Strong Convexity in Stochastic Programs with Deviation Risk Measures

Matthias Claus · Rüdiger Schultz · Kai Spürkel

Received: date / Accepted: date

Abstract We give sufficient conditions for the expected excess and the upper semideviation of recourse functions to be strongly convex. This is done in the setting of two-stage stochastic programs with complete linear recourse and random right-hand side. This work extends results on strong convexity of risk-neutral models.

Keywords Two-Stage Stochastic Programming · Linear Recourse · Strong Convexity · Expected Excess · Semideviation

Mathematics Subject Classification (2000) 90C15 · 90C31

1 Introduction

Stochastic programs with linear recourse belong to those optimization problems which are understood best, both structurally and algorithmically. To some extent, this statement gains its substance from fairly widespread convexity properties.

Matthias Claus
University of Duisburg-Essen, Thea-Leymann-Str. 9, D-45127 Essen, Germany
Tel.: +49 201 1836887
E-mail: matthias.claus@uni-due.de
ORCID iD: 0000-0002-2617-0855

Rüdiger Schultz
University of Duisburg-Essen, Thea-Leymann-Str. 9, D-45127 Essen, Germany
Tel.: +49 201 1836880
E-mail: ruediger.schultz@uni-due.de

Kai Spürkel
University of Duisburg-Essen, Thea-Leymann-Str. 9, D-45127 Essen, Germany
Tel.: +49 201 1836883
E-mail: kai.spuerkel@uni-due.de
This has spurred research into conditions for feasibility and optimality that are verifiable with reasonable effort. Another important topic in this respect is stability of the model under data perturbation, (stochastic) approximation, or statistical estimation. In fact, for these and any type of change the model could be exposed to, it is desirable that the resulting changes in optimal values and sets of feasible or optimal solution points stay small for small model changes.

Risk neutral stochastic programs with linear recourse obey “some basic” convexity and may fulfill a requirement called strong convexity. In the present paper, the accent is on deriving verifiable sufficient conditions for objective functions to be strongly convex. Extending the purely expectation based analysis in [13], we address risk aversion via deviation risk measures.

The paper is organized as follows. In section 2 preliminaries on stochastic programs and strong convexity are collected. Section 3 contains the main results on strong convexity of stochastic programs with deviation risk measures. Section 4 concludes with remarks on consequences for quantitative stability of the considered two-stage models.

2 Preliminaries

2.1 Stochastic Programs with Linear Recourse

The stochastic program with recourse is among the most popular models in optimization under uncertainty, see [5], [14] for introduction into stochastic programming. In the two-stage linear recourse case there is a random linear program

$$
\min \left\{ h^\top x + q^\top y \mid Tx + Wy = z(\omega), \ y \geq 0 \right\}
$$

(1)

with all vectors and matrices of conformable (finite) dimensions. $z(\omega)$ is a random vector on some probability space $(\mathbb{R}^s, B^s, \mu)$ with $B^s$ the Borel $\sigma$-algebra on $\mathbb{R}^s$, and $\mu$ denotes a Borel probability measure on $\mathbb{R}^s$. The optimization problem (1) comes with a typical requirement in stochastic programming, nonanticipativity, which amounts to using only those information that is available upon making decisions.

This specifies the dynamics underlying (1) in the following way: After having chosen a first-stage decision $x$, the random vector $z$ is unveiled such that a well-defined optimization problem in $y$ remains. From its set of optimal solutions, a second-stage recourse solution is taken, altogether generating the random optimal value

$$
\phi(x, \omega) := \min_y \left\{ q^\top y \mid Wy = z(\omega) - Tx, \ y \geq 0 \right\}.
$$

(2)

Clearly, when referring to decision making under uncertainty in the present context, it is all about finding a “best possible” decision $x$, the only one to be taken under uncertainty (or incomplete information). Data for making the decision are “compressed” in the real numbers $\phi(x, \omega)$ where $x$ varies in some feasible subset $X$ of $\mathbb{R}^s$ and $z(\omega)$ in the support of the probability measure $\mu$. 
It will be convenient to adopt the perspective about \( \phi(x, \omega) \) of representing a random variable induced by \( x \). Finding a best \( x \) then means finding a best member in the family of random variables (= measurable functions) \( \{ \phi(x, \cdot) \mid x \in X \} \). The whole history of two-stage models in stochastic programming can be embedded into this view. The most straightforward method of ranking random variables is assigning them real numbers and resorting to the usual order relation on \( \mathbb{R} \). The most popular assignment in this respect is taking the expected value which gives the optimization problem

\[
\min \left\{ Q_\mathbb{E}(x) := \int_{\mathbb{R}^s} \phi(x, z) \mu(dz) \mid x \in X \right\}
\]

Historically, models of the above type have marked the advent of stochastic programming, [4], [7], not the least due to convexity properties they obey. Indeed, consider the value function

\[
\varphi(t) := \min_y \{ q^\top y \mid Wy = t, y \geq 0 \}.
\]

(3)

It belongs to the folklore in stochastic programming that, under mild assumptions, namely if the feasible sets in (3) for all \( t \in \mathbb{R}^s \) and the feasible set of the linear programming dual to (3) all are nonempty, one obtains by duality that \( \varphi(t) = \max_{\xi} \{ W^\top \xi \leq q \} t^\top \xi \), confirming convexity as a pointwise maximum of linear functions. Finally, linearity of the integral yields convexity of \( Q_\mathbb{E} \).

### 2.2 Strong Convexity

Convex functions have an important role in the analysis of extremal problems, more specifically and without being exhaustive, the investigation of structures, the design of algorithms, and the implementation of applications. There are various subclasses of convex functions allowing for specific conclusions, see for instance the treatment in Chapter 7 of [3] and in [6].

