CONTINUITY AND STABILITY OF TWO-STAGE STOCHASTIC PROGRAMS WITH QUADRATIC CONTINUOUS RECOURSE

ZHIPING CHEN
Department of Computing Science, School of Mathematics and Statistics
Xi’an Jiaotong University, 710049 Xi’an, Shanxi, China

YOUPAN HAN
School of Science, Xi’an Polytechnic University, 710048 Xi’an, Shanxi, China

Abstract. For two-stage stochastic programs with quadratic continuous recourse where all the coefficients in the objective function and the right-hand side vector in the second-stage constraints vary simultaneously, we firstly show the locally Lipshitz continuity of the optimal value function of the recourse problem, then under suitable probability metric, we derive the joint Lipschitz continuity of the expected optimal value function with respect to the first-stage variables and the probability distribution. Furthermore, we establish the qualitative and quantitative stability results of the optimal value function and the optimal solution set with respect to the Fortet-Mourier probability metric, when the underlying probability distribution is perturbed. Finally, we show the exponential convergence rate of the optimal value sequence when we solve two-stage quadratic stochastic programs by the sample average approximation method.

1. Introduction. Stochastic program provides a powerful tool to treat decision problems under uncertainty, it has been successfully applied to many practical management problems in engineering, telecommunications, transportation, finance and so on. A variety of two- and multi-stage continuous (linear) recourse models and stochastic programs with chance (probabilistic) constraints have been proposed and studied in numerous papers, and there are also a multitude of books about stochastic programs, for instance, [2, 9, 14].

The two-stage stochastic programming with recourse is a flexible and practical mathematical model. It is concerned with problems that require a here-and-now decision on the basis of the given probabilistic information on random parameters. The costs to be minimized consist of the direct cost of the here-and-now or first-stage decision as well as the cost generated by the necessity of taking a recourse or second-stage decision in response to the random environment. The recourse cost is often formulated by means of the expected value with respect to the probability distribution of the involved random parameters. In this way, two-stage models and their solutions depend on the underlying probability distribution. Since this
distribution is often incompletely known in practice, or it has to be approximated for implementable computation, the stability of two-stage stochastic programming models with respect to the probability measure is thus very important. This problem has been studied in a number of papers, see, for example, [16, 21, 22, 23].

While important, ordinary stochastic programming models assume linear costs associated with recourse decisions. The linear recourse function ignores the widely existed risk averse attitude of decision makers, which is actually one of the main reasons for adopting stochastic programming models. As a way to reflect the potential nonlinear preference attached to recourse variables and to better model many important problems in fields like finance and risk management, quadratic stochastic programs with recourse have recently been investigated in several papers. These papers are mainly about the numerical solution of the proposed problem([3, 4, 5, 12, 15] and the modeling approach([10, 18]).

Compared with the rich literature on structural properties and stability of stochastic programs with linear recourse, there are only a few contributions to the two-stage quadratic stochastic programming problem ([18, 21]). A class of quadratic recourses is introduced in the second-stage problem in [18]. In the model, the right-hand side vector and the technology matrix appear in the objective function of the second-stage problem, rather than in the constraints as that in usual linear stochastic programs. The quantitative stability of the optimal value function to the above model is obtained in [21], where only the right-hand side vector in constraints and the linear part of the recourse objective function are perturbed and the coefficient matrix in the second-order term of the recourse objective function is positive definite.

Motivated by traditional two-stage stochastic programs with linear recourse, we consider in this paper the following general two-stage stochastic programming problem with quadratic continuous recourse:

$$\begin{align*}
\min & \quad \frac{1}{2} x^T G x + c^T x + Q(x, P) \\
\text{s.t.} & \quad Ax \leq b, \quad x \in \mathbb{R}^n,
\end{align*}$$

(1)

where $A \in \mathbb{R}^{l \times n}, b \in \mathbb{R}^l, c \in \mathbb{R}^n, G \in \mathbb{R}^{n \times n}, Q(x, P) = \int_{\xi} \Phi(x, \xi) P(d\xi), \Phi$ is the optimal value function of the following quadratic stochastic program:

$$\begin{align*}
\Phi(x, \xi) = \min_{z} & \quad \frac{1}{2} z^T H(\xi) z + u(\xi)^T z \\
\text{s.t.} & \quad Wz \leq h(\xi) - T(\xi)x, \\
& \quad z \in \mathbb{R}^m,
\end{align*}$$

(2)

here $u(\xi) \in \mathbb{R}^m$ and $H(\xi) \in \mathbb{R}^{m \times m}$ is a symmetric matrix. As in papers like [16],[20], we assume that $u(\xi), H(\xi), T(\xi) \in \mathbb{R}^{m \times n}$ and $h(\xi) \in \mathbb{R}^m$ are affine functionals of $\xi$, with $\xi$ being an s-dimensional random vector; $P$ belongs to the set of all Borel probability measures $\mathcal{P}(\Xi)$ on $\Xi$. The feasible solution set of problem (2) is denoted by $F(x, \xi)$, i.e., $F(x, \xi) = \{z | Wz \leq h(\xi) - T(\xi)x\}$ and $X = \{x \in \mathbb{R}^n | Ax \leq b\}$ is the feasible solution set of problem (1).

The model (1)-(2) extends usual two-stage stochastic programs with recourse from the linear recourse to the quadratic recourse. Meanwhile, it generalizes current two-stage quadratic stochastic programs with recourse in [21, 22] to the situation that all the coefficients except the recourse matrix in the quadratic recourse problem are random. After studying continuity properties of the optimal value function of
the second stage problem (2), we will establish the qualitative and quantitative stability of problem (1)-(2). Our results significantly improve the relevant conclusions in [21, 22].

