Week 1

10 March 2021 13:30

- Matrices $\rightarrow$ Multiply them.

(1) $\rightarrow \begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix} \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$

$$= a_1 b_1 + \cdots + a_n b_n$$

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}_{m\times n} \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}_{n\times 1} = \begin{bmatrix} \\ \\ \end{bmatrix}_{m\times 1}$$

$$A = \begin{bmatrix} A_1 \\ \vdots \\ A_m \end{bmatrix} \qquad b = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$$

$$A_1, \ldots, A_m \in \mathbb{R}^{1\times n}$$

$$Ab = \begin{bmatrix} A_1 b \\ \vdots \\ A_m b \end{bmatrix}$$

$A \in \mathbb{R}^{m\times n}$ $\qquad B \in \mathbb{R}^{n\times p}$

$B = \begin{bmatrix} b_1 & \cdots & b_p \end{bmatrix}$; $b_1, \ldots, b_p \in \mathbb{R}^{n\times 1}$

$$AB = \begin{bmatrix} Ab_1 & \cdots & Ab_p \end{bmatrix} \in \mathbb{R}^{m\times p}$$

$\in \mathbb{R}^{m\times 1}$

$A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times n}$

We say that B is an inverse of A if
$AB = I = BA$.

Fact. (Will see later) $AB = I \Rightarrow BA = I$
(this was not clear, a priori)

$\rightarrow$ Functions $f, g : X \longrightarrow X$. ($X \neq \emptyset$ is some set.)
If $(f \circ g)(x) = x \quad \forall x \in X$,
is it necessary that $(g \circ f)(x) = x \quad \forall x \in X$?

No. Find example.

$Ax = b$. — $(*)$ $\quad A \in \mathbb{R}^{m\times n}$, $x \in \mathbb{R}^{n\times 1}$, $b \in \mathbb{R}^{m\times 1}$

If A is upper triangular, it is easy by back-substitution. (Whether consistent or not is also clear.)

Idea: Do operations on both A and b to get something as above.

$\rightarrow$ If $Ax_0 = b$, i.e., x_0 is a particular sol^n,
and $S = \{x \in \mathbb{R}^{n\times 1} : Ax = 0\}$.
Then, all solutions of $(*)$ are precisely of the form $x_0 + s$ for some $s \in S$.

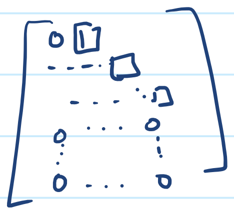
Idea: Row echelon form (REF)

(i) All zero rows at bottom. (Possibly none.)

(No zero row can be above a nonzero row.)

first $\neq 0$ element from left

(2) Pivots should be strictly from left to right as you go from top to bottom.



Week 2

17 March 2021 09:42

Outline:

1. Recall REF. n variables, r pivots ⇒ (n - r) free variables
2. **Ax** = **0** has ***only*** the zero solution ⇔ n = r ← every column has a pivot
3. EROs
4. GEM
5. **Ax** = **0** has ***only*** the zero solution ⇔ any REF of **A** has n non-zero rows
6. Inverse
7. **Ax** = **0** has ***only*** the zero solution ⇔ A is invertible
8. Let **A**, **B** ∈ ℝ^{n × n}. **AB** = **I** ⇔ **BA** = **I**
9. RCF. REF + pivots are 1 + the entries above the pivots are 0s
10. **A** can be transformed to **I** via EROs ⇔ **A** is invertible
11. GJM
12. Linear (in)dependence
13. Row rank
14. Given n column vectors, make a matrix with those as columns and find its row rank r. We know r ≤ n. The vectors are linearly independent ⇔ r = n.
15. EROs don't change row rank. Thus, **A** and REF(**A**) have the same row rank.
16. If **A'** is in REF, then row-rank(**A'**) = number-of-non-zero-rows(**A'**).

$A \in \mathbb{R}^{m \times n}$ (items 1–5)

$A \in \mathbb{R}^{n \times n}$ (items 6–7)

---

3. EROs → Elementary Row operations

Type I : Interchange two rows

Type II : Add a scalar multiple of R_i to R_j where $i \neq j$.

