Strong Convexity and Directional Derivatives of Marginal Values in Two–Stage Stochastic Programming *

Darinka Dentcheva[1], Werner Römisch[1] and Rüdiger Schultz[2]

[1] Humboldt-Universität Berlin, Institut für Angewandte Mathematik, Unter den Linden 6, D-10099 Berlin

[2] Konrad-Zuse-Zentrum für Informationstechnik Berlin, Heilbronner Str. 10, D-10711 Berlin

Abstract. Two-stage stochastic programs with random right-hand side are considered. Verifiable sufficient conditions for the existence of second-order directional derivatives of marginal values are presented. The central role of the strong convexity of the expected recourse function as well as of a Lipschitz stability result for optimal sets is emphasized.

Keywords. Two-stage stochastic programs, directional derivatives of marginal values, strong convexity, sensitivity analysis

1991 Mathematics Subject Classification: 90C15, 90C31

1 Introduction

Consider the following two-stage stochastic program

$$(1.1) \qquad \min\{g(x) + Q_\mu(Ax) : x \in C\}$$

$$(1.2) \qquad Q_\mu(\chi) = \int_{\mathbb{R}^s} \tilde{Q}(z - \chi)\mu(\mathrm{d}z),$$

$$(1.3) \qquad \tilde{Q}(t) = \min\{q^\top y : Wy = t, y \geq 0\}$$

where $g : \mathbb{R}^m \to \mathbb{R}$ is a convex function, $C \subset \mathbb{R}^m$ is a non-empty closed convex set and μ is a Borel probability measure on $\mathbb{R}^s$. Furthermore, $q \in \mathbb{R}^{\overline{m}}$ and $A \in L(\mathbb{R}^m, \mathbb{R}^s), W \in L(\mathbb{R}^{\overline{m}}, \mathbb{R}^s)$. To have (1.1.)-(1.3) well-defined we assume

*This research has been supported by the Schwerpunktprogramm " Anwendungsbezogene Optimierung und Steuerung" of the Deutsche Forschungsgemeinschaft

(A1) $\quad W(\mathbb{R}^{\overline{m}}_{+}) = \mathbb{R}^s$ (complete recourse),

(A2) $\quad M_D := \{u \in \mathbb{R}^s : W^\top u \leq q\} \neq \emptyset$ (dual feasibility),

(A3) $\quad \int\limits_{\mathbb{R}^s} \|z\|\mu(\mathrm{d}z) < +\infty$ (finite first moment).

By linear programming duality, (A1) together with (A2) implies that $\tilde{Q}(t) \in \mathbb{R}$ for all $t \in \mathbb{R}^s$. Due to (A3) also the integral in (1.2) is finite (cf. [10], [24] and the beginning of Section 2).

The model (1.1)-(1.3) is derived from an optimization problem with uncertain data, where some evidence on the probability distribution of the random data is at hand or has been gained on the basis of statistical information. We have a first-stage decision x to be made *here and now* (i.e. before the realization of z), and a second-stage decision (recourse action) y that has to be fixed after the realization of the random parameters. (1.1) then aims at fixing an x that minimizes the sum of the first-stage costs and the expected second-stage costs caused by the corrective action y. Further details and fundamental properties of (two-stage) stochastic programs can be found in [10], [24].

The present paper contributes to the stability analysis of (1.1) if μ (and hence Q_μ) is subjected to perturbations. We consider (1.1) with convex functions $v : \mathbb{R}^s \to \mathbb{R}$ instead of Q_μ and study the optimal (marginal) value φ of (1.1) as a function of v. Resorting to convex perturbations v is motivated by the fact that , given (A1) and (A2), Q_μ is convex for any probability measure μ fulfilling (A3) (cf. [10], [24] and Section 2).

Our investigations focus on second-order directional derivatives of the marginal-value function φ. In [18], [19], [20] such objects are considered for general parametric optimization problems. Lacking smoothness and non-uniqueness of optimal solutions prevent a direct application of the techniques from [19], [20] in the present setting. In contrast to the very general paper [18] we do not utilize a second-order strong stability condition imposed there. Our independent approach uses ideas from [18], [20] and is essentially based on a Lipschitz-stability result for optimal solutions ([16]) and on the strong (strict) convexity of Q_μ ([15],[17]). Accent is placed on ending up with conditions that are verifiable for the problem class (1.1)-(1.3). The issue of first-order directional differentiability of φ in the context of two-stage stochastic programs is essentially settled in [7], [22]. For the reader's convenience we display a central result in this respect. In the general non-linear programming context, first-order directional derivatives of marginal values are addressed in [8], [14], cf. also the references therein.

Our paper is organized as follows: In Section 2 we review improved convexity properties (strict and strong convexity) of Q_μ that were established in [17]. As essential prerequisites, these properties enter Section 3, where we analyze the second-order directional differentiability of Q_μ.

2 Strong Convexity

Given (A1) and (A2), linear programming duality implies

$$\tilde{Q}(t) = \max\{t^\top u : W^\top u \leq q\}. \tag{2.1}$$

Moreover, M_D is also bounded, i.e. it has the vertices $\mathrm{d}_i (i = 1, \ldots, l)$. By (2.1) we obtain

$$\tilde{Q}(t) = \max_{i=1,\ldots,l} \mathrm{d}_i^\top t,$$

i.e. $\tilde{Q}$ is piecewise linear and convex.

