

Çoklu Doğrusal Bağlantı Durumunda Ridge ve Temel Bileşenler Regresyon Analiz Yöntemlerinin Kullanımı

Mehmet TOPAL¹ Ecevit EYDURAN² A.Mutlu YAĞANOĞLU¹
Adem Yavuz SÖNMEZ³ Sıddık KESKİN⁴

¹Atatürk Üniversitesi Ziraat Fakültesi Zootehni Bölümü Erzurum (mtopal@atauni.edu.tr)

²Iğdır Üniversitesi Ziraat Fakültesi Zootehni Bölümü, Iğdır

³Atatürk Üniversitesi Ziraat Fakültesi Su Ürünleri Bölümü, Erzurum

⁴Yüzüncü Yıl Üniversitesi Tıp Fakültesi, Van

Geliş Tarihi : 20.07.2010

Kabul Tarihi : 22.10.2010

ÖZET: Ridge ve temel bileşenler regresyon analiz yöntemleri, çok değişkenli regresyon verilerini analiz etmek için kullanılan istatistik analiz yöntemleridir. Çoklu bağlantı ortaya çıktığında en küçük kareler tahminleri sapmasız olmasına karşın tahminlerin varyansları büyük olduğundan gerçek değerlerinden oldukça uzakta olabilmektedirler. Bir derece yanlı regresyon tahminlerine izin vermek suretiyle ridge ve temel bileşenler regresyon standart hataları indirgenir. Dolayısıyla çoklu bağlantı durumu mevcut olduğunda en küçük kareler metoduna alternatif olarak ridge ve temel bileşenler regresyon metodları kullanılabilir. Bu araştırmada farklı yaşlara sahip 91 adet sazan balığından elde edilen çeşitli vücut ölçüleri kullanılarak karkas ağırlığını tahminleyen bir modelin geliştirilmesi amaçlanmıştır. Vücut ölçüleri arasında çoklu bağlantı durumu ortaya çıkmasından dolayı en küçük kareler regresyonuna alternatif olan ridge ve temel bileşenler regresyon analiz yöntemleri uygulanmış ve aynı veri seti için bu üç metod karşılaştırılmıştır. Karşılaştırma kriteri olarak belirleme katsayısı (R^2), hata kareler ortalamasının karekökü (S), hata kareler ortalaması ve modellerin varyasyon katsayısı kullanılmıştır. Bu kriterlere göre, en iyi uyumu sırasıyla en küçük kareler ($R^2=0.905$, $S=19.587$), ridge ($R^2=0.898$, $S=20.2563$) ve temel bileşenler regresyon ($R^2=0.878$, $S=22.127$) metodlarının verdiği gözlenmiştir. Sonuç olarak, çoklu doğrusal bağlantı durumunda en küçük kareler metodu kullanmak yerine Ridge ve temel bileşenler regresyon yöntemlerinin kullanılmasının daha doğru olabileceği kanaatine varılmıştır.

Anahtar kelimeler: Çoklu doğrusal bağlantı, Ridge regresyon, Temel bileşenler regresyonu, en küçük kareler metodu

Use of Ridge and Principal Component Regression Analysis Methods in Multicollinearity

ABSTRACT: Ridge and principal component regression analysis methods are statistical analysis techniques that are used to analyze multiple regression data. In the case of Multicollinearity, although Least Squares estimates are unbiased, variances of these estimates are larger and these variances can be farther than real values. With adding a degree of bias to regression estimates, standard errors of Ridge and principal component regression are reduced. Therefore, in the event of Multicollinearity, Ridge and principal component regression methods can be used as an alternative to Least Squares method. This investigation aimed to fit a model in order to estimate carcass weight from various body measurements of 91 cyprinus fish with different ages. As Multicollinearity problems among body measurements were determined, Ridge and principal component regression methods as an alternative to Least Squares method were applied for available data, performances of these three methods for the data were compared with each other. In order to compare effectiveness of these methods, Coefficient of Determination (R^2), Root of Mean Square Error (RMSE), Mean Square Error (MSE), and Coefficient of Variation as comparison criteria were used. According to these criteria, the best fit orders were observed in least squares ($R^2=0.905$, $S=19.587$), Ridge ($R^2=0.898$, $S=20.2563$) and principal component regression ($R^2=0.878$, $S=22.127$), respectively.

