

複数の到着流をもつ単一サーバ待ち行列

滝根 哲哉

1 はじめに

待ち行列理論は、共有資源に対する利用要求が確率的に発生するという仮定の下で、資源競合問題を抽象化した数学モデルの構築と解析に関する理論である。通信網への応用を意識したとき、待ち行列理論は通信トラヒック理論とも呼ばれており、電話網の設計問題を解決するための枠組として、20世紀初頭に研究が開始されている。それ以降、パケット交換網、衛星通信、LAN、狭帯域ISDN、広帯域ISDN(ATM)などの新しい通信網の出現に伴い新しいモデルが導入され、理論の発展が加速されてきた。

近年の通信網の高速化に伴い、従来、個別網に収容されていた多様なメディア(音声、動画、データなど)を単一の通信網に収容する統合サービス網の技術開発が進められている。統合サービス網に収容されるトラヒックは、メディアあるいはコンテンツによって、それぞれ固有の特徴を持っており、それゆえ、通信網内に設置された交換機では、性質の異なるトラヒックが多重化されて伝送されることになる。

このような技術動向に刺激され、複数の到着流をもつ待ち行列の研究がこの10数年来、精力的に行われてきた[10]。これらは固定長の伝送単位をもつATM網に刺激され、近年、急速に研究の進んだ離散時間待ち行列モデルに関する研究と、このような状況を一般的に記述した連続時間待ち行列モデルに関する研究に分けることができる。本稿では、連続時間モデルに関して最新的话题を紹介する。離散時間モデルに関しては[10]を参照。

2 連続時間待ち行列モデル

複数の到着流をもつ待ち行列を考察する際には、重畳後の到着過程に対する解析的取り扱いが容易になるように各到着流を記述できると都合が良い。各到着流が独立なポワソン過程に従う場合、重畳後の到着過程もまたポワソン過程となるが、通信網における応用では、必ずしも伝送単位の発生がポワソン過程に従うとは仮定できず、より広い範囲の到着過程を記述できるモデルが必要となる。

そのような要請を満たす到着過程の1つが、2.1章で紹介するマルコフ型到着過程(MAP: Markovian ar-

rival process)[3]である。MAPには、独立なMAPの重畳がMAPで表現できることや、あらゆる定常単点過程を任意の精度で近似できるなどの利点がある。MAPと等価な到着過程は他にも幾つか提案されているが、特にMAPは待ち行列の解析という観点から最も簡潔な表現をもつため、急速に浸透した。

到着過程がMAPに従う単一サーバ待ち行列は、通常、アルゴリズム的解法の一つである行列解析法(matrix analytic method)によって解析される。単独の待ち行列に対するアルゴリズム的解法は、昨年、OR誌で企画された「待ち行列研究の新しい潮流」[5,6]でも紹介されているように、システム内容数に対応する離散確率変数とシステムの状態変化を記述するために必要な情報を保持する補助的な離散確率変数の組からなる2変数マルコフ連鎖に対して、その構造を利用して定常解を求める数値アルゴリズムと見ることができる。

本稿の後半では、サービス時間分布の異なる複数のMAP到着流を収容する単一サーバ待ち行列に対する筆者の研究成果を紹介する。このモデルの興味深い点は、2.1章で紹介するような従来の行列解析法の枠組では取り扱うことが出来ないことである。このような待ち行列の最も単純な例は例えば以下のようなものである。

例 2つの到着流を収容する単一サーバ待ち行列

到着流1は平均到着間隔が λ_1^{-1} のポワソン到着、サービス時間分布関数 $h_1(x)$ をもつとする。一方、到着流2は平均到着間隔が λ_2^{-1} の2ステージーラン分布に従い、サービス時間分布関数 $h_2(x)$ をもつとする。客は先着順サービス規律(FIFO)に従ってサービスされる。待ち部屋の容量は十分に大きく、到着した客は全てシステム内に入ることができる。

この待ち行列における到着過程を図で表現すると以下のようなになる。

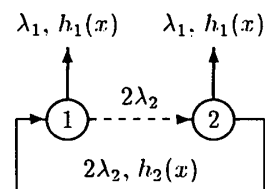


図1: ポワソン到着と2ステージーラン到着の重畳もし、 $h_1(x) = h_2(x)$ ならば、 n 番目の客の離脱直

たきね てつや 京都大学大学院情報学研究所数理工学専攻
〒606-8501 京都市左京区吉田本町
E-mail: takine@kuamp.kyoto-u.ac.jp

