数理解析研究所講究録
第 1584 巻 2008 年 21-24

一般化チェビシェフ不等式とその最適化への応用

北原　知就，水野　眞治，中田　和秀

東京工業大学　社会理工学研究科　経営工学専攻
東京都目黒区大岡山 2-12-1
連絡先：kitahara.t.ab@m.titech.ac.jp

概要

一般化チェビシェフ不等式 Vandenberghe et al. (2007) とは，平均ベクトルと分散共分散行列が既知な確率変数がある 2 次不等式によって記述される領域に入る確率の下限を与えるものである．本研究では一般化チェビシェフ不等式の最適化への 2 つの応用を提案する． 1 つ目の 2 次ミニマックス判別問題は平均，共分散がわかっている 2 つのクラスを判別するための 2 次関数を求める問題である．このとき，各クラスに様々な分布を考えたときの最悪の場合の判別率が最良なものを求める．2 つ目は，平均，共分散が与えられたとき，(1)確率変数の分布形によらずそこに入る確率がある要求水準以上である，(2)体積がなるべく小さい，という 2 つの条件を満たす楕円を決定する問題(確率制約付最小楕円決定問題)である．本研究によって，各問題に対して興味深い理論的性質，数値実験の結果が示される．

1. はじめに

世の中の多くの事象は確率変数によって記述される．しかし，その確率変数についての完全な情報を得ることは不可能であり，我々が知りうるのは，その平均や分散といったごく一部の情報である．それにも関わらず，その変数によって制御される事象が起こる確率を見積もらなくてはならないことは多々ある．

1 次元の確率変数 X があり，その平均を μ，分散を σ^2 とする．今，k を 1 より大きいスカラーとする．このとき，次の関係が成り立つことが知られている．

$$\Pr\{|X-\mu|<k\sigma\}\geq 1-\frac{1}{k^2} \quad (1)$$

(1)はチェビシェフの不等式と呼ばれる．ここで，

$$\Pr(X)=\begin{cases}1/2k^2 & (X=\mu-k\sigma)\\ 1-1/k^2 & (X=\mu)\\ 1/2k^2 & (X=\mu+k\sigma)\end{cases}$$

という確率変数を考えると，その平均は μ，分散は σ^2 であり，かつ

$$\Pr\{|X-\mu|<k\sigma\}=1-\frac{1}{k^2}$$

となる．このように，実際に下限を達成する確率変数が存在するという意味で(1)は厳密な不等式である．

最近は，1 次元のチェビシェフ不等式を多次元に拡張した一般化チェビシェフ不等式が Vandenberghe et al. (2007) などで盛んに研究されている．一般化チェビ

シェフ不等式については 2 節で詳説するが,その重要な点として,次の 2 点が挙げられる.

1. (1)に見られるように,1 次元のチェビシェフ不等式において確率を見積もる領域が平均を含む区間であるのに対し,一般化チェビシェフ不等式ではより広い領域を許す.
2. (1)のように確率の下限を陽に与えることはできないが,関連する半正定値計画問題を解くことで確率の下限とそれを達成する分布を構成することができる.ここで半正定値計画問題とは,近年の内点法アルゴリズムの発展により効率的に解くことができるようになった重要な数理計画問題である.

このような最近の研究の動向を踏まえ,著者らは一般化チェビシェフ不等式の 2 つの応用を研究した.一つは 2 次ミニマックス判別問題である.これは,平均ベクトルと分散共分散が既知な 2 つのクラスが与えられたとき,様々な分布を考えたときの最悪の場合の判別率が最良である 2 次判別関数を決定する問題である.この問題はパラメータ付半正定値計画問題として定式化されるが,著者らは最適な2次判別関数の中に線形のものが含まれるという興味深い性質を明らかにした.

二つ目の応用は確率制約付最小楕円決定問題である.これは確率変数の平均ベクトルと分散共分散行列が与えられたときに,そこに入る確率が前もって決められた要求水準を満たし,かつ体積が最小である楕円を決める問題である.ロバスト最適化等との関連から,このような問題を考えることは重要である.この問題については,数値実験の結果から,最適な楕円の形状が分散共分散行列から決定されるという興味深い現象が観察された.

