確率的計画法とサンプル情報

森田　浩

1. はじめに

われわれが意思決定を行なうときには，さまざまな制約条件とともに不確実性も考慮しなければならないことが多い．したがって，計画問題がいろいろな分野で応用されるとき，その大部分のものは係数のいずれかの中に不確実性を含んでいることになる．この不確実性を確率的な変動として扱っているのが確率的計画法である．標本空間ΩとA上の確率測度Pをもつ確率空間 (Ω, A, P) を考えるとき，確率的計画法の一般形は次のような期待値関数によってに表わすことができる．

$$(1)\quad \begin{cases} \text{最小化} & E\{f_0(x,\omega)\}=\int_\Omega f_0(x,\omega)P(d\omega), \\ \text{条件} & E\{f_i(x,\omega)\}=\int_\Omega f_i(x,\omega)P(d\omega)\le 0, \\ & \qquad i=1,2,\cdots,m, \\ & x\in X\subset R^n,\ \omega\in\Omega. \end{cases}$$

ただしXはR^nの閉部分集合であり，$f_0: R^n\times\Omega\to R\cup\{-\infty,+\infty\}$と$f_i: R^n\times\Omega\to R, i=1,2,\cdots,m$はランダム下半連続関数である．この一般形は確率的計画法の等価確定モデルとして広く使われているリコースモデルや機会制約条件モデルなども含んでいる[7]．Pが既知ならば問題 (1) は定義可能だが，多くの実際の問題ではPについての完全情報が与えられるとは必ずしも限らない．つまり確率変動に関する不完全情報のもとでの確率的計画問題を考えなければならないことが起こるのである．

このような状況で問題 (1) の最適解を求めるための1つの方法は，観測値データにもとづいて求められたPの推定値あるいは近似値をPの代わりに用い，そのときに得られる確率的計画問題の最適解から推定する方法である．大きさ$\nu(<\infty)$のサンプルにもとづいて求められたPの推定値をP^νとするとき，不完全情報のもとでの確率的計画問題

$$(2)\quad \begin{cases} \text{最小化} & E^\nu\{f_0(x,\omega)\}=\int_\Omega f_0(x,\omega)P^\nu(d\omega), \\ \text{条件} & E^\nu\{f_i(x,\omega)\}=\int_\Omega f_i(x,\omega)P^\nu(d\omega)\le 0, \\ & \qquad i=1,2,\cdots,m, \\ & x\in X\subset R^n,\ \omega\in\Omega. \end{cases}$$

が得られるが，この問題(2)の最適解から問題(1)の最適解を推定しようとするのである．この方法においては，得られたサンプルは問題(1)の最適解を推定するにあたってどの程度の価値を持っているのか，サンプルの大きさはいくつにするのが適当か，問題(1)と問題(2)の最適解にどのような関係があるか，などといったことに興味が持たれるであろう．

確率的計画法におけるサンプル情報の価値の考え方は Bracked & Soland［2］が示し，さらに Jagannathan［8］がリコースモデルと機会制約条件モデルに対するサンプル情報の価値，およびそれにもとづく最適サンプル数について示している．また不完全情報のもとでの“最適解”を与えるいくつかの確率計画モデルも確率的線形計画問題に対して提案されている．真の最適解あるいは最適値の取り得る範囲の予測［3］，ゲーム理論的なミニマックスアプローチ［4, 10］などである．さらにサンプル数が十分に大きくなったとき，ある条件下で問題(2)の最適解が問題(1)の最適解に収束すること［5,6］や逐次的にサンプルが得られる場合の adaptive な確率計画モデル［9］などの提案もなされている．

本稿ではここに挙げた確率計画モデルの中から，サンプル情報を取り入れたいくつかの確率的線形計画モデルを紹介する．

2. ミニマックスアプローチ

次の線形計画問題(3)において，いくつかの係数が確率変数であるとする．

$$(3)\quad \begin{cases} \text{最大化} & c'x, \\ \text{条件} & Ax=b,\ x\ge 0. \end{cases}$$

