



**T.C.**

**ORDU ÜNİVERSİTESİ**  
**FEN BİLİMLERİ ENSTİTÜSÜ**

**ÇOKLU BAĞLANTI DURUMUNDA YANLI REGRESYON**  
**YÖNTEMLERİNİN İNCELENMESİ**

**DİLDAR AYŞE DERMAN**

**YÜKSEK LİSANS TEZİ**  
**ZOOTEKNİ ANABİLİM DALI**

**ORDU 2019**

**T.C.**  
**ORDU ÜNİVERSİTESİ**  
**FEN BİLİMLERİ ENSTİTÜSÜ**  
**ZOOTEKNİ ANABİLİM DALI**

**ÇOKLU BAĞLANTI DURUMUNDA YANLI REGRESYON  
YÖNTEMLERİNİN İNCELENMESİ**

**DİLDAR AYŞE DERMAN**

**YÜKSEK LİSANS TEZİ**

**ORDU 2019**

## TEZ ONAY

Dildar Ayşe DERMAN tarafından hazırlanan "ÇOKLU BAĞLANTI DURUMUNDA YANLI REGRESYON YÖNTEMLERİNİN İNCELENMESİ" adlı tez çalışmasının savunma sınavı 25.02.2019 tarihinde yapılmış ve jüri tarafından oy birliği / oy çokluğu ile Ordu Üniversitesi Fen Bilimleri Enstitüsü ZOOTEKNİ ANABİLİM DALI YÜKSEK LİSANS TEZİ olarak kabul edilmiştir.

Danışman  
Dr. Öğr. Üyesi Yeliz KAŞKO ARICI

Jüri Üyeleri

Danışman  
Dr. Öğr. Üyesi Yeliz KAŞKO ARICI  
Biyoistatistik ve Tıbbi Bilişim Bölümü,  
Ordu Üniversitesi

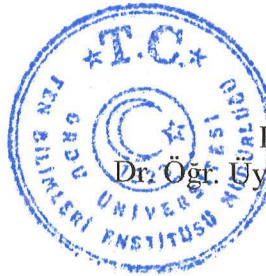
Üye  
Prof. Dr. Sezai ALKAN  
Zootekni Bölümü, Ordu Üniversitesi

Üye  
Doç. Dr. Özgür KOŞKAN  
Zootekni Bölümü, Isparta Uygulamalı  
Bilimler Üniversitesi

İmza



27 / 06 / 2019 tarihinde enstitüye teslim edilen bu tezin kabulü, Enstitü Yönetim Kurulu'nun 28 / 06 / 2019 tarih ve 2019 / 206 sayılı kararı ile onaylanmıştır.



Enstitü Müdürü  
Dr. Öğr. Üyesi Mehmet Sami GÜLER

## TEZ BİLDİRİMİ

Tez yazım kurallarına uygun olarak hazırlanan ve kullanılan intihal tespit programının sonuçlarına göre; bu tezin yazılmasında bilimsel ahlak kurallarına uyulduğunu, başkalarının eserlerinden yararlanılması durumunda bilimsel normlara uygun olarak atıfta bulunulduğunu, tezin içerdiği yenilik ve sonuçların başka bir yerden alınmadığını, kullanılan verilerde herhangi bir tahrifat yapılmadığını, tezin herhangi bir kısmının bu üniversite veya başka bir üniversitedeki başka bir tez çalışması olarak sunulmadığını beyan ederim.

**DİLDAR AYŞE DERMAN**

Not: Bu tezde kullanılan özgün ve başka kaynaktan yapılan bildirişlerin, çizelge, şekil ve fotoğrafların kaynak gösterilmeden kullanımı, 5846 sayılı Fikir ve Sanat Eserleri Kanunundaki hükümlere tabidir.

## ÖZET

### ÇOKLU BAĞLANTI DURUMUNDA YANLI REGRESYON YÖNTEMLERİNİN İNCELENMESİ

DİLDAR AYŞE DERMAN

ORDU ÜNİVERSİTESİ FEN BİLİMLERİ ENSTİTÜSÜ

ZOOTEKNİ ANABİLİM DALI

YÜKSEK LİSANS TEZİ 69 SAYFA

(TEZ DANIŞMANI: DR. ÖĞR. ÜYESİ YELİZ KAŞKO ARICI)

Bu tez çalışmasının amacı, bağımsız değişkenler arasında çoklu bağlantı olması durumunda en küçük kareler regresyonu yerine kullanılması önerilen yanlı regresyon yöntemlerinin incelenmesidir. Bu amaçla, bir bağımlı değişken ile aralarında çoklu bağlantı olduğu belirlenen altı bağımsız değişkenin yer aldığı gerçek bir veri seti kullanılmıştır. Bazı tavuk yumurtası kalite parametrelerini içeren veri setinde, yumurta ak ağırlığı ve yumurta sarı ağırlığı bağımlı değişkenler olarak, diğer kalite parametreleri (yumurta eni, yumurta boyu, yumurta ağırlığı, şekil indeksi, kabuk ağırlığı, kabuk kalınlığı) ise bağımsız değişkenler olarak kabul edilmiştir. Hem yumurta ak ağırlığının hem de yumurta sarı ağırlığının bağımlı değişken olduğu iki adet regresyon modeli oluşturulmuştur. Yanlı regresyon yöntemi olarak, Ridge regresyon analizi, temel bileşenler regresyon analizi ve kısmi en küçük kareler regresyon analizi yöntemleri kullanılmış ve bu yöntemlerin sonuçları en küçük kareler regresyonu ile karşılaştırılmıştır. Performans kriteri olarak, tahmin edilen katsayıların standart hata değerleri, hata kareler ortalaması (HKO), düzeltilmiş belirleme katsayısı ( $R^2$ -düz) ve Akaike Bilgi Kriteri (AIC) kullanılmıştır. En küçük kareler regresyon yönteminde daha düşük HKO ve daha yüksek  $R^2$ -düz değerleri elde edilmesine rağmen, tahminlenen regresyon katsayılarının standart hata değerlerinin daha yüksek olduğu belirlenmiştir. Çoklu bağlantı durumunda, en küçük kareler regresyonu yerine yanlı regresyon yöntemlerinin kullanılmasının tahmin edilen regresyon katsayılarının standart hatalarını normalleştirdiği, dolayısıyla daha güvenilir sonuçlar verdiği ortaya konulmuştur.

**Anahtar Kelimeler:** Çoklu Bağlantı, En Küçük Kareler Regresyonu, Ridge Regresyon Analizi, Kısmi En Küçük Kareler Regresyon Analizi, Temel Bileşenler Regresyon Analizi

## ABSTRACT

### EXAMINATION OF BIASED REGRESSION METHODS UNDER MULTICOLLINEARITY

DİLDAR AYŞE DERMAN

ORDU UNIVERSITY INSTITUTE OF NATURAL AND APPLIED  
SCIENCES

ANIMAL SCIENCE

MASTER'S THESIS, 69 PAGES

(SUPERVISOR: ASSIST. PROF. DR. YELİZ KAŞKO ARICI)

The aim of this thesis study is to investigate the biased regression methods that are suggested to be used instead of least squares regression method in case of happening multicollinearity between independent variables. For this purpose, a real data set including six dependent variables, which are identified as having multicollinearity, is used with a dependent variable. In the data set including some chicken egg quality parameters, egg albumen weight and egg yolk weight were accepted as dependent variables, whereas other quality parameters (egg width, egg height, egg weight, shape index, shell weight, shell thickness) were accepted as independent variables. Two regression models were made up, in which both the egg albumen weight and the egg yolk weight were dependent variables. Ridge regression analysis, principal component regression analysis and partial least squares regression analysis were used as the biased regression and the results of these methods were compared with the least squares regression. The standard errors of the coefficient estimates (SEE), mean squared error (MSE), adjusted coefficient of determination ( $R^2$ -adj) and Akaike Information Criterion (AIC) were used as performance criteria. Although lower MSE and higher  $R^2$ -adj values were obtained in the least squares regression, the coefficient estimates had higher standard error values. In case of multicollinearity, the use of regression methods instead of the least squares regression revealed that regression coefficients gave more reliable results because of normalizing the SEE.

**Keywords:** Multicollinearity, Least Square Regression, Ridge Regression Analysis, Partial Least Squares Regression Analysis, Principal Components Regression Analysis

## TEŞEKKÜR

Tez çalışmamın tüm aşamalarında her daim bilgi ve deneyimleriyle yoluma ışık tutan, akademik birikimlerini benimle sabırla paylaşan kıymetli danışman hocam Sayın Dr. Öğr. Üyesi Yeliz KAŞKO ARICI'ya sonsuz teşekkürlerimi sunarım.

Tez çalışmamın temelini oluşturan veri setini kullanmam için izin veren ve olumlu bakış açısı ile desteğini hiç esirgemeyen değerli hocam Sayın Prof. Dr. Sezai ALKAN'a teşekkürü bir borç bilirim.

Manevi ve maddi desteklerini her an üzerimde hissettiğim kıymetli babam, annem, kardeşim ve eşime çok teşekkür ederim.

Ayrıca tez yazım aşamasında yanımda olan ve tezimi tamamlamama izin veren oğlum sevgili Mete'ye çok teşekkür ederim.

## İÇİNDEKİLER

	<u>Sayfa</u>
<b>TEZ BİLDİRİMİ</b> .....	<b>I</b>
<b>ÖZET</b> .....	<b>II</b>
<b>ABSTRACT</b> .....	<b>III</b>
<b>TEŞEKKÜR</b> .....	<b>IV</b>
<b>İÇİNDEKİLER</b> .....	<b>V</b>
<b>ŞEKİL LİSTESİ</b> .....	<b>VII</b>
<b>ÇİZELGE LİSTESİ</b> .....	<b>VIII</b>
<b>SİMGELER ve KISALTMALAR LİSTESİ</b> .....	<b>IX</b>
<b>1. GİRİŞ</b> .....	<b>1</b>
<b>2. GENEL BİLGİLER</b> .....	<b>4</b>
2.1 En Küçük Kareler (OLS) Regresyonu .....	<b>4</b>
2.1.1 En Küçük Kareler (OLS) Regresyonunun Varsayımları .....	<b>5</b>
2.1.2 Gauss-Markov Teoremi .....	<b>7</b>
2.2 Çoklu Bağlantı (Multicollinearity).....	<b>7</b>
2.2.1 Çoklu Bağlantının Nedenleri.....	<b>8</b>
2.2.2 Çoklu Bağlantıyı Belirleme Yöntemleri .....	<b>8</b>
2.2.3 Çoklu Bağlantı Sorununun Ortadan Kaldırılması .....	<b>10</b>
2.3 Ridge Regresyon (RR) .....	<b>11</b>
2.3.1 Ridge Regresyon ile En Küçük Kareler Regresyonunun İlişkisi .....	<b>13</b>
2.3.2 Ridge Regresyonun Varyansı.....	<b>13</b>
2.3.3 Ridge Regresyon Hata Kareler Ortalaması .....	<b>13</b>
2.3.4 Ridge Regresyonda k Parametresinin Tespit Edilmesi .....	<b>14</b>
2.3.5 Ridge İzi Yöntemi .....	<b>15</b>
2.4 Temel Bileşenler Regresyonu (PCR).....	<b>15</b>
2.4.1 Özdeğerler ve Özvektörler .....	<b>16</b>
2.4.2 Temel Bileşenlerin Belirlenmesi.....	<b>17</b>
2.4.3 Temel Bileşen Sayısının Belirlenmesi .....	<b>19</b>
2.5 Kısmi En Küçük Kareler Regresyonu (PLS) .....	<b>19</b>
2.5.1 Doğrusal Olmayan Yinelemeli En Küçük Kareler Algoritması (NIPALS).....	<b>21</b>
2.5.2 PLS Analizinin İstatistiksel Olarak Esinlenilmiş Değişikliğinin Basit Bir Uygulaması (SIMPLS).....	<b>22</b>
2.5.3 Evrensel Kısmi En Küçük Kareler Algoritması (UNIPALS) .....	<b>23</b>
2.5.4 Çekirdek Algoritması .....	<b>23</b>
2.5.5 Örnek-Uzaklık Kısmi En Küçük Kareler Algoritması (SAMPLS).....	<b>23</b>
2.5.6 Özvektör Algoritması.....	<b>24</b>
2.6 Önceki Çalışmalar .....	<b>24</b>
<b>3. MATERYAL ve YÖNTEM</b> .....	<b>31</b>
3.1 Materyal .....	<b>31</b>
3.2 Yöntem.....	<b>31</b>
3.2.1 Çalışma Değişkenleri .....	<b>31</b>
3.2.2 Regresyon Yöntemlerinin Performans Göstergeleri .....	<b>31</b>
<b>4. BULGULAR ve TARTIŞMA</b> .....	<b>33</b>
4.1 Bulgular.....	<b>33</b>
4.1.1 Çalışma Değişkenlerine Ait Tanıtıcı İstatistik Değerleri .....	<b>33</b>
4.1.2 Yumurta Ak Ağırlığı Bağımlı Değişkenine Ait Bulgular .....	<b>34</b>



4.1.2.1 En Küçük Kareler Analiz Sonuçları.....	34
4.1.2.2 Ridge Regresyon Analiz Sonuçları .....	36
4.1.2.3 Temel Bileşenler Regresyon Analiz Sonuçları .....	41
4.1.2.4 Kısmi En Küçük Kareler Regresyon Sonuçları .....	45
4.1.3 Sarı Ağırlığı Bağımlı Değişkeni İçin Elde Edilen Sonuçlar .....	46
4.1.3.1 En Küçük Kareler Analiz Sonuçları.....	46
4.1.3.2 Ridge Regresyon Analiz Sonuçları .....	47
4.1.3.3 Temel Bileşenler Regresyon Analiz Sonuçları .....	53
4.1.3.4 Kısmi En Küçük Kareler Regresyon Analiz Sonuçları.....	57
4.2 Tartışma.....	59
5. SONUÇ ve ÖNERİLER.....	61
6. KAYNAKLAR .....	62
7. ÖZGEÇMİŞ.....	69

## ŞEKİL LİSTESİ

	<u>Sayfa</u>
Şekil 2.1 Bir Tahmin Edicide Varyans ve Yanlılık .....	12
Şekil 4.1 Ak Ağırlık İçin Ridge İzi Grafiği .....	36
Şekil 4.2 Ak Ağırlık İçin Ridge Regresyon VIF Grafiği .....	37
Şekil 4.3 Ak Ağırlık İçin Beta İzi Grafiği .....	42
Şekil 4.4 Ak Ağırlık İçin Temel Bileşenler VIF Grafiği .....	42
Şekil 4.5 Sarı Ağırlık İçin Ridge İzi Grafiği.....	48
Şekil 4.6 Sarı Ağırlık İçin Ridge Regresyon VIF Grafiği .....	49
Şekil 4.7 Sarı Ağırlık İçin Beta İzi Grafiği .....	54
Şekil 4.8 Sarı Ağırlık İçin Temel Bileşenler VIF Grafiği.....	55

## ÇİZELGE LİSTESİ

	<u>Sayfa</u>
Çizelge 3.1 Çalışma Değişkenleri.....	31
Çizelge 4.1 Çalışmada Kullanılan Değişkenlere Ait Tanıtıcı İstatistikler.....	33
Çizelge 4.2 Değişkenler Arasındaki Korelasyon Katsayıları .....	34
Çizelge 4.3 En Küçük Kareler Varyans Analizi Tablosu .....	34
Çizelge 4.4 En Küçük Kareler Çoklu Regresyon Analizi Sonuçları .....	35
Çizelge 4.5 Değişkenlere Ait Korelasyonların Özdeğerleri ve Koşul İndeksleri .....	35
Çizelge 4.6 Standartlaştırılmış Ridge Regresyon Değerleri .....	38
Çizelge 4.7 Ridge Regresyon k Analiz Tablosu .....	39
Çizelge 4.8 Ridge Katsayısı $k=0.004$ İçin Ridge Regresyon ve En Küçük Kareler Karşılaştırılması .....	40
Çizelge 4.9 Ridge Regresyon Katsayıları .....	40
Çizelge 4.10 Ridge Regresyon Varyans Analizi Tablosu.....	41
Çizelge 4.11 Temel Bileşenler Regresyon Analiz Sonuçları.....	41
Çizelge 4.12 Temel Bileşenler Varyans Büyütme Değerleri.....	43
Çizelge 4.13 Değişkenlerin Seçilen Temel Bileşenler İle Standardize Edilmiş Katsayıları .....	43
Çizelge 4.14 Temel Bileşenler Katsayı Seçimi.....	43
Çizelge 4.15 Gözardı Edilen 2 Temel Bileşenle Hesaplanan Regresyon Katsayıları	44
Çizelge 4.16 Temel Bileşenler Regresyonu Varyans Analizi Tablosu.....	45
Çizelge 4.17 Kısmi En Küçük Kareler Regresyonu Varyans Analizi Tablosu .....	45
Çizelge 4.18 Kısmi En Küçük Kareler Model Seçimi ve Modelin Geçerliliği .....	46
Çizelge 4.19 En Küçük Kareler Varyans Analizi Tablosu .....	46
Çizelge 4.20 En Küçük Kareler Çoklu Regresyon Analizi Sonuçları .....	47
Çizelge 4.21 Korelasyonların Özdeğerleri ve Koşul İndeks Değerleri.....	47
Çizelge 4.22 Standartlaştırılmış Ridge Regresyon Değerleri .....	50
Çizelge 4.23 Ridge Regresyon k Analiz Tablosu .....	51
Çizelge 4.24 Ridge Katsayısı $k=0.004$ İçin Ridge Regresyon ve En Küçük Kareler Karşılaştırılması .....	52
Çizelge 4.25 Ridge Regresyon Katsayıları .....	52
Çizelge 4.26 Ridge Regresyon Varyans Analizi Tablosu.....	53
Çizelge 4.27 Temel Bileşenler Regresyon Analiz Sonuçları.....	53
Çizelge 4.28 Temel Bileşenler Varyans Büyütme Değerleri.....	55
Çizelge 4.29 Değişkenlerin Seçilen Temel Bileşenler İle Standardize Edilmiş Katsayıları .....	56
Çizelge 4.30 Temel Bileşenler Katsayı Seçimi.....	56
Çizelge 4.31 Gözardı Edilen 2 Temel Bileşenle Hesaplanan Regresyon Katsayıları	56
Çizelge 4.32 Temel Bileşenler Regresyonu Varyans Analizi Tablosu.....	57
Çizelge 4.33 Kısmi En Küçük Kareler Regresyonu Varyans Analizi Tablosu .....	58
Çizelge 4.34 Kısmi En Küçük Kareler Model Seçimi ve Modelin Geçerliliği .....	58
Çizelge 4.35 Ak Ağırlık İçin Analiz Sonuçlarının İncelenmesi .....	59
Çizelge 4.36 Sarı Ağırlık İçin Analiz Sonuçlarının İncelenmesi.....	59

## SİMGELER ve KISALTMALAR LİSTESİ

---

<b>HKO</b>	: Hata Kareler Ortalaması
<b>k</b>	: Ridge Regresyon Parametresi
<b>KI</b>	: Koşul İndeksi
<b>n</b>	: Gözlem Sayısı
<b>NIPALS</b>	: Doğrusal Olmayan Yinelemeli En Küçük Kareler Algoritması
<b>OLS</b>	: En Küçük Kareler
<b>p</b>	: Değişken Sayısı
<b>PCR</b>	: Temel Bileşenler Regresyonu
<b>PLS</b>	: Kısmi En Küçük Kareler Regresyonu
<b>RR</b>	: Ridge Regresyon
<b>SAMPLS</b>	: Örnek- Uzaklık Kısmi En Küçük Kareler Algoritması
<b>SIMPLS</b>	: PLS Yönteminin İstatistiksel Olarak Esinlenilmiş Değişikliğinin Basit Bir Uygulaması
<b>UNIPALS</b>	: Evrensel Kısmi En Küçük Kareler Algoritması
<b>Y</b>	: Bağımlı Değişken
<b>X</b>	: Bağımsız Değişken
<b>VIF</b>	: Varyans Büyütme Faktörü

---

## 1. GİRİŞ

Bağımlı değişken; diğer değişkenlerden etkilenen, açıklanan değişken iken, bağımsız değişken; açıklayan, neden olan değişkendir (Alpar, 2003). Araştırmacılar bağımlı ve bağımsız değişkenler arasındaki ilişkiyi matematiksel modelle açıklamayı hedeflerler (Alpar, 2003). Regresyon analizi, aralarında sebep sonuç ilişkisi olan iki veya daha fazla değişken arasındaki ilişkiyi modellemek için kullanılan istatistiksel analiz yöntemidir (Vural, 2007). Bir bağımlı ve bir bağımsız değişkenden oluşan regresyon modellerinde, matematiksel modelle ilişki ortaya konarak bilinen bağımsız değişken ile bilinmeyen bağımlı değişken tahmin edilebilir (Alpar, 2003). Yani, bağımlı değişkeni açıklayan bir bağımsız değişken varsa basit regresyon analizi denir (Tunç, 2018). Bir bağımsız değişken ile bir bağımlı değişken arasındaki ilişki, doğrusal ve doğrusal olmayan ilişkiler olarak ikiye ayrılır. Doğrusal ilişkilerde, bağımsız değişkenin bağımlı değişkene etkisi doğrusaldır ve toplanabilir yapıdadır. Doğrusal olmayan ilişkide, bağımsız değişkenin bağımlı değişkene etkisi toplanabilir yapıda değildir (Tunç, 2018). Bağımlı değişkeni etkileyen birden çok bağımsız değişken varlığında çoklu regresyon analizi oluşur. Çoklu regresyon analizi ile; bağımlı değişken ile bağımsız değişkenler arasındaki ilişkiler model doğrultusunda incelenir, bağlantılar kurulur (Alpar, 2003). Elde edilecek model doğrultusunda önemli amaçlar şunlardır;

1. Bağımlı değişkenin hangi bağımsız değişkenlerden etkilendiğini bulmak ve etkileyenler içerisinde hangi değişkenlerin etkisinin daha fazla olduğunu araştırmaktır.
2. Bağımsız değişkenler vasıtasıyla mümkün olan en yüksek isabet derecesi ile bağımlı değişkenin değerini tahmin etmektir.

Çoklu regresyon analizinde kullanılan en yaygın yöntem “en küçük kareler (OLS) regresyonu”dur. En küçük kareler regresyonu, hataların normal dağılım gösterdiği, değişken ve hata hakkındaki varsayımların sağlanması durumunda geçerlilik kazanır (Yıldırım, 2010). Varsayımların sağlanamaması durumunda hesaplamaların ve oluşturulan regresyon denklemlerinin istatistiksel olarak geçerliliği olmamaktadır. Çünkü, elde edilen regresyon denkleminde yapılacak tahminlerin tutarsız olma ihtimali artar (Şahinler, 1997).

Bağımsız değişkenler arasında bir veya daha fazla doğrusal bağıntı olması çoklu bağlantı (multicollinearity) sorununu oluşturur (Alpar, 2003). Bağımsız değişkenler arasındaki yüksek korelasyon ile ortaya çıkan çoklu bağlantı sorunu, çoklu regresyon modelleri için varsayımlarını ihlal eden önemli bir husustur. Çoklu bağlantı sorununu giderebilmek için en küçük kareler regresyonu yerine yanlı tahmin edicilere sahip olan “yanlı regresyon yöntemleri”nin kullanılması önerilmektedir.

Çoklu bağlantı sorunu ile karşılaşıldığında en küçük kareler regresyonu yerine önerilen yanlı regresyon yöntemlerinden biri “Ridge regresyon analizi”dir. Ridge regresyon analizi, hatayı en aza indirmeyi amaçlayan regresyon yöntemi olup, genellikle modeldeki iki yada daha fazla bağımsız değişken arasında yüksek dereceden ilişki olması durumunda kullanılır. Bu yöntem uygulanırken ilk olarak bağımsız değişkenler standartlaştırılır (Karadavut ve ark., 2005).

Yanlı regresyon yöntemlerinden bir diğeri “temel bileşenler regresyon analizi”dir. Temel bileşenler regresyonu; 'her doğrusal regresyon modelinin bir dik bağımsız değişkenler topluluğuna dayanarak yeniden açıklanabileceği' temeli üzerine oluşturulmuş bir yöntemdir. Oluşturulan bu yeni ve dik açıklayıcı (bağımsız) değişkenler, orijinal değişkenlerin doğrusal bileşeni olarak elde edilirler ve bağımsız değişkenlerin temel bileşenlerini adını alırlar (Alpar, 2003). Temel bileşenler regresyonunun amacı, veri indirgemesi ve veri tahminlemesini analiz etmektir (Foong, 2007).

Çoklu bağlantı sorunu ile baş etmek için kullanılan yanlı regresyon yöntemlerinden bir diğeri ise “kısmi en küçük kareler regresyon analizi”dir. Bu yöntemde, boyut indirgeme temel alınarak, kısmi en küçük kareler analizi ile çoklu doğrusal regresyon analizinin bir arada kullanılmasıyla oluşur. Kısmi en küçük kareler regresyonu, çoklu bağlantı sorununu ortadan kaldırmada değişkenlerin gözlem sayısından çok, gözlem sayısının değişken sayısından çok olduğu durumlarda kullanılabilen bir yöntemdir (Bulut ve Alın, 2009).

