

STABILITY IN STOCHASTIC PROGRAMMING WITH RECOURSE. CONTAMINATED DISTRIBUTIONS

J. DUPAČOVÁ

Charles University, Prague, Czechoslovakia

Received 1 December 1983

Revised manuscript received 27 July 1984

In the paper, stability of the optimal solution of a stochastic program with recourse with respect to small changes of the underlying distribution of random coefficients is considered. As a tool, contamination of the given distribution by another one is suggested and the original stability problem is thus reduced to that with linearly perturbed objective function. The theory of perturbed Kuhn-Tucker points and strongly regular equations is used to get explicit formulas for Gâteaux differentials of optimal solutions under different assumptions. Possible exploitation of the results for further robustness studies is indicated.

Key words: Stochastic Programming, Incomplete Information, Stability, Robustness, Contamination, Simple Recourse Problem.

Consider the following stochastic programming problem:

Maximize $f(x; F) = E_F\{c(x) - \varphi(x; z)\}$ on the set \mathcal{X} where $\mathcal{X} \subset \mathbb{R}^n$ is a nonempty closed convex set of admissible solutions, $c: \mathcal{X} \rightarrow \mathbb{R}^1$ is a given function, F is a given joint probability distribution of a random vector z on $(\mathcal{Z}, \mathcal{B}_Z)$, $\mathcal{Z} \subset \mathbb{R}^l$, $\varphi: \mathcal{X} \times \mathcal{Z} \rightarrow \mathbb{R}^1$ is a given nonnegative function such that $\varphi(x; z)$ are measurable for all $x \in \mathcal{X}$. (1)

An example of (1) is when a nonlinear program

$$\begin{aligned} &\text{maximize} && c(x) \\ &\text{subject to} && g_k(x; z) \geq 0, \quad 1 \leq k \leq m, \quad x \in \mathcal{X}, \end{aligned}$$

contains random parameters in $g_k(x; z)$, $1 \leq k \leq m$, and the decision $x \in \mathcal{X}$ has to be chosen before the values of these random parameters are observed. The function $\varphi(x; z)$ evaluates the loss corresponding to the case that the chosen $x \in \mathcal{X}$ does not fulfil the constraints $g_k(x; z) \geq 0$, $1 \leq k \leq m$ for the observed values of the random parameters.

The essential results concerning the objective function in (1) are summarized in the following lemma (see e.g. [10]):

Lemma. Let $\varphi: \mathcal{X} \times \mathcal{Z} \rightarrow \mathbb{R}^1$ be Lipschitz continuous on the set \mathcal{X} for an arbitrary $z \in \mathcal{Z}$ and let the Lipschitzian constant $k_1(z)$ be integrable with respect to F . Let the gradient

$\nabla_x \varphi(x; z)$ exist for $x \in \mathcal{X}$ almost surely with respect to F and let $E_F \varphi(x; z)$ be finite at least for one point $x \in \mathcal{X}$. Then $E_F \varphi(x; z)$ is Lipschitz continuous on \mathcal{X} , the gradient $\nabla_x E_F \varphi(x; z)$ exists for $x \in \mathcal{X}$ and

$$\nabla_x E_F \varphi(x; z) = E_F \nabla_x \varphi(x; z). \quad (2)$$

Remark 1. Under assumptions of the Lemma, the existence of the expectation

$$f(x; F) = E_F \{c(x) - \varphi(x; z)\}$$

is evidently guaranteed for all $x \in \mathcal{X}$. Under the additional assumption that $\varphi(\cdot; z)$ is convex on \mathcal{X} for all $z \in \mathcal{Z}$ and c is concave, then the function $f(x; F)$ is concave, differentiable on \mathcal{X} .

As in stochastic linear programming, the optimal solution $x(F)$ of (1) (provided that it exists) depends on the assumed distribution F . In many real-life situations, however, the assumption of a completely known distribution F is hardly acceptable and the solution of (1) should be thus at least supplemented by a proper stability study with respect to F . In the robust case, a small change in the distribution F should cause only a small change of the optimal solution. In the preceding papers [4, 5], the first attempts were made to study stability of the optimal solution $x(F)$ of (1) with respect to the distribution F and its parameters through completing the approaches developed for nonlinear programming stability studies by suitable statistical methods. In this paper, local behaviour of $x(F)$ will be studied via t -contamination of F by a distribution G belonging to a specified set of distributions (see [3, 4] for special cases), i.e., instead of F , distributions of the form

$$F_t = (1-t)F + tG, \quad 0 \leq t \leq 1, \quad (3)$$

will be considered. In (3), F_t is called *distribution F t -contaminated by distribution G* and for our purpose, the Gâteaux differential

$$dx(F; G - F) = \lim_{t \rightarrow 0^+} \frac{x(F + t(G - F)) - x(F)}{t}$$

of the optimal solution $x(F)$ at F in the direction of $G - F$ is of importance.

