CHƯƠNG 6: TỐC ĐỘ LƯỢNG TIN RIÊNG

TS. TRỊNH VĂN CHIẾN (SOICT-HUST)

ĐẶT VẤN ĐỀ

- Theo lý thuyết về xấp xỉ tiệm cận: khi tung một đồng xu và xem xét 2 mặt của nó

$$\left(x_1,\ldots,x_n\right)\approx 2^{-nH(X)}$$

- Mở rộng trường hợp phức tạp: Chia bài

$$\left(x_1,\ldots,x_n\right)\approx ????$$

CHUỖI MARKOV

- Chuỗi Markov dùng để thiết lập sự phụ thuộc của một chuỗi các biến ngẫu nhiên
- Xem xét một quá trình ngẫu nhiên

$$\{X_1, ..., X_n\}, X_i \in \tilde{X}$$

- Trạng thái tiếp theo chỉ phụ thuộc vào trạng thái trước đó

$$p(x_{n+1} | x_n, ..., x_1) = p(x_{n+1} | x_n)$$

- Xác suất chuyển tiếp: Từ trạng thái $i \rightarrow j$ với xác suất $p_{i,j}$
  - Ta có $p(x_{n+1}) = \sum_{x_n} p(x_n) p(x_{n+1} | x_n)$
  - Do đó $p(x_1, ..., x_n) = p(x_1) p(x_2 | x_1) ... p(x_n | x_{n-1})$

CHUỖI MARKOV ẨN

- Ứng dụng rộng rãi trong nhận dạng giọng nói, nhận dạng chữ viết tay, và học máy
- Không quan sát được chuỗi Markov $(X_1,\dots,X_n)$
- Quan sát chuỗi ngẫu nhiên $(Y_1,\dots,Y_n)$

$$Y_i \sim p(y_i \mid x_í)$$

- Xây dựng mô hình xác suất

$$p(x^n, y^n) = p(x_1)\prod_{i=1}^{n-1} p(x_{i+1} \mid x_i)\prod_{i=1}^{n} p(y_i \mid x_i)$$

CHUỖI MARKOV ĐỘC LẬP THỜI GIAN

- Một chuỗi Markov độc lập thời gian nếu xác suất có điều kiện $p(x_n \mid x_{n-1})$ không phụ thuộc vào *n*, nghĩa là

$$p\left(X_{n+1} = b \mid X_n = a\right) = p\left(X_2 = b \mid X_1 = a\right)$$

- Đối với chuỗi Markov độc lập thời gian, ma trận chuyển trạng thái được xây dựng như sau:

$$P = \begin{bmatrix} P_{11} & \dots & P_{1n} \\ \dots & \dots & \dots \\ P_{n1} & \dots & P_{nn} \end{bmatrix}$$

VÍ DỤ (1)

- Dự báo thời tiết với 02 trạng thái: Nắng (N) và Mưa (M)

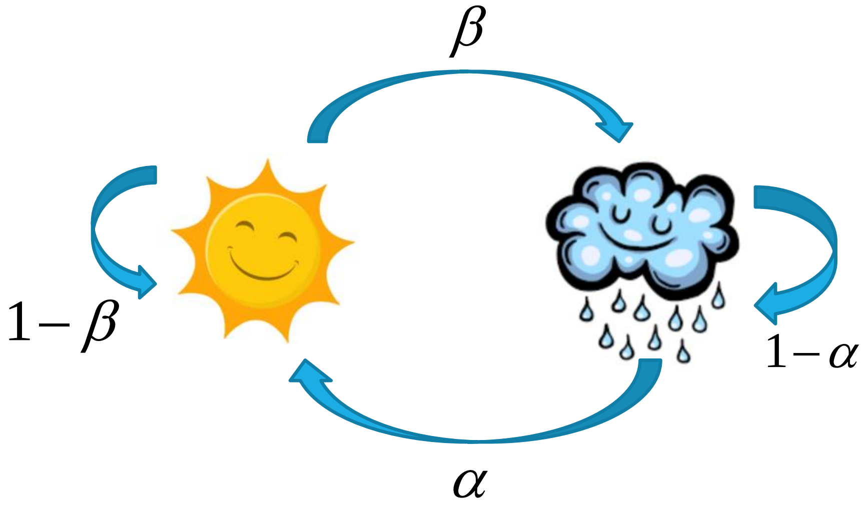
$$\tilde{X} = \{N, M\}$$

- Xác suất có điều kiện:

$$p(N \mid N) = 1 - \beta,\ p(M \mid M) = 1 - \alpha,\ p(M \mid N) = \beta,\ p(N \mid M) = \alpha$$

- Ma trận chuyển trạng thái

$$P = \begin{bmatrix} 1-\beta & \beta \\ \alpha & 1-\alpha \end{bmatrix}$$



VÍ DỤ (2)

- Xác suất quan sát chuỗi NNMM

$$p(NNMM) = p(N)\,p(N \mid N)\,p(M \mid N)\,p(M \mid M) = p(N)(1-\beta)\beta\left(1-\alpha\right)$$

- Một số câu hỏi:
  - Chuỗi sự kiện trên sẽ thay đổi như thế nào sau nhiều ngày quan sát?
  - Chuỗi sự kiện nào sẽ nằm trong tập tiêu biểu?
  - Xác suất của một chuỗi trong tập tiêu biểu là bao nhiêu?

