İSTATİSTİKSEL VERİ ANALİZİNDE TEMEL BİLEŞENLERİN YERİ

YÜKSEK LİSANS TEZİ
(İSTATİSTİK)

REŞAT KASAP
ŞUBAT 1987

GAZİ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

Bu tezin Yüksek Lisans tezi olarak uygun olduğunu onaylarım.

Danışman

Yrd. Doç. Dr. Fikri Öztürk

Sınav Jürisi

Başkan: ______ Prof. Dr. Özkan ÜNVER

Üye : ______ Prof. Dr. Altekin ESİN

Üye : ______ Yrd. Doç. Dr. Fikri ÖZTÜRK

Bu Tez Gazi Üniversitesi Fen Bilimleri Enstitüsü Tez Yazım Esaslarına Uygundur.

ÖZET

Reşat Kasap
Yüksek Lisans Tezi
Gazi Üniversitesi Fen Bilimleri Enstitüsü
Şubat 1987

Bu çalışmada, istatistiksel veri analizinde karşılaşılan Temel Bilişenler (TB), Temel Bileşenler Analizi (TBA) ve Temel Bileşenler Regresyonu (TBR) adları altında incelenmesine çalışılmıştır. TBA, $\underline{X}$ tesadüfi vektörünü, bileşenleri ilişkisiz ve varyanslarının büyüklüklerine göre sıralanmış yeni bir tesadüfi vektöre dönüşümü olarak tanımlanmıştır. TBR, çoklubağlantı sorununda incelenmiştir.

ABSTRACT

THE ROLE OF PRINCIPAL COMPONENTS
IN THE STATISTICAL DATA ANALYSIS

Reşat Kasap
M.S. Thesis
Gazi University
Institute of Science and Technology
February 1987

In this study we tried examining the principal Components (PC) faced in the statistical data analysis under the names of Principal Components Analysis (PCA) and Principal Components Regression (PCR). PCA has been defined as a transformation of a random vector $\underline{X}$ in to a new random vector whose components are uncorrelated and, ordered according to their size of variances. And PCR has been examined in the sense of multicollinearity.

TEŞEKKÜR

Bu çalışmanın gerçekleşmesinde bana yön veren ve konunun işlenmesinde büyük katkısı olan hocam, Sayın Yrd.Doç.Dr. Fikri ÖZTÜRK'e en derin saygı ve şükranlarımı sunarım.

İÇİNDEKİLER

Sayfa

GİRİŞ

İstatistiksel veri analizinde karşılaşılan Temel Bileşenler (TB), Temel Bileşenler Analizi (TBA) ve Temel Bileşenler Regresyonu (TBR) adları altında incelenmeye çalışılacaktır.

Çok değişkenli istatistiksel analizin temelini oluşturan TBA, çalışmanın önemli bir kesimini teşkil etmekte olup, tekniğin esası ve oluşumu gösterilmeye çalışılacaktır. Pek çok, çok değişkenli analiz çalışmalarında araştırma konusu olmuş TBA hakkında söylenenlerin bir araya getirilmesine çalışılacaktır.

Lineer modellerde çoklubağlantı yapısının belirlenmesi ve yanlı kestirim ile bu sorunun incelenmesine çalışılacaktır. Bu konuda yapılmış pek çok çalışmalar olmakla beraber sadece TB açısından çoklubağlantının incelenmesi ve önemli yanlı kestirimlerden biri olan TB kestiricisi ayrı bir inceleme konusu olmuştur.

Bu nedenle çalışmanın amacı, az sayıda yeni değişkenler (temel bileşenler)'i türeten ve bilgi kaybını minumum yapmaya çalışan en uygun istatistiksel veri analizlerinden TBA'ni incelemek ve çoklubağlantı yapısını özdeğerlerle belirleme ölçütleri ile yanlı kestirim olan TBR'nunu çoklubağlantı sorununun incelenmesinde göstermektir.

Çok değişkenli analizlerle ilgili çalışmalar 1900 yıllarına dayanır. Çok değişkenli analizler, çok sayıda değişkeni az sayıda lineer bileşimlere indirgeyerek, değişkenler arasındaki karmaşık ilişkilerin yorumlanmasında kolaylık sağlayan istatistiksel analizlerdir.

Temel bileşenler analizi, Karl Pearson (1901) tarafından betimlenmiş ve Hotteling (1933) tarafından geliştirilmiştir.

Girshick (1939) ve Anderson (1958), temel bileşenleri, dağılımlarını ve örnek özelliklerini ele alarak geliştirmişlerdir.

Morrison (1967) temel bileşenlerin geometrik yorumlanmasına ve diğer bazı özelliklerine değinmiştir.

Çokdeğişkenli analizlerin kullanımı, özellikle bilgisayar kullanımının gelişmesinden sonra daha yaygın hale gelmiştir. Cooley ve Lohnes (1971) temel bileşenlerin pratikte yorumlanmasını da vererek, bilgisayara uyarlanmasını göstermişlerdir.

Öngel (1975) doçentlik tezinde faktör çözümüne giriş olarak, TBA'ni ele alarak iki uygulama ile açıklamıştır.

Matematiksel teknikleri kullanarak, çok değişkenli analizlere geçişleri Green (1976) bir arada incelemiştir.

Srivastava v.d. (1979), temel bileşenlerin en çok olabilirlik kestirimlerini ve asimptotik dağılımlarına değinmiştir.

TBA'nin kuramını ve bir uygulamasını Ersoy (1981) çalışmasında vermiştir.

Shurt (1985), TBA'nin kuramsal yapısının temelini teşkil eden yaklaşımı, genelleştirilmiş en küçük kareler ve özdeğerleri kullanarak göstermiştir.

Massy (1965), temel bileşenler regresyonu çoklubağlantı sorununun halledilmesinde kullanıldığını söylemiştir.

Silvey (1969), çoklubağlantı kavramını tanımlamış ve çoklubağlantıyı incelemek için özdeğerlerin ve özvektörlerin kullanımını önermiştir.

Çoklubağlantının özdeğerlerle belirlenmesi ölçütleri, Marquard (1970) ve Mason (1975) tarafından verilmiştir.

Marquard (1970), Temel bileşenler kestiricisi ile ilgili çalışmalar yapmıştır.

Hill v.d. (1977), TBR'da bileşen seçim ölçütlerini göstermişlerdir.

Bölüm I 'de; TBA başlığı altında, yığında TB'in tanımı, değişkenler normal dağıldığında TB'in yorumu, TB'in en çok olabilirlik kestirimi ve TB'in örnek özellikleri hakkında bilgi verilmiştir. Ayrıca birinci temel bileşenin yorumu ve pratikte TB'in yorumu incelenmiştir. Bölümün sonunda bir örnekle tekniğin uygulaması yapılmıştır.

Bölüm II 'de; lineer modellerde temel bileşenleri anlatırken, ilk önce bazı tanımlar yapıldıktan sonra çoklubağlantının özdeğerlerle belirlenmesi ölçütleri verilmiştir. TBR başlığı altında; genelleştirilmiş invers ve kısıtlanmış EKKK olarak TB kestiricisi ve özellikleri ile son olarak da, çıkarılacak bileşenlerin seçim ölçütü verilmiştir. Ayrıca bölümün sonunda, Hald'ın çimento verilerinden oluşan bir örnekle tekniğin uygulaması yapılmıştır.

Bölüm III ́de ise, bölüm I ve bölüm II de verilen tekniklerin, üç değişkenli bir örnek üzerinde uygulaması yapılmıştır.

BÖLÜM I

TEMEL BİLEŞENLER ANALİZİ

I.1. TEMEL BİLEŞENLERİN TANIMI

$\underline{X}$:pxl tesadüfi vektörü Σ kovaryans matrisine sahip olsun. Burada sadece varyans ve kovaryanslarla ilgileneceğimizden $E(\underline{X})=\underline{0}$ olduğunu varsayacağız. Bundan başka fikirlerin ve işlemlerin oluşumunda, $\underline{X}$ in gerçek dağılımı, kovaryans matrisi hariç ilgili değildir. Bununla birlikte, eğer $\underline{X}$ normal dağılmış ise, temel bileşenler daha fazla yorum getirebilir. Aşağıdaki işlemlerde Σ tekil (singular) yani pozitif yarı tanımlı katlı köklere sahip olma durumlarınıda kapsar.

Amacımız, $\underline{X}$ vektörünün bir lineer dönüşümü olan

$$Y_1 = \underline{w}'\underline{X} \qquad (I.1)$$

tesadüfi değişkenini, $\|\underline{w}\|=1$ kısıtlaması altında maksimum varyansa sahip olacak şekilde belirlemektir. Y_1 in varyansı

$$E(\underline{w}'\underline{X}\underline{X}'\underline{w}) = \underline{w}'\Sigma\,\underline{w}$$

olmak üzere Lagrange fonksiyonu

$$\phi(\underline{w},\lambda) = \underline{w}'\Sigma\,\underline{w}-\lambda(\underline{w}'\underline{w}-1) \qquad (I.2)$$

olsun, burada λ Lagrange çarpanıdır. $\underline{w}$ ye göre kısmi türev alınıp sıfıra eşitlenirse

$$\frac{\partial\phi}{\partial\underline{w}} = 2\Sigma\underline{w}-2\lambda\underline{w}=\underline{0} \qquad (I.3)$$

olur. Buradan

$$(\Sigma-\lambda I)\underline{w} = \underline{0} \qquad (I.4)$$

elde edilir. Bu homojen denklem sisteminin $\underline{w} = \underline{0}$ dan başka çözümü olması için

$$|\Sigma-\lambda I| = \underline{0} \qquad (I.5)$$

olmalıdır. (I.5) fonksiyonu λ nın p. dereceden bir polinomudur. Bu yüzden (I.5) ifadesinin p tane kökü vardır; bu kökler $\lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_p$ olsun. (I.4) ifadesini soldan $\underline{w}'$ ile çarparsak

$$\underline{w}'\Sigma\, \underline{w} = \lambda \underline{w}'\underline{w} = \lambda \qquad (I.6)$$

elde ederiz. Buradan (I.1) de verilen $\underline{w}'\underline{X}$ in varyansı λ_1 olur. Böylece maksimum varyans için (I.4) ifadesindeki en büyük λ_1 kökünü kullanacağız. $\underline{w}_1$, $(\Sigma-\lambda_1 I)\underline{w}=\underline{0}$ ın λ_1'e karşılık gelen normlanmış bir çözümü olsun. O zaman $Y_1 = \underline{w}_1'\underline{X}$ maksimum varyanslı lineer bir bileşimdir.

Şimdi, Y_1 ile ilişkisiz tüm lineer bileşimler arasında en büyük varyansa sahip $Y_2 = \underline{w}'\underline{X}$, $\|\underline{w}\|=1$ değişkenini araştıralım. Y_1 ile Y_2 nin ilişkisiz olması

$$\begin{aligned} E(Y_1 \cdot Y_2) &= E[(\underline{w}_1'\underline{X})(\underline{w}'\underline{X})] \qquad (I.7) \\ &= E(\underline{w}'\underline{X}\underline{X}'\underline{w}_1) \\ &= \underline{w}'\Sigma\, \underline{w}_1 \\ &= \lambda_1 \underline{w}'\underline{w}_1 = 0 \end{aligned}$$

olmasıdır. Böylece $\underline{w}'\underline{X}$ ve Y_1 hem istatistiksel anlamda ilişkisiz hem de geometrik anlamda $\underline{w}_1'\underline{w} = 0$ olduğundan ortogonal (dik) dir. Şimdi λ ve t_1 Lagrange çarpanları olmak üzere

$$\phi_2(\underline{w},\lambda,t_1) = \underline{w}'\Sigma\,\underline{w}-\lambda(\underline{w}'\underline{w}-1)-2t_1\underline{w}'\Sigma\,\underline{w}_1 \qquad (I.8)$$

ifadesini maksimum etmek istiyoruz. Kısmi türevi alınıp, sıfıra eşitlenirse

$$\frac{\partial\phi_2}{\partial\underline{w}} = 2\Sigma\,\underline{w}-2\lambda\underline{w}-2t_1\Sigma\,\underline{w}_1 = \underline{0} \qquad (I.9)$$

olup, bunu soldan $\underline{w}_1'$ ile çarparsak (I.7) den

$$\begin{aligned} 0 &= 2\underline{w}_1'\Sigma\,\underline{w}-2\lambda\underline{w}_1'\underline{w}-2t_1\underline{w}_1'\Sigma\,\underline{w}_1 \qquad (I.10)\\ &= -2t_1\lambda_1 \end{aligned}$$

elde ederiz. Bu yüzden, $t_1=0$ ve $\underline{w}$ (I.4)ü sağlamalı, λ da (I.5)i sağlamalıdır. $\lambda_{(2)}$, $(\Sigma-\lambda_{(2)}I)\underline{w} = \underline{0}$, $\underline{w}'\underline{w} = 1$ ve (I.7) yi sağlayan bir $\underline{w}$ vektörü var olacak şekilde $\lambda_1,\ldots,\lambda_p$ lerin en büyüğü olsun, bu vektöre $\underline{w}_2$ diyelim. Bunu karşı gelen lineer bileşim $Y_2= \underline{w}_2'\underline{X}$ olacaktır. (Bunun sonucu $\lambda_{(2)}= \lambda_2$ olduğunu göstermiş olacak ve $\lambda_{(1)}= \lambda_1$ olarak tanımlayalım.)

