线性代数 第14讲

10月27日

第三章第2讲 正交矩阵和QR分解

上一讲要点回顾

正交矩阵

QR分解

Household变换（反射变换）

内积

定义3.1.2(内积) 定义 R^n 上的两个向量 a,b 的内积为实数 $a^T b$,

即如果 $a=\begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}, b=\begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$, 则 a,b 的内积为 $a_1b_1+\cdots+a_nb_n$.

命题 3.1.3 向量内积满足如下性质:

1. **对称性**: $\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}=\boldsymbol{b}^{\mathrm{T}}\boldsymbol{a}$;
2. **双线性性**: $\boldsymbol{a}^{\mathrm{T}}(k_1\boldsymbol{b}_1+k_2\boldsymbol{b}_2)=k_1\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}_1+k_2\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}_2, (k_1\boldsymbol{a}_1+k_2\boldsymbol{a}_2)^{\mathrm{T}}\boldsymbol{b}=k_1\boldsymbol{a}_1^{\mathrm{T}}\boldsymbol{b}+k_2\boldsymbol{a}_2^{\mathrm{T}}\boldsymbol{b}$;
3. **正定性**: $\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a}\geqslant 0$, 且 $\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a}=0$ 当且仅当 $\boldsymbol{a}=\boldsymbol{0}$.

向量 $a=\begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}$ 的长度定义为: $\|a\|=\sqrt{a^T a}=\sqrt{a_1^{\ 2}+\cdots+a_n^{\ 2}}$,

$\|a-b\|$ 称为向量 a,b 间的距离。

向量的夹角与正交

$|a^T b| \le \|a\| \cdot \|b\|$, 等号成立当且仅当 a,b 线性相关(或是说共线)。

根据Cauchy-Schwarz不等式, $|a^T b| \le \|a\| \cdot \|b\|$, 对于非零向量 a,b, $\left|\dfrac{a^T b}{\|a\| \cdot \|b\|}\right| \le 1$,

定义非零向量 a,b 的夹角为 $\arccos \dfrac{a^T b}{\|a\| \cdot \|b\|}$.

如果向量 a,b 满足 $a^T b = 0$, 称二者正交或垂直，记为 $a \perp b$. 零向量与任何向量都正交。

定义 3.1.8 (正交向量组) 设 $a_1, \cdots, a_r$ 是 $\mathbb{R}^n$ 中的向量组，如果这些向量都非零且两两正交，则称该向量组为**正交向量组**. 特别地，如果正交向量组中的向量都是单位向量，则称其为**正交单位向量组**.

命题 3.1.9 正交向量组线性无关.

定义 3.1.10 (标准正交基) 设 $\mathcal{M}$ 是 $\mathbb{R}^n$ 的子空间，如果它的一组基是正交向量组，则称之为 $\mathcal{M}$ 的一组**正交基**；如果它的一组基是正交单位向量组，则称之为 $\mathcal{M}$ 的一组**标准正交基**.

Gram-Schmidt 正交化

将任意一组基改造为标准正交基

设 $\mathcal{M}$ 是 $\mathbb{R}^n$ 的子空间，从 $\mathcal{M}$ 的任意一组基 $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_r$ 出发，

$$\tilde{\boldsymbol{q}}_1 = \boldsymbol{a}_1,$$

$$\tilde{\boldsymbol{q}}_2 = \boldsymbol{a}_2 - \frac{\tilde{\boldsymbol{q}}_1^{\mathrm{T}} \boldsymbol{a}_2}{\tilde{\boldsymbol{q}}_1^{\mathrm{T}} \tilde{\boldsymbol{q}}_1} \tilde{\boldsymbol{q}}_1,$$

$$\tilde{\boldsymbol{q}}_3 = \boldsymbol{a}_3 - \frac{\tilde{\boldsymbol{q}}_1^{\mathrm{T}} \boldsymbol{a}_3}{\tilde{\boldsymbol{q}}_1^{\mathrm{T}} \tilde{\boldsymbol{q}}_1} \tilde{\boldsymbol{q}}_1 - \frac{\tilde{\boldsymbol{q}}_2^{\mathrm{T}} \boldsymbol{a}_3}{\tilde{\boldsymbol{q}}_2^{\mathrm{T}} \tilde{\boldsymbol{q}}_2} \tilde{\boldsymbol{q}}_2,$$

$$\vdots$$

$$\tilde{\boldsymbol{q}}_r = \boldsymbol{a}_r - \frac{\tilde{\boldsymbol{q}}_1^{\mathrm{T}} \boldsymbol{a}_r}{\tilde{\boldsymbol{q}}_1^{\mathrm{T}} \tilde{\boldsymbol{q}}_1} \tilde{\boldsymbol{q}}_1 - \cdots - \frac{\tilde{\boldsymbol{q}}_{r-1}^{\mathrm{T}} \boldsymbol{a}_r}{\tilde{\boldsymbol{q}}_{r-1}^{\mathrm{T}} \tilde{\boldsymbol{q}}_{r-1}} \tilde{\boldsymbol{q}}_{r-1}.$$

为了得到标准正交基，只要再把正交基中的每个向量都单位化即可：$\boldsymbol{q}_i = \frac{\tilde{\boldsymbol{q}}_i}{\|\tilde{\boldsymbol{q}}_i\|}$.