後のシステム内総客数 Q_n とその時点でのアーラン分布がいずれのステージにあるかという確率変数 S_n の組 (Q_n, S_n) は 2 変数マルコフ連鎖をなす。しかし、 $h_1(x) \neq h_2(x)$ ならば、客の到着とサービス時間分布の間に相関が生じるため、総客数 Q_n と S_n の組はマルコフ連鎖を構成しない。実際、客数の挙動を表すマルコフ連鎖を構築しようとすれば、サービス中ならびに待っている個々の客がそれぞれ、どちらの到着流から来たかという情報を完全に保持する必要がある¹。

この例から分かるように、サービス時間分布の異なる複数の非ポワソン到着流を収容する単一サーバ待ち行列では、客数の挙動を表現するマルコフ連鎖は極めて複雑な構造をもつため、そのような定式化による解析はほぼ不可能である。以下では、最初に M/G/1 型マルコフ連鎖とその解法に触れた後、例で挙げたような待ち行列モデルの扱いが可能な枠組を、解析手法に関する着眼点を中心に紹介する。

2.1 M/G/1 型マルコフ連鎖と MAP [3,4]

単一サーバ待ち行列において、 n 番目の客の離脱直後のシステム内総客数を Q_n 、 n 番目の客のサービス時間中に到着する客数を A_n で表すと

$$Q_{n+1} = \max(Q_n - 1, 0) + A_{n+1}, \quad n = 0, 1, \dots \quad (1)$$

が成立する。特に到着率 λ_k のポワソン到着、独立同一なサービス時間分布 $h_k(x)$ をもつ独立な P 個の到着流の重畳を収容する場合、

$$\lambda = \sum_{k=1}^P \lambda_k, \quad h(x) = \sum_{k=1}^P \lambda_k h_k(x) / \lambda$$

とすれば、この待ち行列は到着率 λ のポワソン到着、独立同一なサービス時間分布 $h(x)$ をもつ M/GI/1 待ち行列と見なすことができ、結果として $\{A_n\}$ は独立同一分布に従う確率変数列となる。この場合、 $a_k = \Pr(A_n = k)$ とすると、 a_k は

$$a_k = \int_0^\infty e^{-\lambda x} \frac{(\lambda x)^k}{k!} dh(x), \quad k = 0, 1, \dots$$

で与えられる。よって $\{Q_n\}$ は遷移確率行列

$$\begin{pmatrix} a_0 & a_1 & a_2 & a_3 & \dots \\ a_0 & a_1 & a_2 & a_3 & \dots \\ 0 & a_0 & a_1 & a_2 & \dots \\ 0 & 0 & a_0 & a_1 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \quad (2)$$

をもつマルコフ連鎖で表現される。平均サービス時間を μ^{-1} としたとき、 $\rho = \lambda/\mu < 1$ ならば、このマルコフ連鎖は定常状態確率をもつ。離脱直後に k 人の客がいる定常状態確率を q_k ($k = 0, 1, \dots$) とすると、

$q_0 = 1 - \rho$ であり、 q_k ($k = 1, 2, \dots$) は

$$q_k = \left(q_0 \bar{a}_k + \sum_{j=1}^{k-1} q_j \bar{a}_{k-j+1} \right) / (1 - \bar{a}_1)$$

より順次計算される²。ただし $\bar{a}_k = \sum_{j=k}^\infty a_j$ である。

上記の例では、ポワソン到着が仮定されていた。この仮定を取り除き、より一般的な分布を扱うためのモデルとして相型分布 (phase-type distribution) やマルコフ型到着過程 (MAP) がある。これらはいずれも、有限状態をもつ連続時間マルコフ連鎖を考え、状態遷移が起こったとき、予め決められた確率で客が到着するというものである。例えば MAP は以下のように定義される [3]。状態集合 $\mathcal{M} = \{1, \dots, M\}$ をもつ既約な有限状態連続時間マルコフ連鎖 $\{S(t)\}$ を考える。各状態 $i \in \mathcal{M}$ での滞在時間は平均 η_i^{-1} の指数分布に従う。状態 i での滞在終了後、確率 $p_{i,j}$ ($j \in \mathcal{M}, j \neq i$) で到着を伴わず状態 j へ遷移する。また、確率 $q_{i,j}$ ($j \in \mathcal{M}$) で 1 人の客の到着を伴って状態 j へ遷移する。ただし

$$\sum_{j \in \mathcal{M}, j \neq i} p_{i,j} + \sum_{j \in \mathcal{M}} q_{i,j} = 1$$

