

YILDIZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

CEBİRSEL ÖZDEĞER PROBLEMLERİNDE
PERTURBASYON TEORİSİ

139708-

Matematik Müh. Abdullah KÖKSAL

FBE Matematik Mühendisliği Anabilim Dalı Matematik Mühendisliği Programında
Hazırlanan

YÜKSEK LİSANS TEZİ

Tez Danışmanı : Prof. Tahir ŞİŞMAN

Prof. Dr. Mehmet AHLATGİOĞLU

Prof. Dr. Mustafa SIVRI

Y.Ö. YÜKSEK LİSANS TEZİ
FEN BİLİMLERİ ENSTİTÜSÜ
MATHEMATICS

İSTANBUL, 2003

İÇİNDEKİLER

	Sayfa
SİMGE LİSTESİ.....	iv
ÖNSÖZ.....	v
ÖZET.....	vi
ABSTRACT.....	vii
1 GİRİŞ.....	1
2 TANIMLAR.....	2
2.1 Matris Tanımları.....	2
2.2 Transpoze Matrislerin Özdeğerleri ve Özvektörleri	3
2.3 Ayrık Özdeğerler.....	5
2.4 Benzerlik Dönüşümleri.....	7
2.5 Katlı Özdeğerler ve Genel Matrisler için Kanonik Formlar.....	9
2.6 Özvektörlerin Eksik (Defective) Sistemi.....	10
2.7 Jordan (Klasik) Kanonik Form.....	12
2.8 Temel Bölenler.....	13
2.9 Bir Matrisin Karakteristik Polinomunun Eşlik (Companion) Matrisi.....	14
2.10 Derogatory Olmayan Matrisler.....	15
2.11 Frobenius (Rasyonel) Kanonik Form.....	17
2.12 Jordan ve Frobenius Kanonik Form Arasındaki İlişki.....	18
2.13 Denk Dönüşümler.....	19
2.14 Lamda Matrisleri.....	20
2.15 Temel İşlemler.....	21
2.16 Smith's Kanonik Form.....	21
2.17 Bir Lamda Matrisinin k satırlı Minörlerinin OBEB 'i.....	24
2.18 $(A-\lambda I)$ nın İnvariant Çarpanları.....	25
2.19 Üçgen Kanonik Form.....	26
2.20 Hermitian ve Simetrik Matrisler.....	27
2.21 Hermitian Matrislerin Temel Özellikleri.....	28
2.22 Kompleks Simetrik Matrisler.....	30
2.23 Uniter Dönüşümler ile Üçgen Forma Dönüşümler.....	30
2.24 Kuadratik Formlar.....	31
2.25 Pozitif Tanımlılık için Gerek ve Yeter Şartlar.....	32
2.26 Bir Vektörün Minimum Polinomu.....	33
2.27 Bir Matrisin Minimum Polinomu.....	34
2.28 Cayley-Hamilton Teoremi.....	36
2.29 Minimum Polinom ve Kanonik Form Arasındaki İlişki.....	37
2.30 Genelleştirilmiş Vektörler.....	40
2.31 Temel Benzerlik Dönüşümleri.....	41
2.32 Elementer Matrislerin Özellikleri.....	43
2.33 Temel Benzerlik Dönüşümleri ile Üçgen Kanonik Forma İndirgeme.....	43
2.34 Temel Uniter Dönüşümler.....	45
2.35 Temel Uniter Hermitian Matrisleri.....	46
2.36 Temel Uniter Dönüşümler ile Üçgen Forma İndirgeme.....	48
2.37 Normal Matrisler.....	50
2.38 Değişmeli Matrisleri.....	51

2.39	AB -nin Özdeğerleri.....	53
2.40	Vektör ve Matris Normları.....	54
2.41	Matris Normları.....	55
2.42	Euclidian ve Spectrallar.....	56
2.43	Normlar ve Limitler.....	58
2.44	Sonsuz Matris Serilerinden Kaçınma.....	61

3. PERTURBASYON

TEORİSİ.....	62
3.1 Perturbasyon Teorisinin Temelleri	62
3.2 Özdeğerlerin Sürekliliği Üzerine Ostrowski Teoremi.....	63
3.3 Cebirsel Fonksiyonlar.....	64
3.4 Nümerik Örnekler.....	66
3.5 Basit Özdeğerler için Perturbasyon Teorisi.....	67
3.6 Özvektörlerin Perturbasyonu.....	68
3.7 Lineer Temel-Bölenlere Sahip Matrisler.....	69
3.8 Özdeğerlerin Birinci Mertebeden Perturbasyonu.....	70
3.9 Özdeğerlerin Birinci Mertebeden Perturbasyonu.....	71
3.10 Yüksek Mertebeden Perturbasyonlar.....	72
3.11 Katlı Özdeğerler.....	73
3.12 Gerschgorin Teoremleri.....	73
3.13 Gerschgorin Teoremleri Üzerine Kurulu Perturbasyon Teorisi.....	75
3.14 Lineer Olmayan Bölenlerin Genel Dağılımına Tekabül eden Perturbasyon.....	84
3.15 Jordan Kanonik Formundan Özdeğerlerin Perturbasyon Teoremi.....	84
3.16 Katlı Özdeğerlere Karşılık Gelen Özvektörlerin Perturbasyonu (Lineer Temel-Bölenler).....	86
3.17 Perturbasyon Teoreminin Sınırları.....	87
3.18 s_i -ler Arasındaki İlişki.....	88
3.19 Hesaplana bir Problemin Kondüsyonu.....	89
3.20 Kondüsyon Sayıları.....	90
3.21 Özdeğer Problemine Karşılık Gelen Spectral Kondüsyon Sayısı.....	90
3.22 Spectral Kondüsyon Sayısının Özellikleri.....	92
3.23 Kondüsyon Sayılarının Değişmezlik Özelliği.....	93
3.24 Çok Kötü Kondüsyonlu Matrisler.....	94
3.25 Reel Simetrik Matrisler için Perturbasyon Teorisi.....	97
3.26 Simetrik-Olmayan Perturbasyonlar.....	97
3.27 Simetrik Perturbasyonlar.....	98
3.28 Klasik Teknikler.....	99
3.29 Uniter Ranklı Simetrik Matris.....	102
3.30 Özdeğerlerin Ektremal Özellikleri.....	103
3.31 Özdeğerlerin Minimax Karakterizasyonu.....	104
3.32 İki Simetrik Matrisin Toplamının Özdeğerleri.....	107
3.33 Pratik Uygulamalar.....	108
3.34 Minimax Prensibinin Diğer Uygulamaları.....	109
3.35 Ayırma Teoremi.....	110

KAYNAKLAR.....	111
----------------	-----

ÖZGEÇMİŞ.....	112
---------------	-----

SİMGE LİSTESİ

λ	Matris özdeğeri
I	Birim matris
A	$n \times m$ boyutlu matris
H	Hermitian matris
A^H	A' 'nın Hermitian matrisi
A^T	A' 'nın Transpoze matrisi
$\det(A)$	A matrisinin determinantı
$\text{diag}(A)$	Diagonal A matrisi
$\ A\ $	A matrisinin normu



ÖNSÖZ

Tez danışmanım Prof. Dr. Tahir ŞİŞMAN' a ve çalışmamda büyük katkıları olan Prof. Dr. Abdullah YILDIZ' a teşekkürü bir borç bilirim.



ÖZET

Özdeğer ve özvektörlerin hesaplanması esnasında ortaya çıkan hatalar, perturbasyon teoremleri ile incelenmiştir. Ayrıca bu yöntemler ile ulaşılabilecek en iyi sınırlar üzerinde de çalışılmıştır.

Perturbasyon teoremleri içerisinde 3 ana metod incelenmiştir. Bunlar Ostrowski teoremi, Basit Perturbasyon teorisi ve Gerschgorin teoremi dir.

Ostrowski Teoremi özdeğerlerin perturbasyonunun üst sınırı üzerinde inceleme yapmaktadır. Fakat bu yöntem, ancak bazı kısıtlı koşullar altında iyi sonuçlar vermektedir. Basit Perturbasyon teorisi matrisin karakteristik denklemi üzerine yakınsak kuvvet serilerinin uygulanması üzerine, Gerschgorin teoremi ise Jordan kanonik form üzerine kuruludur.

Perturbasyon teoremlerinin sınırlarının incelenmesiyle birlikte kondüsyon sayısı belirlenmektedir. Spectral kondüsyon sayısı adı da verilen bu sayı ile matrisler iyi konumlu kötü konumlu ve çok kötü konumlu matrisler olarak sınıflandırılmaktadır.

Anahtar Kelimeler: Matris, Özdeğer, Özvektör, Özproblem, Perturbasyon.

ABSTRACT

The errors observed in calculations of eigenvalues and eigenvectors are examined by perturbation theorems. Also, the upper boundaries that can be reached through the methods have been studied.

Three basic methods of perturbation theorems have been analyzed. These are: Ostrowski's theorem, Simple perturbation theorem and Gerschgorin's theorem.

Ostrowsky's theorem studies on upper boundaries of the eigenvalues' perturbation. But, this method can provide satisfactory solutions only under some limited conditions. Simple perturbation theorem is based on application of convergent power series on matrix' characteristic equation. Gerschgorin theorem is structured on Jordan's canonical form. The condition number is determined through examination of perturbation theorem limits.

The number which is also named as the spectral condition number classifies whether the matrixes are well conditioned, ill conditioned and very ill conditioned, or not.

Keywords: Matrix, Eigenvalue, Eigenvector, Eigenproblem, Perturbation



1. GİRİŞ

Son yıllarda matris yöntemleri üzerinde yapılan çalışmaların fizik, mekanik, yöneylem ve ekonomi gibi uygulamalı bilimlerde verimli bir ortam sağlaması, matrislerin özdeğer ve özvektörlerinin önemini artırmıştır. Bununla birlikte özdeğer ve özvektörlerin hesaplama yöntemleri üzerinde araştırma alanları genişlemiş ve bunların hesaplanması esnasında ortaya çıkan hataların, özproblem üzerindeki etkilerini inceleyen yeni yöntemler ortaya konulmuştur.

Bu kitap, matrislerin özdeğer ve özvektörlerinin hata analizinin, perturbasyon yöntemleri ile incelenmesi üzerinde durmaktadır. İki bölümden oluşan kitabın ilk bölümünde temel matris tanımları ve işlemleri üzerinde durulmuştur. İkinci bölümde ise özdeğer ve özvektörün hesaplanması esnasında ortaya çıkan hataların kaynağı ve bunların Perturbasyon Teorisi ile ilişkilendirilmesi üzerinde durulmuştur. Ayrıca üç ana Perturbasyon Teorisi ile bu hataların özdeğer ve özvektörler üzerindeki etkileri incelenmiştir.



2. TANIMLAR

2.1 Matris Tanımları

Temel cebirsel özdeğer problemi, sıfır çözümü olmayan n homojen lineer denklem sistemi

$$Ax = \lambda x \quad (2.1)$$

in λ değerlerinin bulunması işlemidir. Bu denklem

$$(A - \lambda I)x = 0 \quad (2.2)$$

formunda da yazılabilir. Keyfi λ değerleri için bu denklem sistemi yalnızca $x=0$ çözümüne sahiptir.

Bu denklemde bulunan $(A - \lambda I)$ matrisinin *tekil-olmaması* (*non-singular*) durumunda, aşikar-olmayan (*non-trivial*) bir çözümü vardır.

$$\det(A - \lambda I) = 0 \quad (2.3)$$

Denklem (2.3) ün sol tarafının determinantı açık (*explicit*) *polinom* denklemini vermesi için aşağıdaki şekilde yazılabilir:

$$\alpha_0 + \alpha_1 \lambda + \dots + \alpha_{n-1} \lambda^{n-1} + (-1)^n \lambda^n = 0 \quad (2.4)$$

Denklem (2.4) A matrisinin karakteristik denklemi ve denklemin sol tarafı da karakteristik polinom olarak adlandırılır. λ^n başkatsayısının sıfırdan farklı olması ve kompleks düzlemde çalıştığımızı kabul edersek, denklem daima n köke sahiptir. Genelde A matrisi reel olmasına rağmen kökler kompleks olabilir ve ayrıca kökler n sayısına kadar herhangi bir katlılıkta olabilir.

Bu n kök, A 'nın özdeğerleri (*eigenvalues*, *latent roots*, *characteristic values* veya *proper values*) adını alır. Herhangi bir λ özdeğerine karşılık olarak (2.2) denklemi en az bir tane *sıfır-olmayan* x vektör çözümü vardır. Bu vektöre, λ değerine karşılık gelen özvektör (*eigenvector*, *latent vector*, *characteristic vector* veya *proper vector*) adını alır. Eğer $(A - \lambda I)$ matrisinin *rankı* $(n-1)$ den küçük ise, (2.2) denklemini sağlayan birden fazla bağımsız vektör vardır. Eğer x,

(2.2)' nin bir çözümü ise, herhangi bir k değeri için kx de (2.2)'nin bir çözümüdür. Ayrıca, $(A-\lambda I)$ rankı $(n-1)$ olduğunda λ -ya bağlı özvektör de sabit bir çarpana göre keyfidir. Bu çarpan, özvektörün bazı istenen nümerik özelliklere sahip olması için, uygun bir şekilde seçilir. Böyle vektörler *normalize* edilmiş vektör olarak adlandırılır. En çok kullanılan normalizasyon formları aşağıdaki gibidir:

- (i) Bileşenlerinin *modüllerinin* karelerinin toplamı, 1'e eşit olan vektörler,
- (ii) En büyük bileşenin modülünün 1'e eşit olan vektörler,
- (iii) Bileşenlerinin modüllerinin toplamı 1'e eşit vektörler.

(i) ve (iii) deki normalizasyon, keyfi birim kompleks çarpanına açılımında dezavantajları vardır. Buradan itibaren anlatılan çoğu işlemlerde özvektörler, normalize edilmemiş durumdadırlar. Bu tür vektörler, en büyük elemanın modüllerinin 0,1 ve 1,0 arasında olması için, 10 (veya 2) un kuvveti ile çarpılır. Bu şekilde muamele edilen vektörleri *standart hale getirilmiş vektörler* olarak adlandıracğız. (Prager, 1988)

2.2 Transpoze Matrislerin Özdeğerleri ve Özvektörleri

Bir matrisin transpozisinin özdeğer ve özvektörlerinin Genel Teoride önemli bir yeri vardır. A^T -nin özdeğerleri, sıfır-olmayan çözüme sahip A nın λ özdeğerlerine eşittir.

$$A^T y = \lambda y \quad (2.5)$$

$$\det(A^T - \lambda I) = 0 \quad (2.6)$$

Ve bir matrisin determinantı, transpozisinin determinantına eşit olduğundan, A^T -nin özdeğerleri A -nin özdeğerlerine eşittir. Fakat genellikle özvektörler farklıdır. Bu sebeple, A^T -nin λ_i ye karşılık gelen özvektörlerini y_i ile temsil edeceğiz. Böylelikle aşağıdaki denklem yazılabilir:

$$A^T y_i = \lambda_i y_i \quad (2.7)$$

Denklem (2.5) aşağıdaki formda da yazılabilir:

$$y_i^T A = \lambda_i y_i^T \quad (2.8)$$

Öyle ki bazen y_i^T , A -nın λ_i özdeğerine karşılık gelen *sol-özvektörü* (*left-hand*), aşağıdaki gibi x_i vektöründe λ_i -ye karşılık gelen A nın *sağ-özvektörü* (*right-hand*) olarak adlandırılır.

$$Ax_i = \lambda_i x_i \quad (2.9)$$

Transpoze matrisin özdeğerlerinin önemi esas olarak aşağıda belirtilmektedir.

Eğer x_i , A -nın λ_i -ye karşılık gelen bir özvektörü ise ve y_j A^T -nin λ_j -ye karşılık gelen özvektörü ise, aşağıdaki denklem yazılır:

$$x_i^T y_j = 0 \quad (\lambda_i \neq \lambda_j) \quad (2.10)$$

(2.9) denklemi,

$$x_i^T A^T = \lambda_i x_i^T \quad (2.11)$$

şeklinde ifade edilirse, aşağıdaki denklem elde edilir:

$$A^T y_j = \lambda_j y_j \quad (2.12)$$

(2.10)' nin y_j ile sağdan çarpımının (*Post-multiplying*) ve (2.12) in x_i^T ile soldan çarpımının (*Pre-multiplying*) farkı ile aşağıdaki denklem elde edilir

$$0 = \lambda_i x_i^T y_j - \lambda_j x_i^T y_j \quad (2.13)$$

x_i ve/veya y_j kompleks vektör olabileceğinden $x_i^T y_j$ bir *iç-çarpım* (*inner-product*) değildir. Eğer (2.14) denklemi elde ediliyor ise, (2.15) denklemi elde edilemez.

$$x_i^T y_j = y_j^T x_i \quad (2.14)$$

$$x_i^T y_j = \overline{(y_j^T x_i)} \quad (2.15)$$

Eğer x kompleks ise (2.16) denklemini elde edilir.

$$x^T x = 0 \quad (2.16)$$

Halbuki asıl *iç-çarpım* her sıfır olmayan x vektörü için sıfır dan büyüktür.

2.3 Ayrık-Özdeğerler

n boyutlu bir A matrisinin n özdeğerinin de ayrık olması durumunda, özsystemlerin genel teorisi en basit haldedir ve ilk olarak bu durumu inceleyeceğiz. Özdeğerler $\lambda_1, \lambda_2, \dots, \lambda_n$ ile temsil edeceğiz. Denklem (2.2)'nin gerçekte, her bir λ_i değeri için en az bir çözümü vardır ve böylece, $x_1, x_2, \dots, x_n$ ile temsil edilen özvektör kümesinin varlığını kabul edebiliriz. Keyfi çarpanlar dışında bu vektörlerin her birinin *tek* olduğunu gösterelim.

İlk olarak vektörlerin lineer bağımsız olmaları gerektiğini gösterelim. r , lineer-bağımlı vektör sayısı olsun. Vektörleri $x_1, x_2, \dots, x_n$ şeklinde numaralandırabiliriz. Bu durumda aşağıdaki ilişki sağlanır:

$$\sum_{i=1}^r \alpha_i x_i = 0 \quad (2.17)$$

Burada α_i -lerin hiç biri sıfır değildir. Tanımdan x_i -lerin hiç biri *sıfır-vektör* olmadığından, $r > 2$ olduğu açıktır. Denklem (2.17) in A ile soldan çarpım ile aşağıdaki denklem elde edilir.

$$\sum_{i=1}^r \alpha_i \lambda_i x_i = 0 \quad (2.18)$$

Denklem (2.17) in, λ_r ile çarpımı ve (2.18)' nin farkı ile aşağıdaki denklem elde edilir:

$$\sum_{i=1}^{r-1} \alpha_i (\lambda_r - \lambda_i) x_i = 0 \quad (2.19)$$

Buradaki başkatsayıların hiçbiri sıfır değildir. Denklem (2.19), $x_1, x_2, \dots, x_{r-1}$ in lineer bağımlı olduğunu gösterir ki bu hipotez ile çelişkidir. Böylelikle n özvektör bağımsızdır ve *n-boyutlu uzayı* germişlerdir. Sonuç olarak bunlar keyfi herhangi bir vektörü ifade etmek için kullanılabilirler.

Elde edilen bu sonuçtan, özdeğerlerin tekliği görülür. Eğer λ_1 e karşılık gelen z_1 gibi ikinci bir özvektör var ise, aşağıdaki şekilde ifade edilebilir

$$z_1 = \sum_1^n \alpha_i x_i \quad (2.20)$$

Burada α_i -lerden en az biri sıfırdan farklıdır. (2.20) denkleminin A ile çarpımı aşağıdaki denklemi verir:

$$\lambda_1 z_1 = \sum_1^n \alpha_i \lambda_i x_i \quad (2.21)$$

(2.20) denkleminin λ_1 ile çarpımı aşağıdaki denklemi verir:

$$0 = \sum_2^n \alpha_i (\lambda_i - \lambda_1) x_i \quad (2.22)$$

ve x_i -lerin *lineer-bağımsız* olmalarından dolayı aşağıdaki denklem elde edilir:

$$\begin{aligned} \alpha_i (\lambda_i - \lambda_1) &= 0 & (i=2,3,\dots,n) \\ \alpha_i &= 0 & (i=2,3,\dots,n) \end{aligned} \quad (2.23)$$

sıfır olmayan α_1 ise, z_1 in x_1 in katı olması demektir.

Benzer olarak A^T -nin özvektörlerinin *tek* ve *lineer-bağımsız* olduklarını gösterebiliriz. Daha önce aşağıdaki ifade ispat edilmişti.

$$x_i^T y_j = 0 \quad (\lambda_i \neq \lambda_j) \quad (2.24)$$

Öyle ki iki vektör kümesi de *bi-ortogonal system* dir. Buradan aşağıdaki denklemi elde edilir:

$$x_i^T y_i \neq 0 \quad (i=1,2,\dots,n) \quad (2.25)$$

x_i eğer y_i -ye ortogonal ise $y_1, y_2, \dots, y_n$ vektörlerine de ortogonal olmalı ve böylece tüm n -boyutlu uzaya ortogonal olmalı. Bu, x_i -ler *sıfır-uzayı*na ait olmadığından dolayı imkansızdır.

Denklem (2.24) ve (2.25) x_i veya y_i terimleriyle ifade edilen keyfi bir v vektörünün açık bir ifadesini elde etmemize imkan verir. Eğer

$$v = \sum_{i=1}^n \alpha_i x_i \quad (2.26)$$

yazılırsa aşağıdaki denklem elde edilir

$$y_i^T v = \sum_{i=1}^n \alpha_i y_j^T x_i = \alpha_j y_j^T x_j \quad (2.27)$$

Böleni sıfır olmadığından dolayı her durumda

$$\alpha_j = y_j^T v / y_j^T x_j \quad (2.28)$$

denkleme ulaşılır.

2.4 Benzerlik Dönüşümleri

Eğer x_i ve y_j -deki keyfi çarpanları aşağıdaki gibi seçersek

$$y_i^T x_i = 1 \quad (i=1,2,\dots,n) \quad (2.29)$$

(4.20) ve (2.19) denklemleri, i. satırı y_i^T olan Y^T matrisinin, i. kolonu x_i olan X matrisinin tersi olduğunu gösterir. (2.30) denkleminin n tanesi (2.31) denklemindeki gibi matris formunda yazılabilir.

$$A x_i = \lambda_i x_i \quad (2.30)$$

$$A X = X \text{diag}(\lambda_i) \quad (2.31)$$

Buradan X in tersinin mevcut ve Y^T ye eşit olduğunu görürüz. Denklem (2.31), aşağıdaki denklemi verir:

$$X^{-1} A X = Y^T A X = \text{diag}(\lambda_i) \quad (2.32)$$

A matrisinin *tekil-olmayan* bir H matrisi ile, $H^{-1} A H$ dönüşümü teorik ve pratik açıdan temel bir öneme sahiptir. A ile $H^{-1} A H$ *benzer* şeklinde ifade edilirken, bu dönüşüm *benzerlik-dönüşümü* (*similarity transforms*) olarak adlandırılır. Ayrıca $H A H^{-1}$ inde A nın *benzerlik-*

dönüşümü olduğu açıktır. A'nın özdeğerleri ayrık olduğu zaman, A -yı köşegen forma dönüştüren bir *benzerlik-dönüşümünün* var olduğunu ve bu *dönüşüm-matrisinin*, A'nın özvektörlerine eşit sayıda kolona sahip olduğunu gösterdik. Karşıt olarak eğer (2.33) denklemi elde edilebilir ise, (2.34) denklemi de elde edilir.

$$H^{-1}AH = \text{diag}(\mu_i) \quad (2.33)$$

$$AH = H \text{diag}(\mu_i) \quad (2.34)$$

Son denklem μ_i -nin A'nın bir özdeğeri olduğunu ve H'nin i. kolonunun μ_i ye karşılık gelen öz vektörü olduğunu gösterir.

Bir matrisin özdeğerleri, bir benzerlik-dönüşümü altında *invariant*tir.

$$Ax = \lambda x \quad (2.35)$$

$$H^{-1}Ax = \lambda H^{-1}x$$

$$H^{-1}A(HH^{-1})x = \lambda H^{-1}x \quad (2.36)$$

$$(H^{-1}AH)H^{-1}x = \lambda H^{-1}x$$

Özdeğerlerde böylece korunur ve özvektörler H^{-1} ile çarpılır.

Çoğu nümerik metotlar, genel bir forma sahip herhangi bir A matrisinin öz sistemlerine daha kolay ulaşmak için, özel bir forma sahip olan öz sistemi daha kolay bulunabilir B matrisine indirgeme işlemini göz önünde bulundurur.

Benzerlik, *geçişli (transitive)* özelliğe sahiptir.

$$B = H_1^{-1}AH_1 \quad \text{ve} \quad C = H_2^{-1}BH_2 \quad (2.37)$$

$$C = H_2^{-1}H_1^{-1}AH_1H_2 = (H_1H_2)^{-1}A(H_1H_2) \quad (2.38)$$

Bir matrisin özel bir forma indirgenmesi, basit *benzerlik-dönüşüm* dizisi ile olmaktadır.

2.5 Katlı Özdeğerler ve Genel Matrisler için Kanonik Formlar

Bir yada daha fazla katlı özdeğerlere sahip bir matris için özvektörler sisteminin yapısı ayrık özdeğerli matrislerin yapısından biraz daha karışıktır. A matrisini köşegen forma dönüştüren benzerlik-dönüşümü hala geçerli olabilir. Eğer bu geçerli ise, tekil-olmayan H matris için aşağıdaki denkleme sahip oluruz

$$H^{-1}AH = \text{diag}(\lambda_i) \quad (2.39)$$

Bu durumda λ_i , A'nın uygun bir katlılıkta özdeğerleri olmalıdır.

$$H^{-1}(A - \lambda I)H = H^{-1}AH - \lambda H^{-1}IH = \text{diag}(\lambda_i - \lambda) \quad (2.40)$$

Her iki tarafın determinantını alarak aşağıdaki denkleme ulaşılır:

$$\det(H^{-1})\det(A - \lambda I)\det(H) = \prod(\lambda_i - \lambda) \quad (2.41)$$

$$\det(A - \lambda I) = \prod(\lambda_i - \lambda) \quad (2.42)$$

(6.1) denklemini aşağıdaki formda yazılırsa

$$AH = H\text{diag}(\lambda_i) \quad (2.43)$$

H'nin kolonlarının A'nın özvektörlerine eşit olduğu görülür. H tekil olmadığından, kolonları bağımsızdır. Örneğin, λ_1 çift kat bir kök ise bu durumda H'nin ilk iki kolonu x_1 ve x_2 ile ifade edilir.

$$Ax_1 = \lambda_1 x_1, \quad Ax_2 = \lambda_1 x_2 \quad (2.44)$$

Burada x_1 ve x_2 bağımsızdır. Denklem (2.44) dan dolayı, x_1 ve x_2 tarafından gerilmiş bir altuzaydaki herhangi bir vektör de özvektördür.

$$A(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 \lambda_1 x_1 + \alpha_2 \lambda_1 x_2 = \lambda_1 (\alpha_1 x_1 + \alpha_2 x_2) \quad (2.45)$$

Hala *n-boyutlu uzayı* geren özvektörler seçilebilir ve keyfî bir vektörü ifade etmek için baz olarak kullanılabilirler.

Aşağıdaki gibi ifade edilen köşegen bir Λ matrisi e_1, e_2, e_3, e_4, e_5 gibi 5 bağımsız özvektöre sahiptir.

$$\Lambda = \begin{vmatrix} 1 & & & & \\ & 1 & & & \\ & & 2 & & \\ & & & 2 & \\ & & & & 3 \end{vmatrix} \quad (2.46)$$

e_1 ve e_2 -nin herhangi bir kombinasyonu, $\lambda=1$ e karşılık gelen bir özvektördür ve e_3 ile e_4 ün herhangi bir kombinasyonu $\lambda=2$ ye karşılık gelen bir özvektördür. Λ matrisine benzer herhangi bir matris için, aşağıdaki gibi tekil-olmayan bir H matrisi vardır

$$H^{-1}AH = \Lambda \quad (2.47)$$

ve A nın özvektörleri $He_1, He_2, He_3, He_4, He_5$ dir. Burada ilk ikisinin herhangi bir bileşimi $\lambda=1$ e karşılık gelen bir özvektör iken 3. ile 4. nün herhangi bir bileşimi $\lambda=2$ ye karşılık gelen özvektördür.

2.6 Özvektörlerin Eksik (Defective) sistemi

Bir köşegen matrise benzeyen fakat katlı özdeğerlere sahip bir matrisin özellikleri ayrıntı özdeğerli bir matrisin özelliklerine benzer. Bununla birlikte katlı özdeğere sahip matrislerin hepsi bu tipte değildir.

$$C_2(a) = \begin{vmatrix} a & 1 \\ 0 & a \end{vmatrix} \quad (2.48)$$

(7.1) gibi tanımlanan $C_2(a)$ matrisi $\lambda=a$ çift katlı özdeğere sahiptir. Eğer x, x_1 ve x_2 bileşenleri ile $C_2(a)$ nın bir özvektörü ise bu durumda (2.49) denklemini elde ederiz

$$\begin{aligned} 0x_1 + x_2 &= 0 \\ 0x_1 + 0x_2 &= 0 \end{aligned} \quad (2.49)$$

Bu denklemlerden $x_2=0$ sonucuna ulaşılır. Böylece $\lambda=a$ -ya karşılık gelen e_1 gibi sadece bir tane özvektör vardır. $C_2(a)$, $b \rightarrow a$ limitinde aşağıdaki gibi tanımlanan bir B_2 matrisi şeklinde düşünülebilir.

$$B_2 = \begin{vmatrix} a & 1 \\ 0 & b \end{vmatrix} \quad (2.50)$$

$b \neq a$ için bu matris a ve b gibi iki özdeğere sahiptir ve bunlara karşı gelen özvektörlerde aşağıdaki gibidir.

$$\begin{vmatrix} 1 \\ 0 \end{vmatrix} \text{ ve } \begin{vmatrix} 1 \\ b-a \end{vmatrix} \quad (2.51)$$

$b \rightarrow a$ limitinde 2 özdeğer ve özvektör çakışırlar.

Daha genel olarak eğer $C_r(a)$ -yı aşağıdaki gibi tanımlanırsa

$$C_1(a) = [a], \quad C_r(a) = \begin{vmatrix} a & 1 & & & \\ & a & 1 & & \\ & & a & 1 & \\ . & . & . & . & . \\ & & & a & 1 \\ & & & & a & 1 \end{vmatrix} \quad (2.52)$$

bu durumda $C_r(a)$, r katlı bir özdeğer olan a -ya sahip olurken, sadece bir tane $x=e_1$ özvektörüne sahip olur. Bu $[C_r(a)-aI]$ -nin rankının $(r-1)$ olması gerçeğinden gelir. (Sağ üst köşede bulunan $(r-1)$ mertebeden matrisin determinantı 1'e eşittir.). Ayrıca transpoze matris $[C_r(a)]^T$ sadece bir özvektöre, $y=e_r$, sahiptir. Buradan, ayrı özdeğerli matrislerin özvektörleri için elde edilen çözüme karşıt olarak, doğrudan aşağıdaki denklem elde edilir:

$$x^T y = e_1^T e_r = 0 \quad (2.53)$$

$C_r(a)$ matrisi, $(r>1)$, benzerlik-dönüşümüyle köşegen matrise dönüştürülemez. Eğer (2.54) ve (2.55) denklemlerini sağlayan bir H matrisi var ise λ_i , $C_r(a)$ nın katlı özdeğeri a -ya eşit olmalıdır.

$$H^{-1} C_r H = \text{diag}(\lambda_i), \quad (2.54)$$

$$C_r H = H \text{diag}(\lambda_i), \quad (2.55)$$

H tekil değildir, çünkü bu kolonlar lineer bağımsızdır. Böyle bir H matrisinin var olmadığı görülür.

$C_r(a)$ matrisi r. mertebeden basit *Jordan-altmatrisi* (*simple Jordan submatrix*) olarak veya r. mertebeden basit *klasik-altmatris* (*simple classical submatrix*) olarak adlandırılır.

2.7 Jordan (Klasik) Kanonik Form

Katlı özdeğerlere sahip bir matrisin *köşegen-matrise* benzemediği görülmektedir. A matrisini dönüştürebilecek, benzerlik-dönüşümüyle örtüşen daha yakın bir formun yapısı izleyen teoremden elde edilir.

A, n. mertebeden, $m_1, m_2, \dots, m_r$ katlı $\lambda_1, \lambda_2, \dots, \lambda_r$ gibi r adet ayrık özdeğeri olan ve (8.1) denklemini sağlayan bir matris olsun.

$$m_1 + m_2 + \dots + m_r = n \quad (2.56)$$

Bu durumda benzerlik-dönüşümünü sağlayan öyle bir H matrisi vardır ki, $H^{-1}AH$ matrisi, köşegen boyunca izole edilmiş *basit Jordan-altmatris*lere sahip geri kalan yerlerde sıfır olan bir formdadır.

λ_i ile ilişkili altmatrislerin mertebelerinin toplamı m_i ye eşittir. $H^{-1}AH$, *basit-Jordan alt-matrislerinin direk-toplamı* olarak ifade edilebilir.

Böylece beş özdeğeri λ_1 e, bir özdeğeri λ_2 ye eşit olan (2.57) de tanımlanan C matrisine benzer altıncı mertebeden bir matris, beşi λ_1 ve biri λ_2 olan köklere sahip (2.58) de tanımlanan matrise benzemez.

$$C = \begin{vmatrix} C_2(\lambda_1) & & & \\ & C_2(\lambda_1) & & \\ & & C_1(\lambda_1) & \\ & & & C_1(\lambda_2) \end{vmatrix} \quad (2.57)$$

$$\begin{vmatrix} C_3(\lambda_1) & & \\ & C_2(\lambda_1) & \\ & & C_1(\lambda_2) \end{vmatrix} \quad (2.58)$$

Basit Jordan-altmatrislerden oluşan matrise Jordan-kanonik form veya A matrisinin klasik-kanonik formu (classical canonical form) olarak adlandırılır.

Özel bir durum olarak eğer *Jordan-kanonik form* içindeki *Jordan-altmatrisler* birinci mertebesinde iseler, bu formlar köşegen hale gelirler. Eğer bir matrisin özdeğerleri ayrık ise, *Jordan-kanonik form* $\text{diag}(\lambda_i)$ olan bir matris olduğu daha önceden gösterilmişti.

