数理解析研究所講究録
28巻 1967年 145-158

確率的動的計画における一二の注意

東京女子大　小河原正已

I. 確率的予測の利用について

§1.1 確率的予測論から.

$x=(x_1, x_2, \cdots, x_n)$ および z を正規母集団 $N(\mu, \sigma^2)$ からの大きさそれぞれ $n>1$ および 1 の独立な標本とする. $\bar{x}=n^{-1}\sum x_i$, $s^2=n^{-1}\sum(x_i-\bar{x})^2$ とすると

(1.1) $$T=\frac{z-\bar{x}}{s}\sqrt{\frac{n-1}{n+1}}$$

は $\theta=(\mu, \sigma^2)$ に無関係に, 自由度 $n-1$ の t 分布 $g_{n-1}(t)$ にしたがう.

$$\int_{-\infty}^{t_{n-1}(\alpha)} g_{n-1}(t)\,dt=\alpha \qquad (0<\alpha<1)$$

とすると

(1.2) $$P\{t_{n-1}(\alpha)<T<t_{n-1}(\beta)\}=P\left\{\bar{x}+t_{n-1}(\alpha)s\sqrt{\frac{n+1}{n-1}}<z<\bar{x}+t_{n-1}(\beta)s\sqrt{\frac{n+1}{n-1}}\right\}$$
$$=\beta-\alpha \qquad (0<\alpha<\beta<1)$$

これは母数の信頼区間と同様な意味で, z の <信頼区間> を与える. x を固定し, $\alpha=0$ ($t_{n-1}(\alpha)=-\infty$) とおけば (1.2) から z の分

布函数が導かれる．それは t_{n-1} を自由度 $n-1$ の t 分布にしたがう確率変数とするときの

$$(1.3) \qquad z = \bar{x} + t_{n-1} s \sqrt{\frac{n+1}{n-1}}$$

の分布と同じである．

次に $(\bar{x}, s^2)$ は (μ, σ^2) に対する十分統計量系であるから (μ, σ^2) の信頼分布 (fiducial distribution) $\phi(\mu, \sigma^2 | \bar{x}, s^2)$ が存在する．一方 z は正規分布 $f(z | \mu, \sigma^2)$ にしたがうから

$$(1.4) \qquad \psi(z | \bar{x}, s^2) = \int_{-\infty}^{\infty} \int_{0}^{\infty} f(z | \mu, \sigma^2) \phi(\mu, \sigma^2 | \bar{x}, s^2) d\mu d\sigma^2$$

$$= \frac{1}{B(\frac{1}{2}, \frac{n-1}{2})} \{(n+1)s^2\}^{-\frac{1}{2}} \left[1 + \frac{(z-\bar{x})^2}{(n+1)s^2}\right]^{-\frac{n}{2}}$$

これは (1.3) の分布にほかならない (1.3) の分布を z の確率的予測とよぶのであるが，この考えは一般の時系列に対しても，ある仮定のもとに拡張されるのである．(Ogawara [1])

定義 1.（第一種の確率的予測）

確率過程の確率法則は未知のパラメタの系 θ によって定まるものとする　x, x' は観測値系列の部分集合で $x \cap x' = \phi$ とし，z を予測対象とする．ここで x' を固定するとき，統計量 $T = t(z, x; x')$ の分布は θ に無関係と仮定する．このとき，さらに x も固定して，T の（x を固定しないときの）分布から，z の形式的分布 $\Psi(z | x, x')$ を導く．これを z の第一種の確率的予測 (stochastic prediction) という

定義2（第二種の確率的予測）

x'を与えたときの条件付尤度函数が

$$L(x;\theta|x')=g(t;\theta|x')\,h(x|x'),\quad t=t(x),\quad \theta\in\Theta$$

と書けるとき，$0<\int_{\Theta} g(t;\theta|x')d\theta<\infty$ ならば

$$\phi(\theta|t,x')=g(t;\theta|x')\Big/\int_{\Theta} g(t;\theta|x')d\theta$$

とおく．x''を与えたときの，zの分布函数を$F(z;\theta|x'')$とするとき，

(1.5) $$\Psi(z|t,x',x'')=\int_{\Theta} F(z;\theta|x'')\,\phi(\theta|t,x')\,d\theta$$

