

**BAĞIMSIZ BİLEŞENLER ANALİZİ İLE ÇOKLU
BAĞLANTI SORUNUNA BİR YAKLAŞIM**

**AN APPROACH TO MULTICOLLINEARITY PROBLEM
WITH INDEPENDENT COMPONENTS ANALYSIS**

NURBANU BURSA

PROF. DR. HÜSEYİN TATLIDİL

Tez Danışmanı

Hacettepe Üniversitesi

Lisansüstü Eğitim-Öğretim ve Sınav Yönetmeliğinin

İstatistik Anabilim Dalı için Öngördüğü

DOKTORA TEZİ olarak hazırlanmıştır.

NURBANU BURSA'nın hazırladığı "Bağımsız Bileşenler Analizi ile Çoklu Bağlantı Sorununa Bir Yaklaşım" adlı bu çalışma aşağıdaki jüri tarafından İSTATİSTİK ANABİLİM DALI'nda DOKTORA TEZİ olarak kabul edilmiştir.

Prof. Dr. Mehmet Akif BAKIR

Başkan



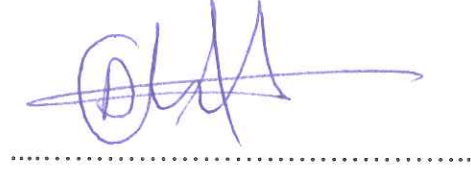
Prof. Dr. Hüseyin TATLIDİL

Danışman



Prof. Dr. Duru KARASOY

Üye



Doç. Dr. Gamze ÖZEL KADILAR

Üye



Dr. Öğr. Üyesi Elif Özge ÖZDAMAR

Üye



Bu tez Hacettepe Üniversitesi Fen Bilimleri Enstitüsü tarafından DOKTORA TEZİ olarak / /..... tarihinde onaylanmıştır.

Prof. Dr. Menemşe GÜMÜŞDERELİOĞLU

Fen Bilimleri Enstitüsü Müdürü

ETİK

Hacettepe Üniversitesi Fen Bilimleri Enstitüsü, tez yazım kurallarına 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 ahlak kurallarına uygun olarak sunduğumu,
- başkalarının eserlerinden yararlanması durumunda ilgili eserlere bilimsel normlara uygun olarak atıfta bulunduğumu,
- atıfta bulunduğum eserlerin tümünü kaynak olarak gösterdiğimi,
- kullanılan verilerde herhangi bir değişiklik yapmadığımı,
- ve bu tezin herhangi bir bölümünü bu üniversite veya başka bir üniversitede başka bir tez çalışması olarak sunmadığımı

beyan ederim.

21 / 06 / 2019



NURBANU BURSA

YAYINLANMA FİKRİ MÜLKİYET HAKLARI BEYANI

Enstitü tarafından onaylanan lisansüstü tezimin/raporumun tamamını veya herhangi bir kısmını, basılı (kağıt) ve elektronik formatta arşivleme ve aşağıda verilen koşullarla kullanıma açma iznini Hacettepe Üniversitesine verdiğimi bildiririm. Bu izinle Üniversiteye verilen kullanım hakları dışındaki tüm fikri mülkiyet haklarım bende kalacak, tezimin tamamının ya da bir bölümünün gelecekteki çalışmalarda (makale, kitap, lisans ve patent vb.) kullanım hakları bana ait olacaktır.

Tezin kendi orijinal çalışmam olduğunu, başkalarının haklarını ihlal etmediğimi ve tezimin tek yetkili sahibi olduğumu beyan ve taahhüt ederim. Tezimde yer alan telif hakkı bulunan ve sahiplerinden yazılı izin alınarak kullanması zorunlu metinlerin yazılı izin alarak kullandığımı ve istenildiğinde suretlerini Üniversiteye teslim etmeyi taahhüt ederim.

Yükseköğretim Kurulu tarafından yayımlanan "*Lisansüstü Tezlerin Elektronik Ortamda Toplanması, Düzenlenmesi ve Erişime Açılmasına İlişkin Yönerge*" kapsamında tezim aşağıda belirtilen koşullar haricince YÖK Ulusal Tez Merkezi / H. Ü. Kütüphaneleri Açık Erişim Sisteminde erişime açılır.

- Enstitü / Fakülte yönetim kurulu kararı ile tezimin erişime açılması mezuniyet tarihimden itibaren 2 yıl ertelenmiştir.
- Enstitü / Fakülte yönetim kurulu gerekçeli kararı ile tezimin erişime açılması mezuniyet tarihimden itibaren ay ertelenmiştir.
- Tezim ile ilgili gizlilik kararı verilmiştir.

21 / 06 / 2019



NURBANU BURSA

ÖZET

BAĞIMSIZ BİLEŞENLER ANALİZİ İLE ÇOKLU BAĞLANTI SORUNUNA BİR YAKLAŞIM

Nurbanu BURSA

Doktora, İstatistik Bölümü

Tez Danışmanı: Prof. Dr. Hüseyin TATLIDİL

Haziran 2019, 123 sayfa

Tez çalışmasında ilk olarak, çok değişkenli istatistiksel bir yöntem olmasına rağmen istatistikte pek fazla bilinmeyen daha çok mühendislik alanında kullanılan Bağımsız Bileşenler Analizi konusu ayrıntılı bir şekilde ele alınmış ve konuyla ilgili kısıtlı Türkçe literatüre katkıda bulunulmuştur.

İkinci olarak ise çoklu doğrusal regresyon analizinde sık karşılaşılan sorunlardan biri olan ve varsayım bozulmalarına neden olarak regresyon modeli üzerinde olumsuz etkilere sebebiyet veren çoklu bağlantı sorununa, Bağımsız Bileşenler Analizi ile yeni bir çözüm önerisi getirilmiştir. Böylece analizin mevcut kullanımlarının yanı sıra farklı amaçlar için de bir araç olarak kullanılabilmesi gösterilmiştir. Bu bağlamda, çoklu bağlantı sorununun çözümünde kullanılan Temel Bileşenler Regresyonu ve Kısmi En Küçük Kareler

Regresyonu gibi yanlı regresyon yöntemlerinin işleyişine benzeyen, Bağımsız Bileşenler Analizi'ne dayalı yeni bir yöntem geliştirilmiştir. Yöntemde, literatürden farklı olarak çoklu doğrusal regresyon analizine dahil edilecek bağımsız bileşenlerin seçiminde, entropi alanındaki bir kavram olan ortak bilgi miktarının kullanımı önerilmiştir. Geliştirilen yöntemin performansı, yapay ve gerçek veriler üzerinde ayrı ayrı değerlendirilmiştir.

Sonuç olarak, tez kapsamında önerilen Bağımsız Bileşenler Analizi'ne dayalı yöntem ile çoklu bağlantı sorununa çözüm sunulabileceği, regresyon analizindeki açıklanan (bağımlı) değişken için diğer yanlı regresyon yöntemlerine göre daha düşük hata miktarı ile tahmin ve öngörü yapılabileceği ve ayrıca regresyon katsayılarının daha düşük standart hatalar ile tahmin edilebileceği belirlenmiştir.

Anahtar Kelimeler: Bağımsız Bileşenler Analizi, Temel Bileşenler Analizi, Regresyon Analizi, Çoklu Bağlantı Sorunu, Temel Bileşenler Regresyonu, Bağımsız Bileşenler Regresyonu, Shannon Entropisi, Ortak Bilgi Miktarı.

ABSTRACT

AN APPROACH TO MULTICOLLINEARITY PROBLEM WITH INDEPENDENT COMPONENTS ANALYSIS

Nurbanu BURSA

Doctor of Philosophy, Department of Statistics

Supervisor: Prof. Dr. Huseyin Tatlidil

June 2019, 123 pages

In the first part of the thesis, although it is a multivariate statistical method, Independent Components Analysis which is mostly used in the engineering field, which is not known much in statistics, has been discussed in detail and has been contributed to the limited Turkish literature.

Secondly, a new solution was proposed by Independent Components Analysis to the multicollinearity problem, which is one of the most common problems in multiple linear regression analysis, which caused the assumption distortions and caused negative effects on the regression model. Thus, it has been shown that the analysis can be used as a tool for different purposes in addition to its current uses. In this context, a new method has been developed based on Independent Components Analysis, which is similar to the

operation of biased regression methods such as Principal Components Regression and Partial Least Squares Regression used to solve multicollinearity problem. The novel contribution of this method to the literature is that the use of mutual information, a concept in the field of entropy, is proposed in the selection of independent components to be included in multiple linear regression analysis. The performance of the developed method was evaluated separately on artificial and real datasets.

As a result, it was determined that the proposed method based on Independent Components Analysis can be used to solve the multicollinearity problem, estimation and prediction can be made with lower error amount than other biased regression methods for the explained (dependent) variable in the regression analysis, and regression coefficients can be estimated with lower standard errors.

Keywords: Independent Components Analysis, Principal Component Analysis, Regression Analysis, Multicollinearity Problem, Principal Component Regression, Independent Component Regression, Shannon Entropy, Mutual Information.

TEŐEKKÜR

Her zaman sabrı ve içten tavırlarıyla yanımda olan, manevi desteęini hiçbir zaman esirgemeyen, bilgi ve tecrübelerini paylaşarak bana yol gösteren ve beni bir meslektaşı olarak yetiřtiren deęerli danıřmanım Prof. Dr. Hüseyin TATLIDİL'e,

Tez izleme komiteleri boyunca katkı ve eleřtirileriyle beni her defasında düşünmeye ve daha iyi bir çalıřma ortaya koymaya sevk eden deęerli hocalarım Prof. Dr. Mehmet Akif BAKIR ve Prof. Dr. Duru KARASOY'a,

Yardımlarıyla her zaman yanımda olan kadim dostum Gizem ULUSOY'a,

Tüm yoğunluęuna raęmen sorularımı her defasında yanıtız bırakmayan Doç. Dr. Semra TÜRKAN'a,

Farkında olmadan sürece en büyük desteęi veren Doç. Dr. Sercan OKUTUCU'ya,

Bugün sahip olduęum her şeyimi borçlu olduęum canım annem Mürüvvet BURSA, sevgili ablam Ayřegül ÜLKER ve sevgili halam Nurdan BURSA'ya,

Lisansüstü eğitimime bařladıęım günden bu yana bursiyeri olmaktan gurur duyduęum, daima bilime katkıda bulunan TÜBİTAK kurumuna,

en içten teőekkürlerimi sunarım.

İÇİNDEKİLER

ÖZET	i
ABSTRACT	iii
TEŞEKKÜR	v
İÇİNDEKİLER	vi
ÇİZELGELER DİZİNİ	viii
ŞEKİLLER DİZİNİ	x
SİMGELER VE KISALTMALAR	xii
1. GİRİŞ	1
2. BAĞIMSIZ BİLEŞENLER ANALİZİ İLE İLGİLİ GENEL BİLGİLER.....	3
2.1. Bağımsız Bileşenler Analizinin Tarihçesi ve Uygulama Alanları.....	3
2.2. Bağımsız Bileşenler Analizine Öncülük Eden İki Problem	5
2.2.1. Çok Değişkenli Verinin Temsili Problemi	5
2.2.2. Kör Kaynak Ayrıştırma Problemi	7
2.3. Bağımsız Bileşenler Analizi için Gerekli Bazı Matematiksel ve İstatistiksel Kavramlar	8
2.3.1. Çok Değişkenli İstatistiksel Kavramlar	8
2.3.2. Optimizasyon Teorisi ile İlgili Temel Kavramlar	24
2.3.3. Bilgi Kuramı ile İlgili Temel Kavramlar	27
3. BAĞIMSIZ BİLEŞENLER ANALİZİ	34
3.1. Bağımsız Bileşenler Analizi Modeli	35
3.1.1. Bağımsız Bileşenler Analizinin Varsayımları ve Kısıtları	36
3.1.2. Bağımsız Bileşenler Analizinin Belirsizlikleri	38
3.2. Bağımsız Bileşenlerin Elde Edilmesinde Kullanılan Yaklaşımlar ve Algoritmalar	38
3.2.1. Bağımsız Bileşenler Analizi Öncesinde Kullanılacak Verinin Hazırlanması	39

3.2.2. Bağımsız Bileşenlerin Tahmini	40
3.2.3. FastICA Algoritması	45
3.2.4. Bağımsız Bileşenler Analizi ile Temel Bileşenler Analizi'nin Karşılaştırılması ve Bir Uygulama	48
4. BAĞIMSIZ BİLEŞENLER ANALİZİ İLE İSTATİSTİKSEL BİR SORUNUN ÇÖZÜMÜ	61
4.1. Çoklu Doğrusal Regresyon Analizi	61
4.2. Çoklu Bağlantı Sorunu, Etkileri ve Çözüm Yolları.....	63
4.3. Bağımsız Bileşenler Analizi ile Çoklu Bağlantı Sorununun Çözümüne Dair Literatür Taraması.....	68
4.4. Bağımsız Bileşenler Analizi ile Çoklu Bağlantı Sorununun Çözümü.....	72
5. UYGULAMA	76
5.1. Yapay Veri Uygulamaları	76
5.1.1. Yapay Veri Uygulaması-1	76
5.1.2. Yapay Veri Uygulaması-2.....	91
5.2. Gerçek Veri Uygulamaları	99
5.2.1. Gerçek Veri Uygulaması-1	100
5.2.2. Gerçek Veri Uygulaması-2.....	104
5.2.3. Gerçek Veri Uygulaması-3.....	107
5.2.4. Gerçek Veri Uygulaması-4.....	110
6. SONUÇLAR ve TARTIŞMA	113
7. KAYNAKLAR	116
EKLER	124
EK 1 - Tezden Üretilmiş Bildiriler.....	124
EK 2 - Tez Çalışması Orjinallik Raporu.....	125
ÖZGEÇMİŞ	127

ÇİZELGELER DİZİNİ

Çizelge 3.1. Bağımsız Bileşenler ile Orijinal Bileşenler Arasındaki Pearson Korelasyon Katsayıları	60
Çizelge 3.2. Temel Bileşenler ile Orijinal Bileşenler Arasındaki Pearson Korelasyon Katsayıları	60
Çizelge 5.1. Yapay Veri Uygulaması-1 İçin Hoeffding Test Sonuçları.....	78
Çizelge 5.2. Bir Bağımsız Bileşenli Eğitim Kümesi Hata Kareler Ortalaması Sonuçları	83
Çizelge 5.3. Bir Bağımsız Bileşenli Test Kümesi Hata Kareler Ortalaması Sonuçları..	84
Çizelge 5.4. Diğer Bağımsız Bileşenlerin (Birli) Eğitim Kümesi Hata Kareler Ortalaması Sonuçları	85
Çizelge 5.5. Diğer Bağımsız Bileşenlerin (Birli) Test Kümesi Hata Kareler Ortalaması Sonuçları	86
Çizelge 5.6. İki Bağımsız Bileşenli Eğitim Kümesi Hata Kareler Ortalaması Sonuçları	87
Çizelge 5.7. İki Bağımsız Bileşenli Test Kümesi Hata Kareler Ortalaması Sonuçları ..	88
Çizelge 5.8. Diğer Bağımsız Bileşenlerin (İkili) Eğitim Kümesi Hata Kareler Ortalaması Sonuçları	89
Çizelge 5.9. Diğer Bağımsız Bileşenlerin (İkili) Test Kümesi Hata Kareler Ortalaması Sonuçları	90
Çizelge 5.10. Yapay Veri Uygulaması-2 İçin Hoeffding Test Sonuçları.....	91
Çizelge 5.11. İki Bağımsız Bileşenli Eğitim Kümesi Hata Kareler Ortalaması Sonuçları	96
Çizelge 5.12. İki Bağımsız Bileşenli Test Kümesi Hata Kareler Ortalaması Sonuçları	97
Çizelge 5.13. Üç Bağımsız Bileşenli Eğitim Kümesi Hata Kareler Ortalaması Sonuçları	98
Çizelge 5.14. Üç Bağımsız Bileşenli Test Kümesi Hata Kareler Ortalaması Sonuçları	99
Çizelge 5.15. Gerçek Veri Uygulaması-1’de Değişkenler Arası Korelasyon Matrisi ..	101
Çizelge 5.16. Gerçek Veri Uygulaması-1’de Eğitim ve Test Kümeleri İçin VIF ve Koşul İndeksi Değerleri	102
Çizelge 5.17. Bağımsız Bileşenler ile Açıklanan Değişken Arasındaki Ortak Bilgi Miktarı ve Korelasyon Katsayıları.....	103
Çizelge 5.18. Güçlü Çoklu Bağlantı Durumu için Hata Kareler Ortalaması Sonuçları	103

Çizelge 5.19. Gerçek Veri Uygulaması-2’de Değişkenler Arası Korelasyon Matrisi ..	105
Çizelge 5.20. Gerçek Veri Uygulaması-2’de Eğitim ve Test Kümeleri İçin VIF ve Koşul İndeksi Değerleri	106
Çizelge 5.21. Bağımsız Bileşenler ile Açıklanan Değişken Arasındaki Ortak Bilgi Miktarı ve Korelasyon Katsayıları	107
Çizelge 5.22. Zayıf Çoklu Bağlantı Durumu İçin Hata Kareler Ortalaması Sonuçları .	107
Çizelge 5.23. Gerçek Veri Uygulaması-3’te Değişkenler Arası Korelasyon Matrisi ...	108
Çizelge 5.24. Gerçek Veri Uygulaması-3’te Eğitim ve Test Kümeleri İçin VIF ve Koşul İndeksi Değerleri	109
Çizelge 5.25. Bağımsız Bileşenler ile Açıklanan Değişken Arasındaki Ortak Bilgi Miktarı ve Korelasyon Katsayıları	109
Çizelge 5.26. Orta Düzey Çoklu Bağlantı Durumu İçin Hata Kareler Ortalaması Sonuçları	110
Çizelge 5.27. Gerçek Veri Uygulaması-4’te Değişkenler Arası Korelasyon Matrisi ...	111
Çizelge 5.28. Gerçek Veri Uygulaması-4’te VIF ve Koşul İndeksi Değerleri	111
Çizelge 5.29. Bağımsız Bileşenler ile Açıklanan Değişken Arasındaki Ortak Bilgi Miktarı ve Korelasyon Katsayıları	112
Çizelge 5.30. Çoklu Bağlantı Durumu İçin Jackknife Standart Hatasının Tahmin Sonuçları	112

ŞEKİLLER DİZİNİ

Şekil 2.1 Ortak Bilginin Şekilsel Gösterimi	33
Şekil 3.1 Kokteyl Parti Problemi.....	34
Şekil 3.2. Temel Bileşenler Analizi ve Bağımsız Bileşenler Analizi ile Elde Edilen Bileşenler İçin Bir Örnek	49
Şekil 3.3. S Matrisi İçin Üretilen s_1 Serisi.....	51
Şekil 3.4. S Matrisi İçin Üretilen s_2 Serisi	51
Şekil 3.5. S Matrisi İçin Üretilen s_3 Serisi	51
Şekil 3.6. S Matrisi İçin Üretilen s_4 Serisi	52
Şekil 3.7. S Matrisi İçin Üretilen s_5 Serisi	52
Şekil 3.8. S Matrisi İçin Üretilen s_6 Serisi	52
Şekil 3.9. Oluşturulan X Matrisindeki Gözlenen x_1 Serisi	53
Şekil 3.10. Oluşturulan X Matrisindeki Gözlenen x_2 Serisi.....	53
Şekil 3.11. Oluşturulan X Matrisindeki Gözlenen x_3 Serisi	53
Şekil 3.12. Oluşturulan X Matrisindeki Gözlenen x_4 Serisi.....	54
Şekil 3.13. Oluşturulan X Matrisindeki Gözlenen x_5 Serisi.....	54
Şekil 3.14. Oluşturulan X Matrisindeki Gözlenen x_6 Serisi.....	54
Şekil 3.15. Bağımsız Bileşenler Analizi Sonucunda Tahmin Edilen Bağımsız Bileşen-1	55
Şekil 3.16. Bağımsız Bileşenler Analizi Sonucunda Tahmin Edilen Bağımsız Bileşen-2	56
Şekil 3.17. Bağımsız Bileşenler Analizi Sonucunda Tahmin Edilen Bağımsız Bileşen-3	56
Şekil 3.18. Bağımsız Bileşenler Analizi Sonucunda Tahmin Edilen Bağımsız Bileşen-4	56
Şekil 3.19. Bağımsız Bileşenler Analizi Sonucunda Tahmin Edilen Bağımsız Bileşen-5	57
Şekil 3.20. Bağımsız Bileşenler Analizi Sonucunda Tahmin Edilen Bağımsız Bileşen-6	57
Şekil 3.21. Temel Bileşenler Analizi Sonucunda Tahmin Edilen Temel Bileşen-1	57

Şekil 3.22. Temel Bileşenler Analizi Sonucunda Tahmin Edilen Temel Bileşen-2	58
Şekil 3.23. Temel Bileşenler Analizi Sonucunda Tahmin Edilen Temel Bileşen-3	58
Şekil 3.24. Temel Bileşenler Analizi Sonucunda Tahmin Edilen Temel Bileşen-4	58
Şekil 3.25. Temel Bileşenler Analizi Sonucunda Tahmin Edilen Temel Bileşen-5	59
Şekil 3.26. Temel Bileşenler Analizi Sonucunda Tahmin Edilen Temel Bileşen-6	59
Şekil 5.1. $n = 500$ İçin s_1 Bileşeni	79
Şekil 5.2. $n = 500$ İçin s_2 Bileşeni	79
Şekil 5.3. $n = 500$ İçin y Değişkeni	80
Şekil 5.4. $n = 500$ ve $\sigma = 0.05$ İçin x_1 Değişkeni	80
Şekil 5.5. $n = 500$ ve $\sigma = 0.05$ İçin x_2 Değişkeni	80
Şekil 5.6. $n = 500$ ve $\sigma = 0.05$ İçin x_3 Değişkeni	81
Şekil 5.7. $n = 500$ ve $\sigma = 0.05$ İçin x_4 Değişkeni	81
Şekil 5.8. $n = 500$ ve $\sigma = 0.05$ İçin x_5 Değişkeni	81
Şekil 5.9. $n = 500$ ve $\sigma = 0.05$ İçin x_6 Değişkeni	82
Şekil 5.10. $n = 1000$ İçin s_1 Bileşeni	92
Şekil 5.11. $n = 1000$ İçin s_2 Bileşeni	92
Şekil 5.12. $n = 1000$ İçin s_3 Bileşeni	93
Şekil 5.13. $n = 1000$ İçin y Değişkeni	93
Şekil 5.14. $n = 1000$ ve $\sigma = 0.9$ İçin x_1 Değişkeni	93
Şekil 5.15. $n = 1000$ ve $\sigma = 0.9$ İçin x_2 Değişkeni	94
Şekil 5.16. $n = 1000$ ve $\sigma = 0.9$ İçin x_3 Değişkeni	94
Şekil 5.17. $n = 1000$ ve $\sigma = 0.9$ İçin x_4 Değişkeni	94
Şekil 5.18. $n = 1000$ ve $\sigma = 0.9$ İçin x_5 Değişkeni	95

SİMGELER VE KISALTMALAR

Simgeler

A	Karışım Matrisi
β	Regresyon Katsayısı
D	Köşegen Matris
$E(.)$	Beklenen Değer
Σ	Kovaryans Matrisi
F	Birikimli Dağılım Fonksiyonu
λ	Özdeğer
$H(.)$	Entropi Fonksiyonu
I	Ortak Bilgi Miktarı
I	Birim Matris
J	Jakobiyen Matris
μ	Ortalama Vektörü
R^2	Belirtme Katsayısı
R	Korelasyon (İlişki) Matrisi
σ	Standart Sapma
S	Bağımsız Bileşenler Matrisi
U	Bağımsız Bileşenler Matrisinin Tahmini
W	Ayrıştırma Matrisi (\mathbf{A}^{-1} matrisinin tahmini)
x	Raslantı Değişkeni
x	Raslantı Vektörü
X	Veri Matrisi
Z	Beyaz Veri Matrisi

Kısaltmalar

bas	Basıklık Katsayısı
BBA	Bağımsız Bileşenler Analizi
BBA-OBM	Ortak Bilgi Miktarı ile Bağımsız Bileşenlerin Seçildiği Regresyon Analizi
BBA-K	Korelasyon Katsayısı ile Bağımsız Bileşenlerin Seçildiği Regresyon Analizi
BBR	Bağımsız Bileşenler Regresyonu
cos	Kosinüs Fonksiyonu
ÇDR	Çoklu Doğrusal Regresyon
det	Determinant
EKK	En Küçük Kareler
FastICA	Hızlı Sabit-Nokta Algoritması
FOBI	Dördüncü Dereceden Kör Belirleme
HKO	Hata Kareler Ortalaması
JADE	Özdeğer Matrislerinin Ortak Yaklaşım ile Köşegenleştirilmesi
KEKR	Kısmi En Küçük Kareler Regresyonu
Kİ	Koşul İndeksi
köş	köşegen
rem	Remainder (Kalıntı) Fonksiyonu
RR	Ridge Regresyonu
sign	İşaret Fonksiyonu
sin	Sinüs Fonksiyonu
SOBI	İkinci Dereceden Kör Belirleme
TBA	Temel Bileşenler Analizi
TBR	Temel Bileşenler Regresyonu
VIF	Varyans Şişme Değerleri

1. GİRİŞ

Bağımsız Bileşenler Analizi (BBA); çok değişkenli veriyi, altında yatan esas yapıyı daha görünür yapacak bir şekilde dönüştüren istatistiksel bir yöntemdir. BBA ile çok değişkenli verinin, istatistiksel olarak birbirinden bağımsız bileşenlerin doğrusal ya da doğrusal olmayan birleşimleri şeklinde ifade edilmesi sağlanır (Hyvärinen, Karhunen ve Oja, 2001).

Gerçek yaşamdaki problemleri etkin bir şekilde çözmesi sebebiyle BBA, son yıllarda sinyal işleme ve makine öğrenmesindeki bir dizi soruna uygulanan standart bir veri analizi tekniği haline gelmiştir (Shlens, 2014). BBA, bu alanlarda yaygın olarak kullanılmasına ve esasında çok değişkenli istatistiksel bir yöntem olmasına rağmen istatistik alanında pek fazla bilinmemesi nedeniyle, Temel Bileşenler Analizi (TBA) kadar yoğun olarak kullanılmamaktadır. Özellikle istatistik alanındaki Türkçe literatür incelendiğinde, konuyla ilgili Özdamar'ın (2009) doktora tez çalışması haricinde bir çalışma bulunmadığı görülmektedir. Bu durumun oluşmasına;

- Mevcut çalışmalarda BBA'nın arkasında yatan istatistiksel yapının tam olarak ön plana çıkarılmamasının,
- TBA gibi benzer yöntemlerle olan farkının yeterince iyi açıklanmamasının ve
- BBA uygulamalarının çoğu zaman mühendislik alanına özgü verilerle sınırlı kalmasının

neden olduğu düşünülmektedir. Bu sebeple doktora tez çalışması, yukarıda bahsi geçen eksikliklerin giderilmesi ve kısıtlı Türkçe literatüre katkıda bulunulması amacıyla BBA üzerine kurulmuştur. Bu bağlamda çalışmada, BBA'nın istatistik alanındaki bilinirliğinin artırılması ve bunun için de istatistikte sık karşılaşılan mevcut bir soruna BBA'nın araç olarak kullanıldığı yeni bir çözüm önerisinin sunulması hedeflenmiştir. Böylece, BBA'nın sinyal işleme ve makine öğrenmesi alanlarından başka, istatistik alanı içerisinde de farklı kullanım amaçlarının olabileceği gösterilerek yeni çalışmalara öncülük edileceği düşünülmektedir.

Çalışmanın izleyen ikinci bölümünde öncelikle, BBA modelinin gelişim süreci ve uygulama alanları hakkında genel bilgiler aktarılmıştır. Sonrasında, BBA'nın daha iyi anlaşılabilmesi için gerekli bazı matematiksel ve istatistiksel kavramlara yer verilmiştir. Bölüm sonunda ise BBA'nın niçin ortaya çıktığının mantıksal çerçevesinin daha iyi kavranabilmesi amacıyla BBA'nın çözüm sunduğu iki farklı problem incelenmiştir.

Üçüncü bölümde; temel BBA modeli olarak bilinen modelden, modelin özelliklerinden, modelin çözümü için kullanılan algoritmalarından bahsedilmiştir. Ayrıca bu bölümde, çok değişkenli istatistiksel yöntemlerden biri olan ve işleyiş açısından BBA'ya benzeyen TBA'nın, BBA ile karşılaştırılmasına ve örnek bir uygulamaya da yer verilmiştir.

Dördüncü bölümde, çoklu doğrusal regresyon analizinde en çok karşılaşılan sorunlardan biri olan çoklu bağlantı sorunundan ve bu sorunun çözümünde kullanılan mevcut yöntemlerden bahsedildikten sonra, BBA'nın çoklu bağlantı sorununun çözümü için nasıl kullanılabileceğine ve konu ile ilgili literatür taramasına yer verilmiştir. Bu bağlamda, çoklu bağlantı sorununun çözümünde kullanılan yanlı regresyon yöntemlerinden Temel Bileşenler Regresyonu (TBR) ve Kısmi En Küçük Kareler Regresyonu'na (KEKR) işleyiş olarak benzeyen BBA'ya dayalı yeni bir yöntem önerilmiştir. Önerilen yöntemde, literatürden farklı olarak bağımsız bileşenlerin seçimi için entropi alanındaki bir kavram olan ortak bilgi miktarı kullanılmıştır.

Beşinci bölümde, çoklu bağlantı sorununa sahip yapay ve gerçek veriler üzerinde uygulamalar gerçekleştirilmiştir. Uygulamalarda çoklu bağlantı sorununu çözmek için kullanılan yöntemler ile BBA'ya dayalı önerilen yeni yöntemin performansları karşılaştırılmıştır.

Altıncı ve son bölümde ise uygulamalar kapsamında elde edilen sonuçlar hakkında genel değerlendirmeler yapılmış ve tezin katkılarına değinilmiştir.

2. BAĞIMSIZ BİLEŞENLER ANALİZİ İLE İLGİLİ GENEL BİLGİLER

Bu bölümün ilk kısmında, BBA'nın tarihçesine ve gün geçtikçe gelişen uygulama alanlarına ait örneklerle yer verilmiştir. BBA, çok değişkenli istatistiksel bir yöntem olduğu için sonraki kısımda, konu içerisinde ihtiyaç duyulabilecek istatistiksel bağımsızlık, yüksek dereceden istatistikler gibi bazı çok değişkenli istatistiksel kavramlar hatırlatılmıştır. Son kısımda ise BBA'nın çıkış noktasını oluşturan iki farklı problem ele alınmış ve sonraki bölümde anlatılacak olan BBA modeli için bir zemin hazırlanmıştır.

2.1. Bağımsız Bileşenler Analizinin Tarihçesi ve Uygulama Alanları

Henüz kullanılmaya başlandığı yıllarda BBA ismiyle anılmayan bu yöntem, ilk kez Héroult, Jutten ve Ans (1985) tarafından, kasılan kasların hareketlerinin kodlanması için kullanılmıştır. Bu çalışmada sinir sisteminin, kasların kasılmasıyla danişmansız (gözetimsiz) bir öğrenme gerçekleştirerek kasa gelen uyarının açısını ve hızını belirlediği gösterilmiştir. Ancak; elde edilen sonuçların Fransızca yayımlanması nedeniyle konu, bu dönemde uluslararası literatürde yeterince yaygınlaşmamış ve konunun etkisi sadece Fransız araştırmacılarla sınırlı kalmıştır. Sinyal işleme ve yapay sinir ağları üzerinde çalışan Fransız araştırmacılar tarafından sonraki yıllarda yayımlanan çalışmalarda bu teknik, Héroult-Jutten modeli adıyla yer bulmuştur. Jutten ve Héroult'un (1991a, 1991b, 1991c) ilerleyen yıllarda İngilizce olarak yayımladıkları çalışmaları ile konu, yavaş yavaş tüm dünya tarafından bilinir hale gelmiştir.

BBA adı ilk kez Comon (1994) tarafından yayımlanan bir çalışmada kullanılmıştır. Bu çalışmada ayrıca, BBA'nın ilk kez kapsamlı bir matematiksel formülasyonuna da yer verilmiştir. 1990'lı yılların ortasından itibaren BBA konusuna olan ilgi artmış; farklı BBA modelleri, bu BBA modellerinin çözümünde kullanılacak yeni algoritmalar geliştirilmiş ve BBA'nın kör kaynak ayırtırmadan başka farklı amaçlar için de kullanılabilceği gösterilmeye başlanmıştır. 2000'li yılların başından itibaren ise konuyla ilgili kitaplar yayımlanmış ve "Independent Component Analysis and Signal Separation" adı ile konuya özel kongreler düzenlenmeye başlanmıştır.

BBA günümüzde,

- Medikal görüntüleme; beynin yaydığı sinyallerin kaynaklarının bulunması ve ultrasonografi sinyallerinin ayrıştırılması,

(Örnek olarak bkz. Artoni, Delorme ve Makeig, 2019; Tierney, Wilkes ve Byram, 2019; Baker ve ark., 2019)

- Jeolojide; sismik dalgaların incelenmesi ve jeolojik haritalama,

(Örnek olarak bkz. Albert ve Bowman, 2018; de Lauro, Petrosino ve Falanga, 2018; Cohen-Waeber ve ark., 2018)

- Hata tespitinde; kalite kontrol süreçlerinde karşılaşılan hataların belirlenmesi,

(Örnek olarak bkz. Garcia-Bracamonte ve ark., 2018; Yu ve ark., 2018; Li ve Yan, 2018)

- Görüntü işleme; görüntülerdeki öznelik çıkarımı ile görüntülerin netleştirilmesi ve karakterize edilmesi,

(Örnek olarak bkz. Lahaw, Essaidani ve Seddik, 2018; Wang ve Guo, 2019; Koush ve ark., 2019)

- Telekomünikasyonda; ses sinyallerinin ya da radyo dalgalarının ayrıştırılması,

(Örnek olarak bkz. Jiang ve ark., 2017; Wang ve ark., 2018; Aveta ve ark., 2018)

- Ekonometride; finansal serilerin hangi faktörlerin birleşiminin etkisiyle gerçekleştiğinin tespiti ve zaman serilerinin tahmini,

(Örnek olarak bkz. Gouriéroux, Monfort ve Renne, 2017; Chowdhury, Chakravarty ve Hossain, 2018; Chen ve ark., 2019)

- Veri ve metin madenciliğinde; boyut indirgeme ve kümeleme,

(Örnek olarak bkz. Witten ve ark. 2016; Thomas, Zhu ve Romagnoli, 2018; Gultepe ve Makrehchi, 2018)

- Biyoenformatik ve genetikte; mikrodizin analizi ile gen ekspresyonundaki esas bileşenlerin belirlenmesi,

(Örnek olarak bkz. Zhou ve Altman, 2018; Kamal ve ark., 2018; Ghosh ve ark., 2019)

- Kimyada; gıdaların üretim ya da analiz aşamasında kullanılan yakın kızılötesi (NIR) spektroskopisindeki bileşenlerin analizi

(Örnek olarak bkz. Liu ve ark., 2018; Alves ve ark., 2019; Delaporte ve ark., 2019)

gibi amaçlar için yaygın olarak kullanılmaktadır.

Görüldüğü üzere BBA ortaya çıkışından günümüze kadar, hemen hemen bütün bilim dallarına hitap eden geniş bir uygulama yelpazesine sahip hale gelmiştir.

2.2. Bağımsız Bileşenler Analizine Öncülük Eden İki Problem

BBA; çok değişkenli verilerin altta yatan faktörlerini ya da bileşenlerini bulmak için kullanılan istatistiksel bir yöntemdir. Hem istatistiksel olarak birbirinden bağımsız hem de normal olmayan bileşenleri araması, BBA'yı diğer yöntemlerden ayıran en önemli özellik olarak ön plana çıkmaktadır. BBA'nın asıl çıkış noktası; kör kaynak ayrıştırma problemine çözüm bulmak gibi gözükse de, temelinde çok değişkenli verinin daha iyi bir temsiline bulunması yatmaktadır (Hyvärinen, Karhunen ve Oja, 2001).

2.2.1. Çok Değişkenli Verinin Temsili Problemi

İstatistik ve ilgili alanlardaki problemlerden biri de, çok değişkenli verinin uygun bir temsiline ya da gösteriminin nasıl bulunacağıdır. Burada temsilden kasıt; veriyi, esas yapısı daha görünür (ulaşılır) olacak şekle dönüştürmektir. İyi bir temsil; veri madenciliği, açıklayıcı veri analizi, sinyal işleme gibi pek çok yöntemin esas amacıdır (Hyvärinen, Karhunen ve Oja, 2001).

Çok değişkenli verinin temsili problemini bir örnek üzerinde açıklamak için, birlikte gözlemlenen ve birkaç değişkenden oluşan bir veri kümesi düşünölsün. Değişkenlerin sayısı; p , gözlemlerin sayısı; T , veriler; $x_i(t)$, $i=1, \dots, p$ ve $t=1, \dots, T$ ile gösterilsin. Burada p ve T boyutları çok büyük olabilir. “Veri kümesinin, $k \leq p$ olmak üzere p boyutlu uzaydan k boyutlu uzaya dönüşümünü sağlayan ve dönüştürölmüş değişkenlerin (verideki esas yapıyı anlatan altta yatan faktörlerin ya da bileşenlerin) verideki saklı bilgiyi verdiği fonksiyon ne olabilir?” sorusunun cevabı, çok değişkenli verinin iyi bir temsiline verecektir. Bu soruyu matematiksel olarak ifade etmek için, Eşitlik 2.1 kullanılabilir:

$$y_i(t) = \sum_j w_{ij} x_j(t) \quad i = 1, \dots, k \quad j = 1, \dots, p \quad (2.1)$$

Burada, yorum ve hesaplama kolaylığı açısından doğrusal fonksiyonlar göz önünde bulundurulmuştur. Bu yüzden y_i ile gösterilen her bileşen, gözlenen değişkenlerin doğrusal bir birleşimi olarak ifade edilmiştir. Burada w_{ij} 'ler, gözlenen değişkenlerin temsilini ifade eden ağırlık katsayılarıdır. Aynı problem, matris-vektör gösterimi ile de Eşitlik 2.2'deki gibi ifade edilebilir:

$$\begin{pmatrix} y_1(t) \\ y_2(t) \\ \cdot \\ \cdot \\ y_p(t) \end{pmatrix} = \mathbf{W} \begin{pmatrix} x_1(t) \\ x_2(t) \\ \cdot \\ \cdot \\ x_p(t) \end{pmatrix} \quad (2.2)$$

Eşitlik 2.2'de her bir $x_i(t)$, $t = 1, \dots, T$ bir rasgele değişkenin örneklemidir. \mathbf{W} matrisi, y_i dönüştürülmüş bileşenlerinin istatistiksel özellikleri kullanılarak belirlenebilir (Hyvärinen, Karhunen ve Oja, 2001).

\mathbf{W} matrisini belirlemek için kullanılan bir yöntem, y_i bileşenlerinin sayısını sınırlamak ve y_i 'nin verideki bilgiyi olabildiğince fazla içermesini sağlayacak şekilde y_i 'yi belirlemektir. Bu yöntem aynı zamanda, Temel Bileşenler Analizi'ne ve Faktör Analizi'ne öncülük eder (Hyvärinen, Karhunen ve Oja, 2001).

\mathbf{W} matrisinin belirlenmesinde kullanılan bir başka yöntem ise bağımsızlıktır. Buna göre, y_i bileşenleri istatistiksel olarak bağımsız olacak şekilde \mathbf{W} matrisi belirlenmelidir. Başka bir ifadeyle, herhangi bir bileşenin değeri, diğer bileşenlerin değerleri hakkında bilgi vermemelidir. Aslında FA ve TBA'da elde edilen faktörler ve bileşenler de bağımsızdır; ancak bu kısmen doğrudur. Çünkü bu yöntemler, verinin normal dağıldığını varsayar. Veri normal dağılıyorsa bağımsız bileşenleri elde etmek kolaydır. Bunun nedeni, normal dağılan verilerde ilişkisiz bileşenlerin her zaman bağımsız olmasıdır. Gerçekte ise veri, çoğu zaman normal dağılım göstermez. Bu da BBA'nın çıkış noktasını oluşturur (Hyvärinen, Karhunen ve Oja, 2001).

2.2.2. Kör Kaynak Ayırıştırma Problemi

Çok değişkenli verinin iyi bir temsilini bulma problemi, sinyal işleme alanında da çok karşılaşılan bir problemdir. Sinyal işlemede karşılaşılan bu problemi bir örnek üzerinde açıklamak için bazı fiziksel nesnelere ya da kaynaklar tarafından birtakım sinyallerin yayıldığı bir durum düşünölsün. Bu fiziksel kaynaklar; beynin farklı alanlarından yayılan elektrik sinyalleri, aynı odada konuşan insanlardan yayılan konuşma sinyalleri veya cep telefonlarından yayılan radyo dalgaları olabilir. Böyle bir ortamda pek çok alıcı olduğu varsayılırsa, bu alıcılar farklı pozisyonlarda olduğu için her kayıt, orijinal kaynak sinyallerinin farklı ağırlıklarla karışımından oluşur. Bu ağırlıklar, alıcıların kaynak sinyallere olan uzaklıklarına bağlı olarak değişir (Hyvärinen, Karhunen ve Oja, 2001).

Benzer bir durumu matematiksel olarak ifade etmek için sadece 3 tane kaynak sinyal ve 3 tane de gözlenen sinyal olduğu varsayılınsın. t zamanında gözlenen sinyaller $x_1(t)$, $x_2(t)$, $x_3(t)$ ve orijinal sinyaller de $s_1(t)$, $s_2(t)$, $s_3(t)$ ile gösterilsin. $x_i(t)$, Eşitlik 2.3'teki gösterildiği gibi kaynaklar ile alıcılar arasındaki uzaklığa bağlı olan a_{ij} katsayıları ile $s_i(t)$ 'lerin ağırlıklı toplamlarına eşittir. Eşitlik 2.3'teki a_{ij} 'ler esasında karışım ağırlıklarının veren sabit katsayılardır. Fiziksel karışım sisteminin tüm özellikleri bilinmediğinden a_{ij} değerlerinin de bilinmediği varsayılır. Ayrıca s_i kaynak sinyalleri de doğrudan kaydedilemedikleri için bilinmemektedir.

$$\begin{aligned}x_1(t) &= a_{11}s_1(t) + a_{12}s_2(t) + a_{13}s_3(t) \\x_2(t) &= a_{21}s_1(t) + a_{22}s_2(t) + a_{23}s_3(t) \\x_3(t) &= a_{31}s_1(t) + a_{32}s_2(t) + a_{33}s_3(t)\end{aligned}\tag{2.3}$$

Gözlenen $x_1(t)$, $x_2(t)$, $x_3(t)$ karışımlarından, kaynak sinyallerin bulunmaya çalışılması kör kaynak ayırıştırma problemi olarak adlandırılmaktadır. Kör kelimesinin kullanılma sebebi, kaynak sinyaller hakkında herhangi bir şey bilinmemesidir. a_{ij} karışım katsayılarından oluşan matrisin tersi alınarak \mathbf{W} matrisi bulunabilir ve s_i 'ler Eşitlik 2.4'teki gibi ayrıştırılabilir:

$$\begin{aligned}s_1(t) &= w_{11}x_1(t) + w_{12}x_2(t) + w_{13}x_3(t) \\s_2(t) &= w_{21}x_1(t) + w_{22}x_2(t) + w_{23}x_3(t) \\s_3(t) &= w_{31}x_1(t) + w_{32}x_2(t) + w_{33}x_3(t)\end{aligned}\tag{2.4}$$

Bu durumda problem, verinin iyi bir temsilini bulan Eşitlik 2.1'e benzemektedir. Buradaki her bir $x_i(t)$, $t = 1, \dots, T$ sinyali, x_i raslantı değişkeninin bir örnekleme olarak düşünülebilir (Hyvärinen, Karhunen ve Oja, 2001).

Kaynak ayrıştırmada w_{ij} katsayılarının tahmini için genel istatistiksel özellikler kullanılır ve sinyallerin istatistiksel olarak birbirlerinden bağımsız olmaları yeterlidir. Bu sebeple, sadece istatistiksel bağımsızlık bilgisini kullanan algoritmalar ile \mathbf{W} katsayılar matrisi tahmin edilebilir. Eşitlik 2.5'teki y_1, y_2, y_3 sinyalleri bağımsız ise orijinal s_1, s_2, s_3 kaynak sinyallerine eşittirler.

$$\begin{aligned} y_1(t) &= w_{11}x_1(t) + w_{12}x_2(t) + w_{13}x_3(t) \\ y_2(t) &= w_{21}x_1(t) + w_{22}x_2(t) + w_{23}x_3(t) \\ y_3(t) &= w_{31}x_1(t) + w_{32}x_2(t) + w_{33}x_3(t) \end{aligned} \quad (2.5)$$

Bu nedenle, kaynak ayrıştırma probleminde kaynak sinyaller, veri kümesinin bağımsız bileşenleri olarak da adlandırılır (Hyvärinen, Karhunen ve Oja, 2001).

2.3. Bağımsız Bileşenler Analizi için Gerekli Bazı Matematiksel ve İstatistiksel Kavramlar

BBA'nın tam olarak anlaşılabilmesi için, özellikle istatistikte ve matematikte yer alan bazı temel kavramların gözden geçirilmesi önem arz etmektedir. Bu sebeple bu kısım, çok değişkenli istatistiksel kavramlara, optimizasyon teorisine ve bilgi kuramıyla ilgili temel tanımlara ayrılmıştır. Tez çalışmasında değişkenler; küçük harflerle, vektörler; kalın küçük harflerle, matrisler ise kalın büyük harflerle gösterilmiştir.

2.3.1. Çok Değişkenli İstatistiksel Kavramlar

BBA'nın temeli, çok değişkenli verinin temsiline dayandığı için analizde raslantı değişkenleri yerine raslantı vektörleri ile işlem yapılmaktadır. Bu sebeple bu kısımda, raslantı vektörlerine dair istatistiksel özelliklerden bahsedilmiştir.

2.3.1.1. Raslantı Vektörünün Dağılımı

\mathbf{x} , Eşitlik 2.6'daki¹ gibi $x_i, i = 1, 2, \dots, n$ sürekli raslantı değişkeninden oluşan bir raslantı vektörü olmak üzere, \mathbf{x} 'in birikimli dağılım fonksiyonu Eşitlik 2.7'deki gibi tanımlanmıştır:

$$\mathbf{x} = (x_1, x_2, \dots, x_n)^T \quad (2.6)$$

$$F_{\mathbf{x}}(\mathbf{x}_0) = P(\mathbf{x} \leq \mathbf{x}_0) \quad (2.7)$$

Eşitlik 2.7; \mathbf{x} vektörünün her bir bileşeninin, \mathbf{x}_0 vektörünün bileşenlerinden küçük ya da eşit olma olasılığını vermektedir ve tek değişkenli birikimli dağılım fonksiyonuyla benzer özelliklere sahiptir. Her bir bileşen için azalmayan bir fonksiyon olup, aldığı değerler $0 \leq F_{\mathbf{x}}(\mathbf{x}) \leq 1$ aralığındadır. \mathbf{x} vektörünün bütün bileşenleri sonsuza yaklaşırken $F_{\mathbf{x}}(\mathbf{x})$, üst sınırı olan 1'e ulaşır ve herhangi bir bileşeni $x_i \rightarrow -\infty$ iken $F_{\mathbf{x}}(\mathbf{x}) = 0$ olur.

\mathbf{x} 'in çok değişkenli olasılık yoğunluk fonksiyonu $f_{\mathbf{x}}(\mathbf{x})$, $F_{\mathbf{x}}(\mathbf{x})$ 'in \mathbf{x} vektörünün bütün bileşenlerine göre türevinin alınmasıyla Eşitlik 2.8'deki gibi elde edilir. Benzer şekilde $f_{\mathbf{x}}(\mathbf{x})$ 'ten $F_{\mathbf{x}}(\mathbf{x})$ 'e de Eşitlik 2.9'daki gibi bir geçiş mevcuttur:

$$f_{\mathbf{x}}(\mathbf{x}_0) = \left. \frac{\partial}{\partial x_1} \frac{\partial}{\partial x_2} \dots \frac{\partial}{\partial x_n} F_{\mathbf{x}}(\mathbf{x}) \right|_{\mathbf{x}=\mathbf{x}_0} \quad (2.8)$$

$$F_{\mathbf{x}}(\mathbf{x}_0) = \int_{-\infty}^{x_0} f_{\mathbf{x}}(\mathbf{x}) d\mathbf{x} = \int_{-\infty}^{x_{0,1}} \int_{-\infty}^{x_{0,2}} \dots \int_{-\infty}^{x_{0,n}} f_{\mathbf{x}}(\mathbf{x}) dx_n \dots dx_2 dx_1 \quad (2.9)$$

Tek değişkenli olasılık yoğunluk fonksiyonunda olduğu gibi, çok değişkenli olasılık yoğunluk fonksiyonunda da Eşitlik 2.10 geçerlidir:

$$\int_{-\infty}^{\infty} f_{\mathbf{x}}(\mathbf{x}) d\mathbf{x} = 1 \quad (2.10)$$

n boyutlu \mathbf{x} raslantı vektörü ile m boyutlu \mathbf{y} raslantı vektörünün bileşik dağılım fonksiyonu ise Eşitlik 2.11'den yola çıkılarak Eşitlik 2.12'deki gibi bulunmaktadır.

