Operations Research Letters 46 (2018) 543–547

Contents lists available at ScienceDirect

Operations Research Letters

journal homepage: www.elsevier.com/locate/orl

Quantitative stability of multistage stochastic programs via calm modifications

Jie Jiang, Zhiping Chen *

School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an, China

ARTICLE INFO

Article history:
Received 26 April 2018
Accepted 30 August 2018
Available online 24 September 2018

Keywords:
Multistage stochastic programming
Calm modification
Filtration distance
Quantitative stability

ABSTRACT

In this paper, we revisit the quantitative stability of multistage stochastic programs. Different from the single calm modification used in Küchler (2008), we introduce two types of calm modifications which leads to a much simpler proof and tighter upper bound for the difference of optimal values of multistage stochastic programs under different stochastic processes than those of Küchler (2008). In addition, we avoid those restrictive assumptions in Küchler (2008) and the filtration distance in Heitsch et al. (2006). Finally, we illustrate our results with two numerical examples.

1. Introduction

We consider the following T-stage ($T \geq 2$) stochastic linear program (see [6,12]):

$$\begin{aligned} \inf_{x_1 \in D_1} \langle c_1, x_1 \rangle + \mathbb{E}\Bigg[\inf_{\boldsymbol{x}_2 \in D_2(x_1, \boldsymbol{\xi}_2)} \langle c_2(\boldsymbol{\xi}_2), \boldsymbol{x}_2 \rangle \\ + \mathbb{E}\Big[\inf_{\boldsymbol{x}_3 \in D_3(\boldsymbol{x}_2, \boldsymbol{\xi}_3)} \langle c_3(\boldsymbol{\xi}_3), \boldsymbol{x}_3 \rangle \\ + \cdots + \mathbb{E}\Big[\inf_{\boldsymbol{x}_T \in D_T(\boldsymbol{x}_{T-1}, \boldsymbol{\xi}_T)} \langle c_T(\boldsymbol{\xi}_T), \boldsymbol{x}_T \rangle \Big] \Big] \Bigg], \end{aligned} \tag{1}$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product in finite dimensional Hilbert space. $\boldsymbol{\xi} = (\boldsymbol{\xi}_t)_{t=1}^T \in \mathcal{L}_T(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^{sT})$ is a $\mathbb{R}^{sT}$-valued stochastic process defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, with finite Tth order absolute moments, and $\boldsymbol{x} = (\boldsymbol{x}_t)_{t=1}^T$ is the sequence of decision variables. We use bold letters, for example $\boldsymbol{\xi}$ or $\boldsymbol{x}$, to denote random vectors in contrast to their realizations ξ or x. The corresponding filtration of $\boldsymbol{\xi}$ is $\{\mathcal{F}_t\}_{t=1}^T$, defined by $\mathcal{F}_t = \sigma(\boldsymbol{\xi}^t)$ for $t = 1, 2 \ldots, T$, here $\boldsymbol{\xi}^t := (\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \ldots, \boldsymbol{\xi}_t)$ and especially $\boldsymbol{\xi}^T = \boldsymbol{\xi}$. The similar notation is adopted for other variables. Of course, we have that $\{\emptyset, \Omega\} = \mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \cdots \subseteq \mathcal{F}_T = \mathcal{F}$. Specially, $\mathcal{F}_1 = \{\emptyset, \Omega\}$ indicates that $\boldsymbol{\xi}_1 = \xi_1$ is deterministic. For $t = 1, 2, \ldots, T$, we use Ξ_t and Ξ^t to denote the support sets of $\boldsymbol{\xi}_t$ and $\boldsymbol{\xi}^t$, respectively. The corresponding probability measures are denoted by $\mathbb{P}_t$ and $\mathbb{P}^t$. Cost vectors $c_1 \in \mathbb{R}^n$ and $c_t : \Xi_t \to \mathbb{R}^n, t = 2, \ldots, T$, are affinely linear mappings with respect to ξ_t. $D_1 \subseteq \mathbb{R}^n$ and the feasible solution multifunctions $D_t : X_{t-1} \times \Xi_t \rightrightarrows \mathbb{R}^n$ are defined by

$$D_t(x_{t-1}, \xi_t) = \{x_t \in X_t \subseteq \mathbb{R}^n : A_t x_t + B_t(\xi_t) x_{t-1} = h_t(\xi_t)\}, \tag{2}$$

where X_t are nonempty polyhedral sets; the recourse matrix $A_t \in \mathbb{R}^{m \times n}$, the technology matrix $B_t : \Xi_t \to \mathbb{R}^{m \times n}$ and the right-hand side vector $h_t : \Xi_t \to \mathbb{R}^m$ are also affinely linear mappings with respect to ξ_t for $t = 2, 3, \ldots, T$. The affine linearity of $c(\xi) := (c_1(\xi_1), c_2(\xi_2), \ldots, c_T(\xi_T))$, means that

$$\|c(\xi) - c(\hat{\xi})\| \leq K \|\xi - \hat{\xi}\|, \;\; \|c(\xi)\| \leq K \max\{1, \|\xi\|\}$$

hold for some $K \geq 1$ and any $\xi, \hat{\xi} \in \Xi^T$.

