

MİMAR SİNAN GÜZEL SANATLAR ÜNİVERSİTESİ ★
FEN BİLİMLERİ ENSTİTÜSÜ

**FONKSİYONEL REGRESYON MODELLERİNDE KİSMİ EN KÜÇÜK
KARELER YÖNTEMİ ÜZERİNE YENİ YAKLAŞIMLAR**

DOKTORA TEZİ

Hazırlayan: Semanur SARIÇAM

Anabilim Dalı: İSTATİSTİK

Programı: İSTATİSTİK

Tez Danışmanı: Prof. Dr. Barış AŞIKGİL

Eş Danışman: Doç. Dr. Ufuk BEYAZTAŞ

ARALIK 2022

MİMAR SİNAN GÜZEL SANATLAR ÜNİVERSİTESİ ★
FEN BİLİMLERİ ENSTİTÜSÜ

**FONKSİYONEL REGRESYON MODELLERİNDE KİSMİ EN KÜÇÜK
KARELER YÖNTEMİ ÜZERİNE YENİ YAKLAŞIMLAR**

DOKTORA TEZİ

Hazırlayan: Semanur SARIÇAM

Anabilim Dalı: İSTATİSTİK

Programı: İSTATİSTİK

Tez Danışmanı: Prof. Dr. Barış AŞIKGİL

Eş Danışman: Doç. Dr. Ufuk BEYAZTAŞ

ARALIK 2022

..... tarafından hazırlanan adlı bu
tezin tezi olarak uygun olduğunu onaylarım.

.....
Tez Yöneticisi

Bu çalışma, jürimiz tarafından Anabilim Dalında
..... tezi olarak kabul edilmiştir.

Başkan : _____

Üye : _____

Üye : _____

Üye : _____

Üye : _____

Bu tez, Mimar Sinan Güzel Sanatlar Üniversitesi Enstitüsü tez
yazım kurallarına uygundur.

Mimar Sinan Güzel Sanatlar Üniversitesi Lisansüstü Tez Yazım Kılavuzuna uygun olarak hazırladığım bu tez çalışmada;

- Tez içindeki bütün bilgi ve belgeleri akademik kurallar çerçevesinde elde ettiğimi,
- Görsel, işitsel ve yazılı tüm bilgi ve sonuçları bilimsel etik kurallarına uygun olarak sunduğumu,
- Başkalarının eserlerinden yararlanılması durumunda ilgili eserlere bilimsel normlara uygun olarak atıfta bulunduğumu,
- Atıfta bulunduğum eserlerin tümünü kaynak olarak gösterdiğimi,
- Kullanılan verilerde herhangi bir değişiklik yapmadığımı,
- Ücret karşılığı başka kişilere yazdırmadığımı (dikte etme dışında), uygulamalarımı yaptırmadığımı,
- Bu tezin herhangi bir bölümünü bu üniversite veya başka bir üniversitede başka bir tez çalışması olarak sunmadığımı

beyan ederim.

ÖNSÖZ

Yüksek lisanstan doktora eğitim süreci boyunca tanıdığım Mimar Sinan Güzel Sanatlar Üniversitesi Fen - Edebiyat Fakültesi İstatistik Bölümü Öğretim Üyesi Prof. Dr. Barış AŞIKGİL'e ve proje aracılığıyla tanıma fırsatı bulduğum eş danışmanım Marmara Üniversitesi Fen Fakültesi İstatistik Bölümü Öğretim Üyesi Doç. Dr. Ufuk BEYAZTAŞ'a tez çalışması boyunca beni teşvik ettikleri için ve hiçbir zaman ilgi ve desteklerini esirgemedikleri için ayrıca bu alanla tanışma fırsatına vesile oldukları için sonsuz şükranlarımı sunarım.

Tez Savunma Jüri Hocalarımdan Mimar Sinan Güzel Sanatlar Üniversitesi Fen - Edebiyat Fakültesi İstatistik Bölümü Öğretim Üyeleri Prof. Dr. Gülay BAŞARIR'a, ve Prof. Dr. Eylem DENİZ HOWE'a, Marmara Üniversitesi Fen Fakültesi İstatistik Bölümü Öğretim Üyesi Deniz İNAN'a ve Yıldız Teknik Üniversitesi Fen – Edebiyat Fakültesi İstatistik Bölümü Öğretim Üyesi Doç. Dr. Atıf Ahmet EVREN'e değerli ve kıymetli yorum ve görüşleri için teşekkürü bir borç bilirim.

Yüksek lisanstan itibaren tanıdığım Mimar Sinan Güzel Sanatlar Üniversitesi Fen - Edebiyat Fakültesi İstatistik Bölümü Öğretim Üyeleri Prof. Dr. Aydın ERAR'a, Doç. Dr. Ayça ÇAKMAK PEHLİVANLI'ya, Öğr. Gör. Dr. Damla İLTER'e, Arş. Gör. Metin YANGIN'a, Arş. Gör. Turgut ÖZALTINDIŞ'e ve Arş. Gör. Zeynep BAL'a güzel yorumları ve destekleri için teşekkür ederim.

Proje vasıtasıyla tanıma fırsatı bulduğum Marmara Üniversitesi Fen Fakültesi İstatistik Bölümü Öğretim Üyeleri Prof. Dr. Müjgan TEZ'e, Prof. Dr. Birsen EYGİ ERDOĞAN'a, Prof. Dr. Ufuk YOLCU'ya ve Doç. Dr. Özge CAGCAG YOLCU'ya her zaman değerli görüşleri için en candan duygularıyla teşekkür ederim.

Bu tez 120F270 No'lu Türkiye Bilimsel ve Teknolojik Araştırma Kurumu (TÜBİTAK) projesi kapsamında gerçekleştirilmiştir. Tüm katkıları için TÜBİTAK'a teşekkür ederim.

Son olarak en büyük minnettarlığım, maddi ve manevi olarak desteklerini esirmeyen beni bugünlere getiren her zaman yanımda olan sevgili babama, anneme, ablama ve kardeşlerime ve beni bu çalışmamda motive ederek yardımlarını esirgemeyen değerli dostlarıma sonsuz teşekkür ederim.

Yeğenim Rüzgar Alp'e...

FONKSİYONEL REGRESYON MODELLERİNDE KİSMİ EN KÜÇÜK KARELER YÖNTEMİ ÜZERİNE YENİ YAKLAŞIMLAR

ÖZET

Teknolojik gelişmeler, gözlemleri zaman, uzay ve grid noktaları gibi bir süreklilik üzerinden elde edilen karmaşık ve yüksek boyutlu veri yapılarına yol açmıştır. Geleneksel matematiksel/istatistiksel yöntemler, çoklu bağlantı gibi yüksek boyutlu verilerdeki yaygın problemler nedeniyle bu tür verileri analiz edememektedir. Bu nedenle, fonksiyonel verileri analiz etmek için geçerli istatistiksel yöntemlere duyulan ihtiyaç artmıştır. Sonuç olarak, fonksiyonel veri analizine (FDA) olan ilgi yıllar içinde önemli ölçüde artmış ve kimya, meteoroloji, tıp ve finans gibi çeşitli bilim alanlarında genel bir çerçeve haline gelmiştir.

FDA'da, yanıt ve tahmin değişkenleri arasındaki fonksiyonel bağlantıyı araştırmak için, yanıt veya tahmin değişkenlerinden en az birinin sonsuz boyutlu rastgele eğrilerden oluştuğu fonksiyonel doğrusal regresyon modeline (FLRM) özel bir ilgi gösterilmiştir. FLRM'deki temel amaç, sonsuz boyutlu regresyon katsayısı fonksiyonunu tahmin etmektir. Bu tahmin yöntemleri maksimum olabilirlik, maksimum cezalandırılmış olabilirlik, en küçük kareler, temel bileşen regresyonu ve kısmi en küçük kareler regresyonudur. Diğerleri arasında, kısmi en küçük kareler, gelişmiş parametre tahmini ve tahmin doğruluğu üretir. Bununla birlikte, hemen hemen tüm mevcut fonksiyonel kısmi en küçük kareler yöntemleri, SIMPLS ve NIPALS gibi olağan algoritmalara dayanmaktadır. Bu algoritmalar sezgisel ve tutarlı olmakla birlikte, kararlı parametre tahminleri üretmeyebilirler ve/veya yüksek boyutlu fonksiyonel verileri analiz ederken hesaplama açısından yoğun olabilirler.

Bu tezde amacımız, skaler - fonksiyon regresyon altında regresyon katsayısı fonksiyonunu verimli bir şekilde tahmin etmek için iki hibrit fonksiyonel kısmi en küçük kareler yöntemi önermektir. Önerilen yöntemlerde, sonsuz boyutlu fonksiyonel tahminciler ilk önce bir temel genişletme tekniği kullanılarak sonlu boyutlu bir uzaya yansıtılır. Ardından, skorun yeniden ortogonalizasyonuna ve vektör yüklemesine dayanan iki kısmi en küçük kareler algoritması, skaler yanıt ile fonksiyonel tahmin edicilerin temel katsayıları arasındaki doğrusal ilişkiyi tahmin etmek için kullanılır. Sonlu örneklem performansı ve hesaplama hızı için bir dizi Monte Carlo simülasyon çalışması ve bir şeker süreci veri seti kullanılarak sonuçlar, mevcut fonksiyonel kısmi en küçük kareler yöntemleriyle karşılaştırıldığında, önerilen yöntemlerin daha az hesaplama süresi ile kararlı parametre tahminleri ve öngörüler ürettiğini ortaya koyuldu.

Anahtar Kelimeler: Fonksiyonel veri analizi, skaler - fonksiyon regresyon modeli, fonksiyonel kısmi en küçük kareler, ortogonalleştirme, SIMPLS, NIPALS.

NEW APPROACHES TO THE PARTIAL LAST SQUARE METHOD IN FUNCTIONAL REGRESSION MODELS

ABSTRACT

Technological developments have led to complex and high-dimensional data structures whose observations are obtained over a continuum, such as time, space, and grid points. Traditional mathematical/statistical methods fail to analyze such kind of data because of common problems in high-dimensional data such as multicollinearity. Therefore, the need for valid statistical methods to analyze functional data has increased. As a result, the interest in functional data analysis (FDA) has increased significantly over the years, and it has become a general framework in various fields of science, such as chemometry, meteorology, medicine, and finance.

In FDA, a particular interest has been devoted to the functional linear regression model (FLRM), in which at least one of the response or predictor variables consist of infinite-dimensional random curves, to explore the functional connectivity between response and predictor variables. The main aim in FLRM is to estimate the infinite-dimensional regression coefficient function. Several estimation strategies, such as, maximum likelihood, maximum penalized likelihood, least squares, principal component regression, and partial least squares regression. Among others, the partial least squares produces improved parameter estimation and prediction accuracy. However, almost all the available functional partial least squares methods are based on the usual algorithms, such as SIMPLS and NIPALS. While these algorithms are intuitive and consistent, they may not produce stable parameter estimates and/or they may be computationally intensive when analyzing high-dimensional functional data.

In this thesis, our aim is to propose two modified functional partial least squares methods to efficiently estimate the regression coefficient function under the scalar-on-function regression. In the proposed methods, the infinite-dimensional functional predictors are first projected onto a finite-dimensional space using a basis expansion method. Then, two partial least-squares algorithms, based on re-orthogonalization of the score and loading vectors, are used to estimate the linear relationship between scalar response and the basis coefficients of the functional predictors. The finite-sample performance and computing speed are evaluated using a series of Monte Carlo simulation studies and a sugar process dataset. Our results revealed that, compared with available functional partial least squares methods, our proposed methods produce stable parameter estimates and predictions with less computing time.

Key Words: Functional data analysis, scalar-on-function regression model, functional partial least squares, bidiagonalization, SIMPLS, NIPALS.

İÇİNDEKİLER

	Sayfa
ÖNSÖZ	v
ÖZET	vi
SUMMARY	vii
İÇİNDEKİLER	viii
ÇİZELGE LİSTESİ.....	x
ŞEKİL LİSTESİ.....	xi
KISALTMALAR	xii
1. GİRİŞ	1
2. FONKSİYONEL VERİ ANALİZİ	8
2.1 Matematiksel Bazı Gösterimler ve İsimlendirme	9
2.1.1 Hilbert Uzayları	9
2.1.2 Hilbert Uzaylarında Operatörler	10
2.1.3 L^2 Uzayı	11
2.1.4 Stokastik Süreçler	12
2.1.5 Karhunen - Loeve Teoremi	13
2.2 Temel Açılımlar İçin Genişletme Teknikleri	14
2.2.1 Fourier Temel Açılımı	15
2.2.2 B-Spline Temel Açılımı	15
2.2.3 Diğer Faydalı Temel Açılımlar	16
3. FONKSİYONEL DOĞRUSAL REGRESYON	17
3.1 Fonksiyonel Doğrusal Regresyon Modelleri.....	17
3.1.1 Skaler – Fonksiyon Doğrusal Regresyon Modelleri	17
3.1.2 Fonksiyon – Fonksiyon Doğrusal Regresyon Modelleri	19
3.1.3 Fonksiyon – Skaler Doğrusal Regresyon Modelleri	22
3.2 Tahmin Yöntemleri	23
3.2.1 En Küçük Kareler Yöntemi	23
3.2.2 En Çok Olabilirlik Yöntemi	24
3.2.3 Cezalandırılmış En Çok Olabilirlik Yöntemi	25

4. FONKSİYONEL DOĞRUSAL REGRESYON MODELİNDE KİSMİ EN KÜÇÜK KARELER ÜZERİNE YENİ YAKLAŞIMLAR	27
4.1 Temel Genişletme Kullanılarak Regresyon Katsayısı Tahmini.....	31
4.2 Çift Köşegenleştirme ve PLS	33
5. UYGULAMA	41
5.1 Monte – Carlo Deneyleri	41
5.2 Şeker İşlem Verisi	25
6. SONUÇ VE ÖNERİLER.....	50
KAYNAKLAR	52
EKLER	57
ÖZGEÇMİŞ.....	67

ÇİZELGE LİSTESİ

Sayfa

Çizelge 5.1 : Sayısal kesinlik karşılaştırması: Hesaplanan MSE ve MISE değerleri. **43**

Çizelge 5.2 : Şekere işlem verileri için PLS yöntemlerinin MSE değerleri. **48**



ŞEKİL LİSTESİ

- Şekil 5.1** : Oluşturulan skaler yanıtın, fonksiyonel tahmin edicinin ve regresyon parametre fonksiyonunun grafiksel gösterimi. **42**
- Şekil 5.2** : K , h , ve n değerlerini artırmak için SIMPLS, NIPALS ve Bidiag2 tabanlı FPLS Yöntemleri için tahmini hesaplama süreleri. **44**
- Şekil 5.3** : Uyarım dalga boylarında $[230,240,255,290,305,325,340] = [X_1(t), X_2(t), X_3(t), X_4(t), X_5(t), X_6(t), X_7(t)]$ elde edilen 268 kül içeriği ve şeker numunelerinin emisyon spektrumlarının grafiksel gösterimi. **46**
- Şekil 5.3** : $[K, h] = [25, 15]$ (birinci satır) ve $[K, h] = [250, 50]$ (ikinci satır) olduğunda $\langle X(t), \hat{B}(t) \rangle$ 'ye karşı Y grafiği. **49**

KISALTMALAR

Bidiag1	: Alt Çift Köşegenleştirme
Bidiag2	: Üst Çift Köşegenleştirme
FDA	: Fonksiyonel Veri Analizi
FLR	: Fonksiyonel Doğrusal Regresyon
FLRM	: Fonksiyonel Doğrusal Regresyon Modeli
FoFRM	: Fonksiyon – Fonksiyon Regresyon Modeli
FoSRM	: Fonksiyon – Skaler Regresyon Modeli
FPC	: Fonksiyonel Temel Bileşen
FPLS	: Fonksiyonel Kısmi En Küçük Kareler
LS	: En Küçük Kareler
ML	: En Çok Olabilirlik
NIPALS	: Doğrusal Olmayan Yinelemeli Kısmi En Küçük Kareler
PLS	: Kısmi En Küçük Kareler
PML	: Cezalandırılmış En Çok Olabilirlik
SIMPLS	: PLS yönteminin İstatistiksel olarak İlham Edilen Modifikasyonunun Doğrudan Uygulanması
SoFRM	: Skaler – Fonksiyon Regresyon Modeli

1. GİRİŞ

Fonksiyonel Veri Analizi (Functional Data Analysis, FDA) sürekli olarak gözlemlenen fonksiyonların eğrileri üzerindeki bilgilerin analizi ile ilgilenmektedir. Veri toplamada ve depolamadaki teknolojik gelişmelerden dolayı FDA, istatistikte önem kazanmakta olup ekonometri, finans, meteoroloji, biyoloji, jeoloji, psikoloji, tıp, kemometri, spor bilimleri gibi pek çok alanda kullanımı artmıştır (Ramsay ve Silverman, 1997; Essomba, 2015; Beyaztaş ve Shang, 2020a). Ramsay ve Silverman (1997, 2002), Ferraty ve Vieu (2006), Horvath ve Kokoszka (2012) ve Cuevas (2014) FDA araçlarında bir dizi teorik gelişme ve uygulamayı tartışmışlardır.

Fonksiyonel veri analizinin amaçları, aslında diğer herhangi bir istatistik veri analizi araçlarının amaçlarıyla aynı olmakla birlikte, Ramsay ve Silverman (2005), bu temel amaçları aşağıdaki şekilde tanımlamışlardır:

- Verilerin daha fazla analize yardımcı olacak şekillerde dönüştürülmesi,
- Verinin karakteristik özelliklerini görselleştirmek için vurgulanması,
- Veriler arasındaki önemli örüntü ve değişim kaynaklarının incelenmesi,
- Bağımlı veya bağımsız değişken bilgisi kullanılarak bağımlı değişkendeki değişimin açıklanması,
- Aynı ya da farklı fonksiyonların yer aldığı iki veya daha fazla veri kümesini belirli varyasyon türlerine göre karşılaştırılması.

Bu bağlamda, klasik istatistikte boyut indirgeme, sınıflama, modelleme amacıyla kullanılan temel bileşenler analizi, diskriminant analizi, regresyon analizi gibi bazı yöntemler fonksiyonel forma uyarlanarak analiz edilmektedir.

Fonksiyonel veri analizi kavramı ilk olarak Ramsay (1982) ve Ramsay ve Dalzell (1991) tarafından ortaya atıldığı bilinmesine rağmen, aslında bu alan oldukça eskiye dayanmaktadır (Grenander, 1950; Rao, 1958; Wang ve diğ., 2015). Ramsay (1982),

klasik istatistiksel teknikleri fonksiyonel veriye uyarlayarak tanımlayıcı istatistikler, en küçük kareler yaklaşımı, temel bileşenler analizi ve kanonik korelasyon analizi yöntemlerini geniş çerçevede açıklamıştır. Besse ve Ramsay (1986), fonksiyonlardan oluşan veri setinde klasik temel bileşenler analizinin türev bilgisi dahilinde satır uzayının metrik uzaya eşdeğer olduğunu, Ramsay ve Dalzell (1991), çok değişkenli veri analizi araçlarından doğrusal regresyon analizi ve temel bileşenler analizi yöntemlerinin sonsuz boyutlu uzaya nasıl uygulanabileceğini göstermişlerdir. Silverman (1996), fonksiyonel temel bileşenler analizinde genellikle düzgünleştirme kullanıldığını belirtmiş ve bunun için pürüzlülüğü hesaba katan yeni bir düzgünleştirme parametresi ile bir yaklaşım önermiştir. Manteiga ve Vieu (2007), fonksiyonel verilerin faktöriyel analizi, fonksiyonel değişkenlerle regresyon ve eğrilerin sınıflandırılması, fonksiyonel veri analizi ve parametrik olmayan istatistikler hakkında detaylı bilgi vermişlerdir. Shang (2014), fonksiyonel temel bileşenler analizinin bir incelemesini ve bunun açıklayıcı analizde, modellemede, tahminde ve fonksiyonel verilerin sınıflandırılmasında kullanımını önermiştir.

Değişkenlerin fonksiyonel olma ve olmama durumlarında iki değişken arasındaki ilişkiyi açıklayabilmek için fonksiyonel doğrusal regresyon modeli (Functional Linear Regression Model, FLRM) ortaya atılmıştır. Öncelikle hem bağımlı değişkenin hem de bağımsız değişkenin fonksiyonel formdaki modeli ileri sürülmüştür. Ardından bağımsız değişkenin fonksiyonel, bağımlı değişkenin skaler ve bağımlı değişkenin fonksiyonel, bağımsız değişkenin skaler formdaki regresyon modelleri ele alınmıştır (Ramsay ve Dalzell, 1991; Ramsay ve Silverman, 1997; Cardot ve diğ., 1999). Escabias ve diğ. (2004), fonksiyonel lojistik regresyon modeli ile parametre fonksiyonunu tahmin etmiştir. İkili yanıt değişkenin çoklu doğrusallık ve yüksek boyutluluk gibi problemleri beraberinde getirdiğinden dolayı fonksiyonel lojistik regresyonun temel bileşenler analizi ile birlikte kullanılması önerilmiştir (Escabias ve diğ., 2005). Literatürde fonksiyonel lojistik regresyon üzerinde pek çok çalışma yer almaktadır (Ratcliffe ve diğ., 2002; Aguilera ve diğ., 2008).