¹ Eşitlikteki T , transpoz işlemi göstermektedir. Çalışmada tüm vektörler, aksi belirtilmediği sürece sürekli raslantı değişkenlerinden oluşan sütun vektörü olarak ele alınmıştır.

Burada \mathbf{x}_0 ve \mathbf{y}_0 , sırasıyla \mathbf{x} ve \mathbf{y} raslantı vektörleri ile aynı boyuta sahip sabit vektörlerdir.

$$F_{\mathbf{x},\mathbf{y}}(\mathbf{x}_0, \mathbf{y}_0) = P(\mathbf{x} \leq \mathbf{x}_0, \mathbf{y} \leq \mathbf{y}_0) \quad (2.11)$$

$$F_{\mathbf{x},\mathbf{y}}(\mathbf{x}_0, \mathbf{y}_0) = \int_{-\infty}^{\mathbf{x}_0} \int_{-\infty}^{\mathbf{y}_0} f_{\mathbf{x},\mathbf{y}}(\xi, \eta) d\eta d\xi \quad (2.12)$$

Ayrıca, \mathbf{x} ve \mathbf{y} raslantı vektörlerinin bileşik olasılık yoğunluk fonksiyonu $f_{\mathbf{x},\mathbf{y}}(\mathbf{x}_0, \mathbf{y}_0)$, bileşik dağılım fonksiyonu $F_{\mathbf{x},\mathbf{y}}(\mathbf{x}_0, \mathbf{y}_0)$ 'nin \mathbf{x} ve \mathbf{y} raslantı vektörlerinin tüm bileşenlerine göre türevinin alınması ile tanımlanabilir. \mathbf{x} ve \mathbf{y} raslantı vektörlerinin marjinal yoğunluk fonksiyonları ise bileşik olasılık yoğunluk fonksiyonunun diğer raslantı vektörüne göre integrali alınarak Eşitlik 2.13 ve Eşitlik 2.14'teki gibi elde edilir (Hyvärinen, Karhunen ve Oja, 2001).

$$f_{\mathbf{x}}(\mathbf{x}) = \int_{-\infty}^{\infty} f_{\mathbf{x},\mathbf{y}}(\xi, \eta) d\eta \quad (2.13)$$

$$f_{\mathbf{y}}(\mathbf{y}) = \int_{-\infty}^{\infty} f_{\mathbf{x},\mathbf{y}}(\xi, \eta) d\xi \quad (2.14)$$

2.3.1.2. Beklenen Değer ve Momentler

\mathbf{x} raslantı vektöründen herhangi bir $g(\mathbf{x})$ türetilsin. $g(\mathbf{x})$ burada skaler, vektör ya da matris olabilir. $g(\mathbf{x})$ 'in beklenen değeri Eşitlik 2.15'teki gibi elde edilir:

$$E\{g(\mathbf{x})\} = \int_{-\infty}^{\infty} g(\mathbf{x}) f_{\mathbf{x}}(\mathbf{x}) d\mathbf{x} \quad (2.15)$$

İki farklı raslantı vektörünün bileşik dağılımlarının beklenen değeri ise

$$E\{g(\mathbf{x}, \mathbf{y})\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(\mathbf{x}, \mathbf{y}) f_{\mathbf{x},\mathbf{y}}(\mathbf{x}, \mathbf{y}) d\mathbf{y} d\mathbf{x} \quad (2.16)$$

şekindedir.

Genellikle \mathbf{x} raslantı vektörünün olasılık yoğunluk fonksiyonu bilinmez; ancak \mathbf{x} raslantı vektöründen elde edilmiş olan K büyüklüğünde $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_K$ örnekleme bilinir. Raslantı vektörünün olasılık yoğunluk fonksiyonunun bilinmemesine rağmen beklenen değerin, bu örneklem değerleri kullanılarak Eşitlik 2.17'deki gibi tahmin edilmesi mümkündür:

$$E\{g(\mathbf{x})\} \approx \frac{1}{K} \sum_{j=1}^K g(\mathbf{x}_j) \quad (2.17)$$

Beklenen değerin en büyük avantajı, Eşitlik 2.17'de de gösterildiği gibi olasılık yoğunluk fonksiyonları açısından tanımlanmış olsalar bile, doğrudan verilerden de tahmin edilebilir olmalarıdır. Benzer şekilde \mathbf{x} ve \mathbf{y} raslantı vektörlerinin bileşik dağılımları için K büyüklüğünde $(\mathbf{x}_1, \mathbf{y}_1), (\mathbf{x}_2, \mathbf{y}_2), \dots, (\mathbf{x}_K, \mathbf{y}_K)$ örneklem çiftinin bilinmesi durumunda da beklenen değer,

$$E\{g(\mathbf{x}, \mathbf{y})\} \approx \frac{1}{K} \sum_{j=1}^K g(\mathbf{x}_j, \mathbf{y}_j) \quad (2.18)$$

biçiminde elde edilebilir.

Beklenen değerin sahip olduğu bazı önemli özellikler mevcuttur. Bunlardan bazıları aşağıda verilmiştir:

- Doğrusallık

$\mathbf{x}_i; i = 1, \dots, m$ farklı raslantı vektörleri kümesini, $a_i; i = 1, \dots, m$ de rasgele olmayan bazı skaler katsayıları gösterebilir. Bu durumda Eşitlik 2.19 geçerlidir.

$$E\left\{\sum_{i=1}^m a_i \mathbf{x}_i\right\} = \sum_{i=1}^m a_i E\{\mathbf{x}_i\} \quad (2.19)$$

- Doğrusal Dönüşüm

\mathbf{x} , $m \times 1$ boyutlu raslantı vektörünü, \mathbf{A} ve \mathbf{B} ise sırasıyla $k \times m$ boyutlu ve $m \times l$ boyutlu matrisleri gösterebilir. Bu durumda Eşitlik 2.20 yazılabilir.

$$E\{\mathbf{Ax}\} = \mathbf{A}E\{\mathbf{x}\}, \quad E\{\mathbf{xB}\} = E\{\mathbf{x}\}\mathbf{B} \quad (2.20)$$

- Değişmezlik

$\mathbf{y} = g(\mathbf{x})$, \mathbf{x} raslantı vektörünün vektör değerli bir fonksiyonu olsun.

$$\int_{-\infty}^{\infty} \mathbf{y} f_{\mathbf{y}}(\mathbf{y}) d\mathbf{y} = \int_{-\infty}^{\infty} g(\mathbf{x}) f_{\mathbf{x}}(\mathbf{x}) d\mathbf{x} \quad (2.21)$$

Eşitlik 2.21'den görüldüğü üzere, integraller farklı olasılık yoğunluk fonksiyonları için uygulanmalarına rağmen sonuç $E\{\mathbf{y}\} = E\{g(\mathbf{x})\}$ olmaktadır (Hyvärinen, Karhunen ve Oja, 2001).

- **Ortalama Vektörü ve Korelasyon Matrisi**

\mathbf{x} raslantı vektörünün momentleri², \mathbf{x} raslantı vektörünü karakterize etmekte kullanılan beklenen değerlerdir. \mathbf{x} raslantı vektörünün birinci dereceden momenti, ortalama vektörüdür³ ve Eşitlik 2.22'deki gibi ifade edilir:

$$\boldsymbol{\mu}_{\mathbf{x}} = E\{\mathbf{x}\} = \int_{-\infty}^{\infty} \mathbf{x} f_{\mathbf{x}}(\mathbf{x}) d\mathbf{x} \quad (2.22)$$

$f_{x_i}(x_i)$, \mathbf{x} 'in i . bileşenin marjinal olasılık yoğunluk fonksiyonu olmak üzere, n boyutlu $\boldsymbol{\mu}_{\mathbf{x}}$ vektörünün her bir bileşeni Eşitlik 2.23'teki gibi hesaplanır:

$$\mu_{x_i} = E\{x_i\} = \int_{-\infty}^{\infty} x_i f_{\mathbf{x}}(\mathbf{x}) d\mathbf{x} = \int_{-\infty}^{\infty} x_i f_{x_i}(x_i) dx_i \quad (2.23)$$

Bir diğer önemli momentler kümesi ise \mathbf{x} raslantı vektörünün bileşen çiftleri arasındaki korelasyonlardan (ilişkilerden) oluşur. \mathbf{x} 'in i . ve j . bileşenleri arasındaki r_{ij} korelasyonu, ikinci dereceden moment kullanılarak

$$r_{ij} = E\{x_i x_j\} = \int_{-\infty}^{\infty} x_i x_j f_{\mathbf{x}}(\mathbf{x}) d\mathbf{x} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_i x_j f_{x_i, x_j}(x_i, x_j) dx_i dx_j \quad (2.24)$$

biçiminde hesaplanır. Korelasyonlar, negatif ya da pozitif olabilir.

r_{ij} korelasyon çiftlerinden oluşan $n \times n$ boyutlu korelasyon matrisi ise

² Sonraki bölümlerde momentler konusu daha ayrıntılı olarak anlatılacaktır.

³ Gerektiği takdirde veriden, tahmin edilen ortalama vektörünün çıkarılması ile verinin 0 ortalamalı olması sağlanır. Bu özellik, BBA'da verinin hazırlanmasında bir önışlem adımı olarak çoğu kez kullanılmaktadır.

$$\mathbf{R}_x = E\{\mathbf{xx}^T\} \quad (2.25)$$

şeklindedir. Korelasyon matrisinin bazı önemli özellikleri şöyledir:

- Simetrik bir matristir: $\mathbf{R}_x = \mathbf{R}_x^T$
- Yarı pozitif tanımlıdır: $\mathbf{a}^T \mathbf{R}_x \mathbf{a} \geq 0$ (\mathbf{a} , n boyutlu vektör)
- \mathbf{R}_x 'in bütün özdeğerleri gerçel ve negatif değildir.

- **Kovaryans Matrisi**

\mathbf{x} 'in $n \times n$ boyutlu kovaryans matrisi Eşitlik 2.26'daki gibidir:

$$\Sigma_x = E\{(\mathbf{x} - \boldsymbol{\mu}_x)(\mathbf{x} - \boldsymbol{\mu}_x)^T\} \quad (2.26)$$

Kovaryans matrisinin elemanları olan σ_{ij} 'ler, Eşitlik 2.27'de de görüldüğü gibi r_{ij} korelasyonlarına karşılık gelen merkezsiz momentlerdir.

$$\sigma_{ij} = E\{(x_i - \mu_i)(x_j - \mu_j)\} \quad (2.27)$$

Kovaryans matrisi, korelasyon matrisi ile aynı özelliklere sahiptir. Her ikisi de ikinci dereceden istatistikleri kullanarak raslantı değişkenleri arasındaki ilişkiyi ölçer.

Korelasyon matrisi, beklenen değer özelliklerinden yararlanarak kovaryans matrisi cinsinden Eşitlik 2.28'deki gibi yazılabilir. Ayrıca ortalama vektörü $\boldsymbol{\mu}_x = \mathbf{0}$ olduğunda, korelasyon ve kovaryans matrislerinin birbirine eşit olacağı Eşitlik 2.28'den görülmektedir.

$$\mathbf{R}_x = \Sigma_x + \boldsymbol{\mu}_x \boldsymbol{\mu}_x^T \quad (2.28)$$

Bileşik dağılımlar için ise sırasıyla Eşitlik 2.29 ve Eşitlik 2.30'daki çapraz korelasyon ve çapraz kovaryans matrisleri kullanılır. Eğer \mathbf{x} ve \mathbf{y} raslantı vektörlerinin ortalama vektörleri $\mathbf{0}$ ise çapraz korelasyon ve çapraz kovaryans matrisleri birbirine eşit olur.

$$\mathbf{R}_{xy} = E\{\mathbf{xy}^T\} \quad (2.29)$$

$$\Sigma_{xy} = E\{(\mathbf{x} - \boldsymbol{\mu}_x)(\mathbf{y} - \boldsymbol{\mu}_y)^T\} \quad (2.30)$$

Burada, \mathbf{x} ve \mathbf{y} raslantı vektörlerinin boyutları farklı olabilir. Bunun için çapraz korelasyon ve çapraz kovaryans matrislerinin kare matris olması gerekmektedir. Ayrıca, çapraz korelasyon ve kovaryans matrislerinin tanımları gereği Eşitlik 2.31 de yazılabilir.

$$\mathbf{R}_{\mathbf{xy}} = \mathbf{R}_{\mathbf{yx}}^T, \quad \Sigma_{\mathbf{xy}} = \Sigma_{\mathbf{yx}}^T \quad (2.31)$$

Aynı boyutlu \mathbf{x} ve \mathbf{y} raslantı vektörlerinin toplamlarının kovaryans matrisi ise Eşitlik 2.32'deki gibidir (Hyvärinen, Karhunen ve Oja, 2001):

$$\Sigma_{\mathbf{x+y}} = \Sigma_{\mathbf{x}} + \Sigma_{\mathbf{xy}} + \Sigma_{\mathbf{yx}} + \Sigma_{\mathbf{y}} \quad (2.32)$$

2.3.1.3. İlişkisizlik ve Beyazlık

\mathbf{x} ve \mathbf{y} raslantı vektörleri arasında korelasyon yok ise Eşitlik 2.33'te görülen çapraz kovaryans matrisleri $\mathbf{0}$ matrisine eşittir. Bu durum, ilişkisizlik ya da korelasyonsuzluk olarak adlandırılır ve Eşitlik 2.34'teki gibi yazılabilir:

$$\Sigma_{\mathbf{xy}} = E\{(\mathbf{x} - \boldsymbol{\mu}_{\mathbf{x}})(\mathbf{y} - \boldsymbol{\mu}_{\mathbf{y}})^T\} = \mathbf{0} \quad (2.33)$$

$$\mathbf{R}_{\mathbf{xy}} = E\{\mathbf{xy}^T\} = E\{\mathbf{x}\}E\{\mathbf{y}^T\} = \boldsymbol{\mu}_{\mathbf{x}}\boldsymbol{\mu}_{\mathbf{y}}^T \quad (2.34)$$

İlişkisizlik, x ve y raslantı değişkenleri için ise

$$\sigma_{xy} = E\{(x - \mu_x)(y - \mu_y)\} = 0 \quad (2.35)$$

$$r_{xy} = E\{xy\} = E\{x\}E\{y\} = \mu_x\mu_y \quad (2.36)$$

şeklinde ifade edilebilir. \mathbf{x} raslantı vektörünün bileşenleri karşılıklı olarak korelasyonsuz ise bu durumda kovaryans matrisi Eşitlik 2.37'deki gibi olur:

$$\Sigma_{\mathbf{x}} = E\{(\mathbf{x} - \boldsymbol{\mu}_{\mathbf{x}})(\mathbf{x} - \boldsymbol{\mu}_{\mathbf{x}})^T\} = \mathbf{D} \quad (2.37)$$

Burada $\mathbf{D} = \text{köş}(\sigma_{x_1}^2, \sigma_{x_2}^2, \dots, \sigma_{x_n}^2)$, $n \times n$ boyutlu köşegen matristir ve elemanları \mathbf{x} 'in x_i bileşenlerinin $\sigma_{x_i}^2 = E\{(x_i - \mu_{x_i})^2\}$ varyanslarından oluşur.

$\mathbf{0}$ ortalama ve birim kovaryans (ve bu yüzden korelasyon) matrisine (sabit bir varyans terimiyle çarpılmış da olabilir) sahip raslantı vektörleri, beyazdırlar. Beyaz raslantı vektörleri için Eşitlik 2.38'deki koşullar sağlanır. Burada \mathbf{I} , $n \times n$ boyutlu birim matristir.

$$\boldsymbol{\mu}_x = \mathbf{0}, \mathbf{R}_x = \boldsymbol{\Sigma}_x = \mathbf{I} \quad (2.38)$$

Matrislerin beyaz hale getirilmesine, beyazlatma (whitening) denir. Beyazlık, ilişkisizliğin özel bir halidir ve beyazlık özelliği, ortogonal dönüşümlerden etkilenmemektedir. Bu özellik, matematiksel olarak aşağıdaki gibi ifade edilebilir:

Eşitlik 2.39'daki gibi $n \times n$ boyutlu \mathbf{T} matrisiyle tanımlanan ortogonal dönüşüm, \mathbf{x} raslantı vektörüne uygulansın. Ortogonal bir \mathbf{T} matrisi, n boyutlu uzayda normları ve uzaklıkları koruyan bir döndürmeyi (koordinat eksenlerinin değişimi) belirtir.

$$\mathbf{y} = \mathbf{T}\mathbf{x}, \mathbf{T}^T\mathbf{T} = \mathbf{T}\mathbf{T}^T = \mathbf{I} \quad (2.39)$$

\mathbf{x} , beyaz bir raslantı vektörü ise Eşitlik 2.40 ve Eşitlik 2.41 elde edilir:

$$\boldsymbol{\mu}_y = E\{\mathbf{T}\mathbf{x}\} = \mathbf{T}E\{\mathbf{x}\} = \mathbf{T}\boldsymbol{\mu}_x = \mathbf{0} \quad (2.40)$$

$$\boldsymbol{\Sigma}_y = \mathbf{R}_y = E\{\mathbf{T}\mathbf{x}(\mathbf{T}\mathbf{x})^T\} = \mathbf{T}E\{\mathbf{x}\mathbf{x}^T\}\mathbf{T}^T = \mathbf{T}\mathbf{R}_x\mathbf{T}^T = \mathbf{T}\mathbf{T}^T = \mathbf{I} \quad (2.41)$$

Eşitlik 2.40 ve Eşitlik 2.41'e göre \mathbf{y} vektörü de beyazdır. Dolayısıyla, beyazlık özelliğinin ortogonal dönüşümler altında korunduğu söylenebilir (Hyvärinen, Karhunen ve Oja, 2001). Beyazlatma işlemi, BBA'da oldukça kullanışlı bir ön hazırlık aşamasıdır.

2.3.1.4. İstatistiksel Bağımsızlık

BBA'nın temelini oluşturan anahtar kavram, istatistiksel bağımsızlıktır. Basitlik açısından ilk olarak iki farklı x ve y raslantı değişkenleri ele alınsın. Bu iki raslantı değişkeni birbirinden bağımsız ise birinin değerini bilmek, diğerinin değeriyle ilgili herhangi bir bilgi vermez. Matematiksel olarak ise istatistiksel bağımsızlık, olasılıklar ile tanımlanabilir. Eğer Eşitlik 2.42'deki koşul sağlanıyorsa " x ve y bağımsızdır." denir.

$$f_{x,y}(x, y) = f_x(x)f_y(y) \quad (2.42)$$

Buna eşdeğer olarak, eşitlikteki olasılık yoğunluk fonksiyonları çarpanlarına ayrılabilen birikimli dağılım fonksiyonlarıyla da değiştirilebilir. Ayrıca, bağımsız değişkenler Eşitlik

2.43'teki temel özelliği sağlarlar. Burada $g(x)$ ve $h(y)$, integrali alınabilir herhangi iki fonksiyondur.

$$\begin{aligned}
 E\{g(x)h(y)\} &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x)h(y)f_{x,y}(x,y)dx dy \\
 &= \int_{-\infty}^{\infty} g(x)f_x(x)dx \int_{-\infty}^{\infty} h(y)f_y(y)dy \\
 &= E\{g(x)\}E\{h(y)\}
 \end{aligned} \tag{2.43}$$

Eşitlik 2.43, istatistiksel bağımsızlığın ilişkisizlikten çok daha güçlü bir özellik olduğunu ortaya çıkarır. Çünkü daha önce ilişkisizliği açıklayan Eşitlik 2.34, $g(x)$ ve $h(y)$ 'nin doğrusal fonksiyonlar olduğu özel durum için Eşitlik 2.41'deki bağımsızlık özelliğinden elde edilmişti. Ancak raslantı değişkenlerinin bileşik olasılık yoğunluk fonksiyonu normal dağılıma sahip ve ilişkisizler ise aynı zamanda bağımsızlık özelliğini de sağlarlar.

İki raslantı değişkenininin bağımsızlığı, raslantı vektörleri için de genelleştirilebilir. $\mathbf{x}, \mathbf{y}, \mathbf{z}, \dots$; farklı boyutlara sahip raslantı vektörleri olsun. $\mathbf{x}, \mathbf{y}, \mathbf{z}, \dots$ için bağımsızlık koşulu Eşitlik 2.44'ten ve sağladıkları temel özellik de Eşitlik 2.45'ten görülmektedir (Hyvärinen, Karhunen ve Oja, 2001):

$$f_{\mathbf{x}, \mathbf{y}, \mathbf{z}, \dots}(\mathbf{x}, \mathbf{y}, \mathbf{z}, \dots) = f_{\mathbf{x}}(\mathbf{x})f_{\mathbf{y}}(\mathbf{y})f_{\mathbf{z}}(\mathbf{z})\dots \tag{2.44}$$

$$E\{g_{\mathbf{x}}(\mathbf{x})g_{\mathbf{y}}(\mathbf{y})g_{\mathbf{z}}(\mathbf{z})\dots\} = E\{g_{\mathbf{x}}(\mathbf{x})\}E\{g_{\mathbf{y}}(\mathbf{y})\}E\{g_{\mathbf{z}}(\mathbf{z})\}\dots \tag{2.45}$$

2.3.1.5. Koşullu Dağılımlar ve Bayes Kuralı

“ \mathbf{y}_0 sabit değerine sahip \mathbf{y} raslantı vektörü verildiğinde, \mathbf{x} raslantı vektörünün olasılık yoğunluk fonksiyonu nedir?” veya “ \mathbf{x}_0 sabit değerine sahip \mathbf{x} raslantı vektörü verildiğinde, \mathbf{y} raslantı vektörünün olasılık yoğunluk fonksiyonu nedir?” sorularının cevaplarını koşullu dağılımlar verir ve

$$f_{\mathbf{x}|\mathbf{y}}(\mathbf{x}|\mathbf{y}) = \frac{f_{\mathbf{x},\mathbf{y}}(\mathbf{x},\mathbf{y})}{f_{\mathbf{y}}(\mathbf{y})}, \quad f_{\mathbf{y}|\mathbf{x}}(\mathbf{y}|\mathbf{x}) = \frac{f_{\mathbf{x},\mathbf{y}}(\mathbf{x},\mathbf{y})}{f_{\mathbf{x}}(\mathbf{x})} \tag{2.46}$$

şeklinde ifade edilirler. Burada \mathbf{y}_0 ya da \mathbf{x}_0 genellikle, \mathbf{y} ve \mathbf{x} vektörlerindeki belirli bir değerdir. \mathbf{x} ve \mathbf{y} raslantı vektörleri istatistiksel olarak birbirlerinden bağımsız ise $f_{\mathbf{x}|\mathbf{y}}(\mathbf{x}|\mathbf{y})$ koşullu dağılımı \mathbf{x} 'in koşulsuz $f_{\mathbf{x}}(\mathbf{x})$ dağılımına, $f_{\mathbf{y}|\mathbf{x}}(\mathbf{y}|\mathbf{x})$ koşullu dağılımı da \mathbf{y} 'nin koşulsuz $f_{\mathbf{y}}(\mathbf{y})$ dağılımına eşittir.

Eşitlik 2.46'dan \mathbf{x} ve \mathbf{y} 'nin bileşik dağılımı Eşitlik 2.47'deki gibi yeniden yazılabilir. Eşitlik 2.47'den hareketle Bayes Kuralı olarak bilinen Eşitlik 2.48 elde edilir. Eşitlik 2.48 ile \mathbf{y} 'nin $f_{\mathbf{y}}(\mathbf{y})$ önsel dağılımı bilindiğinde, \mathbf{x} raslantı vektörünün verilen özel bir değeri için \mathbf{y} 'nin $f_{\mathbf{y}|\mathbf{x}}(\mathbf{y}|\mathbf{x})$ sonsal dağılımı hesaplanabilir.

$$f_{\mathbf{x},\mathbf{y}}(\mathbf{x},\mathbf{y}) = f_{\mathbf{x}|\mathbf{y}}(\mathbf{x}|\mathbf{y})f_{\mathbf{y}}(\mathbf{y}) = f_{\mathbf{y}|\mathbf{x}}(\mathbf{y}|\mathbf{x})f_{\mathbf{x}}(\mathbf{x}) \quad (2.47)$$

$$f_{\mathbf{y}|\mathbf{x}}(\mathbf{y}|\mathbf{x}) = \frac{f_{\mathbf{x}|\mathbf{y}}(\mathbf{x}|\mathbf{y})f_{\mathbf{y}}(\mathbf{y})}{f_{\mathbf{x}}(\mathbf{x})} \quad (2.48)$$

Ayrıca, koşullu dağılımlar için koşullu beklenen değerler de

$$E\{g(\mathbf{x},\mathbf{y})|\mathbf{y}\} = \int_{-\infty}^{\infty} g(\xi,\mathbf{y})f_{\mathbf{x}|\mathbf{y}}(\xi|\mathbf{y})d\xi \quad (2.49)$$

$$E\{g(\mathbf{x},\mathbf{y})\} = E\{E\{g(\mathbf{x},\mathbf{y})|\mathbf{y}\}\} \quad (2.50)$$

biçiminde bulunabilir (Hyvärinen, Karhunen ve Oja, 2001).

2.3.1.6. Çok Değişkenli Normal Dağılım

$n \times 1$ boyutlu \mathbf{x} raslantı vektörü normal dağılıma sahip ise \mathbf{x} 'in olasılık yoğunluk fonksiyonu Eşitlik 2.51'deki gibidir:

$$f_{\mathbf{x}}(\mathbf{x}) = \frac{1}{(2\pi)^{n/2} (\det \Sigma_{\mathbf{x}})^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_{\mathbf{x}})^T \Sigma_{\mathbf{x}}^{-1}(\mathbf{x}-\boldsymbol{\mu}_{\mathbf{x}})\right) \quad (2.51)$$

$\boldsymbol{\mu}_{\mathbf{x}}$; \mathbf{x} 'in ortalama vektörünü, $\Sigma_{\mathbf{x}}$ ise kovaryans matrisini göstermektedir.

Çok değişkenli normal dağılım için bazı özellikler şu şekilde sıralanabilir:

- Çok deęişkenli normal daęılımının yazılabilmesi için \mathbf{x} raslantı vektörünün sadece birinci dereceden ($\boldsymbol{\mu}_x$) ve ikinci dereceden ($\boldsymbol{\Sigma}_x$) istatistiklerinin bilinmesi yeterlidir. Bütün yüksek dereceden momentleri sadece $\boldsymbol{\mu}_x$ ve $\boldsymbol{\Sigma}_x$ 'e baęlıdır ve dięer momentlerin bilinmesi yeni bir bilgi getirmez. Bu sebeple, normal daęılan bir veri için birinci ve ikinci dereceden istatistiksel bilgiye dayalı yöntemler yeterlidir.

- Eęer \mathbf{x} raslantı vektörü normal daęılıyor ise \mathbf{x} 'in doęrusal bir dönüřümü olan $\mathbf{y} = \mathbf{A}\mathbf{x}$ de $\boldsymbol{\mu}_y = \mathbf{A}\boldsymbol{\mu}_x$ ortalama vektörü ve $\boldsymbol{\Sigma}_y = \mathbf{A}\boldsymbol{\Sigma}_x\mathbf{A}^T$ kovaryans matrisi ile normal daęılır.

- Bileşik normal daęılıma sahip raslantı vektörlerinin marjinal ve kořullu daęılımları da normal daęılım gösterir (Hyvärinen, Karhunen ve Oja, 2001).

Bilindięi üzere raslantı deęişkenleri ilişkisiz olmalarına raęmen baęımsız olmayabilirler. Ancak çok deęişkenli normal daęılıma sahip raslantı vektörünün ilişkisiz olan bileşenleri, aynı zamanda baęımsızdır (Eaton, 2007). Çok deęişkenli normal daęılımın kovaryans matrisi $\boldsymbol{\Sigma}_x$ köşegen bir matris deęil ise \mathbf{x} raslantı vektörünün bileşenleri birbirleriyle ilişkilidir (Hyvärinen, Karhunen ve Oja, 2001).

2.3.1.7. Merkezi Limit Teoremi

Baęımsız ve aynı daęılıma sahip k tane standart raslantı deęişkenin toplamının ya da aritmetik ortalamasının daęılımı, $k \rightarrow \infty$ iken 0 ortalamalı ve birim varyanslı normal daęılıma yakınsar. Bu sonuç, merkezi limit teoremi olarak bilinir. Merkezi limit teoremini raslantı vektörleri için de düşünmek mümkündür. Eęer k tane \mathbf{z}_i raslantı vektörleri, ortak bir $\boldsymbol{\mu}_z$ ortalama vektörü ve $\boldsymbol{\Sigma}_z$ kovaryans matrisi ile baęımsız ve aynı daęılıma sahip ise $k \rightarrow \infty$ için Eşitlik 2.52'de görülen \mathbf{y}_k raslantı vektörü, $\mathbf{0}$ ortalama ve $\boldsymbol{\Sigma}_z$ kovaryans matrisi ile çok deęişkenli normal daęılıma sahiptir.

$$\mathbf{y}_k = \frac{1}{\sqrt{k}} \sum_{i=1}^k (\mathbf{z}_i - \boldsymbol{\mu}_z) \quad (2.52)$$

2.3.1.8. Dönüşümlerin Olasılık Yoğunluk Fonksiyonu

n boyutlu \mathbf{x} ve \mathbf{y} raslantı vektörleri arasında Eşitlik 2.53'teki gibi bir dönüşüm olsun ve bu dönüşümün Eşitlik 2.54'teki gibi elde edilebilir tek bir tersi olsun.

$$\mathbf{y} = g(\mathbf{x}) \quad (2.53)$$

$$\mathbf{x} = g^{-1}(\mathbf{y}) \quad (2.54)$$

Bu durumda \mathbf{y} 'nin olasılık yoğunluk fonksiyonu Eşitlik 2.55'teki gibi \mathbf{x} 'in olasılık yoğunluk fonksiyonundan elde edilir.

$$f_y(\mathbf{y}) = \frac{1}{|\det \mathbf{J}g(g^{-1}(\mathbf{y}))|} f_x(g^{-1}(\mathbf{y})) \quad (2.55)$$

Eşitlik 2.55'te $\mathbf{J}g$, Jakobiyen matristir ve elemanları Eşitlik 2.56'daki gibidir. $g_j(\mathbf{x})$ ise $g(\mathbf{x})$ vektör fonksiyonunun j . bileşenidir.

$$\mathbf{J}g(\mathbf{x}) = \begin{bmatrix} \frac{\partial g_1(\mathbf{x})}{\partial x_1} & \frac{\partial g_2(\mathbf{x})}{\partial x_1} & \dots & \frac{\partial g_n(\mathbf{x})}{\partial x_1} \\ \frac{\partial g_1(\mathbf{x})}{\partial x_2} & \frac{\partial g_2(\mathbf{x})}{\partial x_2} & \dots & \frac{\partial g_n(\mathbf{x})}{\partial x_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_1(\mathbf{x})}{\partial x_n} & \frac{\partial g_2(\mathbf{x})}{\partial x_n} & \dots & \frac{\partial g_n(\mathbf{x})}{\partial x_n} \end{bmatrix} \quad (2.56)$$

Eğer \mathbf{x} ve \mathbf{y} arasındaki dönüşüm doğrusal ve tekil değil ise başka bir ifadeyle; $\mathbf{y} = \mathbf{A}\mathbf{x}$ ve $\mathbf{x} = \mathbf{A}^{-1}\mathbf{y}$ şeklinde ise Eşitlik 2.55, Eşitlik 2.57'ye dönüşür. Daha önce de belirtildiği gibi Eşitlik 2.55'teki \mathbf{x} raslantı vektörü çok değişkenli normal dağılıma sahip ise \mathbf{y} raslantı vektörü de çok değişkenli normal dağılıma sahip olur (Hyvärinen, Karhunen ve Oja, 2001).

$$f_y(\mathbf{y}) = \frac{1}{|\det \mathbf{A}|} f_x(\mathbf{A}^{-1}\mathbf{y}) \quad (2.57)$$

2.3.1.9. Yüksek Dereceden İstatistikler

Yüksek dereceden istatistikler, iki ve daha yüksek dereceden istatistikleri kapsamaktadır. Genel olarak BBA'da momentler ve kümülanlar, yüksek dereceden istatistikler olarak kullanılmaktadırlar.

- **Momentler**

x raslantı değişkeninin j . merkezsiz olmayan momenti ve j . merkezsiz momenti sırasıyla Eşitlik 2.58 ve Eşitlik 2.59'daki gibidir.

$$\alpha_j = E\{x^j\} = \int_{-\infty}^{\infty} \zeta^j f_x(\zeta) d\zeta, \quad j=1,2,\dots \quad (2.58)$$

$$m_j = E\{(x - \alpha_1)^j\} = \int_{-\infty}^{\infty} (\zeta - \mu_x)^j f_x(\zeta) d\zeta, \quad j=1,2,\dots \quad (2.59)$$

Birinci merkezsiz olmayan moment α_1 ; x 'in ortalaması, ikinci merkezsiz olmayan moment α_2 ise x 'in ortalama gücüdür. Birinci merkezsiz moment anlamsız iken, ikinci merkezsiz moment ise varyansa eşittir. Üçüncü merkezsiz moment olan çarpıklık, olasılık yoğunluk fonksiyonunun simetrikliği hakkında bilgi verir. Dördüncü dereceden yüksek momentler ve istatistikler ise uygulamalarda nadiren kullanılır. Bazı BBA algoritmalarında kolay hesaplanmasından dolayı dördüncü merkezsiz olmayan moment $\alpha_4 = E\{x^4\}$ kullanılmaktadır. Birçok uygulamada, kullanışlı özelliklerinden dolayı dördüncü merkezsiz moment $\mu_4 = E\{(x - \mu_x)^4\}$ yerine basıklık olarak adlandırılan dördüncü dereceden istatistik kullanılmaktadır. Ortalamanın sıfır olduğu durum için basıklık,

$$bas(x) = E\{x^4\} - 3[E\{x^2\}]^2 \quad (2.60)$$

şeklinindedir. Eşitlik 2.60 yerine Eşitlik 2.61'deki normalize edilmiş basıklık da kullanılabilir.

$$bas(x) = \frac{E\{x^4\}}{[E\{x^2\}]^2} - 3 \quad (2.61)$$

Beyazlatılmış veriler için $E\{x^2\} = 1$ olacağından, basıklığın her iki versiyonu da beyaz veriler için Eşitlik 2.62'ye dönüşür. Bu sebeple, beyaz veriler için x 'in dağılımını karakterize ederken basıklık yerine, dördüncü moment kullanılabilir.

$$bas(x) = \tilde{\kappa}(x) = E\{x^4\} - 3 \quad (2.62)$$

x ve y raslantı değişkenleri istatistiksel olarak bağımsız ise basıklıkları için Eşitlik 2.63'teki toplanabilirlik özelliği geçerlidir. Bu özellik, dördüncü merkezsiz olmayan moment için geçerli olmadığından momentler yerine kümülanlar daha kullanışlıdır.

$$bas(x + y) = bas(x) + bas(y) \quad (2.63)$$

Herhangi bir skaler β parametresi için basıklık aynı zamanda Eşitlik 2.64'teki özelliğe de sahiptir.

$$bas(\beta x) = \beta^4 bas(x) \quad (2.64)$$

Basıklık, raslantı değişkeninin normal dağılıp dağılmadığını belirlemek için kullanılacak en basit istatistiktir.⁴ x raslantı değişkeni normal dağılıma sahip ise basıklık değeri 0 olur. Eğer x raslantı değişkeni negatif basıklık değerine sahip ise normal dağılıma göre daha az basık (subgaussian), pozitif basıklık değerine sahip ise de normal dağılıma göre daha dik (supergaussian) bir dağılıma sahip olur (Hyvärinen, Karhunen ve Oja, 2001).

- **Kümülanlar**

x 'in gerçek değerli, 0 ortalamalı, $f_x(x)$ olasılık yoğunluk fonksiyonuna sahip sürekli bir raslantı değişkeni olduğu varsayalım. x 'in birinci karakteristik fonksiyonu, olasılık

⁴ Basıklık, BBA'da elde edilen bağımsız bileşenlerin normal olup olmadığının tespitinde sıkça kullanılmaktadır.

yoğunluk fonksiyonunun sürekli Fourier dönüşümü şeklinde Eşitlik 2.65'teki gibi yazılabilir:

$$\varphi(\omega) = E\{\exp(i\omega x)\} = \int_{-\infty}^{\infty} \exp(i\omega x) f_x(x) dx \quad (2.65)$$

Eşitlikte $i = \sqrt{-1}$ ve ω , x 'e karşılık gelen dönüştürülmüş değişkendir. Her olasılık dağılımının tek bir karakteristik fonksiyonu mevcuttur ve bu ifadenin tersi de doğrudur. Karakteristik fonksiyonun Taylor serisi açılımından

$$\varphi(\omega) = \int_{-\infty}^{\infty} \left(\sum_{k=0}^{\infty} \frac{x^k (i\omega)^k}{k!} \right) f_x(x) dx = \sum_{k=0}^{\infty} E\{x^k\} \frac{(i\omega)^k}{k!} \quad (2.66)$$

eşitliğine ulaşılır. Bu açılımın katsayıları x 'in merkezsiz olmayan $E\{x^k\}$ momentlerini vermektedir. Bu nedenle karakteristik fonksiyon, moment çıkarıcı fonksiyon olarak da adlandırılmaktadır.

x 'in genellikle ikinci karakteristik fonksiyonu ya da kümülanant çıkarıcı fonksiyonu kullanılmaktadır. Bu fonksiyon, birinci karakteristik fonksiyonun doğal logaritması alınarak bulunur.

$$\phi(\omega) = \ln(\varphi(\omega)) = \ln(E\{\exp(i\omega x)\}) \quad (2.67)$$

x 'in k 'nıncı kümülanantlarına (κ_k), ikinci karakteristik fonksiyonun Taylor serisi açılımından

$$\phi(\omega) = \sum_{k=0}^{\infty} \kappa_k \frac{(i\omega)^k}{k!} \quad (2.68)$$

$$\kappa_k = (-i)^k \left. \frac{d^k \phi(\omega)}{d\omega^k} \right|_{\omega=0} \quad (2.69)$$

eşitlikleriyle ulaşılır. x , 0 ortalamalı bir raslantı değişkeni ise ilk 4 kümülanantı Eşitlik 2.70'teki gibi olacaktır.

$$\kappa_1 = 0, \quad \kappa_2 = E\{x^2\}, \quad \kappa_3 = E\{x^3\}, \quad \kappa_4 = E\{x^4\} - 3[E\{x^2\}]^2 \quad (2.70)$$

Görüldüğü üzere ilk üç kümülanant, ilk üç momente eşit iken dördüncü kümülanant ise basırlığa eşittir.

Eğer \mathbf{x} , $f_{\mathbf{x}}(\mathbf{x})$ şeklinde olasılık yoğunluk fonksiyonuna sahip bir raslantı vektörü olsaydı, karakteristik fonksiyonu yine olasılık yoğunluk fonksiyonunun Fourier dönüşümüyle bulunabilirdi.

$$\varphi(\boldsymbol{\omega}) = E\{\exp(i\boldsymbol{\omega}\mathbf{x})\} = \int_{-\infty}^{\infty} \exp(i\boldsymbol{\omega}\mathbf{x})f_{\mathbf{x}}(\mathbf{x})d\mathbf{x} \quad (2.71)$$

Eşitlik 2.71’de $\boldsymbol{\omega}$, \mathbf{x} ile aynı boyuta sahip bir satır vektörüdür. İntegral, \mathbf{x} ’in tüm bileşenleri için hesaplanır. \mathbf{x} ’in momentleri, birinci karakteristik fonksiyon $\varphi(\boldsymbol{\omega})$ ’nın Taylor serisi açılımındaki katsayılarından, kümülanları ise ikinci karakteristik fonksiyon $\phi(\boldsymbol{\omega}) = \ln(\varphi(\boldsymbol{\omega}))$ ’nın Taylor serisi açılımındaki katsayılarından bulunur. Çok değişkenli durumda, kümülanlar genellikle çapraz-kümülanlar olarak adlandırılır.

\mathbf{x} raslantı vektörü $\mathbf{0}$ ortalama vektörüne sahip ise ikinci, üçüncü ve dördüncü dereceden kümülanları

$$\begin{aligned} k\ddot{u}m(x_i, x_j) &= E\{x_i x_j\}, \quad k\ddot{u}m(x_i, x_j, x_k) = E\{x_i x_j x_k\}, \\ k\ddot{u}m(x_i, x_j, x_k, x_l) &= E\{x_i x_j x_k x_l\} - E\{x_i x_j\}E\{x_k x_l\} - E\{x_i x_k\}E\{x_j x_l\} - E\{x_i x_l\}E\{x_j x_k\} \end{aligned} \quad (2.72)$$

şeklinde dir. Burada ikinci kümülan, aynı zamanda ikinci dereceden momente de eşit olduğu için x_i ve x_j değişkenleri arasındaki kovaryansı ve korelasyonu vermektedir. Üçüncü kümülan, üçüncü momente eşittir. Dördüncü kümülan ise dördüncü dereceden momentten farklıdır.

Kümülanlar, momentlerin toplamı şeklinde de ifade edilebildiklerinden, momentler ve kümülanlar aynı istatistiksel bilgiyi içerirler. Ancak kümülanların, momentlerden farklı bazı özellikleri de bulunmaktadır. Bu özellikler aşağıda sıralanmıştır:

- \mathbf{x} ve \mathbf{y} aynı boyutlu istatistiksel olarak bağımsız raslantı vektörleri ise $\mathbf{z} = \mathbf{x} + \mathbf{y}$ toplamının kümülanı, \mathbf{x} ve \mathbf{y} ’nin kümülan toplamlarına eşittir. Bu özellik, ikiden fazla bağımsız raslantı vektörleri için de geçerlidir.

- \mathbf{x} raslantı vektörü çok değişkenli normal dağılıma sahip ise üçüncü ve daha yüksek dereceden kümülanlarının hepsi sıfırdır. Bu yüzden yüksek dereceden kümülanlar, raslantı vektörlerinin aynı ortalama vektörü ve kovaryans matrisine sahip normal dağılan raslantı vektörlerinden farklılıklarını ölçer.

Yüksek dereceden momentler ve kümülanlar, ikinci dereceden istatistiklere göre daha fazla örnekleme ihtiyaç duyarlar. Ayrıca verideki aykırı değerlere de daha duyarlıdırlar (Hyvärinen, Karhunen ve Oja, 2001).

2.3.2. Optimizasyon Teorisi ile İlgili Temel Kavramlar

BBA'nın asıl amacı bağımsız bileşenlerin ortaya çıkarılmasını sağlayan ayrıştırma matrisi \mathbf{W} 'yu tahmin etmektir. Ancak bağımsız bileşenlerin tahmininde iki bilinmeyenli bir eşitlik kullanıldığı için \mathbf{W} 'yu doğrudan tahmin etmek mümkün değildir. Bunun yerine ceza (cost, objective, contrast) fonksiyonlarının optimizasyonu (minimize ya da maksimize edilmesi) ile tahmin edilirler. Bu sebeple bu kısımda, bağımsız bileşenler elde edilirken optimizasyon teorisinde karşılaşılabilecek bazı temel kavramlara yer verilmiştir.

2.3.2.1. Vektör ve Matris Gradyanları

Gradyan, çok değişkenli bir fonksiyonun her bir değişkenine göre alınan türevleri ile yazılan bir vektördür. Eşitlik 2.73'teki gibi m tane değişkenden oluşan skaler değerli türevlenebilir g fonksiyonu ele alalım

$$g = g(w_1, \dots, w_m) = g(\mathbf{w}) \quad (2.73)$$

Bu fonksiyonun \mathbf{w} 'ya göre vektör gradyanı,

$$\frac{\partial g}{\partial \mathbf{w}} = \begin{pmatrix} \frac{\partial g}{\partial w_1} \\ \vdots \\ \frac{\partial g}{\partial w_m} \end{pmatrix} \quad (2.74)$$

şekindedir ve ∇g ya da $\nabla_{\mathbf{w}} g$ olarak da gösterimleri mevcuttur. g fonksiyonunun ikinci dereceden gradyanının yazılması da mümkündür.

$$\frac{\partial^2 g}{\partial \mathbf{w}^2} = \begin{pmatrix} \frac{\partial^2 g}{\partial w_1^2} & \dots & \frac{\partial^2 g}{\partial w_1 w_m} \\ \vdots & & \vdots \\ \frac{\partial^2 g}{\partial w_m w_1} & \dots & \frac{\partial^2 g}{\partial w_m^2} \end{pmatrix} \quad (2.75)$$

Eşitlik 2.75 aynı zamanda $g(\mathbf{w})$ fonksiyonunun Hessian matrisi olarak da adlandırılmaktadır. İkinci dereceden kısmi türevlerden oluşan $m \times m$ boyutlu Hessian matrisi her zaman simetriktir. Bahsi geçen kavramlar vektör değerli fonksiyonlar için de genelleştirilebilir. n tane vektörden oluşan türevlenebilir g fonksiyonunun gradyanı Eşitlik 2.77'deki gibidir. Bu eşitlik aynı zamanda $\mathbf{J}g$ şeklinde gösterilen Jakobiyen matris olarak da bilinmektedir ve bu matrisin i 'ninci sütunu, $g_i(\mathbf{w})$ 'nin gradyan vektörüne eşittir (Hyvärinen, Karhunen ve Oja, 2001).

$$g(\mathbf{w}) = \begin{pmatrix} g_1(\mathbf{w}) \\ \vdots \\ g_n(\mathbf{w}) \end{pmatrix} \quad (2.76)$$

$$\frac{\partial g}{\partial \mathbf{w}} = \begin{pmatrix} \frac{\partial g_1}{\partial w_1} & \dots & \frac{\partial g_n}{\partial w_1} \\ \vdots & & \vdots \\ \frac{\partial g_1}{\partial w_m} & \dots & \frac{\partial g_n}{\partial w_m} \end{pmatrix} \quad (2.77)$$

Elemanları, $m \times n$ boyutlu $\mathbf{W} = (w_{ij})$ matrisinden oluşan skaler değerli

$$g = g(\mathbf{W}) = g(w_{11}, \dots, w_{ij}, \dots, w_{mn}) \quad (2.78)$$

şeklindeki g fonksiyonunun \mathbf{W} 'ya göre matris gradyanı,

$$\frac{\partial g}{\partial \mathbf{W}} = \begin{pmatrix} \frac{\partial g}{\partial w_{11}} & \dots & \frac{\partial g}{\partial w_{1n}} \\ \vdots & & \vdots \\ \frac{\partial g}{\partial w_{m1}} & \dots & \frac{\partial g}{\partial w_{mn}} \end{pmatrix} \quad (2.79)$$

biçimindedir. Vektör ve matris gradyanlarına örnek olarak aşağıdaki fonksiyonlar verilebilir:

- $g(\mathbf{w}) = \mathbf{a}^T \mathbf{w}$ ise $\frac{\partial \mathbf{a}^T \mathbf{w}}{\partial \mathbf{w}} = \mathbf{a}$
- $g(\mathbf{w}) = \mathbf{w}^T \mathbf{A} \mathbf{w}$ ise $\frac{\partial \mathbf{w}^T \mathbf{A} \mathbf{w}}{\partial \mathbf{w}} = \mathbf{A} \mathbf{w} + \mathbf{A}^T \mathbf{w}$
- $g(\mathbf{w}) = \mathbf{w}^T \mathbf{A} \mathbf{w}$ ve \mathbf{A} simetrik bir matris ise $\frac{\partial^2 \mathbf{w}^T \mathbf{A} \mathbf{w}}{\partial \mathbf{w}^2} = 2\mathbf{A}$
- $g(\mathbf{W}) = \det \mathbf{W}$ ise $\frac{\partial \det \mathbf{W}}{\partial \mathbf{W}} = \text{adj}(\mathbf{W})^T$
- $g(\mathbf{W}) = \log |\det \mathbf{W}|$ ise $\frac{\partial \log |\det \mathbf{W}|}{\partial \mathbf{W}} = \frac{1}{|\det \mathbf{W}|} \frac{\partial |\det \mathbf{W}|}{\partial \mathbf{W}} = (\mathbf{W}^T)^{-1}$

2.3.2.2. Çok Değişkenli Fonksiyonların Taylor Serisi Açılımları

Skaler değerler alan w değişkeninin $g(w)$ fonksiyonunun Taylor serisi açılımı

$$g(w') = g(w) + \frac{dg}{dw}(w' - w) + \frac{1}{2} \frac{d^2g}{dw^2}(w' - w)^2 + \dots \quad (2.80)$$

şeklinindedir. Benzer bir açılım m tane değişkenin $g(\mathbf{w}) = g(w_1, \dots, w_m)$ fonksiyonu ve matris değişkenli $g(\mathbf{W})$ fonksiyonu için de sırasıyla Eşitlik 2.81 ve Eşitlik 2.82'deki gibi yazılabilir (Hyvärinen, Karhunen ve Oja, 2001):

$$g(\mathbf{w}') = g(\mathbf{w}) + \left(\frac{\partial g}{\partial \mathbf{w}} \right)^T (\mathbf{w}' - \mathbf{w}) + \frac{1}{2} (\mathbf{w}' - \mathbf{w})^T \frac{\partial^2 g}{\partial \mathbf{w}^2} (\mathbf{w}' - \mathbf{w}) + \dots \quad (2.81)$$

$$g(\mathbf{W}') = g(\mathbf{W}) + i_z \left(\frac{\partial g}{\partial \mathbf{W}} \right)^T (\mathbf{W}' - \mathbf{W}) + \dots \quad (2.82)$$

2.3.2.3. Gradyan Azaltma ile Matris Optimizasyonu

Gradyan azaltma için $L(\mathbf{w})$ ceza fonksiyonu ile minimizasyon amacının güdüldüğü bir optimizasyon problemi düşünölsün. $\mathbf{w}(0)$ gibi bir başlangıç noktası için $L(\mathbf{w})$

hesaplanır ve negatif gradyanın doğrultusunda uygun bir uzaklık kadar hareket ettirilir. Ulaşılan her nokta için işlem tekrar edilir ve iteratif olarak $L(\mathbf{w})$ fonksiyonu güncellenerek minimum sonuca ulaşılır. Bahsi geçen güncelleme kuralı $t = 1, 2, \dots$ için

$$\mathbf{w}(t) = \mathbf{w}(t-1) - \alpha(t) \left. \frac{\partial L(\mathbf{w})}{\partial \mathbf{w}} \right|_{\mathbf{w}=\mathbf{w}(t-1)} \quad (2.83)$$

şeklinde dir. Eşitlik 2.83'te $\alpha(t)$ negatif eğim doğrultusundaki uygun uzaklığı göstermekte ve genellikle öğrenme katsayısı olarak adlandırılmaktadır. Eğer $\mathbf{w}(t) - \mathbf{w}(t-1)$ farkı $\Delta \mathbf{w}$ olarak gösterilirse Eşitlik 2.83, Eşitlik 2.84'teki gibi de ifade edilebilir:

$$\Delta \mathbf{w} = -\alpha \frac{\partial L(\mathbf{w})}{\partial \mathbf{w}} \quad (2.84)$$

İterasyon, $\|\mathbf{w}(t) - \mathbf{w}(t-1)\|$ arasındaki Öklit uzaklığı daha önceden belirlenmiş küçük bir tolerans seviyesine yakınsayana kadar devam eder.

