

求解随机广义垂直线性互补问题的随机近似方法

杨妍娇, 王奕菲, 张 杰*

辽宁师范大学数学学院, 辽宁 大连

收稿日期: 2023年3月11日; 录用日期: 2023年4月7日; 发布日期: 2023年4月13日

摘 要

近几年随机广义垂直线性互补问题的求解方法不断完善。本文提出了一种新型的求解随机广义垂直线性互补问题(SEVLCP)的方法, 即随机近似(SA)算法。基于Fischer-Burmeister函数的性质, 先将随机广义垂直线性互补问题转化为无约束极小化问题, 再利用随机近似算法进行求解。本文详细讨论了原问题的重新构造过程, 并提出了一种有效求解的迭代格式, 以及在适当的条件下, 得到了所提出方法的全局收敛结果。

关键词

随机广义垂直线性互补问题, 随机近似方法, Fischer-Burmeister函数

Stochastic Approximation Approaches to Stochastic Extended Vertical Complementarity Problem

Yanjiao Yang, Yifei Wang, Jie Zhang*

School of Mathematics, Liaoning Normal University, Dalian Liaoning

Received: Mar. 11th, 2023; accepted: Apr. 7th, 2023; published: Apr. 13th, 2023

Abstract

In recent years, the solution methods of stochastic extended vertical linear complementarity problems have been continuously improved. In this paper, a new method for solving stochastic extended

*通讯作者。

vertical linear complementarity problem (SEVLCP) is proposed, namely stochastic approximation (SA) methods. Based on the properties of the Fischer-Burmeister function, the stochastic extended vertical linear complementarity problem is reformulated in terms of the unconstrained minimization problem, and then solved by the stochastic approximation methods. This paper discusses the reformulation process of the original problem in detail, and proposes an iterative scheme for effective solving, and obtains the global convergence results of the proposed method under appropriate conditions.

Keywords

Stochastic Extended Vertical Linear Complementarity Problem, Stochastic Approximation Approaches, Fischer-Burmeister Function

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

本文研究如下的随机广义垂直线性互补问题：求 $x \in \mathbb{R}^n$ ，使得

$$\min \{F_1(x), F_2(x), F_3(x)\} = 0, \quad (1)$$

$$F_i(x) = E[M_i(\xi(\omega))]x + E[q_i(\xi(\omega))], i = 1, 2, 3,$$

其中, $M_i(\cdot): \mathbb{R}^m \rightarrow \mathbb{R}^{n \times n}$, $q_i(\cdot): \mathbb{R}^m \rightarrow \mathbb{R}^n$, $i = 1, 2, 3$ 为随机映射, $\xi: \Omega \rightarrow \Xi \subset \mathbb{R}^m$ 是概率空间 (Ω, F, P) 上的随机向量, E 表示数学期望, \min 表示极小化函数。在本文中, 假设 $M_i(\xi(\omega)), q_i(\xi(\omega)), i = 1, 2, 3$ 是 ω 的可测函数并满足

$$E \left[\sum_{i=1}^3 \|M_i(\xi(\omega))\| + \|q_i(\xi(\omega))\| \right] < +\infty,$$

为简单起见, 我们把 $\xi(\omega)$ 记为 ξ , 但要与上下文中 Ξ 中的确定性向量 ξ 区分开。

广义垂直线性互补问题在工业工程、经济管理、博弈论和网络中的应用已经成为数学规划中一个成熟而富有成果的学科, 在许多文献中都研究了广义垂直线性互补问题的解和算法的存在性, 如[1]。为了将现实中的不确定因素考虑其中, 带有随机变量的垂直线性互补问题受到了学者们的广泛关注。求解问题(1)的关键在于处理 $M_i(\xi)$ 和 $q_i(\xi)$ 的期望值, 如果能够准确地计算出问题(1)所涉及的数学期望, 则问题就转化为确定型广义随机线性互补问题, 这已经具有了成熟的求解方法, 如对数指数正则法[2]和人工神经网络法[3]等。本文假设问题(1)中的期望值难于计算或计算成本很高, 因此选用随机近似方法将其转化成确定型问题进行求解和近似。