本稿ではS^nはn次元実対称行列の集合を表し,$\Re^n$はn次元実ベクトルの集合を表す.

2. 一般化チェビシェフ不等式

n次元の確率変数Xがあり,その平均ベクトル,分散共分散行列(μ,Σ)が既知であるとする.本稿では以下このことを$X\sim(\mu,\Sigma)$と表記する.いま,有限個の 2 次不等式制約によって定まる次の領域

$$T=\{x\in\Re^n \mid x^TA_ix+2b_i^Tx+c_i<0,\quad \forall i=1,\dots,m\} \quad (2)$$

が与えられているとする.ここで,$(A_i,b_i,c_i)\in S^n\times\Re^n\times\Re\,(i=1,\dots m)$であり,各$A_i$は半正定値であるとは限らない.Vandenberghe et al. (2007) は$X\sim(\mu,\Sigma)$なる確率変数がTに入る確率の下限,すなわち$\inf_{X\sim(\mu,\Sigma)}\Pr\{X\in T\}$が次の半正定値計画問題の最適値であることを示した:

$$\begin{aligned}
&\min\ 1-\sum_{i=1}^m\lambda_i\\
&\text{s.t}\ \begin{bmatrix} Z_i & z_i\\ z_i^T & \lambda_i\end{bmatrix}\succeq O\ (i=1,\dots,m),\\
&\qquad \sum_{i=1}^m\begin{bmatrix} Z_i & z_i\\ z_i^T & \lambda_i\end{bmatrix}\preceq\begin{bmatrix}\Sigma+\mu\mu^T & \mu\\ \mu^T & 1\end{bmatrix},\\
&\qquad \mathrm{tr}(A_iZ_i)+2b_i^Tz_i+c_i\lambda_i\geq 0,\ i=1,\dots,m.
\end{aligned} \quad (3)$$

ここで，$Z_i, z_i, \lambda_i\ (i=1,\ldots,m)$ が変数である．確率の下限を達成する分布は(3)の最適解から計算することができる．

3. 2次ミニマックス判別問題

平均ベクトル，分散共分散行列が $(\mu_i, \Sigma_i)\ (i=1,2)$ である2群があるとする．以下では，各分散共分散行列は正定値であるとする．サンプル x が与えられたとき，2次判別関数 $q(z)=z^T Az+2b^T z+c$ を用いて判別する（ここで，$(A,b,c)\in S^n\times\Re^n\times\Re$，$A$ は凸でなくても良い）．すなわち，$q(x)<0$ ならば第1群に，$q(x)>0$ ならば第2群に判別し，$q(x)=0$ ならばどちらに判別しても良いとする．判別関数が与えられたとき，第1群のサンプルが第2群に誤って判別される確率は第1群の平均ベクトル，分散共分散行列だけではなく，第1群の分布形にも依存する．そこで，平均ベクトルが μ_1，分散共分散行列が Σ_1 であるような全ての分布を考えたときの誤判別率の上限を考え，それを

$$\alpha_{l12}=\sup_{x\sim(\mu_1,\Sigma_1)}\Pr\{q(x)\geq 0\}$$

と表す．同様に，第2群のサンプルを第1群に誤って判別する確率の上限 α_{l21} は，

$$\alpha_{l21}=\sup_{x\sim(\mu_2,\Sigma_2)}\Pr\{q(x)\leq 0\}$$

と表される．そして，判別関数 $q(z)$ の最悪の場合の誤判別率 α_l を

$$\alpha_l=\max\{\alpha_{l12},\alpha_{l21}\} \qquad (1)$$

と定義する．2次ミニマックス判別問題は，最悪の場合の誤判別率を最小にする2次判別関数を求める問題である．ここに，判別関数が線形である線形ミニマックス判別問題は Lanckriet et al. (2002)で研究されていることを注記しておく．詳細は割愛するが，著者らは Kitahara et al. (2007)で(2)の双対問題を利用して2次ミニマックス判別問題が次のパラメータ付半正定値計画問題に帰着できることを明らかにした：

