II.TEMEL BİLEŞENLER ANALİZİ

$p > 1$ değişken sayısı olmak üzere, $\underline{\boldsymbol{X}}' = \begin{bmatrix} X_1 & X_2 & \dots & X_p \end{bmatrix}$ değişkenler vektörü verilsin. $E(\underline{\boldsymbol{X}}) = \underline{\boldsymbol{\mu}}$ ve $Cov(\underline{\boldsymbol{X}}) = \boldsymbol{\Sigma}$ olduğunu kabul edelim. $\underline{\boldsymbol{\mu}}$ ve $\boldsymbol{\Sigma}$, $\underline{\boldsymbol{X}}$ değişkenler vektörünün temsil ettiği kitlenin parametreleri olduğundan genelde bilinmezler. Bu durumda en çok olabilirlik tahmin yöntemi ile örneklemden tahmin edilmeye çalışılır. Bu amaçla ilgili kitleden n birimlik bir rastgele örnek çekilir. Örnek birimleri $\underline{\boldsymbol{X}}_1, \underline{\boldsymbol{X}}_2, \dots, \underline{\boldsymbol{X}}_n$ gözlem vektörleri olmak üzere;

$\underline{\boldsymbol{\mu}}$ parametresinin en çok olabilirlik tahmin edicisi $\underline{\overline{\boldsymbol{X}}} = \frac{1}{n}\sum_{i=1}^{n} \underline{\boldsymbol{X}}_i$ örnek ortalama vektörü iken $\boldsymbol{\Sigma}$ parametresinin en çok olabilirlik tahmin edicisi $\boldsymbol{S} = \frac{1}{n-1}\sum_{i=1}^{n}(\underline{\boldsymbol{X}}_i - \underline{\overline{\boldsymbol{X}}})(\underline{\boldsymbol{X}}_i - \underline{\overline{\boldsymbol{X}}})'$ örnek varyans-kovaryans matrisidir. Burada i-nci gözlem vektörü $\underline{\boldsymbol{X}}'_i = \begin{bmatrix} X_{1i} X_{2i} & \dots & X_{pi} \end{bmatrix}$, $i = \overline{1, n}$ olup, p-boyutlu uzayda bir nokta ile temsil edilebilir. Bütün gözlemler p-boyutlu uzayda birer nokta ile temsil edildiğinde ortaya bir noktalar kümesi (bulutu) çıkacaktır.

Çok değişkenli istatistiksel analiz teknikleri genel olarak p tane değişkeni ya da bu p tane değişken üzerinde ölçümü yapılan$n$ tane birimi (birey veya nesne) ele almakta ve ölçümlerden elde edilen veri yapılarını incelemektedir. Örneğin;

- ✓ Değişkenlerden birisini ya da bir kaçını açıklanan (bağımlı) değişken, geriye kalanları açıklayıcı (bağımsız) değişken olarak ele alıp, açıklayıcı değişkenler yardımıyla açıklanan değişkenleri açıklamaya çalışmak (Regresyon analizi: Çoklu veya çok değişkenli lineer regresyon)
- ✓ Değişkenleri iki ruba ayırarak, her bir gruptaki değişkenlerin lineer fonksiyonları olacak şekilde yeni değişkenler bulmaya çalışmak (Kanonik korelasyon analizi)
- ✓ Bu p tane değişkeni bağımlı değişkenler kabul edip, bir ya da birden fazla faktörler yardımıyla, faktörlerin bağımlı değişkenler üzerine etkisini incelemek (Çok değişkenli varyans analizi)
- ✓ Uzayda noktalarla temsil edilen gözlemlerin birbirlerine benzer olanlarını gösteren küme veya sınıfları belirlemek (Kümeleme analizi, sınıflama analizi)
- ✓ Birbirlerine benzer gözlemlerin belirlediği sınıfları birbirinden ayıran fonksiyonları bulmak ve bu fonksiyonlar yardımıyla yeni gözlemleri bu sınıflara atamak (Diskriminant analizi)

$X_1, X_2, \dots, X_p$ ile gösterilen bu değişkenlerden bazılarının birbiri ile ilişkili (bağımlı) olması ve değişken sayısının çok fazla, yani p'nin çok büyük olması durumunda, bahsedilen bu analizlerin uygulanmasında araştırıcıya bir takım sorunlar çıkarabilmektedir.

- ✓ Değişken sayısının çok fazla olması hem işlem yükünü ve maliyeti arttırır hem de elde edilen sonuçların yorumlanmasında güçlükler ortaya çıkartabilir.
- ✓ p (bağımlı değişken sayısı) >n (gözlem sayısı) ise regresyon analizinde çözümleme etkisiz ve hatta imkânsız olabilir. Öte yandan $p \cong s.d_{hata}$ iken varyans analizinde uygulanan testin gücü düşer, eğer $p > s.d_{hata}$ ise test uygulanamaz.
- ✓ Bağımsız (açıklayıcı) değişkenlerin yüksek derecede ilişkili olması durumunda regresyon analizinde regresyon parametrelerinin tahminleri tutarsız olabilir. Çünkü bu durumda çoklu bağlantı sorununun ortaya çıkması muhtemeldir.

Bahsedilen bu tür sorunların giderilmesinde başvurulacak istatistiksel tekniklerin en önemlisi temel bileşenler analizi (TBA) dir.

TBA, başlangıç sistemindeki$p$ tane X_j değişkenlerinin varyans-kovaryans veya korelasyon yapısını daha az sayıda ve bu değişkenlerin doğrusal bileşimleri olan yeni değişkenlerle ifade etme yöntemidir. Diğer bir ifade ile aralarında orta ya da yüksek derecede korelasyonlar olan p tane X_j değişkenlerinin açıkladığı yapıyı (varyans-kovaryans veya korelasyon) daha az sayıda ($k \leq p$) ve başlangıç değişkenlerin doğrusal fonksiyonları olan yeni değişkenlerle ifade etme yöntemine TBA denir.

