講座

待ち行列研究の新しい潮流 (5)

待ち時間分布の漸近的特性

牧本 直樹, 小林 和朝

1. はじめに

前回[14]では，通信システムの高度化・複雑化による入力過程モデルの多様化について解説があった．その流れは，ポアソン過程からマルコフ型到着過程に至るマルコフ型の入力過程と，長期依存性を表現するための入力過程の2つに分類することができる．当然のことながら，このような入力過程モデルの多様化は，それを受け入れる待ち行列モデルの解析にも大きな影響を与える．前者については，[15]に解説されている行列解析法の拡張と精密化を促し，後者については大偏差理論のようにこれまで馴染みのなかった解析方法の導入をもたらした．

一方，通信システムの進歩は解析方法だけでなく性能評価尺度にも影響を及ぼしている．その1つが微小セル損失率の評価であろう．短く分割されたデータを高速回線で送信するシステムでは，サービスの種類に応じて10^{-9}以下という微小なセル損失率を品質保証する．このような微小確率の評価には，平均待ち時間などの評価とは異なる方法が必要であり，90年代に入ってからはATM(Asynchronous Transfer Mode)の設計や評価からの要請を受けて幅広い研究がなされている．以下では，徐々に整いつつある微小確率評価のための基本的な枠組みを紹介する．

2. マルコフ型待ち行列の漸近特性

よく知られているように，M/M/1待ち行列の系内数Nと待ち時間Wの定常分布はそれぞれ

$$\begin{aligned} \mathrm{P}(N=n) &= (1-\rho)\rho^n, \quad n=0,1,\cdots, \\ \mathrm{P}(W>x) &= \rho e^{-(\mu-\lambda)x}, \quad x>0 \end{aligned}$$

で与えられる[3]．したがって，系内数と待ち時間の裾分布（あるレベルを超える確率）は幾何的あるいは指数的に減衰することがわかる．このような性質は，到着過程やサービス時間分布がより一般的な場合でも成立するだろうか．本節では，1節で述べた第1の流れに沿ってマルコフ型の待ち行列モデルに対してこの問題を考える．基本的なマルコフ型待ち行列には，モデル化を工夫することによってGI/M/1型かM/G/1型のマルコフ連鎖として表せるものが多いが，実はこれらの構造が漸近特性と密接な関わりを持っている．そのため，以下では構造的マルコフ連鎖の漸近特性に対する一般論を中心に考えていくことにする．なおGI/M/1型やM/G/1型マルコフ連鎖については，この連載中の記事[15]や[10]を参照されたい．

2.1 GI/M/1型マルコフ連鎖

$\boldsymbol{X}_n=(L_n,I_n)$を推移確率行列

$$\boldsymbol{P}=\begin{pmatrix} \boldsymbol{B}_1 & \boldsymbol{B}_0 & \boldsymbol{O} & \boldsymbol{O} & \cdots \\ \boldsymbol{B}_2 & \boldsymbol{A}_1 & \boldsymbol{A}_0 & \boldsymbol{O} & \cdots \\ \boldsymbol{B}_3 & \boldsymbol{A}_2 & \boldsymbol{A}_1 & \boldsymbol{A}_0 & \cdots \\ \boldsymbol{B}_4 & \boldsymbol{A}_3 & \boldsymbol{A}_2 & \boldsymbol{A}_1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{1}$$

に支配される2次元のエルゴード的なGI/M/1型マルコフ連鎖とする．このマルコフ連鎖の定常状態確率を$\pi_{(n,k)}$，レベルnに対する確率ベクトルを$\boldsymbol{\pi}_n=(\pi_{(n,k)})$で表すと，公比行列$\boldsymbol{R}$が存在して

$$\boldsymbol{\pi}_n=\boldsymbol{\pi}_1\boldsymbol{R}^{n-1}, \quad n=2,3,\cdots \tag{2}$$

となる．非負行列$\boldsymbol{R}$の最大固有値をη，ηに対応する左右の固有ベクトルをそれぞれ$\boldsymbol{\ell}$，$\boldsymbol{r}^\top$とすると[4]，Perron–Frobeniusの定理から

$$\boldsymbol{R}^n=\eta^n\boldsymbol{r}^\top\boldsymbol{\ell}+o(\eta^n), \quad n\to\infty \tag{3}$$

まきもと なおき 筑波大学大学院経営政策科学研究科
〒112–0012 文京区大塚 3–29–1
こばやし かずとも NEC C&C メディア研究所
〒216–8555 川崎市宮前区宮崎 4–1–1

