

基于修正的Carroll函数的概率约束优化问题的对偶算法

冯千千

辽宁师范大学数学学院, 辽宁 大连

收稿日期: 2024年10月27日; 录用日期: 2024年11月21日; 发布日期: 2024年11月28日

摘 要

本文用连续可微非凸函数描述的概率约束分析非线性随机优化问题。为此描述了潜在概率函数的水平集的切锥和法锥, 并在此基础上, 提出 p -有效点的定义, 形成问题的一阶和二阶最优性条件, 基于 p -有效点生成的概率函数的水平集, 通过修正的Carroll函数生成一个对偶算法。

关键词

概率约束, p -有效点, 修正的Carroll函数, 对偶算法

A Dual Algorithm for Probabilistic Constrained Optimization Problems Based on Modified Carroll Functions

Qianqian Feng

School of Mathematics, Liaoning Normal University, Dalian Liaoning

Received: Oct. 27th, 2024; accepted: Nov. 21st, 2024; published: Nov. 28th, 2024

Abstract

In this paper, probabilistic constraints described by continuously differentiable non-convex functions are used to analyze nonlinear stochastic optimization problems. To this end, the tangent and normal cones of the level set of potential probability functions are described, and on this basis, the definition of p -effective points is proposed to form the first and second order optimality conditions of the problem. Based on the water-level set of probability functions generated by p -effective points, a dual algorithm is generated by the modified Carroll function.

文章引用: 冯千千. 基于修正的 Carroll 函数的概率约束优化问题的对偶算法[J]. 应用数学进展, 2024, 13(11): 5100-5105. DOI: 10.12677/aam.2024.1311492

Keywords

Probability Constraint, p -Effective Point, Modified Carroll Function, Dual Algorithm

Copyright © 2024 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

目前, 很多有重要价值的实际问题都属于概率约束优化问题, 因而概率约束优化问题的研究具有重要的理论意义和应用价值。

如风险优化中的风险值计算问题, 含新能源电力系统调度问题, 移动网络中的路线定位处理问题等均可建模为下述概率约束优化问题:

$$\begin{aligned} & \min f(x) \\ & \text{s.t. } \Pr[g(x) \geq Y] \geq p, \\ & \quad x \in X, \end{aligned} \quad (\text{PCOP})$$

其中 $f: \mathbb{R}^n \rightarrow \mathbb{R}$, $g(x) = (g_1(x), \dots, g_m(x))$, $g_i: \mathbb{R}^n \rightarrow \mathbb{R}^m$, 是二次连续可微函数, $X \subset \mathbb{R}^n$ 是一个闭凸集, $Y \in \mathbb{R}^m$ 是随机的, 概率 $p \in (0, 1)$ 是置信水平, $\Pr$ 代表概率。

对于 $p \in (0, 1)$, Y 的分布函数的水平集为

$$\mathcal{L}_p = \{y \in \mathbb{R}^m : \Pr[Y \leq y] \geq p\}.$$

由分布函数的单调性和右连续性易知, 集合 $\mathcal{L}_p$ 是非空且闭的, 问题(PCOP)可表示为

$$\begin{aligned} & \min f(x) \\ & \text{s.t. } g(x) \in \mathcal{L}_p, \\ & \quad x \in X. \end{aligned} \quad (1)$$

非线性 Lagrange 函数是经典的 Lagrange 函数的修正形式, 它关于乘子向量或约束函数是非线性函数; 基于非线性 Lagrange 函数建立的求解非线性优化问题的对偶方法即为非线性 Lagrange 方法。由于对偶方法对原始变量的可行性没有限制。因此非线性 Lagrange 方法在求解约束优化问题中扮演着重要的角色。

2. 概率约束优化问题(PCOP)的凸近似形式

对于 $y \in \mathbb{R}^m$, Y 的分布函数为:

$$F_Y(y) = \Pr[Y \leq y].$$

定义 1 [1] 对于 $p \in (0, 1)$, $v \in \mathbb{R}^m$ 被称为概率分布函数 F_Y 的一个 p -有效点, 如果 $F_Y(v) \geq p$, 并且不存在 $y \leq v$, $y \neq v$, 使得 $F_Y(y) \geq p$ 。