The present paper is devoted to one of these subclasses, namely strongly convex functions in recourse stochastic programs. A function \( f : \mathbb{R}^d \to \mathbb{R} \) defined on some convex set \( V \) is called strongly convex (with modulus \( \kappa \) on \( V \)) if there exists a real number \( \kappa > 0 \) such that for all \( x, x' \in V \) and all \( \lambda \) with \( 0 \leq \lambda \leq 1 \)

\[
f(\lambda x' + (1 - \lambda)x) \leq \lambda f(x') + (1 - \lambda) f(x) - \frac{\kappa}{2} \lambda(1 - \lambda)\|x' - x\|^2.
\]

The literature of strong convexity appears somewhat scattered. The already mentioned books [3] and [6], the early reference [8] and the section on background material in [14] all may serve as supplementary resources. In what follows we collect some material on strong convexity to indicate its importance and this way motivate our research on verification of strong convexity for a substantial class of stochastic programs.

Let \( f : \mathbb{R}^d \to \mathbb{R} \) be a convex function. Then \( f \) is strongly convex with modulus \( \kappa \) if and only if \( F(x) := f(x) - \frac{\kappa}{2} \|x\|^2 \) is convex, see [14] Page 393 for
a proof. Assume for notational convenience that \( f \) is also differentiable, and write down the following characterization of convexity for the function \( F \) at some \( x \in \mathbb{R}^d \)

\[
F(x') - F(x) \geq F'(x)(x' - x)
\]

for all \( x, x' \in \mathbb{R}^d \).

Substituting the definition of \( F \) yields

\[
f(x') - f(x) - \frac{\kappa}{2} \|x'\|^2 + \frac{\kappa}{2} \|x\|^2 \geq (f'(x) - \kappa x^T)(x' - x)
\]

for all \( x, x' \in \mathbb{R}^d \).

After simple formula manipulation one obtains

\[
f(x') \geq f(x) + f'(x)(x' - x) + \frac{\kappa}{2} \|x' - x\|^2
\]

for all \( x, x' \in \mathbb{R}^d \). (4)

This justifies the following interpretation: While, at any point of the interior of its domain, a convex function is globally supported by an affine function, a strongly convex function is so by a quadratic function with postively definite quadratic form.

Fixing \( x, x' \in \mathbb{R}^d \), letting \( x' \) take the role of \( x \) in (4), and adding up (4) implies the following characterization of strong convexity in terms of strong monotonicity of the gradient

\[
(f'(x') - f'(x))(x' - x) \geq \kappa \|x' - x\|^2
\]

for all \( x, x' \in \mathbb{R}^d \).

If point \( x \) in (4) is a minimizer of \( f \) then the gradient in (4) vanishes, leaving

\[
f(x') \geq f(x) + \frac{\kappa}{2} \|x' - x\|^2
\]

for all \( x, x' \in \mathbb{R}^d \) (5)

what often is called the quadratic growth condition.

Beside the fact that the quadratic growth condition implies uniqueness of the minimizer it is the possibility of estimating distances of arguments by differences of objective function values that is technically attractive. In particular this holds for estimating speed of convergence for minimizing sequences, for deriving quantitative stability of optimization problems, and for detecting second-order asymptotics in the Sample Average Approximation Method.

3 Sufficient Conditions for Strong Convexity of Recourse Functions

Before taking a look at the risk-neutral setting let us observe a general feature: namely using (3) one obtains

\[
Q_E(x) = \int_{\mathbb{R}^s} \varphi(z - Tx) \mu(dz).
\]

Thus, for \( Q_E \) to be strongly convex it is necessary that the null space of the matrix \( T \) is \( \{0\} \). Since this assumption is unacceptably strong, the subsequent studies will be carried out with respect to the transformed vector \( Tx \in \mathbb{R}^s \)
which we shall refer to as $x$ henceforth for the sake of simplicity. This turns
the current object of study into

$$Q_E(x) = \int_{\mathbb{R}^s} \varphi(z - x) \mu(dz).$$

with some Borel-measure $\mu$ and $\varphi(t) = \min\{q^T y \mid Wy = t, y \geq 0\}$. In \cite{13} the
following result regarding strong convexity of $Q_E$ is shown:

**Theorem 31** Assume that the following conditions are satisfied:

A1 For every $t$ there exist some $y \geq 0$ such that $Wy = t$. (Complete recourse)
A2 There exists some $\xi$ with $W^\top \xi < q$. (Strengthened sufficiently expensive
recourse)
A3 $\|z\|$ is $\mu$-integrable. (Finite first moments)
A4 $\mu$ has a density $\theta$ with respect to the Lebesgue-measure and there exists a
convex open set $V$, constants $r, \rho > 0$ such that $\theta \geq r$ a.s. on $V + B_\rho(0)$.

Then $Q_E$ is strongly convex on $V$.

The main idea of the proof is to show monotonicity of the gradient of
$Q_E$ which involves calculating the a.e. existing derivative under the integral
sign and rewriting $Q_E'$ as a Stieltjes-integral. Invoking the integration-by-parts
formula for Stieltjes-integrals (which here boils down to rearranging a finite
number of terms) yields an expression which can be estimated conveniently.
The key arguments exploit the special structure of the linearity complex of
$\varphi$ to which the following proposition from \cite{15} relates to. We shall use these
results throughout the text without further reference.

**Lemma 32** Assume A1 and that there exists some $\xi$ such that $W^\top \xi \leq q$. It
is well known from Linear Programming that under these two conditions the
linear programs

$$\min\{q^T y \mid Wy = t, y \geq 0\} \quad \text{and} \quad \max\{t^\top z \mid W^\top t \geq q\}$$

are solvable for all $t$, their optimal values coincide and the polyhedron $M_D := \{z \mid W^\top z \geq q\}$ is the convex hull of its finitely many vertices \{d_i \mid i \in I := \{1, \ldots, N\}\}. We also have

