline Regresyon Modelleri	116
5.2.2.1. Kısıtlandırılmamış Kareli Spline Regresyon Modeli	116
5.2.2.2. Kısıtlandırılmış Kareli (Kuadratik) Spline Regresyon Modeli	118
5.2.2.2.1. Sabit Katsayıları Kısıtlandırılmış Kareli Spline Regresyon Modeli .	118
5.2.2.2.2. Sabit Katsayıları Kısıtlandırılmış ve Düğüm Noktasında Birinci Derece Türevleri Eşit Olan Kareli Spline Regresyon Modeli.....	120
5.2.3. Kübik Spline Regresyon Modeli	122
5.2.3.1. Kısıtlandırılmamış Kübik Spline Regresyon Modeli	122
5.2.3.2. Kısıtlandırılmış Kübik Spline Regresyon Modeli	125
5.2.3.2.1. Sabit Katsayıları Kısıtlandırılmış Kübik Spline Regresyon Modeli .	125
5.2.3.2.2. Sabit Katsayıları Kısıtlandırılmış ve Düğüm Noktasında Birinci Derece Türevleri Eşit Olan Kübik Spline Regresyon Modeli.....	128
5.2.3.2.3. Sabit Katsayıları Kısıtlandırılmış ve Düğüm Noktasında İkinci Derece Türevleri Eşit Olan Kübik Spline Regresyon Modeli	130
5.3. POLİNOM REGRESYON MODELLERİ.....	132
5.3.1. Basit Doğrusal Regresyon Modeli	132
5.3.2. Kareli Regresyon Modeli	134
5.3.3. Kübik Regresyon Modeli	135

5.4. MODELLERİN KARŞILAŞTIRILMASI.....	137
SONUÇ VE DEĞERLENDİRME	139
KAYNAKÇA	143
EKLER.....	162
ÖZGEÇMİŞ.....	163

ÖZET**DOKTORA TEZİ****TÜRKİYE’DE İHRACATIN İTHALATI KARŞILAMA ORANLARININ
SPLİNE REGRESYON MODELLERİ YARDIMIYLA ARAŞTIRILMASI****Ömer ALKAN****Tez Danışmanı: Prof. Dr. Erkan OKTAY****2013, 178 sayfa****Jüri: Prof. Dr. Erkan OKTAY (Danışman)****Prof. Dr. Aşır GENÇ****Prof. Dr. Alaattin KIZILTAN****Doç. Dr. M. Suphi ÖZÇOMAK****Yrd. Doç. Dr. Emrah TALAŞ**

Spline, basitçe parçalı polinomlardan oluşan bir eğri olarak tanımlanabilir. Birçok farklı alanda kullanılan splinelar, regresyon analizinde de kullanılmaktadır. Bu çalışmada, spline fonksiyonların doğrusal, kareli ve kübik spline regresyon modellerinde kullanımı incelenmiş ve en küçük kareler yöntemiyle verilerden spline parametrelerinin nasıl tahmin edileceği gösterilmiştir.

Spline regresyonda, düğüm noktalarının sayısının ve konumunun belirlenmesi önemli bir sorundur. Bu çalışmada öncelikle, Belirlilik Katsayısı (R^2) temel alınarak spline regresyonda düğüm noktasının tahminine ilişkin bir yöntem önerilmiştir. Türkiye’nin 1923-2010 yılları arası ihracatın ithalatı karşılama oranlarının dağılımında önerilen yöntemle düğüm noktaları belirlenmiş ve belirlenen düğüm noktalarına göre doğrusal, kareli ve kübik spline regresyon modelleri kurulmuştur.

Spline regresyon modellerin polinom regresyon modellere göre daha iyi sonuçlar verdiği belirlenmiştir. En düzgün spline regresyon modellerinde, kareli spline regresyon modelinin ihracatın ithalatı karşılama oranları dağılımını en iyi açıklayan model olduğu tespit edilmiştir.

Anahtar Kelimeler: Spline Regresyon, Düğüm Noktası, İhracat, İthalat

ABSTRACT**Ph. D. DISSERTATION****THE STUDY OF IMPORT SUPPLY RATES OF EXPORT IN TURKEY
WITH THE HELP OF SPLINE REGRESSION MODELS****Ömer ALKAN****Advisor: Prof. Dr. Erkan OKTAY****2013, Page: 178****Jury: Prof. Dr. Erkan OKTAY (Advisor)****Prof. Dr. Aşır GENÇ****Prof. Dr. Alaattin KIZILTAN****Assoc. Prof. Dr. M. Suphi ÖZÇOMAK****Assist. Prof. Dr. Emrah TALAŞ**

Spline can simply be defined as a curve made up of piecewise polynomials. Splines used in many different fields are used in regression analyses as well. In this study, the use of spline functions in the models of linear, quadratic and cubic spline regression is studied and with the method of the ordinary least squares, it is shown how spline parameters are estimated from the data.

In the spline regression, the determination of the number and place of knots is an important problem. In this study, primarily on the basis of coefficient of determination (R^2), a method related to the estimation of knot in the spline regression is proposed. With the method suggested in the distribution of import supply rates of export of Turkey within 1923 – 2010, knots are determined and in accordance with the specified knots, linear, quadratic and cubic spline regression models are established.

It is determined that spline regression models in contrast to polynomial regression models gave better results. In the most smooth spline regression models, it is confirmed that quadratic spline regression model is the most explanatory model in the distribution of import supply rates of export.

Key Words: Spline Regression, Knot, Export, Import

KISALTMALAR DİZİNİ

RSS	:	Hata (Artık) Kareler Toplamı
ESS	:	Regresyon Kareler Toplamı
TSS	:	Genel Kareler Toplamı
R^2	:	Belirlilik Katsayısı
$\bar{R}^2$	:	Düzeltilmiş Belirlilik Katsayısı
AIC	:	Akaike Bilgi Kriteri
SIC	:	Schwarz Bilgi Kriteri
PRESS	:	Öngörülen Hata Kareler toplamı
R^2_{pred}	:	PRESS Belirlilik Katsayısı
C_p	:	Mallows'un C_p Kriteri
S	:	Modelin Hatası
MSE	:	Hata Kareler Ortalaması
RMSE	:	Hata Kareler Ortalaması Karekökü
MAE	:	Ortalama Mutlak Hata
MAPE	:	Ortalama Mutlak Hata Yüzdesi
RE	:	Oransal Hata
CRM	:	Hata Kümesinin Katsayısı
Eİ	:	Etkinlik İndeksi

ŞEKİLLER DİZİNİ

Şekil 1.1. Mekanik Spline Örnekleri ve Ağırlıklar	17
Şekil 1.2. Matematikte Kullanılan Spline Fonksiyonu	18
Şekil 1.3. Doğrusal Spline Fonksiyonu	23
Şekil 1.4. Kareli Spline Fonksiyonu	24
Şekil 1.5. Kübik Spline Fonksiyonu	26
Şekil 2.1. Doğrusal Spline Regresyon ve Muhtemel Düğüm Noktası Alanı	56
Şekil 2.2. Kareli Spline Regresyon ve Muhtemel Düğüm Noktası Alanı	56
Şekil 2.3. Kübik Spline Regresyon ve Muhtemel Düğüm Noktası Alanları	56
Şekil 3.1. Bir Regresyon Katsayısının Mevsimsel Model Eğrisi	77
Şekil 3.2. Parçalı polinom regresyon	82
Şekil 3.3. Splineların Kısmi Sıralaması	89
Şekil 5.1. İlk Adım Sonrası R^2 Değerlerinin Dağılımı	105
Şekil 5.2. İkinci Adım Sonrası R^2 Değerleri Dağılımı.....	106
Şekil 5.3. Üçüncü Adım Sonrası R^2 Değerleri Dağılımı	107
Şekil 5.4. Dördüncü Adım Sonrası R^2 Değerleri Dağılımı.....	108
Şekil 5.5. Beşinci Adım Sonrası R^2 Değerleri Dağılımı.....	109
Şekil 5.6. Altıncı Adım Sonrası R^2 Değerleri Dağılımı.....	110
Şekil 5.7. 1923-2010 Yılları Arası İhracatın İthalatı Karşılama Oranlarının Dağılımı ve Düğüm Noktaları	111
Şekil 5.8. Kısıtlandırılmamış Doğrusal Spline Regresyonla Tahmin Edilen ve Gözlenen İhr./İth. Oranı Dağılımı	114
Şekil 5.9. Kısıtlandırılmış Doğrusal Spline Regresyonla Tahmin Edilen ve Gözlenen İhr./İth. Oranı Dağılımı	115
Şekil 5.10. Kısıtlandırılmamış Kareli Spline Regresyonla Tahmin Edilen ve Gözlenen İhr./İth. Oranı Dağılımı	118
Şekil 5.11. Sabit Katsayıları Kısıtlandırılmış Kareli Spline Regresyonla Tahmin Edilen ve Gözlenen İhr./İth. Oranı Dağılımı.....	120
Şekil 5.12. Sabit Katsayıları Kısıtlandırılmış ve Düğüm Noktasında Birinci Derece Türevleri Eşit Olan Kareli Spline Regresyonla Tahmin Edilen ve Gözlenen İhr./İth. Oranı Dağılımı	122

Şekil 5.13. Kısıtlandırılmamış Kübik Spline Regresyonla Tahmin Edilen ve Gözlenen İhr./İth. Oranı Dağılımı.....	125
Şekil 5.14. Sabit Katsayıları Kısıtlandırılmış Kübik Spline Regresyonla Tahmin Edilen ve Gözlenen İhr./İth. Oranı Dağılımı	127
Şekil 5.15. Sabit Katsayıları Kısıtlandırılmış ve Düğüm Noktasında Birinci Derece Türevleri Eşit Olan Kübik Spline Regresyonla Tahmin Edilen ve Gözlenen İhr./İth. Oranı Dağılımı	130
Şekil 5.16. Sabit Katsayıları Kısıtlandırılmış ve Düğüm Noktasında İkinci Derece Türevleri Eşit Olan Kübik Spline Regresyonla Tahmin Edilen ve Gözlenen İhr./İth. Oranı Dağılımı	132
Şekil 5.17. Basit Doğrusal Regresyonla Tahmin Edilen ve Gözlenen İhr./İth. Oranı Dağılımı	133
Şekil 5.18. Kareli Regresyonla Tahmin Edilen ve Gözlenen İhr./İth. Oranı Dağılımı.	135
Şekil 5.19. Kübik Regresyonla Tahmin Edilen ve Gözlenen İhr./İth. Oranı Dağılımı.	136

TABLOLAR DİZİNİ

Tablo 5.1. İlk Adım Sonrası R^2 Değerleri	105
Tablo 5.2. İkinci Adım Sonrası R^2 Değerleri	106
Tablo 5.3. Üçüncü Adım Sonrası R^2 Değerleri	107
Tablo 5.4. Dördüncü Adım Sonrası R^2 Değerleri.....	108
Tablo 5.5. Beşinci Adım Sonrası R^2 Değerleri	109
Tablo 5.6. Altıncı Adım Sonrası R^2 Değerleri	110
Tablo 5.7. Kısıtlandırılmamış Doğrusal Spline Regresyonda Varyans Analizi.....	113
Tablo 5.8. Kısıtlandırılmamış Doğrusal Spline Regresyon Modelinde Tahminlerin Çıktısı.....	113
Tablo 5.9. Kısıtlandırılmış Doğrusal Spline Regresyonda Varyans Analizi.....	115
Tablo 5.10. Kısıtlandırılmış Doğrusal Spline Regresyon Modelinde Tahminlerin Çıktısı	115
Tablo 5.11. Kısıtlandırılmamış Kareli Spline Regresyonda Varyans Analizi.....	116
Tablo 5.12. Kısıtlandırılmamış Kareli Spline Regresyon Modelinde Tahminlerin Çıktısı	117
Tablo 5.13. Sabit Katsayıları Kısıtlandırılmış Kareli Spline Regresyonda Varyans Analizi	119
Tablo 5.14. Sabit Katsayıları Kısıtlandırılmış Kareli Spline Regresyon Modelinde Tahminlerin Çıktısı	119
Tablo 5.15. Sabit Katsayıları Kısıtlandırılmış ve Düğüm Noktasında Birinci Derece Türevleri Eşit Olan Kareli Spline Regresyonda Varyans Analizi.....	121
Tablo 5.16. Sabit Katsayıları Kısıtlandırılmış ve Düğüm Noktasında Birinci Derece Türevleri Eşit Olan Kareli Spline Regresyon Modelinde Tahminlerin Çıktısı.....	121
Tablo 5.17. Kısıtlandırılmamış Kübik Spline Regresyonda Varyans Analizi.....	123

Tablo 5.18. Kısıtlandırılmamış Kübik Spline Regresyon Modelinde Tahminlerin Çıktısı	124
Tablo 5.19. Sabit Katsayıları Kısıtlandırılmış Kübik Spline Regresyonda Varyans Analizi	126
Tablo 5.20. Sabit Katsayıları Kısıtlandırılmış Kübik Spline Regresyon Modelinde Tahminlerin Çıktısı.....	127
Tablo 5.21. Sabit Katsayıları Kısıtlandırılmış ve Düğüm Noktasında Birinci Derece Türevleri Eşit Olan Kübik Spline Regresyonda Varyans Analizi	128
Tablo 5.22. Sabit Katsayıları Kısıtlandırılmış ve Düğüm Noktasında Birinci Derece Türevleri Eşit Olan Kübik Spline Regresyon Modelinde Tahminlerin Çıktısı	129
Tablo 5.23. Sabit Katsayıları Kısıtlandırılmış ve Düğüm Noktasında İkinci Derece Türevleri Eşit Olan Kübik Spline Regresyonda Varyans Analizi	131
Tablo 5.24. Sabit Katsayıları Kısıtlandırılmış ve Düğüm Noktasında İkinci Derece Türevleri Eşit Olan Kübik Spline Regresyon Modelinde Tahminlerin Çıktısı.....	131
Tablo 5.25. Basit Doğrusal Regresyonda Varyans Analizi.....	133
Tablo 5.26. Basit Doğrusal Regresyon Modelinde Tahminlerin Çıktısı	133
Tablo 5.27. Kareli Regresyonda Varyans Analizi.....	134
Tablo 5.28. Kareli Regresyon Modelinde Tahminlerin Çıktısı.....	134
Tablo 5.29. Kübik Regresyonda Varyans Analizi.....	136
Tablo 5.30. Kübik Regresyon Modelinde Tahminlerin Çıktısı.....	136
Tablo 5.31. Model Seçim Kriterleri	137
Tablo 5.32-a. Modellerin Karşılaştırılması	138
Tablo 5.33-b. Modellerin Karşılaştırılması	138

ÖNSÖZ

Bu tez çalışmasında önümüzü açan, maddi/manevi katkılarıyla bize her türlü desteği sağlayan, değerli bilgileri ve önerileri ile yardımlarını esirgemeyen kıymetli ve saygı değer hocam Prof. Dr. Erkan OKTAY'a,

Tez çalışması için geçici görevle gittiğim Selçuk Üniversitesi Fen Fakültesi İstatistik Bölümü'nde bana her türlü kolaylığı sağlayan ve bu tezin ortaya çıkmasında çok önemli katkıları olan değerli hocam İstatistik Bölüm Başkanı Prof. Dr. Aşır GENÇ'e, Selçuk Üniversitesi'nde bulunduğum süre içinde beni kendilerinden biri olarak gören ve bana her konuda yardımcı olan İstatistik Bölümü araştırma görevlileri ve öğretim üyelerine,

Yardımlarına her zaman başvurduğum kıymetli arkadaşlarım Arş. Gör. Miraç EREN, Arş. Gör. Gökhan ERKAL ve Arş. Gör. Abdulkerim KARAASLAN'a ve önerileri ile tez çalışmasına katkı sağlayan Ekonometri Bölümü öğretim üyeleri ve arkadaşlarıma,

Tez çalışmam süresince hiçbir zaman desteklerini esirgemeyen değerli eşim ve aileme teşekkürlerimi borç bilirim.

GİRİŞ

Regresyon analizi, değişkenler arasındaki ilişkiyi modellemek ve araştırmak için yaygın olarak kullanılan istatistiksel bir tekniktir. Regresyon uygulamaları sayısal verilerden oluşur. Mühendislik, ziraat, tıp, sosyoloji, kimya, fizik, biyoloji, iktisat, işletme gibi sosyal-fen ve mühendislik bilimlerin bütün alanlarında kullanılır.

Hangi tür olursa olsun, regresyon yöntemlerinin iki temel amacı vardır. Birincisi, regresyon yöntemleri bağımsız ve bağımlı değişkenler arasındaki ilişkiyi incelemenin ve bunu ortaya koymanın bir yolunu gösterir. İkinci olarak, henüz gerçekleşmeyen gözlemlerin tahminini verir. Bunlara ilaveten üçüncü bir amaç olarak da regresyon yöntemleri eğri tahmini için kullanılmaktadır.

Regresyon yöntemleri, bağımsız değişken/değişkenlerdeki değişikliklerin bir fonksiyonu olarak bağımlı değişkendeki değişiklikleri modellemek için kullanılmaktadır. Standart regresyon yöntemleri parametrik modeller sınıfında olup, modellerin oluşturulması aşamasında güçlü parametrik varsayımları içermektedir. Diğer yandan parametrik olmayan regresyon yöntemlerde ise varsayımların geçerliliğine veya sağlanmasına ilişkin herhangi bir şey yapılmaz. Parametrik ve parametrik olmayan regresyon yöntemleri arasında yer alan yarı parametrik regresyon yöntemleri, çok genel varsayımları (örneğin bağımlı ve bağımsız değişken arasındaki ilişkinin belirli bir şekil ile sınırlı olmaksızın bazı durumlarda sürekli veya düzgün olması gibi) dikkate alarak iki yöntem arasında bir denge oluşturur.

Zaman serileriyle modelleme çoğunlukla ekonomik verileri içerir. Ekonomik zaman serileri genellikle, çok sayıda iç/dış şoklar ve diğer değişkenlerden etkilenmektedir. Ekonomik hayat sürekli olmasına rağmen, savaşlar, doğal olaylar, ekonomik, sosyal veya siyasi krizler gibi bazı olaylar genellikle, o dönemlere ait toplanan verilerde ani ve yapısal değişikliklere sebep olur. Zaman serilerinde bu yapısal değişiklikleri analize dâhil etmek ve zaman serilerinin sürekliliğini sağlamak için spline fonksiyonları kullanılır. Yani yapısal değişikliği olduğu ekonomik bir modeli düzgün bir şekilde kurmak için spline fonksiyonu uygun bir yöntemdir.

Spline fonksiyonları, interpolasyon, tahmin ve eğri uydurmada, türev ve kısmi türev denklemlerinde, istatistik alanında ve ayrıca jeoloji, meteoroloji, biyomedikal

bilimler, sosyoloji ve ekonomi vb. gibi uygulamalı arařtırmalarda kullanıldıđı modern sayısal analizin birok alanında nem kazanmıřtır.

İstatistikte splinelerin rol, verilere uydurulacak eđriyi semek iin uygun bir fonksiyon sınıfı oluřturmaktır. Uydurulacak eđri, iki deđiřken arasındaki iliřkide rassal dalgalanmayı ok fazla yansıtmaksızın iliřkinin temel zelliklerini gstermelidir. İstatistikte splinelerin bir uygulaması da, standart regresyon tekniklerin kullanılarak veriye dzgn eđriler tahmin edilmesidir. Spline, regresyon analizlerinde kullanmak iin faydalı ve uygun bir fonksiyon tipidir. Spline regresyon, eđrinin biiminin nceden belirlenmediđi durumda parametrik olmayan bir tekniktir ancak aynı zamanda eđriyi tahmin etmek iin dođrusal regresyon teknikleri kullanıldıđı zaman parametrik bir tekniktir. Bylece istatistikte splineler sıklıkla yarı parametrik bir teknik olarak adlandırılır.

Bu alıřmada, Trkiye'nin 1923-2010 yılları arasındaki ihracatın ithalatı karřılama oranları dađılımının (Bkz. Ek:1) modellenmesinde spline regresyon modelleri kullanılmıřtır. Dađılımda dđm noktalarının sayısı ve konumunun belirlenmesi iin dođrusal spline regresyonların kullanıldıđı belirlilik katsayısına (R^2) dayalı bir yntem nerilmiřtir. Spline regresyon modellerinin kurulabilmesi iin nerilen yntemle gerekli olan dđm noktaları belirlenmiřtir. R^2 'nin kullanılmasının sebebi, bađımsız deđiřken ile bađımlı deđiřken arasında yapısal deđiřiklikler olduđunda R^2 deđerleri de deđiřiklik gsterir. Buradan hareketle, R^2 deđerlerindeki nemli deđiřikliklere (artıř veya azalıřlara) gre dđm noktaları belirlenmiřtir.

Yapılan tez alıřması beř blmden oluřmaktadır. Birinci blmde, spline ile ilgili dnyada ve Trkiye'de yapılan alıřmaların literatr arařtırması yapılmıř, spline kavramı ve fonksiyonları, spline hesaplamaları, spline gsterimleri ve eřitleri ele alınmıřtır.

İkinci blmde, spline regresyonda dđm noktasının belirlenmesine iliřkin yapılan alıřmalarla ilgili literatr arařtırması yapılmıř, spline regresyonda dđm noktası tanımı, sabit ve deđiřken dđm noktası kavramları ve dđm noktasının belirlenmesinde kullanılan yntemler ele alınmıřtır.

nc blmde, spline regresyon modellerle yakın iliřkili glge deđiřkenli regresyon ve kesikli paralı regresyon, uygulamada kullanılan dođrusal, kareli ve kbik

spline regresyon modelleri ve farklı yazarlar tarafından geliştirilen kübik spline regresyon modelleri ele alınmıştır.

Dördüncü bölümde, spline regersyonda model seçimiyle ilgili kriterler incelenmiştir.

Beşinci bölümde ise, uygulama olarak ihracatın ithalatı karşılama oranları dağılımında, önerilen yöntemle bilinmeyen düğüm noktası sayısı ve konumlarının nasıl belirlendiği ve belirlenen düğüm noktalarına göre spline regresyon modellerinin kurulması ele alınmıştır.

Son olarak da çalışmanın sonunda elde edilen sonuçlar belirtilmiş ve bir değerlendirme yapılmıştır.

BİRİNCİ BÖLÜM

SPLİNE KAVRAMI

1.1. LİTERATÜR ARAŞTIRMASI

1.1.1. Dünyada Yapılan Çalışmalar

Reinsch (1967)¹, kübik splineleri kullanarak düzgünleştirme için (smoothing) bir algoritma geliştirmiştir.

Fuller (1969)², üretim yüzeylerini (production surfaces) ve zaman serisi trendlerini tahmin etmek için parçalı polinomların kullanımını tanımlamış ve göstermiştir. Fuller'in çalışmasında parçalı polinom eğrileri süreklidir. Aynı zamanda parçalı polinomların birleşme noktalarında türevleri de süreklidir. Fuller çalışmasında bu tür polinomları “Grafted Polinom” olarak adlandırmıştır.

Greville (1969)³, polinom spline fonksiyonlarının temel özelliklerini ele almış ve doğal spline fonksiyonlar üzerinde durmuştur. Splineların gösterimi, splineların değişkenlik (varyasyonel) özellikleri, doğrusal fonksiyonlarla ilişkili en iyi tahminler ve düzgünleştirici splinelar da tartışılmıştır.

Gujarati (1970)⁴, iki parçalı doğrusal regresyonda katsayılar arasındaki eşitlik durumlarını belirlemek için chow testi ve gölge değişken yaklaşımını karşılaştırmıştır. Uygulama, 1946-1954 ve 1955-1963 yılları arasındaki gelir ve tasarruf ilişkisi incelenmiş ve her iki metotta da benzer sonuçlar elde edilmiştir.

Poirier (1973)⁵, yapısal değişikliğin belirlenmesi amacıyla kübik spline fonksiyonların regresyonda kullanımını göstermiş ve kurulan regresyon modelini en küçük kareler yöntemiyle tahmin etmiştir. Uygulama örneğinde 1911-1971 yılları boyunca yapılan Indianapolis 500 yarışlarında birinci gelen yarışmacıların hızları ele alınmıştır. I. Dünya Savaşı (1917-1918) ve II. Dünya Savaşı (1942-1945) boyunca

¹ Christian H. Reinsch, “Smoothing by Spline Functions”, *Numerische Mathematik*, 10(3), 1967, 177-183.

² Wayne A. Fuller, “Grafted Polynomials as Approximating Functions”, *Australian Journal of Agricultural Economics*, 13(1), 1969, 35-46.

³ T.N.E. Greville, “Introduction to Spline Functions”, (Ed. T.N.E. Greville), *In Theory and Applications of Spline Functions*, (pp. 1-35), Academic Press, New York 1969.

⁴ Damador Gujarati, “Use of Dummy Variables in Testing for Equality Between Sets of Coefficients in Two Linear Regressions”, *The American Statistician*, 24(1), 1970, 50-52.

⁵ Dale J. Poirier, “Piecewise Regression Using Cubic Spline”, *Journal of the American Statistical Association*, 68(343), 1973, 515-524.

yarıřlar yapılmadıđından, bu savařların yarıřların performansını etkileyip etkilemediđi incelenmiřtir. Analiz sonucu savař dönemlerinde yapısal bir deđiřiklik olduđu yani her iki savařın da yarıř performansları üzerinde etkisi olduđu tespit edilmiřtir.

Poirier ve Garber (1974)⁶, Kore ve Vietnam Savařlarının ABD’deki sekiz havayolu řirketinin kar oranlarında önemli etkisi olup olmadıđını incelemek için dođrusal spline fonksiyonlarını kullanmıřtır. Yani yapılan regresyon analizinde matematiksel olarak zaman boyutu spline fonksiyonlarla gösterilmiřtir. Regresyon analizi sonucu her iki savařın da firmaların kar oranları üzerinde olumlu etkisinin olduđu tespit edilmiřtir.

Wold (1974)⁷, deneysel verilerin eđri tahmininde, polinom ve üstel fonksiyonlara alternatif olarak kullanılabilecek spline fonksiyonların avantajlarını göstermiřtir. Düşüm noktası sayısı ve konumunu belirlenmesinde dikkat edilecek hususları dört maddede anlatmıř ve kübik/kareli spline fonksiyonların uygulamalarını yapmıřtır.

Poirier (1975a⁸, 1975b⁹), ekonomik uygulamalarda iki boyutlu spline fonksiyonları (bilinear spline) çalışmıřtır. Bireylerin yař ve eğitim durumlarına göre aldıkları ücretlerdeki deđiřim incelenmiřtir. Aynı yıl yaptıđı başka bir çalışmada da, Cobb-Douglas üretim fonksiyonuna spline fonksiyonu uyarlamıřtır.

Barth, A. Kraft ve J. Kraft (1976)¹⁰, likitide tuzađını test etmek için kübik spline fonksiyonlarını kullanmıřtır.

Buse ve Lim (1977)¹¹, kübik spline regresyonun kısıtlandırılmıř en küçük kareli regresyonun özel bir durumu olduđunu göstermiřtir. Kısıtlandırılmıř en küçük kareli regresyonun avantajları anlatılırken, teorik olarak iki yöntemin eřit olduđunun ispatı da gösterilmiřtir.

⁶ Dale J. Poirier and Steven G. Garber, “The Determinants of Aerospace Profit Rates”, *Southern Economic Journal*, 41(2), 1974, 228-238.

⁷ Svante Wold, “Spline Functions in Data Analysis”, *Technometrics*, 16(1), 1974, 1-11.

⁸ Dale J. Poirier, “On the Use of Bilinear Splines in Economics”, *Journal of Econometrics*, 3(1), 1975, 23-34.

⁹ Dale J. Poirier, “On the Use of Cobb-douglas Splines”, *International Economic Review*, 16(3), 1975, 733-744.

¹⁰ James Barth, Arthur Kraft and John Kraft, “Estimation of the Liquidity Trap Using Spline Functions”, *The Review of Economics and Statistics*, 58(2), 1976, 218-222.

¹¹ A. Buse and L. Lim, “Cubic Splines as a Special Case of Restricted Least Squares”, *Journal of the American Statistical Association*, 72(357), 1977, 64-68.

Gnad (1977)¹², regresyon modellerinde parametrik splinelerin kullanımını göstermiştir. Parametrik splinelerin teorik alt yapısını ve matris gösterimlerini inceledikten sonra, matrislerin en küçük kareler yöntemiyle çözümünü ele almıştır. Uygulama örneğinde doğrusal, kareli ve kübik spline regresyon modellerini kullanmıştır.

MecNail, Trussell ve Turner (1977)¹³, splinelerin demografik verilerde kullanımını göstermiştir. Bu amaçla, kübik ve kuintik (beşinci dereceden) spline fonksiyonlarını kullanarak İtalya'daki kadınların yaşa göre doğum oranlarını incelemiştir.

Suits, Mason ve Chan (1978)¹⁴, spline fonksiyonların en küçük kareli regresyonla herhangi bir istatistiksel paket programında nasıl tahmin edileceğini anlaşılır bir şekilde göstermiştir. Çalışma özellikle kübik spline regresyon üzerine yoğunlaşmıştır. İki düğüm noktalı doğrusal ve kübik spline regresyonun teorik alt yapısı anlatıldıktan sonra, düğüm noktasının ikiden fazla olduğu ve bağımsız değişkenlerin birden fazla olduğu durumlar için kübik spline regresyonun genelleştirilmiş hali de verilmiştir.

Smith (1979)¹⁵, kesikli (truncated) üstel taban veya (+) fonksiyonlarını kullanarak sabit düğüm noktalı spline regresyonların teorisini ele almış ve farklı dereceli polinom parçalarından oluşan karma spline regresyon modellerini göstermiştir. Spline regresyon modelleri sabit düğüm noktalı alındığından, bu modellerin herhangi bir istatistik paket programında en küçük kareler yöntemiyle tahmin edilebileceğini belirtmiştir.

Robb (1980)¹⁶, düğüm noktalarının 0, 3, 6, 9 ve 12. ay olarak belirlendiği kareli spline fonksiyonlarla mevsimselliği hesaplamıştır. 1956-1975 yılları arasında Kanada'daki kadınların işsizlik oranı ile asal yaştaki çalışabilir erkeklerin işsizlik oranı arasındaki ilişkiyi incelemiştir. Yetişkin kadınların (25 yaş ve üstü) aylık işsizlik oranı bağımlı değişken alınırken asal yaştaki çalışabilir erkeklerin (25-44) işsizlik oranı ve

¹² W. Gnad, "Regression by Spline Functions", *Empirical Economics*, 2(2), 1977, 69-77.

¹³ Donald R. MecNail, T. James Trussell and John C. Turner, "Spline Interpolation of Demographic Data", *Demography*, 14(2), 1977, 245-252.

¹⁴ Daniel B. Suits, Andrew Mason and Louis Chan, "Spline Functions Fitted by Standart Regression Methods", *Review of Economics and Statistics*, 60(1), 1978, 132-139.

¹⁵ Patricia L. Smith, "Splines as a Useful and Convenient Statistical Tool", *The American Statistician*, 33(2), 1979, 57-62.

¹⁶ A. Leslie Robb, "Accounting for Seasonality with Spline Functions", *The Review of Economics and Statistics*, 62(2), 1980, 321-323.

dört parçalı kareli spline fonksiyonda kullanılan mevsimsel değişkenler (3, 6, 9. ay iç düğüm noktaları) bağımsız değişken olarak alınmıştır. Yapılan analizde düğüm noktaları anlamlı bulunmuş ve dört parçalı kuadratik spline regresyonla bu ilişkideki mevsimsel örüntü (seasonal pattern) tespit edilmiştir.

Prunty (1983)¹⁷, birçok verinin parçalı doğrusal fonksiyonlarla uygun bir şekilde modellenebileceğini açıklamıştır. Parçalı doğrusal fonksiyonların genel biçimlerini ve özelliklerini göstererek deneysel verilerle örnek bir uygulama örneği vermiştir.

Wegman ve Wright (1983)¹⁸, parametrik ve parametrik olmayan spline gösterimlerini ele aldıktan sonra, splinelerin istatistik alanındaki kullanımlarını araştırmıştır.

Eubank (1984)¹⁹, splinela ilgili literatür araştırmasından sonra splinelerin regresyon analizindeki önemini göstermiştir. Polinom regresyon ve spline regresyonların teorisi hakkında bilgi verip, spline regresyonda sabit ve değişken düğüm noktası durumlarını ele almıştır.

Kaishev (1984)²⁰, spline regresyon problemlerinin çözümüne yönelik SADP adlı bir bilgisayar paket programı geliştirmiştir.

Anderson (1985)²¹, Detroit bölgesi nüfus yapısında 1965-1985 yılları arasında meydana gelen değişiklikleri belirlemek için nüfus sayımı verileri ve kübik spline regresyon modellerini kullanmıştır.

Devlin ve Weeks (1986)²², lojistik regresyon modellere spline fonksiyonların uyarlamasını göstermiştir. Bunun için Stone'nun uyarlamalı (additive) splineleri kullanılmıştır.

¹⁷ Lyle Prunty, "Curve Fitting with Smooth Functions that are Piecewise-Linear in The Limit", *Biometrics*, 39(4), 1983, 857-866.

¹⁸ Edward J. Wegman and Ian W. Wright, "Splines in Statistics", *Journal of the American Statistical Association*, 78(382), 1983, 351-365.

¹⁹ Randall L. Eubank, *Approximate Regression Models and Splines*, The Office of Naval Research, (Technical Report No. 180), Southern Methodist University, USA/Texas 1984.

²⁰ V. K. Kaishev, "A Computer Package for Solving Spline Regression", T. Havranek, Z. Sidak and M. Novak (Ed.), *COMPSTAT 1984: Proceedings in Computational Statistics* (pp. 409-414), Wien: Physica-Verlag 1984.

²¹ John E. Anderson, "The Changing Structure of a City: Temporal Changes in Cubic Spline Urban", *Journal of Regional Science*, 25(3), 1985, 413-425.

Marcus (1987/1988)²³, yapısal değişikliğin olduğu zaman serisi verilerinde spline fonksiyonları ve gölge değişkenlerin kullanımını incelemiştir. Gölge değişkenlerin aksine spline fonksiyonların kullanımının daha avantajlı olduğunu belirtmiştir.

Allsop ve Weisberg (1988)²⁴, 1984 yılında başkanlık seçim kampanyası süresince seçmenlerin parti kimliğinde devam eden değişiklikler üzerine yaptıkları araştırmada, spline regresyon kullanımının bir örneğini göstermişlerdir. Araştırmada, 19 Haziran, 12 Eylül ve 10 Ocak olmak üzere üç düğüm noktası belirlenmiştir. Spline düğüm noktaları kullanılarak, kırılmaların olmadığı genel bir eğri elde etmek için düğüm noktalarının birleştirildiği dört doğrusal regresyon tahmin edilmiştir.

McDonald ve Studden (1990)²⁵, aracın yaptığı birikimli kilometreye göre otomotiv emisyonu (hidrokarbon, karbonmonoksit ve azot oksitleri) arasındaki ilişkiyi doğrusal spline regresyon ve doğrusal regresyon ile incelemiştir. Analiz sonucunda doğrusal spline regresyonun daha iyi sonuç verdiği belirlenmiştir.

Marsh ve Sindone (1991)²⁶, konutların bulundukları konumlarının konut satış fiyatları üzerindeki etkisini tahmin etmek için üç boyutlu polinom splineları kullanmıştır. İstatistiksel olarak anlamlı değişkenlerin seçiminde adımsal regresyon önerilmiştir.

Speyrer ve Ragas (1991)²⁷, 1971 ve 1986 yılları arasında New Orleans ve Louisiana’da satılan 2000 evi dikkate alarak, sel riski ve zorunlu deprem sigortasının emlak fiyatları üzerindeki etkisini incelemişlerdir. Şehir dışındaki ve şehir merkezindeki bölgeler de spline değişken olarak regresyon analizine dâhil edilmiştir.

²² Thomas F. Devlin and Barbara J. Weeks, “Spline Functions for Logistic Regression Modeling”, *SAS Institute Inc (Ed.), Proceedings of the Eleventh Annual SAS Users Group International Conference*, (pp. 646-649), Atlanta, Georgia 1986.

²³ Richard D. Marcus, “How to Deal with Structural Changes in The Data”, *The Journal of Business Forecasting Methods & Systems*, 6(9), Winter 1987/1988, 14-16.

²⁴ Dee Allsop and Herbert F. Weisberg, “Measuring Change in Party Identification in an Election Campaign”, *American Journal of Political Science*, 32(4), 1988, 996-1017.

²⁵ Gary C. McDonald and William J. Studden, “Design Aspects of Regression-Based Ratio Estimation”, *Technometrics*, 32(4), 1990, 417-424.

²⁶ Lawrence C. Marsh and Anthony B. Sindone, “A Calibration Technique for Estimating the Effect of Location on the Values of Residuals Properties”, *The Property Tax Journal*, 10(2), 1991, 261-276.

²⁷ Janet F. Speyrer and Wade R. Ragas, “Housing Prices and Flood Risk: An Examination Using Spline Regression”, *The Journal of Real Estate Finance and Economics*, 4(4), 1991, 395-407.

Horowitz, Loughran ve Savin (1996)²⁸, aylık ortalama dönüşler ve doğrusal spline regresyon yöntemlerini kullanarak beklenen dönüşler ile firma büyüklüğü arasındaki ilişkiyi incelemiştir. Yani beklenen dönüşlerin büyük firmalardan ziyade küçük firmalarda daha fazla olup olmadığı araştırılmıştır. Analiz sonucu firma büyüklüğü ve beklenen dönüş arasında sistematik bir ilişki olmadığı tespit edilmiştir.

Koo (1997)²⁹, süreksiz regresyon fonksiyonlarını tahmin etmek için doğrusal spline regresyonu kullanmıştır. Sıçrama noktaları yakın regresyon fonksiyonların tahmini için düğüm noktası birleştirme yöntemini sunmuştur.

Kalyanam ve Shively (1998)³⁰, düzensiz fiyat etkilerinin tahminine olanak sağlayan hiyerarşik bayezyen model çerçevesinde stokastik spline regresyon yaklaşımını kullanmıştır. Uygulamada beş farklı markadaki katı margarin ve dört farklı markadaki ton balığının düzensiz fiyat hareketleri tahmin edilmiştir.

Strawczynski (1998)³¹, farklı vergi dilimlerini hesaplamak için parçalı doğrusal spline regresyonu kullanmıştır.

Tarpey (2000)³², polinom regresyon ve spline regresyon modellerle şişelerin yüksekliği ve hacmi arasındaki ilişkiyi incelemiştir. Uygulamada parametrik spline kullanılmıştır.

Aggrey (2002)³³, tavukların büyüme eğrilerini belirlediği çalışmasında spline regresyon modeliyle doğrusal olmayan büyüme modellerini (Richards, Gompertz ve Logistic) karşılaştırmıştır. Analiz sonucunda, uyum iyiliği kriterlerinden R^2 (Belirlilik Katsayısı) değerlerine göre kıyaslama yapıldığında doğrusal olmayan büyüme modellerinin spline regresyon modelinden daha iyi olduğu belirlenmiştir.

²⁸ Joel L. Horowitz, Tim Loughran, and N. E. Savin, "A Spline Analysis of the Small Firm Effect: Does Size Really Matter?", *EconWPA, Econometrics*, 9, 1996, 1-26, Erişim Tarihi: 11.10.2011, <http://129.3.20.41/eps/em/papers/9608/9608001.pdf>.

²⁹ Ja-Yong Koo, "Spline Estimation of Discontinuous Regression Functions", *Journal of Computational and Graphical Statistics*, 6(3), 1997, 266-284.

³⁰ Kirthi Kalyanam and Thomas S. Shively, "Estimating Irregular Pricing Effects: A Stochastic Spline Regression Approach", *Journal of Marketing Research*, 35(1), 1998, 16-29.

³¹ Michel Strawczynski, "Social Insurance and The Optimal Piecewise Linear Income Tax", *Journal of Public Economics*, 69(3), 1998, 371-388.

³² Thaddeus Tarpey and John Holcomb, "Spline Bottles", *The American Statistician*, 54(2), 2000, 129-135.

³³ S. E. Aggrey, "Comparison of Three Nonlinear and Spline Regression Models for Describing Chicken Growth Curves", *Poultry Science*, 81, 2002, 1782-1788.

Unser (2002)³⁴, bilgisayar destekli tasarım ve bilgisayar grafiklerinde yaygın bir şekilde kullanılan splineların tıbbi görüntüleme uygulamalarında kullanımını araştırmıştır.

Muniz, Galindo ve Garcia (2003)³⁵, Barcelona Bölgesiyle ve Anderson (1982)³⁶ ise, ABD'nin 30 şehriyle ilgili nüfus yoğunluğu çalışmalarında kübik spline nüfus yoğunluk fonksiyonlarını kullanmıştır. Her iki çalışmada da Suits ve ark. (1978) tarafından geliştirilen kübik spline regresyon kullanılarak ağırlıklandırılmış en küçük kareli yöntemle tahmin edilmiştir. Kübik spline fonksiyonların üstel fonksiyonlara göre daha iyi sonuç verdiği tespit edilmiştir.

Brusilovskiy (2004)³⁷, SAS programını kullanarak iki değişkenli regresyonda bağımlı değişkenin modellenmesinde parçalı regresyonu temel alan bir yöntem önermiştir.

N. Breaz ve D. Breaz (2005)³⁸, veri analizinde spline fonksiyonların kullanımı üzerine çalışma yapmıştır. Spline interpolasyonu, parametrik en küçük kareli spline ve parametrik olmayan cezalı en küçük kareli splinelar incelenmiştir.

Guo ve White (2005)³⁹, kübik spline regresyonun ve Dunfield ve Read (1972)⁴⁰, kübik spline fonksiyonun kimya alanında uygulamasını göstermiştir.

Hurley, Hussey, McKeown ve Addy (2006)⁴¹, simülasyonla bir bağımlı ve bir bağımsız değişkenli beş veri yapısı oluşturarak her veri yapısı için doğrusal regresyon,

³⁴ Michael Unser, "Splines: A Perfect Fit for Medicine Imaging", Milan Sonka, J. Michael Fitzpatrick (Ed.), *Proceedings of the SPIE International Symposium on Medical Imaging* (pp. 225-236), Spie, San Diego 2002.

³⁵ Ivan Muniz, Anna Galindo and Miguel Angel Garcia, "Cubic Spline Population Density Functions City Delimitation: The Case of Barcelona", *Urban Studies*, 40(7), 2003, 1303-1321.

³⁶ John E. Anderson, "Cubic-Spline Urban-Density Functions", *Journal of Urban Economics*, 12(2), 1982, 155-167.

³⁷ Eugene Brusilovskiy, "The Piecewise Regression Model as a Response Modeling Tool", *Conference Proceedings: NESUG 2004*, Erişim Tarihi: 08.01.2012, <http://www.nesug.org/html/Proceedings/nesug04/an/an09.pdf>.

³⁸ Nicoleta Breaz and Daniel Breaz, "On the Spline Functions in Data Analysis Framework", *International Conference on Theory and Applications of Mathematics and Informatics ICTAMI2005* (ss. 197-210), Acta Universitatis Apulensis, ISSN 1582-5329, Alba Lulia, Romania 2005, pp. 197-210.

³⁹ Qingzhi Guo and Ralph E. White, "Cubic Spline Regression for the Open-Circuit Potential Curves of a Lithium-Ion Battery", *Journal of The Electrochemical Society*, 152(2), 2005, 343-350.

⁴⁰ Lawrence G. Dunfield and John F. Read, "Determination of Reaction Rates by the Use of Cubic Spline Interpolation", *The Journal of Chemical Physics*, 57(5), 1972, 2178-2183.

⁴¹ Deborah Hurley, James Hussey, Robert McKeown and Cheryl Addy, *An Evaluation of Splines in Linear Regression* [Poster], *SAS Conference Proceedings: SAS Users Group International 31*, San Francisco (March 2006).

polinom regresyon (kübik ve kareli) ve spline regresyon (doğrusal, kareli ve kübik) olmak üzere altı regresyon modeli kurmuştur. Yapılan analiz sonuçlarına göre en uygun modellerin spline regresyon modelleri olduğu tespit edilmiştir.

Doorn, Verhoef ve Bijmolt (2007)⁴², tutum ve davranış arasındaki ilişkinin her zaman doğrusal olmadığını, aşırı tutumlardan dolayı tutum ve davranış arasında doğrusal olmayan bir ilişkinin de söz konusu olduğunu belirtmiştir. Bu amaçla, doğrusal, kareli, üstel ve spline model olmak üzere dört farklı model kurulmuş ve karşılaştırma yapılmıştır. Modeller R^2 (Belirlilik Katsayısı), Akaike Bilgi Kriteri (AIC) ve Bayezyen Bilgi Kriteri (BIC)'ne göre karşılaştırılmış ve spline modelin en etkin model olduğu tespit edilmiştir.

García, Coronel, Martinez ve Duque (2007)⁴³, Meksika'da bir firmanın 1 yıllık süt satışı verilerini analiz etmek ve öngöründe bulunmak için doğrusal, kareli, kübik regresyon ve kübik spline regresyon modellerini kullanmıştır. Verilerin dağılımı doğrusal, kareli ve kübik regresyon modellerle tek bir parçada, kübik spline regresyon model için de iki düğüm noktası belirlenerek sürekli üç parçada tahmin edilmiştir. Modeller R^2 (Belirlilik Katsayısı) ve MSE (Hata Kareler Ortalaması-Mean Square Error) değerlerine göre karşılaştırılmış ve kübik spline regresyon modelin en etkin model olduğu tespit edilmiştir.

Mulla (2007)⁴⁴, istatistiksel teknik olarak faydalı bir araç olan spline regresyon modellerin klinik araştırmalarında kullanımı gösterilmiştir. Uygulamada kareli spline regresyon kullanılarak serum albümin ve hastane ölüm oranı riski arasındaki ilişki araştırılmıştır.

Kulkarni ve Narendranath (2008)⁴⁵, polinom regresyon ve spline regresyon modellerle tarımsal üretim tahmini yapmıştır. Bu amaçla, Andhra Pradesh eyaletinde 1956-1998 yılları arasındaki pirinç ve tahıl üretim verileri kullanılarak polinom ve

⁴² Jenny van Doorn, Peter C. Verhoef. and Tammo H. A. Bijmolt, "The Importance of Non-Linear Relationships Between Attitude and Behaviour in Policy Research", *Journal of Consumer Policy*, 30(2), 2007, 75-90.

⁴³ Jorge L. García, Nancy A. Coronel, Erwin A. Martínez and Julian A. Duque, "Using Spline for Demand Forecasting", V. G. Almanza (Ed.), *Second International Congress in Engineering and Technology: October 2008*, (pp. 37-56), Chihuahua, Mexico: Ciudad Juarez 2007.

⁴⁴ Z. D. Mulla, "Spline Regression in Clinical Research", *West Indian Med J.*, 56(1), 2007, 77-79.

⁴⁵ B.S. Kulkarni and V.V. Narendranath, "Spline Regression Models for Forecasting of Agricultural Production", *The Journal of Research Angrau*, 36(2&3), 2008, 28-38.

spline regresyon modelleri kurulmuştur. Kurulan modellerle 1999 ve 2002 yılları için tahminde bulunulmuş ve spline regresyon modellerin daha iyi sonuç verdiği tespit edilmiştir.

Chau, Bonfiglioli, Chey ve Bauman (2009)⁴⁶, fiziksel aktivite, şişmanlık ve tütün kullanımı gibi sağlık konularına Avustralya medyasının verdiği önemi ölçmek için doğrusal spline regresyon modelini kullanmıştır. Bu amaçla, 1986-2006 yılları arasında fiziksel aktivite, şişmanlık ve sigara kullanımı gibi konularda yapılan/yayınlanan tüm çalışmaların yıllara göre dağılımı elde edilmiştir. 1991, 1996 ve 2001 yılları düğüm noktası alınarak üç doğrusal spline regresyon modeli kurulmuştur.

Chernobai, Reibel ve Carney (2009)⁴⁷, Los Angeles bölgesinde yapımı yeni biten karayolunun emlak fiyatları üzerindeki etkisini araştırmışlardır. Bu amaçla, spline regresyon modellerini kullanarak evlerin karayolundan uzaklığının fiyatlar üzerindeki etkisini incelemişlerdir.

Hirvonen (2009)⁴⁸, kantil ve spline regresyon yöntemleriyle kuşaklararası kazancın sürekliliğinde, bireylerin eğitimi, yetenekleri ve sağlığı arasında fark olup olmadığını araştırmıştır. Spline regresyon analizi sonucunda babanın kazancının oğlunun kazancını doğrudan etkilediği ve etkisinin eğitimden daha fazla olduğu belirlenmiştir.

Xi, Lee, Teng ve Lin (2009)⁴⁹, Taiwan’da bankacılık hizmetlerinde performans nitelik düzeyi ve müşteri memnuniyeti arasındaki ilişkiyi modellemek için istatistiksel bir teknik olarak parçalı regresyonun kullanımını göstermişlerdir. Önerilen yöntem hem kolay hem de geleneksel yöntemlere göre daha etkindir.

⁴⁶ Josephine Chau, Catriona Bonfiglioli, Tien Chey, Adrian Bauman, “The Cinderella of Public Health News: Physical Activity Coverage in Australian Newspapers, 1986-2006”, *Australian and New Zealand Journal of Public Health*, 33(2), 2009, 189-192.

⁴⁷ Ekaterina Chernobai, Michael Reibel and Michael Carney, “Nonlinear Spatial and Temporal Effects of Highway Construction on House Prices”, *The Journal of Real Estate Finance and Economics*, 42(3), 2009, 348-370.

⁴⁸ Lalaina Hirvonen, *The Effect of Children on Earnings Using Exogenous Variation in Family Size: Swedish Evidence*, Working Paper Series 2, Swedish Institute for Social Research, Stockholm University 2009.

⁴⁹ Jun Yang Xi, Chung-Liang Lee, Hsiao-Lin Teng and Shao-Bin Lin, “Piecewise Regression: A Tool for Understanding The Relationship between the Attribute-Level Performance and Customer Satisfaction”, *International MultiConference of Engineers and Computer Scientists*, (March 2009), Hong Kong.

Sarvida ve Baguio (2010)⁵⁰, zaman serisi verileri trendinde yapısal değişiklik olduğu zaman farklı bir metot olarak spline regresyon modellerin etkinliğini araştırmıştır. Doğrusal ve kareli spline regresyon modelleri kurulmuş ve modellerin karşılaştırılmasında RMSE (Root Mean Squared Error) ve AIC (Akaike Information Criterion) kullanılmıştır.

Sun ve Liu (2010)⁵¹, spline regresyonla, online açık artırmada satıcının itibarının(reputation) fiyatları nasıl etkilediğini incelemiştir. Açık artırma fiyatı ve satıcının itibarı arasında doğrusal olmayan bir ilişki olduğu düşünülmüştür.

Marcel, Răzvan ve Cristian (2011)⁵², para ve maliye politikalarında emlak ve bina vergilerinin önemli bir yeri olduğunu belirtmiştir. Bu amaçla, yerel yönetimlerin vergi politikalarını gerçekleştirmesinde spline fonksiyonların kullanımını göstermiştir. Oradea Belediyesi'nin 2009 Şubat ayına ait emlak vergisi ve bina vergisi göstergeleri alınarak bir aylık veri oluşturulmuştur. Bu veride her hafta bir düğüm noktası alınarak doğrusal spline regresyon modeli kurulmuş ve 2009 Mart ayı için tahminlerde bulunulmuştur.

1.1.2. Türkiye’de Yapılan Çalışmalar

Hakverdi (1995)⁵³, parçalı doğrusal regresyon analizinde düğüm noktalarının belirlenmesini araştırmıştır. Bu amaçla, simülasyonla oluşturulan gözlem setinde düğüm noktalarının belirlenmesinde R^2 (Belirlilik Katsayısı), Cp İstatistiği ve Cook Uzaklığı yaklaşımı kullanılmıştır.

Koldere (1997)⁵⁴, kesikli doğrusal spline regresyonu kullanarak 1968-1965 yılları arasında Gayri Safi Milli Hâsıla’ya bağlı olarak Türkiye’nin ithalatında yapısal değişim olup olmadığını incelemiştir. Düğüm noktası olarak önce 1980 yılı daha sonra da 1990 yılı alınmıştır. Farklı kırılma noktalarına göre kurulan spline regresyon modelinde eğim

⁵⁰ Rex L. Sarvida and Carolina B. Baguio, “On the Efficiency of Spline Regression for Time Series Data” [Bildiri], *11th National Convention on Statistics*, (September 2010), Mandaluyong City: EDSA Shangri-La Hotel.

⁵¹ Chia-Hung Sun and Kang E. Liu, “Information Asymmetry and Small Business in Online Auction Market”, *Small Business Economics*, 34(4), 2010, 433-444.

⁵² Boloş Marcel, Pop Răzvan and Otgon Cristian, “Using Spline Functions for the Substantiation of Tax Policies by Local Authorities”, *Annals of the University of Oradea: Economic Science*, 1(1), 2011, 336-347.

⁵³ Kazım Hakverdi, *Parçalı Lineer Regresyon*, (Yayımlanmamış Yüksek Lisans Tezi), Gazi Üniversitesi Fen Bilimleri Enstitüsü, Ankara 1995.

⁵⁴ Yasemin Koldere, *Spline Fonksiyonları ve Yapısal Değişim: Türkiye İthalatı Üzerine Bir Uygulama*, (Yayımlanmamış Yüksek Lisans Tezi), Marmara Üniversitesi Sosyal Bilimler Enstitüsü, İstanbul 1997.

farkı anlamlı olduğundan iki dönem için ithalat fonksiyonunun anlamlı olduğu tespit edilmiştir.

Akkartal (2000)⁵⁵, kübik spline regresyon modeliyle Türkiye’de kişi başına düşen otomobil miktarlarını incelemiştir. Bu amaçla, Türkiye’nin 1960-1994 yılları arasındaki kişi başına düşen otomobil miktarları verisi, 1971 ve 1980 düğüm noktalarına göre Suits ve ark. (1978) tarafından geliştirilen kübik spline regresyon modeliyle analiz edilmiştir.

Genç, Kınacı, Oturanç, Kurnaz, Bilir ve Özbaltı (2002a⁵⁶, 2002b⁵⁷), günlük ortalama güneş radyasyon verilerini analiz etmek için Poirier (1973) tarafından geliştirilen kübik spline regresyonu kullanmıştır. Bu amaçla, İzmir ilinin 1994-1998 yılları arasındaki güneş ışını verileri kübik spline regresyon modelle analiz edilmiş ve kurulan modelin güvenilirliği test edilmiştir.

Yerdelen (2002)⁵⁸, iktisadi zaman serilerinde doğrusal spline fonksiyonlarını gecikmeli modellerle birlikte ele almıştır. Bu amaçla, Türkiye’nin 1970-1999 yılları arasındaki istihdam değerlerine bağlı olarak Gayri Safi Milli Hâsıla’daki değişimi incelemiştir. 1981 öncesi ve sonrası olmak üzere yapılan incelemede, ikinci dönemde istihdamın bir ve iki dönem gecikmeli değerlerinin de GSMH üzerinde etkili olduğu tespit edilmiştir.

Kurnaz, Kınacı, Oturanç, Genç ve Bilir (2003)⁵⁹, günlük ortalama sıcaklık verilerinin analizi için kübik spline tipi bir model kullanmıştır. Bu amaçla, Konya ili Akşehir ilçesinin 1997-2002 yılları arasındaki sıcaklık verileri kübik spline fonksiyonlarla analiz edilmiştir.

⁵⁵ Erkut Akkartal, *Splayn Fonksiyonları ve Türkiye’de Kişi Başına Düşen Otomobil Miktarları Üzerine Bir Uygulama*, (Yayımlanmamış Doktora Tezi), İstanbul Üniversitesi Sosyal Bilimler Enstitüsü, İstanbul 2000.

⁵⁶ Aşır Genç, Galip Oturanç, İsmail Kınacı ve Aydın Kurnaz, “Güneş Radyasyon Verilerinin Kübik Spline Fonksiyonları Yardımıyla Modellenmesi” [Bildiri], *Hacettepe Üniversitesi İstatistik Günleri Sempozyumu*, (Mayıs 2002), Ankara.

⁵⁷ Aşır Genç, İsmail Kınacı, Galip Oturanç, Aydın Kurnaz, Şefik Bilir ve Necdet Özbaltı, “Statistical Analysis of Solar Radiation Data Using Cubic Spline Functions”, *Energy Sources*, 24(12), 2002, 1131-1138.

⁵⁸ Ferda Yerdelen, “Gecikmeli Modellerde Spline Fonksiyonun Kullanımı”, *İktisat Fakültesi Mecmuası*, 52(2), 2002, 47-64.

⁵⁹ Aydın Kurnaz, İsmail Kınacı, Galip Oturanç, Aşır Genç ve Şefik Bilir, “Günlük Sıcaklık Verilerinin Kübik Spline Fonksiyonları Yardımıyla Modellenmesi”, H. K. Öztürk, M. Atılğan, A. Yılcı ve Ö. Atalay (Haz.), *I. Ege Enerji Sempozyumu ve Sergisi: 22-24 Mayıs 2003*, (ss.464-468), Pamukkale Üniversitesi Mühendislik Fakültesi, Denizli.

Şener (2006)⁶⁰, iki değişkenli regresyon analizinde regresyon doğrusunun bir veya birden çok noktada ani değişim gösterdiği durumlarda, kırılma noktaları ve parametrelerin tahmin edilmesinde SAS programını kullanmıştır. Bu amaçla, uygulamada simülasyonla oluşturulan tek kırılma noktalı iki doğrulu ve iki kırılma noktalı üç doğrulu veriler için regresyon modelleri kurulmuştur.

Şahin (2009)⁶¹, kesikli ve sürekli parçalı regresyonların tarımsal alanlarda yapılan çalışmalarda kullanımını göstermiştir. Tarımsal uygulamalarla ilgili dokuz farklı veri setine parçalı regresyonlar uygulanmış ve Gompertz, Logistic, Schnute, Richard, Orskov, Gamma, Üssel gibi klasik modellerle karşılaştırılmıştır. Kübik spline regresyonun diğer modeller göre daha iyi sonuç verdiği tespit edilmiştir.

Demirel (2010⁶², 2011⁶³), İzmir ili sıcaklık verilerini doğrusal spline regresyon modelleriyle incelemiştir. Bu amaçla, Devlet Meteoroloji Enstitüsü'nden alınan İzmir iline ait 34 yıllık sıcaklık verileri, düğüm noktası konumlarının bilinmesi ve bilinmemesi durumları için ayrı ayrı incelenmiştir.

Berberoğlu (2011⁶⁴, 2012⁶⁵), kübik spline regresyon ve kısıtlandırılmış en küçük kareli regresyon modelleriyle Türkiye'nin 1987-2008 yılları boyunca kriz dönemlerinde döviz kuru oranlarındaki yapısal kırılmaları incelemiştir.

1.2. İNTERPOLASYON KAVRAMI

İnterpolasyon kelimesi elementer anlamda bir fonksiyonun tablo halinde verilmiş değerlerinden hareketle, bu fonksiyonun bu aralıkta bilinmeyen değerlerinin hesaplanması işlemidir. İnterpolasyon bir $y(x)$ fonksiyonunun, $x_0, x_1, \dots, x_n$ noktalarındaki değerleri bilinirken, diğer noktalardaki değerlerinin bilinmesi

⁶⁰ Emral Şener, *Bazı Kırık Hat (Broken-Line) Regresyon Moellerinin SAS İstatistik Paket Programı Kullanılarak Çözümlemesi*, (Yayımlanmamış Yüksek Lisans Tezi), Mustafa Kemal Üniversitesi Fen Bilimleri Enstitüsü, Antakya 2006.

⁶¹ Mustafa Şahin, *Parçalı Regresyonlar ve Tarımsal Alanlarda Kullanımı*, (Yayımlanmamış Doktora Tezi), Kahramanmaraş Sütçü İmam Üniversitesi Fen Bilimleri Enstitüsü, Kahramanmaraş 2009.

⁶² Neslihan Demirel, "İzmir İli Sıcaklık Verilerinin Spline Regresyon ile Modellenmesi" [Bildiri], 7. İstatistik Günleri Sempozyumu, (Haziran 2010), ODTÜ/Ankara.

⁶³ Neslihan Demirel, "İzmir İli Sıcaklık Verilerinin Regresyon Eğrileri İle Modellenmesi", *İstatistik Araştırma Dergisi*, 8(2), 2011, 68-75.

