运筹学与控制论

DOI:10.11721/cqnuj20230105

求解随机互补问题的可微罚方法*

黄章乙，赵 勇，傅 璐

(重庆交通大学 数学与统计学院，重庆 400074)

摘要:【目的】研究随机互补问题的期望值模型。【方法】借鉴罚函数法和样本均值逼近方法求解随机互补问题的期望值模型。【结果】在适当假设下，证明了样本均值逼近问题最优解的收敛性和收敛率。【结论】所得结果为研究随机互补问题提供了新的思路。

关键词:罚函数法；随机互补问题；样本均值逼近方法；收敛性

中图分类号:O221.5　　**文献标志码**:A　　**文章编号**:1672-6693(2023)01-0133-06

考虑如下互补问题：求解 $\boldsymbol{x}\in\mathbf{R}^n$，使得：

$$\boldsymbol{x}\leqslant 0, F(\boldsymbol{x})\leqslant 0, \boldsymbol{x}^{\mathrm{T}}F(\boldsymbol{x})=0, \tag{1}$$

其中 $F:\mathbf{R}^n\to\mathbf{R}^n$ 为向量值函数。互补问题因被应用于数学规划、控制理论、运筹学和博弈论等多个领域而受到人们广泛关注[1-7]。由于现实生活中的很多实际问题涉及随机因素，如果忽略这些随机因素的存在将会导致决策失误。因此，从理论和实际应用方面而言，有必要研究随机互补问题。

考虑随机互补问题如下：求解 $\boldsymbol{x}^*\in\mathbf{R}^n$，使得：

$$\boldsymbol{x}^*\leqslant 0, F(\boldsymbol{x}^*,\boldsymbol{\omega})\leqslant 0, (\boldsymbol{x}^*)^{\mathrm{T}}F(\boldsymbol{x}^*,\boldsymbol{\omega})=0, \boldsymbol{\omega}\in\Omega, \text{a. s.}。 \tag{2}$$

其中：$\boldsymbol{\omega}$ 是概率空间 (Ω,F,P) 上的随机向量，$F:\mathbf{R}^n\times\Omega\to\mathbf{R}^n$ 为向量值函数，a. s. 是 almost surely 的缩写，表示几乎确定。由于随机因素的存在，一般情况下很难找到一个公共解 $\boldsymbol{x}^*$，使得问题(2)对几乎所有的 $\boldsymbol{\omega}\in\Omega$ 成立。于是，传统的分析和计算互补问题的方法很大程度上无法直接求解随机互补问题。因此，研究随机互补问题的首要任务是建立对应的恰当确定性模型，进而得到某种合理的解决方案。目前，学者们已研究了随机互补问题的多种确定性模型[8-21]，例如：CVaR 约束随机优化模型[8]、期望值(EV)模型[9]、期望残差极小化模型(ERM)[10-13]、随机均衡约束模型(SMPEC)[14-16]、随机规划模型[17]、鲁棒优化模型[18-19]和基于风险度量的 CVaR 优化模型[20]等。

本文考虑随机互补问题(2)的期望值模型，即求解 $\boldsymbol{x}^*\in\mathbf{R}^n$，使得：

$$\boldsymbol{x}^*\leqslant 0, \mathrm{E}[F(\boldsymbol{x}^*,\boldsymbol{\omega})]\leqslant 0, (\boldsymbol{x}^*)^{\mathrm{T}}\mathrm{E}[F(\boldsymbol{x}^*,\boldsymbol{\omega})]=0, \tag{3}$$

其中 E 表示关于 $\boldsymbol{\omega}$ 的概率分布的数学期望。受到文献[7]的启发，利用罚函数法求解随机互补问题(3)，即通过可微 $l_{1/p}$-罚函数方法把问题(3)转化为如下非线性方程：

$$\varphi(\boldsymbol{x},\rho):=\begin{bmatrix}\rho\boldsymbol{x}_1\mathrm{E}[F_1(\boldsymbol{x},\boldsymbol{\omega})]+[\mathrm{E}[F_1(\boldsymbol{x},\boldsymbol{\omega})]]_+^q\\ \rho x_2\mathrm{E}[F_2(\boldsymbol{x},\boldsymbol{\omega})]+[\mathrm{E}[F_2(\boldsymbol{x},\boldsymbol{\omega})]]_+^q\\ \vdots\\ \rho\boldsymbol{x}_n\mathrm{E}[F_n(\boldsymbol{x},\boldsymbol{\omega})]+[\mathrm{E}[F_n(\boldsymbol{x},\boldsymbol{\omega})]]_+^q\end{bmatrix}=0, \boldsymbol{x}\in C, \tag{4}$$

其中：$\rho>0$ 为惩罚参数，且 $q=1+\dfrac{1}{p}(p>1)$，$C:=\{\boldsymbol{x}\in\mathbf{R}^n\mid\boldsymbol{x}\leqslant 0\}$。若函数 $\mathrm{E}[F_i(x,\omega)]$是连续可微的，则复合函数$[\mathrm{E}[F_i(\boldsymbol{x},\boldsymbol{\omega})]]_+^q$ 在 $q>1$ 时也是一阶连续可微的。为了求解方程组(4)，考虑如下最小二乘问题：

$$\min_{x\in C}\Phi(\boldsymbol{x},\rho)=\|\varphi(\boldsymbol{x},\rho)\|^2。 \tag{5}$$

在本文中，记 $\mathrm{D}(A,B)$为集合 A 到集合 B 的偏差。

---

* **收稿日期**:2022-07-30　**修回日期**:2022-10-26　**网络出版时间**:2023-02-22 16:06

资助项目:国家自然科学基金(No. 12001072)

