线性代数笔记

Kevin Y.

School of Physical Science,
University of Science and Technology of China
yuhongfei@mail.ustc.edu.cn

Friday 6th September, 2024

School of Physical Sciences, University of Science and Technology of China

2023 级 02 系 6 班，Kevin Yu ，yuhongfei@mail.ustc.edu.cn

本资料为个人整理笔记，仅供复习参考，不适用于直接学习。任何错误或疏漏，请指正。

First release, June 2024

Contents

第一章　矩阵

1.1　矩阵的定义与基本概念

1.2 CR 分解、矩阵的相抵与秩

1.2.1 矩阵的 CR 分解

矩阵 $\mathbf{A} \in \mathbb{F}^{m\times n}$，则存在矩阵 $\mathbf{C} \in \mathbb{F}^{m\times r}, \mathbf{R} \in \mathbb{F}^{r\times n}$ $(r = \operatorname{rank}\mathbf{A})$，使得

$$\mathbf{A} = \mathbf{C}\mathbf{R}$$

称为矩阵的 **CR** 分解，$\mathbf{C}, \mathbf{R}$ 分别为 $\mathbf{A}$ 的线性无关列组成的矩阵和行阶梯型矩阵。

CR 分解表示矩阵 $\mathbf{A}$ 的秩、行秩、列秩均相等。

显然存在可逆矩阵 $\mathbf{P} \in \mathbb{F}^{m\times m}, \mathbf{Q} \in \mathbb{F}^{n\times n}$，使得 $\mathbf{C} = \mathbf{P}\begin{pmatrix} \mathbf{I}_r & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{pmatrix}, \mathbf{R} = \begin{pmatrix} \mathbf{I}_r & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{pmatrix}\mathbf{Q}$。因此矩阵 $\mathbf{A}$ 可写成

$$\mathbf{A} = \mathbf{P}\begin{pmatrix} \mathbf{I}_r & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{pmatrix}\mathbf{Q}$$

1.2.2 矩阵的相抵

定义 1.1 设 $\mathbf{A}, \mathbf{B} \in \mathbb{F}^{m\times n}$，则 $\mathbf{A}$ 和 $\mathbf{B}$ 相抵当且存在可逆方阵 $\mathbf{P} \in \mathbb{F}^{m\times m}, \mathbf{Q} \in \mathbb{F}^{n\times n}$，使得

$$\mathbf{B} = \mathbf{P}\mathbf{A}\mathbf{Q}$$

矩阵的相抵标准型

由 **CR** 分解可知，任意矩阵相抵于 $\begin{pmatrix} \mathbf{I}_r & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{pmatrix}$，因此矩阵相抵的标准型为

$$\begin{pmatrix} \mathbf{I}_r & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{pmatrix}$$

矩阵相抵的性质

相抵关系满足自反性、对称性、传递性。

定理 1.1 矩阵相抵的充要条件是秩相等。

因此**秩是矩阵的相抵不变量**。

1.2.3 矩阵的秩

相抵标准型的应用

常见秩的不等式:

1.

$$\operatorname{rank}\begin{pmatrix}\mathbf{A} & \mathbf{C}\\ \mathbf{O} & \mathbf{B}\end{pmatrix} \geqslant \operatorname{rank}\mathbf{A} + \operatorname{rank}\mathbf{B}, \quad \text{when } \mathbf{C} = \mathbf{O}$$

2. 线性变换只能维度不变或降低:

$$\operatorname{rank}(\mathbf{AB}) \leqslant \min\{\operatorname{rank}\mathbf{A}, \operatorname{rank}\mathbf{B}\}$$

3. Frobenius 不等式:

$$\operatorname{rank}(\mathbf{AC}) + \operatorname{rank}(\mathbf{CB}) \leqslant \operatorname{rank}\mathbf{C} + \operatorname{rank}(\mathbf{ACB})$$

4. 加法不等式:

$$\operatorname{rank}(\mathbf{A} + \mathbf{B}) \leqslant \operatorname{rank}\mathbf{A} + \operatorname{rank}\mathbf{B}$$

5. 拆项不等式:

$$\operatorname{rank}(\mathbf{I} - \mathbf{AB}) \leqslant \operatorname{rank}(\mathbf{I} - \mathbf{A}) + \operatorname{rank}(\mathbf{I} - \mathbf{B})$$

