《数值计算方法》课程

矩阵的特征值和奇异值

奇异值分解定理 (非常重要)

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分解与QR算法：

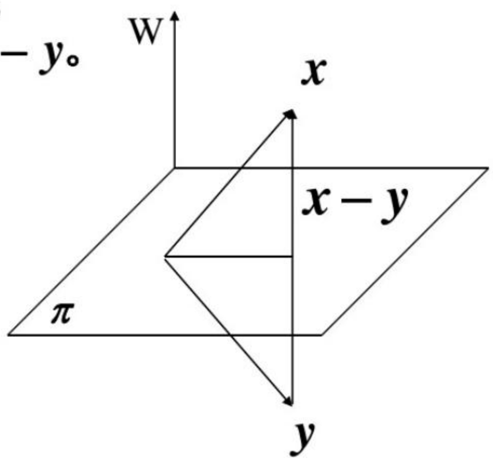
A=QR

豪斯霍尔德变换 （镜面映射法）

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$U = x - y$。



$$A\bar{Q}_0 = \bar{Q}_1 R_1,$$
$$A\bar{Q}_1 = \bar{Q}_2 R_2,$$
$$A\bar{Q}_2 = \bar{Q}_3 R_3,$$
$$\vdots$$

$$A_0 \equiv AQ_0 = Q_1 R_1',$$
$$A_1 \equiv R_0' Q_1 = Q_2 R_2',$$
$$A_2 \equiv R_2' Q_2 = Q_3 R_3',$$
$$\vdots$$

课程回顾

- **QR分解**

  构造矩阵H1使得

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

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

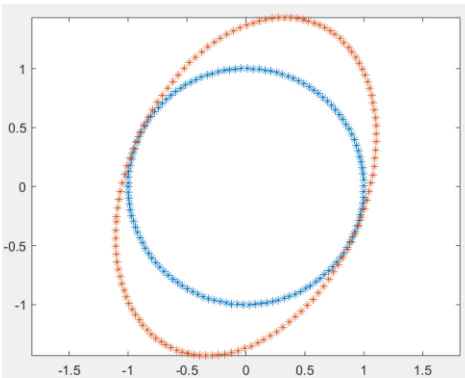
$$
\left[\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \hline 0 & 0 & & \\ 0 & 0 & \multicolumn{2}{c}{\hat{H}_3} \end{array}\right]
\left[\begin{array}{cc|c} \times & \times & \times \\ 0 & \times & \times \\ \hline 0 & 0 & \times \\ 0 & 0 & \times \end{array}\right]
=
\left[\begin{array}{cc|c}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

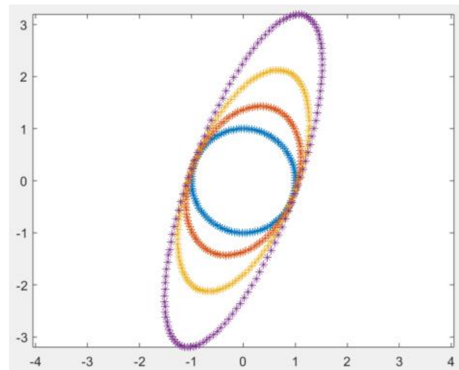
$$
H_3H_2H_1A = R,
$$

$$
A = H_1H_2H_3R = QR,
$$

奇异值分解

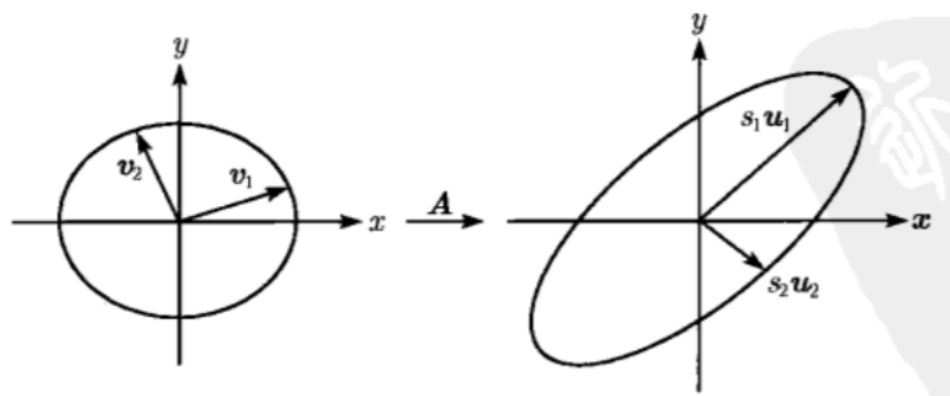
奇异值向量





对每一个矩阵***A***，存在标准正交基 ***U***, ***V***及非负数***s***, 使得

$$
\begin{aligned}
A\boldsymbol{v}_1 &= s_1\boldsymbol{u}_1,\\
A\boldsymbol{v}_2 &= s_2\boldsymbol{u}_2,\\
&\vdots\\
A\boldsymbol{v}_n &= s_n\boldsymbol{u}_n.
\end{aligned}
$$



特征向量的拓展

奇异值分解

奇异值与奇异向量

定义 设$A \in R^{m\times n}$，若$A^T A$的特征值为

$$\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_r > \lambda_{r+1} = \ldots = \lambda_n = 0$$

称$\sigma_i = \sqrt{\lambda_i}\,(i = 1, 2, \ldots, r)$为矩阵$A$的奇异值.

