DISTRIBUTION SENSITIVITY ANALYSIS FOR STOCHASTIC PROGRAMS WITH RECOURSE

Jinde Wang

November 1982
WP-82-133

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
A-2361 Laxenburg, Austria

ABSTRACT

In this paper we study the stability of solutions to stochastic programming problems with recourse and show the Lipschitz continuity of optimal solutions as well as the associated Lagrange multipliers with respect to the distribution function.

DISTRIBUTION SENSITIVITY ANALYSIS
FOR STOCHASTIC PROGRAMS WITH RECOURSE

Jinde Wang

University of Nanjing, Nanjing, China
and
University of Kentucky, Lexington, U.S.A.

1. INTRODUCTION

We consider *stochastic programs with recourse* [1] of the type

$\underset{\sim}{P}$ find $x \in R^{n_1}$ such that $g_i(x) \leq 0$, $i = 1,\dots,s$;

$g_i(x) = 0$, $i = s+1,\dots,m$

and $Z(x) = cx + E[Q(x,\omega)]$ is minimized, where

$$(1.1) \qquad Q(x,\omega) = \left[\inf_{y \in R_+^{n_2}} qy \mid Wy = \xi(\omega) - Tx\right] \quad ;$$

here $R_+^{n_2}$ denotes the positive orthant. By $\xi(\omega)$ we denote a sample of the random vector ξ defined on the probability space (Ω, F, P) and with values in R^{m_2}. All other quantities that appear in the formulation of P are fixed (nonstochastic). We assume that P is a stochastic program with *complete recourse* [1, Section 6], i.e. that the linear system

$$(1.2) \qquad Wy = t \quad , \quad y \in R_+^{n_2}$$

is solvable for any $t \in R^{m_2}$.

The purpose of this paper is to study the sensitivity of the optimal value, the optimal solution and associated Kuhn-Tucker multipliers, to changes in the distribution of random variables of the problem. We shall consider perturbations in the space of random vectors ξ defined on (Ω, F, P); to indicate this dependence of Z on ξ we write $Z(x,\xi)$ for the objective function of P_ξ. We begin by reviewing and extending some of the properties of Z (Section 2) that lead us to certain stability results for the set of optimal solutions (Section 3) and the Lipschitz continuity of the infima (Section 4). Sensitivity of the Kuhn-Tucker points is further analyzed in Section 5. An example is given in Section 6 to illustrate the results.

Dupačova [2] obtains distribution sensitivity results for stochastic programs, assuming that Z is twice differentiable. Here we identify the class of problems for which Z inherits second order differentiability properties. We also obtain distribution sensitivity results when Z satisfies much weaker differentiability conditions, a case that covers all continuous distribution functions used in practice.

2. ANALYTICAL PROPERTIES OF THE OBJECTIVE

To set the stage we start with some well known facts about the function Z [1, Section 7], [3], we then refine Lipschitz continuity and first order differentiability results, and finally derive second order differentiability properties. Eventually, this will allow us to apply the tools of Nonsmooth Analysis to analyze the sensitivity of stochastic programs. A key role is played by the lemma below. A *finite closed polyhedral complex* is any finite collection H of closed convex polyhedra, called cells of H, such that

(i) if C is a cell of H, then every closed face of C is a member of H;

(ii) if C_1 and C_2 are distinct cells of H, then either they are disjoint, or one is a face of the other, or their intersection is a common face.

We are interested in a complicial decomposition of the parameter-space of a linear program with varying right-hand sides. This turns out to be a decomposition in simplicial convex cones of the positive hull of the points generated by the columns $A^1,\dots,A^k$ of the technology matrix A, i.e.

$$\text{pos } A = \text{pos } (A^1,\dots,A^k) = \left\{t \mid t = \sum_{\ell=1}^{k} A^{\ell} x_{\ell} , x_{\ell} \geq 0\right\} .$$

2.1 Basis Decomposition Theorem [4]. *Let*

$$P(t) = \inf [cx \mid Ax = t , x \geq 0] ,$$

where the matrix A *and the vector* d *are fixed. Then*

(i) $P(t) < +\infty$ *if and only if* t *lies in* pos A;

(ii) *either* P *is bounded on* pos A *or* $P \equiv -\infty$ *on* pos A;

(iii) *when* P *is bounded, there exists a decomposition of* pos A *into a finite closed polyhedral complex* H *whose cells are simplicial cones (with vertex at the origin) and a one-to-one correspondence between the one dimensional cells of* H *and selected columns of* A *which generate them such that*

(a) *the closed* m*-dimensional cells of* H *cover* pos A, *and*

(b) *the* m *columns of* A *associated with the edges of a* m*-dimensional cell* C *of* H *constitute an optimal basis for all* t *in* C.

Applying this result to Q, and recalling that pos W = R^{m_2} since we have complete recourse (1.2), we see that Q is either identically $-\infty$ or finite on $R^{n_1} \times \Omega$. We assume throughout that Q is finite. Finiteness of Z follows from an integrability condition as we show next.

2.2 Proposition [1, Section 7]. If ξ *is summable then* Z *is finite on* R^{n_1}.

Proof. Let $\{W_{(j)} , j = 1,\dots,J\}$ be the finite collection of sub-matrices of W prescribed by the Basis Decomposition Theorem, each one determining the optimal basis when

$$\xi(\omega) - Tx \in \text{pos } W_{(j)} \quad ,$$

or equivalently when $\omega \in \Xi_{(j)}(x)$ where

$$(2.3) \qquad \Xi'_{(j)}(x) = \{\omega \mid W_{(j)}^{-1}[\xi(\omega) - Tx] \geq 0\} \quad .$$

From this dissection $\{\Xi'_{(j)}(x), j = 1,\ldots,J\}$ we obtain a partition by setting

$$\Xi_{(j)}(x) = \Xi'_{(j)}(x) \setminus \bigcup_{k=1}^{j-1} \Xi'_{(k)}(x) \quad .$$

We have that

$$(2.4) \qquad Z(x) = cx + \sum_{j=1}^{J} \int_{\Xi_{(j)}(x)} q_{(j)} W_{(j)}^{-1} [\xi(\omega) - Tx] P(d\omega)$$

where $q_{(j)}$ is the subvector of a q corresponding to the submatrix $W_{(j)}$ of W. The integrands are linear in ξ and thus finite, from which follows the finiteness of Z since there are only a finite number of terms. □

Since Z is finite and convex [1, Theorem 7.6] it follows that it is locally Lipschitz. In fact Z is Lipschitz [1, Theorem 7.7]. We want to go one step further and show that the Lipschitz constant is independent of ξ. We also establish that $\xi \mapsto Z(x,\xi) : L^1_{m_2} = L^1(\Omega,F,P;R^{m_2}) \to R$ is also Lipschitz continuous. In view of the above this will allow to conclude that $(x,\xi) \to Z(x,\xi)$ is Lipschitz, i.e. jointly in x and ξ.

2.5 Proposition. *The function* $(x,\xi) \mapsto Z(x,\xi)$ *is Lipschitz continuous on* $R^n \times L^1_{m_2}$.

Proof. We first observe that for all $\xi \in L^1_{m_2}$

$$(2.6) \qquad |Z(x,\xi) - Z(x^0,\xi)| \leq L|x - x^0|$$

with L independent of ξ. A detailed argument would go as the proof of Theorem 7.7 in [1]; recall naturally that here q and T are fixed.

