《数值计算方法》课程

矩阵的特征值和奇异值

QR分解与QR算法

胡建芳

（研究方向：计算机视觉）

http://sdcs.sysu.edu.cn/content/5143

计算机学院

课程回顾

QR分解：

A=QR

施密特正交化（Schmidt Orthogonalization）

设 $\alpha_1, \alpha_2, \cdots, \alpha_m\ (m \leq n)$ 是 R^n 中的一个线性无关向量组，若令

$$\beta_1 = \alpha_1$$

$$\beta_2 = \alpha_2 - \frac{\langle \alpha_2, \beta_1 \rangle}{\langle \beta_1, \beta_1 \rangle}\beta_1$$

$$\cdots$$

$$\beta_m = \alpha_m - \frac{\langle \alpha_m, \beta_1 \rangle}{\langle \beta_1, \beta_1 \rangle}\beta_1 - \frac{\langle \alpha_m, \beta_2 \rangle}{\langle \beta_2, \beta_2 \rangle}\beta_2 - \cdots - \frac{\langle \alpha_m, \beta_{m-1} \rangle}{\langle \beta_{m-1}, \beta_{m-1} \rangle}\beta_{m-1}$$

则 $\beta_1, \beta_2, \cdots, \beta_m$ 就是一个 正交向量组，若再令

$$e_i = \frac{\beta_i}{\|\beta_i\|}(i = 1, 2, \cdots, m)$$

就得到一个标准正交向量组 $e_1, e_2, \cdots, e_m$ ， 且该向量组与 $\alpha_1, \alpha_2, \cdots, \alpha_m$ 等价。 [1]

任意一个满秩实矩阵A，都可以唯一地分解A=QR，其中Q为正交矩阵，R是具有正对角元的上三角矩阵。

豪斯霍尔德变换的QR分解

豪斯霍尔德变换

定义 设非零向量$W \in R^n, W = (w_1, w_2, \cdots, w_n)^T$,
且满足条件$\|W\|_2 = 1$,形如

$$H = I - 2WW^T$$

的n阶方阵称为初等反射阵,或称为*Householder*
变换阵.

$$H = \begin{bmatrix} 1-2w_1^2 & -2w_1w_2 & \cdots & -2w_1w_n \\ -2w_2w_1 & 1-2w_2^2 & \cdots & -2w_2w_n \\ \cdots & \cdots & \cdots & \cdots \\ -2w_nw_1 & -2w_nw_2 & \cdots & 1-2w_n^2 \end{bmatrix}$$

豪斯霍尔德变换的QR分解

- **豪斯霍尔德变换**

例：$W=\left(\frac{1}{\sqrt{2}}\quad 0\quad \frac{1}{\sqrt{2}}\right)^T\in R^3, \|W\|_2=1$

$$H=I-2WW^T=I-2\begin{bmatrix}\frac{1}{\sqrt{2}}\\0\\\frac{1}{\sqrt{2}}\end{bmatrix}\left(\frac{1}{\sqrt{2}}\quad 0\quad \frac{1}{\sqrt{2}}\right)$$

