

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

- (a) Show that if $\{v_1, \dots, v_k\}$ is a basis for the nullspace of C , then $\{Pv_1, \dots, Pv_k\}$ is a basis for the nullspace of B .
- (b) Prove that $\text{rank}(B) = \text{rank}(C)$.
- (c) Let A be any $m \times p$ matrix. Suppose that the linear system $Ax = b$ is consistent for every $b \in \mathbb{R}^m$. Prove that the linear system $A^T y = 0$ has only the trivial solution.

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

$$CP^{-1}v = P^{-1}Bv = P^{-1}0 = 0,$$

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

$$c_1v_1 + \dots + c_kv_k = P^{-1}v.$$

Hence,

$$v = P(c_1v_1 + \dots + c_kv_k) = c_1Pv_1 + \dots + c_kPv_k.$$

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

$$v = d_1Pv_1 + \dots + d_kPv_k,$$

then

$$P^{-1}v = P^{-1}(d_1Pv_1 + \dots + d_kPv_k) = d_1v_1 + \dots + d_kv_k.$$

Since $\{v_1, \dots, v_k\}$ is a basis for the nullspace of C , the two representations of $P^{-1}v$ are the same, i.e., $c_i = d_i$, $i = 1, \dots, k$.

Since every vector in the nullspace of B can be uniquely written as a linear combination of $\{Pv_1, \dots, Pv_k\}$, we conclude that $\{Pv_1, \dots, Pv_k\}$ is a basis for the nullspace of B .

- (b) By (a), if $\text{nullity}(C) = k$, then $\text{nullity}(B) = k$; so $\text{rank}(B) = m - k = \text{rank}(C)$.
- (c) The system $Ax = b$ is consistent if and only if b belongs to the column space of A . If $Ax = b$ is consistent for all $b \in \mathbb{R}^m$, then the column space of A is \mathbb{R}^m . So

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

Therefore, $A^T y = 0$ has only the trivial solution. □

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

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

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

Proof. (a) Suppose that $P^{-1}AP = D_1$ and $P^{-1}BP = D_2$ are diagonal matrices. Then

$$\begin{aligned} AB &= (PD_1P^{-1})(PD_2P^{-1}) = PD_1D_2P^{-1} = PD_2D_1P^{-1} \\ &= (PD_2P^{-1})(PD_1P^{-1}) = BA. \end{aligned}$$

(b) Let v_1, \dots, v_n be eigenvectors of A associated to the n distinct eigenvalues. Then they are linearly independent. Let

$$P = \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}.$$

Then P is invertible and $P^{-1}AP$ is a diagonal matrix. By assumption, v_1, \dots, v_n are also eigenvectors of B . Then $P^{-1}BP$ is also a diagonal matrix. By (a), $AB = BA$.

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

$$A(Bv) = BA v = B(\lambda v) = \lambda(Bv).$$

Then Bv belongs to the eigenspace of A associated to λ .

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

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

2018/2019 SEMESTER 1

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

- (i) D has n rows.
- (ii) k is at least n .
- (iii) D is of full rank.

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

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

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

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

2(c) Let X, Y be square matrices of the same order. Prove the following statements.

- (i) $X^T X = 0$ if and only if $X = 0$.
- (ii) $XY = 0$ if and only if $X^T XY = 0$.

Proof. (i) If $X = 0$, it is clear that $X^T X = 0$.

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

(ii) If $XY = 0$, it is clear that $X^T XY = 0$.

Conversely, if $X^T XY = 0$, then $(XY)^T (XY) = Y^T X^T XY = 0$; by (i) $XY = 0$. □

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

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

Proof. (i) Suppose 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,

$$n \geq \dim E_{\lambda_1} + \dots + \dim E_{\lambda_n} \geq \underbrace{1 + \dots + 1}_n = n.$$

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

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

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

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

(iii) Let v_i be an eigenvector of A associated to λ_i , $i = 1, \dots, n$. Then v_1, \dots, v_n are linearly independent. Let

$$Q = \begin{pmatrix} v_1 & \dots & v_n \end{pmatrix}.$$

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

2017/2018 SEMESTER 2

5. Let A be an $n \times n$ matrix.

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

(c) Let v_1, v_2, v_3 be orthonormal vectors in \mathbb{R}^n . Suppose $w \in \mathbb{R}^n$. Define

$$b_1 = (Av_1) \cdot w, \quad b_2 = (Av_2) \cdot w, \quad b_3 = (Av_3) \cdot w, \quad q = b_1 v_1 + b_2 v_2 + b_3 v_3.$$

Calculate $v_1 \cdot q$, $v_2 \cdot q$ and $v_3 \cdot q$.

