Chapter 5 in a glance:

Let T be a linear operator whose standard matrix is A with size n x n. Then, a nonzero vector x is said to be an **eigenvector** of A and T if there exists a scalar λ such that

$$T(\mathbf{x}) = A\mathbf{x} = \lambda\mathbf{x} \quad .$$

The scalar λ is called an **eigenvalue** of T (or A).

A scalar t is an eigenvalue of A if and only if t satisfies the **characteristic equation** of A:

$$\det\ (A - tI) = 0$$

A matrix A is **diagonalizable** (i.e., $\exists P$ invertible, D diagonal such that $A = PDP^{-1}$) if and only if A has n linearly independent eigenvectors.

Chapter 6 review: (page 1 of 2)

1) A subset $\mathcal{S}$ of $\mathcal{R}^n$ is said to be an **orthogonal** set if

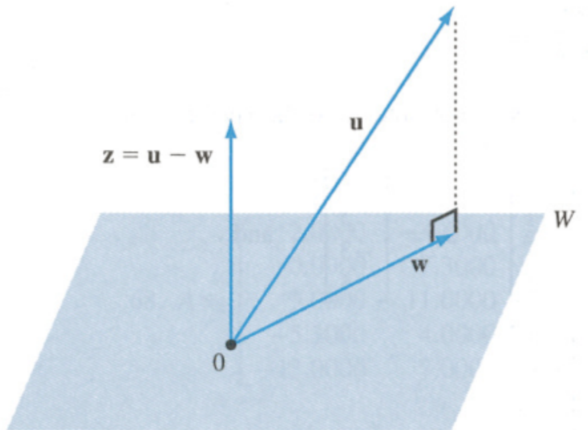
$$\forall \mathbf{u}_1, \mathbf{u}_2 \in \mathcal{S}, \mathbf{u}_1 \neq \mathbf{u}_2 \Rightarrow \mathbf{u}_1 \cdot \mathbf{u}_2 = 0.$$

An orthogonal set without zero vectors is linearly independent.

2) For a subspace W of $\mathcal{R}^n$, an **orthogonal basis** can be found by starting with any basis and performing **Gram-Schmidt Process**.

3) An **orthonormal** basis is an orthogonal basis whose vectors have unit norms.

4) Let W be a subspace of $\mathcal{R}^n$ with an orthonormal basis $\{\mathbf{w}_1, \mathbf{w}_2, \cdots, \mathbf{w}_k\}$. Then any vector $\mathbf{u} \in \mathcal{R}^n$ can be uniquely decomposed as $\mathbf{u} = \mathbf{w} + \mathbf{z}$ where $\mathbf{w} \in W$ and $\mathbf{z} \in W^{\perp}$.



orthogonal complement

5) The **orthogonal projection** $\mathbf{w}$ can be found as

$$\mathbf{w} = (\mathbf{u} \cdot \mathbf{w}_1)\mathbf{w}_1 + (\mathbf{u} \cdot \mathbf{w}_2)\mathbf{w}_2 + \cdots + (\mathbf{u} \cdot \mathbf{w}_k)\mathbf{w}_k$$

Chapter 6 review: (page 2 of 2)