⁶⁴ Bahar Berberoğlu and C. Necat Berberoğlu, "Modeling the Structural Shifts in Real Exchange Rate with Cubic Spline Regression (CSR). Turkey 1987-2008", *International Journal of Business and Social Science*, 2(17), 2011, ss.60-66.

⁶⁵ Bahar Berberoğlu, "A Note on the Predicted Responses in Cubic Splines", *Chiana-USA Business Review*, 11(1), 2012, 30-36.

problemidir⁶⁶. Yani interpolasyon, gerçek fonksiyonu bilmeksizin bilinen veri noktaları arasında bir fonksiyon değerini hesaplamak için kullanılır. İnterpolasyon yöntemleri, global ve parçalı interpolasyon olmak üzere iki temel kategoriye ayrılır⁶⁷.

1.2.1. Global İnterpolasyon

Bu yöntem, tüm veri noktalarının tahmin edilmesinde tek bir denklemin kullanılmasını temel almaktadır. Yani bu denklem genellikle yüksek dereceli polinom denklemdir. Bu yöntemle düzgün eğriler elde edilmesine rağmen, salınım etkisinin olması ve ara noktalarda aşırı sapma olmasından dolayı bu yöntemler bazı uygulamalar için pek uygun değildir⁶⁸. Yaklaşım yöntemleri içerisinde polinom yaklaşımları önemli bir yer tutmaktadır. Ancak, nokta sayısının artmasıyla elde edilecek interpolasyon polinomunun derecesi de artar. Ayrıca, noktalar dışında interpolasyon polinomunun gerçek fonksiyondan göstereceği büyük sapmalarda dikkate alındığında, polinom interpolasyonu yerine başka tekniklere ihtiyaç olduğu anlaşılmaktadır⁶⁹.

1.2.2. Parçalı İnterpolasyon

Bu yöntemlerde bilinen veri noktalarının her bir çifti arasında düşük dereceli polinomlar kullanılmaktadır. Birinci dereceli polinom kullanılırsa, doğrusal interpolasyon olarak adlandırılır. İkinci ve üçüncü dereceli polinomlar kareli ve kübik spline olarak adlandırılır. Daha yüksek dereceli spline ile daha düzgün eğriler elde edilir. Parçalı interpolasyonda salınım etkisi ve ara noktalarda aşırı sapma daha azdır⁷⁰.

1.3. SPLİNE TANIMI VE MAHİYETİ

Spline fonksiyonun en ilkel hali olan mekanik spline, verilen noktalardan geçmesi istenen eğriler çizmek için üzerine birkaç ağırlığın yerleştirildiği ince, elastik bir çubuktan oluşur⁷¹. Mühendisler, mimarlar ve grafikerler tarafından kullanılan mekanik spline, genelde düzgün eğriler çizmek için kullanılan ahşap bir çubuk veya metal bir şerittir⁷². Doğu İngiltere lehçesinden bir kelime olan⁷³ spline, sözlük tanımında bina

⁶⁶ Mustafa Bayram, *Nümerik Analiz*, Birsan Yayınevi Ltd. Şti., İstanbul 2009, s.93.

⁶⁷ C. J. C. Kruger, "Constrained Cubic Spline Interpolation for Chemical Engineering Applications", Korf Technology, Erişim Tarihi: 01.01.2012, <http://www.korf.co.uk/spline.pdf>, s.1.

⁶⁸ Kruger, p.1.

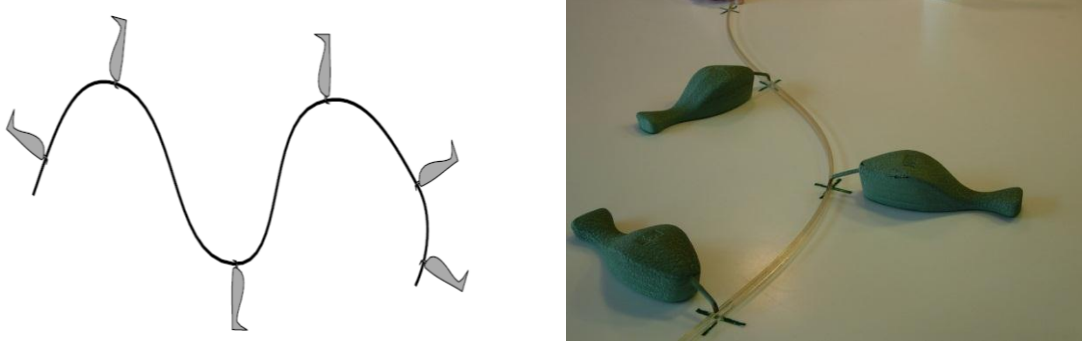
⁶⁹ Fatih Baş, *Spline Fonksiyonları ve Uygulamaları*, (Yayımlanmamış Yüksek Lisans Tezi), Sakarya Üniversitesi Fen Bilimleri Enstitüsü, Sakarya 2007, s.1.

⁷⁰ Kruger, s.1.

⁷¹ N. Breaz and D. Breaz, s.197.

⁷² Hongyu Wang, "A Survey on Spline Theory and Applications", Center for Visual Computing (CVC)

inşaatında kullanılan ahşap ince bir şerit demektir. Spline olarak adlandırılan uzun ince ahşap şeritler, belirlenen belli noktalar arasında düzgün eğriler çizmek için kullanılmaktaydı⁷⁴. Mekanik spline örneği Şekil 1.1’de gösterilmiştir⁷⁵.



Şekil 1.1. Mekanik Spline Örnekleri ve Ağırlıklar

Splineların matematik alanında ilk kullanımı Schoenberg tarafından gösterilmiştir. Matematikte kullanılan spline, özünde mekanik splinea benzerdir. Noktalar sayısal verileri ve ağırlıklar da polinomların katsayılarını göstermektedir⁷⁶. Splinelar genellikle, eğri veya doğru parçalarının birbirlerinden ayrı oluşturulduğu ve daha sonra bu parçaların birleştirildiği parçalı polinom olarak tanımlanır⁷⁷. Yani spline yönteminde, bağımlı değişkene göre bağımsız değişken aralığı alt parçalara bölünüp, düğüm noktası olarak adlandırılan bölünme noktalarında belirli durumlara göre birleştirilerek sürekli bir eğri elde edilir⁷⁸.

and Department of Computer Science, Stony Brook University, New York 2006, Erişim Tarihi: 12 Ekim 2011, http://www.cs.sunysb.edu/~wanghy/papers/hongyu_rpe.pdf, s.19.

⁷³ Richard H. Bartels, John C. Beatty and Brian A. Barsky, *An Introduction to Splines for use in Computer Graphics and Geometric Modeling*, Morgan Kaufmann Publishers, California 1987, s.vii.

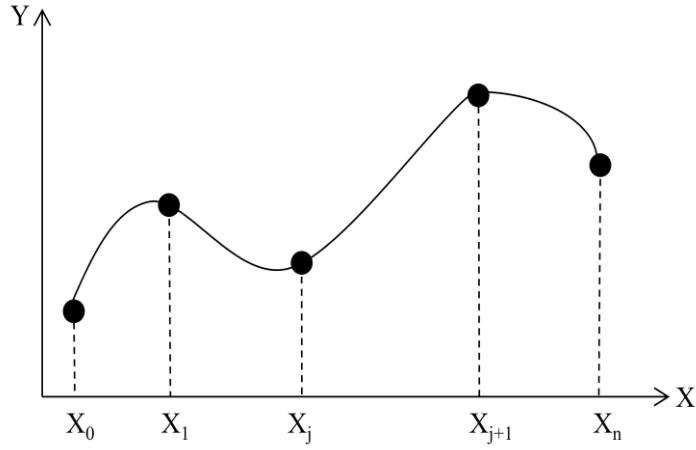
⁷⁴ Wegman ve Wright, s.351.

⁷⁵ Henry Packard Moreton, *Minimum Curvature Variation Curves, Networks, and Surfaces for Fair Free-Form Shape Design*, Ph. D. Disertation, University of California, Berkeley 1992, s.34.

⁷⁶ Isaac Jacob Schoenberg, "Contributions to the Problem of Approximation of Equidistant Data by Analytic Functions", *Quarterly of Applied Mathematics*, 4, 1946, ss.45–99 and 112–141.

⁷⁷ Thomas P. Ryan, *Modern Regression Methods*, John Wiley & Sons, New York 1997, s.334.

⁷⁸ George A. F. Seber and Alan J. Lee, *Linear Regression Analysis*, Second Edition, John Wiley & Sons, New Jersey 2003, s.173.



Şekil 1.2. Matematikte Kullanılan Spline Fonksiyonu

Splinelar genellikle d dereceli parçalı polinomlar olarak tanımlanmaktadır. Parçalı polinomların birleştiği noktalarda, fonksiyon değerleri ve $d-1$.türevleri aynıdır⁷⁹. Buna göre Şekil 1.2’de görüldüğü üzere, $x_0, x_1, \dots, x_n$ düğümlerine sahip d . dereceden bir spline fonksiyonu şu özelliklere sahiptir⁸⁰.

- a) Her bir $[x_j, x_{j+1}]$ aralığında $S(x)$ polinomunun derecesi, $\text{der}(S) \leq d$ ’dir.
- b) $[x_0, x_n]$ aralığında $S(x)$ ’in $(d-1)$. türevi süreklidir.

Splinelar, eğrilerin sürekli olmasının önemli olduğu uygulamalar için mükemmel bir araçtır⁸¹. Doğrusal olmayan, seyrek ve yapılandırılmamış verilerin tahmininde hassas bir model oluşturmak için çok geniş bir uygulama alanına sahiptir⁸².

Spline fonksiyonu, belli süreklilik özelliklerin sağlandığı durumlarda farklı polinom parçalarının düğüm noktalarında birleştirildiği parçalı bir polinomdur. Spline mümkün olan en düzgün parçalı polinomlardan oluşur⁸³. Splinelar genellikle, sürekli ve düzgün olma koşulunu sağlayan doğrular veya eğrilerdir⁸⁴. Splineların düz bir doğru olması zorunlu değildir. Spline ikinci veya üçüncü dereceden bir polinom olarak da tercih edilir ve fonksiyonun birinci ve ikinci türevlerinin sürekli olması sağlanır⁸⁵.

⁷⁹ Gregory and Serono, s.2.

⁸⁰ Galip Oturanç, Aydın Kurnaz, Mehmet E. Kiriş ve Yıldırım Keskin, *Sayısal Analiz*, 2.Baskı, Dizgi Ofset, Konya 2008, s.194.

⁸¹ Wang, s.18.

⁸² Kommineni, Basu and Vemuri, s.242.

⁸³ Randall L. Eubank, *Nonparametric Regression and Spline Smoothing*, Second Edition, Marcel Dekker, New York 1999, s.281.

⁸⁴ Hurley, Hussy, McKeow and Addy, s.1.

⁸⁵ Pindyck and Rubinfeld, s. 118.

1.4. SPLİNELARIN ÖNEMİ

Veri analizinde sıklıkla karşılaşılan sorunlardan biri de, polinom bir modelle tahmin yapılmak istendiğinde, tek bir polinomla tahmin yapılmasının yeterli olmamasıdır⁸⁶. Çünkü bazı pratik uygulamalar için, düşük mertebeli polinomlar bağımlı değişkene doğru öngörü modelleri sağlamada yeteri kadar esnek değildir⁸⁷. Ayrıca, düşük mertebeli polinomlarda doğrusal en küçük kareli regresyon tekniği doğrusal olmayan bir ilişkiyi modellemek için yeterli değildir. Çok yüksek mertebeli polinom regresyonlarda ise, modelleme boyutlarının artması ve çoklu doğrusal bağlantı gibi sorunlar çıkar⁸⁸. Polinom regresyonda, model tahmini ayrık gözlemlerden olumsuz olarak aşırı bir şekilde etkilenmektedir⁸⁹.

Bu olumsuzlukları giderecek alternatif yaklaşım ise, aranan fonksiyonu temsil edecek şekilde verilen noktaların kümesine daha düşük dereceden parçalı polinomlarla yaklaşmak ve işlem kolaylığı sağlamaktır. Bu tür özelliklere sahip parçalı polinomlar spline fonksiyonlarıdır⁹⁰. Polinomun derecesi arttıkça doğal salınım etkisi arttığı için, spline fonksiyonların kullanımı doğal salınım etkisini azaltmaktadır⁹¹. Mesela zaman serisi verileri her zaman tek bir düzgün polinomla tahmin edilemez. Çoğunlukla, politik veya yapısal değişiklikler, tek bir basit polinomla etkili bir şekilde tahmin edilemeyen veri noktalarında beklenmedik değişiklikler (regresyon eğrisi yönünde değişiklikler olması gibi) meydana getirir⁹².

Polinomlar doğru üzerinde sonlu bir aralıkta tanımlanan fonksiyonların tahmin edilmesinde uzun zamandan beri kullanılmaktadır. Başlangıçta polinom interpolasyonları ve kesikli Taylor serisi olarak adlandırılan Taylor polinomları

⁸⁶ Norman R. Draper, Irwin Guttman and Peter Lipow, "All-Bias Designs for Spline Functions Joined at the Axes", *Journal of the American Statistical Association*, 72(358), 1977, s.424.

⁸⁷ David Woods and Susan Lewis "All-Bias Designs for Polynomial Spline Regression Models", *Australian & New Zealand Journal of Statistics*, 48(1), 2006, s. 50.

⁸⁸ Balaji Kommineni, Shubhankar Basu and Ranga Vemuri, "A Spline Based Regression Technique on Interval Valued Noisy Data", *Sixth International Conference on Machine Learning and Applications: 13-15 December 2007*, University of Cincinnati, Ohio/USA, s.241.

⁸⁹ Steve Horvath, *Weighted Network Analysis Applications in Genomics and Systems Biology*, Springer Science+Business Media, New York 2011, s.358.

⁹⁰ Aynur Canıvar, *Küçük Spline Fonksiyonlar Yardımıyla Lineer Olmayan Kısmi Türevli Diferensiyel Denklemlerin Sayısal Çözümleri*, (Yayımlanmamış Yüksek Lisans Tezi), Osmangazi Üniversitesi Fen Bilimleri Enstitüsü, Eskişehir 2003, s.14.

⁹¹ Girdhar G. Agarwal and W. J. Studden, *An Algorithm for Selection of Design and Knots in the Response Curve Estimation By Spline Functions*, Technical Report No. 78-15, Purdue University, USA: Indiana 1978, s.2.

⁹² Marsh, *Estimating Spline Knots in Time Series Polynomial Regression Models*, s.2.

kullanılmaktaydı. Daha sonra en küçük kareler ve minimax yaklaşımlarının kullanımı yaygınlaşmıştır. Bunlar yaklaşım hatalarını en aza indirerek sonuç belirlemektedir⁹³.

Dalgalı dağılım gösteren verilerin tahmininde spline fonksiyonları ve diğer parçalı polinomların kullanımı geniş bir uygulama alanına sahiptir ve birçok analizin de konusu olmaktadır⁹⁴. Polinom splinelerin kullanımı, parametrik olmayan modelleme için etkili bir yaklaşım sağlar. Polinom splineler, lojistik ve poisson regresyon gibi genelleştirilmiş regresyon, en küçük kareli regresyon, olasılık yoğunluk ve koşullu olasılık yoğunluk fonksiyon tahminini de içeren geniş bir aralıktaki istatistiksel problemlere uygulanabilir⁹⁵.

Spline fonksiyonları bağımlı ve bağımsız değişkenler arasındaki ilişkinin fonksiyonel biçimi bilinmediği zaman, verileri tahmin etmek veya modellemek için sıklıkla kullanılmaktadır⁹⁶. Spline fonksiyonları oluşturan parçalı polinom fonksiyonlar, yaygın olarak kullanılmakta olan tahmin fonksiyonlarıdır. Parçalı polinomlar, ilgili bilgisayar programlarında kolaylıkla birleştirilebilir. Bundan dolayı, özellikle deneysel verilerin veya modellenen eğrilerin tahmininde parçalı polinomların kullanımı uygundur⁹⁷. Basit veya çoklu regresyon analizinde kullanılan splinelerle, bağımsız değişkenlerin doğrusal olmayan etkileri araştırılır⁹⁸.

1.5. SPLİNE FONKSİYONLARI

Spline fonksiyon yaklaşımları, çeşitli problemlerin çözümünde en iyi yaklaşımdır. Son yıllarda spline fonksiyonları hem pratik hem de teorik çalışmalarda yaygın bir şekilde kullanılmaktadır⁹⁹. Spline fonksiyonları, tahmin fonksiyonların özel bir sınıfı olup özellikle, sayısal analizin uygulama aşamalarında yoğun bir şekilde

⁹³ G. M. Phillips and P. J. Taylor, "Algorithms for Spline and Other Approximations to Functions and Data", *Computer Physics Communications*, 73(1-3), 1992, s.1.

⁹⁴ Alberto M. Mier Muth and Alan S. Willsky, *A Sequential Method for Spline Approximation with Variable Knots*, Technical Report (ESL-P-759), M.I.T. Electronic Systems Laboratory, Mexico City 1977, s.2.

⁹⁵ Jianhua Z. Huang, "Local Asymptotics for Polynomial Spline Regression", *The Annals of Statistics*, 31(5), 2003, s.1600.

⁹⁶ Smith, *Curve Fitting and Modeling with Splines Using Statistical Variable Selection Techniques*, s.1.

⁹⁷ Carl de Boor and John R. Rice, *Least Squares Cubic Spline Approximation I-Fixed Knots*, Technical Report CSD-TR 20, Computer Sciences, Purdue University, Indiana 1968, s.1.

⁹⁸ Martin Gregory and Merck Serono, "A Collection of SAS Macros to Calculate Odds Ratios Using Spline Regression", *SAS: PhUSE Proceedings (October 12-15)*, Manchester, UK 2008, Erişim Tarihi: 12.06.2011, http://lexjansen.com/cgi-bin/xsl_transform.php?x=phuse2008&s=proceedings&c=phuse.

⁹⁹ Arshad Khan, Islam Khan and Tariq Aziz, "A Survey on Parametric Spline Function Approximation", *Applied Mathematics and Computation*, 171(2), 2005, s. 984.

kullanılmaktadır. Çünkü güçlü yakınsaklık özelliklerinden dolayı interpolasyon ve eğrilere yaklaşımda etkili bir araç olarak kullanılır¹⁰⁰.

Spline fonksiyonların sahip olduğu özellikler aşağıdaki gibidir¹⁰¹;

- ✓ Spline fonksiyon uzayları, çok uygun tabanlı (baz) fonksiyonlara sahip olan sonlu boyutlu doğrusal uzaylardır.
- ✓ Spline fonksiyonları, nispeten daha düzgün fonksiyonlardır.
- ✓ Spline fonksiyonların elle hesaplanması, bilgisayarlarda değerlendirilmesi ve kaydedilmesi kolaydır.
- ✓ Spline fonksiyonların türevleri ve integralleri de spline fonksiyondur.
- ✓ Spline fonksiyonları, polinomların bütün özelliklerine de sahiptir.
- ✓ Tahmin teorisi ve sayısal analizde splinelerin kullanımıyla ortaya çıkan çeşitli matrisler, uygun işaret (sign) ve determinantsal özelliklere sahiptir.
- ✓ Spline fonksiyonun işaret yapısı ve şekli, katsayılarının işaretine bağlı olabilir.
- ✓ $[a, b]$ aralığı üzerindeki (a: ilk düğüm noktası, b: sonuncu düğüm noktası) her sürekli fonksiyon yeteri kadar düğüm noktası belirlenerek mertebesi sabit olan m. mertebeden spline fonksiyonla iyi bir şekilde tahmin edilebilir.
- ✓ Düzgün fonksiyonların tahmini için tam yakınsama oranları spline fonksiyonlarla verilebilir.
- ✓ Düşük mertebeli splineler çok esnektir ve genellikle polinomlardaki gibi salınım göstermezler.

Splineler sıklıkla kullanılmasına rağmen, birden fazla bağımsız değişken kullanıldığı zaman bazı sakıncaları vardır. Düğüm noktası pozisyonlarının yerleşimini önerebilecek önsel bir bilgi olmadığında normal bir splinenin kullanımını doğrulamak zordur¹⁰². En yaygın olarak kullanılan üç tür spline fonksiyonu vardır.

1.5.1. Doğrusal Spline Fonksiyonu

Birinci dereceden en basit spline olup, doğru parçalarının birleştirilmesiyle oluşturulur¹⁰³. Doğrusal splineler, bazı yönlerden yüksek dereceli spline fonksiyonlara

¹⁰⁰ Barth, Arthur Kraft and John Kraft, s.219.

¹⁰¹ Larry L. Schumaker, *Spline Functions: Basic Theory*, Third Edition, Cambridge University Press, New York 2007, s.5.

¹⁰² Ryan, s.341.

¹⁰³ http://people.stat.sfu.ca/~cschwarz/Consulting/Trinity/Phase2/TrinityWorkshop/Workshop-material-Simon/Intro_to_splines/intro_to_splines_notes.pdf, s.8.

göre avantajlı sayılabilir. İlk olarak, temel fonksiyona parçalı doğrusal fonksiyonlarla yaklaşıldığı zaman, fonksiyon başlangıçta düzdür. Ani düşüşlerden sonra (bir eşik değeri etkisi oluşur) fonksiyon tekrar düz devam eder. Doğrusal splineda, bu etkiyi yakalamak için biri düşüşün başlangıcında ve diğeri sonunda olmak üzere iki kırılma noktası yeterlidir. Ancak kareli splineda bu etkiyi yakalamak için üç kırılma noktası gereklidir. Konkav(içbükey) bir parça oluştuğundan, birincisi konkav parça oluşturan düşüşün başlangıcında, ikincisi ortasında ve üçüncüsü sonundadır. İkinci olarak, yüksek dereceli spline fonksiyonlar kullanıldığı zaman, sınır değeri problemi olduğu için kısmi bir çözüm olarak splinelerin uç noktalarda doğrusal olması istenir. Başlangıçta doğrusal splineler kullanılırsa buna gerek kalmaz¹⁰⁴.

Doğrusal spline fonksiyonu, veri noktalarının alt aralıklara bölünmesiyle her bir aralık için bu aralıktaki değerlerden geçen bir doğrusal fonksiyon kullanılması esasına dayanmaktadır. Bu fonksiyonlar,

$$S(x) = \begin{cases} S_0(x) = y_0 + m_0(x - x_0), & x \in [x_0, x_1] \\ S_1(x) = y_1 + m_1(x - x_1), & x \in [x_1, x_2] \\ \vdots \\ S_j(x) = y_j + m_j(x - x_j), & x \in [x_j, x_{j+1}] \\ \vdots \\ S_{n-1}(x) = y_{n-1} + m_{n-1}(x - x_{n-1}), & x \in [x_{n-1}, x_n] \end{cases} \quad (1.1)$$

biçiminde ifade edilir. $x_j \leq x \leq x_{j+1}$ aralığındaki her bir doğrunun eğimi,

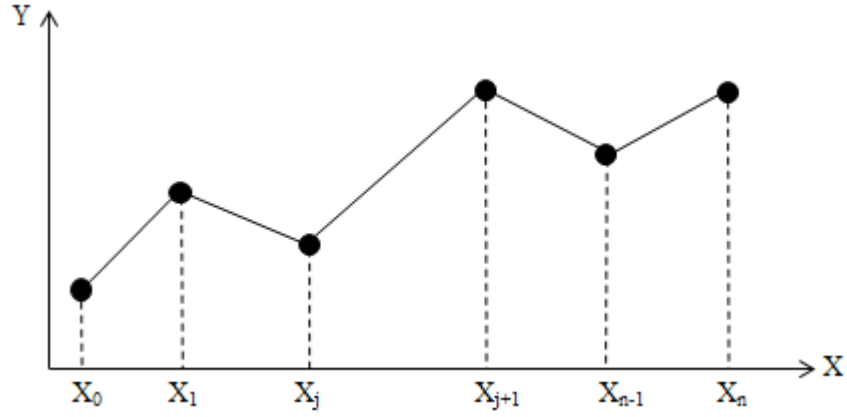
$$m_j = \frac{y_{j+1} - y_j}{x_{j+1} - x_j}$$

eşitliği ile bulunur. Şekil 1.3'te, eşitlik 1.1'in grafiksel gösterimi ele alınmıştır.

Böylece fonksiyonun herhangi bir $x_i \in [x_0, x_n]$ noktasının bulunduğu aralık tespit edilerek bu aralıkta fonksiyon değeri $S(x_i) = S_j(x_i) = y_j + m_j(x_i - x_j)$ yardımı ile hesaplanabilir¹⁰⁵.

¹⁰⁴ Kevin J. Fox, "Non-Parametric Estimation of Technical Progress", *Journal of Productivity Analysis*, 10(3), 1998, s.240.

¹⁰⁵ Oturanç, Kurnaz, Kiriş ve Keskin, s.194.



Şekil 1.3. Doğrusal Spline Fonksiyonu

1.5.2. Kareli (Kuadratik) Spline Fonksiyon

Kareli spline, doğrusal splinedan daha düzgündür ve kübik splinelardan daha kolay bir şekilde yorumlanır. Bundan dolayı da kareli splineların epidemiyoloji (toplumdaki hastalık, kaza ve sağlıkla ilgili durumların dağılımını, görülme sıklıklarını ve bunları etkileyen belirteçleri inceleyen bir tıp bilimi dalı¹⁰⁶) alanında kullanımı daha yaygındır¹⁰⁷.

Kareli spline fonksiyonda, $[x_0, x_n]$ aralığının her alt aralığındaki noktalardan geçen fonksiyonlar, ikinci dereceden polinom olarak seçilmektedir¹⁰⁸.

Bu durumda $S_j(x) = a_jx^2 + b_jx + c_j$ yazılacak olursa, kareli doğrusal fonksiyon,

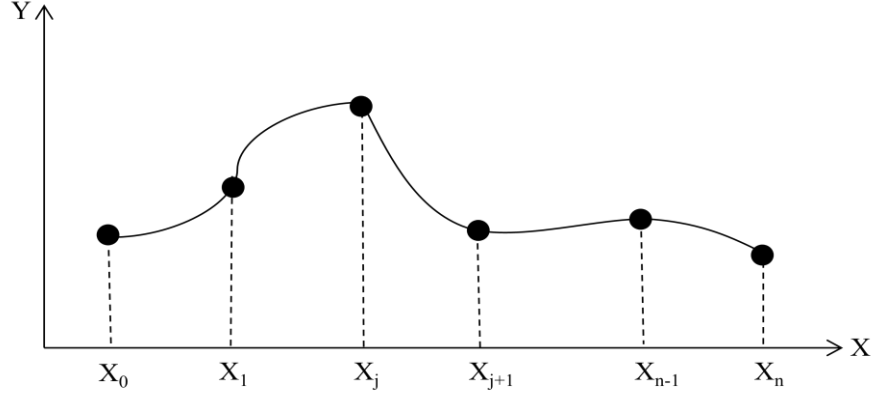
$$S(x) = \begin{cases} S_0(x) = a_0x^2 + b_0x + c_0, & x \in [x_0, x_1] \\ S_1(x) = a_1x^2 + b_1x + c_1, & x \in [x_1, x_2] \\ \vdots \\ S_j(x) = a_jx^2 + b_jx + c_j, & x \in [x_j, x_{j+1}] \\ \vdots \\ S_{n-1}(x) = a_{n-1}x^2 + b_{n-1}x + c_{n-1}, & x \in [x_{n-1}, x_n] \end{cases} \quad (1.2)$$

biçiminde ifade edilir. Şekil 1.4'te eşitlik 1.2'nin grafiksel gösterimi ele alınmıştır.

¹⁰⁶ <http://tr.wikipedia.org/wiki/Epidemiyoloji>, Erişim Tarihi: 03.04.2012.

¹⁰⁷ Mulla, s.78.

¹⁰⁸ Oturanç, Kurnaz, Kiriş ve Keskin, s.195.



Şekil 1.4. Kareli Spline Fonksiyonu

$S_j(x) = a_j x^2 + b_j x + c_j$ biçimindeki kareli spline fonksiyonunda 3 bilinmeyen vardır. Eşitlik 1.2’de $(n + 1)$ nokta için n tane aralık söz konusudur. Bundan dolayı $3n$ tane bilinmeyen hesaplanması gerekmektedir. Bu bilinmeyenlerin hesaplanması için $3n$ tane denkleme veya koşula ihtiyaç vardır¹⁰⁹. Buna göre,

1. İç düğüm noktalarında fonksiyon değerleri eşit olduğundan fonksiyon süreklidir. Yani,

$$S_{j-1}(x_j) = a_{j-1}x_j^2 + b_{j-1}x_j + c_{j-1} = y_j$$

$$S_j(x_j) = a_j x_j^2 + b_j x_j + c_j = y_j$$

biçiminde ifade edilir. Burada, $j = 1, 2, \dots, n - 1$ dir. Her bir nokta için iki denklem elde edildiğinden $n - 1$ nokta için $2n - 2$ denklem elde edilir.

2. Kareli spline fonksiyonunda ilk ve son fonksiyonlar uç noktalardan geçmektedir. Yani,

$$S_0(x_0) = a_0 x_0^2 + b_0 x_0 + c_0 = y_0$$

$$S_{n-1}(x_n) = a_{n-1} x_n^2 + b_{n-1} x_n + c_{n-1} = y_n$$

biçiminde ifade edilir. Burada, 2 denklem elde edilir.

3. İç düğüm noktalarında fonksiyonun birinci türev değerleri eşittir. Yani, $S'_{j-1}(x_j) = S'_j(x_j)$ ’ dir.

$$S'_{j-1}(x_j) = a_{j-1} x_j^2 + b_{j-1} x_j + c_{j-1} = 2a_{j-1} x_j + b_{j-1}$$

¹⁰⁹ Melis Zorşahin, *Polinom ve Polinom Olmayan Kübik Spline Fonksiyonlar Yardımıyla Bazı Kısmi Diferensiyel Denklemlerin Sayısal Çözümleri*, (Yayımlanmamış Yüksek Lisans Tezi), Osmangazi Üniversitesi Fen Bilimleri Enstitüsü, Eskişehir 2009, ss.3-4.

$$S'_j(x_j) = a_j x_j^2 + b_j x_j + c_j = 2a_j x_j + b_j$$

biçiminde ifade edilir. Burada, $j = 1, 2, \dots, n-1$ dir. Bu eşitlik $n-1$ nokta için sağlanacağından $n-1$ denklem elde edilir.

Böylece toplam $2n-2+2+n-1=3n-1$ denklem elde edilir. $3n$ bilinmeyeninin hesaplanması için 1 adet denkleme ihtiyaç vardır. Dolayısıyla son koşul olarak,

4. İlk noktada ikinci türevin 0(sıfır) olduğu kabul edilir. Yani, $a_0 = 0$ olup, ilk iki noktanın bir doğru ile birleştirildiği anlamına gelir.

Sonuç olarak, $3n$ bilinmeyenli denklem sistemini çözmek için $3n$ bağıntı elde edilmiş olur.

1.5.3. Kübik Spline Fonksiyon

Kübik spline, ilk olarak James Ferguson tarafından ele alınmıştır¹¹⁰. Kübik spline, manuel bir eğri çizme aracı olarak mühendisler tarafından yaygın bir şekilde kullanılan mekanik splinea yakın matematiksel bir yaklaşımdır. Kübik spline ile daha düzgün eğriler elde edilmektedir¹¹¹. Kübik spline, tüm veri noktalarından geçerek düzgün bir eğri elde edilmesine olanak sağlayan interpolasyon yöntemidir¹¹². Kübik splinelar, en yaygın kullanılan spline fonksiyonları olup, verileri tahmin eden düzgün fonksiyonlardır. Kübik splinelar, interpolasyon için kullanıldığı zaman, Lagrange ve Hermite interpolasyonları gibi yüksek dereceli polinom interpolasyonlarında görülen salınım etkisi göstermez¹¹³.

Matematiksel olarak her $[x_j, x_{j+1}]$ alt aralığında öyle kübik $S(x)$ fonksiyonları kurulabilir ki $[x_0, x_n]$ aralığında sonuçta oluşturulan $S(x)$ parçalı fonksiyonu birinci ve ikinci mertebeden türevleri $[x_0, x_n]$ aralığında sürekli olur. Burada birinci mertebeden türevinin sürekli olması ile fonksiyonun keskin köşelere sahip olmaması, ikinci

¹¹⁰ David Salomon, *Curves and Surfaces for Computer Graphics*, Springer Science+Business Media, New York 2006, s.141.

¹¹¹ Khan, A., Khan, I. and Aziz, s. 984.

¹¹² Irving K. Hwang, "Using SAS-GLM in Kinetic Modeling", SAS Institute Inc (Ed.), *Proceedings of the Third Annual SAS Users Group International Conference*, Las Vegas, Nevada 1978, s.124.

¹¹³ M. Luz Gamiz, K. B. Kulasekera, Nikolaos Limnios and Bo Henry Lindqvist, *Applied Nonparametric Statistics in Reliability*, Springer-Verlag, London 2011, s.202.

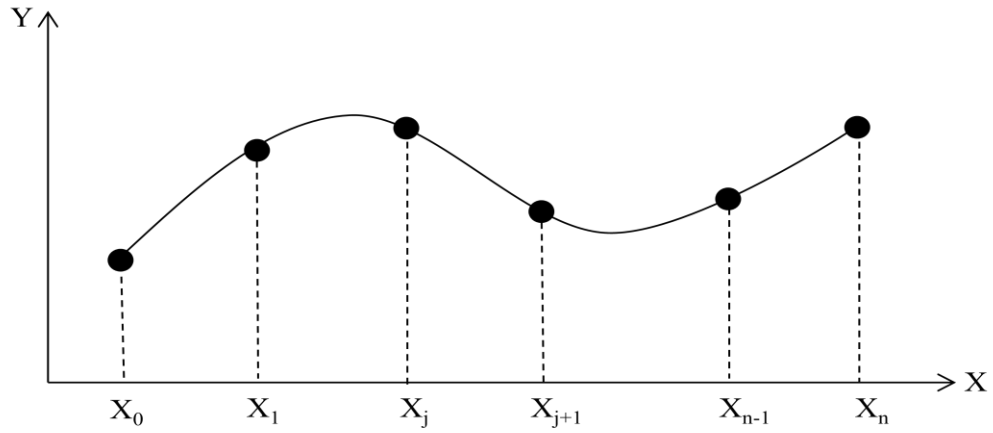
mertebeden türevinin sürekli olması ile fonksiyonun her noktasında eğim açısının tanımlanmış olması sağlanır¹¹⁴.

Kübik spline fonksiyonda, $[x_0, x_n]$ aralığının her alt aralığındaki noktalardan geçen fonksiyonlar, üçüncü dereceden polinom olarak seçilmektedir.

Bu durumda $S_j(x) = a_jx^3 + b_jx^2 + c_jx + d_j$ yazılacak olursa, kübik spline fonksiyonu,

$$S(x) = \begin{cases} S_0(x) = a_0x^3 + b_0x^2 + c_0x + d_0, & x \in [x_0, x_1] \\ S_1(x) = a_1x^3 + b_1x^2 + c_1x + d_1, & x \in [x_1, x_2] \\ \vdots & \vdots \\ S_j(x) = a_jx^3 + b_jx^2 + c_jx + d_j, & x \in [x_j, x_{j+1}] \\ \vdots & \vdots \\ S_{n-1}(x) = a_{n-1}x^3 + b_{n-1}x^2 + c_{n-1}x + d_{n-1}, & x \in [x_{n-1}, x_n] \end{cases} \quad (1.3)$$

biçiminde ifade edilir. Şekil 1.5'te eşitlik 1.3'ün grafiksel gösterimi ele alınmıştır.



Şekil 1.5. Kübik Spline Fonksiyonu

$S_j(x) = a_jx^3 + b_jx^2 + c_jx + d_j$ biçimindeki kübik spline fonksiyonunda 4 bilinmeyen vardır. Eşitlik 1.3'te $(n + 1)$ nokta için n tane aralık söz konusudur. Bundan dolayı $4n$ tane bilinmeyen hesaplanması gerekmektedir. Bu bilinmeyenlerin hesaplanması için $4n$ tane denkleme veya koşula ihtiyaç vardır. Buna göre,

1. $x \in [x_j, x_{j+1}]$ için, her bir kübik polinom,

¹¹⁴ Oturaç, Kurnaz, Kiriş ve Keskin, s.198.

$$S(x) = S_j(x) = a_j(x - x_j)^3 + b_j(x - x_j)^2 + c_j(x - x_j) + d_j$$

şeklinde kübik bir polinomdur.

2. İç düğüm noktalarında fonksiyon değerleri eşit olduğundan fonksiyon süreklidir. Her bir nokta için iki denklem elde edildiğinden $n - 1$ nokta için $2n - 2$ denklem elde edilir.

3. Spline fonksiyonu her bir düğüm noktasından geçmektedir. Burada, 2 denklem elde edilir.

4. İç düğüm noktalarında fonksiyonun birinci türev değerleri eşittir. Bu durum $n - 1$ nokta için sağlanacağından $n - 1$ denklem elde edilir.

5. İç düğüm noktalarında fonksiyonun ikinci türev değerleri eşittir. Yani, $S''_{j-1}(x_j) = S''_j(x_j)$ dir.

$$S''_{j-1}(x_j) = a_{j-1}x_j^3 + b_{j-1}x_j^2 + c_{j-1}x_j + d_{j-1} = 6a_{j-1}x_j + 2b_{j-1}$$

$$S''_j(x_j) = a_jx_j^3 + b_jx_j^2 + c_jx_j + d_j = 6a_jx_j + 2b_j$$

biçiminde ifade edilir. Burada, $j = 1, 2, \dots, n - 1$ dir. Bu eşitlik $n - 1$ nokta için sağlanacağından $n - 1$ denklem elde edilir.

Böylece toplam $2n - 2 + 2 + n - 1 + n - 1 = 4n - 2$ denklem elde edilir. $4n$ bilinmeyen hesaplanması için 2 adet denkleme ihtiyaç vardır. Dolayısıyla son koşul olarak,

6. İlk ve son noktalarda ikinci türevin 0(sıfır) olduğu kabul edilir. Yani, $S''(x_0) = S''(x_n) = 0$ dır. Bu durum uç noktalarda fonksiyonun doğrusal olacağı anlamına gelir.

Sonuç olarak, $4n$ bilinmeyenli denklem sistemini çözmek için $4n$ bağıntı elde edilmiş olur.

1.5.3.1. Kübik Spline Katsayılarının Bulunmasında Birinci Yaklaşım

Birinci yaklaşımda kübik spline fonksiyonun katsayıları bulunurken Langrange interpolasyon formülünden yararlanılmaktadır. Buna göre, $S(x)$ fonksiyonu parçalı

kübik fonksiyon olduğu için, bu fonksiyonun ikinci türevi $S''(x)$ $[x_0, x_n]$ aralığında parçalı doğrusaldır¹¹⁵. Herhangi bir $[x_{j-1}, x_j]$ aralığında,

$[x_{j-1}, S''(x_{j-1})], [x_j, S''(x_j)]$ noktalarından geçen doğru denklemi doğrusal lagrange interpolasyon formülüne göre,

$$\frac{S_j''(x) - S''(x_{j-1})}{x - x_{j-1}} = \frac{S''(x_j) - S''(x_{j-1})}{x_j - x_{j-1}} \quad (1.4)$$

eşitliğiyle bulunur¹¹⁶.

Eşitlik 1.4'ün $S_j''(x)$ 'e göre düzenlenmesi sonucu,

$$\begin{aligned} S_j''(x)(x_j - x_{j-1}) - S''(x_{j-1})(x_j - x_{j-1}) &= S''(x_j)(x - x_{j-1}) - S''(x_{j-1})(x - x_{j-1}) \\ S_j''(x)(x_j - x_{j-1}) &= S''(x_j)(x - x_{j-1}) + S''(x_{j-1})(x_j - x_{j-1} - x + x_{j-1}) \\ S_j''(x)(x_j - x_{j-1}) &= S''(x_j)(x - x_{j-1}) + S''(x_{j-1})(x_j - x) \\ S_j''(x) &= S''(x_j) \frac{x - x_{j-1}}{x_j - x_{j-1}} + S''(x_{j-1}) \frac{x_j - x}{x_j - x_{j-1}} \end{aligned} \quad (1.5)$$

eşitliği elde edilir.

Eşitlik 1.5'te $S''(x_j) = M_j$, $S''(x_{j-1}) = M_{j-1}$ ve $h_j = x_j - x_{j-1}$ ile gösterilsin¹¹⁷.

Eşitlik 1.5 düzenlenirse,

$$S_j''(x) = M_{j-1} \frac{x_j - x}{h_j} + M_j \frac{x - x_{j-1}}{h_j} \quad (1.6)$$

eşitliği elde edilir. Eşitlik 1.6'nın iki kez integrali alınır,

$$\begin{aligned} S_j'(x) &= \int M_{j-1} \frac{x_j - x}{h_j} dx + \int M_j \frac{x - x_{j-1}}{h_j} dx \\ S_j'(x) &= -M_{j-1} \frac{(x_j - x)^2}{2h_j} + M_j \frac{(x - x_{j-1})^2}{2h_j} + A \end{aligned} \quad (1.7)$$

¹¹⁵ John H. Mathews, *Numerical Methods for Mathematics, Science, and Engineering*, Second Edition, Prentice Hall, New Jersey 1992, s.287.

¹¹⁶ William Weatherspoone, *Spline Analysis and an Application*, Master's Thesis, University of Texas, El Paso 1975, s.6.

¹¹⁷ J. H. Ahlberg, E. N. Nilson and J. L. Walsh, *The Theory of Splines and Their Applications*, Academic Press, New York 1967, s.10.

$$S_j(x) = \int -M_{j-1} \frac{(x_j-x)^2}{2h_j} dx + \int M_j \frac{(x-x_{j-1})^2}{2h_j} dx + \int A dx$$

$$S_j(x) = M_{j-1} \frac{(x_j-x)^3}{6h_j} + M_j \frac{(x-x_{j-1})^3}{6h_j} + Ax + B \quad (1.8)$$

eşitliği elde edilir. Hesaplamalarda kolaylık sağlamak için,

$$Ax + B = A \left(\frac{x_j - x}{h_j} \right) + B \left(\frac{x - x_{j-1}}{h_j} \right)$$

alınarak¹¹⁸ A ve B integral sabitleri olmak üzere,

$$S_j(x) = M_{j-1} \frac{(x_j-x)^3}{6h_j} + M_j \frac{(x-x_{j-1})^3}{6h_j} + A \frac{x_j-x}{h_j} + B \frac{x-x_{j-1}}{h_j} \quad (1.9)$$

eşitliği bulunur. Bu fonksiyonun her bir aralığın uç noktalarından geçmesi koşulu¹¹⁹ kullanılarak A ve B integral sabitleri bulunur. Yani $S_j(x_{j-1}) = y_{j-1}$ ve $S_j(x_j) = y_j$ koşulları¹²⁰ kullanılır. Buna göre, eşitlik 1.9'da x yerine sırasıyla x_{j-1} ve x_j yazılarak A ve B integral sabitleri elde edilir.

$$\begin{aligned} S_j(x_{j-1}) &= M_{j-1} \frac{(x_j-x_{j-1})^3}{6h_j} + M_j \frac{(x_{j-1}-x_{j-1})^3}{6h_j} + A \frac{x_j-x_{j-1}}{h_j} + B \frac{x_{j-1}-x_{j-1}}{h_j} \\ y_{j-1} &= M_{j-1} \frac{h_j^2}{6} + A \rightarrow A = y_{j-1} - \frac{M_{j-1}h_j^2}{6} \end{aligned} \quad (1.10)$$

bulunur.

$$\begin{aligned} S_j(x_j) &= M_{j-1} \frac{(x_j-x_j)^3}{6h_j} + M_j \frac{(x_j-x_{j-1})^3}{6h_j} + A \frac{x_j-x_j}{h_j} + B \frac{x_j-x_{j-1}}{h_j} \\ y_j &= M_j \frac{h_j^2}{6} + B \rightarrow B = y_j - \frac{M_j h_j^2}{6} \end{aligned} \quad (1.11)$$

bulunur. A ve B sabitleri eşitlik 1.9'da yerine yazılırsa,

$$\begin{aligned} S_j(x) &= M_{j-1} \frac{(x_j-x)^3}{6h_j} + M_j \frac{(x-x_{j-1})^3}{6h_j} + \left(y_{j-1} - \frac{M_{j-1}h_j^2}{6} \right) \frac{x_j-x}{h_j} + \left(y_j - \frac{M_j h_j^2}{6} \right) \frac{x-x_{j-1}}{h_j} \end{aligned} \quad (1.12)$$

¹¹⁸ Duygu Dönmez, *Diferansiyel Denklemlerin B-Spline Metodu ile Çözümleri Hakkında*, (Yayımlanmamış Yüksek Lisans Tezi), Celal Bayar Üniversitesi Fen Bilimleri Enstitüsü, Manisa 2008, s.7.

¹¹⁹ Bayram, s.150.

¹²⁰ Oturanç, Kurnaz, Kiriş ve Keskin, s.200.

eşitliği elde edilir. Eşitlik 1.12’de M_{j-1} ve M_j olmak üzere sadece iki bilinmeyen vardır. Eğer bu bilinmeyenler bulunursa, $[x_0, x_n]$ aralığında spline fonksiyonunu oluşturmak için, eşitlik 1.12 kullanılabilir. M_{j-1} ve M_j bilinmeyenlerini bulmak için $S'(x)$ ’ in sürekliliğini kullanarak iç düğüm noktalarında $S'_{j+1}(x_j) = S'_j(x_j)$ koşulundan faydalanılır. Buna göre eşitlik 1.12’nin türevi alınır,

$$\begin{aligned}
 S'_j(x) &= -M_{j-1} \frac{(x_j-x)^2}{2h_j} + M_j \frac{(x-x_{j-1})^2}{2h_j} + \frac{y_{j-1} - \frac{M_{j-1}h_j^2}{6}}{h_j} + \frac{y_j - \frac{M_jh_j^2}{6}}{h_j} \\
 S'_j(x) &= -M_{j-1} \frac{(x_j-x)^2}{2h_j} + M_j \frac{(x-x_{j-1})^2}{2h_j} + \frac{y_j - \frac{M_jh_j^2}{6} - y_{j-1} + \frac{M_{j-1}h_j^2}{6}}{h_j} \\
 S'_j(x) &= -M_{j-1} \frac{(x_j-x)^2}{2h_j} + M_j \frac{(x-x_{j-1})^2}{2h_j} + \frac{y_j - y_{j-1}}{h_j} - \frac{M_j - M_{j-1}}{6} h_j
 \end{aligned} \tag{1.13}$$

eşitliği bulunur. $S''_j(x)$ ve $S_j(x)$ fonksiyonları $[x_0, x_n]$ aralığında sürekli olduğundan, $S'(x)$ fonksiyonunun da bu aralıkta sürekli olması gerekir. Dolayısıyla $[x_{j-1}, x_j]$ ve $[x_j, x_{j+1}]$ aralıklarının ortak noktalarındaki türevlerinin soldan ve sağdan limitleri bulunur¹²¹.

$x = x_j$ noktasında $S'_{j+1}(x_j)$ ve $S'_j(x_j)$ türevleri hesaplanır. Öncelikle $S'_j(x_j)$ hesabı yani soldan limit için eşitlik 1.13’te x yerine x_j yazılarak,

$$\begin{aligned}
 S'_j(x_j) &= -M_{j-1} \frac{(x_j-x_j)^2}{2h_j} + M_j \frac{(x_j-x_{j-1})^2}{2h_j} + \frac{y_j - y_{j-1}}{h_j} - \frac{M_j - M_{j-1}}{6} h_j \\
 S'_j(x_j) &= M_{j-1} \frac{h_j}{6} + M_j \frac{h_j}{3} + \frac{y_j - y_{j-1}}{h_j}
 \end{aligned} \tag{1.14}$$

eşitliği elde edilir. Benzer şekilde, $S'_{j+1}(x_j)$ hesabı yani sağdan limit için eşitlik 1.13’te x yerine x_j ve j yerine $j + 1$ yazılarak,

$$\begin{aligned}
 S'_{j+1}(x_j) &= -M_j \frac{(x_{j+1}-x_j)^2}{2h_{j+1}} + M_{j+1} \frac{(x_j-x_j)^2}{2h_{j+1}} + \frac{y_{j+1} - y_j}{h_{j+1}} - \frac{M_{j+1} - M_j}{6} h_{j+1} \\
 S'_{j+1}(x_j) &= -M_j \frac{h_{j+1}}{3} - M_{j+1} \frac{h_{j+1}}{6} + \frac{y_{j+1} - y_j}{h_{j+1}}
 \end{aligned} \tag{1.15}$$

eşitliği elde edilir.

¹²¹ Canıvar, s.21.

Düğüm noktasında $S'_j(x_j) = S'_{j+1}(x_j)$ fonksiyonları birbirine eşittir. Bu fonksiyonların birbirlerine eşitlenmesiyle,

$$M_{j-1} \frac{h_j}{6} + M_j \frac{h_j}{3} + \frac{y_j - y_{j-1}}{h_j} = -M_j \frac{h_{j+1}}{3} - M_{j+1} \frac{h_{j+1}}{6} + \frac{y_{j+1} - y_j}{h_{j+1}}$$

$$\frac{h_j}{6} M_{j-1} + \frac{2(h_j + h_{j+1})}{6} M_j + \frac{h_{j+1}}{6} M_{j+1} = \frac{y_{j+1} - y_j}{h_{j+1}} - \frac{y_j - y_{j-1}}{h_j}$$

$$h_j M_{j-1} + 2(h_j + h_{j+1}) M_j + h_{j+1} M_{j+1} = \frac{6}{h_{j+1}} (y_{j+1} - y_j) - \frac{6}{h_j} (y_j - y_{j-1}) \quad (1.16)$$

eşitliği elde edilir¹²².

Eşitlik 1.16'da $(n - 1)$ tane iç düğüm noktası yazılırsa $(n + 1)$ tane bilinmeyen ikinci türev içeren $(n - 1)$ tane denklem elde edilir¹²³.

1.5.3.2. Kübik Spline Katsayılarının Bulunmasında İkinci Yaklaşım

Bu yaklaşımda doğrudan kübik spline fonksiyonun 1-5 özelliklerini kullanarak, kübik spline fonksiyon katsayılarının bulunması için denklemlerin yazılması amaçlanmaktadır¹²⁴. Buna göre, her noktadan geçen kübik spline fonksiyonu $j = 0, 1, \dots, n$ için,

$$S(x_j) = S_j(x) = y_j = a_j(x - x_j)^3 + b_j(x - x_j)^2 + c_j(x - x_j) + d_j \quad (1.17)$$

biçiminde ifade edilir.

Eşitlik 1.17'de tanımlanan $S_j(x)$ fonksiyonları (x_j, y_j) ve (x_{j+1}, y_{j+1}) noktalarından geçer¹²⁵. Dolayısıyla eşitlik 1.17'de x yerine x_j yazılırsa,

$$S_j(x_j) = y_j = a_j(x_j - x_j)^3 + b_j(x_j - x_j)^2 + c_j(x_j - x_j) + d_j$$

$$S_j(x_j) = y_j = d_j \quad (1.18)$$

bulunur. Kübik spline fonksiyonun iç noktalarda sürekli olduğu koşuluna göre,

$$S_j(x_j) = S_{j-1}(x_j), \quad j = 1, 2, \dots, n - 1 \quad (1.19)$$

¹²² Ahlberg, Nilson and Walsh, s.10.

¹²³ Bayram, s.151.

¹²⁴ Dursun Aydın, *Semiparametrik Regresyon Modellemede Splayn Düzeltme Yaklaşımı ile Tahmin ve Çıkarsamalar*, (Yayımlanmamış Doktora Tezi), Anadolu Üniversitesi Fen Bilimleri Enstitüsü, Eskişehir 2005, s.31.

¹²⁵ Meltem Doğan, *q-Integer Noktalarında B-Spline Fonksiyonu ve Özellikleri*, (Yayımlanmamış Yüksek Lisans Tezi), Muğla Üniversitesi Fen Bilimleri Enstitüsü, Muğla 2008, s.12-14.

eşitliği bulunmaktadır. Buna göre eşitlik 1.17'de x yerine x_j ve j yerine $j - 1$ yazılırsa,

$$S_{j-1}(x_j) = a_{j-1}(x_j - x_{j-1})^3 + b_{j-1}(x_j - x_{j-1})^2 + c_{j-1}(x_j - x_{j-1}) + d_{j-1} \quad (1.20)$$

eşitliği elde edilir. Eşitlik 1.19'a göre $j = 1, 2, \dots, n - 1$ için,

$$d_j = a_{j-1}(x_j - x_{j-1})^3 + b_{j-1}(x_j - x_{j-1})^2 + c_{j-1}(x_j - x_{j-1}) + d_{j-1} \quad (1.21)$$

eşitliği yazılabilir¹²⁶. Eşitlik 1.17'nin birinci dereceden türevi alındığında,

$$S'_j(x) = 3a_j(x - x_j)^2 + 2b_j(x - x_j) + c_j \quad (1.22)$$

eşitliği elde edilir. Eşitlik 1.22'de x yerine x_j yazılırsa,

$$S'_j(x_j) = 3a_j(x_j - x_j)^2 + 2b_j(x_j - x_j) + c_j$$

$$S'_j(x_j) = c_j \quad (1.23)$$

bulunur. Kübik spline fonksiyonun iç noktalarda birinci derece türevlerinin eşit olduğu koşuluna göre, ,

$$S'_j(x_j) = S'_{j-1}(x_j), \quad j = 1, 2, \dots, n - 1 \quad (1.24)$$

eşitliği bulunmaktadır. Buna göre eşitlik 1.22'de x yerine x_j ve j yerine $j - 1$ yazılırsa,

$$S'_{j-1}(x_j) = 3a_{j-1}(x_j - x_{j-1})^2 + 2b_{j-1}(x_j - x_{j-1}) + c_{j-1} \quad (1.25)$$

eşitliği elde edilir. Eşitlik 1.24'e göre $j = 1, 2, \dots, n - 1$ için,

$$c_j = 3a_{j-1}(x_j - x_{j-1})^2 + 2b_{j-1}(x_j - x_{j-1}) + c_{j-1} \quad (1.26)$$

eşitliği yazılabilir. Eşitlik 1.22'nin ikinci dereceden türevi alındığında,

$$S''_j(x) = 6a_j(x - x_j) + 2b_j \quad (1.27)$$

eşitliği elde edilir. Eşitlik 1.27'de x yerine x_j yazılırsa,

$$S''_j(x_j) = 6a_j(x_j - x_j) + 2b_j$$

$$S''_j(x_j) = 2b_j = m_j \rightarrow b_j = \frac{m_j}{2} \quad (1.28)$$

¹²⁶ Rama B. Bhat and Snehashish Chakraverty, *Numerical Analysis in Engineering*, Alpha Science International, Pangbourne 2004, s.102-104.

eşitliği bulunur. Kübik spline fonksiyonun iç noktalarda ikinci derece türevlerinin eşit olduğu koşuluna göre, ,

$$S_j''(x_j) = S_{j-1}''(x_j), \quad j = 1, 2, \dots, n-1 \quad (1.29)$$

eşitliği bulunmaktadır. Buna göre eşitlik 1.27'de x yerine x_j ve j yerine $j-1$ yazılırsa,

$$S_{j-1}''(x_j) = 6a_{j-1}(x_j - x_{j-1}) + 2b_{j-1} \quad (1.30)$$

eşitliği elde edilir. Eşitlik 1.29'a göre $j = 1, 2, \dots, n-1$ için,

$$\begin{aligned} 2b_j &= 6a_{j-1}(x_j - x_{j-1}) + 2b_{j-1} \\ b_j &= 3a_{j-1}(x_j - x_{j-1}) + b_{j-1} \end{aligned} \quad (1.31)$$

eşitliği yazılabilir.

Eşitlik 1.31'de indis değerinin 1 artırılmasıyla,

$$b_{j+1} = 3a_j(x_{j+1} - x_j) + b_j \quad (1.32)$$

eşitliği elde edilir¹²⁷. Eşitlik 1.32'nin a_j katsayısına göre yeniden düzenlenmesi sonucu,

$$\begin{aligned} b_{j+1} - b_j &= 3a_j(x_{j+1} - x_j) \\ a_j &= \frac{b_{j+1} - b_j}{3(x_{j+1} - x_j)} \rightarrow a_j = \frac{\frac{m_{j+1}}{2} - \frac{m_j}{2}}{3(x_{j+1} - x_j)} \rightarrow a_j = \frac{m_{j+1} - m_j}{6(x_{j+1} - x_j)} \end{aligned} \quad (1.33)$$

eşitliği bulunur. c_j katsayısı ise, eşitlik 1.21'in d_{j+1} 'e göre düzenlenmesiyle bulunabilir.

Yani,

$$d_{j+1} = a_j(x_{j+1} - x_j)^3 + b_j(x_{j+1} - x_j)^2 + c_j(x_{j+1} - x_j) + d_j \quad (1.34)$$

ifadesi elde edilir. Eşitlik 1.34, c_j katsayısına göre düzenlenirse,

$$\begin{aligned} c_j(x_{j+1} - x_j) &= (d_{j+1} - d_j) - a_j(x_{j+1} - x_j)^3 - b_j(x_{j+1} - x_j)^2 \\ c_j &= \frac{(d_{j+1} - d_j)}{(x_{j+1} - x_j)} - \frac{m_{j+1} - m_j}{6(x_{j+1} - x_j)}(x_{j+1} - x_j)^2 - b_j(x_{j+1} - x_j) \\ c_j &= \frac{(y_{j+1} - y_j)}{(x_{j+1} - x_j)} - \frac{m_{j+1} - m_j + 3m_j}{6}(x_{j+1} - x_j) \end{aligned}$$

¹²⁷ Jacek Ronda and Graeme John Oliver, *Introduction to Numerical Methods with Matlab Procedures*, 2010, Erişim Tarihi: 22.02.2012, http://home.agh.edu.pl/~pwojnaro/dydaktyka/mn/numerical_methods_book_2011_04_05.pdf, s.53-56.

$$c_j = \frac{y_{j+1}-y_j}{x_{j+1}-x_j} - \frac{m_{j+1}+2m_j}{6}(x_{j+1}-x_j) \quad (1.35)$$

eşitliği elde edilir. Sonuç olarak $h_j = x_{j+1} - x_j$ alınırsa, a_j, b_j, c_j ve d_j katsayılarını bulmak için,

$$a_j = \frac{m_{j+1}-m_j}{6h_j} \quad (1.36)$$

$$b_j = \frac{m_j}{2} \quad (1.37)$$

$$c_j = \frac{y_{j+1}-y_j}{h_j} - \frac{m_{j+1}+2m_j}{6}h_j \quad (1.38)$$

$$d_j = y_j \quad (1.39)$$

denklemleri kullanılır. Burada h_j değeri a_j ve c_j değerlerine, m_j değeri de a_j, b_j ve c_j değerlerine bağlıdır. d_j katsayısını y_j 'ye eşit olduğu bilinmektedir. m_j 'nin elde edilebilmesi için eşitlik 1.26'daki c_j denkleminin $j + 1$ indisine göre düzenlemesi gerekecektir. Buna göre,

$$c_{j+1} = 3a_j h_j^2 + 2b_j h_j + c_j \quad (1.40)$$

biçiminde ifade edilir. Eşitlik 1.40'ta, eşitlik 1.36, 1.37 ve 1.38'deki ifadeler yerine yazılarak m_j bilinmeyenlerini bulmak için yeni denklem sistemi oluşturulur.

$$\begin{aligned} \frac{y_{j+2}-y_{j+1}}{h_{j+1}} - \frac{m_{j+2}+2m_{j+1}}{6}h_{j+1} &= 3\frac{m_{j+1}-m_j}{6h_j}h_j^2 + 2\frac{m_j}{2}h_j + \frac{y_{j+1}-y_j}{h_j} - \frac{m_{j+1}+2m_j}{6}h_j \\ \frac{m_{j+2}}{6}h_{j+1} + \frac{m_{j+1}}{3}h_{j+1} + \frac{m_{j+1}}{2}h_j - \frac{m_j}{2}h_j + m_j h_j - \frac{m_{j+1}}{6}h_j - \frac{m_j}{3}h_j &= \frac{y_{j+2}-y_{j+1}}{h_{j+1}} - \frac{y_{j+1}-y_j}{h_j} \\ \frac{m_j}{6}h_j + \frac{2m_{j+1}}{6}(h_{j+1} + h_j) + \frac{m_{j+2}}{6}h_{j+1} &= \frac{6}{h_{j+1}}(y_{j+2} - y_{j+1}) - \frac{6}{h_j}(y_{j+1} - y_j) \end{aligned} \quad (1.41)$$

eşitliği elde edilir. Görüldüğü üzere ikinci yaklaşımla bulunan eşitlik 1.41, eşitlik 1.16 ile aynıdır.