6. 转置保秩性:

$$\operatorname{rank}(\mathbf{A}^\top) = \operatorname{rank}\mathbf{A} = \operatorname{rank}(\mathbf{A}\mathbf{A}^\top) = \operatorname{rank}(\mathbf{A}^\top\mathbf{A})$$

1.3 LU 分解与线性方程组

1.4 QR 分解与线性空间

1.5 $\mathbf{TJT}^{-1}$ 分解、矩阵的相似与 Jordan 标准形

1.5.1 Jordan 标准形

定义 1.2 对于任意 $\lambda \in \mathbb{C}, m \in \mathbb{N}$，形如

$$\begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & 1 & \cdots & 0 & 0 \\ 0 & 0 & \lambda & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \\ 0 & 0 & 0 & \cdots & \lambda & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda \end{pmatrix}$$

的 m 阶方阵称为 *Jordan* 块，记做 $\mathbf{J}_m(\lambda)$，其中 m 表示它的阶数，λ 是它的对角线元也是其特征值。

Jordan 形矩阵 $\mathbf{J}$ 满足:

$$\mathbf{J} = \mathrm{diag}(\mathbf{J}_{m_1}(\lambda_1), \mathbf{J}_{m_2}(\lambda_2) \cdots, \mathbf{J}_{m_k}(\lambda_k))$$

定理 1.2 任意一个复方阵 $\mathbf{A}$ 都相似与一个 *Jordan* 形矩阵 $\mathbf{J}$，它的每个 *Jordan* 块的特征值都是 $\mathbf{A}$ 的特征值。若不计 *Jordan* 块的排列顺序，则 $\mathbf{J}$ 唯一且由 $\mathbf{A}$ 的特征值和秩唯一确定。

对于 Jordan 块 $\mathbf{J}_m(0)$，我们有

$$\mathrm{rank}(\mathbf{J}_m(0)^k) = \begin{cases} m-k,\ k \leqslant m \\ 0,\ k > m \end{cases}$$

1.5.2 矩阵的相似

定义 1.3 设 $\mathbf{A}, \mathbf{B} \in \mathbb{F}^{n\times n}$，则 $\mathbf{A}$ 和 $\mathbf{B}$ 相合当且仅当存在可逆方阵 $\mathbf{T} \in \mathbb{F}^{n\times n}$，使得

$$\mathbf{B} = \mathbf{T}^{-1}\mathbf{A}\mathbf{T}$$

矩阵的相似标准型

任意矩阵相似于一个 Jordan 形，因此矩阵的相似标准型为

$$\mathbf{J} = \mathrm{diag}(\mathbf{J}_{m_1}(\lambda_1), \mathbf{J}_{m_2}(\lambda_2) \cdots, \mathbf{J}_{m_k}(\lambda_k))$$

矩阵相似的性质

相似关系满足自反性、对称性、传递性。

定理 1.3 相似的矩阵具有相同的特征多项式与特征值。

如果一个方阵相似于对角阵则称该方阵可对角化。

定理 1.4 数域 $\mathbb{F}$ 上的 n 阶方阵 $\mathbf{A}$ 相似与对角阵的充分必要条件是 $\mathbf{A}$ 有 n 个线性无关的特征向量。

引理 1.1 设 $\mathbf{A}$ 是数域 $\mathbb{F}$ 上的 n 阶方阵，则属于 $\mathbf{A}$ 的不同特征值的特征向量是线性无关的。

推论 1.1 如果矩阵 $\mathbf{A}$ 的 n 个特征值两两不同，则 $\mathbf{A}$ 相似于对角阵。

不是所有矩阵都可对角化，但它总可相似于一个上三角阵。

定理 1.5 任何一个 n 阶复方阵 $\mathbf{A}$ 都可以相似于一个上三角阵，且该上三角阵的主对角线上的元素都是 $\mathbf{A}$ 的特征值。

