Statistical Inference of Two-Stage Linear Second-Order Conic Stochastic Programs

Qingsong Duan, Liwei Zhang

School of Mathematical Sciences, Dalian University of Technology, Dalian Liaoning
Email: duanqs@mail.dlut.edu.cn, lwzhang@dlut.edu.cn

Received: Jun. 24th, 2018; accepted: Jul. 13th, 2018; published: Jul. 20th, 2018

Abstract

In this paper, we consider a linear second-order conic optimization problem and in which all parameters are perturbed and random variables. We demonstrate that the optimal value function can be expressed as a min-max optimization problem over compact convex set, and we present the asymptotic distribution of an SAA estimator of the optimal value for a two-stage program whose second stage problem is a second-order conic programming problem.

Keywords

Second-Order Conic Optimization, Two-Stage Stochastic Problem, SAA Estimator Asymptotic Distribution

线性二阶锥两阶段随机规划问题的统计推断

段庆松，张立卫

大连理工大学，数学科学学院，辽宁 大连
Email: duanqs@mail.dlut.edu.cn, lwzhang@dlut.edu.cn

收稿日期：2018年6月24日；录用日期：2018年7月13日；发布日期：2018年7月20日

摘 要

在本篇文章中，我们考虑一类带有线性二阶锥约束的两阶段随机规划问题，该问题的全部参数都是随机变量。我们将原问题的最优值函数改写为一个包含紧致凸约束集合的极小极大问题，利用第二阶段问题的 Lagrange对偶性质，得到其最优值函数的样本均值近似(SAA)估计的渐近分布。
关键词
二阶锥优化，两阶段随机规划问题，SAA估计的渐近分布

1. 引言

经典的线性两阶段优化问题的模型如下:

$$\begin{align*}
\min_{x \in \mathbb{R}^n} & \quad d^Tx + E\left[\left(\theta(x, \xi)\right)^T\right] \\
\text{s.t.} & \quad Ax = b, x \geq 0,
\end{align*}$$

其中 $x \in \mathbb{R}^n, y \in \mathbb{R}^m$ 分别是第一阶段和第二阶段问题的决策变量，$d \in \mathbb{R}^n, \xi = \{c, W, T, h\}$ 是随机变量，其中 $c \in \mathbb{R}^n, W \in \mathbb{R}^{n \times m}, T \in \mathbb{R}^{m \times n}, h \in \mathbb{R}^l$. 当 $\xi$ 服从的概率分布进行扰动时，$\theta(x, \xi)$ 的连续性在研究线性二阶锥问题的稳定性分析中起着关键的作用。很多文章研究过包含线性二阶锥约束的优化问题，然而所有变量都是随机的情况很少被考虑，只有少数的例外：Römisch 和 Wets 在[1]中得到了最优值函数 $\theta(x, \xi)$ 的 Lipschitz 连续性，Han 和 Chen 在[2]中研究了所含参数全是随机变量的线性规划的连续性质。

在本文中，我们研究带有线性二阶锥约束的两阶段随机规划问题，即 $\theta(x, \xi)$ 是如下问题的最优值函数:

$$\begin{align*}
\min_{y \in \mathbb{R}^m} & \quad c^Ty \\
\text{s.t.} & \quad a_i^Ty + q_i^Tx - b_i \geq \|B'_i\|_2, i = 1, \ldots, l \quad (1.2)
\end{align*}$$

学者们对于此问题的最优值函数和最优解集映射进行了定性和定量的稳定性分析[3]，并研究了其扰动问题和对偶问题最优解集的上半连续性以及最优解集的 Hadamard 方向可微性[4]。在此基础上，利用该问题的样本均值近似（SAA）估计的渐近分布来研究其统计推断。

SAA 方法是研究一般随机规划问题的常用方法。考虑问题

$$\min_{x \in \mathbb{R}^n} \mathbb{E}[F(x, \xi)]$$

其中 $X$ 是 $\mathbb{R}^n$ 空间的一个非空闭子集，函数 $F : X \times \Xi \to \mathbb{R}^l$, $\xi$ 是概率分布 $P$ 在支撑集合 $\Xi \subseteq \mathbb{R}^l$ 上产生的随机变量。如果目标函数 $F(x, \xi)$ 是某个问题的最优值函数，那么问题(1.3)就可以被考虑为一个两阶段随机规划问题的数学模型。
机规划问题。我们假设期望值函数 $f(x)$ 是有定义的，并且对于所有 $x \in X$，是有限的，即对于每一个 $x \in X, \xi \in \Xi$，最优值 $F(x, \xi)$ 取得有限值。SAA 方法的思想是选取随机变量 $\xi$ 的 $N$ 个样本值得到的 $N$ 个历史观察值。对于 $x \in X$，期望值函数 $f(x)$ 可以由 $F(x, \xi), j = 1, \ldots, N$ 的平均值估计得出。那么问题 (1.3) 可以被如下问题近似：