Now for fixed x, the function

$$t \mapsto Q'(x,t) = \inf \left[qy \mid Wy = t - Tx \, , \, y \in R_+^{n_2} \right]$$

is linear on a finite number of polyhedra:

$$S_{(j)}(x) = \left\{ t \mid W_{(j)}^{-1}(t - Tx) \geq 0 \right\} \subset R^{m_2} \quad ,$$

and is given by

$$Q'(x,t) = q_{(j)} W_{(j)}^{-1}(t - Tx)$$

for $W_{(j)}$ and $q_{(j)}$ as defined in the proof of Proposition 2.2, the assertions following again from the Basis Decomposition Theorem 2.1. Hence with M independent of t, we have

$$|Q'(x,t) - Q'(x,t_0)| \leq M|t - t_0| \quad .$$

Substituting $\xi(\cdot)$ for t and $\xi_0(\cdot)$ for t_0 we get

$$\text{(2.7)} \qquad |Z(x,\xi) - Z(x,\xi_0)| \leq \int_\Omega |Q'(x,\xi(\omega)) - Q'(x,\xi_0(\omega)| P(d\omega)$$

$$\leq M \int_\Omega |\xi(\omega) - \xi_0(\omega)| P(d\omega) = M\| \xi - \xi_0 \|_{L^1} \quad .$$

The joint Lipschitz continuity of Z is an immediate consequence of (2.6) and (2.7). □

Next we turn to differentiability properties of $Z(\cdot,\xi)$. If the distribution of $\tilde{\xi}$ is absolutely continuous, i.e. is given by a density function f, then Z is continuously differentiable [4, Chapter III, Theorem 12] and its gradient is given by

$$\nabla Z(x) = c - \sum_{j=1}^{J} q_{(j)} W_{(j)}^{-1} P(\Xi_{(j)}(x)) \quad ,$$

as follows from (2.4), or equivalently

(2.8) $$\nabla Z(x) = c - \sum_{j=1}^{J} q_{(j)} W_{(j)}^{-1} \int_{S_{(j)}(x)} f(t)dt$$

with $S_{(j)}(x) = \{t | W_{(j)}^{-1}(t - Tx) \geq 0\}$ as before. Note that for $j = i,\ldots,J$, the $S_{(j)}(x)$ are translates of convex polyhedral cones, viz.

$$S_{(j)}(x) = W_{(j)}^{-1} Tx + \{t | W_{(j)}^{-1} t \geq 0\} \quad .$$

We attain higher order differentiability properties for Z through the study of the analytical properties of the integrals $\int_{S_{(j)}(x)} f(t)dt$ that define ∇Z.

2.9 Lemma. *Let* $\chi = (\chi_1,\ldots,\chi_n) \in R^n$ *and*

$$I(\chi) = \int_{-\infty}^{\chi_1} dt_1 \ldots \int_{-\infty}^{\chi_n} dt_n h(t_1,\ldots,t_n) \quad ,$$

where h *is continuous, nonnegative and such that* $I: R^n \to R_+$ *is bounded above. Moreover suppose that all* (n-1) *iterated integrals converge uniformly with respect to the remaining variable, e.g. with* i = 1 *and*

$$i(t_1,M_2,\ldots,M_n) = \int_{M_2}^{\chi_2} dt_2 \ldots \int_{M_n}^{\chi_n} dt_n h(t_1,t_2,\ldots,t_n) \quad ,$$

$i(t_1,M_2,\ldots,M_n)$ *converges uniformly (with respect to* t_1*) to* $i(t_1,-\infty,\ldots,-\infty)$. *Then* I *is continuously differentiable.*

Proof. We prove the lemma for n = 2, the general case is obtained by induction. We have

$$I(\chi_1,\chi_2) = \int_{-\infty}^{\chi_1} dt_1 \int_{-\infty}^{\chi_2} dt_2 h(t_1,t_2) \quad .$$

From the definition of I we have that

$$\frac{\partial}{\partial \chi_1} I(\chi_1,\chi_2) = \int_{-\infty}^{\chi_2} h(\chi_1,t_2)dt_2$$

and

$$\frac{\partial}{\partial \chi_2} I(\chi_1,\chi_2) = \int_{-\infty}^{\chi_1} h(t_1,\chi_2)dt_1 \quad .$$

The uniform convergence of the marginals implies that $\frac{\partial}{\partial \chi_i} I$ is continuous for $i = 1,2$. Also the mixed second order partials are continuous, since

$$\frac{\partial}{\partial \chi_1} \frac{\partial}{\partial \chi_2} I(\chi_1,\chi_2) = h(\chi_1,\chi_2) \quad .$$

With this we obtain the uniform continuity of $\frac{\partial I}{\partial \chi_1}(\chi_1,\cdot)$ with respect to χ_1, and similarly for $\frac{\partial I}{\partial \chi_2}(\cdot,\chi_2)$. From this it follows that both $\frac{\partial I}{\partial \chi_1}$ and $\frac{\partial I}{\partial \chi_2}$ are continuous in χ_1, χ_2 jointly. □

2.10 Theorem. *Suppose the random variable* ξ *has a density function* f *such that* $t \mapsto f(Bt)$ *satisfies the assumptions of* Lemma 2.9 *for any invertible matrix* B. *Then* $Z(\cdot,\xi)$ *is twice continuously differentiable.*

Proof. By (2.8) we know that it is sufficient to prove that all the integrals

$$I_{(j)}(x) = \int_{S_{(j)}(x)} f(t)dt = \int_{W_{(j)}^{-1}t \geq W_{(j)}^{-1}Tx} f(t)dt$$

are continuously differentiable.

If for some j, $W_{(j)}$ is the unit matrix, then by lemma (2.9), integral $I_{(j)}(x)$ is twice continuously differentiable.

If $W_{(j)}$ is not a unit matrix, then by coordinate transformation $\tau = W_{(j)}^{-1}t$ the integral $I_{(j)}(x)$ becomes

$$I_j(x) = \int_{\tau \geq Tx} f(W_{(j)}\tau)W_{(j)}d\tau$$

$$= \int_{(Tx)_1}^{\infty} d\tau_1 \ldots \int_{(Tx)_{m_2}}^{\infty} f(W_{(j)}\tau)W_{(j)}d\tau_{m_2}$$

which allows us to apply Lemma 2.9.

Thus $Z(\cdot,\xi)$ is twice continuously differentiable. □

Sometimes it is not so easy to verify the uniform convergence of the iterated integrals. The following theorem which shows the Lipschitz continuity of $\nabla Z(x)$ without demanding uniform convergence for the integrals, would be very useful in practice.

