1. HAFTA

TEMEL BİLEŞENLER ANALİZİ

Temel bileşenler analizi, orijinal değişkenlerin birkaç doğrusal birleşiminin varyans-kovaryans yapısının açıklanması ile ilgilenir. Temel bileşenler analizi ile boyut indirgeme yapılır.

Kitle Temel Bileşenleri

Cebirsel olarak temel bileşenler $X_1, X_2, ..., X_p$ gibi p tane rasgele değişkenin doğrusal birleşimleridir. Geometriksel olarak, bu doğrusal birleşimler koordinat eksenleri $X_1, X_2, ..., X_p$ olan orijinal sistemin döndürülmesiyle elde edilen yeni koordinat eksenlerin belirlenmesiyle açıklanabilir. Yeni eksenler değişkenler arasındaki kovaryans yapısıyla elde edilir.

Temel bileşenler $X_1, X_2, ..., X_p$ rasgele değişkenlerinin veya $\underline{X}' = (X_1, X_2, ..., X_p)$ rasgele vektörünün varyans-kovaryans matrisi $\mathbf{\Sigma}$ veya korelasyon matrisi $\boldsymbol{\rho}$ dan elde edilir.

$\underline{X}' = (X_1, X_2, ..., X_p)$ rasgel vektörünün özdeğerleri $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_p \geq 0$ olan $\mathbf{\Sigma}$ varyans-kovaryans matrisine sahip olsun.

$$Y_1 = \underline{l}_1' \underline{X} = l_{11}X_1 + l_{21}X_2 + ... + l_{p1}X_p$$
$$Y_2 = \underline{l}_2' \underline{X} = l_{12}X_1 + l_{22}X_2 + ... + l_{p2}X_p$$
$$\vdots$$
$$Y_p = \underline{l}_p' \underline{X} = l_{1p}X_1 + l_{2p}X_2 + ... + l_{pp}X_p$$

doğrusal birleşimlerini göz önüne alalım. Buradan,

$$Var(Y_i) = \underline{l}_i' \Sigma \underline{l}_i \quad , \quad i = 1, 2, ..., p$$

$$Cov(Y_i, Y_k) = \underline{l}_i' \Sigma \underline{l}_k \; ; \quad i, k = 1, 2, ..., p$$

olarak elde edilir.

$Y_1, Y_2, ..., Y_p$ temel bileşenleri birbirleriyle ilişkisiz ve varyans değerlerine göre sıralanmaktadır. Yani $Var(Y_1) = \underline{l}_1' \Sigma \underline{l}_1$ maksimumdur. Açıkça $Var(Y_1) = \underline{l}_1' \Sigma \underline{l}_1$ her hangi bir sabit ile çarpılarak artırılabilir. Bu belirsizliğin giderilmesi için katsayı vektörleri birim uzunlukta olacak biçimde belirlenmeye çalışılır.

Böylece birinci temel bileşen, $\underline{l}_1' \underline{l}_1 = 1$ kısıtına göre $Var(Y_1) = Var(\underline{l}_1' \underline{X})$ değerini maksimum yapan $\underline{l}_1' \underline{X}$ doğrusal birleşimidir. İkinci temel bileşen, $\underline{l}_2' \underline{l}_2 = 1$ ve $Cov(Y_1, Y_2) = Cov(\underline{l}_1' \underline{X}, \underline{l}_2' \underline{X}) = 0$ kısıtlarına göre $Var(Y_2) = Var(\underline{l}_2' \underline{X})$ değerini ikinci sırada maksimum yapan $\underline{l}_2' \underline{X}$ doğrusal birleşimidir. Buradan, i inci temel bileşen, $\underline{l}_i' \underline{l}_i = 1$ ve $Cov(Y_i, Y_k) = Cov(\underline{l}_i' \underline{X}, \underline{l}_k' \underline{X}) = 0$, $k < i$ kısıtlarına göre $Var(Y_i) = Var(\underline{l}_i' \underline{X})$ değerini i inci sırada maksimum yapan $\underline{l}_i' \underline{X}$ doğrusal birleşimidir.