Disregarding the constraints (i.e., taking $\mathcal{X} = \mathbb{R}^n$), the optimal solution $x(F_t)$ of the program

$$\text{maximize } f(x; F_t) = E_{F_t} \{c(x) - \varphi(x; z)\} \quad (4)$$

should fulfil the system of n equations

$$\Psi(x; F_t) = 0$$

where (for F, G fixed) $\Psi: \mathbb{R}^n \times \langle 0, 1 \rangle \rightarrow \mathbb{R}^n$ and its components

$$\psi_j(x; F_t) = \frac{\partial}{\partial x_j} f(x; F_t), \quad 1 \leq j \leq n,$$

are assumed to exist for all j . Obviously

$$\Psi(x; F_t) = \Psi(x; F) + t[\Psi(x; G) - \Psi(x; F)], \quad 0 \leq t \leq 1.$$

Using the implicit function theorem, the Gâteaux differential $dx(F; G - F)$ can be computed under suitable differentiability and regularity assumptions; taking into account that $\psi(x(F); F) = 0$, we get

$$dx(F; G - F) = -D^{-1}\Psi(x(F); G)$$

where

$$D = \left(\frac{\partial \psi_j(x(F); F)}{\partial x_k} \right) = \left(\frac{\partial^2 f(x(F); F)}{\partial x_j \partial x_k} \right), \quad 1 \leq j, k \leq n.$$

To obtain the Gâteaux differential $dx(F; G - F)$ of the optimal solution of (1), we shall use the theory of perturbed Kuhn–Tucker points and strongly regular equations developed in [12], [13]. In principle, it is possible to get Gâteaux differentials of optimal solutions for probabilistic constrained programs using similar tools.

The knowledge of the Gâteaux differential of $x(F)$ at F in the direction of $G - F$ is useful not only for the first order approximation of the optimal solutions corresponding to distributions belonging to a neighbourhood of F but also for deeper statistical conclusions on robustness, namely, in connection with statistical properties of the estimate $x(F_\nu)$ of $x(F)$, which is based on the empirical distribution F_ν . For the special choices $G = \delta_u$ (degenerated distributions concentrated at one point u), the Gâteaux differential $dx(F; \delta_u - F)$ corresponds to the influence curve $\Omega_F(u)$ widely used in asymptotic statistics. Different characteristics of $\Omega_F(y)$ suggested in [9] measure the effect of contamination of the data by gross errors, the local effect of rounding or grouping of the observations, etc.

We shall concentrate upon obtaining formulas for the Gâteaux differentials under different assumptions leaving the detailed investigation of the statistical aspects to a forthcoming paper. We shall start with the general constrained case with

$$\mathcal{X} = \{x \in \mathbb{R}^n: g_i(x) \geq 0, 1 \leq i \leq m, h_r(x) = 0, 1 \leq r \leq p\};$$

the Lagrange function and the Kuhn–Tucker points will be denoted by

$$L(w; F) = f(x; F) + \sum_{i=1}^m u_i g_i(x) + \sum_{r=1}^p v_r h_r(x),$$

$$w(F) = [x(F), u(F), v(F)] \text{ and } I(F) = \{i: g_i(x(F)) = 0\}.$$

Theorem 1. *For the program*

$$\text{maximize } f(x; F) := E_F\{c(x) - \varphi(x; z)\} \quad \text{on the set } \mathcal{X} \tag{5}$$

assume

$$(i) \quad \mathcal{X} = \{x \in \mathbb{R}^n: g_i(x) \geq 0, 1 \leq i \leq m, h_r(x) = 0, 1 \leq r \leq p\} \neq \emptyset, \tag{6}$$

$g_i, 1 \leq i \leq m, h_r, 1 \leq r \leq p$, are twice continuously differentiable,

(ii) $c: \mathcal{X} \rightarrow \mathbb{R}^1$ is twice continuously differentiable.

