1999年度日本オペレーションズ・リサーチ学会
秋季研究発表会

EMアルゴリズムによるMMPPパラメータ推定

長岡技術科学大学 *山田 雅裕 YAMADA Masahiro
長岡技術科学大学 中川 健治 NAKAGAWA Kenji

1 はじめに

MMPP(Markov Modulated Poisson Processes) は，ATM ネットワークにおける入力トラヒックの近似モデルとして用いられている．到着率，状態遷移率という2種類のパラメータに依存しており，これらパラメータを正確に決定することは ATM ネットワークにおける有効なトラヒック制御を可能とする．本報告では EM アルゴリズムを用いた MMPP パラメータ推定問題について検討する．従来の方法 [1] では，事象の到着時間間隔を観測データとして用い，MMPP への EM アルゴリズムの直接的な適用によりパラメータ推定を行う方法が提案されている．しかし MMPP は連続時間確率過程であり，時間パラメータに依存するため，直接的な適用は複雑な計算を伴う．そこで MMPP から離散時間確率過程へのモデル変換を考える．従来の研究 [2] では，モデルを MMPP から MMBP(Markov Modulated Bernoulli Processes) へ変換し，EM アルゴリズムを適用するパラメータ推定法が提案されている．この方法は，到着時間間隔を微小時間スロット h を用いて時間軸を離散化し，2値データ (h 中に事象の到着があれば1，なければ0を出力) を生成することにより観測データの変換を行い，EM アルゴリズムを適用してパラメータを推定する．しかしこの変換において，2値データ数が非常に多くなり，メモリ消費，計算量が多くなるという問題点がある．本研究ではこの問題を解決するために，MMPP を MMGP(Markov Modulated Geometric Processes) に変換し，EM アルゴリズムを適用してパラメータ推定を行う方法を提案する．シミュレーションによる従来法 [1]，[2] との比較を行い，提案法の有効性を示す．

2 EMアルゴリズム

EM アルゴリズムは，データに欠測値が存在した場合に，観測データと隠れ変数からなる完全データを考え，完全データの尤度関数の条件付き期待値を計算し，モデルパラメータの最尤推定を行う方法である．EM アルゴリズムには完全データの尤度関数の条件付き期待値を計算する E-step と最尤推定法を行う M-step がある．

3 MMPPパラメータ推定

本研究の提案法は，複雑な計算を避け高速なパラメータ推定を行うために，時間パラメータに依存し解析の複雑な連続時間 MMPP から，微小時間スロット h を用いて，時間パラメータに依存せず解析の容易な離散時間 MMGP へモデルを変換し，MMGP において EM アルゴリズムを適用してパラメータ推定を行う．まず MMPP から MMGP へのモデル変換について示し，その後 EM アルゴリズムを用いたパラメータ推定法について示す．

3.1 MMGP

MMGP は，状態 i における h 中に事象が到着する確率 b_i の幾何分布に従って事象が到着し，状態遷移確率 r_{ij} の離散時間マルコフ連鎖に従って状態を遷移するモデルである．

3.2 観測データ変換

MMPP から MMGP への観測データの変換を行う．S 状態 MMPP の N 個の到着時間間隔を微小時間スロット h により時間離散化し，到着間隔に含まれる時間スロット数の系列 $l_1^N = \{l_1, l_2, \cdots, l_N\}$ を生成する．提案法では離散後のデータ数は N 個である．しかし従来法 [2] の観測データである2値データの数は $\sum_{n=1}^{N} l_n$ であり，l_n は1以上の整数であるため，N よりも多くなる．観測データ数と比例して計算量が増加するため，提案法は従来法 [2] よりも少ない計算量で計算が可能である．

3.3 パラメータ変換

MMPP から MMGP へのパラメータの変換を次式の近似式を用いて行う．

$$b_i \simeq \lambda_i h \tag{1}$$

$$r_{ij} \simeq \frac{\sigma_{ij}}{\lambda_i} \tag{2}$$

ここで λ_i は MMPP の状態 i における到着率，σ_{ij} は MMPP の状態 i から j への状態遷移率を表す．3.4 節の

パラメータ推定法により MMGP のパラメータを推定した後, (1), (2) を用いて MMPP パラメータへ戻す.

3.4 EM アルゴリズムパラメータ推定

S 状態 MMGP において, 観測データ系列 l_1^N から EM アルゴリズムを適用してパラメータ b_i, r_{ij} を推定する.

