MATRIX

矩阵论及其应用

主讲老师：徐跃东

2022-2023学年第二学期

矩阵分析

- 向量范数
- 矩阵范数
- 序列极限
- 矩阵幂级数
- 矩阵函数
- 矩阵微积分
- 矩阵函数的应用

矩阵函数

定义：设$\boldsymbol{f(z)}$是复变量的解析函数，$\boldsymbol{f(z)=\sum_{k=0}^{\infty} a_k z^k}$的收敛半径$R$。如果矩阵$A\in C^{n\times n}$的谱半径$\rho(A)<R$, 则称

$$f(A)=\sum_{k=0}^{\infty} a_k A^k$$

为A的矩阵函数。

矩阵函数

例：整个复平面绝对收敛的函数

$$e^{\lambda}=1+\frac{\lambda}{1!}+\frac{\lambda^2}{2!}+\frac{\lambda^3}{3!}+\cdots$$

$$\cos\lambda=1-\frac{\lambda^2}{2!}+\frac{\lambda^4}{4!}-\cdots$$

$$\sin\lambda=\lambda-\frac{\lambda^3}{3!}+\frac{\lambda^5}{5!}-\cdots$$

例：无论A为何矩阵

$$e^{A}=I+\frac{1}{1!}A+\frac{1}{2!}A^2+\frac{1}{3!}A^3+\cdots\cdots$$

$$\cos A=I-\frac{1}{2!}A^2+\frac{1}{4!}A^4-\cdots\cdots$$

$$\sin A=A-\frac{1}{3!}A^3+\frac{1}{5!}A^5-\cdots\cdots$$

矩阵函数

例：整个复平面绝对收敛的函数

$$e^{\lambda}=1+\frac{\lambda}{1!}+\frac{\lambda^2}{2!}+\frac{\lambda^3}{3!}+\cdots$$

$$\cos\lambda=1-\frac{\lambda^2}{2!}+\frac{\lambda^4}{4!}-\cdots$$

$$\sin\lambda=\lambda-\frac{\lambda^3}{3!}+\frac{\lambda^5}{5!}-\cdots$$

例：无论A为何矩阵

$$e^{A}=I+\frac{1}{1!}A+\frac{1}{2!}A^2+\frac{1}{3!}A^3+\cdots\cdots$$

$$\cos A=I-\frac{1}{2!}A^2+\frac{1}{4!}A^4-\cdots\cdots$$

$$\sin A=A-\frac{1}{3!}A^3+\frac{1}{5!}A^5-\cdots\cdots$$

矩阵函数

基本性质：

(1) $e^{kA}\cdot e^{lA}=e^{(k+l)A}$

(2) $(e^{A})^{-1}=e^{-A}$

(3) 若$AB=BA$，则 $e^{A}\cdot e^{B}=e^{B}\cdot e^{A}=e^{A+B}$

(4) $\frac{d}{dt}(e^{tA})=Ae^{tA}=e^{tA}A$

(5) $\frac{d}{dt}(\cos A\,t)=-A\sin A\,t=-\sin A\,t\cdot A$

(6) $\frac{d}{dt}(\sin A\,t)=A\cos A\,t=\cos A\,t\cdot A$

矩阵函数

典型的误解：

数域元素存在： $e^{\lambda_1}\cdot e^{\lambda_2}=e^{\lambda_2}\cdot e^{\lambda_1}=e^{\lambda_1+\lambda_2}$

而 $e^{A}\cdot e^{B}, e^{B}\cdot e^{A}, e^{A+B}$ 往往互不相等。

验证： $A=\begin{pmatrix}1 & 1\\ 0 & 0\end{pmatrix}$ $B=\begin{pmatrix}1 & -1\\ 0 & 0\end{pmatrix}$

矩阵函数

例：设A为反对称矩阵，证明e^{A}为正交矩阵。

例：设 $A=\begin{bmatrix} \frac{3}{2} & 0 & 2 \\ 1 & \frac{-3}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}$，

讨论 lnA 是否有意义

矩阵函数的计算

➢ **约当标准型法**

设 A 为一个 n 阶矩阵，J 为其Jordan标准形，则

$$A = PJP^{-1}$$

且有 $f(A) = a_n A^n + a_{n-1}A^{n-1} + \cdots + a_1 A + a_0 I$

$$= a_n(PJP^{-1})^n + a_{n-1}(PJP^{-1})^{n-1} + \cdots + a_0 I$$

$$= P(a_n J^n + a_{n-1}J^{n-1} + \cdots + a_0 I)P^{-1}$$

$$= Pf(J)P^{-1}$$

矩阵函数的计算

➢ 约当标准型法

$$J_i(\lambda_i)=\begin{bmatrix}\lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i\end{bmatrix}_{d_i\times d_i}\quad (i=1,2,\cdots,r)$$

$$J_i^k(\lambda_i)=\begin{bmatrix}\lambda_i^k & c_k^1\lambda_i^{k-1} & \cdots & c_k^{d_i-1}\lambda_i^{k-d_i+1} \\ & \lambda_i^k & \ddots & \vdots \\ & & \ddots & c_k^1\lambda_i^{k-1} \\ & & & \lambda_i^k\end{bmatrix}_{d_i\times d_i}$$

矩阵函数的计算

➢ 约当标准型法

$$f(J_i)=\begin{bmatrix} f(\lambda_i) & f'(\lambda_i) & \cdots & \dfrac{1}{(d_i-1)!}f^{(d_i-1)}(\lambda_i) \\ & f(\lambda_i) & \ddots & \vdots \\ & & \ddots & f'(\lambda_i) \\ & & & f(\lambda_i) \end{bmatrix}_{d_i\times d_i}$$

矩阵函数的计算

➢ 约当标准型法

例　已知多项式 $f(x)=x^4-2x^3+x-1$

与矩阵 $A=\begin{bmatrix} 3 & 0 & 8 \\ 3 & -1 & 6 \\ -2 & 0 & -5 \end{bmatrix}$，求 $f(A)$

矩阵函数的计算

➢ 约当标准型法

解：首先求出矩阵的 A 的Jordan标准形 P 及其相似变换矩阵

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

$$P=\begin{bmatrix}0 & 4 & 1\\ 1 & 3 & 0\\ 0 & -2 & 0\end{bmatrix}$$

$$P^{-1}=\begin{bmatrix}0 & 1 & 3/2\\ 0 & 0 & -1/2\\ 1 & 0 & 2\end{bmatrix}$$

那么有

矩阵函数的计算

- 约当标准型法

解：$f(A)=Pf(J)P^{-1}$

$$=\begin{bmatrix}0 & 4 & 1\\ 1 & 3 & 0\\ 0 & -2 & 0\end{bmatrix}\begin{bmatrix}f(-1) & 0 & 0\\ 0 & f(-1) & f'(-1)\\ 0 & 0 & f(-1)\end{bmatrix}\begin{bmatrix}0 & 1 & 3/2\\ 0 & 0 & -1/2\\ 1 & 0 & 2\end{bmatrix}$$

$$=\begin{bmatrix}f(-1)+4f'(-1) & 0 & 8f'(-1)\\ 3f'(-1) & f(-1) & 6f'(-1)\\ -2f'(-1) & 0 & f(-1)-4f'(-1)\end{bmatrix}$$

$$=\begin{bmatrix}-35 & 0 & -72\\ -27 & 1 & -54\\ 18 & 0 & 37\end{bmatrix}$$

矩阵函数的计算

➢ 待定系数法

计算矩阵函数 $\boldsymbol{f}(\boldsymbol{A})$ 是基于每个矩阵 $\boldsymbol{A}$ 存在最小多项式；

假设$\boldsymbol{A} \in \boldsymbol{F}^{n\times n}$ 的最小多项式是：

$$m_A(\lambda) = b_0 + b_1\lambda + b_2\lambda^2 + \cdots + \lambda^r$$

多项式 $\boldsymbol{f}(\lambda) = \boldsymbol{q}(\lambda) \circ \boldsymbol{m}_{\boldsymbol{A}}(\lambda) + \boldsymbol{\gamma}(\lambda)$。由于$\boldsymbol{m}_{\boldsymbol{A}}(\lambda)$=0，则

$$\boldsymbol{f}(\boldsymbol{A}) = \boldsymbol{\gamma}(\boldsymbol{A})$$

矩阵函数的计算

➢ 待定系数法

将$\boldsymbol{A}$的最小多项式表示为：

$$m_A(\lambda) = (\lambda - \lambda_1)^{r_1}(\lambda - \lambda_2)^{r_2} \cdots (\lambda - \lambda_s)^{r_s}$$

$$\sum_{i=1}^{s} r_i = r \leq n$$

其中 $\lambda_1 \lambda_2 \cdots \lambda_s$ 是 $\boldsymbol{A}$ 的互异的特征值。

矩阵函数的计算

➢ 待定系数法

定义 $f(\lambda)$ 在 A 的谱上确定：设 A 的最小多项式是 $m_A(\lambda)$ ，如果复函数 $f(\lambda)$ 在A的谱 $\{\lambda_1, \lambda_2, \cdots, \lambda_s\}$ 上有下述确定的值：

$$f(\lambda_j), f'(\lambda_j), \cdots, f^{(r_j-1)}(\lambda_j) \qquad (1 \le j \le s)$$