1.6 $\mathbf{Q}\Lambda\mathbf{Q}^\top$ 分解、矩阵的相合与二次型

1.6.1 $\mathbf{Q}\Lambda\mathbf{Q}^\top$ 分解

命题 1.1 实对称方阵的特征值均为实数。

命题 1.2 实对称方阵的不同特征值对应的特征向量彼此正交。

在已知上述命题的情况下，可以由数学归纳法得出，任意实对称方阵 $\mathbf{A} \in \mathbb{R}^{n\times n}$，存在正交矩阵 $\mathbf{Q} \in \mathbb{R}^{n\times n}$，使得

$$\mathbf{A} = \mathbf{Q}\Lambda\mathbf{Q}^\top$$

称为矩阵的 $\mathbf{Q}\Lambda\mathbf{Q}^\top$ 分解，$\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$，其中 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 是 $\mathbf{A}$ 的特征值。

1.6.2 矩阵的相合

定义 1.4 设 $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n\times n}$，则 $\mathbf{A}$ 和 $\mathbf{B}$ 相合当且存在可逆矩阵 $\mathbf{P} \in \mathbb{R}^{n\times n}$，使得

$$\mathbf{B} = \mathbf{P}^\top\mathbf{A}\mathbf{P}$$

本章后续主要讨论实对称矩阵的相合。

矩阵的相合标准型

由二次型的标准型及二次型与实对称矩阵的对应关系可知,任意矩阵相合于 $\begin{pmatrix} I_r & & \\ & -I_s & \\ & & \mathbf{O} \end{pmatrix}$，其中 $r + s = \mathrm{rank}\,\mathbf{A}$，分别被称为实对称阵 $\mathbf{A}$ 的**正、负惯性指数**，分别与 $\mathbf{A}$ 的正、负特征值数量相等。因此实对称矩阵的相合标准型为

$$\begin{pmatrix} I_r & & \\ & -I_s & \\ & & \mathbf{O} \end{pmatrix}$$

矩阵相合的性质

相合关系满足自反性、对称性、传递性。

定理 1.6 任意 n 阶实对称方阵 $\mathbf{A}$ 相合于对角阵。

定理 1.7 任意 n 阶实对称方阵 $\mathbf{A}$ 与 $\mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$ 相合且变化矩阵为正交方针。即存在正交方阵，使得

$$\mathbf{P}^\top\mathbf{A}\mathbf{P} = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$$

其中 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 是 $\mathbf{A}$ 的特征值。

由实对称矩阵的相合标准型可知**实对称阵的相合不变量为正、负惯性指数**。

1.7 UΣV 分解

设 $\mathbf{A} \in \mathbb{F}^{m\times n}$，则存在正交矩阵 $\mathbf{U} \in \mathbb{F}^{m\times m}, \mathbf{V} \in \mathbb{F}^{n\times n}, \Sigma \in \mathbb{F}^{m\times n}$，使得

$$\mathbf{A} = \mathbf{U}\Sigma\mathbf{V}$$

其中 $\Sigma = \begin{pmatrix} \Lambda & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{pmatrix}$。$\Lambda = \mathrm{diag}(\sigma_1, \sigma_2, \cdots, \sigma_r)$，$\sigma_1 \geqslant \sigma_2 \geqslant \cdots \geqslant \sigma_r$ 称为 $\mathbf{A}$ 的奇异值。

第二章　线性空间及变换

2.1　线性空间

2.2 线性变换

定义 2.1 设 V, V" 是 $\mathbb{F}$ 上的两个线性空间，若映射 $\mathscr{A}: V \to V'$，$\forall \alpha, \beta \in V,\ \lambda \in fF$，满足：

$$\mathscr{A}(\alpha + \beta) = \mathscr{A}(\alpha) + \mathscr{A}(\beta)$$

$$\mathscr{A}(\lambda\alpha) = \lambda\mathscr{A}(\alpha)$$

则称 $\mathscr{A}$ 是线性空间 V 到线性空间 V' 的线性映射。特别地，如果 $V = V'$，则称 $\mathscr{A}$ 为线性空间 V 上的一个线性变换。

