线性代数 (B1)

童伟华　管理科研楼 1205 室 [1]
E-mail: tongwh@ustc.edu.cn

[1] 数学科学学院　中国科学技术大学

2021-2022 学年第二学期　MATH1009.08

§6 欧几里德空间的定义

线性空间：集合 + 加法与数乘 + 八条运算规律

$\mathbb{R}^3$：有度量（长度、夹角等）的线性空间 $\Rightarrow$ n 维的有度量线性空间，即 n 维欧几里德空间

关键问题：如何将三维空间的度量推广到一般的 n 维线性空间?

$\forall \boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^3$, $\langle \boldsymbol{a}, \boldsymbol{b} \rangle = \| \boldsymbol{a} \| \| \boldsymbol{b} \| \cos\theta$

$\Rightarrow \| \boldsymbol{a} \| = \sqrt{\langle \boldsymbol{a}, \boldsymbol{a} \rangle}$, $\theta = \arccos \frac{\langle \boldsymbol{a}, \boldsymbol{b} \rangle}{\sqrt{\langle \boldsymbol{a}, \boldsymbol{a} \rangle \langle \boldsymbol{b}, \boldsymbol{b} \rangle}}$

容易看出：$\mathbb{R}^3$ 中向量的长度、夹角可以通过内积来定义！

思考：一般线性空间中，如何引入内积？内积需要满足哪些性质？

内积满足哪些本质的性质？$\Rightarrow$ 线性性、对称性及正定性。

定义 6.1

设 V 是实数域 $\mathbf{R}$ 上的线性空间，如果 V 中任意两个向量 $\mathbf{a}$ 和 $\mathbf{b}$ 都按某一法则对应于一个实数，记作 $(\mathbf{a},\mathbf{b})$，且满足

(1) **对称性**：$\forall \mathbf{a},\mathbf{b}\in V$，有 $(\mathbf{a},\mathbf{b})=(\mathbf{b},\mathbf{a})$；

(2) **线性性**：$\forall \lambda\in\mathbb{R},\mathbf{a},\mathbf{b},\mathbf{c}\in V$，有

$$(\lambda\mathbf{a},\mathbf{b})=\lambda(\mathbf{a},\mathbf{b}),$$
$$(\mathbf{a}+\mathbf{b},\mathbf{c})=(\mathbf{a},\mathbf{c})+(\mathbf{b},\mathbf{c});$$

(3) **正定性**：$\forall \mathbf{a}\in V$，有 $(\mathbf{a},\mathbf{a})\geq 0$，等号成立当且仅当 $\mathbf{a}=\mathbf{0}$。

则称 $(\mathbf{a},\mathbf{b})$ 为向量 $\mathbf{a}$ 和 $\mathbf{b}$ 的**内积**，定义了内积的实数域 $\mathbf{R}$ 上的线性空间 V 称为**欧几里得** (Euclid) **空间**，简称**欧氏空间**。

$\Rightarrow \forall\, \lambda_1, \lambda_2 \in \mathbb{R}, \boldsymbol{a}, \boldsymbol{b} \in V$，有

$$(\lambda_1 \mathbf{a} + \lambda_2 \mathbf{b}, \mathbf{c}) = \lambda_1(\mathbf{a}, \mathbf{c}) + \lambda_2(\mathbf{b}, \mathbf{c})$$

$$(\mathbf{a}, \lambda_1 \mathbf{b} + \lambda_2 \mathbf{c}) = \lambda_1(\mathbf{a}, \mathbf{b}) + \lambda_2(\mathbf{a}, \mathbf{c})$$

$$\Rightarrow \Big(\sum_{i=1}^{k} \lambda_i \mathbf{a}_i, \sum_{j=1}^{m} \mu_j \mathbf{b}_j\Big) = \sum_{i=1}^{k}\sum_{j=1}^{m} \lambda_i \mu_j (\mathbf{a}_i, \mathbf{b}_j)$$

$$(\mathbf{a}, \mathbf{0}) = (\mathbf{a}, \lambda \mathbf{0}) = \lambda(\mathbf{a}, \mathbf{0})$$

$$\Rightarrow (\lambda - 1)(\mathbf{a}, \mathbf{0}) = 0, \ \forall \lambda \in \mathbb{R}$$

$$\Rightarrow (\mathbf{a}, \mathbf{0}) = 0$$

§6.1.2 欧几里德空间的性质

$$
\begin{aligned}
&\forall \lambda \in \mathbb{R}, \boldsymbol{a}, \boldsymbol{b} \in V \Rightarrow (\lambda \boldsymbol{a} + \boldsymbol{b}, \lambda \boldsymbol{a} + \boldsymbol{b}) \geqslant 0 \\
&\Rightarrow (\boldsymbol{a}, \boldsymbol{a})\lambda^2 + 2(\boldsymbol{a}, \boldsymbol{b})\lambda + (\boldsymbol{b}, \boldsymbol{b}) \geqslant 0 \\
&\Rightarrow |2(\boldsymbol{a}, \boldsymbol{b})|^2 \leqslant 4(\boldsymbol{a}, \boldsymbol{a})(\boldsymbol{b}, \boldsymbol{b}) \\
&\Rightarrow |(\boldsymbol{a}, \boldsymbol{b})| \leqslant \sqrt{(\boldsymbol{a}, \boldsymbol{a})(\boldsymbol{b}, \boldsymbol{b})}
\end{aligned}
$$

定理 6.1 (Cauchy-Schwarz 不等式)

设 V 是欧氏空间，$(\cdot,\cdot)$ 是 V 的内积，则对 V 中的任意两个向量 $\mathbf{a}$ 和 $\mathbf{b}$，有

$$|(\mathbf{a}, \mathbf{b})| \leqslant \sqrt{(\mathbf{a}, \mathbf{a})(\mathbf{b}, \mathbf{b})}.$$

§6.1.2 欧几里德空间的性质

内积 $\Rightarrow$ 范数

定义 6.2

设 V 是欧氏空间，$(\cdot,\cdot)$ 是 V 的内积，对于任意的 $\mathbf{a}\in V$，称

$$|\mathbf{a}| = \sqrt{(\mathbf{a},\mathbf{a})}$$

称为 $\mathbf{a}$ 的*长度*或*模*。

三角不等式：$\forall\,\mathbf{a},\mathbf{b}\in V$ 有

$$|\mathbf{a}+\mathbf{b}| \leqslant |\mathbf{a}|+|\mathbf{b}|$$

$$\Leftrightarrow |\mathbf{a}+\mathbf{b}|^2 = (\mathbf{a}+\mathbf{b},\mathbf{a}+\mathbf{b}) = (\mathbf{a},\mathbf{a})+2(\mathbf{a},\mathbf{b})+(\mathbf{b},\mathbf{b})$$
$$\leqslant |\mathbf{a}|^2+2|\mathbf{a}||\mathbf{b}|+|\mathbf{b}|^2 = (|\mathbf{a}|+|\mathbf{b}|)^2$$

