6. DOĞRUSAL REGRESYON MODELİNE MATRİS YAKLAŞIMI

$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + u_i$ $(i = 1, 2, \dots, n)$ gibi çok açıklayıcı değişkene sahip bir model, aşağıdaki gibi bir eşanlı denklem modelini göstermektedir.

$$
\begin{aligned}
Y_1 &= \beta_0 + \beta_1 X_{11} + \beta_2 X_{21} + \dots + \beta_k X_{k1} + u_1 \\
Y_2 &= \beta_0 + \beta_1 X_{12i} + \beta_2 X_{22} + \dots + \beta_k X_{k2} + u_2 \\
&\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\
Y_n &= \beta_0 + \beta_1 X_{1n} + \beta_2 X_{2n} + \dots + \beta_k X_{kn} + u_n
\end{aligned}
\tag{6.1}
$$

Bu modelin matrislerle ifadesi ise aşağıdaki gibidir.

$$
\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & X_{11} & X_{21} & \dots & X_{k1} \\ 1 & X_{12} & X_{22} & & X_{k2} \\ & \vdots & & \ddots & \vdots \\ 1 & X_{1n} & X_{2n} & \cdots & X_{kn} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \text{veya}
$$

Y = Xβ + u

Burada

$\mathbf{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}$ nx1 boyutlu bağımlı değişken gözlemleri vektörü,

$\mathbf{X} = \begin{bmatrix} 1 & X_{11} & X_{21} & \dots & X_{k1} \\ 1 & X_{12} & X_{22} & & X_{k2} \\ & \vdots & & \ddots & \vdots \\ 1 & X_{1n} & X_{2n} & \cdots & X_{kn} \end{bmatrix}$ nxk boyutlu açıklayıcı değişken verileri matrisi,

$\boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix}$ kx1 boyutlu katsayılar vektörü ve $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$ nx1 boyutlu hata terimleri vektörüdür.

6.1 Klasik Doğrusal Regresyon Modelinin Varsayımlarının Matrisle Gösterilmesi

1. Hata terimlerinin beklenen değeri sıfırdır ($E(u_i) = 0$) varsayımın matrisle gösterimi $E(\mathbf{u}) = \mathbf{0}$ şeklindedir.

$$E\begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} = \begin{bmatrix} E(u_1) \\ E(u_2) \\ \vdots \\ E(u_n) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$ anlamına gelmektedir.

2. Hata terimlerinin varyansı sabittir ve aralarındaki kovaryans sıfırdır ($E(u_iu_j) = 0$, $i \neq j$ ve $E(u_iu_j) = \sigma^2$, $i=j$ iken) varsayımın matrisle gösterimi $E(\mathbf{uu}') = \sigma^2\mathbf{I}$ dir.

$$E(\mathbf{uu}') = E\begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}\begin{bmatrix} u_1 & u_2 & \dots & u_n \end{bmatrix} = E\begin{bmatrix} u_1^2 & u_1u_2 & \dots & u_1u_n \\ u_2u_1 & u_2^2 & \dots & u_2u_n \\ \vdots & \vdots & \ddots & \vdots \\ u_nu_1 & u_nu_2 & \dots & u_n^2 \end{bmatrix}$$

$$= \begin{bmatrix} E(u_1^2) & E(u_1u_2) & \dots & E(u_1u_n) \\ E(u_2u_1) & E(u_2^2) & \dots & E(u_2u_n) \\ \vdots & \vdots & \ddots & \vdots \\ E(u_nu_1) & E(u_nu_2) & \dots & E(u_n^2) \end{bmatrix}$$

$E(u_i) = 0$ olduğundan $Var(u_i) = E\{u_i - E(u_i)\}^2 = E\{u_i\}^2$ ve

$Cov(u_iu_i) = E\{[u_i - E(u_i)][u_j - E(u_j)]\} = E\{u_iu_j)$. Dolayısıyla,

$$E(\mathbf{uu}') = \begin{bmatrix} Var(u_1^2) & Cov(u_1u_2) & \dots & Cov(u_1u_n) \\ Cov(u_2u_1) & Var(u_2^2) & \dots & Cov(u_2u_n) \\ \vdots & \vdots & \ddots & \vdots \\ Cov(u_nu_1) & Cov(u_nu_2) & \dots & Var(u_n^2) \end{bmatrix}$$

$$= \begin{bmatrix} \sigma^2 & 0 & \dots & 0 \\ 0 & \sigma^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma^2 \end{bmatrix} = \sigma^2\begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix} = \sigma^2\mathbf{I}$$ olmaktadır.