[3]到着率λ，サービス率μ，トラヒック密度を$\rho=\lambda/\mu$とする．

[4]$\eta\boldsymbol{\ell}=\boldsymbol{\ell}\boldsymbol{R}$，$\eta\boldsymbol{r}^\top=\boldsymbol{R}\boldsymbol{r}^\top$，$\boldsymbol{\ell}\boldsymbol{r}^\top=1$とする．ただし$\boldsymbol{r}^\top$は$\boldsymbol{r}$の転置を表す．

と展開できて，(2) と (3) より

$$\boldsymbol{\pi}_n = \eta^{n-1}(\boldsymbol{\pi}_1 \boldsymbol{r}^\top)\boldsymbol{\ell} + o(\eta^n), \quad n \to \infty \tag{4}$$

が得られる．

待ち行列モデルとの対応を考えると，レベルnは系内数がn人の状態の集まりであるから

$$\mathrm{P}(N = n) = \boldsymbol{\pi}_n \mathbf{1}^\top = c\,\eta^n + o(\eta^n), \quad n \to \infty \tag{5}$$

$$c = \eta^{-1}(\boldsymbol{\pi}_1 \boldsymbol{r}^\top)(\boldsymbol{\ell}\mathbf{1}^\top) \tag{6}$$

となり，系内数の裾分布は漸近的には幾何的に減衰することが示される．したがって，減衰率ηと定数cが待ち行列モデルのパラメータから同定できれば，(5) は大きなnに対する小さな確率$\mathrm{P}(N = n)$を評価するための有用な情報を与えてくれる．

話をGI/M/1型マルコフ連鎖に戻すと，ηとcのうちηは一般に推移確率行列から明示的に決定されることが知られている．行列$\{\boldsymbol{A}_i\}$の母関数行列

$$\boldsymbol{A}(z) = \sum_{i=0}^{\infty} z^i \boldsymbol{A}_i, \quad 0 \le z \le 1$$

が存在するとして，$\boldsymbol{A}(z)$の最大 (Perron–Frobenius) 固有値をzの関数と考えて$PF(z)$で表す．このとき，ηは方程式

$$\eta = PF(\eta), \quad 0 < \eta < 1 \tag{7}$$

を満たす唯一の解として与えられる．これに対して，(6) からcを計算するには$\boldsymbol{\pi}_1$が必要となるが，$\boldsymbol{\pi}_1$は$\boldsymbol{A}_i$だけでなく$\boldsymbol{B}_i$にも依存しているため，特殊なモデルを除いてはcを明示的に定めることは難しい．

例 2.1 GI/PH/c 待ち行列に対して到着直前の状態に着目して GI/M/1 型マルコフ連鎖を構成し，上の結果を適用する．到着間隔とサービス時間分布の Laplace–Stieltjes 変換 (LST) をそれぞれ$\alpha(s)$，$\beta(s)$で表し，$\kappa > 0$を

$$\log \alpha(\kappa) + \log \beta(-c\kappa) = 0, \quad \kappa > 0$$

を満たす解[5]とすると，$\eta = \alpha(\kappa)$として

$$\mathrm{P}(N = n) \;=\; c_1 \eta^n + o(\eta^n), \quad n \to \infty, \tag{8}$$

$$\mathrm{P}(W > x) \;=\; c_2 e^{-\kappa x}, \quad x \to \infty \tag{9}$$

が得られる．ここで重要なのは，一般の GI/M/1 型マルコフ連鎖では$PF(z)$の方程式 (7) でηを定めたが，GI/PH/c というモデルの構造を利用することで，$\alpha(s)$や$\beta(s)$など待ち行列モデルのパラメータを使ってκやηがより直接的に表されるという点である．なお，この結果はサービス時間分布がサーバによって異なる場合へも拡張されている [11]. □

[5]到着間隔の裾分布が指数的に減衰すれば一般に存在して一意である．

2.2 M/G/1 型マルコフ連鎖

次に，推移確率行列

$$\boldsymbol{P} = \begin{pmatrix} \boldsymbol{B}_0 & \boldsymbol{B}_1 & \boldsymbol{B}_2 & \boldsymbol{B}_3 & \cdots \\ \boldsymbol{A}_0 & \boldsymbol{A}_1 & \boldsymbol{A}_2 & \boldsymbol{A}_3 & \cdots \\ \boldsymbol{O} & \boldsymbol{A}_0 & \boldsymbol{A}_1 & \boldsymbol{A}_2 & \cdots \\ \boldsymbol{O} & \boldsymbol{O} & \boldsymbol{A}_0 & \boldsymbol{A}_1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{10}$$