As a result, it was concluded that, use of Ridge and principal component regression analysis methods could be truer instead of least squares method under Multicollinearity problem

Key Words: Multicollinearity, Ridge regression, principal component regression, least squares method.

GİRİŞ

Bir bağımlı (Y) ve birden fazla bağımsız değişken (X_1, X_2, \dots, X_n) arasındaki fonksiyonel ilişki çoklu regresyon analizi ile incelenir. Regresyon analizinde amaç bağımsız değişkenlerden bağımlı değişkeni tahminleyebilen en iyi modelin oluşturulması veya bağımlı değişkenin hangi bağımsız değişkenlerden daha çok etkilendiğini tespit etmektir. Bu amaç için oluşturulacak modeldeki parametrelerin tahmini için en çok kullanılan yöntemlerden birisi hata kareler toplamının minimum olacak şekilde modeldeki parametrelerin tahmin edilmesini sağlayan en küçük kareler tahmin yöntemidir. Çoklu regresyon analizinin uygulanabilmesi için (en küçük kareler tahmin metoduna göre parametrelerin tahmin edilebilmesi

için), bazı varsayımların geçerli olması gerekir. Bu varsayımlar; bağımsız değişkenler arasında çoklu bağlantının olmaması, hata terimine ait (ϵ) dağılımın normal olması, hata teriminin varyansı tüm X değerleri için sabit olmalı, hata terimi değerleri arasında bir ilişkinin (otokorelasyonun) olmaması yani herhangi bir ϵ_i değeri kendisinden önceki ϵ_j değerine bağımlı olmaması, hata terimi (ϵ_i) ile bağımsız değişkenler (X_i) arasındaki kovaryans sıfıra eşit olmalı yani ϵ_i ile bağımsız değişkenler arasında bir ilişki olmaması şeklinde ifade edilebilir (Akkaya ve Pazarlıoğlu, 1998; Neter vd. 1989). Çoklu doğrusal regresyon analizinde en çok karşılaşılan problemlerden biri bağımsız değişkenler arasındaki doğrusal ilişkinin bulunması yani çoklu doğrusal

bağlantı durumunun olmasıdır. Bağımsız değişkenlerden biri diğer bağımsız değişken veya değişkenlerin bir doğrusal fonksiyonu olarak ifade ediliyorsa bağımsız değişkenler arasında doğrusal ilişkinin varlığından söz edilir.

Çoklu doğrusal bağlantı, regresyon katsayılarının yanlış tahmin edilmesine, regresyon katsayılarının standart hatalarının abartılı çıkmasına buna bağlı olarak güven aralıklarının büyümesine ve t test değerinin küçülmesine sebep olabilir. Standart hatanın büyümesi istatistiksel olarak önemli olan regresyon katsayılarının önemsiz çıkmasına böylece yanlış sonuç elde edilmesine sebep olabilir. Çoklu doğrusal bağlantının olup olmadığını tespit etmek için (Hines ve Montgomery, 1990; Alpar, 1997) a) bağımsız değişkenler arasındaki korelasyon matrisinden faydalanılır. Eğer değişkenler arasındaki korelasyon katsayısının mutlak değeri 1'e yakınsa o bağımsız değişkenler arasında çoklu doğrusal bağlantının olduğu ifade edilir. b) Varyans şişirme değerleri (VIF); bu değerler büyüdükçe (VIF değerleri ≥ 10) ilgili bağımsız değişkenler arasında çoklu bağlantının varlığından bahsedilebilir. VIF değeri büyüdükçe değişkenler arasındaki çoklu doğrusal bağlantı o derecede artar. Genelde VIF değeri 10 üzerinde olduğunda o değişkenler arasında çoklu bağlantının varlığı kabul edilir. c) Bağımsız değişkenlere ait korelasyon matrisinin determinanı 0 ile 1 arasında değişir ve çoklu doğrusal bağlantının belirlenmesinde kullanılabilir. Korelasyon matrisinin determinanı 1 olduğunda değişkenler arasında doğrusal bağımlılığın olmadığı, ancak korelasyon matrisinin determinanı 0 olduğunda ise değişkenler arasında çoklu doğrusal bağlantının olduğu ifade edilir. Determinant değeri küçüldükçe çoklu bağlantının derecesi de o oranda artar. d) Korelasyon matrisinin bir veya daha fazla karakteristik kökü (özdeğerleri) sıfır veya sıfıra yakın ise çoklu bağlantının olduğu söylenir. e) Özdeğerlerin en büyüğünün en küçüğüne oranı da ($\lambda_{\max} / \lambda_{\min}$) çoklu doğrusal bağlantının bir ölçüsü olarak kullanılır ve bu oran büyüdükçe çoklu doğrusal bağlantının derecesi de artar. Genelde ($\lambda_{\max} / \lambda_{\min}$) oranı 10'dan küçük olduğunda veri setinde ciddi bir çoklu bağlantının olmadığı ifade edilir. f) Regresyon katsayılarının t değerlerinin hiçbiri anlamlı olmadığı halde regresyon modeline ait olan F istatistiğinin anlamlı olması veri setinde çoklu doğrusal bağlantının varlığının bir göstergesidir.