随机近似方法可追溯到 Robbins 和 Monro 在 1951 年发表的论文[4], 自此以后随机近似算法被广泛用于解决随机问题, 如文献[5] [6] [7]。经典的随机近似算法分析在文献[8]中进行了详细的介绍。该算法的一个重要改进是由 Polyak [9]和 Polyak 和 Juditsky [10]提出的, 利用已知步长的平均值来获得更长的步长。在本文中, 将利用该方法求解随机广义垂直线性互补问题(1), 证明了其收敛性。

问题(1)是随机线性互补问题的一个延伸, 当 $i = 1, E[M_0(\xi(\omega))] = I, E[q_0(\xi(\omega))] = 0$ 时, 就简化为随机线性互补问题: 寻找 $x \in \mathbb{R}^n$, 使得

$$x \geq 0, F_1(x) \geq 0, F_1(x)^T x = 0, \quad (2)$$

$$F_1(x) = E[M_1(\xi(\omega))]x + E[q_1(\xi(\omega))],$$

它是随机非线性互补问题的一个特例，可以追溯到 King 和 Rockafellar [11]。因此，可以利用随机近似算法对垂直互补问题进行求解。

2. 问题的重新构造

我们将问题(1)重新表述为一个随机非光滑方程组，然后通过最小平方最小化来求解方程组。首先回顾 Fischer-Burmeister 函数 $\phi: \mathbb{R}^2 \rightarrow \mathbb{R}$

$$\phi(a, b) = \sqrt{a^2 + b^2} - a - b.$$

利用该函数有

$$\min\{a, b\} = \phi(a, b),$$

基于上述函数提出求解广义垂直线性互补问题的函数如下：

$$\min\{a, b, c\} = \phi(\phi(a, b), c).$$

设

$$\varphi(a, b, c) = \phi(\phi(a, b), c).$$

问题(1)等价于下列随机方程组

$$G(x) = \begin{pmatrix} \varphi(F_{11}(x), F_{21}(x), F_{31}(x)) \\ \vdots \\ \varphi(F_{1n}(x), F_{2n}(x), F_{3n}(x)) \end{pmatrix} = 0 \quad (3)$$

式(1)和式(3)的解集是重合的，使用 Fischer-Burmeister 函数的主要好处之一是它在任何地方都是半光滑的，并且在除原点以外的任何点上都是连续可微的，并且该函数是全局 Lipschitz 连续的。对于随机垂直线性互补问题(1)的解，提出求解随机方程组(3)的随机近似算法是很自然的[5]。在接下来的内容中，我们考虑基于(1)的最小化重构的 SA 方法。令

$$\min_x \Psi(x) = \frac{1}{2} \|G(x)\|^2 = \frac{1}{2} \sum_{j=1}^n \varphi(F_{1j}(x), F_{2j}(x), F_{3j}(x))^2. \quad (4)$$

为方便下面的证明，不妨令

$$g_j(x) = \varphi(F_{1j}(x), F_{2j}(x)), j = 1, \dots, n,$$

$$\psi(a, b) = \phi(a, b)^2,$$

则

$$\min \Psi(x) = \frac{1}{2} \sum_{j=1}^n \phi(g_j(x), F_{3j}(x))^2 = \frac{1}{2} \sum_{j=1}^n \psi(g_j(x), F_{3j}(x)). \quad (5)$$

考虑以下的迭代格式求解上述问题

$$x^{k+1} = x^k - a_k (d^k + \omega^k) \quad (6)$$

其中，

$$d^k = -\frac{1}{2} \nabla \Psi(x^k),$$

ω^k 是随机变量为确定值 ξ^k 时得到的 d^k 的近似值的随机误差。例如，如果 ξ^k 是 $\xi(\omega)$ 的一个样本，那么选择

$$d^k + \omega^k = -\frac{1}{2} \nabla \Psi(x^k, \xi^k).$$

在[12]中利用这种搜索方向建立了求解非线性互补问题的无导数迭代格式。下面的结果表明，在适当的条件下， d^k 是 Ψ 在 x^k 处的下降方向，并且当 $d^k = 0$ 时 x^k 是问题(4)的一个解。