を第二種の確率的予測（*fiducial prediction*）という．

以上2つの定義がはじめに示した例のように一致するためにはある条件が必要である（Ogawara[1]）これらの概念は結局（条件付）2標本理論に属するもので Kitagawa[2]の理論で述べられている予測の概念と同様ではあるが，後者は十分統計量の予測を取扱っており，前者は一般には条件付独立性を導入している点など若干異なるように思われる．しかしいずれにしても十分統計量が重要な役割をすることは明かで，両者の関連についての考察は後日にゆずりたい．

定義1（または定義2）による予測の精度は必ずしも最善とはいえない．すなわちzの予測（分布）の分散が常に最小であるとは保証されない．それで大標本の場合には

定義3（近似的確率的予測）

xを与えたときのzの条件付分布を$f(z;\theta|x)$とする．xによるθの推定を$\hat{\theta}$とするとき，$f(z;\hat{\theta}|x)$をもつて，zの確率的予測とすることができる．

しかし，これと$f(z;\theta|x)$との差に関する精密な議論は一般には困難のようである．

なお，R.A.Fisher [3] は上記の定義1の考え方の方を *fiducial prediction* とよんでいる．

例1．正規分布にしたがう単純マルコフ過程

$$x=\{x_{t-1}, x_{t-3}, \cdots, x_{t-2N+1}\}.$$

$$x'=\{x_t, x_{t-2}, \cdots, x_{t-2N}\}$$

$$z=x_{t+1}$$

とし，$\bar{x}_1=N^{-1}\sum x_{t-2\nu+1}$，$\bar{x}_2=N^{-1}\sum x_{t-2\nu}$．

$$x_\nu'=x_{t-2\nu+1}-\bar{x}_1,\quad x_\nu''=x_{t-2\nu}-\bar{x}_2\ (1\le\nu\le N),\quad x_0''=x_t-\bar{x}_2$$

とおき，$\hat{c}=\sum x_\nu' x_\nu''/\sum(x_\nu'')^2$ とおけば，定義1によつて

$$(1.6)\qquad z=\bar{x}_1+\hat{c}x_0''+t_{N-2}\left[\frac{1}{N-2}\sum(x_\nu'-\hat{c}x_\nu'')^2\left(\frac{N+1}{N}+\frac{(x_0'')^2}{\sum(x_\nu'')^2}\right)\right]^{1/2}$$

ただしt_{N-2}は自由度$N-2$のt分布にしたがう確率変数である

$$z=\{x_{t+1}, x_{t+2}, \cdots, x_{t+K}\}$$

の場合も同様である．独立系列の場合は(1.3)となる．

例2．多項分布

K種の事象がn回の試行中それぞれ$t_1,\cdots,t_K$回起きたとき，

さらに m 回の試行で起るべき回数を $z_1, \cdots, z_K$ とすると，

$$(1.7)\quad \psi(z_1, \cdots; z_K \mid t_1, \cdots; t_K) = \frac{(n+K-1)!\, m! \prod_{j=1}^{K} (t_j + z_j)!}{(n+m+K-1)! \prod_{j=1}^{K} t_j! \prod_{j=1}^{K} z_j!}$$

によって $z_1, \cdots, z_K$ の *fiducial prediction* が与えられる.

これは一様事前分布を仮定した場合の Bayesian prediction (Fisher [3]) と一致している.

とくに $m=1$ のときは

$$(1.8)\quad \psi(z_j \mid t_j) = \begin{cases} (t_j+1)/(n+K) & z_j = 1 \text{のとき} \\ 1-(t_j+1)/(n+K) & z_j = 0 \text{のとき} \end{cases} \quad (j=1, 2, \cdots, K)$$

これは推定値 $\hat{p}_j = (t_j+1)/(n+K)$ を与えるものと考えてよい.