$$=\begin{bmatrix}0&0&-1\\0&1&0\\-1&0&0\end{bmatrix}$$

豪斯霍尔德变换的QR分解

- **豪斯霍尔德变换**

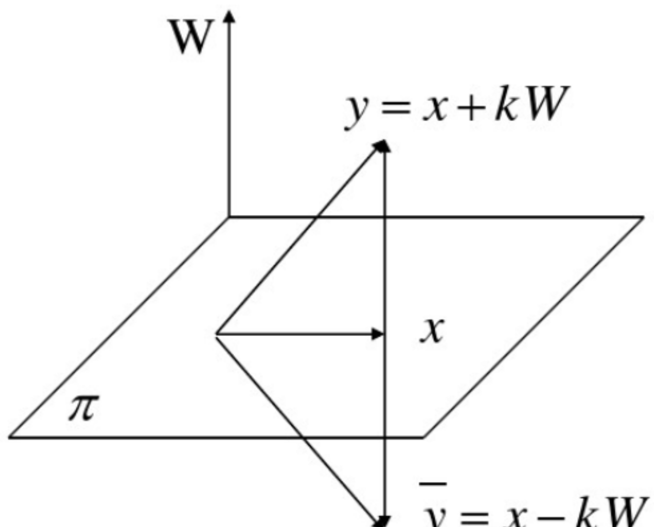
（3）镜映射－几何意义

平面π方程 $W^T x = 0 \quad \forall x \in \pi$

若 $x \in \pi \quad , Hx = (I - 2WW^T)x = x - 2WW^T x = x$

若 $y \notin \pi \quad , Hy = H(x + kW) = x + k(I - 2WW^T)W$

$$= x + kW - 2kWW^T W = x - kW = \overline{y}$$



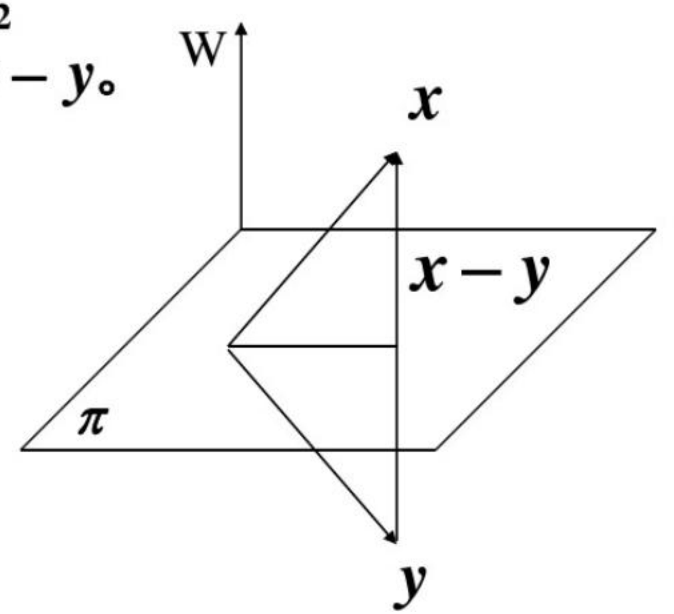
豪斯霍尔德变换的QR分解

豪斯霍尔德变换

定理 设两个不相等的n维向量$x, y \in R^n, x \neq y$,
但$\|x\|_2 = \|y\|_2$,则存在householder阵

$$H = I - 2\frac{UU^T}{\|U\|_2^2}$$

使$Hx = y$，其中$U = x - y$。



豪斯霍尔德变换的QR分解

豪斯霍尔德变换

证：若设$W=\dfrac{U}{\|U\|_2}$，则有$\|W\|_2=1$，因此

$$H=I-2WW^T=I-2\frac{UU^T}{\|U\|_2^2}$$

$$=I-2\frac{(x-y)}{\|x-y\|_2^2}(x^T-y^T)$$

$$Hx=x-2\frac{(x-y)}{\|x-y\|_2^2}(x^T-y^T)x$$

$$=x-2\frac{(x-y)(x^Tx-y^Tx)}{\|x-y\|_2^2}$$

因为$\|x-y\|_2^2=(x^T-y^T)(x-y)=2(x^Tx-y^Tx)$

代入上式后即得到$Hx=y$

$$\boxed{\begin{aligned} \because\ x^Tx&=y^Ty,\\ x^Ty&=y^Tx \end{aligned}}$$

豪斯霍尔德变换的QR分解

- **豪斯霍尔德变换**

1. *Householder*变换可以将给定的向量变为一个与任一个$e_i \in R^n (i=1,2,\cdots,n)$同方向的向量。

即：$\forall x=(x_1,x_2,\cdots,x_n)^T \in R^n, x \neq 0$，可构造$H$阵，

使$Hx = y = -\sigma_i e_i = (0,\ldots,0,-\sigma_i,0,\cdots,0)^T \in R^n$

其中$\sigma_i = sign(x_i)\|x\|_2 = sign(x_i)(\sum_{k=1}^{n} x_k^2)^{\frac{1}{2}}$，

$$sign(x_i) = \begin{cases} 1 & x_i \geq 0 \\ -1 & x_i < 0 \end{cases}$$

$U = x - y = x + \sigma_i e_i = (x_1,\cdots,x_i+\sigma_i,\cdots,x_n)^T$，

豪斯霍尔德变换的QR分解

- **豪斯霍尔德变换**

构造初等反射阵

$$H = I - 2WW^T = I - 2\frac{UU^T}{\|U\|^2} = I - \frac{1}{\rho}UU^T$$

有$Hx = y = -\sigma_i e_i$

其中

$$\rho = \frac{1}{2}U^T U = \frac{1}{2}(x_1^2 + \ldots + (x_i + \sigma_i)^2 + \cdots + x_n^2)$$

$$= \frac{1}{2}(2x_i\sigma_i + 2\sigma_i^2) = \sigma_i(x_i + \sigma_i)$$

