Journal of Mathematical Economics 2 (1975) 349–370. © North-Holland Publishing Company

STOCHASTIC CONVEX PROGRAMMING: KUHN–TUCKER CONDITIONS

R. T. ROCKAFELLAR*

University of Washington, Seattle, Wash. 98195, U.S.A.
University of Grenoble, Grenoble, France

R. J.-B. WETS†

University of Kentucky, Lexington, Ky. 40506, U.S.A.

Received September 1974

Optimality criteria are derived for stochastic programs with convex objective and convex constraints. The problem consists in selecting $x_1 \in R^{n_1}$ and $x_2 \in \mathscr{L}^\infty(S, \Sigma, \sigma; R^{n_2})$ so as to satisfy the constraints and minimize total expected cost, where σ is a probability measure. The (basic) Kuhn–Tucker conditions are obtained in terms of conditions on the existence of saddle points of a Lagrangian associated with the stochastic program. We also give an interpretation of these results in terms of equilibrium theory with particular emphasis on a non-standard price system associated with the restriction that the (first stage) decision x_1 must be chosen independent of the random elements of the problem.

1. Introduction

In this paper we obtain Kuhn–Tucker conditions for the following stochastic program: find $x_1 \in R^{n_1}$ and $x_2 \in \mathscr{L}^\infty_{n_2} = \mathscr{L}^\infty(S, \Sigma, \sigma; R^{n_2})$ satisfying

$$x_1 \in C_1 \quad \text{and} \quad f_{1i}(x_1) \leqq 0, \qquad \text{for} \quad i = 1, \ldots, m_1, \tag{1}$$

satisfying almost surely (a.s.)

$$x_2(s) \in C_2 \quad \text{and} \quad f_{2i}(s, x_1, x_2(s)) \leqq 0, \qquad \text{for} \quad i = 1, \ldots, m_2, \tag{2}$$

and minimizing subject to these conditions

$$f_{10}(x_1) + \textstyle\int_S f_{20}(s, x_1, x_2(s))\sigma(\mathrm{d}s). \tag{3}$$

Here (S, Σ, σ) is a probability space. The sets $C_1 \subset R^{n_1}$ and $C_2 \subset R^{n_2}$ are closed, convex and non-empty. The functions $f_{1i}(\cdot)$ for $i = 0, 1, \ldots, m_1$, and

*Supported in part by the Air Force Office of Scientific Research under grant AF-AFOSR-72-2269.

†Supported in part by the National Science Foundation under grant GP-31551.

the functions $f_{2i}(s, \cdot, \cdot)$ for $i = 0, 1, \ldots, m_2$, are finite and convex on R^{n_1} and $R^{n_1} \times R^{n_2}$, respectively. The space $\mathscr{L}_n^p$ denotes the usual Lebesgue space of R^n-valued functions over (S, Σ, σ) with $\mathscr{L}_n^\infty$ the class of measurable essentially bounded functions. Moreover the functions $f_{2i}(\cdot, x_1, x_2)$ are measurable on S, summable for $i = 0$ and bounded for $i = 1, \ldots, m_2$. This implies that for every $x_1 \in R^{n_1}$ and $x_2 \in \mathscr{L}_{n_2}^\infty$, the functions $f_{2i}(s, x_1, x_2(s))$ are measurable in s, summable for $i = 0$ and essentially bounded for $i = 1, \ldots, m_2$. We derived duality results for this problem in Rockafellar–Wets (forthcoming – a).

Problem (1), . . ., (3) is a 'static' version of the more common formulation 'á la dynamic programming' for stochastic programs with recourse: find $x_1 \in R^{n_1}$ such that

$$x_1 \in C_1 \quad \text{and} \quad f_{1i}(x_1) \leqq 0, \qquad \text{for} \quad i = 1, \ldots, m_2, \tag{4}$$

and minimizing

$$f_{10}(x_1) + \int_S q(s, x_1)\sigma(\mathrm{d}s), \tag{5}$$

where

$$q(s, x_1) = \inf_{x_2 \in C_2} \{f_{20}(s, x_1, x_2) \mid f_{2i}(s, x_1, x_2) \leqq 0 \quad \text{for} \quad i = 1, \ldots, m_2\}. \tag{6}$$

The basic difference between problems (1), . . ., (3) and (4), . . ., (6) is that in the latter one it is made explicit that the *recourse decision* x_2 will be selected only *after* the random elements of the problem have been observed whereas the former problem requires the selection at the outset of a *recourse function* $s \mapsto x_2(s)$ which will specify all decisions for every possible value s. In the literature devoted to decision-making under risk such functions are also called decision rules [Blackwell–Girshick (1954), Gartska–Wets (1974)], policies [Blackwell (1962), Dynkin (1974)], feedback control functions or simply stochastic controls [Fleming–Rishel (1975)].