2.11 Theorem. *Suppose the density function* $f(t)$ *is such that every one-dimensional marginal density function is bounded on any finite interval and this property is preserved under invertible linear transformation* $t \to Bt$. *Then* $\nabla Z(x)$ *is Lipschitz continuous everywhere.*

Proof. First we assume $I_{(j)}(x)$ is in standard form, then

$$I_{(j)}(x) = \int_{(Tx)_1}^{\infty} dt_1 \int_{(Tx)_2}^{\infty} dt_2 \cdots \int_{(Tx)_{m_2}}^{\infty} f(t_1,\ldots,t_{m_2})dt_{m_2}$$

Denoting Tx by $\chi = (\chi_1,\chi_2,\ldots,\chi_{m_2})'$, we have

$$I_{(j)}(\chi) = \int_{\chi_1}^{\infty} dt_1 \int_{\chi_2}^{\infty} dt_2 \cdots \int_{\chi_{m_2}}^{\infty} f(t_1,t_2,\ldots,t_{m_2})dt_{m_2} \quad .$$

Let $\Delta I_{(j)}^{(1)}$ be the variation associated to $\Delta\chi_1$. Then

$$(2.12) \quad |\Delta I_{(j)}^{(1)}| = \Big|\int_{\chi_1}^{\infty} dt_1 \int_{\chi_2^0}^{\infty} dt_2 \cdots \int_{\chi_{m_2}^0}^{\infty} f(t_1,\ldots,t_{m_2})dt_{m_2}$$

$$- \int_{\chi_1^0}^{\infty} dt_1 \int_{\chi_2^0}^{\infty} dt_2 \cdots \int_{\chi_{m_2}^0}^{\infty} f(t_1,\ldots,t_{m_2})dt_{m_2}\Big|$$

$$= \Big|\int_{\chi_1}^{\chi_1^0}\Big[\int_{\chi_2^0}^{\infty} dt_2 \cdots \int_{\chi_{m_2}^0}^{\infty} f(t_1,\ldots,t_{m_2})dt_{m_2}\Big]dt_1\Big|$$

$$\leq \Big|\int_{\chi_1}^{\chi_1^0}\Big[\int_{-\infty}^{\infty} dt_2 \cdots \int_{-\infty}^{\infty} f(t_1,\ldots,t_{m_2})dt_{m_2}\Big]dt_1\Big| = \Big|\int_{\chi_1}^{\chi_1^0} \tilde{f}_1(t_1)dt_1\Big|$$

where $\tilde{f}_1(t_1)$ is the one-dimensional marginal density function on axis t_1, and for $i = 2,\ldots,m_2$, $\tilde{f}_i(t_i)$ are defined similarly.

Since on any finite interval $[\chi_1^{(1)}, \chi_1^{(2)}]$, of which χ_1^0 is an interior point, $\tilde{f}_1(t_1)$ is bounded and inequality (2.12) holds for any $(\chi_2^0,\ldots,\chi_{m_2}^0)$, we have

$$|\Delta I_{(j)}^{(1)}| \leq K_{(j)}^{(1)}(\chi_1^0)|\Delta\chi_1|$$

where $K_{(j)}^{(1)}(\chi_i^0) = \max\{\tilde{f}_1(t_1), t_1 \in [\chi_1^{(1)}, \chi_1^{(2)}]\}$.

In a similar way we can get

$$|\Delta I_{(j)}^{(i)}| \leq K_{(j)}^{(i)}(\chi_i^{(0)}) \cdot |\Delta\chi_i|$$

for any $(\chi_1^{(0)},\chi_2^{(0)},\ldots,\chi_{i-1}^{(0)},\chi_{i+1}^{(0)},\ldots,\chi_{m_2}^{(0)})$, where $K_{(j)}^{(i)}(\chi_i^{(0)}) = \max\{\tilde{f}_i(t_i), t_i \in [\chi_i^{(1)}, \chi_i^{(2)}]\}$.

Therefore for the variation $\Delta I_{(j)}$ associated to any $\Delta\chi$ we have

$$\begin{aligned}\Delta I_{(j)} &= I_{(j)}(\chi_1,\ldots,\chi_{m_2}) - I_{(j)}(\chi_1^{(0)},\ldots,\chi_{m_2}^{(0)}) \\ &= \sum_{i=1}^{m_2}\Big[I_{(j)}(\chi_1,\ldots,\chi_{i-1},\chi_i,\chi_{i+1}^{(0)},\ldots,\chi_{m_2}^{(0)}) \\ &\quad - I_{(j)}(\chi_1,\ldots,\chi_{i-1},\chi_i^{(0)},\chi_{i+1}^{(0)},\ldots,\chi_{m_2}^{(0)})\Big] \\ &\leq \sum_{i=1}^{m_2} K_{(j)}^{(i)}(\chi_i^{(0)}) \cdot |\Delta\chi_i| \leq K_j(\chi^0) \cdot |\Delta\chi| ,\end{aligned}$$

where

$$K_j(\chi^0) = \max\{K_{(j)}^{(i)}(\chi_i^{(0)}), i=1,2,\ldots,m_2\} \text{ and } |\Delta\chi| = \left[\sum_{i=1}^{m_2}|\Delta\chi_i|^2\right]^{1/2}.$$

This inequality implies $I_{(j)}(\chi)$ is Lipschitz continuous in χ, thus also Lipschitz continuous in x.

By using transformation $t = W_{(j)}\tau$ and paying attention to the second part of the assumption one can get Lipschitz continuity of $I_{(j)}(x)$ for all $j = 1,2,\ldots,J$. This yields the Lipschitz continuity of $\nabla Z(x)$. ■

3. STABILITY OF THE SOLUTION SET

We study the multifunction $A: L^1_{m_2} \rightrightarrows R^{n_1}$ defined by

$$A(\xi) = \arg\min_{x \in S} Z(x,\xi) = \arg\min F(x,\xi)$$

where

$$S = \{x \mid g_i(x) \leq 0 \ , \ i=1,\ldots,s \ ; \ g_i(x) = 0 \ , \ i=s+1,\ldots,m_1\}$$

is the feasibility region and

$$F(x,\xi) = \begin{cases} Z(x,\xi) & \text{if } x \in S, \\ +\infty & \text{otherwise.} \end{cases}$$

The functions g_i are taken to be continuous and thus S is closed and $F(\cdot,\xi)$ is lower semicontinuous. We rely on the theory of epi-convergence to derive stability results for the optimality set A as a function of ξ.

For convenience of the reader we review the definition and the main implications of epi-convergence here, for further details the reader is referred to [6], [7].

3.1 **Definition**. *The sequence* $\{f^\nu : R^n \to \bar{R}, \nu=1,2,\ldots\}$ *is said to epiconverge to* $f : R^n \to \bar{R}$ *at* x, *if*

(a) *for all subsequences* $\{\nu_k, k=1,2,\ldots\}$ *and* $x^k \to x$, *we have*

$$\liminf f^{\nu_k}(x^k) \geq f(x) \quad ,$$

(b) $\limsup f^\nu(x^\nu) \leq f(x)$ *for some sequence* $x^\nu \to x$.

If this holds for every $x \in R^n$, *the* f^ν *are said to epi-converge to* f.

3.2 **Theorem**. [6] *Suppose* $\{f^\nu\}$ *is a sequence, epi-converges to* f. *Then for every subsequence* $\{f^{\nu_k}, k=1,2,\ldots\}$ *we have*

$$\limsup (\arg\min f^{\nu_k}) \subset \arg\min f$$

i.e. *if for* $k = 1,2,\ldots, x^k \in \arg\min f^{\nu_k}$ *and* $x^k \to x$, *then* $x \in \arg\min f$.

