Forth Recitation Class
Linear Algebra

YAO Shaoxiong

UM-SJTU Joint Institute

April 3, 2019

Table of contents

Why Norm?

Motivation

In ***infinite*** dimensional linear space, we do not have ***countable*** basis.

Solution

We define ***norm*** as an ***equivalent*** relation. The infinite sum of countable elements is defined considered to be equivalent to a vector,

$$\|u - \sum_{n=1}^{\infty} v_n\| \to 0, \text{ then } u = \sum_{n=1}^{\infty} v_n.$$

Why Norm?

Solution(Contd.)

Even better, this representation satisfies,

$$u^{(1)} = \sum_{n=1}^{\infty} v_n^{(1)}, \quad u^{(2)} = \sum_{n=1}^{\infty} v_n^{(2)},$$

then

$$k_1 u^{(1)} + k_2 u^{(2)} = \sum_{n=1}^{\infty} k_1 v_n^{(1)} + k_2 v_n^{(2)}.$$

Problem

For arbitrary sequence, the summation **may not exist**! We can only allow reasonable summation. Those are ***Cauthy sequences***.

Norm Induced by Inner Product

Inner Product

Inner product is **conjugate symmetry, linear in the first compomenent** and **positive-definite**.

The Cauchy-Schwarz Inequality

For $x, y \in V$,

$$|(x, y)| \leq \sqrt{(x, x)}\sqrt{(y, y)}.$$

Comment:

From Cauchy-Schwarz inequality, the induced norm $\|x\| = \sqrt{(x, x)}$ is actually a norm.

Banach Space and Hilbert Space

Completeness

A normed linear space X is called ***complete*** if any Cauchy sequence converges.

Banach Sapce

A complete normed linear space is called a ***Banach space***.

Hilbert Sapce

A complete inner product linear space is called a ***Hilbert space***.

Properties For Inner Product

Pythagoras's Theorem.

Let $(V, \langle\cdot,\cdot\rangle)$ be an inner product space and M some subset of V. Let $z = x + y$, where $x \in M$ and $y \in M^{\perp}$. Then

$$||z||^2 = ||x||^2 + ||y||^2.$$

Parallelogram Rule

Let V be a real or complex vector space. Then every norm on V, if it is induced by some inner product, then it satisfies

$$||x + y||^2 + ||x - y||^2 = 2(||x||^2 + ||y||^2)$$

for all $x, y \in V$.

Properties For Inner Product

Example

Check for vector space $C([a, b])$, the function defined as

$$(f, g) = \int_a^b f(x)g(x)dx,$$

is actually a inner product.

Solution

(i) Linearity. For $\lambda, \mu \in \mathbb{R}$, $(\lambda f_1 + \mu f_2, g) = \lambda(f_1, g) + \mu(f_2, g)$ from the linearity of integral.

(ii) Positive defined. If $(f, f) = 0$, we will have $f(x) = 0$ for $x \in [a, b]$. Why? f is a continuous function so that if there is $f(x_0) > 0$, there is an interval $(x_0 - \epsilon, x_0 + \epsilon)$ such that $f(x) > 0$, so the integral will larger than $2\epsilon \inf_{x \in (a,b)} |f(x)|^2$.

Orthogonal and Orthonormal Elements

Definition

If for two elements $x, y \in X$, we have $(x, y) = 0$, then they are ***orthogonal***.

Orthonormal Elements

For a tuple of elements $(v_1, ..., v_n) \in X$, if

$$(v_j, v_k) = \begin{cases} 1 & \text{for } j = k \\ 0 & \text{for } j \neq k \end{cases} \qquad j, k = 1, ..., n,$$

they are called ***orthonormal*** elements.

Theorem

Orthonormal elements are linearly independent.

Orthogonal Complements and Direct Sum

Question: If we know $V = \text{span}\{v_1, ..., v_m\}$, how to find $V^\perp$?

Solution: Write $A = [v_1 \ \cdots \ v_m]$, find $\ker A^T$.

Example

For

$$V = \text{span}\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \right\},$$

find $V^\perp$.

Orthogonal Complements and Direct Sum

Question:

If $V \oplus V^{\perp} = X$, how to find $x = x_1 + x_2$ with $x_1 \in V$, $x_2 \in V^{\perp}$?

Solution:

We have an orthonormal basis $\mathcal{B}_1 = (u_1, ..., u_m)$ for V,

$$x_1 = (u_1, x)u_1 + \cdots + (u_m, x)u_m,$$

then $x_2 = x - x_1$.

Note:

(i) Note that $proj_V x = x_1$.

(ii) If $x \in V$, $proj_V x = x$; if $x \in V^{\perp}$, $proj_V x = 0$.

Orthogonal Complements and Direct Sum

Question:

If $V_1 \oplus V_2 = X$,

- $V_1 \cap V_2 = \emptyset$? **No.** $V_1 \cap V_2 = \{0\}$. Why?
- $V_1 \cup V_2 = X$? **No.**
- How to find $f \notin V_1 \cup V_2$?

Relation for Dimension

For $V_1 \oplus V_2 = X$,

$$\dim V_1 + \dim V_2 = \dim X.$$

Orthogonal Matrix

Definition

If the columns a matrix are orthonormal elements, matrix Q is called an orthogonal matrix.

Properties

If a matrix Q is an orthogonal $n \times m$ matrix, we have

$$Q^T Q = I_m.$$

Theorem

For $n \times n$ matrix Q. The matrix Q is orthogonal if and only if $Q^T = Q^{-1}$.

Orthogonal Matrix

Example

Check that

$$Q = \frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$$

is an orthogonal matrix.

Gram-Schimit Algorithm

Quesiton:

Given $V \subset X$, how to find an orthonormal basis? Assume we define this subspace using $(v_1, ..., v_n)$ as a basis.

