Chapter 2

December 13, 2016

1 Direct Sum

Definition *Let V be a vector space and let M_1, M_2, ..., M_k are subspaces of V. The* ***sum*** *of these subspaces M is defined as*

$$M = \{m \in V : \quad m = m_1 + m_2 + \ldots + m_k \quad \textit{where } m_i \in M_i, \quad i = 1, 2, \ldots, k\}.$$

Theorem *The sum of subspaces is also a subspace of V.*

Proof:

Definition *Let M_1, M_2, ..., M_k be subspaces of a vector space V. These subspaces are* *said to be* ***linearly independent*** *if,*

$$m = m_1 + m_2 + \ldots + m_k = 0, \quad \textit{where } m_i \in M_i \quad \textit{implies}$$

$$m_1 = m_2 = \ldots = m_k = 0 \quad \textit{for} \quad i = 1, 2, \ldots, k$$

Definition *Let M_1, M_2, ..., M_k be subspaces of a vector space and also let*

- $M = M_1 + M_2 + \ldots + M_k$
- *M_1, M_2, ..., M_k are linearly independent*

Then M is said to be the ***direct sum*** *of subspaces M_1, M_2, ..., M_k and denoted by $M = M_1 \oplus M_2 \oplus \ldots \oplus M_k$*

Definition *If $M = V$ (the linear space itself) then $V = M_1 \oplus M_2 \oplus \ldots \oplus M_k$ is called the direct sum decomposition of V.*

Example: Let V=$\mathbb{R}^4$, $x = [x_1, x_2, x_3, x_4]^T \in \mathbb{R}$.

Definition *Let V be an inner product space. Two subspaces M_1 and M_2 are said to be* ***orthogonal*** *if,*

$$\langle m_1, m_2 \rangle = 0 \quad \forall m_1 \in M_1, m_2 \in M_2.$$

Orthogonality is denoted as $M_1 \perp M_2$

Definition *Let $M = M_1 \oplus M_2 \oplus \ldots \oplus M_k$ and let $M_i \perp M_i$ for all $i \neq j$. Then M is said to be **orthogonal direct sum** of subspaces $M_1, M_2, \ldots, M_k$.*

Symbolically,

Definition *Let M be a subspace of an inner product space V. The **orthogonal complement** $M^\perp$ of the subspace M is defined as*

$$M^\perp := \{x \in V : \langle x, m \rangle = 0 \forall m \in M\}.$$

Theorem *$M^\perp$ is itself a subspace.*

Proof:

Example: $V = \mathbb{R}^3$, $M = Span(\left\{ \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\})$, $\quad M^\perp =?$

Theorem *Let V be an inner product space and M is a subspace of V. V can always be written as the direct sum of a subspace and its orthogonal complement, i.e., we always have* $V = M \oplus M^{\perp}$.

Proof:

2 Projection Theorem

Theorem *"Projection Theorem"*

Let H be a Hilbert space (inner product space, complete w.r.t the norm induced by the inner product) and let M be a finite dimensional subspace of H. For any $x \in H$, the following minimization problem has a solution.

$$\min_{m \in M} \|x - m\|$$

(i.e., we can find the closest vector to x lying in the subspace M).

Proof:

Remark: $m^* = x_1$ can be interpreted as the "best approximation" of x chosen from the vectors in M. Vector x_2 can be interpreted as the "error in the approximation". This error must be orthogonal to the subspace.

Computational aspects of the projection theorem: Suppose we are given a basis for M. That is, $M = Span(\{v_1, v_2, \dots v_k\})$. Given $x \in H \supset M$, we want to figure out $x_1 \in M$, where $x = x_1 + x_2$ for $x_2 \in M^{\perp}$.

Example: Let H be the space of square integrable functions with domain $[\pi, \pi]$ with inner product $\langle f_1, f_2 \rangle = \int_{-\pi}^{\pi} f_1(t)\overline{f_2(t)}dt$. Let M be the subspace $M = Span\{\frac{e^{jkt}}{2\pi}\}_{k=-N}^{N}$. Note that dimension of M is $2N+1$ and the basis set is orthonormal.

$$\langle f_n, f_m \rangle = \int_{-\pi}^{\pi} e^{j(n-m)t}dt = \delta_{n,m}$$

Now, let $g \in H$ be an arbitrary vector (a function). Then $g = g_1 + g_2$, where $g_1 \in M$ and $g_2 \in M^{\perp}$.

Note that, $g_1(t)$ is the **best approximation** to $g(t)$ within the subspace M. $g_1(t)$ turns out to be the finite Fourier series representation of $g(t)$. As $N \to \infty$ we obtain the **Fourier series** representation.

