模式识别与计算机视觉：第四次作业

人工智能学院 221300066 季千焜

2025 年 5 月 21 日

1 习题一

(a)

对于每个数据点 x_i，其局部线性嵌入（LLE）的权重 w_{ij} 通过最小化重构误差并满足约束条件得到。具体步骤如下：

1. 确定最近邻：找到 x_i 的 k 个最近邻点，记作 $\{x_j \mid j \in N(i)\}$，其中 $N(i)$ 为邻居索引集合。

2. 构建局部协方差矩阵：定义矩阵 $C_i \in \mathbb{R}^{k\times k}$，其元素为：

$$C_{jk} = (x_j - x_i)^\top (x_k - x_i)$$

其中 $j, k \in N(i)$。

3. 求解权重：在约束条件 $\sum_{j\in N(i)} w_{ij} = 1$ 下，最小化重构误差 $e_i = \left\| x_i - \sum_{j\in N(i)} w_{ij} x_j \right\|^2$。该优化问题的解为：

$$w_i = \frac{C_i^{-1}\mathbf{1}}{\mathbf{1}^\top C_i^{-1}\mathbf{1}}$$

其中 $\mathbf{1}$ 为全 1 向量，分母为归一化因子以保证权重和为 1。

目标函数和约束条件为：

$$\min_{w_i} \left\| \sum_{j\in N(i)} w_{ij}(x_j - x_i) \right\|^2 \quad \text{s.t.} \quad \sum_{j\in N(i)} w_{ij} = 1$$

将目标函数展开为二次型：

$$e_i = w_i^\top C_i w_i$$

引入拉格朗日乘数 λ，构造拉格朗日函数：

$$\mathcal{L} = w_i^\top C_i w_i + \lambda \left(\mathbf{1}^\top w_i - 1 \right)$$

对 w_i 求导并令导数为零：

$$\frac{\partial \mathcal{L}}{\partial w_i} = 2C_i w_i + \lambda \mathbf{1} = 0 \implies w_i = -\frac{\lambda}{2} C_i^{-1} \mathbf{1}$$

代入约束条件 $\mathbf{1}^\top w_i = 1$：

$$-\frac{\lambda}{2} \mathbf{1}^\top C_i^{-1} \mathbf{1} = 1 \implies \lambda = -\frac{2}{\mathbf{1}^\top C_i^{-1} \mathbf{1}}$$

最终得到权重：

$$w_i = \frac{C_i^{-1} \mathbf{1}}{\mathbf{1}^\top C_i^{-1} \mathbf{1}}$$

该解保证了重构误差最小且权重归一，满足局部线性关系的假设。

(b)

i. 旋转不变性

设旋转后的数据点为 $x_i' = Qx_i$，其中 Q 是正交矩阵（$QQ^\top = I$）。对于邻域点 x_j 和 x_i，其差异为：

$$x_j' - x_i' = Qx_j - Qx_i = Q(x_j - x_i)$$

局部协方差矩阵 C_i' 的元素为：

$$C_{jk}' = (x_j' - x_i')^\top (x_k' - x_i') = (Q(x_j - x_i))^\top Q(x_k - x_i) = (x_j - x_i)^\top Q^\top Q(x_k - x_i) = (x_j - x_i)^\top (x_k - x_i) = C_{jk}$$

因此 $C_i' = C_i$，权重公式为：

$$w_i' = \frac{(C_i')^{-1} \mathbf{1}}{\mathbf{1}^\top (C_i')^{-1} \mathbf{1}} = \frac{C_i^{-1} \mathbf{1}}{\mathbf{1}^\top C_i^{-1} \mathbf{1}} = w_i$$

故旋转后权重不变。

ii. 平移不变性

设平移后的数据点为 $x_i' = x_i + t$。邻域差异为：

$$x_j' - x_i' = (x_j + t) - (x_i + t) = x_j - x_i$$

局部协方差矩阵 $C'_i = C_i$，因此权重 $w'_i = w_i$，平移不影响权重。

iii. 缩放不变性

设缩放后的数据点为 $x'_i = sx_i$，其中 $s \neq 0$。邻域差异为：

$$x'_j - x'_i = sx_j - sx_i = s(x_j - x_i)$$

局部协方差矩阵 C'_i 的元素为：

$$C'_{jk} = (s(x_j - x_i))^\top (s(x_k - x_i)) = s^2(x_j - x_i)^\top (x_k - x_i) = s^2 C_{jk}$$

