II.3 Önemli TB Sayısını Belirleme

Her bir uygulamada verileri etkili bir şekilde özetlemek için kaç tane TB alınması gerektiğine, bir diğer ifade ile önemli TB sayısına karar verilmelidir. Bu amaçla uygulanabilecek yöntemler:

1) Analitik Yöntemler

2) İstatistiksel Yöntemler

3) Grafik Yöntemi

olarak sınıflandırılmaktadır.

II.3.1 Analitik Yöntemler

Bu yöntemler varyans açıklama oranlarına dayalı olan yöntemlerdir.

i) Toplam varyansı açıklama oranı: İlk k tane $(k < p)$ TB tarafından toplam varyansı açıklama oranı %80'i geçiyorsa, bu TB'ler önemli geriye kalanlar $(g = p - k$ tane$)$ önemsiz kabul edilir. Böylece ilk k tane TB alınıp, geriye kalan $g = p - k$ tane TB ihmal edilebilir.

ii) 1'den büyük olan öz değer sayısı: Özellikle korelasyon matrisinden elde edilen TB'ler için, tüm öz değerlerin ortalamasından (yani $\frac{\left(\sum_{j=1}^{p} \lambda_j\right)}{p} = 1$) büyük olan öz değer sayısı kadar TB'ler önemli, geriye kalanlar önemsiz kabul edilebilir. Çünkü her bir öz değer bir TB'nin varyansıdır. Buna göre varyansı 1'den büyük olan TB'ler önemlidir.

II.3.2 İstatistiksel Yöntemler

Bu yöntemler önemli temel bileşen sayısını hipotez testi ile belli bir güven seviyesinde belirlemeye çalışan yöntemlerdir. Büyük varyanslı TB'lerin anlamlılığını test etmek için kitleye ait son g tane öz değerin önemsiz olup olmadığı test edilir. Kitleye ait öz değerler $\gamma_1 > \gamma_2 > \cdots > \gamma_k > \cdots > \gamma_p > 0$ olmak üzere öz değerler TB'lerin varyansları olduğundan önemli TB sayısı ile önemli öz değer sayısı eş anlamlı olarak kullanılabilir. Bu anlamda önemli TB ya da özdeğer sayısı belirlemede hipotez testine dayalı iki ayrı yöntem kullanılmaktadır.

i) Anderson'ın Genel Olabilirlik Oran Testi

Kabul edelim ki $k < p$ olmak üzere kitleye ait TB'lerden bir diğer ifade ile öz değerlerden ilk k tanesi $(\gamma_1, \gamma_2, \dots, \gamma_k)$ önemli olsun. Bu durumda geriye kalan $g = p - k$ tane TB ya da bunların varyansları olan öz değerlerin $(\gamma_{k+1}, \gamma_{k+2}, \dots, \gamma_{k+g})$ önemsiz olduğunu düşünebiliriz. Bu öz değerler kitleye ait parametreler olduğundan ya Σ matrisinin ya da ρ matrisinin öz değerleri olacaktır. Söz konusu düşüncenin doğru olup olmadığını bir hipotez testi ile kontrol edelim. Test edilecek hipotezler:

$H_0: \gamma_{k+1} = \gamma_{k+2} = \dots = \gamma_{k+g} = 0$

$H_1: \gamma_{k+1} > 0 \ \ (p = k + g)$

şeklinde kurulur. H_0 hipotezini test etmek için gerekli olan test istatistiği hipotezde yer olan parametrelerin tahmin edicileri olan ve örnek varyans kovaryans matrisinin (S) ya da örnek korelasyon matrisinin (R) öz değerlerine bağlıdır. Bu öz değerler $\lambda_1 > \lambda_2 > \cdots > \lambda_k > \cdots > \lambda_p > 0$ olmak üzere; test istatistiği;

$$\chi^2 = -(n-1)\sum_{j=k+1}^{k+g} ln\lambda_j + (n-1)gln\left[\frac{\sum_{j=k+1}^{k+g}\lambda_j}{g}\right] \sim \chi^2_{\frac{(g-1)(g+2)}{2}}$$

olarak tanımlıdır. Karar kuralı ise α önem seviyesinde kritik değer $\chi_t^2 = \chi^2_{\frac{(g-1)(g+2)}{2};\alpha}$ olmak üzere, eğer $\chi^2 > \chi_t^2$ ise H_0 ret edilir, aksi takdirde yani $\chi^2 \leq \chi_t^2$ ise H_0 ret edilemez.

Eğer birinci aşama sonucunda H_0 hipotezi ret edilirse ilk k tane öz değerin (TB'nin) yanı sıra $(k+1)$-nci öz değerin (TB'nin) de önemli olduğu belirlenmiş olur. Geriye kalan öz değerler (TB'ler) içerisinde başka önemli olanların olup olmadığı yeni hipotezler oluşturularak test işlemi tekrarlanır. Bu test süreci ardışık olarak H_0 hipotezinin kabul edildiği aşamaya kadar devam eder. H_0 hipotezinin kabul edildiği aşamada işlem durdurulur ve bu aşamada geriye kalan öz değerlerin (TB'lerin) önemsiz olduğuna, bunlardan öncekilerin ise önemli olduğuna karar verilir. Böylece ileriye doğru genişleme algoritması diyebileceğimiz bu algoritma ile başlangıçta kabul edilen sayıdan daha fazla önemli TB sayısı belirlenmiş olur.

Eğer birinci aşama sonucunda H_0 hipotezi kabul edilirse geriye kalan $g = p - k$ tane öz değerin (TB'nin) önemsi olduğu kesin olarak belirlenmiş olur. Hatta önemli olduğu kabul edilen ilk k tane öz değer (veya TB) grubu içerisinde de önemsiz olan bulunabilir. Bu grup içerisinde gerçekten önemsiz olan öz değer (ya da TB) olup olmadığını belirlemek için bu gruptaki öz değerlerden en küçüğü olan λ_k önemsizler grubuna verilerek yeni hipotezler oluşturulur ve test işlemi tekrarlanır. Bu test süreci ardışık olarak H_0 hipotezinin ret edildiği aşamaya kadar devam eder. H_0 hipotezinin ret edildiği aşamada işlem durdurulur ve bu aşamada test edilen öz değerin (veya TB'nin) önemli olduğuna geldiği grubun içerisinde kalması gerektiğine karar verilir. Böylece geriye doğru daralma algoritması diyebileceğimiz bu algoritma ile başlangıçta kabul edilen sayıdan ($k'dan$) daha az ya da başlangıçta kabul edilen sayı kadar önemli TB sayısı belirlenmiş olur.

Kolaylık olması bakımından bu test işlemi sürecine birinci aşamada $H_0: \gamma_{p-1} = \gamma_p = 0$; $H_1: \gamma_{p-1} > 0$ ile başlayıp geriye doğru daralma algoritmasını uygulayarak önemli öz değer (veya TB) sayısı belirlenebilir. Eğer H_0 hipotezi ret edilirse işlem biter ve önemli öz değer (TB) sayısının $k = p - 1$ olduğu kesinleşmiş olur. Eğer H_0 hipotezi kabul edilirse, o zaman yeni hipotezler $H_0: \gamma_{p-2} = \gamma_{p-1} = \gamma_p = 0$; $H_1: \gamma_{p-2} > 0$ şeklinde oluşturulup test edilir. Böylece süreç geriye doğru en az bir k değeri için yeni H_0 hipotezi ret edilinceye kadar tekrarlanır. H_0 hipotezinin ret edildiği aşamada süreç durdurulur ve bu aşamada test edilen öz değer (TB) ile birlikte bu öz değerden büyük olanların önemli olduğuna karar verilir. Sonuç olarak ilk k tane TB'nin önemli olduğu belirlenmiş olur, çünkü ilk k tane öz değer ilk k tane TB'nin varyansıdır.

ii) Büyük Varyanslı TB'lerin Anlamlılık Testi

Bu test işleminin uygulama süreci Anderson'ın genel olabilirlik oran testi ile tamamen aynı olup, tek farklılık hipotezleri test ederken kullanılacak olan test istatistiğindedir. Bu test işlemi için test istatistiği;

$$U = \left(n - \frac{2p+11}{6}\right)\left(gln\bar{\lambda} - \sum_{j=k+1}^{p} ln\lambda_j\right) \sim \chi^2_{\frac{(g-1)(g+2)}{2}}$$

şeklindedir. Burada $\bar{\lambda}$: geriye kalan $g = p - k$ tane öz değerin ortalaması olup,

$\overline{\lambda} = \frac{1}{g}\sum_{j=k+1}^{p} \lambda_j$

eşitliği ile bulunur.

iii) Bartlett Testi

$\underline{X} : p \times 1$ başlangıç değişkenler vektörü, $\underline{Z}: p \times 1$ standart değişkenler vektörü ve $Cov(\underline{Z}) = Kor(\underline{X}) = R : p \times p$ olmak üzere, R matrisinin öz değer-öz vektör çiftleri $j = 1, 2, \dots, p$ için $(\lambda_j\,,\ \underline{a}_j)$ olsun. $\Lambda: p \times p, R$ nin özdeğerler matrisi ve $A: p \times p; R$ nin özvektörler matrisi olmak üzere spectral ayrışım gereğince;

$$R = A'\Lambda A = \sum_{j=1}^{p} \lambda_j\, \underline{a}_j \underline{a}_j' = \sum_{j=1}^{p} R_j \quad \dots\dots(*)$$

yazılabilir. Burada $j = 1, 2, \dots, p$ için $R_j = \lambda_j \underline{a}_j \underline{a}_j': p \times p$ matrisler olup, her bir R_j matrisi, R matrisinin bir parçasını oluşturmaktadır.

