

T.C.
İNÖNÜ ÜNİVERSİTESİ
SOSYAL BİLİMLERİ ENSTİTÜSÜ



DİNAMİK DOĞRUSAL MODELLER İLE

TÜRKİYE'YE AİT GAYRİ SAFİ YURTIÇİ
HÂSILA TAHMİNİ

YÜKSEK LİSANS TEZİ

DANIŞMAN

Doç.Dr. Bahadır YÜZBAŞI

HAZIRLAYAN

Çetin GÖRÜR

MALATYA-2019

T.C.
İNÖNÜ ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ

DİNAMİK DOĞRUSAL MODELLER İLE
TÜRKİYE'YE AİT GAYRİ SAFİ YURTİÇİ
HÂSILA TAHMİNİ

YÜKSEK LİSANS TEZİ

Hazırlayan:

Çetin GÖRÜR

Tez Danışmanı:

Doç. Dr. Bahadır YÜZBAŞI

MALATYA-2019

T.C
İNÖNÜ ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ
DİNAMİK DOĞRUSAL MODELLER İLE TÜRKİYE'YE AİT GAYRİ SAFİ
YURTIÇİ HÂSILA TAHMİNİ
YÜKSEK LİSANS TEZİ

DANIŞMAN

Doç. Dr. Bahadır YÜZBAŞI

HAZIRLAYAN

Çetin GÖRÜR

Jürimiz. 14.01.2019 tarihinde yapılan savunma sonucunda bu yüksek lisans tezini (oybirliği/~~oy çokluğu~~) ile başarılı bularak Ekonometri Anabilim dalında yüksek lisans tezi olarak kabul etmiştir.

Jüri Üyelerinin Unvan Ad Soyadı

1. Doç. Dr. Bahadır YÜZBAŞI (Danışman)
2. Doç. Dr. Tayfur BAYAT
3. Doç. Dr. Gökhan GÖKDERE

İmzası



İNÖNÜ ÜNİVERSİTESİ SOSYAL BİLİMLER ENSTİTÜSÜ YÖNETİM
Kurulunun.....tarih ve sayılı kararıyla onaylanmıştır.

Prof. Dr. Mehmet KUBAT

Sosyal Bilimler Enstitüsü Müdürü

ONUR SÖZÜ

Doç. Dr. Bahadır YÜZBAŞI' ın danışmanlığında yüksek lisans tezi olarak hazırladığım “**DİNAMİK DOĞRUSAL MODELLER İLE TÜRKİYE’YE AİT GAYRİ SAFİ YURTIÇİ HÂSILA TAHMİNİ**” başlıklı bu çalışmanın, bilimsel ahlak ve geleneklere aykırı düşecek bir yardıma başvurmaksızın tarafımdan yazıldığını ve yararlandığım bütün yapıtların hem metin içinde hem de kaynakçada yöntemine uygun biçimde gösterilenlerden oluştuğunu belirtir, bunu onurumla doğrularım.

Tarih:

Ad-Soyad: Çetin GÖRÜR

İmza:

BİLDİRİM

Doç. Dr. Bahadır YÜZBAŞ'ın danışmanlığında hazırladığım **“DİNAMİK DOĞRUSAL MODELLER İLE TÜRKİYE'YE AİT GAYRİ SAFİ YURTİÇİ HÂSILA TAHMİNİ”** başlıklı doktora tezimin tamamen kendi çalışmam olduğunu ve her alıntıya kaynak gösterdiğimi taahhüt eder, doktora tezimin kâğıt ve elektronik kopyalarının İnönü Üniversitesi Sosyal Bilimler Enstitüsü arşivlerinde aşağıda belirttiğim koşullarda saklanmasına izin verdiğimi onaylarım.

- Tezimin tamamı her yerden erişime açılabilir.
- Tezim sadece İnönü Üniversitesi yerleşkelerinden erişime açılabilir.
- Tezimin yıl süreyle erişime açılmasını istemiyorum. Bu sürenin sonunda uzatma için başvuruda bulunmadığım takdirde, tezimin tamamı her yerden erişime açılabilir.

Tarih:

Çetin GÖRÜR

ÖNSÖZ

Tez çalışmam süresince bilgi ve öneri anlamında hiçbir zaman desteğini esirgemeyen ve bu tezin oluşmasında büyük rol oynayan değerli tez danışmanım Doç. Dr. Bahadır YÜZBAŞI'na,

Tez hazırlama aşamasında bilgi, tecrübe ve desteklerini benden esirgemeyen değerli hocalarım Dr. Öğr. Üyesi Şakir İŞLEYEN, Dr. Öğr. Üyesi Yener ALTUN ve Dr. Öğr. Üyesi Yıldırım DEMİR'e

Çalışmalarım sırasında bana yardımcı olan değerli arkadaşlarım Emsal Çağla AVCU ve Tolga YILMAZ'a

Bana maddi ve manevi her türlü desteği veren aileme,
Sonsuz teşekkür ederim.

ÖZET

Ekonomik tahmin her zaman işletmeler ve hükümetler için finansal politika ve stratejiyi formüle ederken bir karar verme aracı olarak önemli bir rol oynamaktadır. Bununla birlikte, gerçek verilerle uğraşırken (sadece iktisatla ilgili olanı değil, aynı zamanda sosyoloji, biyoloji, mühendislik vb.), nadiren ilgilenilen konu hakkında yeterli bilgiye sahip olduğumuz olur. Analiz etmek istediğimiz sistemi tanımlayan kesin ve deterministik bir modelin mevcut olduğu durumlarda bile, bazı açıklayıcı değişkenleri atlamak veya ölçüm hataları yapmak her zaman mümkündür. Böylece, araştırılan konuyu analiz ederken bir belirsizlik durumu her zaman mevcut olacaktır. Bu nedenle yapılan tahminleri kesin bir gösterge olarak göremeyiz.

Son yıllarda, zaman serisi analizi amaçları için durum uzay modellerinin(DUM) uygulanmasına yönelik artan bir ilgi vardır. Bu kategoriye ait modeller, bir dinamik sistem ve rastgele bir bozukluk bileşeni olarak belirli bir zaman serisini açıklar. Güçlü olasılıksal yapıları, geniş bir uygulama yelpazesi için kullanılabilecek esnek bir çerçeve sunar. Bu esneklik, DUM'ların bir alt sınıfını, yani dinamik doğrusal modelleri(DLM) seçmemize neden olan olmaktadır.

Küreselleşen dünyada ülkelerin ekonomik gücünü gösteren önemli değişkenlerden biri de gayri safi yurtiçi hasılasıdır(GSYİH). Güçlü bir ekonomiye sahip olan ülkelere ait GSYİH değerleri yüksektir. Gelişmemiş ülkelere genel olarak GSYİH verileri düşüktür.

Bu tezde, DUM'ların bir alt sınıfı olan DLM'ler açıklanıp zaman serilerinde kullanımı tanıtılmaktadır. Dinamik doğrusal modeller ile Türkiye'ye ait GSYİH verileri kullanılarak gelecek yıllar için tahminlerde bulunmaktadır. Bu amaç doğrultusunda, uygulama ve tahminler için son dönemlerde çok amaçlı ve kullanışlı olan R programı kullanılmaktadır.

Anahtar Kelimeler: Dinamik Doğrusal Modeller(DLM), Durum Uzay Modelleri(DUM), Zaman Serisi, Durağanlık, Gayri Safi Yurtiçi Hasıla(GDP)

ABSTRACT

The economic forecast always plays an important role as a decision-making tool when formulating financial policy and strategy for businesses and governments. However, when dealing with real data (not just about economics, but also sociology, biology, engineering, etc.), we rarely know enough about the topic of interest. Even if there is a precise and deterministic model that defines the system we want to analyze, it is always possible to skip some explanatory variables or make measurement errors. Thus, an amount of uncertainty will always be available when analyzing the subject being investigated. Therefore, we cannot see the estimates made as a definitive science.

In recent years, there has been an increasing interest in the application of state space models (DUM) for time series analysis purposes. The models of this category describe a specific time series as a dynamic system and a random component of disorder. Its powerful probability structures offer a flexible framework that can be used for a wide range of applications. This flexibility causes us to choose a subclass of DUMs, ie, dynamic linear models (DLM).

One of the important variables that show the economics strength of the countries in the globalized world is the gross domestic product (GDP). GDP values of countries with a strong economy are high. In underdeveloped countries, GDP figures are generally low.

In this thesis, a subclass of DUMs, DLMs are explained and their use in time series is introduced. With dynamic linear models using GDP data of Turkey are making predictions for the coming years. For this purpose, the R program, which is very useful and useful in recent years, is used for applications and predictions.

Key Words: Dynamic Linear Models (DLM), State Space Models (DUM), Time Series, Stationarity, Gross Domestic Product

İÇİNDEKİLER DİZİNİ

Kabul Onay	ii
Onur Sözü	iii
Bildirim	iv
Ön Söz	v
Özet	vi
Abstract	vii
İçindekiler Dizini	viii
Çizelge ve Tablo Dizini	x
Grafik ve Şekil Dizini	xi
Kısaltmalar Dizini	xii
GİRİŞ	1

BİRİNCİ BÖLÜM

DİNAMİK DOĞRUSAL MODELLERİ VE DURUM UZAY MODELLERİNE GENEL BAKIŞ

1.Dinamik Doğrusal Modelleri ve Durum Uzay Modellerine Genel Bakış.....	3
---	---

İKİNCİ BÖLÜM

2. DOĞRUSAL REGRASYON MODELİNDE BAYES ÇIKARIMI VE DURUM UZAY MODELLERİ

2.1. Doğrusal Regresyon Modelinde Bayes Çıkarımı	9
2.2. Durum Uzay Modelleri	10

ÜÇÜNCÜ BÖLÜM

DİNAMİK DOĞRUSAL MODELLER

3.Dinamik Doğrusal Modeller.....	13
3.1. R Programı ve dlm	15
3.1.1. dlm Paketinde Dinamik Doğrusal Modeller	16

3.2. Durum Hesaplama ve Tahmini.....	18
3.3.Filtreleme.....	20
3.3.1. Dinamik Doğrusal Modeller için Kalman Filtresi	22
3.3.2. Eksik Gözlemlerle Filtreleme	27
3.3.3. Düzleştirici(Smoother).....	28
3.3.4. Dinamik Doğrusal Modeller için Kalman Düzleştiricisi (Kalman Smoother).....	30
3.3.5. DLM Modelleri İçin Tahmin	32

DÖRDÜNCÜ BÖLÜM

4. DİNAMİK DOĞRUSAL MODELLER İLE TÜRKİYE'YE AİT GAYRİ SAFİ YURTİÇİ HÂSILA TAHMİNİ

4.1. Ekonomik Büyüme.....	37
4.2. Türkiye Ait GSYİH'nin DLM ile Tahmini	37
4.2.1. GSYİH için n. Düzeyde Bir DLM Kurulumu.....	38
4.2.2. GSYİH İçin Kalman Filtreleme	39
4.2.3. GSYİH için Kalman Düzleştirici	40
4.2.4. Tahmin Modelinin Karşılaştırılması	42
4.2.5. GSYİH İçin Tahmin.....	43

BEŞİNCİ BÖLÜM

DEĞERLENDİRME VE SONUÇ

5. Değerlendirme ve Sonuç	46
KAYNAKÇA	48
EKLER	51
ÖZGEÇMİŞ	60

ÇİZELGE VE TABLO DİZİNİ

Çizelge 3.1.1.1.: R Programında Dinamik Doğrusal Modellere Ait Bazı Fonksiyonlar	17
Tablo 4.2.1.1.: DLM Çıktı Sonuçları	38
Tablo 4.2.4.1.: Model 1 ve Model 2 Karşılaştırma Sonuçları	42
Tablo 4.2.4.2.: TheilU İstatistik Sonuçları	43



GRAFİK VE ŞEKİL DİZİNİ

Grafik 1.1.:	Aile Gıda Harcamaları, Çeyrek Veriler (1996Q1 - 2005Q4).....	4
Grafik 1.2.:	2000Q1 - 2010Q4 Çeyreklik İngiltere'deki Gaz Tüketimi.....	4
Grafik 1.3.:	Ahswan'daki Nil Nehri'nin Yıllık Akış Ölçümleri, 1871-1970.....	5
Grafik 1.4.:	İngiltere’de Aylık Tavuk Fiyatları	6
Grafik 4.2.1.:	1968-2017 Türkiye GSYİH Verileri	38
Grafik 4.2.2.1.:	1968-2017 Türkiye GSYM Verileri İçin Kalman Filtreleme	39
Grafik 4.2.3.1.:	1968-2017 Türkiye GSYİH Verileri İçin Kalman Düzleştirici	40
Grafik 4.2.3.2.:	1968-2017 Türkiye GSYİH Verileri İçin Kalman Filtreleme ve Düzleştirici	41
Grafik 4.2.3.3.:	1968-2017 Türkiye GSYİH Verileri İçin Dönemsel Değişim	41
Grafik 4.2.5.1.:	1968-2017 Türkiye GSYİH Gerçek Verileri ile Tahmin Verilerinin Gösterimi	44
Grafik 4.2.5.2.:	1968-2017 Türkiye GSYİH Kestirim Grafiği	44
Grafik 4.2.5.3.:	2018-2029 Türkiye GSYİH Tahmini Verileri	45
Şekil 2.2.1.:	Durum Uzay Modeli İçin Bağımlılık Yapısı	12
Şekil 3.3.5.1.:	$Y_{1:t}$ 'den y_{t+k} 'ya Bilgi Akışı.....	33

KISALTMALAR DİZİNİ

AR-GE	Araştırma Geliştirme
ARMA	Otoregresif Hareketli Ortalama Süreci
DLM	Dinamik Doğrusal Modeller
DUM	Durum Uzay Modelleri
GSMH	Gayri Safi Milli Hâsıla
GSYİH	Gayri Safi Yurtiçi Hâsıla
MAD	Ortalama Mutlak Sapma
MAPE	Ortalama Mutlak Yüzde Hataları
MSE	Ortalama Kare Hataları
TDA	Tekil Değer Ayırıştırması

GİRİŞ

Dinamik doğrusal modellerin(DLM) istatistiksel tarihi 1880 yıllarına dayanmakta olmasına rağmen 1960'lı yılların başlarında, dinamik sistemlerin izlenmesi ve kontrol edilmesi için mühendislikte geliştirilmiştir. DLM alanındaki ilk popüler uygulamalar, Apollo ve Polaris havacılık programlarında meydana gelmiştir ancak son yıllarda dinamik doğrusal modeller ve daha genel olarak durum uzay modelleri, biyolojiden ekonomiye, çok geniş bir alandaki uygulamalarla, büyük bir popülerite kazanmıştır. Bunların yanı sıra jeofizik bilimden genetiğe kadar mühendislik ve kalite kontrolü alanlarında kullanılmıştır. Dinamik doğrusal modelleri kullanan istatistiksel zaman serileri analizi, 1970-80'lerde büyük ölçüde gelişmiştir ve zamanla durum uzay modelleri(DUM) günümüzde ilgi odağı haline gelmiştir. Dinamik doğrusal modeller, durağan olmayan zaman serilerinin veya yapısal değişikliklerin modellenmesinde çok daha fazla esneklik sağlar ve genellikle daha kolay yorumlanmasını sağlar. Ayrıca daha genel durum alanı modeller sınıfını, analizi Gauss olmayan ve doğrusal olmayan dinamik sistemlere uygular. Dinamik doğrusal modelleri tahmin etmek için farklı yaklaşımlar vardır bunlara örnek olarak en küçük kareler veya maksimum olabilirlik örnek olarak gösterilebilir.

Kalman (1960), Bayesian yaklaşımına uygun olduğunu söyleyebileceğimiz dinamik doğrusal modellerin bazı temel kavramlarının altını çizmektedir. İlk adım deterministikten stokastik bir sisteme geçiyor; unutulmuş değişkenler, ölçüm hataları veya kusurlar nedeniyle daima mevcut olan belirsizlik olasılıkla açıklanmaktadır. Sonuç olarak, ilgili miktarlar (özellikle, sistemin t zamanında durumu) tahmini, mevcut bilgi dikkate alındığında, koşullu dağılımı hesaplanarak çözülür. Bu, Bayesci çıkarımda genel ve temel bir kavramdır. Dinamik doğrusal modeller, bir dinamik sistemin çıkışını, örneğin bir zaman serisini, rastgele hatalardan etkilenmeyen ve gözlemlenemeyen durum sürecinin bir fonksiyonu olarak tarif etme fikrine dayanmaktadır. Gizli değişkenler üzerinde koşullanarak, verilerdeki zamansal bağımlılığı modellemenin bu yolu, basit ve son derece güçlüdür. Dinamik doğrusal modellerin bir diğer önemli avantajı, hesaplamaların yineleyici bir şekilde yapılabilmesidir, yani tüm geçmişin depolanmasını gerektirmeksizin, ilgilinin koşullu dağılımları güncellenebilir. Bu, verilerin zaman içinde sıralı olarak gelmesi ve güncel çıkarımın gerekli olması

durumunda son derece avantajlıdır ve ihtiyaç duyulan depolama kapasitesinin azaltılması, büyük veri setleri için daha da önemli hale gelmektedir.

Bu tezin amacı, DLM ve DUM modellerini tanıtır, Türkiye için elde edilen 1968-2017 yılları arasındaki Gayri Safi Yurtiçi Hasıla(GSYİH) verilerini kullanarak 2018-2029 yılları arasında meydana gelecek değişimi DLM yöntemini kullanarak tahmin etmektir. Bu amaç doğrultusunda, Türkiye'nin GSYİH'sine ait veriler milyon dolar cinsinden Dünya Bankası resmi web adresinden alındı. Elde edilen veriler R programına yüklenip veriler DLM olarak tanıtıldıktan sonra analizler yapıldı.

Bu tez toplamda beş bölümden oluşmaktadır. İlk bölümde Dinamik Doğrusal Modelleri ve Durum Uzay Modellerine Genel Bakış adı altında, DLM ve DUM modellerinin tarihi, kullanım alanları ve bu modeller hakkında genel bilgiler verilmektedir. İkinci bölümde, doğrusal regresyon modellerinde Bayes çıkarımı ve DUM tanıtılıp bu modeller hakkında genel bilgiler verilmektedir. Üçüncü bölümde, Dinamik Doğrusal Modeller adı altında, DLM'lerin tanıtımı yapılmaktadır. Ayrıca bu bölümde yer alan, R programında DLM, Durum Hesaplama ve Tahmini, Filtreleme, DLM için Kalman Filtresi, Eksik Gözlemlerle Filtreleme, Düzleştirici(Smoother), DLM için Kalman Düzleştirici ve DLM için Tahmin başlıkları sırasıyla açıklanmaktadır. Dördüncü bölümde ise, Türkiye'ye Ait GSYİH'nin DLM İle Tahmini başlığı altında Türkiye'ye ait GSYİH verileri kullanılarak DLM uygulama kısmında üçüncü bölümde yer alan başlıkların sırasıyla uygulaması yapılmaktadır. Son bölüm olan beşinci bölümde ise, sonuç ve değerlendirme kısmı yer almaktadır.