(iii) The distribution F on $(\mathcal{X}, \mathcal{B}_{\mathcal{X}})$ and the function $\varphi: \mathcal{X} \times \mathcal{Z} \rightarrow \mathbb{R}^1$ fulfil the assumptions of the Lemma and the mean value $E_F \varphi(x; z)$ is twice continuously differentiable.

(iv) For the program (5) with \mathcal{X} given by (6), Kuhn-Tucker conditions of the first and second order, the linear independence condition and the strict complementarity conditions are fulfilled for $w(F) = [x(F), u(F), v(F)] \in \mathbb{R}^n \times \mathbb{R}_+^m \times \mathbb{R}^p$, and the matrix

$$C = \nabla_{xx}^2 L(w(F); F)$$

is nonsingular.

(v) There is a neighbourhood $\mathcal{O}(x(F)) \subset \mathbb{R}^n$ on which φ and the distribution G on $(\mathcal{Z}, \mathcal{B}_{\mathcal{Z}})$ fulfil the assumptions of the Lemma and $E_G \varphi(x; z)$ is twice continuously differentiable on $\mathcal{O}(x(F))$.

Then: (a) There is a neighbourhood $\mathcal{O}(w(F)) \subset \mathbb{R}^n \times \mathbb{R}_+^m \times \mathbb{R}^p$, a real number $t_0 > 0$ and a continuous function $w: (0, t_0) \rightarrow \mathcal{O}(w(F))$, $w(0) = w(F)$ such that for any $t \in (0, t_0)$, $w(t) = [x(t), u(t), v(t)]$ is the Kuhn-Tucker point of

$$\max_{x \in \mathcal{X}} f(x; F_t) := E_{F_t} \{c(x) - \varphi(x; z)\} \quad (7)$$

for which the second order sufficient condition, the linear independence condition and the strict complementarity conditions are fulfilled.

(b) The Gâteaux differential $dx(F; G - F)$ of the isolated local maximizer $x(F)$ of (5), (6) in the direction of $G - F$ is given by

$$dx(F; G - F) = -D^{-1} \nabla_x L(w(F); G), \quad (8)$$

where

$$D^{-1} = [I - C^{-1} P (P^T C^{-1} P)^{-1} P^T] C^{-1}, \quad (9)$$

$$P = [\nabla_x g_i(x(F)), i \in I(F), \nabla_x h_r(x(F)), 1 \leq r \leq p]$$

and I is the n -dimensional unit matrix.

Proof. The first assertion of Theorem 1 can be proved by means of the implicit function theorem as in [12, Theorem 2.1]. To prove the second assertion, we shall use the implicit function theorem once more. (See also [7] for a similar approach.)

For the sake of simplicity assume that $I(F) = \{1, \dots, s\}$, denote by $\bar{u} \in \mathbb{R}^s$ the projection of u into R^s and define

$$\Psi: \mathbb{R}^n \times \mathbb{R}^s \times \mathbb{R}^p \times \langle 0, 1 \rangle \rightarrow \mathbb{R}^n \times \mathbb{R}^s \times \mathbb{R}^p,$$

a vector valued function whose components are

$$\psi_j(x, \bar{u}, v; t) = \frac{\partial f(x; F_t)}{\partial x_j} + \sum_{i=1}^s u_i \frac{\partial g_i(x)}{\partial x_j} + \sum_{r=1}^p v_r \frac{\partial h_r(x)}{\partial x_j}, \quad 1 \leq j \leq n,$$

$$\psi_{n+i}(x, \bar{u}, v; t) = g_i(x), \quad 1 \leq i \leq s,$$

$$\psi_{n+s+i}(x, \bar{u}, v; t) = h_r(x), \quad 1 \leq r \leq p.$$

Under our specification of $I(F)$, the system $\Psi(x, \bar{u}, v; t) = 0$ together with $u_i = 0$, $s + 1 \leq i \leq m$, forms the local Kuhn-Tucker conditions of the first order for problem (7) with $t < t_0$ and according to (i), (ii), (iii), (v), there exists a neighbourhood $\mathcal{O}(x(F))$ on which $\Psi(x, \bar{u}, v; t)$ is continuously differentiable with respect to all variables. The matrix

$$D(x, \bar{u}, v; t) = \begin{pmatrix} \frac{\partial \Psi(x, \bar{u}, v; t)}{\partial x_j}, 1 \leq j \leq n, & \frac{\partial \Psi(x, \bar{u}, v; t)}{\partial u_i}, 1 \leq i \leq s, & \\ & & \frac{\partial \Psi(x, \bar{u}, v, t)}{\partial v_r}, 1 \leq r \leq p \end{pmatrix}$$