Bu şekilde devam edildiğinde (r + 1). adımda Y_1, $Y_2,\ldots,Y_r$ ile ilişkisiz bütün $\underline{w}'\underline{X}$, $\|\underline{w}\|=1$, lineer bileşimler arasında maksimum varyansa sahip olacak şekilde bir $\underline{w}$ vektörü bulmak istiyoruz. Yani,

$$\begin{aligned} 0 &= E\,[(\underline{w}'\underline{X})Y_j] = E(\underline{w}'\underline{X}\underline{X}'\underline{w}_j) \qquad (I.11)\\ &= \underline{w}'\Sigma\,\underline{w}_2 = \lambda_{(j)}\underline{w}'\underline{w}_j \qquad j = 1,2,\ldots,r \end{aligned}$$

olacak şekilde; λ ve $t_1,\ldots,t_r$ Lagrange çarpanları olmak üzere

$$\phi_{r+1}(\underline{w},\lambda,t_1,\ldots,t_r) = \underline{w}'\Sigma\,\underline{w}-\lambda(\underline{w}'\underline{w}-1)-2\sum_{j=1}^{r} t_j\underline{w}'\Sigma\,\underline{w}_j \quad (I.12)$$

yi maksimum etmek istiyoruz. (I.12) nin $\underline{w}$ ye göre kısmi türevi alınıp sıfıra eşitlenirse

$$\frac{\partial\phi_{(r+1)}}{\partial\underline{w}} = 2\Sigma\,\underline{w}-2\lambda\underline{w}-2\sum_{j=1}^{r} t_j\Sigma\,\underline{w}_j = \underline{0} \qquad (I.13)$$

olup, bu eşitliği soldan $\underline{w}_j$ ile çarparsak

$$0 = 2\underline{w}_j'\Sigma\,\underline{w}-2\lambda\underline{w}_j'\underline{w}-2t_j\underline{w}_j'\Sigma\,\underline{w}_j \qquad (I.14)$$

elde ederiz. $\lambda_{(j)}\neq 0$ ise $2t_j\lambda_{(j)}= 0$ ve $t_j= 0$ dır. Eğer $\lambda_{(j)}= 0$ ise $\Sigma\,\underline{w}_j= \lambda_{(j)}\underline{w}_j= 0$ ve (I.13) deki toplam sıfır olur. Böylece $\underline{w}$, (I.4) ifadesini sağlamalı ve λ da (I.5) ifadesini sağlamalı.

λ_{r+1}, $(\Sigma-\lambda_{(r+1)}I)\underline{w}=\underline{0}$, $\underline{w}'\underline{w}=1$ ve (I.11) ifadesini sağlayan bir $\underline{w}$ vektörü var olacak şekilde, $\lambda_1,\ldots,\lambda_p$ lerin en büyüğü olsun. Bu vektörü $\underline{w}_{r+1}$ ile ve karşı gelen lineer bileşimi $Y_{r+1}=\underline{W}'_{r+1}\,\underline{X}$ ile gösterelim $\lambda_{r+1}=0$ ve $\lambda_j=0$ ise $\underline{w}_j'\Sigma\underline{w}_{r+1}=0$ olması, $\underline{w}_j'\underline{w}_{r+1}=0$ olmasını gerektirmez. $\underline{w}_{r+1}$ sıfır olan λ_j li $\underline{w}_j$ ve $\underline{w}_{r+1}$ in doğrusal bir bileşimi olarak alınabilir. Bu şekilde alınan $\underline{w}_{r+1}$ $(j=1,\ldots,r)$ tüm $\underline{w}_j$ lerle ortogonaldır. Bu yöndem $\underline{w}'\underline{w}=1$, ve (I.11) ifadelerini sağlayan bir $\underline{w}$ vektörü bulunamayacağı (m+1). adıma kadar uygulanır. $\underline{w}_1,\ldots,\underline{w}_m$ lineer bağımsız olduklarından m=p dır veya $m < p$ dır.

Şimdi $m < p$ eşitsizliği bizi bir çelişkiye götürdüğünü görelim. $m < p$ ise $\underline{w}_i'\underline{e}_j=0$, $\underline{e}_i'\underline{e}_j=\delta_{ij}$ olacak şekilde p-m tane vektör yani $\underline{e}_{m+1},\ldots,\underline{e}_p$ vektörleri mevcuttur. $(\underline{e}_{m+1},\ldots,\underline{e}_p) = M$ olsun. Şimdi $\lambda_1=\theta$ ile

$M\underline{c} = \sum_{i=m+1}^{p} c_i \underline{e}_i$ (I.4) ifadesinin bir çözümü olacak şekilde bir p-m bileşenli $\underline{c}$ vektörü ve bir θ sayısının var olduğunu göstereceğiz. $|M'\Sigma M-\theta I|=0$ ın bir kökü ve $M'\Sigma M\underline{c}=\theta\underline{c}$ yi sağlayan bir köke karşı gelen bir $\underline{c}$ yi gözönüne alalım. $\Sigma M\underline{c}$, $\underline{w}_1,\ldots,\underline{w}_m$ 'e ortogonal bir vektördür. Bunun için $\underline{e}_{m+1},\ldots,\underline{e}_p$ tarafından gerilmiş uzayda bir vektördür ve $M\underline{g}$ olarak yazılabilir (Burada $\underline{g}$, (p-m) bileşenli bir vektördür.) $\Sigma M\underline{c} = M\underline{g}$ soldan M' ile çarpılarak $M'\Sigma M\underline{c} = M'M\underline{g} = \underline{g}$ elde edilir. Böylece $\underline{g}=\theta\underline{c}$ olur ve $\Sigma(M\underline{c}) = \theta(M\underline{c})$ olur. O zaman $(M\underline{c})'\underline{X}$, $(j=1,\ldots,m)$ olmak üzere $\underline{w}_j'\underline{X}$ ile ilişkisiz olur. Bu yeni bir $\underline{w}_{m+1}$ vektörünün var olduğunu ifade eder. Bu durum $m<p$ varsayımımızla çelişeceğinden $m=p$ olmalıdır.[1]

$W = (\underline{w}_1,\underline{w}_2,\ldots,\underline{w}_p)$ ve

$$\Lambda = \begin{bmatrix} \lambda_{(1)} & 0 & \ldots & 0 \\ 0 & \lambda_{(2)} & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & \lambda_{(p)} \end{bmatrix} \qquad \text{(I.15)}$$

gösterimiyle $\Sigma\underline{w}_r = \lambda_{(r)}\underline{w}_r$ denklemleri matris formunda

$$\Sigma W = W\Lambda \qquad \text{(I.16)}$$

olarak yazılabilir ve $r\neq s$ olmak üzere $\underline{w}_r'\underline{w}_r=1$ ve $\underline{w}_r'\underline{w}_s=0$ denklemleri

$$W'W = I \qquad \text{(I.17)}$$

olarak yazılabilir. (I.16) ve (I.17) den

$$W'\Sigma W = \Lambda \qquad \text{(I.18)}$$

elde edilir.

$$|\Sigma-\lambda I| = |W'|.|\Sigma-\lambda I|.|W| \qquad (I.19)$$

$$= |W'\Sigma W - \lambda W'W| = |\Lambda-\lambda I|$$

$$= \prod_{j=1}^{p} (\lambda_j-\lambda)$$

den dolayı bu ifadenin kökleri Λ nın köşegen elamanları olduğunu görürüz. $\lambda_{(1)}=\lambda_1, \lambda_{(2)}=\lambda_2, \ldots, \lambda_{(p)}=\lambda_p$ dir.

Şimdi, buraya kadar söylenenleri bir teoremle özetleyelim.

Teorem I.1 $\underline{X}$:pxl tesadüfi vektörü için $E(\underline{X})=\underline{0}$ ve $E(\underline{X}\underline{X}')=\Sigma$ olsun. Kovaryans matrisi

$$E(\underline{Y}\underline{Y}') = \Lambda = \begin{bmatrix} \lambda_1 & 0 \ldots 0 \\ 0 & \lambda_2 \ldots 0 \\ \vdots & \vdots \quad\; \vdots \\ 0 & 0 \ldots \lambda_p \end{bmatrix} \qquad (I.20)$$

ve $\lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_p$, olacak şekilde

$$\underline{Y} = W'\underline{X} \qquad (I.21)$$

ortogonal lineer dönüşümü vardır. W nin r. sütunu $\underline{w}_r$, $(\Sigma-\lambda_r I)\underline{w} = \underline{0}$'ı sağlar. $\underline{Y}$ nin r. bileşeni, $Y_r = \underline{w}_r'\underline{X}$ $Y_1, Y_2, \ldots, Y_{r-1}$ ile ilişkisiz bütün normlanmış lineer bileşimler içinde en büyük varyansa sahiptir.

$\underline{Y}$ vektörü, $\underline{X}$ in temel bileşenler vektörü olarak tanımlanır. $Y_1, Y_2, \ldots, Y_p$ dönüşümleri yapıldığında, Y_1 en büyük varyanslı katsayılar vektörü normlanmış $\underline{X}$ in lineer bileşimidir.[1]

Sonuç I.1. $\lambda_{r+1}=\ldots=\lambda_{r+m}=d$ (yani, d katlı bir özdeğer) olduğunu varsayalım. O zaman $(\Sigma-dI)$, (p-m) ranklıdır. Ayrıca, $W^{\star}=(\underline{w}_{r+1}, \ldots, \underline{w}_m)$ ortogonal bir matris ile sağdan çarpılması dışında tek olarak belirlenir.[1]

Bilindiği gibi tek değişkenli dağılımın varyansı σ^2 nin çok değişkenli karşılığı Σ kovaryans matrisidir. Başka bir çok değişkenli karşılığı ise, çok değişkenli dağılımın "genelleştirilmiş varyansı" denilen Σ nın determinantı, $|\Sigma|$ dır.

Teorem I.2 Temel bileşenler vektörünün genelleştirilmiş varyansı, orjinal vektörün genelleştirilmiş varyansıdır ve temel bileşenlerin varyansları toplamı orjinal değişkenlerin varyansları toplamına eşittir.

İspat: $\underline{X}$ tesadüfi vektör ve $\underline{Y}=W\underline{X}$ temel bileşenler vektörü olmak üzere, $E(\underline{Y})=\underline{0}$ ve $E(\underline{Y}\underline{Y}')=W\Sigma W'$ dir. $\underline{Y}$ nin genelleştirilmiş varyansı

$$|W\Sigma W'|=|W|.|\Sigma|.|W'|=|\Sigma|.|WW'|=|\Sigma| \qquad (I.22)$$

olmak üzere $\underline{X}$ in genelleştirilmiş varyansına eşittir. $\underline{Y}$ nin bileşenlerinin varyansları toplamı

$$\sum_{i=1}^{p} E(Y_i^2) = tr(W\Sigma W') = tr(\Sigma W'W) \qquad (I.23)$$

$$= tr(\Sigma I) = tr(\Sigma)$$

$$= \sum_{i=1}^{p} E(X_i^2)$$

olur.

TBA de küçük varyanslı bileşenlerin, toplam varyans üzerinde etkisi ihmal edilebilir, bundan dolayı bileşen sayısı p den küçük olabilir ($r \leq p$ gibi). $r<p$ iken ilk r temel bileşen $(\underline{w}_1'\underline{X},\ldots,\underline{w}_r'\underline{X})$, $\underline{X}$:pxl in yerine kullanıldığında bilgi kaybı olacaktır. Bu kaybın minumum olması istenir.

I.2. $\underline{X}:p\times 1 \sim N(\underline{0},\Sigma)$ DAĞILIMLI OLDUĞUNDA TEMEL BİLEŞENLERİN YORUMU

$\underline{X}:p\times 1 \sim N(\underline{0},\Sigma)$, rank $(\Sigma)=p$ olsun. $\underline{X}$ in olasılık yoğunluk fonksiyonu

$$f(\underline{x})=\frac{1}{(2\pi)^{\frac{1}{2}p}\,|\Sigma|^{\frac{1}{2}}}\,e^{-\frac{1}{2}\underline{x}'\Sigma^{-1}\underline{x}}\ ,\quad \underline{x}\in \mathbb{R}^p \qquad (I.24)$$

olmak üzere sabit yoğunluklu noktaların kümesi olan $\{\underline{x}\in \mathbb{R}^p\ ;\ \underline{x}'\Sigma^{-1}\underline{x}=c\ ,\ c>0\}$ elipsoidlerini göz önüne alalım. Bu elipsoidler aynı merkez ve temel (ana) eksenli iç içe elipsoidlerdir. Böyle bir elipsoidin temel eksenleri $g(\underline{x})=\underline{x}'\underline{x}\ ,\ \underline{x}\in\{\underline{x}\ :\ \underline{x}'\Sigma^{-1}\underline{x}=c\}$ fonksiyonunun durağan noktalarını belirleyen $\underline{x}$ vektörlerin doğrultularıdır. Lagrange çarpanları yöntemiyle durağan noktaları bulalım. Lagrange fonksiyonu

$$\phi(\underline{x}\ ,\ \lambda)=\underline{x}'\underline{x}-\lambda\underline{x}'\Sigma^{-1}\underline{x} \qquad (I.25)$$

olmak üzere

$$\frac{\partial\phi}{\partial\underline{x}}=2\underline{x}-2\lambda\Sigma^{-1}\underline{x}=0 \qquad (I.26)$$