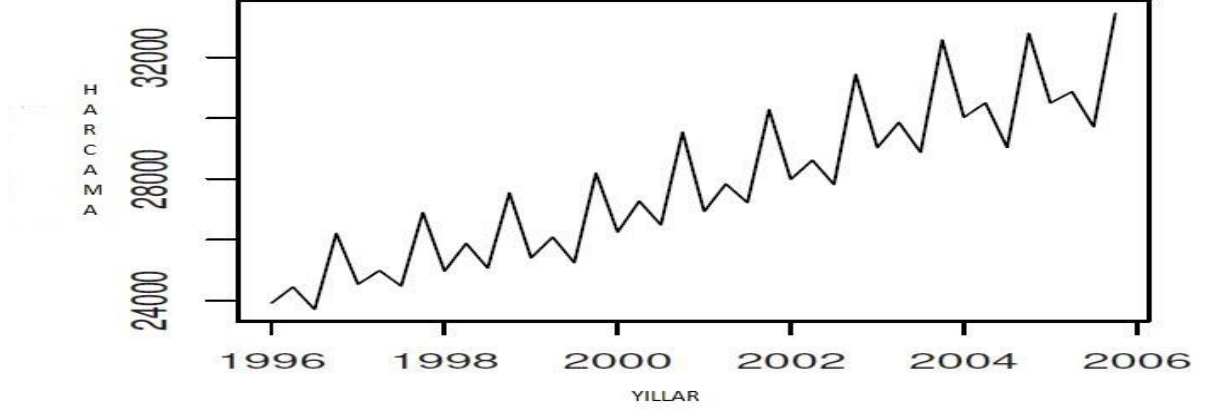
BİRİNCİ BÖLÜM

1. DİNAMİK DOĞRUSAL MODELLERİ VE DURUM UZAY MODELLERİNE GENEL BAKIŞ

Gaussian tarafından ele alınan durum uzay modelleri, aynı zamanda dinamik doğrusal modeller olarak adlandırılmaktadır. Son yıllarda zaman serileri analizinde durum uzay modellerinin uygulanmasına artan bir ilgi vardır; Harvey (1989), West ve Harrison (1997), Durbin ve Koopman (2001), Kunsch (2001) ve Migon vd. (2005) bu alanda çalışmalar yapmışlardır. Durum uzay modelleri, rastgele problemler tarafından bozulan dinamik bir sistemin çıkışı olarak bir zaman dizisini dikkate alır. Durum uzay modelleri, bir zaman dizisinin, trend, mevsimsel veya regresif bileşenler gibi birkaç bileşenin kombinasyonu olarak doğal bir yorumuna izin verir. Aynı zamanda, çok geniş bir uygulama yelpazesi için esnek bir çerçeve sunan mükemmel ve güçlü bir olasılıksal yapıya sahiptir. Bu alandaki hesaplamalar, yinelemeli algoritmalar ile uygulanabilir. Tahmin etme problemleri, mevcut bilgiler ışığında, ilgili miktarların koşullu dağılımını yinelemeli olarak hesaplayarak çözülür. Bu anlamda, Bayesian yapısı çerçevesinde oldukça doğal bir şekilde düzenli hale getirilirler.

Durum uzay modelleri, durağan olmayan, yapısal değişimler ve düzensiz kalıpların varlığında, tek değişkenli veya çok değişkenli zaman serilerini modellemek için kullanılabilir. Zaman serileri analizinde durum uzay modellerinin olası uygulamaları için bir yapı geliştirmek amacıyla, Grafik 1.1.'de çizilen veriler dikkate alınabilir. Grafik 1.1. Giovanni Petris vd. (2009) çalışmalarında R ile Dinamik Doğrusal Modeller adlı kitabından alınmıştır. Bur da yer alan veriler, İtalya'ya ait aile gıda harcamaları çeyrek verileridir (1996Q1 - 2005Q4). Bu zaman serisi oldukça tahmin edilebilir görünüyor, çünkü zaman içinde davranışlarını düzenli olarak tekrarlıyor. Biraz daha fazla değişkenlik gösteren bir trend ve oldukça düzenli bir mevsimsel bileşen görüyoruz. Aslında, temel zaman serileri analizi, incelenen verinin davranışında makul bir düzenlilik bulma olasılığına dayanmaktadır. Yani seriler, zaman içinde düzenli bir kalıbı tekrarlamaya eğilimli ise, gelecekteki davranışları tahmin etmek daha kolay olmaktadır.

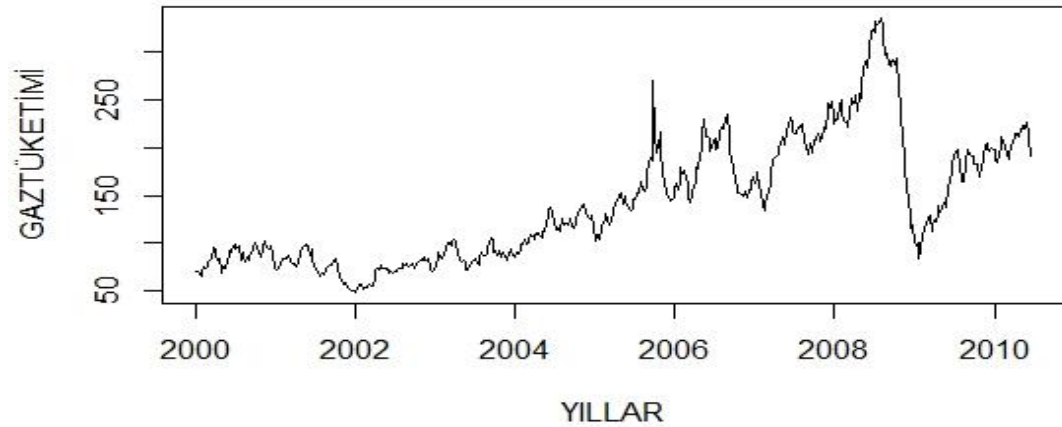
Grafik 1.1. Aile Gıda Harcamaları, Çeyrek Veriler (1996Q1 - 2005Q4).
<http://con.istat.it> adresinden veriler elde edilebilir.



Grafik 1.1. incelendiğinde değişken bir trendin, zaman içinde düzenli bir tekrarın ve düzenli mevsimsel bir değişimin olduğu görülmektedir. Bu özelliklere sahip verilerin olduğu bir zaman serisinde geleceğe dair tahminlerde bulunmak daha kolaydır. Şekilde aile gıda harcamalarında genel olarak bir artışın olduğu görülmektedir.

Aşağıda yer alan Grafik 1.2. 2000'den 2010'a kadar İngiltere'nin çeyreklik gaz tüketimini göstermektedir (veriler, UKgas olarak R programında mevcuttur). Bur da çizilenler gibi zaman serileri için işler daha karmaşık haldedir.

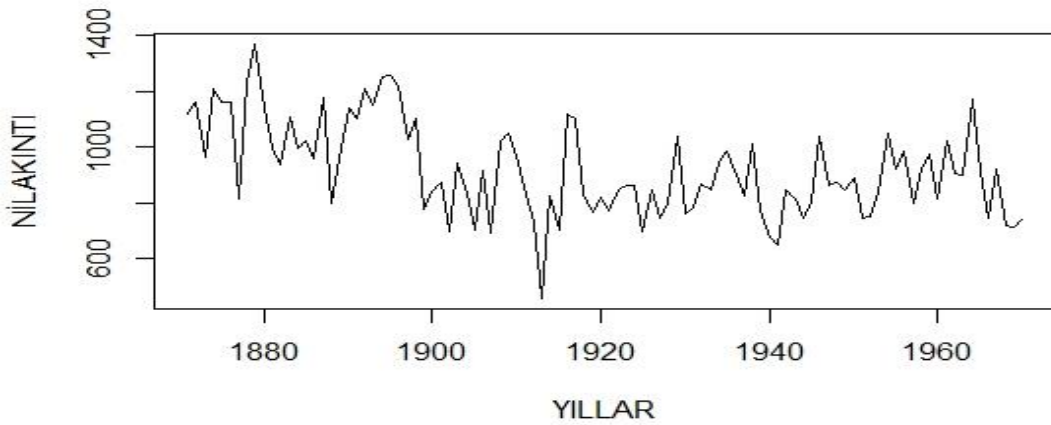
Grafik 1.2. 2000Q1- 2010Q4 Çeyreklik İngiltere'deki Gaz Tüketimi



Grafik 1.2. incelendiğinde mevsimsel bileşende bir değişiklik olduğunu açıkça görüyoruz. Bu verilere ait bir zaman serisinde geleceğe dair tahmin de bulunmak daha zordur. Şekilde mevsimsel olarak gaz tüketiminin büyük oranda artış veya düşüş yaşadığı görülmektedir.

Grafik 1.3. iyi çalışılmış bir veri setini göstermektedir: 1871'den 1970'e kadar Ashwan'daki Nil nehrinin yıllık akışının ölçümlerine ait seri seviye değişimlerini göstermektedir (veriler ve grafik, dlm paketi olarak R programında mevcuttur). Ashwan'ın ilk barajının yapımı 1898'de başladı; ikinci büyük baraj ise 1971'de tamamlandı. Grafik 1.1. ve 1.2.'ye nispeten daha karmaşık olan ve dalgalanmaların sık olduğu Grafik 1.3. için geleceğe dair tahminlerin gittikçe zorlandığını görmekteyiz. Şekillerden anlaşılacağı üzere, geleceğe dair tahminlerde bulunmak için veya sağlıklı sonuçlarda bulunabilmek için daha esnek zaman serisi modellerine ihtiyaç duymaya başlıyoruz, bu da temel sistemin sabit bir modelini ve kararlılığını kabul etmiyor, ancak değişim noktalarını veya yapısal kırılmaları içerebilen bir modelin olabileceği görülmektedir.

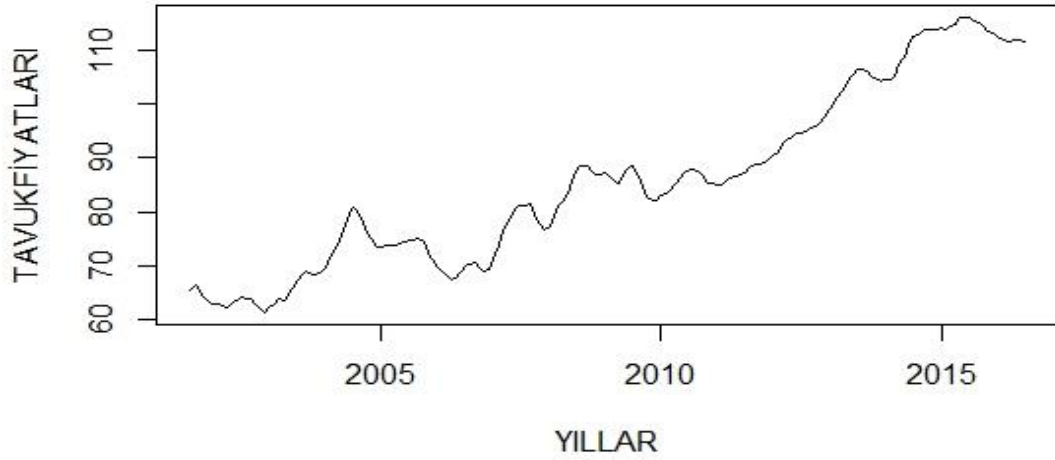
Grafik 1.3. Ahswan'daki Nil Nehri'nin Yıllık Akış Ölçümleri, 1871-1970



Grafik 1.3. analiz edildiğinde mevsimsel bileşende bir değişiklik olduğunu açıkça görüyoruz. Bu verilere ait bir zaman serisinde geleceğe dair tahmin de bulunmak daha zordur. Şekilde baraj yapımından sonra yıllık nehir akışında bir düşüşün olduğu ve yıllara göre değişiklik gösterdiği görülmektedir.

Daha düzensiz verilere sahip olan Grafik 1.4. veriler, R programı için kullanılan “Asts” paketi indirilebilir. Bur da yer alan veriler, İngiltere’de aylık tavuk fiyatlarını göstermektedir (01.2001-12.2016). Bu seri açıkça durağan olmayan ve oldukça düzensiz bir seri olduğu görülmektedir.

Grafik 1.4. İngiltere’de Aylık Tavuk Fiyatları



Grafik 1.4. analiz edildiğinde; serinin durağan olmadığı ve istikrarsız bir artışın olduğu gözlemlenmektedir. Durağan olmayan bir verinin analiz edilmesinin sağlıksız sonuçlar vereceği birçok çalışmada yer almaktadır. Gaussian tarafından Dinamik doğrusal modeller olarak adlandırılan Durum uzay modellerinin bazı avantajları;

- Durağan olmayan zaman serilerinin ARMA modelleri ile analizi, en azından durağanlık elde etmek için verilerin ön dönüşümü gerektirir bu dönüşüm veri kaybına neden olmaktadır. Durum uzay modellerinde böyle bir dönüşüm gerekmediğinden bu veri kaybı engellenmiş olur.
- Ancak, ortalama düzeyde değişkenlik, yapısal kırılma ve ani sıçramalarda istikrarsızlık gösteren verileri doğrudan analiz etmemize olanak tanıyan modellere sahip olmak daha doğal sonuçlar elde etmemizi sağlamaktadır.
- Durum uzay modelleri, özel bir durum olarak ARMA modellerini içerir, ancak verilerin ön dönüşümü gerektirmeksizin durağan olmayan zaman serilerine uygulanabilir.

- Fakat burada önemli olan başka bir durum söz konusudur. Örneğin, ekonomik verilerle işlem yapılırken kullanılan tek değişkenli bir zaman serisi modeli oldukça sınırlıdır. Grafik 1.4.'te yer alan tek değişkenli zaman serisi modeli yüksek frekanslı veriler, hızlıca düzensizliklere, yapısal kırılmalara veya sıçramalara uyum sağladığı için tatmin edici olabilir. Fakat bunun yanı sıra, piyasaları etkileyen ekonomik ve sosyo-politik değişkenleri daha derin ve daha geniş bir araştırma yapılmadan ani değişiklikler önceden tahmin edilemez. Bu durumda değişiklikleri tahmin etmek hiçte kolay değildir. Bunun çözümü, modele gerekli olan değişkenleri dâhil etmekle olmaktadır. Durum uzay modellerinde regresyon terimleri oldukça doğaldır ve durum uzay modelleri genellikle çok değişkenli zaman serileri için formüle edilebilir, şeklinde sıralanabilir.

Varlık problemi stokastik süreçler ve zaman serileri teorisinde her zaman temel ve ilgi çekici bir konu olmasına rağmen, durum uzay modelleri altmışlı yılların başında mühendislikten esinlenerek meydana gelmiştir. Kolmogorov (1941), bu sorunu, Wold (1938) tarafından önerilen bir temsil kullanarak, sabit stokastik süreçler için incelemiştir. Wiener (1949) sürekli stokastik süreçleri üzerinde çalışarak, Wiener-Hopf integral denkleminin çözümünü tahmin etme problemini çözmeye çalışmıştır. Soruna yeni bir bakış, Kalman (1960) tarafından, Bode-Shannon rasgele süreçlerin temsilini ve dinamik sistemlerin “durum geçişi” analiz yöntemini kullanarak verilmiştir. Kalman'ın çözümü, Kalman filtresi (Kalman; 1960; Kalman ve Bucy; 1963) olarak bilinen durağan ve durağan olmayan rasgele süreçler için geçerli sayılmıştır. Bu yöntemler zamanla diğer alanlarda hızla popülerlik kazanmış, tarımdan ekonomiye kadar geniş bir yelpazeye yayılmış problemlere uygulanmıştır. Lauritzen (2002) ve Plackett (1950) erken bir dönemde bu yöntemleri istatistik literatürüne kazandırmalarına rağmen bu yöntemlerin önemi istatistikçiler tarafından ancak daha sonra fark edilmiştir. Bu gecikmenin bir nedeni, Kalman filtresindeki çalışmanın çoğunlukla mühendislik literatüründe yayınlanmış olmasıdır. Başka bir neden de istatistikçilerin bu konu hakkında yeterince gayret göstermedikleri şeklinde söylenebilir.

Durum uzay modelleri, 70'li yıllarda zaman dizisi literatüründe (Akaike; 1974; Harrison ve Stevens; 1976) ortaya çıkmış ve seksenlerde tamamen yer edinmiştir

(Harvey; 1989; West ve Harrison; 1997; Aoki; 1987) son yıllarda ilgi odağı haline gelen durum uzay modelleri, bir yandan zaman serileri analizine çok uygun modellerin geliştirilmesine, aynı zamanda moleküler biyoloji veya genetik gibi daha geniş bir uygulama yelpazesine ve diğer yandan hesaplama araçlarının geliştirilmesine kadar uzanmaktadır.



İKİNCİ BÖLÜM

2. DOĞRUSAL REGRESYON MODELİNDE BAYES ÇIKARIMI VE DURUM UZAY MODELLERİ

2.1. Doğrusal Regresyon Modelinde Bayes Çıkarımı

Dinamik doğrusal modeller, regresyon katsayılarının zaman içinde değişmesine izin verildiğinde standart doğrusal regresyon modelinin genelleştirilmesi olarak kabul edilebilir. Doğrusal regresyon modeli, Y değişkenini açıklayıcı değişkenler x ile ilişkilendirmek için en popüler araçtır ve aşağıdaki şekilde tanımlanır;

$$Y_t = x_t' \beta + \varepsilon_t, \quad t = 1, \dots, n, \quad \varepsilon_t \stackrel{iid}{\sim} N(0, \sigma^2) \quad (2.1.1)$$

Bu denklemde Y_t bir rastgele değişkendir ve x_t ve β p boyutlu vektörlerdir. Temel formülasyonunda, x değişkenleri deterministik veya eksojen olarak kabul edilir. Stokastik regresyonda ise x rassal değişkenlerdir. İkinci durumda her bir t için bir rastgele $(p + 1)$ boyutlu vektör (Y_t, X_t) vardır. Bu durumda ortak dağılım belirtilmeli ve doğrusal regresyon modeli türetilmelidir. Bunu yapmanın bir yolu ortak dağılımın aşağıdaki şekilde yer alan modelin Gaussian olduğunu varsaymaktır.

$$\begin{bmatrix} Y_t \\ X_t \end{bmatrix} | \mu, \Sigma \sim N(\mu, \Sigma), \quad \mu = \begin{bmatrix} \mu_y \\ \mu_x \end{bmatrix}, \quad \Sigma = \begin{bmatrix} \Sigma_{yy} & \Sigma_{yx} \\ \Sigma_{xy} & \Sigma_{xx} \end{bmatrix}$$

Çok değişkenli Gauss dağılımının özelliklerinden olan ortak dağılımı X_t için bir marjinal modele ve Y_t için $X_t = x_t$ olarak verilen aşağıdaki şekilde verildiği gibi koşullu modelle ayrıştırabiliriz;

$$X_t | \mu, \Sigma \sim N(\mu_x, \Sigma_{xx}),$$
$$Y_t | x_t, \Sigma \sim N(\beta_1 + x_t' \beta_2, \sigma^2),$$

burada;

$$\beta_2 = \Sigma_{xx}^{-1} \Sigma_{xy},$$

$$\beta_1 = \mu_y - \mu_x' \beta_2,$$

$$\sigma^2 = \Sigma_{yy} - \Sigma_{yx}\Sigma_{xx}^{-1}\Sigma_{xy}$$

Eğer önceki (μ, Σ) dağılımı, marjinal modelin ve koşullu modelin parametrelerinin bağımsız olduğu şekildeyse, $(Y_t, X_t, \beta, \Sigma)$ dağılımında bir kesinti olur. Başka bir deyişle, eğer bizim ilgimiz esas olarak Y değişkeninde ise dikkatimizi koşullu modele vermeliyiz. Bu durumda regresyon modeli, verilen Y_t (β, Σ) ve x_t 'nin koşullu dağılımını açıklar. Model 2.1'i aşağıdaki şekilde yazabiliriz;

$$Y|X, \beta, V \sim N_n(X\beta, V) \quad (2.1.2)$$

burada $Y = (Y_1, \dots, Y_n)$ ve X , $n \times p$ matrisidir. Denklem (2.1), bir çapraz kovaryans matrisini ($V = \sigma^2 I_n$) yani Y_t 'lerin aynı varyans σ^2 ile koşullu olarak bağımsız olduğunu ifade etmektedir.

basit bir formun hareketi

$$\theta_t = \theta_{t-1} + v + w_t, \quad w_t \sim N(0, \sigma_w^2) \quad (2.1.2.3)$$

burada v , bilinen bir nominal hızdır, w_t ortalaması sıfır ve varyansı bilinen σ_w^2 ile bir Gaussian rasgele hatasıdır. Böylece, hedefin bilinmeyen pozisyonunu ardışık zaman noktalarında açıklayan bir $(\theta_t: t = 1, 2, \dots)$ işlemine sahibiz. Gözlem denklemini şimdi

$$Y_t = \theta_t + \varepsilon_t, \quad \varepsilon_t \stackrel{iid}{\sim} N(0, \sigma^2) \quad (2.1.2.4)$$

şeklinde yazabiliriz.