Bağımsız özvektörlerin toplam sayısı *Jordan-kanonik form* içindeki altmatrislerin sayısına eşittir. Böylelikle, (2.57) denkleminde tanımlanan, C matrisine dönüşebilen bir A matrisi dört bağımsız özvektöre sahiptir. C -nin özvektörleri, e_1, e_3, e_5 ve e_6 dir ve A -daki karşılıkları ise He_1, He_3, He_5 ve He_6 dir. Burada 5 katlı λ_1 özdeğerine karşılık sadece 3 özvektör vardır.

İspatın metodu, daha sonraki bölümlerde anlatacağımız metotlarla pek bir ilgisi olmadığından, *Jordan-kanonik formunun* varlığını ve *tekliğini* (uniqueness) ispat etmeyeceğiz. r. mertebeden basit *Jordan-altmatrislerinde*, süper-köşegen üzerinde yayılması için 1'e eşit elemanlar aldık. (2.51) denklem formuna *üst-Jordan form* ve tersine de *alt-Jordan form* olarak adlandırılırlar. Bu elemanların birim değerlerinin seçiminin özel bir önemi bulunmamaktadır. Daha açık olarak, bir *köşegen-matris* ile uygun bir benzerlik-dönüşümüyle bu elemanlara keyfî *sıfır-olmayan* değerler verilebilir.

2.8 Temel Bölenler

C, A nın *Jordan-kanonik formu* olduğu durumda, $(C-\lambda I)$ matrisini ele alalım. Örnek olarak, (2.57) denklemindeki C matrisine karşılık gelen aşağıdaki matrisi inceleyelim

$$(C-\lambda I) = \begin{vmatrix} C_2(\lambda_1-\lambda) & & & \\ & C_2(\lambda_1-\lambda) & & \\ & & C_1(\lambda_1-\lambda) & \\ & & & C_1(\lambda_2-\lambda) \end{vmatrix} \quad (2.59)$$

Genelde, $(C-\lambda I)$, $C_r(\lambda_i-\lambda)$ formunun matrislerinin *direk-toplamıdır*. $(C-\lambda I)$ *kanonik-form* unun altmatrislerinin determinantları, A matrisinin *temel-bölenleri* (elementary divisors) olarak adlandırılır. Böylelikle, (2.57) denklemindeki C matrisine benzer herhangi bir matrisin temel-bölenleri, aşağıdaki gibidir:

$$(\lambda_1 - \lambda)^2, (\lambda_1 - \lambda)^2, (\lambda_1 - \lambda), (\lambda_2 - \lambda) \quad (2.60)$$

Bir matrisin temel-bölenlerinin çarpımı, onun *karakteristik-polinomuna* eşittir. Eğer *Jordan-kanonik form* köşegen ise, bu durumda temel-bölenler aşağıdaki gibidir ve hepside *lineer* dir.

$$(\lambda_1 - \lambda), (\lambda_2 - \lambda), \dots, (\lambda_n - \lambda) \quad (2.61)$$

Ayrık özdeğerli bir matris daima *lineer temel-bölenlere* sahiptir. Fakat, bir matris bir yada daha fazla katlı özdeğerlere sahipse o zaman matris *lineer temel-bölenlere* sahip olabilirde olmayabilirde.

Eğer bir matris bir yada daha fazla *lineer-olmayan temel-bölenlere* sahip ise bunun manası, matrisin *Jordan-kanonik formunun* içindeki altmatrislerin en az bir tanesi *1'den* daha büyük mertebededir. Böylece bu matris *n* bağımsız özvektörden daha az özvektöre sahiptir. Böyle matrislere *eksik-matris (defective-matris)* adı verilir.

2.9 Bir Matrisin Karakteristik Polinomunun Eşlik (Companion) Matrisi

A'nın karakteristik denklemi aşağıdaki şekilde ifade edilsin:

$$(-1)^n [\lambda^n - p_{n-1} \lambda^{n-1} - p_{n-2} \lambda^{n-2} - \dots - p_0] = 0 \quad (2.62)$$

Bu durumda aşağıdaki C matrisinin karakteristik denkleminin, A'nın karakteristik denklemine özdeş olduğu kolaylıkla görülür.

$$C = \begin{vmatrix} p_{n-1} & p_{n-2} & \dots & p_1 & p_0 \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & 0 \end{vmatrix} \quad (2.63)$$

C matrisi, A matrisinin karakteristik polinomunun *eşlik-matrisi (companion matrix)* olarak adlandırılır. C ve A aynı karakteristik polinoma sahip olduğundan genelde bunların benzer olup olmadığını sormak oldukça doğaldır. Bunun cevabı bir sonraki paragrafta verilecektir. Bu *eşlik-matrisini* 3 farklı yapı içerisinde verebiliriz. Örneğin, aşağıdaki denkleme sahip olabiliriz:

$$\begin{pmatrix} 0 & 0 & 0 & \dots & 0 & p_0 \\ 1 & 0 & 0 & \dots & 0 & p_1 \\ 0 & 1 & 0 & \dots & 0 & p_2 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & 0 & \dots \\ 0 & 0 & 0 & \dots & 1 & p_{n-1} \end{pmatrix} \quad (2.64)$$

Veya p_i elemanı, ilk kolon veya son satır içerisinde, süper-köşegen üzerindeki *birim* elemanlar ile birlikte alınabilir. Bu tipteki matrisleri *eşlik-matrisi* olarak adlandırılır.

2.10 Derogatory Olmayan Matrisler

A matrisinin $\lambda_1, \lambda_2, \dots, \lambda_n$ gibi ayrık özdeğerlere sahip olduğunu kabul edelim. Bu durumda A matrisinin $\text{diag}(\lambda_i)$ ye benzer olduğunu biliyoruz. Şimdi $\text{diag}(\lambda_i)$ -nin *eşlik-matrise* benzer olduğunu gösterelim. Bunu C yi $\text{diag}(\lambda_i)$ ye dönüştüren matrisin sergilenmesiyle ispat edeceğiz. Genel durum, $n=4$ hali ile yeterince gösterilmiş olur.

Aşağıdaki şekilde tanımlanan bir H matrisini göz önüne alalım:

$$H = \begin{pmatrix} \lambda_1^3 & \lambda_2^3 & \lambda_3^3 & \lambda_4^3 \\ \lambda_1^2 & \lambda_2^2 & \lambda_3^2 & \lambda_4^2 \\ \lambda_1 & \lambda_2 & \lambda_3 & \lambda_4 \\ 1 & 1 & 1 & 1 \end{pmatrix} \quad (2.65)$$

Eğer λ_i -ler farklı ise bu matris *tekil-olmayan* bir matristir. Buradan aşağıdaki denklem elde edilir:

$$H \text{diag}(\lambda_i) = \begin{pmatrix} \lambda_1^4 & \lambda_2^4 & \lambda_3^4 & \lambda_4^4 \\ \lambda_1^3 & \lambda_2^3 & \lambda_3^3 & \lambda_4^3 \\ \lambda_1^2 & \lambda_2^2 & \lambda_3^2 & \lambda_4^2 \\ \lambda_1 & \lambda_2 & \lambda_3 & \lambda_4 \end{pmatrix} \quad (2.66)$$

$$\lambda_i^4 = p_3 \lambda_i^3 + p_2 \lambda_i^2 + p_1 \lambda_i + p_0 \quad (2.67)$$

olduğundan

$$\begin{vmatrix} p_3 & p_2 & p_1 & p_0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{vmatrix} H = \begin{vmatrix} \lambda_1^4 & \lambda_2^4 & \lambda_3^4 & \lambda_4^4 \\ \lambda_1^3 & \lambda_2^3 & \lambda_3^3 & \lambda_4^3 \\ \lambda_1^2 & \lambda_2^2 & \lambda_3^2 & \lambda_4^2 \\ \lambda_1 & \lambda_2 & \lambda_3 & \lambda_4 \end{vmatrix} \quad (2.68)$$

ilişkisi bulunur. Buradan da

$$CH = H \text{ diag}(\lambda_i), \quad (2.69)$$

elde edilir. C matrisinin $\text{diag}(\lambda_i)$ ye benzer olduğu görülür.

Şimdi katlı özdeğerlere sahip olduğumuz duruma dönelim. Eğer A, λ_i ye karşılık gelen birden fazla bağımsız özvektöre sahip ise bu durumda C ye benzer olmadığını görürüz. Sol alt köşedeki $(n-1)$ mertebeden matrisin determinantı 1 olduğundan, $(C-\lambda I)$ 'nin rankı $(n-1)$ dir. λ_i ye karşılık gelen x_i özvektörü, aşağıdaki gibi verilir:

$$x_i^T = [\lambda_i^{n-1}, \lambda_i^{n-2}, \dots, \lambda_i, 1] \quad (2.70)$$

$$(C - \lambda_i I) x_i = 0 \quad (2.71)$$

Bir matrisin, *karakteristik-polinomunun eşlik-matrisine* benzemesi için, o matrisin *Jordan-kanonik formunun* sadece bir tane, her bir ayrık özdeğer ile bağlantılı *Jordan-altmatrislerine* sahip olması gerekmektedir. Oldukça etkili olan bu koşulu şimdi ispat edelim. Bunu basit bir durum ile göstermek yeterli olacaktır.

$(\lambda_1 - \lambda)^2$, $(\lambda_3 - \lambda)$, $(\lambda_4 - \lambda)$ *temel-bölenlere* sahip 4. merteben bir matrisi göz önünde alalım. Böyle bir A_4 matrisi aşağıdaki gibi verilir

$$A_4 = \begin{vmatrix} \lambda_1 & 1 & & \\ 0 & \lambda_1 & & \\ & & \lambda_3 & \\ & & & \lambda_4 \end{vmatrix} \quad (2.72)$$

Eğer (2.73) deki gibi tanımlanan H matrisi $\lambda_1, \lambda_3, \lambda_4$ ayrık özdeğerlere sahip ise, *tekil-olmayan* bir matristir. Böylece (2.74) denklemi elde edilir.

$$H = \begin{vmatrix} \lambda_1^3 & 3\lambda_1^2 & \lambda_3^3 & \lambda_4^3 \\ \lambda_1^2 & 2\lambda_1 & \lambda_3^2 & \lambda_4^2 \\ \lambda_1 & 1 & \lambda_3 & \lambda_4 \\ 1 & 0 & 1 & 1 \end{vmatrix} \quad (2.73)$$

$$HA_4 = \begin{vmatrix} \lambda_1^4 & 4\lambda_1^3 & \lambda_3^4 & \lambda_4^4 \\ \lambda_1^3 & 3\lambda_1^2 & \lambda_3^3 & \lambda_4^3 \\ \lambda_1^2 & 2\lambda_1 & \lambda_3^2 & \lambda_4^2 \\ \lambda_1 & 1 & \lambda_3 & \lambda_4 \end{vmatrix} \quad (2.74)$$

$\lambda_1, f(\lambda)=0$ karakteristik denklemin çift kat kökü olduğundan, $f'(\lambda_1)=0$ elde edilir.

$$4\lambda_1^3 = 3p_3\lambda_1^2 + 2p_2\lambda_1 + p_1 \quad (2.75)$$

$$CH = HA_4 \quad (2.76)$$

Benzer bir yol olarak, n. mertebeden bir matris ile bağlantılı r. mertebeden bir *Jordan-altmatris*lerini inceleyebiliriz. Eğer λ_1 ilgili özdeğer ise H, r kolona sahiptir.

$$\begin{vmatrix} \lambda_1^{n-1} & \binom{n-1}{1} \lambda_1^{n-2} & \dots & \binom{n-1}{r-1} \lambda_1^{n-r} \\ \lambda_1^{n-2} & \binom{n-2}{1} \lambda_1^{n-3} & & \binom{n-2}{r-1} \lambda_1^{n-r-1} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_1 & 1 & \dots & 0 \\ 1 & 0 & \dots & 0 \end{vmatrix} \quad (2.77)$$

Böylece her ayrıık λ_i ye karşılık gelen sadece bir tane Jordan matrisinin var olduğu, her bir ayrıık λ_i ye karşılık gelen özvektörlerinde sadece bir tane olduğu ve bu matrisin *karakteristik-*

polinomunun, eşlik-matrisine benzeyen matris olduğu ispat edilmiş olur. Böyle matrisler *derogatory olmayan matris* olarak adlandırılırlar.

2.11 Frobenius (Rasyonel) Kanonik Form

Herhangi bir i için λ_i ile bağlantılı birden fazla *Jordan-altmatris* (ve böylece birden fazla özvektör) var ise, böyle matrisler *derogatory* olarak adlandırılırlar.

Bir *derogatory-matrisin*, onun karakteristik polinomunun eşlik-matrisine benzer olmadığını gördük. Şimdi de kanonik formun, bir *derogatory matris* için eşlik-matrise benzer olup olmadığını araştıralım.

r . mertebeden bir *Frobenius-matris* B_r -yi

$$B_r = \begin{vmatrix} b_{r-1} & b_{r-2} & \dots & b_1 & b_0 \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & 1 & 0 \end{vmatrix} \quad (2.78)$$

formunda tanımlayalım. Genel Matrisler için Temel Teorem aşağıdaki gibi verilir:

Her bir A matrisi, benzerlik-dönüşümü ile, $B_{r1}, B_{r2}, \dots, B_{rs}$ şeklinde ifade edilebilen s -tane Frobenius matrisinin *direk-toplamına* indirgenebilir. Her bir B_{ri} matrisinin *karakteristik-polinomu*, tüm önceki B_{rj} -lerin *karakteristik-polinomu* bölür. Bir *derogatory olmayan matris* durumunda, $s=1$ ve $r_1=n$ olur. Frobenius matrislerinin *direk-toplamı* *Frobenius-kanonik form* (*rational canonical form*) olarak adlandırılır. *Rasyonel kanonik form* ismi, A nın elemanlarının cisminde rasyonel olan dönüşümler tarafından elde edilebilir olması gerçeğinden gelmektedir.

Burada *Frobenius-kanonik formun* varlığını gösteren ispat verilmeyecektir. Bunun yerine, *Jordan-kanonik formu* ile olan ilişkisi gösterilecektir.

2.12 Jordan ve Frobenius Kanonik Form Arasındaki İlişki

İki kanonik form arasındaki ilişkiyi bir örnek ile gösterelim. 10. mertebeden olan aşağıdaki gibi *temel-bölenlere* sahip bir A matrisini ele alalım.

$$\begin{pmatrix} (\lambda_1 - \lambda)^3 & (\lambda_2 - \lambda)^2 & (\lambda_3 - \lambda) \\ (\lambda_1 - \lambda) & (\lambda_2 - \lambda)^2 & \\ (\lambda_1 - \lambda) & & \end{pmatrix} \quad (2.79)$$

Bunun *Jordan-kanonik formu* aşağıdaki gibi yazılır:

$$\begin{vmatrix} C_3(\lambda_1) & & \\ & C_2(\lambda_2) & \\ & & C_1(\lambda_3) \end{vmatrix} = \begin{vmatrix} G_1 & & \\ & G_2 & \\ & & G_3 \end{vmatrix} \quad (2.80)$$

G_1 matrisi olarak her özdeğerdeki en yüksek mertebeden *Jordan-kanonik altmatrisleri* beraberce gruplandırıldı. Diğer bir sonraki en büyük mertebe G_2 ve G_3 v.s.

Bu nedenle her bir G_i matrisi, sadece bir tane λ_i ile bağlantılı *Jordan-altmatrise* sahiptir. Ve bu nedenle, benzerlik-dönüşümü H_i tarafından, uygun mertebeli tek bir Frobenius matrisine indirgenebilir. Böylece örnek aşağıdaki gibi olur.

$$\begin{aligned} H_1^{-1} G_1 H_1 &= B_6 \quad (6. \text{ mertebeden Frobenius matris}) \\ H_2^{-1} G_2 H_2 &= B_3 \quad (3. \text{ mertebeden Frobenius matris}) \\ H_3^{-1} G_3 H_3 &= B_1 \quad (1. \text{ mertebeden Frobenius matris}) \end{aligned} \quad (2.81)$$

Böylece

$$\begin{vmatrix} H_1^{-1} & & \\ & H_2^{-1} & \\ & & H_3^{-1} \end{vmatrix} \begin{vmatrix} G_1 & & \\ & G_2 & \\ & & G_3 \end{vmatrix} \begin{vmatrix} H_1 & & \\ & H_2 & \\ & & H_3 \end{vmatrix} = \begin{vmatrix} B_6 & & \\ & B_3 & \\ & & B_1 \end{vmatrix} \quad (2.82)$$

elde edilir.

B_6 nın karakteristik polinomu $(\lambda_1 - \lambda)^3(\lambda_2 - \lambda)^2(\lambda_3 - \lambda)$, B_3 ün karakteristik polinomu $(\lambda_1 - \lambda)(\lambda_2 - \lambda)^2$ ve B_1 in karakteristik polinomu $(\lambda_1 - \lambda)$ şeklindedir. Bu polinomların her biri, onun bir önkisiminin çarpanıdır. *Frobenius-kanonik form* içindeki altmatrislerin *karakteristik-polinomlarının* çarpanları bir matrisin temel-bölenleri olduğu aşıkardır. Eğer A matrisi reel ise, bu durumda herhangi bir kompleks özdeğer, *eşlenik-çifti (conjugate pairs)* olarak meydana

gelir. Böylece her bir G_i açık olarak reel olacaktır. Bu durum, *Frobenius-kanonik formun*, A cisminin rasyonel dönüşümleriyle elde edilir.

2.13 Denk Dönüşümler

Bir matrisin özdeğerlerinin benzerlik-dönüşümü altında invariant olduğu daha önce görülmüştü. Şimdi kısaca Genel bir dönüşüm sınıfını göz önüne alalım. Eğer tekil-olmayan (kare matrisler) aşağıdaki gibi P ve Q matrisleri mevcut ise, B matrisi A matrisine *denk-matris* denir.

$$PAO = B \quad (2.83)$$

A ve B matrislerinin kare olmasına gerek yoktur ancak aynı boyutta olması gerekmektedir. *Binet-Cauchy* teoreminden, denk denklemlerin aynı ranklı olması gerekmektedir.

Lineer denklem takımını çözmede bizim verdiğimiz denklik kavramı temeldir.

Eğer x

$$Ax = b \quad (2.84)$$

denklemini sağlarsa, açıkça P uygun bir matris olmak üzere

$$PAx = Pb \quad (2.85)$$

denklemini de sağlar. Ayrıca, eğer P *tekil-olmayan* bir matris ise (2.85) ün çözümü (2.84) - ninde çözümüdür. ((2.85)ün P^{-1} ile çarpımıyla görülebileceği gibi.)

Şimdi Q -nun *tekil-olmayan* bir matris olduğu zaman aşağıdaki gibi değişken değişimi yapalım.

$$x = Q y \quad (2.86)$$

Bu durumda aşağıdaki denklemi elde ederiz:

$$PAQy = Pb \quad (2.87)$$

Ve (2.87) de, y -nin herhangi bir çözümünden, (2.84) deki x in bir çözümünü veya tersini türetiriz. Böylece (2.84) ve (2.87) denklemler eşdeğer denklemler olarak göz önünde bulundurulurlar.

2.14 Lamda Matrisleri

Şimdi denklik kavramını, elemanları λ_i -nin polinomları olan matrislere genişletelim. Böyle matrisler λ -matrisleri olarak adlandırılırlar. Eğer $A(\lambda)$, derecesi k dan büyük olmayan polinom elemanlarından oluşan $(m \times n)$ -lik bir matris ise, bu durumda

$$A(\lambda) = A_k \lambda^k + A_{k-1} \lambda^{k-1} + \dots + A_0 \quad (2.88)$$

yazılabilir. Burada $A_k, A_{k-1}, \dots, A_0$, elemanları λ -ya bağımlı olmayan $(m \times n)$ boyutlu matrislerdir. Bir λ -matrisinin rankı, λ nın *sıfır-olmayan* polinomundan olan minörlerinin en yüksek derecesidir.

Eğer, (2.89) de olduğu gibi, λ dan bağımsız *sıfır-olmayan* determinanta sahip $P(\lambda)$ ve $Q(\lambda)$ gibi kare matrisler varsa, $A(\lambda)$ ve $B(\lambda)$ λ -matrisleri *denk-matris (equivalent-matrix)* olarak adlandırılırlar.

$$P(\lambda)A(\lambda)Q(\lambda) = B(\lambda) \quad (2.89)$$

$P(\lambda)$, λ dan bağımsız *sıfır-olmayan* determinanta sahip ise. bu durumda $[P(\lambda)]^{-1}$ de aynı özelliğe sahip bir λ -matristir.

2.15 Temel İşlemler

Şimdi, aşağıdaki tipte, tekil-olmayan bir λ -matris $P(\lambda)$ -yı inceleyelim.

- | | | | | | |
|-------|-----------------------|----------------------|--|-----------------------|------|
| (i) | $p_{ii}(\lambda) = 1$ | $(i \neq p, q);$ | $p_{pq}(\lambda) = p_{qp}(\lambda) = 1;$ | $p_{ij}(\lambda) = 0$ | yada |
| (ii) | $p_{ii}(\lambda) = 1$ | $(i = 1, \dots, n);$ | $p_{pq}(\lambda) = f(\lambda);$ | $p_{ij}(\lambda) = 0$ | yada |
| (iii) | $p_{ii}(\lambda) = 1$ | $(i \neq p);$ | $p_{pp}(\lambda) = k \neq 0;$ | $p_{ij}(\lambda) = 0$ | yada |

(i) tipinde olan bir matrisin determinantının -1, (ii) tipinde olan bir matrisin determinantının +1, (iii) tipinde olan bir matrisin determinantının ise k olduğu görülür.

$A(\lambda)$ nın (i), (ii), (iii) tipli bir matrisler ile sırasıyla soldan çarpımı aşağıdaki etkilere sahiptir:

- (i) p ve q satırları yer değiştirir,
- (ii) $f(\lambda)$ kere q satırı p satırına eklenir,
- (iii) p satırı *sıfır-olmayan* bir k sabiti ile çarpılır.

Sağdan çarpım kolonlar üzerinde aynı etkiye sahiptir. $A(\lambda)$ nın bu tür dönüşümleri, *temel-işlemler (elementary-operations)* olarak adlandırılır. (Aktaş Z., Öncül H. Ural, 1984)

2.16 Smith's Kanonik Form

Denk λ -matrislerin temel teoremi aşağıdaki gibidir:

Rankı r olan her $A(\lambda)$ λ -matrisi, *temel-dönüşümler* ile

$$D = \left| \begin{array}{ccccccc} E_1(\lambda) & & & & & & \\ & E_2(\lambda) & & & & & \\ & & \ddots & & & & \\ & & & \ddots & & & \\ & & & & E_r(\lambda) & & \\ & & & & & 0 & \\ & & & & & & \ddots \\ & & & & & & & 0 \end{array} \right| \quad (2.90)$$

gibi verilen *denk-matris* D köşegen forma dönüştürülebilir. Burada her bir $E_i(\lambda)$, λ -ya göre bir *monic polinom* dur. Yani $E_i(\lambda)$, $E_{i+1}(\lambda)$ -yı böler. (*Monic polinom*larda, en büyük dereceli terimin baş katsayısı 1 dir.) D -nin diğer tüm değerleri sıfırdır.

İspatı aşağıdadır. Temel işlemler serisi ile köşegen bir form oluşturalım. Herhangi bir adımda, ilgili matrisin (i,j) elemanını $a_{ij}(\lambda)$ olarak tanımlayalım.

$a_{ij}(\lambda)$, $A(\lambda)$ da en düşük dereceli elemanlarından bir tanesi olsun. 1. ve i. satırlar ile, 1. ve j. kolonların yer değişimi ile, bu eleman, $a_{11}(\lambda)$ yapılır. Böylece

$$\begin{aligned} a_{i1}(\lambda) &= a_{11}(\lambda)q_{i1}(\lambda) + r_{i1}(\lambda) \quad (i=2,3,\dots,m) \\ a_{1j}(\lambda) &= a_{11}(\lambda)q_{1j}(\lambda) + r_{1j}(\lambda) \quad (j=2,3,\dots,n) \end{aligned} \quad (2.91)$$

yazılabilir. Burada $r_{i1}(\lambda)$ ve $r_{1j}(\lambda)$ kalanlardır ve dereceleri $a_{11}(\lambda)$ nın derecesinden daha küçüktür. Eğer $r_{i1}(\lambda)$ ve $r_{1j}(\lambda)$ sıfır değil ise, $r_{k1}(\lambda)$ nın sıfır olmadığını farz edelim. 1. satırı $q_{k1}(\lambda)$ kere k. satırdan çıkarıp, 1. ile k. satırların yerlerini değiştirelim. $a_{11}(\lambda)$ bu durumda *sıfır-olmayan* $r_{k1}(\lambda)$ -ya dönüşür. Ve bu pozisyonundan bir önceki elemandan daha düşük dereceli olur.

Bu yöntemin devamı ile, $a_{11}(\lambda)$ elemanı hiçbir zaman sıfır olmaz ve derecesi düzgün bir biçimde azalır. Nihai olarak geçerli $a_{11}(\lambda)$, $(a_{11}(\lambda), \lambda)$ dan bağımsız hale dönüşünceye kadar yada daha erken bir adıma kadar, tüm geçerli $a_{i1}(\lambda)$ ve $a_{1j}(\lambda)$ -ları bölmelidir.

Bu gerçekleştiği takdirde matris, birinci kolonun uygun çarpanı ile 2 den m. satıra kadar ve birinci kolonun uygun çarpanı ile 2. kolondan n. kolona kadar farkının alınmasıyla

$$\left| \begin{array}{cc} a_{11}(\lambda) & 0 \\ 0 & A_2(\lambda) \end{array} \right| \quad (2.92)$$

şekle indirgenir. Bu adımda $a_{11}(\lambda)$ $A_2(\lambda)$ nın tüm elemanlarını böler. Aksi bir durum olarak herhangi bir $a_{ij}(\lambda)$ elemanı için

$$a_{ij}(\lambda) = a_{11}(\lambda)q_{ij}(\lambda) + r_{ij}(\lambda) \quad (r_{ij}(\lambda) \neq 0) \quad (2.93)$$

kabul edelim. Daha sonra, i. satırı ilk kolona ekleyerek, bir önceki işlemi tekrar edelim. Böylece düşük dereceli, *sıfır-olmayan* bir $a_{11}(\lambda)$ -ya sahip (2.92) formu elde edilir. Bu işleme devam ederek, $A_2(\lambda)$ nın her bir elemanını bölen $a_{11}(\lambda)$ -ya sahip (2.92) formuna ulaşırsınız. Bu, ya $a_{11}(\lambda)$, λ -dan bağımsız hale gelince yada daha erken bir adımda meydana gelir.

Eğer $A_2(\lambda)$ sıfır değilse bu matrise, $A(\lambda)$ -ya biraz önce uygulanan benzer işlem uygulanır. Bu, sadece (2.92) matrisinin 2 den m ye kadar olan satırlar ve 2 den n e kadar olan kolonlar üzerinde gerçekleşir. (birinci satır ve n . kolon bu işlemlerden etkilenmezler.) Böylece

$$\left| \begin{array}{cc|c} a_{11}(\lambda) & 0 & \\ 0 & a_{22}(\lambda) & 0 \\ \hline & 0 & A_3(\lambda) \end{array} \right| \quad (2.94)$$

elde edilir. Burada $a_{11}(\lambda)$, $a_{22}(\lambda)$ -yı ve $a_{22}(\lambda)$ $A_3(\lambda)$ nın tüm elemanlarını böler.

Bu yol üzerinde devam ederek aşağıdaki forma ulaşılır:

$$\left| \begin{array}{ccc|c} a_{11}(\lambda) & & & \\ & a_{22}(\lambda) & & \\ & & \ddots & \\ & & & 0 \\ \hline & & & a_{ss}(\lambda) \\ & 0 & & 0 \end{array} \right| \quad (2.95)$$

elde edilir. Burada işlem, $s=m$ veya $s=n$ veya $A_{s+1}(\lambda)$ sıfır-matris olana kadar devam ettirir. Sonuçta her bir $a_{ii}(\lambda)$ ardışığını böler. Bununla birlikte, bir λ -matrisinin rankı, tekil-olmayan matrislerle çarpımı altında invariant olacağından, $s=r$ elde edilir. $a_{ii}(\lambda)$ sıfır olmadığından, uygun sabitlerle satırların çarpımıyla, polinomların her birinin en yüksek dereceden kuvvetlerinin başkatsayıları 1 yapılabilir. $a_{ii}(\lambda)$ bu durumda monic polinom, $E_i(\lambda)$ -ya dönüşür ve her bir $E_i(\lambda)$, $E_{i+1}(\lambda)$ -yı böler. Bu teoremi ispatlar. Burada sadece sıradan işlemlerin kullanıldığı açıktır. *D temel-dönüşüm* ile elde edildiğinden, determinantları λ dan bağımsız $P(\lambda)$ ve $Q(\lambda)$ tekil-olmayan matrisler için

$$P(\lambda)A(\lambda)Q(\lambda) = D \quad (2.96)$$

elde edilir. (2.91) in D formu, denk λ -matrisler için *Smith's-kanonik form* olarak bilinir. (Bronson , 1984).

2.17 Bir Lamda Matrisinin k satırlı Minörlerinin OBEB 'i

Smith's-kanonik formun k satırlı minörlerini göz önüne alalım. Eğer $k > r$ ise k satırlı tüm minörleri 0 dır. Diğer taraftan sadece *sıfır-olmayan* k satırlı minörleri, $E_i(\lambda)$ nın k . ile çarpımlarıdır. Her bir $E_i(\lambda)$, $E_{i+1}(\lambda)$ nın çarpımları olduğundan, k satırlı minörlerinin OBEB'i $G_k(\lambda)$, $E_1(\lambda)E_2(\lambda)\dots E_k(\lambda)$ olur. Böylece $G_0(\lambda)=1$ yazılırsa

$$E_i(\lambda)=G_i(\lambda)/G_{i-1}(\lambda) \quad (i=1,2,\dots,r) \quad (2.97)$$

elde edilir. Şimdi bir matrisin (*birim-matris* olması için en büyük kuvvete sahip λ nın başkatsayısını alacağız) k satırlı minörlerinin OBEB'nin, *eşdeğer-dönüşüm* altında invariant olduğunu göstereceğiz.

Eğer

$$P(\lambda)A(\lambda)Q(\lambda) = B(\lambda) \quad (2.98)$$

sağlıyorsa, $B(\lambda)$ nın herhangi k satırlı minörü, ($P_\alpha, A_\beta, Q_\gamma$; $P(\lambda), A(\lambda), Q(\lambda)$ nın k satırlı minörleri olmak üzere) $P_\alpha, A_\beta, Q_\gamma$ formunun terimlerinin toplamı şeklinde ifade edilebilir. Böylece $A(\lambda)$ nın k satırlı minörlerinin OBEB -i $B(\lambda)$ nın k satırlı minörünü böler.

$$A(\lambda)=[P(\lambda)]^{-1}B(\lambda)[Q(\lambda)]^{-1}$$

ve $[P(\lambda)]^{-1}$ ile $[Q(\lambda)]^{-1}$ λ -matrisleri olduğundan, $B(\lambda)$ nın k satırlı minörlerinin OBEB'i $A(\lambda)$ nın k satırlı minörünü böler. Her iki OBEB eşit olmalıdır. Tüm eşit λ -matrisleri böylece aynı *Smith's-kanonik forma* sahiptirler. $E_i(\lambda)$ polinomları $A(\lambda)$ nın invariant polinomları yada invariant çarpanları olarak adlandırılırlar. (Bronson , 1984).

2.18 $(A-\lambda I)$ -nın Invariant Çarpanları

Eğer A bir kare matris ise, $(A-\lambda I)$, rankı n olan bir λ -matristir. Böylece, ardışıkları tarafından her biri bölünen $E_1(\lambda), E_2(\lambda)\dots E_n(\lambda)$ gibi, n *invariant-çarpana* sahiptir. Şimdi $(A-\lambda I)$ nın *invariant-çarpanlarını*, *Frobenius-kanonik formundaki* Frobenius matrislerinin determinantları ile ilişkilendireceğiz. B yi A nın *Frobenius-kanonik formu* olduğunu kabul edelim. Bu durumda

$$H^{-1}AH=B \quad (2.99)$$

$$H^{-1}(A-\lambda I)H = B-\lambda I \quad (2.100)$$

gibi bir H matris vardır. H^{-1} ve H matrisleri *tekil-olmayan* ve λ dan bağımsız olduklarından, $(B-\lambda I)$, $(A-\lambda I)$ -ya denktir.

Şimdi $(B-\lambda I)$ nın invariant faktörlerini göz önüne alalım. Bunların genel formu, basit bir örnek ile daha etkili olarak gösterilir. $(B-\lambda I)$ nın

$$\begin{vmatrix} p_3-\lambda & p_2 & p_1 & p_0 \\ 1 & -\lambda & 0 & 0 \\ 0 & 1 & -\lambda & 0 \\ 0 & 0 & 1 & -\lambda \end{vmatrix} = \begin{vmatrix} B_1-\lambda I & & \\ & B_2-\lambda I & \\ & & B_3-\lambda I \end{vmatrix} \quad (2.101)$$

$$\begin{vmatrix} q_2-\lambda & q_1 & q_0 \\ 1 & -\lambda & 0 \\ 0 & 1 & -\lambda \end{vmatrix} = \begin{vmatrix} r_1-\lambda & r_0 \\ 1 & -\lambda \end{vmatrix}$$

formda olduğunu kabul edelim. §14 den, $\det(B_1-\lambda I)$, $\det(B_2-\lambda I)$, $\det(B_3-\lambda I)$ dizisinde bu polinomların her biri ardışıkları tarafından bölünür. Sıfır olmayan minörlerin incelenmesiyle i satırlı minörlerinin OBEB'i olan $G_i(\lambda)$

$$\begin{aligned} G_0(\lambda) &= 1 & (tanımdan); \\ G_1(\lambda) &= G_2(\lambda) = \dots = G_6(\lambda) = 1 \\ G_7(\lambda) &= \det(B_3-\lambda I) \\ G_8(\lambda) &= \det(B_3-\lambda I)\det(B_2-\lambda I) \\ G_9(\lambda) &= \det(B_3-\lambda I)\det(B_2-\lambda I)\det(B_1-\lambda I) \end{aligned} \quad (2.102)$$

gibi verilir. Böylece invariant-çarpanlar

$$\begin{aligned} E_1(\lambda) &= E_2(\lambda) = \dots = E_6(\lambda) = 1 \\ E_7(\lambda) &= \det(B_3-\lambda I) \\ E_8(\lambda) &= \det(B_2-\lambda I); \\ E_9(\lambda) &= \det(B_1-\lambda I). \end{aligned} \quad (2.103)$$

olur.