Type III : Multiplying a row with a non-zero scalar

---

4. GEM → Gauss Elimination Method

↳ Algo to convert a matrix into an REF using EROs.

---

5. # non-zero rows of A' = # pivots of A' (A' is in REF)

5 follows from 2.

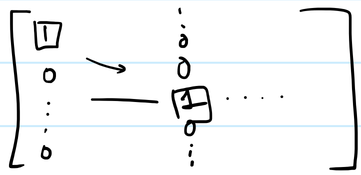
---

6. If $A \in \mathbb{R}^{n \times n}$, then $B \in \mathbb{R}^{n \times n}$ is ~~an~~ the inverse of A if $AB = I = BA$.

9. RCF if (i) it is REF

(ii) it has all pivots as 1

(iii) everything above pivot are also 0



RCF is unique. (REF need not be.)

10. A is invertible $\Leftrightarrow$ RCF of A is I

$\Leftrightarrow$ A can be transformed to I via EROs

11. Take $A \in \mathbb{R}^{n \times n}$.

Make the augmented matrix

$$[A \mid I]$$

performs EROs to make A into its RCF (so same operations on I as well)

$$[A' \mid B]$$

If A is inv., then $A' = I$ and $B = A^{-1}$.

If A is not inv., then $A' \neq I$.

11. Linear dependence

- $S \subset \mathbb{R}^{n \times 1}$ (or $\mathbb{R}^{1 \times n}$) (possibly infinite)

- S is linearly dependent if there exist (distinct) $v_1, \dots, v_s \in S$ and $\alpha_1, \dots, \alpha_s \in \mathbb{R}$, not all zero such that

$$\alpha_1 v_1 + \cdots + \alpha_s v_s = \mathbf{0}$$

in $\mathbb{R}^{n\times 1}$ (or $\mathbb{R}^{1\times n}$)

- For example, if $\alpha_1 \neq 0$ and $n \geq 2$, then

$$v_1 = -\frac{1}{\alpha_1}(\alpha_2 v_2 + \cdots + \alpha_s v_s).$$

- if $\mathbf{0} \in S$, then S is lin. dep.

  Take $n=1$, $v_1 = \mathbf{0}$, $\alpha_1 = 1 \neq 0$.

  Then, $1 \cdot \mathbf{0} = \mathbf{0}$.

- If $S = \{v\}$ and $v \neq \mathbf{0}$. Then, S is not lin. dep.

- if $S = \emptyset$, then S is not lin. dep.

- S is linearly independent if S is not linearly dependent.

- $\emptyset$ is lin. indep. $\{v\}$ is lin indep iff $v \neq \mathbf{0}$.

---

13. row-rank (A) = maximum no. of lin. indep rows of A.

if $A = \mathbf{0}$, then row-rank$(A) = 0$.

$$\text{row-rank} \begin{bmatrix} [1 & 1] \\ 2 & 2 \end{bmatrix} = 1$$

this is lin indep

$\{[1 \quad 1], [2 \quad 2]\}$ is lin. dep.

15. In general, row-rank (A) = row-rank (A')
where A' is an REF of A.

Week 4

31 March 2021 10:47

Outline:

1. Linear transformations
2. Model example
3. M^E_F(T)
4. Composite
5. Null space, image space (relate with A, T_A)
6. Eigen(value, vector, space)
7. Characteristic polynomial
8. Algebraic, geometric multiplicity
9. Similarity of square matrices
10. When is B ~ A?
11. Diagonalisable, how do we get P?

1. V, $W \longrightarrow$ vector spaces over $\mathbb{K}$

($\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$)

A linear transformation from V to W is a function

$$T : V \longrightarrow W$$

with the following properties:

$$\text{(i)} \quad T(v_1 + v_2) = T(v_1) + T(v_2) \qquad \forall\, v_1, v_2 \in V,$$
$$\text{(ii)} \quad T(\alpha v) = \alpha \cdot T(v) \qquad \forall\, \alpha \in \mathbb{K},\ \forall\, v \in V.$$

