矩阵分析与应用

第十讲　矩阵分析及其应用之二

信息与通信工程学院

吕旌阳

本讲主要内容

- 矩阵函数的一般定义
- 矩阵函数的性质
- 矩阵的微分和积分

三、矩阵函数的一般定义

展开式 $f(z)=\sum c_k z^k$， $\left(|z|<r, r>0\right)$ ，要求

（1） $f^{(k)}(0)$ 存在 $\left(k=0,1,2,\cdots\right)$

（2） $\lim\limits_{k\to\infty}\dfrac{f^{(k+1)}(\xi)}{(k+1)!}\xi^{k+1}=0 \quad \left(|z|<r\right)$

对于一元函数 $f(z)=\dfrac{1}{z}$ 等，还不能定义矩阵函数。

基于矩阵函数值的**Jordan**标准形算法，拓宽定义

矩阵函数的一般定义

设 $P^{-1}AP = J = \mathrm{diag}\left(J_1,\cdots,J_s\right)$，$J_i = \lambda_i I + I^{(1)}$

如果 $f(z)$ 在 λ_i 处有 $m_i - 1$ 阶导数，令

$$f\left(J_i\right) = \sum_{k=0}^{\infty} c_k J_i^k = f\left(\lambda_i\right) I + \frac{f'\left(\lambda_i\right)}{1!} I^{(1)} + \cdots + \frac{f^{(m_i-1)}\left(\lambda_i\right)}{(m_i-1)!} I^{(m_i-1)}$$

$$f\left(A\right) = \sum_{k=0}^{\infty} c_k A^k = P \bullet \sum_{k=0}^{\infty} c_k J^k \bullet P^{-1} = P \bullet diag\left(f(J_1),\cdots,f(J_s)\right) \bullet P^{-1}$$

称 $f(A)$ 为对应于 $f(z)$ 的矩阵函数

[注] ① 拓宽定义不要求$f(z)$能展为“z”的幂级数，但要求在A的特征值 λ_i （重数为m_i）处有 m_i-1 阶导数,后者较前者弱！

② 当能够展为“z”的幂级数时,矩阵函数的拓宽定义与级数原始定义是一致的.

例9：$A=\begin{bmatrix}2&1&&\\&2&1&\\&&2&1\\&&&2\end{bmatrix}, f(z)=\dfrac{1}{z}$，求 $f(A)$

解：$f(z)=\dfrac{1}{z}, f'(z)=-z^{-2}, f''(z)=2z^{-3}, f'''(z)=-6z^{-4}$

$$f(A)=f(J)$$

$$=f(2)\bullet I+f'(2)\bullet I^{(1)}+\frac{f''(2)}{2!}\bullet I^{(2)}+\frac{f'''(2)}{3!}\bullet I^{(3)}$$

$$=\begin{bmatrix}0.5&-0.25&0.125&-0.0625\\&0.5&-0.25&0.125\\&&0.5&-0.25\\&&&0.5\end{bmatrix}$$

例10： $A=\begin{bmatrix}1&1&0\\0&1&0\\0&0&2\end{bmatrix}$, $f(z)=\sqrt{z}$ ，求 $f(A)$

解： $f(z)=\sqrt{z}, f'(z)=\dfrac{1}{2\sqrt{z}}$

$$J_1=\begin{bmatrix}1&1\\&1\end{bmatrix}: \quad f(J_1)=f(1)\bullet I+f'(1)\bullet I^{(1)}=\begin{bmatrix}1&1/2\\&1\end{bmatrix}$$

$$J_2=[2]: \quad f(J_2)=f(2)\bullet I=\left[\sqrt{2}\right]$$

$$f(A)=f(J)=\begin{bmatrix}f(J_1)&\\&f(J_2)\end{bmatrix}=\begin{bmatrix}1&1/2&0\\&1&0\\&&\sqrt{2}\end{bmatrix}$$