Cauchy-Schwarz 不等式 $\Rightarrow$ 三角不等式

§6.1.2 欧几里德空间的性质

范数 $\Rightarrow$ 距离

$$d(\mathbf{a},\mathbf{b}) = |\mathbf{a}-\mathbf{b}|,\ \forall \mathbf{a},\mathbf{b} \in V$$

距离满足如下性质:

- **对称性**：$d(\mathbf{a},\mathbf{b}) = d(\mathbf{b},\mathbf{a})$;
- **正定性**：$d(\mathbf{a},\mathbf{b}) \geq 0$，等号成立当且仅当 $\mathbf{a}=\mathbf{b}$;
- **三角不等式**：$d(\mathbf{a},\mathbf{b}) \leq d(\mathbf{a},\mathbf{c}) + d(\mathbf{c},\mathbf{b})$。

推广的线路：内积 $\Rightarrow$ 范数 $\Rightarrow$ 距离，是否存在其它方案？（泛函分析、拓扑学）

如果推广向量之间的夹角？Cauchy-Schwarz 不等式

$$\Rightarrow -1 \leqslant \frac{(\boldsymbol{a},\boldsymbol{b})}{|\boldsymbol{a}||\boldsymbol{b}|} \leqslant 1,\ \forall \boldsymbol{a},\boldsymbol{b} \neq \mathbf{0} \in V$$

定义 6.3

对于欧氏空间 V 中两个非零向量 $\mathbf{a}$ 和 $\mathbf{b}$，定义它们之间的*夹角*为

$$\theta = \arccos \frac{(\mathbf{a},\mathbf{b})}{|\mathbf{a}||\mathbf{b}|}. \tag{1}$$

特别地，当 $(\mathbf{a},\mathbf{b}) = 0$ 时，称向量 $\mathbf{a}$ 和 $\mathbf{b}$ *正交*或*垂直*，记作 $\mathbf{a} \perp \mathbf{b}$。

§6.1.2 欧几里德空间的性质

例 6.1

设 $V=\mathbf{R}^n$，对于任意两个向量

$$\mathbf{a}=(a_1,a_2,\ldots,a_n),\quad \mathbf{b}=(b_1,b_2,\ldots,b_n)$$

定义

$$(\mathbf{a},\mathbf{b})=a_1b_1+a_2b_2+\cdots+a_nb_n,$$

容易验证它是 $\mathbf{R}^n$ 上的一个内积。此时，$\mathbf{R}^n$ 中向量 $\mathbf{a}$ 的模为

$$|\mathbf{a}|=\sqrt{a_1^2+a_2^2+\cdots+a_n^2},$$

两个向量 $\mathbf{a}$ 与 $\mathbf{b}$ 的夹角为

$$\theta=\arccos\frac{a_1b_1+a_2b_2+\cdots+a_nb_n}{\sqrt{a_1^2+a_2^2+\cdots+a_n^2}\sqrt{b_1^2+b_2^2+\cdots+b_n^2}}。$$

对应这种内积，Cauchy-Schwarz 不等式为

$$|a_1b_1+\cdots+a_nb_n|\leqslant\sqrt{a_1^2+\cdots+a_n^2}\sqrt{b_1^2+\cdots+b_n^2}。$$

例 6.2

设 $C[a,b]$ 是闭区间 $[a,b]$ 上实的连续函数的全体。对于任意的 $f,g\in C[a,b]$，定义

$$(f,g)=\int_a^b f(x)g(x)\mathrm{d}x,$$

容易验证，它满足内积定义中的 (1) 和 (2)。如果

$$(f,f)=\int_a^b f^2(x)\mathrm{d}x=0,$$

则由连续函数的性质可知 $f\equiv 0$，所以 (3) 也满足。对应的 Cauchy-Schwarz 不等式为

$$\left|\int_a^b f(x)g(x)\mathrm{d}x\right|\leqslant\sqrt{\int_a^b f^2(x)\mathrm{d}x}\sqrt{\int_a^b g^2(x)\mathrm{d}x}。$$

特别地，对于 $C[-\pi,\pi]$，容易验证三角函数

$$1,\cos x,\sin x,\cos 2x,\sin 2x,\ldots,\cos nx,\sin nx$$

在上述内积之下是两两正交的。

$\Rightarrow$ 无穷维线性空间 $C[a,b]$ 上的傅立叶分析

与描述线性变换的想法类似，如果知道内积在线性空间 V 的一组基上的值，那么内积在任意向量上的取值也就清楚了。

设 V 是有限维的欧氏空间，$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n$ 是它的一组基。任取 V 中任意两个向量 $\boldsymbol{\alpha}$ 和 $\boldsymbol{\beta}$，

$$\boldsymbol{x} = x_1\boldsymbol{\alpha}_1 + x_2\boldsymbol{\alpha}_2 + \cdots + x_n\boldsymbol{\alpha}_n,$$

$$\boldsymbol{y} = y_1\boldsymbol{\alpha}_1 + y_2\boldsymbol{\alpha}_2 + \cdots + y_n\boldsymbol{\alpha}_n,$$

$$\Rightarrow (\boldsymbol{x}, \boldsymbol{y}) = \Big(\sum_{i=1}^{n} x_i\boldsymbol{\alpha}_i, \sum_{j=1}^{n} y_j\boldsymbol{\alpha}_j\Big) = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i y_j(\boldsymbol{\alpha}_i, \boldsymbol{\alpha}_j)$$

$$= \begin{pmatrix} x_1 & x_2 & \cdots & x_n \end{pmatrix} \begin{pmatrix} (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_1) & (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2) & \cdots & (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_n) \\ (\boldsymbol{\alpha}_2, \boldsymbol{\alpha}_1) & (\boldsymbol{\alpha}_2, \boldsymbol{\alpha}_2) & \cdots & (\boldsymbol{\alpha}_2, \boldsymbol{\alpha}_n) \\ \vdots & \vdots & \ddots & \vdots \\ (\boldsymbol{\alpha}_n, \boldsymbol{\alpha}_1) & (\boldsymbol{\alpha}_n, \boldsymbol{\alpha}_2) & \cdots & (\boldsymbol{\alpha}_n, \boldsymbol{\alpha}_n) \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

记 $G=(g_{ij})_{n\times n}$，其中 $g_{ij}=(\boldsymbol{\alpha}_i,\boldsymbol{\alpha}_j)$，$i,j=1,2,\dots n$，称矩阵 G 是内积 $(\cdot,\cdot)$ 在基 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\dots,\boldsymbol{\alpha}_n$ 下的**度量矩阵**。