En küçük varyansa sahip olması sebebiyle yansız lineer tahmin ediciler arasında en küçük kareler regresyonu farklı yöntemler ile oluşturulmuş regresyon modellerinin karşılaştırılmasının en basit yoludur. Ancak bu durum en küçük kareler regresyonu ile elde edilen varyanstan daha küçük varyans elde edilemeyeceğini garanti etmez. Bu

sebeple HKO deęeri kk olan model daha tercih edilir (Coşkuntuncel, 2010). Regresyon modellerinin karşılaştırılmasında HKO'nun yanı sıra tahminlenen regresyon katsayılarının standart hataları, modele ait düzeltilmiş belirleme katsayısı ( $R^2$ -düz), ve Akaike Bilgi Kriteri (AIC) kullanılmaktadır. Bu performans göstergeleri için; her zaman standart hata, HKO ve AIC deęerlerinin kk, ( $R^2$ -düz) deęerinin ise yksek olması istenilmektedir (Shibata, 1981).

Bu tez alıřmasında, yanlı tahmin yntemleri olarak isimlendirilen Ridge regresyon analizi, temel bileşenler regresyon analizi, kısmi en kk kareler regresyon analizi yntemlerinin incelenmesi amaçlanmıştır. ncelikle oklu baęlantı sorunu ve tespit etme yntemleri aıklanmış sonrasında ise sz konusu yntemler aıklanmıştır. Yntemlerin oklu baęlantı sorunu karşılarındaki durumlarının incelenmesi amacıyla tavuk yumurta kalite parametreleri olarak adlandırılan deęişkenlerden oluřturulan gerek bir veri seti kullanılmıştır. Yntemlerin deęerlendirilmesinde tahminlenen katsayıların standart hatalarının yanı sıra Hata Kareler Ortalaması (HKO), modele ait düzeltilmiş belirleme katsayısı ( $R^2$ -düz) ve Akaike Bilgi Kriteri (AIC) kullanılmıştır.

## 2. GENEL BİLGİLER

### 2.1 En Küçük Kareler (OLS) Regresyonu

En küçük kareler varsayımları sağlandığında regresyon analizinde en yaygın kullanılan tahmin edicidir. En küçük kareler regresyonu, normallik, sabit varyanslılık, sapan değer içermeme gibi varsayımların sağlandığı durumlarda güvenilir tahminler elde eder (Neter, 1996; Fox, 1997; Alma ve Vupa, 2008).

En küçük kareler, bağımsız değişkenlerin birim değeri değişirken, bağımlı değişkenin birim değerinin varyansının sabit kalacağını varsaymaktadır (Gujarati, 1995; Orhunbilge, 2000; Tarı, 2006; Yamak ve Köseoğlu, 2006; Albayrak, 2008).

En küçük kareler regresyonunun amacı; hata terimlerinin kareleri toplamını en aza indirmeye yaparak, modeli en uygun konuma getirmektir (Eker, 2011).

$$Q_{(OLS)}(b) = \sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_i - \bar{y}_i)^2 \quad (2.1)$$

Yukarıdaki fonksiyonda en küçük kareler regresyonu sayesinde gerçek  $y_i$  değerleri ile tahmin edilen  $\bar{y}_i$  değerleri arasındaki farkın kareleri toplamı yani b katsayısı tahmin edilerek sonucun minimum olması hedeflenir.

Hata terimlerinin kareleri,

$$\sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_i - \bar{y}_i)^2 = \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i)^2 \quad (2.2)$$

ise bu ifadeyi en küçük yapan tahmin değerleri ( $\beta_0, \beta_1$ ),  $\beta_0$  ve  $\beta_1$ 'e göre türev alınıp sıfıra eşitlenerek bulunur.

$$\frac{\partial \sum e_i^2}{\partial \beta_0} = -2 \sum e_i = 0 \quad \text{ise} \quad \sum e_i = 0 \quad (2.2a)$$

$$\frac{\partial \sum e_i^2}{\partial \beta_1} = -2 \sum e_i x_i = 0 \quad \text{ise} \quad \sum e_i x_i = 0 \quad (2.2b)$$

$\beta_0, \beta_1$  : Regresyon katsayı değerleridir.



$\beta_0$ , regresyonun y eksenini kestiği noktayı ifade eder ve sabit değerdir.  $\beta_1$ , bağımsız değişkende bir birim değişiklik olduğunda bağımlı değişkendeki değişimi ifade eder (Tunç, 2018).

En küçük kareler regresyonunun özellikleri aşağıda belirtilmiştir:

1. Hata terimlerinin beklenen değeri (ortalaması) sıfırdır.

$$\sum e_i = 0 \text{ olduğu için } \sum e_i/n = 0 \text{ olur.}$$

2. Hata terimleri  $y_i$  tahminleri ile ilişkisizdir.

3. Hata terimleri  $x_i$  ile ilişkisizdir.

$$\sum e_i x_i = 0 \text{ dır.}$$

En küçük kareler regresyonu, varsayımlarının sağlanamaması durumunda yanlış tahminler yapmaktadır. Bu durumda anlamlılık testleri gerekliliğini yitirmektedir (Orhunbilge, 2000).

### 2.1.1 En Küçük Kareler (OLS) Regresyonunun Varsayımları

En küçük kareler regresyonunun varsayımları, gerçek regresyon modelinin ideal durumunu açıklar. En küçük karelerin iyi bir tahmin edici olması için varsayımları karşılanmalıdır. En küçük kareler regresyonunun varsayımlarını bilmek, sonuçların sapmasızlığını koruyabilmek açısından önemlidir. En küçük kareler regresyonunun varsayımları aşağıda belirtilmiştir:

**1. Regresyon modeli doğrusaldır.** Yani iki değişkenli modelde;  $y_i = \beta_0 + \beta_1 x_i$  iken; birden çok bağımsız değişken olan modelde;  $y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_p x_{pi}$  olarak ifade edilir. İki regresyon modelinde de katsayılar doğrusaldır fakat x ve y değişkenleri doğrusal olmayabilir.

**2. Hata terimlerinin ( $e_i$ ) ortalaması sıfırdır.** Yani iki değişkenli model de;  $\sum e_i/x_i = 0$  ve çok değişkenli modelde;  $E(e_i / x_{1i}, x_{2i}, x_{pi})$  olarak ifade edilir.

**3. Varyans sabittir.** Yani iki değişkenli modelde;

$$\text{Var}(e_i/x_i) = E[e_i - E(e_i)/x_i]^2 = E[e_i/x_i]^2 = \sigma^2 \quad (E(e_i) = 0 \text{ dır.})$$

Yukarıdaki denklem çok değişkenli model içinde geçerlidir.  $\sigma^2$  sabit sayıya eşittir.

Yani, her  $x_i$  değeri için y'nin varyansının sabit olduğunu gösterir.

**4. Hata terimlerinin arasında otokorelasyon yoktur.** En küçük karelerde  $x_i$  ve  $x_j$  gibi iki  $x$  değeri için  $e_i$  ve  $e_j$  arasında korelasyon sıfırdır. İki değişkenli model için geçerli olan çok değişkenli modelde de geçerlidir. Bu varsayımın sağlanamaması durumunda otokorelasyon sorunu oluşur. Otokorelasyon sorununu ortaya çıkaran durumlar;

- ✓ Bağımsız değişkenlerin modeli yeteri kadar açıklamaması,
- ✓ Gözlem sayısının yetersizliği,
- ✓ Modelin uygun olmaması,
- ✓ Bağımsız değişkenler arasında güçlü ilişki olması durumlarıdır.

Yanlış model seçimi yada alınması gereken bağımsız değişkenin modele dahil edilmemesinden kaynaklanan duruma 'sahte otokorelasyon' denir. Hata terimleri arasında gerçek ilişki olmasından kaynaklanan duruma ise 'gerçek otokorelasyon' denir (Albayrak, 2014).

Otokorelasyon sorununun tespiti grafik yöntemi, Durbin-Watson testleri ve Von-Neumann testleri ile yapılır (Tunç, 2018).

Otokorelasyon varlığında en küçük kareler regresyon analizi sonuçlarında varyans-kovaryans matrisi geçerliliği ortadan kalkar (Albayrak, 2014).

**5. Bağımsız değişkenler arasında kovaryans sıfırdır.** İki değişkenli model de;

$$\begin{aligned} \text{Cov}(e_i, x_i) &= E[e_i - E(e_i)][x_i - E(x_i)] & (2.3) \\ &= E[e_i (x_i - E(x_i))] \\ &= E[e_i x_i - e_i \cdot E(x_i)] \\ &= E(e_i x_i) - E(x_i) E(e_i) \\ &= E(e_i x_i) \\ &= 0 \end{aligned}$$

sonucu elde edilir. Yani, bağımsız değişkenler arasında anlamlı ilişki yoktur (Albayrak, 2014).

**6. X değerleri tekrarlanan örneklerde değişmez.** X değerleri rastgele seçilen değişkenler değildir. Çünkü örnek seçiminde her  $x$  değerine karşılık gelen  $y$

değerlerinden biri rastgele seçilir. Yani, bağımsız değişkenler sabit, bağımlı değişken rastgele olmalıdır.

**7. Gözlem sayısı, değişken sayısından büyük olmalıdır.**

**8. Model doğru tanımlanmalı, tüm değişkenler dikkate alınmalıdır.**

**9. Bağımsız değişkenler arasında ilişki olmamalıdır.** Bu varsayımın yerine gelmemesi durumunda çoklu bağlantı sorunu oluşmaktadır.

### **2.1.2 Gauss-Markov Teoremi**

Gauss-Markov teoremine göre, doğrusal regresyon analizinde varsayımlar sağlandığında en küçük kareler regresyon analizi en sapmasız, en yansız, doğrusal ve en küçük varyansa sahip tahmin edicidir (Kızılırmak, 2016).

En küçük kareler regresyonu bir bağımlı değişkene etki eden birden fazla bağımsız değişken olduğunda da yani çoklu regresyon analizinde de 'bağımsız değişkenler arasında çoklu bağlantı olmaması' durumunda tutarlı sonuçlar verir. Ancak çoklu doğrusal regresyon analizinde bağımsız değişkenler arasında doğrusal ilişki varsa çoklu bağlantı sorunu ortaya çıkar. Bu durumda bağımsız değişkenlerin korelasyonu ve standart hataları artar. Sonuçta, regresyon analizi yanlış yorumlanır (Kaşko, 2007).

### **2.2 Çoklu Bağlantı (Multicollinearity)**

Çoklu regresyon analizlerinde bazen bağımsız değişkenler arasında bir ilişki gündeme gelmektedir. Bağımsız değişkenler arasında bu şekilde doğrusal bir ilişkinin oluşması ile regresyon modelinin güvenilirliği azalır, regresyon denklemi ile ulaşılan sonuçlar gerçek değerlerinden farklılık gösterir (Kaşko, 2007). Oysa çoklu regresyon analizlerinde bağımsız değişkenler arasında ilişkili olmaması sağlanması gereken varsayımdır. Bu varsayımın sağlanmaması durumunda, yani bağımsız değişkenler arasında doğrusal ilişkinin oluşması çoklu bağlantı (multicollinearity) sorununu oluşturur (Alpar, 2003).

Bağımsız değişkenler arasında doğrusal ya da doğrusala yakın ilişki olmasına çoklu bağlantı denir. Çoklu bağlantının olması halinde bağımsız değişkenlerin bağımlı değişkenler üzerindeki etkisini değerlendirmek zor olabilmektedir (Pimentel ve ark., 2007; Eker, 2011). Yani regresyon katsayılarının varyans ve kovaryansları artmakta,  $R^2$  değeri yüksek olmasına rağmen t testi sonuçları güvenli olmamaktadır (Gujarati,

1995). Bağımsız değişkenler arasındaki korelasyonun 0.80'in üzerinde olması çoklu bağlantı sorununun oluştuğunu gösterir (Garson, 2006; Kaşko, 2007).

### 2.2.1 Çoklu Bağlantının Nedenleri

Çoklu bağlantının nedenleri aşağıdaki gibi sıralanabilir;

- ✓ Örneğin popülasyonu yeteri kadar temsil etmemesi,
- ✓ Bağımsız değişkenler arasında gerçekten ilişki olması,
- ✓ Gözlem sayısının değişken sayısından küçük olması ( $n < p+1$ ),
- ✓ Bir değişken grubunun indirgenmesi ile yapılan yanlışlık,
- ✓ Araştırmacıdan kaynaklanan hataların oluşması durumlarıdır (Alpar, 2003; Kaşko, 2007).

### 2.2.2 Çoklu Bağlantıyı Belirleme Yöntemleri

Çoklu bağlantıyı oluşturan bir çok sebep olabilir. En başta iki değişken arasındaki korelasyon katsayısının 1'e yakın olması (0.80'den büyükse) çoklu bağlantıyı gösterir.

Bir başka sebep, regresyon katsayılarının büyüklüğü ve işaretidir. Regresyon katsayılarına ait regresyon modelindeki katsayıların büyüklüğü ve işaretlerin beklenenin tersi olması çoklu bağlantı oluşmasına sebep olur.

Bağımsız değişkenin regresyon modeline eklenmesi ya da regresyon modelinden çıkarılması ile regresyon katsayılarında istenilmeyen değişiklik meydana gelir. Bu da çoklu bağlantıya sebep olur (Pamukçu, 2010).

Çoklu bağlantının belirlenmesi kadar derecesinin belirlenmesi de büyük önem taşımaktadır (Kaşko, 2007). Çoklu bağlantının derecesinin ölçütleri;

**Bağımsız değişkenlere ilişkin korelasyon katsayılarının belirlenmesi:** İki değişken arasındaki basit korelasyon katsayısı anlamlı fakat kısmi korelasyon katsayısı anlamsız ise çoklu bağlantı sorunundan bahsedilir. Fakat bu yöntem her zaman güvenilir sonuçlar vermeyebilir. Yani, kısmi korelasyon katsayılarının yüksek olması durumunda da çoklu bağlantı oluşabilir (Pamukçu, 2010).

**Korelasyon matrisiyle belirlenmesi:** Çoklu bağlantının tespit edilmesinde uygulanması en kolay yöntemdir (Pamukçu, 2010).

Korelasyon matrisi ile çoklu bağlantı belirlenirken bağımsız değişkenler;

$$s_{ij} = \frac{(x_{ij} - \bar{x}_j)}{\sqrt{\sum_{j=1}^p (x_{ij} - \bar{x}_j)^2}} \quad (2.4)$$

Denklem 2.4 ile  $X'X$  standartlaştırılır.  $X'X$  standartlaştırılmış korelasyon matrisindeki köşegen dışı  $r_{ij}$ 'nin elemanları kontrol edilir.  $r_{ij}$ ;  $x_i$  ve  $x_j$  arasındaki açının kosinüsüdür (Farrar ve Glauber, 1967; Pamukçu, 2010).  $|r_{ij}|$  1 veya 1'e yakınsa çoklu bağlantı sorunundan bahsedilir (Pamukçu, 2010). Bağımsız değişkenler arasındaki korelasyon katsayısı yüksek ( $r > 0.80$ ) ise, çoklu bağlantı probleminin varlığından bahsedilir (Neter ve ark., 1990; Gujarati, 1995; Eker, 2011).

**Tolerans değeri ve varyans büyütme faktörü:** Bağımsız değişkenlerin korelasyon matrisinin tersinin köşegen öğelerine varyans büyütme değerleri denir. Varyans büyütme faktörü ile çoklu bağlantı probleminden dolayı varyansların gerçek değerlerden ne kadar uzaklaştığı tespit edilir.  $VIF_i$  şeklinde gösterilir.  $VIF_i$  değerleri tolerans değeri ile ilgili olarak eşitlik 2.5'teki gibi hesaplanır.

$$VIF_i = 1 / (1 - R_i^2) \quad (2.5)$$

Varyans büyütme faktörü değerlerinin 10'un üzerinde olmasının sebebi güçlü çoklu bağlantıdır. Böyle bir durumda regresyon katsayıları gözden geçirilmelidir (Albayrak, 2005; Meko, 2006; Kaşko, 2007).  $VIF_i$ 'ler iki ve daha fazla çoklu bağlantının varlığını belirlemede kullanılabilir en iyi yöntemdir (Alpar, 2003).

Çoklu bağlantının belirlenmesinde diğer yöntem tolerans değerlerini hesaplamaktır. Tolerans değeri,

$$T = 1 - R_j^2 \quad (2.6)$$

şeklinde hesaplanır (Pamukçu, 2010). Varyans büyütme faktörü ve tolerans değeri denklemleri incelendiğinde, varyans büyütme faktörü değerinin büyük, tolerans değerinin küçük çıktığı durumlarda çoklu bağlantı görülür.

**Özdeğer ve özvektörlerin incelenmesi:** Özvektörler, değişkenlerin doğrusal bileşenleri olup, değişken sayısı kadardır. Özdeğerler ise özvektörler aracılığıyla belirlenen varyansı ifade eder (Kaşko, 2007). Çoklu bağlantı problemi yoksa özdeğerler 1'e eşittir. Özdeğerlerden en az birinin 1'den farklı olması yada en az birinin 0'a yakın olması çoklu bağlantı problemini oluşturur. Fakat bütün özdeğerlerin tek tek incelenmesi zor olduğu için bunun yerine Virod ve Ulah, (1981) yapmış oldukları çalışmada en büyük özdeğer ile en küçük özdeğere dayalı koşul indeksi önermişlerdir. Özdeğerler

$\lambda_{\max} = \lambda_1 > \lambda_2 > \lambda_3 > \dots > \lambda_p = \lambda_{\min}$  şeklinde gösterilirse,

Koşul İndeksi (KI),

$$KI = \sqrt{\frac{\lambda_{\max}}{\lambda_{\min}}} \quad (2.7)$$

şeklinde hesaplanır.

$KI < 10$  ise çoklu bağlantı problemi önemli seviyede değildir.

$10 < KI < 30$  ise çoklu bağlantı orta seviyededir.

$30 < KI$  ise çoklu bağlantı şiddetlidir ve birden fazla çoklu bağlantı vardır (Pagel ve Lunneborg, 1985; Gujarati, 1995; Rathert ve ark., 2011).

**F ve t testlerinin karşılaştırılması:** F ve t testleri regresyon modelinin anlamlılığı hakkında bilgi verir. Regresyon katsayılarına ait t testi değerlerinin tamamı anlamsız iken, F testinin anlamlı çıkması çoklu bağlantının varlığını belirtir. Ancak çoklu bağlantının önemli düzeyde olduğu durumlarda bir çok veri bu şekilde tutarlı sonuç sergileyemediği için kuşkulu bir yöntemdir (Pamukçu, 2010).

### 2.2.3 Çoklu Bağlantı Sorununun Ortadan Kaldırılması

Çoklu bağlantı probleminin ortadan kaldırılması çoklu bağlantının oluşma sebebine, derecesine, regresyonun kullanım amacına ve çoklu bağlantı sebebi etmenlerin önemine gibi faktörlere bağlıdır (Canküyer ve Sönmez, 1996; Kaşko, 2007). Çoklu bağlantı probleminin çözümünde en etkin yol; modeldeki değişkenleri çıkarmadan regresyon katsayılarını yanlış tahmin etmektir (Kaşko, 2007). Çoklu bağlantının giderilmesi için araştırmacılar önerilerde bulunmuşlardır (Gujarati, 1995; Neter, 1990; Albayrak, 2005). Bunlar;

**a) Örnek genişliğinin artırılması:** Gözlem sayısı artırılarak varyansların küçülmesi sağlanabilir.

**b) Bağımsız değişkenlerin regresyon modelinden çıkarılması:** Aşamalı olarak daha az etkili olan bağımsız değişkenler çıkarılabilir. Eğer çoklu bağlantı sorunu örneklemeden kaynaklanıyorsa bu yol kullanılmamalıdır. Çünkü gerçekte önemli olan bir değişken başka değişkenlerle yapay ilişkili gözüküp yanlış model oluşmasına sebep olabilir (Kaşko, 2007).

**c) Bağımsız değişkenlerin birleştirilmesi:** Birbiriyle bağlantısı bulunan bağımsız değişkenler birleştirilerek tek bir değişken elde edilir ve regresyon modeline tek bir değişken olarak konulabilir. Çok tercih edilmeyen bir yöntemdir.

**d) Yanlı tahmin yöntemlerinin kullanılması:** En küçük kareler regresyonu bağımsız değişkenler arasındaki ilişki olmaması durumunda en küçük varyanslı istenilen sonuçları vermektedir. Fakat bağımsız değişkenler arasında ilişki söz konusu olduğunda en küçük kareler güvenilir sonuçlar vermemektedir. Yanlı tahmin yöntemlerinin kullanılmasındaki amaç, yanlılık terimi kullanarak küçük varyanslı tahmin edici bulmaktır (Kaşko, 2007).

Yanlı tahmin yöntemleri arasında en yaygın olarak kullanılan yöntemler; Ridge regresyon analizi, temel bileşenler regresyon analizi, kısmi en küçük kareler regresyon analizidir.

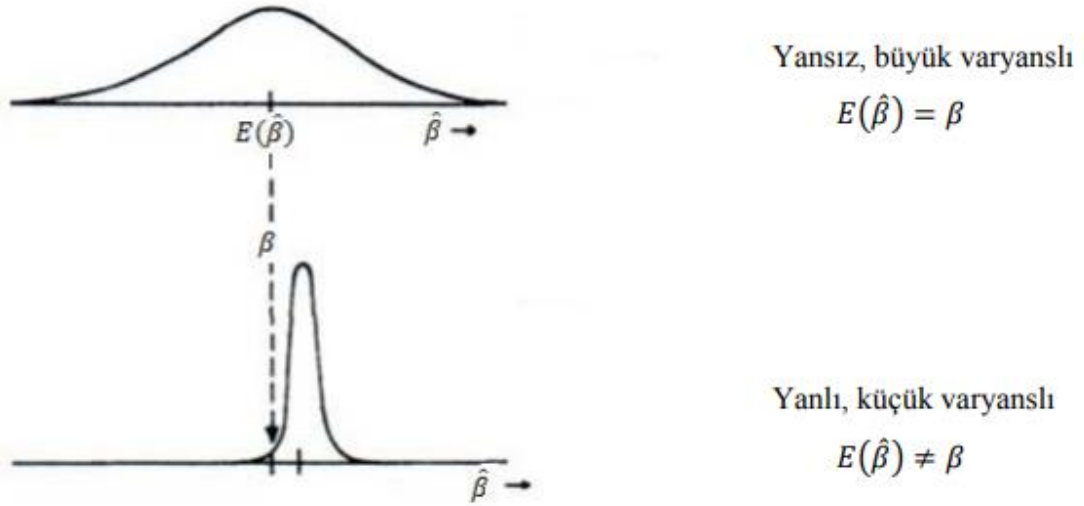
### **2.3 Ridge Regresyon (RR)**

Hoerl ve Kennard, (1970) çoklu bağlantı problemi ile karşılaşıldığında en küçük kareler regresyonuna alternatif olarak Ridge regresyon analizini önermişlerdir. Ridge regresyon, çoklu bağlantı sorunu varlığında en küçük kareler regresyonundan daha küçük varyanslı parametre tahminlerinin elde edilmesini ve gereksiz değişkenlerin modelden çıkarılmasını sağlar (Karadavut ve ark., 2005). Bağımsız değişkenlerin birbirleri üzerindeki etkilerini en aza indirmeyi ve doğru tahminler yapmayı hedefler (Üçkardeş ve ark., 2012). Hoerl ve Kennard, (1970) tarafından geliştirilen Ridge regresyon analizinin amaçları;

1. Çoklu bağlantı sorunu oluştuğunda katsayılar da meydana gelen değişikliklerin grafik ile gösterilmesi,
2. Bağımsız değişkenler arasında ilişki gözlemlendiğinde en küçük kareler regresyonundan daha küçük varyanslı sonuçlar olması,
3. Regresyon modelinde gereksiz olduğu belirlenen değişkenlerin çıkarılması olarak belirlenmiştir (Polat, 2010).

Ridge regresyon analizi en küçük kareler regresyonunun basamaklarının tekrarlanması ile oluşur. Fakat Ridge regresyon analizinin en küçük karelerden farkı,  $k$  Ridge parametresinin ilave edilmesidir (Yıldırım, 2010). Ridge regresyon  $k$  yanlılık sabitinin

eklenmesi ile yanlış tahminler yapmasına karşın, tahminlerin varyanslarını azaltmaktadır. Şekil 2.1’de tahmin ediciye ait varyans ve yanlışlık gösterilmiştir (Pamukçu, 2010).



**Şekil 2.1** Bir Tahmin Edicide Varyans ve Yanlılık

En küçük karelerin regresyon modeli,

$$Y = \beta X + e \quad (2.8)$$

şeklindedir. Burada;

---

Y	: Bağımlı değişkeni,
$\beta$	: Regresyon doğrusunun Y eksenini kestiği noktayı,
X	: Bağımsız değişkeni,
e	: Tesadüfi hatayı ifade eder.

---

Gerekli dönüşümler yapıldıktan sonra regresyon denklemi,

$$\beta = (X'X)^{-1} X'Y \quad (2.9)$$

şeklinde bulunur.

