*D.E.Ü. İ.İ.B.F. Dergisi*
Cilt: 7, Sayı: 2, 1992, ss.226-239.

# ÇOKLU DOĞRUSAL BAĞLANTI VE ORTAYA ÇIKARDIĞI SONUÇLAR

Mustafa GÜNEŞ (*) Ercan BALDEMİR (**)

*ÖZET*

***Bu çalışmanın amacı çoklu doğrusal bağlantı durumunu araştırmak ve paremetrelerle katsayılarının güvenilirliğini tartışmaktır.***

***Çalışmada çoklu doğrusal bağlantının sebep olduğu sonuçlar kısaca anlatıldıktan sonra paremetrelerin beklenen değerleri ve varyansları elde edilmiştir. Daha sonra tam çoklu doğrusal bağlantı durumu incelenmiş ve öncelikle de iki açıklayıcı değişkenden oluşan model tartışılmış ve daha sonra da $b_2$ ve $b_3$ parametre tahminlerinin çözümlerinin tanımsız olduğu görülmüştür. Bu bulgulardan sonra, ikiden fazla açıklayıcı değişkene sahip modellerde korelasyon katsayılarının çoklu doğrusal bağlantının varlığını anlamada yetersiz kaldığı görülmüştür.***

***Çalışmanın son bölümünde de, yüksek seviyeli çoklu doğrusal bağlantı durumu tartışılmış ve daha sonra da herhangi iki açıklayıcı değişken arasındaki korelasyon katsayısının 1'e yakın olması nedeniyle, katsayı tahminlerinin varyans ve kovaryanslarının oldukça büyük olduğu görülmüştür.***

## 1- Giriş

Çoklu Regresyon Varsayımlarından biri de (X'X) matrisi, rankı k olan (k*k) lık bir kare matrisi ve dolayısıyla $(X'X)^{-1}$'in hesaplanabildiği varsayımıdır. Yani X matrisi (n*k) tertibinde dir. Modelde yer alan parametrelerin tahmin edilebilmesi için gözlem sayısının parametre sayısından fazla (n>k) olması gerekir. n=k olduğunda yalnız örneğe bağlı kesin bir ilişki, n<k olduğunda ise sonsuz sayıda kesin ilişki elde edilir. Meselâ, $Y=b_0+b_1X\varepsilon$ modelinde iki parametre bulunmakta olup; gözlem sayısı 2 olduğunda bu noktadan geçen sonsuz sayıda doğrusal ilişki bulunur. İhtimalli ilişkilerle ilgilendiğimiz

(*) Yrd.Doç.Dr.D.E.Ü.İ.İ.B.F., Ekonometri Bölümü
(**) Arş.Gör.D.E.Ü.İ.İ.B.F., Ekonometri Bölümü

için her iki durumda arzu edilmez. X matrisinin rankının k'ya eşit olması ise bu matrisin vektörlerinin birbirinden bağımsız olduğunu gösterir. Bu durum b vektörünün tahmin edici olan $b=(X'X)^{-1}X'Y$ formülündeki $(X'X)^{-1}$ matrisinin bulunmasını mümkün kılar. En az bir vektör diğer bir vektöre (veya vektörlere) doğrusal olarak bağımlı ise X matrisinin rankı k'dan küçük olur ve $|X'X|=$ olur. O zaman $(X'X)^{-1}$ matrisi var olmayacaktır. İşte iki veya daha fazla bağımsız değişkenin kendi aralarında sıkı bir şekilde bağımlı olmaları neticesinde ortaya çıkan bu durum "Çoklu Doğrusal Bağlantı (Multicallinearity)" olarak tanımlanmaktadır Bu durum daha çok zaman serileri verilerinde görülmekte ve Klasik Doğrusal Regresyon Modeli'nden bir sapmayı ifade etmektedir. (Ertek, 1987; 139)

Klasik Normal Doğrusal Regresyon Modeli varsayımlarında açıklayıcı değişkenlerden hiçbirinin diğer bir açıklayıcı değişken veya değişkenlerle karşılıklı sıkı ilişkisinin olmaması arzu edilir. Bu varsayım ihlal edildiğinde Çoklu Doğrusal Bağlantı ortaya çıkar. Diğer taraftan, açıklayıcı değişkenler birbirlerinden bağımsız olduğunda Çoklu Doğrusal Bağlantının yokluğundan söz edilir. Açıklayıcı değişkenler arasındaki bu durumlar çeşitli Çoklu Doğrusal Bağlantı dereceleriyle tanımlanır. En önemli özellik, yüksek dereceli Çoklu Doğrusal Bağlantı durumudur ki, bu da, bir açıklayıcı değişkenler kombinasyonuyla karşılıklı çok sıkı ilişkisinden kaynaklanır.

