

MIT 18.06 线性代数

第十六讲：投影矩阵和最小二乘法

整理自 MIT 18.06 课程及相关公开笔记

引言

本讲在上一节投影概念的基础上，进一步深入研究投影矩阵的性质及其在最小二乘法（Least Squares）中的应用。最小二乘法的本质即将目标向量投影到矩阵的列空间中寻找最优解。

1 投影矩阵

1.1 投影矩阵 P

回顾上一节推出的投影矩阵公式：

$$P = A(A^T A)^{-1} A^T$$

对于任意向量 $\mathbf{b}$ ，投影矩阵 P 的作用是将 $\mathbf{b}$ 投影到矩阵 A 的列空间 $C(A)$ 中，得到投影向量 $\mathbf{p} = P\mathbf{b}$ 。如果向量 $\mathbf{b}$ 本身就在矩阵 A 的列空间内，即存在向量 $\mathbf{x}$ 使得 $A\mathbf{x} = \mathbf{b}$ ，则对其进行投影：

$$P\mathbf{b} = A(A^T A)^{-1} A^T A\mathbf{x} = A\mathbf{x} = \mathbf{b}$$

这表明，当向量已经位于列空间时，投影的结果等于该向量本身。

如果向量 $\mathbf{b}$ 垂直于 A 的列空间，即 $\mathbf{b}$ 位于 A 的左零空间 $N(A^T)$ 内时，则 $A^T \mathbf{b} = \mathbf{0}$ 。此时投影结果为：

$$P\mathbf{b} = A(A^T A)^{-1} (A^T \mathbf{b}) = A(A^T A)^{-1} \mathbf{0} = \mathbf{0}$$

这说明，垂直于列空间的分量在投影过程中会被完全消除。

对于一般的向量 $\mathbf{b}$ ，它总可以被分解为两个正交分量：

$$\mathbf{b} = \mathbf{p} + \mathbf{e}$$

其中， $\mathbf{p} = P\mathbf{b}$ 是 $\mathbf{b}$ 在列空间 $C(A)$ 中的分量，而 $\mathbf{e} = \mathbf{b} - \mathbf{p}$ 是 $\mathbf{b}$ 在左零空间 $N(A^T)$ 中的分量。

1.2 另一个投影矩阵 $I - P$

对于误差向量 $\mathbf{e} = \mathbf{b} - \mathbf{p}$ ，可以用矩阵形式表示为：

$$\mathbf{e} = \mathbf{b} - P\mathbf{b} = (I - P)\mathbf{b}$$

矩阵 $(I - P)$ 同样具有投影矩阵的全部性质，它负责将任意向量投影到 A 的左零空间中。它可以被视为 A 的列空间的正交补空间投影矩阵。

2 最小二乘法与线性回归

2.1 问题背景与正规方程

最小二乘法是投影理论在数据拟合中的经典应用。假设存在三个数据点 $(1, 1), (2, 2), (3, 2)$ ，试图寻找一条最接近这三点的直线方程 $y = \beta + \alpha x$ 。如果能完美穿过三点，则应满足方程组：

$$\beta + \alpha(1) = 1$$

$$\beta + \alpha(2) = 2$$

$$\beta + \alpha(3) = 2$$

将其写成矩阵形式 $A\mathbf{x} = \mathbf{b}$ ：

$$\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} \beta \\ \alpha \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$$

不难发现，这三个点不在同一条直线上（方程组矛盾），因此原方程无解。最小二乘法的目标即寻找

最优解 $\hat{\mathbf{x}} = \begin{bmatrix} \hat{\beta} \\ \hat{\alpha} \end{bmatrix}$ ，使得误差的平方和最小。

将 $\mathbf{b}$ 投影到 A 的列空间中，利用正规方程 $A^T A \hat{\mathbf{x}} = A^T \mathbf{b}$ 进行求解：

$$A^T A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 3 & 6 \\ 6 & 14 \end{bmatrix}$$

$$A^T \mathbf{b} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 5 \\ 11 \end{bmatrix}$$

代入正规方程：

$$\begin{bmatrix} 3 & 6 \\ 6 & 14 \end{bmatrix} \begin{bmatrix} \hat{\beta} \\ \hat{\alpha} \end{bmatrix} = \begin{bmatrix} 5 \\ 11 \end{bmatrix}$$