(i) $\varphi(t) = \max_{i \in I} d_i^\top t$.
(ii) $\varphi(t) = d_i^\top t$ for all $t \in K_i := \{z \mid (d_i - d_j)^\top z \geq 0 \forall j \in I\}$.
(iii) $\bigcup_{i \in I} K_i = \mathbb{R}^s$ and each intersection $K_i \cap K_j$ with $i \neq j$ is a common closed
face of dimension strictly less than $s$.
(iv) $\dim(K_i \cap K_j) = s - 1$ if and only if $d_i$ and $d_j$ are adjacent vertices of $M_D$.
(v) Every $K_i$ is a finite union of simplicial cones which can be written as
images of $\mathbb{R}^s_+$ under linear transformations induced by basis submatrices of $W$. 
We will extend the results concerning strong convexity of risk-neutral objective functions in [13] to risk-averse stochastic programs. More specifically, we will consider the expected excess over target $\eta$

$$Q_{EE}(x) := \int \max\{\eta, \varphi(z - x)\} \mu(dz)$$

and the upper semideviation

$$Q_{D+}(x) := \int \max\{Q_E(x), \varphi(z - x)\} \mu(dz).$$

Due to the similar structure of $Q_{EE}$ and $Q_{D+}$ it will be convenient to consider the function

$$Q_g(x) := \int \max\{g(x), \varphi(z - x)\} \mu(dz)$$

where $g$ is a continuously differentiable, convex function, and derive a convenient formula for $[Q'_g(x + u) - Q'_g(x)]u$ first which is then analyzed further for the special cases $g \equiv \eta$ and $g = Q_E$. Note that under assumptions $A1 - A3$ the functions $Q_g(x)$ are well defined and convex on all of $\mathbb{R}^s$.

3.1 Strong convexity of $Q_{EE}$

In order to give a formula for $[Q'_g(x + u) - Q'_g(x)]u$ in a concise way, we will use the following notations:

$$y_0(x, u) := g'(x)u, \quad y_i(u) := -d_i^*u, i \in I$$

$$I(x, u)(\tau) := \{i \in I \cup \{0\} \mid y_i(x, u) \leq \tau\}$$

$$M_g(\varphi)(x) := \{z \mid g(x) > \varphi(z - x)\}, \quad M_g(\varphi)(x) := \{z \mid g(x) < \varphi(z - x)\}$$

$$M_0(x) := M_g(\varphi)(x), \quad M_i(x) := M_g(\varphi)(x) \cap (x + K_i), i \neq 0.$$ 

We shall work with an additional assumption to ensure differentiability of $Q_g$:

A5 The sets $\{z \mid 0 = \varphi(z - x)\}$ have $\mu$-measure zero for every $x$.

Geometrically speaking, A5 holds if and only if 0 is not a vertex of $M_D$.

**Lemma 33** Assume that A1-A5 hold. Then for all $x, x + u \in V$ the following equation holds:

$$[Q'_g(x + u) - Q'_g(x)]u = \int_{\mathbb{R}} \mu\left( \bigcup_{i \in I(x, u)\{0\}} M_i(x) \setminus \bigcup_{i \in I(x + u, u)\{0\}} M_i(x + u) \right) d\tau.$$  

(7)
Proof First note that the sets
\[ \{ z \mid g(x) = \varphi(z - x) \} \]
with \( g(x) \neq 0 \) all have \( \mu \)-measure zero which is a direct consequence of conditions A1 and A2. The case \( g(x) = 0 \) is covered by assumption A5. In connection with A4 we can calculate \( Q'_g(x)u \) by differentiation under the integral sign as
\[ Q'_g(x)u = \mu(M_{g > \varphi}(x))g'(x)u - \sum_{i \neq 0} \mu(M_i(x)) d_i^T u, \quad (8) \]
Following the argument in [1], we see that for fixed \( x \) and \( u \) we can rewrite (8) as the expected value of a discrete probability distribution with values 
\[ y_0 := g'(x)u, \quad y_i := -d_i^T u \] and probabilities 
\[ \pi_0 := \mu(M_{g > \varphi}(x)), \quad \pi_i := \mu(M_i(x)) \] respectively. By transformation to the induced measure and using integration by parts for Riemann-Stieltjes integrals we arrive at
\[ [Q'_D(x + u) - Q'_D(x)]u = \int (F_x(\tau) - F_{x+u}(\tau)) d\tau, \]
where \( F_x = \sum_{i \in I_{x,u}(\tau)} \pi_i \) is the cdf of the discrete distribution just defined. Note that the boundary terms in the integration by parts-formula cancel out because \( (F_x - F_{x+u})(\tau) = 0 \) for sufficiently big \( |\tau| \).

For proving the lemma we will demonstrate that
\[ F_x(\tau) - F_{x+u}(\tau) = \mu(\bigcup_{i \in I(x,u)(\tau)} M_i(x) \setminus \bigcup_{i \in I(x+u,u)(\tau)} M_i(x + u)) \]
which can be done by showing that for all \( \tau \)
\[ \bigcup_{i \in I(x+u,u)(\tau)} M_i(x + u) \subseteq \bigcup_{i \in I(x,u)(\tau)} M_i(x). \quad (9) \]
To see why this sufficient, note that by lemma [32] and A4
\[ F_x(\tau) = \sum_{i \in I(x,u)(\tau)} \mu(M_i(x)) = \mu(\bigcup_{i \in I(x,u)(\tau)} M_i(x)) \]
and thus, if (9) holds,
\[ F_x(\tau) - F_{x+u}(\tau) = \sum_{i \in I(x,u)(\tau)} \mu(M_i(x)) - \sum_{i \in I(x+u,u)(\tau)} \mu(M_i(x + u)) \]
\[ = \mu(\bigcup_{i \in I(x,u)} M_i(x) \setminus \bigcup_{i \in I(x+u,u)} M_i(x + u)). \]
From the proof in [13, theorem 2.2] it is known that
\[ \bigcup_{i \in I(x+u,u)(\tau) \setminus \{0\}} u + K_i \subseteq \bigcup_{i \in I(x,u)(\tau) \setminus \{0\}} K_i \quad (10) \]
and we shall use this result to simplify the proof of inclusion 11.