四、矩阵函数的性质

级数定义或拓宽定义给出的矩阵函数具有下列性质：

(1) $f(z)=f_1(z)+f_2(z) \Rightarrow f(A)=f_1(A)+f_2(A)$

$$f^{(l)}(\lambda_i)=f_1^{(l)}(\lambda_i)+f_2^{(l)}(\lambda_i)$$

$$\Rightarrow f^{(l)}(J_i)=f_1^{(l)}(J_i)+f_2^{(l)}(J_i)$$

$$f(A)=P\bullet\left\{\begin{bmatrix} f_1(J_1) & & \\ & \ddots & \\ & & f_1(J_s)\end{bmatrix}+\begin{bmatrix} f_2(J_1) & & \\ & \ddots & \\ & & f_2(J_s)\end{bmatrix}\right\}\bullet P^{-1}$$

$$=f_1(A)+f_2(A)$$

(2) $f(z)=f_1(z)\bullet f_2(z)$

$$\Rightarrow f(A)=f_1(A)\bullet f_2(A)=f_2(A)\bullet f_1(A)$$

$$f_1(J_i)\bullet f_2(J_i)=\left[f_1\bullet I+f_1'\bullet I^{(1)}+\frac{f_1''}{2!}\bullet I^{(2)}+\cdots+\frac{f_1^{(m_i-1)}}{(m_i-1)!}\bullet I^{(m_i-1)}\right]\bullet$$

$$\left[f_2\bullet I+\frac{f_2'}{1!}\bullet I^{(1)}+\frac{f_2''}{2!}\bullet I^{(2)}+\cdots+\frac{f_2^{(m_i-1)}}{(m_i-1)!}\bullet I^{(m_i-1)}\right]$$

$$=(f_1f_2)\bullet I+\frac{f_1'f_2+f_1f_2'}{1!}\bullet I^{(1)}+\frac{f_1''f_2+2f_1'f_2'+f_1f_2''}{2!}\bullet I^{(2)}+\cdots$$

$$=(f_1f_2)\bullet I+\frac{(f_1f_2)'}{1!}\bullet I^{(1)}+\frac{(f_1f_2)''}{2!}\bullet I^{(2)}+\cdots+\frac{(f_1f_2)^{(m_i-1)}}{(m_i-1)!}\bullet I^{(m_i-1)}$$

$$=f(J_i)$$

$$f(A)=P\begin{bmatrix} f(J_1) & & \\ & \ddots & \\ & & f(J_s) \end{bmatrix}P^{-1}$$

$$=P\begin{bmatrix} f_1(J_1) & & \\ & \ddots & \\ & & f_1(J_s) \end{bmatrix}P^{-1}\bullet P\begin{bmatrix} f_2(J_1) & & \\ & \ddots & \\ & & f_2(J_s) \end{bmatrix}P^{-1}$$

$$= f_1(A)\bullet f_2(A)$$

4、矩阵的微分和积分

定义: 如果矩阵 $A(t)=\left(a_{ij}(t)\right)_{m\times n}$,的每一个元素 $a_{ij}(t)$ 是变量t 的可微函数，则$A(t)$关于t的导数(微商)定义为

$$\frac{dA(t)}{dt}=\left(a'_{ij}(t)\right)_{m\times n}，\quad 或者 \quad A'(t)=\left(a'_{ij}(t)\right)_{m\times n}$$

定理8：设 $A(t),B(t)$ 可导，则有

（1）$$\frac{d}{dt}\left[A(t)+B(t)\right]=\frac{d}{dt}A(t)+\frac{d}{dt}B(t)$$

（2）$A_{m\times n}, f(t)$ 可导 $$\frac{d}{dt}\left[f(t)A(t)\right]=f'(t)A(t)+f(t)A'(t)$$

（3）$A_{m\times n},B_{n\times l}:$ $$\frac{d}{dt}\left[A(t)B(t)\right]=A'(t)B(t)+A(t)B'(t)$$

证明：（3）左 $$=\frac{d}{dt}\left(\sum_k a_{ik}(t)b_{kj}(t)\right)_{m\times l}$$

$$=\left(\sum_k a'_{ik}(t)b_{kj}(t)+\sum_k a_{ik}(t)b'_{kj}(t)\right)_{m\times l}$$

$$=\left(\sum_k a'_{ik}(t)b_{kj}(t)\right)_{m\times l}+\left(\sum_k a_{ik}(t)b'_{kj}(t)\right)_{m\times l}$$ =右

