统计计算

2016 年 3 月 10 日

目录

目录

目录

QR 分解的定义

设 $\mathbf{A} \in \mathbb{C}^{m\times n}$. 如果 $\mathbf{A}$ 有分解

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

其中 $\mathbf{Q} \in \mathbb{C}^{m\times n}$ 为列正交矩阵, $\mathbf{R} \in \mathbb{C}^{n\times n}$ 为上三角矩阵, 那么称该分解为 $\mathbf{A}$ 的 QR 分解.

QR 分解的存在唯一性

定理

设 $\mathbf{A} \in \mathbb{C}^{m\times n}$. 若 $m \geq n$, 则存在列正交矩阵 $\mathbf{Q} \in \mathbb{C}^{m\times n}$ 和上三角矩阵 $\mathbf{R} \in \mathbb{C}^{n\times n}$, 使得 $\mathbf{A} = \mathbf{QR}$.

若 $m = n$, 则 $\mathbf{Q}$ 为酉阵.

进一步, 如果 $\mathbf{A}$ 非奇异, 可以选取 $\mathbf{R}$ 为具有正对角元素的上三角矩阵, 并且在这种情形下, 因子 $\mathbf{Q}$ 和 $\mathbf{R}$ 都是唯一的.

求解线性方程组 $\mathbf{Ax} = \mathbf{b}$.

目录

对于一般情形的 $\mathbf{A}$, 可用 Gram-Schmidt 算法的推广形式给出 $\mathbf{A}$ 的 QR 分解式.

设 $\mathbf{A} = [\mathbf{x}_1, \cdots, \mathbf{x}_n] \in \mathbb{R}^{m\times n}$ 为满列秩矩阵, 且有如下 QR 分解:

$$\mathbf{A} = [\mathbf{q}_1, \mathbf{q}_2, \cdots, \mathbf{q}_n] \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ & r_{22} & \cdots & r_{2n} \\ & & \ddots & \vdots \\ & & & r_{nn} \end{bmatrix}.$$

由上式两边各个列向量对应相等，可得向量方程组

$$\begin{cases} \mathbf{x}_1 &= r_{11}\mathbf{q}_1, \\ \mathbf{x}_2 &= r_{12}\mathbf{q}_1 + r_{22}\mathbf{q}_2, \\ \cdots & \cdots \quad \cdots \\ \mathbf{x}_n &= r_{1n}\mathbf{q}_1 + \cdots + r_{nn}\mathbf{q}_n. \end{cases} \tag{1}$$

由方程组 (1) 的第一个方程可求出 r_{11} 和 $\mathbf{q}_1$;
用 $\mathbf{q}_1^{\mathrm{H}}$ 左乘第二个方程，可求出 $r_{12} = \mathbf{q}_1^{\mathrm{H}}\mathbf{x}_2$ 以及 r_{22} 和 $\mathbf{q}_2$;
依此类推，最后用 $\mathbf{q}_1^{\mathrm{H}}, \cdots, \mathbf{q}_{n-1}^{\mathrm{H}}$ 左乘最后一个方程，依次可求出 $r_{1n}, \cdots, r_{n-1,n}$ 以及 r_{nn} 和 $\mathbf{q}_n$.

Gram-Schmidt 算法 (续)

计算顺序

$$
\begin{aligned}
& r_{11} \to \mathbf{q}_1 \\
\to & r_{12} \to r_{22} \to \mathbf{q}_2 \\
\to & \cdots \\
\to & r_{1n} \to r_{2n} \to \cdots \to r_{nn} \to \mathbf{q}_n
\end{aligned}
$$

设 $\mathbf{A} = [\mathbf{x}_1, \cdots, \mathbf{x}_n]$ 为满列秩矩阵.

