date423cs → 検索

神経回路網特論
2010 年 7 月 2 日

補足メモ

<u>先週</u>

1. 問題の説明した.
2. 動的計画法 (Dynamic Programming) を利用すし解決する方法を説明した.
3. その他，必要な数理的知識：
   正規分布，条件付き確率，同時確率，ベイズの公式，事後確率最大化

<u>今週，来週</u>： 基礎知識を復習しコードを書く

1. もとのデータ（$x_0, \cdots, x_{499}$）をマルコフ過程により生成する.
2. 標準正規分布（ガウス分布）にしたがうデータをコンピュータで生成し，$y_0, \cdots, y_{499}$ を生成する.

   これは何のためか → 自分で問題を作って解き，どれくらい正しく復元できるか試すため.

3. コード作成の際の注意点： データ構造とアルゴリズム

   データ構造を決めてコードを書く．原理ももちろん理解する（裏面参照）.

   以下は単なる例：

   ```
   int X[500];
   double Y[500];
   double C[500][2];
   int S[500][2];
   ```

   変数の説明：

   C[10][1]： 状態遷移を繰り返し $x_{10} = 1$ という状態にたどり着いた系列のうち，最も尤もらしい系列の発生確率（正確には，のようなもの）.

   S[10][1]： $x_{10} = 1$ であると仮定した場合に求めた x_9 の最適な値.

4. コードが完成したら，講義の web ページにあるデータを入力し，復元結果が，望ましい値と一致しているか確かめる.
5. データは 500 個ある.
6. この課題は「分かった」か「分からない」かはっきり区別が付きます.「分かったような気がする」というのはありません.「分かった」人に聞いてよいので，必ず自分で「わかった」と納得できるまで考えてください.

目的： ノイズを含んだデータ $y_0, y_1, \cdots$ から，もとの波形 $x_0, x_1, \cdots$ を復元する

与えられているもの

1. データ： $\vec{y} = (y_0, y_1, \cdots, y_{499})$
2. もとのデータ $x_0, x_1, \cdots$ の発生機構：$p_0 = 0.5, p_{00} = 0.99, p_{01} = 0.01, p_{11} = 0.97, p_{10} = 0.03$.

問題： これをもとに，最ももっともらしい解 $\vec{x}_{\mathrm{MAP}} = (x_0, x_1, \cdots, x_n)$ を探す

$$\vec{x}_{\mathrm{MAP}} = \underset{\vec{x}}{\operatorname{argmax}} \operatorname{Prob}(\vec{x} \mid \vec{y}) \tag{1}$$

※記号 argmax の使用例： $f(x) = -(x-1)^2$ を考える．$\tilde{x} = \operatorname{argmax}_x f(x) = 1$

$$\operatorname{Prob}(x_0, \cdots, x_T \mid y_0, \cdots, y_T) \quad = \quad \frac{\operatorname{Prob}(x_0, \cdots, x_T,\ y_0, \cdots, y_T)}{\operatorname{Prob}(y_0, y_1, \cdots, y_T)} \tag{2}$$

分母は $\vec{y}$ のみに関する確率． これは与えられているもので変えようがない $\to$ 分母は定数．最大値を求めるのには関係がない．したがって

$$\vec{x}_{\mathrm{MAP}} = \underset{\vec{x}}{\operatorname{argmax}} \operatorname{Prob}(\vec{x} \mid \vec{y}) = \underset{\vec{x}}{\operatorname{argmax}}\ \operatorname{Prob}(x_0, \cdots, x_T,\ y_0, \cdots, y_T) \tag{3}$$

$$= \underset{x_0, x_1, \cdots, x_T}{\operatorname{argmax}} \prod_{n=0}^{T} \operatorname{Prob}(y_n | x_n) \operatorname{Prob}(x_0) \prod_{n=1}^{T} \operatorname{Prob}(x_n | x_{n-1}) \tag{4}$$

これは以下の絵をイメージして考えると理解しやすい．

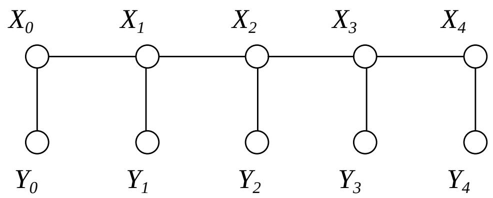

図 1: 確率変数の依存性．$y_0, y_1, \cdots,$ は与えられている．$x_0, x_1, \cdots \in \{0, 1\}$ の値を決めれば枝の値は決まる．n 個のノードがあれば 2^n とおりある．n が大きい場合は...

目的と解決策

1. 各枝には正の値がついている．値は枝の両端のノードの値に依存して決まる．
2. グラフを構成する枝すべてについて，枝の値を掛け算した結果を最大にしたい
3. 水平方向の枝の値： $\operatorname{Prob}(x_n | x_{n-1}) \geqq 0$，具体的には $p_{00}, p_{01}, p_{11}, p_{10}$ の値．
4. 垂直方向の枝の値： $\operatorname{Prob}(y_n | x_n) = \dfrac{1}{\sqrt{2\pi}\sigma} \exp\left\{ -\dfrac{(y_n - x_n)^2}{2\sigma^2} \right\} \Delta y \geqq 0$
5. 初期状態の確率は $\operatorname{Prob}(x_0) = 0.5$ つまり $p_0 = p_1 = 0.5$ にしておく．
6. 実際に動的計画法を用いて最も尤もらしい値をもとめるときは log をとって計算する．
7. 大小を判定することだけが重要な場合，無駄なことを計算しない効率のよいコードを書く．
8. x_n に関係ない項は共通なので無視できる．

$$\log\left(\frac{1}{\sqrt{2\pi}\sigma} \exp\left\{ -\frac{(y_n - x_n)^2}{2\sigma^2} \right\} \Delta y \right) = -\log(\sqrt{2\pi}\sigma) - \frac{(y_n - x_n)^2}{2\sigma^2} + \log(\Delta y)$$