1 正交性和最小二乘法

内积、长度和正交性

如果 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 是 $\mathbb{R}^n$ 中的向量，则可以将 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 作为 $n\times 1$ 矩阵. 转置矩阵 $\boldsymbol{u}^T$ 是 $1\times n$ 矩阵，且矩阵乘积 $\boldsymbol{u}^T\boldsymbol{v}$ 是一个 1×1 矩阵，我们将其记为一个不加括号的实数（标量）. 数 $\boldsymbol{u}^T\boldsymbol{v}$ 称为 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 的**内积**，通常记作 $\boldsymbol{u}\cdot\boldsymbol{v}$，也称为**点积**

定理 1 设 $\boldsymbol{v}$,$\boldsymbol{u}$ 和 $\boldsymbol{w}$ 是 $\mathbb{R}^n$ 中的向量,c 是一个数，那么

a. $\boldsymbol{u}\cdot\boldsymbol{v}=\boldsymbol{v}\cdot\boldsymbol{u}$

b. $(\boldsymbol{u}+\boldsymbol{v})\cdot\boldsymbol{w}=\boldsymbol{u}\cdot\boldsymbol{w}+\boldsymbol{v}\cdot\boldsymbol{w}$

c. $(c\boldsymbol{u})\cdot v=c(\boldsymbol{u}\cdot\boldsymbol{v})=\boldsymbol{u}\cdot(c\boldsymbol{v})$

d. $\boldsymbol{u}\cdot\boldsymbol{u}\geqslant 0$, 并且 $\boldsymbol{u}\cdot\boldsymbol{u}=0$ 成立的充分必要条件是 $\boldsymbol{u}=\boldsymbol{0}$

定义 向量 $\boldsymbol{v}$ 的**长度**（或**范数**）是非负数 $||\boldsymbol{v}||$, 定义为

$$||\boldsymbol{v}||=\sqrt{\boldsymbol{v}\cdot\boldsymbol{v}}=\sqrt{v_1^2+v_2^2+\cdots+v_n^2}\ \text{且}\ ||\boldsymbol{v}||^2=\boldsymbol{v}\cdot\boldsymbol{v}$$

长度为 1 的向量称为**单位向量**. 如果把一个非零向量除以其自身的长度，即乘 $\dfrac{1}{||\boldsymbol{v}||}$，就可以得到一个单位向量，即 $\boldsymbol{u}=\dfrac{\boldsymbol{v}}{||\boldsymbol{v}||}$

把向量 $\boldsymbol{v}$ 化成单位向量 $\boldsymbol{u}$ 的过程，称为向量 $\boldsymbol{v}$ 的**单位化**

定义 如果 $\boldsymbol{u}\cdot\boldsymbol{v}=0$，则 $\mathbb{R}^n$ 中的两个向量 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 是（相互）**正交的**

定理 2 (毕达哥拉斯（勾股）定理)
两个向量 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 正交的充分必要条件是 $||\boldsymbol{u}+\boldsymbol{v}||^2=||\boldsymbol{u}||^2+||\boldsymbol{v}||^2$

如果向量 $\boldsymbol{z}$ 与 $\mathbb{R}^n$ 的子空间 W 中的任意向量都正交，则称 $\boldsymbol{z}$ **正交于 W**

与子空间 W 正交的向量 $\boldsymbol{z}$ 的全体组成的集合称为 W 的**正交补**，记作 $W^{\perp}$

1. 向量 $\boldsymbol{x}$ 属于 $W^{\perp}$ 的充分必要条件是向量 $\boldsymbol{x}$ 与生成空间 W 的任一向量都正交
2.$W^{\perp}$ 是 $\mathbb{R}^n$ 的一个子空间

定理 3 *假设 A 是 $m\times n$ 矩阵，那么 A 的行空间的正交补是 A 的零空间，且 A 的列空间的正交补是 A^T 的零空间:*

$$(\mathrm{Row}\,A)^{\perp}=\mathrm{Nul}\,A \text{ 且 } (\mathrm{Col}\,A)^{\perp}=\mathrm{Nul}\,A^T$$

正交集

$\mathbb{R}^n$ 中的向量集合 $\{\boldsymbol{u}_1,\cdots,\boldsymbol{u}_p\}$ 称为**正交集**，如果集合中的任意两个不同向量都正交，即当 $i\neq j$ 时，$\boldsymbol{u_i}\cdot\boldsymbol{u_j}=0$

定理 4 *如果 $S=\{\boldsymbol{u}_1,\cdots,\boldsymbol{u}_p\}$ 是由 $\mathbb{R}^n$ 中非零向量构成的正交集，那么 S 是线性无关集，因此构成 S 所生成的子空间的一组基*