Kabul edelimki önemli TB (ya da özdeğer) sayısı $k, (k < p)$ olsun. Bu durumda geriye kalan $g = p - k$ tane öz değerin (TB'nin) önemsiz olduğu düşünülür. Böylece (*) eşitliği gereğince, önemli özdeğerler (TB'ler) için ;

$$R_h = \sum_{j=1}^{k} R_j = \lambda_1 \underline{a}_1 \underline{a}_1' + \lambda_2 \underline{a}_2 \underline{a}_2' + \cdots + \lambda_k \underline{a}_k \underline{a}_k'$$

matrisi ve önemsiz özdeğerler (TB'ler) için;

$$R_g = \sum_{j=k+1}^{k+g} R_j = \lambda_{k+1} \underline{a}_{k+1} \underline{a}_{k+1}' + \lambda_{k+2} \underline{a}_{k+2} \underline{a}_{k+2}' + \cdots + \lambda_{k+g} \underline{a}_{k+g} \underline{a}_{k+g}'\,, (k + g = p)$$

matrisi yazılabilir. Bu matrislerden R_h: hipotez için tanımlanan korelasyon matrisi ve R_g: Kalıntı (Artık) matrisi adını alır.

Eğer R_g artık matrisinin istatistiksel anlamda bir sıfır matrisi olduğu gösterilebilirse, önemli özdeğer (TB) sayısı olarak k sayısı alınabilir. Bu durumu bir hipotez testi ile inceleyebiliriz. Test edilecek hipotezler;

$H_0: R_g = [o];\ \ (\gamma_{k+1} = \gamma_{k+2} = \ \dots = \gamma_{k+g} = 0)$

$H_1: R_g \neq [o]\ \ , \gamma_{k+1} > 0\ \ (p = k + g)$

şeklinde kurulur. Test istatistiği;

$$\chi^2 = -\left[(n-1) - \frac{2p+5}{6} \times \frac{2k}{3}\right] ln(U_g) \sim \chi^2_{\frac{(g-1)(g+2)}{2}}$$

eşitliği ile verilir. Burada;

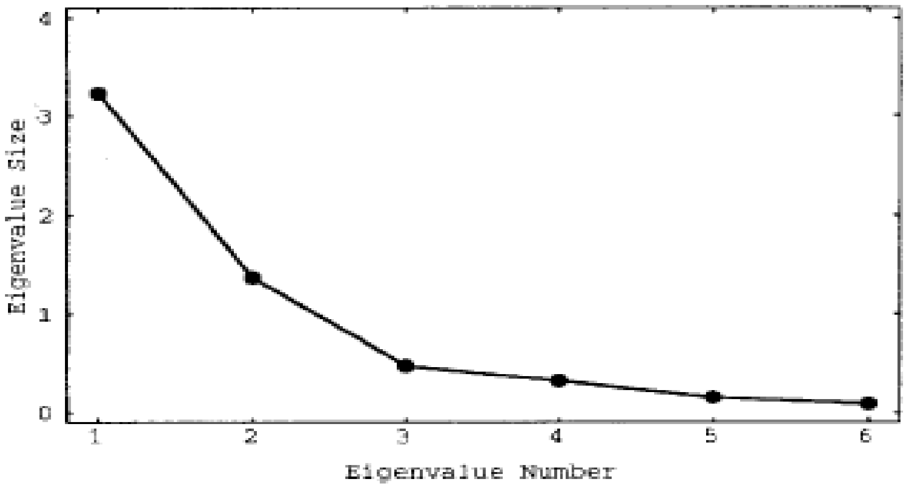
$$U_g = \frac{|R| g^g}{\left(\prod_{j=1}^{k} \lambda_j\right)\left(p - \sum_{j=1}^{k} \lambda_j\right)^g} \quad , \ 0 \leq U_g \leq 1 \text{ şeklinde tanımlıdır.}$$

Karar kuralı ise α önem seviyesinde kritik değer $\chi_t^2 = \chi^2_{\frac{(g-1)(g+2)}{2};\,\alpha}$ olmak üzere, eğer $\chi^2 > \chi_t^2$ ise H_0 ret edilir, aksi takdirde yani $\chi^2 \leq \chi_t^2$ ise H_0 ret edilemez.

H_0 hipotezinin ret ya da kabul kararına göre Anderson'ın genel olabilirlik oran testinde olduğu gibi bir döngü oluşturularak test işlemi önemli özdeğer (veya TB) sayısı belirleninceye kadar tekrarlanır.

II.3.3 Scree Graph (Dik Yamaç Grafik) Yöntemi

Koordinat düzleminde yatay eksen üzerinde öz değerlerin indis numaralarını $(j = 1, 2, \dots, p)$ ve düşey eksende öz değer büyüklüklerini gösterelim. $j = 1, 2, \dots, p$ için düzlemde $(j\,,\ \lambda_j)$ noktaları belirlenir ve bu noktalar birleştirilerek bir grafik çizilir. Elde edilen grafikte en yüksek (hızlı) düşüşün ortaya çıktığı nokta (dirsek noktası, kırılma noktası) önemli ve önemsiz TB'leri ayıran noktadır. Bu noktaya ve bu noktanın solunda kalan noktalara karşılık gelen TB'ler önemli, sağında kalan noktalara karşılık gelen TB'ler önemsizdir. Örneğin aşağıdaki şekilde yatay eksen üzerindeki 3 noktası dirsek noktası olacağından önemli TB sayısı $k = 3$ alınabilir.



Şekil: Screegraph (Grafik) Yöntemi ile Önemli TB Sayısının Belirlenmesi

<u>Örnek:3</u> $p = 6$ değişken üzerinde $n = 60$ birimin ölçüm değerleri aşağıdadır. Bu verilerden örnek varyans kovaryans matrisi ve korelasyon matrisi ile bu matrislere ait özdeğer ve özvektörler aşağıda verilmiştir. Buna göre ;

a) Bu veriye TBA uygulamanın gerekli olup olmadığına %5 önem seviyesinde karar veriniz?

b) TB'leri başlangıç değişkenlerinin $(X_t; t = 1,2, \dots ,6)$ ve standart değişkenlerin $(Z_t; t = 1,2, \dots ,6)$ doğrusal fonksiyonları olarak elde ediniz?

c) Önemli TB sayısını analitik yöntemlerle belirleyiniz?

d) Önemli TB sayısını istatistiksel yöntemler ile belirleyiniz? (önem seviyesini %5 alınız)

e) Önemli TB sayısını grafik yöntem ile belirleyiniz?

f) Önemli olan TB'ler ile X_t değişkenlerinin ve Z_t standart değişkenlerinin korelasyonlarını hesaplayınız?

g) Önemli olduğuna karar verdiğiniz TB'ler için TB skorlarını bulunuz? (Hem başlangıç verileri hem de standart veriler için)

