数值分析内部课程习题解答

第三章

SCAU DataHub

2025 年 4 月 18 日

目录

1 Fitting functions to data 2

2 The normal equations 2

3 The QR factorization 4

4 Computing QR factorizations 5

1 Fitting functions to data

Solution 1.1. 根据局部极小充分条件来推导.

首先明确局部极小的充分条件，对函数 $f(x)$，若 x^* 使得 $f'(x^*)=0$ 且 $f''(x^*)>0$，则 x^* 一定是 $f(x)$ 的局部极小点.

题目中给出的条件已经说明 x^* 是 $f(x)$ 的局部极小点了，需要我们证明的是，x^* 还是 $[f(x)]^2$ 的局部极小点. 令 $g(x)=[f(x)]^2$，结合题设条件，只需验证 $g'(x^*)=0$ 且 $g''(x^*)>0$ 即可证明.

Solution 1.2. 代码见 Gitee 仓库.

Solution 1.3. 代码见 Gitee 仓库.

Solution 1.4. 代码见 Gitee 仓库.

Solution 1.5. 代码见 Gitee 仓库.

作变换 $z=\frac{t^2}{1+t^2}$ 后，拟合效果比原先更好的原因是，原先以 $1,t,t^2,t^3$ 作为基是无界的，而改用 $1,z,z^2,z^3$ 作为基是有界的，而且函数图像更像 tanh，所以效果更好. 这启发我们做数据拟合的时候选一组合适的基非常重要.

Solution 1.6. 代码见 Gitee 仓库.

Solution 1.7. 代码见 Gitee 仓库.

Solution 1.8. 对等式两边取倒数，转化为拟合线性函数.

观察拟合函数的形式，对等式两边取倒数，有

$$\underbrace{\frac{1}{y(t)}}_{=:z}\approx\frac{t+b}{a}=\underbrace{\frac{b}{a}}_{=:c_1}+\underbrace{\frac{1}{a}}_{=:c_2}t,$$

按照上式定义 z,c_1,c_2，就转化成了 z 关于 t 的线性函数，只需要拟合 $z=c_1+c_2t$，然后将 c_1,c_2 转化回 a,b 即可，具体公式为

$$a=\frac{1}{c_2},\quad b=\frac{c_1}{c_2}.$$

Solution 1.9. 对等式两边取倒数，转化为拟合线性函数.

观察拟合函数的形式，对等式两边取倒数，有

$$\underbrace{\frac{1}{y(t)}}_{=:z}\approx\frac{at+b}{t}=\underbrace{a}_{=:c_1}+\underbrace{b}_{=:c_2}\underbrace{\frac{1}{t}}_{=:x},$$

按照上式定义 z,x,c_1,c_2，就转化成了 z 关于 x 的线性函数，剩余步骤同上题.

2 The normal equations

Solution 2.1. 按照正规方程的形式求解.

最小二乘问题可以转化为求解正规方程 $\mathbf{A}^T\mathbf{A}x = \mathbf{A}^T\mathbf{b}$，于是计算

$$\mathbf{A}^T\mathbf{A} = \begin{bmatrix} 8 & -6 \\ -6 & 6 \end{bmatrix}, \quad \mathbf{A}^T\mathbf{b} = \begin{bmatrix} -10 \\ 6 \end{bmatrix},$$

求解线性系统，得到

$$\mathbf{x} = \begin{bmatrix} -2 \\ -1 \end{bmatrix}.$$

Solution 2.2. 按照伪逆的定义一步步算.

最终结果为

$$\mathbf{A}^+ = \frac{1}{14}\begin{bmatrix} 1 & -2 & 3 \end{bmatrix}.$$

Solution 2.3. 按照矩阵对称的定义容易证得.

对任意的 $m \times n$ 矩阵 $\mathbf{A}$，有

$$(\mathbf{A}^T\mathbf{A})^T = \mathbf{A}^T\mathbf{A},$$

故 $\mathbf{A}^T\mathbf{A}$ 是对称矩阵.

Solution 2.4. 按照可逆矩阵的性质容易证得.

由于题设 $\mathbf{A}$ 是一个可逆的方阵，所以有

$$\mathbf{A}^+ = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T = \mathbf{A}^{-1}(\mathbf{A}^T)^{-1}\mathbf{A}^T = \mathbf{A}^{-1}.$$

Solution 2.5. 根据本节的定理 3.2.4 可得到.

(a) 因为 $\mathbf{A}^T\mathbf{A}$ 是非奇异的，所以根据伪逆的定义容易得到.

(b) 代码见 Gitee 仓库.

Solution 2.6. 根据最小二乘问题和伪逆的定义可以得到.