に支配されるエルゴード的なM/G/1型マルコフ連鎖を考える．2.1 節と同様に，定常状態確率ベクトルを$\boldsymbol{\pi}_n = (\pi_{(n,k)})$で表す．GI/M/1 型では特別な条件が無くても (4) のように$\boldsymbol{\pi}_n$は幾何的に減衰するが，(10) の構造からM/G/1型では少なくとも$\boldsymbol{A}_i$や$\boldsymbol{B}_i$が幾何的に減衰する必要がある．また，[15] に述べられているように，GI/M/1型と異なりM/G/1型では (2) の行列幾何形式解のような陽表現を持たないため，$\boldsymbol{\pi}_n$の漸近特性を調べるには別の方法が必要となる．

$\boldsymbol{P}$の構造から，$\boldsymbol{\pi}_n$が満たす平衡方程式は

$$\boldsymbol{\pi}_n = \boldsymbol{\pi}_0 \boldsymbol{B}_n + \sum_{i=1}^{n+1} \boldsymbol{\pi}_i \boldsymbol{A}_{n+1-i}, \quad n = 0, 1, \cdots$$

で与えられる．両辺の母関数ベクトルをとって形式的に整理すると

$$\boldsymbol{\pi}(z) = \boldsymbol{\pi}_0[z\boldsymbol{B}(z) - \boldsymbol{A}(z)][z\boldsymbol{I} - \boldsymbol{A}(z)]^{-1} \tag{11}$$

となる．ただし，

$$\boldsymbol{\pi}(z) = \sum_{i=0}^{\infty} z^i \boldsymbol{\pi}_i,$$

$$\boldsymbol{A}(z) = \sum_{i=0}^{\infty} z^i \boldsymbol{A}_i, \quad \boldsymbol{B}(z) = \sum_{i=0}^{\infty} z^i \boldsymbol{B}_i$$

とし，$\boldsymbol{I}$は単位行列を表す．2.1 節と同様に，$PF(z)$を$\boldsymbol{A}(z)$の最大固有値とし，

$$\eta^{-1} = PF(\eta^{-1}), \quad 0 < \eta < 1 \tag{12}$$

を満たすηが存在するとしよう．(11), (12) より，$z \uparrow \eta^{-1}$のとき$\boldsymbol{\pi}(z) \uparrow \infty$となるが，$\boldsymbol{\pi}(z)$の定義からこれは適当なベクトル$\boldsymbol{u}$に対して

$$\boldsymbol{\pi}_n = c\,\eta^n\,\boldsymbol{u} + o(\eta^n), \quad n \to \infty \tag{13}$$

となることを意味する．以上はやや直観的な説明であるが，基本的には(12)の解ηが求まり，さらに$\boldsymbol{B}(\eta)$が存在すれば(13)の幾何的な減衰が示される．GI/M/1型と比較すると，(1)と(10)の間の双対的な関係がηを決める方程式(7)と(12)に反映されていることがわかる．なお，この議論の詳細はFalkenberg[5]を参照されたい．

例 2.2 MAP/GI/1待ち行列の退去直後の時点に着目して構成される隠れマルコフ連鎖はM/G/1型となる．MAPに関して到着を伴う（伴わない）推移速度行列を$\boldsymbol{D}$ ($\boldsymbol{C}$)で表し[6]，$(s\boldsymbol{I}-\boldsymbol{C})^{-1}\boldsymbol{D}$の最大固有値を$\alpha(s)$，サービス時間分布のLSTを$\beta(s)$とする．(12)の解$\eta$は，方程式

$$\log\alpha(\kappa)+\log\beta(-\kappa)=0, \quad \kappa>0$$

の解κを用いて$\eta=\alpha(\kappa)$で与えられる．このようにして定まるηとκに対して，系内数と待ち時間の裾分布は(8), (9)の漸近的な性質を満たす．なおMAP/GI/1においては，サービス時間の裾分布が指数的に減衰して$\beta(s)$が適当な領域で定義されていれば，必要な条件はほぼ満たされていると考えてよい． □

