Homework #0　　17.11.29

(1) 组合数的证明.（使用数学归纳法）

证. 1) 当 $N=1$ 时,

$C(N,k)=C(1,k)=1$　　$k=0,1$ 原式成立

2) 假设 $N=n-1$ 时, 原等式成立. 即

$$C(N,k)=C(n-1,k)=\frac{(n-1)!}{k!(n-k-1)!}\quad 0\le k\le n-1$$

3) 则当 $N=n$ 时.

$$C(N,k)=C(n,k)=C(n-1,k)+C(n-1,k-1)$$

$$=\frac{(n-1)!}{k!(n-k-1)!}+\frac{(n-1)!}{(k-1)!(n-k)!}$$

$$=\frac{n!}{k!(n-k)!}\cdot\left(\frac{n-k}{n}+\frac{k}{n}\right)\text{（凑项）}$$

$$=\frac{n!}{k!(n-k)!}=C(n,k)\quad 0\le k\le n$$

(2) 1) $X\sim B(10,1/2)$　$P(X=4)=C_{10}^4\times(1/2)^4\times(\frac{1}{2})^6=C_{10}^4/2^{10}$

2) 题目的意思为: 从52张扑克中选出5张, 使其选出的结果可以表示成"XXXYY"的形式的概率大小.

a. 总的可能数为 C_{52}^5.

b. 从13类扑克(52张)中选出2类 A_{13}^2

c. 从其中一类中的4张选出3张 C_4^3

d. 同样从另一类选出2张 C_4^2

则 $P=(A_{13}^2\times C_4^3\times C_4^2)/C_{52}^5$.

(3) 条件概率，需要注意的是分母的至少有一次是"L头"。

所以 $P(B|A)=P(AB)/P(A)=(\frac{1}{2})^3/(1-(1/2)^3)=1/7$

其中 A 表示至少有一次是"L头"

B 表示三次都是 L头

(4) 贝叶斯判.及全概率公式

先上公式 $P(B|A)=\frac{P(AB)}{P(A)}=\frac{P(B)\cdot P(A|B)}{\sum_{i=0}^{2} P(A|B_i)\cdot P(B_i)}$

设 A 为 $|x|=1$. B 为 $x<0$, 即 $x=-1$.

$\therefore P(B|A)=\frac{P(B)P(A|B)}{P(A)}=\frac{P(AB)}{P(A)}=\frac{P(AB)}{P(A)}$

$=(\frac{1}{2}\times\frac{1}{4})\div(\frac{1}{2}\times\frac{1}{8}+\frac{1}{2}\times1/4)=2/3$

(5) $\max P(A\cap B)=\min(P(A),P(B))=0.3$

$\min P(A\cap B)=\max(0,P(A)+P(B)-1)=0.$

$\max P(A\cup B)=\min(1,P(A)+P(B))=0.7$

$\min P(A\cup B)=\max(P(A),P(B))=0.4$

(6) 证明. $\sigma_x^2=\frac{1}{N-1}\sum_{n=1}^{N}(x_n-\bar{x})^2$

$=\frac{1}{N-1}\sum_{n=1}^{N}(x_n^2-2x_n\bar{x}+\bar{x}^2)$

$=\frac{1}{N-1}(\sum_{n=1}^{N}x_n^2-\sum_{n=1}^{N}2\cdot x_n\cdot\bar{x}+\sum_{n=1}^{N}\bar{x}^2)$

$=\frac{1}{N-1}(\sum_{n=1}^{N}x_n^2-2\cdot N\cdot\bar{x}\cdot\bar{x}+N\bar{x}^2)$

$=\frac{1}{N-1}(\sum_{n=1}^{N}x_n^2-N\bar{x}^2)$

$=\frac{N}{N-1}(\frac{1}{N}\sum_{n=1}^{N}x_n^2-\bar{x}^2).$

(7) 分布函数、概率密度、高斯分布. 《概率论与数理统计》Sec5

先证明 $Z=X+Y$ 仍然服从高斯分布. $P_{26}\sim P_{27}$

且若 X,Y 相互独立, $X\sim N(\mu_1,\sigma_1^2)$, $Y\sim N(\mu_2,\sigma_2^2)$

则 $Z=X+Y$ 有 $Z\sim N(\mu_1+\mu_2,\sigma_1^2+\sigma_2^2)$. 推广到有限个独立正态随机变量仍然成立.