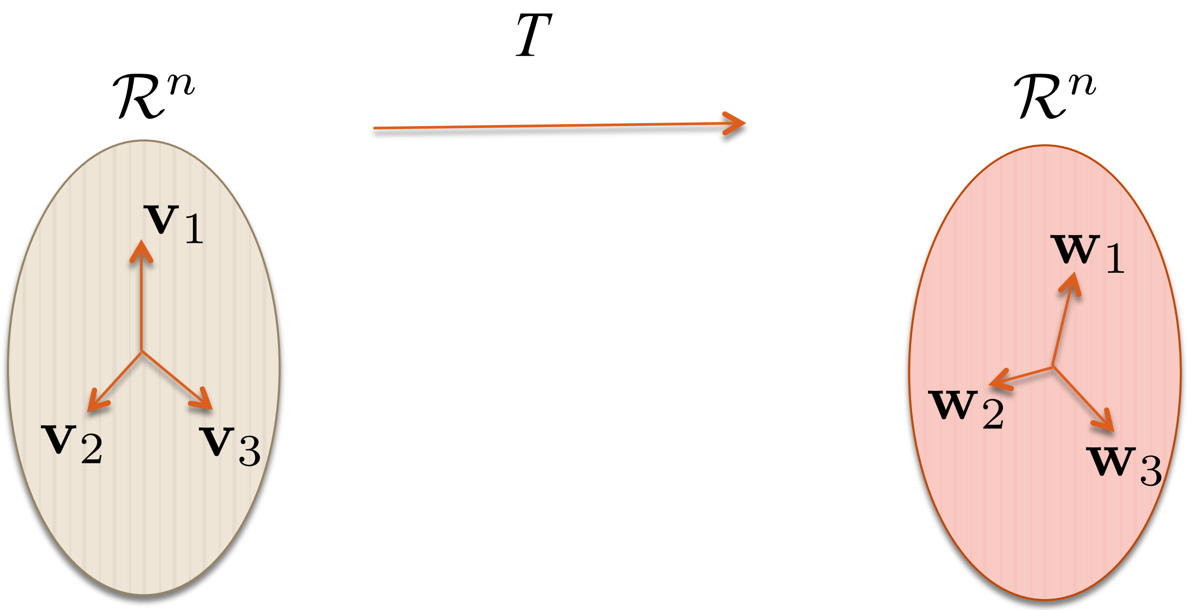
6) Another way of find the orthogonal projection of $\mathbf{u} \in \mathcal{R}^n$ on a subspace W (using a basis $\mathcal{B} = \{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_k\}$which is not required to be orthogonal or orthonormal):

$$\mathbf{w} = C(C^TC)^{-1}C^T\mathbf{u}$$

where C (n x k) contains the basis vectors of W.

7) An n x n matrix Q is called "**orthogonal**" if its columns form an **orthonormal basis** of $\mathcal{R}^n$. ($Q^TQ = I$)

8) A **linear operator** T on $\mathcal{R}^n$ is called **orthogonal** if its standard matrix is **orthogonal**. It is also "**norm-preserving**" ($||T(\mathbf{u})||^2 = ||\mathbf{u}||^2, \forall \mathbf{u} \in \mathcal{R}^n$) and preserving dot products ($\mathbf{u} \cdot \mathbf{v} = T(\mathbf{u}) \cdot T(\mathbf{v}), \forall \mathbf{u}, \mathbf{v} \in \mathcal{R}^n$).



What is a "good" basis?

1) From the point of view of a linear operator:
2) From the point of view of a subspace (or a vector space): orthogonal, or even orthonormal.

Section 6.6 Symmetric Matrices

Definition (in Chapter 2)

A square matrix A is called a **symmetric matrix** if $A^T = A$.

In this section, we will study some interesting properties of any symmetric matrix A.

In particular, we would like to learn properties of eigenvalues and eigenvectors of a symmetric matrix A.

Questions:
(1) Is an eigenvalue of A always real?
(2) Are any two eigenvectors of A corresponding to distinct eigenvalues always orthogonal?
(3) Is a symmetric matrix A always diagonalizable?

Section 6.6 Symmetric Matrices

Example: Consider $A = A^T = \begin{bmatrix} a & b \\ b & c \end{bmatrix} \in R^{2\times 2}$

$$\Rightarrow \det(tI_2 - A) = t^2 - (a+c)t + ac - b^2.$$

Since $(a+c)^2 - 4(ac - b^2) = (a-c)^2 + 4b^2 \geq 0$,

A has two real eigenvalues.