对于最小二乘问题 $\min \|\mathbf{b} - \mathbf{A}\mathbf{x}\|_2$，就是要找一个 $\mathbf{x}$ 使得 $\mathbf{A}\mathbf{x}$ 尽可能地接近 $\mathbf{b}$，且 $\mathbf{A}\mathbf{x}$ 是 $\mathbf{A}$ 的值域的向量 (这是必然的，$\mathbf{A}\mathbf{x} \in \mathrm{Range}(\mathbf{A})$). 根据正规方程和伪逆的定义可知，$\mathbf{x} = \mathbf{A}^+\mathbf{b}$ 就是最小二乘问题的解，即 $\mathbf{A}\mathbf{A}^+\mathbf{b}$ 就是那个尽可能接近 $\mathbf{b}$ 且在 $\mathbf{A}$ 的值域的向量.

Solution 2.7. 根据算法中的各个步骤逐个拆解.

渐进复杂度中 $2mn^2$ 来自于 $\mathbf{N} = \mathbf{A}^T\mathbf{A}$，而 $\frac{1}{3}n^3$ 来自于 Cholesky 分解. 其余的步骤的复杂度都是小于 3 阶的.

Solution 2.8. 代码见 Gitee 仓库.

其中 part (b) 出现 $\kappa(\mathbf{A}_\beta) \to \infty$ 的原因是，当 β 逐渐减小时，$\sin^2(t) + \cos^2(\beta t) \approx 1$，导致 $\mathbf{A}$ 的列向量线性相关 (矩阵奇异)，所以条件数增大.

Solution 2.9. 代码见 Gitee 仓库.

3 The QR factorization

Solution 3.1. 利用标准正交矩阵的性质.

对任意的 $\|\mathbf{x}\|_2 = 1$，有

$$\|\mathbf{Qx}\|_2^2 = (\mathbf{Qx})^T\mathbf{Qx} = \mathbf{x}^T\mathbf{Q}^T\mathbf{Qx} = \mathbf{x}^T\mathbf{x} = \|\mathbf{x}\|_2^2 = 1,$$

因此 $\|\mathbf{Qx}\|_2 = 1$，进而有

$$\|\mathbf{Q}\|_2 = \max_{\|\mathbf{x}\|_2=1} \|\mathbf{Qx}\|_2 = 1.$$

Solution 3.2. 根据正交矩阵的定义以及前面 ONC 矩阵的性质可证. 证明略.

其中第 3 条性质可以结合前面定理 3.2.4 证明.

Solution 3.3. 代码参考 Gitee 仓库.

Solution 3.4. 根据题中的 Hint 可证. 证明略.

可参考两个定理，一个是矩阵 $\mathbf{R}$ 不满秩等价于行列式 $|\mathbf{R}| = 0$，另一个是矩阵乘积的行列式等于矩阵行列式的乘积，即 $|\hat{\mathbf{Q}}\hat{\mathbf{R}}| = |\hat{\mathbf{Q}}||\hat{\mathbf{R}}|$.

Solution 3.5. 根据矩阵伪逆的定义以及正交矩阵的性质可以得到. 推导略.

Solution 3.6. 根据矩阵伪逆的定义以及正交矩阵的性质可以得到. 推导略.

Solution 3.7. 代码参考 Gitee 仓库.

Solution 3.8. 根据正交矩阵的性质可以证明.

(a) 结合第 6 题的结论可证，证明略.

(b) 根据正交矩阵的性质可证.

$\mathbf{P}^2 = \mathbf{PP} = \hat{\mathbf{Q}}\hat{\mathbf{Q}}^T\hat{\mathbf{Q}}\hat{\mathbf{Q}}^T = \hat{\mathbf{Q}}(\hat{\mathbf{Q}}^T\hat{\mathbf{Q}}) = \hat{\mathbf{Q}}\mathbf{I}\hat{\mathbf{Q}}^T = \hat{\mathbf{Q}}\hat{\mathbf{Q}}^T = \mathbf{P}$

(c) 证明 $\mathbf{u}, \mathbf{v}$ 正交，即证明 $\mathbf{u}^T\mathbf{v} = 0$，代入 $\mathbf{u}, \mathbf{v}$ 的表达式，有

$$\mathbf{u}^T\mathbf{v} = (\mathbf{Px})^T(\mathbf{I} - \mathbf{P})\mathbf{x} = (\mathbf{Px})^T\mathbf{x} - (\mathbf{Px})^T\mathbf{Px} = \mathbf{x}^T\mathbf{P}^T\mathbf{x} - \mathbf{x}^T\mathbf{P}^T\mathbf{Px},$$

