1 Probabity and Statistics

1.1 (combinatorics)

Proof: To prove the equation $C_n^k = \frac{n!}{k!(n-k)!}$, **mathematical induction can be applied there.**

S1 when $n = 0$, k can only value as 0 or 1.

In those cases,$C_n^k = 1 = \frac{n!}{k!(n-k)!}$ holds

S2 Assume that it holds when $n = t$, as $C_n^k = \frac{n!}{k!(n-k)!}$.
Then for $n = t+1$

$$C_{t+1}^k = C_t^k + C_t^{k-1} = \frac{t!}{k!(t-k)} + \frac{t!}{(k-1)!(t-k+1)!} = \frac{t!}{k!(t-k+1}(t-k+1+k) = \frac{(t+1)!}{k!(t+1-k)!}$$

hence the equation holds when $n = t+1$

So

$$C_n^k = \frac{n!}{k!(n-k)!}$$

for $n \geq 1$ and $0 \leq k \leq n$ □

1.2 (counting)

Let random variable X represent the number of the heads, and $X \sim B(10, \frac{1}{2})$, then

$$P(X = 4) = C_{10}^4(\frac{1}{2})^4(\frac{1}{2})^6 = \frac{105}{512}$$

Let event A be the event of getting a full house(XXXYY) from a deck of 52 cards, then

$$P(A) = \frac{A_4^2 \cdot C_{13}^3 \cdot C_{13}^2}{C_{52}^5} = \frac{429}{4165}$$

1.3 (conditional probability)

Let event A represent the event of all three tosses resulting in head, while event B represent the event of at least one of the tosses resulting in head. Then

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

1.4 (Bayes theorem)

$$\begin{aligned} P(X<0) &= \frac{P(bit=1, X=-1)}{P(bit=0, X=1) + P(bit=1, X=-1)} \\ &= \frac{\frac{1}{2}\times\frac{1}{4}}{\frac{1}{2}\times\frac{1}{8}+\frac{1}{2}\times\frac{1}{4}} = \frac{2}{3} \end{aligned}$$

1.5 (union/intersection)

As applications of the principle of Inclusion-exclusion,

$$P(A \cap B) = P(A) + P(B) - P(A \cup B)$$

and

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Hence

$$P(A \cap B)_{\max} = \min[P(A), P(B)] = P(A) = 0.3$$
$$P(A \cap B)_{\min} = \max[0, P(A) + P(B) - 1] = 0$$
$$P(A \cup B)_{\max} = \min[P(A) + P(B), 1] = 0.7$$
$$P(A \cup B)_{\min} = \max[P(A), P(B)] = P(B) = 0.4$$

2 Linear Algebra

2.1 (rank)

$$r\begin{pmatrix} 1 & 2 & 1 \\ 1 & 0 & 3 \\ 1 & 1 & 2 \end{pmatrix} = 2$$

2.2 (inverse)

$$\begin{pmatrix} 0 & 2 & 4 \\ 2 & 4 & 2 \\ 3 & 3 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 0 & 2 & 3 \\ 2 & 4 & 3 \\ 4 & 2 & 1 \end{pmatrix}$$

2.3 (eigenvalue/eignevectors)

Table 1: Eigenvalues and their respective eigenvectors

α_1=4	$p_1 = [\ -1 \quad -2 \quad 1\ ]^T$
α_2=2	$p_2 = [\ -1 \quad 0 \quad 1\ ]^T$
α_3=2	$p_3 = [\ -1 \quad 1 \quad 0\ ]^T$

2.4 (singular value decomposition)

(a) For Σ is a rectangular diagonal matrix, so is $\Sigma^\dagger$

$$\Sigma\Sigma^\dagger{}_{[i][j]} = \sum_{k=1}^{m} \Sigma_{[i][k]} \cdot \Sigma^\dagger{}_{[k][j]}$$

Apparently, $\Sigma\Sigma^\dagger$ will be an $n \times n$ identity matrix. Since U and V are both unitary matrix,

$$\begin{cases} UU^T = U^T U = I \\ VV^T = V^T V = I \end{cases}$$

$$\begin{aligned} MM^\dagger M &= (U\Sigma V^T)(V\Sigma^\dagger U^T)(U\Sigma V^T) \\ &= U\Sigma\Sigma^\dagger\Sigma V^T \\ &= U\Sigma V^T \\ &= M \end{aligned}$$

(b)

$$MM^{\dagger} = (U\Sigma V^T)(V\Sigma^{\dagger}U^T) = U\Sigma\Sigma^{\dagger}U^T = UU^T = I$$

Analogously,

$$M^{\dagger}M = I$$

So,

$$M^{-1} = M^{\dagger}$$

2.5 (PD/PSD)

(a)

$$\mathbf{x}^T ZZ^T \mathbf{x} = (\mathbf{x}Z)(\mathbf{x}Z)^T \geq 0$$

So ZZ^T is a postive semi-defined matrix.