In the last decade, the stability analysis of multistage stochastic programs has been investigated in a number of works, see, for example, [2,4,5,10,12] and the references therein. The quantitative stability results have a significant impact on suitable methods for approximating the underlying continuous data process, which in return make it possible for us to solve original multistage stochastic programs by solving large scale deterministic optimization problems.

In early work [3], the authors studied the quantitative stability by assuming implicitly that filtrations of the original and approximate stochastic processes are consistent. To extend this result to a general situation where the filtration is also perturbed, Römisch and his coauthors employed the so-called filtration distance to describe the variation of filtrations in [4]. The main reason for introducing the filtration distance was that they adopted the feasible solutions in the α-level set to describe the optimal values under different stochastic processes. Then, Eichhorn and Römisch extended in [2] the risk-neutral result in [4] to the risk-averse case with polyhedral risk measures introduced in [1]. Considering

* Corresponding author.
E-mail addresses: jiangjiecq@stu.xjtu.edu.cn (J. Jiang), zchen@xjtu.edu.cn (Z. Chen).

https://doi.org/10.1016/j.orl.2018.08.007

the computational difficulty of filtration distances, Küchler used in [5] some strong recursive assumptions to avoid the filtration distance, and obtained the quantitative stability assertion for a class of measurable perturbations. For more information on this topic, we refer readers to [7,10,11] and the references therein.

In this paper, we aim at simplifying the proof and strengthening the quantitative stability conclusion in [5] by adopting two types of calm modifications. Our quantitative stability result also avoids the tricky filtration distance. For these purposes, we adopt the following notation in what follows. $\|\xi^t\| := \max_{1\le i\le t}\|\xi_i\|$, here $\|\xi_i\|$ is the Euclidean norm in $\mathbb{R}^s$ for $1\le i\le t$ and $t=1,2,\ldots,T$. Analogously, we define $\|x^t\| = \max_{1\le i\le t}\|x_i\|$ and $\|x_i\|$ is the Euclidean norm in $\mathbb{R}^n$ for $1\le i\le t$ and $t=1,2,\ldots,T$. For sets $S_1,S_2\subseteq\mathbb{R}^n$, $d(S_1,S_2) := \sup_{s_1\in S_1} d(s_1,S_2)$, here $d(s_1,S_2) := \inf_{s_2\in S_2}\|s_1-s_2\|$ and $d_H(S_1,S_2) := \max\{d(S_1,S_2),d(S_2,S_1)\}$. $\upsilon(\boldsymbol{\xi})$ denote the optimal value of problem (1) under the stochastic process $\boldsymbol{\xi}$.

We need the following Lipschitzian results about D_t, $t=2,\ldots,T$, which can be found in [8, Example 9.35].

Proposition 1.1. *For $D_t(x_{t-1},\xi_t)$, $t=2,\ldots,T$, defined in (2), the following assertions hold:*

$$d_H(D_t(x_{t-1},\xi_t),D_t(\hat{x}_{t-1},\xi_t)) \le B\max\{1,\|\xi_t\|\}\left\|\hat{x}_{t-1}-x_{t-1}\right\|,$$

$$d_H(D_t(x_{t-1},\xi_t),D_t(x_{t-1},\hat{\xi}_t)) \le B\max\{1,\|x_{t-1}\|\}\|\hat{\xi}_t-\xi_t\|$$

for some constant $B>0$.

If we define $F(\boldsymbol{x},\boldsymbol{\xi}) = \sum_{t=1}^T \langle c_t(\boldsymbol{\xi}_t),\boldsymbol{x}_t\rangle$, model (1) can be equivalently rewritten as (see, for example, [4,5,8])

$$\min\{\mathbb{E}[F(\boldsymbol{x},\boldsymbol{\xi})] : \boldsymbol{x}\in\mathcal{D}(\boldsymbol{\xi})\},$$

where $\mathcal{D}(\boldsymbol{\xi})$ is a collection of decision processes $\boldsymbol{x}=(\boldsymbol{x}_1,\boldsymbol{x}_2,\ldots,\boldsymbol{x}_T)$ with $x_1\in D_1$ and measurable mappings $\boldsymbol{x}_t\in D_t(\boldsymbol{x}_{t-1},\boldsymbol{\xi}_t)$ for $t=2,\ldots,T$.

