线性代数

Worked Examples for Lecture 9

LECTURER: 杨启哲　　LAST MODIFIED: 2024 年 4 月 30 日

1. 假设现在有 9 个测量结果 $(1,0),(2,0),\cdots(9,0),(10,40)$，即前面 9 个测量结果的第二个分量都是 0，而第十个测量结果的第二个分量是 40。现在请寻找一条水平直线 $y=C$ 去尽可能的拟合这些测量结果，使得下列误差尽可能的小，这里令 e_i 是直线 $y=C$ 和第 i 个测量结果的误差。

   (1) 最小二乘 (least squares)：$E=\sum_{i=1}^{10} e_i^2$

   (2) 最小最大误差 (least maximum error)：$E_\infty=\max_{i\in[10]}|e_i|$.

   (3) 最小绝对误差 (least sum of errors)：$E_1=\sum_{i=1}^{10}|e_i|$.

分析

- 转换成方程的表述是指考虑下列的方程组：

$$\begin{bmatrix}1\\ \vdots\\ 1\end{bmatrix}\begin{bmatrix}x\end{bmatrix}=\begin{bmatrix}0\\ \vdots\\ 0\\ 40\end{bmatrix}\quad(A\mathbf{x}=\mathbf{b})$$

  找到一个近似解 C 使得需要满足的误差最小。
- 这道题其实想说明的是，对于不同的误差度量，我们可能会得到不同的拟合结果，这也是我们在实际问题中需要考虑的问题。
- 按照范数的概念，实际上这分别是 2-范数，∞-范数和 1-范数。而我们课上着重讲的是最小二乘实际上考虑的是 2-范数，也就是 $\|\mathbf{b}-A\mathbf{x}\|$ 的长度 (欧氏距离)。
- 最小二乘的求解实际上通过上一点我们可知道，本质上就是求到 A 的列空间上的投影（此时误差的长度最小），所以利用投影的方式做就可以了。

解答.

- 计算 $A^TA=10$，$A^T\mathbf{b}=40$，从而此时的 $C=4$，即 $y=4$ 是最佳拟合直线，其对应的误差为：

$$E=\sum_{i=1}^{9}(0-4)^2+(40-4)^2=1440$$

- 注意到其误差表示为：

$$E_\infty=\max\{|0-C|,|40-C|\}$$

  所以此时 $C=20$ 是误差最小的时候，其误差为 $E_\infty=20$，此时最佳拟合直线为 $y=20$.
- 注意到其误差表示为：

$$E_1=\sum_{i=1}^{9}|0-C|+|40-C|=9|C|+|40-C|$$

从而当 $C<0$ 时误差递减，当 $C>40$ 时误差递增，所以此时 $C=0$ 是误差最小的时候，其误差为 $E_1=40$，此时最佳拟合直线为 $y=0$.

□

2. 对下述 5 个点 $(-2,0),(-1,0),(0,1),(1,0),(2,0)$ 找到一条最为接近的抛物线 (parabola) $y=ax^2+bx+c$(这里采用的是最小二乘的误差)。

分析

事实上，对于任何的多项式拟合，如果确定了多项式的次数，我们都可以转换成 $A\mathbf{x}=\mathbf{b}$ 的形式。一般来说，假设有 n 个点 (x_i,y_i)，我们想拟合成一个 n 次多项式 $f(x)=a_nx^n+\cdots+a_1x+a_0$，其实际上是在求下列线性方程组的近似解：

$$\begin{bmatrix}1 & x_1 & \cdots & x_1^n\\ \vdots & \vdots & \ddots & \vdots\\ 1 & x_n & \cdots & x_n^n\end{bmatrix}\begin{bmatrix}a_0\\ \vdots\\ a_n\end{bmatrix}=\begin{bmatrix}y_1\\ \vdots\\ y_n\end{bmatrix}$$

注意到此时左面的矩阵是已知的，并不是未知的。

当然当次数高了以后，计算会变得复杂，但有关矩阵的乘法包括逆的计算我们可以借助计算机来完成。本质还是要知道计算方法。

解答. 令 $y=ax^2+bx+c$，则 (a,b,c) 实际上是下列方程的最优近似解：

$$\begin{bmatrix}1 & -2 & 4\\ 1 & -1 & 1\\ 1 & 0 & 0\\ 1 & 1 & 1\\ 1 & 2 & 4\end{bmatrix}\begin{bmatrix}c\\ b\\ a\end{bmatrix}=\begin{bmatrix}0\\ 0\\ 1\\ 0\\ 0\end{bmatrix}\ (A\mathbf{x}=\mathbf{b})$$