## 2. ÇOKLU DOĞRUSAL BAĞLANTININ SEBEPLERİ

Çoklu Doğrusal Bağlantının detaylı tartışmasına başlamadan önce şu iki noktayı açıklığa kavuşturmakta fayda vardır.

1) Çoklu Doğrusal Bağlantı bir çeşit değil derece meselesidir. Anlamlı fark çoklu doğrusal bağlantının varlığı ve yokluğu arasında değil, onun çeşitli dereceleri arasındadır.

2) Çoklu Doğrusal Bağlantının nonstokastik olduğu farzedilen açıklayıcı değişkenlerin durumunu göstermesinden dolayı, o popülasyonun değil örneğin bir özelliğidir.

Dolayısıyla çoklu doğrusal bağlantının testi yerine eğer istersek onun herhangi özel bir örnekteki derecesini ölçebiliriz. (Kmenta, 1984; 431)

Çoklu Doğrusal Bağlantı çeşitli nedenlerle ortaya çıkabilir. İlk olarak, iktisadi değişkenlerin zaman içinde birlikte değişme eğilimleri vardır. Birçok iktisadi büyüklükler aynı faktörlerden etkilenirler ve bunların etkileri bir kez devreye girdikten sonra, iktisadi değişkenler zaman içinde kabaca aynı davranış kalıbına uyarlar. Meselâ hızlı gelişme dönemlerinde temel iktisadi büyüklükler, bazıları gecikmeli de olsa hep birden büyürler. Böylece gelir, tüketim, tasarruf, yatırım, fiyatlar ve istihdam değerleri iktisadi genişleme dönemlerinde yükselir, daralma dönemlerinde ise düşerler. Zaman serilerindeki büyüme ve genel eğilim faktörleri Çoklu Doğrusal Bağlantının en ciddi sebepleridir. İkinci nokta ise bazı açıklayıcı değişkenlerin gecikmeli değerlerinin ilişkide ayrı birer faktör olarak kullanılmasıdır. Bir değişkenin ardışık değerlerinin birbiriyle bağlantılı olması tabiidir. Meselâ içinde bulunan dönemin geliri, kısmen önceki yılların gelirine bağımlılık gösterir. Dolayısıyla geçikmeli modellerde Çoklu Doğrusal Bağlantı bulunması hemen hemen kesin gibidir. (Koutsoyiannis, 1989; 237)

İktisatçılar genellikle deney yapamayan bilim adamaları olarak tanımlanırlar. Deney yapabilen bilim adamları kontrol edilebilir değişkenleri izleyerek, yeterli örnekler elde edip, çeşıtli şartlar altında arzulanan neticelere doğrudan ulaşabilirler. Fakat deney yapamayan bilim adamları ise önceden toplanmış olan bilgileri, verileri değerlendirir. Bu veriler genellikle toplum veya çeşitli ajanslar tarafından toplanmıştır. Sonuç olarak iktisatçıların kullandıkları veriler ve deney kalıpları onların seçimi değildir. Gözlem imkânları sınırlıdır, değişken çeşitleri kısıtlıdır. Bu yüzden seçilen örnekler yeterli bilgi veremez ve araştırmaları kesin delillerle desteklemek çok zor olduğundan sorulara tam cevap alınamaz. Bilgi, veri yetersizliği, buna bağlı olarak istatistiki sonuçların kesin olamaması Çoklu Doğrusal Bağlantıların oluşmasına neden olur.

## 2.1 Çoklu Doğrusal Bağlantılı Modeller

Deney modellerinin yetersiz olması ve değişkenler arası ilişkilerin varlığı, deney değişkenlerinin bir matris notasyonu içinde sunulmasını zorunlu kılar. Bunun örnekleri çoktur. Meselâ, hane halkı giderleri servet ve gelirinin bir fonksiyonu olup bu açıklayıcı değişkenlerin pozitif bir etkisi vardır. Az servetli ailelerin geliri az, çok servetli ailelerin geliri çoktur. Burada servet ve gelir "çoklu doğrusal bağlantı" dır.

Çoklu Doğrusal Bağlantıya sahip modellerde şu hususlara dikkat edilmelidir.

1) Değişkenlerin ayrı ayrı, deney üzerindeki etkilerini kesin olarak tanımlamak çok zordur. Regresyon katsayılarının, bağıntılı değişkenlerin etkilerini yansıtmasından dolayı, diğer bütün veriler sabit olarak tutulur. Katsayıları tanıyabilme eğilimi daha güçlü ve Çoklu Doğrusal Bağlantı için daha önemlidir. Bilinmeyen parametreler ve etkilenen değişkenler arası güçlü korelasyonlar değişkenler arasında var olabilecek kesin neticelerin oluşmasını engeller.