因此 $C'_i = s^2 C_i$，其逆矩阵为 $(C'_i)^{-1} = \frac{1}{s^2} C_i^{-1}$。权重计算为：

$$w'_i = \frac{(C'_i)^{-1}\mathbf{1}}{\mathbf{1}^\top (C'_i)^{-1}\mathbf{1}} = \frac{\frac{1}{s^2} C_i^{-1}\mathbf{1}}{\frac{1}{s^2}\mathbf{1}^\top C_i^{-1}\mathbf{1}} = \frac{C_i^{-1}\mathbf{1}}{\mathbf{1}^\top C_i^{-1}\mathbf{1}} = w_i$$

故缩放后权重不变。

(c)

局部几何保持：

优化目标 $\min \sum_{i=1}^n \left\| y_i - \sum_{j=1}^n w_{ij} y_j \right\|^2$ 要求低维表示 y_i 的局部线性重构误差最小，其中权重 w_{ij} 来自原始高维空间的局部结构。由于权重 w_{ij} 在低维空间中被保留，邻域点之间的线性组合关系得以维持，从而保持了局部几何结构。

约束条件的作用：

1. 零均值约束 $\sum_{i=1}^n y_i = 0$：消除平移自由度，确保低维表示以原点为中心，避免因整体平移导致的解不唯一。
2. 单位协方差约束 $\sum_{i=1}^n y_i y_i^\top = I$：消除缩放和旋转自由度。该约束强制低维表示的协方差矩阵为单位矩阵，即各维度方差为 1 且互不相关，从而固定缩放和旋转方向，保证解的唯一性。

自由度消除与剩余影响：

- 平移：通过零均值约束完全消除。
- 缩放：通过单位协方差约束完全消除。
- 旋转：单位协方差约束隐含了各向同性条件，允许整体旋转，但旋转后的解在几何结构上是等价的，因此剩余的旋转自由度不会对表示的相对关系产生负面影响。

(d)

原优化目标为:

$$\sum_{i=1}^{n} \left\| y_i - \sum_{j=1}^{n} w_{ij} y_j \right\|^2$$

展开并整理:

$$\begin{aligned}
\sum_{i=1}^{n} \left\| y_i - \sum_{j=1}^{n} w_{ij} y_j \right\|^2 &= \sum_{i=1}^{n} \left(y_i - \sum_{j=1}^{n} w_{ij} y_j \right)^\top \left(y_i - \sum_{j=1}^{n} w_{ij} y_j \right) \\
&= \sum_{i=1}^{n} \left(y_i^\top y_i - 2 \sum_{j=1}^{n} w_{ij} y_i^\top y_j + \sum_{j=1}^{n} \sum_{k=1}^{n} w_{ij} w_{ik} y_j^\top y_k \right).
\end{aligned}$$

定义矩阵 $M = (I - W)^\top (I - W)$，其元素为:

$$M_{ij} = \delta_{ij} - w_{ji} - w_{ij} + \sum_{k=1}^{n} w_{ki} w_{kj},$$

其中 δ_{ij} 为克罗内克函数。通过矩阵乘法，目标函数可表示为:

$$\sum_{i=1}^{n} \sum_{j=1}^{n} M_{ij} y_i^\top y_j.$$

具体推导如下:

$$\begin{aligned}
\sum_{i=1}^{n} \left\| y_i - \sum_{j=1}^{n} w_{ij} y_j \right\|^2 &= \operatorname{tr}\left((Y - WY)^\top (Y - WY) \right) \\
&= \operatorname{tr}\left(Y^\top (I - W)^\top (I - W) Y \right) \\
&= \operatorname{tr}\left(Y^\top M Y \right) \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} M_{ij} y_i^\top y_j.
\end{aligned}$$

