Theory of Eigensystems

Let A be a square matrix. The nonzero vector $x \in \mathbb{C}^n$ is called an **eigenvector** of $A \in \mathbb{C}^{n\times n}$ if Ax is a multiple of x,

$$Ax = \lambda x, \quad x \neq 0. \tag{1}$$

(λ, x) is called an **eigenpair.**

Easy to show: Eigenvalues must satisfy

$$\det(\lambda I - A) = 0.$$

The set of all eigenvalues of A is called the **spectrum** of A and will be denoted by

$$\Lambda(A) = \{\lambda_1, \ldots, \lambda_n\}.$$

A subspace $\mathcal{X}$ of $\mathbb{C}^n$ is an **invariant subspace** for A if $A\mathcal{X} \subseteq \mathcal{X}$, that is, $x \in \mathcal{X}$ implies $Ax \in \mathcal{X}$.

Theorem (Thm. 1)

Let the columns of $X \in \mathbb{C}^{n \times p}$, $p \leq n$, form a basis for a subspace $\mathcal{X}$ of $\mathbb{C}^n$. Then $\mathcal{X}$ is an invariant subspace for A if and only if $AX = XB$ for some $B \in \mathbb{C}^{p \times p}$. When the latter equation holds, the spectrum of B is contained within that of A.

Proof:

Similarity, Unitary Similarity

Let $A, B \in \mathbb{C}^{n\times n}$. The matrices A and B are **similar** if there exists a nonsingular matrix P such that

$$B = P^{-1}AP. \tag{2}$$

This is called a **similarity transformation** and P is the **transforming matrix**.

Theorem (Thm. 2)

Let A and B be similar, say $B = P^{-1}AP$. Then A and B have the same eigenvalues, and x is an eigenvector of A with associated eigenvalue λ if and only if $P^{-1}x$ is an eigenvector of B with associated eigenvalue λ.

A and B are said to be **unitarily similar** if there is a unitary matrix U such that $B = U^*AU$. If A and B are real, then they are said to be **orthogonally similar** if there is a real, orthogonal matrix U such that $B = U^TAU$.

If a matrix A is similar to a diagonal matrix then A is said to be **diagonalizable** or **simple**.

How is Similarity Used in Solving Problems?

The **similarity method** is a strategy frequently used for solving problems. Here is an outline of the method.

- Step 1 : Choose a matrix B similar to A for which the problem is easier to solve.
- Step 2 : Solve the problem using the matrix B instead of A (the B-problem).
- Step 3 : Interpret the solution to the B-problem in terms of the matrix A.

Example

Given $A = \begin{bmatrix} 1 & 4 \\ 3 & 2 \end{bmatrix}$, find each entry in A^{1010}.

Example

In certain problems in economics the state of a system is described by a matrix $S_n = I + A + A^2 + \cdots + A^n$ at time n where A is a given matrix. Use the similarity method to investigate the behaviour of the system in the "long run", i.e., when n is large.

Canonical Forms

Issai Schur (1875–1941) may have asked:

What is the simplest form a square matrix can take under similarity transformations?

Particularly important: ***Unitary transforms.***

Theorem (Schur's theorem, Thm. 3)

Let $A \in \mathbb{C}^{n\times n}$. Then there exists a unitary matrix U and an upper triangular matrix T such that

$$T = U^{-1}AU = U^{*}AU.$$

Proof:

The unitary similarity transformation $T = U^{-1}AU$ can also be written

$$A = UTU^*,$$

called the **Schur decomposition** of A.

- Since $\det(U)\det(U^*) = \det(UU^*) = 1$,

$$\det(\lambda I - A) = \det(\lambda I - T) = \prod_{i=1}^{n}(\lambda - t_{ii}),$$

  the diagonal elements of T are the eigenvalues of A.
- Schur decomposition is not unique.
- The columns of U are called **Schur vectors**.
- With the exception of u_1, the Schur vectors are not, in general, eigenvectors of A.

Diagonalizable Matrices

Is any matrix A diagonalizable?

In other words can we always find a nonsingular P such that $P^{-1}AP$ is diagonal?

Class of Matrices Unitarily Similar to a Diagonal Matrix:

A matrix A is **normal** if $AA^* = A^*A$.

