5. Let $\boldsymbol{B}$ be any $m \times m$ matrix. Suppose $\boldsymbol{P}$ is an invertible $m \times m$ matrix and that $\boldsymbol{C} = \boldsymbol{P}^{-1}\boldsymbol{B}\boldsymbol{P}$.

(a) Show that if $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$ is a basis for the nullspace of $\boldsymbol{C}$, then $\{\boldsymbol{P}\boldsymbol{v}_1, \ldots, \boldsymbol{P}\boldsymbol{v}_k\}$ is a basis for the nullspace of $\boldsymbol{B}$.

(b) Prove that $\operatorname{rank}(\boldsymbol{B}) = \operatorname{rank}(\boldsymbol{C})$.

(c) Let $\boldsymbol{A}$ be any $m \times p$ matrix. Suppose that the linear system $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ is consistent for every $\boldsymbol{b} \in \mathbb{R}^m$. Prove that the linear system $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{y} = \boldsymbol{0}$ has only the trivial solution.

Proof. (a) If $\boldsymbol{v}$ belongs to the nullspace of $\boldsymbol{B}$, i.e., $\boldsymbol{B}\boldsymbol{v} = \boldsymbol{0}$. Then

$$\boldsymbol{C}\boldsymbol{P}^{-1}\boldsymbol{v} = \boldsymbol{P}^{-1}\boldsymbol{B}\boldsymbol{v} = \boldsymbol{P}^{-1}0 = 0,$$

i.e., $\boldsymbol{P}^{-1}\boldsymbol{v}$ belongs to the nullspace of $\boldsymbol{C}$. There exist unique constants $c_1, \ldots, c_k$ such that

$$c_1\boldsymbol{v}_1 + \cdots + c_k\boldsymbol{v}_k = \boldsymbol{P}^{-1}\boldsymbol{v}.$$

Hence,

$$\boldsymbol{v} = \boldsymbol{P}(c_1\boldsymbol{v}_1 + \cdots + c_k\boldsymbol{v}_k) = c_1\boldsymbol{P}\boldsymbol{v}_1 + \cdots + c_k\boldsymbol{P}\boldsymbol{v}_k.$$

On the other hand, if there exist $d_1, \ldots, d_k \in \mathbb{R}$ such that

$$\boldsymbol{v} = d_1\boldsymbol{P}\boldsymbol{v}_1 + \cdots + d_k\boldsymbol{P}\boldsymbol{v}_k,$$

then

$$\boldsymbol{P}^{-1}\boldsymbol{v} = \boldsymbol{P}^{-1}(d_1\boldsymbol{P}\boldsymbol{v}_1 + \cdots + d_k\boldsymbol{P}\boldsymbol{v}_k) = d_1\boldsymbol{v}_1 + \cdots + d_k\boldsymbol{v}_k.$$

Since $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$ is a basis for the nullspace of $\boldsymbol{C}$, the two representations of $\boldsymbol{P}^{-1}\boldsymbol{v}$ are the same, i.e., $c_i = d_i$, $i = 1, \ldots, k$.

Since every vector in the nullspace of $\boldsymbol{B}$ can be uniquely written as a linear combination of $\{\boldsymbol{P}\boldsymbol{v}_1, \ldots, \boldsymbol{P}\boldsymbol{v}_k\}$, we conclude that $\{\boldsymbol{P}\boldsymbol{v}_1, \ldots, \boldsymbol{P}\boldsymbol{v}_k\}$ is a basis for the nullspace of $\boldsymbol{B}$.

(b) By (a), if $\operatorname{nullity}(\boldsymbol{C}) = k$, then $\operatorname{nullity}(\boldsymbol{B}) = k$; so $\operatorname{rank}(\boldsymbol{B}) = m - k = \operatorname{rank}(\boldsymbol{C})$.

(c) The system $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ is consistent if and only if $\boldsymbol{b}$ belongs to the column space of $\boldsymbol{A}$. If $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ is consistent for all $\boldsymbol{b} \in \mathbb{R}^m$, then the column space of $\boldsymbol{A}$ is $\mathbb{R}^m$. So

$$\operatorname{nullity}(\boldsymbol{A}^{\mathrm{T}}) = m - \operatorname{rank}(\boldsymbol{A}^{\mathrm{T}}) = m - \operatorname{rank}(\boldsymbol{A}) = m - m = 0.$$

Therefore, $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{y} = \boldsymbol{0}$ has only the trivial solution. $\square$

6. Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be $n \times n$ matrices.

(a) Show that if there is an invertible $n \times n$ matrix $\boldsymbol{P}$ such that $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ and $\boldsymbol{P}^{-1}\boldsymbol{B}\boldsymbol{P}$ are both diagonal matrices, then $\boldsymbol{A}\boldsymbol{B} = \boldsymbol{B}\boldsymbol{A}$.

(b) Assume that $\boldsymbol{A}$ has n distinct eigenvalues. Show that if every eigenvector of $\boldsymbol{A}$ is an eigenvector of $\boldsymbol{B}$ (possibly associated with a different eigenvalue), then $\boldsymbol{AB} = \boldsymbol{BA}$.

(c) Assume again that $\boldsymbol{A}$ has n distinct eigenvalues. Show that if $\boldsymbol{AB} = \boldsymbol{BA}$, then every eigenvector of $\boldsymbol{A}$ is an eigenvector of $\boldsymbol{B}$ (possibly associated with a different eigenvalue).

Proof. (a) Suppose that $\boldsymbol{P}^{-1}\boldsymbol{AP} = \boldsymbol{D}_1$ and $\boldsymbol{P}^{-1}\boldsymbol{BP} = \boldsymbol{D}_2$ are diagonal matrices. Then

$$\begin{aligned}\boldsymbol{AB} &= (\boldsymbol{PD}_1\boldsymbol{P}^{-1})(\boldsymbol{PD}_2\boldsymbol{P}^{-1}) = \boldsymbol{PD}_1\boldsymbol{D}_2\boldsymbol{P}^{-1} = \boldsymbol{PD}_2\boldsymbol{D}_1\boldsymbol{P}^{-1} \\ &= (\boldsymbol{PD}_2\boldsymbol{P}^{-1})(\boldsymbol{PD}_1\boldsymbol{P}^{-1}) = \boldsymbol{BA}.\end{aligned}$$

(b) Let $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ be eigenvectors of $\boldsymbol{A}$ associated to the n distinct eigenvalues. Then they are linearly independent. Let

$$\boldsymbol{P} = \begin{pmatrix}\boldsymbol{v}_1 & \cdots & \boldsymbol{v}_n\end{pmatrix}.$$

Then $\boldsymbol{P}$ is invertible and $\boldsymbol{P}^{-1}\boldsymbol{AP}$ is a diagonal matrix. By assumption, $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ are also eigenvectors of $\boldsymbol{B}$. Then $\boldsymbol{P}^{-1}\boldsymbol{BP}$ is also a diagonal matrix. By (a), $\boldsymbol{AB} = \boldsymbol{BA}$.

(c) Let $\boldsymbol{v}$ be an eigenvector of $\boldsymbol{A}$ associated to the eigenvalue λ, i.e., $\boldsymbol{Av} = \lambda\boldsymbol{v}$ with $\boldsymbol{v} \neq \boldsymbol{0}$. Since $\boldsymbol{AB} = \boldsymbol{BA}$, we have

$$\boldsymbol{A}(\boldsymbol{Bv}) = \boldsymbol{BAv} = \boldsymbol{B}(\lambda\boldsymbol{v}) = \lambda(\boldsymbol{Bv}).$$

Then $\boldsymbol{Bv}$ belongs to the eigenspace of $\boldsymbol{A}$ associated to λ.

Since $\boldsymbol{A}$ has n distinct eigenvalues, the dimension of each eigenspace is 1. In particular, the eigenspace of $\boldsymbol{A}$ associated to λ is $\operatorname{span}\{\boldsymbol{v}\}$. Hence, $\boldsymbol{Bv} = \mu\boldsymbol{v}$ for some constant μ.

Note that $\boldsymbol{v} \neq \boldsymbol{0}$. Then $\boldsymbol{v}$ is an eigenvector of $\boldsymbol{B}$ associated to the eigenvalue μ. □

2018/2019 Semester 1

1(b) Suppose $\boldsymbol{D}$ is a matrix with k columns such that the linear system $\boldsymbol{Dx} = \boldsymbol{r}$ is consistent for all vectors $\boldsymbol{r} \in \mathbb{R}^n$. For each of the statements below, determine if the statement is true. Justify your answer.

(i) $\boldsymbol{D}$ has n rows.
(ii) k is at least n.
(iii) $\boldsymbol{D}$ is of full rank.

Solution. Note that $\boldsymbol{D}$ and $\boldsymbol{Dx} = \boldsymbol{r}$ have the same number of rows. Then $\boldsymbol{D}$ has n rows.

The system $\boldsymbol{Dx} = \boldsymbol{r}$ is consistent if and only if $\boldsymbol{r}$ belongs to the column space of $\boldsymbol{D}$. Since $\boldsymbol{Dx} = \boldsymbol{r}$ is consistent for all $\boldsymbol{r} \in \mathbb{R}^n$, the column space of $\boldsymbol{D}$ is $\mathbb{R}^n$. So $\operatorname{rank}(\boldsymbol{D}) = n$.

From $\operatorname{rank}(\boldsymbol{D}) \leq k$, we have $n \leq k$.

Therefore, (i), (ii) and (iii) are all true. □

2(c) Let $\boldsymbol{X}, \boldsymbol{Y}$ be square matrices of the same order. Prove the following statements.

(i) $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X} = \boldsymbol{0}$ if and only if $\boldsymbol{X} = \boldsymbol{0}$.
(ii) $\boldsymbol{X}\boldsymbol{Y} = \boldsymbol{0}$ if and only if $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{Y} = \boldsymbol{0}$.

Proof. (i) If $\boldsymbol{X} = \boldsymbol{0}$, it is clear that $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X} = \boldsymbol{0}$.

Conversely, suppose $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X} = \boldsymbol{0}$. Let $\boldsymbol{v}_i$ be the i^{th} column of $\boldsymbol{X}$. Then the (i,j)-entry of $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}$ is $\boldsymbol{v}_i^{\mathrm{T}}\boldsymbol{v}_j$. In particular, its (i,i)-entries $\boldsymbol{v}_i^{\mathrm{T}}\boldsymbol{v}_i = \|\boldsymbol{v}_i\|^2 = 0$. Hence, $\boldsymbol{v}_i = \boldsymbol{0}$ for all i. So $\boldsymbol{X} = \boldsymbol{0}$.