The embedding of a dynamic problem into a 'static' version by an extension of the decision space is a useful procedure when seeking necessary and/or sufficient conditions for optimality, see for example Kushner (1972). The restriction of x_2 to $\mathscr{L}_{n_2}^\infty$ in (1), . . ., (3) does however create some question as to the equivalence of the two problems. It is shown in Rockafellar–Wets (forthcoming – a) that under a rather weak regularity condition (certainly satisfied if C_2 is bounded) both problems are indeed equivalent. In particular x_1^0 solves (4), . . ., (6) if and only if there exists a sequence of feasible solutions $\{(x_1^0, x_2^k)\}$ with

$$f_{10}(x_1^0) + \int_S f_{20}(s, x_1^0, x_2^k(s))\sigma(\mathrm{d}s),$$

converging to the infimum of (1), . . ., (3). In particular, we have that

$$\lim_{k \to +\infty} \int_S f_{20}(s, x_1^0, x_2^k(s))\sigma(\mathrm{d}s) = \int_S q(s, x_1^0)\sigma(\mathrm{d}s).$$

It is possible, as we shall see, to associate with the original problem (1), . . ., (3) a *Lagrangian* L on $X \times Y$ defined by

$$\begin{aligned} L(x, y) &= L_1(x_1, y_1) + \int_S L_2(s, x_1, x_2(s), y_2(s))\sigma(\mathrm{d}s), \\ & \qquad\qquad\qquad\qquad \text{if} \quad x \in X_0, \quad y \in Y_0, \\ &= -\infty, \qquad\qquad\quad\; \text{if} \quad x \in X_0, \quad y \notin Y_0, \\ &= +\infty, \qquad\qquad\quad\; \text{if} \quad x \notin X_0, \end{aligned} \tag{7}$$

where

$$\begin{aligned} &X = R^{n_1} \times \mathscr{L}^{\infty}_{n_2}, \qquad Y = R^{m_1} \times \mathscr{L}^{1}_{m_2}, \\ &X_0 = \{x \in X \mid x_1 \in C_1, \quad x_2(s) \in C_2 \quad \text{a.s.}\}, \\ &Y_0 = \{y \in Y \mid y_1 \geqq 0, \quad y_2(s) \geqq 0 \quad \text{a.s.}\}, \end{aligned}$$

and the functions L_1, L_2 are given by

$$L_1(x_1, y_1) = f_{10}(x_1) + \sum_{1}^{m_1} y_{1i} f_{1i}(x_1), \tag{8}$$

and

$$L_2(s, x_1, x_2, y_2) = f_{20}(s, x_1, x_2) + \sum_{1}^{m_2} y_{2i} f_{2i}(s, x_1, x_2). \tag{9}$$

Of interest are the saddle points of L. A pair $(\bar{x}, \bar{y})$ is a saddle point of L with respect to minimization in x and maximization in y, if for any point $(x, y) \in X \times Y$ we have that

$$L(\bar{x}, y) \leqq L(\bar{x}, \bar{y}) \leqq L(x, \bar{y}). \tag{10}$$

From the general theory of convex optimization [Rockafellar (1974a)], as shown in section 2, we know that the saddle points of L characterize pairs of optimal solutions to (1), . . ., (3) and an associated dual problem which can be interpreted in terms of a price system relative to perturbation of the constraints (2) and (3).

It will be shown that the following conditions are necessary and sufficient for a pair $(\bar{x}, \bar{y})$ to be a saddle point of L on $X \times Y$:

Basic Kuhn–Tucker conditions

(a) $\bar{x}_1 \in C_1$, *and for* $i = 1, \ldots, m_1$, *one has* $\bar{y}_{1i} \geqq 0$, $f_{1i}(\bar{x}_1) \leqq 0$ *and* $\bar{y}_{1i} f_{1i}(\bar{x}_1) = 0$;

(b) *for almost all* s, $\bar{x}_2(s) \in C_2$, *and for* $i = 1, \ldots, m_2$, *one has* $\bar{y}_{2i}(s) \geqq 0$, $f_{2i}(s, \bar{x}_1, \bar{x}_2(s)) \leqq 0$ *and* $\bar{y}_{2i}(s) f_{2i}(s, \bar{x}_1, \bar{x}_2(s)) = 0$; *moreover there exists a function* $\rho \in \mathscr{L}^1_{n_1} = \mathscr{L}^1((S, \Sigma, \sigma); R^{n_1})$ *such that*

(c) *the expression*

$$L_1+x_1\cdot\int\rho = f_{10}(x_1)+\sum_1^{m_1}\bar{y}_{1i}f_{1i}(x_1)+x_1\cdot(\int\rho(s)\sigma\,(\mathrm{d}s))$$

attains its minimum in x_1 *over the set* C_1 *at* $\bar{x}_1$;

(d) *for almost all s, the expression*

$$L_2-x_1\rho = f_{20}(s, x_1, x_2)+\sum_1^{m_2}\bar{y}_{2i}(s)f_{2i}(s, x_1, x_2)-x_1\cdot\rho(s)$$

attains its minimum in (x_1, x_2) *over the set* $R^{n_1}\times C_2$ *at* $(\bar{x}_1, \bar{x}_2(s))$.