2) Bilinmeyen parametreler sıfırdan farklı olarak oluşmayabilir. Bu değişkenlerin etkisiz olduğundan değil, etkilerinin tam olarak hesaplanamamasından kaynaklanır. Bunun yüksek $R^2$ veya F değerleri belirli deney güçlerini gösterir.

3) Parametre tahminlerini yaparken toplamlara çok dikkat edilmesi ve bazı gözlemlerin denklemden çıkarılması durumunda hassas olunması gerekir.

4) Tek tek değişkenlerin ayrılmasındaki zorluklara rağmen, doğru neticeler almak yine mümkündür. Bu ise çoklu bağlantı dizilişinin deneydeki değişkenlerle aynı aralıkta olması ve ele alınan örneğinde bu aralığa ait olması ile gerçekleşir.

Bu alanda yapılan araştırmalar, bir problemin en iyi çözümümün bol bilgi toplanmasından geçtiğini göstermiştir. Bu bilgi yeni veriler, teorik ilişkiler veya istatistiki neticeler olabilir. (Wiley, 897)

En küçük karelerin uygulanabilmesi için çok önemli bir şart, açıklayıcı değişkenlerin kendi aralarında tam bir doğrusal bağlantıya sahip olmamalarıdır. ($r_{x_i x_j} \neq 1$). Çoklu Doğrusallık terimi, açıklayıcı değişkenler arasında doğrusal yada doğrusala yakın ilişkilerin varlığını ifade etmekte kullanılır. Eğer açıklayıcı değişkenler arasında tam bir doğrusal bağlantı varsa, yani, bu değişkenler için korelasyon katsayısı bire eşitse, parametreler belirlenemez hale gelir, herbir parametre için ayrı ayrı sayısal değerler bulmak imkansızlaşır ve en küçük kareler yöntemi işlemez olur. Diğer taraftan, eğer açıklayıcı değişkenler arasında hiçbir bağlantı yoksa, yani, bu değişkenler için korelasyon katsayısı sıfıra eşitse, bunlara dikey (ortogonal) değişkenler denir ki bu değişkenler ortak varyansı sıfır olan değişkenlerdir. ($S_{x_i x_j}/n=0$). Bu katsayıların tahmininde, hiç olmazsa çoklu doğrusallık bakımından, hiçbir problemin olmadığı anlamına gelir. Aslında dikey X'ler durumunda çoklu regresyon çözümlemesine gitmeye bile gerek yoktur çünkü herbir $b_j$ parametresi, ilgili değişkenle Y arasındaki basit bir regresyonla tahmin edilebilir: $Y=f(x_i)$. (Koutsoyiannis, 1989; 236)

Uygulamalarda bu gibi durumlarla pek sık karşılaşılmaz. Çünkü çoğu iktisadi değişkenin zaman içinde birbirlerine bağlı olmaları sebebiyle açıklayıcı değişkenler arasında azda olsa çoklu doğrusal bağlantı bulunur. Bu durumda her açıklayıcı değişken çifti için basit korelasyon katsayısı sıfır ile bir arasında bir değer alacak ve çoklu doğrusal bağlantı problemi, parametre tahminlerinin doğruluğunu ve kararlılığını bozabilecektir.

## 3- ÇOKLU DOĞRUSAL BAĞLANTININ ORTAYA ÇIKARDIĞI SONUÇLAR:

Kuvvetli Çoklu Doğrusal Bağlantı halinde En Küçük Kareler Tahmincilerinin, en iyi doğrusal sapmasız tahminci özelliklerini muhafaza etmelerine rağmen, Çoklu Doğrusal Bağlantı arttıkça, her bağımsız değişkenin bağımlı değişken üzerindeki net etkisini gösteren regresyon katsayısı değeri elde edilemeyecektir. Çoklu Doğrusal Bağlantının pratikte ortaya çıkardığı sakıncaları şöyle sıralayabiliriz.

1- Regresyon katsayılarının değerlerinin belirsiz olması

2- Regresyon katsayılarının varyanslarının ve dolayısıyla güven aralıklarının büyümesi

3- Hesaplanan t değerlerinin küçülmesi

4- $R^2$ nin olduğundan fazla büyümesi

5- $b_i$ tahminleri ve standart hatalarının, verilerdeki küçük değişiklerden önemli ölçüde etkilenmeleridir. (Akkaya, 1989; 433)

Eğer açıklayıcı değişkenler arasındaki ilişki tamsa ($rx_ix_j=1$) o zaman;

a) Katsayıların tahminleri belirlenemez. Yani eğer açıklayıcı değişken $x_1$ ve $x_2$ arasındaki $x_2=kx_1$ şeklinde tam bir ilişki varsa o zaman $\hat{b}_1$ ve $\hat{b}_2$ formüllerinde $x_2$ yerine $kx_1$ değeri koyduğumuzda her iki sonuçta $\frac{0}{0}$ bulunur ki bu da belirsizdir.

b) Bu tahminlerin standart hataları sonsuz büyük olur. Yine aynı şekilde $x_2=kx_1$ ise bunu $\hat{b}_1$ ve $\hat{b}_2$ nin varyansları formüllerinde $x_2$ yerine yine $kx_1$ yazdığımızda payda sıfır çıkacağı için her iki sonuçta $\infty$ olacaktır.