(ii) If $\boldsymbol{X}\boldsymbol{Y} = \boldsymbol{0}$, it is clear that $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{Y} = \boldsymbol{0}$.

Conversely, if $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{Y} = \boldsymbol{0}$, then $(\boldsymbol{X}\boldsymbol{Y})^{\mathrm{T}}(\boldsymbol{X}\boldsymbol{Y}) = \boldsymbol{Y}^{\mathrm{T}}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{Y} = \boldsymbol{0}$; by (i) $\boldsymbol{X}\boldsymbol{Y} = \boldsymbol{0}$. □

5(b) Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be square matrices of order n. Suppose $\boldsymbol{AB} = \boldsymbol{BA}$ and $\boldsymbol{A}$ has n distinct eigenvalues.

(i) Show that each eigenspace of $\boldsymbol{A}$ has dimension 1.
(ii) Show that if $\boldsymbol{u}$ is an eigenvector of $\boldsymbol{A}$, then $\boldsymbol{u}$ is also an eigenvector of $\boldsymbol{B}$.
(iii) Show that $\boldsymbol{A}$ and $\boldsymbol{B}$ are simultaneously diagonalizable, i.e., there exists an invertible matrix $\boldsymbol{P}$ such that $\boldsymbol{PAP}^{-1}$ and $\boldsymbol{PBP}^{-1}$ are diagonal.

Proof. (i) Suppose $\boldsymbol{A}$ has n distinct eigenvalues $\lambda_1, \dots, \lambda_n$. Then $\dim E_{\lambda_i} \geq 1$, $i = 1, \dots, n$. On the other hand,

$$n \geq \dim E_{\lambda_1} + \cdots + \dim E_{\lambda_n} \geq \underbrace{1 + \cdots + 1}_{n} = n.$$

We must have $\dim E_{\lambda_i} = 1$ for all $i = 1, \dots, n$.

(ii) If $\boldsymbol{u}$ is an eigenvector of $\boldsymbol{A}$ associated to the eigenvalue λ, then $\boldsymbol{Au} = \lambda\boldsymbol{u}$ with $\boldsymbol{u} \neq \boldsymbol{0}$. So

$$\boldsymbol{A}(\boldsymbol{Bu}) = \boldsymbol{BAu} = B(\lambda\boldsymbol{u}) = \lambda(\boldsymbol{Bu}).$$

It follows that $\boldsymbol{Bu}$ belongs to the eigenspace $E_{\boldsymbol{A},\lambda}$ of $\boldsymbol{A}$ associated to the eigenvalue λ. By (i) $\dim E_{\boldsymbol{A},\lambda} = 1$; so $E_{\boldsymbol{A},\lambda} = \operatorname{span}\{\boldsymbol{u}\}$. Then $\boldsymbol{Bu} = \mu\boldsymbol{u}$ for some constant μ. We conclude that $\boldsymbol{u}$ is an eigenvector of $\boldsymbol{B}$ associated to the eigenvalue μ.

(iii) Let $\boldsymbol{v}_i$ be an eigenvector of $\boldsymbol{A}$ associated to λ_i, $i = 1, \dots, n$. Then $\boldsymbol{v}_1, \dots, \boldsymbol{v}_n$ are linearly independent. Let

$$\boldsymbol{Q} = \begin{pmatrix} \boldsymbol{v}_1 & \cdots & \boldsymbol{v}_n \end{pmatrix}.$$

Then $\boldsymbol{Q}$ is invertible and $\boldsymbol{Q}^{-1}\boldsymbol{AQ}$ is a diagonal matrix. By (ii), each $\boldsymbol{v}_i$ is also an eigenvector of $\boldsymbol{B}$. Then $\boldsymbol{Q}^{-1}\boldsymbol{BQ}$ is also a diagonal matrix. Let $\boldsymbol{P} = \boldsymbol{Q}^{-1}$. Then $\boldsymbol{P}$ is invertible and both $\boldsymbol{PAP}^{-1}$ and $\boldsymbol{PBP}^{-1}$ are diagonal matrices. □

5. Let $\boldsymbol{A}$ be an $n \times n$ matrix.

(a) Show that for any $\boldsymbol{u}, \boldsymbol{w} \in \mathbb{R}^n$, $(\boldsymbol{A}\boldsymbol{u}) \cdot \boldsymbol{w} = \boldsymbol{u} \cdot (\boldsymbol{A}^{\mathrm{T}}\boldsymbol{w})$.

(c) Let $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3$ be orthonormal vectors in $\mathbb{R}^n$. Suppose $\boldsymbol{w} \in \mathbb{R}^n$. Define

$$b_1 = (\boldsymbol{A}\boldsymbol{v}_1) \cdot \boldsymbol{w}, \quad b_2 = (\boldsymbol{A}\boldsymbol{v}_2) \cdot \boldsymbol{w}, \quad b_3 = (\boldsymbol{A}\boldsymbol{v}_3) \cdot \boldsymbol{w}, \quad \boldsymbol{q} = b_1\boldsymbol{v}_1 + b_2\boldsymbol{v}_2 + b_3\boldsymbol{v}_3.$$

Calculate $\boldsymbol{v}_1 \cdot \boldsymbol{q}$, $\boldsymbol{v}_2 \cdot \boldsymbol{q}$ and $\boldsymbol{v}_3 \cdot \boldsymbol{q}$.

(d) Using the same definitions as in Part(c), show that for every $\boldsymbol{v} \in \operatorname{span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$, $(\boldsymbol{A}\boldsymbol{v}) \cdot \boldsymbol{w} = \boldsymbol{v} \cdot \boldsymbol{q}$.

Proof. (a) $(\boldsymbol{A}\boldsymbol{u}) \cdot \boldsymbol{w} = (\boldsymbol{u}\boldsymbol{A})^{\mathrm{T}}\boldsymbol{w} = (\boldsymbol{u}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}})\boldsymbol{w} = \boldsymbol{u}^{\mathrm{T}}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{w}) = \boldsymbol{u} \cdot (\boldsymbol{A}^{\mathrm{T}}\boldsymbol{w})$.

(c) Note that $\boldsymbol{v}_i \cdot \boldsymbol{v}_i = 1$ and $\boldsymbol{v}_i \cdot \boldsymbol{v}_j = 0$ for $i \neq j$. Then

$$\boldsymbol{v}_i \cdot \boldsymbol{q} = \boldsymbol{v}_i \cdot (b_1\boldsymbol{v}_1 + b_2\boldsymbol{v}_2 + b_3\boldsymbol{v}_3) = b_1(\boldsymbol{v}_i \cdot \boldsymbol{v}_1) + c_2(\boldsymbol{v}_i \cdot \boldsymbol{v}_2) + c_3(\boldsymbol{v}_i \cdot \boldsymbol{v}_3) = b_i, \quad i = 1,2,3.$$

(d) Let $\boldsymbol{v} \in \operatorname{span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$. Then there exist constant c_1, c_2, c_3 such that

$$\boldsymbol{v} = c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + c_3\boldsymbol{v}_3.$$

Recall that $\boldsymbol{v}_i \cdot \boldsymbol{q} = b_i = (\boldsymbol{A}\boldsymbol{v}_i) \cdot \boldsymbol{w} = \boldsymbol{v}_i \cdot (\boldsymbol{A}^{\mathrm{T}}\boldsymbol{w})$. Then

$$\begin{aligned}
\boldsymbol{v} \cdot \boldsymbol{q} &= (c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + c_3\boldsymbol{v}_3) \cdot \boldsymbol{q} = c_1(\boldsymbol{v}_1 \cdot \boldsymbol{q}) + c_2(\boldsymbol{v}_2 \cdot \boldsymbol{q}) + c_3(\boldsymbol{v}_3 \cdot \boldsymbol{q}) \\
&= c_1[\boldsymbol{v}_1 \cdot (\boldsymbol{A}^{\mathrm{T}}\boldsymbol{w})] + c_2[\boldsymbol{v}_2 \cdot (\boldsymbol{A}^{\mathrm{T}}\boldsymbol{w})] + c_3[\boldsymbol{v}_3 \cdot (\boldsymbol{A}^{\mathrm{T}}\boldsymbol{w})] \\
&= (c_1\boldsymbol{v}_1) \cdot (\boldsymbol{A}^{\mathrm{T}}\boldsymbol{w}) + (c_2\boldsymbol{v}_2) \cdot (\boldsymbol{A}^{\mathrm{T}}\boldsymbol{w}) + (c_3\boldsymbol{v}_3) \cdot (\boldsymbol{A}^{\mathrm{T}}\boldsymbol{w}) \\
&= (c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + c_3\boldsymbol{v}_3) \cdot (\boldsymbol{A}^{\mathrm{T}}\boldsymbol{w}) \\
&= \boldsymbol{v} \cdot (\boldsymbol{A}^{\mathrm{T}}\boldsymbol{w}) = (\boldsymbol{A}\boldsymbol{v}) \cdot \boldsymbol{w}.
\end{aligned}$$

□

6. Let $\boldsymbol{A}$ be an $n \times n$ matrix.

(a) Suppose λ is an eigenvalue of $\boldsymbol{A}$ and suppose $\boldsymbol{v}$ is an eigenvector associated with λ. Show that for any $m > 0$, $\boldsymbol{A}^m\boldsymbol{v} = \lambda^m\boldsymbol{v}$.

(b) Let λ be a real number which is an eigenvalue of $\boldsymbol{A}$. Show that if $m > 0$ and $\boldsymbol{A}^m = \boldsymbol{I}$, then $\lambda = \pm 1$.

(c) Suppose $\boldsymbol{P}$ and $\boldsymbol{B}$ are $n \times n$ matrices such that $\boldsymbol{P}$ is invertible and $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{B}$. Show that if $\boldsymbol{B}^2 = \boldsymbol{I}$, then $\boldsymbol{A}^2 = \boldsymbol{I}$.

(d) Assume that $\boldsymbol{A}$ is a symmetric matrix. Show that if $m > 0$ and $\boldsymbol{A}^m = \boldsymbol{I}$, then $\boldsymbol{A}^2 = \boldsymbol{I}$.

