7.3 奇异值分解

7.3 奇异值分解

回顾：正交矩阵

定义 如果 n 阶矩阵 A 满足 $\mathrm{A}^T\mathrm{A}=\mathrm{I}$，即 $\mathrm{A}^{-1}=\mathrm{A}^T$，则 A 称之为正交矩阵。

A 为正交矩阵 $\Leftrightarrow$
A 的列向量都是单位向量，且两两正交

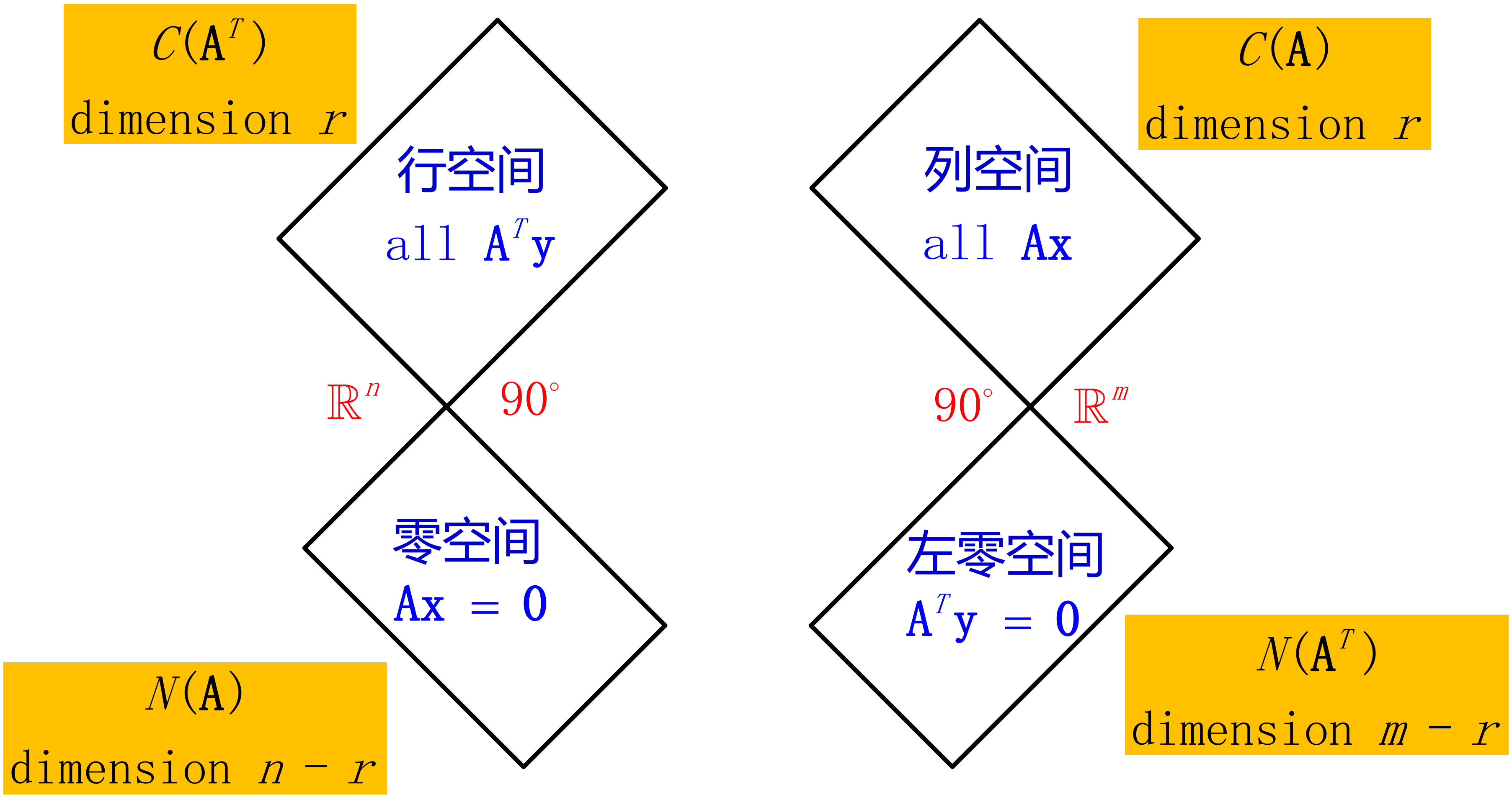
$$\begin{pmatrix} \mathbf{a}_1^T \\ \mathbf{a}_2^T \\ \vdots \\ \mathbf{a}_n^T \end{pmatrix}(\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_n)=\mathrm{I}\Rightarrow \mathbf{a}_i^T\mathbf{a}_j=\begin{cases}1, & \text{当 } i=j, \\ 0, & \text{当 } i\neq j\end{cases}\quad (i,j=1,2,\cdots,n),$$

