

线性最小二乘问题

July 22, 2019

最小二乘问题

- 线性最小二乘问题
- 总体最小二乘问题
- 约束最小二乘问题
- ...

最小二乘问题在统计学, 最优化问题, 材料与结构力学, 信号与图像处理等方面都有着广泛的应用, 是计算数学的一个重要研究分支, 也是一个活跃的研究领域.

本讲主要介绍求解线性最小二乘问题的三种直接法.

1 引言

考虑线性最小二乘问题

$$\min_{x \in \mathbb{R}} \|Ax - b\|_2^2 \quad (1)$$

其中 $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$. 问题 (1) 的解称为最小二乘解.

- 当 $m = n$ 且 A 非奇异时, 这就是一个线性方程组, 解为 $x = A^{-1}b$;
- 当 $m < n$ 时, 未知量个数大于约束个数, 此时我们称问题 (1) 为欠定 (或亚定) 的
- 当 $m > n$ 时, 约束个数大于未知量个数, 此时我们称问题 (1) 为超定的

为了讨论方便, 本讲总是假定 A 是满秩的.

1.1 欠定方程组

若 $m < n$, 则 $Ax = b$ 存在无穷多个解, 这时我们通常寻求最小范数解, 于是问题就转化为下面的约束优化问题

$$\min_{Ax=b} \frac{1}{2} \|x\|_2^2 \quad (2)$$

对应的Lagrange 函数为

$$\mathcal{L}(x, \lambda) = \frac{1}{2}\|x\|_2^2 + \lambda^\top (Ax - b)$$

其中 $\lambda = [\lambda_1, \lambda_2, \dots, \lambda_m]^\top$ 是Lagrange 乘子.

此时优化问题 (2) 的解就是 $\mathcal{L}(x, \lambda)$ 的鞍点, 即下面方程组的解:

$$\frac{\partial \mathcal{L}}{\partial x} = x + A^\top \lambda = 0, \frac{\partial \mathcal{L}}{\partial \lambda} = Ax - b = 0.$$

写成矩阵形式为

$$\begin{bmatrix} I & A^\top \\ A & 0 \end{bmatrix} \begin{bmatrix} x \\ \lambda \end{bmatrix} = \begin{bmatrix} 0 \\ b \end{bmatrix}$$

如果 A (行) 满秩, 即 $\text{rank}(A) = m$, 则系数矩阵非奇异, 上述方程组存在唯一解.

1.2 超定方程组

当 $m > n$ 时, 线性方程组 $Ax = b$ 的解可能不存在. 记

$$J(x) = \|Ax - b\|_2^2$$

易知 $J(x)$ 是关于 x 的二次凸函数. 因此, x_* 是问题 (1) 的解当且仅当 x_* 是 $J(x)$ 的稳定点.

令其一阶导数为零, 可得

$$A^\top Ax - A^\top b = 0$$

于是 (1) 就转化为一个 (半正定) 线性方程组.

本讲我们主要讨论超线性最小二乘问题的求解.

2 初等变换矩阵

2.1 初等矩阵

2.2 Gauss 变换

2.3 Householder 变换

2.4 Givens 变换

2.5 正交变换的舍入误差分析

矩阵计算的一个基本思想就是把较复杂的问题转化为等价的较简单的, 易于求解问题. 而完成这个转化的基本工具就是初等变换矩阵, 其中常用的有三个: Gauss 变换, Householder 变换和 Givens 变换.

2.1 初等矩阵

我们考虑初等矩阵

$$E(u, v, \tau) = I - \tau uv^*$$

其中 $u, v \in \mathbb{C}^n$ 是非零向量, τ 是一个非零复数. 事实上, $E(u, v, \tau)$ 是单位矩阵的一个秩 1 扰动.

定理 设 $E(u, v, \tau)$ 是一个初等矩阵, 我们有

$$(1) \det(E(u, v, \tau)) = 1 - \tau v^* u$$

(2) 若 $1 - \tau v^* u \neq 0$, 则 $E(u, v, \tau)$ 非奇异, 且

$$(E(u, v, \tau))^{-1} = E(u, v, \tau)$$

其中 $\gamma = \frac{\tau}{\tau v^* u - 1}$

2.2 Gauss 变换

设 $l_j = [0, \dots, 0, l_{j+1,j}, \dots, l_{n,j}]^T, j = 1, 2, \dots, n$, 则 Gauss 变换定义为

$$L(l_j) \triangleq E(l_j, e_j, -1) = I + l_j e_j^T = \begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & l_{j+1,j} & 1 & \\ & & \vdots & & \ddots \\ & & l_{n,j} & & & 1 \end{bmatrix}$$

向量 l_j 称为 Gauss 向量, 由定理 2.1 可知

$$\det(L(l_j)) = 1, \quad (L(l_j))^{-1} = E(l_j, e_j, 1) = E(-l_j, e_j, -1) = L(-l_j)$$

Gauss 变换主要用于矩阵的 LU 分解

2.3 Householder 变换

定义 我们称矩阵

$$H = I - \frac{2}{v^* v} v v^* = I - \frac{2}{\|v\|_2^2} v v^*, \quad 0 \neq v \in \mathbb{C}^n \quad (3)$$

为 Householder 矩阵 (或 Householder 变换, 或 Householder 反射), 向量 v 称为 Householder 向量. 我们通常将矩阵 (3) 记为 $H(v)$

从几何上看, Householder 变换就是一个关于超平面 $\text{span}\{v\}^\perp$ 的反射, 对任意一个向量 $x \in \mathbb{C}^n$, 可将其写为

$$x = \frac{v^* x}{v^* v} v + y \triangleq \alpha v + y$$

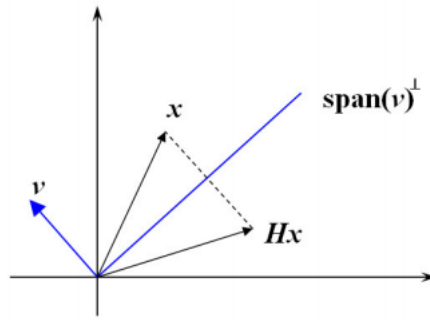


图 2.1 Householder 变换的几何意义

其中 $\alpha v \in \text{span}\{v\}, y \in \text{span}\{v\}^\perp$. 则

$$Hx = x - \frac{2}{v^*v}vv^*x = x - 2\alpha v = -\alpha v + y$$

即 Hx 与 x 在 $\text{span}\{v\}^\perp$ 方向有着相同的分量, 而在 v 方向的分量正好相差一个符号. 也就是说, Hx 是 x 关于超平面 $\text{span}\{v\}^\perp$ 的镜面反射, 见图 2.1. 因此, Householder 矩阵也称为反射矩阵.

Householder 矩阵的几个基本性质

定理 设 $H \in \mathbb{C}^{n \times n}$ 是一个 Householder 矩阵, 则 (1) $H^* = H$, 即 H Hermite 的;

(2) $H^*H = I$, 即 H 是酉矩阵;

(3) $H^2 = I$, 所以 $H^{-1} = H$;

(4) $\det(H) = -1$;

(5) H 有两个互异的特征值: $\lambda = 1$ 和 $\lambda = -1$, 其中 $\lambda = 1$ 的代数重数为 $n - 1$

Householder 矩阵的重要应用: 将一个向量除第一个元素以外的所有元素都化为零

引理 设 $x, y \in \mathbb{C}^n$ 为任意两个互异的向量, 则存在一个 Householder 矩阵 $H(x)$ 使得 $y = H(v)x$ 的充要条件是 $\|x\|_2 = \|y\|_2$ ($x^*y \in \mathbb{R}$ (取 $v = x - y$ 即可))

定理 设 $x = [x_1, x_2, \dots, x_n]^T \in \mathbb{R}^n$ 是一个非零向量, 则存在 Householder 矩阵 H 使得 $Hx = \alpha e_1$, 其中 $\alpha = \|x\|_2$ (或 $\alpha = -\|x\|_2$), $e_1 = [1, 0, \dots, 0]^T \in \mathbb{R}^n$

设 $x = [x_1, x_2, \dots, x_n]^T \in \mathbb{R}^n$ 是一个实的非零向量, 下面讨论如何计算定理中的 Householder 矩阵 $H(v)$. 由引理的证明过程可知

$$v = x - \alpha e_1 = [x_1 - \alpha, x_2, \dots, x_n]^T$$

在实际计算中, 为了尽可能地减少舍入误差, 我们通常避免两个相近的数做减法运算, 否则就会损失有效数字. 因此, 我通常取

$$\alpha = -\text{sign}(x_1) \cdot \|x\|_2$$

事实上, 我们也可以取 $\alpha = \text{sign}(x_1) \|x\|_2$ 但此时为了减少舍入误差, 我们需要通过下面的公式来计算 v 的第一个分量 v_1

$$\alpha = \text{sign}(x_1) \|x\|_2$$

$$v_1 = x_1 - \alpha = \frac{x_1^2 - \|x\|_2^2}{x_1 + \alpha} = \frac{-(x_2^2 + x_3^2 + \cdots + x_n^2)}{x_1 + \alpha}$$

$$v_1 = \begin{cases} x_1 - \alpha, & \text{if } \text{sign}(x_1) < 0 \\ \frac{-(x_2^2 + x_3^2 + \cdots + x_n^2)}{x_1 + \alpha}, & \text{otherwise} \end{cases}$$