Gradyan azaltmasında, öğrenme katsayısının seçimi en önemli unsurdur. Çünkü, $\alpha(t)$ için çok küçük bir değer seçilmesi algoritmanın yavaş yakınsamasına neden olurken, çok büyük bir değer seçilmesi de yakınsamaya engel olmaktadır (Hyvärinen, Karhunen ve Oja, 2001).

2.3.3. Bilgi Kuramı ile İlgili Temel Kavramlar

BBA, farklı disiplinler tarafından geliştirildiği için, farklı yaklaşımlara sahip BBA algoritmaları mevcuttur. Bunlardan bazıları da entropi tabanlıdır. Bu sebeple bu kısımda, ileride BBA algoritmalarından bahsedilirken bilgi kuramı ile ilgili karşılaşılabilecek bazı temel kavramlara yer verilmiştir.

2.3.3.1. Entropi ve Entropi Fonksiyonu

Herhangi bir olayın ortaya çıkma olasılığı, belli seviyede bu olayın gerçekleşip gerçekleşmeyeceği hakkındaki bir belirsizliğin göstergesidir. Başka bir ifadeyle, yapılan

bir deneyde A olayı p olasılığı ile ortaya çıkıyor ise bu deneyde A olayı gerçekleşebilir de, gerçekleşmeyebilir de denilebilir. Bu durumda A olayının gerçekleşmesi hakkında bir belirsizlik söz konusudur. Örneğin iki ya da ikiden fazla duruma sahip bir fiziksel sistem ele alınsın ve bu sistemin verilmiş durumlarda bulunma olasılıkları p_1, p_2, \dots, p_n olsun. Eğer sözü edilen fiziksel sisteme örnek olarak zar atış deneyi ele alınırsa bu sisteme, altı duruma sahip bir fiziksel sistem gibi bakılabilir. Her durumun gerçekleşme olasılığının eşit kabul edilmesiyle sözü edilen olasılık, $1/6$ olur.

Fiziksel sistemler hakkındaki belirsizliğin göstergesi olarak entropi kavramı kullanılmaktadır. Entropi, olasılıklar yardımıyla olayların ortaya çıkma belirsizliğini belli ölçüde ifade eden bir büyüklüktür. Başka bir deyişle, bir olayın gerçekleşme olasılığı o olay için belirsizliğin bir göstergesi iken entropi, fiziksel sistemi oluşturan tüm olaylar için belirsizliğin bir ifadesi olarak kullanılır (Şamilov, 2015).

Entropi ilk olarak termodinamikte enerjiyle ilgili olarak ortaya çıkmıştır. Toplanabilirlik, depolanabilirlik, aktarılabilirlik ve kabul edilebilirlik enerjinin özelliklerindedir. Aynı özelliklerin bilgi kavramına da ait olduğunu fark eden Shannon (1948), bilgiyi belirsizlikle eş tutmuş ve fizik bilimciler tarafından belirsizliğin ölçüsü olarak adlandırılan entropi ile bilgi kuramı arasındaki ilişkiyi görerek entropi kavramını “İletişimin Matematiksel Teorisi” başlıklı çalışmasında kullanmıştır. Shannon (1948) çalışmasında, bir mesajın içerisindeki belirsizliği olasılık kavramı ile ilişkilendirmiş ve bilgi miktarını ölçmüştür.

Sistemin entropisi denildiğinde, tüm durumların beklenen değeri anlaşılır ve x sisteminin entropisi $H(x)$ ile gösterilir.

x kesikli raslantı değişkeni ya da sistemi, x_1, x_2, \dots, x_n gösterilen değerleri $p_i = P(x = x_i)$, $i = 1, 2, \dots, n$ olasılıkları ile alıyor ise x sisteminin iki tabanına göre entropisi,

$$H(x) = -\sum_{i=1}^n p_i \log p_i \quad (2.85)$$

şeklinde. Genellikle entropi fonksiyonunda, 2 tabanında logaritma kullanılmakla birlikte e tabanında logaritmanın (\ln) kullanımı da yaygındır. Entropi birimi, 2 tabanının kullanılması durumunda “binary digit” kavramının kısaltması olan “bit”, e tabanının kullanılması durumunda “natural digits”’in kısaltması olan “nats” olmaktadır (Bursa ve Özel Kadılar, 2016).

Entropi, aynı zamanda beklenen değer cinsinden de

$$H(x) = -E[\log p(x)] = E[-\log p(x)] = E\left[\log \frac{1}{p(x)}\right] \quad (2.86)$$

biçiminde yazılabilir.

Sürekli bir sistem ya da raslantı değişkeni için kullanılan entropi, diferansiyel entropi olarak adlandırılmaktadır. $f(x)$ yoğunluk fonksiyonuna sahip bir x raslantı değişkeni ya da sistemi için diferansiyel entropi,

$$H(x) = -\int f(x) \log f(x) dx = E[-\log f(x)] \quad (2.87)$$

biçiminde tanımlıdır. \mathbf{x} sürekli raslantı vektörü için de diferansiyel entropiyi bulmak mümkündür.

$$H(\mathbf{x}) = -\int f(\mathbf{x}) \log f(\mathbf{x}) d\mathbf{x} \quad (2.88)$$

Eğer \mathbf{x} raslantı vektörü, $\mathbf{y} = f(\mathbf{x})$ şeklinde bir dönüşüme sahip ise dönüşümün entropisi Eşitlik 2.89 ve Eşitlik 2.90’daki gibi yazılabilir.

$$H(\mathbf{y}) = -E(\log f(\mathbf{y})) \quad (2.89)$$

$$H(\mathbf{y}) = H(\mathbf{x}) + E(\log |\det Jf(\mathbf{x})|) \quad (2.90)$$

Örneğin, $\mathbf{y} = \mathbf{W}\mathbf{x}$ şeklinde bir dönüşüm fonksiyonu ise \mathbf{y} ’nin entropisi $H(\mathbf{y}) = H(\mathbf{x}) + \log |\det \mathbf{W}|$ şeklindedir (Hyvärinen, Karhunen ve Oja, 2001).

2.3.3.2. Bileşik ve Koşullu Entropi Fonksiyonları

Bir önceki başlıkta tek bir raslantı değişkeninin ya da sisteminin entropi fonksiyonu tanımlanmıştı. Bu başlıkta ise (x, y) biçimindeki raslantı değişken çifti ya da iki sistem için bileşik entropi kavramı ele alınmıştır. Bileşik sistemin entropi fonksiyonu, kesikli ve sürekli sistemler için sırasıyla,

$$H(x, y) = E[-\log p(x, y)] = -\sum \sum p(x, y) \log p(x, y) \quad (2.91)$$

$$H(x, y) = E[-\log f(x, y)] = -\int \int f(x, y) \log f(x, y) dx dy \quad (2.92)$$

olarak tanımlanmaktadır.

(x, y) gibi bileşik bir sistemde y sisteminin durumu bilindiğinde, x sisteminin koşullu entropisi kesikli ve sürekli sistemler için sırasıyla,

$$H(x|y) = E[-\log p(y|x)] = -\sum \sum p(x, y) \log p(y|x) \quad (2.93)$$

$$H(x|y) = E[-\log f(x, y)] = -\int \int f(x, y) \log f(x, y) dy \quad (2.94)$$

biçiminde yazılabilir.

(x, y) bileşik sisteminde, y sistemi ile x sistemi bağımsız ise bu iki sistemin bileşik entropisi, sistemlerin entropi toplamları olan $H(x, y) = H(x) + H(y)$ 'ye eşit olur. Ayrıca zincir kuralının bir gereği olarak (x, y) bileşik sisteminin entropisi için

$$\begin{aligned} H(x, y) &= H(x) + H(y|x) \\ &= H(y) + H(x|y) \end{aligned} \quad (2.95)$$

yazılabilir⁵ (Şamilov, 2015).

⁵ $H(x|y) \neq H(y|x)$ 'dir. Ancak; $H(x) - H(x|y) = H(y) - H(y|x)$ eşitliği yazılabilir.

2.3.3.3. Negentropi

Entropi, belirsizlik ölçümünün yanı sıra normal dağılmamanın bir ölçüsü olarak da kullanılmaktadır. Sıfır ortalama ve birim varyansa sahip raslantı değişkenleri arasında normal dağılan değişkenin entropisi her zaman maksimumdur. Bu bilgiden hareketle geliştirilen negentropi (negatif entropi) terimi, Σ kovaryans matrisi ile normal dağılan \mathbf{x}_{normal} raslantı vektörü için Eşitlik 2.96'daki gibi hesaplanır:

$$H(\mathbf{x}_{normal}) = \frac{n}{2}(1 + \log 2\pi) + \frac{1}{2} \log |\det \Sigma| \quad (2.96)$$

Negentropi, her zaman pozitif değerler alır ve sadece normal dağılan değişkenler için sıfır değerini alır (Özdamar, 2009).

2.3.3.4. Bilgi Miktarının Entropi ile Ölçülmesi

Fiziksel bir sistemin durumlarının gerçekleşmesinde rasgelelik var ise bu sistem, belli bir entropiye (belirsizliğe) sahiptir. Sistemin sahip olduğu durumların gerçekleşme olasılıklarının değişimi, sistemin entropisinin değişimi anlamına gelmektedir. Entropinin değişimi, yeni bir kavram doğurur. Bu kavram, bilgi (enformasyon) kavramıdır.

n duruma sahip x fiziksel sisteminin ilk entropisi $H(x)$, $p = (p_1, p_2, \dots, p_n)$, bazı etkenlerle durum olasılıklarındaki değişim sonucunda son entropisi ise $H(\tilde{x})$, $\tilde{p} = (\tilde{p}_1, \tilde{p}_2, \dots, \tilde{p}_n)$ olsun.

$$I = H(x) - H(\tilde{x}) \quad (2.97)$$

biçiminde ifade edilen farka, sistem hakkında bilgi denir ve birimi, entropi birimiyle aynıdır.

Görüldüğü gibi bilgi, sistemin entropi değişiminin bir ifadesidir ve $I > 0$, $I < 0$ ve $I = 0$ olabilir.

- $I > 0$ ise bilgi pozitifdir ve sistemin entropisi I kadar düşmüştür.

- $I < 0$ ise bilgi negatiftir ve sistemin entropisi I kadar artmıştır.
- $I = 0$ ise bilgi miktarı sıfırdır ve sistemin entropisi değişmemiştir (Şamilov, 2015).

2.3.3.5. Ortak Bilgi Miktarı

x sisteminin belirsizliğinin giderilmesi için gerekli bilgi miktarının I olduğu bilinmektedir. Ancak bazen x sisteminin doğrudan gözlenememesinden dolayı bu sisteme bağlı olan y sistemini ele almak daha elverişli olabilir. Bu kısımda, y sistemini gözlemlemekle elde edilen bilgi miktarının x sistemi için ne anlama geldiği incelenmiştir.

x sisteminin entropisi $H(x)$ ise bu sistemin belirsizliğinin aşılması için gerekli bilgi miktarı $I = H(x)$ ile ifade edilir. y sistemi dikkate alınmış ise koşullu entropi kavramına göre y sisteminin x hakkında içerdiği entropi miktarı $H(x|y)$ 'dir. Böylece y sisteminin gözlenmesi ile x sisteminin entropisinde gerçekleşen azalma miktarı başka bir deyişle, y sisteminin x sistemi hakkında verdiği bilgi miktarı,

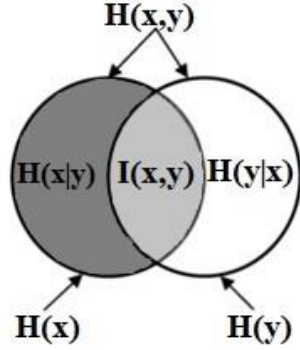
$$I(x,y) = H(x) - H(x|y) \quad (2.98)$$

şeklinde ifade edilir. Eşitlik 2.98, aynı zamanda ortak (mutual) bilgi miktarı olarak da adlandırılmaktadır ve Şekil 2.1'deki gibi gösterilebilir. Ortak bilgi miktarı benzer şekilde, x sisteminin bilinmesinin y sisteminin entropisinde sağladığı azalış miktarı olarak da Eşitlik 2.99'daki gibi yazılabilir.

$$I(x,y) = H(y) - H(y|x) \quad (2.99)$$

Ortak bilgi miktarı değerinin, esasında koşullu entropi değerlerinin büyüyüp küçülmesine bağlı olduğu görülmektedir. Bu nedenle $I(x,y)$ bir bağımlılık ölçüsü ya da iki raslantı değişkeni arasındaki birliktelik derecesinin değeri olarak da kullanılmaktadır. Ancak; $I(x,y)$ 'de korelasyon katsayısında olduğu gibi bir yön kavramı yoktur. Sadece her iki değişkene ilişkin bilgilerin olduğu alanı ifade etmektedir. Bu açıdan korelasyon

katsayısına göre daha geneldir. Çünkü korelasyon katsayısı iki raslantı değişkeni arasındaki doğrusal uzaklıklara bakarken, ortak bilgi miktarı ise doğrusal olmayan uzaklıklara da bakar.



Şekil 2.1 Ortak Bilginin Şekilsel Gösterimi

Ortak bilgi miktarını, x ve y kesikli ya da sürekli raslantı değişkenleri için bileşik olasılık fonksiyonu $p(x, y)$ ya da $f(x, y)$, marjinal olasılık fonksiyonları ise $p(x)$, $p(y)$ ya da $f(x)$, $f(y)$ olmak üzere olasılık yoğunluk fonksiyonları ile de yazmak mümkündür.

$$I(x, y) = E \left[\log \frac{p(x, y)}{p(x)p(y)} \right] = \sum_{x \in \mathbb{R}} \sum_{y \in \mathbb{R}} p(x, y) \log \frac{p(x, y)}{p(x)p(y)} \quad (2.100)$$

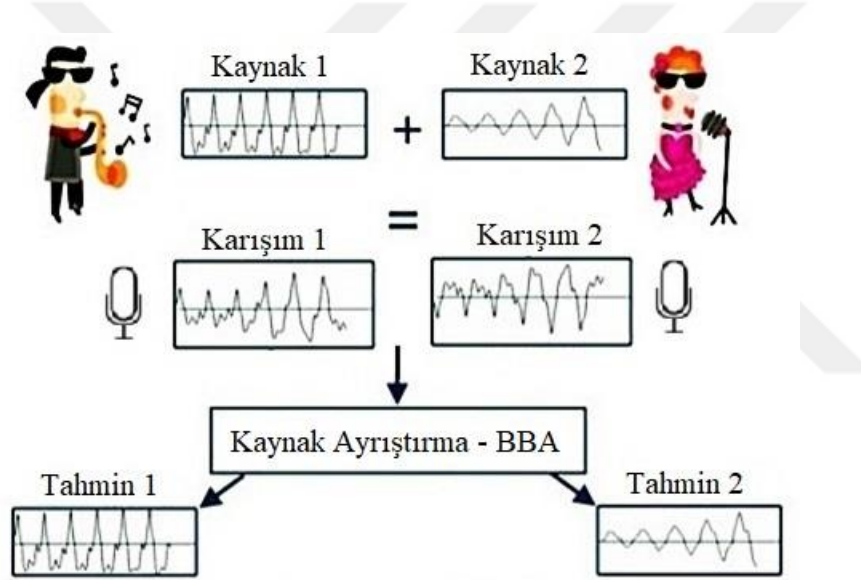
$$I(x, y) = E \left[\log \frac{f(x, y)}{f(x)f(y)} \right] = \int_{x \in \mathbb{R}} \int_{y \in \mathbb{R}} f(x, y) \log \frac{f(x, y)}{f(x)f(y)} dx dy \quad (2.101)$$

$I(x, y) = I(y, x)$ olduğu için ortak bilgi miktarı simetriktir. Ayrıca her zaman pozitifdir (Hyvärinen, Karhunen ve Oja, 2001; Bursa ve Özel Kadılar, 2016).

3. BAĞIMSIZ BİLEŞENLER ANALİZİ

Sinyal işleme, makine öğrenmesi ve diğer mühendislik bilimleri açısından BBA; gözlenen çok değişkenli bir sinyale gizlenmiş, birbirinden bağımsız ve normal dağılmayan kaynak sinyalleri ayırmak için kullanılan bir tekniktir (Hyvärinen ve Oja, 2000). Bu bağlamda BBA en iyi şekilde, kör kaynak ayırtırmaya örnek bir problem olan kokteyl parti problemi ile açıklanmaktadır.

Kokteyl parti probleminde aynı odada bulunan birden çok kişinin seslerinin ayırt edilmesi amaçlanmaktadır. Bunun için iki kişinin aynı odada olduğu ve iki adet mikrofonun da bu kişilerin seslerini kaydettiği düşünülün.



Şekil 3.1 Kokteyl Parti Problemi

Problem, mikrofonların kaydettiği karışık seslerden, her bir kişinin konuşmasını bağımsız bir şekilde ortaya çıkarmaktır. İki sinyal karışımından (mikrofondan) BBA ile her bir kişinin konuşması, başka bir deyişle iki adet orijinal ses ya da kaynak sinyali elde edilebilir. Sadece kaynak sinyaller birbirinden bağımsız olduğunda bu karışımları ayırmak mümkün olduğu için BBA'nın temel varsayımı, kaynak sinyallerin birbirinden bağımsız olduğudur. Karışımlar, kokteyl parti probleminde olduğu gibi seslerden meydana gelebileceği gibi, radyo dalgalarından ya da görüntülerden de meydana gelebilir (Stone, 2004).

İstatistiksel açıdan BBA ise sırasıyla, temel bileşenler ve faktörler olarak adlandırılan daha az sayıdaki gözlenmemiş değişkenler açısından ilişkili değişkenler arasındaki değişkenliği tanımlayan TBA ve FA'nın daha gelişmiş bir hali olarak düşünülebilir. Bu yaklaşımlar, sadece kovaryans ya da korelasyon matrisi gibi ikinci dereceden istatistiklerin analizine dayanırlarken BBA, raslantı değişkenlerinin basıklık, çarpıklık gibi daha yüksek dereceden istatistiklerini kullanır. FA'ya benzer bir şekilde, BBA'da da bir gizli değişken modeli oluşturulur. Modelde gözlenen değişkenlerin, karşılıklı olarak bağımsız ve normal dağılmayan değişkenlerin karışımları oldukları varsayılır (Nordhausen ve Oja, 2018).

3.1. Bağımsız Bileşenler Analizi Modeli

Temel BBA modeli Eşitlik 3.1'deki gibi ifade edilmektedir:

$$x_i = a_{i1}s_1 + a_{i2}s_2 + \dots + a_{in}s_n, \quad i = 1, \dots, n \quad (3.1)$$

Burada, n tane gözlenen x_1, x_2, \dots, x_n raslantı değişkenleri, a_{ij} 'ler $i, j = 1, \dots, n$ gerçek katsayılar olmak üzere, n tane s_1, s_2, \dots, s_n raslantı değişkeninin doğrusal bir kombinasyonu olarak modellenmiştir. Eşitlikte, s_j 'ler bağımsız bileşenler başka bir deyişle, doğrudan gözlemlenemediklerinden dolayı gizli değişkenlerdir ve tanım gereği istatistiksel olarak birbirinden bağımsızdırlar. Ayrıca karışım katsayıları (ya da ağırlıklar) a_{ij} 'ler de bilinmemektedir. BBA'nın amacı, gözlenen x_i raslantı değişkenlerini kullanarak karışım katsayılarını ve bağımsız bileşenleri tahmin etmektir.

BBA modelini vektör-matris gösterimiyle de ifade etmek mümkündür. Eşitlik 3.3'te, \mathbf{x} raslantı vektörünün; x_1, x_2, \dots, x_n karışımlarından, \mathbf{s} raslantı vektörünün; s_1, s_2, \dots, s_n 'lerden oluştuğu görülmektedir. \mathbf{A} matrisi ise a_{ij} 'lerden oluşmaktadır. Eşitlik 3.2'de ise \mathbf{a}_i 'ler \mathbf{A} matrisinin sütunlarına karşılık gelmektedir.

$$\mathbf{x} = \sum_{i=1}^n \mathbf{a}_i s_i \quad (3.2)$$

$$\mathbf{x} = \mathbf{A}\mathbf{s} \quad (3.3)$$

p tane \mathbf{x}^T raslantı vektörünün bulunduğu durum için BBA modeli de Eşitlik 3.4'teki gibi ifade edilir:

$$\mathbf{X}_{p \times n} = \mathbf{A}_{p \times p} \mathbf{S}_{p \times n} \quad (3.4)$$

Görüldüğü üzere temel BBA modelinde hata terimi bulunmamaktadır. Bu yüzden bazı araştırmacılar tarafından temel BBA modeli, gürültüsüz BBA olarak da adlandırılmaktadır. Hata terimi kullanılmayan BBA modelinin çoğu uygulama için yeterli olduğu görülmüştür (Hyvärinen, Karhunen ve Oja, 2001).

Eşitlik 3.4'te yer alan BBA modelinde \mathbf{A} ve \mathbf{S} matrisleri, yalnızca \mathbf{X} matrisi kullanılarak tahmin edilir. Bunun için öncelikle, karışım matrisinin tersinin (\mathbf{A}^{-1}) tahmini olan ve ayırıştırma matrisi olarak adlandırılan \mathbf{W} matrisi bulunur. Sonrasında, \mathbf{W} matrisi kullanılarak \mathbf{U} matrisi elde edilir. \mathbf{U} ; $k \leq p$ olmak üzere $k \times n$ boyutlu \mathbf{S} matrisinin tahmini başka bir ifadeyle, bağımsız bileşenler matrisinin tahminidir.

$$\mathbf{U} = \mathbf{W}\mathbf{X} \quad (3.5)$$

Karışım matrisinin ve bağımsız bileşenlerin elde edilmesinde normal olmamanın maksimizasyonu, en çok olabilirlik yöntemi, ortak bilgi miktarının minimizasyonu, tensörsel yöntemler, doğrusal olmayan ilişkisizleştirme ve doğrusal olmayan TBA gibi çok farklı tahmin yöntemleri kullanılmaktadır.

3.1.1. Bağımsız Bileşenler Analizinin Varsayımları ve Kısıtları

BBA modelinin tahmin edilebilmesi için bazı varsayım ve kısıtların sağlanması gerekmektedir. Bunlar aşağıda sıralanmıştır:

- *Bağımsız bileşenlerin birbirlerinden istatistiksel olarak bağımsız oldukları varsayılır.*

Başka bir deyişle bağımsız bileşenlerin bileşik olasılık yoğunluk fonksiyonu, marjinal olasılık yoğunluk fonksiyonlarının çarpımı olan $f(\mathbf{s}) = f(s_1, s_2, \dots, s_p) = \prod_j f(s_j)$

şeklinde yazılabilmelidir. Bu, modelin tahmin edilebilmesi için en temel varsayımdır. Bazı uygulamalar için bağımsızlığın fiziksel bir yorumu da kullanılabilir. Eğer bileşenler fiziksel olarak ayrı ve etkileşime girmeyen nesnelere tarafından yaratılıyorsa, istatistiksel olarak bağımsız kabul edilebilirler (Hyvärinen, 2013).

- *Bağımsız bileşenlerin dağılımı normal değildir.*

BBA'yı diğer klasik yöntemlerden ayıran en önemli varsayımdır. BBA modelinin tahmin edilebilmesi için yüksek dereceden bilgilere ihtiyaç vardır. Ancak bilindiği üzere, normal dağılımın yüksek dereceden kümülanları (basıklık, çarpıklık gibi) sifra eşittir ve bu yüzden normal dağılan değişkenler için BBA'yı uygulamak mümkün olamamaktadır. Bu durumun ispatına aşağıda yer verilmiştir:

İSPAT: Beyaz bir veri kümesinden elde edilen iki bağımsız bileşenin normal dağıldığı varsayalım. Bu durumda, olasılık yoğunluk fonksiyonları Eşitlik 3.6'daki gibidir:

$$f(s_1, s_2) = \frac{1}{2\pi} \exp\left(-\frac{s_1^2 + s_2^2}{2}\right) = \frac{1}{2\pi} \exp\left(-\frac{\|s\|^2}{2}\right) \quad (3.6)$$

Veri kümesi beyaz olduğu için aynı zamanda \mathbf{A} matrisi ortogondur ve $\mathbf{A}^{-1} = \mathbf{A}^T$ olduğu bilinmektedir. Bu özellik göz önünde bulundurularak x_1 ve x_2 'ye göre Eşitlik 3.6

$$f(x_1, x_2) = \frac{1}{2\pi} \exp\left(-\frac{\|\mathbf{A}^T \mathbf{x}\|^2}{2}\right) |\det \mathbf{A}^T| \quad (3.7)$$

biçiminde yeniden yazılabilir. \mathbf{A} matrisi ortogondur olduğu için aynı zamanda $\|\mathbf{A}^T \mathbf{x}\|^2 = \|\mathbf{x}\|^2$ ve $|\det \mathbf{A}| = 1$ 'dir. Böylece Eşitlik 3.7, Eşitlik 3.8'e dönüşür.

$$f(x_1, x_2) = \frac{1}{2\pi} \exp\left(-\frac{\|\mathbf{x}\|^2}{2}\right) \quad (3.8)$$

Eşitlik 3.6 ve Eşitlik 3.8'den bağımsız bileşenlerin dağılımıyla, gözlenen değişkenlerin (karışımların) dağılımlarının aynı olduğu görülmektedir. Bu sebeple, karışım matrisini belirlemek mümkün değildir. Kısacası \mathbf{A} matrisi, normal dağılan bağımsız bileşenler için belirlenmemektedir.

- *Karışım matrisinin, tersi alınabilir kare bir matris olduğu varsayılır.*

Bu varsayım şart olmamakla birlikte, genellikle hesaplama kolaylıkları açısından tercih edilmektedir (Hyvärinen, Karhunen ve Oja, 2001).

3.1.2. Bağımsız Bileşenler Analizinin Belirsizlikleri

- *Bağımsız bileşenlerin varyansları hesaplanamaz.*

Bunun nedeni, modelde hem \mathbf{S} 'nin hem de \mathbf{A} 'nın bilinmemesidir. \mathbf{A} karışım matrisi bilinmediği için bağımsız bileşen, Eşitlik 3.9'daki gibi herhangi bir α_i katsayısıyla çarpıldığında bu katsayı, \mathbf{A} 'nın \mathbf{a}_i sütunuyla etkisiz hale gelecektir. Bu nedenle, bağımsız bileşenlerin işaretleri de belirsizdir.

$$\mathbf{x} = \sum \left(\frac{1}{\alpha_i} \mathbf{a}_i \right) (s_i \alpha_i) \quad (3.9)$$

- *Bağımsız bileşenlerin önem sırası belirlenemez.*

Bunun nedeni de yine modelin iki bilinmeyenli olmasıdır. Elde edilen herhangi bir bileşen, birinci bileşen olarak kabul edilebilir (Hyvärinen ve Oja, 2000).

3.2. Bağımsız Bileşenlerin Elde Edilmesinde Kullanılan Yaklaşımlar ve Algoritmalar

Çoğu BBA algoritması modelin tahminini, 'önişleme' ve 'tahmin' olmak üzere iki aşamada gerçekleştirir. Önişleme aşamasında veri üzerinde bazı dönüşümler uygulanır ve böylece hesaplama kolaylığı sağlanarak zamandan tasarruf sağlanır. Tahmin aşamasında ise hazır hale getirilmiş veriden, seçilen optimizasyon yöntemlerinden biri ile bağımsız bileşenler ortaya çıkarılır (Tharwat, 2018).

3.2.1. Bağımsız Bileşenler Analizi Öncesinde Kullanılacak Verinin Hazırlanması

3.2.1.1. Verinin Merkezileştirilmesi

BBA için kullanılacak algoritmada kolaylık açısından gözlenen (karışım) değerlerin ve bağımsız bileşenlerin sıfır ortalamaya sahip olması istenir. Bu özellik sağlanmıyor ise bazı önişlemler ile sağlatılabilir. Bunun için BBA uygulanmadan önce gözlenen değerler, Eşitlik 3.10 ile ortalamalarından farkları alınarak 0 ortalamalı hale getirilirler.

$$\mathbf{x} = \mathbf{x} - E\{\mathbf{x}\} \quad (3.10)$$

Böylece bağımsız bileşenler de 0 ortalamalı hale getirilmiş olur. Bu önişlem sonrası karışım matrisi ise aynı kalır.

$$E\{\mathbf{s}\} = E\{\mathbf{A}^{-1}\mathbf{x}\} = \mathbf{A}^{-1} E\{\mathbf{x}\} \quad (3.11)$$

İstenirse 0 ortalamalı veriden karışım matrisi ve bağımsız bileşenler tahmin edildikten sonra, elde edilen 0 ortalamalı bağımsız bileşenlere $\mathbf{A}^{-1} E\{\mathbf{x}\}$ teriminin eklenmesiyle orijinal bağımsız bileşenlere geçiş yapılabilir.

3.2.1.2. Verinin Beyazlatılması

Beyazlatma işlemi, verileri ilişkisiz ve birim varyanslı hale getirmek için gerçekleştirilir. 0 ortalamalı bir raslantı vektörü (\mathbf{x}) eğer beyaz ise, bileşenleri ilişkisiz ve kovaryans matrisi Σ , birim matrise (\mathbf{I}) eşittir.

$$\Sigma = E\{\mathbf{xx}^T\} = \mathbf{I} \quad (3.12)$$

Beyaz olmayan verilerin, dönüşüm matrisi kullanılarak beyazlatılması ve böylece birinci ve ikinci derece istatistiklerin etkisinden arındırılması mümkündür. Örneğin \mathbf{x} gözlenen vektörü, doğrusal bir dönüşümle beyaz \mathbf{z} vektörüne dönüştürülebilir.

$$\mathbf{z} = \mathbf{V}\mathbf{x} \quad (3.13)$$

Beyazlatma için çok farklı dönüşümler mevcuttur. Bunlardan en çok kullanılanı kovaryans matrisinin özdeğerlerinin ayrıştırılmasıdır. Burada Λ , $E\{\mathbf{xx}^T\}$

özdeğerlerinin ortogonal matrisi, $\mathbf{D} = \text{köş}(d_1, \dots, d_n)$ ise özdeğerlerden oluşan köşegen bir matristir.

$$E\{\mathbf{xx}^T\} = \mathbf{\Lambda D \Lambda}^T \quad (3.14)$$

Eşitlik 3.14 doğrultusunda beyazlatma matrisi \mathbf{V} , Eşitlik 3.15'deki gibi oluşturulabilir.

$$\mathbf{V} = \mathbf{\Lambda D}^{-1/2} \mathbf{\Lambda}^T \quad (3.15)$$

Beyazlatma için özdeğerlerin ayrıştırılmasından başka bir diğer yol ise TBA'nın kullanılmasıdır.

BBA modelindeki veri beyazlatıldıktan sonra

$$\mathbf{z} = \mathbf{Vx} = \mathbf{VAs} = \tilde{\mathbf{A}}\mathbf{s} \quad (3.16)$$

şeklinde ifade edilebilir.

Beyazlatma sonucu elde edilen yeni $\tilde{\mathbf{A}} = \mathbf{VA}$ karışım matrisi ortogondur. Ortogonallik özelliđi, Eşitlik 3.17'de görölmektedir.

$$E\{\mathbf{zz}^T\} = \tilde{\mathbf{A}} E\{\mathbf{ss}^T\} \tilde{\mathbf{A}}^T = \tilde{\mathbf{A}} \tilde{\mathbf{A}}^T = \mathbf{I} \quad (3.17)$$

Bu özellik sayesinde, karışım matrisi aranırken sadece ortogondal matris uzayı araştırılacaktır. Başka bir deyişle, orijinal \mathbf{A} matrisinin n^2 tane elemanın tahmin edilmesi yerine, ortogondal $\tilde{\mathbf{A}}$ matrisinin $n(n-1)/2$ tane elemanın belirlenmesi yeterli olacaktır. Böylece BBA algoritmaları için problemin karmaşıklığı da azalacaktır (Hyvärinen, Karhunen ve Oja, 2001).

3.2.2. Bağımsız Bileşenlerin Tahmini

BBA'da amaç, öncelikle karışım matrisinin tersi olan ayrıştırma matrisini (\mathbf{W}) tahmin etmek ve sonrasında bu matrisi beyazlatılmış veriler üzerine uygulayarak bağımsız bileşenleri elde etmektir.

Verilerin önışleme aşaması tamamlandıktan sonra \mathbf{W} matrisi, üç farklı bağımsızlık yaklaşımı kullanılarak tahmin edilebilir. Bu yaklaşımlardan ilki, normallik varsayımına

dayanır ve basıklık ya da negentropi değerlerini kullanarak bağımsız bileşenleri, normal olmamalarını maksimum yapacak şekilde belirler. İkinci yaklaşımda, karşılıklı bilginin minimize edilmesi hedeflenirken, son yaklaşımda ise en çok olabilirlik tahmini kullanılmaktadır. Bağımsız bileşenler elde edilirken hangi yaklaşımın kullanılacağı belirlendikten sonra, hangi algoritma ile yaklaşımın optimizasyonunun sağlanacağına karar verilir.

3.2.2.1. Normal Olmamanın Ölçümü

Bağımsız bileşenler, normal olmamaları maksimize edilerek ortaya çıkarılabilir. Değişkenlerin normallikten ne kadar uzak olduklarını ölçmek için basıklık ve negentropi değerleri kullanılır.

- **Basıklık**

Ayrırma fonksiyonu, elde edilecek bağımsız bileşenlerin basıklığını maksimum yapacak şekilde belirlenir.

Örnek olarak, $\mathbf{X} = \mathbf{AS}$ şeklinde gösterilen BBA modelindeki \mathbf{s}_1 ve \mathbf{s}_2 olmak üzere iki bağımsız bileşen ortaya çıkarılacak olsun ve \mathbf{s}_1 , \mathbf{s}_2 ve \mathbf{S} 'nin aynı zamanda birim varyansa sahip oldukları varsayalım. Bağımsız bileşenlerin basıklığı, basıklık katsayısının toplama özelliği kullanılarak Eşitlik 3.19'daki gibi yazılabilir

$$\mathbf{U} = \mathbf{WX} = \mathbf{WAS} = \mathbf{QS} = \mathbf{q}_1\mathbf{s}_1 + \mathbf{q}_2\mathbf{s}_2 \quad (3.18)$$

$$bas(\mathbf{S}) = bas(\mathbf{q}_1\mathbf{s}_1) + bas(\mathbf{q}_2\mathbf{s}_2) = \mathbf{q}_1^4 bas(\mathbf{s}_1) + \mathbf{q}_2^4 bas(\mathbf{s}_2) \quad (3.19)$$

Ayrıca bağımsız bileşenler birim varyanslı olduğu için \mathbf{q}_1 ve \mathbf{q}_2 için $E(\mathbf{U}^2) = \mathbf{q}_1^2 + \mathbf{q}_2^2 = 1$ kısıtı mevcuttur. Bu kısıt, \mathbf{Q} 'nun iki boyutlu uzayda birim çember olması gerektiği anlamına gelmektedir. Böylece BBA'nın amacı birim çember üzerinde $bas(\mathbf{S}) = \mathbf{q}_1^4 bas(\mathbf{s}_1) + \mathbf{q}_2^4 bas(\mathbf{s}_2)$ eşitliğini maksimize etmeye dönüşür. Optimal, yani maksimum çözüm, \mathbf{Q} 'dan biri 0, diğeri de -1 ya da +1 olduğunda elde edilir ve bu optimal çözümler, $\pm\mathbf{s}_i$ bağımsız bileşenlerini verir.

Veri matrisi beyaz bir matris (\mathbf{Z}) ise bu kez bağımsız bileşenler, $\mathbf{U} = \mathbf{WZ}$ eşitliğindeki bağımsız bileşenlerin basıklığı maksimize edilerek elde edilirler. Bu durumda \mathbf{W} ve \mathbf{Z} birim varyanslı oldukları için bağımsız bileşenlerin basıklıkları $bas(\mathbf{U}) = E\{(\mathbf{WZ})^4\} - 3$ şeklinde olacaktır (Hyvärinen, Karhunen ve Oja, 2001; Tharwat, 2018).

Basıklığın maksimize edilmesi için kullanılan en temel algoritma, gradyan algoritmasıdır.

\mathbf{U} 'nun basıklığının gradyanı $\frac{\partial |bas(\mathbf{W}^T \mathbf{Z})|}{\partial \mathbf{W}} = cE\{\mathbf{Z}(\mathbf{W}^T \mathbf{Z})^3\}$ 'dır ve c burada bir sabittir.

Her iterasyonda, başlangıçta rasgele olarak belirlenmiş olan ağırlık vektörü $\mathbf{w}_{yeni} = \mathbf{w}_{eski} + \eta E(\mathbf{Z}(\mathbf{w}_{eski} \mathbf{Z}))$ şeklinde yenilenir. Burada η öğrenme katsayısıdır. Basıklık, birim çember üzerinde optimize edildiği için elde edilen ağırlık vektörü sonrasında normuna bölünerek $\mathbf{w}_{yeni}^n = \mathbf{w}_{yeni} / \|\mathbf{w}_{yeni}\|$ güncellenir. Algoritma, yakınsama sağlanana kadar çalışmaya devam eder (Tharwat, 2018).

- **Negentropi**

Basıklığın aykırı değerlere karşı duyarlı olmasından dolayı çoğu zaman normal olmamanın bir ölçüsü olarak basıklık yerine hesaplanması daha karmaşık ve güç olsa da negentropi tercih edilmektedir. Negentropinin hesaplanmasındaki zorluğun altında yatan en önemli sebep, olasılık yoğunluk fonksiyonunun çoğu zaman parametrik olmayan tahmininin bulunması gerekliliğidir. Bu nedenle yaklaşık negentropi değerini hesaplamak için yüksek dereceden kümülanların kullanıldığı farklı yaklaşımlar ortaya konmuştur.

$$J(u) \approx \frac{1}{12} E(u^3)^2 + \frac{1}{48} bas(u)^2 \quad (3.20)$$

Eşitlik 3.20'de u raslantı değişkeninin 0 ortalamalı ve birim varyanslı olduğu varsayılmaktadır. Ancak bu eşitlikte de basıklık bulunduğu için aykırı değerlerden etkilenme olabileceği için, Hyvärinen (1999) tarafından maksimum entropi temeline dayanan başka bir yaklaşım önerilmiştir:

$$J(u) \approx \sum_{i=1}^p k_i (E(G_i(u))) - E(G_i(v))^2 \quad (3.21)$$

Eşitlik 3.21’de k_i ; pozitif sabitleri, ν ; sıfır ortalama ve birim varyans ile normal dağılan bir raslantı değişkenini ve G_i ise kuadratik fonksiyonları temsil eder. G_i fonksiyonu olarak $1 \leq a_1 \leq 2$ (a_1 , genellikle 1’e eşittir) olacak şekilde genellikle Eşitlik 3.22’deki fonksiyonlar seçilir (Tharwat, 2018).

$$G_1(u) = \frac{1}{a_1} \log \cosh a_1 u, \quad G_2(u) = -\exp(-u^2/2) \quad (3.22)$$

Basıklıkta olduğu gibi negentropinin maksimize edilmesi için de gradyan algoritması kullanılarak Eşitlik 3.21 maksimize edilebilir.

3.2.2.2. Ortak Bilgi Miktarının Minimizasyonu

Bağımsız bileşenler arasındaki karşılıklı ortak bilgiyi minimize etmek, BBA tahmini için iyi bilinen yaklaşımlardan biridir. Daha önce ikinci bölümde bahsedildiği gibi bilgi kuramına ait bir kavram olan ortak bilgi miktarı, raslantı değişkenleri arasındaki bağımlılığı ölçer. BBA’da, farklı bileşenler arasındaki ortak bilgi minimuma indirilerek bağımsız bileşenler elde edilebilir.

$\mathbf{U} = \mathbf{W}\mathbf{X}$ ve $H(\mathbf{U}) = H(\mathbf{X}) + \log|\mathbf{W}|$ olan BBA modelinde bağımsız bileşenler arasındaki ortak bilgi miktarı, Eşitlik 3.23’teki gibi bulunur:

$$\begin{aligned} I(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_p) &= \sum_{i=1}^p H(\mathbf{u}_i) - H(\mathbf{U}) \\ &= \sum_{i=1}^p H(\mathbf{u}_i) - H(\mathbf{X}) - \log|\det \mathbf{W}| \end{aligned} \quad (3.23)$$

Bu eşitlikte \mathbf{U} beyaz bir matris olduğunda, $E(\mathbf{U}\mathbf{U}^T) = \mathbf{W}E(\mathbf{X}\mathbf{X}^T)\mathbf{W}^T = \mathbf{I}$ özelliğinden dolayı $\det \mathbf{W}E(\mathbf{X}\mathbf{X}^T)\mathbf{W}^T = (\det \mathbf{W})(\det E(\mathbf{X}\mathbf{X}^T))(\det \mathbf{W}^T) = \det \mathbf{I} = 1$ eşitliği geçerli olacaktır. Bunun sonucu olarak da $(\det \mathbf{W})$ bir sabit olacağı için ortak bilgi miktarı

$$I(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_p) = C - \sum_i J(\mathbf{u}_i) \quad (3.24)$$

şeklinde negentropi cinsinden yazılabilecektir. Burada C sabiti, \mathbf{W} ’ya bağlı değildir. Eşitlik 3.24’ün minimize edilmesiyle bağımsız bileşenler elde edilebilir. Ayrıca bu

eşitlikten negentropinin maksimize edilmesinin, ortak bilgi miktarının minimize edilmesi ile ilişkili olduğu da açıkça görülmektedir (Hyvärinen, Karhunen ve Oja, 2001).

Normal olmamanın maksimize edildiği yöntemlerde tahmin edilen bağımsız bileşenler ilişkisiz hale getirilmeye zorlanırken, ortak bilgi miktarının minimize edilmesinde ise bu kısıtlama gerekli değildir.

3.2.2.3. En Çok Olabilirlik Tahmini

İstatistikte en çok kullanılan tahmin yöntemlerinden biri olan bu yöntem, BBA'da ortaya çıkarılan bağımsız bileşenler (\mathbf{U}) için en uygun olan ayrıştırma matrisini (\mathbf{W}) tahmin etmek için kullanılır. $\mathbf{X} = \mathbf{AS}$ şeklindeki BBA modeli için \mathbf{X} 'in olasılık yoğunluk

fonksiyonu $f_{\mathbf{X}}(\mathbf{X}) = \frac{f_{\mathbf{S}}(\mathbf{S})}{|\det \mathbf{A}|} = |\det \mathbf{W}| f_{\mathbf{S}}(\mathbf{S})$ şeklindedir. Bağımsız bileşenlerin olasılık

yoğunluk fonksiyonu $f_{\mathbf{S}}(\mathbf{S}) = f_1(\mathbf{s}_1) f_2(\mathbf{s}_2) \dots f_p(\mathbf{s}_p) = \prod_i f_i(\mathbf{s}_i)$ şeklinde olduğu için

\mathbf{X} 'in olasılık yoğunluk fonksiyonu aynı zamanda Eşitlik 3.25'teki gibi de yazılabilir.

$$f_{\mathbf{X}}(\mathbf{X}) = |\det \mathbf{W}| \prod_i f_i(\mathbf{s}_i) = |\det \mathbf{W}| \prod_i f_i(\mathbf{w}_i \mathbf{X}) \quad (3.25)$$

\mathbf{X} 'in n gözlemden oluştuğu durumda \mathbf{W} 'nin $L(\mathbf{W})$ olabilirlik ve log-olabilirlik fonksiyonları ise

$$L(\mathbf{W}) = \prod_i \prod_j |\det \mathbf{W}| f_j(\mathbf{w}_j \mathbf{x}(i)) \quad (3.26)$$

$$\ln L(\mathbf{W}) = \sum_{j=1}^p \ln f_j(\mathbf{w}_j \mathbf{x}(i)) + n \ln |\det \mathbf{W}| \quad (3.27)$$

şeklindedir. Eşitlik 3.27'yi maksimize eden \mathbf{W} değerleri kullanılarak bağımsız bileşenler tahmin edilir.

Log-olabilirlik fonksiyonu, \mathbf{W} 'nun bir fonksiyonu olarak yazılabileceği gibi bağımsız bileşenlerin bir fonksiyonu olarak da yazılabilir. Ancak çok daha karmaşık olduğu için bu yönteme yer verilmemiştir (Tharwat, 2018).

Herhangi bir x raslantı değişkeninin ortalaması $E(x) = \frac{1}{n} \sum_{i=1}^n x_i \Rightarrow \sum_{i=1}^n x_i = nE(x)$ olduğundan dolayı Eşitlik 3.27, Eşitlik 3.28'deki gibi de yazılabilir:

$$\frac{1}{n} \ln L(\mathbf{W}) = E \sum_{j=1}^p \ln f_j(\mathbf{w}_j \mathbf{X}) + \ln |\det \mathbf{W}| \quad (3.28)$$

Eşitliğin ilk terimi aynı zamanda $E \sum_{j=1}^p \ln f_j(\mathbf{w}_j \mathbf{X}) = -\sum_{j=1}^p H(\mathbf{w}_j \mathbf{X})$ olduğu için olabilirlik fonksiyon değeri ve ortak bilgi miktarı yaklaşık olarak birbirine eşittir. Aralarında işaret ve sabit terim kadar fark vardır. Bu nedenle en çok olabilirlik ile ortak bilgi miktarının kullanıldığı tahmin yaklaşımları aynı bağımsız bileşenleri verirler. Her ikisi için de normal olmamanın maksimize edildiği yöntemlerdeki gibi gradyan algoritması kullanılabilir. Ancak normal olmamayı maksimize eden yöntemlerle (negentropi ve basıklık) bağımsız bileşenler tek tek tahmin edilebilirken, bu durum karşılıklı bilgi miktarının ya da en çok olabilirlik fonksiyonunun kullanıldığı durumlar için geçerli değildir (Tharwat, 2018).

