Lecture IV

Key Issues:

Real symmetric matrices and canonical forms

Symmetric Matrices

Recall that a symmetric matrix $A = \left(a_{ij}\right)$ satisfies: $\quad a_{ij} = a_{ji}, \quad \forall 1 \le i, j \le n.$

It is a real symmetric matrix if, additionally, all a_{ij}'s are real.

Notation:

$$A = A^T, \; A \in \mathbb{R}^{n \times n}.$$

Fact 1 about Symmetric Matrices

The eigenvalues of a real symmetric matrix are always real.

Proof of Fact 1

By contradiction, assume that a real symmetric A has a complex eigenvalue, say, λ. Then,

$$Ax = \lambda x \quad \Rightarrow \quad A\bar{x} = \bar{\lambda}\bar{x}, \text{ or } \bar{x}^T A = \bar{\lambda}\bar{x}^T.$$

because A is symmetric. This further implies that

$$\bar{x}^T Ax = \lambda \bar{x}^T x \;\text{ and }\; \bar{x}^T Ax = \bar{\lambda}\bar{x}^T x.$$

$$\Rightarrow \; 0 = \left(\lambda - \bar{\lambda}\right) x^T \bar{x}$$

$$\Rightarrow \left(\lambda - \bar{\lambda}\right) = 0, \;\text{ a contradiction.}$$

Fact 2 about Symmetric Matrices

For any real symmetric matrix, its eigenvectors associated with <u>distinct</u> eigenvalues are orthogonal.

Remarks:

- Two vectors $x,\ y \in \mathbb{R}^n$ are orthogonal if $x^T y = 0$.
- *Orthogonal vectors are linearly independent.*

Proof of Fact 2

For a real symmetric A, consider a pair of eigenvectors (x, y) associated with distinct eigenvalues λ, μ, i.e.,

$$Ax = \lambda x \quad \text{and} \quad Ay = \mu y.$$

This further implies that

$$y^T Ax = \lambda y^T x \quad \text{and} \quad x^T Ay = \mu x^T y.$$

$$A \text{ symmetric} \;\Rightarrow\; y^T Ax = \left(y^T Ax\right)^T = x^T Ay$$

$$\Rightarrow\; 0 = (\lambda - \mu)\, x^T y$$

$$\Rightarrow x^T y = 0, \quad \text{as wished.}$$

Canonical Form – First Pass

Consider a real symmetric matrix $A \in \mathbb{R}^{n\times n}$,

with *distinct* (real, by Fact 1) eigenvalues $\{\lambda_i\}_{i=1}^{n}$.

Then, there is an orthogonal matrix O, i.e., $O^T O = I$,

such that

$$O^T A O = diag(\lambda_i) \triangleq \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix}.$$

Constructive Proof

For each eigenvalue λ_i, take an eigenvector x^i, which has unit norm, i.e., $\left\|x^i\right\| = \sqrt{(x^i)^T x^i} = 1.$

Define a matrix O as:

$$O \triangleq \left(x^1, \; \ldots, \; x^n\right) \in \mathbb{R}^{n \times n}$$

Then, $$O^T = \begin{bmatrix} \left(x^1\right)^T \\ \vdots \\ \left(x^n\right)^T \end{bmatrix} \in \mathbb{R}^{n \times n}$$

Constructive Proof (cont'd)

It is directly checked using Fact 2 that $O^T O = I$, i.e., O is an orthogonal matrix.

In addition, $O^T A O = diag(\lambda_i) \triangleq \Lambda.$

Exercise

Compute the eigenvalues λ_1, λ_2 of

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}$$

and find a transformation matrix O s.t.

$$O^T A O = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}.$$

What if *A* is *not necessarily* symmetric

Answer:

Yes! As long as the eigenvalues are mutually distinct, there is a nonsingular matrix P such that

$$P^{-1}AP = diag(\lambda_i), \text{ denoted } A \sim diag(\lambda_i).$$

$However$, this P may not be orthogonal.

Remark: A non symmetric matrix may not be diagonanizable.

Show that the following matrix

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

is not diagonalizable.

Comment

Two similar matrices have the same eigenvalues. So, if

$A \sim diag(\lambda_i)$, i.e., $P^{-1}AP = diag(\lambda_i)$,

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

However, the converse is *not* true.

Exercise

Two matrices having the same eigenvalues may <u>not</u> be similar.

Show that the following matrix

$$A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

is *not* diagonizable. In other words, it is *not* similar to

$$\Lambda = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Question (Necessity and Sufficiency):

When is a matrix similar to a diagonal matrix?