(1) $r_{11} = |\mathbf{x}_1|$, $\mathbf{q}_1 = \mathbf{x}_1 / r_{11}$;

(2) 对 $i = 2, \cdots, n$ 计算

$$\begin{cases} r_{ki} = \mathbf{q}_k^{\mathrm{H}} \mathbf{x}_i, \quad k = 1, \cdots, i-1, \\ r_{ii} = \left| \mathbf{x}_i - \sum_{k=1}^{i-1} r_{ki} \mathbf{q}_k \right|, \\ \mathbf{q}_i = \left(\mathbf{x}_i - \sum_{k=1}^{i-1} r_{ki} \mathbf{q}_k \right) \Big/ r_{ii}. \end{cases}$$

经典 Gram-Schmidt 正交化算法的主要缺点是其数值性能不好.

例: 设

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1+\varepsilon \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{x}_2 = \begin{bmatrix} 1 \\ 1 \\ 1+\varepsilon \\ 1 \end{bmatrix}, \quad \mathbf{x}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1+\varepsilon \end{bmatrix},$$

其中 $0 < |\varepsilon| \ll 1$. 假设机器计算精度使得在计算机上 $1+\varepsilon$ 可以精确表示, 但 $1+2\varepsilon+\varepsilon^2$ 被舍入成 $1+2\varepsilon$.

由 Gram-Schmidt 算法可得未经单位化的向量

$$\mathbf{q}_1 = \begin{bmatrix} 1 \\ 1+\varepsilon \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{q}_2 = \begin{bmatrix} 0 \\ -\varepsilon \\ \varepsilon \\ 0 \end{bmatrix}, \quad \mathbf{q}_3 = \begin{bmatrix} 0 \\ -\varepsilon \\ 0 \\ \varepsilon \end{bmatrix}.$$

可以计算 $\dfrac{\mathbf{q}_1^{\mathrm{T}}\mathbf{q}_2}{|\mathbf{q}_1||\mathbf{q}_2|} \approx 0$, 且 $\dfrac{\mathbf{q}_1^{\mathrm{T}}\mathbf{q}_3}{|\mathbf{q}_1||\mathbf{q}_3|} \approx 0$, 但是 $\dfrac{\mathbf{q}_2^{\mathrm{T}}\mathbf{q}_3}{|\mathbf{q}_2||\mathbf{q}_3|} = \dfrac{1}{2}$.

考虑到经典 Gram-Schmidt 正交化算法的上述不足, 人们提出一种修正的 Gram-Schmidt 正交化算法, 也是利用向量方程组 (1), 但与前述算法计算顺序不一样的是, 修正的算法依次求出方程组 (1) 中等式右边的各列.

记 $\mathbf{x}_i^{(1)} = \mathbf{x}_i (i = 1, \cdots, n)$.

同 Gram-Schmidt 过程, 由方程组 (1) 中第一个方程可求出 r_{11} 和 $\mathbf{q}_1$.

用 $\mathbf{q}_1^{\mathrm{H}}$ 左乘其它 $n-1$ 个方程, 利用正交条件可求出 $r_{1k}(k = 2, \cdots, n)$. 于是得到各个方程右边第一项.

令

$$\mathbf{x}_k^{(2)} = \mathbf{x}_k^{(1)} - r_{1k}\mathbf{q}_1, \quad k = 2, \cdots, n,$$

那么方程组 (1) 化为

$$\begin{cases} \mathbf{x}_2^{(2)} = r_{22}\mathbf{q}_2, \\ \mathbf{x}_3^{(2)} = r_{23}\mathbf{q}_2 + r_{33}\mathbf{q}_3, \\ \cdots \quad \cdots \quad \cdots \\ \mathbf{x}_n^{(2)} = r_{2n}\mathbf{q}_2 + \cdots + r_{nn}\mathbf{q}_n. \end{cases} \tag{2}$$

用上面相同的方法可求出方程组 (2) 中各个方程右边第一项 (即方程组 (1) 中各个方程右边第二项), 依次继续下去, 即可得到 $\mathbf{Q}$ 和 $\mathbf{R}$.

修正的 Gram-Schmidt 算法 (续)

计算顺序

$$
\begin{aligned}
& r_{11} \to \mathbf{q}_1 \to r_{12} \to \cdots \to r_{1,n-1} \to r_{1n} \\
\to & r_{22} \to \mathbf{q}_2 \to r_{23} \to \cdots \to r_{2n} \\
\to & \cdots \\
\to & r_{nn} \to \mathbf{q}_n
\end{aligned}
$$

修正的 Gram-Schmidt 算法 (续)

设 $\mathbf{A} = [\mathbf{x}_1, \cdots, \mathbf{x}_n]$ 满列秩. 记 $\mathbf{x}_i^{(1)} = \mathbf{x}_i$.