veya

$$\underline{x}=\lambda\Sigma^{-1}\underline{x} \qquad (I.27)$$

elde edilir. Sağdan Σ ile çarpılırsa

$$\Sigma\underline{x}=\lambda\underline{x} \qquad (I.28)$$

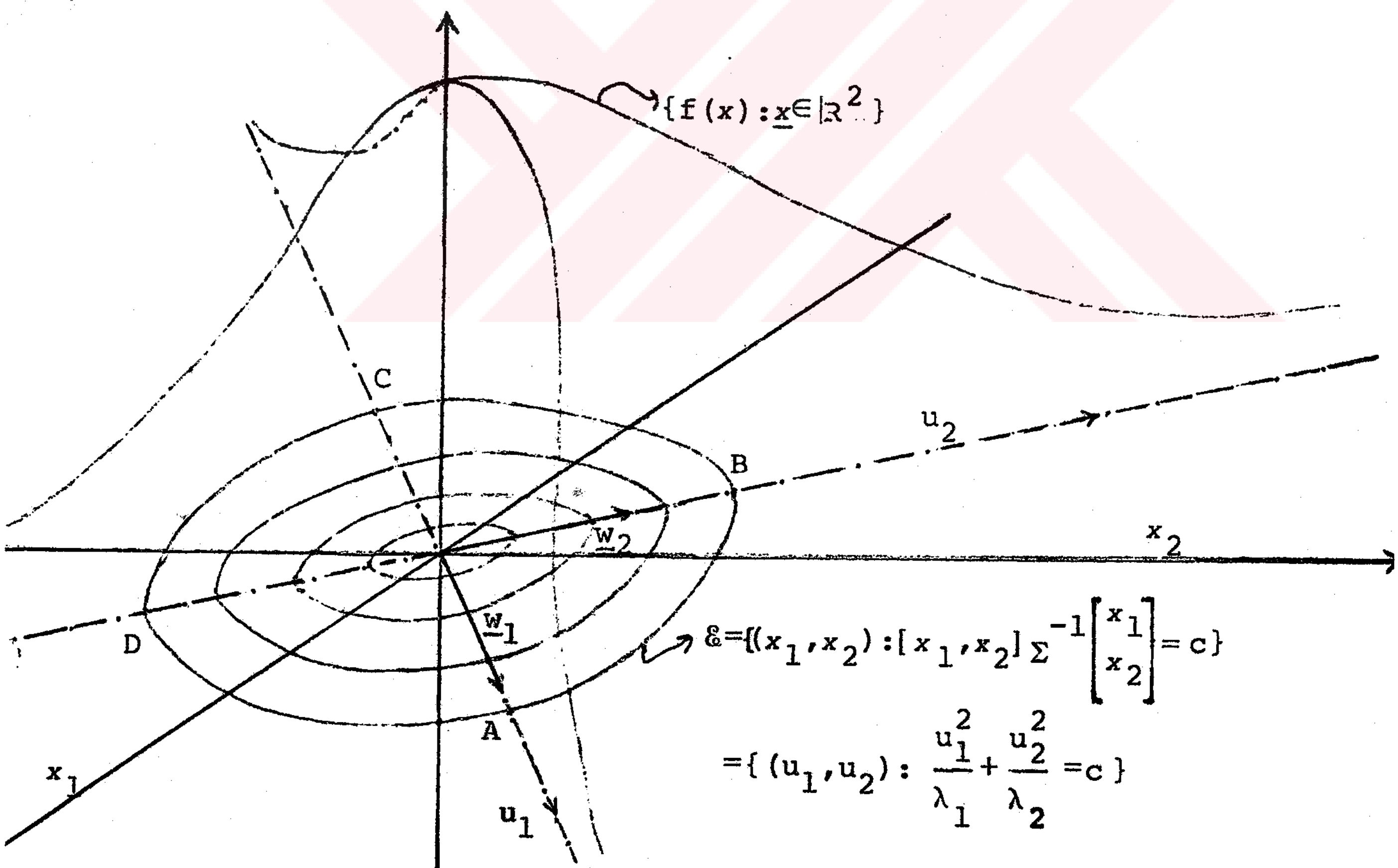
olur. Bu eşitlik (I.4) ile aynıdır ve aynı işlemler geliştirilebilir. Böylece $\underline{w}_1,\ldots,\underline{w}_p$ vektörleri elipsoidin temel (ana) eksenlerini verir. $\underline{u}=W\underline{x}$ dönüşümü koordinat eksenlerinin döndürülmesidir. Bu dönüşümle elde edilen yeni eksenler elipsoidin temel eksenleri yönündedir. Yeni koordinat sisteminde elipsoidin denklemi

$$\underline{u}'\Lambda^{-1}\underline{u} = \sum_{j=1}^{p} \frac{u_j^2}{\lambda_j} = c \qquad \text{(I.29)}$$

dir. Böylece temel eksenlerin uzunluğu $2\sqrt{\lambda_i c}$ dir.

Şimdi p boyutlu uzayda söylenenleri, p=2 olduğunda şekil olarak gösterelim.

$$\underline{X} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \sim N\left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix} \right\}, \ \text{rank}(\Sigma) = 2 \ \text{olsun.}$$



Şekil I.1

$\underline{X}$: 2x1 in olasılık yoğunluk fonksiyonu

$$f(\underline{x})=(2\pi)^{-1}|\Sigma_{2x2}|^{-\frac{1}{2}}\, e^{-\frac{1}{2}[x_1 x_2]\begin{bmatrix}\sigma_{11} & \sigma_{12}\\ \sigma_{21} & \sigma_{22}\end{bmatrix}^{-1}\begin{bmatrix}x_1\\ x_2\end{bmatrix}},\ \underline{x}\in\mathbb{R}^2$$

olmak üzere, yukarıdaki açıklamaların ışığında

$$\mathcal{E}=\{\underline{x}\in\mathbb{R}^2\ ;\ [x_1 x_2]\begin{bmatrix}\sigma_{11} & \sigma_{12}\\ \sigma_{21} & \sigma_{22}\end{bmatrix}^{-1}\begin{bmatrix}x_1\\ x_2\end{bmatrix}=c\}$$

ve

$$g(\underline{x})=x_1^2+x_2^2\ ,\ \underline{x}\in\mathcal{E}$$

olsun.

$\mathcal{E}$ elipsinin durağan noktaları şekil I.1 deki A, B, C ve D noktalarıdır. Bu durumda sözü geçen elipsoidler birer elips olmak üzere eksenleri AC ve BD doğruları üzerindedir. Bu doğrular aynı zamanda Σ nın özvektörlerinin doğrultularıdır.

I.3. TEMEL BİLEŞENLERİN EN ÇOK OLABİLİRLİK KESTİRİCİLERİ

$N(\underline{\mu},\Sigma)$ dağılımından N birimlik tesadüfi bir örnek çekildiğini varsayalım. S örnek kovaryans matrisi

$$S=\frac{1}{N-1}\sum_{i=1}^{N}(\underline{x}_i-\overline{\underline{x}})(\underline{x}_i-\overline{\underline{x}})' \qquad (I.30)$$

olmak üzere, S'in özdeğerleri $L_1>L_2>\ldots>L_p$ ve bunlara gelen özvektörler $\underline{V}_1,\ldots,\underline{V}_p$ olup, $(S-L_jI)\underline{V}_j$ ve $\underline{V}_j'\underline{V}_j=1$ dir.

Teorem I.3 $\underline{w}_1,\ldots,\underline{w}_p$, Σ nın farklı $\lambda_1>\lambda_2>\ldots>\lambda_p$ gibi özdeğerlerine karşılık gelen özvektörler olsun. Benzer şekilde, $\underline{V}_1,\ldots,\underline{V}_p$ ve $L_1>L_2>\ldots>L_p$ sırasıyla, Σ nın en çok olabilirlik kestiricisi S nin, özvektörleri ve özdeğerleri olsun. O zaman $\underline{w}_j$ lerin en çok olabilirlik kestiricileri $\underline{V}_j$ ler ve λ_j lerin en çok olabilirlik kestiricileri L_j ler $(j=1,\ldots,p)$ dir.[2]

Teorem I.4 $\lambda_1,\ldots,\lambda_r$, Σ nın $q_1,\ldots,q_r$ katlı özdeğerleri ise, o zaman λ_k nın en çok olabilirlik kestiricisi

$$\hat{\lambda}_k=\frac{1}{q_k}\,\frac{n}{N}\sum_{j=q_{(k-1)}+1}^{q_k} L_j \quad , \; k=1,\ldots,r \text{ ve } n=N-1 \qquad (I.31)$$

dir.[2]

I.4. TEMEL BİLEŞENLERİN ÖRNEK ÖZELLİKLERİ

Σ kovaryans matrisinin farklı özdeğerleri $\lambda_1>\lambda_2>\ldots>\lambda_p>0$ ve bunlara karşılık gelen özvektörler, $\underline{w}_1,\underline{w}_2,\ldots,\underline{w}_p$ olsun. N birimlik tesadüfi örneğimiz, $\underline{X}_1,\underline{X}_2,\ldots,\underline{X}_N$ olmak üzere, örnek kovaryans matrisi $n=N-1$ serbestlik derecesinde S olsun. S nin özdeğerleri, $L_1>L_2>\ldots>L_p>0$ ve bunlara karşılık gelen özvektörler ise, $\underline{V}_1,\underline{V}_2,\ldots,\underline{V}_p$ olsun.

Girshick ve Anderson "n" çok büyük olduğunda aşağıdaki sonuçların sağlandıklarını göstermişlerdir.[3]

(1) L_j, ona karşılık gelen $\underline{V}_j$ vektörünün elamanlarından bağımsız dağılmıştır.

(2) $\lim_{n\to\infty} \sqrt{n}(L_j-\lambda_j) \sim N(0, 2\lambda_j^2)$ dağılımına sahiptir.

(3) $\sqrt{n}(\underline{V}_j-\underline{w}_j)$, sıfır ortalama vektörlü ve kovaryans matrisi

$$\lambda_j \sum_{h=1}^{p} \frac{\lambda_h}{(\lambda_h-\lambda_j)^2} \underline{w}_h \underline{w}_h' \qquad (I.32)$$

olan çoklu normal dağılıma göre dağılmıştır.

(4) $\underline{V}_i$ nin r. elamanıyla $\underline{V}_j$ nin s. elamanları arasındaki kovaryans

$$-\frac{\lambda_i \lambda_j}{n(\lambda_i-\lambda_j)^2} w_{si} w_{rj}, \quad i \neq j \qquad (I.33)$$

dir.

Şimdi yukarıdaki sonuçların ışığında yığın özdeğerlerine ve özvektörlerine ilişkin hipotez testlerini ve güven aralıklarını verelim.

Yığın kovaryans matrisinin j. özdeğerine ilişkin güven aralığı 2. sonuçtan çıkar. "n" büyük olduğunda

$$\sqrt{\frac{n}{2}}(L_j-\lambda_j)/\lambda_j \sim N(0,1)$$

dir. Buradan

$$P\left(\frac{n}{2\lambda_j^2}(L_j-\lambda_j)^2 \leq z_{\alpha/2}^2\right) = \alpha \qquad (I.34)$$

$$P\left(\frac{L_j}{1+z_{\alpha/2}\sqrt{\frac{2}{n}}} \leq \lambda_j \leq \frac{L_j}{1-z_{\alpha/2}\sqrt{\frac{2}{n}}}\right) = 1-\alpha \qquad (I.35)$$

dir.

Σ nın r tane orta büyüklükteki özdeğerlerinin eşitliğini test etmek istediğimizde hipotez testi

$$H_o: \lambda_{q+1}=\ldots=\lambda_{q+r} \qquad (I.36)$$

H_1: Orta büyüklükte bazı özdeğerler farklıdır

biçimindedir. Burada q tane büyük özdeğer p-q-r tane küçük özdeğer vardır. Test istatistiğimiz ise

$$\chi^2= -n \sum_{j=q+1}^{q+r} \ln L_j + nr\ln \frac{\sum_{j=q+1}^{q+r} L_j}{r} \qquad (I.37)$$

dir. (I.36) nolu hipotez doğru olduğunda istatistik büyük n'ler için $\frac{1}{2}r(r+1)-1$ serbestlik derecesiyle χ^2(kikare) dağılımına sahiptir.[3]

Σ nın farklı λ_j özdeğerlerine karşılık gelen özvektörlerin bazı özel $\underline{w}_{jo}$ özvektörlerine eşit olup olmadığı test edilebilir. Asimptotik hipoz testi,

$$H_o: \underline{w}_j=\underline{w}_{jo} \qquad (I.38)$$

$$H_1: \underline{w}_j\neq\underline{w}_{jo}$$

olsun. 3. sonuç ve onun kovaryans matrisi (I.32) den test istatistiği

$$\chi^2=n(L_j\underline{w}'_{jo}S^{-1}\underline{w}_{jo}+\frac{1}{L_j}\underline{w}'_{jo}S\,\underline{w}_{jo}-2) \qquad (I.39)$$

olup, p-1 serbestlik derecesine sahiptir.[3]

Şimdi, alınmayan bileşenlere karşılık gelen özdeğerlerin (küçük özdeğerlerin) yığın değerlerinin sıfır olup olmadığını test etmeye çalışalım. Burada değişken değerleri standartlaştırıldığında S=R olduğu göz önüne alınmalıdır.

Bartlett tarafından yapılan bu testte geçmeden önce, bu test için kullanılacak artık veya hata (residual or error) matrisi kavramına değinelim. Söz gelimi j. (ya da birinci) özdeğer için

$$R_j \underline{V}_j = L_j \underline{V}_j \tag{I.40}$$

veya

$$R_j = L_j \underline{V}_j \underline{V}_j' \tag{I.41}$$

olsun. Burada R_j, mümkün olan p sayıdaki bileşenden sadece j. (söz gelimi birinci) si alınmış olsa idi, bu bileşenle R nin elde edilen R_j(veya R_1) lik kısmıdır. Birinci bileşenin açıklamadığı artık (veya hata) $R_{A(1)}$ olarak gösterilirse,

$$R_{A(1)} = R - R_1 \tag{I.42}$$

dir. Daha sonra ikinci bir bileşenin alınması halinde ise

$$R_{A(1.+2.)} = R - (R_1 + R_2)$$

olur. $R = R_1 + R_2 + \ldots + R_p$ olduğu göz önüne alınırsa, "r" sayıda bileşen alındığında artık matris

$$R_{A(1.+2.+\ldots+r.)} = R - \sum_{j=1}^{r} R_j = \sum_{j=r+1}^{p} R_j \tag{I.43}$$

olur.[4]

"r" bileşenin alınmasından sonra geriye kalan artık matrisin, yığın değeri sıfır olan bir matristen gelip gelmediğinin testi, artık matris için

$$H_o: P = \underline{0} \tag{I.44}$$

$$H_1: P \neq \underline{0}$$

hipotezi olup, artık matris elamanları istatistiksel olarak sıfırdan farklı olmadığını ifade eder. Yani Bartlett testi R (veya S) matrisinin p-r sayıdaki öz-değerler (küçük özdeğerler)inin yığın değerlerinin sıfır olduğunu test ediyor demektir.

Serbestlik derecesi (p-r)(p-r-1)/2 olan Bartlett tarafından yaklaşık χ^2(ki-kare) dağılımına sahip olan istatistik ise

$$\chi^2 = -[(N-1)-\frac{1}{6}(2p+5)-\frac{2}{3}r]\log_e Q_{p-r} \qquad (I.45)$$

burada

$$Q = \frac{|R|}{\{\prod_{j=1}^{r} L_j[(p-\sum_{j=1}^{r} L_j)/(p-r)]^{p-r}\}}$$

dir. Burada, N birim sayısı, p değişken (veya bileşen) sayısı, "r" alınan bileşen sayısını göstermektedir.[5]

I.5. BİRİNCİ TEMEL BİLEŞENİN YORUMU

$\underline{X}$:pxl $\sim$ ($\underline{0}$,Σ) dağılımından N birimlik tesadüfi bir örnek çekildiğini varsayalım. Gözlemlerin analiz edilecek verilerini içeren Nxp tipindeki veri matrisi

$$X = \begin{bmatrix} x_{11} & x_{12} \cdots & x_{1j} \cdots & x_{1p} \\ x_{21} & x_{22} \cdots & x_{2j} \cdots & x_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ x_{k1} & x_{k2} \cdots & x_{kj} \cdots & x_{kp} \\ \vdots & \vdots & \vdots & \vdots \\ x_{N1} & x_{N2} \cdots & x_{Nj} \cdots & x_{Np} \end{bmatrix}$$

olsun. Bu matrisin x_{kj} elamanı, k. bireyin j. değişken üzerindeki aldığı nicel değeri ifade eder. Burada gözlemler ortalamalardan sapmalar biçiminde olsun. Bu verilere ait örnek kovaryans matrisi S nin farklı özdeğerleri $\ell_1, \ell_2, \ldots, \ell_p$ ve bunlara karşılık gelen özvektörleri $\underline{v}_1, \ldots, \underline{v}_p$ olsun. Şimdi $\mathbb{R}^p$ de bulunan $(x_{k1}, x_{k2}, \ldots, x_{kp}) = \underline{x}^k$, $k=1,2,\ldots,N$ noktaların kendisine dik uzaklıkların kareleri toplamını minumum yapan D doğrusu, S nin en büyük özdeğeri ℓ_1 ve özvektörü $\underline{v}_1$ e karşılık gelebileceğini göstermeye çalışacağız.

