1. $[\underline{A}x]_i = \sum_{j=1}^{n} a_{ij}x_j$

$B = A^HA = (A^HA)^H = B^H$

2. 运算法则

$\frac{dA}{dt} = \dot{A} = \begin{bmatrix} \frac{da_{11}}{dt} & \cdots & \\ \vdots & \ddots & \frac{da_{mn}}{dt} \end{bmatrix}$

$\frac{d(AB)}{dt} = \frac{dA}{dt}B + A\frac{dB}{dt}$ $\quad \frac{de^{At}}{dt} = Ae^{At}$

$\int A\,dt = \begin{bmatrix} \int a_{11}dt & \cdots \\ \vdots & \ddots \end{bmatrix}$

3. 奇异矩阵：$Ax=0$ 存在非零解

(向量组线性相关)

4. $(A_r + jA_i)(x_r + jx_i) = b_r + jb_i$

$\left[\begin{matrix} A_r & -A_i \\ A_i & A_r \end{matrix} \middle| \begin{matrix} b_r \\ b_i \end{matrix}\right] \to \left[\begin{matrix} I_n & 0 & x_r \\ 0 & I_n & x_i \end{matrix}\right]$

5. $\forall$ 任意. $\ni$ 使得 $\exists$ 存在

6. 线性映射：叠加性 & 齐次性

$T(av+bw) = aT(v) + bT(w)$

7. 内积、$\langle x,y\rangle = x^Hy$ 或 $\int_{*} x^*(t)y(t)\,dt$

加权内积、x^HGy G为正定H矩阵

L_0 L_1 L_2 范数 $\quad L_\infty$ (最大绝对值)

↑ 元素的

L_p 范数 $(\sum |x_i|^p)^{\frac{1}{p}}$

8. 矩阵内积、$\langle A,B\rangle = \mathrm{vec}(A)^H\mathrm{vec}(B) = \mathrm{tr}(A^HB)$

诱导范数 $\|A\| = \max\{\frac{\|Ax\|}{\|x\|}, x \neq 0\}$

$\|A\|_{spec} = \|A\|_p\ (p=2) \Rightarrow$ 最大奇异值

$p=1$ 最大列和 ，$p=\infty$ 最大行和

9. "元素形式"范数

L_1 L_2(Frobenius) L_∞

10. 统计

相关矩阵 $R_x = E\{x(\xi)x^H(\xi)\} \in H$

自协方差 $C_x = \mathrm{Cov}(x,x) = E\{(x-\mu)(x-\mu)^H\}$

$= R_x - \mu\mu^H$

互相关 $R_{xy} = E(xy^H)$

互协方差 $C_{xy} = R_{xy} - \mu_x\mu_y^H$

△复高斯的 $E(xx^H) = \sigma I$ $\quad$ 实虚独立

但 $E(xx^T) = 0$ $\quad$ 因为 $E(x_{Rk}^2(t))$

$= E(x_{Ik}^2(t)) = \frac{1}{2}\sigma^2$

11. 二次型 $f(x) = x^HAx \in \boxed{R}$

(其中 $A \in H$) $\underline{x \neq 0}$

$f(x) > 0 \Leftrightarrow$ 所有特征值为正

$\mathrm{tr}(A) = \sum \lambda$

$\det(A) = \prod \lambda$

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

$d|X| = |X|\,\mathrm{tr}(X^{-1}dX)$

$d|F(X)| = |F(X)|\,\mathrm{tr}\left(F^{-1}(X)\,dF(X)\right)$

---

1. 左逆矩阵 $LA = I$ $\quad$ 伪逆

$L = (A^HA)^{-1}A^H$ $\quad m > n$

Moore-Penrose 广义逆.

$AA^{\dagger}A = A$ $\quad A^{\dagger}AA^{\dagger} = A^{\dagger}$

$AA^{\dagger}$、$A^{\dagger}A \in H$

(1) 超定方程 $Ax=b$.

$m > n$ 满列秩

$x = (A^HA)^{-1}A^Hb$ 最小二乘解.

(2) 欠定方程 $Ax=b$

$m < n$ 满行秩

~~$x = A^H(A^H$~~ $x = A^H(AA^H)^{-1}b$

最小二范数解.