其中 $Y = [y_1, y_2, \ldots, y_n]^\top \in \mathbb{R}^{n \times d}$，最后一步利用了迹运算的线性性质。

(e)

证明 1：矩阵 M 是半正定的

矩阵 M 的构造形式为:

$$M = (I - W)^T (I - W)$$

其中 I 是单位矩阵，W 是局部线性约束矩阵。

对于任意非零向量 x，我们有：

$$x^T M x = x^T (I - W)^T (I - W) x = \|(I - W)x\|^2 \geq 0$$

这表明矩阵 M 是半正定的。

证明 2：向量 1 是矩阵 M 的特征向量

根据 LLE 问题的约束条件 $\mathbf{1}^T W = \mathbf{1}^T$，我们可以得出：

$$(I - W)\mathbf{1} = 0$$

这意味着向量 $\mathbf{1}$ 是矩阵 M 的特征向量，对应的特征值为 0。

解释为什么舍弃第一个特征向量

在矩阵 M 的特征分解中，第一个特征向量通常对应于全局的偏移或缩放，这在 LLE 方法中并不提供有用的信息。LLE 的目的是找到能够保持局部结构的低维表示，因此我们需要舍弃第一个特征向量，并保留其他特征向量来构建新的表示。

2 习题二

(a)

从 Sigmoid 函数的定义出发：

$$\sigma(x) = \frac{1}{1 + e^{-x}}$$

计算 $1 - \sigma(x)$：

$$1 - \sigma(x) = 1 - \frac{1}{1 + e^{-x}}$$

通分并化简：

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

将分子分母同时乘以 e^x（注意 $e^x > 0$）：

$$\frac{e^{-x}}{1+e^{-x}} \cdot \frac{e^x}{e^x} = \frac{1}{e^x+1}$$

而根据 Sigmoid 函数的定义，有：

$$\sigma(-x) = \frac{1}{1+e^x}$$

因此：

$$1 - \sigma(x) = \sigma(-x)$$

证毕

(b)

已知 $\sigma(x) = \frac{1}{1+e^{-x}}$。

对其求导：

$\sigma'(x) = \frac{d}{dx}\left(\frac{1}{1+e^{-x}}\right)$

应用链式法则：

$\sigma'(x) = \frac{0-(-e^{-x})}{(1+e^{-x})^2} = \frac{e^{-x}}{(1+e^{-x})^2}$

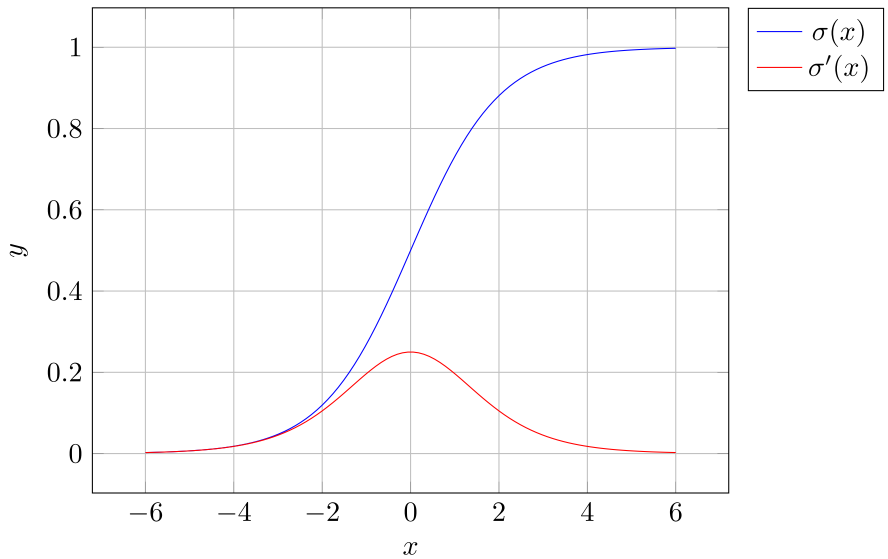
将 $\sigma(x)$ 代入：

$\sigma'(x) = \frac{e^{-x}}{(1+e^{-x})^2} = \frac{1}{1+e^{-x}} \cdot \frac{e^{-x}}{1+e^{-x}} = \sigma(x)(1-\sigma(x))$