TBA'nin üç temel amacı vardır. Bunlar:

i) Değişkenler arasındaki bağımlılık yapısını ortadan kaldırmak (çoklu bağlantıyı gidermek)

ii) Boyut (veri) indirgemesi yapmak

iii) Başka istatistiksel analizler için veri hazırlamak

şeklinde ifade edilir. TBA bu üç amacı gerçekleştirirken başlangıç sistemine ait toplam varyansı da büyük oranda açıklamaya özen gösterir. Başlangıç sistemine ait toplam varyans:

$$\sigma^2_{Top.} = \text{İ}z(\mathbf{\Sigma}) = \sigma_{11} + \sigma_{22} + \cdots + \sigma_{pp} \quad ,(\mathbf{\Sigma} - \text{biliniyorsa}) \tag{2.1}$$

$$\hat{\sigma}^2_{Top} = \text{İ}z(\boldsymbol{S}) = s_{11} + s_{22} + \cdots + s_{pp} \quad ,(\mathbf{\Sigma} - \text{bilinmiyorsa}) \tag{2.2}$$

eşitlikleri ile verilir.

TBA'ne duyulan ihtiyacın gerekçeleri şu şekilde açıklanabilir:

i) Bir oluşumu ya da istatistiksel olayı tanımlamada p tane değişken kullanılmışsa, bu değişkenlerin belirlediği toplam varyansı ifade etmek üzere k sayıda temel bileşen bulmak, böylece daha az sayıda değişken ile çalışarak p-boyutlu uzay yerine k-boyutlu ($k \leq p$) bir uzayda çalışmak, yani boyut indirgemek amaçlanır. Burada $p$tane orijinal değişkenler arasında yüksek ya da orta düzeyde korelasyonlar olabilir. Durum böyle iken, bu değişkenlerin belirlediği toplam varyansı ifade etmek üzere aralarında korelasyon bulunmayan k sayıda doğrusal bileşen (temel bileşen) ile orijinal durumu açıklamak mümkündür.

ii) Çok değişkenli istatistiksel analiz tekniklerinde genellikle değişkenler arasında önemli düzeyde yüksek kovaryans ya da korelasyon olması arzu edilmez. Kovaryans veya korelasyonlardan veri setinin arındırılarak kullanılması daha uygun olur. Bu sebeple p tane ilişkili değişkeni bu değişkenlerin doğrusal bileşenleri olan ve aralarında korelasyon bulunmayan yeni yapay değişkenlerle (temel bileşenlerle) ifade etmek mümkündür.

iii) TB'ler bizzat kendileri bir sonuç olmaktan ziyade, sonuç almayı sağlayıcı özelliğe sahiptirler. Çünkü TB'ler daha kapsamlı incelemeler için bir ara adım özelliği taşırlar. Çoklu lineer regresyon uygulanması istenilen fakat lineer regresyon varsayımlarından birisi olan çoklu bağlantı şartının yerine gelmemesi nedeniyle lineer regresyon analizinin uygulanamadığı durumlarda çoklu lineer regresyona bir giriş değeri olabilir. TB'ler, genellikle birimlerin bazen de değişkenlerin sınıflandırılmasında kullanılan kümeleme analizinde bir veri olabileceği gibi, ölçeklenmiş (ağırlıklandırılmış) temel bileşenler veya faktör analizi modelleri için kovaryans matrisinin bir faktörleşmesi olabilirler.

II.1 Temel Bileşenler Analizinin Gerekliliği

$X_1, X_2, \ldots, X_p$ değişkenlerinin veya $\underline{\boldsymbol{X}}' = \begin{bmatrix} X_1 & X_2 & \ldots & X_p \end{bmatrix}: p \times 1$ boyutlu değişkenler vektörünün oluşturduğu kitle için kitle ortalama vektörü $E(\underline{\boldsymbol{X}}) = \underline{\boldsymbol{\mu}}$ ve kitle varyanskovaryans matrisi $Cov(\underline{\boldsymbol{X}}) = \boldsymbol{\Sigma}$ olsun. Bu durumda $\underline{\boldsymbol{X}}$ değişkenler vektörünün standart değişkenler vektörü;

$\boldsymbol{\Sigma}^{1/2} = K\text{öş}\left[\sigma_1,\ \sigma_2,\ \ldots, \sigma_p\right]$ olmak üzere

$$\underline{\boldsymbol{Z}} = \begin{bmatrix} Z_1 \\ Z_2 \\ . \\ . \\ . \\ Z_p \end{bmatrix} = \left(\boldsymbol{\Sigma}^{1/2}\right)^{-1} \left(\underline{\boldsymbol{X}} - \underline{\boldsymbol{\mu}}\right) \tag{2.3}$$

şeklinde tanımlanır. Burada $k = 1, 2, \dots, p$ için $\sigma_k = \sqrt{\sigma_{kk}}$ olup, X_k değişkenine ait standart sapmayı göstermektedir. Eğer $\underline{\boldsymbol{\mu}}$ ve $\boldsymbol{\Sigma}$ bilinmiyorsa standart değişkenler vektörü; $\mathbf{S}^{1/2} = K\text{öş}\left[s_1,\ s_2,\ \dots, s_p\right]$ olmak üzere

$$\underline{\boldsymbol{Z}} = \left(\mathbf{S}^{1/2}\right)^{-1}\left(\underline{\boldsymbol{X}} - \underline{\overline{\boldsymbol{X}}}\right) \tag{2.4}$$

dönüşümü ile elde edilir. Burada $k = 1, 2, \dots, p$ için $s_k = \sqrt{s_{kk}}$ olup, X_k değişkenine ait örnek standart sapmasını göstermektedir.