命题 2.1: 假设 F 是全局 Lipschitz 连续的并且二次连续可微，存在正的常数 C_1, C_2, C_3 使得

$$\max_j (\|x\| + C_1) \|\nabla^2 F_{ij}(x)\| \leq C_2, i = 1, 2, 3, j = 1, \dots, n,$$

$$\max_j (\|x\| + C_1) \|\nabla^2 g_j(x)\| \leq C_3, j = 1, \dots, n, \forall x \in \mathbb{R}^n.$$

那么， $\nabla \Psi$ 在 \mathbb{R}^n 上是全局 Lipschitz 连续的。

证: 由定义(5)可知

$$\nabla \Psi(x) = \sum_{j=1}^n [\nabla_1 \psi(g_j(x), F_{3j}(x)) \nabla g_j(x) + \nabla_2 \psi(g_j(x), F_{3j}(x)) \nabla F_{3j}(x)],$$

要证 $\nabla \Psi$ 是全局 Lipschitz 连续的，只需证 $\nabla_1 \psi(g_j(x), F_{3j}(x)) \nabla g_j(x)$ 和 $\nabla_2 \psi(g_j(x), F_{3j}(x)) \nabla F_{3j}(x)$ 分别是全局 Lipschitz 连续的。

由([5], Proposition 5.1)的证明有 $\nabla_2 \psi(g_j(x), F_{3j}(x)) \nabla F_{3j}(x)$ 是全局 Lipschitz 连续的。接下来证明 $\nabla_1 \psi(g_j(x), F_{3j}(x)) \nabla g_j(x)$ 是全局 Lipschitz 连续的。当且仅当它的雅可比矩阵包含在

$$\Omega \equiv \nabla g_j(x) \partial \nabla_1 \psi(g_j(x), F_{3j}(x))^T + \nabla_1 \psi(g_j(x), F_{3j}(x)) \nabla^2 g_j(x)$$

中， Ω 是有界的。首先，由([5], Lemma 5.2)知 $\partial \nabla_1 \psi(g_j(x), F_{3j}(x))$ 有界。其次，由

$$g_j(x) = \phi(F_{1j}(x), F_{2j}(x)),$$

得

$$\nabla g_j(x) = \nabla \phi(F_{1j}(x), F_{2j}(x)) = \nabla_1 \phi(F_{1j}(x), F_{2j}(x)) \nabla F_{1j}(x) + \nabla_2 \phi(F_{1j}(x), F_{2j}(x)) \nabla F_{2j}(x).$$

因为

$$|\nabla_1 \phi(a, b)| = \left| \frac{a}{\sqrt{a^2 + b^2}} - 1 \right| \leq 1, |\nabla_2 \phi(a, b)| = \left| \frac{b}{\sqrt{a^2 + b^2}} - 1 \right| \leq 1,$$

从而 $\nabla_1 \phi(F_{1j}(x), F_{2j}(x))$ 和 $\nabla_2 \phi(F_{1j}(x), F_{2j}(x))$ 是有界的。并且 $F_i(x), i = 1, 2, 3$ 是 Lipschitz 连续的，则 $\nabla F_{ij}(x)$ 有界，所以 $\nabla g_j(x)$ 有界。

又

$$\begin{aligned} & \|\nabla_1 \psi(g_j(x), F_{3j}(x))\| \\ &= 2 \|\phi(g_j(x), F_{3j}(x)) \nabla_1 \phi(g_j(x), F_{3j}(x))\| \\ &\leq 4 \max \{ \|g_j(x)\|, |F_{3j}(x)| \} \\ &\leq 8 \max \{ \|F_{1j}(x)\|, |F_{2j}(x)|, |F_{3j}(x)| \} \\ &= O(\|x\| + C_1). \end{aligned}$$