である。このとき、 (i, j) 要素 $C_{i,j}$ 、 $D_{i,j}$ がそれぞれ

$$C_{i,j} = \eta_i p_{i,j} \quad (j \neq i), \quad C_{i,i} = -\eta_i, \quad D_{i,j} = \eta_i q_{i,j}$$

で与えられる $M \times M$ 行列 C, D を用いて、この到着過程を表現 (C, D) をもつ MAP という。定義より D の (i, j) 要素は到着を伴う状態 i から状態 j への遷移率を表し、 C の非対角 (i, j) 要素は到着を伴わない状態 i から状態 j への遷移率を表す。例えば、前述のポワソン流と 2 ステージアーラン流の重畳の場合、

$$C = \begin{pmatrix} -\lambda_1 - 2\lambda_2 & 2\lambda_2 \\ 0 & -\lambda_1 - 2\lambda_2 \end{pmatrix}, \quad D = \begin{pmatrix} \lambda_1 & 0 \\ 2\lambda_2 & \lambda_1 \end{pmatrix} \quad (3)$$

となる (図 1 参照)。以下では、 $S(t)$ の値を時刻 t における MAP の相と呼ぶ。なお、 $N(t)$ を時間間隔 $(0, t]$ の間に到着した客数とすると、 $\exp[(C+zD)t]$ の (i, j) 要素は

$$E \left[z^{N(t)} 1_{\{S(t)=j\}} \mid S(0)=i \right]$$

を表す。ここで $1_{\{x\}}$ は事象 x の指示関数である。また、MAP の相の定常確率ベクトルを π としたとき、 π は

$$\pi(C+D) = 0, \quad \pi e = 1$$

より決定され、これを用いて平均到着率 λ は次式で与えられる。

$$\lambda = \pi D e$$

¹ サービス時間が最も扱いやすい (平均の異なる) 指数分布に従う場合でも同様である

² 定常な M/G/1 待ち行列の場合、離脱直後の客数分布は任意時点における客数分布に等しい

ここで e は全ての要素が 1 の列ベクトルである。

さて, (1) で考えた単一サーバ待ち行列における Q_n に対して, ポワソン到着の仮定を取り除き, 重畳後の客の到着が表現 (C, D) をもつ MAP に従うと仮定する³. ただし, 客のサービス時間分布は全て独立同一で, 分布関数 $h(x)$ に従うとする (これを MAP/GI/1 待ち行列という). このとき, A_n は独立な確率変数列ではないが, 離脱直後の MAP の相が与えられると, 過去の履歴とは独立に A_n の確率分布が定まる. そこで MAP/GI/1 待ち行列において S_n を n 番目の離脱直後の MAP の相を表すとする. このとき $\{(Q_n, S_n)\}$ は 2 変数マルコフ連鎖となり, 遷移確率行列は

$$\begin{pmatrix} B_0 & B_1 & B_2 & B_3 & \dots \\ A_0 & A_1 & A_2 & A_3 & \dots \\ O & A_0 & A_1 & A_2 & \dots \\ O & O & A_0 & A_1 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \quad (4)$$

で与えられる. ここで A_k ($k = 0, 1, \dots$) は

$$A^*(z) = \sum_{k=0}^{\infty} A_k z^k = \int_0^{\infty} \exp[(C + zD)x] dh(x)$$

を満たす非負の $M \times M$ 行列で与えられ⁴, $B_k = (-C)^{-1} D A_k$ となる. A_k の (i, j) 要素はサービスの開始時点において MAP の状態 S_n が i であったという条件の下で, サービス時間内に k 人の客が到着し, サービスの終了時点での MAP の状態 S_{n+1} が j である確率を表している. また B_k に現れる行列 $(-C)^{-1} D$ の (i, j) 要素は, ある時点で MAP の状態が i であったという条件の下で, 次の到着が起こった直後の MAP の状態が j である確率を表している.

