Lecture VII

- The Jordan Canonical Form
- Examples and Applications

Review of Canonical Forms

- If A is an $n \times n$ matrix with distinct eigenvalues, then there exists a nonsingular matrix P s.t.

$$P^{-1}AP = \mathrm{diag}(\lambda_i).$$

- If A is Hermitian (possibly having repeated eigenvalues), $\exists$ a unitary matrix U such that

$$U^*AU = \mathrm{diag}(\lambda_i).$$

A Motivating Example

As stated previously, not every matrix can be transformed into a canonical diagonal form. For example,

$$A = \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}, \quad b \neq 0$$

cannot be transformed into a diagonal matrix.

Jordan Canonical Form

Theorem (Jordan):

Let A be an $n \times n$ matrix whose different eigenvalues are $\lambda_1, \ \ldots, \ \lambda_s$ with multiplicities $m_1, \ldots, m_s$:

$$\det\left(\lambda I - A\right) = \prod_{i=1}^{s} \left(\lambda - \lambda_i\right)^{m_i}$$

Then, A is transformable into a Jordan canonical form.

i.e., $\exists$ nonsingular P such that

$$P^{-1} A P = \text{blockdiag}\left(\Lambda_i\right) \doteq J$$

Theorem (Jordan), cont'd:

$$P^{-1}AP = \text{blockdiag}\left(\Lambda_i\right) \doteq J$$

where

$$\Lambda_i = \begin{pmatrix} \lambda_i & 0 & \cdots & \cdots & 0 \\ 1 & \lambda_i & \ddots & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots 1 & \lambda_i & 0 \\ 0 & 0 & \cdots & \cdots\ 1 & \lambda_i \end{pmatrix}$$

Jordan Block

Comments

- *In some texts*, J^T is used as Jordan form.

- Different Jordan blocks, say Λ_i, Λ_j may be associated with the same eigenvalues.

- The number $\bar{s}$ of Jordan blocks: $s \leq \bar{s} \leq n$.

Illustration via 3x3 matrices

If a 3×3 matrix A has an eigenvalue λ_1 of multiplicity three, then it may be reduced into one of the following Jordan forms:

$$J_1 = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_1 \end{pmatrix}, \quad J_2 = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 1 & \lambda_1 \end{pmatrix},$$

$$J_3 = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 1 & \lambda_1 & 0 \\ 0 & 1 & \lambda_1 \end{pmatrix}.$$

Remark 1

The distinct Jordan forms $\left(J_i,\ J_k\right),\ i \neq k,$ are not similar to each other.

Remark 2

When each Jordan block $\Lambda_i(\lambda_i)$ in the Jordan form J is one-dimensional (i.e. $n_i = 1$) and $s = n$, the Jordan matrix J becomes diagonal.

Application to Matrix Analysis of Differential Equations

Given a set of 1st-order differential equations

$$\dot{x}(t) = Ax(t), \quad x(0) \in \mathbb{R}^n,$$

applying the transformation $y = P^{-1}x$ yields:

$$\dot{y}(t) = \left(P^{-1}AP\right)y(t) \;:=\; Jy(t).$$

$$\Leftrightarrow \;\; \dot{y}^i(t) = \Lambda_i y^i(t), \;\; y^i \in \mathbb{R}^{m_i}, \; y \doteq \begin{bmatrix} y^1 \\ \vdots \\ y^{\bar{s}} \end{bmatrix}.$$

Comment

So, with the help of Jordan canonical form, solving differential equations can be reduced down to solving ***lower-order (disjoint!)*** differential equations.

(see a forthcoming lecture.)

Principal Vectors

In order to develop a constructive proof for Jordan's Theorem, let's introduce the notion of principal vector, or generalized eigenvector, which is a generalization of eigenvector.

Principal Vectors

A (possibly zero) vector p is a principal vector of grade $g \geq 0$ belonging to the eigenvalue λ_i if

$$\left(\lambda_i I - A\right)^g p = 0,$$

for which g is the smallest non-negative integer.

Examples

- The vector $p = 0$ is the principal vector of grade 0.

- The (nonzero) eigenvectors are the principal vectors of grade 1.

Motivating Question

In case of transformation to diagonal canonical form, i.e., $P^{-1}AP = \operatorname{diag}(\lambda_i)$, the columns of P are linearly independent eigenvectors.

What about the matrix P in Jordan form?
How to construct P from principal vectors?

Linear Spaces

Define the linear space composed of all principal vectors of grade $\leq g$ belonging to λ_i :

$$P_g(\lambda_i) = \left\{ p \mid \ (\lambda_i I - A)^g \ p = 0 \right\}$$

i.e., the null space of $(\lambda_i I - A)^g$.