称 $f(\lambda)$ 在 A 的谱上确定。

矩阵函数的计算

➢ 待定系数法

定理 设 $f(\lambda)$ 和 $g(\lambda)$ 是两个复多项式，两者的次数和系数均可以不同，$A \in P^{n\times n}$，则 $f(A)=g(A)$ 的充分必要条件是 $f(\lambda)$ 和 $g(\lambda)$ 在A的谱上的值完全相同。

$$m_A(\lambda)=(\lambda-\lambda_1)^{n_1}(\lambda-\lambda_2)^{n_2}\cdots(\lambda-\lambda_s)^{n_s} \qquad \sum_{i=1}^{s} n_i = m$$

$$g(\lambda)=c_0+c_1\lambda+c_2\lambda^2+\cdots+c_{k-1}\lambda^{m-1}$$

谱上等值确定g（λ），则f（A）＝g（A）

矩阵函数的计算

- 待定系数法

例1：设 $A=\begin{bmatrix}1&0&0\\0&2&0\\0&0&3\end{bmatrix}$

求 $f(A)$ 的多项式表示并且计算

$$e^{tA}, \sin\frac{\pi}{4}A, \cos\frac{\pi}{4}A$$

解：容易观察出该矩阵的最小多项式为

$$m(x)=(x-1)(x-2)(x-3)$$

矩阵函数的计算

➢ 待定系数法

这是3次多项式，从而存在一个2次的多项式

$$p(x) = a_2x^2 + a_1x + a_0$$

$$\begin{cases} f(\lambda) = g(\lambda)\big|_{\lambda=\lambda_1,\lambda_2,\dots} \\ f(A) = g(A) \end{cases}$$

且满足 $p(1) = f(1), p(2) = f(2), p(3) = f(3)$

于是存在

$$f(1) = a_2 + a_1 + a_0$$

$$f(2) = 4a_2 + 2a_1 + a_0$$

$$f(3) = 9a_2 + 3a_1 + a_0$$

$$f(\lambda) = q(\lambda)\underbrace{m(\lambda)}_{\text{最小多项式}} + p(\lambda)$$

矩阵函数的计算

➢ 待定系数法

$$a_0 = f(3) - 3f(2) + 3f(1)$$
$$a_1 = -\frac{1}{2}(3f(3) - 8f(2) + 5f(1))$$
$$a_2 = \frac{1}{2}(f(3) - 2f(2) + f(1))$$

$$f(A) = a_2A^2 + a_1A + a_0I = \begin{bmatrix} f(1) & 0 & 0 \\ 0 & f(2) & 0 \\ 0 & 0 & f(3) \end{bmatrix}$$

矩阵函数的计算

- 待定系数法

当 $f(x)=e^{tx}$ 时，可得

$$f(1)=e^{t},f(2)=e^{2t},f(3)=e^{3t}$$

于是有

$$e^{tA}=\begin{bmatrix} e^{t} & 0 & 0 \\ 0 & e^{2t} & 0 \\ 0 & 0 & e^{3t} \end{bmatrix}$$

当$f(x)=sin\frac{\pi}{4}x$ 时，可得

$$f(1)=\frac{\sqrt{2}}{2},f(2)=1,f(3)=\frac{\sqrt{2}}{2}$$

矩阵函数的计算

➢ 待定系数法

故有

$$\sin\frac{\pi}{4}A=\begin{bmatrix}\frac{\sqrt{2}}{2} & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & \frac{\sqrt{2}}{2}\end{bmatrix}$$

类似地有

$$\cos\frac{\pi}{4}A=\begin{bmatrix}\frac{\sqrt{2}}{2} & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & -\frac{\sqrt{2}}{2}\end{bmatrix}$$

矩阵函数的计算

➢ 待定系数法

例2：设 $A=\begin{bmatrix}1 & 0 & 0\\ 0 & 2 & 1\\ 0 & 0 & 2\end{bmatrix}$

求$f(A)$ 的多项式表示并且计算

$$e^{tA}, \qquad \sin\pi A, \qquad \cos\frac{\pi}{4}A$$

解：容易观察出该矩阵的最小多项式为

$$m(x)=(x-1)(x-2)^2$$

这是一个3次多项式，从而存在一个次数为2的多项式

矩阵函数的计算

➢ 待定系数法

$$p(x) = a_2x^2 + a_1x + a_0$$

$$p(1) = f(1), p(2) = f(2), p'(2) = f'(2)$$

$$f(1) = a_2 + a_1 + a_0$$
$$f(2) = 4a_2 + 2a_1 + a_0$$
$$f'(2) = 4a_2 + a_1$$

解得

$$a_0 = 2f'(2) - 3f(2) + 4f(1)$$
$$a_1 = -3f'(2) + 4f(2) - 4f(1)$$
$$a_2 = f'(2) - f(2) + f(1)$$

矩阵函数的计算

➢ 待定系数法

矩阵多项式为：

$$f(A)=a_2A^2+a_1A+a_0I=\begin{bmatrix} f(1) & 0 & 0 \\ 0 & f(2) & f'(2) \\ 0 & 0 & f(2) \end{bmatrix}$$

$$f(x)=e^{tx} \qquad f(1)=e^t, f(2)=e^{2t}, f'(2)=te^{2t}$$

$$e^{tA}=\begin{bmatrix} e^t & 0 & 0 \\ 0 & e^{2t} & te^{2t} \\ 0 & 0 & e^{2t} \end{bmatrix}$$

矩阵函数的计算

➢ 待定系数法

当$f(x) = sin\pi x$ 时，可得

$$f(1) = 0, f(2) = 0, f'(2) = \pi$$

故有

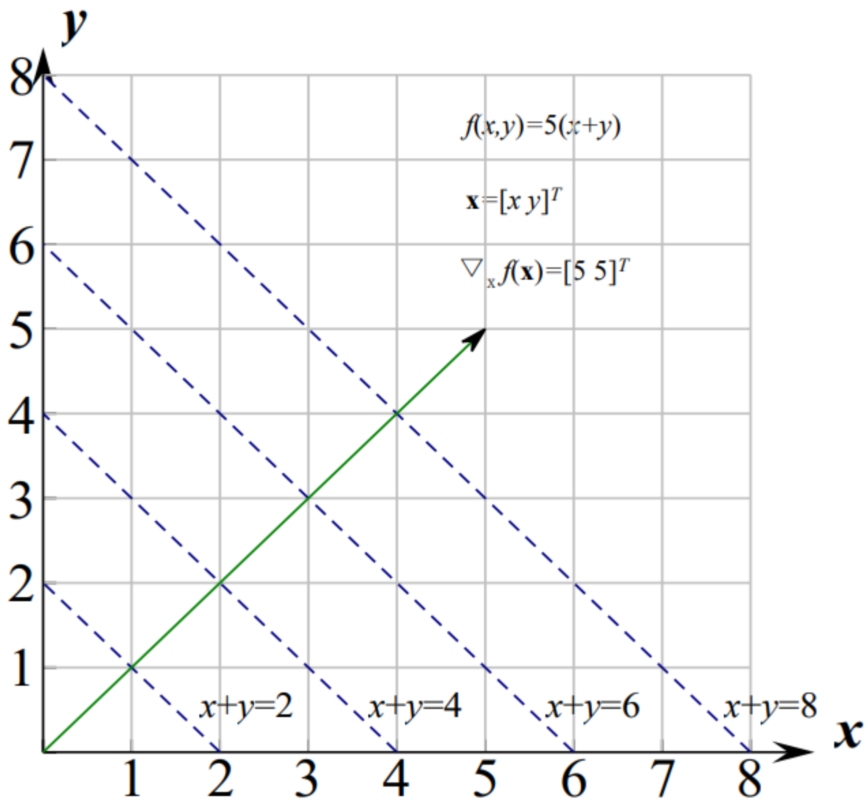
$$\sin \pi A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & \pi \\ 0 & 0 & 0 \end{bmatrix}$$