MAP/GI/1 における遷移確率行列 (4) は M/GI/1 から得られる遷移確率行列 (2) をブロック化したものと見ることができる. さらに, この 2 変数マルコフ連鎖 $\{(Q_n, S_n)\}$ は, (i) 境界部分を除いて空間的に同質 ($Q_n \geq 1$ のとき, 客数が $k-1$ 人増える遷移確率は A_k で表現される), かつ (ii) Q_n は 1 回の遷移で高々 1 つしか減らない, という特徴をもつ. 2 変数マルコフ連鎖がこのような特徴をもつ (すなわち遷移確率行列が式 (4) のような形をとる) とき, M/GI/1 型マルコフ連鎖と呼ばれる. M/GI/1 型マルコフ連鎖はこれらの特徴を利用して, 定常確率を計算するアルゴリズムが知られている. 詳細は [4, 16] を参照.

2.2 MAP/GI/1 の定常客数分布 [3,13]

以下では MAP 固有の性質を利用した MAP/GI/1 待ち行列の客数解析を簡単に紹介する. 定常状態において, 客の離脱直後の客数ならびに MAP の相をそれ

³MAP の重畳は MAP で表現できることに注意, 2.3 章参照
⁴ A_k の計算アルゴリズムに関しては [8] 参照

ぞれ Q, S^* とする. さらに $q_{k,j} = \Pr(Q = k, S^* = j)$ を j 番目の要素にもつ列ベクトルを q_k で表す. このとき, 遷移確率行列 (4) ならびに $B_k = (-C)^{-1} D A_k$ より $q^*(z) = \sum_{k=0}^{\infty} z^k q_k$ は次式を満たすことが分かる.

$$q^*(z)[zI - A^*(z)] = q_0(-C)^{-1}(C + zD)A^*(z)$$

一方, Y を任意時点におけるシステム内容数とし, S を MAP の相を表すとする. このとき $y_{k,j} = \Pr(Y = k, S = j)$ を j 番目の要素にもつ列ベクトルを y_k とすると, $y^*(z) = \sum_{k=0}^{\infty} z^k y_k$ は

$$y^*(z)(C + zD) = \lambda(z-1)q^*(z)$$

を満たすことが知られている [3,11]. これらより

$$y^*(z)[zI - A^*(z)] = (z-1)y_0 A^*(z)$$

を得る. この式は y_k ($k = 0, 1, \dots$) が, 遷移確率行列

$$\begin{pmatrix} A_0 & A_1 & A_2 & A_3 & \dots \\ A_0 & A_1 & A_2 & A_3 & \dots \\ O & A_0 & A_1 & A_2 & \dots \\ O & O & A_0 & A_1 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

をもつ M/GI/1 型マルコフ連鎖の定常解と同一であることを示している [13]. それゆえ, 通常の M/GI/1 型マルコフ連鎖に対する解法を用いれば定常解を得ることができる. 特に, MAP/GI/1 の場合には, この解法で中心的な役割を果たす確率行列 G は

$$G = \int_0^{\infty} \exp[(C + DG)x] dh(x) \quad (5)$$

を満たす. この確率行列 G の不変確率ベクトル g を用いて y_0 は $y_0 = (1 - \rho)g$ で与えられ, y_k ($k = 1, 2, \dots$) は

$$y_k = \left(y_0 \bar{A}_k + \sum_{j=1}^{k-1} y_j \bar{A}_{k-j+1} \right) (I - \bar{A}_1)^{-1}$$

により順次計算される. ただし

$$\bar{A}_k = \sum_{j=k}^{\infty} A_j G^{j-k}, \quad k = 1, 2, \dots$$

である. 待ち時間に関しては 2.4 章を参照.

2.3 複数の到着流の記述 [1,3,7]

表現 $(\tilde{C}_1, \tilde{D}_1)$ ならびに $(\tilde{C}_2, \tilde{D}_2)$ をもつ独立な 2 つの MAP の重畳を考える. それぞれの MAP の相がとる状態集合を M_1, M_2 とすると, 重畳後の到着過程において相がとる状態集合 M はこれらの直積集合 $M_1 \times M_2$ となる. すなわち重畳された到着過程は, 表現 (C, D) をもつ MAP となる. ただし

$$C = \tilde{C}_1 \otimes I_2 + I_1 \otimes \tilde{C}_2, \quad D = \tilde{D}_1 \otimes I_2 + I_1 \otimes \tilde{D}_2$$