定理9：设 $A_{n\times n}$ 为常数矩阵，则有

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

（2） $$\frac{d}{dt}\cos(tA)=-A\bullet\sin(tA)=-\sin(tA)\bullet A$$

（3） $$\frac{d}{dt}\sin(tA)=A\bullet\cos(tA)=\cos(tA)\bullet A$$

证明：（1） $e^{tA}=\sum_{n=0}^{\infty}\frac{1}{n!}(tA)^n$ 绝对收敛

$$\left(e^{tA}\right)_{ij}=\delta_{ij}+\frac{t}{1!}(A)_{ij}+\frac{t^2}{2!}\left(A^2\right)_{ij}+\cdots+\frac{t^k}{k!}\left(A^k\right)_{ij}+\cdots$$ 绝对收敛

$$\frac{d}{dt}\left(e^{tA}\right)_{ij}=0+\left(A\right)_{ij}+\frac{t}{1!}\left(A^2\right)_{ij}+\cdots+\frac{t^{k-1}}{(k-1)!}\left(A^k\right)_{ij}+\cdots$$

绝对收敛

$$\frac{d}{dt}e^{tA}=A+\frac{t}{1!}A^2+\cdots+\frac{t^{k-1}}{(k-1)!}A^k+\cdots$$

绝对收敛

$$=\begin{cases} A\left[I+\frac{t}{1!}A+\cdots+\frac{t^{k-1}}{(k-1)!}A^{k-1}+\cdots\right] & =Ae^{tA} \\ \left[I+\frac{t}{1!}A+\cdots+\frac{t^{k-1}}{(k-1)!}A^{k-1}+\cdots\right]A & =e^{tA}A \end{cases}$$

定义: 如果矩阵 $A(t)=\left(a_{ij}(t)\right)_{m\times n}$ 的每一个元素 $a_{ij}(t)$ 在 $[t_0,t]$ 上可积，称$A(t)$可积，记为

$$\int_{t_0}^{t} A(\tau)d\tau=\left(\int_{t_0}^{t} a_{ij}(\tau)d\tau\right)_{m\times n}$$

（1）$\int_{t_0}^{t}\left[A(\tau)+B(\tau)\right]d\tau=\int_{t_0}^{t}A(\tau)d\tau+\int_{t_0}^{t}B(\tau)d\tau$

（2）A为常数矩阵：$\int_{t_0}^{t}\left[A\bullet B(\tau)\right]d\tau=A\bullet\left[\int_{t_0}^{t}B(\tau)d\tau\right]$

B为常数矩阵：$\int_{t_0}^{t}\left[A(\tau)\bullet B\right]d\tau=\left[\int_{t_0}^{t}A(\tau)d\tau\right]\bullet B$

（3）设 $a_{ij}(t)\in C[t_0,t_1]$ 则：$\frac{d}{dt}\int_{t_0}^{t}A(\tau)d\tau=A(t)$

（4）设 $a_{ij}'(t)\in C[t_0,t_1]$, 则：$\int_{t_0}^{t_1}A'(\tau)d\tau=A(t_1)-A(t_0)$

其它微分概念

函数对矩阵的导数(包括向量)

定义:设$X=(\xi_{ij})_{m\times n}$,mn元函数

$$f(X)=f(\xi_{11},\xi_{12},\cdots,\xi_{1n},\cdots,\xi_{m\times n})$$

定义$f(X)$对矩阵X的导数为

$$\frac{df}{dX}=\left(\frac{\partial f}{\partial \xi_{ij}}\right)_{m\times n}=\begin{bmatrix}\frac{\partial f}{\partial \xi_{11}} & \cdots & \frac{\partial f}{\partial \xi_{1n}} \\ \vdots & & \vdots \\ \frac{\partial f}{\partial \xi_{m1}} & \cdots & \frac{\partial f}{\partial \xi_{mn}}\end{bmatrix}$$

例11: $x=\begin{bmatrix}\xi_1\\ \vdots\\ \xi_n\end{bmatrix}$: $f(x)=f(\xi_1,\xi_2,\cdots,\xi_n)$ $\quad \dfrac{df}{dx}=\begin{bmatrix}\dfrac{\partial f}{\partial \xi_1}\\ \vdots\\ \dfrac{\partial f}{\partial \xi_n}\end{bmatrix}$