矩阵分析

- 向量范数
- 矩阵范数
- 序列极限
- 矩阵幂级数
- 矩阵函数
- 矩阵微积分
- 矩阵函数的应用

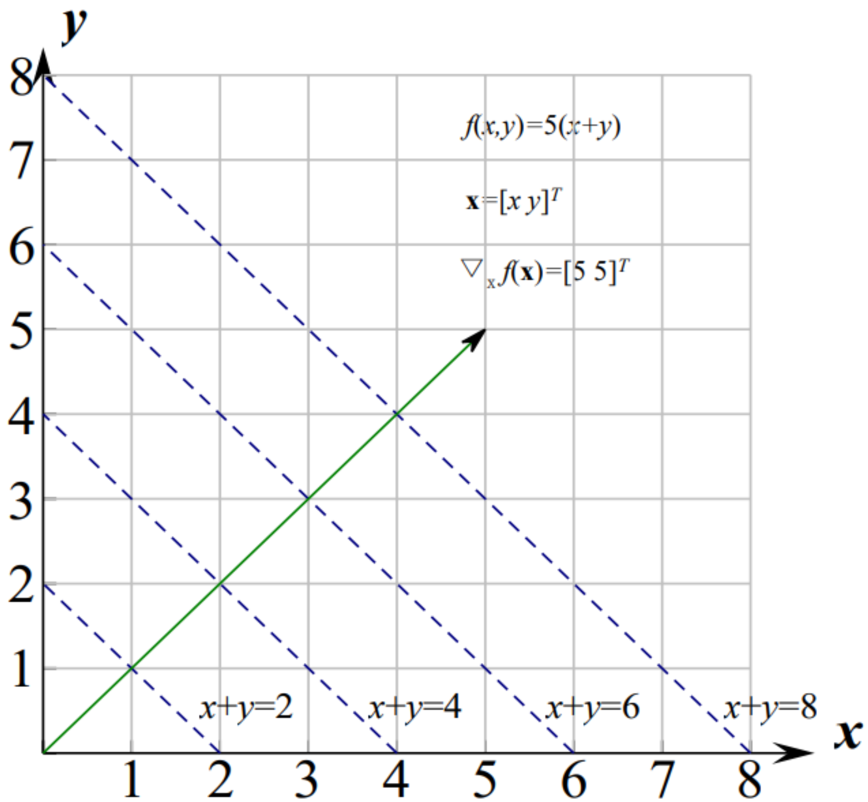
矩阵微积分

➢ **例1：** 函数 $f(x,y) = 5(x+y)$ 的梯度



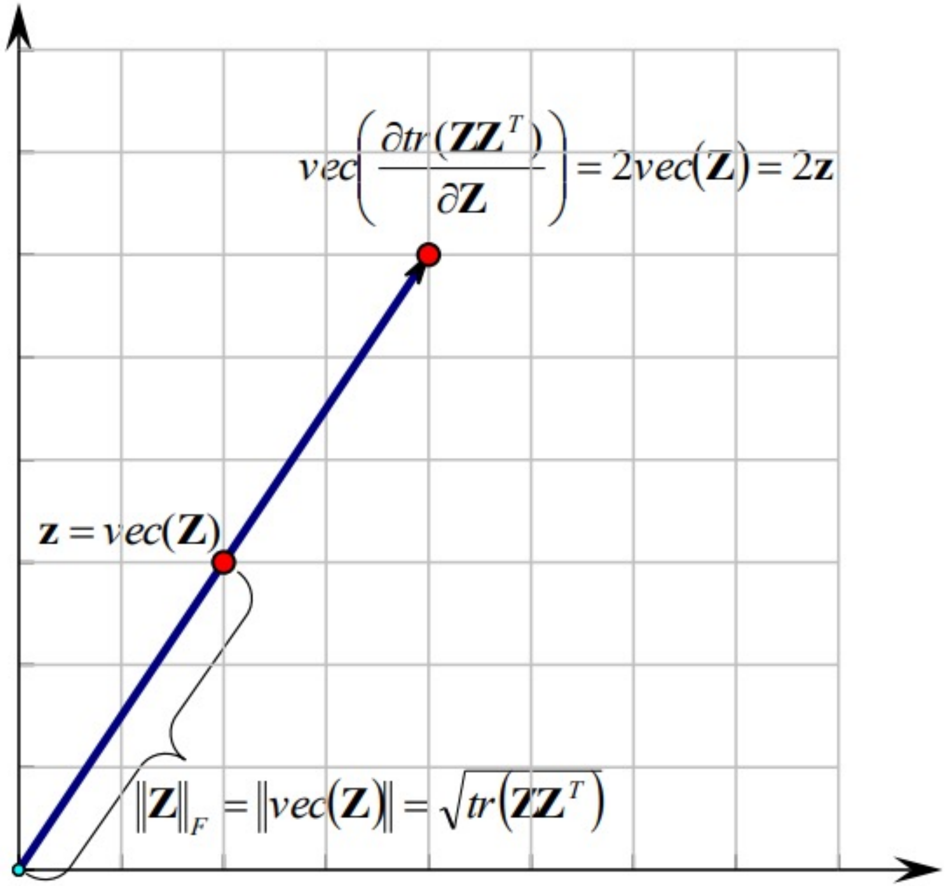
矩阵微积分

➢ **例2：** 函数 $f(x,y) = 0.5(x^2 + y^2)$ 的梯度



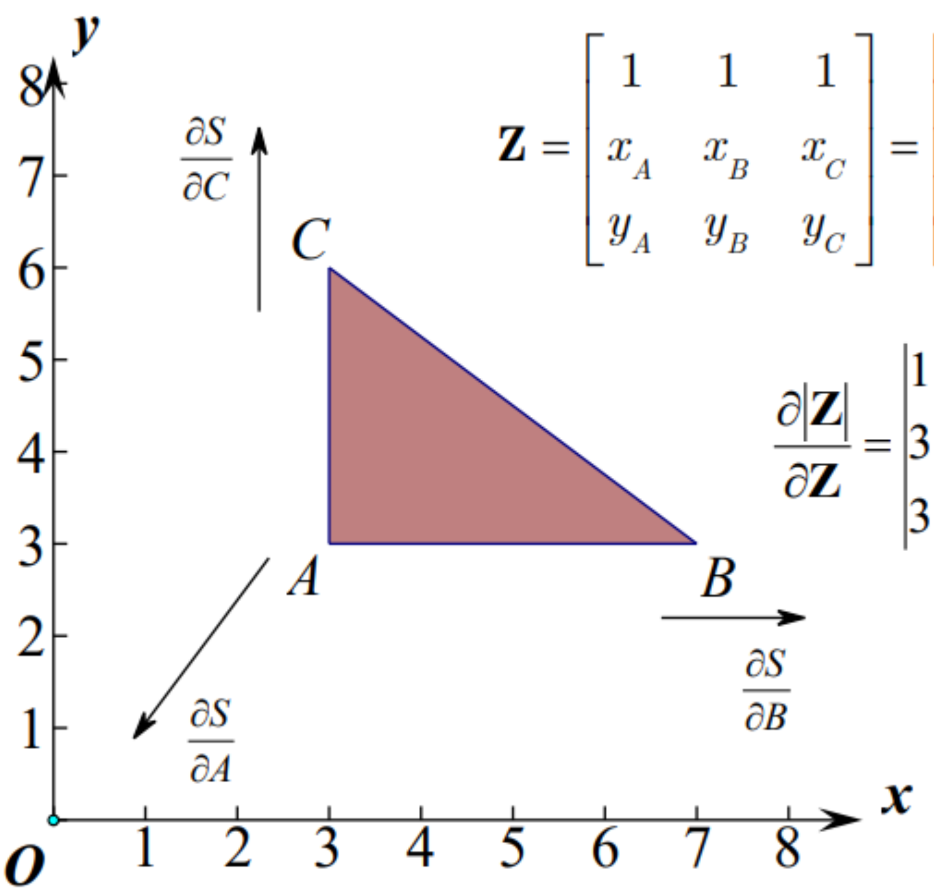
矩阵微积分

- 例3：迹函数相对于矩阵的梯度 $\frac{\partial\left(tr\left(\mathbf{Z}\mathbf{Z}^T\right)\right)}{\partial\mathbf{Z}}=\frac{\partial\left(tr\left(\mathbf{Z}^T\mathbf{Z}\right)\right)}{\partial\mathbf{Z}}=2\mathbf{Z}$



矩阵微积分

➢ **例4：** 行列式相对于矩阵的梯度 $\frac{\partial|\mathbf{Z}|}{\partial\mathbf{Z}}=|\mathbf{Z}|\left(\mathbf{Z}^{-1}\right)^{T}$



$$\mathbf{Z}=\begin{bmatrix}1 & 1 & 1\\ x_A & x_B & x_C\\ y_A & y_B & y_C\end{bmatrix}=\begin{bmatrix}1 & 1 & 1\\ 3 & 7 & 3\\ 3 & 3 & 6\end{bmatrix}$$

$$S=\frac{1}{2}|\mathbf{Z}|=\frac{1}{2}\left\|\begin{bmatrix}1 & 1 & 1\\ 3 & 7 & 3\\ 3 & 3 & 6\end{bmatrix}\right\|$$

$$\frac{\partial|\mathbf{Z}|}{\partial\mathbf{Z}}=\begin{vmatrix}1 & 1 & 1\\ 3 & 7 & 3\\ 3 & 3 & 6\end{vmatrix}\left(\begin{bmatrix}1 & 1 & 1\\ 3 & 7 & 3\\ 3 & 3 & 6\end{bmatrix}^{-1}\right)^{T}=\begin{bmatrix}33 & -9 & -12\\ -3 & 3 & 0\\ -4 & 0 & 4\end{bmatrix}$$

$$\frac{\partial S}{\partial\mathbf{Z}}=\begin{bmatrix}0 & 0 & 0\\ \frac{\partial S}{\partial x_A} & \frac{\partial S}{\partial x_B} & \frac{\partial S}{\partial x_C}\\ \frac{\partial S}{\partial y_A} & \frac{\partial S}{\partial y_B} & \frac{\partial S}{\partial y_C}\end{bmatrix}=\frac{1}{2}\begin{bmatrix}0 & 0 & 0\\ -3 & 3 & 0\\ -4 & 0 & 4\end{bmatrix}$$