2. 直和 $A \oplus B = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$

Hadamard 积、$(A * B)_{ij} = a_{ij}b_{ij}$

Kronecker 积

3. $(A + xy^H)^{-1} = A^{-1} - \frac{A^{-1}xy^HA^{-1}}{1 + y^HA^{-1}x}$

---

1. 反Hermitian 矩阵 $R^H = -R$

2. 下三角 $(\forall a_{ij} = 0, i<j)$

严格下三角 $(\forall a_{ij} = 0, i \le j)$

单位下三角 $(a_{ii} = 1)$

3. 相似矩阵 $B = S^{-1}AS$ $(n \times n)$

S是非奇异 $n \times n$ $\Leftrightarrow B \sim A$

$\det(B) = \det(A) = \prod \lambda$

$\mathrm{tr}(B) = \mathrm{tr}(A)$

相似阵特征值相同 △

特征向量为线性变换关系 △

请记住：$\det(AB) = \det(A)\det(B)$ △

---

1. Jacobian (1) 标量函数

$D_x = \frac{\partial}{\partial x^T} = [\frac{\partial}{\partial x_1} \cdots \frac{\partial}{\partial x_m}]$

对于矩阵 ① $D_X = \frac{\partial}{\partial X^T} = \begin{bmatrix} \frac{\partial}{\partial x_{11}} & \frac{\partial}{\partial x_{m1}} \\ \frac{\partial}{\partial x_{1n}} & \frac{\partial}{\partial x_{mn}} \end{bmatrix} \in \underline{\underline{R}}^{n \times m}$

② $D_{\mathrm{vec}X} = \frac{\partial}{\partial \mathrm{vec}(X)^T} = [\frac{\partial}{\partial x_{11}} \cdots \frac{\partial}{\partial x_{m1}} \cdots \frac{\partial}{\partial x_{1n}} \cdots \frac{\partial}{\partial x_{mn}}]$

(2) 矩阵函数

~~$D_X f(\underline{X})$~~ $D_XF(\underline{X}) = \begin{bmatrix} \frac{\partial f_{11}}{\partial \mathrm{vec}X^T} \\ \vdots \\ \frac{\partial f_{p1}}{\partial \mathrm{vec}X^T} \\ \vdots \\ \frac{\partial f_{1q}}{\partial} \\ \frac{\partial f_{pq}}{\partial \mathrm{vec}X^T} \end{bmatrix} \in \underline{\underline{R}}^{pq \times mn}$

2. 梯度函数 (算子)

$\nabla_x = \frac{\partial}{\partial x} = [\frac{\partial}{\partial x_1} \cdots \frac{\partial}{\partial x_m}]^T$

$\nabla_{\mathrm{vec}\underline{X}} = \frac{\partial}{\partial \mathrm{vec}\underline{X}}$

$\nabla_{\underline{X}} f(\underline{X}) = [\text{直接对应}] \in m \times n$

$\dot{\underline{x}} = -\nabla_x f(x)$ $\quad \dot{\underline{X}} = -\nabla_{\mathrm{vec}\underline{X}} f(\underline{X})$ 梯度流

$\nabla_{\underline{X}} F(\underline{X}) = [D_XF(\underline{X})]^T$

---

3. 正则化：改善奇异性

求解奇异/病态方程组

cond(A) 减小

反正则化：对满秩(列)

但存在噪声或误差，

引入一个小的负扰动

$-\lambda I$，使 A^HA 去干扰.

求解

1. 条件数：刻画线性方程(组)时，

误差经过矩阵A的传播扩大为解向

量的误差的程度，因此是衡量线性

方程数值稳定性的一个重要指标.

2. 外罚函数对可行集以外的所有点进行惩罚

求出的解均满足约束条件，是不等式约束优化

问题的精确(不等式)解，是一种最优设计方案；

内罚阻挡了可行集边界的点，是原始优化问题

的近似解，是一种次优设计方案。

外罚可用不可行点启动，收敛慢；内罚要求初

始点满足约束，选择较困难，但收敛性能好.

全部 $\quad$ (确实) 只满足严格不等式

$f_i(x) < 0$ (无二)

3. $\frac{\partial x^TAx}{\partial x^T} = x^T(A^T + A)$

$d[\mathrm{tr}(U)] = \mathrm{tr}[d(U)]$

$d(UV) = (dU)V + U(dV)$ △

$df(\underline{x}) = \frac{\partial f(x)}{\partial \underline{x}^T}d\underline{x} = \mathrm{tr}(\underline{\underline{A}}\,d\underline{x})$

A为 $D_x f(\underline{x})$

(1) 矩阵变元函数同理.