无论怎样选取 α , 我们都有 $H = I - \beta vv^*$ 其中

$$\beta = \frac{2}{v^*v} = \frac{2}{(x_1 - \alpha)^2 + x_2^2 + \cdots + x_n^2} = \frac{2}{2\alpha^2 - 2\alpha x_1} = -\frac{1}{\alpha v_1}$$

算法 2.1 计算 Householder 向量

%Given $x \in \mathbb{R}^n$, compute β, v such that $Hx = \|x\|_2 e_1$ with $H = I - \beta vv^*$

```

1: function [ $\beta, v$ ] = house( $x$ )
2:  $n = \text{length}(x)$  (here  $\text{length}(x)$  denotes the dimension of  $x$ )
3:  $\sigma = x_2^2 + x_3^2 + \cdots + x_n^2$ 
4:  $v = x$ 
5: if  $\sigma = 0$  then
6:     if  $x_1 < 0$  then
7:          $v_1 = 2x, \beta = 2/v_1^2$ 
8:     else
9:          $v_1 = 0, \beta = 0$ 
10:    end if
11: else
12:     $\alpha = \sqrt{x_1^2 + \sigma}$  %  $\alpha = \|x\|_2$ 
13:    if  $x_1 < 0$  then
14:         $v_1 = x_1 - \alpha$ 
15:    else
16:         $v_1 = -\alpha/(x_1 + \alpha)$ 
17:    end if
18:     $\beta = 2/(v_1^2 + \alpha)$ 
19: end if

```

总运算量大约为 $3n$, 且具有很好的数值稳定性

在实际计算时, 我们可以将向量 v 单位化, 使得 $v_1 = 1$, 这样, 我们就无需为 v 另外分配空间, 而是将 $v(2:n)$ 存放在 $x(2:n)$ 中, 因为变换后的向量 x 除第一个分量外, 其它都为零.

为了避免可能产生的溢出, 可事先将 x 单位化, 即令 $x = x/\|x\|_2$

Householder 变换的运算量

设 $A \in \mathbb{R}^{m \times n}$, $H = I - \beta vv^* \in \mathbb{R}^m$, 则

$$HA = (I - \beta vv^*)A = A - \beta vv^*A = A - \beta v(A^*v)^*$$

因此, 在做 **Householder** 变换时, 并不需要生成 **Householder** 矩阵, 只需要 **Householder** 向量即可. 上面矩阵相乘的总运算量大约为 $4mn$.

2.4 Givens 变换

$$G(i, j, \theta) = \begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & c & & s \\ & & & \ddots & \\ & & -s & & c \\ & & & & \ddots & \\ & & & & & 1 \end{bmatrix} \in \mathbb{R}^{n \times n} \quad (\theta \in [0, 2\pi], i \leq j)$$

Givens 变换 (或 **Givens** 旋转, 或 **Givens** 矩阵), 其中 $c = \cos(\theta)$, $s = \sin(\theta)$

定理: $G(i, j, \theta)$ 是正交矩阵, 且 $\det(G(i, j, \theta)) = 1$

左乘 **Givens** 矩阵: 只会影响第 i 行和第 j 的元素.

右乘 **Givens** 矩阵: 只会影响第 i 和第 j 列的元素.

例 设 $x = [x_1, x_2]^T \in \mathbb{R}^2$, , 则存在一个 **Givens** 变换 $G = \begin{bmatrix} c & s \\ -s & c \end{bmatrix} \in \mathbb{R}^{2 \times 2}$ 使得

$Gx = [r, 0]^T$, 其中 c, s 和 r 的值如下:

- 若 $x_1 = x_2 = 0$, 则 $c = 1, s = 0, r = 0$;
- 若 $x_1 = 0$ 但 $x_2 \neq 0$, 则 $c = 0, s = x_2/|x_2|, r = |x_2|$
- 若 $x_1 \neq 0$ 但 $x_2 = 0$, 则 $c = \text{sign}(x_1), s = 0, r = |x_1|$;
- 若 $x_1 \neq 0$ 且 $x_2 \neq 0$, 则 $c = x_1/r, s = x_2/r, r = \sqrt{x_1^2 + x_2^2}$.

通过 Givens 变换, 我们可以将向量 $x \in \mathbb{R}^2$ 的第二个分量化为 0.

事实上, 对于任意一个向量 $x \in \mathbb{R}^n$, 我们都可以通过 Givens 变换将其任意一个位置上的分量化为 0. 更进一步, 我们也可以通过若干个 Givens 变换, 将 x 中除第一个分量外的所有元素都化为 0.

算法 2.2 Givens 变换

% Given $x = [a, b]^\top$, compute c, s such that $Gx = [r, 0]^\top$ where $r = \|x\|_2$

```

1: function  $[c, s] = \text{givens}(a, b)$ 
2: if  $b = 0$  then
3:     if  $a \geq 0$  then
4:          $c = 1, s = 0$ 
5:     else
6:          $c = -1, s = 0$ 
7:     end if
8: else
9:     if  $|b| > |a|$  then
10:         $\tau = a/b, \quad s = \text{sign}(b)/\sqrt{1 + \tau^2}, \quad c = s\tau$ 
11:    else
12:         $\tau = b/a, \quad c = \text{sign}(a)/\sqrt{1 + \tau^2}, \quad s = c\tau$ 
13:    end if
14: end if

```

2.5 正交变换的舍入误差分析

引理 设 $P \in \mathbb{R}^{n \times n}$ 是一个精确的 Householder 或 Givens 变换, \tilde{P} 是其浮点运算近似, 则

$$\text{fl}(\tilde{P}A) = P(A + E), \quad \text{fl}(A\tilde{P}) = (A + F)P$$

其中 $\|E\|_2 = \mathcal{O}(\varepsilon_u) \cdot \|A\|_2, \|F\|_2 = \mathcal{O}(\varepsilon_u) \cdot \|A\|_2$

这说明对一个矩阵做 Householder 变换或 Givens 变换是向后稳定的.

考虑对矩阵 A 做一系列的正交变换, 则有

$$\text{fl}\left(\tilde{P}_k \cdots \tilde{P}_1 A \tilde{Q}_1 \cdots \tilde{Q}_k\right) = P_k \cdots P_1 (A + E) Q_1 \cdots Q_k$$

其中 $\|E\|_2 = \mathcal{O}(\varepsilon_u) \cdot (k\|A\|_2)$, 这说明整个计算过程是向后稳定的

一般地, 假设 X 是一个非奇异的线性变换, \tilde{X} 是其浮点运算近似. 当 X 作用到 A 上时, 我们有

$$\text{fl}(\tilde{X}A) = XA + E = X(A + X^{-1}E) \triangleq X(A + F)$$

其中 $\|E\|_2 = \mathcal{O}(\varepsilon_u) \cdot \|XA\|_2 \leq \mathcal{O}(\varepsilon_u) \cdot \|X\|_2 \cdot \|A\|_2$, 故

$$\|F\|_2 = \|X^{-1}E\|_2 \leq \mathcal{O}(\varepsilon_u) \cdot \|X^{-1}\|_2 \cdot \|X\|_2 \cdot \|A\|_2 = \mathcal{O}(\varepsilon_u) \cdot \kappa_2(X) \cdot \|A\|_2$$

因此, 舍入误差将被放大 $\kappa_2(X)$ 倍. 当 X 是正交变换时, $\kappa_2(X)$ 达到最小值 1, 这就是为什么在浮点运算中尽量使用正交变换的原因.

3 QR 分解

3.1 QR 分解的存在唯一性

3.2 基于 MGS 的 QR 分解

3.3 基于 Householder 变换的 QR 分解

3.4 列主元 QR 分解

3.5 基于 Givens 变换的 QR 分解

3.6 QR 分解的稳定性

3.1 QR 分解的存在唯一性

定理 (QR 分解) 设 $A \in \mathbb{C}^{m \times n} (m \geq n)$. 则存在一个单位列正交矩阵 $Q \in \mathbb{C}^{m \times n}$ 和一个上三角矩阵 $R \in \mathbb{C}^{n \times n}$, 使得

$$A = QR \tag{4}$$

证明: 设 $A = [a_1, a_2, \dots, a_n] \in \mathbb{C}^{m \times n}$

若 A 列满秩, 即 $\text{rank}(A) = n$. 则 QR 分解 (4) 就是对 A 的列向量组进行 Gram-Schmidt 正交化过程的矩阵描述. 具体过程见下面的算法.