Standart değişkenler vektörü için ortalama vektörü; $E\left(\underline{\boldsymbol{Z}}\right) = \underline{\mathbf{0}} : p \times 1$ iken; $Kor\left(\underline{\boldsymbol{X}}\right) = Cov\left(\underline{\boldsymbol{Z}}\right) = \boldsymbol{R}: p \times p$ boyutlu korelasyon matrisidir (Gösteriniz?). Bu durumda standart değişkenler sistemi için toplam varyans; $\sigma^2_{Top.} = \text{İ}z(\boldsymbol{R}) = p$ dir. Eğer $\underline{\boldsymbol{X}}$ ya da $\underline{\boldsymbol{Z}}$'nin değişkenleri arasında tam ya da tama yakın bağımsızlık varsa, yani $\boldsymbol{R} = \boldsymbol{I}_p$ oluyorsa, bu durumda başlangıç sistemine $\left(\left[X_1 X_2\ \dots\ X_p\right]\right.$ veya $\left.\left[Z_1 Z_2\ \dots\ Z_p\right]\right)$TBA'nin uygulanmasına gerek yoktur. $\boldsymbol{R} = \boldsymbol{I}_p$ olması, başlangıç sistemine ait $\boldsymbol{X} = \left[\underline{\boldsymbol{X}}_1\ \underline{\boldsymbol{X}}_2\ \dots\ \underline{\boldsymbol{X}}_n\right]: p \times n$ veri matrisinden elde edilen $\mathbf{Z} = \left[\underline{\boldsymbol{Z}}_1\ \underline{\boldsymbol{Z}}_2\ \dots\ \underline{\boldsymbol{Z}}_n\right]: p \times n$ standartlaştırılmış veri matrisindeki gözlem noktalarının ($\underline{\boldsymbol{Z}}_i, i = \overline{1, n}$) p-boyutlu uzayda bir küre oluşturması anlamına gelmektedir. $\boldsymbol{R} = \boldsymbol{I}_p$ olması aynı zamanda $\boldsymbol{\Sigma} = K\text{öş}\left[\sigma_{11},\ \sigma_{22},\ \dots, \sigma_{pp}\right]$ (veya $\boldsymbol{S} = K\text{öş}\left[s_{11},\ s_{22},\ \dots, s_{pp}\right]$) olması anlamına geleceğinden, bu durum başlangıç sistemine ait gözlem noktalarının ($\underline{\boldsymbol{X}}_i, i = \overline{1, n}$) p-boyutlu uzayda bir elipsoid meydana getirdiğini gösterir.

Deneysel veriler için hangi durumlarda $\boldsymbol{R} = \boldsymbol{I}_p$ kabul edileceği bir hipotez testi ile kontrol edilebilir. Bu amaçla uygulanan teste Küresellik Testi denir. Küresellik testi için hipotezler:

$$H_0: \boldsymbol{R} = \boldsymbol{I}_p$$

$$H_1: \boldsymbol{R} \neq \boldsymbol{I}_p \tag{2.5}$$

şeklide kurulur.H_0 hipotezini test etmek için test istatistiği;

$$\chi^2 = -\left[n - 1 - \frac{2p+5}{6}\right] ln(|\boldsymbol{R}|) \sim \chi^2_{\frac{p(p-1)}{2}} \tag{2.6}$$

dir. Karar kuralı; α önem seviyesinde H_1'e göre kritik değer $\chi^2_t = \chi^2_{\frac{p(p-1)}{2};\alpha}$ olmak üzere, eğer $\chi^2 > \chi^2_t$ ise H_0 hipotezi ret edilir. Bu durumda başlangıç sistemine TBA'nin uygulanması gereklidir. Eğer $\chi^2 \leq \chi_t^2$ ise H_0 hipotezi ret edilemez. Bu durumda başlangıç sistemine TBA'nin uygulanması gerekmez.

II.2 Temel Bileşenlerin Elde Edilmesi

Temel bileşenlerin elde edilmesinde, başlangıç sistemine ait değişkenlerin ölçüm birimlerinin aynı türden olup olmaması ve kitle varyans-kovaryans matrisinin bilinip bilinmemesi dikkate alınır.

i) Başlangıç sistemine ait değişkenler aynı türden ölçüm birimlerine sahip ve kitle varyans-kovaryans matrisi biliniyorsa, elde edilecek olan temel bileşenler Kitle temel bileşenleri adını alır ve $\mathbf{\Sigma}$ matrisine ait bilgiden yararlanarak elde edilir. Bu temel bileşenler $X_k, (k = \overline{1,p})$ başlangıç değişkenlerinin lineer fonksiyonları olacaktır.

ii) Başlangıç sistemine ait değişkenler aynı türden ölçüm birimlerine sahip ve kitle varyans-kovaryans matrisi bilinmiyorsa, elde edilecek olan temel bileşenler örnek temel bileşenleri adını alır ve $\mathbf{S}$ matrisine ait bilgiden yararlanarak elde edilir. Bu temel bileşenler $X_k, (k = \overline{1,p})$ başlangıç değişkenlerinin lineer fonksiyonları olacaktır.

iii) Başlangıç sistemine ait değişkenler farklı türden ölçüm birimlerine sahipse, temel bileşenler $\boldsymbol{\rho}$ veya $\mathbf{R}$ matrisine ait bilgiden yararlanarak elde edilir. Bu temel bileşenler $Z_k, (k = \overline{1,p})$ standart değişkenlerinin lineer fonksiyonları olarak elde edilecektir.