2.2. Durum Uzay Modelleri

$(Y_t)_{t \geq 1}$ şeklinde bir zaman seri düşünelim, ortak sonlu boyutlu dağılımların belirlenmesi ($Y_1, \dots, Y_T$) herhangi bir $t \geq 1$ için kolay bir durum değildir. Özellikle, zaman serileri uygulamalarında bağımsızlık veya değiş tokuş edilebilirlik varsayımları nadiren desteklenebilmektedir, çünkü bunlar esasen zamanla ilgisiz durumdadır. Markoviya bağımlılığı, zaman içinde kesin bir rolün olduğu Y_t 'ler arasındaki en basit bağımlılık biçimidir. Markov zincirini $(Y_t)_{t \geq 1}$ 'nin herhangi bir $t > 1$ için,

$$\pi(y_t | y_{1:t-1}) = \pi(y_t | y_{t-1})$$

şeklinde yazabiliriz. Bu, tüm gözlemlerin $t - 1$ zamanına kadar taşıdığı Y_t ile ilgili bilgilerin sadece y_{t-1} tarafından taşınan bilgilerle tamamen aynı olduğu anlamına gelir. Bu durumu Y_t ve $y_{1:t-2}$ 'nin y_{t-1} olması şartıyla koşullu olarak bağımsız olması halinde de söyleyebiliriz. Bir Markov zinciri için sonlu boyutlu birleşim dağılımlarını aşağıda yer aldığı şekilde de yazabiliriz;

$$\pi(y_{1:t}) = \pi(y_1) \cdot \prod_{j=2}^t \pi(y_j | y_{j-1})$$

Bununla birlikte gözlemler için birçok uygulamada kullanılması uygun olmayan bir Markov yapısının olduğu varsayılmaktadır. Durum uzay modelleri, gözlemler için daha karmaşık modeller tanımlamak için bir Markov zincirinin nispeten basit bağımlılık yapısı üzerine inşa edilir. Bir durum uzay modelinde, durum süreci olarak adlandırılan gözlenemeyen bir Markov zinciri (θ_t) olduğunu ve Y_t , θ_t 'nin kesin olmayan bir ölçümü olduğu varsayılmaktadır. Mühendislik uygulamalarında genellikle Y_t çıkışı üreten fiziksel olarak gözlemlenebilir bir sistemin durumu açıklanmaktadır. Öte yandan, ekonometrik uygulamalarda θ_t çoğu zaman yararlı bir yoruma sahip olabilen gizli bir yapıdır. Her durumda, (θ_t) gözlemlenebilir zaman serilerinin (Y_t) olasılık dağılımını belirleme görevini kolaylaştıran yardımcı bir zaman dizisi olarak düşünebilir.

Formal olarak bir durum uzay modeli, aşağıdaki varsayımları karşılayan, bir R^p değerli zaman serisinden (θ_t , $t = 0, 1, \dots$) ve bir R^m değerli zaman serisinden (Y_t , $t = 1, 2, \dots$) oluşur.

Varsayımlar;

(A.1) (θ_t) bir Markov zinciridir.

(A.2) Şartlı olarak (θ_t), Y_t 'ler bağımsızdır ve Y_t yalnızca θ_t , 'ye bağlıdır.

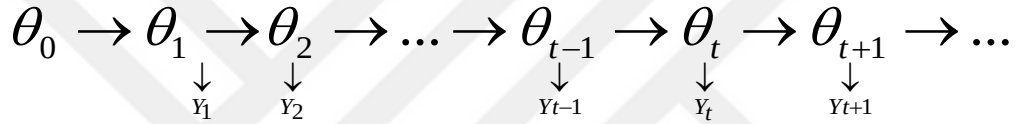
(A.1) ve (A.2) 'nin sonucu olarak bir başlangıç uzaysal modelinin tamamen, ilk $\pi(\theta_0)$ ve koşullu yoğunluklar olan $\pi(\theta_t | \theta_{t-1})$ ve $\pi(Y_t | \theta_t)$ ile belirtilmiş olmasıdır. Aslında, herhangi bir $t > 0$ için,

$$\pi(\theta_{0:t}, y_{1:t}) = \pi(\theta_0) \cdot \prod_{j=1}^t \pi(\theta_j | \theta_{j-1}) \pi(y_j | \theta_j) \quad (2.2.1)$$

şeklinde yazılabilir. (2.2.1.) 'den, koşullandırma veya marjinalleştirme yoluyla, herhangi bir diğer ilişkili dağılım türetilebilir. Örneğin, $y_{1:t}$ gözlemlerinin orta yoğunluğu, (2.2.1) 'de θ_j 'lerin bütünleştirilmesiyle elde edilebilir. Bununla birlikte, bu şekilde (2.2.1) 'in basit ürün formunun kaybolduğu dikkate alınmalıdır.

Cowell vd.(1999) bir durum uzay modelinin varsaydığı bilgi akışını Şekil 2.2.1'de görüldüğü gibi ele almışlardır. Şekildeki grafik, yönlendirilmiş bir döngüsüz grafiğin özel bir durumudur. Modelin grafiksel gösterimi, bir durum uzay modelinde meydana gelen rastgele değişkenlerin koşullu bağımsızlık özelliklerini ortaya çıkarmak için kullanılabilir.

Şekil 2.2.1. Durum Uzay Modeli İçin Bağımlılık Yapısı



Aslında, rastgele değişkenlerden oluşan A ve B değişkenlerinin üçüncü bir değişkeni olan C, koşullu olarak bağımsız olarak gösterilebilir ancak ve ancak C değişkeni, A ve B değişkeninin ayırır yani A değişkenindeki bir değişkeni B'ye bağlayan herhangi bir yol C'den geçer. Önceki ifadede Şekil 2.6'daki okların, grafiğin her iki yönde de çevrilebilecek tersi olmayan kenarları olarak kabul edilmesi gerektiği dikkate alınmalıdır. Örnek olarak, y_t ve $(\theta_{0:t-1}, y_{1:t-1})$ 'nin θ_t verildiğinde koşullu olarak bağımsız olduğunu göstermek için Şekil 2.6'yı kullanacağız. Basit bir şekilde ispatlamaya çalıştığımızda, y_t 'yi önceki y_s 'den ($s < t$) birine veya θ_s , ($s < t$), durumlarından birine sahip olan herhangi bir yolun θ_t 'den geçmesi gerektiğini gözlemlemekten ibarettir dolayısıyla $\{\theta_t\}$, $\{\theta_{0:t-1}, y_{1:t-1}\}$ ve $\{\theta_t\}$ 'yi ayırır. Bu durumu;

$$\pi(y_t | \theta_{0:t-1}, y_{1:t-1}) = \pi(y_t | \theta_{0:t}) \text{ şeklinde ifade edebiliriz.}$$

Benzer şekilde, θ_t ve $(\theta_{0:t-2}, y_{1:t-1})$, θ_{t-1} verildiğinde koşullu olarak bağımsız olduğunu gösterebilir ve koşullu dağılımlar cinsinden aşağıdaki şekilde ifade edilebilir;

$$\pi(\theta_t | \theta_{0:t-1}, y_{1:t-1}) = \pi(\theta_t | \theta_{t-1}).$$

ÜÇÜNCÜ BÖLÜM

3. DİNAMİK DOĞRUSAL MODELLER

Dinamik doğrusal model, doğrusal ve Gaussian olmak üzere genel durum uzay modelinin özel bir durumu olarak sunulur. Dinamik doğrusal modeller için, iyi bilinen Kalman filtresi tarafından tahmin ve tahminler tekrarlanabilir. Bir dinamik doğrusal model (DLM), $t = 0$ zamanında p -boyutlu durum vektörü için bir normal ön dağıtım ile belirtilir,

$$\theta_0 \sim N_p(m_0, C_0), \quad (3.1)$$

her $t \geq 1$ zamanı için bir çift denklemle birlikte,

$$Y_t = F_t \theta_t + v_t, \quad v_t \sim N_m(0, V_t), \quad (3.2)$$

$$\theta_t = G_t \theta_{t-1} + w_t, \quad w_t \sim N_p(0, W_t), \quad (3.3)$$

burada G_t ve F_t bilinen matrisler (sırasıyla $p \times p$ ve $m \times p$ değerleri), $(v_t)_{t \geq 1}$ ve $(w_t)_{t \geq 1}$ sıfır ve bilinen varyans matrisleri $(V_t)_{t \geq 1}$ ve $(W_t)_{t \geq 1}$ ile bağımsız Gauss rasgele vektörlerinin sırasıyla iki bağımsız dizisidir. Denklem (3.2) gözlem denklemi olarak adlandırılırken (3.3) durum denklemi veya sistem denklemdir. Ayrıca, θ_0 'ın v_t ve w_t 'den bağımsız olduğu varsayılmaktadır. Bir DLM'nin $Y_t | \theta_t \sim N(F_t \theta_t, V_t)$ ve $\theta_t | \theta_{t-1} \sim N(G_t \theta_{t-1}, W_t)$ ile önceki bölümde yer alan (A.1) ve (A.2)'nin varsayımlarını karşıladığını gösterebiliriz.

Denklem (3.1)'in aksine, θ_0 için bir önceki dağılım ile genel bir durum uzay modeli, gözlem ve evrim denklemleri ile birlikte belirtilebilir,

$$Y_t = h_t(\theta_t, v_t),$$

$$\theta_t = g_t(\theta_{t-1}, w_t),$$

burada h_t ve g_t isteğe bağlı fonksiyonlar için kullanılmaktadır. Doğrusal durum uzay modelleri, doğrusal işlevler olarak gt ve ht 'yi belirtir ve Gaussian doğrusal modeller, Gauss dağılımlarının varsayımlarını ekler. Normallik varsayımı birçok uygulamada mantıklıdır ve merkezi limit teorem argümanlarıyla doğrulanabilir. Ancak, modelleme aykırı değerleri için ağır kuyruklu hatalar veya ayrık zaman serilerinin işlenmesi için

dinamik genelleştirilmiş doğrusal model gibi birçok önemli uzantı vardır. Normallik varsayımı kaldırılırken meydana gelecek sorunlar, ek hesaplama zorluklarıdır.

Zaman serileri analizi için DLM kullanımı bu kısımda yer almaktadır. Tek değişkenli bir zaman serisi ($Y_t : t = 1, 2, \dots$) için en basit model,

$$\begin{aligned} Y_t &= \mu_t + v_t, & v_t &\sim N(0, V), \\ \mu_t &= \mu_{t-1} + w_t, & w_t &\sim N(0, W), \end{aligned} \quad (3.4)$$

tarafından tanımlanan rastgele yürüyüş artı gürültü modeli olarak adlandırılır. Hata dizileri (v_t) ve (w_t), hem içlerinde hem de bunların arasında bağımsızdır. Bu $m = p = 1$, $\theta_t = \mu_t$ ve $F_t = G_t = 1$ 'e sahip bir DLM'dir. Dinamizmde hız bulunmadığında, sezgisel olarak, net bir trend ya da mevsimsel değişim göstermeyen zaman serileri için uygundur. Gözlemler (Y_t), rasgele bir yürüyüşle tanımlanan zaman içinde rastgele değişimlere maruz kalan bir seviye μ_t 'nin gürültülü gözlemleri olarak modellenir. Bu nedenle modele yerel seviye modeli de denir. Eğer $W = 0$ ise, sabit ortalama modele geri dönmüş olunur. Rastgele yürüyüşün (μ_t) durağan olmadığı göz ardı edilmemelidir. DLM'ler durağan olmayan zaman serilerini modellemek için kullanılabilir. Aksine ARMA modelleri, durağanlığa ulaşmak için verilerin ön dönüşümü gerektirir. Biraz daha detaylandırılmış bir model, yerel seviye modeli ile aynı gözlem denklemine sahip olan doğrusal olmayan büyüme modeli veya yerel doğrusal eğilimdir fakat dinamikler arasında zamanla değişen bir eğim içeren μ_t için;

$$\begin{aligned} Y_t &= \mu_t + v_t, & v_t &\sim N(0, V), \\ \mu_t &= \mu_{t-1} + \beta_{t-1} + w_{t,1}, & w_{t,1} &\sim N(0, \sigma_\mu^2) \\ \beta_t &= \beta_{t-1} + w_{t,2}, & w_{t,2} &\sim N(0, \sigma_\beta^2) \end{aligned}$$

ilgisiz hatalar v_t , $w_{t,1}$ ve $w_{t,2}$ ile yukardaki şekilde yazılır.

$$\theta_t = \begin{bmatrix} \mu_t \\ \beta_t \end{bmatrix}, \quad G = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad W = \begin{bmatrix} \sigma_\mu^2 & 0 \\ 0 & \sigma_\beta^2 \end{bmatrix}, \quad F = [1 \ 0]$$

şekliyle bir Dinamik Doğrusal Modeldir. Sistem, varyans σ_μ^2 ve varyans σ_β^2 'nin sıfır olmasına izin verir. Bu örneklerde G_t , F_t matrisleri ve V_t , W_t matrislerinin sabit olduğu

göz ardı edilmemelidir. Bu durumda modelin zamanla değişmez olduğu söylenir. Özellikle, popüler Gaussian ARMA modelleri, özel DLM vakaları olarak elde edilebilir. Hannan ve Deistler(1988), Gaussian ARMA ve DLM modellerinin zamanla değişmeyen durumda eşdeğer olduğunu söylemektedirler.

DLM'ler, doğrusal regresyon modelinin bir genellemesi olarak kabul edilebilir ve zaman değişken regresyon katsayılarına izin verir. Basit, statik doğrusal regresyon modeli, bir Y değişkeni ve bir rassal olmayan açıklayıcı değişken x arasındaki ilişkiyi;

$$Y_t = \theta_1 + \theta_2 x_t + \varepsilon_t, \quad \varepsilon_t \stackrel{iid}{\sim} N(0, \sigma^2) \text{ olarak tanımlar.}$$

Burada $(Y_t, x_t, t = 1, 2, \dots)$ 'in zaman içinde gözlemlendiği şekliyle düşünebiliriz. x ve y arasındaki doğrusal olmayan fonksiyonel ilişkinin zamanla değişen regresyon parametrelerine izin verilmesi, çalışılan süreçteki yapısal değişimleri ve bazı değişkenlerin ihmalini modelleyebilir. Basit bir dinamik doğrusal regresyon modeli

$$Y_t = \theta_{t,1} + \theta_{t,2} x_t + \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma_t^2) \text{'yi varsayar,}$$

daha ileri bir denklem ile;

$$\theta_t = G_t \theta_{t-1} + w_t, \quad w_t \sim N_2(0, W_t) \text{ şeklinde modellenir.}$$

$F_t = [1, x_t]$ ve $\theta_t = (\theta_{t,1}, \theta_{t,2})$ durumu bir Dinamik Doğrusal Modelidir. Özel bir örnek olarak, $G_t = I$, birim matrisi, $\sigma_t^2 = \sigma^2$ ve her t için $w_t = 0$ olduğu bir durumda, basit statik doğrusal regresyon modelini elde etmiş oluruz.

3.1. R Programı ve *dlm*

R programı, istatistiksel yöntemlerin ve grafik çizim tekniklerinin birçoğunu kullanarak kullanıcılara güçlü altyapı ve ara yüz sunan bir bilgisayar programı olmasının yanı sıra R programı, araştırma alanlarına göre kategorize edilmiş alt paketler, geniş kütüphanesi ile kullanıcısına yeni fonksiyonlar ve ilave özelliklerin dahil edilmesiyle geliştirilebilir bir programlama dilidir. “R Geliştirme Çekirdek Ekibi” tarafından geliştirilmesi devam edilen programın kurucuları Yeni Zellanda Auckland Üniversitesinden Ross Ihaka ve Robert Gentleman’dır. R Programlama dili aynı zamanda S programlama dilinin açık kaynak kodlu sürümüdür. S programlama dili

arasında farklılıklar olmasına rağmen S programlama dili için yazılan kodlar R programlama dili ortamında da çalışabilmektedir. Diğer programlama dillerinin aksine R programlama dilinin büyük bir bölümü farklı kullanıcılar tarafından yazılmasına rağmen aynı düzende yazılır. Bu nedenle algoritmik seçimlerin takibi kolaylaşır. İleri seviye kullanıcılar kullanım kolaylığını sağlamak veya yeni özellikler dahil etmek için R programındaki nesnelere C programlama dili ile eklentiler yazabilirler. Petris (2010) R paketi, dinamik doğrusal modeller kullanarak *dml* için Bayesci çıkarsama için tümleşik bir ortam sağlamaktadır. Bilinmeyen model parametrelerinin hem Bayesci çıkarsama hem de en çok olabilirlik öngörüsü için gerekli tüm araçları *dml* paketi kapsamaktadır. *dml* paketi kullanıcının çoğunlukla olabilirlik tahmini için kullanılır. *dml* paketinin Kalman filtrelemesi ve düzleştirme için de fonksiyonları bulunmaktadır. Bu paketin esas gücü simülasyon altyapısı ile Bayesci çıkarsama için sağladığı araçlardadır.

3.1.1. *dml* Paketinde Dinamik Doğrusal Modeller

Dinamik Doğrusal Modeller, *dml* paketinde sınıf özneliğine sahip listelenmiş listeler olarak temsil edilir bu da onları “*dml*” sınıfının nesneleri haline getirir. Sınıf *dml*'nin nesneleri sabit veya zamanla değişen DLM'leri temsil edebilir. Sabit bir DLM, F , V , G , W , C_0 matrisleri ve m_0 vektörü verildiğinde tamamen belirtilir. R'de, bu bileşenler sırasıyla, FF , V , GG , W , $C0$ ve $m0$ elemanları olarak bir *dml* nesnesinde depolanır. Ekstratör ve değiştirme fonksiyonları, modelin belirli bölümlerine kullanıcı olarak erişmek ve değiştirmek için mevcuttur. Paket ayrıca, modeli ayrıntılı şekilde incelediğimiz minimum girişten DLM'lerin belirli sınıflarını oluşturan çeşitli işlevler de sağlar. Genel tek değişkenli veya çok değişkenli DLM, *dml* fonksiyonu kullanılarak belirtilebilir. Bu fonksiyon bileşenlerinden bir *dml* nesnesi oluşturur bu da girdi üzerinde bazı akıl yürütme kontrollerini gerçekleştirir, örneğin tutarlılık için matrislerin boyutlarını test etme gibi. Girdi, adlandırılmış argümanlarla veya bağımsız argümanlar olarak bir liste şeklinde verilebilir.

dml paketinde yer alan fonksiyonlar oldukça kullanışlıdır. Uygulamada yaygın olarak karşılaşılan polinomiyal modeller, *dml* paketinde basitleştirilmiş ve bu tür modellerin belirtilmesi için *dmlModPoly* ile gösterilen özel bir fonksiyon tanımlanmıştır. *dmlModPoly*'ye ek olarak, *dml* paketi dinamik doğrusal modellerin

standart türlerini yaratmak amacıyla başka fonksiyonlar da içermektedir. Özel modellerin R’da tanıtılması için kullanılabilen bazı fonksiyonlar aşağıda yer alan çizelge 3.1.1.1.’de özetlenmiştir.

Çizelge 3.1.1.1. R Programında Dinamik Doğrusal Modellere Ait Bazı Fonksiyonlar

Fonksiyon	Model
<i>dmlModARMA</i>	<i>dmlModARMA(ar = NULL, ma = NULL, sigma2 = 1, dV, m0, C0)</i>
<i>dmlModPoly</i>	<i>dmlModPoly(order = 2, dV = 1, dW = c(rep(0, order - 1), 1), m0 = rep(0, order), C0 = 1e+07 * diag(nrow = order))</i>
<i>dmlModReg</i>	<i>dmlModReg(X, addInt = TRUE, dV = 1, dW = rep(0, NCOL(X) + addInt), m0 = rep(0, length(dW)), C0 = 1e+07 * diag(nrow = length(dW)))</i>
<i>dmlModSeas</i>	<i>dmlModSeas(frequency, dV = 1, dW = c(1, rep(0, frequency - 2)), m0 = rep(0, frequency - 1), C0 = 1e+07 * diag(nrow = frequency - 1))</i>
<i>dmlModTrig</i>	<i>dmlModTrig(s, q, om, tau, dV = 1, dW = 0, m0, C0)</i>
<i>dmlFilter</i>	<i>dmlFilter(y, mod, debug = FALSE, simplify = FALSE)</i>
<i>dmlSmooth</i>	<i>dmlSmooth(y, ...)</i> ## Default S3 method: <i>dmlSmooth(y, mod, ...)</i> ## S3 method for class 'dmlFiltered' <i>dmlSmooth(y, ..., debug = FALSE)</i>
<i>dmlForecast</i>	<i>dmlForecast(mod, nAhead = 1, method = c("plain", "svd"), sampleNew = FALSE)</i>
<i>dmlRandom</i>	<i>dmlRandom(m, p, nobs = 0, JFF, JV, JGG, JW)</i>

Çizelge 3.1.1.1.'de yer alan fonsiyonların işlevleri; *dmlModARMA*, belirtilen tek değişkenli veya çok değişkenli ARMA sürecini temsil eden bir sınıf DLM nesnesi oluşturur. *dmlModARMA* çok değişkenli durum için de tanımlanabilirken; diğer yaratıcı fonsiyonlar tek değişkenli gözlemler için kullanışlıdır. *dmlModPoly* fonsiyonu bir N 'inci düzen polinom DLM oluşturur. *dmlModReg* fonsiyonu, doğrusal regresyon modelinin bir DLM temsilini oluşturur. *dmlModSeas* fonsiyonu, mevsimsel bileşenin DLM temsilini oluşturur. *dmlModTrig* fonsiyonu, belirli bir periyodik bileşeni temsil eden bir DLM oluşturur. *dmlFilter* fonsiyonu, varyans, kovaryans matrisleri ile birlikte durum vektörlerinin filtrelenmiş değerlerini hesaplamak için Kalman filtresini uygular. Varsayılan olarak işlev, "*dmlFiltered*" sınıfının bir nesnesini döndürür. "DlmFiltre" sınıfındaki nesneler için artıklar ve *tsdiag* yöntemleri mevcuttur. *dmlSmooth* fonsiyonu, varyans, kovaryans matrisleri ile birlikte, durum vektörlerinin düzeltilmiş değerlerini hesaplamak için Kalman'ın düzleştirme modelini uygular. *dmlForecast* fonsiyonu, gelecekteki gözlemlerin ve sistem durumlarının beklenen değerini ve varyansını değerlendirir. Gelecekteki gözlemlerin ve sistem durumlarının dağıtımından bir örnek de üretebilir. *dmlRandom* fonsiyonu, "DLM" sınıfının rastgele (sabit veya zamanla değişen) bir nesnesini ve bunlardan gelen durumları ve gözlemleri oluşturur, şeklindedir.