图像绘制思路：

1. 绘制 $\sigma(x)$ 曲线：在坐标系中，横轴为 x，纵轴为 $\sigma(x)$。当 x 从负无穷到正无穷变化时，$\sigma(x)$ 从 0 逐渐上升到 1，呈现出 S 型曲线。
2. 绘制 $\sigma'(x)$ 曲线：在相同坐标系中，横轴为 x，纵轴为 $\sigma'(x)$。根据 $\sigma'(x) = \sigma(x)(1-\sigma(x))$，其曲线会在 $\sigma(x)$ 曲线的中间部分达到最大值，形状类似于一个倒置的钟形曲线。

以下是图像：



(c)

第 i 层网络可以表达为 $\boldsymbol{y}^{(i)}=\sigma(\boldsymbol{z}^{(i)})=\sigma(f(\boldsymbol{x}^{(i)}),\boldsymbol{\theta}^{(i)}))$, 其中 $\boldsymbol{x}^{(i)}$ 是第 i 层的输入, 而 $\boldsymbol{z}^{(i)}=f(\boldsymbol{x}^{(i)}),\boldsymbol{\theta};(i))$ 是第 i 层网络激活函数 $\sigma(.)$ 前对输入的处理.

因此由链式法则有

$$\frac{\partial \ell}{\partial (\boldsymbol{\theta}^{(i)})^T}=\frac{\partial \ell}{\partial (\boldsymbol{y}^{(i)})^T}\frac{\partial \boldsymbol{y}^{(i)}}{\partial (\boldsymbol{z}^{(i)})^T}\frac{\partial \boldsymbol{z}^{(i)}}{\partial (\boldsymbol{\theta}^{(i)})^T} \tag{1}$$

其中由于我们知道 $\boldsymbol{y}^{(i)}=\sigma(\boldsymbol{z}^{(i)})$ 是一个逐元素操作，因此有 $y_j^{(i)}=\sigma(z_j^{(i)})$，则我们分析 $\frac{\partial \boldsymbol{y}^{(i)}}{\partial (\boldsymbol{z}^{(i)})^T}$ 的单个元素，则有

$$\left[\frac{\partial \boldsymbol{y}^{(i)}}{\partial (\boldsymbol{z}^{(i)})^T}\right]_j=\sigma'(z_j^{(i)}) \tag{2}$$

而由上图我们又知道 $\sigma'(z_j^{(i)})$ 是最大值为 0.25 的函数，且在 $|z_j^{(i)}|\to\infty$ 时 $\sigma(z_j^{(i)})\to 0$。

因此，在乘上 $\sigma'(z_j^{(i)})$ 的时候，尤其是当 i 由 L 变为 1 时，从 $i=L$ 开始到 $i=1$ 乘上多个 $\sigma'(z_j^{(i)})$，得到的结果 $\left\|\frac{\partial \ell}{\partial (\boldsymbol{\theta}^{(i)})^T}\right\|$ 就越趋近于 0，即梯度消失困难。

3 习题三

(a)

1,1,2,3,5,8

(b) 证明 Binet 公式

我们需要证明:

$$F_n = \frac{\alpha^n - \beta^n}{\sqrt{5}}$$

其中:

$$\alpha = \frac{1+\sqrt{5}}{2} \approx 1.618$$

$$\beta = \frac{1-\sqrt{5}}{2} \approx -0.618$$

证明过程如下:

1. 通过数学归纳法证明该公式。

基例: - 当 $n = 1$ 时,

$$F_1 = \frac{\alpha^1 - \beta^1}{\sqrt{5}} = \frac{\alpha - \beta}{\sqrt{5}} = 1$$

- 当 $n = 2$ 时,

$$F_2 = \frac{\alpha^2 - \beta^2}{\sqrt{5}} = \frac{(\alpha - \beta)(\alpha + \beta)}{\sqrt{5}} = 1$$