Necessary and Sufficient Condition for the Canonical Diagonal Form

- An $n \times n$ matrix A is similar to a diagonal matrix iff A has n linearly independent eigenvectors.

- When A has n distinct eigenvalues, it is similar to a diagonal matrix.

Proof

First, note that Statement 2 follows from Statement 1 and a result proved previously.

Assume A is similar to a diagonal matrix $\Lambda = diag\left\{\lambda_i\right\}$.

Then, $\exists P$ nonsingular s.t. $P^{-1}AP = \Lambda$.

Let $P = \left(p^1\ p^2\ \ \dots\ p^n\right)$, with $\left\{p^i\right\}$ linearly independent.

$$AP = P\Lambda \quad \Rightarrow \quad Ap^i = \lambda_i p^i,\ \forall i = 1, 2, \dots., n$$

implying that p^i is an eigenvector for eigenvalue λ_i.

Proof (cont'd)

Conversely, assume that A has n linearly independent eigenvectors $\left\{ p^i \right\}_{i=1}^{n}$, i.e., $Ap^i = \lambda_i p^i$.

Then, $P = \left(p^1 \; p^2 \; \ldots \; p^n \right)$ is nonsingular and satisfies (by direct computation) that

$$P^{-1}AP = \Lambda.$$

Comment

From the proof of Part 1, it follows that the following is an equivalent condition for diagonalization of A:

$$\dim N\left(A-\lambda_1 I\right)+\cdots+\dim N\left(A-\lambda_k I\right)=n$$

where

$\lambda_1,\ldots,\lambda_k$ are the distinct eigenvalues of A, $k \le n$.

An Example

Bring the matrix $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$

into a diagonal form.

The eigenvalues of A are $\lambda_1 = -j,\ \lambda_2 = j$.

As it can be directly checked, the associated independent eigenvectors are:

$$c^1 = \begin{bmatrix} 1 \\ j \end{bmatrix} \quad \text{and} \quad c^2 = \begin{bmatrix} 1 \\ -j \end{bmatrix}.$$

Then, $P = \left(c^1 \quad c^2 \right)$, implying that

$$P^{-1} A P = \begin{pmatrix} -j & 0 \\ 0 & j \end{pmatrix}.$$

Diagonalizable Matrix

A matrix is said to be "diagonalizable", if it is similar to a diagonal matrix.

Are the following statements true or false:

(1) Two diagonalizable matrices always commute.

(2) The block-diagonal matrix

$$B = block\ diag\left\{B_i\right\},\ B_i \in \mathbb{R}^{n_i \times n_i}$$

is diagonalizable if and only if each B_i is diagonalizable.

Let's stop for a short review…

- **Review of the results on nontrivial solutions to homogeneous equations:**

$$\boxed{Ax = 0, \quad A \in \mathbb{R}^{m\times n}, \quad x \in \mathbb{R}^n.}$$

- **How about inhomogeneous systems?**

A Quiz?

- Any set of vectors $x^i \in \mathbb{R}^n$, with $1 \le i \le N$, are always linearly dependent, if $N > n$.

Real and Symmetric Matrices

- **The eigenvalues are always real.**
- **Eigenvectors associated with distinct eigenvalues are always orthogonal.**
- **Any matrix *with no repeated eigenvalues* is diagonalizable.**
- **How to transform a real and symmetric matrix into a diagonal form?**

A General Result for General Symmetric Matrices

For any real and symmetric matrix $A \in \mathbb{R}^{n\times n}$,

there always exists an orthogonal matrix, say O,

$O^T O = I$, such that

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

Special case: A Trivial Example

$$A = \begin{pmatrix} a_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & a_n \end{pmatrix}$$

Clearly, the identity matrix is an orthogonal matrix.

Before proving this general and fundamental result, let us introduce some useful tools.

The Gram-Schmidt Orthogonalization Process

Question :

How to generate a set of mutually orthogonal vectors $\left\{ y^i \right\}_{i=1}^{N}$ *successively*,
from a set of N real linearly independent n-dimensional vectors $\left\{ x^i \right\}_{i=1}^{N}$?

Let us start with a set of real-valued vectors $\left\{x^i\right\}_{i=1}^N$. Here is the systematic procedure.