故 $Z\sim N(-1,5)$.

线性代数部分.

(1) 通过矩阵初等变换后. 可保留的非零行数目.

$$\begin{bmatrix}1&2&1\\1&0&3\\1&1&2\end{bmatrix}\underset{r_3-r_1}{\overset{r_2-r_1}{\sim}}\begin{bmatrix}1&2&1\\0&-2&2\\0&-1&1\end{bmatrix}\underset{r_3-r_2}{\overset{r_2\div 2}{\sim}}\begin{bmatrix}1&2&1\\0&-1&1\\0&0&0\end{bmatrix}$$

即 $\mathrm{rank}(A)=2$.

(2) A 可逆 $\Leftrightarrow$ 存在可逆矩阵 P, 使得 $PA=E$,

由于 $PA=B \Leftrightarrow \begin{cases}PA=B\\PE=P\end{cases} \Leftrightarrow P(A,E)=(B,P) \Leftrightarrow (A,E)\sim(B,P)$

令 $B=F$ 为 A 的行最简形. 那么若 $F=E$, 则 A 可逆

并由 $PA=E$ 知 $P=A^{-1}$. 另外 $A^{-1}=\frac{A^*}{|A|}$

$$\left(\begin{array}{ccc|ccc}0&2&4&1&0&0\\2&4&2&0&1&0\\3&3&1&0&0&1\end{array}\right)\left(\begin{array}{ccc|ccc}1&2&1&0&1/2&0\\0&1&2&1/2&0&0\\0&-3&-2&0&-3/2&1\end{array}\right)\left(\begin{array}{ccc|ccc}1&0&-3&-1&1/2&0\\0&1&2&1/2&0&0\\0&0&4&3/2&-3/2&1\end{array}\right)$$

$$\left(\begin{array}{ccc|ccc}1&0&0&1/8&-\frac{5}{8}&\frac{3}{4}\\0&1&0&-1/4&\frac{3}{4}&-1/2\\0&0&1&3/8&-3/8&1/4\end{array}\right)$$

(5) A为n阶矩阵，如果数λ与n维非零向量x，使得$Ax=\lambda x$成立，那么，λ为矩阵A的特征值，非零向量x为A的对应于特征值λ的特征向量。其也可写成$(A-\lambda E)x=0$，这是n个未知数n个方程的齐次线性方程组，有非零解的充分必要条件是系数行列式$|A-\lambda E|=0$，矩阵A的特征方程。其在复数范围内恒有解，解个数为方程的次数，因此n阶矩阵A在复数范围内有n个特征值。

且有 $\lambda_1+\lambda_2+\cdots+\lambda_n=a_{11}+a_{22}+\cdots+a_{nn}$

$\lambda_1 * \lambda_2 \cdots \lambda_n = |A|$.

解：特征多项式为

$$|A-\lambda E| = \begin{vmatrix} 3-\lambda & 1 & 1 \\ 2 & 4-\lambda & 2 \\ -1 & -1 & 1-\lambda \end{vmatrix} = \begin{vmatrix} \lambda-2 & 0 & \lambda-2 \\ 2 & 4-\lambda & 2 \\ 1 & 1 & \lambda-1 \end{vmatrix}$$

$$= (\lambda-2)^2(4-\lambda)=0$$

$\therefore \lambda_1=2 \quad \lambda_2=2 \quad \lambda_3=4$

$\lambda_1=2=\lambda_2$时，对应的特征向量为 $p=-c_1\begin{bmatrix}1\\0\\0\end{bmatrix}-c_2\begin{bmatrix}0\\1\\0\end{bmatrix}$