Over the past few years, the sample average approximation (SAA) method has been increasingly investigated and recognized as one of the effective methods for solving stochastic programs, see discussions in [13, 17, 24, 25], and references therein. The basic idea of SAA is to replace the original problem by approximating the expected function value with the sample average value. The main benefit of SAA is that one may avoid the computation of the expected values which are often multiple integrations. The SAA approach for stochastic programs with linear recourse has been well-studied; see, for instance, [24, 26]. In particular, Shapiro established in [26] the exponential convergence of optimal solutions of the two-stage SAA problem by studying an equivalent SAA of the optimal value function of the second-stage problem. As an extension to the above result and an application of our stability results, we will discuss the exponential convergence of the optimal value function of problem (1)-(2).

The rest of the paper is organized as follows. In section 2, we obtain the locally Lipschitz continuity of the optimal value function of problem (2). Then, we establish the Lipschitz continuity of the expected optimal value function of problem (2). We investigate in section 3 the qualitative and quantitative stabilities of the optimal value function and the optimal solution set of the two-stage quadratic stochastic programming problem (1)-(2). When the problem (1)-(2) is approximated by its corresponding sample average, we show the exponential convergence rate of the corresponding optimal value sequence in section 4. Section 5 provides a brief conclusion.

2. Continuity of two-stage stochastic programming with quadratic continuous recourse. To ensure that problem (1)-(2) is well-defined, we impose the following assumptions:

(A1) for any $(\xi, x) \in \Xi \times \mathbb{R}^n$, there exists a $z \in \mathbb{R}^m$ such that $Wz \leq h(\xi) - T(\xi)x$.
(A2) for arbitrary $z \in \mathbb{R}^m$, there exists a $v \geq 0$ such that $H(\xi)z + WTv + u(\xi) = 0$.
(A3) $0 \in \text{int}(D)$, here D is the convex hull formed by the row vectors of $W$.
(A4) $P \in \mathcal{P}_3(\Xi)$, i.e., $P \in \mathcal{P}(\Xi)$ and $\int_{\Xi} \|\xi\|^3 P(d\xi) < +\infty$.

Remark 1. Conditions (A1) and (A2) correspond to the complete recourse and dual feasibility in the stochastic programming literature [23], these two conditions ensure that the optimal value function of problem (2) is finite; Assumption (A3) means that the feasible solution set $F(x, \xi)$ of the problem (2) is bounded, furthermore, it ensures that the corresponding optimal solution set is also bounded.

When the feasible solution set $F(x, \xi)$ is a bounded set, we have the following equivalent description about its boundedness ([7][7]).

Proposition 1. $F(x, \xi)$ is bounded if and only if $0 \in \text{int}(D)$.

Furthermore, if $F(x, \xi)$ is a bounded set, an upper bound on the norms of its feasible solutions can be derived from [8] as follows.

Theorem 1. If $F(x, \xi)$ is a bounded set, then for all $z \in F(x, \xi)$, we have

$$\|z\| \leq \frac{\max_{i \leq l} \{h(\xi) - T(\xi)x\}}{d(0, \text{bd}(D))} \leq \frac{\|h(\xi) - T(\xi)x\|_\infty}{d(0, \text{bd}(D))}.$$
The following theorem establishes the local continuity of the optimal value function of problem (2).

**Theorem 2.** For arbitrary \( \xi \), suppose \((A_1)\) and \((A_2)\) are satisfied and \( 0 \in \text{int}(D) \). Then the optimal value function of problem (2) is locally Lipschitz continuous.

**Proof.** \((A_1)\) and \((A_2)\) ensure that, for arbitrary \( x \in X \), \( \Phi(x, \cdot) \) is real-valued. Suppose that \( \xi' \) is in some neighborhood of \( \xi \), \( y(x) \) is an optimal solution of problem (2) with respect to \( \xi \) and \( y(x)' \) is a feasible solution of problem (2) with respect to \( \xi' \). We have the following estimation.

\[
\Phi(x, \xi') - \Phi(x, \xi) \leq \frac{1}{2} y(x)'^T H(\xi') y(x)' + u(\xi')^T y(x)' - \frac{1}{2} y(x)^T H(\xi) y(x) - u(\xi)^T y(x)
\]

\[
= \frac{1}{2} y(x)'^T H(\xi') y(x)' - \frac{1}{2} y(x)^T H(\xi) y(x) + \frac{1}{2} y(x)'^T H(\xi) y(x) - \frac{1}{2} y(x)^T H(\xi) y(x) - \frac{1}{2} y(x)'^T H(\xi) y(x) + \frac{1}{2} y(x)'^T H(\xi) y(x) - \frac{1}{2} y(x)'^T H(\xi) y(x) + \frac{1}{2} y(x)'^T H(\xi) y(x) - \frac{1}{2} y(x)'^T H(\xi) y(x)
\]

\[
+ y(x)'^T (u(\xi') - u(\xi)) + u(\xi)^T (y(x)' - y(x)).
\]

Since \( H(\xi) \) is a symmetric matrix, there exists an orthogonal matrix \( P(\xi) \) such that

\[
P(\xi)^T H(\xi) P(\xi) = \Lambda(\xi),
\]

where \( \Lambda(\xi) \) is a diagonal matrix whose diagonal elements are eigenvalues of \( H(\xi) \). Therefore

\[
y(x)'^T H(\xi) y(x)' - y(x)^T H(\xi) y(x)
\]

\[
= y(x)'^T P(\xi)^T \Lambda(\xi) P(\xi) y(x)' - y(x)^T P(\xi)^T \Lambda(\xi) P(\xi) y(x)
\]

\[
= (P(\xi) y(x)')^T \Lambda(\xi) P(\xi) y(x)' - (P(\xi) y(x))^T \Lambda(\xi) P(\xi) y(x)
\]

\[
= (P(\xi)(y(x) + y(x)'))^T \Lambda(\xi) P(\xi)(y(x)' - y(x))
\]

\[
= (y(x) + y(x)')^T H(\xi)(y(x)' - y(x)).
\]