また，とくに二項分布 $(K=2)$ の場合には

$$(1.9)\quad \psi(z_j \mid t_j) = \begin{cases} (t_j+1)/(n+2) & z_j = 1 \text{のとき} \\ 1-(t_j+1)/(n+2) & z_j = 0 \text{のとき} \end{cases} \quad (j=1, 2)$$

状態 $e_1, e_2, \cdots, e_K$ のマルコフ連鎖 $x(t)$ の実現が $x(n-j) = x_{n-j}$ $(j \geq 0)$ のとき，$y_j = x_{n-2j+1}$，$x'_j = x_{n-2j}$ $(j \geq 1)$ とする. $\{x'_j\}$ の中で $x_n = e_r$ と同じものの個数を N_r，$(x'_j = e_r, y_j = e_s)$ なる組の数を N_{rs} $(s=1, 2, \cdots, K)$ とすれば $x(n+1)$ の *fiducial prediction* は

$$(1.10)\quad \hat{p}_{rs} = (N_{rs}+1)/(N_r+K) \quad (s=1, \cdots, K)$$

例3．ポアソン過程

過去 n 期間に t 回生起したとき，次の1期に起る回数を Z と

すれば定義2によって

$$(1.11)\quad \psi(z\,|\,t)=\left(\frac{n}{n+1}\right)^{t+1}\binom{t+z}{z}\left(\frac{1}{n+1}\right)^{z}\qquad (z=0,1,2,\cdots)$$

§1.2 初期決定問題

たとえば，多次元配分過程において，制約条件を

$$\sum_{j=1}^{N} a_{ij}(x_j)\leq c_i \qquad i=1,2,\cdots,M.$$

とし，確率変数 z_j の分布函数を $\Phi_j(z)$ とすると

$$(1.12)\quad f_N(\{c_i\})=\underset{x_1,\cdots,x_N}{\mathrm{Max}}\sum_{j=1}^{N}\int g_j(x_j,z)\,d\Phi_j(z)$$

$$=\underset{\{a_{iN}(x_N)\leq c_i\}}{\mathrm{Max}}\left[\int g_N(x_N,z)\,d\Phi_N(z)+f_{N-1}\left(\{c_i-a_{iN}(x_N)\}\right)\right].$$

ここで z_j の分布函数が未知の場合に，過去において，z_j の実現値が直接または間接に知られているならば，それによる確率的予測を $\Phi_j(z)$ とすればよい．

§1.3 適応制御過程

x_t：状態変数，y_t：制御変数，r_t：定常確率過程　のとき過程の構造を

$$x_{t+1}=g(x_t,y_t,r_t)\qquad (t=0,1,\cdots,N)$$

効用函数を

$$u_t(x_t, y_t, x_{t+1}) = u_t(x_t, y_t, g(x_t, y_t, r_t)) \equiv v_t(x_t, y_t, r_t)$$

とする．$x_0 = c$ から出発して，$U_N = \sum_{t=0}^{N} v_t$ の期待値を最大にするのが目的である．ここで $\{r_t\}$ は $\{x_t\}$，$\{y_t\}$ には無関係で，その確率法則は未知であるが，g の構造と，x_t，y_t およびその過去の値から，r_t の実現値が求められるものとする．その標本系列 $\{r_t, r_{t-1}, \cdots, r_0, r_{-1}, \cdots\}$ から r_t の確率的予測が求められるとし，それを $\Psi_\tau(r)$ $(\tau = 0, 1, \cdots; t)$ とする．この予測はデータがふえるにしたがって改善されてゆく．この確率的予測による期待値を $\tilde{E}$ で表わす．

$$f_N(c) = \underset{\{y_0, \cdots, y_N\}}{Max} \tilde{E}\, U_N = \underset{\{y_0, \cdots, y_N\}}{Max} \sum_{t=0}^{N} \int v_t(x_t, y_t, r)\, d\Psi_t(r) \quad (N \geq 1)$$

ここで

$$f_0(c) = \underset{y_0}{Max} \int v_0(c, y_0, r)\, d\Psi_0(r)$$

$$f_1(c) = \underset{y_0}{Max} \int \left[v_0(c, y_0, r_0) + f_0(g(c, y_0, r_0)) \right] d\Psi_0(r_0)$$

ただし

$$f_0(g(c, y_0, r_0)) = \underset{y_1}{Max} \int v_1(g(c, y_0, r_0), y_1, r_1)\, d\Psi_1(r_1)$$