Çoklu bağlantı sorunu oluştuğunda bağımsız değişkenler arasındaki ilişkinin yüksek olması  $X'X$  matrisinin varyansını büyütür.  $X'X$  matrisinin büyümesi sonucunda önemli olan parametre değerleri önemsiz çıkar. Bu sorunu ortadan kaldırmak için denklemde yer alan  $X'X$  matrisinin köşegen elemanlarına pozitif k sabiti eklenerek bu matrisin varyansının küçültülmesi hedeflenir (Hoerl ve Kennard, 1970).



Ridge regresyon için parametre denklemi 2.10 numaralı eşitlikteki gibi elde edilir.

$$\beta^* = (X'X + kI)^{-1}X'Y \quad 0 \leq k \leq 1 \quad (2.10)$$

şeklinde elde edilir (Hoerl, 1962).

### 2.3.1 Ridge Regresyon ile En Küçük Kareler Regresyonunun İlişkisi

En küçük kareler tahmincisine ait denklemin:

$\beta = (X'X)^{-1} X'Y$  şeklinde hesaplandığı daha önce belirtilmiştir. Bu denklem,

$X'X \beta = X'Y$  şeklinde de yazılabilir.

Ridge regresyon parametre denkleminde  $X'Y$  yerine  $X'X \beta$  yazıldığında;

$$\beta^* = (X'X + kI)^{-1}X'X \beta \quad (2.11)$$

denkleminde elde edilir.

$X'X$  matrisinin tersinin tersi kendisine eşit olduğundan, Ridge regresyon denklemi:

$$\beta^* = [I + k(X'X)^{-1}]^{-1} \beta \quad (2.12)$$

şeklini alır.

$$Z = [I + k(X'X)^{-1}]^{-1} \text{ olarak tanımlanırsa,} \quad (2.13)$$

$$\beta^* = Z \beta \quad (2.14)$$

haline dönüşür. Ridge regresyon en küçük kareler regresyonunun basamaklarının tekrarı ile elde edilmiştir (Kurtuluş, 2001).

### 2.3.2 Ridge Regresyonun Varyansı

Ridge regresyon analizine ait varyans denklemi:

$$V(\beta^*) = V(X'X + kI)^{-1}X'Y \quad (2.15)$$

$$= \sigma^2 (X'X + kI)^{-1} X'X (X'X + kI)^{-1} \quad (2.16)$$

şeklindedir.

### 2.3.3 Ridge Regresyon Hata Kareler Ortalaması

Hata kareler ortalaması (HKO),  $\beta^*$  dan  $\beta$ 'ya olan uzaklıktır (Kurtuluş, 2001).

$$HKO(\beta^*) = E[\beta^* - \beta]^2 = V[\beta^* + E(\beta^* - \beta)] \quad (2.17)$$

yada;

$$HKO(\beta^*) = V(\beta^*) + (\beta^*)' \text{nin yanlılığı}^2 \text{ dir.} \quad (2.18)$$

$$\text{Yanlılık } (\beta^*) = E[(\beta^*)] - \beta = k \beta' (X'X + kI)^{-1} \beta \quad (2.19)$$

şeklinde hesaplanır. Daha önce belirlenen Ridge regresyon varyans denklemi ile yanlılık denklemi birleştirildiğinde:

$$HKO = \sigma^2 \text{ İz}[(X'X + kI)^{-1} X'X(X'X + kI)^{-1}] + k^2 \beta' (X'X + kI)^{-2} \beta \quad (2.20)$$

denklemi ile Ridge regresyona ait hata kareler ortalaması bulunur.

Denklemden yer alan İz; bir kare matrisinin köşegen üstündeki değerlerinin toplamıdır (Pamukçu, 2010; Kuvat, 2018).

### 2.3.4 Ridge Regresyonda k Parametresinin Tespit Edilmesi

Ridge regresyon analizi için k seçimi önemlidir. Yanlılığı en aza indiren ve hata kareler ortalamasını azaltan k sabitini bulmak için birçok yöntem geliştirilmiştir (Sinan, 2003). Bunlardan bazıları aşağıdaki gibidir;

- ✓ Ridge izi yöntemi
- ✓ Hoerl ve Kennard yöntemi
- ✓ Lee ve Campbell yöntemi
- ✓ McDonald ve Galarneau yöntemi
- ✓ Lawless ve Wong yöntemi

Hoerl ve Kennard, k sabitinin tek başına anlamlı olmadığını fakat çoklu bağlantı durumunda en küçük karelerden daha tutarlı sonuç verdiğini belirtmişlerdir. Ridge parametresi k sabitinin seçimi için en uygun yöntemin Ridge izi olduğunu ifade etmişlerdir (Kuvat, 2018).

Anderson, (1998) Ridge regresyon analizindeki k sabitinin belirlenmesinde 'varyans büyütme faktörü değerlerinin 1'e yaklaşımını belirleyen değer alınabilir' demiştir.

Ridge regresyon parametresi k sabitinin belirlenmesi özdeğerler yöntemine de dayanır. k sabitinin belirlenmesi ile analizin durağanlaştığı noktalar veya özdeğerin 1'e en yakın olduğu nokta bulunur. Kurtuluş (2001) tarafından önerilen k sabitinin belirlenmesinde koşul indeksinden faydalanılarak,

$$k \leq \frac{\lambda_{\max} - 100\lambda_{\min}}{99} \quad k \neq 0 \quad (2.21)$$

denklemini bulunmuştur. Bu eşitlik ile k sabitinin VIF değerinin 1'e en yakın olduğu nokta belirlenir (Anderson, 1998). k sabitinin 1'e yaklaştığı tahminlerin yanlılığının artmasına rağmen varyansları azalır.

### 2.3.5 Ridge İzi Yöntemi

Ridge izi, çoklu doğrusal bağlantıyı açıkça görmek ve regresyon katsayılarının kararsızlıklarını belirlemek için başvurulan grafiksel yöntemdir (Sinan, 2003). Ridge izi grafiğinde, Ridge regresyon katsayıları düşey ekseninde, k değerleri yatay ekseninde bulunur. Ridge izi grafiğinde k değerleri arttıkça, Ridge regresyon katsayılarının hata kareler ortalaması azalır. En küçük kareler tahmin edicisinin regresyon katsayılarının hata kareler ortalamasından daha küçük sonuçlar veren Ridge tahmin edicisi bulmak için k değerindeki artışa rağmen regresyon katsayılarının durağanlaştığı noktadaki k değeri Ridge parametresi olarak seçilir (Hoerl ve Kennard, 1970).

k değeri seçilirken katsayılar ve hata kareler toplamı istenmeyen büyüklüğe ulaşmamalıdır. Ayrıca en küçük kareler regresyon analizi ile işaret değışen katsayılar k parametresi ile işaret değıştirerek uygun hale getirilir (Sinan, 2003).

k parametresi için belirlenen aralık:

$$m = y - \sum_{i=1}^n \lambda_i (\lambda_i + k)^{-1} \quad (2.22)$$

şeklinde bulunur. Burada:

---

y : Bağımlı değışken sayısı,  
k : Ridge parametresidir.

---

Yukarıdaki denklem 2.22'ye göre; k=0 için m=0 dır. k→∞ m→n dir. Yani m, k parametresinin sınırlarının [0, ∞) aralığından, [0, n] aralığına girmesini sağlar. Bu şekilde Ridge izi belirlenir (Sinan, 2003).

Ridge izi kolay ve anlaşılır olmasından dolayı kullanımı yaygın olan yöntemdir.

### 2.4 Temel Bileşenler Regresyonu (PCR)

Temel bileşenler regresyonunun geçmişı, Beltrami, (1873) ve Jordan, (1874) tarafından bağımsız elde edilen "Tekil Değer Ayrışımı"na kadar uzanmaktadır (Stewart, 1993; Pamukçu, 2015). Günümüzde kullanılan temel bileşenler regresyon

analizi Pearson, (1901) tarafından bulunmuştur (Pamukçu, 2015). Temel bileşenler regresyon analizinin tekniğinin bilgisayara uygulanmasını ve yorumlanmasını Cooley ve Lohnes, (1971) yapmışlardır (Pamukçu, 2015).

Temel bileşenler regresyonunun kullanımı, bağımsız değişkenlerin sayısının gözlem sayısından büyük olması ve bağımsız değişkenlerin arasında yüksek ilişki olması durumlarında avantajlıdır. Çünkü bu gibi durumlarda, temel bileşenler regresyonu açıklayıcı değişkenlerin sayısını indirgeyebilir ve regresyon katsayıları tahminini daha doğru hale getirilebilir (Rencher, 2002). Temel bileşenler regresyon analizi, çoklu bağlantı sorunu ile karşılaşıldığında, her doğrusal regresyon modelinin bir dik açıklayıcı değişkenler kümesi ile yeniden açıklanması esas alır (Özkan, 2009).

Temel bileşenler regresyonunda amaç, çok sayıda birarada bulunan değişkenlerin boyut indirgeyerek anlamlı daha az sayıda değişkene dönüştürülmesidir. Bu indirgeme sonunda aralarında korelasyon olmayan bağımsız değişkenler elde edilir (Alkan, 2008). Boyut indirgeme yöntemi olan temel bileşenler regresyon analizinde elde edilebilecek maksimum değişken sayısı orjinal değişken sayısı kadardır. Temel bileşenler regresyonu, değişkenler arasındaki varyansı maksimum açıklayabilecek yeni değişkenler üretmeyi hedefler (Karakuş, 2011).

#### **2.4.1 Özdeğerler ve Özvektörler**

Matrisle ifade edilen sayılar ve vektörlerdir (Pamukçu, 2015).  $A$ ,  $p \times p$  boyutlu bir matris olsun.  $\lambda$  skaler (sayı ve birimle belirtilen),  $x$  vektörel (büyüklüğü, yönü ve doğrultusu olan) büyüklükler olmak üzere ( $x \neq 0$ ):

$Ax = \lambda x$  eşitliği sağlanıyorsa  $x$ ,  $A$  matrisinin özvektörü,  $\lambda$  matrisinin özdeğeridir. Yani;  $x$ , her bir  $\lambda$  özdeğerlerine karşılık gelen özvektördür.

Özdeğerler ve özvektörler, bir matrisi analiz ederken kullanılan özayrışımı imkanı sağlarlar. Özayrışımı, matrislerin içindeki bulunduğu fonksiyonların en büyük ve en küçük değerlerini bulmak açısından önemlidir. Temel bileşenler regresyon analizi, bir korelasyon veya kovaryans özayrışımından bulunmaktadır (Pamukçu, 2015).

Özdeğer ve özvektörlere ait özellikler aşağıda sıralanmıştır:

- ✓ Özdeğerler sıfır değerini alırken, özvektörler sıfır olamaz.
- ✓ Özdeğer sıfır olduğunda ( $\lambda=0$ ),  $A$  matrisinin ( $p \times p$ ) tersi alınamaz.

Bir A matrisinin ( $p \times p$ ) tersinin alınabilmesi için tüm özdeğerler sıfırdan farklı olmalıdır.

Bir A matrisinin ( $p \times p$ ) her özdeğerine karşılık gelen özvektörlerin oluşturduğu kümeye **öz uzay** denir.

**Özdeğerlerin sayısının bulunması:**  $\lambda$ ,  $p \times p$  boyutlu A matrisinin özdeğeri ve x özvektörü olmak üzere:

$x \neq 0$ ,  $\det(A - \lambda I_p) = 0$  ise A matrisinin en fazla p tane farklı özdeğeri vardır. Yani, A matrisinin özdeğerini bulabilmek için  $(A - \lambda I_p)x = 0$  denkleminin çözüm kümesi sonsuz olmalıdır.

**Özdeğerlerin bulunması:** Bir  $p \times p$  boyutlu A matrisinin özdeğerleri aşağıdaki basamaklar takip edilerek bulunur:

- ✓  $\det(A - \lambda I_p) = 0$  hesaplanır.
- ✓  $\det(A - \lambda I_p) = 0$  için denklemin kökleri bulunur. p tane olduğu için A matrisinin özdeğerleri p tanedir.
- ✓ Her özdeğer için  $(A - \lambda I_p)x = 0$  denkleminde özvektör bulunur.

**Köşegen matrisi:** Boyutu  $p \times p$  olan A matrisinin  $p \times 1$  boyutlu, p tane özvektörleri  $x_1, x_2, \dots, x_p$ 'dir. Özvektör matrisi (P), bu özvektörleri sütunlarında bulunduran matristir.  $P^{-1}AP$  ise özdeğer matrisidir.

$$P = (x_1, x_2, \dots, x_p)$$

$$P^{-1}AP = \Lambda = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_p \end{bmatrix} \quad (2.23)$$

Bir matrisin köşegeninin bulunabilmesi için özvektörleri doğrusal bağımsız olmalıdır (Pamukçu, 2015).

#### 2.4.2 Temel Bileşenlerin Belirlenmesi

Temel bileşenler regresyon analizinde ilk olarak değişkenler standardize edilmektedir. Bunun için hem bağımlı hem bağımsız değişkenlerin kendi ortalamalarından farkları alınır ve standart sapmalarına bölünür. Sonrasında yapılacak tüm hesaplamalar bu standardize veriler ile yapılır. Temel bileşenler regresyonunda bağımsız (açıklayıcı)

değişkenler temel bileşenlerine dönüştürülür ve aşağıdaki eşitlik elde edilir (Hintze, 2007; Topal ve ark., 2010).

$$X'X = PDP' = Z'Z \quad (2.24)$$

---

$X'X$  : Bağımsız değişkenlerin korelasyon matrisini,  
 $D$  :  $X'X$ 'in özdeğerlerinin bir köşegen matrisi (Temel bileşenleri tanımlar)  
 $P$  :  $X'X$ 'in özvektör matrisini,  
 $Z$  : Veri matrisini ifade eder (Hintze, 2007; Topal ve ark., 2010).

---

Temel Bileşenler (P) ortogonal olduğundan  $P'P=I$  dir.

$X_1, X_2, \dots, X_n$  gibi bağımsız X değişkenlerinin ağırlıklı ortalamaları  $Z_1, Z_2, \dots, Z_n$  gibi değişkenlere dönüştürülür. Bu değişkenler temel bileşen olduğu için aralarındaki korelasyon sıfırdır. Güçlü çoklu bağlantı problemine sebep olan özdeğeri küçük temel bileşenler hesaplanmadan çıkarılır. Bu durumda Y'nin regresyon analizi yapıldığında çoklu bağlantı problemi ortadan kaldırılmış olur. X'e geri dönüştürülen sonuçlar ile regresyon katsayıları tahmin edilir. Regresyon katsayıları tahmini aşağıdaki gibi;

$$A=(Z'Z)^{-1}Z'Y = D^{-1}Z'Y \quad (2.25)$$

bulunur. Bu eşitlik en küçük kareler regresyonudur. A ve  $\beta$  ile iki regresyon modeli ilişkilendirilebilir;

$$A=P'\beta \quad (2.26a)$$

$$\beta=PA \quad (2.26b)$$

A'nın sıfıra eşit ögesi temel bileşenler analizinden çıkarılabilir. Temel bileşenler analizinin adımları şu şekildedir;

- ✓ X matrisi için temel bileşenler regresyon analizi uygulanır ve temel bileşenler (Z) elde edilir.
- ✓ A en küçük kare tahmini için Z üzerine Y'nin regresyon analizi uygulanır.
- ✓ A'nın en son verisi sıfıra eşitlenir ve  $\beta=PA$  eşitliği ile hesaplanan katsayılar orijinal ölçeğine dönüştürülür (Albayrak, 2005; Hintze, 2007; Topal ve ark., 2010).

Temel bileşenler regresyon analizinde sonuçlar Ridge regresyon analizinde olduğu gibi yanlı olur fakat varyansın küçültülmesinden dolayı yanlılıktaki büyüklük dengelenir. Yani temel bileşenler regresyon analizinin hata kareler ortalaması en küçük kareler regresyonuna göre daha küçüktür (Hintze, 2007; Topal ve ark., 2010).

Ridge regresyon analizinde belirleyici olan k yanlılık sabitinin seçiminde yaşanan belirsizlikten, temel bileşenler regresyon analizinde modelden çıkarılacak temel bileşenlerin sayısı daha tutarlı sonuçlar verir. Yani çoklu doğrusal bağlantıyı ortadan kaldırmak temel bileşenler regresyon analizindeki özdeğeri küçük olan temel bileşenlerin analizden çıkarılması daha kesin sonuçlar verir (Albayrak, 2005; Al-Hassan ve Al-Kassab, 2009; Aswani ve ark., 2011).

### **2.4.3 Temel Bileşen Sayısının Belirlenmesi**

Temel bileşenler regresyonunda, bileşen sayısının belirlenmesi en önemli husustur. Temel bileşenler belirlendikten sonra bileşen sayısı bazı yöntemler ile belirlenir. Temel bileşen sayısını belirlemek için kullanılan yöntemler aşağıdaki gibi sıralanabilir;

1. En basit ve kullanışlı yöntem, toplam bileşen sayısının açıklanan toplam varyasyonun 2/3'ünden küçük olan özdeğerlerin toplanması şeklindedir. Rencher, (2002) açıklanan toplam varyasyon oranını % 80 olarak bildirmiştir (Eker, 2011).
2. Jolliffe Kriterine göre, özdeğeri 0.70 ve daha büyük olan temel bileşenler alınmalıdır.
3. Cattell, (1966) tarafından önerilen grafik yönteminde temel bileşenlerin özdeğerlerinin grafikleri çizilir. Grafik incelemesinde, tüm bileşenlerin durağanlaşmaya başladığı noktadan itibaren temel bileşenler alınır.
4. Temel bileşen sayısının belirlendiği bir diğer yöntem; özdeğeri birden büyük olan temel bileşenlerin belirlenmesidir (Küçükönder ve ark., 2004; Çamdeviren ve ark., 2005).

### **2.5 Kısmi En Küçük Kareler Regresyonu (PLS)**

Kısmi en küçük kareler regresyonu, kısmi en küçük kareler analizi ile çoklu doğrusal regresyon analizinden oluşan istatistiksel bir yöntemdir (Bulut ve Alın, 2009). Kısmi en küçük kareler regresyonu ilk olarak Herman Wold tarafından 1966 yılında ekonomik ve sosyal olayları modellemek için kullanılmıştır. Kemometri literatüründe önemli bir yer edinmiştir. Kowalski tarafından 1979 yılında kimya alanında kullanılmaya başlanmıştır. Kısmi en küçük kareler regresyonu istatistiksel özellikleri, Wold (1984), Manne (1987), Naes ve Martens (1986), Helland (1988), Stone ve Brook (1990) gibi istatistikçi ve matematikçilerle anlamlandırılmıştır (Bulut, 2011). 1980

yılından günümüze kadar birbiriyle doğrusal ilişkili X ve Y değişkenlerinden oluşan kısmi en küçük kareler regresyonu ile regresyon analizinin uygulanmasının zor olduğu karmaşık veri setlerinin çözümlenir (Polat, 2009).

Temel bileşenler regresyonu ile kısmi en küçük kareler regresyonu benzer yapıdadır. İki yöntemde de açıklayıcı değişkenler yeni bileşenlere dönüştürülür. Elde edilen bu yeni bileşenlerle bağımlı değişkene en küçük kareler yöntemi analizi yapılır. Fakat bu iki yöntemin farkı; temel bileşenler regresyonunda bileşenler bağımsız değişkenlerden elde edilirken, kısmi en küçük kareler regresyonunda bağımsız değişkenlerle beraber bağımlı değişken ile maksimum kovaryans oluşturacak bileşenler üretmektir (Naes ve Martens, 1986; Bulut, 2011).

Kısmi en küçük kareler regresyonu ilişkili verilere karşı duyarsızdır ve çok sayıda bağımsız değişken için işlem yapabilir (Bulut, 2011).

Model oluşturulurken daha az sayıda bileşen hesaplanır ve modelin yeterliliğine bileşen sayısı karar verir. Modele daha fazla bileşen eklenmesi modelin tahmin yeteneğinin artmasını sağlar (Bulut, 2011).

Kısmi en küçük kareler regresyonunda bileşenler y bağımlı değişkeni ile yüksek düzeyde ilişkilidir. Bu sebeple bileşen sayısı temel bileşenler regresyonuna göre daha azdır. Kısmi en küçük kareler regresyon modeli, temel bileşenler regresyon modeline göre daha az durağandır. Yani, daha az bileşen kullanıldığı için daha az varyans açıklanır (Bulut, 2011).

Kısmi en küçük kareler regresyonu algoritmalarla hesaplanır. Algoritmaların kullanılmasının amacı, hesaplama hızını artırmaktır. Algoritmalarla varyans-kovaryans matrisleri ve ilişki matrisleri hesaplanmaktadır (Lindgren ve Rännar, 1998).

Kısmi en küçük kareler regresyonunda kullanılan algoritmaların tarihsel gelişimi aşağıdaki gibidir (Kondylis, 2006).

Ortogonal skorlar kısmi en küçük kareler regresyonu (Wold ve ark., 1984).

Ortogonal yükler kısmi en kareler regresyonu (Martens, 1985).

Özvektör algoritması (Höskuldsson, 1988).

Helland algoritması (Helland, 1988).



SIMPLS algoritması (De Jong, 1993).

SAMPLS algoritması (Bush ve Nachbar, 1993).

Çekirdek algoritması (Rännar, 1994).

UNIPALS algoritması (Glen, 1998).

### 2.5.1 Doğrusal Olmayan Yinelemeli En Küçük Kareler Algoritması (NIPALS)

Klasik kısmi en küçük kareler algoritması olarak bilinen bu algoritma 1966 yılında Herman Wold tarafından geliştirilmiştir. Wold 1975 yılında kendi geliştirdiği bu algoritmaya katkıda bulunmuştur. Bu algoritmanın amacı, kovaryans matrisini maksimum yapacak bileşenleri bulmaktır. Algoritmada bileşenler aynı anda elde edilemeyeceğinden dolayı bileşenler için yineleme yöntemi uygulanır. Her aşamada tek bir bileşen ve bu bileşene ait ağırlık ve yük değerleri hesaplanır. Eğer bağımlı değişken bir tane ise yineleme yöntemine gerek kalmaz. Algoritma istenilen bileşen sayısı elde edilince ya da X matrisi sıfırlanınca sonlanır (Polat, 2009).

NIPALS algoritmasında,  $N \times K$  boyutlu X bağımsız (açıklayıcı) değişkenler matrisi ve  $N \times P$  boyutlu Y bağımlı değişkenler matrisidir (Bulut ve Alın, 2009).

Burada,

---

K : Bağımsız değişken sayısını,  
P : Bağımlı değişken sayısını verir.

---

Algoritmada a bileşen sayısı olup  $a = 1, 2, \dots, A$  dır. İlk adımda orijinal matrisler

$(X_1 = X, Y_1 = Y)$ ' nin kullanıldığı algoritmanın basamakları aşağıdaki gibidir.

1. Çok sayıda bağımlı değişken var ise, bu değişkenlerden oluşan Y matrisinin en yüksek varyansa sahip olan sütunu ya da ilk sütunu, tek bağımlı değişken var ise o değişken sütununa eşit olur.  $\mathbf{u}_a = \mathbf{y}'$  dir.
2. X'in Y'nin ilgili bileşeni  $\mathbf{u}_a$  üzerine regresyonundan X ve u arasındaki kovaryansı maksimum yapan w ağırlık vektörü  $\mathbf{w}_a = \mathbf{X}'_a \mathbf{u}$  ( $\mathbf{v}'_a \mathbf{v}_a$ ) elde edilir.
3.  $\mathbf{w}_a / |\mathbf{w}_a|$  ile boyu 1 olacak şekilde ölçeklendirilir.
4.  $\mathbf{t}_a = \mathbf{X}_a \mathbf{w}_a$  ile X'in ilgili bileşeni  $\mathbf{t}_a$ ,  $\mathbf{w}_a$  ağırlık vektörü ile X'in doğrusal kombinasyonu şeklinde hesaplanır.

