# 矩阵的奇异值分解

董波
数学科学学院
大连理工大学

- 对于方阵,利用其特征值和特征向量可以刻画矩阵的结构。
- 对长方阵情形,这些方法已经不适用.而矩阵的奇异值分解理论能改善这种情况。

# 矩阵分解

Jordan分解：$A = TJT^{-1}$

$T$可逆阵，　$J$若当标准型

Schur分解：$A = URU^H$

$U$酉矩阵，　$R$上三角矩阵

分解：　　$A = UDV^H$

$U, V$酉矩阵，$D$对角矩阵

变换矩阵

可逆$\longrightarrow$酉矩阵

宽松$\longrightarrow$严格

标准型：

双对角$\longrightarrow$上三角

严格$\longrightarrow$宽松

# 构造性证明

假定矩阵 $A \in C^{m\times n}$ $\quad \operatorname{rank}(A) = r$

$A^H A$ 为Hermite半正定矩阵，特征值均为实数且非负

$$\operatorname{rank}(A^H A) = \operatorname{rank}(A)$$

由Schur定理的推论，存在 $n$ 阶酉阵 $V$，使得

$$\left(U^H A V\right)^H \left(U^H A V\right) = \begin{pmatrix} (\sqrt{\lambda_1})^2 & & \\ & \ddots & \\ & & (\sqrt{\lambda_n})^2 \end{pmatrix} = \begin{pmatrix} \Sigma^2 & 0 \\ 0 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix}^H \begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix}$$

# 矩阵U的构造

矩阵划分：

$$V=\begin{pmatrix}V_1 & V_2\end{pmatrix},\quad V_1\in C^{n\times r}, V_2\in C^{n\times(n-r)}\qquad U=\begin{pmatrix}U_1 & U_2\end{pmatrix},\quad U_1\in C^{m\times r}, U_2\in C^{m\times(m-r)}$$

目标：寻找矩阵$U_1,U_2$使得

$$\begin{pmatrix}\boldsymbol{\Sigma} & 0\\ 0 & 0\end{pmatrix}=\boldsymbol{U^H AV}=\begin{pmatrix}U_1^H\\ U_2^H\end{pmatrix}A\begin{pmatrix}V_1 & V_2\end{pmatrix}=\begin{pmatrix}U_1^H AV_1 & U_1^H AV_2\\ U_2^H AV_1 & U_2^H AV_2\end{pmatrix}$$

$$U_1^H AV_1=\boldsymbol{\Sigma}\qquad U_2^H AV_1=0$$

$$U_1^H AV_2=0\qquad U_2^H AV_2=0$$

由于

$$\boldsymbol{V^H A^H AV}=\begin{pmatrix}V_1^H\\ V_2^H\end{pmatrix}\boldsymbol{A^H}A\begin{pmatrix}V_1 & V_2\end{pmatrix}=\begin{pmatrix}V_1^H\boldsymbol{A^H}AV_1 & V_1^H\boldsymbol{A^H}AV_2\\ V_2^H\boldsymbol{A^H}AV_1 & V_2^H\boldsymbol{A^H}AV_2\end{pmatrix}=\begin{pmatrix}(AV_1)^H AV_1 & (AV_1)^H AV_2\\ (AV_2)^H AV_1 & (AV_2)^H AV_2\end{pmatrix}=\begin{pmatrix}\boldsymbol{\Sigma^2} & 0\\ 0 & 0\end{pmatrix}$$

故

$$(AV_1)^H AV_1=\boldsymbol{\Sigma^2}\quad (AV_2)^H AV_2=\boldsymbol{0}\quad (AV_1)^H AV_2=\boldsymbol{0}\quad (AV_2)^H AV_1=\boldsymbol{0}$$

从而 $U_1=\left(\boldsymbol{\Sigma}^{-1}(AV_1)^H\right)^H$ $\qquad U_2$为$U_1$的标准正交补

# 奇异值分解定理

假定$\boldsymbol{A}\in\boldsymbol{C}^{m\times n}$，且 $\operatorname{rank}(A)=r$，则存在$m$阶、$n$阶 酉阵$U$、$V$ 使得

$$\boldsymbol{A}=\boldsymbol{U}\begin{pmatrix}\boldsymbol{\Sigma} & 0\\ 0 & 0\end{pmatrix}\boldsymbol{V}^{H}$$

其中 $\boldsymbol{\Sigma}=\operatorname{diag}(\sigma_1,\ldots,\sigma_r),\ \sigma_i\neq 0,(i=1,2,\cdots,r).$