1. 对 $i = 1, \cdots, n-1$ 计算

$$\begin{cases} r_{ii} = \left|\mathbf{x}_i^{(i)}\right|, \\ \mathbf{q}_i = \mathbf{x}_i^{(i)} / r_{ii}, \\ r_{ik} = \mathbf{q}_i^{\mathrm{H}} \mathbf{x}_k^{(i)}, \quad k = i+1, \cdots, n, \end{cases}$$

   以及

$$\mathbf{x}_k^{(i+1)} = \mathbf{x}_k^{(i)} - r_{ik}\mathbf{q}_i, \quad k = i+1, \cdots, n.$$

2. 计算 $r_{nn} = \left|\mathbf{x}_n^{(n)}\right|, \mathbf{q}_n = \mathbf{x}_n^{(n)} / r_{nn}$.

修正的 Gram-Schmidt 算法 (续)

续上例. 用修正的 Gram-Schmidt 算法可得

$$\mathbf{q}_1 = \begin{bmatrix} 1 \\ 1+\varepsilon \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{q}_2 = \begin{bmatrix} 0 \\ -\varepsilon \\ \varepsilon \\ 0 \end{bmatrix}, \quad \mathbf{q}_3 = \begin{bmatrix} 0 \\ -\varepsilon/2 \\ -\varepsilon/2 \\ \varepsilon \end{bmatrix}.$$

于是 $\mathbf{q}_2^{\mathrm{T}}\mathbf{q}_3 = 0$, 即 $\mathbf{q}_2$ 与 $\mathbf{q}_3$ 正交.

作业

给出上述两个例子的详细计算过程.

算法方面的注解

1. 虽然上面的两种算法在数学上是等价的，但从数值计算方面来说，修正的 Gram-Schmidt 算法具有更多优点．特别是当 $\mathbf{x}_1, \cdots, \mathbf{x}_n$ 中某些向量接近平行时，修正算法比原算法得到的向量组更接近于正交．由此可见，计算顺序的不同有时会大大影响计算结果．

2. 尽管修正算法有时也会得到不太理想的正交阵，但经验结果表明：对大多数问题而言，修正算法的计算精度及稳定性比原算法更优．

算法方面的注解

1. 虽然上面的两种算法在数学上是等价的，但从数值计算方面来说，修正的 Gram-Schmidt 算法具有更多优点. 特别是当 $\mathbf{x}_1, \cdots, \mathbf{x}_n$ 中某些向量接近平行时，修正算法比原算法得到的向量组更接近于正交. 由此可见，计算顺序的不同有时会大大影响计算结果.
2. 尽管修正算法有时也会得到不太理想的正交阵，但经验结果表明：对大多数问题而言，修正算法的计算精度及稳定性比原算法更优.

问　题?

目录

Householder 变换的定义

在平面 $\mathbb{R}^2$ 中将非零向量 $\mathbf{x}$ 映射为关于 $\mathbf{e}_1$ 轴对称 (换一种表达为关于 "与 $\mathbf{e}_2$ 轴正交的直线" 对称) 的向量 $\mathbf{y}$ 的变换, 称为关于 $\mathbf{e}_1$ 轴的镜像 (反射) 变换.

设 $\mathbf{x} = [x_1, x_2]^{\mathrm{T}}$, 则有

$$\mathbf{y} = \begin{bmatrix} x_1 \\ -x_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \mathbf{x} = (\mathbf{I} - 2\mathbf{e}_2\mathbf{e}_2^{\mathrm{T}})\mathbf{x}.$$

Householder 变换定义 (续)

一般地, 将向量 $\mathbf{x}$ 映射为关于 “与单位向量 $\mathbf{u}$ 正交的直线” 对称的向量 $\mathbf{y}$ 的镜像变换可描述为

$$\mathbf{y} = \mathbf{x} - 2\mathbf{u}(\mathbf{u}^{\mathrm{T}}\mathbf{x}) = (\mathbf{I} - 2\mathbf{u}\mathbf{u}^{\mathrm{T}})\mathbf{x}.$$

设单位向量 $\mathbf{u} \in \mathbb{R}^n$, 称

$$\mathbf{H}_{\mathbf{u}} = \mathbf{I} - 2\,\mathbf{u}\mathbf{u}^{\mathrm{T}}$$

为 Householder 矩阵 (初等反射矩阵).