Moreover in this case we also have

$$\lim (\inf f^{\nu_k}) = \inf f \quad .$$

With this result we can obtain the next theorem.

3.3 <u>Theorem</u>. *If* ξ *converges to* ξ_0 *in the* L^1*-sense, then* $F(x,\xi)$ *epi-converges to* $F(x,\xi_0)$ *and hence*

$$\limsup_{\xi\to\xi_0} A(\xi) \subset A(\xi_0) \quad .$$

Proof. By Theorem 3.2 it suffices to verify conditions (a) and (b) of Definition 3.1. Because $L^1_{m_2}$ is separable, we may restrict ourselves to convergent sequences ξ^ν. Let x_0 be a point of S. For all subsequences $\{\nu_k, k=1,2,\ldots\}$ and $x^k \to x_0$, by Proposition 2.5, we have

$$\liminf_{k\to\infty} F(x^k,\xi^{\nu_k}) = \lim_{k\to\infty} F(x^k,\xi^{\nu_k}) = F(x_0,\xi_0)$$

if $x^k \in S$ for all $k \geq K$ where K is a certain integer, and

$$\liminf F(x^k,\xi^{\nu_k}) = +\infty > F(x_0,\xi_0)$$

if $x^k \notin S$ for all $k \geq K$. This implies (a) holds at $x_0 \in S$. Again by Proposition 2.5, we have a sequence $x^\nu \equiv x_0$ for all such that

$$\limsup_{\nu\to\infty} F(x^\nu,\xi^\nu) = \lim_{\nu\to\infty} F(x^\nu,\xi^\nu) = F(x_0,\xi_0)$$

i.e. (b) is satisfied at $x_0 \in S$.

If $x_0 \notin S$, then $F(x_0,\xi_0) = F(x_0,\xi^\nu) = +\infty$ for all $\nu = 1,2,\ldots$ Since S is a closed set, x_0 must be an interior point of S^c, the complement of S. Then there is a neighborhood of x_0 such that at all points x in this neighborhood $F(x,\xi_0) = F(x,\xi^\nu) = +\infty$ for all $\nu = 1,2,\ldots$ Thus (a) and (b) hold trivially. □

4. LIPSCHITZ CONTINUITY OF THE INFIMA FUNCTION

From epi-convergence, Theorem 3.2 and Theorem 3.3, follows that if for $\nu = 1,2,\ldots$

$$\xi^\nu \overset{L^1}{\to} \xi_0 \quad , \qquad x^\nu \in A(\xi^\nu) \quad \text{and} \quad x^\nu \to x_0$$

then $x_0 \in A(\xi_0)$ and

$$\inf F(x,\xi_0) =: \varphi(\xi_0) = \lim_{\nu\to\infty} [\varphi(\xi^\nu) := \inf F(x,\xi^\nu)] \quad .$$

This means that the infima function φ is continuous at ξ_0. In the present setting we have already shown that z is jointly Lipschitzian in (x,ξ), under some additional mild assumptions we can get Lipschitz continuity of $\varphi(\xi)$.

4.1. Theorem. *Suppose ξ converges to ξ_0 in the L^1-sense and the solution set $A(\xi_0)$ is compact. Then the infima function $\varphi(\xi)$ is Lipschitz continuous in ξ in the L^1-sense, i.e.*

$$|\varphi(\xi) - \varphi(\xi_0)| \le L \cdot \|\xi - \xi_0\|_{L^1}$$

for some constant L.

Proof. The proof involves three major steps.

i) We prove that $\varphi(\xi) > -\infty$ for all ξ such that $\|\xi - \xi_0\| < \delta$ for some $\delta > 0$.

First, note that since $A(\xi_0)$ is compact, $Z(x,\xi_0)$ is a convex function, hence $\varphi(\xi_0)$ is finite and $Z(x,\xi_0) \to +\infty$ as $|x| \to +\infty$.

Suppose the assertion i) is not true, then there is a sequence $\{\xi^\nu\}$ converging to ξ_0 (in L^1-sense) and having $\varphi(\xi^\nu) = -\infty$. Hence for each ν there exists a sequence $\{x_k^\nu, k=1,2,\ldots\}$ such that $Z(x_k^\nu,\xi^\nu)$ tends to $-\infty$ as $k\to\infty$. Obviously $|x_k^\nu| \to \infty$ as $k\to\infty$. Or equivalently for each ν there is an integer $K_1(\nu)$ such that for all $k \ge K_1(\nu)$ it follows that

$$(4.2) \qquad Z(x_k^\nu,\xi^\nu) < -\Delta$$

where Δ is any given positive large number.

On the other hand, as pointed out above, for each ν there exists an integer $K_2(\nu)$ such that

$$(4.3) \qquad Z(x_k^\nu, \xi_0) > \Delta \quad \text{for all } k \geq K_2(\nu) \quad .$$

Combining (4.2) and (4.3) we obtain

$$Z(x_k^\nu, \xi_0) - Z(x_k^\nu, \xi^\nu) > 2\Delta$$

for all $k \geq K(\nu) = \max[K_1(\nu), K_2(\nu)]$ for each ν.

This conflicts with the Lipschitz continuity of Z in (x,ξ), of Proposition 2.5. Thus assertion i) must be true.

ii) We prove that $Z(x,\xi)$ attains its minimum when $\|\xi - \xi_0\|_{L^1} < \delta$.

Suppose for some sequence ξ^ν, converging to ξ_0 in the L^1-sense, the assertion fails. Then for each ν there is a sequence $\{x_i^\nu, i=1,2,...\}$ such that $|x_i^\nu| \to \infty$ as $i \to \infty$ and

$$\lim_{i\to\infty} Z(x_i^\nu, \xi^\nu) = \varphi(\xi^\nu) \quad .$$

By step i) $\varphi(\xi^\nu)$ is finite when $\nu \geq N_1$ for some number N_1. This implies that for a given $\varepsilon_1 > 0$, there is a large number $I(\nu)$ such that

$$(4.4) \qquad \varphi(\xi^\nu) < Z(x_i^\nu, \xi^\nu) \leq \varphi(\xi^\nu) + \varepsilon_1 \quad .$$

By Lipschitz continuity of Z, for some $\varepsilon_2 > 0$ there exists N_2 such that for all $\nu \geq N_2$

$$(4.5) \qquad Z(x_i^\nu, \xi_0) \leq Z(x_i^\nu, \xi^\nu) + \varepsilon_2$$

and for some $\varepsilon_3 > 0$ and $\chi_0 \in A(\xi_0)$

$$(4.6) \qquad Z(x_0, \xi_0) \geq Z(x_0, \xi^\nu) + \varepsilon_3 \quad .$$

Let $N = \max(N_1, N_2)$. The last three inequalities lead to

$$\begin{aligned} Z(x_i^\nu, \xi_0) &\le Z(x_i^\nu, \xi^\nu) + \varepsilon_2 \\ &\le \varphi(\xi^\nu) + \varepsilon_1 + \varepsilon_2 \\ &\le Z(x_0, \xi^\nu) + \varepsilon_1 + \varepsilon_2 \le Z(x_0, \xi_0) + \varepsilon_3 + \varepsilon_1 + \varepsilon_2 \end{aligned}$$

for all $\nu \ge N$ and $i \ge I(\nu)$.