单位向量

两两正交

7.3 奇异值分解

回顾：四个子空间



7.3 奇异值分解

回顾：相似矩阵和相似变换

定义

设 A, B 都是 n 阶矩阵，若有可逆矩阵 P，使 $P^{-1}AP = B$，则称 B 是 A 的相似矩阵，或是矩阵 A 和 B 相似。
该运算称为对 A 进行相似变换。可逆矩阵 P 被称为把 A 变成 B 的相似变换矩阵。

相似矩阵 Similar matrix	相似变换 Similarity transformation

7.3 奇异值分解

回顾：对称矩阵的性质

$$A = P\Lambda P^{-1}$$

普通情况

$$A = ?\,\Lambda\,?$$

对称矩阵

定理

每个对称矩阵都分解成 $A = Q\Lambda Q^T$, Λ 为包括 A 特征值的对角矩阵，Q 的列由 A 标准正交的特征向量组成。

$A = Q\Lambda Q^{-1} = Q\Lambda Q^T, Q^{-1} = Q^T$

7.3 奇异值分解

什么是SVD?

n x n

$$A = P\Lambda P^{-1}$$

普通情况

$$A = Q\Lambda Q^{T}$$

对称矩阵

---

m x n

$$A = U\Sigma V^{T}\ ?$$

m x n　　m x m　　m x n　　n x n

7.3 奇异值分解

(1) SVD 产生基本子空间中的含若干 $\mathbf{v}$ 向量和 $\mathbf{u}$ 向量的标准正交基

(2) $\Rightarrow$ A 转换成对角矩阵 Σ，并满足 $A\mathbf{v}_i = \sigma_i\mathbf{u}_i$: σ_i 为奇异值

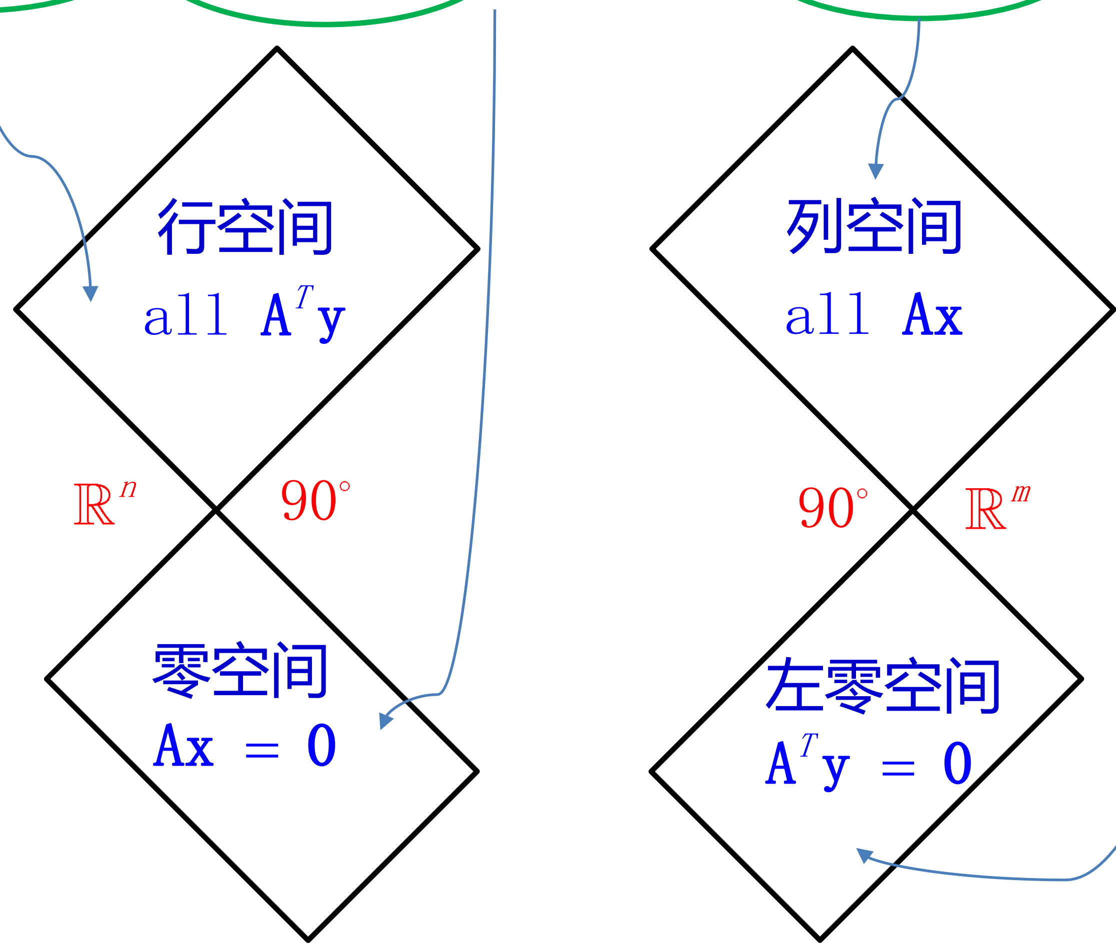
(3) 两基对角化 $A = U\Sigma V^T$ 信息多于 $A = X\Lambda X^{-1}$