Proof. (a) By definition, $\boldsymbol{A}\boldsymbol{v} = \lambda\boldsymbol{v}$. Suppose that $\boldsymbol{A}^k\boldsymbol{v} = \lambda^k\boldsymbol{v}$ for some $k > 0$. Then

$$\boldsymbol{A}^{k+1}\boldsymbol{v} = \boldsymbol{A}(\boldsymbol{A}^k\boldsymbol{v}) = \boldsymbol{A}(\lambda^k\boldsymbol{v}) = \lambda^k(\boldsymbol{A}\boldsymbol{v}) = \lambda^k(\lambda\boldsymbol{v}) = \lambda^{k+1}\boldsymbol{v}.$$

By induction, $\boldsymbol{A}^m\boldsymbol{v} = \lambda^m\boldsymbol{v}$ for every positive integer m.

(b) Let $\boldsymbol{v}$ be an eigenvector of $\boldsymbol{A}$ associated to the eigenvalue λ. Then $\boldsymbol{Av} = \lambda\boldsymbol{v}$. By (a),

$$\lambda^m \boldsymbol{v} = \boldsymbol{A}^m \boldsymbol{v} = \boldsymbol{I}\boldsymbol{v} = \boldsymbol{v}.$$

Note that $\boldsymbol{v} \neq \boldsymbol{0}$. We must have $\lambda^m = 1$. Since $\lambda \in \mathbb{R}$, $\lambda = \pm 1$.

(c) If $\boldsymbol{B}^2 = \boldsymbol{I}$, then

$$\boldsymbol{A}^2 = (\boldsymbol{PBP}^{-1})^2 = \boldsymbol{PB}^2\boldsymbol{P}^{-1} = \boldsymbol{PIP}^{-1} = \boldsymbol{PP}^{-1} = \boldsymbol{I}.$$

(d) Suppose that $\boldsymbol{A}$ is symmetric. Then $\boldsymbol{A}$ is diagonalizable. There exists an invertible matrix $\boldsymbol{P}$ such that $\boldsymbol{P}^{-1}\boldsymbol{AP} = \boldsymbol{D}$, where $\boldsymbol{D}$ is a diagonal matrix whose diagonal entries are the eigenvalues of $\boldsymbol{A}$. If $\boldsymbol{A}^m = \boldsymbol{I}$ for some $m > 0$, then by (b) the eigenvalues of $\boldsymbol{A}$ are ± 1. It follows that $\boldsymbol{D}^2 = \boldsymbol{I}$. By (c), $\boldsymbol{A}^2 = \boldsymbol{I}$. □

2017/2018 SEMESTER 1

5. Let $\boldsymbol{C}$ be a square matrix.

(a) Show that the nullspace of $\boldsymbol{C}$ is a subset of the nullspace of $\boldsymbol{C}^2$.
(b) If $\mathrm{rank}(\boldsymbol{C}^2) = \mathrm{rank}(\boldsymbol{C})$, show that the nullspace of $\boldsymbol{C}^2$ is equal to the nullspace of $\boldsymbol{C}$.

Proof. (a) Let $\boldsymbol{v}$ be a vector in the nullspace of $\boldsymbol{C}$. Then $\boldsymbol{Cv} = \boldsymbol{0}$ and thus

$$\boldsymbol{C}^2\boldsymbol{v} = (\boldsymbol{CC})\boldsymbol{v} = \boldsymbol{C}(\boldsymbol{Cv}) = \boldsymbol{C}\boldsymbol{0} = \boldsymbol{0}.$$

So $\boldsymbol{v}$ is also a vector in the nullspace of $\boldsymbol{C}^2$.

(b) By (a), the nullspace of $\boldsymbol{C}$ is a subspace of the nullspace of $\boldsymbol{C}^2$. Suppose $\boldsymbol{C}$ has order n. Then

$$\mathrm{nullity}(\boldsymbol{C}) = n - \mathrm{rank}(\boldsymbol{C}) = n - \mathrm{rank}(\boldsymbol{C}^2) = \mathrm{nullity}(\boldsymbol{C}^2).$$

We conclude that the nullspace of $\boldsymbol{C}$ equals the nullspace of $\boldsymbol{C}^2$. □

6. Let $\boldsymbol{A}$ be an $n \times n$ matrix. For each $\lambda \in \mathbb{R}$, we define a linear transformation $T_\lambda : \mathbb{R}^n \to \mathbb{R}^n$ such that

$$T_\lambda(\boldsymbol{u}) = \boldsymbol{Au} - \lambda\boldsymbol{u} \quad \text{for} \quad \boldsymbol{u} \in \mathbb{R}^n.$$

(a) Write down the standard matrix for T_λ.
(b) For any $\lambda, \mu \in \mathbb{R}$, show that

$$(\boldsymbol{A} - \lambda\boldsymbol{I})(\boldsymbol{A} - \mu\boldsymbol{I}) = (\boldsymbol{A} - \mu\boldsymbol{I})(\boldsymbol{A} - \lambda\boldsymbol{I}).$$

(c) Suppose $\boldsymbol{A}$ is diagonalizable and the eigenvalues of $\boldsymbol{A}$ are $\lambda_1, \lambda_2, \ldots, \lambda_k$.
   (i) If $\boldsymbol{v}$ is an eigenvector of $\boldsymbol{A}$, say, $\boldsymbol{Av} = \lambda_i\boldsymbol{v}$ for some i, show that

$$(\boldsymbol{A} - \lambda_1\boldsymbol{I})(\boldsymbol{A} - \lambda_2\boldsymbol{I}) \cdots (\boldsymbol{A} - \lambda_k\boldsymbol{I})\boldsymbol{v} = \boldsymbol{0}.$$

   (ii) Define $S = T_{\lambda_1} \circ T_{\lambda_2} \circ \cdots \circ T_{\lambda_k}$. Prove that S is the zero transformation.

Proof. (a) $T_\lambda(\boldsymbol{u}) = \boldsymbol{A}\boldsymbol{u} - \lambda\boldsymbol{u} = \boldsymbol{A}\boldsymbol{u} - \lambda\boldsymbol{I}\boldsymbol{u} = (\boldsymbol{A} - \lambda\boldsymbol{I})\boldsymbol{u}$ for every $\boldsymbol{u} \in \mathbb{R}^n$. Then the standard matrix for T_λ is $\boldsymbol{A} - \lambda\boldsymbol{I}$.

(b) $(\boldsymbol{A} - \lambda\boldsymbol{I})(\boldsymbol{A} - \mu\boldsymbol{I}) = \boldsymbol{A}^2 - A(\mu\boldsymbol{I}) - (\lambda\boldsymbol{I})\boldsymbol{A} + (\lambda\boldsymbol{I})(\mu I) = \boldsymbol{A}^2 - (\lambda+\mu)\boldsymbol{A} + \lambda\mu I$.

$(\boldsymbol{A} - \mu\boldsymbol{I})(\boldsymbol{A} - \lambda\boldsymbol{I}) = \boldsymbol{A}^2 - \boldsymbol{A}(\lambda\boldsymbol{I}) - (\mu\boldsymbol{I})\boldsymbol{A} + (\mu\boldsymbol{I})(\lambda\boldsymbol{I}) = \boldsymbol{A}^2 - (\lambda+\mu)\boldsymbol{A} + \lambda\mu\boldsymbol{I}$. Therefore,

$$(\boldsymbol{A} - \lambda\boldsymbol{I})(\boldsymbol{A} - \mu\boldsymbol{I}) = (\boldsymbol{A} - \mu\boldsymbol{I})(\boldsymbol{A} - \lambda\boldsymbol{I}).$$

(c) If $\boldsymbol{A}\boldsymbol{v} = \lambda_i\boldsymbol{v}$, then $(\boldsymbol{A} - \lambda_i\boldsymbol{I})\boldsymbol{v} = \boldsymbol{A}\boldsymbol{v} - \lambda_i\boldsymbol{v} = \boldsymbol{0}$. By (b),

$$\begin{aligned}
&(\boldsymbol{A} - \lambda_1\boldsymbol{I})(\boldsymbol{A} - \lambda_2\boldsymbol{I})\cdots(\boldsymbol{A} - \lambda_k\boldsymbol{I})\boldsymbol{v} \\
&= (\boldsymbol{A} - \lambda_1\boldsymbol{I})\cdots(\boldsymbol{A} - \lambda_{i-1}\boldsymbol{I})(\boldsymbol{A} - \lambda_{i+1}\boldsymbol{I})\cdots(\boldsymbol{A} - \lambda_k\boldsymbol{I})(\boldsymbol{A} - \lambda_i\boldsymbol{I})\boldsymbol{v} \\
&= (\boldsymbol{A} - \lambda_1\boldsymbol{I})\cdots(\boldsymbol{A} - \lambda_{i-1}\boldsymbol{I})(\boldsymbol{A} - \lambda_{i+1}\boldsymbol{I})\cdots(\boldsymbol{A} - \lambda_k\boldsymbol{I})\boldsymbol{0} = \boldsymbol{0}.
\end{aligned}$$

Since $\boldsymbol{A}$ is diagonalizable, it has n linearly independent eigenvectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$.

The standard matrix for S is $(\boldsymbol{A} - \lambda_1\boldsymbol{I})\cdots(\boldsymbol{A} - \lambda_k\boldsymbol{I})$. Then for each $\boldsymbol{v}_i$,

$$\boldsymbol{S}(\boldsymbol{v}_i) = (\boldsymbol{A} - \lambda_1\boldsymbol{I})\cdots(\boldsymbol{A} - \lambda_k\boldsymbol{I})\boldsymbol{v}_i = \boldsymbol{0}.$$

For any $\boldsymbol{v} \in \mathbb{R}^n$, there exist constant $c_1, \ldots, c_n$ such that $\boldsymbol{v} = c_1\boldsymbol{v}_1 + \cdots + c_n\boldsymbol{v}_n$. Then

$$S(\boldsymbol{v}) = c_1 S(\boldsymbol{v}_1) + \cdots + c_n S(\boldsymbol{v}_n) = \boldsymbol{0}.$$

Therefore, S is the zero transformation. □

2016/2017 Semester 2

6(a) Let $\boldsymbol{A}$ be a square matrix of order n. Let $\boldsymbol{M}_{ij}$ be the matrix of order $n-1$ obtained from $\boldsymbol{A}$ by deleting the i^{th} row and the j^{th} column. Prove that if $\boldsymbol{A}$ is invertible, then at least n of the matrices $\boldsymbol{M}_{ij}$ are invertible.

