信息论第一次小测解答

中国科学技术大学《信息论 A》006125.01 班助教组

2024 年 3 月 13 日

第 1 题

证明下列关于熵的链式法则：

- （基本形式）对于两个随机变量 X，Y，我们有

$$H(X,Y) = H(X) + H(Y|X) = H(Y) + H(X|Y). \tag{1}$$

- （条件形式）对于三个随机变量 X，Y，Z，我们有：

$$H(X,Y|Z) = H(X|Z) + H(Y|X,Z). \tag{2}$$

- （一般形式）对于 n 个随机变量 $X_1, X_2, ..., X_n$，我们有：

$$H(X_1, X_2, \ldots, X_n) = \sum_{i=1}^{n} H(X_i|X_{i-1}, \ldots, X_1), \tag{3}$$

  其中 X_0 可以视为一个退化的随机变量，比如说一个常数。

证明. 讲义中已经给出了(1)的证明，我们将其附在下面，基于此给出其他形式的证明。

$$\begin{aligned}
H(X,Y) &= \mathbf{E}\left[\log \frac{1}{P_{X,Y}(X,Y)}\right] \\
&= \mathbf{E}\left[\log \frac{1}{P_X(X)P_{Y|X}(Y|X)}\right] \\
&= \mathbf{E}\left[\log \frac{1}{P_X(X)} + \log \frac{1}{P_{Y|X}(Y|X)}\right] \\
&= \mathbf{E}\left[\log \frac{1}{P_X(X)}\right] + \mathbf{E}\left[\log \frac{1}{P_{Y|X}(Y|X)}\right] \\
&= H(X) + H(Y|X).
\end{aligned}$$

关于链式法则条件形式(2)的证明：

$$H(X,Y|Z) = \mathbf{E}\left[\log \frac{1}{P_{X,Y|Z}(X,Y|Z)}\right]$$

$$
\begin{aligned}
&= \mathbf{E}\left[\log \frac{1}{P_{X|Z}(X|Z)P_{Y|X,Z}(Y|X,Z)}\right] \\
&= \mathbf{E}\left[\log \frac{1}{P_{X|Z}(X|Z)} + \log \frac{1}{P_{Y|X,Z}(Y|X,Z)}\right] \\
&= \mathbf{E}\left[\log \frac{1}{P_{X|Z}(X|Z)}\right] + \mathbf{E}\left[\log \frac{1}{P_{Y|X,Z}(Y|X,Z)}\right] \\
&= H(X|Z) + H(Y|X,Z).
\end{aligned}
$$

关于链式法则一般形式(3)的证明：

$$
\begin{aligned}
H(X_1, X_2, ..., X_n) &= H(X_1) + H(X_2, X_3, ..., X_n|X_1) && (4) \\
&= H(X_1) + H(X_2|X_1) + H(X_3, X_4, ..., X_n|X_2, X_1) && (5) \\
&= \sum_{i=1}^{n} H(X_i|X_{i-1}, ..., X_1), && (6)
\end{aligned}
$$

其中(4)使用了链式法则的基本形式，(5)使用了链式法则的条件形式，(6)不断使用链式法则的条件形式即可。

也可以使用如下条件概率的连续展开：

$$
P(X_1, ..., X_n) = \prod_{i=1}^{n} P(X_i|X_{i-1}, ...X_1),
$$

从而得到：

$$
\begin{aligned}
H(X_1, ..., X_n) &= \mathbf{E}\left[\log \frac{1}{P(X_1, ..., X_n)}\right] = \mathbf{E}\left[\log \frac{1}{\prod_{i=1}^{n} P(X_i|X_{i-1}, ...X_1)}\right] \\
&= \sum_{i=1}^{n} \mathbf{E}\left[\log \frac{1}{P(X_i|X_{i-1}, ...X_1)}\right] = \sum_{i=1}^{n} H(X_i|X_{i-1}, ..., X_1).
\end{aligned}
$$

□

第 2 题

证明互信息满足下列基本性质：

- （对称性）

$$
I(X;Y) = I(Y;X) \tag{7}
$$

- （自信息）

$$
I(X;X) = H(X) \tag{8}
$$

- （互信息的分解）

$$
\begin{aligned}
I(X;Y) &= H(X) - H(X|Y) && (9) \\
&= H(Y) - H(Y|X) && (10) \\
&= H(X) + H(Y) - H(X,Y) && (11)
\end{aligned}
$$