Bu matris hata terimlerinin varyans-kovaryans matrisidir. Ana köşegen üzerindeki öğeler varyansı, diğer öğeler kovaryansı gösterir. Simetrik matristir.

$E(\mathbf{uu}') = \sigma^2\mathbf{I}$ varsayımı iki alt varsayım içerir.

2.i Sabit varyans (homoscedasticity) veya değişmeyen varyans varsayımı: Hata terimlerinin varyansı sabittir ve aynıdır. $Var(u_i) = \sigma^2$ Bu varsayım sağlanmadığında *değişen varyans* (heteroscedasticity) *sorunu* vardır.

2.ii Ardışık bağımlılık (autocorrelation) yok varsayımı: Hata terimleri arasındaki kovaryans sıfırdır. $Cov(u_iu_j) = 0$ $i,j = 1, \dots, n$ $i \neq j$ Bu varsayım sağlanmadığında *ardışık bağımlılık sorunu* ortaya çıkar.

2i varsayımı ile Y_i'lerin varyansları da belirlenmiş olmaktadır. Vektör olarak ifade edersek

$$Var(\mathbf{Y}) = E\{[\mathbf{Y} - E(\mathbf{Y})]'[\mathbf{Y} - E(\mathbf{Y})]\} = E\{[\mathbf{Y} - \mathbf{X\beta})]'[\mathbf{Y} - \mathbf{X\beta}]\}$$

$$= E\{[\mathbf{X\beta} + \mathbf{u} - \mathbf{X\beta})]'[\mathbf{X\beta} + \mathbf{u} - \mathbf{X\beta}]\} = E\{\mathbf{u}'\mathbf{u}\} = \sigma_u^2 I$$

Bu $Var(Y_i) = \sigma_u^2$ anlamına gelmektedir.

3. X matrisindeki açıklayıcı değişkenler rassal değişkenler olmayıp her örneklemde aynı değeri alan değişkenlerdir. Bu durumda X değişkenleri ile hata terimleri arasında bir ilişki de olamaz. Açıklayıcı değişkenlerin rassal değişkenler olması kabul edilirse, bu durumda X'ler ile u'nun ilişkisiz olduğunu açıkça varsaymak gerekir. Her iki durumda da varsayım şu şekildedir: $Cov(X_{ij}u_i) = 0$ veya matris ifadesiyle $E(\mathbf{X}'\mathbf{u}) = 0$

4. X matrisinin aşaması (rank) denklemde tahmin edilen katsayı adedi k'ya eşittir ve k da veri sayısı n'den küçüktür: $rank(\mathbf{X}) = k < n$

$rank(\mathbf{X}) = k$ olması, X verilerini temsil eden vektörlerin birbirinden bağımsız olması anlamına gelmektedir. Bu varsayım sağlanmadığında *tam çoklu doğrusallık* (perfect multicollinearity) sorunu var demektir.

$k < n$ sınırlaması, *serbestlik derecesi* (degrees of freedom) üzerine bir sınırlama getirmektedir. Tahmin sonuçlarının serbestlik derecesi bakımından güvenilir olabilmesi için SD en az 12 olmalıdır.

5. Hata terimleri normal dağılıma sahiptir. Birinci ve ikinci varsayımların geçerli olduğu kabul edilirse, varsayım şu şekildedir: $\mathbf{u} \sim N(0, \sigma^2\mathbf{I})$

Bu varsayım t ve F-testi gibi hipotez testlerinin uygulanmasında gereklidir.

6.2 EKK tahmini

Genel Doğrusal Model

$Y_i = \beta_1 + \beta_2 X_{i2} + \beta_3 X_{i3} + \ldots + \beta_k X_{ik} + u_i$, $i = 1 \ldots n$ veya $\mathbf{Y = X\beta + u}$ modelinde EKK yöntemi $\hat{\mathbf{u}}'\hat{\mathbf{u}} = [\hat{u}_1 \quad \hat{u}_2 \quad \ldots \quad \hat{u}_n]\begin{bmatrix}\hat{u}_1\\ \hat{u}_2\\ \vdots\\ \hat{u}_n\end{bmatrix} = \hat{u}_1^2 + \hat{u}_2^2 + \cdots + \hat{u}_n^2 = \sum_{i=1}^{n}\hat{u}_i^2$ yi minimize eder.