***Proposition:** If $A = A^T \in \mathscr{R}^{n\times n}$, then all eigenvalues of A are real.

Proof Let $A\mathbf{x} = \lambda\mathbf{x}$, where $\mathbf{x} = [\, x_1 \;\; x_2 \;\; \cdots \;\; x_n \,]^T \neq \mathbf{0}$. Since $a_{ij} = a_{ji}$,

$$\begin{aligned} \mathbf{x}^{*T} A\mathbf{x} &= \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} x_i^* x_j \\ &= \sum_{i=1}^{n} a_{ii}|x_i|^2 + \sum_{i=1}^{n}\sum_{j=1}^{i-1} a_{ij}\left(x_i^* x_j + x_j^* x_i\right) \end{aligned}$$

Theorem 6.14

If $\mathbf{u}, \mathbf{v} \in \mathcal{R}^n$ are eigenvectors of a **symmetric matrix** $A \in \mathcal{R}^{n \times n}$ that correspond to distinct eigenvalues, then $\mathbf{u}$ and $\mathbf{v}$ are orthogonal.

Proof

Theorem 6.14

If $\mathbf{u}, \mathbf{v} \in \mathcal{R}^n$ are eigenvectors of a **symmetric matrix** $A \in \mathcal{R}^{n\times n}$ that correspond to distinct eigenvalues, then $\mathbf{u}$ and $\mathbf{v}$ are orthogonal.

Proof Let $\mathbf{u}, \mathbf{v} \in \mathcal{R}^n$ be eigenvectors of A corresponding to eigenvalues λ, μ, respectively.

$$\Rightarrow A\mathbf{u}\cdot\mathbf{v} = \lambda\mathbf{u}\cdot\mathbf{v} = \lambda(\mathbf{u}\cdot\mathbf{v})$$

$$= \mathbf{u}\cdot A^T\mathbf{v} = \mathbf{v}\cdot A\mathbf{v} = \mathbf{u}\cdot\mu\mathbf{v} = \mu(\mathbf{u}\cdot\mathbf{v}).$$

$$\Rightarrow (\lambda - \mu)\,\mathbf{u}\cdot\mathbf{v} = 0 \Rightarrow \mathbf{u}\cdot\mathbf{v} = 0 \text{ since } \lambda - \mu \neq 0.$$

Theorem 6.14' (*)

Consider a matrix $A \in \mathcal{C}^{n\times n}$ that satisfies $A^H = A$. If $\mathbf{u}, \mathbf{v} \in \mathcal{C}^n$ are eigenvectors of A that correspond to distinct eigenvalues, then $\mathbf{u}$ and $\mathbf{v}$ are orthogonal.

* ***Proof*** Follow the proof for **Theorem 6.14**, with $\mathcal{R}$ changed to $\mathcal{C}$ and A^T changed to A^H.

Theorem 6.15

For a matrix $A \in \mathcal{R}^{n \times n}$, A is **symmetric** (i.e., $A = A^T$) if and only if there is an **orthonormal basis** for $\mathcal{R}^n$ consisting of eigenvectors of A, in which case there exists an **orthogonal** matrix P and a **diagonal** matrix D such that $P^T A P = D$.

Theorem 6.15

For a matrix $A \in \mathcal{R}^{n\times n}$, A is **symmetric** (i.e., $A = A^T$) if and only if there is an **orthonormal basis** for $\mathcal{R}^n$ consisting of eigenvectors of A, in which case there exists an **orthogonal** matrix P and a **diagonal** matrix D such that $P^T AP = D$.

Proof Sufficiency ("if"): $A = (P^T)^{-1}DP^{-1} = (P^{-1})^{-1}DP^T = PDP^T$

$$\Rightarrow A^T = (PDP^T)^T = (P^T)^T DP^T = PDP^T = A.$$

Necessity* ("only if"): By induction on n.

The necessity obviously holds for $n = 1$.

Assume the necessity holds for $n \geq 1$, and consider $A \in \mathcal{R}^{(n+1)\times(n+1)}$.

A has an eigenvector $\mathbf{b}_1 \in \mathcal{R}^{n+1}$ corresponding to a real eigenvalue λ.

