1.1 证明：① $N=1$ 时 $k=0$ 或 $k=1$，$C(N,k)=C(1,0)=C(1,1)=1$ 结论成立

② 假设 $N=t$ 时结论成立，则

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

③ $N=t+1$ 时

$$\begin{aligned}C(t+1,k)&=C(t,k)+C(t,k-1)\\&=\frac{t!}{k!(t-k)!}+\frac{t!}{(k-1)!(t-(k-1))!}\quad(\text{应用②})\\&=\frac{(t+1)!}{k!(t+1-k)!}\left(\frac{t+1-k}{t+1}+\frac{k}{t+1}\right)\\&=\frac{(t+1)!}{k!(t+1-k)!}\quad(0\le k\le t)\end{aligned}$$

$$C(t+1,k)=1\quad(k=t+1)$$

故 $N=t+1$ 时结论亦成立，由①②③知，对 $N\ge1$，结论均成立

1.2 解：① $X\sim B(10,\frac{1}{2})$

$$P(X=4)=C_{10}^4(\tfrac{1}{2})^4(1-\tfrac{1}{2})^6=\frac{C_{10}^4}{2^{10}}$$

② 52张牌任意抽5张的组合数为 C_{52}^5

型如 XXXYY，先选取X、Y，A_{13}^2

再在4张X中选取3张，C_4^3，在4张Y中选取2张，C_4^2

故 $p=\dfrac{A_{13}^2C_4^3C_4^2}{C_{52}^5}$

1.3 解：$P(A|B)=\dfrac{P(AB)}{P(B)}=\dfrac{(\frac{1}{2})^3}{1-(\frac{1}{2})^3}=\dfrac{1}{7}$

1.4 解：$P(X<0\,|\,|X|=1)=P(X=-1\,|\,|X|=1)=P(bit=1\,|\,|X|=1)$

$$=\frac{P(bit=1,|X|=1)}{P(|X|=1)}=\frac{P(bit=1,X=-1)}{P(bit=1,X=-1)+P(bit=0,X=1)}$$

$$=\frac{\frac{1}{2}\times\frac{1}{4}}{\frac{1}{2}\times\frac{1}{4}+\frac{1}{2}\times\frac{1}{8}}=\frac{2}{3}$$

1.5 解：

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

$\min P(A\cap B)=\min[0,P(A)+P(B)-1]=0$

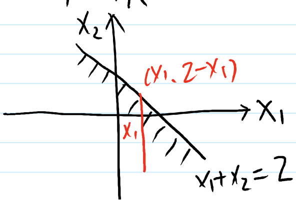
$\max P(A\cup B)=\min[P(A)+P(B),1]=0.7$

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

1.6 解：

$$\begin{aligned}\sigma_x^2&=\frac{1}{N-1}\sum_{n=1}^N(x_n-\bar{x})^2\\&=\frac{1}{N-1}\sum_{n=1}^N(x_n^2-2x_n\bar{x}+\bar{x}^2)\\&=\frac{1}{N-1}\left(\sum_{n=1}^Nx_n^2-2\bar{x}\sum_{n=1}^Nx_n+N\bar{x}^2\right)\\&=\frac{1}{N-1}\left(\sum_{n=1}^Nx_n^2-2N\bar{x}^2+N\bar{x}^2\right)\\&=\frac{N}{N-1}\left(\frac{1}{N}\sum_{n=1}^Nx_n^2-\bar{x}^2\right)\end{aligned}$$

1.7. 解：

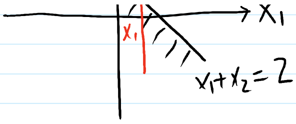


$$P(Z\le z)=P(X_1+X_2\le z)=\iint_B f(x_1,x_2)\,dx_1dx_2$$

$$=\int_{-\infty}^{\infty}\left(\int_{-\infty}^{z-x_1}f(x_1,x_2)\,dx_2\right)dx_1$$

Z的密度函数为上式求导

$$l(z)=\int_{-\infty}^{\infty}f(x_1,z-x_1)\,dx_1$$



z的密度函数为上式求导
$$l(z)=\int_{-\infty}^{\infty} f(x_1, z-x_1)\,dx_1$$
$$=\int_{-\infty}^{\infty} f(x, z-x)\,dx$$
$$=\int_{-\infty}^{\infty} f_1(x) f_2(z-x)\,dx \quad (x_1, x_2 独立)$$
$x_1 \sim N(\mu_1, \sigma_1^2)$ $\quad x_2 \sim N(\mu, \sigma_2^2)$ 代入上式
$$l(z)=\frac{1}{2\pi\sigma_1\sigma_2}\int_{-\infty}^{\infty} \exp\left[-\frac{1}{2}\left(\frac{(x-\mu_1)^2}{\sigma_1^2}+\frac{(z-x-\mu_2)^2}{\sigma_2^2}\right)\right]dx$$
$$=\left(\sqrt{2\pi(\sigma_1^2+\sigma_2^2)}\right)^{-1}\exp\left[-\frac{1}{2}(z-\mu_1-\mu_2)^2/(\sigma_1^2+\sigma_2^2)\right]$$
故 $z \sim N(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2)$