II.2.1 Kitle Temel Bileşenleri

$\underline{\boldsymbol{X}}' = \left[X_1 X_2 \ \dots \ X_p\right]: p \times 1$ boyutlu rastgele değişkenleri vektörünün oluşturduğu kitle için kitle ortalama vektörü $E(\underline{\boldsymbol{X}}) = \underline{\boldsymbol{\mu}}: p \times 1$, kitle varyans-kovaryans matrisi $Cov(\underline{\boldsymbol{X}}) = \boldsymbol{\Sigma} : p \times p$ ve kitle korelasyon matrisi $Kor(\underline{\boldsymbol{X}}) = cov(\underline{\boldsymbol{Z}}) = \boldsymbol{\rho} : p \times p$ olsun. $X_k, (k = \overline{1,p})$ değişkenlerinin aynı türden (veya farklı türden) ölçüm birimlerine sahip olduğunu ve $\mathbf{\Sigma}$ matrisinin (veya $\boldsymbol{\rho}$ matrisinin) de biliniyor olduğunu kabul edelim. Eğer $X_k, (k = \overline{1,p})$ değişkenleri aynı türden ölçüm birimlerine sahipse, TB'ler $\boldsymbol{\Sigma}$ matrisine ait bilgi kullanılarak elde edilir. Eğer $X_k, (k = \overline{1,p})$ değişkenleri farklı türden ölçüm birimlerine sahipse, TB'ler $\boldsymbol{\rho}$ matrisine ait bilgi kullanılarak elde edilir. TB'ler, cebirsel anlamda birbirleri ile ilişkili olan $X_k, (k = \overline{1,p})$ değişkenlerinin [veya $Z_k, (k = \overline{1,p})$ değişkenlerinin] doğrusal fonksiyonları olacağından, j-nci TB; eğer $\mathbf{\Sigma}$ matrisine ait bilgiler kullanılacaksa

$$Y_j = \underline{\boldsymbol{a}}'_j \underline{\boldsymbol{X}} = a_{j1} X_1 + a_{j2} X_2 + \dots + a_{j\,p} X_p \ , \ j = 1, 2, \dots, p \qquad (2.7)$$

şeklinde, eğer $\boldsymbol{\rho}$ matrisine ait bilgiler kullanılacaksa

$$Y_j = \underline{\boldsymbol{a}}'_j\underline{\boldsymbol{Z}} = a_{j1}Z_1 + a_{j2}Z_2 + \cdots + a_{j\,p}Z_p \ , \ j = 1, 2, \dots, p \tag{2.8}$$

şeklinde yazılabilir. Bu durumda j-nci TB'nin varyansı; Eşitlik (2.7)'ye göre;

$$Var(Y_j) = Var(\underline{\boldsymbol{a}}'_j\underline{\boldsymbol{X}}) = \underline{\boldsymbol{a}}'_jVar(\underline{\boldsymbol{X}})\underline{\boldsymbol{a}}_j = \underline{\boldsymbol{a}}'_j\boldsymbol{\Sigma}\underline{\boldsymbol{a}}_j, j = 1, 2, \dots, p \tag{2.9}$$

iken, Eşitlik (2.8)'e göre:

$$Var(Y_j) = Var(\underline{\boldsymbol{a}}'_j\underline{\boldsymbol{Z}}) = \underline{\boldsymbol{a}}'_jVar(\underline{\boldsymbol{Z}})\underline{\boldsymbol{a}}_j = \underline{\boldsymbol{a}}'_j\boldsymbol{\rho}\underline{\boldsymbol{a}}_j, j = 1, 2, \dots, p \tag{2.10}$$

olarak elde edilir. Ayrıca $j \neq t$ olmak üzere j-nci TB $Y_j = \underline{\boldsymbol{a}}'_j\underline{X}$ (veya $Y_j = \underline{\boldsymbol{a}}'_j\underline{Z}$) ile t-nci TB $Y_t = \underline{\boldsymbol{a}}'_t\underline{\boldsymbol{X}}$ (veya $Y_t = \underline{\boldsymbol{a}}'_t\underline{\boldsymbol{Z}}$) arasındaki kovaryans, $\boldsymbol{\Sigma}$ matrisine ait bilgiler kullanıldığında;

$$Cov(Y_j , Y_t) = Cov(\underline{\boldsymbol{a}}'_j\underline{\boldsymbol{X}} , \underline{\boldsymbol{a}}'_t\underline{\boldsymbol{X}}) = \underline{\boldsymbol{a}}'_jCov(\underline{\boldsymbol{X}})\underline{\boldsymbol{a}}_t = \underline{\boldsymbol{a}}'_j\boldsymbol{\Sigma}\underline{\boldsymbol{a}}_t , j \neq t = 1, 2, \dots, p \tag{2.11}$$

iken, $\boldsymbol{\rho}$ matrisine ait bilgiler kullanıldığında;

$$Cov(Y_j , Y_t) = Cov(\underline{\boldsymbol{a}}'_j\underline{\boldsymbol{Z}} , \underline{\boldsymbol{a}}'_t\underline{\boldsymbol{Z}}) = \underline{\boldsymbol{a}}'_jCov(\underline{\boldsymbol{Z}})\underline{\boldsymbol{a}}_t = \underline{\boldsymbol{a}}'_j\boldsymbol{\rho}\underline{\boldsymbol{a}}_t , j \neq t = 1, 2, \dots, p \tag{2.12}$$

olacaktır.

Eşitlik (2.7) veya Eşitlik (2.8) ile tanımlanan TB'lerin elde edilmesinde izlenecek yol şu şekildedir:

1-nci TB: $Y_1 = \underline{\boldsymbol{a}}'_1\underline{\boldsymbol{X}}$ (veya $Y_1 = \underline{\boldsymbol{a}}'_1\underline{\boldsymbol{Z}}$) ile gösterilirse, bu TB $\underline{\boldsymbol{a}}'_1\underline{\boldsymbol{a}}_1 = 1$ yan şartı altında en büyük varyansa sahip olan Y_j , ($j = 1, 2, \dots, p$) doğrusal bağıntısı olmalıdır.

2-nci TB: $Y_2 = \underline{\boldsymbol{a}}'_2\underline{\boldsymbol{X}}$ (veya $Y_2 = \underline{\boldsymbol{a}}'_2\underline{\boldsymbol{Z}}$) ile gösterilirse, bu TB $\underline{\boldsymbol{a}}'_2\underline{\boldsymbol{a}}_2 = 1$ ve $\underline{\boldsymbol{a}}'_2\underline{\boldsymbol{a}}_1 = 0$ yan şartları altında Y_1'den sonra ikinci en büyük varyansa sahip olan Y_j , ($j = 1, 2, \dots, p$) doğrusal bağıntısı olmalıdır.

