

基于修正的指数函数的概率约束问题的对偶算法

张 微*, 张瑛珏, 王禹钧

辽宁师范大学数学学院, 辽宁 大连

收稿日期: 2024年10月25日; 录用日期: 2024年11月18日; 发布日期: 2024年11月27日

摘 要

本文通过连续可微的非凸函数所形成的概率约束, 来分析概率约束问题。描述了潜在的概率函数的水平集的切锥和法锥。进一步, 基于 p -有效点的概念, 形成这些问题的一阶和二阶最优性条件。对于离散分布函数的这种情况, 产生一个基于修正的指数函数的对偶算法来解决概率约束问题。

关键词

概率约束, p -有效点, 修正的指数函数, 对偶算法

A Dual Algorithm for Probabilistic Constraint Problems Based on Modified Exponential Functions

Wei Zhang*, Yingjue Zhang, Yujun Wang

College of Mathematics, Liaoning Normal University, Dalian Liaoning

Received: Oct. 25th, 2024; accepted: Nov. 18th, 2024; published: Nov. 27th, 2024

Abstract

In this paper, the problem of probability constraints is analyzed by means of the probability constraints formed by continuously differentiable non-convex functions. The tangent and normal cones of the level set of potential probability functions are described. Further, based on the concept of p -efficient points, the first and second order optimality conditions of these problems are formed. For

*通讯作者。

this case of the discrete distribution function, a dual algorithm based on the modified exponential function is generated to solve the probability constraint problem.

Keywords

Probability Constraints, p -Efficient Point, Modified Exponential 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

近年来, 概率约束优化问题成为备受关注的研究热点之一。该类问题在随机优化领域具有重要的理论意义和应用价值, 已经广泛地应用于很多实际问题中。

很多实际问题可以通过概率约束优化问题进行解决, 如现金匹配、风险优化、供应链管理、水资源管理等, 该类问题均可建模为:

$$\begin{aligned} \min & f(x) \\ \text{s.t. } & \Pr[g(x) \geq Y] \geq p, \\ & x \in \mathcal{D}, \end{aligned} \quad (1)$$

其中, $f: \mathbb{R}^n \rightarrow \mathbb{R}$, $g: \mathbb{R}^n \rightarrow \mathbb{R}^m$ 是二次连续可微函数, 且

$$g(x) = (g_1(x), g_2(x), \dots, g_m(x)).$$

$\mathcal{D} \subset \mathbb{R}^n$ 是一个闭凸集, $Y \in \mathbb{R}^m$ 是随机变量, $p \in (0, 1)$ 为置信水平, $\Pr$ 代表概率。

对于概率约束优化问题, 国内外众多学者给出过不同的研究方法, Ruszczyński 等学者[1]在 2000 年关于“包含整数值随机变量的概率约束优化问题”作出研究, 基于 p -有效点的定义, 推导出各种等价问题的形式, 又引入 r -凹离散概率分布的概念解决“具有离散随机变量的概率约束优化问题”; Ruszczyński [2]在 2002 年针对“具有概率约束的随机优化问题”进行研究, 这些约束涉及具有离散分布的随机变量, 提出解决概率约束优化问题的“有效不等式和一般迭代算法”; Ruszczyński 等人[3]在 2004 年, 引入了 p -有效点的概念后, 将其应用于得到带有概率约束的非线性随机优化问题的等价问题以及最优性的充分必要条件中, 从而提出了“概率约束问题的对偶方法”; Luedtke [4]在 2010 年针对联合概率约束线性规划的可行域非凸的问题, 提出一个“混合整数规划公式”。Ruszczyński 等人[5]又在 2012 年又提出了增广 Lagrange 方法。

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。因此其在求解约束优化问题中扮演着重要的角色。Hestenes [6]与 Powell [7]在 1969 年各自独立提出了近似增广 Lagrange 函数来求解具有等式约束的非线性优化问题。Boggs & Tolle [8]在 1980 年引入了一类新的乘子法, 发展了对偶理论, 对于具有不等式约束问题, Bertsekas [9]在 1982 年提出了非凸规划问题的修正的指数 Lagrange 函数, Polyak [10]在 1992 年给出了修正的 Frish 函数和修正的 Carroll 函数, 提出相应的对偶算法。鉴于非线性优化问题对偶算法的有效性, 本文基于修正的指数函数求解概率约束优化问题(1)的对偶算法。