3.2. Durum Hesaplama ve Tahmini

Durum uzay modellerinin büyük esnekliği, büyük uygulamalı problemler yelpazesindeki kapsamlı uygulamalarının bir nedenidir. Tabii ki, herhangi bir istatistiksel uygulamada olduğu gibi, çok önemli ve genellikle zor bir adım dikkatli bir model belirtimidir. Sezgisel bir anlama sahip olan durum modellerinde birçok problemin statistikçi ve uzmanlar birlikte bir durum uzay modeli oluşturabilirler ve uzman bilgisi, durum denklemindeki geçiş olasılıklarını belirlemek, durum uzayının boyutunu belirlemek vb. durumlar için kullanılabilir. Bununla birlikte, çoğu zaman model oluşturma önemli bir zorluk olabilir. Fiziksel olarak yorumlanabilir durumların net bir şekilde tanımlanması mümkün olmayabilir veya durum alanı temsili benzersiz olamaz, durum uzay modeli çok büyük ve zayıf bir şekilde tanımlanabilir ya da model çok karmaşıktır. Burada, verilen $\pi(y_t|\theta_t)$ ve $\pi(\theta_t|\theta_{t-1})$ modellerini yoğunluklarının belirtildiğini varsayarak, hesap ve tahmin için temel tekrarları şeklinde ele alıyoruz.

Belirli bir durum uzay modeli için, temel görevler gözlemlenmemiş hallerde çıkarım yapmak ya da gözlem sırasının bir kısmına dayanarak gelecekteki gözlemleri tahmin etmektir. Tahmin ve tahmin, mevcut bilgiler dikkate alındığında, ilgili miktarların koşullu dağılımlarının hesaplanmasıyla çözülür.

Durum vektörünü tahmin etmek için, koşullu yoğunluklu $\pi(\theta_s|y_{1:t})$ hesaplamak için. Filtreleme sorunları ($s = t$ olduğunda), durum tahmini ($s > t$) ve düzeltme ($s < t$) arasında ayırım yapılır. Bu, filtreleme ve düzeltme arasındaki farkın altını çizmeye değeridir. Filtreleme probleminde, verilerin zaman içinde ardışık olarak gelmesi beklenir. Uygulanan pek çok problemde durum böyledir. Örneğin, hareket eden bir nesneyi veya bir kişinin tahmin günlük etmesi gereken finansal uygulamaları izleme problemini düşünün, faiz oranlarının yapısı, cari tahminlerin güncellenmesi ile yeni veriler piyasa üzerinde ertesi gün gözlenmektedir. Bu durumlarda, t vektörüne kadar olan gözlemlere dayanarak durum vektörünün mevcut değerini tahmin etmek ve yeni veriler $t + 1$ zamanında kullanıma sunulduğunda tahminlerin güncellemesi için bir tahmin her zaman kolaylık sağlar ve bunun için bir işlem gerekir. Filtreleme problemini çözmek için $\pi(\theta_t|y_{1:t})$ koşullu yoğunluk hesaplanır. Bir DLM'de Kalman filtresi, yeni veriler kullanıma sunuldukça durum vektöründeki güncel çıkarımı güncellemek için formül sağlar. Bu filtreleme yoğunluğu $\pi(\theta_t|y_{1:t})$ 'den $\pi(\theta_{t+1}|y_{1:t+1})$ 'ye geçmek içindir.

Düzeltilme ya da retrospektif analiz problemi, $y_1, \dots, y_t$ durumunda, $1, \dots, t$ sırasındaki durum dizisini tahmin etmekten oluşur. Birçok uygulamada, belli bir süre için bir zaman dizileri üzerinde gözlemler vardır ve geriye dönük olarak gözlemlerin altında yatan sistemin davranışını incelemek istenir. Örneğin, ekonomik çalışmalarda araştırmacı, belirli bir yıl boyunca tüketimin veya bir ülkenin gayri safi yurtiçi hasılasına sahip olabilir ve sistemin sosyo-ekonomik davranışını geriye dönük olarak anlamak isteyebilir. Yumuşatma problemi $\theta_{1:t}$ verilen $y_{1:t}$ 'nin koşullu dağılımı hesaplanarak çözülür. Filtrelemeye gelince, tekrarlayıcı bir algoritma olarak uygulanabilir.

Nitekim, zaman serilerinde analiz tahmini genellikle ana görevdir, durum tahmini ise gelecekteki gözlemlerin değerini tahmin etmek için sadece bir adımdır. Bir adımli öngörme için, yani $y_{1:t}$ 'ye dayanarak bir sonraki gözlemi Y_{t+1} tahmin eden, ilk

önce bir sonraki vektörün θ_{t+1} değerinin tahmin edilmesi ve ardından, bu tahmine dayanarak Y_{t+1} için tahmin hesaplanmaktadır. Bir adımlı öncelikli durum tahmini yoğunluğu $\pi(\theta_{t+1}|y_{1:t})$ 'dir ve θ_t 'nin filtreleme yoğunluğuna dayanmaktadır. Bundan, bir adımlık $\pi(y_{t+1}|y_{1:t})$ öncül yoğunluğu elde edilir.

3.3.Filtreleme

Filtreleme tarihi, iletişim mühendisliğinden daha eski olmasına rağmen Wiener ve Kolmogorov'un birbirlerinden bağımsız olarak geliştirdikleri kestirim teoremi, sinyal artı gürültü modellerindeki sinyallerin ayrıştırılmasına yönelik önemli bir çalışmadır ve filtreleme teorisinin kuramsal temellerinin meydana geldiği önemli bir teoremdir. Wiener ve Kolmogorov'a ait bu teoremi daha sonra Kalman (1960), Kalman ve Bucy(1960) yaptıkları çalışmalarda durum-konum modellerini kullanarak, dinamik doğrusal sistemlerin kestirimi için akışlı bir filtre sistemi oluşturmuşlardır. Kalman, problemi en uygun filtreyi tasarlamayı olanaklı hale getiren diferansiyel denklemlerle meydana getirmiştir. Bu sonuç, o dönemde kontrol mühendisleri tarafından kapsamlı olarak roketlerden kimyasal tesislere kadar hemen her şeyin kontrolünde kullanılmış; yapılan çalışmaların çoğu ağırlıklı olarak mühendislik alanlarında ilgi görmüştür. Young (1969), Kalman denklemlerinin Plackett'ın (1950) çalışmasının bir uzantısı olarak elde edilebileceğini tespit etmiştir; bu aynı zamanda, Kalman filtresiyle en küçük kareler prensibinin eşdeğerliğinin de bir yöntemi olarak kabul görür. Kalman'ın modeli üzerinde çalışan Ho, Lee (1964) ve Aoiki (1967) ise, bu modele Bayesci yaklaşımı ekleyerek, dinamik doğrusal süreçlerin akışlı kestirimi için kuramsal bir dayanak elde etmişlerdir. Aşağıdaki önerme, genel durum uzay modeli için filtreleme tekrarlarının bir sunumunu içerir.

Önerme 3.3.1.(Filtreleme tekrarları). Aşağıdaki ifadeler, (A.1) - (A.2)) ile tanımlanan genel durum uzay modeli için geçerlidir.

- i. Durumlar için tek adımlı ön tahmin yoğunluğu filtrelenmiş yoğunluk $\pi(\theta_{t-1}|y_{1:t-1})$ den aşağıda yer alan model gibi hesaplanabilir;

$$\pi(\theta_t|y_{1:t-1}) = \int \pi(\theta_t|\theta_{t-1})\pi(\theta_{t-1}|y_{1:t-1})d\theta_{t-1}. \quad (3.3.a)$$

- ii. Gözlemler için tek adımlı ön tahmin yoğunluğu, durumlar için öngörü yoğunluğundan aşağıdaki şekilde hesaplanabilir;

$$\pi(y_t|y_{1:t-1}) = \int \pi(y_t|\theta_t)\pi(\theta_t|y_{1:t-1})d\theta_t \quad (3.3.b)$$

- iii. Filtreleme yoğunluğu, yukarıdaki yoğunluklardan aşağıdaki gibi hesaplanabilir;

$$\pi(\theta_t|y_{1:t}) = \frac{\pi(y_t|\theta_t)\pi(\theta_t|y_{1:t-1})}{\pi(y_t|y_{1:t-1})} \quad (3.3.c)$$

İspat. Modelin şartlı bağımsızlık özelliklerine büyük ölçüde dayanır ve bu da şekil 2.2.1.'deki grafikten çıkarılabilir.

(İ) 'i kanıtlamak için; θ_t 'nin $Y_{1:t-1}$ 'den bağımsızlığından yola çıkarak, verilen θ_{t-1} . Böylece;

$$\begin{aligned} \pi(\theta_t|y_{1:t-1}) &= \int \pi(\theta_{t-1}, \theta_t|y_{1:t-1})d\theta_{t-1} \\ &= \int \pi(\theta_t|\theta_{t-1}, y_{1:t-1})\pi(\theta_{t-1}|y_{1:t-1})d\theta_{t-1} \\ &= \int \pi(\theta_t|\theta_{t-1})\pi(\theta_{t-1}|y_{1:t-1})d\theta_{t-1} \end{aligned}$$

şeklinde olur.

(İİ) 'i kanıtlamak için; Y_t 'nin $Y_{1:t-1}$ 'den bağımsızlığından yola çıkarak, verilen θ_t .

Böylece;

$$\begin{aligned} \pi(y_t|y_{1:t-1}) &= \int \pi(y_t, \theta_t|y_{1:t-1})d\theta_t \\ &= \int \pi(y_t|(\theta_t, y_{1:t-1}))\pi(\theta_t|y_{1:t-1})d\theta_t \\ &= \int \pi(y_t|(\theta_t))\pi(\theta_t|y_{1:t-1})d\theta_t \end{aligned}$$

şeklinde olur.

(İİİ) 'i kanıtlamak için; Bayes kuralından ve Y_t 'nin $Y_{1:t-1}$ 'den bağımsızlığından yola çıkarak, verilen θ_t .

Böylece;

$$\pi(\theta_t|y_{1:t}) = \frac{\pi(\theta_t|(y_{1:t-1})\pi(y_t|\theta_t, y_{1:t-1}))}{\pi(y_t|y_{1:t-1})} = \frac{\pi(\theta_t|(y_{1:t-1})\pi(y_t|\theta_t))}{\pi(y_t|y_{1:t-1})}$$

şeklinde olur.

Önceki önermenin sağladığı bir adım önde tahmin dağılımından, durum için ve gözlem için k -adımlar önde tahmin dağılımları formüllere göre yinelemeli olarak hesaplanabilir bu formüller aşağıda verildiği gibi gösterilebilir.

$$\pi(\theta_{t+k}|y_{1:t}) = \int \pi(\theta_{t+k}|\theta_{t+k-1})\pi(\theta_{t+k-1}|y_{1:t}))d\theta_{t+k-1}$$

ve

$$\pi(y_{t+k}|y_{1:t}) = \int \pi(y_{t+k}|\theta_{t+k})\pi(\theta_{t+k}|y_{1:t}))d\theta_{t+k}.$$

Bu arada, bu özyinelemeler aynı zamanda $\pi(\theta_t|y_{1:t})$ geçmiş gözlemlerde yer alan bilgileri özetlediğini göstermektedir ($y_{1:t}$, herhangi bir $k > 0$ için Y_{t+k} 'yi tahmin etmek için yeterlidir).

3.3.1. Dinamik Doğrusal Modeller için Kalman Filtresi

Kalman filtresi başta havacılık alanında olmak üzere pek çok alanda sıklıkla kullanılan bir yöntemdir. Aslında filtre olarak isimlendirilmiş olsa da mevcut ve önceki durumlara göre bir sonraki durumu tahmin eden bir algoritmadır. Kalman filtresi, Amerikan Apollo uzay programının da temel taşlarından biridir. Genel olarak kalman filtresi, gürültülü veriler üzerinden sistemin gelecek durumu hakkında tahminde bulunur. Genel olarak var olan sonuçlar, ilke olarak filtreleme ve tahmin problemlerini çözer. Fakat, genelde ilgili koşullu dağılımların gerçek hesaplaması hiç de kolay bir iş değildir. DLM'ler, genel özyinelemeli bağıntıların önemli derecede basitleştiklerinde önem taşırlar. Bu durumda, çok değişkenli Gauss dağılımı hakkında standart sonuçları kullanarak, rasgele vektörün $(\theta_0, \theta_1, \dots, \theta_t, Y_1, \dots, Y_t)$ herhangi bir $t \geq 1$ için Gauss dağılımına sahip olduğu kolayca kanıtlanabilir. Marjinal ve koşullu dağılımların da Gaussian olduğu izlenir. Tüm ilgili dağıtımlar Gaussian olduğu için, bunların araçları ve varyansları ile tamamen belirlenir. DLM'ler için filtreleme sorununun çözümü ünlü

Kalman filtresi tarafından verilir. Kalman filtresi, mevcut durumu tahmin etmek için geçmiş ve güncel gözlemleri kullanır. Bir DLM için Kalman Filtresi aşağıdaki şekilde yazılır.

Önerme 3.3.1.1.(Kalman Filtreleme). Bölüm 3'te tanımlanan DLM'yi ele alalım;

$$\theta_{t-1}|y_{1:t-1} \sim N(m_{t-1}, C_{t-1})$$

Daha sonra aşağıdaki ifadeler geçerlidir.

- i. $y_{1:t+1}$ aşağıda yer alan parametrelerle verilen θ_t 'in tek adımlı öngörülü dağılımı Gaussian'dır.

$$\begin{aligned} a_t &= E(\theta_t|y_{1:t-1}) = G_t m_{t-1} \\ R_t &= Var(\theta_t|y_{1:t-1}) = G_t C_{t-1} G_t' + W_t \end{aligned} \quad (3.3.1.a)$$

- ii. $y_{1:t+1}$ aşağıda yer alan parametrelerle verilen Y_t 'in tek adımlı öngörülü dağılımı Gaussian'dır.

$$\begin{aligned} f_t &= E(Y_t|y_{1:t-1}) = F_t a_t, \\ Q_t &= Var(Y_t|y_{1:t-1}) = F_t R_t F_t' + V_t \end{aligned} \quad (3.3.1.b)$$

- iii. $y_{1:t}$ aşağıda yer alan parametrelerle verilen θ_t 'in filtreleme dağılımı Gaussian'dır.

$$\begin{aligned} m_t &= E(\theta_t|y_{1:t}) = a_t + R_t F_t' Q_t^{-1} e_t, \\ C_t &= Var(\theta_t|y_{1:t}) = R_t - R_t F_t' Q_t^{-1} F_t R_t \end{aligned} \quad (3.3.1.c)$$

Burada $e_t = Y_t - f_t$ tahmini hata terimidir.

İspat. Marjinal ve koşullu dağılımların Gauss olduğu yerlerde, (2.2.1) tarafından verilen rastgele vektörü olan $(\theta_0, \theta_1, \dots, \theta_t, Y_1, \dots, Y_t)$ ortak bir dağılıma sahiptir. Çok değişkenli Normal dağılımdaki standart sonuçlardan, $(\theta_0, \theta_1, \dots, \theta_t, Y_1, \dots, Y_t)$ 'in ortak dağılımının

herhangi bir $t \geq 1$ için Gaussian olduğunu izler. Tahmini dağıtımlar ve filtreleme dağılımları Gauss'tur. Bu neden, araçlarını ve varyanslarını hesaplamak için yeterlidir.

Yukarıda yer alan (i) $\theta_t|y_{1:t-1} \sim N(a_t, R_t)$ olsun. Burada yer alan a_t ve R_t 'nin kanıtı için denklem (3.3.) kullanarak aşağıdaki şekilde elde edilebilir.

$$\begin{aligned} a_t &= E(\theta_t|y_{1:t-1}) = E(E(\theta_t|\theta_{t-1}, y_{1:t-1})|y_{1:t-1}) \\ &= (G_t \theta_{t-1}|y_{1:t-1}) = G_t m_{t-1} \end{aligned}$$

ve

$$\begin{aligned} R_t &= Var(\theta_t|y_{1:t-1}) \\ &= E(Var(\theta_t|\theta_{t-1}, y_{1:t-1})|y_{1:t-1}) + Var(E(\theta_t|\theta_{t-1}, y_{1:t-1})|y_{1:t-1}) \\ &= W_t + G_t C_{t-1} G_t' \end{aligned}$$

Yukarıda yer alan (ii) $Y_t|y_{1:t-1} \sim N(f_t, Q_t)$ olsun. Burada yer alan f_t ve Q_t 'nin kanıtı için denklem (3.2.) kullanarak aşağıdaki şekilde elde edilebilir.

$$\begin{aligned} Q_t &= Var(Y_t|y_{1:t-1}) \\ &= E(Var(Y_t|\theta_t, y_{1:t-1})|y_{1:t-1}) + Var(E(Y_t|\theta_t, y_{1:t-1})|y_{1:t-1}) \\ &= V_t + F_t R_t F_t' \end{aligned}$$

Yukarıda yer alan (iii)'yi önerme 3.3.1.(iii)' de yer alan denklemde, mevcut özel duruma uyarlayabiliriz. DLM durumunda tüm dağılımlar Gauss'tur ve problem lineer model için Bayes çıkarım problemi ile aynıdır bunu önceki bir $N(a_t, R_t)$ Gaussian'ı izleyen bir regresyon parametresi θ_t ile aşağıda yer aldığı gibi gösterebiliriz,

$$Y_t = F_t \theta_t + v_t, \quad v_t \sim N(0, V_t)$$

burada V_t bilinendir.

$$\begin{aligned} \theta_t|y_{1:t} &\sim N(m_t, C_t), \\ m_t &= a_t + R_t F_t' Q_t^{-1} (Y_t - F_t a_t), \\ C_t &= R_t - R_t F_t' Q_t^{-1} F_t R_t. \end{aligned}$$

Denklemlerini kullanarak; Kalman filtresi, $\theta_0 \sim N(m_0, C_0)$ 'ten başlayarak $\pi(\theta_1|y_1)$ 'yi hesaplayarak ve yeni veriler kullanıma girdiğinde yinelemeli olarak ilerleyerek yordayıcı ve süzgeç dağılımlarını hesaplamamıza olanak sağlar. m_0 'in ifadesi, sezgisel tahmin düzeltme formuna sahiptir; “filtre ortalaması, yeni gözlemin tahmininden ne kadar farklı olduğuna bağlı olarak, artı ortalama a_t düzeltme tahminine eşittir”. Düzeltme teriminin ağırlığı kazanç matrisi ile aşağıdaki şekilde verilir.

$$K_t = R_t F_t' Q_t^{-1}.$$

Böylece, mevcut veri noktası Y_t 'nin ağırlığı, gözlem varyansı V_t 'ye (Q_t aracılığıyla) ve $R_t = \text{Var}(\theta_t|y_{1:t-1}) = G_t C_{t-1} G_t' + W_t$ 'ye bağlıdır.