Proof. Method 1: Let $\boldsymbol{A} = (a_{ij})$. Then $\det(\boldsymbol{A})$ can be found by expansion along the i^{th} row:

$$\det(\boldsymbol{A}) = (-1)^{i+1} a_{i1} \det(\boldsymbol{M}_{i1}) + \cdots + (-1)^{i+n} a_{in} \det(\boldsymbol{M}_{in}).$$

Since $\boldsymbol{A}$ is invertible, $\det(\boldsymbol{A}) \neq 0$; so at least one of $\det(\boldsymbol{M}_{ij}) \neq 0$.

For each $i = 1, \ldots, n$, there exists one invertible submatrix $\boldsymbol{M}_{ij}$. So there are in total at least n invertible submatrices $\boldsymbol{M}_{ij}$.

Method 2: Note that $\boldsymbol{A}[\mathbf{adj}(\boldsymbol{A})] = \det(\boldsymbol{A})\boldsymbol{I}$. If $\boldsymbol{A}$ is invertible, then $\mathbf{adj}(\boldsymbol{A})$ is also invertible. In particular, $\mathbf{adj}(\boldsymbol{A})$ does not have zero columns.

Since the (j,i)-entry of $\mathbf{adj}(\boldsymbol{A})$ is $(-1)^{i+j}\det(\boldsymbol{M}_{ij})$, for each $i = 1, \ldots, n$, there exists at least one $\boldsymbol{M}_{ij}$ such that $\det(\boldsymbol{M}_{ij}) \neq 0$. Therefore, there are at least n invertible submatrices $\boldsymbol{M}_{ij}$. □

6(b) Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be square matrices of the same order.

(i) Prove that the nullspace of $\boldsymbol{B}$ is a subspace of the nullspace of $\boldsymbol{AB}$.

(ii) Using (i) prove that

$$\text{nullity}(\boldsymbol{A}) + \text{nullity}(\boldsymbol{B}) \geq \text{nullity}(\boldsymbol{AB}).$$

Proof. (i) Let $\boldsymbol{v}$ be a vector in the nullspace of $\boldsymbol{B}$, i.e., $\boldsymbol{Bv} = \boldsymbol{0}$. Then $(\boldsymbol{AB})\boldsymbol{v} = \boldsymbol{A}(\boldsymbol{Bv}) = \boldsymbol{A0} = \boldsymbol{0}$. So $\boldsymbol{v}$ is also a vector in the nullspace of $\boldsymbol{AB}$. Therefore, the nullspace of $\boldsymbol{B}$ is a subspace of the nullspace of $\boldsymbol{AB}$.

(ii) Suppose $\boldsymbol{A}$ and $\boldsymbol{B}$ have order n. Let $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$ be a basis for the nullspace of $\boldsymbol{B}$ and extend it to a basis $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k, \boldsymbol{w}_1, \ldots, \boldsymbol{w}_m\}$ for the nullspace of $\boldsymbol{AB}$.

If $c_1\boldsymbol{Bw}_1 + \cdots + c_m\boldsymbol{Bw}_m = \boldsymbol{0}$ for some constants $c_1, \ldots, c_m$, then

$$\boldsymbol{B}(c_1\boldsymbol{w}_1 + \cdots + c_m\boldsymbol{w}_m) = \boldsymbol{0},$$

which implies that $c_1\boldsymbol{w}_1 + \cdots + c_m\boldsymbol{w}_m$ belongs to the nullspace of $\boldsymbol{B}$, that is,

$$c_1\boldsymbol{w}_1 + \cdots + c_m\boldsymbol{w}_m = d_1\boldsymbol{v}_1 + \cdots + d_k\boldsymbol{v}_k$$

for some constants $d_1, \ldots, d_k$. Since $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k, \boldsymbol{w}_1, \ldots, \boldsymbol{w}_m\}$ is linearly independent, we must have $c_1 = \cdots = c_m = d_1 = \cdots = d_k = 0$. Hence, $\{\boldsymbol{Bw}_1, \ldots, \boldsymbol{Bw}_m\}$ is linearly independent.

Note that $\boldsymbol{A}(\boldsymbol{Bw}_i) = (\boldsymbol{AB})\boldsymbol{w}_i = 0$, $\boldsymbol{Bw}_i$ belongs to the nullspace of $\boldsymbol{A}$. Hence,

$$\text{span}\{\boldsymbol{Bw}_1, \ldots, \boldsymbol{Bw}_m\} \subseteq \text{nullspace of } \boldsymbol{A}.$$

Therefore, $\text{nullity}(\boldsymbol{A}) \geq m$, and thus

$$\text{nullity}(\boldsymbol{A}) + \text{nullity}(\boldsymbol{B}) \geq m + k = \text{nullity}(\boldsymbol{AB}). \qquad \square$$

2016/2017 Semester 1

5. Let $\boldsymbol{A}$ be a square matrix of order n.

(a) Let $E = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ be the standard basis for $\mathbb{R}^n$. Show that $\boldsymbol{Ae}_j =$ the j^{th} column of $\boldsymbol{A}$.

(b) Suppose $\boldsymbol{A}^m = \boldsymbol{0}$ and $\boldsymbol{A}^{m-1} \neq \boldsymbol{0}$ for some integer $m \geq 2$.

  (i) Show that there exists at least one vector $\boldsymbol{u} \in \mathbb{R}^n$ such that $\boldsymbol{A}^{m-1}\boldsymbol{u} \neq 0$.

  (ii) Show that $\{\boldsymbol{u}, \boldsymbol{Au}, \ldots, \boldsymbol{A}^{m-1}\boldsymbol{u}\}$ is linearly independent where $\boldsymbol{u}$ is the vector obtained in Part (i).

(c) Prove that if $\boldsymbol{A}^{n+1} = \boldsymbol{0}$, then $\boldsymbol{A}^n = \boldsymbol{0}$.

Proof. (a) Let $\boldsymbol{v}_j$ be the j^{th} column of $\boldsymbol{A}$. Then $\boldsymbol{A} = \begin{pmatrix}\boldsymbol{v}_1 & \cdots & \boldsymbol{v}_n\end{pmatrix}$.

On the other hand, $\boldsymbol{A} = \boldsymbol{AI} = \boldsymbol{A}\begin{pmatrix}\boldsymbol{e}_1 & \cdots & \boldsymbol{e}_n\end{pmatrix} = \begin{pmatrix}\boldsymbol{Ae}_1 & \cdots & \boldsymbol{Ae}_n\end{pmatrix}$. Therefore, $\boldsymbol{Ae}_j = \boldsymbol{v}_j$.

(b) Since $\boldsymbol{A}^{m-1} \neq \boldsymbol{0}$, not all columns of $\boldsymbol{A}^{m-1}$ are zero vectors. Suppose that the j^{th} column of $\boldsymbol{A}^{m-1}$ is nonzero. Then $\boldsymbol{A}^{m-1}\boldsymbol{u} \neq \boldsymbol{0}$ by taking $\boldsymbol{u} = \boldsymbol{e}_j$.

Assume that $\boldsymbol{u}, \boldsymbol{A}\boldsymbol{u}, \ldots, \boldsymbol{A}^{m-1}\boldsymbol{u}$ are linearly dependent. Then there exist constant $c_1, \ldots, c_m$ which are not all zero such that

$$c_1\boldsymbol{u} + c_2\boldsymbol{A}\boldsymbol{u} + \cdots + c_m\boldsymbol{A}^{m-1}\boldsymbol{u} = \boldsymbol{0}.$$

Suppose that $c_1 = \cdots = c_{k-1} = 0$ and $c_k \neq 0$. Then

$$c_k\boldsymbol{A}^{k-1}\boldsymbol{u} + c_{k+1}\boldsymbol{A}^{k}\boldsymbol{u} + \cdots + c_m\boldsymbol{A}^{m-1}\boldsymbol{u} = \boldsymbol{0}.$$

So

$$\begin{aligned}
\boldsymbol{0} &= \boldsymbol{A}^{m-k}(c_k\boldsymbol{A}^{k-1}\boldsymbol{u} + c_{k+1}\boldsymbol{A}^{k}\boldsymbol{u} + \cdots + c_m\boldsymbol{A}^{m-1}\boldsymbol{u}) \\
&= c_k\boldsymbol{A}^{m-1}\boldsymbol{u} + c_{k+1}\boldsymbol{A}^{m}\boldsymbol{u} + \cdots + c_m\boldsymbol{A}^{2m-k-1}\boldsymbol{u} \\
&= c_k\boldsymbol{A}^{m-1}\boldsymbol{u} + \boldsymbol{0} + \cdots + \boldsymbol{0} \\
&= c_k\boldsymbol{A}^{m-1}\boldsymbol{u}.
\end{aligned}$$

Since $\boldsymbol{A}^{m-1}\boldsymbol{u} \neq 0$, we must have $c_k \neq 0$, contradicting our assumption.

Therefore, $\boldsymbol{u}, \boldsymbol{A}\boldsymbol{u}, \ldots, \boldsymbol{A}^{m-1}\boldsymbol{u}$ are linearly independent.

(c) Suppose that $\boldsymbol{A}^{n+1} = \boldsymbol{0}$. Assume that $\boldsymbol{A}^{n} \neq \boldsymbol{0}$. Then by (b), there exists $\boldsymbol{u} \in \mathbb{R}^n$ such that $\{\boldsymbol{u}, \boldsymbol{A}\boldsymbol{u}, \ldots, \boldsymbol{A}^{n}\boldsymbol{u}\}$ is linearly independent. However, in $\mathbb{R}^n$ any set of $n+1$ vectors must be linearly dependent. We obtain a contradiction. Therefore, $\boldsymbol{A}^{n} = \boldsymbol{0}$. □

6(b) Let $\boldsymbol{C}$ be a symmetric matrix of order n with a characteristic polynomial

$$(\lambda - \lambda_1)(\lambda - \lambda_2)\cdots(\lambda - \lambda_n),$$

where $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$. Prove that for any nonzero vector $\boldsymbol{x} \in \mathbb{R}^n$, $\lambda_1 \leq \dfrac{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{C}\boldsymbol{x}}{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}} \leq \lambda_n$.