利用度量矩阵，内积可以表示为

$$(\boldsymbol{x},\boldsymbol{y})=X^{\mathrm{T}}GY,$$

其中 $X=(x_1,\dots,x_n)^{\mathrm{T}}$，$Y=(y_1,\dots,y_n)^{\mathrm{T}}$。

容易看出，度量矩阵是实对称矩阵，并且具有如下性质：

$$\boldsymbol{x}G\boldsymbol{x}\geqslant 0,\quad \text{且等号成立当且仅当 } \boldsymbol{x}=\boldsymbol{0},$$

称满足上述性质的实对称方阵 G 为**正定方阵**。

事实上，若取定 n 维欧几里德空间 V 的一组基，则

$$V \text{ 上的内积 } (\cdot,\cdot) \stackrel{1-1}{\longleftrightarrow} \text{正定的 } n \text{ 阶实对称方阵 } G。$$

线性空间的基是不唯一的，问题:

(1) 欧氏空间的内积在不同基下对应的度量矩阵之间有什么关系?

(2) 是否存在一组特殊的基，使得欧氏空间的内积对应的度量矩阵尽可能的简单?

设欧氏空间 V 的一组基为: $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n$，内积 $(\cdot,\cdot)$ 在此基下的度量矩阵为 G，现有 V 的另一组基: $\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \ldots, \boldsymbol{\eta}_n$，内积 $(\cdot,\cdot)$ 在这组基下的度量矩阵为 $\overline{G}$，则对于 $\boldsymbol{a}, \boldsymbol{b} \in V$，有

$$(\boldsymbol{a}, \boldsymbol{b}) = \begin{pmatrix} a_1 & \cdots & a_n \end{pmatrix} G \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} \bar{a}_1 & \cdots & \bar{a}_n \end{pmatrix} \overline{G} \begin{pmatrix} \bar{b}_1 \\ \vdots \\ \bar{b}_n \end{pmatrix},$$

其中 $\boldsymbol{a} = a_1\boldsymbol{\alpha}_1 + \cdots + a_n\boldsymbol{\alpha}_n = \bar{a}_1\boldsymbol{\eta}_1 + \cdots + \bar{a}_n\boldsymbol{\eta}_n$,

$\boldsymbol{b} = b_1\boldsymbol{\alpha}_1 + \cdots + b_n\boldsymbol{\alpha}_n = \bar{b}_1\boldsymbol{\eta}_1 + \cdots + \bar{b}_n\boldsymbol{\eta}_n$。

记基 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n$ 到基 $\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \ldots, \boldsymbol{\eta}_n$ 的过渡矩阵为 P，即

$$(\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \ldots, \boldsymbol{\eta}_n) = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n)P,$$

$$\Rightarrow \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = P \begin{pmatrix} \overline{a}_1 \\ \vdots \\ \overline{a}_n \end{pmatrix}, \quad \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = P \begin{pmatrix} \overline{b}_1 \\ \vdots \\ \overline{b}_n \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} \overline{a}_1 & \cdots & \overline{a}_n \end{pmatrix} P^{\mathrm{T}} G P \begin{pmatrix} \overline{b}_1 \\ \vdots \\ \overline{b}_n \end{pmatrix} = \begin{pmatrix} \overline{a}_1 & \cdots & \overline{a}_n \end{pmatrix} \overline{G} \begin{pmatrix} \overline{b}_1 \\ \vdots \\ \overline{b}_n \end{pmatrix}$$

$$\Rightarrow P^{\mathrm{T}} G P = \overline{G}$$

定义 6.4

设 $A, B \in F^{n\times n}$，若存在数域 F 上的 n 阶可逆方阵 P，使得 $B = P^{\mathrm{T}}AP$，则称 A 与 B（在数域 F 上）**相合**，记作 $A \sim B$。

容易验证:

(1) $A \sim A$;

(2) $A \sim B \Rightarrow B \sim A$;

(3) $A \sim B, B \sim C \Rightarrow A \sim C$。

$\Rightarrow$ 相合为等价关系!

$\Rightarrow$ 欧氏空间内积在不同基下的度量矩阵彼此相合!

定义 6.5

如果 V 中有一组两两正交的非零向量 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_r$，即

$$(\boldsymbol{\alpha}_i, \boldsymbol{\alpha}_j) = |\boldsymbol{\alpha}_i||\boldsymbol{\alpha}_j|\delta_{ij},\ i,j = 1, 2, \ldots, r,$$

则称该向量组为**正交向量组**。由正交向量组构成的基称为**正交基**。由单位向量组构成的正交基称为**标准正交基**。

规定：单独一个非零向量也构成正交向量组。

命题 6.2

欧氏空间中的正交向量组 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_r\} \Rightarrow \boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_r$ 线性无关。

$\Rightarrow$ 内积在正交基下的度量矩阵为对角阵，在标准正交基下的度量矩阵为单位阵。

n 维欧氏空间的标准正交基是否存在？回答是肯定的。

设欧氏空间 V 的一组基：$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n \Rightarrow$ 构造一组标准正交基?

$$\boldsymbol{e}_1 = \frac{\boldsymbol{\alpha}_1}{|\boldsymbol{\alpha}_1|},$$

$$\boldsymbol{\beta}_2 = \boldsymbol{\alpha}_2 - (\boldsymbol{\alpha}_2, \boldsymbol{e}_1)\boldsymbol{e}_1,\ \boldsymbol{e}_2 = \frac{\boldsymbol{\beta}_2}{|\boldsymbol{\beta}_2|},$$

$$\vdots$$

$$\boldsymbol{\beta}_k = \boldsymbol{\alpha}_k - \sum_{i=1}^{k-1}(\boldsymbol{\alpha}_k, \boldsymbol{e}_i)\boldsymbol{e}_i,\ \boldsymbol{e}_k = \frac{\boldsymbol{\beta}_k}{|\boldsymbol{\beta}_k|},\ k = 1, 2, \ldots, n,$$

$\Rightarrow \boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$ 为一组标准正交基，上述构造性算法称为 Schmidt **正交化方法**

定理 6.3 (Schmidt 正交化)

从 n 维欧氏空间 V 的任意一组基出发，可以构造一组标准正交基。

§6.3 欧几里德空间的同构

n 维线性空间之间是彼此同构的！ $\Rightarrow$ n 维欧几里德空间是否同构?