第一作者简介:黄章乙，男，研究方向为随机优化，E-mail：huangzhangyi1997@163.com

网络出版地址:https://kns.cnki.net/kcms/detail/50.1165.N.20230222.1346.020.html

1 解的存在性

本节将在一定假设下讨论优化问题(5)解的存在性。首先介绍将用到的一些定义和假设。

定义 1[3] 如果存在正常数 μ 使得对任意 $\boldsymbol{x},\boldsymbol{y}\in\mathbf{R}^n$,有:$\max\limits_{1\leqslant i\leqslant n}(\boldsymbol{x}_i-\boldsymbol{y}_i)(f_i(\boldsymbol{x})-f_i(\boldsymbol{y}))\geqslant\mu\|\boldsymbol{x}-\boldsymbol{y}\|^2$,则称 $f:\mathbf{R}^n\to\mathbf{R}^n$ 是一致 P-函数。

假设 1 (A1)$\mathrm{E}[F(\boldsymbol{x},\boldsymbol{\omega})]$在 $\mathbf{R}^n$ 上为一致 P-函数;(A2)对任意的 $\boldsymbol{\omega}\in\Omega$,函数 $F(\boldsymbol{x},\boldsymbol{\omega})$是连续可微的,且对任何固定的 $\boldsymbol{x}\in C$,$\mathrm{E}[F(\boldsymbol{x},\boldsymbol{\omega})]$有限;(A3)存在可积函数 $\kappa_1(\omega):\Omega\to\mathbf{R}_+$,使得对任意的 $\boldsymbol{x},\boldsymbol{y}\in C$ 和任意的 $\boldsymbol{\omega}\in\Omega$ 有:$\|F(\boldsymbol{x},\boldsymbol{\omega})-F(\boldsymbol{y},\boldsymbol{\omega})\|\leqslant\kappa_1(\boldsymbol{\omega})\|\boldsymbol{x}-\boldsymbol{y}\|$。

定理 1 若假设(A1)~(A2)成立,则对任意的 $\rho>0$,问题(5)的水平集有界。

证明 假设对给定的 $\rho>0$,存在非负数 $\bar{c}$ 使得问题(5)的水平集无界,则存在序列$\{\boldsymbol{x}^k\}\subseteq C$ 满足$\lim\limits_{k\to\infty}\|\boldsymbol{x}^k\|=\infty$,使得:$\Phi(\boldsymbol{x}^k,\rho)\leqslant\bar{c}$。

定义指标集 $I\subseteq\{1,\cdots,n\}$为:$I:=\{i\mid\{x_i^k\}$无界$\}$。由于$\{\boldsymbol{x}^k\}$无界,则 $I\neq\varnothing$。定义序列$\{\boldsymbol{y}^k\}$:$y_i^k:=\begin{cases}0, & \text{如果 } i\in I\\ x_i^k, & \text{如果 } i\notin I\text{。}\end{cases}$ 由 $\boldsymbol{y}^k,\boldsymbol{x}^k$ 的定义以及假设(A1),则存在 $j\in I$ 和 $\mu>0$,使得:

$$\mu\sum_{i\in I}(x_i^k)^2=\mu\|\boldsymbol{x}-\boldsymbol{y}\|^2\leqslant(x_j^k-y_j^k)(\mathrm{E}[F_j(\boldsymbol{x}^k,\boldsymbol{\omega})]-\mathrm{E}[F_j(\boldsymbol{y}^k,\boldsymbol{\omega})])\leqslant$$
$$\max_{i\in I}(x_i^k-y_i^k)(\mathrm{E}[F_i(\boldsymbol{x}^k,\boldsymbol{\omega})]-\mathrm{E}[F_i(\boldsymbol{y}^k,\boldsymbol{\omega})])=x_l^k(\mathrm{E}[F_l(\boldsymbol{x}^k,\boldsymbol{\omega})]-\mathrm{E}[F_l(\boldsymbol{y}^k,\boldsymbol{\omega})])。\tag{6}$$

其中:$l\in I$ 是使得最后一个等式取到最大值的指标之一。由于 $l\in I$,且$\{\boldsymbol{x}^k\}\subseteq C$,不失一般性,假设当 $k\to\infty$时,

$$\{x_l^k\}\to-\infty。\tag{7}$$

对(6)式两边同时除以 x_l^k,则:

$$\mathrm{E}[F_l(\boldsymbol{x}^k,\boldsymbol{\omega})]-\mathrm{E}[F_l(\boldsymbol{y}^k,\boldsymbol{\omega})]\leqslant\mu x_l^k。\tag{8}$$

显然,当 $k\to\infty$时,(8)式右边趋近于负无穷。由假设(A2)和 $\boldsymbol{y}^k$ 的定义可得 $\mathrm{E}[F(\boldsymbol{y}^k,\boldsymbol{\omega})]$有界。于是,由(8)式可知:

$$\mathrm{E}[F_l(\boldsymbol{x}^k,\boldsymbol{\omega})]\to-\infty,k\to\infty。\tag{9}$$

因此,结合(7)及(9)式,有:

$$\bar{c}\geqslant\Phi(\boldsymbol{x}^k,\rho)=\sum_{i=1}^n\varphi_i^2(\boldsymbol{x}^k,\rho)\geqslant(\rho x_l^k\mathrm{E}[F_l(\boldsymbol{x}^k,\boldsymbol{\omega})]+[\mathrm{E}[F_l(\boldsymbol{x}^k,\boldsymbol{\omega})]]_+^q)^2\to\infty,k\to\infty。\tag{10}$$