(d) Using the same definitions as in Part(c), show that for every $v \in \text{span}\{v_1, v_2, v_3\}$, $(Av) \cdot w = v \cdot q$.

Proof. (a) $(Au) \cdot w = (uA)^T w = (u^T A^T) w = u^T (A^T w) = u \cdot (A^T w)$.

(c) Note that $v_i \cdot v_i = 1$ and $v_i \cdot v_j = 0$ for $i \neq j$. Then

$$v_i \cdot q = v_i \cdot (b_1 v_1 + b_2 v_2 + b_3 v_3) = b_1 (v_i \cdot v_1) + b_2 (v_i \cdot v_2) + b_3 (v_i \cdot v_3) = b_i, \quad i = 1, 2, 3.$$

(d) Let $v \in \text{span}\{v_1, v_2, v_3\}$. Then there exist constant c_1, c_2, c_3 such that

$$v = c_1 v_1 + c_2 v_2 + c_3 v_3.$$

Recall that $v_i \cdot q = b_i = (Av_i) \cdot w = v_i \cdot (A^T w)$. Then

$$\begin{aligned} v \cdot q &= (c_1 v_1 + c_2 v_2 + c_3 v_3) \cdot q = c_1 (v_1 \cdot q) + c_2 (v_2 \cdot q) + c_3 (v_3 \cdot q) \\ &= c_1 [v_1 \cdot (A^T w)] + c_2 [v_2 \cdot (A^T w)] + c_3 [v_3 \cdot (A^T w)] \\ &= (c_1 v_1) \cdot (A^T w) + (c_2 v_2) \cdot (A^T w) + (c_3 v_3) \cdot (A^T w) \\ &= (c_1 v_1 + c_2 v_2 + c_3 v_3) \cdot (A^T w) \\ &= v \cdot (A^T w) = (Av) \cdot w. \end{aligned}$$

□

6. Let A be an $n \times n$ matrix.

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

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

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

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

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

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

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

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

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

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

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

$$\mathbf{A}^2 = (\mathbf{P}\mathbf{B}\mathbf{P}^{-1})^2 = \mathbf{P}\mathbf{B}^2\mathbf{P}^{-1} = \mathbf{P}\mathbf{I}\mathbf{P}^{-1} = \mathbf{P}\mathbf{P}^{-1} = \mathbf{I}.$$

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

2017/2018 SEMESTER 1

5. Let \mathbf{C} be a square matrix.

(a) Show that the nullspace of \mathbf{C} is a subset of the nullspace of \mathbf{C}^2 .

(b) If $\text{rank}(\mathbf{C}^2) = \text{rank}(\mathbf{C})$, show that the nullspace of \mathbf{C}^2 is equal to the nullspace of \mathbf{C} .

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

$$\mathbf{C}^2\mathbf{v} = (\mathbf{C}\mathbf{C})\mathbf{v} = \mathbf{C}(\mathbf{C}\mathbf{v}) = \mathbf{C}\mathbf{0} = \mathbf{0}.$$

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

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

$$\text{nullity}(\mathbf{C}) = n - \text{rank}(\mathbf{C}) = n - \text{rank}(\mathbf{C}^2) = \text{nullity}(\mathbf{C}^2).$$

We conclude that the nullspace of \mathbf{C} equals the nullspace of \mathbf{C}^2 . \square

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

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

(a) Write down the standard matrix for T_λ .

(b) For any $\lambda, \mu \in \mathbb{R}$, show that

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

(c) Suppose \mathbf{A} is diagonalizable and the eigenvalues of \mathbf{A} are $\lambda_1, \lambda_2, \dots, \lambda_k$.