(4) $A = U\Sigma V^T = \mathbf{u}_1\sigma_1\mathbf{v}_1^T + \cdots + \mathbf{u}_r\sigma_r\mathbf{v}_r^T$

$$U = \begin{pmatrix}\mathbf{u}_1 & \cdots & \mathbf{u}_r & \mathbf{u}_{r+1} & \cdots & \mathbf{u}_m\end{pmatrix}, V = \begin{pmatrix}\mathbf{v}_1 & \cdots & \mathbf{v}_r & \mathbf{v}_{r+1} & \cdots & \mathbf{v}_n\end{pmatrix}$$

7.3 奇异值分解

$$\mathbf{V}=\left(\mathbf{v}_1\ \cdots\ \mathbf{v}_r\ \mathbf{v}_{r+1}\ \cdots\ \mathbf{v}_n\right),\quad \mathbf{U}=\left(\mathbf{u}_1\ \cdots\ \mathbf{u}_r\ \mathbf{u}_{r+1}\ \cdots\ \mathbf{u}_m\right)$$



7.3 奇异值分解

m x m

正交矩阵

n x n

正交矩阵

$$A = U\Sigma V^T$$

任意矩阵

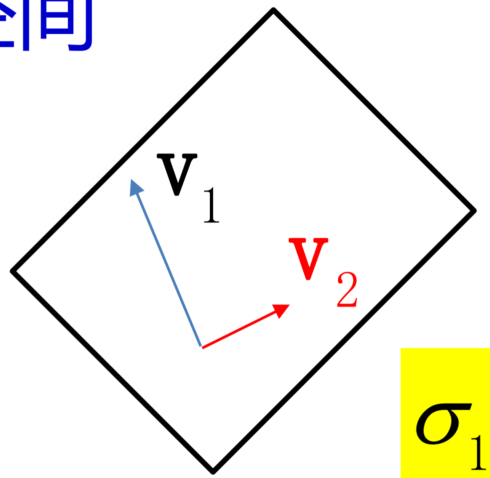
m x n

对角矩阵

m x n

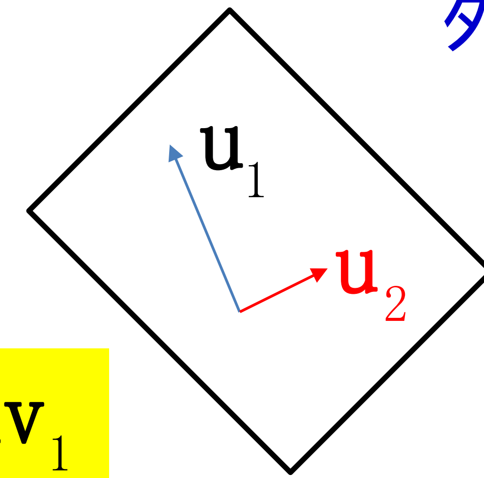
如何获得SVD?