Householder 变换的性质

1. $\mathbf{H}^{\mathrm{T}} = \mathbf{H}$ (对称矩阵);
2. $\mathbf{H}^{\mathrm{T}}\mathbf{H} = \mathbf{I}$ (正交矩阵);
3. $\mathbf{H}^2 = \mathbf{I}$ (对合矩阵);
4. $\mathbf{H}^{-1} = \mathbf{H}$ (自逆矩阵);
5. $\det \mathbf{H} = -1$;
6. 分块对角矩阵 $\mathrm{diag}(\mathbf{I}_m, \mathbf{H}, \mathbf{I}_n)$ 为 Householder 矩阵.

引理

对任意非零向量 $\mathbf{x} \in \mathbb{R}^n (n > 1)$，存在 Householder 矩阵 $\mathbf{H}$，使得 $\mathbf{Hx} = |\mathbf{x}|\mathbf{e}_1$，其中 $\mathbf{e}_1 \in \mathbb{R}^n$ 为标准单位向量.

证明　若 $\mathbf{x} = |\mathbf{x}|\mathbf{e}_1$，取与 $\mathbf{x}$ 正交的单位列向量 $\mathbf{u}$，则有

$$\mathbf{H_u x} = (\mathbf{I} - 2\,\mathbf{u}\mathbf{u}^{\mathrm{H}})\mathbf{x} = \mathbf{x} - 2\,\mathbf{u}\langle \mathbf{u}, \mathbf{x} \rangle = \mathbf{x} = |\mathbf{x}|\mathbf{e}_1.$$

若 $\mathbf{x} \neq |\mathbf{x}|\mathbf{e}_1$，令

$$\mathbf{u} = \frac{\mathbf{x} - |\mathbf{x}|\mathbf{e}_1}{\left|\mathbf{x} - |\mathbf{x}|\mathbf{e}_1\right|}.$$

由 $\left|\mathbf{x}-|\mathbf{x}|\mathbf{e}_1\right|^2 = 2\left\langle \mathbf{x}-|\mathbf{x}|\mathbf{e}_1, \mathbf{x}\right\rangle$ 可知

$$\begin{aligned}
\mathbf{H_u x} &= \mathbf{x} - 2\mathbf{u}\langle \mathbf{u}, \mathbf{x}\rangle \\
&= \mathbf{x} - 2\left\langle \mathbf{x}-|\mathbf{x}|\mathbf{e}_1, \mathbf{x}\right\rangle \frac{\mathbf{x}-|\mathbf{x}|\mathbf{e}_1}{\left|\mathbf{x}-|\mathbf{x}|\mathbf{e}_1\right|^2} \\
&= \mathbf{x} - (\mathbf{x}-|\mathbf{x}|\mathbf{e}_1) \\
&= |\mathbf{x}|\mathbf{e}_1.
\end{aligned}$$

$\square$

Householder 变换 (续)

定理

任意非奇异矩阵 $\mathbf{A}$ 均可通过左连乘一系列 Householder 矩阵化为上三角矩阵.

证明　设 $\mathbf{A}^{(0)} = \mathbf{A} \in \mathbb{R}^{n\times n}$.