定义 6.6

实数域 $\mathbb{R}$ 上的两个欧氏空间 V 和 V' 称为**同构的**，如果存在一个从 V 到 V' 的一一映射 $\sigma : V \to V'$，满足

(1) $\sigma(\lambda\boldsymbol{\alpha} + \mu\boldsymbol{\beta}) = \lambda\sigma(\boldsymbol{\alpha}) + \mu\sigma(\boldsymbol{\beta})$;

(2) $(\sigma(\boldsymbol{\alpha}), \sigma(\boldsymbol{\beta})) = (\boldsymbol{\alpha}, \boldsymbol{\beta})$,

其中 $\forall \boldsymbol{\alpha}, \boldsymbol{\beta} \in V$, $\lambda, \mu \in \mathbb{R}$。

容易验证欧氏空间之间的同构关系是等价关系！

§6.3 欧几里德空间的同构

设 V 是 n 维欧氏空间，取 V 的一组标准正交基：$\boldsymbol{e}_1, \boldsymbol{e}_2, \dots, \boldsymbol{e}_n$，则

$$\forall \boldsymbol{\alpha} \in V \Rightarrow \boldsymbol{\alpha} = x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2 + \cdots + x_n\boldsymbol{e}_n$$

令 $\sigma(\boldsymbol{\alpha}) = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n \Rightarrow$ 映射 $\sigma : V \to \mathbb{R}^n$ 是同构映射!

定理 6.4

两个有限维欧氏空间同构的充分必要条件是它们的维数相同。

欧氏空间的结构完全被它的维数所决定!

§6.4.1 正交变换与正交矩阵

研究欧氏空间中满足特殊几何性质的线性变换

正交变换：保持内积不变的变换

定义 6.7

设 V 是一个 n 维的欧氏空间，$\mathcal{A}$ 是 V 上的一个线性变换，如果 $\mathcal{A}$ 保持 V 的内积不变，即 $\forall \boldsymbol{a}, \boldsymbol{b} \in V$ 都有

$$(\mathcal{A}\boldsymbol{a}, \mathcal{A}\boldsymbol{b}) = (\boldsymbol{a}, \boldsymbol{b}),$$

则称 $\mathcal{A}$ 是 V 上的**正交变换**。

§6.4.1 正交变换与正交矩阵

定理 6.5

设 V 是一个 n 维的欧氏空间，$\mathcal{A}$ 是 V 上的一个线性变换，则 $\mathcal{A}$ 为正交变换当且仅当下列两个条件之一成立

(1) $\mathcal{A}$ 保持任意向量的模长不变；

(2) $\mathcal{A}$ 将标准正交基变换为标准正交基。

思考：正交变换的矩阵表示满足什么代数条件?

§6.4.1 正交变换与正交矩阵

欧氏空间 V 有一组标准正交基 $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$， V 上线性变换 $\mathcal{A}$ 在这组基下的表示矩阵为 A，即

$$(\mathcal{A}\boldsymbol{e}_1, \mathcal{A}\boldsymbol{e}_2, \ldots, \mathcal{A}\boldsymbol{e}_n) = (\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n)A$$

$$\Rightarrow \mathcal{A}\boldsymbol{e}_i = \sum_{l=1}^{n} a_{li}\boldsymbol{e}_l,\ \mathcal{A}\boldsymbol{e}_j = \sum_{k=1}^{n} a_{kj}\boldsymbol{e}_k$$

$$\Rightarrow (\mathcal{A}\boldsymbol{e}_i, \mathcal{A}\boldsymbol{e}_j) = (\sum_{l=1}^{n} a_{li}\boldsymbol{e}_l, \sum_{k=1}^{n} a_{kj}\boldsymbol{e}_k) = \sum_{l=1}^{n}\sum_{k=1}^{n} a_{li}\, a_{kj}(\boldsymbol{e}_l, \boldsymbol{e}_k)$$

$$= \sum_{k=1}^{n} a_{ki}\, a_{kj} = \delta_{ij},\ \forall i, j = 1, 2, \ldots, n \ \Leftrightarrow\ A^{\mathrm{T}}A = I$$

定义 6.8

如果实方阵 A 满足

$$A^{\mathrm{T}}A = I \quad 或 \quad A^{-1} = A^{T},$$

则称方阵 A 为**正交矩阵**。

§6.4.1 正交变换与正交矩阵

$$A^{\mathrm{T}}A = I \Rightarrow \det(A) = \pm 1,\ A^{-1} = A^{\mathrm{T}}$$

n 阶实方阵 A 为正交方阵 $\Leftrightarrow$ A 的行向量组（列向量组）构成 $\mathbb{R}^{n\times n}$ 的标准正交基

定理 6.6

欧氏空间中的线性变换 $\mathcal{A}$ 是正交变换的充分必要条件是 $\mathcal{A}$ 在标准正交基下的矩阵 A 是正交矩阵。

§6.4.1 正交变换与正交矩阵

命题 6.7

设 V 是 n 维欧氏空间，则

(1) 单位变换是正交变换；

(2) 两个正交变换的复合仍然是正交变换；

(3) 正交变换一定可逆，其逆变换也是正交变换。

记 $O(V)$ 为 n 维欧氏空间 V 上正交变换全体所构成的集合
$\Rightarrow$ 正交变换群

记 $GL(V)$ 为 n 维欧氏空间 V 上可逆变换全体所构成的集合
$\Rightarrow$ 一般线性群

$\mathcal{A}$ 为正交变换 $\Rightarrow \det(\mathcal{A}) = \pm 1$

定义 6.9

如果正交变换 $\mathcal{A}$ 在一组基下的矩阵行列式为 1，则称 $\mathcal{A}$ 为**第一类变换**。如果正交变换 $\mathcal{A}$ 在一组基下的矩阵行列式为 -1，则称 $\mathcal{A}$ 为**第二类变换**。

在 $\mathbb{R}^2$ 中，旋转变换是第一类正交变换，第二类正交变换的实例是反射变换。

§6.4.1 正交变换与正交矩阵

定理 6.8

设 $\mathcal{A}$ 是欧氏空间 V 上的正交变换，则

(1) *$\mathcal{A}$ 的特征值的模长都为 1。特别地，$\mathcal{A}$ 的实特征值（如果存在的话）只能是 1 或者 -1；*

(2) *如果 V 的维数是奇数且 $\mathcal{A}$ 是第一类正交变换，则 $\mathcal{A}$ 一定存在值为 1 的特征值。*

几何解释：三维欧氏空间 $\mathbb{R}^3$ 中任何一个旋转变换（第一类变换）一定保持一个向量不变，即旋转轴。

§6.4.1 正交变换与正交矩阵

例 6.3

在二维平面上任取一点 O 作为原点，将平面上每个点 P 与向量 $\overrightarrow{OP}$ 对应起来，将平面看成二维欧氏空间 V。设 $\mathcal{A}$ 是 V 上的正交变换，则