$\Rightarrow$ $\mathbf{b}_1 \neq \mathbf{0}$, and $\exists$ an orthonormal basis $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_{n+1}\}$ for $\mathcal{R}^{n+1}$ by the **Extension Theorem** and Gram-Schmidt Process.

Let $B = [\ \mathbf{b}_1 \ \ \mathbf{b}_2 \ \ \cdots \ \ \mathbf{b}_{n+1}\ ]$. $\Rightarrow B$ is orthogonal and

$$B^T A B = \begin{bmatrix} \mathbf{b}_1^T \\ \mathbf{b}_2^T \\ \vdots \\ \mathbf{b}_{n+1}^T \end{bmatrix} \begin{bmatrix} A\mathbf{b}_1 & A\mathbf{b}_2 & \cdots & A\mathbf{b}_{n+1} \end{bmatrix} = \left[\begin{array}{c|ccc} \mathbf{b}_1^T A\mathbf{b}_1 & \mathbf{b}_1^T A\mathbf{b}_2 & \cdots & \mathbf{b}_1^T A\mathbf{b}_{n+1} \\ \hline \mathbf{b}_2^T A\mathbf{b}_1 & \mathbf{b}_2^T A\mathbf{b}_2 & \cdots & \mathbf{b}_2^T A\mathbf{b}_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{b}_{n+1}^T A\mathbf{b}_1 & \mathbf{b}_{n+1}^T A\mathbf{b}_2 & \cdots & \mathbf{b}_{n+1}^T A\mathbf{b}_{n+1} \end{array}\right]$$

$$= \left[\begin{array}{c|c} \lambda & \mathbf{0}^T \\ \hline \mathbf{0} & S \end{array}\right], \text{ since } \mathbf{b}_1^T A\mathbf{b}_1 = \lambda \mathbf{b}_1^T \mathbf{b}_1 = \lambda \text{ and } \mathbf{b}_j^T A\mathbf{b}_1 = \mathbf{b}_1^T A\mathbf{b}_j = 0\ \forall j \neq 1.$$

$S = S^T \in \mathcal{R}^{n\times n} \Rightarrow \exists$ an orthogonal $C \in \mathcal{R}^{n\times n}$ and a diagonal $L \in \mathcal{R}^{n\times n}$ such that $C^T S C = L$ by the induction hypothesis.

$$\Rightarrow \underbrace{\begin{bmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & C^T \end{bmatrix} B^T}_{\text{orthogonal } P^T} \underbrace{A B \begin{bmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & C \end{bmatrix}}_{\text{orthogonal } P} = \begin{bmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & C^T \end{bmatrix} \begin{bmatrix} \lambda & \mathbf{0}^T \\ \mathbf{0} & S \end{bmatrix} \begin{bmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & C \end{bmatrix} = \begin{bmatrix} \lambda & \mathbf{0}^T \\ \mathbf{0} & C^T S C \end{bmatrix} = \underbrace{\begin{bmatrix} \lambda & \mathbf{0}^T \\ \mathbf{0} & L \end{bmatrix}}_{\text{diagonal } D}$$

Theorem 6.15'

For a matrix $A \in \mathcal{C}^{n\times n}$, A is **Hermitian** (i.e., $A = A^H$) if and only if there is an orthonormal basis for $\mathcal{C}^n$ consisting of eigenvectors of A, in which case there exists an **unitary** matrix P and a diagonal matrix D such that $P^H AP = D$.

Proof Follow the proof for **Theorem 6.15**, with $\mathcal{R}$ changed to $\mathcal{C}$ and $(\cdot)^T$ changed to $(\cdot)^H$. Note B and C are unitary, and $L \in \mathcal{R}^{n\times n}$.

Finding an orthonormal basis consisting of eigenvectors of A where $A = A^T \in \mathcal{R}^{n\times n}$ or $A = A^H \in \mathcal{C}^{n\times n}$:

(1) Compute all distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_k$ of A.