这与假设矛盾,故结论得证。 证毕

注 1 若问题(5)的水平集非空有界,则问题(5)的最优解集非空有界。

引理 1[7] 若假设(A1)~(A2)成立,令 $\bar{\boldsymbol{x}}(\rho)$为问题(5)的全局最优解,$\boldsymbol{x}'$为问题(3)的解,则存在常数 $\lambda>0$,使得:$\|\bar{\boldsymbol{x}}(\rho)-\boldsymbol{x}'\|^2\leqslant\lambda\rho^p$。

引理 2[7] 若假设(A1)~(A3)成立,若 $\bar{\boldsymbol{x}}(\rho)$为问题(5)的局部最优解且有不等式成立:$\mathrm{E}[F(\bar{\boldsymbol{x}}(\rho),\boldsymbol{\omega})]\leqslant0$,则 $\bar{\boldsymbol{x}}(\rho)$是问题(3)的解。

2 近似问题最优解和稳定点的收敛性

由于问题(5)中目标函数涉及数学期望,在实际中可能随机变量的分布未知或者难以得到期望值函数的显式表达式,因此采用样本均值逼近方法处理数学期望。于是,选取 $\boldsymbol{\omega}$ 的独立同分布样本 $\boldsymbol{\omega}_1,\cdots,\boldsymbol{\omega}_N$,则得到问题(5)的如下近似问题:

$$\min_{x\in C}\Phi^N(\boldsymbol{x},\rho)=\|\varphi^N(\boldsymbol{x},\rho)\|^2,\tag{11}$$

其中:$\varphi^N(\boldsymbol{x},\rho):=\begin{pmatrix}\rho\boldsymbol{x}_1F_1^N(\boldsymbol{x})+[F_1^N(\boldsymbol{x})]_+^q\\ \rho\boldsymbol{x}_2F_2^N(\boldsymbol{x})+[F_2^N(\boldsymbol{x})]_+^q\\ \vdots\\ \rho\boldsymbol{x}_NF_n^N(\boldsymbol{x})+[F_n^N(\boldsymbol{x})]_+^q\end{pmatrix}$,$F_j^N(\boldsymbol{x})=\dfrac{1}{N}\sum\limits_{i=1}^NF_j(\boldsymbol{x},\boldsymbol{\omega}_i),j=1,\cdots,n$。

在本节中均假设在给定的 $\rho>0$ 下分析近似问题的收敛性。下面给出本节将用到的假设:

假设 2　(A4)当 N 足够大时，对任意的 $i=1,2,\cdots,n$，若 $\lim\limits_{\|\boldsymbol{x}\|\to\infty} \mathrm{E}[F_i(\boldsymbol{x},\boldsymbol{\omega})]=\infty(-\infty)$，则有：$\lim\limits_{\|\boldsymbol{x}\|\to\infty} F_i^N(\boldsymbol{x})=\infty(-\infty)$；(A5)对任意的 $\boldsymbol{\omega}\in\Omega$，$\nabla_x F(\boldsymbol{x},\boldsymbol{\omega})$ 关于 $\boldsymbol{x}$ 连续。

引理 3[22]　设 X 是 $\mathbf{R}^n$ 的一个非空紧子集，若：(i) 对于任意的 $\boldsymbol{x}\in X$，对几乎所有的 $\boldsymbol{\omega}\in\Omega$，函数 $F(\cdot,\boldsymbol{\omega})$ 连续；(ii) $F(\boldsymbol{x},\boldsymbol{\omega})$ 被一个可积函数控制；(iii) 样本 $\boldsymbol{\omega}_1,\cdots,\boldsymbol{\omega}_N$ 独立同分布，则 $\mathrm{E}[F(\boldsymbol{x},\boldsymbol{\omega})]$ 在 X 上是有限且连续的，且 $F^N(\boldsymbol{x})=\dfrac{1}{N}\sum\limits_{i=1}^{N}F(\boldsymbol{x},\boldsymbol{\omega}_i)$ 在 X 上依概率 1 一致收敛于 $\mathrm{E}[F(\boldsymbol{x},\boldsymbol{\omega})]$。

定理 2　若假设(A1)、(A2)和(A4)成立，则问题(11)的水平集有界。

证明　利用反证法证明。对给定的 $\rho>0$，假设存在 $\bar{c}\geqslant 0$，使得问题(11)的水平集 $L_{\Phi^N}(\bar{c})=\{\boldsymbol{x}\in C\,|\,\Phi(\boldsymbol{x},\rho)\leqslant\bar{c}\}$，无界。则存在序列 $\{\boldsymbol{x}^k\}\subset C$ 且 $\lim\limits_{k\to\infty}\|\boldsymbol{x}^k\|=\infty$，使得：

$$\Phi^N(\boldsymbol{x}^k,\rho)\leqslant\bar{c}。\tag{12}$$

于是，利用定理 1 类似的方法定义指标集 I，再结合假设(A1)，当 $k\to\infty$ 时，不失一般性，取 $\{x_l^k\}\to-\infty$，有：

$$\{\mathrm{E}[F_l(\boldsymbol{x}^k,\boldsymbol{\omega})]\}\to-\infty。\tag{13}$$

再根据假设(A4)可以得到：

$$\{F_l^N(\boldsymbol{x}^k)\}\to-\infty。\tag{14}$$

于是，结合(12)及(14)式知：

$$c\geqslant\Phi^N(\boldsymbol{x}^k,\rho)=\sum_{i=1}^{n}(\varphi_i^N(\boldsymbol{x}^k,\rho))^2\geqslant(\rho x_l^k F_l^N(\boldsymbol{x}^k)+[F_l^N(\boldsymbol{x}^k)]_+^q)^2\to\infty,k\to\infty。\tag{15}$$