$$= \begin{pmatrix} \nabla_{xx}^2 L(w; F_t) & P \\ P^T & 0 \end{pmatrix}$$

with $w = (x, \bar{u}, 0, v)$ and $0 \in \mathbb{R}^{m-s}$ is nonsingular for $w = w(F)$ according to (iv). This implies the existence of continuous right-hand derivatives of $\bar{w}(t) = [x(t), \bar{u}(t), v(t)]$ at 0:

$$\frac{d\bar{w}(0^+)}{dt} = -D(\bar{w}(0); 0)^{-1} \frac{\partial}{\partial t} \Psi(\bar{w}(0); 0) \tag{10}$$

where

$$\begin{aligned} & \frac{\partial}{\partial t} \Psi(\bar{w}(0); 0) \\ &= \begin{pmatrix} \nabla_x f(x(F); G) - \nabla_x f(x(F); F) \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} \nabla_x f(x(F); G) + \sum_{i=1}^s u_i(F) \nabla_x g_i(x(F)) + \sum_{r=1}^p v_r(F) \nabla_x h_r(x(F)) \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} \nabla_x L(w(F); G) \\ 0 \end{pmatrix}. \end{aligned} \tag{11}$$

Formula (8), (9) follows from (10), (11) by inversion of the block matrix $D(\bar{w}(0); 0) = \begin{pmatrix} C & P \\ P^T & 0 \end{pmatrix}$; the first equality in (11) follows from $\nabla_x L(w(F); F) = 0$.

Remark. Due to the fact that (7) is a special type of a linearly perturbed nonlinear program, the Gâteaux differentials of $x(F)$ at F both in the direction of $G-F$ and in the direction of G are equal: $dx(F; G-F) = dx(F; G)$.

For \mathcal{X} polyhedral we get, as a special case of Theorem 1,

Theorem 2. *Let in the problem*

$$\max_{x \in \mathcal{X}} f(x; F) = E_F \{c(x) - \varphi(x; z)\} \tag{12}$$

the following assumptions be fulfilled:

(i) $\mathcal{X} = \{x \in \mathbb{R}^n : Px = p, x \geq 0\} \neq \emptyset$, $P(r, n)$, $p \in \mathbb{R}^r$ are a given matrix of rank r and a given vector; let the vertices of \mathcal{X} be nondegenerate.

(ii) $c: \mathcal{X} \rightarrow \mathbb{R}^1$ is twice continuously differentiable.

(iii) The distribution F on $(\mathcal{X}, \mathcal{B}_{\mathcal{X}})$ and the function $\varphi: \mathcal{X} \times \mathcal{Z} \rightarrow \mathbb{R}^1$ fulfil the assumptions of the Lemma and the mean value $E_F \varphi(x; z)$ is twice continuously differentiable.

(iv) There exists a Kuhn-Tucker point $[x(F); \pi(F)]$ for (12) such that the second order sufficient condition and the strict complementarity conditions are fulfilled. For $J = \{j: x_j(F) > 0\}$, the matrix

$$C_J = \left(\frac{\partial^2 f(x(F); F)}{\partial x_j \partial x_k} \right), \quad j, k \in J, \quad (13)$$

is nonsingular.

(v) There is a neighbourhood $\mathcal{C}(x(F)) \subset \mathbb{R}^n$ on which the function φ and the distribution G on $(\mathcal{X}, \mathcal{B}_{\mathcal{X}})$ fulfil assumptions of the Lemma and $f(x; G)$ is twice continuously differentiable on $\mathcal{C}(x(F))$.

Then (a) There are neighbourhoods $\mathcal{C}_1(x(F)) \subset \mathcal{C}(x(F))$, $\mathcal{V}(\pi(F)) \subset \mathbb{R}^r$, a real number $t_0 > 0$ and continuous functions

$$\begin{aligned} x: (0, t_0) &\rightarrow \mathcal{C}_1(x(F)), & x(0) &= x(F), \\ \pi: (0, t_0) &\rightarrow \mathcal{V}(\pi(F)), & \pi(0) &= \pi(F), \end{aligned}$$

such that for any $t \in (0, t_0)$, $[x(t); \pi(t)]$ is a Kuhn-Tucker point for the problem

$$\max_{x \in \mathcal{X}} f(x; F_t) = E_{F_t} \{c(x) - \varphi(x; z)\} \quad (14)$$

with $F_t = (1-t)F + tG$, $0 \leq t \leq 1$. The second order sufficient condition and the strict complementarity conditions are fulfilled for $[x(t); \pi(t)]$ and

$$\begin{aligned} x_j(F_t) &= x_j(t) = 0, & j &\notin J, \\ x_j(F_t) &= x_j(t) > 0, & j &\in J. \end{aligned}$$