Sonuç 1. $\underline{X}' = (X_1, X_2, ..., X_p)$ rasgele vektörünün varyans-kovaryans matrisi $\mathbf{\Sigma}$'nın özdeğer ve özvektör çiftleri $(\lambda_1, \underline{e}_1), (\lambda_2, \underline{e}_2), ..., (\lambda_p, \underline{e}_p)$ olmak üzere i inci temel bileşen

$$Y_i = \underline{e}_i' \underline{X} = e_{1i} X_1 + e_{2i} X_2 + ... + e_{pi} X_p \quad ; \quad i = 1, 2, ..., p$$

ile verilir. Burada $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_p \geq 0$ ve

$$Var(Y_i) = \underline{e}_i' \Sigma \underline{e}_i = \underline{e}_i' \lambda_i \underline{e}_i = \lambda_i \underline{e}_i' \underline{e}_i = \lambda_i \quad ; \quad i = 1, 2, ..., p$$

$$Cov(Y_i, Y_k) = \underline{e}_i' \Sigma \underline{e}_k = \underline{e}_i' \lambda_k \underline{e}_k = \lambda_k \underline{e}_i' \underline{e}_k = 0 \; ; \quad i, k = 1, 2, ..., p, \; i \neq k$$

dır. Eğer bazı λ_i ler eşit ise ilişkili katsayı vektörleri $\underline{e}_i$'ler ve Y_i'lerin seçimi tek değildir.

Sonuç 2. $\underline{X}' = (X_1, X_2, ..., X_p)$ rasgele vektörünün varyans-kovaryans matrisi $\mathbf{\Sigma}$'nın özdeğer ve özvektör çiftleri $(\lambda_1, \underline{e}_1), (\lambda_2, \underline{e}_2), ..., (\lambda_p, \underline{e}_p)$ olsun ve temel bileşenler $Y_1 = \underline{e}_1' \underline{X}, \; Y_2 = \underline{e}_2' \underline{X}, ..., Y_p = \underline{e}_p' \underline{X}$ olmak üzere

$$\begin{aligned}\sigma_{11}+\sigma_{22}+...+\sigma_{pp} &= \sum_{i=1}^{p} Var(X_i) \\ &= \lambda_1+\lambda_2+...+\lambda_p \\ &= \sum_{i=1}^{p} Var(Y_i)\end{aligned}$$

dir.

İspat: $\sigma_{11}+\sigma_{22}+...+\sigma_{pp}=tr(\Sigma)$ dır. $\mathbf{\Lambda}$, $\mathbf{\Sigma}$ nın özdeğerlerinden oluşan diagonal bir matris ve $\mathbf{P}$, $\mathbf{\Sigma}$ nın özdeğerlerine karşılık gelen birim özvektörlerden oluşan ortogonal bir matris olsun Yani $\mathbf{P}=(\underline{e}_1,\underline{e}_2,...,\underline{e}_p)$ ve $\mathbf{PP'}=\mathbf{P'P}=\mathbf{I}$'dir. Spektral ayrışımdan, $\Sigma=\mathbf{P\Lambda P'}$ dır. Buradan,

$$\begin{aligned}tr(\Sigma) &= tr(\mathbf{P\Lambda P'}) \\ &= tr(\Lambda\mathbf{P'P}) \\ &= tr(\Lambda) \\ &= \lambda_1+\lambda_2+...+\lambda_p\end{aligned}$$

dir. Böylece

$$\sum_{i=1}^{p} Var(X_i) = tr(\Sigma) = tr(\Lambda) = \sum_{i=1}^{p} Var(Y_i)$$

dir. Bu sonuç toplam kitle varyansının $\sigma_{11}+\sigma_{22}+...+\sigma_{pp}=\lambda_1+\lambda_2+...+\lambda_p$ oluğunu ve buradan *k* ıncı temel bileşenin varyansının toplam varyansa oranının

$$\frac{\lambda_k}{\sigma_{11}+\sigma_{22}+...+\sigma_{pp}}=\frac{\lambda_k}{\lambda_1+\lambda_2+...+\lambda_p},\quad k=1,2,...,p$$

olduğunu gösterir.