Whenever the optimal solutions x to (1), . . ., (3) can be characterized in terms of these conditions, it opens up the possibility of developing dual methods of computation that take advantage of a 'decomposition' into independent finite dimensional problems of minimization.

Of course, the characterization of optimal solutions in terms of 'Kuhn–Tucker conditions' also has theoretical significance for mathematical economics. Up to now, no such general characterization has been derived, even for the case of stochastic *linear* programming. [The 'weaker' duality results of Wets (1970) for the linear model are given an economic perspective in Gartska (1973).] In section 4, we give an interpretation of the multipliers $(\bar{y}_1, \bar{y}_2)$ appearing in the basic Kuhn–Tucker conditions as 'equilibrium prices' to be associated with the constraints (2) and (3), respectively. We then proceed to seek the interpretation to give the 'new' multiplier ρ and it is shown that it corresponds to a price to be paid for *non-anticipativity*, which simply means that x_1 must be selected before the random events are observed. In the original problem (1), . . ., (3) this is reflected by the fact that x_1 does not depend on s. To our knowledge, this is the first time that multipliers associated with the non-anticipativity restrictions are introduced *explicitly* in the optimality conditions for stochastic optimization problems. These multipliers play a fundamental role in the theory of stochastic optimization. This becomes even more apparent when studying N-stage problems [Rockafellar (1974b)] or when any attempt is made to fit the theory associated with stochastic optimization problems in an abstract framework [Wets (1974)].

There is an 'extended' form of the basic Kuhn–Tucker conditions appearing in Rockafellar–Wets (forthcoming – b) resulting from an extended Lagrangian, also associated with the problem at hand, but including singular functionals as dual variables. In another paper (forthcoming – c) we show that these singular multipliers essentially correspond to *induced constraints* on x_1. These induced constraints are generated by the implicit restriction: x_1 must be selected in such a way that there exists $x_2\in\mathscr{L}^\infty_{n_2}$ with $x_2(s)\in C_2$ and $f_{2i}(s, x_1, x_2(s))\leq 0$ for $i = 1, \ldots, m_2$, for almost all s. Finally if these induced constraints are introduced explicitly in the problem, it is shown that these basic Kuhn–Tucker conditions provide optimality criteria for *all* two-stage stochastic convex programs [Rockafellar–Wets (forthcoming – c, section 4 and Theorem 1)].

2. The Lagrangian

We developed a duality theory (forthcoming – a) for (1), . . ., (3) by embedding the problem in a class of 'perturbed' problems. Let

$$U = R^{m_1} \times \mathscr{L}^{\infty}_{m_2}$$

be the perturbation space. This is the natural choice for perturbation of the constraints, since from our assumptions (forthcoming – a, section 1) it follows that the functions

$$x_1 \mapsto f_{1i}(x_1) \quad \text{and} \quad (x_1, x_2) \mapsto f_{2i}(\cdot, x_1, x_2(\cdot))$$

have their values in R and $\mathscr{L}^{\infty}$, respectively. The function

$$F: X \times U \to]-\infty, +\infty]$$

is defined as follows. If $x = (x_1, x_2) \in X$ and $u = (u_1, u_2) \in U$ satisfy

$$x_1 \in C_1 \quad \text{and} \quad f_{1i}(x_1) \leqq u_{1i}, \qquad \text{for} \quad i = 1, \ldots, m_1, \tag{11}$$

and almost surely

$$x_2(s) \in C_2 \quad \text{and} \quad f_{2i}(s, x_1, x_2(s)) \leqq u_{2i}, \qquad \text{for} \quad i = 1, \ldots, m_2, \tag{12}$$

then $F(x, u)$ is the expected cost (6):

$$F(x, u) = f_{10}(x_1) + \int_S f_{20}(s, x_1, x_2(s))\sigma(\mathrm{d}s); \tag{13}$$

otherwise $F(x, u) = +\infty$. The original problem is thus equivalent to finding x in X which minimizes $F(x, 0)$.

Depending on how a space Y is paired with the space U, this system of perturbations will lead to different forms of the Lagrangian [Rockafellar (1974a)]. As before (forthcoming – a) we take here

$$\langle u, y \rangle = u_1 \cdot y_1 + \int_S u_2(s) \cdot y_2(s)\sigma(\mathrm{d}s), \qquad \text{for} \quad y \in Y = R^{m_1} \times \mathscr{L}^1_{m_2}. \tag{14}$$

Then according to the general theory [Rockafellar (1974a)], the Lagrangian associated with this system of perturbation is the function

$$L: X \times Y \to [-\infty, +\infty]$$

defined by (7). To show this it suffices to establish that L and F are conjugates in a sense made precise by the following proposition, whose proof was sketched out in Rockafellar–Wets (forthcoming – a).