(i) If \mathbf{v} is an eigenvector of \mathbf{A} , say, $\mathbf{A}\mathbf{v} = \lambda_i\mathbf{v}$ for some i , show that

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

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

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

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

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

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

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

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

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

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

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

$$S(\mathbf{v}) = c_1S(\mathbf{v}_1) + \cdots + c_nS(\mathbf{v}_n) = \mathbf{0}.$$

Therefore, S is the zero transformation. □

2016/2017 SEMESTER 2

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

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

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

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

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

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

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

6(b) Let A and B be square matrices of the same order.

(i) Prove that the nullspace of B is a subspace of the nullspace of AB .

(ii) Using (i) prove that

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

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

(ii) Suppose A and B have order n . Let $\{v_1, \dots, v_k\}$ be a basis for the nullspace of B and extend it to a basis $\{v_1, \dots, v_k, w_1, \dots, w_m\}$ for the nullspace of AB .

If $c_1 Bw_1 + \dots + c_m Bw_m = 0$ for some constants c_1, \dots, c_m , then

$$B(c_1 w_1 + \dots + c_m w_m) = 0,$$

which implies that $c_1 w_1 + \dots + c_m w_m$ belongs to the nullspace of B , that is,

$$c_1 w_1 + \dots + c_m w_m = d_1 v_1 + \dots + d_k v_k$$

for some constants d_1, \dots, d_k . Since $\{v_1, \dots, v_k, w_1, \dots, w_m\}$ is linearly independent, we must have $c_1 = \dots = c_m = d_1 = \dots = d_k = 0$. Hence, $\{Bw_1, \dots, Bw_m\}$ is linearly independent.

Note that $A(Bw_i) = (AB)w_i = 0$, Bw_i belongs to the nullspace of A . Hence,

$$\text{span}\{Bw_1, \dots, Bw_m\} \subseteq \text{nullspace of } A.$$

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

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

□

2016/2017 SEMESTER 1

5. Let A be a square matrix of order n .

(a) Let $E = \{e_1, \dots, e_n\}$ be the standard basis for \mathbb{R}^n . Show that Ae_j is the j^{th} column of A .

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

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

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

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

Proof. (a) Let v_j be the j^{th} column of A . Then $A = \begin{pmatrix} v_1 & \dots & v_n \end{pmatrix}$.

On the other hand, $A = AI = A \begin{pmatrix} e_1 & \dots & e_n \end{pmatrix} = \begin{pmatrix} Ae_1 & \dots & Ae_n \end{pmatrix}$. Therefore, $Ae_j = v_j$.

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

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

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

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

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

So

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

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

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

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

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

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

where $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$. Prove that for any nonzero vector $\mathbf{x} \in \mathbb{R}^n$, $\lambda_1 \leq \frac{\mathbf{x}^T \mathbf{C} \mathbf{x}}{\mathbf{x}^T \mathbf{x}} \leq \lambda_n$.

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

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

$$\mathbf{x} = (\mathbf{x} \cdot \mathbf{v}_1)\mathbf{v}_1 + \dots + (\mathbf{x} \cdot \mathbf{v}_n)\mathbf{v}_n,$$

and

$$\mathbf{x}^T \mathbf{x} = \mathbf{x} \cdot \mathbf{x} = (\mathbf{x} \cdot \mathbf{v}_1)^2 + \dots + (\mathbf{x} \cdot \mathbf{v}_n)^2.$$

We also have

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

and

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

Therefore,

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

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

$$\lambda_1 \leq \frac{\mathbf{x}^T \mathbf{C} \mathbf{x}}{\mathbf{x}^T \mathbf{x}} \leq \lambda_n. \quad \square$$

2015/2016 SEMESTER 2

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

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

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

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

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

and

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

So

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

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

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

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

2015/2016 SEMESTER 1

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

Proof. Suppose $W = \text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$. View each \mathbf{v}_i as a row vector and set $\mathbf{A} = \begin{pmatrix} \mathbf{v}_1 \\ \vdots \\ \mathbf{v}_k \end{pmatrix}$. Then

W is the row space of \mathbf{A} .