Eğer *p* değeri büyük olduğunda, ilk birkaç temel bileşen için bu oran %80 veya %90 elde edilirse, fazla bilgi kaybı olmadan bu bileşenler orijinal *p* değişkenin yerini alır. Ayrıca orijinal değişkenlerin hangi temel bileşenler üzerinde daha etkin olduğuna da bakılabilir. Bunun için orijinal değişkenler ile temel bileşenler arasındaki korelasyonlara bakmak gerekir.

Sonuç 3. Eğer $Y_1=\underline{e}_1'\underline{X},\ Y_2=\underline{e}_2'\underline{X},...,Y_p=\underline{e}_p'\underline{X}$, $\mathbf{\Sigma}$ varyans-kovaryans matrisinden elde edilen temel bileşenler ise

$$\rho_{X_k,Y_i} = \frac{e_{ki}\sqrt{\lambda_i}}{\sqrt{\sigma_{kk}}} \quad ; \quad i,k = 1,2,...,p$$

X_k rasgele değişkeni ile Y_i temel bileşeni arasındaki korelasyon katsayısıdır. Burada $(\lambda_1,\underline{e}_1),(\lambda_2,\underline{e}_2),...,(\lambda_p,\underline{e}_p)$ ' ler $\boldsymbol{\Sigma}$ 'nın özdeğer ve birim özvektör çiftleridir.

İspat: $Corr(X_k,Y_i) = \rho_{X_k,Y_i} = \dfrac{Cov(X_k,Y_i)}{\sqrt{Var(X_k)}\sqrt{Var(Y_i)}} \quad ; \quad i,k = 1,2,...,p$

dir. $\underline{l}'_k = (0,...,0,1,0,...,0)$ olsun.

Böylece $\underline{l}_k{}' \underline{X} = X_k$

dir ve

$$\begin{aligned} Cov(X_k,Y_i) &= Cov(\ \underline{l}_k{}' \underline{X},\ \underline{e}_i{}' \underline{X}) \\ &= \underline{l}_k{}' \Sigma \underline{e}_i \\ &= \underline{l}_k{}' \lambda_i \underline{e}_i \\ &= \lambda_i \underline{l}_k{}' \underline{e}_i \\ &= \lambda_i e_{ki} \end{aligned}$$

olarak elde edilir.

Ayrıca

$$Var(Y_i) = \lambda_i$$

ve

$$Var(X_k) = \sigma_{kk}$$

olduğundan

$$\rho_{X_k,Y_i} = \frac{Cov(X_k,Y_i)}{\sqrt{Var(X_k)}\sqrt{Var(Y_i)}}$$
$$= \frac{\lambda_i e_{ki}}{\sqrt{\lambda_i}\sqrt{\sigma_{kk}}} \qquad ; \quad i,k = 1,2,...,p$$
$$= \frac{\sqrt{\lambda_i} e_{ki}}{\sqrt{\sigma_{kk}}}$$

dir.

Örnek 1: $\underline{X}' = (X_1, X_2, X_3)$ rasgele vektörüne ilişkin varyans- kovaryans matrisi

$$\Sigma = \begin{bmatrix} 1 & -2 & 0 \\ -2 & 5 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

olsun. $\sum$ varyans-kovaryans matrisini göz önüne alarak

a) Temel bileşenleri elde ediniz.
b) Temel bileşenlerin varyanslarını bulunuz.
c) Temel bileşenler arasındaki kovaryansları bulunuz.
d) Temel bileşenlerin toplam değişimi açıklama oranlarını hesaplayınız.
e) Temel bileşenler ile rasgele değişkenler arasındaki korelasyonları bulunuz.

Çözüm 1:

a) Temel bileşenlerin bulunabilmesi için öncelikle $\sum$ varyans-kovaryans matrisinin özdeğer ve birim özvektörleri elde edilmelidir.