Consequences:

$$\text{(i)} \quad T(\underset{\substack{\shortparallel \\ V}}{\mathbf{0}}) = \underset{\substack{\shortparallel \\ W}}{\mathbf{0}}$$

(ii) For all $\alpha_1, \dots, \alpha_s \in \mathbb{K}$ and $v_1, \dots, v_s \in V$:

$$T(\alpha_1 v_1 + \dots + \alpha_s v_s) = T(\alpha_1 v_1) + \dots + T(\alpha_s v_s)$$
$$= \alpha_1 T(v_1) + \dots + \alpha_s T(v_s).$$

2. Let $A \in \mathbb{R}^{m \times n}$. This gives a linear transformation

$$T_A : \mathbb{R}^{n \times 1} \longrightarrow \mathbb{R}^{m \times 1}$$

defined as

$$I_A(x) = Ax$$

3. $\underline{M^E_F(T)}$.

Let $T: \overset{\text{dim}=n}{V} \longrightarrow \overset{\text{dim}=m}{W}$ be a lin. transf.

Fix ordered bases E of V and F of W.

Say, $E = (v_1, \ldots, v_n)$ and
$F = (w_1, \ldots, w_m)$.

The matrix $M = M^E_F(T)$ is defined as:

EW

(i) Compute $T(v_1)$ and write it as a lin. combination of F. (Can do this since F is a basis of W.)

(This combination is unique.)

$$T(v_1) = a_{11} w_1 + a_{21} w_2 + \cdots + a_{m1} w_m.$$

The first column of M is $\begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}$.

(ii) Do the same for $T(v_2)$.

$\vdots$

(n) Do it for $T(v_n)$.

$M^{\text{dom}}_{\text{codomain}}$

4. v. spaces $V \xrightarrow{T} W \xrightarrow{S} U$

4. v. spaces $V \xrightarrow{T} W \xrightarrow{S} U$

(ordered!) bases $E \subseteq V$, $F \subseteq W$, $G \subseteq U$

T and S are lin. transf.

Note $S \circ T : V \to U$ is also linear. (Check!)

$$M_G^E(S \circ T) = M_G^F(S) M_F^E(T).$$

5. $T : V \to W$ lin. transf.

$$\mathcal{N}(T) := \{ v \in V : T(v) = \underset{W}{\underset{\Uparrow}{0}} \} \subseteq V$$

vector subspaces of V and W

$$\mathcal{I}(T) := \{ w \in W : \exists v \in V \text{ s.t. } T(v) = w \} \subseteq W$$

If $V = \mathbb{R}^{n \times 1}$, $W = \mathbb{R}^{m \times 1}$, $A \in \mathbb{R}^{m \times n}$, then

$$\mathcal{N}(T_A) = \mathcal{N}(A) \text{ and}$$
$$\mathcal{I}(T_A) = \mathcal{C}(A).$$

6. Let $A \in \mathbb{K}^{m \times n}$.

Suppose $v \in \mathbb{K}^{n \times 1} \setminus \{0\}$ and $\lambda \in \mathbb{K}$ is such that

$$Av = \lambda v.$$

Then, v is called an eigenvector of A and λ an eigenvalue.

The eigenspace of λ is defined as

$$\mathcal{N}(A - \lambda I) = \{ v \in \mathbb{R}^{n \times 1} : Av = \lambda v \}.$$

All eigenvectors along with $\mathbf{0}$.

7. Let $p_A(t) := \det(A - tI)$.

This is the characteristic polynomial of A.

<u>Thm.</u> $\lambda \in \mathbb{K}$ is an e-val of $A \Leftrightarrow p_A(\lambda) = 0$.

8. geometric multiplicity of $\lambda := \dim(\mathcal{N}(A - \lambda I))$
algebraic multiplicity of $\lambda :=$ largest m s.t. $(t-\lambda)^m$ is a factor of $p_A(t)$.

9. Let $A, B \in \mathbb{K}^{n \times n}$.

$A \sim B \overset{\text{def}^n}{\Longleftrightarrow} \exists P \in \mathbb{K}^{n \times n}$ invertible such that $P^{-1} A P = B$

<u>Check</u>: $\sim$ is an equivalence relation.