FDA araçlarından biri olan fonksiyonel doğrusal regresyon modeli (Functional Linear Regression Model, FLRM) literatürde büyük bir ilgi görmüştür. Bu kapsamda

parametre tahminlerinde pek çok yöntem yer almaktadır ve ilk olarak Ramsay ve Silverman (1997) tarafından, B-spline, Fourier, Gauss ve dalgacık tabanlı genel genişletme teknikleri önerilmiştir. Daha sonra, Marx ve Eilers (1999), bağımlı değişkenin skaler, bağımsız değişkenin fonksiyonel formda olduğu modelde (SoFRM) fonksiyonel nesnelere yansıtılmak için B-spline tabanını kullanmışlar ve ardından, regresyon parametresini tahmin etmek için P-spline çerçevesi ile cezalandırılmış log-olasılığını kullanmışlardır. B-spline tabanını ve Ridge regresyonu kullanan Cardot ve diğ. (2003), $\beta(t)$ regresyon parametre fonksiyonu için bir tahmin edici önermişlerdir. Matsui ve diğ. (2008), SoFRM'nin çok değişkenli versiyonunu tahmin etmek için maksimum olabilirlik (Maximum Likelihood, ML) ve cezalandırılmış maksimum olabilirlik (Penalized Maximum Likelihood, PML) yöntemleriyle birlikte radyal temel fonksiyonları değerlendirmişlerdir. Goldsmith ve diğ. (2011), SoFRM için fonksiyonel tahmin edicinin B-spline bazında yaklaşık olduğu ve regresyon parametre fonksiyonunun cezalı spline regresyon kullanılarak tahmin edildiği hızlı-uyumlu (fast-fitting) yöntemler geliştirmişlerdir. Zhao ve diğ. (2012), fonksiyonel tahmin ediciyi sonlu boyutlu uzaya yansıtılmak için bir dalgacık temeli ve regresyon parametre fonksiyonunu yaklaşık olarak tahmin etmek için bir LASSO (Least Absolute Shrinkage and Selection Operator, En Küçük Mutlak Küçülme ve Seçim Operatörü) tahmincisi kullanmıştır. Bununla birlikte, genel temel fonksiyon genişletme yöntemleri, fonksiyonel regresyon katsayısına yaklaşmak için çok sayıda temel fonksiyon gerektirebilir, bu da modelin aşırı uyumuna ve düşük tahmin doğruluğuna neden olmaktadır. Ek olarak, LS (Least Squares, En Küçük Kareler) ve ML gibi geleneksel tahmin tekniklerinin kullanımı, çoklu bağlantı sorunu nedeniyle kararsız tahminler üretebilir (Matsui ve diğ., 2009). Ayrıca PML, cezalı spline regresyon ve Ridge regresyon gibi düzleştirme tekniklerine dayalı tahmin yöntemlerinin kullanılması hesaplama süresini önemli ölçüde uzatmaktadır (Beyaztaş ve Shang, 2020b).

Fonksiyonel temel bileşenler (Functional Principal Component, FPC) regresyonu, SoFRM'deki fonksiyonel nesnelere sonlu boyutlu ortonormal FPC tabanlarına yansıtılmaktadır. FPC regresyonu, ortogonal gizli bileşenleri çıkarırken fonksiyonel

tahminciler arasındaki deneysel kovaryansı kullanan veri odaklı bir yöntemdir. Başka bir deyişle, FPC'leri oluştururken ampirik kovaryans fonksiyonlarından çıkarılan fonksiyonel tahmin edicilerin bilgileri kullanılır, bu da FPC'yi genel temel fonksiyon genişletme tekniklerinden daha bilgilendirici yapar. FPC regresyon yöntemi ortogonal gizli bileşenler ürettiği için, SoFRM'deki çoklu bağlantı problemini atlar ve hesaplama süresi açısından verimli parametre tahminleri üretebilir. Bu nedenle, SoFRM'deki birçok tahmin yöntemi, FPC regresyonuna dayalı olarak oluşturulmuştur (Yao, 2007; Hall ve Horowitz, 2007; Lee ve Park, 2012; Maronna ve Yohai, 2013; Goldsmith ve Scheipl, 2014; Wang ve diğ., 2016; Shin ve Lee, 2016; Reiss ve diğ., 2017; Kalogridis ve Aelst, 2019). Öte yandan, fonksiyonel tahmin edicilerin kovaryansındaki bilgilerin çoğu, birkaç FPC tarafından yakalanır ve bunlar, yanıt değişkenini modellemek için çok önemli değildir. Temel fonksiyonlar ve fonksiyonel tahmin ediciler arasındaki etkileşimi açıklayan en önemli terimlerin tümü veya bir kısmı, daha yüksek dereceli ana bileşenlerden gelebilir (Delaigle ve Hall, 2012). Fonksiyonel kısmi en küçük kareler (Functional Partial Least Squares, FPLS) regresyonu, fonksiyonel nesnelere ortonormal gizli bileşenlere ve karşılık gelen puanlara yansır. FPC regresyonundan farklı olarak, FPLS regresyonu, SoFRM'yi sonlu boyutlu uzaya yansıtırken hem yanıt hem de öngörücü değişkenlerin bilgilerini kullanır, yani FPLS bileşenleri, yanıt ve fonksiyonel öngörücü arasındaki kovaryansı maksimize ederek elde edilir. FPC'lerle karşılaştırıldığında, FPLS daha az terimle daha bilgilendirici bileşenler üretmektedir (Delaigle ve Hall, 2012). Ayrıca Aguilera ve diğ. (2010), FPLS'nin FPC'den daha etkin parametre tahminleri ürettiğini göstermişlerdir. Bu nedenle, SoFRM'yi tahmin etmek için genellikle FPC yerine FPLS tercih edilmektedir (Saporta, 1981; Preda ve Saporta, 2005; Preda ve diğ., 2007; Reiss ve Ogden, 2007; Kramer ve diğ., 2008; Aguilera ve diğ., 2010; Delaigle ve Hall, 2012; Aguilera ve diğ., 2013, 2016; Yu ve diğ., 2016; Febrero-Bande ve diğ., 2017; Kalogridis ve Aelst, 2019).

FPLS yöntemleri, yinelemeli öz-birleştirme gerektiren SIMPLS (The Statistically Inspired Modification of The Partial Least Squares, Kısmi En Küçük Karelerin İstatistiksel Olarak İlham Edilen Modifikasyonu) ve NIPALS (Non-Linear Iterative Partial Least Squares, Doğrusal Olmayan Yinelemeli Kısmi En Küçük Kareler)

(çoğunlukla SIMPLS) olmak üzere iyi bilinen iki algoritmaya dayalı olarak oluşturulmaktadır. SIMPLS ile karşılaştırıldığında, NIPALS algoritması daha fazla hesaplama süresi gerektirmektedir (Björck ve Indahl, 2017; Beyaztaş ve Shang, 2020b) ve bu nedenle, seyrek ve/veya büyük ölçekli veri kümelerinde hesaplama açısından uygun olmayabilir. Öte yandan, SIMPLS, NIPALS'den daha hızlı bir algoritma olmasına rağmen, genellikle regresyon katsayılarını tahmin etmede ve model uydurmada sayısal olarak etkin sonuçlar vermemektedir (Andersson, 2009; Björck ve Indahl, 2017). Bu sonuçlar ışığında, SIMPLS ve NIPALS algoritmalarına dayalı olarak oluşturulan FPLS regresyon yöntemleri, genel temel genişletme fonksiyonları ve FPC regresyonundan daha iyi çıkarımlar sağlarken, bu yöntemlerin kullanımı üç durumda yetersiz kalabilmektedir. 1) Fonksiyonel tahmin edici ultra yoğun eğrilerden oluştuğunda (yani eğri başına zaman noktası sayısı çok büyük olduğunda) regresyon parametre fonksiyonu için geliştirilmiş bir tahmin ve SoFRM'nin geliştirilmiş bir tahmini elde etmek için, fonksiyonel tahmin edicileri sonlu boyutlu bir uzaya yansıtmak nispeten çok sayıda temel fonksiyona ihtiyaç duyulabilir. 2) Skaler yanıtta değişimi daha iyi açıklamak için modelde çok sayıda fonksiyonel tahmin ediciye ihtiyaç vardır. Bu durumda, modelin boyutu, yalnızca bir fonksiyonel öngörücü duruma kıyasla daha büyük hale gelir. 3) 1 ile 2'nin birleşimi. Bu gibi durumlarda, NIPALS tabanlı FPLS yöntemi, hesaplama açısından olanaksız olabilirken, SIMPLS tabanlı FPLS yöntemi, kararsız regresyon katsayısı tahminleri ve model uydurma sonuçları üretebilir.

Fonksiyonel veri analizine ülkemizde 2000'li yıllardan itibaren ilgi gösterilmeye başlanmıştır. Keser ve Ertaş (2007) çalışmalarında İstanbul Menkul Kıymetler Borsasına (İMKB) ait 01.10.2006 ile 31.12.2006 tarihleri arasındaki 30 hisse senedinin 61 ayık noktada gözlenen kapanış fiyatları üzerinde düzgünleştirilmiş fonksiyonel temel bileşenler analizi ile İMKB 30 endeksine BIST farklı bir bakış açısı sunmuşlardır. Özçomak ve Gündüz (2014), Borsa İstanbul'da işlem gören şirketlerin kapanış fiyatları ile işlem miktarları arasındaki ilişkinin fonksiyonel kanonik korelasyon ile analizi üzerine çalışmışlardır. Çalışmalarında BIST'de işlem gören şirketlere ait kapanış fiyatı ve işlem miktarı verileri ele alınarak kapanış fiyatı ile işlem

miktarı arasında kuvvetli bir ilişki olduğunu göstermişlerdir. Keser (2010) bir başka çalışmada Ege Bölgesi yağış verilerini kullanarak fonksiyonel veri analizi ile bir inceleme yapmıştır. Bu çalışmada, 22 farklı meteoroloji istasyonun 2000-2005 yıllarına ait aylık ortalama yağış verileri arasındaki değişim yapısını ortaya koymak için düzgünleştirilmiş fonksiyonel temel bileşenler analizini kullanmıştır. Şafak (2019), fonksiyonel veri analizi ile sağlık alanına özgü bir çalışma gerçekleştirmiştir. Uygulamada, meme kanseri nedeniyle radikal mastektomi ameliyatı olan ve lenf bezleri alınan 57 kadının ameliyattan sonraki fleksiyon açıklığı ölçümleri alınarak, iki farklı tedavi yöntemi fonksiyonel t-testi ile analiz edilmiştir. Erdoğan (2020), Türkiye’de yaşa özel doğurganlık hızının tahmini için 2001 - 2018 yıllarına ait veri seti üzerinde parametrik olmayan düzleştirme yöntemleri ile sağlam istatistik yöntemlerini kullanarak ve Lee-Carter yöntemini genelleştirerek yeni bir yöntem önermiştir. Denizli (2016), İspanya’nın Cadiz bölgesinde bulunan Guadalquivir nehir ağzında biriken katı madde oranının tahmin edilmesi için fonksiyonel doğrusal regresyon modellerini (FDRM) uygulayarak, bu modeller ile klasik istatistiksel modellerden elde edilen sonuçları kıyaslamıştır. Sonuçlara göre, FDRM’nin tahmin etmede iyi bir alternatif olduğu yorumu yapılmıştır.

Bu çalışmanın geri kalanı aşağıdaki gibi yapılandırılmıştır. İkinci bölümde, fonksiyonel veri yapısı, fonksiyonel veri için matematiksel kavramlar, genişletme teknikleri ve model tahmini hakkında teorik bilgiler ele alınmıştır.

Üçüncü bölümde, fonksiyonel veri analizi araçlarından fonksiyonel doğrusal regresyon modelleri tanıtılmıştır. Diğer bir ifadeyle, bağımlı ve bağımsız değişkenlerin fonksiyonel olma durumlarına göre üç farklı regresyon model yapısından bahsedilmiştir.

Dördüncü bölümde, bağımlı değişkeninin skaler açıklayıcı değişkeninin fonksiyonel olduğu regresyon modeli üzerinde regresyon katsayı fonksiyonunu verimli bir şekilde tahmin etmek için hibrit iki fonksiyonel kısmi en küçük kareler yöntemi detaylı olarak açıklanmıştır.

Beşinci bölümde, çeşitli Monte-Carlo simülasyon çalışmaları ve kemometrik veri analizleri gerçekleştirilmiş ve sonuçlar mevcut yöntemlerle karşılaştırılmıştır.

Son bölümde, elde edilen sonuçlar genel anlamda yorumlanmıştır. Ayrıca, burada sunulan metodolojinin nasıl genişletilebileceğine dair bazı fikirlerle önerilerde bulunulmuştur.



2. FONKSİYONEL VERİ ANALİZİ

Pek çok alandaki veriler, fonksiyonel olarak tanımlanan bir süreçle bize gelmektedir. FDA’da gözlemlenen veriler, klasik istatistikteki nokta olarak tanımlanan gözlemlerin aksine farklı noktalar temelinde bir fonksiyonu tanımlamaktadır. Diğer bir deyişle, her eğri tek bir fonksiyonel gözlemden oluşmaktadır ve bu gözlemler Hilbert uzayında tanımlıdırlar. Klasik istatistikte kullanılan yöntemlerin fonksiyonel veri analizinde karşılığı bulunmaktadır.

FDA’nın ilk adımında fonksiyonel bir veriyi oluşturmak için uygun bir temel fonksiyon genişletme teknikleri yardımıyla belirlenmesi gerekmektedir. Böylece, belirlenen temel fonksiyon, fonksiyonel gözlemin doğrusal bir kombinasyonu olarak ifade edilmektedir. Daha açık bir şekilde ifade edecek olursak; T kapalı ve sınırlı bir aralık olmak üzere $t, \{t_1, t_2, \dots, t_j\} \in T$ üzerinde gözlemlenen ve genelde ayrık zaman noktaları olarak tanımlanan bir argümanı gösterebiliriz. O zaman, X_i fonksiyonel gözlem şu şekilde ifade edilir:

$$X_i(t) = \sum_{k=1}^{\infty} c_{ik} \phi_k(t) \quad (2.1)$$

$$\hat{X}_i(t) = \sum_{k=1}^K c_{ik} \phi_k(t), \quad \forall t \in T \quad (2.2)$$

Burada, $\hat{X}_i(t)$, $X_i(t)$ ’in yaklaşımıdır ve $K \rightarrow \infty$ iken $X_i(t)$ ’ye yakınsar. $\phi_k(t)$, ($k = 1, \dots, K$) k -ıncı temel fonksiyon ve c_{ik} , karşılık gelen katsayısıdır. $Y = [Y_1, Y_2, \dots, Y_N]'$, N fonksiyonel verinin vektörü olsun. O zaman, fonksiyonel bir model aşağıdaki gibi yazılır:

$$Y_{ij} = X_i(t_j) + \varepsilon_{ij}, \quad 1 \leq j \leq J, 1 \leq i \leq N \quad (2.3)$$

$Y_{ij}, X_i(t_j)$ stokastik sürecin gürültülü bir gözlemi ve ε_{ij} , i -inci fonksiyonel veri ile ilişkili 0 ortalama ve σ_i^2 varyanslı rastgele bir hata terimidir.

Fonksiyonel değerleri sürekli zaman içinde gözlemlemek oldukça zor hatta imkansız bir durumdur. Temel açılımlar yardımıyla genişletme teknikleri fonksiyonel verilere iyi bir yaklaşım sağlamaktadır. Çünkü temel fonksiyonlar, verileri üreten süreçle aynı temel özelliklere sahip olmaktadır. Ayrıca, hesaplamalarda ve analizlerde ham verideki gürültüyü en aza indirmektedir. Temel fonksiyon tekniklerindeki en önemli görev en uygun temel fonksiyonun belirlenmesidir (Ramsay ve Silverman, 2005; Beyaztaş ve Shang, 2020a, 2020b).

2. 1 Matematiksel Bazı Gösterimler ve İsimlendirme

Klasik istatistikte, gözlemler skaler değerlerden oluşmaktadır ve doğrunun üzerinde tanımlıdır. Çok değişkenli istatistikteki gözlemler ve \mathbb{R}^p Öklid uzayının değerleri ise vektörlerden oluşmaktadır, burada p , gözlenen vektörlerin boyutudur. Fonksiyonlar her ne kadar t_j noktalarında gözlemleniyor olsa da, bunlar sonsuz boyutlu uzayda ele alınmaktadır. Bu bölümde, fonksiyonel verilerin nasıl bir uzayda yaşadığına ve fonksiyonel analiz ile fonksiyonel veri analizi arasındaki bağlantının daha rahat anlaşılabilmesi için bazı matematiksel gösterimlere ve teoremlere yer verilmiştir (Kokoszka ve Reimherr, 2017).

2.1.1 Hilbert Uzayları

Tanım (İç Çarpım Uzayı): H bir vektör uzayı olsun. H üzerinde bir iç çarpım, $\langle \cdot, \cdot \rangle: H \times H \rightarrow \mathbb{R}$ ve her $u, v, w \in H$ ve $\alpha, \beta \in \mathbb{R}$ için,

1. $\langle u, u \rangle \geq 0$ ve $\langle u, u \rangle = 0 \Leftrightarrow u = 0_H$ (Pozitiflik).
2. $\langle u, v \rangle = \overline{\langle v, u \rangle}$ (Eşlenik simetriklik).
3. $\langle \alpha u + \beta w, v \rangle = \alpha \langle u, v \rangle + \beta \langle w, v \rangle$ (Doğrusallık).

özelliklere sahip bir fonksiyondur. O zaman, bu özellikler altında, $(H, \langle \cdot, \cdot \rangle)$ çiftine iç çarpım uzayı olarak tanımlanır.

Tanım (Cauchy – Schwartz Eşitsizliği): Eğer $(H, \langle \cdot, \cdot \rangle)$ çifti iç çarpım uzayı ise o zaman, $\forall u, v \in H$ için aşağıdaki Cauchy - Schwartz eşitsizliği bu durumda geçerli olmaktadır:

$$|(u, v)|^2 \leq \langle u, u \rangle \langle v, v \rangle$$

Tanım (Norm): Eğer $(H, \langle \cdot, \cdot \rangle)$ iç çarpım uzayı ise, o zaman aşağıdaki gibi

$$\|v\| := \sqrt{\langle v, v \rangle} \quad \forall v \in H$$

ifade edilen $\|\cdot\|: H \rightarrow \mathbb{R}$ fonksiyonu bir norm olarak tanımlanır. Bu durumda, $(H, \langle \cdot, \cdot \rangle)$ çifti normlu uzay olarak tanımlanır. Ayrıca, tüm normlu uzaylar metrik uzayı olarak da ifade edilir.

Tanım (Tam): Eğer normlu H uzayında her $\{v_n\}$ Cauchy dizisi için $v \in H$, öyle ki $v_n \rightarrow v$, var ise H iç çarpım uzayı tamdır.

Verilen tanımlar bilgisinde, H tam bir iç çarpım uzayı ise H uzayına Hilbert uzayı denir.

2.1.2 Hilbert Uzaylarında Operatörler

Tanım (Hilbert Schmidt Operatörü): $\|\cdot\|$ normu üreten $\langle \cdot, \cdot \rangle$ iç çarpım ile ayrılabilir bir H Hilbert uzayını düşünelim ve H üzerinde sınırlı (sürekli) lineer operatörlerin uzayını \mathcal{L} ile gösterelim. O zaman, operatörün normu aşağıdaki gibi tanımlanır:

$$\|\Psi\|_{\mathcal{L}} = \sup\{\|\Psi(x)\|_{\mathcal{L}}: \|x\| \leq 1\}$$

Eğer $\{e_n\}$ ve $\{f_n\}$ iki ortonormal dizi ve sıfıra yakınsayan bir gerçekte dizi $\{\lambda_j\}$ var ise $\Psi: H \rightarrow H$ operatörü kompakttır ve

$$\Psi(x) = \sum_{i=1}^{\infty} \lambda_j \langle x, e_i \rangle f_i, \quad x \in H$$

$$= \lim_{m \rightarrow \infty} \sum_{i=1}^m \lambda_j \langle x, e_i \rangle f_i, \quad x \in H$$

denklemini Ψ 'nin tekil deęer ayrışımı (Singular Value Decomposition, SVD) olarak temsil edilir. $\sum_{i=1}^{\infty} \lambda_i^2 < \infty$ saęlayan Ψ kompakt operatörü ise Hilbert - Schmidt operatörü olarak tanımlanır. S , Hilbert - Schmidt operatörlerinin uzayı,

$$\langle \Psi_1, \Psi_2 \rangle_S = \sum_{i=1}^{\infty} \langle \Psi_1(e_i), \Psi_2(e_i) \rangle$$

skaler çarpımla ayrılabilir bir Hilbert uzayı olduęu söylenir. Burada, $\{e_i\}$ keyfi bir ortonormal temeldir. $\Psi \in \mathcal{L}$ operatörünün;

- Eęer $\langle \Psi(x), y \rangle = \langle x, \Psi(y) \rangle$, $x, y \in H$ ise simetrik
- Eęer $\langle \Psi(x), x \rangle \geq 0$, $x \in H$ ise pozitif yarı tanımlı

olduęu söylenir. Simetrik pozitif yarı tanımlı bir Ψ Hilbert - Schmidt operatörü,

$$\Psi(x) = \sum_{i=1}^{\infty} \lambda_i \langle x, e_i \rangle e_i, \quad x \in H$$

denklemini ile Ψ 'nin özfonksiyonları olan $\{e_i\}$ ortonormal dizisi ile ayrışımını saęlar. Yani, $\Psi(e_i) = \lambda_i e_i$ 'dir.