X_1	X_2	X_3	X_4	X_5	X_6	X_1	X_2	X_3	X_4	X_5	X_6
15,5	60,0	21,1	10,3	13,4	12,4	14,9	56,5	20,4	7,4	13,0	12,0
15,4	59,7	20,0	12,8	14,5	11,3	15,4	57,5	19,5	10,5	13,8	11,5
15,1	59,7	20,2	11,4	14,1	12,1	15,3	55,4	19,2	9,7	13,3	11,5
14,3	56,9	18,9	11,0	13,4	11,0	14,6	56,0	19,8	8,5	12,0	11,5
14,8	58,0	20,1	9,6	11,1	11,7	16,2	56,3	19,5	11,5	14,5	11,8
15,2	57,5	18,5	9,9	12,8	11,4	14,6	58,0	19,9	13,0	13,4	11,5
15,4	58,0	20,8	10,2	12,8	11,9	15,9	56,7	18,7	10,8	12,8	12,6
16,3	58,0	20,1	8,8	13,0	12,9	14,7	55,8	18,7	11,1	13,9	11,2
15,5	57,0	19,6	10,5	13,9	11,8	15,5	58,5	19,4	11,5	13,4	11,9
15,0	56,5	19,6	10,4	14,5	12,0	16,1	60,0	20,3	10,6	13,7	12,2
15,5	57,2	20,0	11,2	13,4	12,4	15,2	57,8	19,9	10,4	13,5	11,4
15,5	56,5	19,8	9,2	12,8	12,2	15,1	56,0	19,4	10,0	13,1	10,9
15,7	57,5	19,8	11,8	12,6	12,5	15,9	59,8	20,5	12,0	13,6	11,5
14,4	57,0	20,4	10,2	12,7	12,3	16,1	57,7	19,7	10,2	13,6	11,5
14,9	54,8	18,5	11,2	13,8	11,3	15,7	58,7	20,7	11,3	13,6	11,3
16,5	59,8	20,2	9,4	14,3	12,2	15,3	56,9	19,6	10,5	13,5	12,1
15,5	56,1	18,8	9,8	13,8	12,6	15,3	56,9	19,5	9,9	14,0	12,1
15,3	55,0	19,0	10,1	14,2	11,6	15,2	58,0	20,6	11,0	15,1	11,7
14,5	55,6	19,3	12,0	12,6	11,6	16,6	59,3	19,9	12,1	14,6	12,1
15,5	56,5	20,0	9,9	13,4	11,5	15,5	58,2	19,7	11,7	13,8	12,1
15,2	55,0	19,3	9,9	14,4	11,9	15,8	57,5	18,9	11,8	14,7	11,8
15,3	56,5	19,3	9,1	12,8	11,7	16,0	57,2	19,8	10,8	13,9	12,0
15,3	56,8	20,2	8,6	14,2	11,5	15,4	57,0	19,8	11,3	14,0	11,4
15,8	55,5	19,2	8,2	13,0	12,6	16,0	59,2	20,8	10,4	13,8	12,2
14,8	57,0	20,2	9,8	13,8	10,5	15,4	57,6	19,6	10,2	13,9	11,7
15,2	56,9	19,1	9,6	13,0	11,2	15,8	60,3	20,8	12,4	13,4	12,1
15,9	58,8	21,0	8,6	13,5	11,8	15,4	55,0	18,8	10,7	14,2	10,8
15,5	57,3	20,1	9,6	14,1	12,3	15,5	58,4	19,8	13,1	14,5	11,7
16,5	58,0	19,5	9,0	13,9	13,3	15,7	59,0	20,4	12,1	13,0	12,7
17,3	62,6	21,5	10,3	13,8	12,8	17,3	61,7	20,7	11,9	13,3	13,3

$Cov(\underline{X}) = S$ ve $Kor(\underline{X}) = Cov(\underline{Z}) = R$ olmak üzere;

$$S = \begin{bmatrix} 0,37017 & 0,59966 & 0,14881 & 0,04441 & 0,10712 & 0,20932 \\ 0,59966 & 2,66080 & 0,80931 & 0,66116 & 0,09873 & 0,37988 \\ 0,14881 & 0,80931 & 0,45826 & 0,01127 & -0,01322 & 0,11984 \\ 0,04441 & 0,66116 & 0,01127 & 1,47372 & 0,25220 & -0,05438 \\ 0,10712 & 0,09873 & -0,01322 & 0,25220 & 0,48801 & -0,03559 \\ 0,20932 & 0,37988 & 0,11984 & -0,05438 & -0,03559 & 0,32368 \end{bmatrix}$$

$$R = \begin{bmatrix} 1,000 & 0,604 & 0,361 & 0,060 & 0,252 & 0,605 \\ 0,604 & 1,000 & 0,733 & 0,334 & 0,087 & 0,409 \\ 0,361 & 0,733 & 1,000 & 0,014 & -0,028 & 0,311 \\ 0,060 & 0,334 & 0,014 & 1,000 & 0,297 & -0,079 \\ 0,252 & 0,087 & -0,028 & 0,297 & 1,000 & -0,090 \\ 0,605 & 0,409 & 0,311 & -0,079 & -0,090 & 1,000 \end{bmatrix}$$

Kovaryans matrisinin (S) özdeğer ve özvektörleri:

λ_j	$\underline{a}_1$	$\underline{a}_2$	$\underline{a}_3$	$\underline{a}_4$	$\underline{a}_5$	$\underline{a}_6$
3,3482	0,205	0,137	0,427	-0,443	0,163	-0,730
1,3788	0,875	0,216	-0,086	0,132	0,331	0,231
0,477	0,262	0,228	-0,117	0,374	-0,776	-0,352
0,3256	0,320	-0,894	-0,172	-0,175	-0,164	-0,109
0,1555	0,063	-0,223	0,864	0,361	-0,108	0,238
0,0895	0,128	0,184	0,141	-0,696	-0,473	0,470

Korelasyon matrisinin (R) özdeğer ve özvektörleri:

λ_j	$\underline{a}_1$	$\underline{a}_2$	$\underline{a}_3$	$\underline{a}_4$	$\underline{a}_5$	$\underline{a}_6$
2,5649	0,510	-0,005	-0,448	-0,033	-0,625	-0,384
1,3681	0,561	0,084	0,323	0,022	-0,221	0,725
0,9352	0,464	-0,145	0,473	0,470	0,306	-0,474
0,678	0,143	0,664	0,313	-0,594	0,090	-0,282
0,3206	0,108	0,646	-0,466	0,489	0,312	0,129
0,1332	0,423	-0,336	-0,394	-0,430	0,601	0,084

Çözüm

a) Hipotezler:

$H_0: R = I_6$ veya $H_0: \Sigma = K\ddot{o}\c{s}[\sigma_{11}\ ,\ \sigma_{22},\ \sigma_{33}\ ,\ \sigma_{44}\ ,\ \sigma_{55}\ ,\ \sigma_{66}\ ]$

$H_1: R \neq I_6$ $H_1: \Sigma \neq K\ddot{o}\c{s}[\sigma_{11}\ ,\ \sigma_{22},\ \sigma_{33}\ ,\ \sigma_{44}\ ,\ \sigma_{55}\ ,\ \sigma_{66}\ ]$

Test istatistiği; $\chi^2 = -\left[n - 1 - \frac{2P+5}{6}\right] ln|R| \sim \chi^2_{p(p-1)/2}$

(veya $\chi^2 = -\left[n - 1 - \frac{2P+5}{6}\right] ln|S| \sim \chi^2_{p(p-1)/2}$)

$p = 6\ ; n = 60\ ;\ |R| = \prod_{j=1}^{6} \lambda_j = 0{,}095\ ;\ \ |S| = \prod_{j=1}^{6} \lambda_j = 0{,}00998$ olup, test istatistiğinin alabileceği değer;

$\chi^2 = -\left[59 - \frac{17}{6}\right] ln(0{,}095) = 132{,}21$ (veya $\chi^2 == -\left[59 - \frac{17}{6}\right] ln(0{,}00998) = 258{,}77$)

dir. Kritik değer $\alpha = 0{,}05$ için $\chi^2_{\frac{p(p-1)}{2};\alpha} = \chi^2_{15;0{,}05} = 24{,}996$ dır. Buna göre 132,21>24,996 (veya 258,77>24,996) olduğundan H_0 hipotezi ret edilir. Yani bu verilere TBA uygulanmalıdır.

b) <u>Başlangıç değişkenlerinin doğrusal fonksiyonları olarak elde edilen TB'ler</u>: Bu TB'ler örnek kovaryans matrisinin özvektörleri kullanılarak elde edilen TB'ler olup;

$Y_j = \underline{a}_j'\underline{X}\ \ j = 1, 2, \ldots, 6$ biçimindedir. Buna göre:

1-nci TB: $Y_1 = \underline{a}_1'\underline{X} = 0{,}205X_1 + 0{,}875X_2 + 0{,}262X_3 + 0{,}320X_4 + 0{,}063X_5 + 0{,}128X_6$

2-nci TB: $Y_2 = \underline{a}_2'\underline{X} = 0{,}137X_1 + 0{,}216X_2 + 0{,}228X_3 - 0{,}894X_4 - 0{,}223X_5 + 0{,}184X_6$

3-ncü TB: $Y_3 = \underline{a}_3'\underline{X} = 0{,}427X_1 - 0{,}086X_2 - 0{,}117X_3 - 0{,}172X_4 + 0{,}864X_5 + 0{,}141X_6$

4-ncü TB: $Y_4 = \underline{a}_4'\underline{X} = -0{,}443X_1 + 0{,}132X_2 + 0{,}374X_3 - 0{,}175X_4 + 0{,}361X_5 - 0{,}696X_6$

5-nci TB: $Y_5 = \underline{a}_5'\underline{X} = 0{,}163X_1 + 0{,}331X_2 - 0{,}776X_3 - 0{,}164X_4 - 0{,}108X_5 - 0{,}473X_6$

6-ncı TB: $Y_6 = \underline{a}_6'\underline{X} = -0{,}730X_1 + 0{,}231X_2 - 0{,}352X_3 - 0{,}109X_4 + 0{,}238X_5 + 0{,}470X_6$

şeklinde bulunur.