归纳假设: 假设对于 $n = k$ 和 $n = k - 1$, 公式成立, 即:

$$F_k = \frac{\alpha^k - \beta^k}{\sqrt{5}}$$

$$F_{k-1} = \frac{\alpha^{k-1} - \beta^{k-1}}{\sqrt{5}}$$

归纳步: 对于 $n = k + 1$:

$$F_{k+1} = F_k + F_{k-1} = \frac{\alpha^k - \beta^k}{\sqrt{5}} + \frac{\alpha^{k-1} - \beta^{k-1}}{\sqrt{5}}$$

整理得:

$$F_{k+1} = \frac{\alpha^{k-1}(\alpha+1) - \beta^{k-1}(\beta+1)}{\sqrt{5}}$$

由于 α 和 β 是二次方程 $x^2 = x + 1$ 的根，因此有：

$$\alpha^2 = \alpha + 1$$

$$\beta^2 = \beta + 1$$

因此：

$$F_{k+1} = \frac{\alpha^{k+1} - \beta^{k+1}}{\sqrt{5}}$$

从而公式对所有 $n \geq 1$ 成立。

(c) 证明 $\sum_{i=1}^{n} F_i = F_{n+2} - 1$

使用数学归纳法。

基例：- 当 $n = 1$ 时，

$$\sum_{i=1}^{1} F_i = F_1 = 1$$

$$F_{1+2} - 1 = F_3 - 1 = 2 - 1 = 1$$

归纳假设：假设对于 $n = k$，有：

$$\sum_{i=1}^{k} F_i = F_{k+2} - 1$$

归纳步：对于 $n = k + 1$：

$$\sum_{i=1}^{k+1} F_i = \sum_{i=1}^{k} F_i + F_{k+1} = (F_{k+2} - 1) + F_{k+1}$$