1.5.3.3. Kübik Spline Fonksiyonda Sınır Değerleri Sorunu

İki veya daha fazla koşul, $[a, b]$ aralığında tanımlı parçanın sınır noktalarında fonksiyonun türev değerleri veya spline değerleri üzerinde kısıtlamalar şeklinde verilir. En yaygın kullanılan sınır koşulları aşağıdaki gibidir¹²⁸.

Birinci Tip Sınır Koşulu

$[a, b]$ aralığında tanımlı parçanın sınır noktalarında, fonksiyonun birinci türev değerleri verilir.

$$S'(a) = f'(a), \quad S'(b) = f'(b)$$

İkinci Tip Sınır Koşulu

$[a, b]$ aralığında tanımlı parçanın sınır noktalarında, fonksiyonun ikinci türev değerleri verilir.

$$S''(a) = f''(a), \quad S''(b) = f''(b)$$

Üçüncü Tip Sınır Koşulu

$[a, b]$ aralığında tanımlı parçanın sınır noktalarında fonksiyonun birinci ve ikinci türev değerleri eşittir.

$$S'(a) = S'(b), \quad S''(a) = S''(b)$$

Bu koşullar periyodik olarak da adlandırılır. Fonksiyon periyodik olduğu zaman doğal olarak koşulların yerine getirilmesi gereklidir. Fonksiyonun periyodu $T = b - a$ ya eşittir.

Dördüncü Tip Sınır Koşulu

$$S'''(y, x_1 - 0) = S'''(y, x_1 + 0), \quad S'''(y, x_{m-1} - 0) = S'''(y, x_{m-1} + 0)$$

şeklindedir. Genel olarak spline fonksiyonun üçüncü türevi iç düğüm noktalarında kesiksizdir. Ancak dördüncü tip sınır koşulu kullanılarak kırılmaların sayısı düşürülebilir. Bu durumda elde edilen spline fonksiyon, $[x_0, x_1]$ ve $[x_{m-1}, x_m]$ aralıklarında tanımlı parçalar üzerinde sürekliliğini kaybetmeden üç kez türevlenebilir.

¹²⁸ Eugene V. Shikin and Alexander I. Plis, *Handbook on Splines for the User*, CRC Press, Boca Raton, Florida 1995, s.14.

Doğal Kübik Spline

En yaygın şekilde kullanılan sınır değer koşullarından biri de doğal kübik splinedir. Doğal kübik splinede doğal kelimesi, tahmin edilen splinenin en küçük ve en büyük düğüm noktalarında ikinci türevlerinin sıfır olduğu varsayımını ifade etmektedir¹²⁹.

$$S''(a) = S''(b) = 0$$

şeklinde ifade edilir.

1.6. SPLİNE GÖSTERİMİ

Spline fonksiyonların farklı doğrusal model gösterimleri bulunmaktadır. En çok kullanılan gösterimler, kesikli üstel fonksiyon, B-spline ve spline interpolasyon gösterimidir¹³⁰.

1.6.1. Kesikli Üstel Fonksiyon Gösterimi (+ Fonksiyon)

Kesikli üstel fonksiyon (truncated power function) gösterimi, kolay anlaşılması ve istatistiksel hipotezlerin test edilmesi bakımından en sade yaklaşımlardan biridir. Katsayıları kolay yorumlanması açısından diğer gösterimlere göre daha iyidir¹³¹. Kesikli üstel fonksiyon $(X - \xi)_+^n$ biçiminde gösterilir ve $X > \xi$ ise $(X - \xi)^n$, $X \leq \xi$ ise 0 değerini alır. Burada X gözlem değerleri, ξ ise düğüm noktasıdır. Eğer $n = 0$ değerini alırsa bu fonksiyon basamak (heaviside) fonksiyon veya adım fonksiyonu olarak adlandırılır. Bu gösterim ilk olarak Schoenberg ve Whitney (1953) tarafından verilmiştir¹³².

1.6.2. B-Spline

Eğriler ve yüzeyler için B-spline yöntemler ilk olarak 1940'lı yıllarda ortaya çıkmıştır. 1970'li yıllarda yoğun olarak çalışılan B-splinela ilgili en dikkat çekici çalışma R. Riesenfeld tarafından yapılmıştır. B-spline “basis spline” kelimelerinin

¹²⁹ Siddhartha Chib and Edward Greenberg, “Additive Cubic Spline Regression with Dirichlet Process Mixture Errors”, *Journal of Econometrics*, 156(2), 2010, s.324.

¹³⁰ Hina Mchta and Thomas Capizzi “Evaluating Linear Model Representations of Cubic Splines Using Proc Reg”, SAS Institute Inc (Ed.), *Proceedings of the Seventh Annual SAS Users Group International Conference*, San Francisco, California 1982, s.562.

¹³¹ Koo, s.268.

¹³² Greville, s.2.

kısaltılmış halidir¹³³. B-spline (basis-splines), Bezier eğrisinin genelleştirilmiş halidir. B-splinelar matematiksel olarak, diğer spline türlerinden daha karmaşıktır. B-splinelar, k. mertebeden ve B adet iç düğüm noktası olan fonksiyon şeklinde tanımlanır. B-spline polinomun derecesi mertebenin bir eksiğidir ($k-1=n$)¹³⁴.

Tekrarlı B-splinenin genel formu için, sıfırıncı dereceden B-Spline¹³⁵,

$$B_i^0(X) = \begin{cases} 1, & t_i \leq x \leq t_{i+1} \\ 0, & \text{Diğer Yerler} \end{cases}$$

olmak üzere, n. dereceden B-spline,

$$B_i^n(x) = \left(\frac{x - t_i}{t_{i+n} - t_i} \right) B_i^{n-1}(x) + \left(\frac{t_{i+n+1} - x}{t_{i+n+1} - t_{i+1}} \right) B_{i+1}^{n-1}(x) \quad (1.42)$$

n: Polinom derecesi $i = 0, \pm 1, \pm 2, \dots$ t : Düğüm noktası

biçiminde yazılabilir.

B-splinelarda düğüm noktaları eşit aralıklı olduğu zaman eşitlik 1.42 daha basit bir şekilde yazılabilir. Düğüm noktalarının herhangi bir eşit aralıklı sistemi, değişkenin doğrusal bir değişikliği yapılarak elde edilebildiği için, $t_i = i$ seçilir. Buna göre, tekrarlı ilişki¹³⁶,

$$B_i^n(x) = \frac{1}{n} (x - i) B_i^{n-1}(x) + \frac{1}{n} (i + n + 1 - x) B_{i+1}^{n-1}(x) \quad (1.43)$$

olarak yazılabilir.

B-spline fonksiyonlarının genel özellikleri aşağıdaki gibidir¹³⁷.

- ✓ Sıfırıncı dereceden B-splineda (B_i^0), tüm i ve x değerleri için $B_i^0(X) \geq 0$ eşitsizliği vardır.
- ✓ Sıfırıncı dereceden B-spline, bütün sayı doğrusu üzerinde sağdan süreklidir.
- ✓ Bütün $x \in R$ noktaları için, $\sum_{i=-\infty}^{\infty} B_i^n(x) = \sum_{i=-\infty}^{\infty} B_i^0(x) = 1$ eşitliği vardır.
- ✓ Eğer $n \geq 1$ ve x ise (t_i, t_{i+k+1}) aralığında değilse, $B_i^n(x) = 0$ değerini alır.

¹³³ Salomon, s.251.

¹³⁴ Jeffrey S. Racine, "A Primer on Spline Regression", 2011, Erişim Tarihi: 05 Eylül 2011, http://mirror.fcaglp.unlp.edu.ar/pub/CRAN/web/packages/crs/vignettes/spline_primer.pdf.

¹³⁵ Carl De Boor, *A Practical Guide to Splines*, Springer Verlag, New York 1978, s.131.

¹³⁶ Phillips and Taylor, s.8.

¹³⁷ Nurhan Özkan, *Splayn Teorisi ve Diferensiyel Denklemlere Uygulamaları*, (Yayımlanmamış Yüksek Lisans Tezi), Ondokuz Mayıs Üniversitesi Fen Bilimleri Enstitüsü, Samsun 2006, ss. 34-41.

1.6.3. İnterpolasyon Gösterimi

Çok sayıda veri noktalarına bir tek eğri ile yaklaşmak büyük kolaylıklar sağlasa da bazı durumlarda büyük hatalara neden olabilir. Ayrıca bu amaç için kullanılan Newton, Hermite ve Lagrange interpolasyon polinomlarının derecesi nokta sayısı arttıkça artacağından, bu tür polinomlarla yapılacak işlemler zorlaşır. Bu gibi durumlarda, daha esnek ve alternatif bir yöntem olarak¹³⁸, art arda gelen iki veri arasında birinci, ikinci ya da üçüncü dereceden fonksiyonlarla yaklaşımın yapıldığı spline interpolasyon yöntemi önerilmektedir. Spline interpolasyonu veri noktalarını çeşitli aralıklara bölerek, her bir aralıkta daha küçük dereceden polinomlarla yaklaşım yapma esasına dayanır¹³⁹.

1.7. SPLİNE TÜRLERİ

Farklı alanlarda kullanılan çeşitli spline türleri vardır. En yaygın kullanılan spline türleri regresyon spline ve düzgünleştirici (smoothing) splinedir. Ayrıca, başka bir spline türü de uygulama alanında pek kullanımı olmasa da teorik olarak çalışılan çok değişkenli splinelardır.

1.7.1. Regresyon Spline

Bu tür splinelar parametrik spline olarak da adlandırılmaktadır¹⁴⁰. Düğüm noktalarının seçimi oldukça önemli olduğu kesikli üstel fonksiyon gösterimini temel alan bu spline türünde, düğüm noktalarının sayısı nispeten daha azdır ve parametre tahminleri en küçük kareler yöntemiyle yapılmaktadır¹⁴¹. Regresyon spline yöntemleri, kısıtlı ve kısıtlandırılmamış modellerin katsayı tahminlerinin daha kolay ve hızlı bir şekilde hesaplanmasından dolayı basit en küçük kareler çözümlerini içermektedir. Ayrıca, regresyon splinelar parametrik olmayan regresyon modellerin tahmini için çekirdek (kernel) yöntemlere bir alternatiftir¹⁴².

¹³⁸ Endre Süli and David F. Mayers, *An Introduction to Numerical Analysis*, Cambridge University Press, New York 2003, s.292.

¹³⁹ Bayram, s.144.

¹⁴⁰ Shih-Huang Chan, *Polynomial Spline Regression with Unknown Knots and AR(1) Errors*, Ph. D. Dissertation, Ohio State University, Ohio 1989, s.2.

¹⁴¹ Lee, s.647.

¹⁴² Shujie Ma, Jeffrey S. Racine and Lijian Yang, "Spline Regression in the Presence of Categorical Predictors", Working Paper, Michigan State University, East Lansing, Michigan 2011, s.1.

Regresyon spline özellikleri aşağıdaki gibi sıralanabilir¹⁴³.

✓ Parametrik splinelar, parçalı polinomlardır ve tahmin ediciler belirlendikten sonra mevcut herhangi bir regresyon programı kullanılarak tahmin edilebilirler. Spline regresyon kesikli çıktılar için lojistik modeller, yaşam çözümlemesi modelleri ve çoklu doğrusal regresyona benzerdir.

✓ Spline fonksiyonlar için regresyon katsayıları en çok olabilirlik veya en küçük kareler gibi standart teknikler kullanılarak tahmin edilir ve istatistiksel çıkarım kolaylıkla yapılabilir. Tahmin edilen regresyon fonksiyonu için güven aralığı sınırları standart teoriyle elde edilebilir.

✓ Spline fonksiyonu, nihai modelde tahmin ediciyi göstermek için kullanılabilir. Parametrik olmayan yöntemler bir tahmin denklemi vermez.

1.7.2. Düzgünleştirici Spline (Smoothing Spline)

Regresyon fonksiyonlarını modellemek için parçalı polinomların kullanımı 1960'lı yıllara kadar gitmektedir. Ancak parametrik olmayan regresyondaki kullanımı olan regresyon spline düzgünleştirme ise nispeten daha yenidir. Regresyon spline düzgünleştirme aynı zamanda polinom spline düzgünleştirme ve B-spline düzgünleştirme olarak da bilinir¹⁴⁴.

Düzgünleştirme sorunu aslında, istatistiksel regresyon problemidir veya en küçük karelerle eğri uydurmadır¹⁴⁵. İlk olarak düzgünleştirme yaklaşımının temel fikri Whittaker (1923)'ün çalışmasına dayanmaktadır. Daha sonra düzgünleştirici splinelar Schoenberg (1964) ve Reinsch (1967) tarafından geliştirilmiştir¹⁴⁶. Düzgünleştirici spline, regresyon spline ve cezalı en küçük kareler yöntemine farklı bir yaklaşımdır. Temel mantığı, hata (artık) kareler toplamını minimize eden düzgün bir fonksiyon bulmaktır¹⁴⁷. Spline düzgünleştirme yöntemi iki ya da daha üst düzeyli polinomlar

¹⁴³ Frank E. Harrell, *Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression and Survival Analysis*, Springer-Verlag, New York 2001, s.26.

¹⁴⁴ Matt P. Wand, "A Comparison of Regression Spline Smoothing Procedures", *Computational Statistics*, 15(4), 2000, s.443-444.

¹⁴⁵ John F. Monahan, *Numerical Methods of Statistics*, Second Edition, Cambridge University Press, New York 2011, s.151.

¹⁴⁶ Holger Dette, Viatcheslav B. Melas and Andrey Pepelyshev, "Optimal Designs for Free Knot Least Squares Splines", *Statistica Sinica*, 18, 2008, s.1047.

¹⁴⁷ Fan Jianqing and Yao Quiwei, *Nonlinear Time Series: Nonparametric and Parametric Methods*, Springer-Verlag, New York 2003, s.251.

kullanıldığında elde edilen ve parametrik olmayan hata varyansının kontrol edilmesini sağlayan bir yöntemdir¹⁴⁸.

Düzgünleştirici splineda, düğüm noktası sayısı çok olabilir ancak parametrelerin tahmini en küçük kareler yöntemiyle yapılmaz. Düğüm noktası seçiminin önemi nispeten daha azdır. Düzgünleştirici splineda parametrelerin tahmini daha çok cezalı en küçük kareler yöntemiyle yapılmaktadır¹⁴⁹.

Buna göre, bir bağımlı değişken (y) ve bu bağımlı değişkenle ne tür bir ilişki içerisinde olduğu bilinmeyen bir bağımsız değişkenin (x) yer aldığı parametrik olmayan regresyon modeli aşağıdaki gibi tanımlanmaktadır¹⁵⁰.

$$y_i = f(x_i) + \varepsilon_i \quad a < x_1 < \dots < x_n < b \quad (1.44)$$

Eşitlik 1.2’de,

$f \in C^2[a, b]$: İki kez türevlenebilir bilinmeyen bir pürüzsüz fonksiyon

$(y_i)_{i=1}^n$: Bağımlı değişkene ait gözlem değerleri

$(x_i)_{i=1}^n$: Parametrik olmayan bağımsız değişkene ait gözlem değerleri

$(\varepsilon_i)_{i=1}^n$: Bağımsız ve normal olarak dağılan, σ^2 ortak varyanslı ve sıfır ortalamalı rassal hata terimidir.

Parametrik olmayan regresyonun amacı bilinmeyen $f(x)$ fonksiyonunu tahminle belirlemektir. Burada dikkate alınması gereken husus, gözlenen verilerden 1.44 modeline uygun fonksiyonun tahminidir. $f(x)$ fonksiyonunu tahmin etmek için en yaygın olarak kullanılan yöntemlerden biri de basit doğrusal regresyondur. Doğrusal regresyonda $f(x)$ fonksiyonu $\hat{f}(x) = a + bx$ biçiminde tahmin edilmektedir. Burada a sabit terim iken, b eğim katsayısıdır. $\hat{f}(x) = a + bx$ biçimindeki bütün fonksiyonlar için amaç,

$$RSS(f) = \sum_{i=1}^n [y - f(x)]^2 \quad (1.45)$$

¹⁴⁸ Barış Kaki, Abdullah Yeşilova ve Celaleddin Şen, “Yarı Parametrik Regresyon Yönteminin Hayvancılıkta Kullanılması”, 2004, Erişim Tarihi: 11 Ocak 2012, http://4uzbk.sdu.edu.tr/4UZBK/BGB/4UZBK_097.pdf, s.628.

¹⁴⁹ Lee, s.648.

¹⁵⁰ Aydın, *Semiparametrik Regresyon Modellemeye Splayn Düzeltme Yaklaşımı ile Tahmin ve Çıkarsamalar*, s.38.

ifadesi ile verilen hata kareler toplamını minimum yapmaktır. Eşitlik 1.45 modeline uygun gözlem verileri için f fonksiyonu yaklaşık olarak doğrusalsa tahmin için doğrusal regresyon yaklaşımı etkin olabilir. f fonksiyonu doğrusal değilse veri tahmini başarılı olamaz. Dolayısıyla, tahminde başarılı olabilmek için, doğrusal regresyonda ileri sürülen tahmin koşulları değiştirilmeli ve değişen eğilimli f fonksiyonları üzerinde hata kareler toplamının minimizasyonu dikkate alınmalıdır.

Bu durumda, hata kareler toplamını minimum yapacak fonksiyonlar kümesine katı parametrik kısıtlamalar yüklemeksizin, eğrinin pürüzsüzlüğü için bir ceza uygulanabilir. Yani burada uygulanacak olan yöntem cezalı en küçük kareler yöntemidir.

$$S(f) = \sum_{i=1}^n [y - f(x)]^2 + \lambda \int_a^b [f''(x)]^2 dx \quad (1.46)$$

Eşitlik 1.46'da amaç $S(f)$ cezalı en küçük kareler kriterini minimum yapmaktır. Burada ilk terim hata kareler toplamını gösterir ve bu ifade uyumdan yoksunluğu cezalandırır. Yani, uyumun verilere yakınlığını ölçer. İkinci terim pürüzlülük cezasını gösterir ve bu pürüzlülüğe ceza yükler. Yani, fonksiyondaki eğriliği cezalandırır. İkinci terimdeki λ ise düzgünleştirme parametresidir. Düzgünleştirme parametresi $\int_a^b [f''(x)]^2 dx$ ile ölçümlenen eğrinin pürüzlülüğü ve $\sum_{i=1}^n [y - f(x)]^2$ ile ölçümlenen verilere uyumunu dengeler.

λ parametresi 0 dan $+\infty$ 'a değişirken, çözüm interpolasyondan doğrusal modele doğru değişir. Şayet $\lambda = \infty$ alınırsa, eşitlik 1.46 sabit eğimli doğrusal regresyon uyumu üretir. $\lambda = 0$ alınırsa tümüyle esnek eğimli bir interpolasyon uyumuna karşı gelir¹⁵¹.

λ arttıkça düşük varyanslı ancak yanlış uyumlar elde edilirken, λ azaldıkça yüksek varyanslı ancak yansız uyumlar elde edilir. λ parametresi çok küçük ya da çok büyük değerler almadığında, ara değerler söz konusu olduğunda veriye uygulanan

¹⁵¹ Aydın, *Semiparametrik Regresyon Modellemeye Splayn Düzeltme Yaklaşımı ile Tahmin ve Çıkarımlar*, s.39-41.

düzgünleştirmenin ölçüsüne yorumlanabilir nitelikte bir katkısı olmaz. Bu problemin çözümü için λ parametresi serbestlik derecesi ile dönüştürülebilir¹⁵².

Ayrıca, düzgünleştirici splineda düğüm noktalarında birinci ve ikinci türevler sürekli olup, üçüncü türevde süreksiz olabilir. Fonksiyonda $(-\infty, x_0)$ ve (∞, x_n) aralıklarında ikinci türev sıfırdır ve böylece bu veri aralıklarının dışında fonksiyon doğrusaldır¹⁵³.

Düzgünleştirme parametresinin seçiminde kullanılan çapraz geçerlilik (cross-validation) yöntemi Wahba ve arkadaşları tarafından önerilmiş ve yaptıkları birçok çalışmada yöntem geliştirilmiştir. Çapraz geçerlilik yönteminin temel mantığı veri noktalarını bir kere dışarda tutarak, geriye kalan verilerle dışarda tutulan verileri en iyi tahmin eden λ değerini seçmektir¹⁵⁴.

1.7.3. Çok Değişkenli Spline (Multivariate Spline)

Çok değişkenli spline fonksiyonları, s-boyutlu veri setlerinin yorumlanmasında ve farklı parametrelerle tanımlanan matematiksel bir modelin gösteriminde kullanılmaktadır. Çok değişkenli splinelar aynı zamanda, araba ve uçak gövdeleri gibi karmaşık parçaları modelleme tasarım süreci için CAGD’de (Computer Aided Geometric Design) önemli bir rol oynamaktadır¹⁵⁵. İnce tabakalı spline ve Tensor çarpım spline yaygın olarak kullanılan çok değişkenli spline türleridir.

1.7.3.1. İnce Tabakalı Spline (Thin Plate Spline)

İnce tabakalı splinelar, cezalı en küçük kareler yaklaşımıyla çok değişkenli düzgünleştirme için temel araçtır¹⁵⁶. Aslında ince tabakalı splinelar genellikle, basit parametrik olmayan veya yarı parametrik modelleri tahmin etmek için gereksizdir. İnce tabakalı splinelar öncelikle, birden fazla boyutlu verileri düzgünleştirmek için kullanılır.

¹⁵² Seda Bağdatlı, *Semiparametrik Regresyon ve Bir Uygulama*, (Yayımlanmamış Yüksek Lisans Tezi), İstanbul Ticaret Üniversitesi Fen Bilimleri Enstitüsü, İstanbul 2010, s.16.

¹⁵³ B. W. Silverman, “Some Aspects of the Spline Smoothing Approach to Non-Parametric Regression Curve Fitting”, *Journal of the Royal Statistical Society. Series B (Methodological)*, 47(1), 1985, s.3.

¹⁵⁴ B. W. Silverman, “A Fast and Efficient Cross-Validation Method for Smoothing Parameter Choice in Spline Regression”, *Journal of the American Statistical Association*, 79(387), 1984, s.584.

¹⁵⁵ Russell Maw, *A Survey of Application of Spline Functions to Statistics*, Master’s Thesis, East Tennessee State University, Johnson City 2001, s.31.

¹⁵⁶ Chong Gu, “Multivariate Spline Regression”, (Ed. Michael G. Schimek), *Smoothing and Regression Approaches, Computation and Application*, John Wiley&Sons, New York 2000, s.350.

Fakat, aynı zamanda Markov Zinciri Monte Carlo yöntemiyle spline modelleri tahmin etmek için de kullanılmaktadır.

İnce tabakalı splinelar, düzgünleştirici splineların daha genel bir versiyonudur. Yani spline tahmininde yerleştirilen ceza herhangi bir türev mertebesi üzerinden yerleştirilebilir. İnce tabakalı spline aşağıdaki gibi ifade edilmektedir ¹⁵⁷.

$$SS(f, \lambda) = \sum_{i=1}^n y_i - f(x_i)^2 + \lambda \int_{x_i}^{x_n} f^m(x_i)^2 dx \quad (1.47)$$

Burada m , $2m > d$ olması koşuluyla m . türevi göstermektedir. d ise bağımsız değişken x_i 'nin boyutudur.

1.7.3.2. Tensor Çarpım Spline (Tensor Product Spline)

Bir boyutlu durum tensor çarpım spline yapılarının kullanımıyla çoklu boyutlara genişletilebilir. Tek değişkenli B-splineları kullanarak m değişkenli bir fonksiyonun çok değişkenli bir tensor çarpım spline ifadesi aşağıdaki gibi gösterilebilir ¹⁵⁸.

$$f(x_1, x_2, \dots, x_m) = \sum_{i_1=1}^{n_1} \dots \sum_{i_m=1}^{n_m} a_{i_1 \dots i_m} B_{i_1, k_1, t_1}(x_1) \dots B_{i_m, k_m, t_m}(x_m) \quad (1.48)$$

Bu eşitlik m boyutlu uzayda bir tensor çarpım fonksiyonunu göstermektedir ¹⁵⁹. Tek değişkenli splineların herhangi bir türü B-splineların yerinde kullanılabilir. Burada $B_{i,k,t}$, t düğüm noktası serisi için k mertebeli i . B-Spline, $a_{i_1 \dots i_m}$ ise katsayılarıdır.

¹⁵⁷ Luke Keele, *Semiparametric Regression for the Social Sciences*, John Wiley & Sons, London 2008, s.74.

¹⁵⁸ Balaji Kommineni, *Spline Center and Range Regression and its Application to Variation Aware Performance Macromodeling of Analog Circuits*, University of Cincinnati, Andhra Pradesh/India 2005, s.33.

¹⁵⁹ Erişim Tarihi: 4 Temmuz 2012, <http://www.iue.tuwien.ac.at/phd/khalil/node31.html>.

İKİNCİ BÖLÜM

SPLİNE REGRESYONDA DÜĞÜM NOKTASI

Quandt (1958)¹⁶⁰, en fazla bir tane sıçrama noktasının olduğu iki farklı sistemi içeren kesikli doğrusal regresyon modelinin parametrelerini ve doğrusal regresyon sistemindeki bilinmeyen değişim noktasının konumunu tahmin etmek için maksimum olabilirlik tahmin yöntemini önermiştir. Ayrıca sistemde değişikliğin ortaya çıktığı alternatif hipotezine karşın, sistemde değişikliğin olmadığı hipotezini test etmek için olabilirlik oran testi ve F dağılımına dayalı küçük örneklem testini önermiştir. Robison (1964)¹⁶¹, basit doğrusal regresyon yerine iki farklı sistemi içeren kesikli polinom regresyonların parametrelerini tahmin etmek ve polinom regresyonların birleşme noktalarının bilindiği ve bilinmediği durumlar için maksimum olabilirlik metodunu kullanmıştır. Polinom regresyon modellerin kesişim noktası için güven aralığının nasıl belirleneceğini göstermiştir. Her iki yazar da parçalı regresyon modelleriyle ilgili çalışmalarında tek bir değişme noktası üzerinde durmuşlardır.

Hudson (1966)¹⁶², genel modeli iki parçalı regresyon biçiminde tanımlayarak, parçalı regresyonların birleşme noktasını en küçük kareli çözümle bulmak için algoritmik bir yaklaşım geliştirmiştir. Hataların normal dağıldığı varsayıldığında en küçük kareler tahmin edicisi her zaman maksimum olabilirlik tahmin edicisiyle aynı sonucu verdiğinden en küçük kareler tekniğini kullanmıştır. Regresyonların birleşme noktasının farklı durumlarını dikkate alarak 1.tip, 2.tip ve 3.tip birleşme noktası olmak üzere sınıflandırma yapmıştır. Yani birleşme noktalarının x eksenindeki veri noktasında veya böyle noktalar arasında olup olmadığına ya da regresyonların birleşme noktalarında farklı katsayılarla sahip olup olmadığına bağlı olarak sınıflandırma yapmıştır.

¹⁶⁰ Richard E. Quandt, "The Estimation of the Parameters of a Linear Regression System Obeying Two Separate Regimes", *Journal of the American Statistical Association*, 53(284), 1958, 873-880.

¹⁶¹ D. E. Robison, "Estimates for the Points of Intersection of Two Polynomial Regressions", *Journal of the American Statistical Association*, 59(305), 1964, 214-224.

¹⁶² Derek J. Hudson, "Fitting Segmented Curves Whose Join Points Have to be Estimated", *Journal of the American Statistical Association*, 61(316), 1966, 1097-1129.

De Boor and Rice (1968)¹⁶³, düğüm noktalarının değişken olduğu durumlarda farklı düğüm noktalarını alarak en iyi sonucu elde etmeye çalışan bir algoritma geliştirmiştir.

Bellman ve Roth (1969)¹⁶⁴, birleşme noktaları bilinmeyen düzgün doğru parçalarını ard arda sıralayarak sürekli bir optimal eğri tahmini yapmak için dinamik programlama teorisini kullanan bir teknik önermişlerdir. Bu teknikle mümkün olan en iyi tahminin yapılabileceğini sayısal bir örnekle göstermişlerdir.

Hinkley (1969¹⁶⁵, 1971¹⁶⁶), bir bağımsız değişkenli iki parçalı doğrusal regresyon modelinde Hudson'un metodolojisini kullanarak bilinmeyen birleşme noktalarıyla ilgili olarak tahmin ve sonuç çıkarma problemini çalışmıştır. İki parçalı doğrusal regresyon modelinde kesişim noktalarının güven bölgelerini, maksimum olabilirlik tahminlerini ve güven aralıklarını elde etmek için yöntemler geliştirmiştir. Geliştirilen yöntemler monte carlo çalışmasıyla örnek bir veri seti üzerinde gösterilmiştir.

McGee ve Carleton (1970)¹⁶⁷, hiyerarşik kümeleme ve standart regresyon teorisini birleştirerek bir yöntem geliştirmişlerdir. Parçalı regresyonu incelemiş ve önerdikleri yöntemle parçalı regresyonun gerçek verilerle kullanımını göstermişlerdir. Finansal verilerle yapılan uygulama örneğinde, 01.11.1967 ve 30.11.1969 yılları arası aylık satış hacimleri dikkate alınarak, New York ve Amerikan Menkul Kıymetler Borsalarının Boston Menkul Kıymetler Borsası üzerindeki etkisi araştırılmıştır. Hiyerarşik kümeleme analiziyle birlikte parçalı doğrusal regresyon çözümü sonucu dört parçalı bir model belirlenmiştir. Ayrıca, Monte Carlo simülasyon çalışmasıyla da parçalı regresyonla ilgili örneklem dağılım sorunları incelenmiştir.

¹⁶³ Carl de Boor and John R. Rice, *Least Squares Cubic Spline Approximation II-Variable Knots*, Technical Report CSD TR 21, USA: Purdue University 1968.

¹⁶⁴ Richard Bellman and Robert Roth, "Curve Fitting by Segmented Straight Lines", *Journal of the American Statistical Association*, 64(327), 1969, 1079-1084.

¹⁶⁵ D. V. Hinkley, "Inference about the Intersection in Two-Phase Regression", *Biometrika*, 56(3), 1969, 495-504.

¹⁶⁶ D. V. Hinkley, "Inference in Two-Phase Regression", *Journal of the American Statistical Association*, 66(336), 1971, 736-743.

¹⁶⁷ Victor E. McGee and Willard T. Carleton, "Piecewise Regression", *Journal of the American Statistical Association*, 65(331), 1970, 1109-1124.

Bacon ve Watts (1971)¹⁶⁸, tek bir düğüm noktasının olduğu ancak düğüm noktasının bilinmediği iki farklı doğrusal ilişkiyi içeren regresyon modellerinde, iki doğrusal sistem arasında düzgün geçişe olanak sağlayan bayezyen bir tahmin yöntemi önermiştir.

Gallant ve Fuller (1973)¹⁶⁹, düğüm noktalarının sayısının ve polinom derecelerinin bilindiği ancak düğüm noktası konumlarının bilinmediği parçalı polinomların birleşme noktalarının tahmin edilmesi üzerine çalışma yapmışlardır. Süreklilik ve türev koşulları kullanılarak parçalı regresyon modelinin doğrusal olmayan regresyon modeline dönüştürülmesi için parametreler yeniden belirlenmiştir. Daha sonra regresyon parametrelerinin en küçük kareler tahminlerini elde etmek için Hartley'in Modifiye Edilmiş Gauss-Newton Metodunu kullanarak doğrusal olmayan regresyon modelini tahmin etmişlerdir. Ayrıca bu yaklaşımın uygulamaları için hipotez testini ve bazı asimptotik özellikleri de göstermişlerdir.

Goldfeld ve Quandt (1973)¹⁷⁰, tek bir sıçramanın olduğu zamana veya diğer değişkenlere bağlı kesikli (switching) regresyon modellerinde sıçrama noktasının ve regresyon parametrelerinin tahmini için maksimum olabilirlik tahmin yöntemini göstermiştir.

Halpern (1973)¹⁷¹, parçalı doğrusal regresyonda düğüm noktalarının sayısı ve konumlarını belirlemek için bayezyen yöntemleri kullanmıştır.

Ertel ve Fowlkes (1976)¹⁷², doğrusal spline ve çok parçalı regresyonda düğüm noktası sayısının ve konumlarının belirlenmesi için hata kareler toplamına dayalı algoritma geliştirmiştir. Uygulama örneğinde, 1929-1973 yılları arasında AT&T (Amerika'nın uluslararası telekomünikasyon şirketi) şirketinin ortalama yıllık hisse senedi fiyatlarının hisse senedi başına temerrüt faiz oranı ve hisse senedi başına

¹⁶⁸ David W. Bacon and Donald G. Watts, "Estimating the Transition between Two Intersecting Straight Lines", *Biometrika*, 58(3), 1971, 525-534.

¹⁶⁹ A. R. Gallant and Wayne A. Fuller, "Fitting Segmented Polynomial Regression Models Whose Join Points have to be Estimated", *Journal of the American Statistical Association*, 68(341), 1973, 144-147.

¹⁷⁰ Stephen M. Goldfeld and Richard E. Quandt, "The Estimation of Structural Shifts by Switching Regressions", *Annals of Economic and Social Measurement*, 2(4), 1973, 475-485.

¹⁷¹ Elkan F. Halpern, "Bayesian Spline Regression When the Number of Knots is Unknown", *Journal of the Royal Statistical Society Series B*, 35(2), 1973, 347-360.

¹⁷² James E. Ertel and Edward B. Fowlkes, "Some Algorithms for Linear Spline and Piecewise Multiple Linear Regression", *Journal of the American Statistical Association*, 71(355), 1976, 640-648.

kazanılan kazanca göre değişimi incelenmiştir. 1959 yılında modelde anlamlı bir değişiklik olduğu belirlenmiş ve iki parça halinde analiz yapılmıştır.

Hawkins (1976)¹⁷³, düğüm noktalarının bilinmediği parçalı regresyon modellerinin tahmin edilmesinde dinamik programlama ve hiyerarşik çözüm yöntemlerini kullanmıştır. Her iki yöntem de parçaların doğrusal olduğu duruma göre uygulanmıştır.

Muth ve Willsky (1977)¹⁷⁴, spline regresyonda düğüm noktasının belirlenmesi için genelleştirilmiş olabilirlik oranına dayalı bir yaklaşım önermiştir. Önerilen yöntemde doğrusal spline kullanılmış ve tek bir düğüm noktasının olduğu varsayımına göre işlemler yapılmıştır.

Agarwall ve Studden (1978)¹⁷⁵, düğüm noktalarının seçimi için doğrusal splineları kullanarak IMSE (Intergrated Mean Square Error)'ye dayalı bir algoritma geliştirmiştir.

Park (1978)¹⁷⁶, birçok uygulamada tek bir polinom fonksiyonla bağımlı değişkenin yeterince açıklanamadığını belirtmiştir. Yani, tek bir polinom foksiyonla tüm deney alanı üzerinde yeterli tahmin sağlanamaz. Bu durumlarda regresyon modelinin parçalı polinom regresyon model olarak düşünülmesi gerektiğini izah etmiştir. Bilinmeyen parametrelerin tahminlerinde D-optimal model kriterini önermiştir.

Maronna ve Yohai (1978)¹⁷⁷, iki değişkenli bir regresyonda beklenmedik değişimlerle ilgili olarak ortalamadaki sistematik değişikliğin belirlenmesi için olabilirlik oran testine dayalı bir test istatistiği önermiştir. Önerilen test istatistiğinin asimptotik özellikleri de gösterilmiştir.

¹⁷³ Douglas M. Hawkins, "Point Estimation of the Parameters of Piecewise Regression Models", *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 25(1), 1976, 51-57.

¹⁷⁴ Muth and Willsky, ss.1-26.

¹⁷⁵ Girdhar G. Agarwal and W. J. Studden, *An Algorithm for Selection of Design and Knots in the Response Curve Estimation By Spline Functions*, Technical Report No. 78-15, Purdue University, USA: Indiana 1978.

¹⁷⁶ Sung H. Park, "Experimental Designs for Fitting Segmented Polynomial Regression Models", *Technometrics*, 20(2), 1978, 151- 154.

¹⁷⁷ Ricardo Maronna and Victor J. Yohai, "A Bivariate Test for the Detection of A Systematic Change in Mean", *Journal of the American Statistical Association*, 73(363), 1978, 640-645.

Lerman (1980)¹⁷⁸, birleşme noktaları bilinmeyen parçalı regresyonları tahmin etmek için Grid-Search Metodunu önermiş ve Hudson (1966), Hinkley (1969,1971) tarafından çalışılan standart tekniklerle kıyaslama yapmıştır.

Tishler ve Zang (1981a¹⁷⁹, 1981b¹⁸⁰), sürekli bir bağımlı değişkenli ve bilinmeyen geçiş noktalı doğrusal parçalı regresyonun parametrelerini tahmin etmek için maksimum olabilirlik metoduna dayalı bir algoritma geliştirmişlerdir. Önerilen yöntem de olabilirlik fonksiyonunun analitik türevleri kullanılmış ve iki parçalı regresyon modellerinde değişim noktalarının maksimum veya minimum olduğu durumlar ele alınmıştır. Regresyon doğrusunun sürekli olması için türevlenemeyen maksimum veya minimum noktalara göre modeller oluşturulmuştur.

Smith (1982)¹⁸¹, spline regresyonda istatistiksel değişken seçim tekniklerinin uygulamasını göstermiştir. B-spline tabanını kullanarak geriye doğru düğüm noktası eleme algoritması ve regresyon spline tabanını kullanarak düğüm noktası seçimiyle ilgili algoritma geliştirmiştir.

Marsh, L. C. (1983¹⁸², 1986¹⁸³), spline regresyonda düğüm noktalarının sayısı ve konumlarının belirlenmesi için basit bir yöntem göstermiştir. Uygulanan yöntem için SAS programında adımsal (stepwise) regresyonu temel alan bir algoritma geliştirilmiştir.

Marsh (1987)¹⁸⁴, zaman serili polinom regresyon modellerinde adımsal regresyonla tahmin edilen düğüm noktalarının konumlarının doğruluğunu belirlemek için Gauss-Newton yaklaşımını kullanmıştır.

¹⁷⁸ P. M. Lerman, "Fitting Segmented Regression Models by Grid Search", *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 29(1), 1980, 77-84.

¹⁷⁹ Asher Tishler and Israel Zang, "A New Maximum Likelihood Algorithm for Piecewise Regression", *Journal of the American Statistical Association*, 76(376), 1981a, 980-987.

¹⁸⁰ Asher Tishler and Israel Zang, "A Maximum Likelihood Method for Piecewise Regression Models With a Continuous Dependent Variable", *Journal of the Royal Statistical Society (Applied Statistics)*, 30(2), 1981b, 116-124.

¹⁸¹ Patricia L. Smith, *Curve Fitting and Modeling with Splines Using Statistical Variable Selection Techniques*, NASA Contractor Report 166034, Langley Research Center, Virginia 1982.

¹⁸² Lawrence C. Marsh, "On Estimating Spline Regression", SAS Institute Inc. (Ed.), *Proceedings of the Eighth Annual SAS Users Group International Conference*, (pp. 723-728), New Orleans, Louisiana 1983.

¹⁸³ Lawrence C. Marsh, "Estimating the Number of and Location of Knots in Spline Regression", *Journal of Applied Business Research*, 2(3), 1986, 60-70.

¹⁸⁴ Lawrence C. Marsh, "Estimating Spline Knots in Time Series Polynomial Regression Models", *The Institute of Management Sciences and Operations Research Society of America*, (October 1987), St. Louis.

Foley ve Nielson (1989)¹⁸⁵, parametrik spline eğrilerde düğüm noktası dizilerinin seçimi için kontrol noktalarının geometrik durumunu dikkate alan yeni bir metot göstermiştir.

Friedman ve Silverman (1989)¹⁸⁶, parçalı doğrusal spline modellerde verilerle düğüm noktası konumunu belirlemek için ileri adımsal (forward stepwise) seçim yöntemini temel alan bir algoritma önermiştir.

Marsh, Maudgal ve Raman (1990)¹⁸⁷, SAS programında parçalı doğrusal ve yüksek dereceli spline regresyon modellerini tahmin etmek için birleşme noktalarının önceden bilinip bilinmediği durumlara göre kısıtlandırılmış regresyon ve doğrusal olmayan regresyon metotlarını göstermişlerdir. Birleşme noktasının önceden bilindiği durumlarda kısıtlandırılmış en küçük kareler yöntemini, önceden bilinmediği durumlarda doğrusal olmayan metotlar için Gauss-Newton ve Newton-Raphson algoritmalarını kullanmışlardır.

Larson (1992)¹⁸⁸, düğüm noktası sayısının bilindiği ancak konumlarının bilinmediği doğrusal spline regresyonda en küçük kareler yöntemiyle düğüm noktasının konumunu belirlemeye çalışmıştır. Bilgisayar simülasyonu ile elde edilen veriler üzerinde yöntemin uygulaması gösterilmiştir.

Dupuis (1994)¹⁸⁹, doğrusal spline regresyonda düğüm noktalarının konumu ve sayısının belirlenmesi için, Laplace transform tekniklerini kullanarak yeni bir yaklaşım geliştirmiştir.

Toraichi ve Kamada (1995)¹⁹⁰, eşit aralıklı verilerin en düzgün interpolasyonuna olanak sağlayan periyodik kareli spline fonksiyonun düğüm noktası konumlarını belirlemeye çalışmıştır.

¹⁸⁵ Thomas A. Foley and Gregory M. Nielson, "Knot Selection for Parametric Spline Interpolation", T. Lyche and L. L. Schumaker (Ed.), *Mathematical Methods in Computer Aided Geometric Design*, (pp.261-262), Academic Press, New York 1989.

¹⁸⁶ Jerome H. Friedman and Bernard W. Silverman, "Flexible Parsimonious Smoothing and Additive Modeling", *Technometrics*, 31(1), 1989, 3-21.

¹⁸⁷ Lawrence C. Marsh, Manjula Maudgal and Jaishankar Raman, "Alternative Methods of Estimating Piecewise Linear and Higher Order Regression Models using SAS Software", SAS Institute Inc. (Ed.), *Proceedings of the Fifteenth Annual SAS Users Group International Conference*, (pp. 523-527), Nashville, Tennessee 1990.

¹⁸⁸ Harold J. Larson, "Least Squares Estimation of Linear Splines with Unknown Knot Locations", *Computational Statistics & Data Analysis*, 13(1), 1992, 1-8.

¹⁸⁹ Deppie Janice Dupuis, *Knots in Spline Regression: Estimation and Inference Using Laplace Transform Techniques*, Ph. D. Disertation, University of New Brunswick, New Brunswick/Canada 1994.

Smith ve Kohn(1996)¹⁹¹, spline regersyonda düğüm noktalarını tercih etmek için Bayesyen değişken seçim yaklaşımını geliştirmiştir.

Lindstrom (1999)¹⁹², optimal düğüm noktası konumlarını bulmak için cezalı düğüm noktası konumuna dayalı yeni bir tahmin yöntemi göstermiştir. Modeller içinde düğüm noktası sayısını tercih etmek için Genelleştirilmiş Çapraz Geçerlilik (GCV) ve Akaike Bilgi Kriteri (AIC) kullanılmıştır.

Kim, Fay, Feuer ve Midthune (2000)¹⁹³, doğrusal spline regresyonda düğüm noktalarının sayısını tahmin etmek için permutasyon testlerini kullanmıştır.

Lee (2002)¹⁹⁴, spline regresyonda en iyi düğüm noktası setinin seçimi için genelleştirilmiş çapraz geçerlilik ve adımsal düğüm noktası seçim yöntemini kullanmıştır. Seçilen düğüm noktalarına göre en küçük karelerle regresyon tahmini yapılmıştır.

Ruppert (2002)¹⁹⁵, düğüm noktası sayılarının cezalı splineların performansı üzerindeki etkisini çalışmıştır. Düğüm noktası sayılarının otomatik seçimi için iki algoritma önermiştir.

Muggeo (2003)¹⁹⁶, konumu bilinmeyen bir veya daha fazla düğüm noktasının modelin parametresi olarak yer aldığı regresyon modellerinde parçalı terimlerin tahminini yapmıştır. Tahmin için doğrusallaştırma tekniği kullanılarak alternatif bir yaklaşım önerilmiştir.

Mao ve Zhao (2003)¹⁹⁷, spline regresyon modellerinde optimal düğüm noktası sayısının belirlenmesi için genelleştirilmiş çapraz geçerliliği kullanmıştır. Daha sonra

¹⁹⁰ Kazuo Toraichi and Masaru Kamada, “Knot Positions for the Smoothest Periodic Quadratic Spline Interpolation of Equispaced Data”, *Linear Algebra and Its Applications*, 221, 1995, 245-251.

¹⁹¹ Michael Smith and Robert Kohn, “Nonparametric Regression Using Bayesian Variable Selection”, *Journal of Econometrics*, 75(2), 1996, 317-344.

¹⁹² Mary J. Lindstrom, “Penalized Estimation of free-Knot Splines”, *Journal of Computational and Graphical Statistics*, 8(2), 1999, 333-352.

¹⁹³ Hyune-Ju Kim, Michael P. Fay, Eric J. Feuer and Douglas N. Midthune, “Permutation Tests for Joipoint Regression with Applications to Cancer Rates”, *Statistics in Medicine*, 19(3), 2000, 335-351.

¹⁹⁴ Thomas C. M. Lee, “On Algorithms for Ordinary Least Squares Regression Spline Fitting: A Comparative Study”, *Journal of Statistical Computation and Simulation*, 72(8), 2002, 647-663.

¹⁹⁵ David Ruppert, “Selecting the Number of Knots for Penalized Splines”, *Journal of Computational and Graphical Statistics*, 11(4), 2002, 735-757.

¹⁹⁶ Vito M. R. Muggeo, “Estimating Regression Models with Unknown Break-Points”, *Statistics in Medicine*, 22, 2003, 3055-3071.

¹⁹⁷ Wenxin Mao and Linda H. Zhao, “Free-Knot Polynomial Splines with Confidence Intervals”, *Journal Of The Royal Statistical Society Series B*, 65(4), 2003, 901-919.

bilinmeyen düğüm noktasının konumlarını maksimum olabilirlik tahmin yöntemini kullanarak doğrusal olmayan en küçük kareler çözümüyle elde etmiştir.

Toms ve Lesperance (2003)¹⁹⁸, ekolojik eşik değerlerini modellemek için parçalı regresyon kullanımını göstermişlerdir. Çalışmada dört farklı model kıyaslanarak, kırılma noktası tahmini için üç farklı biçimde güven aralığı hesaplanmıştır.

Bessaoud, Daures ve Molinari (2005)¹⁹⁹, lojistik regresyon modeli için değişken düğüm noktalı splinelerla yaşa bağlı olarak tüp bebek sahibi olabilme olasılığını incelemiştir. Spline fonksiyonların derecesi ve düğüm noktalarının sayısını model seçim yöntemleri kullanılarak belirlenmiştir.

Petunin, Djomina, Klyushin ve Savkina (2006)²⁰⁰, Çernobil kazasından etkilenen bir grubun radyasyona maruz kalma durumuna bağlı olarak kanser olasılığı üzerinde radyasyonun etkisini analiz etmek için spline regresyon modellerini kullanmıştır. İki düğüm noktalı doğrusal splinelerla yeni bir metot önerilmiştir. Düğüm noktalarının belirlenmesi için de bir algoritma geliştirilmiştir.

Yao ve Lee (2008)²⁰¹, cezalı spline regresyon tahmininde düğüm noktası konumlarının belirlenmesi için basit ve etkili bir yöntem olan eşit aralıklı (equally-spaced) metotları önermiştir.

Kim, Yu ve Feuer (2009)²⁰², doğrusal spline regresyonda düğüm noktası sayısını tahmin etmek için kullanılan permutasyon yönteminin asimtotik etkinliğini incelemiş ve Bayezyen Bilgi Kriteri (BIC), Akaike Bilgi Kriteri (AIC) ve Genelleştirilmiş Çapraz Geçerliliğe (GCV) dayalı yöntemlerle karşılaştırılmıştır.

¹⁹⁸ Judith D. Toms and Mary L. Lesperance, "Piecewise Regression: A Tool For Identifying Ecological Thresholds", *Ecology*, 84(8), 2003, 2034-2041.

¹⁹⁹ F. Bessaoud, J.P. Daures, N. Molinari, "Free Knot Splines for Logistic Models and Threshold Selection", *Computer Methods and Programs in Biomedicine*, 77(1), 2005, 1-9.

²⁰⁰ Yu. I. Petunin, E. A. Djomina, D. A. Klyushina and M. Yu. Savkina, "Spline-Regression Estimation of The Influence of Exposure Dose on Cancer Probability", *Cybernetics and Systems Analysis*, 42(3), 2006, 449-455.

²⁰¹ Fang Yao and Thomas C. M. Lee, "On Knot Placement for Penalized Spline Regression", *Journal of the Korean Statistical Society*, 37(3), 2008, 259-267.

²⁰² Hyune-Ju Kim, Binbing Yu, and Eric J. Feuer, "Selecting The Number of Change-Points in Segmented Line Regression", *Statistica Sinica*, 19(2), 2009, 597-609.

2.1. DÜĞÜM NOKTASININ TANIMI

Spline literatüründe değişim noktaları düğüm noktaları olarak adlandırılır²⁰³. Spline regresyon modellerinin oluşturulması için düğüm noktalarının belirlenmesi temel bir adımdır²⁰⁴. Bağımlı ve bağımsız değişken arasındaki ilişkiler doğrusal, eğrisel ve kübik şeklinde olabilir. Aynı bağımsız değişken üzerinde ilişkinin şeklinin değiştiği bağımsız değişken değerine düğüm noktası denir²⁰⁵. Düğüm noktası basitçe tanımlanacak olursa, bağımlı ve bağımsız değişken arasındaki ilişkide değişiklikleri ortaya çıkaran ve regresyon doğrusunu parçalara bölen noktalardır²⁰⁶. Bir başka ifadeyle, doğru parçalarının birleşme noktalarındaki değişken değerleri düğüm noktası olarak adlandırılır²⁰⁷. Düğüm noktası, veri kümesindeki tanım aralıklarından birinin sonu, bir diğerinin başlangıcıdır. Dolayısıyla düğüm noktası fonksiyonunu yönünün veya hareketinin değiştiği yerdir²⁰⁸.

Düğüm noktalarından önceki ve sonraki değerler için modeller ayrı ayrı çözülür ve tahmin edilen parçalar birleştirilir. Mesela bir düğüm noktası varsa, bütün gözlemler n ve birinci döneme ait gözlem sayısı n_1 , ikinci döneme ait gözlem sayısı n_2 olmak üzere, $n_1 + n_2 = n$ 'dir²⁰⁹.

Spline regresyonda polinom parçalarının derecesi, düğüm noktalarının konumu ve sayısı değişebilir²¹⁰. Optimal düğüm noktasına karar verme durumunun zor olması ve henüz ortak kabul gören bir kriter de bulunmamasına rağmen, bir çok durumda

²⁰³ G.A.F. Seber and C.J Wild, *Nonlinear Regression*, John Wiley&Sons, New Jersey 2003, s.481

²⁰⁴ Kulkarni and Narendranath, s.29.

²⁰⁵ Yasemin Özfalcı, *Çokdeğişkenli Uyarlanabilir Regresyon Kesitleri: MARS*, (Yayımlanmamış Yüksek Lisans Tezi), Gazi Üniversitesi Fen Bilimleri Enstitüsü, Ankara 2008, s.16.

²⁰⁶ Chein-Hesing Sung, *Estimation Of A Modified Linear Spline Regression: Theory and Application*, Ph. D. Dissertation, Wayne State University, Detroit 1985, s.3.

²⁰⁷ Devlin and Weeks, s.646.

²⁰⁸ Birce Ünal, *Çok Değişkenli Uyarlamalı Regresyon Uzanımları*, (Yayımlanmamış Yüksek Lisans Tezi), Hacettepe Üniversitesi Fen Bilimleri Enstitüsü, Ankara 2009, s.10.

²⁰⁹ Yerdelen, s.49.

²¹⁰ Gregory and Serono, s.2.

konumları iyi belirlenmiş birkaç düğüm noktası yeterli olabilir²¹¹. Verileri düzeltmede spline fonksiyonları kullanmanın başarısı uygun düğüm noktaları seçimine bağlıdır²¹².

Düğüm noktası sayısını artırma, genellikle spline fonksiyonun veriye daha iyi uydurulmasını sağlamasına rağmen, birkaç düğüm noktalı spline düğüm noktası çok olan splinedan genellikle daha düzgündür²¹³.

Doğrusal bir zaman serisi modelinde durağan olmayan eğim katsayısını incelemek için gerek duyulan istatistiksel araçlar olası değişikliklere bağlıdır. Tek bir değişimin olduğu regresyon denkleminde, değişikliğin tam konumu test edilmek istendiği zaman, önceki ve sonraki dönemlere ait hata kareler ortalaması karşılaştırılabilir²¹⁴.

2.2. DÜĞÜM NOKTASININ ÖNCEDEN BİLİNMESİ (SABİT DÜĞÜM NOKTASI)

Spline regresyon modelde düğüm noktasının konumu ve sayısı önceden biliniyorsa, her bir aralıkta doğrusal bir biçim olacağı için regresyon katsayılarının asimtotik özellikleri klasik doğrusal modellerle aynıdır²¹⁵. Düğüm noktaları bilinen splineları tahmin etmek, düğüm noktası bilinmeyen spline tahmin etmekten daha kolaydır. Çünkü düğüm noktalarının sayısı ve konumu önceden bilindiği zaman, regresyon denkleminin parametre tahminleri en küçük kareler yöntemiyle kolaylıkla tahmin edilebilir²¹⁶.

2.3. DÜĞÜM NOKTASI KONUMUNUN ÖNCEDEN BİLİNMEMESİ (DEĞİŞEN DÜĞÜM NOKTASI)

Çalışmaların çoğu düğüm noktası konumlarının sabit olduğu yani önceden bilindiği durumlar için yapılmışken, tahminlerin doğruluğunu artırmak için değişken

²¹¹ Martin Gregory, Hanno Ulmer, Karl P. Pfeiffer, Stefan Lang and Alexander M. Strasak, "A Set of SAS Macros for Calculating and Displaying Adjusted Odds Ratios (With Confidence Intervals) for Continuous Covariates in Logistic B-Spline Regression Models", *Computer Methods and Programs in Biomedicine*, 92, 2008, s.112.

²¹² Robert E. Smith, Joseph M. Price and Lona M. Howser, *A Smoothing Algorithm Using Cubic Spline Functions*, NASA Technical Note D-7397, National Aeronautics and Space Administration, Washington D.C.1974, s.1.

²¹³ Hurley, Hussy, McKeow and Addy, s.1.

²¹⁴ John U. Farley and Melvin J. Hinich, "A Test for a Shifting Slope Coefficient in a Linear Model", *Journal of the American Statistical Association*, 65(331), 1970, s.1320.

²¹⁵ Chan, s.12.

²¹⁶ Rudolf Freund, Ramon Littell and Lee Creighton, *Regression Using JMP*, SAS Institute, North Carolina 2003, s.192.

düğüm noktası durumları da dikkate alınmaktadır²¹⁷. Spline regresyonda düğüm noktalarının konumu ve sayısının tercihi önemli ve zor bir sorundur. Bundan dolayı düğüm noktası ve sayısının seçimi iyi çalışılması gereken konulardan biridir. En iyi spline regresyon modelini oluşturmak için düğüm noktalarının konumu ve sayısı iyi belirlenmelidir²¹⁸.

Değişken düğüm noktaları spline yaklaşımının etkinliğini artırırken, iki önemli dezavantajı da bulunmaktadır. Birincisi, düğüm noktaları için en iyi konumları bulmak çok zordur. Düğüm noktaları modele doğrusal olmayan bir biçimde girer ve hata kareler yüzeyi toplamında çoğu düğüm noktası vektörüne (tekrarlı düğüm noktalarının da olduğu) uygun olan sayısız yerel optimum (local optima) bulunur. İkincisi ise, tekrarlı düğüm noktalarını içeren optimum düğüm noktası vektörü tahmin edilen eğrinin düzgün olmayan bir yapıda olmasına sebep olur²¹⁹.

Düğüm noktası konumunun önceden bilinmediği olaylar için düğüm noktası sayısının bilindiği ve düğüm noktası sayısının bilinmediği gibi iki farklı durum söz konusudur. Düğüm noktası sayısının bilindiği ancak konumlarının bilinmediği durumda regresyon modeli doğrusal olmayan regresyon olur²²⁰. Düğüm noktasının kaç tane olacağı verilerin görsel olarak incelenmesiyle belirlenebiliyorsa, standart doğrusal olmayan en küçük kareler yöntemiyle düğüm noktası konumları tahmin edilebilir. Düğüm noktası sayıları görsel olarak belirlenemezse en iyi düğüm noktası yerini belirlemek için genel olarak kullanılan iki yaklaşım vardır. Ya problem, türev tabanlı optimizasyon algoritmaları genel amacın performansını daha iyi hale getirmek için değiştirilir ya da problemi çözmek için bir özel amaç optimizasyon algoritması geliştirilir. Özel amaçlı optimizasyon algoritmalarına adımsal (stepwise) düğüm noktası seçim algoritmaları örnek verilebilir²²¹.

Düğüm noktalarının konumunun bilinmediği durumlarda, düğüm noktaları verilerden tahmin edilebilecek parametreler olarak ele alınır²²².

²¹⁷ Muth and Willsky, s.2.

²¹⁸ Agarwal and Studden, s.4.

²¹⁹ Lindstrom, s.334.

²²⁰ Marsh and Cormier, s.5.

²²¹ Lindstrom, ss.335-336.

²²² Eubank, Eubank, *Approximate Regression Models and Splines*, s.34.

2.4. DÜĞÜM NOKTASININ BELİRLENMESİNDE KULLANILAN YÖNTEMLER

Düğüm noktası konumunun belirlenmesi ve sayısının seçimi çeşitli yöntemler ile gerçekleştirilebilir.

2.4.1. Noktasal Dağılımdan (Serpilme Diyagramı) Görsel Olarak Belirleme

Herhangi bir veri seti için düğüm noktalarının konum ve sayısının tercihi araştırmacının deneyim ve önsezisine göre de belirlenebilir²²³. Verinin noktasal dağılımı görsel olarak incelenerek düğüm noktaları belirlenebilir. Yani birçok durumda gözlenen veri noktaları, regresyonda kullanılan düğüm noktaları model için en iyi tahmini veren noktalar olmasa bile, eğrideki değişikliğe göre görsel olarak belirlenen düğüm noktaları kullanıldığı zaman istenilen doğrulukla tahmin elde edilebilir²²⁴.

Doğrusal splineler için düğüm noktaları, verilerin eğim katsayısında bir değişme olduğu noktalarda belirlenir. Kareli splineler için düğüm noktaları, verideki dönüm noktalarında veya yakın maksimum, minimum noktalarında belirlenir. Kübik splineler için düğüm noktaları, verideki dönüm noktalarına yakın olacak şekilde belirlenir. Düğüm noktaları birden fazla uç noktada (maksimum veya minimum) değildir ve herhangi iki düğüm noktası arasında bir dönüm noktası ortaya çıkar²²⁵.

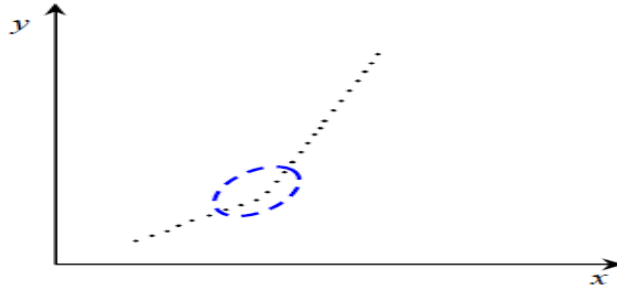
Noktasal dağılımı incelerken dikkat edilecek husus, ani yön değiştirmelerin olduğu bölgelerin muhtemel düğüm bölgeleri olmasıdır. Yani regresyon doğrusunun eğiminin değiştiği noktalardır. Ayrıca parçalamada kullanılacak fonksiyon da, noktasal dağılımın neresinde düğüm noktası/noktalarının oluşturulacağı konusunda belirleyici olacaktır. Buna göre, bazı noktasal dağılımlara ait muhtemel düğüm noktalarının olabileceği alanlar daire içerisinde Şekil 2.1, Şekil 2.2 ve Şekil 2.3'te gösterilmiştir²²⁶.