分别计算 A^TA 和 $A^T\mathbf{b}$ 可得：

$$A^TA=\begin{bmatrix}5 & 0 & 10\\ 0 & 10 & 0\\ 10 & 0 & 34\end{bmatrix},\ A^T\mathbf{b}=\begin{bmatrix}1\\ 0\\ 0\end{bmatrix}$$

进一步解方程 $A^TA\mathbf{x}=A^T\mathbf{b}$ 可得 $\mathbf{x}=(\frac{34}{70},0,-\frac{10}{70})$，从而拟合的抛物线为 $y=\frac{34}{70}x^2-\frac{10}{70}$. □

3. 对下列向量 $\mathbf{a},\mathbf{b},\mathbf{c}$，给出 $Gram-Schmidt$ 正交化后得到的正交向量 $\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3$。

$$\mathbf{a}=\begin{bmatrix}1\\ -1\\ 0\end{bmatrix},\ \mathbf{b}=\begin{bmatrix}2\\ 0\\ -2\end{bmatrix},\ \mathbf{c}=\begin{bmatrix}3\\ -3\\ 3\end{bmatrix}$$

分析

$Gram-Schmidt$ 正交化的本质就是投影，然后作正交补的拆分。大家可以想象一下，假设你已经有了 $\mathbf{q}_1,\cdots,\mathbf{q}_k$ 这 k 个正交的向量，其张成的向量空间为 Q_k，现在 $\mathbf{a}_{k+1}$ 是不属于 Q_k 的一个向量，那么构造一个 $\mathbf{q}_{k+1}$ 和 $\mathbf{q}_1,\cdots,\mathbf{q}_k$ 都正交的方式是这样的：

(1) 求 $\mathbf{a}_{k+1}$ 到 Q_k 的投影 $\mathbf{q}'_k$。

(2) 由投影的概念，其误差 $\mathbf{a}_{k+1}-\mathbf{q}'_k$ 与 $\mathbf{q}_{k+1}$ 和 $\mathbf{q}_1,\cdots,\mathbf{q}_k$ 都正交，这也就是 $\mathbf{q}_{k+1}$。

(3) 同时我们还要注意到由于 $\mathbf{q}_i$ 之间都是正交的，所以求到 Q_k 的投影实际上就是求到每个 $\mathbf{q}_i$ 的投影之和。

解答.

- 令 $\mathbf{q}_1=\mathbf{a}_1=\begin{bmatrix}1\\-1\\0\end{bmatrix}$
- 求 $\mathbf{b}$ 到 $\mathbf{q}_1$ 的投影 $\mathbf{p}_1=\frac{\mathbf{q}_1\mathbf{q}_1^T}{\mathbf{q}_1^T\mathbf{q}_1}\mathbf{b}=\begin{bmatrix}1\\-1\\0\end{bmatrix}$，所以我们有：

$$\mathbf{q}_2=\mathbf{b}-\mathbf{p}_1=\begin{bmatrix}1\\1\\-2\end{bmatrix}$$

- 求 $\mathbf{c}$ 分别到 $\mathbf{q}_1$ 和 $\mathbf{q}_2$ 的投影 $\mathbf{p}_1=\frac{\mathbf{q}_1\mathbf{q}_1^T}{\mathbf{q}_1^T\mathbf{q}_1}\mathbf{c}=\begin{bmatrix}3\\-3\\0\end{bmatrix}$，$\mathbf{p}_2=\frac{\mathbf{q}_2\mathbf{q}_2^T}{\mathbf{q}_2^T\mathbf{q}_2}\mathbf{c}=\begin{bmatrix}-1\\-1\\2\end{bmatrix}$，所以我们有：

$$\mathbf{q}_3=\mathbf{c}-\mathbf{p}_1-\mathbf{p}_2=\begin{bmatrix}1\\1\\1\end{bmatrix}$$

□

4. 请给出所有的既是下三角矩阵又是正交矩阵的 $n\times n$ 矩阵。

分析

首先我们明确一下两种矩阵的定义：

- 下三角矩阵：对于一个 $n\times n$ 的矩阵 A，如果对于任意的 $i<j$ 都有 $A(i,j)=0$，即主对角线的右上方全是 0.
- 正交矩阵：对于一个 $n\times n$ 的矩阵 A，其列向量是两两正交且都是单位向量。(或者说 $A^TA=I$)

所以考虑矩阵第 n 列，其形式是：

$$\begin{bmatrix}0\\\vdots\\0\\a\end{bmatrix}$$

由于其必须是个单位向量，所以 $a=\pm1$。依次往左推即可得到相应答案。

解答. 其形式为如下的对角矩阵：

$$Q = diag(\pm 1, \cdots, \pm 1) = \begin{bmatrix} \pm 1 & 0 & \cdots & 0 \\ 0 & \pm 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \pm 1 \end{bmatrix}$$

□