1'e eşit olmayan $(A-\lambda I)$ nın *invariant-çarpanları*, *Frobenius-kanonik formdaki alt-matrislerin* karakteristik polinomlarına eşit olduğundan dolayı, yukarıda elde ettiğimiz sonuç, genel bir sonuçtur. Benzerlikle λ denkleğinin arasındaki temel ilişki aşağıdaki gibidir.

A nın B ye benzer olması için gerek ve yeter koşul $(A-\lambda I)$ ve $(B-\lambda I)$ nın denk olmasıdır. Bu nedenle $(A-\lambda I)$ -nin *invariant-çarpanları* bazen A nın *benzerlik-invariantları* olarak adlandırılır.

Yine §14 ten, eğer $\lambda_1, \lambda_2, \dots, \lambda_s$ A nın ayrık özdeğerleri olmak üzere

$$E_i(\lambda) = (\lambda_1 - \lambda)^{i_1} (\lambda_2 - \lambda)^{i_2} \dots (\lambda_s - \lambda)^{i_s}, \quad (2.104)$$

yazarsak sağ taraftaki çarpanlar A nın *temel-bölenleri* olur. (Seymor, 1978)

2.19 Üçgen Kanonik Form

Bir matrisin *Jordan-kanonik formuna* indirgenmesi ile bu matrisin özdeğerler ve özvektörler hakkındaki tüm bilgi elde edilebilir. Bu form *üçgen-kanonik formunun* özel bir durumudur. Bu son hal teorik yönden ve pratik hesaplamalar açısından çok daha basittir. Bunun önemi aşağıdaki teoremin sonucudur.

Her bir matris bir *üçgen-matrise* benzerdir. Yani köşegeni altındaki her bir elemanı sıfır olan bir matrise. Bu tür bir matrise *alt-üçgen matris* adı verilir. Jordan formu gibi alternatif bir dönüşümde köşegenin altındaki tüm elemanları sıfır yapar (*üstüçgen-matris*).

Elde edilen bu matrisin köşegen elemanlarının uygun çarpanları matrisin özdeğerleridir. Eğer bu elemanlar μ_i ise, *üçgen-matrisin karakteristik-polinomu* $(\mu_i - \lambda)$ olur ve *karakteristik-polinomun benzerlik-dönüşümü* altında invariant olduğu daha önce gösterilmişti.

2.20 Hermitian ve Simetrik Matrisler

Reel ve simetrik matrisler, pratik hesaplamalarda ve §31 de de göreceğimiz gibi, özel bir önemi vardır. Bunu simetrik ve kompleks elemanlara sahip tüm matrisler için genişletebiliriz. Ancak bu genişletme küçük bir anlam taşır. Çünkü bu genişletilmiş matrisler, reel simetrik matrislerin çok önemli özelliklerinin çoğuna sahip değildirler. En kullanışlı genişletme, $(P+iQ)$ formulu sınıflara olan genişletmedir. Burada P reel ve simetrik, Q ise reel ve *ters-simetrik* tir. *Ters-simetrik (Skew-simetrik)* bir matrisin tanımı aşağıdaki gibidir:

$$Q^T = -Q \quad (2.105)$$

Bu sınıftan olan matrisler *Hermitian-matris* olarak adlandırılırlar. Eğer A bir *Hermitian-matris* ise tanımdan

$$\bar{A}^T = A \quad (2.106)$$

elde ederiz. Burada $\bar{A}$, A -nın kompleks eşlenik elemanlarını tanımlamaktadır. $\bar{A}^T$ matrisi sıklıkla A^H ile ifade edilir ve *Hermitian transpose matris* olarak adlandırılır. Benzer olarak x^H -de, x kolon vektörünün kompleks eşlenik elemanlarından oluşan satır vektörü ifade eder. Bu notasyon uygundur ve bir avantaja sahiptir – şöyle ki - Hermitian matrisler için ispat edilen sonuçlar, reel ve simetrik matrisler için A^H , A^T ile ve x^H , x^T ile değiştirilerek uygun sonuçlara çevrilebilir. Eğer a bir skaler ise a^H , a nın kompleks eşlenik olarak tanımlanabilir.

$$\begin{aligned} (A^H)^H &= A \\ (x^H)^H &= x \\ (ABC)^H &= C^H B^H A^H \end{aligned} \quad (2.107)$$

Eğer A Hermitian ise

$$A^H = A \quad (2.108)$$

olur.

$y^H x$, y ve x in iç-çarpımıdır ve

$$\overline{y^H x} = (y^H x)^H = x^H y. \quad (2.109)$$

i

ç-çarpım $x^H x$, tüm x -ler için *sıfır-vektörü hariç* reel ve pozitif bir sayıdır. Çünkü x in elemanlarının modülünün karelerinin toplamıdır. (Wilkinson J.H, 1965)

2.21 Hermitian Matrislerin Temel Özellikleri

Bir Hermitian matris in özdeğerlerinin hepsi reel dir. Eğer,

$$Ax = \lambda x \quad (2.110)$$

ise bu durumda

$$x^H A x = \lambda x^H x. \quad (2.111)$$

olur.

Şimdi *sıfır-olmayan* bir x için $x^H x$ in reel ve pozitif olduğunu hatırlayalım. Ayrıca, $x^H A x$ reel bir skaler olduğundan

$$(x^H A x)^H = x^H A^H x^{HH} = x^H A x \quad (2.112)$$

yazılabilir. Böylece λ reeldir. Reel simetrik matrisler için reel özdeğerler, reel özvektörler manasına gelir. Fakat bir kompleks Hermitian matrisin özvektörleri genelde kompleksir.

Eğer $Ax = \lambda x$ ise bir Hermitian matris için

$$A^H x = \lambda x \quad (2.113)$$

yazılabilir ve her iki tarafında kompleks eşleniğini alırsak:

$$A^T \bar{x} = \lambda \bar{x} \quad (2.114)$$

elde edilir.

Bir Hermitian matrisin transpozisinin özdeğerleri kendi özdeğerlerinin kompleks eşleniğidir. §3 ün notasyonundan

$$y_i = \bar{x}_i \quad (2.115)$$

$y_i^T x_i$ sayısı Hermitian matrisler için $x_i^H x_i$ ye dönüşür. Eğer bir Hermitian matris ayrık özdeğerlere sahip ise özvektörleri

$$x_i^H x_j = 0 \quad (i \neq j). \quad (2.116)$$

sağlar:

Eğer x_i aşağıdaki gibi normalize edilirse:

$$x_i^H x_i = 1 \quad (2.117)$$

(2.116) ve (2.117) denklemleri, özvektörler tarafından yapılandırılan X matrisinin

$$\begin{aligned} X^H X &= I \\ X^H &= X^{-1} \end{aligned} \quad (2.118)$$

sağladığını gösterir. (Wilkinson J.H., 1965)

Denklem (2.118) u sağlayan bir matris *uniter-matris* (*unitary-matrix*) olarak adlandırılır. Reel bir *uniter-matris* *ortogonal-matris* olarak adlandırılır.

Hermitian ve reel simetrik matrisler için aşağıdaki ifadeleri ispatlandı.

Eğer A Hermitian ise aşağıdaki gibi U *uniter-matrisi* mevcuttur.

$$(i) \quad U^H A U = \text{diag}(\lambda_i) \quad (\text{reel } \lambda_i)$$

Eğer A reel ve simetrik ise, aşağıdaki gibi bir *ortogonal-matris* vardır:

$$(ii) \quad U^T A U = \text{diag}(\lambda_i) \quad (\text{reel } \lambda_i)$$

(iii) Hermitian matris eksik (defective) olamaz.

(i) ve (ii) Hermitian matrislerin en önemli özellikleridir.

Buradan Hermitian bir matrisin temel-bölenlerin tümünün lineer olduğu sonucuna ulaşılır.

Eğer bir Hermitian matris katlı özdeğerlere sahip ise bu matris *derogatory* dir.

2.22 Kompleks Simetrik Matrisler

Reel simetrik matrislerin önemli özelliklerinden hiçbirinin kompleks simetrik matrislerde de geçerli olmadığı

$$A = \begin{vmatrix} 2i & 1 \\ 1 & 0 \end{vmatrix} \quad (2.119)$$

gibi tanımlanan A matrisi göz önünde bulundurarak görülür.

Karakteristik-denkle $\lambda^2 - 2i\lambda - 1$ dir. Öyle ki $\lambda=i$ 2 katlı özdeğedir. Karşılık gelen özvektörlerin x_1, x_2 bileşenleri

$$ix_1 + x_2 = 0 \quad x_1 - ix_2 = 0 \quad (2.120)$$

sağlar. Öyle ki sadece bir özvektör vardır ve $(1, -i)$ bileşenlerine sahiptir.

Burada $\lambda=i$ ile ilişkili bir kuadratik bölenin var olduğunu gösterelim. Gerçekte *Jordan-kanonik form* aşağıdaki gibidir:

$$\begin{vmatrix} i & 1 \\ 0 & i \end{vmatrix} \quad (2.121)$$

ve

$$\begin{vmatrix} 2i & 1 \\ 1 & 0 \end{vmatrix} \begin{vmatrix} 1 & 0 \\ -i & 1 \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ -i & 1 \end{vmatrix} \begin{vmatrix} i & 1 \\ 0 & i \end{vmatrix} \quad (2.122)$$

sahip oluruz.

2.23 Üniter Dönüşümler ile Üçgen Forma Dönüşümler

Bir Hermitian matris üniter benzerlik-dönüşümü ile *köşegen-matrise* indirgenebilir. Herhangi bir matris *üniter benzerlik-dönüşümü* ile *üçgen-forma* indirgenebilir. *Üçgen-form* derhal özdeğerleri verdiğiinden ve özvektörler aynı basitlik ile bulunabildiğinden, bu sonuç pratik bir öneme sahiptir. Çünkü *üniter-dönüşüm* istenilen nümerik özelliklere sahiptir.

2.24 Kuadratik Formlar

Reel simetrik A matrisi ile

$$x^T A x = \sum_{i,j=1}^n a_{ij} x_i x_j, \quad (2.123)$$

gibi tanımlanan n tane $x_1, x_2, \dots, x_n$ değişkenlerinin $x^T A x$ kuadratik formu arasında ilişki kurulabilir.

Burada x , x_i bileşenlerinden oluşan bir vektördür. Her bir kuadratik forma karşılık

$$x^T A y = \sum_{i,j=1}^n a_{ij} x_i y_j = y^T A x \quad (2.124)$$

gibi tanımlanan *bilineer-form* $x^T A y$ yi tanımlayabiliriz. Bazı tip kuadratik formlar özellikle önemlidir.

A matrisi ile bir kuadratik form aşağıdaki gibi tanımlara sahiptir:

$$\left. \begin{array}{l} \text{pozitif-tanımlı} \\ \text{negatif tanımlı} \\ \text{negatif-olmayan tanımlı} \\ \text{pozitif-olmayan tanımlı} \end{array} \right\} x^T A x \left\{ \begin{array}{l} > 0 \\ < 0 \\ \geq 0 \\ 0 \end{array} \right. \quad x \neq 0 \quad (2.125)$$

Bir pozitif-tanımlı kuadratik forma sahip bir matris, pozitif-tanımlı matris (positive-definite matrix) olarak adlandırılır. (Gould S.H, 1974)

$x^T A x$ in *pozitif-tanımlı* olması için, gerek ve yeter şart, A nın tüm özdeğerlerinin pozitif olmasıdır.

$$R^T A R = \text{diag}(\lambda_i) \quad (2.126)$$

şeklinde R ortogonal matrisi var olduğundan

$$x = Rz; \quad z = R^T x \quad (2.127)$$

$$x^T A x = z^T R^T A R z = \sum_{i=1}^n \lambda_i z_i^2 \quad (2.128)$$

yazılabilir.

$\sum \lambda_i z_i^2$ -nin pozitif olması için yeter ve gerek şart tüm *sıfır-olmayan* z için tüm λ_i -lerin pozitif olmasıdır. (2.127) den *sıfır-olmayan* z ye karşılık gelen *sıfır-olmayan* x -leri görebiliriz. Böylece sonuç elde edilir. Bir matrisin determinantı, onun özvektörlerinin çarpımı olduğundan, *pozitif-tanımlı matrisin* determinantı pozitif olmalıdır.

(2.128) ve (2.129) denklemlerinden, ortogonal R matrisinin – ve dolayısıyla A nın özvektörlerinin- bulunması probleminin

$$\sum_{i,j} a_{ij} x_i x_j = 1 \quad (2.129)$$

kuadratik ifadenin temel eksenlerinin belirlenmesi ile aynı olduğu ve λ_i -nin temel eksenlerin kare kökünün tersi olduğu sonucuna varırız. Bir katlı özdeğer asal eksenlerinin belirsizliğine sebep olur. Tipik bir örnek 3 boyutlu uzayda bir elipsoidin iki eşit asal eksene sahip olmasıdır.

2.25 Pozitif Tanımlılık için Gerek ve Yeter Şartlar

$x^T A_r x$ ile $x^T A x$ ifadesinde $x_{r+1}, x_{r+2}, \dots, x_n$ -ler sıfır alınarak elde edilen kısmı göstereceğiz. $x^T A_r x$ böylece r adet $x_1, x_2, x_3, \dots, x_r$ değişkenin bir kuadratik formdur. Bu formun A_r matrisi, A nın r . Mertebeden ön temel altmatrisidir. Eğer $x^T A x$ *pozitif-tanımlı* ise, $x^T A_r x$ de *pozitif-tanımlı*dır. Eğer sıfır olmayan $x_1, x_2, \dots, x_r$ değerleri için $x^T A_r x$ pozitif değilse

$$x^T = (x_1, x_2, \dots, x_r, 0, \dots, 0) \quad (2.130)$$

de pozitif değildir. Böylece *ön asli-minörlerin* (leading prensible minör) tümü pozitif olan bir matris pozitif olur.

Benzer bir metot ile *pozitif-tanımlı* bir matrisin herhangi ön temel altmatrisinin *pozitif-tanımlı* ve böylece pozitif bir determinanta sahip olduğu gösterilir. Bununla birlikte, n pozitif *asli-minörlerin* varlığı da *pozitif-tanımlılık* için yeterli bir koşuldur. İspat tümevarım ile yapılacaktır.

($n-1$) mertebeden bir matris için bunun doğru olduğunu kabul edelim. Bu durumda n . mertebeden bir matris için de doğru olduğu ispatlanacak. Şimdi aksi olarak, A , n *asli-minörleri* pozitif olan n . mertebeden bir matris olsun ve A *pozitif-tanımlı* olmasın. Eğer $R, R^T A R = \text{diag}(\lambda_i)$ için ortogonal bir matris ise $x = Rz$ yazarak

$$x^T A x = z^T R^T A R z = \sum \lambda_i z_i^2 \quad (2.131)$$

elde edilir. Hipotezden A *pozitif-tanımlı* değil ve böylece $(\lambda_i) = \det(A) > 0$ olduğundan, negatif λ_i kadar r çift sayı olmalıdır. İlk r ve r ye kadar olanların negatif olduğunu kabul ederek

$$z_{r+1} = z_{r+2} = \dots = z_n = x_n = 0 \quad (2.132)$$

denklem kümesini göz önünde bulunduralım. z_i, x_i -nin lineer fonksiyonları olduğundan, n bilinmeyenli $x_1, x_2, \dots, x_n$ de $(n-r+1)$ lineer homojen denkleme sahip oluruz. Böylece en az bir tane *sıfır-olmayan* x çözümü vardır ki bunu $(x_1, x_2, \dots, x_{n-1}, 0)$ şeklinde ifade edebiliriz. Karşılığı olan z de sıfır değildir ve bunu $(z_1, z_2, \dots, z_r, 0, 0, \dots, 0)$ şeklinde ifade edebiliriz. Bu x ve z için

$$x^T A x = \sum_1^n \lambda_i z_i^2 = \sum_1^r \lambda_i z_i^2 < 0 \quad (2.133)$$

yazılabilir. Bu, $x_n=0$ olduğundan, $x^T A_{n-1} x$ in $(x_1, x_2, \dots, x_{n-1})$ vektörü için negatif olduğunu gösterir. Bu hipotezimiz ile çelişkidir. Sonuç açık olarak, 1. mertebeden matris için doğrudur ve böylece herhangi mertebeden matris için de geçerli olur.

Aynı şekilde bir kuadratik form ile bir reel simetrik matrisin ilişkilendirilmesinde yaptığımız gibi, Hermitian bir matris A ile

$$x^H A x = \sum_{i,j=1}^n a_{ij} x_i x_j \quad (2.134)$$

$$(x^H A x)^H = x^H A^H x^{HH} = x^H A x \quad (2.135)$$

Hermitian formunu ilişkilendirebiliriz. Böylece $x^H A x$ reeldir.

2.26 Bir Vektörün Minimum Polinomu

Bir matrisin özvektörleri, x ve Ax arasında lineer bir ilişkinin var olmasından dolayı karakterize edilebilirler. Keyfi bir b vektörü için hiçbir ilişki mevcut değildir. Fakat eğer, $b, Ab, A^2b, \dots, A^r b$, dizisi yapılandırıldığında, bu vektörleri bağımlı yapacak bir minimum m değeri vardır ve $m < n$ dir. Bu ilişki

$$(A^m + c_{m-1} A^{m-1} + \dots + c_0 I) b = 0 \quad (2.136)$$

$$(A^r + d_{r-1} A^{r-1} + \dots + d_0 I) b = 0 \quad (r < m) \quad (2.137)$$

gibi yazılabilir. (2.137) in sol tarafına karşılık gelen $c(\lambda)$ *monic-polynomu*, A -ya göre b -nin *minimum-polinom* olarak adlandırılır. Eğer $s(A)$, $s(A)b=0$ için A -daki diğer herhangi bir polinom ise, bu durumda $c(A)$, $s(A)$ -yı bölmelidir. Aksi takdirde, Euclidian algoritmasından

$$p(A)s(A)-q(A)c(A) = r(A) \quad (2.138)$$

gibi $p(A)$, $q(A)$, $r(A)$ polinomları mevcuttur. Burada $r(A)$, $c(A)$ dan daha düşük derecedendir. Fakat (2.138) den $r(A)b=0$ olur ve bu $c(A)$ nın, en düşük dereceden polinom olması hipotezine ters düşer.

Minimum polinom *tek* dir. Eğer bir m . dereceden ikinci *monic-polinom* için $d(A)b=0$ ise, bu durumda $[c(A)-d(A)]b=0$ ve $c(A)-d(A)$ nın derecesi m den küçüktür. $c(A)$ nın derecesi A -ya bağlı b -nin *derecesi* olarak adlandırılır. (Wilkinson J.H, 1965)

2.27 Bir Matrisin Minimum Polinomu

Benzer bir yol olarak, $I, A, A^2, \dots, A^r$ matris dizisini göz önüne alalım. *Lineer-bağımlı* olan bu matris kümesi için, r -nin en küçük bir s değeri vardır. Bu ilişki aşağıdaki şekilde yazılabilir:

$$A^s + m_{s-1}A^{s-1} + \dots + m_0I = 0 \quad (2.139)$$

$$A = \begin{vmatrix} p_1 & p_0 \\ 1 & 0 \end{vmatrix} \quad (2.140)$$

$$A^2 = \begin{vmatrix} p_1^2 + p_0 & p_1 p_0 \\ p_1 & p_0 \end{vmatrix} \quad (2.141)$$

$$\begin{aligned} A^2 - p_1 A - p_0 I &= \begin{vmatrix} p_1^2 + p_0 & p_1 p_0 \\ p_1 & p_0 \end{vmatrix} - \begin{vmatrix} p_1^2 & p_1 p_0 \\ p_1 & p_0 \end{vmatrix} - \begin{vmatrix} p_0 & 0 \\ 0 & p_0 \end{vmatrix} \\ &= \begin{vmatrix} 0 & 0 \\ 0 & 0 \end{vmatrix} \end{aligned} \quad (2.142)$$

(2.143) den dolayı daha düşük mertebelerin ilişkisi yoktur.

$$A + c_0 I = \begin{pmatrix} p_1 + c_0 & p_0 \\ 1 & c_0 \end{pmatrix} \quad (2.143)$$

ve herhangi bir c_0 için bu *sıfır-matris* değildir.

(2.139) denkleminin sol tarafına karşılık gelen $m(\lambda)$ polinomuna A matrisinin *minimal-polinomu* adı verilir. §28' te olduğu gibi, bunun *tek* olduğu ve $p(A)=0$ için diğer $p(\lambda)$ polinomlarını böldüğü gösterebilir.

$m(A)b=0$ olduğundan, A -ya göre b vektörünün minimum polinomu $c(\lambda)$, $m(\lambda)$ -yı böler. $c_1(\lambda), c_2(\lambda), \dots, c_n(\lambda)$ A -ya göre herhangi $e_1, e_2, \dots, e_n$ -nin minimum polinomları, $g(\lambda)$ de $c_i(\lambda)$ nın *OKEK* i olsun. Bu durumda

$$g(A)e_i = 0 \quad (i = 1, \dots, n) \quad (2.144)$$

elde edilir. Böylece $g(A)=0$ olur. Bu herhangi bir vektör için $g(A)b=0$ olduğunu gösterir. $g(\lambda)$, tüm vektörlerin minimum polinomlarının *OKEK*'idir.

$g(\lambda)=m(\lambda)$ olduğundan ve tanımdan $m(A)=0$ dır. Böylece $m(A)e_i=0$ olur ve e_i -nin minimum polinomlarının *OKEK*'i. $m(\lambda)$ -ya bölünür ve $g(\lambda)$, $m(\lambda)$ -yı böler. Fakat tanımdan, $m(\lambda)$, $m(A)=0$ nın en düşük dereceden polinomudur.

Bu teorem $b, Ab, A^2b, \dots$ kolon vektörleri sisteminden geliştirilecektir. Benzer bir yol ile, $b^T, b^T A, b^T A^2, \dots$ satır vektör sistemini göz önünde bulundurarak da, A -ya göre b^T -nin minimum polinomunu belirlenebilir. Burada A -ya göre b^T -nin minimum polinomu, A^T -ya göre b polinomununki ile aynıdır.

Eğer $e_1^T, e_2^T, \dots, e_n^T$ -nin minimum polinomlarının *OKEK* i $h(\lambda)$ ile ifade edilirse, bu durumda $h(A)$ nın *sıfır-matris* olduğunu ve böylece $h(\lambda)$ nın $m(\lambda)$ ile eşit olduğu görülür. İster satır ister kolon vektörleri göz önünde bulundursun, A nın aynı minimum polinomuna ulaşılır.

2.28 Cayley-Hamilton Teoremi

A'nın her bir kuvvetinde n^2 eleman olduğundan, A'nın minimum polinomunun n^2 den daha büyük dereceden olamayacağı açıktır. Eğer, A'nın *karakteristik-polinomu* aşağıdaki gibi verilirse

$$\lambda^n + c_{n-1}\lambda^{n-1} + \dots + c_0 = 0 \quad (2.145)$$

(2.146) denklemi elde edilir:

$$A^n + c_{n-1}A^{n-1} + \dots + c_0I = 0 \quad (2.146)$$

Bunun manası bir matrisin minimum polinomunun n den daha büyük mertebeden olamayacağıdır.

Bunu ispat etmek için, bir A matrisinin B adjointini tanıtacağız. Bu (i,j) elemanı için a_{ji} -nin kofaktörüne sahip olan matris demektir. Böylece

$$AB = \det(A)I \quad (2.147)$$

elde edilir.

Bu son durum her matris için doğrudur. Aynı zamanda özel olarak $(\lambda I - A)$ matrisi için de doğrudur. $(\lambda I - A)$ -nın $B(\lambda)$ adjointi, bir λ matrisidir. Buradaki her bir eleman, derecesi (n-1) veya λ -nın kinden daha küçük dereceden bir polinomdur. Çünkü $(\lambda I - A)$ nın (n-1) mertebeden bir minörüdürler. Böylece

$$B(\lambda) = B_{n-1}\lambda^{n-1} + B_{n-2}\lambda^{n-2} + \dots + B_0 \quad (2.148)$$

yazılabilir. Burada B_i -ler sabit elemanlı (n x n) mertebeden matrislerdir. (2.147) sonucunu $(\lambda I - A)$ -ya uygulayarak, aşağıdaki denkleme ulaşılır:

$$\begin{aligned} (\lambda I - A)(B_{n-1}\lambda^{n-1} + \dots + B_0) &= \det(\lambda I - A)I \\ &= (\lambda^n + c_{n-1}\lambda^{n-1} + \dots + c_0)I \end{aligned} \quad (2.149)$$

Bu sonuç λ –ya göre özdeşliktir ve λ -nın aynı dereceden kuvvetlerini eşitleyerek

$$\begin{aligned}
 B_{n-1} &= I \\
 B_{n-2} - AB_{n-1} &= c_{n-1}I \\
 B_{n-3} - AB_{n-2} &= c_{n-2}I \\
 &\dots\dots\dots \\
 -AB_0 &= c_0I
 \end{aligned} \tag{2.150}$$

elde edilir. Bu denklemlerin $A^n, A^{n-1}, \dots, I$ ile çarpımları aşağıdaki ifadeyi verir:

$$0 = A^n + c_{n-1}A^{n-1} + \dots + c_0I \tag{2.151}$$

Her kare matris böylece kendisinin karakteristik denklemini sağlar. Bu sonuç Cayley-Hamilton teoremi olarak bilinir. Minimum polinomlar karakteristik denklemi böldüğü ve böylece ilkinin n den daha büyük olmayan dereceden olduğu sonucuna ulaşılır.

2.29 Minimum Polinom ve Kanonik Form Arasındaki İlişki

§28 deki görüş bazen *Frobenius-kanonik formun* üretilmesi içinde kullanılır. Bu matrisin minimum polinom ile onun *Frobenius* ve *Jordan-kanonik formları* arasındaki ilişkiyi anlamak için önemli bir noktadır.

İlk olarak bütün benzer matrisler aynı minimum polinoma sahip olduğuna dikkat edelim. Eğer $H^{-1}AH=B$ ise

$$H^{-1}A^2H = H^{-1}AHH^{-1}AH = B^2 \tag{2.152}$$

ve benzer olarak

$$H^{-1}A^rH = B^r \tag{2.153}$$

$$H^{-1}f(A)H = f(B)$$

verir. Burada $f(A)$, A nın herhangi bir polinomudur. Eğer $f(A)=0$ ise, $f(B)=0$ olur. Terside geçerlidir. Böylece A nın minimum polinomunu, B -nin minimum polinomunu böler. Terside geçerlidir. Bu durumda bu polinomlar özdeşler.

Şimdi denklem (2.135) ile verilen r . mertebeden B_r Frobenius matrisini göz önüne alalım. Aşağıdaki ifadeler elde edilir:

$$\begin{aligned} e_{s+1}^T B_r &= e_s^T \quad (s = 1, 2, \dots, r-1) \\ e_1^T B_r &= (b_{r-1}, b_{r-2}, \dots, b_0) \end{aligned} \quad (2.154)$$

$$e_r^T (B_r)^k = e_{r-k}^T \quad (k = 0, 1, \dots, r-1) \quad (2.155)$$

verir.

$k < r-1$ için

$$e_r^T (a_k B_r^k + \dots + a_0 I) = (0, \dots, 0, a_k, a_{k-1}, \dots, a_0) \quad (2.156)$$

yazılır. a_i -ler sıfır olmadığı müddetçe vektörün sağ tarafı sıfır olamaz. Böylece B_r -ye göre e_r^T -nin minimum polinomları r den daha düşük dereceden olamaz.

$$e_r^T (B_r^r) = e_r^T (b_{r-1} B_r^{r-1} + b_{r-2} B_r^{r-2} + \dots + b_0 I) \quad (2.157)$$

yazılır. Bu durumda e_r^T -nin minimum polinomları B_r -nin *karakteristik-polinomu* ve B_r -nin minimum polinomu onun *karakteristik-polinomudur*. §29 $r = 2$ için iyi bir örnektir.

Genel bir *Frobenius-kanonik form* B

$$B = \begin{vmatrix} B_{r_1} & & & \\ & B_{r_2} & & \\ & & \ddots & \\ & & & B_{r_{s-1}} \end{vmatrix} \quad (2.158)$$

gibi yazılabilir. Burada her bir B_r -nin *karakteristik-polinomu* onun ardışıkları tarafından bölünür.

B deki herhangi $f(B)$ polinomu için aşağıdaki ifade elde edilir:

$$f(B) = \begin{vmatrix} f(B_{r_1}) & & & \\ & f(B_{r_2}) & & \\ & & \ddots & \\ & & & f(B_{r_s}) \end{vmatrix} \quad (2.159)$$

Eğer $f(B)$ sıfır-matris ise, $f(B_{r_i})$ -nin her biri sıfır-matris olur. Bu durumda yalnız ve yalnız $f(\lambda)$ B_{r_i} nin karakteristik polinomu olan $f_i(\lambda)$ -nin bir çarpımı ise, $f(B_{r_i})$ sıfır-matristir. Ve her bir $f_i(\lambda)$, $f_1(\lambda)$ -yı böldüğünden, $f_1(\lambda)$, B -nin minimum polinomudur. (Wilkinson J.H, 1965)

İspat, eğer yalnız ve yalnız bir matrisin *Frobenius-kanonik* formunda sadece bir tane Frobenius matris var ise, yani derogatory olmayan matris ise, matrisin polinomunun, onun karakteristik polinomuna eşit olduğunu gösterir. Buda eğer özdeğerler ayrık ise de geçerlidir.

Şimdi *Jordan-kanonik* formuna dönerek

$$C_3(a) = \begin{vmatrix} a & 0 & 0 \\ 1 & a & 0 \\ 0 & 1 & a \end{vmatrix} \quad (2.160)$$

Jordan matrisini ele alalım. Kısa bir gösterim olması için matrisimizi C ile ifade edelim.

$$C^2 = \begin{vmatrix} a^2 & 0 & 0 \\ 2a & a^2 & 0 \\ 1 & 2a & a^2 \end{vmatrix} \quad C^3 = \begin{vmatrix} a^3 & 0 & 0 \\ 3a^2 & a^3 & 0 \\ 3a & 3a^2 & a^3 \end{vmatrix} \quad (2.161)$$

(Burada *alt-Jordan-kanonik formu* olarak kabul edeceğiz). C^2 -nin, C ve I nin bir lineer kombinasyonu olmadığı bir gerçektir. Böylece $C_3(a)$ nın minimum polinomunun derecesi 2 den büyüktür. Buradan minimum polinom, $(a-\lambda I)^3$ karakteristik-polinomudur. Eğer C *Jordan-kanonik formuna* sahip ise C , *Jordan-altmatrislerinin direk-toplamıdır*.

Bu form içerisindeki altmatrisleri, $C_{ij}(\lambda_i)$ ile λ_i -nin belirli değerlerine karşılık gelen herhangi bir altmatris olarak ifade edeceğiz. *Frobenius formda* olduğu gibi, C -nin bir polinomu $f(C)$ -nin, $f[C_{ij}(\lambda_i)]$ matrislerinin *direk-toplamı* olduğu görülür. Böylece, eğer yalnız ve yalnız $f[C_{ij}(\lambda_i)]$ sıfır-matris ise $f(C)$ sıfırdır. Eğer λ_i -ye bağlı en yüksek dereceden *Jordan matrisi* $C_{i1}(\lambda_i)$ ile ifade edilirse, bu durumda, C -nin minimum polinomu $(\lambda-\lambda_i)^{i1}$ olmalıdır. §14 teki

Frobenius form ile Jordan form arasındaki ilişki, Jordan ve benzer Frobenius-kanonik formların hepsinin aynı minimum polinomlara sahip olduklarını onaylar.

2.30 Genelleştirilmiş Vektörler

Lineer temel bölenlere sahip bir A matrisi için tüm n -boyutlu uzayı geren n özvektör vardır. Eğer A, *lineer-olmayan temel-bölenlere* sahip ise, bu doğru değildir. Çünkü, n den daha az bağımsız özvektör vardır. N boyutlu uzayı geren vektörler kümesini oluşturmak uygundur. Bunların seçimi öyle olmalıdır ki lineer bölen halinde yukarıdakini versin. Şimdi özvektörlerin X matrisinin kolonları gibi alınabileceği durumu inceleyelim.

$$X^{-1}AX = \text{diag}(\lambda_i) \quad (2.162)$$

Matrisin *lineer-olmayan bölenlere* sahip olmadığı zaman bir genişletme, A'nın *Jordan-kanonik forma* indirgeyen X matrisinin n kolonunu *baz-vektörler* olarak almaktır.

Bu vektörler önemli bir ilişki sağlar. 8. mertebeden basit bir örnek ile durum daha iyi anlaşılacaktır. A matrisini

$$AX = \begin{vmatrix} C_3(\lambda_1) & & & \\ & C_2(\lambda_1) & & \\ & & C_2(\lambda_2) & \\ & & & C_1(\lambda_3) \end{vmatrix} \quad (2.163)$$

olarak kabul edelim. Bu durumda eğer $x_1, x_2, \dots, x_8$ X'in kolonları ise

$$\begin{aligned} Ax_1 &= \lambda_1 x_1 + x_2, & Ax_4 &= \lambda_1 x_4 + x_5, & Ax_6 &= \lambda_2 x_6 + x_7, & Ax_8 &= \lambda_3 x_8 \\ Ax_2 &= \lambda_1 x_2 + x_3, & Ax_5 &= \lambda_1 x_5, & Ax_7 &= \lambda_2 x_7 \\ Ax_3 &= \lambda_1 x_3 \end{aligned} \quad (2.164)$$

bunları eşitleyerek

$$\begin{aligned} (A - \lambda_1 I)^3 x_1 &= 0, & (A - \lambda_1 I)^2 x_4 &= 0, & (A - \lambda_2 I)^2 x_6 &= 0, & (A - \lambda_3 I) x_8 &= 0 \\ (A - \lambda_1 I)^2 x_2 &= 0, & (A - \lambda_1 I) x_5 &= 0, & (A - \lambda_2 I) x_7 &= 0 \\ (A - \lambda_1 I) x_3 &= 0 \end{aligned} \quad (2.165)$$

elde ederiz. Böylece her bir vektör

$$(A - \lambda_i I)^j x_k = 0 \quad (2.166)$$

eşitliğini sağlar. (2.166) denklemini sağlayan $(A - \lambda_i I)$ -nın daha düşük derecelerinde sağlamayan bir vektör, λ_i ye karşılık gelen j . dereceden genelleştirilmiş vektör (*principal vector of grade j corresponding to λ_i*) olarak adlandırılır. $(\lambda - \lambda_i)^s$, A -ya göre x_k -nin minimum polinomudur. Herhangi bir özvektör, 1. dereceden genelleştirilmiş-vektördür ve X matrisinin kolonları, A nın tüm genelleştirilmiş-vektörlerine karşılık gelir. X in kolonları bağımsız olduğundan, n -boyutlu uzayı geren genelleştirilmiş-vektörler kümesi mevcuttur. Genelde *temel-vektörler* tek değildir. *Jordan-kanonik formundaki* bir matris için, $e_1, e_2, \dots, e_n$ vektörleri *genelleştirilmiş-vektörler* kümesini yapılandırır.