$$
\min_{x \in X} \left\{ \frac{1}{N} \sum_{j=1}^{N} F(x, \xi^j) \right\}.
$$


本篇文章中，我们给出了问题 (1.2) 的 Lagrange 对偶问题，从而将原始问题的最优值函数改写成一个极大极小问题。利用最优值函数的 Hadamard 方向可微性和 Delta 定理得到了线性二阶锥两阶段随机规划问题最优值函数的 SAA 估计的近似分布。

### 2. 最优值函数的统计推断

首先分析线性二阶锥两阶段随机规划问题的扰动问题的性质。给定参数 $u = (c, A, Q, B, b, x)$，当其收敛到 $\bar{u} = (\bar{c}, \bar{A}, \bar{Q}, \bar{B}, \bar{b}, \bar{x})$ 时，我们分析最优值函数 $f(\cdot, \cdot)$ 的性质。其扰动问题 $P(\bar{x}, \bar{\xi})$ 可以表示为如下形式：

$$
\min_{y} \bar{c}^T y \\
\text{s.t. } g(y, \bar{x}, \bar{\xi}) \in \Theta
$$

其中 $g(y, \bar{x}, \bar{\xi}) = (g^1(y, \bar{x}, \bar{\xi}), \cdots, g^l(y, \bar{x}, \bar{\xi}))$, $\Theta = \Theta_{i, \alpha} \times \cdots \times \Theta_{i, \alpha}$。设 $f(y, \bar{u}) = \bar{c}^T y$。令 $\Phi(\bar{u})$ 表示问题 (2.1) 的可行集合，记为

$$
\Phi(\bar{u}) = \left\{ y \in \mathbb{R}^m : g_i(y, \bar{x}, \bar{\xi}) \in \Theta_{i, \alpha}, i = 1, \ldots, l \right\}.
$$

同时令 $Y^*(\bar{u})$ 表示问题 (2.1) 的最优解集。定义

$$
\Psi(\bar{u}, \alpha) = \Phi(\bar{u}) \cap \text{lev}_{\text{stat}} f(\cdot, \bar{u})
$$

其中

$$
\text{lev}_{\text{stat}} f(\cdot, \bar{u}) = \left\{ y \in \mathbb{R}^m : f(y, \bar{u}) \leq \alpha \right\}, \alpha \in \mathbb{R}.
$$

我们可以推导出线性二阶锥优化问题 (2.1) 的 Lagrange 对偶问题。问题 (2.1) 的 Lagrange 函数可以写为

$$
L(y, \bar{x}; \bar{u}) = \bar{c}^T y - \sum_{i=1}^{l} \xi^T(x - \bar{\xi}^i)
$$

其中 $\bar{\xi}^i = \left( \bar{\xi}^1, \cdots, \bar{\xi}^i \right)$，线性算子 $\bar{\Lambda} : \mathbb{R}^m \rightarrow \mathbb{R}^{l \times l} \times \cdots \times \mathbb{R}^{l \times l}$ 定义如下

$$
\bar{\Lambda} = (\bar{B}^T y, \bar{a}^T y; \cdots; \bar{B}^T y, \bar{a}^T y).
$$

则问题 (2.1) 的 Lagrange 对偶可写为

$$
\max \sum_{i=1}^{l} \xi^T(x - \bar{\xi}^i)
$$

s.t. $\bar{c} - \bar{\Lambda}^T \bar{\xi} = 0$

$$
\bar{\xi} \in \Theta
$$

DOI: 10.12677/aam.2018.77105 878 应用数学进展
段庆松，张立卫

这里 $\tilde{A}^*$ 是 $\tilde{A}$ 的伴随矩阵且 $\tilde{A}^* \tilde{A}$ 可由如下的形式计算

$$\tilde{A}^* \tilde{A} = \sum_{i=1}^{l} [\tilde{B}^T, \tilde{a}] \tilde{a}^i.$$ 

令 $\phi(\tilde{A}, \tilde{u})=\sum_{i=1}^{l} \tilde{a}^i (\tilde{b}_i - q^T x)$ 和 $\Lambda^*(\tilde{u})$ 分别表示问题(2.2)最优值函数以及最优解。问题(2.2)的可行集合可以由

$$\{ \tilde{x} = (\tilde{x}^1, \ldots, \tilde{x}^l) \in \Theta : \tilde{x} - \tilde{A}^* \tilde{A} = 0 \}$$