Proof. Since $\boldsymbol{C}$ is symmetric, there exists an orthonormal basis $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ for $\mathbb{R}^n$ such that $\boldsymbol{v}_i$ is an eigenvector of $\boldsymbol{C}$ associated to the eigenvalue λ_i.

Let $\boldsymbol{x} \in \mathbb{R}^n$. Then

$$\boldsymbol{x} = (\boldsymbol{x}\cdot\boldsymbol{v}_1)\boldsymbol{v}_1 + \cdots + (x\cdot\boldsymbol{v}_n)\boldsymbol{v}_n,$$

and

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x} = \boldsymbol{x}\cdot\boldsymbol{x} = (\boldsymbol{x}\cdot\boldsymbol{v}_1)^2 + \cdots + (\boldsymbol{x}\cdot\boldsymbol{v}_n)^2.$$

We also have

$$\boldsymbol{C}\boldsymbol{x} = (\boldsymbol{x}\cdot\boldsymbol{v}_1)\boldsymbol{C}\boldsymbol{v}_1 + \cdots + (\boldsymbol{x}\cdot\boldsymbol{v}_1)\boldsymbol{C}\boldsymbol{v}_n = \lambda_1(\boldsymbol{x}\cdot\boldsymbol{v}_1)\boldsymbol{v}_1 + \cdots + \lambda_n(\boldsymbol{x}\cdot\boldsymbol{v}_n)\boldsymbol{v}_n,$$

and

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{C}\boldsymbol{x} = \boldsymbol{x}\cdot(\boldsymbol{C}\boldsymbol{x}) = \lambda_1(\boldsymbol{x}\cdot\boldsymbol{v}_1)^2 + \cdots + \lambda_n(\boldsymbol{x}\cdot\boldsymbol{v}_n)^2.$$

Therefore,

$$\lambda_1(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}) = \lambda_1(\boldsymbol{x}\cdot\boldsymbol{v}_1)^2 + \cdots + \lambda_1(\boldsymbol{x}\cdot\boldsymbol{v}_n)^2 \le \boldsymbol{x}^{\mathrm{T}}\boldsymbol{C}\boldsymbol{x} \le \lambda_n(\boldsymbol{x}\cdot\boldsymbol{v}_1)^2 + \cdots + \lambda_n(\boldsymbol{x}\cdot\boldsymbol{v}_n)^2 = \lambda_n(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}).$$

If $\boldsymbol{x} \neq \boldsymbol{0}$, then $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x} \neq 0$, and we have

$$\lambda_1 \le \frac{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{C}\boldsymbol{x}}{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}} \le \lambda_n. \qquad \square$$

2015/2016 Semester 2

5(b) Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a linear operator such that its standard matrix is diagonalizable. Prove that $\mathrm{R}(T) = \mathrm{R}(T \circ T)$ and $\mathrm{Ker}(T) = \mathrm{Ker}(T \circ T)$.

Proof. If the standard matrix $\boldsymbol{A}$ is diagonalizable, then there exists a basis $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ for $\mathbb{R}^n$ consisting of eigenvectors of $\boldsymbol{A}$, i.e., $\boldsymbol{A}\boldsymbol{v}_i = \lambda_i \boldsymbol{v}_i$ for eigenvalue λ_i of $\boldsymbol{A}$.

If $\boldsymbol{v} \in \mathrm{Ker}(T)$, then $T(\boldsymbol{v}) = \boldsymbol{0}$, and thus $T \circ T(\boldsymbol{v}) = T(T(\boldsymbol{v})) = T(\boldsymbol{0}) = \boldsymbol{0}$, i.e., $\boldsymbol{v} \in \mathrm{Ker}(T \circ T)$.

Conversely, let $\boldsymbol{v} \in \mathrm{Ker}(T \circ T)$, there exist unique constants $c_1, \ldots, c_n$ such that $\boldsymbol{v} = c_1\boldsymbol{v}_1 + \cdots + c_n\boldsymbol{v}_n$. So

$$T(\boldsymbol{v}) = c_1 T(\boldsymbol{v}_1) + \cdots + c_n T(\boldsymbol{v}_n) = c_1\lambda_1\boldsymbol{v}_1 + \cdots + c_n\lambda_n\boldsymbol{v}_n,$$

and

$$T \circ T(\boldsymbol{v}) = T(T(\boldsymbol{v})) = c_1\lambda_1 T(\boldsymbol{v}_1) + \cdots + c_n\lambda_n T(\boldsymbol{v}_n) = c_1\lambda_1^2\boldsymbol{v}_1 + \cdots + c_n\lambda_n^2\boldsymbol{v}_n = \boldsymbol{0}.$$

It follows that $c_i\lambda_i^2 = 0$ for all i. Hence, $(c_i\lambda_i)^2 = c_i(c_i\lambda_i^2) = 0$, and thus $c_i\lambda_i = 0$ for all i. So

$$T(\boldsymbol{v}) = c_1\lambda_1\boldsymbol{v}_1 + \cdots + c_n\lambda_n\boldsymbol{v}_n = \boldsymbol{0};$$

that is, $\boldsymbol{v} \in \mathrm{Ker}(T)$. Therefore, $\mathrm{Ker}(T) = \mathrm{Ker}(T \circ T)$. In particular, $\mathrm{nullity}(T) = \mathrm{nullity}(T \circ T)$.

If $\boldsymbol{v} \in \mathrm{R}(T \circ T)$, then there exists $\boldsymbol{u} \in \mathbb{R}^n$ such that $\boldsymbol{v} = T \circ T(\boldsymbol{v}) = T(T(\boldsymbol{v}))$, and thus $\boldsymbol{v} \in \mathrm{R}(T)$. Hence, $\mathrm{R}(T \circ T) \subseteq \mathrm{R}(T)$. Moreover, since

$$\dim \mathrm{R}(T \circ T) = \mathrm{rank}(T \circ T) = n - \mathrm{nullity}(T \circ T) = n - \mathrm{nullity}(T) = \mathrm{rank}(T) = \dim \mathrm{R}(T),$$

we conclude that $\mathrm{R}(T \circ T) = \mathrm{R}(T)$. $\square$

6(b) Let $V = \mathrm{span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \boldsymbol{v}_4\}$ be a vector space such that $\boldsymbol{v}_i$ are unit vectors for all i and $\boldsymbol{v}_i \cdot \boldsymbol{v}_j < 0$ if $i \neq j$.

(i) Show that no two vectors among $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \boldsymbol{v}_4\}$ are linearly dependent.

(ii) Prove that $\dim(V) \ge 3$.

Proof. (i) Assume that $\boldsymbol{v}_i$ and $\boldsymbol{v}_j$ are linearly dependent ($i \neq j$). Then $\boldsymbol{v}_j = c\boldsymbol{v}_i$ for a constant c. By assumption, $\boldsymbol{v}_i \cdot \boldsymbol{v}_j = c < 0$.

Choose $k \neq i, j$. Then $\boldsymbol{v}_i \cdot \boldsymbol{v}_k < 0$, but this would imply that $\boldsymbol{v}_j \cdot \boldsymbol{v}_k = c(\boldsymbol{v}_i \cdot \boldsymbol{v}_k) > 0$, contradicting the assumption. Hence, $\boldsymbol{v}_i$ and $\boldsymbol{v}_j$ are linearly dependent.

(ii) Assume that $\dim(V) \leq 2$. By (i), $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ are linearly independent; so $\dim(V) \geq 2$. Hence, $\dim(V) = 2$, and $\{\boldsymbol{v}_1, \boldsymbol{v}_2\}$ is a basis for V.

Write $\boldsymbol{v}_3 = c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2$ for constants c_1 and c_2. Then $c_1 \neq 0$, $c_2 \neq 0$, and

$$\boldsymbol{v}_1 \cdot \boldsymbol{v}_3 = c_1 + c_2(\boldsymbol{v}_1 \cdot \boldsymbol{v}_2) < 0 \quad \text{and} \quad \boldsymbol{v}_2 \cdot \boldsymbol{v}_3 = c_1(\boldsymbol{v}_1 \cdot \boldsymbol{v}_2) + c_2 < 0.$$

Multiply the first inequality by $\boldsymbol{v}_1 \cdot \boldsymbol{v}_2 < 0$ to get

$$c_1(\boldsymbol{v}_1 \cdot \boldsymbol{v}_2) + c_2(\boldsymbol{v}_1 \cdot \boldsymbol{v}_2)^2 > 0.$$

So

$$c_2(\boldsymbol{v}_1 \cdot \boldsymbol{v}_2)^2 > -c_1(\boldsymbol{v}_1 \cdot \boldsymbol{v}_2) > c_2;$$

that is, $c_2[(\boldsymbol{v}_1 \cdot \boldsymbol{v}_2)^2 - 1] > 0$. By Cauchy-Schwarz inequality, $|\boldsymbol{v}_1 \cdot \boldsymbol{v}_2| \leq \|\boldsymbol{v}_1\| \, \|\boldsymbol{v}_2\| = 1$; so $c_2 < 0$. Then $c_1(\boldsymbol{v}_1 \cdot \boldsymbol{v}_2) > -c_2 > 0$ implies that $c_1 < 0$. However, we would have

$$\boldsymbol{v}_3 \cdot \boldsymbol{v}_4 = c_1(\boldsymbol{v}_1 \cdot \boldsymbol{v}_4) + c_2(\boldsymbol{v}_2 \cdot \boldsymbol{v}_4) > 0,$$

a contradiction. Therefore, $\dim(V) \geq 3$. □

2015/2016 Semester 1

4(b) Let W be a subspace of $\mathbb{R}^n$ and $W^\perp = \{\boldsymbol{w} \in \mathbb{R}^n \mid \boldsymbol{w} \text{ is orthogonal to } W\}$. Prove that $\dim(W) + \dim(W^\perp) = W$.

Proof. Suppose $W = \operatorname{span}\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$. View each $\boldsymbol{v}_i$ as a row vector and set $\boldsymbol{A} = \begin{pmatrix} \boldsymbol{v}_1 \\ \vdots \\ \boldsymbol{w}_k \end{pmatrix}$. Then W is the row space of $\boldsymbol{A}$.