例 12.5 求矩阵

$$\boldsymbol{A} = \begin{bmatrix} 0 & -\frac{1}{2} \\ 3 & 0 \\ 0 & 0 \end{bmatrix}$$

的奇异值和奇异向量

$$\boldsymbol{A}\boldsymbol{v}_1 = \boldsymbol{A}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = 3\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = s_1\boldsymbol{u}_1,$$

$$\boldsymbol{A}\boldsymbol{v}_2 = \boldsymbol{A}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix} = s_2\boldsymbol{u}_2$$

奇异值分解

- **奇异值分解（SVD）定理**

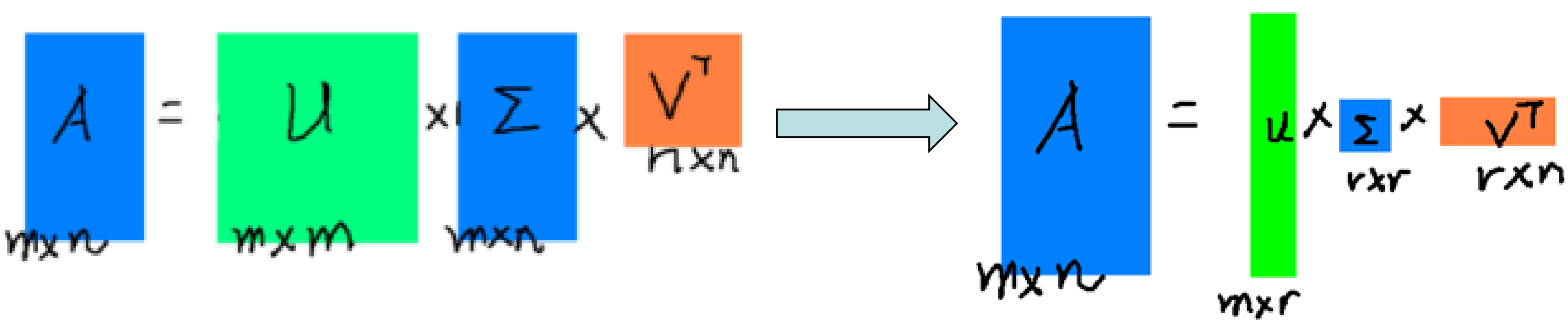
定理 设$A \in R^{m\times n}$，秩$(A)=r$,则存在m阶正交阵U和n阶正交阵V，使得

$$U^T AV = \Sigma = \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix}$$

其中$\Sigma_r = diag(\sigma_1, \sigma_2, \ldots, \sigma_r)$,且$\sigma_1 \ge \sigma_2 \ge \ldots \ge \sigma_r > 0$

称$A = U\Sigma V^T = U\begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix}V^T$为矩阵$A$的奇异值分解,

$\Sigma = \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix}$为矩阵$A$的奇异值矩阵。



奇异值分解

奇异值分解（SVD）定理

证明 $\because$ 秩$(A)=r, A^TA$为半正定阵，故A^TA的特征值设为

$\sigma_1^2,\sigma_2^2,\ldots,\sigma_n^2$，且$\sigma_1\ge\sigma_2\ge\ldots\ge\sigma_r>0,\sigma_{r+1}=\cdots=\sigma_n=0$

于是存在n阶正交阵V，使A^TA正交相似于对角阵，即

$$V^T\left(A^TA\right)V=diag(\sigma_1^2,\sigma_2^2,\cdots,\sigma_r^2,0,\cdots,0)$$

其中$V=\left[v_1,\cdots,v_r,v_{r+1},\cdots,v_n\right]$，将$V$分块为$V=\left[V_1,V_2\right]$

其中$V_1=\left[v_1,\cdots,v_r\right], V_2=\left[v_{r+1},\cdots,v_n\right]$

于是有$\begin{bmatrix}V_1^T\\V_2^T\end{bmatrix}\left(A^TA\right)\left[V_1,V_2\right]=\begin{bmatrix}V_1^T\left(A^TA\right)V_1 & V_1^T\left(A^TA\right)V_2\\V_2^T\left(A^TA\right)V_1 & V_2^T\left(A^TA\right)V_2\end{bmatrix}$

$$=diag(\sigma_1^2,\sigma_2^2,\cdots,\sigma_r^2,0,\cdots,0)$$

$$\therefore V_1^T\left(A^TA\right)V_1=diag(\sigma_1^2,\sigma_2^2,\cdots,\sigma_r^2)=\Sigma_r^2$$

$$\Rightarrow \Sigma_r^{-1}V_1^T\left(A^TA\right)V_1\Sigma_r^{-1}=I_r$$

$$V_2^T\left(A^TA\right)V_2=0,\Rightarrow\left(AV_2\right)^T\left(AV_2\right)=0\Rightarrow AV_2=0$$

