第三讲　线性最小二乘问题

— QR 分解

什么是 QR 分解

$$\boxed{A = QR}$$

QR 分解

QR 分解是将一个矩阵分解一个正交矩阵 (酉矩阵) 和一个三角矩阵的乘积. QR 分解被广泛应用于线性最小二乘问题的求解和矩阵特征值的计算.

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 $Q \in \mathbb{C}^{m\times n}$ (即 $Q^*Q = I_{n\times n}$) 和一个上三角矩阵 $R \in \mathbb{C}^{n\times n}$, 使得

$$\boxed{A = QR} \tag{3.1}$$

若 A 列满秩, 则存在一个具有正对角线元素的上三角矩阵 R 使得 (3.1) 成立, 且此时 QR 分解唯一, 即 Q 和 R 都唯一.

(板书)

算法 1.1 Gram-Schmidt 过程

$r_{11} = \|a_1\|_2$
$q_1 = a_1 / r_{11}$
for $j = 2$ to n **do**
    $q_j = a_j$
    for $i = 1$ to $j - 1$ **do**
        $r_{ij} = q_i^* a_j$ % q_i^* 表示共轭转置
        $q_j = q_j - r_{ij} q_i$
    end for
    $r_{jj} = \|q_j\|_2$
    $q_j = q_j / r_{jj}$
end for

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

$$\boxed{A = QR},$$

其中 $R = \begin{bmatrix} R_{11} \\ 0 \end{bmatrix} \in \mathbb{C}^{m\times n}$ 是上三角矩阵.

% Q 是方阵, R 是长方阵.

A 不是满秩矩阵情形

存在一个置换矩阵 P, 使得 AP 的前 l 列是线性无关的, 其中 $l = \mathrm{rank}(A)$.

因此我们可以对 AP 进行 QR 分解.

A 不是满秩矩阵情形

存在一个置换矩阵 P, 使得 AP 的前 l 列是线性无关的, 其中 $l = \text{rank}(A)$.

因此我们可以对 AP 进行 QR 分解.

推论 设 $A \in \mathbb{C}^{m\times n}$ $(m \geq n)$, 且秩为 l $(0 \leq l \leq n)$. 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.

A 不是满秩矩阵情形

存在一个置换矩阵 P, 使得 AP 的前 l 列是线性无关的, 其中 $l = \mathrm{rank}(A)$.

因此我们可以对 AP 进行 QR 分解.

推论 设 $A \in \mathbb{C}^{m\times n}$ $(m \geq n)$, 且秩为 l $(0 \leq l \leq n)$. 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.

✍ 上述结论也可简化为

$$AP = \tilde{Q} \begin{bmatrix} R_{11} & R_{12} \end{bmatrix}, \quad \text{其中} \quad \tilde{Q} \in \mathbb{C}^{m\times l}$$

几点注记

✍ 通过 GS 正交化过程实现的 QR 分解中, Q 和 R 都可能是复矩阵.

若 A 是实矩阵, 则 Q 和 R 都可以是实矩阵.

✍ 如果 A 是非奇异的方阵, 则 QR 分解也可以用来求解线性方程组 $Ax = b$.

✍ 基于 GS 正交化的 QR 分解算法 1.1 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 $4m^2n - 4mn^2 + 4n^3/3$ 运算量.

怎么计算 QR 分解

三种方案

MGS， Householder， Givens

方案一：基于 MGS

在证明 QR 分解的存在性时, 利用了 Gram-Schmidt 正交化过程.

但由于数值稳定性方面的原因, 在实际计算中, 不能直接采用 Gram-Schmidt 过程, 取而代之的是 **修正的 Gram-Schmidt 过程** (modified Gram-Schmidt process), 即 **MGS**.

算法 1.2 基于 MGS 的 QR 分解

Set $R = [r_{ik}] = 0_{n\times n}$ (the $n \times n$ zero matrix)
if $a_1 = 0$ **then**
    $q_1 = 0$
else
    $r_{11} = \|a_1\|_2$, $q_1 = a_1/\|a_1\|_2$
end if
for $k = 2$ to n **do**
    $q_k = a_k$
    for $i = 1$ to $k - 1$ **do**   % MGS 过程, 注意与 GS 的区别
        $r_{ik} = q_i^{\intercal} q_k$
        $q_k = q_k - r_{ik} q_i$
    end for
    if $q_k \neq 0$ **then**
        $r_{kk} = \|q_k\|_2$
        $q_k = q_k / r_{kk}$
    end if
end for

方案二：基于 Householder 变换

Householder 变换的重要应用: 将任何一个非零变量 $x \in \mathbb{R}^n$ 转化成 $\pm\|x\|_2 e_1$

方案二：基于 Householder 变换

Householder 变换的重要应用: 将任何一个非零变量 $x \in \mathbb{R}^n$ 转化成 $\pm\|x\|_2 e_1$

不失一般性, 考虑 $m = n$ 时的情形.