由 $\det\mathbf{A}^{(0)} \neq 0$ 知 $\mathbf{A}^{(0)}$ 的第 1 列 $\mathbf{x}^{(1)} \neq \mathbf{0}$. 令 $a_{11}^{(1)} = |\mathbf{x}^{(1)}|$. 由引理知存在 Householder 矩阵 $\mathbf{H}_1$, 使得

$$\mathbf{H}_1\mathbf{x}^{(1)} = |\mathbf{x}^{(1)}|\mathbf{e}_1^{(n)}, \quad \mathbf{e}_1^{(n)} \in \mathbb{R}^n,$$

从而

$$\mathbf{H}_1\mathbf{A}^{(0)} = \left[\begin{array}{c|ccc} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} \\ \hline 0 & & & \\ \vdots & & \mathbf{A}^{(1)} & \\ 0 & & & \end{array}\right].$$

由 $\det\mathbf{A}^{(1)} \neq 0$ 知 $\mathbf{A}^{(1)}$ 的第 1 列 $\mathbf{x}^{(2)} \neq \mathbf{0}$. 令 $a_{22}^{(2)} = \left|\mathbf{x}^{(2)}\right|$. 由引理知存在 Householder 矩阵 $\mathbf{H}_2$, 使得

$$\mathbf{H}_2\mathbf{x}^{(2)} = \left|\mathbf{x}^{(2)}\right|\mathbf{e}_1^{(n-1)}, \quad \mathbf{e}_1^{(n-1)} \in \mathbb{R}^{n-1},$$

从而

$$\mathbf{H}_2\mathbf{A}^{(1)} = \left[\begin{array}{c|ccc} a_{22}^{(2)} & a_{23}^{(2)} & \cdots & a_{2n}^{(2)} \\ \hline 0 & & & \\ \vdots & & \mathbf{A}^{(2)} & \\ 0 & & & \end{array}\right].$$

如此继续下去, 得到 $\mathbf{A}^{(i)}$ 以及 Householder 矩阵 $\mathbf{H}_i (i = 1, \cdots, n-2)$.

由 $\det \mathbf{A}^{(n-2)} \neq 0$ 知 $\mathbf{A}^{(n-2)}$ 的第 1 列 $\mathbf{x}^{(n-1)} \neq \mathbf{0}$. 令 $a_{n-1,n-1}^{(n-1)} = \left|\mathbf{x}^{(n-1)}\right|$. 由引理知存在 Householder 矩阵 $\mathbf{H}_{n-1}$, 使得

$$\mathbf{H}_{n-1}\mathbf{x}^{(n-1)} = \left|\mathbf{x}^{(n-1)}\right|\mathbf{e}_1^{(2)}, \quad \mathbf{e}_1^{(2)} \in \mathbb{R}^2,$$

从而

$$\mathbf{H}_{n-1}\mathbf{A}^{(n-2)} = \begin{bmatrix} a_{n-1,n-1}^{(n-1)} & a_{n-1,n}^{(n-1)} \\ 0 & a_{nn}^{(n-1)} \end{bmatrix}.$$

令

$$\mathbf{H} = \begin{bmatrix} \mathbf{I}_{n-2} & \mathbf{O} \\ \mathbf{O} & \mathbf{H}_{n-1} \end{bmatrix} \cdots \begin{bmatrix} \mathbf{I}_2 & \mathbf{O} \\ \mathbf{O} & \mathbf{H}_3 \end{bmatrix} \begin{bmatrix} 1 & \mathbf{0}^{\mathrm{T}} \\ \mathbf{0} & \mathbf{H}_2 \end{bmatrix} \mathbf{H}_1,$$

则 $\mathbf{H}$ 为 $n-1$ 个 Householder 矩阵的乘积, 且

$$\mathbf{HA} = \begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1,n-1}^{(1)} & a_{1n}^{(1)} \\ & a_{22}^{(2)} & \cdots & a_{2,n-1}^{(2)} & a_{2n}^{(2)} \\ & & \ddots & \vdots & \vdots \\ & & & a_{n-1,n-1}^{(n-1)} & a_{n-1,n}^{(n-1)} \\ & & & & a_{nn}^{(n-1)} \end{bmatrix}.$$

$\square$

Householder 变换 (续)

1. 在上述定理中, $\mathbf{H}$ 为有限个 Householder 矩阵的乘积, 因而是正交阵.
2. 令 $\mathbf{R} = \mathbf{HA}$, 则 $\mathbf{A} = \mathbf{H}^{\mathrm{T}}\mathbf{R}$ 为 $\mathbf{A}$ 的 $\mathbf{QR}$ 分解.
3. 上述定理的证明过程实际上给出了用 Householder 变换方法求矩阵 QR 分解的一种算法.
4. 上述定理中 $\mathbf{A}$ 的条件可以削弱为满列秩情形. 详细内容参见教材第 237 页第 3, 4 段以及定理 2.1.

