

言語処理のための機械学習入門

5. 系列ラベリング

5.1 準備

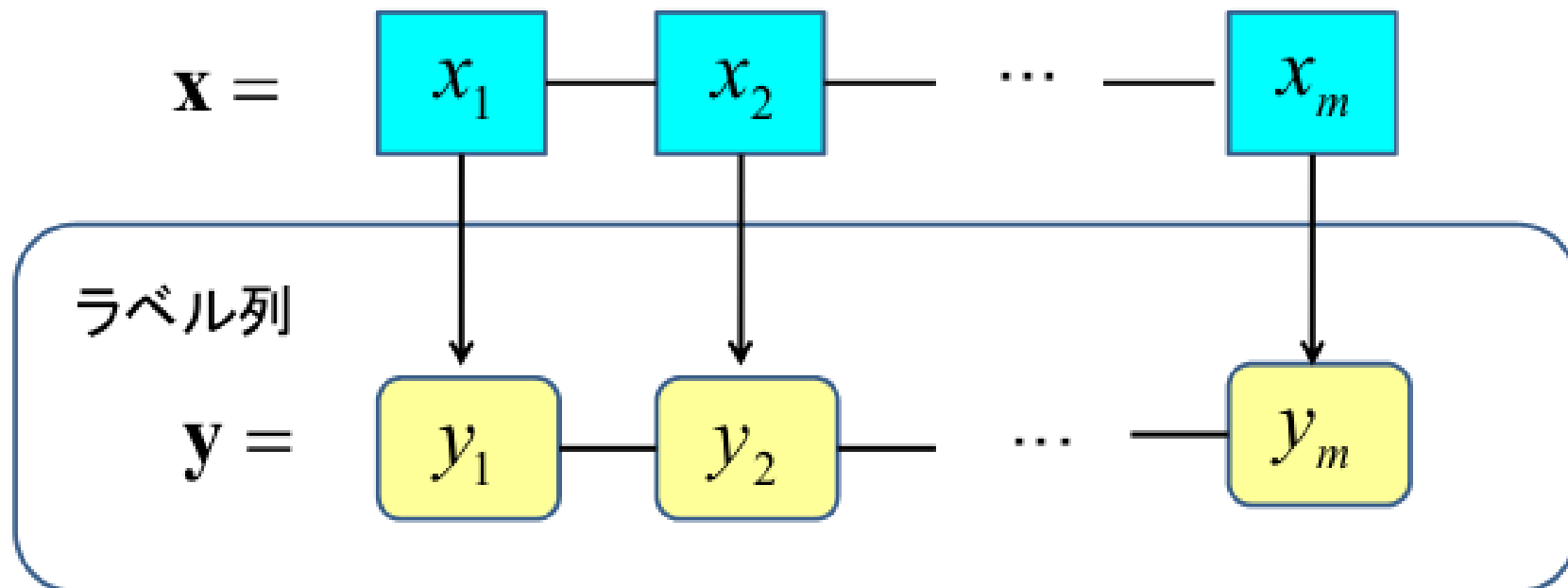
5.2 隠れマルコフモデル

5.3 通常のカテゴリ器の逐次適用

新納浩幸

系列ラベリング問題

観測系列



$y_i \in Y$ (ラベル集合)

観測系列からラベル列を推定する問題

隠れマルコフモデル

Hidden Markov Model (HMM)

(x_i, y_i) は (x_{i-1}, y_{i-1}) のみに依存

詳細に言うと $\left\{ \begin{array}{l} x_i \text{ は } y_i \text{ のみに依存} \\ y_i \text{ は } y_{i-1} \text{ のみに依存} \end{array} \right.$

$$\begin{aligned} P(\mathbf{x}, \mathbf{y}) &= \prod_i P(x_i, y_i | x_{i-1}, y_{i-1}) \\ &= \prod_i P(x_i, |y_i) P(y_i | y_{i-1}) \end{aligned}$$

パラメータ推定

訓練データ（観測できた系列データの集合）

$$D = \{(\mathbf{x}^{(1)}, \mathbf{y}^{(1)}), (\mathbf{x}^{(2)}, \mathbf{y}^{(2)}), \dots, (\mathbf{x}^{(N)}, \mathbf{y}^{(N)})\}$$

最尤法

$$\begin{aligned} L &= \sum_{i=1}^N \log P(\mathbf{x}^{(i)}, \mathbf{y}^{(i)}) \\ &= \sum_{i=1}^N \sum_j \left(\log P(x_j^{(i)} | y_j^{(i)}) + \log P(y_j^{(i)} | y_{j-1}^{(i)}) \right) \\ &= \sum_{x,y} n((x, y), D) \log p_{x|y} + \sum_{y,y'} n((y', y), D) \log q_{y|y'} \end{aligned}$$