- Contains the important subclasses of Hermitian and unitary matrices.
- Schur's theorem takes nice form: the triangular matrix T is diagonal.
- This special form of Schur's theorem is called the *spectral theorem*.

Theorem (Spectral theorem, Thm. 4)

Let $A \in \mathbb{C}^{n\times n}$. Then A is normal if and only if there is a unitary matrix U and a diagonal matrix Λ such that

$$A = U\Lambda U^*.$$

Theorem (Spectral theorem, Thm. 4)

Let $A \in \mathbb{C}^{n \times n}$. Then A is normal if and only if there is a unitary matrix U and a diagonal matrix Λ such that

$$A = U \Lambda U^*.$$

Proof: ($\Rightarrow$) Let $A = UTU^*$ be the Schur decomposition of A. If A is normal then T is normal (see Exercise 4). Since a normal and triangular matrix is diagonal (see Exercise 8), T is diagonal.

Theorem (Spectral theorem, Thm. 4)

Let $A \in \mathbb{C}^{n \times n}$. Then A is normal if and only if there is a unitary matrix U and a diagonal matrix Λ such that

$$A = U \Lambda U^*.$$

Proof: ($\Rightarrow$) Let $A = UTU^*$ be the Schur decomposition of A. If A is normal then T is normal (see Exercise 4). Since a normal and triangular matrix is diagonal (see Exercise 8), T is diagonal.

($\Leftarrow$) If $A = U\Lambda U^*$ with U unitary and Λ diagonal then, since diagonal matrices commute,

$$\begin{aligned} AA^* &= (U\Lambda U^*)(U\Lambda^* U^*) \\ &= U\Lambda\Lambda^* U^* \\ &= U\Lambda^*\Lambda U^* \\ &= (U\Lambda^* U^*)(U\Lambda U^*) = A^*A. \end{aligned}$$ □

Theorem (Thm. 5)

$A \in \mathbb{C}^{n\times n}$ is normal if and only if it has n orthogonal eigenvectors.

Theorem (Thm. 5)

$A \in \mathbb{C}^{n\times n}$ is normal if and only if it has n orthogonal eigenvectors.

Proof: ($\Rightarrow$) From spectral Theorem, normal matrices have an orthonormal basis of eigenvectors.

Theorem (Thm. 5)

$A \in \mathbb{C}^{n\times n}$ is normal if and only if it has n orthogonal eigenvectors.

Proof: ($\Rightarrow$) From spectral Theorem, normal matrices have an orthonormal basis of eigenvectors.

($\Leftarrow$) Suppose $u_1, \dots, u_n$ is an orthonormal basis of $\mathbb{C}^n$ consisting of eigenvectors of A:

$$Au_j = \lambda_j u_j, \quad j = 1, \dots, n.$$

Let $U = [\, u_1 \quad \cdots \quad u_n \,]$ and $\Lambda = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$. Then $AU = U\Lambda$ or equivalently $A = U\Lambda U^*$ and A is normal by the spectral Theorem. □

Examples of Normal Matrices

The following matrices satisfy $AA^* = A^*A$:

- **Hermitian** matrices, i.e., $A^* = A$,
- **symmetric** matrices, i.e., $A^T = A$,
- **unitary** matrices, i.e., $AA^* = A^*A = I$,
- **orthogonal** matrices, i.e., $AA^T = A^TA = I$,
- **skew-Hermitian** matrices, i.e., $A^* = -A$,
- **skew-symmetric** matrices, i.e., $A^T = -A$.

Matrices Similar to a Diagonal Matrix

Recall: Normal matrices are unitarily diagonalizable.

Question: Which matrices are diagonalizable in general, i.e., $P^{-1}AP$ is diagonal with P not necessarily unitary?

Theorem (Thm. 6)

A matrix $A \in \mathbb{C}^{n \times n}$ is diagonalizable if and only if A has n linearly independent eigenvectors.

Proof:

Matrices with distinct eigenvalues

Theorem (Thm. 7)

A matrix with distinct eigenvalues is diagonalizable.

Matrices with distinct eigenvalues

Theorem (Thm. 7)

A matrix with distinct eigenvalues is diagonalizable.