Standart değişkenlerin doğrusal fonksiyonları olarak elde edilen TB'ler: Bu TB'ler örnek kovaryans matrisinin özvektörleri kullanılarak elde edilen TB'ler olup;
$Y_j = \underline{a}_j'\underline{Z} \;\; j = 1, 2, \ldots, 6$ biçimindedir. Buna göre:
1-nci TB: $Y_1 = \underline{a}_1'\underline{Z} = 0{,}510Z_1 + 0{,}561Z_2 + 0{,}464Z_3 + 0{,}143Z_4 + 0{,}108Z_5 + 0{,}423Z_6$
2-nci TB: $Y_2 = \underline{a}_2'\underline{Z} = -0{,}005Z_1 + 0{,}084Z_2 - 0{,}145Z_3 + 0{,}664Z_4 + 0{,}646Z_5 - 0{,}336Z_6$
3-ncü TB: $Y_3 = \underline{a}_3'\underline{Z} = -0{,}448Z_1 + 0{,}323Z_2 + 0{,}473Z_3 + 0{,}313Z_4 - 0{,}466Z_5 - 0{,}394Z_6$
4-ncü TB: $Y_4 = \underline{a}_4'\underline{Z} = -0{,}033Z_1 + 0{,}022Z_2 + 0{,}470Z_3 - 0{,}594Z_4 + 0{,}489Z_5 - 0{,}430Z_6$
5-nci TB: $Y_5 = \underline{a}_5'\underline{Z} = -0{,}625Z_1 - 0{,}221Z_2 + 0{,}306Z_3 + 0{,}090Z_4 + 0{,}312Z_5 + 0{,}601Z_6$
6-ncı TB: $Y_6 = \underline{a}_6'\underline{Z} = -0{,}384Z_1 + 0{,}725Z_2 - 0{,}474Z_3 - 0{,}282Z_4 + 0{,}129Z_5 + 0{,}084Z_6$
şeklinde bulunur.

c) Önemli TB sayısının analitik yöntemlere göre belirlenmesi:
i) Varyans açıklama oranı ile
TB'lerin örnek varyans kovaryans matrisinden veya korelasyon matrisinden elde edilmesi durumuna göre ayrı ayrı değerlendirilecektir; Başlangıç sistemine ait toplam varyans tahmini $\hat{\sigma}^2_{Top} = \text{İ}z(S) = 5{,}7746$ iken, standart değişkenler sistemine göre toplam varyans tahmini $\hat{\sigma}^2_{Top} = \text{İ}z(R) = p = 6$ dır. Böylece her bir TB için varyanslar, varyans açıklama oranlar ile kümülatif varyans açıklama oranları aşağıdaki tabloda verilmiştir.

Başlangıç değişkenlerine göre			**Standart değişkenlere göre**		
$V(Y_j) = \boldsymbol{\lambda_j}$	$VAO(Y_j) = \dfrac{\boldsymbol{\lambda_j}}{\text{İ}z(S)}$	Küm$(VAO(Y_j))$	$V(Y_j) = \boldsymbol{\lambda_j}$	$VAO(Y_j) = \dfrac{\boldsymbol{\lambda_j}}{\text{İ}z(R)}$	Küm$(VAO(Y_j))$
3,3482	0,58	0,58	2,5649	0,43	0,43
1,3788	0,24	0,82	1,3681	0,23	0,66
0,4770	0,08	0,90	0,9352	0,16	0,82
0,3256	0,06	0,96	0,6780	0,11	0,93
0,1555	0,03	0,99	0,3206	0,05	0,98
0,0895	0,01	1,00	0,1332	0,02	1,00

- ✓ Örnek varyans kovaryans matrisine ait bilgi kullanılarak başlangıç değişkenlerinin doğrusal fonksiyonları olarak elde edilen TB'lerden ilk iki tanesi (Y_1 ve Y_2) birlikte toplam varyansın %82'sini açıklamaktadır. Bu sebeple önemli TB sayısı $k = 2$ olarak alınabilir. Bu durumda başlangıç sisteminde 6 değişkenin açıkladığı yapıyı, varyansta %18'lik bir bilgi kaybı ile ilk iki TB'ni kullanarak açıklamak mümkündür.
- ✓ Korelasyon matrisine ait bilgi kullanılarak standart değişkenlerin doğrusal fonksiyonları olarak elde TB'lerden ilk iki tanesi (Y_1 ve Y_2) birlikte toplam varyansın %66'sını açıklamakta iken ilk üç tanesi (Y_1, Y_2 ve Y_3) birlikte toplam varyansın %82'sini açıklamaktadır. Buna göre standart değişkenler sisteminde 6 değişkenin açıkladığı yapıyı, varyansta yine %18'lik bir bilgi kaybı ile ilk üç TB'ni kullanarak açıklamak mümkündür. Böylece ilk üç TB, önemli TB sayısı, yani $k = 3$ alabiliriz.

ii) $\lambda_j > 1$ olan özdeğer sayısı ile
Önemli TB sayısına karar vermede bu kriterin kullanılması, TB'ler korelasyon matrisine ait bilgi kullanılarak elde edildiğinde tercih edilmektedir. Yani TB'ler standart değişkenlerin doğrusal fonksiyonları olarak elde edilmişse önemli TB sayısına karar vermede bu analitik yöntemi kullanabiliriz.

Korelasyon matrisinde bu kritere uyan özdeğer sayısı iki (λ_1 ve λ_2) olduğundan önemli TB sayısı $k = 2$ olarak alınabilir. Ancak bu iki TB toplam varyansın sadece %66'sını karşılamaktadır. Hem toplam varyanstan açıklanan kısmın düşük olması hem de üçüncü özdeğerin bire çok yakın ve toplam varyansı açıklamada %16'lık gibi önemli bir orana sahip olması nedeniyle $k = 3$ olarak güncellenebilir. Bu durumda ilk üç TB birlikte toplam varyansın %82'sini açıklayacaktır.

d) Önemli TB sayısının istatistiksel yöntemlere göre belirlenmesi:
i)<u>Anderson'ın Genel Olabilirlik Oran Testi ile</u>

İlk olarak örnek varyans kovaryans matrisinden elde edilen TB'ler için önemli TB sayısını belirleyelim. $p = 6$ olduğundan kitle kovaryans matrisine ait özdeğerler $\gamma_1 > \gamma_2 > \dots, \gamma_6 \geq 0$ parametreleri olsun. Bu yöntemde test işlemine son iki TB'nin ya da özdeğerin önemliliği test edilerek başlanır ve bu aşamada önemli TB sayısının $k = 4$ olduğu kabul edilmiş olur. Geriye kalan $g = p - k = 2$ özdeğerin/TB'nin önemliliği test edilecektir. Buna göre test edilecek hipotezler;

$H_0: \gamma_5 = \gamma_6 = 0$ (son iki TB önemsiz)
$H_1: \gamma_5 > 0$ (beşinci TB önemli)

şeklinde oluşturulur. Test istatistiği ise;

$$\chi^2 = -(n-1)\sum_{j=k+1}^{k+g} ln\lambda_j + (n-1)g ln\left[\frac{\sum_{j=k+1}^{k+g}\lambda_j}{g}\right] \sim \chi^2_{\frac{(g-1)(g+2)}{2}}$$

dir. Burada $n = 60, k = 4, g = 2, p = k + g = 6$ olup H_0 doğru iken test istatistiğinin alabileceği değer;

$$\chi^2 = -59(ln\lambda_5 + ln\lambda_6) + 59 * 2 * ln\left[\frac{\lambda_5+\lambda_6}{2}\right] = -59[\ln(0{,}1555) + \ln(0{,}0895)] + 118\ln\left[\frac{0{,}1555+0{,}0895}{2}\right] = 4{,}445$$

olur. Kritik değer $\alpha = 0{,}05$ için $\chi^2_t = \chi^2_{\frac{(g-1)(g+2)}{2};\alpha} = \chi^2_{2;0{,}05} = 5{,}991$ olup, $\chi^2 = 4{,}445 < \chi^2_t = 5{,}991$ olduğundan H_0 hipotezi ret edilemez. Buna göre son iki TB kesin olarak önemsizdir.