3.2.3. FastICA Algoritması

Hyvärinen (1999) tarafından geliştirilen hızlı sabit-nokta algoritması ya da FastICA algoritması, BBA modelindeki bağımsız bileşenleri ortaya çıkarır. FastICA, normal olmamanın maksimize edildiği, ortak bilgi miktarının minimize edildiği veya en çok olabilirlik tahmin edicisinin kullanıldığı tahmin yöntemlerinin hepsinin optimizasyonunda kullanılabilir.

Bağımsız bileşenlerin elde edilmesinde FastICA algoritması dışında gecikmeli korelasyon matrisi gibi ikinci dereceden istatistikleri kullanan SOBI (Second-Order Blind Identification-İkinci Dereceden Kör Belirleme), dördüncü dereceden istatistikleri kullanan FOBI (Fourth-Order Blind Identification - Dördüncü Dereceden Kör Belirleme),

entropinin maksimizasyonuna dayanan Infomax, ortak köşegenleştirmeyi⁶ temel alan JADE (Joint Approximation Diagonalization of Eigenmatrices) algoritmaları da yaygın olarak kullanılmaktadır (Nordhausen ve Oja, 2018). Literatürde bahsi geçen BBA algoritmalarını, etkinlik ve hız açısından karşılaştıran çalışmalar mevcuttur. Örnek olarak bkz. Matthias, Hauesien ve Ivanova, 2009; Naik, 2011; Dharmapranı ve ark., 2016; Sahonero-Alvarez ve Calderon; 2017.

Tez çalışmasında, uygulamalarda en çok tercih edilen algoritmaların başında geldiği için FastICA algoritması kullanılmıştır. Diğer algoritmaların ve FastICA'nın çeşitli varyasyonlarının uygulamaları için hem MATLAB, hem de R ortamında kullanılacak paketler bulunmaktadır. Örnek olarak R (R Core Team, 2019) için bkz. JADE paketi (Miettinen, Nordhausen ve Taskinen, 2017), fICA paketi (Miettinen ve ark., 2017), ica paketi⁷ (Helwig, 2015), fastICA paketi (Marchini, Heaton ve Ripley, 2017).

FastICA, en azından karesel yakınsama hızına sahip olduğu için gradyan temelli algoritmalara göre daha hızlıdır. Ayrıca algoritmada öğrenme katsayısı gibi önceden belirlenmesi gereken değerler olmadığı için uygulanması gayet kolaydır. FastICA, bağımsız bileşenlerin hepsini teker teker (deflasyon yaklaşımı) veya aynı anda (simetrik yaklaşım) tahmin edebilir. Bunu yaparken genellikle negentropi değerlerini kullanır (Naik ve Kumar, 2011). Çalışmanın uygulama kısmında da negentropiyi temel alan FastICA kullanıldığı için bu bölümde sadece negentropi temelli FastICA'nın algoritma adımlarına yer verilmiştir.

3.2.3.1. Bağımsız Bileşenlerin Teker Teker Bulunmasına İlişkin Negentropi Temelli FastICA Algoritmasının Adımları

Eşitlik 3.22'de de gösterilen G fonksiyonunun türevleri $1 \leq a_1 \leq 2$ (a_1 , genellikle 1'e eşittir) olacak şekilde Eşitlik 3.29 ve Eşitlik 3.30'daki gibi g ile gösterilir ve bu fonksiyonlardan biri seçilir.

⁶ Bir kare matris kümesinin ortak köşegenleştirilmesi, matrisleri mümkün olduğunca köşegen yapan temel ortonormal değişimin bulunmasından oluşur.

⁷ Aynı zamanda çalışmanın uygulama kısımlarında bağımsız bileşenlerin elde edilmesinde kullanılan pakettir.

$$g_2(u) = \tanh a_1 u \quad g'_1(u) = a_1 (1 - \tanh a_1 u) \quad (3.29)$$

$$g_2(u) = u \exp(-u^2/2) \quad g'_2(u) = (1 - u^2) \exp(-u^2/2) \quad (3.30)$$

Buna göre algoritma adımları aşağıdaki gibidir (Hyvärinen, Karhunen ve Oja, 2001):

- 1) Veri kümesini önce merkezileştirip daha sonra beyazlatarak \mathbf{z} 'yi elde et.
- 2) Tahmin edilecek bağımsız bileşenlerin sayısını (m) belirleyip, p 'yi 1'e eşitle.
- 3) Rasgele olarak birim normlu bir \mathbf{w}_p başlangıç vektörü seç.
- 4) Eşitlik 3.29 ya da Eşitlik 3.30'u kullanarak $\mathbf{w}_p = E(\mathbf{z}g(\mathbf{w}_p \mathbf{z})) - E(g'(\mathbf{w}_p \mathbf{z}))\mathbf{w}_p$ hesapla.
- 5) $\mathbf{w}_p^o = \mathbf{w}_p - \sum_{j=1}^{p-1} (\mathbf{w}_p \mathbf{w}_j) \mathbf{w}_j$ ortogonalleştirmesini yap.
- 6) $\mathbf{w}_p^n = \mathbf{w}_p^o / \|\mathbf{w}_p^o\|$ normlaştırma işlemini yap.
- 7) Yakınsama sağlanmaz ise 4. adıma geri dön.
- 8) p 'yi $p+1$ 'e eşitle ve $p \leq m$ ise 3. adıma dön.

3.2.3.2. Aynı Anda Birden Fazla Bağımsız Bileşenlerin Bulunmasına İlişkin Negentropi Temelli FastICA Algoritması

Birden fazla bağımsız bileşeni aynı anda tahmin etmek için $\mathbf{w}_1, \dots, \mathbf{w}_n$ vektörleriyle daha önceki başlıkta verilen algoritmanın birçok kez çalıştırılması ve her iterasyondan sonra $\mathbf{w}_1, \dots, \mathbf{w}_n$ vektörlerinin ortogonalleştirilmesi gerekmektedir. Vektörlerin ortogonalleştirilmesinin amacı, farklı vektörlerin aynı maksimum noktaya yakınsamasını engellemektir.

Algoritma adımları aşağıdaki gibidir (Hyvärinen, Karhunen ve Oja, 2001):

- 1) Veri kümesini önce merkezileştirip daha sonra beyazlatarak \mathbf{z} 'yi elde et.
- 2) Tahmin edilecek bağımsız bileşenlerin sayısını (m) belirle.

- 3) Rasgele olarak birim normlu vektörlerden oluşan bir $\mathbf{W} = (\mathbf{w}_1, \dots, \mathbf{w}_m)^T$ başlangıç matrisini seç.
- 4) Eşitlik 3.29 ya da Eşitlik 3.30'u kullanarak $\mathbf{w}_i = E(\mathbf{z}g(\mathbf{w}_i\mathbf{z})) - E(g'(\mathbf{w}_i\mathbf{z}))\mathbf{w}_i$ $i = 1, \dots, m$ hesapla.
- 5) $\mathbf{W}^o = (\mathbf{W}\mathbf{W}^T)^{-1/2}\mathbf{W}$ ortogonalleştirmesini yap.
- 6) $\mathbf{W}^n = \mathbf{W}^o / \|\mathbf{W}^o\|$ normlaştırma işlemini yap.
- 7) Yakınsama sağlanamaz ise 4. adıma geri dön.

3.2.4. Bağımsız Bileşenler Analizi ile Temel Bileşenler Analizi'nin Karşılaştırılması ve Bir Uygulama

Bölüm 3'e başlarken BBA'nın, istatistik alanında yaygın olarak kullanılan TBA'nın gelişmiş bir hali olarak görülebileceğinden bahsedilmişti. Ancak şimdiye kadar anlatılan kısımlardan da anlaşılacağı üzere her ne kadar analizler arasında benzerlikler bulunsa da, birbirlerinden çok farklı yönleri de mevcuttur. Bu açıdan konunun daha iyi anlaşılabilmesi için analizlerin benzer ve farklı yönlerinin ayrı bir başlık altında sıralanmasının faydalı olacağı düşünülmüş ve bu yönler aşağıdaki gibi listelenmiştir:

- TBA ve BBA'nın her ikisi de istatistiksel bir dönüşüm tekniğidir.
- TBA, ölçülen verilerin ikinci dereceden istatistiklerinde yer alan (kovaryans ya da korelasyon değerleri) bilgiyi kullandığı için normal dağılıma dayalıdır. BBA ise verinin normal olmayan özelliklerinden yararlandığı için yüksek dereceden istatistiklerinde (basıklık) yer alan bilgiyi kullanır. Bu nedenle TBA, sadece değişkenler arası korelasyonları yok edebilir; ancak yüksek dereceden bağımlılıkları yok edemez. Buna karşın BBA, hem korelasyonları hem de yüksek dereceden bağımlılıkları yok eder.
- Veriler normal dağılımlı ise TBA yeterli iken, normal dağılımlı değil ise daha yüksek dereceden istatistikler zorunlu olacağı için BBA'nın kullanımına ihtiyaç duyulur.

- TBA’da elde edilen temel bileşenler (korelasyon ya da kovaryans matrisinin özvektörleri) birbirine diktir. BBA’da elde edilen bileşenler ise Şekil 3.2’de görüldüğü gibi birbirine dik değildir.



Şekil 3.2. Temel Bileşenler Analizi ve Bağımsız Bileşenler Analizi ile Elde Edilen Bileşenler İçin Bir Örnek

- TBA’da verideki varyansı en iyi açıklayan temel bileşenler bulunur ve birinci temel bileşen varyansın en büyük kısmını açıklar. İkinci temel bileşen de varyansın ikinci en büyük kısmını açıklar. Ayrıca birinci temel bileşen ile ikinci temel bileşen birbirine dik olduğu için ilişkisizdirler. Başka bir deyişle, TBA varyansı maksimize eder ve bu nedenle TBA’da elde edilen temel bileşenler varyansın maksimum olduğu yönlere karşılık gelir. Buna karşın BBA, ortaya çıkarılacak bileşenlerin normal olmamasını maksimize ederek birbirinden bağımsız bileşenlerin elde edilmesini sağlar. Böylece BBA vektörleri Şekil 3.2’deki gibi verinin eksenlerine karşılık gelir.

- TBA’da elde edilen temel bileşenler için önem sıralaması yapılabildiğinden bazı bileşenler diğerlerine göre daha önemlidir. Ancak BBA’da böyle bir sıralama yapılamadığı için elde edilen tüm bağımsız bileşenlerin eşit öneme sahip olduğu düşünülür.

- TBA’nın esas amacı, verinin daha düşük boyutlarda temsilini sağlamak iken, BBA’nın asıl amacı ise verinin olabildiğince bağımsız dik vektörlerle temsil edilmesini

sağlamaktır. Kısacası bir anlamda, TBA veriyi bir araya toplayıp sıkıştırırken BBA ise verinin ayrışmasını sağlar (Shlens, 2014).

Bu bölümde ayrıca TBA ile BBA'nın aynı veri kümesi üzerinde karşılaştırıldığı ve BBA'nın daha iyi sonuç verdiği bir uygulamaya da yer verilmiştir. Uygulamada, 512 gözlemden ($t=1,2,\dots,512$) oluşan altı tane birbirinden farklı ve normal dağılmayan seriler üretilmiştir. Serilerin üretildiği fonksiyonlar Eşitlik 3.31'den Eşitlik 3.36'ya kadar verilmiştir. Ayrıca Şekil 3.3'ten Şekil 3.8'e kadar üretilen bu serilerin grafikleri de görülmektedir. Serilerin üretiminde kullanılan fonksiyonlar için Mutihac ve Vun Helle'nin (2004) çalışmaları baz alınmıştır.

$$s_1(t) = 2 \sin(t/149) \cos(t/8) + 0.2(\text{uniform} \sim [0,1]) \quad (3.31)$$

$$s_2(t) = \text{sign}(\sin(12t + 9 \cos(2/29))) + 0.1(\text{uniform} \sim [0,1]) \quad (3.32)$$

$$s_3(t) = (\text{rem}(t,79) - 17) / 23 + 0.1(\text{uniform} \sim [0,1]) \quad (3.33)$$

$$s_4(t) = ((\text{rem}(t,23) - 11) / 9)^5 + 0.1(\text{uniform} \sim [0,1]) \quad (3.34)$$

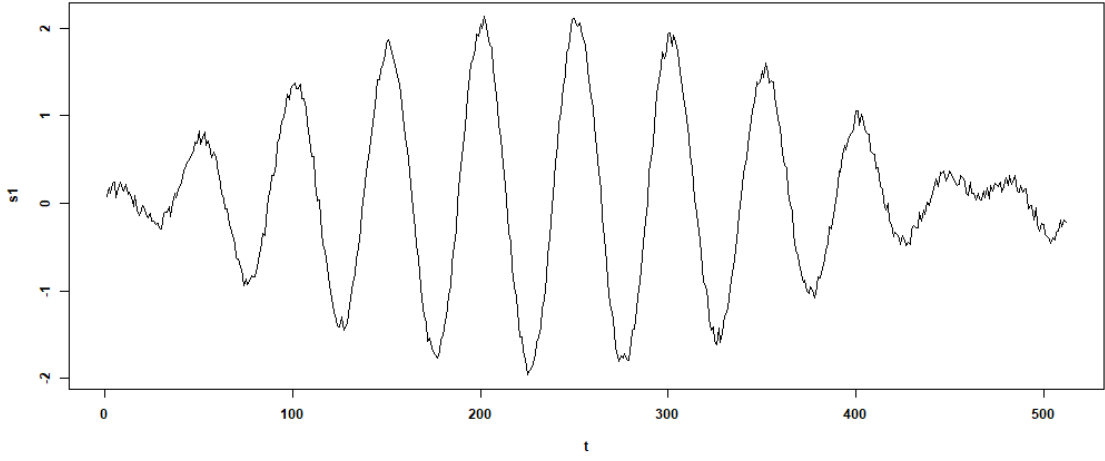
$$s_5(t) = 5 \exp(-t/121) \cos(37t) + 0.1(\text{uniform} \sim [0,1]) \quad (3.35)$$

$$s_6(t) = ((2(\text{uniform} \sim [0,1]) < 0.5) - 1) \log(\text{uniform} \sim [0,1]) \quad (3.36)$$

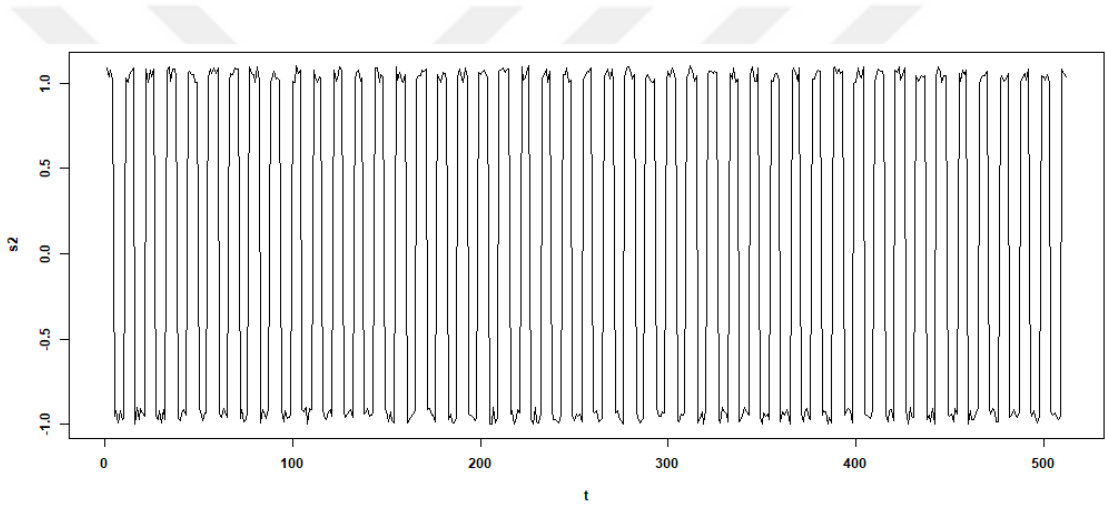
Üretilen bu serilerin bağımsız bileşen olarak kullanılmasıyla gözlenen \mathbf{X} matrisi Eşitlik 3.38'de yer alan \mathbf{A} karışım matrisi kullanılarak oluşturulmuştur. \mathbf{X} matrisinin gözlemlerinin grafikleri de Şekil 3.9'dan itibaren görülmektedir.

$$\mathbf{X} = \mathbf{AS} \quad (3.37)$$

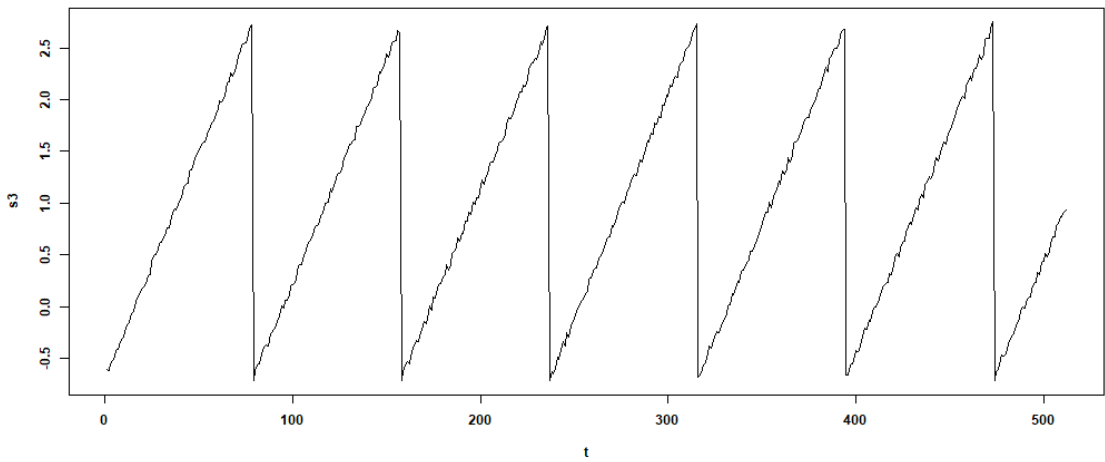
$$\mathbf{A} = \begin{bmatrix} 5.09 & 2.46 & 0.94 & 9.90 & 4.68 & 8.74 \\ 9.56 & 9.32 & 3.16 & 4.33 & 8.37 & 1.47 \\ 4.05 & 1.07 & 5.14 & 9.83 & 0.64 & 8.60 \\ 5.53 & 0.10 & 5.03 & 6.26 & 2.87 & 9.64 \\ 4.07 & 1.36 & 2.56 & 6.06 & 6.28 & 0.92 \\ 1.90 & 4.62 & 3.59 & 1.61 & 1.97 & 1.47 \end{bmatrix} \quad (3.38)$$



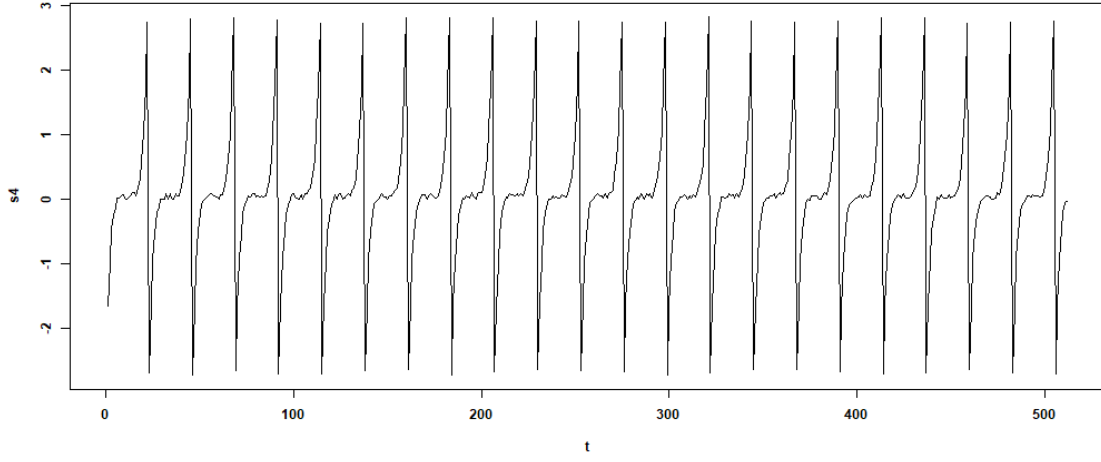
Şekil 3.3. S Matrisi İçin Üretilen s_1 Serisi



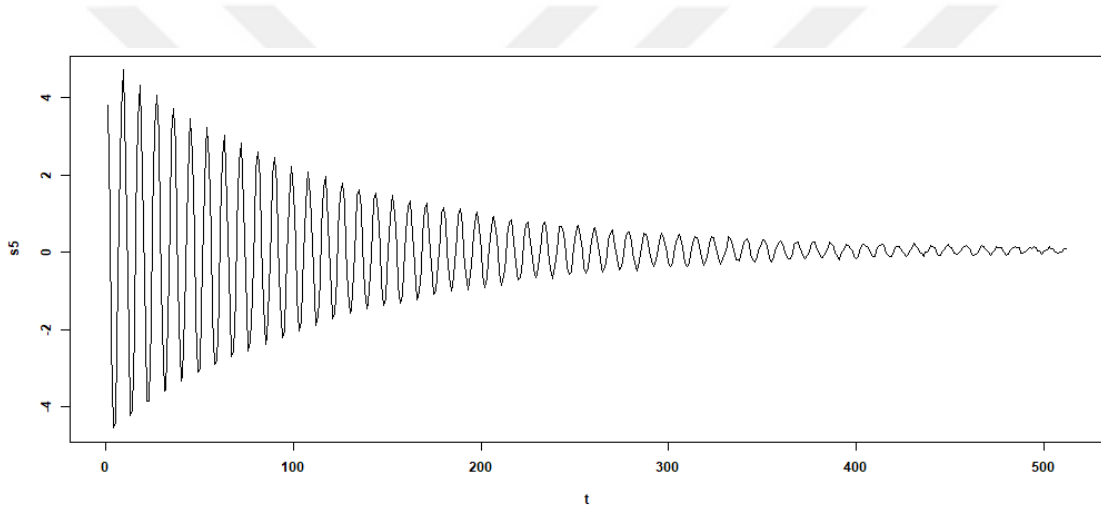
Şekil 3.4. S Matrisi İçin Üretilen s_2 Serisi



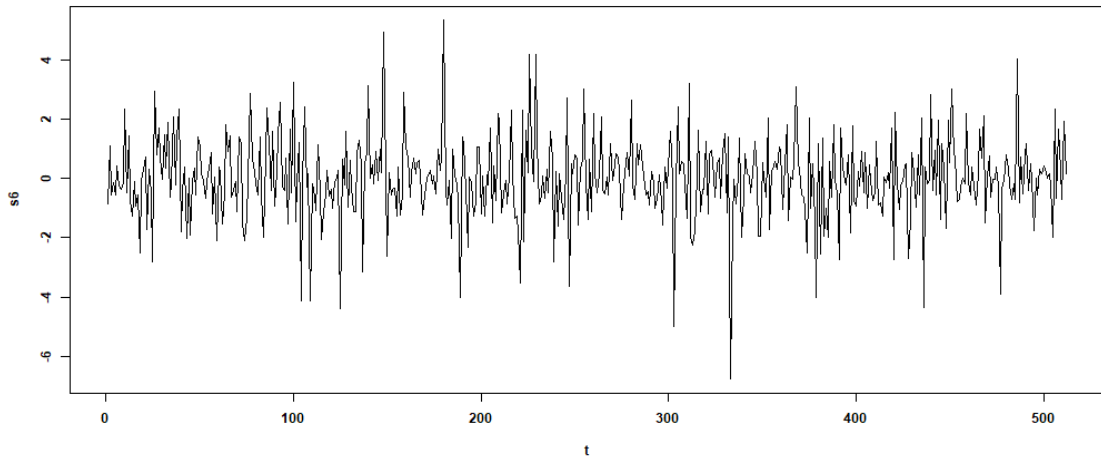
Şekil 3.5. S Matrisi İçin Üretilen s_3 Serisi



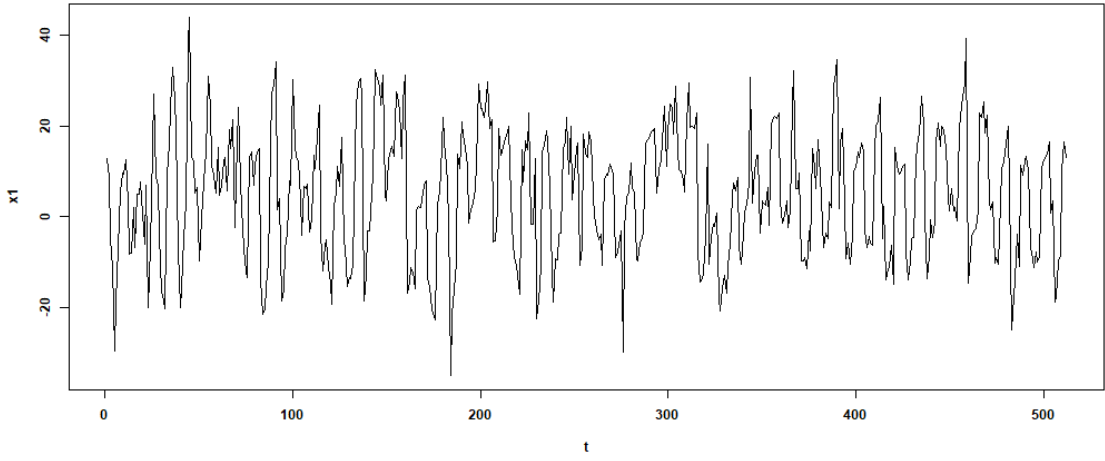
Şekil 3.6. S Matrisi İçin Üretilen s_4 Serisi



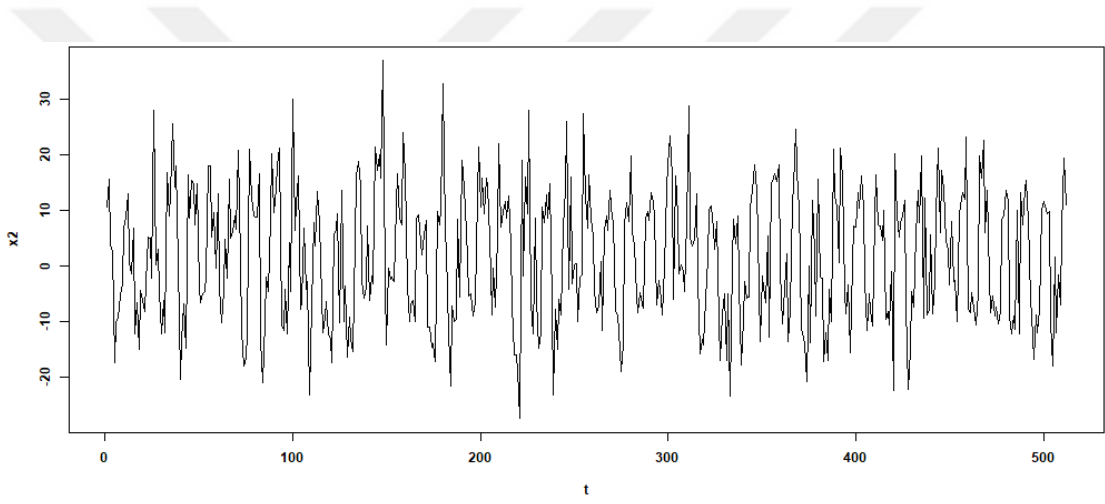
Şekil 3.7. S Matrisi İçin Üretilen s_5 Serisi



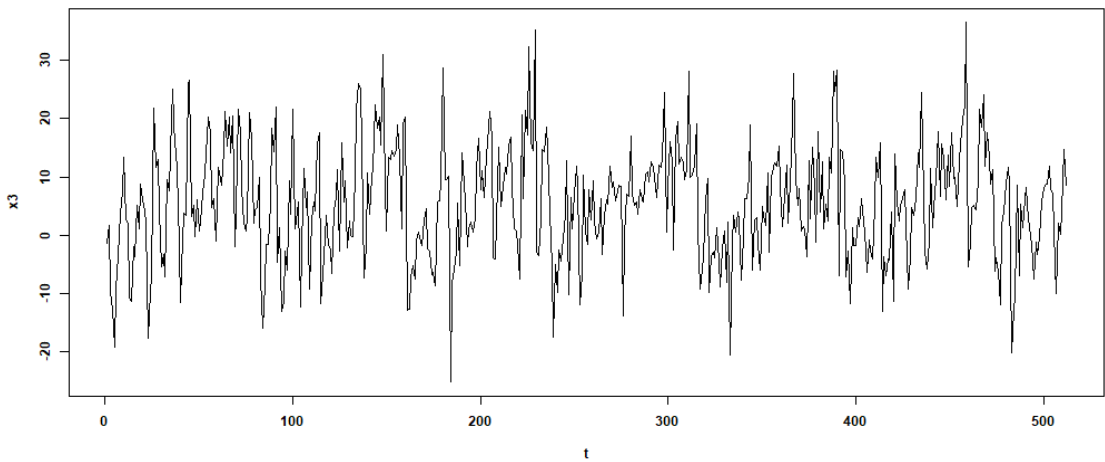
Şekil 3.8. S Matrisi İçin Üretilen s_6 Serisi



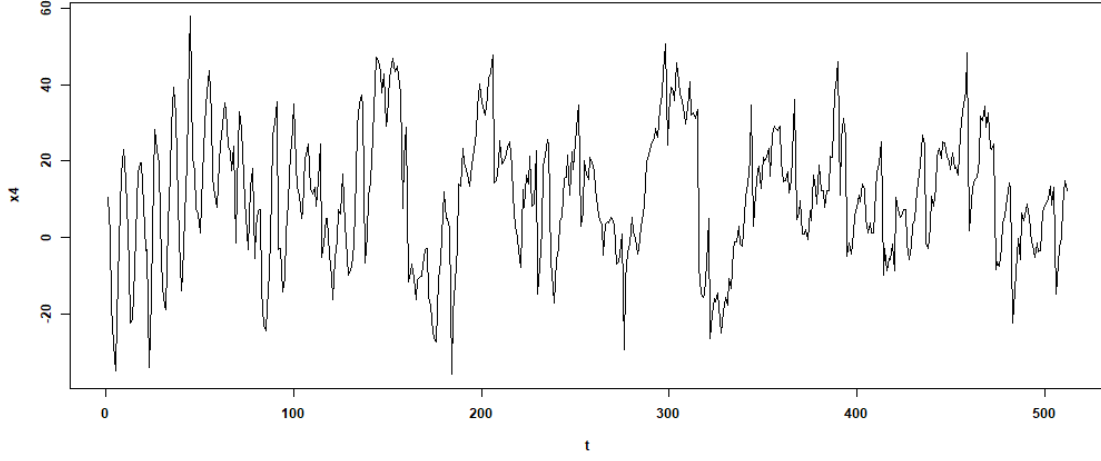
Şekil 3.9. Oluşturulan \mathbf{X} Matrisindeki Gözlenen x_1 Serisi



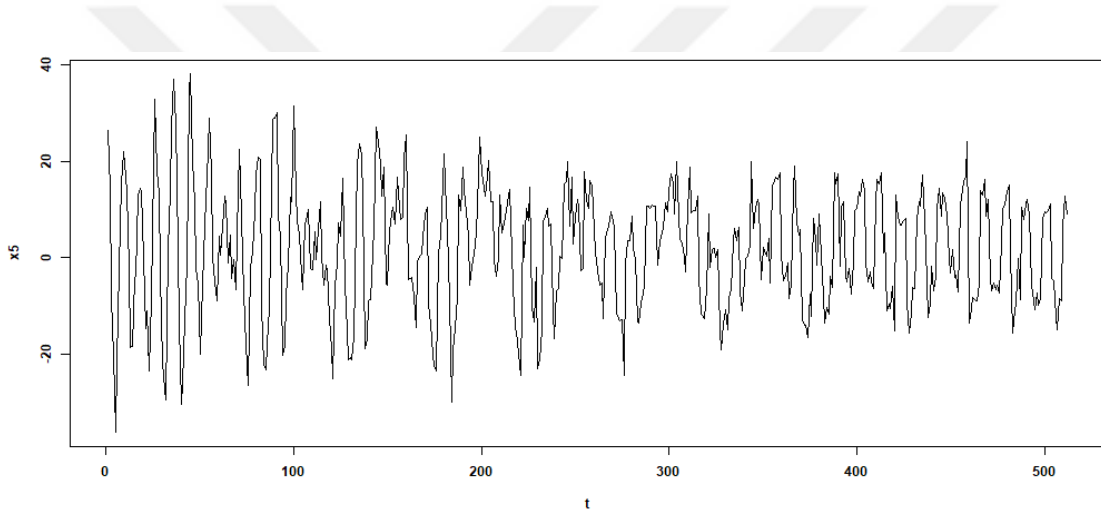
Şekil 3.10. Oluşturulan \mathbf{X} Matrisindeki Gözlenen x_2 Serisi



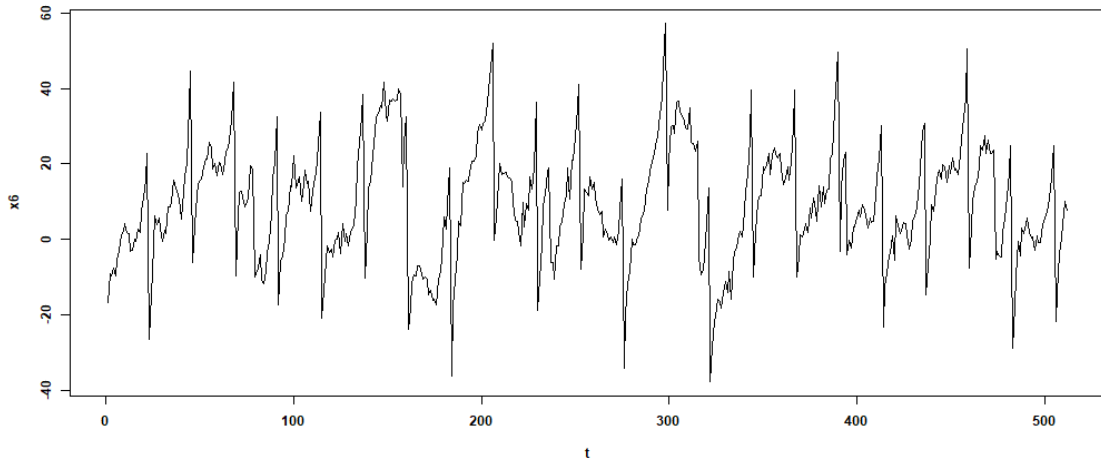
Şekil 3.11. Oluşturulan \mathbf{X} Matrisindeki Gözlenen x_3 Serisi



Şekil 3.12. Oluşturulan X Matrisindeki Gözlenen x_4 Serisi



Şekil 3.13. Oluşturulan X Matrisindeki Gözlenen x_5 Serisi



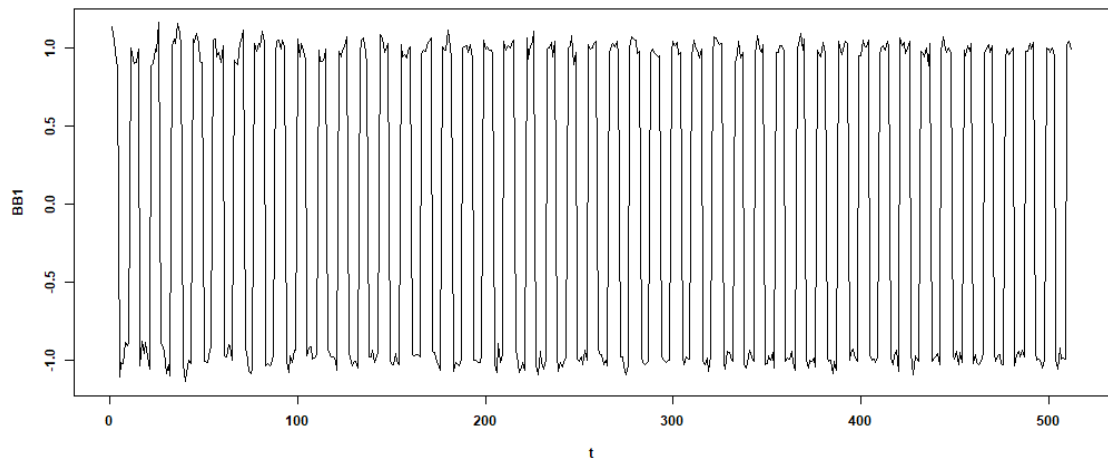
Şekil 3.14. Oluşturulan X Matrisindeki Gözlenen x_6 Serisi

Sonrasında, oluşturulan bu gözlemler BBA ve TBA ile orijinal bileşenlerine ayrılmaya çalışılmıştır. FastICA algoritmasının kullanıldığı BBA ve TBA sonucu elde edilen \hat{A} matrisinin tahminleri sırasıyla Eşitlik 3.39 ve Eşitlik 3.40'ta görülmektedir.

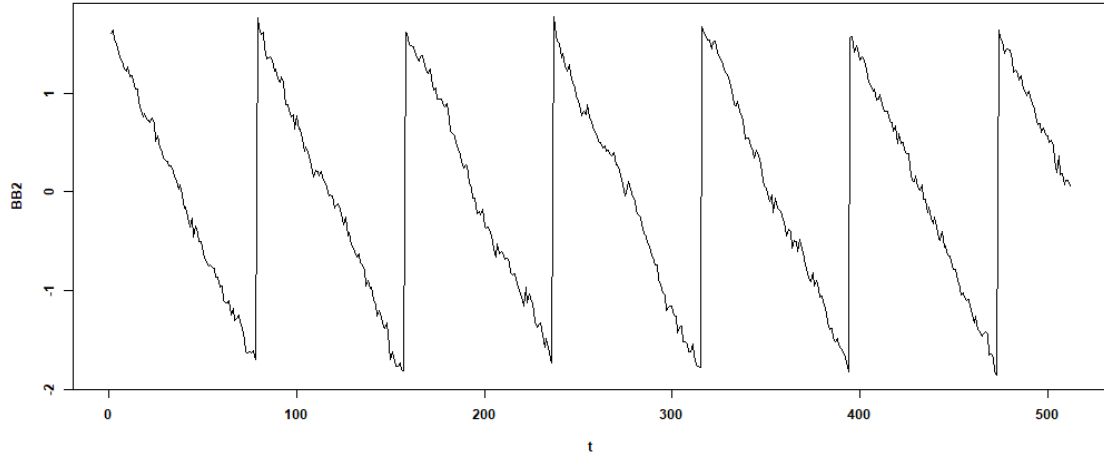
$$\hat{A}_{BBA} = \begin{bmatrix} 9.88 & 9.44 & 3.51 & 4.83 & 8.48 & 2.11 \\ 4.09 & 1.05 & 4.97 & 9.94 & 0.69 & 8.72 \\ -5.61 & -0.16 & -4.99 & -6.19 & -2.92 & -9.61 \\ 4.68 & 2.17 & 0.60 & 9.08 & 4.45 & 7.97 \\ -4.55 & -1.24 & -2.93 & -7.06 & -7.24 & -0.98 \\ -2.28 & -6.38 & -5.04 & -1.79 & -2.47 & -1.57 \end{bmatrix} \quad (3.39)$$

$$\hat{A}_{TBA} = \begin{bmatrix} -0.45 & -0.36 & -0.42 & -0.42 & -0.41 & -0.37 \\ 0.13 & 0.61 & -0.10 & -0.36 & 0.34 & -0.60 \\ -0.20 & 0.34 & 0.68 & -0.29 & -0.53 & 0.07 \\ 0.13 & 0.52 & -0.59 & -0.06 & -0.37 & 0.48 \\ -0.55 & 0.32 & -0.06 & 0.71 & -0.11 & -0.26 \\ -0.65 & 0.10 & 0.00 & -0.30 & 0.53 & 0.44 \end{bmatrix} \quad (3.40)$$

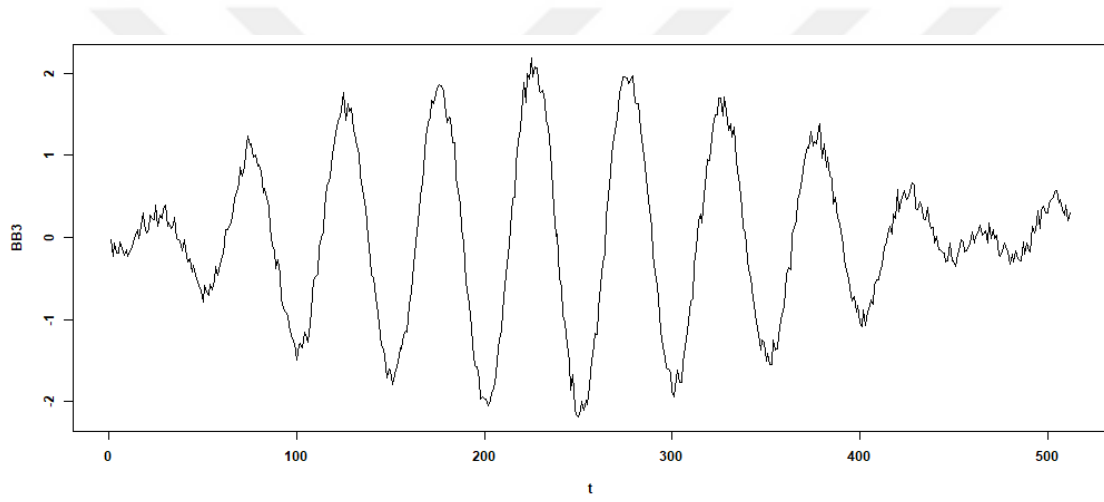
BBA ile tahmin edilen orijinal bileşenler Şekil 3.15'ten itibaren, TBA ile tahmin edilen orijinal bileşenler ise Şekil 3.21'den itibaren verilmiştir. Grafiklerin hepsi birlikte incelendiğinde, BBA'nın S matrisinin bileşenleriyle neredeyse aynı bileşenleri (serileri) bulduğu ve bu yüzden de TBA'ya göre normal dağılım göstermeyen orijinal bileşenleri bulmada daha başarılı olduğu görülmektedir.