定义　$\mathbb{R}^n$ 中子空间 W 的一个**正交基**是 W 的一个基，也是正交集

定理 5　假设 $\{\boldsymbol{u}_1, \cdots, \boldsymbol{u}_p\}$ 是 $\mathbb{R}^n$ 中子空间 W 的正交基，对 W 中的每个向量 $\boldsymbol{y}$，线性组合 $\boldsymbol{y} = c_1\boldsymbol{u}_1 + \cdots + c_p\boldsymbol{u}_p$ 中的权可以由 $c_j = \dfrac{\boldsymbol{y} \cdot \boldsymbol{u_j}}{\boldsymbol{u_j} \cdot \boldsymbol{u_j}}$（$j = 1, \cdots, p$）计算

对 $\mathbb{R}^n$ 中给出的非零向量 $\boldsymbol{u}$，考虑 $\mathbb{R}^n$ 中一个向量 $\boldsymbol{y}$ 分解为两个向量之和的问题，一个向量是向量 $\boldsymbol{u}$ 的倍数，另一个向量与 $\boldsymbol{u}$ 正交. 我们期望写成

$$\boldsymbol{y} = \hat{\boldsymbol{y}} + \boldsymbol{z} \tag{1.1}$$

其中 $\hat{\boldsymbol{y}} = \alpha\boldsymbol{u}$，$\alpha$ 是一个数，$\boldsymbol{z}$ 是一个垂直于 $\boldsymbol{u}$ 的向量. 对给定数 α，记 $\boldsymbol{z} = \boldsymbol{y} - \alpha\boldsymbol{u}$，则方程(1.1)可以满足. 那么 $\boldsymbol{y} - \hat{\boldsymbol{y}}$ 和 $\boldsymbol{u}$ 正交的充分必要条件是

$$(\boldsymbol{y} - \alpha\boldsymbol{u}) \cdot \boldsymbol{u} = \boldsymbol{y} \cdot \boldsymbol{u} - (\alpha\boldsymbol{u}) \cdot \boldsymbol{u} = \boldsymbol{y} \cdot \boldsymbol{u} - \alpha(\boldsymbol{u} \cdot \boldsymbol{u}) = 0$$

也就是满足方程(1.1)且 $\boldsymbol{z}$ 与 $\boldsymbol{u}$ 正交的充分必要条件是 $\alpha = \dfrac{\boldsymbol{y} \cdot \boldsymbol{u}}{\boldsymbol{u} \cdot \boldsymbol{u}}$ 和 $\hat{\boldsymbol{y}} = \operatorname{proj}_L \boldsymbol{y} = \dfrac{\boldsymbol{y} \cdot \boldsymbol{u}}{\boldsymbol{u} \cdot \boldsymbol{u}} \cdot \boldsymbol{u}$. 向量 $\hat{\boldsymbol{y}}$ 称为 $\boldsymbol{y}$ **在** $\boldsymbol{u}$ **上的正交投影**，向量 $\boldsymbol{z}$ 称为 $\boldsymbol{y}$ **与** $\boldsymbol{u}$ **正交的分量**

如果集合 $\{\boldsymbol{u}_1, \cdots, \boldsymbol{u}_p\}$ 是由单位向量构成的正交集，那么它是一个**单位正交集**

如果 W 是一个由单位正交集合生成的子空间，那么 $\{\boldsymbol{u}_1, \cdots, \boldsymbol{u}_p\}$ 是 W 的**单位正交基**

定理 6　一个 $m \times n$ 矩阵 U 具有单位正交列向量的充分必要条件是 $U^TU = I$

定理 7 假设 U 是一个具有单位正交列的 $m \times n$ 矩阵，且 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 是 $\mathbb{R}^n$ 中的向量，那么

a. $||U\boldsymbol{x}|| = ||\boldsymbol{x}||$

b. $(U\boldsymbol{x}) \cdot (U\boldsymbol{y}) = \boldsymbol{x} \cdot \boldsymbol{y}$

c. $(U\boldsymbol{x}) \cdot (U\boldsymbol{y}) = 0$ 的充分必要条件是 $\boldsymbol{x} \cdot \boldsymbol{y} = 0$

正交投影

对给定向量 $\boldsymbol{y}$ 和 $\mathbb{R}^n$ 中子空间 W，存在属于 W 的向量 $\hat{\boldsymbol{y}}$ 满足:

(1) W 中有唯一向量 $\hat{\boldsymbol{y}}$，使得 $\boldsymbol{y} - \hat{\boldsymbol{y}}$ 与 W 正交

(2) $\hat{\boldsymbol{y}}$ 是 W 中唯一最接近 $\boldsymbol{y}$ 的向量

定理 8 (正交分解定理)