View $\boldsymbol{w}$ as a column vector. Then

$$\begin{aligned} \boldsymbol{w} \in W^\perp &\Leftrightarrow \boldsymbol{w} \text{ is orthogonal to } \boldsymbol{v}_i \text{ for all } i = 1, \ldots, k \\ &\Leftrightarrow \boldsymbol{v}_i \cdot \boldsymbol{w} = \boldsymbol{v}_i\boldsymbol{w} = 0 \text{ for all } i = 1, \ldots, k \\ &\Leftrightarrow \boldsymbol{A}\boldsymbol{w} = \begin{pmatrix} \boldsymbol{v}_1 \\ \vdots \\ \boldsymbol{v}_k \end{pmatrix} \boldsymbol{w} = \boldsymbol{0} \\ &\Leftrightarrow \boldsymbol{w} \in \text{nullspace of } \boldsymbol{A}. \end{aligned}$$

In other words, $W^\perp$ is the nullspace of $\boldsymbol{A}$. Therefore,

$$\dim(W) + \dim(W^\perp) = \operatorname{rank}(\boldsymbol{A}) + \operatorname{nullity}(\boldsymbol{A}) = n.$$ □

5. Let $\boldsymbol{A}$ be an $n \times n$ matrix such that $\boldsymbol{A}^n = \boldsymbol{0}$. Suppose there exists a nonzero vector $\boldsymbol{v} \in \mathbb{R}^n$ such that $\boldsymbol{A}^{n-1}\boldsymbol{v} \neq \boldsymbol{0}$.

(b) Prove that $\{\boldsymbol{v}, \boldsymbol{A}\boldsymbol{v}, \dots, \boldsymbol{A}^{n-1}\boldsymbol{v}\}$ is a basis for $\mathbb{R}^n$.

(c) Let $\boldsymbol{P} = \begin{pmatrix} \boldsymbol{A}^{n-1}\boldsymbol{v} & \cdots & \boldsymbol{A}\boldsymbol{v} & \boldsymbol{v} \end{pmatrix}$ which is an invertible matrix of order n. Show that

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \ddots & 1 \\ 0 & 0 & \cdots & \cdots & 0 \end{pmatrix}.$$

Proof. (b) Assume that $\boldsymbol{v}, \boldsymbol{A}\boldsymbol{v}, \dots, \boldsymbol{A}^{n-1}\boldsymbol{v}$ are linearly independent. Then there exists constants $c_1, \dots, c_n$ which are not all zero such that

$$c_1\boldsymbol{v}_1 + c_2\boldsymbol{A}\boldsymbol{v} + \cdots + c_n\boldsymbol{A}^{n-1}\boldsymbol{v} = \boldsymbol{0}.$$

Let k be the smallest index such that $c_k \neq 0$. Then

$$c_k\boldsymbol{A}^{k-1}\boldsymbol{v} + c_{k+1}\boldsymbol{A}^k\boldsymbol{v} + \cdots + c_n\boldsymbol{A}^{n-1}\boldsymbol{v} = \boldsymbol{0}.$$

Pre-multiplication of $\boldsymbol{A}^{n-k}$ yields

$$\boldsymbol{0} = c_k\boldsymbol{A}^{n-1}\boldsymbol{v} + c_{k+1}\boldsymbol{A}^n\boldsymbol{v} + \cdots + c_n\boldsymbol{A}^{2n-k-1}\boldsymbol{v} = c_k\boldsymbol{A}^{n-1}\boldsymbol{v}.$$

Since $\boldsymbol{A}^{n-1}\boldsymbol{v} \neq 0$, we must have $c_k = 0$, a contradiction.

Therefore, $\{\boldsymbol{v}, \boldsymbol{A}\boldsymbol{v}, \dots, \boldsymbol{A}^{n-1}\boldsymbol{v}\}$ is a linearly independent subset of $\mathbb{R}^n$ with exactly n vectors. Hence, it is a basis for $\mathbb{R}^n$.

(c) We have proved that $S = \{\boldsymbol{A}^{n-1}\boldsymbol{v}, \dots, \boldsymbol{A}\boldsymbol{v}, \boldsymbol{v}\}$ is a basis for $\mathbb{R}^n$. Then $\boldsymbol{P}$ is the transition matrix from S to the standard basis E. For any $\boldsymbol{u} \in \mathbb{R}^n$,

$$\boldsymbol{P}[\boldsymbol{u}]_S = \boldsymbol{u} \quad \text{and} \quad \boldsymbol{P}[\boldsymbol{A}\boldsymbol{u}]_S = \boldsymbol{A}\boldsymbol{u}.$$

Then

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}[\boldsymbol{u}]_S = \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{u} = [\boldsymbol{A}\boldsymbol{u}]_S.$$

Note that for each $k = 1, \dots, n$, $[\boldsymbol{A}^{n-k}\boldsymbol{v}]_S = \boldsymbol{e}_k$. Then

$$\begin{aligned} \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} &= \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}\begin{pmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 & \cdots & \boldsymbol{e}_n \end{pmatrix} \\ &= \begin{pmatrix} \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}[\boldsymbol{A}^{n-1}\boldsymbol{v}]_S & \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}[\boldsymbol{A}^{n-2}\boldsymbol{v}] & \cdots & \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}[\boldsymbol{v}]_S \end{pmatrix} \\ &= \begin{pmatrix} [\boldsymbol{A}^n\boldsymbol{v}]_S & [\boldsymbol{A}^{n-1}\boldsymbol{v}]_S & \cdots & [\boldsymbol{A}\boldsymbol{v}]_S \end{pmatrix} \\ &= \begin{pmatrix} \boldsymbol{0} & \boldsymbol{e}_1 & \cdots & \boldsymbol{e}_{n-1} \end{pmatrix}. \end{aligned}$$

□

6. Let $\boldsymbol{A}$ be an invertible matrix of order n such that for any nonzero vectors $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n$, the angle between $\boldsymbol{u}$ and $\boldsymbol{v}$ is always equal to the angle between $\boldsymbol{A}\boldsymbol{u}$ and $\boldsymbol{A}\boldsymbol{v}$.

(a) Let $\boldsymbol{A} = \begin{pmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_n \end{pmatrix}$, where $\boldsymbol{a}_i$ is the i^{th} column of $\boldsymbol{A}$. Show that $\{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n\}$ is an orthogonal basis for $\mathbb{R}^n$.

(b) Prove that $\boldsymbol{A} = c\boldsymbol{P}$ for some scalar c and orthogonal matrix $\boldsymbol{P}$.

Proof. (a) Since $\boldsymbol{A}$ is invertible, $\boldsymbol{a}_i \neq \boldsymbol{0}$ for all i. For $i \neq j$, the angle between $\boldsymbol{e}_i$ and $\boldsymbol{e}_j$ is 90°. Hence, the angle between $\boldsymbol{A}\boldsymbol{e}_i = \boldsymbol{a}_i$ and $\boldsymbol{A}\boldsymbol{e}_j = \boldsymbol{a}_j$ is also 90°. Then

$$\frac{\boldsymbol{a}_i \cdot \boldsymbol{a}_j}{\|\boldsymbol{a}_i\| \, \|\boldsymbol{a}_j\|} = \cos(90^\circ) = 0,$$

which implies that $\boldsymbol{a}_i \cdot \boldsymbol{a}_j = 0$. Therefore, $\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n\}$ is an orthogonal basis for $\mathbb{R}^n$.

(b) For $i \neq j$, $(\boldsymbol{e}_i + \boldsymbol{e}_j) \cdot (\boldsymbol{e}_i - \boldsymbol{e}_j) = 0$, i.e., the angle between $\boldsymbol{e}_i + \boldsymbol{e}_j$ and $\boldsymbol{e}_i - \boldsymbol{e}_j$ is 90°. By assumption, the angle between

$$\boldsymbol{A}(\boldsymbol{e}_i + \boldsymbol{e}_j) = \boldsymbol{a}_i + \boldsymbol{a}_j \quad \text{and} \quad \boldsymbol{A}(\boldsymbol{e}_i - \boldsymbol{e}_j) = \boldsymbol{a}_i - \boldsymbol{a}_j$$

is also 90°. Hence,

$$\frac{(\boldsymbol{a}_i + \boldsymbol{a}_j) \cdot (\boldsymbol{a}_i - \boldsymbol{a}_j)}{\|\boldsymbol{a}_i + \boldsymbol{a}_j\| \, \|\boldsymbol{a}_i - \boldsymbol{a}_j\|} = \cos(90^\circ) = 0.$$

It follows that

$$(\boldsymbol{a}_i + \boldsymbol{a}_j) \cdot (\boldsymbol{a}_i - \boldsymbol{a}_j) = \|\boldsymbol{a}_i\|^2 - \|\boldsymbol{a}_j\|^2 = 0,$$

that is, $\|\boldsymbol{a}_i\| = \|\boldsymbol{a}_j\|$. Let $c = \|\boldsymbol{a}_i\|$ and $\boldsymbol{v}_i = \|\boldsymbol{a}_i\|/c$. Then $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ is an orthonormal basis for $\mathbb{R}^n$. Let $\boldsymbol{P} = \begin{pmatrix} \boldsymbol{v}_1 & \cdots & \boldsymbol{v}_n \end{pmatrix}$. Then $\boldsymbol{P}$ is an orthogonal matrix and

$$\boldsymbol{A} = \begin{pmatrix} c\boldsymbol{v}_1 & \cdots & c\boldsymbol{v}_n \end{pmatrix} = c\begin{pmatrix} \boldsymbol{v}_1 & \cdots & c\boldsymbol{v}_n \end{pmatrix} = c\boldsymbol{P}.$$

□

2014/2015 Semester 2

4(b) Let $\boldsymbol{M}$ and $\boldsymbol{N}$ be two $n \times n$ matrices. Suppose $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ is a set of linearly independent eigenvectors for both $\boldsymbol{M}$ and $\boldsymbol{N}$. Then $\boldsymbol{M}\boldsymbol{N} = \boldsymbol{N}\boldsymbol{M}$.