Proof: Assume A's e'vals $\lambda_1, \ldots, \lambda_n$ are distinct and assume e'vecs $x_1, \ldots, x_n$ are linearly dependent so that $\sum_{i=1}^{n} \alpha_i x_i = 0$ with $\alpha_k \neq 0$ for some k. We may assume $\alpha_n \neq 0$. Then

$$0 = (A - \lambda_1 I) \sum_{i=1}^{n} \alpha_i x_i = \sum_{i=2}^{n} \alpha_i (\lambda_i - \lambda_1) x_i.$$

Matrices with distinct eigenvalues

Theorem (Thm. 7)

A matrix with distinct eigenvalues is diagonalizable.

Proof: Assume A's e'vals $\lambda_1, \ldots, \lambda_n$ are distinct and assume e'vecs $x_1, \ldots, x_n$ are linearly dependent so that $\sum_{i=1}^n \alpha_i x_i = 0$ with $\alpha_k \neq 0$ for some k. We may assume $\alpha_n \neq 0$. Then

$$0 = (A - \lambda_1 I) \sum_{i=1}^{n} \alpha_i x_i = \sum_{i=2}^{n} \alpha_i (\lambda_i - \lambda_1) x_i.$$

$$0 = (A - \lambda_2 I) \sum_{i=2}^{n} \alpha_i (\lambda_i - \lambda_1) x_i = \sum_{i=3}^{n} \alpha_i (\lambda_i - \lambda_1)(\lambda_i - \lambda_2) x_i.$$

Matrices with distinct eigenvalues

Theorem (Thm. 7)

A matrix with distinct eigenvalues is diagonalizable.

Proof: Assume A's e'vals $\lambda_1, \ldots, \lambda_n$ are distinct and assume e'vecs $x_1, \ldots, x_n$ are linearly dependent so that $\sum_{i=1}^{n} \alpha_i x_i = 0$ with $\alpha_k \neq 0$ for some k. We may assume $\alpha_n \neq 0$. Then

$$0 = (A - \lambda_1 I) \sum_{i=1}^{n} \alpha_i x_i = \sum_{i=2}^{n} \alpha_i (\lambda_i - \lambda_1) x_i.$$

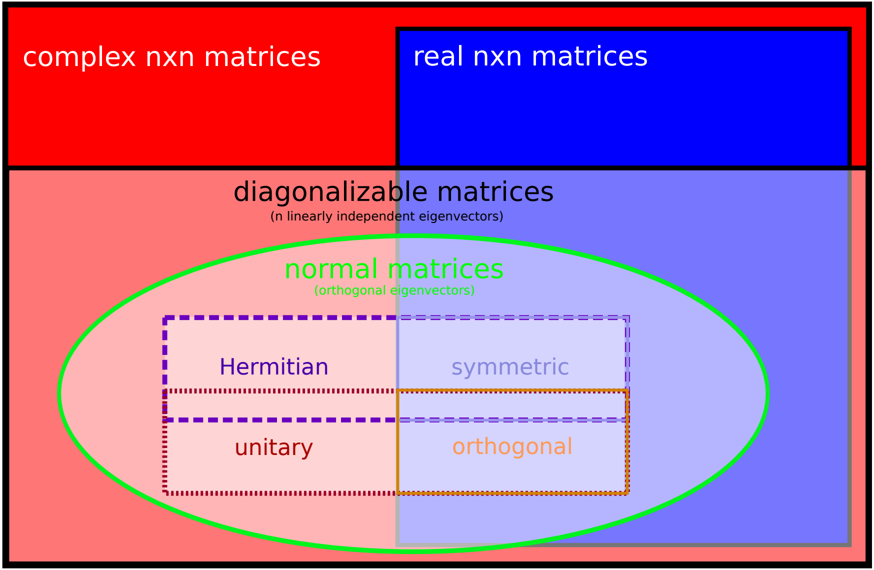
$$0 = (A - \lambda_2 I) \sum_{i=2}^{n} \alpha_i (\lambda_i - \lambda_1) x_i = \sum_{i=3}^{n} \alpha_i (\lambda_i - \lambda_1)(\lambda_i - \lambda_2) x_i.$$

Continuing similar multiplications we obtain that

$$0 = \alpha_n (\lambda_n - \lambda_{n-1})(\lambda_n - \lambda_{n-2}) \ldots (\lambda_n - \lambda_1) x_n.$$

Contradiction since the $\lambda_i \neq \lambda_j$, $\alpha_n \neq 0$ and $x_n \neq 0$. Hence A has n l. i. eigenvectors and it is diagonalizable by Thm. 6.