Çoklu doğrusal bağlantının giderilmesi için veriye yeni değişkenler eklenebilir veya veri setinde çoklu bağlantılı değişkenlerden bir veya bir kaç modelden çıkarılabilir. Ancak değişkenlerin modelden çıkarılmasıyla modele gerçekten katkı yapan bir değişken çıkarılmış olabilir ki bu da bilgi kaybına sebep olur. Modeldeki değişkenleri çıkarmadan regresyon katsayılarını tahmin etmek için

genellikle çoklu doğrusal bağlantıya en küçük kareler yönteminden daha az duyarlı olan yöntemler kullanılır. Bu yöntemlerin başında yanlış tahminler veren ridge regresyon ve temel bileşenler regresyon yöntemlerinin kullanılması önerilir.

Bu çalışmada, sazan balıklarında çeşitli vücut ölçüleri kullanılarak karkas ağırlığını tahminleyen bir model tespit edilmesi amaçlanmıştır. Ancak vücut ölçüleri arasında çoklu bağlantı durumu ortaya çıktığı için en küçük kareler regresyonuna alternatif olan ridge ve temel bileşenler regresyon analiz yöntemleri uygulanmış ve aynı veri setinde bu üç metod karşılaştırılmıştır.

MATERYAL VE YÖNTEM

Materyal

Çalışmada kullanılan balıklar, Aras Nehri'ni oluşturan iki önemli sistemden birisi olan Hasankale çayından yakalanmıştır. Hasankale çayı Erzurum Hasankale ilçesi dolaylarında Deveburnu mevkiinden doğup Pasinler ovasında Kurnuç ve Müceldi dereleri ile birleşip Hasankale Çayını oluşturur ve Köprüköy dolaylarında Aras'a katılır.

Capoeta capeta capoeta örnekleri 6kg'lık 12x12mm göz açıklığına sahip serpmeye ağlar vasıtasıyla yakalanmıştır. Yakalanan balıklarda meristik ölçümler laboratuvarında boylama tahtası ve elektronik kumpas vasıtasıyla yapılmıştır. Yaş tayini dorsal yüzgeç ile lateral çizgi arasındaki pullardan alınan örneklerde yaş halkalarının sayımı vasıtasıyla yapılmıştır. Herhangi bir karışıklığa meyil verilmemesi için her bir balıktan en az 3 pul yaş tayininde incelenmiştir (Nikolsky 1963; Ambrose 1989; Lagler 1967). 1+, 2+, 3+, 4+, 5+ yaş gurubun dahil balıklar sırasıyla 1, 2, 3, 4, 5 yaş guruplarına dahil edilmiştir. Ağırlık ölçümler 0,01gr'a duyarlı hassas terazide yapılmıştır. Bu balıklardan, karkas ağırlığı, total boy (cm), standart boy (cm), baş boyu (cm), vücut yüksekliği (cm) ve yaş (yıl) gibi ölçümler alınmıştır.

Verilerin istatistik analizi NCSS paket programında yapılmıştır.

Yöntem

Regresyon eşitliği matris formunda

$$Y = X\beta + e \quad (1.1)$$

şeklinde yazılır. Eşitlikte Y bağımlı değişken; X bağımsız değişken; β regresyon katsayılarını ve e hatayı göstermektedir.