Another way to reformulate the multistage stochastic linear program (1) is the dynamic programming method. Concretely, let $Q_t : X_{t-1}\times\Xi^t\to\mathbb{R}$ denote the recourse function at the tth stage, which is defined recursively by

$$Q_t(x_{t-1},\xi^t) = \inf_{x_t\in D_t(x_{t-1},\xi_t)} \langle c_t(\xi_t),x_t\rangle + \mathbb{E}\left[Q_{t+1}(x_t,\boldsymbol{\xi}^{t+1})|\boldsymbol{\xi}^t=\xi^t\right] \quad (3)$$

for $t=T,T-1,\ldots,1$, here $Q_{T+1}:=0$ and $x_0:=1$. Then problem (1) is equivalent to

$$\min_{x_1\in D_1}\langle c_1,x_1\rangle + \mathbb{E}\left[Q_2(x_1,\boldsymbol{\xi}^2)\right].$$

Of particular interest in this paper, we consider the following measurable perturbation of $\boldsymbol{\xi}$.

Definition 1.2 (*Approximation of Stochastic Process, [5]*). A stochastic process $\tilde{\boldsymbol{\xi}}$ on the probability space $(\Omega,\mathcal{F},\mathbb{P})$ is called an approximation of $\boldsymbol{\xi}$, if there exist measurable mappings $f_t:\Xi^t\to\Xi_t$ for $t=1,2,\ldots,T$, such that the following conditions are satisfied:
(a) $\tilde{\boldsymbol{\xi}}_t = f_t(\boldsymbol{\xi}^t)$ for $t=1,2,\ldots,T$;
(b) $f^T(\Xi^T)\subseteq\Xi^T$;
(c) $f_1(\xi_1)=\xi_1$ for every $\xi_1\in\Xi_1$;
(d) $f^T(\boldsymbol{\xi}^T)\in\mathcal{L}_T(\Omega,\mathcal{F},\mathbb{P};\mathbb{R}^{sT})$.
Here, $f^t(\xi^t) := (f_1(\xi_1),f_2(\xi^2),\ldots,f_t(\xi^t))$ for $t=1,2,\ldots,T$.

To guarantee that the feasible solution set in each stage under the perturbed stochastic process $\tilde{\boldsymbol{\xi}}$ is nonempty, we need the following commonly used assumption.

Assumption 1 (*Relatively Complete Recourse Locally Around $\boldsymbol{\xi}$*). There exists a $\delta>0$ such that for any perturbed stochastic process $\tilde{\boldsymbol{\xi}}$ with $\|\tilde{\boldsymbol{\xi}}-\boldsymbol{\xi}\|\le\delta$, $x_1\in D_1$ and $\boldsymbol{x}_\tau\in D_\tau(x_{\tau-1},\tilde{\boldsymbol{\xi}}_\tau)$, $\tau=2,\ldots,t-1$, $D_t(\boldsymbol{x}_{t-1},\tilde{\boldsymbol{\xi}}_t)$ is nonempty $\mathbb{P}^t$-a.s. for $t=2,3,\ldots,T$.

The relatively complete recourse assumption is widely used in the stability analysis of stochastic programs, see the review [9] for two-stage stochastic programming problems and [2,4] for the multistage case.

To continue our discussion, in the same way as that in [5, Assumption 2.3], we introduce the following growth condition.

Assumption 2. There exists a positive number $C\ge 1$ such that, for every measurable mapping $\boldsymbol{x}_{t-1}:\Xi^{t-1}\to X_{t-1}$, there exists an optimal solution $\boldsymbol{x}_t(\xi^t)$ to problem (3) such that

$$\left\|\boldsymbol{x}_t(\xi^t)\right\| \le C\max\left\{1,\left\|\boldsymbol{x}_{t-1}(\xi^{t-1})\right\|\right\}\cdot\max\left\{1,\left\|\xi^t\right\|\right\},\ \ \mathbb{P}^t-\text{a.s.} \quad (4)$$

for $t=2,3,\ldots,T$. Specially, we have $\|x_1\|\le C$ for any $x_1\in D_1$.