Proof. Let $\boldsymbol{P} = \begin{pmatrix} \boldsymbol{v}_1 & \cdots & \boldsymbol{v}_n \end{pmatrix}$. Then $\boldsymbol{P}$ is invertible such that $\boldsymbol{P}^{-1}\boldsymbol{M}\boldsymbol{P} = \boldsymbol{D}_1$ and $\boldsymbol{P}^{-1}\boldsymbol{N}\boldsymbol{P} = \boldsymbol{D}_2$ are diagonal matrices. Then

$$\begin{aligned} \boldsymbol{M}\boldsymbol{N} &= (\boldsymbol{P}\boldsymbol{D}_1\boldsymbol{P}^{-1})(\boldsymbol{P}\boldsymbol{D}_2\boldsymbol{P}^{-1}) = \boldsymbol{P}\boldsymbol{D}_1\boldsymbol{D}_2\boldsymbol{P}^{-1} \\ &= \boldsymbol{P}\boldsymbol{D}_2\boldsymbol{D}_1\boldsymbol{P}^{-1} = (\boldsymbol{P}\boldsymbol{D}_2\boldsymbol{P}^{-1})(\boldsymbol{P}\boldsymbol{D}_1\boldsymbol{P}^{-1}) = \boldsymbol{N}\boldsymbol{M}. \end{aligned}$$

□

2014/2015 Semester 1

5(b) Given hat $F : \mathbb{R}^3 \to \mathbb{R}^3$ is a linear transformation, P is a plane in $\mathbb{R}^3$ given by the equation $x + y + z = 0$, and ℓ is a line in $\mathbb{R}^3$ given by the set $\{(t, t, t) \mid t \in \mathbb{R}\}$.

Suppose F maps the plane P onto the line ℓ and maps the line ℓ to the origin. Show that the linear transformation F^2 (i.e., $F \circ F$) is the zero transformation.

Proof. Note that P has a basis $\{(-1,1,0),(-1,0,1)\}$ and ℓ has a basis $\{(1,1,1)\}$, Let

$$\boldsymbol{v}_1 = (-1,1,0), \quad \boldsymbol{v}_2 = (-1,0,1) \quad \text{and} \quad \boldsymbol{v}_3 = (1,1,1).$$

Since $\det\begin{pmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix} = 3 \neq 0$, $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ is a basis for $\mathbb{R}^3$.

By assumption, $F(\boldsymbol{v}_i) \in \ell$, and thus $F \circ F(\boldsymbol{v}_i) = F(F(\boldsymbol{v}_i)) = 0$, $i = 1,2$. We also have

$$F \circ F(\boldsymbol{v}_3) = F(F(\boldsymbol{v}_3)) = F(\mathbf{0}) = \mathbf{0}.$$

Any vector $\boldsymbol{v} \in \mathbb{R}^3$ can be written as $\boldsymbol{v} = c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + c_3\boldsymbol{v}_3$ for some constants c_1, c_2, c_3. Then

$$F \circ F(\boldsymbol{v}) = c_1 F \circ F(\boldsymbol{v}_1) + c_2 F \circ F(\boldsymbol{v}_2) + c_3 F \circ F(\boldsymbol{v}_3) = \mathbf{0}.$$

Therefore, $F \circ F$ is the zero transformation. $\square$

5(c) Let $T : \mathbb{R}^3 \to \mathbb{R}^2$ be a linear transformation. Suppose $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3\}$ is a basis for $\mathbb{R}^3$, and $\{T(\boldsymbol{u}_1), T(\boldsymbol{u}_2), T(\boldsymbol{u}_3)\}$ spans $\mathbb{R}^2$. Show that the standard matrix of T is of full rank.

Proof. Let $\boldsymbol{A}$ be the standard matrix for T. Then $\boldsymbol{A}$ is of size 2×3.

Note that $\mathrm{R}(T) \supseteq \operatorname{span}\{T(\boldsymbol{u}_1), T(\boldsymbol{u}_2), T(\boldsymbol{u}_3)\} = \mathbb{R}^2$. We have $\mathrm{R}(T) = \mathbb{R}^2$. Then

$$\operatorname{rank}(\boldsymbol{A}) = \operatorname{rank}(T) = 2.$$

It follows that $\boldsymbol{A}$ has full rank. $\square$

2013/2014 Semester 2

2(a)(iii) Suppose $\boldsymbol{Ax} = \boldsymbol{b}$ is an inconsistent linear system. Prove that for all $k \neq 0$, $k \in \mathbb{R}$, the linear system $\boldsymbol{Ax} = k\boldsymbol{b}$ is also inconsistent. If $\boldsymbol{v}$ is a least squares solution for $\boldsymbol{Ax} = \boldsymbol{b}$, is $k\boldsymbol{v}$ a least squares solution for $\boldsymbol{Ax} = k\boldsymbol{b}$? Justify your answer.

Proof. Assume that $\boldsymbol{Ax} = k\boldsymbol{b}$ is consistent for some $k \neq 0$. Then there exists a vector $\boldsymbol{v}$ such that $\boldsymbol{Av} = k\boldsymbol{b}$. But then

$$\boldsymbol{A}(k^{-1}\boldsymbol{v}) = k^{-1}(\boldsymbol{Av}) = k^{-1}(k\boldsymbol{v}) = \boldsymbol{v}.$$

So $k^{-1}\boldsymbol{v}$ is a solution to $\boldsymbol{Ax} = \boldsymbol{b}$, contradicting the inconsistency of the system.

Suppose that $\boldsymbol{v}$ is a least squares solution for $\boldsymbol{Ax} = \boldsymbol{b}$. Then $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{Av} = \boldsymbol{b}$. So

$$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}(k\boldsymbol{v}) = k(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{Av}) = k(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}) = \boldsymbol{A}^{\mathrm{T}}(k\boldsymbol{b}).$$

Hence, $k\boldsymbol{v}$ is a least squares solution to $\boldsymbol{Ax} = k\boldsymbol{b}$. $\square$

2013/2014 SEMESTER 1

Let $\boldsymbol{Y}$ be a diagonalizable matrix of order n. Suppose $\lambda_1, \lambda_2, \ldots, \lambda_k$ are the distinct eigenvalues of $\boldsymbol{Y}$ with eigenspaces $E_{\lambda_1}, E_{\lambda_2}, \ldots, E_{\lambda_k}$ and $S_{\lambda_1}, S_{\lambda_2}, \ldots, S_{\lambda_k}$ are the corresponding bases for $E_{\lambda_1}, E_{\lambda_2}, \ldots, E_{\lambda_k}$. Prove that

$$E_{\lambda_1} + E_{\lambda_2} + \cdots + E_{\lambda_k} = \mathbb{R}^n.$$

Proof. If $d_i = \dim(E_{\lambda_i})$, write $S_{\lambda_i} = \{\boldsymbol{v}_{i1}, \ldots, \boldsymbol{v}_{id_i}\}$.

Suppose that $\boldsymbol{Y}$ is diagonalizable. Then $d_1 + \cdots + d_k = n$, and $S = S_{\lambda_1} \cup \cdots \cup S_{\lambda_k}$ is a basis for $\mathbb{R}^n$. Every vector $\boldsymbol{u} \in \mathbb{R}^n$ can be written as

$$\boldsymbol{u} = c_{11}\boldsymbol{v}_{11} + \cdots + c_{1d_1}\boldsymbol{v}_{1d_1} + \cdots + c_{k1}\boldsymbol{v}_{k1} + \cdots + c_{kd_k}\boldsymbol{v}_{kd_k},$$

for some constants c_{ij}. Let $\boldsymbol{v}_i = c_{i1}\boldsymbol{v}_{i1} + \cdots + c_{id_i}\boldsymbol{v}_{id_i}$. Then $\boldsymbol{v}_i \in E_{\lambda_i}$ and

$$\boldsymbol{u} = \boldsymbol{v}_1 + \cdots + \boldsymbol{v}_k.$$

Therefore, $\mathbb{R}^n = E_{\lambda_1} + \cdots + E_{\lambda_k}$. □

2012/2013 SEMESTER 2

3. Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be square matrices of the same order. Let $\boldsymbol{x}$ be an eigenvector of $\boldsymbol{AB}$ associated with eigenvalue λ.

(i) If $\lambda \neq 0$, show that $\boldsymbol{Bx}$ is an eigenvector of $\boldsymbol{BA}$ with eigenvalue λ.

(ii) If $\lambda = 0$, is $\boldsymbol{Bx}$ an eigenvector of $\boldsymbol{BA}$ with eigenvalue λ? Justify your answer.

Proof. (i) It is given that $\boldsymbol{ABx} = \lambda\boldsymbol{x}$ with $\boldsymbol{x} \neq \boldsymbol{0}$. Then

$$(\boldsymbol{BA})(\boldsymbol{Bx}) = \boldsymbol{B}(\boldsymbol{ABx}) = \boldsymbol{B}(\lambda\boldsymbol{x}) = \lambda(\boldsymbol{Bx}).$$

If $\lambda \neq 0$, then $\boldsymbol{ABx} = \lambda\boldsymbol{x} \neq 0$, and thus $\boldsymbol{Bx} \neq 0$. Hence, $\boldsymbol{Bx}$ is an eigenvector of $\boldsymbol{BA}$ associated to the eigenvalue λ.

(ii) If $\lambda = 0$, the statement may not be true. For example, let $\boldsymbol{B} = \boldsymbol{0}$, then $\boldsymbol{Bx} = \boldsymbol{0}$ for any $\boldsymbol{x}$; in particular, $\boldsymbol{Bx}$ cannot be an eigenvector. □

4(b) Let $\boldsymbol{A}$ be a square matrix of order n such that for any $\boldsymbol{u} \in \mathbb{R}^n$,

$$\|\boldsymbol{Au}\| = \|\boldsymbol{u}\|.$$

(i) Prove that $\boldsymbol{Au} \cdot \boldsymbol{Av} = \boldsymbol{u} \cdot \boldsymbol{v}$ for any $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n$.

(ii) Using (i) or otherwise, prove that $\boldsymbol{A}$ is an orthogonal matrix.