Bu şekilde devam ederek;

j −nci TB: $Y_j = \underline{\boldsymbol{a}}'_j\underline{\boldsymbol{X}}$ (veya $Y_j = \underline{\boldsymbol{a}}'_j\underline{\boldsymbol{Z}}$) ile gösterilirse, bu TB $\underline{\boldsymbol{a}}'_j\underline{\boldsymbol{a}}_j = 1$ ve $t = 1, 2, \dots, (j - 1)$ olmak üzere tüm (j , t) ikilileri için $\underline{\boldsymbol{a}}'_j\underline{\boldsymbol{a}}_t = 0$ yan şartları altında $Y_1, Y_2, \dots, Y_{(j-1)}$'den sonra j-nci en büyük varyansa sahip olan Y_j , ($j = 1, 2, \dots, p$) doğrusal bağıntısı olmalıdır.

Bu yolla oluşturulacak olan p tane TB için doğal olarak p-nci TB olan $Y_p = \underline{\boldsymbol{a}}'_p\underline{\boldsymbol{X}}$ (veya $Y_p = \underline{\boldsymbol{a}}'_p\underline{\boldsymbol{Z}}$), $\underline{\boldsymbol{a}}'_p\underline{\boldsymbol{a}}_p = 1$ ve $t = 1, 2, \dots, (p - 1)$ olmak üzere her (p , t) ikilileri için $\underline{\boldsymbol{a}}'_p\underline{\boldsymbol{a}}_t = 0$ yan şartları altında $Y_1, Y_2, \dots, Y_{(p-1)}$'den sonra j-nci en büyük varyansa bir diğer ifadeyle en küçük varyansa sahip olan Y_p doğrusal bağıntısıdır.

Bu şekilde elde edilen TB'lerin hepsi birlikte düşünüldüğünde, Eşitlik (2.7) ile gösterilen TB'ler matris formunda;

$$\underline{\boldsymbol{Y}} = \boldsymbol{A}\underline{\boldsymbol{X}} \tag{2.13}$$

şeklinde veya Eşitlik (2.8) ile gösterilen TB'ler matris formunda;

$$\underline{\boldsymbol{Y}} = \boldsymbol{A}\underline{\boldsymbol{Z}} \tag{2.14}$$

şeklinde ifade edilebilir. Burada $\underline{\boldsymbol{Y}}' = [Y_1 \quad Y_2 \quad \dots \quad Y_p]$: $1 \times p$ boyutlu TB'ler vektörü iken

$$\boldsymbol{A} = \begin{bmatrix} \underline{\boldsymbol{a}}_1' \\ \underline{\boldsymbol{a}}_2' \\ \vdots \\ \underline{\boldsymbol{a}}_p' \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ & & \vdots & \\ a_{p1} & a_{p2} & \cdots & a_{pp} \end{bmatrix}$$: $p \times p$ boyutlu TB'ler katsayılar matrisidir.

TB'lerin elde edilmesindeki yan şartlar dikkate alındığında; $\underline{\boldsymbol{a}}'_j \underline{\boldsymbol{a}}_t = \begin{cases} 1, & j = t \ ise \\ 0, & j \neq t \ ise \end{cases}$, $(j, t = 1, 2, \dots, p)$ olduğundan, her bir TB için katsayı vektörleri birim-normal vektörlerdir. Yani bu vektörler birim uzunluğa sahip ve birbirlerine dik (ortoganal) vektörlerdir. Bu sebeple $\boldsymbol{A}$ matrisi bir ortogonal matristir. Buna göre birbileri ile ilişkili olan değişkenlerin $(X_k, k = 1, 2, \dots, p$ veya $Z_k, k = 1, 2, \dots, p)$ oluşturduğu sistemden, TB'leri elde edebilmek için $\underline{X}$ (veya $\underline{Z}$) değişkenler vektörünü soldan bir ortogonal matris ile çarpmak gerekir. Bu takdirde elde edilecek olan TB'ler birbirleri ile ilişkisiz olacaktır.

Eşitlik (2.13) [veya (2.14)] ile elde edilen TB'ler vektörünün ortalama vektörü ile varyans-kovaryans matrisi sırasıyla (kitle için);

$E(\underline{\boldsymbol{Y}}) = \boldsymbol{A}E(\underline{\boldsymbol{X}}) = \boldsymbol{A}\underline{\boldsymbol{\mu}}$ ve

$$Cov(\underline{\boldsymbol{Y}}) = Cov(\boldsymbol{A}\underline{\boldsymbol{X}}) = \boldsymbol{A}cov(\underline{\boldsymbol{X}})\boldsymbol{A}' = \boldsymbol{A\Sigma A}' = \begin{bmatrix} \sigma_{Y_1}^2 & 0 & \cdots & 0 \\ 0 & \sigma_{Y_2}^2 & \cdots & 0 \\ & & \vdots & \\ 0 & 0 & \cdots & \sigma_{Y_p}^2 \end{bmatrix} \tag{2.15}$$

iken, $E(\underline{\boldsymbol{Y}}) = \boldsymbol{A}E(\underline{\boldsymbol{Z}}) = \underline{\boldsymbol{0}}$ ve

$$Cov(\underline{\boldsymbol{Y}}) = Cov(\boldsymbol{A}\underline{\boldsymbol{X}}) = \boldsymbol{A}cov(\underline{\boldsymbol{X}})\boldsymbol{A}' = \boldsymbol{A}\boldsymbol{\rho}\boldsymbol{A}' = \begin{bmatrix} \sigma_{Y_1}^2 & 0 & \cdots & 0 \\ 0 & \sigma_{Y_2}^2 & \cdots & 0 \\ & \vdots & & \\ 0 & 0 & \cdots & \sigma_{Y_p}^2 \end{bmatrix} \tag{2.16}$$

olacaktır. Özellikle TB'lerin kovaryans matrislerinin elde edilmesinde $\boldsymbol{A}$ matrisinin ortogonal olduğu dikkate alınmıştır.