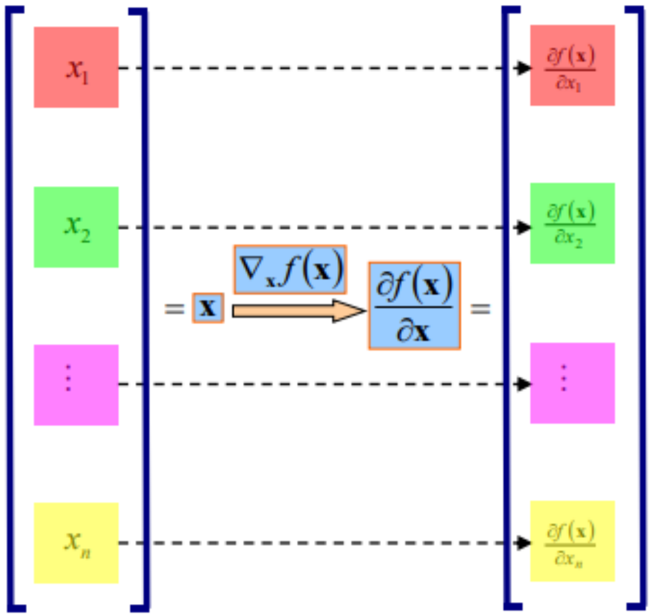
矩阵微积分

- 实值标量函数对于实向量的梯度

$$\nabla_{\mathbf{x}} f(\mathbf{x}) = \left[\frac{\partial f(\mathbf{x})}{\partial x_1}, \frac{\partial f(\mathbf{x})}{\partial x_2}, \cdots, \frac{\partial f(\mathbf{x})}{\partial x_n}\right]^T = \frac{\partial f(\mathbf{x})}{\partial \mathbf{x}}$$

- 以列向量为自变量的标量函数，其对于自变量的梯度仍然为一阶数相同的列向量

- 梯度的每个分量代表着函数在该分量方向上的变化率



矩阵微积分

- 实值向量函数对于实向量的梯度

$$\mathbf{f}(\mathbf{x})=[f_1(\mathbf{x}),f_2(\mathbf{x}),\cdots,f_m(\mathbf{x})]$$

$$\frac{\partial \mathbf{f}(\mathbf{x})}{\partial \mathbf{x}}=\left[\frac{\partial f_1(\mathbf{x})}{\partial \mathbf{x}},\frac{\partial f_2(\mathbf{x})}{\partial \mathbf{x}},\cdots,\frac{\partial f_m(\mathbf{x})}{\partial \mathbf{x}}\right]=\begin{bmatrix}\frac{\partial f_1(\mathbf{x})}{\partial x_1} & \frac{\partial f_2(\mathbf{x})}{\partial x_1} & \cdots & \frac{\partial f_m(\mathbf{x})}{\partial x_1}\\ \frac{\partial f_1(\mathbf{x})}{\partial x_2} & \frac{\partial f_2(\mathbf{x})}{\partial x_2} & \cdots & \frac{\partial f_m(\mathbf{x})}{\partial x_2}\\ \vdots & \vdots & \ddots & \vdots\\ \frac{\partial f_1(\mathbf{x})}{\partial x_n} & \frac{\partial f_2(\mathbf{x})}{\partial x_n} & \cdots & \frac{\partial f_m(\mathbf{x})}{\partial x_n}\end{bmatrix}=\nabla_{\mathbf{x}}\mathbf{f}(\mathbf{x})$$

- 向量函数对于向量的求导，相当于向量函数中的每一个分量函数对向量求导
- 行向量函数对列向量自变量求导形成矩阵；列向量函数对行向量自变量求导也可以形成矩阵

矩阵微积分

➢ 例： $\mathbf{f}(\mathbf{x})=\mathbf{x}=[x_1,x_2,\cdots,x_n]^T$

$$\frac{\partial \mathbf{f}(\mathbf{x})}{\partial \mathbf{x}^T}=\frac{\partial \mathbf{x}}{\partial \mathbf{x}^T}=\mathbf{I}_{n\times n}$$

$$\frac{\partial (\mathbf{f}(\mathbf{x}))^T}{\partial \mathbf{x}}=\frac{\partial \mathbf{x}^T}{\partial \mathbf{x}}=\mathbf{I}_{n\times n}$$

$$\frac{\partial \mathbf{f}(\mathbf{x})}{\partial \mathbf{x}}=\frac{\partial \mathbf{x}}{\partial \mathbf{x}}=\left(vec\left(\mathbf{I}_{n\times n}\right)\right)$$

$$\frac{\partial \left(\mathbf{f}(\mathbf{x})\right)^T}{\partial \mathbf{x}^T}=\frac{\partial \mathbf{x}^T}{\partial \mathbf{x}^T}=vec\left(\mathbf{I}_{n\times n}\right)^T$$

矩阵微积分

➢ 例： $\boldsymbol{f(x) = Ax}$

$$\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x} = \begin{bmatrix} \mathbf{A}(1,:)\mathbf{x} \\ \mathbf{A}(2,:)\mathbf{x} \\ \vdots \\ \mathbf{A}(n,:)\mathbf{x} \end{bmatrix}$$

$$\frac{\partial \mathbf{f}(\mathbf{x})}{\partial \mathbf{x}^T} = \frac{\partial (\mathbf{A}\mathbf{x})}{\partial \mathbf{x}^T} = \mathbf{A}$$

矩阵微积分

- 例：$f(x) = x^T A x$

$$f(\mathbf{x}) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} x_i x_j$$

与 x_k 相关的项 $\sum_{j=1}^{n} a_{kj} x_k x_j + \sum_{i=1}^{n} a_{ik} x_i x_k$

$$\frac{\partial f(\mathbf{x})}{\partial x_k} = \sum_{j=1}^{n} a_{kj} x_j + \sum_{i=1}^{n} a_{ik} x_i = \mathbf{A}(k,:)\mathbf{x} + \mathbf{A}(:,k)^T \mathbf{x} = \left(\mathbf{A}(k,:) + \mathbf{A}(:,k)^T\right)\mathbf{x}$$

$$\frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{A}^T \mathbf{x}$$

矩阵微积分

- 常用梯度公式及求导法则（针对向量）

  - 线性法则 $$\frac{\partial(c_1 f(\mathbf{x})+c_2 g(\mathbf{x}))}{\partial \mathbf{x}}=c_1\frac{\partial f(\mathbf{x})}{\partial \mathbf{x}}+c_2\frac{\partial g(\mathbf{x})}{\partial \mathbf{x}}$$

  - 乘积法则 $$\frac{\partial f(\mathbf{x})g(\mathbf{x})}{\partial \mathbf{x}}=g(\mathbf{x})\frac{\partial f(\mathbf{x})}{\partial \mathbf{x}}+f(\mathbf{x})\frac{\partial g(\mathbf{x})}{\partial \mathbf{x}}$$

  - 商法则 $$\frac{\partial(f(\mathbf{x})/g(\mathbf{x}))}{\partial \mathbf{x}}=\frac{1}{g^2(\mathbf{x})}\left(g(\mathbf{x})\frac{\partial f(\mathbf{x})}{\partial \mathbf{x}}-f(\mathbf{x})\frac{\partial g(\mathbf{x})}{\partial \mathbf{x}}\right)$$

  - 链式法则 $$\frac{\partial f(\mathbf{g}(\mathbf{x}))}{\partial \mathbf{x}}=\frac{\partial \mathbf{g}^T(\mathbf{x})}{\partial \mathbf{x}}\frac{\partial f(\mathbf{g})}{\partial \mathbf{g}}$$

矩阵微积分

➢ 常用梯度公式及求导法则

函数 $f(\mathbf{x})=c$

$$\frac{\partial c}{\partial \mathbf{x}}=\begin{bmatrix}0\\0\\\vdots\\0\end{bmatrix}=\mathbf{0}$$

函数 $f(\mathbf{x})=\mathbf{a}^T\mathbf{x}$

$$\frac{\partial \mathbf{a}^T\mathbf{x}}{\partial \mathbf{x}}=\frac{\partial \mathbf{x}^T\mathbf{a}}{\partial \mathbf{x}}=\mathbf{a}$$

函数 $f(\mathbf{x})=\mathbf{x}^T\mathbf{A}\mathbf{y}$

$$\frac{\partial \mathbf{x}^T\mathbf{A}\mathbf{y}}{\partial \mathbf{x}}=\mathbf{A}\mathbf{y}$$

