

無限隠れマルコフモデルによるバスケットボールの戦術分析

慶應義塾大学大学院理工学研究科 竹原 大翼
慶應義塾大学理工学部 小林 景

1 はじめに

近年プロスポーツにもデータ解析が多く用いられるようになり、最適な戦術を過去のデータに基づいて決定することが一般的になったが、戦術そのものを理解するために過去のデータを分析することも競技水準の向上のために重要である。

そこで今回はバスケットボールにおいて、ある1チームの戦術を理解するために過去のデータを分析することを試みる。この時プレイバイプレイデータと呼ばれるデータを用いて解析する。プレイバイプレイデータはある1試合で起こった事象が時系列的に記録されたもので、第何クォーター残り時間何分何秒の時に、どちらのチームのどの選手が何をしたかと、その時の両チームの得点が記されている。

本研究ではプレイバイプレイデータにタグ付けることで各プレイの使い方を分類を考える。隠れマルコフモデル (HMM) を用いた場合、学習において状態 (隠れ変数) の取りうる値の数を事前に設定しなくてはならない。この設定手段として赤池 (1973) による赤池情報量規準 (AIC) を用いた方法 [2] が考えられるが、AIC はパラメータが識別可能な場合にのみ妥当性を持つため、HMM には一般に用いることができない。そこでこの問題の回避策として階層ディリクレ過程を用いた無限隠れマルコフモデル (iHMM) による解析が提案されている [1]。このモデルでは状態数が無限なものを考え、パラメータの事後分布からマルコフチェーンモンテカルロ法によって状態をサンプリングして、同時に状態数も決めることができる。今回はこの無限隠れマルコフモデルを用いて各プレイをタグ付けする。

2 手法および結果

HMM は隠れ状態、観測変数の二変数の確率モデルであって隠れ状態 (s_t) はマルコフ連鎖に従い、観測される変数 (y_t) の分布は隠れ状態に依存するよ

うなモデルである。iHMM は HMM の状態の遷移に関するパラメータの事前分布に階層ディリクレ過程を考えるようなベイズモデルであり、次式のように表される。

$$G_0 \sim \text{DP}(\gamma, G), \quad \pi_k | G_0 \sim \text{DP}(\beta, G_0), \quad \phi_k \sim H, \\ s_t | s_{t-1} \sim \text{Categorical}(\pi_{s_{t-1}}), \quad y_t | s_t \sim F(\phi_{s_t}).$$

ここで $\text{DP}(\gamma, G)$ は強度パラメータ γ 、基底分布 G のディリクレ過程を、 $F(\phi_k)$ はパラメータ ϕ_k によって特徴付けられた分布を表している。

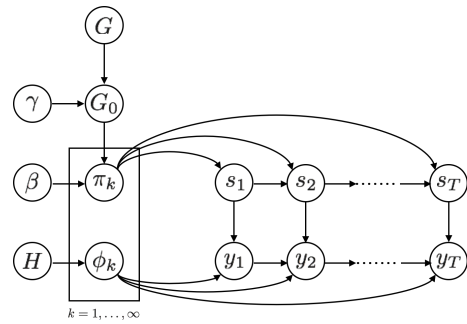


図1 無限隠れマルコフモデルのグラフィカルな表現

今回は観測される変数 (y_t) をある1チームの1シーズン全試合のプレイバイプレイデータとして、各プレイに対して隠れ変数 (s_t) を事後分布からサンプリングすることで各プレイにタグ付けをして分類し、そのプレイの目的を分析することを試みる。実際に戦術を分析した詳細な結果は当日に報告する。

参考文献

- [1] Beal, M. J., Ghahramani, Z. and Rasmussen, C. E. (2002), The infinite hidden Markov model. In Advances in neural information processing systems 577-584.
- [2] 池田思朗 (1993), 隠れマルコフモデルの生成に関する研究, 修士論文, 東京大学, 計数工学科.