例12: $A=(a_{ij})_{n\times n}$, $x=\begin{bmatrix}\xi_1\\ \vdots\\ \xi_n\end{bmatrix}$: $f(x)=x^T Ax$, 求 $\dfrac{df}{dx}$

$$f(x)=\xi_1\sum_{j=1}^{n}a_{1j}\xi_j+\cdots+\xi_k\sum_{j=1}^{n}a_{kj}\xi_j+\cdots+\xi_n\sum_{j=1}^{n}a_{nj}\xi_j$$

$$\frac{\partial f}{\partial \xi_k}=\xi_1 a_{1k}+\cdots+\xi_{k-1}a_{k-1,k}+\left(\sum_{j=1}^{n}a_{kj}\xi_j+\xi_k a_{kk}\right)$$

$$+\xi_{k+1}a_{k+1,k}+\cdots+\xi_n a_{nk}\quad=\sum_{j=1}^{n}a_{kj}\xi_j+\sum_{i=1}^{n}a_{ik}\xi_i$$

$$\therefore \frac{df}{dx} = \left(A + A^T\right)x$$

如果 $A = A^T$ ，有 $\frac{df}{dx} = 2Ax$

例13: $X=\left(\xi_{ij}\right)_{n\times n}$: $f(X)=\left[\mathrm{tr}(X)\right]^2$ 求 $\left.\dfrac{df}{dX}\right|_{X=I_n}$

解:

$$f(X)=\left(\xi_{11}+\xi_{22}+\cdots+\xi_{nn}\right)^2$$

$$\frac{df}{dX}=2\left(\xi_{11}+\xi_{22}+\cdots+\xi_{nn}\right)I_n$$

$$\left.\frac{df}{dX}\right|_{X=I_n}=2nI_n$$

例14: $A \in R^{m\times n}, b \in R^n$，若 $x \in R^n$ 使得 $\|Ax-b\|_2 = \min,$

则 $A^T Ax = A^T b$

解: $f(x) = \|Ax-b\|_2^2 = (Ax-b)^T (Ax-b)$

$$= x^T A^T Ax - 2b^T Ax + b^T b$$

$$g(x) = b^T Ax = b_1 \sum_{j=1}^{n} a_{1j}\xi_j + \cdots + b_m \sum_{j=1}^{n} a_{mj}\xi_j$$

$$\frac{dg}{dx} = \begin{bmatrix} \frac{\partial g}{\partial \xi_1} \\ \vdots \\ \frac{\partial g}{\partial \xi_n} \end{bmatrix} = \begin{bmatrix} b_1 a_{11} + \cdots + b_m a_{m1} \\ \vdots \\ b_1 a_{1n} + \cdots + b_m a_{mn} \end{bmatrix} = A^T b$$

$$\frac{df}{dx} = 2A^T Ax - 2A^T b = 0 \quad \Rightarrow A^T Ax = A^T b$$

【注】$r(A^T A) = r(A) \Rightarrow r(A^T A | A^T b) = r(A^T A) \Rightarrow A^T Ax = A^T b$ 有解

5、函数矩阵对矩阵的导数

定义:设 $X=\left(\xi_{ij}\right)_{m\times n}$, $f_{kl}(X)=f_{kl}\left(\xi_{11},\xi_{12},\cdots,\xi_{1n},\cdots,\xi_{m\times n}\right)$

$$F=\begin{bmatrix} f_{11} & \cdots & f_{1s} \\ \vdots & & \vdots \\ f_{r1} & \cdots & f_{rs} \end{bmatrix},\qquad \frac{\partial F}{\partial \xi_{ij}}=\begin{bmatrix} \frac{\partial f_{11}}{\partial \xi_{ij}} & \cdots & \frac{\partial f_{1s}}{\partial \xi_{ij}} \\ \vdots & & \vdots \\ \frac{\partial f_{r1}}{\partial \xi_{ij}} & \cdots & \frac{\partial f_{rs}}{\partial \xi_{ij}} \end{bmatrix},$$

定义 $$\frac{dF}{dX}=\begin{bmatrix} \frac{\partial F}{\partial \xi_{11}} & \cdots & \frac{\partial F}{\partial \xi_{1n}} \\ \vdots & & \vdots \\ \frac{\partial F}{\partial \xi_{m1}} & \cdots & \frac{\partial F}{\partial \xi_{mn}} \end{bmatrix}$$