(2) Determine the corresponding eigenspaces $\mathcal{E}_1, \mathcal{E}_2, \ldots, \mathcal{E}_k$.

(3) Get an orthonormal basis $\mathcal{B}_i$ for each $\mathcal{E}_i$.

(4) $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2 \cup \cdots \cup \mathcal{B}_k$ is an orthonormal basis for A.

Example: $A = \begin{bmatrix} 2 & -2 \\ -2 & 5 \end{bmatrix}$ has eigenvalues $\lambda_1 = 6$ and $\lambda_2 = 1$, with corresponding eigenspaces $\mathcal{E}_1 = \text{Span}\{[\ -1\ \ 2\ ]^T\}$ and $\mathcal{E}_2 = \text{Span}\{[\ 2\ \ 1\ ]^T\}$, respectively.

$\Rightarrow \mathcal{B}_1 = \{[\ -1\ \ 2\ ]^T/\sqrt{5}\}$ and $\mathcal{B}_2 = \{[\ 2\ \ 1\ ]^T/\sqrt{5}\}$

$\Rightarrow P^T A P = D$, where $P = \frac{1}{\sqrt{5}}\begin{bmatrix} -1 & 2 \\ 2 & 1 \end{bmatrix}$ and $D = \begin{bmatrix} 6 & 0 \\ 0 & 1 \end{bmatrix}$.

Example:

$$A = \begin{bmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{bmatrix}$$

Eigenvalues of A are $\lambda_1 = 2$ and $\lambda_2 = 8$.

$\Rightarrow \mathcal{E}_1 = \text{Span}\{[\,-1\;\;1\;\;0\,]^T, [\,-1\;0\;1\,]^T\}$ and

$\mathcal{E}_2 = \text{Span}\{[\,1\;\;1\;\;1\,]^T\}$.

$\Rightarrow$ Can apply the Gram-Schmidt process to find orthonormal bases for $\mathcal{E}_1$ and $\mathcal{E}_2$.

$$\Rightarrow \mathcal{B}_1 \cup \mathcal{B}_2 = \left\{ \begin{bmatrix} \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}, \begin{bmatrix} \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \\ \frac{-2}{\sqrt{6}} \end{bmatrix}, \begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix} \right\}$$

$\Rightarrow P^T A P = D$, where

$$P = \begin{bmatrix} \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{bmatrix} \quad \text{and} \quad D = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 8 \end{bmatrix}.$$

Example: Conic sections and quadratic forms

$$ax^2 + 2bxy + cy^2 + dx + ey + f = 0$$

1. Circle / ellipse
2. parabola
3. hyperbola

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \qquad \frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$$

How to determine, in general, the type of conic sections based on coefficients a, b, and c?

The associate quadratic form $ax^2 + 2bxy + cy^2$ of the quadratic form $ax^2 + 2bxy + cy^2 + dx + ey + f$ can be expressed as

$$\mathbf{v}^T A\mathbf{v} = (P\mathbf{v}')^T A(P\mathbf{v}') = (\mathbf{v}')^T P^T AP\mathbf{v}' = (\mathbf{v}')^T D\mathbf{v}' = \lambda_1(x')^2 + \lambda_2(y')^2,$$

with $\mathbf{v} = [\, x \;\; y \,]^T = P\mathbf{v}' = [\, x' \;\; y' \,]^T$,

$$A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}, \text{ and } P^T AP = D = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}.$$

$\Rightarrow$ Easy to judge the nature of the corresponding conic section.