2.31 Temel Benzerlik Dönüşümleri

Genel formdaki bir matrisin kanonik formlardan birine veya herhangi bir yoğunlaştırılmış forma, genellikle, basit benzerlik-dönüşüm zincirleri ile ulaşılmaktadır. §17 de bazı *temel-dönüşümler* incelenmişti. Fakat bunları bir şekilde genelleştirmek ve kitabın kalan kısımlarına bağlı kalarak her bir tipi standart bir notasyon olarak tanıtmak daha uygun olacaktır.

Bundan sonra bahsedilen matrislerden bazıları *temel-dönüşüm* ve *temel-matrisler* üzerine kurulu dönüşümler şeklinde ifade edilecektir.

(i) I_{ij} matrisleri, i . satır ve j . kolonları hariç birim matrise eşittir. i, j kolonları

$$\begin{array}{cc|c} \text{i. kolon} & \text{j. kolon} & \\ \hline 0 & 1 & \text{i. satır} \\ 1 & 0 & \text{j. satır} \end{array} \quad (2.167)$$

gibidir. $I_{ii} = I$, $I_{ij} I_{ij} = I$ ve I_{ij} kendisinin tersidir ve aynı zamanda ortogonaldır. I_{ij} ile soldan çarpım i ile j satırlarının ve sağdan çarpım, i ile j kolonlarının yer değişimine sebep olur. I_{ij} matrisi ile bir *benzerlik-dönüşümü* böylece i ve j satırları ile, i ve j kolonlarının yer değişimine sebep olur.

(ii) $P_{\alpha_1, \alpha_2, \dots, \alpha_n}$ matrisleri

$$p_{i i} = 1, \quad p_{ij} = 0 \quad (2.168)$$

gibi tanımlanır. Burada $(\alpha_1, \alpha_2, \dots, \alpha_n)$ $(1,2,\dots,n)$ -nin permutasyonlarıdır. Bu matrislere *permutasyon-matrisleri* olarak adlandırılırlar. Bir *permutasyon-matrisindeki* her bir kolon ve satırda bir tane *birim-eleman* vardır. Bir permutasyon matrisin transpozesi de bir permutasyon matrisidir.

$$(P_{\alpha_1, \alpha_2, \dots, \alpha_n})^{-1} = (P_{\alpha_1, \alpha_2, \dots, \alpha_n})^T \quad (2.169)$$

Her bir permutasyon matris, I_{ij} tipinde bir matrisin çarpımı gibi ifade edilebilir. P ile soldan çarpım, başlangıç satır α_i -nin i. satıra ve P^T ile sağdan çarpım başlangıç kolon α_i -nin i. kolonuna aktarılmasına sebep olur.

(iii) R_j matrisleri

$$(-r_{i1}, -r_{i2}, \dots, -r_{i, i-1}, 1, -r_{i, i+1}, \dots, -r_{in}) \quad (2.170)$$

satırı hariç birim matrise eşittir.

A'nın R_i ile soldan çarpımı, i. satırdan diğer satırların çarpımının çıkarılmasına sebep olur. Diğer taraftan sağdan çarpım, i. kolonun çarpımının diğer kolonlardan çıkartılmasına sebep olur. Karşılık gelen R_i , her bir $-r_{ij}$ elemanlarının, $+r_{ij}$ elemanları ile değiştirilmiş aynı yapıda bir matristir. A'nın $R^{-1}AR$ benzerlik dönüşümü, A'nın i. kolonunun uygun çarpımlarının farkı ve i. satırın diğer tüm satırlara eklenmesi ile elde edilir.

İlk adımda i. kolon hariç hepsi değişirken, ikinci adımda sadece i. satır değişmektedir.

(iv) S_i matrisleri, R_i -nin transpozesi ile elde edilir. S_i -nin i. kolonu bir satır vektörü olarak

$$(-s_{1i}, -s_{2i}, \dots, -s_{i-1, i}, 1, -s_{i+1, i}, \dots, -s_{ni}) \quad (2.171)$$

gibi yazılır.

(v) M_i matrisleri, i. satırı

$$(0, 0, \dots, 0, 1, -m_{i, i+1}, \dots, -m_{in}) \quad (2.172)$$

şeklinde olan diğer yerleri birim matris gibi olan matrislerdir. M_i ayrıca, $r_{i1}, \dots, r_{i,i-1}$ elemanlarının hepsi sıfır olan bir R_i matrisidir.

(vi) N_i matrisi, M_i -nin transpozesiyle elde edilir. i . kolon bir satır vektörü

$$(0, 0, \dots, 0, 1, -n_{i+1,i}, \dots, -n_{ni}) \quad (2.173)$$

gibi yazılır. N_i ilk elemanları sıfır olan bir S_i dir.

(vii) M_{ij} matrisleri (i,j) elemanı $-m_{ij}$, geri kalan elemanlar ise birim matrise eşittir. M_{ij} böylece önceki matrislerin özel bir formudur.

2.32 Elementer Matrislerin Özellikleri

Aşağıda temel matrislerin önemli özellikleri yazılmıştır:

- (i) $M_{n-1}M_{n-2}\dots M_1$ çarpanı uniter köşegeninde 1 elemana sahip *üstüçgen-matris*leridir. $j>i$ için (i,j) pozisyonundaki eleman, $-m_{ij}$ -dir. m_{pq} elemanlarının hiçbir çarpanı bulunmaz.
- (ii) $M_1, M_2, \dots, M_{n-1}$ çarpımı uniter köşegeninde 1 bulunan bir *üstüçgen-matristir*. Fakat m_{pq} -nun çarpımları diğer elemanları etkilediğinden, yapısı diğer çarpımlardan çok daha komplekstir.
- (iii) Benzer olarak, $N_1N_2\dots N_{n-1}$ ve $N_{n-1}N_{n-2}\dots N_1$ köşegeninde 1 bulunan *altıçgen-matris*lerdir. $i>j$ için (i,j) deki elemanlar $-n_{ij}$ dir. Geri kalanları $M_1M_2\dots M_{n-1}$ çarpanlarından daha kompleks bir yapıya sahiptir.

2.33 Temel Benzerlik Dönüşümleri ile Üçgen Kanonik Forma İndirgenme

Herhangi bir matris benzerlik-dönüşümleri ile üçgen-forma indirgenebileceğini göstereceğiz. Bu matris bir grup elementer matrislerin çarpımıdır. İspat tümevarım ile yapılır.

$(n-1)$ mertebeden bir matris için bunun doğru olduğunu kabul edelim. A n . mertebeden bir matris ve λ bir özdeğer olsun. Böylece en az bir tane λ -ya karşılık gelen

$$Ax = \lambda x \quad (2.174)$$

gibi bir x özvektörü mevcuttur.

Şimdi herhangi bir tekil-olmayan X matrisi için XAX^{-1} in özdeğeri λ dir. Xx , e_i olacak şekilde X i nasıl seçileceğini inceleyelim.

$$x^T = (x_1, x_2, \dots, x_{r-1}, 1, x_{r+1}, \dots, x_n) \quad (2.175)$$

Burada r . eleman x in sıfır-olmayan ilk elemanıdır ve bunu 1 yapmak için keyfi çarpanlar seçilebilir.

$$(I_{1r}x)^T = (1, y_2, y_3, \dots, y_n) = y^T \quad (2.176)$$

Burada y_i -ler x_i -nin mertebesindedir. $I_{1,1} = I$ olduğundan, $r = 1$ durumunda kapsanmaktadır. Eğer y yi, $n_{i1}=y_i$ olan bir N_1 matrisi ile soldan çarpılırsa, ilki hariç hepsi sıfır olacaktır. Böylece

$$N_1 y = N_1 I_{1r} x = e_i \quad (2.177)$$

olur ve $N_1 I_{1r} A I_{1r} N_1^{-1}$ matrisi λ özdeğerine ve karşılık gelen özvektör e_i e sahip olur. Bu

$$\left| \begin{array}{cc} \lambda & b^T \\ 0 & B \end{array} \right| \quad (2.178)$$

gibidir. Burada B $(n-1)$ mertebeden bir kare matristir. Hipotezden

$$HBH^{-1} = T \quad (2.179)$$

gibi temel matrislerin çarpımı olan bir H kare matris vardır. Burada T bir *üçgen-form*dur.

$$\left| \begin{array}{cc} 1 & 0 \\ 0 & H \end{array} \right| \left| \begin{array}{cc} \lambda & b^T \\ 0 & B \end{array} \right| \left| \begin{array}{cc} 1 & 0 \\ 0 & H^{-1} \end{array} \right| = \left| \begin{array}{cc} \lambda & b^T H^{-1} \\ 0 & HBH^{-1} \end{array} \right| = \left| \begin{array}{cc} \lambda & b^T H^{-1} \\ 0 & T \end{array} \right| \quad (2.180)$$

Böylece sonuç, $n=1$ için doğru olduğundan genel durum içinde geçerlidir.

2.34 Temel Uniter Dönüşümler

Son bölümde temel matrisleri, *sıfır-olmayan* genel vektörü ke_1 -li ifadeye döndürmek için kullandık. Aynı etkiye sahip iki temel sınıftan oluşan temel uniter matrislerden ilkinin tanımlayalım.

Aşağıdaki gibi tanımlanan R matrisini göz önüne alalım:

$$\begin{aligned} r_{ii} &= e^{i\alpha} \cos\theta & r_{ij} &= e^{i\beta} \sin\theta \\ r_{ji} &= e^{i\gamma} \sin\theta & r_{jj} &= e^{i\delta} \cos\theta \\ r_{pp} &= 1 \quad (p \neq i,j) & r_{pq} &= 0 \end{aligned} \quad (2.181)$$

Burada α, γ ve δ reeldirler. Eğer aşağıdaki koşul sağlanırsa $R^H R$ formuyla da doğrulanacağı üzere-bu matris uniterdir.

$$\alpha - \gamma \equiv \beta - \delta + \pi \pmod{2\pi} \quad (2.182)$$

$$\gamma = \pi - \beta, \quad \delta = -\alpha \quad (2.183)$$

$$\begin{aligned} r_{ii} &= e^{i\alpha} \cos\theta, & r_{ij} &= e^{i\beta} \sin\theta \\ r_{ji} &= -e^{-i\beta} \sin\theta, & r_{jj} &= e^{-i\alpha} \cos\theta \end{aligned} \quad (2.184)$$

Bu kısıtlamalar altındaki sınıfı *düzlemsel-dönme* (*plane-rotation*) olarak ifade edeceğiz.

(2.184) den

$$r_{ii} = r_{ii}^*, \quad r_{ij} = -s^*, \quad r_{ji} = -s, \quad r_{jj} = c \quad (2.185)$$

yazılabilir. Burada $|c|^2 + |s|^2 = 1$ dir.

R^H , aşağıdaki ifadeyi sağlayan α ve β ne karşılık gelen bir *düzlemsel-dönme*dir.

$$\alpha = -\alpha, \quad \beta = \beta + \pi \quad (2.186)$$

Bir *düzlemsel-dönme*nin terside böylece bir *düzlemsel-dönme*dir. $\alpha = \beta = 0$ iken matris reel ve ortogonaldır. Aynı zamanda (i,j) düzleminde, açısı doğrultusundaki bir dönme ile ilişkisi açıktır.

Eğer x herhangi bir vektör i ve j verilsin. Öyle bir *düzlemsel-dönme* seçebiliriz ki $(Rx)_i$ reel olsun ve negatif olmasın ve $(Rx)_j$ ise sıfır olsun. Açık olarak tüm $s \neq i, j$ için $(Rx)_s = x_s$ olur.

$$(Rx)_j = -sx_i + cx_j \quad (2.187)$$

Eğer $r^2 = |x_i|^2 + |x_j|^2$ yazılırsa, elde edilen r sıfır değildir ve aşağıdaki ifade yazılabilir:

$$s = x_j / r \quad c = x_i / r \quad (2.188)$$

Burada r pozitif olarak seçilebilir. Böylece $(Rx)_j = 0$ olur ve

$$(Rx)_i = cx_i + sx_j = |x_i|^2 / r + |x_j|^2 / r = r > 0 \quad (2.189)$$

elde edilir. Eğer $r = 0$ ise bu durumda $c=1$ ve $s=0$ alınabilir. Eğer $(Rx)_i$ reel istenmezse, bu durumda (2.182) denklemindeki α parametreleri düşürülebilir. Eğer x_i reel ise (2.188) c -nin değerini reel olarak verir.

x vektörü, $(1,2)$, $(1,3)$, ..., $(1,n)$ düzlemlerinde, $(n-1)$ *düzlemsel-dönmeleri* ile soldan çarpımıyla ke_1 -e dönüştürülebilir. $(1,i)$ düzlemindeki rotasyon, i . elemanı sıfıra indirgeyecek şekilde seçilir. k , reel ve negatif olmayacak şekilde seçilebilir.

$$k = (|x_1|^2 + |x_2|^2 + \dots + |x_n|^2)^{1/2} = (x^H x)^{1/2} \quad (2.190)$$

Eğer, x ve y , $(x^H x) = (y^H y)$ eşitliğini sağlıyorsa, aşağıdaki gibi, rotation matrisinin bir S çarpanı mevcuttur.

$$Sx = y \quad (2.191)$$

Bu, x ve y -nin, *düzlemsel-dönme* matrislerin bir çarpanı tarafından ke_1 e dönüştürülebilmesinden ve bir *düzlemsel-dönme* matrisinin tersinin de bir *düzlemsel-dönme* matrisi olmasından gelir.

2.35 Temel Uniter Hermitian Matrisleri

Temel uniter matrislerin ikinci sınıfı

$$P = I - 2\omega\omega^H, \quad \omega^H\omega = 1 \quad (2.192)$$

gibi tanımlanan P matrisinden oluşur.

$$\begin{aligned} PP^H &= (I - 2\omega\omega^H)(I - 2\omega\omega^H) \\ &= I - 4\omega\omega^H + 4\omega(\omega^H\omega)\omega^H = I \end{aligned} \quad (2.193)$$

olduğundan matrisler *uniter* ve *Hermitian* dır.

Bundan böyle bu tür matrisleri temel Hermitian matrisleri olarak ifade edeceğiz. ω bir reel vektör olduğunda, P ortogonal ve simetriktir.

Eğer y ve x

$$y = Px = x - 2\omega(\omega^Hx) \quad (2.194)$$

sağlıyor ise

$$y^Hy = x^HP^HPx = x^Hx \quad (2.195)$$

$$x^Hy = x^HPx \quad (2.196)$$

elde edilir ve (39.5) in sağ tarafı, P Hermitian olduğundan, reeldir. Böylece bir x ve y verildiğinde, eğer $x^Hx = y^Hy$ ve x^Hy reel değil ise, $y = Px$ olacak şekilde bir P bulunmayabilir. Bu şartlar sağlandığında P -yi aşağıdaki gibi bulabiliriz.

Eğer $x=y$ ise, $\omega^Hx = 0$ olacak şekilde uygun bir P verecek bir ω vardır. Aksi takdirde (39.3) den, herhangi bir uygun ω , $y-x$ yönündedir.

$$\omega = e^{i\alpha} (y-x) / [(y-x)^H(y-x)]^{1/2} \quad (2.197)$$

$e^{i\alpha}$ çarpanının P ye bir etkisi yoktur. Uygun bir ω hemen gösterilebilir.

$$(I - 2\omega\omega^H)x = x - \frac{2(y-x)(y-x)^Hx}{(y-x)^H(y-x)}, \quad (2.198)$$

$$\begin{aligned} 2(y-x)^Hx &= 2(y^Hx - x^Hx) \\ &= y^Hx + x^Hy - x^Hx - y^Hy \\ &= -(y-x)^H(y-x) \end{aligned} \quad (2.199)$$

Kabulden x^Hy reel ve $x^Hx = y^Hy$ dir.

$$(I - 2\omega\omega^H)x = x + (y-x) = y \quad (2.300)$$

Uygun bir P matrisinin soldan çarpımı ile, x -i ke_1 (k reel bir sayı olmak üzere) şekline genelde çevrilemez. Eğer x in bileşenleri $r_1 e^{i\theta_1}$ ile ifade edilirse, x aşağıdaki gibi bir y -ye çevrilebilir:

$$y = \pm (x^Hx)^{1/2} e^{i\theta_1} e_1 \quad (2.301)$$

Çünkü x^Hy reel dir. Uygun bir P için

$$\omega = (x-y) / [(x-y)^H(x-y)]^{1/2} \quad (2.302)$$

yazılabilir. Eğer $r_1=0$ ise θ_1 keyfi olarak alınabilir. Ve eğer $\theta_1=0$ olarak alınırsa, karşılık gelen P matrisi, x i, e_1 in reel bir çarpanına çevirir.

Genellikle P matrisinin, karşılık gelen ω nın 0 dan $n-1$ e kadar r -nin bazı değerleri için ilk r pozisyonundaki sıfır bileşenlerine sahip olması ile ilgileneceğiz. Bir x vektörünün herhangi bir P ile soldan çarpımı x -in ilk r bileşenini değiştirmez.

Eğer x reel ise, (2.301) ve (2.302) deki y ve ω da reeldir ve böylece P de reel olur. $e^i P$ formundaki uniter matrisler ile, keyfi bir vektör ke_1 formuna indirgenebilir. (k reeldir)

2.36 Temel Uniter Dönüşümler ile Üçgen Forma İndirgeme

Daha önce bahsedilen (§23,25) ortogonal *uniter-dönüşümler* ile kanonik forma indirgenmesi konusundaki iki sonucu ispat edelim. İlk olarak herhangi bir matrisin bir uniter dönüşüm ile üçgen forma indirgenebileceğini gösterelim. İspat tümevarım ile yapılacaktır.

(n-1) mertebeden bir matris için bunun doğru olduğunu kabul edelim. A, n mertebeden bir matris olsun ve λ ile x , A'nın özdeğer ve özvektörleri olsun. Öyle ki

$$Ax = \lambda x \quad (2.203)$$

§38 ve §40 dan aşağıdaki gibi bir uniter matris P vardır:

$$Px = ke_1 \quad (2.304)$$

P, temel Hermitian matrislerden bir tanesi veya (n-1) *düzlemsel-dönme* matrislerin çarpımı olsun. (41.1) den

$$\begin{aligned} PAx &= P\lambda x & PA(P^H P)x &= \lambda Px \\ (PAP^H)ke_1 &= \lambda ke_1 & (PAP^H)e_1 &= \lambda e_1 \quad k \neq 0 \end{aligned} \quad (2.305)$$

olur. PAP^H matrisi aşağıdaki gibi $A^{(1)}$ formunda olur.

$$A^{(1)} = \begin{vmatrix} \lambda & b^T \\ 0 & A_{n-1} \end{vmatrix} \quad (2.306)$$

Burada A_{n-1} (n-1) mertebesinde. Şimdi hipotezden

$$RA_{n-1}R^H = T_{n-1} \quad (2.307)$$

gibi bir R uniter matrisi vardır. Burada T_{n-1} *üçgen-matristir*. Böylece

$$\begin{vmatrix} 1 & 0 \\ 0 & R \end{vmatrix} A^{(1)} \begin{vmatrix} 1 & 0 \\ 0 & R^H \end{vmatrix} = \begin{vmatrix} \lambda & b^T R^H \\ 0 & T_{n-1} \end{vmatrix} \quad (2.308)$$

ve (2.308) nın sağ tarafındaki *üçgen-matristir*. Çünkü

$$\begin{vmatrix} 1 & 0 & P \\ 0 & R & \end{vmatrix}$$

uniter olduğundan sonuç kurulmuştur. Eğer A ve onun özdeğerleri reel ise bu durumda, özvektörlerinde reel olacağı ve sonuç olarak P matrisinin de ortogonal olacağı görülür. Burada dönüşümün tümünde ortogonal matrisler kullanılmaktadır. Eğer A bir Hermitian ise, bu

durumda *üçgen-matris* köşegendir. Görüleceği gibi uniter bir R matrisi vardır ve T *üçgen-matristir*.

$$RAR^H = T \quad (2.309)$$

Soldaki matris Hermitian matristir ve böylece T de bir Hermitian ve köşegendir.

2.37 Normal Matrisler

Herhangi bir Hermitian matris için aşağıdaki gibi bir uniter matrisin mevcut olduğu ispat edildi.

$$RAR^H = D \quad (2.310)$$

Burada D köşegen ve reeldir. Bu sınıfı D -nin kompleks olduğu duruma genişletelim. Diğer bir deyişle, D köşegen ve R uniter iken $R^H D R$ ile ifade edilen A matrisi göz önüne alalım. Eğer

$$A = R^H D R \quad (2.311)$$

ise

$$AA^H = R^H D R R^H D^H R = R^H (D D^H) R \quad (2.312)$$

$$A^H A = R^H D^H R R^H D R = R^H (D^H D) R \quad (2.313)$$

Böylece $DD^H = D^H D$ olduğundan, $AA^H = A^H A$ olur. Karşıt olarak eğer

$$AA^H = A^H A \quad (2.314)$$

kullanılırsa, A (2.311) de olduğu gibi bir çarpanlara ayrılabilir. Herhangi bir A matrisi $R^H T R$ formunda ifade edilebilir ki burada R uniter ve T *üstüçgen-matristir*. Böylece

$$R^H T R R^H T^H R = R^H T^H R R^H T R \quad (2.315)$$

$$R^H T T^H R = R^H T^H T R$$

elde edilir. Ayrıca R -nin soldan çarpımı ve R^H -nin sağdan çarpımı ile

$$T T^H = T^H T \quad (2.316)$$

elde edilir. Köşegen elemanlar eşitleyerek, T -nin tüm köşegen dışındaki elemanlar sıfır bulunur. Böylece T köşegendir. $R^H D R$ formuna getirilen matris sınıfı, $A A^H = A^H A$ için aynıdır. Böyle matrisler *normal-matrisler* olarak adlandırılır. (2.314) denkleminde, normal matrisler aşağıdaki özelliklere sahiptir:

$$\begin{aligned} A^H &= A && \text{(Hermitian matrisler)} \\ A^H &= -A && \text{(Ters-Hermitian matrisler)} \\ A^H &= A^{-1} && \text{(Uniter matrisler)} \end{aligned}$$

2.38 Değişmeli Matrisler

Normal matrislerin A ve A^H matrisleri yer değiştiren özelliğe sahiptirler. Aşağıdaki ilişkiden

$$A = R^H D R \tag{2.317}$$

$$A R^H = R^H D \tag{2.318}$$

$$A^H R^H = R^H D^H \tag{2.319}$$

A ve A^H matrisleri böylece ortak bir özvektörler sistemine sahiptir. Yani, R^H -in kolonlarıyla yapılandırılmışlardır.

Eğer herhangi A ve B matrisleri ortak bir özvektörler sistemine sahip ise bu durumda $AB=BA$ olduğunu göstereceğiz. Eğer özvektörlerin bu sistemi H matrisinin kolonlarını yapılandırıyorsa

$$A = H D_1 H^{-1} \quad B = H D_2 H^{-1} \tag{2.320}$$

$$AB = H D_1 H^{-1} H D_2 H^{-1} = H D_1 D_2 H^{-1} \tag{2.321}$$

$$BA = H D_2 H^{-1} H D_1 H^{-1} = H D_2 D_1 H^{-1} \tag{2.322}$$

elde edilir. Köşegen-matrisler yerdeğiştirmeli matris olduğundan, $AB=BA$ olur.

Karşıt olarak, eğer A ve B yerdeğiştirmeli matris ve *lineer temel-bölenlere* sahip ise, A ve B ortak özvektörler sistemini paylaşırlar. A nın özvektörlerini $h_1, h_2, \dots, h_n$ olduğu kabul edilirse bunlar H matrisini yapılandırır. Bu durumda

$$H^{-1}AH = \text{diag}(\lambda_i) \quad (2.323)$$

olur ki burada λ_i ayrıık olabilir de olmayabilir de. Eđer izole edilmiş $\lambda_1=\lambda_2=\lambda_3$, $\lambda_4=\lambda_5$ ve λ_6 , λ_7 ve λ_8 için sekizinci mertebeden bir matrisi göz önüne alınırsa, denklem (2.323)

$$H^{-1}AH = \begin{vmatrix} \lambda_1 I_3 & & & & & & & \\ & \lambda_4 I_2 & & & & & & \\ & & \lambda_6 & & & & & \\ & & & \lambda_7 & & & & \\ & & & & \lambda_8 & & & \end{vmatrix} \quad (2.324)$$

gibi yazılabilir. Burada $Ah_i = \lambda_i h_i$ olduğundan

$$BAh_i = \lambda_i Bh_i \quad (2.325)$$

$$A(Bh_i) = \lambda_i (Bh_i) \quad (2.326)$$

bulunur. Bu, eđer h_i , λ_i ye karşılık gelen A nın bir özvektörü ise, Bh_i -nin, λ_i -ye karşılık gelen özvektörlerin alt uzayına içinde olduğunu gösterir. Böylece

$$\begin{aligned} Bh_1 &= p_{11}h_1 + p_{21}h_2 + p_{31}h_3 & Bh_6 &= \mu_6 h_6 \\ Bh_2 &= p_{12}h_1 + p_{22}h_2 + p_{32}h_3 & Bh_7 &= \mu_7 h_7 \\ Bh_3 &= p_{13}h_1 + p_{23}h_2 + p_{33}h_3 & Bh_8 &= \mu_8 h_8 \\ Bh_4 &= q_{11}h_4 + q_{21}h_5 & & \\ Bh_5 &= q_{12}h_4 + q_{22}h_5 & & \end{aligned} \quad (2.327)$$

Bu denklemler aşağıdaki formda tek bir matris formunda yazılabilirler.

$$H^{-1}BH = \begin{vmatrix} P & & & & & & & \\ & Q & & & & & & \\ & & \mu_6 & & & & & \\ & & & \mu_7 & & & & \\ & & & & \mu_8 & & & \end{vmatrix} \quad (2.328)$$

Şimdi B *linear temel-bölenlere* sahip olduğundan, (2.328) denkleminin sağındaki matris ve P ile Q içinde doğru olmalıdır. İkinci ve üçüncü dereceden K ve L matrisleri böylece

$$AK^{-1}PK = \begin{vmatrix} \mu_1 & & & & \\ & \mu_2 & & & \\ & & \mu_3 & & \end{vmatrix} \text{ ve } L^{-1}QL = \begin{vmatrix} \mu_4 & & & & \\ & \mu_5 & & & \\ & & \mu_6 & & \\ & & & \mu_7 & \\ & & & & \mu_8 \end{vmatrix} \quad (2.329)$$

gibidir. Buradan

$$G = \begin{vmatrix} K & & \\ & L & \\ & & I_3 \end{vmatrix} \quad (2.330)$$

$$G^{-1}H^{-1}BHG = \text{diag}(\mu_i) \quad (2.331)$$

elde edilir. Diğer taraftan (2.328) den

$$G^{-1}H^{-1}AHG = \text{diag}(\lambda_i) \quad (2.332)$$

bulunur. Böylece B ve A ortak özvektörler sistemini paylaşırlar, yani HG -nin kolonlarıdır.

2.39 AB -nin Özdeğerleri

Her kare A ve B için AB ve BA nın özdeğerleri aynıdır. İspat aşağıdadır.

$$\begin{vmatrix} I & O \\ -B & \mu I \end{vmatrix} = \begin{vmatrix} \mu I & A \\ B & \mu I \end{vmatrix} = \begin{vmatrix} \mu I & A \\ O & \mu^2 I - BA \end{vmatrix} \quad (2.333)$$

$$\begin{vmatrix} \mu I & -A \\ O & I \end{vmatrix} = \begin{vmatrix} \mu I & A \\ B & \mu I \end{vmatrix} = \begin{vmatrix} \mu^2 I - AB & O \\ B & \mu I \end{vmatrix} \quad (2.334)$$

(2.333) ve (2.334) nin her iki tarafının determinantlarını alarak

$$\begin{vmatrix} \mu I & A \\ B & \mu I \end{vmatrix} = X \quad (2.335)$$

$$\mu^n \det(X) = \mu^n \det(\mu^2 I - BA) \quad (2.336)$$

$$\mu^n \det(X) = \mu^n \det(\mu^2 I - AB) \quad (2.337)$$

elde edilir. (2.336) ve (2.337) denklemlerinde $\mu^2 = \lambda$ yazarak

$$\det(\lambda I - BA) = \det(\lambda I - AB) \quad (2.338)$$

bulunur. Bu AB ve BA nın aynı özdeğerlere sahip olduğunu gösterir.

Aynı ispat, eğer A $m \times n$ ve B $n \times m$ -lik matrisler ise, AB ve BA nın, $|m-n|$ ekstra sıfır özdeğerlere sahip yüksek dereceden olanların çarpımı hariç, aynı özdeğerlere sahip olduğunu gösterir.

2.40 Vektör ve Matris Normları

Bir vektör veya matrisin büyüklüğü hakkında tüm bilgileri veren ve kompleks bir sayının modülü gibi bir rol oynayan tek bir sayının var olması vektör ve matris hakkında bazı temel bilgileri vermesi açısından oldukça faydalıdır. Bu amaç için, *norm* olarak adlandırılan, bir matris veya vektörün elemanlarının bazı fonksiyonlarını kullanacağız.

Bir vektörün normu $\|x\|$ olarak ifade edilecek ve aşağıdaki koşulları sağlayacaktır:

$$\begin{aligned} \text{(i)} \quad \|x\| &> 0, & (x \neq 0) \\ \text{(ii)} \quad \|kx\| &= |k| \|x\| \\ \text{(iii)} \quad \|x + y\| &\leq \|x\| + \|y\| \\ \text{(iv)} \quad \|x - y\| &\geq \|x\| - \|y\| & ((i) \text{ ve } (ii) \text{ den}) \end{aligned} \quad (2.339)$$

Biz burada sadece üç basit vektör normu kullanacağız.

$$\|x\|^p = (|x_1|^p + |x_2|^p + \dots + |x_n|^p)^{1/p} \quad (p = 1, 2, \infty) \quad (2.340)$$

Burada $\|x\|$ ile $\max |x_i|$ aynı anlamdadır. $\|x\|_2$, x vektörünün Euclidian uzunluğudur. Bunların haricinde normlar hakkında aşağıdaki ifadelerde yazılabilir:

$$x^H x = \|x\|_2^2 \quad (2.341)$$

$$\begin{aligned} |x^H y| &\leq \sum |x_i| |y_i| \\ &\leq \|x\|_2 \cdot \|y\|_2 \end{aligned} \quad (2.342)$$

$$x^H y = \|x\|_2 \|y\|_2 e^{i\theta} \cos \theta \quad (2.343)$$

x ve y reel olduğu zaman, θ iki vektör arasındaki açıyı temsil eder.

Benzer olarak bir A matrisinin normu $\|A\|$ şeklinde ifade edilecek ve bu normlar aşağıdaki koşulları sağlayacaklardır:

$$\begin{aligned}
(i) \quad & \|A\| > 0 && \text{eğer } A \neq 0 \text{ ise} \\
(ii) \quad & \|kA\| = |k| \cdot \|A\| && \text{her hangi bir kompleks skaler için} \\
(iii) \quad & \|A+B\| \leq \|A\| + \|B\| && \\
(iv) \quad & \|AB\| \leq \|A\| \|B\| &&
\end{aligned} \tag{2.344}$$

2.41 Matris Normları

$\sup_{x \neq 0} \|Ax\| / \|x\|$ şeklinde tanımlanan vektör normu ile bir A matrisi arasında bir ilişki kurulabilir.

(2.339) in (ii) ilişkisinden, bunun $\sup_{\|x\|=1} \|Ax\|$ -e eşit olduğu görülür. Bu sayı açık olarak, A

matrisinin bir fonksiyonudur ve bu, bir matris normunun istenilen koşullarını sağlayacak şekilde kolaylıkla değiştirilebilir. Buna *vektör normundan üretilen matris normları* adı verilir.

$$\|A\| = \sup_{x \neq 0} \|Ax\| / \|x\| \tag{2.345}$$

$$\|Ax\| \leq \|A\| \cdot \|x\| \tag{2.346}$$

Yukarıdaki ifadeler, tüm *sıfır-olmayan* x ve x=0 için doğrudur.

Bu şekilde verilen normlar için, tanımdan, $\|I\|=1$ eşitliği vardır. Dahası

$$\|Ax\| = \|A\| \|x\|$$

gibi bir sıfır olmayan vektör daima vardır. Böylece $\sup \|Ax\| / \|x\|$ yerine $\max \|Ax\| / \|x\|$ yazılabilir. $\|x\|_p$ den elde edilen norm $\|A\|_p$ ile gösterilir. Bu norm aşağıdaki koşulları sağlar:

$$\|A\|_1 = \max_j \sum_i |a_{ij}| \tag{2.347}$$

$$\|A\|_\infty = \max_i \sum_j |a_{ij}| \tag{2.348}$$

$$\|A\|_2 = (\lambda^{\max}(A^H A))^{1/2} \tag{2.349}$$

İlk iki sonuç açıktır ve $\|A\|_1 = \|AH\|_\infty$ olur. Üçüncüsü aşağıda olduğu gibi ispatlanması gerekir.

$A^H A$ matrisi Hermitian dır ve özdeğerleri, $x^H A^H A x = (Ax)^H (Ax) > 0$ olduğundan negatif değildir. Bunun özdeğerlerini δ_2 ile ifade edelim. Burada $\delta_1^2 > \delta_2^2 > \dots > \delta_n^2 > 0$ dır.

Tanımdan

$$\|A\|_2 = \max_{x \neq 0} \|Ax\|_2 / \|x\|_2 \quad (2.350)$$

$$\|A\|_2^2 = \max \|Ax\|_2^2 / \|x\|_2^2 = \max(x^H A^H A x / x^H x) \quad (2.351)$$

elde edilir.