5.  $\mathbf{t}_a$  bileşeninin Y'yi açıklayan  $\mathbf{c}_a$  ağırlık vektörü  $\mathbf{c}_a = \mathbf{Y}'_a \mathbf{t}_a / (\mathbf{t}'_a \mathbf{t}_a)$  ile Y'nin  $\mathbf{t}_a$  üzerine regresyonundan bulunur.
6.  $\mathbf{c}_a / |\mathbf{c}_a|$  ile boyu 1 olmalıdır.
7. Y için ilgili bileşen  $\mathbf{u}_{a(\text{yeni})}$ ,  $\mathbf{c}_a$  ağırlık vektörü ile Y'nin kombinasyonunu oluşturacak şekilde  $\mathbf{Y}_a \mathbf{c}_a / (\mathbf{c}'_a \mathbf{c}_a)$  hesaplanır.
8. 2. basamaktaki  $\mathbf{u}_a$  değeri ile 7. Basamaktaki  $\mathbf{u}_{a(\text{yeni})}$  değeri arasında yakınsama olup olmadığına bakılır. Bu yakınsamanın  $10^{-6}$  gibi sifıra çok yakın bir değer olması ile tespit edilir. Yakınsama sağlanırsa sonraki basamaklara geçilir, sağlanmazsa  $\mathbf{u}_{a(\text{yeni})}$  değeri  $\mathbf{u}_a$  yerine konularak algoritmaya devam edilir.
9. X'in ilgili bileşeni  $\mathbf{t}_a$  üzerine regresyonundan bileşenin bağımsız değişken üzerindeki etkisini ifade eden yük vektörü  $\mathbf{p}_a$ ,  $\mathbf{X}'_a \mathbf{t}_a / (\mathbf{t}'_a \mathbf{t}_a)$  ile bulunur.
10. Y'nin ilgili bileşeni  $\mathbf{u}_a$  üzerine regresyonundan, bileşenin bağımlı değişken üzerindeki etkisini ifade eden yük vektörü  $\mathbf{q}_a$ ,  $\mathbf{Y}'_a \mathbf{u}_a / (\mathbf{u}'_a \mathbf{u}_a)$  ile bulunur.
11. X ve Y için bileşenler ayrı hesaplandığı için bileşenler arasında zayıf bir ilişki mevcuttur. Bunu ortadan kaldırmak için her bir bileşen için Y'nin ilgili bileşeni  $\mathbf{u}_a$ 'nın X'in ilgili bileşeni  $\mathbf{t}_a$  üzerine regresyonundan elde edilen  $\mathbf{b}_a$  katsayısı  $\mathbf{b}_a = \mathbf{u}'_a \mathbf{t}_a / (\mathbf{t}'_a \mathbf{t}_a)$  ile hesaplanır.
12. Elde edilen bileşenler ve yükler değişkenleri modelde kullanılır. Bağımsız değişken  $\mathbf{X} = \mathbf{TP}'$  ve bağımlı değişken  $\mathbf{Y} = \mathbf{BTC}'$  ile modellenir. Bir sonraki bileşeni elde etmek için kullanılan  $\mathbf{X}_{a+1}$  ve  $\mathbf{Y}_{a+1}$  artık matrisleri  $\mathbf{X}_{a+1} \rightarrow \mathbf{X}_a - \mathbf{t}_a \mathbf{p}'_a$  ve  $\mathbf{Y}_{a+1} \rightarrow \mathbf{Y}_a - \mathbf{b}_a \mathbf{t}'_a$  ile hesaplanır.

Algoritmaya değişkenlerdeki değişim büyük ölçüde açıklanıncaya kadar devam edilir. Algoritma en az sayıda bileşen sayısını verir (Bulut ve Alın, 2009).

### 2.5.2 PLS Analizinin İstatistiksel Olarak Esinlenilmiş Değişikliğinin Basit Bir Uygulaması (SIMPLS)

SIMPLS algoritması 1993 yılında De Jong tarafından önerilmiştir. Bu algoritmanın farkı sonuçları indirgenmiş X matrisleri yerine orijinal X matrislerinin kombinasyonlarını kullanarak hesaplama yapmaktır. Bu sebepten NIPALS

algoritmasından farklı sonuçlar verir. Ancak sonuçlar arasındaki farklılık oldukça azdır (Bulut, 2011).

SIMPLS algoritmasında X matrislerinin orijinal kombinasyonları kullanıldığı için NIPALS algoritmasına göre daha hızlıdır (De Jong, 1993). Bağımlı değişkenler tek boyutlu ise sonuçlar iki algoritmada benzerlik gösterir. Fakat çok değişkenli durum söz konusu olduğunda sonuçlar farklılık gösterir (De Jong, 1993; Bulut, 2011).

### **2.5.3 Evrensel Kısmi En Küçük Kareler Algoritması (UNIPALS)**

Glen tarafından 1998 yılında bulunan bu algoritma kısmi en küçük karelerin hesaplanmasında daha küçük varyans-kovaryans matrislerini kullanan ilk algoritmadır.

NIPALS algoritmasında olduğu gibi orijinal veri matrisinin indirgenmesiyle sonuçlar elde ettiğinden UNIPALS VE NIPALS algoritmalarının sonuçları benzerlik gösterir (Lindgren ve Rännar, 1998).

### **2.5.4 Çekirdek Algoritması**

Değişken sayısının gözlem sayısından ya da gözlem sayısının değişken sayısından çok olduğu veri setlerinde NIPALS algoritması ile hesaplama yapmak zordur. Bu durumlarda kullanılması için Lindgren tarafından çekirdek algoritması geliştirilmiştir. İkinci çekirdek lgoritması Rännar, (1994) tarafından geliştirilmiştir. İki algoritma oldukça benzerdir ancak aralarındaki en önemli fark Rännar'ın önerdiği çekirdek algoritmasında açıklayıcı değişken sayısı gözlem sayısından büyük olduğunda daha tutarlı sonuç vermektedir. Her iki çekirdek algoritması da çoklu bağımlı değişken olması durumunda iyi sonuçlar verir.

### **2.5.5 Örnek-Uzaklık Kısmi En Küçük Kareler Algoritması (SAMPLS)**

Bu algoritma Bush ve Nachbar, (1993) tarafından özel olarak çok fazla açıklayıcı değişkenin ve az sayıda gözlem değerinin olduğu durumlar için önerilmiştir. SAMPLS algoritmasında bağımlı değişken sayısı bir olmalıdır. Hesaplama açısından SAMPLS, NIPALS algoritmasına göre daha üstündür. SAMPLS algoritması diğer algoritmalarındaki gibi ağırlık ve yük matrisleri vermez. Ağırlık ve yük matrislerini vermediği için değişkenler arasında korelasyon bilgisi kaybolur. Bu da SAMPLS algoritmasının önemli dezavantajıdır. Bu algorithmada indirgeme işlemi X ve Y

matrislerinin birisini seçme avantajı sağlamaktadır. Bağımlı değişken sayısı bir tane olduğundan bağımlı değişkeni indirgemek daha hızlı sonuç verir. Bu özelliği klasik algoritmaya benzediği için sonuçlarda benzerlik gösterir (Bush ve Nachbar, 1993).

### **2.5.6 Özvektör Algoritması**

Özvektör algoritması çekirdek algoritmaya benzer bir yapıya sahiptir. Özvektör algoritmasında kısmi en küçük kareler bileşenlerinin sayısı kadar olan en büyük özdeğerlere karşılık gelen özvektörler hesaplanır. Bu algoritma da indirgenme olmadığından büyük problemlerde çözüm üretmez (Polat, 2009; Bulut, 2011).

## **2.6 Önceki Çalışmalar**

Ridge regresyon yönteminin Hoerl ve Kennard, (1970) tarafından bulunduğu kabul edilir. Ridge regresyon analizi ile tahmin edilen regresyon katsayılarının, en küçük kareler regresyonuyla yapılan tahminlerden daha küçük hata kareler ortalamasına sahip oldukları belirlenmiş ve önerilmiştir.

Mahajan ve ark., (1977) çalışmalarında veri setinde çoklu bağlantı sorunu olması durumunda en küçük kareler regresyonu ile tahmin edilen pazarlama modellerine ait regresyon katsayılarının yanlış sonuçlara götüreceğini belirtmişlerdir. Bu amaçla çalışmalarında Ridge regresyon analizini kullanmışlardır.

Askin ve Montgomery, (1980) genişletilmiş veri kümeleri ile güçlendirilmiş sağlam tahmin ediciler tespit etmişlerdir. Böylece regresyon yöntemlerini birleştirme amacına ulaşmışlardır.

Kidwell ve Brown, (1982) çalışmalarında Ridge regresyon analizini yapay veriler ile kullanmışlardır. Bulgular, tahminleyicilerin ortogonal (dik) olmadığı durumda, yani çoklu bağlantı sorunu varlığında Ridge regresyonun en küçük kareler regresyonundan farklı sonuçlar verdiğini göstermiştir.

Liu, (1993) yaptığı çalışma ile farklı tahmincilerin birleştirilmesi ile onların avantajlarını da bir araya getirmeyi amaçlamıştır. Bu düşünce ile Stein tipi tahmincisi ile Ridge tahmincisini birleştirerek Liu tahmincisini bulmuştur. Liu regresyon analizini, en küçük kareler regresyonu ile çeşitli veri setleri aracılığıyla karşılaştırmıştır.

Kurtuluş, (2001) yapmış olduđu bir alıřmada, bađımsız (aıklayıcı) deđiřkenler arasında oklu bađlantı olması durumunda ortaya ıkan sorunları gidermek iin geliřtirilmiř olan Ridge regresyon analizi zerine alıřmıř ve bu yntemi en kk kareler regresyonu ile karřılařtırmıřtır.

Ortabař, (2001) alıřmasında oklu dođrusal bađlantı sorununu ortadan kaldırmak iin kullanılan temel bileřenler regresyon analizi ile Ridge regresyon analizi incelemiřtir. Yanlı regreyon yntemlerinin oklu bađlantı problemini ortadan kaldırdıđını, standart hatası daha kk hata kareler ortalamalı tahminlerin bulunduđunu belirlemiřtir.

Graham, (2003) yaptıđı alıřmada oklu bađlantı problemine karřılık eřitli istatistiksel yntemlerin kullanımını gstermiřtir. alıřmasında gerek ekolojik veriler kullanmıřtır. Graham bu yaptıđı ile gerek ekolojik verilere ait oklu regresyon modellerinde oklu bađlantılı durumların yorumlanmasını kolaylařtırmıřtır.

Ergneř, (2004) yaptıđı alıřmada en kk kareler regresyonunu Ridge regresyon analizi ile karřılařtırmıřtır. oklu bađlantı sorununun yok edilmesi veya en aza indirilmesi iin Ridge regresyon analizinin kullanımının gerekliliđini anlatmıřtır. Gvenilir, sađlam sonular iin en kk kareler regresyonu yerine Ridge regresyon analizini nermiřtir.

Albayrak, (2005) yaptıđı alıřmada en kk kareler regresyonunu, Ridge regresyon analizi ve temel bileřenler regresyon analizi karřılařtırmıřtır. Karřılařtırma sonucunda, bađımsız deđiřkenler arasındaki yksek oklu bađlantı probleminden dolayı Ridge regresyon ve temel bileřenler regresyonu, en kk kareler regresyonuna kıyasla daha dřk standart hataya sahip olup, en kk kareler regresyonundan daha tutarlı tahminler sađladıđını tespit etmiřtir.

Karadavut ve ark., (2005) yapmış oldukları alıřmada nohut (*cicer arietinum L.*) bitkisinin verimini etkileyen bazı zellikleri en kk kareler regresyonu, Ridge regresyon analizi ve M- regresyon analizi ile karřılařtırmayı hedeflemiřlerdir. Nohut bitkisinin tane ađırlıđını etkileyen deđiřkenlere en kk kareler regresyonu uygulamıřlardır. Analiz sonularında oklu bađlantı problemini tespit ettikten sonra Ridge regresyon ve M- regresyon analizleri uygulamıřlardır. Elde ettikleri parametre tahminleri karřılařtırılmıř ve analiz sonularına gre M- regresyon analizinin diđer yntemlerden daha uygun ve tutarlı olduđu tespit edilmiřtir.

Yolacan ve ark., (2005) çalışmalarında çoklu bağlantı sorunu olan ekonomik verilere Ridge regresyon ve yapay sinir ağıları algoritmalarını uygulamış ve iki yöntemi karşılaştırmalı olarak yorumlamışlardır.

Çankaya ve ark., (2006) çoklu doğrusal regresyon yönteminde parametre tahminlerinin karşılaştırılması üzerine çalışmalar yapmışlardır. Araştırmanın sonucunda, aykırı değerler tespit edildiğinde en küçük kareler regresyonunun en yüksek belirtme katsayısına sahip olduğunu belirtmişlerdir. Aykırı değerlerin ve yüksek uç değerlerin veri setinden uzaklaştırılmaması, etkilerinin araştırılması yeni yöntemler ile tutarlı sonuçlar bulunması tavsiye edilmiştir.

Karakaş, (2008) tarafından yapılan çalışmada çoklu doğrusal bağlantının sebepleri, çoklu bağlantının nasıl belirleneceği incelenmiştir. Çalışmada Liu ve Ridge regresyon yöntemleri başta olmak üzere yanlı regresyon yöntemleri özellikleri ile birlikte incelenmiştir. Uygulama çalışmasında yanlı regresyon yöntemlerinden Ridge regresyon ve Liu regresyon kullanılarak bir istihdam modelinin tahmini amaçlanmıştır. Çalışmanın sonucunda en küçük kareler, Ridge, Liu regresyon analizleri karşılaştırılmıştır. Liu regresyon parametrelerinin anlamlılığı açısından daha uygun olduğuna karar verilmiştir.

Pamukçu, (2010) çalışmasında sistolik kan basıncının (SKB) tahmini için yanlı regresyon yöntemlerini incelemeyi amaçlamıştır. Bu amaç doğrultusunda, 1 Ekim 2009-22 Aralık 2009 tarihleri arasında Fırat Üniversitesi Tıp Merkezi Kardiyoloji Polikliniğine gelen 128 hastadan yaş, boy, karın çevresi, glikoz, HDL, LDL, üre, kreatinin, potasyum, HGB, HCT ve diastolik kan basıncı gibi açıklayıcı değişkenler elde edilmiştir. Yanlı regresyon analizlerinden Ridge regresyon analizinde çoklu bağlantının giderildiği gözlemlenmiştir. Sonuçta, diastolik kan basıncı, karın çevresi ve kreatin değişkenlerinin, sistolik kan basıncının tahmininde kullanılabileceğine, temel bileşenler regresyonu analizi ve Ridge regresyon analizi ile daha tutarlı tahminler elde edildiği tespit edilmiştir. Temel bileşenler regresyonu ve Ridge regresyonu sonuçları kıyaslandığında ise, temel bileşenler regresyonunun daha iyi tahminler yaptığı bulunmuştur.

Tüylüoğlu ve Albayrak, (2010) çalışmalarında Türkiye'de hayat pahalılığını ölçmek amacıyla 2008 yılına ait 26 ili içeren TÜİK tarafından seçilmiş 375 ürün ve hizmetin

fiyatlarına temel bileşenler regresyonu uygulayarak, iller için standartlaştırılmış mekansal tüketici fiyat endekslerini hesaplamışlardır. Sonuçlara göre en pahalı iller İstanbul, Ankara, İzmir iken; en ucuz il grupları Şanlıurfa-Diyarbakır, Gaziantep-Adıyaman-Kilis, Mardin-Batman-Şırnak-Siirt'tir. Araştırmacılar çalışmalarında Ridge regresyon yöntemi ile illerin genel fiyat düzeyleri ile diğer alt gruplar arasındaki ilişkileri incelemişlerdir. Hayat pahalılığını belirleyen faktörlerin sırasıyla giyim, mobilya ve ev eşyası, diğer ürün ve hizmetler ile konut harcamaları olduğunu tespit etmişlerdir.

Yıldırım, (2010) çalışmasında en küçük kareler regresyonu, Ridge regresyon analizi ve Robust regresyon analizini incelemeyi, analiz yaptıktan sonra sonuçları karşılaştırmayı ve aykırı değerlerin etkilerini belirlemeyi amaç edinmiştir. Araştırmalarının sonucunda regresyon analizine ait belirtme katsayısı ve hata kareler ortalaması değerlerinin aykırı değer varlığından etkilendiğini, aykırı değer olduğu durumda ve veri setinden çıkarıldıktan sonra parametre tahminlerinden elde edilen t, F değerlerinin farklılık gösterdiğini, aykırı değerlerin çoklu bağlantı problemini en aza indirdiğini tespit etmiştir.

Rathert ve ark., (2011) çalışmalarında Japon bıldırcınlarında yumurta kalite özelliklerinin tahmin edilmesinde temel bileşenler regresyonu ile en küçük kareler regresyonunu karşılaştırıp incelemişlerdir. Araştırmada yaş aralığı 20-24 hafta arasında olan, seleksiyona uğramamış, rastgele çiftleşmiş Japon bıldırcın yumurtaları kullanılmıştır. Sonuç olarak, tespit edilen çoklu bağlantı probleminin temel bileşenler regresyonu ile çözülebildiği görülmüştür.

Aşkın, (2011) çalışmasında sol atriyum çapının, temel bileşenler regresyonu, kısmi en küçük kareler regresyonu ve yapay sinir ağları yöntemleri ile tahminlerinin yapılmasını amaçlamıştır. Bu amaçla Fırat Üniversitesi Tıp Fakültesi Kardiyoloji Polikliniğine başvuran 127 hipertansif hastanın ekokardiyografi raporlarını prospektif olarak toplamıştır. Temel bileşenler regresyon yöntemi, kısmi en küçük kareler regresyon analizi ve yapay sinir ağları yöntemleri ile topladığı bu verileri analiz etmiştir. Analiz neticesinde, yapay sinir ağları yönteminin, sol atriyum çapının tahmininde daha iyi sonuç verdiğini tespit etmiştir.

Eker, (2011) çalışmasında regresyon analizinde çoklu bağlantı probleminin tespit edilmesi durumunda en küçük kareler regresyonunun varsayımları sağlayamamasından dolayı, çoklu bağlantı problemini tespit ettikten sonra Ridge regresyon ve temel bileşenler regresyon analizlerini incelemeyi hedeflemiştir. Bu amaçla, Ondokuz Mayıs Üniversitesi Araştırma ve Uygulama Çiftliği'nde yetiştirilen 85 tane Karayaka kuzusundan cidago yüksekliği, sağrı yüksekliği, vücut uzunluğu, göğüs genişliği, göğüs çevresi, göğüs derinliği, ön, orta, arka sağrı genişliği gibi vücut ölçüleri ile canlı ağırlık ölçüleri alınmıştır (Veriler süttten kesim döneminde alınmıştır). En küçük kareler regresyonu ile yapılan analizler sonucunda ön sağrı ve orta sağrı genişlikleri için çoklu bağlantı problemi tespit edilmiştir. Verilere SPSS, MINITAB ve NCSS gibi farklı istatistik programları ile Ridge regresyon ve temel bileşenler regresyonu uygulanmış; bağımsız değişkenler üzerindeki çoklu bağlantının etkileri ortadan kaldırılmıştır. Araştırmacı yaptığı analizler sonucunda, çoklu bağlantı probleminin varlığında en küçük kareler regresyonu yerine Ridge regresyon ve temel bileşenler regresyon analizlerini kullanmayı önermiştir.

Duran, (2011) çalışmasında parametrik ve semiparametrik regresyon modellerinde çoklu bağlantı problemi varlığında kullanılabilir yanlı tahmin edici yöntemleri incelemiştir. Parametrik olmayan regresyon yöntemlerinde bağımsız değişken sayısı arttıkça tahminler gerçeklikten uzaklaşacağı için daha tutarlı sonuçlar için alternatif yöntem olarak semiparametrik regresyon modellerini önermiştir. Semiparametrik regresyon modelleri parametrik ve parametrik olmayan regresyonların bir kombinasyonudur. Araştırmacı çalışmasında semiparametrik regresyon modellerinde kullanılan Speckman ve cezalı en küçük kareler regresyonu yöntemlerine dayalı yeni Liu regresyon analizine göre tahmin ediciler üretmiştir. Simülasyon çalışması sonucunda Liu regresyon analizinin standart hata kareler ortalaması ölçütüne göre diğer yöntemlerden üstünlüğünü ispatlamıştır. Çoklu bağlantının güçlü olması durumunda üstünlüğün belirginleştiğini tespit etmiştir.

Demirci, (2014) çalışmasında genetik algoritma yaklaşımı ile Ridge regresyon parametresi olan  $k$  sabitinin belirlenmesini incelemiştir. Uygun  $k$  değerinin tespiti için Ridge izi analizi, iteratif Ridge analizi ve yönlendirilmiş Ridge analizi ve genetik algoritma iki farklı veriye uygulanmış ve sonuçlar karşılaştırılmıştır. Karşılaştırma



sonucunda Ridge regresyonu parametresi olan  $k$  değeri genetik algoritma ile bir çözüm kümesinden oluşan çok sayıda değer ile incelenmiştir.

Asar, (2015) çalışmasında Liu tipi lojistik regresyon tahmin edicileri üzerinde analizler yapmıştır. Binari lojistik regresyon yönteminde çoklu bağlantı sorununun varlığı varyansı şişirmekte ve analizlerin performansı düşmektedir. Bu yüzden çoklu bağlantı problemini ortadan kaldırmak için lojistik regresyon kullanılmıştır. Çalışmasında, Monte Carlo simülasyonunu kullanmış ve daha önceki yöntemlerle hata kareler ortalamaları kriteri bakımından performansı karşılaştırma yapmıştır. Çalışmasının sonucunda, parametrelerin seçimi için tekrarlı (iterative) bir yöntem kullanılmasını önermiştir.

Kırdemir, (2017) farklı istatistik paket programlarıyla Ridge regresyon analizini incelemiştir. Araştırma için ilköğretim 6. sınıfta okuyan 40 öğrenciyi rasgele seçerek, matematik dersindeki başarıya türkçe, fen, sosyal ve dil alanlarındaki başarının etkisini tespit etmeye çalışmıştır. SPSS, NCSS, R studio programları ile araştırma ayrıntılı incelenmiştir. Araştırmanın sonucunda, açıklayıcı değişkenler arasında güçlü çoklu doğrusal bağlantı olduğu zaman, Ridge regresyonun tutarlı, geçerli, durağan ve kuramsal beklentilere uygun tahminler sağladığı saptamıştır.

Kuvat, (2018) çalışmasında Ridge regresyon analizi için sağlam parametre bulmayı amaçlamıştır. En iyi  $k$  parametresini bulmak için önerilen  $k$  parametrelerinden en iyi performansa sahip olanları seçmiştir ve neticede en uygun  $k$  parametresini elde etmiştir. Diğer parametrelerle elde ettiği bu parametreyi karşılaştırmıştır. En küçük hata kareler ortalaması bakımından karşılaştırma yapıldığında, önerilen  $k$  parametresinin diğerlerine göre daha fazla sonuç verdiğini tespit etmiştir.

Tunç, (2018) yaptığı bir çalışmada en küçük kareler regresyon analizi ve temel bileşenler regresyon analizini karşılaştırmayı hedeflemiştir. Bu amaçla, çoklu bağlantının derecesinin ve örnek genişliğinin yeterliliğinin tespiti için simülasyonla iki farklı veri grubu türetmiştir. Birinci veri grubu, çoklu bağlantıya sahip 10 veri seti; ikinci veri grubu örnek genişliği farklı 10 veri setinden oluşmuştur. Simülasyonla türetilen bu verilerde çoklu bağlantının yaygın olduğunu tespit etmiş ve tüm verilere en küçük kareler regresyon analizi ve temel bileşenler regresyonu uygulamıştır. Uygulama sonucunda, en küçük kareler regresyon analizi ile elde edilen katsayıların,

temel bileşenler regresyonunda elde edilen katsayılardan işaret ve büyüklük olarak farklı olduğunu gözlemlemiştir. Ayrıca, temel bileşenler regresyon analizi sonucunda elde edilen katsayıların standart hatalarının daha düşük olduğunu tespit etmiştir.

### 3. MATERYAL ve YÖNTEM

#### 3.1 Materyal

Bu çalışmanın materyalini, basit tesadüfi örnekleme yolu ile seçilen 100 adet yumurtanın kalite parametreleri olarak adlandırılan değişkenlerine ait veriler oluşturmaktadır. Yumurtalar altlıklı yer sisteminde yetiştirilen yaklaşık 40 haftalık Beç tavuklarından toplanmıştır. Yumurtalar kalite ölçümünden önce 24 saat oda sıcaklığında bekletilmiştir.

#### 3.2 Yöntem

##### 3.2.1 Çalışma Değişkenleri

Çalışmada yumurta iç özelliklerinden ak ağırlığı (AA) ve sarı ağırlığı (SA) dış özelliklerinden ise yumurta ağırlığı (YA), yumurta eni (YE), yumurta boyu (YB), kabuk ağırlığı (KA), kabuk kalınlığı (KK) ve şekil indeksi (Şİ) özellikleri kalite parametreleri olarak dikkate alınmıştır. Çalışmada değişken olarak dikkate alınan yumurta kalite parametreleri Çizelge 3.1’de verilmiştir.

##### Çizelge 3.1 Çalışma Değişkenleri

---

Ak Ağırlığı (g)	= Yumurta Ağırlığı - (Sarı Ağırlığı + Kabuk Ağırlığı)
Kabuk Kalınlığı (mm)	= (Sivri uç+Orta kısım+Küt uç)/3
Şekil İndeksi	= 100 * (Yumurta Genişliği / Yumurta Uzunluğu)
Sarı Ağırlığı (g)	
Yumurta Ağırlığı (g)	
Yumurta Eni (mm)	
Yumurta Boyu (mm)	
Kabuk Ağırlığı (g)	

---

##### 3.2.2 Regresyon Yöntemlerinin Performans Göstergeleri

Çalışmada kullanılan regresyon yöntemlerinin incelenmesinde; tahminlenen katsayıların standart hatalarının yanı sıra Hata Kareler Ortalaması (HKO), modele ait düzeltilmiş belirleme katsayısı ( $R^2$ -düz) ve Akaike Bilgi Kriteri (AIC) kullanılmıştır. Bu kriterler için; her zaman standart hata, HKO ve AIC değerlerinin küçük, ( $R^2$ -düz) değerinin ise yüksek olması istenilmektedir (Shibata, 1981).