From Theorem 1, we have

\[
\|y(x)\| \leq C_1 \max \{1, \|x\|\} \|h(\xi') - T(\xi') x\| \quad d(0, \text{bd}(D))
\]

where \( C_1 \) denotes the equivalent constant between \( \| \cdot \|_\infty \) and the Euclidean norm and \( C_2 \) is a constant relying on \( C_1 \) and the affine linearity of \( u(\xi'), T(\xi') \). Similarly, we obtain the following estimation for \( \|y(x)\|\):

\[
\|y(x)\| \leq C_2 \max \{1, \|x\|\} \max \{1, \|\xi\|\} \quad d(0, \text{bd}(D))
\]
Substituting all the above estimates into (3), we obtain

\[
\Phi(x, \xi') - \Phi(x, \xi) \\
\leq \frac{1}{2} y(x)'(H(\xi') - H(\xi)y(x)') + \frac{1}{2}(P(\xi)(y(x) + y(x)'))' + \Lambda(\xi)
\]

where

\[
P(\xi)(y(x)' - y(x)) + y(x)'(u(\xi') - u(\xi)) + u(\xi')(y(x)' - y(x))
\]

\[
\leq \frac{1}{2} \|y(x)\|^2 + \frac{1}{2}(y(x) + y(x)')' + \|H(\xi)\|
\]

\[
\|y(x)' - y(x)\| + \|y(x)\| \|u(\xi') - u(\xi)\| + \|u(\xi)\| \|y(x)' - y(x)\|
\]

\[
= \|y(x)\| \left( \frac{1}{2} \|y(x)\|^2 + \|H(\xi)\| \right) + \|u(\xi') - u(\xi)\| \\
+ \|y(x)\| \left( \frac{1}{2} \|y(x)\|^2 + \|u(\xi)\| \right) \|y(x)' - y(x)\|
\]

\[
\leq \mathcal{K} \max\{1, \|\xi\|^2\} \max\{1, \|x\|^2\} \|\xi - \xi'\|
\]

\[
+ \tilde{K} \max\{1, \|\xi\|^2, \|\xi'\|^2\} \max\{1, \|x\|\} \|y(x)' - y(x)\|
\]

for some \(\mathcal{K}, \tilde{K} < \infty\). As for the last inequality in the above expression, we used the affine linearity of \(H(\xi), u(\xi)\) and the boundedness of the feasible solution set.

From Theorem 2.2 in [11], there exists a constant \(\beta\) such that

\[
\|y(x)' - y(x)\| \leq \beta \max\{1, \|x\|\} \|\xi - \xi'\|.
\]

Let \(\tilde{K}_1 = \max\{\mathcal{K}, \tilde{K}, \beta\}\), the above results imply

\[
\Phi(x, \xi') - \Phi(x, \xi) \leq \tilde{K}_1 \max\{1, \|x\|^2\} \max\{1, \|\xi\|^2, \|\xi'\|^2\} \|\xi - \xi'\|
\]

Similarly, we have

\[
\Phi(x, \xi) - \Phi(x, \xi') \leq \tilde{K}_1 \max\{1, \|x\|^2\} \max\{1, \|\xi\|^2, \|\xi'\|^2\} \|\xi - \xi'\|
\]

\[\Box\]

To prove the Lipschitz continuity of the expected optimal value function \(Q(x, P)\) of problem (2), we need to describe the difference between two probability distributions. Therefore, the probability metric is adopted here. In stochastic programs, there are many probability metrics, for instance, the Kolmogorov-Smirnov probability metric, the Wasserstein probability metric and the Fortet-Mourier probability metric. Suppose that the stochastic variable \(g(\xi)\) satisfies the following inequality:

\[
|g(\xi) - g(\xi')| \leq \max\{1, \|\xi\|, \|\xi'\|\}^{p-1} \|\xi - \xi'\|, \quad \forall \, \xi, \xi' \in \Xi,
\]

where \(p \geq 1\). Let \(\mathcal{F}_p(\Xi)\) denote all the \(g(\xi)\)s satisfying the above inequality, then the Fortet-Mourier probability metric is defined as:

\[
\zeta_p(\mu, \nu) = \sup_{g \in \mathcal{F}_p(\Xi)} \left| \int_{\Xi} g(\xi) \mu(d\xi) - \int_{\Xi} g(\xi) \nu(d\xi) \right|.
\]

From Theorem 2, we know that the optimal value function of the second-stage problem (2) is locally Lipschitz continuous; moreover, when \(X\) is a bounded set, i.e., there is a constant \(r\) such that \(\|x\| \leq r, \forall x \in X\). Then there exists a constant \(K_1 = \tilde{K}_1 \max\{1, r^2\} > 0\) such that:

\[
|\Phi(x, \xi) - \Phi(x, \xi')| \leq K_1 \max\{1, \|\xi\|^2, \|\xi'\|^2\} \|\xi - \xi'\|
\]

From the definition of \(\mathcal{F}_p(\Xi)\), we have \(\frac{1}{\mathcal{K}} \Phi(x, \xi) \in \mathcal{F}_p(\Xi)\). So, we adopt the third order Fortet-Mourier probability metric. Under this probability metric, we establish
the Lipschitz continuity of the expected optimal value function of the second-stage problem (2) with respect to the first-stage decision variable and the probability distribution.

**Theorem 3.** Suppose that $(A_1)$-$(A_4)$ are satisfied and $X$ is a bounded set, then the expected optimal value function $Q(x, P)$ is Lipschitz continuous.