注意到，$\mathbf{P} = \hat{\mathbf{Q}}\hat{\mathbf{Q}}^T$，故

$$\mathbf{P}^T\mathbf{P} = (\hat{\mathbf{Q}}\hat{\mathbf{Q}}^T)^T\hat{\mathbf{Q}}\hat{\mathbf{Q}}^T = \hat{\mathbf{Q}}\hat{\mathbf{Q}}^T\hat{\mathbf{Q}}\hat{\mathbf{Q}}^T = \hat{\mathbf{Q}}\hat{\mathbf{Q}}^T = \mathbf{P},$$

因此

$$\mathbf{u}^T\mathbf{v} = \mathbf{x}^T\mathbf{P}^T\mathbf{x} - \mathbf{x}^T\mathbf{P}^T\mathbf{Px} = \mathbf{x}^T\mathbf{P}^T\mathbf{x} - \mathbf{x}^T\mathbf{Px} = 0,$$

得到 $\mathbf{u}, \mathbf{v}$ 正交.

4 Computing QR factorizations

Solution 4.1. 关键要理解 Householder 变换的几何意义.

Householder 变换矩阵定义为 $\mathbf{P} = \mathbf{I} - 2\mathbf{v}\mathbf{v}^T$，所以我们找 $\mathbf{P}$ 其实是在找 $\mathbf{v}$. 结合图 3.4.1，向量 $\mathbf{v}$ 就像一个镜子，我们希望将向量 $\mathbf{z} = [2, 9, -6]^T$ 反射到 x_1 轴上，变成 $\mathbf{Pz} = \|\mathbf{z}\|[1, 0, 0]^T$.

第一步需要确定 $\mathbf{v}$ 的方向，即垂直于 $\mathbf{z}$ 和 $\mathbf{Pz}$ 的对称轴，不妨先记为 $\mathbf{w}$，于是有

$$\mathbf{w} = \underbrace{\|\mathbf{z}\|\begin{bmatrix}1\\0\\0\end{bmatrix}}_{\text{变换后}} - \underbrace{\mathbf{z}}_{\text{变换前}} = \begin{bmatrix}11\\0\\0\end{bmatrix} - \begin{bmatrix}2\\9\\-6\end{bmatrix} = \begin{bmatrix}9\\-9\\6\end{bmatrix}.$$

第二步是单位化，即

$$\mathbf{v} = \frac{\mathbf{w}}{\|\mathbf{w}\|}.$$

具体结果见 Gitee 仓库.

Solution 4.2. 根据矩阵对称和正交的定义可证.

矩阵对称是指 $\mathbf{P}^T = \mathbf{P}$，显然有

$$\mathbf{P}^T = (\mathbf{I} - 2\mathbf{v}\mathbf{v}^T)^T = \mathbf{I} - 2(\mathbf{v}\mathbf{v}^T)^T = \mathbf{I} - 2\mathbf{v}\mathbf{v}^T = \mathbf{P}.$$

矩阵正交是指 $\mathbf{P}^T\mathbf{P} = \mathbf{I}$，结合 $\mathbf{P}$ 的对称性，展开表达式

$$\mathbf{P}^T\mathbf{P} = \mathbf{PP} = (\mathbf{I} - 2\mathbf{v}\mathbf{v}^T)(\mathbf{I} - 2\mathbf{v}\mathbf{v}^T) = \mathbf{I} - 4\mathbf{v}\mathbf{v}^T + 4(\mathbf{v}\mathbf{v}^T)(\mathbf{v}\mathbf{v}^T),$$

注意到 $\mathbf{v}$ 是标准正交向量，即 $\mathbf{v}^T\mathbf{v} = 1$，所以上式右端可化为

$$\mathbf{I} - 4\mathbf{v}\mathbf{v}^T + 4\mathbf{v}(\mathbf{v}^T\mathbf{v})\mathbf{v}^T = \mathbf{I} - 4\mathbf{v}\mathbf{v}^T + 4\mathbf{v}\mathbf{v}^T = \mathbf{I}.$$

Solution 4.3. 参考图 3.4.1，理解几何意义.

(a) 当 $\mathbf{u}$ 转到与 $\mathbf{v}$ 平行的方向时，$\mathbf{Pu} = -\mathbf{u}$，所以取 $\mathbf{u} = \mathbf{v}$ 或 $\mathbf{v}$ 的若干倍都可以.

(b) 当 $\mathbf{u}$ 转到与 $\mathbf{v}$ 垂直的方向时，$\mathbf{Px} = \mathbf{x}$. 推广到 n 维空间中时，垂直则变成了正交，即要满足 $\mathbf{v}^T\mathbf{x} = 0$，设 $\mathbf{v}$ 在 $\mathbb{R}^n$ 中的一个 m 维子空间 $\mathbb{R}^m$，则与 $\mathbf{v}$ 正交的向量只能落在 $n - m$ 维的子空间 $\mathbb{R}^n \setminus \mathbb{R}^m$ 中，所以满足 $\mathbf{Px} = \mathbf{x}$ 的线性无关的向量至多有 $n - m$ 个.