View \mathbf{w} as a column vector. Then

$$\mathbf{w} \in W^\perp \Leftrightarrow \mathbf{w} \text{ is orthogonal to } \mathbf{v}_i \text{ for all } i = 1, \dots, k$$

$$\Leftrightarrow \mathbf{v}_i \cdot \mathbf{w} = \mathbf{v}_i \mathbf{w} = 0 \text{ for all } i = 1, \dots, k$$

$$\Leftrightarrow \mathbf{A}\mathbf{w} = \begin{pmatrix} \mathbf{v}_1 \\ \vdots \\ \mathbf{v}_k \end{pmatrix} \mathbf{w} = \mathbf{0}$$

$$\Leftrightarrow \mathbf{w} \in \text{nullspace of } \mathbf{A}.$$

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

$$\dim(W) + \dim(W^\perp) = \text{rank}(\mathbf{A}) + \text{nullity}(\mathbf{A}) = n. \quad \square$$

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

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

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

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

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

$$c_1v + c_2Av + \cdots + c_nA^{n-1}v = 0.$$

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

$$c_kA^{k-1}v + c_{k+1}A^k v + \cdots + c_nA^{n-1}v = 0.$$

Pre-multiplication of A^{n-k} yields

$$0 = c_kA^{n-1}v + c_{k+1}A^n v + \cdots + c_nA^{2n-k-1}v = c_kA^{n-1}v.$$

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

Therefore, $\{v, Av, \dots, A^{n-1}v\}$ 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 $S = \{A^{n-1}v, \dots, Av, v\}$ is a basis for \mathbb{R}^n . Then P is the transition matrix from S to the standard basis E . For any $u \in \mathbb{R}^n$,

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

Then

$$P^{-1}AP[u]_S = P^{-1}Au = [Au]_S.$$

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

$$\begin{aligned} P^{-1}AP &= P^{-1}AP \begin{pmatrix} e_1 & e_2 & \cdots & e_n \end{pmatrix} \\ &= \begin{pmatrix} P^{-1}AP[A^{n-1}v]_S & P^{-1}AP[A^{n-2}v]_S & \cdots & P^{-1}AP[v]_S \end{pmatrix} \\ &= \begin{pmatrix} [A^n v]_S & [A^{n-1}v]_S & \cdots & [Av]_S \end{pmatrix} \\ &= \begin{pmatrix} 0 & e_1 & \cdots & e_{n-1} \end{pmatrix}. \end{aligned}$$

□

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

- (a) Let $A = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \end{pmatrix}$, where a_i is the i^{th} column of A . Show that $\{a_1, a_2, \dots, a_n\}$ is an orthogonal basis for \mathbb{R}^n .
- (b) Prove that $A = cP$ for some scalar c and orthogonal matrix P .

Proof. (a) Since A is invertible, $a_i \neq 0$ for all i . For $i \neq j$, the angle between e_i and e_j is 90° . Hence, the angle between $Ae_i = a_i$ and $Ae_j = a_j$ is also 90° . Then

$$\frac{a_i \cdot a_j}{\|a_i\| \|a_j\|} = \cos(90^\circ) = 0,$$

which implies that $a_i \cdot a_j = 0$. Therefore, $\{a_1, \dots, a_n\}$ is an orthogonal basis for \mathbb{R}^n .

- (b) For $i \neq j$, $(e_i + e_j) \cdot (e_i - e_j) = 0$, i.e., the angle between $e_i + e_j$ and $e_i - e_j$ is 90° . By assumption, the angle between

$$A(e_i + e_j) = a_i + a_j \quad \text{and} \quad A(e_i - e_j) = a_i - a_j$$

is also 90° . Hence,

$$\frac{(a_i + a_j) \cdot (a_i - a_j)}{\|a_i + a_j\| \|a_i - a_j\|} = \cos(90^\circ) = 0.$$

It follows that

$$(a_i + a_j) \cdot (a_i - a_j) = \|a_i\|^2 - \|a_j\|^2 = 0,$$

that is, $\|a_i\| = \|a_j\|$. Let $c = \|a_i\|$ and $v_i = \|a_i\|/c$. Then $\{v_1, \dots, v_n\}$ is an orthonormal basis for \mathbb{R}^n . Let $P = \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}$. Then P is an orthogonal matrix and

$$A = \begin{pmatrix} cv_1 & \cdots & cv_n \end{pmatrix} = c \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix} = cP. \quad \square$$

2014/2015 SEMESTER 2

- 4(b)** Let M and N be two $n \times n$ matrices. Suppose $\{v_1, v_2, \dots, v_n\}$ is a set of linearly independent eigenvectors for both M and N . Then $MN = NM$.