Kalman filtresi;

$$\mu_t|y_{1:t-1} \sim N(m_{t-1}, R_t = C_{t-1} + W),$$

$$Y_t|y_{1:t-1} \sim N(f_t = m_{t-1}, Q_t = R_t + V),$$

$$\mu_t|y_{1:t} \sim N(m_t = m_{t-1} + K_t e_t, C_t = K_t V),$$

burada $K_t = R_t/Q_t$ ve $e_t = Y_t - f_t$, şeklindedir. İşlem Y_t 'e ait davranışın, genellikle sinyal-gürültü oranı olarak adlandırılan iki hata varyansı olan $r = W/V$ arasındaki orandan büyük ölçüde etkilendiğinin altını çizmek gerekir. Bunu görmenin en iyi yolu, farklı V ve W değerleri için Y_t 'in bazı yörüngelerini simüle etmektir. Bu, tahmin ve tahmin mekanizmasının yapısına yansır. Burada $m_t = K_t y_t + (1 - K_t)m_{t-1}$, y_t ve m_{t-1} 'nin ağırlıklandırılmış ortalamasıdır. $K_t = R_t/Q_t = (C_{t-1} + W)/(C_{t-1} + W + V)$ ağırlığı, mevcut gözlem olan y_t 'nin aynı zamanda adaptif katsayı olarak adlandırılır ve $0 < K_t < 1$ 'yi karşılar. Verilen herhangi bir C_0 için sinyal-gürültü r küçük ise, K_t küçüktür ve y_t az ağırlık alır. Eğer karşı uçta, $V = 0$ ise $K_t = 1$ ve $m_t = y_t$ 'ye sahibiz demektir, buda en son veri noktası tarafından bir adımlı tahmin yapılabileceği anlamına gelir. W ve V 'nin farklı nispi büyüklüklerinin filtrelenmiş dağılımın ortalamasını nasıl etkilediğine ve tek adımlı tahminlerine dair R programında yapılmış örnek uygulama bölümünde mevcuttur.

Yukarıda yer alan yinelemeli güncelleme formülleri kullanılarak ön dağılım varyanslar C_t 'nin (aynı zamanda R_t ve Q_t 'nin) değerlendirilmesi görüldüğü gibi

mümkün olabilir fakat asimetrik olmayan ve hatta negatif belirli hesaplanmış varyans matrislerine yol açabilen nümerik dengesizlikten muzdariptir. Bu sorunun üstesinden gelmek için alternatif olarak istikrarlı algoritmalar geliştirilmiştir. Buna örnek olarak, İstatistik literatüründe Morf ve Kailath (1975) ve Anderson ve Moore (1979)'un çalışmalarında en çok kullanılan C_t 'nin karekökünün sıralı güncellemesi için formül sağlayan karekök filtre yer almaktadır.

Özellikle gözlemsel gürültünün küçük bir varyansa sahip olduğu durumlarda, karekök filtresinin bile sayısal kararlılık problemlerini beraberinde getirdiğini ve negatif hesaplanmış varyanslara yol açtığını Oshman ve Bar-Itzhack (1986) ve Wang vd. (1992) belirtmektedirler. Aynı zamanda karekök filtresine alternatif olarak, daha sağlam bir algoritma olduklarını varsaydıkları, C_t 'nin tekil değer ayrıştırmasını (TDA) ardışık olarak güncellemeye dayanan bir algoritma önermektedirler. Açıkçası, TDA tabanlı filtre karekök filtre olarak görülebilir. Aslında $A = UD^2U'$, bir varyans matrisinin TDA'sıysa, DU' , A 'nın kareköküdür. Fakat, standart karekök filtreleme algoritmalarına kıyasla TDA tabanlı olan filtreleme tipik olarak daha stabildir.

Kalman filtresi *dml* paketinde *dmlFilter* işleviyle gerçekleştirilir. Argümanlar, sayısal bir vektör, matris veya zaman dizisi biçimindeki veriler ve *mod*, *dml* sınıfı bir nesne veya bir *dml* nesnesine zorlanabilir bir liste şeklindeki verilerdir. Yukarıda belirtilen sayısal stabilite nedenlerden dolayı, hesaplamalar C_t ve R_t varyans matrislerinin TDA'sında gerçekleştirilir. Buna göre, her bir t için bir dikey matris $U_{C,t}$ ve bir $D_{C,t}$ vektörünü ve benzer şekilde R_t için bir vektör $C_t = U_{C,t} \text{diag}(D_{C,t}^2) U_{C,t}'$ olan çıktısını sağlar.

DmlFilter tarafından üretilen çıktı, “*dmlFiltered*” sınıf öznitelikli bir liste, orijinal verilere ve modele ek olarak (y ve *mod* bileşenleri), öngörülü ve filtrelenmiş dağıtımların (a ve m bileşenleri) ve tahmin edici ve süzölmüş dağılımların ($U.R, D.R, U.C$ ve $D.C$ bileşenleri) varyanslarının TDA'sını içerir. Kolaylık sağlamak için, çıktı listesinin bileşen f 'si kullanıcıya bir adımlı öngörme sağlar. Bileşen $U.C$, yukarıdaki $U_{C,t}$ 'nin matrislerinin bir listesidir. $D.C$ ise C_t 'lerin TDA'sında yer alan $D_{C,t}$ 'sinin vektörleri olan satır tarafından depolanan bir matristir. Bu durum benzer şekilde $U.R$ ve $D.R$ içinde geçerlidir. Yardımcı fonksiyon *dmlSvd2var*, TDA'lardan

varyansları yeniden oluşturmak için kullanılabilir. Daha önce anlatıldığı gibi uygulama için R programından *dlm* paketi kullanılır ve bu pakette daha uygun olan *dlmModPoly* kullanılmaktadır. Daha önce anlattığımız gibi, W ve V 'nin göreceli büyüklüğü, kazanç matrisine giren önemli bir faktördür. Buda, duruma göre durum öncesi güncellemenin beklenmedik gözlemlere ne kadar duyarlı olduğunu belirler. Yerel seviye modelinde sinyal gürültü oranı W / V 'nin rolünü göstermek için yapılacak olan uygulamanın gerçek seviyesini tahmin etmek amacıyla, belirgin bir şekilde farklı sinyal-gürültü oranına sahip iki model kullanırız. İki model için filtrelenmiş değerler daha sonra karşılaştırılabilir.

3.3.2. Eksik Gözlemlerle Filtreleme

Uygulanan veri analizinde, bir veya daha fazla eksik gözlem içeren bir zaman dizisi ile uğraşmak sürekli bir durumdur. Çok değişkenli zaman serilerinde, eksik gözlemler iki farklı tipte olabilir, tamamen eksik ve kısmen eksik gözlemler. İlk tip, gözlem vektörünün bir zaman t 'de mevcut olmaması durumunda ortaya çıkan tiptir. İkinci durumda ise, gözlem vektörünün sadece bazı bileşenleri mevcut değildir. Bu durum, örneğin, birkaç ülkede bir dizi hisse senedi endeksinin kapanış fiyatlarının bir günlük zaman serisini dikkate alırken ortaya çıkabilir. Eğer gün (t), A ülkesinde tatil ama B ülkesinde değil ise, o gün için ülke A endeksi için kapanış fiyatı tanımlanmamıştır yani eksiktir fakat B ülkesi endeksinin kapanış fiyatı ise normal olarak kaydedilir. Açıkça, tek değişkenli bir zaman dizisi için bir gözlem eksiktir. Bu durumu durum uzay modellerinin yapısı, eksik gözlemlerin filtreleme özümlemesinde kolaylıkla yerleştirilebileceği şekildedir. Bu kısımda tamamen eksik gözlemler durumu ele alınmaktadır. Bu doğrultuda R düzenin takiben, özel bir değeri NA olan bir eksik gözlemi ele almaktayız. Eğer t zamanında gözlem eksik ise, $y_t = NA$ ve y_t herhangi bir bilgi taşımamaktadır, böylece denklem aşağıda yer aldığı gibidir.

$$\pi(\theta_t | y_{1:t}) = \pi(\theta_t | y_{1:t-1}) \quad (3.3.2.1)$$

Bu durumda, t zamanında filtreleme dağılımı, $t - 1$ zamanında sadece bir adım önde tahminsel dağılımdır filtreleme tekrarında bir (3.3.2.1) ile (3.3.c) değiştirilmelidir. Bu durumda bir DLM için, $\theta_t | y_{1:t-1} \sim N(a_t, R_t)$, yapılması gereken $m_t = a_t$ ve $C_t = R_t$ 'yi ayarlamaktır. $t + 1$ 'den itibaren standart süzme özyinelemesi her zamanki gibi

devam eder ve elde edilen y_{t+1} eksiksizdir. Resmi olarak, bir DLM'de $y_t = NA$ 'ya sahip olmanın, $F_t = 0$ veya V_t ayarı ile aynı olduğu göz ardı edilmemelidir. İlk durumda y_t herhangi bir şekilde θ_t ile bağlantılı değildir, ikincisinde gözlem ise θ_t hakkında anlamlı bilgiler sağlama konusunda tamamen güvenilmez olduğu için çok gürültülüdür. Her iki şekilde de $K_t = 0$ kazanç matrisi elde edilir ve sonuç olarak $m_t = a_t$ ve $C_t = R_t$ olur.

Şimdi m boyutlu gözlem vektörleri olan bir alan uzay modeli düşünelim, $m > 1$. y_t bileşenlerinin hepsi değil, bazılarının eksik olduğunu varsayalım. Bu durumda y_t vektörü, θ_t hakkında bazı bilgiler verir ancak tüm bu bilgiler eksiksiz bileşenlerde bulunur. $\tilde{y}_t$, y_t 'nin sadece eksik olmayan bileşenleri içeren bir vektör olsun. Daha sonra 4.3.1.'de yer alan filtreleme tekrarında, $\pi(y_t|\theta_t)$, $\pi(\tilde{y}_t|\theta_t)$ ile ve $\pi(y_t|y_{1:t-1})$, $\pi(\tilde{y}_t|y_{1:t-1})$ ile değiştirilmelidir. $\tilde{m}_t$ tarafından $\tilde{y}_t$ 'nin boyutunu belirten ve $\tilde{m}_t$ 'nin m matris M_t ile m yi kimlik matrisinden çıkarılmasıyla elde edilen $\tilde{m}_t$ 'yi y_t eksik bileşenlerine karşılık gelen satırlar olarak dikkate alınmalıdır, böylece $\tilde{y}_t = M_t y_t$ denklemi elde edilir. Aslında y_t yerine $\tilde{y}_t$ 'yi gözlemledik, önceki $N(a_t, R_t)$ ve son $N(m_t, C_t)$ 'nin güncellenmesinde dikkate alınması gereken doğru gözlem denklemi aşağıda verildiği gibidir.

$$\tilde{y}_t = \tilde{F}_t + \tilde{v}_t, \quad \tilde{v}_t \sim N(0, \tilde{V}_t)$$

burada, $\tilde{F}_t = M_t + F_t$ ve $\tilde{V}_t = M_t + V_t + M_t'$ şeklindedir. *dmlFilter* işlevi, NA 'ları içeren verileri doğru filtreleme dağıtımlarının anlarını hesaplayarak kabul eder.

3.3.3. Düzleştirici(Smoother)

Durum uzay modellerinin cazip özelliklerinden biri de yeni veriler ortaya çıktıkça, tahmin ve tahminlerin sıralı olarak uygulanabilmesidir. Düzeltmenin amacı, örneğin gözlemlerin altında yatan sosyo-ekonomik nedeni incelemek, sistemin yapısını geriye dönük olarak mevcut gözlemler temelinde yeniden yapılandırmaktır. Bununla birlikte zaman serileri analizinde, genellikle belli bir süre için Y_t üzerinde gözlemler vardır($t = 1, \dots, T$) ve sistemin davranışını geriye dönük olarak yeniden yapılandırmak, gözlemlerin altında yatan sosyo-ekonomik yapıyı ya da fiziksel durumu incelemek istemektedir. Bu durumda, verilmiş $y_{1:T}$ 'in koşullu dağılımını hesaplamak

için, herhangi bir $t < T$ için, filtreleme dağılımından $\pi(\theta_T | y_{1:T})$ başlayarak ve tüm devletlerin geçmişini geriye doğru tahmin etmek için geriye dönük özyinelemeli bir algoritma kullanılabilir. Aşağıdaki önerme, genel durum uzay modeli için düzleştirme tekrarlarının bir sunumunu içerir.

Önerme 3.3.3.1.(Düzleştirme tekrarları). Aşağıdaki ifadeler, (A.1) - (A.2) ile tanımlanan genel durum uzay modeli için geçerlidir.

- i. Koşullu $y_{1:T}$, durum sırası $(\theta_0 \dots \theta_T)$ aşağıda verilen model tarafından geri geçiş olasılıkları vardır;

$$\pi(\theta_t | \theta_{t+1}, y_{1:T}) = \frac{\pi(\theta_{t+1} | \theta_t) \pi(\theta_t | y_{1:t})}{\pi(\theta_{t+1} | y_{1:t})}$$

- ii. Verilen $y_{1:T}$ θ_t 'in yumuşatma dağılımlarını, $\pi(\theta_T | y_{1:T})$ 'den başlayarak aşağıda verildiği gibi t 'deki geriye doğru özyinelemeye göre hesaplanabilir.

$$\pi(\theta_t | y_{1:T}) = \int \frac{\pi(\theta_{t+1} | \theta_t)}{\pi(\theta_{t+1} | y_{1:t})} \pi(\theta_{t+1} | y_{1:T}) d\theta_{t+1}.$$

İspat. (İ) 'i kanıtlamak için; θ_t ve $y_{1:T}$ koşullu olarak bağımsız θ_{t+1} verilir. Ayrıca, θ_{t+1} ve $y_{1:T}$ koşullu olarak bağımsız θ_t verilir. (Bunu göstermek için şekil 3.1 kullanılabilir.) Bayes formülünü kullanarak;

$$\begin{aligned} \pi(\theta_t | \theta_{t+1}, y_{1:T}) &= \pi(\theta_t | \theta_{t+1}, y_{1:t}) \\ &= \frac{\pi(\theta_t | y_{1:t}) \pi(\theta_{t+1} | \theta_t, y_{1:t})}{\pi(\theta_{t+1} | y_{1:t})} \\ &= \frac{\pi(\theta_t | y_{1:t}) \pi(\theta_{t+1} | \theta_t)}{\pi(\theta_{t+1} | y_{1:t})} \end{aligned}$$

modelleri elde edilir.

(İİ) 'i kanıtlamak için; $\pi(\theta_t | \theta_{t+1}, y_{1:T})$, θ_{t+1} ile ilgili olarak aşağıdaki şekilde marjinalleştirilmelidir;

$$\pi(\theta_t | y_{1:T}) = \int \pi(\theta_t, \theta_{t+1} | y_{1:T}) d\theta_{t+1}$$

$$\begin{aligned}
&= \int \pi(\theta_{t+1}|y_{1:T})\pi(\theta_t|\theta_{t+1}, y_{1:T})d\theta_{t+1} \\
&= \int \pi(\theta_{t+1}|y_{1:T}) \frac{\pi(\theta_{t+1}|\theta_t)\pi(\theta_t|y_{1:T})}{\pi(\theta_{t+1}|y_{1:T})} d\theta_{t+1} \\
&= \pi(\theta_t|y_{1:t}) \int \pi(\theta_{t+1}|\theta_t) \frac{\pi(\theta_{t+1}|y_{1:T})}{\pi(\theta_{t+1}|y_{1:t})} d\theta_{t+1}
\end{aligned}$$

şeklinde ifade edilebilir.

3.3.4. Dinamik Doğrusal Modeller için Kalman Düzleştiricisi (Kalman Smoother)

Dinamik lineer modelleri incelerken, ön dağılım parametrelerin alternatif “özyinelemeli” ifadesini sağlamak yararlıdır. Kalman düzleştiricisi kullanılırken aşağıda yer alan denklemlerden yararlanılır.

$$C_n = (X'V^{-1}X + C_0^{-1})^{-1} = C_0 - C_0X'(XC_0X' + V)^{-1}XC_0 \quad (3.3.4.1)$$

yukarıdaki denklemi kullanarak, ön dağılım beklenti aşağıda yer aldığı gibi m_n olarak ifade edilebilir.

$$m_n = m_0 + C_0X'(XC_0X' + V)^{-1}(y - Xm_0) \quad (3.3.4.2)$$

Yukarıda yer alan her iki denklem, aşağıda kullanacağımız modeller için yazıldı.

4. bölümde tanımlanan bir DLM modeli için; $\theta_{t+1}|y_{1:T} \sim N(s_{t+1}, S_{t+1})$ daha sonra $\theta_t|y_{1:T} \sim N(s_t, S_t)$ burada,

$$s_t = m_t + C_t G'_{t+1} R_{t+1}^{-1} (s_t - a_{t+1}),$$

$$S_t = C_t - C_t G'_{t+1} R_{t+1}^{-1} (R_{t+1} - S_{t+1}) R_{t+1}^{-1} G_{t+1} C_t$$

modelleri kullanılabilir. Çok değişkenli Gauss dağılımının özelliklerinden, θ_t verilen $y_{1:t}$ 'in koşullu dağılımı Gauss'tur. Bundan durum, ortalama ve varyansın hesaplanmasını mümkün kılar. Öte yandan aşağıda yer alan modelleri elde ederiz

$$s_t = E(\theta_t|y_{1:T}) = E(E(\theta_t|\theta_{t-1}, y_{1:T})|y_{1:T})$$

$$S_t = Var(\theta_t|y_{1:T}) = Var(E(\theta_t|\theta_{t+1}, y_{1:T}) + E(Var(\theta_t|\theta_{t+1}, y_{1:T})|y_{1:T})).$$

θ_t ve $Y_{t+1:T}$ şartlı olarak, bağımsız θ_{t+1} verilir. Buradan $\pi(\theta_t|\theta_{t+1}, y_{1:T}) = \pi(\theta_t|\theta_{t+1}, y_{1:t})$ sonucu elde edilir. Bu dağılımı hesaplamak için Bayes formülünü kullanabiliriz. $\pi(\theta_{t+1}|\theta_t, y_{1:t}) = \pi(\theta_{t+1}|\theta_t)$ olasılığının durum denklemi ile ifade edildiğini, ve aşağıda yer alan denklem olduğu göz ardı edilmemelidir

$$\theta_{t+1}|\theta_t \sim N(G_{t+1}\theta_t, W_{t+1}).$$

Önceki $\pi(\theta_t|y_{1:t})$, $N(m_t, C_t)$ şeklinde tanımlanmıştı. Yukarıda yer alan denklemleri (3.3.4.1.) ve (3.3.4.2.) kullanarak, aşağıda yer alan denklemler elde edilir;

$$\begin{aligned} E(\theta_t|\theta_{t+1}, y_{1:t}) &= m_t + C_t + G'_{t+1}(G_{t+1}C_tG'_{t+1} + W_{t+1})^{-1} + (\theta_{t+1} - G_{t+1}m_t) \\ &= m_t + C_tG'_{t+1}R_{t+1}^{-1}(\theta_{t+1} - a_{t+1}) \end{aligned}$$

bunu takip eden

$$\begin{aligned} s_t &= E(E(\theta_t|\theta_{t+1}, y_{1:t})|y_{1:T}) = m_t + C_tG'_{t+1}R_{t+1}^{-1}(s_{t+1} - a_{t+1}) \\ S_t &= Var(E(E(\theta_t|\theta_{t+1}, y_{1:t})|y_{1:T}) + E(Var(\theta_t|\theta_{t+1}, y_{1:t})|y_{1:T})) \\ &= C_t - C_tG'_{t+1}R_{t+1}^{-1}G_{t+1}C_t + C_tG'_{t+1}R_{t+1}^{-1}S_{t+1}R_{t+1}^{-1}G_{t+1}C_t \\ &= C_t - C_tG'_{t+1}R_{t+1}^{-1}(R_{t+1} - S_{t+1})R_{t+1}^{-1}G_{t+1}C_t, \end{aligned}$$

burada varsayımlar tarafından;

$$E(\theta_t|\theta_{t+1}, y_{1:t}) = s_{t+1} \text{ ve } Var(\theta_t|y_{1:t}) = S_{t+1}$$

şeklinde tanımlanır.