²²³ Ian Barrodale and Andrew Young, "A Note on Numerical Procedures for Approximation by Spline Functions", *Computer Journal*, 9(3), s.319.

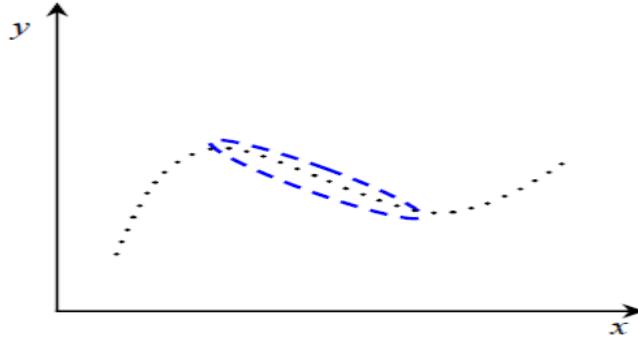
²²⁴ Guo and White, s.344.

²²⁵ Randall L. Eubank, *Spline Smoothing and Nonparametric Regression*, Marcel Dekker, New York 1988, s.357.

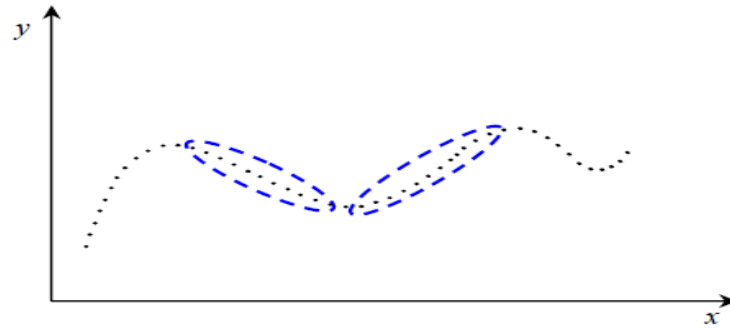
²²⁶ Şahin, ss.16-17.



Şekil 2.1. Doğrusal Spline Regresyon ve Muhtemel Düğüm Noktası Alanı



Şekil 2.2. Kareli Spline Regresyon ve Muhtemel Düğüm Noktası Alanı



Şekil 2.3. Kübik Spline Regresyon ve Muhtemel Düğüm Noktası Alanları

2.4.2. Adımsal (Stepwise) Regresyon Yöntemi

Adımsal (stepwise) yöntemler, düğüm noktalarının sayısı ve konumunu belirlemek için sıklıkla kullanılmaktadır. Adımsal regresyonla mantığı aynıdır. Spline regresyonda kullanılan adımsal yöntemde, en iyi alt değişkenler kümesini bulmak yerine, en iyi düğüm noktası sayısı ve konumları araştırılır²²⁷.

Spline regresyonda adımsal seçim ilk olarak Smith (1982) tarafından önerilmiştir²²⁸. Daha sonra Stone ve ark. (1997)²²⁹ tarafından bu yöntem geliştirilmiştir.

²²⁷ Katharine Lynn Gray, *Comparision of Trend Detection Methods*, Ph. D. Disertation in Mathematical Sciences, University of Montana, Missoula 2007, s.11.

²²⁸ Wand, s.446.

Marsh(1986), spline düğüm noktalarının sayısını ve konumunu tahmin etmek için SAS programında adımsal seçimi temel alan bir spline regresyon algoritması geliştirmiştir. Algoritmada eşitlik 2.1'deki n.dereceden ve k düğüm noktalı polinom spline regresyonun genel gösterimi kullanılarak düğüm noktası sayısı ve konumları belirlenmiştir.

$$y = \sum_{j=0}^n B_{0j}x^j + \sum_{i=1}^k \sum_{j=0}^n B_{ij}(x - t_i)_+^j + \varepsilon \quad (2.1)$$

Eşitlik 2.1'de y herhangi bir sürekli değişken, x ise açıklayıcı değişkeni ifade etmektedir. t_i , düğüm noktasıdır ($i = 1, 2, \dots, k$). k , özel x değerlerinin oluşturduğu düğüm noktaları kümesidir. B 'ler regresyon katsayıları ve ε ise regresyon varsayımlarının sağlandığı hata terimidir²³⁰.

Süreklilik kısıtlaması olmayan bu model, kısıtlandırılmamış gölge değişkenli model olarak adlandırılır. Her bir parçada kısıtlandırılmamış sabit terim ve katsayı değerleri vardır.

Örneğin 25 gözlemlik bir veri seti için kübik spline modelin istenildiği varsayılın. Eşitlik 2.1'e göre kısıtlandırılmamış gölge değişkenli kübik spline modeli aşağıdaki gibi yazılabilir.

$$y = B_{00} + B_{01}x + B_{02}x^2 + B_{03}x^3 + \sum_{i=1}^{25} \sum_{j=0}^3 B_{ij}(x - t_i)_+^j D_i + \varepsilon \quad (2.2)$$

Eşitlik 2.2'de düğüm noktalarının konumu için 1'den 25'e kadar bütün tamsayılar üzerinde inceleme yapılabilir. Her bir gözlem potansiyel düğüm noktası olarak ele alınır. Yani $t_i = 1$ ve $t_i = 25$ arasında 25 uygun düğüm noktası konumu vardır. Her bir olası düğüm noktası konumu için gölge değişkenler seti oluşturulur. $D_i = 1$ ve $D_i = 25$ arasında 25 gölge değişken oluşturulur. Burada $x < t_i$ ise, $D_i = 0$ ve $x \geq t_i$ ise $D_i = 1$ değerini alır. Ayrıca eşitlik 2'de üç grup spline değişken seti oluşturulur. Bunlar, doğrusal spline değişkenler $(x - t_i)D_i$, kareli spline değişkenler $(x - t_i)^2 D_i$ ve kübik spline değişkenler $(x - t_i)^3 D_i$ 'dir. Eşitlik 2.2'de çok sayıda katsayı vardır. Yani bu denklemde potansiyel olarak tahmin edilmesi gereken $104 \{(25 \times 4) + 4\}$ katsayı

²²⁹ C.J. Stone, M. Hansen, C. Kooperberg and Y.K. Truong, "Polynomial Splines and Their Tensor Products in Extended Linear Modeling", *The Annals of Statistics*, 25(4), 1997, 1371-1425.

²³⁰ Marsh,, "Estimating the Number of and Location of Knots in Spline Regression", s.61-62.

vardır. Adımsal (stepwise) seçim yöntemi, bu katsayılarından istatistiksel olarak anlamlı olanları seçmek için kullanılır.

SAS programında eşitlik 2.2 için programlama koşulları tanımlandıktan sonra, adımsal seçim yöntemi sonucu istatistiksel bakımdan anlamlı olanların seçilmesiyle veriyi tahmin etmek için farklı dereceli polinomların herhangi bir kombinasyonu kurulabilir²³¹.

2.4.3. Eşit Aralıklı Yöntem (Equally Spaced Method)

Bu yöntemde, düğüm noktası olarak $[a, b]$ gibi ilgilenilen aralıkta K adet eşit aralıklı nokta alınır. Yani, K düğüm noktaları,

$$\xi_r = a + (b - a)r/(K + 1), \quad r = 1, 2, \dots, K \quad (2.3)$$

olarak tanımlanır.

Bu düğüm noktalarını yerleştirme yöntemi tasarlanan zaman noktalarından (the design time points)²³² bağımsızdır. Bu yöntem genellikle, tasarlanan zaman noktalarının $[a, b]$ gibi ilgilenilen aralıkta eşit olarak dağıldığı düşünüldüğü zaman uygulanır²³³.

2.4.4. Eşit Aralıklı Örneklem Kantilleri Yöntemi (Equally Spaced Sample Quantiles Method)

Bu yöntemde, düğüm noktası olarak tasarlanan zaman noktalarının $(t_i, i = 1, 2, \dots, n)$ eşit aralıklı örneklem kantilleri kullanılır. $t_{(1)}, t_{(2)}, \dots, t_{(n)}$ tasarlanan zaman noktalarının sıra istatistikleri olsun. Buna göre, düğüm noktaları,

$$\xi_r = t_{(1 + \lceil rn/(K+1) \rceil)}, \quad r = 1, 2, \dots, K \quad (2.4)$$

olarak tanımlanır. Bu yöntem, tasarlanan zaman noktaları eşit dağıldığı zaman yaklaşık olarak eşit aralıklı yöntemle aynıdır²³⁴.

2.4.5. Diğer Yöntemler

Literatür araştırmasında da belirtildiği üzere düğüm noktası sayısı ve konumunun belirlenmesi için, farklı araştırmacılar tarafından En Küçük Kareler ve En Çok

²³¹ Marsh., “Estimating the Number of and Location of Knots in Spline Regression”, s.64-66.

²³² Bağımsız değişkenin zaman olarak alındığı durumlarda, zaman değerlerini belirtmektedir.

²³³ Hulin Wu, and Jin-Ting Zhang, *Nonparametric Regression Methods for Longitudinal Data Analysis Mixed Effects Modeling Approaches*, John Wiley&Sons, New Jersey 2006, s.53.

²³⁴ Wu, and Zhang, s.53.

Olabilirlik Yöntemi, Genelleştirilmiş Olabilirlik, Dinamik programlama, Gauss-Newton Metodu, Hata Kareler Toplamına dayalı algoritmalar, Hiyerarşik Çözüm Yöntemleri, Laplace Transform Teknikleri, Akaike Bilgi Kriteri, Mallow'un C_p Kriteri, Çapraz Geçerlilik, Genelleştirilmiş Çapraz Geçerlilik, Bayesyen Seçim, Adımsal Yöntem, Permutasyon Testleri gibi farklı yöntemler kullanılmıştır.

ÜÇÜNCÜ BÖLÜM

SPLİNE REGRESYON MODELLERİ

3.1. SPLİNE REGRESYONA GİRİŞ

Bazı durumlarda bağımlı değişken üzerindeki bağımsız değişkenin etkisi doğrusal değildir. Böyle problemler için basit doğrusal regresyon modeli uygulanamaz. Çünkü yapılan tahminler zayıf olur ve model varsayımları sağlanmaz. Dolayısıyla eğrisel ilişkinin değerlendirilmesinde, polinom ve doğrusal olmayan regresyon gibi spline regresyon yöntemleri de kullanılmaktadır²³⁵. Matematik alanında yoğun olarak kullanılan spline, regresyonda da kullanılmaktadır. Spline regresyon yöntemlerinde spline, regresyon modellerine göre hesaplanmaktadır²³⁶. Spline regresyon modelleri bir zaman doğrusundaki bükümleri ve sapmaları düzgünleştirmek ve tahmin etmek için kullanılan genel bir tekniktir. Spline regresyon modelleri karmaşık ve anlaşılması zor bir şey gibi görülebilmesine rağmen, aslında model üzerinde birkaç basit kısıtlamanın bulunduğu gölge değişkenli modellerdir²³⁷. Spline regresyon modelleri, gölge değişkenli modellere uygun bir alternatiftir. Bir modelin sabit katsayısı ve eğimini değiştirmek için gölge değişken kullanmak genellikle, regresyon doğrusunda bir kırılma olduğu zaman ortaya çıkmaktadır. Mesela bir firmanın büyük bir fabrikası kapanırsa, bu firmanın üretiminde önemli bir düşüş olur. Ancak daha güç farkedilen değişiklikler, esas itibariyle regresyon doğrusunda bir kırılma olmaksızın regresyon doğrusunun yönünün değiştiği noktada ortaya çıkabilir. Mesela, konut yakıtı vergisi ile ilgili bir değişikliğin yapılma zamanı konut yakıtı talebinde yapısal bir dönüm noktasını gösterebilir. Ancak talepte ani bir düşüş (veya artış) ile sonuçlanması beklenmeyebilir²³⁸.

Spline regresyon modelleri, gölge değişkenler kullanıldığı zaman ortaya çıkan sıçramaları ortadan kaldırarak bir zaman serisinin sürekli olmasını sağlar. Yapısal değişim içeren veri setine uygulandığı zaman tahminler için daha iyi sonuçlar elde edilir. Mesela bir şirketin üçer aylık veya yıllık satış tahminlerinin temeli geçmiş yıllara

²³⁵ Miroslav Kaps and William R. Lamberson, *Biostatistics for Animal Science*, 2nd Edition, MPG Books Group, London 2009, s.185.

²³⁶ Dupuis, s.1.

²³⁷ Lawrence C. Marsh and David R. Cormier, *Spline Regression Models*, Sage Publications, California 2002, s.1.

²³⁸ Marsh, "Estimating the Number of and Location of Knots in Spline Regression", s.60.

ait satış verilerine dayanır. Regresyon analizi kullanılarak, gelecekle ilgili öngörude bulunabilmek için trendler tahmin edilir. Satışlarda önemli yapısal değişiklikler ortaya çıktığı zaman, zamana bağlı satış tahminleri için spline regresyon yöntemleri önemli bir araçtır²³⁹.

Spline modeller, bağımsız (açıklayıcı) değişkenle bağımlı (yanıt) değişkeni arasındaki ilişkide değişimleri de dikkate alarak daha iyi bir tahmin sağlayabilir²⁴⁰. Spline regresyon yöntemleri, anlaşılması kolay olduğundan modellerin tahmininde yaygın bir şekilde kullanılmaktadır²⁴¹. Spline regresyon modellerini kullanmak birçok yönden avantajlıdır²⁴². Spline regresyon modelleri;

✓ Yüksek dereceden bir polinom fonksiyonla ifade edilebilecek noktasal bir dağılımın, daha küçük dereceden polinomlar kullanılarak ifade edilmesini sağlar.

✓ Herhangi bir aralık için, “x” değerine karşılık gelen “y” değeri tahmin edilmek istendiğinde, ilgili düğümler arasında tanımlanan düşük dereceli polinom kullanılarak aynı veri setinin diğer parçalarına uydurulan fonksiyonlarından bağımsız olarak bu tahmin yapılabilir.

✓ Spline regresyon modellerle noktasal dağılıma daha duyarlı ve daha esnek fonksiyonlar elde edilebilmektedir. Bu da noktasal dağılımlarda gözlem değerlerine yakın doğru veya eğrilerin elde edilmesi anlamına gelecektir. Yani verilere, gözlem noktalarını en iyi temsil eden fonksiyon uydurulmuş olur.

Özellikle spline regresyon modellerde doğrusallık varsayımlarının hiçbirinin sağlanması gerekmediği için, spline regresyon modelleri geleneksel parametrik regresyon yaklaşımlarına alternatif bir yöntem olarak önerilmektedir²⁴³. Farklı konumlarda ayrı düzgünleştirme dereceleriyle regresyon eğrisini tahmin edebilme avantajından dolayı, spline modellerin regresyon modelleri arasında farklı bir yeri vardır²⁴⁴.

Sürekli bir değişken farklı gruplara veya parçalara ayrıldığı zaman iki farklı durum ortaya çıkar. Ani kırılmaların olması durumunda modelin tam bir kesikli regresyon

²³⁹ Marcus, s.14.

²⁴⁰ Hurley, Hussy, McKeow and Addy, s.1.

²⁴¹ Mao and Zhao, s.901.

²⁴² Şahin, s.11-12.

²⁴³ Gregory, Ulmer, Pfeiffer, Lang and Strasak, s.110.

²⁴⁴ Holger Dette and Viatcheslav B. Melas, “A Note on All-Bias Designs with Applications in Spline Regression Models”, *Journal of Statistical Planning and Inference*, 140(7), 2010, s.2042.

olarak kalıp kalmayacağı ya da modelin spline regresyon modele dönüştürülüp dönüştürülmeyeceğidir.²⁴⁵ Spline regresyon modelleri kullanıldığı zaman, düğüm noktalarında düzgün eğim katsayılarının olması sağlanır ve verilerin sürekliliği korunur²⁴⁶. Yani spline regresyon modelle, orijinal fonksiyonun sürekli olmasını sağlayan değişiklikler yapılır²⁴⁷.

Regresyon analizlerinde, gölge değişkenli modeller, kesikli parçalı modeller ve spline modeller arasında yakın bir ilişki vardır²⁴⁸.

3.2. GÖLGE DEĞİŞKENLİ REGRESYON

Regresyon analizinde sıklıkla kullanılan en faydalı araçlardan biri de gölge değişkendir²⁴⁹. Gölge değişken, geleneksel olarak ırk, cinsiyet ve meslek gibi sayısal bir ölçekle ölçülemeyen nitel değişkenleri regresyon analizine dâhil etmek için kullanılan basit ve faydalı bir yöntemdir²⁵⁰. Çünkü sosyal bilimlerde araştırmacılar gölge değişkenleri sıklıkla, davranış, tutum ve sosyo-ekonomik özellikler üzerinde etkili olabilecek ırk farklılıkları, cinsiyet veya bölgesel farklılıklar gibi değişkenleri açıklamak için kullanırlar²⁵¹.

Gölge değişkenli modeller gözlenen veriler için iki veya daha fazla sınıflandırmanın yapıldığı ayrı regresyon doğrularını içerir. Gözlemler belirli nominal, sıralı ve sürekli değişkenlerin değerlerine göre gruplara bölünür. Mesela nominal değişken değeri cinsiyet olabilir. Bu durumda gözlemler, erkek ve bayan gurubu olmak üzere iki ayrı regresyon doğrusu oluşturulur. Aynı şekilde, gelir gibi sürekli bir değişken için önceden belirlenmiş bir kesim noktası verildiğinde, gözlemler düşük gelir sınıfı ve yüksek gelir sınıfı şeklinde iki ayrı regresyon doğrusu veya düşük, orta ve yüksek gelir sınıfları için üç ayrı regresyon doğrusu oluşturulabilir²⁵².

²⁴⁵ Mars and Cormier, s.7.

²⁴⁶ Marcus, s.15.

²⁴⁷ Marsh and Sindone, s.268.

²⁴⁸ Mars and Cormier, s.6.

²⁴⁹ William H. Greene, *Econometric Analysis*, Fifth Edition, Prentice Hall, New Jersey 2003, s.316.

²⁵⁰ Daniel B. Suits, "Use of Dummy Variables in Regression Equations", *Journal of the American Statistical Association*, 52(280), 1957, s.548.

²⁵¹ Melissa A. Hardy, *Regression With Dummy Variables*, Sage Publications, California 1993, s.2.

²⁵² Mars and Cormier, s.6-7.

Nitel değişkenler için gölge değişken kullanıldığı gibi bir zaman serisinde meydana gelen yapısal değişiklikler için de gölge değişken kullanılabilir. Yapısal değişiklik öncesi için sıfır ve sonrası için bir değeri alan gölge değişken tanımlanır²⁵³.

Trend adı verilen değişkenlerdeki uzun devre eğilimleri küçük değişiklikler gösterebilir. Bu değişiklikler modelin yapısında önemli değişiklikler meydana getirebilir. Değişkenlerde büyük artış veya azalışlar meydana getiren faktörler genellikle bir olay veya köklü bir değişiklik sonucunda değişkenin trendini değiştirir. Değişkenlerin trendinde önemli sayılabilecek artış veya azalış oluşturan değişikliklere yapısal değişiklik adı verilmektedir²⁵⁴.

Gölge değişkenli regresyon, spline regresyonun alternatifi olarak görülse de iki önemli dezavantajı vardır. Birincisi, gölge değişkenler keskin bir şekilde süresiz parçalı modeller oluşturur. İkincisi ise, süresizlikten dolayı basit doğrusal ve uyarlamalı fonksiyonlara yaklaşımda bu yöntemin etkili olmamasıdır²⁵⁵.

Gölge değişkenli regresyon modellerinde, ya regresyon modelinin sabitindeki değişmeyi, ya regresyon modelinin eğimindeki değişmeyi ya da hem regresyon modelinin sabiti hem de regresyon modelinin eğimindeki değişmeyi açıklamak gibi üç farklı durum için gölge değişken kullanılır²⁵⁶.

Bu üç farklı durum regresyon modelleriyle aşağıdaki gibi gösterilebilir²⁵⁷. Gölge değişken D ile ifade edilirse, gölge değişken kullanımı ile ilgili olarak, belirtilen gözlem, ilgili özelliğe sahip değilse gölge değişken “0”, ilgili özelliğe sahipse gölge değişken “1” değerini alır. Buna göre,

Sadece sabitteki değişmeyi açıklamak için kurulacak tek kategorili gölge değişkenli toplamsal regresyon modeli,

$$Y_i = \alpha + \beta X_i + \gamma D_i + \varepsilon_i \quad i = 1, 2, 3, \dots, t \quad (3.1)$$

$$D = 0 \text{ için, } Y_i = \alpha + \beta X_i + \varepsilon_i$$

²⁵³ Marcus, s.14.

²⁵⁴ Aydın Karakoca, *Regresyon Modellerinde Yapısal Değişiklik ve Model Seçimi*, (Yayımlanmamış Yüksek Lisans Tezi), Selçuk Üniversitesi Fen Bilimleri Enstitüsü, Konya 2003, s.19.

²⁵⁵ Fox, , s.239.

²⁵⁶ Robert S. Pindyck and Daniel L. Rubinfeld, *Econometric Models and Economic Forecasts*, Third Edition, McGraw-Hill, New York 1991, s.117.

²⁵⁷ Karakoca, s.33.

$$D = 1 \text{ için, } Y_i = (\alpha + \gamma) + \beta X_i + \varepsilon_i$$

şeklinde ifade edilir.

Sadece eğimdeki değişmeyi açıklamak için kurulacak tek kategorili gölge değişkenli çarpımsal regresyon modeli,

$$Y_i = \alpha + \beta X_i + \gamma D_i X_i + \varepsilon_i \quad i = 1, 2, 3, \dots, t \quad (3.2)$$

$$D = 0 \text{ için, } Y_i = \alpha + \beta X_i + \varepsilon_i$$

$$D = 1 \text{ için, } Y_i = \alpha + (\beta + \gamma) X_i + \varepsilon_i$$

şeklinde ifade edilir.

Sabit ve eğimdeki değişmeyi açıklamak için kurulacak tek kategorili gölge değişkenli birleşimli (toplamsal ve çarpımsal) regresyon modeli,

$$Y_i = \alpha + \beta X_i + \gamma D_i + \delta D_i X_i + \varepsilon_i \quad i = 1, 2, 3, \dots, t \quad (3.3)$$

$$D = 0 \text{ için, } Y_i = \alpha + \beta X_i + \varepsilon_i$$

$$D = 1 \text{ için, } Y_i = (\alpha + \gamma) + (\beta + \delta) X_i + \varepsilon_i$$

şeklinde ifade edilir.

Eşitlik 3.3'e göre, γ istatistik bakımdan anlamlıysa yani γ katsayısı sıfırdan farklıysa ikinci regresyon denkleminin sabiti $\alpha + \gamma$ ile elde edilir. Bu durumda ilk regresyon denkleminin sabiti de α olur. Diğer taraftan, γ istatistiki bakımdan anlamlı bulunmazsa her iki denklemin ortak sabit değeri α olur. Regresyon denklemleri eğimleri bakımından incelendiğinde, δ istatistik bakımdan anlamlıysa yani δ katsayısı sıfırdan farklıysa ikinci regresyon denkleminin eğimi $\beta + \delta$ ile elde edilir. Bu durumda ilk regresyon denkleminin eğim katsayısı da β olur. Diğer taraftan, δ istatistiki bakımdan anlamlı bulunmazsa her iki denklemin ortak eğim katsayı değeri β olur²⁵⁸.

3.3. KESİKLİ PARÇALI REGRESYON

Bir veri setinin daha iyi temsil edilebilmesi için, verinin alt kümelerinin bir dizisine polinomlar uydurulabilir. Bu polinom tipinin dezavantajı, tüm veri seti için polinomların ve türevlerinin sürekliliği sağlanamaz²⁵⁹. Kesikli regresyon ve spline

²⁵⁸ Gujarati, s.51.

²⁵⁹ Smith, Price and Howser, s.1.

regresyon arasındaki önemli farklılık, kesikli regresyon yöntemleri ani veya beklenmedik değişiklikler üzerinde dururken, spline yöntemler ani değişiklikler arasındaki en düzgün geçişler üzerinde durmaktadır. Bundan dolayı splinelar sürekli fonksiyonlar olarak tanımlanı²⁶⁰.

3.3.1. Parametreleri En Küçük Kareler Yöntemiyle Tahmin Edilen Kesikli Regresyon

Veri noktaları $(\underline{x}_1, y_1), (\underline{x}_2, y_2), \dots, (\underline{x}_n, y_n)$ biçiminde ifade edilsin. $\underline{X}$ ile $\underline{Y}$ arasında farklı aralıklar üzerinde belirlenmiş

$$Y_t = E[\underline{Y}|\underline{X} = \underline{x}_t] + \varepsilon_t \quad (3.4)$$

biçiminde bir regresyon fonksiyonu ele alınsın²⁶¹. Burada $E[\underline{Y}|\underline{X}] = f(\underline{x}, \underline{\theta})$ fonksiyonu farklı aralıklarda tanımlanmış farklı eğrilerin birleştirilmesiyle oluşturulmuştur ve

$$f(\underline{x}, \underline{\theta}, \underline{\alpha}) = \begin{cases} f_1(\underline{x}, \underline{\theta}_1), & x \leq \alpha_1 \\ f_2(\underline{x}, \underline{\theta}_2), & \alpha_1 < x \leq \alpha_2 \\ \vdots & \vdots \\ f_D(\underline{x}, \underline{\theta}_D), & \alpha_{D-1} < x \end{cases} \quad (3.5)$$

biçimindedir²⁶².

α ile belirlenmiş aralıkların uç noktaları bilinmemekte ve bu noktalara düğüm noktası adı verilmektedir. D alt modeli ortak noktaların ya da düğüm noktalarının uç noktalarında modelin safhasını belirtir. Bu problem aynı zamanda parçalı regresyon problemi olarak da bilinir. Bu tür modeller, sistemdeki parçalı modellerin sayısının küçük olması, sistemin davranışının bir doğru fonksiyonu olarak iyi tanımlanmış olması ve gözlemlerde ani değişimler olması durumları için amaçlanır.

Eşitlik 3.5 ile verilen model parçalı regresyon modeli olarak bilinir. Burada, düzgün regresyon ilişkisinden daha ziyade regresyondaki yapısal değişimler ya da ani değişimleri tanımlayan model üzerinde durulmaktadır.

²⁶⁰ Mars and Cormier, s.8.

²⁶¹ Aşır Genç ve Ark., "Parçalı Regresyon Kullanılarak Güneş Radyasyon Verilerinin Değerlendirilmesi", *VI. Türk-Alman Enerji Sempozyumu: 21-24 Haziran*, Ege Üniversitesi, İzmir 2001, s.272.

²⁶² Seber and Wild, s.433.

Sıralı $x_1 < x_2 < \dots < x_n$ verileri için,

$$y_i = E[y|x_i] = \begin{cases} \beta_{10} + \beta_{11} x + \beta_{12} x^2, & x \leq \alpha \\ \beta_{20} + \beta_{21} x + \beta_{22} x^2, & x > \alpha \end{cases} \quad (3.6)$$

biçimindeki model göz önüne alınsın. Eğer $E[y|x_i]$ sürekli olması kısıtlaması yoksa α değişim noktasının tahmini beklenmez. Bu durumda $i=K$ model sayısı olmak üzere model,

$$y_i = \begin{cases} \beta_{10} + \beta_{11} x_i + \beta_{12} x_i^2 + \varepsilon_{1i} & , i = 1, 2, \dots, K \\ \beta_{20} + \beta_{21} x_i + \beta_{22} x_i^2 + \varepsilon_{2i} & , i = K + 1, \dots, n \end{cases} \quad (3.7)$$

şeklinde olur. Eğer K parametre değeri önceden biliniyorsa problem basit lineer regresyon problemine dönüşür. Eğer K bilinmiyorsa sonuç çıkarım işlemlerinde birçok sıkıntılar ortaya çıkabilir. Burada $\varepsilon_{di} \sim N(0, \sigma_d^2)$, $\sigma_1^2 = \sigma_2^2$ ve ε_1 ve ε_2 bağımsız oldukları varsayılacaktır. Bu durumda $\underline{\beta}_1$ ve $\underline{\beta}_2$, σ_1^2 ve σ_2^2 bilinmeyen parametreleri tahmin edilecektir.

Matris formunda, $\underline{y}_1^{(K)}: K \times 1$ ve $\underline{y}_2^{(K)}: (n - K) \times 1$ olsun. Ayrıca,

$$\underline{y} = \begin{pmatrix} \underline{y}_1^{(K)} \\ \underline{y}_2^{(K)} \end{pmatrix}, \quad X = \begin{pmatrix} X_1^{(K)} & 0 \\ 0 & X_2^{(K)} \end{pmatrix}$$

$$\underline{\beta} = \begin{pmatrix} \underline{\beta}_1 \\ \underline{\beta}_2 \end{pmatrix}, \quad \underline{\varepsilon} = \begin{pmatrix} \varepsilon_1^{(K)} \\ \varepsilon_2^{(K)} \end{pmatrix}$$

olmak üzere regresyon modeli,

$$\underline{y} = X\underline{\beta} + \underline{\varepsilon} \quad (3.8)$$

biçiminde olur. Model parametresi $\underline{\beta}'$ 'nın en küçük kareler tahmin edicisi

$$\hat{\underline{\beta}}^{(K)} = (X'X)^{-1}X'\underline{y} \quad (3.9)$$

ile tahmin edilir²⁶³. Buna göre,

$$\hat{\underline{\beta}}_d^{(K)} = \left(X_d^{(K)'} X_d^{(K)} \right)^{-1} X_d^{(K)'} \underline{y}_d^{(K)}, \quad d = 1, 2 \quad (3.10)$$

²⁶³ Genç ve Ark., s.273.

$$\hat{\sigma}_1^2(K) = \frac{\hat{S}_1^{(K)}}{K}$$

$$\hat{\sigma}_2^2(K) = \frac{\hat{S}_2^{(K)}}{n - K}$$

$$\hat{S}_d^{(K)} = \left(\underline{y}_d^{(K)} - X_d^{(K)} \hat{\underline{\beta}}_d^{(K)} \right)' \left(\underline{y}_d^{(K)} - X_d^{(K)} \hat{\underline{\beta}}_d^{(K)} \right) \quad (3.11)$$

modelin hata kareler toplamı $\sigma_1^2 = \sigma_2^2$ olduğundan

$$\hat{\sigma}^2(K) = \frac{\hat{S}_1^{(K)} + \hat{S}_2^{(K)}}{n}$$

dır²⁶⁴. Bu durumda

$$\hat{\underline{\theta}}_d \sim N_{P_d}(\underline{\theta}_d, \hat{\sigma}^2(X_d'X_d)^{-1}) \quad (3.12)$$

şeklinde bir dağılım gösterir ve

$$\hat{\sigma}^2 = \frac{S(\hat{\underline{\theta}})}{n - p}$$

olarak elde edilir. Buradan,

$$\sqrt{n}(\hat{\underline{\theta}} - \underline{\theta}^*) \sim N_p(\underline{0}, \hat{\sigma}^2(X'X)^{-1}) \quad (3.13)$$

yazılabilir. Nihai olarak parametreler ve model hakkında bir istatistiksel sonuç çıkarımı yapmak mümkün olmaktadır²⁶⁵.

3.3.2. Parametreleri En Çok (Maksimum) Olabilirlik Yöntemiyle Tahmin Edilen Kesikli Regresyon

Kesikli regresyon modellerinde, regresyon parametrelerinin yanı sıra düğüm (kesim) noktası da maksimum olabilirlik fonksiyonu kullanılarak hesaplanabilir²⁶⁶.

Değişkenlerin zamana göre sıralandığı ve tek bir düğüm noktasının olduğu iki parçalı regresyon modeli ilk olarak Quandt(1958) tarafından ele alındı. Buna göre iki parçalı regresyon,

$$Y_i = \beta_{10} + \beta_{11}x_i + \varepsilon_{1i}, \quad i = 1, 2, 3, \dots, t$$

²⁶⁴ Seber and Wild, s.439-440.

²⁶⁵Ufuk Karadavut ve ark. "Parçalı Regresyon Yardımı ile Bitki Boyu-Zaman İlişkisi Parametrelerinin Tahmini", *Sakarya Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, 8(1), 2004, s.115-116.

²⁶⁶ Koldere, s. 57.

$$Y_i = \beta_{20} + \beta_{21}x_i + \varepsilon_{2i} \quad i = t + 1, \dots, n \quad (3.14)$$

biçiminde ifade edilmektedir. Burada ε_{1i} ve ε_{2i} hata terimlerinin normal dağılımlı olduğu, sıfır ortalamalı ve σ_1^2 ve σ_2^2 varyansları ile bağımsız olarak dağıtıldıkları varsayılmaktadır²⁶⁷.

ε_1 ve ε_2 için normal dağılım fonksiyonları,

$$\begin{aligned} f(Y_i) &= \frac{1}{\sqrt{2\pi}\sigma_1} \exp \left[-\left(\frac{1}{2\sigma_1^2} \right) (Y_i - \beta_{10} - \beta_{11}x_i)^2 \right] \\ f(Y_i) &= \frac{1}{\sqrt{2\pi}\sigma_2} \exp \left[-\left(\frac{1}{2\sigma_2^2} \right) (Y_i - \beta_{20} - \beta_{21}x_i)^2 \right] \end{aligned} \quad (3.15)$$

şeklinde yazılır. Normal dağılım fonksiyonlarından, ilk parçayı oluşturan t adet gözlem ve ikinci parçayı oluşturan n-t adet gözlem için olabilirlik fonksiyonları,

$$\begin{aligned} f(Y_1, Y_2, \dots, Y_t) &= \left(\frac{1}{\sqrt{2\pi}\sigma_1} \right)^t \exp \left(-\frac{1}{2\sigma_1^2} \sum_{i=1}^t (Y_i - \beta_{10} - \beta_{11}x_i)^2 \right) \\ f(Y_{t+1}, Y_{t+2}, \dots, Y_n) &= \left(\frac{1}{\sqrt{2\pi}\sigma_2} \right)^{n-t} \exp \left(-\frac{1}{2\sigma_2^2} \sum_{i=t+1}^n (Y_i - \beta_{20} - \beta_{21}x_i)^2 \right) \end{aligned}$$

biçiminde yazılabilir²⁶⁸. Tüm örneklemin olabilirlik fonksiyonu ise,

$$\begin{aligned} f(Y_1, \dots, Y_t, Y_{t+1}, \dots, Y_n) &= \left(\frac{1}{\sqrt{2\pi}\sigma_1} \right)^t \left(\frac{1}{\sqrt{2\pi}\sigma_2} \right)^{n-t} \exp \left(-\frac{1}{2\sigma_1^2} \sum_{i=1}^t (Y_i - \beta_{10} - \beta_{11}x_i)^2 - \right. \\ &\quad \left. - \frac{1}{2\sigma_2^2} \sum_{i=t+1}^n (Y_i - \beta_{20} - \beta_{21}x_i)^2 \right) \end{aligned} \quad (3.16)$$

biçiminde yazılır²⁶⁹.

Olabilirlik fonksiyonunun logaritması alındığında,

$$\begin{aligned} L &= -t \log(\sqrt{2\pi}\sigma_1) - (n-t) \log(\sqrt{2\pi}\sigma_2) - \left(\frac{1}{2\sigma_1^2} \right) \sum_{i=1}^t (Y_i - \beta_{10} - \beta_{11}x_i)^2 - \\ &\quad - \left(\frac{1}{2\sigma_2^2} \right) \sum_{i=t+1}^n (Y_i - \beta_{20} - \beta_{21}x_i)^2 \end{aligned}$$

ifadesi elde edilir. Bu eşitlikte gerekli işlemler yapılırsa,

²⁶⁷ R. E. Quandt, "Tests of the Hypothesis that a Linear Regression System Obeys Two Separate Regimes", *Journal of the American Statistical Association*, 55(290), 1960, s. 324.

²⁶⁸ Quandt, The Estimation of the Parameters of a Linear Regression System Obeying Two Separate Regimes, s.874.

²⁶⁹ S. M. Goldfeld and R. E. Quandt, *Nonlinear Methods in Econometrics*, North-Holland Publishing Company, Amsterdam 1972, s.259.

$$L = -t \log \sqrt{2\pi} - t \log \sigma_1 - n \log \sqrt{2\pi} - n \log \sigma_2 + t \log \sqrt{2\pi} + t \log \sigma_2 - \left(\frac{1}{2\sigma_1^2}\right) \sum_{i=1}^t (Y_i - \beta_{10} - \beta_{11}x_i)^2 - \left(\frac{1}{2\sigma_2^2}\right) \sum_{i=t+1}^n (Y_i - \beta_{20} - \beta_{21}x_i)^2$$

olur. Nihai olarak,

$$L = -n \log \sqrt{2\pi} - t \log \sigma_1 - (n - t) \log \sigma_2 - \left(\frac{1}{2\sigma_1^2}\right) \sum_{i=1}^t (Y_i - \beta_{10} - \beta_{11}x_i)^2 - \left(\frac{1}{2\sigma_2^2}\right) \sum_{i=t+1}^n (Y_i - \beta_{20} - \beta_{21}x_i)^2 \quad (3.17)$$

biçiminde yazılır. Bu eşitlik aynı zamanda,

$$\ln L = -\frac{n}{2} \ln 2\pi - \frac{t}{2} \ln \sigma_1^2 - \frac{n-t}{2} \ln \sigma_2^2 - \left(\frac{1}{2\sigma_1^2}\right) \sum_{i=1}^t (Y_i - \beta_{10} - \beta_{11}x_i)^2 - \left(\frac{1}{2\sigma_2^2}\right) \sum_{i=t+1}^n (Y_i - \beta_{20} - \beta_{21}x_i)^2 \quad (3.18)$$

şeklinde de ifade edilebilir.

Burada t olabilirlik fonksiyonunu maksimum yapan değerdir²⁷⁰. β_{10} , β_{11} , β_{20} ve β_{21} 'e göre kısmi türevler alınıp sıfıra eşitlenerek $\widehat{\beta}_{10}$, $\widehat{\beta}_{11}$, $\widehat{\beta}_{20}$ ve $\widehat{\beta}_{21}$ tahminleri elde edilir. Daha sonra σ_1 ve σ_2 'ye göre kısmi türev alınıp sıfıra eşitlenir. Daha sonra $\widehat{\beta}_{10}$, $\widehat{\beta}_{11}$, $\widehat{\beta}_{20}$ ve $\widehat{\beta}_{21}$ tahminleri yerine yazılırsa $\hat{\sigma}_1^2$ ve $\hat{\sigma}_2^2$ elde edilir. Elde edilen değerler eşitlik 3.17'de yerine yazılırsa, n gözlem için maksimum olabilirlik logaritması aşağıdaki gibi hesaplanır²⁷¹.

$$L = -n \log \sqrt{2\pi} - t \log \hat{\sigma}_1 - (n - t) \log \hat{\sigma}_2 - \frac{n}{2} \quad (3.19)$$

²⁷⁰ George G. Judge, William E. Griffiths, R. Carter Hill, Helmut Lütkepohl and Tsoung-Chao Lee, *The Theory and Practise of Econometrics*, Second Edition, John Wiley&Sons, New York 1985, s.804.

²⁷¹ Quandt, "The Estimation of the Parameters of a Linear Regression System Obeying Two Separate Regimes", s.875.

3.4. SPLİNE REGRESYON ÇEŞİTLERİ

Spline regresyon modelleri yüksek mertebeli polinom regresyon, doğrusal olmayan ve üstel modellere iyi bir alternatif olarak kullanılabilir. Verilerin yapısına bağlı olarak uygun eğriler tahmin edilmesinde basit regresyon modellerinden daha esnektir²⁷². Spline regresyon modelleri daha esnektir ancak modellerin etkinliği düğüm noktası sayısı, konumu ve polinomun derecesine bağlıdır. Polinom derecesinin tanımlanması, verilere en uygun tahmini elde etmek için aralıkların özelliklerine uygun polinom fonksiyon türünü belirlemektir²⁷³.

Spline regresyonlar, bir regresyon modeline göre hesaplanan splinelardır²⁷⁴. Kısıtlandırılmamış polinom spline genel formu aşağıdaki gibidir²⁷⁵.

$$P(x) = \sum_{i=0}^n a_i x^i + \sum_{i=1}^k \sum_{j=0}^n b_{ij} (x - \xi_i)_+^j + u_t \quad (3.20)$$

Eşitlik 3.20’de P, $i = 0, 1, \dots, k$ olmak üzere her bir aralıkta (ξ_i, ξ_{i+1}) n.dereceden sıralı bir polinomdur. Ayrıca P polinomunun düğüm noktalarında $(n - 1)$. dereceden türevleri süreklidir.

Bu modelde her bir spline parçanın kendine ait sabit terimi ve herhangi bir düğüm noktasında bir veya daha üstü türevlerde süreksizliğe yol açabilen eğim katsayıları vardır²⁷⁶.

Verilerden spline parametrelerinin doğrudan nasıl tahmin edileceği aşağıda gösterilmektedir²⁷⁷. Kısıtlandırılmış spline genel denklemini;

$$S^n(x) = \sum_{j=0}^n a_j x^j + \sum_{t=1}^k d_t (x - \xi_t)_+^n + u_t \quad (3.21)$$

şeklinde yazmak mümkündür. Burada,

a : Bağımsız değişkenlerin parametre katsayısı

²⁷² Aggrey, s.1787.

²⁷³ Sarvida and Baguio, s.2.

²⁷⁴ B. Berberoğlu and C. N. Berberoğlu, s.62.

²⁷⁵ Smith, “Splines As a Useful and Convenient Statistical Tool”, s.58.

²⁷⁶ Dennis Markov, *Information Content in Stock Market Technical Patterns: A Spline Regression Approach*, Ph. D. Dissertation, University of Notre Dame, Indiana 2002, s.42.

²⁷⁷ Gnad, s.70-71.

x : Bağımsız değişken

ξ : Düğüm noktası

d : Düğüm noktası parametre katsayısı

u : Tüm regresyon varsayımlarının sağlandığı hata terimi

n : Spline fonksiyonun derecesi' dir.

Ayrıca,

$$(x - \xi_t)_+^n = \begin{cases} 0, & x \leq \xi_t \\ x - \xi_t, & x > \xi_t \end{cases} \quad \text{koşulu vardır.} \quad (3.22)$$

Burada düğüm noktaları verilmiştir. Yani ξ_t biliniyor.

a_j ve d_t ' ler spline fonksiyon parametreleridir. Görüldüğü üzere spline fonksiyonu bu parametrelerde doğrusaldır. Vektörler ve matrisler tanımlanacak olursa,

$$a = \begin{bmatrix} a_0 \\ a_0 \\ \vdots \\ \vdots \\ a_n \end{bmatrix} \quad ((n+1) \times 1) \text{ boyutlu vektör} \quad d = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ \vdots \\ d_{k-1} \end{bmatrix} \quad ((k-1) \times 1) \text{ boyutlu vektör}$$

$$E = \begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^n \\ 1 & x_2 & x_2^2 & \cdots & x_2^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_m & x_m^2 & \cdots & x_m^n \end{bmatrix} \quad ((m \times n) + 1) \text{ boyutlu matris}$$

$$E_+ = \begin{bmatrix} (x_1 - \xi_1)_+^n & (x_1 - \xi_2)_+^n & \cdots & (x_1 - \xi_{k-1})_+^n \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ (x_m - \xi_1)_+^n & (x_m - \xi_2)_+^n & \cdots & (x_m - \xi_{k-1})_+^n \end{bmatrix} \quad (m \times (k-1)) \text{ boyutlu matris}$$

$$u = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ \vdots \\ u_m \end{bmatrix} \quad (m \times 1) \text{ boyutlu vektör} \quad \eta = \begin{bmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \vdots \\ \eta_m \end{bmatrix} \quad (m \times 1) \text{ boyutlu vektör}$$

E ve E_+ matrisleri ve η vektörü m gözlemden oluşturulur. Regresyon modeli,

$$\eta_i = S_{\Delta}^n(x_i) + u_i \quad (3.23)$$

biçiminde kurulur. Eşitlik 3.23, matris notasyonuyla,

$$\eta = \Xi a + \Xi_+ d + u \quad (3.24)$$

biçiminde gösterilir. Eşitlik 3.24'te $b = \begin{bmatrix} a \\ d \end{bmatrix}$ ve $X = (\Xi, \Xi_+)$ olarak tanımlanırsa regresyon modeli,

$$\eta = Xb + u \quad (3.25)$$

haline getirilir. Sonuç olarak 3.25 nolu eşitlik doğrusal regresyon modelinden başka bir şey değildir. Model parametresi en küçük kareler tahmin edicisi,

$$\hat{b} = (X'X)^{-1}X'\eta \quad (3.26)$$

ile tahmin edilir.

Klasik doğrusal normal regresyon modelinin genel varsayımları bu teoremin bütün durumlarında uygulanabilir. $\hat{b}$ en iyi doğrusal yansız tahmin edicidir ve normal dağılımlıdır.

$$E(\hat{b}) = b \quad (3.27)$$

$$Var(\hat{b}) = \hat{\sigma}^2(X'X)^{-1} \quad (3.28)$$

olmaktadır.

3.4.1. Doğrusal Spline Regresyon

Doğrusal spline regresyon modeli eğim katsayılarının doğrudan yorumlanmasına olanak sağlar. Ancak uyum iyiliği bazı verilerde veya belli bir aralık üzerinde zayıf olabilir. Diğer bir dezavantajı da düğüm noktalarında türevlerinin süreksiz olmasıdır²⁷⁸.

Doğrusal spline, sürekli parçalı doğrusal bir fonksiyondur. Yani düz doğru parçalarından oluşan sürekli bir fonksiyondur²⁷⁹. Doğrusal splinenin temel mantığı herhangi bir sürekli fonksiyonun parçalı doğrusal fonksiyonlarla isteğe bağlı olarak iyi bir şekilde tahmin edilebilmesidir²⁸⁰. Parçalı doğrusal spline regresyon modeli, bağımsız

²⁷⁸ Sung, s.14.

²⁷⁹ Horowitz, Loughran, and Savin, s.2.

²⁸⁰ Horowitz, Loughran, and Savin, s.8.

değişken eksenini üzerinde farklı aralıklar için farklı doğrusal ilişkileri öneren bir modeldir²⁸¹.

Parçalı doğrusal modeller aslında, spline fonksiyonların özel bir durumudur. Parçalı doğrusal spline regresyonda, tahmin edilecek regresyon doğrusunun sürekli olması gerektiği kısıtlamasıyla eğimdeki değişmelerin belirlenmesine olanak sağlanır²⁸². Spline regresyon doğrusal olduğunda, düğüm noktalarında fonksiyonların eşit olması sağlanır. Fakat fonksiyonların eğimleri farklı olabilir²⁸³.

Eşitlik 3.29’da kısıtlamasız doğrusal spline regresyon modelinin genel formu gösterilmiştir.

$$Y(x) = \sum_{i=0}^1 a_i x_t^i + \sum_{i=1}^k \sum_{j=0}^1 b_{ij} (x_t - \xi_i)_+^j D_{ij} + e_t \quad (3.29)$$

Eşitlik 3.29’daki genel forma göre, bir düğüm noktalı yani iki parçalı süreksiz doğrusal spline regresyon modeli²⁸⁴,

$$Y(x) = a_1 + a_2 x_t + b_{10} (x_t - \xi_1)_+^0 D_{10} + b_{11} (x_t - \xi_1)_+^1 D_{11} + e_t$$

$$Y(x) = a_1 + a_2 x_t + b_{10} D_{10} + b_{11} (x_t - \xi_1) D_{11} + e_t$$

biçiminde yazılır. Bu eşitlik,

$$Y(X) = \begin{cases} a_1 + a_2 x_t, & x \leq \xi \\ a_1 + a_2 x_t + b_{10} D_{10} + b_{11} (x_t - \xi_1) D_{11}, & x > \xi \end{cases} \quad (3.30)$$

şeklinde parçalı olarak eşitlik 3.30’daki gibi yazılabilir. Burada $x \leq \xi$ ise $D = 0$ ve $x > \xi$ ise $D = 1$ değerini alır.

Eşitlik 3.29’da belirtilen model üzerinde bazı süreklilik kısıtlamaları uygulanarak katsayıları azaltılmış bir model geliştirilebilir. Dolayısıyla b_{10} katsayısı çıkarılarak eşitlik 3.31’deki gibi model sürekli hale getirilir²⁸⁵.

²⁸¹ William Mendenhall and Terry Sincich, *A Second Course in Statistics: Regression Analysis*, Fifth Edition, Prentice Hall, New Jersey 1996.

²⁸² Pindyck and Rubinfeld, s. 117.

²⁸³ Marsh ve Cormier, s.4.

²⁸⁴ Mendenhall and Sincich, s.455.

²⁸⁵ Marsh, “Estimating the Number of and Location of Knots in Spline Regression”, s.62.

$$Y(x) = \sum_{i=0}^1 a_i x_t^i + \sum_{i=1}^k b_{i1} (x_t - \xi_i)_+ D_{i1} + e_t \quad (3.31)$$

Eşitlik 3.31'deki genel forma göre, bir düğüm noktalı yani iki parçalı sürekli doğrusal spline regresyon modeli,

$$Y(x) = a_1 + a_2 x_t + b_{11} (x_t - \xi_1)_+ D_{11} + e_t$$

biçiminde yazılır. Bu eşitlik,

$$Y(X) = \begin{cases} a_1 + a_2 x_t, & x \leq \xi \\ a_1 + a_2 x_t + b_{11} (x_t - \xi_1) D_{11}, & x > \xi \end{cases} \quad (3.32)$$

şeklinde parçalı olarak eşitlik 3.32'deki gibi yazılabilir. Burada $x \leq \xi$ ise $D = 0$ ve $x > \xi$ ise $D = 1$ değerini alır²⁸⁶.

3.4.2. Kareli (Kuadratik) Spline Regresyon

Kareli spline regresyonun, doğrusal spline regresyona göre avantajı, dağılım grafiğinde dip ve tepe noktaların tahmininde daha iyi olmasıdır²⁸⁷. Spline regresyon kareli olarak alındığında, düğüm noktalarında hem fonksiyonlar hem de fonksiyonların eğimleri eşittir. Fakat eğimlerin değişim oranı farklıdır²⁸⁸. Eşitlik 3.33'te kısıtlamasız kareli spline regresyon modelinin genel formu gösterilmiştir.

$$P(x) = \sum_{i=0}^2 a_i x_t^i + \sum_{i=1}^k \sum_{j=0}^2 b_{ij} (x_t - \xi_i)_+^j D_{ij} + e_t \quad (3.33)$$

Eşitlik 3.33'deki genel forma göre, bir düğüm noktalı yani iki parçalı süreksiz kareli spline regresyon modeli,

$$Y(x) = a_1 + a_2 x_t + a_3 x_t^2 + b_{10} (x_t - \xi_1)_+^0 D_{10} + b_{11} (x_t - \xi_1)_+^1 D_{11} + b_{12} (x_t - \xi_1)_+^2 D_{12} + e_t$$

$$Y(x) = a_1 + a_2 x_t + a_3 x_t^2 + b_{10} D_{10} + b_{11} (x_t - \xi_1) D_{11} + b_{12} (x_t - \xi_1)_+^2 D_{12} + e_t$$

²⁸⁶ Rudolf J. Freund, William J. Wilson and Ping Sa, *Regression Analysis: Statistical Modeling of a Response Variable*, Second Edition, Academic Press, San Diego 2006, s.279.

²⁸⁷ David Ruppert, Matt P. Wand and Raymond J. Carroll, *Semi Parametric Regression*, Cambridge University Press, New York 2003, s.68.

²⁸⁸ Marsh ve Cormier, s.4.

biçiminde yazılır. Bu eşitlik,

$$Y(X) = \begin{cases} a_1 + a_2x_t + a_3x_t^2, & x \leq \xi \\ a_1 + a_2x_t + a_3x_t^2 + b_{10}D_{10} + b_{11}(x_t - \xi_1)D_{11} + b_{12}(x_t - \xi_1)_+^2D_{12}, & x > \xi \end{cases}$$

şeklinde parçalı olarak yazılabilir. Burada $x \leq \xi$ ise $D = 0$ ve $x > \xi$ ise $D = 1$ değerini alır.

Eşitlik 3.33'te belirtilen model üzerinde bazı süreklilik kısıtlamaları uygulanarak katsayıları azaltılmış bir model geliştirilebilir. Dolayısıyla b_{10} katsayısı çıkarılarak eşitlik 3.34'deki gibi model sürekli hale getirilir.

$$P(x) = \sum_{i=0}^2 a_i x^i + \sum_{i=1}^k \sum_{j=1}^2 b_{ij} (x - \xi_i)_+^j \quad (3.34)$$

Eşitlik 3.34'deki genel forma göre, bir düğüm noktalı yani iki parçalı sabiti kısıtlandırılmış sürekli kareli spline regresyon modeli,

$$Y(x) = a_1 + a_2x_t + a_3x_t^2 + b_{11}(x_t - \xi_1)D_{11} + b_{12}(x_t - \xi_1)_+^2D_{12} + e_t$$

biçiminde yazılır. Bu eşitlik,

$$Y(X) = \begin{cases} a_1 + a_2x_t + a_3x_t^2 & x \leq \xi \\ a_1 + a_2x_t + a_3x_t^2 + b_{11}(x_t - \xi_1)D_{11} + b_{12}(x_t - \xi_1)_+^2D_{12} & x > \xi \end{cases}$$

şeklinde parçalı olarak yazılabilir. Burada $x \leq \xi$ ise $D = 0$ ve $x > \xi$ ise $D = 1$ değerini alır²⁸⁹. Eşitlik 3.34'te belirtilen model sürekli hale getirilmiştir ancak birinci dereceden türevlerin eşit olmamasından dolayı düğüm noktalarında keskin köşeler oluşabilir. Bu keskin köşeler b_{11} katsayısının sıfır yapılmasıyla düzgünleştirilebilir. Yani her bir düğüm noktasında birinci türevlerin eşit olması koşuluyla düğüm noktalarında keskinlik azaltılır. Sonuç olarak eşitlik 3.35'deki gibi bir model elde edilir²⁹⁰.

$$P(x) = \sum_{i=0}^2 a_i x^i + \sum_{i=1}^k b_{i2} (x - \xi_i)_+^2 \quad (3.35)$$

²⁸⁹ Freund, Wilson and Sa, s.280.

²⁹⁰ Marsh, "Estimating the Number of and Location of Knots in Spline Regression", s.62.

Eşitlik 3.35'deki genel forma göre, bir düğüm noktalı yani iki parçalı sabiti kısıtlandırılmış ve düğüm noktasında birinci derece türevli sürekli kareli spline regresyon modeli,

$$Y(x) = a_1 + a_2x_t + a_3x_t^2 + b_{12}(x_t - \xi_1)_+^2 D_{12} + e_t$$

biçiminde yazılır. Bu eşitlik,

$$Y(X) = \begin{cases} a_1 + a_2x_t + a_3x_t^2, & x \leq \xi \\ a_1 + a_2x_t + a_3x_t^2 + b_{12}(x_t - \xi_1)_+^2 D_{12}, & x > \xi \end{cases}$$

şeklinde parçalı olarak yazılabilir. Burada $x \leq \xi$ ise $D = 0$ ve $x > \xi$ ise $D = 1$ değerini alır.

3.4.2.1. Kareli Spline Regresyonla Mevsimselliğin Hesaplanması

Robb(1980), kareli spline regresyon modelle mevsimselliğin hesaplanmasına ilişkin bir model geliştirmiştir. Buna göre, regresyon modelinin eşitlik 3.36'daki gibi olduğu varsayılın²⁹¹.

$$Y_t = A(s) + B(s)X_t + e_t \quad (3.36)$$

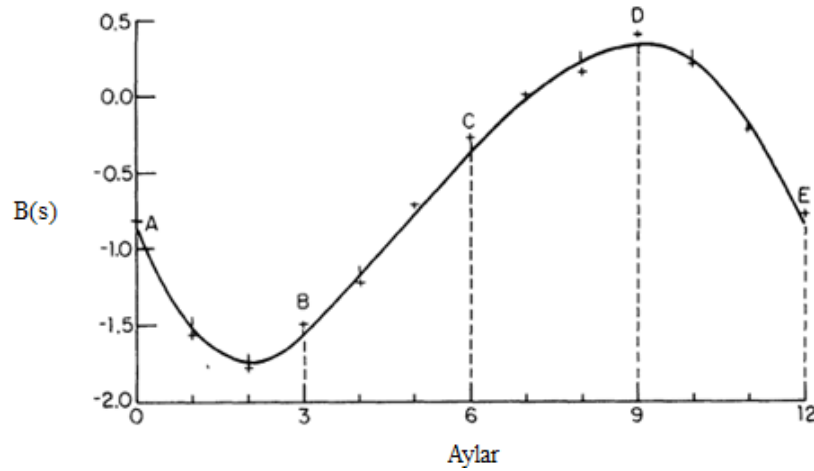
$$t = 1, 2, \dots, 12 * k$$

$$s = 1, 2, \dots, 12$$

Burada s indeksi ayları ve t indeksi ise k adet yıldaki aylara ait gözlemleri göstermektedir. Y bağımlı değişkenin ve X bağımsız değişkenin gözlem değerlerini göstermektedir. A ve B tahmin edilecek parametreler olup, yılın ayına göre değişmesine olanak sağlar. e ise regresyon modelinin hata terimidir.

B parametresi mevsimlerle birlikte değişmekte ve A parametresi ise sabit tutulmaktadır. B(s)'nin Şekil 3.1'de verilen durum uygun olduğu varsayılın. Şekil 3.1'de (+) işareti ile gösterilen noktalar 12 ayrı katsayının tahminleri olarak düşünülebilir. Şekil 3.1'de gösterilen yılın dört ayrı parçasının her birine polinom tahmin edilmesiyle birlikte B, C, D ve E(=A) düğüm noktalarında sürekliliğin ve düzgünlüğün sağlanması gerekmektedir. Bu amaçla ikinci dereceden kareli polinom uygulanmıştır. Böylece B(s) fonksiyonu eşitlik 3.37'deki gibi yazılır.

²⁹¹ Robb, s.321.



Şekil 3.1. Bir Regresyon Katsayısının Mevsimsel Model Eğrisi

$$B(s) = a_0 + a_1s + a_2s^2 + [b_0 + b_1(s - 3) + b_2(s - 3)^2]D_3 + [c_0 + c_1(s - 6) + c_2(s - 6)^2]D_6 + [d_0 + d_1(s - 9) + d_2(s - 9)^2]D_9 \quad (3.37)$$

Eşitlik 3.37’de D_3 , D_6 ve D_9 gölge değişkenlerdir. Buna göre, $s \geq 3$ ise $D_3 = 1$, diğer durumlarda $D_3 = 0$, $s \geq 6$ ise $D_6 = 1$, diğer durumlarda $D_6 = 0$ ve $s \geq 9$ ise $D_9 = 1$, diğer durumlarda $D_9 = 0$ olur. $B(s)$ ’nin B,C,D ve E düğüm noktalarında sürekli olması gerektiğinden ilave kısıtlamalar gerekmektedir. Öncelikle eşitlik 3.37’de²⁹²,

$$b_0 = 0 \quad (3.38)$$

$$c_0 = 0 \quad (3.39)$$

$$d_0 = 0 \quad (3.40)$$

$$a_0 = a_0 + 12a_1 + 144a_2 + 9b_1 + 81b_2 + 6c_1 + 36c_2 + 3d_1 + 9d_2 \quad (3.41)$$

Kısıtlamaları belirlenir. Çünkü b_0 , c_0 ve d_0 sabitlerinin sıfıra eşitlenmesiyle regresyon eğrisinin B, C ve D düğüm noktalarında sürekli olması sağlanır. E noktası yılın son ayı aralık ($s = 12$) ve yeni yılın başlangıcı ($s = 0$) arasındaki nokta olduğu için, sürekliliğin sağlanması amacıyla $s = 12$ alınarak eşitlik 3.37’de yerine koyulursa eşitlik 3.41’deki kısıtlama elde edilir. Böylece tahmin edilecek parametre sayısı 12’den 8’e düşer.

²⁹² Robb, s.322.

