Lecture V

- The higher dimensional case of general real symmetric matrices

- Extension and Applications

The General Case

We already proved the result with $N = 2$.

By induction, let us assume that for each integer $1 \le k \le N$, we can find an orthogonal matrix O_k which reduces a real symmetric matrix $A_k = \left(a_{ij}\right)_{k\times k}$ to the diagonal form:

$$O_k^T A_k O_k = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_k \end{pmatrix}$$

The General Case: Goal

We want to find an orthogonal matrix O_{N+1} of order $N+1$, which reduces a real symmetric matrix $A_{N+1} = \left(a_{ij}\right)_{(N+1)\times(N+1)}$ to the diagonal form:

$$O_{N+1}^T A_{N+1} O_{N+1} = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_{N+1} \end{pmatrix}$$

Systematic Procedure

Let us name the rows of $A_{N+1} = \left(a_{ij}\right)_{(N+1)\times(N+1)}$ as:

$$A_{N+1} \doteq \begin{pmatrix} a^1 \\ a^2 \\ \vdots \\ a^{N+1} \end{pmatrix}$$

and take an eigenvalue λ_1 and its associated normalized eigenvector x^1.

Systematic Procedure (cont'd)

By means of the Gram-Schmidt orthogonalization process, we can form an orthogonal matrix O_1 whose first column is the given $x^1 = y^1$:

$$O_1 = \left(y^1,\ y^2,\ \ldots,\ y^{N+1}\right)$$

Then, as shown in Case $N = 2$, it holds:

$$O_1^T A_{N+1} O_1 = \begin{pmatrix} \lambda_1 & b_{12} & \cdots & b_{1,N+1} \\ 0 & & & \\ \vdots & & A_N & \\ 0 & & & \end{pmatrix},\quad A_N \in \mathbb{R}^{N\times N}$$

Exercise

Can you prove the above identity?

Answer

Carrying out the multiplication, we have

$$A_{N+1}O_1 = \begin{pmatrix} \langle a^1, x^1 \rangle & \dots & \langle a^1, x^{N+1} \rangle \\ \vdots & \ddots & \vdots \\ \langle a^{N+1}, x^1 \rangle & \cdots & \langle a^{N+1}, x^{N+1} \rangle \end{pmatrix}$$

$$= \begin{pmatrix} \lambda_1 x_{11} & \langle a^1, x^2 \rangle & \dots & \langle a^1, x^{N+1} \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_1 x_{1,N+1} & \langle a^{N+1}, x^2 \rangle & \cdots & \langle a^{N+1}, x^{N+1} \rangle \end{pmatrix}$$

Answer (cont'd)

Since O_1 is an orthogonal matrix, it follows that

$$O_1^T A_{N+1} O_1 = \begin{pmatrix} \lambda_1 & b_{12} & \cdots & b_{1,N+1} \\ 0 & & & \\ \vdots & & A_N & \\ 0 & & & \end{pmatrix}, \quad A_N \in \mathbb{R}^{N\times N}$$

Comment 1

Furthermore, since $O_1^T A_{N+1} O_1$ must be symmetric, we have $b_{1j} = 0,\ j = 2, \ldots, N+1,$ and A_N is symmetric. Thus,

$$O_1^T A_{N+1} O_1 = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & A_N & \\ 0 & & & \end{pmatrix}, \quad A_N = A_N^T.$$

Comment 2

Given the identity

$$O_1^T A_{N+1} O_1 = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & A_N & \\ 0 & & & \end{pmatrix}, \quad A_N = A_N^T$$

we conclude that the eigenvalues of A_N *must* be $\lambda_2,\ \lambda_3,\ \ldots,\ \lambda_{N+1},$ the remaining eigenvalues of A_{N+1}.

Systematic Procedure (cont'd)

By induction, there is an orthogonal matrix O_N which reduces A_N to diagonal form. Form the $(N+1)$-dimensional matrix

$$S_{N+1} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & O_N & \\ 0 & & & \end{pmatrix}$$

Clearly, S_{N+1} is also orthogonal, i.e., $S_{N+1}^T S_{N+1} = I$.

Systematic Procedure (cont'd)

It can be directly checked that

$$S_{N+1}^{T}\left(O_1^T A_{N+1} O_1\right) S_{N+1} = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_{N+1} \end{pmatrix}$$

or, equivalently,

$$O_{N+1}^{T} A_{N+1} O_{N+1} = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_{N+1} \end{pmatrix}$$

with $O_{N+1} = O_1 S_{N+1}$.

Formal Statement of the Main Result

Let $A \in \mathbb{R}^{n\times n}$ be a real symmetric matrix. Then, it may be transformed into a diagonal form by using an orthogonal matrix O so that

$$O^T AO = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix}$$

where $\{\lambda_i\}_{i=1}^{n}$ are the eigenvalues of A.