- 可表示为

$$\frac{dF}{dX}=\left(\frac{1}{dX}\right)\otimes dF$$

例15: $x=\begin{bmatrix}\xi_1\\ \vdots\\ \xi_n\end{bmatrix}$, $F(x)=\left[f_1(x), f_2(x),\cdots, f_l(x)\right]$

$$\frac{dF}{dx}=\begin{bmatrix}\dfrac{\partial f_1}{\partial \xi_1} & \cdots & \dfrac{\partial f_l}{\partial \xi_1}\\ \vdots & & \vdots\\ \dfrac{\partial f_1}{\partial \xi_n} & \cdots & \dfrac{\partial f_l}{\partial \xi_n}\end{bmatrix}$$

例16: $A=\left(a_{ij}\right)_{n\times n}$, $\quad x=\begin{bmatrix}\xi_1\\ \vdots\\ \xi_n\end{bmatrix}$ $\quad Ax=\begin{bmatrix}\sum_{j=1}^{n} a_{1j}\xi_j\\ \vdots\\ \sum_{j=1}^{n} a_{nj}\xi_j\end{bmatrix}$

$$\frac{d(Ax)}{dx^T}=\begin{bmatrix}\dfrac{d(Ax)}{d\xi_1} & \cdots & \dfrac{d(Ax)}{d\xi_n}\end{bmatrix}=\begin{bmatrix}a_{11} & \cdots & a_{1n}\\ \vdots & & \vdots\\ a_{n1} & \cdots & a_{nn}\end{bmatrix}=A$$

矩阵分析的应用

$$\begin{cases} \xi_1'(t) = a_{11}\xi_1(t) + a_{12}\xi_2(t) + \cdots + a_{1n}\xi_n(t) + b_1(t) \\ \xi_2'(t) = a_{21}\xi_1(t) + a_{22}\xi_2(t) + \cdots + a_{2n}\xi_n(t) + b_2(t) \\ \qquad\qquad\qquad\cdots \\ \xi_n'(t) = a_{n1}\xi_1(t) + a_{n2}\xi_2(t) + \cdots + a_{nn}\xi_n(t) + b_n(t) \end{cases}$$

$$A = \begin{bmatrix} a_{11} & \cdots & a_{12} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}, x(t) = \begin{bmatrix} \xi_1(t) \\ \vdots \\ \xi_n(t) \end{bmatrix}, b(t) = \begin{bmatrix} b_1(t) \\ \vdots \\ b_n(t) \end{bmatrix}, c(t) = \begin{bmatrix} c_1(t) \\ \vdots \\ c_n(t) \end{bmatrix}$$

齐次微分方程： $x'(t) = A \bullet x(t)$

非齐次微分方程： $x'(t) = A \bullet x(t) + b(t)$

齐次微分方程的解法

定理10：齐次方程 $x'(t)=A\bullet x(t)$ 满足 $x(t_0)=x_0$ 的解存在并且唯一

证：存在性　设 $x(t)=e^{(t-t_0)A}x_0$，则

$$x'(t)=Ae^{(t-t_0)A}x_0=A\bullet x(t) \qquad x(t_0)=e^{O}x_0=x_0$$

唯一性　设 $x(t)$ 满足 $x'(t)=A\bullet x(t), x(t_0)=x_0$

$$x'(t)-Ax(t)=0 \Rightarrow e^{-tA}x'(t)+e^{-tA}(-A)x(t)=0$$

$$\Rightarrow\left[e^{-tA}x(t)\right]'=0 \quad \Rightarrow e^{-tA}x(t)=c \quad \Rightarrow x(t)=e^{tA}c$$

因为 $x(t_0)=x_0$, 所以 $x_0=e^{t_0A}c \Rightarrow c=e^{-t_0A}x_0$

因此 $x(t)=e^{tA}e^{-t_0A}x_0=e^{(t-t_0)A}x_0$

例1：设 $A=\begin{bmatrix}1 & 1 & 0\\ & 1 & 0\\ & & 2\end{bmatrix}$ 求 $x'(t)=A\bullet x(t)$ 的通解