Daha düzgün sürekli bir regresyon eğrisi elde etmek için $B(s)$ fonksiyonunun düğüm noktalarında parçalı fonksiyonların birinci dereceden türevlerinin eşit olması sağlanır. Bunun için de,

$$b_1 = 0 \quad (3.42)$$

$$c_1 = 0 \quad (3.43)$$

$$d_1 = 0 \quad (3.44)$$

$$a_1 = a_1 + 24a_2 + 18b_2 + 12c_2 + 6d_2 \quad (3.45)$$

ilave kısıtlamaları gerekmektedir. Burada da eşitlik 3.37'nin s 'ye göre birinci dereceden türevi alınarak, $s = 12$ değeri yerine koyulursa eşitlik 3.45 kısıtlaması elde edilir. Böylece tahmin edilecek parametre sayısı 8'den 4'e düşer.

Eşitlik 3.41 ve 3.45'deki kısıtlamalarda gerekli düzenlemeler yapılırsa,

$$0 = 12a_1 + 144a_2 + 9b_1 + 81b_2 + 6c_1 + 36c_2 + 3d_1 + 9d_2 \quad (3.46)$$

$$0 = 24a_2 + 18b_2 + 12c_2 + 6d_2 \quad (3.47)$$

eşitlikleri elde edilir. Eşitlik 3.42-3.44'deki kısıtlamalar gereği, eşitlik 3.46'da b_1 , c_1 ve d_1 sıfır alınıp, a_2 yerine de $(-18b_2 - 12c_2 - 6d_2)/24$ ifadesi yazılırsa²⁹³,

$$0 = 12a_1 + 144 \left(\frac{-18b_2 - 12c_2 - 6d_2}{24} \right) + 81b_2 + 36c_2 + 9d_2 \quad (3.48)$$

$$0 = 12a_1 - 27b_2 - 36c_2 - 27d_2 \quad (3.49)$$

$$a_1 = 2,25b_2 + 3c_2 + 2,25d_2 \quad (3.50)$$

eşitliği elde edilir. a_2 ifadesi ise,

$$a_2 = -\frac{18}{24}b_2 - \frac{12}{24}c_2 - \frac{6}{24}d_2 \quad (3.51)$$

$$a_2 = -\frac{18}{24}b_2 - \frac{12}{24}c_2 - \frac{6}{24}d_2 \quad (3.52)$$

biçiminde elde edilir.

Tüm kısıtlamaların eşitlik 3.37'de yerine yazılmasıyla,

$$B(s) = a_0 + (2,25b_2 + 3c_2 + 2,25d_2)s + (-0,75b_2 - 0,5c_2 - 0,25d_2)s^2 + b_2(s-3)^2D_3 + c_2(s-6)^2D_6 + d_2(s-9)^2D_9 \quad (3.53)$$

²⁹³ Robb, s.322.

eşitliği elde edilir. Eşitlik 18, b_2 , c_2 ve d_2 ortak parantezinde yeniden düzenlenirse,

$$B(s) = a_0 + b_2[2,25s - 0,75s^2 + (s - 3)^2 D_3] + c_2[3s - 0,5s^2 + (s - 6)^2 D_6] + d_2[2,25s - 0,25s^2 + (s - 9)^2 D_9] \quad (3.54)$$

eşitliği elde edilir. Eşitlik 3.54'teki köşeli parantez içi terimler sırasıyla S_3 , S_6 ve S_9 olarak kısaltılır ve Eşitlik 3.36'da yerine koyulursa,

$$Y_t = A + a_0X_t + b_2(S_3X_t) + c_2(S_6X_t) + d_2(S_9X_t) + e_t \quad (3.55)$$

biçiminde olur. Eşitlik 3.55'teki bilinmeyen beş parametre en küçük kareler yöntemiyle tahmin edilebilir. Alternatif olarak, mevsimsel örüntü eğimden ziyade sabitle ilişkili olsaydı, tahmin edilecek denklem aşağıdaki gibi olacaktı.

$$Y_t = A_0 + B_2(S_3) + C_2(S_6) + D_2(S_9) + BX_t + e_t \quad (3.56)$$

3.4.3. Kübik Spline Regresyon

Kübik spline regresyonun kareli regresyona göre avantajı, dağılım grafiğinde dip ve tepe noktaların tahmininde kareli spline regresyondan daha esnektir²⁹⁴. Kübik spline regresyon, doğrusal olmayan veri setlerine daha iyi tahminler yapmak için uygun tekniklerden biridir²⁹⁵. Yüksek mertebeli splinelar, veriye daha esnek ve çok düzgün bir şekilde regresyon eğrisi uydurmak için kullanılır²⁹⁶. Bunlardan biri olan kübik splinelar, düğüm noktalarında ikinci türevleri sürekli olan parçalı kübik polinomlardır²⁹⁷. Kübik splinelar yaygın bir şekilde kullanılmaktadır ve kübik splinelardan başka daha yüksek mertebeli splineların kullanımı pek tercih edilmemektedir²⁹⁸.

Spline regresyon kübik olarak alındığında, fonksiyonların eğimleri ve eğimlerin değişim oranlarıyla birlikte fonksiyonlar kısıtlandırılır. Verilere uygun bir tahmin yapılırken, eğrinin çok düzgün ve esnek olması istenildiği durumlarda kübik splinelar daha yaygın bir şekilde kullanılmaktadır²⁹⁹. Kübik splineları kullanan parametrik

²⁹⁴ Keele, s.55.

²⁹⁵ Sung, s.14.

²⁹⁶ Kommineni, Basu and Vemuri, s.242.

²⁹⁷ Chib and Greenberg, s.324.

²⁹⁸ Phillips and Taylor, s.3.

²⁹⁹ Marsh ve Cormier, s.4.

olmayan regresyon yöntemi, eğri tahmini için etkin, esnek ve yaygın bir şekilde uygulanabilirliği olan bir yaklaşımdır³⁰⁰.

Kübik spline regresyon, temelde kübik spline interpolasyon yöntemiyle aynıdır. Kübik spline regresyon ve kübik spline interpolasyon yöntemi arasındaki farklılık, düğüm noktası sayısının kübik spline regresyonda az olmasıdır³⁰¹.

Eşitlik 3.57’de kısıtlamasız kübik spline regresyon modelinin genel formu gösterilmiştir.

$$P(x) = \sum_{i=0}^3 a_i x^i + \sum_{i=1}^k \sum_{j=0}^3 b_{ij} (x - \xi_i)_+^j \quad (3.57)$$

Eşitlik 3.57’deki genel forma göre, bir düğüm noktalı yani iki parçalı süreksiz kübik spline regresyon modeli,

$$Y(x) = a_1 + a_2 x_t + a_3 x_t^2 + a_4 x_t^3 + b_{10} (x_t - \xi_1)_+^0 D_{10} + b_{11} (x_t - \xi_1)_+^1 D_{11} + b_{12} (x_t - \xi_1)_+^2 D_{12} + b_{13} (x_t - \xi_1)_+^3 D_{13} + e_t$$

$$Y(x) = a_1 + a_2 x_t + a_3 x_t^2 + a_4 x_t^3 + b_{10} D_{10} + b_{11} (x_t - \xi_1) D_{11} + b_{12} (x_t - \xi_1)^2 D_{12} + b_{13} (x_t - \xi_1)^3 D_{13} + e_t$$

biçiminde yazılır. Bu eşitlik,

$$Y(x) = \begin{cases} a_1 + a_2 x_t + a_3 x_t^2 + a_4 x_t^3, & x \leq \xi \\ a_1 + a_2 x_t + a_3 x_t^2 + a_4 x_t^3 + b_{10} D_{10} + b_{11} (x_t - \xi_1) D_{11} + b_{12} (x_t - \xi_1)^2 D_{12} + b_{13} (x_t - \xi_1)^3 D_{13}, & x > \xi \end{cases}$$

şeklinde parçalı olarak yazılabilir. Burada $x \leq \xi$ ise $D = 0$ ve $x > \xi$ ise $D = 1$ değerini alır.

Eşitlik 3.57’de belirtilen model üzerinde bazı süreklilik kısıtlamaları uygulanarak katsayıları azaltılmış bir model geliştirilebilir. Dolayısıyla b_{10} katsayısı çıkarılarak eşitlik 3.58’deki gibi model sürekli hale getirilir.

$$P(x) = \sum_{i=0}^3 a_i x^i + \sum_{i=1}^k \sum_{j=1}^3 b_{ij} (x - \xi_i)_+^j \quad (3.58)$$

³⁰⁰ Silverman, “Some Aspects of the Spline Smoothing Approach to Non-Parametric Regression Curve Fitting”, s.1.

³⁰¹ Guo and White, s.344.

Eşitlik 3.58'deki genel forma göre, bir düğüm noktalı yani iki parçalı sabiti kısıtlandırılmış sürekli kübik spline regresyon modeli,

$$Y(x) = a_1 + a_2x_t + a_3x_t^2 + a_4x_t^3 + b_{11}(x_t - \xi_1)D_{11} + b_{12}(x_t - \xi_1)_+^2D_{12} + b_{13}(x_t - \xi_1)_+^3D_{13} + e_t$$

biçiminde yazılır. Bu eşitlik,

$$Y(X) = \begin{cases} a_1 + a_2x_t + a_3x_t^2 + a_4x_t^3, & x \leq \xi \\ a_1 + a_2x_t + a_3x_t^2 + a_4x_t^3 + b_{11}(x_t - \xi_1)D_{11} + b_{12}(x_t - \xi_1)_+^2D_{12} + b_{13}(x_t - \xi_1)_+^3D_{13}, & x > \xi \end{cases}$$

şeklinde parçalı olarak yazılabilir. Burada $x \leq \xi$ ise $D = 0$ ve $x > \xi$ ise $D = 1$ değerini alır.

Eşitlik 3.58'de belirtilen model sürekli hale getirilmiştir ancak birinci dereceden türevlerin eşit olmamasından dolayı düğüm noktalarında keskin köşeler oluşabilir. Bu keskin köşeler b_{11} katsayısının sıfır yapılmasıyla düzgünleştirilebilir. Yani her bir düğüm noktasında birinci türevlerin eşit olması koşuluyla düğüm noktalarında keskinlik azaltılır. Sonuç olarak eşitlik 3.59'daki gibi bir model elde edilir.

$$P(x) = \sum_{i=0}^3 a_i x^i + \sum_{i=1}^k \sum_{j=2}^3 b_{ij} (x - \xi_i)_+^j \quad (3.59)$$

Eşitlik 3.59'daki genel forma göre, bir düğüm noktalı yani iki parçalı sabiti kısıtlandırılmış ve düğüm noktasında birinci derece türevli sürekli kübik spline regresyon modeli,

$$Y(x) = a_1 + a_2x_t + a_3x_t^2 + a_4x_t^3 + b_{12}(x_t - \xi_1)_+^2D_{12} + b_{13}(x_t - \xi_1)_+^3D_{13} + e_t$$

biçiminde yazılır. Bu eşitlik,

$$Y(X) = \begin{cases} a_1 + a_2x_t + a_3x_t^2 + a_4x_t^3, & x \leq \xi \\ a_1 + a_2x_t + a_3x_t^2 + a_4x_t^3 + b_{12}(x_t - \xi_1)_+^2D_{12} + b_{13}(x_t - \xi_1)_+^3D_{13}, & x > \xi \end{cases}$$

şeklinde parçalı olarak yazılabilir. Burada $x \leq \xi$ ise $D = 0$ ve $x > \xi$ ise $D = 1$ değerini alır.

Eşitlik 3.59'daki modelde düğüm noktalarında ikinci dereceden türevler eşit değildir. İkinci dereceden türevlerin eşitlenmesiyle regresyon doğrusundaki pürüzlülük daha düzgün hale getirilir. Dolayısıyla b_{12} katsayısının sıfır yapılmasıyla daha düzgün

regresyon doğrusu elde edilir. Sonuç olarak eşitlik 3.60'daki gibi en düzgün kübik spline regresyon modeli bulunmuş olur.

$$P(x) = \sum_{i=0}^3 a_i x^i + \sum_{i=1}^k b_{i3} (x - \xi_i)_+^3 \quad (3.60)$$

Eşitlik 3.60'daki genel forma göre, bir düğüm noktalı yani iki parçalı sabiti kısıtlandırılmış ve düğüm noktasında ikinci derece türevli sürekli kübik spline regresyon modeli,

$$Y(x) = a_1 + a_2 x_t + a_3 x_t^2 + a_4 x_t^3 + b_{13} (x_t - \xi_1)_+^3 D_{13} + e_t$$

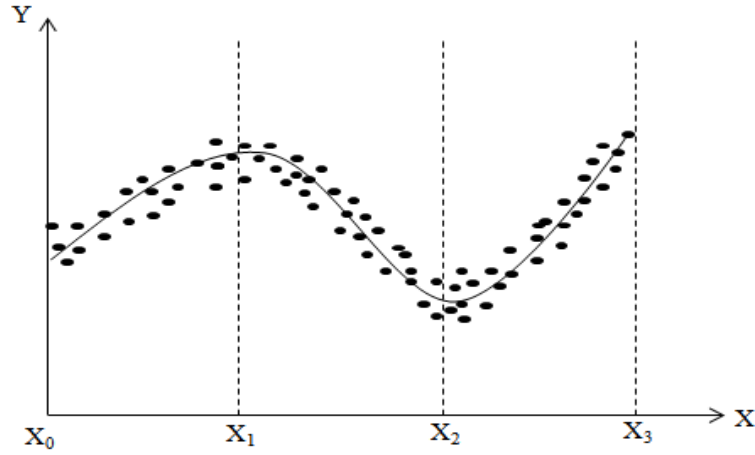
biçiminde yazılır. Bu eşitlik,

$$Y(X) = \begin{cases} a_1 + a_2 x_t + a_3 x_t^2 + a_4, & x \leq \xi \\ a_1 + a_2 x_t + a_3 x_t^2 + a_4 x_t^3 + b_{13} (x_t - \xi_1)_+^3 D_{13}, & x > \xi \end{cases}$$

şeklinde parçalı olarak yazılabilir. Burada $x \leq \xi$ ise $D = 0$ ve $x > \xi$ ise $D = 1$ değerini alır.

3.4.3.1. Suits ve Ark. Tarafından Geliştirilen Kübik Spline Regresyon

Doğrusal ve kareli spline regresyona göre daha esnek olan kübik spline regresyon birçok uygulama için yeterlidir. Kübik spline regresyon yapısı Şekil 3.2'de gösterilmiştir. X eksenini üç parçaya ayırmıştır.



Şekil 3.2. Parçalı Polinom Regresyon

Üç aralıkta tanımlanan kübik spline regresyon modeli aşağıdaki gibi oluşturulur. Suits ve ark. tarafından önerilen kübik spline regresyon modelinde düğüm noktaları arasındaki uzaklıkların eşit olduğu varsayılmaktadır³⁰².

$$Y = [a_1 + b_1(X - X_0) + c_1(X - X_0)^2 + d_1(X - X_0)^3]D_1 + [a_2 + b_2(X - X_1) + c_2(X - X_1)^2 + d_2(X - X_1)^3]D_2 + [a_3 + b_3(X - X_2) + c_3(X - X_2)^2 + d_3(X - X_2)^3]D_3 + u \quad (3.61)$$

Burada, X_0 , X_1 , X_2 ve X_3 düğüm noktaları, D_i , gölge değişken, a_i , b_i , c_i ve d_i tahmin edilecek parametre katsayıları ve u ise tüm regresyon varsayımlarının sağlandığı hata terimidir.

Eşitlik 3.61, kısıtlanmamış (kesikli) kübik spline regresyon modelidir. Arzu edilen kübik spline regresyon modelini elde etmek için ilave kısıtlar gerekmektedir. Buna göre, Eşitlik 3.62 ve 3.65'te, a_i üzerindeki kısıtlamayla düğüm noktalarında sağ ve sol aralıktaki fonksiyon değerlerinin eşit ve fonksiyonun sürekli olması sağlanır. Eşitlik 3.63 ve 3.66'da, b_i üzerindeki kısıtlamayla düğüm noktalarında sağ ve sol aralıktaki fonksiyonların birinci türev değerlerinin eşit olması sağlanır. Aynı şekilde, eşitlik 3.64 ve 3.67'de, c_i üzerindeki kısıtlamayla düğüm noktalarında sağ ve sol aralıktaki fonksiyonların ikinci türev değerlerinin de eşit olması sağlanır.

$$a_2 = a_1 + b_1(X_1 - X_0) + c_1(X_1 - X_0)^2 + d_1(X_1 - X_0)^3 \quad (3.62)$$

$$b_2 = b_1 + 2c_1(X_1 - X_0) + 3d_1(X_1 - X_0)^2 \quad (3.63)$$

$$c_2 = c_1 + 3d_1(X_1 - X_0) \quad (3.64)$$

$$a_3 = a_2 + b_2(X_2 - X_1) + c_2(X_2 - X_1)^2 + d_2(X_2 - X_1)^3 \quad (3.65)$$

$$b_3 = b_2 + 2c_2(X_2 - X_1) + 3d_2(X_2 - X_1)^2 \quad (3.66)$$

$$c_3 = c_2 + 3d_2(X_2 - X_1) \quad (3.67)$$

Eşitlik 3.65-3.67'deki kısıtlarının a_1 , b_1 , c_1 ve d_1 cinsinden yeniden düzenlenmesi sonucu, eşitlik 3.68-3.70 elde edilir.

³⁰² Suits, Mason ve Chan, s.134.

$$a_3 = a_1 + b_1(X_1 - X_0) + c_1(X_1 - X_0)^2 + d_1(X_1 - X_0)^3 + [b_1 + 2c_1(X_1 - X_0) + 3d_1(X_1 - X_0)^2](X_2 - X_1) + [c_1 + 3d_1(X_1 - X_0)](X_2 - X_1)^2 + d_2(X_2 - X_1)^3$$

$$a_3 = a_1 + b_1(X_1 - X_0) + c_1(X_1 - X_0)^2 + d_1(X_1 - X_0)^3 + b_1(X_2 - X_1) + 2c_1(X_1 - X_0)(X_2 - X_1) + 3d_1(X_1 - X_0)^2(X_2 - X_1) + c_1(X_2 - X_1)^2 + 3d_1(X_1 - X_0)(X_2 - X_1)^2 + d_2(X_2 - X_1)^3 \quad (3.68)$$

$$b_3 = b_1 + 2c_1(X_1 - X_0) + 3d_1(X_1 - X_0)^2 + 2[c_1 + 3d_1(X_1 - X_0)](X_2 - X_1) + 3d_2(X_2 - X_1)^2$$

$$b_3 = b_1 + 2c_1(X_1 - X_0) + 3d_1(X_1 - X_0)^2 + 2c_1(X_2 - X_1) + 6d_1(X_1 - X_0)(X_2 - X_1) + 3d_2(X_2 - X_1)^2 \quad (3.69)$$

$$c_3 = c_1 + 3d_1(X_1 - X_0) + 3d_2(X_2 - X_1) \quad (3.70)$$

Eşitlik 3.62-3.67'deki kısıtların eşitlik 3.61'de yerine yazılmasıyla eşitlik 3.71 elde edilir.

$$Y = [a_1 + b_1(X - X_0) + c_1(X - X_0)^2 + d_1(X - X_0)^3]D_1 + \{[a_1 + b_1(X_1 - X_0) + c_1(X_1 - X_0)^2 + d_1(X_1 - X_0)^3] + [b_1 + 2c_1(X_1 - X_0) + 3d_1(X_1 - X_0)^2](X - X_1) + [c_1 + 3d_1(X_1 - X_0)](X - X_1)^2 + d_2(X - X_1)^3\}D_2 + \{[a_1 + b_1(X_1 - X_0) + c_1(X_1 - X_0)^2 + d_1(X_1 - X_0)^3 + b_1(X_2 - X_1) + 2c_1(X_1 - X_0)(X_2 - X_1) + 3d_1(X_1 - X_0)^2(X_2 - X_1) + c_1(X_2 - X_1)^2 + 3d_1(X_1 - X_0)(X_2 - X_1)^2 + d_2(X_2 - X_1)^3] + [b_1 + 2c_1(X_1 - X_0) + 3d_1(X_1 - X_0)^2 + 2c_1(X_2 - X_1) + 6d_1(X_1 - X_0)(X_2 - X_1) + 3d_2(X_2 - X_1)^2](X - X_2) + [c_1 + 3d_1(X_1 - X_0) + 3d_2(X_2 - X_1)](X - X_2)^2 + d_3(X - X_2)^3\}D_3 + u \quad (3.71)$$

Eşitlik 3.71'de D_1 , D_2 ve D_3 köşeli parantezlerinin açılarak yeniden düzenlenmesiyle eşitlik 3.72 elde edilir.

$$Y = a_1D_1 + b_1(X - X_0)D_1 + c_1(X - X_0)^2D_1 + d_1(X - X_0)^3D_1 + a_1D_2 + b_1(X_1 - X_0)D_2 + c_1(X_1 - X_0)^2D_2 + d_1(X_1 - X_0)^3D_2 + b_1(X - X_1)D_2 + 2c_1(X_1 - X_0)(X - X_1)D_2 + 3d_1(X_1 - X_0)^2(X - X_1)D_2 + c_1(X - X_1)^2D_2 + 3d_1(X_1 - X_0)(X - X_1)^2D_2 + d_2(X - X_1)^3D_2 + a_1D_3 + b_1(X_1 - X_0)D_3 + c_1(X_1 - X_0)^2D_3 + d_1(X_1 - X_0)^3D_3 + b_1(X_2 - X_1)D_3 + 2c_1(X_1 - X_0)(X_2 -$$

$$\begin{aligned}
& X_1)D_3 + 3d_1(X_1 - X_0)^2(X_2 - X_1)D_3 + c_1(X_2 - X_1)^2D_3 + 3d_1(X_1 - \\
& X_0)(X_2 - X_1)^2D_3 + d_2(X_2 - X_1)^3D_3 + b_1(X - X_2)D_3 + 2c_1(X_1 - X_0)(X - \\
& X_2)D_3 + 3d_1(X_1 - X_0)^2(X - X_2)D_3 + 2c_1(X_2 - X_1)(X - X_2)D_3 + 6d_1(X_1 - \\
& X_0)(X_2 - X_1)(X - X_2)D_3 + 3d_2(X_2 - X_1)^2(X - X_2)D_3 + c_1(X - X_2)^2D_3 + \\
& 3d_1(X_1 - X_0)(X - X_2)^2D_3 + 3d_2(X_2 - X_1)(X - X_2)^2D_3 + d_3(X - X_2)^3D_3 + \\
& u \tag{3.72}
\end{aligned}$$

Eşitlik 3.72’de b terimlerinin bir araya getirilerek düzenlemesiyle eşitlik 3.73 elde edilir.

$$\begin{aligned}
& b_1(X - X_0)D_1 + b_1(X_1 - X_0)D_2 + b_1(X - X_1)D_2 + b_1(X_1 - X_0)D_3 + \\
& b_1(X_2 - X_1)D_3 + b_1(X - X_2)D_3 \\
& b_1(X - X_0)D_1 + b_1D_2(X_1 - X_0 + X - X_1) + b_1D_3(X_1 - X_0 + X_2 - X_1 + X - \\
& X_2) \\
& b_1(X - X_0)D_1 + b_1(X - X_0)D_2 + b_1(X - X_0)D_3 = b_1(X - X_0)(D_1 + D_2 + D_3) \\
& b_1 \text{ terimi: } b_1(X - X_0) \tag{3.73}
\end{aligned}$$

Eşitlik 3.72’de c terimlerinin bir araya getirilip düğüm noktaları arasındaki uzaklıkların da eşit olduğu dikkate alınarak düzenlemesiyle eşitlik 3.74 elde edilir.

$$\begin{aligned}
& c_1(X - X_0)^2D_1 + c_1(X_1 - X_0)^2D_2 + 2c_1(X_1 - X_0)(X - X_1)D_2 + c_1(X - \\
& X_1)^2D_2 + c_1(X_1 - X_0)^2D_3 + 2c_1(X_1 - X_0)(X_2 - X_1)D_3 + c_1(X_2 - X_1)^2D_3 + \\
& 2c_1(X_1 - X_0)(X - X_2)D_3 + 2c_1(X_2 - X_1)(X - X_2)D_3 + c_1(X - X_2)^2D_3 \\
& c_1(X - X_0)^2D_1 + c_1D_2\{(X_1 - X_0)^2 + 2(X_1 - X_0)(X - X_1) + (X - X_1)^2\} + \\
& c_1D_3\{(X_1 - X_0)^2 + 2(X_1 - X_0)(X_2 - X_1) + (X_2 - X_1)^2 + 2(X_1 - X_0)(X - \\
& X_2) + 2(X_2 - X_1)(X - X_2) + (X - X_2)^2\} \\
& c_1 \text{ terimi:} \\
& c_1\{(X - X_0)^2D_1 + ((X_1 - X_0) + (X - X_1))^2D_2 + (2(X_1 - X_0) \text{ veya } 2(X_2 - \\
& X_1) + (X - X_2))^2D_3\} \tag{3.74}
\end{aligned}$$

Eşitlik 3.72’de d_1 terimlerinin bir araya getirilip düğüm noktaları arasındaki uzaklıkların da eşit olduğu dikkate alınarak düzenlemesiyle eşitlik 3.75 elde edilir.

$$\begin{aligned}
& d_1(X - X_0)^3 D_1 + d_1(X_1 - X_0)^3 D_2 + 3d_1(X_1 - X_0)^2(X - X_1)D_2 + \\
& 3d_1(X_1 - X_0)(X - X_1)^2 D_2 + d_1(X_1 - X_0)^3 D_3 + 3d_1(X_1 - X_0)^2(X_2 - X_1)D_3 + \\
& 3d_1(X_1 - X_0)(X_2 - X_1)^2 D_3 + 3d_1(X_1 - X_0)^2(X - X_2)D_3 + 6d_1(X_1 - X_0)(X_2 - \\
& X_1)(X - X_2)D_3 + 3d_1(X_1 - X_0)(X - X_2)^2 D_3
\end{aligned}$$

d_1 terimi:

$$\begin{aligned}
& d_1\{(X - X_0)^3 D_1 + [(X_1 - X_0)^3 + 3(X_1 - X_0)^2(X - X_1) + 3(X_1 - X_0)(X - \\
& X_1)^2]D_2 + [7(X_1 - X_0)^3 \text{ veya } 7(X_2 - X_1)^3 + 9(X_1 - X_0)^2(X - X_2) \text{ veya } 9(X_2 - \\
& X_1)^2(X - X_2) + 3(X_1 - X_0)(X - X_2)^2]D_3\} \quad (3.75)
\end{aligned}$$

Eşitlik 3.72’de d_2 terimlerinin bir araya getirilip düğüm noktaları arasındaki uzaklıkların da eşit olduğu dikkate alınarak düzenlemesiyle eşitlik 3.76 elde edilir.

$$\begin{aligned}
& d_2(X - X_1)^3 D_2 + d_2(X_2 - X_1)^3 D_3 + 3d_2(X_2 - X_1)^2(X - X_2)D_3 + \\
& 3d_2(X_2 - X_1)(X - X_2)^2 D_3
\end{aligned}$$

d_2 terimi:

$$\begin{aligned}
& d_2\{(X - X_1)^3 D_2 + [(X_2 - X_1)^3 + 3(X_2 - X_1)^2(X - X_2) + 3(X_2 - X_1)(X - \\
& X_2)^2]\} \quad (3.76)
\end{aligned}$$

Eşitlik 3.72’de tüm terimlerin düzenlenmesiyle birlikte yeni durum eşitlik 3.77’deki gibi elde edilir.

$$\begin{aligned}
Y = & a_1 + b_1(X - X_0) + c_1\{(X - X_0)^2 D_1 + ((X_1 - X_0) + (X - X_1))^2 D_2 + \\
& (2(X_1 - X_0) \text{ veya } 2(X_2 - X_1) + (X - X_2))^2 D_3\} + d_1\{(X - X_0)^3 D_1 + \\
& [(X_1 - X_0)^3 + 3(X_1 - X_0)^2(X - X_1) + 3(X_1 - X_0)(X - X_1)^2]D_2 + \\
& [7(X_1 - X_0)^3 \text{ veya } 7(X_2 - X_1)^3 + 9(X_1 - X_0)^2(X - X_2) \text{ veya } 9(X_2 - X_1)^2(X - \\
& X_2) + 3(X_1 - X_0)(X - X_2)^2]D_3\} + d_2\{(X - X_1)^3 D_2 + [(X_2 - X_1)^3 + \\
& 3(X_2 - X_1)^2(X - X_2) + 3(X_2 - X_1)(X - X_2)^2]\} + d_3(X - X_2)^3 D_3 + u \quad (3.77)
\end{aligned}$$

Daha önce de belirtildiği üzere düğüm noktaları arasındaki uzaklıklar eşitti. Buna göre, düğüm noktaları arasındaki uzaklık $(X_1 - X_0) = (X_2 - X_1) = (X_3 - X_2) = m$ gibi bir değer olsun. Eşitlik 3.77’de m değerleri yerine yazılırsa eşitlik 3.78 elde edilir³⁰³.

³⁰³ Suits, Mason ve Chan, s.134.

$$\begin{aligned}
Y = & a_1 + b_1(X - X_0) + c_1 \left\{ (X - X_0)^2 D_1 + (m + (X - X_1))^2 D_2 + \right. \\
& \left. (2m + (X - X_2))^2 D_3 \right\} + d_1 \{ (X - X_0)^3 D_1 + [m^3 + 3m^2(X - X_1) + \\
& 3m(X - X_1)^2] D_2 + [7m^3 + 9m^2(X - X_2) + 3m(X - X_2)^2] D_3 \} + d_2 \{ (X - \\
& X_1)^3 D_2 + [m^3 + 3m^2(X - X_2) + 3m(X - X_2)^2] D_3 \} + d_3 (X - X_2)^3 D_3 + u \quad (3.78)
\end{aligned}$$

Eşitlik 3.78'deki beş değişkenden oluşan kübik spline regresyon, aslında Y 'nin çoklu bir regresyonudur. Pratikte kullanmak için daha kolay olan ve daha fazla sayıdaki aralık için genelleştirme yapılmasına olanak sağlayan bir denklem elde etmek için eşitlik 3.78'de birkaç ilave işlem yapılır. Bu amaçla, $X \geq X_i$ ise $D_i^* = 1$ diğer durumlarda $D_i^* = 0$ olmak üzere D_1^* ve D_2^* gibi yeni gölge değişkenler seti oluşturulur. Yani, $X \geq X_1$ ise $D_1^* = 1$, diğer durumlarda $D_1^* = 0$ değerini alır. Aynı şekilde, $X \geq X_2$ ise $D_2^* = 1$, diğer durumlarda $D_2^* = 0$ değerini alır.

Bu notasyonların uygulanmasıyla eşitlik 3.78'e eş değer olan eşitlik aşağıdaki gibi ifade edilir.

$$Y = a_1 + b_1(X - X_0) + c_1(X - X_0)^2 + d_1(X - X_0)^3 + (d_2 - d_1)(X - X_1)^3 D_1^* + (d_3 - d_2)(X - X_2)^3 D_2^* \quad (3.79)$$

Eşitlik 3.78'deki gibi 3.79'da beş değişkenli çoklu bir regresyondur. Regresyon katsayıları a_1 , b_1 , c_1 ve d_1 doğrudan hesaplanabilir. d_2 ve d_3 ise son iki katsayıdan hesaplanabilir.

Dörtten fazla düğüm noktası yani üçten fazla aralık olduğu zaman, eşitlik 3.79'daki spline regresyon modeli uygulanabilir. Buna göre, k iç düğüm noktası sayısı olmak üzere, X ekseninde, $X_0, X_1, \dots, X_{k+1}$ noktalarında tanımlı $k + 1$ aralıklarına uydurulan kübik spline regresyon modelinin genelleştirilmiş hali aşağıdaki gibi yazılır³⁰⁴.

$$Y = a_1 + b_1(X - X_0) + c_1(X - X_0)^2 + d_1(X - X_0)^3 + \sum_{i=1}^k (d_{i+1} - d_i)(X - X_i)^3 D_i^*$$

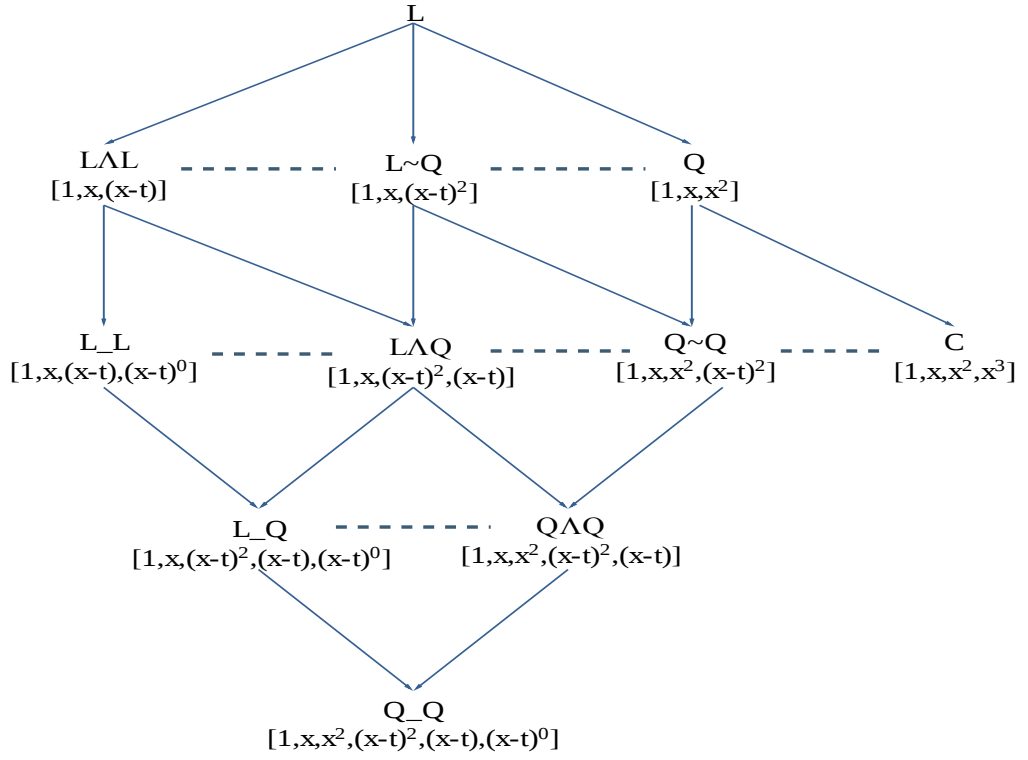
³⁰⁴ Suits, Mason ve Chan, s.135.

3.4.4. Splinelerin Kısmi Sıralaması

Düzgünleştirme düzeylerini birleştirme, hem doğrusal hem de kareli terimlerin aynı modelde olduğu doğrusal-kareli spline model gibi bir karma spline regresyon modeli oluşturabilir. Şekil 3.3'te, iki parçalı yani tek bir düğüm noktasının olduğu doğrusal ve kareli spline regresyon modellerinin farklı varyasyonları gösterilmiştir. Yani doğrusal ve kareli spline regresyon modellerinin farklı durumlar için kısmi sıralaması yapılmıştır³⁰⁵.

Şekil 3.3'e göre, L basit doğrusal, Q kareli ve C ise kübik polinom regresyon modellerini göstermektedir. LAL, sabit terimi kısıtlandırılmış sürekli doğrusal spline regresyonu, L~Q, ilk parçanın doğrusal ve ikinci parçanın kareli olduğu, aynı zamanda da düğüm noktasında birinci derece türevin sürekli olduğu spline regresyonu göstermektedir. L_L, kesikli olan kısıtlandırılmamış doğrusal spline regresyonu, LAQ, ilk parçanın doğrusal ve ikinci parçanın kareli olduğu ancak düğüm noktasında birinci derece türevinde kesikli olan spline regresyonu ve Q~Q, en düzgün kareli spline regresyonu (sabit kısıtlandırılmış ve düğüm noktasında birinci derece türevi sürekli) göstermektedir. L_Q, ilk parçanın doğrusal ve ikinci parçanın kareli olduğu kesikli spline regresyonu, QAQ, sadece sabiti kısıtlandırılmış sürekli kareli spline regresyonu ve Q_Q ise kısıtlandırılmamış kesikli spline regresyon modelini göstermektedir.

³⁰⁵ Smith, "Splines As a Useful and Convenient Statistical Tool", s.60.



L: Doğrusal, Q: Kareli, C: Kübik polinomlar

_: Kısıtlandırılmamış spline

Λ : Sürekli spline

$\sim$: Düğüm noktasında birinci derece türevi de sürekli olan spline

Şekil 3.3. Splineların Kısmi Sıralaması

DÖRDÜNCÜ BÖLÜM

SPLİNE REGRESYONDA MODEL SEÇİMİ

4.1. MODEL SEÇİM KRİTERLERİ

Uygulamalarda genellikle, değişik sayıda bağımsız değişken içeren farklı modeller tahmin edilerek bunlar arasında, bazı kriterlere göre en tatmin edici sonuçları veren model, en uygun model olarak seçilir³⁰⁶. Modellerin performansı, gözlenen değerlerle verileri test için elde edilen tahmini değerlerin birbirlerine ne kadar yakın olduğuylla ilişkilidir³⁰⁷.

4.1.1. Belirlilik Katsayısı

Yaygın şekilde kullanılan belirlilik katsayısı, regresyon modelinin yeterliliğinin bir ölçümüdür. Belirlilik katsayısı R^2 ile gösterilir ve bağımlı değişkende meydana gelen değişmelerin, modeldeki bağımsız değişken/değişkenler tarafından açıklanabilen oranını verir. R^2 'nin değeri 0 ile 1 arasında değişir. R^2 değeri büyüdükçe, Y 'deki değişimin modelin bağımsız değişkenleri ile açıklanan yüzdesi de büyür ve regresyon doğrusunun, gözlemlere uyumunun iyiliği artar³⁰⁸. Dolayısıyla, R^2 değerinin 1'e yakın olması, tahmin edilen regresyon doğrusunun verilere uyumunun daha iyi olduğunu gösterir. R^2 değeri³⁰⁹,

$$R^2 = 1 - \frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2} = \frac{\sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2}$$

$$R^2 = 1 - \frac{RSS}{TSS(RSS + ESS)} = \frac{ESS}{TSS(RSS + ESS)}$$

eşitliğiyle hesaplanır. Burada;

RSS : Hata kareler toplamı (Açıklanamayan değişkenlik)

ESS : Regresyon kareler toplamı (Açıklanan değişkenlik)

TSS : Genel kareler toplamı

³⁰⁶ Recep Tarı, *Ekonometri*, İkinci Baskı, Alfa Basım Yayın Dağıtım Ltd. Şti., İstanbul 1999, s.332.

³⁰⁷ Dursun Aydın, "A comparison of the Nonparametric Regression Models Using Smoothing Spline and Kernel Regression", *International Journal of Mathematical, Physical and Engineering Sciences*, 2(2), 2008, s.255.

³⁰⁸ Tarı, s.78.

³⁰⁹ Damador N. Gujarati, *Basic Econometrics*, 4th Edition, McGraw-Hill, New York 2003, s.538.

Y : Gözlenen değer

$\hat{Y}$: Tahmin edilen değer

$\bar{Y}$: Gözlenen değerlerin ortalamasıdır.

Regresyon kareler toplamı ve hata kareler toplamı matrislerle³¹⁰,

$$ESS = \hat{\beta}'X'Y' = Y'[X(X'X)^{-1}X']Y$$

$$RSS = Y'[I - X(X'X)^{-1}X']Y$$

biçiminde gösterilir. Burada Y , $(n \times 1)$ boyutlu bağımlı değişken vektörü ve X , $(n \times p)$ boyutlu bağımsız değişken/değişkenler matrisidir.

R^2 yaygın olarak kullanılan model seçim kriteri olsa da, bazı sorunlarla karşılaşmaktadır³¹¹.

✓ Verilen örnek büyüklüğü içinde tahmini değerler gerçek değerlere ne kadar yakın bulunsalar da, gelecek tahmininde bu garantiyi sağlamak mümkün olmayabilir.

✓ R^2 'lerin karşılaştırılabilmesi için modellerin fonksiyonel yapısının ve tahmin edicilerinin aynı olması gerekmektedir.

✓ Modele alınan bağımsız değişken sayısı arttıkça R^2 değeri de artmaktadır. Bu yöntemle maksimum R^2 'ye ulaşılabilir. Fakat bu durum gelecek tahmin hata varyansının da yükselmesine neden olacaktır. Bu nedenle model seçim kriteri olarak alınması her zaman yeterli değildir.

4.1.2. Düzeltilmiş R^2

Regresyonun standart hatasını minimize etme üzerine dayanır. Modele ilave bağımsız değişken eklenmesi sonucu R^2 değerinde bir artış söz konusu olmaktadır. Bunun nedeni de modele alınan yeni değişkenlerle birlikte serbestlik derecesi kaybının hesaba dâhil edilmemesidir³¹². Bu amaçla Henry Theil, bağımsız değişkenlerin katkısıyla ortaya çıkan R^2 değerindeki yükselmeyi önlemek için $\bar{R}^2$ sembolü ile gösterilen düzeltilmiş R^2 'yi geliştirmiştir. Düzeltilmiş R^2 ,

³¹⁰ Hakverdi, s.6.

³¹¹ Meltem Şengün Ucal, "Ekonometrik Model Seçim Kriterleri Üzerine Kısa Bir İnceleme", *C. Ü. İktisadi ve İdari Bilimler Dergisi*, 7(2), s.44.

³¹² Tarı, s.333.

$$\bar{R}^2 = 1 - \frac{RSS/(n-k)}{TSS/(n-1)} = 1 - (1 - R^2) \frac{(n-1)}{(n-k)}$$

eşitliği ile hesaplanır³¹³.

RSS: Hata kareler toplamı (Açıklanamayan değişkenlik)

TSS: Genel kareler toplamı

n: Gözlem sayısı

k: Bağımsız değişken sayısıdır.

Düzeltilmiş belirleme katsayısı her zaman belirleme katsayısından düşük çıkar. Bağımsız değişken sayısının artmasına karşın örnek büyüklüğü ve belirleme katsayısının düşmesi durumunda bu iki değer arasındaki fark azalmaktadır. Benzer şekilde bağımsız değişken sayısı sabit kaldığında örnek büyüklüğü ve belirleme katsayısının yükselmesi durumunda da bu iki değer arasındaki fark azalmaktadır. Sonuç olarak örnek büyüklüğünü de dikkate alarak hesaplanan düzeltilmiş belirleme katsayısı, özellikle doğrusal olmayan dağılımlardan elde edilen denklemlerin karşılaştırılmasında belirleme katsayısına göre daha çok tercih edilmektedir³¹⁴.

4.1.3. Bilgi Kriterleri

Bilgi kriterlerinde tesadüfî değişkenin doğru dağılıma yakınlığı, modelin gerçekliği ve doğru dağılım arasındaki fark ile ölçülmektedir. Yani, doğru ve alternatif modeller için *Y*'nin beklenen değerleri arasındaki farkın tahmini ile alternatif modelin doğru modele yaklaşması sağlanmaktadır. Bilgi kriterleri öngörü hata kareler ortalamasını minimize etmeye dayanır³¹⁵. Bilgi kriterlerinin en genel özelliği, iki veya daha fazla model kıyaslandığında, en düşük bilgi kriteri değerine sahip olan modelin seçilmesidir³¹⁶.

³¹³ Gujarati, s.537.

³¹⁴ Şahin, s.42.

³¹⁵ Selahattin Güriş ve Ebru Çağlayan, *Ekonometri Temel Kavramlar*, 3. Basım, Der Yayınları, İstanbul 2010, s.638.

³¹⁶ Gujarati, s.537.

4.1.3.1. Akaike Bilgi Kriteri (AIC)

Modele ilave deęiřken eklendięinde, bu deęiřkenlerin modelde yapacaęı etkiye bir sınırlama getirme fikri, AIC daha farklı bir řekilde ele alınmıřtır. Temel mantık $\bar{R}^2$ ile aynıdır ancak AIC'ye getirilen sınırlama daha katıdır.

Akaike Bilgi Kriterinin dikkat çekici özellikleri ařaęıdaki gibidir.

✓ AIC, sadece örneklem için deęil aynı zamanda regresyon modelinin örneklem dıřındaki gelecek tahmini performansı için de kullanılır.

✓ AIC, yuvalanmıř (nested) ve yuvalanmamıř (non-nested) modellerin ikisinde de kullanılır.

✓ AR modelinde gecikme uzunluęunun belirlenmesi için de kullanılmaktadır.

AIC kriteri,

$$AIC = e^{2m/n} \frac{\sum \hat{u}_i^2}{n} = e^{2m/n} \frac{RSS}{n} \quad (4.1)$$

eřitlięiyle elde edilir. Burada $2m/n$ sınırlama faktörüdür. Matematiksel olarak iřlem kolaylıęı saęlamak için 4.1 eřitlięinin her iki tarafının da logaritması alınır,

$$\ln AIC = \left(\frac{2m}{n} \right) + \ln \left(\frac{RSS}{n} \right) \quad (4.2)$$

eřitlięi elde edilir³¹⁷. (4.2) eřitlięi daha açık bir řekilde ařaęıdaki gibi de ifade edilebilir³¹⁸.

$$AIC = \frac{2m}{n} + \log \left(\frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{n} \right) \quad (4.3)$$

Y : Gözlenen(gerçek) deęer

$\hat{Y}$: Tahmin edilen deęer

n : Gözlem sayısı

m : Sabit terim dâhil parametre sayısıdır.

³¹⁷ Gujarati, s.537.

³¹⁸ Çaędař Hakan Aladaę ve Erol Eğrioęlu, “En İyi Yapay Sınır Aęı Mimarisinin Belirlenmesinde Uyarlamalı Aęırlıklı Bilgi Kriterinin Kullanılması”, Türkiye İstatistik Kurumu (Haz.), 19. İstatistik Arařtırma Sempozyumu Bildiriler Kitabı: 6-7 Mayıs, Ankara 2010, s.51.

4.1.3.2. Schwarz Bilgi Kriteri (SIC)

Schwarz bilgi kriteri, bayesyen işlemlerle elde edildiği için literatürde schwarz bayesyen bilgi kriteri veya bayesyen bilgi kriteri olarak da adlandırılır³¹⁹. Schwarz bilgi kriteri,

$$SIC = n^{m/n} \frac{\sum \hat{u}_i^2}{n} = n^{m/n} \frac{RSS}{n} \quad (4.4)$$

eşitliğiyle elde edilir. Burada m/n sınırlama faktörüdür. Aynı şekilde matematiksel olarak işlem kolaylığı sağlamak için 4.4 eşitliğinin her iki tarafının da logaritması alınır,

$$\ln SIC = \left(\frac{m}{n}\right) \ln n + \ln \left(\frac{RSS}{n}\right) \quad (4.5)$$

eşitliği elde edilir³²⁰. (4.5) eşitliği daha açık bir şekilde aşağıdaki gibi de ifade edilebilir³²¹.

$$SIC = \frac{m \log(n)}{n} + \log \left(\frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{n} \right) \quad (4.6)$$

Y : Gözlenen(gerçek) değer

$\hat{Y}$: Tahmin edilen değer

n : Gözlem sayısı

m : Sabit terim dâhil parametre sayısıdır.

Schwarz bilgi kriterinin dikkat çekici özellikleri aşağıdaki gibidir³²².

- ✓ SIC’de özünde AIC’ye benzemektedir.
- ✓ Sınırlama getirmede AIC’ye göre biraz daha katıdır.
- ✓ SIC değeri, her zaman AIC değerinden küçük çıkmaktadır.
- ✓ Schwarz bilgi kriteri de AIC gibi, sadece örneklem için değil aynı zamanda regresyon modelinin örneklem dışındaki gelecek tahmini performansı için de kullanılır.

³¹⁹ Herman J. Bierens, *Information Criteria and Model Selection*, 2006, Erişim Tarihi: 27 Haziran 2012, Pennsylvania State University Ağ Sitesi: <http://econ.la.psu.edu/~hbierens/INFORMATIONCRIT.PDF>.

³²⁰ Gujarati, s.538.

³²¹ Aladağ ve Eğrioğlu, s.52.

³²² Ucal, s.47.

4.1.4. PRESS İstatistiği

Allen (1974) tarafından önerilen PRESS istatistiği, farklı modellerin tahmin performansını karşılaştırmak için kullanılmaktadır³²³. Düzgünleştirme parametresinin seçiminde kullanılan çapraz geçerlilik yöntemi regresyon modellerindeki *PRESS* kriterine denktir³²⁴. Bu kriter, regresyon analizinde model seçimi için en çok kullanılan kriterlerden biridir. PRESS istatistiği modelin tahmin gücünün iyi bir göstergesidir. PRESS artığı, bir veri noktası hariç tutularak model tahmin edildiği zaman, hariç tutulan veri noktasının tahmin edilen değeriyle gözlenen değeri arasındaki farktır. PRESS istatistiği PRESS artıklarının kareleri toplamına eşittir. Bu hata kareler toplamı olduğu için, iyi bir modelde PRESS istatistiğinin küçük olması beklenir³²⁵. PRESS istatistiği aşağıdaki gibi hesaplanabilir³²⁶.

$$PRESS = \sum_{i=1}^n (Y_i - \hat{Y}_{i,-1})^2$$

Y_i : Gözlenen(gerçek) değer

$\hat{Y}_{i,-1}$: Örneklemdeki i . Y_i gözlemi olmaksızın tahmin edilen regresyon denkleminde elde edilen Y_i 'nin tahmini değeridir.

Ayrıca, PRESS istatistiği matrislerle,

$$PRESS = \sum_{i=1}^n e_{i,-1}^2$$

$$e_{i,-1} = \frac{Y_i - \hat{Y}_i}{1 - (X - \bar{X})'(X'X)^{-1}(X - \bar{X})}$$

$$e_{i,-1} = \frac{e_i}{1 - p_{ii}}$$

ve

³²³ David M. Allen, "The Relationship between Variable Selection and Data Augmentation and a Method for Prediction", *Technometrics*, 16(1), 1974, ss.125-127.

³²⁴ Elif Öztürk, *Parametrik Olmayan ve Yarı Parametrik Panel Veri Modelleri: Çevresel Kuznets Eğrisinin Analizi*, (Yayımlanmamış Doktora Tezi), Marmara Üniversitesi Sosyal Bilimler Enstitüsü, İstanbul 2011, s.76.

³²⁵ Hurley, Hussy, McKeow and Addy, s.2.

³²⁶ Gürış ve Çağlayan, s.642.

$$PRESS = \sum_{i=1}^n \left(\frac{e_i}{1 - p_{ii}} \right)$$

olarak hesaplanır. Hesaplanan PRESS artıkları en küçük kareler artıklarından büyüktür.

4.1.5. R^2_{pred} Değeri

Bu istatistik değeri, regresyon modelinin öngörü kapasitesinin bir göstergesini verir. PRESS istatistiğine bağlı olarak R^2_{pred} değeri aşağıdaki gibi hesaplanır³²⁷.

$$R^2_{pred} = 1 - \frac{PRESS}{\sum_{i=1}^n (Y_i - \bar{Y})^2}$$

Y : Gözlenen(gerçek) değer

$\bar{Y}$: Gözlenen değerlerin ortalaması

n : Gözlem sayısıdır.

4.1.6. Mallows'un C_p Kriteri

Bu kriter öngörü ortalama hata karesini minimize etmeye dayanır. C_p istatistiği değerinin küçük olması beklenir³²⁸. C_p kriteri düzeltilmiş R^2 istatistiğiyle, düzeltilmiş R^2 de R^2 istatistiği ile yakından ilişkilidir. Yani değişimleri birbirlerine bağlıdır³²⁹.

Modelin sabit teriminin de dâhil olduğu k adet bağımsız değişken olduğu varsayılan bir modelde, p adet bağımsız değişken seçilerek ($p \leq k$) hata kareler toplamının hesaplanmasıyla Mallows'un model seçim kriteri³³⁰,

$$C_p = \frac{RSS_p}{\hat{\sigma}^2} - (n - 2p)$$

$$C_p = \frac{RSS_p}{\left(\frac{\sum_{i=1}^n e_i^2}{n - p} \right)} - (n - 2p)$$

eşitliğiyle bulunabilir³³¹.

³²⁷ Raymond H. Myers, *Classical and Modern Regression with Applications*, Second Edition, Duxbury Press, Pacific Grove, CA 1990, s.171.

³²⁸ Güriş ve Çağlayan, s.634.

³²⁹ Norman R. Draper and Harry Smith, *Applied Regression Analysis*, Third Edition, John Wiley&Sons, New York 1998, s.332.

³³⁰ Gujarati, s.538.

³³¹ Hakverdi, s.25.

RSS_p : Hata (artık) kareler toplamı (Residual Sum of Squares)

p : Modeldeki bağımsız değişken sayısı

n : Gözlem sayısı

$\hat{\sigma}^2$: Anakütle varyansının (σ^2) tahmin edicisidir, yani hata kareler ortalamasıdır.

4.1.7. Modelin Hatası (S)

Modelin standart hatası, gerçek değerlerle tahmini değerler arasındaki farkın ölçüleriyle bulunur. Standart sapmayla benzerlik gösterir. Standart hata aşağıda belirtildiği gibi hesaplanmaktadır. Standart hatası daha küçük olan model değişkenler arasındaki ilişkiyi daha iyi temsil etmektedir³³².

$$S_{YX} = \sqrt{\frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{n - 1 - m}}$$

Y : Gözlenen(gerçek) değer

$\hat{Y}$: Tahmin edilen değer

n : Gözlem sayısı

m : Parametre sayısıdır.

4.1.8. Hata Kareler Ortalaması (MSE)

Model değerlendirme ve seçim kriteri olarak hata kareler ortalaması da kullanılabilir. Hata kareler toplamının serbestlik derecesine bölünmesiyle elde edilir. Hesaplanan modeller için de en küçük MSE değerine sahip olan model seçilir. MSE değeri, aşağıdaki gibi hesaplanır³³³.

$$MSE = \sum_{i=1}^n \frac{(Y_i - \hat{Y}_i)^2}{n - p}$$

Y : Gözlenen(gerçek) değer

$\hat{Y}$: Tahmin edilen değer

³³² Alaaddin Başar ve Erkan Oktay, *Uygulamalı İstatistik 2 Kısa Teorik Bilgiler ve Çözülmüş Problemler*, Aktif Yayınevi, Erzurum 2004.

³³³ Ufuk Karadavut, Alper Sinan, Aşı Genç, Çetin Palta, Aydın Karakoca ve Şeref Aksoyak, “Dağdaş Buğday Bitkilerinin Büyüme Eğrilerinin Belirlenmesinde Model Seçimi”, *4. İstatistik Kongresi: 8-12 Mayıs 2005*, Belek/Antalya, s.259.

p : Parametre sayısı

n : Gözlem sayısıdır.

Parametre sayısı arttıkça hata kareler toplamı değeri azalır. Buna bağlı olarak hata kareler ortalaması değeri de azalır ancak bu süreklilik devam etmez. Çünkü hata kareler ortalaması ilave parametre sayısına bağlı olarak başlangıçta azalır, sonra sabit kalır ve sonunda artış olur. Bunun sebebi de modele ilave parametre sayısının eklenmesiyle hata kareler toplamındaki azalmanın serbestlik derecesi kaybını karşılayamamasıdır³³⁴.

4.1.9. Hata Kareler Ortalaması Karekökü (RMSE)

Hata kareler ortalaması karekökü sıklıkla, bir model tarafından tahmin edilen değerlerle gözlenen değerler arasındaki farklılığın ölçümü için kullanılmaktadır. Aşağıdaki gibi hesaplanmaktadır³³⁵.

$$RMSE = \sqrt{\sum_{i=1}^n \frac{(Y_i - \hat{Y}_i)^2}{n}}$$

Ayrıca hata kareler ortalamasının karekökünün örneklemdeki gözlenen değerlerin ortalamasına oranıyla, örneklem hacminin RMSE üzerindeki etkisi azaltılmaya çalışılır. Buna göre, örneklem hacminin etkisinin azaltıldığı %RMSE (Percentage Root Mean Square Error) değeri aşağıdaki gibi hesaplanır.³³⁶

$$\%RMSE = \frac{100}{\bar{Y}} * \sqrt{\sum_{i=1}^n \frac{(Y_i - \hat{Y}_i)^2}{n}}$$

Y : Gözlenen(gerçek) değer

$\hat{Y}$: Tahmin edilen değer

n : Gözlem sayısı

p : Parametre sayısı

³³⁴ Montgomery, Peck and Vining, s.334.

³³⁵ Aydın, "A comparison of the Nonparametric Regression Models Using Smoothing Spline and Kernel Regression", s.255.

³³⁶ Keith Loague and Richard E. Green, "Statistical and Graphical Methods for Evaluating Solute Transport Models: Overview and Application", *Journal of Contaminant Hydrology*, 7, 1991, s.58.

$\bar{Y}$: Gözlenen değerlerin ortalamasıdır.

%RMSE istatistiğinin en küçük değeri sıfırdır. %RMSE değeri en düşük olan model seçilir.

4.1.10. Ortalama Mutlak Hata (MAE) ve Ortalama Mutlak Hata Yüzdesi (MAPE)

Ortalama mutlak hata, tahmin edilen değerlerin gözlenen değerlerden ne kadar uzakta olduğunu gösteren istatistiksel bir ölçümdür. Çoğunlukla zaman serilerinde kullanılmaktadır. Ortalama mutlak hata aşağıdaki gibi hesaplanmaktadır³³⁷.

$$MAE = \frac{1}{n} \sum_{i=1}^n |Y_i - \hat{Y}_i|$$

Ortalama mutlak hata yüzdesi ise, tahminlerin nispi(oransal) doğruluğunu belirlemek için kullanılır. Ortalama mutlak hata yüzdesi, tahminlerin gerçek değerlerden ne kadar sapma gösterdiğini yüzde olarak ölçer. Yüzdeler sezgisel olarak anlaşılır ölçümlerdir. MAPE istatistiği aşağıdaki gibi hesaplanmaktadır³³⁸.

$$MAPE = \frac{100}{n} \sum_{i=1}^n \frac{|Y_i - \hat{Y}_i|}{|Y_i|}$$

Y : Gözlenen(gerçek) değer

$\hat{Y}$: Tahmin edilen değer

n : Gözlem sayısıdır.

4.1.11. Oransal Hata (Relative Error)

Regresyon artıklarını temel alan oransal hata büyüdükçe, model zayıflamaktadır. Oransal hata aşağıdaki gibi hesaplanmaktadır³³⁹.

$$E = \frac{\sum_{i=1}^n \frac{(\hat{Y}_i - Y_i)}{Y_i} * 100}{n}$$

³³⁷ Erişim Tarihi: 27 Haziran 2012, http://en.wikipedia.org/wiki/Mean_absolute_error.

³³⁸ Marcus, s.15.

³³⁹ Ufuk Karadavut, Aşır Genç, Abdurrahman Tozluca ve Çetin Palta, "Silajlık ve Danelik Mısırlarda Kuru Madde Birikiminin Bazı Matematiksel Büyüme Modelleri ile Analizi", *Tarım Bilimleri Dergisi*, 16, 2010, s.92.

Y : Gözlenen değer

$\hat{Y}$: Tahmin edilen değer

n : Gözlem sayısıdır.

4.1.12. Hata Kümesinin Katsayısı (CRM)

Hata kümesinin katsayısı (Coefficient of Residual Mass), gözlenen değerler toplamının tahmin edilen değerler toplamından farkının gözlenen değerler toplamına oranlanmasıyla elde edilir. CRM istatistiği, pozitif veya negatif olabilir. CRM deri, -1 ile +1 arasında değişmektedir. Negatif bir değer tahmin edilen değerlerin çoğunluğunun gözlenen değerlerden büyük olduğunu gösterir. Pozitif bir değer de tahmin edilen değerlerin çoğunluğunun gözlenen değerlerden küçük olduğunu gösterir. Ancak sıfıra en yakın olan değer tercih edilmektedir. CRM istatistiği aşağıdaki gibi hesaplanmaktadır³⁴⁰.

$$CRM = \frac{\sum_{i=1}^n Y_i - \sum_{i=1}^n \hat{Y}_i}{\sum_{i=1}^n Y_i}$$

Y : Gözlenen değer

$\hat{Y}$: Tahmin edilen değer

n : Gözlem sayısıdır.