Recap



complex nxn matrices
real nxn matrices
diagonalizable matrices
(n linearly independent eigenvectors)
normal matrices
(orthogonal eigenvectors)
Hermitian
symmetric
unitary
orthogonal

Application: Principal Component Analysis

Assume that one makes m measurements for each of n objects, and collect these data in columns $b_1, \dots, b_n \in \mathbb{R}^m$.

Example: $m = 3$ measurements, $n = 8$ people

	P1	P2	P3	P4	P5	P6	P7	P8
age	22	30	23	23	22	21	22	21
weight	10.4	12.2	10.5	10.9	9.0	12.5	11.5	10.2
shoe size	7	8	7	7	8	8	9	7

Question: If one wanted to distinguish these 8 people by a linear combination of the 3 measurements, what would be a best possible combination?

Application: Principal Component Analysis

1. Form the data matrix $B \in \mathbb{R}^{m \times n}$:

$$B = \begin{bmatrix} 22 & 30 & 23 & 23 & 22 & 21 & 22 & 21 \\ 10.4 & 12.2 & 10.5 & 10.9 & 9.0 & 12.5 & 11.5 & 10.2 \\ 7 & 8 & 7 & 7 & 8 & 7 & 9 & 7 \end{bmatrix}$$

2. Substract the mean for each row, $\widehat{B} = B - B[1, \ldots, 1]^T / n$:

$$\widehat{B} = \begin{bmatrix} -1 & 7 & 0 & 0 & -1 & -2 & -1 & -2 \\ -0.5 & 1.3 & -0.4 & 0 & -1.9 & 1.6 & 0.6 & -0.7 \\ -0.5 & 0.5 & -0.5 & -0.5 & 0.5 & -0.5 & 1.5 & -0.5 \end{bmatrix}$$

3. Form the symmetric **covariance matrix** $\widehat{C} = \frac{1}{n-1} \widehat{B}\widehat{B}^T$:

$$\widehat{C} = \begin{bmatrix} 8.571 & 1.300 & 0.571 \\ 1.300 & 1.303 & 0.085 \\ 0.571 & 0.085 & 0.571 \end{bmatrix}$$

4. Compute the eigenvectors and eigenvalues of $\widehat{C}$:

$$\widehat{C}U = UD, \quad U^T U = I, \quad D = \mathrm{diag}(\lambda_1, \dots, \lambda_m), \quad \lambda_j \geq \lambda_{j+1}.$$

In our example (age, weight, shoe size):

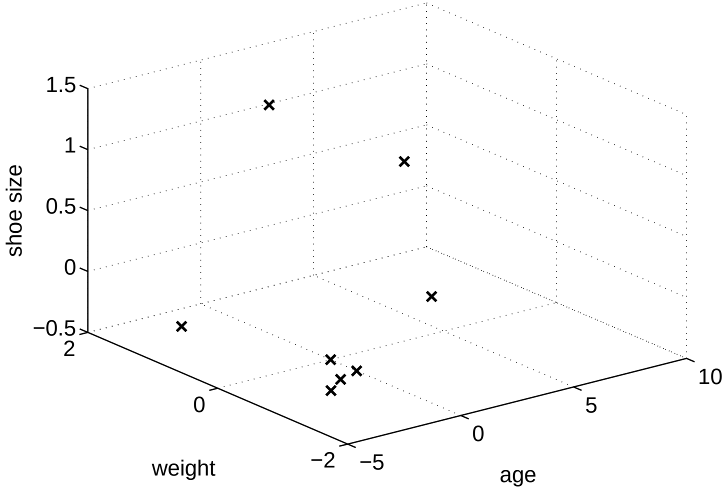
$$U = \begin{bmatrix} 0.982 & 0.169 & -0.073 \\ 0.170 & -0.985 & 0.012 \\ 0.070 & 0.024 & 0.997 \end{bmatrix} \text{ and } D = \begin{bmatrix} 8.84 & & \\ & 1.08 & \\ & & 0.53 \end{bmatrix}.$$

The columns of $U = [u_1, \dots, u_m]$ are orthogonal directions in which the **principal components** $u_j^T b$ have the largest possible variance for all data columns b of B, with the **variance** given by λ_j.