特征值特征向量的意义

- 豪斯霍尔德变换

例 已知向量$x=(2,0,2,1)^T$，试构造$Householder$阵，使$Hx=Ke_3$，其中$e_3=(0,0,1,0)^T\in R^4, K\in R$。

解： $\sigma_3=sign(x_3)\|x\|_2=\sqrt{4+0+4+1}=3$，因$x_3=2>0$，

故取$K=-\sigma_3=-3$　于是$y=-\sigma_3 e_3=Ke_3=(0,0,-3,0)^T$，

$U=x-y=(2,0,5,1)^T$，$\rho=\sigma_3(\sigma_3+x_3)=3(3+2)=15$

$$\rho=\frac{1}{2}U^TU$$

$$H=I-\frac{1}{\rho}UU^T=\frac{1}{15}\begin{bmatrix}11&0&-10&-2\\0&1&0&0\\-10&0&-10&-5\\-2&0&-5&14\end{bmatrix}$$

豪斯霍尔德变换的QR分解

- **QR分解**

  构造矩阵H1使得

$$
H_1A = H_1\begin{bmatrix} \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \end{bmatrix} = \begin{bmatrix} \times & \times & \times \\ 0 & \times & \times \\ 0 & \times & \times \\ 0 & \times & \times \end{bmatrix}
$$

$$
\left[\begin{array}{c|ccc} 1 & 0 & 0 & 0 \\ \hline 0 & & & \\ 0 & & \hat{H}_2 & \\ 0 & & & \end{array}\right]\left[\begin{array}{c|cc} \times & \times & \times \\ \hline 0 & \times & \times \\ 0 & \times & \times \\ 0 & \times & \times \end{array}\right] = \left[\begin{array}{c|cc} \times & \times & \times \\ \hline 0 & \times & \times \\ 0 & 0 & \times \\ 0 & 0 & \times \end{array}\right]
$$

$$
H_3H_2H_1A = R,
$$

$$
A = H_1H_2H_3R = QR,
$$

$$
\left[\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \hline 0 & 0 & & \\ 0 & 0 & \multicolumn{2}{c}{\hat{H}_3} \end{array}\right]\left[\begin{array}{cc|c} \times & \times & \times \\ 0 & \times & \times \\ \hline 0 & 0 & \times \\ 0 & 0 & \times \end{array}\right] = \left[\begin{array}{cc|c} \times & \times & \times \\ 0 & \times & \times \\ \hline 0 & 0 & \times \\ 0 & 0 & 0 \end{array}\right]
$$

豪斯霍尔德变换的QR分解

QR分解

用 Householder 反射求 $\boldsymbol{A}=\begin{bmatrix}3 & 1\\ 4 & 3\end{bmatrix}$ 的 $\boldsymbol{QR}$ 分解.