It means that if i and ν are large enough $\{Z(x_i^\nu, \xi_0)\}$ is dominated by some constant and again is in contradiction with the assumptions. Thus assertion ii) holds.

iii) We now show Lipschitz continuity of $\varphi(\xi)$.

From steps i) and ii) we have

$$\varphi(\xi) = Z(x(\xi), \xi) \quad \text{for all } \xi \text{ such that } \|\xi - \xi_0\| < \delta.$$

Then it follows that

$$|\varphi(\xi) - \varphi(\xi_0)| = |Z(x(\xi), \xi) - Z(x_0, \xi_0)|$$

$$\le \begin{cases} |Z(x_0, \xi) - Z(x_0, \xi_0)| & \text{if } Z(x(\xi), \xi) \ge Z(x_0, \xi_0) \\ |Z(x(\xi), \xi) - Z(x(\xi), \xi_0)| & \text{if } Z(x(\xi), \xi) < Z(x_0, \xi_0) \end{cases} .$$

In both cases, by the Lipschitz property of Z, we have

$$|\varphi(\xi) - \varphi(\xi_0)| \le L \cdot \|\xi - \xi_0\|_{L1} \quad . \quad \square$$

5. DIFFERENTIABILITY AND LIPSCHITZ CONTINUITY OF OPTIMAL SOLUTIONS AND LAGRANGIAN MULTIPLIERS

Let us turn now to distribution sensitivity analysis. Suppose we have a collection of random variables $\xi(\omega, \lambda)$, $\lambda \in \Lambda \subset R^P$. Corresponding to each random variable $\xi(\omega, \lambda)$ there is an optimization problem $P(\lambda)$ of the following type:

$$P(\lambda): \begin{cases} \text{Min } Z(x,\lambda) = cx + E\{\min\limits_{y\geq 0} qy \mid Wy = \xi(\omega,\lambda) - Tx\} \\ \text{s.t.} \quad g_i(x) \leq 0 \quad i = 1,2,\ldots,s, \\ \qquad g_i(x) = 0 \quad i = s+1,\ldots,m_1. \end{cases}$$

To those problems, in which $Z(x,\lambda)$ and for $i = 1,2,\ldots,m_1$, $g_i(x)$ are twice continuously differentiable, we can apply the classical result of sensitivity analysis, see [8], and get the following proposition.

5.1. <u>Proposition</u>. *Assume*

(a) the random vector $\xi(\omega,\lambda)$ *has a continuous density function* $D(t,\lambda)$ *such that* $Z(x,\lambda)$ *is twice continuously differentiable around the point* (x_0,λ_0), *here* λ^0 *is a point in* Λ *and* x^0 *is an optimal solution of problem* $P(\lambda^0)$;

(b) the constraint functions $g_i(x)$, $i = 1,2,\ldots,m_1$, *are in* C^2 *and the gradients* $\{\nabla g_i(x^0), i\in I_1; \nabla_i g(x^0), i = s+1,\ldots,m_1\}$ *are linearly independent, where* $I_1 = \{i \mid 1 \leq i \leq s, g_i(x^0) = 0\}$;

(c) (x^0,u^0,v^0) *is a Kuhn-Tucker point, satisfying the following optimality condition:*

$$\nabla_x L(x^0,u^0,v^0,\lambda^0) = 0 \text{ where } L(x,u,v,\lambda) = Z(x,\lambda) + \sum_{i=1}^{s} u_i g_i(x) + \sum_{i=s+1}^{m_1} v_i g_i(x)$$

$$u = (u_1,\ldots,u_s)', \quad v = (v_{s+1},\ldots,v_{m_1})'$$

$$u_i^0 g_i(x^0) = 0 \qquad i = 1,\ldots,s$$

$$u_i^0 \geq 0 \qquad i = 1,\ldots,s$$

$$g_i(x^0) \leq 0 \qquad i = 1,\ldots,s, \quad g_i(x^0) = 0 \quad i = s+1,\ldots,m_1 ,$$

$y^T\nabla^2_{xx} L(x^0,u^0,v^0,\lambda^0)Y > 0$ for all $y \in R^{n_1}$ such that

$$y^T \nabla g_i(x^0) \leq 0 \qquad i \in I_1$$

$$y^T \nabla g_i(x^0) = 0 \qquad 1 \leq i \leq s, \quad u_i^0 > 0$$

$$y^T \nabla g_i(x^0) = 0 \qquad i = s+1, \ldots, m_1 \quad ;$$

(d) strict complementary slackness holds, i.e. $u_i^0 > 0$, $i \in I_1$.
Then

i) the solution $K(\lambda) = (x(\lambda), u(\lambda), v(\lambda))$ *of the equations*

$$\nabla_x L(x,u,v,\lambda) = 0$$

$$u_i g_i(x) = 0 \qquad 1 \leq i \leq s$$

$$g_i(x) = 0 \qquad s+1 \leq i \leq m_1$$

is differentiable at $\lambda = \lambda^0$ *with* $K(\lambda^0) = (x^0, u^0, v^0)$ *and* $x(\lambda)$ *is the optimal point of problem* $P(\lambda)$*, while* $u(\lambda)$, $v(\lambda)$ *are the associated Lagrange multipliers.*

ii) the infima function $\Phi(\lambda)$ *is differentiable at* λ^0.

Remark: The assumption (a) that $z(x,\lambda)$ is twice continuously differentiable is not so stringent; it will be satisfied by most continuous distribution functions (i.e. the ones with continuous density function). In Section 6, an example is given for illustration.

When the density function is not continuous, the differentiability can not be asserted. But for most useful distribution functions $\nabla_x L$ is Lipschitzian. Thus we can apply some results of Nonsmooth Analysis to our problem.

5.2. Definition. [9] *The generalized derivative* (Jacobian) *of a Lipschitz continuous function* $f(x) : R^n \to R^n$ *at point* x^0*, denoted by* $\partial f(x^0)$*, is defined as the convex hull of all matrices* M *of the form*

$$M = \lim_{x^i \to x^0} Jf(x^i)$$

where x^i *converges to* x^0 *and* f *is differentiable at* x^i, $Jf(x^i)$ *is the Jacobian.*

Similarly one can define generalized partial derivatives of a Lipschitz continuous function.

5.3. Definition. [9] *The generalized partial derivative of a Lipschitz continuous function* $f(x_1,x_2)$ *from* $R^n \times R^m$ *to* R^n *with respect to* x_1 *at point* (x_1^0,x_2^0) *is defined by*

$$\partial x_1 f(x_1^0,x_2^0) = \overline{co}\left\{\lim_{\substack{x_1^i \to x_1^0 \\ x_2^i \to x_2^0}} \frac{\partial f}{x_1}(x_1^i,x_2^i)\right\} .$$

5.4. Definition. [8],[9]. *If every element of* $\partial f(x^0)$ *is of maximal rank,* $\partial f(x^0)$ *is said to be surjective. If every element of* $\partial_{x_1} f(x_1^0,x_2^0)$ *is of maximal rank,* $\partial_{x_1} f(x_1^0,x_2^0)$ *is said to be surjective.*

5.5. Lipschitz Implicit Function Theorem. [10] *Suppose that* U *is an open subset of* $R^n \times R^m$, (x_1^0,x_2^0) *lies in* U, f *is a Lipschitz continuous mapping on* U *with values in* R^n *and*

$$f(x_1^0,x_2^0) = 0 \quad .$$

If $\partial_{x_1} f(x_1^0,x_2^0)$ *is surjective, then for some open set* $V \subset R^m$ *with* $x_2^0 \in V$ *and some mapping* $g : V \to R^n$,

(1) g *is Lipschitz continuous,*

(2) $g(x_2^0) = x_1^0$

(3) $(g(x_2),x_2) \in U$ *and* $f(g(x_2),x_2) = 0$ *for every* x_2 *in* V.