Kalman düzleştiricisi, $\theta_T|y_{1:T} \sim N(s_T, S_T = C_T)$ durumunda $t = T - 1$ 'den başlayarak $\theta_t|y_{1:T}$ dağılımlarını, daha sonra $\theta_t|y_{1:T}$ ($t = T - 2, T - 3, v. b$) dağılımlarının geriye doğru işlemini hesaplamayı sağlar. Düzleştirici yinelemesinin, sadece filtreleme ve Kalman filtresi kullanılarak elde edilen bir adım önde tahmin edici anlar aracılığıyla verilere bağlı olduğuna dikkat edilmelidir. Bu nedenle, bir zaman dizisinde eksik gözlemler varsa, bu, filtreleme tekrarlama gerçekleştirilirken hesaba katılmalıdır ancak düzeltme yinelemesinde ek bir ayarlama gerekmemektedir. *dlmSmooth* işlevi, *dlmFiltered* sınıfının bir nesnesinden başlayarak R'deki genellikle *dlmFilter* tarafından üretilen çıktı hesaplamalarını gerçekleştirir. Alternatif olarak,

kullanıcı verileri ve modeli sağlayabilir, bu durumda *dmlFilter* içten olarak adlandırılır. *dmlSmooth* kendi TDA'sı açısından verilen s ve $U.S, D.S$ ve onların varyansları bileşenleri ile bir liste döndürür.

DLM'ler sistemin durumu hakkında öğrenme süreci dinamik bir ortamda, yani durumun bu konuda bilgi sahibi olduğu gibi değiştiği bir ortamda gerçekleşir. Bu nedenle, zaman $t - 1$ 'den zaman T 'ye filtreleme varyansının güncellenmesinde çelişkili süreç devam vardır; bir yandan gözlem y_t , θ_{t-1} hakkında yeni bilgiler oluşturmaktadır ancak bu arada sistemin durumu θ_t tarafından taşınan ek belirsizlik ile θ_t 'ye dönüşür. Bu ek belirsizlik, varyans $W_t = W$ ile temsil edilir. C_0 büyükse, durum hakkında önceden tahminine fazla güven duyulmaz, o zaman ilk gözlemler çok bilgilendirici ve C_t üzerindeki etkileri filtreleme varyansının genel bir azalmasına neden olan durumun dinamiklerinden çok daha önemlidir. Bununla birlikte, daha fazla veri toplandıkça, bir ek gözlemin sistemin durumu hakkındaki bilgiler üzerindeki etkisi azalır ve bir noktada, ek varyans W ile temsil edilen bilgi kaybıyla tam olarak dengelenmektedir. O andan itibaren C_t esasen sabit kalmaktadır.

3.3.5. DLM Modelleri İçin Tahmin

$y_{1:t}$ ile, Y_{t+k} gözlemlerin veya θ_{t+k} durum vektörlerinin gelecekteki değerlerini tahmini önemli bir konudur ve istatistikçiler tarafından büyük ilgi görmektedir. Durum uzay modelleri, hesaplamaların özyinelemeli (recursive) formu, tek adımlı tahminleri hesaplamayı ve bunları yeni veriler ortaya çıktıkça ardışık olarak güncellemeyi doğal kılar. Bu, verilerin günlük olarak tahmini hisse senedi fiyatlarında olduğu gibi ya da hareketli bir hedefi takip etmede sıralı olarak ulaştığı uygulamalı sorunlara olan ilgiyi açıkça göstermektedir; ancak modelin performansını kontrol etmek için bir araç olarak, bir adımlı tahminler genellikle “örnek olarak” hesaplanır.

Bir DLM için, durumlar ve gözlemler için tek adımlı ön tahmin dağılımları, filtreleme bölümünde gösterildiği gibi Kalman filtresinin bir yan ürünü olarak elde edilir. R'de tek adımlı tahminler $f_t = E(Y_t | y_{1:t-1})$ *dmlFilter* işlevinin çıkışında sağlanır. Gözlemin bir adım ileriye yönelik tahminlerinin her bir t değeri için f_t , m_{t-1} ortalama filtrelemenin doğrusal bir fonksiyonu olduğu için, kazanç matrisinin büyüklüğü, f_t nin beklenmedik bir y_{t-1} gözlemin ne kadar duyarlı olduğunu belirlemek için aynı rolü

oyunar. Rastgele yürüyüş artı gürültü modeli söz konusu olduğunda bu özellikle belirgindir, çünkü bu durumda, $f_t = m_{t-1}$ dir.

Pek çok uygulamada, gelecekte biraz daha fazla pay görmeye ilgi duyuyor ve k adımları için serinin davranışının olası senaryolarını sağlıyoruz. Burada, t zamana kadar olan verilere dayanarak, durumların koşullu dağılımlarının ve gözlemlerinin gelecekteki bir $t + k$ zamandaki ortalamaları ve varyansları için özyinelemeli (recursive) formülleri sunulmaktadır. Modelin Markovian doğası göz önüne alındığında, zamana göre filtreleme dağılımı, modelin gelecekteki gelişimi için bir başlangıç dağılımı gibi davranır. Daha detaylı anlatmak gerekirse, şimdiki ve gelecekteki $(\theta_{t+k})_{k \geq 0}$ durumların ve gelecekteki $(Y_{t+k})_{k \geq 1}$ gözlemlerin ortak dağılımı, koşullu dağılımlara sahip bir $\pi((\theta_{t+k} | (\theta_{t+k-1})$ ve $\pi(y_{t+k} | (\theta_{t+k})$ durum uzay modelidir ve başlangıç dağılımı $\pi((\theta_t | (y_{1:t}))$ dir. Verilerin sağladığı gelecekle ilgili bilgiler, bu dağıtımda yer almaktadır. Özellikle bir DLM için, veriler yalnızca m_t elde etmek için ortalama $\pi((\theta_t | (y_{1:t}))$ kullanıldığından, m_t nin öngörme amaçları için yeterli olan verilerin bir özeti sağladığı anlaşılmaktadır. Değişkenler arasında bağımlılık yapısını temsil eden Şekil 2.2.1.'e bakarak bu konuda daha fazla sezgiye sahip olunabilir. $Y_{1:t}$ 'den y_{t+k} ya kadar olan yolun şekil 3.3.5.1'deki gibi olduğunu söyleyebiliriz bu da $Y_{1:t}$ verisinin θ_t hakkında bilgi sağladığını gösterir, bu da θ_{t+k} 'ye kadar gelecekteki durum evrimi hakkında ve sonuç olarak Y_{t+k} ya kadar bilgi verir. Elbette, k büyüdükçe, sisteme daha fazla belirsizlik girer ve tahminler daha az ve daha az kesin olacaktır.

Şekil.3.3.5.1. $Y_{1:t}$ 'den y_{t+k} 'ya Bilgi Akışı

$$\begin{array}{ccccccc} \theta_t & \longrightarrow & \theta_{t+1} & \longrightarrow & \dots & \longrightarrow & \theta_{t+k} \\ \downarrow & & & & & & \downarrow \\ Y_{1:t} & & & & & & Y_{t+k} \end{array}$$

Önerme 3.3.5.1., durumlar için tahmin dağılımlarını, genel durum uzay modeli için gözlemleri hesaplamak için özyinelemeli (recursive) formüller sağlar.

Önerme 3.3.5.1.(Tahmin yineleme). Herhangi bir $k > 0$ ve (A.1) - (A.2) ile tanımlanan genel durum uzay modeli için aşağıdaki ifadeler sağlanır.

i) Durumun k -adımları-öncesi tahmini dağılımı:

$$\pi(\theta_{t+k}|y_{1:t}) = \int \pi(\theta_{t+k}|\theta_{t+k-1})\pi(\theta_{t+k-1}|y_{1:t})d\theta_{t+k-1}.$$

(ii) Gözlemin k-adımları-öncesinde tahmini dağılımı:

$$\pi(y_{t+k}|y_{1:t}) = \int \pi(y_{t+k}|\theta_{t+k})\pi(\theta_{t+k}|y_{1:t})d\theta_{t+k}.$$

İspat. Modelin koşullu bağımsızlık özelliklerini kullanarak;

i. Modelin koşullu bağımsızlık özelliklerini kullanarak,

$$\begin{aligned}\pi(\theta_{t+k}|y_{1:t}) &= \int \pi(\theta_{t+k}, \theta_{t+k-1} | y_{1:t})d\theta_{t+k-1} \\ &= \int \pi(\theta_{t+k}, \theta_{t+k-1} | y_{1:t})\pi(\theta_{t+k-1}|y_{1:t})d\theta_{t+k-1} \\ &= \int \pi(\theta_{t+k}, \theta_{t+k-1})\pi(\theta_{t+k-1}|y_{1:t})d\theta_{t+k-1}\end{aligned}$$

elde edilir.

ii. Burada ispat yapılırken yine modellerin koşullu bağımsızlık özelliklerine dayanmaktadır. Bu durumda,

$$\begin{aligned}\pi(y_{t+k}|y_{1:t}) &= \int \pi(y_{t+k}, \theta_{t+k} | y_{1:t})d\theta_{t+k} \\ &= \int \pi(y_{t+k}, \theta_{t+k} | y_{1:t})\pi(\theta_{t+k}|y_{1:t})d\theta_{t+k} \\ &= \int \pi(y_{t+k}, \theta_{t+k})\pi(\theta_{t+k}|y_{1:t})d\theta_{t+k}\end{aligned}$$

elde edilir.

DLM'ler için önerme 3.3.5.1. daha spesifik bir şekil alır, çünkü tüm integraller açık bir şekilde hesaplanabilir. Ancak, tüm tahmin dağılımları Gaussian olduğu için filtreleme ve düzeltme için olduğu gibi ortalamalarını ve varyanslarını hesaplamak yeterlidir. Önerme 3.3.5.2. bunları hesaplamak için özyinelemeli (recursive) formüller içerir. İlk önce bazı denklemleri tanıtmamız gerekiyor. Bu durumda $k \geq 1$ için,

$$a_t(k) = E(\theta_{t+k}|y_{1:t}), \quad (3.3.5.a)$$

$$R_t(k) = Var(\theta_{t+k}|y_{1:t}), \quad (3.3.5.b)$$

$$f_t(k) = E(Y_{t+k}|y_{1:t}), \quad (3.3.5.c)$$

$$Q_t(k) = Var(\theta_{t+k}|y_{1:t}), \quad (3.3.5.d)$$

olarak tanımlanır.

Önerme 3.3.5.2. Bölüm 3'te tanımlanan DLM için, $a_t(0) = m_t$ ve $R_t(0) = C_t$ olsun.

Bu takdirde $k \geq 1$ için aşağıdaki durumlar sağlanır.

i. $y_{1:t}$ ile verilen θ_{t+k} nin dağılımı

$$a_t(k) = G_{t+k}a_{t,k-1},$$

$$R_t(k) = G_{t+k}R_{t,k-1}G'_{t+k} + W_{t+k};$$

ile Gaussian'dır.

ii. $y_{1:t}$ ile verilen Y_{t+k} nin dağılımı

$$f_t(k) = F_{t+k}a_t(k),$$

$$Q_t(k) = F_{t+k}R_t(k)F'_{t+k} + V_{t+k}$$

ile Gaussian'dır.

İspat. Daha önce de belirttiğimiz gibi, tüm koşullu dağılımlar Gaussian'dır. Bu nedenle, sadece formülleri ve varyansları veren formülleri kanıtlamamız gerekmektedir.

3.3.1.1.'e göre sonuç $k = 1$ için sağlanır. $k > 1$ için,

$$a_t(k) = E(\theta_{t+k}|y_{1:t}) = E(E(\theta_{t+k}|y_{1:t}, \theta_{t+k-1})|y_{1:t}),$$

$$= E(G_{t+k}\theta_{t+k-1}|y_{1:t}) = G_{t+k}a_{t,k-1},$$

$$R_t(k) = Var(\theta_{t+k}|y_{1:t}) = Var(E(\theta_{t+k}|y_{1:t}, \theta_{t+k-1})|y_{1:t})$$

$$+ E(Var(\theta_{t+k-1}|y_{1:t}, \theta_{t+k-1})|y_{1:t}) = G_{t+k}a_{t,k-1}$$

$$= G_{t+k}R_{t,k-1}G'_{t+k} + W_{t+k}$$

$$\begin{aligned}
f_t(k) &= E(Y_{t+k}|y_{1:t}) = E(E(Y_{t+k}|y_{1:t}, \theta_{t+k})|y_{1:t}) \\
&= E(F_{t+k}\theta_{t+k})|y_{1:t} = F_{t+k}a_t(k), \\
Q_t(k) &= Var(\theta_{t+k}|y_{1:t}) = Var(E(Y_{t+k}|y_{1:t}, \theta_{t+k})|y_{1:t}) \\
&\quad + E(Var(Y_{t+k}|y_{1:t}, \theta_{t+k})|y_{1:t}) \\
&= F_{t+k}R_t(k)F'_{t+k} + V_{t+k},
\end{aligned}$$

şeklinde yazılabilir.

Verilerin, son gözlemin yapıldığı noktada sadece ortalama filtreleme dağılımı üzerinden tahmini dağılımlara girdiğine dikkat edilmeli. *DlmForecast* işlevi, gözlemlerin ve durumların tahmini dağılımlarının ortalamalarını ve varyanslarını hesaplar. İsteğe bağlı olarak, gelecekteki durumların ve gözlemlerin bir örneğini çizmek için kullanılabilir. *DlmForecast*'in ana argümanı, *dlmFiltered* sınıfının bir nesnesidir. Alternatif olarak, bir veri *dlm* sınıfı (ya da uygun adlandırılmış bileşenlere sahip bir liste) olabilir, burada *m0* ve *C0* bileşenleri, veri göz önüne alındığında gözlem periyodunun sonunda durum vektörünün ortalaması ve varyansı olarak yorumlanır yani, onlar son (en son) filtreleme dağılımının ortalaması ve varyansıdır.

DÖRDÜNCÜ BÖLÜM

4. DİNAMİK DOĞRUSAL MODELLER İLE TÜRKİYE'YE AİT GAYRİ SAFİ YURTTİÇİ HÂSILA TAHMİNİ

4.1. Ekonomik Büyüme

Ekonomik büyüme, globalleşen dünyada ülkelerin rekabet hızına ayak uydurabilmeleri için önemli bir kavramdır. Ekonomik göstergelerden biri olan GSYİH gelişmemiş ülkelerde düşüktür. Bir ülkenin ekonomik gücünü gösteren 9 temel faktör vardır. Bu faktörler;

- Gayri Safi Yurtiçi Hâsıla(GSYİH)
- Kişi başına düşen Gayri Safi Milli Hâsıla(GSMH)
- Gelir Dağılımı(Adil gelir dağılımı)
- Eğitim
- Sağlık
- Nüfus
- Tasarruf ve Yatırım Düşüklüğü
- Sektörel Yapı
- Kentleşme Oranı ve Düalist Yapı

şeklinde sıralanabilir.

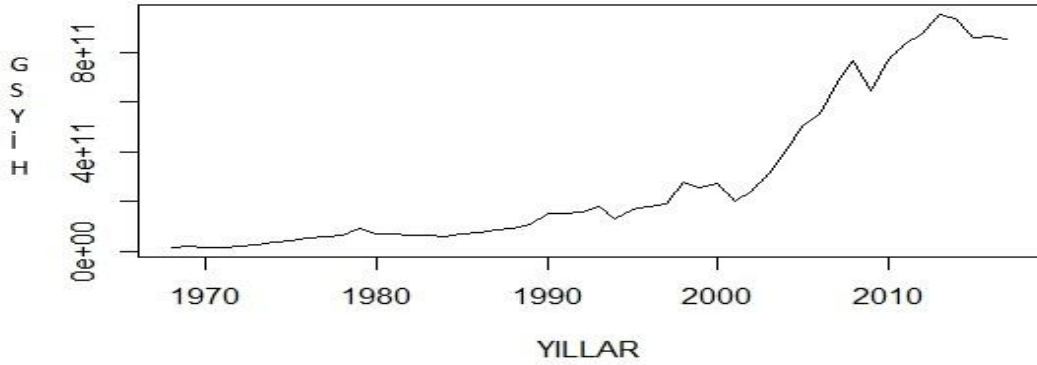
4.2. Türkiye Ait GSYİH'nin DLM ile Tahmini

Gayrisafi yurt içi hâsıla (GSYİH), bir ülkenin ekonomik büyüklüğünün birkaç ölçütünden biridir. GSYİH, GSMH'den farklı olarak, bir ülke sınırları içerisinde belli bir zaman içinde, üretilen tüm nihai mal ve hizmetlerin para birimi cinsinden değeridir. Bir ülkenin yüksek bir GSYİH'ye sahip olması, ülkenin rekabet dünyasında yer almasını ve ekonomik olarak güçlü olmasını sağlar. Düşük bir GSYİH'ye sahip olan ülkeler genellikle gelişmemiş ve dışa bağımlı olan ülkelerdir.

Bu bölümde Türkiye'nin GSYİH gücüne dair bir analiz yapılmaktadır. Bu amaç doğrultusunda, Türkiye'ye ait 1968-2017 yılları arasında yer alan GSYİH yıllık verileri Dünya Bankası(Worldbank) resmi web adresinden alındı. Veriler milyon dolar bazında verilmektedir. Yıllık veriler R programına aktarıldıktan sonra, DLM ile tahmin yapılması için veriler DLM olarak tanıtıldıktan sonra analizler yapıldı. Yapılan analiz

ve sonuçları aşağıda başlıklar halinde verilmiştir. Uygulamada kullanılan kodlar ekler bölümünde verilmiştir.

Grafik 4.2.1. 1968-2017 Türkiye GSYİH Verileri



Grafik 4.2.1. 1968-2017 yılları arasında Türkiye'ye ait GSYİH verilerinin grafiksel gösterimini sunmaktadır.

4.2.1. GSYİH için n. Düzeyde Bir DLM Kurulumu

Bu kısımda ileriki başlıklarda uygulayacağımız, Filtreleme, Düzleştirme ve Tahmin için gerekli özellikler yer almaktadır. Sabit bir DLM'nin tamamen açıklanabilmesi için, F , V , G , W , C_0 matrisleri ve vektör m_0 değerlerinin bilinmesi gerekmektedir. R paketinde bu değerler FF , V , GG , W , $C0$, ve $m0$ olarak elde edilir. Burada $m0$ bize örnek öncesi durum vektörünün beklenen değerini, $C0$ ise örnek öncesi durum vektörünün varyansını verir. Bu kısımda gözlem gürültü varyansı olan V , sistem gürültü varyansı W ve U.C ile birlikte tahmin hata varyansların TDA'sını veren D.C değerleri elde edilir. Ayrıca, TDA'lardan yeniden uygun varyans elde etmemizi sağlayan $d1mSvd2var$ bu bölümde elde edilir. Aşağıda bu bölümde elde edilen çıktılar yer almaktadır.

Tablo 4.2.1.1. DLM Çıktı Sonuçları

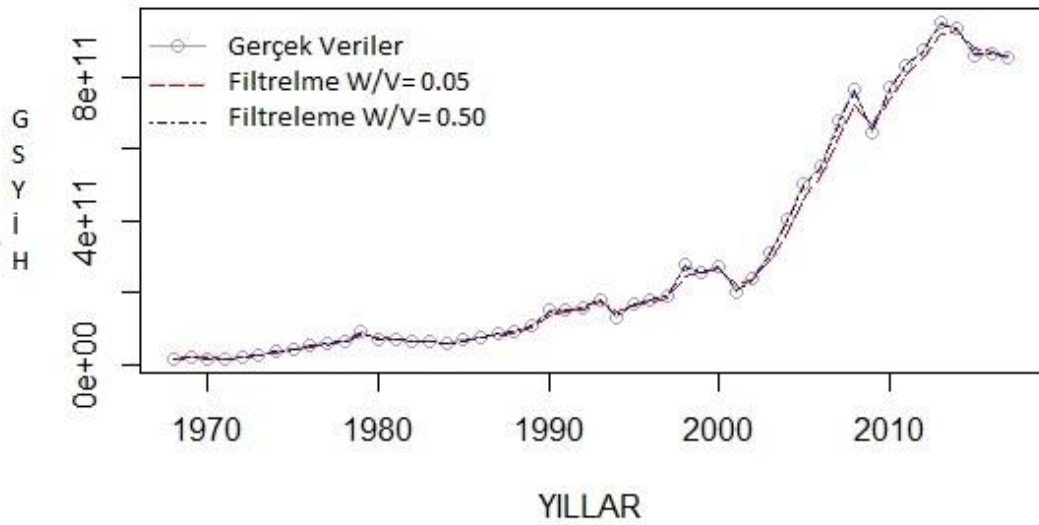
$m0$	$C0$	FF	V	GG	W
0.000000e+	1.000000e+0	1.000000e+	2.741148e-	1.000000e+	2.409779e+

Tablo 4.3.1.'de görüldüğü üzere gözlem gürültü varyansı $2.741148e-0$ ve sistem gürültü varyansı ise $2.409779e+2$ 'dir.