非负实数$\sigma_i$称为矩阵$A$的非零奇异值。

奇异值分解亦可写为

$$\boldsymbol{A}=\boldsymbol{U}\begin{pmatrix}\boldsymbol{\Sigma} & 0\\ 0 & 0\end{pmatrix}\boldsymbol{V}^{H}=\begin{pmatrix}U_1 & U_2\end{pmatrix}\begin{pmatrix}\boldsymbol{\Sigma} & 0\\ 0 & 0\end{pmatrix}\begin{pmatrix}V_1^H\\ V_2^H\end{pmatrix}=U_1\boldsymbol{\Sigma}V_1^H$$

$\boldsymbol{A}=\boldsymbol{U}_1\Sigma\boldsymbol{V}_1^H$ 称为矩阵$A$ 约化的奇异值分解。

## 奇异值酉不变性

设 $A$、$B \in \mathbf{C}^{m\times n}$, 如果存在$m$阶、$n$阶酉阵$U$、$V$, 使得

$$\boldsymbol{A}=\boldsymbol{U}\boldsymbol{B}\boldsymbol{V}^{H},$$

则矩阵$A$、$B$的奇异值相同。

证: 由 $\boldsymbol{U}^{H}\boldsymbol{A}\boldsymbol{V}=\boldsymbol{B}$ , 则有

$$\boldsymbol{B}^{H}\boldsymbol{B}=\left(\boldsymbol{U}^{H}\boldsymbol{A}\boldsymbol{V}\right)^{H}\left(\boldsymbol{U}^{H}\boldsymbol{A}\boldsymbol{V}\right)=\boldsymbol{V}^{H}\boldsymbol{A}^{H}\left(\boldsymbol{U}\boldsymbol{U}^{H}\right)\boldsymbol{A}\boldsymbol{V}=\boldsymbol{V}^{H}\left(\boldsymbol{A}^{H}\boldsymbol{A}\right)\boldsymbol{V}$$

即 $\boldsymbol{B}^{H}\boldsymbol{B}$ 与 $\boldsymbol{A}^{H}\boldsymbol{A}$ 相似, 故它们具有相同的特征值, 命题得证。

# 左右奇异向量

$$A = U_1 \Sigma V_1^H \quad \Rightarrow \quad U_1^H A = \Sigma V_1^H, \quad u_i^H A = \sigma_i v_i^H, i = 1,2,\cdots,r$$

$U$的列向量$u_1,u_2,\cdots,u_m$称为矩阵$A$的与奇异值$\sigma_i$对应的**左奇异向量**

$$A = U_1 \Sigma V_1^H \quad \Rightarrow \quad AV_1 = U_1 \Sigma, \qquad Av_i = \sigma_i u_i, i = 1,2,\ldots,r$$

$V$的列向量$v_1,v_2,\cdots,v_n$称为矩阵$A$的与奇异值$\sigma_i$对应的**右奇异向量**.

| 长方阵的奇异值<br>左右奇异向量 | ⟺ | 方阵的特征值<br>左右特征向量 |
| --- | --- | --- |

# 左右奇异向量（续）

$$\boldsymbol{A}\boldsymbol{A}^H\boldsymbol{U}=\boldsymbol{U}\begin{pmatrix}\boldsymbol{\Sigma} & 0\\ 0 & 0\end{pmatrix}V^H\boldsymbol{V}\begin{pmatrix}\boldsymbol{\Sigma} & 0\\ 0 & 0\end{pmatrix}\boldsymbol{U}^H\boldsymbol{U}=\boldsymbol{U}\begin{pmatrix}\boldsymbol{\Sigma}^2 & 0\\ 0 & 0\end{pmatrix}$$

$$\boldsymbol{A}^H\boldsymbol{A}\boldsymbol{V}=\boldsymbol{V}\begin{pmatrix}\boldsymbol{\Sigma} & 0\\ 0 & 0\end{pmatrix}\boldsymbol{U}^H\boldsymbol{U}\begin{pmatrix}\boldsymbol{\Sigma} & 0\\ 0 & 0\end{pmatrix}\boldsymbol{V}^H\boldsymbol{V}=\boldsymbol{V}\begin{pmatrix}\boldsymbol{\Sigma}^2 & 0\\ 0 & 0\end{pmatrix}$$

左奇异向量 $\boldsymbol{u}_1,\boldsymbol{u}_2,\cdots,\boldsymbol{u}_m$ 为$AA^H$的单位正交特征向量

右奇异向量 $\boldsymbol{v}_1,\boldsymbol{v}_2,\cdots,\boldsymbol{v}_n$ 为$A^HA$的单位正交特征向量.