である。ここで \otimes はクロネッカ積であり、 I_i は M_i の要素数と同じ次元の単位行列である。特に D を構成している2つの要素はそれぞれの到着流からの到着を表現していることに注意する。

さて、複数の到着流を明示的に記述するためにマーク付きMAPが提案されている[1]。これは、到着を伴う遷移率を表現する行列 D を各到着流毎に個別に用意するというものである。一般に、マーク付きMAPを用いて P 個の到着流を記述する際、MAP にならって (C, D_1, \dots, D_P) で表現される。ここで D_k ($k = 1, \dots, P$) の (i, j) 要素は、 k 番目の到着流からの客がMAPの相の i から j への遷移に伴って到着する率を与えている。行列 C の定義はMAPの場合と同じである。例えば上の例では、 $D_1 = \tilde{D}_1 \otimes I_2$ 、 $D_2 = I_1 \otimes \tilde{D}_2$ となる。また、1章の例において、ポワソン流を到着流1、2ステージアーラン流を到着流2としたとき、以下のように記述される(図1, (3)参照)。

$$D_1 = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_1 \end{pmatrix}, \quad D_2 = \begin{pmatrix} 0 & 0 \\ 2\lambda_2 & 0 \end{pmatrix}$$

$N_k(t)$ を時間間隔 $(0, t]$ の間に到着流 k から到着した客数とすると $\exp[(C + z_1 D_1 + \dots + z_P D_P)t]$ の (i, j) 要素は

$$E \left[z_1^{N_1(t)} \dots z_P^{N_P(t)} 1_{\{S(t)=j\}} \mid S_0 = i \right]$$

となる。

マーク付きMAPの表現はサービス時間に関する情報を含まない。前述の例で示したような、各到着流がそれぞれ異なるサービス時間分布をもつ場合を扱うとき、それらを明示的に表し、かつ簡潔な(例えば図1と等価なものを表現する)記法が必要となる。そこでマーク付きMAPを拡張し、 $(C, D_1(x), \dots, D_P(x))$ の形をもつ表現を導入する。ここで $D_k(x)$ ($k = 1, \dots, P$) の (i, j) 要素は、サービス時間が x 以下であるような k 番目の到着流からの客が、MAPの相の i から j への遷移に伴って到着する率を与えている。行列 C の定義はマーク付きMAPの場合と同じである。例えば図1の到着流は以下のように表現される。

$$D_1(x) = \begin{pmatrix} \lambda_1 h_1(x) & 0 \\ 0 & \lambda_1 h_1(x) \end{pmatrix}, \quad D_2(x) = \begin{pmatrix} 0 & 0 \\ 2\lambda_2 h_2(x) & 0 \end{pmatrix}$$

特に $D_k = D_k(\infty)$ ($k = 1, \dots, P$) に注意する。

2.4 系内仕事量分布と待ち時間 [7]

本章以降は、表現 $(C, D_1(x), \dots, D_P(x))$ の到着流を収容するFIFO単一サーバ待ち行列の解析を考える。まず、この待ち行列におけるMAPの相の定常確率分布 π は

$$\pi \left(C + \sum_{k=1}^P D_k \right) = 0, \quad \pi e = 1$$

より決定される。さらに、これを用いて、この待ち行列の利用率 ρ は

$$\rho = \pi \int_0^\infty x dD(x) e$$

で与えられる。ただし $D(x) = D_1(x) + \dots + D_P(x)$ である。以下では $\rho < 1$ を仮定する⁵。

最初に、この待ち行列の系内仕事量の挙動を考える。ある時点における系内仕事量とは、その時点以降、到着がなかったと仮定したとき、システムが空になるまでの時間で与えられる。

もし、一旦到着した客はサービスが終了するまで退去することがなく、かつ仕事がある限りサーバは稼働し続けるとすると、系内仕事量はサービスの順序とは無関係な量となる(これを仕事量保存型の待ち行列という)。この場合、時刻 t における系内仕事量 V_t は $V_t > 0$ のとき率1で減少し、客の到着があると、その客のサービス時間分だけ上にジャンプする。

さて、時刻0で系内仕事量が x であり、かつ、MAPの相が i であるという条件の下で、これ以降初めてシステムが空になった時点でのMAPの相が j である確率を (i, j) 要素にもつ行列 $P(x)$ を考える。定義より $o(\delta x)$ の項を無視すると

$$P(x + \delta x) = P(x) \left[I + C\delta x + \int_0^\infty dD(y) \delta x P(y) \right]$$