(1) 当 $\det(\mathcal{A})=1$ 时，$\mathcal{A}$ 是绕原点的旋转，在 V 的任意一组标准正交基下的表示矩阵 $A=\begin{pmatrix}\cos\alpha & -\sin\alpha\\ \sin\alpha & \cos\alpha\end{pmatrix}$；

(2) 当 $\det(\mathcal{A})=-1$ 时，$\mathcal{A}$ 是关于过原点的某条直线 l 的轴对称，在 V 的某一组标准正交基下的表示矩阵 $A=\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}$。

例 6.4

设 $\mathbb{R}^3$ 是建立了直角坐标系的欧式空间，$\mathcal{A}$ 是 $\mathbb{R}^3$ 上的正交变换且 $\det(\mathcal{A}) = 1$。求证：$\mathcal{A}$ 是绕某条过原点的直线的旋转。

§6.4.2 对称变换和对称矩阵

定义 6.10

设 V 是 n 维欧氏空间，$\mathcal{A}$ 是 V 上的线性变换。如果 $\mathcal{A}$ 满足

$$(\mathbf{a}, \mathcal{A}\mathbf{b}) = (\mathcal{A}\mathbf{a}, \mathbf{b}), \ \forall \boldsymbol{a}, \boldsymbol{b} \in V,$$

则称 $\mathcal{A}$ 是 V 上的对称变换。

定理 6.9

设 $\mathcal{A}$ 是欧氏空间 V 上的线性变换，则 $\mathcal{A}$ 是对称变换 $\Leftrightarrow$ $\mathcal{A}$ 在 V 的任何一组标准正交基下的表示矩阵 A 是实对称方阵。

§6.4.2 对称变换和对称矩阵

定理 6.10

设 $\mathcal{A}$ 是欧氏空间 V 上的对称变换，则 $\mathcal{A}$ 的不同特征值对应的特征向量相互正交。

推论 6.1

实对称阵 A 的属于不同特征值的特征向量必正交。

§6.4.3 实对称矩阵的对角化

命题 6.11

实对称矩阵 A 的特征值都是实数。

定理 6.12

对于任意 n 阶实对称阵 A，存在一个 n 阶正交矩阵 T，使得 $T^{-1}AT$ 为对角矩阵。

例 6.5

设 $A=\begin{pmatrix}1&2&2\\2&1&2\\2&2&1\end{pmatrix}$，求正交阵 T，使 $T^{-1}AT$ 为对角阵。

§6.4.3 实对称矩阵的对角化

例 6.6

设 3 阶实对称矩阵 A 的各行元素之和均为 3，向量 $\boldsymbol{\alpha}_1=(-1,2,-1)^{\mathrm{T}}$，$\boldsymbol{\alpha}_2=(0,-1,1)^{\mathrm{T}}$ 是线性方程组 $A\boldsymbol{x}=\boldsymbol{0}$ 的两个解。

(1) 求 A 的特征值与特征向量；
(2) 求正交矩阵 Q 和对角阵 D，使得 $Q^{\mathrm{T}}AQ=D$；
(3) 求 A 及 $(A-\frac{3}{2}I)^6$，其中 I 为 3 阶单位矩阵。

§6.4.3 实对称矩阵的对角化

例 6.7

设 $A=\begin{pmatrix} 0 & -1 & 4 \\ -1 & 3 & a \\ 4 & a & 0 \end{pmatrix}$，存在正交矩阵 Q 使得 $Q^{\mathrm{T}}AQ$ 为对角阵。若 Q 的第一列为 $\frac{1}{\sqrt{6}}(1,2,1)^{\mathrm{T}}$，求 a，Q。

例 6.8

设 A 为 3 阶实对称矩阵，$\mathrm{rank}(A)=2$ 且

$$A\begin{pmatrix}1 & 1\\ 0 & 0\\ -1 & 1\end{pmatrix}=\begin{pmatrix}-1 & 1\\ 0 & 0\\ 1 & 1\end{pmatrix}。$$

(1) 求矩阵 A 的特征值与特征向量；

(2) 求矩阵 A。

§6.5.1 酉空间的基本概念

欧氏空间 $\mathbb{R}^n$ （$F=\mathbb{R}$） $\Rightarrow$ 酉空间 $\mathbb{C}^n$ （$F=\mathbb{C}$），内积如何推广?

定义 6.11

设 V 是复数域 $\mathbb{C}$ 上的线性空间，如果 V 内任意两个向量 $\boldsymbol{a}$ 和 $\boldsymbol{b}$ 都按某一法则对应于一个复数，记作 $(\boldsymbol{a},\boldsymbol{b})$，且满足

(1) **共轭对称性**：$\forall \boldsymbol{a},\boldsymbol{b}\in V$，有 $(\boldsymbol{a},\boldsymbol{b})=\overline{(\boldsymbol{b},\boldsymbol{a})}$;

(2) **线性性**：$\forall \lambda\in\mathbb{C},\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}\in V$，有

$$(\boldsymbol{a},\lambda\boldsymbol{b})=\lambda(\boldsymbol{a},\boldsymbol{b}),$$

$$(\boldsymbol{a},\boldsymbol{b}+\boldsymbol{c})=(\boldsymbol{a},\boldsymbol{b})+(\boldsymbol{a},\boldsymbol{c});$$

(3) **正定性**：$\forall \boldsymbol{a}\in V$，有 $(\boldsymbol{a},\boldsymbol{a})\geq 0$，等号成立当且仅当 $\boldsymbol{a}=\boldsymbol{0}$，

则称 $(\boldsymbol{a},\boldsymbol{b})$ 为 $\boldsymbol{a}$ 和 $\boldsymbol{b}$ 的**内积**，定义了内积的复数域 $\mathbf{C}$ 上的线性空间 V 称为**酉空间**。

§6.5.1 酉空间的基本概念

容易看出：$(\lambda \boldsymbol{a}, \boldsymbol{b}) = \overline{\lambda}(\boldsymbol{a}, \boldsymbol{b})$，$(\boldsymbol{a}+\boldsymbol{b}, \boldsymbol{c}) = (\boldsymbol{a}, \boldsymbol{c}) + (\boldsymbol{b}, \boldsymbol{c})$

$$\Rightarrow \left(\sum_{i=1}^{m} \lambda_i \boldsymbol{a}_i, \sum_{j=1}^{n} \mu_j \boldsymbol{b}_j\right) = \sum_{i=1}^{m}\sum_{j=1}^{n} \overline{\lambda}_i \mu_j (\boldsymbol{a}_i, \boldsymbol{b}_j)$$