Remark 1.3. Assumption 2 holds automatically when X_t, $1\le t\le T$, are bounded. From Assumption 2, there exists a subset $\bar{\Xi}^t\subseteq\Xi^t$ with $\mathbb{P}^t(\Xi^t\setminus\bar{\Xi}^t)=0$, such that for any $\xi^t\in\bar{\Xi}^t$, we have

$$\begin{aligned}\left\|x_t(\xi^t)\right\| &\le C\max\left\{1,\left\|x_{t-1}(\xi^{t-1})\right\|\right\}\cdot\max\left\{1,\left\|\xi^t\right\|\right\}\\ &= C\max\left\|x_{t-1}(\xi^{t-1})\right\|\cdot\max\left\{1,\left\|\xi^t\right\|\right\}\\ &\le C^2\max\left\{1,\left\|x_{t-2}(\xi^{t-2})\right\|\right\}\cdot\max\left\{1,\left\|\xi^{t-1}\right\|\right\}\\ &\quad\cdot\max\left\{1,\left\|\xi^t\right\|\right\}\\ &\le C^2\left\|x_{t-2}(\xi^{t-2})\right\|\cdot\max\left\{1,\left\|\xi^t\right\|\right\}^2.\end{aligned}$$

Then, we recursively obtain

$$\left\|\boldsymbol{x}_t(\xi^t)\right\| \le C^t\max\left\{1,\left\|\boldsymbol{\xi}^t\right\|\right\}^{t-1},\ \ \mathbb{P}^t-\text{a.s.},\ \ t=2,\ldots,T. \quad (5)$$

2. Main results

We present our main results about quantitative stability of multistage stochastic linear programs when the original stochastic process is perturbed by an approximation defined in Definition 1.2. To this end, we introduce the following two types of calm modifications.

Definition 2.1 (*Calm Modifications*). For an optimal solution $\boldsymbol{x}^*$ under stochastic process $\boldsymbol{\xi}$ satisfying the growth condition (4), we call

(i) $\bar{\boldsymbol{x}}^*(\hat{\boldsymbol{\xi}}) = (\bar{x}_1^*,\bar{\boldsymbol{x}}_2(\hat{\boldsymbol{\xi}}^2),\ldots,\bar{\boldsymbol{x}}_T(\hat{\boldsymbol{\xi}}^T))$ the class I calm modification under stochastic process $\hat{\boldsymbol{\xi}}$, if it is defined by

$$\bar{x}_1^* = x_1^*,\ \bar{\boldsymbol{x}}_t^*(\hat{\boldsymbol{\xi}}^t)\in\operatorname{argmin}_{z\in D_t(\bar{\boldsymbol{x}}_{t-1}^*(\hat{\boldsymbol{\xi}}^{t-1}),\hat{\boldsymbol{\xi}}_t)}\left\|z-\boldsymbol{x}_t^*(f^t(\boldsymbol{\xi}^t))\right\|,\quad t=2,\ldots,T;$$

(ii) $\bar{\bar{\boldsymbol{x}}}^*(\hat{\boldsymbol{\xi}}) = (\bar{\bar{x}}_1^*,\bar{\bar{\boldsymbol{x}}}_2(\hat{\boldsymbol{\xi}}^2),\ldots,\bar{\bar{\boldsymbol{x}}}_T(\hat{\boldsymbol{\xi}}^T))$ the class II calm modification under stochastic process $\hat{\boldsymbol{\xi}}$, if it is defined by

$$\bar{\bar{x}}_1^* = x_1^*,\ \bar{\bar{\boldsymbol{x}}}_t^*(\hat{\boldsymbol{\xi}}^t)\in\operatorname{argmin}_{z\in D_t(\bar{\bar{\boldsymbol{x}}}_{t-1}^*(\hat{\boldsymbol{\xi}}^{t-1}),\hat{\boldsymbol{\xi}}_t)}\left\|z-\boldsymbol{x}_t^*(\boldsymbol{\xi}^t)\right\|,\quad t=2,\ldots,T.$$

The class I calm modification can also be found in [5], where it is called the 'calm modification'. It is known from Assumption 1 that for a sufficiently small perturbation $\hat{\boldsymbol{\xi}}$ with $\|\hat{\boldsymbol{\xi}}-\boldsymbol{\xi}\|\le\delta$, the class I calm modification always exists, so does the class II calm modification when $\|\hat{\boldsymbol{\xi}}-f(\boldsymbol{\xi})\|\le\delta$. In addition, we know from [8, Theorem 14.37] that both $\bar{\boldsymbol{x}}^*$ and $\bar{\bar{\boldsymbol{x}}}^*$ can be selected to be measurable. From the viewpoint of measurability, we know that $\{f_t\}_{t=1}^T$ should be measurable too. If we consider a general perturbation to the stochastic process $\boldsymbol{\xi}$ as that in [4], it might be impossible to select a measurable class I calm modification and the complex filtration distance has to be adopted. In [5, Example A.3], the author illustrated that the measurability of f is indispensable