对于 $p \in (0, 1)$, 记

$$\mathcal{E} = \{v \mid v \text{ 是 } F_Y(y) \text{ 的 } p\text{-有效点}\},$$

对于 $v^j \in \mathcal{E}$, $j \in J$, 定义

$$K_j = v^j + \mathbb{R}_+^m,$$

这里的 J 是一个任意集合。

命题 1 [1] $\mathcal{L}_p = \bigcup_{j \in J} K_j$ 。

用 $\mathcal{S}_{m+1}$ 表示 $\mathbb{R}^{m+1}$ 中的单纯形, 即

$$\mathcal{S}_{m+1} = \left\{ \alpha \in \mathbb{R}^{m+1} : \sum_{i=1}^{m+1} \alpha_i = 1, \alpha_i \geq 0, i = 1, \dots, m+1 \right\},$$

p -有效点集的凸包 $co\mathcal{E}$ 表示如下:

$$co\mathcal{E} = \left\{ \sum_{i=1}^{m+1} \alpha_i v^{j_i} : \alpha \in \mathcal{S}_{m+1}, j_i \in J \right\}.$$

命题 2 [1] 对于每一个 $p \in (0, 1)$, 集合 $co\mathcal{L}_p$ 是闭的。

定义 2 设 S 为非空凸集, $x \in S$, 若 x 不能用 S 中的两个不同点的凸组合表示, 即不存在 $x_1 \in S$, $x_2 \in S$, $x_1 \neq x_2$, 使

$$x = \alpha x_1 + (1 - \alpha) x_2, \quad 0 < \alpha < 1,$$

成立, 则称 x 为 S 的一个极点。

命题 3 [1] 对于每个点 $p \in (0, 1)$, 有

$$co\mathcal{L}_p = co\mathcal{E} + \mathbb{R}_+^m,$$

$co\mathcal{L}_p$ 的极点集 $E \neq \emptyset$, 且 $E \subset co\mathcal{E}$ 。

对于任意 $p \in (0, 1)$, $\mathcal{L}_p$, $co\mathcal{L}_p$ 和 $\mathcal{E}$ 有下界[1]。

引理 1 [2] $co\mathcal{L}_p = coE + \mathbb{R}_+^m$ 。

由此, 问题(1)近似于如下问题:

$$\begin{aligned} & \min f(x) \\ & \text{s.t. } g(x) \in co\mathcal{L}_p, \\ & \quad x \in X. \end{aligned} \quad (2)$$

3. 最优性条件

记问题(2)的可行域为:

$$\Phi = \{x \mid g(x) \in co\mathcal{L}_p, x \in X\}.$$

$\nabla f(x)$ 为 f 在 x 处的梯度。 $\mathcal{J}g(x)$ 为 g 在 x 处的 Jacobian。

令 $\hat{x} \in \Phi$, 则存在 $\hat{w} \in \mathcal{E}$, $\hat{y} \in \mathbb{R}_+^m$ 使得 $g(\hat{x}) = \hat{w} + \hat{y}$, 记

$$I_0(\hat{x}) = \{i \mid \hat{y}_i = 0, 1 \leq i \leq m\}$$

是 $\hat{x}$ 处的积极约束, $\mathcal{K}_X(\hat{x})$ 是集合 X 在 $\hat{x} \in X$ 处的可行方向锥, 向量 $\delta \in \mathcal{K}_X(\hat{x})$ 和一个 p -有效点 ω_i^0 的凸组合有如下关系:

$$\forall i \in I_0(\hat{x}), g_i(\hat{x}) + \langle \nabla g_i(\hat{x}), \delta \rangle > \omega_i^0. \quad (3)$$

条件(3)等价于问题(2)在可行点 $\hat{x}$ 处的 Robinson's 约束规范条件。

命题 4 [2] 在可行点 $\hat{x} \in \Phi$ 处的约束规范(3)条件下, 记

$$T_{\Phi}(\hat{x}) = \{\delta \in T_X(\hat{x}) : \exists \omega \in \mathcal{E}, \exists \beta > 0 : g_i(\hat{x}) + \langle \nabla g_i(\hat{x}), \beta \delta \rangle \geq \omega_i \quad \forall i \in I_0(\hat{x})\}$$