Clearly,

$$P_0(\lambda_i) \subset P_1(\lambda_i) \subset P_2(\lambda_i) \subset \cdots$$

An Interesting Result

Let A be an $n \times n$ matrix with the distinct eigenvalues $\lambda_1, \ldots, \lambda_s$, $1 \le s \le n$, with multiplicities $m_1, \ldots, m_s$.

Then, every vector $x \in \mathbb{R}^n$ can be written as

$$x = p^1 + p^2 + \cdots + p^s$$

where p^i is a uniquely defined principal vector associated with λ_i of grade $\le m_i$.

Comment 1

A special, but interesting, case is when there are n linearly independent eigenvectors, say, $c^1, \ \ldots, \ c^n$. In this case, $\exists \xi_i$ scalars s.t.

$$x = \xi_1 c^1 + \cdots + \xi_n c^n := p^1 + \cdots + p^n.$$

Comment 2

Its proof relies upon the well-known Cayley-Hamilton theorem; see any standard matrix or linear algebra textbook.

Example

Consider the matrix $A = \begin{pmatrix} 1 & 0 & 0 \\ 7 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$.

* Compute its eigenvalues and the associated eigenvectors.
* Can each column be written as a linear combination of eigenvectors?
* Show that each column can be written as a unique representation of principal vectors.

Answer

- $\lambda_1 = 1 \ (m_1 = 2), \ \ \lambda_2 = 2 \ (m_2 = 1).$
- *The* eigenvectors of λ_1 are of the form

  $\xi \times \mathrm{col}(0,1,0)$, ξ any nonzero scalar.

  The eigenvectors of λ_2 are of the form

  $\xi \times \mathrm{col}(0,0,1)$, ξ any nonzero scalar.
- $P_2(\lambda_1) = \{ p^1 \mid p^1 = col(\alpha, \beta, 0) \}$

  $P_1(\lambda_2) = \{ p^2 \mid p^2 = col(0, 0, \gamma) \}.$

Cayley-Hamilton Theorem Revisited

For any $n \times n$ matrix A,

$$\rho(A) = A^n + \alpha_1 A^{n-1} + \cdots + \alpha_{n-1} A + \alpha_n I = O$$

where $\rho(\lambda)$ is the characteristic polynomial of A, i.e.,

$$\rho(\lambda) = \det(\lambda I - A) = \lambda^n + \sum_{i=1}^{n} \alpha_i \lambda^{n-i}.$$

Example

Consider $A = \begin{pmatrix} -7 & -4 \\ 8 & 5 \end{pmatrix}$. Verify that

1) The characteristic polynomial $\rho_A(\lambda)$ is:

$$\rho_A(\lambda) = \lambda^2 + 2\lambda - 3.$$

2) $\rho_A(A) = A^2 + 2A - 3I = 0 \triangleq \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$

Another Proof

Define the $n \times n$ matrix of signed cofactors:

$$C(\lambda) = \text{cof}(\lambda I - A).$$

Then, using $M(\text{cof}\, M)^T = (\det M)I$,

$$(\lambda I - A)C^T(\lambda) = \rho(\lambda)I.$$

In addition,

$$C^T(\lambda) = \lambda^{n-1}C_0 + \cdots + \lambda C_{n-1} + C_n$$

for constant matrices C_i's.

Proof (cont'd)

By identification of the coefficients of equal powers of λ gives

$$C_0 = I$$
$$C_1 - AC_0 = \alpha_1 I$$
$$\vdots$$
$$C_{n-1} - AC_{n-2} = \alpha_{n-1} I$$
$$-AC_{n-1} = \alpha_n I.$$

Multiplying the first eq. by A^n, the second by A^{n-1},...,

and then adding them up leads to: $O = \rho(A) I$.

Question

How to compute principal vectors for a given matrix?

A Motivating Example

Consider a 2×2 Jordan block $J = \begin{pmatrix} \lambda & 0 \\ 1 & \lambda \end{pmatrix}$.

Denote $P = \left[x^1 \;\; x^2 \right]$ that transforms A into J.

Namely, $P^{-1}AP = J$. So, we have $AP = PJ$, or

$$A\left[x^1 \;\; x^2 \right] = \left[x^1 \;\; x^2 \right] \begin{pmatrix} \lambda & 0 \\ 1 & \lambda \end{pmatrix}$$

$\Rightarrow Ax^2 = \lambda x^2$, so x^2 is an eigenvector;

$(A - \lambda I)x^1 = x^2$, so x^1 is a principal vector (of grade 2).

