

# BÀI 3 - PHÂN TÍCH THÀNH PHẦN ĐỘC LẬP ICA

Hạn nộp bài: hết ngày 10/04/2024.

## 1 Bài tập lý thuyết

**Bài 1.** (Một số khái niệm của lý thuyết thông tin)

Xét hai biến ngẫu nhiên rời rạc  $X$  và  $Y$  có tập xác định lần lượt là  $D_X$  và  $D_Y$ .

- *Entropy* của  $X$  cho bởi

$$H(X) = \mathbb{E}[-\log p(X)] = - \sum_{x \in D_X} p(x) \log p(x).$$

- *Entropy đồng thời (Joint Entropy)* của  $X$  và  $Y$  cho bởi

$$H(X, Y) = - \sum_{x \in D_X} \sum_{y \in D_Y} p(x, y) \log p(x, y).$$

- *Entropy có điều kiện (Conditional Entropy)* của  $Y$  khi biết  $X$  cho bởi

$$H(Y | X) = - \sum_{x \in D_X} \sum_{y \in D_Y} p(x, y) \log \frac{p(x, y)}{p(x)}.$$

- *Thông tin tương hỗ (Mutual Information)* của  $X$  và  $Y$  cho bởi

$$I(X; Y) = \sum_{x \in D_X} \sum_{y \in D_Y} p(x, y) \log \left( \frac{p(x, y)}{p(x)p(y)} \right).$$

Hãy chứng minh các tính chất sau:

- $I(X; Y) \geq 0$ ;
- $I(X; Y) = H(X) - H(X | Y) = H(Y) - H(Y | X)$ ;
- $I(X; Y) = H(X) + H(Y) - H(X, Y)$ ;
- $I(X; Y) = H(X, Y) - H(X | Y) - H(Y | X)$ .

**Bài 2.** Trong quá trình thiết lập ICA, ta có biểu thức

$$H(Wx) = H(x) + \log |\det W|.$$

Hãy chứng minh biểu thức trên.

## 2 Bài tập thực hành

**Bài 3.** Cho đoạn code

```
n_samples = 2000
time = np.linspace(0, 8, n_samples)

s1 = np.sin(2 * time)
s2 = np.sign(np.sin(3 * time))
s3 = np.random.normal(size=n_samples)
S = np.c_[s1, s2, s3]

A = np.array([[0.5, 0.1, 0.3], [0.2, 0.8, 0.5], [0.1, 0.3, 0.7]])
X = np.dot(S, A.T)
```

- a) Đọc hiểu đoạn code trên.  $S$  và  $X$  là gì? Hãy trực quan hoá chúng.
- b) Dùng ICA để tái tạo lại  $S$  từ  $X$ . Trực quan hoá kết quả và nhận xét.
- c) Dùng PCA để tái tạo lại  $S$  từ  $X$ . Trực quan hoá, nhận xét và so sánh với câu trên.

**Bài 4.** Dữ liệu ở [Kaggle - ICA musical](#) chứa 3 đoạn ghi âm (ICA mix 1, 2, 3). Ta sẽ dùng ICA để tái tạo lại các đoạn âm gốc theo các bước sau:

- a) Đọc dữ liệu và trực quan hoá. Phát các đoạn ghi âm này.
- b) Dùng ICA để tái tạo lại các đoạn âm gốc.
- c) Trực quan hoá kết quả. Phát các đoạn âm kết quả và nhận xét.