解得：

$$\hat{\beta} = \frac{2}{3}, \quad \hat{\alpha} = \frac{1}{2}$$

最终得到的最优拟合直线方程为： $y = \frac{2}{3} + \frac{1}{2}x$ 。

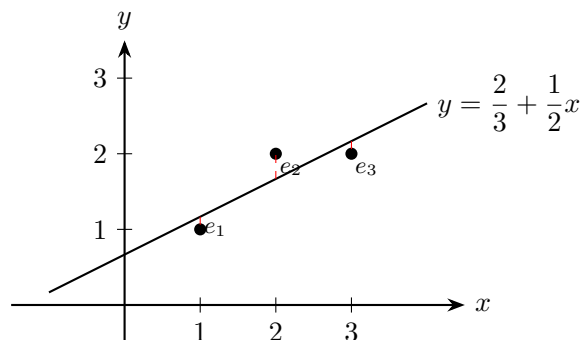


图 1：数据点到最优拟合直线的垂直距离（误差）

2.2 几何解释与拟合结果

从几何角度看，拟合过程等价于向量 $\mathbf{b} = (1, 2, 2)^T$ 投影到由矩阵 A 的列向量 $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ 和 $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ 所张成的二维列空间中。投影后的向量 $\mathbf{p}$ 即为拟合直线上三个点对应的纵坐标：

$$\mathbf{p} = A\hat{\mathbf{x}} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 2/3 \\ 1/2 \end{bmatrix} = \begin{bmatrix} 7/6 \\ 5/3 \\ 13/6 \end{bmatrix}$$

对应的误差向量 $\mathbf{e}$ 为实际数据点与拟合值的偏差：

$$\mathbf{e} = \mathbf{b} - \mathbf{p} = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} - \begin{bmatrix} 7/6 \\ 5/3 \\ 13/6 \end{bmatrix} = \begin{bmatrix} -1/6 \\ 2/6 \\ -1/6 \end{bmatrix}$$

通过计算可以验证， $\mathbf{p}$ 与 $\mathbf{e}$ 正交（点积为 0），且误差向量 $\mathbf{e}$ 垂直于 A 的两个列向量，这证明了最小二乘法的投影本质。

2.3 偏导数方法

除了矩阵投影法，还可以利用微积分中的偏导求解。最小化误差平方和函数：

$$E(\alpha, \beta) = e_1^2 + e_2^2 + e_3^2 = (\beta + \alpha - 1)^2 + (\beta + 2\alpha - 2)^2 + (\beta + 3\alpha - 2)^2$$

分别对 β 和 α 求偏导数，并令其等于 0：

$$\begin{aligned} \frac{\partial E}{\partial \beta} &= 2(\beta + \alpha - 1) + 2(\beta + 2\alpha - 2) + 2(\beta + 3\alpha - 2) = 0 \\ \frac{\partial E}{\partial \alpha} &= 2(\beta + \alpha - 1) + 4(\beta + 2\alpha - 2) + 6(\beta + 3\alpha - 2) = 0 \end{aligned}$$

化简得到与正规方程完全相同的线性方程组：

$$3\beta + 6\alpha = 5$$

$$6\beta + 14\alpha = 11$$

这证明了通过微积分（最小化距离平方和）和线性代数（寻找投影）得到的解是完全一致的。

总结

本讲系统展示了投影理论在最小二乘法中的应用：

- 通过投影矩阵 $P = A(A^T A)^{-1} A^T$ ，任意向量 $\mathbf{b}$ 可以被分解为列空间分量 $\mathbf{p} = P\mathbf{b}$ 与左零空间分量 $\mathbf{e} = (I - P)\mathbf{b}$ 。
- 对于无解的线性方程组，最小二乘法通过求解正规方程 $A^T A\hat{\mathbf{x}} = A^T \mathbf{b}$ 可以得到误差平方和最小的最优解。
- 矩阵 $A^T A$ 可逆的充要条件是 A 的列向量线性无关。当 A 的列向量是标准正交向量组时，计算会极大简化。