With $u_1 = [0.982, 0.170, 0.070]^T$ we find that "age" is the best separator for our 8 people, and the artificial variable

$$0.982 \times \text{age} + 0.170 \times \text{weight} + 0.070 \times \text{shoe size}$$

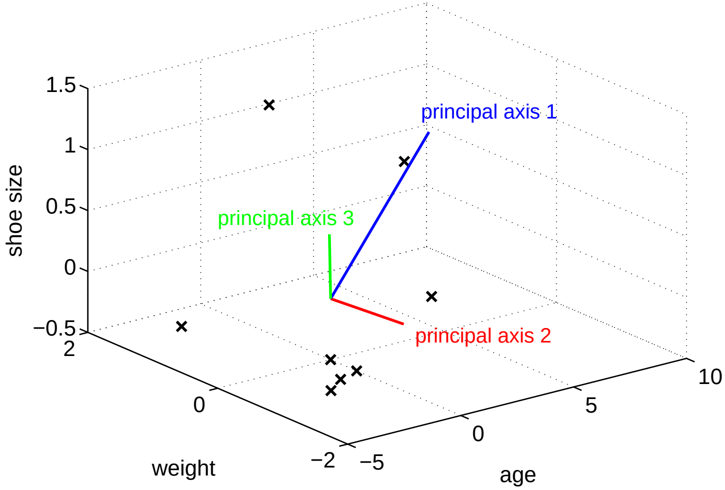
has a largest possible variance of $\lambda_1 = 8.84$.



With $u_1 = [0.982, 0.170, 0.070]^T$ we find that "age" is the best separator for our 8 people, and the artificial variable

$$0.982 \times \text{age} + 0.170 \times \text{weight} + 0.070 \times \text{shoe size}$$

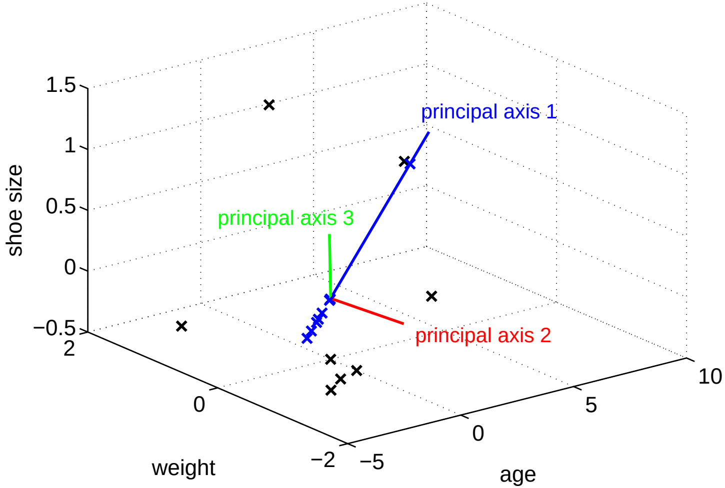
has a largest possible variance of $\lambda_1 = 8.84$.



With $u_1 = [0.982, 0.170, 0.070]^T$ we find that "age" is the best separator for our 8 people, and the artificial variable

$$0.982 \times \text{age} + 0.170 \times \text{weight} + 0.070 \times \text{shoe size}$$

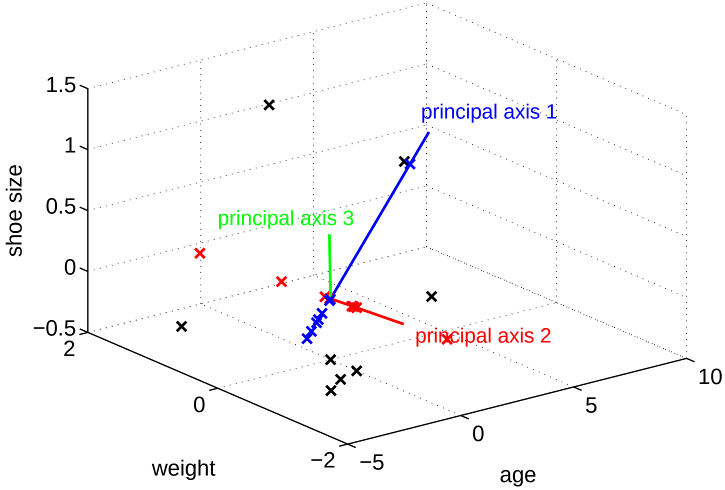
has a largest possible variance of $\lambda_1 = 8.84$.