Together with (A3) this implies that Q_μ is a real-valued convex function on $I\!R^s$. It is natural to ask for stronger properties of Q_μ. Concerning smoothness there are sufficient conditions for Q_μ to be (twice) continuously differentiable: If μ has a density, then Q_μ is continuously differentiable and

$$\nabla Q_\mu(\chi) = \sum_{i=1}^{l} (-\mathrm{d}_i)\mu(\chi + K_i), \tag{2.2}$$

where $K_i\,(i = 1, \ldots, l)$ denotes the normal cone to M_D at d_i (cf. [10], [24]). The function Q_μ is twice continuously differentiable if $\mu \circ B$ has a continuously differentiable distribution function for any nonsingular transformation $B \in L(I\!R^s, I\!R^s)$ (for details consult [11], [23] and [15]).

In the present paper we focus on improved convexity properties. Recall that Q_μ is strictly convex if the convexity inequality holds strictly for different arguments; Q_μ is called strongly convex on a convex subset $V \subset I\!R^s$ if there exists a constant $\kappa > 0$ such that for all $\chi_1, \chi_2 \in V, \lambda \in [0, 1]$

$$Q_\mu(\lambda\chi_1 + (1-\lambda)\chi_2) \leq \lambda Q_\mu(\chi_1) + (1-\lambda)Q_\mu(\chi_2) - \kappa\lambda(1-\lambda)||\chi_1 - \chi_2||^2.$$

The strong convexity of Q_μ on a convex subset V is equivalent to the strong monotonicity of the gradient ∇Q_μ on V and to the positive definiteness of the Hessian $\nabla^2 Q_\mu$ on V, i.e. $\langle \nabla^2 Q_\mu(\chi)h, h\rangle \geq 2\kappa||h||^2$ for all $\chi \in V, h \in I\!R^s$ (cf. [12]).

Let us consider two illustrative examples to provide some initial insight into the situation.

Example 2.1 *Let* $\tilde{Q}(t) = \min\{y^+ + y^- : y^+ - y^- = t,\ y^+ \geq 0, y^- \geq 0\}$ *and* $\mu \in \mathcal{P}(I\!R)$ *be given by the density*

$$\Theta(\tau) = \begin{cases} 4|\tau| & \text{if } -\frac{1}{2} \leq \tau \leq \frac{1}{2} \\ 0 & \text{otherwise} \end{cases}.$$

Linear programming duality yields

$$\tilde{Q}(t) = \max\{t\,u : -1 \leq u \leq 1\} = |t|.$$

A straightforward calculation then provides

$$Q_\mu(\chi) = \begin{cases} \frac{4}{3}|\chi|^3 + \frac{1}{3} & \textit{if } -\frac{1}{2} \le \chi \le \frac{1}{2} \\ |\chi| & \textit{otherwise} \end{cases} .$$

Two conclusions can be drawn from this representation. Firstly, Q_μ is piecewise linear and, hence, not strictly convex outside the support of μ. Secondly, inside the support of μ the function Q_μ is strictly convex but not strongly convex since the second derivative vanishes at $\chi = 0$.

Example 2.2 *Let* $W = \begin{pmatrix} 1 & 0 & -1 & 0 \\ 1 & 1 & 1 & -1 \end{pmatrix}$, $q = (1,-1,1,1)^\top$ *and assume that* $\mu \in \mathcal{P}(I\!R^2)$ *fulfils* (A3) *and has a density.*

It is easy to see that (A1) *is fulfilled and that* $M_D = conv\left((-2,-1)^\top, (2,-1)^\top\right)$. *However, in view of* (2.2) *the second component of* $\nabla Q_\mu(\chi)$ *is identical -1 for all* $\chi \in I\!R^s$. *Therefore,* ∇Q_μ *cannot be strictly monotone and* Q_μ *is not strictly convex.*

The following theorems (proved in [17]) give positive answers with respect to the improved convexity of Q_μ. Roughly speaking, it suffices to eliminate the pathologies encountered above.

Theorem 2.3 *Assume* (A1), (A3) *and*

(A2)* $\quad$ $int\, M_D \neq \emptyset$,

(A4) $\quad$ μ *has a density.*

Then Q_μ is strictly convex on any open convex subset of the support of μ.

Theorem 2.4 *Assume* (A1), (A2)* , (A3) *and*

(A4)* $\quad$ *there exists an open convex set $V \subset I\!R^s$, constants $r > 0, \rho > 0$, and a density Θ_μ of μ such that*

$$\Theta_\mu(\tau) \ge r \quad \textit{for all} \quad \tau \in I\!R^s \quad \textit{with}\ \mathrm{dist}(\tau, V) \le \rho.$$

Then Q_μ is strongly convex on V.