$\mathbb{R}^p$'de, D bir doğru ve $\underline{x}$ bir nokta olmak üzere $d(\underline{x}, D)$ ifadesi $\underline{x}$ in D ye olan ortogonal uzaklığını göstersin. $\mathbb{R}^p$ de verilen noktalar $\underline{x}^1, \ldots, \underline{x}^N$ olmak üzere bunların dik uzaklıklarının kareleri toplamı olan

$$T = \sum_{k=1}^{N} d(\underline{x}^k, D)^2 \qquad (I.46)$$

yi minumum yapacak D doğrusunu arıyoruz.

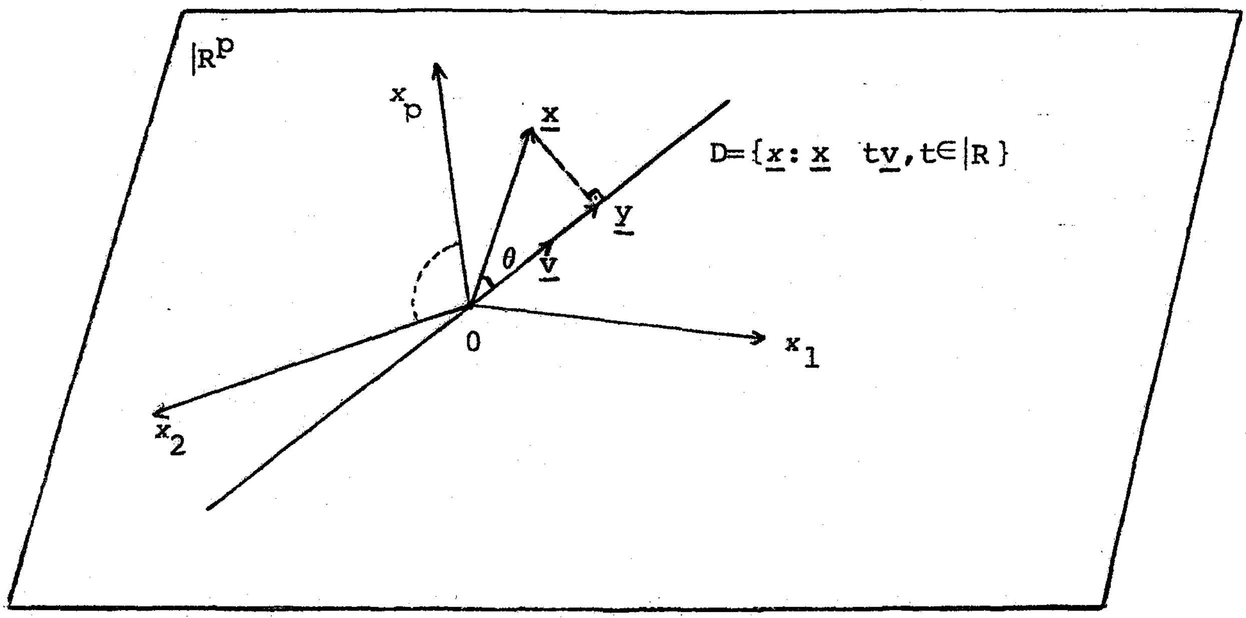
Öncelikle aranan doğrunun noktaların ağırlık merkezinden geçmesi gerektiğini gösterelim. O halde $\sum_{k=1}^{N} \underline{x}^k = \underline{0}$ durumu merkezileştirilmiş değerleri ifade etmek üzere; $\|\underline{v}\| = 1$ koşulu altında,

$$D = \{\underline{x} : \underline{x} = t\underline{v} \ , \ t \in \mathbb{R}\} \qquad (I.47)$$

doğrusu orjinden geçen bir doğru olsun. $\mathbb{R}^p$ deki her $\underline{x}$ için $[d(\underline{x}, D)]^2$ hesaplanacaktır. Şekil I.2 den, $\underline{y}$ vektörü $\underline{x}$ in D üzerindeki izdüşümü olmak üzere

$$\begin{aligned} \|\underline{y}\| &= \|\underline{x}\| . \cos\theta \\ &= \|\underline{x}\| \, \|\underline{v}\| \cos\theta \\ &= \langle \underline{x}, \underline{v} \rangle \end{aligned}$$

dır.



Şekil I.2

Pisagor Teoremi'nden, şekil I.2 göz önüne alındığında

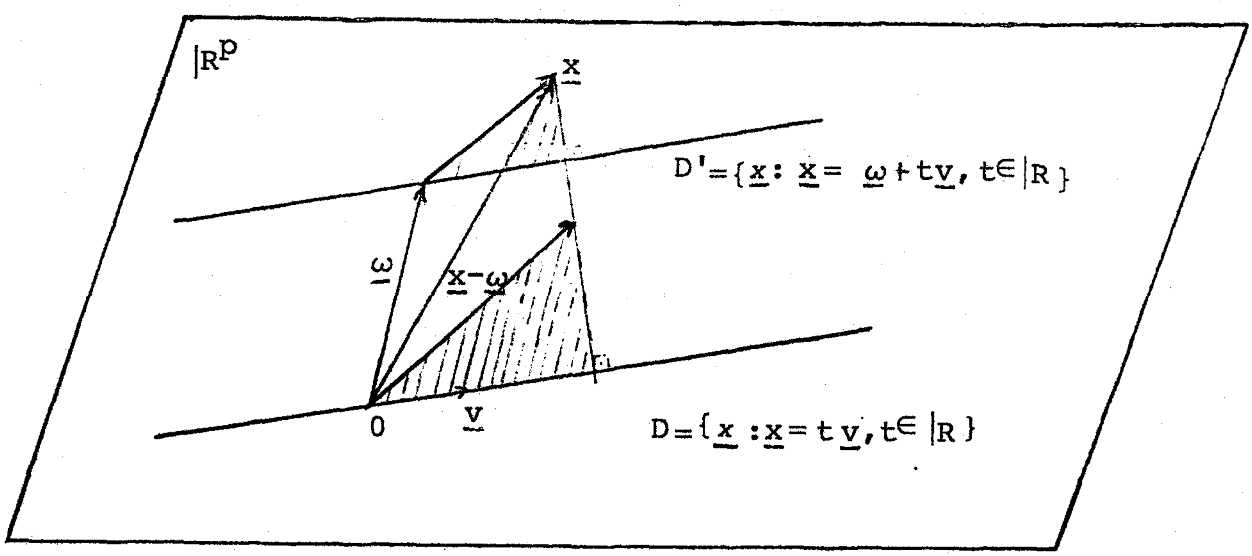
$$d(\underline{x},D)^2=\|\underline{x}\|^2-<\underline{x},\underline{v}>^2 \qquad (I.48)$$

olacaktır.

Şimdi $\mathbb{R}^p$ de orjinden geçmeyen D ye paralel bir D' doğrusunu gözönüne alalım. $\underline{\omega}$:pxl tipinde bir vektör olmak üzere

$$D'=\{\underline{x}:\underline{x}=\underline{\omega}+t\underline{v},\ t\in\mathbb{R}\} \qquad (I.49)$$

olsun. Aradığımız uzaklık şekil I.3 de görüldüğü gibi,



Şekil I.3

$$d(\underline{x},D')= d(\underline{x}-\underline{\omega},D) \qquad (I.50)$$

dır.[6] Öyleyse,

$$d(\underline{x},D')^2= d(\underline{x}-\underline{\omega}, D)^2=\|\underline{x}-\underline{\omega}\|^2-<(\underline{x}-\underline{\omega}),\underline{v}>^2 \qquad (I.51)$$

dır. Buradan

$$\begin{aligned} d(\underline{x}-\underline{\omega},D)^2 &=<(\underline{x}-\underline{\omega}),(\underline{x}-\underline{\omega})>-[<\underline{x},\underline{v}>-<\underline{\omega},\underline{v}>]^2 \\ &=<\underline{x},\underline{x}>-<\underline{x}-\underline{\omega}>+<\underline{\omega},\underline{\omega}>-<\underline{\omega},\underline{x}>-[<\underline{x},\underline{v}>^2 \\ &\qquad +<\underline{\omega},\underline{v}>^2-2<\underline{x},\underline{v}><\underline{\omega},\underline{v}>] \\ &=\|\underline{x}\|^2-<\underline{x},\underline{v}>^2+\|\underline{\omega}\|^2-<\underline{\omega},\underline{v}>^2-2<\underline{x},\underline{\omega}>+2<\underline{x},\underline{v}><\underline{\omega},\underline{v}> \\ &=d(\underline{x},D)^2+d(\underline{\omega},D)^2-2<\underline{x},[\underline{\omega}-<\underline{\omega},\underline{v}>\underline{v}]> \qquad (I.52) \end{aligned}$$

elde edilir. Burada, $\underline{\omega}-<\underline{\omega},\underline{v}>\underline{v}=\underline{a}$ dersek,

$$\begin{aligned} \sum_{k=1}^{N} d(\underline{x}^k,D')^2 &= \sum_{k=1}^{N} d(\underline{x}^k,D)^2+Nd(\underline{\omega},D)^2-2\sum_{k=1}^{N}<\underline{x}^k,\underline{a}> \\ &= \sum_{k=1}^{N} d(\underline{x}^k,D)^2+Nd(\underline{\omega},D)^2-2<(\sum_{k=1}^{N}\underline{x}^k),\underline{a}> \end{aligned}$$

yazılıp, $\sum_{k=1}^{N}\underline{x}^k=\underline{0}$ olduğu hatırlanırsa

$$\sum_{k=1}^{N} d(\underline{x}^k,D')^2= \sum_{k=1}^{N} d(\underline{x}^k,D)^2+Nd(\underline{\omega},D)^2 \qquad (I.53)$$

olur. Böylece

$$\sum_{k=1}^{N} d(\underline{x}^k,D')^2\geq \sum_{k=1}^{N} d(\underline{x}^k,D)^2 \qquad (I.54)$$

dır. Yani istenilen doğrunun ağırlık merkezinden geçmesi gerektiğini göstermiş olduk.

D doğrusunun belirlenmesi problemi, D nin doğrultusunun, yani $\underline{v}$ vektörünün belirlenmesine indirgenmiş oldu. $\|\underline{v}\| = 1$ kısıtlaması altında (I.46) daki optimizasyon problemini göz önüne alalım.

$$T = \sum_{k=1}^{N} d(\underline{x}^k, D)^2 = \sum_{k=1}^{N} \|\underline{x}^k\| - \sum_{k=1}^{N} <\underline{x}^k, \underline{v}>^2 \tag{I.55}$$

yi minimum etmek demek, $\|\underline{v}\| = 1$ kısıtı altında

$$f(\underline{v}) = \sum_{k=1}^{N} <\underline{x}^k, \underline{v}>^2 \tag{I.56}$$

yı maksimum etmek demektir. $C = X'X$ olmak üzere

$$f(\underline{v}) = \|X\underline{v}\| = \underline{v}'C\underline{v} \tag{I.57}$$

olarak yazılır.[7] Böylece f bir karesel (quadratic) formdur.

Şimdi simetrik bir matris olan C yi köşegenleştirelim. L = köşg. $(\ell_1, \ell_2, \ldots, \ell_p)$ ve V ortogonal matris olmak üzere

$$L = V'CV \tag{I.58}$$

olup, L nin gerçel değerleri için $\ell_1 > \ell_2 > \ldots > \ell_p > 0$ olduğu bilinmektedir.[8] Şimdi L ye karşılık gelen karesel form

$$g(\underline{u}) = \underline{u}'L\underline{u} = \sum_{j=1}^{p} \ell_j^2 u_{jj}^2 \tag{I.59}$$

olur. Eşitlik (I.58)'den

$$C = VLV'$$

$$\underline{v}'C\underline{v} = \underline{u}'L\underline{u}$$

$$\underline{v}'V\mathsf{L}V'\underline{v}=\underline{u}'\mathsf{L}\underline{u}$$

$$f(\underline{v})=g(V'\underline{v}) \tag{I.60}$$

bulunur. Bu yüzden

$$\text{mak.}\{f(\underline{v}):\|\underline{v}\|=1\}=\text{mak.}\{g(\underline{u}):\|\underline{u}\|=1,\ \underline{u}=V'\underline{v}\} \tag{I.61}$$

olur. Öyleyse

$$\text{mak.}\{g(\underline{u})=\sum_{j=1}^{p}\ell_j^2 u_{jj}^2=\underline{u}'\mathsf{L}\underline{u}\ ;\underline{u}'\underline{u}=\|\underline{u}\|=\sum^{p} u_{jj}^2=1\}$$

problemi için Lagrange fonksiyonu kullanılarak

$$\phi=\underline{u}'\mathsf{L}\underline{u}-\ell_j(\underline{u}'\underline{u}-1)$$

ve

$$\frac{\partial\phi}{\partial\underline{u}}=2\mathsf{L}\underline{u}-2\ell_j\underline{u}=\underline{0}$$

olup,

$$2\mathsf{L}\underline{u}-2\ell_j\underline{u}=\underline{0}$$

$$(\mathsf{L}-\ell_j)\,\underline{u}=\underline{0} \tag{I.62}$$

veya

$$\mathsf{L}\underline{u}=\ell_j\underline{u} \tag{I.63}$$

olur. Görüldüğü gibi (I.62) veya (I.63) eşitliklerine göre ℓ_j , L nin özdeğeri ve $\underline{u}$ da özvektörüdür. Bu formda optimum olan $g(\underline{u})$, en büyük özdeğer ℓ_1 ve buna karşılık gelen

$$\underset{p\times p}{\begin{bmatrix}\ell_1 & 0\ldots\ldots 0\\ 0 & \ell_2\ldots 0\\ \vdots & \vdots \quad\quad \vdots\\ 0 & 0\ldots\ldots\ell_p\end{bmatrix}}\underset{p\times 1}{\begin{bmatrix}1\\0\\\vdots\\0\end{bmatrix}}=\ell_1\underset{p\times 1}{\begin{bmatrix}1\\0\\\vdots\\0\end{bmatrix}};\quad \underline{u}=\underset{p\times 1}{\begin{bmatrix}1\\0\\\vdots\\0\end{bmatrix}}=\underline{e}_1 \tag{I.64}$$

ilk standart temel vektör ile maksimum değere ulaşır. Bundan dolayıdır ki, f nin maksimum değeri,

$$\underline{v} = V\underline{u} = \begin{bmatrix} v_{11} & v_{12} \cdots & v_{1p} \\ v_{21} & v_{22} \cdots & v_{2p} \\ \vdots & \vdots & \vdots \\ v_{p1} & v_{p2} \cdots & v_{pp} \end{bmatrix}_{pxp} \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}_{px1} = \begin{bmatrix} v_{11} \\ v_{21} \\ \vdots \\ v_{p1} \end{bmatrix}_{px1} \qquad (I.65)$$

olduğu zaman görülür. Bu da V'nin ilk sütunu ve ℓ_1 e karşılık gelen özvektörü $\underline{v} = V\underline{e} = V\underline{u}$ dır. O halde bizim aradığımız doğrunun doğrultusu X'X matrisinin en büyük özdeğerine karşılık gelen özvektörün belirlediği doğrultudur.