ここでAは$m\times n$行列，bはm次元列ベクトル，cとxはn次元列ベクトルである．非負で凸のリコース関数をϕとするとき，問題(3)のリコースモデルは次のようになる．

$$(4)\quad |\ \text{最大化}\quad E_F\{c'x-\phi(x;A,b)\}.$$

確率変数の分布Fが既知であれば，問題(4)を非線形の

もりた　ひろし　大阪市立大学　商学部
〒558 大阪市住吉区杉本 3—3—138

確定問題として扱うことができるが，ここでは分布Fについて不完全な情報しか与えられていない場合を考える．すなわち，不完全情報のために，分布Fについては“ある分布の集合$\mathscr{F}$に属している”ということだけしかわかっていないものとする．簡単のためbのみが確率変数であるとすると，たとえば，集合$\mathscr{F}$はサンプル情報にもとづいて，(5)式のようにモーメント値を規定することによって定義されたり［4］，(6)式のように分布Fがある特定の分布形にしたがうということだけがわかっているときにその分布パラメータを推定することによって定義されたりする［9］．

(5) $\mathscr{F}=\{F \mid E_F b_i=\mu_i, V_F b_i=\sigma_i^2, 1\leq i\leq m\}$

(6) $\mathscr{F}=\{F \mid (\mu, \sigma^2)\in S_\alpha, F \text{ は正規分布}\}$

ここで，μ_iとσ_i^2はサンプルから計算されたb_iの平均と分散で，S_αは正規分布の分布パラメータの信頼領域（信頼係数α）である．この$F\in\mathscr{F}$なる条件のもとで，次のミニマックス問題(7)の最適解を問題(4)に対するミニマックス解と呼ぶ．

(7) $\left|\ \text{最大化}\ \min_{F\in\mathscr{F}} E_F\{c'x-\phi(x;A,b)\}.\right.$

ミニマックス解は，直面している不確実性の状況において想定される最悪の状況下で最適化された解であると見ることができる．

(5)式の集合$\mathscr{F}$に対するミニマックス問題(7)はモーメント問題と呼ばれている．これは確率変数の分布形がわかっていないとき，平均や分散などのモーメント値をサンプル情報から与えることで分布Fの属している集合$\mathscr{F}$を定めるモデルである．このとき，規定されたモーメントの数をkとすると，いわゆる“最悪の分布F^*”は高々$k+1$個の点の離散分布となる［4］．例として，bのみを確率変数とした単純リコース問題を考えると，リコース関数は次のように与えられる．

(8) $\phi(x;b)=\sum_{i=1}^{m} w_i(A_ix-b_i)^+,\ \ w_i\geq 0$

ただし，$(z)^+=\max(z,0)$，A_iはAの第i行ベクトルである．このとき問題(7)における決定変数xに関する最大化と分布Fに関する最小化の操作の交換が可能となり，さらにリコース関数の値の“最悪の場合”の期待値は次のような凸関数となる．

(9) $$\max_{F\in\mathscr{F}} E_F\phi(x;b)=\sum_{i=1}^{m}\frac{w_i}{2}\left\{\sqrt{\sigma_i^2+(A_ix-\mu_i)^2}-(A_ix-\mu_i)\right\}$$

(6)式の集合$\mathscr{F}$に対するミニマックス問題(7)では，確率変数の分布形がたとえば正規分布というようにわかっているとき，まずサンプル情報からその分布パラメータ（平均と分散）を推定する．そして分布パラメータをある信頼係数で推定された信頼領域に制限することで，分布Fの属している集合$\mathscr{F}$を定める．例として，bのみを確率変数とした2次リコース問題を考えると，リコース関数(10)およびその期待値(11)は次のように与えられる．

(10) $\phi(x;b)=\sum_{i=1}^{m} w_i(A_ix-b_i)^2,\ \ w_i\geq 0,$

(11) $E_F\phi(x,b)=\sum_{i=1}^{m} w_i\{(A_ix-\mu_i)^2+\sigma_i^2\}$

正規分布に対しては平均μと分散σ^2の信頼領域は

(12) $$S_\alpha=\{(\mu_i,\sigma_i^2),\ i=1,2,\cdots,m \mid \sum_{i=1}^{m}\frac{(\mu_i-\bar{\mu}_i)^2}{s_i^2}\leq\frac{m(N-1)}{N(N-m)}F_{\alpha/2}(m,N-m),\ \frac{(N-1)s_i^2}{\chi^2_{\beta/2}(N-1)}\leq\sigma_i^2\leq\frac{(N-1)s_i^2}{\chi^2_{1-\beta/2}(N-1)},\ i=1,2,\cdots,m\}.$$

