NATIONAL UNIVERSITY OF SINGAPORE
Department of Mathematics
MA1101R Linear Algebra I

2018-2019 (Semester 1) **Tutorial 10**

1. Let $\boldsymbol{A}$ be a symmetric matrix. If $\boldsymbol{u}$ and $\boldsymbol{v}$ are two eigenvectors of $\boldsymbol{A}$ associated with eigenvalues λ and μ, respectively, where $\lambda \neq \mu$, show that $\boldsymbol{u} \cdot \boldsymbol{v} = 0$ by following the following steps.

   (a) Show that $\boldsymbol{v}^T\boldsymbol{A} = \mu\boldsymbol{v}^T$.

   Since $\boldsymbol{v}$ is an eigenvector of $\boldsymbol{A}$ associated with eigenvalue μ, we have $\boldsymbol{A}\boldsymbol{v} = \mu\boldsymbol{v}$. By $\boldsymbol{A} = \boldsymbol{A}^T$, $\boldsymbol{v}^T\boldsymbol{A} = \boldsymbol{v}^T\boldsymbol{A}^T = (\boldsymbol{A}\boldsymbol{v})^T = \mu\boldsymbol{v}^T$.

   (b) Show that $\boldsymbol{v}^T\boldsymbol{A}\boldsymbol{u} = \mu\boldsymbol{v} \cdot \boldsymbol{u}$.

   By Part (a), $\boldsymbol{v}^T\boldsymbol{A}\boldsymbol{u} = (\boldsymbol{v}^T\boldsymbol{A})\boldsymbol{u} = \mu\boldsymbol{v} \cdot \boldsymbol{u}$.

   (c) Show that $\boldsymbol{v}^T\boldsymbol{A}\boldsymbol{u} = \lambda\boldsymbol{v} \cdot \boldsymbol{u}$.

   Since $\boldsymbol{u}$ is an eigenvector of $\boldsymbol{A}$ associated with eigenvalue λ, we have $\boldsymbol{A}\boldsymbol{u} = \lambda\boldsymbol{u}$. Then $\boldsymbol{v}^T\boldsymbol{A}\boldsymbol{u} = \boldsymbol{v}^T(\boldsymbol{A}\boldsymbol{u}) = \lambda\boldsymbol{v} \cdot \boldsymbol{u}$.

   (d) Show that $\boldsymbol{u} \cdot \boldsymbol{v} = 0$.

   By Parts (b) and (c), $\lambda\boldsymbol{v} \cdot \boldsymbol{u} = \mu\boldsymbol{v} \cdot \boldsymbol{u}$. As $\mu \neq \lambda$, we have $\boldsymbol{v} \cdot \boldsymbol{u} = 0$.

2. Let

$$\boldsymbol{A} = \begin{pmatrix} b & a & a \\ a & b & a \\ a & a & b \end{pmatrix}.$$

   (a) Show that $\boldsymbol{A}$ has eigenvalues $b - a$ and $2a + b$.

   Since

$$\begin{aligned} &\det(\lambda\boldsymbol{I} - \boldsymbol{A}) = \lambda^3 - 3b\lambda^2 - 3a^2\lambda + 3b^2\lambda - 2a^3 + 3a^2b - b^3 \\ =&(\lambda - (b-a))^2(\lambda - (2a+b)), \end{aligned}$$

   the eigenvalues are $b - a$ and $2a + b$.

   (b) Find an orthogonal basis of the eigenspace E_{b-a}.

   For $\lambda = b - a$, the linear system $(\lambda\boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{0}$ is

$$\begin{pmatrix} -a & -a & -a \\ -a & -a & -a \\ -a & -a & -a \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Solving the linear system, by the assumption $a \neq 0$, we have

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = t \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} + s \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$$

where t and s are arbitrary parameters. So $\{(-1, 1, 0)^T, (-1, 0, 1)^T\}$ is a basis for E_{b-a}. By Gram-Schmidt Process, we obtain an orthonormal basis

$$\{(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0)^T, (-\frac{1}{\sqrt{6}}, -\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}})^T\}$$

for E_{b-a}.

If $a = 0$, we may choose the standard basis.

(c) Find an orthogonal basis of the eigenspace E_{2a+b}.