奇异值分解

奇异值分解（SVD）定理

令$U_1 = AV_1\Sigma_r^{-1}$,则有

$$U_1^TU_1=\Sigma_r^{-1}V_1^TA^TAV_1\Sigma_r^{-1}=\Sigma_r^{-1}\left(\Sigma_r^2\right)_1\Sigma_r^{-1}=I_r$$

这说明U_1的r个列向量是单位正交向量组。

在R^m中可由U_1的r个列向量扩充为一组标准正交基，即存在

$U_2=\left[u_{r+1},\cdots,u_m\right]$使$U=\left[U_1,U_2\right]$为$m$阶正交阵。

于是有$U^TAV=\begin{bmatrix}U_1^T\\U_2^T\end{bmatrix}A\left[V_1,V_2\right]=\begin{bmatrix}U_1^TAV_1 & U_1^TAV_2\\U_2^TAV_1 & U_2^TAV_2\end{bmatrix}$

由$AV_2=0$得$U_1^TAV_2=0$，$U_2^TAV_2=0$

由$U_1=AV_1\Sigma_r^{-1}$得$U_1^TAV_1=\Sigma_r^{-1}V_1^TA^TAV_1=\Sigma_r^{-1}\Sigma_r^2=\Sigma_r$

由$AV_1=U_1\Sigma_r$得$U_2^TAV_1=U_2^TU_1\Sigma_r=0$, $(\because U_2^TU=0)$

即 $U^TAV=\begin{bmatrix}\Sigma_r & 0\\0 & 0\end{bmatrix}$

奇异值分解

奇异值分解（SVD）定理

求A的奇异值分解的具体步骤:

1、求$A^T A$的特征值，用$A^T A$的正特征值得到A的奇异值和Σ_r;

2、由$A^T A$的特征值求出对应的特征向量，正交化后得到n阶正交阵$V=\left[V_1, V_2\right]\in R^{n\times n}$;

3、计算$U_1 = AV_1\Sigma_r^{-1}$，再扩充为R^m中的标准正交基，得到正交阵$U\in R^{m\times m}$.

奇异值分解

奇异值分解（SVD）定理

例： 设$A=\begin{bmatrix}1&1\\1&1\\0&0\end{bmatrix}$，求$A$的奇异值分解.

解： $A^TA=\begin{bmatrix}2&2\\2&2\end{bmatrix}$，　　秩$(A^TA)=1$，

其特征值$\lambda_1=4,\lambda_2=0$，

A的奇异值$\sigma_1=\sqrt{4}=2$，

A的奇异值矩阵为$\Sigma=\begin{bmatrix}2&0\\0&0\\0&0\end{bmatrix}=\begin{bmatrix}\Sigma_r&0\\0&0\\0&0\end{bmatrix}$

$\Sigma_r=2$

奇异值分解

奇异值分解（SVD）定理

解$x_1+x_2=0$，得$u_2=\left(\frac{1}{\sqrt{2}},-\frac{1}{\sqrt{2}},0\right)^T$，$u_3=(0,0,1)^T$，

$$得正交矩阵U=\begin{bmatrix}\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0\\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0\\ 0 & 0 & 1\end{bmatrix}$$

$$矩阵A的奇异值分解是A=U\Sigma V^T=U\begin{bmatrix}\Sigma_r & 0\\ 0 & 0\end{bmatrix}V^T$$

$$=\begin{bmatrix}\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0\\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}2 & 0\\ 0 & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}}\end{bmatrix}$$

奇异值分解

- **奇异值分解（SVD）定理**

  计算过程很复杂，怎么简单点计算

定理 12.10 设 A 是 $m\times n$ 矩阵，那么存在 R^n 和 R^m 的两个正交基 $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_n\}$，$\{\boldsymbol{u}_1,\cdots,\boldsymbol{u}_m\}$，以及实数 $s_1\geqslant\cdots\geqslant s_n\geqslant 0$ 使得对于 $1\leqslant i\leqslant\min\{m,n\}$，$\boldsymbol{A}\boldsymbol{v}_i=s_i\boldsymbol{u}_i.\boldsymbol{V}=[\boldsymbol{v}_1|\cdots|\boldsymbol{v}_n]$ 的列，即右奇异向量，是 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 的正交特征向量组；$\boldsymbol{U}=[\boldsymbol{u}_1|\cdots|\boldsymbol{u}_m]$ 的列，即左奇异向量，是 $\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}$ 的正交特征向量组.

如果A 是对称矩阵，分解是怎样的?

对于m远远大于n，怎么计算AA’的特征向量?

奇异值分解

- **奇异值分解（SVD）定理**

A′ A和AA′ ，这两个矩阵特征向量有什么联系?

定理：x是 A′ A的特征向量，则Ax是 AA′ 的特征向量。

奇异值分解

- **作业：**

  无

群名称：数值计算方法
群　号：1132838842

矩阵特征值

THE END