Sonuçta bulduğumuz en büyük özdeğer ve ona karşılık gelen özvektör, daha önce söz edilen S nin en büyük özdeğeri ℓ_1 ve buna karşılık gelen $\underline{v}_1$ özvektörüdür. Burada dikkat edilirse, ℓ_1 birinci temel bileşenin varyansı, $\underline{v}_p$ ise birinci temel bileşene ait lineer bileşinin katsayılar vektörü olduğu görülür.

TB için buraya kadar yapılan çalışmalardan sonra, karşılaşılan güçlüklerden biri TB in yorumlanmasıdır. Şimdi, bu konuyla ilgili yapılan işlemleri, pratikte TB in yorumu başlığı altında inceleyeceğiz.

I.6. PRATİKTE TEMEL BİLEŞENLERİN YORUMU

Pratikte TB in yorumu, TB in değişkenlerle olan korelasyon katsayıları yardımıyla adlandırılmaları ve birimlerin TB itibariyle büyüklüklerine göre sıralanması olarak göreceğiz.

Şimdi, değişkenler ile bileşenler arasındaki korelasyon katsayılarını gösteren yapı matrisini görelim.

$\underline{Z} = (Z_1,\ldots,Z_p)' \sim (\underline{0},\Sigma)$, burada

$$\Sigma = \begin{bmatrix} 1 & \rho_{12} \cdots \rho_{1p} \\ & 1 \cdots\cdots \rho_{2p} \\ & \vdots \\ & 1 \end{bmatrix} \qquad (I.66)$$

olup, $\underline{Z}$ standartlaştırılmış değişkenlerin vektörü ve bunlardan elde edilen temel bileşenler vektörü $\underline{Y}$ olsun. Bileşenlerin ortalaması, $E(Y_j) = E(\underline{w}_j'\underline{Z}) = \underline{w}_j'E(\underline{Z}) = 0$ ve standart sapması, $\sqrt{var(Y_j)} = \sqrt{\lambda_j}$ olmak üzere

$$Y_j^{\star} = \frac{(Y_j - E(Y_j))}{\sqrt{var(Y_j)}} = \frac{Y_j}{\sqrt{\lambda_j}}, \quad j=1,2,\ldots,p \qquad (I.67)$$

değişken standartlaştırılmış j. temel bileşendir. Bu durum p tane bileşen için

$$\begin{bmatrix} Y_1^{\star} \\ Y_2^{\star} \\ \vdots \\ Y_p^{\star} \end{bmatrix} = \begin{bmatrix} 1/\sqrt{\lambda_1} & 0 \cdots\cdots\cdots & 0 \\ 0 & 1/\sqrt{\lambda_2} \cdots & 0 \\ \vdots & \vdots & \vdots \\ 0 & 0 \cdots\cdots\cdots & 1/\sqrt{\lambda_p} \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_p \end{bmatrix}$$

olup, matris gösterimiyle

$$\underline{Y}^{\star} = \Lambda^{-\frac{1}{2}}\underline{Y}$$

$$= \Lambda^{-\frac{1}{2}} W'\underline{Z} \qquad (I.68$$

biçimini alır. Burada $\underline{Z} = W\Lambda^{\frac{1}{2}}\underline{Y}^{\star}$ yazılır. Bu dönüşümdeki

$\mathring{A} = W\Lambda^{\frac{1}{2}}$ matrisine yapı matrisi denir. Bu matris için

$$\begin{aligned}
\mathring{A} &= W\Lambda^{\frac{1}{2}} \\
&= W\Lambda^{\frac{1}{2}} W'W\Lambda^{\frac{1}{2}}\Lambda^{-\frac{1}{2}} \\
&= W\Lambda W'W\Lambda^{-\frac{1}{2}} \\
&= \Sigma W\Lambda^{-\frac{1}{2}} \\
&= E(\underline{Z}\underline{Z}')W\Lambda^{-\frac{1}{2}} \\
&= E(\underline{Z}\underline{Y}^{\star\prime})
\end{aligned}$$

$$= \begin{bmatrix} \rho_{Z_1Y_1^\star} & \rho_{Z_1Y_2^\star} \cdots\cdots & \rho_{Z_1Y_p^\star} \\ \rho_{Z_2Y_1^\star} & \rho_{Z_2Y_2^\star} \cdots\cdots & \rho_{Z_2Y_p^\star} \\ \vdots & \vdots & \vdots \\ \rho_{Z_pY_1^\star} & \rho_{Z_pY_2^\star} \cdots\cdots & \rho_{Z_pY_p^\star} \end{bmatrix} \qquad (I.69)$$

yazılır.[5]

$\mathring{A}$ matrisi TBA bakımından önemlidir. Çünkü, hangi değişkenlerin hangi bileşen ile ilişkili olduğunu göstererek değişkenlerin ortak niteliğini saptar. Bu da bileşenlerin tanımlanmasına ve onu adlandırmaya yardımcı olur. Burada değişkenliğin ortak niteliğini saptarken R korelasyon matrisine bakılmasına dikkat edilmelidir.[9]

Standartlaştırılmış herhangi bir değişken Z_j

$$Z_j = \rho_{Z_jY_1^\star} Y_1^\star + \rho_{Z_jY_2^\star} Y_2^\star + \ldots + \rho_{Z_jY_p^\star} Y_p^\star, j=1,2,..,p \quad (I.70)$$

olarak yazılabilir. Bu TBA modeli olup, p tane değişken için p tane bileşen alınarak, toplam varyansı açıklayacağı varsayılmaktadır.[4] Ayrıca yukarıda yazılan yapı matrisi, N birimlik tesadüfi bir örnekten tahmin edilmek istendiğinde, $A = VL^{1/2}$ ile bulunur. Bu durumda (I.70), örnekten bulunan değerler cinsinden $Z_j = r_{Z_jY_1^\star} Y_1^\star + r_{Z_jY_2^\star} Y_2^\star + \ldots + r_{Z_jY_p^\star} Y_p^\star$ olarak yazılabilir.

Şimdi, N birim (gözlem)lik tesadüfi bir örnek için yapılan bir TBA uygulamasında bileşenlerin adlandırılmasından sonra, birim (gözlem)lerin bileşenler itibariyle değerleri bulunarak, büyüklüklerine göre sıralanması istenebilir.[10] Öyleyse standartlaştırılmış gözlemler matrisi

$$Z = \begin{bmatrix} Z_{11} & Z_{12} \ldots & Z_{1p} \\ Z_{21} & Z_{22} \ldots & Z_{2p} \\ \vdots & \vdots & \vdots \\ Z_{N1} & Z_{N2} \ldots & Z_{Np} \end{bmatrix} \qquad (I.71)$$

ve

$$Z'Z = VLV' \qquad (I.72)$$

olmak üzere temel bileşenlerin gözlem matrisi $Y = ZV$ olsun. Buradan temel bileşenlerin standartlaştırılmış gözlem matrisi

$$Y^\star = \begin{bmatrix} 1/\sqrt{\ell_1} & 0 \ldots\ldots & 0 \\ 0 & 1/\sqrt{\ell_2} \ldots & 0 \\ \vdots & \vdots & \vdots \\ 0 & 0 \ldots\ldots & 1/\sqrt{\ell_p} \end{bmatrix} Y$$

$$= L^{-\frac{1}{2}} ZV$$

$$= L^{-\frac{1}{2}} V'Z$$

$$= B'Z$$

$$= ZB \qquad (I.73)$$

olup, bu matrise aynı zamanda bileşen değerleri matrisi denir.[5]

Y^{*} matrisi bulunduktan sonra matrisin her bir sütun vektörü, her bir bileşenin gözlemler itibariyle aldıkları değerleri verir. Bu vektörlerin elamanları büyükten küçüğe doğru sıralanarak yorumlamaya yardımcı olduğundan TBA bakımından önemlidir.

TBA uygulamalarında karşılaşılan zorluklardan biri de ölçüm birimlerinin etkileridir. S matrisinde tüm veya bazı değişkenlerin sonuçları, ölçüm birimleri değişkenliği içerir. Bunun sonucunu, özdeğerler ve özvektörlerde gözlemek zordur. Bununla birlikte önemli ve önemsiz değişkenlerin rolu değişir, yani birbirlerinin yerini alabilirler. Böylece küçük varyanslı olan, büyük varyanslı durumuna geçebilir. Bundan dolayı temel bileşenler analizinde yalnızca, tüm değişkenler aynı birimle ölçüldüğünde veya en azından karşılaştırılabilir birimlerle ölçüldüğünde kullanılabilir. Bu zorluktan kurtulmanın yolu, TB'e karşılık gelen korelasyon matrisini dikkate almaktır.[2]

Şimdi, TBA için söylenenleri bir örnek üzerinde gösterelim.

Örnek I.1 Toprağın yapısında bulunan;

X_1: Çamur % si

X_2: Kil % si

X_3: Organik madde % si

X_4: pH (asitlik ve bazlık derecesi)

değişkenler olsun.[11] Bu değişkenleri içeren, değişkenliğin büyük bir kısmını açıklayan yeni bileşenler bulmaya çalışacağız. Ayrıca bileşenlerle ilgili diğer özelliklerde incelenmeye çalışılacaktır. Bunun için inceleme alanı olan topraktan 20 numune alınmış olup, bu numuneler üzerinden değişkenlere ait veriler aşağıdadır.

Numuneler	X_1	X_2	X_3	X_4
1	13.0	9.7	1.5	6.4
2	10.0	7.5	1.5	6.5
3	20.6	12.5	2.3	7.0
4	33.8	19.0	2.8	5.8
5	20.5	14.2	1.9	6.9
6	10.0	6.7	2.2	7.0
7	12.7	5.7	2.9	6.7
8	36.5	15.7	2.3	7.2
9	27.1	14.3	2.1	7.2
10	25.5	12.9	1.9	7.3
11	26.5	14.9	2.4	6.7
12	22.3	8.4	4.0	7.0
13	30.8	7.4	2.7	6.4
14	25.3	7.0	4.8	7.3
15	31.2	11.6	2.4	6.5
16	22.7	10.1	3.3	6.2
17	31.2	9.6	2.4	6.0
18	13.2	6.6	2.0	5.8
19	11.1	6.7	2.2	7.2
20	20.7	9.6	3.1	5.9
ortalamalar	22.235	10.505	2.535	6.650
standart sapmalar	8.34369	3.71745	0.80212	0.51248

Korelasyon matrisi

$$R=\begin{bmatrix} 1. & 0.6743 & 0.2131 & 0.0242 \\ & 1. & -0.1960 & 0.0131 \\ & & 1. & 0.0787 \\ & & & 1. \end{bmatrix}$$

olur. R nin özdeğerler matrisi

$$L=\begin{bmatrix} 1.676 & 0. & 0. & 0. \\ & 1.146 & 0. & 0. \\ & & 0.960 & 0. \\ & & & 0.218 \end{bmatrix}$$

dır. ℓ_j lere bakıldığında

$$\ell_1 > \ell_2 > \ell_3 > \ell_4 > 0$$

olduğu görülür. Burada

$$\sum_{j=1}^{4} \ell_j = 1.676 + 1.146 + 0.960 + 0.218 = 4$$

ve

$$tr(R) = \sum_{j=1}^{p} \ell_j = 4$$

dır.

Ayrıca

$$|R| = \prod_{j=1}^{p} \ell_j = 0.402$$

dir. Bileşenler tarafından açıklanan varyans payları, sırasıyla

$$\frac{\ell_1}{p}, \ldots, \frac{\ell_4}{p} \text{ veya } \frac{\ell_1}{\sum_{j=1}^{4} \ell_j}, \ldots, \frac{\ell_4}{\sum_{j=1}^{4} \ell_j} \text{ den}$$

Özdeğerler	Açıklanan varyans payları %	Birikimli açıklanan pay %
1.676	41.90	41.90
1.146	28.65	70.55
0.960	24.00	94.55
0.218	5.45	100.00

olur. Özdeğerlerden sonra onlara karşılık gelen özvektörler matrisi

$$V=[\underline{v}_1,\underline{v}_2,\underline{v}_3,\underline{v}_4]=\begin{bmatrix} 0.710 & 0.702 & 0.025 & 0.042 \\ 0.182 & -0.241 & 0.836 & 0.459 \\ -0.147 & 0.111 & -0.423 & 0.887 \\ -0.664 & 0.661 & 0.349 & -0.026 \end{bmatrix}$$

olarak bulunur. Bu matris bir ortogonal matris olup $V'V=I$ ve $|V|=1$ koşullarını sağlar.