2. 概率约束优化问题(1)的近似形式

定义 2.1 [3] 问题(1) Y 的分布函数为

$$F_Y(y) = \Pr[y \in \mathbb{R}^m : Y \leq y].$$

定义 2.2 [3] 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

$$\mathcal{E} = \{v^j \mid v^j \text{ 是 } p\text{-有效点}, j \in J\}$$

是 Y 的 p -有效点集, 其中 J 是一个任意集合。

对于 $p \in (0,1)$, 由于分布函数的单调性和右连续性易知, Y 的分布函数的水平集

$$\mathcal{L}_p = \{y \in \mathbb{R}^m : F_Y(y) \geq p\}$$

是非空闭集, 水平集 $\mathcal{L}_p$ 的凸包 $\text{co}\mathcal{L}_p$ 也是闭的[3]。

则问题(1)表示为

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

定义

$$K_j = v^j + \mathbb{R}_+^m, v^j \in \mathcal{E}, j \in J.$$

命题 2.1 [5] 水平集可表示为

$$\mathcal{L}_p = \bigcup_{j \in J} K_j.$$

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 $\text{co}\mathcal{E}$ 表示如下:

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

那么, $\mathcal{L}_p$ 的凸包 $\text{co}\mathcal{L}_p$ 也有下述表示。

命题 2.2 [5] 对于每个 $p \in (0,1)$, 有 $\text{co}\mathcal{L}_p = \text{co}\mathcal{E} + \mathbb{R}_+^m$ 。

对于任意 $p \in (0,1]$, $\mathcal{E}$, $\mathcal{L}_p$, $\text{co}\mathcal{E}$ 和 $\text{co}\mathcal{L}_p$ 有下界[3]。

$\text{co}\mathcal{L}_p$ 中的极点是 p -有效点。记 $\mathcal{E}_0$ 为 $\text{co}\mathcal{L}_p$ 的极点集, $\mathcal{E}_0$ 非空, 且它包含于 p -有效点集 $\mathcal{E}$ [5]。

引理 2.1 [5] 集合 $\text{co}\mathcal{L}_p$ 有如下表示:

$$\text{co}\mathcal{L}_p = \text{co}\mathcal{E}_0 + \mathbb{R}_+^m,$$

问题(2)中 $\mathcal{L}_p$ 凸化后, 问题(2)转化为下述问题

$$\begin{aligned} \min & f(x) \\ \text{s.t. } & g(x) \in \text{co}\mathcal{L}_p, \\ & x \in \mathcal{D}. \end{aligned} \quad (3)$$

3. 求解问题(3)的对偶算法

记问题(3)的可行域为

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

设 $K_{\mathcal{D}}(\hat{x})$, $N_{\mathcal{D}}(\hat{x})$, $T_{\mathcal{D}}(\hat{x})$ 分别表示 $\mathcal{D}$ 在 $\hat{x}$ 处的可行方向锥, 法锥, 切锥。

给定可行点 $\hat{x} \in \Phi$, 则存在 $\hat{\omega} \in \text{co}\mathcal{L}$, $\hat{y} \in \mathbb{R}_+^m$, 使得 $g(\hat{x}) = \hat{\omega} + \hat{y}$ 。记 $\hat{x}$ 处的积极约束指标集为

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

对于向量 $\delta \in K_{\mathcal{D}}(\hat{x})$, 一个 p -有效点 ω_i^0 的凸组合存在

$$\forall i \in I_0(\hat{x}), g_i(\hat{x}) + \langle \nabla g_i(\hat{x}), \delta \rangle > \omega_i^0. \quad (4)$$

条件(4)等价于问题(3)在可行点 $\hat{x}$ 处的 Robinson's 约束规范条件[5]。

定义 Lagrange 函数

$$L(x, u) = f(x) - \langle u, g(x) \rangle.$$

定理 3.1 [5] 设 $\hat{x}$ 是问题(3)的一个局部最优解, 且在 $\hat{x}$ 处满足条件(4), 则存在一个向量 $\hat{u} \in \mathbb{R}_+^m$, 使得

$$-\nabla_x L(\hat{x}, \hat{u}) \in N_{\mathcal{D}}(\hat{x}), \quad -\hat{u} \in N_{\text{co}\mathcal{L}_p}(g(\hat{x})), \quad (5)$$

记 $\Lambda(\hat{x})$ 为满足(5)中出现的 Lagrange 乘子的集合。记 $\mathcal{J}g(x)$ 为 g 在 x 处的 Jacobian。