算法 3.1 Gram-schmidt Process

1: $r_{11} = \|a_1\|_2$

2: $q_1 = a_1/r_{11}$

3: for $j = 2$ to n do


```

4:    $q_j = a_j$ 
5:   for  $i = 1$  to  $j - 1$  do
6:        $r_{ij} = q_i^* a_j$ 
7:        $q_j = q_j - r_{ij} q_i$ 
8:   end for
9:    $r_{jj} = \|q_j\|_2$ 
10:   $q_j = q_j / r_{jj}$ 
11: end for

```

如果 A 不是列满秩, 我们可以做类似的正交化过程:

- 如果 $a_1 = 0$ 则令 $q_1 = 0$; 否则令 $q_1 = a_1 / \|a_1\|_2$;
- 对于 $j = 2, 3, \dots$ 计算 $\tilde{q}_j = a_j - \sum_{i=1}^{j-1} (q_i^* a_j) q_i$.
如果 $\tilde{q}_j = 0$ 则令 $q_j = 0$, 否则令 $q_j = \tilde{q}_j / \|\tilde{q}_j\|_2$

于是我们有

$$A = QR$$

其中 $Q = [q_1, q_2, \dots, q_n]$ 列正交 (但不是单位列正交, 列向量中可能有零向量). 这里的 $R = [r_{ij}]_{n \times n}$ 是上三角矩阵, 定义如下

$$r_{ij} = \begin{cases} q_i^* a_j, & \text{for } i \leq j \\ 0, & \text{for } i > j \end{cases}$$

易知, 如果 Q 的某一行 $q_k = 0$, 则 R 中对应的第 k 行全为 0

设 $\text{rank}(A) = l < n$, 则 $\text{rank}(Q) = l$ 即 Q 只有 l 个非零列, 不妨设为 $q_{i_1}, q_{i_2}, \dots, q_{i_l}$, 他们构成 \mathbb{C}^m 的一个单位正交向量组, 将其扩展成 \mathbb{C}^m 中的一组标准正交基, 即

$$q_{i_1}, q_{i_2}, \dots, q_{i_l}, \tilde{q}_1, \dots, \tilde{q}_{m-l}$$

然后我们用 \tilde{q}_1 替换 Q 中的第一个零列, 用 \tilde{q}_2 替换 Q 总的第二个零列, 依次类推, 将 Q 中的所有零列都替换掉. 将最后得到的矩阵记为 \tilde{Q} , 于是 $\tilde{Q} \in \mathbb{C}^{m \times n}$ 单位列正交. 由于 \tilde{Q} 中的新添加的列向量正好与 R 中的零行相对应, 所以

$$\tilde{Q}R = QR = A$$

这就是 A 的 QR 分解

满秩矩阵 QR 分解的存在唯一性

定理 若 A 列满秩, 并要求 R 的对角线元素都为正, 则 A 的 QR 分解存在且唯一.

† 若 A 是实矩阵, 则所有运算都是实运算, 因此 Q 和 R 都是实矩阵.

† 有时也将 QR 分解定义为: 存在酉矩阵 $Q \in \mathbb{C}^{m \times m}$ 使得

$$A = Q \begin{bmatrix} R_{11} \\ 0 \end{bmatrix}$$

其中 $R_{11} \in \mathbb{R}^{n \times n}$ 为上三角矩阵.

若 A 不满秩, 存在置换矩阵 P , 使得 AP 的前 l 列线性无关, 其中 $l = \text{rank}(A)$. 对 AP 进行 QR 分解, 可得:

推论 设 $A \in \mathbb{C}^{m \times n} (m \geq n)$ 则存在一个置换矩阵 P , 使得

$$AP = Q \begin{bmatrix} R_{11} & R_{12} \\ 0 & 0 \end{bmatrix}_{n \times n}$$

其中 $Q \in \mathbb{C}^{m \times n}$ 单位列正交, $R_{11} \in \mathbb{C}^{l \times l}$ 是非奇异上三角矩阵.

† 上述结论可简化为

$$AP = Q \begin{bmatrix} R_{11} & R_{12} \end{bmatrix}$$

其中 $Q \in \mathbb{C}^{m \times n}$ 单位列正交, $R_{11} \in \mathbb{C}^{l \times l}$ 是非奇异上三角矩阵.

推论 (满秩分解) 设 $A \in \mathbb{C}^{m \times n}$, 且 $\text{rank}(A) = l \leq \min\{m, n\}$, 则存在满秩矩阵 $F \in \mathbb{C}^{m \times l}$ 和 $G \in \mathbb{C}^{l \times n}$, 使得

$$A = FG$$

† 如果 A 是非奇异方阵, 则 QR 分解可用来求解线性方程组 $Ax = b$

† 基于 G-S 正交化的 QR 分解的运算量大约为 $2mn^2$.

在后面, 我们会介绍基于 Householder 变换的 QR 分解, 在不需要计算 Q 的情况下, 运算量大约为 $2mn^2 - 2n^3/3$, 如果需要计算 Q , 则需另外大约 $2mn^2 - 2n^3/3$ 运算量.

3.2 基于 MGS 的 QR 分解

由于数值稳定性方面的原因, 在实际计算中, 我们一般不采用 Gram-Schmidt 过程, 取而代之的是修正的 Gram-Schmidt 过程, 本算法的运算量大约为 $2mn^2$

算法 3.2 基于 MGS 的 QR 分解

% Given $A \in \mathbb{R}^{m \times n}$, compute $Q = [q_1, \dots, q_n]$ and R such that $A = QR$

1. Set $R = [r_{ik}] = 0_{n \times n}$ (the $n \times n$ zero matrix)

2. if $a_1 = 0$ then

```

3.       $q_1 = 0$ 
4. else
5.       $r_{11} = \|a_1\|_2$ 
6.       $q_1 = a_1 / \|a_1\|_2$ 
7. end if
8. for  $k = 2$  to  $n$  do
9.       $q_k = a_k$ 
10.   for  $i = 1$  to  $k - 1$  do
11.        $r_{ik} = q_i^T q_k$ 
12.        $q_k = q_k - r_{ik} q_i$ 
13.   end for
14.   if  $q_k \neq 0$  then
15.        $r_{kk} = \|q_k\|_2$ 
16.        $q_k = q_k / r_{kk}$ 
17.   end if
18. end for

```

† MGS 得到的 QR 分解中, $Q \in \mathbb{R}^{n \times n}$, $R \in \mathbb{R}^{n \times n}$

3.3 基于 Householder 变换的 QR 分解

由 Householder 变换的性质可知, 我们可以将任何一个非零变量 $x \in \mathbb{R}^n$ 转化成 $\|x\|_2 e_1$, 即除第一个元素外, 其他都为零.

假定 $m = n$, 即 $A \in \mathbb{R}^{n \times n}$ 令 $H_1 \in \mathbb{R}^{n \times n}$ 为 Householder 变换, 满足

$$H_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{bmatrix} = \begin{bmatrix} r_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

于是

$$H_1 A = \left[\begin{array}{c|ccc} r_1 & \tilde{a}_{12} & \cdots & \tilde{a}_{1n} \\ \hline 0 & & & \\ \vdots & & \tilde{A}_2 & \\ 0 & & & \end{array} \right], \tilde{A}_2 \in \mathbb{R}^{(n-1) \times (n-1)}$$

同样,构造 Householder 变换 $\tilde{H}_2 \in \mathbb{R}^{(n-1) \times (n-1)}$, 使得

$$\tilde{H}_2 \tilde{A}_2 = \left[\begin{array}{c|ccc} r_1 & \tilde{a}_{12} & \cdots & \tilde{a}_{1n} \\ \hline 0 & & & \\ \vdots & & \tilde{A}_2 & \\ 0 & & & \end{array} \right], \tilde{A}_2 \in \mathbb{R}^{(n-1) \times (n-1)}$$

令 $H_2 = \begin{bmatrix} 1 & 0 \\ 0 & \tilde{H}_2 \end{bmatrix} \in \mathbb{R}^{n \times n}$, 则

$$H_2 H_1 A = \left[\begin{array}{cc|ccc} r_1 & \tilde{a}_{12} & \tilde{a}_{13} & \cdots & \tilde{a}_{1n} \\ 0 & r_2 & \tilde{a}_{23} & \cdots & \tilde{a}_{2n} \\ \hline 0 & 0 & & & \\ \vdots & \vdots & & \tilde{A}_3 & \\ 0 & 0 & & & \end{array} \right]$$

不断重复上述过程,这样,我们就得到一系列的矩阵

$$H_k = \begin{bmatrix} I_{k-1} & 0 \\ 0 & \tilde{H}_k \end{bmatrix}, \quad k = 2, 3, \dots, n-1$$

使得

$$H_{n-1} \cdots H_2 H_1 A = \left[\begin{array}{cccc} r_1 & \tilde{a}_{12} & \cdots & \tilde{a}_{1n} \\ 0 & r_2 & \cdots & \tilde{a}_{2n} \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & r_n \end{array} \right] \triangleq R$$

令 $Q = (H_{n-1} \cdots H_2 H_1)^{-1} = H_1 H_2 \cdots H_{n-1}$, 则

$$A = (H_{n-1} \cdots H_2 H_1)^{-1} R = QR$$

如果不需要生成 Q , 则运算量大约为 $2mn^2 - 2/3n^3$.