行空间



$\mathbb{R}^n$

列空间



$\mathbb{R}^m$

$$\sigma_1 \mathbf{u}_1 = A\mathbf{v}_1$$

$$\sigma_2 \mathbf{u}_2 = A\mathbf{v}_2$$

$$\cdots$$

一组正交基 → 另一组正交基

7.3 奇异值分解

如何获得SVD?

$\cdots \longrightarrow$ **矩阵表达形式**

$$A\begin{pmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_r \end{pmatrix} = \begin{pmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_r \end{pmatrix} \begin{pmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_r \end{pmatrix}$$

$$\Downarrow$$

$$\underbrace{A}_{m \times n}\underbrace{V_r}_{n \times r} = \underbrace{U_r}_{m \times r}\underbrace{\Sigma_r}_{r \times r}$$

如何获得SVD?

$$A\begin{pmatrix} v_1 & v_2 & \cdots & v_r \end{pmatrix} = \begin{pmatrix} u_1 & u_2 & \cdots & u_r \end{pmatrix}\begin{pmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_r \end{pmatrix}$$

$$\Downarrow$$

$$A\begin{pmatrix} v_1 & v_2 & \cdots & v_r & v_{r+1} & \cdots & v_n \end{pmatrix} =$$

$$\begin{pmatrix} u_1 & u_2 & \cdots & u_r & u_{r+1} & \cdots & u_m \end{pmatrix}\begin{pmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_r & \\ & & & \\ & & & \end{pmatrix}$$

$$\overset{m \times n}{A}\,\overset{n \times n}{V} = \underset{m \times m}{U}\,\underset{m \times n}{\Sigma}$$

7.3 奇异值分解

如何获得SVD?

$$AV = U\Sigma \Rightarrow AVV^{-1} = U\Sigma V^{-1} = U\Sigma V^T$$

V 标准正交矩阵

$$A = U\Sigma V^T$$

$$A^TA = (U\Sigma V^T)^T(U\Sigma V^T) = V\Sigma^T U^T U\Sigma V^T = V\Sigma^T\Sigma V^T$$

$\sigma_i^2 = \lambda_i$ of A^TA

7.3 奇异值分解

例1 将矩阵 $A = \begin{pmatrix} 3 & 0 \\ 4 & 5 \end{pmatrix}$ 奇异值分解。

解

$$A^T A = \begin{pmatrix} 25 & 20 \\ 20 & 25 \end{pmatrix} \Rightarrow \sigma_1^2 = \lambda_1 = 45, \quad \sigma_2^2 = \lambda_2 = 5$$

$$\begin{pmatrix} 25 & 20 \\ 20 & 25 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = 45\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 25 & 20 \\ 20 & 25 \end{pmatrix}\begin{pmatrix} -1 \\ 1 \end{pmatrix} = 5\begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

$$\mathbf{v}_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \mathbf{v}_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 1 \end{pmatrix}, \qquad \text{(注：标准正交)}$$

解 ……

$$\mathbf{A}\mathbf{v}_1=\sigma_1\mathbf{u}_1 \Rightarrow \mathbf{u}_1=\frac{\mathbf{A}\mathbf{v}_1}{\sigma_1}=\frac{\begin{pmatrix}3&0\\4&5\end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}}{\sqrt{45}}=\frac{1}{\sqrt{90}}\begin{pmatrix}3\\9\end{pmatrix}=\frac{1}{\sqrt{10}}\begin{pmatrix}1\\3\end{pmatrix}$$

$$\mathbf{A}\mathbf{v}_2=\sigma_2\mathbf{u}_2 \Rightarrow \mathbf{u}_2=\frac{\mathbf{A}\mathbf{v}_2}{\sigma_2}=\frac{\begin{pmatrix}3&0\\4&5\end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix}-1\\1\end{pmatrix}}{\sqrt{5}}=\frac{1}{\sqrt{10}}\begin{pmatrix}-3\\1\end{pmatrix}$$

$$\Rightarrow \mathbf{A}=\mathbf{U}\Sigma\mathbf{V}^T \text{ with } \mathbf{U}=\frac{1}{\sqrt{10}}\begin{pmatrix}1&-3\\3&1\end{pmatrix},\ \Sigma=\begin{pmatrix}\sqrt{45}&0\\0&\sqrt{5}\end{pmatrix},\ \mathbf{V}=\frac{1}{\sqrt{2}}\begin{pmatrix}1&-1\\1&1\end{pmatrix}$$