Proof. Let $P = \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}$. Then P is invertible such that $P^{-1}MP = D_1$ and $P^{-1}NP = D_2$ are diagonal matrices. Then

$$\begin{aligned} MN &= (PD_1P^{-1})(PD_2P^{-1}) = PD_1D_2P^{-1} \\ &= PD_2D_1P^{-1} = (PD_2P^{-1})(PD_1P^{-1}) = NM. \end{aligned} \quad \square$$

2014/2015 SEMESTER 1

- 5(b)** Given that $F: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ 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

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

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

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

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

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

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

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

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

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

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

$$\text{rank}(\mathbf{A}) = \text{rank}(T) = 2.$$

It follows that \mathbf{A} has full rank. □

2013/2014 SEMESTER 2

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

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

$$\mathbf{A}(k^{-1}\mathbf{v}) = k^{-1}(\mathbf{A}\mathbf{v}) = k^{-1}(k\mathbf{b}) = \mathbf{b}.$$

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

Suppose that \mathbf{v} is a least squares solution for $\mathbf{A}\mathbf{x} = \mathbf{b}$. Then $\mathbf{A}^T \mathbf{A}\mathbf{v} = \mathbf{A}^T \mathbf{b}$. So

$$\mathbf{A}^T \mathbf{A}(k\mathbf{v}) = k(\mathbf{A}^T \mathbf{A}\mathbf{v}) = k(\mathbf{A}^T \mathbf{b}) = \mathbf{A}^T (k\mathbf{b}).$$

Hence, $k\mathbf{v}$ is a least squares solution to $\mathbf{A}\mathbf{x} = k\mathbf{b}$. □

2013/2014 SEMESTER 1

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

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

Proof. If $d_i = \dim(E_{\lambda_i})$, write $S_{\lambda_i} = \{v_{i1}, \dots, v_{id_i}\}$.

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

$$u = c_{11}v_{11} + \dots + c_{1d_1}v_{1d_1} + \dots + c_{k1}v_{k1} + \dots + c_{kd_k}v_{kd_k},$$

for some constants c_{ij} . Let $v_i = c_{i1}v_{i1} + \dots + c_{id_i}v_{id_i}$. Then $v_i \in E_{\lambda_i}$ and

$$u = v_1 + \dots + v_k.$$

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

2012/2013 SEMESTER 2

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

- (i) If $\lambda \neq 0$, show that Bx is an eigenvector of BA with eigenvalue λ .
- (ii) If $\lambda = 0$, is Bx an eigenvector of BA with eigenvalue λ ? Justify your answer.

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

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

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

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

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

$$\|Au\| = \|u\|.$$

- (i) Prove that $Au \cdot Av = u \cdot v$ for any $u, v \in \mathbb{R}^n$.
- (ii) Using (i) or otherwise, prove that A is an orthogonal matrix.

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

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

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

Subtraction yields

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

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

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

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

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

Then the columns of \mathbf{A} form an orthonormal basis for \mathbb{R}^n . Hence, \mathbf{A} is an orthogonal matrix. \square

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

$$v = (v - Av) + Av,$$

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 A is diagonalizable.

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

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

$$\text{tr}(A) = \text{tr}(APP^{-1}) = \text{tr}(P^{-1}AP) = \text{tr}(D) = r = \text{rank}(A). \quad \square$$

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

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

Define $V - W$ by

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

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

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

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

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

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

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

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

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

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

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

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

where $\mathbf{b} = (c_1, \dots, c_n)^T$. It is clear that $\mathbf{a} \neq \mathbf{0}$ and $\mathbf{b} \neq \mathbf{0}$; otherwise $\mathbf{A} = \mathbf{0}$. □