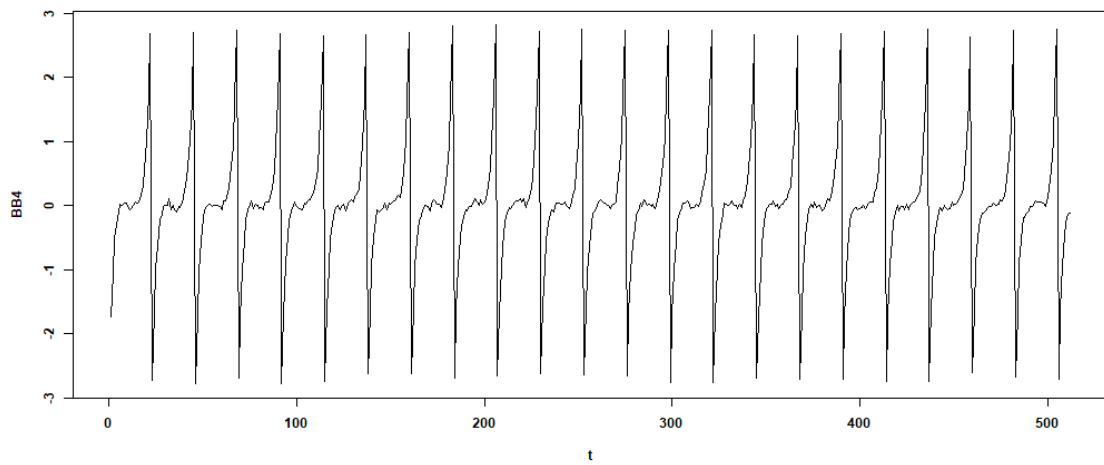
Şekil 3.15. Bağımsız Bileşenler Analizi Sonucunda Tahmin Edilen Bağımsız Bileşen-1



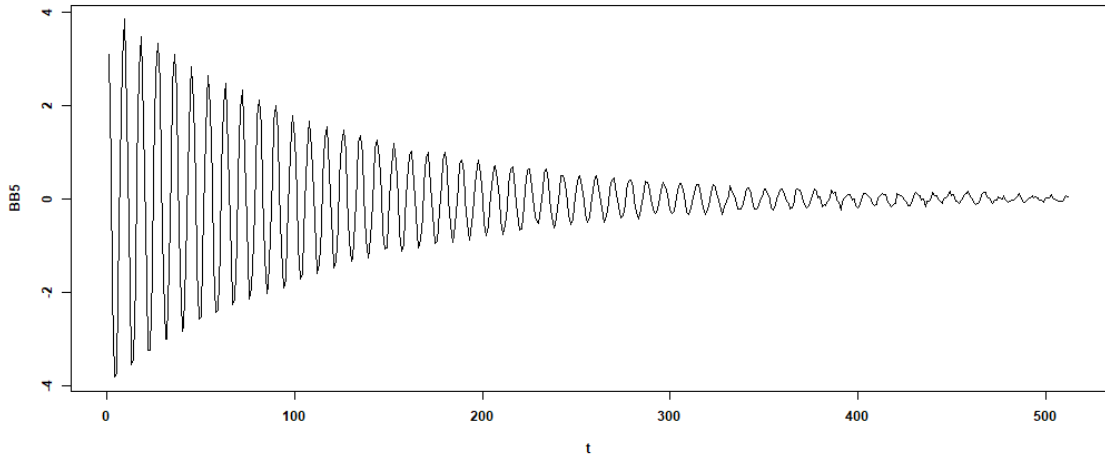
Şekil 3.16. Bağımsız Bileşenler Analizi Sonucunda Tahmin Edilen Bağımsız Bileşen-2



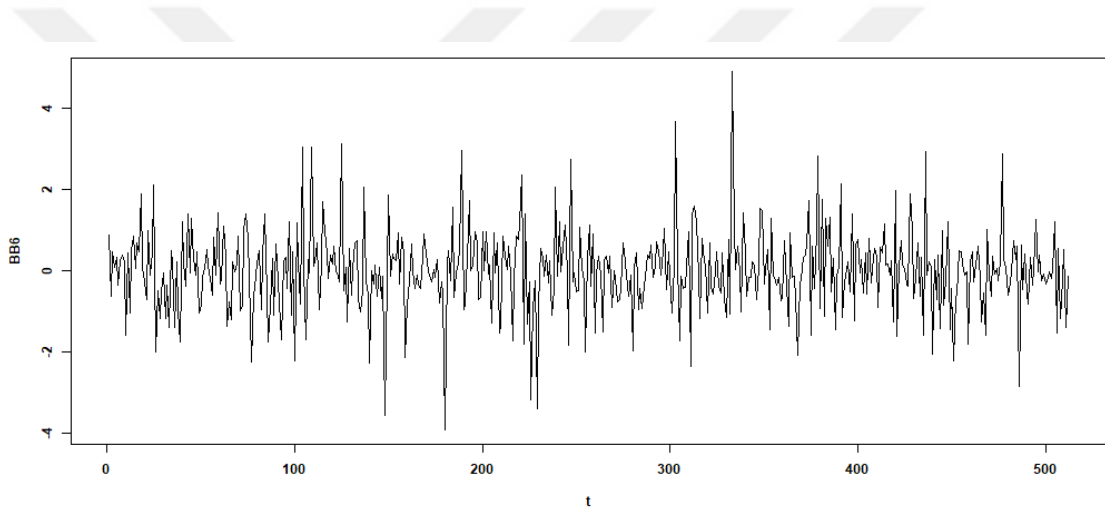
Şekil 3.17. Bağımsız Bileşenler Analizi Sonucunda Tahmin Edilen Bağımsız Bileşen-3



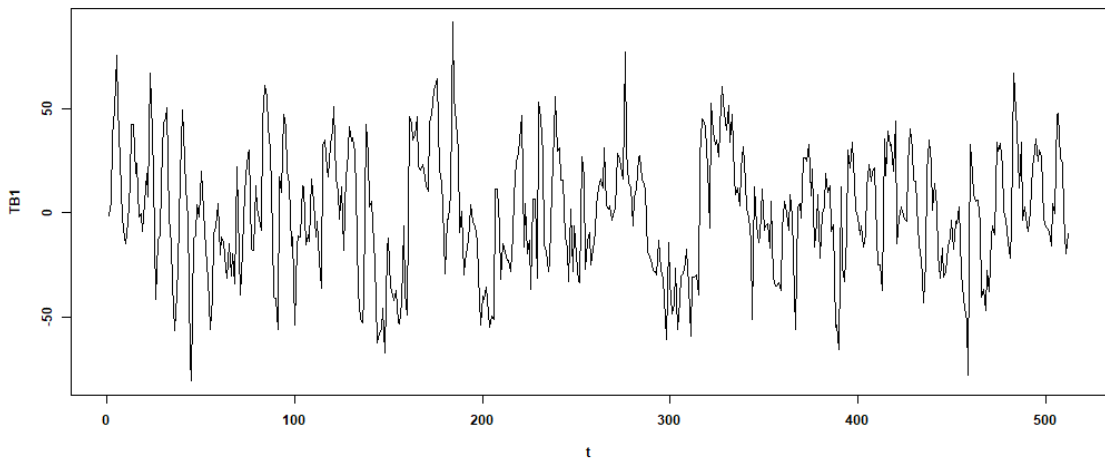
Şekil 3.18. Bağımsız Bileşenler Analizi Sonucunda Tahmin Edilen Bağımsız Bileşen-4



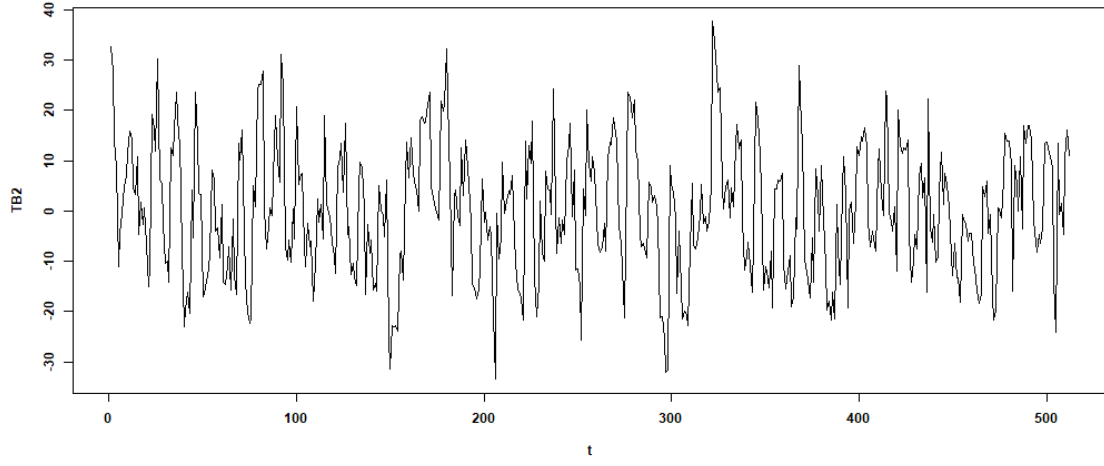
Şekil 3.19. Bağımsız Bileşenler Analizi Sonucunda Tahmin Edilen Bağımsız Bileşen-5



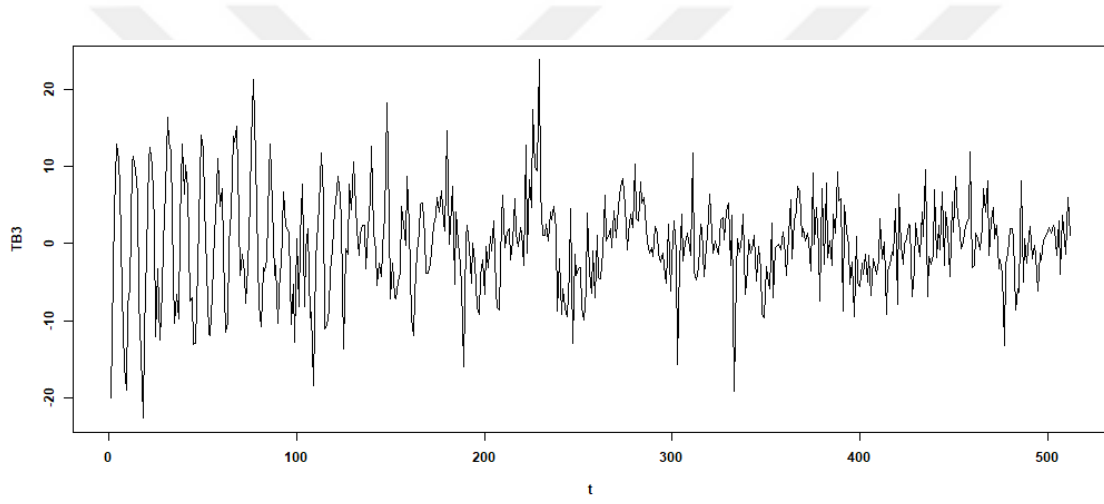
Şekil 3.20. Bağımsız Bileşenler Analizi Sonucunda Tahmin Edilen Bağımsız Bileşen-6



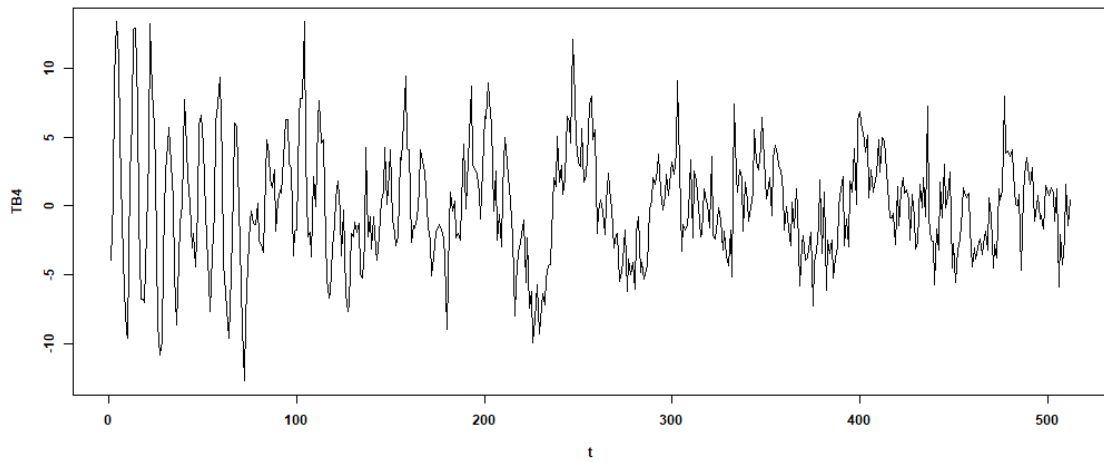
Şekil 3.21. Temel Bileşenler Analizi Sonucunda Tahmin Edilen Temel Bileşen-1



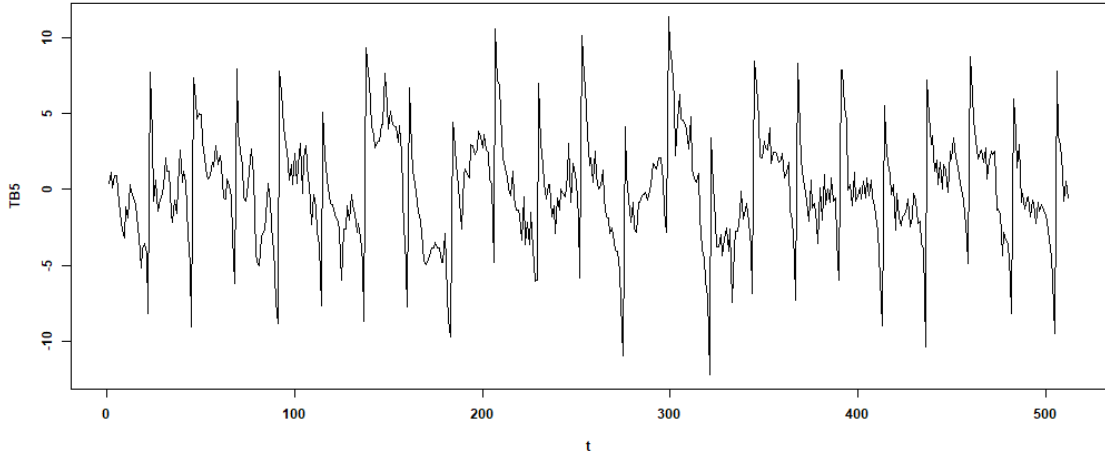
Şekil 3.22. Temel Bileşenler Analizi Sonucunda Tahmin Edilen Temel Bileşen-2



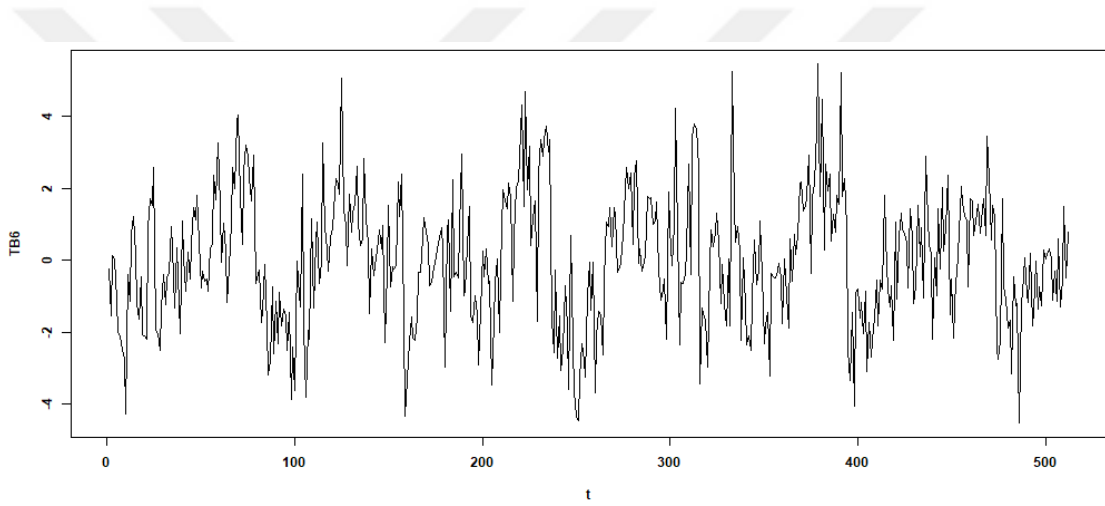
Şekil 3.23. Temel Bileşenler Analizi Sonucunda Tahmin Edilen Temel Bileşen-3



Şekil 3.24. Temel Bileşenler Analizi Sonucunda Tahmin Edilen Temel Bileşen-4



Şekil 3.25. Temel Bileşenler Analizi Sonucunda Tahmin Edilen Temel Bileşen-5



Şekil 3.26. Temel Bileşenler Analizi Sonucunda Tahmin Edilen Temel Bileşen-6

BBA'nın orijinal bileşenleri bulmada TBA'ya göre daha başarılı olduğu, Çizelge 3.1 ve Çizelge 3.2'deki korelasyon matrislerinden de görülmektedir.

Çizelge 3.1 incelendiğinde BBA ile tahmin edilen bileşenlerin S matrisindeki bileşenler ile neredeyse %100'lük bir ilişki içerisinde olduğu gözlenmektedir. Çizelgeye göre, BBA ile tahmin edilen 3. bağımsız bileşen (BB3); s_1 'e, 1. bağımsız bileşen (BB1); s_2 'ye, 2. bağımsız bileşen (BB2); s_3 'e, 4. bağımsız bileşen (BB4); s_4 'e, 5. bağımsız bileşen (BB5); s_5 'e, 6. bağımsız bileşen (BB6) ise s_6 'ya karşılık gelmektedir. BB2 ve BB3 karşılık

geldikleri bileşenler ile negatif bir ilişkiye sahip oldukları için Şekil 3.16 ve Şekil 3.17'nin sırasıyla, Şekil 3.5 ve Şekil 3.3'ün tam tersi oldukları görülmektedir.

Çizelge 3.1. Bağımsız Bileşenler ile Orijinal Bileşenler Arasındaki Pearson Korelasyon Katsayıları

	s_1	s_2	s_3	s_4	s_5	s_6
BB1	0.001	0.999	0.029	0.037	-0.012	0.013
BB2	-0.050	-0.005	-0.997	0.003	0.007	0.036
BB3	-0.997	0.004	0.056	0.002	-0.001	-0.049
BB4	0.002	0.013	-0.016	0.999	-0.005	-0.071
BB5	-0.005	-0.029	-0.001	0.005	0.999	0.026
BB6	0.059	0.020	-0.032	-0.008	0.022	-0.999

Çizelge 3.2 incelendiğinde ise TBA ile tahmin edilen bileşenlerin S matrisindeki bileşenler ile çok yüksek ilişkilere sahip olmadıkları göze çarpmaktadır. Örneğin, TBA ile tahmin edilen 1. temel bileşen (TB1); s_1, s_2, s_3 ve s_4 ile neredeyse aynı miktarda ilişkilere sahiptir. Bu yüzden de TBA sonucu elde edilen temel bileşenlerin hangi orijinal bileşenlere karşılık geldiği tespit edilememektedir.

Çizelge 3.2. Temel Bileşenler ile Orijinal Bileşenler Arasındaki Pearson Korelasyon Katsayıları

	s_1	s_2	s_3	s_4	s_5	s_6
TB1	-0.456	-0.486	-0.437	-0.435	-0.342	-0.214
TB2	-0.250	0.674	-0.432	-0.348	0.176	0.310
TB3	-0.307	0.091	0.239	0.197	-0.658	0.599
TB4	0.492	0.364	-0.324	0.151	-0.616	-0.384
TB5	0.484	-0.094	0.321	-0.771	-0.142	0.264
TB6	-0.399	0.400	0.597	-0.182	-0.135	-0.531

4. BAĞIMSIZ BİLEŞENLER ANALİZİ İLE İSTATİSTİKSEL BİR SORUNUN ÇÖZÜMÜ

Bu bölümde; öncelikle Çoklu Doğrusal Regresyon (ÇDR) modelinden ve bu modelde karşılaşılabilecek sorunlardan biri olan çoklu bağlantı sorunundan bahsedilmiştir. Ardından bu sorunun çözümü için kullanılan mevcut yöntemler anlatılmış ve yeni bir çözüm yöntemi olarak BBA'nın kullanımı önerilmiştir. Konuyla ilgili literatür taraması verildikten sonra önerilen yöntemin nasıl işleyeceğine dair bilgiler sunulmuştur.

4.1. Çoklu Doğrusal Regresyon Analizi

Regresyon analizi, değişkenler arasındaki nedenselliği modellemek, incelemek ve tahmin ya da öngörü yapmak için kullanılan istatistiksel bir yöntemdir (Montgomery, Peck ve Vining, 2013). Regresyon analizi sonucunda açıklanan (bağımlı)⁸ değişken ile açıklayıcı (bağımsız) değişkenler arasındaki bağıntıyı belirtmek için uygun bir denklem bulunur. Analizde kullanılacak olan değişkenlerin sayılarına, türlerine ve dağılımlarına göre farklı regresyon modelleri mevcuttur. Bu kısımda birden fazla açıklayıcı değişken içeren ÇDR modeli ele alınmıştır.

y açıklanan değişkeni ve bu değişken ile ilişkili p tane açıklayıcı değişkenden oluşan ÇDR modeli Eşitlik 4.1.'de verilmiştir. Regresyon katsayıları olarak bilinen β_j , $j = 0, 1, \dots, p$ parametreleri; x_i ($i \neq j$) açıklayıcı değişkenleri sabit tutulduğu zaman x_j açıklayıcı değişkenindeki 1 birimlik değişime karşılık gelen y değişkenindeki beklenen değişimi göstermektedirler. Bu sebeple, kısmi regresyon katsayıları olarak da adlandırılırlar.

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p + \varepsilon \quad (4.1)$$

Eşitlik 4.1'in Eşitlik 4.2'deki gibi matris şeklinde gösterimi de mevcuttur. Burada \mathbf{y} ; açıklanan değişken için $n \times 1$ boyutlu gözlem vektörü, \mathbf{X} ; $(p+1) \times n$ boyutlu açıklayıcı

⁸ Çalışmada, BBA'dan elde edilen bağımsız bileşenler ile karışmaması için regresyon analizindeki bağımsız değişkenler yerine açıklayıcı değişkenler, bağımlı değişken yerine de açıklanan değişken kullanımı tercih edilmiştir.

değişkenler matrisi⁹, β ; $(p+1) \times 1$ boyutlu regresyon katsayıları vektörü ve ϵ ise $n \times 1$ boyutlu rasgele hatalar vektörüdür.

$$\mathbf{y} = \mathbf{X}^T \beta + \epsilon \quad (4.2)$$

Regresyon katsayılarının tahmini için genellikle hata kareler toplamının minimize edildiği En Küçük Kareler (EKK) yöntemi kullanılır. β 'nın EKK tahminleri Eşitlik 4.3'te görülmektedir. Bu eşitlik için $(\mathbf{X}\mathbf{X}^T)$ matrisinin tersi bulunabilmeli yani $\det \mathbf{X}\mathbf{X}^T \neq 0$ olmalıdır.

$$\hat{\beta} = (\mathbf{X}\mathbf{X}^T)^{-1} \mathbf{X}\mathbf{y} \quad (4.3)$$

Bu yöntemle elde edilen \mathbf{y} 'ye karşılık gelen tahmin değerleri Eşitlik 4.4'teki gibidir.

$$\hat{\mathbf{y}} = \mathbf{X}^T \hat{\beta} \quad (4.4)$$

\mathbf{y} ile $\hat{\mathbf{y}}$ arasındaki fark ise $(\mathbf{e} = \mathbf{y} - \hat{\mathbf{y}})$ artıkları vermektedir.

EKK tahmin edicisi $\hat{\beta}$, β 'nın en iyi, doğrusal, yansız tahmin edicisidir. $\hat{\beta}$ tahmin edicisinin bazı istatistiksel özellikleri Eşitlik 4.5 ve Eşitlik 4.6'da verilmiştir. Eşitlik 4.5'te $E(\epsilon) = \mathbf{0}$ ve $(\mathbf{X}\mathbf{X}^T)^{-1} \mathbf{X}\mathbf{X}^T = \mathbf{I}$ 'dir. Ayrıca Eşitlik 4.6'da $(\mathbf{X}\mathbf{X}^T)^{-1} \mathbf{X}$, sabitler matrisidir ve \mathbf{y} 'nin varyansı da $\sigma^2 \mathbf{I}$ 'dir.

$$\begin{aligned} E(\hat{\beta}) &= E\left[(\mathbf{X}\mathbf{X}^T)^{-1} \mathbf{X}\mathbf{y}\right] = E\left[(\mathbf{X}\mathbf{X}^T)^{-1} \mathbf{X}(\mathbf{X}^T \beta + \epsilon)\right] \\ &= E\left[(\mathbf{X}\mathbf{X}^T)^{-1} \mathbf{X}\mathbf{X}^T \beta + (\mathbf{X}\mathbf{X}^T)^{-1} \mathbf{X}\epsilon\right] = \beta \end{aligned} \quad (4.5)$$

$$\begin{aligned} Var(\hat{\beta}) &= Var\left[(\mathbf{X}\mathbf{X}^T)^{-1} \mathbf{X}\mathbf{y}\right] = (\mathbf{X}\mathbf{X}^T)^{-1} \mathbf{X} Var(\mathbf{y}) \left[(\mathbf{X}\mathbf{X}^T)^{-1} \mathbf{X}\right]^T \\ &= \sigma^2 (\mathbf{X}\mathbf{X}^T)^{-1} \mathbf{X}\mathbf{X}^T (\mathbf{X}\mathbf{X}^T)^{-1} = \sigma^2 (\mathbf{X}\mathbf{X}^T)^{-1} \end{aligned} \quad (4.6)$$

⁹ Açıklayıcı değişkenler matrisi \mathbf{X} ; literatürde genellikle $n \times (p+1)$ boyutlu olarak ele alınır ve buna bağlı olarak ÇDR modeli de $\mathbf{y} = \mathbf{X}\beta + \epsilon$ şeklinde yazılır. Ancak çalışmada, BBA ile paralel bir kullanım tercih edildiği için \mathbf{X} açıklayıcı değişkenler matrisi $(p+1) \times n$ olarak belirlenmiş ve dolayısıyla da ÇDR modeli olarak $\mathbf{y} = \mathbf{X}^T \beta + \epsilon$ kullanılmıştır.

EKK yöntemi ile oluşturulan bir ÇDR modelinin, geçerli ve yeterli olabilmesi için aşağıda sıralanan varsayımları sağlaması gerekmektedir:

- Açıklanan değişken ile açıklayıcı değişkenler arasındaki ilişki, doğrusal ya da en azından yaklaşık doğrusaldır.
- Hata terimi sıfır ortalamaya sahiptir.
- Hata terimi, sabit σ^2 varyansına sahiptir
- Hatalar ilişkisizdir.
- Hatalar normal dağılmaktadır.
- Açıklayıcı değişkenler birbirleriyle ilişkisizdir.

Bu varsayımlardan bir ya da birkaçının sağlanmaması durumu, varsayım bozulmaları olarak adlandırılmaktadır. Varsayım bozulmaları; hata terimleri arasında ilişki olması durumunda “özilişki”, hata terimlerinin varyanslarının eşit olmaması durumunda “değişen varyanslılık”, hataların normal olmaması durumunda “normal olmama” ve açıklayıcı değişkenlerin ilişkili olması durumunda da “çoklu bağlantı” sorunu olarak ortaya çıkmaktadır. Tez çalışmasında, varsayım bozulmalarından sadece çoklu bağlantı sorunu ele alınmıştır (Rawlings, Pantula ve Dickey, 2001).

4.2. Çoklu Bağlantı Sorunu, Etkileri ve Çözüm Yolları

Açıklayıcı değişkenler arasında yakın doğrusal bağımlılıklar olduğu zaman çoklu bağlantı sorununun var olduğu söylenir. Temel olarak çoklu bağlantı, X açıklayıcı değişkenler matrisinin satırları arasındaki doğrusal bağımlılıktır. X 'in satırları bağımsız olmadıkça her bir veri kümesi bir dereceye kadar çoklu bağlantıdan etkilenmektedir.

Çoklu bağlantı; yanlış veri toplama yönteminden, model üzerindeki fiziksel kısıtlamalardan, yanlış model seçiminden ya da gözlemlerden çok daha fazla sayıda açıklayıcı değişkene sahip olunması nedeniyle aşırı tanımlanmış modelden kaynaklı olabilir (Montgomery, Peck ve Vining, 2013).

Model anlamlı iken regresyon katsayılarının anlamsız çıkması, bir açıklayıcı değişkenin ya da gözlemin modele eklenmesi ya da çıkarılması durumunda regresyon katsayılarındaki büyük değişimler, modelde çoklu bağlantı probleminin olabileceğinin göstergeleridir. Bu göstergelerin yanı sıra çoklu bağlantının varlığını ve derecesini tespit edebilmek için kullanılan farklı yöntemlerden bazıları aşağıda sıralanmıştır:

- Korelasyon Matrisi

Açıklayıcı değişken çiftleri arasında basit korelasyon katsayısı 0.80 ve üzerinde ise çoklu bağlantı olabileceği düşünülür. Korelasyon katsayısının 1'e yaklaşması ise güçlü çoklu bağlantıyı işaret eder.

- R_j^2 , Tolerans ve Varyans Şişme Değerleri

R_j^2 değeri, x_j 'nin diğer $p-1$ sayıda açıklayıcı değişken üzerinden modellendiğinde elde edilen belirtme katsayısıdır. Bu katsayının 1'e yaklaşması, x_j ile diğer açıklayıcı değişkenler arasında güçlü çoklu bağlantı olabileceğinin göstergesidir. Belirtme katsayısının 1'den çıkarılmasıyla tolerans değeri elde edilir. Bu değer 0'a yakın olması, x_j 'nin diğer açıklayıcı değişkenlerle çoklu bağlantılı olduğunu gösterir. Regresyon modelindeki her bir terimin varyans şişme değeri (VIF) ise açıklayıcı değişkenler arasındaki bağımlılıkların o terimin varyansı üzerindeki birleşik etkilerini ölçmektedir. VIF değerleri, açıklayıcı değişkenlere ilişkin korelasyon matrisinin tersinin köşegen öğeleri ile elde edilir. Eşitlik 4.8'deki gibi Tolerans değerinin kullanılmasıyla da VIF değerlerinin hesaplanması mümkündür.

$$Tolerans_j = (1 - R_j^2) \quad (4.7)$$

$$VIF_j = \frac{1}{(1 - R_j^2)} \quad (4.8)$$

VIF_j değerinin 1'e yakın olması x_j ile diğer açıklayıcı değişkenler arasında ilişki olmadığını göstergesidir. 2'den büyük VIF değerleri çoklu bağlantının bir sorun yaratabileceğini gösterirken 10'dan büyük değerler ise çoklu bağlantının ciddi sorunlar

yaratabileceğini gösterir. VIF değerinin çok büyük olması, çoklu bağlantının güçlü olduğunu bir göstergesidir.

- Koşul İndeksi:

Koşul indeksi (Kİ) açıklayıcı değişkenlerin korelasyon matrisinin özdeğerleri yardımıyla Eşitlik 4.9'daki gibi hesaplanır. Koşul indeksinin 15 ve 30 arasında olması, çoklu bağlantının bir problem yaratabileceğini gösterir. 30'dan büyük değerler ise çoklu bağlantının ciddi bir sorun yaratacağını gösterir.

$$KI = \sqrt{\frac{\lambda_{\max}}{\lambda_{\min}}} \quad (4.9)$$

Çoklu bağlantı sorunu olan bir modelde regresyon katsayıları, olması gereken değer ve işaretten farklı tahmin edilebilmektedir. Ayrıca regresyon katsayılarının standart hataları da olması gerekenden daha büyük tahmin edildiği için katsayıların önemliliklerini gösteren t değerlerini de etkilemektedir. Bu sebeple çoklu bağlantı durumunda açıklayıcı değişkenlerin anlamlılığına ilişkin yanlış sonuçlar elde edilebilir. Özellikle tam çoklu bağlantı durumunda regresyon katsayıları hesaplanamamakta ve standart hataları da sonsuz olmaktadır (Alpar, 2013; Montgomery, Peck ve Vining, 2013).

Regresyon analizinde önemli bir yer tutan çoklu bağlantı sorununun giderilmesi için önerilen pek çok yöntem mevcuttur. Bu yöntemlerden ilki, ek veri toplanmasıdır. Ancak bu yöntem ekonomik kısıttan ya da çalışılan sürecin sonrasında tekrar örnekleme uygun olmamasından dolayı her zaman mümkün olamamaktadır. İkinci bir yöntem ise, değişken seçimidir. Ancak bu yolla analiz dışı bırakılan değişkenlerin bağımlı değişkeni iyi açıklayan değişkenler olması durumunda, değişken seçimi de iyi bir yöntem olmaktan çıkmaktadır. Bir başka yöntem ise, yanlış tahmin tekniklerinin kullanılmasıdır. Bu yöntem, çoklu bağlantı sorununu gidermek için kullanımı en çok önerilen yöntemdir.

Bilindiği üzere EKK tahmin edicileri yansız tahmin ediciler içinde en küçük varyansa sahiptir. Ancak çoklu bağlantı durumunda bu en küçük varyans, kabul edilemez büyüklükte olabilir. EKK tahmin edicisinin yansız olması koşulundan vazgeçilerek daha

düşük varyansa sahip tahminler, yanlı tahmin teknikleri ile elde edilebilir (Rawlings, Pantula ve Dickey, 2001). Yanlı tahmin tekniklerinin başında özgün değişkenler yerine bunların dik dönüşümlerinin kullanıldığı Temel Bileşenler Regresyonu (TBR) ve korelasyon matrisinin köşegen elemanlarına küçük bir pozitif sayı eklenerek tahmin varyanslarının küçültüldüğü Ridge Regresyonu (RR) gelmektedir. Bir diğer yanlı tahmin yöntemi ise, TBR ile benzer bir teknik kullanan Kısmi En Küçük Kareler Regresyonu'dur (KEKR). TBR'de sadece açıklayıcı değişkenlere TBA uygulanırken, KEKR'de ise hem açıklanan hem de açıklayıcı değişkenlere eş zamanlı bir ayrıştırma uygulanır.

Tez çalışmasında yukarıda bahsedilen yöntemlerden hareketle, BBA sonucu elde edilen bağımsız bileşenlerin de çoklu bağlantı sorununu çözebileceği, hatta mevcut diğer yöntemlerin aksine ilişki yerine daha güçlü istatistiksel bir özellik olan bağımsızlığı esas aldığı için daha iyi sonuçlar vereceği düşünülmüştür. Bu amaçla, BBA ile hem açıklayıcı değişkenler arasındaki ilişkinin yok edildiği hem de açıklanan değişkenin göz önünde bulundurulduğu TBR ve KEKR'ye benzer bir mantıkla çoklu bağlantı sorunu çözülmeye çalışılmıştır. Bir sonraki kısımda da, bu konuyla ilgili yapılmış olan çalışmalar incelenmiştir. Ancak öncesinde, uygulama bölümünde önerilen modelin, TBR ve KEKR'ye benzerliğinden dolayı sadece bu yöntemlerle karşılaştırılması yapıldığı için bu yöntemlerin işleyişine ilişkin hatırlatıcı bilgilere yer verilmiştir.

TBR, merkezileştirilmiş ve ölçeklendirilmiş açıklayıcı değişkenler matrisine (\mathbf{Z} , $p \times n$ boyutlu) TBA uygulanması üzerine inşa edilmiştir. Eşitlik 4.10'da, TBA ile elde edilen \mathbf{T} temel bileşenleri görülmektedir.

$$\mathbf{T} = \mathbf{V}^T \mathbf{Z} \quad (4.10)$$

TBR'de orijinal değişkenler yerine, TBA ile elde edilen temel bileşenlerin yani orijinal değişkenlerin dik dönüşümleri üzerine Eşitlik 4.12'deki gibi EKK yöntemi uygulanır ve $\hat{\boldsymbol{\gamma}} = (\mathbf{T}\mathbf{T}^T)^{-1} \mathbf{T}\mathbf{y} = \mathbf{L}^{-1}\mathbf{T}\mathbf{y}$ şeklindeki regresyon katsayıları tahmin edilir.

$$\mathbf{y} = \mathbf{Z}^T \boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (4.11)$$

$$\mathbf{y} = \mathbf{T}^T \boldsymbol{\gamma} + \boldsymbol{\varepsilon} \quad (4.12)$$

Dönüşüm matrisi ortogonal bir matris ($\mathbf{V}\mathbf{V}^T = \mathbf{I}$) olduğu için aynı zamanda Eşitlik 4.13 de yazılabilir.

$$\mathbf{Z}^T \boldsymbol{\beta} = \mathbf{Z}^T \mathbf{V}\mathbf{V}^T \boldsymbol{\beta} = \mathbf{T}^T \boldsymbol{\gamma} \quad (4.13)$$

Görüldüğü üzere TBR ile elde edilen temel bileşenlerin regresyon katsayılarından orijinal değişkenlerin regresyon katsayılarına $\boldsymbol{\beta} = \mathbf{V}\boldsymbol{\gamma}$ eşitliği ile geçiş yapılabilir. Bu nedenle tüm temel bileşenler kullanıldığında, EKK ile aynı tahmin $\hat{\boldsymbol{\beta}} = \mathbf{V}\hat{\boldsymbol{\gamma}}$ sonuçlarına ulaşılır. Ancak tüm temel bileşenlerin regresyon analizinde kullanılması çoklu bağlantı sorununu ortadan kaldırmaz. Zaten TBR'nin arkasında yatan temel düşünce de çoklu bağlantıya sebep olan boyutları yok etmektir ki, bunu da açıklayıcı değişkenler matrisinin korelasyon matrisinden elde edilen özdeğerlerin (λ_j) çok küçük olduğu boyutları analiz dışı bırakarak yapar. Örneğin s tane temel bileşen yok edildiğinde genellikle bunlar, en küçük özdeğere sahip s tane temel bileşendir ve regresyon analizi için geriye kalan g tane temel bileşeni kullanır ($s + g = p$). Böyle bir durumda regresyon katsayıları, Eşitlik 4.14'teki gibi elde edilir.

$$\hat{\boldsymbol{\beta}}_{(g)} = \mathbf{V}_{(g)} \hat{\boldsymbol{\gamma}}_{(g)} \quad (4.14)$$

$(p \times 1) \quad (p \times g) \quad (g \times 1)$

Ancak temel bileşenler yok edilirken dikkatli olmak gerekir. Çünkü yok edilen her temel bileşen, tahminler üzerinde kısıtlama ve yan miktarında bir artış getirir. Uygulamada genellikle, varyansı büyütebilecek yeterince küçük özdeğerlere sahip temel bileşenler ve tahmin edilen regresyon katsayısı sıfırdan farklı olmayan temel bileşenler yok edilir. (Rawlings, Pantula ve Dickey, 2001).

KEKR'de, açıklayıcı değişkenler ile açıklanan değişken ya da değişkenler arasındaki kovaryansı maksimum yapan birbirine dik gizli değişkenler bulunur. Her iki değişken bloğu arasındaki ilişki, kısmi en küçük kareler skor vektörleri ile modellenir. Merkezileştirilmiş değişken blokları, Eşitlik 4.15 ve 4.16'daki gibi ayrıştırılır.

$$\mathbf{X}^T = \mathbf{TP}^T + \mathbf{E} \quad (4.15)$$

$$\mathbf{Y}^T = \mathbf{UQ}^T + \mathbf{F} \quad (4.16)$$

Eşitliklerde \mathbf{X} ; $p \times n$ boyutlu açıklayıcı değişkenler matrisini, \mathbf{Y} ; $p \times m$ boyutlu açıklanan değişkenler matrisini, \mathbf{T} ve \mathbf{U} ; $n \times p$ boyutlu skor matrislerini, \mathbf{P} ; $n \times p$ ve \mathbf{Q} ; $m \times p$ boyutlu yükler matrisini, \mathbf{E} ; $n \times p$ ve \mathbf{F} ; $m \times p$ boyutlu artık matrislerini göstermektedir. \mathbf{T} matrisinin \mathbf{t} , \mathbf{U} matrisinin \mathbf{u} skor vektörleri arasındaki kovaryansı $[\text{cov}(\mathbf{X}^T \mathbf{w}, \mathbf{Y}^T \mathbf{c})]^2 = [\text{cov}(\mathbf{t}, \mathbf{u})]^2$ maksimum olacak şekilde NIPALS algoritması ile \mathbf{w} ve \mathbf{c} ağırlıkları bulunur. Ayrıca bu maksimizasyonu sağlayan \mathbf{t} , \mathbf{u} skor bileşenleri ile \mathbf{p} ve \mathbf{q} yük vektörleri hesaplanır. \mathbf{X} 'in skorları, \mathbf{Y} 'nin iyi birer tahmin edicisi oldukları için hesaplanan değerler yardımıyla Eşitlik 4.16, $\mathbf{Y}^T = \mathbf{TQ}^T + \mathbf{G}$ şeklinde de yazılabilir. Böylece KEKR katsayıları Eşitlik 4.17'deki gibi elde edilebilir.

$$\mathbf{Y}^T = \mathbf{TQ}^T + \mathbf{G} = \mathbf{X}^T \mathbf{WQ}^T + \mathbf{G} = \mathbf{X}^T \mathbf{B} + \mathbf{F} \quad (\mathbf{B} = \mathbf{WQ}^T) \quad (4.17)$$

İstenilen sayıda bileşen elde edildiğinde ya da açıklanan değişkendeki değişimin büyük bir kısmı açıklandığında algoritma sonlanır (Polat, 2009; Bulut, 2011).

KEKR'de NIPALS algoritmasından başka kullanılacak çekirdek algoritması, özvektör algoritması, SIMPLS, SAMPLS, UNIPALS gibi çok çeşitli algoritmalar da mevcuttur. KEKR'de yeterli bileşen sayısına karar verilirken genellikle çapraz geçerlilik yöntemi kullanılmaktadır (Bulut, 2011).

4.3. Bağımsız Bileşenler Analizi ile Çoklu Bağlantı Sorununun Çözümüne Dair Literatür Taraması

Bu bölümde; 'regresyon', 'çoklu bağlantı', 'Bağımsız Bileşenler Analizi' anahtar kelimelerinin kullanıldığı genel bir literatür taraması yapılmıştır. Literatür taramasının sonucunda BBA'nın bir regresyon tekniği olarak ele alındığı çalışmalara daha çok kimyanın kemometri¹⁰ alanında rastlanılmıştır.

¹⁰ Kimyasal verilerin istatistik, matematik ve bilgisayar algoritmaları ile işlenmesi Kemometri bilim dalını ortaya çıkarmıştır.

Konuyla ilgili ilk çalışma, Gustafsson (2005) tarafından yapılmıştır. Çalışmada öncelikle, ÇDR modellerinin kimya alanında standart bir yöntem olarak kullanıldığından; ancak çoğu zaman gözlem sayısının değişken sayısından küçük olması, açıklayıcı değişkenler arasında yüksek ilişkilerin var olması ve deneysel hataların mevcut olması gibi nedenlerden dolayı oluşturulan modellerin kötü uyum sergileme ya da gözlemlere duyarlı çözümler üretme sorunlarına sahip olduğundan bahsedilmiştir. Bu gibi durumlarda kullanılan gizli değişken temelli yöntemlerden TBR ve KEKR yöntemlerine alternatif olarak BBA'ya dayalı iki yeni regresyon algoritması yazar tarafından önerilmiştir. Önerilen algoritmalarından bir tanesi TBR'deki TBA'yı BBA ile değiştirmekte, diğeri de KEKR'deki gibi sadece açıklayıcı değişkenleri değil, açıklanan değişkeni de kullanarak gizli değişkenleri tahmin etmektedir. Çalışmada, uniform ve normal dağılımlardan üretilen ilişkili değişkenler üzerinde, farklı örneklem büyüklükleri ve farklı ölçüm hataları için yöntemlerin karşılaştırmaları yapılmıştır. Karşılaştırılan tüm yöntemler doğru tahminler verirken, yalnızca BBA'ya dayalı yöntem ile kimyasal olarak yorumlanabilen gizli değişkenler elde edildiği sonucuna varılmıştır.

Shao ve ark. (2006) ise çalışmalarında, Bağımsız Bileşenler Regresyonu (BBR) adını ilk kez kullanmışlardır. Tütün bitkisinin NIR spektrumlarından elde edilen örneklerdeki genel bileşenlerin (toplam şeker, toplam nitrojen, nikotin ve su) tahmini üzerine bir uygulama gerçekleştirmişler ve BBR ile TBR'den elde ettikleri sonuçları karşılaştırmışlardır. Her iki yöntemde de aynı sayıda bağımsız bileşen ya da temel bileşen ile aynı açıklayıcılık gücüne ve aynı tahmin hatasına sahip modeller elde etmişlerdir. Ancak bağımsız bileşenler ortogonalliğe göre daha güçlü bir koşul ve yüksek dereceden istatistik olan bağımsızlığı esas aldığı için, temel bileşenlere göre daha fazla kimyasal açıklama imkanı sunmuştur. Bu yargıya, elde edilen temel bileşenlerin ve bağımsız bileşenlerin açıklayıcı değişkenler ile aralarındaki ilişki katsayıları hesaplanarak varılmıştır. Sonuç olarak BBA'nın karmaşık kimyasal veri kümelerinde nicel ve nitel bilgileri almak için umut vadeden bir araç olduğunu söylemişlerdir.

Kaneko, Arakawa ve Funatsu (2008) çalışmalarına; ÇDR, TBR, KEKR gibi çok değişkenli tekniklerin kimoinformatikteki pek çok problemin çözümünde kullanıldığına dair bilgilerle bir giriş yapmışlardır. Regresyon analizinde kullanılan açıklayıcı

değişkenlerden önemli bileşenlerin ortaya çıkarılması için BBA kullanımını ve buna bağlı olarak da Bağımsız Bileşenler Analizi-Çoklu Doğrusal Regresyon (BBA-ÇDR) yöntemini önermişlerdir. Normal dağılımdan ilişkili olarak üretilen iki açıklayıcı değişken ve bir açıklanan değişken ile benzetim çalışması yapmışlardır. BBA-ÇDR yöntemi, KEKR'ye göre çok daha doğru tahmin sonuçları vermiştir. Bunun yanı sıra gerçek bir veri kümesi olan 346 değişken ve 1290 gözlemden oluşan sulu çözünürlük veri kümesi üzerinde de KEKR, BBA-ÇDR ve Genetik Algoritma-KEKR yöntemlerinin tahmin ve açıklayıcılık güçlerini karşılaştırmış ve en iyi sonuçları BBA-ÇDR yönteminden almışlardır. Ek olarak BBA-ÇDR yönteminden elde edilen regresyon katsayıları, her tanımlayıcı maddenin (değişkenin) sulu çözünürlüğe olan katkısını belirttiği için ayrıntılı bir şekilde yorum yapabilme şansı da vermiştir.

Dai ve ark. (2013), regresyon analizinin baraj deformasyonunun istatistiksel olarak modellenmesinde yoğun olarak kullanıldığından bahsetmişlerdir. Ancak regresyon analizinde kullanılan değişkenler arasındaki çoklu bağlantı sorununun zaman zaman bu alanda yanlış modeller sunabildiğini ve bu durumun çözümü için kullanılan geleneksel bir yöntem olan adımsal regresyonun, deformasyona neden olan bazı çevresel faktörleri göz ardı edebildiğini belirtmişlerdir. TBR ve KEKR yöntemlerinin de baraj deformasyonunu fiziksel olarak yorumlamada yetersiz kalmalarından dolayı çalışmada, baraj deformasyonunu modellemek ve deformasyonun fiziksel kökenlerini belirlemek için BBR'nin kullanımını önermişlerdir. 6 açıklayıcı değişken ve bir açıklanan değişken ile genişliği 400 olan bir örneklem üzerinde gerçekleştirilen benzetim çalışmasının sonrasında önerdikleri yöntemi, Çin'in Hunan eyaletindeki Wuqiangxi Barajının deformasyonunu modellemek için kullanmışlardır. Sonuç olarak, BBR'nin barajın deformasyonunu sadece doğru bir şekilde modellemediği, aynı zamanda ortaya çıkarılmış bağımsız bileşenler aracılığıyla deformasyonu etkileyen fiziksel faktörlerin (sıcaklık, su seviyesi, zaman) tanımlanmasına da yardımcı olduğunu göstermişlerdir.

Chakrabarty ve Sarkar (2013) çalışmalarında, Hint Borsasındaki (Nifty) oynaklık üzerine ekonomik değişkenlerin etkisini araştırmışlardır. Çalışmada oynaklığı hesaplarken GARCH(1,1) modelini kullanmışlar ve 12 adet makroekonomik değişkenin hesaplanan bu GARCH oynaklığı üzerindeki etkisini ÇDR ile incelenmişlerdir. Ancak bu 12

değişkenden 7'sinin çoklu bağlantı sorunundan etkilendiği görülmüştür. Bu değişkenlerin hepsini analiz dışı bırakmak mümkün olmadığı için çoklu bağlantıyı yok etmek amacıyla bu 7 değişkene BBA uygulanmıştır. Değişkenler normal dağılım göstermediği için TBA yerine BBA tercih edilmiştir. Bu 7 değişkenin bağımsız bileşenleri elde edildikten sonra tüm değişkenlere tekrar doğrusal regresyon analizi uygulanmıştır. Ancak her ne kadar çalışmada bahsedilmese de, regresyon modelinde kullanılan değişkenlerin bir kısmının orijinal değişken olması, diğer bir kısmının ise BBA'dan elde edilen bağımsız bileşenler olmasının elde edilen katsayıların yorumlanmasında bir sorun yaratacağı aşikârdır.

Zhan, Dai ve Zhou (2014) çalışmalarında, baraj deformasyonu analizinde regresyon yöntemlerinin çok fazla kullanıldığını ancak deformasyona neden olan değişkenler arasında bağımlılık problemi olduğu için her bir değişkenin gerçek katkısının ayırt edilemediğini belirtmişlerdir. Dai ve ark.'ın (2013) çalışmasına benzer şekilde çalışmalarında, bağımlılığı yok etmek ve istatistiksel olarak bağımsız değişkenleri elde etmek için BBA'nın kullanımını önermişlerdir. Yapay bir veri kümesi üzerinde EKK regresyonu, TBR ve BBR ile elde ettikleri sonuçları karşılaştırmışlardır. Ayrıca gerçek bir veri kümesi üzerinde de bir uygulama gerçekleştirmişlerdir. Sonuç olarak BBR'nin bağımlılık problemini çözebildiğini ve bağımlı değişkene daha iyi açıklama getirdiğini göstermişlerdir. Ayrıca BBA'nın deformasyon analizi alanındaki uygulamalarda büyük potansiyeller içerdiği yorumunda bulunmuşlardır.

Haque, de Souza ve Rahman (2017) çalışmalarında, su talebinin doğru modellenmesinin ve tahmin edilmesinin zorluklarından bahsetmişlerdir. Bunun nedeni olarak da mevcut verinin yapısını ve kalitesini, havza modellemesindeki coğrafi farklılıkları göstermişlerdir. Su talebi tahmininde çeşitli tekniklerin benimsendiği, ancak BBR'nin uygulamasının daha önce yapılmadığı üzerinde durmuşlardır. Bu yüzden çalışmada, orta vadeli kentsel su talebi tahmini için ilk kez BBR yönteminin kullanımı Brezilya'nın, Aquidauana şehrinden alınan veriler üzerinde gerçekleştirilmiştir. Ayrıca geliştirilen BBR modelinin performansı, ÇDR ve TBR ile karşılaştırılmıştır. BBR modelinin, su talebini modellemede diğer iki modelden daha iyi performans gösterdiği görülmüştür. Böylece BBR yönteminin, su talebi modellerini başarıyla geliştirme potansiyeline sahip

olduğu ve kabul edilen yöntemin su talebi tahmin modeli geliştirmek için diğer ülkelere de uygulanabilir olduğu sonuçlarına varılmıştır.

4.4. Bağımsız Bileşenler Analizi ile Çoklu Bağlantı Sorununun Çözümü

Bilindiği gibi, TBA ile p boyutlu uzayı çok iyi tanımlayan yeni dik (ilişkisiz) temel bileşenler ya da özvektörler elde edilir. Analizin amacı, p tane değişkenin açıkladığı bilginin k tane yeni değişkenle açıklanmasıdır ($k \leq p$). Bu amaç doğrultusunda p boyutlu uzaydaki toplam varyans (özdeğerler toplamı), her biri özvektörlerle tanımlanan yeni değişkenlerle ifade edilmektedir. En büyük varyans birinci özvektöre, en küçük varyans sonuncu özvektöre ait olduğundan birinci temel bileşen en fazla, diğer temel bileşenler ise gittikçe azalan oranlarda toplam açıklanan varyansa katkıda bulunur. Böylece az sayıda bileşen ile toplam varyansın büyük bir kısmı açıklanabilmektedir (Tatlıdil, 2002; Jolliffe, 2011).

TBR ise doğrusal regresyon modelinin, TBA sonucunda özgün bağımsız değişkenlerin doğrusal birleşimi olarak elde edilen temel bileşenlerin kullanımıyla yeniden açıklanabileceği gerçeğine dayanarak oluşturulmuştur. Bu noktadan hareketle, BBA sonucu elde edilen bağımsız bileşenlerin de doğrusal regresyonda özgün değişkenler yerine kullanılabileceği çalışma kapsamında incelenmiştir. Zira bilindiği üzere istatistiksel bağımsızlık, ilişkisizliğe göre çok daha güçlü bir kavramdır ve bağımsız bileşenler elde edilirken yüksek dereceden istatistikler kullanılır. Temel bileşenler elde edilirken ise kovaryans ya da korelasyon gibi ikinci dereceden istatistikler kullanılır. Bu nedenle de BBA, TBA'nın başarısız olduğu durumlarda daha iyi sonuçlar veren güçlü bir yöntemdir. Ancak TBA'da, elde edilen temel bileşenler toplam varyansı açıklayıcılık güçlerine göre sıralanabilirken, BBA'da analizin doğası gereği elde edilen bağımsız bileşenler için bir sıralama yapılamamaktadır. Bu sebeple, TBR'de hangi temel bileşenlerin seçileceği önem seviyelerine göre belli iken BBA'nın kullanıldığı regresyon yönteminde ise bir belirsizlik bulunmaktadır. Ancak TBR'de de başka bir sorun mevcuttur: Regresyon modeline dahil edilecek temel bileşenlerin seçiminde açıklanan değişkenin göz ardı edilmesi. Zaman içerisinde bu problem fark edilmiş ve bunun için açıklanan değişkenin de göz önünde bulundurulduğu KEKR yöntemi geliştirilmiştir.