7.3 奇异值分解

伪逆矩阵

方阵 $A_{n\times n}$ 可逆 $\Rightarrow$ 存在 A^{-1}

一般矩阵 $A_{m\times n}$?可逆? $\Rightarrow$ 存在 $A^{+}_{n\times m}$

$$A^{+} = V\Sigma^{+}U^{T} = \begin{pmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_r & \cdots & \mathbf{v}_n \end{pmatrix} \begin{pmatrix} \sigma_1^{-1} & & \\ & \ddots & \\ & & \sigma_r^{-1} \\ & & & \end{pmatrix} \begin{pmatrix} \mathbf{u}_1 & \cdots & \mathbf{u}_r & \cdots & \mathbf{u}_m \end{pmatrix}^{T}$$

$$A = U\Sigma V^{T}$$

SVD 分解 → 图像处理

Matlab 代码

```
a=imread( 'xxx.png');
    a=double(a);
    [U S V]=svd(a);
```

② 50

```
re=U(:,1:50)*S(1:50,1:50)*S(:,1:50)';
figure;
imshow(mat2gray(re));
imwrite(mat2gray(re),'xxx50.jpg')
```

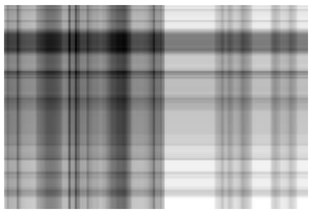
① 10

```
re=U(:,1:10)*S(1:10,1:10)*V(:,1:10)';
figure;
imshow(mat2gray(re));
imwrite(mat2gray(re),'xxx10.jpg')
```

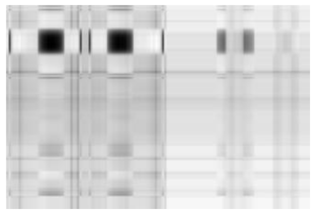
③ 100

```
re=U(:,1:100)*S(1:100,1:100)*V(:,1:100)';
figure;
imshow(mat2gray(re));
imwrite(mat2gray(re),'xxx100.jpg')
```

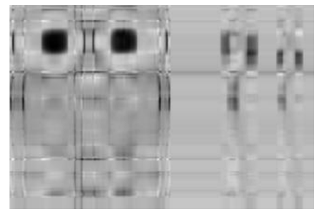
SVD 分解 → 图像处理



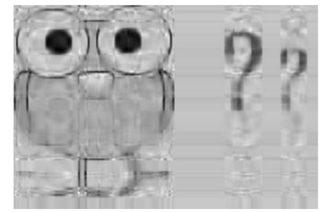
1



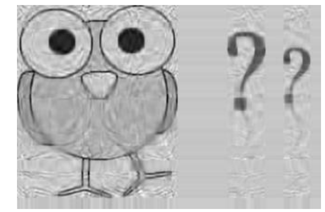
2



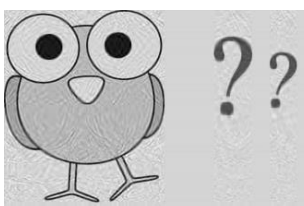
5



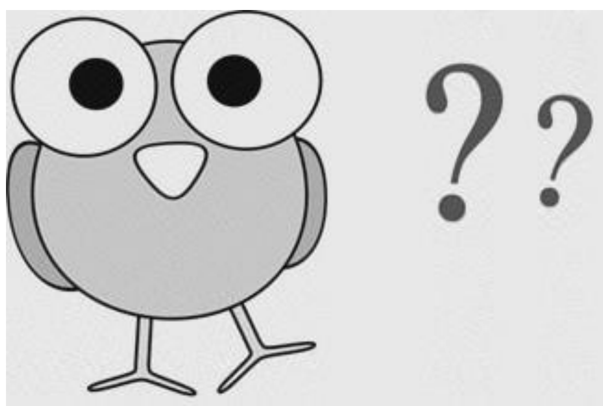
10



20



50



100

原图

- **LAPACK** — **L**inear **A**lgebra **PACK**age

- SVD 的应用领域
  - **图像处理：压缩、去噪**
  - **科学与工程参数估算、反问题**
  - **金融、化工(主成分分析 PCA)**
  - 信号处理
  - 机器学习
  - 气象学
  - …….