矩阵 Q 的计算

可通过下面的算法实现

$$Q = I_n, \quad Q = QH_k, \quad k = 1, 2, \dots, n-1$$

若保留了所有的 Householder 向量, 则 Q 可以通过下面的[向后累积法](#)实现:

$$Q = I_n, \quad Q = H_k Q, \quad k = n-1, n-2, \dots, 1$$

优点: 一开始 Q 会比较稀疏, 随着迭代的进行, Q 才会慢慢变满

运算量大约为 $4(m^2n - mn^2 + \frac{1}{3}n^3)$.

† 若 $m > n$ 则由 Householder 变换得到的 QR 分解中, $Q \in \mathbb{R}^{m \times m}$, $R \in \mathbb{R}^{m \times n}$

算法 3.3 基于 Householder 变换的 QR 分解

%Given $A \in \mathbb{R}^{m \times n}$, compute $Q \in \mathbb{R}^{m \times m}$, $R \in \mathbb{R}^{m \times n}$ such that $A = QR$ % The upper triangular part of R is stored in the upper triangular part of A

```

1: Set  $Q = I_{m \times m}$ 
2: for  $k = 1$  to  $n$  do
3:      $x = A(k : m, k)$ 
4:      $[\beta, v_k] = \text{house}(x)$ 
5:      $v_k = v_k / \|v_k\|_2$ 
6:      $A(k : m, k : n) = (I_{m-k+1} - 2v_k v_k^T) A(k : m, k : n)$ 
7:      $Q(1 : k-1, k : m) = Q(1 : k-1, k : m) (I_{m-k+1} - 2v_k v_k^T)$ 
8:      $Q(k : m, k : m) = Q(k : m, k : m) (I_{m-k+1} - 2v_k v_k^T)$ 
9: end for

```

上面的算法只是一个简单描述, 并没有考虑运算量问题.

† 在实际计算时, 我们通常会保留所有的 Householder 向量.

由于第 k 步中 \hat{H}_k 所对应的 Householder 向量 v_k 的长度为 $m-k+1$, 因此我们先把 v_k 单位化, 使得 v_k 的第一元素为 1, 这样就只要存储 $v_k(2 : \text{end})$, 共 $m-k$ 个元素.

这样, 我们就可以把所有的 Householder 向量存放在 A 的严格下三角部分, 而 A 的上三角部分仍然存放 R .

在计算 Q 时采用向后累积的方法, 所以总的运算量大约为 $4m^2n - 2mn^2 + 2n^3$. 若 $m = n$, 则运算量大约为 $8n^3$.

† 我们也可以考虑分块 Householder QR 分解, 以便充分利用 3 级 BLAS 运算, 提高计算效率.

3.4 列主元 QR 分解

当 A 不是满秩时, 我们可以进行列主元 QR 分解.

定理 (列主元 QR 分解) 设 $A \in \mathbb{C}^{m \times n} (m \geq n)$, 且 $\text{rank}(A) = k < n$, 则存在置换矩阵 P ,

正交矩阵 $Q \in \mathbb{C}^{m \times m}$, 使得

$$AP = Q \begin{bmatrix} R_{11} & R_{12} \\ 0 & 0 \end{bmatrix}_{m \times n}$$

其中 $R_{11} \in \mathbb{C}^{k \times k}$ 是非奇异上三角矩阵, 且对角线元素满足 $r_{11} \geq r_{22} \geq \cdots \geq r_{kk} > 0$

列主元 QR 分解的实现过程与 QR 分解基本类似, 只是每一步需要选个列主元, 同时做一个列交换

假设经过 l 步后, 我们得到下面的分解

$$AP^{(l)} = Q^{(l)} \begin{bmatrix} R_{11}^{(l)} & R_{12}^{(l)} \\ 0 & R_{22}^{(l)} \end{bmatrix} \triangleq Q^{(l)} R^{(l)} P^{(l)} \quad (Q^{(l)})^\top AP^{(l)} = R^{(l)}$$

其中 $P^{(l)}$ 为置换矩阵, $Q^{(l)}$ 为正交矩阵, $R_{12}^{(l)} \in \mathbb{R}^{(l) \times (l)}$ 非奇异上三角矩阵. 下面进行第 $l+1$ 步:

1. 计算 $R_{22}^{(l)}$ 所有列的范数, 若都为 0, 则算法结束, 此时必有 $l = k$.
2. 若 $l < k$, 则 $R_{22}^{(l)} \neq 0$. 设范数最大的列为第 i_{l+1} 列 (若有相等的, 取其中一列即可), 范数为 $r_{l+1, l+1}$. 若 $i_{l+1} \neq 1$, 则交换 $R^{(l)}$ 的第 $l+1$ 列与第 i_{l+1} 列, 并记相应的置换矩阵为 P_{l+1} .
3. 以 $R_{22}^{(l)}$ 的第 1 列构造 Householder 变换 \bar{H}_{l+1}
令 $H_{l+1} = \text{Blkdiag}(I_l, \bar{H}_{l+1})$, $P^{(l+1)} = P^{(l)} P_{l+1}$. 则

$$H_{l+1} (Q^{(l)})^\top AP^{(l+1)} = H_{l+1} R^{(l)} P_{l+1} = \begin{bmatrix} R_{11}^{(l)} & \tilde{R}_{12}^{(l)} \\ 0 & \tilde{R}_{22}^{(l)} \end{bmatrix} \triangleq R^{(l+1)}$$

其中 $\tilde{R}_{22}^{(l)}$ 的第一列除第一个元素外, 其余都是零, 且该元素等于 $r_{l+1, l+1}$, 即

$$R^{(l+1)} = \left[\begin{array}{cc|c} R_{11}^{(l)} & * & * \\ 0 & r_{l+1, l+1} & * \\ \hline 0 & 0 & * \end{array} \right] \triangleq \begin{bmatrix} R_{11}^{(l+1)} & R_{12}^{(l+1)} \\ 0 & R_{22}^{(l+1)} \end{bmatrix}$$

其中 $R_{11}^{(l+1)} \in \mathbb{R}^{(l+1) \times (l+1)}$ 为非奇异上三角矩阵
记 $Q^{(l+1)} \triangleq Q^{(l)} H_{l+1}^\top$, 则

$$AP^{(l+1)} = Q^{(l+1)} R^{(l+1)}$$

依此类推, 直到第 k 步, 即可得 A 的列主元 QR 分解.
矩阵 R_{11} 的对角线元素的递减关系可由列主元的选取方法推出.

3.5 基于 Givens 变换的 QR 分解

我们同样可以利用 Givens 变换来做 QR 分解.

设 $A \in \mathbb{R}^{n \times n}$, 构造 Givens 变换 G_{21} , 作用在 A 的最前面的两行上, 使得

$$G_{21} \begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \\ \vdots \\ a_{n1} \end{bmatrix} = \begin{bmatrix} \tilde{a}_{11} \\ 0 \\ a_{31} \\ \vdots \\ a_{n1} \end{bmatrix}$$

构造 Givens 变换 G_{31} , 作用在 $G_{21}A$ 的第 1 行和第 3 行上, 将 a_{31} 化为零. 由于 G_{31} 只改变第 1 行和第 3 行的值, 所以第二行的零元素维持不变. 以此类推, 我们可以构造一系列的 Givens 变换 $G_{41}, G_{51}, \dots, G_{n1}$ 使得 $G_{n1} \cdots G_{21}A$ 的第一列中除第一个元素外, 其它元素都化为零, 即

$$G_{n1} \cdots G_{21}A = \begin{bmatrix} * & * & \cdots & * \\ 0 & * & \cdots & * \\ \vdots & \vdots & & \vdots \\ 0 & * & \cdots & * \end{bmatrix}$$

下面我们可以对第二列进行类似的处理. 构造 Givens 变换 $G_{32}, G_{42}, \dots, G_{n2}$ 将第二列的第 3 至第 n 个元素全化为零, 同时保持第一列不变.

其他列也做类似的处理. 最后, 通过构造 $\frac{1}{2}n(n-1)$ 个 Givens 变换, 将转化成一个上三角矩阵 R , 即

$$R = G_{n,n-1} \cdots G_{21}A, \square A = (G_{n,n-1} \cdots G_{21})^\top R \triangleq QR$$

† 与 Householder 变换一样, 在进行 Givens 变换时, 我们不需要显式地写出 Givens 矩阵.

† 对于稠密矩阵而言, 基于 Givens 变换的 QR 分解的运算量比 Householder 变换要多很多.