Yukarıda bahsi geçen BBA ile ilgili belirsizlik için bir önceki başlıkta sunulan konuyla ilgili literatür genel olarak tekrar değerlendirildiğinde, birçok çalışmanın açıklayıcı değişkenlere BBA uyguladıktan sonra elde ettikleri bağımsız bileşenlerden hangilerinin regresyon modeline dahil edilmesi gerektiği konusunda yeterli bir bilgi vermediği, bazılarının bütün bağımsız bileşenlerle yapılabilecek tüm kombinasyonları denediği ve içlerinden tahmin hatasını minimum yapan kombinasyonu seçtikleri (işlem süresini uzatan bir yöntemdir), bazılarının ise açıklanan değişken ile en yüksek Pearson korelasyon katsayısına sahip olan bağımsız bileşenleri seçtiği görülmüştür. Korelasyon katsayısını kullanan çalışmaların KEKR yöntemine benzer bir bakış açısıyla açıklanan değişkeni de göz önünde bulundurdıkları görülmektedir. Bilindiği üzere, Pearson korelasyon katsayısı sadece doğrusal ilişkilere odaklanmaktadır. Bağımsız bileşenler, açıklayıcı değişkenlerin doğrusal bir birleşimleri olduğu için ve açıklanan değişken de açıklayıcı değişkenlerin doğrusal bir birleşimi olduğu için, bağımsız bileşenler ile açıklanan değişken arasında da doğrusal bir ilişki olduğu varsayılmaktadır. Ancak; açıklanan değişken ile açıklayıcı değişken arasındaki Pearson korelasyon katsayısının yüksek olması, bu açıklayıcı değişkenin iyi bir açıklama gücüne sahip olacağı anlamına gelmemektedir. Çünkü regresyon analizinde göz önünde bulunduran, açıklanan değişken ile açıklayıcı değişkenler arasındaki ilişki değil, nedenselliktir. Bu nedenle çalışmada, doğrusal ilişkinin yanı sıra karmaşık ilişkileri de göz önünde bulunduran bir ölçütün kullanılmasının daha doğru sonuçlar vereceği düşünülerek korelasyon katsayısı yerine iki sistemin taşıdığı ortak bilgi miktarının önerilmiştir. Korelasyon katsayısı yerine ortak bilgi miktarının kullanıldığı çalışmalar incelendiğinde, özellikle ilişkisel bağlantılar yerine nedensellik ortaya çıkarılmak istendiğinde ortak bilgi miktarının tercih edildiği görülmektedir (Pearl, 2009). Bu açıdan, regresyon analizinde açıklanan değişkeni modelleyecek bağımsız bileşenlerin seçiminde ortak bilgi miktarının kullanımının ne kadar doğru bir tercih olduğu da görülmektedir.

Ortak bilgi miktarına göre bağımsız bileşenlerin seçiminin yapıldığı regresyon analizinin aşamaları aşağıda sıralanmıştır:

1) Açıklayıcı değişkenlerden oluşan $\mathbf{X}_{p \times n}$ veri matrisi (p değişken sayısı, n gözlem sayısı) öncelikle BBA analizinin hazırlık aşamasının tamamlanması için merkezileştirilir

ve beyazlatılır. Böylece ortalaması $\mathbf{0}$, kovaryans matrisi de birim matris (\mathbf{I}) olan $\mathbf{Z}_{p \times n}$ matrisi elde edilir.

$$\mathbf{Z} = \mathbf{\Lambda}^{-1/2} \mathbf{T}^T \mathbf{X} = \mathbf{QX} \quad (4.19)$$

Eşitlik 4.19'da $\mathbf{Q} = \mathbf{\Lambda}^{-1/2} \mathbf{T}$, beyazlatma matrisidir. $\mathbf{\Lambda}$, $E(\mathbf{XX}^T)$ kovaryans matrisinin özdeğerlerinden oluşan köşegen matris, \mathbf{T} de $E(\mathbf{XX}^T)$ kovaryans matrisinin özvektörlerinden oluşan ortogonal matristir.

2) Veri matrisinin beyaz hali olan \mathbf{Z} matrisinden, BBA'da kullanılan algoritmalarından biri olan FastICA algoritması ile bağımsız bileşenlerden oluşan $\mathbf{U}_{p \times n}$ matrisi tahmin edilir.

$$\mathbf{U} = \mathbf{WZ} = \mathbf{WQX} \quad (4.20)$$

Buradaki \mathbf{W} karışım matrisi, $E(\mathbf{UU}^T) = E(\mathbf{WZ}(\mathbf{WZ})^T) = \mathbf{WE}(\mathbf{ZZ}^T)\mathbf{W}^T = \mathbf{I}$ özelliğinden dolayı ortogonal bir matristir.

3) Bir önceki aşamada tahmin edilen \mathbf{u} bağımsız bileşen vektörlerinin, açıklanan değişken ile olan ortak bilgi miktarları hesaplanır. Ortak bilgi miktarlarına göre büyükten küçüğe doğru sıralanan bağımsız bileşenlerden ilk k tanesi ($k \leq p$) seçilir.

4) Seçilen bağımsız bileşenler, regresyon analizinde açıklayıcı değişken olarak kullanılır ve regresyon katsayıları aşağıdaki eşitlikler yardımıyla hesaplanır:

$$\begin{aligned} \mathbf{y} &= \mathbf{X}^T \hat{\boldsymbol{\beta}}_{x \rightarrow y} + \boldsymbol{\varepsilon} & \mathbf{y} &= \mathbf{U}_k \hat{\boldsymbol{\beta}}_{u \rightarrow y} + \boldsymbol{\varepsilon} \\ \hat{\boldsymbol{\beta}}_{x \rightarrow y} &= (\mathbf{XX}^T)^{-1} \mathbf{Xy} & \hat{\boldsymbol{\beta}}_{u \rightarrow y} &= (\mathbf{U}_k \mathbf{U}_k^T)^{-1} \mathbf{U}_k \mathbf{y} \\ \mathbf{y} &= \mathbf{X}^T \hat{\boldsymbol{\beta}}_{x \rightarrow y} & \mathbf{y} &= \mathbf{U}_k^T \hat{\boldsymbol{\beta}}_{u \rightarrow y} \end{aligned} \quad (4.21)$$

Eşitlik 4.21'den hareketle, $\mathbf{X}^T \hat{\boldsymbol{\beta}}_{x \rightarrow y} = \mathbf{U}_k^T \hat{\boldsymbol{\beta}}_{u \rightarrow y}$ olduğu için $\hat{\boldsymbol{\beta}}_{x \rightarrow y} = (\mathbf{WQ})^T \hat{\boldsymbol{\beta}}_{u \rightarrow y}$ olarak yazılabilir. Böylece, bağımsız bileşenlerin regresyon katsayılarından orijinal değişkenlerin regresyon katsayılarına geçiş yapılabilir.

EKK regresyonuyla elde edilen β katsayılarının yansız ve $\hat{\sigma}^2(\mathbf{XX}^T)^{-1}$ varyanslı olmalarından hareketle bağımsız bileşenlerin regresyon katsayılarının beklenen değerini ve varyansını da hesaplamak mümkündür.

$$E(\hat{\beta}_{u \rightarrow y}) = E\left(\left((\mathbf{WQ})^T\right)^{-1} \beta_{x \rightarrow y}\right) = \left((\mathbf{WQ})^T\right)^{-1} E(\beta_{x \rightarrow y}) = \left((\mathbf{WQ})^T\right)^{-1} \beta \quad (4.22)$$

$$\begin{aligned} \text{Var}(\hat{\beta}_{u \rightarrow y}) &= \text{Var}\left(\left((\mathbf{WQ})^T\right)^{-1} \beta_{x \rightarrow y}\right) = \left((\mathbf{WQ})^T\right)^{-1} \text{Var}(\beta_{x \rightarrow y}) \left(\left((\mathbf{WQ})^T\right)^{-1}\right)^T \\ &= \left((\mathbf{WQ})^T\right)^{-1} \sigma^2 (\mathbf{XX}^T)^{-1} (\mathbf{WQ})^{-1} \\ &= \sigma^2 \left((\mathbf{WQ})^T\right)^{-1} (\mathbf{UU}^T)^{-1} (\mathbf{WQ})^{-1} \end{aligned} \quad (4.23)$$

5. UYGULAMA

Bu bölümde; regresyon analizindeki açıklayıcı değişkenlere BBA uygulandıktan sonra elde edilen bağımsız bileşenlerin, açıklanan değişken ile sahip oldukları ortak bilgi miktarlarına göre seçtikleri ve sonrasında da regresyon analizinde değişken olarak yer aldıkları BBA-OBM olarak kısaltılan yöntemin performansı değerlendirilmiştir. Bu amaçla, çoklu bağlantıya sahip iki tane yapay veri ve dört tane de gerçek veri kümesi üzerinde uygulamalar gerçekleştirilmiştir. Önerilen BBA-OBM yönteminin sonuçları; KEKR, TBR ve bileşenlerin ortak bilgi miktarı yerine Pearson korelasyon katsayısı ile seçtikleri BBA-K yöntemlerinin sonuçları ile karşılaştırılmıştır.

5.1. Yapay Veri Uygulamaları

Yapay veriler üretilirken Lee, Quin ve Lee (2006) ile Haykin'in (2008) çalışmalarından esinlenilmiştir. $n = 100, 500, 1000, 5000$ şeklindeki farklı örneklem büyüklüklerinin 3'te 2'si eğitim, 3'te 1'i ise test kümesi olarak ayrılmıştır. Farklı düzeylerde hata terimlerinin de eklendiği veri kümelerine; önerilen BBA-OBM yöntemi ve KEKR, TBR, BBA-K yöntemleriyle regresyon analizi uygulanmıştır. Her durum ve yöntem için eğitim kümesinin tahmin, test kümesinin de öngörü sonuçlarının performansı hata kareler ortalaması (HKO) ile karşılaştırılmıştır.

$$HKO = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (5.1)$$

5.1.1. Yapay Veri Uygulaması-1

Öncelikle $\mathbf{S} = [s_1 \quad s_2]$ ve $k = 1, 2, \dots, n$ olmak üzere Eşitlik 5.2 ve Eşitlik 5.3'te verilmiş olan dağılımlardan iki farklı bağımsız bileşen üretilmiştir.

$$s_1(k) = \sin(0.5k) \quad (5.2)$$

$$s_2(k) = \text{uniform} \sim [-1, 1] \quad (5.3)$$

Her ne kadar farklı dağılımlardan üretilmiş olsalar da, üretilen bileşenlerin istatistiksel olarak birbirlerinden bağımsız olup olmadıklarının kontrolü, parametrik olmayan bir test

olan Hoeffding Bağımsızlık Testi ile yapılmıştır. Hoeffding (1948) tarafından önerilen bu testte amaç, marjinal dağılımların çarpımı ile ampirik iki değişkenli dağılım arasındaki mesafeyi hesaplayarak veri kümelerinin bağımsızlığını sınamaktır. Yokluk hipotezi x ve y raslantı değişkenlerinin ya da veri kümelerinin birbirinden bağımsız olduğu şeklinde kurulur.

$$H_0 : F(x, y) = F(x)F(y), \text{ tüm } (x, y) \text{ çiftleri için}$$

$$H_1 : F(x, y) \neq F(x)F(y), \text{ tüm } (x, y) \text{ çiftleri için}$$

Teste, n gözlemden oluşan x raslantı değişkenine R_i , $i = 1, 2, \dots, n$ sıra sayıları ve n gözlemden oluşan y raslantı değişkenine de T_i , $i = 1, 2, \dots, n$ sıra sayıları verilerek başlanır. Daha sonra $i = 1, 2, \dots, n$ için c_i değeri hesaplanır. c_i değeri, $x_j < x_i$ ve $y_j < y_i$ olan (x_j, y_j) çiftlerinin sayısıdır. Raslantı değişkenlerinde aynı değeri alan gözlemler yok ise başka bir deyişle, ortak sıra sayısı yok ise c_i değeri hesaplanırken Eşitlik 5.4, ortak sıra sayısı var ise de Eşitlik 5.5 kullanılır (Hollander, Wolfe ve Chicken, 2014).

$$c_i = \sum_{j=1}^n \phi(x_j, x_i) \phi(y_j, y_i) \text{ ve } \phi(a, b) = \begin{cases} 1, & a < b \text{ ise} \\ 0, & \text{diğer durumlarda} \end{cases} \quad (5.4)$$

$$c_i = \sum_{\substack{j=1 \\ a \neq 1}}^n \phi^*(x_j, x_i) \phi^*(y_j, y_i) \text{ ve } \phi^*(a, b) = \begin{cases} 1, & a < b \text{ ise} \\ \frac{1}{2}, & a = b \text{ ise} \\ 0, & \text{diğer durumlarda} \end{cases} \quad (5.5)$$

Sonrasında D_1 , D_2 ve D_3 değerleri ile Hoeffding'in D istatistiği Eşitlik 5.9'daki gibi hesaplanır:

$$D_1 = \sum_{i=1}^n (R_i - 1)(R_i - 2)(T_i - 1)(T_i - 2) \quad (5.6)$$

$$D_2 = \sum_{i=1}^n (R_i - 2)(T_i - 2)c_i \quad (5.7)$$

$$D_3 = \sum_{i=1}^n c_i (c_i - 1) \quad (5.8)$$

$$D = \frac{D_1 - 2(n-2)D_2 + (n-2)(n-3)D_3}{n(n-1)(n-2)(n-3)(n-4)} \quad (5.9)$$

Hoeffding'in D istatistiği, $\left[-\frac{1}{60}, \frac{1}{30}\right]$ aralığında değerler alır. Olasılık değerleri

$$d_\alpha = \frac{1}{30} \sqrt{\frac{2(n^2 + 5n - 32)}{9n(n-1)(n-3)(n-4)\alpha}} \quad \text{için } P(D > d_\alpha) \leq \alpha \text{ olarak bulunur ise yokluk}$$

hipotezi α düzeyinde reddedilir (de Siqueira Santos ve ark., 2013).

Üretilen veri kümelerine her bir örneklem büyüklüğü göz önünde bulundurularak Hoeffding Bağımsızlık Testi uygulandığında elde edilen olasılık değerleri Çizelge 5.1'de özetlenmiştir.

Çizelge 5.1. Yapay Veri Uygulaması-1 İçin Hoeffding Test Sonuçları

Hoeffding Bağımsızlık Testi	n=100	n=500	n=1000	n=5000
(s_1, s_2)	$p = 0.377$	$p = 0.669$	$p = 0.316$	$p = 0.578$

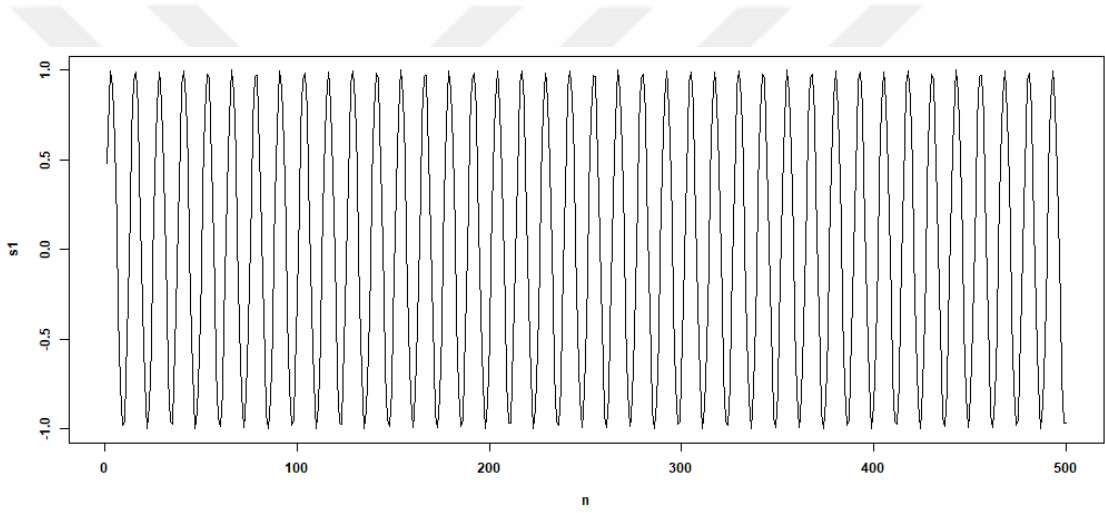
Bütün p olasılık değerleri $\alpha = 0.05$ 'ten büyük çıktığı için üretilen s_1 ve s_2 bileşenlerinin tüm örneklem büyüklüklerinde istatistiksel olarak birbirinden bağımsız oldukları %95 güven düzeyinde söylenebilir.

Üretilen bağımsız bileşenlerin istatistiksel olarak birbirlerinden bağımsız olduklarının gösterilmesinden sonra, temelde iki bağımsız bileşenin etkisiyle oluştuğu varsayılan açıklanan değişkenin gerçek değerleri, $\mathbf{y} = 2\mathbf{s}_1 + 7\mathbf{s}_2$ eşitliği ile üretilmiştir. Sonrasında, doğrudan bu iki bağımsız değişkenin gözlenemediği, onların yerine altı tane açıklayıcı değişkenin gözlendiği bir durum oluşturmak için bu iki bağımsız bileşen kullanılarak $\mathbf{X} = \mathbf{A}\mathbf{S}$ eşitliği ile $\mathbf{X} = [\mathbf{x}_1 \ \mathbf{x}_2 \ \mathbf{x}_3 \ \mathbf{x}_4 \ \mathbf{x}_5 \ \mathbf{x}_6]$ şeklindeki açıklayıcı değişkenlerin

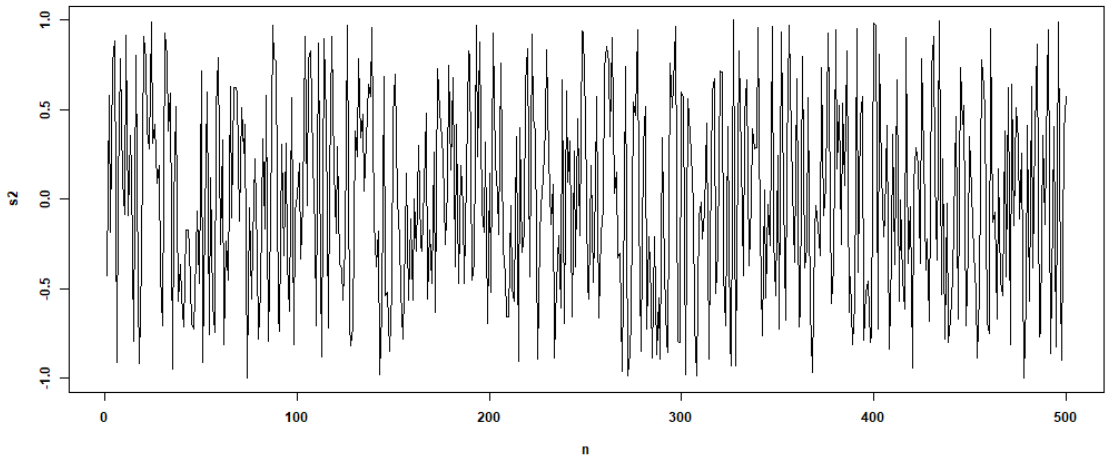
verileri üretilmiştir. Üretim aşamasında Eşitlik 5.10'da verilen \mathbf{A} matrisi değerleri kullanılmıştır.

$$\mathbf{A} = \begin{bmatrix} 0.86 & -0.55 & 0.17 & -0.63 & 0.89 & 0.27 \\ 0.39 & 0.65 & 0.32 & 0.12 & -0.97 & -0.14 \end{bmatrix} \quad (5.10)$$

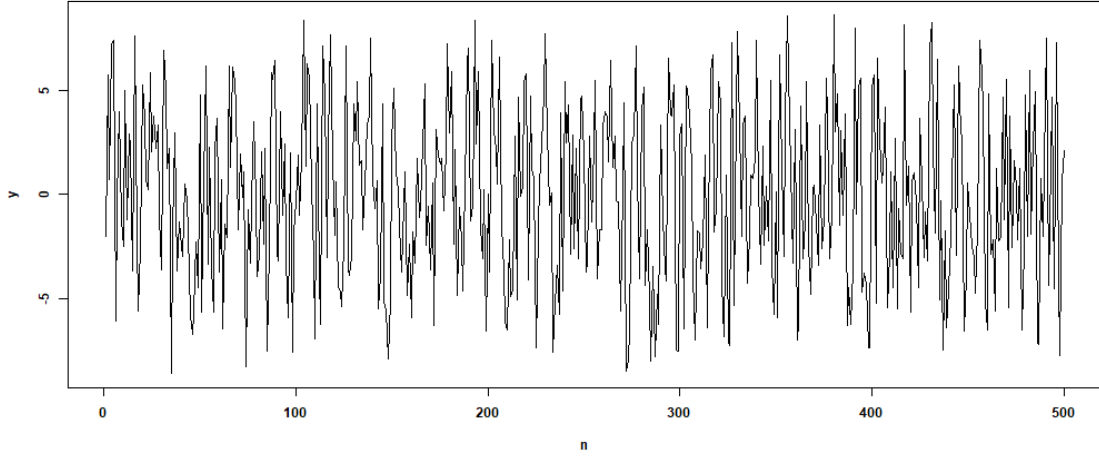
Ayrıca üretilen açıklayıcı değişkenlere normal dağılımdan üretilmiş olan 0 ortalama ve farklı standart sapmalara (σ) sahip hata terimleri de eklenmiştir. Örnek olarak, $n = 500$ için üretilen $s_1, s_2, y, x_1, x_2, x_3, x_4, x_5, x_6$ 'nın grafikleri sırasıyla Şekil 5.1'den itibaren verilmiştir.



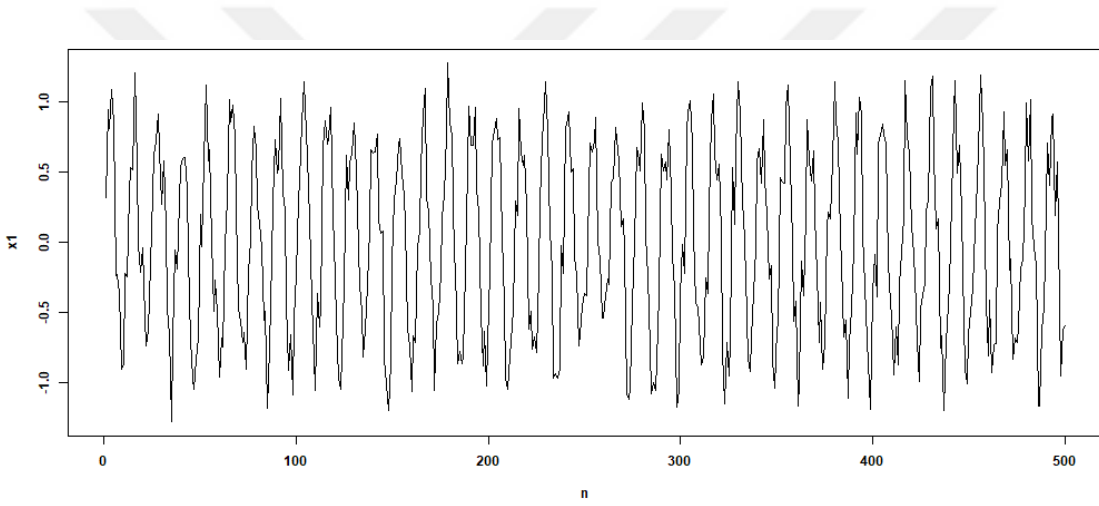
Şekil 5.1. $n = 500$ İçin s_1 Bileşeni



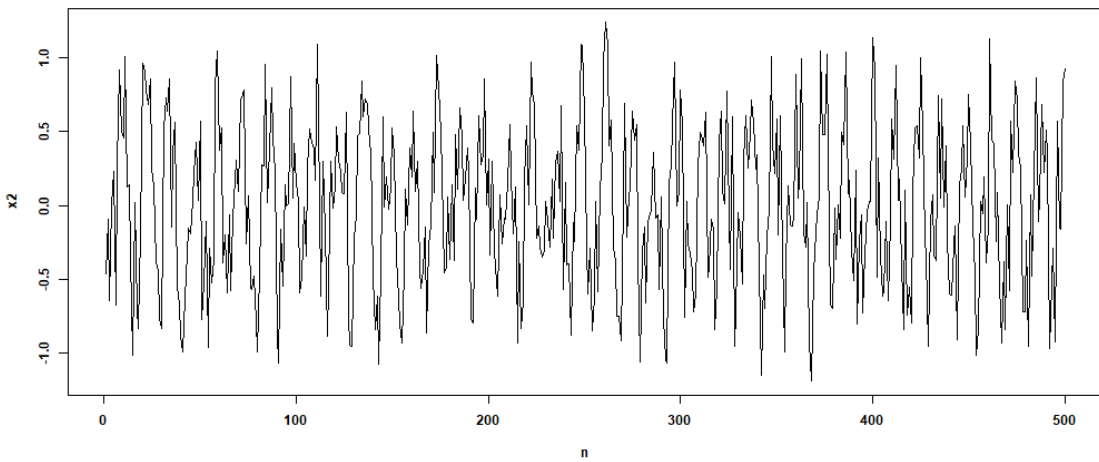
Şekil 5.2. $n = 500$ İçin s_2 Bileşeni



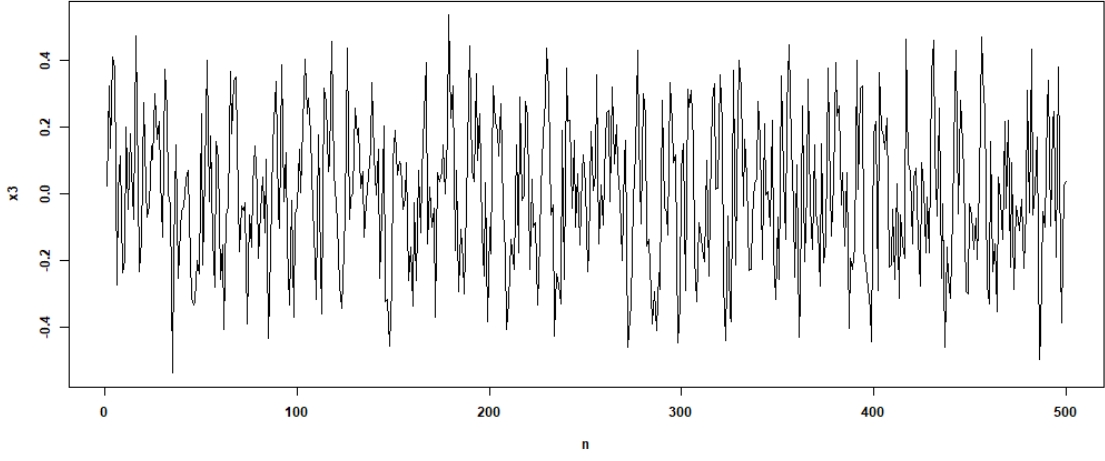
Şekil 5.3. $n = 500$ İçin y Değişkeni



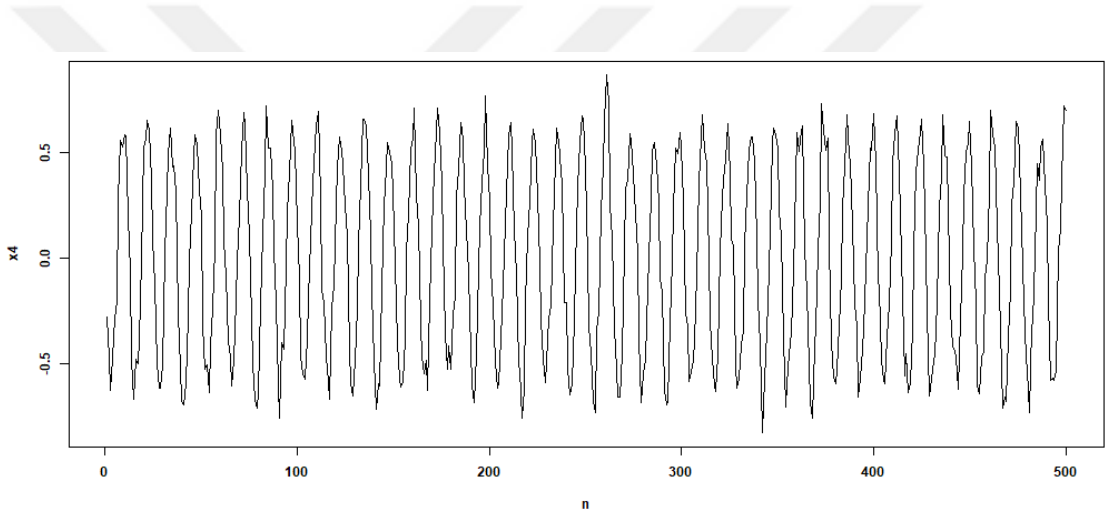
Şekil 5.4. $n = 500$ ve $\sigma = 0.05$ İçin x_1 Değişkeni



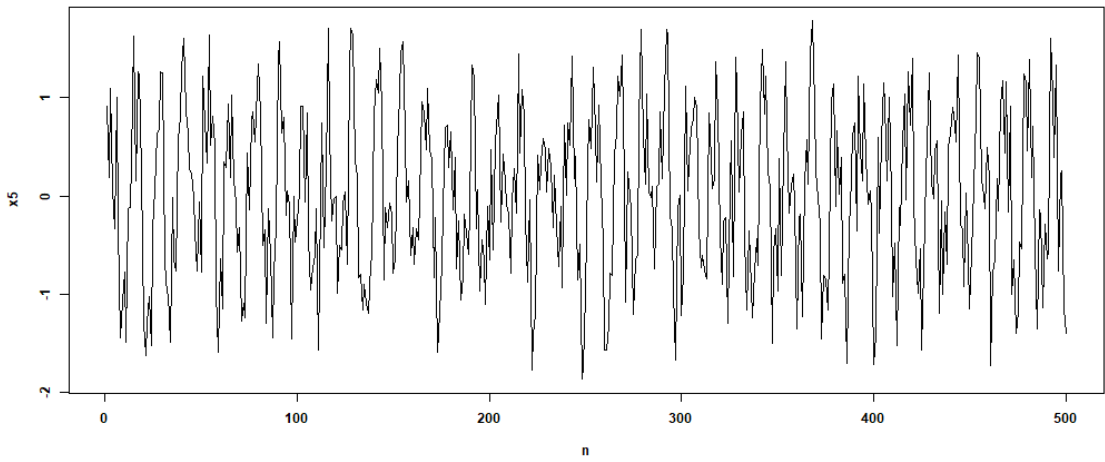
Şekil 5.5. $n = 500$ ve $\sigma = 0.05$ İçin x_2 Değişkeni



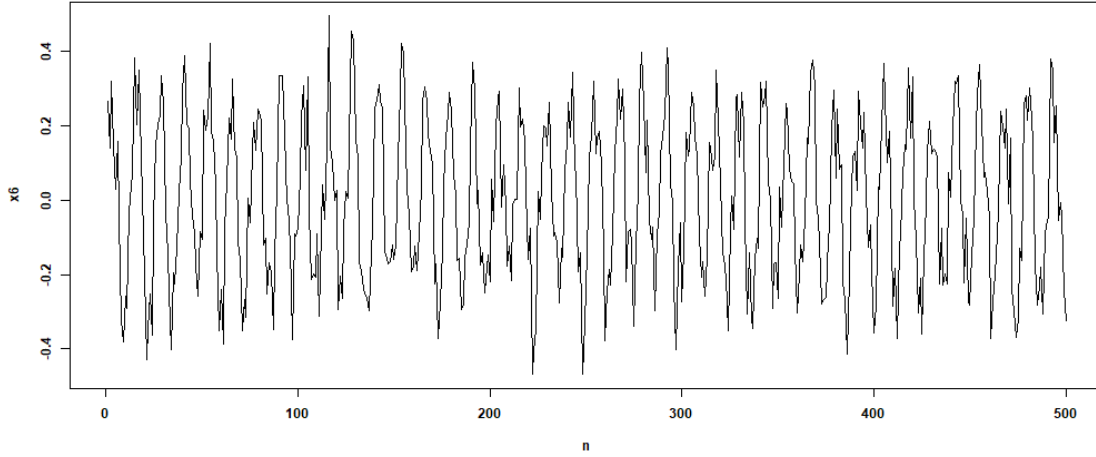
Şekil 5.6. $n = 500$ ve $\sigma = 0.05$ İçin x_3 Değişkeni



Şekil 5.7. $n = 500$ ve $\sigma = 0.05$ İçin x_4 Değişkeni



Şekil 5.8. $n = 500$ ve $\sigma = 0.05$ İçin x_5 Değişkeni



Şekil 5.9. $n = 500$ ve $\sigma = 0.05$ için x_6 Değişkeni

Uygulamada, karşılaştırılan tüm yöntemlerde belli bir standardı yakalamak için hepsinde de aynı sayıda bağımsız bileşen kullanımı tercih edilmiştir. Açıklanan değişken, yapı itibarıyla iki bağımsız bileşenden oluştuğundan dolayı sonuçlar, bir bileşenin ve iki bileşenin olduğu iki farklı durum için ayrı ayrı elde edilmiştir. Öncelikle uniform dağılımdan üretilmiş olan s_2 bağımsız bileşeni, açıklanan değişken içerisinde bir hata terimi gibi düşünülmüş ve tahmin edilen bileşenlerden sadece bir tanesi ile y değişkeni açıklanmaya çalışılmıştır. KEKR ve TBR için toplam varyans açıklama oranı en yüksek bağımsız bileşen seçilirken, BBA-OBM için BBA uygulandıktan sonra elde edilen bağımsız bileşenlerden y değişkeni ile en yüksek ortak bilgi miktarına sahip bağımsız bileşen¹¹ ve BBA-K için de BBA uygulandıktan sonra elde edilen bağımsız bileşenlerden y değişkeni ile en yüksek Pearson korelasyon katsayısına sahip bağımsız bileşen seçilmiştir.

Çizelge 5.2.'de eğitim kümesi için tahmin edilen bağımsız bileşenlerden bir bağımsız bileşen ile y değişkeni açıklanmaya çalışıldığında elde edilen HKO'lar görülmektedir. Çizelgede, BBA-OBM ve BBA-K yöntemlerinin her ikisi de aynı bileşenleri seçtikleri için aynı HKO sonuçlarını vermişlerdir. Çizelge 5.2 genel olarak incelendiğinde BBA'ya dayalı yöntemlerin diğer yöntemlere göre daha düşük HKO'lara sahip oldukları ve örneklem büyüklüğü arttıkça da HKO değerlerinin azaldıkları görülmektedir. En yüksek

¹¹ Çizelgelerde parantez içerisinde hangi bileşenin seçildiği görülmektedir.

HKO değerlerine TBR yöntemi sahiptir. KEKR yöntemi ise özellikle standart sapmanın küçük olduğu durumlarda TBR'ye göre daha düşük HKO'lar vermektedir. Ayrıca BBA'ya dayalı yöntemlerin standart sapmanın büyüklüğünden neredeyse hiç etkilenmedikleri görülmektedir.

Çizelge 5.2. Bir Bağımsız Bileşenli Eğitim Kümesi Hata Kareler Ortalaması Sonuçları

	σ	KEKR	TBR	BBA-OBM	BBA-K
	0.01	3.456687	7.466903	2.454838 (s ₂)	2.454838 (s ₂)
	0.05	3.632650	7.517366	2.454632 (s ₂)	2.454632 (s ₂)
n=100	0.10	4.119004	7.610726	2.458550 (s ₂)	2.458550 (s ₂)
	0.50	4.904762	8.537590	2.456162 (s ₃)	2.456162 (s ₃)
	0.70	5.233858	8.533501	2.459426 (s ₃)	2.459426 (s ₃)
	0.90	6.530640	8.502792	2.459004 (s ₃)	2.459004 (s ₃)
	0.01	3.818067	7.314321	2.157434 (s ₂)	2.157434 (s ₂)
	0.05	3.994806	7.346429	2.157378 (s ₂)	2.157378 (s ₂)
n=500	0.10	4.434819	7.415903	2.157155 (s ₂)	2.157155 (s ₂)
	0.50	6.092425	8.570291	2.861852 (s ₃)	2.861852 (s ₃)
	0.70	6.377712	8.726106	2.157574 (s ₃)	2.157574 (s ₃)
	0.90	6.580801	8.766895	2.157464 (s ₃)	2.157464 (s ₃)
	0.01	3.649667	8.147447	2.027997 (s ₂)	2.027997 (s ₂)
	0.05	3.857672	8.167834	2.028069 (s ₂)	2.028069 (s ₂)
n=1000	0.10	4.362556	8.215545	2.028080 (s ₂)	2.028080 (s ₂)
	0.50	6.671310	9.079196	2.028357 (s ₃)	2.028357 (s ₃)
	0.70	7.307406	9.269358	2.028045 (s ₃)	2.028045 (s ₃)
	0.90	7.816983	9.330156	2.027864 (s ₃)	2.027864 (s ₃)
	0.01	3.389468	8.241597	1.980053 (s ₂)	1.980053 (s ₂)
	0.05	3.641415	8.257634	1.980056 (s ₂)	1.980056 (s ₂)
n=5000	0.10	4.200503	8.298145	1.980064 (s ₂)	1.980064 (s ₂)
	0.50	6.697190	8.895673	1.980079 (s ₃)	1.980079 (s ₃)
	0.70	7.356542	9.087617	1.980060 (s ₃)	1.980060 (s ₃)
	0.90	7.859917	9.149830	1.980054 (s ₃)	1.980054 (s ₃)

Çizelge 5.3'te ise aynı durum için test kümesinden elde edilen HKO'lar görülmektedir. Beklendiği gibi test kümesi HKO sonuçları, eğitim kümesi HKO sonuçlarına göre çok daha yüksektir. Burada da benzer şekilde en düşük HKO'lara BBA'ya dayalı yöntemlerle ulaşılmış ve yöntemlerin HKO sonuçları standart sapma düzeylerinden neredeyse hiç etkilenmemiştir. TBR ise en kötü HKO sonucunu veren yöntem olmuştur.

Çizelge 5.3. Bir Bağımsız Bileşenli Test Kümesi Hata Kareler Ortalaması Sonuçları

	σ	KEKR	TBR	BBA-OBM	BBA-K
	0.01	24.64854	53.51748	6.147735	6.147735
	0.05	24.82909	53.32287	6.147524	6.147524
n=100	0.10	26.83503	53.09843	6.159422	6.159422
	0.50	36.77541	45.31775	6.151068	6.151068
	0.70	39.02795	45.98506	6.160182	6.160182
	0.90	43.21219	46.20633	6.159835	6.159835
	0.01	28.52354	57.39673	7.052714	7.052714
	0.05	29.33675	57.40981	7.052524	7.052524
n=500	0.10	31.86766	57.49611	7.051823	7.051823
	0.50	42.40438	50.55419	10.20879	10.20879
	0.70	43.56209	51.32241	7.053122	7.053122
	0.90	44.25607	51.63916	7.052833	7.052833
	0.01	22.87033	57.15809	7.468976	7.468976
	0.05	24.27931	57.29893	7.469229	7.469229
n=1000	0.10	27.41646	57.49307	7.469283	7.469283
	0.50	41.18531	52.15117	7.469869	7.469869
	0.70	44.20671	52.81335	7.469242	7.469242
	0.90	46.53793	53.08055	7.468623	7.468623
	0.01	19.77993	56.07744	7.525511	7.525511
	0.05	21.32804	56.14802	7.525535	7.525535
n=5000	0.10	24.66415	56.26432	7.525591	7.525591
	0.50	39.56548	51.04816	7.525621	7.525621
	0.70	42.96535	51.70986	7.525564	7.525564
	0.90	45.52991	52.01574	7.525524	7.525524

Her ne kadar BBA'ya dayalı yöntemler diğer yöntemlere göre en iyi HKO sonuçlarını verse de, bu yöntemlerin ortak bilgi miktarını ya da korelasyon katsayısını kullanarak tahmin edilen bağımsız bileşenlerden doğru bağımsız bileşenleri seçip seçmediklerinin sağlanması da çalışma kapsamında ayrıca yapılmıştır.

Çizelge 5.4 ve Çizelge 5.5'te sırasıyla eğitim kümesi ve test kümesi için "BBA'ya dayalı yöntemlerde diğer bağımsız bileşenler seçilmiş olsaydı ne kadarlık bir HKO elde edilirdi?" sorusunun cevabı bulunmaktadır. Her iki çizelgeden de hem eğitim kümesinde hem de test kümesinde en düşük HKO'ya sahip olan bağımsız bileşenlerin seçildiği görülmektedir.

Çizelgelerdeki X işareti, açıklayıcı değişkenlere BBA uygulandığında bulunamayan bağımsız bileşenleri göstermektedir. Bilindiği gibi altı tane açıklayıcı değişken üretildiği için BBA ile altı tane bağımsız bileşen bulunmaya çalışılmış ancak zaman zaman matrislerin tekil olmasından dolayı üç ya da dört tane bağımsız bileşen bulunabilmiştir.

Çizelge 5.4. Diğer Bağımsız Bileşenlerin (Birli) Eğitim Kümesi Hata Kareler Ortalaması Sonuçları

	σ	s_1	s_2	s_3	s_4	s_5	s_6
	0.01	6.110959	2.454838	8.528559	X	X	X
	0.05	6.111191	2.454632	8.528533	X	X	X
n=100	0.10	6.107326	2.458550	8.528480	X	X	X
	0.50	8.528583	6.109611	2.456162	8.54632	X	X
	0.70	8.528512	6.106419	2.459426	X	X	X
	0.90	8.528507	6.106845	2.459004	X	X	X
	0.01	6.650667	2.157434	8.791206	X	X	X
	0.05	6.650710	2.157378	8.791219	X	X	X
n=500	0.10	6.650890	2.157155	8.791262	X	X	X
	0.50	8.158877	6.578577	2.861852	X	X	X
	0.70	8.791192	6.650541	2.157574	X	X	X
	0.90	8.791202	6.650641	2.157464	X	X	X
	0.01	7.424477	2.027997	9.389398	9.393318	X	X
	0.05	7.424433	2.028069	9.389370	X	X	X
n=1000	0.10	7.424480	2.028080	9.389313	X	X	X
	0.50	9.389430	7.424085	2.028357	X	X	X
	0.70	9.389204	7.424624	2.028045	X	X	X
	0.90	9.389480	7.424529	2.027864	X	X	X
	0.01	7.280518	1.980053	9.241444	X	X	X
	0.05	7.280526	1.980056	9.241433	9.250908	X	X
n=5000	0.10	7.280543	1.980064	9.241408	9.250093	X	X
	0.50	9.241400	7.280536	1.980079	9.248420	X	X
	0.70	9.241421	7.280534	1.980060	9.154908	X	X
	0.90	9.241438	7.280523	1.980054	8.177585	X	X

Çizelge 5.5. Diğer Bağımsız Bileşenlerin (Birli) Test Kümesi Hata Kareler Ortalaması Sonuçları

	σ	s_1	s_2	s_3	s_4	s_5	s_6
	0.01	45.47535	6.147735	45.82758	X	X	X
	0.05	45.47914	6.147524	45.82749	X	X	X
n=100	0.10	45.46021	6.159422	45.82820	X	X	X
	0.50	45.46167	45.47106	6.151068	45.68702	X	X
	0.70	45.82823	45.45597	6.160182	X	X	X
	0.90	45.82816	45.45804	6.159835	X	X	X
	0.01	51.99128	7.052714	51.84469	X	X	X
	0.05	51.99151	7.052524	51.84503	X	X	X
n=500	0.10	51.99246	7.051823	51.84555	X	X	X
	0.50	45.69648	51.68062	10.20879	X	X	X
	0.70	51.84447	51.99077	7.053122	X	X	X
	0.90	51.84463	51.99128	7.052833	X	X	X
	0.01	52.26929	7.468976	53.07929	53.61160	X	X
	0.05	52.26908	7.469229	53.07904	X	X	X
n=1000	0.10	52.26930	7.469283	53.07867	X	X	X
	0.50	53.07946	52.26751	7.469869	X	X	X
	0.70	53.07789	52.26991	7.469242	X	X	X
	0.90	53.07992	52.26946	7.468623	X	X	X
	0.01	50.39061	7.525511	52.28167	X	X	X
	0.05	50.39065	7.525535	52.28155	52.38933	X	X
n=5000	0.10	50.39072	7.525591	52.28130	52.53129	X	X
	0.50	52.28121	50.39069	7.525621	52.40044	X	X
	0.70	52.28143	50.39068	7.525564	52.60795	X	X
	0.90	52.28160	50.39063	7.525524	49.11093	X	X

İkinci durumda ise bu kez, iki bağımsız bileşen ile y değişkeni açıklanmaya çalışılmıştır. KEKR ve TBR için toplam varyans açıklama oranı en yüksek iki bağımsız bileşen seçilirken, BBA-OBM için BBA uygulandıktan sonra elde edilen bağımsız bileşenlerden y değişkeni ile en yüksek ortak bilgi miktarına sahip iki bağımsız bileşen ve BBA-K için de BBA uygulandıktan sonra elde edilen bağımsız bileşenlerden y değişkeni ile en yüksek Pearson korelasyon katsayısına sahip iki bağımsız bileşen seçilmiştir.

Çizelge 5.6.'da eğitim kümesi, Çizelge 5.7.'de de test kümesi için iki bileşen ile y değişkeni açıklanmaya çalışıldığında elde edilen HKO'lar görülmektedir. Yine burada da

ortak bilgi miktarına ve korelasyon katsayısına dayalı BBA yöntemleri birbirleriyle aynı bağımsız bileşenleri seçmişlerdir. Eğitim kümesi sonuçlarında iki tane istisna dışında BBA'ya dayalı yöntemlerin daha iyi sonuçlar verdiği görülmektedir. Ancak özellikle standart sapmanın küçük olduğu durumlarda KEKR'nin test kümesinde diğer tüm yöntemlere göre daha düşük HKO'ya sahip olduğu görülmektedir. Standart sapma oranı büyüdükçe ise BBA'ya dayalı yöntemler diğer yöntemlere göre daha düşük HKO'ya sahip olmaktadır. Bir önceki duruma benzer şekilde burada da, BBA'ya dayalı yöntemler standart sapmanın büyümesinden ya da küçülmesinden neredeyse hiç etkilenmemektedir.

Çizelge 5.6. İki Bağımsız Bileşenli Eğitim Kümesi Hata Kareler Ortalaması Sonuçları

	σ	KEKR	TBR	BBA-OBM	BBA-K
	0.01	0.017784	0.017941	0.018619 (s ₁ , s ₂)	0.018619 (s ₁ , s ₂)
	0.05	0.390620	0.490181	0.018645 (s ₁ , s ₂)	0.018645 (s ₁ , s ₂)
n=100	0.10	1.035893	2.462953	0.018698 (s ₁ , s ₂)	0.018698 (s ₁ , s ₂)
	0.50	0.504934	8.007239	0.018595 (s ₂ , s ₃)	0.018595 (s ₂ , s ₃)
	0.70	3.040491	8.003704	0.018666 (s ₂ , s ₃)	0.018666 (s ₂ , s ₃)
	0.90	3.994199	8.001377	0.018671 (s ₂ , s ₃)	0.018671 (s ₂ , s ₃)
	0.01	0.019889	0.020092	0.008448 (s ₁ , s ₂)	0.008448 (s ₁ , s ₂)
	0.05	0.432004	0.552609	0.008435 (s ₁ , s ₂)	0.008435 (s ₁ , s ₂)
n=500	0.10	1.167156	2.668415	0.008392 (s ₁ , s ₂)	0.008392 (s ₁ , s ₂)
	0.50	2.841159	7.890315	0.640776 (s ₂ , s ₃)	0.640776 (s ₂ , s ₃)
	0.70	4.206860	7.961756	0.008462 (s ₂ , s ₃)	0.008462 (s ₂ , s ₃)
	0.90	4.650814	7.996717	0.008451 (s ₂ , s ₃)	0.008451 (s ₂ , s ₃)
	0.01	0.019783	0.019978	0.031538 (s ₁ , s ₂)	0.031538 (s ₁ , s ₂)
	0.05	0.432898	0.546710	0.031566 (s ₁ , s ₂)	0.031566 (s ₁ , s ₂)
n=1000	0.10	1.204199	2.619762	0.031624 (s ₁ , s ₂)	0.031624 (s ₁ , s ₂)
	0.50	2.355232	8.620194	0.031506 (s ₂ , s ₃)	0.031506 (s ₂ , s ₃)
	0.70	3.786068	8.733204	0.041347 (s ₂ , s ₃)	0.041347 (s ₂ , s ₃)
	0.90	4.272730	8.787505	0.031457 (s ₂ , s ₃)	0.031457 (s ₂ , s ₃)
	0.01	0.019513	0.019692	0.009564 (s ₁ , s ₂)	0.009564 (s ₁ , s ₂)
	0.05	0.428079	0.531674	0.009574 (s ₁ , s ₂)	0.009574 (s ₁ , s ₂)
n=5000	0.10	1.211353	2.495824	0.009600 (s ₁ , s ₂)	0.009600 (s ₁ , s ₂)
	0.50	1.748404	8.619305	0.009607 (s ₂ , s ₃)	0.009607 (s ₂ , s ₃)
	0.70	3.287346	8.739617	0.009587 (s ₂ , s ₃)	0.009587 (s ₂ , s ₃)
	0.90	3.824328	8.797209	0.009570 (s ₂ , s ₃)	0.009570 (s ₂ , s ₃)

Çizelge 5.7. İki Bağımsız Bileşenli Test Kümesi Hata Kareler Ortalaması Sonuçları

	σ	KEKR	TBR	BBA-OBM	BBA-K
	0.01	0.020273	0.020447	4.951591	4.951591
	0.05	0.398496	0.512152	4.951816	4.951816
n=100	0.10	1.070673	4.446456	4.951684	4.951684
	0.50	1.247361	54.15470	4.951467	4.951467
	0.70	18.79631	54.47726	4.951489	4.951489
	0.90	29.38164	54.46000	4.951571	4.951571
	0.01	0.019092	0.019288	5.691196	5.691196
	0.05	0.491038	0.675399	5.697830	5.697830
n=500	0.10	1.705129	7.159632	5.706231	5.706231
	0.50	18.82710	58.49202	9.044689	9.044689
	0.70	31.98008	59.41639	5.856496	5.856496
	0.90	36.22648	59.76372	5.923085	5.923085
	0.01	0.018293	0.018478	5.686489	5.686489
	0.05	0.522969	0.711013	5.749748	5.749748
n=1000	0.10	1.915941	6.970380	5.750038	5.750038
	0.50	12.19653	57.72575	5.749674	5.749674
	0.70	23.62813	58.69299	5.755705	5.755705
	0.90	27.75977	59.04177	5.749462	5.749462
	0.01	0.018469	0.018642	5.610067	5.610067
	0.05	0.508358	0.671327	5.610120	5.610120
n=5000	0.10	1.884144	6.137711	5.610232	5.610232
	0.50	7.821385	55.56751	5.610271	5.610271
	0.70	18.76452	56.62624	5.610181	5.610181
	0.90	23.02070	57.00421	5.610095	5.610095

Bir bağımsız bileşenin seçildiği duruma benzer şekilde iki bağımsız bileşenin seçildiği durum için de BBA'ya dayalı yöntemlerin en düşük HKO'ya sahip bağımsız bileşenleri doğru bir şekilde seçip seçmediğinin kontrolü eğitim ve test kümesi için ayrı ayrı yapılmıştır. Çizelge 5.8 ve Çizelge 5.9'dan hem eğitim hem de test kümeleri için doğru bağımsız bileşenlerin seçildiği görülmektedir.