7.4 条件优化

例1 求$Q(x)=9x_1^2+4x_2^2+3x_3^2$，在限制条件$\mathbf{x}^T\mathbf{x}=1$下的最大值和最小值。

解 由于x_2^2和x_3^2是非负的，注意到

$$4x_2^2 \le 9x_2^2 \quad 和 \quad 3x_3^2 \le 9x_3^2$$

所以当$x_1^2+x_2^2+x_3^2=1$时，

$$\begin{aligned} Q(x) &= 9x_1^2+4x_2^2+3x_3^2 \\ &\le 9x_1^2+9x_2^2+9x_3^2 \\ &= 9(x_1^2+x_2^2+x_3^2) \\ &= 9 \end{aligned}$$

例1 求$Q(x)=9x_1^2+4x_2^2+3x_3^2$,在限制条件$x^Tx=1$下的最大值和最小值。

解 因此当x是单位向量时，$Q(x)$ 的最大值不超过9. 更进一步， 当$x=(1,0,0)$ 时，$Q(x)=9$. 从而9是$Q(x)$ 在$x^Tx=1$条件下的最大值。为求出$Q(x)$ 的最小值，注意到

$$9x_1^2\geq 3x_1^2,4x_2^2\geq 3x_2^2$$

因此当$x_1^2+x_2^2+x_3^2=1$时，

$$Q(x)\geq 3x_1^2+3x_2^2+3x_3^2=3(x_1^2+x_2^2+x_3^2)=3$$

当$x_1=0,x_2=0,x_3=1$时,$Q(x)$=3，从而3是$Q(x)$在$x^Tx=1$ 条件下的最小值。

例2

设$A=\begin{bmatrix}3 & 0\\0 & 7\end{bmatrix}$,当$x$属于$R^2$时，$Q(x)=x^TAx$,图1是$Q$ 的图形表示，图2表示圆柱体内的一部分，圆柱与曲面的截面是点集(x_1,x_2,z),表示在$x_1^2+x_2^2=1$ 情况下的$z=Q(x_1,x_2)$.这些点的高度值是$Q(x)$ 的约束值，从几何意义上来看，条件优化问题确定的是截面曲线上最高点和最低点的位置。

7.4 条件优化

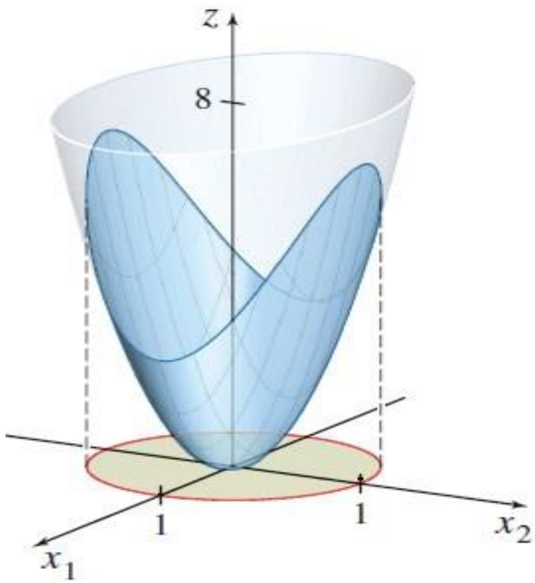


FIGURE 1 $z = 3x_1^2 + 7x_2^2$.

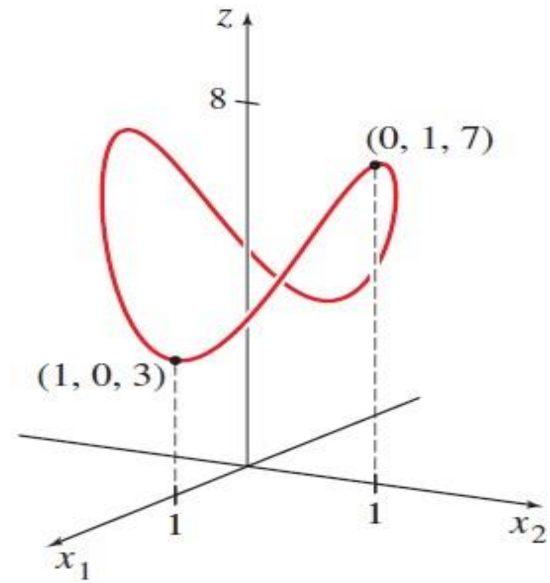


FIGURE 2 The intersection of $z = 3x_1^2 + 7x_2^2$ and the cylinder $x_1^2 + x_2^2 = 1$.