となる．ただし，Nはサンプル数，$\bar{\mu}$とs^2はサンプル平均とサンプル分散，$F_\alpha(m,n)$は自由度対(m,n)のF分布のα%点，$\chi^2_\beta(n)$は自由度nのカイ2乗分布のβ%点$(\alpha=\beta^{1/m}/2)$である．このときリコース関数の値の“最悪の場合”の期待値は

(13) $$\max_{F\in\mathscr{F}} E_F\phi(x;b)=\sum_{i=1}^{m} w_i\left\{\left(\frac{\lambda(A_ix-\bar{\mu}_i)}{\lambda-w_is_i^2}\right)^2+\frac{(N-1)s_i^2}{\chi^2_{1-\beta/2}(N-1)}\right\}$$

となるが，かっこ内の第2項は決定変数xによらない定数であるから，以下のミニマックス問題では省略することができる．ただし，λはラグランジェ乗数で次の式を満足していなければならない．

(14) $$\sum_{i=1}^{m}\frac{w_i^2s_i^2(A_ix-\bar{\mu}_i)^2}{(\lambda-w_is_i^2)^2}=\frac{m(N-1)}{N(N-m)}F_{\alpha/2}(m,N-m)$$

$$\lambda>\max_{i\in I} w_is_i^2,\ \ I=\{i \mid A_ix-\bar{\mu}_i\neq 0\}$$

このときミニマックス解を与える問題

(15) $$\left|\begin{array}{ll}\text{最大化} & c'x-\sum_{i=1}^{m} w_i\left(\dfrac{\lambda(A_ix-\bar{\mu}_i)}{\lambda-w_is_i^2}\right)^2,\\ \text{条件} & (14).\end{array}\right.$$

は凸計画とはならないが，$\lambda\geq 3/2\,\lambda_{\max}$（$\lambda_{\max}$は$w_is_i^2$の最大値）に限れば$x$と$\lambda$に関する凸計画となる．また，$\lambda$をパラメータとみたとき，問題(15)は$x$に関する凸計画となり，その時の最適解$x^*(\lambda)$は$\lambda$に関して連続性を持つ．以上から，数値解法によりミニマックス解$x^*(\lambda^*)$

を求めることができる[10].

確率的線形計画問題で信頼領域を用いて集合$\mathscr{F}$を定めるミニマックスモデルは，目的関数の係数を推定する問題[11]や制約式の係数行列の中にある未知係数を推定する問題[12]についても考察されている.

3. 最適解の推定

問題(2)の最適解x^νはサンプルの大きさνが大きくなるにつれて問題(1)の最適解x^*に収束することが望まれる．次の定理は，問題$\min E^\nu\{f(x,\omega)\}$の最適解$x^\nu$が問題$\min E\{f(x,\omega)\}$の最適解$x^*$に収束することに関するものである.

定理1 [6] すべての$x\in X$に対して$\omega\mapsto f(x,\omega)$が$\Omega$上の連続関数であり，かつ，すべての$\omega\in\Omega$に対して$x\mapsto f(x,\omega)$が$R^n$上の下半連続関数であるとする．さらに，サンプリングによって得られる測度P^νの列$\{P^\nu\}_{\nu=0}^\infty$がPに分布収束し，かつ，$f(x,\cdot)$-tight である，すなわち，すべての$x\in X$と$\varepsilon>0$に対して

(16) $$\int_{\Omega\setminus K_\varepsilon}|f(x,\omega)|P^\nu(d\omega)<\varepsilon,\ \ \nu=0,1,\cdots$$

となるコンパクト集合$K_\varepsilon\subset\Omega$が存在するとする．このとき

(17) $$\int_\Omega f(x,\omega)P(d\omega)=\operatorname*{epi-lim}_{\nu\to\infty}\int_\Omega f(x,\omega)P^\nu(d\omega)$$