With $u_1 = [0.982, 0.170, 0.070]^T$ we find that “age” is the best separator for our 8 people, and the artificial variable

$$0.982 \times \text{age} + 0.170 \times \text{weight} + 0.070 \times \text{shoe size}$$

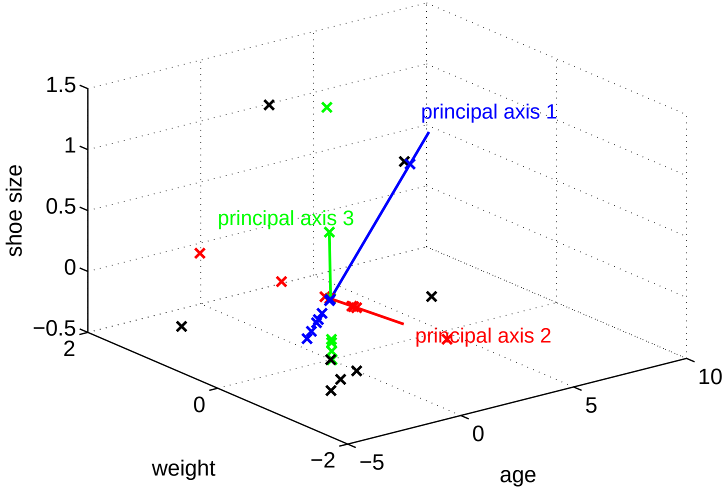
has a largest possible variance of $\lambda_1 = 8.84$.



With $u_1 = [0.982, 0.170, 0.070]^T$ we find that “age” is the best separator for our 8 people, and the artificial variable

$$0.982 \times \text{age} + 0.170 \times \text{weight} + 0.070 \times \text{shoe size}$$

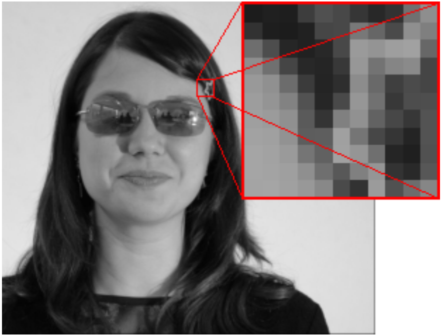
has a largest possible variance of $\lambda_1 = 8.84$.



Application: Face Recognition

The **eigenface method** by L. Sirovich and M. Kirby (1987) and M. Turk and A. Pentland (1991) is based on principal component analysis.

The (long) vectors $b_1, \ldots, b_n$ correspond to n different images, and the entries correspond to gray-scale values of pixels:



$$\Longrightarrow \begin{bmatrix} 0.9 \\ 0.3 \\ 0.5 \\ 0.6 \\ 0.3 \\ \vdots \end{bmatrix}.$$

The eigenface method then proceeds as explained before:

1. Form data matrix $B = [b_1, \ldots, b_n]$ of images.
2. Substract the mean image (average face)

$$\widehat{B} = B - B[1, \ldots, 1]^T/n.$$

3. Form the symmetric **covariance matrix**

$$\widehat{C} = \frac{1}{n-1}\widehat{B}\widehat{B}^T.$$

4. Compute orthogonal eigenvectors $u_1, u_2, \ldots$ of $\widehat{C}$. These are called **eigenfaces**. The principal component $u_1^T b_j$ has the largest variance for all images $b_1, \ldots, b_n$.
5. This can be used for face recognition: find an image b_j "closest" to a **test image** t by comparing $u_1^T t$ and $u_1^T b_j$.

Nondiagonalizable matrices

Recall:

Theorem (Thm. 6)

$A \in \mathbb{C}^{n\times n}$ is diagonalizable iff A has n linearly independent eigenvectors.

Theorem (Thm. 7)

A matrix with distinct eigenvalues is diagonalizable.

What can we say about matrices not similar to diagonal matrices?