Özdeğerler için : $\det(\sum - \lambda I) = 0$ olacak şekildeki $\lambda \in \mathbb{R}$ değerlerini bulmalıyız.

$$\det\begin{bmatrix} 1-\lambda & -2 & 0 \\ -2 & 5-\lambda & 0 \\ 0 & 0 & 2-\lambda \end{bmatrix} = 0$$

$\lambda_1 = 5.8284$, $\lambda_2 = 2$ ve $\lambda_3 = 0.17$ bulunur.

(Özdeğerler büyükten küçüğe sıralanır : $\lambda_1 \geq \lambda_2 \geq \lambda_3$)

Özdeğerlere karşılık gelen özvektörler : $\sum \underline{x} = \lambda \underline{x}$ eşitliğinden elde edilir. Örneğin , $\lambda_1 = 5.8284$ özdeğerine karşılık gelen özvektörü bulalım:

$$\begin{bmatrix} 1 & -2 & 0 \\ -2 & 5 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 5.8284 \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

denklem sistemi açık formda yazalım :

$$\begin{aligned} x_1 - 2x_2 &= 5.8284x_1 \\ -2x_1 + 5x_2 &= 5.8284x_2 \\ 2x_3 &= 5.8284x_3 \end{aligned}$$

Son eşitlikten $x_3 = 0$ olduğu açıktır. 1. ve 2. denklemler yeniden düzenlendiğinde ikisinin de $4.8284x_1 + 2x_2 = 0$ denklemine eşit oldukları görülecektir. Bu durumda denklemin sonsuz çözümü vardır. Bunlardan biri $x_1 = 1$ alınırsa, $x_2 = -2.4142$ bulunur. Bu durumda

$$\left\| \underline{x} \right\| = \sqrt{(1)^2 + (-2.4142)^2 + (0)^2} = 2.6131$$

$$\underline{e}'_1 = [1 \; -2.4142 \; 0] \,/\, 2.6131 \cong [0.383 \; -0.924 \; 0]$$

Özdeğerler ve birim özvektörler aşağıdaki tabloda verilmiştir:

$$\lambda_1 = 5.8284 \cong 5.83 \qquad \lambda_2 = 2 \qquad \lambda_3 = 0.17$$

$$\underline{e}'_1 = [0.383 \; -0.924 \; 0] \quad \underline{e}'_2 = [0 \;\; 0 \; 1] \qquad \underline{e}'_3 = [0.924 \; -0.383 \; 0]$$

Buradan temel bileşenler:

$$Y_1 = \begin{bmatrix} 0.383 & -0.924 & 0 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} = 0.383\,X_1 - 0.924X_2 + 0X_3$$

$$Y_2 = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} = 0\,X_1 + 0X_2 + 1X_3 = X_3$$

$$Y_3 = \begin{bmatrix} 0.924 & -0.383 & 0 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} = 0.924X_1 - 0.383X_2 + 0X_3$$

elde edilir. Varyans - Kovaryans matrisine bakıldığında, X_3 rasgele değişkeni diğer rasgele değişkenlerle ilişkisiz olduğundan, temel değişkenlerden birinin X_3 olacağı zaten biliniyordu.

b) Temel bileşenlerin varyanslarının bulunması.

1. Yol :

$$\begin{aligned} Var\left(Y_1\right) &= Var\left(0.383\,X_1 - 0.924X_2\right) \\ &= \left(0.383\right)^2 Var\left(X_1\right) + \left(0.924\right)^2 Var\left(X_2\right) - 2\left(0.383\right)\left(0.924\right)Cov\left(X_1, X_2\right) \\ &= \left(0.383\right)^2 \left(1\right) + \left(0.924\right)^2 Var\left(5\right) - 2\left(0.383\right)\left(0.924\right)\left(-2\right) \\ &= 5.83 \\ &= \lambda_1 \end{aligned}$$

$$\begin{aligned} Var\left(Y_2\right) &= Var\left(X_3\right) \\ &= 2 \\ &= \lambda_2 \end{aligned}$$

$$\begin{aligned} Var\left(Y_3\right) &= Var\left(0.924\,X_1 - 0.383X_2\right) \\ &= \left(0.924\right)^2 Var\left(X_1\right) + \left(0.383\right)^2 Var\left(X_2\right) - 2\left(0.924\right)\left(0.383\right)Cov\left(X_1, X_2\right) \\ &= \left(0.924\right)^2 \left(1\right) + \left(0.383\right)^2 \left(5\right) - 2\left(0.924\right)\left(0.383\right)\left(-2\right) \\ &= 0.17 \\ &= \lambda_3 \end{aligned}$$