Proposition 1. The following formulas are valid:

$$L(x, y) = \inf_{u \in U} \{F(x, u) + \langle u, y \rangle\} \quad \textit{for all} \quad (x, y) \in X \times Y; \tag{15}$$

$$F(x, u) = \sup_{y \in Y} \{L(x, y) - \langle u, y \rangle\} \quad \textit{for all} \quad (x, u) \in X \times U. \tag{16}$$

Proof. The formulas hold trivially if $x \notin X_0$, so henceforth we consider only a fixed x in X_0. The infimum in (15) is by definition the infimum in u of the linear expression

$$u_1 \cdot y_1 + \int_S u_2(s) y_2(s) \sigma(\mathrm{d}s) + K(x_1, x_2),$$

for all $u_1 \in R^{m_1}$ and $u_2 \in \mathscr{L}^{\infty}_{m_2}$ satisfying

$$u_{1i} \geqq f_{1i}(x_1), \qquad \text{for} \quad i = 1, \ldots, m_1, \tag{17}$$

and almost surely

$$u_{2i}(s) \geqq f_{2i}(s, x_1, x_2(s)), \qquad \text{for} \quad i = 1, \ldots, m_2; \tag{18}$$

and where K is a constant depending on the fixed x in X_0:

$$K(x_1, x_2) = f_{10}(x_1) + \int_S f_{20}(s, x_1, x_2(s)) \sigma(\mathrm{d}s).$$

If $y \in Y_0$ it is easily verified that this infimum is

$$L_1(x_1, y_1) + \int_S L_2(s, x_1, x_2(s)) \sigma(\mathrm{d}s),$$

with L_1 and L_2 given by (8) and (9), whereas if $y \notin Y_0$, the infimum is $-\infty$. This establishes (15). On the other hand, for $x \in X_0$ the supremum in (16) is by definition the supremum in y of the linear form:

$$K(x_1, x_2) + \sum_1^{m_1} y_{1i}(f_{1i}(x_1) - u_{1i})$$
$$+ \int_S \sum_1^{m_2} y_{2i}(s)(f_{2i}(s, x_1, x_2(s)) - u_{2i}(s)) \sigma(\mathrm{d}s),$$

for all $y_1 \in R^{m_1}$ and $y_2 \in \mathscr{L}^{\infty}_{m_2}$, satisfying

$$y_{1i} \geqq 0, \qquad \text{for} \quad i = 1, \ldots, m_1,$$

and almost surely

$$y_{2i}(s) \geqq 0, \quad \text{for} \quad i = 1, \ldots, m_2.$$

If $y \notin Y_0$ then $L(x, y) = -\infty$ and consequently $L(x, y) - \langle x, y \rangle = -\infty$ which cannot be a supremum, since one verifies easily that if $y \in Y_0$ the supremum is $+\infty$ if (u_1, u_2) does not satisfy conditions (17) and almost surely (18); otherwise it is equal to $K(x_1, x_2)$. This is exactly the description of $F(x, u)$ given by (11), ..., (13).

One observes that formula (16) can be interpreted as saying that the function $y \mapsto -L(x, -y)$ is the (convex) conjugate on Y of the function $u \mapsto F(x, u)$ on U. In view of the fact that for any $x \in X_0$ the above functions are not identically $+\infty$ it follows directly from Proposition 1 that if $x \in X_0$, the functional $u \mapsto F(x, u)$ is weakly lower semicontinuous, convex and not identically $+\infty$ and the functional $y \mapsto L(x, y)$ is weakly upper semicontinuous, concave and not identically $-\infty$. In fact we (forthcoming – a, Proposition 3) have established a stronger assertion which we reproduce here.

Corollary. The functional $(x, u) \to F(x, u)$ is lower semicontinuous, convex and not identically $+\infty$ on $X \times U$. The lower semicontinuity is not only with respect to the normable topology but also with respect to the weak topology on $X \times U$ induced by the pairing

$$\langle x, v \rangle + \langle y, u \rangle = x_1 \cdot v_1 + \int_S x_2(s) \cdot v_2(s)\sigma(\mathrm{d}s) + y_1 \cdot u_1 + \int_S y_2(s) \cdot u_2(s)\sigma(\mathrm{d}s),$$

with

$$v = (v_1, v_2) \in V = R^{n_1} \times \mathscr{L}^1_{n2}.$$

The dual problem resulting from this Lagrangian can be expressed as:

P Find $x \in X$ such that $f(x)$ is minimized

where

$$f(x) = \sup_{y \in Y} L(x, y), \tag{19}$$

and a corresponding dual

D Find $y \in Y$ such that $g(x)$ is maximized

where

$$g(x) = \inf_{x \in X} L(x, y). \tag{20}$$

It is easy to see that **P** is equivalent to the original problem (1), ..., (3) since from (16) it follows that $f(x) = F(x, 0)$ and finding x minimizing $F(x, 0)$ is exactly the original problem (in its unperturbed form).