For $\lambda = 2a + b$, the linear system $(\lambda \boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{0}$ is

$$\begin{pmatrix} 2a & -a & -a \\ -a & 2a & -a \\ -a & -a & 2a \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Solving the linear system, by the assumption $a \neq 0$, we have

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = t \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

where t is arbitrary parameters. So $\{(1, 1, 1)^T\}$ is a basis for E_{2a+b}. By Gram-Schmidt Process, we obtain an orthonormal basis $\{(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}})^T\}$ for E_{2a+b}.

If $a = 0$, we may choose the standard basis.

(d) Find a matrix $\boldsymbol{P}$ that orthogonally diagonalize $\boldsymbol{A}$. and determine $\boldsymbol{P}^T\boldsymbol{A}\boldsymbol{P}$.

Form the matrix $\boldsymbol{P}$ by using the orthonormal bases as the column vectors:

$$\boldsymbol{P} = \begin{pmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{pmatrix} \text{ and } \boldsymbol{P}^T\boldsymbol{A}\boldsymbol{P} = \begin{pmatrix} b-a & 0 & 0 \\ 0 & b-a & 0 \\ 0 & 0 & 2a+b \end{pmatrix} = \begin{pmatrix} b & a & a \\ a & b & a \\ a & a & b \end{pmatrix}.$$

3. Let

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

(a) Compute $\boldsymbol{A}^4$.

Following Question (2), we have

$$\boldsymbol{P}^T\boldsymbol{A}\boldsymbol{P} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{pmatrix},$$

where

$$\boldsymbol{P} = \begin{pmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{pmatrix}$$

is obtained in Question (2). By $\boldsymbol{P}^T = \boldsymbol{P}^{-1}$,

$$\boldsymbol{A}^4 = \boldsymbol{P} \begin{pmatrix} 1^4 & 0 & 0 \\ 0 & 1^4 & 0 \\ 0 & 0 & 4^4 \end{pmatrix} \boldsymbol{P}^T = \begin{pmatrix} 86 & 85 & 85 \\ 85 & 86 & 85 \\ 85 & 85 & 86 \end{pmatrix}.$$

Remark: To obtain A^4, we can compute the matrices multiplication. Or by Question (2), note that

$$\boldsymbol{P} \begin{pmatrix} b-a & 0 & 0 \\ 0 & b-a & 0 \\ 0 & 0 & 2a+b \end{pmatrix} \boldsymbol{P}^T = \begin{pmatrix} b & a & a \\ a & b & a \\ a & a & b \end{pmatrix}.$$

Applying this matrix equation to our case, let

$$\begin{pmatrix} b-a & 0 & 0 \\ 0 & b-a & 0 \\ 0 & 0 & 2a+b \end{pmatrix} = \begin{pmatrix} 1^4 & 0 & 0 \\ 0 & 1^4 & 0 \\ 0 & 0 & 4^4 \end{pmatrix}.$$

We establish a linear system

$$\begin{cases} b - a = 1 \\ 2a + b = 256 \end{cases}$$

Solve for $a = 85$ and $b = 86$ and then obtain the same result.

Note that the second method heavily relies on the symmetry of the matrix $\boldsymbol{A}$, which is a shortcut to compute $\boldsymbol{A}^4$. However, it is not always the case.

(b) Find a matrix $\boldsymbol{B}$ such that $\boldsymbol{B}^2 = \boldsymbol{A}$.

Similar to Part (a), since

$$\begin{pmatrix} \pm 1 & 0 & 0 \\ 0 & \pm 1 & 0 \\ 0 & 0 & \pm 2 \end{pmatrix}^2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{pmatrix}$$

we may take

$$\boldsymbol{B} = \boldsymbol{P} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \boldsymbol{P}^T = \begin{pmatrix} \frac{4}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{4}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{4}{3} \end{pmatrix}.$$

4. A square matrix $(a_{ij})_{n\times n}$ is called a stochastic matrix if all the entries are non-negative and the sum of entries of each column is 1, i.e. $a_{1i} + a_{2i} + \cdots + a_{ni} = 1$ for $i = 1, 2, \ldots, n$.

   Let $\boldsymbol{A}$ be a stochastic matrix.

   (a) Show that $(1, 1, \ldots, 1)^T$ is an eigenvector of $\boldsymbol{A}^T$.