设矩阵 $A \in \mathbb{R}^{n\times n}$, 令 $H_1 \in \mathbb{R}^{n\times n}$ 为一个 Householder 变换, 满足

$$H_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{bmatrix} = \begin{bmatrix} r_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \Longrightarrow \quad H_1 A = \left[\begin{array}{c|ccc} r_1 & \tilde{a}_{12} & \cdots & \tilde{a}_{1n} \\ \hline 0 & & & \\ \vdots & & \tilde{A}_2 & \\ 0 & & & \end{array}\right],$$

其中 $\tilde{A}_2 \in \mathbb{R}^{(n-1)\times(n-1)}$.

构造一个 Householder 变换 $\tilde{H}_2 \in \mathbb{R}^{(n-1)\times(n-1)}$, 将 $\tilde{A}_2$ 的第一列中除第一个元素外的所有元素都化为 0, 可得

$$\tilde{H}_2 \tilde{A}_2 = \left[\begin{array}{c|ccc} r_2 & \tilde{a}_{23} & \cdots & \tilde{a}_{2n} \\ \hline 0 & & & \\ \vdots & & \tilde{A}_3 & \\ 0 & & & \end{array}\right].$$

令

$$H_2 = \begin{bmatrix} 1 & 0 \\ 0 & \tilde{H}_2 \end{bmatrix} \quad \Longrightarrow \quad H_2 H_1 A = \left[\begin{array}{cc|ccc} r_1 & \tilde{a}_{12} & \tilde{a}_{13} & \cdots & \tilde{a}_{1n} \\ 0 & r_2 & \tilde{a}_{23} & \cdots & \tilde{a}_{2n} \\ \hline 0 & 0 & & & \\ \vdots & \vdots & & \tilde{A}_3 & \\ 0 & 0 & & & \end{array}\right].$$

不断重复上述过程. 这样, 我们就得到一系列的矩阵

$$H_k = \begin{bmatrix} I_{k-1} & 0 \\ 0 & \tilde{H}_k \end{bmatrix}, \quad k = 1, 2, 3, 4, \ldots, n-1$$

使得

$$H_{n-1} \cdots H_2 H_1 A = \begin{bmatrix} r_1 & \tilde{a}_{12} & \cdots & \tilde{a}_{1n} \\ 0 & r_2 & \cdots & \tilde{a}_{2n} \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & r_n \end{bmatrix} \triangleq R.$$

由于 Householder 变换都是正交矩阵, 因此 $H_1, H_2, \ldots, H_{n-1}$ 也都是正交矩阵. 令

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

则 Q 也是正交矩阵, 且

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

矩阵 Q 的计算

矩阵 Q 可通过下面的算法实现

$$\begin{cases} Q = I_n, \\ Q = QH_k, \quad k = 1, 2, \ldots, n-1. \end{cases}$$

✍ 如果 $m > n$, 我们仍然可以通过上面的过程进行 QR 分解, 只是最后我们得到一个正交矩阵 $Q \in \mathbb{R}^{m\times m}$ 和一个上三角矩阵 $R \in \mathbb{R}^{m\times n}$, 使得 $A = QR$.

✍ 如果不需要生成 Q, 则基于 Householder 变换的 QR 分解的总运算量大约为 $2mn^2 - 2n^3/3$.

算法 1.3 基于 Householder 变换的 QR 分解 (MATLAB)

% The upper triangular part of R is stored in the upper triangular part of A

Set $Q = I_{m\times m}$

for $k = 1$ to n **do**

$\quad x = A(k:m, k)$

$\quad [\beta,\ v_k] = \textbf{House}(x)$

$\quad A(k:m, k:n) = (I_{m-k+1} - \beta v_k v_k^{\mathsf{T}})A(k:m, k:n)$

$\quad\qquad = A(k:m, k:n) - \beta v_k\big(v_k^{\mathsf{T}} A(k:m, k:n)\big)$

$\quad Q(:, k:m) = Q(:, k:m)(I_{m-k+1} - \beta v_k v_k^{\mathsf{T}})$

$\quad\qquad = Q(:, k:m) - \beta\big(Q(:, k:m)v_k\big)v_k^{\mathsf{T}}$

end for

✍ 上面的算法只是关于利用 Householder 变换来实现 QR 分解的一个简单描述, 并没有考虑运算量问题.

方案三：基于 Givens 变换

我们同样可以利用 Givens 变换来做 QR 分解.

(板书)

方案三：基于 Givens 变换

我们同样可以利用 Givens 变换来做 QR 分解.

(板书)

✍ 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. 因此基于 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 $A \in \mathbb{R}^{m\times n}$，其中 $m > n$，我们仍然可以通过 Givens 变换进行 QR 分解.