2.1.3 L^2 Uzayı

$L^2 = L^2(T)$ uzayı, $\int_T x^2(t) dt < \infty$ saęlayan $T = [a, b]$ üzerinde tanımlanan x ölçülebilir reel deęerli fonksiyonların kümesidir. Dięer bir deyişle,

$$\langle x, y \rangle = \int_T x(t)y(t) dt$$

iç çarpımı ile L^2 uzayı ayrılabilir bir Hilbert uzayıdır. Aşğıdaki gibi tanımlanan,

$$\Psi(x)(t) = \int_T \Psi(t, s)x(s) ds, \quad x \in L^2$$

integral operatörleri L^2 uzayında önemli operatörlerdir. Bu tür operatörler ancak ve ancak,

$$\iint_T \Psi^2(t, s) dt ds < \infty$$

ise Hilbert – Schmidt operatörüdür. Bu durumda,

$$\|\Psi\|_S^2 = \iint_T \Psi^2(t, s) dt ds$$

koşulunu sağlamaktadır. Eğer $\Psi(s, t) = \Psi(t, s)$ ve $\iint_T \Psi(t, s)x(t)x(s) dt ds \geq 0$ ise Ψ integral operatörü simetrik ve pozitif tanımlıdır. Bu nedenle Ψ operatörü;

$$\Psi(t, s) = \sum_{i=1}^{\infty} \lambda_i e_i(t) e_i(s), \quad L^2(T \times T)$$

biçiminde yazılabilir ve Mercer Teoremi olarak bilinmektedir.

2.1.4 Stokastik Süreçler

$(\Omega, \mathcal{F}, \mathbb{P})$ bir olasılık uzayı olduğunu varsayalım. Burada, Ω rastgele deneylerin yapıldığı örneklem uzayı; \mathcal{F} , Ω üzerinde alt kümelerinin bir σ -cebiri ve \mathbb{P} , \mathcal{F} üzerinde bir olasılık ölçüsüdür. $X = \{X_t, t \in T\}$, $(\Omega, \mathcal{F}, \mathbb{P})$ üzerinde tanımlı rastgele bir değişkeninin beklenen değeri ve varyansı sırasıyla aşağıdaki gibidir:

$$E(X) = \int_{\Omega} X(w) d\mathbb{P}(w)$$

$$Var(X) = E[(X - E(X))^2]$$

$L^2(\Omega, \mathcal{F}, \mathbb{P})$, Ω üzerinde kare integrallenebilir rastgele değişkenlerin Hilbert uzayını gösterir ve

$$L^2(\Omega, \mathcal{F}, \mathbb{P}) = \{X: \Omega \rightarrow \mathbb{R}: \int_{\Omega} |X(w)|^2 d\mathbb{P}(w) < \infty\}$$

biçiminde ifade edilir. $\{Y_t\}_{t \in T}$ keyfi bir stokastik süreç olsun. O zaman,

$$Y_t = E(X_t) + X_t$$

$E(X_t) = 0, \forall t \in T$ stokastik süreci merkezileştirilir. Burada, $X_t = Y_t - E(X_t)$. Genellik kaybı olmadan, FDA'da merkezileştirilmiş stokastik süreçler üzerinde durulmaktadır.

2.1.5 Karhunen – Loeve Teoremi

$X, 0$ ortalamalı kare sürekli stokastik bir süreç olsun. O zaman,

$$X_t = \sum_i^{\infty} x_i e_i(t), \quad \forall t \in T$$

burada, x_i katsayılar $x_i(w) = \int_T X_t(w) e_i(t) dt$ ile hesaplanmaktadır. Ayrıca,

1. $E(x_i) = 0, \quad \forall i \in \mathbb{N}$.
2. $E(x_i x_j) = \delta_{ij} \lambda_j, \quad \forall i, j \in \mathbb{N}$.
3. $Var(x_j) = \lambda_i, \quad \forall i \in \mathbb{N}$.

koşullar sağlanmalıdır. Burada, $\delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$ ve (λ_j, e_i) dizisi Mercer Teoreminde tanımlanır. Yukarıdaki denklem $L^2(\Omega, \mathcal{F}, \mathbb{P})$ uzayına düzgün bir şekilde yakınsamaktadır. Bu teorem ile sınırlı bir aralıktaki L^2 uzayının elemanları olan özfonksiyonlar herhangi bir stokastik sürecin, sonsuz doğrusal bir kombinasyonu olarak temsil edilebilmektedir. Böylece, özfonksiyonların sonsuz doğrusal kombinasyonu, temel fonksiyonların sonlu toplamı olarak yazılabilmektedir. Yani, (2.2) denklemindeki $\phi_k(t)$ temel fonksiyonlar ile c_k katsayıları Karhunen – Loeve teoremindeki özfonksiyonlar ve özdeğerler ile yer değiştirilerek elde edilmektedir (Bosq, 2000; Ramsay ve Silverman, 2005; Brezis, 2010; Essomba, 2015; Kokoszka ve Reimherr, 2017).

2.2 Temel Açılımlar İçin Genişletme Teknikleri

FDA’da ilk adımda, gözlemlenen bir verinin fonksiyonel gözlem olarak temsil edilebilmesi için örneklem eğrilerini fonksiyonel forma dönüştürmek gerekmektedir. Örneklem eğrilerinin, kare integrallenebilir fonksiyonların $L^2(T)$ Hilbert uzayına ait olan $X = \{X(t): t \in T\}$ stokastik bir süreç olduğunu varsayalım. O zaman, (2.1)’deki denklem,

$$X_i(t) \approx c_i' \phi(t), \quad \forall t \in T, i = 1, \dots, N \quad (2.4)$$

biçiminde yazılmaktadır. Burada, $c_i = [c_{i1}, \dots, c_{iK}]'$ katsayıları ağırlığı oluştururken her bir $\phi(t) = [\phi_1(t), \dots, \phi_K(t)]'$ temel fonksiyonların vektörü olarak ifade edilir. Denklemin matris gösterim şekli aşağıdaki gibidir:

$$X_i(t_i) \approx \Phi(t_i) c_i, \quad \forall t \in T \quad (2.5)$$

Burada, $\Phi(t_i) = \begin{pmatrix} \phi_1(t_{i1}) & \dots & \phi_K(t_{i1}) \\ \vdots & \ddots & \vdots \\ \phi_1(t_{ij}) & \dots & \phi_K(t_{ij}) \end{pmatrix}$ her t_i noktasında değerlendirilen temel

fonksiyonların bir $J \times K$ boyutlu matrisidir (Ramsay ve Silverman, 2005; Essomba, 2015).

Genişletmenin boyutu, K olarak tanımlanan temel fonksiyonların sayısına bağlıdır ve sonsuz boyutlu uzaya ait olduğu için sonlu boyutlu uzaya yansıtılarak çalışılması sağlanmaktadır. Genişletmede temel fonksiyon sayısının seçimi FDA için oldukça dikkat ve önem gerektiren bir durumdur. Örneğin, K sayısının küçük olması hesaplamada kolaylık sağlamaktadır (Denizli, 2016). Fonksiyonel verilere uygulanabilen pek çok genişletme tekniği bulunmaktadır. Bu tekniklerden en sık kullanılan iki temel açılım aşağıdaki gibidir:

2.2.1 Fourier Temel Açılımı

Bir T aralığında tanımlanan periyodik fonksiyonlar için en uygun temel, ϕ_K 'ların Fourier temeli aldığı biçim olup aşağıdaki gibidir:

$$\phi_0(t) = \frac{1}{\sqrt{T}}, \quad \phi_{2r-1}(t) = \frac{\sin(rwt)}{\sqrt{T/2}}, \quad \phi_{2r}(t) = \frac{\cos(rwt)}{\sqrt{T/2}}, \quad r = 1, \dots, \frac{K-1}{2}$$

Burada, K temel fonksiyon sayısıdır ve Fourier temeline göre hesaplama yapmak için K tek sayı olmalıdır. Frekans w , $T = 2\pi/w$ aralığının periyodunu ve uzunluğunu belirler. $\phi(t)$ fonksiyon vektörü, t_j , $j = 1, \dots, J$ için ayrık zaman noktalarında değerlendirilen $\phi(t) = [\phi_0(t), \phi_1(t), \phi_2(t), \dots, \phi_{2r}(t)]'$ formuna sahiptir (Beyaztaş ve Shang, 2020a).

Fourier temel açılımı, güçlü yerel özelliklerin olmadığı ve eğriliğin her yerde aynı düzende devam etme eğiliminde olduğu durumlarda son derece kullanışlıdır. Bu temel açılım periyodik olduğu için uzun zaman serilerinde hesaplama verimliliği açısından kolaylık sağlamaktadır. Bu nedenle eşit aralıklı ve periyodik yapılarda bu temel genişletme tekniği tercih edilmelidir (Ramsay ve Silverman, 2005).

2.2.2 B-Spline Temel Açılımı

De Boor (2001) tarafından geliştirilen B-Spline temel fonksiyonları, periyodik olmayan veriler için tercih edilen aynı zamanda daha az sayıda fonksiyon sayısı gerektiren durumlarda kullanılabilmesi ve hızlı hesap edilebilmesinden dolayı oldukça avantajlı bir genişletme tekniğidir (Ramsay ve Silverman, 2005; Hovarth ve Kokoszka, 2012). Spline temeller, kesme noktalarında düzgün bir şekilde bağlanan ve her aralıkta m mertebeden bir parçalı fonksiyondur. Burada, belirtilen sayıda kesme noktası belirtilen sıradaki spline fonksiyonlarının doğrusal kombinasyonları şeklinde ifade edilir. Başka bir deyişle, $\tau = \{\tau_1, \dots, \tau_{m+1}\}$, temelin tanımlandığı T aralığı, m alt aralığın (düğüm) bir kesme noktaları dizisini, $\tau_0 < \tau_1 < \dots < \tau_m$ gösterebilir. Bu durumda, k 'inci B-Spline temel fonksiyonu,

$$\phi_k(t) = B_{k,m}(t)$$

olarak tanımlanır. B-spline tabanlı fonksiyonlar oluşturmak için, düğüm dizisinin $[\tau_0, \tau_m]$ sınırlarının dışında fazladan düğümlere ihtiyaç vardır. $\dots < \tau_{-1} < \tau_0 < \tau_1 < \dots < \tau_m < \tau_{m+1} < \dots$ eşitsizliği arttırılmış düğüm dizisi olarak tanımlanır ise sabit B-Spline temeller aşağıdaki gibi tanımlanır:

$$B_{k,1}(t) = \begin{cases} 1, & \tau_k \leq t \leq \tau_{k+1} \\ 0, & \text{diğer durumlarda} \end{cases}$$

Bu fonksiyonlara Haar temel fonksiyonlar olarak da isimlendirilmektedir. Sabit B-Spline açılımlar yardımıyla, yüksek mertebeden B-Spline temeli aşağıdaki özyineleme yoluyla elde edilmektedir:

$$B_{k,m}(t) = \frac{t - T_k}{T_{k+m-1} - T_k} B_{k,m-1}(t) + \frac{T_{k+m} - t}{T_{k+m} - T_{k+1}} B_{k+1,m-1}(t), \quad t \in T, m \geq 2$$

B-Spline temel fonksiyonların sayısı, mertebesine ve düğüm sayısına bağlıdır. Bu tekniğin temel avantajı esnek ve hızlı hesaplanmasıdır (Beyaztas ve Shang, 2020a).

2.2.3 Diğer Faydalı Temel Açılımlar

Literatürde önemli bir dizi başka temel açılımlar da yer almaktadır. Son yıllardaki farklı durumlar beraberinde veri analizinde farklı temel açılımların kullanımını gerektirmektedir. Örneğin, matematiksel olarak oldukça büyük ve karmaşık bir yapının meydana gelmesinden dolayı Fourier serisinin frekansa özgü yaklaşık gücünü ve spline temel açılımların zaman veya uzamsal olarak yerleştirilmiş özellikleriyle birleştiren dalgacıklar oluşturulmaktadır. Ayrıca, bazı basit temel açılımlara da ilginin son derece önem kazandığı görülmektedir. Bu diğer faydalı temel açılımlar, dalga tabanlı temel açılım, Gaussian radyal temel açılım, sabit temel açılımı, üstel temel açılım, güç temel açılımı, polinom temel açılımları, çokgen (poligonal) temel açılımı, adım-fonksiyon temel açılımı, tek terimli temel açılım, ampirik ve tasarımcı temel açılımlarından oluşmaktadır (Ramsay ve Silverman, 2005; Ramsay ve diğ., 2009).

3. FONKSİYONEL DOĞRUSAL REGRESYON

Bu bölümde, Fonksiyonel Doğrusal Regresyon (Functionoanal Linear Regression, FLR) ile ilgili modellerden ve ilgili parametre tahmin yöntemlerinden bahsedilecektir. Bilindiği üzere klasik istatistikte değişkenler arasında doğrusal bir ilişki kurmak ve bu ilişki ile önkestimler elde etmek doğrusal regresyon ile gerçekleşmektedir. FLR ile birçok analizde olduğu gibi veriyi fonksiyonel forma dönüştürerek analiz yapılmaktadır. Klasik doğrusal regresyondaki amaç gibi FLR, değişimin ne kadarının diğer değişkenler tarafından açıklanabileceğini göz önünde bulundurarak, bir skaler ya da fonksiyonel yanıtın değişkenliğini keşfetmeyi amaçlar.

3.1 Fonksiyonel Doğrusal Regresyon Modelleri

Standart doğrusal regresyon modelinde bağımlı ve bağımsız değişken(ler) skaler değerlere sahip iken Fonksiyonel Doğrusal Regresyon Modellerinde (Functional Linear Regression Models, FLRM) bağımlı ve/veya bağımsız değişkenlerinin fonksiyonel olma durumlarına göre üç farklı regresyon modeli tanımlanmaktadır ve bu modeller sırasıyla aşağıdaki gibidir:

3.1.1 Skaler - Fonksiyon Doğrusal Regresyon Modelleri

Skaler – Fonksiyon Doğrusal Regresyon modeli (Scalar on Function Regression Model, SoFRM) bağımlı değişkenin skaler, bağımsız değişkenin fonksiyonel değerlerden oluştuğu bir modeldir. Daha açık bir şekilde ifade edersek, $\{Y_i, i = 1, \dots, N\}$, skaler yanıtların bir N -vektörü ve T_m üzerinde tanımlı $\{X_{im}(t), m = 1, \dots, M\}$, M -fonksiyonel öngörücüler olsun. $X_{im}(t)$, fonksiyonları, genişletme teknikleri kullanılarak elde edilmektedir. O zaman, bir SoFRM aşağıdaki gibidir:

$$Y_i = \beta_0 + \sum_{m=1}^M \int X_{im}(t) \beta_m(t) dt + \varepsilon_i, \quad \forall i, m \quad (3.1)$$

Burada, β_0 , orijini ayarlayan olağan kesme terimidir, $\beta_m(t)$ katsayı fonksiyonlarıdır ve ε_i , 0 ortalama ve σ_i^2 varyans ile bağımsız ve normal olarak dağılılan hata terimleridir. $X_{im}(t)$ fonksiyonel tahmin edicileri ise şu şekilde ifade edilmektedir:

$$X_{im}(t) = \sum_{k=1}^{K_m^\alpha} c_{imk} \phi_{mk}(t) = \mathbf{c}_{im}^T \boldsymbol{\phi}_m(t), \quad \forall t \in T_m \quad (3.2)$$

Ayrıca $\beta_m(t)$ katsayı fonksiyonları,

$$\beta_m(t) = \sum_{l=1}^{K_m^\beta} b_{ml} \psi_{ml}(t) = \mathbf{b}_m^T \boldsymbol{\psi}_m(t), \quad \forall t \in T_m \quad (3.3)$$

biçiminde verilebilir. Burada, K_m^β temel fonksiyonları $\{\psi_{m1}(t), \dots, \psi_{mK_m^\beta}(t)\}$ doğrusal kombinasyonları ile temsil edilmektedir. (3.1) denkleminde (3.2) ve (3.3) denklemleri yerine yazılarak aşağıdaki sonuç elde edilmektedir:

$$\begin{aligned} Y_i &= \beta_0 + \sum_{m=1}^M \int \mathbf{c}_{im}^T \boldsymbol{\phi}_m(t) \boldsymbol{\psi}_m^T(t) \mathbf{b}_m dt + \varepsilon_i \\ &= \beta_0 + \sum_{m=1}^M \mathbf{c}_{im}^T \mathbf{J}_{\phi\psi}^{(m)} \mathbf{b}_m + \varepsilon_i \end{aligned} \quad (3.4)$$

Burada, $\mathbf{J}_{\phi\psi}^{(m)} = \int \boldsymbol{\phi}_m(t)\boldsymbol{\psi}_m^T(t)dt$, $K_m^x \times K_m^\beta$ çarpım matrisidir. Denklem (3.4) matris formunda aşağıdaki şekilde yeniden yazılabilir:

$$\mathbf{Y} = \mathbf{Z}\mathbf{B} + \boldsymbol{\epsilon} \quad (3.5)$$

Burada,

$$\mathbf{Z} = \begin{pmatrix} \mathbf{z}_1^T \\ \vdots \\ \mathbf{z}_N^T \end{pmatrix} = \begin{pmatrix} 1 & \mathbf{c}_{11}^T \mathbf{J}_{\phi\psi}^{(1)} & \dots & \mathbf{c}_{1M}^T \mathbf{J}_{\phi\psi}^{(M)} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \mathbf{c}_{N1}^T \mathbf{J}_{\phi\psi}^{(1)} & \dots & \mathbf{c}_{NM}^T \mathbf{J}_{\phi\psi}^{(M)} \end{pmatrix}$$

$$\mathbf{B} = \begin{pmatrix} \beta_0 \\ \mathbf{b}_1 \\ \vdots \\ \mathbf{b}_M \end{pmatrix}$$

\mathbf{Y} , $N \times 1$ vektörlü skaler yanıtlar vektörü; \mathbf{Z} , fonksiyonel ortak değişkenlerin $N \times (\sum_{m=1}^M K_m^x + 1)$ boyutlu matrisi ve \mathbf{B} , fonksiyonel katsayıların $(\sum_{m=1}^M K_m^\beta + 1) \times 1$ boyutlu vektörü ve $\boldsymbol{\epsilon}$, $N \times 1$ boyutlu hata terimleri vektörüdür.