这与假设矛盾，故问题(11)的水平集有界。　　证毕

注 2　若问题(11)的水平集非空有界，则问题(5)的最优解集非空有界。

引理 4　若假设(A2)～(A3)成立，$\varphi^N(\boldsymbol{x},\rho)$ 在任意紧集 $X\subset C$ 上依概率 1 一致收敛于 $\varphi(\boldsymbol{x},\rho)$。

证明　由于 $\rho>0$ 给定，则对任意的 $\boldsymbol{x}\in X$，存在 $M_1\geqslant 0$，使对任意的 $i=1,2,\cdots,n$，有 $\rho|x_i|\leqslant M_1$ 成立，再由假设(A2)～(A3)和引理 3，当 $N\to\infty$ 时，$F^N(\boldsymbol{x})$ 和 $(F^N(\boldsymbol{x}))^q$ 在 X 上分别依概率 1 一致收敛于 $\mathrm{E}[F(\boldsymbol{x},\boldsymbol{\omega})]$ 及 $\mathrm{E}[F^q(\boldsymbol{x},\boldsymbol{\omega})]$，即对任意的 $\varepsilon>0$，存在 N_0，使得当 $N>N_0$ 时，有：

$$\sup_{\boldsymbol{x}\in X}\max_{1\leqslant i\leqslant n}|F_i^N(\boldsymbol{x})-\mathrm{E}[F_i(\boldsymbol{x},\boldsymbol{\omega})]|\leqslant\frac{\varepsilon}{2M_1\sqrt{n}},\text{w. p. 1.},\tag{16}$$

$$\sup_{\boldsymbol{x}\in X}\max_{1\leqslant i\leqslant n}|(F_i^N(\boldsymbol{x}))^q-\mathrm{E}[F_i^q(\boldsymbol{x},\boldsymbol{\omega})]|\leqslant\frac{\varepsilon}{2\sqrt{n}},\text{w. p. 1.}。\tag{17}$$

这里 w. p. l. 表示 with probability 1，根据 $\varphi^N(\boldsymbol{x},\rho)$ 和 $\varphi(\boldsymbol{x},\rho)$ 的定义得：

$$\sup_{\boldsymbol{x}\in X}\|\varphi^N(\boldsymbol{x},\rho)-\varphi(\boldsymbol{x},\rho)\|=\sup_{\boldsymbol{x}\in X}\left\|\begin{matrix}\rho\boldsymbol{x}_1F_1^N(\boldsymbol{x})+[F_1^N(\boldsymbol{x})]_+^q-\rho\boldsymbol{x}_1\mathrm{E}[F_1(\boldsymbol{x},\boldsymbol{\omega})]-[\mathrm{E}[F_1(\boldsymbol{x},\boldsymbol{\omega})]]_+^q\\ \rho\boldsymbol{x}_2F_2^N(\boldsymbol{x})+[F_2^N(\boldsymbol{x})]_+^q-\rho\boldsymbol{x}_2\mathrm{E}[F_2(\boldsymbol{x},\boldsymbol{\omega})]-[\mathrm{E}[F_2(\boldsymbol{x},\boldsymbol{\omega})]]_+^q\\ \vdots\\ \rho\boldsymbol{x}_nF_n^N(\boldsymbol{x})+[F_n^N(\boldsymbol{x})]_+^q-\rho\boldsymbol{x}_n\mathrm{E}[F_n(\boldsymbol{x},\boldsymbol{\omega})]-[\mathrm{E}[F_n(\boldsymbol{x},\boldsymbol{\omega})]]_+^q\end{matrix}\right\|=$$

$$\sup_{\boldsymbol{x}\in X}\left\|\begin{matrix}\rho\boldsymbol{x}_1(F_1^N(\boldsymbol{x})-\mathrm{E}[F_1(\boldsymbol{x},\boldsymbol{\omega})])+([F_1^N(\boldsymbol{x})]_+^q-[\mathrm{E}[F_1(\boldsymbol{x},\boldsymbol{\omega})]]_+^q)\\ \rho\boldsymbol{x}_2(F_2^N(\boldsymbol{x})-\mathrm{E}[F_2(\boldsymbol{x},\boldsymbol{\omega})])+([F_2^N(\boldsymbol{x})]_+^q-[\mathrm{E}[F_2(\boldsymbol{x},\boldsymbol{\omega})]]_+^q)\\ \vdots\\ \rho\boldsymbol{x}_n(F_n^N(\boldsymbol{x})-\mathrm{E}[F_n(\boldsymbol{x},\boldsymbol{\omega})])+([F_n^N(\boldsymbol{x})]_+^q-[\mathrm{E}[F_n(\boldsymbol{x},\boldsymbol{\omega})]]_+^q)\end{matrix}\right\|\leqslant$$

$$\sup_{\boldsymbol{x}\in X}\sqrt{n}\left\|\begin{matrix}\rho\boldsymbol{x}_1(F_1^N(\boldsymbol{x})-\mathrm{E}[F_1(\boldsymbol{x},\boldsymbol{\omega})])+([F_1^N(\boldsymbol{x})]_+^q-[\mathrm{E}[F_1(\boldsymbol{x},\boldsymbol{\omega})]]_+^q)\\ \rho\boldsymbol{x}_2(F_2^N(\boldsymbol{x})-\mathrm{E}[F_2(\boldsymbol{x},\boldsymbol{\omega})])+([F_2^N(\boldsymbol{x})]_+^q-[\mathrm{E}[F_2(\boldsymbol{x},\boldsymbol{\omega})]]_+^q)\\ \vdots\\ \rho\boldsymbol{x}_n(F_n^N(\boldsymbol{x})-\mathrm{E}[F_n(\boldsymbol{x},\boldsymbol{\omega})])+([F_n^N(\boldsymbol{x})]_+^q-[\mathrm{E}[F_n(\boldsymbol{x},\boldsymbol{\omega})]]_+^q)\end{matrix}\right\|_\infty\leqslant$$

