Volume: 05 Number: 02 Page: 1-19 ISSN No: 1303-6319 Cilt: 05 Sayı: 02 Sayfa: 1-19

İÇİLİŞKİ VE GENELLEŞTİRİLMİŞ MAKSİMUM ENTROPİ TAHMİN EDİCİLERİ

Prof. Dr. Altan ÇABUK* **Prof. Dr. Fikri AKDENİZ****

ÖZET

Bu çalışmada, alışılmış gösterimlerle $y = X\beta + u$ *genel lineer regresyon modeli düşünülmüştür. Bir çok uygulamada tasarım matrisi X şiddetli içilişkiye sahip olabilir. İçilişkinin varlğında Ridge regresyon tahmin edicisi* $\hat{\beta}_k = (X'X + kI)^{-1} X'y$ *(Hoerl ve Kennard, 1970) ve Liu tahmin edicisi* $\hat{\beta}_d = (X'X + I)^{-1}(X'y + d\hat{\beta})$ *(Liu, 1993) ya da geliştirilmiş Ridge ve Liu-tipi tahmin ediciler en küçük kareler tahmin edicilerini iyileştirmek amacıyla kullanılmaktadır. Çalışmada, alternatif tahmin etme yöntem bilimi olarak maksimum entropi verilmiş ve temel veri kümesinde kötü koşulluluk olduğunda, genel lineer regresyonda parametreleri tahmin etmek için maksimum entropi yöntemi kullanılmıştır. Genelleştirilmiş maksimum entropi (GME) tahmin edicisi nitelendirilerek, parametre destek matrisleriyle birlikte parametreler üzerine eşitsizlik kısıtlarının koyulduğu tahmin yöntemi geliştirilmiştir. GME tahmin ediciler alternatif tahmin etme yöntemleri (En küçük kareler (EKK), eşitsizlik kısıtlı EKK, Ridge regresyon ve Liu-tip) ile hata kareleri ortalaması (HKO) ölçütüne göre karşılaştırılmıştır. Bu amaç için ABD'de tavuk talebi veri kümesi (Gujarati, 1992) üzerinde tahmin ediciler için nümerik olarak analiz edilmiştir.*

Anahtar kelimeler: Destek Noktası, Eşitsizlik Kısıtlı EKK Tahmin Edicisi, Genelleştirilmiş Maksimum Entropi Tahmin Edicisi, Liu Tahmin Edicisi, Ridge Rregresyon Tahmin Edicisi.

1. GİRİŞ

Genel lineer modelde uygulanan yöntemlerin sağlaması gerekli varsayımlarda veya koşullarda çoğu kez sapmalar görülür. Araştırmacılar, problemi nitelendirerek ve daha incelikli tahmin yöntemleri bularak bu zorlukların üstesinden gelecek yanıtlar arar. Klasik regresyon varsayımlarında düzenli olarak görülen bu sapmaların nedeni, analizde kullanılan veriden kaynaklanır. Bu veri, genel olarak deneysel olmayan, aynı zamanda, araştırmacının kontrolü dışında oluşan gözlemlerden oluşur. Günümüzde, her ne kadar örneklem yöntemleri geliştirilmiş olsa bile, kullanılan veri kümelerinin çoğu örneklem-içi ve örneklem-dışı (bir araya toplama, kodlama, uyarlama ve yerine kullanma gibi) hataların etkisi altında kalır. Klasik regresyon modelini etkileyen problemlerden, durağan olmama veya tahmin edilecek olan parametre sayısının gözlem sayısından daha çok sayıda olması problemi, kötü sunum (ill-posed) problemi olarak tanımlanır.

* Prof. Dr. Çukurova Üniversitesi, İktisadi ve İdari Bilimler Fakültesi, Ekonometri Bölümü, 01330 Adana. e-posta: haltan@cu.edu.tr

** Prof. Dr. Çukurova Üniversitesi, Fen Edebiyat Fakültesi, İstatistik Bölümü, 01330 Adana. e-posta: akdeniz@cu.edu.tr

Diğer yandan, parametre tahminlerinin oldukça kararsız olma problemi ise kötü-koşulluluk (ill-conditioned) problemi olarak tanımlanır.

Gözlemler deneysel olmayan yöntemlere bağlı üretildiklerinde veya veri üretme yöntemi ya da süreci kötü tasarlanmışsa, bu sorunla sıkça karşılaşılır. Parametre tahminlerinin oldukça kararsız olmasına neden olan böyle durumlarda içilişki problemi ortaya çıkar. Uygulamalı çalışmalarda, tam çoklu içilişki problemi ile karşılaşmak yok denecek kadar az olsa bile, önemli derecede çoklu içilişki problemi söz konusu olduğunda, EKK tahmin edicileri büyük varyanslı çıkar. Çoklu içilişkinin varlığı aşağıdaki durumlara neden olabilir:

1) Parametre tahminlerinde büyük sapmalara neden olur. Bu durumda parametre tahminlerinin kararlı olduğunu söylemek olanaklı değildir.
2) Parametre tahminlerinde artan değişkenlik olabileceği gibi yanlış işaretli çıkmasına da neden olur.
3) Parametrelere ilişkin hipotez testleri daha az güce sahip olacağından, hipotez testlerinden çıkarılan sonuçlarda güven eksikliği olacaktır.

Bu nedenlerden dolayı, uygulamalı çalışmalarda çoklu içilişki sorununun olma olasılığının oldukça iyi araştırılması gerekir. Golan vd. (1996, s.128) talep fonksiyonu tahmini çalışmalarında bu önemli problemi özellikle belirtmişlerdir.

Veride ortaya çıkan çoklu içilişki problemi ile ilgili birçok standart yaklaşım vardır. Bu standart yaklaşımların büyük bölümü, problemi yöntemine göre düzenleyen (parametreleri düzeltici) örneklem dışı önsel bilgiyi kullanan yaklaşımlardır. Bu nedenle, çoklu içilişki probleminin üstesinden gelebilmek için elde var olan bilgiden daha fazlasına gereksinim vardır. Tutarlı tahminlere ulaşmayı kolaylaştırmak için, sık sık ek varsayımlara başvurulur. Ancak, bu varsayımların veriler ile tutarlı olup olmadığına bakılmaz (Soofi, 1990).

Çoklu içilişki problemi söz konusu olduğunda, diğer bir tahmin yöntemi zayıf-koşulluluk ve kötü-sunumluluk gibi klasik tahmin problemlerinden etkilenmeyen ve son zamanlarda ortaya atılan maksimum entropi (ME) tahmin yöntemidir. ME prensibi, Shannon (1948) tarafından geliştirilen "bilgi kuramı kavramından" yararlanan Jaynes'in (1957a, b) çalışmasına dayanır. Jaynes "veri ile uyumlu modellerin seçilmesi gerektiğini aksi halde, uygun model seçilmemesi durumunda, kullanılan modellerin daha az bilgi verici olduğunu" önermektedir. ME, bilinmeyen parametrelerin tahminine farklı bir şekilde yaklaşır; ME gereksiz veya gelişigüzel varsayımları veya kısıtları kullanmaksızın, verinin modelde söz sahibi olmasına olanak tanır.