表示。进一步，可以定义

$$\Gamma(\tilde{u}, \alpha) = \{ \tilde{x} = (\tilde{x}^1, \ldots, \tilde{x}^l) \in \Theta : \tilde{x} - \tilde{A}^* \tilde{A} = 0 \}$$

其中

$$lev_{\alpha, \phi}(\tilde{u}, \alpha) = \{ \tilde{x} \in \mathbb{R}^{k+1} \times \cdots \times \mathbb{R}^{l+1} : \phi(\tilde{A}, \tilde{u}) \leq \alpha \}, \alpha \in \mathbb{R}.$$ 

由引理 3.3 和[4]中的引理 2.3 可知，原问题和对偶问题的可行集是水平有界的，即对于某些，$\delta > 0, \alpha \in \mathbb{R}, i \in X$，$\|x_l^p - (c, A, Q, B, b)\| \leq \delta$ 以及有界集合 $B \subset \mathbb{R}^n, H \subset \mathbb{R}^{k+1} \times \cdots \times \mathbb{R}^{l+1}$，有

$$\Psi(\tilde{u}, \alpha) \subset B, \Gamma(\tilde{u}, \alpha) \subset H.$$ 

则由 Lagrange 对偶定理，扰动问题的最优值可以写成下面的极大极小问题：

$$\theta(\tilde{c}, \tilde{A}, \tilde{Q}, \tilde{B}, \tilde{x}) = \max_{\lambda \in \mathbb{R}^{l+1}} \min_{y \in \Theta_{\tilde{A}, \tilde{u}}} L(y, \lambda; \tilde{u}).$$

为了进一步的研究，我们给出关于问题(1.2)和(2.1)的三个假设。

假设 1：$X \subset \mathbb{R}^n$ 是一个非空紧致凸集。

假设 2：对于每个 $x \in X$，问题(1.2)的最优值函数是有限的，且问题(1.2)的最优解集是紧致的。

假设 3：对于每个 $x \in X$，问题(1.2)的 Slater 条件成立，即对于每个 $x \in \mathbb{R}^n$，存在 $y_x$使得

$$g_i^i(y_x, x, \xi) \in \text{int} \, \Theta_{\tilde{A}, \tilde{u}}, i = 1, \ldots, l.$$ 

上式也等价为 $a_i^T y_x + q_i^T x - b_i > \|B^T y_x\|, i = 1, \ldots, l$。

如果假设 3 成立，则对偶问题(2.2)有一个非空子集且问题(1.2)和其对偶问题(2.2)的对偶间隙为 0。文献[4]证明了，原问题和对偶问题的可行集是凸集，假设 1～3 成立时，扰动问题和其对偶问题的 Slater 条件在 $\xi$ 附近都成立，而且两个问题的最优解都存在上半连续的。文章[3]在同样的假设下，得到了关于扰动问题的最优值函数和最优解集映射的定性和定量的稳定性分析结果。

**引理 1**：给定 $(c, A, Q, B, b)$ 和 $x \in X$，若假设 1 ～ 3 成立，则最优值函数 $\theta(\tilde{c}, \tilde{A}, \tilde{Q}, \tilde{B}, \tilde{x})$ 在点 $(c, A, Q, B, b, x)$ 处是 Hadamard 方向可微的。因此我们可以得到 $\theta(u)$ 在 $u$ 处的 Taylor 展开式为

$$\theta(u) = \theta(\tilde{u}) + \inf_{y \in X(\tilde{u})} \sup_{\lambda \in \Lambda_{\tilde{u}}} \sum_{i=1}^{l} \lambda^i (\Delta b_i - \lambda q^T x - q^T \Delta x) + o(\|\Delta u\|),$$

其中 $u = (\tilde{c}, \tilde{A}, \tilde{Q}, \tilde{B}, \tilde{x})$，$\Delta u = \tilde{u} - u$，且 $\|\Delta u\| \leq \delta$。

现在我们考虑第二阶段问题(1.2)的最优值函数的渐近性质，其最优解函数，记为 $\theta(x, \xi)$，其中随机变量 $\xi = (\tilde{c}, \tilde{A}, \tilde{Q}, \tilde{B}, \tilde{x})$。令

$$f(x, \xi) = d^T x + \theta(x, \xi),$$

DOI: 10.12677/aam.2018.77105 879 应用数学进展
则两阶段随机规划问题可以表示为

$$\min \ E[f(x, \xi)]$$

s.t.  \( x \in X \)