基于 Givens 变换的 QR 分解主要用于当矩阵的非零下三角元素相对较少时的情形, 比如对上 Hessenberg 矩阵进行 QR 分解

† 如果 $A \in \mathbb{R}^{m \times n}, m > n$, 仍然可以通过 Givens 变换进行 QR 分解

算法 3.4 基于 Givens 变换的 QR 分解

% Given $A \in \mathbb{R}^{m \times n}$, compute $Q \in \mathbb{R}^{m \times m}$ and $R \in \mathbb{R}^{m \times n}$ such that $A = QR$
 % The upper triangular part of R is stored in the upper triangular part of A

- 1: Set $Q = I_{m \times m}$
- 2: for $k = 1$ to n do
- 3: for $i = k + 1$ to m do
- 4: $[c, s] = \text{givens}(a_{kk}, a_{ik})$

5: $\begin{bmatrix} A(k, k:n) \\ A(i, k:n) \end{bmatrix} = G \begin{bmatrix} A(k, k:n) \\ A(i, k:n) \end{bmatrix}$ where $G = \begin{bmatrix} c & s \\ -s & c \end{bmatrix}$

6: $[Q(1:m, k), Q(1:m, i)] = [Q(1:m, k), Q(1:m, i)]G^T$

7: end for

8: end for

3.6 QR 分解的稳定性

基于 Householder 变换和 Givens 变换的 QR 分解都具有很好的数值稳定性, 基于 MGS 的 QR 分解也是向后稳定的

† 如果需要计算 Q, 则 MGS 的运算量相对较少, 因此当 A 的列向量具有很好的线性无关性时, 我们可以使用 MGS 来计算 QR 分解.

Björck [Björck 1967] 证明了, 通过 MGS 计算的矩阵 Q 满足

$$Q^T Q = I + E_{MGS}, \quad \|E_{MGS}\|_2 \approx \varepsilon_u \kappa_2(A)$$

而 Householder 变换计算的矩阵 Q 满足

$$Q^T Q = I + E_H, \quad \|E_H\|_2 \approx \varepsilon_u$$

因此, 如果正交性至关重要, 建议使用 Householder 变换, 特别是当 A 的列向量接近线性相关时

4 奇异值分解

定理 (SVD) 设 $A \in \mathbb{C}^{m \times n} (m \geq n)$ 则存在酉矩阵 $U \in \mathbb{C}^{m \times m}$ 和 $V \in \mathbb{C}^{n \times n}$, 使得

$$U^* A V = \begin{bmatrix} \Sigma \\ 0 \end{bmatrix}, \text{ or } A = U \begin{bmatrix} \Sigma \\ 0 \end{bmatrix} V^* \quad (5)$$

其中 $\Sigma = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_n) \in \mathbb{R}^{n \times n}$, 且 $\sigma_1 \geq \dots \geq \sigma_n \geq 0$

† 该定理也可以通过 Hermite 半正定矩阵的特征值分解来证明.

定义 分解 (3.6) 称为 A 的**奇异值分解 (SVD)**, $\sigma_1, \sigma_2, \dots, \sigma_n$ 称为 A 的奇异值, 它们是矩阵 $A^* A$ 的特征值的平方根.

† 在不作特别说明的情况下, 我们总是假定 $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n \geq 0$

† 如果 $A \in \mathbb{R}^{m \times n}$ 是实矩阵, 则 U, V 也都可以是实矩阵 [Horn-Johnson1985].

奇异向量

矩阵 $U = [u_1, u_2, \dots, u_m]$ 和 $V = [v_1, v_2, \dots, v_n]$ 的列向量分别称为 A 的**左奇异向量**和**右**

奇异向量, 即存在关系式

$$\begin{aligned} Av_i &= \sigma_i u_i, \quad i = 1, 2, \dots, n \\ A^* u_i &= \sigma_i v_i, \quad i = 1, 2, \dots, n \\ A^* u_i &= 0, \quad i = n+1, n+2, \dots, m \end{aligned}$$

SVD 的其他形式

$$A = U \begin{bmatrix} \Sigma \\ 0 \end{bmatrix} V^* = \sigma_1 u_1 v_1^* + \sigma_2 u_2 v_2^* + \dots + \sigma_n u_n v_n^*$$

记 $U_n = [u_1, u_2, \dots, u_n] \in \mathbb{C}^{m \times n}$, 则 U_n 是单位列正交矩阵, 且

$$A = U_n \Sigma V^* \quad (6)$$

这就是所谓的细 SVD (thin SVD) 或降阶 SVD (reduced SVD), 有的文献将上式称为奇异值分解.

当 $k < n$ 时, 我们称

$$A_k = \sigma_1 u_1 v_1^* + \sigma_2 u_2 v_2^* + \dots + \sigma_k u_k v_k^*$$

为 A 的截断 SVD (truncated SVD).

奇异值分解基本性质

定理 设 $A = U \begin{bmatrix} \Sigma \\ 0 \end{bmatrix} V^*$ 是 $A \in \mathbb{C}^{m \times n} (m \geq n)$ 的奇异值分解, 则

- (1) $A^* A$ 的特征值是 σ_i^2 , 对应的特征向量是 v_i
- (2) AA^* 的特征值是 σ_i^2 和 $m - n$ 个零, 对应的特征向量是 u_i
- (3) $\|A\|_2 = \sigma_1, \|A\|_F = \sqrt{\sigma_1^2 + \sigma_2^2 + \dots + \sigma_n^2}$
- (4) 若 $\text{rank}(A) = r \leq n$, 则 $\text{Ran}(A) = \text{span}\{u_1, u_2, \dots, u_r\}, \text{Ker}(A) = \text{span}\{v_{r+1}, v_{r+2}, \dots, v_n\}$
- (5) 设 $x \in \mathbb{C}^n$, 且 $\|x\|_2 = 1$ 和 $Y \in \mathbb{C}^{n \times n}$ 是酉矩阵, 则 $\sigma_i(X^* A Y) = \sigma_i(A)$
- (6) (酉不变性) $X \in \mathbb{C}^{m \times m}$ 和 $Y \in \mathbb{C}^{n \times n}$ 是酉矩阵, 则 $\sigma_i(X^* A Y) = \sigma_i(A)$

定理 设 $A = U \Sigma V^*$ 是 $A \in \mathbb{C}^{n \times n}$ 的奇异值分解, 则:

- (1) $|\det(A)| = \sigma_1 \sigma_2 \dots \sigma_n$
- (2) 若 A 非奇异, 则 $\|A^{-1}\|_2 = \sigma_n^{-1}, \kappa_2(A) = \sigma_1 / \sigma_n$
- (3) 若 A 是 Hermite 的, 且 $A = U \Lambda U^*$ 是 A 的西特征值分解, 设 $|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_n|$, 则 $A = U \Sigma V$ 是 A 的奇异值分解, 其中 $\sigma_i = |\lambda_i|, v_i = \text{sign}(\lambda_i) u_i$, 若 $\lambda_i = 0$, 则 $v_i = u_i$

(4) 矩阵 $\begin{bmatrix} 0 & A^* \\ A & 0 \end{bmatrix}$ 的特征值是 $\pm\sigma_i$, 对应单位特征向量 $\frac{1}{\sqrt{2}} \begin{bmatrix} v_i \\ \pm u_i \end{bmatrix}$

† 矩阵条件数取决于奇异值, 不是特征值, 见讲义上的例子

低秩逼近

定理 设 $A = U_n \Sigma V^*$ 是 $A \in \mathbb{C}^{m \times n}$ 的奇异值分解. 令

$$A_k = \sum_{i=1}^k \sigma_i u_i v_i^*$$

则 A_k 是

$$\min_{B \in \mathbb{C}^{m \times n}, \text{rank}(B)=k} \|A - B\|_2 \quad (7)$$

的一个解, 且

$$\|A - A_k\|_2 = \sigma_{k+1}$$

此时, 我们称 A_k 是 A 的一个秩 k 逼近

† 对于 Frobenius 范数, 我们有类似的结论.

定理 (Weyl) 设 $A, B \in \mathbb{C}^{m \times n} (m \geq n)$ 且 $\text{rank}(B) = k$, 则有

$$\max_{x \in \text{Ker}(B), \|x\|_2=1} \|Ax\|_2 \geq \sigma_{k+1}(A) \quad (8)$$

和

$$\min_{x \in \text{Ker}(B), \|x\|_2=1} \|Ax\|_2 \leq \sigma_{n-k}(A) \quad (9)$$

因此,

$$\sigma_1(A - B) \geq \sigma_{k+1}(A), \quad \sigma_n(A + B) \leq \sigma_{n-k}(A) \quad (10)$$

且

$$\sigma_{i+j-1}(A) \leq \sigma_i(B) + \sigma_j(A - B), \quad i = 1, 2, \dots, n, j = 1, 2, \dots, n - i + 1 \quad (11)$$

根据 Weyl 定理, 我们可以得到矩阵奇异值的最小最大定理.