Bugüne kadar yapılan çalışmalardan elde edilen gelişmeler sonucunda GME tahmin edicisi ekonomik uygulamalarda yaygın olarak kullanılmaya başlanmıştır. Veri uygulamalı, GME içeren bilimsel kaynaklar genellikle eksik sunumlu problemlere odaklanmıştır (Fraser, 2000; Paris 2001; Paris ve Howitt , 1998; Campbell ve Hill, 2006).

Daha önce yapılan çalışmalardan sadece Golan vd. (1996) yapmış olduğu çalışmada, temel veri kümesinde kötü koşulluluk söz konusu olduğunda, GME'nin uygulanması düşünülmüştür. Golan vd. (1996) yapmış olduğu bu çalışmada, bünyesinde önemli derecede çoklu içilişki sorununun yer aldığı yapay bir veri kümesi oluşturmuştur. Değişik tahmin edicilerden (EKK, En çok olabilirlik (Maksimum likelihood), Ridge regresyon ve kısıtlı en küçük kareler) elde edilen tahminleri GME'den elde edilen tahminlerle karşılaştırdıklarında, en doğru (kesin) tahmini GME'nin verdiğini bulmuştur. Uygulamış oldukları Monte Carlo simulasyonu sonucunda GME için ölçüt olarak kullanılan "hata kareleri ortalaması"nın, diğer geleneksel tahmin edicilerden elde edilen değerlerle karşılaştırıldığında, önemli derecede düşük olduğunu bulmuştur. Böylece Golan vd. (1996), kötü koşulluluk söz konusu olduğunda, geleneksel tahmin edicilere oranla GME'nin daha güçlü olduğunu göstermiştir.

Çalışmada, Golan vd. (1996, s.86-89) tarafından geliştirilen GME tahmin edicisi kullanılmıştır. Golan vd. (1996) yaptığı bir çok örneklem çalışmasında (özellikle gözlemlerde yüksek derecede içilişki olduğunda), GME tahmin edicisinin hem EKK tahmin edicisinden, hem de kısıtlı en küçük kareler (KEKK=RLS) tahmin edicisinden daha küçük bir riske sahip olduğunu göstermişlerdir. (Golan vd., s.132-133). GME tahmin edicisinde tek bir parametreye yapılan kısıt ile ilgili uygulama Fraser (2000) ve Shen ve Perloff (2001)'ın çalışmalarında da vardır.

Bu çalışmada amaç, genel lineer modellerde (GLM), GME tahmin edicisini araştırmaktır. GME tahmin yöntemi parametreler için sınırları belirlemeyi gerektirdiğinden bu çalışmadaki uygulama içinde, hata terimi için uygun destek matrislerinin ve GME parametrelerinin nasıl belirleneceği incelenmiştir. GME parametreleri için belirlenen destek matrisini, önsel bilgiye, düşünülen eşitsizlik kısıtı ya da kısıtlarına uyumlu olarak değiştirerek, GME tahminlerinin duyarlılığı incelenmiştir.

Bu çalışmanın ikinci bölümünde GLM'de yer olan parametreler ve hata terimleri, uygun destek matrisleri kullanılarak, yeniden parametrelenmiş modelde parametre ve hata tahminlerinin elde edilişi verilmiştir. Üçüncü bölümde ABD'de piliç talebi veri kümesi (Gujarati, 1992) üzerinde Ridge ve Liu tahmin edicilerinde önce $\hat{k}$ ve $\hat{d}$ yanlılık parametreleri tahmin edilerek, sırasıyla Ridge ve Liu tahmin edicileri verilmiştir. Dördüncü bölümde, ABD'de tavuk talebi veri kümesi üzerinde GME tahmin edicileri için sayısal uygulama içinde eşitsizlik kısıtlı en küçük kareler (IRLS) tahminleri bulunmuştur. Ayrıca birden fazla parametrenin yer aldığı, çoklu parametre kısıtları ile GME tahminleri verilerek EKK, EKEKK (IRLS), Ridge ve Liu tahmin edicilerle GME tahmin edicileri hata kareleri ortalaması (HKO) ölçütüne göre karşılaştırılmıştır.

2. GENEL LİNEER MODEL (GLM)'DE GME TAHMİNİ

Golan vd. (1996), GLM'de bilinmeyen parametreleri ve hataları birlikte tahmin etmek için GME'yi kullanmıştır. GLM matris formunda (1) nolu eşitlikte verilen biçimde yazılır.

$$y = X\beta + e \tag{1}$$

Bu modelde:
$y : N \times 1$ bağımlı değişken üzerindeki örneklem gözlemler vektörü,
$X : N \times K$ açıklayıcı değişkenler matrisi (tasarım matrisi),
$e : N \times 1$ bilinmeyen hataların vektörü,
$\beta : K \times 1$ bilinmeyen parametrelerin vektörüdür.

Jaynes (1957a, 1957b), maksimum entropinin, kesikli olasılık dağılımında bilinmeyen olasılıkların tahmin edilmesine olanak tanıdığını göstermiştir. Golan vd. (1996), maksimum entropi yöntem bilimini genelleştirmiş, bilinmeyen parametreleri ve bilinmeyen hataları olasılık formunda ifade ederek doğrusal modeli yeniden parametrelemişlerdir. Destek noktaları ile ilgili bilinmeyen olasılıkları tahmin etmek için maksimum entropi kullanılmıştır. Böylece, bilinmeyen parametrelerin ve bilinmeyen hataların her ikisinin de önsel sınırlar içinde ele alındığı varsayılmıştır.

Parametre desteklerinin öncelikli belirlenmesi önsel bilgiye ya da iktisat teorisine dayanır. Her bir parametre ya da hata terimi için en küçük ve en büyük değerlerini belirlemek kolay bir iş değildir, çünkü iktisat teorisi genellikle bu bilgiyi sağlamayabilir. Bir parametre için doğru önsel bilgi olmadığında, parametre destekleri sıfır değer orta değer olacak şekilde geniş tutulur. Golan vd. (1996) bu önemli noktayı tartışmış ve "parametre destekleri için aralığı geniş tutmak riski azaltıcı sonuç verir'' şeklinde sonuçlandırmışlardır.

Her bilinmeyen parametre için, sıfır etrafında simetrik olma zorunluluğu olmayan ancak bilinmeyen parametreleri sınırları içine alan, $M \geq 2$ destek noktalarından oluşan bir matris tanımlansın.

$z_k : M \times 1$ k. parametre için destek vektörünü, $p_k : M \times 1$ k. parametre ile ilgili destek noktaları üzerindeki olasılıkların (veya ağırlıkların) vektörünü göstermek üzere bilinmeyen parametre vektörü β (2) nolu eşitlikteki gibi yazılabilir.

$$\beta = Zp = \begin{bmatrix} z_1' & 0 & . & . & . & 0 \\ 0 & z_2' & . & . & . & 0 \\ . & . & . & & & . \\ . & . & & . & & . \\ . & . & & & . & . \\ 0 & 0 & . & . & . & z_k' \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \\ . \\ . \\ . \\ p_k \end{bmatrix} \tag{2}$$

Burada:
$\beta : K \times 1$ bilinmeyen parametreler vektörünü,
$Z : K \times KM$ destek noktaları matrisini,
$p : KM \times 1$ bilinmeyen ağırlıklar vektörünü gösterir. Öyleki, tüm k'lar için $p_{km} > 0$ ve $p_k' i_M = 1$'dir.