### 3.1. Çoklu Doğrusal Bağlantının Yokluğu

Açıklayıcı değişkenlerin diğerleriyle karşılıklı ilişkisiz olduğu ve dolayısıyla Çoklu Doğrusal Bağlantının olmadığı durumlarda (X'X) matrisi ortagonaldir. Bu da çoklu regresyon modellerinde basit regresyon kullanmakla açıklayıcı değişkenlerin bağmlı değişken üzerindeki etkisinin ölçülebileceğini gösterir. Bu yol görünüşte daha isabetli ve problemsiz bir yol olmasına rağmen, o kadar çok önemli olmasa bile çoklu regresyondaki $\hat{b}_1$ enküçük kareler tahmincisinin sadece;

$$\hat{b}_1=\bar{Y}-\hat{b}_2\bar{X}_2-b_3\bar{X}_3$$

formüllerden hesaplanacağı için ve basit regresyon modeli de buna imkân vermediğinden dolayı $b_1$ sabit tahmincisi elde edilemeyecektir.

Daha önemlisi meyilli tahminlerde $\hat{b}_1$ ve $\hat{b}_2$'ün varyanslarını tahminlemek için basit regresyon kullanılmaktır. Bunu şöyle gösterebiliriz.

$$\mathrm{Var}(\hat{b}_2) = \frac{\sigma^2 \Sigma x_3^2}{\Sigma x_2^2 \Sigma x_3^2 - \left(\Sigma x_2 x_3\right)^2}$$

olduğunu biliyoruz. $\Sigma x_2 x_3=0$ için denklem

$$\mathrm{Var}(\hat{b}_2) = \frac{\sigma^2}{\Sigma x_2^2}$$

olacaktır. Şimdi basit regresyona dayandırılan Var $(\hat{b}_2)$ nin tahmincisi:

$$S_2^2 = \frac{1}{n-2}\Sigma\left[\left(y_i-\bar{y}\right)-b_2\left(x_{i2}-\bar{x}\right)\right]^2$$

iken; $$S_{b_2}^2 = \frac{S_2^2}{\Sigma x_2^2}$$

$S_2^2$ nin beklenen değeri olarak;

$$E\left(S_{2}^{2}\right)=\frac{1}{n-2}E\Sigma\left[\left(b_2\left(x_{i2}-\bar{x}_2\right)\right)-b_3\left(x_{i3}-\bar{x}\right)+\left(\Sigma_i-\bar{\Sigma}\right)-\hat{b}_2\left(x_{i2}-\bar{x}_2\right)\right]^2$$

$$=\frac{1}{n-2}\left[\Sigma x_2^2\operatorname{var}\left(\hat{b}_2\right)+b_3^2-2\Sigma x_2^2\operatorname{var}\left(\hat{b}_2\right)+(n-1)\sigma^2\right]$$

$$=\frac{1}{n-2}\left[n-2\,\sigma^2+b_3^2\,\Sigma x_3^2\right]$$ bulunur.

Buradan;

$$E\left(S_{\hat{b}_2}^2\right)=\operatorname{Var}\left(b_2\right)+\frac{b_3^2\,\Sigma x_3^2}{\Sigma x_2^2(n-2)}$$

olur ve benzer şekilde;

$$E\left(S_{\hat{b}_3}^2\right)=\operatorname{Var}\left(b_3\right)+\frac{b_2^2\,\Sigma x_2^2}{\Sigma x_3^2(n-2)}$$

elde edilir. (Kmenta, 1984; 431)

Bu $\hat{b}_1$ ve $\hat{b}_2$ nin varyanslarının basit regresyon tahmincisinin bir başka eğilime sahip olduğu anlamına gelir Bu sonuç açıklayıcı değişkenlerle karşılıklı ilişkisiz olan herhangi bir değişken yardımıyla regresyon modellerine genelleştirilebilir.

### 3.2. Tam Çoklu Doğrusal Bağlantı Durumu

Tam Çoklu Doğrusal Bağlantıda iki açıklayıcı değişkenli çoklu regresyon modeli için çoklu doğrusallık;

$x_{i1}=a+bx_{i3}$ şeklinde formüle edilebilir.