Eğer $u_1, u_2, \dots, u_n$ $A^H A$ nın özvektörlerinin bir ortonormal sistemi ise

$$x = \sum \alpha_i u_i \quad (2.352)$$

$$(x^H A^H A x) / (x^H x) = (\sum |\alpha_i|^2 \delta_i^2) / (\sum |\alpha_i|^2) \quad (2.353)$$

yazılabilir. $x=u_i$ iken δ_i^2 değerine ulaşılır. *Negatif-olmayan* bir sayı olan δ_i ne, A nın *tekil-değeri* adı verilir. 2-norm genellikle *spectral-norm* olarak adlandırılır.

2.42 Euclidian ve Spectral Normlar

$\|x\|_2$ vektör normu ile uyumlu olan ikinci bir önemli norm daha vardır. Bu $\|A\|_E$ ile ifade edilen, Euclidian veya Schur normudur. Bu norm

$$\|A\|_E = (\sum |a_{ij}|^2)^{1/2} \quad (2.354)$$

gibi tanımlanır ve

$$\|I\|_E = n^{1/2} \quad (2.355)$$

olduğundan, herhangi bir vektör normundan elde edilemez. (Gould S.H, 1974)

Kolayca hesaplanabilir olduğundan, Euclidian norm pratik amaçlar için oldukça kullanışlıdır.

$$\| |A| \|_E = \|A\|_E \quad (2.356)$$

1 ve ∞ matris normları, benzer ilişkileri sağlar.

$$\| |A| \|_2 \neq \|A\|_2 \quad (2.357)$$

$\|A\|_E^2$, $A^H A$ nın izine (*trace*) eşit olduğundan, $\sum \delta_i^2$

$$\|A\|_2 < \|A\|_E < n^{1/2} \|A\|_2 \quad (2.358)$$

yazılabilir. Burada her iki sınırı da alabilir. Böylece

$$\| |A| \|_2 < \| |A| \|_E = \|A\|_E < n^{1/2} \|A\|_2 \quad (2.359)$$

elde edilir. Euclidian normunun uygun olmayan tarafı aşağıdaki ifadelerdir:

$$\|\text{diag}(\lambda_i)\|_E = (\sum |\lambda_i|^2)^{1/2} \quad (2.360)$$

Halbuki

$$\|\text{diag}(\lambda_i)\|_p = \max |\lambda_i| \quad (p=1, 2, \infty) \quad (2.361)$$

Euclidian matris normu ve 2-vektör ve 2-matris normları, uniter dönüşüm altında değişmeyen önemli özelliklere sahiptir. Keyfi bir x ve A ile uniter R için aşağıdaki ifadeler yazılır

$$\|Rx\|_2 = \|x\|_2 \quad (2.362)$$

$$\|Rx\|_2^2 = (Rx)^H Rx = x^H R^H Rx = x^H x = \|x\|_2^2 \quad (2.363)$$

Benzer olarak

$$\|RA\|_2 = \|A\|_2 \quad \text{ve} \quad \|RAR^H\|_2 = \|A\|_2 \quad (2.364)$$

tanımdan hemen elde edilir. RA nın her bir kolonunun Euclidian uzunluğu, A nın karşılık gelen kolonuna eşit olduğundan

$$\|RA\|_K = \|A\|_K \quad (2.365)$$

doğrudur..

$A = R^H \text{diag}(\lambda_i) R$ -nin bir normal matris olduğunu kabul edelim. Bu durumda

$$\begin{aligned} \|AB\|_E &= \|R^H \text{diag}(\lambda_i) RB\|_E = \|\text{diag}(\lambda_i) RB\|_E < \max |\lambda_i| \cdot \|RB\|_E \\ &= \max |\lambda_i| \cdot \|B\|_E \end{aligned} \quad (2.366)$$

yazılabilir.

Bundan sonra her ne zaman bir normdan bahsedilirse, burada tanıtılan normlardan biri bahsedilmiş olacaktır.

2.43 Normlar ve Limitler

Matris dizisinin yakınsaması için basit bir kriter sayısı elde etmek oldukça faydalıdır.

Tanımdan kolaylıkla elde edilen bu sonuçların ispatı üzerinde durulmayacaktır.

$$\begin{aligned} (A) \quad \lim_{r \rightarrow \infty} A^{(r)} &= A & (\text{yalnız ve yalnız } \|A - A^{(r)}\| \rightarrow 0 \text{ ise}) \\ \|A^{(r)}\| &\rightarrow \|A\| & (\|A^{(r)}\| \rightarrow \|A\| \text{ ise}) \end{aligned}$$

$$(B) \quad \lim_{r \rightarrow \infty} A^{(r)} = 0 \quad (\|A\| < 1 \text{ ise})$$

(C) Eğer λ , A nın bir özdeğeri ise $\lambda < \|A\|$ olur.

$$|\lambda| \cdot \|x\| = \|\lambda x\| = \|Ax\| < \|Ax\| < \|A\| \cdot \|x\| \quad (3.367)$$

$Ax = \lambda x$ sıfır olmadığından buradan uygun matris ve vektör normu için

$$|\lambda| < \|A\|$$

bulunur. Buradan

$$\begin{aligned} \|A\|_2^2 &= (A^H A) \quad \text{-nın max özdeğeridir} \\ &< \|A^H A\|_1 < \|A^H\|_1 \cdot \|A\|_1 = \|A\|_\infty \cdot \|A\|_1 \end{aligned} \quad (3.368)$$

(D) Yalnız ve yalnız tüm özvektörler için $|\lambda_i| < 1$ iken $r \rightarrow \infty$ $A^r = 0$ olur. Bu *Jordan-kanonik form*dan gelmektedir. Tipik bir *Jordan-altmatris*den

$$\begin{vmatrix} a & & \\ 1 & a & \\ & 1 & a \end{vmatrix} = a^{r-2} \begin{vmatrix} a^2 & & \\ ra & a^2 & \\ \frac{1}{2} r(r-1) & ra & a^2 \end{vmatrix} \quad (3.369)$$

elde edilir. Ve yalnız ve yalnız $|a| < 1$ iken sıfıra gider. Burada (D) den (C) yi ispatlamak uygundur. Aşağıdaki gibi tanımlanan $\rho(A)$ sayısı, A nın *spectral-çapı* olarak adlandırılır.

$$\rho(A) = \max |\lambda_i| \quad (3.370)$$

$$\rho(A) < \|A\| \quad (3.371)$$

(E) $I + A + A^2 + \dots$ serisi, yalnız ve yalnız $A^r \rightarrow 0$ iken yakınsaktır. Koşulun yeterliliği (D) den gelmektedir. Eğer bu sağlanırsa, A nın özdeğerlerinin tümü birim çember içerisinde bulunmaktadır ve böylece $(I - A)$ tekil değildir.

$$(I - A)(I + A + A^2 + \dots + A^r) = I - A^{r+1} \quad (3.372)$$

$$I + A + A^2 + \dots + A^r = (I - A)^{-1} - (I - A)^{-1} A^{r+1} \quad (3.373)$$

Sağ taraftaki ikinci terim sıfıra gider.

$$I + A + A^2 + \dots \rightarrow (I - A)^{-1}$$

(F) $I + A + A^2 + \dots$ serisinin yakınsaklığının yeterli bir kriteri, A nın normlarından herhangi birinin 1 den küçük olmasıdır. Eğer bu doğru ise (C) den

$$|\lambda_i| < 1$$

yazılır. Böylece (D) den $A^r \rightarrow 0$ olur ve (E) den dolayı seriler yakınsaktır.

$\|A\| > 1$ gibi bir normun varlığı serinin yakınsaklığını imkansızlaştırmadığına dikkat edilmelidir. Örneğin

$$A = \begin{vmatrix} 0.9 & 0 \\ 0.3 & 0.8 \end{vmatrix} \quad (2.374)$$

ise, özdeğerleri 0.9 ve 0.8 olduğundan $A^r \rightarrow 0$ iken $\|A\|_\infty = 1.1$, $\|A\|_1 = 1.2$ ve $\|A\|_E = (1.54)^{1/2}$ olur.

Pratik uygulamalarda bir serinin yakınsaklığını kurma işleminde en küçük normlardan herhangi birini seçme konusunda serbest oluruz.

(E) deki serinin yakınsaklığı için yeterli bir kriter, A'nın herhangi bir benzerlik-dönüşümünün herhangi bir normunun 1 den küçük olmasıdır. Eğer $A = HBH^{-1}$ ise

$$I + A + \dots + A^r = H(I + B + \dots + B^r)H^{-1} \quad (2.375)$$

elde edilir. Eğer $\|B\| < 1$ ise sağ taraftaki seri yakınsaktır.

Orijinal matrisin normlarından daha küçük standart normlar elde etmek için benzerlik-dönüşümü kullanılabilir. (F) de verilen matris için

$$B = \begin{vmatrix} 3 & 0 \\ 0 & 1 \end{vmatrix} \begin{vmatrix} 1/3 & 0 \\ 0 & 1 \end{vmatrix} = \begin{vmatrix} 0.9 & 0 \\ 0.1 & 0.8 \end{vmatrix} \quad (2.376)$$

yazılabilir ve $\|B\|_\infty = 0.9$ elde edilir.

Aşağıda verilen örneği göz önüne alalım:

$$A = \begin{vmatrix} 0.1 & 0 & 0.5 \\ 0.2 & 0.3 & 0 \\ 0.2 & 0.1 & 0.4 \end{vmatrix} \quad (2.377)$$

$\|A\|_{\infty} = 1.1$, $\|A\|_1 = 1.4$, $\|A\|_E = (1.35)^{1/2}$ elde edilir. B

$$B = \begin{vmatrix} 1/2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} A \begin{vmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = \begin{vmatrix} 0.1 & 0 & 0.5 \\ 0.4 & 0.3 & 0 \\ 0.4 & 0.1 & 0.4 \end{vmatrix} \quad (2.378)$$

gibi tanımlanırsa $\|B\|_{\infty} = 0.9$ olduğu görülür. Öyle ki A serisi yakınsaktır. Bu, *köşegen-matris* ve *benzerlik-dönüşümünün* hata analizinde büyük bir öneme sahiptir.

2.44 Sonsuz Matris Serilerinden Kaçınma

Genellikle, normların temel özelliklerini kullanarak matrislerin sonsuz serilerinden kullanımından kaçınılır. Örneğin, eğer $\|A\| < 1$ ise, $\|(I+A)^{-1}\|$ için bir sınır elde edebiliriz.

$\|A\| < 1$ olduğundan, A'nın tüm özdeğerleri sayısal olarak 1 den küçüktür. $(I+A)$, böylece tekil değildir ve $(I+A)^{-1}$ mevcuttur.

$$\begin{aligned} I &= (I+A)^{-1} \cdot (I+A) \\ &= (I+A)^{-1} + (I+A)^{-1} A \end{aligned} \quad (2.379)$$

$$\begin{aligned} \|I\| &> \|(I+A)^{-1}\| - \|(I+A)^{-1} A\| \\ &> \|(I+A)^{-1}\| - \|(I+A)^{-1}\| \cdot \|A\| \end{aligned} \quad (2.380)$$

$$\|(I+A)^{-1}\| < \|I\| / (1 - \|A\|) = (1 - \|A\|)^{-1} \quad \|I\| = 1 \quad (2.381)$$

Alternatif olarak, eğer $\|A\| < 1$ ise, yakınsaklığı garantilenmiş aşağıdaki ifade yazılabilir:

$$(I+A)^{-1} = I - A + A^2 - \dots \quad (2.382)$$

$$\begin{aligned} \|(I+A)^{-1}\| &< \|I\| + \|A\| + \|A\|^2 + \dots \\ &= (1 - \|A\|)^{-1} \quad (\|I\| = 1) \end{aligned} \quad (2.383)$$

3. PERTURBASYON TEORİSİ

3.1 Perturbasyon Teorisinin Temelleri

Bundan önceki bölümde cebirsel özdeğer özproblemlerinin, matematiksel temelleri üzerinde duruldu. Bu bölümde ise bilgisayar üzerinde, özdeğerler ve özvektörlerin hesaplanmasında ihtiyaç duyulan pratik uygulamalar ve hesaplanan değerlerin doğruluğu ve hassalığı ile ilgilenilecek ve problemin formülasyonundan ve çözümünden aktarılan çeşitli hataların etkileri incelenecektir.

Meydana gelen hataların oluşumunda üç ana temel vardır:

(i) Verilen matrisin elemanları doğrudan fiziksel ölçümlerden alınmıştır. Böylece tüm gözlemlerdeki hatalar probleme aktarılır. Verilen A matrisi, gerçek ölçümlere karşılık gelen matrise yaklaşıktır. Bu durumda eğer A matrisinin tüm elemanlarındaki hatalar δ gibi bir sayı ile sınırlandırılmış garantileniyorsa, gerçek matrisin $(A+E)$ olduğu söylenebilir. Burada E aşağıdaki şekilde ifade edilir:

$$|e_{ij}| < \delta \quad (3.1)$$

Pratik bir uygulamada çözüm sadece A'nın özdeğerlerinin değil, $(A+E)$ matris sınıfındaki tüm matrislerin özdeğerlerinin değişim aralığının incelenmesini de içermektedir. Böylece elemanların perturbasyonuna karşılık gelen bir matrisin özdeğerlerinin perturbasyonunun incelenmesi konusuna girmiş oluruz.

(ii) Matrisin elemanları doğrudan matematiksel formüller ile ifade edilebilir. Ancak gerçek matris bilgisayarda ifade edilirken bazı hatalardan kaçınamayız. Eğer matrisin elemanların tümü yada bir kısmı irrasyonel ise bu hata kaçınılmazdır. Bu durumda A matrisi, bilgisayarda

ifade edilebilen birden fazla matrisin toplamı halinde ifade edilebilir. Böylece (i) durumuna benzer bir problem ile karşı karşıya kalınır. Verilen matris $(A+E)$, E ise hata matrisi olacaktır.

(iii) Verilen matrisin bilgisayarda tam olarak ifade edildiğini kabul etsek bile, hesaplamalar ardından üretilen çözümlerin bilgisayarda tam olarak ifade edildiği söylenemez. Genelde, hesaplama yöntemleri $A_1, A_2, A_3, \dots$ ardışık benzerlik-dönüşümlerine ihtiyaç duyar. Bu hesaplamalar esnasında ortaya çıkan yuvarlatma hataları daima bir sonraki adıma taşınır. İlk bakışta, bu şekilde oluşan hatanın, (i) ve (ii) de bahsedilen hatalardan farklı olduğu düşünülebilir. Ancak bu genel olarak doğru değildir. Genellikle hesaplanan A_i matrisleri $(A+E_i)$ matrislerine benzer olarak ifade edilebilir. Burada E_i yuvarlatma hatalarının fonksiyonlarından oluşan küçük elemanlara sahip matristir. Genelde, E_i için kesin bir hata sınırı vardır. Böylece A_i -nin özdeğerleri $(A+E_i)$ matrisinin özdeğerleri olur. Bu ise yine (i) ve (ii) deki gibi elemanların perturbasyonundan oluşan bir matrisin özdeğerlerinin ve özvektörlerinin perturbasyonunun incelenmesi konusudur.

3.2 Özdeğerlerin Sürekliliği Üzerine Ostrowski Teoremi

Özdeğerler için perturbasyonun üst sınırı (1957) yılında Ostrowski tarafından verilmiştir. Ostrowski' nin sonuçları aşağıdaki gibi verilebilir:

A ve B aşağıdaki ilişkiyi sağlayan matrisler olsun:

$$|a_{ij}| < 1, |b_{ij}| < 1 \quad (3.2)$$

Eğer $\lambda \in (A+\epsilon B)$ -nin bir özdeğeri ise, aşağıdaki gibi A -nın bir λ özdeğeri vardır:

$$|\lambda \in -\lambda| < (n+2)(n^2\epsilon)^{1/n} \quad (3.3)$$

Ayrıca A -nın ve $(A+\epsilon B)$ -nin λ ve $\lambda \in$ özdeğerleri aşağıdaki ifade altında ifade edilebilirler:

$$|\lambda \in -\lambda| < 2(n+1)^2(n^2\epsilon)^{1/n} \quad (3.4)$$

Bunların teoride büyük bir öneme sahip olmasına karşın pratik uygulamalarda pek bir yararı yoktur. Aşağıdaki ilişkiyi sağlayan irrasyonel elemanlardan oluşmuş n . mertebeden bir A matrisin olduğunu kabul edelim:

$$|a_{ij}| < 1 \quad (3.5)$$

A matrisinin elemanlarının aşağıdaki gibi 10 haneli alana yuvarlatılması ile elde edilen $(A+\varepsilon B)$ matrisini göz önüne alalım:

$$|e_{ij}| < \frac{1}{2} 10^{-10} \quad (3.6)$$

Ostrowski' nin ikinci sonucundan aşağıdaki ifadeyi elde edilir:

$$n \prod_{i=1}^n (1 + \varepsilon_i) < 20 \quad (3.7)$$

Görüldüğü gibi bu sınır basit norm teorisinden elde ettiğimiz sınırdan daha zayıftır.

$$|\lambda| < \|A\|_{\infty} < 20 \quad (3.8)$$

$$n \prod_{i=1}^n (1 + \|e_i\|) < 20 + 10 \cdot 10^{-10}$$

$$n \prod_{i=1}^n (1 + \|e_i\|) < 40 + 10^{-9} \quad (3.9)$$

Sadece aşağıdaki koşul gerçekleştiği zaman, Ostrowski teoremi basit norm teorisinden daha iyi sonuçlar verir.

$$2(n+1)^2(n^2\varepsilon)^{1/n} < 2n \quad (3.10)$$

$$\varepsilon < \frac{n^{n-2}}{(n+1)^{2n}} - \frac{1}{e^2 n^{n+2}}$$

Bunun manası $n=20$ için, ε -nun 10^{-27} den daha küçük olması gerektiğidir. Ostrowski sonuçlarından yararlı bir sınır elde etmek için ε bundan çok daha küçük olmalıdır.

$\varepsilon^{1/n}$ çarpanı Ostrowski teoremindeki ana zayıf noktalarından biridir. n . dereceden, n özdeğeri a -ya eşit olan, *Jordan-altmatrisi* $C_n(a)$ örneğini inceleyelim. Eğer $(1,n)$ elemanı 0 dan ε -na çevrilirse karakteristik denklem aşağıdaki gibi olur:

$$(a-\lambda)^n + (-1)^{n-1} \varepsilon = 0 \quad (3.11)$$

$$\lambda = a + \omega^r \varepsilon^{1/n} \quad (r=0, \dots, n-1) \quad (3.12)$$

Burada ω birimin n . temel köküdür.

3.3 Cebirsel Fonksiyonlar

İzleyen incelemelerde, cebirsel fonksiyonlar teorisinin iki önemli sonucuna ihtiyaç duyulacaktır. Bunlar ispatsız olarak aşağıda gösterilecektir.

$f(x,y)$ aşağıdaki gibi ifade edilsin:

$$f(x,y) = y^n + p_{n-1}(x)y^{n-1} + p_{n-2}(x)y^{n-2} + \dots + p_1(x)y + p_0(x) \quad (3.13)$$

Burada $p_i(x)$, x in polinomlarıdır. x -in herhangi bir değerine karşılık gelen $f(x,y)=0$ $y_1(x), y_2(x), \dots, y_n(x)$ gibi n köke sahiptir. Buradaki her bir kök uygun katlılığı ile verilmektedir. $f(0,y)=0$ -ın kökleri $y_1(0), y_2(0), \dots, y_n(0)$ şeklinde ifade edilirler. İki teorem aşağıdaki gibidir:

TEOREM 1 : $y_i(0)$, $f(0,y)=0$ fonksiyonunun basit kökleri olsun. Bu durumda, öyle bir pozitif δ_i sayısı mevcuttur ki, aşağıdaki gibi ifade edilen $f(x,y)=0$ fonksiyonunun basit bir kökü $y_i(x)$ mevcuttur ve

$$y_i(x) = y_i(0) + p_{i1}x + p_{i2}x^2 + \dots \quad (3.14)$$

şeklinde seriye açılır.

Burada sağ tarafta bulunan seri, $|x| < \delta_i$ için yakınsaktır.

- Not: (i) Sadece $y_i(0)$, basit kök olduğu kabul edilecek, $f(0,y)=0$ -ın diğer kökleri hakkında hiçbir kabul olmayacaktır.
- (ii) (3.2) denkleminin sağ tarafındaki seriler sonlu olabilir.
- (iii) $x \rightarrow 0$ iken, $y_i(x) \rightarrow y_i(0)$ olur.

TEOREM 2 : Eğer $y_1(0)=y_2(0)=\dots=y_m(0)$, $f(0,y)$ fonksiyonunun m katlı kökleri ise, öyle bir δ pozitif sayısı vardır ki $|x|<\delta$, $f(x,y)=0$ -ın m kökü aşağıdaki özelliklere sahip olur:

Tüm m kök, $m_1, m_2, \dots, m_r$ -nin aşağıdaki koşul altında r grubunun içerisine düşer.

$$\sum m_i = m \quad (3.15)$$

ve m_i grubunun içindeki bu kökler

$$y_1(0) + p_{i1}z + p_{i2}z^2 + \dots \quad (3.16)$$

şeklindedir.

m_i ye karşılık aşağıdaki gibi tanımlanan z -nin değerleri farklıdır.

$$z = x^{m_i^{-1}} \quad (3.17)$$

z değeri tanımlanır.

- Not: (i) (3.16) serisi herhangi bir noktada kesilebilir.
(ii) $m_i=m$ durumunda $r=1$ olur ve m kökün hepsi aynı kesirli seri ile verilebilir.
(iii) $m_i=1$ halinde buna tekabül eden kuvvet serisi kesirli olmayabilir.
(iv) $x \rightarrow 0$ iken, tüm m kök için $y_1(0)$ yakınsar.

3.4 Nümerik Örnekler

(i) $f(x,y) = y^2(1+x) - 3y(1+x^2) + (2+x)$

$$f(0,y) = y^2 - 3y + 2$$

Böylece $y_1(0)=1$ basit bir köktür. $y_1(x)$ aşağıdaki gibi verilir:

$$y_1(x) = [3(1+x^2) - (1 - 12x + 14x^2 + 9x^4)^{1/2}] / 2(1+x)$$

Bu $12|x| + 14|x|^2 + 9|x|^4 < 1$ yakınsak kuvvet serisine açılabilir.

$$(ii) f(x,y)=y^3-y^2-x(1+x)^2$$

$$f(0,y) = y^3 - y^2$$

$$y^1(0) = 1 \text{ basit kök.}$$

$$y^2(0) = 0 \text{ çift kat kök.}$$

$$y_1(x) = 1+x$$

Çift kat kökün varlığı $y_1(x)$ in kesirli (rasyonel) kuvvetlere taşımadığına dikkat edilmesi gerekmektedir.

$y_2(0) = 0$ a karşılık gelen kökler aşağıdaki ifadenin çözümleridir

$$y^2 + yx + x(1+x) = 0$$

$$y = \frac{1}{2} [-x + (-4x - 3x^2)^{1/2}]$$

$$= \frac{1}{2} [-x + 2ix^{1/2}(1 + \frac{3}{4}x)^{1/2}]$$

Böylece kuvvet serisi sadece $x^{1/2}$ -nin tek kuvvetlerini kapsar. Eğer $|x| < 4/3$ ise yakınsaktır.

$$(iii) f(x,y)=y^4-y^2x(1+x)^2 + x^3(1+x)^2$$

$f(0,y) = 0$ denklemi $y = 0$ dört kat köke sahiptir.

$f(x,y) = 0$ kökleri ise aşağıdaki gibidir:

$$y = x^{1/2}(1+x)^{1/2}$$

$$y = x(1+x)^{1/2}$$

$$y = -x(1+x)^{1/2}$$

Teorem2 ' de $m_1 = 2, m_2 = 1, m_3 = 1$ elde edilir.

Üçüncü polinom dört kat sıfıra sahip olmasına rağmen, y -nin kuvvetlerinin katsayılarının perturbasyonları, köklerdeki perturbasyonların hiçbirinin $x^{1/2}$ mertebesinde değildir. Bu çeşit ilgili perturbasyon özdeğer teorisinde yaygındır.

3.5 Basit Özdeğerler için Perturbasyon Teorisi

(3.2) denklemini sağlayan A ve B gibi iki matris ele alalım. λ_1 , A -nın basit kökü olsun. $(A+\epsilon B)$ -nin özdeğerlerini inceleyelim. A -nın karakteristik denklemi aşağıdaki gibi olsun:

$$\det(\lambda I - A) = \lambda^n + c_{n-1}\lambda^{n-1} + c_{n-2}\lambda^{n-2} + \dots + c_0 = 0 \quad (3.18)$$

Bu durumda $(A+\epsilon B)$ -nin karakteristik denklemi de aşağıdaki gibi olur:

$$\det(\lambda I - A - \epsilon B) = \lambda^n + c_{n-1}(\epsilon)\lambda^{n-1} + c_{n-2}(\epsilon)\lambda^{n-2} + \dots + c_0(\epsilon) = 0 \quad (3.19)$$

Burada $c_r(\epsilon)$, ϵ -nın aşağıdaki gibi $(n-r)$ dereceden bir polinomudur.

$$c_r(0) = c_r \quad (3.20)$$

$$c_r(\epsilon) = c_r + c_{r1}\epsilon + c_{r2}\epsilon^2 + \dots + c_{r,n-r}\epsilon^{n-r} \quad (3.21)$$

λ_1 (5.1) in basit kökü olduğundan, Teorem1 'den, (5.2) nin $\lambda_1(\epsilon)$ basit kökü, aşağıda verilen yakınsak kuvvet serisi ile ifade edilebilir.

$$\lambda_1(\epsilon) = \lambda_1 + k_1\epsilon + k_2\epsilon^2 + \dots \quad (3.22)$$

Açık olarak $\epsilon \rightarrow 0$ iken $\lambda_1(\epsilon) \rightarrow \lambda_1$ olur. Aşağıdaki ifade diğer özdeğerlerden bağımsızdır.

$$|\lambda_1(\epsilon) - \lambda_1| = O(\epsilon) \quad (3.23)$$

3.6 Özvektörlerin Perturbasyonu

6. Özdeğerlerin perturbasyonuna dönersek, ilk olarak A -nın λ_1 basit özdeğerine karşılık gelen x_1 vektörünün bileşenleri için açık bir ifade geliştirelim. λ_1 basit özdeğer olduğundan, $(A - \lambda_1 I)$ en az bir tane $(n-1)$ mertebesinden sıfır-olmayan minöre sahiptir. Bunun, $(A - \lambda_1 I)$ matrisinin ilk $(n-1)$ satırı içerisine yayılmış olduğunu kabul edelim. Bu durumda x in bileşenleri aşağıdaki gibi alınabilir:

$$(A_{n1}, A_{n2}, \dots, A_{nn}) \quad (3.24)$$

Burada A_{ni} , $(A - \lambda_1 I)$ matrisinin (n, i) elemanlarının minörlerini ifade etmektedir. Ve böylece $(n-1)$ den daha büyük dereceye sahip olamayan bir λ_1 -in polinomudur.

$(A + \varepsilon B)$ -nin $\lambda_1(\varepsilon)$ basit özdeğerine bunu uygulayalım. A -nın özvektörü x_1 ile, $(A + \varepsilon B)$ -nin özvektörlerin de $x_1(\varepsilon)$ ile ifade edilsin. Açık olarak, $x_1(\varepsilon)$ elemanları $\lambda_1(\varepsilon)$ ve ε -nün polinomlarıdır. $\lambda_1(\varepsilon)$ için kuvvet serisi, yeterince küçük ε için yakınsak olduğundan $x_1(\varepsilon)$ -nün tüm elemanları ε -nün yakınsak kuvvet serisi ile ifade edilebilir. Aşağıdaki ifadeyi yazalım:

$$x_1(\varepsilon) = x_1 + \varepsilon z_1 + \varepsilon^2 z_2 + \dots \quad (3.25)$$

Sağ taraftaki vektör serisinin her bir bileşeni, ε -nün yakınsak kuvvet serisidir. Özdeğerler için (5.6) nın sonucuna uygun olarak, özvektör için de aşağıdaki sonuçlara ulaşılır ve yine ε -nün kesirli kuvvetleri yoktur.

$$|x_1(\varepsilon) - x_1| = O(\varepsilon) \quad (3.26)$$

3.7 Lineer Temel-Bölenlere sahip Matrisler

Eğer A matrisi lineer temel-bölenlere sahip ise sağ ve sol özdeğerlerinin kümesine sahip oluruz. Bunlarda sırasıyla $x_1, x_2, \dots, x_n$ ve $y_1, y_2, \dots, y_n$ vektörleridir ve

$$y_i^T x_j = 0 \quad (i \neq j) \quad (3.27)$$

şartını sağlar. Bunlar, bütün özdeğerleri basit ise tek türdürler. (3.25) nin z_i vektörleri, x vektörleri cinsinden yazılırsa

$$z_i = \sum_{j=1}^n s_{ji} x_j \quad (3.28)$$

olmak üzere

$$x_1(\varepsilon) = x_1 + \varepsilon \sum_{j=1}^n s_{j1} x_j + \varepsilon^2 \sum_{j=1}^n s_{j2} x_j + \dots$$

şeklinde olur. x_1 vektörünün terimleriyle toplayarak aşağıdaki ifadeye ulaşılır:

$$x_1(\varepsilon) = (1 + \varepsilon s_{11} + \varepsilon^2 s_{12} + \dots) x_1 + (\varepsilon s_{21} + \varepsilon^2 s_{22} + \dots) x_2 + \dots + (\varepsilon s_{n1} + \varepsilon^2 s_{n2} + \dots) x_n \quad (3.29)$$

Parantez içindeki n tane kuvvet serisinin yakınsaklığı, (2) deki serinin mutlak yakınsaklığından gelmektedir.

$x_1(\varepsilon)$ -daki çarpan farkı ile ilgilenmediğimize göre yeterince küçük ε için sıfır olmayan $(1+\varepsilon s_{11}+\varepsilon^2 s_{12}+\dots)$ ile bölebiliriz. Şimdi yeni vektör aşağıdaki gibidir:

$$x_1(\varepsilon) = x_1 + (\varepsilon t_{21} + \varepsilon^2 t_{22} + \dots)x_2 + \dots + (\varepsilon t_{n1} + \varepsilon^2 t_{n2} + \dots)x_n \quad (3.30)$$

Parantez içindeki ifadeler yeterince küçük ε için yakınsak kuvvet serisidir.

Bu ilişki, x in bileşenlerinin doğrudan determinantal ifadelerinden elde edildiği kabulü altında elde edildi. Bununla birlikte eğer x_i , aşağıdaki gibi normalize edilmiş yeni x_i ile yer değiştirilirse:

$$\|x_i\|_2 = 1 \quad (3.31)$$

parantez içindeki ifadelerin her birine bir sabit çarpan gelir ve bu t_{ij} içerisine aktarılır. (3.31) denklemini $x_i(\varepsilon)$ normalize değilse bile normalize edilmiş x_i -ler için doğrudur.

3.8 Özdeğerlerin Birinci Mertebeden Perturbasyonu

Şimdi x_i ve y_i terimleri cinsinden birinci dereceden perturbasyonlarının açık ifadelerini elde edelim. İlk olarak s_i sayısını tanıtalım. Bu daha sonra sıkça kullanılacaktır.

$$s_i = y_i^T x_i \quad (i=1,2,\dots,n) \quad (3.32)$$

Burada y_i ve x_i normalize edilmiş sağ ve sol vektörleridir.

Eğer y_i ve x_i reel ise s_i , bu vektörler arasındaki açının kosinüsüdür.

Eğer bazı katlı özdeğerler varsa karşılık gelen x_i ve y_i tek olarak belirlenemez. Bu durumda kullanılan s_i -lerin, x_i ve y_i -lerin bazı belirli seçimlerine karşılık geldiğini kabul edeceğiz. x_i ve y_i basit bir λ_i ye karşılık geldiği zaman bile, birim modülün keyfi bir kompleks bir çarpanı vardır. Fakat bu durumda $|s_i|$ tamamen belirlenmiş olmalıdır.

Herhangi bir durumda aşağıdaki ifade elde edilir:

$$|s_i| = |y_i^T x_i| < \|y_i\|_2 \|x_i\|_2 = 1 \quad (3.33)$$

β_{ij} aşağıdaki ilişki ile tanımlanan bir miktar olsun:

$$\beta_{ij} = y_i^T B x_j \quad (3.34)$$

Böylece $\|B\|_2 < n$ olduğundan, aşağıdaki ifade elde edilir:

$$|\beta_{ij}| = |y_i^T (B x_j)| < \|y_i\|_2 \|B x_j\|_2 < \|B\|_2 \|y_i\|_2 \|x_j\|_2 < n \quad (3.35)$$

Tanımlardan aşağıdaki ifade yazılabilir:

$$(A + \epsilon B) x_1(\epsilon) = \lambda_1(\epsilon) x_1(\epsilon) \quad (3.36)$$

$\lambda_1(\epsilon)$ ve $x_1(\epsilon)$ -nün bileşenleri, yakınsak kuvvet serisi olarak ifade edilebileceğinden, denklemin her iki tarafındaki ϵ -nün aynı kuvvete sahip terimleri eşitlenebilir. ϵ -daki terimlerin eşitlenmesi denklem (3.22) ve (3.30) den aşağıdaki ifadeyi verir:

$$A \left(\sum_{i=2}^n t_{i1} x_i \right) + B x_1 = \lambda_1 \left(\sum_{i=2}^n t_{i1} x_i \right) + k_1 x_1 \quad (3.37)$$

$$\sum_{i=2}^n (\lambda_i - \lambda_1) t_{i1} x_i + B x_1 = k_1 x_1 \quad (3.38)$$

y_1^T ile soldan çarparak ve bunun dik olduğu hatırlanarak, $y_1^T x_i = 0$ ($i \neq 1$) den aşağıdaki ifade elde edilir:

$$k_1 = y_1^T B x_1 / y_1^T x_1 = \beta_{11} / s_1 \quad (3.39)$$

$$|k_1| < n / |s_1| \quad (8.4) \text{ den} \quad (3.40)$$

Yeterince küçük ϵ için, λ_1 in perturbasyonundaki ana terim, $k_1 \epsilon$ olur. Böylece bu özdeğerin hassaslığının s_1 e bağlı olduğu görülür. Maalesef s_1 keyfî olarak küçük olabilir.