算法 1.4 基于 Givens 变换的 QR 分解

% The upper triangular part of R is stored in the upper triangular part of A

```
Set Q = I_{m×m}
for k = 1 to n do
    for i = k+1 to m do
        [c, s]=givens(a_{kk}, a_{ik})
        [A(k, k:n); A(i, k:n)] = G [A(k, k:n); A(i, k:n)]  where G = [c s; -s c]
        [Q(1:m, k), Q(1:m, i)] = [Q(1:m, k), Q(1:m, i)] G^T
    end for
end for
```

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

$$[Q(1:m,k),\, Q(1:m,i)] = [Q(1:m,k),\, Q(1:m,i)]\, G^{\intercal}$$

QR 分解的稳定性

通过 MGS 计算的矩阵 Q 满足

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

而 Householder 变换计算的矩阵 Q 满足

$$Q^{\mathsf{T}} Q = I + E_H \quad \text{其中} \quad \|E_H\|_2 \approx \varepsilon_u.$$

QR 分解的稳定性

通过 MGS 计算的矩阵 Q 满足

$$Q^\intercal Q = I + E_{MGS} \quad \text{其中} \quad \|E_{MGS}\|_2 \approx \varepsilon_u \kappa_2(A).$$

而 Householder 变换计算的矩阵 Q 满足

$$Q^\intercal Q = I + E_H \quad \text{其中} \quad \|E_H\|_2 \approx \varepsilon_u.$$

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.

QR 分解的稳定性

通过 MGS 计算的矩阵 Q 满足

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

而 Householder 变换计算的矩阵 Q 满足

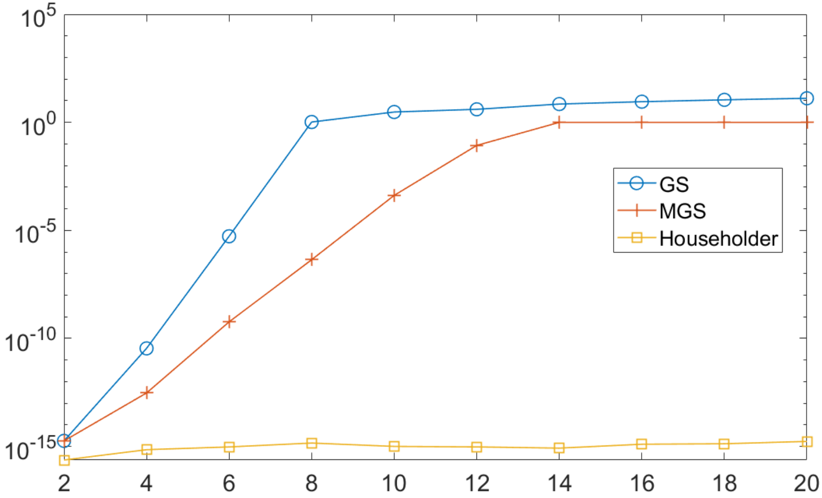
$$Q^{\mathsf{T}}Q = I + E_H \quad \text{其中} \quad \|E_H\|_2 \approx \varepsilon_u.$$

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.

- 当需要计算矩阵 Q 时, 基于 MGS 的 QR 分解的运算量相对较少. 当 A 的列向量具有很好的线性无关性时, 可以使用 MGS 来计算 QR 分解.
- 如果正交性至关重要, 则当 A 的列向量接近线性相关时, 建议使用 Householder 变换.

例 编写程序，分别用 GS, MGS 和 Householder 变换计算 n 阶 Hilbert 矩阵 H 的 QR 分解，并比较三种算法的稳定性，即观察 $\|\tilde{Q}\tilde{R}-H\|_2$ 和 $\|\tilde{Q}^\intercal\tilde{Q}-I\|_2$ 的值，其中 $\tilde{Q}$ 和 $\tilde{R}$ 是计算出来的 QR 分解矩阵因子. (LS_QR_stability.m)

	GS		MGS		Householder	
n	$\|\tilde{Q}\tilde{R}-H\|_2$	$\|\tilde{Q}^\intercal\tilde{Q}-I\|_2$	$\|\tilde{Q}\tilde{R}-H\|_2$	$\|\tilde{Q}^\intercal\tilde{Q}-I\|_2$	$\|\tilde{Q}\tilde{R}-H\|_2$	$\|\tilde{Q}^\intercal\tilde{Q}-I\|_2$
2	0.00e+00	1.81e-15	0.00e+00	1.81e-15	1.24e-16	2.36e-16
4	3.93e-17	3.45e-11	5.55e-17	2.98e-13	2.46e-16	7.08e-16
6	6.21e-17	5.33e-06	2.78e-17	5.81e-10	1.49e-16	9.49e-16
8	7.48e-17	1.03e+00	6.59e-17	4.38e-07	2.57e-16	1.44e-15
10	6.52e-17	3.00e+00	7.49e-17	4.16e-04	6.36e-16	1.00e-15
12	7.87e-17	4.00e+00	7.55e-17	8.52e-02	4.68e-16	9.52e-16
14	6.54e-17	7.00e+00	8.54e-17	9.96e-01	5.71e-16	8.35e-16



10^{5}
10^{0}
10^{-5}
10^{-10}
10^{-15}
2
4
6
8
10
12
14
16
18
20
GS
MGS
Householder