§15 Gram-Schmidt正交化

15.1 引言

设A是$m \times n$ 阶阵，若 $A\mathbf{x} = \mathbf{b}, \mathbf{b} \in \mathbb{R}^m$ 无解，则考虑法方程组

$A^T A\hat{\mathbf{x}} = A^T\mathbf{b}.$

$\mathbf{p} = A\hat{\mathbf{x}}$ 是 $\mathbf{b}$ 在 $C(A)$ 上的投影.

设 $C(A) = C(A')$, 则 $\mathbf{p} = A'\hat{\mathbf{y}}, \hat{\mathbf{y}}$ 是 $A'^T A'\hat{\mathbf{y}} = A'^T\mathbf{b}$ 的解.

因此若 $A'^T A'$ 较简单，则 $\mathbf{p}$ 容易计算.

15.1 引言

例：求 $\mathbf{b}=\begin{pmatrix}-4\\-3\\3\\0\end{pmatrix}$ 在 $A=\begin{pmatrix}1&-2\\1&0\\1&1\\1&3\end{pmatrix}$ 的列空间上的投影.

解： $A^TA=\begin{pmatrix}4&2\\2&14\end{pmatrix}$ $\quad A^TA\hat{\mathbf{x}}=A^T\mathbf{b}\quad\Longrightarrow\mathbf{p}=A\hat{\mathbf{x}}=\dfrac{1}{2}\begin{pmatrix}-7\\-3\\-1\\3\end{pmatrix}$

$\Uparrow$

考虑 $A'=\begin{pmatrix}1&-5\\1&-1\\1&1\\1&5\end{pmatrix}$ $\quad A'^TA'=\begin{pmatrix}4&0\\0&52\end{pmatrix}$ $\quad A'^TA'\hat{\mathbf{y}}=A'^T\mathbf{b}$

A'^TA' 简单，因为 A' 的列相互正交.

$C(A')=C(A)\Longrightarrow$ 将一组基(A 的无关列)换成一组正交的向量 (A' 正交列).

15.1 引言

目标：给定 $V \subset \mathbb{R}^n$ 为一个子空间，$\mathbf{v}_1, \cdots, \mathbf{v}_k$ 是 V 的一组基，把它们变化成一组正交的向量 $\mathbf{w}_1, \cdots, \mathbf{w}_k$ 满足

1. $\mathbf{w}_i^T \mathbf{w}_j = 0,\ i \neq j.$
2. $L(\mathbf{v}_1, \cdots, \mathbf{v}_t) = L(\mathbf{w}_1, \cdots, \mathbf{w}_t),\ 1 \leq t \leq k.$

   $L(\mathbf{v}_1, \cdots, \mathbf{v}_t)$表示 $\mathbf{v}_1, \cdots, \mathbf{v}_t$ 生成的 V 的子空间.

$$L(\mathbf{v}_1, \cdots, \mathbf{v}_t) = \{a_1\mathbf{v}_1 + \cdots + a_t\mathbf{v}_t | a_i \in \mathbb{R}\}.$$

15.2 正交向量组和正交矩阵

定理：设 $\mathbf{v}_1, \cdots, \mathbf{v}_k \in \mathbb{R}^n$ 是非零的 k 个向量，满足 $\mathbf{v}_i^T \mathbf{v}_j = 0,\ i \neq j,$ 则 $\mathbf{v}_1, \cdots, \mathbf{v}_k$ 线性无关.

证明：设 $a_1 \mathbf{v}_1 + \cdots + a_k \mathbf{v}_k = \mathbf{0}, a_i \in \mathbb{R}, 1 \leq i \leq k.$
两边左乘 $\mathbf{v}_1^T$, 则 $a_1 ||\mathbf{v}_1||^2 = 0. \implies a_1 = 0.$
同理 $a_2 = \cdots = a_k = 0.$ 因此 $\mathbf{v}_1, \cdots, \mathbf{v}_k$ 线性无关.

例如: $\mathbb{R}^2$ 中两向量 $(\cos\theta, \sin\theta)^T, (-\sin\theta, \cos\theta)^T$相互正交, 故无关.

定理中 $\mathbf{v}_1, \cdots, \mathbf{v}_k$ 称为正交向量组(orthogonal vectors).