En küçük kareler regresyon katsayıları aşağıdaki eşitlikle tahmin edilir

$$\beta = (X'X)^{-1}X'Y \quad (1.2)$$

Eşitlikte, β ; $(p-1) \times 1$ boyutlu regresyon katsayıları vektörü, $X'X = r_{xx}$; $(p-1) \times (p-1)$ boyutlu bağımsız değişkenlerin korelasyon matrisi, $X'Y$

$=r_{YX}$; $(p-1) \times 1$ boyutlu bağımlı Y değişkeni ile bağımsız X değişkenleri arasındaki korelasyon vektörüdür (Neter vd. 1989). Eşitlik 1.2 aşağıdaki gibi yazılabilir.

$$\beta = r_{XX}^{-1} r_{YX} \quad (1.3)$$

Ridge Regression

Ridge regresyonda parametrelerin tahmin edilmesinde aşağıdaki eşitlik kullanılır

$$\beta^* = (X'X + kI)^{-1} X'Y \quad (1.4)$$

Eşitlikte β^* ; $(p-1) \times 1$ boyutlu ridge regresyon katsayıları vektörü, I ; $(p-1) \times (p-1)$ boyutlu birim matris ve k sabit değer ve genelde $0 \leq k \leq 1$ aralığında bulunur. Ridge parametresi olan k değeri en küçük kareler yönteminden daha küçük hata kareler ortalaması elde edilmesini sağlar (İpek, 2001). Optimum k değerinin seçilmesinde genelde ridge regresyonunun grafiksel gösterimi olan ridge izi ve varyans şişirme değerlerinden faydalanılır. Ridge izinde k nın bir fonksiyonu gibi regresyon katsayıları gösterilir ve düşey eksene regresyon katsayıları yatay eksene ise k değerleri yazılarak oluşturulan grafikte katsayıların stabil olduğu k değeri seçilir. VIF değerine göre k nın seçimi yapılırken genelde VIF 10 dan küçük olduğunda k nın seçilebileceği ifade edilmiştir (Marquardt ve Snee, 1975).

Temel Bileşenler Regresyon Analizi

Modeldeki değişkenleri çıkarmadan regresyon katsayılarını tahmin etmek için genellikle çoklu doğrusal bağlantıya en küçük kareler yönteminden daha az duyarlı olan temel bileşenler regresyon analizi yöntemi kullanılmaktadır. Temel bileşenler analizinde ilk olarak hem bağımlı hem de bağımsız değişkenler kendi ortalamalarından farkları alınıp standart sapmalarına bölünerek standartlaştırılır. Tüm hesaplamalar standart veriler üzerinden yapılır. Temel bileşenler regresyonunu icra etmek için bağımsız değişkenler temel bileşenlerine dönüştürülür ve matematiksel olarak aşağıdaki eşitlik yazılır (Hintze, 2007);

$$X'X = PDP' = Z'Z \quad (1.5)$$

eşitlikte $X'X = r_{XX}$; $(p-1) \times (p-1)$ boyutlu bağımsız değişkenlerin korelasyon matrisidir ve değişkenler kümesine ait özdeğerler veya özvektörler korelasyon veya kovaryans matrisleri kullanılarak hesaplanabilir (Alpar, 1997). D , $X'X$ in özdeğerlerinin bir köşegen matrisidir; P , $X'X$ in

özvektör matrisini; Z , veri matrisini göstermektedir (Hintze, 2007).

Orijinal X değişkenlerinin ağırlıklı ortalamalarını ifade eden yeni Z değişkenleri türetilir. Yani X_1, \dots, X_n bağımsız değişkenleri Z_1, \dots, Z_n gibi değişkenlere dönüştürülür. Bu yeni değişkenler temel bileşen olduğu için temel bileşenler arasındaki korelasyon sıfırdır. Çok küçük özdeğerler hesaplandığında güçlü çoklu bağlantı ortaya çıkacağı için küçük özdeğerlerle ilişkili bileşenler analizden çıkarılır ve genelde özdeğeri çok küçük bir veya iki temel bileşen elde edilir. Özdeğeri düşük temel bileşenler çıkartıldıktan sonra kalan temel bileşenler üzerine Y'nin regresyon analizi yapıldığında çoklu doğrusal bağlantı problemi ortadan kalkmış olur. Daha sonra regresyon katsayılarının tahmin edilmesi için sonuçlar X'in ölçeğine geri dönüştürülür. Elde edilen tahminler yanlış olacaktır fakat bu yanlışlığın büyüklüğünün varyansın küçültülmesiyle telafi edileceği arzu edilmektedir. Yani temel bileşen tahminlerinin hata kareler ortalaması en küçük karelere nazaran daha küçük olması beklenmektedir (Hintze, 2007). Regresyon katsayılarının tahmini aşağıdaki eşitlikle bulunur;