Çalışmada yapılan tüm hesaplamalar SPSS v25 (IBM Inc., Chicago, IL, USA), NCSS 9 (NCSS, LLC., Kaysville, Utah, USA, [ncss.com/software/ncss](http://ncss.com/software/ncss)) ve MINITAB 18 (Minitab Inc., PA, USA). istatistik paket programları kullanılarak yapılmıştır.

## 4. BULGULAR ve TARTIŞMA

### 4.1 Bulgular

#### 4.1.1 Çalışma Değişkenlerine Ait Tanıtıcı İstatistik Değerleri

Çalışmada kullanılan değişkenlere ait tanıtıcı istatistik değerleri; standart sapma, ortalama, en küçük ve en büyük değerler olarak Çizelge 4.1'de verilmiştir. Çalışmada verilerin normal dağılım kontrolü yapılmış ve Kolmogorov-Smirnov testi tüm değişkenlerin normal dağılım gösterdiğini göstermiştir ( $p>0.05$ ).

**Çizelge 4.1** Çalışmada Kullanılan Değişkenlere Ait Tanıtıcı İstatistikler

Değişken	n	Ortalama	Standart Sapma	En Küçük Değer	En büyük Değer
Ak Ağırlığı	100	20.06	1.71	16.67	23.52
Sarı Ağırlığı	100	13.59	0.99	11.69	15.87
Yumurta Ağırlığı	100	40.14	2.35	34.03	45.70
Yumurta Eni	100	37.89	0.86	36.40	40.02
Yumurta Boyu	100	49.47	1.07	46.44	52.68
Şekil İndeksi	100	76.62	1.85	71.73	80.52
Kabuk Ağırlığı	100	6.48	0.78	3.01	8.97
Kabuk Kalınlığı	100	0,54	0.04	0.42	0.69

Her bir bağımsız değişkenin birbirleriyle ve bağımlı değişkenler aralarındaki korelasyon katsayılarına ilişkin Pearson korelasyon katsayıları ve önemlilik düzeyleri Çizelge 4.2'deki korelasyon matrisinde verilmiştir. En yüksek korelasyon kabuk ağırlığı ile kabuk kalınlığı arasında %85.5 olarak hesaplanmış ve istatistiksel olarak önemli bulunmuştur ( $p<0.001$ ). Yumurta eni ve yumurta ağırlığı arasında hesaplanan korelasyon katsayısı ise %85.2 olarak bulunmuştur ( $p<0.001$ ). En düşük korelasyon, bağımlı değişken olan ak ağırlık ile kabuk kalınlığı arasında %0.4 olarak hesaplanmıştır ( $p>0.05$ ). Bağımsız değişkenler arasında gözlenen yüksek korelasyonlar çoklu bağlantı sorununu işaret etmektedir. Ancak çoklu bağlantı sorununun varlığından emin olmak için basit korelasyon katsayıları tek başına yeterli değildir.

**Çizelge 4.2** Değişkenler Arasındaki Korelasyon Katsayıları

Değişken	YA	YE	YB	Şİ	KA	KK	AA	SA
YA	1							
YE	<b>0.852***</b>	1						
YB	0.569**	0.396**	1					
Şİ	0.290**	0.582**	-0.515**	1				
KA	0.658**	0.507**	0.248*	0.250*	1			
KK	0.460**	0.329**	0.143	0.178	<b>0.855***</b>	1		
AA	0.745**	0.681**	0.520**	0.177	0.123	0.004	1	
SA	0.566**	0.446**	0.254*	0.184	0.557**	0.407**	-0.057	1

AA: Ak Ağırlık; KA: Kabuk Ağırlığı; KK: Kabuk Kalınlığı; SA: Sarı Ağırlık; Şİ: Şekil İndeksi; YA: Yumurta Ağırlığı; YB: Yumurta Boyu; YE: Yumurta Eni  
\*:p<0.05, \*\*:p<0.01 \*\*\*:p<0.001

#### 4.1.2 Yumurta Ak Ağırlığı Bağımlı Değişkenine Ait Bulgular

##### 4.1.2.1 En Küçük Kareler Analiz Sonuçları

Çalışmada yumurta ak ağırlığı ve sarı ağırlığı bağımlı değişkenler olarak dikkate alınırken, yumurta eni, yumurta boyu, yumurta ağırlığı, şekil indeksi, kabuk ağırlığı ve kabuk kalınlığı bağımsız değişkenler olarak dikkate alınmıştır. En küçük kareler regresyonu için modele ait varyans analizi sonuçları verilmiştir (Çizelge 4.3). Yapılan varyans analizi sonucuna göre F önemlidir (p<0.01). Tabloya göre R<sup>2</sup> değeri 0.802 bulunmuştur. Fakat çoklu regresyon analizi olduğu için düzeltilmiş R<sup>2</sup> değeri alınır. R<sup>2</sup><sub>d</sub> değeri 0.790'dır. Yani bağımsız değişkenler (yumurta ağırlığı, yumurta eni, yumurta boyu, kabuk ağırlığı, kabuk kalınlığı, şekil indeksi) bağımlı değişkendeki (ak ağırlık) değişimi % 79 açıklamaktadır. Ayrıca yapılan analizlere göre Durbin-Watson testi sonucu 1.861 olarak hesaplanmıştır. Bu, otokorelasyon sorunu olmadığını göstermektedir.

**Çizelge 4.3** En Küçük Kareler Varyans Analizi Tablosu

Model	Serbestlik Derecesi	Kareler Toplamı	Kareler Ortalaması	F	p
Regresyon	6	233.019	38.836	62.916	0.000
Hata	93	57.406	0.617		
Genel	99	290.425			
R <sup>2</sup>	0.802				
R <sup>2</sup> -düz	0.790				

En küçük kareler regresyon sonuçları, anlamlılık değeri ve modelin katsayıları verilmiştir (Çizelge 4.4). Yumurta ağırlığının katsayısı 0.822 bulunmuştur ve

önemlidir ( $p < 0.001$ ). Kabuk ağırlığının katsayısı -1.511 bulunmuştur ve model için önemlidir ( $p < 0.001$ ). Fakat ak ağırlık bağımlı değişkeni ile kabuk ağırlığı ters yönde (negatif) ilişkilidir. Bağımsız değişkenlerden yumurta eni, yumurta boyu ve şekil indeksi bağımsız değişkenlerinin varyans büyütme faktörü değerleri yüksek çıkmıştır ( $VIF > 10$ ). Varyans büyütme faktörleri yüksek olan yumurta eni, yumurta boyu ve şekil indeksi bağımsız değişkenlerinin tolerans değerleri çok küçüktür.

**Çizelge 4.4** En Küçük Kareler Çoklu Regresyon Analizi Sonuçları

Değişken	Katsayı	Standart Hata	t	p	Tolerans Değeri	VIF
Sabit	-252.505	125.061	-2.019	0.046		
YA	0.822	0.088	9.368	<b>0.000</b>	0.146	6.842
YE	-6.243	3.290	-1.898	0.061	0.001	<b>1310.216</b>
YB	4.891	2.502	1.955	0.054	0.001	<b>1155.800</b>
KA	-1.511	0.241	-6.269	<b>0.000</b>	0.174	5.747
KK	2.787	3.653	0.763	0.448	0.248	4.027
Şİ	3.165	1.635	1.935	0.056	0.001	<b>1480.038</b>

Özdeğerler ve koşul indeksi verilmiştir (Çizelge 4.5). Özdeğerler 1'e eşit olduğunda çoklu bağlantı oluşmaz. Fakat özdeğerlerden en az biri sıfıra yakın olduğu için çoklu bağlantı sorunu gözlemlenir. Çizelge 4.5 incelendiğinde; 3, 4, 5, 6 ve 7 numaralı özdeğerlerde çoklu bağlantı sorunu saptanır. Koşul indeksi 30'dan büyük olduğunda birden fazla çoklu bağlantı sorunu oluştuğu için 3, 4, 5, 6 ve 7 güçlü çoklu bağlantıya sebep olmuştur.

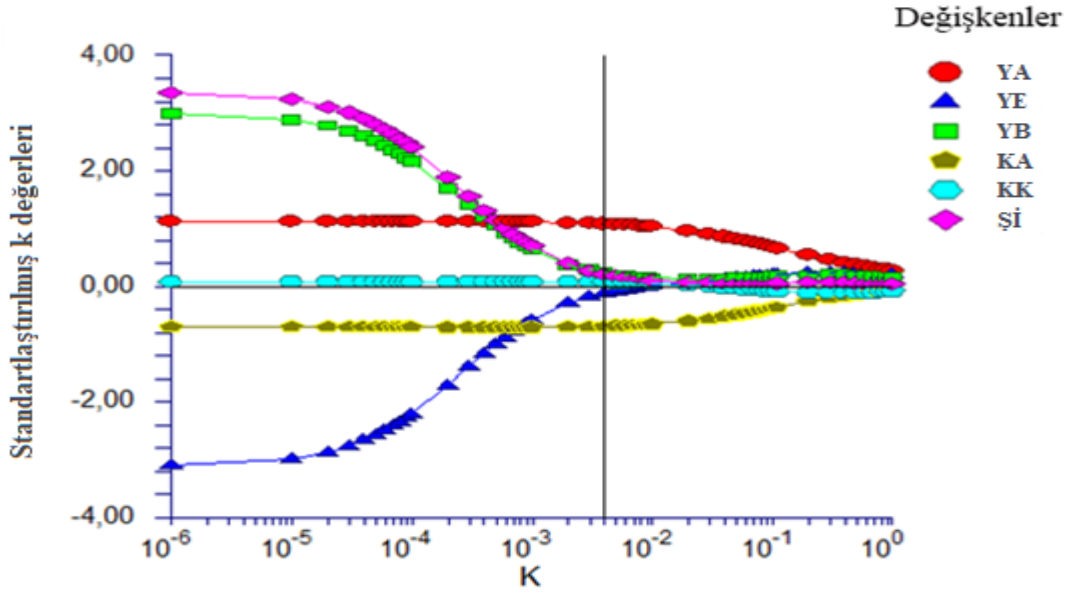
**Çizelge 4.5** Değişkenlere Ait Korelasyonların Özdeğerleri ve Koşul İndeksleri

Numara	Özdeğer	Koşul İndeksi
1	6.984	1.000
2	0.012	24.030
3	0.002	<b>56.809</b>
4	0.001	<b>88.453</b>
5	0.001	<b>104.261</b>
6	0.000	<b>329.251</b>
7	0.000	<b>8385.127</b>

Yumurta kalite değişkenleri için çoklu bağlantı sorununu belirleyen yöntemler incelendiğinde (korelasyon matrisi, varyans büyütme değeri, tolerans, özdeğerler ve koşul indeksi) bağımsız değişkenler arasında çoklu bağlantı sorununun oluştuğu görülmüştür. Yansız olan en küçük kareler regresyonu çoklu bağlantı sorununu ortadan kaldıramadığı için sonuçlar tutarsız olur.

#### 4.1.2.2 Ridge Regresyon Analiz Sonuçları

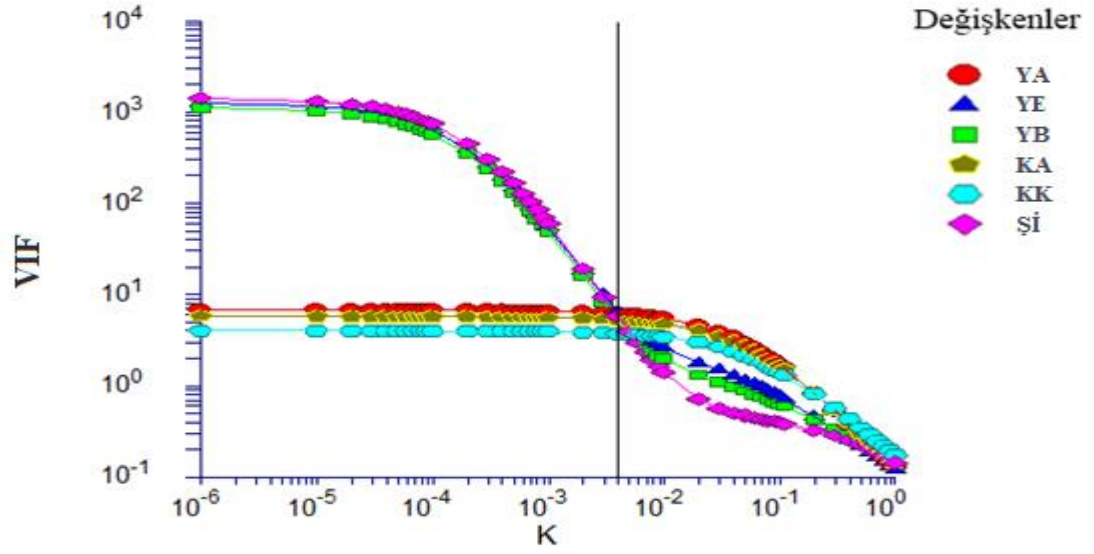
Ak ağırlık bağımlı değişkeni için Ridge izi grafiği verilmiştir (Şekil 4.1). Ridge izi grafiğinde dikey eksende standartlaştırılmış k değerlerinin, yatay eksende k parametresinin olduğu gözlemlenir. k parametresi arttıkça standartlaşan k değerleri durağanlaşır. Şekil 4.1 incelendiğinde,  $k=0.004$  sabitinden sonra regresyon katsayılarının daha durağan hale geldiği görülmektedir.



Şekil 4.1 Ak Ağırlık İçin Ridge İzi Grafiği

Ak Ağırlık İçin Ridge Regresyon VIF Grafiği Şekil 4.2'de verilmiştir. k parametresinin VIF değerlerindeki etkisini gösteren Şekil 4.2 incelendiğinde  $k=0.004$  sabitinden sonra VIF değerlerinin 10'un altında olduğu görülmektedir.





**Şekil 4.2** Ak Ağırlık İçin Ridge Regresyon VIF Grafiği

Standartlaştırılmış Ridge parametresi  $k$  verilmiştir (Çizelge 4.6). Çizelgeye göre yumurta ak ağırlığının pozitif ilişkili olması beklenen bağımsız değişkenler için  $k=0.004$  sabiti seçilerek çoklu bağlantı sorunu ortadan kaldırılmıştır.

**Çizelge 4.6** Standartlaştırılmış Ridge Regresyon Değerleri

<b>k</b>	<b>YA</b>	<b>YE</b>	<b>YB</b>	<b>KA</b>	<b>KK</b>	<b>Şİ</b>
0.001	1.105	-0.5881	0.647	-0.7111	0.0839	0.695
0.002	1.0943	-0.2998	0.3822	-0.7058	0.081	0.3929
0.003	1.085	-0.1865	0.2817	-0.6996	0.0776	0.2768
0.004	1.0761	-0.1249	0.2297	-0.6932	0.0742	0.2156
0.005	1.0675	-0.0855	0.1984	-0.6867	0.0707	0.178
0.006	1.0592	-0.0577	0.1778	-0.6804	0.0674	0.1526
0.007	1.0511	-0.0368	0.1636	-0.6741	0.064	0.1345
0.008	1.0432	-0.0202	0.1534	-0.668	0.0608	0.1209
0.009	1.0355	-0.0066	0.1459	-0.6619	0.0576	0.1103
0.01	1.0279	0.0048	0.1402	-0.656	0.0545	0.102
0.02	0.9601	0.0698	0.1245	-0.6025	0.0268	0.0668
0.03	0.9037	0.1042	0.129	-0.5579	0.0044	0.0577
0.04	0.8559	0.1283	0.1367	-0.5201	-0.0138	0.0548
0.04	0.8559	0.1283	0.1367	-0.5201	-0.0138	0.0548
0.05	0.8149	0.1469	0.1446	-0.4878	-0.029	0.054
0.06	0.7792	0.1619	0.152	-0.4597	-0.0417	0.054
0.07	0.7479	0.1742	0.1585	-0.4351	-0.0524	0.0544
0.08	0.7201	0.1846	0.1643	-0.4132	-0.0614	0.0549
0.09	0.6953	0.1933	0.1695	-0.3938	-0.0692	0.0555
0.1	0.6729	0.2009	0.1739	-0.3763	-0.0758	0.0561
0.11	0.6514	0.2077	0.1781	-0.3594	-0.0819	0.0567
0.2	0.5294	0.2383	0.1972	-0.2643	-0.1083	0.0593
0.3	0.4533	0.247	0.2022	-0.2056	-0.1148	0.0594
0.4	0.404	0.2462	0.2007	-0.168	-0.1131	0.0582
0.5	0.3684	0.2418	0.1966	-0.1414	-0.1084	0.0565
0.6	0.3408	0.2357	0.1914	-0.1213	-0.1027	0.0548
0.7	0.3186	0.2291	0.1858	-0.1054	-0.0968	0.053
0.8	0.3	0.2223	0.1801	-0.0926	-0.091	0.0513
0.9	0.2842	0.2156	0.1745	-0.082	-0.0855	0.0496
1	0.2704	0.2091	0.1691	-0.073	-0.0804	0.0481

Ridge parametresi  $k$ 'ye bağlı olarak değişen parametreler verilmiştir (Çizelge 4.7).  $k=0$  sabiti için  $R^2$  değeri en büyük olduğu için,  $k$  sabiti sıfıra yakın olmalıdır. Sigma, hata kareler ortalamasının kareködür. Sigma, en küçük değerini en küçük kareler regresyonunda aldığı için  $k$  değeri bu değerden çok sapmamalıdır.  $B'B$ , standartlaştırılmış katsayıların kareler toplamıdır ve  $k$  sabitine göre bu değer durağanlaşmalıdır. Ortalama varyans büyütme faktörü, varyans büyütme faktörü değerlerinin ortalamasıdır. En büyük varyans büyütme faktörü,  $k$  sabitinin varyans büyütme değerinin en büyüğünü verir. Bu parametrelerin beklentilerine göre en uygun  $k=0.004$  olarak seçilmiştir.

**Çizelge 4.7** Ridge Regresyon k Analiz Tablosu

<b>k</b>	<b>R<sup>2</sup></b>	<b>SİGMA</b>	<b>B'B</b>	<b>Ort. VIF</b>	<b>En Büyük VIF</b>
0.001	0.795	0.801	2.981	30.502	61.568
0.002	0.792	0.806	2.093	11.747	19.458
0.003	0.790	0.809	1.863	7.320	10.094
<b>0.004</b>	<b>0.789</b>	<b>0.813</b>	<b>1.759</b>	<b>5.590</b>	<b>6.732</b>
0.005	0.787	0.816	1.695	4.721	6.148
0.006	0.785	0.819	1.648	4.212	6.030
0.007	0.784	0.822	1.610	3.879	5.915
0.008	0.782	0.825	1.577	3.644	5.803
0.009	0.780	0.828	1.547	3.466	5.695
0.01	0.779	0.831	1.520	3.326	5.590
0.02	0.765	0.857	1.310	2.610	4.687
0.03	0.752	0.879	1.159	2.231	3.996
0.04	0.741	0.899	1.042	1.955	3.453
0.05	0.732	0.916	0.948	1.738	3.019
0.06	0.722	0.931	0.873	1.562	2.666
0.07	0.714	0.945	0.810	1.417	2.375
0.08	0.706	0.958	0.757	1.294	2.131
0.09	0.699	0.970	0.712	1.190	1.926
0.1	0.692	0.981	0.674	1.101	1.750
0.110	0.685	0.992	0.638	1.018	1.590
0.2	0.637	1.065	0.461	0.620	0.8146
0.3	0.596	1.123	0.366	0.432	0.570
0.4	0.563	1.169	0.309	0.333	0.436
0.5	0.534	1.207	0.268	0.272	0.353
0.6	0.509	1.239	0.237	0.230	0.295
0.7	0.486	1.267	0.212	0.200	0.253
0.8	0.466	1.291	0.191	0.176	0.221
0.9	0.448	1.313	0.174	0.157	0.195
1	0.431	1.333	0.160	0.142	0.174

Sigma: Hata kareler ortalaması karekökü

k=0.004 sabitine göre Ridge regresyon ve en küçük kareler regresyon analizi sonuçları verilmiştir (Çizelge 4.8). Çoklu bağlantı sorununu ortadan kaldıran Ridge regresyon analizinde her değişkenin standart hatasında azalma meydana gelmiştir. Çoklu bağlantıya neden olan yumurta eni, yumurta boyu ve şekil indeksi bağımsız değişkenlerinin standart hatalarında yaklaşık olarak %82 oranında azalma görülmüştür. Ayrıca yapılan analiz sonucunda düzeltilmiş R<sup>2</sup> değeri en küçük kareler (OLS) için 0.790 iken, Ridge regresyon için 0.775 bulunmuştur. Yani Ridge regresyon R<sup>2</sup>'yi çok değiştirmeden daha düşük standart hatalı güvenilir sonuçlar elde etmiştir.

**Çizelge 4.8** Ridge Katsayısı  $k=0.004$  İçin Ridge Regresyon ve En Küçük Kareler Karşılaştırılması

Değişken	Ridge Katsayısı	OLS Katsayısı	Std. Ridge	Std. OLS	Ridge SH	OLS SH
Sabit	-27.204	-247.295				
YA	0.782	0.822	1.076	1.130	0.087	0.088
YE	-0.246	-6.098	-0.125	-3.093	0.243	3.267
YB	0.367	4.782	0.230	2.996	0.179	2.486
KA	-1.511	-1.531	-0.693	-0.702	0.240	0.243
KK	2.951	3.188	-0.074	0.080	3.692	3.688
Şİ	0.199	3.094	0.216	3.356	0.106	1.625
$R^2$	0.789	0.802				
$R^2$ -düz	0.775	0.790				
SİGMA	0.812	0.785				

Aşağıda Ridge regresyon için  $k=0.004$  sabiti için katsayılara göre yumurta ak ağırlığını etkileyen model:

$$-27.20421 + 0.7826706*YA - 0.2461512*YE + 0.3666491*YB - 1.511435*KA - 2.951395*KK + 0.1988136*Şİ \quad (4.1)$$

şeklinde bulunur. Burada;

YA	: Yumurta ağırlığı,
YE	: Yumurta eni,
YB	: Yumurta boyu,
KA	: Kabuk ağırlığı
KK	: Kabuk kalınlığı
Şİ	: Şekil indeksidir.

Ridge regresyona ait katsayılar verilmiştir (Çizelge 4.9). Standart hataların düştüğü, varyans büyütme faktörlerinin 10'un altına düştüğü görülmektedir.

**Çizelge 4.9** Ridge Regresyon Katsayıları

Değişken	Katsayı	Standart Hata	Std. Reg Katsayısı	VIF
Sabit	-27.204			
YA	0.782	0.086	1.076	6.270
YE	-0.247	0.243	-0.124	6.731
YB	0.367	0.179	0.229	5.549
KA	-1.511	0.240	-0.693	5.357
KK	2.951	3.692	0.074	3.787
Şİ	0.199	0.106	0.216	5.843

$k=0.004$  sabiti için Ridge regresyon varyans analizi sonuçları verilmiştir (Çizelge 4.10). Çizelgeye göre yumurta ak ağırlığında meydana gelen değişimin %77.5'i Beç tavuk yumurtasının diğer değişkenleri tarafından açıklanmıştır. Çizelgeye göre Beç tavuk yumurtasında ak ağırlık, Beç tavuk yumurtasına ait diğer değişkenlerle açıklanabilmektedir ( $p<0.05$ ).

**Çizelge 4.10** Ridge Regresyon Varyans Analizi Tablosu

Model	Serbestlik Derecesi	Kareler Toplamı	Kareler Ortalaması	F	p
Sabit	1	40241.960	40241.960		
Regresyon	6	228.994	38.166	57.780	0.000
Hata	93	61.430	0.660		
Toplam	99	290.426	2.933		
SİGMA	0.812				
R <sup>2</sup>	0.789				
R <sup>2</sup> <sub>d</sub>	0.775				

#### 4.1.2.3 Temel Bileşenler Regresyon Analiz Sonuçları

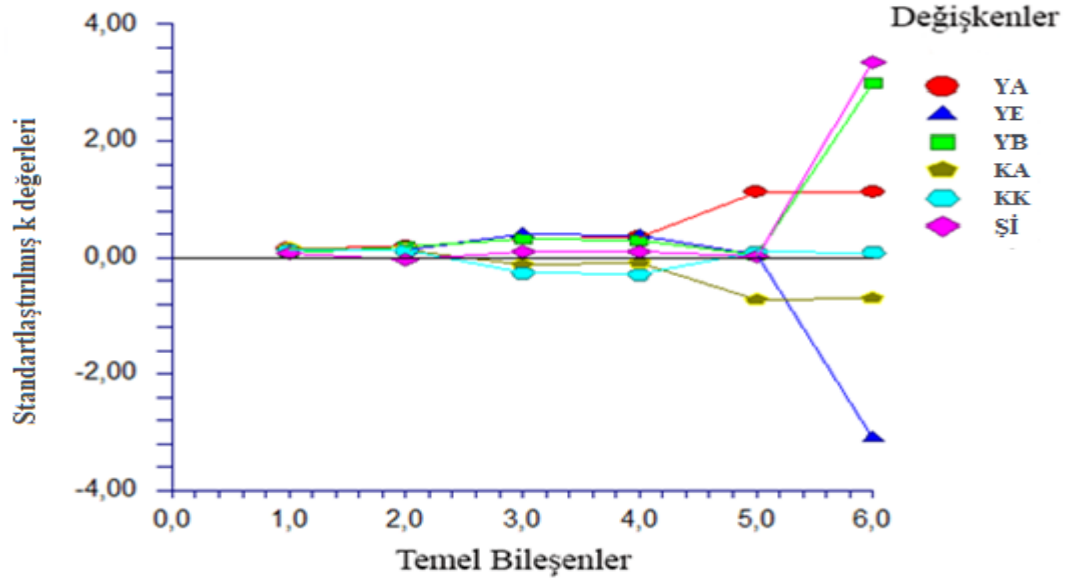
Temel bileşenlere ait regresyon analizi sonuçları verilmiştir (Çizelge 4.11). Çoklu bağlantının oluşmaması için özdeğerler 1'e eşit olmalı ve koşul indeksi 30'dan büyük olmamalıdır. Temel bileşenler regresyonunda çoklu bağlantıyı ortadan kaldırmak için en küçük temel bileşen özvektörler aracılığıyla analizden çıkarılır. Temel bileşenler regresyonunda, temel bileşen sonuçları için iki önemli kriter vardır. Birincisi, özdeğerleri 1'den büyük olan değerler seçilir. İkincisi, toplam varyansın 2/3'ü geçinceye kadar olan kısım alınır (Eker, 2011). Çizelge incelendiğinde bu kriterleri sağlayanların 1 (3.171), 2 (1.554) ve 3 (1.05) olduğu görülmektedir. Çizelgeye göre özdeğeri 0'a en yakın olan ve koşul indeksi 30'dan büyük olan iki temel bileşende doğrusal bağımlılık gözlemlenir.