最常见的线性变换有把每个向量映射为自身的**恒等变换**（或**单位变换**）：

$$\mathscr{E}: \mathscr{E}(x) = x,\ x \in V$$

以及把每个向量都映射为加法单位元的**零变换**：

$$\mathscr{O}: \mathscr{O}(x) = \theta,\ x \in V$$

2.2.1 线性变换的性质

定理 2.1 对于线性空间 V 上的线性变换 $\mathscr{A}$，有如下性质：

1. $\mathscr{A}(\theta) = \theta$；
2. $\mathscr{A}(-\alpha) = -\mathscr{A}(\alpha),\ \forall \alpha \in V$；
3. $\mathscr{A}(\lambda_1\alpha_1 + \lambda_2\alpha_2 + \cdots + \lambda_n\alpha_n) = \lambda_1\mathscr{A}(\alpha_1) + \lambda_2\mathscr{A}(\alpha_2) + \cdots + \lambda_n\mathscr{A}(\alpha_n),\ \forall \lambda_1, \lambda_2, \cdots, \lambda_n \in fF,\ \alpha_1, \alpha_2, \cdots, \alpha_n \in V$；
4. 若 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 是一组线性相关的向量，则 $\mathscr{A}(\alpha_1), \mathscr{A}(\alpha_2), \cdots, \mathscr{A}(\alpha_n)$ 也是一组线性相关的向量。

定义 2.2 设 $\mathscr{A}$ 是线性空间 V 上的线性变换，$\alpha_1, \alpha_2, \cdots, \alpha_n$ 是 V 的一组基，则有：

$$\mathscr{A}(\alpha_1\alpha_2\cdots\alpha_n) := (\mathscr{A}(\alpha_1), \mathscr{A}(\alpha_2), \cdots, \mathscr{A}(\alpha_n)) = (\alpha_1\alpha_2\cdots\alpha_n)\mathbf{A}$$

称 $\mathbf{A}$ **为线性变换** $\mathscr{A}$ **在基** $\alpha_1, \alpha_2, \cdots, \alpha_n$ **下的矩阵**。

上述定义表明，对于确定的一组基，线性变换与变换矩阵是一一对应的，故下面讨论二者的关系：

2.2.2 线性变换与矩阵

线性空间 V 的基 $\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$ 中，向量 α,β 对应的坐标分别为 $\boldsymbol{x},\boldsymbol{y}$，线性变换 $\mathscr{A}$ 对应的矩阵为 $\mathbf{A}$。则 $\beta=\mathscr{A}\alpha$ 对应的矩阵形式就是 $\boldsymbol{y}=\mathbf{A}\boldsymbol{x}$。

对于线性空间 V 上的线性变换 $\mathscr{A}$，由如下定义：

定义 2.3 $\mathscr{A}$ 像的全体称为**像空间**，记为 $\mathrm{Im}(\mathscr{A}):=\{\mathscr{A}(\alpha)|\alpha\in V\}$。
$\mathscr{A}$ 的**核空间**$\mathrm{Ker}(\mathscr{A}):=\{\alpha_V|\mathscr{A}(\alpha)=\theta\}$，$\mathscr{A}$ 的和空间是 V 的子空间。

像空间与核空间与矩阵 $\mathbf{A}$ 的列空间 $C(\mathbf{A})$ 和零空间 $N(\mathbf{A})$ 有如下关系：

$$Im(\mathscr{A})=\{(\alpha_1,\alpha_2,\cdots,\alpha_n)\boldsymbol{y}|\boldsymbol{y}\in C(\mathbf{A})\}$$

$$Ker(\mathscr{A})=\{(\alpha_1,\alpha_2,\cdots,\alpha_n)\boldsymbol{x}|\boldsymbol{x}\in N(\mathbf{A})\}$$

相似的，我们有以下结论：

1. $\dim(\mathrm{Im}(\mathscr{A}))=\mathrm{rank}(\mathbf{A})$;
2. $\dim(\mathrm{Im}(\mathscr{A}))+\dim(\mathrm{Ker}(\mathscr{A}))=\dim V$;