根据命题 2.1 可知 Ω 有界，从而 $\nabla_1 \psi(g_j(x), F_{3j}(x)) \nabla g_j(x)$ 是全局 Lipschitz 连续的。

综上, $\nabla\Psi$ 是全局 Lipschitz 连续的。

3. 收敛性分析

为证明在迭代格式(5)下 Ψ 的收敛性, 作出下列假设:

假设 2.1: 1) 步长 a_k 满足 $a_k > 0, a_k \rightarrow 0, \sum_{k=0}^n a_k = \infty, \sum_{k=0}^n (a_k)^2 < \infty$;

2) 问题 $E[\omega^k | \mathcal{F}_k] = 0$ 。

3) $\sum_{k=1}^{\infty} (a_k)^2 E[\|\omega^k\|^2 | \mathcal{F}_k] < \infty$ 几乎确定成立。

4) $F_i, i=1,2,3$ 在空间 \mathcal{Y} 上是以模量 L 全局 Lipschitz 的。

5) $F_i, i=1,2,3$ 在空间 \mathcal{Y} 上是以模量 σ 全局强单调的。

6) 矩阵 $F = (F_1(x), F_2(x), F_3(x))$ 有行 ω -性质[13]。

我们关于随机近似算法的收敛结果使用了下面的引理, 它是鞅收敛定理的推广。

引理 3.1: [14] 令 \mathcal{F}_k 是 σ -代数的一个递增序列, V_k, α_k, β_k 和 γ_k 是适应于 \mathcal{F}_k 非负随机变量。如果

$$\sum_{k=1}^{\infty} \alpha_k < \infty, \sum_{k=1}^{\infty} \beta_k < \infty,$$

并且

$$E(V_{k+1} | \mathcal{F}_k) \leq (1 + \alpha_k)V_k - \gamma_k + \beta_k,$$

那么 $\{V_k\}$ 是几乎确定收敛的, 以及有 $\sum_{k=1}^{\infty} \gamma_k < \infty$ 几乎确定成立。

定理 3.1: 假设在假设 2.1 中的 1), 2), 3) 成立, 在给定得迭代格式(6)下, 设 $F_i, i=1,2,3$ 都是二次连续可微的, 存在 $t \in (1,2), C_i > 0$ 和一个连续函数 $\sigma_i(x) > 0$, 使得

$$(F_i(y) - F_i(x))^T (y - x) \geq \min\{\sigma_i(x)\|y - x\|^2, C_i\|y - x\|^t\}, \forall y \in \mathbb{R}^n,$$

以及步长满足

$$0 \leq a_k \leq \min\left\{\frac{\sigma(x^k)}{2L}, \frac{1}{2L}\right\},$$

那么由迭代生成的序列 $\{x^k\}$ 几乎确定收敛到随机广义垂直线性互补问题(1)唯一解 x^* 。

证: 由中值定理可知, 在 x^k 和 x^{k+1} 之间存在 y^k 使得

$$\Psi(x^{k+1}) = \Psi(x^k) + \nabla\Psi(y^k)^T (x^{k+1} - x^k).$$

由命题 2.1, $\nabla\Psi$ 是全局 Lipschitz 连续的。因此, 存在常数 $L > 0$, 使得

$$\|\nabla\Psi(y^k) - \nabla\Psi(x^k)\| \leq L\|x^{k+1} - x^k\|.$$

由迭代(6)可得

$$\begin{aligned} & E[\Psi(x^{k+1}) | \mathcal{F}_k] \\ &= E[\Psi(x^k) + \nabla\Psi(y^k)^T (x^{k+1} - x^k) | \mathcal{F}_k] \\ &= E[\Psi(x^k) + a_k \nabla\Psi(y^k)^T (d^k + \omega^k) | \mathcal{F}_k] \\ &= \Psi(x^k) + a_k E[\nabla\Psi(y^k)^T (d^k + \omega^k) | \mathcal{F}_k] \\ &= \Psi(x^k) + a_k E[\nabla\Psi(x^k)^T (d^k + \omega^k) | \mathcal{F}_k] + a_k E[(\nabla\Psi(y^k) - \nabla\Psi(x^k))^T (d^k + \omega^k) | \mathcal{F}_k] \end{aligned}$$