R^n中的三角形，正交投影

在 $\mathbb{R}^n$ 中的三角形的三条边能写成向量 $\boldsymbol{a}$，$\boldsymbol{b}$，$\boldsymbol{a}+\boldsymbol{b}$，两边（长度）之和大于第三边（长度）.

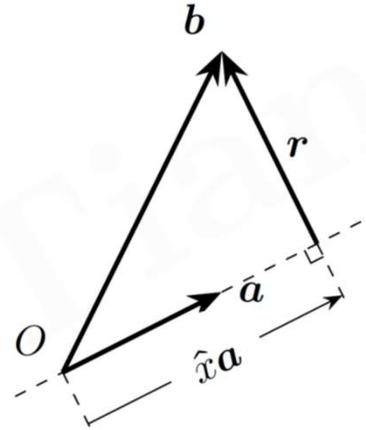


图 3.1.1: 平面向量的逼近

推论 3.1.5 (三角不等式) $\|\boldsymbol{a}+\boldsymbol{b}\| \leqslant \|\boldsymbol{a}\|+\|\boldsymbol{b}\|$，等号成立当且仅当 $\boldsymbol{a},\boldsymbol{b}$ 共线.

定理 3.1.6 (勾股定理) 向量 $\boldsymbol{a},\boldsymbol{b}$ 正交，则 $\|\boldsymbol{a}\pm\boldsymbol{b}\|^2=\|\boldsymbol{a}\|^2+\|\boldsymbol{b}\|^2$.

证. 根据定义，$\|\boldsymbol{a}\pm\boldsymbol{b}\|^2=(\boldsymbol{a}\pm\boldsymbol{b})^{\mathrm{T}}(\boldsymbol{a}\pm\boldsymbol{b})=\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a}\pm2\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}+\boldsymbol{b}^{\mathrm{T}}\boldsymbol{b}=\|\boldsymbol{a}\|^2+\|\boldsymbol{b}\|^2$. □

设 $\boldsymbol{a},\boldsymbol{b}\in \boldsymbol{R}^n$，$\boldsymbol{a}\neq 0$，则有 $\boldsymbol{b}=\dfrac{\boldsymbol{a}^T\boldsymbol{b}}{\boldsymbol{a}^T\boldsymbol{a}}\boldsymbol{a}+\left(\boldsymbol{b}-\dfrac{\boldsymbol{a}^T\boldsymbol{b}}{\boldsymbol{a}^T\boldsymbol{a}}\boldsymbol{a}\right)$

$\boldsymbol{a}^T\left(\boldsymbol{b}-\dfrac{\boldsymbol{a}^T\boldsymbol{b}}{\boldsymbol{a}^T\boldsymbol{a}}\boldsymbol{a}\right)=0$，$\dfrac{\boldsymbol{a}^T\boldsymbol{b}}{\boldsymbol{a}^T\boldsymbol{a}}\boldsymbol{a}$ 称为向量 $\boldsymbol{b}$ 向直线 $\mathrm{span}(\boldsymbol{a})$ 的正交投影。

命题 3.1.7 设 $\boldsymbol{a},\boldsymbol{b}$ 是 $\mathbb{R}^n$ 中的两个向量，$\boldsymbol{a}\neq\boldsymbol{0}$，则 $\left\|\boldsymbol{b}-\dfrac{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}}{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a}}\boldsymbol{a}\right\|=\min\limits_{x\in\mathbb{R}}\|\boldsymbol{b}-x\boldsymbol{a}\|$.

正交矩阵

设 $q_1, q_2, \cdots, q_n$ 是 R^n 的一组标准正交基，记 n 阶方阵 $Q = [q_1 \quad q_2 \quad \cdots \quad q_n]$，

则 $$Q^T Q = \begin{bmatrix} q_1^T \\ q_2^T \\ \vdots \\ q_n^T \end{bmatrix} [q_1 \quad q_2 \quad \cdots \quad q_n] = \begin{bmatrix} q_1^T q_1 & q_1^T q_2 & \cdots & q_1^T q_n \\ q_2^T q_1 & q_2^T q_2 & \cdots & q_2^T q_n \\ \vdots & \vdots & \ddots & \vdots \\ q_n^T q_1 & q_n^T q_2 & \cdots & q_n^T q_n \end{bmatrix} = I_n.$$

定义3.2.1（正交矩阵） 一个 n 阶方阵 Q 如果满足 $Q^T Q = I_n$. 则称 Q 是 n 阶正交矩阵。

命题 3.2.3 两个 n 阶正交矩阵的乘积还是 n 阶正交矩阵.