In addition to these two theorems it is shown in [17] that under (A1)-(A4) the assumption (A2)* is also necessary for the strict convexity of Q. There are instances where (A2)* becomes especially handy: For simple recourse (i.e. $W = (I, -I)$, where I denotes the identity in $I\!R^s$) it is equivalent to $q^+ + q^- > 0$ (componentwise), where $q^\top = (q^{+\top}, q^{-\top}), q^+, q^- \in I\!R^s$; in case $W \in L(I\!R^{s+1}, I\!R^s)$ fulfils (A1) and (A2) is valid, (A2)* is equivalent to $q \notin W^\top(I\!R^s)$ (for details consult [17]).

3 Directional Derivatives of Marginal Values

Consider perturbations

$$(3.1) \qquad \min\{g(x) + v(Ax) : x \in C\}$$

of (1.1), where $v : \mathbb{R}^s \to \mathbb{R}$ is a convex function. We are interested in the directional behaviour of the value function

$$\varphi(v) := \inf\{g(x) + v(Ax) : x \in C\}$$

at Q_μ into convex directions.

Let

$$\psi(v) := \text{argmin}\,\{g(x) + v(Ax) : x \in C\}.$$

The following Lipschitz result on ψ with respect to the Hausdorff distance d_H will become a fundamental prerequisite.

Theorem 3.1 *Assume* (A1)-(A3) *and let* $\psi(Q_\mu)$ *be non-empty, bounded. Let* g *be convex quadratic,* C *be convex polyhedral, and* Q_μ *be strongly convex on some open convex set* $V \supset A(\psi(Q_\mu))$.

Then, for each convex function $v : \mathbb{R}^s \to \mathbb{R}$ *there exist constants* $L > 0, \delta > 0$ *such that*

$$\mathrm{d}_H(\psi(Q_\mu), \psi(Q_\mu + tv)) \leq Lt$$

whenever $0 < t < \delta$.

Proof: The proof splits into two parts: First one has to show that $\psi(Q_\mu + tv) \neq \emptyset$ for $t > 0$ sufficiently small and then the Lipschitz rate has to be established. Guidelines for both parts are given by results in [16] (Proposition 2.3, Theorem 2.4). However, in [16] perturbations of Q_μ are of the type Q_ν with $\nu \in \mathcal{P}(\mathbb{R}^s)$. Therefore, some preparation is needed for drawing conclusions from [16] in case the perturbations are of the type $Q_\mu + tv$.

The analysis in [16] is based on a subgradient distance d which, in the present setting, reads as follows: Let $U = \text{cl}\, U_0$, where U_0 is an open convex bounded set such that $\psi(Q_\mu) \subset U_0$ and $A(U) \subset V$, for convex $Q : \mathbb{R}^s \to \mathbb{R}$ we define

$$(3.2) \qquad \mathrm{d}(Q, Q_\mu; U) := \sup\{\|z^*\| : z^* \in \partial(Q - Q_\mu)(Ax) : x \in U\}$$

where ∂ denotes the Clarke subdifferential ([5]).

Inserting $Q = Q_\mu + tv$ $(t > 0)$ into the above relation yields

$$\mathrm{d}(Q_\mu + tv, Q_\mu; U) := t \cdot \sup\{\|z^*\| : z^* \in \partial v(Ax) : x \in U\}$$

where ∂ specifies to the subdifferential of convex analysis.

Since U is compact, we have

$$L_0 := \sup\{\|z^*\| : z^* \in \partial v(Ax) : x \in U\} < +\infty$$

and

$$d(Q_\mu + tv, Q_\mu; U) = L_0 \cdot t.$$

Re-interpreting Proposition 2.3. and Theorem 2.4 in [16] in terms of the definition (3.2) now provides that $\psi(Q_\mu + tv)$ is non-empty for $t > 0$ sufficiently small and that there exist constants $L > 0, \delta > 0$ such that

$$d_H(\psi(Q_\mu), \psi(Q_\mu + tv)) \leq L \cdot t$$

whenever $0 < t < \delta$. □

Remark 3.2 Note that the above results (and those to follow) remain valid if v is replaced by $v - Q_\mu$. Both the assumptions on g, C and the strong convexity of Q_μ are indispensable in the above theorem. This is illustrated by several examples in [16]. To give an idea we quote the one justifying the polyhedrality assumption on C: Let $m := 2, s := 1, g(x) \equiv 0, A := (1,0), q := (1,1)^\mathsf{T}, W := (1,-1), C := \{x = (x_1, x_2) \in \mathbb{R}^2 : (x_2)^2 \leq x_1\}$ and μ be the uniform distribution on $\left[-\frac{1}{2}, \frac{1}{2}\right]$. Then $Q_\mu(Ax) = x_1^2 + \frac{1}{4}$ for $0 \leq x_1 \leq \frac{1}{2}$ and $\psi(Q_\mu) = \{(0,0)\}$. Let further $v := Q_{\delta_1} - Q_\mu$ where δ_1 is the measure putting unit mass at $z = 1$. Then $d_H(\psi(Q_\mu), \psi(Q_\mu + tv)) \geq \sqrt{\frac{t}{2}}$ for $t > 0$ sufficiently small.

Another precondition for the subsequent second-order analysis consists in a first-order directional differentiability result for φ as obtained, for instance, in [7], [22].