Şimdi, herbir bileşenle değişken arasındaki korelasyon katsayılarını içeren yapı matrisini bulalım. $A=VL^{1/2}$ olmak üzere

$$A=\begin{bmatrix} 0.919 & 0.752 & 0.025 & 0.020 \\ 0.236 & -0.258 & 0.819 & 0.214 \\ -0.190 & 0.119 & -0.415 & 0.414 \\ -0.860 & 0.708 & 0.342 & -0.012 \end{bmatrix}$$

A matrisinden anlaşılacağı gibi X_1(çamur) ve X_4(pH) birinci temel bileşen ile yakın bir ilişki halindedir. Ayrıca bu iki değişken birinci bileşen eksenine daha yakındır. Çünkü değişkenlerin bu bileşen ekseniyle yaptıkları korelasyon (açının cosinus)lar satırlardaki en büyük sayılardır. Aynı bileşen eksenine yakın olan değişkenler ortak nitelikli değişkenlerdir. Birinci bileşenin ortak niteliği, veri matrisindeki değişkenliğin % 41,90 ini açıklar. Birinci ve ikinci bileşen beraber % 70,55 lik kısmını açıklar. % 29.45 lik bilgi kaybıyla istenirse bu dört değişken yerine, onların ortak niteliklerini temsil eden birinci ve ikinci temel bileşenler alınabilir. Bu durumu birinci ve ikinci bileşenler için tablo halinde gösterirsek

	Bileşenler	
Değişkenler	1. Bileşen	2. Bileşen
X_1	0.919	0.752
X_2	0.236	-0.258
X_3	-0.190	0.119
X_4	-0.860	0.708

olur. Şayet bilgi kaybı amaca göre kabul edilebilir düzeyde ise (buna hipotez çalışmaları yardımcı olabilir) dört boyutlu uzay yerine iki boyutlu düzlem ele alınabilir.

Şimdi bulunan TB için örnek özelliklerini inceleyelim. Değişkenliğin en büyük kısmını açıklayan birinci temel bileşen (en büyük özdeğer) için % 95 lik güven aralığı oluşturalım. n = 19, $\ell_1 = 1.676$, $Z_{0.025}=1,96$ değerleri (I.35) eşitliğinde yerine konulursa

$$P\left(\frac{1.676}{1+1.96\sqrt{\frac{2}{19}}} \leq \lambda_1 \leq \frac{1.676}{1-1.96\sqrt{\frac{2}{19}}}\right) = 0.95$$

$$P(0.611 \leq \lambda_1 \leq 4.603) = 0.95$$

olur. İkinci büyük temel bileşen için %95'lik güven aralığını, $n = 19$, $\ell_2 = 1.146$, $Z_{0.025} = 1.96$ değerleri ile oluşturalım.

$$P\left(\frac{1.146}{1+1.96\sqrt{\frac{2}{19}}} \leq \lambda_2 \leq \frac{1.146}{1-1.96\sqrt{\frac{2}{19}}}\right) = 0.95$$

$$P(0.701 \leq \lambda_2 \leq 3.148) = 0.95$$

olur.

Şimdi alınan temel bileşenlerden sonra geride kalan bileşen (küçük özdeğer) lerin yığın değerlerinin sıfır olup olmadığını test edelim. Bunun için ilk önce birinci ve ikinci bileşen alındıktan sonraki artık matrisi bulalım.

$$R_{A(1.+2.)} = R-(R_1+R_2)$$

$$R_{1+2} = A_{4x2}\, A'_{2x4}$$

$$= \begin{bmatrix} 0.919 & 0.752 \\ 0.236 & -0.258 \\ -0.190 & 0.119 \\ -0.860 & 0.708 \end{bmatrix} \begin{bmatrix} 0.919 & 0.236 & -0.190 & -0.860 \\ 0.752 & -0.258 & 0.119 & 0.708 \end{bmatrix}$$

$$= \begin{bmatrix} 1.410 & 0.023 & -0.085 & -0.258 \\ 0.023 & 0.122 & -0.078 & -0.386 \\ -0.085 & -0.078 & 0.050 & 0.170 \\ -0.258 & -0.386 & 0.170 & 1.241 \end{bmatrix}$$

$$R_{A(1.+2.)} = \begin{bmatrix} 0.410 & 0.264 & 0.128 & -0.234 \\ 0.264 & 0.878 & -0.118 & -0.373 \\ 0.128 & -0.118 & 0.950 & -0.091 \\ -0.234 & -0.373 & -0.091 & -0.241 \end{bmatrix}$$

olur. Bu artık matrisin yığın değeri sıfır olan matristen gelip gelmediğini test edelim. Hipotez

H_o: $P = \underline{0}$

H_1: $P \neq \underline{0}$

ise $\alpha = 0.05$ anlamlılık düzeyinde 1 serbestlik derecesiyle (I.45) eşitliğine göre

$$Q = \frac{0.402}{\{(1.676)(1.146)[(4-1.676-1.146)/(4-2)]^{4-2}\}}$$

$$= 0.151$$

ve

$$\chi^2 = -[(20-1) - \frac{1}{6}(2(4)+5) - \frac{2}{3}\,2]\,\ell n 0.151$$

$$= 29.303$$

olur. Bulunan değer ile tablo değeri $\chi^2_{0.05;1} = 3.841$ karşılaştırıldığında H_o red edilir. Yani adı geçen verilerin artık matrisi, $\alpha = 0.05$ anlamlılık ve 1 serbestlik derecesinde, yığın değeri sıfır olan bir matristen gelmediği istatistiksel olarak görülmüştür.

Alınmayan bileşenler için bulunan artık matrisin yığın değeri sıfır olan bir matristen gelmediğini gördükten sonra, bu bileşenlerin (özdeğerlerin) birbirine eşit olup olmadığını test edelim. $n = 19$, $q = 2$ ve $r = 2$ olmak üzere

H_o: $\lambda_3 = \lambda_4$

H_1: $\lambda_3 \neq \lambda_4$

hipotezi, (I.37) eşitliği kullanılarak $\alpha = 0.05$ anlamlılık düzeyinde test edelim.

$$\chi^2 = -19(\ln 0.930 + \ln 0.218) + (19)(2)\ln \frac{0.930+0.218}{2}$$

$$= 9.225$$

değerini, tablo değeri $\chi^2_{0.05;2} = 5.99$ ile karşılaştırıldığında H_o red edilecektir. Yani adı geçen verilerin yığınına ait üçüncü ve dördüncü bileşenler (veya özdeğerler), $\alpha = 0.05$ anlamlılık düzeyinde eşit değildir.

BÖLÜM II

LİNEER MODELLERDE TEMEL BİLEŞENLER

TB'in istatistiksel veri analizindeki yerini incelerken; çok değişkenli istatistiksel analizdeki, TBA ni inceledikten sonra, lineer modellerdeki Temel Bileşenler Regresyonu (TBR)'nu, çoklubağlantı sorununun incelenmesinde göreceğiz. Bunun için ilk önce bazı tanımları verelim.

II.1. BAZI TANIMLAR

II.1.a. Lineer Model

Bu bölümde kullanılacak olan model, $\underline{Y}$:nx1 rasgele değişkenlerin gözlenebilir bir vektörü, X:nxp(n > p) bilinen rasgele olmayan sayıların matrisi $\underline{\beta}$:px1 bilinmeyen paremetrelerin vektörü, $\underline{\varepsilon}$:nx1 aynı dağılıma sahip $E(\underline{\varepsilon}) = 0$, $Cov\ (\underline{\varepsilon}) = \sigma^2 I$ koşulunu sağlayan ve gözlenemeyen rasgele değişkenlerin vektörü olmak üzere

$$\underline{Y} = X\underline{\beta} + \underline{\varepsilon} \qquad (II.1)$$

anlamındadır. Bu şekilde tanımlanan modele genel lineer model veya Gauss-Markov modeli denir.

(II.1) denklemiyle belirlenen modelde $\underline{\varepsilon} \sim N(\underline{0}, \sigma^2 I_n)$ dağılımlı olduğunda modele Genel Lineer Hipotez Modeli denir.[12]

Buna bağlı olarak,

$\underline{Y}$ nin gözlenen değeri $\underline{y}$ olmak üzere

$$\underline{y} - X\underline{\beta} = \underline{e}(\underline{\beta}) \qquad \text{(II.2)}$$

denklem sistemi için

$$\min. \| \underline{e}(\underline{\beta}) \|^2 = (\underline{y} - X\hat{\underline{\beta}})'(\underline{y} - X\hat{\underline{\beta}}) \qquad \text{(II.3)}$$

$$\underline{\beta}\,(\neq 0)$$

koşuluyla belirlenen $\hat{\underline{\beta}}$ vektörüne (II.2) sisteminin EKK çözümü denir.[12]

II.1.b. Çoklubağlantı (Multicollinearity)

p değişken sayısı olmak üzere, X'X matrisinin p ranklı olma varsayımının bozulması yani her bir gözlem için açıklayıcı değişkenler arasında bir ya da birden çok doğrusal bağıntının varlığı tam çoklubağlantı göstergesidir. Tam çoklubağlantıya bir sorun olarak bakılmaz. Güçlü çoklubağlantıda sütunların yaklaşık olarak sıfıra eşit olan lineer bileşimleri olarak karakterize edilir. Daha açık olarak, eğer

$$\sum_{q=1}^{p} w_q \underline{X}_q \approx 0 \qquad \text{(II.4)}$$

olacak şekilde sıfır olmayan $\underline{w}' = (w_1, w_2, \ldots, w_p)$ vektörü varsa, X in sütunları çoklubağlantılıdır denir.[13]

II.1.c. Bir Matrisin Koşul Sayısı

Simetrik bir A matrisi için koşul sayısı

$$k_2(A) = \frac{\underset{j}{\mathrm{Mak}} |\lambda_j|}{\underset{j}{\mathrm{Min}} |\lambda_j|} \qquad \text{(II.5)}$$

olarak tanımlanır.

Lineer modellerde TB'in yerini, çoklubağlantı sorununun incelenmesinde göreceğiz. Bunun için ilk önce özdeğerlerle çoklubağlantının belirlenmesi, daha sonra yanlı kestirimlerden TBR incelenecektir.

II.2. ÇOKLUBAĞLANTININ ÖZDEĞERLERLE BELİRLENMESİ

Çoklubağlantının doğal bir ölçütü X'X matrisinin koşul sayısı olan $k_2(X'X)$ değeridir. $k_2(X'X) < 10$ ise verilerde çok az bir çoklubağlantı vardır. Koşul sayısı 30'a çıkması orta derecede kuvvetli çoklubağlantının varlığını gösterir. $k_2(X'X) > 30$ olması, şiddetli bir çoklubağlantının derecesidir.[14]

Açıklayıcı değişkenler arasındaki çoklubağlantı en iyi şekilde korelasyon matrisinin özdeğerleri ve özvektörlerinin incelenmesiyle belirlenebilir.[14]

X'X matrisinde çoklubağlantı varsa λ_j özdeğerlerinden bir ya da bir kısmı sıfır ya da sıfıra yakın çıkar. X'X in j. özdeğeri λ_j ve buna karşılık gelen özvektör $\underline{V}_j$ olmak üzere

$$\lambda_j = \underline{V}_j' X'X\underline{V}_j = (X\underline{V}_j)'(X\underline{V}_j), \quad j = 1,2,\ldots,p \qquad \text{(II.6)}$$

dir. X'X in son r tane özdeğeri yeterince küçük ise

$$0 \approx (X\underline{V}_j)'(X\underline{V}_j) \Rightarrow X\underline{V}_j \approx 0, \quad j = 1,2,\ldots,r \qquad \text{(II.7)}$$

olacaktır. $\lambda_j \approx 0$ olduğu zaman bunlara karşılık gelen özvektörlerin mutlak değerce büyük olan elemanları çoklubağlantıda içerilen değişkenleri verecektir.[15]

Çoklubağlantıyı belirlemede izlenen bir diğer yol,

$$(X'X)^{-1} = \sum_{j=1}^{p} \lambda_j^{-1} \underline{V}_j \underline{V}_j' \qquad \text{(II.8)}$$

olduğundan $(X'X)^{-1}$ in j. köşegen elamanı, varyans şişirme oran (VIF)ı olarak adlandırılan

$$c_{jj} = \sum_{j=1}^{p} \lambda_j^{-1} V_{ji}^2 \quad , \; i = 1,2,\ldots,p \qquad \text{(II.9)}$$

olarak yazılabilir. Bu c_{jj}'lerden en büyüğünün 10'dan büyük olması zararlı çoklubağlantı olarak nitelendirilmiştir. [15]

Çoklubağlantı sorununun incelenmesinde kullanılan tekniklerden biri, yanlı bir kestirim olan TBR dur. Kendall'ın çoklubağlantı varlığında çok değişkenli TBA nin regresyon analizine dönüştürülebileceği savına dayanarak, EKK Kestiricisine göre yanlı ancak daha küçük AKO (Artık Kareler Ortalama)lı kestirimler TBR ile elde edilir. [15]

II.3. TEMEL BİLEŞENLER REGRESYONU

II.3.a. Genelleştirilmiş İnvers ve Kısıtlanmış EKKK Olarak TB Kestiricisi ve Özellikleri

X'X matrisinin özdeğerli $\lambda_1 > \lambda_2 > \ldots > \lambda_p$ ve bunların köşegen matrisi Λ , özdeğerlere karşılık gelen normlandırılmış özvektörler $\underline{V}_1, \underline{V}_2, \ldots, \underline{V}_p$ ve bunların matrisi V olsun.

V matrisi ile Λ matrisi aşağıdaki gibi parçalansın

$$V = [\underset{p\times r}{V_r} \; \vdots \; \underset{p\times(p-r)}{V_{p-r}}] \qquad \text{(II.10)}$$

$$\Lambda = \begin{bmatrix} \Lambda_r & 0 \\ 0 & \Lambda_{p-r} \end{bmatrix}_{p\times p}$$

Λ^+ matrisi

$$\Lambda^+ = \begin{bmatrix} \Lambda_r^{-1} & 0 \\ 0 & 0 \end{bmatrix} \qquad \text{(II.11)}$$

olmak üzere $(X'X)^+$ matrisi şu şekilde tanımlansın

$$(X'X)^+ = V\Lambda^+V' \qquad \text{(II.12)}$$

$(X'X)^+$ matrisine "belirlenmiş r rankı için X'X in genelleştirilmiş inversi" denilmektedir. [14] Burada

$$(X'X)^+ = V_r\Lambda_r^{-1}V_r' \qquad \text{(II.13)}$$

ve

$$\begin{aligned}(X'X)^+(X'X) &= (V_r\Lambda_r^{-1}V_r')(V_r\Lambda_rV_r') \\ &= V_rV_r' \\ &= I-V_{p-r}V_{p-r}' \end{aligned} \qquad \text{(II.14)}$$

dir.