Solution 4.4. 代码见 Gitee 仓库.

这里的算法是想将 $\mathbf{z}$ 转换到 $\mathbf{e}_1$，公式 3.4.4 不适用于二者相近的情况，所以在函数 3.4.4 的第 12 行做了优化.

Solution 4.5. 巩固目前学到的两种分解，LU 分解和 QR 分解.

(a) 参考定理 3.4.5，对于方阵而言，QR 分解的复杂度近似为 $\frac{4}{3}n^3$，而 LU 分解 (参考定理 2.5.9) 的复杂度近似为 $\frac{2}{3}n^3$. 从复杂度的角度来看，LU 分解要更快，但是从数值稳定性的角度来看，QR 分解更稳定.

(b) 条件数 $\kappa_2(\mathbf{A})$ 定义为
$$\kappa_2(\mathbf{A}) = \|\mathbf{A}^{-1}\|_2\|\mathbf{A}\|_2,$$
根据题设 $\mathbf{A} = \mathbf{QR}$，且 $\mathbf{A}$ 为方阵，故
$$\kappa_2(\mathbf{A}) = \|(\mathbf{QR})^{-1}\|_2\|\mathbf{QR}\|_2 = \|\mathbf{R}^{-1}\mathbf{Q}^{-1}\|_2\|\mathbf{QR}\|_2 = \|\mathbf{R}^{-1}\|_2\|\mathbf{R}\|_2 = \kappa_2(\mathbf{R}),$$
其中用到了正交矩阵的性质 (定理 3.3.6)，$\|\mathbf{QR}\|_2 = \|\mathbf{R}\|_2$.

Solution 4.6. 非方阵的条件数的定义为 $\kappa_2(\mathbf{A}) = \|\mathbf{A}\|_2\|\mathbf{A}^+\|_2$，思路同上题，证明略.

Solution 4.7. 理解 Givens 旋转变换.

(a) Householder 变换考虑镜面反射变换到 $\mathbf{e}_1$，而 Givens 变换考虑旋转变换到 $\mathbf{e}_1$. 旋转矩阵中的 θ 不用求，仅需要将 $\cos(\theta), \sin(\theta)$ 分别看作未知量 x, y，于是有
$$\begin{bmatrix} x & y \\ -y & x \end{bmatrix}\begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} \sqrt{\alpha^2+\beta^2} \\ 0 \end{bmatrix} \Rightarrow \begin{cases} x\alpha + y\beta = \sqrt{\alpha^2+\beta^2} \\ -y\alpha + x\beta = 0 \end{cases},$$
解得
$$x = \frac{\alpha}{\sqrt{\alpha^2+\beta^2}}, \quad y = \frac{\beta}{\sqrt{\alpha^2+\beta^2}},$$
于是就计算出了 $\cos(\theta)$ 和 $\sin(\theta)$ 的表达式.

(b) 代码见 Gitee 仓库.

Solution 4.8. 巩固章节 2.5 的内容.

考虑计算复杂度近似时，仅考虑最高阶的项，可简单理解为考虑循环中的循环，函数 3.4.4 中的核心计算量在第 18 行的循环中.

注意到第 18 行的目的是计算 $\mathbf{R}$，但计算第 k 列时，前面 1 至 $k-1$ 列的 k 至 m 行都已经为 0 了，所以第 17 行并不需要 for j in 1:n，而应该改为 for j in k:n.

因此，对于 18 行内的运算，我们一步步拆分：向量内积包括 $m-k+1$ 次乘法和 $m-k$ 次加法，数乘数包括 1 次乘法，向量乘数包括 $m-k+1$ 次乘法，向量减向量包

括 $m-k+1$ 次乘法. 因此，18 行的总计算量可以写为

$$\begin{aligned}\sum_{k=1}^{n}\sum_{j=k}^{n}4(m-k+1) &= 4\sum_{k=1}^{n}(n-k+1)(m-k+1)\\ &= 4\sum_{k=0}^{n-1}(n-k)(m-k)\\ &= 4\sum_{k=0}^{n-1}mn-(m+n)k+k^2\\ \text{(公式 2.5.1)}\qquad &\approx 4\left(mn^2-(m+n)\frac{n^2}{2}+\frac{n^3}{3}\right)\\ &= 2mn^2-\frac{2}{3}n^3.\end{aligned}$$

Solution 4.9. 根据题设 $m=Kn$ 代入并求极限，答案显然. 证明略.