Let \( z_0 \in \bigcup_{i \in I(x+u, u)} M_i(x + u) \setminus \bigcup_{i \in I(x, u)} M_i(x) \), then we have two cases to consider:

**case 1** \( z_0 \in M_0(x + u) \), from which follows \( g'(x + u)u \leq \tau \). By convexity of \( g \) we conclude \( g'(x)u \leq \tau \), so \( 0 \in I(x, u)(\tau) \) and we have \( z_0 \not\in M_0(x) \) by assumption. We thus have \( g(x) < \varphi(z - x) = d_{i_1}^T (z - x) \) for a suitable index \( i_1 \neq 0 \) which needs to be such that \( -d_{i_1}^T u > \tau \), otherwise \( i_1 \in I(x, u)(\tau) \). On the other hand we have

\[
\begin{align*}
g(x + u) > \varphi(z - x - u) & \geq d_{i_1}^T (z - x - u) = d_{i_1}^T (z - x) - d_{i_1}^T u \\
& > g(x) - d_{i_1}^T u \geq g(x + u) - g'(x + u)u - d_{i_1}^T u \\
& \geq g(x + u) - \tau - d_{i_1}^T u,
\end{align*}
\]

so we have \( -d_{i_1}^T u < \tau \), which is a contradiction. So case 1) is impossible.

**case 2** \( z_0 \in M_{i_0}(x + u) \) for some index \( i_0 \neq 0 \) (which means \( g(x + u) < \varphi(z - x - u) = d_{i_0}^T (z - x - u) \) and \( -d_{i_0}^T u \leq \tau \). By (10) it follows \( z_0 \not\in M_0(x) \) and by assumption we have \( g'(x)u > \tau \), from which we conclude by convexity of \( g \) that also \( g'(x + u)u > \tau \). From here we get with the same arguments as above

\[
\begin{align*}
d_{i_0}^T (z - x) - d_{i_0}^T u = d_{i_0}^T (z - x - u) & = \varphi(z - x - u) > g(x + u) \\
& \geq g(x) + g'(x)u > g(x) + \tau > \varphi(z - x) + \tau \\
& \geq d_{i_0}^T (z - x) + \tau,
\end{align*}
\]

which yields \( -d_{i_0}^T u > \tau \) in contradiction to what was stated above.

For fixed \( x \) and \( u \), the integrand on the right-hand side of (10) is piecewise constant in \( \tau \). The proof of strong convexity will consist of restricting the integration to a suitable subset where the integrand is constant in \( \tau \) and minorized by a quantity that can be estimated easier. The next lemma gives the proof of formula (2.13) in 13, theorem 2.2] concerning estimates of this integrand:

**Lemma 34** Let \( K = \{ u \mid c_j^T u \geq 0, \ j \in J \} \) be a pointed, polyhedral cone in \( \mathbb{R}^n \) (in particular \( |J| < \infty \)) with \( \text{int}(K) \neq \emptyset \). Then there is some \( \alpha > 0 \) such that

\[ \inf_{u \in K} \sup_{j \in J} c_j^T u \geq \alpha \|u\|. \]

**Proof** For arbitrary \( 0 \neq u \in K \) it holds that \( \frac{u}{\|u\|} \in K \) due to \( K \) being a cone, so it is enough to prove that

\[ \inf_{u \in K \atop \|u\| = 1} \sup_{j \in J} c_j^T u = \alpha > 0. \]
From the definition of $K$ it directly follows that $\alpha \geq 0$. If $\alpha$ was 0, we could find a sequence $\{u_n\}_{n \geq 1}$ in the compact set $K_1 := K \cap \{u \mid \|u\| = 1\}$ such that
$$\sup_j c_j^T u_n \to 0,$$
and by passing to a subsequence we had $u_{n_k} \to u_0 \in K_1$ and by continuity $c_j^T u_0 = 0$ for all indices $j$. That however would be a contradiction to the full-dimensionality of $K$ because in particular $\|u_0\| = 1$. □

The main result concerning strong convexity of the expected excess $Q_{EE}$ is given in the next theorem:

**Theorem 35 (Strong convexity for $Q_{EE}$)** Assume that $A1-A5$ hold. Then there exists some $c > 0$ such that for all $\eta < c$ we have strong convexity of the expected excess $Q_{EE}$ over target $\eta$ on the open, convex set $V$ (cf. A4).

Before the proof we shall look at the geometry of the expected excess and give a heuristic reasoning of why the theorem should hold:

The picture shows the one-dimensional case with $\varphi \geq 0$ und $\eta < 0$, so that $Q_{EE} = Q_E$. For fixed $z \in V$ the function $\varphi(z - \cdot)$ can be approximated by a $C^2$-function $\tilde{\varphi}(z - \cdot)$ which has second derivative $\varphi''(z - x) \geq C > 0$ with $x$ close to $z$. With condition A4 in mind we get
$$Q''_E(x) = \int \varphi''(z - x) \mu(\mathrm{d}z) = \int \theta(t) \varphi''(t - x) \mathrm{d}t \geq \int_{B_\rho(x)} \theta(t) \varphi''(t - x) \mathrm{d}t \geq \int_{B_\rho(x)} r C \mathrm{d}t > 0.$$

If $\eta$ is gradually increased, the graph of the integrand in (6), $\max\{\eta, \varphi(\cdot - x)\}$, is truncated as in the next picture:
As long as the "bumps" in the graph lie above the set \( V + B_\rho(0) \), the approximation argument from before yields strong convexity of \( Q_{EE} \), for the approximants of \( \max\{\eta, \varphi(z - \cdot)\} \) will be curved close to the bumps. Note that the best possible modulus of strong convexity will be smaller the greater \( \eta \) is chosen. When \( \eta \) exceeds a certain threshold the graph of \( \max\{\eta, \varphi(z - \cdot)\} \) will be flat in all of \( V + B_\rho(0) \), so strong convexity of \( Q_{EE} \) can no longer be expected to hold.

Proof (Theorem) By lemma 32 (iii) we have \( \mathbb{R}^e = \bigcup_{i \in I} K_i \) with full-dimensional, pointed polyhedral cones \( K_i \). Define \( K_i^+ = \{ z \in K_i \mid d_i^T z \geq 0 \} \) and \( K_i^- = \{ z \in K_i \mid d_i^T z \leq 0 \} \), so we can write

\[
K_i = K_i^+ \cup K_i^-,
\]

and at least one of the two cones \( K_i^+ \) and \( K_i^- \) has nonempty interior.

