拡散近似：離散と連続のはざまで

木村 俊一

1. 拡散近似との出会い

もう20年以上も前のことになるが，大学院修士課程の入学試験に合格してやっと一息つき，本格的に卒論の準備を始めなければ，と考えていた頃，私はその論文 [13] に出会った．論文の著者は，当時IBMワトソン研究所に所属し，現在プリンストン大学教授の小林久志氏であった．1975年9月のことだった．当時，私は京都大学工学部数理工学科計画工学講座（三根研究室）の4年生．助教授になられて間もない大野勝久先生（現在，名古屋工業大学教授）の下で，ネットワーク構造をもつ待ち行列システムの性能評価で「何かできないか」と漠然と考えていたところだった．まだ待ち行列ネットワークに関するテキスト等も出版されておらず，「お題目」は絞られていたのだが，問題そのものの理解や問題解決のための方法論に対する準備は，極めて不十分な状態だった．そんな折，論文作成のノウハウに関して右も左も分からなかった私にとって，小林氏の論文はとても魅力的なものに感じられた．

魅力的な点は2つあった．第1の点は，論文内容の斬新さである．小林氏の論文は，情報通信や生産システムの標準的な性能評価モデルとして知られている，複数のサービス施設がネットワークの形に結合された一般的な待ち行列ネットワークに関するもので，この解析に**拡散近似** (diffusion approximation) の手法を最初に応用したものであった．各サービス施設における客数は，ネットワーク内の他の施設や外部からの客の到着，サービスの終了によって不規則に変化するが，客数である以上，$0, 1, 2, \ldots$ といった非負の整数値しか取り得ない．拡散近似は，複雑でそのままでは解析的に取り扱いが困難な離散値確率過程を，連続値を取る拡散過程（あるいはブラウン運動）で近似する手法で，**流体近似** (fluid approximation) を精緻化した連続近似として，当時すでにNewell [15] らによって，単一の待ち行列システムに対しては精力的な研究が行われていた．待ち行列ネットワークというと，Jacksonネットワークとよばれる客のサービス時間に指数分布を仮定するモデルしか知らなかった私には，一般のサービス時間分布を許す拡散近似は応用上有用なORツールと映ったのである．

第2の魅力的な点は，卒論を期限内にまとめなければならない私にとってはこちらの方がより魅力的に思えたのだが，小林氏の論文で使われている数学に関してかなりの予備知識がすでに私にあったことだった．大学院入試科目である応用物理学の受験勉強が意外な所で役立つものだと感心したのを覚えている．

しかし，私のこういった思い入れは，結果的に卒論で結実することはなかった．大野先生の「拡散近似は卒論では難しすぎるから，修士課程に入ってから勉強しなさい」の一言で，今から考えると別の意味でより難しい，ある医療システムのモデルである階層的な待ち行列ネットワークの解析をする羽目になってしまった．拡散近似の一体何が難しいのかも分からないまま，こうして小林氏の論文は半年間お蔵入りになった．

私が拡散近似と再会したのは，卒論の後始末が済んだ1976年の春のことだった．小林氏の論文の参考文献を手掛かりに過去の研究を遡る過程の中で，大野先生の「拡散近似は卒論では難しすぎる」という言葉の意味が次第に見えてきた．拡散近似には，Newellの成書や小林氏の論文が扱っているOR的あるいは工学的な連続モデルという位置付けとは全く別の，極限定理の証明という純粋に数学的な世界が存在していたのである [4, 20]．これら2つの世界を区別する意味で，本稿では前者を**拡散モデル** (diffusion models)，後者を**拡散極限** (diffusion limits) とよぶことにする．後述するように，拡散極限は重負荷時に拡散モデルの理論的根拠を与えてくれるものの，「近似」という本来の機能に対する関心は非常に希薄で，この点にどうしても馴染めなかった私は，次第に拡散モデルに傾倒してゆ

きむら としかず 北海道大学経済学部
〒060 札幌市北区北9条西7丁目

くことになる. 修士論文, 博士論文, そして現在へと続く私の拡散近似の世界への道はこうして始まったのである.

拡散近似は, 待ち行列に限らず在庫やファイナンス等のORの様々な分野にも幅広く応用されている. しかし, 上に述べた2つの世界が, ある時には相補的に, またある時には対立的に語られるのは, 待ち行列の分野が他を圧倒している. 数学的な厳密さとORへの応用の両面に配慮しながら, 待ち行列における拡散近似の2つの世界を紹介しよう.

2. 離散から連続へ

2.1 確率過程の弱収束

待ち行列において現われる確率過程は, 何らかの形で不連続性を含んでいることが多い. サービス中や行列に並んでいる客の数を表わす過程 (系内客数過程) は, 客の到着や退去に伴い増減する階段関数型のサンプルパスをもっているし, サーバーがこれから処理しなければならない客のサービス時間の総和 (未処理仕事量) も, 鋸刃型のサンプルパスをもっている. 応用上は, 対象とする確率過程のサンプルパスに対しては, 関数空間 $D_E[0,\infty) = \{\omega : [0,\infty) \to E \mid \omega$ はすべての $t \geq 0$ に対して右連続かつすべての $t > 0$ に対して左極限をもつ$\}$ を考えれば十分である. ただし, E は値域集合を表わす. 単一の待ち行列の通常の系内客数過程を扱う場合には, E として実数全体 $\mathbb{R}$ またはその部分集合を, 待ち行列ネットワークや多種類の客を含む場合は, E として多次元の集合を用いる. これに対して, 拡散過程のサンプルパスは連続である. この空間を $C_E[0,\infty)$ ($\subset D_E[0,\infty)$) で表すことにする.

待ち行列システムにおいて現われるある確率過程が, 複雑で解析的に扱うことが難しいものであると仮定しよう. 拡散近似の基本的な考え方は, この確率過程の分布を近似できる拡散過程を見つけ出すことにある. 2項分布を正規分布で近似する古典的結果を思い出して頂ければ容易に想像がつくように, この近似をより正確に表現するためには, 何らかの極限定理を必要とする. この極限定理における基本的な収束概念が**弱収束** (weak convergence) である. 確率過程ではなく確率変数の列に対しては, 弱収束は通常**法則収束** (convergence in distribution) とよばれている. そこでまず準備として, 実数直線上の確率変数の弱収束の定義を復習しておこう.

定義 1 $\mathbb{R}$-値確率変数族 $\{X, X_n; n \geq 0\}$ が与えられ, 分布関数 $P\{X \leq \cdot\}$ のすべての連続点 x において

$$P\{X_n \leq x\} \to P\{X \leq x\} \quad (n \to \infty)$$

が成り立つとき, X_n は X に弱収束するといい, $X_n \xrightarrow{w} X$ と書く.

確率過程列の弱収束を定義するには, $D_E[0,\infty)$ に適当な距離を導入する必要があるが, ここでは本来の定義と等価で定義1と直観的に結び付けることが容易な次の定義を示す.

定義 2 $D_E[0,\infty)$-値確率過程族 $\{X, X_n; n \geq 0\}$ が与えられ, すべての有界連続関数 $h : D_E[0,\infty) \to \mathbb{R}$ に対して

$$E[h(X_n)] \to E[h(X)] \quad (n \to \infty)$$

が成り立つとき, X_n は X に弱収束するといい, $X_n \xrightarrow{w} X$ と書く.

2.2 関数中心極限定理

拡散過程への弱収束を示す極限定理の確率変数版が, よく知られた中心極限定理であることは想像に難くない. したがって, まず最初に初等的な2つの極限定理を思い出すことにしよう.

$\{Z_n; n \geq 0\}$ を独立で同じ分布をもつ (i.i.d.) $\mathbb{R}^d$-値確率変数列とし, $S_n = Z_1 + \cdots + Z_n$, $S_0 = 0$ とおく. **大数の強法則** (strong law of large numbers) は, $E[||Z_n||] < \infty$ ならば, $n \to \infty$ のとき

$$n^{-1} S_n \to \mu, \quad \text{a.s.} \tag{1}$$

が成り立つことを示している. ここで, $||\cdot||$ は $\mathbb{R}^d$ 上のユークリッドノルム, $\mu = E[Z_n]$ を表わす. 他方, **中心極限定理** (central limit theorem) は大数の強法則をさらに精密化したもので, ランダムウォーク $\{S_n; n \geq 0\}$ に対し, $E[||Z_n||^2] < \infty$ ならば, $n \to \infty$ のとき

$$n^{1/2}(n^{-1} S_n - \mu) \xrightarrow{w} \Sigma^{1/2} N(0, I) \tag{2}$$

が成り立つことを示している. ただし, $N(0,I)$ は平均ベクトル0と単位行列の共分散行列をもつ $\mathbb{R}^d$-値多変量正規確率変数を表わし, $\Sigma^{1/2}$ は共分散行列 $\Sigma = E[Z_n^T Z_n] - E[Z_n^T] E[Z_n]$ に対し, $\Sigma = \Sigma^{1/2} (\Sigma^{1/2})^T$ を満たす正則行列を表すものとする. 特に $d = 1$ のとき, 定義1より, (2) はおなじみの形

$$\lim_{n\to\infty} P\left\{\frac{S_n - n\mu}{\sqrt{n}\sigma} \leq x\right\} = \Phi(x) \tag{3}$$

に表せる．ただし，$\sigma^2 = \mathrm{Var}[Z_n]$, $\Phi(\cdot)$ は標準正規分布の分布関数

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-u^2/2} du, \quad x \in \mathbb{R}$$

である．

(1) と (2) の確率過程版を見つけ出すために，$\overline{X}_n(t) \equiv n^{-1} S_{[nt]}$を考えることにする．ここで，$[x]$ は x 以下の最大の整数を表すものとする．このとき，極限定理 (1) は，$n \to \infty$ のとき

$$\overline{X}_n(t) \to \mu t, \quad \text{a.s.} \tag{4}$$

と書き直すことができる． (4) は**大数の関数強法則** (functional strong law of large numbers: FSLLN) とよばれ，流体近似の理論的根拠を与えている．また，中心極限定理については，(2) と同じ尺度変換を用いて，確率過程

$$X_n(t) = n^{1/2}(n^{-1} S_{[nt]} - \mu t) \tag{5}$$

を考えると，(2) は，各 $t \geq 0$ に対して

$$X_n(t) \xrightarrow{\text{w}} \Sigma^{1/2} N(0, t \cdot I) \tag{6}$$

を証明していることに他ならない．X_n が定常独立増分をもつことから，(6) を X_n の有限次元分布へ拡張することができる．すなわち，$0 \leq t_1 < t_2 < \cdots < t_m$ に対し，$n \to \infty$ のとき

$$(X_n(t_1), \ldots, X_n(t_m)) \xrightarrow{\text{w}} (X(t_1), \ldots, X(t_m))$$

が成り立つ．ここで，$(X(t_1), \ldots, X(t_m))$ は平均と分散が，$1 \leq i, j \leq m$ に対して

$$\begin{aligned} &E[X(t_i)] = 0, \\ &E[X^T(t_i) X(t_j)] = \min(t_i, t_j)\, \Sigma \end{aligned} \tag{7}$$

で表される多変量正規確率ベクトルである． (7) によって表される有限次元分布をもつ $D_E[0, \infty)$ 上の唯一の確率過程はブラウン運動過程であるから，$X \in D_E[0, \infty)$ であれば $X \stackrel{\text{d}}{=} \Sigma^{1/2} B$ でなければならない．ここで，B は次の 3 つの性質を満たす $\mathbb{R}^d$ 上の標準ブラウン運動過程である．

(a) $B(\cdot) \in C_E[0, \infty)$
(b) $B(\cdot)$ は定常独立増分をもつ
(c) $B(t) \stackrel{\text{d}}{=} N(0, t \cdot I)$

以上をまとめると，**関数中心極限定理** (functional central limit theorem: FCLT) とよばれる次の結果を得る．

命題 1 ランダムウォーク $\{S_n; n \geq 0\}$ に対し，確率変数列 $\{X_n; n \geq 0\}$ を (5) で定義する．$E[||Z_n||^2] < \infty$ ならば，$D_E[0, \infty)$ において $n \to \infty$ のとき，$X_n \xrightarrow{\text{w}} \Sigma^{1/2} B$ が成り立つ．

明らかに，命題 1 は拡散近似の原型，すなわち

$$S_{[nt]} \stackrel{\text{d}}{\approx} n\mu t + n^{1/2} \Sigma^{1/2} B(t) \tag{8}$$

の形の近似を示唆している．また，命題 1 は，**連続写像原理** (continuous mapping principle) として知られる次の命題と組合せることでさらに強力なものとなる．

命題 2 C_hを写像 $h : D_E[0, \infty) \to \mathbb{R}^d$ の連続点の集合とする．[1] このとき，$X_n \xrightarrow{\text{w}} X$ $(n \to \infty)$ かつ $P\{X \in C_h\} = 1$ であれば，$h(X_n) \xrightarrow{\text{w}} h(X)$ $(n \to \infty)$ が成り立つ．

つまり，いま $Y_n \xrightarrow{\text{w}} Y$を示したいとしよう．$Y_n$ と Y を $P\{X \in C_h\} = 1$ を満たす関数 h を用いて，$X_n = h(Y_n)$ と $X = h(Y)$ に変換する．このとき，従来の理論によって $X_n \xrightarrow{\text{w}} X$が知られているならば，$Y_n$ は Y へ弱収束する．この方法は，組合せ最適化問題の NP-完全性の証明方法と類似する点があり興味深い．

例として，$\mu = 0$ の 1 次元ランダムウォーク $\{S_n; n \geq 0\}$ に対し，時点 n までの最大値 $\overline{S}_n \equiv \max\{S_k : 0 \leq k \leq n\}$ を考えよう．$\overline{S}_n$ は X_n 上で定義された関数 $h(x) = \max\{x(t) : 0 \leq t \leq 1\}$ を用いて，$\overline{S}_n = n^{1/2} h(X_n)$ と表せる．hが任意の $x \in C_E[0, \infty)$に おいて連続であり，ブラウン運動が連続なサンプルパスをもつことから，$X = \sigma B$ のとき，$P\{X \in C_h\} = 1$ が示せる．ゆえに，命題 1 と命題 2 によって，$h(X_n) \xrightarrow{\text{w}} h(\sigma B)$ が成り立ち，近似

$$\overline{S}_n \stackrel{\text{d}}{\approx} \sigma n^{1/2} \max\{B(t) : 0 \leq t \leq 1\}$$

が得られる．

命題 1で示された基本的な FCLT に対し様々な拡張がなされている．これらの拡張の多くは，和を取る確率変数列 $\{Z_n; n \geq 0\}$ が i.i.d. ではなく，かわりにマルコフ連鎖のようなある種の従属性や緩やかな非定常性をもつ場合の部分和過程 $\{S_n; n \geq 0\}$ に対する極限定理を証明することにその特色がある．部分和過程は待ち行列に現れる確率過程において基本的要素で

[1] 関数 h が $x \in D_E[0, \infty)$において連続であるとは，$D_E[0, \infty)$ 上の Skorohod 距離 d に関して $d(x_n, x) \to 0$ となる点列 $\{x_n; n \geq 0\}$ に対し，$||h(x_n) - h(x)|| \to 0$ が成り立つことをいう．

あり, これらの拡張は非常に重要である. さらに, 連続写像原理を適用することで, かなり広いクラスの確率過程間の弱収束を示すことが可能である.

2.3 待ち行列の拡散極限

これまでに述べてきた確率過程に対する極限定理が, 待ち行列にどのように応用されているかを、最も単純な窓口が1つの待ち行列を例にして示そう. 対象となるのは無限の待合い室をもつ $GI/GI/1$ 待ち行列である. 最初の客は窓口が空である時刻 $T_0 = 0$ に到着し, 時間 V_0 のサービスを受ける. 同様に, 第 n 番目の客は時刻 T_n に到着し, 時間 V_n のサービスを受けるものとする. $U_n = T_n - T_{n-1}$ $(n \geq 1)$ とおき, $\{U_n\}$ および $\{V_n\}$ の各列は i.i.d. で互いに独立であると仮定する. また, 簡単のため, $\{U_n\}$ と $\{V_n\}$ は2次までのモーメントをもつと仮定し, 平均を λ^{-1} と μ^{-1}, 分散を σ_A^2 と σ_S^2 でそれぞれ表すことにする.

$GI/GI/1$ 待ち行列において応用上重要な確率過程は, 時刻 t における系内客数 $Q(t)$ および第 n 番目の客の待ち時間 W_n である. トラフィック密度 $\rho = \lambda/\mu < 1$ のときには, 待ち行列は $\{Q(t); t \geq 0\}$ および $\{W_n; n \geq 0\}$ が確率的に有界であるという意味で安定している. しかし, このことは同時に, 確率関数列 $\{(Q(nt) - a_n t)/b_n\}$ または $\{(W_{[nt]} - a_n t)/b_n\}$ が $D_{\mathbb{R}}[0,\infty)$ の非退化過程に弱収束するための列 $\{a_n\}$, $\{b_n\}$ が存在しないことをも意味している. つまり, 定常状態が存在するような状況では, 時間に比例して発散しない確率過程は, (5) の尺度変換を施すと極限が0に縮退してしまうのである.

逆に $\rho > 1$ のときには, これらの確率過程は拡散極限をもつ. すなわち, $D_{\mathbb{R}}[0,\infty)$ において, $n \to \infty$ のとき

$$n^{1/2}(n^{-1}Q(nt) - (\lambda - \mu)t) \xrightarrow{w} (\lambda^3\sigma_A^2 + \mu^3\sigma_S^2)^{1/2}B(t),$$
$$n^{1/2}(n^{-1}W_{[nt]} - (\mu^{-1} - \lambda^{-1})t) \xrightarrow{w} (\sigma_A^2 + \sigma_S^2)^{1/2}B(t)$$

が成り立つ.

では, $\rho = 1$ のときはどうであろうか? 応用上, $\rho = 1$ を特別視することに何の意味もないが, 理論的には $\rho > 1$ の場合と大きく結果が異なる. 紙面の制約から, 詳細については [20] を参照のこと.

これまでの議論は, $GI/GI/1$ 待ち行列を対象にしていたが, 実は $\rho \geq 1$ のときには, 複数窓口待ち行列 $GI/GI/s$ に対しても同様の結果が成り立つ. すべての窓口におけるサービス時間分布 $F(\cdot) = P\{V \leq \cdot\}$ が同一であると仮定すると, 極限を知る上では, $GI/GI/s$ はサービス時間分布が $F(\cdot/s)$ である $GI/GI/1$ と等価になる. したがって, 例えば系内客数過程に対しては, $\rho > 1$ のとき, $D_{\mathbb{R}}[0,\infty)$ において, $n \to \infty$ のとき

$$n^{1/2}(n^{-1}Q(nt) - (\lambda - s\mu)t) \xrightarrow{w} (\lambda^3\sigma_A^2 + s\mu^3\sigma_S^2)^{1/2}B(t)$$

が成り立つ.

複数窓口待ち行列は, 系内客数に依存する待ち行列の特殊な場合と考えることもできる. 出生死滅過程で定式化される一般的な状態依存待ち行列に対する極限定理については [14] を参照のこと. また, 同じ複数窓口でも無限個の窓口をもつ待ち行列において, 到着時間間隔 $\{U_n\}$ を次第に加速したときに如何なる極限をもつか? という研究も行われている [1].

さらに, 最近の研究として, 広帯域ISDNの性能評価において用いられるMMPP (Markov modulated Poisson process) や MAP (Markovian arrival process) といった確率環境下にある到着過程をもつ待ち行列に対し, 少なくとも1つの環境下で $\rho > 1$ であれば, 拡散極限をもつことが証明されている [3].

しかし, こういった「単体」の待ち行列における拡散極限の導出は, 現在の研究の中心課題ではない. 中心は何と言っても待ち行列ネットワークにある. 簡単のため, 外部からの客の到着を許す開いたネットワークに限定して話を進める. (閉じたネットワークに関しては [2] を参照のこと.) d 個のサービス施設のネットワークを考えよう. d 個の施設の各々は, 到着客を先着順にサービスする単一のサーバーから成っている. ネットワーク内の客の経路選択はマルコフ的であると仮定し, 経路行列を $P = (P_{ij})$ で表すことにする. すなわち, P_{ij} は施設 i でサービスを終えた客が直ちに施設 j へ向かう確率を表し, $1 - \sum_{j=1}^{d} P_{ij}$ は施設 i を出た客がそのままネットワークを離れる確率を表す. 議論を簡単にするために, $P_{ii} = 0$ $(1 \leq i \leq d)$ を仮定し, さらに P は既約で $(I - P)^{-1}$ が存在することを仮定する.

$U_i = \{U_i(m); m \geq 1\}$ をネットワーク外部から第 i 施設への客の到着時間間隔列とする. $\{U_i(m)\}$ は i.i.d. で, 平均 λ_i^{-1} と平方変動係数 $a_i^2 < \infty$ (i.e., $a_i^2 = \lambda_i^2 \mathrm{Var}[U_i(m)]$) をもつものとする. また, 施設 i においてサービスされる第 m 番目の客のサービス要求量を $V_i(m)$ で表す. $\{V_i(m)\}$ は i.i.d. で, 平均が1, 平方変動係数 (=分散) $b_i^2 < \infty$ をもつものとする. $\mu_i(n)$

を第 n 近似ネットワーク内の第 i 施設におけるサービス率とすると, $V_i(m)/\mu_i(n)$ はこの施設での第 m 番目の客の実際のサービス時間を表している. これら $2d$ 個の到着時間とサービス要求量の列は互いに独立であると仮定する. このような待ち行列ネットワークは, 一般化 Jackson ネットワークとよばれる.

$Q_{ni}(t)$ で第 n 近似ネットワークにおける時刻 t での第 i 施設における客数を表し, $Q_n(t) = (Q_{n1}(t), \ldots, Q_{nd}(t))$ とおく. $\lambda = (\lambda_1, \ldots, \lambda_d)$ を外部到着率ベクトルとすると, $\lambda(I-P)^{-1} \equiv \mu = (\mu_1, \ldots, \mu_d)$ はネットワークの内部および外部の両方から施設 i に到着する客の有効到着率になる. $\mu(n) = (\mu_1(n), \ldots, \mu_d(n))$ をサービス率ベクトルとし

$$n^{1/2}(\mu(n) - \mu) \to c \quad (n \to \infty) \tag{9}$$

となる c が存在することを仮定する. このとき, 各施設はトラフィック密度 $\rho_i(n) = \mu_i/\mu_i(n)$ $(1 \le i \le d)$ が 1 に近づき, したがってネットワークは「一様に」重負荷の状態にある. さらに, 初期状態に関して

$$n^{-1/2}Q_n(0) \stackrel{w}{\to} Q(0) \quad (n \to \infty) \tag{10}$$

となる $Q(0)$ が存在することを仮定する. (9) と (10) の仮定の下で, 確率過程 $Z = \{Z(t); t \ge 0\}$ を

$$Z(t) = Q(0) + c(I-P)t + \Gamma^{1/2}B(t)$$

で定義する. ただし, 共分散行列 $\Gamma = (\Gamma_{ij})$ は

$$\begin{aligned}\Gamma_{ij} &= (\mu_i^2 b_i^2 + \lambda_i^2 a_i^2)\delta_{ij} - \mu_i b_i^2 P_{ij} - \mu_j b_j^2 P_{ji} \\ &\quad + \sum_{k=1}^{d} \mu_k P_{ki}(\delta_{ij} - P_{kj} + b_k^2 P_{kj})\end{aligned}$$

で与えられる. ただし, $\delta_{ii} = 1$, $\delta_{ij} = 0$ $(i \neq j)$ である. このとき, 次の極限定理を得る [17]: $D_{\mathbb{R}^d}[0,\infty)$ において

$$n^{-1/2}Q_n(nt) \stackrel{w}{\to} Q \equiv f(Z) \quad (n \to \infty)$$

が成り立つ. ここで, $f : D_{\mathbb{R}^d}[0,\infty) \to D_{\mathbb{R}^d_+}[0,\infty)$ は Z を $\mathbb{R}^d$ の非負領域に反射させる働きをもつ写像である. (f の定義については [17] 参照.)

拡散極限 Q は Jackson ネットワークに見られる定性的構造の多くを受け継いでいる. (9) において, もし $c > 0$ であれば, 各施設は客の到着率よりも速いサービス率をもつことになり, ネットワークは安定する. 実際, Q が定常分布をもつための必要十分条件は $c_i > 0$ $(1 \le i \le d)$ であることが知られている [8]. また, 定常分布が積形式をもつための必要十分条件も導出されている [8].

以上では, すべての施設において一様に $\rho_i(n) \to 1$ $(n \to \infty)$ であることを仮定していたが, 応用上はこれはあり得ないだろう. ネットワーク内の施設は $\rho_i \equiv \lim_{n\to\infty} \rho_i(n)$ の値に応じて, 完全ボトルネック施設 $(\rho_i > 1)$, 平衡施設 $(\rho_i = 1)$, 非ボトルネック施設 $(\rho_i < 1)$ の 3 種類に分類できる. これらの施設が混在している場合に対しても極限定理は拡張されている [2]. 非ボトルネック施設では, その行列長は n の関数としては確率的な意味で有界である. したがって, 空間変数が $n^{1/2}$ の単位で測られる尺度変換においては, $n \to \infty$ のとき, 極限における対応する行列長過程は実際上 0 に縮退してしまう. したがって, 非ボトルネック施設の行列長については, 拡散極限は何の情報も与えず, 残りの平衡および完全ボトルネック施設のみの解析に縮小してしまう. これは**状態空間縮約** (state space collapse) とよばれている. 完全ボトルネック施設の拡散極限における挙動は, 単独の $GI/GI/1$ 待ち行列のそれと一致する.

最近, 活発な研究がなされているもう 1 つの拡張は, 客に複数の種類があり, 各施設において優先権付きサービスが行われる場合である [16]. 高い優先権をもつ客は, 低い優先権しかもたない客に比べて速いスピードでネットワークを通過するため, 拡散極限においては高い優先権の客は無視され, やはり状態空間縮約が起こり得る. 多種類の客のいるネットワークに対する拡散極限の詳細については [7] を参照のこと.

2.4 待ち行列の拡散モデル

拡散極限が理論的に如何に正しいものとはいえ, $\rho \ge 1$ でしか使えないということでは応用上役に立たない. そこで, 対象としている確率過程が以下に定義する 2 つの無限小特性量をもつと仮定できる場合, これらによって特徴付けられる拡散過程で元の過程を近似する方法が拡散モデルの基本的な考え方である.

より明確に述べるために, 対象とする待ち行列の確率過程を $J = \{J(t); t \ge 0\}$ で表すことにする. J が $\mathbb{R}^d$ を値域とするとき, **無限小平均** (infinitesimal mean) $\mu : \mathbb{R}^d \to \mathbb{R}^d$ と, **無限小分散** (infinitesimal variance) $\Sigma : \mathbb{R}^d \to \mathbb{R}^{d \times d}$ とよばれる関数をそれぞれ

$$\begin{aligned}E[J(t+\tau) - J(t) \mid J(u); u \le t] &= \mu(J(t))\tau + o(\tau), \\ \mathrm{Cov}[J(t+\tau) - J(t) \mid J(u); u \le t] &= \Sigma(J(t))\tau + o(\tau)\end{aligned}$$

で定義する．Σ が正定値行列のとき，もし J と同じ無限小特性量と連続なサンプルパスをもつマルコフ過程 X が存在するならば，それは拡散過程でなければならないことが知られている．この結果を用いて，$\rho < 1$ に対しても X を J の近似と考えるのが拡散モデルである．このモデリングに際して

- μ と Σ を如何に求めるか？
- J と X の状態空間を如何に対応付けるか？
- J の境界挙動を X に如何に反映させるか？

という問題が生じる．特に，J が系内客数過程のような離散値確率過程の場合，点を区間に対応させる工夫が要求されることになる．また，J の境界挙動を忠実に X に反映させようとすると，X のサンプルパスが境界において不連続となり，拡散過程の枠組みを逸脱してしまうことにもなりかねない．

$GI/GI/s$ 待ち行列に対する拡散モデル [5, 9, 10] は，この 3 つの問題を次第に解決しながら発展し，定常状態の解析に関してはほぼ完成したと言えるだろう．[14] の状態依存待ち行列の拡散極限に対応する拡散モデルも，[11] で 1 つの解決案が示されている．[3] の確率環境下での拡散極限に対応する拡散モデル (e.g., [19]) には，まだ改良の余地がかなり残されており，今後の発展が期待できる．

待ち行列ネットワークに対する拡散モデルは，拡散極限の世界の数学的要求 (特に境界条件) が強すぎて，[13] 以降，3 次元以上の拡散モデルはないに等しい状態にある．拡散極限の世界の研究の興味の対象は待ち行列ネットワークに集中しているため，境界条件を正確にクリアできなければ，拡散極限の研究者が査読者になっている学術雑誌に拡散モデルの論文が採択される可能性は，現状ではかなり低いと考えた方が良い．関連するアプローチとしては，拡散極限から得られる情報を待ち行列ネットワークの分解近似に応用した研究 [6] がある．

3. 連続から離散へ

拡散極限の世界では，離散的な空間から連続的な空間への一方通行の道しか存在しない．連続的な空間での極限過程を元の過程と照らし合わせて評価する姿勢は最初からないのである．この意味では，「近似」という言葉はもともと拡散極限には相応しくないのかもしれない．ところが，拡散モデルの世界では常に元の過程との比較が要求される．最初からモデリングが違うのであって，拡散モデルは (いわゆる「厳密解」を与える) 離散モデルの近似ではない，と突っ張りたい気持ちはあるが，世間はそれを許してくれそうにはない．拡散極限の気楽な立場が羨ましく感じられることもあるが，結局は離散モデルで間に合う程度の問題を扱っていることに，拡散モデルの存在意義が今一つ不確かであることの本質があるような気がする．ファイナンスのように対象の確率的な因果関係がよく分からない問題，過渡解析のように離散モデルでは手に負えない難しい問題にこそ，拡散モデルの真骨頂があるのだろう．

愚痴はこの辺にして，連続的な空間から離散的な空間への帰り道にはどのような問題が待ち受けているのであろうか? 系内客数過程に関しては，状態空間の対応について往路で解決がついていたとしても

- 離散分布を連続分布から如何に抽出するか？

という離散化の問題が本質的に残ることになる．中心極限定理 (2) より，n が十分大きいとき

$$P\{S_n \le x\} \approx \Phi\left(\frac{x - n\mu}{\sigma\sqrt{n}}\right), \quad x \in \mathbb{R}$$

という近似が得られるが，S_n が整数値のみを取る場合には，経験的に**連続補正** (continuity correction)

$$P\{S_n \le x\} \approx \Phi\left(\frac{x + 0.5 - n\mu}{\sigma\sqrt{n}}\right), \quad x \in \mathbb{R}$$

を施した方が，近似が改善されることが知られている．このような「ずらし」を加えた発見的補正 [5]，境界事象 (e.g., 系が空になる，溢れ等) を特別扱いし，境界条件を工夫したり厳密解を適用する補正 [9, 13] 等が提案されている．

もう 1 つ乗り越えなければならない大きな問題は

- 拡散モデルの解を既存の厳密解と如何に整合させるか？

である．例えば，$GI/GI/s$ 待ち行列に対する初期の拡散モデルは，単純な $M/M/s$ 待ち行列に対してさえも厳密解を与えないものであった [5, 9]．このような既知の離散モデルとの整合性の欠如は，拡散モデルの致命的な欠陥として長らく批判を浴びてきた [21]．私自身の最近の研究はこの欠陥の除去を目指したものであり，出生死滅過程あるいは連続時間マルコフ連鎖で定式化できる任意の待ち行列システムの離散モデルと整合する拡散モデル [10, 11] を構築した．

2.4 節と本節において, 拡散モデルを構築する上での5つの問題点を指摘したが, これらの具体的成果については, [12] を参照されたい.

4. エルムの木陰で

私は文部省からの派遣により, プリンストン大学に1996年3月から8月までの5カ月間滞在し研究する機会を得た. 大学内のセミナーで, Bhaskar Sengupta (NEC Research Institute), Ward Whitt (AT&T Laboratories) らの企業に所属する研究者と交流をもてただけでなく, 滞在中に出席したIFIP WG 7.3 Workshop (Lambertville, NJ, May 27-28) では, Erol Gelenbe, Martin Reiser, Yakov Kogan といった拡散近似の先駆者達とも話し合うことができた. 彼等との意見交換の中で鮮明になったことは, 情報通信システムの性能評価においては, 多種類の客をもつ待ち行列ネットワークの過渡解析が重要な研究課題であるという共通の認識であった. 勿論, すでにこの方向での研究は行われているが, 過渡解析となると研究の方法論は自ずと限られてくる. 本質的に多次元の問題であるため, 境界条件の煩雑さを考えると, 拡散近似よりは流体近似の方が当面有望であろう [18]. 拡散極限における状態空間縮約を乗り越えた新たな拡散モデルへの模索が始まっている.

最後になりましたが, プリンストン大学での研究の場を与えて頂いた小林久志教授, および IFIP Workshop での研究発表を勧めて頂いた高木英明教授 (筑波大学) に感謝します.

参考文献

[1] Borovkov, A., "On limit laws for service processes in multi-channel systems," *Siberian Mathematical Journal*, **8** (1967) 746–763.

[2] Chen, H. and Mandelbaum, A., "Stochastic discrete flow networks: diffusion approximations and bottlenecks," *Annals of Probability*, **4** (1991) 1463–1519.

[3] Choudhury, G.L., Mandelbaum, A., Reiman, M.I. and Whitt, W., "Fluid and diffusion limits for queues in slowly changing environments," *Stochastic Models*, to appear, 1997.

[4] Glynn, P.W., "Diffusion approximations," in: *Handbooks in Operations Research and Management Science* (D.P. Heyman and M.J. Sobel, eds.) pp. 145–198, North-Holland, 1990. (木村訳: 「確率モデルハンドブック」第4章, pp. 137–187, 朝倉書店, 1995.)

[5] Halachmi, B. and Franta, W.R., "Diffusion approximation to the multi-server queue," *Management Science*, **24** (1978) 522–529.

[6] Harrison, J.M. and Nguyen, V., "The QNET method for two-moment analysis of open queueing networks," *Queueing Systems*, **6** (1990) 1–32.

[7] Harrison, J.M. and Nguyen, V., "Brownian models of multiclass queueing networks: current status and open problems," *Queueing Systems*, **13** (1993) 5–40.

[8] Harrison, J.M. and Williams, R.J., "Brownian models of open queueing networks with homogeneous customer populations," *Stochastics*, **22** (1987) 77–115.

[9] Kimura, T., "Diffusion approximation for an $M/G/m$ queue," *Operations Research*, **31** (1983) 304–321.

[10] Kimura, T., "An $M/M/s$-consistent diffusion model for the $GI/G/s$ queue," *Queueing Systems*, **19** (1995) 377–397.

[11] Kimura, T., "Birth-death-based diffusion approximations for queues," Discussion Paper Series A, No. 40, Faculty of Economics, Hokkaido University, 1997.

[12] Kimura, T., "Diffusion models for queues in computer/communication systems: a review," in: *Proceedings of SSOR '97* (Hiroshima, August 18–21, 1997).

[13] Kobayashi, H., "Application of the diffusion approximation to queueing networks. I. equilibrium queue distributions; II. nonequlibrium distributions and applications to computer modeling," *Journal of the Association for Computing Machinery*, **21** (1974) 316–328; 459–469.

[14] Mandelbaum, A. and Pats, G., "State-dependent queues: approximations and applications," in: *Stochastic Networks* (F.P. Kelly and R.J. Williams, eds.), pp. 239–282, The IMA Volumes in Mathematics and its Applicatios, Vol. 71, Spriger-Verlag, 1995.

[15] Newell, G.F., *Applications of Queueing Theory*, Chapman and Hall, 1971. (森村・森訳: 「待ち行列理論の応用 — その新しい方法」, サイエンス社, 1973.)

[16] Peterson, W.P., "Diffusion approximations for networks of queues with multiple customer types," *Mathematics of Operations Research*, **16** (1991) 90–118.

[17] Reiman, M.I., "Open queueing networks in heavy traffic," *Mathematics of Operations Research*, **9** (1984) 441–458.

[18] Ren, Q. and Kobayashi, H., "Transient solutions for the buffer behavior in statistical multiplexing," *Performance Evaluation*, **23** (1995) 65–87.

[19] Takahashi, H. and Akimaru, H., "A diffusion model for queues in randomly varying environment," *The Transactions of the IECE of Japan*, **E-69** (1986) 13–20.

[20] Whitt, W., "Heavy traffic limit theorems for queues: a survey," in: *Mathematical Methods in Queueing Theory* (A.B. Clarke, ed.), pp. 307–350, Spriger-Verlag, 1974.

[21] Whitt, W., "Refining diffusion approximations for queues," *Operations Research Letters*, **1** (1982) 165–169.