(b) The vector $dx_j(F; G - F)$ of the components of the Gâteaux differential of the isolated local maximizer $x(F)$ of (12) in the direction of $G - F$ for $j \in J$ is given by

$$dx_j(F; G - F) = -D_J^{-1} \Psi_J(x(F); \pi(F); G) \quad (15)$$

where

$$\Psi_J(x(F); \pi(F); G) = \left(\frac{\partial}{\partial x_j} f(x(F); G) + \sum_{k=1}^r p_{kj} \pi_k(F) \right)_{j \in J},$$

$$D_J^{-1} = [I_J - C_J^{-1} P_J^T (P_J C_J^{-1} P_J^T)^{-1} P_J] C_J^{-1},$$

$$P_J = (p_{kj})_{\substack{1 \leq k \leq r \\ j \in J}}$$

I_J is a unit matrix of dimension $s = \text{card } J$ and C_J is given by (13).

The remaining components of the Gâteaux differential $dx(F; G - F)$ are equal to zero.

In the special case of a simple recourse problem with random right-hand sides and with $\mathcal{X} = \mathbb{R}_+^n$, i.e., for

$$\max_{x \geq 0} E_F \left\{ c^T x - \sum_{i=1}^m q_i \left(\sum_{j=1}^n a_{ij} x_j - b_i \right)^+ \right\}, \tag{16}$$

we have the following theorem:

Theorem 3 [4]. Assume:

- (i) F is an m -dimensional continuous distribution of b for which $E_F b$ exists.
- (ii) The optimal solution $x(F)$ of (16) exists and the strict complementarity conditions hold true. Let $J = \{j : x_j(F) > 0\}$.
- (iii) $q_i > 0$, $1 \leq i \leq m$, $A_j = (a_{ij})$, $1 \leq i \leq m$, $j \in J$ has full column rank.
- (iv) The marginal densities f_i , $1 \leq i \leq m$, are continuous and positive at the points $X_i(F) = \sum_{j \in J} a_{ij} x_j(F)$, $1 \leq i \leq m$, respectively.
- (v) G is an m -dimensional distribution whose marginal distribution functions G_i have continuous derivatives in a neighbourhood of the points $X_i(F) = \sum_{j \in J} a_{ij} x_j(F)$, $1 \leq i \leq m$, respectively.

Then (a) There is a neighbourhood $\mathcal{O}(x(F))$ and a real number $t_0 > 0$ such that the program

$$\max_{x \geq 0} E_{F_t} \left\{ c^T x - \sum_{i=1}^m q_i \left(\sum_{j=1}^n a_{ij} x_j - b_i \right)^+ \right\} \tag{17}$$

with $F_t = (1 - t)F + tG$ has a unique optimal solution $x(F_t) \in \mathcal{O}(x(F))$ for any $0 \leq t < t_0$, $x_j(F_t)$, $j \in J$ are nonzero components of $x(F_t)$ and $x_j(F_t) = 0$ for $j \notin J$.

(b) Components of the Gâteaux differential of the optimal solution $x(F)$ at F in the direction of $G - F$ corresponding to the nonzero components of $x(F)$ are given by

$$dx_j(F; G - F) = (A_j^T K A_j)^{-1} (c_j - A_j^T k) \tag{18}$$

where $c_j = (c_j)_{j \in J}$, $k = (k_i)$, $1 \leq i \leq m$ with

$$k_i = q_i G_i \left(\sum_{h \in J} a_{ih} x_h(F) \right), \quad 1 \leq i \leq m,$$

and

$$K = \text{diag} \left\{ q_i f_i \left(\sum_{h \in J} a_{ih} x_h(F) \right), \quad 1 \leq i \leq m \right\}.$$

Theorem 3 illustrates, inter alia that the assumption of twice continuous differentiability of $f(x; F)$ can be fulfilled in practice. For detailed discussion of this question see [15].