11. $A \in \mathbb{K}^{n \times n}$ is said to be diagonalisable if A is similar to a diagonal matrix.

Proposition

A matrix $\mathbf{A} \in \mathbb{K}^{n\times n}$ is diagonalizable if and only if there is a basis for $\mathbb{K}^{n\times 1}$ consisting of eigenvectors of $\mathbf{A}$. In fact,

$$\begin{aligned} & \mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{D}, \text{ where } \mathbf{P}, \mathbf{D} \in \mathbb{K}^{n\times n} \text{ are of the form} \\ & \mathbf{P} = \begin{bmatrix} \mathbf{x}_1 & \cdots & \mathbf{x}_n \end{bmatrix} \text{ and } \mathbf{D} = \operatorname{diag}(\lambda_1, \dots, \lambda_n) \\ \iff \quad & \{\mathbf{x}_1, \dots, \mathbf{x}_n\} \text{ is a basis for } \mathbb{K}^{n\times 1} \text{ and} \\ & \mathbf{A}\mathbf{x}_k = \lambda_k \mathbf{x}_k \text{ for } k = 1, \dots, n. \end{aligned}$$

Week 5

07 April 2021 13:30

- Diagonalisability

(i) Let $A \in \mathbb{K}^{n \times n}$ and $\lambda \in \mathbb{K}$ be an eigenvalue of A

- alg-mult of λ = $AM(\lambda) :=$ largest $m \in \mathbb{N}$ s.t $(t-\lambda)^m$ divides $p_A(t) = \det(A - tI)$

- geo-mult of λ = $GM(\lambda)$ = nullity $(A - \lambda I)$

Note if λ is an e-val, then $GM(\lambda) \geq 1$, by def^n

- In general, $GM(\lambda) \leq AM(\lambda)$

(ii) Let $\lambda_1, \ldots, \lambda_k \in \mathbb{K}$ be all the eigenvalues of A
Then,
A is diagon'ble $\Leftrightarrow$ $GM(\lambda_1) + \cdots + GM(\lambda_k) = n$
In particular, diagon'ble $\Rightarrow$ $GM(\lambda_i) = AM(\lambda_i)$ $\forall i \in \{1, \ldots, k\}$
Corollary If A has n distinct eigenvalues, then A is diagonalisable.

(Even if $k < n$, the matrix MAY be diagonalisable Eg $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$)

(iii) Procedure for checking diagonalisability of $A \in \mathbb{K}^{n \times n}$

(I) Compute $p_A(t) = \det(A - tI)$

(II) Find all roots $\lambda_1, \ldots, \lambda_k \in \mathbb{K}$ of $p_A(t)$

(III) Compute $GM(\lambda_1), \ldots, GM(\lambda_k)$
[Convert $A - \lambda_1 I$ to an REF to get rank. Use rank-nullity theorem to get nullity $(A - \lambda_1 I) =: GM(\lambda_1)$]

[rank-nullity theorem to get nullity $(A - \lambda_i I) = \mathrm{GM}(\lambda_i)$]

(III) If $\sum_{i=1}^{k} \mathrm{GM}(\lambda_i) = n$, then "diagonalisable",

else, "not diagonalisable"

(IV) Suppose A is diagon'ble, how do we get an invertible $P \in \mathbb{k}^{n \times n}$ s.t. $P^{-1}AP$ is diagonal?

For each $\lambda_1, \ldots, \lambda_k$ as in (II), compute a basis for $\mathcal{N}(A - \lambda_i I)$.

↳ this was the eigenspace of A corresp to λ_i

[Again, convert to an REF and calculate the basic sol^{n}s of $(A - \lambda_i I)x = \mathbf{0}$]

Then, the union of these bases will have n elements, say $v_1, \ldots, v_n \in \mathbb{k}^{n \times 1}$

Construct $P = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} \in \mathbb{k}^{n \times n}$

This is a desired P

(You can have multiple Ps. In fact, take $P' = \alpha P$ for $\alpha \in \mathbb{k} \setminus \{0\}$)

- Inner product

(1) $\langle \, , \, \rangle : \mathbb{k}^{n \times 1} \times \mathbb{k}^{n \times 1} \longrightarrow \mathbb{k}$ satisfying