Now we apply these results to our problem and get the following proposition.

5.6. Theorem. *Assume*

(a) $Z(x,\lambda)$ *and* $\nabla_x Z(x,\lambda)$ *are Lipschitz continuous around* (x^0,λ^0), *where* λ_0 *is a point in* Λ, x^0 *is the optimal point of* $P(\lambda_0)$,

(b) the constraint functions $g_i(x)$, $i = 1,\dots,m_1$, *are in* C^2 *and the gradients* $\{\nabla g_i(x^0)\ i \in I_1 ;\ \nabla_i g(x^0), i = s+1,\dots,m_1\}$ *are linearly independent, where* $I_1 = \{i : 1 \le i \le s,\ g_i(x^0) = 0\}$; $g_i(x)$, $1 \le i \le s$ *are convex functions, and for* $i = s+1,\dots,m_1$, $g_i(x)$ *are affine functions,*

(c) (x^0,u^0,v^0) *is a Kuhn-Tucker point, satisfying the following optimality condition:*

$$(5.7) \qquad \nabla_x L(x^0,u^0,v^0,\lambda^0) = 0 \quad ,$$

where

$$L(x,u,v,\lambda) = Z(x,\lambda) + \sum_{i=1}^{s} u_i g_i(x) + \sum_{i=s+1}^{m_1} v_i g_i(x)$$

$$u = (u_1,\dots,u_s)', \quad v = (v_{s+1},\dots,v_{m_1})'$$

$$u_i^0 g_i(x^0) = 0 \qquad 1 \le i \le s$$

$$u_i^0 \ge 0 \qquad 1 \le i \le s$$

$$g_i(x^0) \le 0 \qquad 1 \le i \le s \ , \quad g_i(x^0) = 0 \qquad s+1 \le i \le m_1$$

for every element M of $\partial_x(\nabla L)\Big|_{(x^0,u^0,u^0,\lambda^0)}$ it holds that $y^T M y > 0$ for all $y \in R^{n_1}$ such that

$$y^T \nabla g_i(x^0) \le 0 \qquad \text{if} \quad i \in I_1$$

$$y^T \nabla g_i(x^0) = 0 \qquad \text{if} \quad 1 \le i \le s \ , \quad u_i^0 > 0$$

$$y^T \nabla g_i(x^0) = 0 \qquad \text{if} \quad i = s+1,\dots,m_1 \quad ;$$

(d) strict complementary slackness holds.

Then

(i) there exists a solution function $K(\lambda) = (x(\lambda),u(\lambda),v(\lambda))$ *to the equations*

$$\nabla_x L(x,u,v,\lambda) = 0$$

$$u_i g_i(x) = 0 \qquad i = 1,\dots,s$$

$$g_i(x) = 0 \qquad i = s+1,\dots,m_1$$

and $K(\lambda^0) = (x^0,u^0,v^0)$. $K(\lambda)$ *is Lipschitz continuous at* λ^0, $x(\lambda)$ *is the optimal solution of* $P(\lambda)$ *and* $u(\lambda)$, $v(\lambda)$ *are the associated Lagrange multipliers.*

(ii) the infima function $\varphi(\lambda)$ *is Lipschitz continuous.*

Proof. (i) Denote the Kuhn-Tucker equation system (5.7) by $H(K,\lambda) = 0$, where K is the vector (x,u,v).

First we show one can apply the Lipschitz implicit function theorem to $H(K,\lambda) = 0$ and then obtain the existence of $K(\lambda)$. Clearly we only need to verify surjectivity of $\partial_K H(K^0,\lambda^0)$, where $K^0 = (x^0,u^0,v^0)$

$$\partial_K H(K^0,\lambda^0) = \lim_{\substack{x^j\to x^0\\ u^j\to u^0\\ v^j\to v^0\\ \lambda^j\to\lambda^0}} \left\{ \left. \begin{matrix} \dfrac{\partial(\nabla_x L)}{\partial x}, \nabla g_1\dots\nabla g_s, & \nabla g_{s+1}\dots\nabla g_{m_1} & \\ u_1\nabla g_1 \quad g_1 & & \\ \vdots \qquad\quad \ddots & & O \\ u_s\nabla g_s \qquad\quad g_s & & \\ \nabla g_{s+1} & & \\ \vdots \qquad\quad O & & O \\ \nabla g_{m_1} & & \end{matrix} \right|_{\substack{x = x^j\\ u = u^j\\ v = v^j\\ \lambda = \lambda^j}} \right\}$$

$$= \left\{ \begin{matrix} \partial_x(\nabla_x L)(x^0,u^0,v^0,\lambda^0), & \nabla g_1(x^0)\dots\nabla g_s(x^0), & \nabla g_{s+1}(x^0)\dots\nabla g_{m_1}(x^0) \\ u_1^0\nabla g_1(x^0) & g_1(x^0) & \\ \vdots & \ddots & O \\ u_s^0\nabla g_s(x^0) & g_s(x^0) & \\ \nabla g_{s+1}(x^0) & & \\ \vdots & O & O \\ \nabla g_{m_1}(x^0) & & \end{matrix} \right\}$$

In the proof of theorem 6 of [11], the author pointed out and proved later that if $\nabla_x^2 L(x^0,u^0,v^0,\lambda^0)$, $g_i(x)$ $i = 1,\dots,m_1$ satisfy the conditions cited here in Proposition 5.1, then the matrix

$$\begin{bmatrix} \nabla_x^2 L(x^0,u^0,v^0,\lambda^0), & \nabla g_1(x^0) \dots \nabla g_s(x^0), & \nabla g_{s+1}(x^0) \dots \nabla g_{m_1}(x^0) \\ u_1 \nabla g_1(x^0) & g_1(x^0) & \\ \vdots & \ddots & O \\ u_s \nabla g_s(x^0) & \quad g_s(x^0) & \\ \nabla g_{s+1}(x^0) & & \\ \vdots & O & O \\ \nabla g_{m_1}(x^0) & & \end{bmatrix}$$

is of maximal rank. In exactly the same way one can prove that under the assumptions in this theorem every element of $\partial_K H(K^0,\lambda^0)$ is of maximal rank, i.e. $\partial_K H(K^0,\lambda^0)$ is surjective.

Thus by the Lipschitz implicit function theorem there is a Lipschitz continuous function $K(\lambda) = (x(\lambda),u(\lambda),v(\lambda))$, satisfying $H(K(\lambda),\lambda) = 0$ for all λ in some neighborhood Λ_1 of λ_0 and $K(\lambda^0) = K^0$.