For arbitrary \( u \neq 0 \) we have \( u \in K_i \) for some \( i \in I \). Consider the case \( u \in K_i^+ \) when \( \text{int}(K_i^+) \neq \emptyset \) first. The other case will be reduced to this one later. Note that under assumptions A1 and A2 \( K_i \) is a polyehdral cone generated by the vectors \( d_i - d_j \), where all \( d_j \) are adjacent to \( d_i \). Lemma 34 can then be applied to yield some \( \alpha > 0 \), which - due to there being only finitely many \( K_i \) - can be chosen independently of \( u \) and \( i \), and some index \( j \) such that

\[
\alpha \| u \| \leq (d_i - d_j)^T u,
\]

from which follows

\[
-d_i^T u + \alpha \| u \| \leq -d_j^T u.
\]

By lemma 33 we get (note that the index sets \( I(x, u)(\tau) \) introduced above lemma 33 do not depend on \( x \) in the case of \( g \equiv \eta \))

\[
|Q'_{EE}(x + u) - Q'_{EE}(x)| u = \int_\mathbb{R} \mu \left( \bigcup_{\ell \in I(u)(\tau)} M_\ell(x) \setminus \bigcup_{\ell \in I(u)(\tau)} M_\ell(x + u) \right) \, d\tau \geq \int_{-\frac{1}{2}d_i^T u}^{-\frac{1}{2}d_i^T u + \alpha \| u \|} \mu \left( \bigcup_{\ell \in I(u)(\tau)} M_\ell(x) \setminus \bigcup_{\ell \in I(u)(\tau)} M_\ell(x + u) \right) \, d\tau. \quad (11)
\]

Because of \( u \in K_i^+ \) we have \( d_i^T u \geq 0 \), so there two cases to consider: \( -d_i^T u \leq 0 \leq -d_i^T u + \alpha \| u \| \) and \( -d_i^T u + \alpha \| u \| \leq 0 \). Given \( -d_i^T u \leq 0 \leq -d_i^T u + \alpha \| u \| \) it has to be \( d_i^T u \geq \frac{1}{2} \alpha \| u \| \) or \( -d_i^T u + \alpha \| u \| \geq \frac{1}{2} \alpha \| u \| \). Assuming the latter we can
Strong Convexity in Stochastic Programs with Deviation Risk Measures

continue (11) as follows:

\[ \geq \int_{0}^{\|u\|} \mu\left( \bigcup_{t \in \tilde{I}(u)(\tau)} M_t(x) \setminus \bigcup_{t \in \tilde{I}(u)(\tau)} M_t(x + u) \right) \, d\tau \]

\[ = \int_{0}^{\|u\|} \mu\left( x + K_i \setminus \bigcup_{-d_i^u < t < -d_i^u} x + u + K_i \right) \, d\tau \]

\[ \geq \int_{0}^{\|u\|} \mu\left( x + K_i \setminus \bigcup_{-d_i^u < t < -d_i^u} x + u + K_i \right) \, d\tau \]

\[ \geq \frac{1}{2} \alpha \|u\| \mu\left( x + K_i \setminus \bigcup_{-d_i^u < t < -d_i^u} x + u + K_i \right). \]

From here on it can be argued as in [13, page 10] that there exists some \( \beta > 0 \) only depending on \( i \) such that

\[ \mu\left( x + K_i \setminus \bigcup_{-d_i^u < t < -d_i^u} x + u + K_i \right) \geq \beta \|u\| \]

which concludes the analysis of the first case.

Going back to the other cases we will always have \( d_i^u \geq \frac{1}{2} \alpha \|u\| \) from here on. We estimate (11) as

\[ \geq \int_{0}^{\min(0, -d_i^u + \alpha \|u\|)} \mu\left( \bigcup_{t \in \tilde{I}(u)(\tau)} M_t(x) \setminus \bigcup_{t \in \tilde{I}(u)(\tau)} M_t(x + u) \right) \, d\tau \]

\[ \geq \int_{0}^{\min(0, -d_i^u + \alpha \|u\|)} \mu\left( M_i(x) \setminus \bigcup_{-d_i^u < t < 0} M_t(x + u) \right) \, d\tau \]

\[ \geq \frac{1}{2} \alpha \|u\| \mu\left( M_i(x) \setminus \bigcup_{-d_i^u < t < 0} M_t(x + u) \right) \]

Note that the last inequality follows because \( -d_i^u < \tau < 0 \) for each \( \tau \) in question and thus

\[ \bigcup_{t \in \tilde{I}(u)(\tau)} M_t(x) \supset M_i(x), \]

\[ \bigcup_{t \in \tilde{I}(u)(\tau)} M_t(x + u) \subset \bigcup_{-d_i^u < t < 0} M_t(x + u), \]

and the sets on the right-hand side do not depend on \( \tau \) anymore.

It remains to be shown that there exists some \( \beta > 0 \) independent of \( x \) and \( u \) such that

\[ \mu\left( M_i(x) \setminus \bigcup_{-d_i^u < t < 0} M_t(x + u) \right) \geq \beta \|u\|. \]  

(12)
Without loss of generality assume $\eta > 0$. It can easily be shown that the set

$$M^+_i(x) := (x + K^+_i) \cap \{z \mid \varphi(z - x) \geq \eta\}$$

is a full-dimensional polyhedron with vertices and with $s - 1$-dimensional facet

$$F := (x + K^+_i) \cap \{z \mid \varphi(z - x) = \eta\}.$$

The situation can be pictured like this:

![Diagram of a polyhedron and facets]

With condition A4 in mind assume $\eta$ is small enough so that for at least one vertex $y_0$ of $M^+_i(x)$ contained in $F$ we have $\|y_0 - x\| = \rho_0 < \rho$. The set

$$F_0 := F \cap B_{\rho - \rho_0}(y_0)$$

then has positive $s - 1$-dimensional measure. We will now show that

$$\tilde{F} := \bigcup_{0 \leq \lambda < 1} \lambda u + F_0 \subset (M^+_i(x) \setminus \bigcup_{-d^I u < 0} M_l(x + u)).$$

The inclusion $\bigcup_{0 \leq \lambda < 1} \lambda u + F_0 \subset M^+_i(x)$ directly follows from $u \in K^+_i$, $K^+_i$ being a convex cone and the definition of $F_0$.