3.1.2 Fonksiyon – Fonksiyon Doğrusal Regresyon Modelleri

Bir diğer FLR modellerinden biri de hem yanıtın hem de açıklayıcı değişkenlerin fonksiyonel olduğu doğrusal regresyon modelidir (Function on Function Regression Model, FoFRM). Diğer bir deyişle, $i = 1, \dots, N$ için $Y_i(t)$ fonksiyonel yanıtlar ve $m = 1, \dots, M$ için $X_{im}(s)$ fonksiyonel tahmin ediciler olsun. Burada, $s \in S$ ve $t \in T$ reel doğru üzerinde kapalı ve sınırlı aralıklar olarak tanımlanır. Bu durumda, fonksiyonel yanıtlar ve tahmin ediciler arasındaki fonksiyonel ilişki aşağıda verildiği gibidir:

$$Y_i(t) = \beta_0(t) + \sum_{m=1}^M \int X_{im}(s) \beta_m(s, t) ds + \varepsilon_i(t), \forall s \in S, \forall t \in T \quad (3.6)$$

$\beta_0(t)$ sabit fonksiyonu, $\beta_m(s, t)$ yanıtı m 'inci tahmin ediciye bağlayan iki değişkenli katsayı fonksiyonu ve $\varepsilon_i(t)$ hata fonksiyonlarıdır. Fonksiyonel tahmin ediciler $X_{im}(s)$ ve fonksiyonel yanıtlar $Y_i(t)$ sırasıyla şu şekilde ifade edilmektedir:

$$X_{im}(s) = \sum_{j=1}^{K_m^x} \tilde{c}_{imj} \phi_{mj}(s) = \tilde{\mathbf{c}}_{im}^T \boldsymbol{\phi}_m(s), \quad \forall s \in S \quad (3.7)$$

$$Y_i(t) = \sum_{k=1}^{K_y} d_{ik} \psi_k(t) = \tilde{\mathbf{d}}_i^T \boldsymbol{\psi}(t), \quad \forall t \in T \quad (3.8)$$

Benzer şekilde, iki değişkenli katsayı fonksiyonu ise şu şekilde tanımlanır:

$$\beta_m(s, t) = \sum_{j,k} b_{mjk} \phi_{mj}(s) \psi_k(t) = \boldsymbol{\phi}_m^T(s) \mathbf{B}_m \boldsymbol{\psi}(t) \quad (3.9)$$

Burada, \mathbf{B}_m , $K_m^x \times K_y$ boyutlu katsayı matrisidir. (3.6)'daki FoFRM'yi sırasıyla aşağıdaki şekilde merkezileştirerek,

$$\begin{aligned} X_{im}^*(s) &= X_{im}(s) - \bar{X}_{im}(s) \\ &= \tilde{\mathbf{c}}_{im}^T \boldsymbol{\phi}(s) - \bar{\mathbf{c}}_{im}^T \boldsymbol{\phi}(s) \end{aligned}$$

$$= \mathbf{c}_{im}^T \boldsymbol{\phi}(s) \quad (3.10)$$

$$Y_i^*(t) = Y_i(t) - \bar{Y}_i(t)$$

$$= \tilde{\mathbf{d}}_i^T \boldsymbol{\psi}(t) - \bar{\mathbf{d}}_i^T \boldsymbol{\psi}(t)$$

$$= \mathbf{d}_i^T \boldsymbol{\psi}(t) \quad (3.11)$$

(3.6) denklemini aşağıdaki gibi elde edilir:

$$Y_i^*(t) = \sum_{m=1}^M \int X_{im}^*(s) \beta_m(s, t) ds + \varepsilon_i^*(t) \quad (3.12)$$

(3.9), (3.10) ve (3.11) denklemlerinden yararlanarak, (3.12) denklemini aşağıdaki biçimde elde edilir:

$$\begin{aligned} \mathbf{d}_i^T \boldsymbol{\psi}(t) &= \sum_{m=1}^M \int \mathbf{c}_{im}^T \boldsymbol{\phi}(s) \boldsymbol{\phi}_m^T(s) \mathbf{B}_m \boldsymbol{\psi}(t) + \varepsilon_i^*(t) \\ &= \sum_{m=1}^M \mathbf{c}_{im}^T \mathbf{J}_{\phi m} \mathbf{B}_m \boldsymbol{\psi}(t) + \varepsilon_i^*(t) \\ &= \mathbf{z}_i^T \mathbf{B} \boldsymbol{\psi}(t) + \varepsilon_i^*(t) \end{aligned} \quad (3.13)$$

Burada, $\mathbf{z}_i = (\mathbf{c}_{i1}^T \mathbf{J}_{\phi 1}, \dots, \mathbf{c}_{iM}^T \mathbf{J}_{\phi M})^T$ $[\sum_{m=1}^M K_m^x]$ boyutlu bir vektördür, $\mathbf{J}_{\phi m} = \int \boldsymbol{\phi}(s) \boldsymbol{\phi}_m^T(s) ds$ $K_m^x \times K_m^x$ boyutlu matrisi ve $\mathbf{B} = (\mathbf{B}_1, \dots, \mathbf{B}_M)^T$ bir $(\sum_{m=1}^M K_m^x \times K_y)$ boyutlu matristir. Yukarıdaki tüm bilgiler göz önüne alınarak tüm gözlemler için FoFRM matris biçiminde aşağıdaki gibi yazılabilir:

$$\mathbf{D}\boldsymbol{\psi}(t) = \mathbf{Z}\mathbf{B}\boldsymbol{\psi}(t) + \boldsymbol{\varepsilon}(t) \quad (3.14)$$

Burada, \mathbf{D} bir $N \times K_y$ boyutlu matris ve \mathbf{Z} bir $N \times (\sum_{m=1}^M K_m^x)$ boyutlu bir matristir (Ramsay ve Silverman, 2005; Beyaztas ve Shang, 2020b).

3.1.3 Fonksiyon – Skaler Doğrusal Regresyon Modelleri

Bir diğer FLRM olarak tanımlanan Fonksiyon – Skaler Regresyon Modeli (Function on Scalar Regression Model, FoSRM) bağımsız değişkenleri skaler, bağımlı değişkeni fonksiyonel değerlerden oluşan bir modeldir. Bir FoSRM modelinde $Y_i(t)$, T üzerinde tanımlı N boyutlu fonksiyonel bir yanıt ve $X_{im}(t)$, M boyutlu skaler tahmin ediciler olsun. Böylece, bu model aşağıdaki gibi ifade edilir:

$$Y_i(t) = \beta_0(t) + \sum_{m=1}^M X_{im}\beta_m(t) + \varepsilon_i(t), \quad t \in T \quad (3.15)$$

Burada $\beta_m(t)$, $M + 1$ boyutlu katsayılar vektörüdür ve $\varepsilon_i(t)$, X 'ten bağımsız bir sıfır ortalamalı rastgele bir hata fonksiyonudur. $\beta_m(t)$ fonksiyonu aşağıdaki gibi tanımlanmaktadır:

$$\beta_m(t) = \sum_{k=1}^K B_{mk}(t)\gamma_{mk} = \mathbf{B}_m^T \boldsymbol{\gamma}_m \quad (3.16)$$

Burada, $\boldsymbol{\gamma}_m = (\gamma_{m1}, \dots, \gamma_{mM})^T$ biçimindedir. O zaman, (3.16) denklemleri kullanılarak (3.15) denklemleri aşağıdaki gibi de verilebilir (Chen ve diğ.,2016; Cai ve diğ., 2022):

$$Y_i(t) = \beta_0(t) + \sum_{m=1}^M X_{im} \sum_{k=1}^K B_{mk}(t)\gamma_{mk} + \varepsilon_i(t)$$

$$= \beta_0(t) + \sum_{m=1}^M X_{im} \mathbf{B}_m^T \boldsymbol{\gamma}_m + \varepsilon_i(t) \quad (3.17)$$

(3.17) denklemi matris biçiminde aşağıda verildiği gibi yazılabilir:

$$\mathbf{Y} = \mathbf{Z}\boldsymbol{\gamma} + \boldsymbol{\epsilon} \quad (3.18)$$

3.2 Tahmin Yöntemleri

Bu bölümde, hem bağımlı hem de bağımsız değişkenlerin fonksiyonel olduğu (3.14) denkleminde verilen regresyon modeli için parametre tahmin yöntemleri üzerinde durulmuştur. Bu tahmin yöntemleri sırasıyla aşağıdaki gibidir:

3.2.1 En Küçük Kareler Yöntemi

\mathbf{B} katsayı matrisini tahmin etmek için Ramsay ve Silverman (2005) tarafından LS yöntemi hata kareler toplamını en aza indirmekten oluşan standart bir yaklaşımdır.

$$\begin{aligned} & \sum_{i=1}^N \int \left[Y_i^*(t) - \sum_{m=1}^M \int X_{im}^*(s) \beta_m(s, t) ds \right]^2 dt \\ &= \int iz \{ (\mathbf{D}\boldsymbol{\psi}(t) - \mathbf{Z}\mathbf{B}\boldsymbol{\psi}(t)) (\mathbf{D}\boldsymbol{\psi}(t) - \mathbf{Z}\mathbf{B}\boldsymbol{\psi}(t))^T \} dt \\ &= \int iz \{ (\mathbf{D} - \mathbf{Z}\mathbf{B}) \boldsymbol{\psi}(t) \boldsymbol{\psi}^T(t) (\mathbf{D} - \mathbf{Z}\mathbf{B})^T \} dt \\ &= iz \{ (\mathbf{D} - \mathbf{Z}\mathbf{B}) \mathbf{J}_\psi (\mathbf{D} - \mathbf{Z}\mathbf{B})^T \} \end{aligned}$$

$$\begin{aligned}
&= iz\{DJ_\psi D^T - DJ_\psi \mathbf{B}^T \mathbf{Z}^T - \mathbf{Z} \mathbf{B} J_\psi D^T + \mathbf{Z} \mathbf{B} J_\psi \mathbf{B}^T \mathbf{Z}^T\} \\
&= iz(DJ_\psi D^T) - 2iz(\mathbf{B} J_\psi D^T \mathbf{Z}) + iz(\mathbf{Z}^T \mathbf{Z} \mathbf{B} J_\psi \mathbf{B}^T) \quad (3.19)
\end{aligned}$$

Daha sonra, (3.19)'da verilen denklemin \mathbf{B} 'ye göre türevi alınıp sonucun sıfıra eşitlenmesiyle \mathbf{B} 'nin LS tahmincisi şu şekilde elde edilir:

$$vec(\widehat{\mathbf{B}}) = (\mathbf{J}_\psi \otimes \mathbf{Z}^T \mathbf{Z})^{-1} vec(\mathbf{Z}^T \mathbf{D} \mathbf{J}_\psi) \quad (3.20)$$

$vec(\widehat{\mathbf{B}})$, \mathbf{B} 'nin $(\sum_{m=1}^M K_m^x) \times K_y$ boyutlu sütun vektörünü ve \otimes Kronecker çarpımını ifade eder.

3.2.2 En Çok Olabilirlik Yöntemi

Bir diğer fonksiyonel regresyon modelinde parametre tahmini için kullanılan yöntem ML yöntemidir. (3.13) denklemindeki modelde $\varepsilon_i^*(t)$ hata fonksiyonlarının temel fonksiyonların doğrusal kombinasyonu ile temsil edilmesi için,

$$\varepsilon_i^*(t) = \sum_{k=1}^{K_y} e_{ik} \psi_k(t) = \mathbf{e}_i^T \boldsymbol{\psi}(t) \quad (3.21)$$

denklemini tanımlanırsa, (3.13) aşağıdaki gibi temsil edilebilir:

$$\mathbf{d}_i^T \boldsymbol{\psi}(t) = \mathbf{z}_i^T \mathbf{B} \boldsymbol{\psi}(t) + \mathbf{e}_i^T \boldsymbol{\psi}(t) \quad (3.22)$$

Burada, $\mathbf{e}_i = (e_{i1}, \dots, e_{iK_y})^T$ bağımsız ve 0 ortalama vektör, Σ varyans - kovaryans matrisi ile aynı normal dağılıma sahip K_y boyutlu bir vektördür.

(3.22) denkleminin her iki tarafı $\boldsymbol{\psi}^T(t)$ çarpılıp ardından T uzayı üzerinde integrali alındığında bu denklem yeniden şu şekilde yazılabilir:

$$\int \mathbf{d}_i^T \boldsymbol{\psi}(t) \boldsymbol{\psi}^T(t) = \int \mathbf{z}_i^T \mathbf{B} \boldsymbol{\psi}(t) \boldsymbol{\psi}^T(t) + \int \mathbf{e}_i^T \boldsymbol{\psi}(t) \boldsymbol{\psi}^T(t)$$

$$\mathbf{d}_i^T \mathbf{J}_\psi = \mathbf{z}_i^T \mathbf{B} \mathbf{J}_\psi + \mathbf{e}_i^T \mathbf{J}_\psi \quad (3.23)$$

Eğer \mathbf{J}_ψ tekil olmayan bir matris ise,

$$\mathbf{d}_i^T = \mathbf{z}_i^T \mathbf{B} + \mathbf{e}_i^T \quad (3.24)$$

denklemini elde edilir. Elde edilen (3.24) denklemini için olasılık yoğunluk fonksiyonu ise şu şekilde yazılabilir:

$$f(\mathbf{Y}_i | \boldsymbol{\theta}) = \frac{1}{(2\pi)^{K_y/2} |\boldsymbol{\Sigma}|^{1/2}} \exp \left\{ -\frac{1}{2} (\mathbf{d}_i - \mathbf{B}^T \mathbf{z}_i)^T \boldsymbol{\Sigma}^{-1} (\mathbf{d}_i - \mathbf{B}^T \mathbf{z}_i) \right\} \quad (3.25)$$

Burada, $\boldsymbol{\theta} = \{\mathbf{B}, \boldsymbol{\Sigma}\}$ 'dir. Böylece, (3.25) denkleminin logaritması alınıp düzenlenmiş hali aşağıdaki gibi elde edilir:

$$l(\mathbf{Y} | \boldsymbol{\theta}) = -\frac{N}{2} \log |\boldsymbol{\Sigma}| - \frac{1}{2} \text{tr} \{ \boldsymbol{\Sigma}^{-1} (\mathbf{D} - \mathbf{ZB})^T (\mathbf{D} - \mathbf{ZB}) \} \quad (3.26)$$

Yukarıdaki denklemin \mathbf{B} ve $\boldsymbol{\Sigma}$ parametrelerine göre türevi alınıp 0 eşitlendiğinde, parametrelerin en çok olabilirlik tahmin edicileri sırasıyla şu şekilde elde edilmektedir:

$$\hat{\mathbf{B}} = (\mathbf{Z}^T \mathbf{Z})^{-1} \mathbf{Z}^T \mathbf{D}, \quad \hat{\boldsymbol{\Sigma}} = \frac{1}{N} (\mathbf{D} - \mathbf{Z}\hat{\mathbf{B}})^T (\mathbf{D} - \mathbf{Z}\hat{\mathbf{B}}) \quad (3.26)$$

3.2.3 Cezalandırılmış En Çok Olabilirlik Yöntemi

LS ve ML yöntemleri bazen kararsız tahmin ediciler üretmektedir. Bu nedenle, ortaya atılan bir diğer yöntem maksimum cezalı olabilirlik yöntemidir. (3.25)'de verilen istatistiksel modelden, cezalandırılmış log – olabilirlik fonksiyonu şu şekilde verilmektedir:

$$l_{\Lambda}(\boldsymbol{\theta}) = l(\mathbf{Y}|\boldsymbol{\theta}) - \frac{N}{2} \text{tr}\{\mathbf{B}^T (\boldsymbol{\Lambda}_M \odot \boldsymbol{\Omega}) \mathbf{B}\} \quad (3.27)$$

Burada $\boldsymbol{\Lambda}_M$, $\lambda_1, \dots, \lambda_M$ düzgünleştirme parametrelerinin $(\sum_{m=1}^M K_m^x) \times (\sum_{m=1}^M K_m^x)$ boyutlu bir matrisi ve $\boldsymbol{\Lambda}_M = \boldsymbol{\lambda}_M \boldsymbol{\lambda}_M^T$, $\boldsymbol{\lambda}_M = (\sqrt{\lambda_1} \mathbf{1}_{K_1^x}, \dots, \sqrt{\lambda_M} \mathbf{1}_{K_M^x})^T$ 'dir. $\boldsymbol{\Omega}$, pozitif yarı tanımlı $(\sum_{m=1}^M K_m^x) \times (\sum_{m=1}^M K_m^x)$ boyutlu bir matris ve \odot , Hadamard çarpımıdır. (3.28) fonksiyonunun \mathbf{B} ve $\boldsymbol{\Sigma}$ parametrelerine göre türevi alınıp 0'a eşitlenmesiyle maksimum cezalı olabilirlik tahmin edicileri sırasıyla aşağıdaki gibi verilir:

$$\text{vec}(\hat{\mathbf{B}}) = \left(\hat{\boldsymbol{\Sigma}}^{-1} \otimes \mathbf{Z}^T \mathbf{Z} + N \mathbf{I}_{K_y} \otimes (\boldsymbol{\Lambda}_M \odot \boldsymbol{\Omega}) \right)^{-1} (\hat{\boldsymbol{\Sigma}}^{-1} \otimes \mathbf{Z}^T) \text{vec}(\mathbf{D}) \quad (3.28)$$

$$\hat{\boldsymbol{\Sigma}} = \frac{1}{N} (\mathbf{D} - \mathbf{Z} \hat{\mathbf{B}})^T (\mathbf{D} - \mathbf{Z} \hat{\mathbf{B}})$$

Ayrıca, \mathbf{D} 'nin cezalandırılmış en çok olabilirlik tahmincisi ise aşağıda tanımlandığı gibidir:

$$\begin{aligned} \text{vec}(\hat{\mathbf{D}}) &= \text{vec}(\mathbf{Z} \hat{\mathbf{B}}) \\ &= \mathbf{S}_{\Lambda} \text{vec}(\mathbf{D}) \end{aligned} \quad (3.30)$$

Burada $\mathbf{S}_{\Lambda} = (\mathbf{I}_{K_y} \otimes \mathbf{Z}) \left(\hat{\boldsymbol{\Sigma}}^{-1} \otimes \mathbf{Z}^T \mathbf{Z} + N \mathbf{I}_{K_y} \otimes (\boldsymbol{\Lambda}_M \odot \boldsymbol{\Omega}) \right)^{-1} (\hat{\boldsymbol{\Sigma}}^{-1} \otimes \mathbf{Z}^T)$, $\text{vec}(\mathbf{D})$ için bir şapka matrisidir. Maksimum cezalı olabilirlik tahmincisi $\hat{\boldsymbol{\theta}} = \{\hat{\mathbf{B}}, \hat{\boldsymbol{\Sigma}}\}$, (3.25) denkleminde yerine koyularak aşağıdaki olasılık yoğunluk fonksiyonu elde edilir (Matsui ve diğ., 2009; Essomba, 2015; Beyaztaş ve Shang, 2020a):

$$f(\mathbf{Y}_i|\boldsymbol{\theta}) = \frac{1}{(2\pi)^{K_y/2} |\hat{\boldsymbol{\Sigma}}|^{1/2}} \exp \left\{ -\frac{1}{2} (\mathbf{d}_i - \hat{\mathbf{B}}^T \mathbf{z}_i)^T \hat{\boldsymbol{\Sigma}}^{-1} (\mathbf{d}_i - \hat{\mathbf{B}}^T \mathbf{z}_i) \right\} \quad (3.31)$$

4. FONKSİYONEL DOĞRUSAL REGRESYON MODELİNDE KISMİ EN KÜÇÜK KARELER YÖNTEMİ ÜZERİNE YENİ YAKLAŞIMLAR

Bu bölümde, I üzerinde tanımlı yanıt değişkeni skaler bağımsız değişkenleri fonksiyonel yapıdaki doğrusal regresyon modeli (SoFRM) üzerinde durulacaktır. Bu çalışmada, SoFRM tahmininde fonksiyonel kısmi en küçük kareler FPLS yönteminin hesaplama hızını ve sayısal kesinliğini geliştirmek için hibrit iki FPLS regresyon yöntemi önerilmektedir. Önerilen yöntemlerde, sonsuz boyutlu fonksiyonel tahminciler ilk önce bir temel genişletme yöntemi kullanılarak sonlu boyutlu bir uzaya yansıtılır. Daha sonra, fonksiyonel tahmin edicilerin temel katsayıları üzerinde skaler yanıtın regresyon problemini tahmin etmek için Manne (1987) tarafından tanıtılan Bidiag1 ve Bidiag2 olmak üzere iki PLS algoritması ele alınarak, tahmin edicilerin deflasyon adımlarını gerektirmeyen Bidiag1 ve Bidiag2 algoritmaları Krylov alt uzayına dayandırılır. Bu nedenle, SIMPLS ve NIPALS algoritmalarından hesaplama açısından daha hızlıdır. Örneğin, Zhou (2021), fonksiyon - fonksiyon regresyon modeli için Krylov alt uzayına dayalı FPLS yönteminin hızlı bir uygulamasını önermiş ve önerilen Krylov alt uzay tabanlı FPLS yönteminin mevcut FPLS yöntemlerinden hesaplama hızı açısından daha iyi performans sergilediğini göstermiştir. Bidiag1 ve Bidiag2 algoritmaları, hem puan hem de yükleme vektörlerinin tam yeniden ortogonalizasyonunu içermektedir ve bu nedenle sayısal kesinlik için hem gerekli hem de yeterli kabul edilmektedir (Björck ve Indahl, 2017).

\mathcal{H} , $\int_I f^2 dt < \infty$ sağlayan sınırlı ve kapalı I , $f : I \rightarrow \mathbb{R}$ aralığında tanımlanan kare integrallenebilir ve reel değerli fonksiyonların ayrılabilir bir $\mathcal{L}_2(I)$ Hilbert uzayını gösterebilir. $\{Y \in \mathbb{R}, X \in \mathcal{L}_2(I)\}$ kitlesinden rastgele bir örnek $\{Y_i, X_i(t) : i = 1, \dots, n\}$ düşünelim. Burada Y , skaler değerli bir yanıt değişkeni ve X , \mathcal{H} 'de tanımlanan elemanlarla \mathcal{L}_2 sürekli stokastik bir süreçtir. Genellik kaybı olmaksızın, hem Y hem de $X(t)$ 'nin ortalama-sıfır süreçler olduğu varsayılır. Yani, $E[Y] = E[X(t)] = 0$ ve $I \in [0, 1]$ 'dir. Böylece, Y ve $X(t)$ arasındaki ilişkiyi araştırmak için kullanılan SoFRM aşağıdaki gibi tanımlanmaktadır:

$$Y = \int_0^1 X(t)\beta(t)dt + \varepsilon \quad (4.1)$$

Burada, $\beta(t) \in \mathcal{L}_2[0,1]$ regresyon parametresi fonksiyonudur ve ε , sıfır ortalama ve σ^2 varyansı ile normal dağılıma sahip rastgele hata terimidir.