Next we show that such a solution $K(\lambda) = (x(\lambda),u(\lambda),v(\lambda))$ satisfies the optimality condition. In fact, since $K(\lambda)$, $g_i(x)$ are continuous and satisfy the equations

$$u_i g_i(x(\lambda)) = 0 \qquad 1 \le i \le s$$

and

$$u_i^0 g_i(x^0) = 0 \qquad 1 \le i \le s$$

then there exists a neighborhood Λ_2 of λ_0 such that for all $\lambda \in \Lambda_1 \cap \Lambda_2$ we have

i) $u_i(\lambda) > 0$ for those i such that $u_i^0 > 0$, and this implies $g_i(x(\lambda)) = 0$

ii) $g_i(x(\lambda)) < 0$ for those i such that $g_i(x^0) < 0$; this implies $u_i(\lambda) = 0$.

By the strict complementary slackness condition any $i = 1,2,\dots,s$ falls in one of the two categories, i.e. for all $i = 1,\dots,s$

(5.7) $\quad u_i(\lambda) \geq 0 \quad$ and $\quad g_i(x(\lambda)) \leq 0$.

Combining (5.7) and equation $H(x,u,\lambda,\lambda) = 0$, we know that for $\lambda \in \Lambda_1 \cap \Lambda_2$ $(x(\lambda),u(\lambda),v(\lambda))$ satisfies the following conditions:

$$\nabla_x L(x(\lambda),\lambda) + \sum_{i=1}^{s} u_i(\lambda)\, g_i(x(\lambda)) + \sum_{i=s+1}^{m_1} v_i(\lambda) \nabla g_i(x(\lambda)) = 0$$

$$g_i(x(\lambda)) \leq 0 \qquad 1 \leq i \leq s$$

$$g_i(x(\lambda)) = 0 \qquad 1 \leq i \leq m_1$$

$$u_i(\lambda) g_i(x(\lambda)) = 0 \qquad 1 \leq i \leq s$$

$$u_i(\lambda) \geq 0 \qquad 1 \leq i \leq s$$

Using the convexity of $Z(x,\lambda)$ and $g_i(x)$, $i = 1,\dots,m_1$, we come to the conclusion that $x(\lambda)$ is the optimal solution of $Z(\lambda)$ and $u(\lambda)$, $v(\lambda)$ are the associated Lagrange multipliers.

(ii) The Lipschitz continuity of infima function $\varphi(\lambda)$ comes out immediately from the equation

$$\varphi(\lambda) = Z(x(\lambda),\lambda) \quad . \quad \square$$

One can get a bound for the Lipschitz constant of $K(\lambda)$ as we make explicit here below.

Consider two points $K(\lambda)$ and $K(\lambda^0)$, λ is in $\Lambda_1 \cap \Lambda_2$. Then

$$H(K(\lambda),\lambda) = 0 = H(K(\lambda^0),\lambda^0) \quad .$$

Thus

(5.8) $\quad H(K(\lambda),\lambda) - H(K(\lambda^0),\lambda) = H(K(\lambda^0),\lambda^0) - H(K(\lambda^0),\lambda)$.

By the Lipschitz mean value theorem, (Section 3 of [9]), we have

$$(5.9) \qquad H(K(\lambda),\lambda) - H(K(\lambda^0),\lambda) = A\cdot(K(\lambda) - K(\lambda^0))$$

for some $A \in \partial_K H([K(\lambda^0),K(\lambda)],\lambda)$, where

$$[K(\lambda^0),K(\lambda)] = \{K(t): K(t) = K(\lambda^0) + t(K(\lambda) - K(\lambda^0)), t \in [0,1]\} .$$

and

$$\partial_K H([K(\lambda^0),K(\lambda)],\lambda) = \overline{co}\left\{ \bigcup_{K(t)} \partial_K H(K(t),\lambda)\right\} .$$

For the right hand side of (5.8) one can easily get

$$(5.10) \qquad |H(K(\lambda^0),\lambda) - H(K(\lambda^0),\lambda^0)| \le \tilde{\ell}\cdot|\lambda - \lambda_0| .$$

From (5.8), (5.9), (5.10) it follows that

$$|K(\lambda) - K(\lambda^0)| \le \frac{\tilde{\ell}}{|A|}\cdot|\lambda - \lambda^0| .$$

When $\partial_K H(K,\lambda)$ is generated by a finite number of elements, it is not so difficult to give an estimation for $|A|$. Fortunately for some useful piecewise continuous density functions we are in this situation as we show in an example in the next section.

6. AN EXAMPLE

In Propositions 5.1 and 5.6 we imposed some conditions on the density function $D(t,\lambda)$. These conditions are met for most of the useful distribution functions which have continuous or piecewise continuous density functions. Here we give an example to illustrate how the conditions can be verified.

Problem $P(\xi)$:

$$\text{Min } z(x,\xi) = x + E\left\{\begin{matrix}\min \\ y_j \ge 0\end{matrix} \quad y_1 + y_2 \;\middle|\; \begin{matrix} y_1 - y_3 = \xi_1 - x \\ y_2 - y_3 = \xi_2 - x\end{matrix}\right\}$$

s.t. $a \le x \le b$

where $\xi = (\xi_1, \xi_2)'$ is a two-dimensional random variable.

(1) ξ is normal with density function

$$D(t_1,t_2) = \frac{1}{2\pi\sigma_1\sigma_2} \exp - \frac{(t_1-\mu_1)^2}{2\sigma_1^2}\left\{-\left[\rho\frac{(t_1-\mu_1)(t_2-\mu_2)}{\sigma_1\sigma^2} + \frac{(t_2-\mu_2)^2}{2\sigma_2^2}\right]\right\}.$$

We have

$$Q(x,\xi(\omega)) = \begin{cases} \xi_1(\omega) + \xi_2(\omega) - 2x & \text{if} \quad \xi_1(\omega)-x \leq 0, \ \xi_2(\omega)-x \geq 0 \\ \xi_2(\omega) - \xi_1(\omega) & \text{if} \quad \xi_1(\omega)-x \leq 0, \ \xi_2(\omega)-\xi_1(\omega) \geq 0 \\ \xi_1(\omega) - \xi_2(\omega) & \text{if} \quad \xi_2(\omega)-x \leq 0, \ \xi_1(\omega)-\xi_2(\omega) \geq 0 \end{cases} .$$

Thus

$$\begin{aligned} Z(x,\xi) = x &+ \int_x^\infty dt_1 \int_x^\infty (t_1+t_2-2x)D(t_1,t_2)dt_2 \\ &+ \int_{-\infty}^{x} dt_1 \int_{t_1}^\infty (t_2-t_1)D(t_1,t_2)dt_2 \\ &+ \int_{-\infty}^{x} dt_2 \int_{t_2}^\infty (t_1-t_2)D(t_1,t_2)dt_2 \quad . \end{aligned}$$

It is not difficult to verify directly from the expression that $Z(x,\xi)$ and $\nabla_x Z(x,\xi)$ are continuously differentiable in $(x,\sigma_1,\sigma_2,\rho,\mu_1,\mu_2)$. Another way to do it, as shown in Section 2, is first to show uniform convergence of the integrals $\int_x^\infty D(t_1,t_2)dt_i$ $i=1,2$, then show that the uniform convergence will be preserved under transformation $t = B\tau$. We are going to do it with a slightly different integral

$$I(x_1,x_2) =$$

$$\int_{x_1}^\infty dt_1 \int_{x_2}^\infty \frac{1}{2\pi\sigma_1\sigma_2} \exp\left\{-\left[\frac{(t_1-\mu_1)^2}{2\sigma_1^2} - \rho\,\frac{(t_1-\mu_1)(t_2-\mu_2)}{\sigma_1\sigma_2} + \frac{(t_2-\mu_2)^2}{2\sigma_2^2}\right]\right\}dt_2 \ .$$