# 例题

例　求矩阵 $A=\begin{pmatrix}1&0&1\\0&1&1\\0&0&0\end{pmatrix}$ 的奇异值分解。

解　求解次序为:　$\sum$, $V$, $V_1, U_1, U$。 计算矩阵

$$A^H A=\begin{pmatrix}1&0&0\\0&1&0\\1&1&0\end{pmatrix}\begin{pmatrix}1&0&1\\0&1&1\\0&0&0\end{pmatrix}=\begin{pmatrix}1&0&1\\0&1&1\\1&1&2\end{pmatrix}$$

令　$\det(\lambda I - A^H A)=\lambda(\lambda-1)(\lambda-3)=0$

则 $A^H A$ 的特征值和$A$的奇异值分别为

$$\lambda_1=3,\quad \lambda_2=1,\quad \lambda_3=0;$$

$$\sigma_1=\sqrt{3},\ \sigma_2=1,\quad \sigma_3=0$$

所以　$\Sigma=\begin{pmatrix}\sqrt{3}&0\\0&1\end{pmatrix}$

## 特征向量　　标准正交

$$\begin{pmatrix} 2 & 0 & -1 \\ 0 & 2 & -1 \\ -1 & -1 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \qquad p_1=\begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} \qquad v_1=\frac{p_1}{\|p_1\|_2}=\frac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}$$

$$\begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & -1 \\ -1 & -1 & -1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \qquad p_2=\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \qquad v_2=\frac{p_2}{\|p_2\|_2}=\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} -1 & 0 & -1 \\ 0 & -1 & -1 \\ -1 & -1 & -2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \qquad p_3=\begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix} \qquad v_3=\frac{p_3}{\|p_3\|_2}=\frac{1}{\sqrt{3}}\begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}$$

$$V=\begin{pmatrix} \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{-1}{\sqrt{2}} & \frac{-1}{\sqrt{3}} \\ \frac{2}{\sqrt{6}} & 0 & \frac{1}{\sqrt{3}} \end{pmatrix}$$

# 约化奇异值分解

因rank($A$)=2，故有

$$V_1=\begin{pmatrix}\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{6}} & \frac{-1}{\sqrt{2}}\\ \frac{2}{\sqrt{6}} & 0\end{pmatrix} \quad U_1=AV_1\Sigma^{-1}=\begin{pmatrix}1&0&1\\0&1&1\\0&0&0\end{pmatrix}\begin{pmatrix}\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{6}} & \frac{-1}{\sqrt{2}}\\ \frac{2}{\sqrt{6}} & 0\end{pmatrix}\begin{pmatrix}\frac{1}{\sqrt{3}} & 0\\ 0 & 1\end{pmatrix}=\begin{pmatrix}\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}}\\ 0 & 0\end{pmatrix}$$

得约化的奇异值分解

$$A=U_1\Sigma V_1^H=\begin{pmatrix}\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}}\\ 0 & 0\end{pmatrix}\begin{pmatrix}\sqrt{3} & 0\\ 0 & 1\end{pmatrix}\begin{pmatrix}\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}}\\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0\end{pmatrix}$$

# 满奇异值分解

计算$U_2$, 使其与$U_1$构成$\mathrm{R}^3$的一组标准正交基，可取 $\boldsymbol{U}_2=\begin{pmatrix}0\\0\\1\end{pmatrix}$ 则

$$\boldsymbol{U}=\begin{pmatrix}\boldsymbol{U}_1 & \boldsymbol{U}_2\end{pmatrix}=\begin{pmatrix}\frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0\\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0\\ 0 & 0 & 1\end{pmatrix}$$

是酉阵, 故矩阵$A$的奇异值分解（满的奇异值分解）为

$$\boldsymbol{A}=\begin{pmatrix}\frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0\\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}\sqrt{3} & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix}\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}}\\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0\\ \frac{-1}{\sqrt{3}} & \frac{-1}{\sqrt{3}} & \frac{1}{\sqrt{3}}\end{pmatrix}$$

# 奇异值分解讨论矩阵的性质

## 像空间、零空间

R($A$)为由$A$的列向量生成的子空间, 称为$A$的值域或像空间. 即

$$R(\boldsymbol{A})=\operatorname{span}\{\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_r\}$$

N($A$)称为$A$的零空间或核空间, 即

$$\mathbf{N}(\boldsymbol{A})=\{\boldsymbol{x}\,|\,\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}\}$$

## 像空间、零空间基底

$$\mathbf{R}(\boldsymbol{A})=\operatorname{span}\{\boldsymbol{u}_1,\boldsymbol{u}_2,\cdots,\boldsymbol{u}_r\},\ N(\boldsymbol{A})=\operatorname{span}\{\boldsymbol{v}_{r+1},\boldsymbol{v}_{r+2},\cdots,\boldsymbol{v}_n\}$$