**Çizelge 4.11** Temel Bileşenler Regresyon Analiz Sonuçları

Numara	Özdeğer	Oran	Kümülatif	Koşul İndeksi
1	<b>3.171</b>	52.850	52.850	1.000
2	<b>1.554</b>	25.900	78.750	2.040
3	<b>1.050</b>	17.500	96.260	3.020
4	0.139	2.310	98.570	22.890
5	0.086	1.430	100.000	36.980
6	0.000	0.000	100.000	12346.480

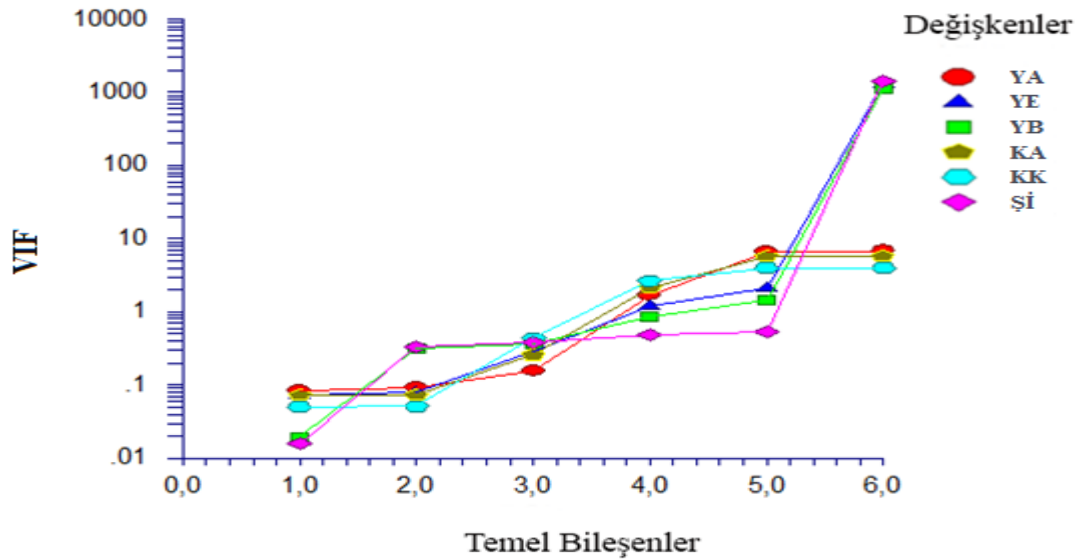
Beta izi grafiği Şekil 4.3'te verilmiştir. Beta izi grafiğinde dikey eksenle standartlaştırılmış betaların, yatay eksenle temel bileşenlerin olduğu gözlemlenir.

Temel bileşenlerin parametresi arttıkça beta değerleri durağanlaşır. Grafikte görüldüğü üzere 5 değerinin üzerinde bağımsız değişkenlerden yumurta eni, yumurta boyu ve şekil indeksi durağanlaşmamıştır.



Şekil 4.3 Ak Ağırlık İçin Beta İzi Grafiği

Şekil 4.4'de temel bileşenlerin varyans büyütme faktörleri verilmiştir. VIF değerlerinin 10'dan küçük olduğunda temel bileşenlerin durağanlaştığı görülür.



Şekil 4.4 Ak Ağırlık İçin Temel Bileşenler VIF Grafiği

Temel bileşenlere ait varyans büyütme değerleri verilmiştir (Çizelge 4.12). Varyans büyütme faktörü 10'dan büyük olan yumurta eni ve şekil indeksi çoklu bağlantı oluşturmaktadır.

**Çizelge 4.12** Temel Bileşenler Varyans Büyütme Değerleri

Temel Bileşen	YA	YE	YB	KA	KK	Şİ
1	0.083	0.073	0.020	0.073	0.051	0.016
2	0.093	0.079	0.324	0.073	0.052	0.337
3	0.159	0.289	0.372	0.260	0.444	0.393
4	1.666	1.183	0.865	2.244	2.739	0.477
5	6.777	2.165	1.480	5.701	4.017	0.534
6	6.841	<b>1310.216</b>	<b>1155.800</b>	5.747	4.027	<b>1480.038</b>

Temel bileşenler regresyonuna ait standardize edilmiş katsayılar verilmiştir (Çizelge 4.13). Bu iki değer çarpıldığında temel bileşenler sonuçları elde edilir.

**Çizelge 4.13** Değişkenlerin Seçilen Temel Bileşenler İle Standardize Edilmiş Katsayıları

Temel Bileşen	YA	YE	YB	KA	KK	Şİ
TB5	1.107	0.051	0.040	-0.712	0.080	0.010
TB6	1.130	-3.167	3.064	-0.693	0.071	3.432

TB: Temel Bileşen

Temel bileşenlere ait katsayı seçimi tablosu aşağıda verilmiştir.  $R^2$  değeri her temel bileşenin bağımlı değişken olan ak ağırlığı açıklama miktarıdır. Çoklu belirtme katsayısı olan  $R^2$  değeri 0 ile 1 arasında değişmektedir. Çizelge 4.14 incelendiğinde özdeğer ile  $R^2$  değerinin paralellik gösterdiği görülmektedir.

**Çizelge 4.14** Temel Bileşenler Katsayı Seçimi

Temel Bileşen	Temel Bileşen Katsayısı	$R^2$	Özdeğer
TB1	0.506	0.277	3.177
TB2	-0.278	0.040	1.554
TB3	1.013	0.364	1.043
TB4	0.099	0.000	0.138
TB5	1.935	0.111	0.086
TB6	9.569	0.007	0.000

TB: Temel Bileşen

Gözardı edilen 2 temel bileşene göre hesaplanan regresyon katsayıları verilmiştir (Çizelge 4.15). Standart hataların küçüldüğü görülmektedir. Ayrıca varyans büyütme

faktörü değerleri küçülmüştür (VIF<10). Yani temel bileşenlerden çoklu bağlantıya sebep olanlar gözardı edilerek çoklu bağlantı sorunu ortadan kaldırılmıştır.

**Çizelge 4.15** Gözardı Edilen 2 Temel Bileşenle Hesaplanan Regresyon Katsayıları

Değişken	Regresyon Katsayısı	Standart Hata	Std. Regresyon Katsayısı	VIF
Sabit	-41.246			
YA	0.257	0.054	0.354	1.666
YE	0.752	0.125	0.381	1.183
YB	0.481	0.086	0.302	0.865
KA	-0.201	0.190	-0.092	2.244
KK	-11.732	3.812	-0.297	2.738
Şİ	0.083	0.037	0.090	0.477

Çizelge 4.13'deki standartlaştırılmış katsayılara göre yumurta ak ağırlığını etkileyen model:

$$-41.24611 + 0.2574323*YA + 0.7520183*YE + 0.4813552*YB - 0.2015728*KA - 11.73254 *KK + 0.08275864*Şİ \quad (4.2)$$

şeklinde bulunur. Modelde,

YA	: Yumurta ağırlığı,
YE	: Yumurta eni,
YB	: Yumurta boyu,
KA	: Kabuk ağırlığı
KK	: Kabuk kalınlığı
Şİ	: Şekil indeksidir.

Temel bileşenler regresyonuna ait varyans analizi sonuçları verilmiştir (Çizelge 4.16). Bağımsız değişkenlerden yumurta eni, yumurta boyu ve şekil indeksi arasında oluşan çoklu bağlantının, temel bileşenler regresyonunda ortadan kaldırıldığı görülmektedir. Bağımsız değişkenler % 66.2 oranında yumurta ak ağırlığını açıklamaktadır.



**Çizelge 4.16** Temel Bileşenler Regresyonu Varyans Analizi Tablosu

<b>Model</b>	<b>SD</b>	<b>Kareler Toplamı</b>	<b>Kareler Ortalaması</b>	<b>F</b>	<b>p</b>
Sabit	1	40241.960	40241.960		
Regresyon	6	198.494	33.082	33.467	0.00
Hata	93	91.931	0.989		
Toplam	99	290.425	2.934		
SİGMA	0.995				
R <sup>2</sup>	0.683				
R <sup>2</sup> -düz	0.662				

#### 4.1.2.4 Kısmi En Küçük Kareler Regresyon Sonuçları

Kısmi en küçük kareler regresyonu varyans analizi verilmiştir (Çizelge 4.17). Minitab 18 paket programında NIPALS algoritması kullanılarak elde edilen analiz sonuçlarına göre elde edilen regresyon modeli istatistiksel olarak önemlidir ( $p < 0.001$ ). Çizelgeye göre düzeltilmiş  $R^2$  değeri 0.780 olarak hesaplanmıştır. Bunun anlamı, bağımlı değişken olan ak ağırlığında meydana gelen değişimi bağımsız değişkenler %78 oranında açıklamaktadır.

**Çizelge 4.17** Kısmi En Küçük Kareler Regresyonu Varyans Analizi Tablosu

<b>Model</b>	<b>Serbestlik Derecesi</b>	<b>Kareler Toplamı</b>	<b>Kareler Ortalaması</b>	<b>F</b>	<b>p</b>
Regresyon	4	230.708	57.671	91.76	0.000
Hata	95	59.717	0.628		
Toplam	99	290.425	2.933		
R <sup>2</sup>	0.794				
R <sup>2</sup> -düz	0.780				

Kısmi en küçük kareler regresyonu için model seçimi ve modelin geçerliliği verilmiştir (Çizelge 4.18). Model seçimi, ideal olan bileşen sayısını belirler. Yumurta ak ağırlığı için seçilen 4 bileşenli model değişiminde dördüncü bileşen bağımsız değişkenlere ait varyansı %97.7 açıklamaktadır. Modelin geçerliğinin ve kestirimdeki başarısının bir ölçüsü olarak kullanılan PRESS istatistiği, regresyon modeli için kestirim hata kareler toplamıdır. Genellikle, PRESS değeri ne kadar küçükse modelin kestirim yeteneği o kadar iyidir (Yeniay ve Göktaş, 2002). Kestirilen  $R^2$  ( $R^2$  pred) modele yeni gözlem eklendiğinde bağımlı değişkenin nasıl etkileneceğini belirler. Yüksek olan kestirilen  $R^2$  değeri modeli daha iyi anlatma özelliğine sahiptir. PRESS, 6 tane PRESS artığının

kareleri toplamını ifade eder.  $R^2$  (pred); yani  $R^2$  kestirim ise, kestirim için yaklaşık bir değer bulunması amacıyla kullanılır.  $R^2$ 'nin kestirim gücüne göre bu modelden yeni gözlemlerin kestirilmesindeki değişkenliğin %76'sı açıklanır.

**Çizelge 4.18** Kısmi En Küçük Kareler Model Seçimi ve Modelin Geçerliliği

Bileşen	X Varyansı	HKT	$R^2$	PRESS	$R^2$ (pred)
1	0.478	133.405	0.540	142.097	0.5107
2	0.712	86.151	0.703	93.515	0.678
3	0.934	79.703	0.725	88.389	0.695
4	0.977	59.717	0.794	68.895	0.762
5		59.682	0.794	70.940	0.755
6		57.406	0.802	68.962	0.762

HKT: Hata kareler toplamı

Kısmi en küçük kareler regresyonunda birden fazla Y değişkeni kullanılabilir. Birden fazla Y değişkeni birbirleriyle ilişkili ise ayrı model kurulmasına gerek yoktur, analiz tüm Y değişkenleri için yapılır. Bu kısmi en küçük kareler regresyonunun avantajıdır.

#### 4.1.3 Sarı Ağırlığı Bağımlı Değişkeni İçin Elde Edilen Sonuçlar

##### 4.1.3.1 En Küçük Kareler Analiz Sonuçları

En küçük kareler regresyon analizi için varyans analizi sonuçları verilmiştir (Çizelge 4.19). Yapılan varyans analizi sonucuna göre F önemlidir ( $p < 0.01$ ). Tabloya göre  $R^2$  değeri 0.412 bulunmuştur. Fakat çoklu regresyon analizi olduğu için düzeltilmiş  $R^2$  değeri dikkate alınır.  $R^2_d$  değeri 0.374'tür. Yani bağımsız değişkenler (yumurta ağırlığı, yumurta eni, yumurta boyu, kabuk ağırlığı, kabuk kalınlığı, şekil indeksi) bağımlı değişkendeki (sarı ağırlık) değişimi %37.4 açıklamaktadır. Ayrıca Durbin-Watson 1.861 olarak hesaplanmış ve otokorelasyon olmadığını göstermiştir.

**Çizelge 4.19** En Küçük Kareler Varyans Analizi Tablosu

Model	Serbestlik Derecesi	Kareler Toplamı	Kareler Ortalaması	F	p
Regresyon	6	40.277	6.713	10.875	0.000
Hata	93	57.406	0.617		
Genel	99	97.684			
$R^2$	0.412				
$R^2$ -düz	0.374				

En küçük kareler regresyon analizi için regresyon analizi sonuçları, anlamlılık değeri ve modelin katsayıları verilmiştir (Çizelge 4.20). Yumurta ağırlığının katsayısı 0.178 bulunmuştur ve önemlidir ( $p < 0.05$ ). Kabuk ağırlığının katsayısı 0.511 bulunmuştur ve

model için önemlidir ( $p < 0.05$ ). Bağımsız değişkenlerden yumurta eni, yumurta boyu ve şekil indeksi bağımsız değişkenlerinin varyans büyütme faktörü değerleri yüksek çıkmıştır ( $VIF > 10$ ). Varyans büyütme faktörleri yüksek olan yumurta eni, yumurta boyu ve şekil indeksi bağımsız değişkenlerinin tolerans değerleri çok küçüktür.

**Çizelge 4.20** En Küçük Kareler Çoklu Regresyon Analizi Sonuçları

Değişken	Katsayı	Standart Hata	t	p	Tolerans Değeri	VIF
Sabit	252.505	125.061	2.019	0.046		
YA	0.178	0.088	2.034	0.045	0.146	6.842
YE	6.243	3.290	1.898	0.061	0.001	<b>1310.216</b>
YB	-4.891	2.502	-1.955	0.054	0.001	<b>1155.800</b>
KA	0.511	0.241	2.119	0.037	0.174	5.747
KK	-2.787	3.653	-0.763	0.448	0.248	4.027
Şİ	-3.165	1.635	-1.935	0.056	0.001	<b>1480.038</b>

Korelasyonlara ait özdeğerler ve koşul indeksi değerleri verilmiştir (Çizelge 4.21). Özdeğerler 1'e eşit olduğunda çoklu bağlantı oluşmaz. Fakat özdeğerlerden en az biri sıfıra yakın olduğu için çoklu bağlantı sorunu gözlemlenir. Çizelge incelendiğinde; 3, 4, 5, 6 ve 7 numaralı özdeğerlerde çoklu bağlantı sorunu saptanır. Koşul indeksi 30'dan büyük olduğunda birden fazla çoklu bağlantı sorunu oluştuğu için 3, 4, 5, 6 ve 7 güçlü çoklu bağlantıya sahiptir.

**Çizelge 4.21** Korelasyonların Özdeğerleri ve Koşul İndeks Değerleri

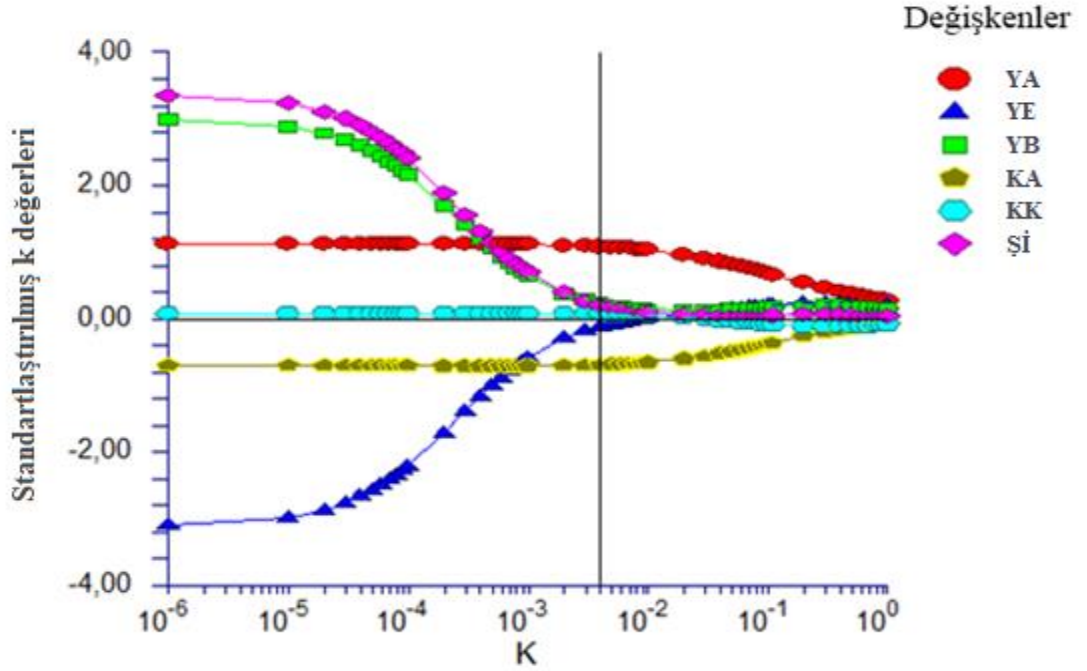
Numara	Özdeğer	Koşul İndeksi
1	6.984	1.000
2	0.012	24.030
3	0.002	<b>56.809</b>
4	0.001	<b>88.453</b>
5	0.001	<b>104.261</b>
6	0.000	<b>329.251</b>
7	0.000	<b>8385.127</b>

Yumurta kalite değişkenleri için çoklu bağlantı varlığını belirleyen yöntemler incelendiğinde (korelasyon matrisi, varyans büyütme değeri, tolerans, özdeğerler ve koşul indeksi) bağımsız değişkenler arasında çoklu bağlantı sorununun oluştuğu görülmüştür. Yansız olan en küçük kareler regresyonu çoklu bağlantı sorununu ortadan kaldıramadığı için sonuçlar tutarsız olur.

#### 4.1.3.2 Ridge Regresyon Analiz Sonuçları

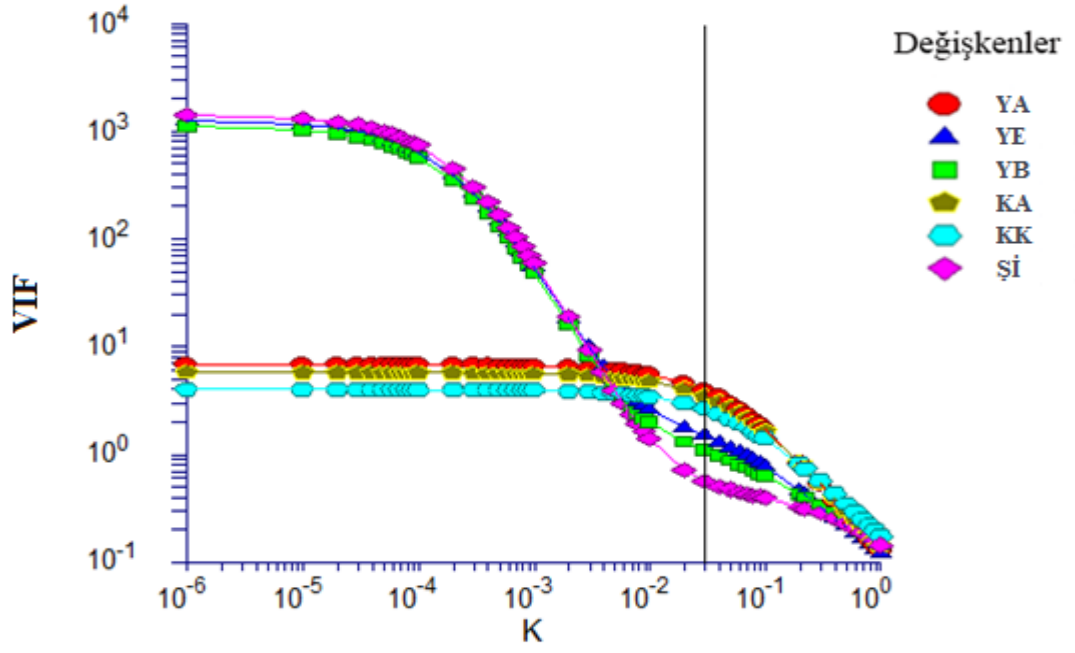
Ridge izi grafiği verilmiştir (Şekil 4.5). Ridge izi grafiğinde dikey ekseninde standartlaştırılmış k değerlerinin, yatay ekseninde k parametresinin olduğu gözlemlenir.

k parametresi arttıkça k değerleri durağanlaşır. Grafik incelendiğinde,  $k=0.004$  sabitinden sonra regresyon katsayıları daha durağan hale gelir.



Şekil 4.5 Sarı Ağırlık İçin Ridge İzini Grafiği

Şekil 4.6'da k parametresinin VIF değerlerindeki etkisi gözlenir.  $k=0.004$  sabitinden sonra VIF değerleri 10'un altındadır.



**Şekil 4.6** Sarı Ağırlık İçin Ridge Regresyon VIF Grafiği

Standartlaştırılmış Ridge parametresi  $k$  verilmiştir (Çizelge 4.22). Grafiğe göre yumurta sarı ağırlığının pozitif ilişkili olması beklenen bağımsız değişkenler için  $k=0.004$  sabiti seçilerek çoklu bağlantı sorunu ortadan kaldırılmıştır.

**Çizelge 4.22** Standartlaştırılmış Ridge Regresyon Değerleri

<b>k</b>	<b>YA</b>	<b>YE</b>	<b>YB</b>	<b>KA</b>	<b>KK</b>	<b>Şi</b>
0.001	0.447	1.024	-1.112	0.446	-0.149	-1.199
0.002	0.448	0.535	-0.650	0.447	-0.149	-0.677
0.003	0.447	0.348	-0.470	0.446	-0.147	-0.475
0.004	0.446	0.249	-0.375	0.445	-0.146	-0.367
0.005	0.444	0.188	-0.315	0.444	-0.144	-0.301
0.006	0.442	0.147	-0.275	0.442	-0.142	-0.255
0.007	0.440	0.118	-0.245	0.441	-0.140	-0.222
0.008	0.438	0.096	-0.222	0.439	-0.138	-0.197
0.009	0.436	0.079	-0.204	0.437	-0.137	-0.177
0.010	0.434	0.066	-0.189	0.435	-0.135	-0.161
0.020	0.416	0.012	-0.116	0.419	-0.118	-0.086
0.030	0.400	0.001	-0.087	0.405	-0.102	-0.058
0.040	0.386	-0.001	-0.069	0.391	-0.089	-0.043
0.050	0.373	0.001	-0.056	0.380	-0.077	-0.033
0.060	0.362	0.005	-0.046	0.369	-0.066	-0.026
0.070	0.351	0.008	-0.037	0.359	-0.056	-0.020
0.080	0.342	0.012	-0.030	0.350	-0.047	-0.015
0.090	0.333	0.016	-0.024	0.341	-0.039	-0.011
0.100	0.325	0.020	-0.019	0.334	-0.032	-0.008
0.200	0.270	0.048	0.013	0.280	0.016	0.010
0.222	0.262	0.052	0.017	0.271	0.023	0.013
0.300	0.239	0.063	0.028	0.248	0.040	0.019
0.400	0.218	0.071	0.036	0.226	0.054	0.023
0.500	0.203	0.077	0.041	0.210	0.062	0.026
0.600	0.191	0.080	0.044	0.197	0.067	0.028
0.700	0.181	0.082	0.046	0.187	0.071	0.029
0.800	0.173	0.083	0.047	0.178	0.073	0.030
0.900	0.166	0.083	0.047	0.170	0.074	0.030
1.000	0.159	0.083	0.048	0.164	0.074	0.030

Ridge parametresi  $k$ 'ye bağlı olarak değişen parametreler Çizelge 4.23'de verilmiştir.  $k=0$  sabiti için  $R^2$  değeri en büyük olduğu için,  $k$  sabiti sifıra yakın olmalıdır. Sigma, hata kareler ortalamasının kareköküdür, en küçük değerini en küçük kareler regresyonunda aldığı için  $k$  değeri bu değerden çok sapmamalıdır. B'B, standartlaştırılmış katsayıların kareler toplamıdır ve  $k$  sabitine göre bu değer durağanlaşmalıdır. Ortalama varyans büyütme faktörü, varyans büyütme faktörü değerlerinin ortalamasıdır. En büyük varyans büyütme faktörü,  $k$  sabitinin varyans büyütme değerinin en büyüğünü verir. Bu parametrelerin beklentilerine göre en uygun  $k=0.004$  olarak seçilmiştir.