First,

$$y^1 := x^1$$

$$y^2 := x^2 + a_{11}x^1$$

where a_{11} is a scalar to be determined so that *inner* product $\left\langle y^1, y^2 \right\rangle \triangleq \left(y^1\right)^T y^2 = 0$

$$\Leftrightarrow \left\langle x^1, x^2 + a_{11}x^1 \right\rangle = 0.$$

$$\left\langle x^1, x^2 + a_{11} x^1 \right\rangle = 0 \quad \Leftrightarrow \quad a_{11} := -\left\langle x^1, x^2 \right\rangle / \left\langle x^1, x^1 \right\rangle$$

with $D_1 := \left\langle x^1, x^1 \right\rangle > 0.$

Next, construct y^3 as:

$$y^3 := x^3 + a_{21} x^1 + a_{22} x^2$$

where a_{21}, a_{22} are scalars to be determined s.t.

$$\left\langle y^3, y^1 \right\rangle = 0, \quad \left\langle y^3, y^2 \right\rangle = 0$$

$\Leftrightarrow$

$$\left\langle y^3, x^1 \right\rangle = 0, \quad \left\langle y^3, x^2 \right\rangle = 0.$$

$$\langle y^3, x^1 \rangle = 0, \quad \langle y^3, x^2 \rangle = 0$$

$\Leftrightarrow$

$$\begin{cases} \langle x^3, x^1 \rangle + a_{21} \langle x^1, x^1 \rangle + a_{22} \langle x^2, x^1 \rangle = 0 \\ \langle x^3, x^2 \rangle + a_{21} \langle x^1, x^2 \rangle + a_{22} \langle x^2, x^2 \rangle = 0 \end{cases}$$

which has a (unique) solution $a_{21},\ a_{22}$ if

$$D_2 := \det \begin{pmatrix} \langle x^1, x^1 \rangle & \langle x^1, x^2 \rangle \\ \langle x^2, x^1 \rangle & \langle x^2, x^2 \rangle \end{pmatrix} \neq 0.$$

By contradiction, assume that

$$D_2 := \det\begin{pmatrix} \langle x^1, x^1 \rangle & \langle x^1, x^2 \rangle \\ \langle x^2, x^1 \rangle & \langle x^2, x^2 \rangle \end{pmatrix} = 0$$

Then, there are two scalars r_1, s_1, not both 0, such that

$$r_1 \langle x^1, x^1 \rangle + s_1 \langle x^1, x^2 \rangle = 0$$

$$r_1 \langle x^2, x^1 \rangle + s_1 \langle x^2, x^2 \rangle = 0$$

$\Rightarrow$

$$\langle x^1, r_1 x^1 + s_1 x^2 \rangle = 0, \quad \langle x^2, r_1 x^1 + s_1 x^2 \rangle = 0$$

$$\left\langle x^1, r_1 x^1 + s_1 x^2 \right\rangle = 0, \quad \left\langle x^2, r_1 x^1 + s_1 x^2 \right\rangle = 0$$

$$\Rightarrow$$

$$\left\langle r_1 x^1 + s_1 x^2, r_1 x^1 + s_1 x^2 \right\rangle = 0$$

$$\Rightarrow \quad r_1 x^1 + s_1 x^2 = 0.$$

Contradiction with x^1, x^2 being linearly independent. Thus,

$$D_2 := \det \begin{pmatrix} \left\langle x^1, x^1 \right\rangle & \left\langle x^1, x^2 \right\rangle \\ \left\langle x^2, x^1 \right\rangle & \left\langle x^2, x^2 \right\rangle \end{pmatrix} \neq 0.$$

So, we have obtained three mutually orthogonal vectors:

$$y^1 := x^1$$

$$y^2 := x^2 + a_{11} x^1$$

$$y^3 := x^3 + a_{21} x^1 + a_{22} x^2$$

Continuing this process, we can find other mutually orthogonal vectors:

$$y^i := x^i + \sum_{k=1}^{i-1} a_{(i-1)k} x^k$$

with the scalars $a_{(i-1)k}$ chosen to achieve the mutual orthogonality condition:

$$\langle y^i, y^j \rangle = 0 \quad \forall i \neq j,$$

or equivalently, $\langle y^i, x^j \rangle = 0, \ \forall 1 \leq j \leq i-1.$

Othonormal Vectors

They are defined as follows:

$$u^i := y^i / \left\| y^i \right\|, \quad i = 1,\ 2,\ \ldots,\ N.$$

It is easy to show that, if $n = N$,

$$O = \left(u^1,\ u^2,\ \ldots,\ u^N \right)$$

is an orthogonal matrix.

An Example

Consider the linearly independent vectors:

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

By means of the Gram-Schmidt process,

find a set of orthonormal vectors $u^1,\ u^2$.

Exercise

Show that if $\{v_1, \cdots, v_k\}$ is a set of k linearly independent vectors in $\mathbb{R}^n$, then there exists an invertible upper triangular matrix $T \in \mathbb{R}^{k \times k}$ such that the matrix $U = VT$ has orthonormal columns.