命题 3.2.4 对 n 阶方阵 Q，以下叙述等价：

1. Q 是正交矩阵，即 $Q^{\mathrm{T}} Q = I_n$；
2. Q 为保距变换，即，对任意 $\boldsymbol{x} \in \mathbb{R}^n$，$\|Q\boldsymbol{x}\| = \|\boldsymbol{x}\|$；
3. Q 为保内积变换，即，对任意 $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$，$Q\boldsymbol{x}$ 与 $Q\boldsymbol{y}$ 的内积等于 $\boldsymbol{x}$ 与 $\boldsymbol{y}$ 的内积.

正交矩阵的性质

命题 3.2.2 对 n 阶方阵 Q，以下叙述等价：

1. Q 是正交矩阵，即 $Q^{\mathrm{T}}Q = I_n$；
2. Q 可逆，且 $Q^{-1} = Q^{\mathrm{T}}$；
3. $QQ^{\mathrm{T}} = I_n$；
4. Q^{T} 是正交矩阵；
5. Q 可逆，且 Q^{-1} 是正交矩阵；
6. Q 的列向量组成 $\mathbb{R}^n$ 的一组标准正交基；
7. Q 的行向量的转置组成 $\mathbb{R}^n$ 的一组标准正交基.

例题选讲

练习 3.2.2 写出元素都是 0 或 1 的所有三阶正交矩阵.

练习 3.2.6 证明，上三角矩阵是正交矩阵时，必是对角矩阵，且对角元素是 ± 1.

练习 3.2.7 对标准基 $\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n$，显然 $\sum_{i=1}^{n} \boldsymbol{e}_i \boldsymbol{e}_i^{\mathrm{T}} = I_n$. 对任意标准正交基 $\boldsymbol{q}_1, \cdots, \boldsymbol{q}_n$，求证 $\sum_{i=1}^{n} \boldsymbol{q}_i \boldsymbol{q}_i^{\mathrm{T}} = I_n$.

练习 3.2.19 设 $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_m$ 是 $\mathbb{R}^n$ 中的 m 个向量，定义矩阵

$$G(\boldsymbol{a}_1, \cdots, \boldsymbol{a}_m) := \begin{bmatrix} \boldsymbol{a}_1^{\mathrm{T}} \boldsymbol{a}_1 & \cdots & \boldsymbol{a}_1^{\mathrm{T}} \boldsymbol{a}_m \\ \vdots & & \vdots \\ \boldsymbol{a}_m^{\mathrm{T}} \boldsymbol{a}_1 & \cdots & \boldsymbol{a}_m^{\mathrm{T}} \boldsymbol{a}_m \end{bmatrix},$$

称为 $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_m$ 的 **Gram 矩阵**. 证明，

1. $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_m$ 是正交单位向量组当且仅当 $G(\boldsymbol{a}_1, \cdots, \boldsymbol{a}_m) = I_m$.
2. Gram 矩阵 $G = G(\boldsymbol{a}_1, \cdots, \boldsymbol{a}_m)$ 是 m 阶对称矩阵，且对任意 $\boldsymbol{x} \in \mathbb{R}^m$，都有 $\boldsymbol{x}^{\mathrm{T}} G \boldsymbol{x} \geqslant 0$.
3. $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_m$ 线性无关当且仅当 $G = G(\boldsymbol{a}_1, \cdots, \boldsymbol{a}_m)$ 可逆，也等价于对任意非零 $\boldsymbol{x} \in \mathbb{R}^m$，都有 $\boldsymbol{x}^{\mathrm{T}} G \boldsymbol{x} > 0$.

QR分解

设 $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n$ 是 $\mathbb{R}^n$ 的一组基，Gram-Schmidt 正交化的计算过程分为两步，

第一步正交化，得到一组正交基

$$
\begin{aligned}
\tilde{\boldsymbol{q}}_1 &= \boldsymbol{a}_1, \\
\tilde{\boldsymbol{q}}_2 &= \boldsymbol{a}_2 - \frac{\tilde{\boldsymbol{q}}_1^{\mathrm{T}} \boldsymbol{a}_2}{\tilde{\boldsymbol{q}}_1^{\mathrm{T}} \tilde{\boldsymbol{q}}_1} \tilde{\boldsymbol{q}}_1, \\
\tilde{\boldsymbol{q}}_3 &= \boldsymbol{a}_3 - \frac{\tilde{\boldsymbol{q}}_1^{\mathrm{T}} \boldsymbol{a}_3}{\tilde{\boldsymbol{q}}_1^{\mathrm{T}} \tilde{\boldsymbol{q}}_1} \tilde{\boldsymbol{q}}_1 - \frac{\tilde{\boldsymbol{q}}_2^{\mathrm{T}} \boldsymbol{a}_3}{\tilde{\boldsymbol{q}}_2^{\mathrm{T}} \tilde{\boldsymbol{q}}_2} \tilde{\boldsymbol{q}}_2, \\
&\vdots \\
\tilde{\boldsymbol{q}}_n &= \boldsymbol{a}_n - \frac{\tilde{\boldsymbol{q}}_1^{\mathrm{T}} \boldsymbol{a}_n}{\tilde{\boldsymbol{q}}_1^{\mathrm{T}} \tilde{\boldsymbol{q}}_1} \tilde{\boldsymbol{q}}_1 - \cdots - \frac{\tilde{\boldsymbol{q}}_{n-1}^{\mathrm{T}} \boldsymbol{a}_n}{\tilde{\boldsymbol{q}}_{n-1}^{\mathrm{T}} \tilde{\boldsymbol{q}}_{n-1}} \tilde{\boldsymbol{q}}_{n-1}.
\end{aligned}
$$