$\mathbf{Y = X\beta + u}$ ilişkisinden $\hat{\mathbf{u}} = \mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}$, $\hat{\mathbf{u}}' = \left(\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}\right)'$ ve böylece

$\hat{\mathbf{u}}'\hat{\mathbf{u}} = \left(\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}\right)'\left(\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}\right) = \mathbf{Y}'\mathbf{Y} - \mathbf{2}\hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{Y} + \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}}$ bulunur**.**

Minimizasyon için bu fonksiyonun $\hat{\beta}$ ya göre türevi sıfıra eşitlenmelidir:

$$\frac{\partial \hat{\mathbf{u}}'\hat{\mathbf{u}}}{\partial \hat{\boldsymbol{\beta}}} = -2\mathbf{X}'\mathbf{y} + 2\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{0}$$

ve bunun sonucu

$\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{X}'\mathbf{Y}$ bulunur.

Son işlemlerin yapılmasında matris cebiri kullanılmıştır. Buradan $\hat{\boldsymbol{\beta}}$ yı bulmak için eşitliğin iki tarafını da **X'X** matrisinin tersiyle çarpmalıyız. Böylece $(\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'\mathbf{X})\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$ ve

$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$

bulunur. Bu EKK tahmin edicisidir.

6.3 $\hat{\boldsymbol{\beta}}$ nın Varyans-Kovaryans Matrisi ve σ_u^2'nun tahmini

$$\hat{\boldsymbol{\beta}}=(\mathbf{X'X})^{-1}\mathbf{X'Y} = (\mathbf{X'X})^{-1}\mathbf{X'}(\mathbf{X}\boldsymbol{\beta} + \mathbf{u}) = (\mathbf{X'X})^{-1}\mathbf{X'X}\boldsymbol{\beta} + (\mathbf{X'X})^{-1}\mathbf{X'u}) = \boldsymbol{\beta} + (\mathbf{X'X})^{-1}\mathbf{X'u} \qquad (6.1)$$

ve böylece

$$E(\hat{\boldsymbol{\beta}}) = E\{\boldsymbol{\beta} + (\mathbf{X'X})^{-1}\mathbf{X'u})\} = E(\boldsymbol{\beta}) + (\mathbf{X'X})^{-1}E(\mathbf{X'u}) = \boldsymbol{\beta} \text{ dır.} \qquad (6.2)$$

Burada üçüncü varsayımdan yararlanılmış ve açıklayıcı değişkenlerin rassal olmayıp, her örneklemde aynı değeri aldığı varsayımıyla X’ler beklenen değerin dışına çıkartılmıştır. Buradan katsayı tahminlerinin varyans-kovaryans tahminleri aşağıdaki gibi bulunabilir.

$$\text{Var,Cov }(\hat{\boldsymbol{\beta}}) = E\{[\hat{\boldsymbol{\beta}}-E(\hat{\boldsymbol{\beta}})][\hat{\boldsymbol{\beta}}-E(\hat{\boldsymbol{\beta}})]'\}$$

$$= E\{[(\mathbf{X'X})^{-1}\mathbf{X'u}][(\mathbf{X'X})^{-1}\mathbf{X'u}]'\} \text{ (çünkü 6.1 ve 6.2'den } \hat{\boldsymbol{\beta}}-E(\hat{\boldsymbol{\beta}}) = (\mathbf{X'X})^{-1}\mathbf{X'u})$$

$$= E\{(\mathbf{X'X})^{-1}\mathbf{X'uu'X}(\mathbf{X'X})^{-1}\} = (\mathbf{X'X})^{-1}\mathbf{X'}\,E(\mathbf{uu'})\mathbf{X}(\mathbf{X'X})^{-1}$$

$$= (\mathbf{X'X})^{-1}\mathbf{X'}\sigma_u^2\mathbf{X}(\mathbf{X'X})^{-1} \qquad \text{(çünkü } E(\mathbf{uu'}) = \sigma_u^2\mathbf{I})$$

$$= \sigma_u^2(\mathbf{X'X})^{-1}\mathbf{X'X}(\mathbf{X'X})^{-1}$$

$$= \sigma_u^2(\mathbf{X'X})^{-1} \qquad \text{(çünkü } \mathbf{X'X}(\mathbf{X'X})^{-1} = \mathbf{I})$$

$$(\mathbf{X'X})^{-1} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k1} & \cdots & a_{kk} \end{bmatrix} \text{ böylece Var, Cov}(\hat{\boldsymbol{\beta}}) = \begin{bmatrix} \sigma_u^2 a_{11} & \sigma_u^2 a_{12} & \cdots & \sigma_u^2 a_{1k} \\ \sigma_u^2 a_{21} & \sigma_u^2 a_{22} & \cdots & \sigma_u^2 a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_u^2 a_{k1} & \sigma_u^2 a_{k1} & \cdots & \sigma_u^2 a_{kk} \end{bmatrix}$$

Bu matrisin asal köşegenindeki öğeler $\hat{\beta}_j$ katsayılarının varyanslarıdır.