Çizelge 5.8. Diğer Bağımsız Bileşenlerin (İkili) Eğitim Kümesi Hata Kareler Ortalaması Sonuçları

σ	s_1, s_2	s_1, s_3	s_1, s_4	s_1, s_5	s_1, s_6	s_2, s_3	s_2, s_4	s_2, s_5	s_2, s_6	s_3, s_4	s_3, s_5	s_3, s_6
0.01	0.018619	6.092340	X	X	X	2.436219	X	X	X	X	X	X
0.05	0.018645	6.092546	X	X	X	2.435987	X	X	X	X	X	X
n=100	0.10	0.018698	6.088628	X	X	X	2.439852	X	X	X	X	X
0.50	6.091016	2.437567	8.527665	X	X	0.018595	6.109384	X	X	2.455988	X	X
0.70	6.087753	2.440759	X	X	X	0.018666	X	X	X	X	X	X
0.90	6.088174	2.440333	X	X	X	0.018671	X	X	X	X	X	X
0.01	0.008448	6.642219	X	X	X	2.148987	X	X	X	X	X	X
0.05	0.008435	6.642276	X	X	X	2.148943	X	X	X	X	X	X
n=500	0.10	0.008392	6.642499	X	X	X	2.148763	X	X	X	X	X
0.50	5.937801	2.221076	X	X	X	0.640776	X	X	X	X	X	X
0.70	6.642079	2.149112	X	X	X	0.008462	X	X	X	X	X	X
0.90	6.642190	2.149013	X	X	X	0.008451	X	X	X	X	X	X
0.01	0.031538	7.392939	7.389965	X	X	1.996459	2.00308	X	X	9.322623	X	X
0.05	0.031566	7.392867	X	X	X	1.996503	X	X	X	X	X	X
n=1000	0.10	0.031623	7.392856	X	X	X	1.996457	X	X	X	X	X
0.50	7.392579	1.996851	X	X	X	0.031506	X	X	X	X	X	X
0.70	7.392892	1.996312	X	X	X	0.041347	X	X	X	X	X	X
0.90	7.393072	1.996407	X	X	X	0.031456	X	X	X	X	X	X
0.01	0.009564	7.270955	X	X	X	1.97049	X	X	X	X	X	X
0.05	0.009574	7.270952	X	X	X	1.970482	X	X	X	X	X	X
n=5000	0.10	0.009600	7.270943	7.191500	X	X	1.970465	1.927990	X	X	9.082562	X
0.50	7.270929	1.970472	9.239791	X	X	0.009607	7.278947	X	X	1.969740	X	X
0.70	7.270947	1.970473	9.115885	X	X	0.009587	7.102216	X	X	1.398590	X	X
0.90	7.270954	1.970484	7.570509	X	X	0.009570	7.155707	X	X	1.335724	X	X

Çizelge 5.9. Diğer Bağımsız Bileşenlerin (İkili) Test Kümesi Hata Kareler Ortalaması Sonuçları

σ	s_1, s_2	s_1, s_3	s_1, s_4	s_1, s_5	s_1, s_6	s_2, s_3	s_2, s_4	s_2, s_5	s_2, s_6	s_3, s_4	s_3, s_5	s_3, s_6
0.01	4.951591	45.48022	X	X	X	6.287969	X	X	X	X	X	X
0.05	4.951816	45.48126	X	X	X	6.287696	X	X	X	X	X	X
n=100	0.10	4.951684	45.46280	X	X	X	6.297263	X	X	X	X	X
0.50	45.47407	6.291363	45.94245	X	X	4.951467	45.53633	X	X	6.160775	X	X
0.70	45.45869	6.299534	X	X	X	4.951489	X	X	X	X	X	X
0.90	45.46071	6.298391	X	X	X	4.951571	X	X	X	X	X	X
0.01	5.691196	51.81675	X	X	X	6.968285	X	X	X	X	X	X
0.05	5.697830	51.81717	X	X	X	6.968188	X	X	X	X	X	X
n=500	0.10	5.706231	51.81868	X	X	X	6.967788	X	X	X	X	X
0.50	45.35384	7.113007	X	X	X	9.044689	X	X	X	X	X	X
0.70	51.81596	6.968595	X	X	X	5.856496	X	X	X	X	X	X
0.90	51.81658	6.968379	X	X	X	5.923085	X	X	X	X	X	X
0.01	5.686489	52.04049	51.98967	X	X	7.389147	7.413400	X	X	53.69526	X	X
0.05	5.749748	52.04002	X	X	X	7.389258	X	X	X	X	X	X
n=1000	0.10	5.750038	53.66242	X	X	X	7.389131	X	X	X	X	X
0.50	52.03888	7.390059	X	X	X	5.749674	X	X	X	X	X	X
0.70	52.03967	7.388765	X	X	X	5.755705	X	X	X	X	X	X
0.90	52.55353	7.389017	X	X	X	5.749462	X	X	X	X	X	X
0.01	5.610067	50.27925	X	X	X	7.482290	X	X	X	X	X	X
0.05	5.610120	50.27986	X	X	X	7.482265	X	X	X	X	X	X
n=5000	0.10	5.610232	50.27968	48.78236	X	X	10.57458	7.658662	X	X	56.86964	X
0.50	50.27956	7.482206	52.25272	X	X	5.610271	50.69282	X	X	7.522140	X	X
0.70	50.27977	7.482239	51.97404	X	X	5.610181	49.87525	X	X	5.764530	X	X
0.90	50.27990	7.482276	51.23672	X	X	5.610095	49.08201	X	X	7.586027	X	X

5.1.2. Yapay Veri Uygulaması-2

Bu uygulama için $\mathbf{S} = [s_1 \ s_2 \ s_3]$ ve $k = 1, 2, \dots, n$ olmak üzere Eşitlik 5.11, Eşitlik 5.12 ve Eşitlik 5.13'te verilmiş olan dağılımlardan üç farklı bağımsız bileşen üretilmiştir.

$$s_1(k) = 2 \cos(0.08k) \sin(0.006k) \quad (5.11)$$

$$s_2(k) = \text{sign}[\sin(0.03k + 9 \cos(0.01k))] \quad (5.12)$$

$$s_3(k) = \text{uniform} \sim [-1, 1] \quad (5.13)$$

Üretilen bileşenlerin istatistiksel olarak birbirlerinden bağımsız olup olmadıklarının kontrolü için uygulanan Hoeffding Bağımsızlık Testi'nin olasılık değerleri Çizelge 5.10'da özetlenmiştir.

Çizelge 5.10. Yapay Veri Uygulaması-2 İçin Hoeffding Test Sonuçları

Hoeffding Testi	n=100	n=500	n=1000	n=5000
(s ₁ , s ₂)	$p = 0.000$	$p = 0.388$	$p = 0.173$	$p = 0.000$
(s ₁ , s ₃)	$p = 0.279$	$p = 0.525$	$p = 0.458$	$p = 0.608$
(s ₂ , s ₃)	$p = 1.000$	$p = 1.000$	$p = 1.000$	$p = 1.000$

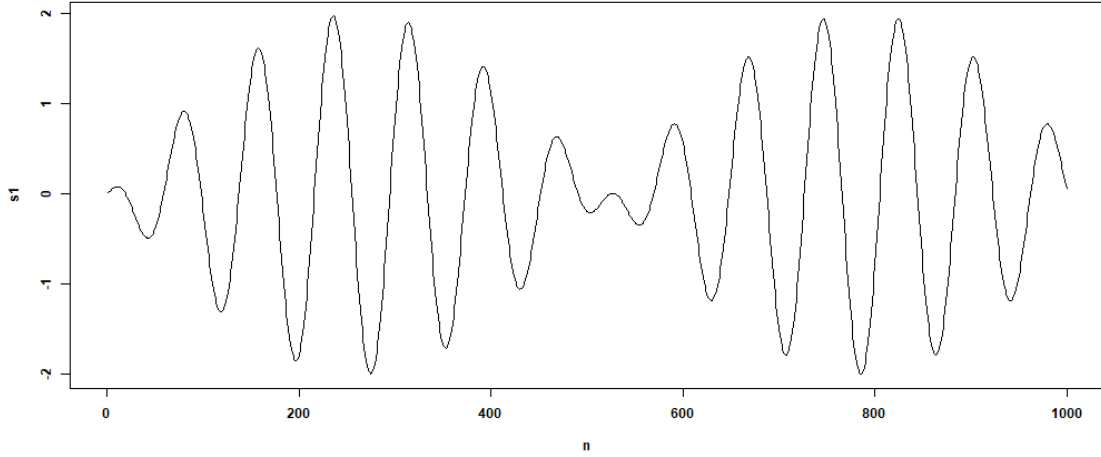
Örneklem büyüklüğünün 100 ve 5000 olduğu durumdaki (s₁, s₂) bileşenleri hariç geriye kalan tüm bileşenlerin test sonuçlarının olasılık değerleri $\alpha = 0.05$ 'ten büyük çıktığı için %95 güven düzeyinde istatistiksel olarak birbirlerinden bağımsız oldukları söylenebilir.

Yapay Veri Uygulaması-1 kısmındaki gibi temelde iki bağımsız bileşen ile oluşturduğu düşünülen açıklanan değişkenin gerçek değerleri, $\mathbf{y} = 2\mathbf{s}_1 + 7\mathbf{s}_2$ eşitliği ile üretilmiştir. Sonrasında, bu iki bağımsız değişkenin doğrudan gözlenemediği, onlar yerine üç bağımsız bileşenin etkisiyle oluşmuş olan beş tane açıklayıcı değişkenin gözlendiği bir durum oluşturulmuştur. $\mathbf{X} = \mathbf{A}\mathbf{S}$ eşitliği kullanılarak $\mathbf{X} = [\mathbf{x}_1 \ \mathbf{x}_2 \ \mathbf{x}_3 \ \mathbf{x}_4 \ \mathbf{x}_5]$

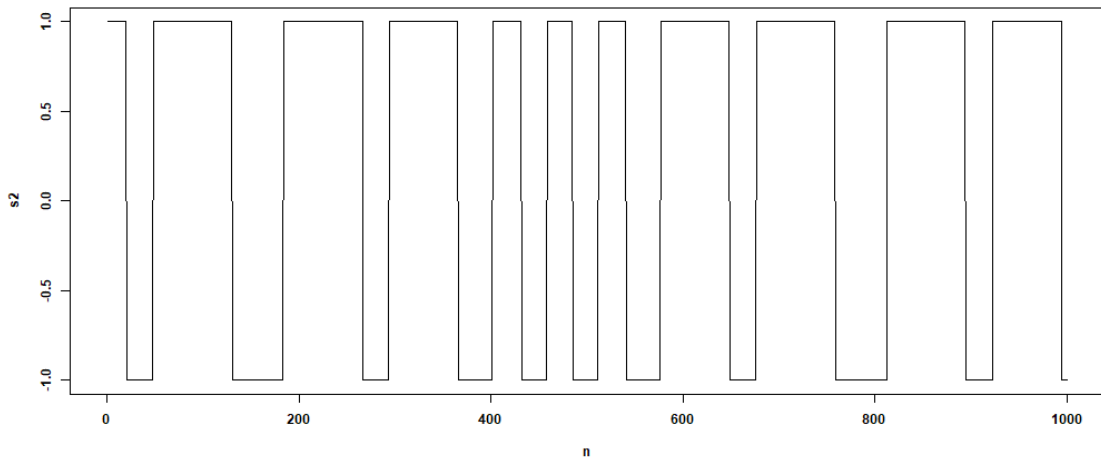
şeklindeki açıklayıcı değişkenlerin verileri üretilmiştir. Üretim aşamasında aşağıdaki A matrisi değerleri kullanılmıştır.

$$A = \begin{bmatrix} 0.86 & -0.55 & 0.17 & -0.33 & 0.89 \\ 0.79 & 0.65 & 0.32 & 0.12 & -0.97 \\ 0.67 & 0.46 & -0.28 & 0.27 & -0.74 \end{bmatrix} \quad (5.14)$$

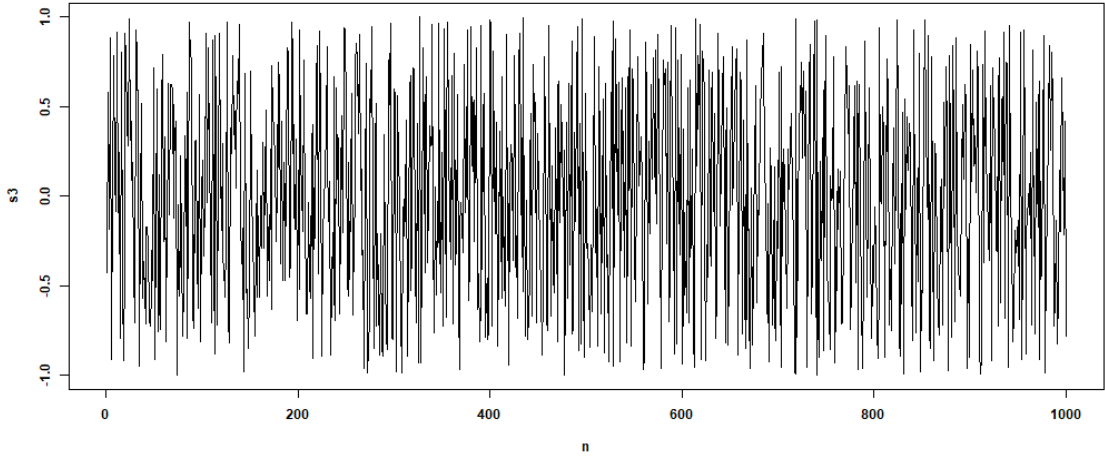
Ayrıca üretilen açıklayıcı değişkenlere normal dağılımdan üretilmiş olan 0 ortalama ve farklı standart sapmalara sahip hata terimleri de eklenmiştir. Örnek olarak, $n = 1000$ için üretilen $s_1, s_2, s_3, y, X_1, X_2, X_3, X_4, X_5$ 'in grafikleri sırasıyla Şekil 5.10'dan itibaren verilmiştir.



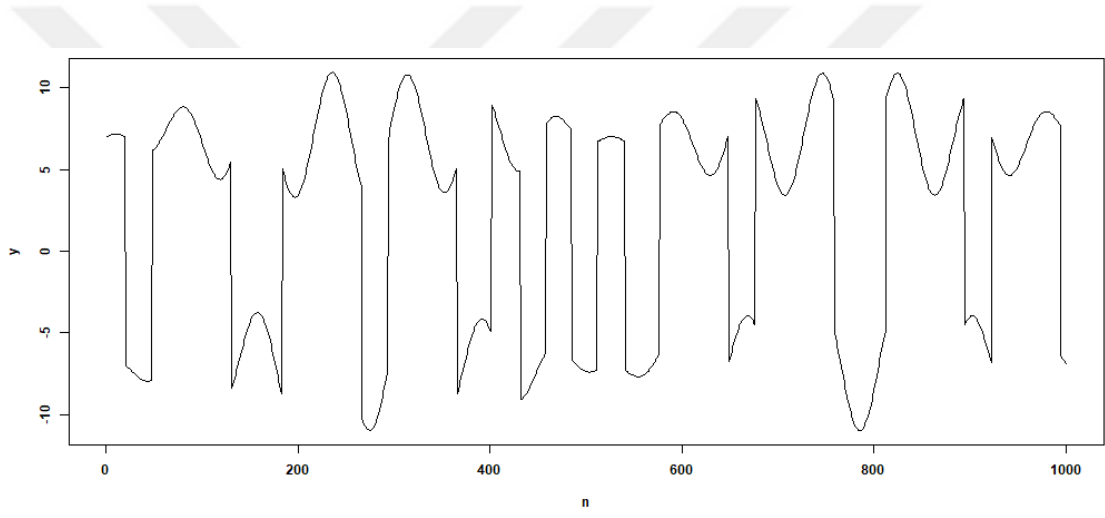
Şekil 5.10. $n = 1000$ İçin s_1 Bileşeni



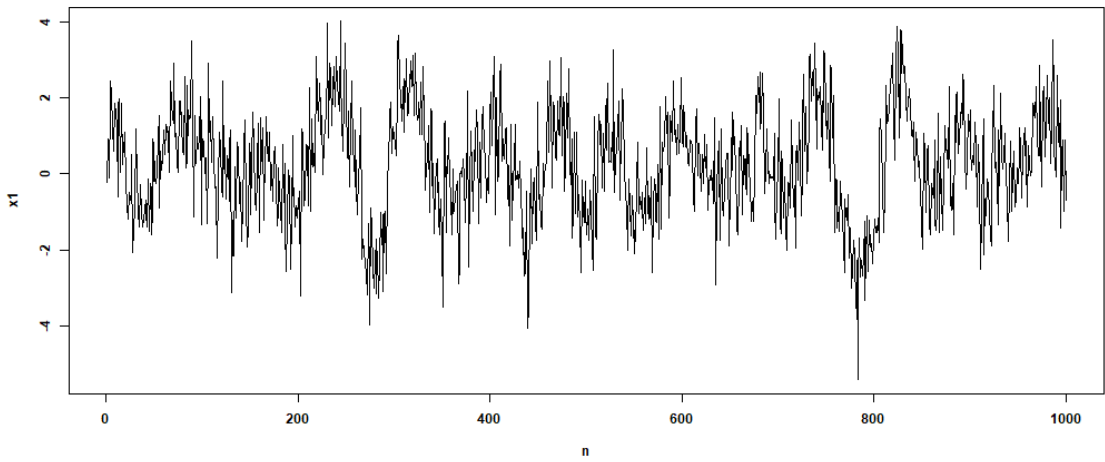
Şekil 5.11. $n = 1000$ İçin s_2 Bileşeni



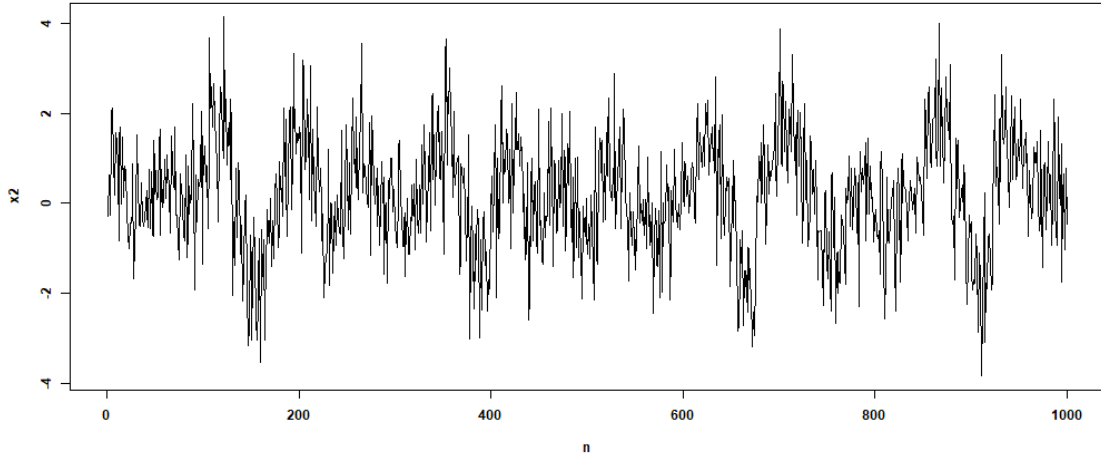
Şekil 5.12. $n = 1000$ İçin s_3 Bileşeni



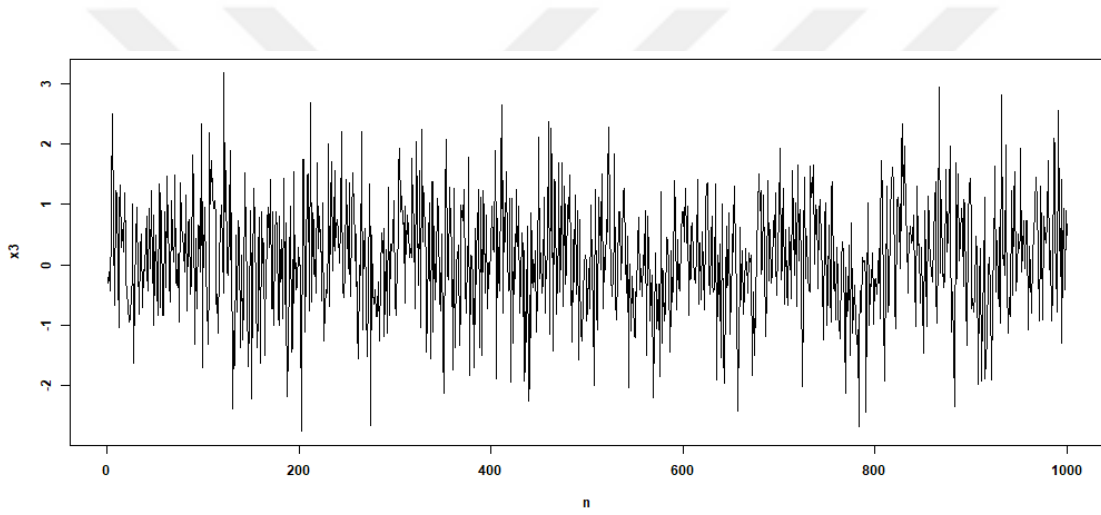
Şekil 5.13. $n = 1000$ İçin y Değişkeni



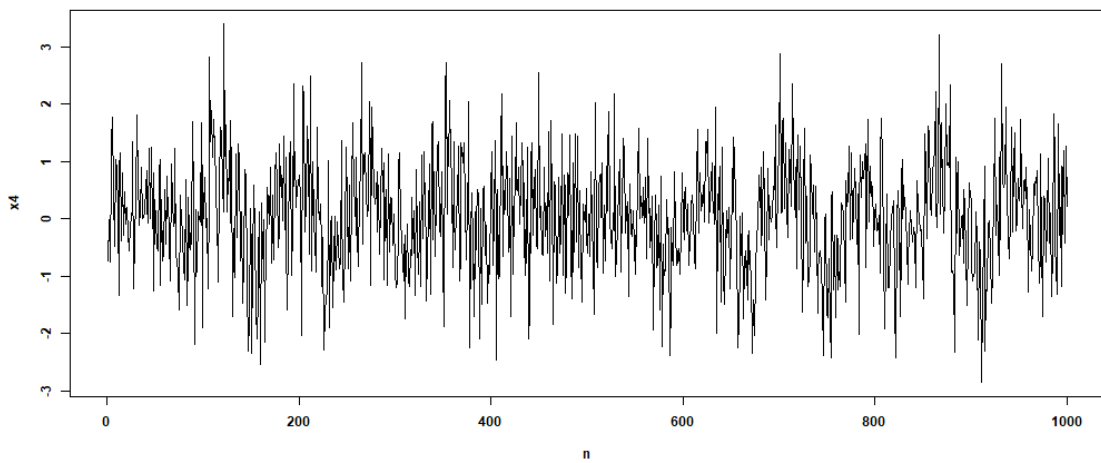
Şekil 5.14. $n = 1000$ ve $\sigma = 0.9$ İçin x_1 Değişkeni



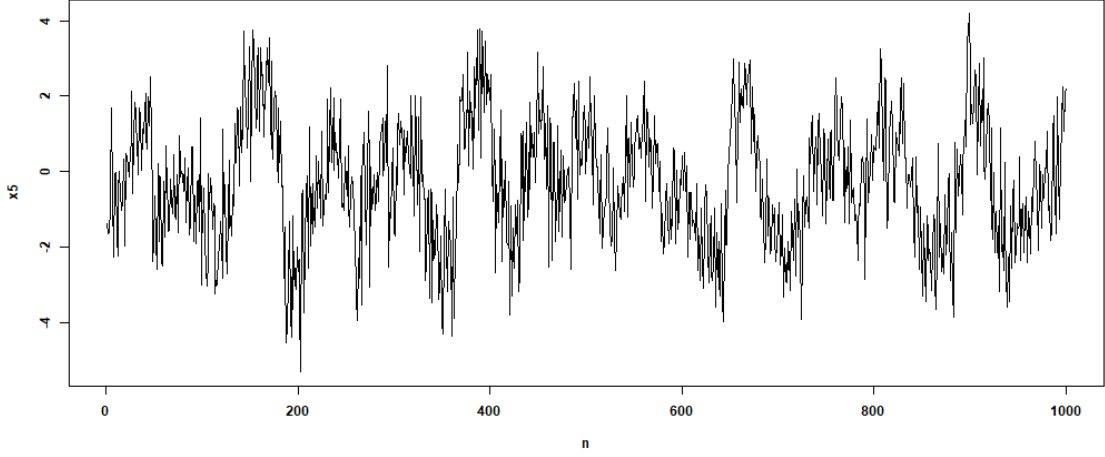
Şekil 5.15. $n = 1000$ ve $\sigma = 0.9$ İçin x_2 Değişkeni



Şekil 5.16. $n = 1000$ ve $\sigma = 0.9$ İçin x_3 Değişkeni



Şekil 5.17. $n = 1000$ ve $\sigma = 0.9$ İçin x_4 Değişkeni



Şekil 5.18. $n = 1000$ ve $\sigma = 0.9$ İçin x_5 Değişkeni

Açıklanan değişken, yapı itibarıyla iki bağımsız bileşenden oluştuğundan dolayı sonuçlar, iki bileşenin ve üç bileşenin olduğu iki farklı durum için ayrı ayrı elde edilmiştir. Öncelikle uniform dağılımdan üretilmiş olan s_3 bağımsız bileşeni, açıklanan değişkeni oluşturan bileşenler içerisinde yer almadığı için tahmin edilen bileşenlerden sadece iki tanesi ile y değişkeni açıklanmaya çalışılmıştır.

Yapay Veri Uygulaması-1'e benzer şekilde KEKR ve TBR için toplam varyans açıklama oranı en yüksek iki bileşen seçilirken, BBA-OBM için BBA uygulandıktan sonra elde edilen bağımsız bileşenlerden y değişkeni ile en yüksek ortak bilgi miktarına sahip iki bileşen ve BBA-K için de BBA uygulandıktan sonra elde edilen bağımsız bileşenlerden y ile en yüksek Pearson korelasyon katsayısına sahip iki bileşen seçilmiştir.

Sırasıyla Çizelge 5.11.'de eğitim kümesi, Çizelge 5.12'de de test kümesi için iki bağımsız bileşen kullanılarak y açıklanmaya çalışıldığında elde edilen HKO sonuçları görülmektedir.

Çizelge 5.11. İki Bağımsız Bileşenli Eğitim Kümesi Hata Kareler Ortalaması Sonuçları

σ	KEKR	TBR	BBA-OBM	BBA-K
0.01	0.016944	0.022497	1.255625 (s ₁ , s ₂)	1.255625 (s ₁ , s ₂)
0.05	0.036883	0.044585	1.242015 (s ₁ , s ₂)	1.242015 (s ₁ , s ₂)
n=100	0.10	0.102558	1.261418 (s ₁ , s ₂)	1.261418 (s ₁ , s ₂)
	0.50	2.296172	1.242161 (s ₂ , s ₃)	1.242161 (s ₂ , s ₃)
	0.70	3.089904	1.241832 (s ₂ , s ₃)	1.241832 (s ₂ , s ₃)
	0.90	3.440282	1.242182 (s ₂ , s ₃)	1.242182 (s ₂ , s ₃)
	0.01	0.051238	0.100803 (s ₁ , s ₂)	0.100803 (s ₁ , s ₂)
	0.05	0.442286	0.428989	0.100701 (s ₁ , s ₂)
n=500	0.10	1.380698	0.090890 (s ₁ , s ₂)	0.090890 (s ₁ , s ₂)
	0.50	1.621594	0.090952 (s ₁ , s ₂)	0.090952 (s ₁ , s ₂)
	0.70	1.968866	0.091029 (s ₁ , s ₃)	0.091029 (s ₁ , s ₃)
	0.90	3.632831	0.092927 (s ₂ , s ₃)	0.092927 (s ₂ , s ₃)
	0.01	0.058772	0.073010	0.002700 (s ₁ , s ₂)
	0.05	0.424898	0.449717	0.002694 (s ₁ , s ₂)
n=1000	0.10	1.410989	0.047931 (s ₁ , s ₂)	0.047931 (s ₁ , s ₂)
	0.50	1.832579	0.047986 (s ₁ , s ₂)	0.047986 (s ₁ , s ₂)
	0.70	2.171710	0.048639 (s ₂ , s ₃)	0.048639 (s ₂ , s ₃)
	0.90	3.885324	0.048067 (s ₂ , s ₃)	0.048067 (s ₂ , s ₃)
	0.01	0.066465	0.084075	0.007461 (s ₁ , s ₂)
	0.05	0.400670	0.427927	0.007459 (s ₁ , s ₂)
n=5000	0.10	1.302667	0.007452 (s ₁ , s ₂)	0.007452 (s ₁ , s ₂)
	0.50	1.712182	0.007459 (s ₁ , s ₂)	0.007459 (s ₁ , s ₂)
	0.70	2.266888	0.007461 (s ₂ , s ₃)	0.007461 (s ₂ , s ₃)
	0.90	4.039988	0.007460 (s ₂ , s ₃)	0.007460 (s ₂ , s ₃)

Çizelge 5.11'e göre ortak bilgi miktarına ve korelasyon katsayısına dayalı BBA yöntemleri, aynı bağımsız bileşenleri seçmişlerdir. Eğitim kümesi sonuçlarında birkaç istisna dışında genel olarak BBA'ya dayalı yöntemlerin daha iyi sonuçlar verdiği görülmektedir. Özellikle küçük örneklem ve küçük standart sapma durumlarında KEKR'nin test kümesinde diğer tüm yöntemlere göre çok daha düşük HKO'lara sahip olduğu görülmektedir. Standart sapma oranı büyüdükçe ise BBA'ya dayalı yöntemler, diğer tüm yöntemlere göre çok daha düşük HKO'ya sahip olmaktadır. Yapay Veri Uygulaması-1'e benzer şekilde burada da BBA'ya dayalı yöntemler standart sapmanın büyümesinden ya da küçülmesinden neredeyse hiç etkilenmemektedir.

Çizelge 5.12. İki Bağımsız Bileşenli Test Kümesi Hata Kareler Ortalaması Sonuçları

	σ	KEKR	TBR	BBA-OBM	BBA-K
n=100	0.01	0.083172	0.112215	11.15032	11.15032
	0.05	0.105746	0.136130	11.09768	11.09768
	0.10	0.145756	0.198735	11.20858	11.20858
	0.50	14.68439	23.70662	11.09777	11.09777
	0.70	19.97880	26.32434	11.09731	11.09731
	0.90	22.56199	27.41845	11.09794	11.09794
n=500	0.01	0.305976	0.386194	1.906702	1.906702
	0.05	0.414992	0.812548	1.906478	1.906478
	0.10	1.821483	2.066046	1.471180	1.471180
	0.50	8.376895	34.42564	1.470846	1.470846
	0.70	14.28411	41.84784	1.470399	1.470399
	0.90	18.53678	43.60336	1.460045	1.460045
n=1000	0.01	0.299670	0.398619	3.813122	3.813122
	0.05	0.629332	0.722241	3.810675	3.810675
	0.10	1.566237	1.800513	4.421576	4.421576
	0.50	8.183824	31.78308	4.421943	4.421943
	0.70	14.10129	38.34169	4.423378	4.423378
	0.90	18.15798	39.83737	4.421987	4.421987
n=5000	0.01	0.350840	0.469152	3.439800	3.439800
	0.05	0.676305	0.791030	3.439825	3.439825
	0.10	1.569307	1.833755	3.440042	3.440042
	0.50	7.752844	32.96077	3.440097	3.440097
	0.70	14.27249	40.56763	3.440013	3.440013
	0.90	19.10308	42.37709	3.440106	3.440106

İkinci durumda ise bu kez üç bağımsız bileşen ile y açıklanmaya çalışılmıştır. KEKR ve TBR için varyans açıklama oranı en yüksek üç bağımsız bileşen seçilirken, BBA-OBM için BBA uygulandıktan sonra elde edilen bağımsız bileşenlerden y ile en yüksek ortak bilgi miktarına sahip üç bağımsız bileşen ve BBA-K için de BBA uygulandıktan sonra elde edilen bağımsız bileşenlerden y değişkeni ile en yüksek Pearson korelasyon katsayısına sahip üç bağımsız bileşen seçilmiştir.

Çizelge 5.13. Üç Bağımsız Bileşenli Eğitim Kümesi Hata Kareler Ortalaması Sonuçları

σ	KEKR	TBR	BBA-OBM	BBA-K
0.01	0.001224	0.001288	1.231171 (s ₁ , s ₂ , s ₃)	0.024455 (s ₁ , s ₂ , s ₄)
0.05	0.004484	0.034000	1.216509 (s ₁ , s ₂ , s ₃)	0.025506 (s ₁ , s ₂ , s ₄)
n=100	0.10	0.008255	0.067086	1.238486 (s ₁ , s ₂ , s ₃)
	0.50	0.602039	1.218103	1.216652 (s ₂ , s ₃ , s ₄)
	0.70	1.029374	1.978205	1.216323 (s ₂ , s ₃ , s ₄)
	0.90	1.342114	2.461043	1.216677 (s ₂ , s ₃ , s ₄)
	0.01	0.016086	0.016319	0.052920 (s ₁ , s ₂ , s ₄)
	0.05	0.172001	0.409775	0.052932 (s ₁ , s ₂ , s ₄)
n=500	0.10	0.072303	1.471686	0.079715 (s ₁ , s ₂ , s ₄)
	0.50	1.114136	1.555522	0.079702 (s ₁ , s ₂ , s ₃)
	0.70	1.576252	2.003966	0.091028 (s ₁ , s ₂ , s ₅)
	0.90	1.917936	2.310199	0.079299 (s ₁ , s ₂ , s ₃)
	0.01	0.016249	0.016465	0.002053 (s ₁ , s ₂ , s ₄)
	0.05	0.225033	0.415270	0.002091 (s ₁ , s ₂ , s ₄)
n=1000	0.10	0.029950	1.599915	0.000145 (s ₁ , s ₂ , s ₄)
	0.50	1.237480	1.835483	0.000147 (s ₁ , s ₂ , s ₃)
	0.70	1.770267	2.351706	0.048639 (s ₂ , s ₃ , s ₅)
	0.90	2.169046	2.696856	0.000146 (s ₁ , s ₂ , s ₃)
	0.01	0.014892	0.015097	0.003102 (s ₁ , s ₂ , s ₃)
	0.05	0.219146	0.381759	0.003101 (s ₁ , s ₂ , s ₃)
n=5000	0.10	0.013909	1.501222	0.003090 (s ₁ , s ₂ , s ₃)
	0.50	1.206893	1.819525	0.004095 (s ₁ , s ₂ , s ₅)
	0.70	1.770843	2.378287	0.006742 (s ₂ , s ₃ , s ₅)
	0.90	2.194623	2.756442	0.003097 (s ₂ , s ₃ , s ₄)

Çizelge 5.13'te eğitim kümesi için üç bağımsız bileşen ile y değişkeni açıklanmaya çalışıldığında elde edilen HKO sonuçları görülmektedir. Çizelgeye göre BBA-OBM ve BBA-K genel olarak birbirlerinden farklı bileşenleri seçmişlerdir. Küçük örneklem ve küçük standart sapma durumlarında KEKR'nin iyi sonuçlar verdiği gözlenmektedir. Genel olarak bakıldığında ise BBA-K'nın diğer yöntemlere göre daha düşük HKO değerleri verdiği; ancak örneklem büyüklüğü arttıkça BBA-OBM'nin BBA-K ile aynı bağımsız bileşenleri seçerek düşük HKO'lara sahip olduğu görülmektedir. Yine burada da en yüksek HKO değerlerini veren TBR yöntemi olmuştur.

Çizelge 5.14'te ise aynı durum için test kümesinden elde edilen HKO'lar görülmektedir. Burada da benzer şekilde küçük örneklemelerde en düşük HKO'lara KEKR ile ulaşılmıştır.

Örnekleme büyüklüğü arttıkça en düşük HKO'lara diğer yöntemlere göre BBA-OBM yöntemi ile daha fazla ulaşılmıştır. Ayrıca BBA temelli yöntemler değişen standart sapma büyüklüklerinden hem eğitim kümesinde hem de test kümesinde neredeyse hiç etkilenmemişlerdir.

Çizelge 5.14. Üç Bağımsız Bileşenli Test Kümesi Hata Kareler Ortalaması Sonuçları

σ	KEKR	TBR	BBA-OBM	BBA-K	
0.01	0.001404	0.001488	12.88171	9.625957	
0.05	0.008174	0.045413	12.86442	9.588943	
n=100	0.10	0.042073	0.298296	12.88658	9.681084
	0.50	2.975597	7.309540	12.86449	9.588816
	0.70	5.315692	12.53870	12.86404	9.588819
	0.90	7.065580	15.98084	12.86467	9.588950
	0.01	0.015793	0.016070	1.790610	3.186457
	0.05	0.449246	0.469247	1.790500	3.194507
n=500	0.10	0.568653	3.792492	1.621020	2.858515
	0.50	8.804613	12.05879	1.621123	2.857900
	0.70	12.04246	15.49668	1.470044	2.857145
	0.90	14.80160	17.81355	1.623970	2.839662
	0.01	0.014416	0.014591	3.826046	4.123393
	0.05	0.442514	0.381501	3.823195	4.123810
n=1000	0.10	0.207690	2.507732	4.221204	4.221204
	0.50	8.588046	12.37666	4.221535	4.221535
	0.70	11.72564	15.81909	4.221941	4.221941
	0.90	14.59390	18.08263	4.221745	4.221745
	0.01	0.013783	0.013968	3.033960	3.033960
	0.05	0.358849	0.366611	3.033961	3.033961
n=5000	0.10	0.091001	2.336479	3.034005	3.034005
	0.50	8.131694	11.86020	3.255576	3.255576
	0.70	11.29361	15.48831	3.387212	3.387212
	0.90	14.25350	17.91040	3.033969	3.033969

5.2. Gerçek Veri Uygulamaları

Gerçek veri uygulamalarının ilk üçünde, farklı alanlardan güçlü, zayıf ve orta düzeyde çoklu bağlantıya sahip olan veri kümeleri üzerinde çalışılmıştır. Bu uygulamalar için veri kümelerinin 3'te 2'si eğitim, 3'te 1'i ise test kümesi olarak ayrılarak önerilen yöntemin ve diğer yöntemlerin tahmin ve öngörü performansları değerlendirilmiştir. Ayrıca çoklu bağlantının açıklayıcı değişkenlerin standart hatalarını artırdığı bilindiği için dördüncü

bir uygulama olarak da önerilen yöntem ile deęişkenlerin tahmin edilen regresyon katsayılarının standart hataları incelenmiş ve mevcut yöntemler ile karşılaştırılması yapılmıştır.

5.2.1. Gerçek Veri Uygulaması-1

İlgili veri North Carolina State Üniversitesinden Dr. Rick Linthurst'un (1979) doktora tezinden alınmıştır. Çalışmanın amacı, Korku Burnu Halici'nde göl otu olan Spartina alterniflor'nun havaya ilişkin biyokütle üretimini etkileyen önemli özellikleri belirlemektir. Her bir bitki örtüsü türü (toplam 45 örnek) için rasgele 5 mevkiden alınan toprağın alt katmanına ait örnekler 14 toprak fizikokimyasalı için aydan aya analiz edilmiştir. Ayrıca her örnek mevki için yer üstü biyokütlesi (BIO- açıklanan deęişken) de her ay ölçülmüştür (Rawlings, Pantula ve Dickey, 2001).

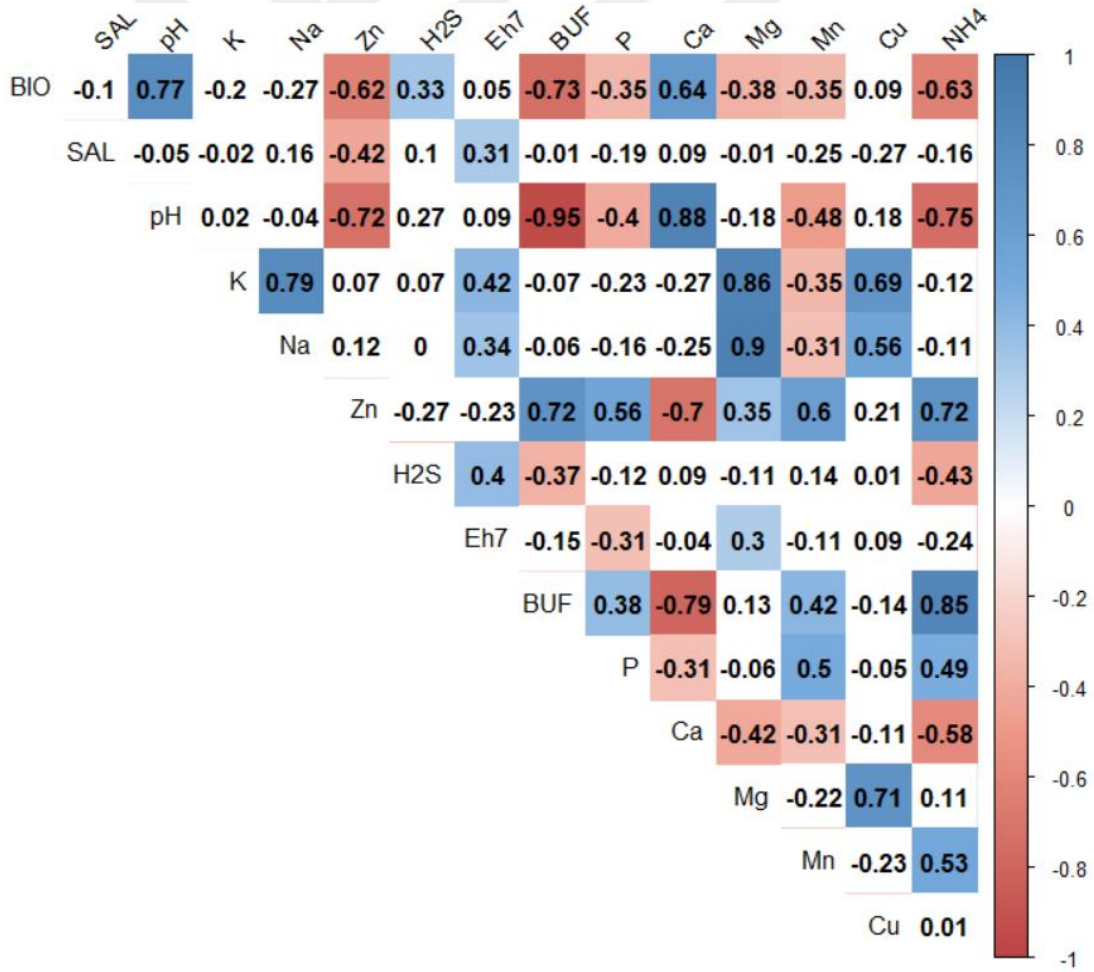
Analizde kullanılan açıklayıcı deęişkenler ve kısaltmaları aşağıdaki gibi sıralanmıştır:

- Tuzluluk oranı (SAL)
- Sudaki çözünmüş pH değeri (pH)
- Potasyum yoğunluğu (p)
- Sodyum yoğunluğu (Na)
- Çinko yoğunluğu (Zn)
- Serbest sülfür molekülleri (H₂S)
- pH 7'de Redox potansiyeli (Eh₇)
- pH 6.6'da yağ asitliği (BUF)
- Fosfor yoğunluğu (P)
- Kalsiyum yoğunluğu (Ca)
- Magnezyum yoğunluğu (Mn)
- Manganez yoğunluğu (Mn)
- Bakır yoğunluğu (Cu)

- Amonyum yoğunluğu (NH₄)

Çizelge 5.15'te analizde kullanılan değişkenler arasındaki korelasyonlar sıcaklık grafiği şeklinde görülmektedir. İstatistiksel olarak %5 anlamlılık düzeyinde önemli bulunan korelasyonlar renkli bir şekilde gösterilmiştir. Korelasyon değerleri üzerindeki rengin koyu maviye yakın olması değişkenler arası pozitif yönlü güçlü bir ilişki olduğunu gösterirken, koyu turuncuya yakın olması da negatif yönlü güçlü bir ilişki olduğunu göstermektedir. Şekle göre en yüksek ilişkiye sahip değişken çifti -0.95'lik bir ilişkiyle pH ve BUF değişkenleridir. Bu ilişkiler haricinde de yüksek ilişkilerin olduğu göze çarpmaktadır.

Çizelge 5.15. Gerçek Veri Uygulaması-1'de Değişkenler Arası Korelasyon Matrisi



Çizelge 5.16. Gerçek Veri Uygulaması-1’de Eğitim ve Test Kümeleri İçin VIF ve Koşul İndeksi Değerleri

Değişkenler	Eğitim Kümesi (2/3)		Test Kümesi (1/3)	
	VIF	Koşul İndeksi	VIF	Koşul İndeksi
SAL	5.16	3.48	30.10	3.71
pH	71.50	4.99	1200.00	6.08
K	46.50	6.85	39.10	9.95
Na	9.76	8.76	359.00	11.62
Zn	16.30	14.07	42.00	15.50
H2S	3.65	16.97	96.90	20.18
EH7	2.40	21.79	11.90	21.81
BUF	33.30	29.69	676.00	40.79
P	1.97	39.51	40.00	47.90
Ca	18.40	46.37	544.00	79.43
Mg	49.00	56.51	865.00	104.35
Mn	8.02	95.83	41.7	187.64
Cu	6.42	110.37	40.0	625.38
NH4	7.98	318.09	36.8	1310.27

Her ne kadar korelasyon matrisinde yüksek ilişkiler tespit edilmiş olsa da çoklu bağlantı sorununun varlığından emin olmak ve ne derece olduğunu belirlemek için her iki veri kümesi için de VIF ve koşul indeksi değerleri hesaplanmıştır. Her iki veri kümesinde de Çizelge 5.16’ya göre 30’un çok daha üzerinde koşul indeksi değerleri (318.09 ve 1310.27) ve 10’nun çok daha üzerinde VIF değerleri (49, 1200) bulunduğu için açıklayıcı değişkenler arası güçlü çoklu bağlantı olduğu söylenebilmektedir.

Güçlü çoklu bağlantının varlığı belirlendikten sonra KEKR, TBR, BBA-OBM ve BBA-K analizleri veriye uygulanmıştır. Yöntemler arasında bir standardı yakalamak ve çok fazla bilgi kaybını önlemek için tüm yöntemlerde tahmin edilen bağımsız bileşenlerden bir tanesi analiz dışı bırakılmıştır. Bu bileşen; KEKR ve TBR için en düşük açıklayıcılık gücüne sahip bileşen, BBA-OBM’de açıklanan değişken ile en düşük ortak bilgi

miktarına sahip bileşen (bkz. Çizelge 5.17) ve BBA-K’da ise açıklanan değişken ile en düşük Pearson korelasyon katsayısına sahip bileşendir.