定理 设 $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n \geq 0$ 是矩阵 $A \in \mathbb{C}^{m \times n} (m \geq n)$ 的奇异值, 则

$$\sigma_k(A) = \min_{\dim(\mathcal{L})=n-k+1} \max_{x \in \mathcal{L}, \|x\|_2=1} \|Ax\|_2, \quad k = 1, 2, \dots, n$$

和

$$\sigma_k(A) = \max_{\dim(\mathcal{L})=k} \min_{x \in \mathcal{L}, \|x\|_2=1} \|Ax\|_2, \quad k = 1, 2, \dots, n$$

其中 \mathcal{L} 表示 \mathbb{C}^n 的一个子空间.

引理 (交错不等式) 设 $A \in \mathbb{C}^{m \times n}$, A_r 是由 A 去除 r 列 (或 r 行) 后得到的子矩阵, 则

$$\sigma_i(A) \geq \sigma_i(A_r) \geq \sigma_{i+r}(A), \quad i = 1, 2, \dots, \min\{m, n\}$$

这里, 当下标 k 大于矩阵的维数时, 我们令 $\sigma_k = 0$. 更进一步, 如果 $B \in \mathbb{C}^{(m-r) \times (n-s)}$ 是 A 的子矩阵, 则

$$\sigma_i(A) \geq \sigma_i(B) \geq \sigma_{i+r+s}(A)$$

引理 设 $A \in \mathbb{C}^{n \times n}$, $1 \leq k \leq n$ 则对任意的单位列正交矩阵 $U_k \in \mathbb{C}^{n \times k}$ 和 $V_k \in \mathbb{C}^{n \times k}$, 有

$$\sigma_i(U_k^* A V_k) \leq \sigma_i(A), \quad i = 1, 2, \dots, k \quad (12)$$

因此,

$$|\det(U_k^* A V_k)| \leq \sigma_1(A) \sigma_2(A) \cdots \sigma_k(A) \quad (13)$$

定理 (Weyl 不等式, 1949) 设 $A \in \mathbb{C}^{n \times n}$ 其奇异值和特征值分别为 $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0$ 和 $|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_n|$ 则

$$|\lambda_1 \lambda_2 \cdots \lambda_k| \leq \sigma_1 \sigma_2 \cdots \sigma_k, \quad k = 1, 2, \dots, n$$

且当 $k = n$ 时, 等号成立.

奇异值扰动分析

定理 设 $A, E \in \mathbb{C}^{m \times n} (m \geq n)$, 则

$$|\sigma_i(A + E) - \sigma_i(A)| \leq \|E\|_2, \quad i = 1, 2, \dots, n$$

下面是上述定理的一个推论, 也是 SVD 的一个重要应用

推论 设 $A \in \mathbb{C}^{n \times n}$, $\|\cdot\|$ 是任一相容矩阵范数, 则对任意的 $\varepsilon > 0$ 总存在矩阵 A_ε , 使得 $\|A - A_\varepsilon\| \leq \varepsilon$, 其中 A_ε 具有互不相同的特征值.

因此, 可对角化矩阵在所有矩阵组成的集合中是稠密的

5 线性最小二乘问题的求解方法

5.1 正规方程法

5.2 QR 分解法

5.3 奇异值分解法

5.1 正规方程法

定理 设 $A \in \mathbb{R}^{m \times n} (m \geq n)$, 则 $x_* \in \mathbb{R}^n$ 是线性最小二乘问题的解当且仅当残量 $r = b - Ax_*$ 与 $\text{Ran}(A)$ 正交, 即 x_* 是下面的正规方程的解

$$A^\top(b - Ax) = 0 \quad \text{或} \quad A^\top A x = A^\top b \quad (14)$$

由于

$$A^\top b \in \text{Ran}(A^\top) = \text{Ran}(A^\top A)$$

因此正规方程 $A^\top A x = A^\top b$ 总是相容的, 即最小二乘解总是存在的, 且当 A 满秩时, 这个解也是唯一的.

定理 设 $A \in \mathbb{R}^{m \times n} (m > n)$, 则 $A^\top A$ 对称正定当且仅当 A 是列满秩的, 即 $\text{rank}(A) = n$, 此时, 线性最小二乘问题的解是唯一的, 其表达式为

$$x = (A^\top A)^{-1} A^\top b$$

当 A 列满秩时, 我们就可以使用 **Cholesky** 分解来求解正规方程, 总的运算量大约为 $mn^2 + 1/3n^3 + \mathcal{O}(n^2)$, 其中大部分的运算量 (mn^2) 是用来计算 $A^\top A$ (由于 $A^\top A$ 对称, 因此只需计算其下三角部分).

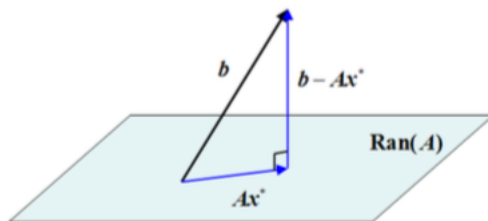
最小二乘的几何含义

我们可以把 b 写成

$$b = Ax_* + r, \text{ 其中 } r \perp \text{Ran}(A) \quad (15)$$

所以 Ax_* 就是 b 在 $\text{Ran}(A)$ 上的正交投影:

Part I title



† 最小二乘解可能并不唯一, 但 b 的上述分解总是唯一的.

最小二乘与鞍点问题

线性最小二乘问题 (3.1) 等价于

$$A^\top r = 0, \quad r = b - Ax$$

即

$$\begin{bmatrix} I & A \\ A^\top & 0 \end{bmatrix} \begin{bmatrix} r \\ x \end{bmatrix} = \begin{bmatrix} b \\ 0 \end{bmatrix} \quad (16)$$

这就是线性最小二乘问题 (3.1) 的**增广方程** (augmented system). 事实上, 方程组 (16) 是下面方程组的一种特殊情形

$$\begin{bmatrix} B & A \\ A^\top & 0 \end{bmatrix} \begin{bmatrix} r \\ x \end{bmatrix} = \begin{bmatrix} f \\ g \end{bmatrix}$$

其中 $B \in \mathbb{R}^{m \times m}$ 对称半正定. 这就是通常所说的鞍点问题. 这个方程组存在唯一解当且仅当 A 列满秩且矩阵 $[B, A]$ 行满秩.

如果 B 对称正定, 则 $r = B^{-1}(f - Ax)$, 代入第二个方程可得

$$A^T B^{-1} Ax = A^T B^{-1} f - g$$

这就是广义的正规方程. 其所对应的广义线性最小二乘问题是

$$\min_{x \in \mathbb{R}^n} \frac{1}{2} \|Ax - f\|_{B^{-1}}^2 + g^T x$$

其中范数 $\|\cdot\|_{B^{-1}}$ 的定义是 $\|x\|_{B^{-1}}^2 = x^T B^{-1} x$.

5.2 QR 分解法

假定 $A \in \mathbb{R}^{m \times n} (m \geq n)$ 满秩, 用三种不同方法来推导

(1) 将 Q 的扩充成一个正交矩阵, 记为 $[Q, \hat{Q}] \in \mathbb{R}^{m \times m}$ 于是有

$$\begin{aligned} \|Ax - b\|_2^2 &= \left\| [Q, \hat{Q}]^T (Ax - b) \right\|_2^2 \\ &= \left\| [Q, \hat{Q}]^T (QRx - b) \right\|_2^2 \\ &= \left\| \begin{bmatrix} Rx - Q^T b \\ -\hat{Q}^T b \end{bmatrix} \right\|_2^2 \\ &= \|Rx - Q^T b\|_2^2 + \|\hat{Q}^T b\|_2^2 \geq \|\hat{Q}^T b\|_2^2 \end{aligned}$$

等号成立当且仅当 $Rx = Q^T b$ 所以最小二乘解为

$$x_* = R^{-1} Q^T b$$

(2) 将 b 写成 $b = (QQ^T + I - QQ^T)b$, 则

$$\begin{aligned} Ax - b &= Ax - (QQ^T + I - QQ^T)b \\ &= (Ax - QQ^T b) - (I - QQ^T)b \end{aligned}$$

由于 QQ^T 是 $\text{Ran}(A)$ 上的正交投影, 因此 $Ax - QQ^T b$ 与 $(I - QQ^T)b$ 正交. 所以

$$\begin{aligned} \|Ax - b\|_2^2 &= \|Ax - QQ^T b\|_2^2 + \|(I - QQ^T)b\|_2^2 \\ &= \|Rx - Q^T b\|_2^2 + \|(I - QQ^T)b\|_2^2 \\ &\geq \|(I - QQ^T)b\|_2^2 \end{aligned}$$

等号成立当且仅当 $Rx = Q^T b$ 所以最小二乘解为

$$x_* = R^{-1} Q^T b$$

(3) 解正规方程. 由定理 5.1 可知, 最小二乘解为

$$\begin{aligned} x_* &= (A^T A)^{-1} A^T b \\ &= (R^T Q^T Q R)^{-1} R^T Q^T b \\ &= (R^T R)^{-1} R^T Q^T b \\ &= R^{-1} Q^T b \end{aligned}$$

† 用 QR 分解来求最小二乘解的运算量大约为 $2mn^2$ (如果采用 Householder 变换的话, 运算量大约为 $2mn^2 - 2/3n^3$). 当 $m \gg n$ 时, 大约为正规方程的两倍. 当 $m = n$ 时, 几乎相同.