が得られる[1]．また

(18) $$\arg\min\int_\Omega f(x,\omega)P^\nu(d\omega)\cap D\neq\phi\ \ (a.s.)$$

(19) $$\{x^*\}=\arg\min\int_\Omega f(x,\omega)P(d\omega)\cap D$$

であるコンパクト集合$D\cap R^n$が存在すれば

(20) $$x^*=\lim_{\nu\to\infty}x^\nu$$

(21) $$\inf\int_\Omega f(x,\omega)P(d\omega)=\lim_{\nu=\infty}\left(\inf\int_\Omega f(x,\omega)P^\nu(d\omega)\right)\ (a.s.)$$

が得られる.

このような枠組みの中で，未知係数をサンプリングによって推定しながら真の最適解の見つける問題を以下に述べる[9]．次の線形計画問題(22)を考える.

(22) $$\left|\begin{array}{ll}\text{最小化} & c'x,\\ \text{条件} & Aa+b\leq 0.\end{array}\right.$$

ここでAとbは未知あるいは確実に知られていない係数であるとする．Aとbの真の値をそれぞれA^*とb^*とするならば，Aとbは密度関数

(23) $$p(t)=\begin{cases}+\infty & t=(A^*,b^*)\\ 0 & t\neq(A^*,b^*)\end{cases}$$

をもつ一点分布にしたがう．$\{x|A^*x+b^*\leq 0\}\neq\phi$と仮定するとき，サンプル情報が得られる度に adaptive な方法で

(24) $$\left|\begin{array}{ll}\text{最小化} & c'x,\\ \text{条件} & E\{Ax+b\}\leq 0\Leftrightarrow A^*x+b^*\leq 0\end{array}\right.$$

の最適解を推定するのが目的である.

Aとbの推定は多変量回帰分析によって行なう．サンプル点x^kにおいて正規分布にしたがう観測雑音u^kを含んだ観測値$y^k=Ax^k+b+u^k$が得られたとき，Aとbの最尤推定量は$(\hat{A},\hat{b})'=(X'X)^{-1}(X'Y)$となる．ただし，$X$はサンプル点の行列($\operatorname{rank}X=n+1$)，$Y$は観測値の行列である．大きさ$\nu$のサンプルから求められる推定量$(\hat{A},\hat{b})$の密度関数$p^\nu(t)$は多変量正規分布を表わし，サンプル数が十分大きくなったときその分散共分散行列がOに近づくならば，$\lim_{\nu\to\infty}p^\nu(t)=p(t)$となり，$(\hat{A},\hat{b})$は$(A^*,b^*)$の一致推定量となる．大きさ$\nu$のサンプルが得られているときの推定値を$(A^\nu,b^\nu)$と表わすと，問題(24)の最適解$x^*$を推定するための問題は次のように表わされる.

(25) $$\left|\begin{array}{ll}\text{最小化} & c'x,\\ \text{条件} & E^\nu\{\hat{A}x+\hat{b}\}\leq 0\Leftrightarrow A^\nu x+b^\nu\leq 0\end{array}\right.$$

問題(25)の最適解x^νを求め，さらに新しいサンプルから更新される推定量が得られる度に問題(25)を解き，最適解を$x^{\nu+1},x^{\nu+2},\cdots$と求めてゆけば，問題(24)の最適解$x^*$に収束する点列$\{x^\nu\}$が得られる．しかし，$x^*$を求めるのが目的であるから，問題(25)の最適解を厳密に求めなくても，x^*へ収束する点列$\{\bar{x}^\nu\}$が構成できれば十分である．さらに，$\bar{x}^\nu$が問題(25)に対して実行可能でx^νに近ければより好ましい.

$\bar{x}^\nu$は問題(25)の実行可能解であるとする．$\bar{x}^\nu$をサンプル点として新たな観測値$y^{\nu+1}$が得られると，推定値は$(A^{\nu+1},b^{\nu+1})$と更新され，

(26) $$\left|\begin{array}{ll}\text{最小化} & c'x,\\ \text{条件} & E^{\nu+1}\{\hat{A}x+\hat{b}\}\leq 0\Leftrightarrow A^{\nu+1}x+b^{\nu+1}\leq 0\end{array}\right.$$

が得られる．問題(26)に対応して，$\bar{x}^\nu$から$\bar{x}^{\nu+1}$を構成するために次の2種類のステップを考える.