3.9 Özvektörlerin Birinci Mertebeden Perturbasyonu

(3.38) ün y_i^T ile soldan çarpımı aşağıdaki ifadeyi verir:

$$(\lambda_i - \lambda_1)t_{i1}s_i + \beta_{i1} = 0 \quad (i = 2, 3, \dots, n) \quad (3.41)$$

Denklem (3.30) den x_1 in perturbasyonundaki birinci dereceden terimleri, aşağıdaki gibi verilir:

$$\varepsilon \left[\frac{\beta_{21}x_2}{(\lambda_1 - \lambda_2)s_2} + \frac{\beta_{31}x_3}{(\lambda_1 - \lambda_3)s_3} + \dots + \frac{\beta_{n1}x_n}{(\lambda_1 - \lambda_n)s_n} \right] \quad (3.42)$$

Payda $(\lambda_1 - \lambda_i)$ çarpanlarına sahip olursa, 26. bölümde gösterileceği üzere, s_1 in haricindeki s_i -lerin varlığı bizi biraz yanıltabilir. Eğer basit özdeğer λ_1 . diğer özdeğerlerden bazılarına yakınsa, λ_1 özdeğerine karşılık gelen özvektör A -nın perturbasyonuna oldukça hassastır. Bu gerçektende böyledir.

λ_i diğer özdeğerlerden iyice uzak ise ve s_i ($i=2, 3, \dots, n$) -lerin hiçbirisi küçük değil ise, bu durumda, x_1 özvektörü A -nın perturbasyonuna oldukça kayıtsızdır.

3.10 Yüksek Mertebeden Perturbasyonlar

(9.1) denklemindeki ε -nın yüksek mertebeden kuvvetlerinin katsayılarının eşitlenmesiyle, yüksek mertebeden perturbasyonların ifadeleri elde edilebilir. ε^2 -nin katsayılarının eşitlenmesiyle aşağıdaki ifade elde edilir:

$$A \left(\sum_{i=2}^n t_{i2}x_i \right) + B \left(\sum_{i=2}^n t_{i1}x_i \right) = k_2x_1 + k_1 \left(\sum_{i=2}^n t_{i1}x_i \right) + \lambda_i \left(\sum_{i=2}^n t_{i2}x_i \right) \quad (3.43)$$

veya

$$\sum_{i=2}^n t_{i2}(\lambda_i - \lambda_1)x_i + B \left(\sum_{i=2}^n t_{i1}x_i \right) = k_2x_1 + k_1 \left(\sum_{i=2}^n t_{i1}x_i \right)$$

y_i^T ile soldan çarpımı aşağıdaki ifadeyi verir:

$$\sum_{i=2}^n t_{1i} \beta_{1i} = k_2 s_1 \quad (3.44)$$

(10.1) den

$$k_2 = \frac{1}{s_1} \sum_{i=2}^n \frac{\beta_{1i} \beta_{1i}}{(\lambda_1 - \lambda_i)} \quad (3.45)$$

aynı şekilde, x_1 in perturbasyonundaki ikinci dereceden terimler elde edilebilir. Fakat bu yüksek dereceden terimlerin çok küçük pratik değeri vardır.

3.11 Katlı Özdeğerler

m katlı λ_1 özdeğerinin perturbasyonuna dönersek, §5 e benzer bir inceleme, §3ün Teorem2' nin kullanımına ihtiyaç duymaktadır. Bu, ε -nun kesirli kuvvet serisinin terimlerine açılabilen bir grup içerisine düşen m özdeğer kümesinin var olacağını gösterir. Eğer sadece bir grup var ise bu durumda özdeğer, bir $\varepsilon^{1/m}$ kuvvet serisi ile ifade edilebilecektir. Fakat eğer birden fazla grup varsa, diğer kesirli kuvvetlere ihtiyaç duyulacaktır. Bazı özel durumlarda, hiçbir kesirli kuvvete ihtiyaç duymayan m grup olabilir ve perturbe edilmiş m özdeğer, $\lambda_i(\varepsilon)$ -nun her biri için aşağıdaki gibi ifade edilir:

$$|\lambda_i(\varepsilon) - \lambda_1| = O(\varepsilon) \quad (3.46)$$

Katlı sıfırlar için karakteristik polinomlar cinsinden analiz yapmak oldukça karışıktır. Çünkü esas matristen uzaklaşmaktadır. Bundan sonraki birkaç bölümde önemli pratik değeri olan daha güçlü bir metot ile pratik analizini yapacağız.

3.12 Gerschgorin Teoremleri

Bu alternatif yaklaşım için Gerschgorin' nin (1931) 2 teoremine ihtiyaç duyulur.

Teorem 3: A matrisinin her bir özdeğeri, merkezi a_{ii} ve yarı çapı $\sum |a_{ij}|$ ($j \neq i$) olan dairesel bir diskin içindedir.

İspat oldukça basittir. λ , A -nın herhangi bir özdeğeri olsun. Bu durumda, aşağıdaki gibi en az bir tane sıfır-olmayan x vardır

$$Ax = \lambda x$$

(3.47)

x vektörünün r . bileşeni en büyük modüle sahip olduğu kabul edilsin. Bu durumda x aşağıdaki gibi normalize edilebilir:

$$x^T = (x_1, x_2, \dots, x_{r-1}, 1, x_{r+1}, \dots, x_n) \quad (3.48)$$

$$|x_i| < 1 \quad (i \neq r) \quad (3.49)$$

(13.1) in her iki tarafındaki r . elemanları eşitleyerek aşağıdaki ifadeye ulaşılır:

$$\sum_{j=1}^n a_{rj} x_j = \lambda x_r = \lambda \quad (3.50)$$

$$|\lambda - a_{rr}| < \sum_{j \neq r} |a_{rj} x_j| < \sum_{j \neq r} |a_{rj}| |x_j| < \sum_{j \neq r} |a_{sj}|$$

ve λ dairesel bir diskin içerisinde bulunur.

İkinci teorem, özdeğerlerin disk boyunca yayılımları hakkında daha detaylı bilgi verir.

Teorem 4 : Eğer Teorem3' ün dairesel çemberinin s tanesi, birbirinden izole edilmiş bağlantılı bir bölge ise, bu durumda bu bölge içerisinde, A -nın s tane özdeğeri vardır.

İspat süreklilik notasyonuna bağlıdır.

$$A = \text{diag}(a_{ii}) + C = D + C \quad (3.51)$$

Burada C matrisi, köşegen dışındaki elemanlardan oluşur. r_i ifadesini aşağıdaki gibi tanımlayalım.

$$r_i = \sum_{j \neq i} |a_{ij}| \quad (3.52)$$

Şimdi

$$0 < \varepsilon < 1 \quad (3.53)$$

şartını sağlayan ε değeri için $(D+\varepsilon C)$ matrisini göz önüne alalım:

$\varepsilon = 0$ için, bu matris D dir. $\varepsilon = 1$ için ise A matrisidir. $(D+\varepsilon C)$ -nin karakteristik polinomunun katsayıları, ε -nun polinomlarıdır ve karakteristik denklemin kökleri, cebirsel fonksiyonlar teorisinden bilindiği gibi, ε -nun *sürekli-fonksiyon*dur. Teorem3' ten, özdeğerlerin hepsi, ε -nun herhangi bir değeri için merkezi a_{ii} , çapı εr_i olan dairesel bir disk içerisine yayılır. Eğer, ε düzgün bir biçimde 0 -dan 1 -e kaydırılırsa özdeğerlerin tümü sürekli bir eğri çizerler.

Genelliği bozmadan, bunun bağlantılı bölgeyi oluşturan ilk s disk olduğunu kabul edebiliriz. Bu durumda, çapı $r_{s+1}, r_{s+2}, \dots, r_n$ olan $(n-s)$ diskleri, çapı $r_1, r_2, \dots, r_s$ olan disklerden izole edilmiş olduğundan, (3.53) aralığındaki tüm ε -lar için, çapı εr_i olan diskler içinde doğrudur. Şimdi $\varepsilon=0$ olduğu zaman, özdeğerler, $a_{11}, a_{22}, \dots, a_{nn}$ olur ve bunların ilk s tanesi ilk s diskin karşılık gelen bölgenin içerisine yayılmıştır. Geri kalan $(n-s)$ bu bölgenin dışında kalır. Bu ε -nun $\varepsilon = 1$ değerine kadar olan tüm değerleri için geçerlidir.

Özel olarak Gerschgorin disklerinden herhangi biri izole edilmiş ise bu disk bir tane özdeğer ihtiva eder. A yerine A^T de içinde bu teori doğrudur.

3.13 Gerschgorin Teoremleri üzerine Kurulu Perturbasyon Teorisi

Şimdi $(A+\varepsilon B)$ matrisini özdeğerlerini, Jordan kanonik formunu kullanarak elde edelim. Burada beş ana durumu tartışacağız.

Durum1. *Lineer temel-bölenler sahip bir matrisin λ_1 basit özdeğerinin perturbasyonu:*

Bu durumda Jordan kanonik formu köşegen form olur. Sütunları, sağ tam özvektör kümesi x_i vektörlerine paralel olan aşağıdaki gibi H matrisi vardır:

$$H^{-1}AH = \text{diag}(\lambda_i) \quad (3.54)$$

Ayrıca burada H^{-1} -nin satırları, y_i^T sol tam özvektörler kümesine paraleldir. Eğer bu vektörler aşağıdaki gibi normalize edilirse H matrisinin i . kolonu x_i ve H^{-1} in i . satırı §8 notasyonundaki y_i^T/s_i olacak şekilde alınabilir. (λ_1 basit olduğu için x_1 ve y_1 birim modül ile tek türdür.)

Burada ifade edilen H ile aşağıdaki ifade derhal görülür:

$$H^{-1}(A+\varepsilon B)H = \text{diag}(\lambda_i) + \varepsilon \begin{vmatrix} \beta_{11}/s_1 & \beta_{12}/s_1 & \dots & \beta_{1n}/s_1 \\ \beta_{21}/s_2 & \beta_{22}/s_2 & \dots & \beta_{2n}/s_2 \\ \dots & \dots & \dots & \dots \\ \beta_{n1}/s_n & \beta_{n2}/s_n & \dots & \beta_{nn}/s_n \end{vmatrix} \quad (3.55)$$

Teorem3 uygulanarak, özdeğerlerin $(\lambda_i + \varepsilon \beta_{ii}/s_i)$ merkezli ve $\varepsilon \sum |\beta_{ij}/s_j|$ ($j \neq i$) çaplı dairesel bir diskin içerisinde bulunduğu gösterilir. (8.4) ilişkisinden, eğer $|b_i| < 1$ ise, i. diskin çapının $n(n-1).\varepsilon/|s_i|$ den daha küçük olduğu görülür. Ayrıca λ_1 basit özdeğer olduğundan, yeterince küçük ε için ilk diskin izole edildiği ve böylece kesin olarak bir tane özdeğer kapsadığı görülür.

Biraz önce elde ettiğimiz sonuç, bizi küçük bir hayal kırıklığına uğrattır. Çünkü, önceki analizden, λ_1 e karşılık gelen özdeğerin merkezi $(\lambda_1 + \varepsilon \beta_{11}/s_1)$ olan dairesel bir diskin içinde olması beklenirdi. Fakat bunun yerine $\varepsilon \rightarrow 0$ limitinde $O(\varepsilon^2)$ çapa sahip bir çember içerisinde konumlandı. Gerschgorin dairesel diskin çapı izleyen yöntem kullanılarak düşürülebilir.

Eğer herhangi bir matrisin i. kolonunu m ile, ve i. satırını $1/m$ ile çarpılırsa, özdeğerler değişmez. Bunu $i=1$ için $m=k/\varepsilon$ alarak, (14.3) deki sağ taraf matrisine uygulayalım.

$$\text{diag}(\lambda_i) + \begin{vmatrix} \varepsilon \beta_{11}/s_1 & \varepsilon^2 \beta_{12}/ks_1 & \varepsilon^2 \beta_{13}/ks_1 & \dots & \varepsilon^2 \beta_{1n}/ks_1 \\ k \beta_{21}/s_2 & \varepsilon \beta_{22}/s_2 & \varepsilon \beta_{23}/s_2 & \dots & \varepsilon \beta_{2n}/s_2 \\ \dots & \dots & \dots & \dots & \dots \\ k \beta_{n1}/s_n & \varepsilon \beta_{n2}/s_n & \varepsilon \beta_{n3}/s_n & \dots & \varepsilon \beta_{nn}/s_n \end{vmatrix} \quad (3.56)$$

(1,1) in dışındaki ilk kolonun elemanları ε dan bağımsız iken (1,1) in dışındaki ilk satırın elemanları, ε^2 çarpanına sahiptir. Tüm diğer elemanlar değişmemiştir. k öyle seçilsin ki, diğer diskler ilk diskin üzerini kaplamaması için yeterince küçük olurken, Gershgorin diski de mümkün olduğunca küçük olsun.

Eğer k aşağıdaki eşitsizliği sağlayan en büyük değere sahip olacak şekilde seçilirse, yeterince küçük ε için bu doğrudur:

$$(3.57)$$

$$|k\beta_{i1}/s_i| < \frac{1}{2} |\lambda_1 - \lambda_i| \quad (i=2,3,\dots,n)$$

(Yukarıda ki denklemden bulunan $\frac{1}{2}$ çarpanı herhangi bir özelliğe sahip değildir. Bu ϵ dan bağımsız ve 1 den küçük herhangi bir sayıda olabilir.)

$$k = \min |(\lambda_1 - \lambda_i)s_i/2\beta_{i1}| \quad (i=2,3,\dots,n) \quad (3.58)$$

(Eğer tüm β_{i1} sıfırsa, $\lambda_1 + \epsilon \cdot \beta_{11}/s_1$ gerçek bir özdeğerdir) k nın bu tanımı ile ilk Gerschgorin diskinin r_1 çapı için aşağıdaki ifadeye ulaşılır:

$$r_1 = \sum_{j=2}^n |\epsilon^2 \beta_{1j}/ks_1| < n(n-1)\epsilon^2/k|s_1| \quad (3.59)$$

$$k^{-1} = \max |2\beta_{i1}/(\lambda_1 - \lambda_i)s_i| < \max |2n/(\lambda_1 - \lambda_i)s_i| \quad (3.60)$$

Burada sınırların $|b_{ij}| < 1$ olacak şekilde, B-nin normalize edilmiş olduğu düşünülecektir.

Sınırlar için gerekli olan ifadeleri nümerik bir örnek ile inceleyelim.

X matrisi aşağıdaki gibi tanımlansın:

$$X = \begin{vmatrix} 0.9 & & \\ & 0.4 & \\ & & 0.4 \end{vmatrix} + 10^{-5} \begin{vmatrix} 0.1234 & 0.4132 & -0.2167 \\ -0.1342 & 0.4631 & 0.1276 \\ 0.1557 & 0.1432 & 0.3125 \end{vmatrix} \quad (3.61)$$

Sağdaki ilk matris çiftkat özdeğere ve *linear-bölenlere* sahiptir.

İlk satır 10^{-9} ile ve ilk kolon 10^{-5} ile çarpılarak X in özdeğerleri aşağıdaki ifadenin özdeğerleri olur.

$$\begin{vmatrix} 0.9 & & \\ & 0.4 & \\ & & 0.4 \end{vmatrix} + \begin{vmatrix} 10^{-5}(0.1234) & 10^{-10}(0.4132) & 10^{-10}(-0.2167) \\ -0.1342 & 10^{-5}(0.4631) & 10^{-5}(0.1276) \\ 0.1567 & 10^{-5}(0.1432) & (0.3125) \end{vmatrix} \quad (3.62)$$

Gershgorin diskleri aşağıdaki özelliklere sahip olur:

$$\begin{array}{ccc}
 \text{merkez} & \text{çap} & \\
 0.9 + 10^{-5}(0,1234) & 10^{-10}(0,4132 + 0,2167) & = 10^{-10}(0,6299) \\
 0.4 + 10^{-5}(0,4631) & 0,1342 + 10^{-5}(0,1276) & \\
 0,4 + 10^{-5}(0,3125) & 0,1567 + 10^{-5}(0,1432) &
 \end{array}$$

İlk disk izole edilmiştir ve sadece bir özdeğere sahiptir. Bu diskin çapı 10^{-10} mertebesinde dir. Diğer iki köşegen elemanlarının eşit olması bu sonucu etkilemez.

Bununla birlikte eğer diğer iki özdeğ erden biri ilkine yaklaşıtırlırsa bu özdeğ erlerin konumları, aleyhte bir durum olarak etkilenir. Aşağıdaki gibi tanımlanan Y matrisini göz önüne alalım:

$$Y = \begin{vmatrix} 0.9 & & \\ & 0.89 & \\ & & 0.4 \end{vmatrix} 10^{-5} \begin{vmatrix} 0.1234 & 0.4132 & -0.2167 \\ -0.1342 & 0.4631 & 0.1276 \\ 0.1567 & 0.1432 & 0.3125 \end{vmatrix} \quad (3.63)$$

İlk iki Gershgorin diski üzerlerini kaplamış olabileceğ inden 10^{-5} çarpanı daha fazla kullanılamaz. 10^{-3} çarpanını kullanarak, Y matrisi aşağıdaki ifadeye dönüşür:

$$\begin{vmatrix} 0.9 & & \\ & 0.89 & \\ & & 0.4 \end{vmatrix} + \begin{vmatrix} 10^{-5}(0.1234) & 10^{-8}(0.4132) & 10^{-8}(-0.2167) \\ 10^{-2}(-0.1342) & 10^{-5}(0.4631) & 10^{-5}(0.1276) \\ 10^{-2}(0.1567) & 10^{-5}(0.1432) & 10^{-5}(0.3125) \end{vmatrix} \quad (3.64)$$

ve ilk Gershgorin diski izole edilmiş olur. Böylece merkezi $0.9+10^{-5}$ ve çapı 10^{-8} (0.6299) olan diskin içerisinde bir tek özdeğ er vardır.

Durum 2: lineer temel-bölenler sahip bir matrisin katlı λ_1 özdeğ erlerinin perturbasyonu

Bu durum $\lambda_1 = \lambda_2 = \lambda_3$ ve $\lambda_4 = \lambda_5$ olan 6. mertebeden bir matrisin basit bir örneğ i ile yeterince iyi gösterilebilir. A matrisi lineer bölenlere sahip olduğ undan, $(A+\epsilon B)$ matrisi aşağıdaki matrise benzerdir:

$$\begin{vmatrix} \lambda_1 & & & & & \\ & \lambda_1 & & & & \\ & & \lambda_1 & & & \\ & & & \lambda_4 & & \\ & & & & \lambda_4 & \\ & & & & & \lambda_6 \end{vmatrix} + \varepsilon \begin{vmatrix} \beta_{11}/s_1 & \beta_{12}/s_1 & \dots & \beta_{16}/s_1 \\ \beta_{21}/s_2 & \beta_{22}/s_2 & \dots & \beta_{26}/s_2 \\ \dots & \dots & \dots & \dots \\ \beta_{61}/s_6 & \beta_{62}/s_6 & \dots & \beta_{66}/s_6 \end{vmatrix} \quad (3.65)$$

λ_6 nın benzeri 1. durumda incelendi. 6. kolon k/ε ile 6.satır ε/k ile çarpılarak özdeğeri, merkezi $\lambda_6 + \varepsilon \cdot \beta_{66} / s_6$ ve $\varepsilon \rightarrow 0$ limitinde çapı $O(\varepsilon^2)$ olan bir diskin içerisine konumlandırılır. Şimdi diğer beş Gershgorin diski ile ilgilenelim. Merkezi $\lambda_i + \varepsilon \beta_{ii} / s_i$ ($i=1,2,3$) olan üç ve yine merkezi $\lambda_4 + \varepsilon \beta_{ii} / s_i$ ($i=4,5$) iki tane disk vardır. Karşılık gelen çapların hepsi ε mertebesindedir. Yeterince küçük ε için üçlü grup diğer ikili gruptan izole edilecektir. Fakat genel olarak gruptaki özel diskler izole edilemezler. Üç katlı λ_1 özdeğerindeki perturbasyonlar, $\varepsilon \rightarrow 0$ iken ε mertebesindedirler. Fakat merkezi $\lambda_1 + \varepsilon \beta_{ii} / s_i$ ($i=1,2,3$) ve çapı ε^2 mertebesinden olan üç diskin her birinde bir özdeğerin var olduğu genelde doğru değildir.

Bununla birlikte bu dairesel üç diskin çapını küçültebiliriz. İlgimizi üç katlı özdeğer üzerine çevirelim ve (3.65) i aşağıdaki gibi ifade edelim:

$$\text{diag}(\lambda_i) + \varepsilon \begin{vmatrix} P & Q \\ R & S \end{vmatrix} \quad (3.66)$$

Burada P,Q,R,S (3x3) mertebesinden matrislerdir. İlk üç satırı ε/k ve ilk üç kolonu k/ε ile çarparak aşağıdaki ifade elde edilir:

$$\text{diag}(\lambda_i) + \begin{vmatrix} \varepsilon P & k^{-1} \varepsilon^2 Q \\ kR & \varepsilon S \end{vmatrix} \quad (3.67)$$

k -yı ε dan bağımsız olarak öyle seçilmelidir ki, ilk üç disk diğerlerinden izole edilmiş olsun. Böylece bazı uygun k değerleri için üç diskin bileşiminde üç özdeğer vardır.

merkez	çap
$\lambda_1 + \varepsilon \beta_{11}/s_1$	$\varepsilon(\beta_{12} + \beta_{13})/ s_1 + \varepsilon^2(\beta_{14} + \beta_{15} + \beta_{16})/k s_1 $
$\lambda_1 + \varepsilon \beta_{22}/s_2$	$\varepsilon(\beta_{21} + \beta_{23})/ s_1 + \varepsilon^2(\beta_{24} + \beta_{25} + \beta_{26})/k s_2 $
$\lambda_1 + \varepsilon \beta_{33}/s_3$	$\varepsilon(\beta_{31} + \beta_{32})/ s_2 + \varepsilon^2(\beta_{34} + \beta_{35} + \beta_{36})/k s_3 $

6. merteben bir matris için çaplardaki bu düşürme yeterli değildir. Fakat daha yüksek dereceden matrisler için daha gerçekçidir. Daha sonra göreceğiz ki (3.66) deki P matrisinin kanonik forma indirgenmesi ile daha kesin sonuçlar elde edeceğiz.

A matrisi katlı özdeğere ve lineer bölönlere sahip iken, karakteristik polinomun katsayılarının hiçbirinde ε -nun kesirli kuvvetleri görölmemektedir.

Durum 3. Bir veya daha fazla lineer-olmayan temel-bölönlere sahip bir matrisin basit bir özdeğerinin perturbasyonu:

A matrisi lineer-olmayan temel-bölönlere sahip olduğundan artık tam özdeğerler kümesine sahip değiliz. İlk olarak, herhangi bir matris için s_i değeri incelenecektir. Aşağıdaki gibi verilen basit bir A matrisini göz önüne alalım:

$$A = \begin{pmatrix} a & 1 \\ 0 & b \end{pmatrix} \quad (3.68)$$

Özdeğerler $\lambda_1 = a$ ve $\lambda_2 = b$ dir ve karşılık gelen sağ ve sol vektörleri aşağıdaki gibidir:

$$\begin{aligned} x_1^T &= (1, 0) & \alpha x_2^T &= (1, b-a) \\ \alpha y_1^T &= (a-b, 1) & y_2^T &= (0, 1) \end{aligned} \quad (3.69)$$

Burada $\alpha = [1 + (a-b)^2]^{1/2}$ eşitliğini sağlar.

$$s_1 = y_1^T x_1 = (a-b)/\alpha \quad s_2 = y_2^T x_2 = (b-a)/\alpha \quad (3.70)$$

elde ederiz ve $b \rightarrow a$ limitinde s_1 ve s_2 sıfıra gider. Biz görürüz ki basit bir özdeğerin perturbasyonunun hassaslığı s_i^{-1} ile orantılıdır ve böylece $b \rightarrow a$ limitinde her iki özdeğerin de çok daha fazla hassaslaştığı görölür. Bununla birlikte b -nin herhangi bir değeri için $s_1^{-1} + s_2^{-1} = 0$ olur. Böylece s_1^{-1} ve s_2^{-1} sayılarının her ikisi de sonsuza giderken bağımsız değildirlir.

Lineer-olmayan temel-bölönlere sahip matrislerin analizi, A nın Jordan kanonik formunun kullanıldığı önceki bölömlere benzerdir. Aşağıdaki gibi *tekil-olmayan* H matrisi olsun:

$$H^{-1}AH = C \quad (3.71)$$

Burada C üst-Jordan kanonik formdur. Burada H matrisinin tüm kolonlarının normalize edildiği düşünülemez. Çünkü bu kolonların skaler çarpanları C de süper-diagonal elemanlarının (8.1) birim olması gereksinmesi ile sınırlanmaktadır. Bununla birlikte, lineer-bölenler ile bağlantılı kolonların normalize edilmiş olduğu kabul edilsin. H^{-1} in normalize edilmiş satırlarından oluşmuş matris G ile ifade edilsin, ve aşağıdaki ifadeleri yazalım:

$$\begin{aligned} g_i^T h_i &= s_i & \text{eğer } g_i \text{ ve } h_i \text{ bir lineer-bölene karşılık geliyorsa.} \\ g_i^T h_i &= t_i & \text{eğer } g_i \text{ ve } h_i \text{ bir lineer-olmayan bölenlere karşılık gelen sistemin bir parçası ise.} \end{aligned}$$

Basit özdeğerin perturbasyonunu basit bir örnek ile inceleyelim. $(\lambda_1 - \lambda)^2$, $(\lambda_3 - \lambda)$, $(\lambda_4 - \lambda)$ temel-bölenlere sahip dördüncü dereceden bir matris göz önüne alalım.

$$H^{-1}(A + \varepsilon B)H = \begin{vmatrix} \lambda_1 & 1 & & \\ 0 & \lambda_1 & & \\ & & \lambda_3 & \\ & & & \lambda_4 \end{vmatrix} + \varepsilon \begin{vmatrix} \beta_{11}/t_1 & \beta_{12}/t_1 & \beta_{13}/t_1 & \beta_{14}/t_1 \\ \beta_{21}/t_2 & \beta_{22}/t_2 & \beta_{23}/t_2 & \beta_{24}/t_2 \\ \beta_{31}/s_3 & \beta_{32}/s_3 & \beta_{33}/s_3 & \beta_{34}/s_3 \\ \beta_{41}/s_4 & \beta_{42}/s_4 & \beta_{43}/s_4 & \beta_{44}/s_4 \end{vmatrix} \quad (3.72)$$

yazabiliriz. Burada

$$\begin{aligned} \beta_{ij} &= g_i^T B h_j & (\text{tüm } i, j \text{ için}) \\ |\beta_{ij}| &< n \end{aligned} \quad (3.73)$$

λ_3 ve λ_4 basit özdeğerlerinin perturbasyonlarının aşağıdaki gibi olup olmadığını araştıralım.

$$\begin{aligned} |\lambda_3(\varepsilon) - \lambda_3| &= O(\varepsilon) \\ |\lambda_4(\varepsilon) - \lambda_4| &= O(\varepsilon) \\ \varepsilon &\rightarrow 0 \end{aligned} \quad (3.74)$$

İlk bakışta, tüm ε -lar için Jordan formundaki birim elemanların varlığı bir dezavantaj gibi görülebilir. Eğer $|\lambda_1 - \lambda_3| < 1$ ise ε -nın herhangi bir değeri için üçüncü disk ilkinden izole edilmiş olmayacaktır. Bununla birlikte eğer (3.72) nın ilk satırını m ve ilk kolonu $1/m$ ile çarpılırsa, aşağıdaki şekle dönüşür:

$$\begin{vmatrix} \lambda_1 & m & & \\ 0 & \lambda_1 & & \\ & & \lambda_3 & \\ & & & \lambda_4 \end{vmatrix} + \varepsilon \begin{vmatrix} \gamma_{11} & \gamma_{12} & \gamma_{13} & \gamma_{14} \\ \gamma_{21} & \gamma_{22} & \gamma_{23} & \gamma_{24} \\ \gamma_{31} & \gamma_{32} & \gamma_{33} & \gamma_{34} \\ \gamma_{41} & \gamma_{42} & \gamma_{43} & \gamma_{44} \end{vmatrix} \quad (3.75)$$

Burada γ_{ij} , ε dan bağımsızdır. Aşağıdaki seçim ifadede ki zorluğu ortadan kaldırır.

$$m = \min \{ \frac{1}{4} |\lambda_1 - \lambda_3|, \frac{1}{4} |\lambda_1 - \lambda_4| \} \quad (3.76)$$

Çapı ε^2 mertebesine getirilerek üçüncü Gershgorin diskini izole etmek için, üçüncü kolonu ile üçüncü satır k/ε ve ε/k ile çarpılır. Bu arada k aşağıdaki gibi seçilir:

$$|k\gamma_{i3}| < \frac{1}{2} |\lambda_3 - \lambda_i| \quad (i \neq 3) \quad (3.77)$$

§15 de olduğu gibi merkezi $\lambda_3 + \varepsilon\gamma_{33}$ ve çapı $\varepsilon^2(|\gamma_{31}| + |\gamma_{32}| + |\gamma_{34}|)$ olan disk, yeterince küçük ε değeri için izole edilir. Böylece sadece bir tek özdeğer vardır. λ_3 basit özdeğerin perturbasyonu aşağıdaki gibi olur:

$$\begin{aligned} \varepsilon \rightarrow 0 \text{ iken} \\ \lambda_3(\varepsilon) - \lambda_3 = \varepsilon\gamma_{33} + O(\varepsilon^2) \end{aligned} \quad (3.78)$$

Şimdi γ_{33} , β_{33}/s_3 olur ve (3.72) den $|\beta_{33}| < n$ bulunur. $(\lambda_1 - \lambda)^2$ bölenlerinin var olması, λ_1 den farklı basit özdeğerin davranışında bir değişiklik yapmaz.

4. Durum: *Derogatory-olmayan bir matrisin lineer-olmayan temel-bölenlerine karşılık gelen özdeğerlerin perturbasyonu:*

Lineer-olmayan bölenler $(\lambda_1 - \lambda)^r$ ile ifade edilsin. Matris derogatory olmadığından $(\lambda_1 - \lambda)$ -ya ihtiyaç duyan başka bir temel-bölen yoktur. Şimdi $\lambda = \lambda_1$ çift kat özdeğerinin perturbasyonunu inceleyelim. (3.72) deki birim eleman, λ_1 elemanı tarafından işgal edilen satırlardan birindedir. Eğer ε ile sıfıra giden bir çapı olan bir Gershgorin diski elde edilecek ise, bazı çarpım işlemi bu eleman üzerinde uygulanmalıdır. 2. kolonu $\varepsilon^{1/2}$ ile 2. satırı $\varepsilon^{1/2}$ ile çarparak, matris toplamı aşağıdaki şekle dönüşür:

$$\begin{vmatrix} \lambda_1 & \varepsilon^{1/2} & & \\ 0 & \lambda_1 & & \\ & & \lambda_3 & \\ & & & \lambda_4 \end{vmatrix} + \begin{vmatrix} \varepsilon & \varepsilon & \varepsilon^{1/2} & \varepsilon \\ \varepsilon & \varepsilon & \varepsilon^{1/2} & \varepsilon^{1/2} \\ \varepsilon & \varepsilon^{1/2} & \varepsilon & \varepsilon \\ \varepsilon & \varepsilon^{1/2} & \varepsilon & \varepsilon \end{vmatrix} \quad (3.79)$$

Burada ikinci matrisin elemanlarında bulunan sabitler yazılmamıştır. $\varepsilon^{1/2}$ deki terimlerin Gerschgorin diskinin çapına baskın olduğu açıktır. Yeterince küçük ε için, mertebesi $\varepsilon^{1/2}$ olan bir çapa sahip λ_1 merkezli bir çemberde iki özdeğer vardır. Benzer olarak bir kübik bölen için matris formları (3.81) formlarından birine 2. ve 3. kolonları $\varepsilon^{1/2}$ ve $\varepsilon^{1/3}$ ile ve ikinci ve üçüncü satırları $\varepsilon^{-1/2}$ ve $\varepsilon^{-1/3}$ ile çarparak dönüştürülür.

$$\begin{vmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 1 \\ 0 & 0 & \lambda_1 \end{vmatrix} + \begin{vmatrix} \varepsilon & \varepsilon & \varepsilon & \varepsilon \\ \varepsilon & \varepsilon & \varepsilon & \varepsilon \\ \varepsilon & \varepsilon & \varepsilon & \varepsilon \\ \varepsilon & \varepsilon & \varepsilon & \varepsilon \end{vmatrix} \quad (3.80)$$

$$\begin{vmatrix} \lambda_1 & \varepsilon^{1/2} \\ 0 & \lambda_1 & \varepsilon^{1/2} \\ 0 & 0 & \lambda_1 \end{vmatrix} + \begin{vmatrix} \varepsilon & \varepsilon^{1/2} & \varepsilon^{1/2} & \varepsilon \\ \varepsilon & \varepsilon & \varepsilon^{1/2} & \varepsilon^{1/2} \\ \varepsilon^{1/2} & \varepsilon^{1/2} & \varepsilon & \varepsilon^{1/2} \\ \varepsilon & \varepsilon^{1/2} & \varepsilon^{1/2} & \varepsilon \end{vmatrix} \quad (3.81)$$

Gershgorin teoremi bu durumda yeterince küçük ε değeri için λ_1 merkezli ve $\varepsilon^{1/3}$ e orantılı çapa sahip disk içerisinde üç tane özdeğerin var olduğunu gösterir.

n. dereceden bir bölen için genel sonuç şimdi daha açıktır. §2 de verilen örnek, eğer B genel bir matris ise, bu durumda $\varepsilon^{1/n}$ çarpanından kaçınılamayacağını gösterir. Tabidir ki, ε mertebesinden veya sıfır mertebesinden özdeğerlerin perturbasyonu için ε mertebesinde özel perturbasyonlar olacaktır.

Durum 5: Birden fazla, en az biri lineer olmayan $(\lambda_i - \lambda_j)$ -yı içeren birden fazla bölen olduğunda, λ_i özdeğerlerinin perturbasyonu:

Bu durum için genel sonuçları vermeden önce, Jordan kanonik forma sahip bir matrisin basit bir örneğini göz önüne alalım.

$$\begin{vmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 1 \\ 0 & 0 & \lambda_1 \end{vmatrix} \quad \begin{vmatrix} \lambda_1 & 1 \\ 0 & \lambda_1 \end{vmatrix} \quad (3.82)$$

Merkezi λ_1 ve $\varepsilon \rightarrow 0$ limitinde çapı $O(\varepsilon^{1/2})$ olan bir disk içerisinde iki tane özdeğerin her zaman olup olmadığını sormak doğaldır. Buna cevap açıkçası hayırdır. Eğer (3.4) ve (5.1) pozisyonlarına ε perturbasyonunu yerleştirirsek karakteristik denklem aşağıdaki gibi olur:

$$(\lambda_1 - \lambda)^5 + \varepsilon^2 = 0 \quad (3.83)$$

Burada tüm perturbasyonlar $\varepsilon^{1/2}$ ile orantılıdır.