(2) $\nabla_X f(X) = A^T$

eg $d\,\mathrm{tr}(AX^{-1}) = \mathrm{tr}\,d(AX^{-1}) = \mathrm{tr}(A\,dX^{-1})$

$= -\mathrm{tr}(AX^{-1}dXX^{-1}) = -\mathrm{tr}(X^{-1}AX^{-1}dX)$

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

4. $\frac{\partial}{\partial z} = (\frac{\partial}{\partial x} - j\frac{\partial}{\partial y})\frac{1}{2}$ $\quad \frac{\partial}{\partial z^*} = \frac{1}{2}(\frac{\partial}{\partial x} + j\frac{\partial}{\partial y})$

$df(z, z^*) = \frac{\partial f}{\partial z^T}dz + \frac{\partial f}{\partial z^H}dz^*$

5. Hessian 矩阵

$H[f(\underline{x})] = \nabla_x[D_x f(x)]$

$= \frac{\partial}{\partial \underline{x}}[\frac{\partial f(x)}{\partial x^T}]$

$d^2 f(x) = dx^T H(\,)\, dx$

不关注复变的极值平稳.

(只关注实变)

1. 无约束优化

$\frac{\partial f}{\partial z^*}\Big|_{z=z_0}=0$

$\begin{bmatrix} H_{z^*,z} & H_{z^*,z^*} \\ H_{z,z} & H_{z,z^*} \end{bmatrix} \succeq 0$

(严格:全局 局部极小)

2. 凸函数

$\Leftrightarrow \frac{\partial^2 f(x)}{\partial x \partial x^T} \succeq 0$ ~~$H_x f(x)$~~

即 $H_x f(x) \succeq 0$

3. $\nabla_{z^*} f(z,z^*)$ 梯度

$\nabla_{\mathrm{vec}\, z^*} f(z,z^*)$

4. Newton法

$f(x+\Delta x)=f(x)+(\nabla f(x))^T\Delta x+\frac{1}{2}(\Delta x)^T\nabla^2 f(x)\Delta x$

$\frac{\partial f(x+\Delta x)}{\partial \Delta x}=\nabla f(x)+\nabla^2 f(x)\Delta x=0$

$\Rightarrow \Delta x=-(\nabla^2 f(x))^{-1}\nabla f(x)$

最速下降法 $\Delta x=-\nabla f(x)$ (沿梯度进行)

5. 约束优化

(1) Lagrange

$\min_x f(x)$ s.t. $Ax=b$

$L(x,\lambda)=f(x)+\lambda^T(Ax-b)$

(2) 罚函数

① 等式约束 $p(x)=\sum_{i=1}^{q}|h_i(x)|^2$ (加性) $\min_x f_0(x)$ s.t. $h_i(x)=0$

都是加性

② 不等式 a. 外罚 $p(x)=\sum_{i=1}^{m}(\max\{0,-f_i(x)\})^r$ $r=1$

s.t. $f_i(x)\geq 0$

$\boxed{\min_x} f_0(x)$ b. 内罚 $p(x)=\sum_{i=1}^{m}\frac{1}{f_i(x)}$ $f_i(x)\to 0$ 则 $p(x)\to\infty$

$p(x)=\sum_{i=1}^{m}\frac{1}{f_i(x)}|\log f_i(x)|$

$p(x)=\sum_{i=1}^{m}e^{\frac{1}{f_i(x)}}$

$h_j(x)=0,\ f_i(x)\leq 0$

③ 混合

a. 外 $\min_x f_0(x)+\rho_1\sum(\max\{0,f_i(x)\})^r+\rho_2\sum|h_j(x)|^2$

b. 内 $\min_x f_0(x)+\rho_1\sum\frac{1}{-f_i(x)}|\log(-f_i(x))|+\rho_2\sum|h_j(x)|^2$

(3) 增广 Lagrange

① $\min_x f_0(x)$ s.t. $h(x)=0$

$L_\rho(x,\lambda)=f_0(x)+\lambda^T h(x)+\rho\|h(x)\|_2^2$

② s.t. $Ax=b,\ Bx\preceq h$ 松弛变量

$L_\rho(x,\lambda,v,s)=f_0(x)+\lambda^T(Ax-b)+v^T(Bx+s-h)+\frac{\rho}{2}(\|Ax-b\|_2^2+\|Bx+s-h\|_2^2)$

1. 普通LS $Ax=b+\Delta b$

$\min\|Ax-b\|_2^2$

$\phi(x)=(Ax-b)^H(Ax-b)$

$x=(A^HA)^{-1}A^Hb$ 满列秩

2. 数据LS (DLS)

~~$(A+\Delta A)x=b$~~

$(A+\Delta A)x=b$

$\min\|\Delta A\|_F^2$ s.t.