Model (4.1)'de, $\beta(t)$ sonsuz boyutlu bir uzaya ait olduğu için $\beta(t)$ regresyon parametre fonksiyonunu tahmin etmek oldukça zor bir durumdur. $\beta(t)$ 'yi tahmin etmek için LS yönteminin kullanılmasının, aşağıdaki gibi integral Wiener - Hopf denkleminde yol açtığı iyi bilinmektedir:

$$E[Y X(t)] = \int_0^1 E[X(t)X(s)]\beta(s)ds \quad (4.2)$$

Burada, $X(t)$ 'nin kovaryans operatörü genellikle tersinir olmadığı için kötü tasarlanmış bir problemdir (Aguilera ve diğ., 2016). (4.2)'deki ters problemi verimli bir şekilde çözmek için birkaç yöntem önerilmektedir. Bu yöntemlerdeki ortak amaç, boyut küçültme yaklaşımının uygulanmasıdır. Bu yaklaşımla beraber, fonksiyonel tahmin edici ve regresyon katsayısı fonksiyonu ilk önce bir temel fonksiyon açılımı yoluyla sonlu boyutlu bir uzaya yansıtılmaktadır. Daha sonra, temel genişletme katsayılarının vektörü üzerindeki skaler tepkinin regresyon modeli, fonksiyonel tahmin edici üzerindeki skaler tepkinin regresyon problemine yaklaşmak için kullanılmaktadır.

(4.1)'deki SoFRM için FPLS yöntemi, fonksiyonel tahmin edicinin doğrusal fonksiyonelleri olarak ortogonal bileşenleri aşağıdaki gibi üretmektedir:

$$\eta = \int_0^1 X(t)w(t)dt$$

Burada $w(t) \in \mathcal{L}_2[0,1]$, skaler yanıt ile fonksiyonel tahmin edici arasındaki kare kovaryansı optimize ederek elde edilen FPLS bileşenlerini tanımlayan bir FPLS ağırlık fonksiyonu olup aşağıdaki gibi verilebilir:

$$\operatorname{argmax}_{w, \|w\|^2=1} \operatorname{Cov}^2 \left(\int_0^1 X(t)w(t)dt, Y \right) \quad (4.3)$$

C_{YX} ve C_{XY} , sırasıyla $X(t)$ 'nin Y 'ye katkısını ve onun ekini aşağıdaki gibi değerlendiren çapraz kovaryans operatörünü gösterir:

$$C_{YX}: \mathcal{L}_2[0,1] \mapsto \mathbb{R}, f \mapsto \int_0^1 \operatorname{Cov}^2(X(t), Y) f(t) dt$$

$$C_{XY}: \mathbb{R} \mapsto \mathcal{L}_2[0,1], x \mapsto f(t) = x \operatorname{Cov}(X(t), Y)$$

$X(t)$ 'nin \mathcal{L}_2 sürekliliğinden, kendine eş, pozitif ve kompakt bir operatör $U = C_{XY} \circ C_{YX}$ olarak tanımlanmaktadır (Preda ve Saporta, 2005). U 'nun spektral analizi,

$$Uw = \lambda w \quad (4.4)$$

biçiminde verilir. (4.4)'deki denklemin bir çözümünü ortonormal FPLS ağırlık fonksiyonları $w(t)$ (özfonksiyon olarak da adlandırılır) ile ilişkili λ ile gösterilen bir dizi pozitif özdeğer sağlamaktadır. Buna göre (4.3)'te tanımlanan maksimizasyon problemi şu şekilde yeniden tanımlanabilir:

$$\max_{w \in \mathcal{L}_2[0,1]} \frac{\langle Uw, w \rangle}{\langle w, w \rangle}$$

λ_1 ile verilen, yani $Uw_1 = \lambda_1 w_1$ ile tanımlanan en büyük özdeğeriyle ilişkili U tarafından üretilen özfonksiyon $w_1(t)$ ile gösterilsin. Böylece, η_1 ile gösterilen ilk FPLS bileşeni aşağıdaki gibi tanımlanır:

$$\eta_1 = \int_0^1 X(t)w_1(t)dt$$

FPLS, yinelemeli bir yöntemdir ve bu nedenle, sonraki FPLS bileşenleri yinelemeli olarak hesaplanmaktadır. Yani, her h yinelemede, FPLS bileşeni, önceki yinelemeden çıkarılan bilgiler dikkate alınarak hesaplanmaktadır. Daha açık bir şekilde ifade edilirse, $X_0 = X$ ve $Y_0 = Y$ olsun. Ek olarak, X_h ve Y_h sırasıyla, h 'inci FPLS bileşeni η_h üzerindeki X_{h-1} 'in doğrusal regresyonlarının kalıntıları ve η_h üzerindeki Y_{h-1} ile aşağıdaki gibi gösterilebilir:

$$X_h(t) = X_{h-1}(t) - \rho_h(t)\eta_h$$

$$Y_h = Y_{h-1} - c_h\eta_h$$

Burada, $\rho_h(t) = E[X_{h-1}(t)\eta_h]/E[\eta_h^2]$ ve $c_h = E[Y_{h-1}(t)\eta_h]/E[\eta_h^2]$. Böylece, h 'inci FPLS bileşeni aşağıdaki hesaplanır:

$$\eta_h = \int_0^1 X_{h-1}(t)w_h(t)dt$$

Burada, $w_h(t) \in \mathcal{L}_2[0,1]$ olup h 'inci FPLS ağırlık fonksiyonu,

$$w_h = \operatorname{argmax}_{w, \|w\|^2=1} \operatorname{Cov}^2 \left(\int_0^1 X_{h-1}(t)w(t)dt, Y_{h-1} \right)$$

biçiminde verilir. Başka bir deyişle, $w_h(t) U_{h-1} w_h = \lambda_h w_h$ 'nin en büyük özdeğeridir. Burada, $U_{h-1} = C_{XY}^{h-1} \circ C_{YX}^{h-1}$ 'dir ve C_{XY}^{h-1} ile C_{YX}^{h-1} sırasıyla, $X_{h-1}(t)$ ve Y_{h-1} 'in çapraz kovaryans operatörleridir.

4.1 Temel Genişletme Kullanılarak Regresyon Katsayısı Tahmini

(4.1) denklemindeki fonksiyonel tahmin değişkeni $X(t)$ 'nin elemanları sonsuz boyutlu bir uzaya aittir. Bununla birlikte, uygulamada sonlu bir ayrık zaman noktalarında gözlemlenmektedirler. Bu nedenle, ayrık olarak gözlemlenen fonksiyonel tahmin edicilerin elemanlarının fonksiyonel formları, SoFRM'yi yerleştirmeden önce ilk olarak yeniden yapılandırılmalıdır. Bu amaçla yaygın olarak kullanılan pratik yaklaşım, bir dizi temel fonksiyon tarafından yayılan sonlu boyutlu bir uzayda fonksiyonel değişkenlerin sonsuz boyutlu formlarına yaklaşıma dayanmaktadır (Ramsay ve Silverman, 1997; Aguilera ve diğ., 2016). Göreceli olarak çok sayıda K temel genişletme fonksiyonu kullanan bu yaklaşımla, fonksiyonel tahmin edici, temel fonksiyonlar $\phi_k(t) \in \mathcal{L}_2[0,1]$ ve taban katsayısı a_k 'in doğrusal bir kombinasyonu olarak aşağıdaki gibi ifade edilmektedir:

$$X(t) = \sum_{k=1}^K a_k \phi_k(t) = \mathbf{a}^T \boldsymbol{\phi}(t)$$

Burada, $\mathbf{a} = [a_1, \dots, a_K]^T$ 'dir. (4.4) denkleminde, FPLS ağırlık fonksiyonları aynı $\boldsymbol{\phi}(t)$ bazında aşağıdaki gibi ifade edilmektedir:

$$w_h(t) = \sum_{k=1}^K w_{hk} \phi_k(t) = \mathbf{w}_h^T \boldsymbol{\phi}(t), \quad h \geq 1$$

Burada $\mathbf{w}_h = [w_{h1}, \dots, w_{hK}]^T$, temel genişletme katsayılarının vektörünü belirtir. Ayrıca, regresyon parametresi fonksiyonu $\beta(t)$, $\boldsymbol{\phi}(t)$ bazında aşağıdaki gibi ifade edilmektedir:

$$\beta(t) = \sum_{k=1}^K \beta_k \phi_k(t) = \boldsymbol{\beta}^T \boldsymbol{\phi}(t)$$

Burada, $\boldsymbol{\beta} = [\beta_1, \dots, \beta_K]^T$.

$\Phi = \int_0^1 \phi(t)\phi^T(t)dt$ $K \times K$ boyutlu simetrik iç çarpım matrisi olsun. $\Phi^{1/2}$ de Φ 'nin karekökünü gösterebilir. Preda ve Saporta (2005) ve Aguilera ve diğ. (2016), $X(t)$ üzerindeki Y 'nin FPLS regresyonunun, rastgele tasarım matrisi $\mathbf{A} = (\Phi)^T \mathbf{a}$ üzerindeki Y 'nin PLS regresyonuna eşdeğer olduğunu, böylece PLS algoritmasının her h adımında her iki yaklaşımın da aynı PLS bileşenleri ürettiğini göstermişlerdir. FPLS algoritmasının h 'inci adımında, temel genişletme katsayıları cinsinden FPLS ağırlık fonksiyonu $w_h(t)$ aşağıdaki gibi elde edilmektedir:

$$w_h = (\Phi^{-1/2})^T \tilde{w}_h$$

Burada \tilde{w}_h , (4.4) denkleminin benzer şekilde,

$$(\Phi^{-1/2})^T C_{h-1} (\Phi^{-1/2}) \tilde{w} = \lambda \tilde{w}, \quad \tilde{w} \in \mathbb{R}^p, \langle \tilde{w}, \tilde{w} \rangle = 1$$

biçiminde tanımlanıp, en büyük özdeğerine karşılık gelen özvektördür ve $C_{h-1}(j, k) = Cov(a_{h-1,j} Y_{h-1}) Cov(a_{h-1,k} Y_{h-1})$, $j, k \in 1, \dots, K$ girişli $K \times K$ boyutlu matristir. Başka bir deyişle, FPLS ağırlık fonksiyonu $w(t)$, aşağıdaki özdeğer/özvektör problemine dayalı olarak temel genişletme katsayılarının sonlu boyutlu uzayında elde edilmektedir:

$$C_{h-1} \Phi w = \lambda w, \quad w \in \mathbb{R}^p, \langle w, \Phi w \rangle = 1$$

Yukarıda sunulan sonuçlar, $X(t)$ üzerindeki Y 'nin FPLS regresyonunun sonlu boyutlu probleminin, \mathbf{A} tahmincisi ve metrik Φ ile sonlu boyutlu PLS regresyonuna indirgenebileceğini göstermektedir. T_h , Y 'nin \mathbf{A} üzerindeki PLS regresyonundan elde edilen birinci h PLS bileşenlerinin matrisini gösterebilir. Böylece, (4.1) modelinin ilk h PLS bileşenleri cinsinden yaklaşık formu aşağıdaki gibi tanımlanmaktadır:

$$\hat{Y}_h = 1\hat{Y}_{0h} + \mathbf{T}_h\hat{Y}_h = 1\hat{Y}_{0h} + \mathbf{A}\Phi\hat{\beta}_h$$

Burada \hat{Y}_{0h} , tahmini kesişme noktasıdır; \hat{Y}_h , Y 'nin \mathbf{T}_h üzerindeki regresyonunun tahmini parametre vektörüdür; $\hat{\beta}_h = [\hat{\beta}_{h1}, \dots, \hat{\beta}_{hK}]^T$ temel katsayının tahmin vektörüdür ve aşağıdaki gibi hesaplanmaktadır:

$$\hat{\beta}_h = (\Phi^{-1/2})^T \mathbf{V}^T \hat{Y}_h$$

Burada, $\mathbf{V} = [\tilde{w}_1, \dots, \tilde{w}_h]^T$ biçimindedir. Son olarak, regresyon parametre fonksiyonu $\beta_h(t)$ aşağıdaki gibi tahmin edilmektedir:

$$\hat{\beta}_h(t) = \sum_{k=1}^K \hat{\beta}_{hk} \phi_k(t)$$

4.2 Çift Köşegenleştirme ve PLS

Model (4.1)'i kestirmek için önerilen tüm mevcut PLS yaklaşımları, PLS algoritmalarına, yani SIMPLS ve NIPALS'e dayanan iki iyi bilinen sıra azaltıcı ortogonal projeksiyonlardır. Bu iki algoritma, tahmin edicinin deflasyon adımlarını gerçekleştirdiği için hesaplama açısından uygun olmayabilirler. Modelde birçok fonksiyonel tahmin edici kullanıldığında veya ayrıık olarak gözlemlenen tahmin edicilerin fonksiyonel formlarını yeniden oluşturmak için birçok temel genişletme fonksiyonu K kullanıldığında uygulanabilirlik olmaz. Bu problemin üstesinden gelmek için bu bölümde, Björck ve Indahl (2017) tarafından önerilen Golub - Kahan (1965) çift köşegenleştirmesine (Bidiag1 ve Bidiag2) dayalı iki alternatif algoritma ele alınmıştır.

SoFRM'nin (4.1) modelindeki PLS yaklaşımı için aşağıdaki doğrusal LS problemini ele alalım:

$$\min \| \mathbf{A}\mathbf{B} - \mathbf{Y} \|_2, \quad \mathbf{A} \in \mathbb{R}^{n \times K}, \mathbf{Y} \in \mathbb{R}^n \quad (4.5)$$

\mathbf{A}^\dagger , \mathbf{A} 'nın sözde tersini gösterebilirsin. O halde, $\mathbf{B}^\dagger = \mathbf{A}^\dagger \mathbf{Y}$, (4.5) denklemi ile verilen minimizasyon probleminin benzersiz LS çözümüdür. Björck (2014) de benzer şekilde, (4.5) denklemi ile verilen probleme ilişkin PLS yaklaşımları \mathbf{B}_m (yaklaşık m 'inci PLS çözümleri) için aşağıdaki alt problemi ele almıştır.

$$\min \| \mathbf{A}\mathbf{B}_m - \mathbf{Y} \|_2, \quad \mathbf{B}_h \in \mathcal{K}_h(\mathbf{A}^T \mathbf{A}, \mathbf{A}^T \mathbf{Y}), h = 1, 2, \dots \quad (4.6)$$

Burada, $\mathcal{K}_m(\mathbf{A}^T \mathbf{A}, \mathbf{A}^T \mathbf{Y}) = \text{span}\{\mathbf{A}^T \mathbf{Y}, (\mathbf{A}^T \mathbf{A})\mathbf{A}^T \mathbf{Y}, \dots, (\mathbf{A}^T \mathbf{A})^{h-1} \mathbf{A}^T \mathbf{Y}\}$ Krylov alt uzaylarını gösterir. $1 \leq m \leq H$ PLS bileşenleri için, (4.6) ile verilen alt problemi tam sıraya sahiptir ve böylece, karşılık gelen PLS çözümü \mathbf{B}_m benzersiz bir şekilde tanımlanmaktadır. Björck'un (2014) Lemma 2.1'i takiben, \mathbf{A} 'nın farklı sıfır olmayan tekil değerlerinin sayısı, Krylov alt uzayı $\mathcal{K}_m(\mathbf{A}^T \mathbf{A}, \mathbf{Y})$ 'nin m derecesine eşittir ve \mathbf{B}_m PLS yaklaşımı, sözde ters çözüm $\mathbf{B}^\dagger = \mathbf{A}^\dagger \mathbf{Y}$ 'ye eşittir. $\mathbf{U} \in \mathbb{R}^{n \times n}$ ve $\mathbf{V} \in \mathbb{R}^{K \times K}$ sırasıyla \mathbb{R}^n ve \mathbb{R}^K için ortogonal tabanlar olan kare matrisleri belirtirse, aşağıdaki denklemler elde edilir:

$$\begin{aligned} \mathbf{A}\mathbf{B} = \mathbf{Y} &\Leftrightarrow \mathbf{U}^T \mathbf{A}\mathbf{B} \\ &= \mathbf{U}^T \mathbf{Y} \Leftrightarrow \mathbf{U}^T \mathbf{A}\mathbf{V}\mathbf{V}^T \mathbf{B} \\ &= \mathbf{B}\mathbf{z} = \mathbf{q} \end{aligned}$$

Burada, $\mathbf{B} = \mathbf{U}^T \mathbf{A}\mathbf{V}$, $\mathbf{z} = \mathbf{V}^T \mathbf{B}$ ve $\mathbf{q} = \mathbf{U}^T \mathbf{Y}$ 'dir. Krylov alt uzayı ise aşağıdaki gibi tanımlanabilir:

$$\mathcal{K}_m(\mathbf{A}^T \mathbf{A}, \mathbf{A}^T \mathbf{Y}) = \mathbf{V}\mathcal{K}_m(\mathbf{B}^T \mathbf{B}, \mathbf{B}\mathbf{q})$$

Regresyon parametresine PLS yaklaşımı $\mathbf{B}_m = \mathbf{V}_m \mathbf{z}_m$ eşitliğinin çözümü olarak elde edilmektedir. Burada \mathbf{z}_m ,

$$\min \|\mathbf{Bz} - \mathbf{q}\|_2, \quad \mathbf{z} \in \mathcal{K}_m(\mathbf{B}^T \mathbf{B}, \mathbf{Bq})$$

biçiminde verilen alt problemin çözümüdür.

Mevcut FPLS yöntemlerinde (4.5) ile verilen minimizasyon problemini çözmek için standart PLS algoritmaları kullanılmaktadır. Örneğin, NIPALS algoritması ele alınsın. $\mathbf{A}_0 = \mathbf{A}$ ve $Y_0 = Y$ başlangıç değerleri ile başlar. Böylece, NIPALS algoritması aşağıdaki gibi tanımlanır:

$$\mathbf{v}_m = \mathbf{A}_{m-1}^T Y_{m-1}; \quad \mathbf{v}_m = \mathbf{v}_m / \|\mathbf{v}_m\|_2, \quad m = 1, 2, \dots$$

$$\mathbf{u}_m = \mathbf{A}_{m-1} \mathbf{v}_m; \quad \mathbf{u}_m = \mathbf{u}_m / \|\mathbf{u}_m\|_2$$

$$(\mathbf{p}_m^T, \boldsymbol{\zeta}_m) = \mathbf{u}_m^T (\mathbf{A}_{m-1}, Y_{m-1})$$

$$(\mathbf{A}_m, Y_m) = (\mathbf{A}_{m-1}, Y_{m-1}) - \mathbf{u}_m (\mathbf{p}_m^T, \boldsymbol{\zeta}_m)$$

Burada, \mathbf{u}_m , \mathbf{v}_m ve \mathbf{p}_m sırasıyla puanı, yükleme ağırlığını ve yükleme vektörlerini belirtir. NIPALS algoritmasında, her adım üç matris ve/veya vektör çarpması ve bir matris deflasyonunu zorunlu kılmaktadır. Kayan nokta (flop) aritmetik işlemleri göz önüne alındığında (Björck, 2014), NIPALS algoritmasının her adımı, PLS faktörü başına $8nK$ flop gerektirmektedir; burada K , temel genişletme fonksiyonlarının sayısını ve n , gözlem sayısını belirtir. Algoritmada, \mathbf{u}_m ortogonal projeksiyonlar, önceki yinelemeden yani, modifiye edilmiş Gram - Schmidt ilişkisi yoluyla da ifade edilebilen $(\mathbf{A}_m, Y_m) = (1 - \mathbf{u}_m \mathbf{u}_m^T)(\mathbf{A}_{m-1}, Y_{m-1})$ 'den elde edilen verilerden çıkarılmaktadır. Algoritmadaki denklemler aşağıdaki gibi tanımlanabilir:

$$\mathbf{A} = \mathbf{U}_m \mathbf{P}_m^T + \mathbf{A}_m, \quad Y = \mathbf{U}_m \mathbf{q}_m + Y_m$$

Burada, $\mathbf{q}_m = [\zeta_1, \dots, \zeta_m]^T$, $\mathbf{P}_m = [\mathbf{p}_1, \dots, \mathbf{p}_m]^T$ ve $\mathbf{U}_m = [\mathbf{u}_1, \dots, \mathbf{u}_m]^T$ biçimindedir. Buna göre, regresyon katsayısının PLS yaklaşımı $\mathbf{B}_m = \mathbf{V}_m^T \mathbf{z}_m$ ile verilir. Burada \mathbf{z}_m ,

$$(\mathbf{P}_m^T \mathbf{V}_m) \mathbf{z}_m = \mathbf{q}_m,$$

$$\mathbf{U}_m^T (\mathbf{Y} - \mathbf{A} \mathbf{V}_m \mathbf{z}_m) = 0$$

denklemleri çözülerek elde edilmektedir. $\mathbf{B}_m = \mathbf{U}_m^T \mathbf{A} \mathbf{V}_m = \mathbf{P}_m^T \mathbf{V}_m$ üst iki köşeli matristir ve aşağıdaki gibi tanımlanır:

$$\mathbf{B}_m = \begin{bmatrix} \rho_1 & \theta_2 & 0 & \cdots & 0 \\ 0 & \rho_2 & \theta_3 & \ddots & \vdots \\ \vdots & 0 & \ddots & \ddots & 0 \\ \vdots & \vdots & \ddots & \rho_{m-1} & \theta_m \\ 0 & \cdots & \cdots & 0 & \rho_m \end{bmatrix} \in \mathbb{R}^{m \times m} \quad (4.7)$$

Burada, \mathbf{B}_m^{-1} geri ikame ile $2m$ floplarda çözülen tam bir üst üçgen matristir. Elden (2004) ve Björck (2014) tarafından kanıtlandığı gibi, NIPALS, Krylov alt uzayları için ortogonal tabanları oluşturan \mathbf{U}_m ve \mathbf{V}_m ortogonal matrislerini aşağıdaki gibi üretmektedir:

$$\mathcal{R}(\mathbf{V}_m) = \mathcal{K}_m(\mathbf{A}^T \mathbf{A}, \mathbf{A}^T \mathbf{Y}), \quad \mathcal{R}(\mathbf{U}_m) = \mathcal{K}_m(\mathbf{A} \mathbf{A}^T, \mathbf{A} \mathbf{Y})$$

Ek olarak, aşağıdaki ilişkilerin üst iki köşeli \mathbf{B}_m için geçerli olduğu gösterilmiştir:

$$\mathbf{A} \mathbf{V}_m = \mathbf{U}_m \mathbf{B}_m, \quad \mathbf{A}^T \mathbf{U}_m = \mathbf{V}_m \mathbf{B}_m^T + \theta_{m+1} \mathbf{v}_m \mathbf{e}_m^T$$

Burada, θ_m normalleştirme sabitidir ve \mathbf{e}_m m 'inci yinelemede elde edilen hatayı belirtir.