$n((x, y), D)$: D の中で x のラベルが y になった回数

$n((y', y), D)$: D の中でラベル y' の次のラベルが y になった回数

$$p_{x|y} = P(x|y) \quad q_{y|y'} = P(y|y')$$

L を最大化する、ラグランジュ乗数法を用いる

$$p_{x|y} = P(x|y) = \frac{n((x, y), D)}{\sum_x n((x, y), D)}$$

$$q_{y|y'} = P(y|y') = \frac{n((y', y), D)}{\sum_y n((y', y), D)}$$

HMM の推論

$\arg \max_y P(\mathbf{x}, \mathbf{y})$ を解く



組み合わせ最適化問題、動的計画法で解く

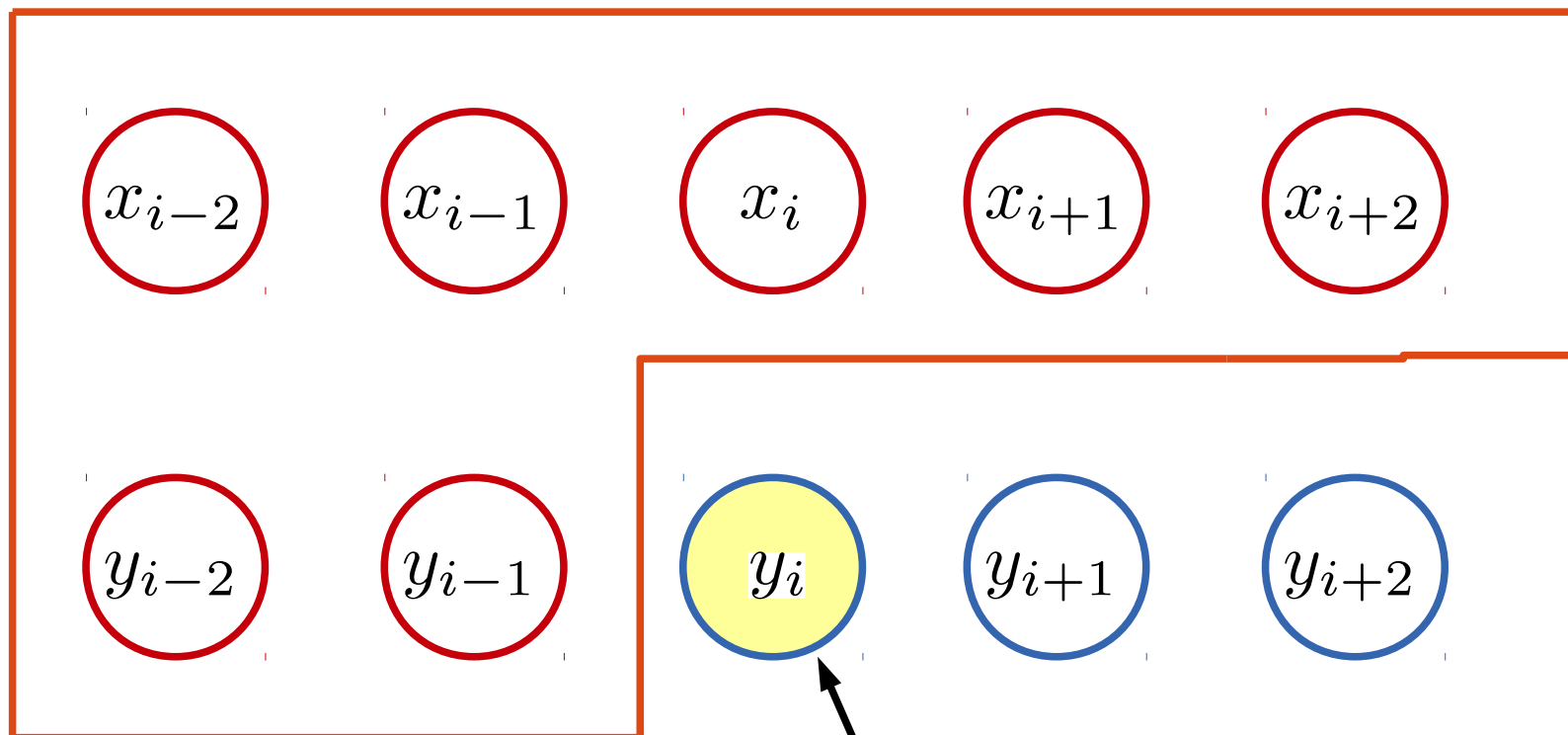


ビタビアルゴリズムが利用可

$$\log P(\mathbf{x}, \mathbf{y}) = \sum_i \log P(x_i, y_i | x_{i-1}, y_{i-1})$$

各 i について y_i を何にするか？

分類器の逐次適用



の情報を利用して これを決める