First, with the help of the transformation

$$t_1' = \frac{t_1 - \mu_1}{\sigma_1} \quad , \qquad t_2' = \frac{t_2 - \mu_2}{\sigma_2}$$

we can assume that $I(x_1,x_2)$ is in the following form

$$I(x_1,x_2) = \int_{x_1}^{\infty} dt_1 \int_{x_2}^{\infty} \frac{1}{2\pi} \exp\left\{-\frac{t_1^2}{2} + \rho t_1 t_2 - \frac{t_2^2}{2}\right\} dt_2 \quad .$$

Now we examine the uniform convergence with respect to t_1 of the integral

$$I_1 = \int_{x_2}^{\infty} \frac{1}{2\pi} e^{-\frac{t_1^2}{2} + \rho t_1 t_2 - \frac{t_2^2}{2}} dt_2$$

$$= \int_{x_2}^{\infty} \frac{1}{2\pi} e^{-\frac{1}{2}(1-\rho^2)t_1^2} \cdot e^{-\frac{1}{2}(t_2-\rho t_1)^2} dt_2 \quad .$$

On any finite interval of t_1, say $\left[t_1^{(1)}, t_1^{(2)}\right]$, we have

$$e^{-\frac{1}{2}(1-\rho^2)t_1^2 - \frac{1}{2}(t_2-\rho t_1)^2} \leq e^{-\frac{1}{2}(1-\rho^2)\bar{t}_1^2} \cdot e^{-\frac{1}{2}(t_2-\rho \bar{\bar{t}}_1)^2}$$

for t_2 large enough, where $\bar{t}_1 = \min\left\{|t_1|, t_1 \in \left[t_1^{(1)}, t_1^{(2)}\right]\right\}$ and $\bar{\bar{t}} = \max\left\{|t_1|, t_1 \in \left[t_1^{(1)}, t_1^{(2)}\right]\right\}$.

Therefore

$$\int_{M}^{\infty} e^{-\frac{1}{2}(1-\rho)t_1^2 - \frac{1}{2}(t_2-\rho t_1)^2} dt_2 \leq \int_{M}^{\infty} e^{-\frac{1}{2}(1-\rho^2)\bar{t}_1^2} \cdot e^{-\frac{1}{2}(t_2-\rho \bar{\bar{t}}_1)^2} dt_2$$

for large positive number M, and the left integral tends to zero uniformly with respect to t_1 on $\left[t_1^{(1)}, t_1^{(2)}\right]$ as $M \to +\infty$ for any $\rho \leq 1$. This is equivalent to I_1 converges uniformly with respect to t_1 on $\left[t_1^{(1)}, t_1^{(2)}\right]$.

By symmetry of t_1, t_2 we obtain uniform convergence of the integral

$$I_2 = \int_{x_1}^{\infty} \frac{1}{2\pi} e^{-\frac{t_1^2}{2} + \rho t_1 t_2 - \frac{t_2^2}{2}} dt_1$$

with respect to t_2 on any finite interval for any $\rho \leq 1$.

Thus the required uniform convergence holds for any σ_1, σ_2, ρ, μ_1, μ_2, i.e., for any covariance matrix V and any mean value, if the integral is in the standard form.

Suppose that $I(x_1,x_2)$ has the following form

$$I(x_1,x_2) = \iint\limits_{Bt \geq x} \frac{1}{2\pi} e^{-\frac{t_1^2}{2} + \rho t_1 t_2 - \frac{t_2^2}{2}} dt_1 dt_2$$

where B is a nonsingular square matrix, with the transformation $\tau = Bt$ we can reduce $I(x_1,x_2)$ into a standard form, with covariance matrix $(B^{-1})^T V B^{-1}$. Then the required uniform covergence is at hand.

(2) ξ is uniformly distributed in the square $[-a,a;-b,b]$, $a>b>0$.

Then

$$Z(x,a,b) = \begin{cases} x + \frac{1}{12ab}(3bx^2 + 3ax^2 - 6abx - 2x^3) + c_1(a,b) & -b \leq x \leq b \\ \frac{x^2}{2a} + c_2(a,b) & -a \leq x < b \\ -x & -a > x \end{cases}$$

where $c_1(a,b)$, $c_2(a,b)$ are polynomials of a and b.

The expression of Z for $x > b$ is omitted here, the analytical behavior of Z there is quite similar.

Then

$$\nabla_x Z(x,a,b) = \begin{cases} 1 + \frac{1}{2ab}(bx + ax - ab - x^2) & -b \leq x \leq b \\ \frac{x}{a} & -a \leq x < -b \\ -1 & -a > x \end{cases}$$

Clearly $\nabla_x Z(x,a,b)$ is differentiable except at points $x = -b$, $x = -a$. At those points the generalized partial derivatives can be computed out easily:

$$\partial_x\left\{\nabla_x Z(-b,a,b)\right\} = \left[\frac{1}{a}, \frac{a+3b}{2ab}\right]$$

and

$$\partial_x\{\nabla_x Z(-a,a,b)\} = [0,\frac{1}{a}] \quad .$$

With this one can easily check whether the conditions imposed in Proposition 5.6 are satisfied.

ACKNOWLEDGMENT

The author wishes to thank Roger Wets for his guidance in preparing this paper.

REFERENCES

[1] Wets, R. 1974. Stochastic programs with recourse: The equivalent deterministic program. SIAM Review 16, 309-339.

[2] Dupačova, J. 1983. Stability in stochastic programming with recourse. Estimated parameters. Mathematical Programming (to appear).

[3] Kall, P. 1966. Qualitative Aussagen zu einigen Problemen der stochastischen Programmierung. Z. Wahrscheinlichkeitstheorie Verw. Gebiete 6, 246-272.

[4] Walkup, D. and R. Wets. 1969. Lifting projections of convex polyhedra. Pacific J. Math. 28, 465-475.

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

[6] Attouch, H. and R. Wets. 1980. Approximation and convergence in nonlinear optimization. In: Nonlinear Programming 4, O.L. Mangasarian, R. Meyer, S. Robinson, eds. Academic Press, New York, 367-394.

[7] Wets, R. 1980. Convergence of convex functions. In: Variational Inequalities and Complementarity Problems, R. Cottle, F. Gianessi and J.-L. Lions, eds. John Wiley, London, 375-403.

[8] Fiacco, A. 1976. Sensitivity analysis for nonlinear programming, using penalty methods. Mathematical Programming 10, 287-311.

[9] Clarke, F.H. 1976. On the inverse function theorem. Pacific J. Math. 64(1).

[10] Pourciau, B.H. 1977. Analysis and optimization of Lipschitz mappings. JOTA 22(3).

[11] Fiacco, A. and M.P. McCormick. 1968. Nonlinear Programming: Sequential Unconstrained Minimization Techniques. John Wiley, New York.