第二步再单位化每个向量，得到标准正交基：$\boldsymbol{q}_i = \frac{\tilde{\boldsymbol{q}}_i}{\|\tilde{\boldsymbol{q}}_i\|}$.

QR分解

设 $A=\begin{bmatrix}\boldsymbol{a}_1 & \cdots & \boldsymbol{a}_n\end{bmatrix}, \widetilde{Q}=\begin{bmatrix}\tilde{\boldsymbol{q}}_1 & \cdots & \tilde{\boldsymbol{q}}_n\end{bmatrix}, Q=\begin{bmatrix}\boldsymbol{q}_1 & \cdots & \boldsymbol{q}_n\end{bmatrix}$，第一步正交化，能得到 $\boldsymbol{a}_i$ 被 $\tilde{\boldsymbol{q}}_1,\cdots,\tilde{\boldsymbol{q}}_i$ 线性表示的表示法．这可用矩阵乘法表示：

$$A=\widetilde{Q}\widetilde{R},\quad \widetilde{R}=\begin{bmatrix}1 & \frac{\tilde{\boldsymbol{q}}_1^{\mathrm{T}}\boldsymbol{a}_2}{\tilde{\boldsymbol{q}}_1^{\mathrm{T}}\tilde{\boldsymbol{q}}_1} & \cdots & \frac{\tilde{\boldsymbol{q}}_1^{\mathrm{T}}\boldsymbol{a}_n}{\tilde{\boldsymbol{q}}_1^{\mathrm{T}}\tilde{\boldsymbol{q}}_1} \\ & 1 & \ddots & \vdots \\ & & \ddots & \frac{\tilde{\boldsymbol{q}}_{n-1}^{\mathrm{T}}\boldsymbol{a}_n}{\tilde{\boldsymbol{q}}_{n-1}^{\mathrm{T}}\tilde{\boldsymbol{q}}_{n-1}} \\ & & & 1\end{bmatrix},$$

第二步单位化可以写成 $\widetilde{Q}=Q\operatorname{diag}(\|\tilde{\boldsymbol{q}}_i\|)$，因此

$$A=Q\operatorname{diag}(\|\tilde{\boldsymbol{q}}_i\|)\widetilde{R}=QR,$$

其中 $R=\operatorname{diag}(\|\tilde{\boldsymbol{q}}_i\|)\widetilde{R}$ 是对角元素都是正数的上三角矩阵，Q 是正交矩阵．注意，$\widetilde{Q}$ 一般不是正交矩阵．

可逆矩阵QR分解的唯一性

定理 3.2.7 (可逆矩阵的 QR 分解) 设 A 是 n 阶可逆矩阵，则存在唯一的分解 $A = QR$，其中 Q 是正交矩阵，R 是对角元都是正数的上三角矩阵.

分解 $A = QR$ 称为矩阵 A 的 **QR 分解**.

观察 Gram-Schmidt 正交化的计算过程，第一步正交化，是一系列对 A 的从左往右的倍加列变换，从而得到 $\widetilde{Q}$. 注意，这种倍加矩阵都是单位上三角矩阵，所以其乘积也是单位上三角矩阵. 因此 $A = \widetilde{Q}\widetilde{R}$，这里 $\widetilde{R}$ 是单位上三角矩阵. 第二步单位化，显然可得 R 是具有正对角元的上三角矩阵.

练习 3.2.13 设向量组 $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_k$ 线性无关，首先令 $\boldsymbol{q}_1$ 为与 $\boldsymbol{v}_1$ 平行的单位向量，然后令 $\boldsymbol{q}_2$ 为二维子空间 $\mathrm{span}(\boldsymbol{v}_1, \boldsymbol{v}_2)$ 中垂直于直线 $\mathrm{span}(\boldsymbol{v}_1)$ 的单位向量，再令 $\boldsymbol{q}_3$ 为三维子空间 $\mathrm{span}(\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3)$ 中垂直于平面 $\mathrm{span}(\boldsymbol{v}_1, \boldsymbol{v}_2)$ 的单位向量，以此类推. 这样得到的 $\boldsymbol{q}_1, \cdots, \boldsymbol{q}_k$ 与 Gram-Schmidt 正交化得到的结果是否一致？如果有区别的话，区别在哪里？从 QR 分解的角度如何解释？