Uygulamada bilinmeyen hatalar için sınırları oluşturmak oldukça zordur. Çalışmada, Pukelsheim (1994)'ün çalışmasından yararlanarak, hata sınırları için sırasıyla $(-3\sigma,3\sigma)$ ve $(-4\sigma,4\sigma)$ aralıkları kullanılmıştır. Bu kuralı kullanmak için σ'nın değeri ya tahmin edilmeli ya da bilinmelidir.

Her bir hata için $J \geq 2$ destek noktaları kümesi tanımlansın. Bu küme, sıfıra göre simetrik ve bilinmeyen hataları sınırlamaktadır. $v_i: J\times 1$ i. hata için destek vektörü ve $w_i: J\times 1$ destek noktaları üzerindeki ağırlıkların ilgili vektörü olsun. Bilinmeyen hata vektörü (3) nolu eşitlikle verilen biçimde yazılabilir.

$$e = Vw = \begin{bmatrix} v_1' & 0 & \cdots & 0 \\ 0 & v_2' & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & v_N' \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_N \end{bmatrix} \tag{3}$$

e: $N\times 1$ rastgele hatalar vektörünü,

V: $N\times NJ$ destek noktaların matrisini,

w: $NJ\times 1$ bilinmeyen ağırlıklar vektörünü gösterir. Öyleki, tüm i'ler için $w_{ij} > 0$ ve $w_i' i_j = 1$'dir. (2) ve (3), eşitliklerinin kullanılmasıyla (1) nolu GLM yeniden parametreleştirilmiş model olarak,

$$y = X\beta + e = XZp + Vw \tag{4}$$

ve

$$\beta = Zp\,, e = Vw \tag{5}$$

biçiminde yazılır. (4) denklemindeki yeniden parametrelemenin kullanılmasıyla standart lineer tahmin problemi, GME problemi olarak aşağıdaki amaç fonksiyonu ile formülleştirilebilir:

$$MaxH(p,w) = -\sum_{m=1}^{M}\sum_{k=1}^{K} p_{km}\ln p_{km} - \sum_{j=1}^{J}\sum_{t=1}^{N} w_{ij}\ln w_{ij} \tag{6}$$

Bu kısıtlı optimizasyon probleminde kısıtlar:

$$y = XZp + Vw \tag{7}$$

$$(I_K \otimes i_M')p = i_K\,, \tag{8}$$

$$(I_N \otimes i_J')w = i_N\,, \tag{9}$$

dir.

Burada;
$\otimes$: Kronoker çarpımı göstermektedir. (7) nolu denklem model kısıtını, (8) ve (9) nolu denklemler ise toplanabilirlik kısıtlarını gösterir. Kısıtlar, K parametrenin ve N hatanın her biri için olasılıklar toplamının 1 olmasını gerektirir. Parametre destek matrisi, blok-köşegen matris ise GME kısıtlı optimizasyon probleminin çözümleri

$$\hat{p}_{km} = \exp(z_{km} x_k' \hat{\lambda}) / \sum_{m=1}^{M} \exp(z_{km} x_k' \hat{\lambda}) \tag{10}$$

ve

$$\hat{w}_{nj} = \exp(v_{nj} \hat{\lambda}_n) / \sum_{j=1}^{J} \exp(v_{nj} \hat{\lambda}_n) \tag{11}$$

olacaktır (Fraser (2000) , Campbell ve Hill (2006)).

3. GENEL LİNEER MODELDE İÇİLİŞKİ OLMASI DURUMUNDA EKK, RİDGE VE LİU TAHMİNLERİ

Bu bölümde, Tablo1'deki veriyi ve (12) nolu denklemi kullanarak 4. bölümde GME tahminleri ile karşılaştırmak üzere EKK ile birlikte Ridge ve Liu yanlı tahminleri hesaplanmıştır. Doğrusal regresyon probleminde GME, Golan vd. (1996) tarafından ifade edilen ampirik Bayes ve Stein-Like tahmin edicilere benzer daraltıcı bir yanlı tahmin edicidir. (1) denklemi ile verilen $y = X\beta + e$ GLM'de β'yı tahmin etmek için genel olarak EKK tahmin edicisi $\hat{\beta} = (X'X)^{-1} X'y$ kullanılır. Açıklayıcı değişkenler arasında içilişki olması durumunda $\hat{\beta}$ kararlı bir tahmin edici değildir. Çalışmada kullanılan (12) nolu model log-lineer tavuk eti talep modelidir.

$$y = X\beta + u = \beta_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + u \tag{12}$$

Modelde yer alan değişkenler aşağıdaki gibi tanımlanmıştır:
$y = \ln Y =$ kişi başına tavuk eti tüketiminin logaritması
$x_2 = \ln X_2 =$ kişi başına harcanabilir reel gelirin logaritması
$x_3 = \ln X_3 =$ tavuk etinin perakende reel satış fiyatının[(1)] logaritması
$x_4 = \ln X_4 =$ domuz etinin perakende reel satış fiyatının logaritması
$x_5 = \ln X_5 =$ sığır etinin perakende reel satış fiyatının logaritması

Modelin değişkenlerine ilişkin gözlemler; (Gujarati, 1992, s.305) Tablo 1'de ve bu değişkenlerin doğal logaritmaları alındıktan sonra gözlemlere ilişkin ortaya çıkan özet istatistikler Tablo 2'de verilmiştir. (12) eşitliğindeki $\beta_2, \beta_3, \beta_4, \beta_5$ parametreleri sırasıyla, gelir, tavuk eti fiyatı, çapraz domuz eti fiyatı ve çapraz sığır eti fiyatı esneklik değerleridir.

(1): Reel fiyatlar, nominal fiyatların gıda için "Tüketici Fiyat Endeksi"ne bölünmesiyle elde edilir.

Bu parametrelerin beklenen işaretleri ise sırasıyla $\beta_2 > 0$, $\beta_3 < 0$, $\beta_4 > 0$ ve $\beta_5 > 0$'dır. Modeldeki $\hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3, \hat{\beta}_4, \hat{\beta}_5$ EKK tahminleri kullanılarak tahmin edilen regresyon fonksiyonu:

$$\underset{(st.hata.)}{\hat{y}} = \underset{(0.156)}{2.190} + \underset{(0.083)}{0.343}\, x_2 - \underset{(0.111)}{0.505}\, x_3 + \underset{(0.099)}{0.149}\, x_4 + \underset{(0.101)}{0.091}\, x_5$$

dır. EKK tahminlerine bakıldığında tüm esneklik değerlerinin beklenen işaretleri sağladığı görülür.