Test for Positive Definiteness

A necessary and sufficient condition for a real symmetric matrix A to be *positive definite* is that all eigenvalues of A are positive.

Indeed,

Recall that a real matrix A is positive definite if

$$x^T A x > 0,\ \forall x \in \mathbb{R}^n, x \neq 0.$$

Then,

$$x^T A x = x^T O \Lambda O^T x, \quad \text{where } \Lambda = diag(\lambda_i)$$

$$= y^T \Lambda y, \qquad \text{where } y = O^T x = (y_i)_{n\times 1}$$

$$= \sum_{i=1}^{n} \lambda_i y_i^2$$

So, the equivalence property follows readily.

Repeated Eigenvalues

As shown previously, if a matrix A has *distinct* eigenvalues, then its associated eigenvectors are linearly independent.

Questions :

- What if A has a repeated eigenvalue λ_1 of (algebraic) multiplicity k?
- Are there always k linearly independent eigenvectors?

Comment

As shown in Lecture IV, the answer is generally negative for a real matrix which is *not* be symmetric.

However, for a real symmetric matrix, we can always find k linearly independent eigenvectors for *any* repeated eigenvalue of multiplicity k.

Indeed,

$\exists$ an orthogonal matrix O such that

$$AO = O\begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix}$$

where $\lambda_1 = \cdots = \lambda_k,\ \lambda_i \neq \lambda_1, \forall i = k+1,\ldots,n.$

$Now,$ the first k columns of O are of course linearly independent, and are eigenvectors associated with λ_1.

In addition,

Any other eigenvector y associated with λ_1 *is* linear combination of these k vectors.

In fact, $y = \sum_{i=1}^{n} c_i x^i$, with x^i the ith column of O.

$c_i = 0, \forall i = k+1, \ \ldots, \ n$ because these eigenvectors x_i are orthogonal with x_j, y, $1 \le j \le k$.

(*see* Lecture IV)

Special Case of Cayley-Hamilton Theorem

As a direct application of the diagonal canonical form, we have

Any real symmetric matrix satisfies its own characteristic equation: $\rho_A(A) = 0,$

where $\rho_A(\lambda) = \det(\lambda I - A) = \lambda^n + \sum_{i=1}^{n} \alpha_i \lambda^{n-i},$

$$\rho_A(A) := A^n + \sum_{i=1}^{n} \alpha_i A^{n-i}, \text{ with } A^0 = I.$$

Application:
Solving Differential Equations

Solving for the solutions of

$$\dot{x} = Ax$$

boils down to an easier problem for

$$\dot{y} = A_c y$$

where A_c is a canonical form of A under a nonsingular transformation

$$P^{-1}: \ x \to \ y = P^{-1}x$$

so that $A_c = P^{-1}AP$ and $x = Py$.

Exercise 1

Solve the following initial-value problem:

$$\dot{x}(t) = \begin{pmatrix} -1 & 3 \\ 3 & -1 \end{pmatrix} x(t), \quad x(0) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

i.e,

$$x(t) = ?, \ \forall t \geq 0.$$

Exercise 2: Extension to Difference Equations

Find an explicit expression for $x_n,\ n = 0,1,...,$ given that

$$x_0 = -1,\ \ x_1 = 2 \text{ and}$$

$$x_n = ax_{n-1} + x_{n-2},\ \ n = 2,3,...$$

Hint

Rewrite the second-order difference equation as:

$$\begin{bmatrix} x_{n-1} \\ x_n \end{bmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & a \end{pmatrix} \begin{bmatrix} x_{n-2} \\ x_{n-1} \end{bmatrix}, \; n = 2, 3, \ldots$$

Equivalently,

$$\xi_n = A\xi_{n-1}, \; n = 1, 2, \ldots$$

with

$$\xi_n := \begin{bmatrix} x_{n-1} \\ x_n \end{bmatrix}.$$

Another Extension

Question : When does there exist an orthogonal matrix O which simultaneously reduces two real symmetric matrices A, B to diagonal form?

Motivational Problem

Solve the 2nd-order differential equation:

$$A\ddot{x}(t) + Bx(t) = 0, \; x \in \mathbb{R}^n$$

where $A, \; B \in \mathbb{R}^{n \times n}$ are symmetric matrices.

Note that such equations often occur in mass-spring problems.

Basic Result

A necessary and sufficient condition for the existence of an orthogonal matrix O such that

$$\begin{cases} O^T AO = diag(\lambda_i) \\ O^T BO = diag(\mu_i) \end{cases}$$

is that A and B commute, i.e. $AB = BA$.

Note: See Section 1.3 of the 2013 textbook of Horn and Johnson for extensions to more than 2 matrices.