设\( \xi^1, \cdots, \xi^N \)是独立同分布的样本，则样本均值近似问题为

$$\min \ f_N(x)$$

s.t.  \( x \in X \)

其中

$$f_N(x) = f\left(x, \frac{1}{N} \sum_{i=1}^{N} \xi^i\right) = d^T x + \theta\left(x, \frac{1}{N} \sum_{i=1}^{N} \xi^i\right).$$

我们用\( \nu \)和\( S^\star \)分别表示原问题(2.3)的最优值和最优解集，同时令\( \hat{\nu}_N, \hat{S}_N \)表示 SAA 问题(2.4)的最优值和最优解集。

假设：假设 \{\( c_A, Q, B, b \)\} 中任意两个元素是相互独立的，\( \xi=(\hat{c}, \hat{A}, \hat{Q}, \hat{B}, \hat{b}) \)的期望值是 \( \mu=(c, A, Q, B, b) \)，即 \( E(\xi) = \mu \)。设\( \xi^1, \cdots, \xi^N \)为独立随机样本，对于\( \xi^i=(\hat{c}^i, \hat{A}^i, \hat{Q}^i, \hat{B}^i, \hat{b}^i) \)，\( i=1, \cdots, N \)以及

$$\hat{\xi}_N = (\hat{c}_N, \hat{A}_N, \hat{Q}_N, \hat{B}_N, \hat{b}_N) = \frac{1}{N} \sum_{i=1}^{N} (\xi^i, \hat{A}^i, \hat{Q}^i, \hat{B}^i, \hat{b}^i).$$

假设

$$\sqrt{N} [\hat{\xi}_N - c] \xrightarrow{d} N(0, \Sigma_c),$$

$$\sqrt{N} [\hat{a}_N - a_i] \xrightarrow{d} N(0, \Sigma_{a_i}), i=1, \cdots, l$$

$$\sqrt{N} [\hat{q}_N - q_i] \xrightarrow{d} N(0, \Sigma_{q_i}), i=1, \cdots, l$$

$$\sqrt{N} [\hat{B}_N - B_i] \xrightarrow{d} N(0, \Sigma_{B_i}), i=1, \cdots, l$$

$$\sqrt{N} [\hat{b}_N - b_i] \xrightarrow{d} N(0, \Sigma_{b_i}), i=1, \cdots, l$$

这里 \( \xrightarrow{d} \) 表示依分布收敛。则

$$\sqrt{N} [\hat{E}_N - E(\xi)] = \sqrt{N} [\hat{E}_N - E(\xi)] \xrightarrow{d} N(0, \Sigma_{\nu}), i=1, \cdots, l,$$

这里 \( \Sigma_{\nu} \) 能被 \( \Sigma_{\nu}^a \) 和 \( \Sigma_{\nu}^b \) 表示。

我们通过 Delta 引理来分析 SAA 最优值 \( \nu_N \) 的一阶渐近性质。

引理 2：((9) Theorem 7.59) 设 \( B_1 \) 和 \( B_2 \) 为 Banach 空间，\( Z_N \) 为 \( B_1 \) 中随机元素序列，\( G \) 是一个从 \( B_1 \) 到 \( B_2 \) 的映射。1）若空间 \( B_1 \) 是可分的，2）映射 \( G \) 在点 \( \mu \in B_1 \) 处是 Hadamard 方向可微的，3）当 \( N \to \infty \) 时，序列 \( \tau_N \) 趋于无穷，序列 \( X_N = \tau_N (Z_N - \mu) \) 依分布收敛于 \( B_1 \) 中的随机元素 \( Z \) 时，则有

$$\tau_N [G(Z_N) - G(\mu)] \xrightarrow{d} G' (\mu; Z).$$

现在我们提出研究 SAA 问题最优值 \( \nu_N \) 渐近性质的主要定理。

定理 1：若假设 4 成立，有

$$N^{1/2} (\hat{\theta}_N - \theta') \xrightarrow{d} \inf_{\beta \in \nabla} \inf_{\alpha \in \Delta} \sup_{x \in \mathcal{L}^\star} \{V(x, y, \lambda)\},$$  (2.5)
这里 $V(x,y,\lambda)$ 是依赖于 $(x,y,\lambda)$ 的随机变量：

$$V(x,y,\lambda) \sim N\left(0, y^T \Sigma y - \sum_{i=1}^L \lambda_{yi}^2 y_i^T \Sigma y + \sum_{i=1}^L \lambda_{yi}^2 \left[ \Sigma y_i^T - x_i^T \Sigma y_0 x \right] \right).$$