矩阵微积分

- 实值函数$f(A)$相对于其自变量$m\times n$矩阵A的梯度定义为

$$\frac{\partial f(\mathbf{A})}{\partial \mathbf{A}}=\begin{bmatrix} \frac{\partial f(\mathbf{A})}{\partial a_{11}} & \frac{\partial f(\mathbf{A})}{\partial a_{12}} & \cdots & \frac{\partial f(\mathbf{A})}{\partial a_{1n}} \\ \frac{\partial f(\mathbf{A})}{\partial a_{21}} & \frac{\partial f(\mathbf{A})}{\partial a_{22}} & \cdots & \frac{\partial f(\mathbf{A})}{\partial a_{2n}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f(\mathbf{A})}{\partial a_{m1}} & \frac{\partial f(\mathbf{A})}{\partial a_{m2}} & \cdots & \frac{\partial f(\mathbf{A})}{\partial a_{mn}} \end{bmatrix}=\nabla_{\mathbf{A}} f(\mathbf{A})$$

- 实值函数相对于矩阵的梯度仍然为原矩阵同阶的矩阵
- 实值函数相对于矩阵的梯度矩阵的每一个分量对应于该函数在矩阵的每一个分量的变化率

矩阵微积分

➢ 例：求$f(A) = x^T A x$ 相对于矩阵A的梯度

$$f(\mathbf{A}) = \mathbf{x}^T \mathbf{A} \mathbf{y} = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} x_i y_j$$

$$\frac{\partial f(\mathbf{A})}{\partial a_{ij}} = x_i y_j$$

$$\frac{\partial f(\mathbf{A})}{\partial \mathbf{A}} = \begin{bmatrix} x_1 y_1 & x_1 y_2 & \cdots & x_1 y_n \\ x_2 y_1 & x_2 y_2 & \cdots & x_2 y_n \\ \vdots & \vdots & \ddots & \vdots \\ x_m y_1 & x_m y_2 & \cdots & x_m y_n \end{bmatrix} = \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix} \begin{bmatrix} y_1 & \cdots & y_n \end{bmatrix} = \mathbf{x}\mathbf{y}^T$$

矩阵微积分

➢ 例：求$f(A) = x^T A x$ 相对于矩阵A的梯度

$$\boxed{f(\mathbf{A}) = \mathbf{x}^T \mathbf{A} \mathbf{y} = vec(\mathbf{A})^T kron(\mathbf{y}, \mathbf{x})}$$

$$\frac{\partial f(vec(\mathbf{A}))}{\partial vec(\mathbf{A})} = kron(\mathbf{y}, \mathbf{x})$$

$$\nabla_{\mathbf{A}} f(\mathbf{A}) = unvec\left(\nabla_{vec(\mathbf{A})} f(vec(\mathbf{A}))\right) = unvec(kron(\mathbf{y}, \mathbf{x})) = \mathbf{x}\mathbf{y}^T$$

$$\boxed{unvec(kron(\mathbf{y}, \mathbf{x})) = \mathbf{x}\mathbf{y}^T}$$

矩阵微积分

- 常用梯度公式及求导法则（针对矩阵）

  - 线性法则 $$\frac{\partial(c_1 f(\mathbf{A})+c_2 g(\mathbf{A}))}{\partial \mathbf{A}}=c_1\frac{\partial f(\mathbf{A})}{\partial \mathbf{A}}+c_2\frac{\partial g(\mathbf{A})}{\partial \mathbf{A}}$$

  - 乘积法则 $$\frac{\partial f(\mathbf{A})g(\mathbf{A})}{\partial \mathbf{A}}=g(\mathbf{A})\frac{\partial f(\mathbf{A})}{\partial \mathbf{A}}+f(\mathbf{A})\frac{\partial g(\mathbf{A})}{\partial \mathbf{A}}$$

  - 商法则 $$\frac{\partial(f(\mathbf{A})/g(\mathbf{A}))}{\partial \mathbf{A}}=\frac{1}{g^2(\mathbf{A})}\left(g(\mathbf{A})\frac{\partial f(\mathbf{A})}{\partial \mathbf{A}}-f(\mathbf{A})\frac{\partial g(\mathbf{A})}{\partial \mathbf{A}}\right)$$

  - 链式法则 $$\frac{\partial g(f(\mathbf{A}))}{\partial \mathbf{A}}=\frac{dg(y)}{dy}\frac{\partial f(\mathbf{A})}{\partial \mathbf{A}}$$

矩阵微分定义

- 对于一个以向量$x=[x_1,x_2,\dots,x_n]^T$为变量的实值函数 $f(x)$，其微分公式定义如下：

$$df(\mathbf{x})=\sum_{i=1}^{n}\frac{\partial f(\mathbf{x})}{\partial x_i}dx_i$$

- 对于一个以$m\times n$阶矩阵X为变量的实值函数$f(x)$，其微分公式定义如下：

$$df(\mathbf{X})=\sum_{i=1}^{m}\sum_{j=1}^{n}\frac{\partial f(\mathbf{X})}{\partial x_{ij}}dx_{ij}$$

矩阵微分定义

➢ 重要公式

$$\boxed{d\mathbf{X}=\begin{bmatrix} dx_{11} & dx_{12} & \cdots & dx_{1n} \\ dx_{21} & dx_{22} & \cdots & dx_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ dx_{m1} & dx_{m2} & \cdots & dx_{mn} \end{bmatrix}=\sum_{i=1}^{m}\sum_{j=1}^{n}\mathbf{e}_i\hat{\mathbf{e}}_j^T dx_{ij}}$$

$$\frac{\partial \mathbf{X}}{\partial x_{ij}}=\mathbf{e}_i\hat{\mathbf{e}}_j^T$$

$$\boxed{df(\mathbf{X})=tr\left(\left(\frac{\partial f(\mathbf{X})}{\partial \mathbf{X}}\right)^T d\mathbf{X}\right)}$$

矩阵微分性质

- 线性法则

$$d(\alpha\mathbf{X})=\alpha d\mathbf{X} \quad d(\mathbf{X}+\mathbf{Y})=d\mathbf{X}+d\mathbf{Y}$$

- 乘法法则

$$d(\mathbf{X}\mathbf{Y})=(d\mathbf{X})\mathbf{Y}+\mathbf{X}(d\mathbf{Y})$$

- 矩阵的逆的微分（要证明）

$$d\left(\mathbf{X}^{-1}\right)=-\mathbf{X}^{-1}(d\mathbf{X})\mathbf{X}^{-1}$$

- 矩阵微分算子和迹算子的可交换性（要证明）

$$d(tr(\mathbf{X}))=tr(d(\mathbf{X}))=\sum_{i=1}^{n}dx_{ii}$$

矩阵微分性质

➢ 例：求 $tr(AXB)$ 相对于矩阵 X 的梯度。其中 A, X, B 分别是$p\times m, m\times n, n\times p$矩阵

方法1：

$$d(tr(\mathbf{AXB})) = tr\left(\mathbf{A}\left(\sum_{i=1}^{m}\sum_{j=1}^{n}\mathbf{e}_i\hat{\mathbf{e}}_j^T dx_{ij}\right)\mathbf{B}\right) = \sum_{i=1}^{m}\sum_{j=1}^{n} tr\left(\mathbf{A}\mathbf{e}_i\hat{\mathbf{e}}_j^T\mathbf{B}\right)dx_{ij}$$

$$= \sum_{i=1}^{m}\sum_{j=1}^{n}\hat{\mathbf{e}}_j^T\mathbf{BA}\mathbf{e}_i dx_{ij} = \sum_{i=1}^{m}\sum_{j=1}^{n}(\mathbf{BA})_{ji}dx_{ij} = \sum_{i=1}^{m}\sum_{j=1}^{n}\left(\mathbf{A}^T\mathbf{B}^T\right)_{ij}dx_{ij}$$

因此，
$$\frac{\partial tr(\mathbf{AXB})}{\partial x_{ij}} = \left(\mathbf{A}^T\mathbf{B}^T\right)_{ij}$$

即，
$$\frac{\partial tr(\mathbf{AXB})}{\partial \mathbf{X}} = \mathbf{A}^T\mathbf{B}^T$$

方法2：

$$d(tr(\mathbf{AXB})) = tr(\mathbf{A}(d\mathbf{X})\mathbf{B}) = tr(\mathbf{BA}(d\mathbf{X}))$$

因此，

$$\frac{\partial tr(\mathbf{AXB})}{\partial \mathbf{X}} = (\mathbf{BA})^T = \mathbf{A}^T\mathbf{B}^T$$

矩阵微分性质

- 例：求 $tr(AX^{-1}B)$ 相对于矩阵 X 的梯度。其中 A, X, B 分别是$p\times m, m\times n, n\times p$矩阵

