Конспект по условной энтропии

- Есть 2 случайные величины $\mathcal{X}, \mathcal{Y}$.
- Пусть известно, что $\mathcal{Y} = y$.
- Тогда:

$$H(\mathcal{X}|\mathcal{Y} = y) = -\sum_{i=1}^{n} p(\mathcal{X} = x_i|\mathcal{Y} = y) \cdot \log(p(\mathcal{X} = x_i|\mathcal{Y} = y))$$

- $H(\mathcal{X}|\mathcal{Y} = y)$ — это среднее количество информации в сообщении о значении X, если до этого было получено сообщение $Y = y$.
- $H(\mathcal{X}|\mathcal{Y} = y)$ — это функция от случайной величины Y, т.е. тоже случайная величина.

Определение (условной энтропии)

Условная энтропия X при условии Y — математическое ожидание $H(\mathcal{X}|\mathcal{Y} = y)$:

$$H(X|Y) = \mathbb{E}[H(\mathcal{X}|\mathcal{Y} = y)] = -\sum_{j=1}^{k} P(\mathcal{Y} = y_j) \sum_{i=1}^{n} p(\mathcal{X} = x_i|\mathcal{Y} = y_j) \cdot \log(p(\mathcal{X} = x_i|\mathcal{Y} = y_j))$$

По формуле совместной вероятности $P(x, y) = P(y)P(x|y)$:

$$H(X|Y) = -\sum_{j=1}^{k} \sum_{i=1}^{n} P(\mathcal{X} = x_i, \mathcal{Y} = y_j) \cdot \log(p(\mathcal{X} = x_i|\mathcal{Y} = y_j))$$

Замечание

- Если X, Y независимы, то для всех i, j: $P(\mathcal{X} = x_i, \mathcal{Y} = y_j) = P(\mathcal{X} = x_i)P(\mathcal{Y} = y_j)$.
- Тогда: $\sum_{j=1}^{k} P(\mathcal{X} = x_i, \mathcal{Y} = y_j) = P(\mathcal{X} = x_i)$.
- Получаем, что $H(X|Y) = H(X)$.

Обобщение на 3 и более случайных величины

$$H(\mathcal{X},\mathcal{Y}) = -\sum_{i=1}^{n}\sum_{j=1}^{k} P(\mathcal{X}=x_i,\mathcal{Y}=y_j)\cdot\log(P(\mathcal{X}=x_i,\mathcal{Y}=y_j))$$

$$H(\mathcal{X},\mathcal{Y}|\mathcal{Z}) = -\sum_{i=1}^{n}\sum_{j=1}^{k}\sum_{t=1}^{m} P(\mathcal{X}=x_i,\mathcal{Y}=y_j,\mathcal{Z}=z_t)\cdot\log\left(P(\mathcal{X}=x_i,\mathcal{Y}=y_j|\mathcal{Z}=z_t)\right)$$

Цепное правило (главная теорема билета)

$$H(X,Y) = H(X) + H(Y|X) = H(Y) + H(X|Y)$$

Для трёх случайных величин:

$$H(X,Y|Z) = H(X|Z) + H(Y|X,Z) = H(Y|Z) + H(X|Y,Z)$$

Доказательство

Из $P(A\cdot B) = P(B)\cdot P(A|B)$ получаем:

$$\log(P(A,B)) = \log(P(B)) + \log(P(A|B))$$

- $\log(P(B))$ превращается в энтропию Y.
- $\log(P(A|B))$ превращается в условную энтропию X при условии Y.

Для условной энтропии:

$$P(A,B|C) = \frac{P(A,B,C)}{P(C)} = \frac{P(B,C)\cdot P(A|B,C)}{P(C)} = P(B|C)\cdot P(A|B,C)$$

$$\log P(X=x_i,Y=y_j|Z=z_t) = \log P(X=x_i|Y=y_j,Z=z_t) + \log P(Y=y_j|Z=z_t)$$

- Первый логарифм превращается в $H(X|Y,Z)$.
- Второй — в $H(Y|Z)$.

$\blacksquare$

Следствие из цепного правила и замечания

Если X, Y — независимы, то $H(X,Y) = H(X) + H(Y)$.