$$A = (Z'Z)^{-1} Z'Y = D^{-1} Z'Y \quad (1.6)$$

Bu eşitlik değişik veri setine uygulanan en küçük kareler regresyonudur. İki regresyon katsayısı seti (A ve β) aşağıdaki formüller kullanılarak ilişkilendirilir;

$$A = P' \beta$$

$$\beta = PA$$

A'nın sıfıra eşit ögesi temel bileşenler analizinden çıkartılabilir. Temel bileşenler regresyonu uygulanırken takip edilmesi gereken adımlar kısaca özetlenirse, ilk olarak X matrisi için temel bileşenler analizi uygulanır ve temel bileşenler (Z) elde edilir, ikinci olarakta A en küçük kare tahminlerini elde etmek için Z üzerine Y'nin regresyon analizi uygulanır, üçüncüsü A'nın en son elementi sıfıra eşitlenir, son olarak $\beta = PA$ eşitliği kullanılarak hesaplanan katsayılar orijinal ölçeğine dönüştürülür (Hintze, 2007; Albayrak, 2005).

BULGULAR VE TARTIŞMA

Analiz sonucunda Sazan balıklarının karkas ağırlığı, total boy, standart boy, baş boyu, vücut yüksekliği ve yaşlarına ait tanımlayıcı istatistik değerleri Çizelge 1'de ve değişkenler arasındaki korelasyon katsayıları ve varyans şişirme (VIF) değerleri Çizelge 2'de verildi.

Çizelge 1: Değişkenlerin tanımlayıcı istatistik değerleri

Değişken	Ortalama	Standart Sapma	Minimum	Maxsimum
Karkas Ağırlığı (KA)	92,51	61,64	6,55	286,4
Total Boy (TB)	21,97	5,42	9,9	33,6
Standart Boy (SB)	18,33	4,80	7,7	28,7
Baş Boyu (BB)	4,16	0,92	2,2	5,9
Vücut Yüksekliği (VY)	4,81	1,04	2,6	6,7
Yaş	2,81	1,13	1,0	5,0

Çizelge 2. Değişkenler arasındaki korelasyon katsayıları ve Varyans Şişirme (VIF) değerleri

	KA	TB	SB	BB	VY	VIF
TB	0.929					29.755
SB	0.888	0.972				28.673
BB	0.907	0.934	0.898			22.160
VY	0.912	0.935	0.902	0.975		22.344
Yaş	0.852	0.872	0.916	0.803	0.811	6.491

Değişkenler arasındaki korelasyon katsayılarının 1'e yakın olması değişkenler arasında doğrusal çoklu bağlantının olduğunu göstermektedir. Korelasyon katsayılar matrisine bakıldığında en küçük korelasyon baş boyu ile vücut yüksekliği arasında (0.803) bulunurken diğer değişkenler arasındaki korelasyon katsayıları genelde 0.90'dan büyük bulunmuştur. Bu da değişkenler arasında çoklu doğrusal bağlantı probleminin olduğunu göstermektedir. Çoklu doğrusal bağlantı olup olmadığını tespit etmede bir diğer yol varyans şişirme değerlerinin kullanılmasıdır. Varyans şişirme değerleri Çizelge 2' de verilmiştir. Total boy, standart boy, baş boyu ve vücut yüksekliğine ait varyans şişirme değerleri 10'dan büyük olduğu için değişkenler arasında çoklu doğrusal bağlantının olduğu görülmektedir. Değişkenler arasında çoklu bağlantının olup olmadığını tespit edilmesinde kullanılan bir diğer yöntem bağımsız değişkenlere ait korelasyon matrisinin özdeğerlerinin kullanılmasıdır. Bağımsız değişkenlere ait korelasyon matrisinin özdeğerleri sırasıyla 4.610, 0.262, 0.084, 0.025 ve 0.019 olarak bulunmuştur. Buna göre son üç özdeğer (karakteristik kök) sıfıra yakın olduğu için değişkenler arasında çoklu bağlantının varlığından söze edilebilir. Ayrıca en büyük özdeğer en küçük özdeğere bölündüğünde 242.63 bulunur. Bu değer 100 den büyük olduğu için yine çoklu bağlantı probleminin olduğuna karar verilir. Çoklu bağlantı probleminin olup olmadığını tespit edilmesinde kullanılan bir diğer yöntem özdeğerlerin terslerinin toplamı, bağımsız değişken sayısından fazla ise bağımsız değişkenler arasında çoklu bağlantının olduğu ifade edilir. Özdeğerlerin terslerinin toplamı 108,57 olarak bulunur ki, bu değer bağımsız değişken sayısından oldukça büyük olduğu için çoklu bağlantı probleminin olduğuna karar verilir. Çoklu bağlantı probleminin olduğu durumlarda çoklu bağlantının