For arbitrary $l$ such that $-d^I_l u < 0$, $0 \leq \lambda < 1$ and $\tilde{z} \in F_0$, which we can write as $y_0 + z'$ with $d^I_l z' = 0$, we have (note that $\tilde{z} - x \in K^+_i$)

$$d^I_l (\lambda u + \tilde{z} - x - u) = (1 - \lambda)(-d^I_l u) + d^I_l (\tilde{z} - x)$$
$$\leq (1 - \lambda)(-d^I_l u) + d^I_l (\tilde{z} - x)$$
$$< d^I_l (y_0 - x) + d^I_l z' = \eta$$

$$d^I_l (\lambda u + \tilde{z} - x - u)$$
so that $\lambda u + \tilde{z} \notin M_t(x + u)$. Thus (14) is shown.

As the last step in the proof we will show that

$$\tilde{F} \subset V + B_\rho(0),$$

(15)

$V$ being the set from condition A4, and that there is some $\beta > 0$ such that

$$\lambda(\tilde{F}) \geq \beta\|u\|.$$

(16)

From (14), (15) and (16) follows (12) as follows:

$$\mu(M_t(x) \setminus \bigcup_{-d_l^t u < 0} M_t(x + u)) \geq \mu(\tilde{F}) = \int_{\tilde{F}} g(t) \, dt \geq \lambda(\tilde{F}) \geq r\lambda(\tilde{F}) \geq r\beta\|u\|.$$

Inclusion (15) holds because $x + \lambda u \in V$ for all $0 \leq \lambda < 1$ and thus by (13)

$$\begin{align*}
\lambda u + F_0 &= (x + \lambda u) + (y_0 - x) - y_0 + F_0 \\
&\subset V + \text{cl}(B_\rho(0)) + B_{\rho - \rho_0}(0) \subset V + B_\rho(0).
\end{align*}$$

Estimate (16) is correct because we have $d_l^t u \geq \frac{1}{2} \alpha\|u\|$, $F_0$ has positive $s - 1$-dimensional volume and the density of $\mu$ is bounded below on $V + B_\rho(0)$ by a positive constant. Applying Cavalieri’s principle yields the desired result.

To get an idea of the geometry behind these arguments, a picture might be helpful:
Consider now the case \( u \in K^-_i \) when \( \text{int}(K^-_i) \neq \emptyset \): It then holds \(-u \in K^+_i\) for some \( i_0 \) with \( \text{int}(K^+_i) \neq \emptyset \). Again invoking lemma 34 yields some \( \alpha > 0 \) and some index \( j_0 \) such that \( d_{j_0} \) and \( d_{i_0} \) are adjacent with

\[
\alpha \|u\| \leq (d_{j_0} - d_{i_0})^T(-u),
\]

from which follows

\[
-d_{j_0}^T u - \alpha \|u\| \geq -d_{j_0}^T u. \tag{17}
\]

By lemma 33 we thus get

\[
[Q'_g(x + u) - Q'_g(x)]u
= \int \mu \left( \bigcup_{l \in I(u)} M_l(x) \right) \bigcap \bigcup_{\tau \in I(u, u)} M_l(x + u) \, d\tau
\geq \int -d_{j_0}^T u \\mu \left( \bigcup_{l \in I(u)} M_l(x) \right) \bigcap \bigcup_{\tau \in I(u, u)} M_l(x + u) \, d\tau.
\]

The integrand can be rewritten as follows (note that for \( i \neq j \) the sets \( M_i(y) \cap M_j(y) \) have \( \mu \)-measure 0 for all choices of \( y \):

\[
\mu \left( \bigcup_{l \in I(u)} M_l(x) \right) \bigcap \bigcup_{\tau \in I(u, u)} M_l(x + u)
= \mu \left( \left( \mathbb{R}^d \right) \bigcup_{\{l \mid y_l(u) > \tau\}} M_l(x) \right) \bigcap \left( \mathbb{R}^d \right) \bigcup_{\{l \mid y_l(u) > \tau\}} M_l(x + u)
= \mu \left( \bigcup_{\{l \mid y_l(u) > \tau\}} M_l(x + u) \right) \bigcap \bigcup_{\{l \mid y_l(u) > \tau\}} M_l(x)
\geq \mu (M_{j_0}(x + u) \cap \bigcup_{\{l \mid y_l(u) > -d_{j_0}^T u\}} M_l(x)).
\]

The last inequality holds because by (17) we have \(-d_{j_0}^T u \leq -d_{j_0}^T u - \alpha \|u\| < \tau < -d_{j_0}^T u\). By renaming \( x' = x + u, \ u' = -u \), observing \( \{l \mid y_l(u) > -d_{j_0}^T u\} = \{l \mid y_l(u') < -d_{j_0}^T (u')\} \) and noting that the lower estimate for the integrand does not depend on \( \tau \) we can continue the estimate for the integral above as

\[
[Q'_g(x' + u') - Q'_g(x')]u' \geq \alpha \|u\| \mu (M_{j_0}(x') \cap \bigcup_{\{l \mid y_l(u', u') < -d_{j_0}^T u'\}} M_l(x' + u')).
\]

where \( x', x' + u' \in V \) with \( u' \in K^+_{i_0} \) such that \( \text{int}(K_{i_0}) \neq \emptyset \). The desired estimate now follows as demonstrated previously after (11).
3.2 Strong convexity of \( Q_{D_+} \)

**Theorem 36 (Strong convexity for the upper semideviation \( Q_{D_+} \))** Under assumptions A1-A5 there exists some continuous function \( C : \mathbb{R}^n \to [0, \infty) \) with \( C(u) \to 0 \) as \( \|u\| \to 0 \), such that \( C(u) > 0 \) for all \( u \neq 0 \) and
\[
[Q'_{D_+}(x + u) - Q'_{D_+}(x)]u \geq C(u) \|u\|^2
\]
for all \( x, x + u \in V \). If we have the additional assumption
\[ A6 \ 0 \leq q \]
then \( Q_{D_+} \) is a strongly convex function on \( V \).