が成立し、これより

$$\frac{d}{dx} P(x) = P(x) \left[C + \int_0^\infty dD(y) P(y) \right]$$

すなわち $P(x) = \exp(Qx)$ を得る。ただし、行列 Q は

$$Q = C + \int_0^\infty dD(x) \exp(Qx)$$

を満たす。この Q はシステムが空である時間区間みに注目した場合のMAPの相の遷移率行列を表している。ここで $Q_0 = C$ とし、

$$Q_{n+1} = C + \int_0^\infty dD(x) \exp(Q_n x), \quad n = 0, 1, \dots$$

としたとき、 Q_n は要素毎に単調に Q へ収束する。もし、全ての到着流が独立同一なサービス時間分布 $h(x)$ を持つならば(すなわちMAP/GI/1ならば)、(5)で与えられる G を用いて $Q = C + DG$ となる。

定常状態における系内仕事量を V とし、MAPの相を S としたとき、 j 番目の要素が $\Pr(V \leq x, S = j)$ で与えられるベクトルを $v(x)$ とする。最初に $v(0)$ を考える。システムが空である確率は $1 - \rho$ であるので $v(0)e = 1 - \rho$ である。よって $\kappa Q = 0$ かつ $\kappa e = 1$ を満たす κ を用いて $v(0) = (1 - \rho)\kappa$ を得る。

⁵これにより定常状態の存在が保証される

次に $v(x)$ の満たす微分積分方程式を導く。微小な時間間隔 δt の間の遷移を考えることにより

$$v(x) = v(x + \delta t)(I + C\delta t) + \int_0^x v(x - y + \delta t)dD(y)\delta t + o(\delta t)$$

を得る。これより

$$\frac{d}{dx}v(x) + v(x)C + \int_0^x v(x - y)dD(y) = 0$$

となる。 $v^*(s)$ を $v(x)$ の Laplace-Stieltjes 変換 (LST) とすると、上の微分方程式ならびに $v(0) = (1 - \rho)\kappa$ より

$$v^*(s)[sI + C + D^*(s)] = (1 - \rho)s\kappa \quad (6)$$

を得る。ただし $D^*(s)$ は $D(x)$ の LST である。特にサービス時間分布が独立同一な分布関数 $h(x)$ (LST $h^*(s)$) をもつ MAP/GI/1 の場合、

$$v^*(s)[sI + C + h^*(s)D] = (1 - \rho)s\kappa$$

となることに注意する。(6) の両辺を微分することにより、定常状態における系内仕事量分布の積率を求めることができる。

次に各到着流から来る客の実待ち時間を考える。客の待ち時間は到着時点における系内仕事量に等しいことに注意する。しかし、客は MAP に従って到着するため、MAP の相によって到着する頻度が異なる。よって定常状態において客が到着時に見る系内仕事量は、MAP の各相における到着率と各相の発生頻度の積の比で重み付けられたものとなる。

定常状態における k 番目 ($k = 1, \dots, P$) の到着流から来る客の待ち時間を W_k とする。さらに定常状態において k 番目 ($k = 1, \dots, P$) の到着流から来た客の到着直後の MAP の相を S_k とする。このとき j 番目の要素が $\Pr(W_k \leq x, S_k = j)$ であるようなベクトル $w_k(x)$ は

$$w_k(x) = \frac{v(x)D_k}{\pi D_k e}$$

で与えられる。さらに、定常状態における k 番目 ($k = 1, \dots, P$) の到着流から来る客のシステム内滞在時間を R_k としたとき、 j 番目の要素が $\Pr(R_k \leq x, S_k = j)$ であるようなベクトル $r_k(x)$ は

$$r_k(x) = \frac{\int_{0-}^x dv(y)D_k(x - y)}{\pi D_k e}$$

で与えられる。

2.5 不変式と待ち行列長分布 [14,15]

最後に待ち行列長分布を考える。最初に述べたように、客数の振舞いを表現するマルコフ連鎖は非常に複雑な構造を持っており、このような接近法では解析的

な結果を得ることが出来ない。ここでは視点を変え、客が FIFO でサービスされるという仮定の下で、待ち行列長分布を系内仕事量分布、あるいは待ち時間分布と関連づけることを考える。以下では k 番目の到着流から来た客をクラス k と呼ぶ。