曲线上的两个最高点在x_1x_2平面之上7 个单位，出现在点$x=0$和$x=\pm1$处，这些点对应A 的特征值7 和特征向量$x=(0,1), -x=(0,-1)$.

7.4 条件优化

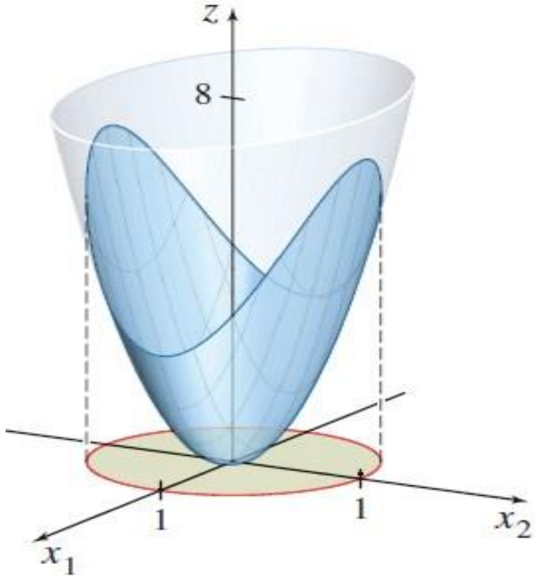


FIGURE 1 $z = 3x_1^2 + 7x_2^2$.

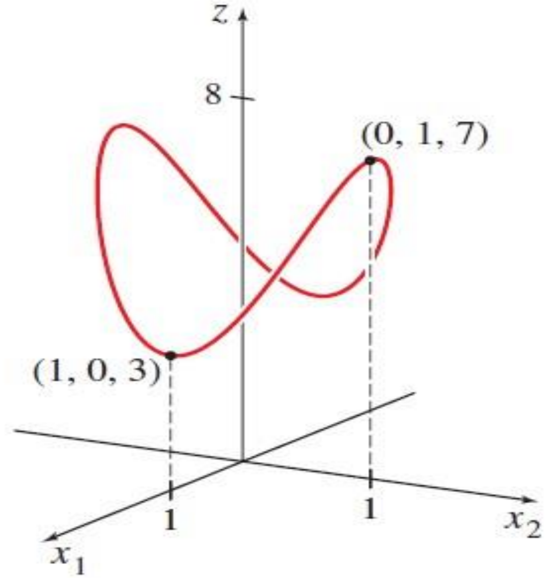


FIGURE 2 The intersection of $z = 3x_1^2 + 7x_2^2$ and the cylinder $x_1^2 + x_2^2 = 1$.

类似的，曲线上的两个最低点在x_1x_2平面之上3 个单位，它们对应特征值3 和特征向量(1, 0) 和(-1 ，0).

7.4 条件优化

定理1

设A是对称矩阵，且

$$m = \min\{x^T Ax : \|x\| = 1\}, \quad M = \max\{x^T Ax : \|x\| = 1\}$$

那么M是A 的最大特征值λ_1，m是A 的最小特征值，如果x是对应M 的单位特征向量u_1，那么$x^T Ax$ 的值等于M，如果x是对应m 的单位特征向量，$x^T Ax$ 的值等于m.

7.4 条件优化

例3

设$A=\begin{bmatrix}3&2&1\\2&3&1\\1&1&4\end{bmatrix}$，求二次型$x^T Ax$在限制条件$x^T x=1$下的最大值，且求一个可以取到该最大值的单位向量。

解 由定理1，我们只需求出A的最大特征值，其多项式是：

$$0=-\lambda^3+10\lambda^2-27\lambda+18=-(\lambda-6)(\lambda-3)(\lambda-1)$$

最大特征值为6.