$$\begin{aligned}
 &= \Psi(x^k) + a_k \nabla \Psi(x^k)^T d^k + 0 + a_k \mathbb{E} \left[\left(\nabla \Psi(y^k) - \nabla \Psi(x^k) \right)^T (d^k + \omega^k) \mid \mathcal{F}_k \right] \\
 &\leq \Psi(x^k) + a_k \nabla \Psi(x^k)^T d^k + a_k \mathbb{E} \left[L \|y^k - x^k\| \cdot \|d^k + \omega^k\| \mid \mathcal{F}_k \right] \\
 &\leq \Psi(x^k) + a_k \nabla \Psi(x^k)^T d^k + La_k^2 \mathbb{E} \left[\|d^k + \omega^k\|^2 \mid \mathcal{F}_k \right] \\
 &\leq \Psi(x^k) + a_k \nabla \Psi(x^k)^T d^k + 2La_k^2 \mathbb{E} \left[\|d^k\|^2 + \|\omega^k\|^2 \mid \mathcal{F}_k \right] \\
 &= \Psi(x^k) + a_k \nabla \Psi(x^k)^T d^k + 2La_k^2 \|d^k\|^2 + 2La_k^2 \mathbb{E} \left[\|\omega^k\|^2 \mid \mathcal{F}_k \right] \\
 &\leq \Psi(x^k) - a_k \|d^k\|^2 + 2La_k^2 \|d^k\|^2 + 2La_k^2 \mathbb{E} \left[\|\omega^k\|^2 \mid \mathcal{F}_k \right] \\
 &= \Psi(x^k) + a_k (2La_k - 1) \|d^k\|^2 + 2La_k^2 \mathbb{E} \left[\|\omega^k\|^2 \mid \mathcal{F}_k \right] \\
 &= \Psi(x^k) - \gamma_k + \beta_k.
 \end{aligned}$$

其中,

$$\begin{aligned}
 \gamma_k &= -a_k (2La_k - 1) \|d^k\|^2, \\
 \beta_k &= 2La_k^2 \mathbb{E} \left[\|\omega^k\|^2 \mid \mathcal{F}_k \right].
 \end{aligned}$$

下面证明对应 $\Psi(x^k)$ 收敛的每个迭代路径 $\{x^k\}$, 有 $\Psi(x^k)$ 收敛到 0。首先根据([5], Theorem 5.1) 的证明可知序列 $\{x^k\}$ 有界, 则存在 $\hat{\sigma} > 0$, 使得 $\sigma(x^k) \geq \hat{\sigma} > 0$ 。由

$$\sum_{k=1}^{\infty} \gamma_k < \infty$$

可得, 当 $k \rightarrow \infty$ 时 $d^k \rightarrow 0$ 。由 d^k 的定义有 $\nabla \Psi(x^k) \rightarrow 0$ 。说明序列 $\{x^k\}$ 收敛到某一点 x^* 。其次证明 x^* 是 $\Psi(x^k)$ 的极小点:

$$\nabla \Psi(x^*) = \sum_{j=1}^n \phi(g_j(x^*), F_{3_j}(x^*)) \nabla \phi(g_j(x^*), F_{3_j}(x^*)) = 0,$$

则

$$\phi(g_j(x^*), F_{3_j}(x^*)) = 0 \text{ 或 } \nabla \phi(g_j(x^*), F_{3_j}(x^*)) = 0.$$

由于

$$\nabla \phi(a, b) = \nabla_1 \phi(a, b) + \nabla_2 \phi(a, b) = \left(\frac{a}{\sqrt{a^2 + b^2}} - 1 \right) + \left(\frac{b}{\sqrt{a^2 + b^2}} - 1 \right),$$