giderilmesi için veri setine yanlı tahmin yöntemlerinden ridge ve temel bileşenler regresyon yöntemleri uygulanabilir. Karkas ağırlığı bağımlı değişken ve total boy, standart boy, baş boyu, vücut yüksekliği ve yaş bağımsız değişken olarak en küçük kareler, temel bileşenler ve ridge regresyon metoduna göre elde edilen çoklu doğrusal regresyon analiz sonuçları Çizelge 3'de verilmiştir.

En küçük kareler metoduna göre karkas ağırlığının tahminlenmesi için oluşturulan çoklu doğrusal regresyon modelinde standartlaştırılmış regresyon katsayılarına göre modele en fazla katkı yapan değişkenlerin sırasıyla total boy, standart boy, yaş, vücut yüksekliği ve baş boyu olduğu gözlenmiştir. Standart boy negatif yönde etki yaparken diğer değişkenler pozitif yönde etki yapmaktadır. Dolayısıyla standart boydaki bir birimlik standart sapmalık artış karkas ağırlığında - 0.71 birimlik standart sapmalık azalışa sebep olacaktır. Oluşturulan modele ait belirleme katsayısı 0.905 olarak bulunmuş ve buna göre modele dahil edilen değişkenlerin karkas ağırlığındaki toplam değişimin %90.5'ini açıkladığı görülmüştür. Temel bileşenler regresyona göre modele en fazla katkı yapan değişkenler sırasıyla baş boyu, vücut yüksekliği, yaş, total boy ve standart boy bulunmuştur. Bu değişkenlerin pozitif yönde etki yaptığı gözlenmiştir. Temel bileşenler regresyon yönteminde ele alınan bağımsız değişkenler karkas ağırlığındaki değişimin %87.8'ini açıklamaktadır. Ridge regresyon yöntemine göre modele en fazla katkı yapan değişkenlerin sırasıyla total boy, standart boy, yaş, vücut yüksekliği ve baş boyu olduğu gözlenmiş olup bu değişkenler karkas ağırlığındaki değişimin %89.8'ini açıklamaktadır. Her üç yöntemde uydurulan modellerde değişkenlere ait regresyon katsayılarının standart hataları ridge regresyon yönteminde diğer yöntemlere göre daha düşük

bulunmuştur. Standart hatanın küçülmesi istatistiksel olarak önemli olan regresyon katsayılarının önemli

çıkmasına ve böylece doğru sonuç elde edilmesine sebep olabilir.

Çizelge 3: En küçük kareler, temel bileşenler ve ridge regresyon yöntemine göre elde edilen sonuçlar

	En Küçük Kareler			Temel Bileşenler Regresyonu			Ridge Regresyon (k=0.005)		
	β	BETA	S(β)	β	BETA	S(β)	β	BETA	S(β)
Sabit	-136.39			-147.09			-138.42		
TB	10.86**	0.96	2.08	1.45*	0.13	0.83	8.91**	0.78	1.74
SB	-9.06**	-0.71	2.30	1.38	0.11	0.91	-6.80**	-0.53	1.93
BB	8.66	0.13	10.55	19.84*	0.30	11,63	10.49	0.16	9.19
VY	12.66	0.21	9.40	14.15	0.24	10.61	13.08	0.22	8.21
Yaş	21.19**	0.39	4.64	11.33**	0.21	4.71	18.92**	0.35	4.45
	R ² =0.905 HKO=383.65 VK=0,2117			R ² =0.878 HKO=489.60 VK=0,2391			R ² =0.898 HKO=410.32 VK=0,2190		