これまでの議論から推察されるように，M/M/1のような単純なモデルでなくとも，(i)到着やサービス過程の一方あるいは両方が有限次元のマルコフ過程によってドライブされていて，(ii)モデルに含まれる一般分布が指数的な減衰を示す，ような待ち行列モデルにおいては待ち時間や系内数の裾分布は指数的な減衰を示す．これは，有限次元マルコフ過程によって持ち込まれるモデルの複雑さ（相関構造）が本質的に短期依存性を持つためである．次節では，モデルの相関構造と裾分布の関係をもう少し詳しく見てみよう．

3. 入力過程の相関構造と漸近特性

1節で触れたように，入力過程モデルの多様化のもう1つの流れは長期依存性を持つ入力過程の導入である．以下では離散時間で話を進めるが，連続時間モデルに対しても同様な議論が成立する．

時点nでの待ち時間をW_n，時点nと$n+1$の間に持ち込まれるサービス要求量からその間に処理可能なサービス量を引いたものをX_nとすると，

$$W_{n+1}=\max(W_n+X_n,\, 0) \tag{14}$$

という関係が成り立つ．以下では，$\{X_n\}$は定常過程であるとし，安定条件として$\mu=\mathrm{E}(X_n)<0$を仮定する．$n\to\infty$としたときの定常状態における待ち時間Wの裾分布は

$$\mathrm{P}(W>x)=\mathrm{P}\left(\sup_{n\geq 0}\{A_n\}>x\right)$$

で与えられる．ここで

$$A_0=0, \quad A_n=\sum_{i=-n}^{1}X_i$$

とする．

$x\to\infty$のときの$\mathrm{P}(W>x)$の挙動は，$\{X_n\}$の相関構造に強く依存するため，以下では2つのケースに分けて考える．$\{X_n\}$の自己相関関数

$$r(k)=\frac{\mathrm{E}((X_0-\mu)(X_k-\mu))}{\mathrm{V}(X_0)}, \quad k=0,1,\cdots$$

が$k\to\infty$のとき指数的に減衰する，すなわち適当な$\gamma\in(0,1)$に対して$r(k)=o(\gamma^k)$を満たすとき$\{X_n\}$は短期依存性を持つといい，$r(k)=O(k^{-\delta})$ $(0<\delta<1)$のようにべき関数的に減衰するならば長期依存性を持つという[7]．一般に，ON-OFFモデルやMAPなどのマルコフ型入力過程は短期依存性を持つのに対し，自己相似過程は長期依存性を持つ[14].

3.1 短期依存性と指数的減衰

まず$\{X_n\}$が短期依存性を持つ場合を考える．

$$\psi(s)=\lim_{n\to\infty}\frac{1}{n}\log\mathrm{E}\left(e^{sA_n}\right) \tag{15}$$

がsのある領域$\mathcal{S}$で定義されており，さらに

$$\psi(\kappa)=0, \quad \kappa>0$$

なるκが$\mathcal{S}$の内点として存在することを仮定する．Glynn and Whitt [7]は，これに加えてある緩やかな条件が成り立てば

$$\lim_{x\to\infty}\frac{1}{x}\log\mathrm{P}(W>x)=-\kappa \tag{16}$$

となることを示した．(16)は，$x\to\infty$のとき

$$\mathrm{P}(W>x)\approx c(x)e^{-\kappa x}+o(e^{-\kappa x}) \tag{17}$$

と近似できることを意味する．$c(x)$は$e^{-\kappa x}$に比べて緩やかに変動する関数で，$\lim_{x\to\infty}x^{-1}\log c(x)=0$を満たす．この結果を例えば(9)と比較すると，$c(x)$の挙動は明らかでないが，本質的に$\mathrm{P}(W>x)$は指数的

[6] 小沢[14]の3.2節を参照．

[7] $\sum_{k=0}^{\infty}r(k)$の収束/発散によって短期/長期依存性を定義する場合もある[1].

に減衰するという点では一致している．このように，入力過程（あるいはサービス過程）がマルコフ的でなくても，その相関構造が短期依存性を持つ場合には，待ち時間の裾分布は指数的な減衰を示す．

3.2 長期依存性を持つモデルの漸近特性

小沢 [14] の4節で述べられている自己相似的な入力過程も含めて，$\{X_n\}$ が長期依存性を持つ場合はどうだろうか．この場合 (15) の右辺は発散してしまうため，3.1節の結果をそのまま適用することはできない．そこで Duffield and O'Connell [4] は大偏差理論を援用することで，長時間依存性を持つ場合の分析を行った．大偏差理論とその適用方法については，次節でより一般的な枠組みの中で説明することにして，ここでは結果を中心に見てみよう．

