高级工程数学第一次作业

51215901023 鲁梦洁

问题：最小二乘解法等价性证明

对于最小二乘问题 $Ax = b$ 有两种解法：第一种方法是转化为 $\left(A^T A\right)x = A^T b$ 求解；第二种解法是转化为一个求最小值的问题 $\min_x \|Ax - b\|_2^2$。求证这两种解法是等价的。

这里给出如下两种证明方式——利用数学运算和求导公式推导证明、利用最小二乘法的几何意义和矩阵投影进行证明。

方法一：利用数学运算和求导公式推导证明

(1) $\min_x \|Ax - b\|_2^2 \Rightarrow \left(A^T A\right)x = A^T b$

$$f(x) = \|Ax - b\|_2^2 = (Ax - b)^T (Ax - b) = \left(x^T A^T - b^T\right)(Ax - b) = x^T A^T Ax - x^T A^T b - b^T Ax + b^T b$$

那么要求 $\min_x f(x)$，那么可以对 $f(x)$ 进行求导

$$\frac{df(x)}{dx} = \frac{df(x)}{dx} = \frac{d\left(x^T A^T Ax\right)}{dx} - \frac{d\left(x^T A^T b\right)}{dx} - \frac{d\left(b^T Ax\right)}{dx} + \frac{d\left(b^T b\right)}{dx}$$

令 $\frac{df(x)}{dx} = 0$ 可以得到 $\frac{df(x)}{dx} = 2A^T Ax - 2A^T b = 0$，即 $\left(A^T A\right)x = A^T b$

(2) $\left(A^T A\right)x = A^T b \Rightarrow \min_x \|Ax - b\|_2^2$

已知 $\left(A^T A\right)x = A^T b$，那么 $\frac{df(x)}{dx} = 2A^T Ax - 2A^T b = 0$。

$$f(x) = \|Ax - b\|_2^2 = (Ax - b)^T (Ax - b) = \left(x^T A^T - b^T\right)(Ax - b) = x^T A^T Ax - x^T A^T b - b^T Ax + b^T b$$

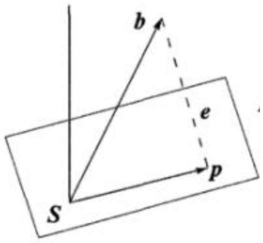
$$\frac{df(x)}{dx} = \frac{df(x)}{dx} = \frac{d\left(x^T A^T Ax\right)}{dx} - \frac{d\left(x^T A^T b\right)}{dx} - \frac{d\left(b^T Ax\right)}{dx} + \frac{d\left(b^T b\right)}{dx} = 2A^T Ax - 2A^T b = 0$$

因此当满足 $\left(A^T A\right)x = A^T b$ 时，$f(x) = \|Ax - b\|_2^2$ 取到最小值，即 $\min_x \|Ax - b\|_2^2$。

方法二：利用最小二乘法的几何意义和矩阵投影进行证明

情况一：如果 $Ax=b$ 有解，那么向量 b 可以表示为矩阵 A 的列向量的线性组合，所以 $Ax-b=0$，即 $\min_x \|Ax-b\|_2^2$ 得到最小值 0。

情况二：如果 $Ax=b$ 无解，设 $A=\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}=\begin{bmatrix} a_1 \\ a_2 \\ \dots \\ a_n \end{bmatrix}$，$C(A)$ 是矩阵 A 的列向量构成的线性子空间，$Ax=b$ 的近似解 $\hat{x}=\begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \\ \dots \\ \hat{x}_n \end{bmatrix}$，$b=\begin{bmatrix} b_1 \\ b_2 \\ \dots \\ b_m \end{bmatrix}$。那么 $Ax=b$ 可以看作寻找一组 m 个 n 元线性方程组的解。那么需要从如下两个方面证明最小二乘解法的等价性。



(1) $\min_x \|Ax-b\|_2^2 \Rightarrow (A^TA)x=A^Tb$

假设 $Ax=b$ 的近似解 $\hat{x}$ 满足 $\min_x \|Ax-b\|_2^2$，也就是说，使得 m 个线性方程误差的平方和最小的一组向量就是 $Ax=b$ 的近似解。设 $A\hat{x}=p$，那么求 $Ax=b$ 的近似解相当于在矩阵 A 的列空间中找到一个向量 p，使得 p 与 b 的误差最小，即 p 是 b 在矩阵 A 的列空间上的投影，如上图所示。因此 $C(A)\perp e$，从而得到 $A^Te=0$。把 $e=b-p$ 代入 $A^Te=0$，得到 $A^Te=A^T\left(b-A\hat{x}\right)=0$，从而得到 $(A^TA)\hat{x}=A^Tb$，即 $\hat{x}=(A^TA)^{-1}A^Tb$。

(2) $(A^TA)x=A^Tb \Rightarrow \min_x \|Ax-b\|_2^2$

假 设 $Ax=b$ 的 近 似 解 $\hat{x}$ 满 足 $(A^TA)\hat{x}=A^Tb$ ， 即 $A^T\left(b-A\hat{x}\right)=0$ ， 所 以

$b - A\hat{x} \perp C(A)$。设 $A\hat{x} = p$，因此 $C(A) \perp p$，即 p 是 b 在矩阵 A 的列空间上的投影。那么求 $Ax = b$ 的近似解相当于在矩阵 A 的列空间中找到一个向量 p，使得 p 与 b 的误差最小。也就是说，使得 m 个线性方程误差的平方和最小的一组向量就是 $Ax = b$ 的近似解，即 $Ax = b$ 的近似解 $\hat{x}$ 满足 $\min_{x} \|Ax - b\|_2^2$。

综上，完成了最小二乘解法等价性证明。