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

$$x(t)=e^{tA}c=\begin{bmatrix}c_1e^t+c_2te^t\\ c_2e^t\\ c_3e^{2t}\end{bmatrix}$$

例2：矩阵函数 e^{tA} 的列向量 $x_1(t),\cdots,x_n(t)$ 构成

齐次方程 $x'(t)=A\bullet x(t)$ 的基础解系

解： e^{tA} 可逆 $\Rightarrow x_1(t),\cdots,x_n(t)$ 线性无关

取 $c=e_j \Rightarrow x_j(t)=e^{tA}c$ 是 $x'(t)=A\bullet x(t)$ 的一个解

通解 $x(t)=e^{tA}c=c_1\bullet x_1(t)+\cdots+c_n\bullet x_n(t)$

非齐次微分方程的解法

方程(1): $x'(t)=A\bullet x(t)$

方程(2): $x'(t)=A\bullet x(t)+b(t)$

$$\left.\begin{array}{l}\tilde{x}(t)\text{是(2)的特解}\\ x(t)\text{是(2)的通解}\end{array}\right\} \Rightarrow \left\{\begin{array}{l}\tilde{x}'(t)=A\bullet\tilde{x}(t)+b(t)\\ x'(t)=A\bullet x(t)+b(t)\end{array}\right.$$

$$\Rightarrow\left[x(t)-\tilde{x}(t)\right]'=A\left[x(t)-\tilde{x}(t)\right]\Rightarrow x(t)-\tilde{x}(t)\text{ 是(1)的解}$$

$$\Rightarrow x(t)-\tilde{x}(t)=c_1\bullet x_1(t)+\cdots+c_n\bullet x_n(t)$$

$$\Rightarrow x(t)=e^{tA}c+\tilde{x}(t)$$

非齐次微分方程的解法

采用常向量变易法求 $\tilde{x}(t)$.设 $\tilde{x}(t)=e^{tA}c(t)$ 满足(2),有

$$Ae^{tA}c(t)+e^{tA}c'(t)=Ae^{tA}c(t)+b(t)$$

$$c'(t)=e^{-tA}b(t)\Rightarrow c(t)=\int_{t_0}^{t}e^{-\tau A}b(\tau)d\tau \quad \text{(原函数之一)}$$

故(2)的通解为 $x(t)=e^{tA}\left[c+\int_{t_0}^{t}e^{-\tau A}b(\tau)d\tau\right]$

特解为 $x(t)\big|_{x(t_0)=x_0}=e^{tA}\left[e^{-t_0A}x_0+\int_{t_0}^{t}e^{-\tau A}b(\tau)d\tau\right]$

[注]当 $t_0=0$ 时，特解 $x(t)\big|_{x(0)=x_0}=e^{tA}\left[x_0+\int_{0}^{t}e^{-\tau A}b(\tau)d\tau\right]$

例3：设 $A=\begin{bmatrix}1&1&0\\&1&0\\&&2\end{bmatrix},b(t)=\begin{bmatrix}1\\0\\e^{2t}\end{bmatrix},x(0)=\begin{bmatrix}-1\\1\\0\end{bmatrix}$

求 $x'(t)=A\bullet x(t)+b(t)$ 满足初始条件 $x(0)$ 的特解

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

$$e^{-\tau A}b(\tau)=\begin{bmatrix}e^{-\tau}\\0\\1\end{bmatrix},\quad \int_0^t e^{-\tau A}b(\tau)d\tau=\begin{bmatrix}1-e^{-t}\\0\\t\end{bmatrix}$$

$$x(t)=e^{tA}\bullet\left[\begin{pmatrix}-1\\0\\1\end{pmatrix}+\begin{pmatrix}1-e^{-t}\\0\\t\end{pmatrix}\right]=\begin{bmatrix}te^t-1\\e^t\\te^{2t}\end{bmatrix}$$

矩阵微分与最优化

最简单的最优化问题是求$f(x)$的极大值和极小值

$$\min_{x \in R} f(x)$$

一般称为无约束的最优化问题

相对于n×1向量x的梯度算子记作∇_x

$$\nabla_x = \left[\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \cdots, \frac{\partial}{\partial x_n}\right]^T = \frac{\partial}{\partial x}$$

n×1实向量x为变元的实标量函数的梯度

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

实标量函数$f(A)$为相对于实矩阵 $A=\left[a_{ij}\right]_{m\times n}$ 的梯度

$$\frac{\partial f(A)}{\partial A}=\left[\frac{\partial f(A)}{\partial a_{ij}}\right]_{m\times n}=\nabla_A f(A)$$