更进一步的，如果 $S^* = \{x\}, Y^*(\bar{p}) = \{y\}, \Lambda^*(\bar{p}) = \{z\}$，我们可以得到

$$N^{[2]} \left( \tilde{V}_n - \nu \right) \to N \left(0, \bar{p}^T \Sigma \bar{p} - \sum_{i=1}^L \bar{z}_{yi}^2 \bar{y}_i^T \Sigma \bar{y} + \sum_{i=1}^L \bar{z}_{yi}^2 \left[ \Sigma \bar{y}_i^T - \bar{x}_i^T \Sigma \bar{y}_0 \right] \right). \quad (2.6)$$

证明：首先我们用引理 2 分析 $N^{[2]} \left( \tilde{f}_n(x) - E (f(x,\xi)) \right)$ 的渐近性质。

设 $B \times \mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R} \times \mathbb{R}$，$B \subset C(X)$ 和 $G : B \to B$

$$G(\xi) = d^T x + \theta(x,\xi)$$

设 $r_n = N^{[2]} Z_n = \hat{z}_n, \mu = \bar{p}$，则可以由假设 4 得到 $r_n (Z_n - \mu) \to Z$，其中

$$Z^e \sim N(0, \Sigma^e),$$
$$Z_i^e \sim N(0, \Sigma_{ij}^e), i = 1, \ldots, l,$$
$$Z^b \sim N(0, \Sigma^b), i = 1, \ldots, l.$$ \quad (2.7)

从而引理 2 中的 1), 2) 可以由假设 4 得出。由引理 1 可知 $G$ 在 $\mu$ 处是 Hadamard 方向可微，从而满足引理 2 中的 3)。因此，我们可以由引理 2 得到

$$N^{[2]} \left[ f(x,\tilde{z}_n) - f(x,\xi) \right] \to G^{*} (\mu, Z).$$

注意到 $G^{*} (\mu, Z) = G^{*} (\mu, Z)$，则由引理 1 可以得到

$$G^{*} (\mu, Z) = \inf \sup \left\{ Z^e - Z_i^e \lambda^e \right\} \left[ \sum_{i=1}^L \lambda_{yi}^2 y_i + \sum_{i=1}^L \lambda_{yi}^2 \left( Z_i^b - Z_i^e \right) \right].$$

设 $V(x,y,\lambda) = \left( Z^e y - \sum_{i=1}^L \lambda_{yi}^2 Z_i^e y_i + \sum_{i=1}^L \lambda_{yi}^2 \left( Z_i^b - Z_i^e \right) \right)$ 则由式(2.7)可以得到

$$V(x,y,\lambda) \sim N \left(0, y^T \Sigma y - \sum_{i=1}^L \lambda_{yi}^2 y_i^T \Sigma y_i + \sum_{i=1}^L \lambda_{yi}^2 \left[ \Sigma y_i^T - x_i^T \Sigma y_0 x \right] \right).$$

等式(2.5)可以由文章[10]中定理 5.7 直接得到。自然地，当 $S^* = \{x\}, Y^*(\bar{p}) = \{y\}, \Lambda^*(\bar{p}) = \{z\}$ 时，我们可以由(2.5)得到(2.6)，证明完毕。

3. 结论

本文基于线性二阶锥两阶段随机规划问题和其对偶问题最优解集的上半连续性，以及最优值函数的 Hadamard 方向可微性，研究了最优值函数的统计推断。利用 Delta 定理，我们推导出所有变量都是随机变量的两阶段随机规划问题的 SAA 估计的渐近分布。

参考文献


Optimization Letters, 9, 1075-1090. https://doi.org/10.1007/s11590-014-0827-6


知网检索的两种方式:

1. 打开知网页面 http://kns.cnki.net/kns/brief/result.aspx?dbPrefix=WWJD
   下拉列表框选择：[ISSN]，输入期刊 ISSN: 2324-7991，即可查询
2. 打开知网首页 http://cnki.net/
   左侧“国际文献总库”进入，输入文章标题，即可查询

投稿请点击：http://www.hanspub.org/Submission.aspx
期刊邮箱：aam@hanspub.org

DOI: 10.12677/aam.2018.77105