- $\langle v, v \rangle \geq 0 \quad \forall v \in \mathbb{k}^{n \times 1}$ and $\langle v, v \rangle = 0 \iff v = \mathbf{0}$

- $\langle u, \alpha v + v' \rangle = \alpha \langle u, v \rangle + \langle u, v' \rangle$

$\forall\, u, v, v' \in \mathbb{K}^{n \times 1}$ and $\alpha \in \mathbb{K}$

- $\langle u, v \rangle = \overline{\langle v, u \rangle}$ (If $\mathbb{K} = \mathbb{R}$, then $\langle u, v \rangle = \langle v, u \rangle$)

- $\|v\| = \sqrt{\langle v, v \rangle}$

(i) Projection let $u, v \in \mathbb{K}^{n \times 1}$

Suppose $v \neq 0$ Then,

$$P_v(u) = \frac{\boxed{\langle v, u \rangle}}{\langle v, v \rangle} v = \frac{\langle v, u \rangle}{\|v\|^2} v$$

Note that this was updated There was originally an error

$$\langle u - P_v(u), v \rangle = 0$$

That is, $(u - P_v(u)) \perp v$

(ii) G-S OP

Start with $(w_1, \ , w_k)$ where $w_1, \ , w_k \in \mathbb{K}^{n \times 1}$

Compute

$$v_1 := w_1,$$
$$v_2 = w_2 - P_{v_1}(w_2),$$
$$v_3 := w_3 - P_{v_2}(w_3) - P_{v_1}(v_3),$$
$$\vdots$$
$$v_k = w_k - P_{v_1}(w_k) - \ - P_{v_{k-1}}(w_k)$$

If some v_i is 0, ignore the $P_{v_i}(w_j)$ term

Then, $(v_1, \ , v_k)$ are orthogonal

Moreover, the cumulatative span (from the beginning) is mantained, i.e.,

$$\operatorname{span}\{w_1, \ , w_j\} = \operatorname{span}\{v_1, \ldots, v_j\} \text{ for all } j \in \{1, \ldots, k\}$$

- Application Suppose $\{w_1, \ldots, w_k\}$ is a basis of some subspace $V \subset \mathbb{K}^{n\times 1}$
  Then, using GSOP, we can an <u>orthogonal</u> basis for V. Further, we can divide by the norm to get an <u>ortho NORMAL</u> basis

- Benefit of orthonormal basis?
  Suppose $(u_1, \ldots, u_k)$ is an orthoNORMAL basis for V
  Let $b \in V$. We know that $\exists\, \alpha_1, \ldots, \alpha_k \in \mathbb{K}$ s.t.

$$b = \alpha_1 u_1 + \cdots + \alpha_k u_k$$

Q How to get $\alpha_1, \ldots, \alpha_k$?
(In general, need to solve a system of equation
NOT very good)

But here, we have it more easily as $\underline{\underline{\alpha_i = \langle u_i, b\rangle}}$

Week 7

21 April 2021 12:48

1. Spectral Theorem
2. Perpendicular complement
3. $\oplus$
4. Projection
5. Best approximation

1. A matrix $A \in K^{n\times n}$ is called normal if $AA^* = A^*A$.

↳ Self-adjoint : $A = A^*$

Skew self-adjoint: $A = -A^*$

Unitary : $AA^* = I = A^*A$ (these are invertible)

Thm. (Spectral theorem for normal matrices)

Let $A \in \mathbb{C}^{n\times n}$. Suppose that A is normal.

Then, $\exists$ a unitary matrix $U \in \mathbb{C}^{n\times n}$ and a diagonal matrix $D \in \mathbb{C}^{n\times n}$ s.t.

$$\underset{\underset{U^*AU}{\parallel}}{U^{-1}AU} = D.$$

→ If $A = A^*$, then $D \in \mathbb{R}^{n\times n}$.

(That is, all eigenvalues of A are real, even if A has non-real entries)

→ If $A = -A^*$, then $iD \in \mathbb{R}^{n\times n}$.

(That is, all e-vals of A are purely imaginary.)