**Proof.** For arbitrary $x, y \in X$ and $\mu, \nu \in P(\Xi)$, we have

$$|Q(x, \mu) - Q(y, \nu)| = |Q(x, \mu) - Q(x, \nu) + Q(x, \nu) - Q(y, \nu)|$$

$$\leq |Q(x, \mu) - Q(x, \nu)| + |Q(x, \nu) - Q(y, \nu)|.$$

For $|Q(x, \nu) - Q(y, \nu)|$, we have

$$|Q(x, \nu) - Q(y, \nu)| = \left| \int_\Xi \Phi(x, \xi) \nu d(\xi) - \int_\Xi \Phi(y, \xi) \nu d(\xi) \right|$$

$$= \int_\Xi |(\Phi(x, \xi) - \Phi(y, \xi)) \nu d(\xi)|.$$

For $\Phi(x, \xi) - \Phi(y, \xi)$, we have the following results.

$$|\Phi(x, \xi) - \Phi(y, \xi)|$$

$$= \frac{1}{2} (z(x)^T H(\xi) z(x) + u(\xi)^T z(x) - z(y)^T H(\xi) z(y) + u(\xi)^T z(y))$$

$$= \frac{1}{2} (P(\xi)(z(x) + z(y))) \Lambda(\xi) P(\xi)(z(x) - z(y)) + u(\xi)^T z(y))$$

$$= \left( \frac{1}{2} \|z(x)\| + \|z(y)\| \right) \|H(\xi)\| + \|u(\xi)\| \|z(x) - z(y)\|.$$

where $z(x), z(y)$ are the optimal solutions of problem (2) with respect to the first-stage variable $x$ and $y$, respectively.

Since the first-stage variables appear only in the right-hand side of the constraints of problem (2), we know from Theorem 2.2 in [11] that there exists a constant $K_2$ such that

$$\|z(x) - z(y)\| \leq K_2 \|x - y\| \|T(\xi)\|.$$

Meanwhile, by adopting the similar proof method as that for Theorem 1, we know that there exists a constant $K_3$, such that

$$\|z(x)\| \leq \frac{K_3}{d(0, bd(D))} \max\{1, \|x\|\} \max\{1, \|\xi\|\}$$

and

$$\|z(y)\| \leq \frac{K_3}{d(0, bd(D))} \max\{1, \|y\|\} \max\{1, \|\xi\|\}.$$  

Therefore, we have from $(A_4)$ that

$$|Q(x, \nu) - Q(y, \nu)| = \int_\Xi |(\Phi(x, \xi) - \Phi(y, \xi)) \nu d(\xi)|$$

$$\leq \int_\Xi K_4 \max\{1, \|\xi\|^3\} \|x - y\| \nu d(\xi)$$

$$= K_5 \|x - y\|.$$
where \( K_4 \) is a constant dependent on \( K_2, K_3, d(0, \text{bd}(D)) \), the boundedness of \( X \) and the affine linearity of \( H(\xi), u(\xi), T(\xi) \), and

\[
K_5 = \int_{\Xi} K_4 \max \{1, \|\xi\|^3\} \nu(d\xi).
\]

Define

\[
\phi(x, \xi) := \frac{1}{K_1} \Phi(x, \xi).
\]

From the definition \( F_p \), it is easy to see that \( \phi(x, \xi) \in F_3(\Xi) \). Therefore, for \( Q(x, \mu) - Q(x, \nu) \), we have the following results.

\[
|Q(x, \mu) - Q(x, \nu)| = \left| \int_{\Xi} \Phi(x, \xi) \mu(d\xi) - \int_{\Xi} \Phi(x, \xi) \nu(d\xi) \right|
\]

\[
= \left| \int_{\Xi} \Phi(x, \xi) (\mu - \nu)(d\xi) \right|
\]

\[
= K_1 \left| \int_{\Xi} \phi(x, \xi) (\mu - \nu)(d\xi) \right|
\]

\[
\leq K_1 \zeta_3(\mu, \nu).
\]

Thus, we obtain the desired results.

3. Stability. We know from Theorem 3 that \( Q(x, \mu) \) is Lipschitz continuous, at the same time, \( \frac{1}{2} x^T G x + c^T x \) is continuous. Thus, the objective function of problem (1) is continuous. According to Theorems 4.2.1 and 4.2.2 in [1], we have the following conclusion.

**Theorem 4.** Suppose that \((A_1) - (A_4)\) are satisfied and \( X \) is a bounded set. Then

(i) the optimal value function \( \vartheta(\cdot) \) is continuous at \( \mu \);

(ii) the optimal solution set \( S(\cdot) \) is Berge upper semi-continuous at \( \mu \).

In order to transform problem (1)-(2) into a minimization problem with respect to the first-stage decision variables, we define \( f_0(x, \xi) : \mathbb{R}^n \times \Xi \rightarrow \mathbb{R} \), as:

\[
f_0(x, \xi) = \begin{cases} \frac{1}{2} x^T G x + c^T x + \Phi(x, \xi), & \text{if } x \in X, F(x, \xi) \neq \emptyset, \\ +\infty, & \text{otherwise.} \end{cases}
\]

Then, problem (1)-(2) can be re-expressed as:

\[
\min \left\{ \int_{\Xi} f_0(x, \xi) P(d\xi), \quad x \in X \right\}. \tag{4}
\]

Firstly, we examine some continuity properties of the function \( f_0(x, \xi) \) by utilizing the conclusions in the last section.