其中 $u_i$ 和 $v_i$ 分别为矩阵$U$和$V$的正交向量。

Hermite矩阵的奇异值为其特征值的绝对值.

$$A^H = A \qquad \sigma(A) = \sqrt{\lambda(A^H A)} = \sqrt{\lambda(A^2)} = \sqrt{\lambda(A)^2} = |\lambda(A)|$$

矩阵$A$的非零奇异值的个数恰为矩阵$A$的秩.

$$rank(A^H A) = rank(AA^H) = rank(A^H) = rank(A)$$

设 $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$ 则

$$\|A\|_2 = \sigma_1 \qquad \|A\|_F = \sqrt{\sigma_1^2 + \sigma_2^2 + \cdots + \sigma_r^2} \qquad |\det(A)| = \prod_{i=1}^{n} \sigma_i$$

$$\sigma_1 = \sqrt{\lambda_{\max}(A^H A)} = \|A\|_2 \qquad \sigma_1^2 + \sigma_2^2 + \cdots + \sigma_r^2 = \mathrm{trace}(A^H A) = \|A\|_F^2$$

$$\prod_{i=1}^{n} \sigma_i = \sqrt{\prod_{i=1}^{n} \sigma_i^2} = \sqrt{\det(A^H A)} = \sqrt{\det(A^H)\det(A)} = \sqrt{\overline{\det(A)}\det(A)} = \sqrt{|\det(A)|^2}$$

# 例题

例：矩阵A的非零奇异值为1,3,5

- A的秩为多少?
- $\|\boldsymbol{A}\|_2$、$\|A\|_F$、$|\det(\boldsymbol{A})|$ 分别为多少?

例：计算矩阵A的奇异值，其中矩阵A为

$$A=\begin{pmatrix}1 & 2 & 1\\ 2 & -1 & 0\\ 1 & 0 & 1\end{pmatrix}$$

# 约化奇异值分解

秩为$r$的 $m\times n$ 阶矩阵$A$可以表示为 $r$ 个秩为1的矩阵的和

$$\boldsymbol{A}=\sigma_1\boldsymbol{u}_1\boldsymbol{v}_1^H+\sigma_2\boldsymbol{u}_2\boldsymbol{v}_2^H+\cdots+\sigma_r\boldsymbol{u}_r\boldsymbol{v}_r^H$$

$$\boldsymbol{A}=\boldsymbol{U}_1\boldsymbol{\Sigma}\boldsymbol{V}_1^H=(\boldsymbol{u}_1,\ldots,\boldsymbol{u}_r)\begin{pmatrix}\sigma_1 & & \\ & \ddots & \\ & & \sigma_r\end{pmatrix}\begin{pmatrix}\boldsymbol{v}_1^H\\ \vdots\\ \boldsymbol{v}_r^H\end{pmatrix}=(\sigma_1\boldsymbol{u}_1,\ldots,\sigma_r\boldsymbol{u}_r)\begin{pmatrix}\boldsymbol{v}_1^H\\ \vdots\\ \boldsymbol{v}_r^H\end{pmatrix}$$

$$=\sigma_1\boldsymbol{u}_1\boldsymbol{v}_1^H+\cdots+\sigma_r\boldsymbol{u}_r\boldsymbol{v}_r^H$$

$$rank(\boldsymbol{u}_i\boldsymbol{v}_i^H)=rank(\boldsymbol{v}_i^H\boldsymbol{u}_i)=1$$

原图

黑白图(k=875)

k=10

k=20

k=40

k=80

# 矩阵奇异值分解几何意义

$$y = Ax = \beta_1 u_1 + \beta_2 u_2$$

$$x = \alpha_1 v_1 + \alpha_2 v_2$$

$$A = (u_1 \quad u_2)\begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix}\begin{pmatrix} v_1^H \\ v_2^H \end{pmatrix}$$

$$Av_1 = \sigma_1 u_1, Av_2 = \sigma_2 u_2$$

$$y = Ax = \alpha_1 Av_1 + \alpha_2 Av_2$$
$$= \alpha_1 \sigma_1 u_1 + \alpha_2 \sigma_2 u_2$$