Bununla birlikte $\varepsilon^{1/2}$ den daha büyük mertebeden olmayan en az bir perturbasyonun olacağı ispat edilebilir. (3.82) matrisinin elemanlarını, ε mertebesinde bütün perturbasyonlarını göz önüne alalım. Karakteristik denklem, $(\lambda_1 - \lambda)$ nın kuvvetlerine açar ve determinantını göz önüne getirirsek sabit terimlerin aşağıdaki yapıda olur:

$$A_2\varepsilon^2 + A_3\varepsilon^3 + A_4\varepsilon^4 + A_5\varepsilon^5 \quad (3.84)$$

ε -dan bağımsız ve diğer ε cinsinden terimler yoktur. Eğer bu denklemin kökleri aşağıdaki gibi verilirse (3.86) ifadesine ulaşılır:

$$(\lambda - \lambda_i) = p_i \quad (3.85)$$

$$\prod_{i=1}^5 p_i = A_2\varepsilon^2 + A_3\varepsilon^3 + A_4\varepsilon^4 + A_5\varepsilon^5 \quad (3.86)$$

Böylece tüm p_i -nin $\varepsilon^{2/5}$ mertebesinde daha büyük olması imkansızdır. λ_i den farklı herhangi bir katlılığa sahip özdeğerlerin ilavesi bu sonucu etkilemez.

3.14 Lineer Olmayan Bölenlerin Genel Dağılımına Tekabül Eden Perturbasyon

$(\lambda_1 - \lambda)$ -daki yüksek dereceli temel-bölenlerin varlığı $(\lambda_1 - \lambda)$ lineer bölenlere karşılık gelen özdeğerlerin de hassaslığını etkiler. Fakat özdeğerler λ_1 den farklı değildir. Aşağıdaki sonuçlar ispatlanabilir.

λ_1 e karşılık gelen temel-bölenleri aşağıdaki gibi olsun:

$$(\lambda_1 - \lambda)^{r_1}, (\lambda_1 - \lambda)^{r_2}, \dots, (\lambda_1 - \lambda)^{r_s} \quad (3.87)$$

Burada

$$r_1 > r_2 > \dots > r_s \quad \sum r_i = t \quad (3.88)$$

Bu durumda yeterince küçük ε için $p_1, p_2, \dots, p_t$ perturbasyonları aşağıdaki ifadeyi sağlar:

$$\prod_{i=1}^t p_i = A_s \varepsilon^s + A_{s+1} \varepsilon^{s+1} + \dots + A_t \varepsilon^t \quad (3.89)$$

Böylece en az bir λ , merkezi λ_1 ve herhangi bir K_1 değeri için çapı $K_1 \varepsilon^{s/t}$ olan dairesel çember içerisinde olur. Diğer taraftan tüm karşılık gelen perturbe edilmiş özdeğerler de merkezi λ_1 ve herhangi bir K_2 değeri için çapı $K_2 \varepsilon^{1/t}$ olan bir disk içerisine yayılmışlardır.

3.15 Jordan Kanonik Formundan Özvektörlerin Perturbasyon Teorisi

Özvektörlerin perturbasyonu Jordan kanonik formu kullanılarak yapılabilir. Problemin tüm bir genellemesi yerine A lineer temel-bölenleri olan durumu ile ilgilenilecektir. Böylece aşağıdaki ifade yazılabilir:

$$H^{-1}AH = \text{diag}(\lambda_i) \quad (3.90)$$

H -nin kolonları, A nın $x_1, x_2, \dots, x_n$ özvektörlerin tanımı kümesidir. $(A + \varepsilon B)$ nın karşılık gelen özvektörlerini $x_1(\varepsilon), x_2(\varepsilon), \dots, x_n(\varepsilon)$ ile, $H^{-1}(A + \varepsilon B)H$ -in özvektörlerini de $z_1(\varepsilon), z_2(\varepsilon), \dots, z_n(\varepsilon)$ ile aşağıdaki koşul altında ifade edeceğiz:

$$x_i(\varepsilon) = H z_i(\varepsilon) \quad (3.91)$$

x_i özvektörlerin tam bir kümesini oluşturduğundan

$$H z_i(\varepsilon) = \alpha_1(\varepsilon) x_1 + \alpha_2(\varepsilon) x_2 + \dots + \alpha_n(\varepsilon) x_n \quad (3.92)$$

$\alpha_i(\varepsilon)$ kümesi vasıtasıyla notasyonu basitleştirmek için §17 de kullanılan altıncı mertebeden bir matrisi göz önüne alalım:

$$H^{-1}(A+\varepsilon B)H = \begin{vmatrix} \lambda_1 & & & & & \\ & \lambda_1 & & & & \\ & & \lambda_1 & & & \\ & & & \lambda_4 & & \\ & & & & \lambda_4 & \\ & & & & & \lambda_6 \end{vmatrix} + \begin{vmatrix} \beta_{11}/s_1 & \beta_{12}/s_1 & \beta_{13}/s_3 & \dots & \beta_{16}/s_1 \\ \beta_{21}/s_2 & \beta_{22}/s_2 & \beta_{23}/s_2 & \dots & \beta_{26}/s_2 \\ \dots & \dots & \dots & \dots & \dots \\ \beta_{61}/s_6 & \beta_{62}/s_6 & \beta_{63}/s_6 & \dots & \beta_{66}/s_6 \end{vmatrix}$$

Eğer λ_6 basit özdeğeri incelendiğinde $z_6(0) = e_6$ olduğu görülür. Karşılık gelen $z_6(\varepsilon)$ -yı en büyük bileşeni 1 olacak şekilde normalize edelim. Yeterince küçük ε değeri için bu altıncı bileşen olmalı. Bunun yerine dört bileşen olduğunu düşünelim ve aşağıdaki ifadeye $z_{46}(\varepsilon)=1$ koyalım:

$$\lambda_6(\varepsilon)[z_{46}(\varepsilon)] = \lambda_4 z_{46}(\varepsilon) + (\varepsilon/s_4) \sum_{i=1}^6 \beta_{4i} z_{i6}(\varepsilon) \quad (3.93)$$

$$\lambda_6 - \lambda_4 = (\varepsilon/s_4) \sum_{i=1}^6 \beta_{4i} z_{i6}(\varepsilon) \quad (3.94)$$

Şimdi $\varepsilon \rightarrow 0$ limitinde sol taraf $\lambda_6 - \lambda_4$ ve sağ taraf sıfır olur. Böylece bir çelişki elde edileceğinden, yeterince küçük ε için $z_{66}(\varepsilon)=1$ elde edilir.

Şimdi $\varepsilon \rightarrow 0$ limitinde herhangi bir K değeri için diğer bileşenlerinin tümünün $K\varepsilon$ dan küçük olup olmadığını inceleyelim.

$$\lambda_6(\varepsilon) z_{i6}(\varepsilon) = \lambda_i z_{i6}(\varepsilon) + (\varepsilon/s_i) \sum_{j=1}^6 \beta_{ij} z_{j6}(\varepsilon) \quad (3.95)$$

$$|\lambda_6(\varepsilon) - \lambda_i| |z_{i6}(\varepsilon)| < (\varepsilon/|s_i|) \sum_{j=1}^6 |\beta_{ij}| \quad (3.96)$$

$$|\lambda_6(\varepsilon) - \lambda_i| > \frac{1}{2} |\lambda_6 - \lambda_i| \quad (3.97)$$

$$|z_{i6}(\varepsilon)| < 2\varepsilon \sum_{j=1}^6 |\beta_{ij}| / |s_i| |\lambda_6 - \lambda_i| \quad (i=1, \dots, 5)$$

$z_{i6}(\varepsilon)$ ($i=1, \dots, 5$) ε mertebesinde olduğundan geliştirilmiş bir sınır elde edilebilir. (3.95) aşağıdaki ifadeyi verir:

$$\{\lambda_6(\varepsilon) - \lambda_i\} z_{i6}(\varepsilon) = \frac{\varepsilon \beta_{i6}}{s_i} + \varepsilon \sum_{j=1}^5 \beta_{ij} z_{j6}(\varepsilon) \quad (3.98)$$

Sağ taraftaki ikinci terim ε^2 mertebesinde.

$$| z_{i6}(\varepsilon) - \varepsilon \beta_{i6}/s_i(\lambda_6 - \lambda_i) | = O(\varepsilon^2) \quad (3.99)$$

Sonuç esas olarak (3.53) dir. Fakat bundan sonra ε^2 terimi için sıkı bir sınırın nasıl elde edileceğini inceleyeceğiz. Bu nümerik hesaplamalarda önemlidir.

3.16 Katlı Özdeğerlere Karşılık Gelen Özvektörlerin Perturbasyonu (Lineer Temel-Bölenler)

Katlı özdeğerlere karşılık gelen özvektörler için daha fazla ispat yapmayacağız. Eğer $z_5(\varepsilon)$, $H^1(A + \varepsilon B)H$ matrisinin $\lambda_5(\varepsilon)$ a karşılık gelen normalize edilmiş özvektörü ise §24 te olduğu gibi $z_5(\varepsilon)$ nın en büyük elemanının 1,2,3 veya 6 olamayacağı ve bunların ε mertebesinde olduğu gösterilebilir. Normalize edilmiş $z_5(\varepsilon)$ aşağıdaki yapılardan birine sahip olmalıdır.

$$[z_{15}(\varepsilon), z_{25}(\varepsilon), z_{35}(\varepsilon), k(\varepsilon), 1, z_{65}(\varepsilon)] \quad (3.100)$$

$$[z_{15}(\varepsilon), z_{25}(\varepsilon), z_{35}(\varepsilon), 1, k(\varepsilon), z_{65}(\varepsilon)] \quad (3.101)$$

Burada $|k(\varepsilon)| < 1$ dir. $(A + \varepsilon B)$ nın karşılık gelen özvektörü $x_5(\varepsilon)$, x_1, x_2, x_3, x_6 doğrultusunda ε mertebesinde bileşenlere sahiptir. Ancak x_4 ve x_5 ile gerilmiş alt uzay bileşenleri de $x_4 + k(\varepsilon)x_5$ veya $k(\varepsilon)x_4 + x_5$ şeklindedir. Bunun üzerinde başka bir iyileştirme bekleyemeyiz. Çünkü bu alt uzaydaki özvektörler A nın gerçek özvektörleridir.

3.17 Perturbasyon Teoreminin Sınırları

Şu ana kadar incelediğimiz perturbasyon teorisi oldukça büyük pratik değere sahip olmasına rağmen, bu bölümün de ana konusu olan, bazı eksik kalan taraflarını inceleyeceğiz.

$\varepsilon \rightarrow 0$ limitinde perturbasyonların doğal yapısına kendimizi konsantre ettik. Fakat aşağıdaki basit örneği göz önüne alalım. Basit bir matris ile inceleyelim:

$$\begin{vmatrix} a & 10^{-10} \\ 0 & a \end{vmatrix} \quad (3.102)$$

Bu matris $(a-\lambda)^2$ lineer-olmayan temel-bölenlere sahiptir ve eğer (2,1) deki elemanın içerisine bir ε perturbasyonu eklenirse, özdeğerler $a \pm (10^{-10}\varepsilon)^{1/2}$ olur. Özdeğerlerin ε -na göre türevi alınırsa $\varepsilon \rightarrow 0$ limitinde özdeğerler de sonsuza gider. Ancak 10^{-10} mertebesinde ki perturbasyonlar bu konunun dışındadır. Eğer perturbasyon matrisi aşağıdaki gibiyse:

$$\begin{vmatrix} \varepsilon_1 & \varepsilon_2 \\ \varepsilon_3 & \varepsilon_4 \end{vmatrix} \quad (3.103)$$

(3.105) matrisinin (3.106) perturbasyon matrisinin terimlerindeki bu problemi göz önüne alalım.

$$\begin{vmatrix} a & 10^{-10} \\ 0 & a \end{vmatrix} + \begin{vmatrix} \varepsilon_1 & \varepsilon_2 \\ \varepsilon_3 & \varepsilon_4 \end{vmatrix} \quad (3.104)$$

$$\begin{vmatrix} a & 0 \\ 0 & a \end{vmatrix} \quad (3.105)$$

(26.4) matrisinin daha fazla bir lineer-olmayan böleni yoktur.

Şimdi özvektörlere dönersek, $x_1, x_2, \dots, x_n$ doğrultusundaki bileşenler içerisine perturbasyonu çözümledik. Eğer x_i ortogonal ise sonuç oldukça tatmin edicidir. Fakat x_i -lerden bazıları hemen hemen lineer-bağımlı ise perturbasyon vektör oldukça küçük olmasına rağmen her x_i -nin doğrultularında oldukça büyük bileşene sahip oluruz. §10 da, λ_1 basit özdeğerine karşılık gelen perturbasyonu aşağıdaki gibi verilmişti:

$$\varepsilon \sum_{i=2}^n \frac{\beta_{i1}}{s_i(\lambda_1 - \lambda_i)} x_i + O(\varepsilon^2) \quad (3.106)$$

Böylece genelde s_i küçük ise x_i -nin doğrultusundaki bileşenlerin oldukça büyük olacağını beklenir.

Eğer aşağıdaki matris göz önüne alınırsa bunun nasıl bir yanlışığa götüreceği görülür:

$$\begin{vmatrix} 2 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1+10^{-10} \end{vmatrix} \quad (3.107)$$

$$s_1=1, s_2 = -10^{-10}/(1+10^{-20})^{1/2}, s_3 = 10^{-10}/(1+10^{-20})^{1/2} \quad (3.108)$$

$\lambda=2$ ye tekabül eden x_1 vektörü, matristeki perturbeye karşı duyarlı değildir. x_2 ve x_3 vektörleri hemen hemen aynıdır (aynı zamanda y_1 ve y_2) ve $\beta_{21}/s_2(\lambda_1-\lambda_2)$ ve $\beta_{31}/s_3(\lambda_1-\lambda_3)$ hemen hemen eşit ve zıt işaretlidirler. Böylece x_2 ve x_3 yönündeki vektörler birbirlerini yok ederler.

3.18 s_i -ler arasındaki ilişki

Bu örnekte $1/s_2$ ve $1/s_3$ oldukça büyük, fakat eşit ve zıt işaretlidirler. $1/s_i$ -nin birinin büyük olması ihtimalini engelleyerek, s_i -nin birbirine bağımlı olduğunu göstereceğiz. Eğer aşağıdaki ifade yazılırsa:

$$\|x_j\|_2 = \|y_j\|_2 = 1 \text{ iken} \quad (3.109)$$

$$x_i = \sum_{j=1}^n \alpha_{ij} y_j, \quad y_i = \sum_{j=1}^n \beta_{ij} x_j$$

$$\begin{aligned} \alpha_{ij} &= x_j^T x_i / x_j^T y_j = x_j^T x_i / s_j, \\ \beta_{ij} &= y_j^T y_i / y_j^T x_j = y_j^T y_i / s_j \end{aligned} \quad (3.110)$$

$$y_i^T x_i = \sum_{j=1}^n \beta_{ij} x_j^T \sum_{j=1}^n \alpha_{ij} y_j = \sum_{j=1}^n \beta_{ij} \alpha_{ij} s_j = \sum_{j=1}^n (x_j^T x_i) (y_j^T y_i) s_{j-1} \quad (3.111)$$

$$s_i = s_i^{-1} + \sum_{j \neq i} (\cos \theta_{ij} \cos \varphi_{ij}) s_j^{-1} \quad (3.112)$$

Burada θ_{ij} , x_i ile x_j arasındaki açıdır. φ_{ij} ise y_i ile y_j arasındaki açıdır. (3.112) denklemi i -nin her bir değeri için doğrudur. Buradan aşağıdaki sonuca ulaşılır:

$$|s_i^{-1}| < |s_i| + \sum_{j \neq i} |(\cos \theta_{ij} \cos \varphi_{ij}) s_j^{-1}| < 1 + \sum_{j \neq i} |s_j^{-1}| \quad (3.113)$$

Bu örnek aşağıdaki koşulları sağlayan üçüncü mertebeden bir matrisin, bir s_i kümesine sahip olamayacağını gösterir.

$$|s_3^{-1}| \geq 1 + |s_1^{-1}| + |s_2^{-1}| \text{ olduğundan } |s_1^{-1}| = 10, |s_2^{-1}| = 10^7, |s_3^{-1}| = 10^8$$

3.19 Hesaplanan bir Problemin Kondüsyonu

Bu bölümde, hesaplanan değerler, datalardaki en küçük değişikliğe bile oldukça hassas ise *kötü-konumlu (ill-conditioned)* olarak isimlendirilen bazı hesaplama yöntemleri ile ilgileneceğiz. Önceki bölümleri bir matrisin özsystemlerinin hassaslığına karar veren faktörlere uygun bir bakış açısı vermişti. Bir matrisin elemanlarındaki en küçük bir değişikliğe bile hassas olan bazı özdeğerler olabilir. Bazıları ise yeterince daha az hassas olabilir. Benzer olarak özvektörlerin bazıları *iyi-konulu (well-conditioned)* bazıları da *kötü-konumlu* olabilir. Aynı zaman da bir özdeğere *kötü-konumlu* olurken karşılık gelen özvektör olmayabilir.

Özdeğer problemlerinin çözümüyle alakalı olarak bir matrisin kötü-konumlu olup olmadığına karar vermek, matrisin tersinin hesaplanmasıyla alakalı koşullara karar vermekten farklıdır. Bir normalize edilmiş matris eğer çok küçük bir özdeğere sahip ise kötü-konumlu olarak göz önünde bulundurulabilir. Fakat bu, özproblemlerin iyi-konumluluk şartı için önemli değildir. k nın uygun bir seçimi ile, $(A - kI)$ matrisini tam olarak tekil yapabiliriz. Fakat $(A - kI)$ nın özsysteminin hassaslığı A nın özsystemi ile aynıdır. Bununla birlikte bu iki hesaplanan problemlerin koşulları arasında bazı bağlantıların olduğunu göstereceğiz.

3.20 Kondüsyon Sayıları

Bir hesaplanan problemle ilgili olarak bir matrisin kondüsyonunu veren bazı sayılara sahip olmak pratik açıdan oldukça önemlidir. Bu tür sayılara *kondüsyon-sayısı (condition-number)* adı verilir. Bu, katsayıların değişimlerine göre çözümün değişiminin hızını belirleyen bir göstergedir. Değişim hızı bu sayı ile orantılıdır.

Daha önce de belirtildiği gibi, sadece özdeğer probleminin hesabı ile sınırlasak bile böyle bir sayı bazı kısıtlamalara sahiptir. Eğer özdeğerlerden herhangi biri çok hassas ise (diğerleri hassas olmasa bile) kondüsyon-sayısı oldukça büyük olmalıdır. Bu kısıtlamalara rağmen lineer elemanter bölenlere sahip tek kondüsyon sayısı yaygın olarak kullanılmaktadır.

3.21 Özdeğer Problemine Karşılık Gelen Spektral Kondüsyon Sayısı

A lineer temel-bölenlere sahip ve H matrisi de aşağıdaki gibi olsun:

$$H^{-1}AH = \text{diag}(\lambda_i) \quad (3.114)$$

Eğer λ $(A+\epsilon B)$ matrisinin özdeğeri ise $(A+\epsilon B-\lambda I)$ matrisi tekildir ve böylece determinantı sıfırdır. Bu durumda aşağıdaki ifade elde edilir:

$$H^{-1}(A + \epsilon B - \lambda I)H = \text{diag}(\lambda_i - \lambda) + \epsilon H^{-1}BH \quad (3.115)$$

Determinantı alınarak (3.115) nın sağ tarafındaki matrisin de tekil olduğu görülür. Burada iki durum vardır:

Durum 1 : $\lambda = \lambda_i$ bazı i değerleri için

Durum 2 : $\lambda \neq \lambda_i$ herhangi i değeri için. Öyle ki aşağıdaki ifade yazılabilir:

$$\text{diag}(\lambda_i - \lambda) + \epsilon H^{-1}BH = \text{diag}(\lambda_i - \lambda)[I + \epsilon \text{diag}(\lambda_i - \lambda)^{-1}H^{-1}BH] \quad (3.116)$$

Tekrar determinant alınarak parantez içindeki matrisin tekil olduğu görülür. Eğer $(I-X)$ tekil ise tüm normlar için $\|X\| > 1$ yazılabilir. Eğer $\|X\| < 1$ ise $(I+X)$ matrisinin özdeğerlerinin hiçbiri sıfır olamaz. Böylece aşağıdaki ifade elde edilir:

$$\|\epsilon \cdot \text{diag}(\lambda_i - \lambda)^{-1}H^{-1}BH\|_2 > 1 \quad (3.117)$$

$$\epsilon \cdot \max_i |(\lambda_i - \lambda)^{-1}| \|H^{-1}\|_2 \|B\|_2 \|H\|_2 > 1 \quad (3.118)$$

$$\min_i |\lambda_i - \lambda| < \epsilon \cdot \|H^{-1}\|_2 \|H\|_2 \|B\|_2 \quad (3.119)$$

i nin en küçük değeri için

$$|\lambda_i - \lambda| < \epsilon \kappa(H) \|B\|_2 \quad (3.120)$$

Burada

$$\kappa(H) = \|H^{-1}\|_2 \|H\|_2 \quad (3.121)$$

olur.

Spektral normuna olan bağımlılığından $\kappa(H)$ genellikle *spektral kondüsyon-sayısı* olarak adlandırılır. A'nın özdeğerlerinin tüm hassaslığı böylece $\kappa(H)$ ın büyüklüğüne bağlı olduğu görülür. Böylece $\kappa(H)$, özdeğer problemine karşılık A'nın bir kondüsyon sayısı gibi göz önüne alınabilir. (Bauer ve Fike (1960)).

(30.6) nın aşağıdaki koşullar altında herhangi bir norm için doğru olduğu görülür:

$$\|\text{diag}(\lambda_i - \lambda_j)^{-1}\| = \max |\lambda_i - \lambda_j|^{-1} \quad (3.122)$$

Bu ifadeler 1-norm ve ∞ -norm için de doğrudur.

Süreklilik kabullerini kullanarak Gershgorin' nin ikinci teoremini isbat etmek için kullanılan metodu kullanarak köklerin yerleri daha da doğru tespit edilebilir. Bu aşağıdaki sonuca götürür. Eğer s tane dairesel disk

$$|\lambda_i - \lambda| = \varepsilon \kappa(H) \|B\|_2$$

diğerlerinden izole edilmiş bir bağlantılı bölgeyi yapılandırıyor, bu bölgede s tane özdeğer vardır. Burada ε -nun küçük olması gerekli değildir.

3.22 Spectral Kondüsyon Sayısının Özellikleri

H matrisi tek türlü olmadığından (özdeğerler ayrık olsa bile, her bir kolon keyfi bir çarpan ile çarpılabilir) bütün H -ler için $\kappa(H)$ nın en küçük değeri A -nın özdeğer problemine göre, A'nın *spektral kondüsyon-sayısı* olur. Herhangi bir durum için aşağıdaki ifadeyi yazalım:

$$\kappa(H) = \|H^{-1}\|_2 \|H\|_2 > \|H^{-1}H\|_2 = 1 \quad (3.123)$$

Eğer A (bölüm 1, §48) normal ise ve özellikle Hermitian veya birim matris ise, H -yi unitary matris olması için alabiliriz ve böylece aşağıdaki eşitliği elde ederiz:

$$\kappa(H) = 1 \quad (3.124)$$

Böylece özdeğer problemi, bir normal matris için daima iyi-konumlu olur. Fakat bu özvektör problemleri için yeterince olması gerekmez.

Şimdi $\kappa(H)$ nın herbir özdeğerin hassaslığını belirleyen $|s_i^{-1}|$ sayısı ile olan ilişkisini göz önüne alalım. λ_i e karşılık gelen sağ ve sol özvektörlerin normalize edilmesi aşağıdaki ifadeleri ile verilir:

$$\begin{aligned} x_i &= He_i / \|He_i\|_2 \\ y_i &= (H^{-1})^T e_i / \|(H^{-1})^T e_i\|_2 \end{aligned} \quad (3.125)$$

$$|s_i| = |y_i^T x_i| = \frac{|e_i^T H^{-1} He_i|}{\|He_i\|_2 \|(H^{-1})^T e_i\|_2} = \frac{1}{\|He_i\|_2 \|(H^{-1})^T e_i\|_2} \quad (3.126)$$

$$\|He_i\|_2 < \|H\|_2 \|e_i\|_2 = \|H\|_2 \quad (3.127)$$

$$\|He_i\|_2 < \|H\|_2 \|e_i\|_2 = \|(H^{-1})^T\|_2 = \|H^{-1}\|_2 \quad (3.128)$$

$$|s_i^{-1}| < \kappa(H) \quad (3.129)$$

Diğer taraftan H -nin kolonları $x_i/s_i^{1/2}$ ve H^{-1} in satırları da $y_i^T/s_i^{1/2}$ olacak şekilde alınabilir. H -nin bu seçimi ile:

$$\kappa(H) = \|H\|_2 \|H^{-1}\|_2 < \|H\|_E \|H^{-1}\|_K = (\sum_{i=1}^n |s_i^{-1}|)^{1/2} (\sum_{i=1}^n |s_i|)^{1/2} = \sum_{i=1}^n |s_i^{-1}| \quad (3.130)$$

Elemanlardaki perturbasyonla alakalı olarak A nın özdeğerlerinin hassaslığının tüm bilgisi için n^3 tane $\lambda_i / \partial a_{ij}$ ifadesine ihtiyaç duyulur. Çünkü A nın her bir elemanlarındaki değişiklikle alakalı olarak her bir özdeğerin hassaslığı oldukça geniş bir aralıkta değişebilir. Pratik çalışmalar için bu büyüklük imkansız sonuçlara ulaştırır. Uygun bir uzlaşma n tane $|s_i^{-1}|$ sayısı ile sağlanır. Bu, özdeğer problemine uygun olarak A nın n *kondüsyon sayısı* olarak ifade edilir. (Wilkinson J.H, 1963)

3.23 Kondüsyon Sayılarının Değişmezlik Özellikleri

κ ve n tane kondüsyon $|s_i^{-1}|$ sayısının her ikisi de üniter benzerlik-dönüşümleriyle alakalı olarak önemli değişmez özelliklere sahiptir. Eğer R üniter matris ise ve aşağıda ki ifadeyi sağlıyor ise

$$B = RAR^H \quad (3.131)$$

$$H^{-1}AH = \text{diag}(\lambda_i) \quad (3.132)$$

aşağıdaki eşitlik bulunur:

$$\begin{aligned} H^{-1}R^HB R^H &= \text{diag}(\lambda_i) \quad \text{veya} \\ (R^H)^{-1}B(R^H) &= \text{diag}(\lambda_i) \end{aligned} \quad (3.133)$$

Böylece B -nin spectral kondüsyon sayısı κ aşağıdaki gibi verilir:

$$\kappa = \|(R^H)^{-1}\|_2 \|R^H\|_2 = \|H^{-1}\|_2 \|H\|_2 = \kappa \quad (3.134)$$

Benzer olarak eğer x_i ve y_i A nın sağ ve sol özvektörleri ise B -nin kiler aşağıdaki gibi verilir:

$$s_i = (v_i^T)^T x_i = v_i^T R^T R x_i = v_i^T R^H R x_i = v_i^T x_i = s_i \quad (3.135)$$

Buradan her bir özdeğerin hassaslığının unitary dönüşümü altında invariant olduğu görülür.

2.24 Çok Kötü-Kondüsyonlu Matrisler

Genelde lineer-olmayan temel-bölenlere karşılık gelen özdeğerler §26 daki kabulleri göz önünde bulundurmamıza rağmen kötü-konumlu olarak düşünmeliyiz. Bununla birlikte bunu kötü-konumlu kabulünün ana yapısı olduğunu düşünmemeliyiz. Eğer özdeğerler ayrık ve iyi yayılmış olsalar bile kötü-konumlu olabilirler. Bunu aşağıdaki örnekle daha iyi gösterebiliriz.

Aşağıdaki gibi tanımlanan 20. mertebeden A matrisini göz önüne alalım:

$$A = \begin{vmatrix} 20 & 20 & & & \\ & 19 & 20 & & \\ & & 18 & 20 & \\ & & & \dots & \\ & & & 2 & 20 \\ & & & & 1 \end{vmatrix} \quad (3.136)$$

Bu bir üçgen-matristir ve özdeğerleri köşegen elemanlardır. Eğer ε değeri (20,1) pozisyonuna eklenirse karakteristik denklem aşağıdaki şekle dönüşür:

$$(20-\lambda)(19-\lambda)\dots(1-\lambda) = 20^{19}\varepsilon \quad (3.137)$$

§9 den, yeterince küçük ε için $\lambda=r$ özdeğerinin perturbasyonu ε nın kesirli kuvvetlerini içermeyeceği görülmüştü. Eğer aşağıdaki ifade yazılırsa:

$$\lambda_r(\varepsilon) - r \sim K_r \varepsilon \quad (3.138)$$

(3.137) den K_r aşağıdaki gibi verilecektir:

$$K_r = 20^{19} (-1)^r / (20-r)!(r-1)! \quad (3.139)$$

Bu sabit r -nin tüm değerleri için büyüktür. En küçük değer K_1 ve K_{20} dir. En büyük değer ise K_{10} ve K_{11} dir.

$$\begin{aligned} -K_1 &= K_{20} : 10^7 \quad (4.31) \\ -K_{11} &= K_{10} : 10^{12} \quad (3.98) \end{aligned} \quad (3.140)$$

K_i o kadar büyüktür ki ε çok küçük olana kadar lineer teori geçerli değildir. $\varepsilon=10^{-10}$ için özdeğerler tablo 1 de gösterilmektedir. Burada 10.5 –e göre simetriklik vardır.

Perturbe edilmiş matrisin özdeğerleri ($\varepsilon=10^{-10}$)

0.99575439	3.9653307 ± i 1.08773570
20.00424561	17.03466930 ± i 1.08773570
2.10924184	5.89397755 ± i 1.94852927
18.89075816	15.10602245 ± i 1.94852927
2.57488140	8.11807338 ± i 2.52918173
18.42511860	12.88192662 ± i 2.52918173
	10.50000000 ± i 2.73339736

Bu tek perturbasyona karşılık gelen özdeğerlerdeki değişiklikler, 10. dereceden temel-bölenlere sahip $C_{10}(a)$ matrisinin özdeğerlerindeki değişiklikler kadar büyüktür. $\varepsilon \rightarrow 0$ limitinde birinci ε mertebesinde ve $\varepsilon^{1/10}$ mertebesinde olacağı gerçeği, perturbasyonları 10^{-10} dan çok daha küçük alınana kadar önemli olmaz. (Wilkinson J.H, 1967)

(3.136) ile verilen matrisin özdeğerlerinin hassaslığı s_i -nin küçük olması gerektiğini gösterir. $\lambda=r$ ye karşılık gelen A nın x_r özvektörü aşağıdaki bileşenlere sahiptir:

$$\left[1; \frac{20-r}{-20}; \frac{(20-r)(19-r)}{(-20)^2}; \dots; \frac{(20-r)!}{(-20)^{20-r}}; 0; \dots; 0 \right] \quad (3.141)$$

y_r de aşağıdaki bileşenlere sahiptir:

$$\left[0; \dots; 0; \frac{(r-1)!}{20^{r-1}}; \dots; \frac{(r-1)(r-2)}{20^2}; \frac{r-1}{20}; 1 \right] \quad (3.142)$$

Bu vektörler 2-normalize vektörleri değildir, fakat $y_r^T x_r$, s_r -nin büyüklüğünün iyi bir tahminini verir.

$$|y_r^T x_r| = (20-r)!(r-1)!/20^{19} \quad (3.143)$$

(3.143) ün sağ tarafının karşılığı (3.144) denkleminin $|K_r|$ değeridir. Ayrıca A ve A^T -nin özvektörleri hemen hemen ortogonaldır. Kesin ortogonalite sadece lineer-olmayan bölenler için bulunmasına rağmen, iyi ayrılmış özdeğerlere sahip bir matrise oldukça yaklaşıyoruz.

A matrisine yakın öyle bir matris vardır ki, bu matris bazı lineer-olmayan temel-bölenlere sahiptir. Gerçekten böyle bir matris (3.142) elemanını perturbe ederek üretebiliriz. E perturbasyonuna karşılık gelen özdeğerler (3.140) denkleminin kökleridir. Şimdi bu denklemin sol tarafındaki fonksiyon $\lambda=10.5$ civarında simetriktir ve 9 maxima ve 10 minimaxa sahiptir.

Eğer ε , 0 dan itibaren artırılırsa 10 ve 11 beraberce 10.5 ile çakışırlar.

$$\left(\frac{1}{2} \frac{3}{2} \frac{1.9}{2} \right)^2 = 20^{19} \varepsilon \quad \varepsilon \approx 7.8 \times 10^{-14} \quad (3.144)$$

Bu çakışma ε yaklaşık 7.8×10^{-14} için olmaktadır. Çünkü $(A-\lambda I)$ perturbe matrisinin rankı keyfi bir λ için 19 dur. Ve 19. mertebeden matris üst tarafta 20^{19} determinantına sahiptir.

Eğer ε daha da artarsa kök 8, 9, 12, 13 hep beraber hareket eder ve ε -nun aynı değeri ile çakışırlar. Bu durumda matris, çift kat özdeğere sahip olur. ve her ikisi de kuadratik bölenlere karşılık gelirler. Benzer olarak ε nın artışı ile, biz kök 6, 7, 14 ve 15 ide berabere getirebiliriz v.b.