为可行集 Φ 在 $\hat{x}$ 处的切锥。

定义 $L(x, u) = f(x) - \langle u, g(x) \rangle$ 。记 $\mathcal{N}_A(z)$ 为集合 A 在 $z \in A$ 处的法锥。

定理 1 [2] 假设 $\hat{x}$ 是问题(2)的一个局部最优解, 且在 $\hat{x}$ 处满足条件(3), 则存在一个向量 $\hat{u} \in \mathbb{R}_+^m$, 使得

$$-\nabla_x L(\hat{x}, \hat{u}) \in \mathcal{N}_X(\hat{x}), \quad -\hat{u} \in \mathcal{N}_{\text{co}\mathcal{E}_p}(g(\hat{x})) \quad (4)$$

如果函数 f 是凸的, g 是一个凹映射, 存在某一向量 $\hat{u} \in \mathbb{R}_+^m$, 使得 $\hat{x} \in \Phi$ 满足条件(4), 则 $\hat{x}$ 是问题(2)的一个全局最优解。

记

$$\Lambda(\hat{x}) = \{\hat{u} | 0 \in \nabla_x L(\hat{x}, \hat{u}) + \mathcal{N}_{\text{co}\mathcal{E}_p}(g(\hat{x}))\}.$$

记切锥

$$T_X(\hat{x}) = \limsup_{t \searrow 0} \frac{X - x}{t}.$$

定理 2 [2] 假设 $\hat{x} \in \Phi$ 满足条件(3)和(4), 对于 $u \in \Lambda(\hat{x})$, 以及对于每个满足条件

$$\langle \nabla f(\hat{x}), \delta \rangle = 0 \text{ 且 } \langle u, \nabla g(\hat{x}) \delta \rangle = 0$$

的非零的 $\delta \in T_X(\hat{x})$, 有

$$\langle \delta, \nabla_x^2 L(\hat{x}, u) \delta \rangle > 0, \quad (5)$$

则 $\hat{x}$ 是问题(2)的一个局部极小解。

记 $I \subset J$ 是问题(2)的 p -有效点集的子集, $|I|$ 表示它的基数。由引理 1 可知,

$$g(x) = \sum_{j \in J} \lambda_j v^j + y_i, \quad v^j \in \mathcal{E}, \quad \lambda \in S_{|I|}, \quad y_i \in \mathbb{R}_+^m.$$

于是, 问题(2)的近似问题如下:

$$\begin{aligned} & \min_{x, \lambda} f(x) \\ & \text{s.t. } g(x) \geq \sum_{j \in I} \lambda_j v^j, \\ & \quad x \in X, \\ & \quad \lambda \in S_{|I|}. \end{aligned} \quad (6)$$

注意, 当 $I = J$ 时, 问题(6)和(2)重合。

4. 求解问题(3)的对偶算法

考虑 Polyak [3] 在 1992 年提出的修正的 Carroll 函数:

$$C(x, u, k) = \begin{cases} f_0(x) - k^{-1} \sum_{i=1}^m u_i \left[1 - (k f_i(x) + 1)^{-1} \right], & x \in \text{int } \Omega_k, \\ +\infty, & x \notin \text{int } \Omega_k, \end{cases}$$

其中 $k > 0$ 是惩罚参数, $\Omega_k = \{x | 1 + k f_i(x) \geq 0, i = 1, \dots, m\}$, 则问题(6)的修正的 Carroll 函数有如下形式:

$$F_{\rho}(x, \lambda, u) = f(x) - \rho^{-1} \sum_{i=1}^m u_i \left[1 - \left(\rho \left(g_i(x) - \sum_{j \in I} \lambda_j v_i^j \right) + 1 \right)^{-1} \right], \quad x \in \text{int } \Omega_{\rho},$$

其中,

$$\Omega_\rho = \left\{ x \mid 1 + \rho \left(g_i(x) - \sum_{j \in I} \lambda_j v_i^j \right) \geq 0, i = 1, \dots, m \right\}.$$

