

じゃんけん bot 案

概要

人 i には固有の手の出し方があるが、これが状態(前のターンでの手、勝敗など)で条件づけたときの確率分布で特徴づけられると考える。これを予想するモデルを考えたい。すなわち、各状態 s , 手 c に対して $P(c|s)$ (大きな行列)を推定したい。

これは結構難しい(と思う)。なぜなら、こちらが観測できるのは手を出した数であり、上の行列を直接観測することはおろか、観測するデータは一人当たり数十回程度であって、一つの成分を推定することすらほとんど不可能なので。

よって、手の出し方について特徴量を見出し、その事前分布を(推定し)活用することが必要そうだなとなる。

モデル

まず次のような設定を考える。

- 観測データ: $(C_t^{(i)}, S_t^{(i)})$: 人 i の t 番目のターンで選んだ手と状態 $i = 1, \dots, n, t = 1, \dots, m$

- データの生成モデル:

- 人 i に固有のパラメータ $\theta^{(i)}$ が混合正規分布から生成される:

$$\theta^{(i)} \sim \text{GMM} \left(\{(\pi_k, \mu_k, \Sigma_k)\}_{k=1}^K \right)$$

- ある行列 T が存在して、人 i の手の出し方は $P^{(i)}(c|s) = \text{softmax}(T\theta)$ の (c, s) 成分、で書けるとする(softmax をとる向きはご想像のとおりです)

このような定式化の下で、次が(本来)考えるべき最尤推定問題である。ここで $Q := \{(\pi_k, \mu_k, \Sigma_k)\}_k$ である。ただし、 ϕ_Q は Q でパラメータづけられた θ の確率密度関数である。

$$\max_T \max_Q \sum_i \log \left[\int \phi_Q(\theta) \prod_t [\text{softmax } T\theta]_{C_t^{(i)}, S_t^{(i)}} d\theta \right]$$

しかし、一人あたりのターン数 m が十分に大きいとき、ラプラス法の観点から

$$\max_T \max_Q \sum_i \log \left[\phi_Q(\hat{\theta}^{(i)}) \prod_t [\text{softmax } T\hat{\theta}^{(i)}]_{C_t^{(i)}, S_t^{(i)}} \right]$$

となる。ここで、 $\hat{\theta}^{(i)} \in \arg \max_{\theta} \sum_t \log [\text{softmax } T\theta]_{C_t^{(i)}, S_t^{(i)}}$ である。つまり、最尤推定する。MAP 推定量でないことに注意する。ラプラス法:

$$\int g(x) e^{mf(x)} dx \sim g(x_0) e^{mf(x_0)} \sqrt{\frac{2\pi}{m|f''(x_0)|}}$$

の対数をとった式で $\frac{1}{2} \log(f''(x_0))$ の項も落としており、 m について関数値(ラプラス法で言う $e^{mf(x)}$ の $f(x)$ のこと)が一定でないので多分本当は正しくない。 m が小さいときにより精度の良い収束先を募集してます(ϕ_Q を活用するとか?)。

また、実戦においては、相手のデータ $\{C_t, S_t\}_t$ を用いて θ を MAP 推定し:

$$\hat{\theta} \in \arg \max_{\theta} \log \phi_Q(\theta) + \sum_t \log [\text{softmax } T\theta]_{C_t, S_t}$$

これを用いて、手の出し方を $P^{(i)}(c|s) = [\text{softmax } T\hat{\theta}]_{c,s}$ で推定するものとする。

上の最尤推定の解き方

EM アルゴリズムみたいなことをする．次の 1,2 を繰り返す．

1. T を固定して Q を推定
2. Q を固定して T を推定

T を固定して $Q = \{(\pi_k, \mu_k, \Sigma_k)\}_{k=1}^K$ を推定

$\theta^{(i)}$ を最尤推定した後に

$$\max_Q \sum_i \log \left[\phi_Q(\hat{\theta}^{(i)}) \prod_t [\text{softmax } T\hat{\theta}^{(i)}]_{C_t^{(i)}, S_t^{(i)}} \right]$$

を解けばよい．これは \prod から後ろの部分は関係ないので，

$$\max_Q \sum_i \log [\phi_Q(\hat{\theta}^{(i)})]$$

という問題になる，

また，ターンが十分に多いとき， $\hat{\theta}^{(i)} \sim \text{GMM}(\{(\pi_k, \mu_k, \Sigma_k)\}_{k=1}^K)$ が期待できるので¹，これは単に混合正規分布のパラメータ $\{(\pi_k, \mu_k, \Sigma_k)\}_{k=1}^K$ を推定する問題と等価である．混合正規分布に対する EM アルゴリズムを回すことにより，これは可能である．

Q を固定して T を推定

$\theta^{(i)}$ を最尤推定した後に

$$\max_T \sum_i \log \left[\phi_Q(\hat{\theta}^{(i)}) \prod_t [\text{softmax } T\hat{\theta}^{(i)}]_{C_t^{(i)}, S_t^{(i)}} \right]$$

を解けばよい． $\hat{\theta}^{(i)}$ は T に依存していることに注意する．ここも， T と $\hat{\theta}^{(i)}$ を交互に求めることにする． T 平面と $\hat{\theta}^{(i)}$ 軸をイメージすると連続関数については割と妥当そうな印象．あるいは T と $\hat{\theta}^{(i)}$ 両方について最適化して $\hat{\theta}^{(i)} = \arg \max \dots$ の曲面に射影するのもありかも．とりあえず，前者のやり方に基づく， $\hat{\theta}^{(i)}$ を固定した上での T の最大化は，

$$\max_T \sum_i \sum_t \log \left([\text{softmax } T\hat{\theta}^{(i)}]_{C_t^{(i)}, S_t^{(i)}} \right)$$

という問題と等価であることがわかる．これくらいの問題なら最適化のライブラリで解くことができる(し，手でもう少し簡単な問題にできそう)．

その他の微小な工夫

T は一意に定まらないので， T の列和をとってできるベクトル v_T について正則化項 $\lambda_1 \|v_T\|_2$ をつける．安定化のため， T, θ の最適化時には， $P^{(i)}(c|s) = [\text{softmax } T\theta]_{c,s}$ についてエントロピー正則化を行っている．

現状の問題点

- θ のスケールが固定されていないので，解が不安定になる

余談

まずこれがうまくいくかわかんないけど，パラメータの事前分布を適当に制限した一般のマルコフ連鎖における確率遷移行列の推定にも使える?? (怪しい)

¹ 漸近正規性をうまく使えると嬉しいんだけど，MAP 推定では漸近正規性の有無は事前分布によってしまわらしい