2. Yol : (Matris işlemleriyle)

$$Var\left(Y_i\right) = Var\left(\underline{e}'_i\,\underline{X}\right) = \underline{e}'_i \sum \underline{e}_i\,, \quad i = 1,2,3$$

$$
\begin{aligned}
Var\left(Y_1\right) &= \underline{e}'_1 \sum \underline{e}_1 \\
&= \begin{bmatrix} 0.383 & -0.924 & 0 \end{bmatrix} \begin{bmatrix} 1 & -2 & 0 \\ -2 & 5 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} 0.383 \\ -0.924 \\ 0 \end{bmatrix} \\
&= \begin{bmatrix} 2,231 & -5,386 & 0 \end{bmatrix} \begin{bmatrix} 0.383 \\ -0.924 \\ 0 \end{bmatrix} \\
&= 5,83
\end{aligned}
$$

elde edilir. Benzer şekilde

$$
\begin{aligned}
Var\left(Y_2\right) &= \underline{e}'_2 \sum \underline{e}_2 \\
&= 2 \\
&= \lambda_2
\end{aligned}
$$

ve

$$
\begin{aligned}
Var\left(Y_3\right) &= \underline{e}'_3 \sum \underline{e}_3 \\
&= 0,17 \\
&= \lambda_3
\end{aligned}
$$

olarak bulunur.

Ayrıca

$$
\begin{aligned}
\sum_{i=1}^{3} Var\left(X_i\right) &= Var\left(X_1\right) + Var\left(X_2\right) + Var\left(X_3\right) \\
&= \sigma_{11} + \sigma_{22} + \sigma_{33} \\
&= 1 + 5 + 2 \\
&= 8
\end{aligned}
$$

$$
\begin{aligned}
\sum_{i=1}^{3} Var\left(Y_i\right) &= Var(Y_1) + Var(Y_2) + Var(Y_3) \\
&= \lambda_1 + \lambda_2 + \lambda_3 \\
&= 5,83 + 2 + 0017 \\
&= 8
\end{aligned}
$$

dir ve buradan

$$
\sum_{i=1}^{3} Var\left(Y_i\right) = \sum_{i=1}^{3} Var\left(X_i\right) = 8
$$

dir.

c) Temel bileşenler arasındaki kovaryansların bulunması.

1. Yol :

$$\begin{aligned} Cov\left(Y_i, Y_k\right) &= Cov\left(\underline{e}'_i \underline{X}, \underline{e}'_k \underline{X}\right) \;;\; i \neq k \\ &= \underline{e}'_i \sum \underline{e}_k \\ &= \underline{e}'_i \lambda_k \underline{e}_k \\ &= \lambda_k \underline{e}'_i \underline{e}_k \\ &= 0 \end{aligned}$$

Burada birim özvektörler bir birine dik olduğundan yukarıdaki sonuç sıfır çıkar.

2. Yol :

$$\begin{aligned} Cov\left(Y_1, Y_2\right) &= Cov\left(0.383\, X_1 - 0.924 X_2, X_3\right) \\ &= \left(0.383\right)\underbrace{Cov\left(X_1, X_3\right)}_{0} - \left(0.924\right)\underbrace{Cov\left(X_2, X_3\right)}_{0} \\ &= 0 \\ Cov\left(Y_1, Y_3\right) &= Cov\left(0.383\, X_1 - 0.924 X_2, 0.924\, X_1 - 0.383 X_2\right) \\ &= \left(0.383\right)\left(0.924\right)\underbrace{Var\left(X_1\right)}_{1} - \left(0.383\right)^2 \underbrace{Cov\left(X_1, X_2\right)}_{-2} \\ &\quad - \left(0.924\right)^2 \underbrace{Cov\left(X_2, X_1\right)}_{-2} - \left(0.924\right)\left(-0.383\right)\underbrace{Var\left(X_2\right)}_{5} \\ &= 0 \end{aligned}$$