(4.1) modeline yaklaşmak için SIMPLS ve NIPALS yerine Björck ve Indahl (2017) tarafından önerilen iki alternatif PLS algoritması (Bidiag1 ve Bidiag2) ele alınmıştır. Yaklaşım (4.1) modeli bağlamında, Björck ve Indahl'in (2017) yaklaşımı, dikkörtgen $\mathbf{A} \in \mathbb{R}^{n \times K}$ matrisinin iki köşeli indirgenmesi için doğrudan bir algoritma sağlamaktadır:

$$U^T AV = \begin{bmatrix} \mathbf{B} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \in \mathbb{R}^{n \times K}$$

Burada, \mathbf{B} bir üst iki köşeli matristir ve \mathbf{U}, \mathbf{V} Householder matrislerinin çıktıları olarak aşağıdaki gibi seçilmektedir (Golub ve Kahan, 1965; Björck, 2014):

$$\mathbf{U} = \mathcal{F}_1 \cdots \mathcal{F}_{K-1}, \quad \mathbf{V} = \mathcal{D}_1 \cdots \mathcal{D}_{n-1}$$

Burada, $\mathcal{D} = \mathcal{D}^{-1} = \mathcal{D}^T$ 'yi sağlayan Householder matrisleri, temel ortogonal (dikey) yansıtıcılarıdır ve aşağıdaki gibi tanımlanmaktadır:

$$\mathcal{D} = \mathbf{I} - 2\boldsymbol{\omega}\boldsymbol{\omega}^T, \quad \|\boldsymbol{\omega}\|_2 = 1$$

Burada $\boldsymbol{\omega}$, Householder matrisidir. Householder matrisleri ile ortogonal dönüşümlere dayalı algoritmaların ideal kararlılık özelliklerini sağladığı kanıtlanmıştır (Higham, 2002; Byers ve Xu, 2008; Smoktunowicz ve Wrobel, 2012). Bu iki köşeli indirgeme yaklaşımı, PLS yaklaşımlarını elde etmek için kolaylıkla kullanılabilir.

Başlangıç olarak, \mathcal{F}_m uygulanırsa, \mathbf{A} 'nın ilk m sütunları aşağıdaki gibi üst iki köşeli forma indirgenmektedir:

$$\mathcal{F}_m \cdots \mathcal{F}_1 \mathbf{A} \mathbf{V}_m = \begin{bmatrix} \mathbf{B}_m \\ \mathbf{0} \end{bmatrix}, \quad \mathbf{V}_m = \mathcal{D}_1 \cdots \mathcal{D}_m \begin{bmatrix} \mathbf{I}_m \\ \mathbf{0} \end{bmatrix} \quad (4.8)$$

Burada, \mathbf{B}_m (4.7)'deki gibidir. Böylece, \mathcal{D}_{m+1} uygulandığında, \mathbf{A} 'nın ayrıştırılmış formunun ilk m satırları üst iki köşegen forma indirgenmektedir:

$$U_m^T \mathbf{A} \mathcal{D}_1 \cdots \mathcal{D}_{m+1} = [\widehat{\mathbf{B}}_m \quad \mathbf{0}], \quad U_m = \mathcal{F}_m \cdots \mathcal{F}_1 \begin{bmatrix} \mathbf{I}_m \\ \mathbf{0} \end{bmatrix} \quad (4.9)$$

Burada, $\widehat{\mathbf{B}}_m = [\mathbf{B}_m \quad \theta_{m+1} \mathbf{e}_m] \in \mathbb{R}^{m \times (m+1)}$ biçimindedir. (4.8) ve (4.9)'un transpozu göz önüne alındığında, aşağıdaki bağıntılar geçerli olmaktadır:

$$\mathbf{A}\mathbf{V}_m = \mathbf{U}_m\mathbf{B}_m \quad (4.10)$$

$$\mathbf{A}^T\mathbf{U}_m = \mathbf{V}_{m+1}\widehat{\mathbf{B}}_m^T = \mathbf{V}_m\mathbf{B}_m^T + \theta_{m+1}\mathbf{v}_{m+1}\mathbf{e}_m^T \quad (4.11)$$

Bidiag2 algoritması ($m = 1$) hesaplama ile başlamaktadır.

$$\theta_1\mathbf{v}_1 = \mathbf{A}^T\mathbf{Y}$$

$$\rho_1\mathbf{u}_1 = \mathbf{A}\mathbf{v}_1$$

Ardından, (4.10) ve (4.11)'deki m 'inci sütunları eşitleyerek, $m \geq 2$ için değerleri hesaplamak için aşağıdaki yineleme algoritması kullanılmaktadır:

$$\theta_m\mathbf{v}_m = \mathbf{A}^T\mathbf{u}_{m-1} - \rho_{m-1}\mathbf{v}_{m-1}$$

$$\rho_m\mathbf{u}_m = \mathbf{A}\mathbf{v}_m - \theta_m\mathbf{u}_{m-1}$$

Burada, üst çift köşegen \mathbf{B} 'nin m 'inci satırı olan normalleştirme sabitleri ρ_m ve θ_m normalleştirme koşullarından elde edilmektedir, yani $\|\mathbf{u}_m\|_2 = \|\mathbf{v}_m\|_2 = 1$. Björck (2014), \mathbf{B}_m 'deki hiçbir çift köşeli ögenin sıfır olmadığını kanıtlamış ve ayrıca, iki köşeli indirgeme yoluyla oluşturulan \mathbf{U}_m ve \mathbf{V}_m matrislerinin sırasıyla Krylov alt uzayları $\mathcal{R}(\mathbf{V}_m) = \mathcal{K}_m(\mathbf{A}^T\mathbf{A}, \mathbf{A}^T\mathbf{Y})$ ve $\mathcal{R}(\mathbf{U}_m) = \mathcal{K}_m(\mathbf{A}\mathbf{A}^T, \mathbf{A}\mathbf{A}^T\mathbf{Y})$ için benzersiz ortonormal tabanlar olduğunu belirtmiştir. Matris formunda Bidiag2 algoritması aşağıdaki gibi ifade edilmektedir:

$$\mathbf{V}_m(\theta\mathbf{e}_1) = \mathbf{A}^T\mathbf{Y}$$

$$\mathbf{A}\mathbf{V}_m = \mathbf{U}_m\mathbf{B}_m \quad (4.12)$$

$$\mathbf{A}^T\mathbf{U}_m = \mathbf{V}_m\mathbf{B}_m^T + \theta_{m+1}\mathbf{v}_{m+1}\mathbf{e}_m^T$$

Burada, \mathbf{e}_1 ve \mathbf{e}_m \mathbb{R}^m 'de ilk ve m 'inci temel vektörlerdir. Indahl (2014), \mathbf{A} için yükleme matrisi $\mathbf{P}_m = \mathbf{A}^T \mathbf{U}_m$ olduğunu kanıtlamıştır. Ayrıca, soldan \mathbf{U}_m^T ile (4.12) çarpımı, $\mathbf{B}_m = \mathbf{P}_m^T \mathbf{V}_m$ biçiminde iki köşeli matrisle sonuçlanmaktadır.

Bidiag1 algoritmasında işlem ($m = 1$) hesaplama ile başlamaktadır.

$$\gamma_1 \mathbf{u}_1 = Y$$

$$\alpha_1 \mathbf{v}_1 = \mathbf{A}^T \mathbf{u}_1$$

Ardından, $m \geq 2$ değerlerini hesaplamak için aşağıdaki özyineleme kullanılmaktadır:

$$\gamma_m \mathbf{u}_m = \mathbf{A} \mathbf{v}_{m-1} - \alpha_{m-1} \mathbf{u}_{m-1}$$

$$\alpha_m \mathbf{v}_m = \mathbf{A}^T \mathbf{u}_m - \gamma_m \mathbf{v}_{m-1}$$

Burada normleştirme sabitleri α_m ve γ_{m+1} \mathbf{G} alt iki köşeli matrisin m 'inci sütununun elemanlarıdır.

Bidiag2'ye benzer şekilde, Bidiag1 algoritması matris formunda aşağıdaki gibi tanımlanmaktadır:

$$\mathbf{U}_{m+1}(\gamma_1 \mathbf{e}_1) = Y$$

$$\mathbf{A} \mathbf{V}_m = \mathbf{U}_{m+1} \mathbf{G}_m$$

$$\mathbf{A}^T \mathbf{U}_{m+1} = \mathbf{V} \mathbf{G}_m^T + \alpha_{m+1} \mathbf{v}_m \mathbf{e}_m^T$$

Bidiag1 algoritmasındaki \mathbf{v}_m vektörleri Bidiag2 algoritmasında sunulanlara eşit olmaktadır. Öte yandan, \mathbf{u}_m vektörleri Krylov alt uzayı $\mathcal{K}_m(\mathbf{A} \mathbf{A}^T, Y)$ için ortogonal tabanlı vektörlerdir (Elden, 2004; Björck, 2014; Björck ve Indahl, 2017).

Bidiag1 ve Bidiag2 algoritmaları tarafından gerçekleştirilen PLS'nin h -adımlarından sonra, regresyon katsayıları $\mathbf{B}_m \mathbf{V}_m = \mathbf{z}_m$ olarak elde edilmektedir. Burada \mathbf{z}_m , Bidiag1 için alt iki köşeli sistemi ve Bidiag2 için üst iki köşeli sistemi aşağıdaki gibi karşılamaktadır:

$$\mathbf{B}_m \mathbf{z}_m = \mathbf{q}_m, \quad \mathbf{G}_m \mathbf{z}_m = \mathbf{q}_m$$

Bu iki köşeli sistemler, geri iterasyon kullanılarak minimum hesaplama çabasıyla çözülebilmektedir (Björck ve Indahl, 2017). NIPALS ve SIMPLS algoritmalarının aksine Bidiag1 ve Bidiag2 algoritmalarında \mathbf{A} için herhangi bir deflasyon yapılmaz. Bunun yerine, geleneksel algoritmalara kıyasla hesaplama maliyetini azaltan yalnızca vektör matrisi çarpımları, yani \mathbf{AV} ve \mathbf{AU} kullanılır. NIPALS algoritmasının aksine, Bidiag1 ve Bidiag2 algoritmaları adım başına $8(n - m)(K - m)$ flop gerektirir. Burada m , PLS bileşenlerinin sayısını göstermektedir.

5. UYGULAMA

Bu bölümde, birkaç Monte - Carlo deneyi ve şeker işleme verilerinin ampirik analizi ile önerilen Bidiag1 ve Bidiag2 tabanlı FPLS yöntemlerinin mevcut referans SIMPLS ve NIPALS tabanlı FPLS yöntemleriyle hesaplama verimliliği ve sayısal hassasiyet açısından karşılaştırması yapılmıştır. Björck ve Indahl (2017), Bidiag1 ve Bidiag2 algoritmalarının çalışmadaki sayısal analiz tarafından onaylanan özdeş sonuçlar ürettiğine dikkat çekmiştir. Björck ve Indahl (2017), Bidiag2'nin biraz daha basit bir uygulamaya sahip olması nedeniyle Bidiag2'nin Bidiag1'e tercih edilen alternatif olabileceği sonucuna varmıştır. Bu nedenle, çalışmada sadece Bidiag2 tabanlı FPLS yönteminin sonuçları sunulmaktadır. Bu çalışmada ele alınan tüm sayısal analizler R 4.1.1 programı kullanılarak elde edilmiştir. Önerilen ve mevcut yöntemler için örnek bir R kodu Ek A kısmında yer almaktadır. Aynı zamanda, örnek R koduna https://github.com/SemanurSaricam/SoFRM_FPLS adresinden ulaşılabilir.

5.1 Monte – Carlo Deneyleri

Monte-Carlo deneyleri boyunca, basit bir SoFRM düşünülür ve fonksiyonel tahmin edicinin unsurlarını oluşturmak için aşağıdaki süreç kullanılır:

$$X(t) = \sum_{k=1}^4 \alpha_k \phi_k(t), \quad t \in [0,1]$$

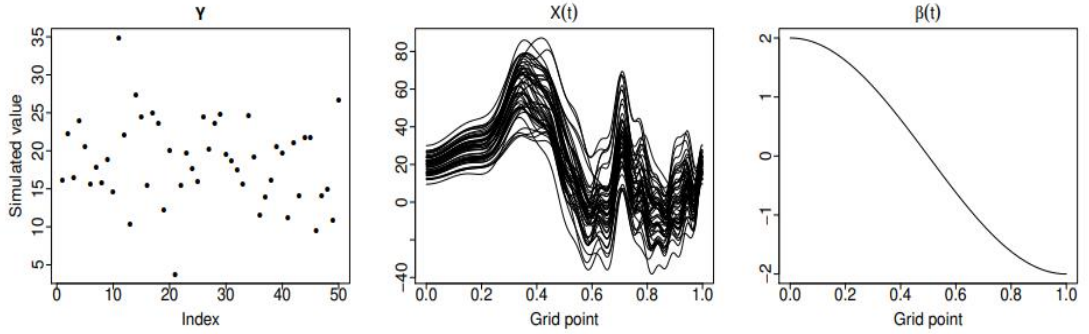
Burada, $\alpha_k \sim \mathcal{N}(4, 3k^{-1/2})$ ve $\phi_k(t) = 4 \sin(k\pi t^2) + \cos(4k\pi t^2)$. Regresyon parametre fonksiyonu, aşağıdaki gibi bir kosinüs fonksiyonu kullanılarak oluşturulur:

$$\beta(t) = 2 \cos(\pi t), \quad t \in [0,1]$$

Ardından, yanıt değişkeninin değerleri aşağıdaki gibi oluşturulur:

$$Y = \int_0^1 X(t)\beta(t)dt + \varepsilon$$

Burada, $\varepsilon \sim \mathcal{N}(0,1)$. Deneysel olarak, fonksiyonel tahmin edicinin elemanları $[0,1]$ aralığında 1001 eşit aralıklı noktada üretilir. Üretilen verinin ve parametre fonksiyonunun grafiksel gösterimi Şekil 1'de sunulmaktadır.



Şekil 5.1 Oluşturulan skaler yanıtın (sol panel), fonksiyonel tahmin edicinin (orta panel) ve regresyon parametre fonksiyonunun (sağ panel) grafiksel gösterimi.

Yöntemlerin sayısal kesinliğini karşılaştırmak amacıyla, fonksiyonel tahmin edici ve skaler yanıt değişkenleri için $n = 1000$ gözlem ve $K = [25, 50, 100, 250, 500, 750]$ kübik B-spline temel fonksiyonlar, fonksiyonel gözlemleri temel genişletme katsayılarının sonlu boyutlu uzayına yansıtmak için kullanılır. Her K için, sözde rastgele 500 Monte-Carlo simülasyon verisi üretilir. Her simülasyonda, önerilen yöntemlerin sayısal kesinliğini değerlendirmek için ilk beş FPLS bileşeni kullanılarak aşağıdaki hata kareler ortalaması (Mean Squared Error, MSE) ve entegre hata kareler ortalaması (Mean Integrated Squared Error, MISE) hesaplanır:

$$MSE = \frac{1}{1000} \sum_{i=1}^{1000} (Y_i - \hat{Y}_i)^2$$

$$MISE = \int_0^1 [\beta(t) - \hat{\beta}(t)]^2 dt$$

500 Monte - Carlo simülasyonu üzerinden hesaplanan ortalama MSE ve MISE değerleri Çizelge 5.1'de sunulmuştur. Sonuçlar, Bidiag2 (ve Bidiag1) tabanlı FPLS yöntemlerinin, NIPALS tabanlı FPLS yöntemiyle aynı MSE ve MISE değerlerini ürettiğini göstermektedir. SIMPLS tabanlı FPLS yöntemi, diğer yöntemlerle karşılaştırıldığında genel olarak daha kötü performans üretmektedir. SIMPLS yönteminin kararsızlığı, başlangıç vektörleri olarak daha yüksek dereceli Krylov vektörlerinin kullanılmasından kaynaklanmaktadır (Björck ve Indahl, 2017).

Çizelge 5.1 Sayısal kesinlik karşılaştırması: Hesaplanan ortalama MSE ve MISE değerleri.

Kriter	K	Yöntem	
		SIMPLS	NIPALS ve Bidiag2
MSE	25	1.0489	1.0189
	50	1.0312	0.9760
	100	0.9767	0.9655
	250	0.9517	0.9474
	500	0.7173	0.7173
	750	0.7173	0.7173
	MISE	25	0.7083
50		0.7173	0.7172
100		0.7175	0.7175
250		0.7172	0.7172
500		1.0149	0.9909
750		1.0244	0.9950

Ayrıca Bidiag2 (ve Bidiag1) tabanlı FPLS yöntemleri SIMPLS ve NIPALS tabanlı FPLS yöntemleriyle hesaplama süreleri açısından karşılaştırılmıştır. Bunu yaparken yöntemlerin hesaplama süreleri sadece B-spline tabanlı genişletme katsayılarının

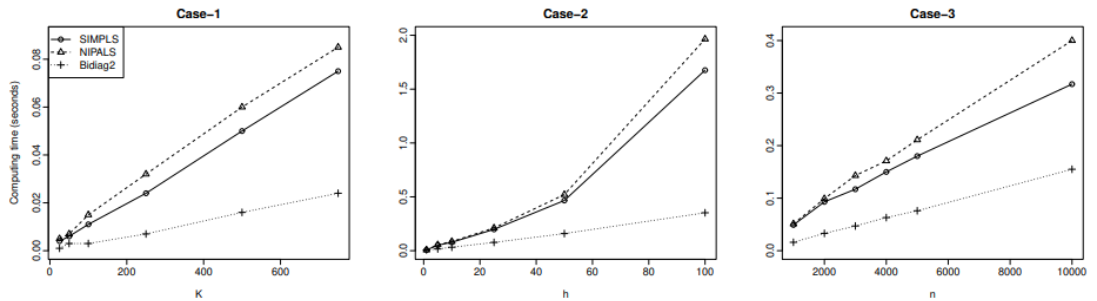
sonlu boyutlu uzayında yapılan hesaplamalar için kaydedilir. Bunun nedeni, fonksiyonel değişkenleri sonlu boyutlu uzaya yansıtma sürecinin tüm yöntemler için aynı olmasıdır. Yöntemlerin hesaplama sürelerini karşılaştırmak için üç durum ele alınmıştır:

Durum-1: $n = 1000$ ve $h = 5$ sabitlenerek, yöntemlerin hesaplama süreleri $K = [25, 50, 100, 250, 500, 750]$ için kaydedilir.

Durum-2: $n = 1000$ ve $K = 500$ sabitlenerek yöntemlerin hesaplama süreleri $h = [1, 5, 10, 25, 50, 100]$ için kaydedilir.

Durum-3: $h = 5$ ve $K = 500$ sabitlenerek, yöntemlerin hesaplama süreleri $n = [1000, 2000, 3000, 4000, 5000, 10000]$ için kaydedilir.

Tüm yöntemler ve durumlar için kaydedilen hesaplama süreleri Şekil 5.2'de sunulmaktadır. Bu şekil, n , K ve h nispeten küçük olduğunda tüm yöntemlerin benzer hesaplama sürelerine sahip olduğunu göstermektedir. Öte yandan, Bidiag2 (ve Bidiag1) tabanlı FPLS yöntemleri, $[n, K, h]$ bileşenlerinden biri büyük olduğunda SIMPLS ve NIPALS tabanlı FPLS yöntemlerinden daha az hesaplama gerektirir. SIMPLS ve NIPALS algoritmaları, Durum-1 ($K = 750$ olduğunda) için Bidiag2 (ve Bidiag1) algoritmasından üç kat, Durum-2 için ($h = 100$ olduğunda) Bidiag2'den (ve Bidiag1) 5,4 kat daha fazla hesaplama süresi gerektirir ve Durum-3 için ($n = 10000$ olduğunda) Bidiag2'den (ve Bidiag1) 2,4 kat daha fazla hesaplama süresi gerektirir.