- （带条件互信息的分解）

$$I(X;Y|Z) = H(X|Z) - H(X|Y,Z) \tag{12}$$
$$= H(Y|Z) - H(Y|X,Z) \tag{13}$$
$$= H(X|Z) + H(Y|Z) - H(X,Y|Z) \tag{14}$$

证明. (7)的证明：

$$I(X;Y) = D(P_{X,Y}||P_X P_Y) = \sum_{(x,y)\in X(\Omega)\times Y(\Omega)} P_{X,Y}(x,y)\log\frac{P_{X,Y}(x,y)}{P_X(x)P_Y(y)}$$
$$= \sum_{(y,x)\in Y(\Omega)\times X(\Omega)} P_{Y,X}(y,x)\log\frac{P_{Y,X}(y,x)}{P_Y(y)P_X(x)} = D(P_{Y,X}||P_Y P_X) = I(Y;X).$$

(8)的证明：

$$I(X;X) = D(P_{X,X}||P_X P_X) = \sum_{(x,x)\in X(\Omega)\times X(\Omega)} P_{X,X}(x,x)\log\frac{P_{X,X}(x,x)}{P_X(x)P_X(x)}$$
$$\overset{(a)}{=} \sum_{x\in X(\Omega)} P_X(x)\log\frac{P_X(x)}{P_X(x)P_X(x)} = \sum_{x\in X(\Omega)} P_X(x)\log\frac{1}{P_X(x)} = H(X).$$

其中 (a) 是由于在 $x_i \neq x_j$ 时，$P_{X,X}(x_i,x_j)\log\frac{P_{X,X}(x_i,x_j)}{P_X(x_i)P_X(x_j)}$ 的值被视作 $0(0\log\frac{0}{0}=0)$，从而我们的支撑集可以从 $X(\Omega)\times X(\Omega)$ 简化为 $X(\Omega)$，并且此时 $P_{X,X}(x_i,x_i) = P_X(x_i)$。

(9)的证明：

$$I(X;Y) = \mathbf{E}\left[\log\frac{P_{X,Y}(X,Y)}{P_X(X)P_Y(Y)}\right] = \mathbf{E}\left[\log\frac{P_X(X)P_{Y|X}(Y|X)}{P_X(X)P_Y(Y)}\right] = \mathbf{E}\left[\log\frac{P_{Y|X}(Y|X)}{P_Y(Y)}\right]$$
$$= \mathbf{E}\left[\frac{1}{\log P_Y(Y)}\right] - \mathbf{E}\left[\log\frac{1}{P_{Y|X}(Y|X)}\right] = H(Y) - H(Y|X).$$

通过对联合概率的不同拆分，有 $I(X;Y) = H(X) - H(X|Y)$，而对三项全部做展开便得到 $I(X;Y) = H(X) + H(Y) - H(X,Y)$.

而事实上，通过链式法则(1)：$H(X,Y) = H(X) + H(Y|X) = H(Y) + H(X|Y)$，我们只需要导出(9)(10)(11)中的任一个，其余两项均可以简单得到。

(12)的证明：

$$I(X;Y|Z) = \mathbf{E}\left[\log\frac{P_{X,Y|Z}(X,Y|Z)}{P_{X|Z}(X|Z)P_{Y|Z}(Y|Z)}\right] = \mathbf{E}\left[\log\frac{P_{X|Z}(X|Z)P_{Y|X,Z}(Y|X,Z)}{P_{X|Z}(X|Z)P_{Y|Z}(Y|Z)}\right]$$
$$= \mathbf{E}\left[\log\frac{P_{Y|X,Z}(Y|X,Z)}{P_{Y|Z}(Y|Z)}\right] = \mathbf{E}\left[\frac{1}{\log P_{Y|Z}(Y|Z)}\right] - \mathbf{E}\left[\log\frac{1}{P_{Y|X,Z}(Y|X,Z)}\right]$$
$$= H(Y|Z) - H(Y|X,Z).$$

其余两个式子可以通过条件形式的链式法则(2)得到，没什么区别，不再赘述。 □

多使用带期望的表示方法，如无必要，无需完全展开成概率求和形式。