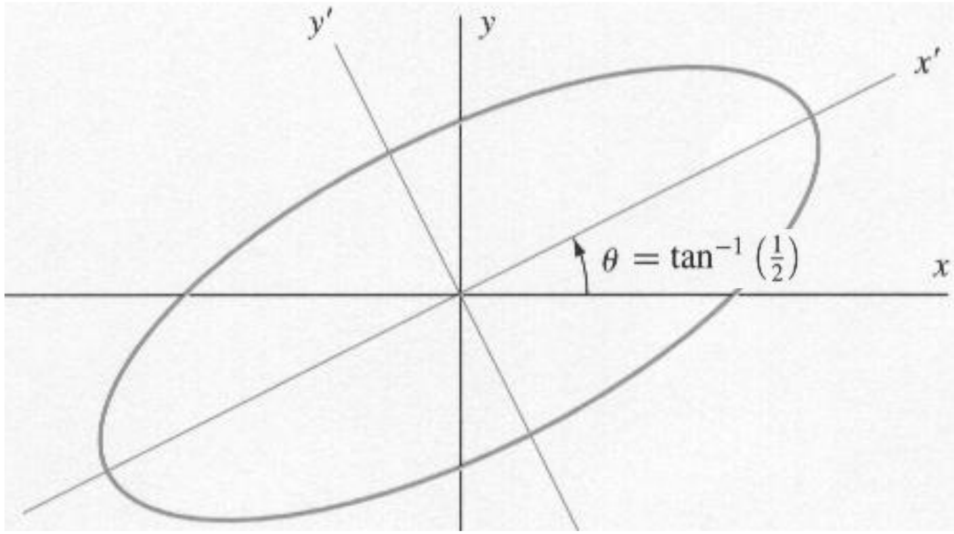
Example: $2x^2 - 4xy + 5y^2 - 36 = 0$

$$\Rightarrow A = \begin{bmatrix} 2 & -2 \\ -2 & 5 \end{bmatrix} \Rightarrow \lambda_1 = 1, \lambda_2 = 6 \text{ and } P = \frac{1}{\sqrt{5}}\begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix}.$$

$$x = \frac{2}{\sqrt{5}}x' - \frac{1}{\sqrt{5}}y'$$
$$y = \frac{1}{\sqrt{5}}x' + \frac{2}{\sqrt{5}}y'.$$

$\Rightarrow$ $2x^2 - 4xy + 5y^2 - 36 = (x')^2 + 6(y')^2 - 36$

$(x')^2 + 6(y')^2 - 36 = 0$ $\Rightarrow$



Note that
$$\underbrace{\begin{bmatrix} \pm 1 & 0 \\ 0 & \pm 1 \end{bmatrix}^T P^T}_{\text{new } P^T} A \underbrace{P \begin{bmatrix} \pm 1 & 0 \\ 0 & \pm 1 \end{bmatrix}}_{\text{new } P} = D = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$

$\Rightarrow$ Can always choose an orthogonal P with $\det P = 1$, a rotation.

Theorem 6.16 (Spectral Decomposition Theorem)

Let A be an $n \times n$ **symmetric** matrix, and let $\{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n\}$ be an orthonormal basis for $\mathcal{R}^n$ consisting of eigenvectors of A with corresponding eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_n$. Then there exist symmetric matrices $P_1, P_2, \cdots, P_n$ such that the following results hold:

(a) $A = \lambda_1 P_1 + \lambda_2 P_2 + \cdots + \lambda_n P_n$. : the spectral decomposition.

(b) rank $P_i = 1$ for all i.

(c) $P_i P_i = P_i$ for all i, and $P_i P_j = O$ if $i \neq j$.

(d) $P_i \mathbf{u}_i = \mathbf{u}_i$ for all i, and $P_i \mathbf{u}_j = \mathbf{0}$ if $i \neq j$.