一般に、表現 (C, D_1, \dots, D_P) をもつマーク付き MAP を收容する定常な待ち行列システムにおいて、 $L_k(t)$ を時刻 t におけるクラス k のシステム内客数とし、 $N_k(t)$ 、 $J_k(t)$ をそれぞれ、時間間隔 $(0, t]$ の間のクラス k の到着客数ならびに離脱客数とする。このとき

$$(L_1(t), \dots, L_P(t)) = (L_1(0), \dots, L_P(0)) + (N_1(t), \dots, N_P(t)) - (J_1(t), \dots, J_P(t))$$

が成立し、 $(N_1(t), \dots, N_P(t))$ が表現 (C, D_1, \dots, D_P) をもつ定常なマーク付き MAP に支配されているとする。このとき、定常状態における $L_k(t)$ の値を L_k 、MAP の相を S とし、

$$l_j(z) = E \left[z_1^{L_1} \cdots z_P^{L_P} 1_{\{S=j\}} \right]$$

を j 番目の要素にもつ結合母関数ベクトルを $l(z)$ とする。また、定常状態において、クラス k の客の離脱直後における $L_n(t)$ の値を $Q_n(k)$ とし、

$$q_{k,j}(z) = E \left[z_1^{Q_1(k)} \cdots z_P^{Q_P(k)} 1_{\{S=j\}} \right]$$

を j 番目の要素にもつ結合母関数ベクトルを $q_k(z)$ とする。このとき、サービス時間分布の構造、サービス規律あるいはサーバ数などに依存せず、非常に広い範囲のモデルに対して

$$l(z) \left(C + \sum_{k=1}^P z_k D_k \right) = \sum_{k=1}^P \lambda_k (z_k - 1) q_k(z) \quad (7)$$

が成立する [15]。

この不変式を用いて待ち行列長分布を導出する。定常状態において、クラス k の待ち客数を X_k 、MAP の相を S としたとき、 j 番目の要素 $x_j(z)$ (ただし $z = (z_1, \dots, z_P)$) が

$$x_j(z) = E \left[z_1^{X_1} \cdots z_P^{X_P} 1_{\{S=j\}} \right]$$

で与えられるような結合母関数ベクトルを $x(z)$ で表す。一方、定常状態において、クラス k の客の待ち部屋からの離脱直後 (サービス開始時点) におけるクラス n の待ち客数を $X_n(k)$ とし、クラス k の客の待ち部屋からの離脱直後における MAP の相を S_k^* としたとき、 j 番目の要素 $x_{k,j}^D(z)$ が

$$x_{k,j}^D(z) = E \left[z_1^{X_1(k)} \cdots z_P^{X_P(k)} 1_{\{S_k^*=j\}} \right]$$

で与えられるような結合母関数ベクトルを $x_k^D(z)$ で表す。クラス k の客の待ち部屋からの離脱直後の待ち客は、FIFO でサービスが行われているため、クラ

ス k の待ち時間の間に到着した客であることを注意すると

$$x_k^P(z) = \int_{0^-}^{\infty} dw_k(x) \exp \left[\left(C + \sum_{i=1}^P z_i D_i \right) x \right]$$

を得る。よって (7) より

$$\begin{aligned} x^*(z) & \left(C + \sum_{k=1}^P z_k D_k \right) \\ & = \sum_{k=1}^P \lambda_k (z_k - 1) \int_{0^-}^{\infty} dw_k(x) \exp \left[\left(C + \sum_{i=1}^P z_i D_i \right) x \right] \end{aligned}$$

を得る。

一方、定常状態において、クラス k のシステム内総客数を Y_k としたとき、 j 番目の要素 $y_j(z)$ が

$$y_j(z) = E \left[z_1^{Y_1} \cdots z_P^{Y_P} 1_{\{S=j\}} \right]$$

で与えられるような結合母関数ベクトルを $y(z)$ で表すと、同様の議論により

$$\begin{aligned} y^*(z) & \left(C + \sum_{k=1}^P z_k D_k \right) \\ & = \sum_{k=1}^P \lambda_k (z_k - 1) \int_{0^-}^{\infty} dr_k(x) \exp \left[\left(C + \sum_{i=1}^P z_i D_i \right) x \right] \end{aligned}$$