Comment

During the Gram-Schmidt process, we proved that the determinants D_k, called *Gramians*, are nonzero. Indeed, we can prove that

$$D_k = \det\left(\left\langle x^i, x^j \right\rangle\right) > 0, \quad 1 \le k \le N,$$

for any set of linearly independent vectors $\left\{x^i\right\}_{i=1}^{k}$.

Indeed,

Leading principle minor

Each Gramian $D_k = \det\left(\left\langle x^i, x^j \right\rangle\right)$ is associated with

a positive-definite quadratic form:

$$Q(u) = \left\langle \sum_{i=1}^{k} u_i x^i,\ \sum_{j=1}^{k} u_j x^j \right\rangle$$

$$= \sum_{i,j=1}^{k} \left\langle x^i, x^j \right\rangle u_i u_j$$

Q positive definite in $u \doteq (u_1, \ldots, u_k) \in \mathbb{R}^k$.

$\Leftrightarrow\ Q(u) \geq 0,$ where equality holds only when $u = 0$.

An Interesting Result

For any positive-definite quadratic form

$$Q = \sum_{i,j=1}^{N} a_{ij} u_i u_j ,$$

the associated determinant

$$D = \det\left(a_{ij}\right)$$

is always positive.

Proof

- First, we prove that $D \neq 0$. By contradiction, assume otherwise, there is a nontrivial solution to

$$\sum_{j=1}^{N} a_{ij} u_j = 0, \quad i = 1, 2, \ldots, N$$

Then, it follows that

$$Q = \sum_{i=1}^{N} u_i \left(\sum_{J=1}^{N} a_{ij} u_j \right) = 0$$

a contradiction.

• Second, we prove that $D > 0$. For $\lambda \in [0,1]$, consider a family of quadratic forms defined as

$$P(\lambda) = \lambda Q + (1-\lambda)\sum_{i=1}^{N} u_i^2 .$$

Clearly, $P(\lambda) > 0$, for all nontrivial u. Then, based on the above analysis, the associated determinants are nonzero.

At $\lambda = 0$, the determinant is $\det I > 0$.

So, by continuity, *at* $\lambda = 1$, the determinant is D which cannot be negative.

General 2x2 Symmetric Matrices

We begin with the two-dimensional case:

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \doteq \begin{pmatrix} a^1 \\ a^2 \end{pmatrix}$$

which is symmetric, i.e., $a_{12} = a_{21}$.

Consider a pair of eigenvalue λ_1 and associated (normalized) eigenvector $x^1 := \begin{pmatrix} x_{11} \\ x_{12} \end{pmatrix}$, *i.e.*

$$Ax^1 = \lambda_1 x^1 \Leftrightarrow \langle a^1, x^1 \rangle = \lambda_1 x_{11}, \quad \langle a^2, x^1 \rangle = \lambda_1 x_{12}$$

General Symmetric Matrices (Cont'd)

Using the Gram-Schmidt process, take a 2×2 orthogonal matrix $O_2 = \left(y^1 \; y^2 \right)$, with $y^1 := x^1$ the given normalized eigenvector.

It will be shown that

$$O_2^T A O_2 = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

General Symmetric Matrices (Cont'd)

First, show that

$$O_2^T A O_2 = O_2^T \begin{pmatrix} \lambda_1 y_{11} & \langle a^1, y^2 \rangle \\ \lambda_1 y_{12} & \langle a^2, y^2 \rangle \end{pmatrix} = \begin{pmatrix} \lambda_1 & b_{12} \\ 0 & b_{22} \end{pmatrix}$$

Then, $b_{12} = 0$ using symmetry;

$$\left(O_2^T A O_2\right)^T = O_2^T A O_2.$$

and $b_{22} = \lambda_2$ because the eigenvalues are unchanged under O.

Exercise 1

Try to reduce the real symmetric matrix

$$A = \begin{pmatrix} 1 & k \\ k & 1 \end{pmatrix}$$

to a diagonal form.

Exercise 2

Define the real bilinear form

$$Q(x,y) = y^T A x = \sum_{i,j=1}^{n} a_{ij} y_i x_j, \qquad \forall x, y \in \mathbb{R}^n$$

that reduces to the inner product when $A = I$.

Prove that Q is symmetric, i.e., $Q(x,y) = Q(y,x)$

if and only if A is symmetric.

See the text (Horn & Johnson, 2nd edition, 2013; page 226)

Homework #4

1. Does the singular matrix

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

have two independent eigenvectors?

2. Show that A and A^T have the same eigenvalues.

Homework #4

3. Show by direct calculation for A and B, 2×2 matrices, that AB and BA have the same characteristic equation.

4. Can you give two matrices that are reducible to the following canonical diagonal matrix

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

Justify your answer.