$$\sqrt{n}\sup_{\boldsymbol{x}\in X}\max_{1\leqslant i\leqslant n}(\rho\|\boldsymbol{x}_i\|\|F_i^N(\boldsymbol{x})-\mathrm{E}[F_i(\boldsymbol{x},\boldsymbol{\omega})]\|+\|[F_i^N(\boldsymbol{x})]_+^q-[\mathrm{E}[F_i(\boldsymbol{x},\boldsymbol{\omega})]]_+^q\|)\leqslant$$

$$\sqrt{n}\sup_{\boldsymbol{x}\in X}\max_{1\leqslant i\leqslant n}(\rho\|\boldsymbol{x}_i\|\|F_i^N(\boldsymbol{x})-\mathrm{E}[F_i(\boldsymbol{x},\boldsymbol{\omega})]\|+\|(F_i^N(\boldsymbol{x}))^q-\mathrm{E}[F_i^q(\boldsymbol{x},\boldsymbol{\omega})]\|)\leqslant$$

$$M_1\sqrt{n}\sup_{\boldsymbol{x}\in X}\max_{1\leqslant i\leqslant n}\|F_i^N(\boldsymbol{x})-\mathrm{E}[F_i(\boldsymbol{x},\boldsymbol{\omega})]\|+\sqrt{n}\sup_{\boldsymbol{x}\in X}\max_{1\leqslant i\leqslant n}\|(F_i^N(\boldsymbol{x}))^q-\mathrm{E}[F_i^q(\boldsymbol{x},\boldsymbol{\omega})]\|。\tag{18}$$

于是,由(16),(17)和(18)式得,当 $N>N_0$ 时,有:

$$\sup_{x\in X}\|\varphi^N(\boldsymbol{x},\rho)-\varphi(\boldsymbol{x},\rho)\|\leqslant M_1\sqrt{n}\frac{\varepsilon}{2M_1\sqrt{n}}+\sqrt{n}\frac{\varepsilon}{2\sqrt{n}}=\varepsilon,\text{w.p.1.}$$ 。

由 $\varepsilon>0$ 的任意性,$\varphi^N(\boldsymbol{x},\rho)$ 在紧集 X 上依概率 1 一致收敛于 $\varphi(\boldsymbol{x},\rho)$。 证毕

为了讨论样本均值逼近问题(11)的最优解集的收敛性,对于给定的 $\rho>0$,记问题(11)和问题(5)的最优解集分别为 $S^N(\rho)$ 和 $S^*(\rho)$,最优值分别为 $\theta^N(\rho)$ 和 $\theta^*(\rho)$。

定理 3 若假设(A1)~(A4)成立,当 $N\to\infty$ 时,问题(11)的最优值 $\theta^N(\rho)$ 依概率 1 收敛到问题(5)的最优值 $\theta^*(\rho)$,且 $\mathrm{D}(S^N(\rho),S^*(\rho))\to 0$。此外,记 $\{\boldsymbol{x}^N(\rho)\}$ 的聚点为 $\overline{\boldsymbol{x}}(\rho)$。如果下列条件满足:(i) 存在紧集 $K\subset C$,使得 $\{\boldsymbol{x}^N(\rho)\}\subset K$, w.p.1;(ii) 对任意的 $\boldsymbol{x}\in K$,函数 $\Phi(\boldsymbol{x},\rho)$ 在 $\overline{\boldsymbol{x}}(\rho)$ 处二阶增长条件成立;(iii) 对任意的 $\boldsymbol{x}\in K$,矩母函数:$\mathcal{M}_F(t):=\mathrm{E}\{\mathrm{e}^{t(F(\boldsymbol{x},\boldsymbol{\omega})-\mathrm{E}[F(\boldsymbol{x},\boldsymbol{\omega})])}\}$,$\mathcal{M}_{[F]_+^q}(t):=\mathrm{E}\{\mathrm{e}^{t[F(\boldsymbol{x},\boldsymbol{\omega})]_+^q-[\mathrm{E}[F(\boldsymbol{x},\boldsymbol{\omega})]]_+^q}\}$ 对任意在 0 附近的 t 是有限的;(iv)矩母函数 $\mathcal{M}_{\kappa_1}(t):=\mathrm{E}\{\mathrm{e}^{t(\kappa_1(\boldsymbol{\omega})-\mathrm{E}[\kappa_1(\boldsymbol{\omega})])}\}$ 和 $\mathcal{M}_{\kappa_2}(t):=\mathrm{E}\{\mathrm{e}^{t(\kappa_2(\boldsymbol{\omega})-\mathrm{E}[\kappa_1(\boldsymbol{\omega})])}\}$ 对任意在 0 附近的 t 是有限的,其中 $\kappa_2(\boldsymbol{\omega})$ 为 $[F(\boldsymbol{x},\boldsymbol{\omega})]_+^q$ 在 K 上的 Lipschitz 常数,则存在正常数 ν 和 β,使得对任意的 $\varepsilon>0$ 有:

$$P\{\|\boldsymbol{x}^N(\rho)-\boldsymbol{x}'\|^2\geqslant\varepsilon+\lambda\rho^p\}\leqslant\nu\mathrm{e}^{-N\beta},$$

其中:$\boldsymbol{x}'$ 为随机互补问题(3)的解,λ 为引理 1 中的正常数。

证明 当假设(A1)~(A4)成立时,则问题(11)的解集都是非空有界的,故存在包含问题(11)解的紧集 K。则由引理 4 得,当 $N\to\infty$ 时,$\varphi^N(\boldsymbol{x},\rho)$ 在紧集 K 上依概率 1 一致收敛于 $\varphi(\boldsymbol{x},\rho)$。于是,由文献[23]的定理 5.3 可得,对任意给定的 $\rho>0$,当 $N\to\infty$,问题(11)的最优值 $\theta^N(\rho)$ 依概率 1 收敛到问题(5)的最优值 $\theta^*(\rho)$,且 $\mathrm{D}(S^N(\rho),S^*(\rho))\to 0$。