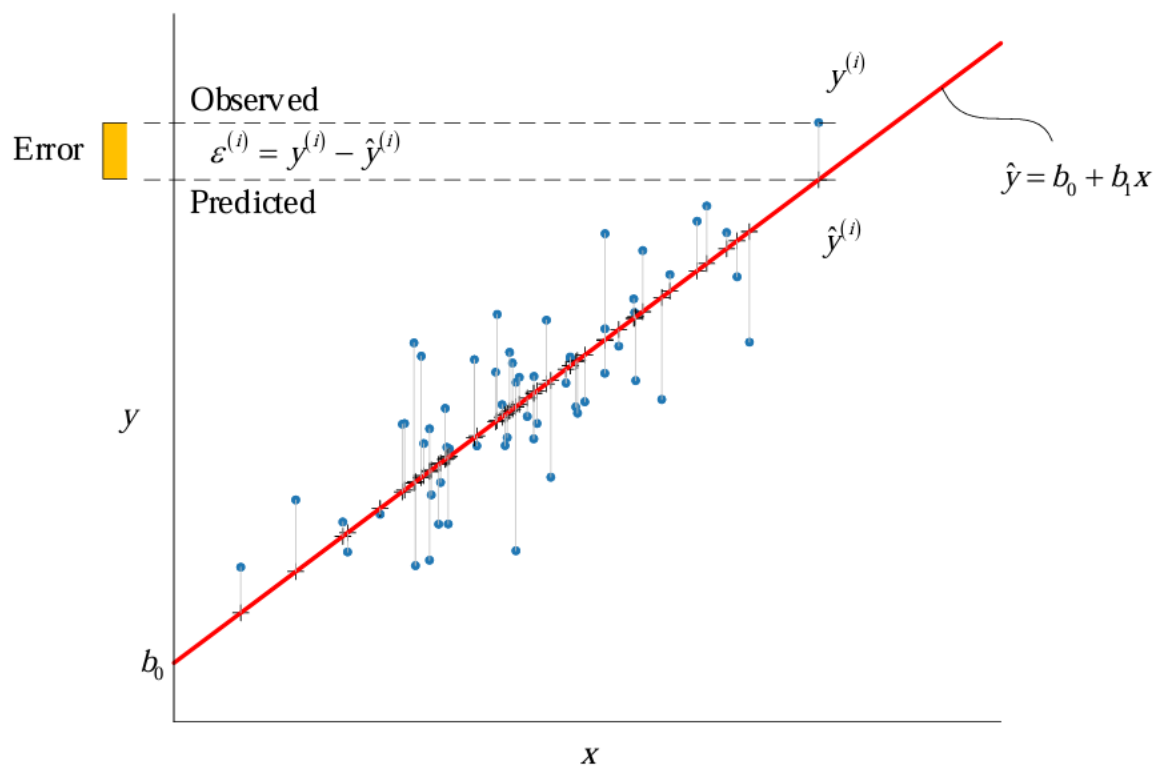


图 2. 简单线性回归中的残差项

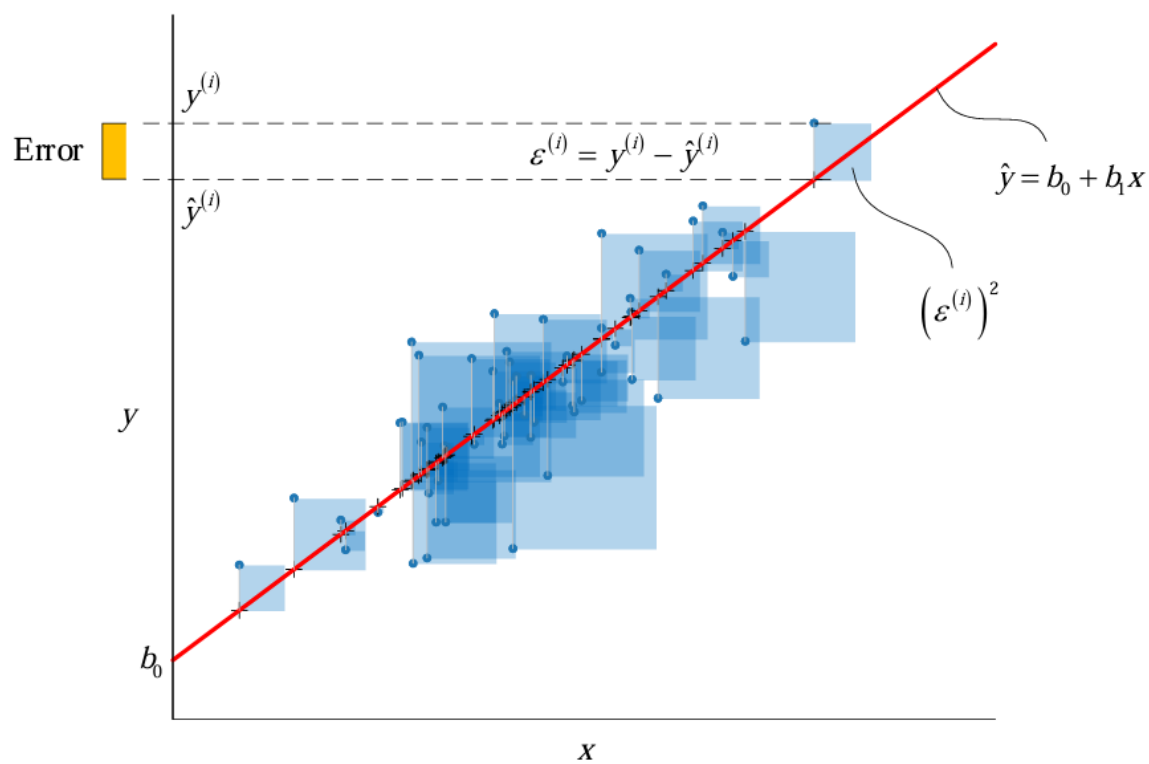


图 6. 残差平方和的几何意义