Burada a ve b sabit sayılar olup, b≠0 dır. Bu durumda örnekteki iki açıklayıcı değişken arasında tam korelasyon vardır. Regresyon katsayılarının en küçük kareler tahmincilerini gözönüne alarak, iki açıklayıcı değişkenli model için en küçük kareler normal denklemleri;

$$\Sigma\, yx_2 = \hat{b}_2 \Sigma x_2^2 + \hat{b}_3 \Sigma x_2 x_3$$

$$\Sigma\, yx_2 = \hat{b}_2 \Sigma x_2 x_3 + \hat{b}_3 \Sigma x_2^2$$

dir. Fakat;

$$\Sigma\, yx_2 = b\Sigma xy_3$$

$$\Sigma x_2^2 = b^2 \Sigma x_3^2$$

$$\Sigma x_2 x_3 = b\Sigma x_3^2$$

olduğunu biliyoruz. Dolayısıyla en küçük kareler normal denklemleri;

$$b\Sigma yx_3 = b(\hat{b}_2 b\Sigma x_3^2 + \hat{b}_3 \Sigma x_3^2$$
$$\Sigma yx_3 = \hat{b}_2 b\Sigma x_3^2 + \hat{b}_3 \Sigma x_3^2 \quad \text{olur.}$$

Bu sonuç bize, birinci normal denklemin b ile çarpılmış ikinci norr al denkleme tamamıyla eşit olduğunu gösterir. Bu yüzden, iki denklem bağımsız değildir ve daha öncede söylediğimiz gibi $b_2$ ve $b_3$ ün çözümleri tanımsızdır.

Şimdi iki değişkenli durumda gösterilemeyen Çoklu Doğrusal Bağlantının özel bir halini anlatmak için üç açıklayıcı değişkenli bir durumu ele alalım. Bu durumda tam çoklu regresyon bağlantının varlığı ise;

$$x_{12} = a + b_3 X_{13} + b_4 X_{14}$$

anlamına gelir. Burada a, $b_3$ ve $b_4$ sabit sayılar $b_3$ ve $b_4$'ün her ikisinin de sıfırdan farklı olduklarını farzedersek, $b_3$ ve $b_4$ arasındaki korelasyon örmek katsayısı;

$$r_{23} = \frac{\Sigma x_2 x_3}{\sqrt{\Sigma x_2^2}\sqrt{\Sigma x_3^2}} = \frac{b_3 \Sigma x_3^2 + b_4 \Sigma x_3 x_4}{\sqrt{b_3^2 \Sigma x_3^2 + b_4^2 \Sigma x_4^2 + 2 b_3 b_4 \Sigma x_3 x_4}\sqrt{\Sigma x_3^2}}$$

$$= \frac{b_3\sqrt{\Sigma x_3^2} + b_4 r_{34}\sqrt{\Sigma x_4^2}}{\sqrt{\left(b_3\sqrt{\Sigma x_3^2} + b_4 r_{34}\sqrt{\Sigma x_4^2}\right)^2 + b_4^2 \Sigma x_4 x_4^2 (1 - r_{34}^2)}}$$

dir. Burada $r_{34}$, $X_3$ ve $X_4$ arasındaki örnek korelasyon katsayısıdır. Benzer şekilde $x_2$ ve $x_4$ arasındaki örnek korelasyon katsayısı;

$$r_{34} = \frac{b_3 r_{34}\sqrt{\Sigma x_3^2} + b_4\sqrt{\Sigma x_4^2}}{\sqrt{\left(b_3 r_{34}\sqrt{\Sigma x_3^2} + b_4\sqrt{\Sigma x_4^2}\right)^2 + b_3^2 \Sigma x_3^2 \left(1 - r_{34}^2\right)}}$$

dir.

Bu sonuçlar açıkça, açıklayıcı değişkenlerin toplam sayısı ikiden fazla olduğunda herhangi iki açıklayıcı değişken arasındaki korelasyonun tam olması gerektiğini göstermektedir. (Kmenta, 1984; 432) Mesela;

$X_{i2} = X_{i3} + X_{i4}$, $\Sigma X_3 X_2 = \Sigma X_4$ ve $r_{34} = -0.5$ olduğunu varsayarsak;

$$r_{23} = \frac{1 - 0.5}{\sqrt{(1 - 0.5)^2 + (1 - 0.5^2)}} = 0.5$$

$$r_{24} = \frac{-0.5 + 1}{\sqrt{(-0.5 + 1)^2 + (1 - 0.5^2)}} = 0.5$$

Bu durumda tam çoklu doğrusal bağlantı vardır ve henüz korelasyon katsayılarının hiçbirisi de kesin değerdeki yarımdan (0.5) büyük değildir. Bu nokta önemlidir, çünkü, iki açıklayıcı değişkenden daha fazla değişken olduğunda sadece korelasyon katsayılarına bakamayacağımız ve mesela tam (veya yüksek) çoklu doğrusal bağlantı olup olmadığını hesaplayamayacağımız anlamına gelir. Diğer taraftan, eğer açıklayıcı değişkenlerin herhangi bir çifti arasındaki korelasyon tamsa, o zaman örnekte olduğu gibi tam çoklu doğrusal bağlantı vardır. $X_2$ ve $X_3$ arasındaki tam korelasyon $X_2$'nin $X_3$'ün tam doğrusal bir fonksiyonu olduğunu ifade eder ve;