Theorem 3.3 *Assume* (A1)-(A3) *and let $\psi(Q_\mu)$ be non-empty, bounded. Then φ is (Gateaux) directionally differentiable at Q_μ in any convex direction $v : \mathbb{R}^s \to \mathbb{R}$ and it holds*

$$\varphi'(Q_\mu; v) := \lim_{t \to 0+} \frac{1}{t}(\varphi(Q_\mu + tv) - \varphi(Q_\mu)) = \min\{v(Ax) : x \in \psi(Q_\mu)\}.$$

Conclusion 3.4 *If Q_μ is strictly convex on some open convex neighbourhood of $A(\psi(Q_\mu))$ (cf. Theorem 3.3) we obtain for all convex $v : \mathbb{R}^s \to \mathbb{R}$*

$$\varphi'(Q_\mu; v) = v(\chi_*) \quad \textit{where} \quad A(\psi(Q_\mu)) = \{\chi_*\}.$$

Remark 3.5 Provided that C is bounded, techniques from [15] (Proposition 2.1) can be utilized to establish that the marginal-value function φ is locally Lipschitzian (with respect to the uniform distance on $A(C)$) at any convex function v. Then, the directional differentiability of φ in the sense of Hadamard (cf. [21]) is a direct consequence of Proposition 3.5 in [21].

In what follows we explore whether φ has second-order directional derivatives at (certain) Q_μ.

Theorem 3.6 *Assume* (A1)-(A3) *and let $\psi(Q_\mu)$ be non-empty, bounded. Let g be convex quadratic, C be convex polyhedral, Q_μ be strongly convex on some open convex neighbourhood of $A(\psi(Q_\mu))$ and twice continuously differentiable at χ_* with $A(\psi(Q_\mu)) = \{\chi_*\}$. Then we have for all convex $v : \mathbb{R}^s \to \mathbb{R}$ and any $x \in \psi(Q_\mu)$*

$$(3.3) \qquad \varphi''(Q_\mu; v) = \lim_{t\to 0+} \frac{1}{t^2}(\varphi(Q_\mu + tv) - \varphi(Q_\mu) - t\varphi'(Q_\mu; v))$$

$$= \inf\{\frac{1}{2}\langle Hy, y\rangle + \frac{1}{2}\langle \nabla^2 Q_\mu(\chi_*)Ay, Ay\rangle + v'(\chi_*; Ay) : y \in S(x)\}$$

where $S(x) = \{y \in \mathbb{R}^m : y \in T_C(x), \nabla g(x)y + \nabla Q_\mu(\chi_*)Ay = 0\}$, $T_C(x) = \liminf_{t\to 0+} \frac{1}{t}(C - x)$ *is the tangent cone to* C *at* x *(with set convergence in Kuratowski's sense* [1]*),* $H := \nabla^2 g(x)$ *and* $v'(\chi_*, .)$ *denotes the directional derivative of* v *at* χ_*.

Observe that, given the function v*, the value of* $\varphi''(Q_\mu; v)$ *is the same for all* $x \in \psi(Q_\mu)$*. Moreover, the infimum in* (3.3) *is attained.*

Proof: Let $v : \mathbb{R}^s \to \mathbb{R}$ be convex and $x \in \psi(Q_\mu)$. Let $L > 0$ and $\delta > 0$ be as in Theorem 3.1 and $t \in (0, \delta)$. By Theorem 3.1 there exists an $x(t) \in \psi(Q_\mu + tv)$ such that $\|x(t) - x\| \le Lt$.

It holds

$$\varphi(Q_\mu + tv) - \varphi(Q_\mu) - t\varphi'(Q_\mu; v) =$$

$$= g(x(t)) + Q_\mu(Ax(t)) + tv(Ax(t)) - g(x) - Q_\mu(Ax) - tv(Ax)$$

$$= \nabla g(x)(x(t) - x) + \frac{1}{2}\langle H(x(t) - x), x(t) - x\rangle +$$

$$+\nabla Q_\mu(Ax)(A(x(t) - x)) + \frac{1}{2}\langle \nabla^2 Q_\mu(Ax)(A(x(t) - x)), A(x(t) - x)\rangle$$

$$+t(v(Ax(t)) - v(Ax)) + o(\|x(t) - x\|^2)$$

where we have used Theorem 3.3 for the first identity and the twice differentiability of g at x and Q_μ at $Ax = \chi_*$ for the second identity, respectively. Moreover, the above remarks imply that $o(\|x(t) - x\|^2) = o(t^2)$. This provides for all $t \in (0, \delta)$

$$\frac{1}{t^2}(\varphi(Q_\mu + tv) - \varphi(Q_\mu) - t\varphi'(Q_\mu; v)) =$$

$$= \frac{1}{t^2}(\nabla g(x)(x(t) - x) + \langle A^\top \nabla Q_\mu(Ax), x(t) - x\rangle) +$$

$$+\frac{1}{2}\langle H(x(t) - x), x(t) - x\rangle +$$

$$+\frac{1}{2}\langle A^\top \nabla^2 Q_\mu(\chi_*)A(\frac{1}{t}(x(t) - x)), \frac{1}{t}(x(t) - x)\rangle +$$

$$+\frac{1}{t}(v((Ax(t)) - v(Ax)) + o(1).$$

The optimality of x implies

$$\langle \nabla g(x) + A^\top \nabla Q_\mu(Ax), x(t) - x\rangle \geq 0.$$

Now take $t_k \underset{k\to\infty}{\longrightarrow} 0+$ in such a way that

$$\liminf_{t\to 0+} \frac{1}{t^2}(\varphi(Q_\mu + tv) - \varphi(Q_\mu) - t\varphi'(Q_\mu; v)) =$$

$$= \lim_{k\to\infty} \frac{1}{t_k^2}(\varphi(Q_\mu + t_k v) - \varphi(Q_\mu) - t_k\varphi'(Q_\mu; v)).$$