**Theorem 5.** Assume that \((A_1) - (A_3)\) are satisfied, \( X \) is a bounded set, i.e., there is a constant \( r \) such that \( \|x\| \leq r, \forall x \in X \). Then for any chosen \( \xi \), there exist constants \( \delta_3(\xi), L_1, L_2, L_3 > 0 \) such that

\[
f_0(x, \xi) - f_0(x, \bar{\xi}) \leq L_1 \max \{1, \|\xi\|, \|\bar{\xi}\|\}^2 \|\xi - \bar{\xi}\|,
\]

\[
f_0(x, \xi) - f_0(\bar{x}, \xi) \leq L_2 \max \{1, \|\xi\|^3\} \|x - \bar{x}\|
\]

\[
f_0(x, \xi) \leq L_3 \max \{1, \|\xi\|^3\}
\]

as long as \( \|\bar{\xi} - \xi\| \leq \delta_3(\xi) \).
\textit{Proof.} \((A_1)\) and \((A_2)\) ensure that \(f_0(x, \cdot)\) is real-valued for every \(x \in X\). For any fixed \(x\), we have

\[
\begin{aligned}
f_0(x, \xi) - f_0(x, \bar{\xi}) &= \frac{1}{2}x^T G x + c^T x + \Phi(x, \xi) - \left(\frac{1}{2}x^T G x + c^T x + \Phi(x, \bar{\xi})\right) \\
&= \Phi(x, \xi) - \Phi(x, \bar{\xi}).
\end{aligned}
\]

From Theorem 2, there exists a constant \(\delta_3(\xi)\) dependent on \(\xi\) such that

\[
\Phi(x, \xi) - \Phi(x, \bar{\xi}) \leq K_1 \max\{1, \|\xi\|^2\} \max\{1, \|\xi\|^2, \|\bar{\xi}\|^2\} \|\xi - \bar{\xi}\|
\]
as long as \(\|\bar{\xi} - \xi\| \leq \delta_3(\xi)\).

Let \(L = K_1 \max\{1, \|\xi\|^2\}\), we have

\[
f_0(x, \xi) - f_0(x, \bar{\xi}) \leq L_1 \max\{1, \|\xi\|, \|\bar{\xi}\|\}^2 \|\xi - \bar{\xi}\|.
\]

For fixed \(\xi\), we have

\[
\begin{aligned}
f_0(x, \xi) - f_0(\bar{x}, \xi) &\leq \frac{1}{2}x^T G x + c^T x + \Phi(x, \xi) - \left(\frac{1}{2}x^T G x + c^T x + \Phi(\bar{x}, \xi)\right) \\
&= \frac{1}{2}x^T G x + c^T x - \frac{1}{2}\bar{x}^T G \bar{x} + c^T \bar{x} + \Phi(x, \xi) - \Phi(\bar{x}, \xi) \\
&= \left(\frac{1}{2}U^T (x + \bar{x})^T \Lambda U + c\right) (x - \bar{x}) + \Phi(x, \xi) - \Phi(\bar{x}, \xi),
\end{aligned}
\]

where \(\Lambda\) is the diagonal matrix whose diagonal elements are eigenvalues of \(G\), \(U\) is an unit orthogonal matrix and satisfies \(U^T GU = \Lambda\).

For the term \(\Phi(x, \xi) - \Phi(\bar{x}, \xi)\), we have from the proof of Theorem 3 that

\[
\Phi(x, \xi) - \Phi(\bar{x}, \xi) \leq K_2 \max\{1, \|\xi\|^3\} \|x - \bar{x}\|.
\]

Let \(L_2 = \max\{r\|G\| + \|c\|, K_2\}\), we obtain

\[
f_0(x, \xi) - f_0(\bar{x}, \xi) \leq L_2 \max\{1, \|\xi\|^3\} \|x - \bar{x}\|.
\]

Finally, we derive an upper bound for the objective function of problem (1)-(2) when the first-stage feasible solution set \(X\) is bounded. From Theorem 1, we have the following estimation for the optimal solution of problem (2):

\[
\|y\| \leq \frac{\max\{(h(\xi) - T(\xi) x), 1 \leq i \leq l\}}{d(0, \text{bd}(D))} = \frac{\bar{L}_3}{d(0, \text{bd}(D))} \max\{1, \|\xi\|\},
\]

where \(\bar{L}_3\) is a constant which depends on the affine linearity of \(T(\xi)\) and \(h(\xi)\) and the boundedness of the first-stage feasible solution set \(X\). With this estimation, we can deduce an upper estimation for the objective function of problem (4). Concretely,

\[
\begin{aligned}
f_0(x, \xi) &= \frac{1}{2}x^T G x + c^T x + \Phi(x, \xi) \\
&= \left(\frac{1}{2}\|G\| \|x\| + \|c\|\right) \|x\| + \left(\frac{1}{2}\|H(\xi)\| \|y\| + \|u(\xi)\|\right) \|y\| \\
&\leq \left(\frac{1}{2}\|G\| \|x\| + \|c\|\right) \|x\| + \left(\frac{1}{2}\|H(\xi)\| + \|u(\xi)\|\right) \max\{1, \|y\|\} \|y\| \\
&\leq \left(\frac{1}{2}\|G\| \|x\| + \|c\|\right) \|x\| + \left(\frac{1}{2}\|H(\xi)\| + \|u(\xi)\|\right) \\
&\quad \max\left\{1, \left(\frac{\bar{L}_3}{d(0, \text{bd}(D))}\right)^2\right\} \max\{1, \|\xi\|\}^2 \\
&\leq L_3 \max\{1, \|\xi\|^3\}.
\end{aligned}
\]
Here $L_3$ is a constant which depends on $\|G\|, \|c\|, \|x\|$, as well as the constants relevant with the affine linearity of $H(\xi), u(\xi), h(\xi)$ and $T(\xi)$, respectively. \hfill \Box

To present our stability results, we introduce the following auxiliary functions. The growth function $\psi_P$ on $R^+$ is defined by

$$
\psi_P(\tau) := \min \left\{ \int_{\Xi} f_0(x, \xi) P(d\xi) - \vartheta(P) : d(x, S(P)) \geq \tau, \ x \in B(P) \right\},
$$

with its inverse function being $\psi_P^{-1}(t) := \sup\{\tau \in R_+ : \psi_P(\tau) \leq t\}$, where $B(P)$ is a neighborhood of the optimal solution set $S(P)$. The condition function $\Psi_P$ is defined by

$$
\Psi_P(\eta) := \eta + \psi_P^{-1}(2\eta), \ \eta \in R_+.
$$

Both functions $\psi_P$ and $\Psi_P$ depend on the data of problem (1)-(2), in particular, on the probability measure $P$. They are lower semi-continuous on $R_+$, $\psi_P$ is nondecreasing, $\Psi_P$ is increasing, and both vanish at 0. One can refer to [19] and [20] for more details about $\psi_P$ and $\Psi_P$.