$L(x,\lambda)=\mathrm{tr}(\Delta A\Delta A^H)+\lambda^H(Ax+\Delta Ax-b)$

$\frac{\partial L}{\partial \Delta A^T}=\Delta A^H+x\lambda^H=0$

$\Rightarrow \lambda=\frac{Ax-b}{x^Hx}$

$\Delta A=-\frac{(Ax-b)x^H}{x^Hx}$

$\Rightarrow J(x)=\|\Delta A\|_F^2=\frac{\|Ax-b\|_2^2}{\|x\|_2^2}$

3. 总体LS (TLS)

$\min\|[\Delta A,\Delta b]\|_F^2$

s.t. $(A+\Delta A)x=b+\Delta b$

$A'=[A,b]$ $x'=\begin{bmatrix}x\\-1\end{bmatrix}$

$\Delta A'=[\Delta A,\Delta b]$

$(A'+\Delta A')x'=0$

$\Rightarrow J(x)=\frac{\|A'x'\|_2^2}{\|x'\|_2^2}=\frac{x'^HA'^HA'x'}{x'^Hx'}$

4. Tikhonov 正则

对于普通LS，A^HA 奇异时，

$J(x)=\|Ax-b\|_2^2+\lambda\|x\|_2^2$

$\frac{\partial J(x)}{\partial x^*}=A^HAx-A^Hb+\lambda x=0$

$\Rightarrow (A^HA+\lambda I)x-A^Hb=0$

$\hat{x}=(A^HA+\lambda I)^{-1}A^Hb$

对角线加载，改善奇异性。

TLS：到直线距离平方和最小

过"零点" $(\bar{x},\bar{y})$ 中心点

普通LS拟和

$D_{LS}=\|(x-\bar{x})+m(y-\bar{y})\|^2$

$\frac{\partial D_{LS}}{\partial m}=0$ 最小 D_{LS}

$d=\frac{\|Ax-b\|_2^2}{\|x\|_2^2+1}$

$d_{min}^2=\lambda_{min}(M^TM)$

$A=\begin{bmatrix}x_1-\bar{x}\\ \vdots \\ x_m-\bar{x}\end{bmatrix}$ $b=\begin{bmatrix}y_1-\bar{y}\\ \vdots \\ y_m-\bar{y}\end{bmatrix}$

$a(x-\bar{x})+b(y-\bar{y})=0$

$k=-\frac{a}{b}$

$M=[A,b]$ $x'=\begin{bmatrix}k\\-1\end{bmatrix}$

$D_{TLS}=\frac{x'^TM^TMx'}{x'^Tx'}$

EVD(M^TM)

1. 特征值

$f(A)=A^n+c_1A^{n-1}+\cdots+c_nI$

特征值为 $f(\lambda)$

e^A 的特征值 $\sum_{i=0}^{\infty}\frac{\lambda^i}{i!}=e^\lambda$

$p(A)=p_nA^n+\cdots+p_0=0$

$p_nA^{n-1}+\cdots+p_0A^{-1}=0$

(没用) $A^{-1}=-\frac{p_nA^{n-1}+\cdots+p_1I}{p_0}$

2. $A=X\Lambda X^{-1}$ Λ: diag(λ)

$A^2=X\Lambda^2X^{-1}$ X是特征向量(矩阵)

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

3. 对称矩阵

$A=A^T \Rightarrow A=Q\Sigma Q^T$

Q为正交矩阵，$\Sigma\in R$ 实

$A=A^H\Rightarrow A=U\Sigma U^H$

U为酉矩阵 $\Sigma\in R$ 实

4. 噪声白化处理

(1) $x_0=x-m_x$ (标准正交变换)

$R_{x_0}=C_x\in H$

$C_x=U_x\Sigma_xU_x^H$

$m_{x_0}=0$

(2) 令 $w=U_x^Hx_0$

$m_w=E(w)=0$

$R_w=C_w=E(ww^H)$

$=U_x^HE(x_0x_0^H)U_x$

$=U_x^HC_xU_x$

$=\Sigma_x$ (即 C_x 特征值)

diag(λ)