By $\|\frac{1}{t_k}(x(t_k) - x)\| \leq L$ for $k \in I\!N$ sufficiently large, there exists a subsequence $\{t_k\}_{k\in I\!N'}$ such that

$$y_k := \frac{1}{t_k}(x(t_k) - x) \underset{k\in I\!N'}{\longrightarrow} y.$$

Now $y \in T_C(x)$ and $x(t_k) = x + t_k y_k$ for all $k \in I\!N'$. Theorem 3.3 yields

$$\begin{aligned} v(Ax) &= \varphi'(Q_\mu; v) = \lim_{\substack{k\to\infty \\ k\in I\!N'}} \frac{1}{t_k}(\varphi(Q_\mu + t_k v) - \varphi(Q_\mu)) \\ &= \lim_{\substack{k\to\infty \\ k\in I\!N'}} \frac{1}{t_k}(g(x + t_k y_k) + (Q_\mu + t_k v)(A(x + t_k y_k)) - g(x) - Q_\mu(Ax)) \\ &= \nabla g(x) y + \nabla Q_\mu(Ax) A y + v(Ax). \end{aligned}$$

The above relation implies

$$\nabla g(x) y + \nabla Q_\mu(Ax) A y = 0,$$

thus $y \in S(x)$.
Therefore

$$\lim_{\substack{k\to\infty \\ k\in I\!N'}} \frac{1}{t_k^2}(\varphi(Q_\mu + t_k v) - \varphi(Q_\mu) - t_k\varphi'(Q_\mu; v)) \geq$$

$$\geq \frac{1}{2} \lim_{\substack{k\to\infty \\ k\in I\!N'}} \Big(\langle H y_k, y_k\rangle + \langle A^\top \nabla^2 Q_\mu(\chi_*) A y_k y_k\rangle +$$

$$+ \frac{1}{t_k}(v(Ax + t_k A y_k) - v(Ax)) \Big)$$

$$= \frac{1}{2}\langle Hy, y\rangle + \frac{1}{2}\langle \nabla^2 Q_\mu(\chi_*) Ay, Ay\rangle + v'(\chi_*; Ay)$$

$$\geq \inf_{y\in S(x)} \left\{ \frac{1}{2}\langle Hy, y\rangle + \frac{1}{2}\langle \nabla^2 Q_\mu(\chi_*) Ay, Ay\rangle + v'(\chi_*; Ay) \right\}.$$

Hence

$$\liminf_{t\to 0+} \frac{1}{t^2}(\varphi(Q_\mu + tv) - \varphi(Q_\mu) - t\varphi'(Q_\mu; v)) \geq$$

$$\geq \inf_{y \in S(x)} \left\{ \frac{1}{2}\langle Hy, y\rangle + \frac{1}{2}\langle \nabla^2 Q_\mu(\chi_*)Ay, Ay\rangle + v'(\chi_*; Ay) \right\}.$$

Now we establish the reverse inequality for the limes superior.

To this end, let $y \in S(x)$ be arbitrary, i.e., in particular , $y \in T_C(x)$. The polyhedrality of C now implies that, given a sequence $\{t_k\}$ with $t_k \to 0+$, we have $x + t_k y \in C$ for sufficiently large k. This allows the following estimate

$$\varphi(Q_\mu + t_k v) - \varphi(Q_\mu) - t_k\varphi'(Q_\mu; v) \leq$$

$$\leq g(x + t_k y) + Q_\mu(A(x + t_k y)) + t_k v(A(x + t_k y)) - g(x) - Q_\mu(Ax) - t_k v(Ax)$$

$$= t_k \nabla g(x)y + \frac{1}{2}t_k^2\langle Hy, y\rangle + t_k \nabla Q_\mu(Ax)Ay$$

$$+\frac{1}{2}t_k^2\langle \nabla^2 Q_\mu(Ax)Ay, Ay\rangle + o(t_k^2) + t_k(v(A(x + t_k y)) - v(Ax))$$

$$= \frac{1}{2}t_k^2\langle Hy, y\rangle + \frac{1}{2}t_k^2\langle \nabla^2 Q_\mu(Ax)Ay, Ay\rangle + o(t_k^2)+$$

$$+t_k(v(A(x + t_k y)) - v(Ax)).$$

The last identity is valid since $y \in S(x)$ implies

$$\nabla g(x)y + \nabla Q(Ax)Ay = 0.$$

Now we obtain

$$\limsup_{k\to\infty} \frac{1}{t_k^2}(\varphi(Q_\mu + t_k v) - \varphi(Q_\mu) - t_k\varphi'(Q_\mu; v))$$

$$\leq \frac{1}{2}\langle Hy, y\rangle + \frac{1}{2}\langle \nabla^2 Q_\mu(Ax)Ay, Ay\rangle + v'(Ax; Ay).$$

Since $y \in S(x)$ was arbitrary, (3.3) is established.