Biraz önce göz önünde bulundurduğumuz matrisin bütün özdeğerlerin bazıları diğerlerinden daha kötü olmasına rağmen, kötü-konumludur. Biz şimdi, geniş olarak değişen kondüsyonlara sahip bir matris kümesinin örneğini vereceğiz. Matrisler aşağıdaki gibi tanımlanan B_n sınıfındandır

$$B_n = \begin{vmatrix} n & (n-1) & (n-2) & \dots & 3 & 2 & 1 \\ (n-1) & (n-1) & (n-2) & \dots & 3 & 2 & 1 \\ (n-2) & (n-2) & (n-2) & \dots & 3 & 2 & 1 \\ & & & & 2 & 2 & 1 \\ & & & & 1 & 1 & 1 \end{vmatrix} \quad (3.145)$$

Bu matrisin en büyük özdeğerleri iyi-konumludur ve en küçükleri çok kötü-konumludur. $n=2$ için s_i -lerin ilk birkaç tanesi birim mertebesinde, son üçü ise 10^{-7} mertebesinde. n -nin değerinin artması ile, en küçük özdeğerler, artan bir şekilde daha kötü kondüsyonlu hale gelir.

Özdeğerlerin bir kısmının, bazı matris elemanlarındaki küçük değişikliklere hassas olması, aşağıdaki şartta açıktır. Matrisin determinanı, matrisin satırlarındaki basit operasyonlar ile görülebileceği gibi, n in tüm değerleri için 1 dir. Eğer $(1,n)$ elemanı $(1+\varepsilon)$ -na kadar değiştirilirse, determinant aşağıdaki şekle dönüşür.

$$1 \pm (n-1)! \varepsilon \quad (3.146)$$

Eğer $\varepsilon=10^{-10}$ ve $n=20$ ise determinant 1 den $(1 - 19/10^{-10})$ a değişir, ki bu, yaklaşık olarak -1.216×10^7 dir. Şimdi bir matrisin determinantının onun özdeğerlerinin çarpımına eşittir. Perturbe edilmiş matrisin özdeğerler inden en az biri, orijinalinden çok farklı olmalıdır.

Tekrar, $n=12$ iken (1,12) elemanındaki 10^{-9} dan daha küçük bir perturbasyon, bir kuadratik bölen üretir. (Gould S.H, 1974)

Bizim biraz önce tartıştığımız tipin kötü-konumlu olması, lineer-olmayan temel-bölenler ile doğrudan ilişkili olandan çok daha önemlidir. Tam lineer-olmayan bölenlere sahip matrisler, pratik çalışmalarda hemen hemen yoktur. Teorik çalışmalarda bile, genellikle küçük tam katsayılı, çok özel bir formun matrisleri ile ilişkilidirler. Eğer, böyle bir matrisin elemanları irrasyonel ise veya bir bilgisayar üzerinde tam tanımlanamıyorsa, yuvarlama hataları genellikle, lineer-olmayan temel-bölenlere sahip olmayan bir matrise taşır. Bazı küçük s_i -li matrisler, diğer taraftan oldukça yaygındır.

3.25 Reel Simetrik Matrisler için Perturbasyon Teorisi

Bundan sonraki bir iki bölümü, reel simetrik matrisler için perturbasyon teorisinin bir incelenmesine ayıracağız. Bu teori genel matrislerden daha basittir çünkü, x_i ve y_i vektör kümesinin ikisi de, eşit olarak alınabilirler (eğer bazı katlı özdeğerler varsa bile) ve böylece tüm s_i -ler 1 dir. Sadece §31 de hatırlattığımız gibi, reel simetrik matrisler için özdeğer problemi, genellikle iyi-konumludur. İzleyen bölümlerde ispat edeceğimiz bazı sonuçlar, derhal, normal matrisler sınıfının tümüne açık olan bir yol içerisine genişler. Diğerleri sadece Hermitian matrisine yada hiçbir şeye genişlemezler. Her bir durumdaki ilgili açılımları ispatsız işaret edeceğiz.

3.26 Simetrik-Olmayan Perturbasyonlar

Burada, bazen simetrik bir matrisin simetrik olmayan pertubasyonu ile ilgileneceğiz ve bağlantılı olarak, A nın simetrik ancak B -nin simetrik olmadığı $(A+\epsilon B)$ matrisinin özsisitemini inceleyeceğiz. (3.134) denkleminde, $(A+\epsilon B)$ -nin her bir özdeğerinin aşağıda verilen dairesel disklerden en az birinin içerisinde bulunduğunu biliyoruz.

$$|\lambda_i - \lambda_j| < \epsilon \|H^{-1}\| \|H\|_2 \|B\|_2 \quad (3.147)$$

ve böylece H ortogonal olarak alındığı için, bu aşağıdaki ifadeye indirgenir

$$|\lambda_i - \lambda_j| < \epsilon \|B\|_2 < n\epsilon \quad (\text{eğer } |b_{ij}| < 1 \text{ ise}) \quad (3.148)$$

Eğer bu disklerin herhangi s tanesi diğerlerinden izole edilmiş bir bağlantılı bölgeyi form ediyor ise bu bölgenin kesinlikle s tane özdeğeri vardır. Bu tüm normal formdaki matrisler için geçerlidir.

Reel simetrik matrislerin bazı durumlarında, genellikle reel B ile ilgileneceğiz. Böylece $(A + \epsilon B)$ -nin bazı kompleks özdeğerleri eşlenik çift gibi meydana çıkmalıdır. Eğer $|\lambda_i - \lambda_j| > 2n\epsilon$ ($j \neq i$) ise bu durumda, i . dairesel disk izole edilmiştir ve sadece bir özdeğer bulunur. Bu özdeğer reel olmalıdır. Böylece eğer, A nın tüm özdeğerleri $2n\epsilon$ -dan daha geniş bir ayrılmaya sahip rasgele simetrik perturbasyon modül olarak ϵ -nu geçmeyen tüm özel elemanlar hala özdeğerleri basit ve reel olarak bırakılır ve özvektörler ise reel ve tam bir sistem oluştururlar.

3.27 Simetrik Perturbasyonlar

Reel simetrik matrislerin özsystemlerinin hesaplanması için en geçerli tekniklerin çoğu, ortogonal benzerlik dönüşümleri kullanımı üzerine kuruludur. Gerçek simetri ardışık dönüşümlerde korunur. Bu sebeple simetrik matrislerin simetrik perturbasyonu ile ilgileneceğiz.

Perturbasyonlar simetrik olduğu zaman perturbe edilmiş matrisler reel özsystemlere sahip olacaktır ve aynı perturbe matrisleri içinde geçerlidir. Orijinal matrisin özdeğerleri ile perturbasyon matris ve perturbed matris arasındaki ilişkiyi araştırmak doğaldır. Bizim buraya kadar elde ettiğimiz sonuçların çoğu, eğer perturbasyonlar küçük değilse zayıftır. Diğer taraftan ispat etmek üzere olduğumuz sonuçların çoğu bu kısıtlamalara bağlı değildir. Böylece biz ϵ ile hazırlarız ve aşağıdaki ifadeyi yazarız

$$C = A + B \quad (3.149)$$

Burada A , B ve böylece C reel ve simetriktir.

3.28 Klasik Teknikler

İki simetrik matrisin toplamının özdeğerlerinin bulunması için minimax prensibi üzerine kurulu oldukça etkili yöntemler vardır. İlk önce çok klasik metot üzerine kurulu bir analiz ile başlayacağız.

Sınırlandırılmış köşegen bir matris için basit bir sonuç elde edelim. X aşağıdaki gibi tanımlanmış simetrik bir matris olsun:

$$X = \begin{vmatrix} \alpha & a^T \\ a & \text{diag}(\alpha_i) \end{vmatrix} \quad (i=1, \dots, n-1) \quad (3.150)$$

Burada a , $(n-1)$ mertebesinden bir vektördür. X in özdeğerlerini α_i ile ilişkilendirmek isteyelim.

Sadece a nın s adet bileşeninin sıfır olmadığı kabul edilsin. Eğer a_j sıfır ise bu durumda α_j X in bir özdeğeri dir. Sadece son $(n-1)$ satırı kapsayan P perturbe matrisinin uygun bir seçimi ile bir Y matrisi aşağıdaki gibi elde edilir.

$$Y = P^T X P = \begin{vmatrix} \alpha & b^T & O \\ b & \text{diag}(\beta_i) & O \\ O & O & \text{diag}(\gamma_i) \end{vmatrix} \quad (3.151)$$

Burada b -nin hiçbir bileşeni sıfır değildir. $\text{diag}(\beta_i)$ s ve $\text{diag}(\gamma_i)$ $(n-1-s)$ mertebesindedir. β_i ve γ_i beraberce α_i -nin yeniden düzenlenmiş halidir. Not edelim ki γ_i yi içeren bazıları veya hepsi katlı olan özvektörler vardır. Eğer varsa $\text{diag}(\gamma_i)$ dirler.

X in özdeğerleri böylece aşağıdaki gibi tanımlanan Z matrisinin özdeğerleriyle beraber γ_i dir.

$$Z = \begin{vmatrix} \alpha & b^T \\ b & \text{diag}(\beta_i) \end{vmatrix} \quad (3.152)$$

Eğer $s=0$ ise Z tek elemanlıdır ve α dır. Böylece X in özdeğerleri değerleri α olan $\text{diag}(\alpha_i)$ dir. Diğer taraftan aşağıdaki gibi olan Z -nin karakteristik polinomunu inceleyelim:

$$(\alpha - \lambda) \prod_{i=1}^s (\beta_i - \lambda) - \sum_{j=1}^s b_j^2 (\beta_i - \lambda) = 0 \quad (3.153)$$

Genelliği bozmadan, β_i arasında sadece t ayrık değer olduğunu ve bunların aşağıdaki gibi $r_1, r_2, \dots, r_t$ çarpanları olduğunu kabul edelim:

$$r_1 + r_2 + \dots + r_t = s \quad (3.154)$$

Eğer β_i -lerin tamamı basit ise $t_2 = s$ olur.

Bu durumda (39.4) ün sol tarafı aşağıdaki çarpana sahip olur:

$$\prod_{i=1}^t (\beta_i - \lambda)^{r_i-1} \quad (3.155)$$

Öyle ki $\beta_i, (r_i-1)$ (3.154) ü $\prod_{i=1}^t (\beta_i - \lambda)^{r_i}$ ile bölerek,

$$0 = (\alpha - \lambda) - \sum_{i=1}^t c_i^2 (\beta_i - \lambda)^{-1} \quad \alpha = f(\lambda) \quad (3.156)$$

denkleminin kökleri olarak elde edilir. Burada her bir c_i^2, β_i ile ilgili r_i tane b_j^2 -nin toplamıdır ve kesin pozitiftir. Farklı β_i -ler azalan sıradadır. $\delta_1, \delta_2, \dots, \delta_{t+1}$ ile ifade edilen $\alpha = f(\lambda)$ nın $t+1$ kökü aşağıdaki ilişkileri sağlar:

$$\begin{aligned} \infty &> \delta_1 \neq \beta_1; & \beta_{i-1} &> \delta_i > \beta_i & (i = 2, 3, \dots, t) \end{aligned} \quad (3.157)$$

Böylece X -in n özdeğeri üç kümenin içerisine düşer:

- (i) α_i nin sıfırlarına karşılık gelen $\gamma_1, \gamma_2, \dots, \gamma_{n-1-s}$ özdeğerleri. Bunlar α_i -nin $n-1-s$ -e eşittirler.
- (ii) Değerleri β_i olan ($i=1, 2, \dots, t$) r_i-1 tane $s-t$ özdeğer. Bunlar α_i -nin $s-t$ -nin ilerisine eşittir.
- (iii) (3.157) ilişkisini sağlayan $t+1$ özdeğer, δ_i ye eşittir. Eğer $t=0$ ise $\delta_1 = \alpha$ dır.

Not edelim ki (i) ve (ii) boş küme olabilirler. Her halükarda bu kümelerin elemanları α da bağımsızdırlar. Eğer X -in özdeğerlerini λ_i ile gösterirsek ve artan sırada sıralarsak α_i leride artan sırada dizersek

$$\lambda_1 > \alpha_1 > \lambda_2 > \alpha_2 > \dots > \alpha_{n-1} > \lambda_n \quad (3.158)$$

elde edilir. Diğer bir deyişle α_i, λ_i -leri zayıf anlamda ayırır. (Wilkinson, Rheinsch, 1971)

Şimdi X den elde edilmiş X_{α} nün özdeğerlerini α yerine α_{α} nı yerleştirerek inceleyelim. X_{α} nün özdeğerleri X özdeğerlerine eşit olacaktır ((i) ve (ii) de olduğu gibi). Bunlar (iii) kümesinde $\delta_{\alpha_1}, \delta_{\alpha_2}, \dots, \delta_{\alpha_{i+1}}$ ile ifade edilsin. Böylece $\lambda > 0$ için aşağıdaki ifade elde edilir:

tüm λ -ları için

$$\frac{df}{d\lambda} = 1 + \sum_1^t c_i^2 / (\beta_i - \lambda)^2 > 1 \quad (3.159)$$

Böylece her $\delta_{\alpha_i} - \delta_i$, 0 ile $\alpha_{\alpha} - \alpha$ arasına yayılacaktır.

$$\delta_{\alpha_i} - \delta_i = m_i(\alpha_{\alpha} - \alpha) \quad (3.160)$$

O halde

$$\begin{aligned} 0 < m_i < 1 \quad \text{ve} \\ \sum m_i &= 1 \end{aligned} \quad (3.161)$$

Eğer $t = 0$ ise $\delta_{\alpha_1} = \alpha_{\alpha}$ ve $\delta = \alpha$ olur.

$$\delta_{\alpha_1} - \delta_1 = \alpha_{\alpha} - \alpha \quad (3.162)$$

$$\delta_{\alpha_i} - \delta_i = m_i(\alpha_{\alpha} - \alpha) \quad (3.163)$$

$$\begin{aligned} 0 < m_i < 1 \\ \sum m_i &= 1 \end{aligned} \quad (3.164)$$

X ve X_{α} nün diğer özdeğerleri eşit olduğundan X ve X_{α} nün n özdeğerinin arasında bir benzerlik kurulmuş oldu. $\lambda_1, \lambda_2, \dots, \lambda_n$ ve $\lambda_{\alpha_1}, \lambda_{\alpha_2}, \dots, \lambda_{\alpha_n}$ -leri aşağıdaki gibi yeniden isimlendirelim:

$$\lambda_i - \lambda_i = m_i(\alpha_i - \alpha) \quad (3.165)$$

Burada (i) ve (ii) kümesinden özdeğerler için $m_i=0$ olur.

3.29 Üniter Ranklı Simetrik Matrisler

Bir önceki paragrafın sonucu aşağıdaki gibi tanımlanmış C -nin özdeğerleri için kullanılabilir:

$$C = A + B \quad (3.166)$$

Burada A ve B simetrik ve B -nin rankı üniterdir. Bu durumda aşağıdaki gibi bir R ortogonal matris vardır:

$$R^T B R = \begin{pmatrix} \rho & O \\ O & O \end{pmatrix} \quad (3.167)$$

ρ , B -nin sıfır-olmayan tek özdeğeridir. Eğer

$$R^T A R = \begin{pmatrix} \alpha & a^T \\ a & A_{n-1} \end{pmatrix} \quad (3.168)$$

yazılırsa bu durumda aşağıdaki gibi (n-1) mertebeden ortogonal bir S matrisi vardır:

$$S^T A_{n-1} S = \text{diag}(\alpha_i) \quad (3.169)$$

Eğer Q aşağıdaki gibi tanımlanırsa:

$$\begin{pmatrix} 1 & O \\ O & S \end{pmatrix}$$

$$Q = R \quad (3.170)$$

Q bir ortogonaldır ve

$$Q^T(A+B)Q \approx \begin{vmatrix} \alpha & b^T \\ b & \text{diag}(\alpha_i) \end{vmatrix} + \begin{vmatrix} \rho & 0 \\ 0 & 0 \end{vmatrix} \quad (3.171)$$

olur.

Burada $b=S^T a$ dır. A ve $(A+B)$ -nin özdeğerleri aşağıdaki matrislerin özdeğerleri olacaktır:

$$\begin{vmatrix} \alpha & b^T \\ b & \text{diag}(\alpha_i) \end{vmatrix} \quad \text{ve} \quad \begin{vmatrix} \alpha+\rho & b^T \\ b & \text{diag}(\alpha_i) \end{vmatrix} \quad (3.171)$$

Eğer bu özdeğerleri azalan mertebeden sıralanırsa λ_i ve $\lambda_i \square$ ile ifade edilirse aşağıdaki sonuca ulaşılır:

$$\begin{aligned} 0 < m_i < 1, \Sigma m_i = 1 & \quad \text{iken} \\ \lambda_i \square -\lambda_i = m_i \rho & \end{aligned} \quad (3.172)$$

Böylece B , A -ya eklendiğinde geri kalan özdeğerlerin tümü B -nin ρ özdeğeri ve sıfır arasında yayılacak bir doğrultuda kaydırılmış olur.

§40 in sonunda bulunan hatırlatmadan ve (3.171) de verilen A -nın dönüşümünü A -nın A_{n-1} temel minörün özdeğerleri, en azından zayıf anlamda A -nin özdeğerlerini ayırılır. Eğer A , k katlı λ_i özdeğerine sahip ise bu durumda A_{n-1} de $k-1$, k , $k+1$ katlılıklarında λ_i özdeğerine sahip olacaktır. (Young, 1971)

3.30 Özdeğerlerin ekstremal özellikleri

Özdeğerlerin minimax karakterizasyonu özelliği verilmeden önce, $x^T x = 1$ kısıtına sahip $x^T A x$ fonksiyonun maksimum değeri M -nin bulunması problemi ile ilgileneceğiz. M aşağıdaki ilişkiyi sağlar:

$$M = \max_{x \neq 0} (x^T A x / x^T x) \quad (3.173)$$

A simetrik bir matris olduğundan, aşağıdaki gibi ortogonal bir R matrisi vardır:

$$R^T A R = \text{diag}(\lambda_i) \quad (3.174)$$

Eğer yeni değişken y aşağıdaki gibi tanımlanırsa:

$$x = R y \quad ; \quad y = R^T x \quad (3.175)$$

aşağıdaki eşitlikler sağlanır:

$$x^T A x = y^T R^T A R y = y^T \cdot \text{diag}(\lambda_i) y = \sum_1^n \lambda_i y_i^2 \quad (3.176)$$

$$x^T x = y^T R^T R y = \sum_1^n y_i^2 \quad (3.177)$$

(3.155) x ve y arasında bire bir ilişki kurduğundan, $\sum_1^n y_i^2 = 1$ koşuluna sahip $\sum_1^n \lambda_i y_i^2$ -nin maksimum değeri orijinal problem ile eşitlenir. Eğer λ_i artmayan sırada olduğu kabul edilirse λ_1 maksimum değerdir. Bu $y=e_1$ eşitliğinde elde edilir. Karşılık gelen x vektörü r_1 olur. R -nin ilk kolonu, λ_1 e karşılık gelen A matrisinin özvektörüdür. (Eğer $\lambda_1 = \lambda_2 = \dots = \lambda_r \neq \lambda_{r+1}$, ise $e_1, e_2, \dots, e_r$ ile gerilmiş olan alt uzaydaki herhangi bir birim vektör y , λ_1 değerini verir.) Benzer olarak λ_n aynı koşul altında $x^T A x$ in minimum değeridir.

Şimdi, x in r_1 -e dik ilave koşuluna sahip, aynı maksimum problemini inceleyelim. Aşağıdaki ilişkiden

$$0 = r_1^T x = r_1^T R y = e_1^T R^T R y = e_1^T y \quad (3.178)$$

y üzerindeki koşul ilk bileşeni sıfıra zorlar. Böylece bu koşula sahip $x^T A x$ in maksimumu λ_2 dir. Bu $y = e^2$ ve $x=r_2$ için elde edilir. Benzer olarak $r_1^T x = r_2^T x = \dots = r_s^T x$ kısıtına sahip $x^T A x$ in maksimumu λ_{s+1} olarak bulunur. Bu $x=r_{s+1}$ için elde edilir.

Özdeğerlerin bu karakterizasyonunun zayıf tarafı her bir λ_s -nin bulunmasının $\lambda_1, \lambda_2, \dots, \lambda_{s-1}$ e karşılık gelen A nın özvektörü $r_1, r_2, \dots, r_{s-1}$ in bilgisi üzerine kurulu olmasıdır. (Young, 1971)

3.31 Özdeğerlerin Minimax Karakterizasyonu

Şimdi yukarıda bahsedilen dezavantaja sahip olmayan özdeğerlerin bir karakterizasyonunu inceleyelim.

Aşağıdaki koşullara sahip $x^T Ax$ -nin maksimumunu göz önüne alalım:

$$\begin{aligned} x^T x &= 1 \\ n_i^T &= 0 \quad n_i \neq 0 \quad (i = 1, 2, \dots, s; s < n) \end{aligned} \quad (3.179)$$

Burada p_i herhangi sıfır-olmayan s vektördür. Diğer bir değişle sadece s tane lineer ilişki sağlayan x in değerleri ile ilgileneceğiz. x -in tüm değerleri için aşağıdaki ifade yazılabilir:

$$\lambda_n < x^T Ax < \lambda_1 \quad (3.180)$$

Öyle ki $x^T Ax$ sınırlıdır ve maksimumu ise p_i -nin n s bileşeninin bir fonksiyonudur. Şimdi şu soruyu soralım: bütün mümkün p_i -lerin s vektörlerinin seçiminde maksimum değerler için kabul edilen minimum değer nedir?

(3.178) ilişkisi aşağıdaki şekle dönüşür:

$$\begin{aligned} y^T y &= 1 \\ q_i^T y &= 0, \quad q_i^T = p_i^T R \neq 0 \end{aligned} \quad (3.181)$$

$p_1, p_2, \dots, p_s$ in herhangi özel seçimlerini göz önüne alalım. Bu q_i kümesinin karşılığını verir Böylece n tane y_i s tane lineer homojen denklemi sağlar. Eğer aşağıdaki kısıtlama ilave edilirse:

$$y_{s+2} = y_{s+3} = \dots = y_n = 0 \quad (3.182)$$

n bilinmeyen $y_1, y_2, \dots, y_n$ içerisinde $n-1$ homojen denklem elde edilir ve en az bir sıfır-olmayan çözüm bulunur.

$$(y_1, y_2, \dots, y_{s+1}, 0, \dots, 0) \quad (3.183)$$

Burada $\sum_{i=1}^{s+1} y_i^2 = 1$ şeklinde normalize edilebilir. Böylece aşağıdaki ifadeyi elde edilir:

$$y^T \text{diag}(\lambda_i) y = \sum_{i=1}^{s+1} \lambda_i y_i^2 > \lambda_{s+1} \quad (3.184)$$

Bu p_i -nin nasıl seçilirse seçilsin bir y -nin ve aşağıdaki koşulu sağlayan bir x in daima olacağını gösterir:

$$x^T A x = y^T \text{diag}(\lambda_i) y > \lambda_{s+1} \quad (3.185)$$

ve böylece herhangi bir p_i seçimi için

$$\max(x^T A x) > \lambda_{s+1} \quad (3.186)$$

elde edilir. Bunun manası

$$\min \max x^T A x > \lambda_{s+1} \quad (3.187)$$

Eğer $p_i = \text{Re}_i$ olarak alınırsa $q_i^T = e_i^T$ olur ve (3.181) ilişkisi aşağıdaki hale dönüşür:

$$y_i = 0 \quad (i = 1, 2, \dots, s) \quad (3.188)$$

Böylece bu özel ilişkilere sahip her hangi bir y için aşağıdaki ifade elde edilir:

$$x^T A x = y^T \text{diag}(\lambda_i) y = \sum_{i=1}^n \lambda_i y_i^2 < \lambda_{s+1} \quad (3.189)$$

$$\max((x^T A x) < \lambda_{s+1} \quad (3.190)$$

(3.189) ve (3.190) ifadesi p_i -nin bu seçimleri ile aşağıdaki ifadeye ulaştırır:

$$\max(x^T A x) < \lambda_{s+1} \quad (3.191)$$

$$y_{s+1} = 1 \quad y_i = 0 \quad (i \neq s+1) \quad (3.192)$$

Böylece x , s lineer ilişkiye sahip olduğunda, $\min \max(x^T A x) \lambda_{s+1}$ olur. Bu sonuç Courant-Fischer teoremi olarak bilinir. Biz vurgulayalım ki minimax değerlerinin kabul edildiği bir x ve en az bir p_i kümesi olduğunu gösterdik.

Benzer bir yöntem ile aşağıda ki ifade ispatlanabilir:

$$x^T x = 1, \quad p_i^T x = 0 \quad (i=1,2,\dots,n-s) \text{ için} \quad (3.193)$$

$$\lambda_s = \max \min(x^T A x)$$

Bu durumda eğer herhangi s tane p_i vektörü varsa, genel olarak maksimumların minimizasyonlarını farklı p_i kümeleri oluşacaktır (p_i -lerin hepsi yada bir kısmı farklı olsun yada olmasın).

$$\lambda_{s+1} < \max(x^T A x) \quad (3.194)$$

$$x^T x = 1 \quad p_i^T x = 0 \quad (i=1,2,\dots,s)$$

3.32 İki Simetrik Matrisin Toplamının Özdeğerleri

Özdeğerlerin minimax karakterizasyonu aşağıdaki ifadeyi sağlayan A , B ve C simetrik matrislerinin özdeğerleri arasında bir ilişki kurmak için kullanılabılır:

$$C = A + B \quad (3.195)$$

A, B, C -nin özdeğerleri sırasıyla α_i , β_i ve γ_i olarak gösterelim ve bütün özdeğerler kümesi artmayan sırada dizilirse minimax teoreminden aşağıdaki ifade elde edilir:

$$\gamma_s = \min \max(x^T C x) \quad (3.196)$$

$$x^T x = 1 \quad p_i^T x = 0 \quad (i=1,2,\dots,s-1)$$

Böylece eğer p_i -lerin keyfi bir kümesi seçersek buna karşılık gelen bütün x_i -ler için

$$\gamma_s < \max(x^T C x) = \max(x^T A x + x^T B x) \quad (3.197)$$

elde ederiz. (Gould S.H. 1974)

Eğer R aşağıdaki gibi ortogonal bir matris ise

$$R^T A R = \text{diag}(\alpha_i) \quad (3.198)$$

ve $p_i = \text{Re}_i$ olarak alınırsa sağlanması istenen şartlar aşağıdaki gibidir:

$$0 = p_i^T x = e_i^T y \quad (i = 1, 2, \dots, s-1) \quad (3.199)$$

p_i -nin bu seçimi ile y -nin ilk $(s-1)$ bileşeni sıfır olur ve (3.197) den aşağıdaki ifade elde edilir:

$$\gamma_s < \max (x^T A x + x^T B x) = \max \left(\sum_{i=s}^n \alpha_i y_i^2 + x^T B x \right) \quad (3.200)$$

Bununla birlikte

$$\sum_{i=s}^n \alpha_i y_i^2 < \alpha_s \quad (3.201)$$

olup keyfi bir x_i için

$$x^T B x < \beta_1 \quad (3.202)$$

olur. p_i -lerin bu seçimine karşılık parantez içindeki ifade $\alpha_s + \beta_1$ den büyük olamaz. Böylece maksimumu $\alpha_s + \beta_1$ den büyük değildir. Buradan aşağıdaki ifadeye yazılır:

$$\gamma_s < \alpha_s + \beta_1 \quad (3.203)$$

$A = C + (-B)$ ve $(-B)$ -nin özdeğerleri artmayan sıradan $-\beta_n, -\beta_{n-1}, \dots, -\beta_1$ olduğundan biraz önce ispat edilen sonuç aşağıdaki ifadeyi verir:

$$\begin{aligned} \alpha_s &< \gamma_s + (-\beta_n) \\ \gamma_s &> \alpha_s + \beta_n \end{aligned} \quad \text{veya} \quad (3.204)$$

(3.203) ve (3.204) ifadesi B, A -ya eklendiğinde özdeğerlerin tümü, B -nin özdeğerlerinin en büyüğü ile en küçüğü arasında yayılan miktarca değişir. Bu rankı unity olan B matrisinin analitik yollarla §41 de ispatladığımız sonuçlarla aynı olur. Not edelim ki küçük perturbelerin etkileri ve sonuçları burada incelenmemiştir.

3.33 Pratik Uygulamalar

Son bölüm pratikte en kullanışlı konulardan biridir. Genellikle B küçük elemanlardan ve bu elemanlar üstten sınırlı olacaktır. Örnek olarak aşağıdaki koşulu inceleyelim:

$$|b_{ij}| < \varepsilon \quad (3.205)$$

$$-n\varepsilon < \beta_n < \beta_1 < n\varepsilon \quad (3.206)$$

$$|\gamma_r - \alpha_r| < n\varepsilon \quad (3.207)$$

Bu simetrik-olmayan perturbasyonlar için §37 nin sonucun dan daha kuvvetlidir. Çünkü, α_i , β , ve γ_i nin dağılımında herhangi bir kısıtlama yoktur. Önceden 45.3 ün α_r özdeğerlerinin $2n\varepsilon$ -dan daha fazla ayrık oldukları ispatlanmıştı.

3.34 Minimax Prensibinin Diğer Uygulamaları

§44 de elde edilen sonuç aşağıdaki sonucu vermesi için genelleştirilebilir:

$$\gamma_{r+s-1} < \alpha_r + \beta_s \quad (r+s-1 < n) \quad (3.208)$$

ispat aşağıdaki gibidir:

Aşağıdaki koşul için en az bir p_i -ler kümesi vardır: gibidir

$$\begin{aligned} p_i^T x &= 0 \\ \max(x^T A x) &= \alpha_r \end{aligned} \quad (i=1,2,\dots,r-1) \text{ koşuluna uyan} \quad (3.209)$$

Aynı zamanda aşağıdaki koşul içinde bir q_i kümesi vardır:

$$\begin{aligned} q_i^T x &= 0 \\ \max(x^T B x) &= \beta_s \end{aligned} \quad (i=1,2,\dots,s-1) \text{ koşuluna uyan} \quad (3.210)$$

Şimdi $p_i^T x = 0$ ve $q_i^T x = 0$ yi sağlayan x kümesini göz önüne alalım. Burada sıfır olmayan bazı x -ler mevcuttur. Çünkü toplam denklem sayısı, (3.208) den dolayı $n-1$ den büyük olmayan $r+s-2$ kadardır. Bu herhangi bir x için aşağıdaki ifadeler yazılabilir:

$$x^T C x = x^T A x + x^T B x < \alpha_r + \beta_s \quad (3.211)$$

$$\max(x^T C x) < \alpha_r + \beta_s \quad (3.212)$$

Burada tüm x , (3.209) ve (3.210) $r+s-2$ lineer denklemin kümesini sağlar.

$$\gamma_{r+s-1} = \min \max(x^T C x) < \alpha_r + \beta_s \quad (3.213)$$

3.35 Ayırma Teoremi

Minimax teoreminin bir uygulama olarak A matrisinin baş temel minörlerinin (*leading prensible minör*) A_{n-1} -in $\lambda_1, \lambda_2, \dots, \lambda_{n-1}$ özdeğerleri A -nın $\lambda_1, \lambda_2, \dots, \lambda_n$ özdeğerlerini ayırır. (41- de bu sonuç analitik metotlarla en küçük temel minörü için gösterilmiştir. Bunun öbürleri içinde doğru olduğu permutasyon matrislerine benzerlik dönüşümleri uygulanarak gösterilebilir.)

$(n-1)$ mertebeden birim vektörlerinin hepsi için x , $x^T A_{n-1} x$ değerleri n . mertebeden kümesi, $x_n=0$ koşuluna bağlı olur. Böylece birim vektörler için $x^T A_n x$ değerleri ile aynıdır.

$$\lambda \square_s = \min \max(x^T A_n x), \quad x^T x = 1 \quad (3.214)$$

$$x_n = 0 \quad p_i^T x = 0 \quad (i = 1, 2, \dots, s-1)$$

Şimdi bu maksimum değer bazı p_i kümeleri için elde edilecektir. Bu p_i kümeleri için, $\lambda \square_s$, (47.1) in ikinci satırındaki s lineer ilişkiye sahip $x^T A_n x$ in maksimum değeridir. Bununla birlikte λ_{s+1} , herhangi s ilişkiye sahip maksimum tarafından minimum olarak alınan değerdir. Böylece

$$\lambda_{s+1} < \lambda \square_s \quad (3.215)$$

Özel olarak şimdi p_i vektörler kümesinin $s-1$ tanesini göz önüne alalım. Birim vektör verilen x lineer koşulları altında $x^T A_n x$ -nin maksimum değerini $f_n(p_i)$ ile gösterelim. $x_n = 0$ ekstra kısıtına sahip maksimumu da $f_{n-1}(p_i)$ ile ifade edelim. Bu durumda aşağıdaki ifadeye ulaşırız:

$$f_{n-1}(p_i) < f_n(p_i) \quad (3.216)$$

ve böylece

$$\min f_{n-1}(p_i) < \min f_n(p_i) \quad (3.217)$$

$$\lambda_s \square < \lambda_s \tag{3.218}$$

olur.

Son altı konunun tüm teorileri derhal Hermitian matrisine genişletilebilir. Sonuçlar ve ispatlar T -nin yerine H konularak elde edilir.



KAYNAKLAR

Aktaş Z., Öncül H. Ural S. (1984), “Sayısal Çözümleme”, ODTÜ Yayınları, Ankara.

Bronson R. Çeviren Hacısalihoğlu H.H. (1984), “Matris İşlemleri”, Schaum serisi – Nobel Yayıncılık, İstanbul.

Seymor L. Çeviren Sadık Dost Yayınevi (1978), “Lineer Cebir”, Schaum Serisi – Güven Kitapevi, Ankara.

Wilkinson J.H. (1965), “The Algebrik Eigenvalue Problem”, Clarendon Press, London.

Gould S.H. (1974), “Variational Methods for Eigenvalue Problems”, University of Toronto Press, London.

Wilkinson J.H. (1963), “Rounding Errors in Algebraic Processes”, Englewood Cliffs, N.J., Prentice-Hall, New York.

Wilkinson J.H. (1967), “The Algebraic Eigenvalue Problem”, Oxford University Press, New York.

Wilkinson J.H. and c. Rheinsch (1971), “Handbook for Automatic Computation II:Lineer Algebra”, Springer-Verlag, New York.

Young D.M. (1971), “Iterative Solution of Large Linear Systems”, Academic Press, New York.

Prager W.H. (1988), “Applied Numerical Linear Algebra”, Englewood Cliffs, Prentice-Hall, New York.

ÖZGEÇMİŞ

Doğum Tarihi 28.04.1973

Doğum yeri Akşehir

Lise 1987-1990 Bahçelievler Lisesi

Lisans 1991-1995 Yıldız Teknik Üniversitesi
Matematik Mühendisliği

Yüksek Lisans 1995-