Proof. (i) For any vectors $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n$,

$$\|\boldsymbol{u}+\boldsymbol{v}\|^2 = (\boldsymbol{u}+\boldsymbol{v})\cdot(\boldsymbol{u}+\boldsymbol{v}) = \|\boldsymbol{u}\|^2 + \|\boldsymbol{v}\|^2 + 2(\boldsymbol{u}\cdot\boldsymbol{v}),$$

$$\|\boldsymbol{u}-\boldsymbol{v}\|^2 = (\boldsymbol{u}-\boldsymbol{v})\cdot(\boldsymbol{u}-\boldsymbol{v}) = \|\boldsymbol{v}\|^2 + \|\boldsymbol{v}\|^2 - 2(\boldsymbol{u}\cdot\boldsymbol{v}).$$

Subtraction yields

$$4(\boldsymbol{u}\cdot\boldsymbol{v}) = \|\boldsymbol{u}+\boldsymbol{v}\|^2 - \|\boldsymbol{u}-\boldsymbol{v}\|^2.$$

For any $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n$,

$$\begin{aligned}
\boldsymbol{A}\boldsymbol{u}\cdot\boldsymbol{A}\boldsymbol{v} &= \frac{1}{4}\left(\|\boldsymbol{A}\boldsymbol{u}+\boldsymbol{A}\boldsymbol{v}\|^2 - \|\boldsymbol{A}\boldsymbol{u}-\boldsymbol{A}\boldsymbol{v}\|^2\right)\\
&= \frac{1}{4}\left(\|\boldsymbol{A}(\boldsymbol{u}+\boldsymbol{v})\|^2 - \|\boldsymbol{A}(\boldsymbol{u}-\boldsymbol{v})\|^2\right)\\
&= \frac{1}{4}\left(\|\boldsymbol{u}+\boldsymbol{v}\|^2 - \|\boldsymbol{u}-\boldsymbol{v}\|^2\right) = \boldsymbol{u}\cdot\boldsymbol{v}.
\end{aligned}$$

(ii) Note that the i^{th} column of $\boldsymbol{A}$ is $\boldsymbol{A}\boldsymbol{e}_i$. We have $\|\boldsymbol{A}\boldsymbol{e}_i\| = \|\boldsymbol{e}_i\| = 1$, and for $i \neq j$,

$$\boldsymbol{A}\boldsymbol{e}_i \cdot \boldsymbol{A}\boldsymbol{e}_j = \boldsymbol{e}_i \cdot \boldsymbol{e}_j = 0.$$

Then the columns of $\boldsymbol{A}$ form an orthonormal basis for $\mathbb{R}$. Hence, $\boldsymbol{A}$ is an orthogonal matrix.

□

4(c) Let $\boldsymbol{A}$ be a square matrix of order n such that $\boldsymbol{A}^2 = \boldsymbol{A}$.

(i) Prove that $\boldsymbol{A}$ is diagonalizable.

(ii) Prove that $\text{rank}(\boldsymbol{A}) = \text{tr}(\boldsymbol{A})$.

Proof. (i) If λ is an eigenvalue of $\boldsymbol{A}$, and $\boldsymbol{v}$ an associated eigenvector, then $\boldsymbol{A}\boldsymbol{v} = \lambda\boldsymbol{v}$; so

$$\lambda\boldsymbol{v} = \boldsymbol{A}\boldsymbol{v} = \boldsymbol{A}^2\boldsymbol{v} = \boldsymbol{A}(\lambda\boldsymbol{v}) = \lambda(\boldsymbol{A}\boldsymbol{v}) = \lambda(\lambda\boldsymbol{v}) = \lambda^2\boldsymbol{v}.$$

Since $\boldsymbol{v} \neq \boldsymbol{0}$, $\lambda = \lambda^2$, and thus $\lambda = 0$ or $\lambda = 1$.

If 0 is not an eigenvalue of $\boldsymbol{A}$, then $\boldsymbol{A}$ is invertible; $\boldsymbol{A}^2 = \boldsymbol{A}$ implies that $\boldsymbol{A} = \boldsymbol{I}$, which is diagonalizable.

If 1 is not an eigenvalue of $\boldsymbol{A}$, then $\boldsymbol{I} - \boldsymbol{A}$ is invertible. Since $\boldsymbol{A}^2 = \boldsymbol{A}$ implies

$$(\boldsymbol{I}-\boldsymbol{A})^2 = \boldsymbol{I} - 2\boldsymbol{A} + \boldsymbol{A}^2 = \boldsymbol{I} - 2\boldsymbol{A} + \boldsymbol{A} = \boldsymbol{I} - A,$$

we have $\boldsymbol{I} - \boldsymbol{A} = \boldsymbol{I}$, i.e., $\boldsymbol{A} = \boldsymbol{0}$, which is diagonalizable.

Suppose 0 and 1 are both eigenvalues of $\boldsymbol{A}$. For any $\boldsymbol{v} \in \mathbb{R}^n$,

$$(\boldsymbol{I}-\boldsymbol{A})(\boldsymbol{A}\boldsymbol{v}) = (\boldsymbol{A}-\boldsymbol{A}^2)\boldsymbol{v} = \boldsymbol{0}\boldsymbol{v} = \boldsymbol{0},$$

which implies that $\boldsymbol{A}\boldsymbol{v} \in E_1$, the eigenspace of $\boldsymbol{A}$ associated to the eigenvalue 1. We also have

$$\boldsymbol{A}(\boldsymbol{v}-\boldsymbol{A}\boldsymbol{v}) = \boldsymbol{A}(\boldsymbol{I}-\boldsymbol{A})\boldsymbol{v} = (\boldsymbol{A}-\boldsymbol{A}^2)\boldsymbol{v} = \boldsymbol{0},$$

which implies that $\boldsymbol{v} - \boldsymbol{A}\boldsymbol{v} \in E_0$, the eigenspace of $\boldsymbol{A}$ associated to the eigenvalue 0. Since

$$\boldsymbol{v} = (\boldsymbol{v} - \boldsymbol{A}\boldsymbol{v}) + \boldsymbol{A}\boldsymbol{v},$$

we conclude that $\dim E_0 + \dim E_1 \geq n$. On the other hand, $\dim E_0 + \dim E_1 \leq n$. Then we must have

$$\dim E_0 + \dim E_1 = n.$$

and $\boldsymbol{A}$ is diagonalizable.

(ii) Let $r = \operatorname{rank}(\boldsymbol{A})$. Since $\operatorname{nullity}(\boldsymbol{A}) = \dim E_0$, $r = n - \dim E_0 = \dim E_1$.

Since $\boldsymbol{A}$ is diagonalizable, there exists an invertible matrix $\boldsymbol{P}$ such that $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{D}$ is a diagonal matrix whose first r diagonal entries are 1, and the remaining are 0. Therefore,

$$\operatorname{tr}(\boldsymbol{A}) = \operatorname{tr}(\boldsymbol{A}\boldsymbol{P}\boldsymbol{P}^{-1}) = \operatorname{tr}(\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}) = \operatorname{tr}(\boldsymbol{D}) = r = \operatorname{rank}(\boldsymbol{A}). \qquad \square$$

2012/2013 Semester 1

2(b)(ii) Let V and W be subspaces of $\mathbb{R}^n$. Recall the definition of $V + W$ as follows

$$V + W = \{\boldsymbol{v} + \boldsymbol{w} \mid \boldsymbol{v} \in V,\ \boldsymbol{w} \in W\}.$$

Define $V - W$ by

$$V - W = \{\boldsymbol{v} - \boldsymbol{w} \mid \boldsymbol{v} \in V,\ \boldsymbol{w} \in W\}.$$

Prove that in general, $V - W = V + W$.

Proof. Every vector in $V + W$ is of the form $\boldsymbol{v} + \boldsymbol{w}$ for $\boldsymbol{v} \in V$, $\boldsymbol{w} \in W$, which can be written as

$$\boldsymbol{v} - (-\boldsymbol{w}), \quad \boldsymbol{v} \in V,\ -\boldsymbol{w} \in W;$$

so $\boldsymbol{v} + \boldsymbol{w} \in V - W$. Conversely, every vector in $V - W$ is of the form $\boldsymbol{v} - \boldsymbol{w}$ for $\boldsymbol{v} \in V$ and $\boldsymbol{w} \in W$, which can be written as

$$\boldsymbol{v} + (-\boldsymbol{w}), \quad \boldsymbol{v} \in V,\ -\boldsymbol{w} \in W;$$

so $\boldsymbol{v} - \boldsymbol{w} \in V + W$. Therefore, $V - W = V + W$. $\square$

4(b)(ii) Prove that an $m \times n$ matrix $\boldsymbol{A}$ has rank 1 if and only if $\boldsymbol{A} = \boldsymbol{a}\boldsymbol{b}^{\mathrm{T}}$ for some nonzero column vectors $\boldsymbol{a}$ and $\boldsymbol{b}$.

Proof. Suppose that $\boldsymbol{A} = \boldsymbol{a}\boldsymbol{b}^{\mathrm{T}}$ for some nonzero column vectors $\boldsymbol{a}$ and $\boldsymbol{b}$. Let a_i and b_i be the i^{th} components of $\boldsymbol{a}$ and $\boldsymbol{b}$ respectively. Since $\boldsymbol{a} \neq \boldsymbol{0}$ and $\boldsymbol{b} \neq \boldsymbol{0}$, $a_i \neq 0$ and $b_j \neq 0$ for some i, j. So the (i, j)-entry of $\boldsymbol{A}$ is $a_i b_j \neq 0$. In particular, $\boldsymbol{A} \neq \boldsymbol{0}$, and thus $\operatorname{rank}(\boldsymbol{A}) \geq 1$.

On the other hand, $\operatorname{rank}(\boldsymbol{A}) \leq \operatorname{rank}(\boldsymbol{a}) = 1$. Hence, $\operatorname{rank}(\boldsymbol{A}) = 1$.

Conversely, suppose that $\operatorname{rank}(\boldsymbol{A}) = 1$, i.e., the column space of $\boldsymbol{A}$ has dimension 1. Let $\{\boldsymbol{a}\}$ be a basis for the column space of $\boldsymbol{A}$. Then the j^{th} column of $\boldsymbol{A}$ is $\boldsymbol{c}_i = c_i\boldsymbol{a}$ for some constant c_i. Hence,

$$\boldsymbol{A} = \begin{pmatrix} c_1\boldsymbol{a} & \cdots & c_n\boldsymbol{a} \end{pmatrix} = \boldsymbol{a}\begin{pmatrix} c_1 & \cdots & c_n \end{pmatrix} = \boldsymbol{a}\boldsymbol{b}^{\mathrm{T}},$$

where $\boldsymbol{b} = (c_1, \ldots, c_n)^{\mathrm{T}}$. It is clear that $\boldsymbol{a} \neq \boldsymbol{0}$ and $\boldsymbol{b} \neq \boldsymbol{0}$; otherwise $\boldsymbol{A} = \boldsymbol{0}$. $\square$