定理 2.2 若线性空间 V 中两组基 $\{\alpha_1,\alpha_2,\cdots,\alpha_n\},\{\beta_1,\beta_2,\cdots,\beta_n\}$，$\mathscr{A}$ 在两组级中的矩阵分别为 $\mathbf{A},\mathbf{B}$，两组基过度矩阵为 $\mathbf{T}$，即 $(\beta_1,\beta_2,\cdots,\beta_n)=(\alpha_1,\alpha_2,\cdots,\alpha_n)\mathbf{T}$，则

$$\mathbf{B}=\mathbf{T}^{-1}\mathbf{A}\mathbf{T}$$

即矩阵 $\mathbf{A}$ 与 $\mathbf{B}$ 相似。

线性映射 $\mathscr{A}$ 的特征多项式、行列式、秩、迹就是对应的矩阵的特征多项式、行列式、秩、迹。下面会证明这四者均只与矩阵特征值有关，且为相似不变量。

2.2.3 特征值与特征向量

线性映射的特征值与特征向量和其对应的矩阵的特征值与特征向量相同（且为相似不变量）。故下面仅考虑矩阵的特征值与特征向量：

定义 2.4 $\mathbf{A}\in\mathbb{F}^{n\times n}$，如果存在 $\lambda\in\mathbb{F}$ 和 $\boldsymbol{x}\in\mathbb{F}^n/\{0\}$，使得

$$\mathbf{A}\boldsymbol{x}=\lambda\boldsymbol{x}$$

则称 λ 为矩阵 $\mathbf{A}$ 的**特征值**，$\boldsymbol{x}$ 为矩阵 $\mathbf{A}$ 的**特征向量**。

定义 2.5 设 $\lambda\in\mathbb{F}$ 是矩阵 $\mathbf{A}\in\mathbb{F}^{n\times n}$ 的特征值，则定义

$$V_A(\lambda):=\{\boldsymbol{x}\in\mathbb{F}^n|\mathbf{A}\boldsymbol{x}=\lambda\boldsymbol{x}\}$$

为矩阵 $\mathbf{A}$ 属于特征值 λ 的特征子空间，特征子空间由所有的特征向量与零向量一起构成。

定义 2.6 设 $\mathbf{A} \in \mathbb{F}^{n\times n}$，称 $p_A(\lambda) = \det(\lambda\mathbf{I} - \mathbf{A})$ 为 $\mathbf{A}$ 的**特征多项式**，$p_A(\lambda) = 0$ 的解就是 $\mathbf{A}$ 的特征值。

定理 2.3 矩阵 $\mathbf{A} \in \mathbb{F}^{n\times n}$ 的特征值为 $\lambda_1, \lambda_2, \cdots, \lambda_n$，则：

1. λ^k 是 $\mathbf{A}^k$ 的特征值，λ 是 $\mathbf{A}^\top$ 的特征值，$\lambda \neq 0$ 时 $\dfrac{1}{\lambda}\det(\mathbf{A})$ 是 $\mathbf{A}^*$ 的特征值；
2. $\mathrm{Tr}(\mathbf{A}) = \sum\limits_{i=1}^{n} \lambda_i$；
3. $\det(\mathbf{A}) = \prod\limits_{i=1}^{n} \lambda_i$。

第三章　内积空间及变换

3.1　内积空间

3.1.1　数组内积空间

定义 3.1 设 $\mathbb{R}^n$ 是 $\mathbb{R}$ 上的 n 维数组空间，对任意列向量 $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^n$，定义他们的内积

$$(\boldsymbol{a}, \boldsymbol{b}) := \boldsymbol{a}^\top \mathbf{G} \boldsymbol{b}$$

其中 $\mathbf{G} \in \mathbb{R}^{n\times n}$，为 *Gram* 矩阵，也称为度量矩阵，是一个 n 阶实对称正定方阵，满足

$$\boldsymbol{a}^\top \mathbf{G} \boldsymbol{a} \geqslant 0, \quad \text{when } \boldsymbol{a} = 0$$

定义了内积的数组空间 $\mathbb{R}^n$ 称为 *Euclid* 空间，也称为欧氏空间。

一般欧氏空间

定义 3.2 设 V 是实数域 $\mathbb{R}$ 上的线性空间，如果 V 中任意两个向量 α, β 都按照某一法则对英语一个实数，记为 (α, β)，且满足:

1. 对称性：$(\alpha, \beta) = (\beta, \alpha)$；
2. 线性性：$(\lambda\alpha, \beta) = \lambda(\alpha, \beta)$ 和 $(\alpha + \beta, \gamma) = (\alpha, \gamma) + (\beta, \gamma)$；
3. 正定性：$(\alpha, \alpha) \geqslant 0, \quad \text{when } \alpha = \theta$，其中 θ 是 V 中的加法单位元。