Çizelge 5.17. Bağımsız Bileşenler ile Açıklanan Değişken Arasındaki Ortak Bilgi Miktarı ve Korelasyon Katsayıları

Bağımsız Bileşenler	Ortak Bilgi Miktarı	Korelasyon
s ₁	0.213476	-0.0027868
s ₂	0.309705	0.2370226
s ₃	0.505721	-0.1834112
s ₄	0.341213	0.4434943
s ₅	0.291787	0.2455409
s ₆	0.213659	0.1970826
s ₇	0.249830	-0.3787109
s ₈	0.362441	0.2809784
s ₉	0.225376	0.1285829
s ₁₀	0.159472	0.0718991
s ₁₁	0.254771	0.2246259
s ₁₂	0.200049	0.2417396
s ₁₃	0.239295	-0.3034499
s ₁₄	0.267835	-0.1916895

Analizler sonucu elde edilen HKO sonuçları Çizelge 5.18’de görülmektedir. Buna göre eğitim kümesinde en düşük HKO, KEKR ile yakalanırken test kümesinde en düşük HKO, BBA-OBM yöntemi ile yakalanmıştır.

Çizelge 5.18. Güçlü Çoklu Bağlantı Durumu için Hata Kareler Ortalaması Sonuçları

	KEKR (13B)	TBR (13 TB)	BBA-OBM (13 BB – S10)	BBA-K (13 BB – S1)
Eğitim-HKO	55372.98	55772.79	57531.77	55375.61
Test-HKO	249040.4	254534.60	248143.80	256935.00

5.2.2. Gerçek Veri Uygulaması-2

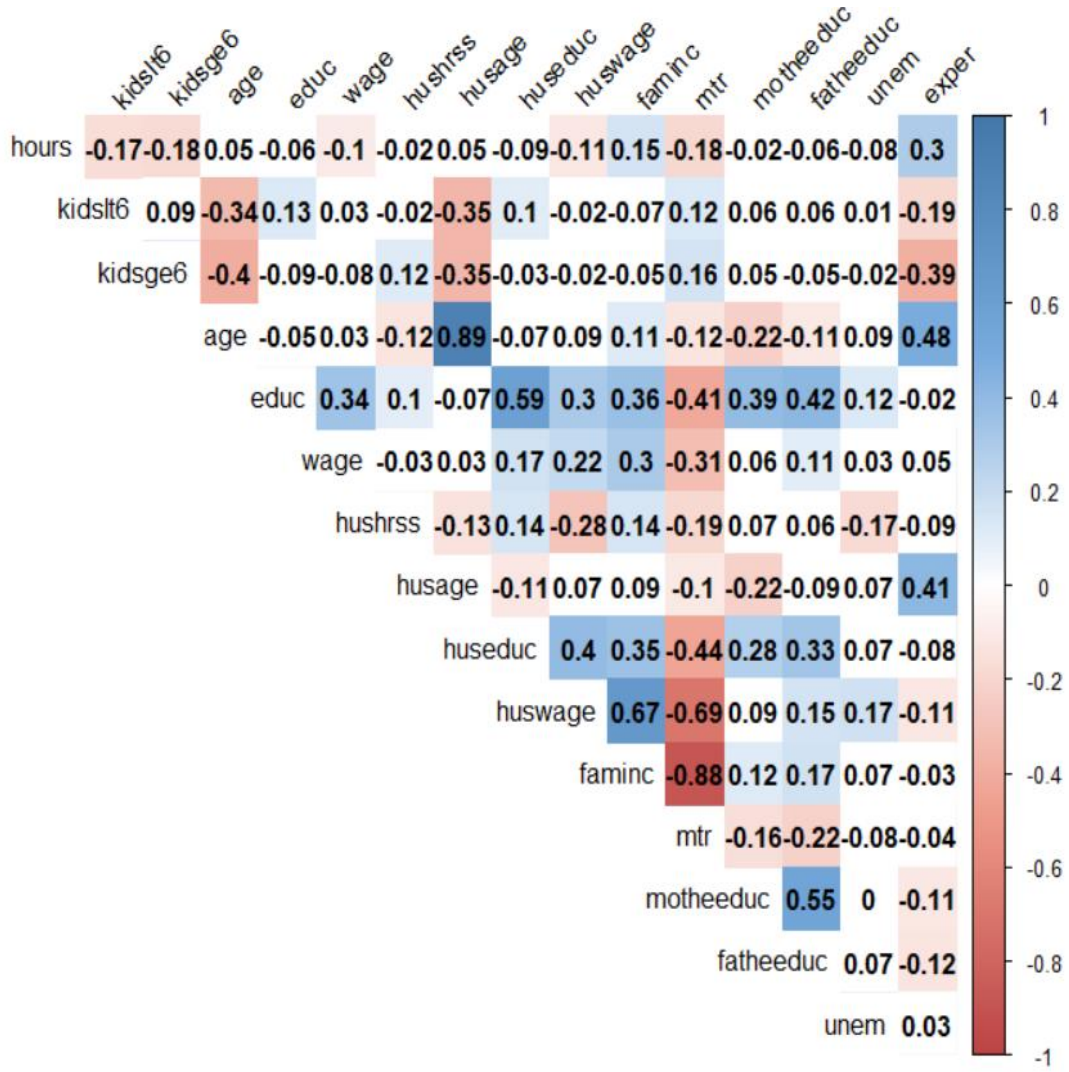
İlgili veri, Mroz (1987) tarafından çeşitli sosyoekonomik değişkenlerin evli kadınların işgücü piyasasındaki çalışma süreleri (hours) üzerindeki etkilerini belirlemeye çalışmak amacıyla toplanmıştır.

Veriler, 1975 yılı için 753 evli kadına ait yatay-kesit verileridir. Ancak içlerinden 329 tanesi çalışmadığı için analizde sadece 429 adet evli ve çalışan kadınların verileri kullanılmıştır.

Analizde kullanılan açıklayıcı değişkenler ve kısaltmaları aşağıdaki gibi sıralanmıştır:

- 6 yaşın altındaki çocuk sayısı (kidslt6)
- 6 ve 18 yaşları arasındaki çocuk sayısı (kidsge6)
- Yıl bazında kadının yaşı (age)
- Kazançlardan tahmin edilen ücret (wage)
- Kocanın saatlik çalışma süresi (hushrs)
- Kocanın yıl bazında eğitim süresi (huseduc)
- Kocanın saatlik ücreti (huswage)
- 1975'te aile geliri (faminc)
- Kadın için federal marjinal vergi oranı (mtr)
- Annenin yıl bazında eğitim süresi (motheeduc)
- Babanın yıl bazında eğitim süresi (Fatheeduc)
- İkamet edilen bölgedeki işsizlik oranı (unem)
- Fiili işgücü piyasası deneyimi (exper)

Çizelge 5.19. Gerçek Veri Uygulaması-2’de Değişkenler Arası Korelasyon Matrisi



Çizelge 5.19’a göre en yüksek ilişkiye sahip değişken çifti 0.89’luk bir ilişkiyle age ve husage değişkenleridir. Bu değişken çifti haricinde de -0.88, -0.69 gibi yüksek ilişkilere sahip değişken çiftleri göze çarpmaktadır.

Her ne kadar korelasyon matrisinde yüksek ilişkiler tespit edilmiş olsa da VIF ve koşul indeksi değerleri incelenerek çoklu bağlantı sorunuyla ilgili daha ayrıntılı bilgiler elde edilebilir. Çizelge 5.20’ye göre her iki veri kümesinde de VIF değerlerinin 2-10 arasında (6.92, 9.31) olması ve koşul indekslerinin de yine çok büyük olmaması (6.69, 7.63) sebebiyle açıklayıcı değişkenler arası zayıf da olsa bir çoklu bağlantı sorunu olduğu söylenebilir.

Çizelge 5.20. Gerçek Veri Uygulaması-2’de Eğitim ve Test Kümeleri İçin VIF ve Koşul İndeksi Değerleri

Değişkenler	Eğitim Kümesi (2/3)		Test Kümesi (1/3)	
	VIF	Koşul İndeksi	VIF	Koşul İndeksi
kidslt6	1.21	1.11	1.43	1.16
kidsge6	1.47	1.52	1.56	1.60
age	6.52	1.68	5.24	1.71
educ	2.01	1.87	2.23	1.82
wage	1.25	1.88	1.32	1.97
hushrss	1.95	2.04	2.04	2.01
husage	6.08	2.09	4.73	2.15
huseduc	1.93	2.17	2.09	2.34
huswage	3.61	2.29	4.31	2.59
faminc	5.10	2.57	5.93	2.82
mtr	6.92	2.89	9.31	3.28
motheeduc	1.56	3.26	1.85	3.32
fatheeduc	1.58	4.46	1.85	4.68
unem	1.10	6.03	1.14	5.88
exper	1.60	6.69	1.54	7.63

Zayıf çoklu bağlantının varlığı belirlendikten sonra KEKR, TBR, BBA-OBM ve BBA-K analizleri veriye uygulanmıştır. Gerçek Veri Uygulaması-1’e benzer şekilde yöntemler arasında belli bir standardı yakalamak ve çok fazla bilgi kaybını önlemek için tüm yöntemlerde KEKR ve TBR için en düşük açıklayıcılık gücüne sahip bağımsız bileşen, BBA-OBM’de açıklanan değişken ile en düşük ortak bilgi miktarına sahip bağımsız bileşen (bkz. Çizelge 5.21) ve BBA-K’da ise açıklanan değişken ile en düşük Pearson korelasyon katsayısına sahip bağımsız bileşen analiz dışı bırakılmıştır. Analizler sonucu elde edilen HKO’lar Çizelge 5.22’de görülmektedir. Buna göre yine eğitim kümesinde en düşük HKO, KEKR ile yakalanırken test kümesinde en düşük HKO ise BBA-OBM yöntemi ile yakalanmıştır.

Çizelge 5.21. Bağımsız Bileşenler ile Açıklanan Değişken Arasındaki Ortak Bilgi Miktarı ve Korelasyon Katsayıları

Bağımsız Bileşenler	Ortak Bilgi Miktarı	Korelasyon
S ₁	0.132199	0.111885
S ₂	0.289684	0.278757
S ₃	0.215554	0.038798
S ₄	0.201058	-0.250021
S ₅	0.095975	0.034613
S ₆	0.146824	0.041689
S ₇	0.135013	0.327551
S ₈	0.253467	-0.173553
S ₉	0.156594	0.206283
S ₁₀	0.136644	-0.020119
S ₁₁	0.126402	0.005250
S ₁₂	0.144675	-0.015231
S ₁₃	0.096817	0.028159
S ₁₄	0.079744	-0.099371
S ₁₅	0.080933	-0.086386

Çizelge 5.22. Zayıf Çoklu Bağlantı Durumu İçin Hata Kareler Ortalaması Sonuçları

	KEKR (14 B)	TBR (14 TB)	BBA-OBM (14 BB – S14)	BBA-K (13 BB – S11)
Eğitim-HKO	336005.5	336007.5	341157.9	336019.9
Test-HKO	535456.7	535818.9	533715.0	534782.5

5.2.3. Gerçek Veri Uygulaması-3

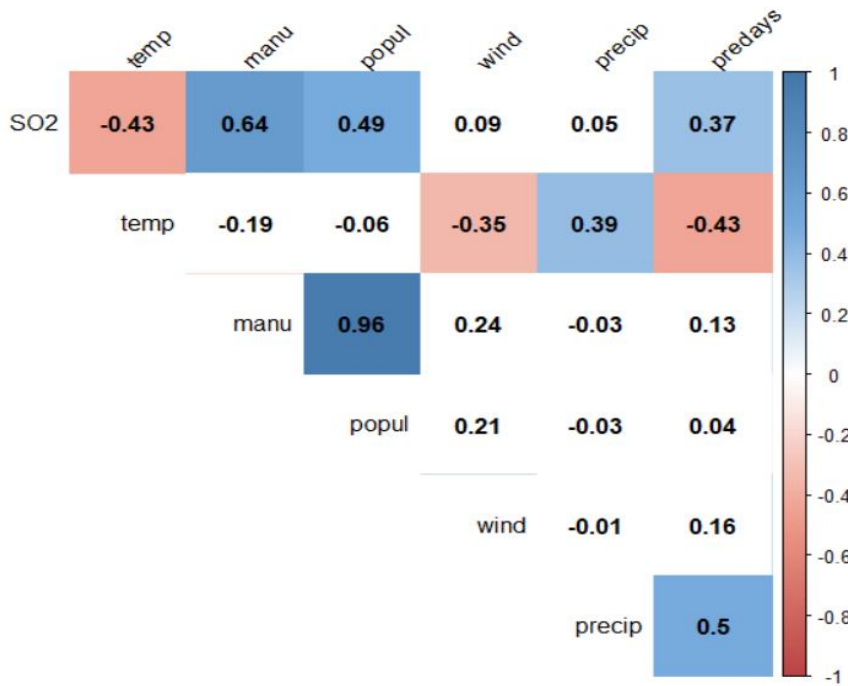
İlgili veri, R programı içerisinde HSAUR3 paketinde bulunan Amerika'ya ait bir hava kirliliği verisidir (R Core Team, 2019). Bu veri kümesi iklim ve nüfus gibi değerleri kullanarak bölgedeki hava kirliliğinin bir göstergesi olan kükürdioksit (SO₂) yoğunluğunu tahmin etmek amacıyla kullanılmaktadır.

Analizde kullanılan açıklayıcı değişkenler ve kısaltmaları aşağıdaki gibi sıralanmıştır:

- Yıllık ortalama sıcaklık (temp)
- 20’den fazla çalışanı olan üretim yerlerinin sayısı (manu)
- Nüfus (popul)
- Yıllık ortalama rüzgar hızı (wind)
- Yıllık ortalama yağış (precip)
- Bir yılda ortalama yağışlı gün sayısı (predays)

Çizelge 5.23’e göre en yüksek ilişkiye sahip değişken çifti, 0.96’lık bir ilişkiyle manu ve popul değişkenleridir. Açıklayıcı değişkenler arasındaki bir sonraki ikinci en yüksek ilişki ise 0.5 ile precip ve predays değişkenleri arasındadır.

Çizelge 5.23. Gerçek Veri Uygulaması-3’te Değişkenler Arası Korelasyon Matrisi



Çizelge 5.24’e göre her iki veri kümesinde de VIF değerlerinin 10’nun üzerinde ya da 10’a yakın olması (17.20, 9.10) ve koşul indekslerinin de yine 10’a yakın (8.77, 7.77)

değerler olması sebebiyle açıklayıcı değişkenler arası çoklu bağlantı düzeyi, orta olarak adlandırılmıştır.

Çizelge 5.24. Gerçek Veri Uygulaması-3'te Eğitim ve Test Kümeleri İçin VIF ve Koşul İndeksi Değerleri

Değişkenler	Eğitim Kümesi (2/3)		Test Kümesi (1/3)	
	VIF	Koşul İndeksi	VIF	Koşul İndeksi
temp	4.47	1.17	5.91	1.23
manu	17.20	1.45	9.04	1.66
popul	16.40	1.49	9.10	1.80
wind	1.29	1.75	2.07	3.02
precip	4.66	4.85	3.13	5.87
predays	2.59	8.77	6.76	7.77

Çizelge 5.25. Bağımsız Bileşenler ile Açıklanan Değişken Arasındaki Ortak Bilgi Miktarı ve Korelasyon Katsayıları

Bağımsız Bileşenler	Ortak Bilgi Miktarı	Korelasyon
s ₁	0.441761	-0.665170
s ₂	0.364126	-0.311517
s ₃	0.221755	-0.433524
s ₄	0.212238	-0.155755
s ₅	0.179988	-0.062210
s ₆	0.318931	-0.210381

Orta düzey çoklu bağlantıya sahip verilere KEKR, TBR, BBA-OBM ve BBA-K analizleri uygulanmıştır. Diğer uygulamalara benzer şekilde yöntemler arasında belli bir standardı yakalamak ve çok fazla bilgi kaybını önlemek için tüm yöntemlerde KEKR ve TBR için en düşük açıklayıcılık gücüne sahip bağımsız bileşen, BBA-OBM'de açıklanan değişken ile en düşük ortak bilgi miktarına sahip bağımsız bileşen (bkz. Çizelge 5.25) ve BBA-K'da ise açıklanan değişken ile en düşük Pearson korelasyon katsayısına sahip bağımsız bileşen, analiz dışı bırakılmıştır.

Çizelge 5.26. Orta Düzey Çoklu Bağlantı Durumu İçin Hata Kareler Ortalaması Sonuçları

	KEKR (5B)	TBR (5 TB)	BBA-OBM (5 BB – S5)	BBA-K (5 BB – S5)
Eğitim-HKO	156.7649	225.612	133.747	133.747
Test-HKO	466.6779	513.231	361.416	361.416

Analizler sonucu elde edilen HKO'lar Çizelge 5.26'da görülmektedir. Buna göre eğitim kümesinde ve test kümesinde en düşük HKO değerlerine BBA temelli yöntemler ile ulaşılmıştır. Hem ortak bilgi miktarı hem de korelasyon katsayısı temelli BBA'larda aynı bağımsız bileşen analiz dışı bırakıldığı için her iki yöntemde de aynı HKO değerleri elde edilmiştir.

5.2.4. Gerçek Veri Uygulaması-4

Çoklu bağlantı sorununun, regresyon katsayılarının standart hatalarını yükseltebileceği bilindiği için bu kısımdaki gerçek veri uygulaması, standart hataların incelenmesine ayrılmıştır.

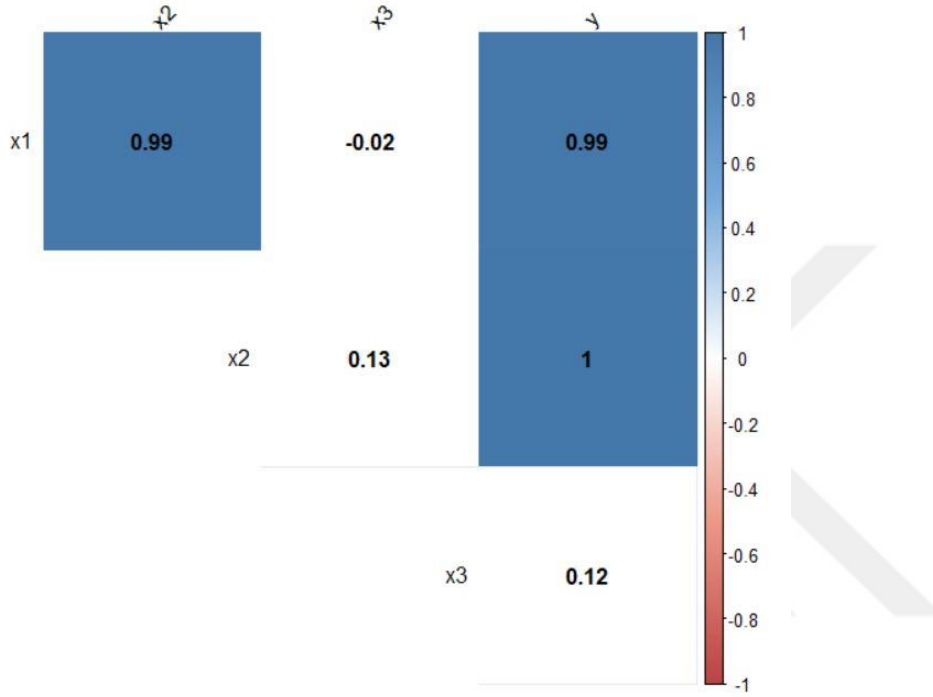
KEKR'de regresyon katsayılarının tahminlerinin dağılımı bilinmediğinden dolayı standart hatalar doğrudan bulunamamakta, Jackknife gibi tahmin edicilerle varyansın tahmin edilmesinden sonra bulunabilmektedirler. Bu nednele, tüm yöntemlerde elde edilen regresyon katsayılarının standart hataları birbirleriyle karşılaştırılırken aralarında bir paralellik olması açısından hepsinin tahmininde Jackknife yöntemi kullanılmıştır.

Orijinal Jackknife varyans tahmin edicisi; n gözlem sayısı, $\bar{\beta}_{-i}$; i . gözlem dışarıda bırakıldığında tahmin edilen regresyon katsayısı ve $\bar{\beta}$; $\bar{\beta}_{-i}$ 'lerin ortalaması olmak üzere Eşitlik 5.15'teki gibi hesaplanmaktadır (Wu, 1986; Martens ve Martens, 2000):

$$\frac{(n-1)}{n} \sum_{i=1}^n (\bar{\beta}_{-i} - \bar{\beta})^2 \quad (5.15)$$

Analizde kullanılan veri, NCSS programının kullanıcı rehberi içerisinde TBR anlatımında kullanılan bir veri kümesidir. Amaç y 'yi, X_1 , X_2 ve X_3 'ü kullanarak tahmin etmektir (Hintze, 2007).

Çizelge 5.27. Gerçek Veri Uygulaması-4'te Değişkenler Arası Korelasyon Matrisi



Çizelge 5.27'e göre X_1 ve X_2 değişkenleri arasında neredeyse %100'lük bir ilişki mevcuttur. Hem ilişkinin yüksekliği hem de Çizelge 5.28'deki VIF ve koşul indeksi değerlerinin çok büyük olması, açıklayıcı değişkenler arasında güçlü çoklu bağlantı olduğunu göstermektedir.

Çizelge 5.28. Gerçek Veri Uygulaması-4'te VIF ve Koşul İndeksi Değerleri

Değişkenler	VIF	Koşul İndeksi
x_1	477	3.72
x_2	486	6.69
x_3	11.70	134.46

Çizelge 5.29. Bağımsız Bileşenler ile Açıklanan Değişken Arasındaki Ortak Bilgi Miktarı ve Korelasyon Katsayıları

Bağımsız Bileşenler	Ortak Bilgi Miktarı	Korelasyon
s_1	0.554695	-0.984577
s_2	-0.137895	0.148408
s_3	-0.203149	-0.011003

Çoklu bağlantıya sahip verilere ÇDR, KEKR, TBR, BBA-OBM ve BBA-K analizleri uygulanmıştır. Yine diğer uygulamalara benzer şekilde yöntemler arasında belli bir standardı yakalamak ve çok fazla bilgi kaybını önlemek için tüm yöntemlerde KEKR ve TBR için en düşük açıklayıcılık gücüne sahip bağımsız bileşen, BBA-OBM’de açıklanan değişken ile en düşük ortak bilgi miktarına sahip bağımsız bileşen (bkz. Çizelge 5.29) ve BBA-K’da ise açıklanan değişken ile en düşük Pearson korelasyon katsayısına sahip bağımsız bileşen, analiz dışı bırakılmıştır.

Çizelge 5.30. Çoklu Bağlantı Durumu İçin Jackknife Standart Hatasının Tahmin Sonuçları

Değişkenler	ÇDR	KEKR (2 B)	TBR (2 TB)	BBA-OBM (2 BB - S2)	BBA-K (2 BB - S3)
x_1	1.0945	0.6767	0.6828	0.0588	7.8912
x_2	1.0909	0.6632	0.6560	0.1304	0.1007
x_3	0.8347	0.2935	0.2902	0.6880	7.3145

Analizler sonucu Jackknife yöntemi ile elde edilen standart hatalar Çizelge 5.30’da görülmektedir. Buna göre her ikinci değişken hariç diğer değişkenler için en düşük standart hataya BBA-OBM yöntemi ile ulaşılmıştır. TBR ve KEKR yöntemleri benzer sonuçlar vermiş, BBA-K yöntemi ise ÇDR yöntemi ile elde edilen standart hatalardan bile daha yüksek değerler vermiştir.

6. SONUÇLAR VE TARTIŞMA

Tez çalışmasının ana hedeflerinden ilki, çok değişkenli istatistiksel bir yöntem olmasına rağmen istatistikçiler tarafından pek fazla bilinmeyen BBA'nın bilinirliğinin artırılması ve konuya dair kısıtlı Türkçe literatüre katkıda bulunulmasıdır. Bu bağlamda tezde ilk olarak, BBA'nın tarihsel süreci ve uygulama alanları ele alınmıştır. Sonrasında, BBA'nın anlaşılabilmesi için gerekli matematiksel ve istatistiksel kavramlar anlatılmıştır. Temel BBA modeline, modelin özelliklerine, modelin çözümünde kullanılan algoritmalara yer verildikten sonra ise BBA'ya benzer bir yöntem olan ve özellikle istatistik alanında sık kullanılan TBA ile BBA'nın karşılaştırılması örnek bir uygulama çerçevesinde ele alınmıştır. Tezin bahsedilen bu kısımlarının; BBA ile ilgili telekomünikasyondan kimyaya, jeolojiden ekonomiye kadar çok çeşitli alanlarda çalışan ya da çalışacak araştırmacılar için ayrıntılı bir Türkçe kaynak eksikliğini gidereceği düşünülmektedir.

Tez çalışmasında ikinci olarak ise istatistik alanında sık karşılaşılan sorunlardan birine, BBA'nın araç olarak kullanıldığı alternatif bir çözüm yönteminin kazandırılması hedeflenmiştir. Bu amaçla BBA'nın, çoklu doğrusal regresyon analizinde sık karşılaşılan ve tahmin edilen regresyon modelleri üzerinde çeşitli olumsuzluklar yaratan çoklu bağlantı sorununun çözümünde kullanılabileceği üzerinde durulmuştur. BBA'nın, daha önce özellikle kimya alanında benzer kullanımlarına rastlanılmasından dolayı, çalışmaya özgünlük katabilmek amacıyla elde edilen bağımsız bileşenlerden hangilerinin modele dahil edilmesi gerektiği konusunda yeni bir yöntem önerilerek literatüre katkı sağlanmıştır.

Önerilen yöntem, entropi alanında sıkça kullanılan ve iki raslantı değişkeni arasındaki birliktelik derecesini veren ortak bilgi miktarının, korelasyon katsayısı yerine kullanımından oluşmaktadır. Ortak bilgi miktarının korelasyon katsayısı yerine tercih edilmesinde en önemli sebep; regresyon analizinde göz önünde bulundurulanın, açıklanan değişken ile açıklayıcı değişkenler arasındaki ilişki yerine nedensellik olmasıdır. Ortak bilgi miktarı, doğrusal ilişkinin yanı sıra karmaşık ilişkileri de göz önünde bulundurduğu için nedenselliğin ortaya çıkarılmasında korelasyon katsayısına göre daha başarılıdır.

Önerilen yöntemin performansı, çoklu bağlantı sorununu çözmek için kullanılan mevcut diğer yöntemler ile birlikte yapay veri kümelerinde ve gerçek veri kümelerinde karşılaştırmalı olarak değerlendirilmiştir.

Yapay veri kümesi uygulamalarının sonuçlarına göre hem eğitim hem de test kümesinde, önerilen yöntemin; diğer yöntemlere nazaran değişen standart sapma düzeylerinden çok fazla etkilenmediği, örneklem büyüklüğü arttıkça açıklanan değişkenin tahmininde daha düşük HKO'lara ulaştığı görülmüştür. Ayrıca standart sapmanın düşük olduğu durumlarda KEKR'nin iyi sonuçlar verdiği görülse de, standart sapmanın yüksek olduğu durumlarda BBA'ya dayalı yöntemlerin daha iyi sonuçlar verdiği belirlenmiştir. TBR yöntemi ise genel olarak diğer yöntemler içerisinde en kötü sonuçları vermiştir.

Yapay veri uygulamalarında korelasyon katsayısına ve ortak bilgi miktarına dayalı BBA yöntemleri çoğu zaman birbirleriyle aynı sonuçları verseler de farklı sonuç verdikleri durumlar da gözlenmiştir. Farklı sonuç elde edilen durumlar incelendiğinde, korelasyon katsayısı yönteminin eğitim, ortak bilgi miktarı yönteminin ise test kümelerinde daha iyi sonuç verdiği görülmüştür. Ayrıca örneklem büyüklüğü arttıkça her iki yöntemden elde edilen HKO değerlerinin de birbirine yaklaştığı da tespit edilmiştir.

HKO değerlerinin karşılaştırıldığı üç farklı alandaki gerçek veri uygulama sonuçlarına göre farklı düzeydeki çoklu bağlantı durumlarında, önerilen yöntemin diğer yöntemlere göre özellikle test kümesinde daha düşük HKO'lara sahip olduğu görülmüştür. Eğitim kümelerinin sonuçları incelendiğinde ise önerilen yöntemin; orta düzey çoklu bağlantıda diğer yöntemlere göre en düşük, zayıf çoklu bağlantıda kabul edilebilir, güçlü çoklu bağlantıda ise en yüksek HKO'lara sahip olduğu görülmüştür. Karşılaştırılan yöntemlerden KEKR yöntemi, diğer yöntemlere göre güçlü ve zayıf çoklu bağlantı durumlarında eğitim kümesinde daha başarılı sonuçlar sergilerken, TBR yöntemi ise genellikle en kötü sonuçları vermiştir.

Çoklu bağlantı sorununun en önemli etkilerinden biri, regresyon katsayılarının standart hatalarının olması gerekenden daha büyük hale gelmesi olduğu için dördüncü gerçek veri

uygulamasında, regresyon katsayılarının Jackknife yöntemi ile tahmin edilen standart hataları karşılaştırılmıştır. Önerilen yöntem, diğer yöntemler içerisinde en düşük standart hataları vererek bu açıdan da kullanılabilirliğini göstermiştir.

Bu bulgular ışığında tez çalışması, kimyadan inşaat mühendisliğine kadar pek çok alanda uygulama alanına sahip olan regresyon analizinde sık karşılaşılan sorunlardan biri olan çoklu bağlantı sorununun BBA ile çözülebileceğini ve özellikle korelasyon katsayısı yerine ortak bilgi miktarının kullanılması ile regresyon analizi için daha doğru bağımsız bileşenlerin seçilebileceğini ortaya koymuştur. Ayrıca, önerilen yöntem ile regresyon modellerinde daha düşük standart sapmaya sahip parametre tahminlerinin elde edilebileceği ve özellikle geleceğe yönelik kullanılacak regresyon modelleri için daha düşük HKO'lara sahip öngörüler gerçekleştirilebileceği gösterilmiştir.

Genel olarak değerlendirildiğinde, BBA'nın istatistik alanındaki sorunlardan birine başarılı bir şekilde çözüm bulmasının konuyla ilgili istatistik alanındaki yeni çalışmalara öncülük edeceği ve konunun çeşitli disiplinlerde farklı amaçlar için kullanımının yaygınlaşacağı düşünülmektedir.

7. KAYNAKLAR

- Albert, S.A. and Bowman, D.C., Tracking Scattered Signals in the Acoustic Coda Using Independent Component Analysis in a Topographically Complex Setting, *Geophysical Journal International*, 216 (2018) 67-776.
- Alpar, R., Uygulamalı Çok Değişkenli İstatistiksel Yöntemler, Detay Yayıncılık, 2013.
- Alves, F.C.G.B.S., Coqueiro, A., Março, P.H. and Valderrama, P., Evaluation of Olive Oils from the Mediterranean Region by UV-Vis Spectroscopy and Independent Component Analysis, *Food Chemistry*, 273 (2019) 124-129.
- Artoni, F., Delorme, A. and Makeig, S., A Visual Working Memory Dataset Collection with Bootstrap Independent Component Analysis for Comparison of Electroencephalographic Preprocessing Pipelines, *Data in Brief*, 22 (2019) 787-793.
- Aveta, F., Refai, H., Lo Presti, P., Tedder, S.A. and Schoenholz, B.L., Independent Component Analysis for Processing Optical Signals in Support of Multi-User Communication, *Free-Space Laser Communication and Atmospheric Propagation XXX*, San Francisco, USA, 2018.
- Baker, B.T., Abrol, A., Silva, R.F., Damaraju, E., Sarwate A.D., Calhoun, V.D. and Plis, S.M., Decentralized Temporal Independent Component Analysis: Leveraging Fmri Data in Collaborative Settings, *NeuroImage*, 186 (2019) 557-569.
- Bulut, Y.M., Çoklu İç İlişki Durumunda Kısmi En Küçük Kareler Regresyonu ve Alternatif Yöntemlerle Karşılaştırılması, Yüksek Lisans Tezi, Eskişehir Osmangazi Üniversitesi Fen Bilimleri Enstitüsü, Eskişehir, 2011.
- Bursa, N., ve G. Özel Kadılar, Türkiye Kredi Temerrüt Takası Primlerinin Entropi Kavramı ile İncelenmesi, *Euroasian Econometrics, Statistics and Empirical Economics Journal*, 3 (2016) 23-32.
- Chakrabarty, R., and Sarkar, A., The Effect of Economic Indicators on the Volatility of Indian Stock Market: Using Independent Component Regression, *Journal of Contemporary Research in Management*, 8 (2013) 1-22.

- Chen, Y., Niu, L., Chen, R.B. and Q, He., Sparse-Group Independent Component Analysis with Application to Yield Curves Prediction, *Computational Statistics & Data Analysis*, 133 (2019) 76-89.
- Chowdhury, U.N., Chakravarty, S.K. and Hossain, M.T., Short-Term Financial Time Series Forecasting Integrating Principal Component Analysis and Independent Component Analysis with Support Vector Regression, *Journal of Computer and Communications*, 36 (2018) 51-67.
- Cohen-Waeber, J., Bürgmann, R., Chaussard, E., Giannico, C., and Ferretti, A., Spatiotemporal Patterns of Precipitation-Modulated Landslide Deformation from Independent Component Analysis of InSAR Time Series, *Geophysical Research Letters*, 45 (2018) 1878-1887.
- Comon, P., Independent Component Analysis, a New Concept?, *Signal Processing*, 36 (1994) 287-314.
- Dai, W., Liu, B., Ding, X. and Huang, D., Modeling Dam Deformation Using Independent Component Regression Method, *Transactions of Nonferrous Metals Society of China*, 23 (2013) 2194-2200.
- de Lauro, E., Petrosino, S. and Falanga, M., Independent Component Analysis as a Monitoring Tool in Geophysical Environment: The Case of Campi Flegrei (Italy). *IEEE Workshop on Environmental, Energy, and Structural Monitoring Systems (EESMS)*, Salerno, 2018 1-6.
- de Siqueira Santos, S., Takahashi, D.Y., Nakata, A. and Fujita, A., A Comparative Study of Statistical Methods Used to Identify Dependencies Between Gene Expression Signals, *Briefings in Bioinformatics*, 15 (2013) 906-918.
- Delaporte, G., Cladière, M., Bouveresse, D.J.R. and Camel, V., Untargeted Food Contaminant Detection Using UHPLC-HRMS Combined with Multivariate Analysis: Feasibility Study on Tea, *Food Chemistry*, 277 (2019) 54-62.
- Dharmapranjani, D., Nguyen, H.K., Lewis, T.W., DeLosAngeles, D., Willoughby, J.O. and Pope, K.J., A Comparison of Independent Component Analysis Algorithms and Measures to Discriminate Between EEG and Artifact Components, *38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Orlando, USA, 2016 825-828.

- Eaton, M.L. *Multivariate Statistics A Vector Space Approach*. Ohio, USA: Institute of Mathematical Statistics, **2007**.
- Garcia-Bracamonte, J.E., Rangel-Magdaleno, J., Ramirez-Cortes, J.M., Gomez-Gill, P. and Paregrina-Barreto, H., Induction Motors Fault Detection Using Independent Component Analysis on Phase Current Signals, *IEEE International Instrumentation and Measurement Technology Conference*, Houston, USA, **2018** 1-6.
- Ghosh, M., Adhikary, S., Kanti Ghosh, K., Sardar, A., Begum, S. and Sarkar, R., Genetic Algorithm Based Cancerous Gene Identification from Microarray Data Using Ensemble of Filter Methods, *Medical & Biological Engineering & Computing*, **57** (2019) 159-176.
- Gouriéroux, C., Monfort, A. and Renne, J.P., Statistical Inference for Independent Component Analysis: Application to Structural VAR Models, *Journal of Econometrics*, **196** (2017) 111-126.
- Gultepe, E. and Makrehchi, M., Improving Clustering Performance Using Independent Component Analysis and Unsupervised Feature Learning, *Human-centric Computing and Information Sciences*, **2018** 8-25.
- Gustafsson, M.G., Independent Component Analysis Chemically Interpretable Latent Variables in Multivariate Regression, *Journal Chemical Information and Modelling*, **45** (2005) 1244-1255.
- Haque, M.M., deSouza, A. and Rahman, A., Water Demand Modelling Using Independent Component Regression Technique, *Water Resources Management*, **31** (2017) 299-312.
- Haykin, S., *Neural Networks and Learning Machines*, Prentice Hall, **2008**.
- Helwig, N.E., *Ica: Independent Component Analysis*, **2015**.
- Hérault, J., Jutten, C. and Ans, C., Détection de Grandeurs Priitives dans un Message Composite par une Architecture de Calcul Neuromimétique en Apprentissage non Suprèvisè, Nice: Actes du Xème Colloque, Gretsì, **1985** 1017-1022.
- Hintze, J.L., *NCSS User's Guide*, NCSS, **2007**.

- Hoeffding, W., A Non-parametric Test of Independence, *The Annals of Mathematical Statistics*, **1948** 546-557.
- Hollander, M., Wolfe, D., and Chicken, E., *Nonparametric Statistical Methods*, Wiley, **2014**.
- Hyvärinen, A., Fast and Robust Fixed-Point Algorithms for Independent Component Analysis, *IEEE Transactions on Neural Networks*, 10 (**1999**) 626-634.
- Hyvärinen, A., Independent Component Analysis: Recent Advances. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 371 (**2013**) 1-19.
- Hyvärinen, A., Karhunen, J. and Oja, E., *Independent Component Analysis*, USA: John Wiley & Sons, **2001**.
- Hyvärinen, A. and Oja, E., Independent Component Analysis: Algorithms and Applications. *Neural Networks*, 13 (**2000**) 411-430.
- Jiang, C. Zhang, H., Ren, Y., Han, Z., Chen, K. and Hanzo, L., Machine Learning Paradigms for Next-Generation Wireless Networks, *IEEE Wireless Communications*, 24 (**2017**) 98-105.
- Jutten, C., and Héroult, J., Blind Separation of Sources, Part I. An Adaptive Algorithm Based on Neuromimetic Architecture, *Signal Processing*, 24 (**1991a**) 1-10.
- Jutten, C. and Héroult, J., Blind Separation of Sources, Part II: Problems Statement, *Signal Processing*, 24 (**1991b**) 11-20.
- Jutten, C., and Héroult, J., Blind Separation of Sources, Part III: Stability Analysis. *Signal Processing*, 24 (**1991c**) 21-29.
- Jolliffe, I., *Principal Component Analysis*, Springer, **2011**.
- Kamal, M.S., Trivedi, M.C., Alam, J.B., Dey, N., Ashour, A.S., Shi, F. and Tavares, J.M.R., Big DNA Datasets Analysis Under Push Down Automata, *Journal of Intelligent & Fuzzy Systems*, **2018** 1555-1565.
- Kaneko, H., Arakawa, M. and Funatsu, K., Development of a New Regression Analysis Method Using Independent Component Analysis, *Journal Chemical Information and Modelling*, 48 (**2008**) 534-544.

- Koush, Y., Masala, N., Scharnowski, F. and De Ville, D.V., Data-Driven Tensor Independent Component Analysis for Model-Based Connectivity Neurofeedback, *NeuroImage*, 184 (2019) 214-226.
- Lahaw, Z.B., Essaidani, D. and Seddik, H., Robust Face Recognition Approaches Using PCA, ICA, LDA Based on DWT, and SVM Algorithms, 41st International Conference on Telecommunications and Signal Processing, Atina, 2018 1-5.
- Langlois, D., Sylvain C. and Dominique G., An Introduction to Independent Component Analysis: InfoMax and FastICA algorithms, *Tutorials in Quantitative Methods for Psychology*, 6 (2010) 31-38.
- Lee, J.M., Quin, S.J. and Lee, I.B., Fault Detection and Diagnosis Based on Modified Independent Component Analysis, *AIChE Journal*, 52 (2006) 3501-3514.
- Li, Z. and Yan, X., Adaptive Selective Ensemble-Independent Component Analysis Models for Process Monitoring, *Industrial & Engineering Chemistry Research*, 57 (2018) 8240-8252.
- Linthurst, R.A., Aeration, Nitrogen, pH and Salinity as Factors Affecting *Spartina Alterniflora* Growth and Dieback. *Doktora Tezi*, North Carolina State University, USA, 1979.
- Liu, Y., Xu, H., Xia, Z. and Gong, Z., Multi-Spectrometer Calibration Transfer Based On Independent Component Analysis, *Analyst*, 143 (2018) 1274-1280.
- Marchini, J.L., Heaton, C. and Ripley, B.D., fastICA: FastICA Algorithms to Perform ICA and Projection Pursuit, 2017.
- Martens, H. and Martens, M., Modified Jackknife Estimation of Parameter Uncertainty in Bilinear Modelling by Partial Least Squares Regression (PLSR), *Food Quality and Preference*, 11 (2000) 5-16.
- Matthias, K., Haueisen, J. and Ivanova, G., Independent Component Analysis: Comparison of Algorithms for the Investigation of Surface Electrical Brain Activity, *Medical & Biological Engineering & Computing*, 47 (2009) 413-423.
- Miettinen, J., Nordhausen, K., Oja, H. and Taskinen, S., fICA: Classical, Reloaded and Adaptive Fastica Algorithms, 2017.

- Miettinen, J., Nordhausen, K. and Taskinen, S., Blind Source Separation Based on Joint Diagonalization in R: The Packages JADE and BSSasyp, *Journal of Statistical Software*, 76 (2017) 1-31.
- Montgomery, D.C., Peck, E.A. and Geoffrey, G., *Introduction to Linear Regression Analysis*, John Wiley & Sons, 2013.
- Mroz, T.A., The Sensitivity of an Empirical Model of Married Women's Hours of Work to Economic and Statistical Assumptions, *Econometrica*, 55 (1987) 765-799.
- Mutihac, R. and Van Hulle, M.M., Comparison of Principal Component Analysis and Independent Component Analysis for Blind Source Separation, *Romanian Reports in Physics*, 56 (2004) 20-32.
- Naik, G.R., A Comparison of ICA Algorithms in Surface EMG Signal Processing, *International Journal of Biomedical Engineering and Technology*, 6 (2011) 363-374.
- Naik, G.R. and Kumar, D.K., An Overview of Independent Component Analysis and Its Applications, *Informatika*, 35 (2011) 63-81.
- Nordhausen, K. and Oja, H., *Independent Component Analysis: A Statistical Perspective*, Wiley Interdisciplinary Reviews: Computational Statistics, 10 (2018) 1-23.
- Özdamar, E.Ö., EEG Analizinde Bağımsız Bileşenler, Doktora Tezi, Mimar Sinan Üniversitesi Fen Bilimleri Enstitüsü, İstanbul, 2009.
- Pearl, J., *Causality*, Cambridge University Press, 2009.
- Polat, E., Kısmi En Küçük Kareler Regresyonu, Yüksek Lisans Tezi, Hacettepe Üniversitesi Fen Bilimleri Enstitüsü, Ankara, 2009.
- R Core Team, *R: A Language and Environment for Statistical Computing*, 2019.
- Rawlings, J.O., Pantula, S.G. and Dickey, D.A., *Applied Regression Analysis: A Research Tool*, Springer Science & Business Media, 2001.
- Sahonero-Alvarez, G. and Calderon, H., A Comparison of SOBI, FastICA, JADE and Infomax Algorithms, *Proceedings of the 8th International Multi-Conference on Complexity, Informatics and Cybernetics*, Orlando, USA, 2017 17-22.
- Shannon, C.E., A Mathematical Theory of Communication., *Bell System Technical Journal*, 27 (1948) 379-423.

- Shao, X., Wang, W., Hou, Z. and Cai, W., A New Regression Method Based on Independent Component Analysis, *Talanta*, 69 (2006) 676-680.
- Shlens, J., A Tutorial on Independent Component Analysis, 2014, <https://arxiv.org/pdf/1404.2986.pdf> (erişildi: 6 Mart 2018).
- Stone, J.V., Independent Component Analysis: A Tutorial Introduction, MIT Press, Londra, 2004.
- Şamilov, A., Entropi, İnfomasyon ve Entropi Optimizasyon, Nobel Yayınevi, Ankara, 2015.
- Tatlıdil, H., Uygulamalı Çok Değişkenli İstatistiksel Analiz, Hacettepe Üniversitesi, Ankara, 2002.
- Tharwat, A., Independent Component Analysis: An Introduction, Applied Computing and Informatics, 2018 Baskıda.
- Thomas, M.C., Zhu, W. and Romagnoli, J.A., Data Mining and Clustering in Chemical Process Databases for Monitoring and Knowledge Discovery, *Journal of Process Control*, 67 (2018) 160-175.
- Tierney, J.E., Wilkes, D.M. and Byram, B.C., Independent Component Analysis-Based Tissue Clutter Filtering for Plane Wave Perfusion Ultrasound Imaging, *Medical Imaging: Ultrasonic Imaging and Tomography*, San Diego, USA, 2019.
- Wang, C., Xu, Y., Tang, M. and Wang, L., Blind Source Separation Based on Variational Bayesian Independent Component Analysis, *IEEE 3rd Advanced Information Technology, Electronic and Automation Control Conference*, Chongqing, China, 2018 1614-1618.
- Wang, Y. and Guo, Y., A Hierarchical Independent Component Analysis Model for Longitudinal Neuroimaging Studies, *NeuroImage*, 189 (2019) 380-400.
- Witten, I.H., Frank, E., Hall, M.A. and Pal, C.J., *Data Mining: Practical Machine Learning Tools and Techniques*, Elsevier, 2016.
- Wu, C.F.J., *Jackknife, Bootstrap and Other Resampling Methods in Regression Analysis*, *The Annals of Statistics*, 14 (1986) 1261-1295.
- Yu, J., Yoo, J., Jang, J., Park, J.H. and Kim, S., A Novel Hybrid of Auto-Associative Kernel Regression and Dynamic Independent Component Analysis for Fault

Detection in Nonlinear Multimode Processes, *Journal of Process Control*, 68 (2018) 129-144.

Zhan, X., Dai, W. and Zhou, S., Application of Independent Component Regression in Dam Deformation Analysis, *Journal of Information&Computational Sciences*, 11 (2014) 1939-1945.

Zhou, W. and Altman, R.B, Data-Driven Human Transcriptomic Modules, *BMC Bioinformatics*, 19 (2018) 1-25.



EKLER

EK 1 - Tezden Üretilmiş Bildiriler

Bursa, N. ve Tatlıdil, H., Independent Component Analysis: An Alternative Approach for Multicollinearity Problem, 4th International Researchers, Statisticians and Young Statisticians Congress, İzmir, 2018.





HACETTEPE ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
YÜKSEK LİSANS / DOKTORA TEZ ÇALIŞMASI ORJİNALLİK RAPORU

HACETTEPE ÜNİVERSİTESİ
FEN BİLİMLER ENSTİTÜSÜ
İSTATİSTİK ANABİLİM DALI BAŞKANLIĞI'NA

Tarih: 18/07/2019

Tez Başlığı / Konusu: Bağımsız Bileşenler Analizi ile Çoklu Bağlantı Sorununa Bir Yaklaşım

Yukarıda başlığı/konusu gösterilen tez çalışmamın a) Kapak sayfası, b) Giriş, c) Ana bölümler d) Sonuç kısımlarından oluşan toplam 123 sayfalık kısmına ilişkin, 18/07/2019 tarihinde şahsım/tez danışmanım tarafından Turnitin adlı intihal tespit programından aşağıda belirtilen filtrelemeler uygulanarak alınmış olan orijinallik raporuna göre, tezimin benzerlik oranı % 5 'tir.

Uygulanan filtrelemeler:

- 1- Kaynakça hariç
- 2- Alıntılar hariç/dâhil
- 3- 5 kelimedenden daha az örtüşme içeren metin kısımları hariç

Hacettepe Üniversitesi Fen Bilimleri Enstitüsü Tez Çalışması Orijinallik Raporu Alınması ve Kullanılması Uygulama Esasları'nı inceledim ve bu Uygulama Esasları'nda belirtilen azami benzerlik oranlarına göre tez çalışmamın herhangi bir intihal içermediğini; aksinin tespit edileceği muhtemel durumda doğabilecek her türlü hukuki sorumluluğu kabul ettiğimi ve yukarıda vermiş olduğum bilgilerin doğru olduğunu beyan ederim.

Gereğini saygılarımla arz ederim.

18/07/2019

Adı Soyadı: Nurbanu Bursa
Öğrenci No: N13146919
Anabilim Dalı: İstatistik
Programı: İstatistik
Statüsü: Y.Lisans Doktora Bütünleşik Dr.

DANIŞMAN ONAYI

UYGUNDUR.

Prof. Dr. Hüseyin Tatlıdil



ÖZGEÇMİŞ

Kimlik Bilgileri

Adı Soyadı : Nurbanu Bursa

Doğum Yeri : Nazilli

Doğum Tarihi : 27 Şubat 1988

Eğitim

Lise : Nazilli Anadolu Lisesi

Lisans : Başkent Üniversitesi İstatistik ve Bilgisayar Bilimleri Bölümü

Yan Dal : Başkent Üniversitesi Uluslararası Ticaret Bölümü

Yüksek Lisans : Hacettepe Üniversitesi İstatistik Bölümü

Yabancı Dil

İngilizce

İş Deneyimi

2011 -... : Hacettepe Üniversitesi Fen Fakültesi İstatistik Bölümü
Araştırma Görevlisi