一般に

$$(1.13) \quad f_N(c) = \underset{y_0}{Max} \int \left[v_0(c, y_0, r_0) + f_{N-1}(g(c, y_0, r_0)) \right] d\Psi_0(r_0)$$

例1. $r_t = \pm 1$ で独立の場合における，最終状態の函数 $u_N(x_N, y_N, x_{N+1})$ の制御の問題．この場合には $u_t = v_t = 0$

$(t=0,1,\cdots,N-1)$　$t=0$ までに $r=+1$ が m 回，$r=-1$ が n 回起きたことが経験されているときは

$$(1.14)\quad f_N(c,m,n)=\underset{y_0}{\mathrm{Max}}\left[\psi_{mn}f_{N-1}(c_+,m+1,n)+(1-\psi_{mn})f_{N-1}(c_-,m,n+1)\right]$$

ただし　$c_+=g(c,y_0,+1)$,　$c_-=g(c,y_0,-1)$

また (1.4) によって　$\psi_{mn}=(m+1)/(m+n+2)$.

r_t がマルコフ連鎖をなすときは (1.10) を用いればよい.

また r_t が一般的な未知の分布にしたがって独立のときは，近似的に多項分布の場合の (1.8) によることができる.

さらにまた，r_t が正規自己回帰過程のときは (1.6) またはこれを一般にした式が利用される.

例2　マルコフ決定過程.

マルコフ決定過程は推移確率行列は既知として論じられる場合が多いが，実際にはこれを経験的に推定しなければならないであろう.

M種類の決定に対応する推移確率行列を $(p_{ij}(q))$ $(q=1,2\cdots M)$ とするとき，情報の増加に応じて，その推定 $(\tilde{p}_{ij}(q))$ を (1.10) によって計算してゆけばよいわけである.

なお利得や割引率も確率変数と考えるべき場合もあるが，その取扱いは今まで述べた方法に準じてなされよう.

II. 分散の問題に関連して

確率的動的過程に対し，その目的函数の期待値を最適化することが普通に行われているが，その分散を求めることが必要であり，また多くの場合に可能でもある．しかし，目的函数の分布を制御変数の函数として求め，それを制御するのが理想的ではあるが，これは一般には困難である．ここではこれに若干の関連をもつ，一二の特殊な場合をあげる．

§2.1 Bang-bang 制御

定数係数の確率的線形微分方程式（形式的表現による）

$$\frac{d^n x}{dt^n}+a_1\frac{d^{n-1}x}{dt^{n-1}}+\cdots+a_n x(t)=u(t)+\mathcal{E}'(t) \tag{2.1}$$

において，$\mathcal{E}(t)$ は Wiener 過程で $E\left(d\mathcal{E}(t)\right)^2=\sigma^2 dt$，$u(t)$ は制御函数で

$$|u(t)|\leq 1 \tag{2.2}$$

とする．

$$x(0)=x_0^1,\ x'(0)=x_0^2,\ \cdots,\ x^{(n-1)}(0)=x_0^n \tag{2.3}$$

から出発して，最短時間で $x(t)=x'(t)=\quad=x^{(n-1)}(t)=0$ に〈近づく〉ように制御する場合には，形式的には期待値 $Ex(t)$ の制御の場合と同様になる，という意味において，$\mathcal{E}(t)$ を〈直接に〉の影響制御することはできないが，$\mathcal{E}(t)$ は点 $\left(x(t),\ x'(t),\cdots,x^{(n-1)}(t)\right)$ の位置

に影響を与え，$u(t)$ はその点の位置に関係して $+1$ または -1 をとることになる.