Şimdi önce Eşitlik (2.7) ile verilen TB'leri temsilen $Y = \underline{\boldsymbol{a}}'\underline{\boldsymbol{X}}$ doğrusal bağıntısını ele alalım. Bu doğrusal bağıntı için Eşitlik (2.9) gereğince $Var(Y) = \underline{\boldsymbol{a}}'\boldsymbol{\Sigma}\underline{\boldsymbol{a}}$ olacaktır. Birinci TB'ni bulmak için bu varyansı, $\underline{\boldsymbol{a}}'\underline{\boldsymbol{a}} = 1$ yan şartı altında maksimize etmeliyiz. Burada hem amaç fonksiyonu olan $Var(Y)$ hem de yan şart olan $\underline{\boldsymbol{a}}'\underline{\boldsymbol{a}} = 1$ ifadesi ikinci dereceden terimler olduğundan bu optimizasyon probleminin çözümünde Lagrange çarpanları yönteminden yararlanılır. Bu yönteme göre $\lambda > 0$ Lagrange çarpanı olmak üzere Lagrange fonksiyonu;

$$L(\underline{\boldsymbol{a}}, \lambda) = Var(Y) - \lambda(\underline{\boldsymbol{a}}'\underline{\boldsymbol{a}} - 1) = \underline{\boldsymbol{a}}'\boldsymbol{\Sigma}\underline{\boldsymbol{a}} - \lambda(\underline{\boldsymbol{a}}'\underline{\boldsymbol{a}} - 1)$$

olacaktır. Bu fonksiyonun $\underline{\boldsymbol{a}}$ vektörüne göre türevi alınırsa;

$$\frac{\partial L}{\partial \underline{\boldsymbol{a}}} = 2\boldsymbol{\Sigma}\underline{\boldsymbol{a}} - 2\lambda\underline{\boldsymbol{a}} = 2(\boldsymbol{\Sigma} - \lambda\boldsymbol{I}_p)\underline{\boldsymbol{a}} = \underline{\boldsymbol{0}} \;\Rightarrow$$

$$(\boldsymbol{\Sigma} - \lambda\boldsymbol{I}_p)\underline{\boldsymbol{a}} = \underline{\boldsymbol{0}} \tag{2.17}$$

elde edilir. Burada $\underline{\boldsymbol{a}} \neq \underline{\boldsymbol{0}}$ olduğundan $|\boldsymbol{\Sigma} - \lambda\boldsymbol{I}_p| = 0$ olmak zorundadır. Bu eşitliğin sol tarafındaki determinantın değeri λ'ya göre p-nci dereceden bir polinom olup, bu polinomun kökleri (p-tane), $\boldsymbol{\Sigma}$ matrisinin özdeğerleridir. $\boldsymbol{\Sigma}$ matrisi simetrik ve pozitif tanımlı olduğundan, özdeğerleri $\lambda_1 > \lambda_2 > \cdots > \lambda_p \geq 0$ sıralamasına uygundur. Elde edilen λ_j ($j = 1, 2, \dots, p$) özdeğerleri Eşitlik (2.17)'de yerine yazılmak suretiyle oluşturulacak olan; $(\boldsymbol{\Sigma} - \lambda_j\boldsymbol{I}_p)\underline{\boldsymbol{a}} = \underline{\boldsymbol{0}}$ homojen lineer denklem sisteminin çözüm vektörü $\boldsymbol{\Sigma}$ matrisinin λ_j özdeğerine karşılık gelen birim normal özvektörü olacaktır. Bu özvektör $\underline{\boldsymbol{a}}_j$ile gösterilir. Öyle ki; $j, t = 1, 2, \dots, p$ için

$$\underline{\boldsymbol{a}}'_j\underline{\boldsymbol{a}}_t = \begin{cases} 1, & j = t \;\Rightarrow\; \|\underline{\boldsymbol{a}}_j\| = 1 \\ 0 & j \neq t \;\Rightarrow \underline{\boldsymbol{a}}_j \perp \underline{\boldsymbol{a}}_t \end{cases} \tag{2.18}$$

özelliği sağlanır. $\boldsymbol{\Sigma}$ matrisinin özdeğerler matrisi $\boldsymbol{\Lambda} = K\ddot{o}ş[\lambda_1, \lambda_2, \dots, \lambda_p]$ ile özvektörler matrisi de $\boldsymbol{A} = \begin{bmatrix} \underline{\boldsymbol{a}}_1' \\ \underline{\boldsymbol{a}}_2' \\ \vdots \\ \underline{\boldsymbol{a}}_p' \end{bmatrix}$ ile gösterilirse, $\boldsymbol{\Sigma}$ matrisi özdeğerleri ve özvektörleri cinsinden;

$$\mathbf{\Sigma} = \mathbf{A}'\mathbf{\Lambda}\mathbf{A} \tag{2.19}$$

şeklinde yazılabilir. Bu yazılışa **Σ** matrisinin Spektral Ayrışımı denir.

Sonuç olarak Eşitlik (2.7) ve (2.9) dikkate alındığında j-nci TB; $\underline{\boldsymbol{a}}'_j$, **Σ** matrisinin λ_j özdeğerine karşılık gelen birim normal özvektör olmak üzere; $Y_j = \underline{\boldsymbol{a}}'_j\underline{\boldsymbol{X}}$ doğrusal bağıntısıdır. Bu TB'nin varyansı;

$$Var(Y_j) = Var(\underline{\boldsymbol{a}}'_j\underline{\boldsymbol{X}}) = \underline{\boldsymbol{a}}'_j\mathbf{\Sigma}\underline{\boldsymbol{a}}_j = \lambda_j \ , (j = 1, 2, \ldots, p) \tag{2.20}$$