用 Householder 变换求矩阵

$$\mathbf{A} = \begin{bmatrix} 3 & 14 & 9 \\ 6 & 43 & 3 \\ 6 & 22 & 15 \end{bmatrix}$$

的 QR 分解.

解　(1) 对 $\mathbf{A}$ 的第 1 列, 构造 Householder 矩阵:

$$\mathbf{x}^{(1)} = \begin{bmatrix} 3 \\ 6 \\ 6 \end{bmatrix}, \mathbf{x}^{(1)} - |\mathbf{x}^{(1)}|\mathbf{e}_1 = 6\begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}, \mathbf{u}_1 = \frac{1}{\sqrt{3}}\begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix},$$

$$\mathbf{H}_1 = \mathbf{I} - 2\mathbf{u}_1\mathbf{u}_1^{\mathrm{T}} = \frac{1}{3}\begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & -2 \\ 2 & -2 & 1 \end{bmatrix},$$

$$\mathbf{H}_1\mathbf{A} = \begin{bmatrix} 9 & 48 & 15 \\ 0 & 9 & -3 \\ 0 & -12 & 9 \end{bmatrix} = \begin{bmatrix} 9 & * \\ \mathbf{0} & \mathbf{A}^{(1)} \end{bmatrix}.$$

Householder 变换举例

(2) 对 $\mathbf{A}^{(1)} = \begin{bmatrix} 9 & -3 \\ -12 & 9 \end{bmatrix}$ 的第 1 列, 构造 Householder 矩阵:

$$\mathbf{x}^{(2)} = \begin{bmatrix} 9 \\ -12 \end{bmatrix}, \quad \mathbf{x}^{(2)} - |\mathbf{x}^{(2)}|\mathbf{e}_1 = 6\begin{bmatrix} -1 \\ -2 \end{bmatrix},$$

$$\mathbf{u}_2 = \frac{1}{\sqrt{5}}\begin{bmatrix} -1 \\ -2 \end{bmatrix},$$

$$\mathbf{H}_2 = \mathbf{I} - 2\mathbf{u}_2\mathbf{u}_2^{\mathrm{T}} = \frac{1}{5}\begin{bmatrix} 3 & -4 \\ -4 & -3 \end{bmatrix},$$

$$\mathbf{H}_2\mathbf{A}^{(1)} = \begin{bmatrix} 15 & -9 \\ 0 & -3 \end{bmatrix}.$$

(3) 令

$$\mathbf{H} = \begin{bmatrix} 1 & \\ & \mathbf{H}_2 \end{bmatrix} \mathbf{H}_1 = \frac{1}{15} \begin{bmatrix} 5 & 10 & 10 \\ -2 & 11 & -10 \\ -14 & 2 & 5 \end{bmatrix},$$

则有

$$\mathbf{Q} = \mathbf{H}^{\mathrm{T}} = \frac{1}{15} \begin{bmatrix} 5 & -2 & -14 \\ 10 & 11 & 2 \\ 10 & -10 & 5 \end{bmatrix}$$

$$\mathbf{R} = \begin{bmatrix} 9 & 48 & 15 \\ & 15 & -9 \\ & & -3 \end{bmatrix}$$

$\mathbf{A} = \mathbf{QR}.$ □

一般矩阵的 QR 分解

考虑一般的矩阵 $\mathbf{A} \in \mathbb{R}^{m\times n}$, 其中

$$\operatorname{rank}\mathbf{A} = r \leq \min\{m, n\},$$

对 $\mathbf{A}$ 进行 QR 分解的详细步骤见教材第 237 页最后一行至第 238 页第三段.