方法1：$d\left(tr\left(\mathbf{AX}^{-1}\mathbf{B}\right)\right) = tr\left(d\left(\mathbf{AX}^{-1}\mathbf{B}\right)\right) = tr\left(\mathbf{A}\left(d\mathbf{X}^{-1}\right)\mathbf{B}\right)$

$$= tr\left(-\mathbf{A}\left(\sum_{i=1}^{n}\sum_{j=1}^{n}\mathbf{X}^{-1}\mathbf{e}_i\mathbf{e}_j^T\mathbf{X}^{-1}dx_{ij}\right)\mathbf{B}\right) = -\sum_{i=1}^{n}\sum_{j=1}^{n}tr\left(\mathbf{AX}^{-1}\mathbf{e}_i\mathbf{e}_j^T\mathbf{X}^{-1}\mathbf{B}\right)dx_{ij}$$

$$= -\sum_{i=1}^{n}\sum_{j=1}^{n}\mathbf{e}_j^T\mathbf{X}^{-1}\mathbf{BAX}^{-1}\mathbf{e}_i dx_{ij} = -\sum_{i=1}^{n}\sum_{j=1}^{n}\left(\mathbf{X}^{-1}\mathbf{BAX}^{-1}\right)_{ji}dx_{ij}$$

$$= -\sum_{i=1}^{n}\sum_{j=1}^{n}\left(\mathbf{X}^{-T}\mathbf{A}^T\mathbf{B}^T\mathbf{X}^{-T}\right)_{ij}dx_{ij}$$

因此，$\dfrac{\partial tr\left(\mathbf{AX}^{-1}\mathbf{B}\right)}{\partial x_{ij}} = -\left(\mathbf{X}^{-T}\mathbf{A}^T\mathbf{B}^T\mathbf{X}^{-T}\right)_{ij}$

即，$\dfrac{\partial tr\left(\mathbf{AX}^{-1}\mathbf{B}\right)}{\partial \mathbf{X}} = -\mathbf{X}^{-T}\mathbf{A}^T\mathbf{B}^T\mathbf{X}^{-T}$

方法2：

$$d\left(tr\left(\mathbf{AX}^{-1}\mathbf{B}\right)\right) = tr\left(\mathbf{A}\left(d\mathbf{X}^{-1}\right)\mathbf{B}\right)$$

$$= -tr\left(\mathbf{AX}^{-1}\left(d\mathbf{X}\right)\mathbf{X}^{-1}\mathbf{B}\right) = -tr\left(\mathbf{X}^{-1}\mathbf{BAX}^{-1}d\mathbf{X}\right)$$

因此，$\dfrac{\partial tr\left(\mathbf{AX}^{-1}\mathbf{B}\right)}{\partial \mathbf{X}} = -\left(\mathbf{X}^{-1}\mathbf{BAX}^{-1}\right)^T = -\mathbf{X}^{-T}\mathbf{A}^T\mathbf{B}^T\mathbf{X}^{-T}$

矩阵微分性质

➢ 例：求 $tr(X^TX)$ 相对于矩阵 X 的梯度。

方法1：$d\left(tr\left(\mathbf{X}^T\mathbf{X}\right)\right)=tr\left(d\left(\mathbf{X}^T\mathbf{X}\right)\right)=tr\left(\left(d\mathbf{X}^T\right)\mathbf{X}+\mathbf{X}^Td\mathbf{X}\right)$

$$=tr\left(\left(\sum_{i=1}^{m}\sum_{j=1}^{n}\hat{\mathbf{e}}_j\mathbf{e}_i^Tdx_{ij}\right)\mathbf{X}+\mathbf{X}^T\left(\sum_{i=1}^{m}\sum_{j=1}^{n}\mathbf{e}_i\hat{\mathbf{e}}_j^Tdx_{ij}\right)\right)$$

$$=\sum_{i=1}^{m}\sum_{j=1}^{n}tr\left(\hat{\mathbf{e}}_j\mathbf{e}_i^T\mathbf{X}\right)dx_{ij}+\sum_{i=1}^{m}\sum_{j=1}^{n}tr\left(\mathbf{X}^T\mathbf{e}_i\hat{\mathbf{e}}_j^T\right)dx_{ij}$$

$$=\sum_{i=1}^{m}\sum_{j=1}^{n}\mathbf{e}_i^T\mathbf{X}\hat{\mathbf{e}}_jdx_{ij}+\sum_{i=1}^{m}\sum_{j=1}^{n}\hat{\mathbf{e}}_j^T\mathbf{X}^T\mathbf{e}_idx_{ij}$$

$$=2\sum_{i=1}^{m}\sum_{j=1}^{n}\mathbf{e}_i^T\mathbf{X}\hat{\mathbf{e}}_jdx_{ij}=2\sum_{i=1}^{m}\sum_{j=1}^{n}(\mathbf{X})_{ij}dx_{ij}$$

因此，$\dfrac{\partial tr\left(\mathbf{X}^T\mathbf{X}\right)}{\partial x_{ij}}=2(\mathbf{X})_{ij}$. 即 $\dfrac{\partial tr\left(\mathbf{X}^T\mathbf{X}\right)}{\partial \mathbf{X}}=\dfrac{\partial tr\left(\mathbf{X}\mathbf{X}^T\right)}{\partial \mathbf{X}}=2\mathbf{X}$

方法2：

$$d\left(tr\left(\mathbf{X}^T\mathbf{X}\right)\right)=tr\left(d\left(\mathbf{X}^T\mathbf{X}\right)\right)=tr\left(\left(d\mathbf{X}^T\right)\mathbf{X}+\mathbf{X}^Td\mathbf{X}\right)=tr\left(\left(d\mathbf{X}^T\right)\mathbf{X}\right)+tr\left(\mathbf{X}^Td\mathbf{X}\right)$$

而，$tr\left(\left(d\mathbf{X}^T\right)\mathbf{X}\right)=tr\left(\left(d\mathbf{X}^T\right)\mathbf{X}\right)^T=tr\left(\mathbf{X}^Td\mathbf{X}\right)$

因此，$d\left(tr\left(\mathbf{X}^T\mathbf{X}\right)\right)=2tr\left(\mathbf{X}^Td\mathbf{X}\right)$

所以，$\dfrac{\partial tr\left(\mathbf{X}^T\mathbf{X}\right)}{\partial \mathbf{X}}=\dfrac{\partial tr\left(\mathbf{X}\mathbf{X}^T\right)}{\partial \mathbf{X}}=2\left(\mathbf{X}^T\right)^T=2\mathbf{X}$

矩阵微分性质

➢ 关于 $tr(*)$ 的常用公式。

$$1.\ \frac{\partial tr(\mathbf{X})}{\partial \mathbf{X}} = \mathbf{I}_{n\times n}$$

$$2.\ \frac{\partial tr(\mathbf{X}^{-1})}{\partial \mathbf{X}} = -(\mathbf{X}^{-2})^T$$

$$3.\ \frac{\partial tr(\mathbf{XA})}{\partial \mathbf{X}} = \frac{\partial tr(\mathbf{AX})}{\partial \mathbf{X}} = \mathbf{A}^T$$

$$4.\ \frac{\partial tr(\mathbf{X}^T\mathbf{A})}{\partial \mathbf{X}} = \frac{\partial tr(\mathbf{AX}^T)}{\partial \mathbf{X}} = \mathbf{A}$$

$$5.\ \frac{\partial tr(\mathbf{X}^2)}{\partial \mathbf{X}} = 2\mathbf{X}^T$$

$$6.\ \frac{\partial tr(\mathbf{AX}^{-1})}{\partial \mathbf{X}} = -(\mathbf{X}^{-1}\mathbf{AX}^{-1})^T$$

$$7.\ \frac{\partial tr(\mathbf{X}^T\mathbf{AX})}{\partial \mathbf{X}} = (\mathbf{A}+\mathbf{A}^T)\mathbf{X}$$

$$8.\ \frac{\partial tr(\mathbf{XAX}^T)}{\partial \mathbf{X}} = \mathbf{X}(\mathbf{A}+\mathbf{A}^T)$$

$$9.\ \frac{\partial tr(\mathbf{AXX}^T\mathbf{A}^T)}{\partial \mathbf{X}} = 2\mathbf{A}^T\mathbf{AX}$$

$$10.\ \frac{\partial tr(\mathbf{AXX}^T\mathbf{B})}{\partial \mathbf{X}} = (\mathbf{BA}+\mathbf{A}^T\mathbf{B}^T)\mathbf{X}$$

矩阵微分性质

- 关于行列式的梯度
  - 对于方阵*X*，存在

$$\frac{\partial|\mathbf{X}|}{\partial x_{ij}} = C_{ij} \quad 和 \quad \frac{\partial|\mathbf{X}|}{\partial \mathbf{X}} = \mathbf{C} = \left(\mathbf{X}^*\right)^T = |\mathbf{X}|\mathbf{X}^{-T}$$

其微分形式为：

$$d|\mathbf{X}| = tr\left(|\mathbf{X}|\mathbf{X}^{-1}d\mathbf{X}\right)$$

矩阵微分性质

- 例：$|\boldsymbol{AXB}|$关于$\boldsymbol{X}$的梯度

$$
\begin{aligned}
d|\mathbf{AXB}| &= tr\left(\left(|\mathbf{AXB}|(\mathbf{AXB})^{-1}\right)d(\mathbf{AXB})\right)\\
&= tr\left(\left(|\mathbf{AXB}|(\mathbf{AXB})^{-1}\right)\mathbf{A}(d\mathbf{X})\mathbf{B}\right)\\
&= tr\left(\left(|\mathbf{AXB}|\mathbf{B}(\mathbf{AXB})^{-1}\right)\mathbf{A}(d\mathbf{X})\right)
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial|\mathbf{AXB}|}{\partial\mathbf{X}} &= \left(\left(|\mathbf{AXB}|\mathbf{B}\left(\mathbf{AXB}\right)^{-1}\right)\mathbf{A}\right)^{T}\\
&= |\mathbf{AXB}|\mathbf{A}^{T}\left(\mathbf{B}^{T}\mathbf{X}^{T}\mathbf{A}^{T}\right)^{-1}\mathbf{B}^{T}
\end{aligned}
$$