İkinci aşamada önemli olduğu düşünülen ve ilk 4 TB'nin yer aldığı grupta TB'lerin varyanslarının özdeğerler olması ve özdeğerler arasındaki $\lambda_1 > \lambda_2 > \lambda_3 > \lambda_4$ sıralaması gereğince varyansı λ_4 olan dördüncü TB'nin önemli olup olmadığı test edilir. Bu durumda yeni hipotezler:

$H_0:: \gamma_4 = \gamma_5 = \gamma_6 = 0$ (son üç TB önemsiz)
$H_1: \gamma_4 > 0$ (dördüncü TB önemli)
olup, test istatistiği yine aynıdır.

$$\chi^2 = -(n-1)\sum_{j=k+1}^{k+g} ln\lambda_j + (n-1)g ln\left[\frac{\sum_{j=k+1}^{k+g}\lambda_j}{g}\right] \sim \chi^2_{\frac{(g-1)(g+2)}{2}}$$

dir. Burada $n = 60, k = 3, g = 3, p = k + g = 6$ olup H_0 doğru iken test istatistiğinin alabileceği değer;

$\chi^2 = -59(ln\lambda_4 + ln\lambda_5 + ln\lambda_6) + 59 * 3 * ln\left[\frac{\lambda_4+\lambda_5+\lambda_6}{3}\right] = -59[ln(0{,}3256) +$
$\ln(0{,}1555) + \ln(0{,}0895)] + 177 \ln\left[\frac{0{,}3256+0{,}1555+0{,}0895}{3}\right] = 24{,}643$
olur. Kritik değer $\alpha = 0{,}05$ için $\chi_t^2 = \chi^2_{\frac{(g-1)(g+2)}{2};\alpha} = \chi^2_{5;0{,}05} = 11{,}07$ olup, $\chi^2 = 24{,}643 >$ $\chi_t^2 = 11{,}07$ olduğundan H_0 hipotezi ret edilir. Buna göre dördüncü TB önemli olup önemli TB'ler grubunda kalmalıdır. Burada test işlemi durdurulur.

Sonuç olarak örnek kovaryans matrisine ait bilgi kullanılarak elde edilen TB'lerden ilk dört tanesi (Y_1, Y_2, Y_3, Y_4) %95 güvenle önemli TB'lerdir ve bu istatistiksel yönteme göre önemli temel bileşen sayısı $k = 4$'dür. Böylece bu dört TB toplam varyansın %96'sını açıklayacaktır.

Şimdi bu yöntemi örnek korelasyon matrisinden elde edilen TB'ler için önemli TB sayısını belirlemede uygulayalım. $p = 6$ olduğundan kitle kolerasyon matrisine ait özdeğerler $\gamma_1 >$ $\gamma_2 > \ldots, \gamma_6 \geq 0$ parametreleri olsun. Önce son iki TB'nin ya da özdeğerin önemli olup olmadığını test edelim ve bu aşamada önemli TB sayısının $k = 4$ olduğu kabul edelim. Geriye kalan $g = p - k = 2$ özdeğerin/TB'nin önemliliği test edilecektir. Buna göre test edilecek hipotezler;

$H_0: \gamma_5 = \gamma_6 = 0$ (son iki TB önemsiz)
$H_1: \gamma_5 > 0$ (beşinci TB önemli)

şeklinde oluşturulur. Test istatistiği ise;

$$\chi^2 = -(n-1)\sum_{j=k+1}^{k+g} ln\lambda_j + (n-1)gln\left[\frac{\sum_{j=k+1}^{k+g}\lambda_j}{g}\right] \sim \chi^2_{\frac{(g-1)(g+2)}{2}}$$

dir. Burada $n = 60, k = 4, g = 2, p = k + g = 6$ olup H_0 doğru iken test istatistiğinin alabileceği değer;

$\chi^2 = -59(ln\lambda_5 + ln\lambda_6) + 59 * 2 * ln\left[\frac{\lambda_5+\lambda_6}{2}\right] = -59[\ln(0{,}3206) + \ln(0{,}1332)] +$
$118 \ln\left[\frac{0{,}3206+0{,}1332}{2}\right] = 11{,}03$
olur. Kritik değer $\alpha = 0{,}05$ için $\chi_t^2 = \chi^2_{\frac{(g-1)(g+2)}{2};\alpha} = \chi^2_{2;0{,}05} = 5{,}991$ olup, $\chi^2 = 11{,}03 >$ $\chi_t^2 = 5{,}991$ olduğundan H_0 hipotezi ret edilir. Buna göre ilk dört TB ile birlikte beşinci TB'de önemlidir.

İkinci aşamada altıncı TB'nin de önemli olup olmadığı test edilir. Bu durumda yeni hipotezler:

$H_0 :: \gamma_6 = 0$ (altıncı TB önemsiz)
$H_1: \gamma_6 > 0$ (altıncı TB önemli)
olup, test istatistiği yine aynıdır.

$$\chi^2 = -(n-1)\sum_{j=k+1}^{k+g} ln\lambda_j + (n-1)gln\left[\frac{\sum_{j=k+1}^{k+g}\lambda_j}{g}\right] \sim \chi^2_{\frac{(g-1)(g+2)}{2}}$$

dir. Burada $n = 60, k = 5, g = 1, p = k + g = 6$ olup H_0 doğru iken test istatistiğinin alabileceği değer;

$\chi^2 = -59(ln\lambda_6) + 59 * 1 * ln\left[\frac{\lambda_6}{1}\right] = -59[ln(0{,}1332)] + 59 * 1 * \ln[0{,}1332] = 0$

olur. Kritik değer $\alpha = 0{,}05$ için $\chi_t^2 = \chi^2_{\frac{(g-1)(g+2)}{2};\alpha} = \chi^2_{0;0{,}05} = 0$ olup, $\chi^2 = \chi_t^2 = 0$ olduğundan H_0 hipotezi ret edilemez. Buna göre altıncı TB önemsiz olup önemli TB sayısı $k = 5$'dir. Burada test işlemi durdurulur.

Sonuç olarak örnek kolerasyon matrisine ait bilgi kullanılarak elde edilen TB'lerden ilk beş tanesi $(Y_1, Y_2, Y_3, Y_4, Y_5)$ %95 güvenle önemli TB'lerdir ve bu TB'ler toplam varyansın %98'ini açıklamaktadır.

ii) Büyük Varyanslı TB'lerin Anlamlılık Testi ile

İlk olarak örnek varyans kovaryans matrisinden elde edilen TB'ler için önemli TB sayısını belirleyelim. $p = 6$ olduğundan kitle kovaryans matrisine ait özdeğerler $\gamma_1 > \gamma_2 > \dots, \gamma_6 \geq 0$ parametreleri olsun. Bu yöntemde test işlemine son iki TB'nin ya da özdeğerin önemliliği test edilerek başlanır ve bu aşamada önemli TB sayısının $k = 4$ olduğu kabul edilmiş olur. Geriye kalan $g = p - k = 2$ özdeğerin/TB'nin önemliliği test edilecektir. Buna göre test edilecek hipotezler;

$H_0: \gamma_5 = \gamma_6 = 0$ (son iki TB önemsiz)
$H_1: \gamma_5 > 0$ (beşinci TB önemli)

şeklinde oluşturulur. Test istatistiği ise;

$$U = \left(n - \frac{2p+11}{6}\right)\left(gln\overline{\lambda} - \sum_{j=k+1}^{p} ln\lambda_j\right) \sim \chi^2_{\frac{(g-1)(g+2)}{2}}$$

olup, burada $\overline{\lambda} = \frac{1}{g}\sum_{j=k+1}^{p}\lambda_j$ dir. $n = 60, k = 4, g = 2, p = k + g = 6$ olması sebebiyle $\overline{\lambda} = \frac{\lambda_5+\lambda_6}{2} = \frac{0{,}1555+0{,}0895}{2} = 0{,}1225$ bulunur. Böylece test istatistiğinin alabileceği değer;

$U = (60 - \frac{2*6+11}{6})(2ln(0{,}1225) - (\ln(0{,}1555) + \ln(0{,}0895))) = 4{,}231$ olacaktır. Kritik değer $\alpha = 0{,}05$ için $\chi_t^2 = \chi^2_{\frac{(g-1)(g+2)}{2};\alpha} = \chi^2_{2;0{,}05} = 5{,}991$ olup, $U = 4{,}231 < \chi_t^2 = 5{,}991$ olduğundan H_0 hipotezi ret edilemez. Buna göre son iki TB kesin olarak önemsizdir.