15.2 正交向量组和正交矩阵

定义：设 $\mathbf{q}_1, \cdots, \mathbf{q}_n$ 是 n 个列向量，它们是标准正交的(orthonormal)

$$\Longleftrightarrow \mathbf{q}_i^T \mathbf{q}_j = \begin{cases} 0, i \neq j \\ 1, i = j \end{cases}, \forall i, j = 1, \cdots, n.$$

令 $Q = (\mathbf{q}_1, \cdots, \mathbf{q}_n)$，则 $Q^T Q = \begin{pmatrix} \mathbf{q}_1^T \\ \vdots \\ \mathbf{q}_n^T \end{pmatrix} \begin{pmatrix} \mathbf{q}_1 & \cdots & \mathbf{q}_n \end{pmatrix}$

$$= \begin{pmatrix} \mathbf{q}_1^T \mathbf{q}_1 & \cdots & \mathbf{q}_1^T \mathbf{q}_n \\ \vdots & \ddots & \vdots \\ \mathbf{q}_n^T \mathbf{q}_1 & \cdots & \mathbf{q}_n^T \mathbf{q}_n \end{pmatrix} = \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix} = I_n$$

若 Q 是一个方阵，则 $Q^{-1} = Q^T, Q$ 称为正交阵(orthogonal matrix).

15.2 正交向量组和正交矩阵

例：$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$, $\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$.

例：设 $\mathbf{u}$ 是一列向量，$\mathbf{u}^T\mathbf{u} = 1$ ($\mathbf{u}$是单位向量unit vector).

令 $Q = I_n - 2\mathbf{u}\mathbf{u}^T, \mathbf{u} \in \mathbb{R}^n$.

Q 是一个反射矩阵(reflection matrix).

$Q\mathbf{u} = \mathbf{u} - 2\mathbf{u}\mathbf{u}^T\mathbf{u} = -\mathbf{u}$.

若 $\mathbf{v} \perp \mathbf{u}$, 则 $Q\mathbf{v} = \mathbf{v}$.

15.2 正交向量组和正交矩阵

比如，$\mathbf{u}=\begin{pmatrix}0\\0\\1\end{pmatrix}$，则 $Q=\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}-\begin{pmatrix}0&0&0\\0&0&0\\0&0&2\end{pmatrix}=\begin{pmatrix}1&0&0\\0&1&0\\0&0&-1\end{pmatrix}$.

考虑映射 $\mathbb{R}^3\xrightarrow{f}\mathbb{R}^3$

$$\begin{pmatrix}x\\y\\z\end{pmatrix}\longmapsto Q\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}x\\y\\-z\end{pmatrix}$$

f 是关于 $x0y$ 平面的反射变换.

15.2 正交向量组和正交矩阵

注：以上两例均是保长度的变换，即

$$f:\mathbb{R}^3 \longrightarrow \mathbb{R}^3, \forall \mathbf{v} \in \mathbb{R}^3, ||\mathbf{v}|| = ||f(\mathbf{v})||.$$

以后我们将说明具有这种性质的变换对应于正交矩阵.

定理：设 Q 是一个正交阵，则 $\forall \mathbf{x} \in \mathbb{R}^n, ||Q\mathbf{x}|| = ||\mathbf{x}||.$

证明: $||Q\mathbf{x}||^2 = (Q\mathbf{x})^T(Q\mathbf{x}) = \mathbf{x}^T Q^T Q\mathbf{x} = \mathbf{x}^T\mathbf{x} = ||\mathbf{x}||^2.$

15.3 Gram-Schmidt正交化过程

以下考虑本讲的目标问题.

先考虑两个向量 $\mathbf{v}_1, \mathbf{v}_2$ 线性无关，求 $\mathbf{w}_1, \mathbf{w}_2$ 满足

$$\mathbf{w}_1^T \mathbf{w}_2 = 0, L(\mathbf{w}_1) = L(\mathbf{v}_1), L(\mathbf{w}_1, \mathbf{w}_2) = L(\mathbf{v}_1, \mathbf{v}_2).$$

进一步令 $\mathbf{q}_1 = \frac{\mathbf{w}_1}{||\mathbf{w}_1||}, \mathbf{q}_2 = \frac{\mathbf{w}_2}{||\mathbf{w}_2||}$, 即 $\mathbf{q}_1, \mathbf{q}_2$ 是标准正交的.

即

$$\{\mathbf{v}_1, \mathbf{v}_2\} \longrightarrow \{\mathbf{w}_1, \mathbf{w}_2\} \longrightarrow \{\mathbf{q}_1, \mathbf{q}_2\}$$

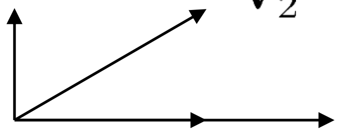
$$L(\mathbf{v}_1, \mathbf{v}_2) \ = L(\mathbf{w}_1, \mathbf{w}_2) \ = L(\mathbf{q}_1, \mathbf{q}_2)$$

15.3 Gram-Schmidt正交化过程

显然 $\mathbf{w}_1 = \mathbf{v}_1, \mathbf{w}_2 = x_1\mathbf{v}_1 + x_2\mathbf{v}_2$.

可以取 $x_2 = 1$, 因为只需 $\mathbf{w}_2 \perp \mathbf{w}_1$, 这样的 $\mathbf{w}_2$ 不唯一.