dir. Gerçekten; Eşitlik (2.19) dikkate alındığında $\boldsymbol{A\Sigma} = \mathbf{AA}'\mathbf{\Lambda}\mathbf{A}$ ve **A** **o**rtogonal olduğundan $\mathbf{AA}' = \boldsymbol{I}_p$ ve böylece $\boldsymbol{A\Sigma} = \mathbf{\Lambda A} \Rightarrow \underline{\boldsymbol{a}}'_j\mathbf{\Sigma} = \lambda_j\underline{\boldsymbol{a}}'_j \, , (j = 1, 2, \ldots, p)$ olduğundan, Eşitlik (2.18)'den $Var(Y_j) = \underline{\boldsymbol{a}}'_j\mathbf{\Sigma}\underline{\boldsymbol{a}}_j = \lambda_j\underline{\boldsymbol{a}}'_j\underline{\boldsymbol{a}}_j = \lambda_j$ bulunur. Buna göre her bir TB, **Σ** matrisinin bir $(\lambda_j \, , \, \underline{\boldsymbol{a}}_j), (\, j = 1, 2, \ldots, p)$ özdeğer-özvektör ikilisi ile ilgilidir. Yani;

$$\left.\begin{array}{l} 1.TB \ \Rightarrow \ Y_1 = \underline{\boldsymbol{a}}'_1\underline{\boldsymbol{X}}\,, \ \ Var(Y_1) = \lambda_1 \\ 2.TB \ \Rightarrow \ Y_2 = \underline{\boldsymbol{a}}'_2\underline{\boldsymbol{X}}\,, \ \ Var(Y_2) = \lambda_2 \\ \qquad\qquad \vdots \\ p.TB \ \Rightarrow \ Y_p = \underline{\boldsymbol{a}}'_p\underline{\boldsymbol{X}}\,, \ \ Var(Y_1) = \lambda_p \end{array}\right\} \tag{2.21}$$

olacaktır. Üstelik $j \neq k = 1, 2, \ldots, p$ için Eşitlik (2.11)'e göre; Eşitlik (2.19) ve (2.18)'den

$Cov(Y_j \, , Y_k) = \underline{\boldsymbol{a}}'_j\mathbf{\Sigma}\underline{\boldsymbol{a}}_k = \lambda_j\underline{\boldsymbol{a}}'_j\underline{\boldsymbol{a}}_k = 0$ elde edilir, yani TB'ler ilişkisizdir. Ayrıca Eşitlik (2.15) ve (2.19) dikkate alındığında **A** matrisinin ortogonal olması sebebiyle TB'ler için kitle varyans kovaryans matrisi:

$$\mathbf{\Sigma}_{\underline{Y}} = \boldsymbol{A\Sigma A}' = \boldsymbol{AA}'\mathbf{\Lambda}\boldsymbol{AA}' = \mathbf{\Lambda} = K\text{öş}[\lambda_1, \lambda_2, \ldots, \ \lambda_p] \tag{2.22}$$

olup, böylece $j = 1, 2, \ldots, p$ için $\sigma^2_{Y_j} = Var(Y_j) = \lambda_j$ elde edilir.

Şimdi de Eşitlik (2.8) ile verilen TB'leri temsilen $Y = \underline{\boldsymbol{a}}'\underline{\boldsymbol{Z}}$ doğrusal bağıntısını ele alalım. Bu doğrusal bağıntı için Eşitlik (2.10) gereğince $Var(Y) = \underline{\boldsymbol{a}}'\boldsymbol{\rho}\underline{\boldsymbol{a}}$ olacaktır. Birinci TB'ni bulmak için bu varyansı, $\underline{\boldsymbol{a}}'\underline{\boldsymbol{a}} = 1$ yan şartı altında maksimize etmeliyiz. Burada hem amaç fonksiyonu olan $Var(Y)$ hem de yan şart olan $\underline{\boldsymbol{a}}'\underline{\boldsymbol{a}} = 1$ ifadesi ikinci dereceden terimler olduğundan bu optimizasyon probleminin çözümünde yine Lagrange çarpanları yönteminden yararlanılır. Bu yönteme göre $\lambda > 0$ Lagrange çarpanı olmak üzere Lagrange fonksiyonu;

$$L(\underline{\boldsymbol{a}}, \lambda) = Var(Y) - \lambda(\underline{\boldsymbol{a}}'\underline{\boldsymbol{a}} - 1) = \underline{\boldsymbol{a}}'\boldsymbol{\rho}\underline{\boldsymbol{a}} - \lambda(\underline{\boldsymbol{a}}'\underline{\boldsymbol{a}} - 1)$$

olacaktır. Bu fonksiyonun $\underline{\boldsymbol{a}}$ vektörüne göre türevi alınırsa;

$$\frac{\partial L}{\partial \underline{\boldsymbol{a}}} = 2\boldsymbol{\rho}\underline{\boldsymbol{a}} - 2\lambda\underline{\boldsymbol{a}} = 2(\boldsymbol{\rho} - \lambda\boldsymbol{I}_p)\underline{\boldsymbol{a}} = \underline{\boldsymbol{0}} \quad \Rightarrow$$

$$(\boldsymbol{\rho} - \lambda\boldsymbol{I}_p)\underline{\boldsymbol{a}} = \underline{\boldsymbol{0}} \tag{2.23}$$

elde edilir. Burada $\underline{\boldsymbol{a}} \neq \underline{\boldsymbol{0}}$ olduğundan $|\boldsymbol{\rho} - \lambda\boldsymbol{I}_p| = 0$ olmak zorundadır. Bu eşitliğin sol tarafındaki determinantın değeri λ'ya göre p-nci dereceden bir polinom olup, bu polinomun kökleri (p-tane), $\boldsymbol{\rho}$ matrisinin özdeğerleridir. $\boldsymbol{\rho}$ matrisi simetrik ve pozitif tanımlı olduğundan, özdeğerleri $\lambda_1 > \lambda_2 > \cdots > \lambda_p \geq 0$ sıralamasına uygundur. Elde edilen λ_j $(j = 1, 2, \dots, p)$ özdeğerleri Eşitlik (2.23)'de yerine yazılmak suretiyle oluşturulacak olan; $(\boldsymbol{\rho} - \lambda_j\boldsymbol{I}_p)\underline{\boldsymbol{a}} = \underline{\boldsymbol{0}}$ homojen lineer denklem sisteminin çözüm vektörü $\boldsymbol{\rho}$ matrisinin λ_j özdeğerine karşılık gelen birim normal özvektörü olacaktır. Bu özvektör $\underline{\boldsymbol{a}}_j$ ile gösterilir. Öyle ki; $j, t = 1, 2, \dots, p$ için

$$\underline{\boldsymbol{a}}'_j\underline{\boldsymbol{a}}_t = \begin{cases} 1, & j = t \;\; \Rightarrow \;\; \|\underline{\boldsymbol{a}}_j\| = 1 \\ 0 & j \neq t \;\; \Rightarrow \underline{\boldsymbol{a}}_j \perp \underline{\boldsymbol{a}}_t \end{cases} \tag{2.24}$$