Tablo 1.Tavuk eti talebi modeli yıllık verileri (1960-1982)

Y	X_2	X_3	X_4	X_5
27.8	397.5	42.2	50.7	78.3
29.9	413.3	38.1	52.0	79.2
29.8	439.2	40.3	54.0	79.2
30.8	459.7	39.5	55.3	79.2
31.2	429.9	37.3	54.7	77.4
33.3	528.6	38.1	63.7	80.2
35.6	560.3	39.3	69.8	80.4
36.4	624.6	37.8	65.9	83.9
36.7	666.4	38.4	64.5	85.5
38.4	717.8	40.1	70.0	93.7
40.4	768.2	38.6	73.2	106.1
40.3	843.3	39.8	67.8	104.8
41.8	911.6	39.7	79.1	114.0
40.4	931.1	52.1	95.4	124.1
40.7	1021.5	48.9	94.2	127.6
40.1	1165.9	58.3	123.5	142.9
42.7	1349.6	57.9	129.9	143.6
44.1	1449.4	56.5	117.6	139.2
46.7	1575.5	63.7	130.9	165.5
50.6	1759.1	61.6	129.8	203.3
50.1	1994.2	58.9	128.0	219.6
51.7	2258.1	66.4	141.0	221.6
52.9	2478.7	70.4	168.2	232.6

Tablo 2'de gözlem değerlerinin logaritmaları için ortalama, minimum, maksimum, standart sapma ve değişim katsayılarının değerleri verilmektedir.

Tablo 2. Tavuk verisi için özet istatistikler

Değişken	Ortalama	Minimum	Maksimum	Standart sapma	Değişim katsayısı
$\ln Y$	3.664	3.325	3.968	0.188	0.051
$\ln X_2$	6.783	5.985	7.815	0.570	0.084
$\ln X_3$	3.847	3.619	4.254	0.222	0.058
$\ln X_4$	4.434	3.926	5.125	0.380	0.086
$\ln X_5$	4.751	4.349	5.449	0.380	0.080

Parametre tahminlerinin istatistiksel anlamlılıklarına bakıldığında ise, gelir esnekliğinin ($\hat{\beta}_2$) ve tavuk eti fiyat esnekliğinin ($\hat{\beta}_3$) anlamlı, ancak diğer yandan, rakip ürünler olan domuz eti çapraz fiyat esnekliğinin ($\hat{\beta}_4$) ve sığır eti çapraz fiyat esnekliğinin ($\hat{\beta}_5$) anlamsız olduğu görülmektedir. Teoride anlamlı olan, ancak uygulamada istatistiksel olarak anlamsız çıkan bu tahminler, tavuk eti talebinin rakip ürünler olan sığır eti ve domuz eti fiyatlarından etkilenmediğini göstermektedir. Bu durumun ortaya çıkması, "açıklayıcı değişkenler arasında çoklu içilişkinin varlığı probleminden kaynaklanabilir mi?" sorusunu araştırmayı gerektirir. Açıklayıcı değişkenler arasında içilişki olması durumunda $\hat{\beta}$ kararlı bir tahmin edici değildir.

Tablo 3'te açıklayıcı değişkenlerin logaritmalarının örneklem korelasyon matrisi verilmektedir.

Tablo 3. Örneklem korelasyon matrisi

	ln$Y$	lnX_2	lnX_3	lnX_4	lnX_5
y= ln$Y$	1.000	0.973	0.804	0.924	0.934
x_2=lnX_2	0.973	1.000	0.907	0.972	0.979
x_3=lnX_3	0.804	0.907	1.000	0.947	0.933
x_4=lnX_4	0.924	0.972	0.947	1.000	0.954
x_5=lnX_5	0.934	0.979	0.933	0.954	1.000

Görüldüğü gibi dört açıklayıcı değişken arasındaki ikili korelasyonlar 0.90'dan büyüktür ve $R^2 = 0.9823$'dür. Bununla birlikte böyle yüksek ikili korelasyonlar talep fonksiyonunun içilişkiden etkilenmesini garanti etmez. Belsley (1991) korelasyon ile içilişkinin aynı anlama gelmediğine dikkat çekmiştir. Açıklayıcı değişken çiftleri arasında korelasyon düşük olduğunda da, veri içilişkiye sahip olabilir. Örnekte açıklayıcı değişkenlerin, birinin geri kalanları üzerindeki regresyon denklemi yazılırsa tüm regresyonlar için R^2 değerlerinin 0.94'ü aştığı görülür (Gujarati, 1992 sayfa 306-307). F-testi uygulanırsa R^{2}'lerin istatistiksel olarak önemli olduğu sonucuna varılır. Bu sonuçlar regresyon denklemindeki her bir açıklayıcı değişkenin diğerleri ile yüksek içilişkili olduğunu gösterir.

Belsley vd. (1980) koşul sayısı = $\kappa = \sqrt{\lambda_{\max} / \lambda_{\min}}$ = 10 civarında olduğunda veride zayıf bir bağımlılık olduğunu; koşul sayısı 30-100 arasında ise orta düzeyden şiddetli bağımlılığa yaklaşıldığını ve koşul sayısı 100'den büyükse ciddi içilişki problemi olduğunu işaret etmiştir. Yukarıda EKK tahmini verilen Gujarati (1992) verisi için $X'X$'in özdeğerleri sırasıyla $\lambda_1 = 2406.9$, $\lambda_2 = 0.47683$, $\lambda_3 = 0.15906$, $\lambda_4 = 0.079068$, $\lambda_5 = 0.016011$; koşul sayısı $\kappa = \sqrt{\lambda_{\max} / \lambda_{\min}} = \sqrt{(2406.9)/0.016011}$ =387.72'dir. Bu sonuç veride şiddetli içilişki olduğunu göstermektedir. İçilişki olduğunda bu sorunu çözmek için Hoerl ve Kennard (1970)'da $\hat{\beta}_k = (X'X + kI)^{-1} X'y$ (k>0) Ridge regresyon tahmin edicisini önermiştir. Liu (1993)'te $\hat{\beta}_d = (X'X + I)^{-1}(X'y + d\hat{\beta}_{OLS})$'yi önermiştir. Model (3)'e dönerek, model kanonik forma indirgendiğinde;

$$y = Z\alpha + e \tag{13}$$

yazılır. Burada;

Q: ortogonal matris ve kolonları $X'X$ in özdeğerlerini oluşturmak üzere, $Z = XQ$ ve $\alpha = Q'\beta$'dır. O halde $Z'Z = Q'X'XQ = \Lambda$ =diag($\lambda_1,...,\lambda_p$), $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_p > 0$ $X'X$'in sıralı özdeğerleridir. (13) modeli için (i) ve (ii) ile iki farklı tahmin edici yazılabilir.

(i) Hoerl ve Kennard (1970)'ın önerdiği Ridge tahmin edicisi:

$$\hat{\beta}_k = (X'X + kI)^{-1} X'y \text{ ve } \hat{\alpha}_k = (\Lambda + kI)^{-1} \Lambda \hat{\alpha}_{OLS} \tag{14}$$

dir. Hata kareler ortalaması :

$$HKO(\hat{\alpha}_k) = \sum_{i=1}^{p} (\lambda_i \sigma^2 + k^2 \alpha_i^2)/(\lambda_i + k)^2 \tag{15}$$

dir (Akdeniz ve Erol, 2003). EKK regresyonundan $\hat{\sigma} = 0.028$ ve

$\hat{\beta}_{OLS} = (2.190, 0.343, -0.505, 0.149, 0.091)'$ olarak bulunur. k için iki farklı tahmin değeri kullanılabilir: Hoerl vd. (1975) tarafından önerilen

$$\hat{k}_{HKB} = p\hat{\sigma}^2 / \hat{\beta}'\hat{\beta} = p\hat{\sigma}^2 / \hat{\alpha}'\hat{\alpha} \tag{16}$$

ve Lawless ve Wang (1976)'ın önerdiği

$$\hat{k}_{LW} = p\hat{\sigma}^2 / \sum_{i=1} \lambda_i \hat{\alpha}_i^2 = p\hat{\sigma}^2 / \hat{\beta}'\Lambda\hat{\beta} \tag{17}$$

formülleri kullanılarak aşağıdaki sonuçlar elde edilir.