在定义了内积的基础上定义模长:

定义 3.3 设 V 是欧氏空间，$(,)$ 是 V 的内积，对于任意 $\alpha \in V$，定义

$$|\alpha| = \sqrt{(\alpha, \alpha)}$$

称为 α 的模，也称为 α 的范数。

一般欧氏空间均有:

定理 3.1 (Cauchy-Schwarz 不等式)

$$|(\alpha, \beta)| \leqslant \sqrt{(\alpha, \alpha)(\beta, \beta)}$$

定义 3.4 (标准正交基) 设 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 是 V 中的一组基，若

$$(\alpha_i,\alpha_j)=\delta_{ij},\ i,j=1,2,\ldots,n$$

则称 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 为 V 的**一组标准正交基**。

定理 3.2 在数组空间 $\mathbb{R}^n$ 中，若 $\boldsymbol{u}_1,\boldsymbol{u}_2,\cdots,\boldsymbol{u}_n$ 是一组标准正交基，则任意 $\boldsymbol{a}\in\mathbb{R}^n$ 都可以表示为:

$$\boldsymbol{a}=(\boldsymbol{a},\boldsymbol{u}_1)\boldsymbol{u}_1+\cdots+(\boldsymbol{a},\boldsymbol{u}_n)\boldsymbol{u}_n$$

定理 3.3 (Schmidt 正交化) 从 n 维欧氏空间 V 中任意一组基 $\alpha_1,\alpha_2,\cdots,\alpha_n$，可以构造一组标准正交基 $\beta_1,\beta_2,\cdots,\beta_n$，使得

$$\operatorname{span}(\alpha_1,\alpha_2,\cdots,\alpha_n)=\operatorname{span}(\beta_1,\beta_2,\cdots,\beta_n)$$

定义 3.5 (空间与向量正交、空间与空间正交) 设 W_1,W_2 是欧氏空间 V 的自空间，对向量 $\alpha\in V$，若 $\alpha\perp\beta.\ \forall\beta\in W_1$，则称 α 子空间 W_1 正交，记为 $\alpha\perp W_1$；若 $\forall\alpha\in W_2,\ \alpha\perp W_1$，则称 W_1 与 W_2 正交，记为 $W_1\perp W_2$。

定理 3.4 (正交补空间) 设 W 是欧氏空间 V 的子空间，则向量集合

$$W^{\perp}:=\{\alpha\in V|\alpha\perp W\}$$

是 V 的自空间，且有 $V=W\oplus W^{\perp}$，称为 W 的**正交补空间**。

定义 3.6 (正交投影) 设 W 是欧氏空间 V 的子空间，$\alpha\in V$，如果向量 $\beta\in W$ 满足

$$\alpha-\beta\perp W^{\perp}$$

，则称 β 是 α 在 W 上的正交投影，记为 $\beta=P\alpha$。

定理 3.5 设 W 是欧氏空间 V 的子空间，$\alpha_1,\alpha_2,\cdots,\alpha_n$ 是 W 的一组标准正交基，则下列结论成立:

1. 对任意 $\alpha\in V$，$P\alpha$ 唯一且

$$P\alpha=(\alpha,\alpha_1)\alpha_1+\cdots+(\alpha,\alpha_n)\alpha_n$$

2. 设 $\alpha\in V$，则对任意 $\beta\in W$，有

$$|\alpha-\beta|\geqslant|\alpha-P\alpha|,\quad \text{when }\beta=P\alpha$$

3.2 内积空间上的变换

3.2.1 正交变换

定义 3.7 设 V 是一个 n 维欧氏空间，$\mathscr{A}$ 是一个线性变换，如果 $\mathscr{A}$ 保持 V 的内积不变，即

$$(\mathscr{A}\alpha, \mathscr{A}\beta) = (\alpha, \beta),\ \forall \alpha, \beta \in V$$