**Çizelge 4.23** Ridge Regresyon k Analiz Tablosu

<b>k</b>	<b>R<sup>2</sup></b>	<b>SİGMA</b>	<b>B'B</b>	<b>Ort. VIF</b>	<b>En Büyük VIF</b>
0.001	0.394	0.798	4.145	30.502	61.568
0.002	0.392	0.799	1.589	11.747	19.458
0.003	0.391	0.800	0.988	7.320	10.094
<b>0.004</b>	<b>0.390</b>	<b>0.801</b>	<b>0.755</b>	<b>5.590</b>	<b>6.732</b>
0.005	0.389	0.801	0.640	4.721	6.148
0.006	0.388	0.802	0.573	4.212	6.030
0.007	0.388	0.802	0.530	3.879	5.915
0.008	0.387	0.802	0.501	3.644	5.803
0.009	0.387	0.803	0.479	3.466	5.695
0.010	0.386	0.803	0.462	3.326	5.590
0.020	0.382	0.806	0.384	2.610	4.687
0.030	0.379	0.808	0.345	2.231	3.996
0.040	0.375	0.810	0.317	1.955	3.453
0.050	0.372	0.812	0.293	1.738	3.019
0.060	0.369	0.814	0.274	1.562	2.666
0.070	0.367	0.816	0.257	1.417	2.375
0.080	0.364	0.817	0.243	1.294	2.131
0.090	0.362	0.819	0.230	1.190	1.926
0.100	0.360	0.820	0.219	1.101	1.750
0.200	0.342	0.832	0.154	0.620	0.846
0.222	0.338	0.834	0.146	0.565	0.752
0.300	0.328	0.840	0.125	0.432	0.570
0.400	0.316	0.848	0.109	0.333	0.436
0.500	0.306	0.854	0.097	0.272	0.353
0.600	0.297	0.860	0.089	0.230	0.295
0.700	0.288	0.865	0.082	0.200	0.253
0.800	0.280	0.870	0.077	0.176	0.221
0.900	0.273	0.874	0.072	0.157	0.195
1.000	0.266	0.878	0.068	0.142	0.174

k=0.004 sabitine göre Ridge ve en küçük kareler regresyon analizi karşılaştırılmıştır (Çizelge 4.24). Çoklu bağlantı sorununu ortadan kaldıran Ridge analizinde standart hatada azalma meydana gelmiştir. Ayrıca yapılan analiz sonucunda düzeltilmiş R<sup>2</sup> en küçük kareler regresyonu için 0.374 iken, Ridge için 0.350 bulunmuştur. Yani Ridge regresyon R<sup>2</sup> yi çok değiştirmeden daha düşük standart hatalı güvenilir sonuçlar elde etmiştir.

**Çizelge 4.24** Ridge Katsayısı  $k=0.004$  İçin Ridge Regresyon ve En Küçük Kareler Karşılaştırılması

Değişken	Ridge Katsayısı	OLS Katsayısı	Std. Ridge	Std. OLS	Ridge SH	OLS SH
Sabit	25.676	252.504				
YA	0.189	0.178	0.450	0.423	0.086	0.087
YE	0.289	6.243	0.252	5.460	0.240	3.290
YB	-0.350	-4.890	-0.379	-5.282	0.176	2.502
KA	0.540	0.510	0.428	0.404	0.236	0.240
KK	-2.941	-2.786	-0.128	-0.122	3.609	3.653
Şİ	-0.198	-3.165	-0.370	-5.919	0.104	1.635
$R^2$	0.389	0.412				
$R^2$ -düz	0.350	0.374				
SİGMA	0.800	0.785				

Ridge regresyonun  $k=0.004$  için katsayıları verilmiştir. Aşağıdaki katsayılara göre yumurta sarı ağırlığını etkileyen model:

$$25.67637 + 0.1898782*YA + 0.2886091*YE - 0.350875*YB + 0.540656*KA - 2.94124*KK - 0.1982489*Şİ \quad (4.3)$$

şeklinde bulunur. Modelde;

YA : Yumurta ağırlığı,  
 YE : Yumurta eni,  
 YB : Yumurta boyu,  
 KA : Kabuk ağırlığı  
 KK : Kabuk kalınlığı  
 Şİ : Şekil indeksidir.

Ridge regresyona ait katsayılar verilmiştir (Çizelge 4.25). Standart hataların düştüğü, varyans büyütme faktörlerinin 10'un altına düştüğü çizelgede görülmektedir.

**Çizelge 4.25** Ridge Regresyon Katsayıları

Değişken	Katsayı	Standart Hata	Std. Reg Katsayısı	VIF
Sabit	25.676			
YA	0.189	0.008	0.450	6.252
YE	0.289	0.240	0.252	6.681
YB	-0.351	0.176	-0.375	5.504
KA	0.540	0.236	0.427	5.291
KK	-2.941	3.609	-0.128	3.776
Şİ	-0.198	0.104	-0.370	5.785



$k=0.004$  sabiti için Ridge regresyon varyans analizi yapılmış ve sonuçları verilmiştir (Çizelge 4.26). Yumurta sarı ağırlığında meydana gelen değişimin %35'i yumurtanın diğer kalite değişkenleri tarafından açıklanmıştır.

**Çizelge 4.26** Ridge Regresyon Varyans Analizi Tablosu

Değişken	Serbestlik Derecesi	Kareler Toplamı	Kareler Ortalaması	F	p
Sabit	1	18480.77	18480.77		
Model	6	37.951	6.325	9.848	0.000
Hata	93	59.723	0.642		
Toplam	99	97.683	0.987		
SİGMA	0.80				
R <sup>2</sup>	0.389				
R <sup>2</sup> -düz	0.350				

#### 4.1.3.3 Temel Bileşenler Regresyon Analiz Sonuçları

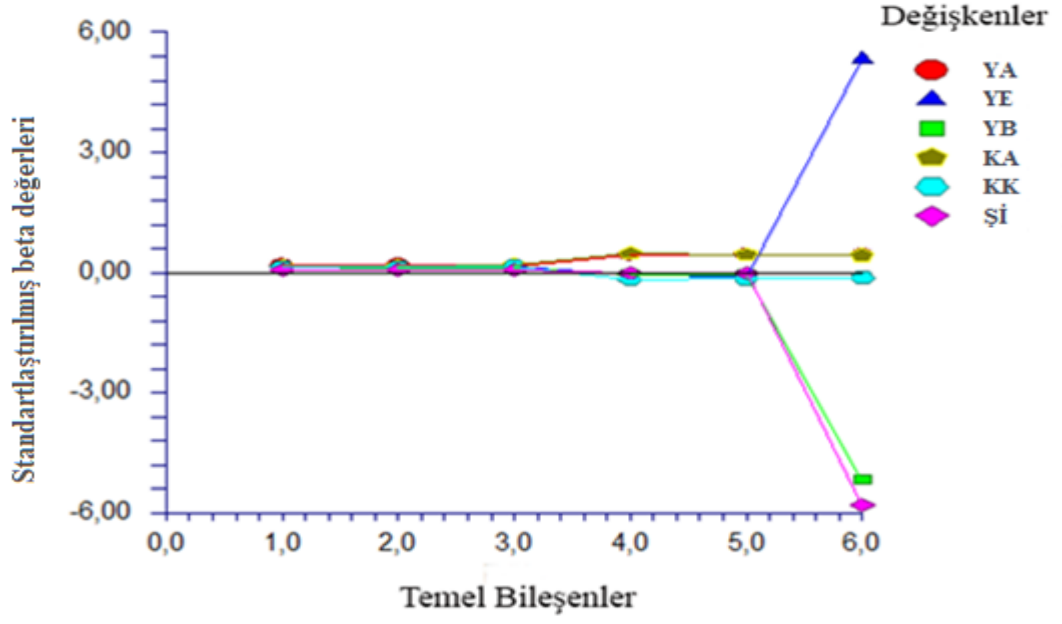
Temel bileşenlere ait regresyon analizi sonuçları verilmiştir (Çizelge 4.27). Çoklu bağlantının oluşmaması için özdeğerler 1'e eşit olmalı ve koşul indeksi 30'dan büyük olmamalıdır. Temel bileşenler regresyonunda çoklu bağlantıyı ortadan kaldırmak için en küçük temel bileşen özvektörler aracılığıyla analizden çıkarılır. Temel bileşenler regresyonunda, temel bileşen sonuçları için iki önemli kriter vardır. Birincisi, özdeğerleri 1'den büyük olan değerler seçilir. İkincisi, toplam varyansın 2/3'ü geçinceye kadar olan kısım alınır (Eker, 2011). Bu kriterleri sağlayanların 1 (3.171), 2 (1.554) ve 3 (1.05) olduğu görülmektedir. Özdeğerleri 0'a yakın olan, aynı zamanda koşul indeksi 30'un üzerinde olan iki bileşende doğrusal bağlantı bulunur.

**Çizelge 4.27** Temel Bileşenler Regresyon Analiz Sonuçları

Numara	Özdeğer	Oran	Kümülatif	Koşul İndeksi
1	<b>3.171</b>	52.850	52.850	1.000
2	<b>1.554</b>	25.900	78.750	2.040
3	<b>1.050</b>	17.500	96.260	3.020
4	0.139	2.310	98.570	22.890
5	0.086	1.430	100.000	36.980
6	0.000	0.000	100.000	12346.480

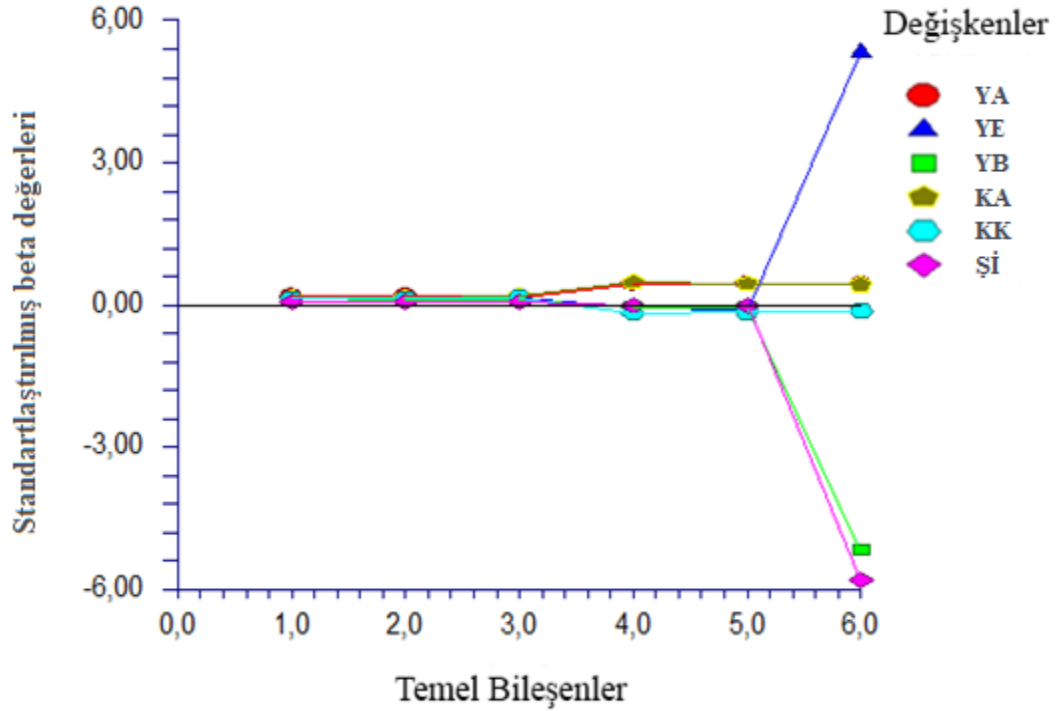
Sarı ağırlığa ait beta izi grafiği verilmiştir (Şekil 4.7). Beta izi grafiğinde dikey eksen standartlaştırılmış betaların, yatay ekseninde temel bileşenlerin olduğu gözlemlenir. Temel bileşenlerin parametresi arttıkça beta değerleri durağanlaşır. Grafikte görüldüğü

üzere 5 değerinde bağımsız değişkenlerden yumurta ağırlığı, kabuk kalınlığı, kabuk ağırlığı durağanlaşmıştır.



Şekil 4.7 Sarı Ağırlık İçin Beta İzi Grafiği

Sarı ağırlığa ait temel bileşenlerin VIF değerleri verilmiştir (Şekil 4.8). VIF değerlerinin 10'dan küçük olduğunda temel bileşenlerin durağanlaştığı görülür.



**Şekil 4.8** Sarı Ağırlık İçin Temel Bileşenler VIF Grafiği

Temel bileşenlere ait varyans büyütme faktörleri verilmiştir (Çizelge 4.28). Yumurta eni, yumurta boyu ve şekil indeksi bağımsız değişkenlerinin VIF değerleri oldukça büyüktür ( $VIF > 10$ ).

**Çizelge 4.28** Temel Bileşenler Varyans Büyütme Değerleri

Temel Bileşen	YA	YE	YB	KA	KK	Şİ
1	0.083	0.072	0.019	0.073	0.051	0.016
2	0.093	0.079	0.324	0.073	0.052	0.336
3	0.159	0.289	0.372	0.259	0.443	0.393
4	1.666	1.183	0.865	2.244	2.738	0.476
5	6.777	2.165	1.480	5.701	4.017	0.534
6	6.841	1310.215	1155.800	5.746	4.027	1480.037

Temel bileşenler regresyonuna ait standardize edilmiş katsayılar verilmiştir (Çizelge 4.29). Bu iki değer çarpıldığında temel bileşenler sonuçları elde edilir.

**Çizelge 4.29** Değişkenlerin Seçilen Temel Bileşenler İle Standardize Edilmiş Katsayıları

Temel Bileşen	YA	YE	YB	KA	KK	Şİ
TB5	0.462	-0.085	-0.073	0.437	-0.137	0.021
TB6	0.423	5.461	-5.282	0.404	-0.122	-5.919

TB: Temel Bileşen

Temel bileşenlere ait katsayı seçimi tablosu aşağıda verilmiştir.  $R^2$  değeri her temel bileşenin bağımlı değişken olan sarı ağırlığı açıklama miktarıdır. Çoklu belirtme katsayısı olan  $R^2$  değeri 0 ile 1 arasında değişmektedir. Çizelge 4.30 incelendiğinde özdeğer ile  $R^2$  değerinin paralellik gösterdiği görülmektedir.

**Çizelge 4.30** Temel Bileşenler Katsayı Seçimi

Temel Bileşen	Temel Bileşen Katsayısı	$R^2$	Özdeğer
TB1	0.325	0.341	3.171
TB2	-0.026	0.001	1.554
TB3	-0.044	0.002	1.043
TB4	0.563	0.044	0.138
TB5	0.066	0.000	0.086
TB6	-9.562	0.023	0.000

TB: Temel Bileşen

Gözardı edilen 2 temel bileşene göre hesaplanan regresyon katsayıları verilmiştir (Çizelge 4.31). Standart hataların küçüldüğü görülmektedir. Ayrıca varyans büyütme faktörü değerleri küçülmüştür ( $VIF < 10$ ). Yani temel bileşenlerden çoklu bağlantıya sebep olanlar gözardı edilerek çoklu bağlantı sorunu ortadan kaldırılmıştır.

**Çizelge 4.31** Gözardı Edilen 2 Temel Bileşenle Hesaplanan Regresyon Katsayıları

Değişken	Regresyon Katsayısı	Standart Hata	Std. Regresyon Katsayısı	VIF
Sabit	10.765			
YA	0.176	0.044	0.418	1.665
YE	0.074	0.100	-0.065	1.183
YB	-0.053	0.069	-0.058	0.865
KA	0.597	0.153	0.473	2.244
KK	-3.643	3.073	-0.159	2.738
Şİ	-0.008	0.029	-0.016	0.476

Çizelge 4.28'deki standartlaştırılmış katsayılara göre yumurta sarı ağırlığını etkileyen model:

$$10.7655 + 0.1762776*YA - 0.07473396*YE - 0.05347648*YB + 0.5977504*KA - 0.364338 *KK - 0.008688837*Şİ \quad (4.4)$$

şeklinde bulunur. Modelde;

---

YA : Yumurta ağırlığı,  
YE : Yumurta eni,  
YB : Yumurta boyu,  
KA : Kabuk ağırlığı  
KK : Kabuk kalınlığı  
Şİ : Şekil indeksidir.

---

Temel bileşenler regresyonu varyans analizi sonuçları verilmiştir (Çizelge 4.32). Bağımsız değişkenlerden yumurta eni, yumurta boyu ve şekil indeksi arasında oluşan çoklu bağlantı, temel bileşenler regresyonunda ortadan kaldırılmıştır. Bağımsız değişkenler % 34.9 oranında yumurta sarı ağırlığını açıklamaktadır.

**Çizelge 4.32** Temel Bileşenler Regresyonu Varyans Analizi Tablosu

---

<b>Model</b>	<b>SD</b>	<b>Kareler Toplamı</b>	<b>Kareler Ortalaması</b>	<b>F</b>	<b>p</b>
Sabit	1	18480.77	18480.77		
Regresyon	6	37.944	6.324	9.844	0.00
Hata	93	59.739	0.642		
Toplam	99	97.683	0.986		
Sigma		0.801			
R <sup>2</sup>		0.388			
R <sup>2</sup> -düz		0.349			

---

#### 4.1.3.4 Kısmi En Küçük Kareler Regresyon Analiz Sonuçları

Kısmi en küçük kareler regresyonu varyans analizi verilmiştir (Çizelge 4.33). Minitab, NIPALS algoritmasını kullanarak analiz sonuçları elde etmiştir. Çizelge 4.33 incelendiğinde, R<sup>2</sup>-düz değerinin 0.304 olduğu görülmektedir (p<0.001). Bağımlı değişken olan sarı ağırlığı, %30.4 oranında açıklamaktadır.

**Çizelge 4.33** Kısmi En Küçük Kareler Regresyonu Varyans Analizi Tablosu

Model	SD	Kareler Toplamı	Kareler Ort.	F	p
Regresyon	1	33.8725	33.8725	51.78	0.000
Hata	98	63.8114	0.6511		
Toplam	99	97.6839			
R <sup>2</sup>	0.346				
R <sup>2</sup> -düz	0.304				

Kısmi en küçük kareler regresyonu için model seçimi ve modelin geçerliliği verilmiştir (Çizelge 4.34). Model seçimi, ideal olan bileşen sayısını belirler. Yumurta sarı ağırlığı için seçilen 4 bileşenli model değişiminde dördüncü bileşen varyansın %52.8'ini açıklamaktadır. Kestirilen R<sup>2</sup> (R<sup>2</sup> pred) modele yeni gözlem eklendiğinde bağımlı değişkenin nasıl etkileneceğini belirler. Yüksek olan kestirilen R<sup>2</sup> değeri modeli daha iyi anlatma özelliğine sahiptir. PRESS, 6 tane PRESS artığının kareleri toplamını ifade eder. R<sup>2</sup> (pred); yani R<sup>2</sup> kestirim ise, kestirim için yaklaşık bir değer bulunması amacıyla kullanılır. R<sup>2</sup>'nin kestirim gücüne göre bu modelden yeni gözlemlerin kestirilmesindeki değişkenliğin %30.9'u açıklanır.

**Çizelge 4.34** Kısmi En Küçük Kareler Model Seçimi ve Modelin Geçerliliği

Bileşen	X Varyansı	HKT	R <sup>2</sup>	PRESS	R <sup>2</sup> (pred)
1	0.528	63.914	0.345	67.417	0.309
2		61.725	0.368	71.497	0.268
3		59.777	0.388	70.488	0.278
4		59.569	0.390	69.827	0.285
5		59.520	0.390	71.407	0.268
6		57.357	0.412	69.466	0.288

HKT: Hata kareler toplamı

Ak ağırlık için en küçük kareler ve yanlı tahmin edici yöntemler karşılaştırılmış ve sonuçları verilmiştir (Çizelge 4.35). Sonuçların güvenilirliği için R<sup>2</sup> değeri yüksek, HKO değeri düşük olmalıdır. Sonuçlar incelendiğinde, en yüksek R<sup>2</sup> değeri en küçük kareler regresyonunda almasına rağmen gösterdiği yanlılıktan dolayı sonuçlar tutarlı değildir. En düşük hata kareler ortalaması en küçük kareler regresyonunda olmasına rağmen, yanlı regresyon yöntemleri değişkenleri standart hataları küçülterek kullanır. Ak ağırlık için yanlı regresyon analizlerinden yüksek düzeltilmiş R<sup>2</sup> ve düşük HKO değeri kısmi en küçük kareler analizinde gözlemlenmiştir.

**Çizelge 4.35 Ak Ağırlık İçin Analiz Sonuçlarının İncelenmesi**

Yöntem	HKO	R <sup>2</sup> -düz	AIC
OLS	0.617	0.790	-42.476
RR	0.660	0.775	-35.739
PCR	0.989	0.662	4.908
PLS	0.628	0.780	-40.709

Sarı ağırlık için en küçük kareler ve yanlı tahmin edici yöntemlerine ait performans kriterleri değerleri Çizelge 4.36’de verilmiştir. Sonuçlar incelendiğinde, en yüksek R<sup>2</sup> değeri en küçük kareler regresyonunda almasına rağmen gösterdiği yanlılıktan dolayı sonuçlar tutarlı değildir. Sarı ağırlık için yanlı regresyon analizlerinden Ridge regresyon ve temel bileşenler regresyonunun hata kareler ortalaması aynı olmasına karşın, Ridge regresyonun düzeltilmiş R<sup>2</sup> değeri yüksektir.

**Çizelge 4.36 Sarı Ağırlık İçin Analiz Sonuçlarının İncelenmesi**

Yöntem	HKO	R <sup>2</sup> -düz	AIC
OLS	0.617	0.374	-42.476
RR	0.642	0.350	-38.660
PCR	0.642	0.349	-31.294
PLS	0.651	0.304	-37.112

## 4.2 Tartışma

Bu tez çalışmasında, Ridge regresyon analizi sonucunda, Ridge parametresi  $k=0.004$  sabitine göre Ridge katsayılarının standart hataları ve VIF değerleri azalmıştır. En küçük kareler regresyon analizi sonucunda; ak ve sarı ağırlık için belirleme katsayısı sırasıyla %80.2 ve %41.2’dir. Ridge regresyon analiz sonuçlarına göre ise; ak ve sarı ağırlık için belirleme katsayısı sırasıyla %78.9 ve %38.9’dur. Yani Ridge regresyon R<sup>2</sup>’de çok fazla değişime sebep olmadan, çölu bağlantı sorununu ortadan kaldırarak güvenilir sonuçlar elde etmiştir. Bu bulgular, Yılmaz (2018)’in *Abra alba* türünün, kabuk ağırlığı ile morfolojik özellikler (kabuk boyu, kabuk eni ve kabuk kalınlığı) arasında çoklu bağlantıyı belirleyip, Ridge regresyon analizi uyguladığı çalışma bulguları ile uyumludur.

Üçkardeş ve arkadaşları (2012), yapmış oldukları çalışmada, yumurta ak indeksinin tahmininde kullanılan diğer kalite özellikleri arasında çoklu bağlantı oluştuğunu tespit etmiş ve verilere Ridge regresyon analizi uygulamışlardır. Çalışmalarında çoklu bağlantı sorunu varlığında, en küçük kareler yerine Ridge regresyon analizinin

kullanılmasının daha tutarlı sonuçlar verdiğini bildirmişlerdir. Bu çalışmanın bulguları ile çalışma bulgularımız benzerlik göstermektedir.