**Proof** Let \( x, x + u \in V \). We apply lemma 33 on \( g = Q_E \) to get
\[
[Q'_{D_+}(x + u) - Q'_{D_+}(x)]u = [Q_{Q_\delta}(x + u) - Q_{Q_\delta}(x)]u
\]
\[
= \int \mu( \bigcup_{l \in I(x, u)(\tau)} M_l(x) \setminus \bigcup_{l \in I(x + u, u)(\tau)} M_l(x + u) ) \, d\tau
\]
\[
\geq \int Q'(x + u)u \mu( \bigcup_{l \in I(x, u)(\tau)} M_l(x) \setminus \bigcup_{l \in I(x + u, u)(\tau)} M_l(x + u) ) \, d\tau. \tag{18}
\]

By A1 and A2 we can apply lemma 33 as we did in the last proof: The fan of cones \( \{K_l\} \) covers \( \mathbb{R}^n \), so \( u \in K_i \) and \(- u \in K_j \) for suitable indices \( i, j \in I \). By lemma 34 we can find some constant \( \alpha > 0 \) independent of \( x \) and \( u \) such that \((d_i - d_j)^\top u \geq \alpha \|u\|\); condition A4 and the full-dimensionality of all \( K_i \) imply that there exists some \( \delta > 0 \) such that for all \( y \in V \) and all indices \( l \in I \) we have \( \mu(y + K_i) \geq \delta \). We thus have
\[
- d_j^\top u - Q_E'(x + u)u = \sum_{i=1}^N \mu(x + u + K_i)(d_i - d_j)^\top u
\]
\[
\geq \mu(x + u + K_i)(d_i - d_j)^\top u \geq \delta \alpha \|u\|,
\]
\[
Q_E'(x)u - (-d_i^\top u) = \sum_{i=1}^N \mu(x + K_i)(d_i - d_i)^\top u
\]
\[
\geq \mu(x + K_i)(d_i - d_i)^\top u \geq \delta \alpha \|u\|. \tag{19}
\]

In 18 it is \( Q_E'(x)u < \tau < Q_E'(x + u)u \), so by 19 we get \( 0, i \in I(x, u)(\tau) \) implying for all \( \tau \) in question the inclusion
\[
x + K_i \subset \bigcup_{l \in I(x, u)(\tau)} M_l(x),
\]
and \( 0, j \notin I(x + u, u)(\tau) \), implying for such \( \tau \) (note that \( M_j(x + u) = (x + u + K_j) \cap M_{Q_{\leq q}}(x + u) \))
\[
\bigcup_{l \in I(x + u, u)(\tau)} M_l(x + u) \subset \mathbb{R}^n \setminus (M_j(x + u) \cup M_{Q_{\geq q}}(x + u))
\]
\[
= \mathbb{R}^n \setminus ((x + u + K_j) \cup M_{Q_{\geq q}}(x + u)).
\]
Combing these two inclusions we can write
\[ \bigcup_{l \in I(x,u)(\tau)} M_l(x) \setminus \bigcup_{l \in I(x,u,u)(\tau)} M_l(x + u) \supset ((x + K_i) \cap (x + u + K_j)) \cup ((x + K_i) \cap M_{Q \geq \varphi}(x + u)). \] (20)

Note that the last expression does not depend on \( \tau \) anymore. Using this and the strong convexity of \( Q \) on \( V \) (cf. theorem 31) we can continue estimating (18) as
\[ \geq \int Q'_{(x+u)\mu} \mu((x + K_i) \cap (x + u + K_j)) \, d\tau \geq \mu((x + K_i) \cap (x + u + K_j)) \cdot C \|u\|^2. \]

Set \( C_i(u) := \mu((x + K_i) \cap (x + u + K_j)) \) and \( C := \min_i C_i \) does the deal for the first claim of the theorem due to the continuity properties of \( \mu \).

Assuming A6 and returning to (18) for estimating the integrand and exploiting strong convexity of \( Q \) again yields
\[ [Q'_{D_i}(x + u) - Q'_{D_i}(x)]u \geq \max \{ \mu((x + K_i) \cap M_{Q \geq \varphi}(x + u)), \mu((x + K_i) \cap (x + u + K_j)) \} \cdot C \|u\|^2. \]

For showing strong convexity of \( Q_{D_i} \) it is thus sufficient to show that there is a constant \( c > 0 \) such that
\[ \max \{ \mu((x + K_i) \cap M_{Q \geq \varphi}(x + u)), \mu((x + K_i) \cap (x + u + K_j)) \} \geq c. \]

Assumption A6 yields \( Q(y) \geq R \) for all \( y \in V \) with some constant \( R > 0 \). Since \( \varphi(0) = 0 \) and \( \varphi \) is continuous, there are \( c_1, r_1 > 0 \) such that for all \( \|u\| < r_1 \) in \( K_i \) it is \( \mu((x + K_i) \cap M_{Q \geq \varphi}(x + u)) \geq c_1 \). Since there are only finitely many \( K_i \) that \( u \) can be contained in, we see that \( c_1 \) can actually be chosen independently of \( i \). Furthermore similar arguments as employed in the proof of lemma 34 show that \( (x + K_i) \cap (x + u + K_j) \geq c_2 > 0 \) for all \( u \) such that \( \|u\| \geq r \). With this the proof is completed.