Let $(\bar{x}, \bar{y})$ be a saddle point of the Lagrangian L on $X \times Y$ (or equivalently on $X_0 \times Y_0$) with respect to minimization of in x and maximization in y. Weak duality between **P** and **D** results directly from the obvious inequality

$$g(y) \leqq f(x), \qquad \text{for all} \quad (x, y) \in X \times Y. \tag{21}$$

On the other hand for $(\bar{x}, \bar{y})$ we have that

$$g(\bar{y}) = L(\bar{x}, \bar{y}) = f(\bar{x}).$$

Thus for $\bar{x}$ to solve **P** and $\bar{y}$ to solve **D** it is sufficient that $(\bar{x}, \bar{y})$ be a saddle point of L. This condition becomes necessary whenever one can also show that

$$\inf \mathbf{P} = \sup \mathbf{D}.$$

This was essentially the burden of Rockafellar–Wets (forthcoming – a, Theorem 3 and its corollaries) when C_1 and C_2 are bounded, and of Rockafellar–Wets (forthcoming – c) where it is shown that, once the so called *induced constraints* have been taken care of, the above equality is in fact valid under a strict feasibility assumption.

Characterization of the saddle points of L will thus yield certain *necessary* and *sufficient* conditions to be satisfied by optimal solutions to **P** and **D**. We already obtained partial results in this direction (forthcoming – a, Theorem 4) in terms of the directional derivatives of the *perturbation function*

$$\phi(u) = \inf_{x \in X} F(x, u).$$

3. The Kuhn–Tucker conditions

We show that the basic Kuhn–Tucker conditions in fact characterize the saddle points (x, y) of L. Note that the saddle point property of $(\bar{x}, \bar{y})$ can be expressed as

$$\max_{y \in Y} L(\bar{x}, y) = L(\bar{x}, \bar{y}) = \min_{x \in X} L(x, \bar{y}),$$

or equivalently, in view of the definition of L, as

$$\max_{y \in Y_0} L(\bar{x}, y) = L(\bar{x}, \bar{y}) = \min_{x \in X_0} L(x, \bar{y}), \qquad \text{for} \quad (\bar{x}, \bar{y}) \in X_0 \times Y_0. \tag{22}$$

Theorem. A pair $(\bar{x}, \bar{y}) \in X \times Y$ is a saddle point of the Lagrangian L if and only if the basic Kuhn–Tucker conditions are satisfied.

Proof. Since a saddle point $(\bar{x}, \bar{y})$ must belong to $X_0 \times Y_0$ it follows that it must satisfy the first two conditions in (a) and first two conditions in (b) of the basic Kuhn–Tucker conditions,

$$\bar{x}_1 \in C_1, \qquad \bar{y}_{1i} \geqq 0, \qquad i = 1, \ldots, m_1,$$

and

$$\bar{x}_2(s) \in C_2 \text{ a.s.}, \qquad \bar{y}_{2i} \geqq 0 \text{ a.s.}, \qquad i = 1, \ldots, m_2.$$

On the other hand for any $\bar{x} \in X_0$, we have that

$$\begin{aligned} \sup_{y \in Y_0} L(\bar{x}, y) = f_{10}(\bar{x}_1) &+ \sum_1^{m_1} \sup_{0 \leqq y_{1i} \in R} y_{1i} f_{1i}(\bar{x}) \\ &+ \int_S f_{20}(s, \bar{x}_1, \bar{x}_2(s)) \sigma(\mathrm{d}s) \\ &+ \sum_1^{m_2} \sup_{0 \leqq y_{2i} \in \mathscr{L}_1{}^1} \int_S y_{2i}(s) f_{2i}(s, \bar{x}_1, \bar{x}_2(s)) \sigma(\mathrm{d}s). \end{aligned} \tag{23}$$

This supremum is $+\infty$ unless $f_{1i}(\bar{x}_1) \leqq 0$ for $i = 1, \ldots, m_1$, and almost surely $f_{2i}(s, \bar{x}_1, \bar{x}_2(s)) \leqq 0$ for $i = 1, \ldots, m_2$. For if for some l, $f_{2l}(s, \bar{x}_1, \bar{x}_2(s)) > 0$ on a set of positive σ-measure, the sequence

$$\{y^k \mid y^k_{1i} = 0 \quad \text{for all } i, \quad y^k_{2i} = 0 \quad \text{for all} \quad i \neq l, \quad y^k_{2l}(s) = k$$
$$\text{if} \quad f_{2l}(s, \bar{x}_1, \bar{x}_2(s)) > 0, \quad y^k_{2l} = 0 \quad \text{otherwise}\}$$

readily shows that the sup must be $+\infty$; and similarly if $f_{1l}(\bar{x}_1) > 0$ for some l. Thus whenever the supremum is finite, for each i we must have that

$$\bar{y}_{1i} f_{1i}(\bar{x}_1) \leqq 0 \qquad \text{and} \qquad \bar{y}_{2i}(s) f_{2i}(s, \bar{x}_1, \bar{x}_2(s)) \leqq 0 \qquad \text{a.s.}$$