例 3.2.8 回顾例 3.1.14，

$$A = \begin{bmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \boldsymbol{a}_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix},$$

利用 Gram-Schmidt 正交化方法得到：

$$\widetilde{Q} = \begin{bmatrix} \tilde{\boldsymbol{q}}_1 & \tilde{\boldsymbol{q}}_2 & \tilde{\boldsymbol{q}}_3 \end{bmatrix} = \begin{bmatrix} 1 & \frac{1}{3} & \frac{1}{2} \\ 1 & \frac{1}{3} & -\frac{1}{2} \\ 1 & -\frac{2}{3} & 0 \end{bmatrix}, \quad \widetilde{R} = \begin{bmatrix} 1 & \frac{2}{3} & \frac{1}{3} \\ & 1 & \frac{1}{2} \\ & & 1 \end{bmatrix},$$

再把正交向量单位化得到：$Q = \begin{bmatrix} \boldsymbol{q}_1 & \boldsymbol{q}_2 & \boldsymbol{q}_3 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & -\frac{2}{\sqrt{6}} & 0 \end{bmatrix}$,

$$R = \operatorname{diag}(\|\tilde{\boldsymbol{q}}_i\|)\widetilde{R} = \begin{bmatrix} \sqrt{3} & & \\ & \frac{\sqrt{6}}{3} & \\ & & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 1 & \frac{2}{3} & \frac{1}{3} \\ & 1 & \frac{1}{2} \\ & & 1 \end{bmatrix} = \begin{bmatrix} \sqrt{3} & \frac{2}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ & \frac{\sqrt{6}}{3} & \frac{1}{\sqrt{6}} \\ & & \frac{1}{\sqrt{2}} \end{bmatrix}.$$

这就是 A 的 QR 分解 $A = QR$.

列满秩矩阵的QR分解

定义 3.2.9 (列正交矩阵) 矩阵 Q，如果满足 $Q^{\mathrm{T}}Q = I_n$，则称为**列正交矩阵**.

显然，Q 是 $m \times n$ 列正交矩阵等价于 Q 的列向量组构成 $\mathbb{R}^m$ 的一个单位向量组，其中有 n 个向量. 此时必有 $m \geqslant n$.

定理 3.2.10 (QR 分解) 对 $m \times n$ 矩阵 A，其中 $m \geqslant n$，则

1. 存在 $m \times n$ 列正交矩阵 Q_1 和具有非负对角元的 n 阶上三角矩阵 R_1，使得 $A = Q_1R_1$；
2. 进一步地，存在 m 阶正交矩阵 Q 和 $m \times n$ 矩阵 R，使得 $A = QR$，其中 $R = \begin{bmatrix} R_1 \\ O \end{bmatrix}$, $Q = \begin{bmatrix} Q_1 & Q_2 \end{bmatrix}$，即 Q 的列向量组由 Q_1 的列向量组扩充而成.

分解 $A = QR$ 称为 A 的 **QR 分解**，$A = Q_1R_1$ 称为 A 的**简化 QR 分解**.

例 3.2.5 给定 2 阶正交矩阵 $Q=\begin{bmatrix}\boldsymbol{q}_1 & \boldsymbol{q}_2\end{bmatrix}$，则 $\boldsymbol{q}_1,\boldsymbol{q}_2$ 是 $\mathbb{R}^2$ 的一组标准正交基. 首先，$\boldsymbol{q}_1,\boldsymbol{q}_2$ 是单位向量，不妨设 $\boldsymbol{q}_1=\begin{bmatrix}\cos\theta\\ \sin\theta\end{bmatrix}$, $\boldsymbol{q}_2=\begin{bmatrix}\cos\varphi\\ \sin\varphi\end{bmatrix}$. 其次，$\boldsymbol{q}_2\perp\boldsymbol{q}_1$，因此 $0=\boldsymbol{q}_1^{\mathrm{T}}\boldsymbol{q}_2=\cos\theta\cos\varphi+\sin\theta\sin\varphi=\cos(\theta-\varphi)$，于是 $\theta-\varphi=\left(k+\frac{1}{2}\right)\pi$. 因此，任意二阶正交矩阵必然具有如下两种形式之一：

$$\begin{bmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{bmatrix},\quad\begin{bmatrix}\cos\theta & \sin\theta\\ \sin\theta & -\cos\theta\end{bmatrix}.\tag{3.2.2}$$

回顾例 1.2.9 ，前者是 $\mathbb{R}^2$ 上旋转变换的表示矩阵，后者是 $\mathbb{R}^2$ 上反射变换的表示矩阵. 这说明 $\mathbb{R}^2$ 上的保距变换只有两种，即旋转和反射.