Comment

Usually, $\left\{x^1,\ x^2\right\}$ is called a Jordan Basis for this 2×2 matrix A. In order words, the JCF transformation matrix P is composed of a Jordan basis, or a set of linearly independent eigenvectors and principal vectors.

General Procedure

Step 1: Solve the characteristic equation

$$\left(A-\lambda I\right)z^{1}=0.$$

Step 2 : For each independent z^{1}, solve

$$\left(A-\lambda I\right)z^{2}=z^{1}$$

where z^{2} clearly solves $\left(A-\lambda I\right)^{2}z^{2}=0.$

Collect only those z^{2} which are linearly independent with the previously found eigenvectors z^{1}.

General Procedure

Step 3: For each independent z^2, solve

$$\left(A-\lambda I\right)z^3 = z^2$$

where z^3 clearly solves $\left(A-\lambda I\right)^3 z^3 = 0.$

Collect only those z^3 which are linearly independent with the previously found vectors $z^1,\ z^2$.

Step 4: Continue in this way till the total number of independent eigenvectors and principal vectors equals to the (algebraic) multiplicity of λ.

General Procedure

Step 4 (cont'd): Denote

$$\left[x^1, x^2, \cdots, x^m\right] = \left[z^m, z^{m-1}, \cdots, z^1\right]$$

and

$$P = \left[x^1, x^2, \cdots, x^m\right].$$

Therefore,

$$P^{-1}AP = J \text{ (associated with eigenvalue } \lambda\text{).}$$

Comments

- **Not any arbitrary choice of linearly independent principal vectors would lead to a correct transformation matrix *P*.**

For example, at Step 2, the linearly independent principal vectors z^2 are chosen according to

$$\left(A - \lambda I\right) z^2 = z^1$$

but NOT :

$$\left(\lambda I - A\right) z^2 = z^1.$$

- **See (the 1960 book of Gantmacher, Vol.1, Chap. VI, Section 8) for another general method of constructing a transformation matrix.**

More on Jordan Basis

Without going into the full details in proving Jordan's Theorem, let's illustrate the concept of Jordan basis and its use in the canonical transformation.

Consider a principal vector v of grade $g = n = 4$. Define:

$$\left.\begin{aligned} x^1 &:= v \\ x^2 &:= (A - \lambda I)x^1 \\ x^3 &:= (A - \lambda I)x^2 \\ x^4 &:= (A - \lambda I)x^3 \end{aligned}\right\} \text{Jordan Basis}$$

Jordan Basis (cont'd)

Then, the 4×4 matrix A can be transformed into the Jordan canonical form:

$$J = \begin{pmatrix} \lambda & 0 & 0 & 0 \\ 1 & \lambda & 0 & 0 \\ 0 & 1 & \lambda & 0 \\ 0 & 0 & 1 & \lambda \end{pmatrix}$$

Principal vector of grade 4

That is,

eigenvector

$$P^{-1}AP = J, \quad P = \left(x^1, x^2, x^3, x^4\right).$$

Comment

If we define $\tilde{P} = \left(x^4, x^3, x^2, x^1\right)$, then A is transformed into the Jordan canonical form J^T, i.e.:

$$\tilde{P}^{-1} A \tilde{P} = J^T = \begin{pmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{pmatrix}.$$

A More Complex Case

If $\left(A-\lambda I\right)$ has rank $n-2,$ i.e. its null space is of dimension 2, then $\exists$ two linearly independent eigenvectors to $\left(A-\lambda I\right)q=0.$

Thus, we need $n-2$ linearly independent principal vectors. In this case, the Jordan basis takes the form $\left\{v^1,v^2,\ldots,v^k\right\}$ and $\left\{u^1,u^2,\ldots,u^l\right\},\;\; k+l=n.$

So, A is transformed into the Jordan canonical form

$$P^{-1}AP = diag\left\{J_1,\ J_2\right\}.$$

Exercise 1

Find a transformation matrix P to bring the following matrix

$$M = \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}, \quad b \neq 0$$

into the Jordan Canonical Form

$$J = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

Exercise 2

Find a transformation matrix to bring the following matrix into a Jordan form:

$$A = \begin{pmatrix} 1 & -1 & 1 & -1 \\ -3 & 3 & -5 & 4 \\ 8 & -4 & 3 & -4 \\ 15 & -10 & 11 & -11 \end{pmatrix}$$

Solution:

$$P = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & -5 & 0 & -5 \\ 0 & 4 & 1 & 5 \\ -1 & 11 & 0 & 12 \end{pmatrix},$$

$$J = \begin{pmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

$\lambda I - A$ becomes (after elementary operations on rows and columns:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \lambda+1 & 0 \\ 0 & 0 & 0 & (\lambda+1)^3 \end{pmatrix}.$$

Therefore, the matrix has two elementary divisors:

$$\lambda+1 \text{ and } (\lambda+1)^3,$$

which give two Jordan blocks, respectively:

$$J_1 = -1, \quad J_2 = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{pmatrix}.$$

See (the 1960 book of Gantmacher, Vol.1, pp.160-164) for the details.