限制条件下$x^T Ax$的最大值，可以在特征值$\lambda=6$ 对应得单位特征向量x处获得，解$(A-6I)x=0$,我们可得向量

例3

设$A=\begin{bmatrix} 3 & 2 & 1 \\ 2 & 3 & 1 \\ 1 & 1 & 4 \end{bmatrix}$，求二次型$x^T Ax$在限制条件$x^T x=1$下的最大值，且求一个可以取到该最大值的单位向量。

解 我们可得特征向量：

$$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \text{ 和 } u_1=\begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}$$

7.4 条件优化

Next

考虑$x^T Ax$的值，此时，x是单位向量且与定理1中提到的特征向量u_1正交。

7.4 条件优化

定理2 设A，λ_1和u_1如定理1所示，在如下条件的限制下

$$x^T x = 1, \quad x^T u_1 = 0$$

$x^T Ax$的最大值是第二大特征值λ_2，且这个最大值可以在x是对应λ_2的特征向量u_2处达到。

Next

例4给出对角阵情形下的证明思路

7.4 条件优化

例4

求$9x_1^2+4x_2^2+3x_3^2$的最大值，其限制条件是$x^Tx=1$和$x^Tu_1=0$，这里$u_1=(1,0,0)$. 注意到u_1是二次型对应的矩阵的最大特征值$\lambda=9$对应的单位特征向量。

解

如果x的坐标为x_1,x_2,x_3,那么限制$x^Tu_1=0$ 简单意味着x_1=0，对这样一个单位向量，$x_2^2+x_3^2=1$ 且

$$
\begin{aligned}
9x_1^2+4x_2^2+3x_3^2 &= 4x_2^2+3x_3^2 \\
&\le 4x_2^2+4x_3^2 \\
&= 4(x_2^2+x_3^2)=4
\end{aligned}
$$

7.4 条件优化

例4

求$9x_1^2+4x_2^2+3x_3^2$的最大值，其限制条件是$x^Tx=1$和$x^Tu_1=0$，这里$u_1=(1,0,0)$. 注意到u_1是二次型对应的矩阵的最大特征值$\lambda=9$对应的单位特征向量。

解

在这样的限制条件下，二次型的最大值不超过4，且这个最大值可在$x=(0,1,0)$处达到，而这就是二次型对应矩阵的第二大特征值对应的特征向量。

7.4 条件优化

定理3

设A是n 阶对称矩阵，且其正交对角化为$A = PDP^{-1}$，将对角矩阵D上的元素重新排列，使得$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$，且$P$的列是其对应的单位特征向量$u_1, \cdots, u_n$,那么对$k = 2, \cdots, n$ 时，在以下限制条件下

$$x^T x = 1, \quad x^T u_1 = 0, \cdots, x^T u_{k-1} = 0$$

$x^T Ax$的最大值是特征值λ_k,且这个最大值在$x = u_k$ 处可以达到。

Q & A

如何获得SVD?

U 是否为标准正交矩阵?

$i \neq j:$

$$\mathbf{u}_i^T\mathbf{u}_j = \left(\frac{\mathbf{Av}_i}{\sigma_i}\right)^T\left(\frac{\mathbf{Av}_j}{\sigma_j}\right) = \frac{\mathbf{v}_i^T\mathbf{A}^T\mathbf{Av}_j}{\sigma_i\sigma_j}$$

$$= \frac{\mathbf{v}_i^T\mathbf{V}\Sigma^T\Sigma\mathbf{V}^T\mathbf{v}_j}{\sigma_i\sigma_j} = \frac{\mathbf{e}_i^T\Sigma^T\Sigma e_j}{\sigma_i\sigma_j} = \text{ZERO}$$

… $\mathbf{u}_i$ 是 $\mathbf{AA}^T$ 的特征向量

6.2 正交集

定义 如果 $m \times n$ 阶矩阵 U 的向量是标准正交的 $\Leftrightarrow \quad U^T U = I.$

证明 设 $U = \begin{bmatrix} u_1 & u_2 & u_3 \end{bmatrix}$，且 $u_i \in \mathbb{R}^m, i = 1, 2, 3.$那么，

$$U^T U = \begin{bmatrix} u_1^T \\ u_2^T \\ u_3^T \end{bmatrix} \begin{bmatrix} u_1 & u_2 & u_3 \end{bmatrix} = \begin{bmatrix} u_1^T u_1 & u_1^T u_2 & u_1^T u_3 \\ u_2^T u_1 & u_2^T u_2 & u_2^T u_3 \\ u_3^T u_1 & u_3^T u_2 & u_3^T u_3 \end{bmatrix}$$

U是正交矩阵当且仅当

$$u_1^T u_2 = u_2^T u_1 = 0,\ u_1^T u_3 = u_3^T u_1 = 0,\ u_2^T u_3 = u_3^T u_2 = 0$$

$$u_1^T u_1 = 1,\ u_2^T u_2 = 1,\ u_3^T u_3 = 1, \text{即} U^T U = I$$