Çalışmada veri setine temel bileşenler regresyonu uygulanarak değişkenlerin seçilen temel bileşenler ile standartlaştırılmış katsayılar elde edilmiştir. Temel bileşenler katsayı seçimi yapıldıktan sonra özdeğerler ve  $R^2$  değerleri paralellik göstermiştir. Göz ardı edilen 2 temel bileşene göre tekrar regresyon katsayıları hesaplanmış ve standart hataları ile VIF değerlerinin düştüğü görülmüştür. Temel bileşenler regresyonu ve kısmi en küçük kareler regresyonu karşılaştırılmıştır. Temel bileşenler regresyon analizi sonucunda ak ve sarı ağırlık için sırasıyla  $R^2$ ; %68.3 ve %38.8'dir. Kısmi en küçük kareler regresyon analizi sonucunda ak ve sarı ağırlık için sırasıyla  $R^2$ ; %79.4 ve %34.6'dır. Elde edilen sonuçlarda yumurta ak ağırlığında kısmi en küçük kareler regresyonu, temel bileşenler regresyonundan daha güvenilir sonuçlar ortaya koymaktadır. Bu sonuç, Akyürek ve Akkol (2018)'un yumurta dış kalite özelliklerini kullanarak iç kalite özelliklerinden yumurta ak ve sarı ağırlığı için bir tahmin modeli oluşturdukları çalışmadan elde edilen bulgular ile uyum göstermektedir



## 5. SONUÇ ve ÖNERİLER

Bağımsız değişkenler arasında çoklu bağlantı sorununun varlığı, en küçük kareler regresyonunun kullanıldığı doğrusal regresyon analizinin güvenilirliğini olumsuz yönde etkilemektedir. En belirgin etki tahmin edilen regresyon katsayılarının standart hatalarında gözlenmektedir. HKO ve  $R^2$  değerlerinde belirgin bir değişim söz konusu olmamasına rağmen, regresyon katsayılarının standart hata değerlerinde belirgin bir büyümeye neden olmaktadır. Bu sebeple, bağımsız değişkenler arasında çoklu bağlantı olması durumunda, en küçük kareler regresyonu yerine “yanlı regresyon yöntemleri” olarak isimlendirilen farklı tahmin edicilere sahip regresyon yöntemlerinin kullanımı önerilmektedir.

Bu tez çalışmasında, yanlı regresyon yöntemleri isimlendirilen “Ridge regresyon analizi”, “temel bileşenler regresyon analizi” ve “kısmi en küçük kareler regresyon analizi” incelenmiştir. Bağımsız değişkenler arasında çoklu bağlantı olduğu belirlenen gerçek bir veri seti kullanılmış ve oluşturulan regresyon modellerinde; katsayılara ait standart hata, HKO,  $R^2$ -düz ve AIC değerleri hesaplanmıştır. Elde edilen bulgular, HKO,  $R^2$ -düz ve AIC değerleri bakımından küçük farklılıklar gözlenmesine rağmen, en küçük kareler regresyon analizi ile tahmin edilen katsayıların standart hatalarının yanlı regresyon yöntemlerine göre daha büyük olduğunu göstermiştir. Kendi aralarında sonuçları benzerlik gösteren yanlı regresyon yöntemleri, çoklu bağlantı varlığından etkilenmeyerek daha tutarlı katsayı tahminleri yapılmasını sağlamıştır.

Sonuç olarak, değişkenler arasında sebep sonuç ilişkilerinin araştırıldığı veri setlerinde değişkenler arasında çoklu bağlantı varsa, en küçük kareler regresyonu yerine yanlı regresyon yöntemlerinin kullanılmasının regresyon katsayılarının standart hatalarını normalleştirdiği dolayısıyla daha güvenilir sonuçlar verdiği ortaya konulmuştur.

## 6. KAYNAKLAR

- Akyürek, S., & Akkol, S. (2018). Yumurta İç Kalite Özelliklerinin Kısmi En Küçük Kareler Regresyonu Kullanılarak Tahmin Edilmesi. *Yüzüncü Yıl Üniversitesi Tarım Bilimleri Dergisi*, 28(4): 473-481.
- Albayrak, A.S. (2005). Çoklu Doğrusal Bağlantı Halinde En Küçük Kareler Tekniğinin Alternatifi Yanlı Tahmin Teknikleri ve Bir Uygulama. *Zonguldak Kara Elmas Üniversitesi Sosyal Bilimler Dergisi*, 1, 105-126.
- Albayrak, A.S. (2014). Otokeralasyon Durumunda En Küçük Kareler Tekniğinin Alternatifi Otoregresyon Teknikleri ve Bir Uygulama. *Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 19(1), 1-20.
- Alma, G.Ö., & Vupa, Ö. (2008). Regresyon Analizinde Kullanılan En Küçük Kareler ve En Küçük Medyan Kareler Yöntemlerinin Karşılaştırılması. *Süleyman Demirel Üniversitesi Fen Edebiyat Fakültesi, Fen Dergisi (E- Dergi)*, 3(2), 219-229.
- Alpar, R. (2003). Uygulamalı Çok Değişkenli İstatistiksel Yöntemlere Giriş, 1. Nobel Yayın Dağıtım, Ankara, 404s.
- Anderson, B. (1998). Scandinavian Evidence on Growth and Age Structure. ESPE 1997 Conference at Upsala University.
- Asar, Y. (2015). Liu Type Logistic Estimators. Selçuk Üniversitesi, Fen Bilimleri Enstitüsü, Doktora Tezi, İstatistik Anabilim Dalı, Konya.
- Aswani, A., Bickel, P., & Tomlin, C. (2011). Regression on Manifolds Estimation of the Exterior Derivative. *Ann. Statist.*, 39(1), 48-81.
- Aşkın, F. (2011). Sol Atriyum Çapının Temel Bileşenler Regresyonu, Kısmi En Küçük Kareler Regresyonu ve Yapay Sinir Ağları ile Tahminlerinin Karşılaştırılması. Yüksek Lisans Tezi, Fen Bilimleri Enstitüsü, İstatistik Anabilim Dalı, Elazığ.
- Aşkın, G.R., & Montgomery, D.C. (1980). Augmented Robust Estimators. *Technometrics*, 22, 333-341.
- Beltrami, E. (1873). Sulle Funzioni Bilineari. *Giornale di Matematiche ad USO Degli Studenti Delle Università*. 11, 98-106.
- Billor, N., & Kırıl, G. (2008). A Comparison of Multiple Outlier Detection Methods for Regression Data, *Communications in Statistics- Simulation and Computation*, 37(3), 521-545.
- Bulut, E., & Alın, A. (2009). Kısmi En Küçük Kareler Regresyon Yöntemi Algoritmalarından Nipals ve PLS – Kernel Algoritmalarının Karşılaştırılması ve Bir Uygulama. *Dokuz Eylül Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 24(2), 127-138.
- Bulut, Y.M. (2011). Çoklu İç İlişki Durumunda Kısmi En Küçük Kareler Regresyonu ve Alternatif Yöntemlerle Karşılaştırılması. Yüksek Lisans Tezi, Osmangazi Üniversitesi, Fen Bilimleri Enstitüsü, İstatistik Anabilim Dalı, Eskişehir.

- Bush, B.L., & Nachbar, R.B. (1993). Sample-Distance Partial Least Squares: PLS Optimized for Many Variables, With Application to CoMFA. Kluwer Academic Publishers, Vol. 7(5), 587-619.
- Canküyer, E., & Sönmez, H. (1996). Regresyon Analizinde Çoklu Doğrusal Bağlantı Sorununun İncelenmesi ve Uygulanması. *Anadolu Üniversitesi Fen Fakültesi Dergisi*, 173-187.
- Cattell, R.B. (1966). The Scree test for the Number of Factors. *Multivariate Behavioral Research*, 1, 245-276.
- Cooley, W.W., & Lohnes, P.R. (1971). *Multivariate Data Analysis* John Wiley and Sons, New York.
- Coşkuntuncel, O. 2010. Sosyal Bilimlerde Yanlı Regresyon Tahmin Edicilerin Kullanılması. *Eğitim ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 1(2), 100-108.
- Çamdeviren, H., Demir, N., Kanık, A., & Keskin, S. (2005). Use of Principal Component Scores in Multiple Linear Regression Models for prediction of *Chlorophyll-a* in reservoirs, *Ecological Modelling* 181, 581-589.
- Çankaya, S., Kayaalp, G.T., Sangun, L., Tahtalı Y., & Akar, M. (2006). A Comparative Study of Estimation Methods for Parameters in Multiple Linear Regression Model. 29(1), 43-47.
- De Jong, S. (1993). SIMPLS: An Alternative Approach to Partial Least Squares Regression, *Chemometrics and Intelligent Laboratory Systems*, 18, 251-263.
- Demirci, M.A. (2014). Ridge Regresyonda Sapma Parametresi k'nın Elde Edilmesinde Genetik Algoritma Yaklaşımı, Yüksek Lisans Tezi, Ondokuz Mayıs Üniversitesi, Samsun, 77s.
- Duran, E. (2011). Regresyon Analizinde Çoklu Bağlantı: Parametrik ve Semiparametrik Tahmin. Gazi Üniversitesi, Fen Bilimleri Enstitüsü Doktora Tezi, İstatistik Anabilim Dalı, Ankara.
- Ergüneş, E. (2004). En Küçük Kareler Yöntemi İle Ridge Regresyon Yönteminin Karşılaştırmalı Olarak İncelenmesi. Yüksek Lisans Tezi, Çukurova Üniversitesi, Fen Bilimleri Enstitüsü, Zootekni Anabilim Dalı, Adana.
- Farrar DE, & Glauber RR. (1967). Multicollinearity in regression analysis: The problem revisited. *The Review of Economics and Statistics*. 49(1), 92-107.
- Foong, N. (2007). An Overview of Biased Estimators, *Journal of Physical Science*, 18(2), 89-106.
- Fox, J. (1997). *Applied Regression Analysis, Linear Models and Related Methods*, 3th, sage publication, 123-240, USA.
- Garson, D. (2006). Testing of assumptions. <http://www.2.chass.ncsu.edu/garson/pa765/assumpt.htm>-(Erişim Tarihi: 18.06.2006).
- Glen, W.G., Sarker, M., Dunn III, W.J., & Scott, D.R. (1989). UNIPALS; Software for Principal Components Analysis and Partial Least Squares Regression. *Tetrahedron Comput. Methodol*, 2, 377-396.

- Graham, M. H. (2003). Confronting Multicollinearity in Ecological Multiple Regression Ecology, 84, 2809–2815.
- Gujarati, D.N. (1995). Basic Econometrics, 3rd Ed., Mc Graw-Hill, New York, 1002.
- Helland, S.I. (2001). Some Theoretical Aspects of Partial Least Squares Regression. Chemometrics and Intelligent Laboratory Systems, 58, 97-107.
- Höskuldsson, A. (1988). PLS Regression Methods. Journal of Chemometrics, 2, 211-228.
- Hintze, J.L. (2007). NCSS Help System. Kaysville, Utah.
- Hochberg, Y., & Tamhane, A.C. (1987). Multiple Comparison procedures. John Wiley & Sons Inc., New York.
- Hoerl, A.E. (1962). Application of Ridge Analysis to Regression Problems. Chem Engg. Proce, 60, 54-59.
- Hoerl, A.E., Kennard, R.W., & Baldwin, K.F. (1975). Ridge Regression: Some Simulations. Commun. Statist. A-Theory Methods, 4, 105-123.
- Hoerl, A.E. & Kennard, R.W., (1970). Ridge Regression: Biased Estimation for Nonorthogonal Problems. Technometrics, 12, 69-82.
- Jolliffe, I.T. (2002). Principal Component Analysis. Springer Series in Statistics, 2. Baskı, New York.
- Jordan, C. (1874). Mémoire Sur Les Formes Bilineaires. *Journal de Mathématiques Pures et Appliquées*, 35-54.
- Karadavut, U., Genç, A., Tozluca, A., Kınacı, İ., Aksoyak, Ş., Palta, Ç. & Pekkör, A. (2005). Nohut (cicer arietinum L.) Bitkisinde Verime Etki Eden Bazı Karakterlerin Alternatif Regresyon Modelleriyle Karşılaştırılması. *Tarım Bilimleri Dergisi*, 11(3), 328-333.
- Karakaya, E. (2011). Ridge ve Liu Tahminleyicilerinin Etkinliklerinin ve Yanlılıklarının Karşılaştırılması. Yüksek Lisans Tezi, Anadolu Üniversitesi, Fen Bilimleri Enstitüsü, İstatistik Anabilim Dalı, Eskişehir.
- Karakaş, S. (2008). Çoklu Doğrusal Bağlantı Problemi ve Yanlı Regresyon Tahminleri. Yüksek Lisans Tezi, İstanbul Üniversitesi, Sosyal Bilimler Enstitüsü, İstanbul.
- Kaşko, Y. (2007). Çoklu Bağlantı Durumunda İkili (Binary) Lojistik Regresyon Modelinde Gerçekleşen I. Tip Hata ve Testin Gücü. Yüksek Lisans Tezi, Ankara Üniversitesi, Fen Bilimleri Enstitüsü, Zootekni Anabilim Dalı, Ankara.
- Kırdemir, Ç. (2017). Ridge Regresyon Yönteminin Farklı Paket Programlarıyla Uygulanması. Yüksek Lisans Tezi, Yıldız Teknik Üniversitesi, Fen Bilimleri Enstitüsü, İstatistik Anabilim Dalı, İstanbul.
- Kızılırmak, B. (2016). Tahmin. [kisisel.ankara.edu.tr/politics.ankara.edu.tr/burca/ekonometri/3-%20Tahmin.pdf](http://kisisel.ankara.edu.tr/politics.ankara.edu.tr/burca/ekonometri/3-%20Tahmin.pdf)-(Erişim Tarihi: 01.12.2018)
- Kidwell, J.S., & Brown, L.H. (1982). Ridge regression as a Technique for Analyzing Models with Multicollinearity. *J Marriage and Family*, Vol. 44(2), 287-299.

- Kondylis, A. (2006). PLS Methods in Regression: Model Assesment and Inference. Ph. D. Thesis, University of Neuchâtel, Switzerland.
- Kurtuluş, M. (2001). Ridge Regresyon Üzerine Bir Çalışma. Gazi Üniversitesi, Fen Bilimler Enstitüsü (Basılmamış), Yüksek Lisans Tezi, Ankara.
- Kuvat, A. (2018). Ridge Regresyonda Sağlam Parametre Bulma. Muğla Sıtkı Koçman Üniversitesi, Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi, İstatistik Anabilim Dalı, Muğla.
- Küçükönder, H., Efe, E., Akyol, E., Şahin, M., & Üçkardeş, F. (2004). Çok Değişkenli İstatistiksel Analizlerin Hayvancılıkta Kullanımı. 4. Ulusal Zootekni Bilim Kongresi, 1-4 Eylül 2004, Isparta, 550-555.
- Lindgren, F., & Rännar, S. (1998). Alternative Partial Least-Squares (PLS) Algorithms, *Perspectives in Drug Discovery and Design*, 12/13/14, 105-113.
- Liu, K. (1993). A New Class of Biased Estimate in Linear Regression. *Communications in Statistics, Theory and Methods*, A (22), 393-402.
- Mahajan, V., Jain, A.K., & Bergier, M. (1977). Parameter Estimation in Marketing Models in the Presence of Multicollinearity: An Application of Ridge Regression. *J Marketing Research*, 14(4), 586-591.
- Marquardt, D.W. (1970). Generalized Invers, Ridge Regression, Biased Linear Estimation and Non Linear Estimation. *Techonometrics*, 12, 591-612.
- Marquardt, D.W., & Snee R.D. (1975). Ridge Regression in Practice. *J Am Stat Assoc*, 29, 3-20.
- Mendeş, M. (2013). Uygulamalı Bilimler İçin İstatistik ve Araştırma Yöntemleri. Kriter Yayınevi, İstanbul, 401s.
- Meko, D. (2006). Multiple Linear Regression. [http://www.ltrr.arizona.edu/~dmeko/notes\\_11.pdf](http://www.ltrr.arizona.edu/~dmeko/notes_11.pdf)-(Erişim Tarihi: 19.05.2006)
- Montgomery, D.C., & Voth. S.R. (1994). Multicollinearity and Leveroge in Mixture Experiments. *Journal of Quality Technology*, 26, 96-108.
- Naes, T., Irgens, C., & Martens, H. (1986). Comparison of Linear Statistical Methods for Calibration of NIR Instriments. *Applied Statistics*, 35(2), 195-206.
- NCSS Statistical System, (2001). User's Guide, Kaysville, NCSS Inc.
- Neter, J., Wasserman, W., & Kutner, M.H. (1990). *Applied Linear Regression Models*. Irwn Inc, Boston.
- Neter, J., Kutner, M., Nachtsheim, C., & Wasserman, W. (1996). *Applied Linear Regression Models*. Irwin, USA.
- Onay, İ. (2003). Ortalama Arteriyel Kan Basıncını Etkileyen Faktörlerin Temel Bileşen Regresyonu İle Belirlenmesi. Yüksek Lisans Tezi, Gazi Üniversitesi, Fen Bilimleri Enstitüsü, Ankara.

- Orhunbilge, N. (2000). Uygulamalı Regresyon ve Korelasyon Analizi. Avcıol-Basım Yayın, İstanbul, 394s.
- Ortabaş, N. (2001). Principal Components in the Problem of the Multicollinearity. Yüksek Lisans Tezi, Dokuz Eylül Üniversitesi, Fen Bilimleri Enstitüsü, İzmir.
- Özdamar, K. (2004). Paket Programlar ile İstatistik Veri Analizi-2 (Çok Değişkenli İstatistikler). Kaan Kitabevi, Eskişehir, 522.
- Özkan, K. (2009). Toprağın Tarla Kapasitesi Değişiminin Toprak Türüne Göre Temel Bileşenler Regresyon Analizi ile Modellenmesi. *Süleyman Demirel Üniversitesi Orman Fakültesi Dergisi*, 2, Issn: 1302-7085, 1.
- Öztürk, İ. (2014). Hayvansal Üretim Verilerinde Çoklu Bağlantı Probleminin Yanlı Regresyon Yöntemi ile Çözümlemesi. *Kahramanmaraş Sütçü İmam Üniversitesi Doğa Bilimleri Dergisi*, 17(3).
- Pamukçu, E. (2010). Sistolik Kan Basıncını Etkileyebilecek Faktörlerin Ridge Regresyon Analizi ile İncelenmesi ve Çoklu Bağlantı Problemi. Yüksek Lisans Tezi, Fırat Üniversitesi, Fen Bilimleri Enstitüsü, İstatistik Anabilim Dalı, Elazığ.
- Pamukçu, E. (2015). Yüksek Boyutlu Kanser Sınıflama Probleminde Bilgi Karmaşıklığı Kriteri ile Aykırı Gözlem Tespiti ve Boyut İndirgeme. Doktora Tezi, Fırat Üniversitesi, Fen Bilimleri Enstitüsü, İstatistik Anabilim Dalı, Elazığ.
- Pagel, M.U., & Lunneberg, C.E. (1985). Empirical Evaluation of Ridge Regression, *Psychological Bulletin*, 97, 342-355.
- Polat, E. (2009). Kısmi En Küçük Kareler Regresyonu. Hacettepe Üniversitesi, Fen Bilimleri Enstitüsü, İstatistik Anabilim Dalı, Yüksek Lisans Tezi. Ankara
- Pearson, K. (1901). On Lines and Planes of Closest Fit to Systems of Points in Space. *Philosophical Magazine*, 2, 559-572.
- Pimentel, E.C.G., Queiroz, S.A., Carvalheiro, R. & Fries, L.A. (2007). Use of Ridge Regression for the Prediction of Early Growth Performance in Crossbred Calves. *Genetics and Molecular Biology*, 30(3), 536-544.
- Rännar, S., Lindgren, F., Geladi, p., & Wold, S. (1994). A PLS Kernel Algorithm For Data Sets With Many Variables and Fewer Objects. Part 1: Theory and Algorithm. *Journal of Chemometrics*, 8, 111-125.
- Rathert, T. Ç., Üçkardeş, F., Nariç, D., & Aksoy, T. (2011). Comparison of Principal Component Regression with the Least Square Method in Prediction of Internal Egg Quality Characteristics in Japanese Quails. *Kafkas Üniversitesi Veterinerlik Fakültesi Dergisi*, 17(5), 687-692.
- Rencher, A.C., 2002, *Methods of Multivariate Analysis*, Wiley, New York.

- Shibata, R. (1989). Statistical Aspects of Model Selection. In Data to Modeling. J.C. Willems (ed). Springer Verlag, Berlin, 216-240.
- Sinan, A. (2003). Lineer Regresyonda Ridge Tahmin Edicileri ve Bir Uygulama. Selçuk Üniversitesi. Yüksek Lisans Tezi, Selçuk Üniversitesi, Fen Bilimleri Enstitüsü, İstatistik Anabilim Dalı, Konya.
- SPSS Inc., (1999). SPSS® Base 10 Application Guide, SPSS Inc., Chicago, USA.
- Stewart, G.W. (1993). On the Early History of the Singular Value Decomposition. *SIAM Review*, 35(4), 551-566.
- Şahinler, S. (1997). Regresyon Analizinde Etkili Gözlemlerin (Influential Observations) Belirlenmesinde Kullanılan İstatistiklerin Karşılaştırmalı Olarak İncelenmesi. Çukurova Üniversitesi, Fen Bilimleri Enstitüsü, Zootekni Anabilim Dalı, Doktora Tezi.
- Tarı, R. (2006). Ekonometri (4. Baskı), Avcı Ofset. İstanbul
- Topal, M., Yağanoğlu, A.M., Eyduran, E., Sönmez, A.Y., & Keskin, S. (2010). Çoklu Doğrusal Bağlantı Durumunda Ridge ve Temel Bileşenler Regresyon Analiz Yöntemlerinin Kullanımı. *Atatürk Üniversitesi Ziraat Fakültesi Dergisi*, 41(1), 53-57.
- Tunç, Z. (2018). En Küçük Kkareler ve Temel Bileşenler Regresyon Analizlerinin Karşılaştırılması. Yüksek Lisans Tezi, İnönü Üniversitesi, Sağlık Bilimleri Enstitüsü, Biyoistatistik ve Tıp Bilişimi Anabilim Dalı, Malatya.
- Tüylüoğlu, Ş., & Albayrak, A.S. (2010). Hayat Pahalılığı ve Türkiye'de İllerin Hayat Pahalılığı Sıralamasını Belirleyen En Önemli Faktörlerin Ridge Regresyon Analiziyle İncelenmesi. *Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 15(2), 63-91.
- Üçkardeş, F., Efe, E., Narinç, D., & Aksoy, T. (2012). Japon Bildircinlarında Yumurta Ak İndeksinin Ridge Regresyon Yöntemiyle Tahmin Edilmesi. *Akademik Ziraat Dergisi*, 1, 11-20.
- Vinod, H.D., Ullah, A. (1981). Recent Advances in Regression Methods. New York, Marcel Dekker.
- Vural, A. (2007). Aykırı Değerlerin Regresyon Modellerine Etkileri ve Sağlam Kestiriciler. Marmara Üniversitesi, Sosyal Bilimler Enstitüsü, Yüksek Lisans Tezi, İstanbul
- Wold, S., Ruhe, A., Wold, H., & Dunn III, W.J. (1984). The Collinearity Problem in Linear Regression. The Partial Least Squares (PLS) Approach to Generalized Inverses, *SIAM J. Sci. Stat. Comput.*, 5(3), 735-743.
- Yamak, R., & Köseoğlu, M. (2006). *Uygulamalı İstatistik ve Ekonometri* (3. Baskı), Celepler Matbaacılık, Trabzon.
- Yeniay, O., & Göktaş, A. (2002). A Comparison of Partial Least Squares Regression with other Prediction Methods. *Hacettepe Journal of Mathematics and Statistics*, 31, 99-111.
- Yıldırım, N. (2010). En Küçük Kareler, Ridge Regresyon ve Robust Regresyon Yöntemlerinde Analiz Sonuçlarına Aykırı Değerlerin Etkilerinin Belirlenmesi.

Yüksek Lisans Tezi, Çukurova Üniversitesi, Fen Bilimleri Enstitüsü, Zootekni Anabilim Dalı, Adana.

Yılmaz, E. (2018). Bir Bivalv Türünün (*Abra alba* (W. Wood, 1802) Kabuk Ağırlığının Ridge Regresyon Yöntemi ile Tahmini. *KSÜ Tarım ve Doğa Dergisi*, 21(4), 593-598.

Yolacan, S., Yazıcı, B., & Ağaoglu, E. (2005). Artificial Neural Network Approach as an Alternative to Ridge Regression on an Economic Problem.4th WSEAS Int. Conf. on Non-Linear Systems and Chaos, Sofia, Bulgaria, 120-124.



## 7. ÖZGEÇMİŞ

Kişisel Bilgiler	
Adı Soyadı	DİLDAR AYŞE DERMAN
Doğum Yeri	ARTVİN
Doğum Tarihi	01.07.1989
Uyruğu	<input checked="" type="checkbox"/> T.C. <input type="checkbox"/> Diğer:
Telefon	05397750773
E-Posta Adresi	<a href="mailto:d.ayse.yavuz@gmail.com">d.ayse.yavuz@gmail.com</a>

Eğitim Bilgileri	
Lisans	
Üniversite	ERZURUM ATATÜRK ÜNİVERSİTESİ
Fakülte	ZİRAAT FAKÜLTESİ
Bölümü	ZİRAAT MÜHENDİSLİĞİ-ZOOTEKNİ A.D.
Mezuniyet Yılı	2012