其它形式的 Householder 变换

前述方法是从矩阵 $\mathbf{A}_{n\times n}^{(0)}$ 开始, 由第 1 列构造 Householder 矩阵产生 $\mathbf{A}_{(n-1)\times(n-1)}^{(1)}$, 依次类推. 教材中给出的方法是产生同阶矩阵 $\mathbf{A}_{n\times n}^{(1)}$.

具体描述为: 用 Householder 变换先将 $\mathbf{A}$ 的第 1 列变为 $[*,0,\cdots,0]^{\mathrm{T}}$; 再将其第 2 列变为 $[*,*,0,\cdots,0]^{\mathrm{T}}$ 且使其第 1 列为零的元素不变; 如此下去, 可将矩阵的第 i 列变为 $[\underbrace{*,\cdots,*}_{i},0,\cdots,0]^{\mathrm{T}}$ 且使前面 $i-1$ 次变换所形成的上三角形保持不变.

其它形式的 Householder 变换 (续)

为实现上述目的, 教材给出了另一种 Householder 变换的定义, 详见第 235 页定义 2.1. (注意要求 $s \neq 0$)

事实上, 两个定义是等价的.

(1) 设 $\mathbf{x} = [x_1, \cdots, x_n]^{\mathrm{T}} \neq \mathbf{0}$. 对任意 $i = 1, \cdots, n$, 若 $s = \sqrt{\sum_{j=i}^{n} x_j^2} \neq 0$, 令

$$\mathbf{z} = \frac{1}{|\mathbf{x}|}[x_1, \cdots, x_{i-1}, -\operatorname{sign}(x_i)s, 0, \cdots, 0]^{\mathrm{T}}.$$

显然 $\mathbf{z}^{\mathrm{T}}\mathbf{z} = \frac{1}{|\mathbf{x}|^2}(\sum_{j=1}^{i-1} x_j^2 + s^2) = 1$, 即 $\mathbf{z}$ 是非零单位向量. 因为存在 $j \geq i$ 使得 $x_j \neq 0$, 所以 $\mathbf{x} \neq |\mathbf{x}|\mathbf{z}$, 从而可令

$$\mathbf{u} = \frac{\mathbf{x} - |\mathbf{x}|\mathbf{z}}{\left|\mathbf{x} - |\mathbf{x}|\mathbf{z}\right|},$$

并构造 Householder 矩阵 $\mathbf{H} = \mathbf{I} - 2\mathbf{u}\mathbf{u}^{\mathrm{T}}$.

两个 Householder 变换定义的等价性

(2) 令 $\beta = \frac{1}{2}\left|\mathbf{x} - |\mathbf{x}|\mathbf{z}\right|^2 \neq 0$, 则

$$\begin{aligned}\beta &= \mathbf{x}^{\mathrm{T}}\mathbf{x} - |\mathbf{x}|\mathbf{z}^{\mathrm{T}}\mathbf{x} \\ &= \left(s^2 + \sum_{j=1}^{i-1} x_j^2\right) - \left(\sum_{j=1}^{i-1} x_j^2 - \mathrm{sign}(x_i)x_i s\right) \\ &= s^2 + \mathrm{sign}(x_i)x_i s \\ &= s^2 + |x_i|s.\end{aligned}$$

令 $\mathbf{v} = \sqrt{2\beta}\mathbf{u}$, 则 $\mathbf{H} = \mathbf{I} - \frac{1}{\beta}\mathbf{v}\mathbf{v}^{\mathrm{T}}$, 且

$$\begin{aligned}\mathbf{v} &= \mathbf{x} - |\mathbf{x}|\mathbf{z} \\ &= [0, \cdots, 0, x_i + \mathrm{sign}(x_i)s, x_{i+1}, \cdots, x_n]^{\mathrm{T}}, \\ \mathbf{H}\mathbf{x} &= |\mathbf{x}|\mathbf{z} = [x_1, \cdots, x_{i-1}, -\mathrm{sign}(x_i)s, 0, \cdots, 0]^{\mathrm{T}}.\end{aligned}$$

Householder 变换的性质

教材中所定义 Householder 矩阵的性质：见教材第 235 – 236 页的性质.

要求课后认真阅读、理解.

问　题?