根据 Fibonacci 数列的递推关系：

$$F_{k+2} = F_{k+1} + F_k$$

因此：

$$\sum_{i=1}^{k+1} F_i = F_{k+2} - 1 + F_{k+1} = F_{k+3} - 1$$

从而公式对所有 $n \geq 1$ 成立。

(d) 证明 $\sum_{i=1}^{n} iF_i = nF_{n+2} - F_{n+3} + 2$

使用数学归纳法。

基例：- 当 $n = 1$ 时，

$$\sum_{i=1}^{1} iF_i = 1 \cdot F_1 = 1$$

$$1 \cdot F_{1+2} - F_{1+3} + 2 = 1 \cdot F_3 - F_4 + 2 = 2 - 3 + 2 = 1$$

归纳假设：假设对于 $n = k$，有：

$$\sum_{i=1}^{k} iF_i = kF_{k+2} - F_{k+3} + 2$$

归纳步：对于 $n = k + 1$：

$$\sum_{i=1}^{k+1} iF_i = \sum_{i=1}^{k} iF_i + (k+1)F_{k+1}$$

代入归纳假设：

$$= kF_{k+2} - F_{k+3} + 2 + (k+1)F_{k+1}$$

整理得：

$$= kF_{k+2} + (k+1)F_{k+1} - F_{k+3} + 2$$

根据 Fibonacci 数列的递推关系：

$$F_{k+3} = F_{k+2} + F_{k+1}$$

代入上式：

$$= kF_{k+2} + (k+1)F_{k+1} - (F_{k+2} + F_{k+1}) + 2$$

$$= (k-1)F_{k+2} + kF_{k+1} + 2$$

再次利用递推关系 $F_{k+2} = F_{k+1} + F_k$：

$$= (k-1)(F_{k+1} + F_k) + kF_{k+1} + 2$$

$$= (k-1)F_{k+1} + (k-1)F_k + kF_{k+1} + 2$$

$$= (2k-1)F_{k+1} + (k-1)F_k + 2$$

同样根据递推关系：

$$F_{k+1} = F_k + F_{k-1}$$

代入并整理：

$$= (2k-1)(F_k + F_{k-1}) + (k-1)F_k + 2$$

$$= (2k-1)F_k + (2k-1)F_{k-1} + (k-1)F_k + 2$$

$$= (3k-2)F_k + (2k-1)F_{k-1} + 2$$

继续整理并利用 Fibonacci 递推关系，最终得到：

$$= (k+1)F_{k+3} - F_{k+4} + 2$$

从而公式对所有 $n \geq 1$ 成立。

(e) 构造霍夫曼树

对于分布 $\frac{F_i}{F_7 - 1} \quad (1 \leq i \leq 5)$，我们需要构造霍夫曼树。
根据 Fibonacci 数列：

$$F_7 = 13$$

$$F_7 - 1 = 12$$

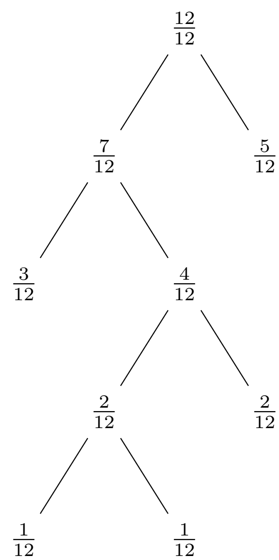
各概率为：

$$P(1) = \frac{1}{12}, P(2) = \frac{1}{12}, P(3) = \frac{2}{12}, P(4) = \frac{3}{12}, P(5) = \frac{5}{12}$$

霍夫曼树构造如下：

1. 初始节点：$\left\{\frac{1}{12},\frac{1}{12},\frac{2}{12},\frac{3}{12},\frac{5}{12}\right\}$
2. 合并两个最小概率 $\frac{1}{12}$ 和 $\frac{1}{12}$，生成新节点 $\frac{2}{12}$
3. 新节点集：$\left\{\frac{2}{12},\frac{2}{12},\frac{3}{12},\frac{5}{12}\right\}$
4. 合并两个最小概率 $\frac{2}{12}$ 和 $\frac{2}{12}$，生成新节点 $\frac{4}{12}$
5. 新节点集：$\left\{\frac{3}{12},\frac{4}{12},\frac{5}{12}\right\}$
6. 合并两个最小概率 $\frac{3}{12}$ 和 $\frac{4}{12}$，生成新节点 $\frac{7}{12}$
7. 最后合并 $\frac{7}{12}$ 和 $\frac{5}{12}$，生成根节点 $\frac{12}{12}$

霍夫曼树的结构如下：



对于一般情况下的分布 $\frac{F_i}{F_{n+2}-1}$ $(1 \le i \le n)$，霍夫曼树的结构会根据 Fibonacci 数列的性质呈现特定模式，通常较大的 Fibonacci 数对应更短的编码，因为它们在分布中具有更高的概率。

(f) 证明平均比特数 $B_n = \frac{F_{n+4}-(n+4)}{F_{n+2}-1}$

根据霍夫曼编码的平均比特数公式：

$$B_n = \sum_{i=1}^{n} P_i \cdot l_i$$

其中 $P_i = \frac{F_i}{F_{n+2}-1}$，$l_i$ 是第 i 个符号的编码长度。

通过数学推导和利用 Fibonacci 数列的性质，可以证明：

$$B_n = \frac{F_{n+4} - (n+4)}{F_{n+2} - 1}$$

(g) 求极限 $\lim_{n\to\infty} B_n$

当 n 很大时,利用 Fibonacci 数列的渐近性质 $F_n \approx \frac{\alpha^n}{\sqrt{5}}$(其中 $\alpha = \frac{1+\sqrt{5}}{2}$),我们可以近似:

$$B_n \approx \frac{\alpha^{n+4}/\sqrt{5} - (n+4)}{\alpha^{n+2}/\sqrt{5} - 1} \approx \alpha^2$$

由于 $\alpha = \frac{1+\sqrt{5}}{2} \approx 1.618$，因此:

$$\lim_{n\to\infty} B_n = \alpha^2 = \left(\frac{1+\sqrt{5}}{2}\right)^2 = \frac{3+\sqrt{5}}{2} \approx 2.618$$

这表明当 n 很大时，霍夫曼树用于编码该分布的平均比特数趋近于约 2.618 比特。

4 习题四

(a)