<u>Special Case:</u> "Application of the projection theorem in $\mathbb{C}^n$". Let $\{m_1, m_2, \ldots, m_k\}$, $k < n$ be a basis for a subspace M of $\mathbb{C}^n$. That is, $M = Span(\{m_1, m_2, \ldots, m_k\})$. Given an arbitrary vector $x \in \mathbb{C}^n$, we know that $x = x_1 + x_2$ with $x_1 \in M$, $x_2 \in M^{\perp}$. We also know that x_1 and x_2 are unique. Let $x_1 = \sum_{i=1}^{k} \alpha_i m_i$. Define matrix $B = [m_1 \; m_2 \; \ldots m_k]$

whose columns are basis vectors. Then we can write $x_1 = B\alpha$ for $\alpha = [\alpha_1, \ldots, \alpha_k]^T$.

Remark: An orthogonal projection matrix $P \in \mathbb{C}^{k\times k}$ satisfies:

- $P^* = P$
- $P^2 = P \Rightarrow P^i = P$ for all $i = 1, 2, \ldots$

Remark: In $\mathbb{C}^n$ the standard inner product is $\langle x, y\rangle = y^* x$. In $\mathbb{R}^n$, this boils down to $\langle x, y\rangle = y^T x$

Example: "Orthogonal projection" Find the orthogonal projection of the vector $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ onto the subspace spanned by $\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \right\}$

3 Projection Theorem & Solution of Linear Equations

consider the linear equation expressed as

$$Ax = b \quad \text{where } A \in \mathbb{C}^{m\times n} \quad \& \quad b \in \mathbb{C}^{m\times 1} \quad \& \quad x \in \mathbb{C}^{n\times 1}.$$

Is there a solution to x? If the answer is yes, is it unique?

Remark:

- A solution exists if and only if $b \in R(A)$.
- A solution is unique if and only if $N(A) = 0$.

Example: Let $A = [1\ 1\ 1\ 1]^T$ and $b = [2.2 \quad 1.9 \quad 2.1 \quad 1.8]^T$.

When there is no exact solution, one can try to find the "‘best approximation"’ to a solution. An approximation can be found by minimizing the norm of the error,

$$\min_{x\in\mathbb{C}^n} \|Ax - b\|^2$$

If a solution exists, then $\|Ax - b\|^2 = 0$, otherwise we can find an approximate solution such that $\hat{x} = \arg\min_{x \in \mathbb{C}^n} \|Ax - b\|^2$.

Example: The length x of a metal rod is inaccurately measured four times and the results are recorded as l_1, l_2, l_3, and l_4. What is the best approximation to x?

Example: Consider the following scenario:

Remark: Suppose $\hat{x}$ is a solution of $Ax = b_1$. Suppose that m is any vector in $N(A)$, then $\hat{x} + m$ is another solution.

In the case of non-uniqueness, we are going to look for a solution with the minimum norm.

$$\min_{Ax=b_1} \|x\|$$

Let $\hat{x}^a$ and $\hat{x}^b$ be two solutions to $Ax = b$. We can decompose both solutions uniquely as:

Theorem $N(A)^{\perp} = R(A^{*})$.

Proof:

Example: Consider the previous example:

Summary:

4 Special Cases of $Ax = b$

4.1 Columns of A form a linearly independent set

A is full-column rank.

4.2 Rows of A form a linearly independent set

A is full-row rank.

4.3 Both rows & columns of A form a linearly independent set

A is invertible.

5 Spectral Analysis of Linear Operators

Definition *Let $A : V \to V$ be a linear transformation defined over the vector space V. A subspace M of V is said to be* ***invariant*** *under A if $A(x) \in M$ for all $x \in M$.*

Example: R(A) is invariant under A.

Example: N(A) is invariant under A.

Definition *Powers of a linear operator are defined as,*

$$A^k(x) = \underbrace{A(A(\ldots A(x)\ldots))}_{A \text{ applied } k \text{ times}}$$

By using the above definition polynomials of A can be constructed as linear combinations of powers of A.

Example: $P(A) = \alpha_0 A^n + \alpha_1 A^{n-1} + \ldots + \alpha_{n-1} A + \alpha_n I$.

Exercise: Show that $R(P(A))$ and $N(P(A))$ are invariant under A.

Definition *Let A denote the matrix representation of a linear operator from V to V (A is a square matrix). The* ***eigenvalues*** *of A, denoted by λ_i, are defined as the n (n = dim V) roots of the equation det(sI-A)=0, where det(sI-A) is known as the* ***characteristic polynomial of*** *A.*

Definition *Vector(s) $e_i \in V$ satisfying $e_i \neq 0$ and $Ae_i = \lambda_i e_i$ is called the **eigenvector(s)** of A corresponding to eigenvalue λ_i.*

Example: Let $A : \mathbb{C}^n \to \mathbb{C}^n$ and λ_i be an eigenvalue of A. $N(A - \lambda_i I)$ is invariant under A.