$$\beta_1 = \alpha_1 \sigma_1, \beta_2 = \alpha_2 \sigma_2$$

向量 $x$ 坐标满足

$$\alpha_1^2 + \alpha_2^2 = 1$$ 单位圆

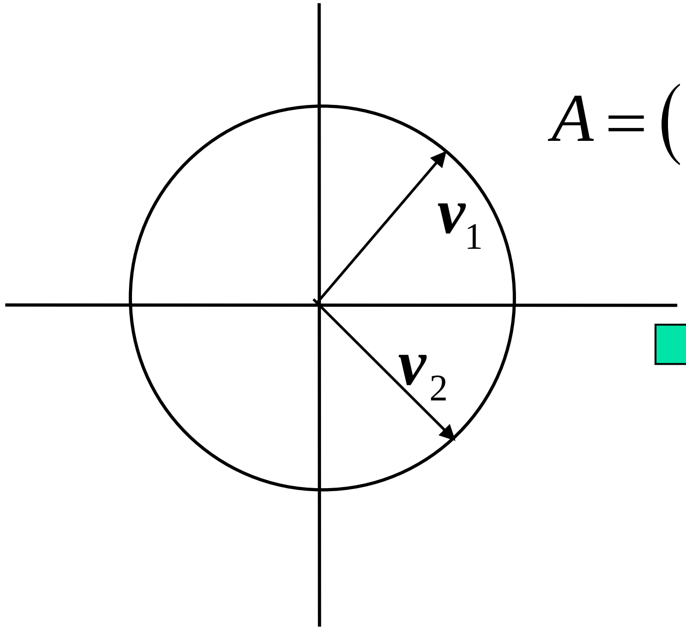
对应向量 $y$ 坐标满足

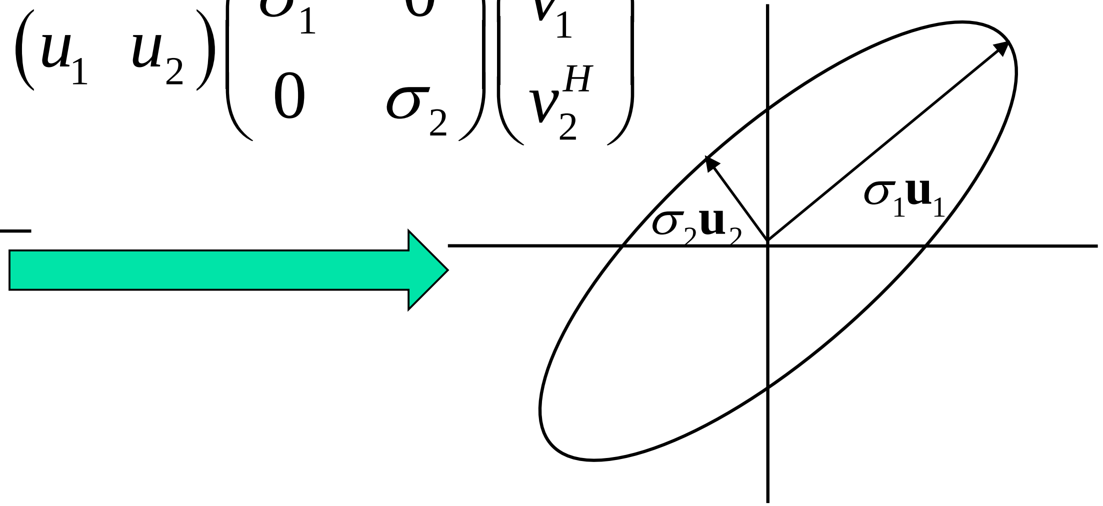
$$\left(\frac{\beta_1}{\sigma_1}\right)^2 + \left(\frac{\beta_2}{\sigma_2}\right)^2 = 1$$ 椭圆

# 矩阵奇异值分解几何意义

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \qquad y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = A \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$A = \begin{pmatrix} u_1 & u_2 \end{pmatrix} \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix} \begin{pmatrix} v_1^H \\ v_2^H \end{pmatrix}$$



# 矩阵奇异值分解几何意义

高维情形

向量 $y = Ax$

$$x = \alpha_1 v_1 + \cdots + \alpha_n v_n,\quad y = \beta_1 u_1 + \cdots + \beta_n u_n$$

仍有 $Av_i = \sigma_i u_i$ $\qquad \beta_1 = \alpha_1\sigma_1, \cdots, \beta_n = \alpha_n\sigma_n$

单位球

$$\alpha_1^2 + \cdots + \alpha_n^2 = 1$$

超椭球：将单位球沿某些正交方向分别以拉伸因子 $\sigma_i$ 拉伸而成的曲面，$u_i$为主半轴，$\sigma_i$为主半轴的长度，恰好是矩阵的奇异值.

$$\left(\frac{\beta_1}{\sigma_1}\right)^2 + \cdots + \left(\frac{\beta_n}{\sigma_n}\right)^2 = 1$$