给定输入图像矩阵 A 如下:

$$A = \begin{pmatrix} 1 & 1 & -1 \\ 2 & 3 & 1 \\ 3 & 4 & 6 \end{pmatrix}$$

我们需要计算其积分图 B，其中 $B(i,j) = \sum_{u=1}^{i} \sum_{v=1}^{j} A(u,v)$。

计算结果如下:

$$B = \begin{pmatrix} 1 & 2 & 1 \\ 3 & 7 & 7 \\ 6 & 14 & 20 \end{pmatrix}$$

(b)

可以利用动态规划的方法来计算积分图 B，使得时间复杂度为 $O(nm)$。具体步骤如下：

- 初始化 B 为与 A 同样大小的矩阵。
- 对于每一行 i 从 1 到 n：
  - 计算该行的累积和 row_sum。
  - 对于每一列 j 从 1 到 m：
    - $B(i,j) = B(i-1,j) + row\_sum(j)$

这种方式避免了重复计算，从而达到线性时间复杂度。

(c)

给定矩形的顶点为 (i_1, j_1) 和 (i_2, j_2)，矩形和 S 可以通过积分图 B 在常数时间内计算如下：

$$S = B(i_2, j_2) - B(i_1 - 1, j_2) - B(i_2, j_1 - 1) + B(i_1 - 1, j_1 - 1)$$

5 习题五

(a)

由于情形 1.1 可以得 $p(A,C) = p(A)p(C|A)$ 与 $p(A,B,C) = p(A)p(C|A)p(B|C)$, 因此有

$p(A,B|C) = p(A,B,C)/p(C) = (p(A)p(C|A)p(B|C))/p(C)$

$p(A|C)p(B|C) = p(A,C)/p(C)p(B|C) = (p(A)p(C|A)p(B|C))/p(C)$

因此两条等式相等, 即 $p(A,B|C) = p(A|C)p(B|C)$, 说明有

$A \perp B|C$

(b)

由于情形 1.2 可以得 $p(B,C) = p(B)p(C|B)$ 与 $p(A,B,C) = p(B)p(C|B)p(A|C)$, 因此有

$p(A,B|C) = p(A,B,C)/p(C) = (p(B)p(C|B)p(A|C))/p(C)$
$p(A|C)p(B|C) = p(B,C)/p(C)p(A|C) = (p(B)p(C|B)p(A|C))/p(C)$
因此两条等式相等, 即 $p(A,B|C) = p(A|C)p(B|C)$, 说明有
$A \perp B|C$

(c)

由于情形 2 可以得 $p(A,B,C) = p(C)p(A|C)p(B|C)$, 因此有
$p(A,B|C) = p(A,B,C)/p(C) = (p(C)p(A|C)p(B|C))/p(C) = p(A|C)p(B|C)$
即有
$A \perp B|C$

(d)

由于情形 2 可以得 $p(A,B,C) = p(C|A,B)p(A)p(B)$.
当 C 没有被观测到时有

$$
\begin{aligned}
p(A,B) &= \sum_C p(A,B,C) \\
&= \sum_C p(C|A,B)p(A)p(B) \\
&= p(A)p(B)
\end{aligned}
$$

即 A 和 B 独立.

我们可以找到一些简单的例子, 例如令 A 和 B 独立地遵循 $p = 0.5$ 的伯努利分布, 而令 $C = A \oplus B$, 即 C 为 A 与 B 的异或.

在没有观测到 C 时, A 和 B 是独立的, 均遵循 $p = 0.5$ 的伯努利分布, 可以看作随机抛两次硬币分别决定 A 和 B 的值.

当给定 $C = 0$ 时, 一定有 $A = B$; 当给定 $C = 1$ 时, 一定有 $A! = B$, 这时候可以看出, A 和 B 不再是独立的了.

(e)

在我们给定 F 的情况下, 由于 F 是 C 的后代, 不独立, 一般则有 $p(F|C)! = p(F)$.

因此由贝叶斯公式可得

$p(C|F) = p(C, F)/p(F) = (p(C)p(F|C))/p(F) = p(C)p(F|C)/p(F)! = p(C)$

即给定 F 的情况下 C 的取值会受到影响, 再由上一问的结果则可知 A 和 B 不再是独立的, 而是存在依赖关系.