1. 実行可能性のためのステップ

$\bar{x}^\nu$は問題(26)に対して実行可能になるとは限らないので，その時には$\bar{x}^\nu$を実行可能領域内の点x_F^νへ射影する.

2. 最適性のためのステップ

実行可能な点x_F^νから$\bar{x}^{\nu+1}$を作るとき，より最適解に近づくようにしたいので内点法のアルゴリズム（アフ

ィンスケーリング法）を用いて

(27) $\bar{x}^{\nu+1}=x_F^{\nu}+\alpha^{\nu}d^{\nu}$

とする．ただし，α^{ν} はステップ幅，d^{ν} は方向ベクトルである．

アフィンスケーリング法では境界上にない内点が必要であるが，1.のステップで得られる点 x_F^{ν} は境界上に位置してしまう．問題(26)では期待値を取った関数が制約条件に使われているが，推定量 $(\hat{A},\hat{b})$ は正規分布にしたがう確率変数であるため，$\bar{x}^{\nu}$ が問題(26)に実行可能でなくても $\hat{A}\bar{x}^{\nu}+\hat{b}\leq 0$ となる確率は0ではない．このことから，点 $\bar{x}^{\nu}$ が $\tilde{A}\bar{x}^{\nu}+\tilde{b}\leq 0$ を満たすように確率変数 $(\hat{A},\hat{b})$ の実現値 $(\tilde{A},\tilde{b})$ を選ぶことは可能であり，この実現値を用いたときの実行可能領域が問題(26)の実行可能領域とできるだけ近くなるように $(\tilde{A},\tilde{b})$ の値を選ぶようにする．簡単のため $\tilde{A}=A^{\nu+1}$ と固定して，$\tilde{b}$ のみを選ぶ場合を考えると，次の b に関する最小化問題によって決められる．

$$(28)\quad \left|\begin{array}{ll}\text{最小化} & (b-b^{\nu})''V[\hat{b}]^{-1}(b-b^{\nu}),\\ \text{条件} & A^{\nu+1}x^{\nu}+b\leq 0.\end{array}\right.$$

これは b の信頼領域の中に $A^{\nu+1}\bar{x}^{\nu}+\tilde{b}\leq 0$ となる $\tilde{b}$ が含まれるようにするとき，その信頼領域の幅が最も小さくなるように $\tilde{b}^*$ を選んでいることを示している．このとき，問題(26)の実行可能領域を以下のように修正する．

$$(29)\quad \begin{aligned}&\{x\,|\,A_i^{\nu+1}x+\tilde{b}_i^*\leq 0\quad \text{for } i\in J,\\ &A_i^{\nu+1}x+b_i^{\nu+1}-\epsilon\leq 0\quad \text{for } i\in J',\\ &A_i^{\nu+1}x+b_i^{\nu+1}\leq 0\quad \text{for } i\notin J\cup J'\}\end{aligned}$$

ただし，$J=\{i\,|\,A_i^{\nu+1}x^{\nu}+b_i^{\nu+1}>0,\tilde{b}_i^*\neq b_i^{\nu+1}\}$,

$J'=\{i\,|\,A_i^{\nu+1}x^{\nu}+b_i^{\nu+1}\leq 0, A_i^{\nu+1}x_F^{\nu}+b_i^{\nu+1}=0\}$ で，ϵ は許容誤差とする．この修正された領域は，問題(26)の実行可能領域に射影された点 x_F^{ν} を実行可能な内点にするだけでなく，問題(26)の条件を推定量にもとづいて緩めていることになる．

図1に解法アルゴリズムを示している．収束性に関しては次の定理が成立する．

定理2 ［9］確率1で

$$(30)\quad \lim_{\nu\to\infty}\{x\,|\,E^{\nu}\{Ax+b\}\leq 0\}=\{x\,|\,E\{Ax+b\}\leq 0\}$$