To prove that the infimum is actually attained, let us denote

$$h(y) := \frac{1}{2}\langle Hy, y\rangle + \ell(Ay),$$

where

$$\ell(\chi) := \frac{1}{2}\langle \nabla^2 Q_\mu(\chi_*)\chi, \chi\rangle + v'(\chi_*, \chi).$$

We will show that the function h is constant on each common direction of recession of h and $S(x)$. Theorem 27.3 in [13] then states that h attains its infimum over $S(x)$.

Let $u \in \mathbb{R}^m$ be a common direction of recession of h and $S(x)$, i.e. $u \in \mathbb{R}^m$ fulfils

$$y + \lambda u \in S(x) \quad \text{and} \quad h(y + \lambda u) \leq h(y)$$

$$\text{for all } \lambda \geq 0 \text{ and all } y \in S(x).$$

Since $S(x)$ is a polyhedral cone, $u \in \mathbb{R}^m$ is a direction of recession of $S(x)$ if and only if $u \in S(x)$.
Let $u \in S(x)$ and $\lambda \geq 0$. It holds

$$h(y+\lambda u) = \frac{1}{2}\langle Hy, y\rangle + \lambda\langle Hu, y\rangle + \frac{\lambda^2}{2}\langle Hu, u\rangle + \ell(Ay + \lambda Au).$$

By assumption, the function ℓ is strongly convex on $\mathbb{R}^s$ and, hence, obeys a unique minimizer $\bar{\chi}$ on $AS(x)$. Moreover, we have, with a suitable constant $\alpha > 0$

$$\ell(\chi) \geq \ell(\overline{\chi}) + \alpha\|\chi - \overline{\chi}\|^2 \quad \text{for all} \quad \chi \in AS(x) \quad ([12]).$$

Therefore,

$$\ell(Ay+\lambda Au) \geq \ell(\overline{\chi}) + \alpha(\|Ay - \overline{\chi}\|^2 + 2\lambda\langle Ay - \overline{\chi}, Au\rangle + \lambda^2\|Au\|^2).$$

For $Au \neq 0$ this implies $\ell(Ay + \lambda Au) \underset{\lambda\to\infty}{\longrightarrow} \infty$. Together with $\langle Hu, u\rangle \geq 0$ we obtain $h(y+\lambda u) \underset{\lambda\to\infty}{\longrightarrow} \infty$, i.e. u is no direction of recession of h.
In case $Au = 0$ and $\langle Hu, u\rangle > 0$ we again obtain $h(y+\lambda u) \underset{\lambda\to\infty}{\longrightarrow} \infty$, showing that u is no direction of recession of h.
It remains to check the case where $Au = 0$ and $\langle Hu, u\rangle = 0$. Then we have $Hu = 0$, yielding $h(y+\lambda u) = \frac{1}{2}\langle Hy, y\rangle$ for all $\lambda \geq 0$, i.e. h is constant in direction u and Theorem 27.3 in [13] works. □

Example 3.7 *Let* $m := s := 1, g(x) \equiv 0, A := 1, C := \mathbb{R}$ *and select* Q_μ *as in Example* 2.1. *Then it holds* $\psi(Q_\mu) = \{0\}, \varphi(Q_\mu) = \frac{1}{3}$. *With* $v(x) := -x \quad (x \in \mathbb{R})$ *we obtain for all* $t \in [0,1]$

$$\psi(Q_\mu + tv) = \{x \in \mathbb{R} : Q'_\mu(x) = t\} = \left\{\frac{1}{2}\sqrt{t}\right\},$$

$$\varphi(Q_\mu + tv) = \frac{1}{3}(1 - t^{\frac{3}{2}})$$

and

$$\varphi'(Q_\mu; v) = \min\{v(x) : x \in \psi(Q_\mu)\} = 0.$$

Thus

$$\frac{1}{t^2}(\varphi(Q_\mu + tv) - \varphi(Q_\mu) - t\varphi'(Q_\mu; v)) = -\frac{1}{3}t^{-\frac{1}{2}}$$

for all $t \in [0,1]$.
Hence, φ *has no second-order directional derivative at* Q_μ *in direction* v. *Note that there exists a neighbourhood of* $\psi(Q_\mu)$ *where* Q_μ *is strictly convex. However, there is no such neighbourhood where* Q_μ *is strongly convex. This shows that the strong convexity in Theorem* 3.6 *is indispensable.*

The next result shows that the estimate for the upper second-order directional derivative of φ, which constitutes the final part of the proof of Theorem 3.6, remains valid under more general hypotheses. Second-order tangents sets to C (cf. [1], [6]) turn out to be essential in this respect. Higher-order sets of such type are studied in [9] in the context of (higher-order) necessary optimality conditions for abstract mathematical programs.