下证存在正常数 ν 和 β,使得 $P\{\|\boldsymbol{x}^N(\rho)-\boldsymbol{x}'\|^2\geqslant\varepsilon+\lambda\rho^p\}\leqslant\nu\mathrm{e}^{-N\beta}$。根据假设 $\Phi(\boldsymbol{x},\rho)$ 在 $\overline{\boldsymbol{x}}(\rho)$ 处的二阶增长条件成立,故存在 $\alpha>0$ 有:

$$\Phi(\boldsymbol{x},\rho)\geqslant\Phi(\overline{\boldsymbol{x}}(\rho),\rho)+\alpha\|\boldsymbol{x}-\overline{\boldsymbol{x}}(\rho)\|^2,\ \forall\,\boldsymbol{x}\in K。\tag{19}$$

由于 $\boldsymbol{x}^N(\rho)\in K$,则:

$$\begin{aligned}\|\boldsymbol{x}^N(\rho)-\overline{\boldsymbol{x}}(\rho)\|^2&\leqslant\frac{1}{\alpha}(\Phi(\boldsymbol{x}^N(\rho),\rho)-\Phi(\overline{\boldsymbol{x}}(\rho),\rho))=\\ &\frac{1}{\alpha}(\Phi(\boldsymbol{x}^N(\rho),\rho)-\Phi^N(\boldsymbol{x}^N(\rho),\rho)+\Phi^N(\boldsymbol{x}^N(\rho),\rho)-\Phi(\overline{\boldsymbol{x}}(\rho),\rho))\leqslant\\ &\frac{1}{\alpha}(\Phi(\boldsymbol{x}^N(\rho),\rho)-\Phi^N(\boldsymbol{x}^N(\rho),\rho)+\Phi^N(\overline{\boldsymbol{x}}(\rho),\rho)-\Phi(\overline{\boldsymbol{x}}(\rho),\rho))\leqslant\\ &\frac{1}{\alpha}(|\Phi^N(\boldsymbol{x}^N(\rho),\rho)-\Phi(\boldsymbol{x}^N(\rho),\rho)|+|\Phi^N(\overline{\boldsymbol{x}}(\rho),\rho)-\Phi(\overline{\boldsymbol{x}}(\rho),\rho)|)\leqslant\\ &\frac{2}{\alpha}\sup_{x\in K}|\Phi^N(\boldsymbol{x},\rho)-\Phi(\boldsymbol{x},\rho)|\leqslant\frac{4M}{\alpha}\sup_{x\in K}\Big|\sum_{i=1}^n(\varphi_i^N(\boldsymbol{x},\rho)-\varphi_i(\boldsymbol{x},\rho))\Big|。\end{aligned}\tag{20}$$

其中:M 为使得对任意的 $i=1,2,\cdots,n$,$\sup_{x\in K}|\varphi_i^N(\boldsymbol{x},\rho)|\leqslant M$ 和 $\sup_{x\in K}|\varphi_i(\boldsymbol{x},\rho)|\leqslant M$ 的正常数。令 $|\varphi_l^N(\boldsymbol{x},\rho)-\varphi_l(\boldsymbol{x},\rho)|=\max_{1\leqslant i\leqslant n}|\varphi_i^N(\boldsymbol{x},\rho)-\varphi_i(\boldsymbol{x},\rho)|$。由于 K 是紧集,对给定的 $\rho>0$,存在 $M_1>0$(不妨假设 $M_1>1$)使得 $|\rho x_i|\leqslant M_1$。因此,由(20)式得到:

$$\begin{aligned}\|\boldsymbol{x}^N(\rho)-\overline{\boldsymbol{x}}(\rho)\|^2&\leqslant\frac{4M}{\alpha}\sup_{x\in K}\Big|\sum_{i=1}^n(\varphi_i^N(\boldsymbol{x},\rho)-\varphi_i(\boldsymbol{x},\rho))\Big|\leqslant\frac{4nM}{\alpha}\sup_{x\in K}|(\varphi_l^N(\boldsymbol{x},\rho)-\varphi_l(\boldsymbol{x},\rho))|=\\ &\frac{4nM}{\alpha}\sup_{x\in K}|\rho\boldsymbol{x}_lE[F_l(\boldsymbol{x},\boldsymbol{\omega})]+[\mathrm{E}[F_l(\boldsymbol{x},\boldsymbol{\omega})]]_+^q-\rho\boldsymbol{x}_lF_l^N(\boldsymbol{x})-[F_l^N(\boldsymbol{x})]_+^q|\leqslant\\ &\frac{4nMM_1}{\alpha}\sup_{x\in K}|\mathrm{E}[F_l(\boldsymbol{x},\boldsymbol{\omega})]-F_l^N(\boldsymbol{x})|+\frac{4nM}{\alpha}\sup_{x\in K}|[\mathrm{E}[F_l(\boldsymbol{x},\boldsymbol{\omega})]]_+^q-[F_l^N(\boldsymbol{x})]_+^q|\leqslant\\ &\frac{8nMM_1}{\alpha}\max\{\sup_{x\in K}|\mathrm{E}[F_l(\boldsymbol{x},\boldsymbol{\omega})]-F_l^N(\boldsymbol{x})|,\sup_{x\in K}|[\mathrm{E}[F_l(\boldsymbol{x},\boldsymbol{\omega})]]_+^q-[F_l^N(\boldsymbol{x})]_+^q|\}。\end{aligned}\tag{21}$$