   Let $\boldsymbol{u}$ be the row matrix $(1, 1, \ldots, 1)$. Then

$$\boldsymbol{uA} = (\sum_{i=1}^{n} a_{i1}, \sum_{i=1}^{n} a_{i2}, \ldots, \sum_{i=1}^{n} a_{in}) = \boldsymbol{u},$$

   that is, $\boldsymbol{A}^T\boldsymbol{u}^T = \boldsymbol{u}^T$. Then $(1, 1, \ldots, 1)^T$ is an eigenvector of $\boldsymbol{A}^T$.

   (b) Show that 1 is an eigenvalue of $\boldsymbol{A}$.

   Since $\boldsymbol{A}^T\boldsymbol{u}^T = \boldsymbol{u}^T$, 1 is an eigenvector of $\boldsymbol{A}^T$. Since $\boldsymbol{A}$ and $\boldsymbol{A}^T$ have the same characteristic polynomials (see Homework 4), 1 is an eigenvalue of $\boldsymbol{A}$.

   (c) Show that $\boldsymbol{A}^k$ for $k \geq 0$ is a stochastic matrix.

   Let $\boldsymbol{u}$ be the row matrix $(1, 1, \ldots, 1)$. By definition, for a square matrix $\boldsymbol{B}$, $\boldsymbol{B}$ is a stochastic matrix if and only if $\boldsymbol{uB} = \boldsymbol{u}$. By $\boldsymbol{uA} = \boldsymbol{u}$, we have

$$\boldsymbol{uA}^k = \boldsymbol{uA}^{k-1} = \cdots = \boldsymbol{u}.$$

   For $k = 0$, $\boldsymbol{A} = I_n$, which is a stochastic matrix. Therefore, for all $k \geq 0$, $\boldsymbol{A}^k$ is a stochastic matrix.

   (d) Let $\boldsymbol{u} = (u_1, u_2, \ldots, u_n)^T$ be an eigenvector of $\boldsymbol{A}^T$ associated an eigenvalue λ and denote

$$u_{\max} = \max\{u_1, u_2, \ldots, u_n\} \text{ and } u_{\min} = \min\{u_1, u_2, \ldots, u_n\},$$

   that is, $u_{\max}$ and $u_{\min}$ are the maximum and minimum elements in the set $\{u_1, u_2, \ldots, u_n\}$, respectively.
   Show that $\lambda u_j \leq u_{\max}$ and $u_{\min} \leq \lambda u_j$ for all $1 \leq j \leq n$.

   Since $\boldsymbol{u} = (u_1, u_2, \ldots, u_n)^T$ be an eigenvector of $\boldsymbol{A}^T$ associated to eigenvalue λ, we have

$$\boldsymbol{A}^T\boldsymbol{u} = (\sum_{i=1}^{n} a_{i1}u_i, \sum_{i=1}^{n} a_{i2}u_i, \ldots, \sum_{i=1}^{n} a_{in}u_i)^T = \lambda(u_1, u_2, \ldots, u_n)^T. \quad (1)$$

   Since all a_{ij} are nonnegative,

$$u_{\min} = (\sum_{i=1}^{n} a_{ij})u_{\min} \leq \sum_{i=1}^{n} a_{ij}u_i \leq (\sum_{i=1}^{n} a_{ij})u_{\max} = u_{\max}.$$

   By Equation (1), $\sum_{i=1}^{n} a_{ij}u_i = \lambda u_j$. Thus for all $1 \leq j \leq n$, $u_{\min} \leq \lambda u_j \leq u_{\max}$..

(e) If λ is an eigenvalue of $\boldsymbol{A}$, then $\lambda \leq 1$.

Since $\boldsymbol{A}$ and $\boldsymbol{A}^T$ have the same characteristic function and λ is an eigenvalue of $\boldsymbol{A}$, λ is also an eigenvalue of $\boldsymbol{A}^T$.

By Part (d), $\lambda u_{\max} \leq u_{\max}$ and $u_{\min} \leq \lambda u_{\min}$. If $u_{\max} \geq 0$, then $\lambda \leq 1$ by $\lambda u_{\max} \leq u_{\max}$; If $u_{\max} < 0$, then $u_{\min} \leq u_{\max} < 0$ and $\lambda \leq 1$ by $u_{\min} \leq \lambda u_{\min}$. In all cases, we have $\lambda \leq 1$.

(f) If λ is an eigenvalue of $\boldsymbol{A}$, then $|\lambda| \leq 1$. (Hint: Apply Question (4e) to $\boldsymbol{A}^2$.)