例：CDMA系统中，有K个用户，第k个用户的扩频波形向量为$\boldsymbol{s}_k(t)$ 。假定用户k的信号幅值为 A_k

在t时刻发送比特为b_k $(+1,-1)$

在基站解扩后，基站的接收信号向量为

$$\boldsymbol{y} = \boldsymbol{RAb} + \boldsymbol{n}$$

其中 $\boldsymbol{A} = \mathbf{diag}\left(A_1, A_2, \cdots, A_K\right), \boldsymbol{b} = \left[b_1, b_2, \cdots, b_K\right]^T$

扩频相关矩阵$\boldsymbol{R}$的元素 $r_{ij} = \int_0^T s_i(t) s_j(t) dt$

设计一个多用户检测器$\boldsymbol{M} = \left[\boldsymbol{m}_1, \boldsymbol{m}_2, \cdots, \boldsymbol{m}_K\right]$，使得

$$\hat{b}_k = \mathrm{sgn}(\boldsymbol{m}_k^T \boldsymbol{y})$$

将K个用户的检测器联合考虑，构造目标函数

$$\boldsymbol{J}(\boldsymbol{M}) = \mathbf{E}\left[\left\|\boldsymbol{b} - \boldsymbol{M}\boldsymbol{y}\right\|_2^2\right]$$

使其最小化，即可得到最优的盲多用户检测器$\boldsymbol{M}$

利用矩阵迹的性质，可得

$$\begin{aligned}\boldsymbol{J}(\boldsymbol{M}) &= \mathbf{E}\left\{\left(\boldsymbol{b} - \boldsymbol{M}\boldsymbol{y}\right)^T\left(\boldsymbol{b} - \boldsymbol{M}\boldsymbol{y}\right)\right\} \\ &= \mathbf{E}\left\{\mathbf{tr}\left[\left(\boldsymbol{b} - \boldsymbol{M}\boldsymbol{y}\right)\left(\boldsymbol{b} - \boldsymbol{M}\boldsymbol{y}\right)^T\right]\right\} \\ &= \mathbf{tr}\left\{\mathbf{E}\left[\left(\boldsymbol{b} - \boldsymbol{M}\boldsymbol{y}\right)\left(\boldsymbol{b} - \boldsymbol{M}\boldsymbol{y}\right)^T\right]\right\} \\ &= \mathbf{tr}\left\{\boldsymbol{cor}\left(\boldsymbol{b} - \boldsymbol{M}\boldsymbol{y}\right)\right\}\end{aligned}$$

其中 $\boldsymbol{cor}\left(\boldsymbol{b} - \boldsymbol{M}\boldsymbol{y}\right) = \mathbf{E}\left[\left(\boldsymbol{b} - \boldsymbol{M}\boldsymbol{y}\right)\left(\boldsymbol{b} - \boldsymbol{M}\boldsymbol{y}\right)^T\right]$是自相关矩阵

在加性噪声与用户信号不相关时有

$$\mathrm{cor}\left(b - My\right) = I + M\left(RA^2R + \sigma^2 R\right)M^T - ARM^T - MRA$$

其中加性噪声的方差为 σ^2

于是目标函数可写作

$$J(M) = \mathrm{tr}\left\{cor\left(b - My\right)\right\}$$

$$= \mathrm{tr}(I) + \mathrm{tr}\left(M\left(RA^2R + \sigma^2 R\right)M^T\right) - \mathrm{tr}\left(ARM^T\right) - \mathrm{tr}\left(MRA\right)$$

利用迹函数的微分公式

$$\frac{\partial \mathrm{tr}\left(M^T B\right)}{\partial M} = \frac{\partial \mathrm{tr}\left(BM^T\right)}{\partial M} = B \qquad \frac{\partial \mathrm{tr}\left(MB\right)}{\partial M} = \frac{\partial \mathrm{tr}\left(BM\right)}{\partial M} = B^T$$

$$\frac{\partial\,\mathrm{tr}\left(MDM^T\right)}{\partial M}=M(D+D^T)$$

因为 $D=RA^2R+\sigma^2R$ 是对称矩阵，所以

$$\frac{\partial J(M)}{\partial M}=2M\left(RA^2R+\sigma^2R\right)-2AR$$

令其为零，即可得

$$M\left(RA^2R+\sigma^2R\right)=AR$$

如果R非奇异，可得最优的多用户检测器为

$$M=A\left(RA^2+\sigma^2I\right)^{-1}$$

作业

- **P163：1、2、5、6**
- **P170：4、5、6**