则称 $\mathscr{A}$ 是 V 内的正交变换。

定理 3.6 下列命题等价：

1. $\mathscr{A}$ 是一个正交变换；
2. $\mathscr{A}$ 保持任意向量的模长不变；
3. $\mathscr{A}$ 将标准正交基变换为标准正交基；
4. $\mathscr{A}$ 在任意一组标准正交基下的矩阵为正交阵。

设 $\mathscr{A}$ 是 V 上的正交变换，则 $\det(\mathscr{A}) = \pm 1$。若 $\det(\mathscr{A}) = 1$，则称 $\mathscr{A}$ 是 V 上的**第一类正交变换**；若 $\det(\mathscr{A}) = -1$，则称 $\mathscr{A}$ 是 V 上的**第二类正交变换**。

命题 3.1 设 $\mathscr{A}$ 是 V 上的第一类正交变换，且 V 的维数是奇数，则 $\mathscr{A}$ 一定有特征值 1。

3.2.2 正交方阵

正交变换在标准正交基下的矩阵为正交阵，下面来考虑一些正交方阵的性质。

定义 3.8 (正交方阵) 若方阵 $\mathbf{Q}$ 满足

$$\mathbf{Q}\mathbf{Q}^\top = \mathbf{Q}^\top\mathbf{Q} = \mathbf{I}$$

则称 $\mathbf{Q}$ 为正交方阵。

定理 3.7 下列命题等价：

1. $\mathbf{Q}$ 是正交方阵；
2. $\mathbf{Q}^\top = \mathbf{Q}^{-1}$；
3. $\mathbf{Q}$ 的行（或列）向量构成 fR^n 的一组标准正交基。

定理 3.8 正交阵有下列性质：

1. 单位真是正交阵；
2. $\mathbf{A}, \mathbf{B}$ 是同阶正交方阵，则 $\mathbf{AB}$ 也是正交方阵；

3. 若 $\mathbf{A}$ 是正交方阵，则 $\mathbf{A}^\top$ 和 $\mathbf{A}^{-1}$ 也是正交方阵；

4. 若 $\mathbf{A}$ 是正交方阵，则 $\det(\mathbf{A}) = \pm 1$;

5. 正交方阵的特征值的模长为 1。

3.2.3 对称变换

定义 3.9 设 V 是一个 n 维欧氏空间，$\mathscr{A}$ 是一个线性变换，如果 $\mathscr{A}$ 满足

$$(\alpha, \mathscr{A}\beta) = (\mathscr{A}\alpha, \beta),\ \forall \alpha, \beta \in V$$

则称 $\mathscr{A}$ 是 V 上的对称变换。

定理 3.9 线性变换 sA 是对称变换的重要条件是 $\mathscr{A}$ 在任意一组标准正交基下的矩阵 $\mathbf{A}$ 是实对称方阵。

3.2.4 实对称方阵

研究实对称方阵就是研究实对称变换。实对称方阵的性质在 $\mathbf{A}\Lambda\mathbf{Q}^\top$ 分解中已经给出。

第四章　二次型

4.1　二次型的定义与分类

定义 4.1 在实数域上，一个含 n 个变元 $x_1, x_2, \cdots, x_n$ 的**二次型**$Q(x_1, x_2, \cdots, x_n)$ 是一个齐次的二次多项式

$$Q(x_1, x_2, \cdots, x_n) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j = \boldsymbol{x}^\top \mathbf{A}\boldsymbol{x}$$

这里 $a_{ij} = a_{ji}$ 为实系数，$\mathbf{A} = (a_{ij})$ 为 n 阶实对称方阵。称矩阵 $\mathbf{A}$ 为二次型 $Q(x_1, x_2, \cdots, x_n)$ 的矩阵，$\mathbf{A}$ 的秩称为二次型的秩。

由实对称方阵的相合标准型可以定义二次型的规范形：

$$Q(x_1, x_2, \cdots, x_n)|_{\boldsymbol{x}=\mathbf{P}\boldsymbol{y}} = y_1^2 + y_2^2 + \cdots + y_r^2 - y_{r+1}^2 - y_{r+2}^2 - \cdots - y_{r+s}^2$$