(15) の代わりに，$n \to \infty$ のとき発散する増加数列 $\{a(n)\}$，$\{v(n)\}$ に対して

$$\overline{\psi}(s) = \lim_{n\to\infty}(v(n))^{-1}\log \mathrm{E}\left(e^{sv(n)A_n/a(n)}\right)$$

がある領域で存在すると仮定し，

$$\overline{\psi}^*(y) = \sup_s\{sy - \overline{\psi}(s)\}$$

と定める．また，適当な増加数列 $h(n)$ に対して

$$g(y) = \lim_{n\to\infty}(h(n))^{-1}v(a^{-1}(n/y))$$

が定義できるものとする ($a^{-1}(z) = \sup\{n\ :\ a(n) \le z\}$)．このとき，大偏差理論における Gärtner–Ellis の定理 [13] から，緩やかな条件の下で

$$\lim_{x\to\infty}\frac{1}{h(x)}\log \mathrm{P}(W > x) = -\inf_{y>0} g(y)\overline{\psi}^*(y) \quad (18)$$

が成立する．(18) から，十分大きな x に対しては

$$\mathrm{P}(W > x) \approx e^{-\kappa h(x)} \quad (19)$$

という漸近特性が得られる ($\kappa = \inf_{y>0} g(y)\overline{\psi}^*(y)$)．

(19) では，$h(x)$ や $g(y)$，$\overline{\psi}^*(y)$ などの関数とモデルとの関連が明示的でないため，以下では具体例を通してモデルパラメータと漸近特性の関係を調べてみよう．なお，$\{X_n\}$ が短時間依存性を持つ場合は，$a(n) = v(n) = n$ と定めることにより $h(n) = n$ となり，(18) から指数的な減衰 (16) が導かれる．

例 3.1 長期依存性を持つ代表的な例として $\{A_n\}$ が（離散時間）fractional Brownian motion (FBM) に従う場合を考えよう．$\{Z_H(t)\}$ を [14] の4.2節で定義される FBM とし，$A_n = Z_H(n) + n\mu$ ($\mu = \mathrm{E}(X_n) < 0$) とする．この場合は $a(n) = n$，$v(n) = n^{2-2H}$ とおいて上の結果を適用すると，

$$\lim_{x\to\infty}\frac{1}{x^{2H-2}}\log \mathrm{P}(W > x) = -\inf_{y>0}\frac{(y-\mu)^2}{2y^{2-2H}}$$

が得られる．したがって，$\{X_n\}$ が長期依存性を持つ場合 ($0.5 < H < 1$) は待ち時間の裾分布がワイブル型となり，短期依存性を持つ場合に比べて長い裾を持つことがわかる． □

4. 多重化における漸近特性

3節では x を増加させたときの待ち時間分布の漸近特性について考えたが，本節では，入力の多重数 L を増加させたときの待ち時間分布の漸近特性を議論する．これは，多数の情報源から送られるデータを高速回線で送信する統計多重されたシステムを想定したものである．情報源から送られるデータ量は時間的に変動するため，多重化によってそれらの増減が打ち消しあうことによる効率化が期待できるが，以下の議論はその効果を定量的に把握するための1つの方向性を示している．

時点 n と $n+1$ との間に L 本の入力によって持ち込まれるサービス要求量からその間に処理可能なサービス量を引いたものを X_n^L とすると，そのときの待ち時間 $\{W_n^L\}$ の変化は

$$W_{n+1}^L = \max(W_n^L + X_n^L, 0)$$

で表される．これは (14) の X_n を X_n^L で置き換えたものである．したがって，$\{X_n^L\}$ を定常過程とし $A_n^L = \sum_{i=-n}^{1} X_i^L$ とすると，定常状態における待ち時間 W^L は

$$\mathrm{P}(W^L > x) = \mathrm{P}\left(\sup_{n\ge 0}\{A_n^L\} > x\right)$$

で与えられる．

ここで，次の仮定をおく．

(A1) $E(X_n^L) < 0$ かつ $\lim_{L\to\infty} E(X_n^L)/L < 0$ とする．

(A2) n に関する増加関数 $a(n)$, $v(n)$ が存在し，

$$\hat{\psi}_n^L(s) = \quad (Lv(n))^{-1}\log \mathrm{E}\left(e^{sA_n^L v(n)/a(n)}\right)$$