$\text{Var}(\hat{\beta}_1) = \sigma_u^2 a_{11}$, $\text{Var}(\hat{\beta}_2) = \sigma_u^2 a_{22}$, … $\text{Var}(\hat{\beta}_k) = \sigma_u^2 a_{kk}$,

Diğer öğeler $\hat{\beta}_j$ katsayıları arasındaki kovaryansları gösterir. $(\mathbf{X'X})^{-1}$ simetrik olduğundan karşılık gelen kovaryanslar eşittir.

$\text{Cov}(\hat{\beta}_1\, \hat{\beta}_2) = \sigma_u^2 a_{12} = \sigma_u^2 a_{21}$ … gibi.

Varyans ve kovaryansların örnek verilerinden tahmin edilebilmesi için σ_u^2'nun tahmin edilmesi gerekir.

Hata terimi için örnekleme varyansı

$$\mathbf{V}\text{ar}(\hat{u}_i) = \hat{\sigma}^2 = \frac{\Sigma \hat{u}_i^{\,2}}{\text{n-k}} = \frac{\mathbf{\hat{u}'\hat{u}}}{\text{n-k}}$$

Eğer hata terimleri tek tek bulunmak istenmiyorsa $\mathbf{\hat{u}'\hat{u}}$ toplamı şu şekilde bulunur:

$$\mathbf{\hat{u}'\hat{u}} = (\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}})'(\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}) = \mathbf{Y'Y} - \mathbf{Y'X}\hat{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}'\mathbf{X'Y} + \hat{\boldsymbol{\beta}}'\mathbf{X'X}\hat{\boldsymbol{\beta}}$$

$(\hat{\boldsymbol{\beta}} = (\mathbf{X'X})^{-1}\mathbf{X'Y}$ olduğundan $\hat{\boldsymbol{\beta}}' = \mathbf{Y'X}(\mathbf{X'X})^{-1})$

$$= \mathbf{Y'Y} - \mathbf{Y'X}(\mathbf{X'X})^{-1}\mathbf{X'Y} - \mathbf{Y'X}(\mathbf{X'X})^{-1}\mathbf{X'Y} + \mathbf{Y'X}(\mathbf{X'X})^{-1}\mathbf{X'X}(\mathbf{X'X})^{-1}\mathbf{X'Y}$$

$$= \mathbf{Y'Y} - \mathbf{Y'X}(\mathbf{X'X})^{-1}\mathbf{X'Y} - \mathbf{Y'X}(\mathbf{X'X})^{-1}\mathbf{X'Y} + \mathbf{Y'X}(\mathbf{X'X})^{-1}\mathbf{X'Y}$$

$$= \mathbf{Y'Y} - \mathbf{Y'X}(\mathbf{X'X})^{-1}\mathbf{X'Y}$$

$$= \mathbf{Y'Y} - \hat{\boldsymbol{\beta}}'\mathbf{X'Y}$$

Demek ki

$$\hat{\sigma}_u^2 = \frac{\mathbf{Y'Y} - \hat{\boldsymbol{\beta}}\mathbf{X'Y}}{\text{n - k}}$$

6.4 Belirlilik Katsayısının ve F İstatistiğinin Matrislerle Gösterimi

Belirlilik katsayısı matrisler kullanılarak

$$R^2 = \frac{\hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{Y} - n\overline{Y}^2}{\mathbf{Y}'\mathbf{Y} - n\overline{Y}^2}$$

formülüyle bulunabilir. Benzer bir şekilde F istatistiğinin formülü aşağıdaki gibidir.

$$F = \frac{R^2/(k-1)}{(1-R^2)/(n-k)}$$

$$F = \frac{(\hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{Y} - n\overline{Y}^2)/(k-1)}{(\mathbf{Y}'\mathbf{Y} - (\hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{Y})/(n-k)}$$