$$Cov\left(Y_2, Y_3\right) = Cov\left(X_3, 0.924\, X_1 - 0.383 X_2\right) = 0$$

d) Temel bileşenlerin toplam değişimi açıklama oranlarının hesaplanması.

- 1. Temel bileşenin toplam varyansa katkısı

$$\begin{aligned} \frac{Var\left(Y_1\right)}{\sum_{i=1}^{3} Var\left(Y_i\right)} &= \frac{\lambda_1}{\lambda_1 + \lambda_2 + \lambda_3} \\ &= \frac{5.83}{8} \\ &= 0.73 \end{aligned}$$

- 2. Temel bileşenin toplam varyansa katkısı

$$\frac{Var\left(Y_2\right)}{\sum_{i=1}^{3} Var\left(Y_i\right)} = \frac{\lambda_2}{\lambda_1 + \lambda_2 + \lambda_3}$$
$$= \frac{2}{8}$$
$$= 0.25$$

- 3. Temel bileşenin toplam varyansa katkısı

$$\frac{Var\left(Y_3\right)}{\sum_{i=1}^{3} Var\left(Y_i\right)} = \frac{\lambda_3}{\lambda_1 + \lambda_2 + \lambda_3}$$
$$= \frac{0.17}{8}$$
$$= 0.02$$

Burada ilk iki temel bileşenin varyanslarının toplamının, toplam varyansa katkısı $0.73 + 0.25 = 0.98$ (oldukça yüksek %98) olduğundan, iki temel bileşen yeterlidir. Bunlar Y_1 ve Y_2 temel bileşenlerdir. Böylece X_1, X_2, X_3 rasgele değişkenleri yerine, Y_1 ve Y_2 temel bileşenleri alınarak boyut indirgeme yapılmış olur. Yani 3 boyuta, 2 boyuta indirgenmiş olur.

e) Temel bileşenler ile rasgele değişkenler arasındaki korelasyonları bulunması.

Temel bileşenler ile rasgele değişkenler arasındaki korelasyonlar

$$\rho_{Y_i,X_k} = \frac{e_{ki}\sqrt{\lambda_i}}{\sqrt{\sigma_{kk}}}\ ,\ i,k = 1,2,\ldots,p$$

ile bulunur.

Örneğin, $\underline{e}'_1 = [e_{11}\ e_{21}\ e_{31}] = [0.383\ -0.924\ 0]$ kullanılarak Y_1 temel bileşeninin X_1, X_2, X_3 rasgele değişkenleri ile arasındaki korelasyonlar :

$$\rho_{Y_1,X_1} = \frac{e_{11}\sqrt{\lambda_1}}{\sqrt{\sigma_{11}}} = \frac{0.383\sqrt{5.83}}{\sqrt{1}} = 0.925$$

$$\rho_{Y_1,X_2} = \frac{e_{21}\sqrt{\lambda_1}}{\sqrt{\sigma_{22}}} = \frac{-0.924\sqrt{5.83}}{\sqrt{5}} = -0.998$$

$$\rho_{Y_1,X_3} = \frac{e_{31}\sqrt{\lambda_1}}{\sqrt{\sigma_{33}}} = \frac{0\sqrt{5.83}}{\sqrt{2}} = 0$$

Benzer şekilde diğer korelasyonlar hesaplandığında aşağıdaki tablo elde edilir:

ρ_{Y_i,X_k}	X_1	X_2	X_3
Y_1	0.925	-0.998	0
Y_2	0	0	1
Y_3	0.381	0.070	0

3. temel bileşen önemsiz olduğundan bu korelasyonlar göz ardı edilebilir.