In particular, we always have an orthogonal (and hence, orthonormal) eigenbasis. The columns of U will form an ortho NORMAL eigenbasis.

2. Let V be an inner product space and $E \subset V$. (E need not be a subspace.)

$$E^{\perp} := \{ v \in V : \langle v, x \rangle = 0 \text{ for all } x \in E \}.$$

(i) $E^{\perp}$ is always a subspace.

(ii) Suppose $Y \subseteq V$ is a subspace. Then, $Y^{\perp}$ is also a subspace. Moreover, $V = Y \oplus Y^{\perp}$.

---

3. Let V be a vector space. Let $U, W \subseteq V$ be subspaces. We write

$$V = U \oplus W$$

if

(i) Every $v \in V$ can be written as $v = u + w$ for some $u \in U$ and $w \in W$,

(ii) the u and w above are unique (depend only on v).

For example, consider $V = \mathbb{R}^{2\times 1}$, $U = \text{span}\{e_1\}$ and $W = \text{span}\{e_2\}$.
Then, $V = U \oplus W$.

Suppose $v \in V$. Then, $v = [a \;\; b]^T$ for some $a, b \in \mathbb{R}$.

(i) $v = \underset{\substack{\parallel \\ u}}{a e_1} + \underset{\substack{\parallel \\ w}}{b e_2}$.

$u \in U$, $w \in W$.

(ii) Suppose $\exists\, u', w'$ s.t. $v = u' + w'$.
Then,

$$\underset{\substack{\cap \\ U}}{u} - u' = w' - \underset{\substack{\cap \\ W}}{w}.$$

$\therefore \quad u - u' \in U \cap W.$ But $U \cap W = \{0\}$.

$$\therefore \quad u - u' = 0 = w' - w.$$

$\therefore \quad u = u'$ & $w = w'$. ∎

→ $V = \mathbb{R}^{2\times 1}$, $U = \mathbb{R}^{2\times 1}$, $W = \text{span}\{e_2\}$.

Then, (i) is true but (ii) is not.

Thus, "$V = U \oplus W$" is not true.

---

4. Suppose V is v. space and $Y, Z \subseteq V$ are subspaces s.t.

$$V = Y \oplus Z.$$

Then, every $v \in V$ can be written

$v = y + z$ for some <u>unique</u> $y \in Y$ and $z \in Z$.

Define $P : V \longrightarrow Y$ by

$P(v) = y$. (where y is as above)

(General Projection.)

Then P is well-defined and linear.

Moreover, $\mathcal{N}(P) = Z$ and $\mathcal{I}(P) = Y$.

By rank-nullity,

$$\dim(V) = \underbrace{\dim(Y)}_{\text{rank}} + \underbrace{\dim(Z)}_{\text{nullity}}.$$

then P is called the orthogonal projection onto Y.

---

3. In general, given a v. space V and subspaces $Y_1, \dots, Y_k \subseteq V$, we write

$$V = Y_1 \oplus \cdots \oplus Y_k$$

if

(i) Every $v \in V$ can be written as
$v = y_1 + \dots + y_k$ for $y_i \in Y_i$ $(1 \le i \le k)$;

(ii) The $y_i \in Y_i$ above are uniquely determined.

Then, once again, we can define the projections

$$P_i : V \longrightarrow Y_i \text{ by } P_i(v) := y_i. \quad (1 \le i \le k)$$

---

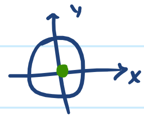
5. Best approximations.

Def^n. Let $E \subseteq \mathbb{K}^{n\times 1}$ be any subset and $b \in \mathbb{K}^{n\times 1}$. $a \in E$ is called a best approximation of b from E if

$$\| b - a \| \le \| b - x \| \quad \text{for all } x \in E.$$

Note: There may be none, one, many, infinitely many best approximations.

$= \{ x \in \mathbb{R}^{2\times 1} : \|x\| = 1 \}$



Then, every $a \in E$ is a best approximation for $b = \mathbf{0} = [0 \;\; 0]^T$.

Others: try yourself.

In this case that E is a subspace, the best approximation (exists and is unique) and is given by $P_E(b)$, where P_E is the orthogonal projection onto E.