Proof (a) Let $P = [\,\mathbf{u}_1\ \ \mathbf{u}_2\ \ \cdots\ \ \mathbf{u}_n\,]$ and $D = \text{diag}[\,\lambda_1\ \ \lambda_2\ \ \cdots\ \ \lambda_n\,]$.

$$\Rightarrow A = PDP^T = P[\,\lambda_1\mathbf{e}_1\ \ \lambda_2\mathbf{e}_2\ \ \cdots\ \ \lambda_n\mathbf{e}_n\,]P^T$$

$$= [\,\lambda_1 P\mathbf{e}_1\ \ \lambda_2 P\mathbf{e}_2\ \ \cdots\ \ \lambda_n P\mathbf{e}_n\,]P^T = [\,\lambda_1\mathbf{u}_1\ \ \lambda_2\mathbf{u}_2\ \ \cdots\ \ \lambda_n\mathbf{u}_n\,]P^T$$

$$= [\,\lambda_1\mathbf{u}_1\ \ \lambda_2\mathbf{u}_2\ \ \cdots\ \ \lambda_n\mathbf{u}_n\,]\begin{bmatrix}\mathbf{u}_1^T \\ \mathbf{u}_2^T \\ \vdots \\ \mathbf{u}_n^T\end{bmatrix} = \lambda_1 \underbrace{\mathbf{u}_1\mathbf{u}_1^T}_{P_1 = P_1^T} + \lambda_2 \underbrace{\mathbf{u}_2\mathbf{u}_2^T}_{P_2 = P_2^T} + \cdots + \lambda_n \underbrace{\mathbf{u}_n\mathbf{u}_n^T}_{P_n = P_n^T}.$$

(b) $\operatorname{rank} P_i = \operatorname{rank} \mathbf{u}_i\mathbf{u}_i^T = 1.$

(c) $P_iP_i = \mathbf{u}_i\mathbf{u}_i^T\mathbf{u}_i\mathbf{u}_i^T = (\mathbf{u}_i^T\mathbf{u}_i)\mathbf{u}_i\mathbf{u}_i^T = 1\cdot\mathbf{u}_i\mathbf{u}_i^T = P_i \quad \forall i;$

$P_iP_j = \mathbf{u}_i\mathbf{u}_i^T\mathbf{u}_j\mathbf{u}_j^T = (\mathbf{u}_i^T\mathbf{u}_j)\mathbf{u}_i\mathbf{u}_j^T = 0\cdot\mathbf{u}_i\mathbf{u}_j^T = O \quad \forall i \neq j.$

(d) $P_i\mathbf{u}_i = \mathbf{u}_i\mathbf{u}_i^T\mathbf{u}_i = (\mathbf{u}_i^T\mathbf{u}_i)\mathbf{u}_i = 1\cdot\mathbf{u}_i \quad \forall i;$

$P_i\mathbf{u}_j = \mathbf{u}_i\mathbf{u}_i^T\mathbf{u}_j = (\mathbf{u}_i^T\mathbf{u}_j)\mathbf{u}_i = 0\cdot\mathbf{u}_i = \mathbf{0} \quad \forall i \neq j.$

Example: $A = \begin{bmatrix} 3 & -4 \\ -4 & -3 \end{bmatrix}$ has eigenvalues $\lambda_1 = 5$ and $\lambda_2 = -5$.

An orthonormal basis consisting of eigenvectors of A is $\mathcal{B} = \{[\ -2\ \ 1\ ]^T/\sqrt{5}, [\ 1\ \ 2\ ]^T/\sqrt{5}\} = \{\mathbf{u}_1, \mathbf{u}_2\}$.

$\Rightarrow \quad P_1 = \mathbf{u}_1\mathbf{u}_1^T = \begin{bmatrix} \frac{4}{5} & -\frac{2}{5} \\ -\frac{2}{5} & \frac{1}{5} \end{bmatrix} \qquad P_2 = \mathbf{u}_2\mathbf{u}_2^T = \begin{bmatrix} \frac{1}{5} & \frac{2}{5} \\ \frac{2}{5} & \frac{4}{5} \end{bmatrix}$

$\Rightarrow \quad A = \lambda_1 P_1 + \lambda_2 P_2$

Homework Set for Section 6.6

- Section 6.6: Problems 15, 18, 19, 21, 23, 25, 43, 47, 48, 55, 56, 59, 61, 64