İkinci aşamada önemli olduğu düşünülen ve ilk 4 TB'nin yer aldığı grupta TB'lerin varyanslarının özdeğerler olması ve özdeğerler arasındaki $\lambda_1 > \lambda_2 > \lambda_3 > \lambda_4$ sıralaması gereğince varyansı λ_4 olan dördüncü TB'nin önemli olup olmadığı test edilir. Bu durumda yeni hipotezler:

$H_0:: \gamma_4 = \gamma_5 = \gamma_6 = 0$ (son üç TB önemsiz)
$H_1: \gamma_4 > 0$ (dördüncü TB önemli)

olup, test istatistiği yine aynıdır.

$$U = \left(n - \frac{2p+11}{6}\right)\left(gln\overline{\lambda} - \sum_{j=k+1}^{p} ln\lambda_j\right) \sim \chi^2_{\frac{(g-1)(g+2)}{2}}$$

burada $\overline{\lambda} = \frac{1}{g}\sum_{j=k+1}^{p}\lambda_j$ dir. $n = 60, k = 3, g = 3, p = k + g = 6$ olması sebebiyle $\overline{\lambda} = \frac{\lambda_4+\lambda_5+\lambda_6}{3} = \frac{0,3256+0,1555+0,0895}{3} = 0,1902$ bulunur. Böylece test istatistiğinin alabileceği değer;

$U = (60 - \frac{2*6+11}{6})(3ln(0,1902) - (\ln(0,3256) + \ln(0,1555) + \ln(0,0895))) = 23,459$ olacaktır. Kritik değer $\alpha = 0,05$ için $\chi_t^2 = \chi^2_{\frac{(g-1)(g+2)}{2};\alpha} = \chi^2_{5;0,05} = 11,07$ olup, $U = 23,459 > \chi_t^2 = 11,07$ olduğundan H_0 hipotezi ret edilir. Buna göre dördüncü TB önemli olup önemli TB'ler grubunda kalmalıdır. Burada test işlemi durdurulur.

Sonuç olarak örnek kovaryans matrisine ait bilgi kullanılarak elde edilen TB'lerden ilk dört tanesi (Y_1, Y_2, Y_3, Y_4) %95 güvenle önemli TB'lerdir ve bu istatistiksel yönteme göre önemli temel bileşen sayısı $k = 4$'dür. Böylece bu dört TB toplam varyansın %96'sını açıklayacaktır.

Şimdi bu yöntemi örnek korelasyon matrisinden elde edilen TB'ler için önemli TB sayısını belirlemede uygulayalım. $p = 6$ olduğundan kitle kolerasyon matrisine ait özdeğerler $\gamma_1 > \gamma_2 > \ldots, \gamma_6 \geq 0$ parametreleri olsun. Önce son iki TB'nin ya da özdeğerin önemli olup olmadığını test edelim ve bu aşamada önemli TB sayısının $k = 4$ olduğu kabul edelim. Geriye kalan $g = p - k = 2$ özdeğerin/TB'nin önemliliği test edilecektir. Buna göre test edilecek hipotezler;

$H_0: \gamma_5 = \gamma_6 = 0$ (son iki TB önemsiz)
$H_1: \gamma_5 > 0$ (beşinci TB önemli)

şeklinde oluşturulur. Test istatistiği ise;

$$U = \left(n - \frac{2p+11}{6}\right)\left(gln\overline{\lambda} - \sum_{j=k+1}^{p} ln\lambda_j\right) \sim \chi^2_{\frac{(g-1)(g+2)}{2}}$$

olup, burada $\overline{\lambda} = \frac{1}{g}\sum_{j=k+1}^{p}\lambda_j$ dir. $n = 60, k = 4, g = 2, p = k + g = 6$ olması sebebiyle $\overline{\lambda} = \frac{\lambda_5+\lambda_6}{2} = \frac{0,3206+0,1332}{2} = 0,2294$ bulunur. Böylece test istatistiğinin alabileceği değer;

$U = (60 - \frac{2*6+11}{6})(2ln(0,2294) - (\ln(0,3206) + \ln(0,1332))) = 11,73$ olacaktır. Kritik değer $\alpha = 0,05$ için $\chi_t^2 = \chi^2_{\frac{(g-1)(g+2)}{2};\alpha} = \chi^2_{2;0,05} = 5,991$ olup, $U = 11,73 > \chi_t^2 = 5,991$ olduğundan H_0 hipotezi ret edilir. Buna göre ilk dört TB ile birlikte beşinci TB'de önemlidir.

İkinci aşamada altıncı TB'nin de önemli olup olmadığı test edilir. Bu durumda yeni hipotezler:

$H_0 :: \gamma_6 = 0$ (altıncı TB önemsiz)
$H_1: \gamma_6 > 0$ (altıncı TB önemli)
olup, test istatistiği yine aynıdır.

$$U = \left(n - \frac{2p+11}{6}\right)\left(gln\overline{\lambda} - \sum_{j=k+1}^{p} ln\lambda_j\right) \sim \chi^2_{\frac{(g-1)(g+2)}{2}}$$

dir. Burada $\overline{\lambda} = \lambda_6 = 0,1332$, $n = 60, k = 5, g = 1, p = k + g = 6$ olup H_0 doğru iken test istatistiğinin alabileceği değer;

$$U = (60 - \frac{2*6+11}{6})(1 * \ln(0,1332) - \ln(0,1332))) = 0$$

olur. Kritik değer $\alpha = 0{,}05$ için $\chi_t^2 = \chi^2_{\frac{(g-1)(g+2)}{2};\alpha} = \chi^2_{0;0{,}05} = 0$ olup, $\chi^2 = \chi_t^2 = 0$ olduğundan H_0 hipotezi ret edilemez. Buna göre altıncı TB önemsiz olup, önemli TB sayısı $k = 5$'dir. Burada test işlemi durdurulur.

Sonuç olarak örnek kolerasyon matrisine ait bilgi kullanılarak elde edilen TB'lerden ilk beş tanesi $(Y_1, Y_2, Y_3, Y_4, Y_5)$ %95 güvenle önemli TB'lerdir ve bu TB'ler toplam varyansın %98'ini açıklamaktadır.

iii) Bartlett Testi ile

Bu yöntem korelasyon matrislerine ait bilgi kullanılarak elde edilen TB'lerden önemli olan TB sayısını belirlemede kullanılır. Bu sebeple burada sadece R matrisinden türetilen TB'lerden önemli olan TB sayısı belirlenecektir. $p = 6$ olduğundan kitle kolerasyon matrisine ait özdeğerler $\gamma_1 > \gamma_2 > \ldots, \gamma_6 \geq 0$ parametreleri olsun. Önce son iki TB'nin ya da özdeğerin önemli olup olmadığını test edelim ve bu aşamada önemli TB sayısının $k = 4$ olduğu kabul edelim. Geriye kalan $g = p - k = 2$ özdeğerin/TB'nin önemliliği test edilecektir. Bu varsayıma göre $j = 1, 2, \ldots, 6$ için $R_j = \lambda_j \underline{a}_j \underline{a}_j'$: 6×6 matrisler ve $R_h = \sum_{j=1}^{k} R_j = \sum_{j=1}^{4} R_j$ (hipotez altında yeniden tanımlanan korelasyon matrisi), $R_g = \sum_{j=k+1}^{k+g} R_j = \sum_{j=5}^{6} Rj$ (kalıntı matrisi) olmak üzere , $R = R_h + R_g$ şeklinde yazılabilir. Buna göre test edilecek hipotezler;

H_0: $R_g = [0]$ ($\gamma_5 = \gamma_6 = 0$, yani son iki TB önemsiz)
H_1: $R_g \neq [0]$ ($\gamma_5 > 0$, yani beşinci TB önemli)

şeklinde oluşturulur. Test istatistiği ise;

$$\chi^2 = -\left[(n-1) - \frac{2p+5}{6} \times \frac{2k}{3}\right] ln\left(U_g\right) \sim \chi^2_{\frac{(g-1)(g+2)}{2}}$$

eşitliği ile verilir. Burada; $n = 60, p = 6, k = 4, g = 2$ ve

$$U_g = \frac{|R| g^g}{\left(\prod_{j=1}^{k} \lambda_j\right)\left(p - \sum_{j=1}^{k} \lambda_j\right)^g} = \frac{(0{,}095)2^2}{(2{,}5649*1{,}3681*0{,}9352*0{,}6780)(6-(2{,}5649+1{,}3681+0{,}9352+0{,}6780))} = 0{,}369$$

olup, yerlerine yazılırsa;

$\chi^2 = -\left(59 - \frac{17}{6} * \frac{8}{3}\right) \ln(0{,}369) = 51{,}288$ bulunur. Kritik değer $\alpha = 0{,}05$ için $\chi_t^2 = \chi^2_{\frac{(g-1)(g+2)}{2};\alpha} = \chi^2_{2;0{,}05} = 5{,}991$ olup, $\chi^2 = 51{,}288 > \chi_t^2 = 5{,}991$ olduğundan H_0 hipotezi ret edilir. Buna göre ilk dört TB ile birlikte beşinci TB'de önemlidir.