を得る。これらの式を用いて、各到着流からの待ち客数ならびにシステム内客数の結合定常確率 $\Pr(X_1 = n_1, \dots, X_P = n_P)$, $\Pr(Y_1 = n_1, \dots, Y_P = n_P)$ を計算する数値アルゴリズムを構築することができる。詳細は [14] を参照。

3 おわりに

複数の到着流をもつ単一サーバ待ち行列に関する従来の手法と、サービス時間分布が到着流毎に異なる場合に関する筆者の研究成果を紹介した。待ち行列は、その名が表す通り客数を中心とした議論が多く、通常、この客数分布を元に待ち時間分布を導出するという手順を経る。しかし、サービス時間分布が到着流毎に異なる場合、客数の挙動を直接解析することは極めて困難であり、客数分布を元にしたアプローチでは、何ら解析の結果を得ることはできない。しかし、系内仕事量、あるいは待ち時間分布を直接解析し、それを利用して客数分布の解析へと向かうアプローチを用いれば、従来、解析が困難と見られていたクラスの問題も解決できる可能性があることを示している。

複数の到着流をもつ待ち行列を考えると、しばしば、各到着流の間に優先度を設け、優先度に従ってサービスを行う場合がある。このような場合についても、2.3 章で紹介した到着流と同様の仮定の下で [7, 8.9.12] において研究が行われている。

一方、通信網への応用では、本稿で紹介した連続時間モデル以外にも、時間軸を離散化した離散時間モデル

に対して研究が進められている。しかしながら、一般に、離散時間モデルでは単位時間内での遷移を表す行列 A_k の構造が隠蔽されることが多く、一部の例外を除いて [2]、従来のアルゴリズム的解法の枠を越えた結果というのはほとんど得られていないのが実情である。

参考文献

- [1] Q.-M. He: "Queues with marked customers." *Adv. Appl. Prob.*, 28, 567-587, 1996.
- [2] F. Ishizaki and T. Takine: "Loss probability in a finite discrete-time queue in terms of the steady state distribution of an infinite queue," to appear in *Queueing Systems*, 1999.
- [3] D. M. Lucantoni K. S. Meier-Hellstern and M. F. Neuts: "A single-server queue with server vacations and a class of non-renewal arrival processes," *Adv. Appl. Prob.* 22, 676-705, 1990.
- [4] M. F. Neuts: *Structured Stochastic Matrices of M/G/1 Type and Their Applications*, Marcel Dekker, Wiley, New York, 1989.
- [5] 高橋幸雄: "待ち行列研究の新しい潮流 (1) — 待ち行列研究の変遷," オペレーションズ・リサーチ, 43 (9), 495-499, 1998.
- [6] 高橋幸雄, 牧本直樹: "待ち行列研究の新しい潮流 (3) 相型分布と行列解析法," オペレーションズ・リサーチ, 43 (11), 618-623, 1998.
- [7] T. Takine and T. Hasegawa: "The workload in the MAP/G/1 queue with state-dependent services: Its application to a queue with preemptive resume priority," *Stoch. Mod.*, 10 (1), 183-204, 1994.
- [8] T. Takine et al.: "Mean waiting times in nonpreemptive priority queues with Markovian arrival and i.i.d. service processes," *Perfor. Eval.*, 20 (1-3), 131-149, 1994.
- [9] T. Takine: "A nonpreemptive priority MAP/G/1 queue with two classes of customers," *JORSJ*, 39 (2), 266-290, 1996.
- [10] 滝根哲哉, 村田正幸: "通信網における待ち行列理論の応用と課題," オペレーションズ・リサーチ, 43 (5), 264-271, 1998.
- [11] T. Takine and Y. Takahashi: "On the relationship between queue lengths at a random instant and at a departure in the stationary queue with BMAP arrivals," *Stoch. Mod.*, 14 (3), 601-610, 1998.
- [12] T. Takine: "The nonpreemptive priority MAP/G/1 queue," to appear in *Opns. Res.*, 1999.
- [13] T. Takine: "A new recursion to compute the queue length in the stationary BMAP/GI/1 queue," submitted for publication.
- [14] T. Takine: "Queue length distribution in a FIFO single-server queue with multiple arrival streams having different service time distributions," submitted for publication.
- [15] T. Takine: "Joint queue length distribution in a stationary queue with multiple arrival streams governed by a Markov chain," submitted for publication.
- [16] 滝根哲哉: "構造化されたマルコフ連鎖と待ち行列," システム/制御/情報, 43 (3), 1999.