$$X_{i2} - a + b_3X_{i3} \quad (b_3 \neq 0)$$

şeklinde ile tanımlamaya karşılık olan yukarıdaki ifadeyi yazabiliriz. Böylece, açıklayıcı değişken sayısı ikiyi geçtiğinde iki açıklayıcı değişken arasındaki tam korelasyon yeterlidir fakat, örnekteki tam çoklu doğrusal bağlantının varlığı için gerekli şart değildir.

Şimdi üç açıklayıcı değişken ihtiva eden regresyon modelinde tam çoklu doğrusal bağlantı şartları altında regresyon katsayılarının en küçük kareler tahmincilerinin ne olacağını görelim.

En küçuk kareler normal denklemleri

$$\Sigma yx_2 = \hat{b}_2\Sigma x_2^2 + \hat{b}_3\Sigma x_2x_3 + \hat{b}_4\Sigma x_2x_4$$
$$\Sigma yx_3 = \hat{b}_2\Sigma x_2x_3 + \hat{b}_3\Sigma x_3^2 + \hat{b}_4\Sigma x_3x_4$$
$$\Sigma yx_4 = \hat{b}_2\Sigma x_2x_4 + \hat{b}_3\Sigma x_3x_4 + \hat{b}_4\Sigma x_4^2$$

dir. Fakat yukarıdaki formülden,

$$\Sigma yx_2 = b_3\Sigma yx_3 + b_4\Sigma yx_4$$
$$\Sigma x_2^2 = b_3^2\Sigma x_3^2 + b_4^2\Sigma x_4^2 + 2b_3b_4\Sigma x_3x_4$$
$$\Sigma x_2x_3 = b_3\Sigma x_3 + b_4\Sigma x_3x_4$$
$$\Sigma x_2x_4 = b_3\Sigma x_3x_4 + b_4\Sigma x_4^2$$

elde ederiz. En küçük kareler normal denklemindeki bu ifadeler yerine koymakla;

$$b_3\Sigma yx_3 + b_4\Sigma yx_4 = \hat{b}_2(b_3^2\Sigma x_3^2 + b_4^2\Sigma x_4^2 + 2b_3b_4\Sigma x_3x_4) +$$
$$\hat{b}_3(b_3\Sigma x_3^2 + b_4\Sigma x_3x_4) + \hat{b}_4(b_3\Sigma x_3x_4 + b_4\Sigma x_4^2)$$
$$\Sigma yx_3 = \hat{b}_2(b_3\Sigma x_3^2 + b_4\Sigma x_3x_4) + \hat{b}_3\Sigma x_3^2 + \hat{b}_4\Sigma x_3x_4$$
$$\Sigma yx_4 = \hat{b}_2(b_3\Sigma x_3x_4 + b_4\Sigma x_4^2) + \hat{b}_3\Sigma x_3x_4 + \hat{b}_4\Sigma x_4^2$$

ifadeleri elde edilir. Böylece birinci normal denklemin, sadece $b_3$ ile çarpılan ikinci normal denkleme ve $b_4$ ile çarpılan normal denkleme eşitliği ölçülür. Bundan dolayı, tam çoklu doğrusal bağlantı altında üç normal denklemde bağımsız değildir ve $b_2$, $b_3$ ve $b_4$ için çözümlenemez. Bu sonuç farklı sayıda açıklayıcı değişkene sahip regresyon modellerine tesmil edilebilir. Regresyon katsayılarının en küçük kareler tahmincileri vektörü,

$$b = (X'X^{-1})\ (X'y)$$

dir. Açıklayıcı değişkenler arasında tam doğrusal bir ilişkinin varlığı, (X'X) kolonlarından birinin başka bir veya daha fazla kolonun tam bir fonksiyonu olduğu anlamına geler. Böylece (X'X) bir tekil matristir ve tersi yoktur. Regresyon katsayılarının en küçük kareler tahmincilerinin tanımsız olduğu verildiğinde, geleneksel yaklaşım, eldeki örneklerden farklı kaynaklardan regresyon katsayıları hakkındaki bilgileri kullanmak olmuştur. Örneğin;

$$Y_i = b_1 + b_2X_{i2} + b_3X_{i3}\varepsilon_j$$

modelinde iki açıklayıcı değişken tamamıyla ilişkilidir. Fakat bildiğimiz gibi, $b3/b2 = k$ olmak üzere k bilinen bir sabit sayı olduğunda, regresyon katsayılarının tahmini mümkün olmaktadır. Regresyon modelinde $b3 = kb2$ yerine koyarsak,