设

$$\boldsymbol{v}=\boldsymbol{\omega}-\boldsymbol{x}=\begin{bmatrix}5\\0\end{bmatrix}-\begin{bmatrix}3\\4\end{bmatrix}=\begin{bmatrix}2\\-4\end{bmatrix},$$

并定义投影矩阵

$$\boldsymbol{P}=\frac{\boldsymbol{v}\boldsymbol{v}^{\mathrm{T}}}{\boldsymbol{v}^{\mathrm{T}}\boldsymbol{v}}=\frac{1}{20}\begin{bmatrix}4 & -8\\ -8 & 16\end{bmatrix}=\begin{bmatrix}0.2 & -0.4\\ -0.4 & 0.8\end{bmatrix},$$

于是

$$\boldsymbol{H}=\boldsymbol{I}-2\boldsymbol{P}=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}-\begin{bmatrix}0.4 & -0.8\\ -0.8 & 1.6\end{bmatrix}=\begin{bmatrix}0.6 & 0.8\\ 0.8 & -0.6\end{bmatrix}.$$

$$\boldsymbol{H}_1\boldsymbol{A}=\begin{bmatrix}0.6 & 0.8\\ 0.8 & -0.6\end{bmatrix}\begin{bmatrix}3 & 1\\ 4 & 3\end{bmatrix}=\begin{bmatrix}5 & 3\\ 0 & -1\end{bmatrix}.$$

两边左乘 $\boldsymbol{H}_1^{-1}=\boldsymbol{H}_1$ 得到

$$\boldsymbol{A}=\begin{bmatrix}3 & 1\\ 4 & 3\end{bmatrix}=\begin{bmatrix}0.6 & 0.8\\ 0.8 & -0.6\end{bmatrix}\begin{bmatrix}5 & 3\\ 0 & -1\end{bmatrix}=\boldsymbol{QR},$$

这里 $\boldsymbol{Q}=\boldsymbol{H}_1^{\mathrm{T}}=\boldsymbol{H}_1$.

豪斯霍尔德变换的QR分解

QR分解

用 Householder 反射求 $A=\begin{bmatrix}1 & -4\\ 2 & 3\\ 2 & 2\end{bmatrix}$ 的 QR 分解.

需要求出把第一列 $x=[1,2,2]$ 变到向量 $\omega=[\|x\|_2,0,0]$ 的 Householder 反射. 设 $v=\omega-x=[3,0,0]-[1,2,2]=[2,-2,-2]$. 参考定理 4.4, 我们有

$$H_1=\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}-\frac{2}{12}\begin{bmatrix}4 & -4 & -4\\ -4 & 4 & 4\\ -4 & 4 & 4\end{bmatrix}=\begin{bmatrix}\frac{1}{3} & \frac{2}{3} & \frac{2}{3}\\ \frac{2}{3} & \frac{1}{3} & -\frac{2}{3}\\ \frac{2}{3} & -\frac{2}{3} & \frac{1}{3}\end{bmatrix},$$

$$H_1A=\begin{bmatrix}\frac{1}{3} & \frac{2}{3} & \frac{2}{3}\\ \frac{2}{3} & \frac{1}{3} & -\frac{2}{3}\\ \frac{2}{3} & -\frac{2}{3} & \frac{1}{3}\end{bmatrix}\begin{bmatrix}1 & -4\\ 2 & 3\\ 2 & 2\end{bmatrix}=\begin{bmatrix}3 & 2\\ 0 & -3\\ 0 & -4\end{bmatrix}.$$

豪斯霍尔德变换的QR分解

QR分解

用 Householder 反射求 $\boldsymbol{A}=\begin{bmatrix}1 & -4\\ 2 & 3\\ 2 & 2\end{bmatrix}$ 的 $\boldsymbol{QR}$ 分解.