对偶算法

步 0 选择一个 p -有效点 v^1 使得存在一点 $\tilde{x} \in X$ 满足 $g(\tilde{x}) > v^1$ 。

令 $J_1 = \{v^1\}$, $k=1$ 。且选择 $u^1 \geq 0$ 。

步 1 解决主问题

$$\min_{x, \lambda} F_{\rho, k}(x, \lambda, u^k), \\ x \in X, \lambda \in S_k$$

设 (x^k, λ^k) 是这个问题的解。

步 2 如果 $k > 1$ 且 $\|u^k - u^{k-1}\| < \varepsilon$, 那么停止, 否则继续步 3。

步 3 计算

$$u_i^{k+1} = u_i^k \frac{1}{\rho \left(g_i(x) - \sum_{j \in I} \lambda_j v_i^j \right) + 1}.$$

步 4 找到一个 p -有效点 v^{k+1} 作为如下问题的解:

$$\min_{z \in \text{co}\mathcal{L}_p} \langle u^{k+1}, z \rangle,$$

令 $J_{k+1} = \{v^{k+1}\} \cup J_k$ 。

步 5 令 $k \rightarrow k+1$, 然后转步 1。

接下来, 我们做如下假设:

A1 存在一个数 $d > 0$, 使得随机向量 Y 的两个实现间的距离大于 d 。

A2 集合 X 是凸的、紧致的。

定理 3 [2] 存在一个有限子集 $I \subseteq J$, 使得问题(2)的局部一阶最优性条件和问题(6)的等价。

定理 4 假设 A1 和 A2, 且问题(2)的可行域非空, 则对于参数 $\rho > 0$, 由对偶算法产生的序列 $\{x^k\}$ 和 $\{u^k\}$ 的聚点 $\hat{x}$, $\hat{u}$, 满足问题(2)的一阶最优性条件, 此外, 若 $\hat{x}$ 满足条件(3)和(5), 那么 $\hat{x}$ 是问题(2)的一个局部极小解。

证明: 由于在 A1 下, 只存在许多有限的 p -有效点, 算法没有产生 p -有效点的非有限序列。因此, 在某次迭 k_0 后, 没有新的 p -有效点在步 4 中出现, 当 X 紧致, $\{x^k\}$ 有一个收敛到子序列的聚点。由于 g 的连续性, $\{g(x^k)\}_{k \in K}$ 有一个聚点 $g(\hat{x})$ 。我们记 $z^k = \sum_{j \in J_k} \lambda_j^k v^j$ 。由于有限个 p -有效点存在, 集合 $\mathcal{E}$ 是有界的且 $\{z_k\}_{k \in K}$ 在 $\text{co}\mathcal{L}_p$ 中有一聚点, 我们记作 $\hat{z}$ 。相似地, 序列 $\{\lambda^k\}$ 有聚点 $\hat{\lambda} \in S_{k_0}$ 且 $\hat{z} = \sum_{j \in J_k} \hat{\lambda}_j^k v^j$ 。我们

可以选择一个迭代序列 $\mathcal{K} \subset \{1, 2, 3, \dots\}$ 以至于这些聚点是相关子序列的极限。记 $J = \{i: \hat{z} < g(\hat{x})\}$ 我们注意到对于足够大的 $k \in \mathcal{K}$, ρ 足够大的时候, 我们得到 $u_i^k = 0$ 。而且, 根据算法步 3, 对于 $i \notin J$, 有

$$u_i^{k+1} = \frac{1}{\rho \left(g_i(x) - \sum_{j \in I} \lambda_j v_i^j \right) + 1}. \quad (7)$$

我们将证明对于 $i \notin J$, $\hat{z}_i = g_i(\hat{x})$ 。假设对于某个 $i \notin J$, $\hat{z}_i > g_i(\hat{x})$ 成立, 那么 $\lim_{k \rightarrow \infty} u_i^k = \infty$ 。此外, k 足够大时, 修正的 Carroll 函数有如下形式:

$$F_\rho(x, \lambda, u) = f(x) - \rho^{-1} \sum_{i=1}^m u_i \left[1 - \left(\rho \left(g_i(x) - \sum_{j \in I} \lambda_j v_i^j \right) + 1 \right)^{-1} \right].$$