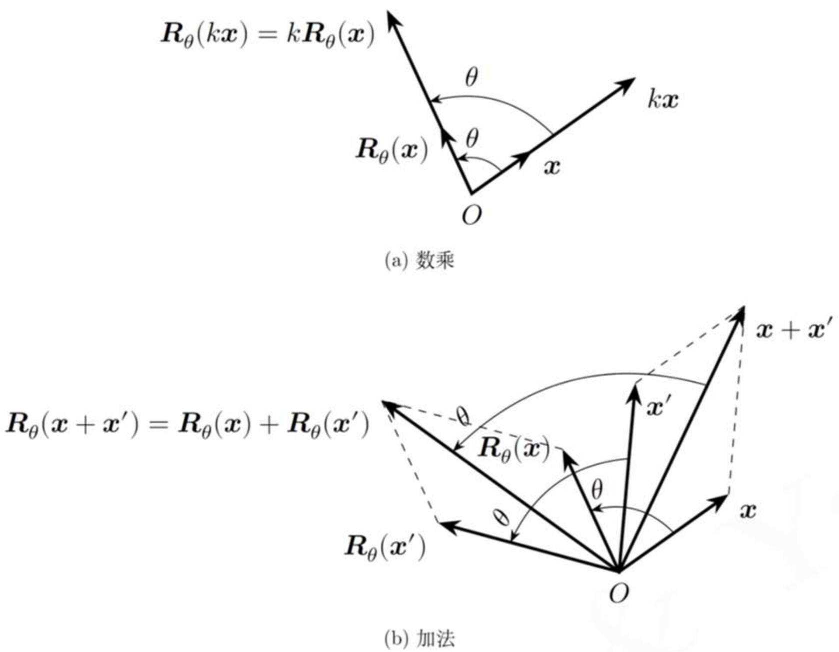


图 1.1.3: 旋转变换

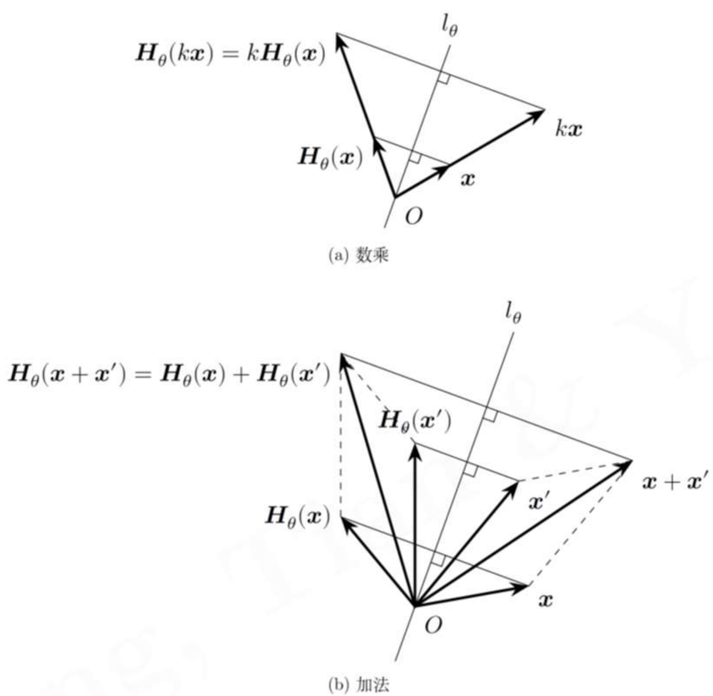


图 1.1.4: 反射变换

反射变换的表示矩阵

两个反射变换的复合是一个旋转变换，转角等于两反射轴夹角的二倍；
反之，任意旋转变换都可以写成两个反射变换的复合.
因此，任意二阶正交矩阵都能写成至多两个反射矩阵的乘积.

考虑反射. 例1.7.8 中反射变换的表示矩阵是 $\boldsymbol{I}_2 - 2\boldsymbol{v}\boldsymbol{v}^{\mathrm{T}}$，其中 $\boldsymbol{v}$ 是反射轴的单位法向量.

$$\begin{bmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{bmatrix} = \boldsymbol{I}_2 - 2 \cdot \begin{bmatrix} \sin\theta \\ -\cos\theta \end{bmatrix} \cdot \begin{bmatrix} \sin\theta & -\cos\theta \end{bmatrix}$$