矩阵微分性质

- 常用梯度的公式

$$1.\ \frac{\partial|\mathbf{X}|}{\partial\mathbf{X}}=|\mathbf{X}|\mathbf{X}^{-T}$$

$$2.\ \frac{\partial|\mathbf{X}^{-1}|}{\partial\mathbf{X}}=-\frac{\mathbf{X}^{-T}}{|\mathbf{X}|}$$

$$3.\ \frac{\partial\log|\mathbf{X}|}{\partial\mathbf{X}}=\mathbf{X}^{-T}$$

$$4.\ \frac{\partial|\mathbf{X}\mathbf{X}^{T}|}{\partial\mathbf{X}}=2|\mathbf{X}\mathbf{X}^{T}|\left(\mathbf{X}\mathbf{X}^{T}\right)^{-1}\mathbf{X}\qquad \operatorname{rank}\left(\mathbf{X}_{m\times n}\right)=m$$

$$5.\ \frac{\partial|\mathbf{X}^{T}\mathbf{X}|}{\partial\mathbf{X}}=2|\mathbf{X}^{T}\mathbf{X}|\mathbf{X}\left(\mathbf{X}^{T}\mathbf{X}\right)^{-1}\qquad \operatorname{rank}\left(\mathbf{X}_{m\times n}\right)=n$$

矩阵微分性质

- 常用梯度的公式

$$6.\ \frac{\partial\left|\mathbf{X}^2\right|}{\partial\mathbf{X}} = 2\left|\mathbf{X}\right|^2\mathbf{X}^{-T} \qquad \operatorname{rank}\left(\mathbf{X}_{m\times m}\right) = m$$

$$7.\ \frac{\partial\left|\mathbf{AXB}\right|}{\partial\mathbf{X}} = \left|\mathbf{AXB}\right|\mathbf{A}^T\left(\mathbf{B}^T\mathbf{X}^T\mathbf{A}^T\right)^{-1}\mathbf{B}^T$$

$$8.\ \frac{\partial\left|\mathbf{X}^T\mathbf{AX}\right|}{\partial\mathbf{X}} = \left|\mathbf{X}^T\mathbf{AX}\right|\left(\mathbf{AX}\left(\mathbf{X}^T\mathbf{AX}\right)^{-1}+\mathbf{A}^T\mathbf{X}\left(\mathbf{X}^T\mathbf{A}^T\mathbf{X}\right)^{-1}\right) \qquad \mathbf{A}\text{不是对称矩阵}$$

$$9.\ \frac{\partial\left|\mathbf{X}^T\mathbf{AX}\right|}{\partial\mathbf{X}} = 2\left|\mathbf{X}^T\mathbf{AX}\right|\mathbf{AX}\left(\mathbf{X}^T\mathbf{AX}\right)^{-1} \qquad \mathbf{A}\text{是对称矩阵}$$

Hessian矩阵

➢ 实值函数$f(x)$对于向量的二阶偏导数

$$\frac{\partial^2 f(\mathbf{x})}{\partial \mathbf{x} \partial \mathbf{x}^T} = \frac{\partial}{\partial \mathbf{x}^T}\left(\frac{\partial f(\mathbf{x})}{\partial \mathbf{x}}\right) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_n} \end{bmatrix}$$

Hessian矩阵

- 实值函数$f(x)$的泰勒展开

方向导数

$$f(\mathbf{x}+\Delta\mathbf{x})=f(\mathbf{x})+\boxed{(\Delta\mathbf{x})^T\nabla_{\mathbf{x}}f(\mathbf{x})}+\boxed{\frac{1}{2}(\Delta\mathbf{x})^T\nabla_{\mathbf{x}}^2f(\mathbf{x})\Delta\mathbf{x}}+\cdots$$

- 实值函数$f(x)$的局部极值条件

$$\nabla_{\mathbf{x}}f(\mathbf{x}_*)=\mathbf{0},\quad \nabla_{\mathbf{x}}^2f(\mathbf{x}_*)\geq\mathbf{0}$$

半正定条件

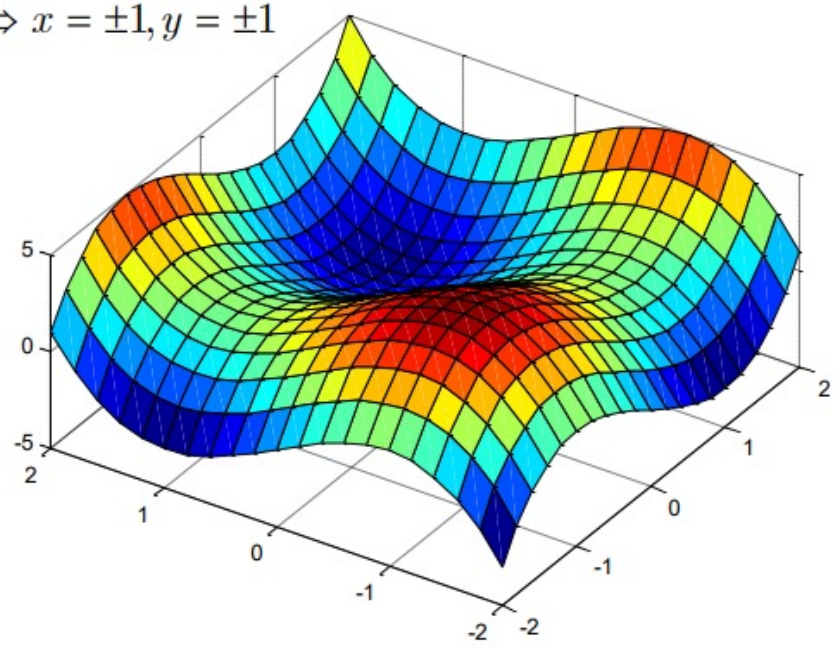
Hessian矩阵

➢ 实值函数$\boldsymbol{f}(\boldsymbol{x})$的局部极值

$$f\left(x,y\right)=x^3+y^3-3x-3y+1$$

$$\nabla_{\mathbf{x}} f\left(\mathbf{x}\right)=\begin{bmatrix}3x^2-3\\3y^2-3\end{bmatrix}\overset{令}{=}\mathbf{0}\Rightarrow x=\pm1,y=\pm1$$

$$\nabla_{\mathbf{x}}^2 f\left(\mathbf{x}\right)=\begin{bmatrix}6x & 0\\0 & 6y\end{bmatrix}$$



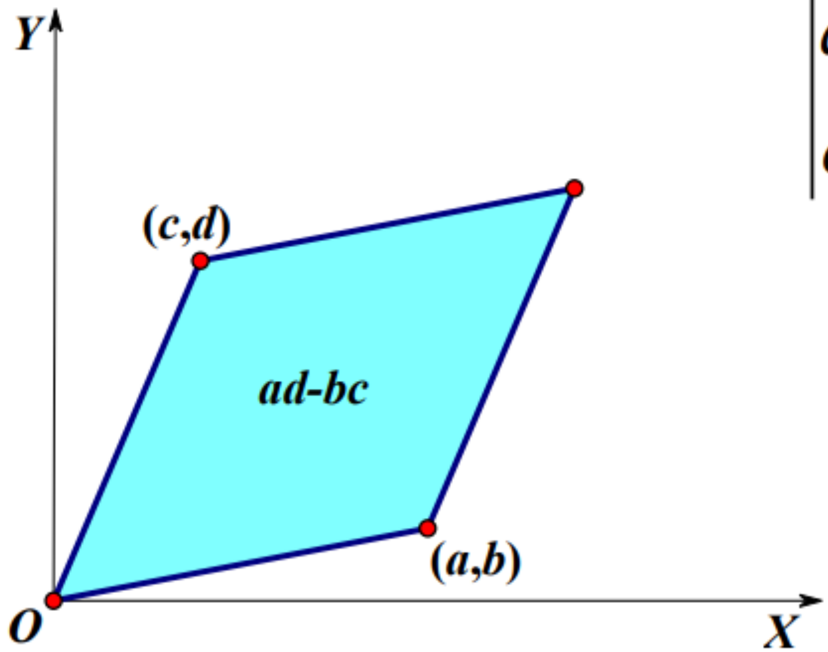
Hessian矩阵

- 实值函数$f(x)$的二阶偏导

$$1.\ \frac{\partial^2 \mathbf{a}^T\mathbf{x}}{\partial\mathbf{x}\partial\mathbf{x}^T} = \mathbf{O}_{n\times n}$$