因此, $\lim_{k \rightarrow \infty} F_\rho(x^k, \lambda^k, u^k) = \infty$ 。另一方面, 设 $\tilde{\lambda} = (1, 0, \dots, 0)^T \in S_{J_k}$, 我们得到 $(\tilde{x}, \tilde{\lambda}) \in X \times S_{J_k}$ 。当 k 足够大时,

$$F_\rho(\tilde{x}, \tilde{\lambda}, u) = f(\tilde{x}) - \rho^{-1} \sum_{i=1}^m u_i \left[1 - \left(\rho \left(g_i(\tilde{x}) - \sum_{j \in I} \tilde{\lambda}_j v_i^j \right) + 1 \right)^{-1} \right] \leq f(\tilde{x}) < F_{\rho, k}(x^k, \lambda^k, u^k)$$

这和 (x^k, λ^k) 是步骤一中主问题的解的结论是矛盾的, 因此 $\hat{z}_i \leq g_i(\hat{x})$, 这表明对于所有的 $i \notin J$, $\hat{z}_i = g_i(\hat{x})$ 。结合(7)式得到, $\{u^k\}$ 有一个聚点 $\hat{u}$, 使得对于所有 $i: \hat{u} > 0$, 有 $\hat{z}_i = g_i(\hat{x})$ 。

步 1 中主问题的最优性条件是:

对于足够大的 $k \in \mathcal{K}$, 我们可以将步 1 中主问题的最优性条件写成

$$-\nabla F(x^k, u^{k+1}) \in \mathcal{N}_X(x^k) \quad (8)$$

$$-V_{J_0}^T u^k \in \mathcal{N}_{S_{k_0}}(\lambda^k) \quad (9)$$

条件(9)意味着

$$0 \geq \langle -u^k, V_{J_0}(\lambda - \lambda^k) \rangle = \langle -u^k, V_{J_0}\lambda - z^k \rangle \quad (10)$$

假设存在一点 $\omega \in \text{co}\Omega_p \setminus \text{co}\{v^j, j \in J_k\}$ 使得 $\langle u^k, \omega - z^k \rangle < 0$, 应用这个假设和(10), 我们得到

$$\langle -u^k, \omega - v^k \rangle = \langle -u^k, \omega - z^k \rangle + \langle -u^k, z^k - v^k \rangle > \langle -u^k, z^k - v^k \rangle \geq 0.$$

不等式 $\langle -u^k, \omega - v^k \rangle > 0$ 和问题 $\min_{z \in \text{co}\Omega_p} \langle u^{k+1}, z \rangle$ 的最优性条件是矛盾的, 这说明 $-u^k \in N_{\text{co}\Omega_p}(v^k)$ 。因此, $-u^k \in N_{\text{co}\Omega_p}(z^k)$, 结合(8), 表明

$$-\nabla F(x^k, u^{k+1}) \in \mathcal{N}_X(x^k) \quad (11)$$

$$-u^{k+1} \in \mathcal{N}_{\text{co}\Omega_p}(z^{k+1}) \quad (12)$$

对 k 取极限, 并将 Ruszczyński [4] 中的引理 2.42 应用到(11)-(12), 我们得到

$$-\nabla F(\hat{x}, \hat{u}) \in \mathcal{N}_X(\hat{x}), \quad -\hat{u} \in \mathcal{N}_{\text{co}\Omega_p}(g(\hat{x})).$$

因此, $\hat{x}$ 满足问题(2)的条件(4), 根据定理 2, $\hat{x}$ 是问题(2)的一个局部极小解。证毕。

参考文献

- [1] Dentcheva, D., Lai, B. and Ruszczyński, A. (2004) Dual Methods for Probabilistic Optimization Problems. *Mathematical Methods of Operational Research*, **60**, 331-346. <https://doi.org/10.1007/s001860400371>
- [2] Dentcheva, D. and Martinez, G. (2011) Augmented Lagrangian Method for Probabilistic Optimization. *Annals of Operations Research*, **200**, 109-130. <https://doi.org/10.1007/s10479-011-0884-5>
- [3] Polyak, R. (1992) Modified Barrier Functions (Theory and Methods). *Mathematical Programming*, **54**, 177-222. <https://doi.org/10.1007/bf01586050>
- [4] Ruszczyński, A. (2006) Nonlinear Optimization. Princeton University Press.