すなわち，$x^1(t)=x(t)$ として $\left(x^0=t,\ \frac{dx^0}{dt}=1\right)$

$$(2.4)\qquad \begin{cases} \dfrac{dx^1}{dt}=x^2(t),\ \cdots,\ \dfrac{dx^{n-1}}{dt}=x^n(t) \\ \dfrac{dx^n}{dt}=-a_1x^n-a_2x^{n-1}-\cdots-a_nx^{(1)}+u+\varepsilon' \end{cases}$$

Pontryagin [5] の記法で

$$(2.5)\qquad H(\psi,x,u,\varepsilon)=\psi_1x^2+\psi_2x^3+\cdots+\psi_{n-1}x^n \\ +\psi_n(-a_1x^n-\cdots-a_nx'+u+\varepsilon')$$

随伴微分方程式は $\left(\frac{d\psi_0}{dt}=0\right)$

$$(2.6)\qquad \begin{cases} \dfrac{d\psi_1}{dt}=a_n\psi_n \\ \dfrac{d\psi_2}{dt}=-\psi_1+a_{n-1}\psi_n \\ \cdots\cdots \\ \dfrac{d\psi_{n-1}}{dt}=-\psi_{n-2}+a_2\psi_n \\ \dfrac{d\psi_n}{dt}=-\psi_{n-1}+a_1\psi_n \end{cases}$$

したがって

$$(2.7)\qquad \frac{d^n\psi_n}{dt^n}-a_1\frac{d^{n-1}\psi_n}{dt^{n-1}}+\cdots+(-1)^na_n\psi_n=0$$

$$(2.8)\qquad \psi_n=c_1e^{\lambda_1t}+\cdots+c_ne^{\lambda_nt}$$

ただし $\lambda_j=\bar{\lambda}_k$ のときは $c_j=\bar{c}_k$，重根がある場合も同様.

かくして，個々の標本函数 $\varepsilon(t)$ に対して

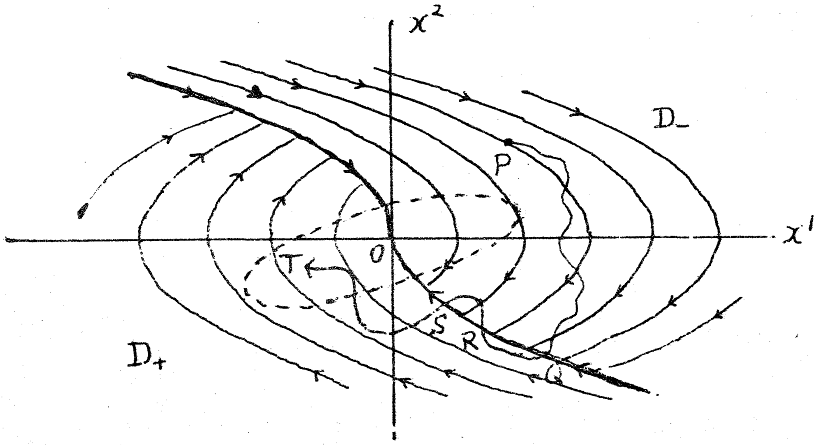
$$(2.9)\qquad \sup_{|u|\le 1} H(\psi, x, u, \varepsilon) = \psi_1 x^2 + \cdots + \psi_{n-1} x^n - \psi_n(a_1 x^n + \cdots + a_n x^1 - \varepsilon') + \sup_{|u|\le 1} \psi_n u$$

ゆえに

$$(2.10)\qquad u(t) = \operatorname{sign} \psi_n$$

例1　$n=2$ の場合

$$(2.11)\qquad \frac{d^2x}{dt^2} = u(t) + \varepsilon'(t) \qquad \text{(形式的に)}$$



解は Pontryagin[5] の例のように, 領域 D_+ では $u(t)=+1$, D_- では $u(t)=-1$ とすればよい　実際には点Pから出発して, たとえば図のように $P \to Q \to R \to S \to T \to \cdots$ のような経路を辿るであろう, そして原点Oに達することはまずない. 実際吾々は $E x^1(t)=0$ $E x^2(t)=0$ となる最小時間 t と, そのための制御を求めたのと同じことをしたという結果になっている.