Example. Let us compute the influence curve corresponding to the case considered in Theorem 3. Having solved the program (16) for the chosen distribution F , the set J , reduced matrices A_J, c_J and the diagonal matrix K are known. Let $u \neq Ax(F)$ be a chosen point and $G = \delta_u$. For the vector k we have $k = q(u; F)$, where

$$\begin{aligned} q_i(u; F) &= q_i \quad \text{if } u_i < X_i(F), \\ &= 0 \quad \text{otherwise.} \end{aligned}$$

The influence curve $\Omega_F(u)$ is given by

$$\Omega_F(u) = (A_J^T K A_J)^{-1} (c_J - A_J^T q(u; F))$$

and to get its characteristics, e.g., the gross-error-sensitivity

$$\gamma^* = \sup_u \|\Omega_F(u)\|$$

means to solve a discrete optimization problem

$$\text{maximize } \sum_k \left[\sum_j \alpha_{kj} \left(c_j - \sum_i \delta_i q_i a_{ij} \right) \right]^2 \quad (19)$$

with zero-one variables $\delta_i, 1 \leq i \leq m$. (In (19), α_{kj} 's denote the corresponding elements of $(A_J^T K A_J)^{-1}$.)

The assumptions of strict complementarity play an essential role in the proofs of Theorems 1, 2 and 3. They guarantee that the interval $\langle 0, t_0 \rangle$ on which $w(t)$ (resp. $[x(t), \pi(t)]$) is the Kuhn-Tucker point of (7) (resp. of (14)) is nonempty. Alternatively, the strict complementarity conditions can be replaced by the strong second order sufficient condition [13] which gives the existence of continuous Kuhn-Tucker points on a nonempty interval $\langle 0, t_0 \rangle$. This approach was applied in imbedding methods [8] and it will be used to get parallel results in our case.

Without assuming the strict complementarity conditions in (5), (6) denote

$$I^+(F) = \{i: g_i(x(F)) = 0 \text{ and } u_i(F) > 0\},$$

$$I^0(F) = \{i: g_i(x(F)) = 0 \text{ and } u_i(F) = 0\}$$

and formulate the strong second order sufficient condition [13]:

For each $y \neq 0$ with

$$y^T \nabla_x g_i(x(F)) = 0, \quad i \in I^+(F), \quad y^T \nabla_x h_r(x(F)) = 0, \quad 1 \leq r \leq p,$$

the inequality $y^T \nabla_{xx}^2 L(x(F), u(F), v(F)) y < 0$ holds true.

Theorem 4. Let assumptions (i)–(iii), (v) of Theorem 1 be fulfilled and the assumption (iv) be replaced by

(iv') For the program (5) with \mathcal{X} given by (6), the linear independence condition and the strong second order sufficient condition are fulfilled for the Kuhn-Tucker point

$$w(F) = [x(F), u(F), v(F)].$$

Then: (a) *There is a neighbourhood $\mathcal{O}(w(F)) \subset \mathbb{R}^n \times \mathbb{R}_+^m \times \mathbb{R}^P$, a real number $t_0 > 0$ and a continuous function*

$$w: (0, t_0) \rightarrow \mathcal{O}(w(F)), \quad w(0) = w(F),$$

such that for any $t \in (0, t_0)$, $w(t) = [x(t), u(t), v(t)]$ is the Kuhn-Tucker point of

$$\max_{x \in \mathcal{X}} f(x; F_t) := E_{F_t}\{c(x) - \varphi(x; z)\}.$$

(b) *The Gâteaux differential $dx(F; G-F)$ is the unique solution of the quadratic program*

$$\begin{aligned} &\text{maximize} \quad \frac{1}{2}x^T \nabla_{xx}^2 L(w(F); F)x + x^T \nabla_x L(w(F); G) \\ &\text{subject to} \quad x^T \nabla_x g_i(x(F)) = 0, \quad i \in I^+(F), \\ &\quad \quad \quad x^T \nabla_x g_i(x(F)) \geq 0, \quad i \in I^0(F), \\ &\quad \quad \quad x^T \nabla_x h_r(x(F)) = 0, \quad 1 \leq r \leq p, \end{aligned} \tag{20}$$

and $du(F; G-F)$, $dv(F; G-F)$ are the unique Lagrange multipliers for (20) related with $dx(F; G-F)$ with zero components $du_i(F; G-F)$ for $i \notin I(F)$.