定理 3.2 [5] 设 $\hat{x} \in \Phi$ 满足条件(4)和(5), 且对于 $u \in \Lambda(\hat{x})$, 以及使得

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

成立的每个非零的 $\delta \in T_{\mathcal{D}}(\hat{x})$, 有下述不等式成立,

$$\langle \delta, \nabla_x^2 L(\hat{x}, u)\delta \rangle > 0, \quad (6)$$

则 $\hat{x}$ 是问题(3)的一个局部极小值。

设 $I \subset J$ 是问题(3)的 p -有效点集中的极点集, $|I|$ 是它的基数, 我们考虑由问题(3)转化成的近似问题:

$$\begin{aligned} \min & f(x) \\ \text{s.t. } & g(x) \geq \sum_{j \in I} \lambda_j v^j, \\ & x \in \mathcal{D}, \lambda \in \mathcal{S}_{|I|}. \end{aligned} \quad (7)$$

当 $I = J$ 时, (7)与(3)重合。

考虑 Bertsekas (1982) [9]提出的修正的指数函数

$$F(x, u, k) = f_0(x) - k^{-1} \sum_{i=1}^m u_i \left[1 - e^{-kf_i(x)} \right];$$

其中, $k > 0$ 是惩罚参数, 则问题(2)的修正的指数函数

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

其中, $F_{\rho, |I|}: \mathbb{R}^m \times \mathcal{S}_{m+1} \times \mathbb{R}^m \rightarrow \mathbb{R}$, 子集 $I \subseteq J$, $\rho > 0$ 是惩罚参数。

对偶算法 下述算法中, 我们用 ε 作为一个准确参数

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

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

步骤 1: 求解主问题

$$\min_{x, \lambda} F_{\rho, k}(x, \lambda, u^k) \\ x \in \mathcal{D}, \lambda \in \mathcal{S}_k.$$

令 (x^k, λ^k) 是这个问题的解。

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

步骤 3: 计算

$$u_i^{k+1} = u_i^k e^{-\rho \left[g_i(x) - \sum_{j \in I} \lambda_j v_j^i \right]}, \quad i = 1, \dots, k.$$

步骤 4: 找到一个 p -有效点 v^{k+1} 作为问题 $\min_{z \in \text{co}\mathcal{L}_p} \langle u^{k+1}, z \rangle$ 的解, 令

$$J_{k+1} = \{v^{k+1}\} \cup J_k.$$

步骤 5: 令 $k \rightarrow k+1$, 然后转到步骤 1。

接下来, 做如下假设:

(A1) 存在一个数 $\kappa > 0$, 使得随机向量 Y 的两个实现间的距离大于 κ ;

(A2) 集合 $\mathcal{D}$ 凸且紧致。

定理 3.3 [5] 存在一个有限子集 $I \subseteq J$, 使得问题(2)与问题(3)的局部一阶最优性条件等价。

因此, 问题(3)在 $(\hat{x}, \hat{\lambda})$ 处关于 $\hat{u}$ 的一阶最优性条件是:

$$-\nabla_x L(\hat{x}, \hat{u}) \in N_{\mathcal{D}}(\hat{x}), \quad -v_i^T \hat{u} \in N_{\mathcal{S}_I}(\hat{\lambda}).$$

定理 3.4 在假设(A1)和(A2)成立的基础上, 假设问题(3)有可行域非空, 则参数 $\rho > 0$ 的修正的指数函数方法产生的序列 $\{x^k\}$ 和 $\{u^k\}$ 的聚点 $\hat{x}$ 和 $\hat{u}$, 满足(3)的一阶最优性条件, 此外, 若 $\hat{x}$ 满足条件(6)和(4), 那么 $\hat{x}$ 是(3)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 -有效点在步骤 4 中产生, 当 $\mathcal{D}$ 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

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

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 $\hat{\lambda} \in \mathcal{S}_{k_0}$ 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, 根据算法步骤 3, 对于 $i \notin J$, 有

$$u_i^{k+1} = u_i^k e^{-\rho \left[g_i(x) - \sum_{j \in I} \lambda_j^k v_j^i \right]}. \quad (8)$$

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, 修正的指数函数有如下形式:

$$F_{\rho,k}(x^k, \lambda^k, u^k) = f(x) - \rho^{-1} \sum_{i=1}^k u_i^k \left[1 - e^{-\rho(g_i(x) - z_i^k)} \right].$$

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

$$F_{\rho,k}(\tilde{x}, \tilde{\lambda}, u^k) = f(\tilde{x}) - \rho^{-1} \sum_{i=1}^k u_i^k \left[1 - e^{-\rho(g_i(\tilde{x}) - v_i^1)} \right],$$

所以

$$F_{\rho,k}(\tilde{x}, \tilde{\lambda}, u^k) \leq f(\tilde{x}) < F_{\rho,k}(x^k, \lambda^k, u^k).$$