が成立する．

定理3 ［9］確率1で

$$(31)\quad \lim_{\nu\to\infty}|c'\bar{x}^{\nu}-c'x^*|\to 0$$

となる．

---

procedure 最適解の推定：
begin
　初期推定値 (A^{ν},b^{ν}) と許容誤差 ϵ を与える；
　x^{ν} を任意に与える；
　repeat
　　begin
　　$\nu\leftarrow\nu+1$ とする；
　　新しい観測値 (x^{ν},y^{ν}) を得て，推定値を更新する；
　　if x^{ν} が実行可能でない
　　　then ステップ1により x_F^{ν} と領域(29)を求める
　　　else $x_F^{\nu}\leftarrow\bar{x}^{\nu}$ とする；
　　ステップ2により $\bar{x}^{\nu+1}\leftarrow x_F^{\nu}+\alpha^{\nu}d^{\nu}$ とする；
　　end
　until $\|\alpha^k d^k\|<\nu$；
　$\bar{x}^{\nu+1}$ が x^* の推定値となる；
end.

図1 解法アルゴリズム

4. おわりに

本稿では2つの話題を中心に確率的計画法における統計的手法について述べた．後者のモデルでは，推定のための問題(26)において機会制約条件を取り入れた問題を考えることによって，最適解の推定値の精度などの評価についても考察しているところである．また，このモデルは確率測度 P の近似を用いた近似解法［1］とも密接に関連しており，より広範囲の問題への拡張も可能であると思われる．

参考文献

［1］ J. Birge and R. J-B Wets, Designing Approximation Schemes for Stochastic Optimization Problems, In Particular for Stochastic Programs with Recourse, *Mathematical Programming Study*, Vol. 27, (1986), 54-102.

［2］ J. Bracken and R. M. Soland, Statistical Decision Analysis of Stochastic Linear Programming Problems, *Naval Research Logistics Quarterly*, Vol. 13, (1966), 205-226.

［3］ T. Cipra, Prediction in Stochastic Linear Programming, *Kybernetika*, Vol. 23, (1987), 214-226.

[4] J. Dupačová, The Minimax Approach to Stochastic Programming and an Illustrative Application, *Stochastics*, Vol.20, (1987), 73-88.

[5] J. Dupačová, Epi-consistency in Restricted Regression Models —The Case of a General Convex Fitting Function, (to appear)

[6] J. Dupačová and R. J-B Wets, Asymptotic Behavior of Statistical Estimators and of Optimal Solutions for Stochastic Optimization Problems, *Annals of Statistics*, Vol.16, (1988), 1517-1549.

[7] Yu. Ermoliev and R. J-B Wets, *Numerical Techniques for Stochastic Optimization*, Springer-Verlag, 1988.

[8] R. Jagannathan, Use of Sample Information in Stochastic Recourse and Chance-Constrained Programming Models, *Management Science*, Vol.31, (1985), 96-108.

[9] H. Morita and H. Ishii, A Stochastic Improvement Method for Stochastic Programming, (to appear).

[10] H. Morita, H. Ishii and T. Nishida, Confidence Region Method for Stochastic Programming Problem, *Journal of Operations Research Society of Japan*, Vol.30, (1987), 218-230.

[11] H. Morita, H. Ishii and T. Nishida, Stochastic Programming with Estimated Objective, *Technology Reports of the Osaka University*, Vol.39, (1989), 1-7.

[12] H. Morita, H. Ishii and T. Nishida, Stochastic Linear Programming Problem with Partially Estimated Constraint, *Mathematica Japonica*, Vol.35, (1990), 551-559.

脚注 1) エピ収束：$g=\text{epi-lim}\, g^{\nu}$ (関数列 $\{g^{\nu}\}$ が g にエピ収束する)⇔

- x に収束するすべての点列 $\{x^{\nu}\}$ に対して

  (A.1) $\liminf_{\nu\to\infty} g^{\nu}(x^{\nu})\geq g(x)$

- x に収束するある点列 $\{x^{\nu}\}$ に対して

  (A.2) $\limsup_{\nu\to\infty} g^{\nu}(x^{\nu})\leq g(x)$