Utilizing the above results, we can establish the quantitative stability results of problem (1)-(2) as follows.

**Theorem 6.** Let conditions $(A_1)$ – $(A_4)$ be satisfied and $X$ be compact, then there exist constants $L > 0, \delta > 0$ such that

$$
|\vartheta(P) - \vartheta(Q)| \leq L\zeta_3(P, Q),
$$

(5)

$$
\emptyset \neq S(Q) \subseteq S(P) + \Psi_P(L\zeta_3(P, Q))B
$$

(6)

for each $Q \in \mathcal{P}_3(\Xi)$ and $\zeta_3(P, Q) < \delta$.

**Proof.** Since the optimal value function of the problem (2) is locally Lipschitz continuous, it is thus lower semi-continuous. We have from [19] that it is random lower semi-continuous (for details, see Example 14.31 in [19]). We have from Theorem 5 that

$$
f_0(x, \xi) \leq C_0 \max \{1, \|\xi\|^3\}
$$

for each pair $(x, \xi) \in X \times \Xi$ and some constant $C_0$, $\int_{\Xi} f_0(x, \xi) Q(d\xi)$ is thus finite due to $(A_4)$. Moreover, it is also lower semi-continuous on $X$. As $X$ is a compact set, the solution set $S(Q)$ is therefore nonempty.

Next, we estimate the difference between the optimal value functions of the original problem and its perturbed problem with respect to the chosen probability metric. From Theorem 5, we have the following estimation for the objective function of problem (4)

$$
f_0(x, \xi) \leq f_0(x, \tilde{\xi}) \leq L \max \{1, \|\xi\|, \|\tilde{\xi}\|\}^2 \|\xi - \tilde{\xi}\|.
$$

Let

$$
f(x, \xi) := \frac{1}{L} \int_{\Xi} f_0(x, \xi) \in \mathcal{F}_3(\Xi), \ \forall \xi \in \Xi.
$$

We have

$$
\sup_{x \in X} \left| \int_{\Xi} f_0(x, \xi)(P - Q)(d\xi) \right| = L \left| \int_{\Xi} f(x, \xi)(P - Q)(d\xi) \right| \leq L\zeta_3(P, Q).
$$

Thus, we obtain the following results.

$$
|\vartheta(P) - \vartheta(Q)| \leq \sup_{x \in X} \left| \int_{\Xi} f_0(x, \xi)(P - Q)(d\xi) \right| \leq L\zeta_3(P, Q).
$$

This completes the proof of the inequality (5).
Finally, we prove the inequality (6). Take \( x \in S(Q) \) and \( \delta = \zeta_3(P,Q) \), choose a feasible solution \( \overline{x} \) of problem (4) such that \( \|x - \overline{x}\| \leq L\delta \). From the above results, the definitions of the growth function and the condition function, we have

\[
\delta(L + L) \geq L\delta + \vartheta(Q) - \vartheta(P) = L\delta + \int_{\Xi} f_0(x,\xi)Q(d\xi) - \vartheta(P)
\]

\[
\geq \int_{\Xi} f_0(\overline{x},\xi)P(d\xi) - \vartheta(P) \geq \psi_P(d(\overline{x},S(P)))
\]

\[
\geq \inf_{y\in x+L\delta B} \psi_P(d(y,S(P))) = \psi_P(d(x,S(P) + L\delta B)).
\]

Therefore, we obtain

\[
d(x, S(P)) \leq L\delta + d(x, S(P) + L\delta B)
\]

\[
\leq L\delta + \psi_P^{-1}(\delta(L + L))
\]

\[
\leq L\delta + \psi_P^{-1}(2L\delta) = \Psi_P(L\delta) = \Psi_P(L\zeta_3(P,Q)).
\]

Since \( S(Q) \) is nonempty, we have shown that \( \emptyset \neq S(Q) \subseteq S(P) + \Psi_P(L\zeta_3(P,Q))B \).

\[\square\]

**Remark 2.** The above results generalize Theorem 23 in [20] from the linear recourse to the quadratic recourse.

4. **Exponential convergence of SAA.** As an application of the continuity and stability results obtained in the last two sections, we investigate the exponential convergence of the approximating optimal value function sequence when we try to solve problem (1)-(2) by the SAA method.

Let \( f_0(x) = E(f_0(x,\xi)) \), \( v(x) = \min_{x \in X} f_0(x) \). Suppose that we have generated \( N \) random sample realizations \( \xi^1, \ldots, \xi^N \) of the random vector \( \xi \), and the expected value \( E(f_0(x,\xi)) \) is approximated by the corresponding sample average

\[
\frac{1}{N} \sum_{j=1}^{N} f_N(x,\xi^j). \]

Then, the problem (4) can be approximated by the following problem, the so-called SAA problem:

\[
v_N(x) =: \min \{ f_N(x) = \frac{1}{N} \sum_{j=1}^{N} f_N(x,\xi^j), \quad x \in X \}. \tag{7}
\]

In what follows, we study the convergence rate of \( v_N(x) \) to \( v(x) \) when the sample size \( N \) tends to infinity. To this end, we firstly introduce the following function: consider a (real valued) random variable \( Y \) with its mean being \( \mu = E[Y] \), its moment-generating function \( M(t) := E[e^{tY}] \) is viewed as an extended real-valued function, i.e., it can take \( +\infty \). Obviously, \( M(t) > 0 \) for all \( t \in R, M(0) = 1 \), and the domain \( \{ t : M(t) < +\infty \} \) of the moment-generating function is an interval containing zero. The conjugate function

\[
I(Z) := \sup_{t \in R} \{ tZ - \Lambda(t) \}
\]

of the logarithmic moment-generating function \( \Lambda(t) := \log M(t) \) is called the large deviation (LD) rate function of \( Y([6]) \). It can be shown that both \( \Lambda(\cdot) \) and \( I(\cdot) \) are convex functions, see [6] for details.