The second inequality follows from the fact that $f_{2i}(s, \bar{x}_1, \bar{x}_2(s)) \leqq 0$ a.s., that $\bar{y} \in Y_0$ implies $\bar{y}_{2i}(s) \geqq 0$ a.s. and that the union of two sets of measure zero is a set of measure zero. Thus the supremum in (23) is actually attained if and only if one has the equations

$$\bar{y}_{1i} f_{1i}(\bar{x}_1) = 0, \qquad \text{for} \quad i = 1, \ldots, m_1,$$

and almost surely

$$\bar{y}_{2i}(s) f_{2i}(s, \bar{x}_1, \bar{x}_2(s)) = 0, \qquad \text{for} \quad i = 1, \ldots, m_2,$$

which are the third conditions of (a) and (b) in the basic Kuhn–Tucker conditions.

Therefore (a) and (b) are equivalent to

$$L(\bar{x}, \bar{y}) = \max_{y \in Y_0} L(\bar{x}, y).$$

To establish the theorem, it remains to show that

$$L(\bar{x}, \bar{y}) = \min_{x \in X_0} L(x, \bar{y}) \tag{24}$$

is equivalent to the existence of a vector valued function $\rho \in \mathscr{L}^1_{n_1}$ such that (c) and (d) hold. Certainly if such a ρ does exist we have that for every $x \in X_0$,

$$\begin{aligned} L(x, \bar{y}) &= L_1(x_1, \bar{y}_1) + x_1 \cdot \int_S \rho(s)\sigma(\mathrm{d}s) \\ &\quad + \int_S [L_2(s, x_1, x_2(s), \bar{y}_2(s)) - x_1 \cdot \rho(s)]\sigma(\mathrm{d}s) \\ &\geqq L_1(\bar{x}_1, \bar{y}_1) + \bar{x}_1 \cdot \int_S \rho(s)\sigma(\mathrm{d}s) \\ &\quad + \int_S [L_2(s, \bar{x}_1, \bar{x}_2(s), \bar{y}_2(s)) - \bar{x}_1 \cdot \rho(s)]\sigma(\mathrm{d}s) \\ &= L(\bar{x}, \bar{y}), \end{aligned}$$

where the inequality results from (c) and (d) since $\bar{x}_1$ minimizes

$$L_1(x_1, \bar{y}_1) + x_1 \int_S \rho(s)\sigma(\mathrm{d}s)$$

on C_1, and $(\bar{x}_1, \bar{x}_2) \in R^{n_1} \times \mathscr{L}^\infty_{n_2}$ minimizes almost surely

$$L_2(s, x_1, x_2(s), \bar{y}_2(s)) - x_1 \cdot \rho(s)$$

on $R^{n_1} \times C_2$. Hence, to complete the proof it suffices to show that the existence of ρ with (c) and (d) is also necessary for $\bar{x}$ to minimize L on X_0. For this purpose we utilize the theory of convex integral functionals and their subgradients [Rockafellar (1971)]. Nothing is harmed if we suppose (redefining $\bar{y}_2$ on a set of measure zero if necessary) that $\bar{y}_2(s) \geqq 0$ for *every* $s \in S$. Let $h_1 : R^{n_1} \to R$ and $h_2 : S \times R^{n_1} \times R^{n_2} \to R$ be defined by

$$h_1(x_1) = f_{10}(x_1) + \sum_1^{m_1} \bar{y}_{1i} f_{1i}(x_1),$$

and

$$h_2(s, x_1, x_2) = f_{20}(s, x_1, x_2) + \sum_1^{m_2} \bar{y}_{2i}(s) f_{2i}(s, x_1, x_2).$$

Since the $\bar{y}_{1i}$ and $\bar{y}_{2i}(s)$ are non-negative and the functions f_{1i} and f_{2i} are convex in x_1 and (x_1, x_2), respectively, the functions h_1 and $h_2(s, \cdot, \cdot)$ are convex (and everywhere finite). Furthermore, since $\bar{y}_2 \in \mathscr{L}^1_{m_2}$ we have that for each $(x_1, x_2) \in R^{n_1} \times R^{n_2}$ the function

$$s \mapsto h_2(s, x_1, x_2)$$

is summable. These properties of h_2 have the following consequences, as established in Rockafellar (1971, Corollary 2A): the expression $h_2(s, x_1(s), x_2(s))$ is summable whenever $x_1(s)$ and $x_2(s)$ are measurable and essentially bounded in s, the functional

$$H_2(x) = \int_S h_2(s, x_1(s), x_2(s))\sigma(\mathrm{d}s)$$

is thus well-defined and finite on $\mathscr{L}^\infty_{n_1} \times \mathscr{L}^\infty_{n_2}$, and of course convex; it is also continuous in the norm topology. Most important, the subgradient set