下面对不等式(21)分两种情况讨论:

a) 若 $\sup_{x\in K}|\mathrm{E}[F_l(\boldsymbol{x},\boldsymbol{\omega})]-F_l^N(\boldsymbol{x})|\geqslant\sup_{x\in K}|[\mathrm{E}[F_l(\boldsymbol{x},\boldsymbol{\omega})]]_+^q-[F_l^N(\boldsymbol{x})]_+^q|$,根据不等式(21),对任意的 $\varepsilon>0$,

有：

$$P\{\|\boldsymbol{x}^N(\rho)-\overline{\boldsymbol{x}}(\rho)\|^2\geqslant\varepsilon\}\leqslant P\left\{\frac{8nMM_1}{\alpha}\sup_{x\in K}|\mathrm{E}[F_l(\boldsymbol{x},\boldsymbol{\omega})]-F_l^N(\boldsymbol{x})|\geqslant\varepsilon\right\}=$$

$$P\left\{\sup_{x\in K}|\mathrm{E}[F_l(\boldsymbol{x},\boldsymbol{\omega})]-F_l^N(\boldsymbol{x})|\geqslant\frac{\alpha\varepsilon}{8nMM_1}\right\}。\tag{22}$$

令 $\eta=\frac{\alpha\varepsilon}{8nMM_1}$，当假设(i)～(iv)成立时，根据文献[23]中的定理 5.1，存在正常数 ν_1 和 β_1，使得对任意大于 0 的 η_1 有：

$$P\{\sup_{x\in K}|\mathrm{E}[F_l(\boldsymbol{x},\boldsymbol{\omega})]-F_l^N(\boldsymbol{x})|\geqslant\eta_1\}\leqslant\nu_1\mathrm{e}^{-N\beta_1}。\tag{23}$$

b) 同理，若 $\sup_{x\in K}|[\mathrm{E}[F_l(\boldsymbol{x},\boldsymbol{\omega})]]_+^q-[F_l^N(\boldsymbol{x})]_+^q|\geqslant\sup_{x\in K}|\mathrm{E}[F_l(\boldsymbol{x},\boldsymbol{\omega})]-F_l^N(\boldsymbol{x})|$，根据文献[23]中的定理 5.1，存在正常数 ν_2 和 β_2，使得对任意大于 0 的 η_2 有：

$$P\{\sup_{x\in K}|[\mathrm{E}[F_l(\boldsymbol{x},\boldsymbol{\omega})]]_+^q-[F_l^N(\boldsymbol{x})]_+^q|\geqslant\eta_2\}\leqslant\nu_2\mathrm{e}^{-N\beta_2}。\tag{24}$$

结合以上两种情况，对任意的 $\varepsilon>0$，存在正常数 ν 和 β，使得：

$$P\{\|\boldsymbol{x}^N(\rho)-\overline{\boldsymbol{x}}(\rho)\|^2\geqslant\varepsilon\}\leqslant\nu\mathrm{e}^{-N\beta}。\tag{25}$$

由引理 1 及(25)式可得：

$$P\{\|\boldsymbol{x}^N(\rho)-\boldsymbol{x}'\|^2\geqslant\varepsilon+\lambda\rho^p\}\leqslant P\{\|\boldsymbol{x}^N(\rho)-\overline{\boldsymbol{x}}(\rho)\|^2+\|\overline{\boldsymbol{x}}(\rho)-\boldsymbol{x}'\|^2\geqslant\varepsilon+\lambda\rho^p\}\leqslant$$

$$P\{\|\boldsymbol{x}^N(\rho)-\overline{\boldsymbol{x}}(\rho)\|^2\geqslant\varepsilon\}\leqslant\nu\mathrm{e}^{-N\beta}。$$

故结论得证。　　证毕

下面讨论近似问题(11)的局部最优解的聚点为问题(3)的解。

定理 4　若假设(A1)～(A5)成立，令$\{\boldsymbol{x}^N(\rho)\}$为问题(11)的局部最优解序列，聚点为 $\overline{\boldsymbol{x}}(\rho)$，若 $F^N(\boldsymbol{x}^N(\rho))\leqslant 0$ 成立，则 $\overline{\boldsymbol{x}}(\rho)$依概率 1 为原问题(3)的解。

证明　不失一般性，由于$\{\boldsymbol{x}^N(\rho)\}$收敛到 $\overline{\boldsymbol{x}}(\rho)$，则 $\overline{\boldsymbol{x}}(\rho)\leqslant 0$。由于 $\boldsymbol{x}^N(\rho)$为问题(11)的局部最优解，则根据文献[7]的命题 3.1，有：

$$\nabla_x\Phi^N(\boldsymbol{x}^N(\rho),\rho)^{\mathrm{T}}(\boldsymbol{x}-\boldsymbol{x}^N(\rho))\geqslant 0,\ \forall\,\boldsymbol{x}\in C。\tag{26}$$

由假设(A1)～(A5)，不难得到：$\lim_{N\to\infty}\nabla_x\Phi^N(\boldsymbol{x}^N(\rho),\rho)=\nabla_x\Phi(\overline{\boldsymbol{x}}(\rho),\rho)$，w. p. 1.。然后对(26)式两边求极限得到：$\nabla_x\Phi(\overline{\boldsymbol{x}}(\rho),\rho)^{\mathrm{T}}(\boldsymbol{x}-\overline{\boldsymbol{x}}(\rho))\geqslant 0,\ \forall\,\boldsymbol{x}\in C$，w. p. 1.。由于 $F^N(\boldsymbol{x}^N(\rho))\leqslant 0$，因此由引理 3 有：$\lim_{N\to\infty}F^N(\boldsymbol{x}^N(\rho))=\mathrm{E}[F(\overline{\boldsymbol{x}}(\rho),\omega)]\leqslant 0$，w. p. 1.。根据引理 2 知，$\overline{\boldsymbol{x}}(\rho)$依概率 1 为问题(3)的解。　　证毕

3 结束语

本文首先借助可微罚函数的方法把随机互补问题转化为优化问题进行求解。然后，在一定条件下讨论所得优化问题解的存在性。最后，借助样本均值逼近方法处理所得优化问题中的数学期望，并讨论近似问题最优解的收敛性和收敛率。所得结果为随机互补问题的求解提供了一种新方法。

参考文献：

[1] FACCHINEI F, PANG J S. Finite-dimensional variational inequalities and complementarity problems[M]. New York: Springer, 2003.