4.1.13. Etkinlik İndeksi (Kapa İstatistiği)

Etkinlik indeksi kapa istatistiği olarak da bilinmektedir. -1 ile +1 arasında değişmektedir. Genel olarak 00-0.40 arasında olan değerler zayıf, 0.40-0.60 arası orta derecede, 0.60-0.80 arası iyi ve 0.80 üzeri ise mükemmel olarak tanımlanmaktadır. Ancak sıfırın altında bir değer alınması durumunda değerlendirme yapılamamaktadır. Etkinlik indeksi aşağıdaki gibi hesaplanmaktadır³⁴¹.

$$EI = \frac{\bar{Y} - \bar{\hat{Y}}}{1 - \bar{\hat{Y}}}$$

³⁴⁰ Jo Smith, Pete Smith and Tom Addiscott, "Quantitative Methods to Evaluate and Compare Soil Organic Matter (SOM) Models", (Ed. David S. Powlson, Pete Smith and Jo U. Smith), *Evaluation of Soil Organic Matter Models*, (pp.181-199), Springer-Verlag, Berlin 1996, s.194.

³⁴¹ Ufuk Karadavut ve Abdurrahman Tozluca, "Çavdar Bitkisinde (*Secale Cereale* L.) Bazı Karakterlerin Büyüme Analizleri", *Bitkisel Araştırma Dergisi*, 2, 2005, s.3.

$\bar{Y}$: Gözlenen değerlerin ortalaması

$\hat{\bar{Y}}$: Tahmin edilen değerlerin ortalamasıdır.

BEŞİNCİ BÖLÜM

İHRACATIN İTHALATI KARŞILAMA ORANI DAĞILIMINDA DÜĞÜM NOKTASININ BELİRLENMESİ VE SPLİNE REGRESYON MODELLERİNİN KURULMASI

5.1. İHRACATIN İTHALATI KARŞILAMA ORANI DAĞILIMINDA ÖNERİLEN YÖNTEMLE DÜĞÜM NOKTALARININ BELİRLENMESİ

5.1.1. 1923-2010 Yılları Türk Dış Ticaretinin Kısaca Değerlendirilmesi

İhracat, ülkenin dış kaynak kullanmadan elde ettiği döviz kazancıdır. Yani yabancı ülkelere döviz karşılığı yapılan satışlardır. İthalat ise, yurtdışında üretilmiş malların, ülkedeki alıcılar tarafından satın alınmasıdır. Dış ticaret, ülkelerin kalkınmalarının sürdürülmesinde ve dünya ülkeleriyle bütünleşmenin sağlanmasında en önemli unsurlardan biri olarak görülmektedir³⁴². Bütün ülkeler uluslararası ticarete katılmaktan dolayı elde edeceği faydayı optimize edecek bazı düzenlemeler yaparlar. Uluslararası ticarete koydukları kısıtlamalar ve düzenlemeler ticaret politikası adı altında toplanır. Ticarete konulan kısıtlamaların en önemlisi tarifelerdir. Ticareti yapılan mal üzerine konulan vergilere tarife denilir. İthal edilen mal üzerinden alınan vergiye ithalat tarifesi adı verilir. İhraç edilen mal üzerinden ödenen vergiye de ihracat tarifesi denir. İthalat tarifesi ihracat tarifesinden daha önemlidir³⁴³.

1923-2010 döneminde Türkiye Cumhuriyeti'nde izlenen dış ticaret politikaları ülkenin temel öncelikleri, sanayileşme politikaları, uluslararası siyasal ve ekonomik konjonktür, doğrultusunda belirlenmiştir. Bu çerçevede dış ticaret politikaları dönem boyunca söz konusu öncelikler ve süregelen siyasal ve ekonomik koşullar açısından değerlendirilerek yeniden şekillendirilmiş ve uygulamaya konulmuştur.

1920'li ve 1930'lu yıllarda Türkiye geri kalmış bir tarım ekonomisi durumundaydı. 1929 Büyük Buhranı, hammadde ihracatçısı ve çoğunlukla tüketim mallarından oluşan sanayi malı ithalatçısı olan ve dünya ticaretine serbest ticaret koşulları altında katılan Türkiye ve Türkiye benzeri tüm azgelişmiş ülkeleri benzer biçimde etkilemiştir. 1929 ile 1931 dış ticarete korumacılığın ve kambiyo

³⁴² Ertekin, M. S. ve Erol, K., "1980 Sonrası Dönemde Türkiye Dış Ticaretinin Genel Bir Analizi", *Anadolu Üniversitesi İ.İ.B.F Dergisi*, 16 (1-2), 2000, 223-254.

³⁴³ Tunç, H., *Uluslararası Ticaret Para ve Finans*, Alfa Kitabevi, İstanbul 2004.

denetimlerinin etkinlik kazandığı, 1932 ise devletçi politikaların uygulanmaya başlandığı yıllar olmuştur. 1930 sonrası sanayileşme çabalarının odağına devlet eliyle oluşturulan teşebbüsler oturtulmuştur. 1930-1945 arasında dış ticaret dengesinde yalnızca 1938 yılında açık vermiştir. Dış ticarete verilen açıktan fazlaya geçilmesinin 1923 ile 1929 yılları arasındaki dönem göz önüne alındığında esas olarak ithalatın kısılmasından kaynaklandığı anlaşılır. 1930 sonrasında izlenen dışa kapalı, korumacı, dış dengeye önem veren politikalar terk edilerek serbest ticaretçi iktisat politikaları uygulanmaya başlanmıştır. İthalatın serbestleştirilmesiyle birlikte 1947'den itibaren ithalatta önemli artışlar yaşanmış ve dış ticarete açık vermeye başlanmıştır³⁴⁴.

1930'lara damgasını vuran korumacı dış ticaret politikaları, 1946 yılında Türk lirasının yüzde 100'den fazla değer kaybetmesi ve liberalizasyon listelerinin saptanarak ithalatta kota uygulamasının sınırlandırılması ile terk edilmeye başlanmıştır. 1953-1961 yılları arasındaki dönemde istikrarlı bir şekilde hareket eden ve büyük ölçüde tarım ürünlerinden oluşan ihracat sayesinde, dış ticaret açıkları daraltılmış ve sürdürülebilir olmuştur.

1960'larda ekonomi politikalarında planlama dönemi başlamış, 1961'de başlayan yeni dönem; ithal ikameci, korumacı dış ticaret politikalarının sürdürüldüğü ve iç piyasaya yönelik bir sanayileşme politikasının benimsendiği bir dönem olmuştur. 1960'larda toplam ihracat içerisinde sanayi ürünlerinin payı yüzde 13-18 arasında değişirken, 1970'lerde yüzde 20-39'lara yükselmiştir. 1962 ile 1976 yılları arasındaki dönemde ithalatın GSMH'deki payı önemli oranda artmış, ihracatın ithalatı karşılama oranı ise aynı şekilde düşmüştür. 1970'lerde artan petrol fiyatlarının dünya ekonomisini bunalıma sürüklediği bir dönemde, kısa dönemli borçlanmalarla büyüyen ve ithalatı artmaya devam eden Türk ekonomisinin, 1977'de dış dengeleri büyük çapta bozulmuştur.

24 Ocak 1980 Türkiye ekonomi politikaları açısından bir dönüm noktasıdır. İhracat destekleri ekonomi politikalarının temel araçlarından biri olmuştur. 1980 sonrası politikalarla ihracat, 1980-1984 arasındaki dönemde yüzde 314 oranında artmış, ancak ithalatta da artış olması sebebiyle dış açıklar 1980 sonrası dönemde de devam etmiştir. 1990'da Türk parasının konvertibilitesi (Milli paraların karşılıklı olarak birbirine

³⁴⁴ Taşkın, M. M., "1923-2003 Döneminde Türkiye Cumhuriyeti'nin Dış Ticaret Politikaları", 2003, Erişim Tarihi: 15 Eylül 2011, www.dtm.gov.tr/dtmadmin/upload.

çevrilebilmesi) ilan edilmiştir. 1990’larda ve 2000’li yıllarda Türkiye’nin dış ticaret politikalarını şekillendiren iki temel bütünleşme Dünya Ticaret Örgütü üyeliği ve Avrupa Birliği ile girilen Gümrük Birliği olmuştur.

5.1.2. Yöntem ve Uygulaması

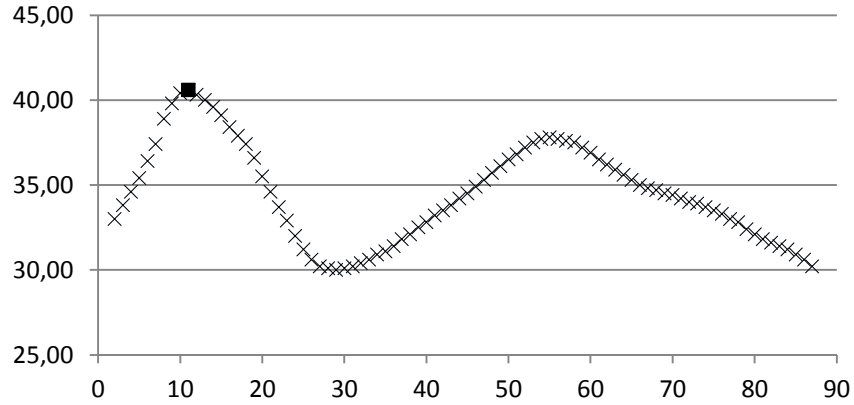
Dağılımda düğüm noktalarının sayısı ve konumunun belirlenmesi için doğrusal spline regresyonların kullanıldığı belirlilik katsayısına (R^2) dayalı bir yöntem önerilmiştir. Spline regresyon modellerinin kurulabilmesi için önerilen yöntemle gerekli olan düğüm noktaları belirlenmiştir.

R^2 ’nin kullanılmasının sebebi, bağımsız değişken ile bağımlı değişken arasında yapısal değişiklikler olduğunda R^2 değerleri de değişiklik gösterir. Buradan hareketle, R^2 değerlerindeki önemli değişikliklere (artış veya azalışlara) göre düğüm noktaları belirlenmiştir.

Uygulamada, 1923-2010 yılları arasında ihracatın ithalatı karşılama oranları dağılımında toplam 88 gözlem bulunmaktadır. Çalışmada, 1923 yılı 1. gözlem, 1924 yılı 2. gözlem, ..., 2010 yılı 88. gözlem olarak ele alınmıştır. Öncelikle her bir gözlem aday düğüm noktası olarak düşünülerek ayrı ayrı doğrusal spline regresyon modelleri kurulmuştur. Yani, dağılımda 2. gözlem düğüm noktası alınarak tek düğüm noktalı doğrusal spline regresyon modeli kurulmuş ve R^2 değeri hesaplanmıştır. Aynı şekilde 3. gözlem düğüm noktası alınarak tek düğüm noktalı spline regresyon modeli kurulmuş ve R^2 değeri hesaplanmıştır. 87. gözleme kadar bu şekilde devam edilerek toplam 86 tane doğrusal spline regresyon modeli kurulmuş ve R^2 değerleri (%) hesaplanmıştır. Hesaplanan R^2 değerleri (%) Tablo 5.1 ve Şekil 5.1’de gösterilmiştir. R^2 değerleri (%) incelendiğinde 11. gözlemden sonra R^2 değerinde düşüş eğilimi başladığı görülmektedir. Bu nedenle, R^2 değerinin değişim gösterdiği 11. gözlem birinci düğüm noktası olarak düşünülmüştür.

Tablo 5.1. İlk Adım Sonrası R^2 Değerleri

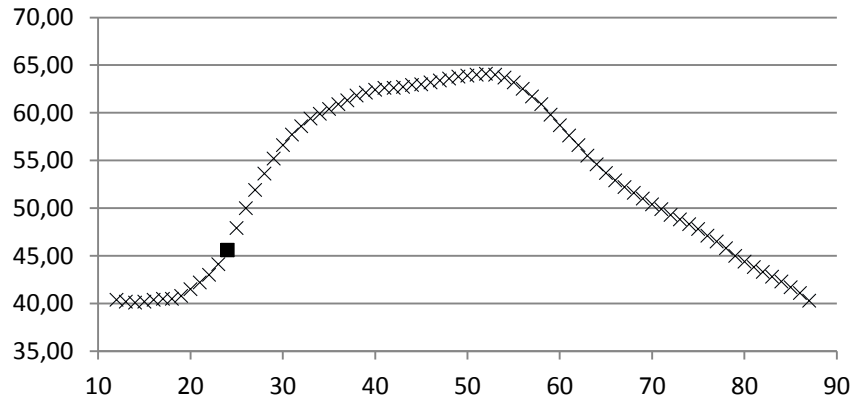
Gözlem No	R^2 (%)	Gözlem No	R^2 (%)	Gözlem No	R^2 (%)	Gözlem No	R^2 (%)	Gözlem No	R^2 (%)
2	33	20	35,5	38	32,1	56	37,7	74	33,7
3	33,8	21	34,6	39	32,5	57	37,6	75	33,5
4	34,6	22	33,7	40	32,8	58	37,5	76	33,3
5	35,4	23	32,9	41	33,2	59	37,2	77	33
6	36,4	24	32	42	33,5	60	36,9	78	32,8
7	37,4	25	31,2	43	33,8	61	36,5	79	32,4
8	38,9	26	30,6	44	34,2	62	36,2	80	32,1
9	39,8	27	30,2	45	34,5	63	35,9	81	31,8
10	40,4	28	30,1	46	34,9	64	35,6	82	31,6
11	40,6	29	30	47	35,3	65	35,3	83	31,4
12	40,3	30	30,1	48	35,7	66	35	84	31,2
13	40	31	30,2	49	36,1	67	34,8	85	30,9
14	39,6	32	30,4	50	36,5	68	34,7	86	30,6
15	39,1	33	30,6	51	36,8	69	34,5	87	30,2
16	38,4	34	30,9	52	37,2	70	34,4		
17	37,9	35	31,1	53	37,5	71	34,2		
18	37,4	36	31,4	54	37,7	72	34		
19	36,6	37	31,8	55	37,8	73	33,9		

**Şekil 5.1.** İlk Adım Sonrası R^2 Değerlerinin Dağılımı

Daha sonra, birinci parçayı oluşturan ilk 10 gözlem analizden çıkartılarak ikinci adıma geçilmiştir. Aynı şekilde 12. gözlem düğüm noktası alınarak tek düğüm noktalı spline regresyon modeli kurulmuş ve R^2 değeri (%) hesaplanmıştır. 87. gözleme kadar bu şekilde devam edilerek toplam 76 tane doğrusal spline regresyon modeli kurulmuş ve R^2 değerleri (%) hesaplanmıştır. Hesaplanan R^2 değerleri (%) Tablo 5.2 ve Şekil 5.2’de gösterilmiştir. R^2 değerleri (%) incelendiğinde 24. gözlemde R^2 değerindeki artış oranının önceki değerlerden fazla olduğu görülmektedir. Bu nedenle, R^2 değerinin değişim gösterdiği 24. gözlem ikinci düğüm noktası olarak düşünülmüştür.

Tablo 5.2. İkinci Adım Sonrası R^2 Değerleri

Gözlem No	R^2 (%)	Gözlem No	R^2 (%)	Gözlem No	R^2 (%)	Gözlem No	R^2 (%)
12	40,4	31	57,7	50	63,9	69	51
13	40,2	32	58,6	51	64	70	50,4
14	40,1	33	59,4	52	64,1	71	49,9
15	40,2	34	59,9	53	64	72	49,3
16	40,4	35	60,4	54	63,7	73	48,8
17	40,5	36	60,9	55	63,2	74	48,3
18	40,5	37	61,3	56	62,5	75	47,8
19	40,8	38	61,8	57	61,7	76	47,1
20	41,5	39	62,1	58	60,9	77	46,5
21	42,2	40	62,4	59	59,8	78	45,8
22	43	41	62,6	60	58,7	79	45
23	44,1	42	62,6	61	57,6	80	44,4
24	45,6	43	62,7	62	56,6	81	43,8
25	47,9	44	62,9	63	55,5	82	43,3
26	50	45	63	64	54,6	83	42,8
27	51,9	46	63,2	65	53,7	84	42,3
28	53,6	47	63,4	66	52,9	85	41,7
29	55,2	48	63,6	67	52,2	86	41,1
30	56,6	49	63,8	68	51,6	87	40,3

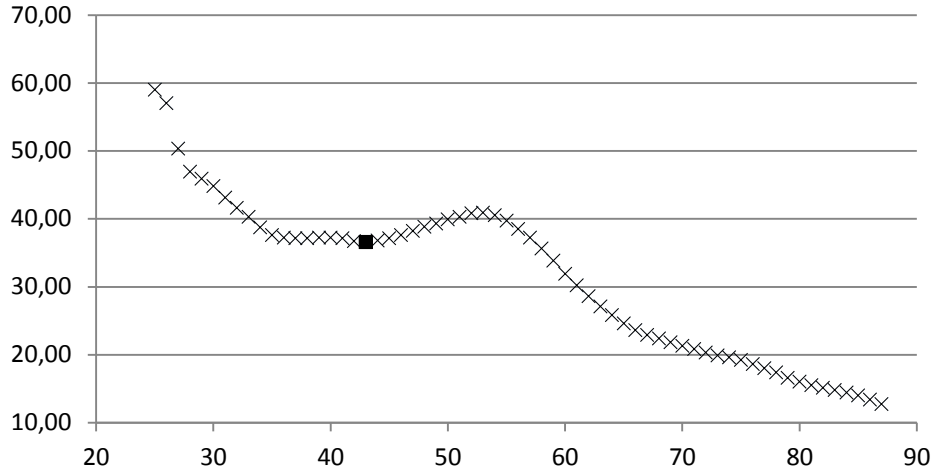
**Şekil 5.2.** İkinci Adım Sonrası R^2 Değerleri Dağılımı

İkinci düğüm noktasına kadar olan 13 gözlem analizden çıkartılarak üçüncü adıma geçilmiştir. Aynı şekilde 25. gözlem düğüm noktası alınarak tek düğüm noktalı spline regresyon modeli kurulmuş ve R^2 değeri (%) hesaplanmıştır. 87. gözleme kadar bu şekilde devam edilerek toplam 63 tane doğrusal spline regresyon modeli kurulmuş ve R^2 değerleri (%) hesaplanmıştır. Hesaplanan R^2 değerleri (%) Tablo 5.3 ve Şekil 5.3'te gösterilmiştir. R^2 değerleri (%) incelendiğinde 43. gözlemde sonra R^2 değerinde artış

eğilimi başladığı görülmektedir. Bu nedenle, R^2 değerinin değişim gösterdiği 43. gözlem üçüncü düğüm noktası olarak düşünülmüştür.

Tablo 5.3. Üçüncü Adım Sonrası R^2 Değerleri

Gözlem No	R^2 (%)	Gözlem No	R^2 (%)	Gözlem No	R^2 (%)	Gözlem No	R^2 (%)
25	59	41	37,1	57	37,2	73	19,9
26	57	42	36,7	58	35,6	74	19,6
27	50,3	43	36,6	59	33,8	75	19,2
28	46,9	44	36,8	60	31,9	76	18,6
29	45,9	45	37,1	61	30,2	77	18
30	44,8	46	37,6	62	28,6	78	17,4
31	43,1	47	38,2	63	27,1	79	16,6
32	41,6	48	38,8	64	25,8	80	16
33	40,3	49	39,3	65	24,6	81	15,5
34	38,7	50	39,9	66	23,6	82	15,1
35	37,6	51	40,3	67	22,9	83	14,8
36	37,2	52	40,8	68	22,4	84	14,4
37	37,1	53	40,9	69	21,8	85	14
38	37,1	54	40,5	70	21,3	86	13,4
39	37,2	55	39,7	71	20,8	87	12,7
40	37,2	56	38,5	72	20,3		



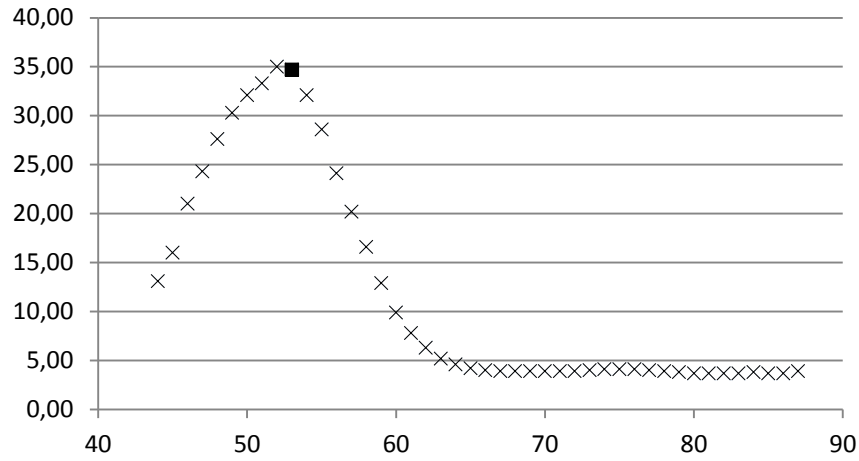
Şekil 5.3. Üçüncü Adım Sonrası R^2 Değerleri Dağılımı

Üçüncü düğüm noktasına kadar olan 13 gözlem analizden çıkartılarak dördüncü adıma geçilmiştir. Aynı şekilde 44. gözlem düğüm noktası alınarak tek düğüm noktalı spline regresyon modeli kurulmuş ve R^2 değeri (%) hesaplanmıştır. 87. gözleme kadar bu şekilde devam edilerek toplam 44 tane doğrusal spline regresyon modeli kurulmuş ve R^2 değerleri (%) hesaplanmıştır. Hesaplanan R^2 değerleri (%) Tablo 5.4 ve Şekil

5.4'te gösterilmiştir. R^2 değerleri (%) incelendiğinde 53. gözlemde sonra R^2 değerinde önemli oranda düşüş eğilimi başladığı görülmektedir. Bu nedenle, R^2 değerinin değişim gösterdiği 53. gözlem dördüncü düğüm noktası olarak düşünülmüştür.

Tablo 5.4. Dördüncü Adım Sonrası R^2 Değerleri

Gözlem No	R^2 (%)	Gözlem No	R^2 (%)	Gözlem No	R^2 (%)
44	13,1	59	12,9	74	4,1
45	16	60	9,9	75	4,1
46	21	61	7,8	76	4,1
47	24,3	62	6,3	77	4
48	27,6	63	5,2	78	3,9
49	30,3	64	4,6	79	3,8
50	32,1	65	4,2	80	3,7
51	33,3	66	4	81	3,7
52	35	67	3,9	82	3,7
53	34,7	68	3,9	83	3,7
54	32,1	69	3,9	84	3,8
55	28,6	70	3,9	85	3,7
56	24,1	71	3,9	86	3,7
57	20,2	72	3,9	87	3,9
58	16,6	73	4		



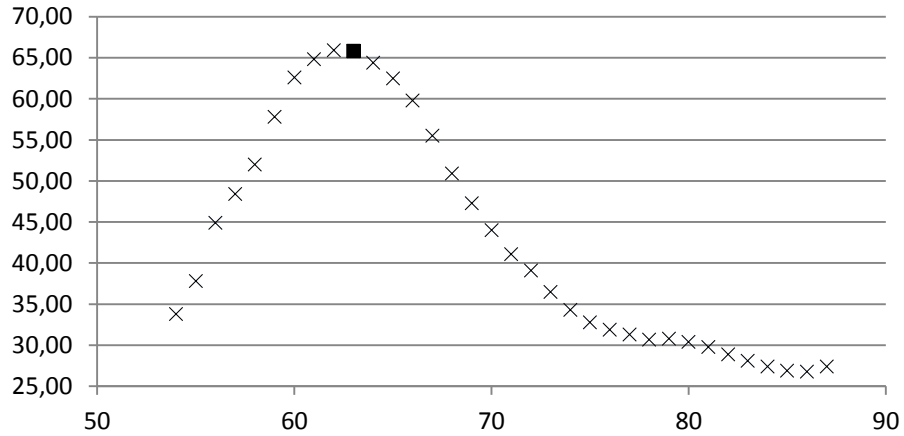
Şekil 5.4. Dördüncü Adım Sonrası R^2 Değerleri Dağılımı

Dördüncü düğüm noktasına kadar olan 10 gözlem analizden çıkartılarak beşinci adıma geçilmiştir. Aynı şekilde 54. gözlem düğüm noktası alınarak tek düğüm noktalı spline regresyon modeli kurulmuş ve R^2 değeri (%) hesaplanmıştır. 87. gözleme kadar bu şekilde devam edilerek toplam 34 tane doğrusal spline regresyon modeli kurulmuş ve R^2 değerleri (%) hesaplanmıştır. Hesaplanan R^2 değerleri (%) Tablo 5.5 ve Şekil

5.5'te gösterilmiştir. R^2 değerleri (%) incelendiğinde 63. gözlemden sonra R^2 değerinde önemli oranda düşüş eğilimi başladığı görülmektedir. Bu nedenle, R^2 değerinin değişim gösterdiği 63. gözlem beşinci düğüm noktası olarak düşünülmüştür.

Tablo 5.5. Beşinci Adım Sonrası R^2 Değerleri

Gözlem No	R^2 (%)	Gözlem No	R^2 (%)	Gözlem No	R^2 (%)
54	33,8	66	59,8	78	30,7
55	37,8	67	55,5	79	30,8
56	44,9	68	50,9	80	30,4
57	48,4	69	47,3	81	29,8
58	52	70	44	82	28,9
59	57,8	71	41,1	83	28,1
60	62,6	72	39,1	84	27,4
61	64,8	73	36,5	85	26,9
62	65,9	74	34,3	86	26,8
63	65,8	75	32,8	87	27,4
64	64,4	76	31,9		
65	62,5	77	31,3		

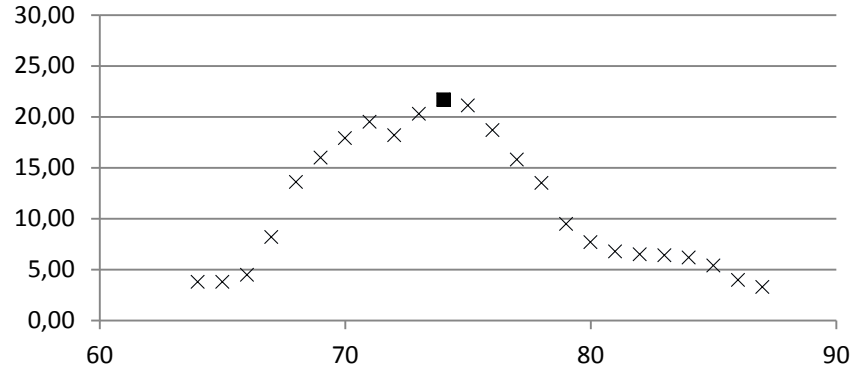


Şekil 5.5. Beşinci Adım Sonrası R^2 Değerleri Dağılımı

Beşinci düğüm noktasına kadar olan 10 gözlem analizden çıkartılarak altıncı adıma geçilmiştir. Aynı şekilde 64. gözlem düğüm noktası alınarak tek düğüm noktalı spline regresyon modeli kurulmuş ve R^2 değeri (%) hesaplanmıştır. 87. gözleme kadar bu şekilde devam edilerek toplam 24 tane doğrusal spline regresyon modeli kurulmuş ve R^2 değerleri (%) hesaplanmıştır. Hesaplanan R^2 değerleri (%) Tablo 5.6 ve Şekil 5.6'da gösterilmiştir. R^2 değerleri (%) incelendiğinde 74. gözlemden sonra R^2 değerinde düşüş eğilimi başladığı görülmektedir. Bu nedenle, R^2 değerinin değişim gösterdiği 74. gözlem altıncı düğüm noktası olarak düşünülmüştür.

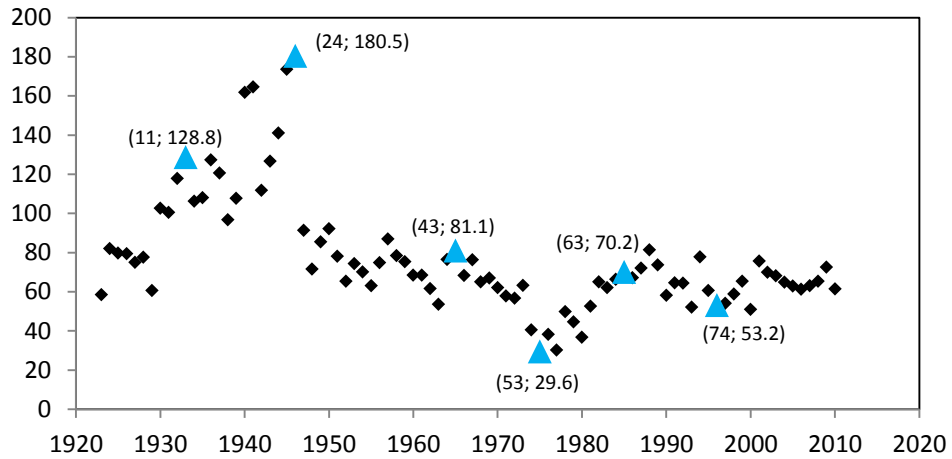
Tablo 5.6. Altıncı Adım Sonrası R^2 Değerleri

Gözlem No	R^2 (%)	Gözlem No	R^2 (%)	Gözlem No	R^2 (%)
64	3,8	72	18,2	80	7,7
65	3,8	73	20,3	81	6,8
66	4,5	74	21,7	82	6,5
67	8,2	75	21,1	83	6,4
68	13,6	76	18,7	84	6,2
69	16	77	15,8	85	5,4
70	17,9	78	13,5	86	4
71	19,5	79	9,5	87	3,3

**Şekil 5.6.** Altıncı Adım Sonrası R^2 Değerleri Dağılımı

5.2. BELİRLENEN DÜĞÜM NOKTALARINA GÖRE SPLİNE REGRESYON MODELLERİNİN KURULMASI

İlk olarak 5.1 başlığı altında doğrusal spline regresyon modelleri kullanılarak R^2 'yi temel alan yöntemle altı düğüm noktası belirlendi. Belirlenen düğüm noktalarının sayısı ve konumları Şekil 5.7'de gösterilmiştir.



Şekil 5.7. 1923-2010 Yılları Arası İhracatın İthalatı Karşılama Oranlarının Dağılımı ve Düşüm Noktaları

1923-2010 yılları arası Türkiye'nin ihracatın ithalatı karşılama oranlarının dağılımı Şekil 5.7 ve Ek 1'de gösterilmiştir. Veri setinde toplam 88 gözlem yer almaktadır. 1923 yılı 1. gözlem, 1924 yılı 2. gözlem, 1925 yılı 3. gözlem,..., 2009 yılı 87. gözlem ve 2010 yılı 88. gözlem olarak alınmıştır.

R^2 'yi temel alan yöntemle belirlenen düşüm noktaları 1933 (11. gözlem), 1946 (24. gözlem), 1965 (43. gözlem), 1975 (53. gözlem), 1985 (63. gözlem) ve 1996 (74. gözlem) yılıdır. Böylelikle, 1923-2010 yılları arasında toplam 6 düşüm noktası belirlenmiştir. Belirlenen düşüm noktalarına göre Şekil 5.7'deki dağılım 7 parçadan oluşmaktadır. Regresyon modelini kurmak üzere belirlenen düşüm noktalarının her biri için gölge değişken oluşturulur. Sonuç olarak 6 düşüm noktası için 6 gölge değişken belirlenecektir. Şekil 5.7'de 11. gözlem için D_{1t} , 24. gözlem için D_{2t} , 43. gözlem için D_{3t} , 53. gözlem için D_{4t} , 63. gözlem için D_{5t} , ve 74. gözlem için D_{6t} , gölge değişkenleri oluşturulur. İlk gölge değişken için $X_t \leq 11$ (1923) olduğu zaman $D_{1t} = 0$, $X_t > 11$ (1923) olduğu zaman $D_{1t} = 1$ alınır. İkinci gölge değişken için $X_t \leq 24$ (1946) olduğu zaman $D_{2t} = 0$, $X_t > 24$ (1946) olduğu zaman $D_{2t} = 1$ alınır. Üçüncü gölge değişken için $X_t \leq 43$ (1965) olduğu zaman $D_{3t} = 0$, $X_t > 43$ (1965) olduğu zaman $D_{3t} = 1$ alınır. Dördüncü gölge değişken için $X_t \leq 53$ (1975) olduğu zaman $D_{4t} = 0$, $X_t > 53$ (1975) olduğu zaman $D_{4t} = 1$ alınır. Beşinci gölge değişken için $X_t \leq 63$ (1985) olduğu zaman $D_{5t} = 0$, $X_t > 63$ (1985) olduğu zaman $D_{5t} = 1$ alınır. Altıncı gölge değişken için $X_t \leq 74$ (1996) olduğu zaman $D_{6t} = 0$, $X_t > 74$ (1996) olduğu zaman $D_{6t} = 1$ alınır. Böylece, gölge değişkenler kullanılarak

$Z_{1t} = D_{1t}(X_t - 11)$ gibi uygun spline değişkenleri oluşturulur³⁴⁵. Yani X_t onbirden küçük olduğu zaman $D_{1t} = 0$ olur ve böylece Z_{1t} asla negatif olmaz. Ayrıca $X_t = 11$ olduğu zaman Z_{1t} sifıra eşit olur. Bu şekilde $Z_{2t} = D_{2t}(X_t - 24)$, $Z_{3t} = D_{3t}(X_t - 43)$, $Z_{4t} = D_{4t}(X_t - 53)$, $Z_{5t} = D_{5t}(X_t - 63)$, ve $Z_{6t} = D_{6t}(X_t - 74)$ gibi spline değişkenleri oluşturulur. Ayrıca kurulacak doğrusal, kareli ve kübik spline regresyon modellerinde duruma göre $Z_{1t}^0 = D_{1t}(X_t - 11)^0$, $Z_{1t} = D_{1t}(X_t - 11)$, $Z_{1t}^2 = D_{1t}(X_t - 11)^2$ ve $Z_{1t}^3 = D_{1t}(X_t - 11)^3$ gibi değişkenlerde modele alınacaktır.

5.2.1. Doğrusal Spline Regresyon Modelleri

5.2.1.1. Kısıtlandırılmamış Doğrusal Spline Regresyon Modeli

Belirlenen düğüm noktalarına göre, oluşturulan gölge değişkenlerle kısıtlandırılmamış doğrusal spline regresyon modeli aşağıdaki gibi kurulur.³⁴⁶

$$Y_t = a_0 + b_0X_t + b_1Z_{1t}^0 + b_2Z_{1t} + \dots + b_{11}Z_{6t}^0 + b_{12}Z_{6t} + e_t \quad (5.1)$$

Oluşturulan regresyon modelinde, a_0 , b_1 , b_3 , b_5 , b_7 , b_9 ve b_{11} kısıtlandırılmamış regresyon modelinin sabitleri, b_0 , b_2 , b_4 , b_6 , b_8 , b_{10} ve b_{12} regresyon modelinin eğim katsayıları veya bağımsız değişkenlerin katsayı değerleri ve e_t ise tüm regresyon varsayımlarının sağlandığı varsayılan hata terimidir. Bu koşullar altında eşitlik 5.1'deki regresyon denklemi üçüncü bölümde anlatıldığı gibi en küçük karelerle tahmin edilebilir. Yapılan analiz sonucu elde edilen sonuçlar aşağıdaki gibidir.

Regresyon eşitliği,

$$\begin{aligned} \text{Tahmini İhr. /İth. Oranı (Y)} = & 2,937 + 0,02678X - 0,17414(X-11)^0 - 0,001131(X- \\ & 11) - 0,81084(X-24)^0 - 0,029612(X-24) + 0,1268(X-43)^0 - 0,015533(X-43) - \\ & 0,1213(X-53)^0 + 0,040049(X-53) + 0,0469(X-63)^0 - 0,027936(X-63) + 0,01487(X-74)^0 \\ & + 0,009964(X-74) + e_t \end{aligned} \quad (5.2)$$

biçiminde oluşturulur.

³⁴⁵ Bundan sonraki bölümlerde regresyon modelleri kurulurken yıllara karşılık gelen gözlem değerleri kullanılacaktır. Regresyon analizlerinde de bağımsız değişken serisi (1923-2010 arası) iki eşit kısma bölünmüş, üstteki değerler (-), alttaki değerler (+) alınmıştır. Yani 1.gözlem -87, 2. gözlem -85, 3. gözlem -83,..., 87. gözlem 85 ve 88. gözlem 87 olarak alınmıştır.

³⁴⁶ $Z_{1t}^0 = Z_{2t}^0 = \dots = Z_{6t}^0 = 1$

Tablo 5.7'deki varyans analizi sonuçlarına bakıldığında F istatistik değeri 36,59 bulunmuş ve kurulan modelin tamamı oldukça anlamlı çıkmıştır ($p < 0,05$). Modelin hata kareler ortalaması (MSE) 0,014 olarak bulunmuştur.

Tablo 5.7. Kısıtlandırılmamış Doğrusal Spline Regresyonda Varyans Analizi

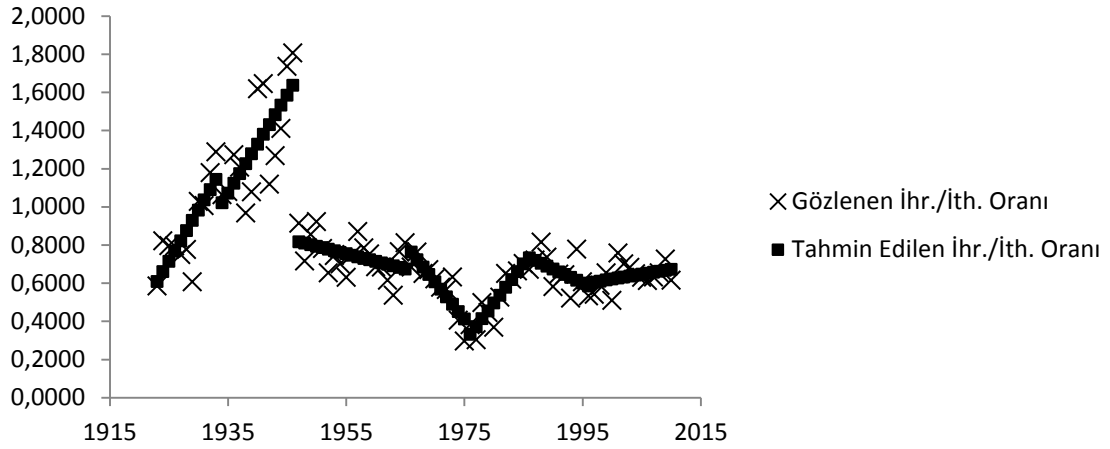
Değişim Kaynağı	Serbestlik Derecesi	Kareler Toplamı	Kareler Ortalaması	F	P
Regresyon	13	6,65877	0,51221	36,59	0,000
Hata	74	1,03587	0,01400		
Toplam	87	7,69464			

Yukarıda belirtilen 5.2 eşitliği Tablo 5.8'e göre yazılmıştır. Tablo 5.8 incelenecek olursa, a_0 , b_1 , b_3 , b_5 , b_7 , b_9 ve b_{11} katsayıları modelin sabitleri olduğundan, düğüm noktalarında regresyon doğrusu sürekli değildir. Regresyon modelinde a_0 , b_0 , b_3 , b_4 , b_6 , b_8 ve b_{10} katsayıları %5 önem düzeyinde anlamlı bulunarak, söz konusu katsayıların sıfırdan farklı olduğu belirlenmiştir. Ayrıca, b_1 , b_2 , b_3 , b_4 , b_6 , b_7 ve b_{10} katsayı değerleri ise negatif bulunmuştur.

Tablo 5.8. Kısıtlandırılmamış Doğrusal Spline Regresyon Modelinde Tahminlerin Çıktısı

Değişkenler	Katsayılar	Katsayı Değerleri	Standart Hata Değerleri	T	P	Güven Sınırları Alt Sınır	Üst Sınır
Modelin İlk Sabiti	a_0	2,937	0,4358	6,74	0,000	2,0687	3,8053
X	b_0	0,02678	0,00564	4,75	0,000	0,0155	0,038
(X-11) ⁰	b_1	-0,17414	0,09643	-1,81	0,075	-0,3663	0,018
(X-11)	b_2	-0,001131	0,007144	-0,16	0,875	-0,0154	0,0131
(X-24) ⁰	b_3	-0,81084	0,08389	-9,67	0,000	-0,978	-0,6437
(X-24)	b_4	-0,029612	0,005037	-5,88	0,000	-0,0396	-0,0196
(X-43) ⁰	b_5	0,1268	0,09622	1,32	0,192	-0,0649	0,3185
(X-43)	b_6	-0,015533	0,006968	-2,23	0,029	-0,0294	-0,0016
(X-53) ⁰	b_7	-0,1213	0,1066	-1,14	0,259	-0,3338	0,0911
(X-53)	b_8	0,040049	0,009211	4,35	0,000	0,0217	0,0584
(X-63) ⁰	b_9	0,0469	0,1034	0,45	0,652	-0,1591	0,2529
(X-63)	b_{10}	-0,027936	0,008616	-3,24	0,002	-0,0451	-0,0108
(X-74) ⁰	b_{11}	0,01487	0,09442	0,16	0,875	-0,1733	0,203
(X-74)	b_{12}	0,009964	0,00687	1,45	0,151	-0,0037	0,0237

Kısıtlandırılmamış doğrusal spline regresyon modeliyle bulunan tahmini ihr./ith. oranları ile gözlenen ihr./ith. oranları dağılımı Şekil 5.8’de verilmiştir.



Şekil 5.8. Kısıtlandırılmamış Doğrusal Spline Regresyonla Tahmin Edilen ve Gözlenen İhr./İth. Oranı Dağılımı

5.2.1.2. Kısıtlandırılmış Doğrusal Spline Regresyon Modeli

Belirlenen düğüm noktalarına göre oluşturulan gölge değişkenlerle kısıtlandırılmış doğrusal spline regresyon modeli aşağıdaki gibi kurulur.

$$Y_t = a_0 + b_0X_t + b_1Z_{1t} + \dots + b_5Z_{5t} + b_6Z_{6t} + e_t \quad (5.3)$$

Oluşturulan regresyon modelinde, a_0 regresyon modelinin sabiti, b_0 , b_1 , b_2 , b_3 , b_4 , b_5 ve b_6 regresyon modelinin eğim katsayıları veya bağımsız değişkenlerin katsayı değerleri ve e_t ise tüm regresyon varsayımlarının sağlandığı varsayılan hata terimidir. Bu koşullar altında eşitlik 5.3’teki regresyon denklemi en küçük karelerle tahmin edilebilir. Yapılan analiz sonucu elde edilen sonuçlar aşağıdaki gibidir.

Regresyon eşitliği,

$$\text{Tahmini İhr. /İth. Oranı (Y)} = 3,2127 + 0,030184X - 0,030973(X-11) - 0,014245(X-24) + 0,005722(X-43) + 0,02308(X-53) - 0,017455(X-63) + 0,005941(X-74) + e_t \quad (5.4)$$

biçiminde oluşturulur.

Tablo 5.9’daki varyans analizi sonuçlarına bakıldığında F istatistik değeri 22,77 bulunmuş ve kurulan modelin tamamı oldukça anlamlı çıkmıştır ($p < 0,05$). Modelin hata kareler ortalaması (MSE) 0,032 olarak bulunmuştur.

Tablo 5.9. Kısıtlandırılmış Doğrusal Spline Regresyonda Varyans Analizi

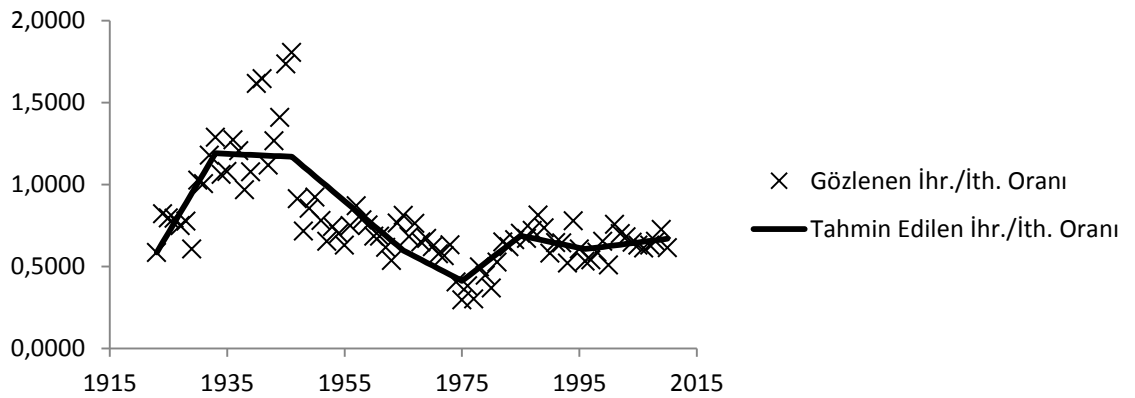
Değişim Kaynağı	Serbestlik Derecesi	Kareler Toplamı	Kareler Ortalaması	F	P
Regresyon	7	5,12324	0,73189	22,77	0,000
Hata	80	2,57139	0,03214		
Toplam	87	7,69464			

Yukarıda belirtilen 5.3 eşitliği Tablo 5.10'a göre yazılmıştır. Tablo 5.10 incelenecek olursa, a_0 modelin sabitidir. Modelin tek bir sabiti olduğu için, düğüm noktalarında regresyon doğrusu süreklidir. Regresyon modelinde a_0 , b_0 , b_1 , b_2 ve b_4 katsayıları %5 önem düzeyinde anlamlı bulunarak, söz konusu katsayıların sıfırdan farklı olduğu belirlenmiştir. Kısıtlandırılmış doğrusal spline regresyonda 1., 2. ve 4. düğüm noktaları anlamlı bulunmuştur. Ayrıca, b_1 , b_2 ve b_5 katsayı değerleri ise negatif bulunmuştur.

Tablo 5.10. Kısıtlandırılmış Doğrusal Spline Regresyon Modelinde Tahminlerin Çıktısı

Değişkenler	Katsayılar	Katsayı Değerleri	Standart Hata Değerleri	T	P	Güven Sınırları Alt Sınır Üst Sınır
Modelin Sabiti	a_0	3,2127	0,509	6,31	0,000	2,1999 4,2257
X	b_0	0,030184	0,006788	4,45	0,000	0,0167 0,0437
(X-11)	b_1	-0,030973	0,009757	-3,17	0,002	-0,0504 -0,0116
(X-24)	b_2	-0,014245	0,005961	-2,39	0,019	-0,0261 -0,0024
(X-43)	b_3	0,005722	0,007222	0,79	0,43	-0,0087 0,0201
(X-53)	b_4	0,02308	0,01008	2,29	0,025	0,003 0,0431
(X-63)	b_5	-0,017455	0,009821	-1,78	0,079	-0,037 0,0021
(X-74)	b_6	0,005941	0,008642	0,69	0,494	-0,0113 0,0231

Kısıtlandırılmış doğrusal spline regresyon modeliyle bulunan tahmini ihr./ith. oranları ile gözlenen ihr./ith. oranları dağılımı Şekil 5.9'da verilmiştir.

**Şekil 5.9.** Kısıtlandırılmış Doğrusal Spline Regresyonla Tahmin Edilen ve Gözlenen İhr./İth. Oranı Dağılımı

5.2.2. Kareli (Kuadratik) Spline Regresyon Modelleri

5.2.2.1. Kısıtlandırılmamış Kareli Spline Regresyon Modeli

Belirlenen düğüm noktalarına göre oluşturulan gölge değişkenlerle aşağıdaki gibi kısıtlandırılmamış kareli spline regresyon modeli kurulur.

$$Y_t = a_0 + b_0X_t + b_1X_t^2 + b_2Z_{1t}^0 + b_3Z_{1t} + b_4Z_{1t}^2 + \dots + b_{17}Z_{6t}^0 + b_{18}Z_{6t} + b_{19}Z_{6t}^2 + e_t \quad (5.5)$$

Oluşturulan regresyon modelinde, a_0 , b_2 , b_5 , b_8 , b_{11} , b_{14} ve b_{17} kısıtlandırılmamış regresyon modelinin sabitleri, b_0 , b_1 , b_3 , b_4 , b_6 , $b_7, \dots, b_{18}$ ve b_{19} regresyon modelinin eğim katsayıları veya bağımsız değişkenlerin katsayı değerleri ve e_t ise tüm regresyon varsayımlarının sağlandığı varsayılan hata terimidir. Bu koşullar altında eşitlik 5.5'teki regresyon denklemi üçüncü bölümde anlatıldığı gibi en küçük karelerle tahmin edilebilir. Yapılan analiz sonucu elde edilen sonuçlar aşağıdaki gibidir.

Regresyon eşitliği,

$$\begin{aligned} \text{Tahmini İhr. /İth. Oranı (Y)} = & 15,624 + 0,3586X + 0,0021544X^2 - 0,1609(X-11)^0 \\ & - 0,07271(X-11) - 0,001137(X-11)^2 - 0,8394(X-24)^0 - 0,06275(X-24) - 0,0007998(X- \\ & 24)^2 - 0,0491(X-43)^0 + 0,00949(X-43) - 0,001712(X-43)^2 - 0,012(X-53)^0 - 0,05754(X- \\ & 53) + 0,001921(X-53)^2 + 0,0081(X-63)^0 - 0,0314(X-63) - 0,000603(X-63)^2 - \\ & 0,0363(X-74)^0 + 0,02506(X-74) - 0,00021(X-74)^2 + e_t \end{aligned} \quad (5.6)$$

biçiminde oluşturulur.

Tablo 5.7'deki varyans analizi sonuçlarına bakıldığında F istatistik değeri 36,59 bulunmuş ve kurulan modelin tamamı oldukça anlamlı çıkmıştır ($p < 0,05$). Modelin hata kareler ortalaması (MSE) 0,014 olarak bulunmuştur.

Tablo 5.11. Kısıtlandırılmamış Kareli Spline Regresyonda Varyans Analizi

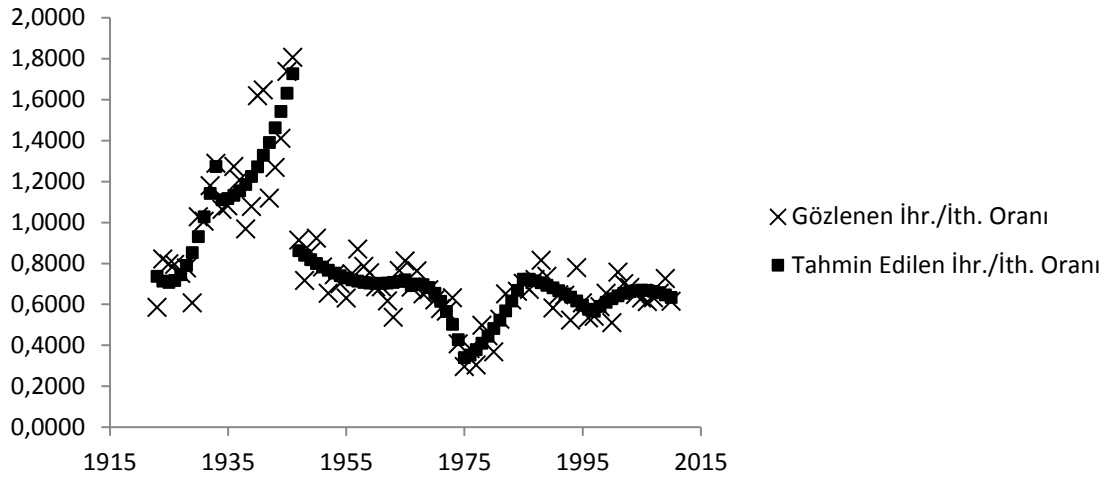
Değişim Kaynağı	Serbestlik Derecesi	Kareler Toplamı	Kareler Ortalaması	F	P
Regresyon	20	6,79373	0,33969	25,26	0,000
Hata	67	0,90091	0,01345		
Toplam	87	7,69464			

Yukarıda belirtilen 5.6 eşitliği Tablo 5.12'ye göre yazılmıştır. Tablo 5.12 incelenecek olursa, a_0 , b_2 , b_5 , b_8 , b_{11} , b_{14} ve b_{17} katsayıları modelin sabitleri olduğundan, düğüm noktalarında regresyon doğrusu sürekli değildir. Regresyon modelinde a_0 , b_0 , b_1 , b_3 , b_5 ve b_6 katsayıları %5 önem düzeyinde anlamlı bulunarak, söz konusu katsayıların sıfırdan farklı olduğu belirlenmiştir. Ayrıca, b_2 , b_3 , b_4 , b_5 , b_6 , b_7 , b_8 , b_{10} , b_{11} , b_{15} , b_{16} , b_{17} ve b_{19} katsayı değerleri ise negatif bulunmuştur.

Tablo 5.12. Kısıtlandırılmamış Kareli Spline Regresyon Modelinde Tahminlerin Çıktısı

Değişkenler	Katsayılar	Katsayı Değerleri	Standart Hata Değerleri	T	P	Güven Sınırları Alt Sınır	Üst Sınır
Modelin İlk Sabiti	a_0	15,624	5,844	2,67	0,009	3,9617	27,2895
X	b_0	0,3586	0,1525	2,35	0,022	0,0542	0,663
X^2	b_1	0,0021544	0,0009897	2,18	0,033	0,0002	0,0041
$(X-11)^0$	b_2	-0,1609	0,1438	-1,12	0,267	-0,448	0,1262
$(X-11)$	b_3	-0,07271	0,02775	-2,62	0,011	-0,1281	-0,0173
$(X-11)^2$	b_4	-0,001137	0,001183	-0,96	0,34	-0,0035	0,0012
$(X-24)^0$	b_5	-0,8394	0,1219	-6,88	0	-1,0827	-0,596
$(X-24)$	b_6	-0,06275	0,01911	-3,28	0,002	-0,1009	-0,0246
$(X-24)^2$	b_7	-0,0007998	0,0006941	-1,15	0,253	-0,0022	0,0006
$(X-43)^0$	b_8	-0,0491	0,1543	-0,32	0,751	-0,357	0,2588
$(X-43)$	b_9	0,00949	0,02995	0,32	0,752	-0,0503	0,0693
$(X-43)^2$	b_{10}	-0,001712	0,001286	-1,33	0,188	-0,0043	0,0009
$(X-53)^0$	b_{11}	-0,012	0,1641	-0,07	0,942	-0,3394	0,3155
$(X-53)$	b_{12}	0,05754	0,03698	1,56	0,124	-0,0163	0,1313
$(X-53)^2$	b_{13}	0,001921	0,001784	1,08	0,285	-0,0016	0,0055
$(X-63)^0$	b_{14}	0,0081	0,1566	0,05	0,959	-0,3046	0,3206
$(X-63)$	b_{15}	-0,0314	0,03393	-0,93	0,358	-0,0991	0,0363
$(X-63)^2$	b_{16}	-0,000603	0,001603	-0,38	0,708	-0,0038	0,0026
$(X-74)^0$	b_{17}	-0,0363	0,1396	-0,26	0,796	-0,3148	0,2423
$(X-74)$	b_{18}	0,02506	0,0264	0,95	0,346	-0,0276	0,0778
$(X-74)^2$	b_{19}	-0,00021	0,001126	-0,19	0,853	-0,0025	0,002

Kısıtlandırılmamış kareli spline regresyon modeliyle bulunan tahmini ihr./ith. oranları ile gözlenen ihr./ith. oranları dağılımı Şekil 5.10'da verilmiştir.



Şekil 5.10. Kısıtlandırılmamış Kareli Spline Regresyonla Tahmin Edilen ve Gözlenen İhr./İth. Oranı Dağılımı

5.2.2.2. Kısıtlandırılmış Kareli (Kuadratik) Spline Regresyon Modeli

5.2.2.2.1. Sabit Katsayıları Kısıtlandırılmış Kareli Spline Regresyon Modeli

Belirlenen düğüm noktalarına göre oluşturulan gölge değişkenlerle aşağıdaki gibi sabit katsayıları kısıtlandırılmış kareli spline regresyon modeli kurulur.

$$Y_t = a_0 + b_0X_t + b_1X_t^2 + b_2Z_{1t} + b_3Z_{1t}^2 + \dots + b_{12}Z_{6t} + b_{13}Z_{6t}^2 + e_t \quad (5.7)$$

Oluşturulan regresyon modelinde, a_0 regresyon modelinin sabiti, b_0 , b_1 , b_2 , b_3 , b_4 , $b_5, \dots$, b_{12} ve b_{13} regresyon modelinin eğim katsayıları veya bağımsız değişkenlerin katsayı değerleri ve e_t ise tüm regresyon varsayımlarının sağlandığı varsayılan hata terimidir. Bu koşullar altında eşitlik 5.7'deki regresyon denklemi üçüncü bölümde anlatıldığı gibi en küçük karelerle tahmin edilebilir. Yapılan analiz sonucu elde edilen sonuçlar aşağıdaki gibidir.

Regresyon eşitliği,

$$\begin{aligned} \text{Tahmini İhr. /İth. Oranı (Y)} = & 9,34 + 0,2X + 0,001159X^2 - 0,03023(X-11) - \\ & 0,001402(X-11)^2 - 0,06105(X-24) + 0,001423(X-24)^2 - 0,04244(X-43) - 0,0017072(X- \\ & 43)^2 + 0,03874(X-53) + 0,001168(X-53)^2 - 0,03295(X-63) - 0,000905(X-63)^2 + \\ & 0,02444(X-74) - 0,0000328(X-74)^2 + e_t \end{aligned} \quad (5.8)$$

biçiminde oluşturulur.

Tablo 5.13'teki varyans analizi sonuçlarına bakıldığında F istatistik değeri 19,35 bulunmuş ve kurulan modelin tamamı oldukça anlamlı çıkmıştır ($p < 0,05$). Modelin hata kareler ortalaması (MSE) 0,023 olarak bulunmuştur.

Tablo 5.13. Sabit Katsayıları Kısıtlandırılmış Kareli Spline Regresyonda Varyans Analizi

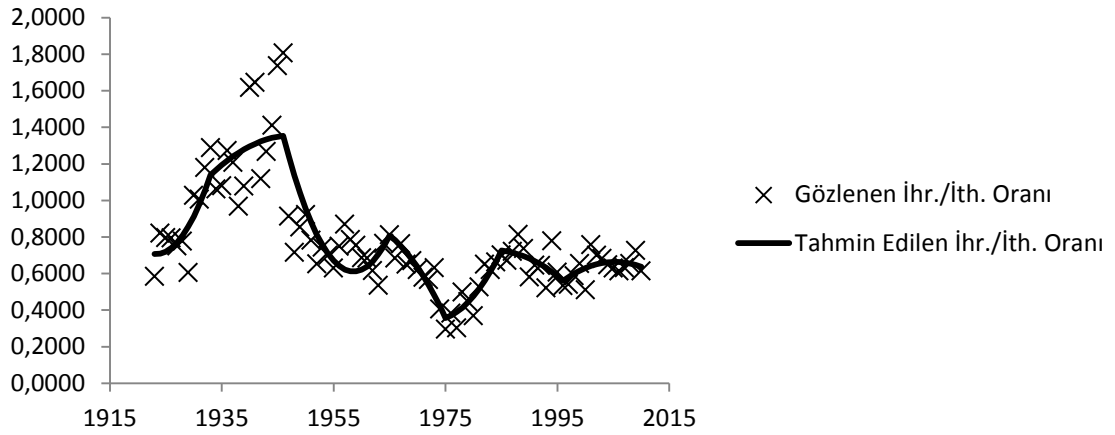
Değişim Kaynağı	Serbestlik Derecesi	Kareler Toplamı	Kareler Ortalaması	F	P
Regresyon	14	6,06109	0,43293	19,35	0,000
Hata	73	1,63355	0,02283		
Toplam	87	7,69464			

Yukarıda belirtilen 5.8 eşitliği Tablo 5.14'e göre yazılmıştır. Tablo 5.14 incelenecek olursa, a_0 modelin sabitidir. Modelin tek bir sabiti olduğu için düğüm noktalarında regresyon doğrusu süreklidir. Regresyon modelinde b_4 ve b_5 katsayıları %5 önem düzeyinde anlamlı bulunarak, söz konusu katsayıların sıfırdan farklı olduğu belirlenmiştir. Ayrıca, b_2 , b_3 , b_4 , b_6 , b_7 , b_{10} , b_{11} ve b_{13} katsayı değerleri ise negatif bulunmuştur.

Tablo 5.14. Sabit Katsayıları Kısıtlandırılmış Kareli Spline Regresyon Modelinde Tahminlerin Çıktısı

Değişkenler	Katsayılar	Katsayı Değerleri	Standart Hata Değerleri	T	P	Güven Sınırları Alt Sınır	Güven Sınırları Üst Sınır
Modelin Sabiti	a_0	9,34	6,739	1,39	0,17	-4,0874	22,7716
X	b_0	0,2	0,1773	1,13	0,263	-0,1533	0,5534
X^2	b_1	0,001159	0,001159	1	0,321	-0,0012	0,0035
(X-11)	b_2	-0,03023	0,03466	-0,87	0,386	-0,0993	0,0388
$(X-11)^2$	b_3	-0,001402	0,001083	-1,29	0,2	-0,0036	0,0008
(X-24)	b_4	-0,06105	0,02206	-2,77	0,007	-0,105	-0,0171
$(X-24)^2$	b_5	0,001423	0,0005314	2,68	0,009	0,0004	0,0025
(X-43)	b_6	-0,04244	0,02769	-1,53	0,13	-0,0976	0,0127
$(X-43)^2$	b_7	-0,0017072	0,0009761	-1,75	0,084	-0,0037	0,0002
(X-53)	b_8	0,03874	0,04013	0,97	0,338	-0,0412	0,1187
$(X-53)^2$	b_9	0,001168	0,001184	0,99	0,327	-0,0012	0,0035
(X-63)	b_{10}	-0,03295	0,03793	-0,87	0,388	-0,1086	0,0426
$(X-63)^2$	b_{11}	-0,000905	0,001104	-0,82	0,415	-0,0031	0,0013
(X-74)	b_{12}	0,02444	0,03086	0,79	0,431	-0,0371	0,086
$(X-74)^2$	b_{13}	-0,0000328	0,0008283	-0,04	0,969	-0,0017	0,0016

Sabit katsayıları kısıtlandırılmış kareli spline regresyon modeliyle bulunan tahmini ihr./ith. oranları ile gözlenen ihr./ith. oranları dağılımı Şekil 5.11'de verilmiştir.



Şekil 5.11. Sabit Katsayıları Kısıtlandırılmış Kareli Spline Regresyonla Tahmin Edilen ve Gözlenen İhr./İth. Oranı Dağılımı

5.2.2.2.2. Sabit Katsayıları Kısıtlandırılmış ve Düğüm Noktasında Birinci Derece Türevleri Eşit Olan Kareli Spline Regresyon Modeli

Belirlenen düğüm noktalarına göre oluşturulan gölge değişkenlerle aşağıdaki gibi sabit katsayıları kısıtlandırılmış ve düğüm noktasında birinci derece türevleri eşit olan kareli spline regresyon modeli kurulur.

$$Y_t = a_0 + b_0X_t + b_1X_t^2 + b_2Z_{1t}^2 + b_3Z_{2t}^2 + \dots + b_6Z_{5t}^2 + b_7Z_{6t}^2 + e_t \quad (5.9)$$

Oluşturulan regresyon modelinde, a_0 regresyon modelinin sabiti, b_0 , b_1 , b_2 , b_3 , b_4 , b_5 , b_6 ve b_7 regresyon modelinin eğim katsayıları veya bağımsız değişkenlerin katsayı değerleri ve e_t ise tüm regresyon varsayımlarının sağlandığı varsayılan hata terimidir. Bu koşullar altında eşitlik 5.9'daki regresyon denklemi üçüncü bölümde anlatıldığı gibi en küçük karelerle tahmin edilebilir. Yapılan analiz sonucu elde edilen sonuçlar aşağıdaki gibidir.

Regresyon eşitliği,

$$\begin{aligned} \text{Tahmini İhr. /İth. Oranı (Y)} = & 7,783 + 0,162X + 0,0009271X^2 - 0,0022534(X- \\ & 11)^2 + 0,0017761(X-24)^2 - 0,0008791(X-43)^2 + 0,0012566 (X-53)^2 - 0,0015021(X-63)^2 \\ & + 0,0011138(X-74)^2 + e_t \end{aligned} \quad (5.10)$$

biçiminde oluşturulur.

Tablo 5.15'teki varyans analizi sonuçlarına bakıldığında F istatistik değeri 26,31 bulunmuş ve kurulan modelin tamamı oldukça anlamlı çıkmıştır ($p < 0,05$). Modelin hata kareler ortalaması (MSE) 0,027 olarak bulunmuştur.

Tablo 5.15. Sabit Katsayıları Kısıtlandırılmış ve Düğüm Noktasında Birinci Derece Türevleri Eşit Olan Kareli Spline Regresyonda Varyans Analizi

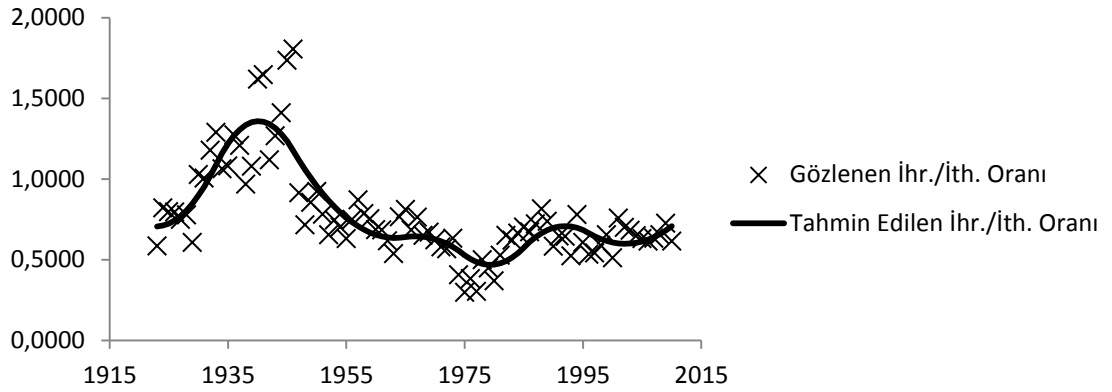
Değişim Kaynağı	Serbestlik Derecesi	Kareler Toplamı	Kareler Ortalaması	F	P
Regresyon	8	5,59457	0,69932	26,31	0,000
Hata	79	2,10007	0,02658		
Toplam	87	7,69464			

Yukarıda belirtilen 5.10 eşitliği Tablo 5.16'ya göre yazılmıştır. Tablo 5.16 incelenecek olursa, a_0 modelin sabitidir. Modelin tek bir sabiti olduğu için, düğüm noktalarında regresyon doğrusu süreklidir. Regresyon modelinde a_0 , b_2 , b_3 , b_4 , b_5 ve b_6 katsayıları %5 önem düzeyinde anlamlı bulunarak, söz konusu katsayıların sıfırdan farklı olduğu belirlenmiştir. Sabit katsayıları kısıtlandırılmış ve düğüm noktasında birinci derece türevleri eşit olan kareli spline regresyonda 1., 2., 3., 4. ve 5. düğüm noktaları anlamlı bulunmuştur. Ayrıca, b_2 , b_4 ve b_6 katsayı değerleri ise negatif bulunmuştur.

Tablo 5.16. Sabit Katsayıları Kısıtlandırılmış ve Düğüm Noktasında Birinci Derece Türevleri Eşit Olan Kareli Spline Regresyon Modelinde Tahminlerin Çıktısı

Değişkenler	Katsayılar	Katsayı Değerleri	Standart Hata Değerleri	T	P	Güven Sınırları Alt Sınır Üst Sınır
Modelin Sabiti	a_0	7,783	3,035	2,56	0,012	1,7416 13,8242
X	b_0	0,162	0,08542	1,9	0,062	-0,0003 0,0021
X^2	b_1	0,0009271	0,0005925	1,56	0,122	-0,0038 -0,0007
$(X-11)^2$	b_2	-0,0022534	0,0007586	-2,97	0,004	0,0012 0,0024
$(X-24)^2$	b_3	0,0017761	0,000297	5,98	0,000	-0,0016 -0,0002
$(X-43)^2$	b_4	-0,0008791	0,0003551	-2,48	0,015	0,0002 0,0024
$(X-53)^2$	b_5	0,0012566	0,0005498	2,29	0,025	-0,0026 -0,0004
$(X-63)^2$	b_6	-0,0015021	0,0005755	-2,61	0,011	-0,0001 0,0023
$(X-74)^2$	b_7	0,0011138	0,0005929	1,88	0,064	-0,008 0,332

Sabit katsayıları kısıtlandırılmış ve düğüm noktasında birinci derece türevleri eşit olan kareli spline regresyon modeliyle bulunan tahmini ihr./ith. oranları ile gözlenen ihr./ith. oranları dağılımı Şekil 5.12'de verilmiştir.



Şekil 5.12. Sabit Katsayıları Kısıtlandırılmış ve Düğüm Noktasında Birinci Derece Türevleri Eşit Olan Kareli Spline Regresyonla Tahmin Edilen ve Gözlenen İhr./İth. Oranı Dağılımı

5.2.3. Kübik Spline Regresyon Modeli

5.2.3.1. Kısıtlandırılmamış Kübik Spline Regresyon Modeli

Belirlenen düğüm noktalarına göre oluşturulan gölge değişkenlerle aşağıdaki gibi kısıtlandırılmamış kübik spline regresyon modeli kurulur.

$$Y_t = a_0 + b_0X_t + b_1X_t^2 + b_2X_t^3 + b_3Z_{1t}^0 + b_4Z_{1t} + b_5Z_{1t}^2 + b_6Z_{1t}^3 + \dots + b_{23}Z_{6t}^0 + b_{24}Z_{6t} + b_{25}Z_{6t}^2 + b_{26}Z_{6t}^3 + e_t \quad (5.11)$$

Oluşturulan regresyon modelinde, a_0 , b_3 , b_7 , b_{11} , b_{15} , b_{19} ve b_{23} kısıtlandırılmamış regresyon modelinin sabitleri, b_0 , b_1 , b_2 , b_4 , b_5 , $b_6, \dots, b_{24}$, b_{25} ve b_{26} regresyon modelinin eğim katsayıları veya bağımsız değişkenlerin katsayı değerleri ve e_t ise tüm regresyon varsayımlarının sağlandığı varsayılan hata terimidir. Bu koşullar altında eşitlik 5.11'deki regresyon denklemi üçüncü bölümde anlatıldığı gibi en küçük karelerle tahmin edilebilir. Yapılan analiz sonucu elde edilen sonuçlar aşağıdaki gibidir.

Regresyon eşitliği,

$$\begin{aligned} \text{Tahmini İhr. /İth. Oranı (Y)} = & 136,47 + 5,105X + 0,06405X^2 + 0,0002679X^3 - \\ & 0,4118(X-11)^0 - 0,07094(X-11) - 0,014603(X-11)^2 - 0,0001387(X-11)^3 - 0,9218(X- \\ & 24)^0 - 0,1019(X-24) - 0,005682(X-24)^2 - 0,0001254(X-24)^3 + 0,058(X-43)^0 - \\ & 0,04422(X-43) + 0,003591(X-43)^2 - 0,0001708(X-43)^3 + 0,0909(X-53)^0 + 0,0577(X- \\ & 53) + 0,00975(X-53)^2 + 0,0000662(X-53)^3 - 0,0003(X-63)^0 - 0,00096(X-63) + \end{aligned}$$

$$0,00093(X-63)^2 + 0,0001336(X-63)^3 - 0,1134(X-74)^0 + 0,0469(X-74) - 0,00306(X-74)^2 + 0,0000086(X-74)^3 + e_t \quad (5.12)$$

biçiminde oluşturulur.

Tablo 5.17'deki varyans analizi sonuçlarına bakıldığında F istatistik değeri 18,17 bulunmuş ve kurulan modelin tamamı oldukça anlamlı çıkmıştır ($p < 0,05$). Modelin hata kareler ortalaması (MSE) 0,014 olarak bulunmuştur.

Tablo 5.17. Kısıtlandırılmamış Kübik Spline Regresyonda Varyans Analizi

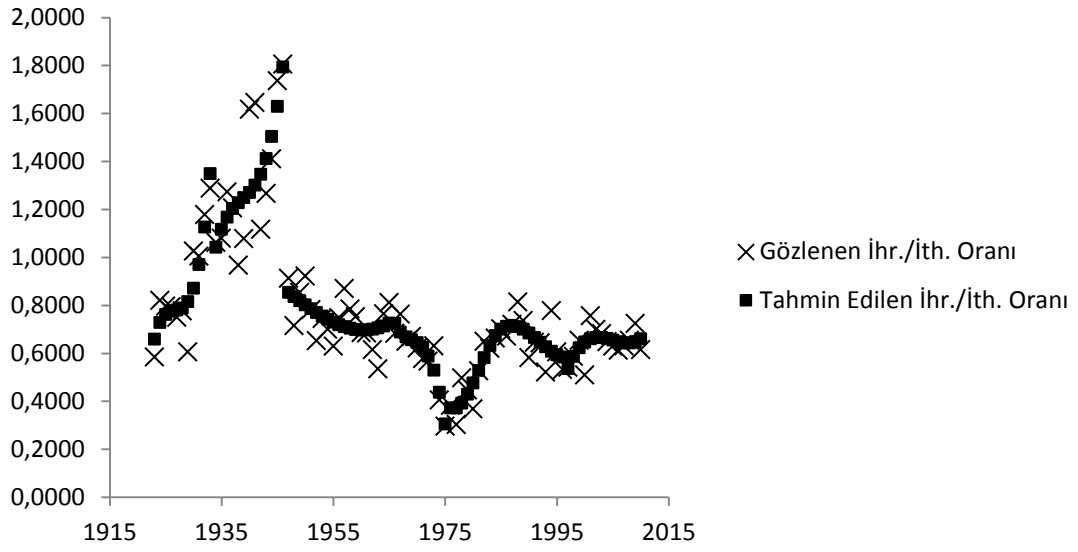
Değişim Kaynağı	Serbestlik Derecesi	Kareler Toplamı	Kareler Ortalaması	F	P
Regresyon	27	6,85620	0,25393	18,17	0,000
Hata	60	0,83844	0,01397		
Toplam	87	7,69464			

Yukarıda belirtilen 5.12 eşitliği Tablo 5.18'e göre yazılmıştır. Tablo 5.18 incelenecek olursa, a_0 , b_3 , b_7 , b_{11} , b_{15} , b_{19} ve b_{23} katsayıları modelin sabitleri olduğundan, düğüm noktalarında regresyon doğrusu sürekli değildir. Regresyon modelinde b_5 , b_7 ve b_8 katsayıları %5 önem düzeyinde anlamlı bulunarak, söz konusu katsayıların sıfırdan farklı olduğu belirlenmiştir. Ayrıca, b_3 , b_4 , b_5 , b_6 , b_7 , b_8 , b_9 , b_{10} , b_{12} , b_{14} , b_{19} , b_{20} , b_{23} ve b_{25} katsayı değerleri ise negatif bulunmuştur.

Tablo 5.18. Kısıtlandırılmamış Kübik Spline Regresyon Modelinde Tahminlerin Çıktısı

Değişkenler	Katsayılar	Katsayı Değerleri	Standart Hata Değerleri	T	P	Güven Sınırları	
						Alt Sınır	Üst Sınır
Modelin İlk Sabiti	a_0	136,47	85,01	1,61	0,114	-33,605	306,4592
X	b_0	5,105	3,334	1,53	0,131	-1,5659	11,7724
X^2	b_1	0,06405	0,04344	1,47	0,146	-0,0229	0,1509
X^3	b_2	0,0002679	0,000188	1,43	0,159	-0,0001	0,0006
$(X-11)^0$	b_3	-0,4118	0,2088	-1,97	0,053	-0,8293	0,0058
$(X-11)$	b_4	-0,07094	0,07194	-0,99	0,328	-0,2148	0,073
$(X-11)^2$	b_5	-0,014603	0,007209	-2,03	0,047	-0,029	-0,0002
$(X-11)^3$	b_6	-0,0001387	0,0002144	-0,65	0,52	-0,0006	0,0003
$(X-24)^0$	b_7	-0,9218	0,1677	-5,5	0,00	-1,2572	-0,5863
$(X-24)$	b_8	-0,1019	0,04731	-2,15	0,035	-0,1965	-0,0073
$(X-24)^2$	b_9	-0,005682	0,004099	-1,39	0,171	-0,0139	0,0025
$(X-24)^3$	b_{10}	-0,0001254	0,0001064	-1,18	0,243	-0,0003	0,0001
$(X-43)^0$	b_{11}	0,058	0,2465	0,24	0,815	-0,4351	0,551
$(X-43)$	b_{12}	-0,04422	0,08882	-0,5	0,62	-0,2219	0,1335
$(X-43)^2$	b_{13}	0,003591	0,008987	0,4	0,691	-0,0144	0,0216
$(X-43)^3$	b_{14}	-0,0001708	0,0002672	-0,64	0,525	-0,0007	0,0004
$(X-53)^0$	b_{15}	0,0909	0,2533	0,36	0,721	-0,4157	0,5975
$(X-53)$	b_{16}	0,0577	0,1019	0,57	0,574	-0,1461	0,2614
$(X-53)^2$	b_{17}	0,00975	0,01148	0,85	0,399	-0,0132	0,0327
$(X-53)^3$	b_{18}	0,0000662	0,000376	0,18	0,861	-0,0007	0,0008
$(X-63)^0$	b_{19}	-0,0003	0,2352	-0,01	0,999	-0,4709	0,4702
$(X-63)$	b_{20}	-0,00096	0,09056	-0,01	0,992	-0,1821	0,1802
$(X-63)^2$	b_{21}	0,00093	0,01	0,09	0,926	-0,0191	0,0209
$(X-63)^3$	b_{22}	0,0001336	0,0003256	0,41	0,683	-0,0005	0,0008
$(X-74)^0$	b_{23}	-0,1134	0,1995	-0,57	0,572	-0,5123	0,2856
$(X-74)$	b_{24}	0,04069	0,06725	0,61	0,547	-0,0938	0,1752
$(X-74)^2$	b_{25}	-0,00306	0,006764	-0,45	0,653	-0,0166	0,0105
$(X-74)^3$	b_{26}	0,0000086	0,0002039	0,04	0,967	-0,0004	0,0004

Kısıtlandırılmamış kübik spline regresyon modeliyle bulunan tahmini ihr./ith. oranları ile gözlenen ihr./ith. oranları dağılımı Şekil 5.13'te verilmiştir.



Şekil 5.13. Kısıtlandırılmamış Kübik Spline Regresyonla Tahmin Edilen ve Gözlenen İhr./İth. Oranı Dağılımı

5.2.3.2. Kısıtlandırılmış Kübik Spline Regresyon Modeli

5.2.3.2.1. Sabit Katsayıları Kısıtlandırılmış Kübik Spline Regresyon Modeli

Belirlenen düğüm noktalarına göre oluşturulan gölge değişkenlerle aşağıdaki gibi sabit katsayıları kısıtlandırılmış kübik spline regresyon modeli kurulur.

$$Y_t = a_0 + b_0X_t + b_1X_t^2 + b_2X_t^3 + b_3Z_{1t} + b_4Z_{1t}^2 + b_5Z_{1t}^3 + \dots + b_{18}Z_{6t} + b_{19}Z_{6t}^2 + b_{20}Z_{6t}^3 + e_t \quad (5.13)$$

Oluşturulan regresyon modelinde, a_0 regresyon modelinin sabiti, $b_0, b_1, b_2, b_3, b_4, b_5, \dots, b_{19}$ ve b_{20} regresyon modelinin eğim katsayıları veya bağımsız değişkenlerin katsayı değerleri ve e_t ise tüm regresyon varsayımlarının sağlandığı varsayılan hata terimidir. Bu koşullar altında eşitlik 5.13'teki regresyon denklemi üçüncü bölümde anlatıldığı gibi en küçük karelerle tahmin edilebilir. Yapılan analiz sonucu elde edilen sonuçlar aşağıdaki gibidir.

Regresyon eşitliği,

$$\begin{aligned} \text{Tahmini İhr. /İth. Oranı (Y)} = & 105,2 + 3,911X + 0,04888X^2 + 0,000204X^3 - \\ & 0,1633(X-11) - 0,001845(X-11)^2 - 0,0003438(X-11)^3 - 0,09093(X-24) + 0,010319(X- \\ & 24)^2 + 0,0000627(X-24)^3 + 0,04651(X-43) + 0,002699(X-43)^2 + 0,0000435(X-43)^3 + \\ & 0,05852(X-53) + 0,004266(X-53)^2 - 0,0000072(X-53)^3 - 0,015(X-63) + 0,000188(X- \\ & 63)^2 + 0,0000487(X-63)^3 + 0,02995(X-74) - 0,000671(X-74)^2 + 0,0000043(X-74)^3 + e_t \end{aligned} \quad (5.14)$$

biçiminde oluşturulur.

Tablo 5.19'daki varyans analizi sonuçlarına bakıldığında F istatistik değeri 15,52 bulunmuş ve kurulan modelin tamamı oldukça anlamlı çıkmıştır ($p < 0,05$). Modelin hata kareler ortalaması (MSE) 0,02 olarak bulunmuştur.

Tablo 5.19. Sabit Katsayıları Kısıtlandırılmış Kübik Spline Regresyonda Varyans Analizi

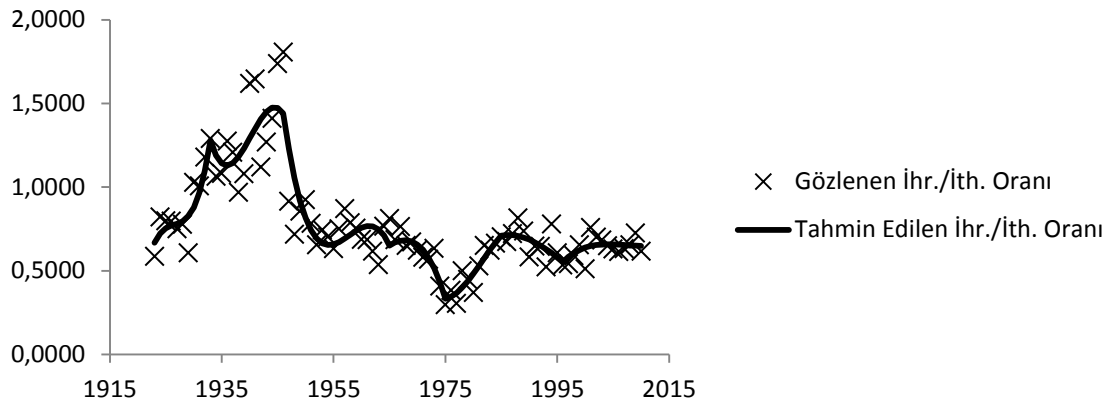
Değişim Kaynağı	Serbestlik Derecesi	Kareler Toplamı	Kareler Ortalaması	F	P
Regresyon	21	6,39902	0,30472	15,52	0,000
Hata	66	1,29562	0,01963		
Toplam	87	7,69464			

Yukarıda belirtilen 5.14 eşitliği Tablo 5.20'ye göre yazılmıştır. Tablo 5.20 incelenecek olursa, a_0 modelin sabitidir. Modelin tek bir sabiti olduğu için düğüm noktalarında regresyon doğrusu süreklidir. Regresyon modelinde sadece b_7 katsayısı %5 önem düzeyinde anlamlı bulunarak, söz konusu katsayının sıfırdan farklı olduğu belirlenmiştir. Ayrıca, b_3 , b_4 , b_5 , b_6 , b_{14} , b_{15} , b_{19} ve b_{20} katsayı değerleri ise negatif bulunmuştur.

Tablo 5.20. Sabit Katsayıları Kısıtlandırılmış Kübik Spline Regresyon Modelinde Tahminlerin Çıktısı

Değişkenler	Katsayılar	Katsayı Değerleri	Standart Hata Değerleri	T	P	Güven Sınırları Alt Sınır	Üst Sınır
Modelin Sabiti	a_0	105,2	96,65	1,09	0,28	-87,804	298,1041
X	b_0	3,911	3,799	1,03	0,307	-3,6764	11,4941
X^2	b_1	0,04888	0,0496	0,99	0,328	-0,0502	0,1479
X^3	b_2	0,000204	0,0002151	0,95	0,346	-0,0002	0,0006
$(X-11)$	b_3	-0,1633	0,07966	-2,05	0,044	-0,3223	-0,0042
$(X-11)^2$	b_4	-0,001845	0,006213	-0,3	0,767	-0,0142	0,0106
$(X-11)^3$	b_5	-0,0003438	0,0002503	-1,37	0,174	-0,0008	0,0002
$(X-24)$	b_6	-0,09093	0,0519	-1,75	0,084	-0,1946	0,0127
$(X-24)^2$	b_7	0,010319	0,003164	3,26	0,002	0,004	0,0166
$(X-24)^3$	b_8	0,0000627	0,00009575	0,65	0,515	-0,0001	0,0003
$(X-43)$	b_9	0,04651	0,06519	0,71	0,478	-0,0836	0,1767
$(X-43)^2$	b_{10}	0,002699	0,005869	0,46	0,647	-0,009	0,0144
$(X-43)^3$	b_{11}	0,0000435	0,0002134	0,2	0,839	-0,0004	0,0005
$(X-53)$	b_{12}	0,05852	0,0933	0,63	0,533	-0,1278	0,2448
$(X-53)^2$	b_{13}	0,004266	0,00741	0,58	0,567	-0,0105	0,0191
$(X-53)^3$	b_{14}	-0,0000072	0,0003282	-0,02	0,983	-0,0007	0,0006
$(X-63)$	b_{15}	-0,015	0,08796	-0,17	0,865	-0,1906	0,1606
$(X-63)^2$	b_{16}	0,000188	0,006765	0,03	0,978	-0,0133	0,0137
$(X-63)^3$	b_{17}	0,0000487	0,0002887	0,17	0,867	-0,0005	0,0006
$(X-74)$	b_{18}	0,02995	0,07112	0,42	0,675	-0,112	0,172
$(X-74)^2$	b_{19}	-0,000671	0,004937	-0,14	0,892	-0,0105	0,0092
$(X-74)^3$	b_{20}	0,0000043	0,0001836	0,02	0,981	-0,0004	0,0004

Sabit katsayıları kısıtlandırılmış kübik spline regresyon modeliyle bulunan tahmini ihr./ith. oranları ile gözlenen ihr./ith. oranları dağılımı Şekil 5.14'te verilmiştir.

**Şekil 5.14.** Sabit Katsayıları Kısıtlandırılmış Kübik Spline Regresyonla Tahmin Edilen ve Gözlenen ihr./İth. Oranı Dağılımı

5.2.3.2.2. Sabit Katsayıları Kısıtlandırılmış ve Düğüm Noktasında Birinci Derece Türevleri Eşit Olan Kübik Spline Regresyon Modeli

Belirlenen düğüm noktalarına göre oluşturulan gölge değişkenlerle aşağıdaki gibi sabit katsayıları kısıtlandırılmış ve düğüm noktasında birinci derece türevleri eşit olan kübik spline regresyon modeli kurulur.

$$Y_t = a_0 + b_0X_t + b_1X_t^2 + b_2X_t^3 + b_3Z_{1t}^2 + b_4Z_{1t}^3 + \dots + b_{13}Z_{6t}^2 + b_{14}Z_{6t}^3 + e_t \quad (5.15)$$

Oluşturulan regresyon modelinde, a_0 regresyon modelinin sabiti, $b_0, b_1, b_2, b_3, b_4, b_5, \dots, b_{13}$ ve b_{14} regresyon modelinin eğim katsayıları veya bağımsız değişkenlerin katsayı değerleri ve e_t ise tüm regresyon varsayımlarının sağlandığı varsayılan hata terimidir. Bu koşullar altında eşitlik 5.15'teki regresyon denklemi üçüncü bölümde anlatıldığı gibi en küçük karelerle tahmin edilebilir. Yapılan analiz sonucu elde edilen sonuçlar aşağıdaki gibidir.

Regresyon eşitliği,

$$\begin{aligned} \text{Tahmini İhr. /İth. Oranı (Y)} = & -97,29 - 3,955X - 0,05257X^2 - 0,0002306X^3 + \\ & 0,011001(X-11)^2 + 0,0000812(X-11)^3 + 0,009629(X-24)^2 + 0,0001607(X-24)^3 + \\ & 0,000816(X-43)^2 + 0,0000203(X-43)^3 + 0,006021(X-53)^2 - 0,00012192(X-53)^3 + \\ & 0,001573(X-63)^2 + 0,00017735(X-63)^3 - 0,001004(X-74)^2 - 0,00005336(X-74)^3 + e_t \end{aligned} \quad (5.16)$$

biçiminde oluşturulur.

Tablo 5.21'deki varyans analizi sonuçlarına bakıldığında F istatistik değeri 19,24 bulunmuş ve kurulan modelin tamamı oldukça anlamlı çıkmıştır ($p < 0,05$). Modelin hata kareler ortalaması (MSE) 0,02 olarak bulunmuştur.

Tablo 5.21. Sabit Katsayıları Kısıtlandırılmış ve Düğüm Noktasında Birinci Derece Türevleri Eşit Olan Kübik Spline Regresyonda Varyans Analizi

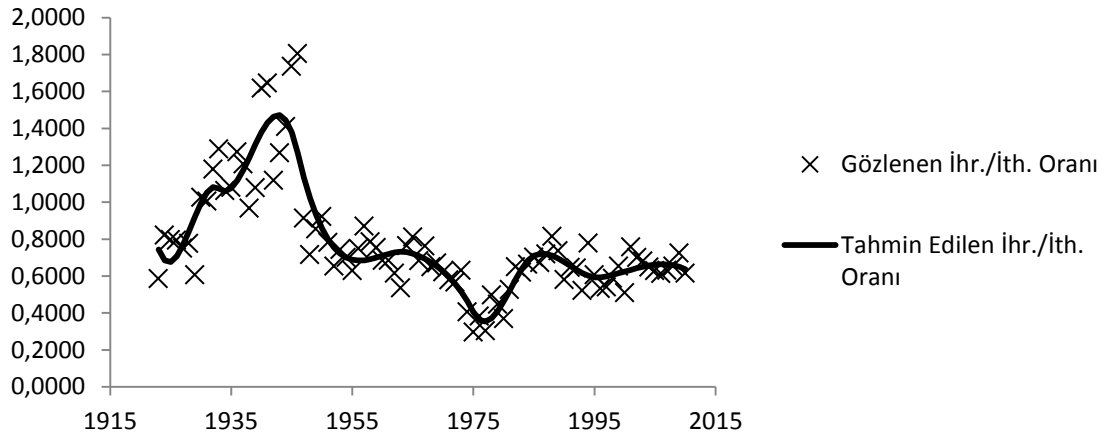
Değişim Kaynağı	Serbestlik Derecesi	Kareler Toplamı	Kareler Ortalaması	F	P
Regresyon	15	6,15804	0,41054	19,24	0,000
Hata	72	1,53660	0,02134		
Toplam	87	7,69464			

Yukarıda belirtilen 5.16 eşitliği Tablo 5.22'ye göre yazılmıştır. Tablo 5.22 incelenecek olursa, a_0 modelin sabitidir. Modelin tek bir sabiti olduğu için, düğüm noktalarında regresyon doğrusu sürekli. Regresyon modelinde b_3 , b_5 ve b_6 katsayı değerleri %5 önem düzeyinde anlamlı bulunarak, söz konusu katsayıların sıfırdan farklı olduğu belirlenmiştir. Ayrıca, a_0 , b_0 , b_1 , b_2 , b_{10} , b_{13} ve b_{14} katsayı değerleri ise negatif bulunmuştur.

Tablo 5.22. Sabit Katsayıları Kısıtlandırılmış ve Düğüm Noktasında Birinci Derece Türevleri Eşit Olan Kübik Spline Regresyon Modelinde Tahminlerin Çıktısı

Değişkenler	Katsayılar	Katsayı Değerleri	Standart Hata Değerleri	T	P	Güven Sınırları Alt Sınır	Üst Sınır
Modelin Sabiti	a_0	-97,29	59	-1,65	0,104	-214,918	20,2975
X	b_0	-3,955	2,378	-1,66	0,101	-8,6972	0,7847
X^2	b_1	-0,05257	0,03179	-1,65	0,103	-0,116	0,0108
X^3	b_2	-0,0002306	0,0001409	-1,64	0,106	-0,0005	0,0001
$(X-11)^2$	b_3	0,011001	0,005015	2,19	0,032	0,001	0,021
$(X-11)^3$	b_4	0,0000812	0,0001222	0,66	0,509	-0,0002	0,0003
$(X-24)^2$	b_5	0,009629	0,002193	4,39	0	0,0053	0,014
$(X-24)^3$	b_6	0,00011607	0,00003504	3,31	0,001	0,000	0,0002
$(X-43)^2$	b_7	0,000816	0,003055	0,27	0,79	-0,0053	0,0069
$(X-43)^3$	b_8	0,0000203	0,00007634	0,27	0,791	-0,0001	0,0002
$(X-53)^2$	b_9	0,006021	0,005411	1,11	0,27	-0,0048	0,0168
$(X-53)^3$	b_{10}	-0,00012192	0,00008103	-1,5	0,137	-0,0003	0,000
$(X-63)^2$	b_{11}	0,001573	0,005127	0,31	0,76	-0,0086	0,0118
$(X-63)^3$	b_{12}	0,00017735	0,0000797	2,23	0,029	0,000	0,0003
$(X-74)^2$	b_{13}	-0,001004	0,00408	-0,25	0,806	-0,0091	0,0071
$(X-74)^3$	b_{14}	-0,00005336	0,00006053	-0,88	0,381	-0,0002	0,0001

Sabit katsayıları kısıtlandırılmış ve düğüm noktasında birinci derece türevleri eşit olan kübik spline regresyon modeliyle bulunan tahmini ihr./ith. oranları ile gözlenen ihr./ith. oranları dağılımı Şekil 5.15'te verilmiştir.



Şekil 5.15. Sabit Katsayıları Kısıtlandırılmış ve Düğüm Noktasında Birinci Derece Türevleri Eşit Olan Kübik Spline Regresyonla Tahmin Edilen ve Gözlenen İhr./İth. Oranı Dağılımı

5.2.3.2.3. Sabit Katsayıları Kısıtlandırılmış ve Düğüm Noktasında İkinci Derece Türevleri Eşit Olan Kübik Spline Regresyon Modeli

Belirlenen düğüm noktalarına göre oluşturulan gölge değişkenlerle aşağıdaki gibi sabit katsayıları kısıtlandırılmış ve düğüm noktasında ikinci derece türevleri eşit olan kübik spline regresyon modeli kurulur.

$$Y_t = a_0 + b_0X_t + b_1X_t^2 + b_2X_t^3 + b_3Z_{1t}^3 + \dots + b_8Z_{6t}^3 + e_t \quad (5.17)$$

Oluşturulan regresyon modelinde, a_0 regresyon modelinin sabiti, $b_0, b_1, b_2, b_3, b_4, b_5, b_6, b_7$ ve b_8 regresyon modelinin eğim katsayıları veya bağımsız değişkenlerin katsayı değerleri ve e_t ise tüm regresyon varsayımlarının sağlandığı varsayılan hata terimidir. Bu koşullar altında 5.17'deki regresyon denklemi üçüncü bölümde anlatıldığı gibi en küçük karelerle tahmin edilebilir. Yapılan analiz sonucu elde edilen sonuçlar aşağıdaki gibidir.

Regresyon eşitliği,

$$\begin{aligned} \text{Tahmini İhr. /İth. Oranı (Y)} = & -46,99 - 2,0784X - 0,02937X^2 - 0,00013554X^3 + \\ & 0,00016552(X-11)^3 - 0,00003207(X-24)^3 + 0,00001366(X-43)^3 - 0,00003116(X-53)^3 + \\ & 0,00003077(X-63)^3 - 0,00001706(X-74)^3 + e_t \end{aligned} \quad (5.18)$$

biçiminde oluşturulur.

Tablo 5.23'teki varyans analizi sonuçlarına bakıldığında F istatistik değeri 18,98 bulunmuş ve kurulan modelin tamamı oldukça anlamlı çıkmıştır ($p < 0,05$). Modelin hata kareler ortalaması (MSE) 0,03 olarak bulunmuştur.

Tablo 5.23. Sabit Katsayıları Kısıtlandırılmış ve Düğüm Noktasında İkinci Derece Türevleri Eşit Olan Kübik Spline Regresyonda Varyans Analizi

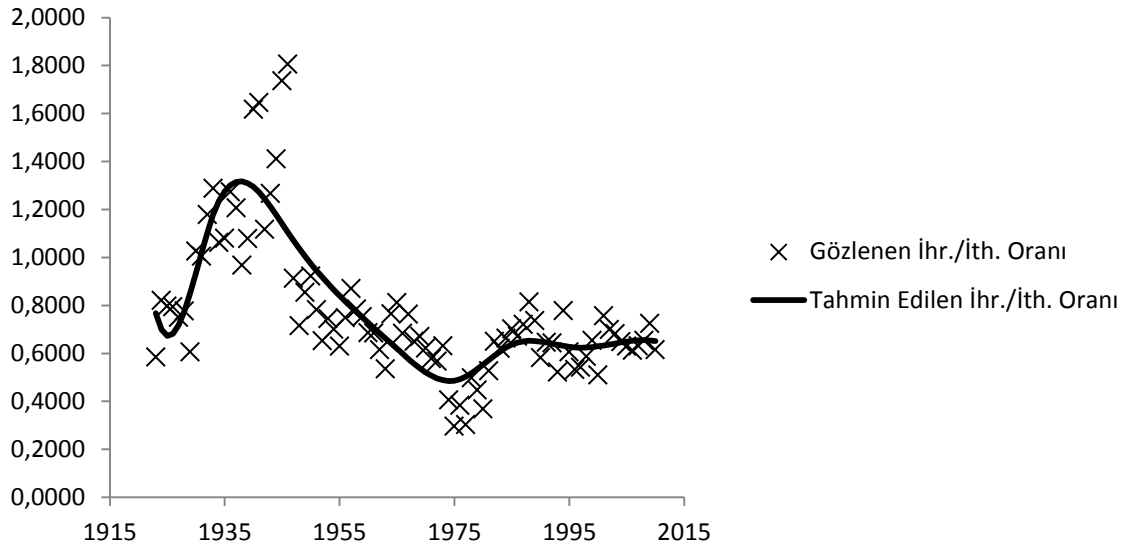
Değişim Kaynağı	Serbestlik Derecesi	Kareler Toplamı	Kareler Ortalaması	F	P
Regresyon	9	5,28249	0,58694	18,98	0,000
Hata	78	2,41215	0,03092		
Toplam	87	7,69464			

Yukarıda belirtilen 5.18 eşitliği Tablo 5.24'e göre yazılmıştır. Tablo 5.24 incelenecek olursa, a_0 modelin sabitidir. Modelin tek bir sabiti olduğu için, düğüm noktalarında regresyon doğrusu süreklidir. Regresyon modelinde a_0 , b_0 , b_1 , b_2 , b_3 ve b_4 katsayı değerleri %5 önem düzeyinde anlamlı bulunarak, söz konusu katsayıların sıfırdan farklı olduğu belirlenmiştir. Sabit katsayıları kısıtlandırılmış ve düğüm noktasında iki derece türevli kübik spline regresyonda 1. ve 2. düğüm noktaları anlamlı bulunmuştur. Ayrıca, a_0 , b_0 , b_1 , b_2 , b_4 , b_6 ve b_8 katsayı değerleri ise negatif bulunmuştur.

Tablo 5.24. Sabit Katsayıları Kısıtlandırılmış ve Düğüm Noktasında İkinci Derece Türevleri Eşit Olan Kübik Spline Regresyon Modelinde Tahminlerin Çıktısı

Değişkenler	Katsayılar	Katsayı Değerleri	Standart Hata Değerleri	T	P	Güven Sınırları Alt Sınır	Üst Sınır
Modelin Sabiti	a_0	-46,99	17,25	-2,72	0,008	-81,3391	-12,6441
X	b_0	-2,0784	0,7522	-2,76	0,007	-3,5761	-0,5811
X^2	b_1	-0,02937	0,01082	-2,71	0,008	-0,0509	-0,0078
X^3	b_2	-0,00013554	0,0000514	-2,64	0,01	-0,0002	0,000
$(X-11)^3$	b_3	0,00016552	0,00005998	2,76	0,007	0,000	0,0003
$(X-24)^3$	b_4	-0,00003207	0,00001413	-2,27	0,026	-0,0001	0,000
$(X-43)^3$	b_5	0,00001366	0,00001511	0,9	0,369	0,000	0,000
$(X-53)^3$	b_6	-0,00003116	0,00002491	-1,25	0,215	-0,0001	0,000
$(X-63)^3$	b_7	0,00003077	0,00002914	1,06	0,294	0,000	0,0001
$(X-74)^3$	b_8	-0,00001706	0,00003936	-0,43	0,666	-0,0001	0,0001

Sabit katsayıları kısıtlandırılmış ve düğüm noktasında ikinci derece türevleri eşit olan kübik spline regresyon modeliyle bulunan tahmini ihr./ith. oranları ile gözlenen ihr./ith. oranları dağılımı Şekil 5.16'da verilmiştir.



Şekil 5.16. Sabit Katsayıları Kısıtlandırılmış ve Düğüm Noktasında İkinci Derece Türevleri Eşit Olan Kübik Spline Regresyonla Tahmin Edilen ve Gözlenen İhr./İth. Oranı Dağılımı

5.3. POLİNOM REGRESYON MODELLERİ

5.3.1. Basit Doğrusal Regresyon Modeli

Basit doğrusal regresyon modeli aşağıdaki gibi kurulur.

$$Y_t = a_0 + b_0X_t + e_t \quad (5.19)$$

Oluşturulan regresyon modelinde, a_0 regresyon modelinin sabiti, b_0 regresyon modelinin eğim katsayısı veya bağımsız değişken katsayı değeri ve e_t ise hata terimidir. Eşitlik 5.19'daki regresyon denklemi üçüncü bölümde anlatıldığı gibi en küçük karelerle tahmin edilebilir. Yapılan analiz sonucu elde edilen sonuçlar aşağıdaki gibidir.

Regresyon eşitliği,

$$\text{Tahmini İhr. /İth. Oranı (Y)} = 0,77567 - 0,0031883X + e_t \quad (5.20)$$

biçiminde oluşturulur.

Tablo 5.25'teki varyans analizi sonuçlarına bakıldığında F istatistik değeri 36,87 bulunmuş ve kurulan modelin tamamı oldukça anlamlı çıkmıştır ($p < 0,05$). Modelin hata kareler ortalaması (MSE) 0,06 olarak bulunmuştur.

Tablo 5.25. Basit Doğrusal Regresyonda Varyans Analizi

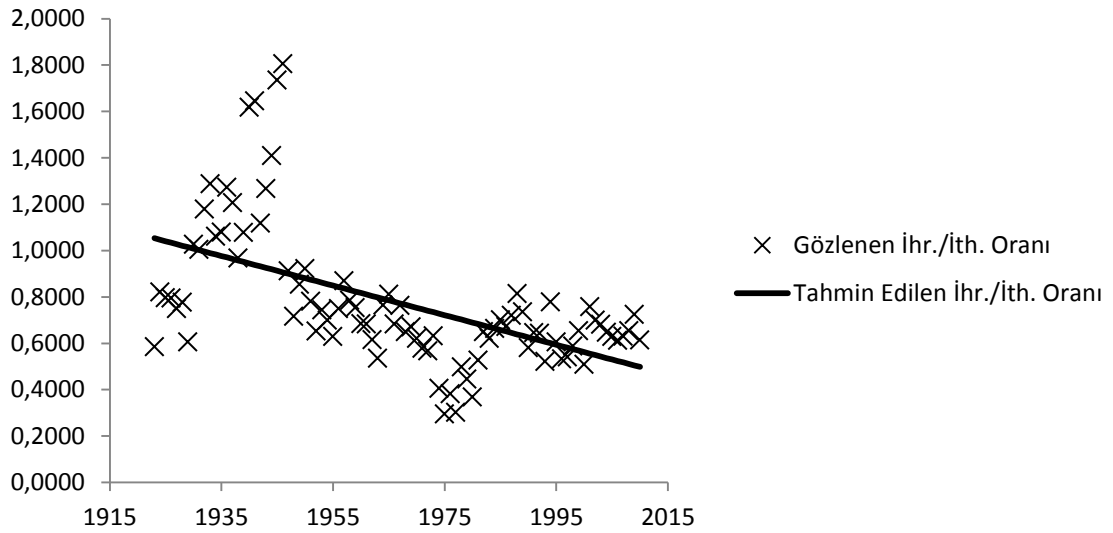
Değişim Kaynağı	Serbestlik Derecesi	Kareler Toplamı	Kareler Ortalaması	F	P
Regresyon	1	2,3088	2,3088	36,87	0,000
Hata	86	5,3858	0,0626		
Toplam	87	7,6946			

Yukarıda belirtilen 5.20 eşitliği Tablo 5.26'ya göre yazılmıştır. Tablo 5.26 incelenecek olursa, a_0 modelin sabiti ve b_0 katsayı değeri %5 önem düzeyinde anlamlı bulunarak, söz konusu katsayıların sıfırdan farklı olduğu belirlenmiştir. Ayrıca, b_0 katsayı değeri ise negatif bulunmuştur.

Tablo 5.26. Basit Doğrusal Regresyon Modelinde Tahminlerin Çıktısı

Değişkenler	Katsayılar	Katsayı Değerleri	Standart Hata Değerleri	T	P	Güven Sınırları Alt Sınır	Üst Sınır
Modelin Sabiti	a_0	0,77567	0,02668	29,08	0,000	0,7226	0,8287
X	b_0	-0,0031883	0,0005251	-6,07	0,000	-0,0042	-0,0021

Basit doğrusal regresyon modeli ile bulunan tahmini ihr./ith. oranları ile gözlenen ihr./ith. oranları dağılımı Şekil 5.17'de verilmiştir.

**Şekil 5.17.** Basit Doğrusal Regresyonla Tahmin Edilen ve Gözlenen İhr./İth. Oranı Dağılımı

5.3.2. Kareli Regresyon Modeli

Kareli regresyon modeli aşağıdaki gibi kurulur.

$$Y_t = a_0 + b_0X_t + b_1X_t^2 + e_t \quad (5.21)$$

Oluşturulan regresyon modelinde, a_0 regresyon modelinin sabiti, b_0 ve b_1 regresyon modelinin eğim katsayıları veya bağımsız değişken katsayı değerleri ve e_t ise hata terimidir. Eşitlik 5.21'deki regresyon denklemi üçüncü bölümde anlatıldığı gibi en küçük karelerle tahmin edilebilir. Yapılan analiz sonucu elde edilen sonuçlar aşağıdaki gibidir.

Regresyon eşitliği,

$$\text{Tahmini İhr. /İth. Oranı (Y)} = 0,73347 - 0,0031883X + 0,00001635X^2 + e_t \quad (5.22)$$

biçiminde oluşturulur.

Tablo 5.27'deki varyans analizi sonuçlarına bakıldığında F istatistik değeri 19,67 bulunmuş ve kurulan modelin tamamı oldukça anlamlı çıkmıştır ($p < 0,05$). Modelin hata kareler ortalaması (MSE) 0,06 olarak bulunmuştur.

Tablo 5.27. Kareli Regresyonda Varyans Analizi

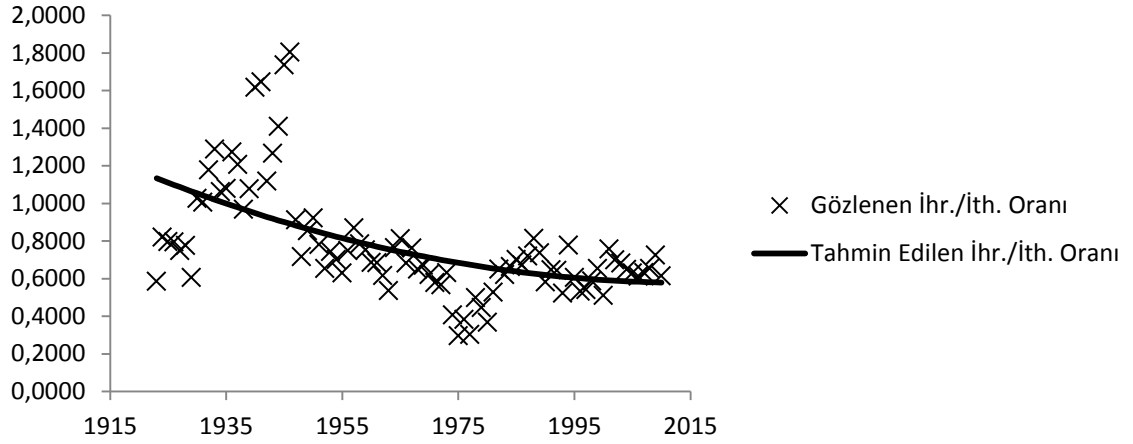
Değişim Kaynağı	Serbestlik Derecesi	Kareler Toplamı	Kareler Ortalaması	F	P
Regresyon	2	2,4341	1,2171	19,67	0,000
Hata	85	5,2605	0,0619		
Toplam	87	7,6946			

Yukarıda belirtilen 5.22 eşitliği Tablo 5.28'e göre yazılmıştır. Tablo 5.28 incelenecek olursa, a_0 modelin sabiti ve b_0 katsayı değeri %5 önem düzeyinde anlamlı bulunarak, söz konusu katsayıların sıfırdan farklı olduğu belirlenmiştir. Ayrıca, b_0 katsayı değeri ise negatif bulunmuştur.

Tablo 5.28. Kareli Regresyon Modelinde Tahminlerin Çıktısı

Değişkenler	Katsayılar	Katsayı Değerleri	Standart Hata Değerleri	T	P	Güven Sınırları	
						Alt Sınır	Üst Sınır
Modelin Sabiti	a_0	0,73347	0,03978	18,44	0,000	0,6544	0,8126
X	b_0	-0,0031883	0,000522	-6,11	0,000	-0,0042	-0,0022
X^2	b_1	0,00001635	0,00001149	1,42	0,158	0,000	0,000

Kareli regresyon modeli ile bulunan tahmini ihr./ith. oranları ile gözlenen ihr./ith. oranları dağılımı Şekil 5.18’de verilmiştir.



Şekil 5.18. Kareli Regresyonla Tahmin Edilen ve Gözlenen İhr./İth. Oranı Dağılımı

5.3.3. Kübik Regresyon Modeli

Kübik regresyon modeli aşağıdaki gibi kurulur.

$$Y_t = a_0 + b_0X_t + b_1X_t^2 + b_2X_t^3 + e_t \quad (5.23)$$

Oluşturulan regresyon modelinde, a_0 regresyon modelinin sabiti, b_0 , b_1 ve b_2 regresyon modelinin eğim katsayıları veya bağımsız değişken katsayı değerleri ve e_t ise hata terimidir. Eşitlik 5.23’teki regresyon denklemi üçüncü bölümde anlatıldığı gibi en küçük karelerle tahmin edilebilir. Yapılan analiz sonucu elde edilen sonuçlar aşağıdaki gibidir.

Regresyon eşitliği,

$$\text{Tahmini İhr. /İth. Oranı (Y)} = 0,73347 - 0,008516X + 0,00001635X^2 + 0,00000115X^3 + e_t \quad (5.24)$$

biçiminde oluşturulur.

Tablo 5.29’deki varyans analizi sonuçlarına bakıldığında F istatistik değeri 25,41 bulunmuş ve kurulan modelin tamamı oldukça anlamlı çıkmıştır ($p < 0,05$). Modelin hata kareler ortalaması (MSE) 0,05 olarak bulunmuştur.

Tablo 5.29. Kübik Regresyonda Varyans Analizi

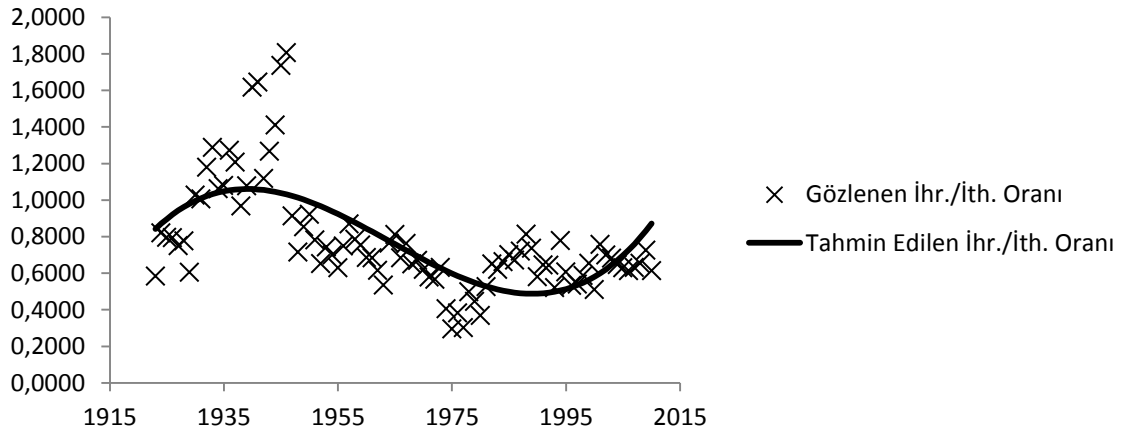
Değişim Kaynağı	Serbestlik Derecesi	Kareler Toplamı	Kareler Ortalaması	F	P
Regresyon	3	3,6607	1,2202	25,41	0,000
Hata	84	4,0339	0,0480		
Toplam	87	7,6946			

Yukarıda belirtilen 5.24 eşitliği Tablo 5.30'a göre yazılmıştır. Tablo 5.30 incelenecek olursa, a_0 modelin sabiti, b_0 ve b_2 katsayı değerleri %5 önem düzeyinde anlamlı bulunarak, söz konusu katsayıların sıfırdan farklı olduğu belirlenmiştir. Ayrıca, b_0 katsayı değeri ise negatif bulunmuştur.

Tablo 5.30. Kübik Regresyon Modelinde Tahminlerin Çıktısı

Değişkenler	Katsayılar	Katsayı Değerleri	Standart Hata Değerleri	T	P	Güven Sınırları Alt Sınır	Üst Sınır
Modelin Sabiti	a_0	0,73347	0,03504	20,93	0,000	0,6638	0,8032
X	b_0	-0,008516	0,00115	-7,4	0,000	-0,0108	-0,0062
X^2	b_1	0,00001635	0,00001012	1,62	0,11	0,000	0,000
X^3	b_2	0,00000115	0,00000023	5,05	0,000	0,000	0,000

Kübik regresyon modeli ile bulunan tahmini ihr./ith. oranları ile gözlenen ihr./ith. oranları dağılımı Şekil 5.19'da verilmiştir.

**Şekil 5.19.** Kübik Regresyonla Tahmin Edilen ve Gözlenen İhr./İth. Oranı Dağılımı

5.4. MODELLERİN KARŞILAŞTIRILMASI

Polinom regresyon modelleri ve en düzgün spline regresyon modelleri içinde hangi modelin daha etkin olduğunu belirlemek için, Tablo 5.31’de belirtilen model seçim kriterleri kullanılmıştır.

Tablo 5.31. Model Seçim Kriterleri

Sıra No	Model Seçim Kriteri	Kısaltması
1	Belirlilik Katsayısı	R^2
2	Düzeltilmiş R^2	$\bar{R}^2$
3	Akaike Bilgi Kriteri	AIC
4	Schwarz Bilgi Kriteri	SIC
5	PRESS Değeri	PRESS
6	R^2_{pred} Değeri	R^2_{pred}
7	Mallows’un C_p Kriteri	C_p
8	Modelin Hatası	S
9	Hata Kareler Ortalaması	MSE
10	Hata Kareler Ortalaması Karekökü	RMSE
11	Yüzde Hata Kareler Ortalaması Karekökü	%RMSE
12	Ortalama Mutlak Hata	MAE
13	Ortalama Mutlak Hata Yüzdesi	MAPE
14	Oransal Hata	RE
15	Hata Kümesinin Katsayısı	CRM
16	Etkinlik İndeksi	Eİ

Model karşılaştırma kriterlerinden, Belirlilik Katsayısı ve Düzeltilmiş R^2 değerlerinin yüksek çıkması, Akaike ve Schwarz Bilgi Kriterlerinin, PRESS istatistiğinin küçük çıkması, PRESS R^2 değerinin yüksek çıkması, Mallows’un C_p Kriteri, Modelin Hatası, Hata Kareler Ortalaması, Yüzde Hata Kareler Ortalaması Karekökü, Ortalama Mutlak Hata, Ortalama Mutlak Hata Yüzdesi ve Oransal Hata değerlerinin küçük çıkması, -1 ile +1 arasında değerler alan Hata Kümesi Katsayısı değerinin sıfıra yakın olması ve pozitif değerleri değerlendirilen Etkinlik İndeksinin ise yüksek olması arzu edilir. Yapılan hesaplamalar sonucunda elde edilen model karşılaştırma kriterleri Tablo 5.32-a ve Tablo 5.32-b’de gösterilmiştir.

Tablo 5.32-a. Modellerin Karşılaştırılması

Modeller/Model Kriterleri	R ²	$\bar{R}^2$	AIC	SIC	PRESS	R ² _{pred} (%)	C _p	S
Doğrusal Spline Regresyon	0,6658	0,6366	-0,5132	-0,2880	3,0483	60,3800	7,0059	0,1793
Kareli Spline Regresyon	0,7271	0,6995	-0,6930	-0,4396	2,5373	67,0200	8,0094	0,1630
Kübik Spline Regresyon	0,6865	0,6504	-0,5317	-0,2502	3,1180	59,4800	9,0126	0,1759
Doğrusal Regresyon	0,3001	0,2919	0,0898	0,1461	5,6405	26,7000	1,0351	0,2503
Kareli Regresyon	0,3164	0,3003	0,0889	0,1734	5,6449	26,6400	1,9838	0,2488
Kübik Regresyon	0,4758	0,4570	-0,1538	-0,0412	4,4119	42,6600	3,0396	0,2191

Tablo 5.33-b. Modellerin Karşılaştırılması

Modeller/Model Kriterleri	MSE	RMSE	%RMSE	MAE	MAPE	RE	CRM
Doğrusal Spline Regresyon	0,0321	0,1793	0,2285	0,1208	15,6207	3,5219	-1,10 ⁻¹⁵
Kareli Spline Regresyon	0,0266	0,1630	0,2078	0,1161	15,5115	3,4077	-3, 10 ⁻¹⁵
Kübik Spline Regresyon	0,0309	0,1758	0,2241	0,1198	15,8322	3,6915	1, 10 ⁻¹³
Doğrusal Regresyon	0,0626	0,2502	0,3189	0,1779	24,8762	8,8757	-4, 10 ⁻¹⁶
Kareli Regresyon	0,0619	0,2488	0,3171	0,1659	22,6838	8,4541	-8, 10 ⁻¹⁶
Kübik Regresyon	0,0480	0,2191	0,2792	0,1611	21,8969	6,1340	-4, 10 ⁻¹⁶

Tablo 5.32-a ve Tablo 5.32-b incelendiğinde, Mallows'un C_p Kriteri ve Hata Kümesi Katsayısı hariç, spline regresyon modellerin polinom regresyon modellere göre daha iyi sonuçlar verdiği görülmektedir. Etkinlik indeksi değerleri negatif çıktığı için dikkate alınmamıştır. En düzgün spline regresyon modellerinde, kareli spline regresyon modelinin ihracatın ithalatı karşılama oranları dağılımını en iyi açıklayan model olduğu tespit edilmiştir. Buna göre, kareli spline regresyon modelin en etkin model olduğu söylenebilir.

SONUÇ VE DEĞERLENDİRME

İhracatın ithalatı karşılama oranları dağılımında, önceden bilinmeyen düğüm noktalarının sayısı ve konumlarının tespit edilmesi için doğrusal spline regresyonları kullanarak R^2 'yi temel alan bir yöntem önerilmiştir. Yöntemin temel mantığı belirtilecek olursa; bağımsız değişken ile bağımlı değişken arasında yapısal değişiklikler olduğunda R^2 değerleri de değişiklik gösterir. Buradan hareketle, R^2 değerlerindeki önemli değişikliklere (artış veya azalışlara) göre düğüm noktaları belirlenmiştir. Öncelikle analizde her bir gözlem aday düğüm noktası olarak düşünülmüş ve buna göre doğrusal spline regresyon modelleri kurulmuş ve R^2 değerleri hesaplanmıştır. Hesaplanan R^2 değerlerinin serpilme diyagramına bakılarak değişim gösterdiği ilk yer belirlenmiş ve buraya tekabül eden gözlem düğüm noktası olarak alınmıştır. Daha sonra ilk düğüm noktasına kadar olan gözlemler analizden çıkartılarak yeniden R^2 değerleri hesaplanmıştır. Hesaplanan R^2 değerlerinin serpilme diyagramına bakılarak ikinci düğüm noktası belirlenmiştir. Bu şekilde devam eden süreç sonucunda 1933, 1946, 1965, 1975, 1985 ve 1996 yılları olmak üzere altı düğüm noktası belirlenmiştir.

1923-2010 yılları arası ihracatın ithalatı karşılama oranlarını dalgalı bir dağılım göstermektedir. Genel olarak dağılımın mahiyeti incelendiğinde, 1923-1929 yılları arasında ihracatın çoğunluğunu tarım ürünleri belirlemekteydi. Bu dönemde ithal ikameci politikalar belirlenerek yerli üretim geliştirilmeye çalışılmıştır. 1926-1929 yıllarında dünya ekonomik buhranına bağlı olarak dış ticarette daralmalar olmuştur. 1930-1939 yılları arasında korumacı ve devletçi politikalar egemen olmuştur. Bu dönemde devlet desteği ile milli bir sanayileşme süreci başlamıştır. Büyük Buhran sonrası uygulanan politikalar ile önemli oranda daralmış olan ithalat II. Dünya Savaşı döneminde iyice azalmış ve dış ticaret dengesi fazla vermeye devam etmiştir. İthalatın iyice azalmasıyla 1946 yılında ihracatın ithalatı karşılama oranı %180,49 olmuştur. Bu nokta ikinci düğüm noktası olarak tespit edilmiştir. Savaş dönemi sonrası serbest ticaretçi iktisat politikaları uygulanmaya başlanmıştır. İthalatın serbestleştirilmesiyle birlikte 1947 yılından itibaren ithalat artmış ve ihracatın ithalatı karşılama oranı azalmaya başlamıştır. 1954-1961 yılları arasında ithalatı kısıtlamaya yönelik politikalar uygulanmıştır. Bu dönemde istikrarlı bir şekilde hareket eden ve büyük ölçüde tarım ürünlerinden oluşan ihracat sayesinde, dış ticaret açıkları daraltılmış ve sürdürülebilir

olmuştur. 1961’de başlayan yeni dönem; ithal ikameci, korumacı dış ticaret politikalarının sürdürüldüğü ve iç piyasaya yönelik bir sanayileşme politikasının benimsendiği bir dönem olmuştur. 1962-1976 yılları arasında ithalata bağımlılık fazlaydı ve ihracatta da durgun bir seyir izleniyordu. Dolayısıyla 1975 yılında ihracatın ithalatı karşılama oranı %29,57 olmuştur. Bu en düşük değer dördüncü düğüm noktası olarak tespit edilmiştir. 24 Ocak 1980 Türkiye ekonomi politikaları açısından bir dönüm noktasıdır. 1980 sonrası politikalarla ihracat, 1980-1984 arasındaki dönemde yüzde 314 oranında artmış, ancak ithalatta da artış olması sebebiyle dış açıklar 1980 sonrası dönemde de devam etmiştir. 1990’da Türk parasının konvertibilitesi (Milli paraların karşılıklı olarak birbirine çevrilebilmesi) ilan edilmiştir. 1990’larda ve 2000’li yıllarda Türkiye’nin dış ticaret politikalarını şekillendiren iki temel bütünleşme Dünya Ticaret Örgütü üyeliği ve Avrupa Birliği ile girilen Gümrük Birliği olmuştur.

Doğrusal spline regresyon modelinde,

Kısıtlandırılmamış doğrusal spline regresyon analizi sonucu, F istatistik değeri 36,59 bulunmuş ve kurulan modelin tamamı oldukça anlamlı çıkmıştır ($p=0$). Regresyon modelinde 14 katsayıdan, a_0 , b_0 , b_3 , b_4 , b_6 , b_8 ve b_{10} katsayıları %5 önem düzeyinde anlamlı bulunarak, söz konusu katsayıların sıfırdan farklı olduğu belirlenmiştir.

Kısıtlandırılmış yani en düzgün doğrusal spline regresyon analizi sonucu, F istatistik değeri 22,77 bulunmuş ve kurulan modelin tamamı oldukça anlamlı çıkmıştır ($p=0$). Regresyon modelinde 8 katsayıdan, modelin sabiti, b_0 , b_1 , b_2 ve b_4 katsayıları %5 önem düzeyinde anlamlı bulunarak, söz konusu katsayıların sıfırdan farklı olduğu belirlenmiştir. Kısıtlandırılmış doğrusal spline regresyonda 1933, 1946 ve 1975 yıllarına ait düğüm noktaları anlamlı bulunmuştur.

Kareli spline regresyon modelinde,

Kısıtlandırılmamış kareli spline regresyon analizi sonucu, F istatistik değeri 36,59 bulunmuş ve kurulan modelin tamamı oldukça anlamlı çıkmıştır ($p=0$). Regresyon modelinde 21 katsayıdan, a_0 , b_0 , b_1 , b_3 , b_5 ve b_6 katsayıları %5 önem düzeyinde anlamlı bulunarak, söz konusu katsayıların sıfırdan farklı olduğu belirlenmiştir.

Sabit katsayıları kısıtlandırılmış kareli spline regresyon analizi sonucu, F istatistik değeri 19,35 bulunmuş ve kurulan modelin tamamı oldukça anlamlı çıkmıştır ($p=0$).

Regresyon modelinde 15 katsayıdan b_4 ve b_5 katsayıları %5 önem düzeyinde anlamlı bulunarak, söz konusu katsayıların sıfırdan farklı olduğu belirlenmiştir.

Sabit katsayıları kısıtlandırılmış ve düğüm noktasında birinci derece türevleri eşit yani en düzgün kareli spline regresyon analizi sonucu, F istatistik değeri 26,31 bulunmuş ve kurulan modelin tamamı oldukça anlamlı çıkmıştır ($p=0$). Regresyon modelinde 9 katsayıdan, a_0 , b_2 , b_3 , b_4 , b_5 ve b_6 katsayıları %5 önem düzeyinde anlamlı bulunarak, söz konusu katsayıların sıfırdan farklı olduğu belirlenmiştir. Sabit katsayıları kısıtlandırılmış ve düğüm noktasında birinci derece türevleri eşit olan kareli spline regresyonda 1933, 1946, 1965, 1975 ve 1985 yıllarına ait düğüm noktaları anlamlı bulunmuştur.

Kübik spline regresyon modelinde,

Kısıtlandırılmamış kübik spline regresyon analizi sonucu, F istatistik değeri 18,17 bulunmuş ve kurulan modelin tamamı oldukça anlamlı çıkmıştır ($p=0$). Regresyon modelinde 28 katsayıdan, b_5 , b_7 ve b_8 katsayıları %5 önem düzeyinde anlamlı bulunarak, söz konusu katsayıların sıfırdan farklı olduğu belirlenmiştir.

Sabit katsayıları kısıtlandırılmış kübik spline regresyon analizi sonucu, F istatistik değeri 15,52 bulunmuş ve kurulan modelin tamamı oldukça anlamlı çıkmıştır ($p=0$). Regresyon modelinde 22 katsayıdan sadece b_7 katsayısı %5 önem düzeyinde anlamlı bulunarak, söz konusu katsayının sıfırdan farklı olduğu belirlenmiştir.

Sabit katsayıları kısıtlandırılmış ve düğüm noktasında birinci derece türevleri eşit olan kübik spline regresyon analizi sonucu, F istatistik değeri 19,24 bulunmuş ve kurulan modelin tamamı oldukça anlamlı çıkmıştır ($p=0$). Regresyon modelinde 16 katsayıdan, b_3 , b_5 ve b_6 katsayı değerleri %5 önem düzeyinde anlamlı bulunarak, söz konusu katsayıların sıfırdan farklı olduğu belirlenmiştir.

Sabit katsayıları kısıtlandırılmış ve düğüm noktasında ikinci derece türevleri eşit yani en düzgün kübik spline regresyon analizi sonucu, F istatistik değeri 18,98 bulunmuş ve kurulan modelin tamamı oldukça anlamlı çıkmıştır ($p=0$). Regresyon modelinde 10 katsayıdan, a_0 , b_0 , b_1 , b_2 , b_3 ve b_4 katsayı değerleri %5 önem düzeyinde anlamlı bulunarak, söz konusu katsayıların sıfırdan farklı olduğu belirlenmiştir. Sabit katsayıları kısıtlandırılmış ve düğüm noktasında iki derece türevli kübik spline regresyonda 1933 ve 1946 yıllarına ait düğüm noktaları anlamlı bulunmuştur.

Görüldüğü üzere, belirlenen düğüm noktalarının en düzgün kareli spline regresyon modelinde daha çok anlamlı olduğu tespit edilmiştir.

Modellerin karşılaştırılması sonucu, Mallows'un C_p Kriteri ve Hata Kümesi Katsayısı hariç, spline regresyon modellerin polinom regresyon modellere göre daha iyi sonuçlar verdiği görülmektedir. En düzgün spline regresyon modellerinde, kareli spline regresyon modelinin ihracatın ithalatı karşılama oranları dağılımını en iyi açıklayan model olduğu tespit edilmiştir.

Şöyleki, en düzgün kareli spline regresyon modelinde, diğer modellere göre Belirlilik Katsayısı (R^2) değeri 0.73, Düzeltilmiş R^2 değeri 0.70 ve PRESS R^2 değeri 0.67 daha yüksek çıktığından, kareli spline regresyon en iyi modeldir. Akaike ve Schwarz Bilgi Kriterleri sırasıyla -0.693 ve -0.440, PRESS istatistiği 2.54, Modelin Hatası 0.16, Hata Kareler Ortalaması 0.027, Hata Kareler Ortalaması Karekökü 0.163, Yüzde Hata Kareler Ortalaması Karekökü 0.208, Ortalama Mutlak Hata 0.116, Ortalama Mutlak Hata Yüzdesi 15.51 ve Oransal Hata değeri 3.41 olarak bulunmuştur. Söz konusu değerler diğer modellere göre daha küçük çıktığından kareli spline regresyon en iyi modeldir. Sonuç olarak, kareli spline regresyon modelin en etkin model olduğu söylenebilir.

KAYNAKÇA

- Agarwal, G.G. and Studden, W.J., *An Algorithm for Selection of Design and Knots in the Response Curve Estimation by Spline Functions*, Technical Report No. 78-15, Purdue University, Indiana 1978.
- Aggrey, S.E., “Comparison of Three Nonlinear and Spline Regression Models for Describing Chicken Growth Curves”, *Poultry Science*, 81, 2002, 1782-1788.
- Ahlberg, J. H., Nilson, E. N. and Walsh, J. L., *The Theory of Splines and Their Applications*, Academic Press, New York 1967.
- Akkartal Erkut, *Splayn Fonksiyonları ve Türkiye’de Kişi Başına Düşen Otomobil Miktarları Üzerine Bir Uygulama*, (Yayımlanmamış Doktora Tezi), İstanbul Üniversitesi Sosyal Bilimler Enstitüsü, İstanbul 2000.
- Aladağ, Ç. H. ve Eğrioğlu, E., “En İyi Yapay Sinir Ağı Mimarisinin Belirlenmesinde Uyarlamalı Ağırlıklı Bilgi Kriterinin Kullanılması”, *Türkiye İstatistik Kurumu (Haz.), 19. İstatistik Araştırma Sempozyumu Bildiriler Kitabı: 6-7 Mayıs*, (ss.49-56), Ankara 2010.
- Allen, D. M., “The Relationship between Variable Selection and Data Agumentation and a Method for Prediction”, *Technometrics*, 16(1), 1974, 125-127.
- Allsop, D. and Weisberg H. F., “Measuring Change in Party Identification in an Election Campaign”, *American Journal of Political Science*, 32(4), 1988, 996-1017.
- Anderson, J. E., “Cubic-Spline Urban-Density Functions”, *Journal of Urban Economics*, 12(2), 1982, 155-167.
- Anderson, J. E., “The Changing Structure of a City: Temporal Changes in Cubic Spline Urban”, *Journal of Regional Science*, 25(3), 1985, 413-425.
- Aydın, D., “A Comparison of the Nonparametric Regression Models Using Smoothing Spline and Kernel Regression”, *International Journal of Mathematical, Physical and Engineering Sciences*, 2(2), 2008, 75-79.
- Aydın Dursun, *Semiparametrik Regresyon Modellemede Splayn Düzeltme Yaklaşımı ile Tahmin ve Çıkarımlar*, (Yayımlanmamış Doktora Tezi), Anadolu Üniversitesi

- Fen Bilimleri Enstitüsü, Eskişehir 2005.
- Bacon, D. W. and Watts, D. G., “Estimating the Transition between Two Intersecting Straight Lines”, *Biometrika*, 58(3), 1971, 525-534.
- Bağdatlı Seda, *Semiparametrik Regresyon ve Bir Uygulama*, (Yayımlanmamış Yüksek Lisans Tezi), İstanbul Ticaret Üniversitesi Fen Bilimleri Enstitüsü, İstanbul 2010.
- Barrodale, I. and Young, A., “A Note on Numerical Procedures for Approximation by Spline Functions”, *Computer Journal*, 9(3), 318-320.
- Bartels, R. H., Beatty, J. C. and Barsky, B. A., *An Introduction to Splines for use in Computer Graphics and Geometric Modeling*, Morgan Kaufmann Publishers, California 1987.
- Barth, J., Kraft, A. and Kraft, J., “Estimation of the Liquidity Trap Using Spline Functions”, *The Review of Economics and Statistics*, 58(2), 1976, 218-222.
- Baş Fatih, *Spline Fonksiyonları ve Uygulamaları*, (Yayımlanmamış Yüksek Lisans Tezi), Sakarya Üniversitesi Fen Bilimleri Enstitüsü, Sakarya 2007.
- Başar, A. ve Oktay, E., *Uygulamalı İstatistik 2 Kısa Teorik Bilgiler ve Çözülmüş Problemler*, Aktif Yayınevi, Erzurum 2004.
- Bayram, Mustafa, *Nümerik Analiz*, Birsan Yayınevi Ltd. Şti., İstanbul 2009.
- Bellman, R. and Roth, R., “Curve Fitting by Segmented Straight Lines”, *Journal of the American Statistical Association*, 64(327), 1969, 1079-1084.
- Berberoğlu, B. and Berberoğlu, C. N., “Modeling the Structural Shifts in Real Exchange Rate with Cubic Spline Regression (CSR). Turkey 1987-2008”, *International Journal of Business and Social Science*, 2(17), 2011, 60-66.
- Berberoğlu, B., “A Note on the Predicted Responses in Cubic Splines”, *Chiana-USA Business Review*, 11(1), 2012, 30-36.
- Bessaoud, F., Daures, J. P. And Molinari, N., “Free Knot Splines for Logistic Models and Threshold Selection”, *Computer Methods and Programs in Biomedicine*, 77(1), 2005, 1-9.
- Bhat, R. B. And Chakraverty, S., *Numerical Analysis in Engineering*, Alpha Science International, Pangbourne 2004.

- Bierens, H. J., *Information Criteria and Model Selection*, 2006, Eriřim Tarihi: 27 Haziran 2012, Pennsylvania State University Ađ Sitesi:
<http://econ.la.psu.edu/~hbierens/INFORMATIONCRIT.PDF>
- Boor, C. de and Rice, J. R., *Least Squares Cubic Spline Approximation I-Fixed Knots*, Technical Report CSD-TR 20, Computer Sciences, Purdue University, Indiana 1968.
- Breaz, N. and Breaz, D., “On the Spline Functions in Data Analysis Framework”, *International Conference on Theory and Applications of Mathematics and Informatics ICTAMI2005* (pp. 197-209), Acta Universitatis Apulensis, ISSN 1582-5329, Alba Lulia, Romania 2005.
- Brusilovskiy, E., “The Piecewise Regression Model as a Response Modeling Tool”, *Conference Proceedings: NESUG 2004*, Eriřim Tarihi: 08.12.2012, <http://www.nesug.org/html/Proceedings/nesug04/an/an09.pdf>
- Buse, A. and Lim, L., “Cubic Splines as a Special Case of Restricted Least Squares”, *Journal of the American Statistical Association*, 72(357), 1977, 64-68.
- Canıvar Aynur, *Kübik Spline Fonksiyonlar Yardımıyla Lineer Olmayan Kısmi Türevli Diferensiyel Denklemlerin Sayısal Çözümleri*, (Yayımlanmamıř Yüksek Lisans Tezi), Osmangazi Üniversitesi Fen Bilimleri Enstitüsü, Eskiřehir 2003.
- Chan Shih-Huang, *Polynomial Spline Regression with Unknown Knots and AR(1) Errors*, Ph. D. Disertation, Ohio State University, Ohio 1989.
- Chau, J., Bonfiglioli, C., Chey, T. and Bauman, A., “The Cinderella of Public Health News: Physical Activity Coverage in Australian Newspapers, 1986-2006”, *Australian and New Zealand Journal of Public Health*, 33(2), 2009, 189-192.
- Chernobai, E., Reibel, M. and Carney, M., “Nonlinear Spatial and Temporal Effects of Highway Construction on House Prices”, *The Journal of Real Estate Finance and Economics*, 42(3), 2009, 348-370.
- Chib, S. and Greenberg, E., “Additive Cubic Spline Regression with Dirichlet Process Mixture Errors”, *Journal of Econometrics*, 156(2), 2010, 322-336.
- De Boor, C., *A Practical Guide to Splines*, Springer Verlag, New York 1978.

- De Boor, C. and Rice, J. R., *Least Squares Cubic Spline Approximation II-Variable Knots*, Technical Report CSD TR 21, USA: Purdue University 1968.
- Demirel N., “İzmir İli Sıcaklık Verilerinin Spline Regresyon ile Modellenmesi” [Bildiri], 7. İstatistik Günleri Sempozyumu, (Haziran 2010), ODTÜ/Ankara.
- Demirel, N., “İzmir İli Sıcaklık Verilerinin Regresyon Eğrileri İle Modellenmesi”, *İstatistik Araştırma Dergisi*, 8(2), 2011, 68-75.
- Detle, H. and Melas, V. B., “A Note on All-Bias Designs with Applications in Spline Regression Models”, *Journal of Statistical Planning and Inference*, 140(7), 2010, 2037-2045.
- Detle, H., Melas, V. B. and Pepelyshev, A., “Optimal Designs for Free Knot Least Squares Splines”, *Statistica Sinica*, 18, 2008, 1047-1062.
- Devlin, T. F. and Weeks, B. J., “Spline Functions for Logistic Regression Modeling”, SAS Institute Inc (Ed.), *Proceedings of the Eleventh Annual SAS Users Group International Conference*, (pp. 646-649), Atlanta, Georgia 1986.
- Doorn, J. V., Verhoef, P. C. and Bijmolt, T. H. A., “The Importance of Non-Linear Relationships between Attitude and Behaviour in Policy Research”, *Journal of Consumer Policy*, 30(2), 2007, 75-90.
- Doğan Meltem, *q-Integer Noktalarında B-Spline Fonksiyonu ve Özellikleri*, (Yayımlanmamış Yüksek Lisans Tezi), Muğla Üniversitesi Fen Bilimleri Enstitüsü, Muğla 2008.
- Dönmez Duygu, *Diferansiyel Denklemlerin B-Spline Metodu ile Çözümleri Hakkında*, (Yayımlanmamış Yüksek Lisans Tezi), Celal Bayar Üniversitesi Fen Bilimleri Enstitüsü, Manisa 2008.
- Draper, N. R., Guttman, I. and Lipow, P., “All-Bias Designs for Spline Functions Joined at the Axes”, *Journal of the American Statistical Association*, 72(358), 1977, 424-429.
- Draper, N. R. and Smith, H., *Applied Regression Analysis*, Third Edition, John Wiley&Sons, New York 1998.
- Dunfield, L.G. and Read, J.F., “Determination of Reaction Rates by the Use of Cubic

- Spline Interpolation”, *The Journal of Chemical Physics*, 57(5), 1972, 2178-2183.
- Dupuis, Depice Janice, *Knots in Spline Regression: Estimation and Inference Using Laplace Transform Techniques*, Ph. D. Disertation, University of New Brunswick, New Brunswick/Canada 1994.
- Ertekin, M. S. ve Erol, K., “1980 Sonrası Dönemde Türkiye Dış Ticaretinin Genel Bir Analizi”, *Anadolu Üniversitesi İ.İ.B.F Dergisi*, 16 (1-2), 2000, 223-254.
- Ertel, J.E. and Fowlkes, E.B., “Some Algorithms for Linear Spline and Piecewise Multiple Linear Regression”, *Journal of the American Statistical Association*, 71(355), 1976, 640-648.
- Eubank, R.L., *Approximate Regression Models and Splines*, The Office of Naval Research, (Technical Report No. 180), Southern Methodist University, USA/Texas 1984.
- Eubank, R.L., *Nonparametric Regression and Spline Smoothing*, Second Edition, Marcel Dekker, New York 1999.
- Eubank, R.L., *Spline Smoothing and Nonparametric Regression*, Marcel Dekker, New York 1988.
- Farley, J. U. and Hinich, M. J., “A Test for a Shifting Slope Coefficient in a Linear Model”, *Journal of the American Statistical Association*, 65(331), 1970, 1320-1329.
- Foley, T. A. and Nielson, G. M., “Knot Selection for Parametric Spline Interpolation”, T. Lyche and L. L. Schumaker (Ed.), *Mathematical Methods in Computer Aided Geometric Design*, (pp.261-262), Academic Press, New York 1989.
- Fox, K. J., “Non-Parametric Estimation of Technical Progress”, *Journal of Productivity Analysis*, 10(3), 1998, 235–250.
- Freund, R., Littell, R. and Creighton, L., *Regression Using JMP*, SAS Institute, North Carolina 2003.
- Freund, R. J., Wilson, W. J. and Sa, P., *Regression Analysis: Statistical Modeling of a Response Variable*, Second Edition, Academic Press, San Diego 2006.

- Friedman, J. H. and Silverman, B. W., “Flexible Parsimonious Smoothing and Additive Modeling”, *Technometrics*, 31(1), 1989, 3-21.
- Fuller, W. A., “Grafted Polynomials as Approximating Functions”, *Australian Journal of Agricultural Economics*, 13(1), 1969, 35–46.
- Gallant, A.R. and Fuller, W. A., “Fitting Segmented Polynomial Regression Models Whose Join Points have to be Estimated”, *Journal of the American Statistical Association*, 68(341), 1973, 144-147.
- Gamiz, M. L., Kulasekera, K. B., Limnios, N. and Lindqvist, B. H., *Applied Nonparametric Statistics in Reliability*, Springer-Verlag, London 2011.
- Garcia, J. L., Coronel, N. A., Martinez, E. A. and Duque, J. A., “Using Spline for Demand Forecasting”, V. G. Almanza (Ed.), *Second International Congress in Engineering and Technology: October 2008*, (pp. 37-56), Chihuahua, Mexico: Ciudad Juarez 2008.
- Genç, A., Oturanç, G., Kınacı, İ. ve Kurnaz, A., “Güneş Radyasyon Verilerinin Kübik Spline Fonksiyonları Yardımıyla Modellenmesi” [Bildiri], *Hacettepe Üniversitesi İstatistik Günleri 2002 Sempozyumu*, (Mayıs 2002), Ankara.
- Genç, A., Kınacı, İ., Kurnaz, A. ve Oturanç, G., “Parçalı Regresyon Kullanılarak Güneş Radyasyon Verilerinin Değerlendirilmesi”, *VI. Türk-Alman Enerji Sempozyumu: 21-24 Haziran*, (ss. 271-280), Ege Üniversitesi, İzmir 2001.
- Genç, A., Kınacı, İ., Oturanç, G., Kurnaz, A., Bilir, Ş. and Özbalta, N., “Statistical Analysis of Solar Radiation Data Using Cubic Spline Functions”, *Energy Sources*, 24(12), 2002, 1131-1138.
- Gnad, W., “Regression by Spline Functions”, *Empirical Economics*, 2(2), 1977, 69-77.
- Goldfeld, S. M. and Quandt, R. E., *Nonlinear Methods in Econometrics*, North-Holland Publishing Company, Amsterdam 1972.
- Goldfeld, S. M. and Quandt, R. E., “The Estimation of Structural Shifts by Switching Regressions”, *Annals of Economic and Social Measurement*, 2(4), 1973, 475-485.
- Gray, Katharine Lynn, *Comparision of Trend Detection Methods*, Ph. D. Disertation, University of Montana, Missoula 2007.

- Greene, W. H, *Econometric Analysis*, Fifth Edition, Prentice Hall, New Jersey 2003.
- Gregory, M. and Serono, M. “A Collection of SAS Macros to Calculate Odds Ratios Using Spline Regression”, *SAS: PhUSE Proceedings (October 12-15)*, Manchester, UK 2008, Eriřim Tarihi: 126.2011, http://lexjansen.com/cgi-bin/xsl_transform.php?x=phuse2008&s=proceedings&c=phuse
- Gregory, M., Ulmer, H., Pfeiffer, K. P., Lang, S. and Strasak, A. M., “A Set of SAS Macros for Calculating and Displaying Adjusted Odds Ratios (With Confidence Intervals) for Continuous Covariates in Logistic B-Spline Regression Models”, *Computer Methods and Programs in Biomedicine*, 92, 2008, 109-114.
- Greville, T.N.E., “Introduction to Spline Functions”, (Ed. T.N.E. Greville), *In Theory and Applications of Spline Functions*, (pp. 1-35), Academic Press, New York 1969.
- Gu, C., “Multivariate Spline Regression”, (Ed. Michael G. Schimek), *Smoothing and Regression Approaches, Computation and Application*, (pp.329-355), John Wiley&Sons, New York 2000.
- Gujarati, D. N, “Use of Dummy Variables in Testing for Equality Between Sets of Coefficients in Two Linear Regressions”, *The American Statistician*, 24(1), 1970, 50-52.
- Gujarati, D. N., *Basic Econometrics*, 4th Edition, McGraw-Hill, New York 2003.
- Guo, Q. And White, R.E., “Cubic Spline Regression for the Open-Circuit Potential Curves of a Lithium-Ion Battery”, *Journal of the Electrochemical Society*, 152(2), 2005, 343-350.
- Güriř, S. ve Çağlayan, E., *Ekonometri Temel Kavramlar*, 3. Basım, Der Yayınları, İstanbul 2010.
- Hakverdi Kazım, *Parçalı Lineer Regresyon*, (Yayımlanmamış Yüksek Lisans Tezi), Gazi Üniversitesi Fen Bilimleri Enstitüsü, Ankara 1995.
- Halpern, E.F., “Bayesian Spline Regression When the Number of Knots is Unknown”, *Journal of the Royal statistical Society Series B*, 35(2), 1973, 347-360.
- Hardy, M. A, *Regression With Dummy Variables*, Sage Publications, California 1993.

- Harrell, F. E., *Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression and Survival Analysis*, Springer-Verlag, New York 2001.
- Hawkins, D. M., "Point Estimation of the Parameters of Piecewise Regression Models", *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 25(1), 1976, 51-57.
- Hinkley, D.V., "Inference about the Intersection in Two-Phase Regression", *Biometrika*, 56(3), 1969, 495-504.
- Hinkley, D.V., "Inference in Two-Phase Regression", *Journal of the American Statistical Association*, 66(336), 1971, 736-743.
- Hirvonen, L., *The Effect of Children on Earnings Using Exogenous Variation in Family Size: Swedish Evidence*, Working Paper Series 2, Swedish Institute for Social Research, Stockholm University 2009.
- Horowitz, J.L., Loughran, T. and Savin, N.E., "A Spline Analysis of the Small Firm Effect: Does Size Really Matter?", *EconWPA, Econometrics*, 9, 1996, 1-26, Erişim Tarihi: 11.10.2011, <http://129.3.20.41/eps/em/papers/9608/9608001.pdf>
- Horvath, S., *Weighted Network Analysis Applications in Genomics and Systems Biology*, Springer Science+Business Media, New York 2011.
- Hudson, D.J., "Fitting Segmented Curves Whose Join Points Have to be Estimated", *Journal of the American Statistical Association*, 61(316), 1966, 1097-1129.
- Huang, J. Z., "Local Asymptotics for Polynomial Spline Regression", *The Annals of Statistics*, 31(5), 2003, 1600-1635.
- Hurley, D., Hussey, J., McKeown, R. and Addy, C., *An Evaluation of Splines in Linear Regression* [Poster], *SAS Conference Proceedings: SAS Users Group International 31*, San Francisco (March 2006).
- Hwang, I. K., "Using SAS-GLM in Kinetic Modeling", SAS Institute Inc (Ed.), *Proceedings of the Third Annual SAS Users Group International Conference*, (pp. 123-126), Las Vegas, Nevada 1978.
- Jianqing, F. and Quiwei, Y., *Nonlinear Time Series: Nonparametric and Parametric Methods*, Springer-Verlag, New York 2003.

- Judge, G. G., Griffiths, W. E., Hill, R. C., Lütkepohl, H. and Lee, T.-C., *The Theory and Practise of Econometrics*, Second Edition, John Wiley&Sons, New York 1985.
- Kaishev, V.K., “A Computer Package for Solving Spline Regression”, T. Havranek, Z. Sidak and M. Novak (Ed.), *COMPSTAT 1984: Proceedings in Computational Statistics* (pp. 409-414), Wien: Physica-Verlag 1984.
- Kaki, B., Yeşilova, A. ve Şen, C., “Yarı Parametrik Regresyon Yönteminin Hayvancılıkta Kullanılması”, 2004, Erişim Tarihi: 11 Ocak 2012, http://4uzbk.sdu.edu.tr/4UZBK/BGB/4UZBK_097.pdf
- Kalyanam, K. and Shively, T. S., “Estimating Irregular Pricing Effects: A Stochastic Spline Regression Approach”, *Journal of Marketing Research*, 35(1), 1998, 16-29.
- Kaps, M. and Lamberson, W. R., *Biostatistics for Animal Science*, 2nd Edition, MPG Books Group, London 2009.
- Karadavut, U., Genç, A., Tozluca, A., Sinan, A., Karakoca, A., Aksoyak, Ş. ve Palta, Ç., “Parçalı Regresyon Yardımı ile Bitki Boyu-Zaman İlişkisi Parametrelerinin Tahmini”, *Sakarya Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, 8(1), 2004, 114-117.
- Karadavut, U., Genç, A., Tozluca, A., ve Palta, Ç., “Silajlık ve Danelik Mısırlarda Kuru Madde Birikiminin Bazı Matematiksel Büyüme Modelleri ile Analizi”, *Tarım Bilimleri Dergisi*, 16, 2010, 89-96.
- Karadavut, U., Sinan, A., Genç, A., Palta, Ç., Karakoca, A. ve Aksoyak, Ş., “Dağdaş Buğday Bitkilerinin Büyüme Eğrilerinin Belirlenmesinde Model Seçimi”, 4. *İstatistik Kongresi: 8-12 Mayıs 2005*, (ss.257-262), Belek/Antalya.
- Karadavut, U. ve Tozluca, A., “Çavdar Bitkisinde (*Secale Cereale* L.) Bazı Karakterlerin Büyüme Analizleri”, *Bitkisel Araştırma Dergisi*, 2, 2005, 1-10.
- Karakoca Aydın, *Regresyon Modellerinde Yapısal Değişiklik ve Model Seçimi*, (Yayımlanmamış Yüksek Lisans Tezi), Selçuk Üniversitesi Fen Bilimleri Enstitüsü, Konya 2003.

- Keele, L., *Semiparametric Regression for the Social Sciences*, John Wiley & Sons, London 2008.
- Khan, A., Khan, I. and Aziz, T., “A Survey on Parametric Spline Funtion Approximation”, *Applied Mathematics and Computation*, 171(2), 2005, 983-1003.
- Kim, H.-J., Fay, M. P., Feuer, E. J. ve Midthune, D. N., “Permutation Tests for Joinpoint Regression with Applications to Cancer Rates”, *Statistics in Medicine*, 19(3), 2000, 335-351.
- Kim, H.-J., Yu, B. And Feuer, E. J., “Selecting The Number of Change-Points in Segmented Line Regression”, *Statistica Sinica*, 19(2), 2009, 597-609.
- Koldere Yasemin, *Spline Fonksiyonları ve Yapısal Değişim: Türkiye İthalatı Üzerine Bir Uygulama*, (Yayımlanmamış Yüksek Lisans Tezi), Marmara Üniversitesi Sosyal Bilimler Enstitüsü, İstanbul 1997.
- Kommineni, B., Basu, S. and Vemuri, R., “A Spline Based Regression Technique on Interval Valued Noisy Data”, *Sixth International Conference on Machine Learning and Applications: 13-15 December 2007*, (pp.241-247), University of Cincinnati, Ohio/USA.
- Kommineni Balaji, *Spline Center and Range Regression and its Application to Variation Aware Performance Macromodeling of Analog Circuits*, University of Cincinnati, Andhra Pradesh/India 2005.
- Koo, J.Y., “Spline Estimation of Discontinuous Regression Functions”, *Journal of Comptational and Graphical Statistics*, 6(3), 1997, 266-284.
- Kruger, C. J. C., “Constrained Cubic Spline Interpolation for Chemical Engineering Applications”, Korf Technology, Erişim Tarihi: 011.2012, <http://www.korf.co.uk/spline.pdf>
- Kulkarni, B. S. and Narendranath, V. V., “Spline Regression Models for Forecasting of Agricultural Production”, *The Journal of Research Angrau*, 36(2&3), 2008, 28-38.
- Kurnaz, A., Kınacı, İ., Oturañç, G., Genç, A. ve Bilir, Ş., “Günlük Sıcaklık Verilerinin Kübik Spline Fonksiyonları Yardımıyla Modellenmesi”, H. K. Öztürk, M.

- Atılğan, A. Yılancı ve Ö. Atalay (Haz.), *I. Ege Enerji Sempozyumu ve Sergisi: 22-24 Mayıs 2003*, (ss.464-468), Pamukkale Üniversitesi Mühendislik Fakültesi, Denizli.
- Larson, H. J., “Least Squares Estimation of Linear Splines with Unknown Knot Locations”, *Computational Statistics & Data Analysis*, 13(1), 1992, 1-8.
- Lee, T. C. M., “On Algorithms for Ordinary Least Squares Regression Spline Fitting: A Comparative Study”, *Journal of Statistical Computation and Simulation*, 72(8), 2002, 647-663.
- Lerman, P.M., “Fitting Segmented Regression Models by Grid Search”, *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 29(1), 1980, 77-84.
- Lindstrom, M. J., “Penalized Estimation of Free-Knot Splines”, *Journal of Computational and Graphical Statistics*, 8(2), 1999, 333-352.
- Loague, K. and Green, R. E., “Statistical and Graphical Methods for Evaluating Solute Transport Models: Overview and Application”, *Journal of Contaminant Hydrology*, 7, 1991, 51-73.
- Ma, S., Racine, J. S. and Yang, L., *Spline Regression in the Presence of Categorical Predictors*, Working Paper, Michigan State University, East Lansing, Michigan 2011.
- Mao, W. and Zhao, L. H., “Free-Knot Polynomial Splines with Confidence Intervals”, *Journal Of The Royal Statistical Society Series B*, 65(4), 2003, 901-919.
- Marcel, B., Razvan, P. and Cristian, O., “Using Spline Functions for the Substantiation of Tax Policies by Local Authorities”, *Annals of the University of Oradea: Economic Science*, 1(1), 2011, 336-347.
- Marcus, R. D., “How to Deal with Structural Changes in the Data”, *The Journal of Business Forecasting Methods & Systems*, 6(9), Winter 1987/1988, 14-16.
- Markov Dennis, *Information Content in Stock Market Technical Patterns: A Spline Regression Approach*, Ph. D. Disertation, University of Notre Dame, Indiana 2002.

- Maronna, R. and Yohai, V. J., "A Bivariate Test for the Detection of A Systematic Change in Mean", *Journal of the American Statistical Association*, 73(363), 1978, 640-645.
- Mars, L. C., "On Estimating Spline Regression", SAS Institute Inc (Ed.), *Proceedings of the Eighth Annual SAS Users Group International Conference*, (pp. 723-728), New Orleans, Louisiana 1983.
- Marsh, L. C., "Estimating the Number of and Location of Knots in Spline Regression", *Journal of Applied Business Research*, 2(3), 1986, 60-70.
- Marsh, L.C., "Estimating Spline Knots in Time Series Polynomial Regression Models", *The Institute of Management Sciences and Operations Research Society of America*, (October 1987), St. Louis.
- Marsh, L.C. and Cormier, D.R., *Spline Regression Models*, Sage Publications, California 2002.
- Marsh, L.C. and Sindone, A.B., "A Calibration Technique for Estimating the Effect of Location on the Values of Residuals Properties", *The Property Tax Journal*, 10(2), 1991, 261-276.
- Marsh, L.C., Maudgal, M. and Raman, J., "Alternative Methods of Estimating Piecewise Linear and Higher Order Regression Models using SAS Software", SAS Institute Inc (Ed.), *Proceedings of the Fifteenth Annual SAS Users Group International Conference*, (pp. 523-527), Nashville, Tennessee 1990.
- Mathews, J. H., *Numerical Methods for Mathematics, Science, and Engineering*, Second Edition, Prentice Hall, New Jersey 1992.
- Mawk Russell, *A Survey of Applications of spline Functions to Statistics*, Master's Thesis, East Tennessee State University, Johnson City 2001.
- McDonald, G. C. and Studden, W.J., "Design Aspects of Regression-Based Ratio Estimation", *Technometrics*, 32(4), 1990, 417-424.
- McGee, Victor E. and Carleton, Willard T., "Piecewise Regression", *Journal of the American Statistical Association*, 65(331), 1970, 1109-1124.
- Mchta, H. and Capizzi, T., "Evaluating Linear Model Representations of Cubic Splines Using Proc Reg", SAS Institute Inc (Ed.), *Proceedings of the Seventh Annual SAS*

- Users Group International Conference*, (pp. 562-567), San Francisco, California 1982.
- MecNail, D. R., Trussell, T. J. and Turner, J. C., “Spline Interpolation of Demographic Data”, *Demography*, 14(2), 1977, 245-252.
- Mendenhall, W. and Sincich, T., *A Second Course in Statistics: Regression Analysis*, Fifth Edition, Prentice Hall, New Jersey 1996.
- Monahan, J. F., *Numerical Methods of Statistics*, Second Edition, Cambridge University Press, New York 2011.
- Moreton, H. P., *Minimum Curvature Variation Curves, Networks, and Surfaces for Fair Free-Form Shape Design*, Ph. D. Disertation, University of California, Berkeley 1992.
- Muggeo, V. M. R., “Estimating Regression Models with Unknown Break-Points”, *Statistics in Medicine*, 22, 2003, 3055-3071.
- Mulla, Z. D., “Spline Regression in Clinical Research”, *West Indian Med J.*, 56(1), 2007, 77-79.
- Muniz, I., Galindo, A. and Garcia, M.A., “Cubic Spline Population Density Functions City Delimitation: The Case of Barcelona”, *Urban Studies*, 40(7), 2003, 1303-1321.
- Muth, A.M.M. and Willsky, A.S., *A Sequential Method for Spline Approximation with Variable Knots*, Technical Report (ESL-P-759), M.I.T. Electronic Systems Laboratory, Mexico City 1977.
- Myers, R. H., *Classical and Modern Regression with Applications*, Second Edition, Duxbury Press, Pacific Grove, CA 1990.
- Oturanç, G., Kurnaz, A., Kiriş, M. E. ve Keskin, Y., *Sayısal Analiz*, 2.Baskı, Dizgi Ofset, Konya 2008.
- Özfallı Yasemin, *Çokdeğişkenli Uyarlanabilir Regresyon Kesitleri: MARS*, (Yayımlanmamış Yüksek Lisans Tezi), Gazi Üniversitesi Fen Bilimleri Enstitüsü, Ankara 2008.

- Özkan Nurhan, *Splayn Teorisi ve Diferensiyel Denklemlere Uygulamaları*, (Yayımlanmamış Yüksek Lisans Tezi), Ondokuz Mayıs Üniversitesi Fen Bilimleri Enstitüsü, Samsun 2006.
- Öztürk Elif, *Parametrik Olmayan ve Yarı Parametrik Panel Veri Modelleri: Çevresel Kuznets Eğrisinin Analizi*, (Yayımlanmamış Doktora Tezi), Marmara Üniversitesi Sosyal Bilimler Enstitüsü, İstanbul 2011.
- Şahin Mustafa, *Parçalı Regresyonlar ve Tarımsal Alanlarda Kullanımı*, (Yayımlanmamış Doktora Tezi), Kahramanmaraş Sütçü İmam Üniversitesi Fen Bilimleri Enstitüsü, Kahramanmaraş 2009.
- Park, S. H., “Experimental Designs for Fitting Segmented Polynomial Regression Models”, *Technometrics*, 20(2), 1978, 151- 154.
- Petunin, Y. I., Djomina, E. A., Klyushin, D. A. and Savkina, M. Y., “Spline-Regression Estimation of The Influence of Exposure Dose on Cancer Probability”, *Cybernetics and Systems Analysis*, 42(3), 2006, 449-455.
- Phillips, G.M. and Taylor, P.J., “Algorithms for Spline and Other Approximations to Functions and Data”, *Computer Physics Communications*, 73(1-3), 1992, 1-21.
- Pindyck, R. S. and Rubinfeld, D. L., *Econometric Models and Economic Forecasts*, Third Edition, Mcgraw-Hill, New York 1991.
- Poirier, D. J., “Piecewise Regression Using Cubic Spline”, *Journal of the American Statistical Association*, 68(343), 1973, 515-524.
- Poirier, D. J. and Garber, S. G., “The Determinants of Aerospace Profit Rates”, *Southern Economic Journal*, 41(2), 1974, 228-238.
- Poirier, D. J., “On the Use of Bilinear Splines in Economics”, *Journal of Econometrics*, 3(1), 1975, 23-34.
- Poirier, D. J., “On the Use of Cobb-douglas Splines”, *International Economic Review*, 16(3), 1975, 733-744.
- Prunty, L., “Curve Fitting with Smooth Functions that are Piecewise-Linear in The Limit”, *Biometrics*, 39(4), 1983, 857-866.
- Quandt, R. E., “The Estimation of the Parameters of a Linear Regression System

- Obeying Two Separate Regimes”, *Journal of the American Statistical Association*, 53(284), 1958, 873-880.
- Quandt, R. E., “Tests of the Hypothesis that a Linear Regression System Obeys Two Separate Regimes”, *Journal of the American Statistical Association*, 55(290), 1960, 324-330.
- Racine, J. S., “A Primer on Spline Regression”, 2011, Erişim Tarihi: 05 Eylül 2011, http://mirror.fcaglp.unlp.edu.ar/pub/CRAN/web/packages/crs/vignettes/spline_primer.pdf.
- Reinsch, C. H., “Smoothing by Spline Functions”, *Numerische Mathematik*, 10(3), 1967, 177-183.
- Robb, A.L., “Accounting for Seasonality with Spline Functions”, *The Review of Economics and Statistics*, 62(2), 1980, 321-323.
- Robison, D. E., “Estimates for the Points of Intersection of Two Polynomial Regressions”, *Journal of the American Statistical Association*, 59(305), 1964, 214-224.
- Ronda, J. and Oliver, G. J., *Introduction to Numerical Methods with Matlab Procedures*, 2010, Erişim Tarihi: 22.2.2012, http://home.agh.edu.pl/~pwojnaro/dydaktyka/mn/numerical_methods_book_2011_04_05.pdf
- Ruppert, D., “Selecting the Number of Knots for Penalized Splines”, *Journal of Computational and Graphical Statistics*, 11(4), 2002, 735-757.
- Ruppert, D., Wand, M. P. and Carroll, R. J., *Semi Parametric Regression*, Cambridge University Press, New York 2003.
- Ryan, T. P., *Modern Regression Methods*, John Wiley & Sons, New York 1997.
- Salomon, D., *Curves and Surfaces for Computer Graphics*, Springer Science+Business Media, New York 2006.
- Sarvida, R. L. and Baguio, C. B., “On the Efficiency of Spline Regression for Time Series Data” [Bildiri], *11th National Convention on Statistics*, (September 2010), Mandaluyong City: EDSA Shangri-La Hotel.

- Schoenberg, I. J., “Contributions to the Problem of Approximation of Equidistant Data by Analytic Functions”, *Quarterly of Applied Mathematics*, 4, 1946, 45–99 and 112–141.
- Schumaker, L. L., *Spline Functions: Basic Theory*, Third Edition, Cambridge University Press, New York 2007.
- Seber, G. A. F and Lee, A. J., *Linear Regression Analysis*, Second Edition, John Wiley & Sons, New Jersey 2003.
- Seber, G.A.F. and Wild, C.J., *Nonlinear Regression*, John Wiley&Sons, New Jersey 2003.
- Shikin, E. V. and Plis, A. I., *Handbook on Splines for the User*, CRC Press, Boca Raton, Florida 1995.
- Silverman, B. W., “A Fast and Efficient Cross-Validation Method for Smoothing Parameter Choice in Spline Regression”, *Journal of the American Statistical Association*, 79(387), 1984, 584-589.
- Silverman, B. W., “Some Aspects of the Spline Smoothing Approach to Non-Parametric Regression Curve Fitting”, *Journal of the Royal Statistical Society. Series B (Methodological)*, 47(1), 1985, 1-52.
- Smith, J., Smith, P. and Addiscott, T., “Quantitative Methods to Evaluate and Compare Soil Organic Matter (SOM) Models”, (Ed. David S. Powlson, Pete Smith and Jo U. Smith), *Evaluation of Soil Organic Matter Models*, (pp.181-199), Springer-Verlag, Berlin 1996.
- Smith, M. and Kohn, R., “Nonparametric Regression Using Bayesian Variable Selection”, *Journal of Econometrics*, 75(2), 1996, 317-344.
- Smith, P. L., “Splines as a Useful and Convenient Statistical Tool”, *The American Statistician*, 33(2), 1979, 57-62.
- Smith, P. L., *Curve Fitting and Modeling with Splines Using Statistical Variable Selection Techniques*, NASA Contractor Report 166034, Langley Research Center, Virginia 1982.
- Smith, R. E., Price, J. M. and Howser, L. M., *A Smoothing Algorithm Using Cubic*

- Spline Functions*, NASA Technical Note D-7397, National Aeronautics and Space Administration, Washington D.C. 1974.
- Speyrer, J. F. and Ragas, W. R., “Housing Prices and Flood Risk: An Examination Using Spline Regression”, *The Journal of Real Estate Finance and Economics*, 4(4), 1991, 395-407.
- Stone, C.J., Hansen, M., Kooperberg, C. and Truong, Y.K., “Polynomial Splines and Their Tensor Products in Extended Linear Modeling”, *The Annals of Statistics*, 25(4), 1997, 1371-1425.
- Strawczynski, M., “Social Insurance and The Optimal Piecewise Linear Income Tax”, *Journal of Public Economics*, 69(3), 1998, 371-388.
- Suits, D. B., “Use of Dummy Variables in Regression Equations”, *Journal of the American Statistical Association*, 52(280), 1957, 548-551.
- Suits, D. B., Mason, A. and Chan, L., “Spline Functions Fitted by Standart Regression Methods”, *Review of Economics and Statistics*, 60(1), 1978, 132-139.
- Sun, C. H. and Liu, K. E., “Information Asymmetry and Small Business in Online Auction Market”, *Small Business Economics*, 34(4), 2010, 433-444.
- Sung Cheing-Hsing, *Estimation of a Modified Linear Spline Regression: Theory and Application*, Ph. D. Disertation, Wayne State University, Detroit/USA 1985.
- Süli, E. and Mayers, D., *An Introduction to Numerical Analysis*, Cambridge University Press, New York 2003.
- Şahin Mustafa, *Parçalı Regresyonlar ve Tarımsal Alanlarda Kullanımı*, (Yayımlanmamış Doktora Tezi), Kahramanmaraş Sütçü İmam Üniversitesi Fen Bilimleri Enstitüsü, Kahramanmaraş 2009.
- Şener Emral, *Bazı Kırık Hat (Broken-Line) Regresyon Moellerinin SAS İstatistik Paket Programı Kullanılarak Çözümlemesi*, (Yayımlanmamış Yüksek Lisans Tezi), Mustafa Kemal Üniversitesi Fen Bilimleri Enstitüsü, Antakya 2006.
- Tarı, Recep, *Ekonometri*, İkinci Baskı, Alfa Basım Yayım Dağıtım Ltd. Şti., İstanbul 1999.
- Tarpey, T. and Holcomb, J., “Spline Bottles”, *The American Statistician*, 54(2), 2000,

129-135.

- Taşkın, M. M., “1923-2003 Döneminde Türkiye Cumhuriyeti’nin Dış Ticaret Politikaları”, 2003, Erişim Tarihi: 15 Eylül 2011, www.dtm.gov.tr/dtmadmin/upload
- Tishler, J. L., and Zang, I., “A New Maximum Likelihood Algorithm for Piecewise Regression”, *Journal of the American Statistical Association*, 76(376), 1981a, 980–987.
- Tishler, J. L., and Zang, I., “A Maximum Likelihood Method for Piecewise Regression Models With a Continuous Dependent Variable”, *Journal of the Royal Statistical Society (Applied Statistics)*, 30(2), 1981b, 116-124.
- Toms, J. D. and Lesperance, M. L., “Piecewise Regression: A Tool for Identifying Ecological Thresholds”, *Ecology*, 84(8), 2003, 2034-2041.
- Toraichi, K. and Kamada, M., “Knot Positions for the Smoothest Periodic Quadratic Spline Interpolation of Equispaced Data”, *Linear Algebra and Its Applications*, 221, 1995, 245-251.
- Tunç, H., *Uluslararası Ticaret Para ve Finans*, Alfa Kitabevi, İstanbul 2004.
- Ucal, M. Ş., “Ekonometrik Model Seçim Kriterleri Üzerine Kısa Bir İnceleme”, *C. Ü. İktisadi ve İdari Bilimler Dergisi*, 7(2), 2006, 41-57.
- Unser, M., “Splines: A Perfect Fit for Medicine Imaging”, Milan Sonka, J. Michael Fitzpatrick (Ed.), *Proceedings of the SPIE International Symposium on Medical Imaging* (pp. 225-236), Spie, San Diego 2002.
- Ünal Birce, *Çok Değişkenli Uyarlamalı Regresyon Uzanımları*, (Yayımlanmamış Yüksek Lisans Tezi), Hacettepe Üniversitesi Fen Bilimleri Enstitüsü, Ankara 2009.
- Wand, M.P., “A Comparison of Regression Spline Smoothing Procedures”, *Computational Statistics*, 15(4), 2000, 443-462.
- Wang, H. “*A Survey on Spline Theory and Applications*”, Center for Visual Computing (CVC) and Department of Computer Science, Stony Brook University, New York 2006, Erişim Tarihi: 12 Ekim 2011,

http://www.cs.sunysb.edu/~wanghy/papers/hongyu_rpe.pdf

Weatherspoone William, *Spline Analysis and an Application*, Master's Thesis, University of Texas, El Paso 1975.

Wegman, J. and Wright, W., "Splines in Statistics", *Journal of the American Statistical Association*, 78(382), 1983, 351-365.

Wold, S., "Spline Functions in Data Analysis", *Technometrics*, 16(1), 1974, 1-11.

Woods, D. and Lewis, S., "All-Bias Designs for Polynomial Spline Regression Models", *Australian & New Zealand Journal of Statistics*, 48(1), 2006, 49-58.

Wu, H. and Zhang, J.-T., *Nonparametric Regression Methods for Longitudinal Data Analysis Mixed Effects Modeling Approaches*, John Wiley&Sons, New Jersey 2006.

Xi, J. Y., Lee, C. L., Teng, H. L., and Lin, S. B., "Piecewise Regression: A Tool for Understanding The Relationship between the Attribute-Level Performance and Customer Satisfaction", *International Multiconference of Engineers and Computer Scientists*, (March 2009), Hong Kong.

Yao, F. and Lee, T. C. M., "On Knot Placement for Penalized Spline Regression", *Journal of the Korean Statistical Society*, 37(3), 2008, 259-267.

Yerdelen, F., "Gecikmeli Modellerde Spline Fonksiyonun Kullanımı", *İktisat Fakültesi Mecmuası*, 52(2), 2002, 47-64.

Zorşahin Melis, *Polinom ve Polinom Olmayan Kübik Spline Fonksiyonlar Yardımıyla Bazı Kısmi Diferensiyel Denklemlerin Sayısal Çözümleri*, (Yayımlanmamış Yüksek Lisans Tezi), Osmangazi Üniversitesi Fen Bilimleri Enstitüsü, Eskişehir 2009.

Erişim Tarihi: 15 Nisan 2011, www.tuik.gov.tr

Erişim Tarihi: 03 Nisan 2012, <http://tr.wikipedia.org/wiki/Epidemiyoloji>

Erişim Tarihi: 27 Haziran 2012, http://en.wikipedia.org/wiki/Mean_absolute_error

Erişim Tarihi: 4 Temmuz 2012, <http://www.iue.tuwien.ac.at/phd/khalil/node31.html>

EKLER

Ek 1. Türkiye’de 1923-2010 Yılları Arası İhracatın İthalatı Karşılama Oranları

Yıllar	İhr./İth. Oranı	Yıllar	İhr./İth. Oranı	Yıllar	İhr./İth. Oranı	Yıllar	İhr./İth. Oranı	Yıllar	İhr./İth. Oranı
1923	58,47	1941	164,51	1959	75,28	1977	30,24	1995	60,59
1924	82,06	1942	111,73	1960	68,51	1978	49,75	1996	53,23
1925	79,64	1943	126,65	1961	68,36	1979	44,60	1997	54,08
1926	79,43	1944	140,97	1962	61,54	1980	36,79	1998	58,74
1927	74,94	1945	173,52	1963	53,53	1981	52,64	1999	65,37
1928	77,63	1946	180,49	1964	76,46	1982	64,98	2000	50,96
1929	60,56	1947	91,28	1965	81,08	1983	62,02	2001	75,69
1930	102,65	1948	71,55	1966	68,29	1984	66,32	2002	69,94
1931	100,49	1949	85,39	1967	76,29	1985	70,16	2003	68,15
1932	117,82	1950	92,21	1968	65,01	1986	67,15	2004	64,76
1933	128,77	1951	78,11	1969	67,00	1987	71,97	2005	62,92
1934	106,18	1952	65,28	1970	62,10	1988	81,35	2006	61,28
1935	107,92	1953	74,37	1971	57,79	1989	73,61	2007	63,08
1936	127,24	1954	70,02	1972	56,64	1990	58,11	2008	65,37
1937	120,64	1955	62,97	1973	63,13	1991	64,59	2009	72,48
1938	96,74	1956	74,87	1974	40,56	1992	64,34	2010	61,42
1939	107,73	1957	86,93	1975	29,57	1993	52,14		
1940	161,69	1958	78,47	1976	38,22	1994	77,81		

Kaynak: TÜİK

ÖZGEÇMİŞ

Kişisel Bilgiler	
Adı ve Soyadı	Ömer ALKAN
Doğum Yeri ve Tarihi	Yusufeli – 057.1982
Eğitim Durumu	
Lisans Öğrenimi	Atatürk Üniversitesi İktisadi ve İdari Bilimler Fakültesi İktisat Bölümü
Y.Lisans Öğrenimi	Atatürk Üniversitesi Sosyal Bilimler Enstitüsü İşletme Anabilim Dalı Sayısal Yöntemler Bilim Dalı
Bildiği Yabancı Diller	İngilizce
Bilimsel Faaliyetler	Kitaplar * OKTAY, E., ÖZEN, Ü., ALKAN, Ö. (2007). <i>Erzurum'da Kredi Kartı Sahipliğini Belirleyen Faktörlerin Araştırılması</i>. Erzurum: Atatürk Üniversitesi Yayınevi. * OKTAY, E., KIZILTAN, A., KÜÇÜKALİ, A., ALKAN, Ö. (2011). <i>Kredi Kartı Sahipliğinde Etkili Olan Sosyo-Ekonomik Faktörlerin Analizi: Atatürk Üniversitesi Öğrencileri Üzerine Bir Uygulama</i>. Erzurum: Eser Ofset Matbaacılık. Makaleler * OKTAY, E., ÖZEN, Ü., ALKAN, Ö. (2009) "Kredi Kartı Sahipliğinde Etkili Olan Faktörlerin Araştırılması: Erzurum Örneği". <i>Dokuz Eylül Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi</i>, 24(2), 1-22 * ELMAS, B., YAKUT, E., ALKAN, Ö. (2011). "İşletmelerin Mali Başarısızlığının Yapay Sinir Ağları ve Lojistik Regresyon Modeli ile Tahmin Edilmesi". <i>Finans Politik & Ekonomik Yorumlar</i>, 48(560), 45-56 * GENÇ, A., OKTAY, E., ALKAN, Ö. (2012). "İhracatın İthalatı Karşılama Oranlarının Parçalı Regresyonlarla Modellenmesi". <i>Atatürk Üniversitesi Sosyal Bilimler Enstitüsü Dergisi</i>, 16(1), 1-15 * OKTAY, E., TALAŞ, E., ALKAN, Ö., GENÇ, A. (2012) "Modeling with Linear Spline Regression of Turkish Tourism Demand". <i>Journal Of Selçuk University Natural And Applied Science</i>, 1(4), 10-22

	Bildiri/Sempozyum/Kongre * OKTAY, E., ALKAN, Ö. "Temel Bileşenler Analizi Yöntemiyle Toplu Taşıma Araçlarından Memnuniyeti Etkileyen Faktörlerin Belirlenmesi: Erzurum İli Örneği", 11. Ekonometri ve İstatistik Sempozyumu, 28-30.2010 Sakarya. * GENÇ, A., ALKAN, Ö. "İhracatın İthalatı Karşılama Oranlarının Parçalı Regresyonlarla Modellenmesi", III. Ulusal Ereğli Kemal Akman Meslek Yüksek Okulu Tebliğ Günleri, 28-29 Nisan 2011, Ereğli-KONYA . * ELMAS, B., ALKAN, Ö., YAKUT, E. "Finansal Başarısızlığın Tahmin Edilmesinde Yapay Sinir Ağları ve Lojistik Regresyon Modelinin Karşılaştırılması", 12. International Symposium On Econometrics Statistics and Operations Research, 26-29.2011 Denizli. * OKTAY, E., ÖZEN, Ü., ALKAN, Ö. "İnternet Bankacılığını Tercih Etkili Olan Faktörlerin Analizi: Atatürk Üniversitesi Örneği", 12. International Symposium On Econometrics Statistics and Operations Research, 26-29.2011 Denizli. * OKTAY, E., TALAŞ, E., ALKAN, Ö., ELMALI, K. "Doğrusal Spline Regresyon ile Turizm Talebinin Analizi", II. Disiplinlerarası Turizm Araştırmaları Kongresi, 12-15.2012 Kemer/ANTALYA. * GENÇ, A., OKTAY, E., ALKAN, Ö. "Spline Regresyonda Düşüm Noktalarının Tahmini Üzerine Bir Çalışma", I.Uluslararası Iğdır Sempozyumu, 19-21.2012 Iğdır. * GENÇ, A., OKTAY, E., ALKAN, Ö. "Erzurum İl Merkezindeki Elektrik Tüketiminin Spline Regresyon Modellerle Analizi", 13th International Conference on Econometrics, Operations Research and Statistics, 24-26 Mayıs 2012 Kıbrıs. *OKTAY, E., KARAASLAN, A., ALKAN, Ö. "Konut Talebinde Etkili Olan Faktörlerin Analizi: Erzurum İl Merkezi'nde Bir Saha Araştırması", 13th International Conference on Econometrics, Operations Research and Statistics, 24-26 Mayıs 2012 Kıbrıs.
--	---

	*OKTAY, E., ALKAN, Ö., KARAASLAN, A. "Multinomial Logit Model Yardımıyla Konut Talep Etme Gerekçeleri Üzerinde Etkili Olan Faktörlerin Analizi", 13th International Conference on Econometrics, Operations Research and Statistics, 24-26 Mayıs 2012 Kıbrıs. *OKTAY, E., GENÇ, A., ALKAN, Ö., ÇELİK, A. "Investigation of Export-Import Ratios in Turkey By Using Spline Regression Models", The 6th International Conference on Applied Statistics, 16-17 November 2012, Bucharest, Romania.
İş Deneyimi	
Projeler	Spline Regresyon Modelleri ve Bir Uygulama, Atatürk Üniversitesi Projesi, BAP: 2011/446
Çalıştığı Kurumlar	*TC Ziraat Bankası İç Kontrolörü (2006-2010) *Atatürk Üniversitesi İktisadi ve İdari Bilimler Fakültesi İşletme Bölümü Araştırma Görevlisi (2010-...)
İletişim	
E-Posta Adresi	oalkan@atauni.edu.tr
Tarih	16.01.2013