4.2.2. GSYİH İçin Kalman Filtreleme

Yerel seviye modelinde sinyal / gürültü oranı W / V 'nin rolünü göstermek için, Türkiye GSYİH 'sinin gerçek seviyesini tahmin etmek için, belirgin bir şekilde farklı sinyal-gürültü oranına sahip iki model kullanarak gerçeklik durumu elde edilmektedir. Filtreleme aşamasında olan gerçek verilere en yakın filtremeyi yakalayabilmektedir. Gerçek verilerin uygun şekilde filtrenmesi tahmin aşamasında sağlıklı sonuçlar elde etmemizi sağlar. Aşağıda yer alan grafik 4.2.2.1. GSYİH için iki modele ait filtreleme sonucunu vermektedir.

Grafik 4.2.2.1. 1968-2017 Türkiye GSYİH Verileri İçin Kalman Filtreleme

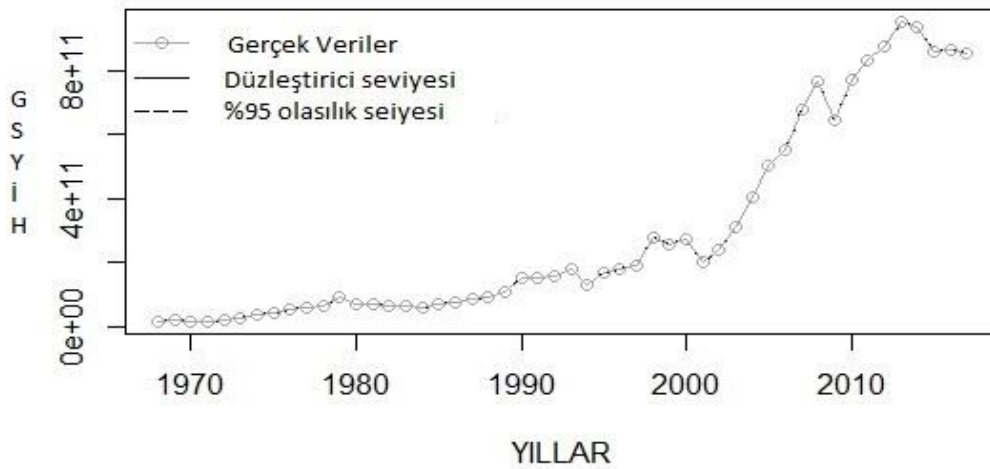


Grafik 4.2.2.1. incelendiğinde, iki modelden kaynaklanan filtrelenmiş seviyeleri göstermektedir. Model 1'den daha büyük bir sinyal-gürültü oranına sahip olan model 2 için filtrelenmiş değerlerin, gerçek verileri daha yakından takip etme eğiliminde olduğu görülmektedir. Gerçek verilerin filtrelenmiş model tarafından yakından takip ediliyor olması yapılacak tahminin sağlıklı sonuçlar vermesini sağlar.

4.2.3. GSYİH için Kalman Düzleştirici

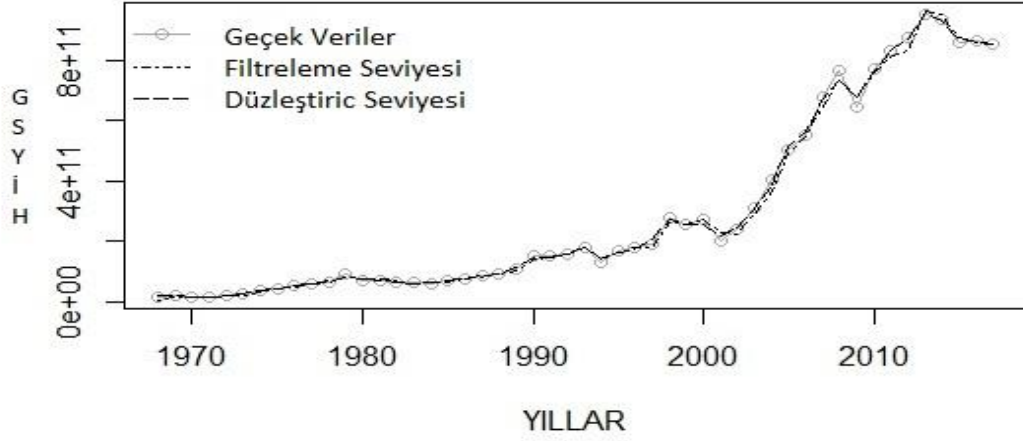
Düzleştirmenin amacı, gözlemlerin altında yatan sosyo-ekonomik nedeni incelemek, sistemin yapısını geriye dönük olarak mevcut gözlemler temelinde yeniden yapılandırmaktır. Bu kısımda, fonksiyon varyans/kovaryans matrisleri ile birlikte, durum vektörlerinin düzeltilmiş değerlerini hesaplamak için Kalman'ı daha esnek bir şekilde uygular. GSYİH için %95 seviyesinde bir düzleştirici uygulanarak analizler yapıldı. Kalman düzleştiricisinde yer alan, D.S, U.S ile birlikte düzeltme hata varyansların TDA'sını verir. Bu kısımda S, durum vektörlerinin düzgünleştirilmiş değerlerin zaman serilerini yansıtır ve ilk gözlemden bir önceki zaman dilimi ile başlar. Düzleştirici seviyesinin gerçek verileri doğru şekilde takip ediyor olması sonuçların sağlıklı olmasını sağlar. Aşağıda yer alan grafik 4.2.3.1.Kalman düzleştiricisini gösterir.

Grafik 4.2.3.1. 1968-2017 Türkiye GSYİH Verileri İçin Kalman Düzleştirici



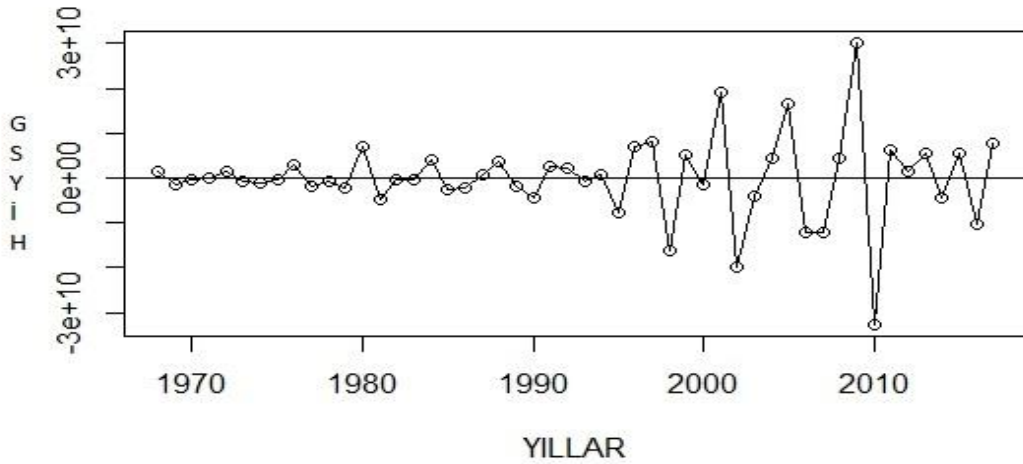
Grafik 4.2.3.1. GSYİH'nın gerçek seviyesi ile ilgili düzleştirilmiş değerler ve % 95 güven aralıklarıyla ilgili veriler arasındaki karşılaştırmayı ve düzeltme etkisini göstermektedir. Burada serilerin çizimlerinden, düzleştirilmiş seviyenin gerçek verilere ne kadar uygun olduğu açık bir şekilde görülmektedir.

Grafik 4.2.3.2. 1968-2017 Türkiye GSYİH Verileri İçin Kalman Filtreleme ve Düzleştirici



Grafik 4.2.3.2., filtreleme ve düzleştirici seviyelerini bir arada göstermektedir. Filtreleme ve düzleştirici seviyelerin gerçek verilere uygun oldukları grafikte açık bir şekilde görülmektedir. Bu uygunluk sonraki yıllarda yapacağımız tahminlerin gerçeğe uygun çıkması için büyük önem taşımaktadır.

Grafik 4.2.3.3. 1968-2017 Türkiye GSYİH Verileri İçin Dönemsel Değişim



Grafik 4.2.3.3. incelendiğinde, Türkiye’ye ait GSYİH’nın dönemsellik değişiminin yaklaşık olarak 1995’li yıllardan sonra olduğu görülmektedir. Bu dönemlerden sonra yükseliş ve düşüşlerde büyük bir fark görülmektedir. Özellikle 1999, 2002 ve 2010 yıllarında bu düşüşlerin daha yüksek görüldüğü söylenebilir. Bunun temel sebebi ise o yıllarda yaşanan krizler gösterilebilir.

4.2.4. Tahmin Modelinin Karşılaştırılması

İki modeli karşılaştırmak için çeşitli tahmin doğruluğu ölçümleri kullanılabilir. Yaygın olarak kullanılan kriterler, sırasıyla aşağıdaki formüllerle tanımlanan ortalama mutlak sapma (MAD), ortalama kare hatası (MSE) ve ortalama mutlak yüzde hatasıdır (MAPE).

$$MAD = \frac{1}{n} \sum_{t=1}^n |e_t|$$

$$MSE = \frac{1}{n} \sum_{t=1}^n e_t^2$$

$$MAPE = \frac{1}{n} \sum_{t=1}^n \frac{|e_t|}{y_t}$$

Aşağıda yer alan tablo 4.2.4.1 tahmin için hangi modelin uygun olduğunu göstermektedir.

Tablo 4.2.4.1. Model 1 ve Model 2 Karşılaştırma Sonuçları

MAD (Model1)	MAD (Model2)	MSE (Model1)	MSE (Model2)	MAPE (Model1)	MAPE (Model2)
377296830	334883524	3.135129e+2	2.48595e+21	0.1704686	0.1563612

Tablo 4.2.4.1. analiz edildiğinde, filtreleme modeli1 için değerlerin genel olarak yüksek olduğunu model 2 için ise daha düşük olduklarını görmekteyiz. Bu nedenle tahminlerde model 2’nin kullanılması daha sağlıklı sonuçlar verecektir. Fakat sadece bu göstergeler ile bu sonuca varılmasının sağlıklı olmayacağı yapılan çalışmalarda

açıklanmaktadır. Bu nedenden dolayı belirli bir modelin tahmin performansını değerlendirmek için kullanılabilecek etkili tahminin Theil(1966)'e ait TheilU istatistiğidir. TheilU eşitsizlik katsayısının küçük bir değere sahip olması yapılacak tahmin için doğruluk oranını artırmaktadır. Aşağıda yer alan denklem TheilU denklemini göstermektedir.

$$U = \sqrt{\frac{\sum_{t=2}^n (y_t - f_t)^2}{\sum_{t=2}^n (y_t - y_{t-1})^2}}$$

Aşağıda yer alan tablo 4.2.4.2., model1 ve model2 için TheilU istatistik sonuçları verilmektedir.

Tablo 4.2.4.2. TheilU İstatistik Sonuçları

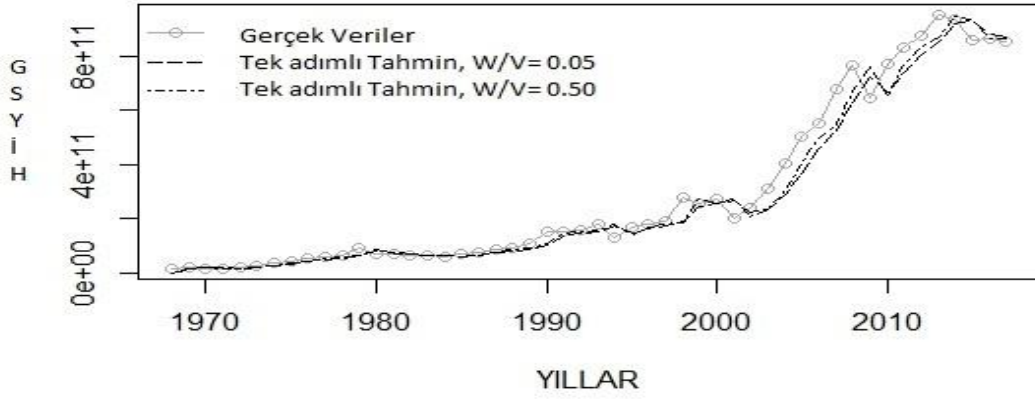
TheilU (Model1)	TheilU (Model2)
1.12149	0.189837

TheilU modele ait MSE değerini tahmini yapılacak değerlerin MSE'si ile karşılaştırmamızı da sağlar. Bu değer model2'e ait MSE değerinden küçük olması ve 0'a yakın olması yapılacak tahminin doğru sonuçlar içerdiğini göstermektedir. Tablo 4.2.4.2. analiz edildiğinde, TheilU değerlerinin modele ait değerlerden çok düşük olduğu ve gerçek değerlere yakın hareket eden model2'nin 0'a yakın olduğunu görmekteyiz. Birden düşük bir TheilU değeri, tahmin edilen modelin ortalama olarak değişiklik yapmadan modelden daha iyi tahminler ürettiği anlamına gelir.

4.2.5. GSYİH İçin Tahmin

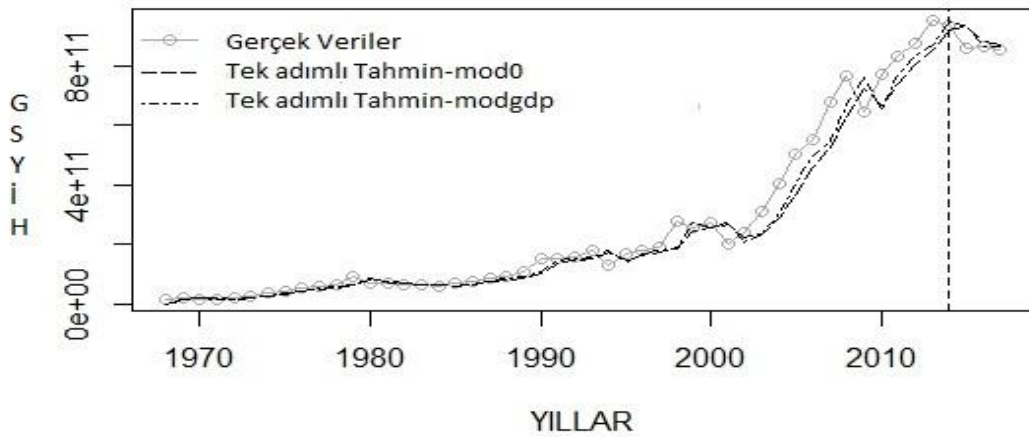
Bu bölümde, Türkiye'ye ait 1968-2017 yılları arası GSYİH verileri kullanılarak 2018-2029 yılları arasında nasıl bir değişimin olacağı tahmin edilmektedir. Tahminin gerçeğe yakın olması amacı ile 2014 yılında kestirim yapıp geriye kalan son üç yıla(2015,2016,2017) göre gelecek tahmini yapılmaktadır. İleriye dönük 12 adımlık bir tahmin yapıp ve yeni veriler 10 tahmin hesap süreci içinde elde edilmektedir. İleriye dönük tahmin süresi ve tahmin hesap süreci isteğe bağlı olarak değiştirilebilir. Aşağıda yer alan grafikler bu adımları göstermektedir.

Grafik 4.2.5.1. 1968-2017 Türkiye GSYİH Gerçek Verileri ile Tahmin Verilerinin Gösterimi



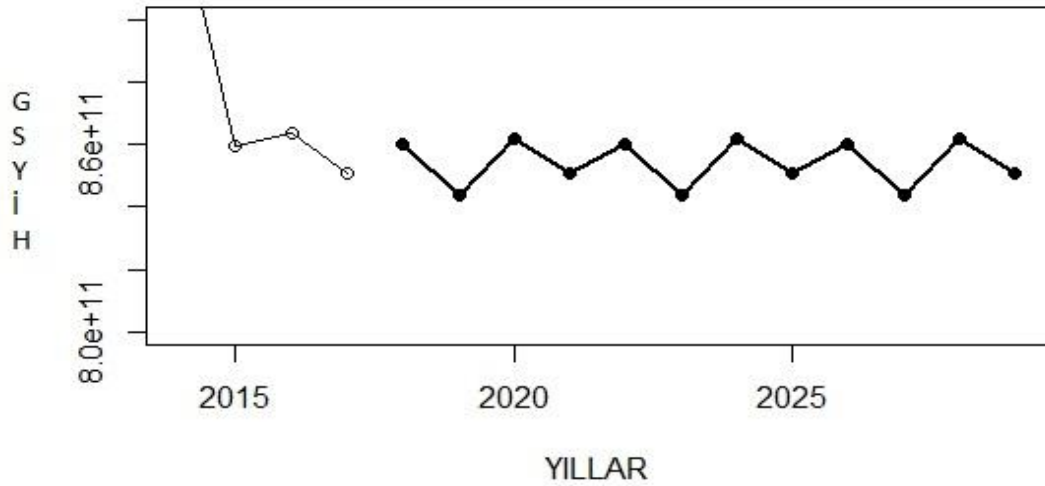
Grafik 4.2.5.1. 196-2017 yılları arasında GSYİH’ye ait gerçek veriler ile tahmini veriler arasındaki yakınlık derecesini göstermektedir. Grafikte toplam üç veri bulunmaktadır. Gri olan hat gerçek verileri, kalın kesikli hat model 1’in tahmini verilerini ve ince kesikli olan hat ise model 2’nin verilerini göstermektedir. Filtreleme kısmında da olduğu gibi burada da model 2’nin gerçek verilere daha yakın hareket ettiği görülmektedir.

Grafik 4.2.5.2. 1968-2017 Türkiye GSYİH Kestirim Grafiği



Grafik 4.2.5.2. 2014 yılını kestirim yılı olarak alınmasını göstermektedir. İleriye dönük yapılacak tahmin son 3 yılı baz alarak yapılmaktadır. Aşağıda yer alan grafik bunu daha detaylı göstermektedir.

Grafik 4.2.5.3. 2018-2029 Türkiye GSYİH Tahmini Verileri



Grafik 4.2.5.3. 2018-2029 yılları arasında Türkiye'ye ait tahmini GSYİH verilerini göstermektedir. Tahmin sonuçları 2014 yılında yapılan kestirimden sonraki üç yıl örnek alınarak DLM tahmincisi tarafından R programı ile oluşturuldu. Grafik analizi yapıldığında önceki 3 yıla göre GSYİH'da düşüşün olduğu görülmektedir. Tahmin yılları aralığında bir dalgalanma görülmektedir. Özellikle 2019, 2024 ve 2028 yıllarında düşüşün olduğu izlenilmektedir.

BEŞİNCİ BÖLÜM

5. DEĞERLENDİRME VE SONUÇ

5 bölümden oluşan bu tezde, ilk olarak DUM ve DLM tarihçesi anlatıldı. İkinci bölümde, doğrusal regresyon modelinde Bayes çıkarımı ve DUM detaylı bir şekilde anlatılıp denklemler ile verildi. Üçüncü bölümde, DLM'ler denklemler ile tanıtıldıktan sonra DLM'ler için yapılan işlemler ayrı başlıklar altında ayrı ayrı verildi. Bu bölümde Kalman filtreleme ve Kalman düzleştirici detaylı şekilde anlatılıp DLM tahminleri için nasıl kullanıldığı verildi. Bölüm 4'te ise bölüm 3'te yer alan DLM ve alt başlıklar Türkiye'ye ait GSYİH'nin tahmini için uygulandı ve sonuçlar ayrı ayrı grafik ve istatistiksel değerler ile değerlendirildi. Son bölüm olan değerlendirme ve sonuç bölümünde ise Türkiye'nin ekonomik olarak güçlü bir ülke olması için bazı önerilerde bulunularak tez çalışması tamamlandı.