$$\partial H_2(x) \subset (\mathscr{L}^\infty_{n_1} \times \mathscr{L}^\infty_{n_2})^*$$

consists of the continuous linear functionals in $\mathscr{L}^\infty_{n_1} \times \mathscr{L}^\infty_{n_2}$ which can be identified with pairs $(q_1, q_2) \in \mathscr{L}^1_{n_1} \times \mathscr{L}^1_{n_2}$ [Rockafellar (1971, Corollary 2C)] satisfying

$$(q_1(s), q_2(s)) \in \partial h_2(s, x_1(s), x_2(s)) \quad \text{a.s.},$$

where $\partial H_2(s, x_1(s), x_2(s))$ is the set of subgradients $R^{n_1} \times R^{n_2}$ of $H_2(s, \cdot, \cdot)$ at $(x_1(s), x_2(s))$. Let

$$J: R^{n_1} \times \mathscr{L}^\infty_{n_2} \to \mathscr{L}^\infty_{n_1} \times \mathscr{L}^\infty_{n_2}$$

be the injection which maps each element of R^{n_1} onto the corresponding constant function in $\mathscr{L}^\infty_{n_1}$. Obviously J is a linear transformation which is continuous with respect to the norm topologies. We have that

$$L(x, \bar{y}) = h_1(x_1) + H_2(Jx), \quad \text{for} \quad x = (x_1, x_2) \in X_0.$$

Introducing the convex function

$$H_1(x) = \begin{cases} h_1(x_1), & \text{if } x \in X_0, \\ +\infty, & \text{if } x \notin X_0, \end{cases}$$

we can write (24) as

$$H_1(\bar{x}) + H_2(J\bar{x}) = \min_{x \in R^{n_1} \times \mathscr{L}^\infty{}_{n_2}} \{H_1(x) + H_2(Jx)\}.$$

This says that $\bar{x}$ is an optimal solution to a certain convex optimization problem of the Fenchel type, as studied in Rockafellar (1966), and since J and H_2 are continuous, it follows from Theorem 1 of that paper that there exists a linear functional $w \in (\mathscr{L}^\infty_{n1} \times \mathscr{L}^\infty_{n2})^*$ such that

$$-J^*w \in \partial H_1(\bar{x}), \tag{25}$$

and

$$w \in \partial H_2(J\bar{x}), \tag{26}$$

where

$$J^* : (\mathscr{L}^\infty_{n1} \times \mathscr{L}^\infty_{n2})^* \to (R^{n_1} \times \mathscr{L}^\infty_{n2})^*$$

is the adjoint of the linear map J.

Relation (26) means, according to properties of the subgradient of H_2 cited above, that w corresponds to a pair $(q_1, q_2) \in \mathscr{L}^1_{n1} \times \mathscr{L}^2_{n2}$ such that

$$(q_1(s), q_2(s)) \in \partial h_2(s, \bar{x}_1, \bar{x}_2(s)) \quad \text{a.s.} \tag{27}$$

Moreover J^*w is in this case the element of $(R^{n_1} \times \mathscr{L}^\infty_{n2})^*$ corresponding to the pair

$$(\textstyle\int_S q_1(s)\sigma(\mathrm{d}s), q_2)$$

in $R^{n_1} \times \mathscr{L}^1_{n2}$. Therefore the relation (25) can be written as

$$H_1(x) \geqq H_1(\bar{x}) - (x_1 - \bar{x}_1) \cdot \textstyle\int_S q_1(s)\sigma(\mathrm{d}s) - \int_S (x_2(s) - \bar{x}_2(s) \cdot q_2(s)\sigma(\mathrm{d}s)$$

for all $x = (x_1, x_2) \in R^{n_1} \times R^{n_2}$.

In other words, recalling the definition of H_1 and X_0, we have

$$h_1(x_1) \geqq h_1(\bar{x}_1) - (x_1 - \bar{x}_1)\textstyle\int_S q_1(s)\sigma(\mathrm{d}s), \qquad \text{for all} \quad x_1 \in C_1, \tag{28}$$

and

$$\textstyle\int_S (x_2(s) - \bar{x}_2(s)) \cdot q_2(s)\sigma(\mathrm{d}s) \geqq 0, \qquad \text{for all} \quad x_2 \in \mathscr{L}^\infty_2 \text{ such that } \quad x_2(s) \in C_2 \quad \text{a.s.} \tag{29}$$

We now claim that (29) implies

$$(x_2 - \bar{x}_2(s)) \cdot q_2(s) \geqq 0 \quad \text{a.s.}, \qquad \text{for all} \quad x_2 \in C_2. \tag{30}$$