PHÂN BỐ DỪNG (1)

- Phân bố dừng (stationary distribution): Một phân bố μ tại các trạng thái sao cho phân bố tại 2 trạng thái n+1 và n là giống nhau
- Ví dụ: Về dự báo thời tiết (nắng, mưa):
  - Giả sử:

$$\mu(N) = \frac{\alpha}{\alpha+\beta}, \mu(M) = \frac{\beta}{\alpha+\beta}, P = \begin{bmatrix} 1-\beta & \beta \\ \alpha & 1-\alpha \end{bmatrix}$$

  - Ta có:

$$p\left(X_{n+1} = N\right) = p(N \mid N)\mu(N) + p(N \mid M)\mu(M)$$

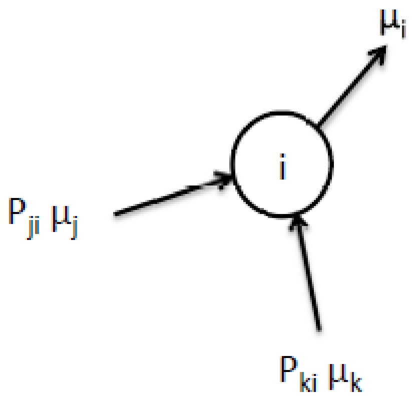
$$= \left(1-\beta\right)\frac{\alpha}{\alpha+\beta} + \alpha\frac{\beta}{\alpha+\beta}$$

$$= \frac{\alpha}{\alpha+\beta} = \mu(N)$$

PHÂN BỐ DỪNG (2)

- Phân bố dừng $\mu P = \mu$
- Phân bố dừng $\mu_i, i = 1, \ldots, |\tilde{X}|$ thỏa mãn

$$\mu_i = \sum_j \mu_j p_{ji}, \quad \sum_{i=1}^{|\tilde{X}|} \mu_i = 1$$



QUÁ TRÌNH DỪNG

- Một quá trình ngẫu nhiên là dừng (stationary process) nếu xác suất xảy ra đồng thời của một tập mẫu bất kì không phụ thuộc vào thời gian

$$p\left(X_1 = x_1, \ldots, X_n = x_n\right) = p\left(X_2 = x_1, \ldots, X_{n+1} = x_n\right)$$

- Ví dụ: Tung đồng xu với 2 mặt 0, 1

$$p\left(X_1 = 0, X_2 = 1\right) = p\left(X_2 = 0, \ldots, X_3 = 1\right) = p(1 - p)$$

TÔC ĐỘ LƯỢNG TIN RIÊNG (1)

- Nếu các biến ngẫu nhiên X_i là độc lập, ta có

$$H\left(X^n\right) = H\left(X_1, \ldots, X_n\right) = \sum_{i=1}^{n} H\left(X_i\right) = nH(X)$$

- Nếu chuỗi biến ngẫu nhiên là phụ thuộc:
  - Lượng tin riêng trung bình mỗi ký hiệu

$$H\left(\tilde{X}\right) = \lim_{n\to\infty} \frac{H\left(X^n\right)}{n}$$

  - Tốc độ phát triển của thông tin

$$H'\left(\tilde{X}\right) = \lim_{n\to\infty} H(X_n \mid X_{n-1}, \ldots, X_1)$$

TỐC ĐỘ LƯỢNG TIN RIÊNG (2)

- Nếu tốc độ phát triển của thông tin tồn tại và X_i là dừng

$$H\left(X_n \mid X_1,\dots,X_{n-1}\right) \le H\left(X_n \mid X_2,\dots,X_{n-1}\right) \le H\left(X_{n-1} \mid X_1,\dots,X_{n-2}\right)$$

- Quan sát:
  - Nếu n tăng thì $H\left(X_n \mid X_1,\dots,X_{n-1}\right)$ giảm
  - Lượng tin riêng luôn không âm: $H\left(X\right) \ge 0$
  - Tốc độ phát triển của thông tin có giới hạn và giới hạn này tồn tại

TÔC ĐỘ LƯỢNG TIN RIÊNG (3)

- Nếu X_i là dừng, $H(\tilde{X}) = H'(\tilde{X})$

$$\frac{1}{n}H\left(X_1,\ldots,X_n\right) = \frac{1}{n}\sum_{i=1}^{n}H(X_i \mid X_{i-1},\ldots,X_1)$$