としたとき，$\hat{\psi}_n(s) = \lim_{L\to\infty}\hat{\psi}_n^L(s)$，$\hat{\psi}(s) = \lim_{n\to\infty}\hat{\psi}_n(s)$ が存在する．ただし $\pm\infty$ も許す．

(A3) $\hat{\psi}_n(s)$ および $\hat{\psi}(s)$ が有限の値をとる s の領域がそれぞれ存在し，領域の内部では微分可能で，かつ点 s が領域の境界に近づくとき $\hat{\psi}_n^L(s)$, $\hat{\psi}(s)$ の微分は発散する．

以上の仮定の下で，Duffield [3] は，$y>0$ に対して

$$\lim_{L\to\infty} L^{-1}\log \mathrm{P}(W>Ly) = -\inf_{n>0}\psi_n^*(y), \qquad (20)$$

$$\begin{aligned}\psi_n^*(y) &= \sup_s\{sy-\psi_n(s)\},\\ \psi_n(s) &= \lim_{L\to\infty} L^{-1}\log \mathrm{E}\left(e^{sA_n^L}\right)\end{aligned}$$

であることを示した[8]．ここで重要なのは，仮定の段階では長期依存過程を含めるために時間スケールを変換する増加関数 $a(n)$, $v(n)$ を必要としたのであるが，これらの関数が待ち時間分布の漸近的な振舞には関係していないことである．

この結果がどのようにして導かれるかを簡単に説明する．(A1) ～ (A3) の仮定の下では，大偏差原理 (Gärtner-Ellis の定理) から A_n^L に関して

$$\lim_{L\to\infty} L^{-1}\log \mathrm{P}\left(A_n^L>Ly\right) = -v(n)\hat{\psi}_n^*(y/a(n)) \qquad (21)$$

$$\hat{\psi}_n^*(y) = \sup_s\{sy-\hat{\psi}_n(s)\}$$

が成り立つ．(21) の右辺は，$v(n)s/a(n)$ を $\tilde{s}$ とおくならば，

$$\begin{aligned}v(n)\hat{\psi}_n^*(y/a(t)) &= v(n)\sup_s\{sy/a(n)-\hat{\psi}_n(s)\}\\ &= \sup_{\tilde{s}}\{\tilde{s}y-\psi_n(\tilde{s})\} = \psi_n^*(y)\end{aligned}$$

となり，$a(n)$, $v(n)$ には依存しない関数 $\psi_n^*(y)$ となる．(21) は，y を固定して考えると $\mathrm{P}\left(A_n^L>Ly\right) = c_n(L)e^{-L\psi_n^*(y)}$, $c_n(L)=e^{o(L)}$ と書くことができる．一方，$\mathrm{P}\left(\sup_{n\ge0}\{A_n^L\}>Ly\right)$ は

$$\begin{aligned}\mathrm{P}\left(\sup_{n\ge0}\{A_n^L\}>Ly\right) &\le \sum_{n=1}^{\infty}\mathrm{P}\left(A_n^L>Ly\right) \qquad (22)\\ \mathrm{P}\left(\sup_{n\ge0}\{A_n^L\}>Ly\right) &\ge \sup_{n>0}\mathrm{P}\left(A_n^L>Ly\right) \qquad (23)\end{aligned}$$

の不等式を満足する．いま，話を簡単にするために，固定した y に対して $\psi_n^*(y)$ が最小値をとる n が一個とし，それを $\hat{n}$ とする．そのとき (22) の左辺は，

$$\begin{aligned}&\sum_{n=1}^{\infty}\mathrm{P}\left(A_n^L>Ly\right)\\ &= e^{-L\psi_{\hat{n}}^*(y)}\left\{c_{\hat{n}}(L)+\sum_{n\ne\hat{n}}c_n(L)e^{-L(\psi_n^*(y)-\psi_{\hat{n}}^*(y))}\right\}\end{aligned}$$

となる．ここで L を大きくすると，$\sum_{n\ne\hat{n}}$ の項は $c_{\hat{n}}(L)$ より先に 0 に近付くため $c_{\hat{n}}(L)e^{-L\psi_{\hat{n}}^*(y)}$ が残る．また，(23) の右辺は L が大きいところでは明らかに $c_{\hat{n}}(L)e^{-L\psi_{\hat{n}}^*(y)}$ となる．以上により，(20) が得られる．$\psi_n^*(y)$ が最小値をとる n が複数ある場合も上限，下限の指数部分は変わらず，(20) が成り立つ．