$\hat{k}_{HKB} = 0.00073233932$ için

$$\hat{\beta}_{HKB} = \begin{bmatrix} 2.1337238 \\ 0.35752892 \\ -0.46839471 \\ 0.12491947 \\ 0.074252764 \end{bmatrix}$$

dir. Tahmin edilmiş hata kareler ortalaması: $HK\hat{O}(\hat{\beta}_{HKB}) = 0.064784018$ olarak bulunur. $\hat{k}_{LW} = 0.000012297786$ için

$$\hat{\beta}_{LW} = \begin{bmatrix} 2.1888132 \\ 0.34282171 \\ -0.50395762 \\ 0.14812854 \\ 0.090805174 \end{bmatrix}$$

dir. Tahmin edilmiş hata kareler ortalaması: $HK\hat{O}(\ \hat{\beta}_{LW}) = 0.063481294$ 'dir.

(ii **)** (1) modeli için Liu (1993) tarafından önerilen Liu tahmin edicisi:

$$\hat{\beta}_d = (X'X + I)^{-1}(X'y + d\hat{\beta}_{OLS}) \tag{18}$$

dir. (13) kanonik modeli için Liu tahmin edicisi:

$$\hat{\alpha}_d = (\Lambda + I)^{-1}(\Lambda + d\hat{\alpha}_{OLS}),\ Q'\beta = \alpha\ , \hat{\beta}_{OLS} = Q\hat{\alpha}_{OLS} \tag{19}$$

dir. O halde

$$\hat{\beta}_d = \begin{bmatrix} 2.1633327 \\ 0.34113666 \\ -0.49515042 \\ 0.14769634 \\ 0.09182490 \end{bmatrix}$$

bulunur.

$$HKO(\hat{\beta}_d) = \mathrm{HKO}(\ \hat{\alpha}_d) = \sum_{i=1}^{p} \left[(\lambda_i + d)^2\sigma^2 + \lambda_i\alpha_i^{\ 2}(1-d)^2\right] / \lambda_i(1+\lambda_i)^2 \tag{20}$$

dir (Akdeniz ve Erol (2003). d' nin HKO'yu minimum yapan optimal değeri (Liu,1993)'te

$$d_{opt} = \left(\sum_{i=1}^{p} \frac{\alpha_i^{\ 2} - \sigma^2}{(1+\lambda_i)^2}\right) / \left(\sum_{i=1}^{p} \frac{\sigma^2 + \lambda_i\alpha_i^{\ 2}}{\lambda(1+\lambda_i)^2}\right) \tag{21}$$

biçiminde vermiştir. β ve σ^2 'nin EKK tahminleri kullanılırsa, tahmin edilmiş hata kareler ortalaması $HK\hat{O}(\hat{\alpha}_d)$ elde edilir. $\hat{d} = 0.98682586$ olduğundan tahmin edilmiş HKO

$HK\hat{O}(\hat{\beta}_d) = 0.062753678$ olarak bulunur. $HKO(\hat{\beta}_k) = HKO(\hat{\alpha}_k)$ ve $HKO(\hat{\beta}_d) = HKO(\hat{\alpha}_d)$ olduğu unutulmamalıdır.

4. PARAMETRELERİN GME İLE TAHMİN EDİLMESİ

GME yöntem biliminin uygulanması için en önemli neden, verinin içilişki yapısına sahip olmasıdır. Böylece içilişki potansiyel bir sorundur. X matrisi "kötü koşullu" olarak düşünülür. Bu bölümde (12) eşitliğindeki parametreler için GME tahminleri elde edilmiş ve üçüncü bölümde elde edilen tahminlerle birlikte tablolar halinde sunulmuştur. Modelin teorik özelliğine uygun parametre kısıtlarına da bu bölümde yer verilmiştir.

Veri kümesinin analizi için MATHEMATICA 5.0 paket programı kullanılmıştır. Ayrıca, GME tahminlerinin elde edilişinde, GAUSS 8 paket programından ve GAUSS 8 paket programının kısıtlı optimizasyon (constrained optimization) modülünden yararlanılmıştır.

4.1 Parametreler Üzerine Kısıt Olmadan GME Tahminleri

Bu bölümde GME kullanılarak modeldeki parametreler tahmin edilmiştir. Bilinmeyen parametreler ve hatalar için destek matrislerinin tanımlanması gerektiğinden, GME tahminleri tek bir küme şeklinde ortaya çıkmamaktadır. (10) ve (11) nolu eşitliklerde gösterildiği gibi GME tahminleri desteklere bağlıdır. Önsel belirlemelere göre GME tahminlerinin duyarlılığını incelemek için farklı parametre ve hata destekleri tanımlanabilir. Önce, parametre destekleri düşünüldüğünde, genel olarak, oldukça iyi önsel bilgi olmadığında ya da katsayının oldukça büyük çıkması beklendiğinde daha geniş sınırlar seçilir. (12) eşitliğindeki parametreler için önsel ortalamalar "0" olarak seçilmiştir. Destekler "0" a göre simetriktir. Bu modelde yok denecek kadar az miktarda önsel bilgiye sahip olunduğu varsayıldığından destekler için ortalaması sıfır olan sınırlar seçilmiştir. İlk olarak parametreler için kısıt olmaksızın önsel ortalamaların sıfır alındığı destek vektörleri Tablo 4'te verilmiştir.

Tablo 4. GME-S3 ve GME-S4 için parametre ve hata destekleri (kısıt yok)

Değişken	Parametre	Parametre desteği	Önsel ortalama
Sabit	β_1	$z_1' = \{-5 \;\; -2.5 \;\; 0 \;\; 2.5 \;\; 5\}$	0
x_2	β_2	$z_2' = \{-1 \;\; -0.5 \;\; 0 \;\; 0.5 \;\; 1\}$	0
x_3	β_3	$z_3' = \{-1 \;\; -0.5 \;\; 0 \;\; 0.5 \;\; 1\}$	0
x_4	β_4	$z_4' = \{-1 \;\; -0.5 \;\; 0 \;\; 0.5 \;\; 1\}$	0

Hata desteklerinin belirlenmesinde gerekli olan σ bilinmemektedir. Bu nedenle tahmin edilmelidir. a) EKK regresyonundan $\hat{\sigma} = 0.028$ bulunmuştur. b) y'nin örneklem standart sapması s =0.188 olarak hesaplanmıştır. Daha büyük olması nedeniyle y'nin örneklem standart sapması kullanılarak 3σ ve 4σ kuralıyla hata destekleri sırasıyla:

3σ	$v' = \{-0.54 \quad -0.27 \quad 0 \quad 0.27 \quad 0.54\}$
4σ	$v' = \{-0.72 \quad -0.36 \quad 0 \quad 0.36 \quad 0.72\}$

olacaktır. Tablo 5'te GME tahminleri ile birlikte parametreler için EKK, Ridge ve Liu tahminleri de verilmiştir. Ayrıca destek noktalarının sayısı için M=5 yerine, M=7 alarak parametre tahminlerinin nasıl etkilendiği de Tablo 5'in son iki kolonunda görülmektedir.