- Ta có $H\left(X_n \mid X_1,\ldots,X_{n-1}\right) \to H'(\tilde{X})$
- Kì vọng Cesaro:
  - Nếu $a_n \to a, b_n = \frac{1}{n}\sum_{i=1}^{n}a_i, b_i \to a$, thì $b_n \to a$
  - Do đó ta có $$\frac{1}{n}H\left(X_1,\ldots,X_n\right) \to H'(\tilde{X})$$

ĐẶC TÍNH XẤP XỈ TIỆM CẬN CHO QUÁ TRÌNH DỪNG

- Dựa vào đặc tính xấp xỉ tiệm cận ta có

$$-\frac{1}{n}\log p\left(X_1,\dots,X_n\right)\to H(\tilde{X})$$

- Một số quan sát:
  - Xác suất : $p(X_1,\dots,X_n)\approx 2^{-nH(X)}$
  - Số chuỗi tiêu biểu trong tập tiêu biểu: $2^{nH(\tilde{X})}$
  - Chúng ta có thể sử dụng $nH(\tilde{X})$ bits để mô tả một chuỗi tiêu biểu

TỐC ĐỘ LƯỢNG TIN RIÊNG CHO CHUỖI MARKOV

- Chuỗi Markov $H(\tilde{X}) = \lim_{n\to\infty} H(X_n | X_{n-1},\ldots, X_1) = \lim_{n\to\infty} H(X_n | X_{n-1}) = H(X_2 | X_1)$
- Theo định nghĩa xác suất chuyển trạng thái

$$p(X_2 = j | X_1 = i) = P_{ij}$$

- Tốc độ lượng tin riêng của chuỗi Markov

$$H(\tilde{X}) = -\sum_{ij} \mu_i P_{ij} \log P_{ij}$$

- Tính tốc độ lượng tin riêng cho chuỗi Markov tương đối dễ:
  - Định nghĩa phân bố dừng μ_i
  - Sử dụng xác suất chuyển trạng thái để tính

$$H(\tilde{X}) = -\sum_{ij} \mu_i P_{ij} \log P_{ij}$$

VÍ DỤ

- Xem xét ví dụ về dự báo thời tiết (Nắng, Mưa)
- Phân bố dừng $\mu(N) = \frac{\alpha}{\alpha+\beta}, \mu(M) = \frac{\beta}{\alpha+\beta}$

$$P = \begin{bmatrix} 1-\beta & \beta \\ \alpha & 1-\alpha \end{bmatrix}$$

- Tốc độ lượng tin riêng

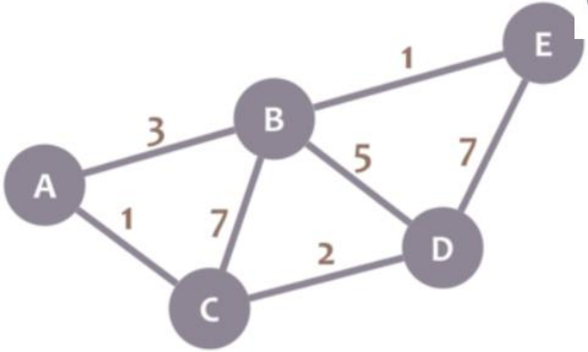
$$H(\tilde{X}) = \frac{\beta}{\alpha+\beta}\left(\alpha\log\alpha + (1-\alpha)\log\left(1-\alpha\right)\right) = \frac{\alpha}{\alpha+\beta}H(\beta) + \frac{\beta}{\alpha+\beta}H\left(\alpha\right) \le H\left(2\frac{\alpha\beta}{\alpha+\beta}\right) \le H\left(\sqrt{\alpha\beta}\right)$$

  - Tốc độ lượng tin riêng đạt cực tại khi α=β=0.5: Hai sự kiện là độc lập với nhau

BƯỚC NGẪU NHIÊN TRÊN ĐỒ THỊ (1)

- Xét một đồ thị vô hướng có m điểm $\{1,\ldots,m\}$
- Cạnh $i \to j$ có trọng số $W_{ij} \geq 0 \left(W_{ij} = W_{ji}\right)$
- Xem xét quá trình bước ngẫu nhiên từ điểm này đến điểm kia trên đồ thị
  - Quá trình bước ngẫu nhiên $X_1, X_2, \ldots$ là một chuỗi các đỉnh của đồ thị (vertices)
  - Với $X_n = i$, bước tiếp theo sẽ được lựa chọn từ các đỉnh lân cận với xác suất

$$P_{ij} = \frac{W_{ij}}{\sum_k W_{ik}}$$



BƯỚC NGẪU NHIÊN TRÊN ĐỒ THỊ (2)

- Đặt

$$W_i = \sum_j W_{ij}, \quad W = \sum_{i,j:i>j} W_{ij}$$

- Phân bố dừng

$$\mu_i = \frac{W_i}{2W}$$

- Có thể kiểm tra phân bố dừng: $\mu P = \mu$