説明からわかるように，L が大きいところでは，

$$\mathrm{P}\left(\sup_{n\ge0}\{A_n^L\}>Ly\right) \approx \sup_{n>0}\mathrm{P}\left(A_n^L>Ly\right)$$

となり，この性質が (20) を導くときの重要なポイントとなっている．これは大偏差理論における，稀な事象は最も可能性の高い起こり方でしか起きない，という性質に対応するものである．

ここで，$I(y)=\inf_{n>0}\psi_n^*(y)$ について議論しておこう．簡単のために $\{X_n^L\}$ が短期依存性を持つ場合を考える．このとき $\kappa=\lim_{y\to\infty}I(y)/y$ が存在し，$\delta=\lim_{n\to\infty}\psi_n(\kappa)$ とおくと $\lim_{y\to\infty}(I(y)-\kappa y)=\delta$ が成り立つ．したがって，十分大きな y に対して $I(y)\approx\kappa y+\delta$ となる．(17) と比較するために，$x=Ly$ とおくと，待ち時間分布 $\mathrm{P}\left(W^L>x\right)$ の指数関数の部分は $e^{-L\delta}e^{-\kappa x}$ となり，$L\delta$ の分だけ更に値が抑えられる．このことは，$\{X_n^L\}$ が長期依存性を持つ場合についても同様に成り立つ．したがって，例えば同じトラヒック密度（負荷）であっても，統計多重されたシステムでは（微小）呼損率がより小さくなることが期待される．ただし，$\{X_n^L\}$ がガウス過程の場合は，多重数 L に依らず $\delta=0$ である．

一方，$y=0$ のときの $I(y)$ は，$I(0)=\psi_1^*(0)$ となり，$I(0)=\lim_{L\to\infty}L^{-1}\log\mathrm{P}\left(X_n^L>0\right)$ となる．これは，待ちが発生する確率 $\mathrm{P}\left(X_n^L>0\right)$ の漸近特性であり，$y=0$ の近傍でも $I(y)$ が有用な情報を持っていることを示している．

例 4.1 $I(y)$ の具体例として，$\{X_n^L\}$ を平均 $L\mu$ の定常ガウス過程の場合を考える．このとき $\{A_n^L\}$ は平均値関数が $Ln\mu$ のガウス過程となる．ここで，A_n^L の分散を $L\sigma_n^2$ とすると $\psi_n(s)=n\mu s+\sigma_n^2s^2/2$, $\psi_n^*(y)=(y-n\mu)^2/(2\sigma_n^2)$ が得られる．したがって $I(y)$ は

$$I(y)=\inf_{n>0}\left\{\frac{(y-n\mu)^2}{2\sigma_n^2}\right\} \qquad (24)$$

で与えられる．

$\{X_n^L\}$ が短期依存性を持つとき，$\lim_{n\to\infty}\sigma_n^2/n=\sigma^2$ であり，このとき，$\lim_{y\to\infty}I(y)/y=\kappa$ となる．一方，$\{X_n^L\}$ が長期依存性を持ち，$\lim_{n\to\infty}\sigma_n^2/n^{2H}=\sigma^2$,

[8] $\inf_{n>0}\psi_n^*(y)$ が不連続であるような y に対しては (20) の等式は成り立たないため，厳密には以下の不等式 $-\inf_{n>0}\psi_n^*(y+0)\le\liminf_{L\to\infty}L^{-1}\log\mathrm{P}(W>Ly)\le\limsup_{L\to\infty}L^{-1}\log\mathrm{P}(W>Ly)\le-\inf_{n>0}\psi_n^*(y)$ で表現される．

$H \in (0.5, 1)$ であるとき，$\lim_{y\to\infty} I(y)/y^{2-2H} = \kappa$ となる． □

例 4.2 例 4.1 において，$\Delta\sigma_n = \sigma_n - \sigma_{n-1}$ として，すべての n に対して $\Delta\sigma_n > \Delta\sigma_{n+1}$ が成り立つと仮定する．ここで，$\phi(n) = -\mu(\sigma_n/\Delta\sigma_n - n)$ とおくと，(24) は $\phi(m) < y \le \phi(m+1)$ となる y に対して

$$I(y) = \frac{(y - m\mu)^2}{2\sigma_m^2} \tag{25}$$

となる．これは，逆に (25) を成り立たせる y の範囲を

$$\frac{y-(m-1)\mu}{\sigma_{m-1}} < \frac{y - m\mu}{\sigma_m} \le \frac{y-(m+1)\mu}{\sigma_{m+1}}$$