$\lambda_3=4$时，特征向量为 $p=c_1\begin{bmatrix}1\\2\\-1\end{bmatrix}$

(6) 证明：奇异值分解.

a) $MM^+M=U\Sigma V^T V\Sigma^+U^T U\Sigma V^T$

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

$=U\Sigma V^T=M$

b) $M^+M=V\Sigma^+U^TU\Sigma V^T=VV^T=I$

故 $M^+=M^{-1}$

(5) 半正定矩阵.

a) 证明: $x^TZZ^Tx = x^TZ(x^TZ)^T \geq 0$

b) 证明: $x^TAx = x^T\lambda x = \lambda x^Tx > 0$ 且 $x^Tx > 0$ 故 $\lambda > 0$

即A的各特征值都大于0

(6) 1) 向量u与x同向时有 u^Tx 取到最大为 $\|u^T\|\|x\| = \|x\|$

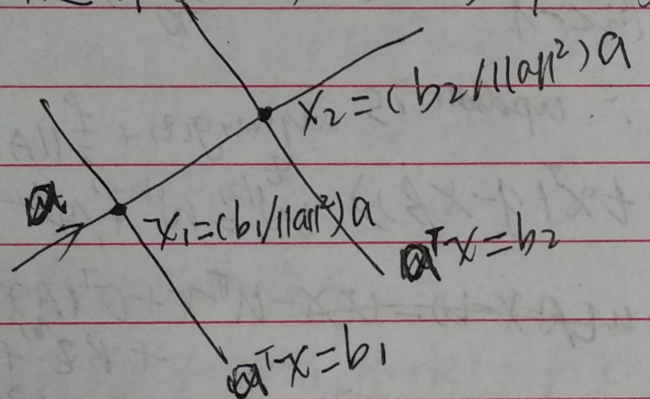
2) 向量u与x反向时有 u^Tx 取到最小为 $-\|u^T\|\|x\| = -\|x\|$

3) u与x垂直时有 $|u^Tx|$ 取到最小为 0

(7) 两个超平面之间的距离为 $|b_1-b_2|/\|a\|_2$ 即为5

两个超平面之间的距离即为两点 x_1, x_2 之间的距离, 即为

$\|x_1-x_2\|_2 = |b_1-b_2|/\|a\|_2$



微积分.

(1) 解: $\frac{df(x)}{dx} = \frac{-2e^{-2x}}{1+e^{-2x}}$　$\frac{\partial g(x,y)}{\partial y} = 2e^{2y} + 6xye^{3xy^2}$

(2) 解: $\frac{\partial f}{\partial v} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial v} = y(-\sin(u+v)) + x(-\cos(u-v))$

$= \cos 2v$

(3) 解: 令 $x-3=t$ 则原式 $= \int_2^7 \frac{2}{t}dt = 2(\ln x)\big|_2^7 = 2(\ln 7 - \ln 2)$

(4) ① $\nabla E = (\frac{\partial E}{\partial u}, \frac{\partial E}{\partial v}) = (2(ue^{v}-2ve^{-u})\cdot(e^{v}+2ve^{-u}),$
$2(ue^{v}-2ve^{-u})\cdot(ue^{v}-2e^{-u}))$

$\therefore \nabla E(1,1) = (2(e^{2}-4e^{-2}), 2(e-2e^{-1})^{2})$.

② $\nabla^2 E = \begin{bmatrix} \frac{\partial^2 E}{\partial u^2}, & \frac{\partial^2 E}{\partial u \partial v} \\ \frac{\partial^2 E}{\partial v \partial u}, & \frac{\partial^2 E}{\partial v^2} \end{bmatrix} = \begin{bmatrix} 2(e^{v}+2ve^{-u})^{2} - 4ve^{-u}(ue^{v}-2ve^{-u}), & 2(ue^{v}-2e^{-u})(e^{v}+2ve^{-u}) + 2(ue^{v}-2ve^{-u})(e^{v}+2e^{-u}) \\ 2(e^{v}+2ve^{-u})(ue^{v}-2e^{-u}) + 2(ue^{v}-2ve^{-u})(e^{v}+2e^{-u}), & 2(ue^{v}-2e^{-u})^{2} + 2(ue^{v}-2ve^{-u})ue^{v} \end{bmatrix}$

$\therefore \nabla^2 E(1,1) = \begin{bmatrix} 2e^{2}+16e+4, & 4e^{2}-16e^{-2} \\ 4e^{2}-16e^{-2}, & 4e^{2}+8e^{-2}-12 \end{bmatrix}$

5) 二元函数的 Talor 展开.