† 通常 QR 算法比较稳定, 是求解最小二乘问题的首选方法, 特别是当 A 条件数较大 (病态) 时.

5.3 奇异值分解法

设 $A \in \mathbb{R}^{m \times n}$ 列满秩, $A = U \begin{bmatrix} \Sigma \\ 0 \end{bmatrix} V^\top$ 是 A 的奇异值分解, $U = [U_n, \tilde{U}]$ 则

$$\begin{aligned} \|Ax - b\|_2^2 &= \left\| U \begin{bmatrix} \Sigma \\ 0 \end{bmatrix} V^\top x - b \right\|_2^2 = \left\| \begin{bmatrix} \Sigma \\ 0 \end{bmatrix} V^\top x - [U_n, \tilde{U}]^\top b \right\|_2^2 \\ &= \left\| \begin{bmatrix} \Sigma V^\top x - U_n^\top b \\ -\tilde{U}^\top b \end{bmatrix} \right\|_2^2 \\ &= \|\Sigma V^\top x - U_n^\top b\|_2^2 + \|\tilde{U}^\top b\|_2^2 \geq \|\tilde{U}^\top b\|_2^2 \end{aligned}$$

等号当且仅当 $\Sigma V^\top x - U_n^\top b = 0$ 时成立, 即

$$x = (\Sigma V^\top)^{-1} U_n^\top b = V \Sigma^{-1} U_n^\top b$$

这就是线性最小二乘问题 (3.1) 的解.

6 广义逆与最小二乘

6.1 广义逆的定义

6.2 广义逆的基本性质

6.3 广义逆的计算

6.4 广义逆与线性最小二乘

6.1 广义逆的定义

广义逆的概念最早由 Moore 于 1920 年提出:

设 $A \in \mathbb{C}^{m \times n}$, 若 $X \in \mathbb{C}^{n \times m}$ 满足

$$AX = P_{\text{Ran}(A)}, \quad XA = P_{\text{Ran}(X)} \quad (17)$$

即 AX 和 XA 分别为 $\text{Ran}(A)$ 和 $\text{Ran}(X)$ 上的正交投影, 则称 X 是 A 的广义逆.

1955 年, Penrose 利用四个矩阵方程给出了广义逆的定义.

定义 设 $A \in \mathbb{C}^{m \times n}$, 若 $X \in \mathbb{C}^{n \times m}$ 满足

$$AXA = A \quad (18)$$

$$XAX = X \quad (19)$$

$$(AX)^* = AX \quad (20)$$

$$(XA)^* = XA \quad (21)$$

则称 X 为 A 的**广义逆** (或 **Moore-Penrose 逆**, 简称 **MP 逆**), 记为 A^\dagger

† 可以证明, 以上两个定义是等价的.

† 若 $A \in \mathbb{C}^{m \times n}$ 非奇异, 则 $A^\dagger = A^{-1}$

定义 设 $A \in \mathbb{C}^{m \times n}$ 则满足矩阵方程 (18)-(21) 的矩阵 $X \in \mathbb{C}^{n \times m}$ 存在唯一, 即广义逆存在且惟一.

广义逆与 SVD

设 $A \in \mathbb{C}^{m \times n}$ $\text{rank}(A) = r > 0$, A 的 SVD 为

$$A = U \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix} V^*, \quad \Sigma_1 = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_r) \in \mathbb{R}^{r \times r}$$

则容易验证

$$A^\dagger = V \begin{bmatrix} \Sigma_1^{-1} & 0 \\ 0 & 0 \end{bmatrix} U^*$$

6.2 广义逆的基本性质

定理 设 $A \in \mathbb{C}^{m \times n}$, 则

- (1) $(A^\dagger)^\dagger = A$
- (2) $(A^T)^\dagger = (A^\dagger)^T, (A^*)^\dagger = (A^\dagger)^*$
- (3) $\text{rank}(A) = \text{rank}(A^\dagger) = \text{rank}(A^\dagger A)$
- (4) $(AA^*)^\dagger = (A^*)^\dagger A^\dagger, (A^*A)^\dagger = A^\dagger (A^*)^\dagger$
- (5) $(AA^*)^\dagger AA^* = AA^\dagger, (A^*A)^\dagger A^*A = A^\dagger A$
- (6) $A^\dagger = (A^*A)^\dagger A^* = A^* (AA^*)^\dagger$ 特别地, 若 A 列满秩, 则 $A^\dagger = (A^*A)^{-1} A^*$. 若 A 行满秩, 则 $A^\dagger = A^* (AA^*)^{-1}$

(7) 若 U, V 是酉矩阵, 则 $(UAV)^\dagger = V^* A^\dagger U^*$

† 一般来说, 当 A, B 是方阵时,

- $(AB)^\dagger \neq B^\dagger A^\dagger$
- $AA^\dagger \neq A^\dagger A$
- $(A^k)^\dagger \neq (A^\dagger)^k$
- A 和 A^\dagger 的非零特征值并不是互为倒数

设 $A \in \mathbb{C}^{m \times n}$, 则

$$\begin{aligned}\text{Ran}(AA^\dagger) &= \text{Ran}(AA^*) = \text{Ran}(A) \\ \text{Ran}(A^\dagger A) &= \text{Ran}(A^* A) = \text{Ran}(A^*) = \text{Ran}(A^\dagger) \\ \text{Ker}(AA^\dagger) &= \text{Ker}(AA^*) = \text{Ker}(A^*) = \text{Ker}(A^\dagger) \\ \text{Ker}(A^\dagger A) &= \text{Ker}(A^* A) = \text{Ker}(A)\end{aligned}$$

6.3 广义逆的计算

- 利用 SVD, 但运算量较大.
- 利用满秩分解

定理 设 $A \in \mathbb{R}^{m \times n}$

- (1) 若 A 是列满秩矩阵, 则 $A^\dagger = (A^* A)^{-1} A^*$
- (2) 若 A 是行满秩矩阵, 则 $A^\dagger = A^* (A A^*)^{-1}$
- (3) 若 A 的秩是 $r \leq \min\{m, n\}$ 且其满秩分解为 $A = FG$ 其中 $F \in \mathbb{R}^{m \times r}, G \in \mathbb{R}^{r \times n}$, 则

$$A^\dagger = G^\dagger F^\dagger = G^* (G G^*)^{-1} (F^* F)^{-1} F^*$$

- 利用 QR 分解

定理 设 $A \in \mathbb{R}^{m \times n} (m \geq n)$ 是列满秩矩阵, 其 QR 分解为 $A = QR$, 其中 $Q \in \mathbb{R}^{m \times n}, R \in \mathbb{R}^{n \times n}$, 则

$$A^\dagger = R^{-1} Q^*$$

- 其他算法

其他比较重要的算法有: Greville 递推算法, Cline 算法等.

6.4 广义逆与线性最小二乘

定理 设 $A \in \mathbb{R}^{m \times n}$ 则线性最小二乘问题的解为

$$x = A^\dagger b + (I - P_{A^\top}) z, \quad \forall z \in \mathbb{R}^n \quad (22)$$

† 通常, 线性最小二乘问题的解 (22) 不是唯一的, 但当 A 列满秩时, $P_{A^\top} = I$, 此时解唯一.

定理 设 $A \in \mathbb{R}^{m \times n}$ 的解集为 \mathcal{S} , 则

$$\min_{x \in \mathcal{S}} \|x\|_2 \quad (23)$$

存在唯一解, 即满足 (23) 的线性最小二乘问题 (3.1) 的解存在且唯一

7 最小二乘扰动分析

定理 设 $A \in \mathbb{R}^{m \times n}$ 且 $\text{rank}(A) = n$ 设 x 是线性最小二乘问题 (3.1) 的解, \tilde{x} 极小化 $\|(A + \delta A)\tilde{x} - (b + \delta b)\|_2$, 则

$$\frac{\|\tilde{x} - x\|_2}{\|x\|_2} \leq \varepsilon \cdot \left\{ \frac{2\kappa_2(A)}{\cos \theta} + \kappa_2^2(A) \tan \theta \right\} + O(\varepsilon^2)$$

其中 $\kappa_2(A) = \sigma_1(A)/\sigma_n(A)$, θ 为 b 与 $\text{Ran}(A)$ 的夹角,

$$\varepsilon \triangleq \max \left\{ \frac{\|\delta A\|_2}{\|A\|_2}, \frac{\|\delta b\|_2}{\|b\|_2} \right\}$$

并假定 $\varepsilon \cdot \kappa_2(A) < 1$ (确保 $A + \delta A$ 满秩, 从而 \tilde{x} 唯一确定) 我们记

$$\kappa_{LS} \triangleq \frac{2\kappa_2(A)}{\cos \theta} + \kappa_2^2(A) \tan \theta$$

这就是最小二乘问题的条件数. 当 $\theta = 0, b \in \text{Ran}(A)$, 此时 $\kappa_{LS} = 2\kappa_2(A)$, 当 $\theta = \pi/2$ 时, $b \perp \text{Ran}(A)$, 此时最小二乘解为 $x = 0$, 而 $\kappa_{LS} = \infty$; 当 $0 < \theta < \pi/2$ 时, $\kappa_{LS} = O(\kappa_2^2(A))$

定义残量 $r = b - Ax$, $\tilde{r} = (b + \delta b) - (A + \delta A)\tilde{x}$, 我们有下面的性质 [Higham2002]

$$\frac{\|\tilde{r} - r\|_2}{\|r\|_2} \leq 1 + 2\varepsilon \cdot \kappa_2(A)$$

当我们使用 QR 分解或 SVD 分解求解最小二乘问题时, 由于采用的是正交变换, 它们都是数值稳定的. 而正规方程涉及求解方程组 $A^\top Ax = A^\top b$, 其精度依赖于条件数 $\kappa_2(A^\top A) = \kappa_2^2(A)$, 因为其误差是以 $\kappa_2^2(A)$ 倍数增长. 因此当 A 的条件数较大时, 正规方程法的精度会大大降低.