(b) Since A is a real symmetric matrix, the eigenvectors of A (presented as v_i, and their respectful eigenvalues are defined as λ_i) can be orthogonal and unitary.
Let $\mathbf{x} = \sum k_i v_i$ and $\mathbf{x}^T = \sum k_i v_i^T$

$$\begin{aligned}&\mathbf{x}^T A\mathbf{x} = (\sum k_i v_i^T)A(\sum k_i v_i)\\ &= (\sum k_i v_i^T)(\sum k_i \lambda_i v_i)\\ &= \sum \lambda_i k_i^2 > 0\end{aligned}$$

Hence, A is postive-defined.

2.6 (inner product)

When $\mathbf{x}$ and $\mathbf{u}$ have the same direction, $\mathbf{x^T u}$ reaches its maximum $\|\mathbf{x}\|$.
When $\mathbf{x}$ and $\mathbf{u}$ have the opposite direction, $\mathbf{x^T u}$ reaches its minimum $\|\mathbf{x}\|$.
When $\mathbf{x}$ and $\mathbf{u}$ are orthogonal, $\left|\mathbf{x^T u}\right|$ reaches its minimun 0.

3 Calculus

3.1 (differential and partial differential)

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

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

3.2 (chain rule)

$$\frac{\partial f}{\partial v} = -y\sin v - x\cos v$$

3.3 (gradient and Hessian)

When $u = 1$ and $v = 1$, the gradient of E

$$\begin{aligned}\nabla E &= -2(e^v + 2ve^{-u})(ue^v - 2ve^{-u})\mathbf{i} + 2(ue^v - 2e^{-u})(ue^v - 2ve^{-u})\mathbf{j}\\ &= 2(4e^{-2} - e^2)\mathbf{i} + 2(2e^{-1} - e)^2\mathbf{j}\end{aligned}$$

Meanwhile, its Hessian matrix $\nabla^2 E =$

$$\begin{pmatrix} 2e^2 - 12 + 16e^{-2} & 0 \\ 0 & 4e^2 - 12 + 8e^{-2} \end{pmatrix}$$

3.4 (Taylor's expansion)

$$
\begin{aligned}
E = & 2(4e^{-2} - e^{2})(u-1) + 2(2e^{-1} - e)^{2}(v-1) \\
& +2e^{2} - 6 + 8e^{-2})(u-1)^{2} + 4(e^{2} - 3 + 4e^{-2})(v-1)^{2} \\
& +o
\end{aligned}
$$

3.5 (optiminzation)

From the inequality of arithmetic and geometric means

$$
Ae^{\alpha} + Be^{-2\alpha} = \frac{1}{2}Ae^{\alpha} + \frac{1}{2}Ae^{\alpha} + Be^{-2\alpha} \geq 3\sqrt[3]{\frac{A}{2}e^{\alpha} \cdot \frac{A}{2}e^{\alpha} \cdot Be^{-2\alpha}}
$$

$$
\geq 3\sqrt[3]{\frac{A^{2}B}{4}}
$$

with equality if and only if $\frac{A}{2}e^{\alpha} = Be^{-2\alpha}$, at that time, $\alpha = \frac{1}{3}(\ln 2B - \mathrm{lnA})$.

3.6 (vector calculus)

$$
\frac{dE}{d\mathbf{w}} = \frac{1}{2}(A\mathbf{w} + A\mathbf{w}) + \mathbf{b} = A\mathbf{w} + \mathbf{b}
$$

Then,

$$
\nabla \mathrm{E}(\mathbf{w}) = \mathrm{A}\mathbf{w}\text{+b}
$$

$$
\frac{d^{2}E}{d\mathbf{w}^{2}} = A
$$

So,

$$
\nabla^{2}\mathrm{E}(\mathbf{w}) = \mathrm{A}
$$