对本题 $z \sim N(-1, 5)$

2.1 解：矩阵的秩描述的是各行/列向量的线性无关性，通过高斯消元法得到非全零行数
$\mathrm{rank}(A)=2$

2.2 解：矩阵的逆可通过 $(A, I) \xrightarrow{行初等变换} (I, A^{-1})$ 得到
$$A^{-1}=\begin{pmatrix} 1/8, & -5/8, & 3/4 \\ -1/4 & 3/4, & -1/2 \\ 3/8 & -7/8 & 1/4 \end{pmatrix}$$

2.3 解：
$$\det(\lambda I-A)=\begin{vmatrix} \lambda-3 & -1 & -1 \\ -2 & \lambda-4 & -2 \\ 1 & 1 & \lambda-1 \end{vmatrix}=(\lambda-4)(\lambda-2)^2=0$$
故特征值 $\lambda_1=4$，$\lambda_2=2$（二重特征值）

求解方程 $(\lambda I-A)x=0$

$\lambda=\lambda_1=4$ 时 $\quad p_1=(1, 2, -1)^T$

$\lambda=\lambda_2=2$ 时 $\quad p_2=(1, 0, -1)$ $\quad p_3=(1, -1, 0)$

2.4 证明：(a) MM^+M
$$=U\Sigma V^T V\Sigma^+ U^T U\Sigma V^T$$
$$=U\Sigma I\Sigma^+ I\Sigma V^T=U\Sigma\Sigma^+\Sigma V^T$$
其中 Σ、Σ^+ 均为对角阵，故
$$\Sigma^+\Sigma[i][j]=\begin{cases}\frac{1}{\Sigma[i][j]}\cdot\Sigma[i][j]=1 & \text{when } \Sigma[i][j]\neq 0 \quad (对角线上) \\ 0 & \text{when } \Sigma[i][j]=0 \quad (对角线外)\end{cases}$$

所以 $\Sigma^+\Sigma=I$， $MM^+M=U\Sigma V^T$

(b) $M^+M=V\Sigma^+U^TU\Sigma V^T=V\Sigma^+\Sigma V^T=VV^T=I$ 故 $M^+=M^{-1}$

2.5 证明：(a) $x^T zz^T x=(x^Tz)(x^Tz)^T \geq 0$

(b) $x^TAx=x^T\lambda x=\lambda x^Tx>0$
$x^Tx>0$ $\Rightarrow \lambda>0$

2.6、解：① u 与 x 方向相同时，有 $u^Tx_{max}=\|x\|$

2.6、解：① u 与 x 方向相同时，有 $u^T x_{max} = \|x\|$

② u 与 x 方向相反时，有 $u^T x_{min} = -\|x\|$

③ u 与 x 方向垂直时，有 $|u^T x|_{min} = 0$

2.7、解：$d = \frac{\Delta b}{|w|} = 5$

3.1 解：$\frac{df(x)}{dx} = \frac{e^{-2x}\cdot(-2)}{1+e^{-2x}} = -\frac{2}{e^{2x}+1}$

$$\frac{\partial g(x,y)}{\partial y} = 2e^{2y} + 6xye^{3xy^2}$$

3.2 解：$\frac{\partial f}{\partial v} = \frac{\partial f}{\partial x}\cdot\frac{\partial x}{\partial v} + \frac{\partial f}{\partial y}\cdot\frac{\partial y}{\partial v}$

$= y(-\sin(u+v)) + x(-\cos(u-v))$

$= -\sin(u-v)\sin(u+v) - \cos(u+v)\cos(u-v) = -\cos 2v$

3.3 解：$\int_5^{10}\frac{2}{x-3}\,dx = 2\int_2^7\frac{1}{x-3}\,d(x-3) = 2\int_2^7\frac{1}{x}\,dx$

$= 2\ln x\Big|_2^7 = 2\ln\frac{7}{2}$

3.4 解：① $\nabla E = \left(\frac{\partial E}{\partial u}, \frac{\partial E}{\partial v}\right)$