由假设 2.1 中 6) 知 F 有行 ω -性质, 存在 $i \in \{1, 2, 3\}$ 使得 $\nabla F_{ij}(x^*) \neq 0$ 。则有

$$\phi(g_j(x^*), F_{3_j}(x^*)) = 0, j = 1, \dots, n.$$

即 $G(x^*) = 0$, 因此 $\Psi(x^*) = 0$ 。 ■

4. 结论

本文利用随机近似方法求解了随机广义垂直线性互补问题, 基于 Fischer-Burmeister 函数将随机广义

垂直线性互补问题转化为一个无约束极小化问题，构造了一个迭代格式，证明了在合适的条件下这种迭代格式所产生的序列收敛性以及原问题的可解性。

参考文献

- [1] Gowda, M.S. and Sznajder, R. (1994) The Generalized Order Linear Complementarity Problem. *SIAM Journal on Matrix Analysis and Applications*, **15**, 779-795. <https://doi.org/10.1137/S0895479892237859>
- [2] Shuang, J.Z., et al. (2013) A Log-Exponential Regularization Method for a Mathematical Program with General Vertical Complementarity Constraints. *Journal of industrial and management optimization*, **9**, 561-577. <https://doi.org/10.3934/jimo.2013.9.561>
- [3] Hou, B., Zhang, J. and Qiu, C. (2022) A Neural Network for a Generalized Vertical Complementarity Problem. *AIMS Mathematics*, **7**, 6650-6668. <https://doi.org/10.3934/math.2022371>
- [4] Robbins, H. and Monro, S. (1951) A Stochastic Approximation Method. *Annals of Mathematical Statistics*, **22**, 400-407. <https://doi.org/10.1214/aoms/1177729586>
- [5] Jiang, H. and Xu, H. (2008) Stochastic Approximation Approaches to the Stochastic Variational Inequality Problem. *IEEE Transactions on Automatic Control*, **53**, 1462-1475. <https://doi.org/10.1109/TAC.2008.925853>
- [6] 庞丽萍, 田琦, 陈爽, 等. 基于投影收缩的 SA 方法求解随机变分不等式问题[J]. 汕头大学学报: 自然科学版, 2015, 30(4): 71-74
- [7] Metivier, M., Priouret, P. and Benveaniste, A. (2012) Adaptive Algorithms and Stochastic Approximations. Springer Publishig, New York, 380-387.
- [8] Chung, K.L. (1953) On a Stochastic Approximation Method. *The Annals of Mathematical Statistics*, **25**, 463-483. <https://doi.org/10.1214/aoms/1177728716>
- [9] Polyak, B. (1990) New Stochastic Approximation Type Procedures. *Automation and Remote Control*, **7**, 98-107.
- [10] Polyak, B. and Juditsky, A. (1992) Acceleration of Stochastic Approximation by Averaging. *SIAM Journal on Control and Optimization*, **30**, 838-855. <https://doi.org/10.1214/aoms/1177728716>
- [11] Zhang, J., He, S.X. and Wang, Q. (2014) A SAA Nonlinear Regularization Method for a Stochastic Extended Vertical Linear Complementarity Problem. *Applied Mathematics & Computation*, **232**, 888-897. <https://doi.org/10.1016/j.amc.2014.01.121>
- [12] Geiger, C. and Kanzow, C. (1996) On the Resolution of Monotone Complementarity Problems. *Computational Optimization & Applications*, **5**, 155-173. <https://doi.org/10.1007/BF00249054>
- [13] Sznajder, R. and Gowda, M.S. (1995) Generalizations of P0 and P Properties; Extended Vertical and Horizontal Linear Complementarity Problems. *Linear Algebra and Its Applications*, **223-224**, 695-715. [https://doi.org/10.1016/0024-3795\(93\)00184-2](https://doi.org/10.1016/0024-3795(93)00184-2)
- [14] Robbins, H. and Siegmund, D. (1971) A Convergence Theorem for Non Negative almost Supermartingales and Some Applications. *Optimizing Methods in Statistics*, **1971**, 233-257. [https://doi.org/10.1016/0024-3795\(93\)00184-2](https://doi.org/10.1016/0024-3795(93)00184-2)