Şimdi genelleştirilmiş invers kestiricisi tanımını yapalım;

$$\hat{\underline{\beta}}^+ = (X'X)^+X'\underline{Y} \qquad \text{(II.15)}$$

kestiricisine $\underline{\beta}$ parametre vektörünün genelleştirilmiş inver kestiricisi denir.

$r = p$ alındığında $(X'X)^+ = (X'X)^{-1}$ olup, $\hat{\underline{\beta}}^+ = \hat{\underline{\beta}}$ dır.

Teorem II.1 $\hat{\underline{\beta}}^+$ kestiricisi $\hat{\underline{\beta}}$ nın bir lineer dönüşümüdür.

İspat: Normal denklemlerde $X'\underline{Y} = X'X\hat{\underline{\beta}}$ olmak üzere (II.12) den

$$\hat{\underline{\beta}}^{+}=(X'X)^{+}X'X\hat{\underline{\beta}} \qquad \text{(II.16)}$$

dir. Buradaki $(X'X)^{+}X'X$ matrisi simetrik ve idenpotent olup, bir dik izdüşüm matrisidir.

$$(X'X)^{+}X'X = V_rV_r' \qquad \text{(II.17)}$$

dır. Burada $\hat{\underline{\beta}}^{+}$, $\hat{\underline{\beta}}$ nın V_r üzerine dik izdüşümü olan vektördür.

Şimdi $\underline{\beta}\in[V_r]$ kısıtlaması altında (II.1) modelini göz önüne aldım.

$$\underline{\beta}\in[V_r] \Leftrightarrow V_{p-r}'\underline{\beta}=0 \qquad \text{(II.18)}$$

olduğundan $V_{p-r}'\underline{\beta}=0$ kısıtlaması altında $\underline{\beta}$ nın kısıtlanmış EKK kestiricisi, $\tilde{\underline{\beta}}=\hat{\underline{\beta}}+S^{-1}R'(RS^{-1}R')^{-1}(\underline{r}-R\hat{\underline{\beta}})$ göz önüne alındığında

$$\begin{aligned}\tilde{\underline{\beta}}&=\hat{\underline{\beta}}+(X'X)^{-1}V_{p-r}(V_{p-r}'(X'X)^{-1}V_{p-r})^{-1}(-V_{p-r}\hat{\underline{\beta}})\\ &=(I-(X'X)^{-1}V_{p-r}\Lambda_{p-r}V_{p-r}')\hat{\underline{\beta}}\\ &=(I-V_{p-r}V_{p-r}')\hat{\underline{\beta}}\\ &=(X'X)^{+}X'X\hat{\underline{\beta}}\\ &=\hat{\underline{\beta}}^{+}\end{aligned} \qquad \text{(II.19)}$$

dır.

$\hat{\underline{\beta}}^{+}$ kestiricisi $\underline{\beta}$ nın $V_{p-r}'\underline{\beta}=0$ parametrik kısıtlaması altında EKKK (En Küçük Kareler Kestiricisi)dir.

Diğer taraftan (II.1) modelinde $Z = XV$ dönüşümü yapılarak

$$\begin{aligned}\underline{Y} &= XVV'\underline{\beta}+\underline{\varepsilon}\\ &= Z\underline{\alpha}+\underline{\varepsilon}\end{aligned} \qquad \text{(II.20)}$$

deki TB türünden $\hat{\underline{\alpha}}$ nın kestiricisi

$$\hat{\underline{\alpha}}=(Z'Z)^{-1}Z'\underline{Y}=\Lambda^{-1}Z'\underline{Y}=\Lambda^{-1}V'X\underline{Y} \qquad \text{(II.21)}$$

dir. Bölüm I'de bahsedilen TB in özellikleri göz önüne alındığında, küçük özdeğerlere karşılık gelen TB in X in toplam değişimine katkısı az olacağını biliyoruz. Sıfır ya da sıfıra yakın özdeğerlerin bileşenleri çıkarıldığında, elde edilen TB kestiricisi

$$\hat{\underline{\alpha}}_{TB}=\begin{bmatrix} I_r & 0 \\ & \\ 0 & 0 \end{bmatrix}\hat{\underline{\alpha}}=\begin{bmatrix} \hat{\underline{\alpha}}_r \\ \\ \underline{0} \end{bmatrix}=\begin{bmatrix} \Lambda_r^{-1} V_r'X'\underline{Y} \\ \\ \underline{0} \end{bmatrix} \qquad \text{(II.22)}$$

olmak üzere

$$\hat{\underline{\beta}}_{TB}=V\hat{\underline{\alpha}}_{TB}=V\begin{bmatrix} \Lambda_r^{-1}V_r'X'\underline{Y} \\ \\ \underline{0} \end{bmatrix}$$

dır. Ayrıca

$$\hat{\underline{\beta}}_{TB}=[\,V_r \vdots V_{p-r}]\begin{bmatrix} \Lambda_r^{-1}V_r'X'\underline{Y} \\ \\ \underline{0} \end{bmatrix}$$

$$=V_r\Lambda_r^{-1}V_r'X'\underline{Y}$$

$$=\hat{\underline{\beta}}^{+} \qquad \text{(II.23)}$$

dır. Görüldüğü gibi $\hat{\underline{\beta}}^{+}$ kestiricisi $\underline{\beta}$ nın bir TB kestiricisidir.

TB için AKT (Artık Kareler Toplamı)

$$AKT_r=\sum_{i=1}^{n}(y_i-\bar{y})^2-\sum_{j=1}^{r}\lambda_j\hat{\alpha}_j^2 \qquad \text{(II.24)}$$

olur.[15]

Şimdi TB kestiricisinin dağılım özelliklerine bakalım, TB kestiricisinin

$$E(\hat{\underline{\beta}}_{TB})=(X'X)^{+}X'E(\underline{Y})$$
$$=(X'X)^{+}X'X\underline{\beta}$$
$$=\underline{\beta}-V_{p-r}V'_{p-r}\underline{\beta}$$
$$=\underline{\beta}-\sum_{j=r+1}^{p}(\underline{V}'_j\underline{\beta})\underline{V}_j \qquad \text{(II.25)}$$

olduğundan $\hat{\underline{\beta}}_{TB}$ kestiricisi yanlıdır. Bireysel parametreler için

$$E(\hat{\beta}_{TB_i})=\beta_i-\sum_{j=r+1}^{p}(\underline{V}'_j\underline{\beta})V_{ji} \qquad \text{(II.26)}$$

dir. Burada V_{ji} , $\underline{V}_j$ nin i. elemanıdır.

$$Cov(\hat{\underline{\beta}}_{TB})=Cov((X'X)^{+}X'X\hat{\underline{\beta}})$$
$$=\sigma^2(X'X)^{+}X'X(X'X)^{-1}X'X(X'X)^{+}$$
$$=\sigma^2(X'X)^{+}$$
$$=\sigma^2\sum_{j=1}^{r}\lambda_j^{-1}\underline{V}_j\underline{V}'_j \qquad \text{(II.27)}$$

dır. Bireysel parametrelerin varyansları

$$Var(\hat{\beta}_{TB_i})=\sigma^2\sum_{j=1}^{r}\lambda_j^{-1}V_{ji} \qquad \text{(II.28)}$$

olur. Sonuçta, $\underline{\varepsilon}\sim N(\underline{0},\sigma^2 I)$ olduğundan

$\hat{\underline{\beta}}_{TB}\sim N((X'X)^{+}X'X\underline{\beta},\sigma^2(X'X)^{+})$ dağılımlıdır.

Tüm kestirimler üzerinden yanlı kestirim ortalaması (YKO) ise

$$YKO(\hat{\underline{\beta}}_{TB})=\sigma^2\sum_{j=1}^{r}\lambda_j^{-1}\underline{V}_j\underline{V}'_j+[\sum_{j=r+1}^{p}(\underline{V}_j\underline{\beta})\underline{V}_j]^2$$

$$=\sigma^2\sum_{j=1}^{r}\lambda_j^{-1}+\sum_{j=r+1}^{p}(\underline{V}'_j\underline{\beta})^2 \qquad \text{(II.29)}$$

şeklinde elde edilir. [13]

II.3.b. Çıkarılacak Bileşenlerin Seçim Ölçütü

Çoklubağlantı tam olduğu zaman $\lambda_j = 0$ 'a ve çoklubağlantı güçlü olması halinde $\lambda_j \approx 0$ 'a karşılık gelen özvektörlerin çıkarılmasıyla TB kestiricileri elde edilebileceğini söylemiştik. Yalnız, $\lambda_j \approx 0$ olduğu durumlarda hangi bileşenlerin çıkarılacağı sorununa iki yol önerilmiştir,

a) En küçük özdeğerlere karşılık gelen özvektörleri çıkarmak,

b) En küçük $\hat{\alpha}_j$ lere karşılık gelen bileşenleri çıkarmak.

Hangi bileşenlerin çıkarılacağına karar vermek için ya da daha önce adı geçen $V'_{p-r}\underline{\beta} = 0$ kısıtlamasının doğruluğunu test etmek için

$$U_r = [(AKT_r - AKT)/(AKT)]\frac{n-p}{p-r} \qquad (II.30)$$

veya

$$U = [\sum_{j=r+1}^{p} \lambda_i \hat{\alpha}_j^2 / (1 - \sum_{j=1}^{p} \lambda_i \hat{\alpha}_j^2)]\frac{n-p}{p-r} \qquad (II.31)$$

İstatistiği verilmiştir. U_r veya $U \sim F_{p-r\,;\,n-p}$ dağılımına sahip olup, tablo değerleriyle karşılaştırması yapılarak çıkarılacak bileşenlerin sayısının bulunabileceği belirtilmiştir. [16]

Şimdi bu bölümde söylenenleri bir örnek üzerinde gösterelim.

Örnek II.1 Hald'ın çimento verilerinden yararlanarak TBR'de söylenenleri göstermeye çalışalım. Portland çimentosunun sertleşmesi sırasında açığı çıkan ısı, çimentonun bileşenlerinin lineer fonksiyonu olarak sunulmuştur.[17]

Bağımsız değişkenler:

X_1: $3CaOAl_2O_3$(tricalcium aliminate)in % ağırlığı

X_2: $3CaOSiO_2$(tricalcium silicate)in % ağırlığı

X_3: $4CaOAl_2O_3Fe_2O_3$(tetracalcium alimuno ferrite)in % ağırlığı

X_4: $2CaSiO_2$(dicalcium silicate)in % ağırlığı

Bağımlı değişken:

Y: Gram başına açığa çıkan ısı miktarı (kalori o-alarak)

Çimento Verileri

	X_1	X_2	X_3	X_4	Y
	7	26	6	60	78.5
	1	29	15	52	74.3
	11	56	8	20	104.3
	11	31	8	47	87.6
	7	52	6	33	95.9
	11	55	9	22	109.2
	3	71	17	6	102.7
	1	31	22	44	72.5
	2	54	18	22	93.1
	21	47	4	26	115.9
	1	40	23	34	83.8
	11	66	9	12	113.3
	10	68	8	12	109.4
Ort.	7.46154	48.15185	11.76923	29.99999	95.42308
Stand. sapma	5.88239	15.56088	6.40513	16.73818	15.04372

Bu veriler için bulunan en küçük kareler kestirimleri ve bunların varyansları

	$\hat{\beta}_j$	$Var(\hat{\beta}_j)$
X_1	1.55216	0.55536
X_2	0.51127	0.52451
X_3	0.10297	0.57029
X_4	-0.14298	0.50337

olsun.

Korelasyon matrisi

$$R = \begin{bmatrix} 1. & 0.22857 & -0.82412 & -0.24544 \\ & 1. & -0.13923 & -0.97296 \\ & & 1. & 0.02953 \\ & & & 1. \end{bmatrix}$$

ve

$$R^{-1} = \begin{bmatrix} 38.53617 & 94.23650 & 41.92868 & 99.90851 \\ & 254.76032 & 105.22158 & 267.89398 \\ & & 46.91811 & 111.28187 \\ & & & 282.88552 \end{bmatrix}$$

olarak bulunur.

R(korelasyon) matrisinin özdeğerler matrisi

$$\Lambda = \begin{bmatrix} 2.23570 & 0. & 0. & 0. \\ & 1.57607 & 0. & 0. \\ & & 0.18660 & 0. \\ & & & 0.00162 \end{bmatrix}$$

ve bunlara karşılık gelen özvektörler matrisi

$$V = \begin{bmatrix} 0.47595 & 0.50898 & 0.67550 & 0.24105 \\ 0.56387 & -0.41194 & -0.31442 & 0.64176 \\ -0.39407 & -0.60497 & 0.63769 & 0.26846 \\ -0.54793 & 0.45123 & -0.19542 & 0.67674 \end{bmatrix}$$

olarak bulunur. Ayrıca

$$X'\underline{Y}=\begin{bmatrix} 0.73071 \\ 0.81624 \\ -0.53466 \\ -0.82629 \end{bmatrix}$$

dir. Verilerden bulunan determinasyon katsayısı $R^2=0.98235$ olsun.

Şimdi parametreleri test edelim. Önce her bir parametrenin sıfırdan farklı olup olmadığı testini yapalım.

$H_o: \beta_j = 0$

$H_1: \beta_j \neq 0 \quad j=1,2,3,4$

hipotezini $\alpha=0,05$ anlamlılık düzeyi ve $n=N-1=12$ serbestlik derecesiyle $t_{N-1}=\frac{\hat{\beta}_j-\beta_j}{\sqrt{var(\hat{\beta}_j)}}$ kullanılarak her bir β_j ye

karşılık bulunan; β_1 için $t_{1H}=2.08282$, β_2 için $t_{2H}=0.70595$, β_3 için $t_{3H}=0.13935$ ve β_4 için $t_{4H}=0.20153$ değerleri, $t_{12;0.025}=|2.179|$ tablo değerleri ile karşılaştırıldığında bütün β_j ler için H_o red edilemeyecek. Bunun anlamı şudur, her bir β_j ayrı ayrı test edildiğinde $\alpha=0.05$ anlamlılık düzeyinde istatistiksel olarak hiç biri anlamlı değildir. İkinci olarak bu parametrelerin hepsinin aynı anda sıfıra eşit olup olmadığını test edelim.