*: P<0.05, **:P<0.01, β : Regresyon katsayıları, BETA: Standardized regresyon katsayıları, S(β): Regresyon katsayılarının standart hatası, R²: Belirleme katsayısı, HKO: Hata kareler ortalaması, VK: Varyasyon Katsayısı

Modelin varyasyon katsayısı, hata kareler ortalamasının karekökünün bağımlı değişkenin ortalamasına bölünmesiyle elde edilir ($VK = \sqrt{HKO/Y}$) ve modele göre elde edilen beklenen değerlerin dağılımı hakkında bilgi verir. Modellerin karşılaştırılmasında varyasyon katsayısı kullanıldığında varyasyon katsayısı küçük olan model daha etkindir ve tercih edilir. Her üç metoda göre elde edilen modellerin varyasyon katsayılarına göre en etkin modeller sırasıyla en küçük kareler, ridge regresyon ve temel bileşenler regresyon metotlarına göre elde edilmiştir. Buna göre Sazan balıklarında karkas ağırlığını tahminlemede en küçük kareler metodu ile elde edilen modele göre beklenen değerlerin varyasyonunun diğer metotlar ile elde edilen modellerin beklenen değerlerinin varyasyonundan daha az olabileceği gözlenmiştir.

Karşılaştırma kriteri olarak belirleme katsayısı, hata kareler ortalaması ve modellerin varyasyon katsayısı kullanıldığında her üç kritere göre de en iyi uyumu en küçük kareler metoduna göre elde edilen modelin verdiği fakat bağımsız değişkenler arasında çoklu doğrusal bağımlılığın olduğu durumlarda en küçük kareler metoduna alternatif olan ridge ve temel bileşenler regresyon yöntemlerinin kullanımı daha uygun olur. Ridge ve temel bileşenler regresyon yöntemleri regresyon denklemindeki katsayıları yanlış tahmin ederek tahminlerin varyanslarını azalttığı için

çoklu bağlantı durumunda her zaman ridge ve temel bileşenler regresyon analizlerinin en küçük kareler yöntemine tercih edilebileceği söylenebilir

KAYNAKLAR

- Akkaya, Ş. ve Pazarlıoğlu, M. V., 1998. Ekonometri I. Anadolu Matbaacılık, İzmir.
- Albayrak, A. S., 2005. Çoklu Doğrusal Bağlantı Halinde En Küçük Kareler Tekniğinin Alternatif Yanlı Tahmin Teknikleri ve Bir Uygulama. ZKÜ Sosyal Bilimler Dergisi, Cilt 1, Sayı 1
- Alpar, R., 1997. Uygulamalı Çok Değişkenli İstatistiksel Yöntemler Giriş-I. Kültür Ofset, Ankara.
- Ambrose J.J.R., 1989. Age determination. In: Larry A, Nielson LA and Johnson DL (eds), Fisheries techniques, Southern Printing Company, Virginia, 301-325.
- Hines, W. W. ve Montgomery, D. C., 1990. Probability and Statistics in Engineering and Management Science. John Wiley & Sons, Inc.
- Hintze, J. L., 2007. NCSS Help System. Kaysville, Utah.
- Lagler, E. D., 1967. Freshwater Fishery Biology, W.M.C. Brown Company Publishers Dobuque Iowa, p.317.
- Marquardt, D.W ve Snee, R.D., 1975. Ridge Regression in Practice. The American Statistician, February 1975, Vol. 29, p3-20.
- Neter, J., Wasserman, W. ve Kutner, M. H., 1989. Applied Linear Regression Models. Irwin Inc. Boston.
- Nikolsky, G.W., 1963. The Ecology of Fishes. Academic Press, London.
- İpek, O., 2001. Ridge Regresyon Üzerine Bir Çalışma. VI. Ulusal Ekonometri ve İstatistik Sempozyumu, Çukurova Üniversitesi, Adana, 20-22 Eylül 2001.