公式为 $f(x_0+h, y_0+k) = f(x_0,y_0) + (h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}) f(x_0,y_0)$
$+ \frac{1}{2!}(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y})^{2} f(x_0,y_0)$
$+ \frac{1}{3!}(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y})^{3} f(x_0,y_0) + \cdots$
$+ \frac{1}{n!}(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y})^{n} f(x_0,y_0) +$

其中 $(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}) f(x_0,y_0) = hf_x(x_0,y_0) + kf_y(x_0,y_0)$

$(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y})^{2} f(x_0,y_0) = h^{2}f_{xx}(x_0,y_0) + k^{2}f_{yy}(x_0,y_0)$
$+ 2hkf_{xy}(x_0,y_0)$

原式 $= E(1,1) + hE_\mu(1,1) + kE_v(1,1) + h^2E_{\mu\mu}(1,1) + k^2E_{vv}(1,1)$
$+ 2hkE_{\mu v}(1,1) + R_2$（泰勒展开）

代入得，其中 $h=\mu, k=v$

$$E(\mu,v) = (e-2e^{-1})^2 + 2(e^2-4e^{-2})\mu + 2(e-2e^{-1})^2 v + (2e^2+16e^{-2}+4)\mu^2 + (4e^2+8e^{-2}-12)v^2 + (8e^2-32e^{-2})\mu v + R_2$$

(6) 对于连续可导函数求极值，可直接使用求导。

设 $f(\alpha) = Ae^{\alpha} + Be^{-2\alpha}$ $\min f(\alpha)$

$$\frac{\partial f(\alpha)}{\partial \alpha} = Ae^{\alpha} + B\cdot(-2)e^{-2\alpha} \quad \therefore \frac{\partial f(\alpha)}{\partial \alpha} = 0 \text{ 得 } \alpha = \frac{1}{3}(\ln 2B - \ln A)$$

7) 向量(矩阵求导)

$$\nabla E(w) = \frac{\partial E}{\partial w} = Aw + b \quad \nabla^2 E(w) = A$$

相关公式 $\frac{\partial AX}{\partial X} = A^T$ $\frac{\partial AX^T}{\partial X} = \frac{\partial X^T A}{\partial X} = A$ $\frac{\partial X^T A X}{\partial X} = 2AX$

$$\frac{\partial A^T X B}{\partial X} = AB^T \quad \frac{\partial A^T X^T B}{\partial X} = BA^T$$

证明可依据定义来证明。

$$\frac{d(\mu v^T)}{dx} = \left(\frac{d\mu}{dx}\right)v^T + \mu\left(\frac{dv^T}{dx}\right) \quad \frac{d(\mu^T v)}{dx} = \left(\frac{d\mu^T}{dx}\right)v + \mu^T\left(\frac{dv^T}{dx}\right)$$

参考：http://voidcn.com/article/p-ponrkmelt-xd.html

(8) 证明：令 $\nabla E(w)=0$ 得 $w=-A^{-1}b$（A为正定）

即得证.

(9) 解：$L(w,\mu)=\frac{1}{2}(w_1^2+2w_2^2+3w_3^2)-\mu(w_1+w_2+w_3-11)$

令 $\frac{\partial L}{\partial w_1}=0 \Rightarrow w_1=\mu$ 同理得 $w_2=\frac{\mu}{2}$ $w_3=\frac{\mu}{3}$

又 $w_1+w_2+w_3=11$ 故 $\mu=6$ 此时 $E(w)=33$