8 最小二乘问题的推广及其应用

8.1 正则化

8.2 加权正则化

8.3 约束最小二乘问题

8.4 应用: 多项式数据拟合

8.5 应用: 线性预测

8.6 应用: 信号处理

8.1 正则化

在求解超定线性方程组时, 我们极小化 $\|Ax - b\|_2^2$ 而对于欠定线性方程组, 我们极小化的是 $\|x\|_2^2$. 两者合起来就是

$$\min_{x \in \mathbb{R}^n \times n} \frac{1}{2} \|Ax - b\|_2^2 + \frac{\alpha}{2} \|x\|_2^2 \quad (24)$$

其中 $\alpha > 0$ 是权系数. 对应的目标函数记为

$$J(x) = \|Ax - b\|_2^2 + \alpha \|x\|_2^2$$

当 $\alpha > 0$ 时, $J(x)$ 是一个严格凸的二次函数, 因此存在唯一的最小值点. 关于 x 求导后, 令其等于零, 可得

$$(A^\top A + \alpha I) x = A^\top b$$

由于 $A^\top A + \alpha I$ 对称正定, 故非奇异. 所以问题 (24) 的唯一解为

$$x = (A^\top A + \alpha I)^{-1} A^\top b$$

8.2 加权正则化

一类应用更广泛的问题是下面的加权正则化问题

$$\min_{x \in \mathbb{R}^n \times n} \frac{1}{2} \|Ax - b\|_2^2 + \frac{\alpha}{2} \|Wx\|_2^2 \quad (25)$$

其中 $W \in \mathbb{R}^{p \times n}$ 是广义加权矩阵, 可以是非负对角矩阵, 对称正定矩阵, 也可以是一般矩阵 (如一阶差分算子, 二阶差分算子等). 注意, W 不一定要是方阵. 经过类似的推导, 问题 (25) 的解满足

$$(A^\top A + \alpha W^\top W) x = A^\top b$$

如果 $A^\top A + \alpha W^\top W$ 非奇异, 则存在唯一解

$$x = (A^\top A + \alpha W^\top W)^{-1} A^\top b$$

8.3 约束最小二乘问题

考虑带有约束的最小二乘问题

$$\begin{aligned} \min_{x \in \mathbb{R}^n \times n} & \frac{1}{2} \|Ax - b\|_2^2 \\ \text{s.t.} & Bx = f \end{aligned} \quad (26)$$

其中 $Bx = f$ 是约束条件. 对应的 Lagrange 函数为

$$J(x) = \frac{1}{2} \|Ax - b\|_2^2 + \lambda^\top (Bx - f)$$

分别对 x 和 λ 求一阶导数, 并令其等于零, 可得

$$\begin{bmatrix} A^T A & B^T \\ B & 0 \end{bmatrix} \begin{bmatrix} x \\ \lambda \end{bmatrix} = \begin{bmatrix} A^T b \\ f \end{bmatrix}$$

如果 $A^T A$ 非奇异, 且 B 行满秩, 则存在唯一解

$$\begin{aligned} \lambda &= \left[B (A^T A)^{-1} B^T \right]^{-1} \left[B (A^T A)^{-1} A^T b - f \right] \\ x &= (A^T A)^{-1} (A^T b - B^T \lambda) \end{aligned}$$

8.4 应用: 多项式数据拟合

已知 n 个点 $\{(t_i, f_i)\}_{i=1}^n$, 寻找一个低次多项式来拟合这些数据. 设拟合多项式为

$$p(x) = a_0 + a_1 t + a_2 t^2 + \cdots + a_m t^m$$

通常 $m \ll n$ 将上述 n 个点代入可得

$$\begin{bmatrix} 1 & t_1 & t_1^2 & \cdots & t_1^m \\ 1 & t_2 & t_2^2 & \cdots & t_2^m \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & t_n & t_n^2 & \cdots & t_n^m \end{bmatrix}_{n \times (m+1)} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_m \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{bmatrix} \quad \text{or} \quad Ax = f$$

其中 $A \in \mathbb{R}^{n \times (m+1)}$, $x = [a_0, a_1, \dots, a_m]^T$, 由于 $m \ll n$ 该方程组是超定的, 解通常是不存在的. 因此, 我们寻找一个近似解, 使得残量 $\|f - Ax\|_2$ 最小, 即求解最小二乘问题

$$\min_{x \in \mathbb{R}^{m+1}} \|f - Ax\|_2^2$$

8.5 应用: 线性预测

预测一个时间序列的未来走向, 一个常用方法就是线性预测. 假定在 t_k 时刻的值 f_k 线性依赖于其前 m 个时刻的值 $f_{k-1}, f_{k-2}, \dots, f_{k-m}$, 即

$$f_k = a_1 f_{k-1} + a_2 f_{k-2} + \cdots + a_m f_{k-m} \quad (27)$$

现在已经测得该时间序列的前 n 个值 $f_i, i = 0, 1, 2, \dots, n-1$, 其中 $n \gg m$. 如何预测其未来的取值? 将现有的数据代入上述关系式可得

$$\begin{bmatrix} f_{m-1} & f_{m-2} & f_{m-3} & \cdots & f_0 \\ f_m & f_{m-1} & f_{m-2} & \cdots & f_1 \\ \vdots & \vdots & \vdots & & \vdots \\ f_{n-2} & f_{n-3} & f_{n-4} & \cdots & f_{n-m-1} \end{bmatrix}_{(n-m) \times m} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix} = \begin{bmatrix} f_m \\ f_{m+1} \\ \vdots \\ f_{n-1} \end{bmatrix} \quad \square Ax = f$$

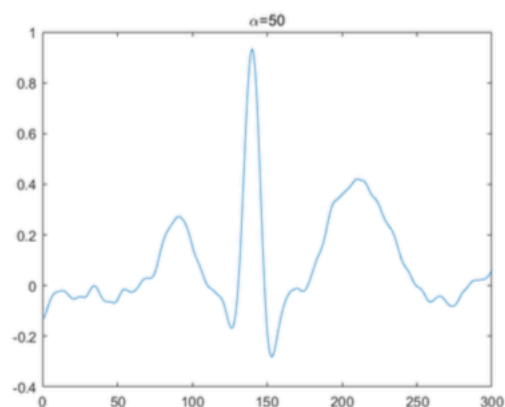
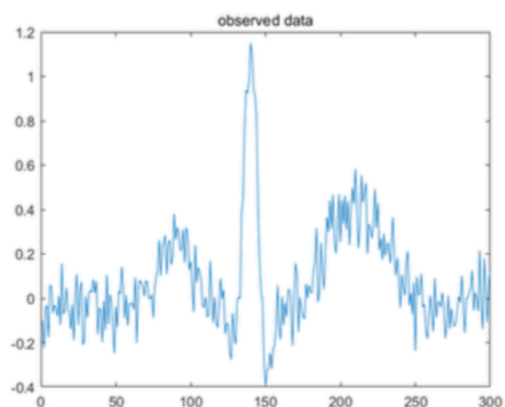
这也是一个超定问题, 因此需要求解最小二乘 $\min_{x \in \mathbb{R}^m} \|f - Ax\|_2^2$

8.6 应用：信号处理

- 信号去噪：在获取数字信号时, 由于各种各样的原因, 最后得到的信号总会带有一定的噪声. 去噪是数字信号和图像处理中的一个基本问题. 其中一个有效方法就是加权最小二乘法.

$$\min_x \frac{1}{2} \|x - b\|_2^2 + \frac{1}{2} \alpha \|Dx\|_2^2$$

其中 D 是离散的二阶导算子.



- 图像去模糊
一般模型:

$$f = K(x) + n$$

其中 x 表示真实图像, f 表示观察到的图像, $K(\cdot)$ 是一个卷积算子, 代表模糊机制, n 代表噪声.

- 图像压缩