Proof of the Necessity

Clearly,

$A = O^T diag(\lambda_i) O$ and $B = O^T diag(\mu_i) O$

commute, because O is orthogonal.

Sufficiency: Sketch of Proof

Case 1: Either A or B has distinct eigenvalues.

Assume that A has distinct eigenvalues. Then,

$$Ax = \lambda x \Rightarrow A\left(Bx\right) = B\left(Ax\right) = \lambda\left(Bx\right).$$

So, Bx, if nonzero, is an eigenvector too,

for the same eigenvalue λ. In other words,

Bx is a multiple of x. As a result,

$$Bx^i = \mu_i x^i, \text{ for each pair } \left(\lambda_i,\ x^i\right).$$

This equality of course also holds if *Bx*=0.

Sufficiency: Sketch of Proof

Case 1 (cont'd)

Now, we observe that A, B have the same eigenvectors x^i, $1 \le i \le n$.

Thus, we can define the orthogonal transformation matrix O as follows:

$$O = \left(x^1, x^2, \cdots, x^n \right)$$

Sufficiency: Sketch of Proof

Case 2: λ_1 repeats k times associated with (linearly independent/orthonormal) eigenvectors $x^1, \cdots, x^k$.

Then, using previous computation, we have

$$Bx^i = \sum_{j=1}^{k} c_{ij} x^j, \quad i = 1, 2, \cdots, k.$$

In addition, $\left\langle x^j, Bx^i \right\rangle = c_{ij} = \left\langle Bx^j, x^i \right\rangle = c_{ji}$.

Now, consider the linear combination $\sum_{i=1}^{k} a_i x^i$.

Sufficiency: Sketch of Proof

Case 2 (cont'd) : We have

$$B(\sum_{i=1}^{k} a_i x^i) = \sum_{i=1}^{k} a_i \left(\sum_{j=1}^{k} c_{ij} x^j \right) = \sum_{j=1}^{k} \left(\sum_{i=1}^{k} c_{ij} a_i \right) x^j .$$

Thus, if we choose a_i so that

$$\sum_{i=1}^{k} c_{ij} a_i = r_1 a_j ,\ j = 1, 2, \cdots, k, \Leftrightarrow \left(r_1 I - C \right) a = 0$$

then we have

$$B\left(\sum_{i=1}^{k} a_i x^i \right) = r_1 \left(\sum_{i=1}^{k} a_i x^i \right)$$

Sufficiency: Sketch of Proof

Case 2 (cont'd) : $B\left(\sum_{i=1}^{k} a_i x^i\right) = r_1\left(\sum_{i=1}^{k} a_i x^i\right)$

implies r_1 is an eigenvalue of B, associated

with eigenvector $\sum_{i=1}^{k} a_i x^i$.

On the other hand, r_1 is an eigenvalue of $C = \left(c_{ij}\right)$

associated with eigenvector $a = col\left(a_i\right)$.

Sufficiency: Sketch of Proof

Case 2 (end):

If T_k is a k-dim. orthogonal transformation reducing C into a diagonal form, then

$$\left[z^1 \cdots z^k \right] = T_k \left[x^1 \cdots x^k \right]$$ is an orthonormal set

with each z^i being common eigenvector for both A and B. $\left(\textit{left as an exercise} \right)$

Exercise 3

Show how to transform the following matrix into a canonical diagonal form, by means of an orthogonal matrix:

$$M = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Normal Matrices:

A generalization of real symmetric matrices

Definition: A square matrix A is said to be normal, if $AA^* = A^*A$.

- If A is normal and α is a scalar, then αA also is normal.
- If A is normal and $B \sim A$, then B also is normal.
- Every unitary matrix is normal.
- Every real symmetric or skew-symmetric matrix is normal.
- Every Hermitian or skew-Hermitian matrix is normal.

Exercise 4

Let $a,\ b$ be constants. Show that $\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$ is normal

and has eigenvalues $a \pm ib$.

Exercise 5

A matrix $A \in \mathbb{C}^{n\times n}$ is conjugate normal if $AA^* = \overline{A^* A}$.

Show that a block-upper-triangular matrix

$$A = \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix} \text{ is conjugate normal}$$

if and only if its diagonal blocks A_{11}, A_{22} are conjugate normal, and $A_{12} = 0$. In particular, an upper triangular matrix is conjugate normal iff it is diagonal.

See the text (2nd ed.) by Horn-Johnson, p. 268.

Homework #5

1. Using the Gram-Schmidt process find a set of mutually orthonormal vectors u^1, u^2, u^3, based on:

$$x^1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix}, \quad x^2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ -1 \end{bmatrix}, \quad x^3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \end{bmatrix}.$$