Tablo 5. Kullanılan veri için EKK, Ridge, Liu ve GME tahminleri (kısıt yok)

Değişken	Parametre	EKK	Ridge $\hat{k}_{HKB}$=0.00073	Liu $\hat{d}$=0.99	GME-S3 M=5	GME-S4 M=5	GME-S3 M=7	GME-S4 M=7
Sabit	β_1	2.190	2.134	2.163	1.511	1.385	1.462	1.338
x_2	β_2	0.343	0.357	0.341	0.263	0.230	0.249	0.219
x_3	β_3	-0.505	-0.468	-0.495	-0.062	-0.001	-0.037	0.018
x_4	β_4	0.149	0.125	0.148	0.053	0.065	0.057	0.071
x_5	β_5	0.091	0.074	0.092	0.078	0.092	0.084	0.096

Tablo 5 incelendiğinde büyüklük açısından GME ve OLS tahminlerinde farklılıklar olduğu ve işaret açısından fark olmadığı görülür. $\beta_1, \beta_2, \beta_4$ için GME tahminleri, EKK tahminlerinden daha küçük elde edilmiştir. β_3 için GME tahmini, EKK tahmininden daha büyüktür. β_5 için EKK tahmini GME-S3 ve GME-S4 tahminleri arasındadır. Ayrıca, Tablo 5'ten görüldüğü gibi tahminlerin normları: $\hat{\beta}'_{EKK}\hat{\beta}_{EKK} =$ 5.199> $\hat{\beta}'_d\hat{\beta}_d = 5.070 > \hat{\beta}'_{Ridge}\hat{\beta}_{Ridge} = 4.922 > \hat{\beta}'_{GME1S3}\hat{\beta}_{GME1S3} = 2.365 > \hat{\beta}'_{GME1S4}\hat{\beta}_{GME1S4} = 1.984$'tür. Tahmin edilmiş HKO'ların incelenmesinden $HK\hat{O}(\hat{\beta}_{EKK}) = 0.0654 >$ $HK\hat{O}(\hat{\beta}_{HKB}) = 0.0648 > HK\hat{O}(\hat{\beta}_{LW}) = 0.0634 > HK\hat{O}(\hat{\beta}_d) = 0.0627 > HK\hat{O}(\hat{\beta}_{GME})$ bulunur. Golan vd. (1996)'da tarafından belirtildiği gibi GME tahmin edicisi en küçük HKO'ya sahiptir.

Katsayıların beklenen işaretleri dikkate alınarak parametre destekleri değiştirildiğinde, bu durumda daha dar parametre sınırları konulabilir. Parametrelerin verilen aralığa düşeceği bilinmektedir.

4.2 Yalnız Parametre İşareti Kısıtlaması ile GME Tahminleri

Bu bölümde her bir katsayının beklenen işareti düşünülerek yeni parametre destekleri verilmiştir. Bu durumda kısıtlı GME'ler K1GME olarak gösterilmiştir. Önsel ortalamaların sıfırdan farklı değerler olduğu Tablo 6'da görülmektedir.

K1GME: Örneklem dışı olarak $\beta_3 \le 0$ bilgisine sahip olunması durumunda β_3 için destek vektörü yalnız negatif değerler olarak alınır. Örneğin, $z'_3 = [-0.8 \quad -0.6 \quad -0.4 \quad -0.2 \quad 0]$

alınabilir. Burada, z_3 $M \times 1$, β_3 için parametre destek vektörüdür. Bu durumda GME tahmini aşağıdaki gibi verilir: M=5 destek noktasının tümünde $\hat{p}_{3m} \geq 0$ olduğundan

$$\hat{\beta}_3 = -0.8\hat{p}_{31} - 0.6\hat{p}_{32} - 0.4\hat{p}_{33} - 0.2\hat{p}_{34} + 0\hat{p}_{35} \leq 0$$

bulunur. $\pm\, 3\sigma$ ve $\pm\, 4\sigma$ kuralıyla hata destekleri aşağıdaki gibidir.

3σ	$v' = \{-0.72 \quad -0.36 \quad 0 \quad 0.36 \quad 0.72\}$	0
4σ	$v' = \{-0.72 \quad -0.36 \quad 0 \quad 0.36 \quad 0.72\}$	0

Tablo 6. K1GME-S3 ve K1GME-S4 için parametre destekleri (yalnız işaret kısıtlı)

$\beta_1 \geq 0, \beta_2 \geq 0,\ \beta_3 \leq 0\ , \beta_4 \geq 0, \beta_5 \geq 0)$

Değişken	Parametre	Parametre desteği	Önsel ortalama
Sabit	β_1	$z_1' = \{0 \quad 0.75 \quad 1.50 \quad 2.25 \quad 3\}$	1.5
x_2	β_2	$z_2' = \{0 \quad 0.15 \quad 0.3 \quad 0.45 \quad 0.6\}$	0.3
x_3	β_3	$z_3' = \{-0.8 \quad -0.6 \quad -0.4 \quad -0.2 \quad 0\}$	- 0.4
x_4	β_4	$z_4' = \{0 \quad 0.05 \quad 0.10 \quad 0.15 \quad 0.20\}$	0.1
x_5	β_5	$z_5' = \{0 \quad 0.04 \quad 0.08 \quad 0.12 \quad 0.16\}$	0.08

Yalnız parametre işaret kısıtlaması ile elde edilen GME tahminleri Tablo 7'de verilmiştir.

Tablo 7. Kullanılan veri için EKK ve GME tahminleri (M=5)

Değişken	Parametre	EKK	K1GME-S3	K1GME-S4
Sabit	β_1	2.190	1.82766	1.78956
x_2	β_2	0.343	0.34437	0.34758
x_3	β_3	-0.505	-0.35048	-0.34723
x_4	β_4	0.149	0.10292	0.10330
x_5	β_5	0.091	0.08214	0.08236

4.3 $\beta_i > \beta_j$ Kısıtlaması ile GME Tahminleri

Tavuk eti talebi değişkeni bireysel olarak hem domuz eti fiyatı değişkeni, hem de sığır eti fiyatı değişkeni için ayrı ayrı EKK kullanılmıştır. Basit bireysel regresyonlar sonucunda sığır eti fiyat değişkeninin, tavuk eti talebini daha büyük pozitif oranda etkilediği sonucu elde edilmiştir. Bu nedenle, bulgulara dayanarak $\beta_5 > \beta_4 > 0$ kısıtlaması düşünüldüğünde parametreler için

$$\begin{bmatrix} \beta_4 \\ \beta_5 \end{bmatrix} = Z.\begin{bmatrix} p_4 \\ p_5 \end{bmatrix} = \begin{bmatrix} z_4' & 0 \\ z_4' & z_5' \end{bmatrix} . \begin{bmatrix} p_4 \\ p_5 \end{bmatrix}$$

eşitliği yazılabilir. Burada Z : $2 \times 2M$ β_4 , β_5 parametreleri ve p_4 , p_5 bilinmeyen olasılıkları ile ilgili destek noktalarının matrisidir. Kısıtların bu kümesi için parametre destekleri K2GME ile aşağıdaki gibi belirlenebilir.