余下的步骤是把向量 $\hat{x}=[-3,-4]$ 变成 $\hat{\omega}=[5,0]$. 从定理 4.4 计算 $\hat{\boldsymbol{H}}_2$ 便得到

$$\begin{bmatrix}-0.6 & -0.8\\ -0.8 & 0.6\end{bmatrix}\begin{bmatrix}-3\\ -4\end{bmatrix}=\begin{bmatrix}5\\ 0\end{bmatrix},$$

$$\boldsymbol{H}_2\boldsymbol{H}_1\boldsymbol{A}=\begin{bmatrix}1 & 0 & 0\\ 0 & -0.6 & -0.8\\ 0 & -0.8 & 0.6\end{bmatrix}\begin{bmatrix}\frac{1}{3} & \frac{2}{3} & \frac{2}{3}\\ \frac{2}{3} & \frac{1}{3} & -\frac{2}{3}\\ \frac{2}{3} & -\frac{2}{3} & \frac{1}{3}\end{bmatrix}\begin{bmatrix}1 & -4\\ 2 & 3\\ 2 & 2\end{bmatrix}=\begin{bmatrix}3 & 2\\ 0 & 5\\ 0 & 0\end{bmatrix}=\boldsymbol{R}.$$

两边左乘 $\boldsymbol{H}_1^{-1}\boldsymbol{H}_2^{-1}=\boldsymbol{H}_1\boldsymbol{H}_2$ 得到 $\boldsymbol{QR}$ 分解:

$$\begin{bmatrix}1 & -4\\ 2 & 3\\ 2 & 2\end{bmatrix}=\boldsymbol{H}_1\boldsymbol{H}_2\boldsymbol{R}=\begin{bmatrix}\frac{1}{3} & \frac{2}{3} & \frac{2}{3}\\ \frac{2}{3} & \frac{1}{3} & -\frac{2}{3}\\ \frac{2}{3} & -\frac{2}{3} & \frac{1}{3}\end{bmatrix}\begin{bmatrix}1 & 0 & 0\\ 0 & -0.6 & -0.8\\ 0 & -0.8 & 0.6\end{bmatrix}\begin{bmatrix}3 & 2\\ 0 & 5\\ 0 & 0\end{bmatrix}$$

$$=\frac{1}{15}\begin{bmatrix}5 & -14 & -2\\ 10 & 5 & -10\\ 10 & 2 & 11\end{bmatrix}\begin{bmatrix}3 & 2\\ 0 & 5\\ 0 & 0\end{bmatrix},$$

豪斯霍尔德变换的QR分解

QR分解唯一性

对于给定的 $m\times n$ 矩阵 $\boldsymbol{A}$, $\boldsymbol{QR}$ 分解不是唯一的. 例如, 定义 $\boldsymbol{D}=\mathrm{diag}(d_1,\cdots,d_m)$, 其中每一个 d_i 或者是 $+1$ 或者是 -1. 于是 $\boldsymbol{A}=\boldsymbol{QR}=\boldsymbol{QDDR}$, 而且我们检验 $\boldsymbol{QD}$ 是正交的, $\boldsymbol{DR}$ 是上三角的.

任意一个满秩实矩阵A，都可以唯一地分解A=QR，其中Q为正交矩阵，R是具有正对角元的上三角矩阵。

豪斯霍尔德变换的QR分解

$A \in R^{n\times n}$， *Schmidt* 正交化方法和*Householder*方法结果一样。

$A \in R^{m\times n}\ (m > n, r(A) = n)$ 　 *Sm* 法和*H*法结果不一样。

*Sm*法 $\underset{m\times n}{\begin{bmatrix} \ \ \ \\ A \\ \ \ \ \end{bmatrix}} = \underset{m\times n}{\begin{bmatrix} \ \ \ \\ Q \\ \ \ \ \end{bmatrix}} \underset{n\times n}{\begin{bmatrix} \ \ \ \\ R \\ \ \ \ \end{bmatrix}}$ 紧凑格式