Proposition 3.8 *Assume* (A1)-(A3) *and let $\psi(Q_\mu)$ be non-empty, bounded. Let Q_μ be strictly convex on some open convex neighbourhood of $A(\psi(Q_\mu))$ and continuously differentiable at χ_* with $A(\psi(Q_\mu)) = \{\chi_*\}$. Assume that Q_μ has a second-order directional derivative at χ_*, i.e. there exist*

$$Q''_\mu(\chi_*; h) = \lim_{t\to 0+} \frac{1}{t^2}(Q_\mu(\chi_* + th) - Q_\mu(\chi_*) - t\nabla Q_\mu(\chi_*)h)$$

for all $h \in \mathbb{R}^s$. Let g be twice continuously differentiable. Then we have for all convex $v : \mathbb{R}^s \to \mathbb{R}$ and $x \in \psi(Q_\mu)$

$$\limsup_{t\to 0+} \frac{1}{t^2}(\varphi(Q_\mu + tv) - \varphi(Q_\mu) - t\varphi'(Q_\mu; v))$$

$$\leq \inf_{y\in S(x)} \inf_{z\in T^2_C(x,y)} \left\{ \nabla g(x)z + \nabla Q_\mu(\chi_*)Az + \frac{1}{2}\langle\nabla^2 g(x)y, y\rangle + Q''_\mu(\chi_*; Ay) + v'(\chi_*; Ay) \right\}$$

where $S(x)$ is given as in Theorem 3.6 *and $T^2_C(x,y)$ is the second-order tangent set to C at $x \in C$ in direction y, i.e.*

$$T^2_C(x,y) = \liminf_{t\to 0+} \frac{1}{t^2}(C - x - ty).$$

Proof:
Let $y \in S(x)$ be arbitrary. If $T^2_C(x,y) = \emptyset$, then the assertion trivially holds. Hence, let $z \in T^2_C(x,y)$. Then, for arbitrary $t_k \to 0+$ there exists a sequence $\{z_k\}$ such that $z_k \to z$ and $x + t_k y + t_k^2 z_k \in C$ for all $k \in \mathbb{N}$. This allows the following estimate

$$\begin{aligned}
&\varphi(Q_\mu + t_k v) - \varphi(Q_\mu) - t_k\varphi'(Q_\mu; v) \\
&\leq g(x + t_k y + t_k^2 z_k) + Q_\mu(A(x + t_k y + t_k^2 z_k)) + t_k v(A(x + t_k y + t_k^2 z_k)) \\
&\quad -g(x) - Q_\mu(Ax) - t_k v(Ax) \\
&= [g(x + t_k y + t_k^2 z_k) - g(x) - t_k\nabla g(x)y] + \\
&\quad +[Q_\mu(A(x + t_k y + t_k^2 z_k)) - Q_\mu(Ax) - t_k\nabla Q_\mu(Ax)Ay] + \\
&\quad +t_k[v(A(x + t_k y + t_k^2 z_k)) - v(Ax)].
\end{aligned}$$

After dividing by t_k^2 the right-hand side converges to (cf. [3], p. 484)

$$\nabla g(x)z + \frac{1}{2}\langle \nabla^2 g(x)y, y\rangle + \nabla Q_\mu(Ax)Az + Q''_\mu(Ax; Ay) + v'(Ax; Ay).$$

Taking infima on the right-hand side yields the assertion. □

An upper bound similar to the above one is given in [2], Proposition 1.

Remark 3.9 The following condition allows to extend the estimate from Proposition 3.8 to the limes inferior $\liminf\limits_{t\to 0+}$: For each $\varepsilon > 0$ there exist $x \in \psi(Q_\mu), y \in S(x)$ such that for arbitrary $t_k \to 0+$ there exists a sequence $\{z_k\}$ such that $z_k \to z, x+t_k y+t_k^2 z_k \in C$ and $g(x+t_k y+t_k^2 z_k)+(Q_\mu+t_k v)(A(x+t_k y+t_k^2 z_k)) \leq \varphi(Q_\mu + t_k v) + \varepsilon t_k^2$ for $k \in \mathbb{N}$ sufficiently large.

This condition is employed in [18], Theorem 4.1, where it is called *second-order strong stability condition.* Verifying it in the context of two-stage stochastic programs is an open problem.

We finally combine the techniques from Theorem 3.6 and Proposition 3.8. In this way, the additional assumptions on g and C can be dropped. However, more implicit hypotheses on $\psi(Q_\mu)$ have to be verified.

Corollary 3.10 *Assume* (A1)-(A3) *and let $\psi(Q_\mu)$ be non-empty, bounded. Let Q_μ be strongly convex on some open convex neighbourhood of $A(\psi(Q_\mu))$ and twice continuously differentiable at χ_* with $A(\psi(Q_\mu)) = \{\chi_*\}$. Let g be twice continuously differentiable and $v : \mathbb{R}^s \to \mathbb{R}$ be convex. Assume that $x \in \psi(Q_\mu)$ has the following properties:*

(i) $\quad \mathrm{d}(x, \psi(Q_\mu + tv)) = O(t),$

(ii) $\quad 0 \in T_C^2(x, y) \quad$ *for all* $y \in T_C(x)$.