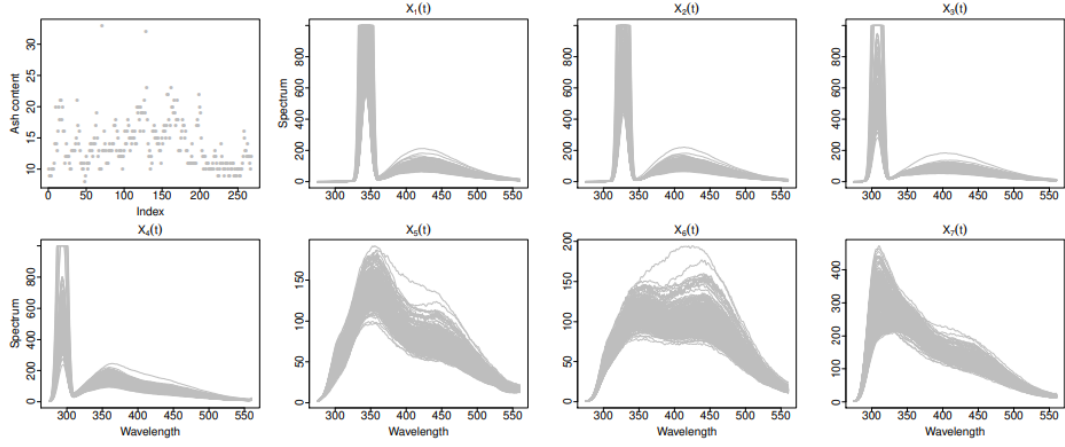


Şekil 5.2 K (sol panel), h (orta panel) ve n (sağ panel) değerlerini artırmak için SIMPLS, NIPALS ve Bidiag2 tabanlı FPLS yöntemleri için tahmini hesaplama süreleri.

Monte-Carlo deney sonuçları, önerilen Bidiag2 (ve Bidiag1) tabanlı FPLS yönteminin, MSE ve MISE açısından kararlı sayısal kesinlik ürettiğini ve mevcut FPLS yöntemlerine kıyasla önemli ölçüde daha az hesaplama süresi gerektirdiğini göstermektedir. FPLS yöntemlerinin çoğunun, sayısal olarak kesin sonuçlar üretemeyen SIMPLS'ye dayalı olduğu göz önüne alındığında, önerilen yöntemlerin SIMPLS tabanlı FPLS yöntemlerine uygun bir alternatif olduğu düşünülebilir. Ek olarak, Şekil 5.3'te sunulan hesaplama süreleri sonuçlarına dayanarak, önerilen yöntemler NIPALS tabanlı FPLS yöntemlerine daha hızlı bir alternatif olarak kabul edilebilir ve özellikle ultra yoğun ve/veya çoklu fonksiyonel tahmin durumları için önerilmektedir.

5.2 Şeker İşlem Verisi

Bu kısımda, Munck ve diğ. (1998) ve Bro (1999) tarafından açıklanan ve http://models.life.ku.dk/Sugar_Process adresinde verilen şeker işlem veri seti kullanılmıştır. Şeker, bu veri setinde tamponlanmamış suda çözündürülmüş ve çözelti, spektro - florometrik olarak ölçülmüştür. Emisyon spektrumları, analiz edilen 268 şeker numunesinin her biri için yedi uyarma dalga boyunda (nm) [230, 240, 255, 290, 305, 325, 340] 275'den 560'a (toplamda 571 dalga boyu) 0.5 nm aralıklarla ölçülmüştür. Ek olarak, 268 şeker numunesinden rafine şekerdeki inorganik safsızlıkların sayısının bir ölçüsü olan bir kalite parametresi olan "kül içeriği" florometrik olarak ölçülmüştür. Daha sonra, Smaga ve Matsui (2018) ve Gartheiss ve diğ. (2013) bu veri seti üzerinde fonksiyonel regresyonda değişken seçimi üzerine çalışmışlardır. Bu çalışmadaki amaç ise, yedi fonksiyonel öngörücü (farklı uyarma dalga boylarında elde edilen şeker numunelerinin emisyon spektrumları) kullanarak kül içeriğini tahmin etmektir. Şeker numunelerinin kül içeriği ve emisyon spektrumunun grafiksel bir gösterimi Şekil 5.3'te sunulmaktadır.



Şekil 5.3 Uyarım dalga boylarında [230, 240, 255, 290, 305, 325, 340] = $[X_1(t), X_2(t), X_3(t), X_4(t), X_5(t), X_6(t), X_7(t)]$ elde edilen 268 kül içeriği ve şeker numunelerinin emisyon spektrumlarının grafiksel gösterimi.

Şeker işlem verilerindeki değişkenler ile aşağıdaki SoFRM'yi dikkate alınız:

$$Y = \sum_{m=1}^7 \int_{275}^{560} X_m(t) \beta_m(t) dt, \quad t \in [275, 560]$$

$$= \int_{275}^{560} \mathbf{X}^T(t) \boldsymbol{\beta}(t) dt$$

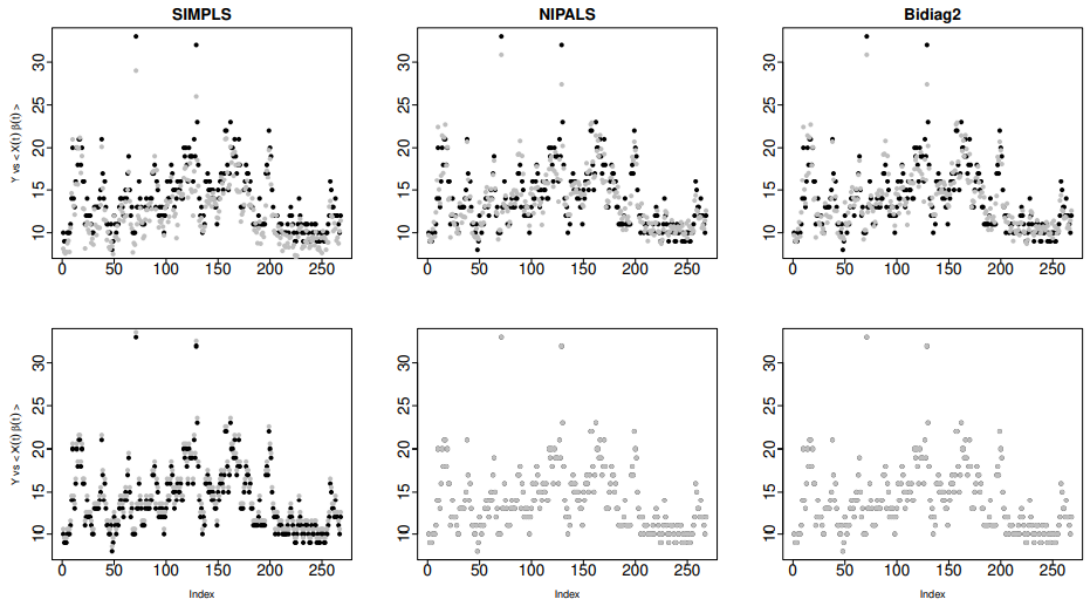
Burada, $\mathbf{X}(t) = [X_1(t), X_2(t), X_3(t), X_4(t), X_5(t), X_6(t), X_7(t)]^T$ ve $\boldsymbol{\beta}(t) = [\beta_1(t), \beta_2(t), \beta_3(t), \beta_4(t), \beta_5(t), \beta_6(t), \beta_7(t)]^T$. Bu veri kümesi için, yukarıdaki modeli tahmin etmede için tüm FPLS yöntemleri kullanılmış ve MSE değerleri, farklı K ve h değerleri kümeleri için hesaplanmıştır. Ayrıca, her K ve h için geçen süreler kaydedilmiştir. Sonuçlar Çizelge 5.2'de sunulmaktadır. Sonuçlara göre, Monte-Carlo deneylerinde olduğu gibi, önerilen Bidiag2 (ve Bidiag1) tabanlı FPLS yöntemi, NIPALS tabanlı FPLS yöntemine benzer MSE değerleri üretmektedir. SIMPLS tabanlı FPLS yöntemi, tüm K ve h değerleri için Bidiag2 ve NIPALS tabanlı FPLS yöntemleriyle karşılaştırıldığında önemli ölçüde daha kötü performans üretmektedir. Fonksiyonel tahmin edicileri yansıtmak için az sayıda temel fonksiyon kullanıldığında

(örneğin, $K = 10$), SIMPLS tabanlı yöntem Bidiag2 ve NIPALS tabanlı yöntemlerden 3.43 ($h = 15$ olduğunda) ve 213.36 ($h = 50$ olduğunda) kat daha büyük MSE değerleri üretir. SIMPLS tabanlı yöntemin tahmin performansı K arttıkça artarken, yine de Bidiag2 ve NIPALS tabanlı yöntemlerin tahmin performansından çok daha kötü tahmin performansına sahiptir. Örneğin, $K = 250$ olduğunda, SIMPLS tabanlı yöntem, Bidiag2 ve NIPALS tabanlı yöntemlerden 2.94 ($h = 15$ olduğunda) ila 3186 ($h = 50$ olduğunda) kat daha büyük MSE değerleri üretir. SIMPLS algoritmasının sayısal kararsızlığı Andersson (2009) ve Björck ve Indahl (2017) tarafından da ifade edilmiş ve SIMPLS algoritmasının kararsızlığın nedeni, başlangıç vektörleri olarak daha yüksek dereceli Krylov vektörlerinin kullanılması olarak belirtilmiştir. Göreceli olarak daha düşük tahmin performansının yanı sıra, SIMPLS tabanlı yöntem ayrıca önerilen Bidiag2 (ve Bidiag1) tabanlı FPLS yönteminden hesaplama açısından daha yavaştır. Önerilen Bidiag2 tabanlı yöntem genellikle en hızlı yöntemdir.

Çizelge 5.2 Şeker işlem verileri için tüm PLS yöntemlerinin hesaplanan MSE değerleri. Kaydedilen hesaplama süreleri (saniye cinsinden) parantez içinde verilmiştir ve en hızlı yöntem/yöntemler koyu renkle işaretlenmiştir.

<i>K</i>	<i>h</i>	Yöntem		
		SIMPLS	NIPALS	Bidiag2
10	5	13.0601(0.002)	3.1774 (0.003)	3.1774 (0.002)
	10	7.4927 (0.003)	1.9937 (0.007)	1.9937 (0.003)
	15	5.9533 (0.009)	1.7325 (0.010)	1.7325 (0.004)
	25	88.6348 (0.010)	1.3863 (0.012)	1.3863 (0.005)
	50	263.2891(0.015)	1.2340 (0.019)	1.2340 (0.009)
25	5	16.4864 (0.003)	3.2498 (0.004)	3.2498 (0.003)
	10	12.7225 (0.004)	2.0418 (0.008)	2.0418 (0.004)
	15	3.7031 (0.007)	1.6610 (0.011)	1.6610 (0.005)
	25	18.7357 (0.013)	0.8592 (0.018)	0.8592 (0.008)
	50	195.0939(0.026)	0.5022 (0.046)	0.5022 (0.017)
50	5	36.5621 (0.007)	4.0226 (0.007)	4.0226 (0.004)
	10	8.9968 (0.010)	2.0560 (0.013)	2.0560 (0.007)
	15	14.4892 (0.013)	1.5631 (0.019)	1.5631 (0.008)
	25	17.2403 (0.019)	0.6636 (0.039)	0.6636 (0.010)
	50	58.2207 (0.042)	0.0463 (0.080)	0.0463 (0.024)
100	5	40.6034 (0.011)	4.1210 (0.014)	4.1210 (0.007)
	10	8.4830 (0.015)	1.9945 (0.032)	1.9945 (0.009)
	15	1.7659 (0.024)	1.5066 (0.042)	1.5066 (0.015)
	25	12.3059 (0.033)	0.4561 (0.068)	0.4561 (0.020)
	50	0.7924 (0.081)	0.0005 (0.147)	0.0005 (0.061)
250	5	40.7240 (0.027)	4.1154 (0.034)	4.1154 (0.015)
	10	10.0299 (0.038)	1.9839 (0.075)	1.9839 (0.022)
	15	4.0632 (0.052)	1.3790 (0.125)	1.3790 (0.029)
	25	22.2482 (0.095)	0.2930 (0.189)	0.2930 (0.041)
	50	0.3186 (0.285)	0.0001 (0.432)	0.0001 (0.120)

Ek olarak, $[K, h] = [25, 15]$ ve $[K, h] = [250, 50]$ iki durum için gözlemlenen ve uydurulan kül içeriği değerlerinin (yani, $\langle X(t), \hat{B}(t) \rangle$)'ye karşı Y grafiklerinin sonuçları Şekil 5.4'te verilmiştir. Bu şekilde önerilen yöntemler SIMPLS tabanlı FPLS yöntemiyle karşılaştırıldığında, önerilen yöntemlerin kül içeriği için geliştirilmiş-uyumlu değerler ürettiğini göstermektedir. NIPALS tabanlı FPLS yöntemi, önerilen yöntemlerden daha fazla hesaplama süresi ile aynı sonuçları üretir.



Şekil 5.4: $[K, h] = [25, 15]$ (birinci satır) ve $[K, h] = [250, 50]$ (ikinci satır) olduğunda $\langle X(t), \hat{B}(t) \rangle$ 'ye karşı Y grafiği.

Grafiklerde, siyah noktalar gözlemlenen kül içeriğini, gri noktalar ise yöntemlere göre uygun kül içeriği değerlerini gösterir. $[K, h] = [250, 50]$ (ikinci sıra) durumunda, NIPALS ve Bidiag2 tabanlı FPLS yöntemleri için gözlemlenen ve uydurulan kül içeriği değerleri örtüşmektedir.

6. SONUÇ VE ÖNERİLER

SoFRM, bir skaler yanıt ile fonksiyonel öngörücü/öngörücüler arasındaki fonksiyonel ilişkiyi araştırmak için genel bir çerçeve haline gelmiştir. Bu modelin parametrelerini tahmin etmek için çeşitli yöntemler önerilmiştir. Bölüm 4'de tartışılan sayısal ve teorik çalışmalar, FPLS dahil olmak üzere boyut küçültme tekniklerinin, mevcut diğer yöntemlerle karşılaştırıldığında SoFRM'nin tahmininde parametre tahminleri ve hesaplama süresi açısından gelişmiş sonuçlar ürettiğini göstermiştir. Mevcut FPLS yöntemleri, iyi bilinen iki algoritma olan NIPALS ve SIMPLS'ye dayanmaktadır. Bu algoritmalar arasında, NIPALS, modelde birçok fonksiyonel tahmin edici kullanıldığında veya ayırık olarak gözlemlenen tahmin edicilerin fonksiyonel formlarını yeniden oluşturmak için çok sayıda temel fonksiyon kullanıldığında, hesaplama açısından mümkün olmayabilir. Öte yandan, SIMPLS ise kararsızlığı nedeniyle parametre tahmininde düşük performans üretebilir.

Skaler - fonksiyon regresyonda regresyon katsayısı fonksiyonunu kararlı ve verimli bir şekilde tahmin etmek için iki hibrit FPLS algoritması sunulmuştur. Önerilen yöntemler, Björck ve Indahl (2017) tarafından tanımlanan yeniden ortogonalleştirilmiş puanlar ve yükleme vektörleri ile Bidiag1 ve Bidiag2 algoritmalarına dayanmaktadır. Önerilen yöntemlerin sayısal kesinliği ve hesaplama verimliliği, bir dizi Monte - Carlo simülasyon çalışması ve şeker işlem verilerinin kemometrik veri analizi yoluyla değerlendirilmiştir. Önerilen yöntemlerin oldukça kararlı parametre ve model uydurma tahminleri ürettiği gösterilmiştir. Sonuçlar ayrıca, mevcut FPLS yöntemleriyle karşılaştırıldığında, önerilen yöntemlerin önemli ölçüde daha az hesaplama süresi gerektirdiğini göstermektedir. Sayısal analizlere göre, SIMPLS tabanlı FPLS yöntemi (en kararsız) en kötü performansı sergilemektedir. Örneğin, şeker işlem verilerinin analizinde, SIMPLS tabanlı yöntem diğer yöntemlere göre 2.94'den 3186 kat kadar daha fazla MSE değeri üretmektedir. Mevcut FPLS yöntemlerinin çoğu, sayısal olarak kesin sonuçlar üretemeyen SIMPLS'ye dayandığından, önerilen yöntemler SIMPLS tabanlı FPLS yöntemine uygun bir alternatif olarak düşünülebilir.

Bölüm 4’te sunulan çalışma, diğer FDA araçlarına genişletilebilir. İlk olarak, önerilen yöntemler, hem yanıt hem de tahmin değişkenlerinin rastgele eğrilerden oluştuğu, fonksiyon - fonksiyon regresyon modellerini (FoFRM) tahmin etmek için kullanılabilir. Sonsuz boyutlu fonksiyon - fonksiyon regresyon modeli, çok değişkenli bir regresyon modeli ile yaklaştırılır. SoFRM ile karşılaştırıldığında, fonksiyon – fonksiyon regresyon modellerinin tahmini, yanıt değişkeninin birden fazla bileşeninin tahmin edilmesini gerektirdiğinden genellikle daha fazla hesaplama süresi gerektirir. Önerilen yöntemler, bu modeli mevcut yöntemlerden daha verimli bir şekilde tahmin etmek için kullanılabilir. Örneğin, bu tür modellerde, belirli bir fonksiyonel tahmin ediciye yaklaşmak için $K = 20$ temel genişletme fonksiyonları kullanılıyorsa, fonksiyonel tahmin edicilerin ikinci dereceden ve etkileşimli etkilerini tahmin etmek için $20 \times 20 = 400$ tensör çarpım bazında genişletme fonksiyonları kullanılır. İkincisi, fonksiyonel tahmin edicilerin kuadratik ve etkileşim etkilerine sahip FRM, bir modelde tahmin edilmesi için önemli ölçüde daha fazla temel katsayı ve parametre gerektirir (Usset ve diğ., 2016; Fuchs ve diğ., 2015; Luo ve Qi, 2019; Matsui, 2020; Sun ve Wang, 2020; Beyaztaş ve Shang, 2021). Bu nedenle, bu tür modeller, tahmin adımı ana etki modellerinden çok daha fazla hesaplama süresi gerektirir. Önerilen yöntemler, bu tür modellerin tahmin aşamasının hesaplama maliyetini azaltmak için kolayca bu modellere genişletilebilir.

KAYNAKLAR

- Aguilera, A. M., Aguilera-Morillo, M. C. ve Preda, C. (2016). Penalized versions of functional pls regression. *Chemometrics and Intelligent Laboratory Systems*, 154, 8–92.
- Aguilera, A. M., Escabias, M., Preda, C. ve Saporta, G. (2010). Using basis expansions for estimating functional PLS regression: Applications with chemometric data. *Chemometrics and Intelligent Laboratory Systems*, 104(2), 289–305.
- Aguilera, A. M., Escabias, M. ve Valderrama, M. (2008). Discussion of different logistic models with functional data. application to systemic lupus erythematosus. *Computational Statistics & Data Analysis*, 53(1):151–163.
- Aguilera, A. M., Escabias, M., Valderrama, M. J. ve Aguilera-Morillo, M. C. (2013). Functional analysis of chemometric data. *Open J Stat*, 3(5):334–343, 2013.
- Andersson, M. (2009). A comparison of nine PLS1 algorithms. *J Chemom*, 23(10):518–529.
- Besse, P. ve Ramsay, J. O. (1986). Principal components analysis of sampled functions. *Psychometrika*, 51(2):285–311.
- Beyaztaş, U. ve Shang, H. L. (2020a). A comparison of parameter estimation in function-on-function regression. *Communications in Statistics - Simulation and Computation*, DOI: 10.1080/03610918.2020.1746340.
- Beyaztaş, U. ve Shang, H. L. (2020b). On function-on-function regression: partial least squares approach. *Environmental and Ecological Statistics*, 27, 95–114.
- Beyaztaş, U. ve Shang, H. L. (2021). A partial least squares approach for function-on-function interaction regression. *Comput Stat*, 36(2):911–939.
- Björck, A. (2014). Stability of two direct methods for bidiagonalization and partial least squares. *SIAM J Matrix Anal Appl*, 35(1):279–291.
- Björck, A. ve Indahl, U. G. (2017). Fast and stable partial least squares modelling: A benchmark study with theoretical comments. *J Chemom*, 31(8):e2898.
- Bosq, D. (2000). *Linear Processes in Function Spaces*. Springer, New York.
- Brezis, H. (2010). *Functional Analysis, Sobolev Spaces and Partial Differential Equations*. Springer.
- Bro, R. (1999). Exploratory study of sugar production using fluorescence spectroscopy and multi-way analysis. *Chemom Intell Lab Syst*, 46(2):133–147.
- Byers, R. ve Xu, H. (2008). A new scaling for newton’s iteration for the polar decomposition and its backward stability. *SIAM J Matrix Anal Appl*, 30(2):822–843.
- Cai, X., Xue, L. ve Cao, J. (2022). Robust estimation and variable selection for function-on-scalar regression. *The Canadian Journal of Statistics*, Vol. 50, No.1, 162–179.