若 W 是 $\mathbb{R}^n$ 的一个子空间，那么 $\mathbb{R}^n$ 中每一个向量 $\boldsymbol{y}$ 可以唯一表示为

$$\boldsymbol{y} = \hat{\boldsymbol{y}} + \boldsymbol{z}$$

其中 $\hat{\boldsymbol{y}}$ 属于 W 而 $\boldsymbol{z}$ 属于 $W^{\perp}$. 实际上，如果 $\{\boldsymbol{u}_1, \cdots, \boldsymbol{u}_p\}$ 是 W 的任意正交基，那么

$$\hat{\boldsymbol{y}} = \frac{\boldsymbol{y} \cdot \boldsymbol{u_1}}{\boldsymbol{u_1} \cdot \boldsymbol{u_1}} \boldsymbol{u}_1 + \cdots + \frac{\boldsymbol{y} \cdot \boldsymbol{u_p}}{\boldsymbol{u_p} \cdot \boldsymbol{u_p}} \boldsymbol{u}_p$$

且 $\boldsymbol{z} = \boldsymbol{y} - \hat{\boldsymbol{y}}$

定理 9 (最佳逼近定理)

假设 W 是 $\mathbb{R}^n$ 的一个子空间，$\boldsymbol{y}$ 是 $\mathbb{R}^n$ 中的任意向量，$\hat{\boldsymbol{y}}$ 是 $\boldsymbol{y}$ 在 W 上的正交投影，那么 $\hat{\boldsymbol{y}}$ 是 W 中最接近 $\boldsymbol{y}$ 的点，也就是

$$||\boldsymbol{y} - \hat{\boldsymbol{y}}|| < ||\boldsymbol{y} - \boldsymbol{v}||$$

对所有属于 W 又异于 $\hat{\boldsymbol{y}}$ 的 $\boldsymbol{v}$ 成立

定理 10 如果 $\{\boldsymbol{u}_1, \cdots, \boldsymbol{u}_p\}$ 是 $\mathbb{R}^n$ 中子空间 W 的单位正交基，那么

$$\operatorname{proj}_W \boldsymbol{y} = (\boldsymbol{y} \cdot \boldsymbol{u}_1)\boldsymbol{u}_1 + (\boldsymbol{y} \cdot \boldsymbol{u}_2)\boldsymbol{u}_2 + \cdots + (\boldsymbol{y} \cdot \boldsymbol{u}_p)\boldsymbol{u}_p$$

如果 $U = [\boldsymbol{u}_1 \quad \boldsymbol{u}_2 \quad \cdots \quad \boldsymbol{u}_p]$，则

$$\operatorname{proj}_W \boldsymbol{y} = UU^T\boldsymbol{y}, \text{ 对所有} \boldsymbol{y} \in \mathbb{R}^n \text{成立}$$

格拉姆-施密特方法

定理 11 (格拉姆-施密特方法)
对 $\mathbb{R}^n$ 的子空间 W 的一个基 $\{\boldsymbol{x}_1, \cdots, \boldsymbol{x}_p\}$，定义

$$\begin{aligned}
\boldsymbol{v}_1 &= \boldsymbol{x}_1 \\
\boldsymbol{v}_2 &= \boldsymbol{x}_2 - \frac{\boldsymbol{x}_2 \cdot \boldsymbol{v}_1}{\boldsymbol{v}_1 \cdot \boldsymbol{v}_1}\boldsymbol{v}_1 \\
\boldsymbol{v}_3 &= \boldsymbol{x}_3 - \frac{\boldsymbol{x}_3 \cdot \boldsymbol{v}_2}{\boldsymbol{v}_2 \cdot \boldsymbol{v}_2}\boldsymbol{v}_2 - \frac{\boldsymbol{x}_3 \cdot \boldsymbol{v}_1}{\boldsymbol{v}_1 \cdot \boldsymbol{v}_1}\boldsymbol{v}_1 \\
&\vdots \\
\boldsymbol{v}_p &= \boldsymbol{x}_p - \frac{\boldsymbol{x}_p \cdot \boldsymbol{v}_{p-1}}{\boldsymbol{v}_{p-1} \cdot \boldsymbol{v}_{p-1}}\boldsymbol{v}_{p-1} - \cdots - \frac{\boldsymbol{x}_p \cdot \boldsymbol{v}_2}{\boldsymbol{v}_2 \cdot \boldsymbol{v}_2}\boldsymbol{v}_2 - \frac{\boldsymbol{x}_p \cdot \boldsymbol{v}_1}{\boldsymbol{v}_1 \cdot \boldsymbol{v}_1}\boldsymbol{v}_1
\end{aligned}$$

那么 $\{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_p\}$ 是 W 的一个正交基. 此外，

$$\operatorname{Span}\{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_k\} = \operatorname{Span}\{\boldsymbol{x}_1, \cdots, \boldsymbol{x}_k\}, \text{ 其中} 1 \leqslant k \leqslant p$$

定理 12 (**QR** 分解)
如果 $m \times n$ 矩阵 A 的列线性无关，那么 A 可以分解为 A=QR，其中 Q 是一个 $m \times n$ 矩阵，其列形成 $\operatorname{Col} A$ 的一个标准正交基，R 是一个 $n \times n$ 上三角可逆矩阵且在对角线上的元素为正数

最小二乘问题