$$Y_i = b_1 + b_2X_{i2} + b_3X_{i3}e_i \text{ veya } Y = b_1 + b_2Z_i + E_i$$

elde ederiz. Burada $Zi = (X_{i2} + kX_{i3})$ ile ölçülür. Bu durumda $b_1$ ve $b_2$'nin en küçük kareler tahmincisini elde edebiliriz ve $b_3$'ün tahminini çıkartabiliriz. Bu tahminciler verilen k değeri üstünden şarta bağlı olarak tahmin şeklinde mutea-la edilebilir. Eğer adaylara benzer şekilde görünen k'nın birkaç değeri varsa, k'nın farklı değerinin üstündeki herbir şarta bağlı olarak bir farklı tahminciler kümesi elde ederiz. (Kmenta 1984; 435) Oranını bildiğimiz şeklinde veya hakkında bazı kesin bilgilerimizin olduğu iki katsayılı regresyon durumlarına nisbeten az rastlanır. Daha sık olarak, farklı bir örnekteki regresyon katsayılarının bir veya daha fazla tahminine sahip olabiliriz.

Çoklu doğrusal bağlantı problemiyle uğraşmanın gayesi, eldeki örneklerin haricindeki kaynaklardan bilgi toplayarak herhangi bir durumda ne yapılması gerektiğini aramaktır. Normalde regresyon modelini ve tahmin prosedürünü tayin etmek için popülasyon hakkında mevcut olan bütün bilgileri kullandığımızı düşünelim: Araştırmanın maliyetinin çok yüksek olduğu durumlar hariç olmak üzere ilgili bilgilerin hepsini araştırmadan önce, çoklu doğrusal bağlantının varlığını beklemek bizim için bir sebep olamaz. Tahmin aşamasına gelindiğinde bütün bilgi kaynakları kullanılmış olacağından, çoklu doğrusal bağlantı için gerekli düzeltme yapılamayacaktır.

### 3.3. Yüksek Dereceli Çoklu Doğrusal Bağlantı

Buraya kadar çoklu doğrusal bağlantının iki uç durumunu analiz etmiş olduk. Bunlar, çoklu doğrusal bağlantının olmadığı ve çoklu doğrusal bağlantının tam olduğu durumlardı.

Her iki durumda pratik uygulamalarda pek bulunmaz, fakat çoğu veriler tam olmasa da biraz çoklu doğrusal bağlantı gösterirler. Bu durumda "yuvarlama hatası" olmaksızın regresyon katsayılarının en küçük kareler tahmincileri ile belirlenen çözümü elde edilir.

Şimdi çoklu doğrusal bağlantının derecesi ile regresyon katsayılarının en küçük kareler tahmincilerinin mahiyeti arasındaki ilişkiyi test edelim. Klasik Doğrusal Regresyon Modelinin temel varsayımları altında regresyon katsayılarının en küçük kareler tahmincileri bütün arzu edilen özelliklere sahiptir. Fakat en küçük kareler tahmincilerinin bu özelliklere sahip olduğunu bilmek, bizim için, eğer onların varyansları sonuç tahminlerinin oldukça güvenilmez olduğu şeklinde ise sadece bir rahatlıktır. Yani, eğer aynı zamanda, bu varyans çok büyükse, tahminlerimizin mümkün olan en küçük varyansa sahip olduklarını bilmek çok faydalı değildir. İki açıklayıcı değişkenli bir regresyon modelini gözönüne alalım. $b_2$ ve $b_3$ 'ün varyansları;

$$\mathrm{Var}(\widehat{b}_2) = \frac{\sigma^2 \Sigma x_3^2}{\Sigma x_2^2 \Sigma x_3^2 - (\Sigma x_2 x_3)^2} = \frac{\sigma^2}{\Sigma x_2^2 (1 - r_{23}^2)}$$

$$\mathrm{Var}(\widehat{b}_3) = \frac{\sigma^2 \Sigma x_2^2}{\Sigma x_2^2 \Sigma x_3^2 - (\Sigma x_2 x_3)^2} = \frac{\sigma^2}{\Sigma x_3^2 (1 - r_{23}^2)}$$

dir ve kovaryansları;

$$\mathrm{Cov}(\hat{b}_2\hat{b}_3) = \frac{-\sigma^2 \Sigma x_2 x_3}{\Sigma x_2^2 \Sigma x_3^2 - (\Sigma x_2 x_3)^2} = \frac{-\sigma^2 r_{23}}{\sqrt{\Sigma x_2^2}\sqrt{\Sigma x_3^2 (1 - r_{23}^2)}}$$ dir.