Proof. The first part follows from [13, Theorem 2.1] and the second one is a variant of [8, Theorem 5].

Assuming strict complementarity condition valid for the optimal solution $dx(F; G-F)$ of (20), one can get the new active set $I(F_t)$ for (7) with t small enough:

$$I(F_t) = I^+(F) \cup \{i \in I^0(F) : dx(F; G-F)^T \nabla_x g_i(x(F)) = 0\}$$

[8, Corollary 1]. For the special case of Theorem 2 we have

$$I^+(F) = \left\{ j : x_j(F) = 0 \text{ and } \frac{\partial}{\partial x_j} f(x(F); F) + \sum_k p_{k_j} \pi_k(F) < 0 \right\},$$

$$I^0(F) = \left\{ j : x_j(F) = 0 \text{ and } \frac{\partial}{\partial x_j} f(x(F); F) + \sum_k p_{k_j} \pi_k(F) = 0 \right\},$$

so that $J = \{1, \dots, n\} - [I^+(F) \cup I^0(F)]$. The program (20) has the form

$$\begin{aligned} &\text{maximize} \quad \frac{1}{2}x^T \nabla_{xx}^2 f(x(F); F)x + x^T \nabla_x f(x(F); G) \\ &\text{subject to} \quad x_j = 0, \quad j \in I^+(F), \quad x_j \geq 0, \quad j \in I^0(F), \end{aligned}$$

and for $t > 0$ small enough, the new set $J(F_t) = \{j : x_j(F_t) > 0\}$ fulfils

$$J \subset J(F_t) \subset J \cup I^0(F).$$

In the simple case where $I^0(F) = \{j_0\}$, we have explicitly

Theorem 5. *Let assumptions (i)-(iii), (v) of Theorem 2 be fulfilled. Assume further the existence of a Kuhn-Tucker point $[x(F), \pi(F)]$ for (12) such that the strong second order sufficient condition is fulfilled and $I^0(F) = \{j_0\}$.*

Then: (a) There exist a neighbourhood $\mathcal{O}_1(x(F))$ and a real number $t_0 > 0$ such that, for $0 \leq t < t_0$,

$$x_j(t) = 0, \quad j \notin J^0 = J \cup I^0(F)$$

and $x(t)$ is the isolated local maximizer of one of the following problems:

$$\max\{f(x; F_t): x \in \mathcal{X}_j\} \quad \text{or} \quad \max\{f(x; F_t): x \in \mathcal{X}_{j^0}\}$$

where for $H \subset \{1, \dots, n\}$,

$$\mathcal{X}_H = \{x \in \mathbb{R}^n: Px = p, x_j = 0, j \notin H\}.$$

(b) Correspondingly, the components of the Gâteaux differential are

$$\begin{pmatrix} dx_H(F; G-F) \\ d\pi(F; G-F) \end{pmatrix} = - \begin{pmatrix} C_H & P_H^T \\ P_H & 0 \end{pmatrix}^{-1} \begin{pmatrix} \Psi_H(x(F), \pi(F); G) \\ 0 \end{pmatrix}$$

where for $H = J$ or J^0

$$C_H = \left(\frac{\partial^2 f(x(F); F)}{\partial x_i \partial x_j} \right)_{i,j \in H}, \quad P_H = (p_{kj})_{\substack{1 \leq k \leq r \\ j \in H}},$$

$$\Psi_H(x(F), \pi(F); G) = \left(\frac{\partial f(x(F); G)}{\partial x_j} + \sum_{k=1}^r p_{kj} \pi_k(F) \right)_{j \in H}$$

The remaining components of $dx(F; G-F)$ for $j \notin H$ equal zero.

By specifying the set \mathcal{G} of distributions G under consideration, the effect of t -contamination of F by distributions belonging to \mathcal{G} on the optimal solution $x(F)$ can be studied. As a rule, $F \in \mathcal{G}$. Typical examples are

1. F uniform distribution of the random vector z on a closed interval $I \subset \mathbb{R}^l$ and \mathcal{G} the set of distributions such that

$$E_G z = E_F z \quad \text{and} \quad P_G(z \in I) = 1 \quad \forall G \in \mathcal{G}. \quad (21)$$

2. The marginal distributions F_i are normal $N(\mu_i, \sigma_i^2)$ and \mathcal{G} is the set of distributions of the random vector z or \mathbb{R}^l such that

$$E_G z_i = \mu_i, \quad \text{var}_G z_i = \sigma_i^2, \quad 1 \leq i \leq l, \quad \forall G \in \mathcal{G}. \quad (22)$$

