2018年矩阵论

一、证明以下变换是否为线性变换

1、x=[x_1, x_2, x_3,], T(x)=[x_1^2, x_1+x_2, x_3]

2、X 是矩阵，T(X) = BXC, B 和 C 是给定矩阵

（利用叠加性和齐次性）

二、正则化的最小二乘解

证明 $J(\boldsymbol{x}) = \frac{1}{2}\left(\|\boldsymbol{Ax}-\boldsymbol{b}\|_2^2 + \lambda\|\boldsymbol{x}\|_2^2\right)$ 的最优解为 $\boxed{\hat{\boldsymbol{x}}_{\mathrm{Tik}} = (\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}+\lambda\boldsymbol{I})^{-1}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{b}}$

三、

1、证明

三、（1）若 $|B| \neq 0$，求证 $A = B^H B$ 是正定矩阵

将其中的 $\mathrm{B^H B}$ 改为 $\mathrm{B}B^H$，证明方法应该是一样的

2、证明反对称矩阵的特征值为 0 或者纯虚数

四、

四、已知 $A=\begin{bmatrix}1 & 0\\ 0 & 1\\ 1 & 1\end{bmatrix}$，求 A 的奇异值以及 A 的右奇异向量

（注意不是右奇异矩阵哦）

五、

1、证明 $\frac{\mathrm{dtr(BX)}}{dX} = \frac{\mathrm{dtr}(X^T B^T)}{dX} = B^T$

例 3.1.3 考查目标函数 $f(\boldsymbol{X}) = \mathrm{tr}(\boldsymbol{XB})$，其中 $\boldsymbol{X}$ 和 $\boldsymbol{B}$ 分别为 $m\times n$ 和 $n\times m$ 实矩阵。首先，矩阵乘积的元素为 $[\boldsymbol{XB}]_{kl} = \sum_{p=1}^{n} x_{kp}b_{pl}$，故矩阵乘积的迹 $\mathrm{tr}(\boldsymbol{XB}) = \sum_{p=1}^{n}\sum_{l=1}^{m} x_{lp}b_{pl}$。于是，利用式 (3.1.27)，易求得

$$\left[\frac{\partial \mathrm{tr}(\boldsymbol{XB})}{\partial \boldsymbol{X}^{\mathrm{T}}}\right]_{ij} = \frac{\partial}{\partial x_{ji}}\left(\sum_{p=1}^{n}\sum_{l=1}^{m} x_{lp}b_{pl}\right) = \sum_{p=1}^{n}\sum_{l=1}^{m}\frac{\partial x_{lp}}{\partial x_{ji}}b_{pl} = b_{ij}$$

即有 $\frac{\partial \mathrm{tr}(\boldsymbol{XB})}{\partial \boldsymbol{X}^{\mathrm{T}}} = \boldsymbol{B}$。又由于 $\mathrm{tr}(\boldsymbol{BX}) = \mathrm{tr}(\boldsymbol{XB})$，故 $n\times m$ Jacobian 矩阵和 $m\times n$ 梯度矩阵分别为

$$\mathrm{D}_{\boldsymbol{X}}\mathrm{tr}(\boldsymbol{XB}) = \mathrm{D}_{\boldsymbol{X}}\mathrm{tr}(\boldsymbol{BX}) = \boldsymbol{B} \quad 和 \quad \nabla_{\boldsymbol{X}}\mathrm{tr}(\boldsymbol{XB}) = \nabla_{\boldsymbol{X}}\mathrm{tr}(\boldsymbol{BX}) = \boldsymbol{B}^{\mathrm{T}} \qquad (3.1.2$$

2、

求实值函数 $f(\boldsymbol{x}) = \boldsymbol{x}^{\mathrm{T}} \boldsymbol{A} \boldsymbol{x}$ 的 Jacobian 矩阵。

$$\mathrm{D}f(\boldsymbol{x}) = \boldsymbol{x}^{\mathrm{T}} \boldsymbol{A} + \boldsymbol{x}^{\mathrm{T}} \boldsymbol{A}^{\mathrm{T}} = \boldsymbol{x}^{\mathrm{T}}(\boldsymbol{A} + \boldsymbol{A}^{\mathrm{T}})$$

六、

八．设 x 为一随机向量，且$x\epsilon\mathbb{R}^{m\mathrm{x}1}$，其均值向量为 m_x，协方差矩阵为 C_x。求对 x 进行迷向圆变换。