其中 r, s 分别为二次型对应的矩阵 $\mathbf{A}$ 的正、负惯性指数，$\mathbf{P}$ 是相合变换对应的可逆矩阵，满足

$$\mathbf{P}^\top \mathbf{A}\mathbf{P} = \begin{pmatrix} I_r & & \\ & -I_s & \\ & & \mathbf{O} \end{pmatrix}$$

4.1.1　二次曲线和曲面的分类

给定平面二次曲线的方程

$$\boldsymbol{x}^\top \mathbf{A}\boldsymbol{x}$$

可以化为标准形式

$$\lambda_1(x')^2 + \lambda_2(y')^2 + 2b_1'x' + 2b_2'y' + c' = 0$$

其中 λ_1, λ_2 为 $\mathbf{A}$ 的特征值，λ_1, λ_2 至少一个不为零。则可对二次曲面做如下分类：

1. **椭圆型**：$\lambda_1\lambda_2 > 0$
2. **双曲型**：$\lambda_1\lambda_2 < 0$
3. **抛物型**：$\lambda_1\lambda_2 = 0$

若上述 $\boldsymbol{x} \in \mathbb{R}^3$，则为二次曲面，其分类形式见教材 P272。

4.2 正定二次型

定义 4.2 设 $Q(x_1, x_2, \cdots, x_n)$ 为实数域上 n 个变量的二次型，若 $\forall x_1, x_2, \cdots, x_n \in \mathbb{R}^n/\{0\}$，都有

$$Q(x_1, x_2, \cdots, x_n) > 0$$

则称 $Q(x_1, x_2, \cdots, x_n)$ 为**正定二次型**，对应的矩阵 $\mathbf{A}$ 称为正定矩阵，矩阵正定简记为 $\mathbf{A} > 0$。

内积空间中的 Gram 矩阵就是一个正定矩阵。考虑到正定二次型与正定矩阵一一对应，故下面通过讨论正定矩阵的性质来研究正定二次型的性质。

4.2.1 正定矩阵

定理 4.1 设 $\mathbf{A}$ 是 n 阶实对称矩阵，则下列命题等价：

1. $\mathbf{A}$ 是正定矩阵；
2. $\mathbf{A}$ 相合于单位阵；
3. $\mathbf{A}$ 的特征值全为正；
4. 存在可逆方阵 $\mathbf{P}$ 使得 $\mathbf{A} = \mathbf{P}^\top \mathbf{P}$；
5. $\mathbf{A}$ 的各阶顺序主子式均为正。

由命题二可知**同一阶正定方阵属于同一个相合等价类**；由命题三可知**正定阵的行列式大于 0**。

4.3 其他类型的二次型

4.3.1 半正定二次型

定义 4.3 设 $Q(x_1, x_2, \cdots, x_n)$ 为实数域上 n 个变量的二次型，若 $\forall x_1, x_2, \cdots, x_n \in \mathbb{R}^n$，都有

$$Q(x_1, x_2, \cdots, x_n) \geqslant 0$$

则称 $Q(x_1, x_2, \cdots, x_n)$ 为**半正定二次型**，对应的矩阵 $\mathbf{A}$ 称为半正定矩阵，矩阵半正定简记为 $\mathbf{A} \geqslant 0$。

4.3.2 半正定矩阵

定理 4.2 设 $\mathbf{A}$ 是 n 阶实对称矩阵，则下列命题等价：

1. $\mathbf{A}$ 是半正定矩阵；
2. $\mathbf{A}$ 相合于 $\begin{pmatrix} I_r & \\ & \mathbf{O} \end{pmatrix}$，其中 $r = \operatorname{rank} \mathbf{A}$；
3. $\mathbf{A}$ 的特征值全非负；
4. 存在实方阵 $\mathbf{P}$ 使得 $\mathbf{A} = \mathbf{P}^\top \mathbf{P}$；
5. $\mathbf{A}$ 的各阶主子式均非负。

与正定矩阵的命题不同，命题五仅要求“$\mathbf{A}$ 的各阶主子式”而非“$\mathbf{A}$ 的各阶顺序主子式“。

4.3.3 负定二次型与半负定二次型

将正定（半正定）二次型中 $Q(x_1, x_2, \cdots, x_n)$ 取负变为 $-Q(x_1, x_2, \cdots, x_n)$ 即为负定（半负定）二次型。相关定理与性质可同理推出。