Practicing Problems for Midterm

1. Compute the eigenvalues of the matrix

$$A = \begin{pmatrix} 7 & -2 \\ 4 & 1 \end{pmatrix}$$

and transform it to one of the canonical forms.

Practicing Problems for Midterm

2. Consider the block diagonal matrix

$$A = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}, \text{ with } A_i \in \mathbb{R}^{n_i \times n_i}, \; n_1 + n_2 = n.$$

Show that the eigenvalues of A are those of A_1 and A_2.

Practicing Problems for Midterm

3. Assume A is a nonsingular matrix. If λ is an eigenvalue of A with eigenvector x, show that λ^{-1} is an eigenvalue of A^{-1}. In addition, give an eigenvector associated with λ^{-1}.

Practicing Problems for Midterm

4. Show that $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ cannot be transformed into a diagonal matrix under any similarity transformation.

Practicing Problems for Midterm

5. For any given 2×2 real orthogonal matrix U, one of the following must hold:

(i) $U = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$ for some θ;

(ii) $U = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$ for some θ.

(Only for those who love math proof!)

Practicing Problems for Midterm

6. Show that J^T is similar to J. That is,

$$\begin{pmatrix} 0 & \cdots & 1 \\ \vdots & \iddots & \vdots \\ 1 & \cdots & 0 \end{pmatrix} J^T \begin{pmatrix} 0 & \cdots & 1 \\ \vdots & \iddots & \vdots \\ 1 & \cdots & 0 \end{pmatrix} = J.$$

Practicing Problems for Midterm

7. Assume that A, D are invertible matrices.

Show that

$$\begin{pmatrix} A & B \\ 0 & D \end{pmatrix}^{-1} = \begin{pmatrix} A^{-1} & -A^{-1}BD^{-1} \\ 0 & D^{-1} \end{pmatrix}.$$

Practicing Problems for Midterm

8. Assume that $A,\ D$ are invertible matrices.

Show that

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}^{-1} = \begin{pmatrix} A^{-1} + A^{-1}BECA^{-1} & -A^{-1}BE \\ -ECA^{-1} & E \end{pmatrix}$$

where E is the inverse of the Schur complement

of A : $\quad E = \left(D - CA^{-1}B\right)^{-1}.$

***Note*: A Very Useful Identity.**

Practicing Problems for Midterm

9. Reduce the following matrix into a canonical diagonal form:

$$A = \begin{pmatrix} M & 0_{2\times 2} \\ 0_{2\times 2} & M \end{pmatrix}$$

where

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

Practicing Problems for Midterm

10. Reduce the following matrix into a Jordan canonical form:

$$A = \begin{pmatrix} 3 & 2 & 1 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

Practicing Problems for Midterm

11. Rank Inequalities (See Horn-Johanson text, page 13)

- Sylvester inequality

  $\forall A \in \mathbb{R}^{m\times k},\ B \in \mathbb{R}^{k\times n},$ we have

  $\left(\mathrm{rank}A + \mathrm{rank}B\right) - k \leq \mathrm{rank}AB \leq \min\left\{\mathrm{rank}A,\ \mathrm{rank}B\right\}.$

- Frobenius inequality

  $\forall A \in \mathbb{R}^{m\times k},\ B \in \mathbb{R}^{k\times p},\ C \in \mathbb{R}^{p\times n},$ we have

  $\mathrm{rank}AB + \mathrm{rank}BC \leq \mathrm{rank}B + \mathrm{rank}ABC$

  with equality iff there are matrices X and Y such that

  $B = BCX + YAB.$

Homework #7

1. For the matrix

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

identify the spaces $P_g(\lambda)$ and the principal vectors of grade 2.

Homework #7

2. Express the following vectors as unique representations of principal vectors found in Problem 1:

$$x = \begin{bmatrix} \sqrt{2} \\ -9 \\ 84 \end{bmatrix}, \quad x = \begin{bmatrix} 0 \\ 9.3 \\ 0 \end{bmatrix}.$$

Homework #7

3. Can you transform the following matrix into a Jordan form:

$$A = \begin{pmatrix} \lambda & \lambda & \lambda \\ 0 & \lambda & \lambda \\ 0 & 0 & \lambda \end{pmatrix}, \quad \lambda \neq 0?$$