[2] FERRIS M C, PANG J S. Engineering and economic applications of complementarity problems[J]. SIAM Review, 1997, 39(4): 669-713.

[3] HARKER P T, PANG J S. Finite-dimensional variational inequality and nonlinear complementarity problems: a survey of theory, algorithms and applications[J]. Mathematical Programming, 1990, 48(1): 161-220.

[4] PANG J S, STEWART D E. Differential variational inequalities[J]. Mathematical Programming, 2008, 113(2): 345-424.

[5] WANG S, YANG X Q, TEO K L. Power penalty method for a linear complementarity problem arising from American option valuation[J]. Journal of Optimization Theory and Applications, 2006, 129(2): 227-254.

[6] HUANG C, WANG S. A power penalty approach to a nonlinear complementarity problem[J]. Operations Research Letters, 2010, 38(1): 72-76.

[7] TIAN B, HU Y, YANG X. A box-constrained differentiable penalty method for nonlinear complementarity problems[J]. Journal

of Global Optimization,2015,62(4):729-747.

[8] XU L,YU B. CVaR-constrained stochastic programming reformulation for stochastic nonlinear complementarity problems[J]. Computational Optimization and Applications,2014,58(2):483-501.

[9] GÜRKAN G,YONCA Ö A,ROBINSON S M. Sample-path solution of stochastic variational inequalities[J]. Mathematical Programming,1999,84(2):313-333.

[10] CHEN X,FUKUSHIMA M. Expected residual minimization method for stochastic linear complementarity problems[J]. Mathematics of Operations Research,2005,30(4):1022-1038.

[11] CHEN X,WETS R J B,ZHANG Y. Stochastic variational inequalities: residual minimization smoothing sample average approximations[J]. SIAM Journal on Optimization,2012,22(2):649-673.

[12] HAMATANI K,FUKUSHIMA M. Pricing American options with uncertain volatility through stochastic linear complementarity models[J]. Computational Optimization and Applications,2011,50(2):263-286.

[13] LING C,QI L,ZHOU G,et al. The SC1 property of an expected residual function arising from stochastic complementarity problems[J]. Operations Research Letters,2008,36(4):456-460.

[14] LIN G H. Combined Monte Carlo sampling and penalty method for stochastic nonlinear complementarity problems[J]. Mathematics of Computation,2009,78(267):1671-1686.

[15] LIN G H,CHEN X,FUKUSHIMA M. Solving stochastic mathematical programs with equilibrium constraints via approximation and smoothing implicit programming with penalization[J]. Mathematical Programming,2009,116(1):343-368.

[16] MATARAMVURA S,OKSENDAL B. Risk minimizing portfolios and HJBI equations for stochastic differential games[J]. Stochastics An International Journal of Probability and Stochastic Processes,2008,80(4):317-337.

[17] WANG M,ALI M M. Stochastic nonlinear complementarity problems:stochastic programming reformulation and penalty-based approximation method[J]. Journal of Optimization Theory and Applications,2010,144(3):597-614.

[18] WU D,HAN J,ZHU J. Robust solutions to uncertain linear complementarity problems[J]. Acta Mathematicae Applicatae Sinica,English Series,2011,27(2):339-352.

[19] ZHOU J,XIU N,CHEN J S. Solution properties and error bounds for semi-infinite complementarity problems[J]. Journal of Industrial & Management Optimization,2013,9(1):99-115.

[20] LIN G,CHEN X. CVaR-based formulation and approximation method for stochastic variational inequalities[J]. Numerical Algebra Control & Optimization,2017,1(1):35-48.

[21] ZHANG C,CHEN X. Stochastic nonlinear complementarity problem and applications to traffic equilibrium under uncertainty [J]. Journal of Optimization Theory and Applications,2008,137(2):277-295.

[22] SHAPIRO A,DENTCHEVA D,RUSZCZYNSKI A. Lectures on stochastic programming: modeling and theory[M]. Philadelphia:Society for Industrial and Applied Mathematics,2014.

[23] SHAPIRO A,XU H. Stochastic mathematical programs with equilibrium constraints,modelling and sample average approximation[J]. Optimization,2008,57(3):395-418.

Operations Research and Cybernetics

A Differentiable Penalty Method for Solving Stochastic Complementarity

HUANG Zhangyi, ZHAO Yong, FU Lu

(College of Mathematics and Statistics, Chongqing Jiaotong University, Chongqing 400074, China)

Abstract: [**Purposes**]It studies the expected value model of stochastic complementarity problems. [**Methods**]Based on the penalty function method and the sample average approximation method, the expected value model for stochastic complementarity problems is solved. [**Findings**]Under appropriate assumptions, the convergence and convergence rate of the optimal solution for the sample average approximation problem are proved. [**Conclusions**]The results provide a new way to study stochastic complementarity problems.

Keywords: penalty function method; stochastic complementarity problems; sample average approximation; convergence

(责任编辑 许 甲)