We denote by

\[
M_x(t) := E\{e^{f_0(x,\xi)} - f_0(x)\}
\]
the moment generating function of the random variable \( f_0(x, \xi) - f_0(x) \). Let us make the following assumption.

(A5) For every \( x \in X \), the moment generating function \( M_x(t) \) is finite for any \( t \) in a neighborhood of zero.

Then we have the following conclusion about the exponential convergence of \( \{v_N(x)\} \).

**Theorem 7.** Let conditions (A1) – (A5) be satisfied and \( X \) be compact, then for any \( \varepsilon > 0 \), there exist constants \( C = C(\varepsilon), \beta = \beta(\varepsilon) \), independent of \( N \), such that

\[
\Pr\{|v_N(x) - v(x)| < \varepsilon\} \geq 1 - C(\varepsilon)e^{-N\beta(\varepsilon)}.
\]

**Proof.** Firstly,

\[
\Pr\{|v_N(x) - v(x)| < \varepsilon\} \geq \Pr\{|f_N(x) - f_0(x)| < \varepsilon, \forall x \in X\}
\]

holds because

\[
|f_N(x) - f_0(x)| < \varepsilon, \forall x \in X
\]

\[
\Leftrightarrow f_0(x) - \varepsilon < f_N(x) < f_0(x) + \varepsilon, \forall x \in X
\]

\[
\Rightarrow \min_{x \in X} f_0(x) - \varepsilon < \min_{x \in X} f_N(x) < \min_{x \in X} f_0(x) + \varepsilon
\]

\[
\Leftrightarrow v(x) - \varepsilon < v_N(x) < v(x) + \varepsilon
\]

\[
\Leftrightarrow |v_N(x) - v(x)| < \varepsilon.
\]

For the \( \Pr\{|f_N(x) - f_0(x)| < \varepsilon, \forall x \in X\} \), we have

\[
\Pr\{|f_N(x) - f_0(x)| < \varepsilon, \forall x \in X\} = 1 - \Pr\{\exists x \in X \text{ s.t. } |f_N(x) - f_0(x)| \geq \varepsilon\}.
\]

By the Cramér’s LD theorem ([6]), we have that, for any \( x \in X \) and \( \varepsilon > 0 \),

\[
\Pr\{f_N(x) - f_0(x) \geq \varepsilon\} \leq e^{-NM_x(\varepsilon)}.
\]

Similarly,

\[
\Pr\{f_N(x) - f_0(x) \leq -\varepsilon\} \leq e^{-NM_x(-\varepsilon)},
\]

and hence

\[
\Pr\{|f_N(x) - f_0(x)| \geq \varepsilon\} \leq e^{-NM_x(-\varepsilon)} + e^{-NM_x(\varepsilon)}.
\]

By assumption (A5), we have that both \( M_x(\varepsilon) \) and \( M_x(-\varepsilon) \) are positive for every \( x \in X \).

For any \( \nu > 0 \), let \( X_\nu = \{x^1, \ldots, x^M\} \) be such that for every \( x \in X \), there exists an \( x^i, i \in \{1, \ldots, M\} \), such that \( \|x - x^i\| \leq \nu \). i.e., \( x^1, \ldots, x^M \) is an \( \nu \)-net in \( X \). We can choose this net in such a way that \( M \leq C(D_M)^N \), where \( D := \sup_{x, x' \in X} \|x' - x\| \) is the diameter of \( X \) and \( C \) is a generic constant. By Theorem 5, we have

\[
f(x^i) - f(x) \leq L\|x^i - x\|,
\]

where \( L = E(L_2 \max\{1, \|\xi\|^3\}) \) is finite due to assumption (A4). Moreover,

\[
f_N(x^i) - f_N(x) \leq L_N\|x^i - x\|,
\]

where \( L_N = \frac{1}{N} \sum_{j=1}^{N} (L_2 \max\{1, \|\xi^j\|^3\}) \). Again, because of assumption (A4), by the Cramér’s LD theorem, we have that for any \( L' > L \), there is a positive constant \( \lambda \) such that

\[
\Pr\{L_N \geq L'\} \leq e^{-N\lambda}.
\]
For an $x \in X$, let $i(x) \in \arg \min_{1 \leq i \leq M} \| x - x_{i(x)} \|$. By the construction of the $\nu$-net, we have that $\| x - x_{i(x)} \| \leq \nu$ for every $x \in X$. Then
\[
|f_N(x) - f_0(x)| \\
\leq |f_N(x) - f_N(x_{i(x)})| + |f_N(x_{i(x)}) - f(x_{i(x)})| + |f(x_{i(x)}) - f_0(x)| \\
\leq L_N \nu + |f_N(x_{i(x)}) - f(x_{i(x)})| + \nu \\
= (L_N + \nu)(|f_N(x_{i(x)}) - f(x_{i(x)})|).
\]
Let us now take an $\nu$-net with such $\nu$ that $L_N \nu = \frac{\varepsilon}{4}$, i.e., $\nu = \frac{\varepsilon}{4L_N}$. Then
\[
\text{Prob}\{|f_N(x) - f_0(x)| \geq \varepsilon\} \leq \text{Prob}\{L_N \nu + |f_N(x_{i(x)}) - f(x_{i(x)})| \geq \frac{3\varepsilon}{4}\}.
\]
Moreover, we have by (10) that
\[
\text{Prob}\{L_N \nu \geq \frac{\varepsilon}{2}\} \leq e^{-N\lambda}
\]
for some $\lambda > 0$, and hence
\[
\text{Prob}\{|f_N(x) - f_0(x)| \geq \varepsilon\} \leq e^{-N\lambda} + \sum_{i=1}^{M} e^{-NM_{x_{i(x)}}(\frac{\varepsilon}{4})} + e^{-NM_{x_{i(x)}}(\frac{\varepsilon}{4})}.
\]
(11)
Since the above choice of the $\nu$-net does not depend on the specific sample realization (although it depends on $N$), and both $M_{x_{i(x)}}(\frac{\varepsilon}{4})$ and $M_{x_{i(x)}}(\frac{\varepsilon}{4})$ are positive, $i = 1, \cdots, M$, we can then deduce (8) from (9) and (11). This completes the proof. \hfill $\Box$

5. Conclusions. In this paper, we have studied the continuity and stability of two-stage quadratic stochastic programs with quadratic recourse. Moreover, we have established the exponential convergence rate of the corresponding optimal value sequence when we solve this stochastic programming problem by the sample average approximation method.