Bu açıkça göstermektedir ki, $r_{23}$ 1'e yakın olduğunda, $b_2$ ve $b_3$'nin varyans ve kovaryansı oldukça büyür. (daha öncede bahsettiğimiz gibi $r_{23} = 1$ olduğunda sonsuz olacaklardır). İki açıklayıcı değişken durumunda $r_{23}$ değeri çoklu doğrusal bağlantının derecesini ölçtüğü için, yukarıdaki ifade, çoklu doğrusal bağlantının derecesi ne kadar yüksekse, $\hat{b}_2$ ve $\hat{b}_3$'nin varyans ve kovaryansı o kadar büyük olacağına işaret eder. İkiden fazla açıklayıcı değişken olduğunda en küçük kareler katsayıları için yukarıdaki formüller;

$$\mathrm{Var}(\hat{b}_k) = \frac{\sigma^2}{\Sigma x_k^2 (1 - R_k^2)} = \frac{\sigma^2}{SS_k} \qquad (k = 2,3,...,K)$$

şeklinde genelleştirilebilir. (Kmenta, 1984; 437) Burada $R_k^2$ bağımlı değişken olarak k'ıncı açıklayıcı değişkenler enküçük kareler regresyonundaki determinat katsayısı ve regresör olarak mevcut kalan açıklayıcı değişkenlerin hepsidir. $SSE_k$'nın hata kareler toplamını gösterdiğini biliyoruz. $1/(1 - R_k)$ miktari ise literatürde "Varyans Enflasyon Faktörü (VIF)" olarak adlandırılır.

## 4 . SONUÇ

Sonuç olarak yüksek dereceli bir çoklu doğrusal bağlantının varlığının, regresyon katsayısının tahmininin tamamiyle kesin olmaması gibi arz edilmeyen bir sonuç göstermesi nedeniyle zararlı olduğu sonucu çıkarılabilir. En küçük kareler tahmincilerinin büyük varyanslarından dolayı bazı beklenmeyen durumlar ortaya çıkabilir. Bununla beraber ya açıklayıcı değişkenlerin küçük bir dağılıma sahip olduğu, yada $\sigma^2$ nin kendi kendine büyük olduğu için, suçu çoklu doğrusal bağlantıya yüklemek istiyorsak, onun derecesi ölçülmelidir. Bu maksat için iki açıklayıcı değişkenli modellerde $r_3$ 'nin değerini kullanabiliriz fakat, ikiden fazla açıklayıcı değişken olduğunda çoklu doğrusal bağlantının derecesini ölçmek daha karmaşık bir hal alır. Bu daha öncede açıkladığımız gibi, yüksek dereceli çoklu doğrusal bağlantının, hatta tam çoklu doğrusal bağlantının varlığı genellikle herhangi iki değişken arasındaki korelasyonun özellikle yüksek olması gerektiği şeklindeki bir düşünceden kaynaklanmaktadır.


## SUMMARY

The main aim of study is to research multicallinearity Condions and to discuss parameters and their coefficient relability.

After, brief ii roducing of results of parameters are obtained. Later, full multicallinearity position is discussed and before, searched on a model consist of two explanatory variables, and then undefined solutions of $b_2$ and $b_3$ estimates are found After those founding, in the case of if a model has explanatory vairable more than two, the correlation coefficients are unsufficient to find out multicallinearitiy.

At the last part of study, high level multicallinearity situations circumtanced and later, because of correlations coefficient between any two explanatory vairables close tovalue of one, variance and covariance of coefficient estimates are obtained mostly large.


## KAYNAKÇA


JOHN Wiley and Sons, **The Theory and Practice of Econometrics,** Second Edition.

KOUTSOYİANNİS A.; **Ekonometri Kuramı, Ekonometri Yöntemlerinin Tanıtımına Giriş,** Ekim - 1989

KMENTA, Jan, **Elementi of Economettics,** Second Edition, Macmillan Publishing Company, New York - 1984

AKKAYA, Şahin, **Ekonometri I,** İzmir 1989

AKALIN Sedat, **Elementer Ekonometri,** E. Ü. Basımevi, İzmir, 1971

AKKAYA, Şahin ve HASGÜR, İbrahim İlhan, **Uygulamalı İstatistik,** Aklıselim Matbaası, İzmir.

DALTA, M., **Ekonometric Methods,** South-Western Publishing Co.

KILIÇBAY, Ahmet, **Ekonometrik Metodlar ve Araştırma,** İstanbul, 1975.

AVRALIOLU, Z., **İstatistik,** 1971.

BALDEMİR, Ercan.; **Çoklu Regresyon Modellerinde Çoklu Doğrusal Baağlantı ve Enflasyon Üzerine Türkiye Uygulaması,** D.E.Ü.S.B.E.Yüksek Lisans Tezi, İzmir, 1989.