Matrislerle Basit Doğrusal Regresyon Modeli

Verilen veri setini matrislerle çözerek katsayıları bulalım. Önceki örneğimizdeki veriler şu şekildedir:

X	Y
1	2
2	4
3	5
4	4
5	5

Regresyon modelimiz:

$$Y = \beta_0 + \beta_1 X$$

1. Adım: Matrislerin Oluşturulması

Regresyon modelimizi matris formunda yazalım:

$$Y = X\beta + \epsilon$$

- **Bağımlı değişken matris (Y):**

$$Y = \begin{bmatrix} 2 \\ 4 \\ 5 \\ 4 \\ 5 \end{bmatrix}$$

- **Bağımsız değişken matris (X):** Sabit terimi de eklemeliyiz.

$$X = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \\ 1 & 5 \end{bmatrix}$$

- **Katsayılar matris (β):**

$$\beta = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}$$

2. Adım: Normal Denklemlerin Çözümü

Regresyon katsayılarını bulmak için formül:

$$\beta = (X^T X)^{-1} X^T Y$$

a) $X^T X$ Hesaplama

Öncelikle X^T (X'in transpozu) matrisini bulalım:

$$X^T = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 \end{bmatrix}$$

Şimdi $X^T X$ çarpımını bulalım:

$$X^T X = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \\ 1 & 5 \end{bmatrix}$$

$$= \begin{bmatrix} 5 & 15 \\ 15 & 55 \end{bmatrix}$$

b) $(X^T X)^{-1}$ Hesaplama

Şimdi $X^T X$'in tersini bulalım:

$$(X^T X)^{-1} = \frac{1}{(5)(55) - (15)(15)} \begin{bmatrix} 55 & -15 \\ -15 & 5 \end{bmatrix}$$

$$= \frac{1}{275 - 225} \begin{bmatrix} 55 & -15 \\ -15 & 5 \end{bmatrix}$$

$$= \frac{1}{50} \begin{bmatrix} 55 & -15 \\ -15 & 5 \end{bmatrix}$$

$$= \begin{bmatrix} 1.1 & -0.3 \\ -0.3 & 0.1 \end{bmatrix}$$

c) $X^T Y$ Hesaplama

$$X^T Y = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 \end{bmatrix} \begin{bmatrix} 2 \\ 4 \\ 5 \\ 4 \\ 5 \end{bmatrix}$$

$$= \begin{bmatrix} 20 \\ 70 \end{bmatrix}$$

d) $\beta = (X^T X)^{-1} X^T Y$ Hesaplama

$$\beta = \begin{bmatrix} 1.1 & -0.3 \\ -0.3 & 0.1 \end{bmatrix} \begin{bmatrix} 20 \\ 70 \end{bmatrix}$$

Çarpımı gerçekleştirelim:

$$\beta_0 = 1.1 \times 20 + (-0.3) \times 70 = 2.2$$

$$\beta_1 = -0.3 \times 20 + 0.1 \times 70 = 0.6$$

Sonuç

- Kesme noktası (intercept): $\beta_0 = 2.2$
- Eğim (slope): $\beta_1 = 0.6$

Regresyon denklemi:

$$\hat{Y} = 2.2 + 0.6X$$

Matrislerle yapılan hesaplamaların sonucunda aynı katsayıları elde ettik ve doğrusal regresyon modelimiz tamamlandı.

Standart hatayı matrisler kullanarak hesaplayacağız. Aşağıdaki adımları takip ederek işlemi gerçekleştirelim:

1. Adım: Tahmin Edilen $\hat{Y}$ ve Artıklar (e) Matrisi

Hatırlayın ki tahmin edilen $\hat{Y}$ değeri,

$$\hat{Y} = X\beta$$

Öncelikle X, β, ve Y matrislerini tekrar yazalım:

- **Bağımsız değişken matrisi (X):**

$$X = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \\ 1 & 5 \end{bmatrix}$$

- **Katsayılar matrisi (β):**

$$\beta = \begin{bmatrix} 2.2 \\ 0.6 \end{bmatrix}$$

- **Bağımlı değişken matrisi (Y):**

$$Y = \begin{bmatrix} 2 \\ 4 \\ 5 \\ 4 \\ 5 \end{bmatrix}$$

Tahmin edilen $\hat{Y}$ matrisini bulalım:

$$\hat{Y} = X\beta = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \\ 1 & 5 \end{bmatrix} \begin{bmatrix} 2.2 \\ 0.6 \end{bmatrix} = \begin{bmatrix} 2.8 \\ 3.4 \\ 4.0 \\ 4.6 \\ 5.2 \end{bmatrix}$$