Nondiagonalizable matrices

Recall:

Theorem (Thm. 6)

$A \in \mathbb{C}^{n \times n}$ is diagonalizable iff A has n linearly independent eigenvectors.

Theorem (Thm. 7)

A matrix with distinct eigenvalues is diagonalizable.

What can we say about matrices not similar to diagonal matrices?

- have less than n linearly independent eigenvectors,
- have multiple eigenvalues.

Example: $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$.

Jordan Canonical Form

What is the simplest form any matrix can take under similarity transform?

Marie Ennemond Camille Jordan (1838–1922)

Jordan Canonical Form

Theorem (Jordan canonical form, Thm. 8)

Any matrix $A \in \mathbb{C}^{n\times n}$ can be expressed in the Jordan canonical form

$$X^{-1}AX = J = \begin{bmatrix} J_1(\lambda_1) & & & \\ & J_2(\lambda_2) & & \\ & & \ddots & \\ & & & J_p(\lambda_p) \end{bmatrix},$$

$$J_k = J_k(\lambda_k) = \begin{bmatrix} \lambda_k & 1 & & \\ & \lambda_k & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_k \end{bmatrix} \in \mathbb{C}^{m_k\times m_k},$$

where X is nonsingular and $m_1 + m_2 + \cdots + m_p = n$.

Jordan Blocks

$$J_k = J_k(\lambda_k) = \begin{bmatrix} \lambda_k & 1 & & \\ & \lambda_k & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_k \end{bmatrix} \in \mathbb{C}^{m_k \times m_k}$$

For example for $m_k = 3$,

$$(J_k(\lambda_k) - \lambda_k I)x = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0 \;\Rightarrow\; x_2 = x_3 = 0$$

so that x is a multiple of $e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$.

Properties

(i) J has p Jordan blocks $\Leftrightarrow$ A has p lin. indep. eigenvectors.

(ii) The **algebraic multiplicity** of a given e'val λ is the sum of the dimensions of the Jordan blocks in which λ appears.

(iii) The **geometric multiplicity** of λ is

- the number of Jordan blocks associated with λ, or
- the number of linearly independent eigenvectors associated with λ or,
- $\dim(\mathrm{null}(A - \lambda I))$.

(iv) An eigenvalue λ is **defective** if it appears in a Jordan block of size greater than 1.
A is **defective** if it has a defective e'val $\Leftrightarrow$ A does not have a complete set of lin. indep. eigenvectors.

(v) The order of the largest Jordan block corresponding to λ is called **index** of λ.

Example 3

Find a Jordan matrix J of a matrix A such that

(a) $p(\lambda) = \det(\lambda I - A) = (\lambda - 1)^3(\lambda - 2)^4$,

(b) $\dim(\mathrm{null}(A - I)) = 2$ and $\dim(\mathrm{null}(A - 2I)) = 3$.

Example 4

Find the Jordan canonical form of

$$A = \begin{bmatrix} 2 & 0 & 0 & 0 \\ -3 & 2 & 0 & 1 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{bmatrix}.$$

Example 5

Determine the Jordan canonical form of a 14×14 matrix A having the following eigenvalues and sequences of ranks:

	$\mathrm{rank}(A - \lambda I)^k$				
k	1	2	3	4	5
$\lambda = 1$	11	10	9	9	9
$\lambda = 2$	12	10	10	10	10
$\lambda = 3$	12	11	10	9	9

Example 5

Determine the Jordan canonical form of a 14×14 matrix A having the following eigenvalues and sequences of ranks:

	$\mathrm{rank}(A - \lambda I)^k$				
k	1	2	3	4	5
$\lambda = 1$	11	10	9	9	9
$\lambda = 2$	12	10	10	10	10
$\lambda = 3$	12	11	10	9	9

Example 5

Determine the Jordan canonical form of a 14×14 matrix A having the following eigenvalues and sequences of ranks:

	$\operatorname{rank}(A - \lambda I)^k$				
k	1	2	3	4	5
$\lambda = 1$	11	10	9	9	9
$\lambda = 2$	12	10	10	10	10
$\lambda = 3$	12	11	10	9	9