$= \left(2(ue^v-2ve^{-u})(e^v+2ve^{-u}),\ 2(ue^v-2ve^{-u})(ue^v-2e^{-u})\right)$

when $u=1,\ v=1$，$\nabla E = (2(e^2-4e^{-2}),\ 2(e-2e^{-1})^2)$

② $\nabla^2 E = \begin{pmatrix} \frac{\partial^2 E}{\partial u^2} & \frac{\partial^2 E}{\partial u\partial v} \\ \frac{\partial^2 E}{\partial v\partial u} & \frac{\partial^2 E}{\partial v^2} \end{pmatrix}$

$$= \begin{pmatrix} 2(e^v+2ve^{-u})^2 - 4ve^{-u}(ue^v-2ve^{-u}), & 2(ue^v-2e^{-u})(e^v+2ve^{-u}) + 2(ue^v-2ve^{-u})(e^v+2e^{-u}) \\ 2(e^v+2ve^{-u})(ue^v-2e^{-u}) + 2(ue^v-2ve^{-u})(e^v+2e^{-u}) & 2(ue^v-2e^{-u})^2 + 2(ue^v-2ve^{-u})ue^v \end{pmatrix}$$

when $u=1,\ v=1$ $\nabla^2 E = \begin{pmatrix} 2e^2+16e^{-2}+4, & 4e^2-16e^{-2} \\ 4e^2-16e^{-2} & 4e^2+8e^{-2}-12 \end{pmatrix}$

3.5. 解：$E(1+h, 1+k) = E(1,1) + (h\frac{\partial}{\partial u}+k\frac{\partial}{\partial v})E(1,1) + (h\frac{\partial}{\partial u}+k\frac{\partial}{\partial v})^2 E(1,1) + R_2$

$= E(1,1) + hE_u(1,1) + kE_v(1,1) + h^2E_{uu}(1,1) + 2hkE_{uv}(1,1) + k^2E_{vv}(1,1) + R_2$

利用3.4中的结果代入上式，并使用 u, v 代替 h, k

得 $E(u,v) = (e-2e^{-1})^2 + 2(e^2-4e^{-2})u + 2(e-2e^{-1})^2 v$

$+ (2e^2+16e^{-2}+4)u^2 + (8e^2-32e^{-2})uv + (4e^2+8e^{-2}-12)v^2 + R_2$

3.6 解：$\frac{df}{d\alpha} = Ae^{\alpha} - 2Be^{-2\alpha} = 0 \Rightarrow \alpha = \frac{1}{3}\ln\frac{2B}{A}$

此时 $f(x)=3\sqrt[3]{\frac{A^2B}{4}}$

3.7 证：① $E=\frac{1}{2}w^TAw+b^Tw$

$$=\frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}w_ia_{ij}w_j+\sum_{i=1}^{d}b_iw_i$$

则 $\frac{\partial E}{\partial w_i}=a_{ii}w_i+\frac{1}{2}\sum_{j\neq i}a_{ij}w_j+\frac{1}{2}\sum_{j\neq i}a_{ji}w_j$

由于 A 为对称阵，故 $a_{ij}=a_{ji}$，$\frac{\partial E}{\partial w_i}=a_{ii}w_i+\sum_{j\neq i}a_{ij}w_j+b_i=\sum_{j=1}^{d}a_{ij}w_j+b_i$

故 $\nabla E=\begin{pmatrix}\sum_{j=1}^{d}a_{1j}w_j+b_1\\ \sum_{j=1}^{d}a_{2j}w_j+b_2\\ \vdots\\ \sum_{j=1}^{d}a_{dj}w_j+b_d\end{pmatrix}=Aw+b$

② $\frac{\partial E^2}{\partial w_i\partial w_j}=a_{ij}$ 所以 $\nabla^2E=A$

3.8 解：$\nabla E(w)=0\Rightarrow w=-A^{-1}b$ 为驻点

又 $\nabla^2E(w)=A$ 为正定矩阵，故该驻点为极小值点

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

$\frac{\partial L}{\partial w_1}=0\Rightarrow w_1=\lambda$ $\quad\frac{\partial L}{\partial w_2}=0\Rightarrow 2w_2=\lambda$ $\quad\frac{\partial L}{\partial w_3}\Rightarrow 3w_3=\lambda$

又 $w_1+w_2+w_3=11$ 故 $\lambda=6$ $w_1=6$ $w_2=3$ $w_3=2$

此时 $\min E(w)=33$

3.10、证明： 假设 $E(w)$ 取到最小值，且 $Aw+b=0$

此时若 $\nabla E(w)+A^T\lambda\neq0$，则 $\nabla E(w)$ 与 A 的行空间不平行，

$\nabla E(w)$ 投影到 A 的行空间的残差 μ 不为零

即 $\mu=(I_d-H)\nabla E(w)$ 不为零，其中 H 为投影矩阵 $A^T(AA^T)^{-1}A$

此时，令 $w'=w-\eta\mu$ $\quad Aw'+b=A(-\eta\mu)=0$ 满足条件

则对于凸函数

$$E(w')-E(w)=\nabla E^T(w)(w'-w)\geq0$$

$$\Rightarrow\nabla E^T(w)(-\eta\mu)\geq0$$

$$\Rightarrow\nabla E^T(w)(I-H)\nabla E(w)\leq0\quad ①$$

向量 $\nabla E(w)$ 与其残差 $(I-H)\nabla E(w)$ 的夹角小于 $90°$

故 $\nabla E^T(w)(I-H)\nabla E(w)>0$，与①矛盾