V 是有限维酉空间：$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n$ 是 V 的一组基

$$\boldsymbol{x} = x_1\boldsymbol{\alpha}_1 + x_2\boldsymbol{\alpha}_2 + \cdots + x_n\boldsymbol{\alpha}_n,\ \boldsymbol{y} = y_1\boldsymbol{\alpha}_1 + y_2\boldsymbol{\alpha}_2 + \cdots + y_n\boldsymbol{\alpha}_n,$$

$$\Rightarrow (\boldsymbol{x}, \boldsymbol{y}) = \begin{pmatrix} \overline{x}_1 & \cdots & \overline{x}_n \end{pmatrix} \begin{pmatrix} (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_1) & \cdots & (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_n) \\ \vdots & \vdots & \vdots \\ (\boldsymbol{\alpha}_n, \boldsymbol{\alpha}_1) & \cdots & (\boldsymbol{\alpha}_n, \boldsymbol{\alpha}_n) \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

在 $\mathbb{C}^n$ 中：$\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$ 标准正交基

$$\boldsymbol{a} = (a_1, a_2, \ldots, a_n)^{\mathrm{T}},\ \boldsymbol{b} = (b_1, b_2, \ldots, b_n)^{\mathrm{T}}$$

$$(\boldsymbol{a}, \boldsymbol{b}) \triangleq \overline{a}_1 b_1 + \overline{a}_2 b_2 + \cdots + \overline{a}_n b_n = \boldsymbol{a}^H \boldsymbol{b}$$

$\Rightarrow \mathbb{C}^n$ 中的**标准内积**

§6.5.2 酉空间的基本性质

内积 $\Rightarrow$ 模或范数 $\Rightarrow$ 距离

$$\| \boldsymbol{a} \| = \sqrt{(\boldsymbol{a}, \boldsymbol{a})}$$

夹角是否成立？答案是否定的，酉空间内的向量之间没有夹角的定义！

Cauchy-Schwarz 不等式：$|(\boldsymbol{a}, \boldsymbol{b})| \leqslant |\boldsymbol{a}||\boldsymbol{b}| \;\not\Rightarrow\; -1 \leqslant \frac{(\boldsymbol{a},\boldsymbol{b})}{|\boldsymbol{a}||\boldsymbol{b}|} \leqslant 1$

酉向量的向量之间是否有正交或垂直的概念？

定义 6.12

酉空间中两个向量 $\boldsymbol{a}$ 和 $\boldsymbol{b}$ 满足 $(\boldsymbol{a}, \boldsymbol{b}) = 0$，则称 $\boldsymbol{a}$ 与 $\boldsymbol{b}$ 相互**正交**或**垂直**，记为 $\boldsymbol{a} \perp \boldsymbol{b}$。

§6.5.2 酉空间的基本性质

定理 6.13

设 V 是 n 维酉空间，则

(1) V 中两两正交的一组非零向量一定是线性无关的；

(2) V 中也存在标准正交基，即存在一组基 $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$，满足

$$(\boldsymbol{e}_i, \boldsymbol{e}_j) = \delta_{ij}, \quad i, j = 1, 2, \ldots, n;$$

(3) 欧氏空间的 Schmidt 正交化过程在酉空间一样有效，即从酉空间中任何一组基出发，可以通过完全一样的 Schmidt 正交化过程，得到一组标准正交基。

在酉空间中，正交变换 $\Rightarrow$ 酉变换，对称变换 $\Rightarrow$ Hermite 变换

§6.5.3 酉变换与酉矩阵

欧氏空间中的正交变换 $\Rightarrow$ 酉空间中的酉变换

定义 6.13

设 $\mathcal{U}$ 是酉空间 V 上的线性变换，如果 $\mathcal{U}$ 保持内积不变，即

$$(\mathcal{U}\boldsymbol{a}, \mathcal{U}\boldsymbol{b}) = (\boldsymbol{a}, \boldsymbol{b}),\ \forall \boldsymbol{a}, \boldsymbol{b} \in V$$

则称 $\mathcal{U}$ 是酉空间 V 上的一个**酉变换**。

酉变换 $\mathcal{U}$ 在酉空间 V 的任一标准正交基下的表示矩阵 U 满足：$U^{\mathrm{H}}U = I$，其中 H 表示共轭转置，即 $U^{\mathrm{H}} \triangleq \overline{(U^{\mathrm{T}})} = (\overline{A})^{\mathrm{T}}$。

定义 6.14

设 U 是一个 n 阶复矩阵，如果它满足

$$U^{\mathrm{H}}U = I \quad 或 \quad U^{-1} = U^{\mathrm{H}}$$

则称 U 为**酉矩阵**。

§6.5.3 酉变换与酉矩阵

定理 6.14

设 $\mathcal{U}$ 是酉空间 V 内的一个线性变换，则下列各命题相互等价

(1) $\mathcal{U}$ 是一个酉变换；

(2) $\mathcal{U}$ 保持向量的模长不变，即 $\forall \boldsymbol{a} \in V$，有 $|\mathcal{U}\boldsymbol{a}| = |\boldsymbol{a}|$；

(3) $\mathcal{U}$ 把酉空间的标准正交基变为标准正交基；

(4) $\mathcal{U}$ 在标准正交基下的表示矩阵是酉矩阵。

$U(V)$: n 维酉空间 V 中所有酉变换的全体所形成的集合 $\Rightarrow$ $U(V)$ 在复合和逆运算之下是封闭的，构成酉群！

$$U^{\mathrm{H}}U = I \Rightarrow \det(U)\overline{\det(U)} = 1 \Rightarrow |\det(U)| = 1$$

$$\Rightarrow \det(U) = e^{i\theta},\ \theta \in \mathbb{R}$$

§6.5.4 Hermite 变换与 Hermite 矩阵

欧氏空间中的对称变换 $\Rightarrow$ 酉空间中的 Hermite 变换

定义 6.15

设 $\mathcal{A}$ 是 n 维酉空间 V 上的线性变换，如果

$$(\boldsymbol{a}, \mathcal{A}\boldsymbol{b}) = (\mathcal{A}\boldsymbol{a}, \boldsymbol{b}),\ \forall \boldsymbol{a}, \boldsymbol{b} \in V$$

则称 $\mathcal{A}$ 是酉空间 V 上的 Hermite *变换*。

Hermite 变换 $\mathcal{A}$ 在酉空间 V 的任一标准正交基下的表示矩阵 A 满足：$A^{\mathrm{H}} = A$。

§6.5.5 规范变换与规范矩阵

代数运算：$A \to A^{\mathrm{T}}(F=\mathbb{R})$，$A \to A^{\mathrm{H}}(F=\mathbb{C}) \Rightarrow$ 几何意义是什么?