Since λ^2 is an eigenvalue of $\boldsymbol{A}^2$ and $\boldsymbol{A}^2$ is a stochastic matrix (see Part (c)), we have $\lambda^2 \leq 1$ as above by Part (e). Hence $|\lambda| \leq 1$.

5. (This is an induction step for proving Remark 6.2.5.3.) Let $\boldsymbol{A}$ be a square matrix of order n. By Theorem 6.2.3, to diagonalize $\boldsymbol{A}$, we need to find n linearly independent eigenvectors.

Suppose we already have $m(< n)$ linearly independent eigenvectors $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_m$, say, $\boldsymbol{A}\boldsymbol{u}_i = \lambda_i \boldsymbol{u}_i$ for $i = 1, 2, \ldots, m$ where $\lambda_1, \lambda_2, \ldots, \lambda_m$ are not necessarily distinct. For a new eigenvalue μ ($\mu \neq \lambda_i$ for $i = 1, 2, \ldots, m$) of $\boldsymbol{A}$, let $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_p\}$ be a basis for the eigenspace E_μ. Prove that $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_m, \boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_p\}$ is linearly independent.

(Hint: Consider the vector equation

$$a_1\boldsymbol{u}_1 + a_2\boldsymbol{u}_2 + \cdots + a_m\boldsymbol{u}_m + b_1\boldsymbol{v}_1 + b_2\boldsymbol{v}_2 + \cdots + b_p\boldsymbol{v}_p = \boldsymbol{0}.$$

By using the property of eigenvectors, show that

$$a_1(\lambda_1 - \mu)\boldsymbol{u}_1 + a_2(\lambda_2 - \mu)\boldsymbol{u}_2 + \cdots + a_m(\lambda_m - \mu)\boldsymbol{u}_m = \boldsymbol{0}.$$

Then make use of the linearly independent assumption on $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_m$, as well as $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_p$, to finish the proof.)

Let a_i and b_j for $1 \leq i \leq m$ and $1 \leq j \leq p$ be the constant such that

$$a_1\boldsymbol{u}_1 + a_2\boldsymbol{u}_2 + \cdots + a_m\boldsymbol{u}_m + b_1\boldsymbol{v}_1 + b_2\boldsymbol{v}_2 + \cdots + b_p\boldsymbol{v}_p = \boldsymbol{0}. \tag{2}$$

Taking the pre-multiplication of $\boldsymbol{A}$ to the both sides of (2), we have

$$\begin{aligned} & a_1\boldsymbol{A}\boldsymbol{u}_1 + a_2\boldsymbol{A}\boldsymbol{u}_2 + \cdots + a_m\boldsymbol{A}\boldsymbol{u}_m + b_1\boldsymbol{A}\boldsymbol{v}_1 + b_2\boldsymbol{A}\boldsymbol{v}_2 + \cdots + b_p\boldsymbol{A}\boldsymbol{v}_p = \boldsymbol{A}\boldsymbol{0} \\ \Rightarrow & a_1\lambda\boldsymbol{u}_1 + a_2\lambda\boldsymbol{u}_2 + \cdots + a_m\lambda\boldsymbol{u}_m + b_1\mu\boldsymbol{v}_1 + b_2\mu\boldsymbol{v}_2 + \cdots + b_p\mu\boldsymbol{v}_p = \boldsymbol{0}. \end{aligned} \tag{3}$$

Adding $-\mu$ times of Equation (2) to Equation (3), we have

$$a_1(\lambda_1 - \mu)\boldsymbol{u}_1 + a_2(\lambda_2 - \mu)\boldsymbol{u}_2 + \cdots + a_m(\lambda_m - \mu)\boldsymbol{u}_m = \boldsymbol{0}.$$

Since $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_m$ are linearly independent,

$$a_1(\lambda_1 - \mu) = a_2(\lambda_2 - \mu) = \cdots = a_m(\lambda_m - \mu) = 0.$$

As $\mu \neq \lambda_i$, $a_i = 0$ for all $1 \leq i \leq m$.

Plugging $a_i = 0$ into (2), we have

$$b_1\boldsymbol{v}_1 + b_2\boldsymbol{v}_2 + \cdots + b_p\boldsymbol{v}_p = \mathbf{0}.$$

Since $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_p$ are linearly independent, we have $b_j = 0$ for all $1 \leq j \leq p$. Therefore $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_m, \boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_p\}$ is linearly independent.