Then $\quad \varphi''(Q_\mu; v) =$

$$= \inf\left\{\frac{1}{2}\langle \nabla^2 g(x)y, y\rangle + \frac{1}{2}\langle \nabla^2 Q_\mu(\chi_*)Ay, Ay\rangle + v'(\chi_*; Ay) : y \in S(x)\right\},$$

where $S(x)$ is given as in Theorem 3.6.

Proof: In view of (i), the same technique as in the proof of Theorem 3.6 applies and one ends up with the right-hand side of the assertion as a lower bound for the $\liminf\limits_{t\to 0+}$. For the $\limsup\limits_{t\to 0+}$, we use Proposition 3.8 and the fact that the necessary optimality condition yields:

$$\nabla g(x)z + \nabla Q_\mu(\chi_*)Az \geq 0 \quad \text{whenever} \quad z \in T_C^2(x, y), y \in T_C(x).$$

(ii) now implies the assertion. □

Acknowledgement:

This paper was partly written during the Workshop " Approximation of Stochastic Optimization Problems" held at the International Institute for Applied Systems Analysis, Laxenburg (Austria) in July 1993. We wish to thank the organizers, Georg Pflug and Andrzej Ruszczyński, for providing excellent working conditions and a stimulating atmosphere.
Further thanks are due to Alexander Shapiro (Georgia Institute of Technology, Atlanta) for valuable comments. Moreover, we are indebted to Alberto Seeger (University of Avignon) for pointing out an incorrectness in an earlier version of this paper.

References

[1] J.-P. Aubin and H. Frankowska, Set-Valued Analysis, Birkhäuser, Boston 1990.

[2] A. Auslender and R. Cominetti, First and second order sensitivity analysis of nonlinear programs under directional constraint qualification conditions, Optimization 21 (1990), 351-363.

[3] A. Ben-Tal and J. Zowe, Directional derivatives in nonsmooth optimization, Journal of Optimization Theory and Applications 47(1985), 483-490.

[4] C. Berge, Espaces Topologiques, Functions Multivoques, Dunod, Paris 1959.

[5] F.H. Clarke, Optimization and Nonsmooth Analysis, Wiley, New York 1983.

[6] R. Cominetti, Metric regularity, tangent sets and second-order optimality conditions, Applied Mathematics and Optimization 21 (1990), 265-287.

[7] J. Dupačová, Stability and sensitivity analysis for stochastic programming, Annals of Operations Research 27 (1990), 115-142.

[8] J.-B. Hiriart-Urruty, Approximate first-order and second-order directional derivatives of a marginal function in convex optimization, Journal of Optimization Theory and Applications 48 (1986), 127-140.

[9] K.-H. Hoffmann and H.J. Kornstaedt, Higher-order necessary conditions in abstract mathematical programming, Journal of Optimization Theory and Applications 26(1978), 533-568.

[10] P. Kall, Stochastic Linear Programming, Springer-Verlag, Berlin 1976.

[11] K. Marti, Approximationen der Entscheidungsprobleme mit linearer Ergebnisfunktion und positiv homogener, subadditiver Verlustfunktion, Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete 31 (1975), 203-233.

[12] B. Poljak, Existence theorems and convergence of minimizing sequences in extremum problems with restrictions, Soviet Math. Dokl. 7 (1966), 72-75 (Dokl. Akad. Nauk SSSR 166 (1966), 287-290).

[13] R. T. Rockafellar, Convex Analysis, Princeton University Press, Princeton 1970.

[14] R.T. Rockafellar, Directional differentiability of the optimal value function in a nonlinear programming problem, Mathematical Programming Study 21 (1984), 213-226.

[15] W. Römisch and R. Schultz, Stability of solutions for stochastic programs with complete recourse, Mathematics of Operations Research 18 (1993), 590-609.

[16] W. Römisch and R. Schultz, Lipschitz stability for stochastic programs with complete recourse, Schwerpunktprogramm "Anwendungsbezogene Optimierung und Steuerung" der DFG, Report No. 408, 1992 and submitted to SIAM Journal Optimization.

[17] R. Schultz, Strong convexity in stochastic programs with complete recourse, Journal of Computational and Applied Mathematics 56 (1994) (to appear).

[18] A. Seeger, Second order directional derivatives in parametric optimization problems, Mathematics of Operations Research 13 (1988), 124-139.

[19] A. Shapiro, Second-order derivatives of extremal-value functions and optimality conditions for semi-infinite programs, Mathematics of Operations Research 10 (1985), 207-219.

[20] A. Shapiro, Sensitivity analysis of nonlinear programs and differentiability properties of metric projections, SIAM Journal Control and Optimization 26 (1988), 628-645.

[21] A. Shapiro, On concepts of directional differentiability, Journal of Optimization Theory and Applications 66 (1990), 477-487.

[22] A. Shapiro Asymptotic analysis of stochastic programs, Annals of Operations Research 30 (1991), 169-186.

[23] J. Wang, Distribution sensitivity analysis for stochastic programs with complete recourse, Mathematical Programming 31 (1985), 286-297.

[24] R.J.-B. Wets, Stochastic programs with fixed recourse: the equivalent deterministic program, SIAM Review 16 (1974), 309-339.