$\mathbf{w}_1^T\mathbf{w}_2 = 0 \implies \mathbf{w}_2 = \mathbf{v}_2 - \dfrac{\mathbf{v}_1^T\mathbf{v}_2}{\mathbf{v}_1^T\mathbf{v}_1}\mathbf{v}_1$, 即 $\mathbf{v}_2$ 减去它在 $\mathbf{v}_1$ 上的投影.



15.3 Gram-Schmidt正交化过程

设 $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ 线性无关，考虑

$$\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\} \xrightarrow{\text{正交化}} \{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\} \xrightarrow{\text{单位化}} \{\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3\}$$

满足 $L(\mathbf{w}_1) = L(\mathbf{v}_1), L(\mathbf{w}_1, \mathbf{w}_2) = L(\mathbf{v}_1, \mathbf{v}_2),$

$L(\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3) = L(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3).$

15.3 Gram-Schmidt正交化过程

已知：$\mathbf{w}_1 = \mathbf{v}_1, \mathbf{w}_2 = \mathbf{v}_2 - \dfrac{\mathbf{v}_1^T \mathbf{v}_2}{\mathbf{v}_1^T \mathbf{v}_1} \mathbf{v}_1.$

设 $\mathbf{w}_3 = x_1 \mathbf{w}_1 + x_2 \mathbf{w}_2 + \mathbf{v}_3$ $\left\{ \begin{array}{l} \mathbf{w}_1^T \mathbf{w}_3 = 0 \implies x_1 = -\dfrac{\mathbf{w}_1^T \mathbf{v}_3}{\mathbf{w}_1^T \mathbf{w}_1} \\ \mathbf{w}_2^T \mathbf{w}_3 = 0 \implies x_2 = -\dfrac{\mathbf{w}_2^T \mathbf{v}_3}{\mathbf{w}_2^T \mathbf{w}_2} \end{array} \right.$

单位化 $\mathbf{q}_1 = \dfrac{\mathbf{w}_1}{||\mathbf{w}_1||}, \mathbf{q}_2 = \dfrac{\mathbf{w}_2}{||\mathbf{w}_2||}, \mathbf{q}_3 = \dfrac{\mathbf{w}_3}{||\mathbf{w}_3||}.$

15.3 Gram-Schmidt正交化过程

对一般情形，首先我们有如下定理：

定理：设 $\alpha_1, \cdots, \alpha_k$ 相互正交，$\mathbf{v} \in L(\alpha_1, \cdots, \alpha_k)$. 则

$$\mathbf{v} = \frac{\alpha_1^T \mathbf{v}}{\alpha_1^T \alpha_1}\alpha_1 + \cdots + \frac{\alpha_k^T \mathbf{v}}{\alpha_k^T \alpha_k}\alpha_k.$$

特别地，若 $\alpha_1, \cdots, \alpha_k$ 标准正交，则

$$\mathbf{v} = (\alpha_1^T \mathbf{v})\alpha_1 + \cdots + (\alpha_k^T \mathbf{v})\alpha_k.$$

15.3 Gram-Schmidt正交化过程

由此定理，设

$$\{\mathbf{v}_1, \cdots, \mathbf{v}_k\} \xrightarrow{\text{正交化}} \{\mathbf{w}_1, \cdots, \mathbf{w}_k\} \xrightarrow{\text{单位化}} \{\mathbf{q}_1, \cdots, \mathbf{q}_k\}$$

且

$$\forall 1 \leq l \leq k, L(\mathbf{v}_1, \cdots, \mathbf{v}_l) = L(\mathbf{q}_1, \cdots, \mathbf{q}_l),$$

则

$$\mathbf{v}_l = (\mathbf{q}_1^T \mathbf{v}_l)\mathbf{q}_1 + \cdots + (\mathbf{q}_l^T \mathbf{v}_l)\mathbf{q}_l.$$

15.3 Gram-Schmidt正交化过程

这给出了另一种方法求 $\mathbf{q}_1, \cdots, \mathbf{q}_k$.

$\mathbf{e}_l$ 为误差向量

$$\mathbf{e}_1 = \mathbf{v}_1 = \mathbf{w}_1, \quad \mathbf{q}_1 = \frac{\mathbf{v}_1}{||\mathbf{v}_1||}$$

$$\mathbf{e}_2 = \mathbf{v}_2 - (\mathbf{q}_1^T\mathbf{v}_2)\mathbf{q}_1 = \mathbf{w}_2, \quad \mathbf{q}_2 = \frac{\mathbf{w}_2}{||\mathbf{w}_2||} = \frac{\mathbf{e}_2}{||\mathbf{e}_2||}$$

$$\mathbf{e}_k = \mathbf{v}_k - (\mathbf{q}_1^T\mathbf{v}_k)\mathbf{q}_1 - \cdots - (\mathbf{q}_{k-1}^T\mathbf{v}_k)\mathbf{q}_{k-1} = \mathbf{w}_k, \quad \mathbf{q}_k = \frac{\mathbf{w}_k}{||\mathbf{w}_k||} = \frac{\mathbf{e}_k}{||\mathbf{e}_k||}$$

这种正交化方法，称为Gram-Schmidt正交化.