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(8)式得到, $\{u^k\}$ 有一个聚点 $\hat{u}$, 使得对于所有 $i: \hat{u} > 0$, 有 $\hat{z}_i = g_i(\hat{x})$ 。

在步骤 1 中主问题的最优性条件是:

$$-\nabla f(x^k) + \sum_{i=1}^k u_i^k e^{-\rho(g_i(x) - \sum_{j \in I} \lambda_j^k v_j^1)} \nabla g_i(x^k) \in N_{\mathcal{D}}(x^k) \quad (9)$$

$$-V_{J_0}^T u^k \in N_{\mathcal{S}_{J_0}}(\lambda^k) \quad (10)$$

对于足够大的 $k \in \mathcal{K}$, 我们可以改写(9)式为

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

条件(10)意味着

$$0 \geq \langle -u^k, V_{J_0}(\lambda - \lambda^k) \rangle = \langle -u^k, V_{J_0} \lambda - z^k \rangle \quad (12)$$

假设存在一点 $\omega \in \text{co}\mathcal{L}_p \setminus \text{co}\{v^j, j \in J_k\}$ 使得 $\langle u^k, \omega - z^k \rangle < 0$, 应用这个假设和(12), 我们得到

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

即不等式 $\langle -u^k, \omega - v^k \rangle > 0$, 这与步骤四中, 问题 $\min_{z \in \text{co}\mathcal{L}_p} \langle u^{k+1}, z \rangle$ 的最优性条件 $-u^k \in N_{\text{co}\mathcal{L}_p}(z)$ 是矛盾的, 这说明 $-u^k \in N_{\text{co}\mathcal{L}_p}(v^k)$ 。因此, $-u^k \in N_{\text{co}\mathcal{L}_p}(z^k)$, 结合(11), 表明

$$-\nabla F(x^k, u^{k+1}) \in N_{\mathcal{D}}(x^k) \quad (13)$$

$$-u^{k+1} \in N_{\text{co}\mathcal{L}_p}(z^{k+1}) \quad (14)$$

在 $k \in \mathcal{K}$ 上取极限, 并应用(06 年的文献)到(13)~(14), 我们得到

$$-\nabla F(\hat{x}, \hat{u}) \in N_{\mathcal{D}}(\hat{x}), \quad -\hat{u} \in N_{\text{co}\mathcal{L}_p}(g(\hat{x})),$$

因此, $\hat{x}$ 满足问题(3)的条件(5), 根据定理 4, 点 $\hat{x}$ 是(3)的局部极小解。□

本文通过运用在 p -有效点的概念下产生的基于修正的指数函数的对偶算法来解决概率约束问题, 目前该方法仍存在不完善之处, 在后续的研究中将进一步探索其可应用的范围, 以及在实际问题中的运用, 望能更好地为社会创造价值。

参考文献

- [1] Dentcheva, D., Prékopa, A. and Ruszczyński, A. (2000) Concavity and Efficient Points of Discrete Distributions in

-
- Probabilistic Programming. *Mathematical Programming*, **89**, 55-77. <https://doi.org/10.1007/pl00011393>
- [2] Dentcheva, D., Prékopa, A. and Ruszczyński, A. (2002) Bounds for Probabilistic Integer Programming Problems. *Discrete Applied Mathematics*, **124**, 55-65. [https://doi.org/10.1016/s0166-218x\(01\)00329-8](https://doi.org/10.1016/s0166-218x(01)00329-8)
 - [3] 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>
 - [4] Luedtke, J., Ahmed, S. and Nemhauser, G.L. (2008) An Integer Programming Approach for Linear Programs with Probabilistic Constraints. *Mathematical Programming*, **122**, 247-272. <https://doi.org/10.1007/s10107-008-0247-4>
 - [5] 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>
 - [6] Hestenes, M.R. (1969) Multiplier and gradient methods. *Journal of Optimization Theory and Applications*, **4**, 303-320. <https://doi.org/10.1007/bf00927673>
 - [7] Powell, M.J.D. (1969) A Method for Nonlinear Constraints in Minimization Problems. In: Fletcher, R. Ed., *Optimization*, Academic Press, 283-298.
 - [8] Boggs, P.T. and Tolle, J.W. (1980) Augmented Lagrangians Which Are Quadratic in the Multiplier. *Journal of Optimization Theory and Applications*, **31**, 17-26. <https://doi.org/10.1007/bf00934785>
 - [9] Bertsekas, D.P. (1982) *Constrained Optimization and Lagrange Multiplier Methods*. Academic Press.
 - [10] Polyak, R. (1992) Modified Barrier Functions (Theory and Methods). *Mathematical Programming*, **54**, 177-222. <https://doi.org/10.1007/bf01586050>