定义 6.16

设 $\mathcal{A}$ 是 n 维酉空间 V 上的一个线性变换，如果存在 V 上的线性变换 $\mathcal{A}^*$，使得

$$(\mathcal{A}\boldsymbol{a}, \boldsymbol{b}) = (\boldsymbol{a}, \mathcal{A}^*\boldsymbol{b}),\ \forall \boldsymbol{a}, \boldsymbol{b} \in V,$$

则称 $\mathcal{A}^*$ 是 $\mathcal{A}$ 的*伴随变换*。

思考：线性变换 $\mathcal{A}^*$ 是否存在？是否唯一？（答案是肯定的!）

$\mathcal{A}$ 在酉空间 V 的任一标准正交基下的表示矩阵分别为 A

$\Rightarrow$ 伴随变换 $\mathcal{A}^*$ 在这组基下的表示矩阵为：A^{H}

§6.5.5 规范变换与规范矩阵

定义 6.17

设 $\mathcal{A}$ 是 n 维酉空间 V 上的一个线性变换，如果 $\mathcal{A}$ 和它的伴随变换 $\mathcal{A}^*$ 可交换，即

$$\mathcal{A}\mathcal{A}^* = \mathcal{A}^*\mathcal{A},$$

则称变换 $\mathcal{A}$ 是一个*规范变换*。规范变换在 V 的任一标准正交基下的矩阵满足

$$AA^{\mathrm{H}} = A^{\mathrm{H}}A,$$

我们称任何满足上式的矩阵为*规范矩阵*。

容易验证：酉变换和 Hermite 变换均为规范变换

$\Rightarrow$ 酉矩阵和 Hermite 矩阵均为规范矩阵

§6.5.5 规范变换与规范矩阵

定理 6.15

设 $\mathcal{A}$ 是 n 维酉空间 V 中的一个规范变换，则

(1) 若 λ 是 $\mathcal{A}$ 的一个特征值, x 是 $\mathcal{A}$ 的属于特征值 λ 的特征向量 $\Rightarrow \overline{\lambda}$ 是伴随变换 $\mathcal{A}^*$ 的特征值，x 是 $\mathcal{A}^*$ 的属于特征值 $\overline{\lambda}$ 的特征向量；

(2) $\mathcal{A}$ 属于不同特征值的特征向量相互正交。

定义 6.18

对于复数域上的 n 阶方阵 A 和 B，如果存在 n 阶酉矩阵 U 使得

$$A = U^{-1}BU,\ U^{-1} = U^H$$

则称 A 酉相似于B。

问题：规范矩阵相似等价类的代表元是什么？

§6.5.5 规范变换与规范矩阵

定理 6.16

设 $\mathcal{A}$ 是 n 维酉空间 V 中任意规范变换，则存在 V 的一组标准正交基，使得 $\mathcal{A}$ 在这组基下的矩阵 A 为对角阵。

$\Leftrightarrow$ 任一规范矩阵 A 都可酉相似于对角阵，即存在酉矩阵 U 使得

$$U^{-1}AU = \mathrm{diag}\{\lambda_1, \lambda_2, \ldots, \lambda_n\},$$

其中 λ_i $(i = 1, 2, \ldots, n)$ 为 A 的特征值。反之，复数域 $\mathbb{C}$ 上任一酉相似于对角阵的矩阵一定是规范矩阵。

定理 6.17

A 为规范方阵 $\Leftrightarrow$ A^{H} 可以表示为 A 的多项式。

定理 6.18

设 $\mathcal{A}$ 是酉空间 V 中任意一个酉变换，则

(1) $\mathcal{A}$ 的特征值一定是模为 1 的复数；

(2) 存在一组标准正交基，使得 $\mathcal{A}$ 在这组基下对应的矩阵为对角阵，即存在一个酉矩阵 U 使得

$$A = U\mathrm{diag}(e^{i\theta_1}, e^{i\theta_2}, \ldots, e^{i\theta_n})U^{\mathrm{H}},\ \theta_j \in \mathbf{R},\ j = 1, 2, \ldots, n。$$

定理 6.19

设 $\mathcal{A}$ 是酉空间 V 中的任意一个 Hermite 变换，则

(1) $\mathcal{A}$ 的特征值一定是实数；

(2) 存在一组标准正交基，使得 $\mathcal{A}$ 在这组基下对应的矩阵为实对角阵，即存在一个酉矩阵 U 使得

$$A = U\operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)U^{\mathrm{H}},$$

其中 $\lambda_i \in \mathbb{R}$ $(i = 1, 2, \ldots, n)$ 是 A 的特征值。

推论 6.2

设 $\mathcal{A}$ 是欧氏空间 V 中的对称变换，则存在 V 的一组标准正交基，使得 $\mathcal{A}$ 在这组基下的表示矩阵为对角阵，即实对称矩阵可正交相似于对角阵。

函数逼近问题

对给定的一组观测数据

$$(x_i, f(x_i)), \quad i = 1, 2, \ldots, m,$$

在函数空间 Φ 中寻找满足函数 φ 使得

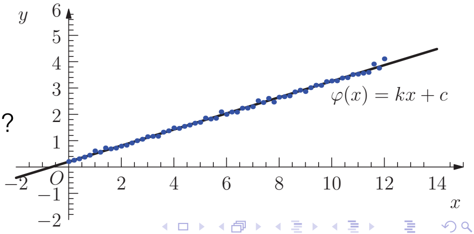
$$\min_{\varphi \in \Phi} \| \varphi - f \|,$$

其中 $\| \cdot \|$ 表示函数空间 Φ 的某种范数。

问题:

(1) 范数 $\| \cdot \|$ 如何选取?

(2) 函数空间 Φ 如何选取?



在泛函分析中，最常用的一类赋范线性空间是 $L^p([a,b])$，即区间 $[a,b]$ 上所有 p 方可积函数的全体。对任意的 $f(x) \in L^p([a,b])$，函数的 p-范数定义如下：

$$||f||_p \triangleq \left(\int_a^b |f(t)|^p dt \right)^{\frac{1}{p}}.$$

问题：对函数逼近问题来说，函数 $f(x)$ 往往是未知的，只有 $f(x)$ 在某些观测点的值是已知的。

$\Rightarrow$ 使用 $L^p([a,b])$ 的离散逼近来刻画 $\varphi(x)$ 与 $f(x)$ 之间的误差或距离，即

$$||\varphi - f||_p \triangleq \left(\sum_{i=1}^{m} |\varphi(x_i) - f(x_i)|^p \right)^{\frac{1}{p}}.$$

在实际计算中，p 常取 1, 2 或 $+\infty$。当 $p=2$ 时，极小化 2-范数的优化问题称为**最小二乘问题**。

函数空间 Φ 一般可依据专业知识或工作经验来选取，例如 Φ 取多项式空间 $\mathbb{P}_n[x]$、样条函数空间 $\mathcal{S}(\mathbb{P}_m[x])$ 或函数空间 $\left\{ae^{bx} \,|\, a, b \in \mathbb{R}\right\}$ 等。