七、

六、已知数据点（2，4)，（2，1)（5，1)，请求出总体最小二乘和一般最小二乘（点到直线的水平距离）的拟合直线，并分析它们 D_{TLS} 和 D_{LS}。

这次考的是把一般最小二乘点到直线水平的距离改成点到垂直的举例（其实我觉得说是竖直举例更合理），记得分清楚 m 写在 x 前还是 y 前面

八、没想到竟然会考作业题哇~~可怕

4.3 考虑方程$\boldsymbol{y} = \boldsymbol{A\theta} + \boldsymbol{e}$，其中$\boldsymbol{e}$为误差向量。定义加权误差平方和 $E_w = \boldsymbol{e}^H \boldsymbol{W} \boldsymbol{e}$，其中$\boldsymbol{W}$为一Hermitian正定矩阵，它对误差起加权作用。

（1）求使E_w最小化的参数向量$\boldsymbol{\theta}$的解。这一解称为$\boldsymbol{\theta}$的加权最小二乘估计。

$$E_w = \boldsymbol{e}^H \boldsymbol{W} \boldsymbol{e} = (\boldsymbol{y} - \boldsymbol{A\theta})^H \boldsymbol{W} (\boldsymbol{y} - \boldsymbol{A\theta})$$

对 $\boldsymbol{\theta}^*$ 求梯度：

$$\begin{aligned}\frac{\partial E_w}{\partial \boldsymbol{\theta}^*} &= \frac{\partial}{\partial \boldsymbol{\theta}^*}(\boldsymbol{y}^H \boldsymbol{W} \boldsymbol{y} - \boldsymbol{y}^H \boldsymbol{W} \boldsymbol{A\theta} - \boldsymbol{\theta}^H \boldsymbol{A}^H \boldsymbol{W} \boldsymbol{y} + \boldsymbol{\theta}^H \boldsymbol{A}^H \boldsymbol{W} \boldsymbol{A\theta}) \\ &= \boldsymbol{A}^H \boldsymbol{W} \boldsymbol{A\theta} - \boldsymbol{A}^H \boldsymbol{W} \boldsymbol{y}\end{aligned}$$

令共轭梯度为0，求得$\boldsymbol{\theta}$的加权最小二乘估计为：

$$\widehat{\boldsymbol{\theta}_{LS}} = (\boldsymbol{A}^H \boldsymbol{W} \boldsymbol{A})^{-1} \boldsymbol{A}^H \boldsymbol{W} \boldsymbol{y}$$

（2）利用$\mathbf{LDL^H}$分解$\mathbf{W = LDL^H}$，证明加权最小二乘准则相当于使误差或数据向量进行预白化。

令 $\boldsymbol{W} = \boldsymbol{LDL^H} = \boldsymbol{LD^{1/2}D^{1/2}L^H}$

将上式带入（1）中$\boldsymbol{\theta}$的加权最小二乘估计解可得：

$$\widehat{\boldsymbol{\theta}_{LS}} = (\boldsymbol{A^H W A})^{-1}\boldsymbol{A^H W y} = (\boldsymbol{A^H LDL^H A})^{-1}(\boldsymbol{A^H LDL^H y})$$
$$= [(\boldsymbol{D^{1/2}L^H A})^H \boldsymbol{D^{1/2}L^H A}]^{-1}(\boldsymbol{D^{1/2}L^H A})^H \boldsymbol{D^{1/2}L^H y}$$
$$= (\boldsymbol{B^H B})^{-1}\boldsymbol{B^H D^{1/2}L^H y}$$

对$\boldsymbol{y} = \boldsymbol{A\theta} + \boldsymbol{e}$两边同时左乘$\boldsymbol{D^{1/2}L^H}$，可得：

$$\boldsymbol{z} = \boldsymbol{D^{1/2}L^H}\mathbf{y} = \boldsymbol{D^{1/2}L^H A\theta} + \boldsymbol{D^{1/2}L^H e} = \boldsymbol{B\theta} + \boldsymbol{\varepsilon}$$
$$E\{\boldsymbol{\varepsilon\varepsilon^H}\} = \boldsymbol{D^{1/2}L^H}E\{\boldsymbol{ee^H}\}\boldsymbol{LD^{1/2}}$$

设$E\{\boldsymbol{ee^H}\} = \boldsymbol{R_e}$，当$\boldsymbol{R_e} = \boldsymbol{L_e D_e L_e^H}$且$\boldsymbol{L_e} = \boldsymbol{L^{-H}}$　$\boldsymbol{D_e} = \boldsymbol{D^{-1}}$时：

$$E\{\boldsymbol{\varepsilon\varepsilon^H}\} = \boldsymbol{D^{1/2}L^H}E\{\boldsymbol{ee^H}\}\boldsymbol{LD^{1/2}} = \boldsymbol{I}$$

所以证明成立。

九、概念题

1、举例写出不等式约束下的外罚函数法和内罚函数法的目标函数

2、条件数表征的意义和改善条件数的方法

3、广义 rayleign 商取到最大值和最小值的条件