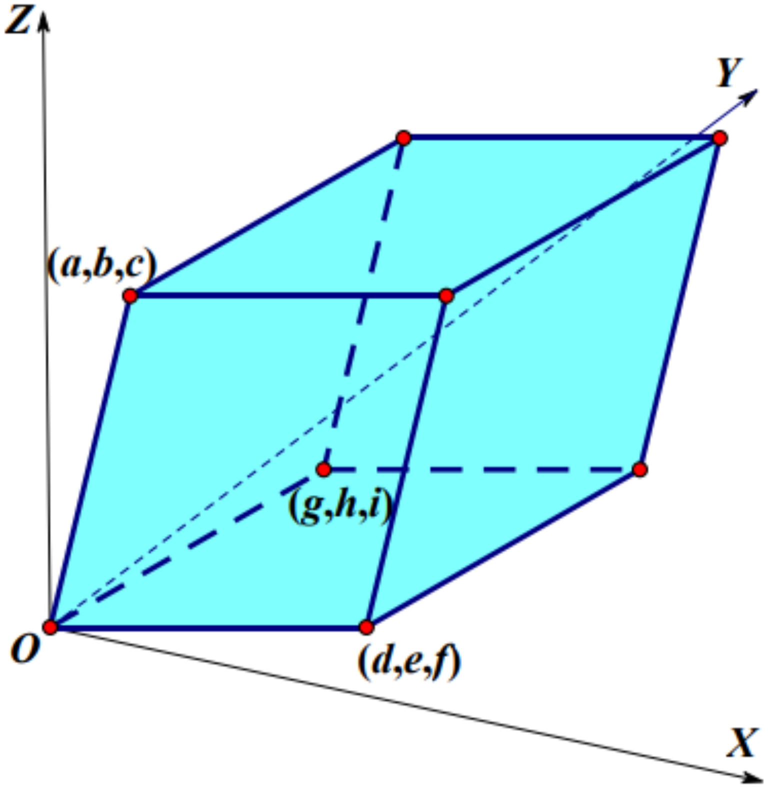
$$2.\ \frac{\partial^2 \mathbf{x}^T\mathbf{A}\mathbf{x}}{\partial\mathbf{x}\partial\mathbf{x}^T} = \mathbf{A} + \mathbf{A}^T$$

行列式几何意义



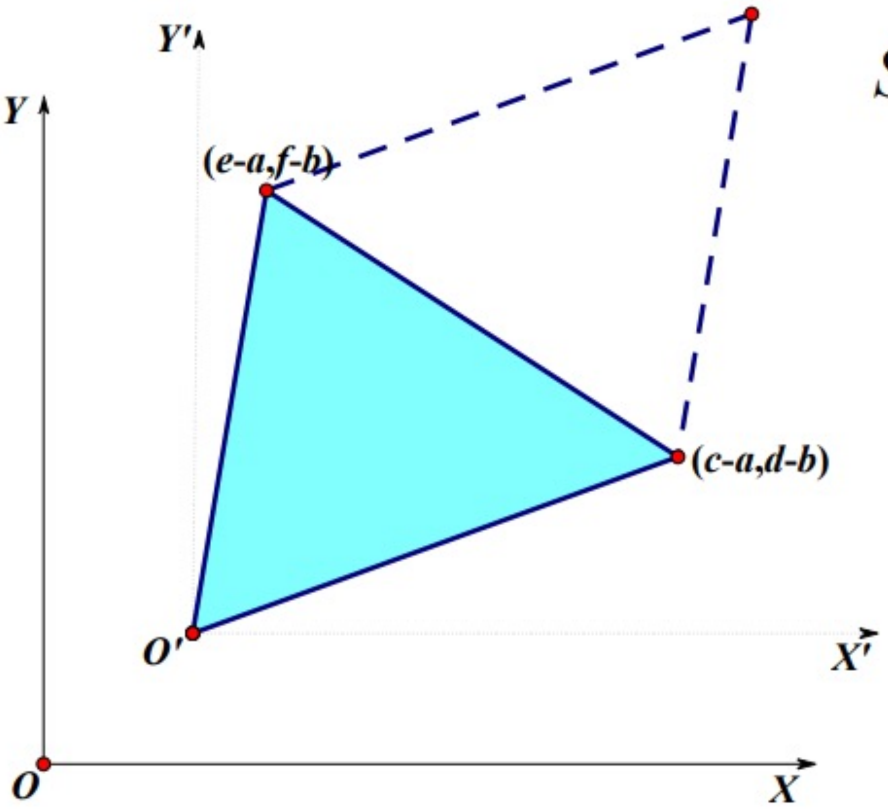
$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

行列式几何意义



$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = aei + bfg + cdh - ceg - bdi - afh$$

行列式几何意义



$$S_{\Delta} = \frac{1}{2} abs\left(\begin{vmatrix} c-a & e-a \\ d-b & f-b \end{vmatrix} \right)$$

$$S_{\Delta} = \frac{1}{2} abs\left(\begin{vmatrix} 1 & 1 & 1 \\ a & c & e \\ b & d & f \end{vmatrix} \right)$$

矩阵分析

- 向量范数
- 矩阵范数
- 序列极限
- 矩阵幂级数
- 矩阵函数
- 矩阵微积分
- 矩阵函数的应用

矩阵微积分应用

微分方程组的一般形式

$$X'(t)=A(t)X(t)+f(t)$$

$$X(t_0)=C_0$$

向量，包含t

齐次：f（t）=0

非齐次：f（t）≠0

常系数：A（t）=A

$$x'(t)=a(t)x(t)+f(t)$$

$$x'(t)=a(t)x(t)$$

$$\frac{dx(t)}{x(t)}=a(t)dt$$

$$\frac{dx(t)}{x(t)}=?$$

二、一阶线性常系数齐次微分方程组

求解： $X'(t)=AX(t)$

$\cancel{X(t_0)=C_0}$

定理5、11 ： 上述方程组的解为：

$$X(t)=e^{A(t-t_0)}x(t_0)$$

矩阵微积分应用

定理：上述方程组的解为：

$$X(t)=e^{A(t-t_0)}x(t_0)$$

$$\frac{d\left(e^{-At}x(t)\right)}{dt} = \underline{(-A)e^{-At}}\,x(t) + e^{-At}\frac{dx(t)}{dt}$$

$$= e^{-At}\cdot\left(-Ax(t)+\dot{x}(t)\right)$$

$$= 0$$

两边求积分 $[t_0, t]$

$$e^{-At}x(t) - e^{-At_0}x(t_0) = 0$$

$$\Rightarrow X(t) = e^{A(t-t_0)}x(t_0) = c\,e^{A(t-t_0)}$$

矩阵微积分应用

定理：上述方程组的解为：

$$X(t) = e^{A(t-t_0)} x(t_0)$$

$$\frac{d\left(e^{-At}x(t)\right)}{dt} = \underline{(-A)e^{-At}}\,x(t) + e^{-At}\frac{dx(t)}{dt}$$

$$= e^{-At}\cdot\left(-Ax(t) + \dot{x}(t)\right)$$

$$= 0$$

两边求积分 $[t_0, t]$

$$e^{-At}x(t) - e^{-At_0}x(t_0) = 0$$

$$\Rightarrow X(t) = e^{A(t-t_0)}x(t_0) = ce^{A(t-t_0)}$$

矩阵微积分应用

例题1　求解　$$\begin{cases} X'(t)=\begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix} X(t) \qquad X(0)=\begin{bmatrix} 1 \\ 2 \end{bmatrix} \end{cases}$$

矩阵微积分应用

一阶线性常系数非齐次微分方程组

求解： $X'(t) = A(t)X(t) + f(t)$

$X(t_0) = C$。

定理：上述方程组的解为：

$$X(t)=e^{A(t-t0)}x(t_0) + \int_{t_0}^{t} e^{A(t-s)} f(s)ds$$

矩阵微积分应用

定理：上述方程组的解为：

$$X(t)=e^{A(t-t0)}x(t_0)+\int_{t_0}^{t}e^{A(t-s)}f(s)ds$$

$$\frac{d\left(e^{-At}\cdot x(t)\right)}{dt}=(-A)e^{-At}x(t)+e^{-At}\cdot x'(t)$$

$$=e^{-At}\cdot\left(\dot{x}(t)-Ax(t)\right)$$

$$=e^{-At}\cdot f(t)$$

$[t_0,t]$ 两端积分:

$$e^{-At}x(t)-e^{-At_0}x(t_0)=\int_{t_0}^{t}e^{-A\tau}f(\tau)\,d\tau$$

$\Rightarrow$ 两端 $\times\ e^{At}$ $\Rightarrow$

$$x(t)=e^{A(t-t_0)}x(t_0)+\int_{t_0}^{t}e^{A(t-\tau)}f(\tau)\,d\tau$$

矩阵微积分应用

例题2　求解

$$\begin{cases} X'(t) = \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix} X(t) + \begin{bmatrix} 1 \\ -1 \end{bmatrix} \\ X(0) = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \end{cases}$$

Thanks