E-step: 状態遷移の軌跡が与えられた時の尤度関数の条件付き期待値 $Q(b,r)$ を計算する.

$$Q(b,r) = \sum_{j=1}^{S}\sum_{i=1}^{S}\sum_{n=1}^{N}\xi_n(i,j)\log r_{ij}$$

$$+\sum_{i=1}^{S}\sum_{n=1}^{N}\gamma_n(i)[(l_n-1)\log\{1-b_i\}+\log b_i] \tag{3}$$

ここで, $\xi_n(i,j)=P(s_n=i,s_{n+1}=j|l_1^N,\Phi^{(k)})$, $\gamma_n(i)=P(s_n=i|l_1^N,\Phi^{(k)})$ であり, $\Phi^{(k)}$ は k 回目の繰り返しで得られたパラメータ ($b_i{}^{(k)}$, $r_{ij}^{(k)}$) の組を表す.

M-step: $Q(b,r)$ をパラメータ b_i, r_{ij} について最大化する.

$$\frac{\partial Q(b,r)}{\partial \Phi} = 0$$

より,

$$b_i^{(k+1)} = \frac{\sum_{n=1}^{N}\gamma_n(i)}{\sum_{n=1}^{N}\gamma_n(i)l_n} \tag{4}$$

$$r_{ij}^{(k+1)} = \frac{\sum_{n=1}^{N-1}\xi_n(i,j)}{\sum_{n=1}^{N-1}\gamma_n(i)} \tag{5}$$

ここで Φ はパラメータ (b_i, r_{ij}) の組を表す. パラメータが収束するまで E-step, M-step を繰り返す. 収束後, 式 (1), (2) を用いて MMPP のパラメータへ逆変換する.

4 シミュレーション結果と考察

2 状態 MMPP において到着時間間隔 2000 個から提案法, 従来法 [1], [2] の 3 つの方法によってパラメータ λ_i, σ_{ij} を推定し, 結果の比較を行った. パラメータ真値は $\lambda_1=100.0$, $\lambda_2=10.0$, $\sigma_{12}=10.0$, $\sigma_{21}=1.0$ とし, 微小時間スロットを $h=0.001$ とする. EM アルゴリズムは繰返しアルゴリズムであるためパラメータの初期値を必要とする. 今回のシミュレーションでは初期値を従来法 [2] の k-means クラスタリング法を用いて決定した. k-means クラスタリング法は到着時間間隔を観測データとし, k-means クラスタリングを用いてパラメータ推定を行う方法である. この方法は推定精度は良くないが, 実行が非常に高速であるという特徴を持つことから, 高速に, かつ真値にある程度近い値を初期値として決定することが出来る.

表 1 に実行回数 30 回での実行時間を示す. λ_i, σ_{ij} の収束展開は, 3 つの方法とも真値に収束しており, 収束精度の点では大きな差異がないため, 省略する. 表 1 の実行時間を比較すると, 提案法が非常に短い時間で計算を行っている事が分かる. 提案法は, 従来法 [1] で必要とする数値積分などの複雑な計算を行う必要が無く, 計算の実行が容易であるため, 従来法 [1] よりも短い実行時間で計算が行えたと考えられる. また提案法は, 従来法 [2] よりも非常に少ない計算量, メモリ消費量で計算を行うことができる. 提案法の観測値数は N, 従来法 [2] の観測値数は $\sum_{n=1}^{N} l_n$ であり l_n は 1 以上の整数であるため, 提案法の観測値数の方が少なくなる. 観測値数と計算量, メモリ消費量が比例関係にあるため, 提案法の計算量, メモリ消費量は従来法 [2] よりも少ない. 今回の実験では提案法の観測値数は従来法 [2] の約 1/50 であり, それに伴い計算量, メモリ消費量が少なくなっている. その結果, 計算時間を短縮出来たものと考えられる.

5 まとめ

MMPP パラメータ推定に EM アルゴリズムを適用する方法について検討した. シミュレーション結果より, 提案法が精度良く, かつ高速にパラメータ推定を行うことを確認した. 今後の課題は, $\lambda=0$ の時に推定精度が悪化する問題点, 実トラヒックに提案法を適用した時のパラメータ推定について検討する.

参考文献

[1] T. Rydén ; "An EM algorithm for estimation in Markov modulated Poisson processes" , Computational Statistics and Data Analysis , vol. 2 , p. 431-447 (1996)

[2] Li Deng and Jon W. Mark ; "Parameter estimation for Markov modulated Poisson processes via the EM algorithm with time descretization" , Telecommunication Systems , vol. 1 , p.321-338 (1993)

表 1: 実行時間の比較

提案法	従来法 [1]	従来法 [2]
2.3[sec]	31.9[sec]	73.3[sec]