15.3 Gram-Schmidt正交化过程

例： $A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$. $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \mathbf{v}_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \mathbf{v}_3 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$.

解： $\mathbf{w}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{q}_1 = \frac{\mathbf{w}_1}{||\mathbf{w}_1||} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}.$

$$\mathbf{w}_2 = \mathbf{v}_2 - (\mathbf{q}_1^T \mathbf{v}_2)\mathbf{q}_1 = \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ 1 \end{pmatrix}, \quad \mathbf{q}_2 = \frac{\mathbf{w}_2}{||\mathbf{w}_2||} = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}.$$

$$\mathbf{w}_3 = \mathbf{v}_3 - (\mathbf{q}_1^T \mathbf{v}_3)\mathbf{q}_1 - (\mathbf{q}_2^T \mathbf{v}_3)\mathbf{q}_2, \quad \mathbf{q}_3 = \frac{\mathbf{w}_3}{||\mathbf{w}_3||} = \frac{1}{\sqrt{3}} \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}.$$

则 $Q = (\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3)$ 是正交阵.

15.4 QR分解

问题：A 和 Q 的关系？

$$A \xrightarrow{\text{行消去}} U, \text{ 则 } A = LU.$$

$$A \underset{\text{(列变换)}}{\xrightarrow{\text{正交化}}} Q, \text{ 则 } A = QR = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{pmatrix} \begin{pmatrix} \frac{2}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & \frac{3}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ 0 & 0 & \frac{2}{\sqrt{3}} \end{pmatrix}.$$

对照正交化公式，$A = QR$, Q 是列正交阵，R 是对角线上为正数的上三角阵，其第 i 个主对角线元素为

$$||\mathbf{w}_i|| = \mathbf{q}_i^T \mathbf{v}_i.$$

15.4 QR分解

应用：

1.设 A 为 $m\times n$ 阶列满秩阵，A 的列线性无关，$A=Q_{m\times n}R_{n\times n}$.

$$A\mathbf{x}=\mathbf{b}\ \Longrightarrow A^TA\hat{\mathbf{x}}=A^T\mathbf{b}\Longleftrightarrow R^TQ^TQR\hat{\mathbf{x}}=R^TQ^T\mathbf{b}$$
$$\Longleftrightarrow R^TR\hat{\mathbf{x}}=R^TQ^T\mathbf{b}\Longleftrightarrow R\hat{\mathbf{x}}=Q^T\mathbf{b}\Longleftrightarrow \hat{\mathbf{x}}=R^{-1}Q^T\mathbf{b}$$

特别地，若 A 的列相互正交，$A=(\alpha_1,\cdots,\alpha_n)$, 则

$$R=diag(||\alpha_1||,\cdots,||\alpha_n||)\Longrightarrow \hat{\mathbf{x}}=(R^{-1})^2A^T\mathbf{b}$$

设 $A\mathbf{x}=\mathbf{b}$ 无解，则 $\mathbf{b}$ 在 $C(A)$ 上的投影为

$$\mathbf{p}=\sum_{i=1}^{n}(\frac{\alpha_i^T\mathbf{b}}{\alpha_i^T\alpha_i})\alpha_i.$$

15.4 QR分解

2. 设 A是可逆方阵， 则 QR 分解是唯一的.

设 $A = Q_1R_1 = Q_2R_2$ 为可逆方阵 A 的两个 QR 分解.

则 $Q_2^{-1}Q_1 = R_2R_1^{-1}$.

$Q_2^{-1}Q_1$ 为正交阵，$R_2R_1^{-1}$ 为上三角阵且对角元素为正.

故 $Q_2^{-1}Q_1 = R_2R_1^{-1} = I_n$.

15.4 QR分解

3.设 $A_{m\times n}$ 列满秩，有 QR 分解 $A=QR.$

$\mathbf{b}\notin C(A)$, 设 $\mathbf{b}$ 在 $C(A)$ 上投影为 $\mathbf{p}, \mathbf{e}=\mathbf{b}-\mathbf{p}.$

则 $(A,\mathbf{b})$ 也是列满秩阵，其 QR 分解如下

$$(A,\mathbf{b})=(Q,\frac{\mathbf{e}}{||\mathbf{e}||})\begin{pmatrix} R & \alpha \\ 0 & ||\mathbf{e}|| \end{pmatrix}, \alpha=Q^T\mathbf{b}.$$