例 1.7.8 回顾例 1.1.10 中的反射变换. 关于直线 $x_2\cos\theta - x_1\sin\theta = 0$ 的反射变换 H_θ，其表示矩阵为 $\begin{bmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{bmatrix}$. 为写出表示矩阵，我们曾考虑 $\boldsymbol{e}_1, \boldsymbol{e}_2$ 的像，这并不容易. 然而注意到反射变换的几何含义，$H_\theta(\boldsymbol{x})$ 和 $\boldsymbol{x}$ 的关系是二者关于该直线对称，因此 $\boldsymbol{x} - H_\theta(\boldsymbol{x}) = 2\boldsymbol{y}$，其中 $\boldsymbol{y}$ 是 $\boldsymbol{x}$ 在直线的法方向（就是与直线垂直的方向）的投影. 根据中学学习的平面向量的知识，我们知道 $\boldsymbol{y} = (\boldsymbol{v}^{\mathrm{T}}\boldsymbol{x})\boldsymbol{v}$，其中 $\boldsymbol{v}$ 是直线的单位法向量. 我们知道 $\boldsymbol{v} = \begin{bmatrix} \sin\theta \\ -\cos\theta \end{bmatrix}$，因此 $H_\theta(\boldsymbol{x}) = \boldsymbol{x} - 2\boldsymbol{v}\boldsymbol{v}^{\mathrm{T}}\boldsymbol{x} = (I_2 - 2\boldsymbol{v}\boldsymbol{v}^{\mathrm{T}})\boldsymbol{x}$，故反射变换的表示矩阵是 $I_2 - 2\boldsymbol{v}\boldsymbol{v}^{\mathrm{T}}$. 可以验证，这和前面含有三角函数的表示矩阵相等. 这个矩阵的优点在于我们不再需要三角函数的参与，因为法向量和单位法向量容易写出，具体说来，直线 $a_1x_1 + a_2x_2 = 0$ 的法向量是 $\begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$. ☺

Household变换

向量 b 向直线 $\mathrm{span}(a)$ 的正交投影 $\dfrac{a^T b}{a^T a} a$

任意 $\mathbb{R}^n$ 中的单位向量 $\boldsymbol{v}$，都唯一决定以其为法向量的超平面 $\mathcal{N}(\boldsymbol{v}^{\mathrm{T}})$.

令 H_v 为一个线性变换,它的表示矩阵为 $I_n - 2vv^T$,则 H_v 是 R^n 上的反射变换,反射面是以 v 为法向量的超平面 $N\left(v^T\right)$,如图3.2.1所示。事实上，对任意 w,

$$w = \left(I_n - vv^T\right) w + vv^T w$$

$$\left(I_n - 2vv^T\right) w = \left(I_n - vv^T\right) w - vv^T w.$$

$$\frac{v^T w}{v^T v} v = v\left(v^T w\right)$$

$vv^T w = \left(v^T w\right) v$ 是 w 向 $\mathrm{span}(v)$ 的投影,

而 $\left(I_n - vv^T\right) w \perp v$,即 w 可以分解为

与 w 共线的向量和与 w 正交的向量之和,

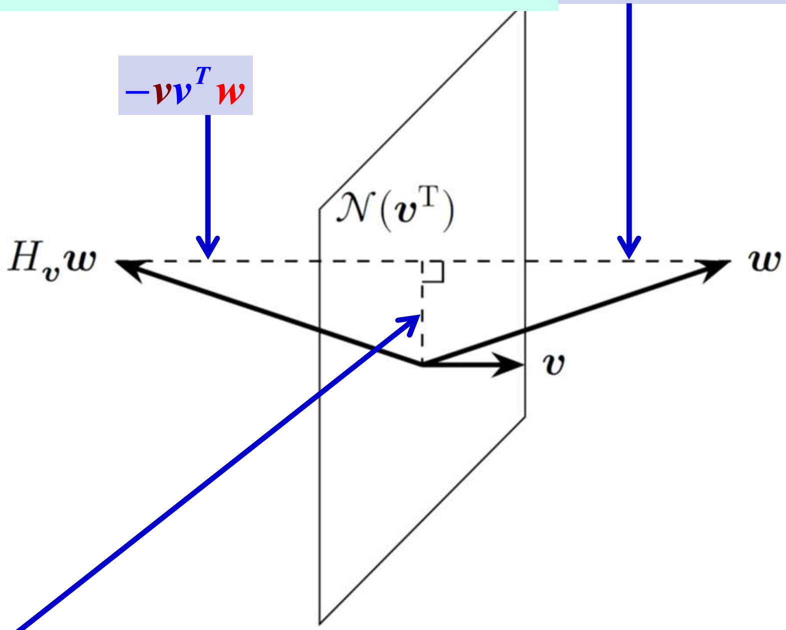
而 $\left(I_n - 2vv^T\right) w$ 是把 w 中的

与 v 共线的成分反向所得到的向量。

可见，此变换描述了反射。

这类形如 H_v 的变换称为Householder变换。

$-vv^T w$



$$w - \frac{v^T w}{v^T v} v = w - vv^T w = \left(I_n - vv^T\right) w$$

图 3.2.1: 高维空间中的反射

命题 3.2.6 给定 $\mathbb{R}^n$ 中向量 $\boldsymbol{x}, \boldsymbol{y}$，满足 $\|\boldsymbol{x}\| = \|\boldsymbol{y}\|$，则存在反射 $\boldsymbol{H_v}$，其中 $\boldsymbol{v} = \dfrac{\boldsymbol{y}-\boldsymbol{x}}{\|\boldsymbol{y}-\boldsymbol{x}\|}$，使得 $\boldsymbol{H_v}(\boldsymbol{x}) = \boldsymbol{y}$.