$\varepsilon(0)=0$ とすると, どこから出発しても, t 時間後には

$$\left\{\ \sigma_1^2(t) = V(x^1) = \sigma^2 t^3/3\right.$$

$$
(2.12) \qquad \begin{cases} \sigma_2^2(t) = V(x^2) = \sigma^2 t \\ \mathrm{cov}(x^1, x^2) = \sigma^2 t^2/2 \\ \rho = \rho(x^1, x^2) = \sqrt{3}/2 \end{cases}
$$

上記の制御によって $Ex^1(t) = Ex^2(t) = 0$ となる時間（最短時間）を t_0 とすると，t_0 時間後に楕円

$$
(2.13) \qquad \frac{(x^1)^2}{\sigma_1^2(t_0)} + \frac{(x^2)^2}{\sigma_2^2(t_0)} - \frac{2\rho x^1 x^2}{\sigma_1(t_0)\sigma_2(t_0)} = (1-\rho^2)\chi_2^2(\alpha)
$$

の中に点が存在する確率は $1-\alpha$ である．ここに $\chi_2^2(\alpha)$ は自由度2の χ^2 分布の $100\alpha\%$ 点である．これがすなわち，はじめに述べた＜近ずく＞の意味である．

§2.2 分散の制御

例2．確率的差分方程式

$$
(2.14) \qquad x_n - \rho x_{n-1} = u_n + \varepsilon_n \quad (n=0,1,\cdots)
$$

において，$\{\varepsilon_n\}$ は直交確率変数列で $E\varepsilon_n = 0$，$V(\varepsilon_n) = \sigma_\varepsilon^2$ とする．$\{u_n\}$ は制御変数列である．$x_0 = a$ とすれば

$$
(2.15) \qquad Ex_n = \rho^n a + \rho^{n-1} u_1 + \cdots + u_n
$$

であって，期待値は適当に制御することができるが

$$
(2.16) \qquad \begin{aligned} V(x_n) &= \frac{1-\rho^{2n}}{1-\rho^2}\sigma_\varepsilon^2 \qquad (|\rho| \neq 1) \\ &= n\sigma_\varepsilon^2 \qquad (|\rho| = 1) \end{aligned}
$$

であって，分散は制御できない．この点は例1と同様である．

次にもし制御変数を

(2.17) $$u_n = c\, x_{n-1} \qquad |c| \le 1$$

とすれば

(2.18) $$E\, x_n = (\rho + c)^n a$$

(2.19) $$V(x_n) = \frac{1-(\rho+c)^{2n}}{1-(\rho+c)^2}\sigma_\varepsilon^2 \qquad (|\rho+c| \neq 1)$$

$$= n\,\sigma_\varepsilon^2 \qquad (|\rho+c| = 1)$$

おのおののnに対し，この分散を最小にする制御は

(2.20) $$\begin{cases} \rho > 1 \text{ のとき} & c = -1 \\ \rho < -1 \text{ のとき} & c = 1 \\ |\rho| \le 1 \text{ のとき} & c = -\rho \end{cases}$$

これはまた $|E\, x_n|$ を最小にする制御でもある.

(2.21) $$x_n - \rho_n x_{n-1} = u_n + \varepsilon_n$$

(2.22) $$u_n = c_n x_{n-1}$$

の場合にも，$E\, x_{n-1}\varepsilon_n = 0$ のときは，$V(x_n) = \sigma_n^2$ とおいて

(2.23) $$\sigma_n^2 = (\rho_n + c_n)^2 \sigma_{n-1}^2 + \sigma_\varepsilon^2$$

であるから，

(2.24) $$\begin{cases} \rho_n > 1 \text{ のとき} & c_n = -1 \\ \rho_n < -1 \text{ のとき} & c_n = 1 \\ |\rho_n| < 1 \text{ のとき} & c_n = -\rho_n \end{cases}$$

参 考 文 献

[1] M. Ogawara: Time series analysis and stochastic prediction. II. Bull. Math. Statist. Vol. 8, No. 3~4, 55-72, Vol. 9, No. 1, 1-9 (1959)

[2] T. Kitagawa: Successive process of statistical inference associated with an additive family of sufficient statistics. Bull. Math. Statist. Vol. 7, No. 3~4. 92-112 (1957)

[3] R.A. Fisher: Statistical methods and scientific inference (1956)

[4] 宮沢光一: 多段決定問題における情報構造. 季刊理論経済学, Vol. 17, No. 2, 62-74 (1966)

[5] L.S. Pontryagin et al.: The mathematical theory of optimal processes (1962).