4 Conclusion

As pointed out in section 2.2 strong convexity implies a quadratic growth condition that is instrumental in stability analysis. To illustrate this, we shall assume that the first-stage cost function is convex quadratic and view
\[ \min_x \left\{ x^T H x + h^T x + \int_{\mathbb{R}^r} \max\{g(x), \varphi(z - x)\} \mu(dz) \mid x \in X \right\} \] (21)
as a parametric problem depending on $\mu$. Note that (1) arises as a special case of (21). Consider the parameter space $(M_1^s, W_1)$, where

$$M_1^s := \left\{ \mu \in \mathcal{P}(\mathbb{R}^s) \mid \int_{\mathbb{R}^s} \max \|z\| \mu(dz) \right\}$$

is the space of Borel probability measures on $\mathbb{R}^s$ with finite moments of first order and

$$W_1(\mu, \nu) := \inf \left\{ \int_{\mathbb{R}^s \times \mathbb{R}^s} \|v - \tilde{v}\| \kappa(d(v, \tilde{v})) \mid \kappa \in \mathcal{P}(\mathbb{R}^s \times \mathbb{R}^s), \kappa \circ \pi_1^{-1} = \mu, \pi_2^{-1} = \nu \right\}$$

denotes the $L_1$-Wasserstein distance. In this setting, the risk neutral case arises if $g \equiv -\infty$ and has been analyzed in [11]. Following the same lines, we shall consider the optimal solution set mapping $\Psi : M_1^s \rightarrow \mathbb{R}^n$ given by

$$\Psi(\mu) := \text{Argmin}_x \left\{ x^\top Hx + h^\top x + \int_{\mathbb{R}^s} \max\{g(x), \varphi(z - Tx)\} \mu(dz) \mid x \in X \right\}.$$

**Proposition 41 (Quantitative stability for $Q_{EE}$ and $Q_{D_+}$) Assume A1-A5 as well as the following**

(a) $H$ is symmetric and positive semidefinite.
(b) $g$ is given by a sufficiently small constant $\eta$ (cf. theorem 35) or by the expectation $Q_{EE}$. In the latter case, additionally assume A6.
(c) $\mu \in M_1^s$ is such that $\Psi(\mu) \neq \emptyset$ is bounded and $T\Psi(\mu) \subseteq V$ (cf. A4).

Then there exist constants $L, \delta > 0$ such that $\Psi(\nu) \neq \emptyset$ and

$$d_H(\Psi(\mu), \Psi(\nu)) \leq LW_1(\mu, \nu)^{\frac{1}{2}}$$

for any $\nu \in M_1^s$ satisfying $W_1(\mu, \nu) \leq \delta$.

**Proof** The proof for the risk neutral case given in [11, theorem 2.7] can be extended with minor modifications whenever $g$ is Lipschitz continuous and the mapping

$$x \mapsto \int_{\mathbb{R}^s} \max\{g(x), \varphi(z - x)\} \mu(dz)$$

is strongly convex on $V$. Thus, the corollary follows from theorem 35 and theorem 36.

**Remark 1** Following the lines of [12, theorem 2.4] the $L_1$-Wasserstein distance $W_1(\mu, \nu)^{\frac{1}{2}}$ in (22) in theorem 41 can be replaced with the subgradient distance

$$d_g(\mu, \nu, \text{cl } V) := \sup\{\|z^*\| \mid z^* \in \partial(Q^\nu_{EE} - Q^\mu_{EE})(Tx), x \in \text{cl } V\},$$

where $\partial(Q^\nu_{EE} - Q^\mu_{EE})$ denotes the Clarke’s subdifferential of

$$x \mapsto \int_{\mathbb{R}^s} \max\{g(x), \varphi(z - x)\} (\nu - \mu)(dz).$$

Furthermore, a generalization of theorem 41 to an appropriate Fortet-Mourier metric (see [11, Chapter 6]) seems possible.
4.1 Future research

The next goal in our research is to prove strong convexity results for a broader class of deviation risk measures that can be represented as convex integral functionals. In this paper we presented a relatively self-contained first-order analysis of two prototypical examples. The geometry of the recourse function played an important role throughout the proofs. A slightly different approach promising good results seems to be a second-order analysis via smooth, strongly convex approximations (see for example [2]) of the recourse-function. With this it seems to be possible to drop assumption A6 completely and take a step towards more involved deviation risk measures. Future research might also involve duality theory of convex integral functionals and generalized (first and second order) derivatives (see [10] for an early reference) as technical tools and a greater emphasis on stability theory.

Acknowledgements The authors would like to thank the Deutsche Forschungsgemeinschaft for supporting the first and second author via the Collaborative Research Center TRR 154.

References

1. Attouch H, Wets R J-B (1993) Quantitative Stability of Variational Systems II. A Framework for Nonlinear Conditioning. SIAM J Optim 3: 359-381
2. Azagra D (2013) Global and fine approximation of convex functions. Proceedings of the London Mathematical Society 107: 799-824
3. Bauschke H, Combettes P (2011) Convex Analysis and Monotone Operator Theory in Hilbert Spaces. Springer, New York
4. Beale E M L (1955) On Minimizing a Convex Function subject to Linear Inequalities. Journal of the Royal Statistical Society, Series B 17: 173-184
5. Birge J R, Louveaux F (2011) Introduction to Stochastic Programming, Second Edition, Springer, New York
6. Borwein, J M, Vanderwerff J D (2010) Convex Functions - Constructions, Characterizations, and Counterexamples. Cambridge University Press, Cambridge
7. Dantzig, G B (1955) Linear Programming under Uncertainty. Management Science 1: 197-206
8. Poljak B (1966) Existence Theorems and Convergence of Minimizing Sequences in Extremum Problems with Restrictions. Dokl. Akad. Nauk SSSR 166: 287 – 290; Soviet Math. Dokl 7 : 72-75
9. Rachev S T (1991) Probability Metrics and the Stability of Stochastic Models, Wiley, Chichester et al.
10. Rockafellar, R T (1968) Integrals which are convex functionals. Pacific J. Math. 24, no. 3: 525-539
11. Römisch W, Schultz R (1991) Stability Analysis for Stochastic Programs. Annals of Operations Research 30: 241-266
12. Römisch W, Schultz R (1996) Lipschitz Stability for Stochastic Programs with Complete Recourse. SIAM J Optim 6(2): 531-547
13. Schultz, R (1994) Strong Concexity in Stochastic Programs with Complete Recourse. Journal of Computational and Applied Mathematics 56: 3-22
14. Shapiro A, Dentcheva D, Ruszczyński A (2014) Lectures on Stochastic Programming - Modeling and Theory. Second Edition, MOS-SIAM, Philadelphia.
15. Walkup, D W, Wets, R J-B (1969) Lifting projections of convex polyhedra. Pacific Journal of Mathematics 28: 465-475