In this context, the extremal distributions belonging to \mathcal{G} are of main interest. For the derivative of the objective function in (7) or (14)

$$\frac{\partial}{\partial t} f(x; F_t) = f(x; G) - f(x; F)$$

we have, for all $G \in \mathcal{G}$,

$$\inf_{G \in \mathcal{G}} f(x; G) - f(x; F) \leq \frac{\partial}{\partial t} f(x; F_t) \leq \sup_{G \in \mathcal{G}} f(x; G) - f(x; F).$$

Let G^* , G^{**} be such that

$$\inf_{G \in \mathcal{G}} f(x; G) = f(x; G^*), \quad \sup_{G \in \mathcal{G}} f(x; G) = f(x; G^{**}).$$

The local changes of $x(F)$ in the direction of G^*-F or $G^{**}-F$ give the extremal local decrease or increase of the optimal value of the objective function $f(x; F)$. The corresponding problem, for $G = G^*$,

$$\max_{x \in \mathcal{X}} f(x; (1-t)F + tG^*) = \max_{x \in \mathcal{X}} [(1-t)f(x; F) + t \inf_{G \in \mathcal{G}} f(x; G)]$$

can be evidently related to the Hodges–Lehman decision rule [14] or to the Nadeau–Theodorescu restricted Bayes strategies [11]. The existence of the extremal distributions G^* , G^{**} has been proved for wide classes of recourse problems and for various sets \mathcal{G} of distributions, e.g., for the sets \mathcal{G} given by (21) and (22). For details see [1, 2, 6].

Acknowledgement

I would like to thank an anonymous referee for suggestions which helped to improve upon an earlier version of the paper.

References

- [1] J. Dupačová, "Minimaxová úloha stochastického lineárního programování a momentový problém", *Ekonomicko-matematický obzor* 13 (1977) 279–307. (Extended abstract: "Minimax approach to stochastic linear programming and the moment problem. Selected results", *Zeitschrift für Angewandte Mathematik und Mechanik* 58 (1978) T466–T467.)
- [2] J. Dupačová, "On minimax decision rule in stochastic linear programming", in: A. Prékopa, ed., *Mathematical methods of operations research 1* (Akademiai Kiadó, Budapest, 1980) pp. 38–48.
- [3] J. Dupačová, "Stability studies in stochastic programs with recourse. The special case", *Zeitschrift für Angewandte Mathematik und Mechanik* 62 (1982) T369–T370.
- [4] J. Dupačová, "Stability in stochastic programming with recourse", *Acta Universitatis Carolinae-Mathematica et Physica* 24 (1983) 23–34.
- [5] J. Dupačová, "Stability in stochastic programming with recourse-Estimated parameters", *Mathematical Programming* 28 (1984) 72–83.
- [6] J. Dupačová, "The minimax approach to stochastic programming and an illustrative application", to appear in *Stochastics*.
- [7] A.V. Fiacco, "Sensitivity analysis for nonlinear programming using penalty methods", *Mathematical Programming* 10 (1976) 287–311.
- [8] H. Gfrerer, J. Guddat and Hj. Wacker, "A globally convergent algorithm based on imbedding and parametric optimization", *Computing* 30 (1983) 225–252.
- [9] F.R. Hampel, "The influence curve and its role in robust estimation", *Journal of the American Statistical Association* 69 (1974) 383–397.
- [10] P. Kall and D. Stoyan, "Solving stochastic programming problems with recourse including error bounds", *Mathematische Operationsforschung und Statistik, Ser. Optimization* 13 (1982) 431–447.
- [11] R. Nadeau and R. Theodorescu, "Restricted Bayes strategies for programs with simple recourse", *Operations Research* 28 (1980) 777–784.

- [12] S.M. Robinson, "Perturbed Kuhn-Tucker points and rates of convergence for a class of nonlinear programming algorithms", *Mathematical Programming* 7 (1984) 1-16.
- [13] S.M. Robinson, "Strongly regular generalized equations", *Mathematics of Operations Research* 5 (1980) 43-62.
- [14] H. Schneeweiss, *Entscheidungskriterien bei Risiko* (Springer-Verlag, Berlin, 1967).
- [15] J. Wang, "Distribution sensitivity analysis for stochastic programs with recourse", to appear in *Mathematical Programming*.