一般地，假设 Φ 为线性空间，函数组 $\{\varphi_j(x)\}_{j=1}^n$ 构成它的一组基，则最小二乘问题可写成

$$\min_{c_1, c_2, \cdots, c_n \in \mathbb{R}} \sum_{i=1}^{m} \left(\sum_{j=1}^{n} c_j \varphi_j(x_i) - f(x_i) \right)^2 .$$

若记 $F(c_1, c_2, \cdots, c_n) = \displaystyle\sum_{i=1}^{m} \left(\sum_{j=1}^{n} c_j \varphi_j(x_i) - f(x_i) \right)^2$，则 $F(c_1, c_2, \cdots, c_m)$ 可视为关于变量 c_1，c_2，$\cdots$，c_m 的多元二次函数。

若记

$$A=\begin{pmatrix}\varphi_1(x_1) & \varphi_2(x_1) & \cdots & \varphi_n(x_1)\\ \varphi_1(x_2) & \varphi_2(x_2) & \cdots & \varphi_n(x_2)\\ \vdots & \vdots & \ddots & \vdots\\ \varphi_1(x_m) & \varphi_2(x_m) & \cdots & \varphi_n(x_m)\end{pmatrix},\ \mathbf{c}=\begin{pmatrix}c_1\\ c_2\\ \cdots\\ c_n\end{pmatrix},\ \mathbf{f}=\begin{pmatrix}f(x_1)\\ f(x_2)\\ \cdots\\ f(x_m)\end{pmatrix},$$

则多元二次函数 $F(c_1,c_2,\cdots,c_n)$ 可表示为

$$\begin{aligned}F(c_1,c_2,\cdots,c_n) &=\| A\mathbf{c}-\mathbf{f}\|_2^2=(A\mathbf{c}-\mathbf{f})^T(A\mathbf{c}-\mathbf{f})\\ &=\mathbf{c}^TA^TA\mathbf{c}-2\mathbf{c}^TA^T\mathbf{f}+\mathbf{f}^T\mathbf{f}.\end{aligned}$$

§6.6 最小二乘法

多元函数 $F(c_1, c_2, \cdots, c_n)$ 取极值的必要条件：

$$\frac{\partial F}{\partial \mathbf{c}} = \begin{pmatrix} \frac{\partial F}{\partial c_1} \\ \frac{\partial F}{\partial c_2} \\ \vdots \\ \frac{\partial F}{\partial c_n} \end{pmatrix} = 2A^TA\mathbf{c} - 2A^T\mathbf{f} = \mathbf{0} \ \Rightarrow \ A^TA\mathbf{c} = A^T\mathbf{f}.$$

事实上，我们可以证明 $A^TA\mathbf{c} = A^T\mathbf{f}$ 也是最小二乘问题解的充分条件。

§6.6 最小二乘法

定义 6.19

设 $A \in \mathbb{R}^{m\times n}$, $\mathbf{f} \in \mathbb{R}^m$，称向量 $\boldsymbol{c}$ 为最小二乘问题 $\min\limits_{\boldsymbol{x}\in\mathbb{R}^n} \| A\boldsymbol{x} - \boldsymbol{f} \|_2^2$ 的解，如果

$$\| A\boldsymbol{c} - \boldsymbol{f} \|_2^2 \leqslant \| A\boldsymbol{x} - \boldsymbol{f} \|_2^2,\ \forall \mathbf{x} \in \mathbb{R}^n$$

成立。

定义 6.20

当 $m \gg n$ 时，称方程组 $A\mathbf{x} = \mathbf{f}$ 为*矛盾方程组*，此时方程个数远远多余未知量的个数方程，是*超定的*；方程组 $A^TA\mathbf{x} = A^T\mathbf{f}$ 称为最小二乘问题的*法方程组*。

思考：如何用欧氏空间的理论来解释最小二乘法?

$A\boldsymbol{x} = (\boldsymbol{a}_1\boldsymbol{a}_2\dots\boldsymbol{a}_n)\boldsymbol{x} = x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + \cdots + x_n\boldsymbol{a}_n$

记 $\mathrm{col}(A) = \mathrm{span}\{\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_n\}$，则最小二乘问题等价于在 $\mathrm{col}(A) \subset \mathbb{R}^m$ 中寻找向量 $\widehat{\boldsymbol{f}}$，使得 $\widehat{\boldsymbol{f}}$ 与 $\boldsymbol{f}$ 距离最近；记 $A\boldsymbol{c} = \widehat{\boldsymbol{f}}$。

$\Rightarrow (\boldsymbol{f} - \widehat{\boldsymbol{f}}) \perp \mathrm{col}(A)$

$\Rightarrow (\boldsymbol{f} - A\boldsymbol{c}) \perp \mathrm{col}(A)$

$\Rightarrow \boldsymbol{a}_i^{\mathrm{T}}(\boldsymbol{f} - A\boldsymbol{c}) = 0,\ i = 1, 2, \dots, n$

$\Rightarrow A^{\mathrm{T}}(\boldsymbol{f} - A\boldsymbol{c}) = \boldsymbol{0}$

$\Rightarrow A^{\mathrm{T}}A\boldsymbol{c} = A^{\mathrm{T}}\boldsymbol{f}$

§6.6 最小二乘法

定理 6.20

设 $A\in\mathbb{R}^{m\times n}$，$\boldsymbol{b}\in\mathbb{R}^m$，则

(1) $A\boldsymbol{x}=\boldsymbol{b}$ 的法方程组 $A^{\mathrm{T}}A\boldsymbol{x}=A^{\mathrm{T}}\boldsymbol{b}$ 恒有解；

(2) $\min\limits_{\boldsymbol{x}\in\mathbb{R}^n}\|A\boldsymbol{x}-\boldsymbol{b}\|_2^2$ 的解 $\Leftrightarrow$ $\boldsymbol{x}\in\mathbb{R}^n$ 是法方程组 $A^{\mathrm{T}}A\boldsymbol{x}=A^{\mathrm{T}}\boldsymbol{b}$ 的解。

定理 6.21

矩阵 $A^{\mathrm{T}}A$ 可逆 $\Leftrightarrow$ A 的列向量组是线性无关的。在此情形下，矛盾方程组 $A\boldsymbol{x}=\boldsymbol{b}$ 有唯一最小二乘解 $\boldsymbol{c}$，且有

$$\boldsymbol{c}=(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}\boldsymbol{b}.$$

历史：最小二乘法是由德国数学家高斯和法国数学家勒让德分别独立提出的，如今被广泛应用于统计学、逼近论和控制论等领域。

Thanks for your attention!