Eigenvectors and Generalized Eigenvectors

$$X^{-1}AX = J \Leftrightarrow AX = XJ \qquad (5)$$

$$J = \begin{bmatrix} J_1(\lambda_1) & & \\ & \ddots & \\ & & J_p(\lambda_p) \end{bmatrix}, \; J_1(\lambda_1) = \begin{bmatrix} \lambda_1 & 1 & \\ & \ddots & 1 \\ & & \lambda_1 \end{bmatrix} \in \mathbb{C}^{m_1 \times m_1},$$

Equating 1st m_1 cols of (5) yields

$$Ax_1 = \lambda_1 x_1, \quad Ax_i = \lambda_1 x_i + x_{i-1}, \; i = 2, \ldots, m_1.$$

Eigenvectors and Generalized Eigenvectors

$$X^{-1}AX = J \Leftrightarrow AX = XJ \qquad (5)$$

$$J = \begin{bmatrix} J_1(\lambda_1) & & \\ & \ddots & \\ & & J_p(\lambda_p) \end{bmatrix}, \; J_1(\lambda_1) = \begin{bmatrix} \lambda_1 & 1 & \\ & \ddots & 1 \\ & & \lambda_1 \end{bmatrix} \in \mathbb{C}^{m_1 \times m_1},$$

Equating 1st m_1 cols of (5) yields

$$Ax_1 = \lambda_1 x_1, \quad Ax_i = \lambda_1 x_i + x_{i-1}, \; i = 2, \ldots, m_1.$$

- Cols $1, m_1 + 1, \ldots, m_1 + m_2 + \cdots + m_{p-1} + 1$ of X are eigenvectors of A and are linearly independent since X is nonsingular.
- The other cols of X are **generalized eigenvectors**.

Eigenvectors and Generalized Eigenvectors

$$X^{-1}AX = J \Leftrightarrow AX = XJ \qquad (5)$$

$$J = \begin{bmatrix} J_1(\lambda_1) & & \\ & \ddots & \\ & & J_p(\lambda_p) \end{bmatrix}, \; J_1(\lambda_1) = \begin{bmatrix} \lambda_1 & 1 & \\ & \ddots & 1 \\ & & \lambda_1 \end{bmatrix} \in \mathbb{C}^{m_1 \times m_1},$$

Equating 1st m_1 cols of (5) yields

$$Ax_1 = \lambda_1 x_1, \quad Ax_i = \lambda_1 x_i + x_{i-1}, \; i = 2, \ldots, m_1.$$

- Cols $1, m_1 + 1, \ldots, m_1 + m_2 + \cdots + m_{p-1} + 1$ of X are eigenvectors of A and are linearly independent since X is nonsingular.
- The other cols of X are **generalized eigenvectors**.

The vectors $x_1, x_2, \ldots, x_{m_1}$ are called a **Jordan chain**. The columns of X form p Jordan chains

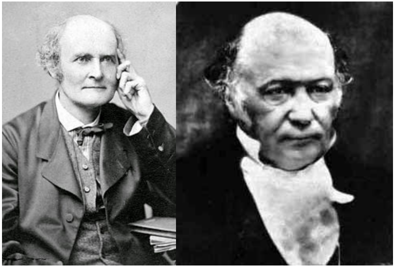
$$\{x_1, \ldots, x_{m_1}\}, \{x_{m_1+1}, \ldots, x_{m_1+m_2}\}, \ldots, \{x_{n-m_p+1}, \ldots, x_n\}.$$

Example 6

Determine the Jordan canonical form, the eigenvectors and generalized eigenvectors of

$$A = \begin{bmatrix} 6 & 2 & 2 \\ -2 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

Cayley–Hamilton Theorem



Arthur Cayley FRS (1821–1895)

Sir William Rowan Hamilton (1805–1865)

Theorem (Thm. 9)

Let p be the characteristic polynomial of an $n \times n$ matrix A. Then $p(A) = O$.

Proof:

The unique monic (leading coeff. $= 1$) polynomial q such that $q(A) = O$ is called the **minimal polynomial** of A.

Theorem (Thm. 10)

Let A be an $n \times n$ matrix with s distinct eigenvalues $\lambda_1, \dots, \lambda_s$. The minimal polynomial of A is

$$q(t) = \prod_{i=1}^{s} (t - \lambda_i)^{n_i},$$

where n_i is the dimension of the largest Jordan block in which λ_i appears ($=$ the index of λ_i).