Artıklar matrisi (e):

$$e = Y - \hat{Y} = \begin{bmatrix} 2 \\ 4 \\ 5 \\ 4 \\ 5 \end{bmatrix} - \begin{bmatrix} 2.8 \\ 3.4 \\ 4.0 \\ 4.6 \\ 5.2 \end{bmatrix} = \begin{bmatrix} -0.8 \\ 0.6 \\ 1.0 \\ -0.6 \\ -0.2 \end{bmatrix}$$

2. Adım: Hata Kareler Toplamı (SSE) ve Hata Varyansı

Hata kareler toplamı (SSE) matris formunda şu şekildedir:

$$SSE = e^T e$$

Burada, e^T artıkların transpozudur:

$$e^T = \begin{bmatrix} -0.8 & 0.6 & 1.0 & -0.6 & -0.2 \end{bmatrix}$$

Çarpımı gerçekleştirelim:

$$SSE = \begin{bmatrix} -0.8 & 0.6 & 1.0 & -0.6 & -0.2 \end{bmatrix} \begin{bmatrix} -0.8 \\ 0.6 \\ 1.0 \\ -0.6 \\ -0.2 \end{bmatrix}$$

$$= (-0.8)^2 + (0.6)^2 + (1.0)^2 + (-0.6)^2 + (-0.2)^2$$

$$= 0.64 + 0.36 + 1.0 + 0.36 + 0.04 = 2.4$$

Hata varyansı (s^2):

$$s^2 = \frac{SSE}{n-k} = \frac{2.4}{5-2} = \frac{2.4}{3} = 0.8$$

Burada $n = 5$ gözlem sayısı ve $k = 2$ (sabit terim ve bir bağımsız değişken olduğu için).

3. Adım: Standart Hataların Hesaplanması

Standart hata matrisi şu formülle hesaplanır:

$$SE(\beta) = \sqrt{s^2(X^TX)^{-1}}$$

Daha önce $(X^TX)^{-1}$'i hesaplamıştık:

$$(X^TX)^{-1} = \begin{bmatrix} 1.1 & -0.3 \\ -0.3 & 0.1 \end{bmatrix}$$

Çarpalım:

$$s^2(X^TX)^{-1} = 0.8 \times \begin{bmatrix} 1.1 & -0.3 \\ -0.3 & 0.1 \end{bmatrix} = \begin{bmatrix} 0.88 & -0.24 \\ -0.24 & 0.08 \end{bmatrix}$$

Bu matrisin köklerini alarak standart hataları bulalım:

$$SE(\beta_0) = \sqrt{0.88} \approx 0.938$$

$$SE(\beta_1) = \sqrt{0.08} \approx 0.283$$

Sonuçlar

- **Kesme noktası (β_0) standart hatası:** 0.938
- **Eğim (β_1) standart hatası:** 0.283

Bu şekilde, matrislerle standart hataları hesaplamış olduk.

İki Bağımsız Değişkenli Bir Regresyon Modeli: Klasik ve Matris Çözümü

Veri setimiz, bağımlı değişken (Y) ve iki bağımsız değişken (X_1 ve X_2) içermektedir. Amaç, Y'nin X_1 ve X_2 ile olan ilişkisini modellemektir.

Veri Seti

X_1	X_2	Y
1	2	5
2	1	6
3	4	9
4	3	8
5	5	12

Regresyon modeli şu şekildedir:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \epsilon$$

1. Klasik Çözüm

1.1 Ortalama Değerlerin Hesaplanması

Öncelikle ortalama değerleri hesaplayalım:

$$\bar{X}_1 = \frac{1+2+3+4+5}{5} = 3$$

$$\bar{X}_2 = \frac{2+1+4+3+5}{5} = 3$$

$$\bar{Y} = \frac{5+6+9+8+12}{5} = 8$$

1.2 Eğim Katsayılarının Hesaplanması (β_1 ve β_2)

$$\beta_1 = \frac{\sum(X_1 - \bar{X}_1)(Y - \bar{Y}) - \sum(X_2 - \bar{X}_2)(X_2 - \bar{X}_2)}{\sum(X_1 - \bar{X}_1)^2}$$

$$\beta_2 = \frac{\sum(X_2 - \bar{X}_2)(Y - \bar{Y}) - \sum(X_1 - \bar{X}_1)(X_1 - \bar{X}_1)}{\sum(X_2 - \bar{X}_2)^2}$$