özelliği sağlanır. $\boldsymbol{\rho}$ matrisinin özdeğerler matrisi $\boldsymbol{\Lambda} = K\text{öş}[\lambda_1, \lambda_2, \dots, \lambda_p]$ ile özvektörler matrisi de $\boldsymbol{A} = \begin{bmatrix} \underline{\boldsymbol{a}}'_1 \\ \underline{\boldsymbol{a}}'_2 \\ \vdots \\ \underline{\boldsymbol{a}}'_p \end{bmatrix}$ ile gösterilirse, $\boldsymbol{\rho}$ matrisi özdeğerleri ve özvektörleri cinsinden;

$$\boldsymbol{\rho} = \mathbf{A}'\boldsymbol{\Lambda}\mathbf{A} \tag{2.25}$$

şeklinde yazılabilir. Bu yazılışa da $\boldsymbol{\rho}$ matrisinin Spektral Ayrışımı denir.

Sonuç olarak Eşitlik (2.8) ve (2.10) dikkate alındığında j-nci TB; $\underline{\boldsymbol{a}}'_j$, $\boldsymbol{\rho}$ matrisinin λ_j özdeğerine karşılık gelen birim normal özvektör olmak üzere $Y_j = \underline{\boldsymbol{a}}'_j\underline{\boldsymbol{X}}$ doğrusal bağıntısıdır. Bu TB'nin varyansı;

$$Var(Y_j) = Var(\underline{\boldsymbol{a}}'_j\underline{\boldsymbol{X}}) = \underline{\boldsymbol{a}}'_j\boldsymbol{\rho}\underline{\boldsymbol{a}}_j = \lambda_j, \; (j = 1, 2, \dots, p) \tag{2.26}$$

dir. Gerçekten; Eşitlik (2.25) dikkate alındığında $\boldsymbol{A}\boldsymbol{\rho} = \mathbf{A}\mathbf{A}'\boldsymbol{\Lambda}\mathbf{A}$ ve **A** **o**rtogonal olduğundan $\mathbf{A}\mathbf{A}' = \boldsymbol{I}_p$ ve böylece $\boldsymbol{A}\boldsymbol{\rho} = \boldsymbol{\Lambda}\mathbf{A} \Rightarrow \underline{\boldsymbol{a}}'_j\boldsymbol{\rho} = \lambda_j\underline{\boldsymbol{a}}'_j, \; (j = 1, 2, \dots, p)$ olduğundan, Eşitlik (2.24)'den $Var(Y_j) = \underline{\boldsymbol{a}}'_j\boldsymbol{\rho}\underline{\boldsymbol{a}}_j = \lambda_j\underline{\boldsymbol{a}}'_j\underline{\boldsymbol{a}}_j = \lambda_j$ bulunur. Buna göre her bir TB, $\boldsymbol{\rho}$ matrisinin bir $(\lambda_j, \underline{\boldsymbol{a}}_j), (j = 1, 2, \dots, p)$ özdeğer-özvektör ikilisi ile ilgilidir. Yani;

$$\left.\begin{array}{l} 1.TB \;\Rightarrow\; Y_1 = \underline{\boldsymbol{a}}'_1\underline{\boldsymbol{Z}}\,,\;\; Var(Y_1) = \lambda_1 \\ 2.TB \;\Rightarrow\; Y_2 = \underline{\boldsymbol{a}}'_2\underline{\boldsymbol{Z}}\,,\;\; Var(Y_2) = \lambda_2 \\ \qquad\qquad\quad \vdots \\ p.TB \;\Rightarrow\; Y_p = \underline{\boldsymbol{a}}'_p\underline{\boldsymbol{Z}}\,,\;\; Var(Y_1) = \lambda_p \end{array}\right\} \tag{2.27}$$

olacaktır. Üstelik $j \neq t = 1, 2, \dots, p$ için Eşitlik (2.12)'ye göre; Eşitlik (2.25) ve (2.24)'den

$Cov(Y_j\,, Y_t) = \underline{\boldsymbol{a}}'_j\boldsymbol{\rho}\underline{\boldsymbol{a}}_t = \lambda_j\underline{\boldsymbol{a}}'_j\underline{\boldsymbol{a}}_t = 0$ elde edilir, yani $\boldsymbol{\rho}$ matrisi kullanılarak standart değişkenlerin doğrusal fonksiyonları şeklinde elde edilen TB'ler de ilişkisizdir. Ayrıca Eşitlik (2.16) ve (2.25) dikkate alındığında **A** matrisinin ortogonal olması sebebiyle TB'ler için kitle varyans kovaryans matrisi:

$$\boldsymbol{\Sigma}_{\underline{Y}} = \boldsymbol{A}\boldsymbol{\rho}\boldsymbol{A}' = \boldsymbol{A}\mathbf{A}'\boldsymbol{\Lambda}\mathbf{A}\boldsymbol{A}' = \boldsymbol{\Lambda} = K\text{öş}\left[\lambda_1, \lambda_2, \dots,\; \lambda_p\right] \tag{2.28}$$

olup, böylece $j = 1, 2, \dots, p$ için $\sigma^2_{Y_j} = Var\left(Y_j\right) = \lambda_j$ elde edilir.