Proof:

Theorem *Let $A \in \mathbb{C}^{n \times n}$ be the matrix representation of a linear transformation $T : \mathbb{C}^n \to \mathbb{C}^n$ with respect to the canonical basis. Suppose that,*

- $\mathbb{C}^n = M_1 \oplus M_2 \oplus \ldots \oplus M_k$
- *Each subspace M_i is invariant under T.*

Let $dim(M_i) = n_i$ and M_i has a basis set $\{b_1^i, b_2^i, \ldots, b_{n_i}^i\} =: B_i$. Then with respect to basis $\{b_1^1, b_2^1, \ldots, b_{n_1}^1;\ \ b_1^2, b_2^2, \ldots, b_{n_2}^2;\ \ \ldots\ \ ;b_1^k, b_2^k, \ldots, b_{n_k}^k\}$, transformation T has a block *diagonal matrix representation.*

$$\bar{A} = \begin{bmatrix} \bar{A}_1 & 0 & \dots & 0 \\ 0 & \bar{A}_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \bar{A}_k \end{bmatrix},$$

where $\bar{A}_i \in \mathbb{C}^{n_i \times n_i}$. *In particular,* $\bar{A}_i = B^{-1}AB$, *where* $B \in \mathbb{C}^{n_i \times n_i}$ *is given by* $B = [B_1, B_2, \dots, B_{n_i}]$

Proof:

Example: Let $A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -2 & -5 & -4 \end{bmatrix}$, $M_1 = Span \left\{ \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\}$, $M_2 = Span \left\{ \begin{bmatrix} 1 \\ -2 \\ 4 \end{bmatrix} \right\}$

i) Is M_1 invariant under A?

ii) Is M_2 invariant under A?

iii) Find $\bar{A}$?

Let A be an $n \times n$ matrix with n distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ Let $e_1, e_2, \ldots, e_n$ be the eigenvectors corresponding to these eigen values, i.e., $Ae_i = \lambda_i e_i$, $i = 1, 2, \ldots, n$.

Claim: The set of eigenvectors $\{e_1, \quad e_2, \quad \ldots, \quad e_n\}$ form a linearly independent set. Furthermore, $N(A - \lambda_i I) = Span(e_i)$ for all i.

Proof:

Theorem *"'Cayley-Hamilton"' Every $n \times n$ matrix satisfies its own characteristic equation, i.e., $d(A) = 0_{n\times n}$.*

Example:

Proof:

Proof (..continued)

Fact: Let $M \in \mathbb{C}^{n \times n}$. For each $\delta > 0$, there exists a matrix $\tilde{M} \in \mathbb{C}^{n \times n}$ with distinct eigenvalues satisfying $\left\| M - \tilde{M} \right\| < \delta$. This is equivalent to stating: "'Matrices with distinct eigenvalues form a dense subset of $\mathbb{C}^{n \times n}$ "'.

Fact: "'*An example*"', Assume $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ has repeated eigenvalues, i.e., $\lambda_1 = \lambda_2$. Then for each $\delta > 0$ one can find numbers $\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4$ satisfying $|\epsilon_i| < \delta$ such that $A_\delta = \begin{bmatrix} a + \epsilon_1 & b + \epsilon_2 \\ c + \epsilon_3 & d + \epsilon_4 \end{bmatrix}$ has distinct eigenvalues, i.e., $\lambda_1^\delta \neq \lambda_2^\delta$

Let us choose a sequence $\{A_n\}_{n=1}^{\infty}$ satisfying $\|A - A_n\| < \frac{1}{n}$ and A_n has distinct eigenvalues. Let us define $d_n(s) := det(sI - A_n)$. Since $det(.)$ is a continuous function one can write $d(A) = \lim_{n \to \infty} d_n(A_n)$.

Note that $d_n(A_n) = 0$ for all n. Hence d(A)=0.

Example:

Example:

6 Minimal Polynomial

Definition *For an $n \times n$ matrix A, the* ***minimal polynomial*** *$m(s)$ is the monic polynomial with smallest degree such that $m(A) = 0_{n\times n}$*

Remark: A monic polynomial has unity as the coefficient of its highest order term.

Theorem *Given $A \in \mathbb{C}^{n\times n}$, let $m(s)$ be its minimal polynomial.*

- *$m(s)$ is unique;*
- *$m(s)$ divides $d(s)$, i.e., there exist a $q(s)$ such that $d(s) = m(s)q(s)$;*
- *Every root of $d(s)$ is also a root of $m(s)$.*

Proof: