マルコフ性の仮定

高橋　幸雄

1. はじめに

マルコフ・モデルは時間パラメータを含む確率モデルの中で最も簡単なものの1つであるため，よく用いられており，使用範囲も非常に広い．自然科学の分野だけでなく社会科学の分野でも人の動きや心の動きにまで応用されて，重要な解析道具の1つとなりつつある．そしてコントロールや各種の数理計画法とも結びついて新しい理論展開も行なわれている．

マルコフ・モデルの記述方法には種々のものがあるが，大別すると，いくつかの独立確率変数を用いてモデルを表現し，その中からマルコフ連鎖となる部分をとりだして分析する場合と，はじめから状態を決め，その間の推移確率を推移図や推移確率行列などの形で表現していく場合とがある（[4] 第5章）．待ち行列や在庫モデルなどでよく用いられているのは前者であり，後述の信頼性における劣化モデル，医療システムの評価モデルや社会科学の分野で用いられる人員計画における昇進モデル，社会移動モデル，マーケティングにおける銘柄転移モデルなどは後者である．

前者の場合，モデルの適合度は主としてモデルを表現するときの確率変数の独立性や分布のあてはまり工合によって決まるが，後者の場合はマルコフ性の仮定がどの程度満足されるかに大きく依存する．マルコフ性はマルコフ・モデルの大もとの仮定でありながら，扱いがむずかしく，直観的な議論で済まされてしまうことが多い．ここでは，このマルコフ性の仮定にまつわる話，特にマルコフ性の仮定が崩れたときの対応に関する話を集めて紹介しよう．

2. 例と問題

ここでの問題意識をはっきりさせるために，簡単な例を2つほどあげておこう．

例1　劣化モデル

機器の劣化の状態を何段階かに分け，劣化の進行状況をそれらの段階を状態とするマルコフ・モデルで記述することがよく行なわれている．故障状態が吸収状態である．もし劣化が進行するだけでよくなることがなければ，推移確率行列 $P=(p(i,j))$ は上三角行列，つまり $p(i,j)=0, j<i$，となる．劣化がある定められた段階まで進んだならば修理や取替えを行なうという保守ルールを用いると，$j<i$ であるようないくつかの $p(i,j)$ を正にすることができ，その結果，故障状態へゆく確率を小さくして，故障による損害を小さくすることができる．

このモデルでは，どうすればマルコフ性が満たされるように状態をとり推移確率を推定することができるかが問題である．多くの場合，マルコフ性が成り立つように劣化の状態を決めることはなかなかむずかしく，たとえ劣化の過程がマルコフ

たかはし　ゆきお　東北大学　経済学部

連鎖で記述しうるものであったとしても，その劣化の具合が観測し得ないものであることもある．このようなときはどうしたらよいであろうか．

例2　診療システム評価モデル

この例として福富和夫氏による胃集団検診の評価モデルがある（[2] または [4]，p. 139）．患者を，健康者，早期がん，進行がん，手おくれがんの4つに大きく分ける．そして，早期がん，進行がんのグループをさらに，未発見時，治療後1年未満，1～2年，2～5年，に細分化して状態をつくる．これに，進行がん患者の回復状態，およびがん死亡，がん以外死亡の2つの吸収状態を加えて状態空間を構成し，推移確率を推定してマルコフ・モデルを作る．集団胃検診をした場合としない場合ではがん発見率が異なるので，それを推移確率に反映させれば，胃集団検診がどの程度の効果があるのかをこのモデルから評価できる．

このモデルでは，推移確率の推定自体かなりむずかしいものであった．特に胃集団検診によらず外来受診によってがんが発見される確率の推定がむずかしかったようである．これは，問題の性質上，このモデルが対象としている集団（患者と健康者の全体）を特定化してその中の各個人の履歴をすべて調べるというマルコフ・モデルでの常道が使えないことによる．したがって，マルコフ性が満たされているかどうかも調べるのは非常にむずかしい．そのため，マルコフ性の仮定があまり効かない分析方法ないしモデル利用法を工夫しなくてはならない．

3.　マルコフ性とその検定

まず，マルコフ性の仮定の復習から始めよう．いまさら書くまでもないが，有限の状態空間 $S=\{1,2,\cdots,s\}$ 上の確率過程 $\{X_t; t=0,1,2,\cdots\}$ がマルコフ性をもつというのは，任意の $t\geqq 1$ と正整数の組 $(i_0,i_1,\cdots,i_{t-1},j)$ に対して，

(1)　$$P\{X_t=j\,|\,X_0=i_0,X_1=i_1,\cdots,X_{t-1}=i_{t-1}\}=P\{X_t=j\,|\,X_{t-1}=i_{t-1}\}$$

が成り立つことである．モデルが斉時的であれば，この右辺はさらに t にもよらず $p(i_{t-1},j)$ と書ける．

いま $\{X_t\}$ と同じ確率構造をもつ互いに独立な m 個のマルコフ連鎖（と思われている確率過程）の履歴のデータ

(2)　$$x_0{}^1,x_1{}^1,\cdots,x_{T_1}{}^1;\cdots;x_0{}^m,x_1{}^m,\cdots,x_{T_m}{}^m$$

が与えられたものとしよう．$n=\sum_k T_k$ とおく．$f(i_0,i_1,\cdots,i_u)$ を $x_t{}^k=i_0, x_{t+1}{}^k=i_1,\cdots,x_{t+u}{}^k=i_u$ となる k と t の組の数，それらの i_u についての和を $f(i_0,i_1,\cdots,i_{u-1},\cdot)$ とする．

もし推移確率 $p(i,j)$ の間に何も特別な関係が仮定されていなければ，推移確率は最尤法によって，

$$\hat{p}(i,j)=f(i,j)/f(i,\cdot)$$

と推定される．モデルに何か構造が入っていて推移確率の間に何かの関係が仮定されている場合でも，適当な方法で $p(i,j)$ を推定することができる．(1)のマルコフ性が満たされているならば，どのように u および $i_0,i_1,\cdots,i_u$ をとっても，

(3)　$$\frac{f(i_0,i_1,\cdots,i_u)}{f(i_0,i_1,\cdots,i_{u-1},\cdot)}\fallingdotseq\hat{p}(i_{u-1},i_u)$$

となっていなければならない．

しかし，ふつうマルコフ性の検定といって行なっているテストでは，すべての u について(3)をチェックすることは稀で，$u=2$ のときの(3)を統計的に検定しているだけである．この検定は，もしマルコフ性が成り立つならば，

(4)　$$\chi^2=\sum_{i,k,j}\frac{\{f(i,k,j)-f(i,k,\cdot)\,\hat{p}(k,j)\}^2}{f(i,k,\cdot)\,\hat{p}(k,j)}$$

は漸近的に χ^2-分布（推移確率に構造の入っていない場合は自由度 $s(s-1)^2$）にしたがうことを利用する．この $u=2$ のときの(3)の検定は，確率過程 $\{X_t\}$ が2次のマルコフ連鎖，つまり対 (X_{t-1},X_t) のなす確率過程が状態空間 $S\times S$ 上のマルコフ連鎖，であるものとして，（1次の）マルコフ性が成り立つかどうかを検定していることに相当する．

4. 弱マルコフ性の仮定とその検定

上でも述べたように，(4)による検定は，マルコフ性の条件(3)のごく一部しかテストしていない．しかし，$u \geqq 3$ のときに(3)を検定するのはデータの量からむずかしいことが多い．というのは，(2)のデータの個数はおよそn個であるが，$(i_0, i_1, \cdots, i_u)$ の組の数は s^{u+1} 個であり，$f(i_0, i_1, \cdots, i_u)$ の値は平均でも n/s^{u+1} 程度にしかならない．そのため，すべての $f(i_0, i_1, \cdots, i_u)$ が適合度検定に耐えうる程度の値（およそ5以上といわれる）となるためには，データの量がよほど多くなければならないからである．

そこで，マルコフ性の仮定のほうをもう少しゆるめて検討可能な程度の条件のモデルにすることはできないかということが考えられる．このために，まず次のことに注意しよう．

一般の確率過程 $\{X_t\}$ において n ステップの推移確率を考えると，

$$
\begin{aligned}
&P\{X_{t+n}=j \mid X_t=i\} \\
&=\sum_{i_1, \cdots, i_{n-1}} P\{X_{t+1}=i_1 \mid X_t=i\} \\
&\qquad \cdot P\{X_{t+2}=i_2 \mid X_t=i, X_{t+1}=i_1\} \\
&\qquad \cdot P\{X_{t+3}=i_3 \mid X_t=i, X_{t+2}=i_2\} \\
&\qquad \cdots \\
&\qquad \cdot P\{X_{t+n}=j \mid X_t=i, X_{t+n-1}=i_{n-1}\}
\end{aligned} \tag{5}
$$

という展開式が成り立つ．この式は確率の乗法法則を使えば簡単に証明できる．したがって，もし $\{X_t\}$ が任意の t, u および i, j, k に対して，

$$
\begin{aligned}
&P\{X_{t+u}=j \mid X_t=i, X_{t+u-1}=k\} \\
&=P\{X_{t+u}=j \mid X_{t+u-1}=k\}=p(k,j)
\end{aligned} \tag{6}
$$

を満たしているならば，その n ステップの推移確率 $P\{X_{t+n}=j \mid X_t=i\}$ はあたかも $\{X_t\}$ が推移確率行列 $P=(p(k,j))$ をもつマルコフ連鎖であるかのようにして計算できる．(6)を満たす確率過程 $\{X_t\}$ を**弱マルコフ連鎖**とよぼう（これは筆者がここで付けた名前で，まだ市民権は獲得していない）．

マルコフ・モデルで通常考える特性量のうち，代表的ないくつかのものは n ステップの推移確率だけから計算できる．たとえば，エルゴード的マルコフ連鎖の定常分布，吸収的連鎖の吸収確率，平均吸収時間，平均訪問回数などがそうである．そこで，もしモデル分析のときに必要な特性量がこのような種類のものであるならば，何も $\{X_t\}$ がマルコフ連鎖でなくとも弱マルコフ連鎖であれば分析は可能である．そして必要な検定も(3)でなく

$$
\frac{f(i, \cdot, \cdots, \cdot, k, j)}{f(i, \cdot, \cdots, \cdot, k, \cdot)} \doteqdot \hat{p}(k,j) \tag{7}
$$

と，はるかに簡単になる．ここで・は対応するインデックスについての和をとったことを表わす．

このように弱マルコフ・モデルは扱える特性量に制約はあるが，条件がゆるいのであてはまる現象も多く，これからの利用が期待される．特に，チェックすべき条件が(7)のように3つの時点における状態にしか関係しないので，検定のさいにもそれほど多くのデータを必要とせず，マルコフ・モデルよりむしろ扱いやすいのではなかろうか．

5. 高次のマルコフ性の検定と状態の細分化

(4)によるマルコフ性の検定でマルコフ性の仮説が棄却された場合，その対応の仕方にはいろいろありうる．1つは，高次のマルコフ性を検討することである．つまり，1次のマルコフ性が棄却されたのならば，2次のマルコフ性は成り立たないかどうか検定し，2次もだめなら3次ではどうか…と続けていくものである．各次数の検定は，(4)と同様の χ^2-適合度検定を用いればよい．しかし次数が高くなるにつれて上で述べたのと同じデータの量不足がネックとなり，実際にはなかなかうまくいかない．

一般に，マルコフ性が成り立たない場合には，状態を細分化してマルコフ性が成り立ちやすくする必要がある．高次のマルコフ・モデルを考えるのはその1つの手段である．状態を，1期前，2期前，…の状態を用いて細分化するのである．しかし，状態の細分化としては，もっと現象に即した方法もありうる．たとえば上の例1の劣化モデ

ルで状態を細分化するのならば，過去の状態によって細分化するよりは，機器の物理的状態，特にそれまで状態の決定に用いていなかった物理的要因を使って状態を細分化したほうが，モデルのあてはまりも良いであろうし，モデルの意味もはっきりする．しかも状態の数も少なくてすむ．ただこの場合の難点は，新しい状態に対応する推移の履歴のデータを新たに得る必要があることである．くわしい観測がなされている場合はよいが，そうでないときはデータをとり直すか，何らかの工夫をしなければならない．

6. マルコフ連鎖の関数としての表現

上で述べたような状況，すなわちマルコフ性を確保するために状態を細分化したいのに，それに対応するデータが得られないときにはどうしたらよいかを考えてみよう．この状況は数学的には次のように表現できる．

$\{Y_t\}$ を状態空間 $\bar{S}=\{1,2,\cdots,\bar{s}\}$ をもつマルコフ連鎖としよう．$\bar{S}$ をいくつかの状態の組に分割する．その分割を $\{A_1,A_2,\cdots,A_s\}$ とする．いま確率過程 $\{X_t\}$ が次の関係で $\{Y_t\}$ と関連づけられているものとしよう．

(8) $X_t=i \qquad Y_t\in A_i$ のとき

このような $\{X_t\}$ を**マルコフ連鎖の関数**とよぶ．上で述べた状況というのは，$\{X_t\}$ の履歴のデータは得られているが $\{Y_t\}$ のデータは得られないという場合である．

$\{X_t\}$ の確率的構造が完全にわかっていれば，(8)を満たすような $\{Y_t\}$ が存在するか，存在するとすればそれはどういうものか，という問いには代数的な議論を用いて答えることができる（Heller［3］または Rosenblatt［5］第3章参照）．しかし，そのためには $\{X_t\}$ に関するデータが無限に必要であり，この方法はあまり実際的でない．

もう少し実際的な議論としては，Basawa & Prakasa Rao［1］第4章に紹介されている Toe の結果がある．ここでは $\{Y_t\}$ の状態空間 $\bar{S}$ およびその分割 $\{A_1,\cdots,A_s\}$ は知られており，$\{Y_t\}$ の推移確率 $q(i,j;\theta)$ はある未知パラメータ θ が決まればすべてわかるものと仮定する．さらに $\{Y_t\}$ は定常状態にあるものとする．すると $\{q(i,j;\theta)\}$ から $\{Y_t\}$ の定常確率 $\omega(j;\theta)$，および $\{X_t\}$ の定常確率

$$\pi(i;\theta)=\textstyle\sum_{j\in A_i}\omega(j;\theta)$$

は θ の関数として計算することができる．そして，いわゆる χ^2 の形

$$\varphi=\sum_{i=1}^{s}\frac{\{f(\cdot,i)-n\pi(i;\theta)\}^2}{n\pi(i;\theta)} \tag{9}$$

を考えると，これが互いに独立な χ^2-分布にしたがう確率変数の1次結合となっていることが示される．そこで(9)を最小にする θ，すなわち

$$\sum_{i=1}^{s}\left\{\frac{f(\cdot,i)}{\pi(i;\theta)}\right\}^2\cdot\frac{d\pi(i;\theta)}{d\theta}=0$$

の解 θ を未知パラメータの推定値として採用する．この方法で求められる推定量 $\hat{\theta}$ は漸近正規性をもつ一致推定量である．

この方法は $\{Y_t\}$ に関するかなりの知識を必要とするが，それでも $\{X_t\}$ のデータから全体の確率構造が決められるところは面白い．もう少し条件がゆるめられれば実際のモデル化の場合にもかなり使えるであろう．この方法の欠点の1つは，定常確率を使用しているため，上の例1や例2のような吸収的モデルには使えないことである．ここでのアイディアを生かして，吸収的モデルでも使える方法の研究が望まれる．

Toe の方法は $\{Y_t\}$ の推移確率の形がわかっていることが前提であった．しかし，実際の場合には，$\{Y_t\}$については皆目見当がつかないことが多い．むしろ $\{X_t\}$ のデータからどの状態を細分化したら良いかの情報を得て，状態空間 $\bar{S}$ と $\{Y_t\}$ を作り出したいというのが本音であろう．残念ながら，そのような方法の研究はまだ全然といってよいほどなされていない．これからの問題である．この問題がむずかしいのは，Basawa & Prakasa Rao の本の中にもあるように，われわれが考えて

いるような状況の下では，(2)の形のデータに対する尤度関数は非常に多くの非線形項の和となってしまうためである．したがって，尤度関数を基礎とした議論はむずかしく，何か発想の転換が要求される．

7. ε-マルコフ・モデル

マルコフ性が満たされない場合の対応策としてもう1つ，ε-マルコフ・モデルがある．これは，例2のように，マルコフ性(1)または弱マルコフ性(6)が完全にではなくある幅をもってなら満たされると考えられる場合に，動的計画法を用いて特性量の上下界を計算しようというものである（[4]第5章）．話はマルコフ・モデルでも弱マルコフ・モデルでもほとんど同じなので，弱マルコフ性の条件をもとにして説明することにしよう．

いま，$\{X_t\}$ は(6)の弱マルコフ性を近似的に満たしているものとする．つまり，(6)の等式は厳密には成り立たないが，$p(k,j)$に少し幅をもたせた

(10) $$\underline{p}(k,j) \leqq P\{X_{t+u}=j \mid X_t=i, X_{t+u-1}=k\} \leqq \bar{p}(k,j)$$

という不等式は満たしているものとする．特に $\bar{p}(k,j)=(1+\varepsilon)\,p(k,j)$, $\underline{p}(k,j)=(1+\varepsilon)^{-1}p(k,j)$ である場合には，(10)を満たす $\{X_t\}$ を**ε-マルコフ連鎖**とよぶ．

$\{X_t\}$ が(10)を満たしているとき，n ステップの推移確率 $P\{X_{t+n}=j \mid X_t=i\}$ の上下界は次のような手続きで計算できる．

δ_jを第j要素は1，他は0の列ベクトル，η_iを第i要素は1，他は0の行ベクトルとする．また $P\{X_{t+u}=l \mid X_t=i, X_{t+u-1}=k\}$ を第 (k,l) 要素にもつ確率行列を P_u とする．すると(5)から，

(11) $$P\{X_{t+n}=j \mid X_t=i\}=\eta_i P_1 P_2 \cdots P_n \delta_j$$

と書ける．第 (k,l) 要素が区間 $[\underline{p}(k,l), \bar{p}(k,l)]$ に入るような確率行列の全体を$\mathscr{D}$としよう．(11)の右辺で各 P_u を$\mathscr{D}$の中で動かしたときの最大値と最小値を求めれば，それが(11)の左辺の上界，下界となっていることは明らかであろう．この上下界は，次のように動的計画法を使ってベクトル$\bar{\delta}_j^{(u)}$および$\underline{\delta}_j^{(u)}$を $u=0,1,2,\cdots$と順次計算することによって求めることができる．

(12) $$\bar{\delta}_j^{(0)}=\delta_j$$
$$\bar{\delta}_j^{(u)}=\max_{P\in\mathscr{D}} P\bar{\delta}_j^{(u-1)}, \quad u=1,2,\cdots$$

$\eta_i\bar{\delta}_j^{(n)}$ が(11)の上界である．下界は，上式で $\bar{\delta}_j^{(u)}$ を $\underline{\delta}_j^{(u)}$, max を min でおきかえた式から $\eta_i\underline{\delta}_j^{(n)}$ で求められる．

吸収的連鎖における吸収確率はnステップ推移確率の極限なので，その上下界は $\bar{\delta}_j^{(n)}$, $\underline{\delta}_j^{(n)}$ の極限として与えられる．したがって，この上下界は，マルコフ決定理論の政策反復法を用いて直接計算することもできる．nステップ推移確率だけから計算できる他の特性量，エルゴード的連鎖の定常分布(相対訪問回数)，吸収的連鎖の平均吸収時間，平均訪問回数についても，(12)を少し変形すれば同様のことがいえる．

この方法の弱点は(10)の上下限 $\bar{p}(k,l)$, $\underline{p}(k,l)$ をデータから推定するうまい方法がまだ見つかっていない点である．そのため今のところこれらの上下限はケースごとに工夫しなければならない．

上にあげた特性量の上下界に関して，理論的に興味ある結果がもう1つ得られている．それらの上下界は $\mathscr{D}$ の中のある確率行列を推移確率行列とする斉時的マルコフ連鎖によって達成されるということである．この事実を使うと，マルコフ・モデルにおける3種の誤差

(i) マルコフ性が成り立たないことによる誤差
(ii) 推移確率の斉時性が成り立たないことによる誤差
(iii) 推移確率の推定誤差

がこれらの特性量に与える影響の程度を比較することができる．推移確率が斉時的でない場合も，推移確率の推定誤差がある場合も，それらの誤差は(10)の形で表現することが可能であり，それらの影響の程度は(12)またはそれと同様の方法で計算できる．特にもし(i), (ii), (iii)の3種の誤差が(10)の意味で同じ程度であれば，それらが上であげた特性

量におよぼす影響はほぼ同程度と結論できる．

おわりに

以上，簡単に紹介したように，マルコフ性が崩れたときの対応に関する研究はまだまだ少ない．マルコフ性が成り立たないのならマルコフ・モデルはさっさと捨ててしまえという意見もあるが，マルコフ・モデルのもつわかりやすさと柔軟さ，特に推移図のもつわかりやすさと柔軟さには捨てがたいものがある．今後も，現実のモデル化を意識したこのような方向の研究が進められることを祈りたい．

引用文献

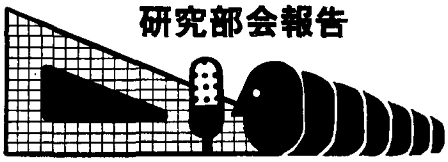
［1］ Basawa, I. V. & B. L. S. Prakasa Rao : *Statistical Inference for Stochastic Processes*. Academic Press, London, 1980

［2］ 福富和夫：システム・モデルによる胃集団検診の評価．オペレーションズ・リサーチ，**21** (1976)，80-86

［3］ Heller, A : On stochastic processes derived from Markov chains. *Ann. Math. Statist.*, **36** (1965)，1286-1291

［4］ 森村英典，高橋幸雄：マルコフ解析，日科技連，1979

［5］ Rosenblatt, M. : *Markov Processes. Structure and Asymptotic Behavior*. Springer-Verlag, Berlin, 1971



環境システム

日時：４月21日(水)16:00～20:00　場所：日科技連
出席者：４名　議題：フローショップスケジューリングにおける Job-base 十分条件，鍋島一郎

経営コンサルタント

• **第26回**　日時：５月８日(土)14:00～17:00
場所：東京都勤労福祉会館　テーマ「経営経済データベースによるコンピュータの活用法」発表者：今村　達（日本ＣＤＣ　コールサービス営業部）

手で提げられるポータブル端末機に電話の送受話器をセットするだけで電話回線で宇宙衛星を介して直接北米ミネアポリスにあるＣＤＣ社の超大型コンピュータを呼びだして使用することにより，グローバルな視点を提供する国際情報データベースを自由に，いつでも，どこでも活用できるようになった．内容は，企業財務・販売情報・金融機関財務情報・証券市況・経済民力統計等々である．

予測とその周辺課題

日時：５月19日(水)18:00～21:00　場所：電力中央研究所　出席者：８名

議題　TIMS 予測文献輪読
予測と白色化フィルター推定について
内容：時系列の予測の問題を，原系列を白色雑音に変換するフィルターの問題としてとらえ，自己回帰等の型について討論を行なった．

環境システム

日時：５月19日(水)18:00～20:00　場所：日科技連
出席者：５名
題目　第１回固定生物膜国際会議に出席して
発表者　有水　彊

生活排水の発生源で所見によって開発された技術を活用して浄化した処理水を，主としてアメリカで開発された大量処理技術と綜合させ広域において土壌浄化技術を実用化するための問題を提示した．

OR/MS とシステム・マネジメント

第１回（４月例会）
日時：４月17日(土)　場所：東京工業大学（長津田キャンパス）出席者：14名

第１回目の研究部会として以下の点について議論が行なわれ，部会メンバーの共通認識を深めた．

(1)本部会の目的，(2)本部会の具体的研究内容，(3)他研究部会との関連，(4)部会の進め方．

以上の点の検討によって，本研究部会は「ＯＲ/ＭＳの研究成果を経営システムのマネジメントに結びつけることを目的とし，ＯＲ/ＭＳと人間心理，行動およびＯＲ/ＭＳと組織の関係の研究，討議，調査分析を行なうこと」を確認した．