İkinci aşamada altıncı TB'nin de önemli olup olmadığı test edilir. Bu durumda yeni hipotezler:

H_0: $R_g = [0]$ ($\gamma_6 = 0$, yani altıncı TB önemsiz)
H_1: $R_g \neq [0]$ ($\gamma_6 > 0$, yani altıncı TB önemli)

olup, test istatistiği yine aynıdır.

$$\chi^2 = -\left[(n-1) - \frac{2p+5}{6} \times \frac{2k}{3}\right] ln\left(U_g\right) \sim \chi^2_{\frac{(g-1)(g+2)}{2}}$$

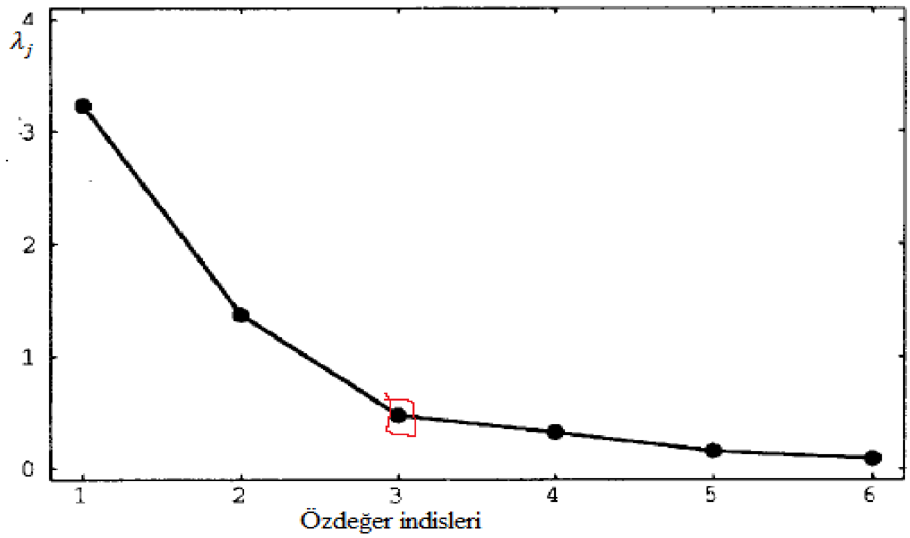
Burada; $n = 60, p = 6, k = 5, g = 1$ ve

$$U_g = \frac{|R|g^g}{\left(\prod_{j=1}^{k}\lambda_j\right)\left(p-\sum_{j=1}^{k}\lambda_j\right)^g} = \frac{(0{,}095)*1}{(2{,}5649*1{,}3681*0{,}9352*0{,}6780*0{,}3206)(0{,}1332)} = 1$$

olup, yerlerine yazılırsa; $\chi^2 = -\left(59 - \frac{17}{6} * \frac{10}{3}\right)\ln(1) = 0$ bulunur. Kritik değer $\alpha = 0{,}05$ için $\chi_t^2 = \chi^2_{\frac{(g-1)(g+2)}{2};\alpha} = \chi^2_{0;0{,}05} = 0$ olup, $\chi^2 = \chi_t^2 = 1$ olduğundan H_0 hipotezi ret edilemez. Buna göre altıncı TB önemsiz olup, önemli TB sayısı $k = 5$'dir. Burada test işlemi durdurulur.

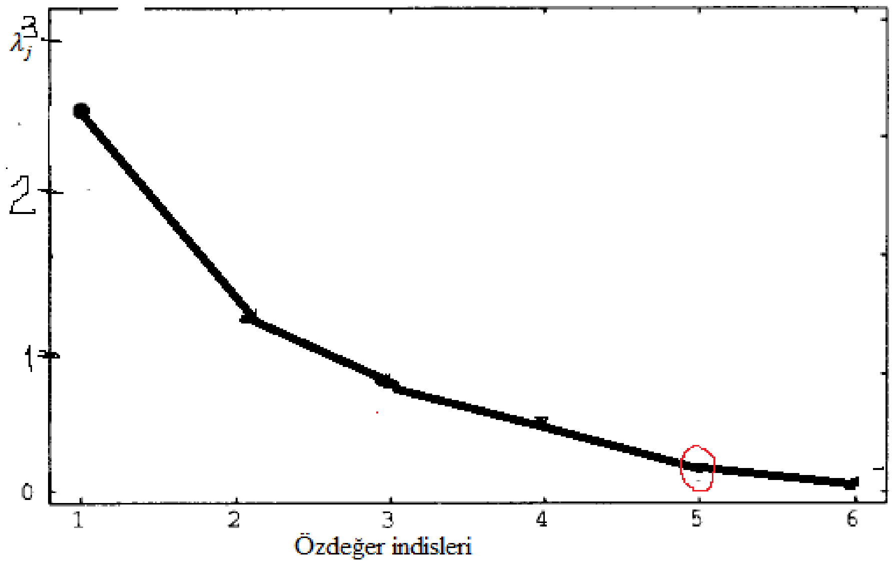
Sonuç olarak örnek kolerasyon matrisine ait bilgi kullanılarak elde edilen TB'lerden ilk beş tanesi $(Y_1, Y_2, Y_3, Y_4, Y_5)$ %95 güvenle önemli TB'lerdir ve bu TB'ler toplam varyansın %98'ini açıklamaktadır.

e) Önemli TB sayısının grafik yöntemle belirlenmesi: Başlangıç sistemi için Scree graph (Dik-Yamaç)



grafiğinde dirsek noktası olarak hızlı düşüşün veya azalmanın son bulduğu nokta özdeğer indisi 3'e karşılık gelen $(3, \lambda_3) = (3\,;\ 0{,}4770)$ noktası alınabilir. Bu durumda ilk üç özdeğer ve böylece ilk üç TB önemli kabul edilir. O halde grafik yönteme göre kovaryans matrisinden elde edilen TB'ler için önemli TB sayısı $k = 3$ olup, bunlar (Y_1, Y_2, Y_3) TB'leridir.

Standart değişkenler sistemi için Scree graph (Dik-Yamaç) grafiğinde dirsek noktası olarak hızlı düşüşün veya azalmanın son bulduğu nokta özdeğer indisi 5'e karşılık gelen $(5, \lambda_5) = (5\,;\ 0{,}3206)$ noktası alınabilir. Bu durumda ilk beş özdeğer ve böylece ilk beş TB önemli kabul edilir. O halde grafik yönteme göre korelasyon matrisinden elde edilen TB'ler için önemli TB sayısı $k = 5$ olup, bunlar $(Y_1, Y_2, Y_3, Y_4, Y_5)$ TB'leridir.



f) Önemli olan TB'ler ile X_t değişkenlerinin korelasyonları: Analitik yöntem ve grafik yönteminde kovaryans matrisinden elde edilen önemli TB'ler Y_1, Y_2, Y_3 olduğundan bu TB'lerden Y_j TB'i ile X_t değişkeni arasındaki korelasyon;

$$r_{Y_j, X_t} = \frac{\sqrt{\hat{\lambda}_j} * \hat{a}_{jt}}{s_t} \quad , j = 1, 2, 3; \; t = 1, 2, \ldots, 6$$

ile bulunur. S matrisinin ilk üç özdeğeri ve bunlara karşılık gelen özvektörleri ile başlangıç değişkenleri ve onlara ait standart sapmalar aşağıdaki tabloda verilmiştir.