$H_o: \beta_1=\beta_2=\beta_3=\beta_4=0$

H_1: En az bir β_j sıfırdan farklıdır

hipotezi $\alpha=0.05$ anlamlılık düzeyinde ve $F=\frac{R^2/(k-1)}{(1-R^2)/(N-k)}$

istatistiğini kullanarak F;k-1, N-k serbestlik derecesiyle test edildiğinde, bulunan $F_H=111.13$ değeri, $F_{4,8;0,05}=3.84$ tablo değeriyle karşılaştırıldığında H_o red edilir. Bunun anlamı şudur, $\alpha=0.05$ anlamlılık düzeyinde sözü edilen parametreler aynı anda istatistiksel olarak sıfıra eşit değildir.

Yapılan bu parametre testlerinin sonuçları çelişkili olması, akla değişkenler arasında çoklubağlantının var olup olmadığı şüphesini getirmektedir. Bu yüzden incelemelerimiz bu doğrultuda olacaktır.

Şimdi daha önce bulunan değerler kullanılarak önce çoklubağlantıyı inceleyelim.

$(VIF)_j$ aynı zamanda $c_{jj}=\sum_{j=1}^{p} \lambda_j^{-1} V_{ji}$, R^{-1} matrisinin köşegen elamanları olduğuna göre maksimum (VIF)=282.88550 için

$$\text{mak}(VIF)_j=282.88550 \geq 10$$

olduğundan şiddetli çoklubağlantı vardır.

Çoklubağlantı derecesini $k_2(X'X)$ yadımıyla da görebiliriz.

$$k_2(X'X)=\frac{\underset{j}{\text{mak}}|\lambda_j|}{\underset{j}{\text{min}}|\lambda_j|}=\frac{2.23570}{0.00162}=1380.06$$

Burada $1380.06 > 30$ olduğundan şiddetli bir çoklubağlantı olduğu tekrar görülür.

Korelasyon matrisinden elde edilen özdeğerlerden sıfıra en yakın olan λ_4 tür. Minimum özdeğere karşılık gelen bileşen

$$Z_4 = 0.24105X_1 + 0.64176X_2 + 0.26846X_3 + 0.67674X_4 \approx 0$$

olup, çoklubağlantıya neden olan değişkenler ise katsayıları mutlak değerce büyük olan X_2 ve X_4dür.

Şimdi, $\lambda_4 \approx 0$ olduğundan, λ_4 ve ona karşılık gelen $\underline{v}_4$ çıkartılarak TB kestirimlerini bulmaya çalışalım.

$$\Lambda_r = \begin{bmatrix} 2.23570 & 0. & 0. \\ & 1.57607 & 0. \\ & & 0.18660 \end{bmatrix}$$

ve

$$V_r = \begin{bmatrix} 0.47595 & 0.50898 & 0.67550 \\ 0.56387 & -0.41394 & -0.31442 \\ -0.39407 & -0.60497 & 0.63769 \\ -0.54793 & 0.45123 & -0.19542 \end{bmatrix}$$

olur. Buradan

$$\hat{\underline{\alpha}}_r = \Lambda_r^{-1} V_r' X' \underline{Y}$$

$$= \begin{bmatrix} 0.65695 \\ 0.00831 \\ 0.04762 \end{bmatrix}$$

ve

$$\hat{\underline{\beta}}^+ = V_r \hat{\underline{\alpha}}_r$$

$$= \begin{bmatrix} 0.34907 \\ 0.35202 \\ -0.23355 \\ -0.36552 \end{bmatrix}$$

olarak bulunur. Kestirimlerin varyansları ise

$Var(\hat{\beta}_{TB})$
0.25134
0.06836
0.23618
0.03934

olur.

BÖLÜM III
UYGULAMA

Şimdi, bölüm I ve bölüm II nin sonunda verilen örneklerden başka, her iki bölümde söylenenleri; üç değişkenli, 10 birimlik veriler için uygulamaya çalışacağız.

Aşağıda verilen verileri ilk önce TBA açısından inceleyelim.

Birimler	X_1	X_2	X_3
1	20	3	5
2	22	5	6
3	18	0	5
4	24	5	6
5	30	5	4
6	25	8	5
7	35	11	5
8	28	8	4
9	40	5	6
10	18	5	4
Ortalamalar	26.0	5.5	5.0
Standart sapmalar	7.3182	2.9907	0.8165

Korelasyon matrisi

$$R = \begin{bmatrix} 1. & 0.4418 & 0.3347 \\ & 1. & -0.2013 \\ & & 1. \end{bmatrix}$$

olup, R nin özdeğerler matrisi

$$L = \begin{bmatrix} 1.4720 & 0. & 0. \\ & 1.1890 & 0. \\ & & 0.3389 \end{bmatrix}$$

dır. Burada

$$\sum_{j=1}^{3} \ell_j = tr(R) = 3$$

ve

$$|R| = \prod_{j=1}^{3} \ell_j = 0.5931$$

dir. Bileşenler tarafından açıklanan varyans payları ve birikimli açıklanan pay ise

Özdeğerler	Açıklanan varyas payları %	Birikimli açıklanan pay %
1.4720	49.07	49.07
1.1890	39.63	88.70
0.3389	11.30	100.00

olur. Özdeğerlere karşılık gelen özvektörler de

$$V = [\underline{v}_1, \underline{v}_2, \underline{v}_3] = \begin{bmatrix} 0.758 & 0.157 & -0.642 \\ 0.560 & -0.628 & 0.582 \\ 0.350 & 0.840 & 0.502 \end{bmatrix}$$

olarak bulunur. Her biri bir bileşenle değişken arasındaki korelasyon katsayılarını içeren yapı matrisi ise

$$A = \begin{bmatrix} 0.9197 & 0.1712 & -0.3737 \\ 0.6794 & -0.6848 & 0.3308 \\ 0.4265 & 0.9159 & 0.2923 \end{bmatrix}$$

olur. A matrisi incelendiğinde, X_1 ve X_2 değişkenleri birinci temel bileşen ile yakın ilişki halindedir. X_1 değişkeni birinci bileşen eksenine daha yakındır.

Çünkü bu değişkenin bileşen ekseniyle yaptığı korelasyon satırdaki en büyük sayıdır. Benzer durum ikinci bileşen ile X_3 değişkeni arasında vardır. Bir bileşene yakın olan değişkenler ortak nitelikli değişkenler olup bileşenin adlandırılmasına yardımcı olur. Birinci bileşenin adlandırılmasına yardımcı olur. Birinci bileşenin ortak niteliği, veri matrisindeki değişkenliğin % 49.07 ini açıklar. İlk iki bileşen beraber % 88.70 lik kısmını açıklar. Amaca uygun olarak istenirse % 11.3 lük bilgi kaybı ile bu üç değişken yerine onların ortak niteliğini temsil eden birinci ve ikinci temel bileşen alınabilir. Şayet bilgi kaybı amaca göre kabul edilebilir düzeyde ise üç boyutlu uzay yerine iki boyutlu düzlem ele alınabilir. Bunu iki bileşen için şu şekilde yazabiliriz.

	Bileşenler	
Değişkenler	1. Bileşen	2. Bileşen
X_1	0.9197	0.1712
X_2	0.6794	-0.6848
X_3	0.4265	0.9159

Yukarıda ortak niteliği belirlenen bu bileşenlerin, birey(birim)lerin bileşen değerlerine göre sıralanışı ise şöyledir. Bu sıralama çalışmanın amacına uygun olarak birimlerin sıralanması önemli ise yapılır. Şimdi bu sıralamayı birinci ve ikinci bileşen için yapalım.

	1. Bileşen Açıklayıcılık payı % 49.07		2. Bileşen Açıklayıcılık payı % 39.63	
Sıra No	Kod No	Bileşen değeri	Kod No	Bileşen değ.
1	7	1.6171	9	1.3153
2	9	1.4712	2	0.9611
3	6	0.3005	1	0.3620
4	5	0.2643	6	-0.5011
5	10	0.2524	10	-0.6898
6	8	0.2075	5	-0.7685
7	3	0.1007	7	-0.8819
8	2	-0.0652	4	-0.8866
9	4	-0.5005	3	-1.2165
10	1	-0.8970	8	-1.3855

Şimdi bulunan TB için örnek özelliklerini inceleyelim. Birinci bileşen (en büyük özdeğer) için % 95 lik güven aralığını oluşturalım. $n=9$, $\ell_1= 1.4720$, $Z_{0.025}=1.96$ değerleri ile

$$P(0.7844 \leqslant \lambda_1 \leqslant 11.9190) = 0.95$$

olur. İkinci temel bileşen için ise

$$P(0.6336 \leqslant \lambda_2 \leqslant 9.6275) = 0.95$$

olarak bulunur.

Alınan temel bileşenlerden sonra geride kalan bileşenlerin yığın değerinin sıfır olup olmadığını test edelim. Bunun için artık matrisi bulalım.

$$R_{1+2} = \begin{bmatrix} 0.8752 & 0.5076 & 0.5491 \\ & 0.9305 & -0.3374 \\ & & 1.0208 \end{bmatrix}$$

den

$$R_{A(1.+2.)} = \begin{bmatrix} 0.1248 & -0.0658 & -0.2144 \\ & 0.0695 & -0.1361 \\ & & -0.0208 \end{bmatrix}$$

olur. Bu artık matrisin yığın değeri sıfır olan matristen gelip gelmediğini, $\alpha = 0.05$ anlamlılık düzeyinde 1 serbestlik derecesiyle test edelim.

H_o: $P=\underline{0}$

H_1: $P\neq\underline{0}$

hipotezi ile

$Q = 0.9997$

ve

$\chi^2 = 0.0074$

olup, bu değer tablo değeri $\chi^2_{0.05;1} = 3.841$ ile karşılaştırıldığında H_o red edilemez. Yani adı geçen verilerin artık matrisi, $\alpha=0.05$ anlamlılık düzeyinde yığın değeri sıfır olan bir matristen geldiği istatistiksel olarak (görülmüştür) reddedilememiştir. Bu da küçük bileşenin yığın değeri sıfıra yakın olduğu anlamındadır.

Aynı verileri Y değişkenini de göz önüne alarak TBR açısından inceleyelim.

Y: 4,4,3,6,4,7,4,3,8,2

Daha önceden bulunan değerler de kullanılarak, önce çoklubağlantı yapısını görmeye çalışalım. Bunun için $k_2(X'X)$ kullanılarak

$$k_2(X'X) = \frac{\underset{j}{\text{mak.}}|\lambda_j|}{\underset{j}{\text{min.}}|\lambda_j|} = \frac{1.4720}{0.3389} = 4.34$$

olup, 4.34 < 10 olduğundan verilerde düşük şiddette çoklubağlantı vardır.

Korelasyon matrisinden elde edilen özdegerlerden sıfıra en yakın olanı λ_4 tür. En küçük özdeğere karşılık gelen bileşen

$$Z_3 = -0.642\ X_1 + 0.582\ X_2 + 0.502\ X_3 \approx 0$$

olup, çokbağlantıya neden olan değişkenler ise katsayıları mutlak değerce büyük olan X_1 ve X_2 dir.

Şimdi, $\lambda_3 \approx 0$ olduğundan, λ_3 ve ona karşılık gelen $\underline{v}_3$ çıkartılarak TB kestirimlerini bulalım.

$$\Lambda_r = \begin{bmatrix} 1.4720 & 0 \\ 0 & 1.1890 \end{bmatrix}$$

ve

$$V_r = \begin{bmatrix} 0.758 & 0.157 \\ 0.560 & -0.628 \\ 0.350 & 0.840 \end{bmatrix}$$

dır.

$$X'\underline{Y} = \begin{bmatrix} 2.2290 \\ -1.2321 \\ 5.8002 \end{bmatrix}$$

den

$$\hat{\underline{\alpha}}_r = \begin{bmatrix} 2.0590 \\ 3.8443 \end{bmatrix} \quad \text{ile} \quad \hat{\underline{\beta}}^{+} = \begin{bmatrix} 2.1640 \\ 1.2612 \\ 3.9499 \end{bmatrix}$$

olarak bulunur.

SONUÇ

Bu çalışmada istatistiksel veri analizinde önemli bir yeri olan temel bileşenler, temel bileşenler analizi ve lineer modellerde temel bileşenler başlıkları altında iki bölümde incelenmiştir.

Temel bileşenler analizi, $\underline{X}$ tesadüfi vektörünü, bileşenleri ilişkisiz ve varyanslarının büyüklüklerine göre sıralanmış yeni bir tesadüfi vektöre dönüşümü olarak tanımlanmıştır. Değişkenler normal dağılımlı olduğunda uygulanacak dönüşüm sonucunda yeni bileşenler, olasılık yoğunluk fonksiyonunun aynı değerli olduğu noktaların geometrik yeri, elipsoidlerin temel eksenlerine karşılık gelir. $\mathbb{R}^p$ de verilen noktaların kendisine dik uzaklıkların kareleri toplamını minimum yapan D doğrusu, S örnek kovaryans matrisinin birinci temel bileşenine karşılık geldiği görülmüştür. Ayrıca bölümün sonunda TBA bir örnek üzerinde gösterilmiştir.

Lineer modellerde temel bileşenler regresyonu kestiricisi, çoklubağlantı sorununun incelenmesinde yanlı bir kestirimdir. Kısıtlanmış EKKK olarak temel bileşenler kestiricisi, $\hat{\underline{\beta}}^+$ olarak verilmiştir. Ayrıca bölümün sonunda verilen örnekten de anlaşıldığı gibi TBR kestirimi, EKKK'ye göre yanlı fakat daha küçük bir kestirim olduğu görülmüştür.

ÖZGEÇMİŞ

ÖZGEÇMİŞ

1961 Rize doğumlu olan araştırmacı, ilk öğrenimini Rize Bahattinpaşa Köyü İlkokulu, orta öğrenimini Samsun Ondokuz Mayıs Lisesi'nde tamamlamıştır. Gazi Üniversitesi Fen-Edebiyat Fakültesi İstatistik Bölümü'nü 1984'de bitirmiş ve bu tarihten Temmuz 1986 tarihine kadar burslu öğrenci olarak kabul edildiği Başbakanlık Devlet İstatistik Enstitüsü'nde Ekonomik İstatistikler Dairesi Fiyat İstatistikleri ve İndeks Şubesi'nde istatistikçi olarak çalışmıştır. Bu tarihten sonra Gazi Üniversitesi Fen-Edebiyat Fakültesi İstatistik Bölümü'nde araştırma görevlisi olarak çalışmaya başlamıştır. Araştırmacı, halen bu görevine devam etmektedir.