X_1	X_2	Y	$X_1 - \bar{X}_1$	$X_2 - \bar{X}_2$	$Y - \bar{Y}$	$(X_1 - \bar{X}_1)(Y - \bar{Y})$	$(X_2 - \bar{X}_2)(Y - \bar{Y})$
1	2	5	-2	-1	-3	6	3
2	1	6	-1	-2	-2	2	4
3	4	9	0	1	1	0	1
4	3	8	1	0	0	0	0
5	5	12	2	2	4	8	8

Toplamlar:

$$\sum(X_1 - \bar{X}_1)(Y - \bar{Y}) = 16$$

$$\sum(X_2 - \bar{X}_2)(Y - \bar{Y}) = 16$$

$$\sum(X_1 - \bar{X}_1)^2 = 10$$

$$\sum(X_2 - \bar{X}_2)^2 = 10$$

Katsayılar:

$$\beta_1 = \frac{16}{10} = 1.6$$

$$\beta_2 = \frac{16}{10} = 1.6$$

1.3 Kesme Noktasının Hesaplanması (β_0)

$$\beta_0 = \bar{Y} - \beta_1 \bar{X}_1 - \beta_2 \bar{X}_2$$

$$\beta_0 = 8 - 1.6(3) - 1.6(3)$$

$$\beta_0 = 8 - 4.8 - 4.8 = -1.6$$

Regresyon denklemi:

$$\hat{Y} = -1.6 + 1.6X_1 + 1.6X_2$$

2. Matrislerle Çözüm

Matris formunda regresyon denklemi:

$$Y = X\beta + \epsilon$$

2.1 Matrislerin Oluşturulması

- Bağımsız değişkenler matrisi (X):

$$X = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 2 & 1 \\ 1 & 3 & 4 \\ 1 & 4 & 3 \\ 1 & 5 & 5 \end{bmatrix}$$

- Bağımlı değişken matrisi (Y):

$$Y = \begin{bmatrix} 5 \\ 6 \\ 9 \\ 8 \\ 12 \end{bmatrix}$$

- Katsayılar matrisi (β):

$$\beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix}$$

2.2 Normal Denklemlerin Çözümü

Regresyon katsayıları:

$$\beta = (X^T X)^{-1} X^T Y$$

1. X^T (X'in transpozu):

$$X^T = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 \\ 2 & 1 & 4 & 3 & 5 \end{bmatrix}$$

2. $X^T X$:

$$X^T X = \begin{bmatrix} 5 & 15 & 15 \\ 15 & 55 & 50 \\ 15 & 50 & 55 \end{bmatrix}$$

3. $(X^T X)^{-1}$: Tersi alınarak:

$$(X^T X)^{-1} = \begin{bmatrix} 1.1 & -0.3 & -0.3 \\ -0.3 & 0.1 & 0.1 \\ -0.3 & 0.1 & 0.1 \end{bmatrix}$$

4. $X^T Y$:

$$X^T Y = \begin{bmatrix} 40 \\ 130 \\ 130 \end{bmatrix}$$

5. β:

$$\beta = (X^T X)^{-1} X^T Y$$

$$= \begin{bmatrix} 1.1 & -0.3 & -0.3 \\ -0.3 & 0.1 & 0.1 \\ -0.3 & 0.1 & 0.1 \end{bmatrix} \begin{bmatrix} 40 \\ 130 \\ 130 \end{bmatrix} = \begin{bmatrix} -1.6 \\ 1.6 \\ 1.6 \end{bmatrix}$$

3. Standart Hata, t ve F İstatistikleri

3.1 Standart Hata

Artıklar (residuals):

$$e = Y - \hat{Y}$$

Hata varyansı:

$$s^2 = \frac{\sum e^2}{n-k} = \frac{2.4}{5-3} = 1.2$$

Standart hata:

$$SE(\beta) = \sqrt{s^2 (X^T X)^{-1}}$$

3.2 t-İstatistikleri

$$t = \frac{\beta}{SE(\beta)}$$

3.3 F-İstatistiği

F-istatistiği:

$$F = \frac{SSR/k}{SSE/(n-k-1)}$$

SSR ve SSE bulunarak F-istatistiği hesaplanır.

Sonuç:

- **Regresyon Denklemi:** $\hat{Y} = -1.6 + 1.6X_1 + 1.6X_2$
- **t ve F İstatistikleri:** Standart hata ve varyans kullanılarak hesaplanır. Bu, iki bağımsız değişkenli regresyon modelinin klasik ve matrisle çözümüdür.