K2GME: $\hat{\beta}_4$ ve $\hat{\beta}_5$ ün her ikisi de pozitif olarak kısıtlandığından β_5 için tahmin,

$\hat{\beta}_5 = \hat{\beta}_4 + z_5' \hat{p}_5 \geq \hat{\beta}_4$ biçimindedir. z_5' destek vektörünün elemanları pozitif olduğundan β_5' in tahmini için önsel ortalama $\hat{\beta}_4 + 0.08$ alınabilir.

Yeniden $\pm 3\sigma$ ve $\pm 4\sigma$ sınırları kullanılarak hata sınırları aşağıda verilmiştir.

3σ	$v' = \{-0.54 \quad -0.27 \quad 0 \quad 0.27 \quad 0.54\}$
4σ	$v' = \{-0.72 \quad -0.36 \quad 0 \quad 0.36 \quad 0.72\}$

K2GME ile verilen kısıtlamaya uygun olarak parametreler için önsel ortalamalar ve parametre destekleri Tablo 8'de verilmiştir.

Tablo 8. K2GME-S3 ve K2GME-S4 için parametre destekleri

$(\beta_1 \geq 0, \beta_2 \geq 0, \ \beta_3 \leq 0 \ , \beta_4 \geq 0, \beta_5 \geq 0)$

Değişken	Parametre	Parametre desteği	Önsel ortalama
Sabit	β_1	$z_1' = \{0 \quad 0.75 \quad 1.50 \quad 2.25 \quad 3\}$	1.5
x_2	β_2	$z_2' = \{0 \quad 0.15 \quad 0.3 \quad 0.45 \quad 0.6\}$	0.3
x_3	β_3	$z_3' = \{-0.8 \quad -0.6 \quad -0.4 \quad -0.2 \quad 0\}$	- 0.4
x_4	β_4	$z_4' = \{0 \quad 0.05 \quad 0.10 \quad 0.15 \quad 0.20\}$	0.1
x_5	β_5	$z_5' = \{0 \quad 0.04 \quad 0.08 \quad 0.12 \quad 0.16\}$	$\hat{\beta}_4$+0.08

$\beta_5 > \beta_4$ kısıtlaması ile IRLS ve GME tahminleri Tablo 9'da verilmiştir.

Tablo 9. Kullanılan veri için EKK, IRLS ve GME tahminleri

Değişken	Parametre	EKK	IRLS	K2GME-S3	K2GME-S4
Sabit	β_1	2.190	2.18141	1.77006	1.71287
x_2	β_2	0.343	0.35291	0.30441	0.30947
x_3	β_3	-0.505	-0.50988	-0.38268	-0.37986
x_4	β_4	0.149	0.11395	0.10010	0.10115
x_5	β_5	0.091	0.11544	0.18009	0.18151

Tablo 9'da bulunan IRLS tahmin edicisinin yanlı fakat kısıtlar doğru olduğu sürece karesel hata kayıp fonksiyonu ölçütüne göre EKK'dan üstün olduğu sonucu elde edilir.

K3GME- Şimdiye kadar elde edilen tahminlerde kişi başı harcanabilir gelir değişkeninin tavuk eti talebini pozitif yönde, hem sığır eti fiyat değişkeninden, hem de domuz eti fiyat değişkeninden daha büyük oranda etkilediği bulgularına ulaşılmıştır. Bu bulgular kullanılarak GME ile $\beta_2 > \beta_5 > \beta_4 > 0$ eşitsizlik kısıtlaması da uygulanabilir. Bu kısıtlama düşünüldüğünde destek matrisi aşağıdaki gibi olacaktır:

$$\beta = Zp = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \beta_5 \end{bmatrix} = \begin{bmatrix} z_1' & 0 & 0 & 0 & 0 \\ 0 & z_2' & 0 & z_4' & z_5' \\ 0 & 0 & z_3' & 0 & 0 \\ 0 & 0 & 0 & z_4' & 0 \\ 0 & 0 & 0 & z_4' & z_5' \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \\ p_5 \end{bmatrix}$$

$\hat{\beta}_{GME} = Z\hat{p}$ eşitliğinin düşünülmesiyle $\hat{\beta}_4 = z_4'\hat{p}_4$, $\hat{\beta}_5 = z_4'\hat{p}_4 + z_5'\hat{p}_5 = \hat{\beta}_4 + z_5'\hat{p}_5 > \hat{\beta}_4$
$\hat{\beta}_2 = z_2'\hat{p}_2 + z_4'\hat{p}_4 + z_5'\hat{p}_5 = z_2'\hat{p}_2 + \hat{\beta}_4 + (\hat{\beta}_5 - \hat{\beta}_4) = z_2'\hat{p}_2 + \hat{\beta}_5 > \hat{\beta}_5 > \hat{\beta}_4$ eşitsizliği yazılır. Burada verilen eşitsizlik kısıtlamasına uygun olarak parametreler için önsel ortalamalar ve destek vektörleri de Tablo 10'da verilmiştir.

Tablo 10. K3GME-S3 ve K3GME-S4 için parametre destekleri

$(\beta_1 \geq 0, \beta_2 \geq 0,\ \beta_3 \leq 0\ , \beta_4 \geq 0, \beta_5 \geq 0)$

Değişken	Parametre	Parametre desteği	Önsel ortalama
Sabit	β_1	$z_1' = \{0 \quad 0.75 \quad 1.50 \quad 2.25 \quad 3\}$	1.5
x_2	β_2	$z_2' = \{0 \quad 0.15 \quad 0.3 \quad 0.45 \quad 0.6\}$	$\hat{\beta}_5 + 0.3$
x_3	β_3	$z_3' = \{-0.8 \quad -0.6 \quad -0.4 \quad -0.2 \quad 0\}$	- 0.4
x_4	β_4	$z_4' = \{0 \quad 0.05 \quad 0.10 \quad 0.15 \quad 0.20\}$	0.1
x_5	β_5	$z_5' = \{0 \quad 0.04 \quad 0.08 \quad 0.12 \quad 0.16\}$	$\hat{\beta}_4$+0.08

K3GME ile verilen eşitsizlik kısıtlamaları ile parametreler için IRLS ve GME tahminleri Tablo 11'de verilmiştir.