Solution

Gram: find $(u_1, ..., u_n)$ orthogonal to each other.

$$
\begin{aligned}
u_1 &:= v_1 \\
u_2 &:= v_2 - \frac{(u_1, v_2)}{(u_1, u_1)} u_1 \\
u_3 &:= v_3 - \frac{(u_1, v_3)}{(u_1, u_1)} u_1 - \frac{(u_2, v_3)}{(u_2, u_2)} u_2
\end{aligned}
$$

Note:

Try to write one equation in matrix form!

Gram-Schimit Algorithm

Solution(Contd.)

Schimit: normalize each vector.

$$b_i = \frac{u_i}{\|u_i\|}, \quad i = 1, ..., n.$$

Note:

We can note that this process can be mixed.

$$\begin{aligned} u_1 &= v_1, & b_1 &= \frac{u_1}{\|u_1\|} \\ u_2 &= v_1 - (b_1, v_1)b_1, & b_2 &= \frac{u_2}{\|u_2\|} \end{aligned}$$

Gram-Schimit Algorithm

Note(Contd.)

Please note that the order is not **unique**, we can simplify calculation by rearranging elements.

Example

Find an orthonormal basis from,

$$\begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} -2 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 18 \\ 0 \\ 0 \end{bmatrix}.$$

QR-factorization

Quesiton:

How to write this process in a matrix form?

Solution

This process can be recorded by ***QR***.

For the basis $(v_1, ..., v_n)$, we write this basis as column vectors of a matrix. We assume we have found an orthonormal basis $(b_1, ..., b_n)$.

$$\begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} = \begin{bmatrix} b_1 & \cdots & b_n \end{bmatrix} \begin{bmatrix} (v_1, b_1) & (v_2, b_1) & \cdots & (v_n, b_1) \\ (v_1, b_2) & (v_2, b_2) & \cdots & (v_n, b_2) \\ \vdots & \vdots & \ddots & \vdots \\ (v_1, b_n) & (v_2, b_n) & \cdots & (v_n, b_n) \end{bmatrix}$$

QR-factorization

Solution(Contd.)

Please note that $(b_1, ..., b_n)$ are produced from $(v_1, ..., v_n)$ **in order**, we have $b_j \perp \mathrm{span}\{v_1, ..., v_{j-1}\}$, so that $(v_i, b_j) = 0$ for $i < j$. The matrix is supposed to be **upper triangular**.

$$\begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} = \begin{bmatrix} b_1 & \cdots & b_n \end{bmatrix} \begin{bmatrix} (v_1, b_1) & (v_2, b_1) & \cdots & (v_n, b_1) \\ 0 & (v_2, b_2) & \cdots & (v_n, b_2) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & (v_n, b_n) \end{bmatrix}$$

QR-factorization

Example

Find QR factorization of

$$A = \begin{bmatrix} 4 & 25 & 0 \\ 0 & 0 & -2 \\ 3 & -25 & 0 \end{bmatrix}$$

Solution

$$QR = \begin{bmatrix} 4/5 & 3/5 & 0 \\ 0 & 0 & -1 \\ 3/5 & -4/5 & 0 \end{bmatrix} \begin{bmatrix} 5 & 5 & 0 \\ 0 & 35 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

QR-factorization

Comment:

Please note that the number of elements usually **smaller** than dimension of the vector space.

Question:

For a $n \times m$ matrix A such that $\text{rank}A < m$, will it be possible to write

$$A = QR$$

where Q is an $n \times m$ matrix with orthonormal columns and R is upper triangular?

QR-factorization

Example

Consider a block matrix

$$A = [A_1 \quad A_2]$$

with linearly independent columns.(A_1 is an $n \times m_1$ matrix and A_2 is an $n \times m_2$ matrix.) Suppose we know the QR factorization of A. Explain how to find QR factorization of A_1.

Matrix Transpose

Definition

If A is an $n \times m$ matrix with each entry (a_{ij}) we define the ***transpose*** of A as an $m \times n$ matrix A^T, with each entry (a_{ji}).

Note:

Please note that we have

$$\operatorname{im} A \oplus \ker A^T = \mathbb{R}^n$$
$$\ker A \oplus \operatorname{im} A^T = \mathbb{R}^m.$$

This is the reason to find $V^\perp$ using $\ker A^\perp$.

Least Square Approximation

Question:

For equation $Ax = b$, if $b \notin \operatorname{im} A$, then there is no solution. How to find a best approximation of b in $\operatorname{im} A$?

Solution

We assume Ax^* corresponds to this best approximation. Then

$$Ax^* = proj_{\text{im A}} b,$$

or in other word

$$Ax^* - b \in (\operatorname{im} A)^\perp,$$

equivalently

$$Ax^* - b \in \ker A^T,$$

equivalently

$$A^T(Ax^* - b) = 0.$$

Least Square Approximation

Definition

For $Ax = b$, the ***normal equaiton*** of $Ax = b$ is

$$A^T A x^* = A^T b.$$

Theorem

If A is an $n \times m$ matrix, then

$$\ker(A) = \ker(A^T A).$$

If $\ker(A) = \{\bar{0}\}$, then $A^T A$ is invertible.

Question:

Why this theorem is helpful?

Least Square Approximation

Theorem

If $\ker(A) = \{\bar{0}\}$, then the linear system

$$Ax = b$$

has unique least-squares solution

$$x^* = (A^T A)^{-1} A^T b.$$

The Matrix of an Orthogonal Projection

If we have A contains a basis of V, we can find the orthogonal projection of b in $\mathrm{im}(A)$ is

$$A(A^T A)^{-1} A^T b.$$

Least Square Approximation

Example

Find the least-squares solution x^* of the system

$$Ax = b,$$

where

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$