- Cardot, H., Ferraty, F. ve Sarda, P. (1999). Functional linear models. *Statistics & Probability Letters* 45, 11–22.
- Cardot, H., Ferraty, F. ve Sarda, P. (2003). Spline estimators for the functional linear model. *Stat Sinica*, 13(3): 2159–2179.
- Chen, Y., Goldsmith, J. ve Ogden, T (2016). Variable selection in function-on-scalar regression. *The ISI's Journal for the Rapid Dissemination of Statistics Research*. DOI: 10.1002/sta4.106.
- Cuevas, A. (2014). A partial overview of the theory of statistics with functional data. *J Stat Plan Inference*, 147: 1–23.
- De Boor, C. (2001) *A Practical Guide to Splines*. Revised Edition. New York: Springer.
- Delaigle, A. ve Hall, P. (2012). Methodology and theory for partial least squares applied to functional data. *Ann Stat*, 40(1):322–352.
- Denizli, N. (2016). *Uzaktan Algılama Verilerine Fonksiyonel Veri Analizi Yaklaşımının Uygulanması*. (Doktora Tezi). Mimar Sinan Güzel Sanatlar Üniversitesi, Fen Bilimleri Enstitüsü, İstanbul.
- Elden, L. (2004). Partial least-squares vs. Lanczos bidiagonalization—I: analysis of a projection method for multiple regression. *Comput Stat Data Anal*, 46(1):11–31.
- Erdoğan, M. (2020). *Fonksiyonel Veri Analizi İle Yaşa Özel Doğurganlık Hızlarının Tahmin Edilmesi*, Yüksek Lisans Tezi, Ankara Üniversitesi, Sağlık Bilimleri Enstitüsü, Ankara.
- Escabias, M., Aguilera, A. M., ve Valderrama, M. J. (2004). Principal component estimation of functional logistic regression: discussion of two different approaches. *Journal of Nonparametric Statistics*, 16(3-4):365–384.
- Escabias, M., Aguilera, A. M., ve Valderrama, M. J. (2005). Modeling environmental data by functional principal component logistic regression. *Environmetrics*, 16(1):95–107.
- Essomba, R.F. (2015). *An Investigation into Functional Linear Regression Modeling*. (Master's Thesis). https://open.uct.ac.za/bitstream/item/16664/thesis_sci_2015_essomba_rene_franck.pdf?sequence=1.
- Febrero-Bande, M., Galeano, P. ve Gonzalez-Manteiga, W. (2017). Functional principal component regression and functional partial least-squares regression: An overview and a comparative study. *Int Stat Rev*, 85(1):61–83.
- Ferraty, F. ve Vieu, P. (2006). *Nonparametric Functional Data Analysis*. Springer, New York.
- Fuchs, K., Scheipl, F. ve Greven, S. (2015). Penalized scalar-on-functions regression with interaction term. *Comput Stat Data Anal*, 81:38–51.
- Gartheiss, J., Maity, A. ve Staicu, A. M. (2013). Variable selection in generalized functional linear models. *Stat*, 2(1):86–101.

- Goldsmith, J., Bobb, J., Crainiceanu, C. M., Caffo, B. ve Reich, D. (2011). Penalized functional regression. *J Comput Graphical Stat*, 20(4):830–851.
- Goldsmith, J. ve Scheipl, F. (2014). Estimator selection and combination in scalar-on-function regression. *Comput Stat Data Anal*, 70:362–372.
- Golub, G. H. ve Kahan, W. (1965). Calculating the singular values and pseudoinverse of a matrix. *SIAM Ser B*, 2(2):205–224.
- Grenander, U. (1950). Stochastic processes and statistical inference. *Arkiv för Matematik*, Vol. 1(3), 195-277.
- Hall, P. Ve Horowitz, J. L. (2007). Methodology and convergence rates for functional linear regression. *Ann Stat*, 35(1):70–91.
- Higham, N. J. (2002). *Accuracy and Stability of Numerical Algorithms*. SIAM, Philadelphia.
- Horvath, L. ve Kokoszka, P. (2012). *Inference for Functional Data with Applications*. Springer, New York.
- Indahl, U. G. (2014). The geometry of PLS1 explained properly: 10 key notes on mathematical properties and some alternative algorithmic approaches to PLS1 modelling. *J Chemom*, 28(3):168–180.
- Kalogridis, I. ve Van Aelst, S. (2019). Robust functional regression based on principal components. *J Multivariate Anal*, 173:393–415.
- Kokoszka, P. ve Reimherr, M. (2017). *Introduction to Functional Data Analysis*. CRC Press, Boca Raton.
- Keser, İ. ve Ertaş, K. (2007). Düzgünleştirilmiş Fonksiyonel Ana Bileşenler Analizi ve Bir Uygulama. *D.E.Ü.İ.İ.B.F. Dergisi Cilt:22 Sayı:1*, ss:1-26.
- Keser, İ. (2010). Ege Bölgesi Yağış Verilerinin Fonksiyonel Veri Analizi İle İncelenmesi. *Dokuz Eylül Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, Cilt:25, Sayı:1, ss.41-67.
- Kramer, N., Boulesteix, A. L. ve Tutz, G. (2008). Penalized partial least squares with applications to b-spline transformations and functional data. *Chemom Intell Lab Syst*, 94(1):60–69.
- Lee, E. R. ve Park, B. U. (2012). Sparse estimation in functional linear regression. *J Multivariate Anal*, 105(1): 1–17.
- Luo, R. ve Qi, X. (2019). Interaction model and model selection for function-on-function regression. *J Comput Graphical Stat*, 28(2):309–322.
- Manne, R. (1987). Analysis of two partial-least-squares algorithms for multivariate calibration. *Chemom Intell Lab Syst*, 2:187–197.
- Manteiga, W. G. ve Vieu, P. (2007). Statistics for functional data. *Computational Statistics & Data Analysis*, 51(10):4788–4792.

- Maronna, R. A. ve Yohai, V. J. (2013). Robust functional linear regression based on splines. *Comput Stat Data Anal*, 65:46–55.
- Marx, B. D. ve Eilers, P. H. C. (1999). Generalized linear regression on sampled signals and curves: A P-spline approach. *Technometrics*, 41(1):1–13.
- Matsui, H. (2020). Quadratic regression for functional response models. *Econ Stat*, 13:125–136.
- Matsui, H., Araki, Y. ve Konishi, S. (2008). Multivariate regression modeling for functional data. *J Data Sci*, 6:313–331.
- Matsui, H., Kawano, S. ve Konishi, S. (2009). Regularized functional regression modeling for functional response and predictors. *J Ind Math*, 1:17–25.
- Munck, L., Nørgaard, L., Engelsen, S. B., Bro, R. ve Andersson, C. A. (1998). Chemometrics in food science—a demonstration of the feasibility of a highly exploratory, inductive evaluation strategy of fundamental scientific significance. *Chemom Intell Lab Syst*, 44(1-2):31–60.
- Özçomak, M. S. ve Gündüz, M. (2014). Borsa İstanbul’da İşlem Gören Şirketlerin Kapanış İle İşlem Miktarları Arasındaki İlişkinin Fonksiyonel Kanonik Korelasyon İle Analizi. *Uluslararası İktisadi ve İdari İncelemeler Dergisi*. DOI:[10.18092/ijeas.35026](https://doi.org/10.18092/ijeas.35026).
- Preda, C. ve Saporta, G. (2005). PLS Regression on a stochastic process. *Comput Stat Data Anal*, 48(1):149–158.
- Preda C., Saporta, G. ve Leveder, C. (2007). PLS Classification for functional data. *Comput Stat*, 22(2): 223–235.
- Ramsay, J. O. (1982). When the data are functions. *Psychometrika*, Vol. 47, No. 4.
- Ramsay, J. O. ve Dalzell, C. J. (1991). Some tools for functional data analysis. *Journal of the Royal Statistical Society. Series B (Methodological)*, Vol. 53, No. 3, p. 539-572.
- Ramsay, J. O., Hooker, G. ve Graves, S. (2009). *Functional Data Analysis with R and MATLAB*. 1st edn, Springer Publishing Company, Incorporated.
- Ramsay, J. O. ve Silverman, B. W. (1997). *Functional Data Analysis*. Springer, New York.
- Ramsay, J. O. ve Silverman, B. W. (2002). *Applied Functional Data Analysis*. Springer, New York.
- Ramsay, J. O. ve Silverman, B. W. (2005). *Functional Data Analysis*. Springer, USA.
- Rao, C. R. (1958). Some statistical methods for comparison of growth curves. *Biometrics*, Vol. 14, No. 1, p. 1-17.
- Ratcliffe, S. J., Heller, G. Z. ve Leader, L. R. (2002). Functional data analysis with application to periodically stimulated foetal heart rate data. I: Functional logistic regression. *Statistics in medicine*, 21(8):1115–1127.
- Reiss, P. T., Goldsmith, J., Shang, H. L. ve Ogden, R. T. (2017). Methods for scalar-on-function regression. *Int Stat Rev*, 85(2):228–249.

- Reiss, P. T. ve Ogden, R. T. (2007). Functional principal component regression and functional partial least squares. *J Am Stat Assoc: Theory Methods*, 102(479):984–996.
- Saporta, G. (1981). Methodes exploratoires d'analyse de donn `ees temporelles. ' Cahiers du B.U.R.O. Universit'e Pierre et Marie Curie, pages 37–38.
- Shang, H. L. (2014). A survey of functional principal component analysis. *AStA Advances in Statistical Analysis*, 98(2):121–142.
- Shin, H. ve Lee, S. (2016). An RKHS approach to robust functional linear regression. *Stat Sinica*, 26:255–272.
- Silverman, B. W. (1996). Smoothed functional principal components analysis by choice of norm. *The Annals of Statistics*, 24(1):1–24.
- Smaga, L. ve Matsui, H. (2018). A note on variable selection in functional regression via random subspace method. *Stat Methods Appl*, 27(3):455–477.
- Smoktunowicz, A. ve Wrobel, I. (2012). Numerical aspects of computing the Moore-Penrose inverse of full column rank matrices. *BIT Comp Sci Numer Math*, 52(2):503–524.
- Sun, Y. ve Wang, Q. (2020). Function-on-function quadratic regression models. *Comput Stat Data Anal*, 142: 106814.
- Şafak, Ç. (2019). Fonksiyonel Veri Analizi ve Sağlık Alanına Özgü Uygulama. (Yüksek Lisans Tezi). Ankara Üniversitesi, Sağlık Bilimleri Enstitüsü, Ankara.
- Usset, J., Staicu, A. M. ve Maity, A. (2016). Interaction models for functional regression. *Comput Stat Data Anal*, 94:317–329.
- Wang, J. L., Chiou, J. M. ve Müller, H. G. (2015). Review of functional data analysis. *Annu. Rev. Statist*: 1-41.
- Wang, J. L., Chiou, J. M., ve Müller, H. G. (2016). Functional data analysis. *Ann Rev Stat Appl*, 3:257–295.
- Yao, F. (2007). Functional principal component analysis for longitudinal and survival data. *Stat Sinica*, 17(3): 965–983.
- Yu, D., Kong, L. ve Mizera, I. (2016). Partial functional linear quantile regression for neuroimaging data analysis. *Neurocomputing*, 195:74–87.
- Zhao, Y., Odgen, R. T. ve Reiss, P. T. (2012). Wavelet-based LASSO in functional linear regression. *J Comput Graphical Stat*, 21(3):600–617.
- Zhou, Z. (2021). Fast implementation of partial least squares for function-on-function regression. *J Multivariate Anal*, 185:104769.

EKLER

EK A: R Fonksiyonunda Yazılan Simülasyon Çalışması



EK A

```
library('fda')
library('expm')
library('pracma')
library('matrixStats')
library('fda.usc')
```

auxiliary_functions.R

```
## FPLS_projection
get_pls_mat = function(data, nbasis, rangeval){
  n = dim(data)[1]
  p = dim(data)[2]
  dimnames(data)=list(as.character(1:n), as.character(1:p))
  grid_points = seq(rangeval[1], rangeval[2], length.out = p)
  bs_basis = create.bspline.basis(rangeval, nbasis = nbasis)
  evalbase = eval.basis(grid_points, bs_basis)
  innp_mat = inprod(bs_basis, bs_basis)
  sqrt_innp_mat = sqrtm(innp_mat)
  fdobj = fdPar(bs_basis, int2Lfd(2), lambda=0)
  pcaobj = smooth.basisPar(grid_points, t(data), bs_basis, Lfdobj=NULL,
lambda=0)$fd
  data_mat = t(pcaobj$coefs) %*% sqrt_innp_mat
  return(list(data_mat = data_mat, sqrt_innp_mat = sqrt_innp_mat,
  eval_base = evalbase))
}
```

```

## Regressin_functions

reg_fun = function(Y, X, nbasis, npls, rangeval, method = c("simpls", "nipals",
"bidiag1", "bidiag2")){

  Y = as.matrix(Y)
  n = dim(Y)[1]

  dmat = get_pls_mat(data = X, nbasis = nbasis, rangeval = rangeval)
  matX = dmat$data_mat

  if(method == "nipals"){
    model_pls = nipalsN(X = matX, Y = Y, a = npls)
  }else if(method == "simpls"){
    model_pls = simpls(X = matX, Y = Y, a = npls)
  }else if(method == "bidiag1"){
    model_pls = bidiag1(X = matX, Y = Y, a = npls)
  }else if(method == "bidiag2"){
    model_pls = bidiag2(X = matX, Y = Y, a = npls)
  }
  coef_pls = model_pls$beta
  fits = matX %*% coef_pls
  bhat = dmat$eval_base %*% (solve(dmat$sqrt_innp_mat) %*% coef_pls)
  return(list(fits = fits, bhat = bhat))
}

```

```

## NIPALS_with_normalization

```

```

nipalsN = function(X,Y,a){
  W = NULL
  T = NULL
  P = NULL
  q = NULL
  for(i in 1:a){

```



```

w = t(X) %*% Y
w = w / norm(w, type = "2")
W = cbind(W, w)
t = X %*% w
t = t / norm(t, type = "2")
T = cbind(T, t)
P = cbind(P, t(X) %*% t)
X = X - t %*% t(P[,i])
q = c(q, t(Y) %*% t)
Y = Y - q[i] * t
}
beta = rowCumsums(t(t(mrdivide(W, triu(t(P) %*% W))) * q))[,a]
return(list(beta=beta))
}

```

```
## Bidiag2_with_reorthogonalization
```

```

bidiag2 = function(X, Y, a){
  B = matrix(0, nrow = a, ncol = 2)
  w = t(X) %*% Y
  w = w / norm(w, type = "2")
  W = w
  t = X %*% w
  rho = norm(t, type = "2")
  t = t / rho
  T = t
  B[1,1] = rho
  d = w / rho
  beta = d %*% (t(t) %*% Y)
  for(i in 2:a){

```

```

w = t(X) %*% t - rho * w
w = w - W %*% (t(W) %*% w)
theta = norm(w, type = "2")
w = w / theta
W = cbind(W, w)
t = X %*% w - theta * t
t = t - T %*% (t(T) %*% t)
rho = norm(t, type = "2")
t = t / rho
T = cbind(T, t)
B[(i-1),2] = theta
B[i,1] = rho
d = (w - theta * d) / rho
beta = cbind(beta, beta[(i-1)] + d %*% (t(t) %*% Y))
}
return(list(beta=beta[,a]))
}

```

```
## Bidiag1_with_reorthogonalization
```

```

bidiag1 = function(X, Y, a){
  beta = NULL
  B = matrix(0, nrow = a, ncol = 2)
  gamma = norm(Y, type = "2")
  t = Y / gamma
  T = t
  w = t(X) %*% t
  alpha = norm(w, type = "2")
  w = w / alpha
  W = w

```

```

d = w
reg = 0
rhubar = alpha
phibar = gamma
for(i in 1:a){
  t = X %*% w - alpha * t
  t = t - T %*% (t(T) %*% t)
  gamma = norm(t, type = "2")
  t = t / gamma
  T = cbind(T, t)
  w = t(X) %*% t - gamma * w
  w = w - W %*% (t(W) %*% w)
  alpha = norm(w, type = "2")
  w = w / alpha
  W = cbind(W, w)
  rho = norm(c(rhubar, gamma), type = "2")
  cos = rhubar / rho
  sin = gamma / rho
  theta = sin * alpha
  rhubar = -cos * alpha
  phi = cos * phibar
  phibar = sin * phibar
  G = matrix(c(cos, sin, sin, -cos), nrow = 2, ncol = 2, byrow = TRUE)
  T[, (i:(i+1))] = T[, (i:(i+1))] %*% G
  B[(i-1), 2] = theta
  B[i, 1] = rho
  reg = reg + d * (phi / rho)
  beta = cbind(beta, reg)
  d = w - d * (theta / rho)
}

```

```

}
return(list(beta=beta[,a]))
}

```

```
## SIMPLS
```

```

simpls <- function(X, Y, a) {
  Y = as.matrix(Y)
  n = nrow(X)
  k = ncol(X)
  m = ncol(Y)
  Ps = matrix(0, k, a)
  Cs = matrix(0, m, a)
  Rs = matrix(0, k, a)
  Ts = matrix(0, n, a)
  mx = apply(X, 2, mean)
  sdx = apply(X, 2, sd)
  X = apply(1:k, function(i) (X[,i]-mx[i]))
  my = apply(Y, 2, mean)
  sdy = apply(Y, 2, sd)
  Y = apply(1:m, function(i) (Y[,i]-my[i]))
  S = t(X)%*%Y
  Snew = S
  for (i in 1:a){
    rs = svd(Snew)$u[,1,drop=FALSE]
    rs = rs/norm(rs,type="F")
    ts = X %*% rs
    tsn = ts/norm(ts,type="F")
    ps = t(X) %*% tsn
    cs = t(Y) %*% tsn
  }
}

```

```

Rs[,i] = rs
Ts[,i] = ts
Ps[,i] = ps
Cs[,i] = cs

Snew = Snew-Ps[,1:i] %*% solve(t(Ps[,1:i]) %*% Ps[,1:i]) %*% t(Ps[,1:i]) %*%
Snew
}
beta = Rs %*% solve(t(Ps) %*% Rs) %*% t(Cs)
return(list(beta = beta))
}

```

```

### data_generation.R

```

```

data_generation = function(n, j){
s = seq(0, 1, length.out = j)
ksi = list()
for(ik in 1:5){
ksi[[ik]] = rnorm(n, 4, sd = (3*ik^(-1/2)))
}
phi = list()
for(ik in 1:5){
phi[[ik]] = 4*sin(ik * pi * s^2 ) + cos(4*ik * pi * s^2)
}
fX = Reduce("+", lapply(1:5, function(k){ksi[[k]] %*% t(phi[[k])}))
Beta_fun = 2*cos(pi * s)
fX = fdata(fX, argvals = s)
Beta_fun = fdata(Beta_fun, argvals = s)
err = rnorm(n, mean=0, sd=1)
Y = inprod.fdata(fX, Beta_fun)
Y = Y + err

```

```

return(list(Y = Y, X = fX$data, Beta = c(Beta_fun$data)))
}

```

simulations.R

```

# Sample size
n = 1000
# Length of function support
j = 1000
# Number of basis functions
nbasis = 25
# Number of PLS components
npls = 5
set.seed(12345)
# Generate the data
sim_dat = data_generation(n = n, j = j)
X = sim_dat$X
Y = sim_dat$Y
beta_true = sim_dat$Beta
# Model fits
simpls_model = reg_fun(Y = Y, X = X, nbasis = nbasis, npls = npls, rangeval =
c(0,1), method = "simpls")
nipals_model = reg_fun(Y = Y, X = X, nbasis = nbasis, npls = npls, rangeval =
c(0,1), method = "nipals")
bidiag1_model = reg_fun(Y = Y, X = X, nbasis = nbasis, npls = npls, rangeval =
c(0,1), method = "bidiag1")
bidiag2_model = reg_fun(Y = Y, X = X, nbasis = nbasis, npls = npls, rangeval =
c(0,1), method = "bidiag2")
# MISE
mean((beta_true - simpls_model$bhat)^2) # 0.708649
mean((beta_true - nipals_model$bhat)^2) # 0.7086455
mean((beta_true - bidiag1_model$bhat)^2) # 0.7086455
mean((beta_true - bidiag2_model$bhat)^2) # 0.7086455
# MSE
mean((Y - simpls_model$fits)^2) # 1.064305
mean((Y - nipals_model$fits)^2) # 1.007794

```

```
mean((Y - bidiag1_model$fits)^2) # 1.007794  
mean((Y - bidiag2_model$fits)^2) # 1.007794
```