から求めることにより導かれる．(25) が意味することは，$0 < y \le \phi(m+1)$ の範囲の y に対して $I(y)$ は μ と $\sigma_1^2, \sigma_2^2, \cdots, \sigma_m^2$ の情報から決定できるということである．ここで重要なのは，$\sigma_1^2, \sigma_2^2, \cdots, \sigma_m^2$ の情報からは，$\sigma_n^2/n \to \sigma^2$ となり短期依存性を持つのか，$\sigma_n^2/n^{2H} \to \sigma^2$, $H \in (0.5, 1)$ となり長期依存性を持つのかが分からないことである．このことは，逆に短期依存性か長期依存性かがわからなくても有限の y に対して $I(y)$ を求めることができることを示している．この点に関しては，文献 [8] に詳しく述べられている． □

以上，$\{X_n^L\}$ が定常過程の場合の多重化における待ち時間分布の漸近特性を議論した．ここでは紙面の関係で省略したが，文献 [9] には非定常な場合でも同様な手法で議論できることが示されている．

5. おわりに

本稿では，情報通信システムにおける微小呼損率評価への応用を念頭において研究が進められている待ち時間分布の漸近特性に関して，基本的な結果を中心に説明を行ってきた．紙面の関係で省略したが，直列型待ち行列を中心とするネットワークモデルの漸近特性についてもこの数年の間にいろいろな問題が解明されつつある．例えば Chang [2], O'Connell [12], Fujimoto et. al [6] などを参照されたい．漸近解析によって得られる結果は微小確率評価のための有用な情報を提供してくれるが，それ自身が呼損率となるわけではないのでそのままで性能評価尺度として使うことはできない．理論的な枠組みの拡充と併せて，実システムへの適用に関する研究が今後重要になると思われる．

参考文献

[1] Beran, J. (1994) *Statistics for Long–Memory Processes*, Chapman & Hall.

[2] Chang, C.–S. (1995) "Sample path large deviations and intree networks," *Queueing Systems*, **20**, 7–36.

[3] Duffield, N. G. (1997) "Economies of scale for long-range dependent traffic in short buffers," *Telecommunication Systems*, **7**, 267-280.

[4] Duffield, N. G. and O'Connell, N. (1995) "Large deviations and overflow probabilities for the general single–server queue, with applications," *Math. Proc. Camb. Phil. Soc.*, **118**, 363–374.

[5] Falkenberg, E. (1994) "On the asymptotic behavior of the stationary distribution of Markov chains of M/G/1 type," *Stoch. Models*, **10**, 75–98.

[6] Fujimoto, K., Takahashi, Y. and Makimoto, N. (1998) "Asymptotic properties of stationary distributions in two–stage tandem queueing systems," *J. Oper. Res. Soc. Japan*, **41**, 118–141.

[7] Glynn, P. W. and Whitt, W. (1994) "Logarithmic asymptotics for steady–state tail probabilities in a single–server queue" *J. Appl. Prob.*, **31A**, 131–156.

[8] Kobayashi, K. and Takahashi, Y. (1998) "Tail probability of a Gaussian fluid queue under finite measurement of input processes," *Performance and Management of Complex Communication Networks*, Hasegawa, Takagi and Takahashi (eds.), Chapman & Hall, 43-58.

[9] Kobayashi, K. and Takahashi, Y. (1998) "Overflow probability for a single-server queue with nonstationary multiplexed input," *Proc. 11th ITC Specialist seminar*, 215-220.

[10] Neuts, M.F. (1995) "Matrix–analytic methods in the theory of queues," *Advances in Queueing*, J. H. Dshalalow, Ed., CRC Press, 265–292.

[11] Neuts, M. F. and Takahashi, Y. (1981) "Asymptotic behavior of the stationary distributions in the GI/PH/c queue with heterogeneous servers," *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, **57**, 441–452.

[12] O'Connell, N. (1997) "Large deviations for departures from a shared buffer," *J. Appl. Prob.*, **34**, 753–766.

[13] Scwartz, A. and Weiss, A. (1995) *Large Deviations for Performance Analysis*, Chapman & Hall.

[14] 小沢 (1998) "いろいろな入力過程モデル," オペレーションズ・リサーチ, **43** (12).

[15] 高橋，牧本 (1998) "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.

[16] 牧本 (1998) "相関をもつランダムウォークの極限分布の漸近的性質と待ち行列モデルへの応用," 応用数理, **8**, 2–13.