Both the qualitative and quantitative stability results are established. Compared to the existing conclusions, we derive the corresponding results for more general problems, that is, all the coefficients in the second-stage quadratic programming program are random except for the recourse matrix. Moreover, except for its symmetry, there are no additional restrictions on the coefficient matrix in the second-order term of the objective function.

Of course, we can consider extending the results in this paper to more complex situations, for example, the involvement of integer variables in the second-stage recourse decision vector, the multistage quadratic stochastic programs.

REFERENCES

[1] B. Bank, J. Guddat, D. Klatte, B. Kummer and K. Tammer, Non-Linear Parametric Optimization, Akademie Verlag, Berlin, 1982.
[2] J. R. Birge and F. Louveaux, Introduction to Stochastic Programming, Springer, New York, 1997.
[3] X. Chen, L. Qi and R. S. Womersley, Newton’s method for quadratic stochastic programs with recourse, J. Comput. Appl. Math., 60 (1995), 29–46.
[4] X. Chen and R. S. Womersley, Random test problems and parallel methods for quadratic programs and quadratic stochastic programs, Optim. Method Softw., 13 (2000), 275–306.
[5] G. M. Cho, Log-barrier method for two-stage quadratic stochastic programming, Appl. Math. Comput., 164 (2005), 45–69.
[6] A. Dembo and O. Zeitouni, Large Deviations Techniques and Applications, Springer-Verlag, New York, 1998.
[7] M. A. Goberna and M. A. López, *Linear Semi-Infinite Optimization*, John Wiley and sons, Chichester, 1998.
[8] Y. Han and Z. Chen, Quantitative stability of full random two-stage stochastic programs with recourse, *Optim. Lett.*, to appear.
[9] P. Kall and S. W. Wallace, *Stochastic Programming*, John Wiley and Sons, Chichester, 1994.
[10] W. K. Klein Haneveld and M. H. Van der Vlerk, Stochastic integer programming: general models and algorithms, *Ann. Oper. Res.*, 85 (1999), 39–57.
[11] O. L. Mangasarian and T. H. Shiau, Lipschitz continuity of solutions of linear inequalities, programs, and complementary problems, *SIAM J. Control Optim.*, 25 (1987), 583–595.
[12] S. Mehrotra and M. G. Özevin, Decomposition-based interior point methods for two-stage stochastic convex quadratic programs with recourse, *Oper. Res.*, 57 (2009), 964–974.
[13] E. L. Plambeck, B. R. Fu, S. M. Robinson and R. Suri, Sample-path optimization of convex stochastic performances functions, *Math. Program.*, 75 (1996), 137–176.
[14] A. Prekopa, *Stochastic Programming*, Kluwer Academic Publishers, Dordrecht, Boston, 1995.
[15] L. Qi and R. S. Womersley, An SQP algorithm for extended linear-quadratic problems in stochastic programming, *Ann. Oper. Res.*, 56 (1995), 251–285.
[16] S. T. Rachev and W. Römisch, Quantitative stability in stochastic programming: the methods of probability metrics, *Math. Oper. Res.*, 27 (2002), 792–818.
[17] S. M. Robinson, Analysis of sample-path optimization, *Math. Oper. Res.*, 21 (1996), 513–528.
[18] R. T. Rockafeller and R. J-B. Wets, A Lagrangian finite generation technique for solving linear-quadratic problems in stochastic programming, *Math. Program. study*, 28 (1986), 63–93.
[19] R. T. Rockafeller and R. J-B. Wets, *Variational Analysis*, Springer, Berlin, 1998.
[20] W. Römisch, Stability of stochastic programming, in *Stochastic Programming: Handbooks in Operations Research and Management Science Vol.10* (eds. A. Ruszczyński, A. Shapiro), North-Holland Publishing Company, Amsterdam, (2003), 483–554.
[21] W. Römisch and R. Schultz, Distribution sensitivity in stochastic programming, *Math. Program.*, 50 (1991), 197–226.
[22] W. Römisch and R. Schultz, Lipschitz stability for stochastic programs with complete recourse, *SIAM J. Optim.*, 6 (1996), 531–547.
[23] W. Römisch and R. J. B. Wets, Stability of ε-approximate solutions to convex stochastic programs, *SIAM J. Optim.*, 18 (2007), 961–979.
[24] A. Shapiro, Monte Carlo sampling methods, in *Stochastic Programming: Handbooks in Operations Research and Management Science Vol.10* (eds. A. Ruszczyński, A. Shapiro), North-Holland Publishing Company, Amsterdam, (2003), 353–425.
[25] A. Shapiro and T. Homem-de-Mello, On rate of convergence of Monte Carlo approximations of stochastic programs, *SIAM J. Optim.*, 6 (1996), 531–547.
[26] A. Shapiro, Complexity of two and multi-stage stochastic programming problems, 2005. Available from: http://www2.isye.gatech.edu/~ashapiro/publications.html.

Received September 2014; 1st revision November 2014; final revision May 2015.

*E-mail address: zchen@mail.xjtu.edu.cn*

*E-mail address: hanyp1999@163.com*