Türkiye'nin GSYİH'sini gösteren gerçek verileri Dünya Bankası resmi web adresinden alındıktan sonra R programına yüklendi. Dördüncü bölümde Kalman filtreleme ve düzleştirici kullanılarak, Türkiye'ye ait GSYİH'nin 2018- 2029 yılları arasında nasıl bir değişim göstereceği tahmin edildi. Filtreleme için kullanılan iki modelden hangisinin uygun olduğunu tespit etmek için, model 1 ve model 2'ye ait MAD, MSE ve MAPE değerleri elde edildi. Bu sonuçlar doğrultusunda model 2'nin gerçek verilere daha uygun hareket ettiği görüldü. Daha sonra yapılan filtrelemenin gelecek için yapılacak tahminlerde gerçeğe yakın olup olmadığını görmek için, TheilU istatistik değeri analiz edildi. TheilU değerinin 0'a yakın bir değer alması, yapılacak tahminin gerçeğe yakınlık durumunu göstermektedir. Elde edilen TheilU değerinin 1'den küçük olduğu tablo 4.2.4.2.'de detaylı şekilde görülmektedir.

Bölüm 4 te yer alan analizlerden sonra grafik 4.2.5.3. Türkiye için 2018- 2029 yılları arasında GSYİH değerlerinin nasıl bir değişim gösterdiği görülmektedir. Genel olarak dalgalı bir seyir izleyen GSYİH değerlerinin önceki yıllara göre bir düşüş yaşadığı grafik 4.2.5.3.'te açık bir şekilde görülmektedir. Bu sonucun gelişmekte olan bir ülke için tehlike arz ettiği düşünülebilir. Güçlü bir ekonomiye sahip olmak ve dünya rekabet pazarında yer alabilmek için yönetimin sağlıklı programlar izlemesi

gerekmektedir. Ülkemiz için aşağıda verilen bazı öneriler güçlü bir ekonomi sağlamak için uygulanabilir.

- Düzenli bir gelir dağılımının sağlanması
- Beşeri sermayelerle üretici yatırımlarına teşvik sağlanması
- Eğitim alanında reformlar yapılması ve teknolojik değişime uyum sağlanması
- Sağlık alanında gelişimlerin takip edilmesi
- Tasarrufların düzenli yapılması ve üretime yönelik yatırımların yapılması
- Sanayi, teknoloji ve tarım alanlarında üretim yapılıp dışa bağımlılığın düşürülmesi
- Turizm alanında ülkenin detaylı bir şekilde tanıtılması ve turizm bölgelerinde turizme dayalı yatırımların artırılması
- AR-GE projelerinin desteklenmesi ve geleceğe yönelik genç nüfusun üretime katkıda bulunmasının sağlanması. Şeklinde öneriler sunulabilir.

KAYNAKÇA

- Akaike, H., (1974) "Markovian representation of stochastic processes and its application to the analysis of autoregressive moving average processes", Annals of the Institute of Statistical Mathematics 26.
- Anderson, B. ve Moore, J., (1979) "Optimal Filtering, Prentice-Hall, Englewood Cliffs".
- Aoki, M., (1979) "Introduction to Stochastic Control Theory, Academic Press", New York.
- Aoki, M., (1987) "State Space Modeling of Time Series", Springer Verlag, New York.
- Barndorff-Nielsen, O., Cox, D. ve Klüppelberg, C., (2001) "Complex Stochastic Systems", Chapman & Hall, London.
- Cowell, R., Dawid, P., Lauritzen, S. ve Spiegelhalter, D., (1999) "Probabilistic networks and expert systems", Springer-Verlag, New York.
- Durbin, J. ve Koopman, S., (2001) "Time Series Analysis by State Space Methods", Oxford University Press, Oxford.
- Harrison, P. ve Stevens, C., (1976) "Bayesian forecasting (with discussion)", Journal of the Royal Statistical Society, Series B 38.
- Harvey, A., (1989) "Forecasting, Structural Time Series Models and the Kalman filter", Cambridge University Press, Cambridge.
- Ho, Y. C. ve Lee, R. C. K., (1964) "A Bayesian Approach to Problems in Stochastic Estimation and Control, Automatic Control", IEEE Transactions on, 9(4).
- Kalman, R., (1960) "A new approach to linear filtering and prediction problems, Trans. of the AMSE", Journal of Basic Engineering (Series D), 82.
- Kalman, R. E., (1960) "A New Approach to Linear Filtering and Prediction Problems", Journal of Basic Engineering, 82(1).

- Kalman, R. E. ve Bucy, R. S., (1960) "New Results in Linear Filtering and Prediction Theory, Trans. of the AMSE", Journal of Basic Engineering (Series D), 83.
- Kalman, R. ve Bucy, R., (1963) "New results in linear filtering and prediction theory, Trans. of the AMSE", Journal of Basic Engineering (Series D), 83.
- Kolmogorov, A., (1941) "Interpolation and extrapolation of stationary random sequences, Bull", Moscow University, Ser. Math. 5.
- Künsch, H., (2001) "State space and hidden Markov models, in O. Barndorff Nielsen, D. Cox and C. Klüppelberg (eds), Complex stochastic systems, Chapman & Hall/CRC, Boca Raton".
- Lauritzen, S. L., (2002) "Thiele: Pioneer in Statistics", Oxford University Press, New York.
- Migon, H., Gamerman, D., Lopez, H. ve Ferreira, M., (2005) "Bayesian dynamic models, in D. Day and C. Rao (eds)", Handbook of Statistics, Vol. 25, Elsevier B.V., chapter 19.
- Morf, M. ve Kailath, T., (1975) "Square-root algorithms for least-squares estimation", IEEE Trans. Automatic Control AC-20.
- Oshman, Y. ve Bar-Itzhack, I., (1986) "Square root filtering via covariance and information eigenfactors", Automatica 22.
- Petris, G., Petrone, S. ve Campagnoli, P., (2009) "Dynamic Linear Models with R", Springer, New York.
- Petris, G., (2010) "An R Package for Dynamic Linear Models", Journal of Statistical Software, 36(12).
- Plackett, R. L., (1950) "Some Theorems in Least Squares", Biometrika, 37.
- Young, F. W., (1969) "Polynomial Conjoint Analysis of Similarities: Definitions for a Special Algorithm", Psychometric Laboratory, University of North Carolina.

- Wang, L., Liber, G. ve Manneback, P., (1992) “Kalman filter algorithm based on singular value decomposition”, Proc. of the 31st Conf. on Decision and Control.
- West, M. ve Harrison, J., (1997) “Bayesian Forecasting and Dynamic Models”, 2nd edn, Springer, New York.
- Wiener, N., (1949) “The Extrapolation, Intepolation and Smoothing of Stationary Time Series”, Wiley, New York.
- Wold, H., (1938) “A Study in the Analysis of Stationary Time Series”, Almquist and Wiksell, Uppsala.
- World Bank Group, (2018) <https://databank.worldbank.org/data/source/world-development-indicators>, (Eriřim tarihi: 14.11.2018).

EKLER

EK 1: Dünya Bankası’ndan alınan Türkiye’ye ait GSYİH verileri Milyon Dolar(1968-2017)

YILLAR	GSYİH	1994	130690000000
1968	17500000000	1995	169486000000
1969	19466666667	1996	181476000000
1970	17086956522	1997	189835000000
1971	16256619964	1998	275769000000
1972	20431095406	1999	255884000000
1973	25724381625	2000	272979000000
1974	35599913836	2001	200252000000
1975	44633707243	2002	238428000000
1976	51280134554	2003	311823000000
1977	58676813687	2004	404787000000
1978	65147022486	2005	501416000000
1979	89394085658	2006	552487000000
1980	68789289566	2007	675770000000
1981	71040020140	2008	764336000000
1982	64546332581	2009	644640000000
1983	61678280115	2010	771902000000
1984	59989909458	2011	832524000000
1985	67234948265	2012	873982000000
1986	75728009963	2013	950579000000
1987	87172789528	2014	934186000000
1988	90852814005	2015	859797000000
1989	107143000000	2016	863722000000
1990	150676000000	2017	851102000000
1991	150028000000		
1992	158459000000		
1993	180170000000		

EK 2: 2018-2029 yıllarına Ait Tahmin Edilen Veriler

YILLAR	GSYİH
2018	860175075047
2019	843858954589
2020	862142373293
2021	851102411223
2022	860175075029
2023	843858954564
2024	862142373252
2025	851102411194
2026	860175074988
2027	843858954560
2028	862142373236
2029	851102411148

EK 3: Tahmin için Kullanılan R Kodları

Model kurulum için gerekli paket ve kütüphane

```
library(dlm)
```

```
app.data <- as.data.frame(gdpp)$GDP
```

```
app.data <- ts(app.data, start = c(1968), frequency = 1)
```

```
plot(app.data, xlab = "YILLAR", ylab = "GSYİH")
```

n. düzeyde bir DLM kurulumu için gerekli kodlar

```
gdpBuild <- function(par) {
```

```
  dlmModPoly(1, dV = exp(par[1]), dW = exp(par[2]))
```

```
}
```

```
gdpMLE <- dlmMLE(app.data, rep(0,2), gdpBuild); gdpMLE$converged
```

```
gdpMod <- gdpBuild(gdpMLE$par)
```

```
str(gdpMLE)
```

```
v <- drop(V(gdpMod))
```

```
w <- drop(W(gdpMod))
```

```
gdpPoly <- dlmModPoly(order = 1, dV = v, dW = w)
```

```
unlist(gdpPoly)
```

```
gdpFilt <- dlmFilter(app.data, gdpPoly)
```

```
str(gdpFilt, 1)
```

```
n <- length(app.data)
```

```
U.C <- gdpFilt$U.C
```

```
D.C <- gdpFilt$D.C
```

```

dlmSvd2var(U.C[[n + 1]], D.C[n + 1, ])

plot(app.data, type='o', col = c("darkgrey"),

      xlab = "YILLAR", ylab = "GSYİH")

### Kalman Filtreleme Kodları ###

mod1 <- dlmModPoly(order = 1, dV = 2, dW = 3)

gdpFilt1 <- dlmFilter(app.data, mod1)

lines(dropFirst(gdpFilt1$m), lty = "longdash", col=2)

mod2 <- dlmModPoly(order = 1, dV = 8, dW = 100)

gdpFilt2 <- dlmFilter(app.data, mod2)

lines(dropFirst(gdpFilt2$m), lty = "dotdash", col=4)

leg <- c("data", paste("filtered, W/V =",

                        format(c(W(mod1) / V(mod1),

                                W(mod2) / V(mod2))))))

legend("topleft", legend = leg,

       col=c("darkgrey", "2", "4", '6'),

       lty = c("solid", "longdash", "dotdash", "dotdash"),

       pch = c(1, NA, NA, NA), bty = "n")

plot(app.data, type='o', col = c("darkgrey"),

      xlab = "YILLAR", ylab = "GSYİH")

### MAD, MSE, MAPE, TheilU, Değerleri###

mean(abs(gdpFilt1$f - app.data))

mean(abs(gdpFilt2$f - app.data))

```

```

mean((gdpFilt1$f - app.data)^2)

mean((gdpFilt2$f - app.data)^2)

mean(abs(gdpFilt1$f - app.data) / app.data)

mean(abs(gdpFilt2$f - app.data) / app.data)

sqrt(sum((gdpFilt1$f - app.data)[-(1:25)]^2) /

      sum(diff(app.data[-(1:2)])^2))

sqrt(sum((gdpFilt2$f - app.data)[-(25:50)]^2) /

      + sum(diff(app.data[-(1:2)])^2))

### Kalman Filtreleme ve Düzleştirici göstergeleri ###

expdFilt <- dlmFilter(app.data , expd.dlm)

lines(dropFirst(expdFilt$m[, 1]), lty = "dotdash")

expdSmooth <- dlmSmooth(expdFilt)

lines(dropFirst(expdSmooth$s[,1]), lty = "longdash")

legend("topleft", col = c("darkgrey", rep("black", 2)),

      lty = c("solid", "dotdash", "longdash"),

      pch = c(1, NA, NA), bty = "n",

      legend = c("data", "filtered level", "smoothed level"))

### Kalman Düzleştirici ###

library(dlm)

gdpSmooth <- dlmSmooth(gdpFilt)

U.S<- gdpSmooth$U.S

D.S<- gdpSmooth$D.S

```

```

s <- gdpSmooth$s

str(gdpSmooth, 1)

drop(dlmSvd2var(U.S[[n + 1]], D.S[n + 1,]))

drop(dlmSvd2var(U.C[[n + 1]], D.C[n + 1,]))

drop(dlmSvd2var(U.S[[n / 2 + 1]], D.S[n / 2 + 1,]))

hwid <- qnorm(0.025, lower = FALSE) *

  sqrt(unlist(dlmSvd2var(U.S, D.S)))

smooth <- cbind(s, as.vector(s) + hwid %o% c(-1, 1))

plot(dropFirst(smooth), plot.type = "s", type = "l",

  lty = c(1, 5, 5), ylab = "GSYİH", xlab = "YILLAR",

  ylim = range(app.data))

lines(app.data, type = "o", col = "darkgrey")

legend("topleft", col = c("darkgrey", rep("black", 2)),

  lty = c(1, 1, 5), pch = c(1, NA, NA), bty = "n",

  legend = c("data", "smoothed level",

    "95% probability limits"))

### Mevsimlik Bileşenler ###

plot(dropFirst(expdSmooth$s[, 3]), type = 'o', xlab = "YILLAR",

  ylab = "GSYİH")

abline(h = 0)

```

```

app.data <- ts(app.data, start = c(1968), frequency = 1)

expd.dlm <- dlm(m0 = rep(0,4), C0 = 1e+07 * diag(4),

               FF = matrix(c(1, 1, 0, 0), nr = 1),

               V = 2.741148e-05,

               GG = bdiag(matrix(1),

                           matrix(c(-1, -1, -1, 1, 0, 0, 0, 1, 0),

                                   nr = 3, byrow = TRUE))),

               W = diag(c(771.35, 86.48, 0, 0), nr = 4))

plot(app.data , xlab = "YILLAR", ylab = "GSYİH", type = 'o',

     col = "darkgrey")

### Tahmin Kısmı ###

a <- window(cbind(app.data, gdpFilt1$f, gdpFilt2$f),

            start = (1968), end = (2017))

plot(a[, 1], type = 'o', col = "darkgrey",

     xlab = "YILLAR", ylab = "GSYİH")

lines(a[, 2], lty = "longdash")

lines(a[, 3], lty = "dotdash")

leg <- c("data", paste("one-step-ahead forecast, W/V =",

                       format(c(W(mod1) / V(mod1),

                               W(mod2) / V(mod2))))))

legend("topleft", legend = leg,

     col = c("darkgrey", "black", "black"),

```

```

lty = c("solid", "longdash", "dotdash"),

pch = c(1, NA, NA, NA), bty = "n")

plot(app.data , xlab = "YILLAR", ylab = "GSYİH", type = 'o',

      col = "darkgrey")

library(dlm)

mod0 <- dlmModPoly(order = 1, dV = v, dW = w)

X <- ts(matrix(mod0$W, nc = 1, nr = length(app.data)),

        start = start(app.data))

window(X, 1968, 2017) <- 12 * mod0$W

modgdp <- mod0

modgdp$X <- X

modgdp$JW <- matrix(1, 1, 1)

gdpFilt <- dlmFilter(app.data, modgdp)

mod0Filt <- dlmFilter(app.data, mod0)

a <- window(cbind(app.data, mod0Filt$f, gdpFilt$f),

            start = (1968), end = (2017)

            plot(a[, 1], type = 'o', col = "darkgrey",

                  xlab= "YILLAR", ylab= "GSYİH")

            lines(a[, 2], lty = "longdash")

            lines(a[, 3], lty = "dotdash")

            abline(v=2014, lty=2)

            leg <- c("data", paste("one-step-ahead forecast -",

```

```

c("mod0", "modgdp"))))

legend("topleft", legend = leg,

      col = c("darkgrey", "black", "black"),

      lty = c("solid", "longdash", "dotdash"),

      pch = c(1, NA, NA), bty = "n")

set.seed(1)

app.data <- ts(app.data, start = c(1968), frequency = 1)

expdFore <- dlmForecast(expdFilt, nAhead = 12, sampleNew = 10)

plot(window(app.data, start = c(1968)), type = 'o',

      xlim = c(2014,2029), ylim = c(8e+11,9e+11),

      xlab = "YILLAR", ylab = "GSYİH")

names(expdFore)

newObs

attach(expdFore)

invisible(lapply(newObs, function(x)

  lines(x, col = "darkgrey",

        type = 'o', pch = 20)))

lines(expdFore$f, type = 'o', lwd = 2, pch = 16)

abline(v = mean(c(time(f)[1], time(app.data)[length(app.data)])),

      lty = "dashed")

detach()

```


ÖZ GEÇMİŞ

Kişisel Bilgiler	İletişim Bilgileri	Eğitim Bilgileri
Ad-Soyad: Çetin GÖRÜR Cinsiyet: Erkek Medeni Durum: Bekâr Doğum Yeri ve Tarihi: Van/1990	Mail: gorurcetin@hotmail.com gorurcetin@gmail.com gorurcetin@yandex.com	İlköğretim: Sempaş İlköğretim Okulu, 2004/Van Lise: Kazım Karabekir Lisesi, 2007/Van Lisans: Çukurova Üniversitesi, Ekonometri(İng.), 2014/Adana Yüksek Lisans: İnönü Üniversitesi, Ekonometri, 2019/Malatya

Sertifika ve Kurslar	Bilgisayar Bilgisi	Yabancı Dil Bilgisi
İnsan Kaynakları Yönetimi	İyi seviyede, Microsoft Office Programları, Eviews, Gauss, R, Stata, SPSS.	İyi seviyede İngilizce YÖKDİL(SOSYAL): 81,250

Yayınlar	Bildiriler
İşleyen, Ş., Altun, Y., Görür, Ç., (2017) “The Causality Relationship between Interest Rate and Income with Investment in USA: 1965-2016”, The Journal of Academic Social Science, 10/12/2017, Sayı 60.	İşleyen, Ş., Altun, Y., Görür, Ç., (2017) “Somut Materyallerle Matematik Öğretiminin Öğrenci Başarısı Üzerindeki Etkisi”, Uluslararası Kültür ve Medeniyet Kongresi 2017, Mardin.

İşleyen, Ş., Altun, Y., Görür, Ç., (2017) “Türkiye’de Tekstil Sektöründe Orta Gelir Tuzağı Analizi”, Journal of Current Researches on Business and Economics, 2017,Cilt7, Sayı 2. Altun, Y., İşleyen, Ş., Görür, Ç., (2018) “Türkiye’de Eğitim ve Sağlık Harcamalarının Ekonomik Büyümeye Etkisi: 1999-2017”, Sosyal Bilimler Enstitüsü Dergisi, The Journal of Social Sciences Institute,2018, Sayı 39, Van. İşleyen, Ş., Altun, Y., Görür, Ç., (2018) “Hizmet İhracatı ve Turizm Harcamalarının Ekonomik Büyümeye Katkısı Üzerine Ampirik Bir Analiz: 1996-2017 Türkiye Örneği”, Anemon Muş Alparslan Üniversitesi Sosyal Bilimler Dergisi 2018 6(6) 953–960, Muş. Yergin, H., Özar, Ş., Görür, Ç., (2018) “İhracat ve İthalattan Alınan Vergilerin Ekonomik Büyüme Üzerinde Etkisine Ait Ampirik Bir Çalışma: 2004-2016 Türkiye Örneği”, Atlas International Refereed Journal On Social Sciences, ISSN:2619-936X.	İşleyen, Ş., Altun, Y., Görür, Ç., (2017) “Sosyal Bilimlerde okutulan Matematik dersine ait öğrenci görüşleri”, IMESET 2017, Bitlis. İşleyen, Ş., Altun, Y., Görür, Ç., (2018) “Geçmişten Günümüze Dünya Ekonomisine Yön Veren Ülkeler ve Sürdürülebilirlik Anlamında Türkiye Ekonomisi Üzerine Bir Mukayese”, UMTEB 2018, Erzurum. İşleyen, Ş., Altun, Y., Görür, Ç., (2018) “Türkiye’de Ekonomik Krizler ve İstihdam Üzerine Bir Değerlendirme”, UMTEB 2018, Erzurum.
--	--