*H*法 $\underset{m\times n}{\begin{bmatrix} \ \ \ \\ A \\ \ \ \ \end{bmatrix}} = \underset{m\times m}{\begin{bmatrix} \ \ \ \\ Q \\ \ \ \ \end{bmatrix}} \underset{m\times n}{\begin{bmatrix} \overline{R} \\ \\ O \end{bmatrix}}$， $\overline{R} \in R^{n\times n}$

豪斯霍尔德变换的QR分解

- **QR算法：一种幂法迭代**

$$\left[A\begin{bmatrix}1\\0\\\vdots\\0\end{bmatrix} \middle| A\begin{bmatrix}0\\1\\\vdots\\0\end{bmatrix} \middle| \cdots \middle| A\begin{bmatrix}0\\0\\\vdots\\1\end{bmatrix} \right] = [\bar{q}_1^1|\cdots|\bar{q}_m^1] \begin{bmatrix} r_{11}^1 & r_{12}^1 & \cdots & r_{1m}^1 \\ & r_{22}^1 & & \vdots \\ & & \ddots & \vdots \\ & & & r_{mm}^1 \end{bmatrix}$$

$$A\bar{Q}_1 = [A\bar{q}_1^1|A\bar{q}_2^1|\cdots|A\bar{q}_m^1]$$

$$= [\bar{q}_1^2|\bar{q}_2^2|\cdots|\bar{q}_m^2] \begin{bmatrix} r_{11}^2 & r_{12}^2 & \cdots & r_{1m}^2 \\ & r_{22}^2 & & \vdots \\ & & \ddots & \vdots \\ & & & r_{mm}^2 \end{bmatrix} = \bar{Q}_2 R_2.$$

迭代多次

$$A\bar{Q}_0 = \bar{Q}_1 R_1,$$

$$A\bar{Q}_1 = \bar{Q}_2 R_2,$$

$$A\bar{Q}_2 = \bar{Q}_3 R_3,$$

$$\vdots$$

豪斯霍尔德变换的QR分解

- **QR算法：一种幂法迭代**

$$
\begin{aligned}
A\bar{Q}_0 &= \bar{Q}_1 R_1, \\
A\bar{Q}_1 &= \bar{Q}_2 R_2, \\
A\bar{Q}_2 &= \bar{Q}_3 R_3, \\
&\vdots
\end{aligned}
\qquad \Longrightarrow \qquad
\begin{aligned}
A_0 &\equiv AQ_0 = Q_1 R_1', \\
A_1 &\equiv R_0' Q_1 = Q_2 R_2', \\
A_2 &\equiv R_2' Q_2 = Q_3 R_3', \\
&\vdots
\end{aligned}
$$

$$
A_{k+1} = R_k Q_k = Q_k^{-1} Q_k R_k Q_k = Q_k^{-1} A_k Q_k = Q_k^T A_k Q_k,
$$

定理 12.4 假设 A 是对称的 $m \times m$ 矩阵，其特征值 λ_i 满足 $|\lambda_1| > |\lambda_2| > \cdots > |\lambda_m|$，非移位 QR 算法线性收敛于 A 的特征向量和特征值. 当 $j \to \infty$ 时，A_j 收敛于特征值在主对角线上的对角矩阵，并且 $\bar{Q}_j = Q_1 \cdots Q_j$ 收敛于一个正交矩阵，它的列是特征向量.

豪斯霍尔德变换的QR分解

- **作业：**

  用豪斯霍尔德变换求下列矩阵的QR分解，并用QR算法求对应的特征值与特征向量。

$$\begin{bmatrix} 3 & -1 & 2 \\ 4 & 1 & 0 \\ -3 & 2 & 1 \\ 1 & 1 & 5 \\ -2 & 0 & 3 \end{bmatrix} \qquad \begin{bmatrix} 4 & 2 & 3 & 0 \\ -2 & 3 & -1 & 1 \\ 1 & 3 & -4 & 2 \\ 1 & 0 & 1 & -1 \\ 3 & 1 & 3 & -2 \end{bmatrix}$$

矩阵特征值

THE END