Tablo 11. Çoklu eşitsizlik kısıtlı tahminler

$\beta_2 > \beta_5 > \beta_4 \geq 0$

Değişken	Parametre	EKK	IRLS	K2GME-S3	K2GME-S4
Sabit	β_1	2.190	1.90879	1.68418	1.58479
x_2	β_2	0.343	0.238528	0.37843	0.39043
x_3	β_3	-0.505	-0.156261	-0.43393	-0.43599
x_4	β_4	0.149	0.003189	0.08093	0.08139
x_5	β_5	0.091	0.153203	0.15198	0.15535

5. TARTIŞMA VE SONUÇ

GLM'nin açıklayıcı değişkenleri arasında önemli derecede çoklu içilişki problemi olduğunda parametre tahminlerinde karasızlık vardır. Bu nedenle, model parametrelerinin kararlı tahminlerini elde etmek gerekir. Bu çalışma, çoklu içilişkinin bulunduğu bir modelde kararlı tahminleri bulmayı amaçlamıştır. Geliştirilmiş yanlı tahmin edicilerden Ridge ve Liu tahmin edicilerine ek olarak, son yıllarda ekonomi kaynaklarında yaygın kullanım alanı bulan genelleştirilmiş maksimum entropi tahmin edicisi kullanılmış ve HKO ölçütüne göre karşılaştırma yapılmıştır. Açıklanan tahmin edicilere ek olarak, teoriyle uyumlu parametre eşitsizlik kısıtlarına da yer verilerek GME ve eşitsizlik kısıtlı EKK (IRLS) tahminleri yapılmıştır. Kullanılan örneklemden elde edilen sonuçlar GME tahminlerinin EKK tahminlerine, Ridge ve Liu tahminlerine yakın olduğunu göstermektedir. Ayrıca, bulunan sonuçlara göre, parametre destekleri değiştirilen orana göre GME tahminleri daha az değişiklikler göstermektedir. Sonuçlar GME tahmin edicileri ve EKK tahmin edicilerinde büyüklük farkı vermekle birlikte işaret farkı vermemiştir.

Tablo 5'teki sonuçlara göre $\left\|\hat{\beta}_{GME}\right\|^2$ en küçüktür. Önsel bilgiye dayalı parametre destekleri uygulandığında elde edilen GME ve EKK tahminleri, önsel bilgiye dayalı olmayan parametre destekleri uygulandığında elde edilen sonuçlarla genel olarak tutarlılık göstermiştir. Beklenildiği gibi, daha geniş hata sınırları kullanıldığında (GME-S4), katsayılar genellikle önsel ortalamalarına doğru çekilir (büzülür). Bu durumda, hatalara daha fazla ağırlık verilmesi, daha düzgün parametre desteklerine bağlı olasılıkların elde edilmesini ifade eder.

Parametre destek noktaları M=5 yerine, Tablo 5'te görüldüğü gibi M=7 alındığında GME ile yapılan incelemede parametre tahminlerinin çok az değiştiği gözlenmiştir.

EKK, eşitsizlik kısıtlı EKK (IRLS), Ridge regresyon tahminleri ve Liu tahmini ile karşılaştırıldığında GME tahminleri daha küçük tahmin edilmiş hata kareleri ortalamasına sahiptir. Hata sınırları $\pm 3\sigma$ ve $\pm 4\sigma$ olarak alındığında, ekonomik olarak anlamlı çoklu kısıtlamalar verildiğinde diagonal olmayan parametre destek matrisinin kullanılmasıyla kısıtlı GME tahminleri Tablo 11'de verilmiştir. Bu durumda GME'nin daraltıcı yapısı

$$\left\|\hat{\beta}_{IRLS}\right\|^2 = 3.7482737 > \left\|\hat{\beta}_{GMES3}\right\|^2 = 3.19798 > \left\|\hat{\beta}_{GMES4}\right\|^2 = 2.285136$$

eşitsizliklerinden de görülmektedir.

Teşekkür. Bu çalışmanın iyileştirilmesinde katkı sağlayan hakemlere, Dil Editörü ve Editör Prof. Dr. Fetih Yıldırım'a teşekkürlerimizi sunuyoruz.

6. KAYNAKLAR

Akdeniz, F. and Erol, H., 2003. Mean squared error matrix comparisons of some biased estimators in linear regression. Communications in Statistics-Theory and Methods 32(12), 2389-2413.

Belsley, D. A., Kuh, E., Welsch, R. E., 1980. Regression diagnostics, New York, Wiley.

Belsley, D. A., 1991. Conditioning diagnostics: Collinearity and weak data in Regression. Wiley Series, New York.

Campbell, R. C. and Carter Hill, R., 2006. Imposing parameter inequality restrictions using the principle of maximum entropy. Journal of Statistical Computation and Simulation 76 (11), 985-1000.

Fraser, I., 2000. An application of maximum entropy estimation: The demand for meat in the United Kingdom. Applied Economics 32, 45-59.

Golan, A., Jodge, G. and Miller, D., 1996. Maximum entropy econometrics. John Wiley and Sons. New York.

Gujarati, D., 1992. Essentials of econometrics. McGraw-Hill International Editions, New York.

Hoerl, A. E. and Kennard, R.W., 1970. Ridge regression: Biased estimation for orthogonal Problems. Technometrics 12, 55-67.

Hoerl, A. E. , Kennard, R., Baldwin, K. F., 1975. Ridge regression: Some simulations. Communications in Statistics-Theory and Methods 4(2), 105-123.

Jaynes, E. T., 1957a. Information theory and statistical mechanics. Physics Review 106, 620-630.

Jaynes, E. T., 1957b. Information theory and statistical mechanics II. Physics Review 108, 171-190.

Lawless, J. F., Wang, P., 1976. A simulation study of Ridge and other regression estimators. Communications in Statistics-Theory and Methods A5, 307-323.

Liu, Kejian,1993. A new class of biased estimate in linear regression. Communications in Statistics Theory and Methods 22, 393-402.

Paris, Q. And Howitt, R. E. 1998. An analysis of ill-posed production problems using maximum entropy. American Journal of Aggricultural Economics, 80, 124-138.

Paris, Q. 2001. MELE: Maximum entropy Leuven estimators. University of California Davis, Working Paper 01-003.

Pukelsheim, F., 1994. The three sigma rule. American Statistician 48, 88-91.

Shannon, C. E., 1948. A mathematical theory of communication. Bell System Technical Journal 27, 379-423.

Shen, E. Z., Perloff, J.M., 2001. Maximum entropy and Bayesian approaches to the ratio problem. Journal of Econometrics, 104, 289-313.

Soofi, E. S.,1990. Effects of collinearity on information about regression coefficients. Journal of Econometrics, 43, 255-274.

MULTICOLLINEARITY AND GENERALIZED MAXIMUM ENTROPY ESTIMATORS

ABSTRACT

In this paper, we have considered the general linear model (GLM) $y = X\beta + u$ in the usual notation. In many applications the design matrix X is frequently subject to severe multicollinearity. In the presence of multicollinearity certain biased estimators, like the ordinary Ridge regression estimator $\hat{\beta}_k = (X'X + kI)^{-1} X'y$ and the Liu estimator $\hat{\beta}_d = (X'X + I)^{-1}(X'y + d\hat{\beta}_{OLS})$ or improved Ridge and Liu-type estimators, are used to outperform the ordinary least squares (OLS) estimates in the linear model. In this paper an alternative estimation methodology, maximum entropy, is given and used to estimate the parameters in a linear regression model when the basic data are ill-conditioned. We described the generalized maximum entropy (GME) estimator and develop a method for imposing parameter inequality restrictions through the GME parameter support matrix. We compared the GME estimator to the alternative estimation methodologies (least squares estimator, inequality restricted least squares (IRLS) estimator, Ridge regression estimator and Liu estimator) analyzed empirically for a US chicken demand data set.

Key words: Generalized Maximum Entropy Estimator, Inequality Restricted Least Squares Estimator, Liu Estimator, Ridge Regression Estimator, Support Point.