这说明，对任意等长的向量 $\boldsymbol{x}, \boldsymbol{y}$，都能找到一个反射，使得 $\boldsymbol{H_v}(\boldsymbol{x}) = \boldsymbol{y}$.
容易验证，$G_{\theta;i,j}$ 和 $H_{\boldsymbol{v}}$ 都是正交矩阵，从而是保距变换.

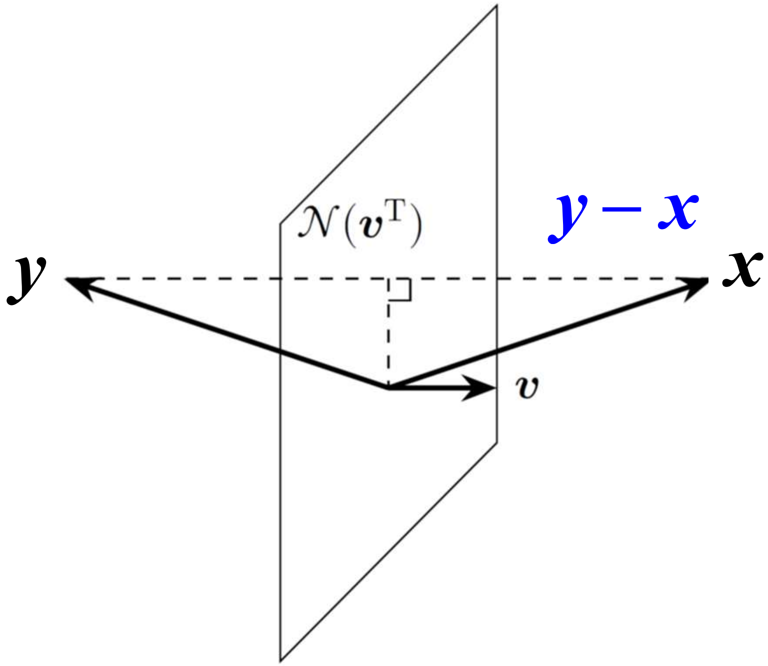


图 3.2.1: 高维空间中的反射

练习3.2.16 证明任意n 阶正交矩阵可以表示成不多于n 个反射的乘积.

练习 2.4.3 求下列矩阵零空间的一组基.

1. $\begin{bmatrix} I_n & I_n \end{bmatrix}$.
2. $\begin{bmatrix} I_n & I_n \\ O & O \end{bmatrix}$.
3. $\begin{bmatrix} I_n & I_n & I_n \end{bmatrix}$.

练习 2.4.11 在平面直角坐标系下给定点 $A(a_1, a_2), B(b_1, b_2), C(c_1, c_2)$，证明，$A, B, C$ 三点不共线当且仅当矩阵 $\begin{bmatrix} a_1 & a_2 & 1 \\ b_1 & b_2 & 1 \\ c_1 & c_2 & 1 \end{bmatrix}$ 可逆.

练习 2.4.19 对 n 阶方阵 A，求证:

1. $A^2 = A$ 当且仅当 $\operatorname{rank}(A) + \operatorname{rank}(I_n - A) = n$.
   提示：考虑 A 和 $I - A$ 的相关的子空间有何关系？
2. $A^2 = I_n$ 当且仅当 $\operatorname{rank}(I_n + A) + \operatorname{rank}(I_n - A) = n$.
   提示：考虑 $I + A$ 和 $I - A$ 的相关的子空间有何关系？

例：用LU三角分解法解

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 2 \\ 3 & 1 & 5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 14 \\ 18 \\ 20 \end{bmatrix}$$

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} 1 & & \\ l_{21} & 1 & \\ l_{31} & l_{32} & 1 \end{bmatrix}\begin{bmatrix} u_{11} & u_{12} & u_{13} \\ & u_{22} & u_{23} \\ & & u_{33} \end{bmatrix} =$$

$$\begin{bmatrix} u_{11} & u_{12} & u_{13} \\ l_{21}u_{11} & l_{21}u_{12} + u_{22} & l_{21}u_{13} + u_{23} \\ l_{31}u_{11} & l_{31}u_{12} + l_{32}u_{22} & l_{31}u_{13} + l_{32}u_{23} + u_{33} \end{bmatrix}$$

解：用分解计算得

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & -5 & 1 \end{bmatrix}\begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & -4 \\ 0 & 0 & -24 \end{bmatrix} = LU.$$

求解 $Ly = \begin{bmatrix} 14 \\ 18 \\ 20 \end{bmatrix} \Rightarrow y = \begin{bmatrix} 14 \\ -10 \\ -72 \end{bmatrix}, \quad Ux = \begin{bmatrix} 14 \\ -10 \\ -72 \end{bmatrix} \Rightarrow x = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}.$

作 业（10月27日）

~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~

练习3.2

1，5，8，10，12，18

11月1日提交

~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~