$$\begin{aligned}
&\min \alpha\\
&\text{s.t}\quad \mathrm{tr}((\Sigma_i+\mu_i\mu_i^T)P_i)+2\mu_i^T q_i+r_i\leq\tau_i\alpha\ (i=1,2)\\
&\qquad \begin{bmatrix} P_i & q_i\\ q_i^T & r_i\end{bmatrix}\succeq O\ (i=1,2),\\
&\qquad \begin{bmatrix} P_1-A & q_1-b\\ (q_1-b)^T & r_1-\tau_1-c\end{bmatrix}\succeq O, \qquad\qquad (4)\\
&\qquad \begin{bmatrix} P_2+A & q_2+b\\ (q_2+b)^T & r_2-\tau_2+c\end{bmatrix}\succeq O,\\
&\qquad \tau_i>0\ (i=1,2),
\end{aligned}$$

2次ミニマックス判別問題に関して，著者らは次の興味深い性質が成立することを示した．

定理1. (Kitahara et al. (2007))線形ミニマックス判別問題の最適解を $l^*(z)=(a^*)^T z+b^*$ とする．このとき，$l^*(z)$ は2次ミニマックス判別問題の最適解である．

定理1が述べるように実際に(4)を解くことなく2次ミニマックス判別問題の最適解が得られるというのは興味深い．また，2次ミニマックス判別問題を考えることにより線形関数より良い最悪の場合の誤判別率を持つ判別関数を得られる可能性があったが，実際にはそのようなことはないというのも注目に値する．

4. 確率制約付最小楕円決定問題

$X \sim (\mu, \Sigma)$ なる n 次元確率変数があるとする．また，α を 0 より大きく 1 より小さい実数とする．このとき，次の 2 条件を満たす楕円を決定することを考える：

1. 確率変数がその楕円内に入る確率の下限は α 以上である．
2. 楕円の体積はなるべく小さい．

この問題はロバスト最適化と呼ばれる不確実性下での意思決定問題と密接な関連がある．

ここでは，簡単のため楕円の中心は確率変数の平均とする．そのような楕円はある n 次元正定値対称行列 P を用いて

$$E = \{x \in \Re^n : \| P^{1/2}(x-\mu) \| < 1\}$$

と表される．すると，確率制約付最小楕円決定問題は次の問題に帰着される：

$$\begin{aligned}
&\max\ \det(P) \\
&\text{s.t}\ \begin{bmatrix} Q & r \\ r^T & s \end{bmatrix} \succeq O, \\
&\quad \begin{bmatrix} Q-P & r+P\mu \\ (r+P\mu)^T & s-\tau-\mu^T P\mu+1 \end{bmatrix} \succeq O, \quad (5) \\
&\quad \tau - \mathrm{tr}((\Sigma+\mu\mu^T)Q) - 2r^T\mu - s \geq \tau\alpha, \\
&\quad P \succeq O, \tau \geq 0.
\end{aligned}$$

この問題における変数は $(P, Q, r, s, \tau) \in S^n \times S^n \times \Re^n \times \Re \times \Re$ である．この問題は Maxdet 問題と呼ばれ，効率的に解けることが知られている．

予備的な数値実験を行った結果，最適な楕円の形状は共分散の逆行列の定数倍によって定まる現象が観察された．この現象の一般性の検証は，今後の課題である．

5. おわりに

本研究では最近盛んに研究されている一般化チェビシェフ不等式の最適化への応用を 2 つ提案した．一つは 2 次ミニマックス判別問題であり，もう一つは確率制約付最小楕円決定問題である．各問題について興味深い理論的性質，数値実験の結果を紹介した．

参考文献

[1] Kitahara, T., Mizuno, S., Nakata, K., 2007. Generalizations of the Linear Minimax Classification Problem. Technical report.

[2] Lanckriet, G.R.G., El Ghaoui, L., Bhattacharyya C., and Jordan, M.I., 2002. A Robust Minimax Approach to Classification. Journal of Machine Learning Research, 3: 555--582.

[3]Vandenberghe, L., Boyd, S. and Comanor, K., 2007.Generalized Chebyshev bounds via semidefinite programming. SIAM Revies, 49: 52—64.