5. 逆向图变换

$y=\Sigma_x^{-\frac{1}{2}}w$

则 $R_y=I$

标准白化处理

6. 主成分分析PCA

$R_x=U\Sigma U^H$

$\hat{x}_j=u_j^Hx$

$\hat{x}=U^Hx$ (前k)

$E\{|\hat{x}_j|^2\}$

$=u_j^HU\Sigma U^Hu_j$

$=\lambda_j$

7. Rayleigh商

$\frac{x^HAx}{x^Hx}=\lambda$

对于TLS

$f(x')=\frac{x'^HA'^HA'x'}{x'^Hx'}$

最小 λ 时的 x'

把最后元素归一化

$x'=-\frac{1}{v(n+1,n+1)}\begin{bmatrix}x\\-1\end{bmatrix}$

8. 数值稳定性

$\frac{\|\delta x\|_2}{\|x\|_2}\leq(\|A\|_2\|A^{-1}\|_2)\frac{\|\delta b\|_2}{\|b\|_2}$

$\frac{\|\delta x\|_2}{\|x+\delta x\|_2}\leq(\|A\|_2\|A^{-1}\|_2)\frac{\|\delta A\|_2}{\|A\|_2}$

$\mathrm{cond}(A)=\frac{\sigma_1}{\sigma_p}$

$\mathrm{cond}(A)=\frac{\sigma_1}{\sigma_p}\geq 1$

改善：

①使用截断SVD

$\frac{\sigma_1}{\sigma_r}<\frac{\sigma_1}{\sigma_p}$

1. SVD

$Y=U\Sigma V^H$ Σ $m\times n$

$=\sum_{i=1}^{r}u_i\sigma_iv_i^H$ r个秩1矩阵

2. YY^H Y^HY

↓EVD ↓EVD

$YY^H=U\Sigma^2U^H$ $Y^HY=V\Sigma^2V^H$

左 $\sigma_i=\sqrt{\lambda_i}$ 右

3. 对于A的 $L=(A^HA)^{-1}A^H$

$L=V[\Sigma_n^{-1}\ 0]U^H$ (列满秩)

$R=A^H(AA^H)^{-1}$

$R=V\begin{bmatrix}\Sigma_m^{-1}\\0\end{bmatrix}U^H$ (行满秩)

4. 秩亏缺时，

$A^\dagger=(U\Sigma V^H)^\dagger=U\Sigma^\dagger V^H$

$\Sigma^\dagger=\begin{bmatrix}\frac{1}{\sigma_1} & & & \\ & \ddots & & \\ & & \frac{1}{\sigma_r} & \\ & & & 0_{0_0}\end{bmatrix}$

5. $\|A\|_2=\sigma_{max}$

$\|A\|_F^2=\mathrm{tr}(A^HA)=\sum_{i=1}^{r}\sigma_i^2$

6. 截断SVD

$Y=U\Sigma V^H$ V是 $n\times r$

$m\times r$ $r\times r$ $r\times n$

7. 低秩逼近 秩r Y 秩p $\hat{Y}$

~~$\|Y-\hat{Y}\|_2=\sigma$~~

$Y-\hat{Y}=\sum_{i=p+1}^{r}u_i\sigma_iv_i^H$

$\|Y-\hat{Y}\|_2=\sigma_{p+1}$

$\|Y-\hat{Y}\|_F=\sqrt{\sigma_{p+1}^2+\cdots+\sigma_r^2}$

① $\frac{\sigma_{p+1}}{\sigma_1}\geq\varepsilon$ 阈值 ② $\frac{\sqrt{\sigma_1^2+\cdots+\sigma_p^2}}{\sqrt{\sigma_1^2+\cdots+\sigma_r^2}}\geq\alpha$

③ ~~$\frac{\|Y-\hat{Y}\|_F}{\sqrt{\sigma_1^2+\cdots+\sigma_r^2}}\geq\alpha$ 阈值~~

对于超定方程LS 更不稳定

$x=(A^HA)^{-1}A^Hb$ $\mathrm{cond}(A^HA)=[\mathrm{cond}(A)]^2=\frac{\sigma_1^2}{\sigma_p^2}$

$\mathrm{cond}(B)=\frac{\sigma_1+\lambda}{\sigma_p+\lambda}$ ②对角线加载，

$B=A+\lambda I$ 引入正则