$\boldsymbol{\lambda_j}$	$\underline{\boldsymbol{a}}_1$	$\underline{\boldsymbol{a}}_2$	$\underline{\boldsymbol{a}}_3$	X_t	s_t
3,3482	0,205	0,137	0,427	X_1	0,608
1,3788	0,875	0,216	-0,086	X_2	1,631
0,477	0,262	0,228	-0,117	X_3	0,677
	0,320	-0,894	-0,172	X_4	1,214
	0,063	-0,223	0,864	X_5	0,699
	0,128	0,184	0,141	X_6	0,569

Önemli TB'ler ile X_t değişkenleri arasındaki korelasyonlar, bu tablodaki bilgiler yardımıyla hesaplanarak aşağıdaki tabloda verilmiştir.

	X_1	X_2	X_3	X_4	X_5	X_6
Y_1	0,617	0,982	0,708	0,482	0,165	0,412
Y_2	0,265	0,156	0,395	-0,865	-0,375	0,380
Y_3	0,485	-0,036	-0,119	-0,098	0,854	0,171

Şimdide önemli olan TB'ler ile Z_t değişkenlerinin korelasyonları bulalım: Analitik yöntemlerde korelasyon matrisinden elde edilen önemli TB'ler Y_1, Y_2, Y_3 olduğundan bu TB'lerden Y_j TB'i ile X_t değişkeni arasındaki korelasyon;

$$r_{Y_j,X_t} = \sqrt{\hat{\lambda}_j} * \hat{a}_{jt}, \qquad j = 1,2,3;\ t = 1,2,\dots,6$$

ile bulunur. R matrisinin ilk üç özdeğeri ve bunlara karşılık gelen özvektörleri ile Standart değişkenler aşağıdaki tabloda verilmiştir.

$\boldsymbol{\lambda_j}$	$\underline{\boldsymbol{a}}_{\mathbf{1}}$	$\underline{\boldsymbol{a}}_{\mathbf{2}}$	$\underline{\boldsymbol{a}}_{\mathbf{3}}$	Z_t
2,5649	0,510	-0,005	-0,448	Z_1
1,3681	0,561	0,084	0,323	Z_2
0,9352	0,464	-0,145	0,473	Z_3
	0,143	0,664	0,313	Z_4
	0,108	0,646	-0,466	Z_5
	0,423	-0,336	-0,394	Z_6

Önemli TB'ler ile Z_t değişkenleri arasındaki korelasyonlar, bu tablodaki bilgiler yardımıyla hesaplanarak aşağıdaki tabloda verilmiştir.

	Z_1	Z_2	Z_3	Z_4	Z_5	Z_6
Y_1	0,817	0,898	0,743	0,229	0,173	0,677
Y_2	-0,006	0,098	-0,170	0,777	0,756	-0,393
Y_3	-0,433	0,312	0,457	0,303	-0,451	-0,381

g) Başlangıç verileri (veya örnek kovaryans matrisi) için V.A.O'nına göre önemli olduğuna karar verilen TB'ler için TB skorlarını bulalım. Bu durumda önemli olan TB sayısı $k = 2$ olup, bu TB'ler Y_1 ve Y_2 idi. Bu TB'ler için temel bileşen skorları aşağıdaki tabloda verilmiştir.

1-nci TB: $Y_1 = \underline{a}_1'\underline{X} = 0{,}205X_1 + 0{,}875X_2 + 0{,}262X_3 + 0{,}320X_4 + 0{,}063X_5 + 0{,}128X_6$

2-nci TB: $Y_2 = \underline{a}_2'\underline{X} = 0{,}137X_1 + 0{,}216X_2 + 0{,}228X_3 - 0{,}894X_4 - 0{,}223X_5 + 0{,}184X_6$

$\boldsymbol{Y_1}$	$\boldsymbol{Y_2}$	$\boldsymbol{Y_1}$	$\boldsymbol{Y_2}$	$\boldsymbol{Y_1}$	$\boldsymbol{Y_2}$
66,9331	9,9795	61,8960	8,4906	64,5425	8,8939
67,0904	6,9675	62,8466	9,8635	62,5988	8,7322
66,7105	8,4600	63,2475	10,2315	67,1243	8,1243
63,4430	7,7605	61,8877	10,6188	64,5422	9,1249
64,3191	10,2335	63,5508	9,0386	65,9236	8,4939
63,7091	9,1130	63,2323	9,3070	63,8185	8,6842
64,9502	9,5966	65,3244	11,1394	63,6318	9,0863
64,6439	10,9513	64,1159	9,6196	65,2321	8,2587
63,9338	8,5888	64,6996	10,5360	66,8449	7,7734
63,4252	8,4047	69,7583	10,8633	65,4261	7,8755
64,4829	8,3193	62,5598	11,5899	64,7158	7,2377
63,1146	10,0075	64,2799	8,6274	64,3853	8,5147
64,8884	8,0263	62,0558	8,9184	64,1768	7,8096
63,8103	9,2483	62,1286	10,4516	66,2886	9,5914
61,7513	7,0851	63,7964	7,4829	64,3295	8,9545
66,4704	10,4352	65,4330	6,5712	67,8121	8,0844
62,8088	9,0073	63,6466	8,4979	61,9086	7,5310
61,8509	8,2465	62,5992	7,3680	66,0682	6,4602
62,7977	6,9931	65,4952	8,1031	66,5049	8,1665
63,3392	9,1647	66,9358	9,5074	69,3057	9,2596

Standart verlerden (veya korelasyon matrisi için V.A.O'nına veya önemli özdeğer sayısına göre önemli olduğuna karar verilen TB'ler için TB skorlarını bulalım. Bu durumda önemli olan TB sayısı $k=3$ olup, bu TB'ler Y_1, Y_2 ve Y_3 idi. Bu TB'ler için temel bileşen skorları aşağıdaki tabloda verilmiştir.

1-nci TB: $Y_1 = \underline{a}_1'\underline{Z} = 0{,}510Z_1 + 0{,}561Z_2 + 0{,}464Z_3 + 0{,}143Z_4 + 0{,}108Z_5 + 0{,}423Z_6$

2-nci TB: $Y_2 = \underline{a}_2'\underline{Z} = -0{,}005Z_1 + 0{,}084Z_2 - 0{,}145Z_3 + 0{,}664Z_4 + 0{,}646Z_5 - 0{,}336Z_6$

3-ncü TB: $Y_3 = \underline{a}_3'\underline{Z} = -0{,}448Z_1 + 0{,}323Z_2 + 0{,}473Z_3 + 0{,}313Z_4 - 0{,}466Z_5 - 0{,}394Z_6$

Y_1	Y_2	Y_3	Y_1	Y_2	Y_3	Y_1	Y_2	Y_3
2,0593	-,7418	1,0805	-1,4098	,3900	-1,3703	-,4876	,1422	,6789
,7650	2,5135	,9988	-1,3004	-1,3326	-,1473	-2,0139	-,1360	,5567
1,0187	,8650	,7112	-,5715	-,3707	-,3834	1,4766	1,0195	1,2553
-2,4791	,7894	,9631	-,6997	-2,2057	-1,7721	,1616	,0965	-,3308
-,8592	-2,7212	2,3398	-1,5858	,5208	1,2931	,8366	,6569	1,2826
-1,7175	-,4942	-,0206	-1,6657	-,5146	,2075	-,3867	-,1994	-,2415
,6057	-1,0940	1,2693	1,2825	-1,2685	,3890	-,4486	-,0438	-,7996
1,4899	-2,1228	-1,0695	,3967	-,3431	-,7310	,6021	1,6331	,0873
-,3458	,3511	-,4280	1,7063	-1,2905	-2,4618	2,0926	1,7416	-,8351
-,7072	,7115	-,7234	5,0957	-,5747	,0056	,4844	,7782	,0478
,4484	-,1581	-,0104	-,8280	-2,4870	,1363	-,1254	1,9752	-1,2375
-,4065	-1,6820	-,2657	-,5647	,4838	-,0508	,4632	,3605	-,6780
,6035	-,5721	,3810	-1,7480	-,4592	-,4754	-,4803	1,0754	,2019
-,5687	-1,3803	1,3178	-2,0595	-2,4096	1,1359	1,9226	-,1801	,2418
-2,6641	1,0641	-,5970	,1422	1,4326	-1,2944	-,3330	,2778	-,2437
2,0963	-,1094	-1,0182	-,5566	1,4280	1,8282	2,2327	,6613	1,4586
-,7068	-,4717	-1,8330	-,2705	-,8004	-1,1537	-2,3394	1,3977	-,7654
-1,7620	,5550	-1,2608	-2,4214	1,1712	-,1352	,5974	2,4164	,3283
-2,0442	,0879	1,2139	,1478	,4968	,2514	1,7765	-,2075	,7693
-,6146	-,3737	,1390	1,9469	-,0157	,0955	4,7208	-,3322	-,3318