To verify this, let D be a countable dense subset of C_2. If $x_2 \in D$, then the function

$$\tilde{x}_2(s) = \begin{cases} x_2, & \text{if} \quad (x_2 - \bar{x}_2(s)) \cdot q_2(s) < 0, \\ \bar{x}_2(s), & \text{if} \quad (x_2 - \bar{x}_2(s)) \cdot q_2(s) \geqq 0, \end{cases}$$

is an element of $\mathscr{L}^{\infty}_{n_1}$ with $\tilde{x}_2(s) \in C_2$ a.s. Hence from (29),

$$0 \leqq \int_S (\tilde{x}_2(s) - \bar{x}_2(s)) \cdot q_2(s)\sigma(\mathrm{d}s)$$
$$= \int_S \min\{0, (x_2 - \bar{x}_2(s)) \cdot q_2(s)\}\sigma(\mathrm{d}s),$$

and consequently

$$(x_2 - \bar{x}_2(s)) \cdot q_2(s) \geqq 0 \quad \text{a.s.}$$

Since this is true for each $x_2 \in D$ and D is countable, there exists a measurable set S' with $\sigma(S') = 1$, such that

$$(x_2 - \bar{x}_2(s)) \cdot q_2(s) \geqq 0, \qquad \text{for all} \quad x_2 \in D \quad \text{and all} \quad s \in S'.$$

Passing to the closure of D for each $s \in S'$, we obtain

$$(x_2 - \bar{x}_2(s)) \cdot q_2(s) \geqq 0, \qquad \text{for all} \quad x_2 \in C \quad \text{and all} \quad s \in S',$$

which is the desired property (30).

To summarize up to this point, we have demonstrated the existence of $(q_1, q_2) \in \mathscr{L}^1_{n_1} \times \mathscr{L}^1_{n_2}$ satisfying (27), (28) and (30). Of course (27) means also by the definition of subgradients that for every $(x_1, x_2) \in R^{n_1} \times R^{n_2}$,

$$h_2(s, x_1, x_2) \geqq h_2(s, \bar{x}_1, \bar{x}_2(s)) + (x_1 - \bar{x}_1) \cdot q_1(s)$$
$$+ (x_2 - \bar{x}_2(s)) \cdot q_2(s) \quad \text{a.s.}$$

The combination of this inequality with (30) yields that almost surely

$$h_2(s, x_1, x_2) - x_1 \cdot q_1(s) \geqq h_2(s, \bar{x}_1, \bar{x}_2(s)) - \bar{x}_1 \cdot q_1(s),$$

for every $(x_1, x_2) \in R^{n_1} \times C_2$. Let $\rho(s) = q_1(s)$. The latter inequality is then precisely condition (d) of the (basic) Kuhn–Tucker conditions. At the same time, we can rewrite (28) as

$$h_1(x_1) + x_1 \cdot \int_S \rho(s)\sigma(\mathrm{d}s) \geqq h_1(\bar{x}_1) + \bar{x}_1 \cdot \int_S \rho(s)\sigma(\mathrm{d}s),$$

for all $x_1 \in C_1$. This is condition (c) of the Kuhn–Tucker conditions, and the proof is therefore done.

There is naturally a subdifferential form of these Kuhn–Tucker conditions. By $\partial_z h(\bar{z}, \bar{v})$ we denote the set of subgradients of the function

$$z \mapsto h(z, \bar{v}),$$

at the point $\bar{z}$.

Basic Kuhn–Tucker conditions (subdifferential form)

(a) $\bar{x}_1 \in C_1$ *and for* $i = 1, \ldots, m_1$, *one has* $\bar{y}_{1i} \geqq 0$, $f_{1i}(\bar{x}_1) \leqq 0$ *and* $\bar{y}_{1i} f_{1i}(\bar{x}_1) = 0$;

(b) *for almost all* s, $\bar{x}_2(s) \in C_2$ *and for* $i = 1, \ldots, m_2$, *one has* $\bar{y}_{2i}(s) \geqq 0, f_{2i}(s, \bar{x}_1, \bar{x}_2(s)) \leqq 0$ *and* $\bar{y}_{2i}(s) f_{2i}(s, \bar{x}_1, \bar{x}_2(s)) = 0$; *moreover there is a function* $\rho \in \mathscr{L}^1_{n_1}$ *such that*

(c_∂) $v - \int \rho(s)\sigma(\mathrm{d}s) \in \partial f_{10}(\bar{x}_1) + \sum_1^{m_1} \bar{y}_{1i} \partial f_{1i}(\bar{x}_1)$, *where the vector* $-v$ *is normal to* C_1 *at* $\bar{x}_1$;

(d_∂) $(\rho(s), v(s)) \in \partial_{(x_1,x_2)} f_{20}(s, \bar{x}_1, \bar{x}_2(s)) + \sum_1^{m_2} \bar{y}_{2i}(s) \partial_{(x_1,x_2)} f_{2i}(s, \bar{x}_1, \bar{x}_2(s))$ *a.s., where* $-v(s)$ *is normal to* C_2 *at* $\bar{x}_2(s)$ *a.s.*

Corollary A. A pair $(\bar{x}, \bar{y}) \in X \times Y$ *is a saddle point of the Lagrangian L if and only if the subdifferential form of the basic Kuhn–Tucker conditions is satisfied.*

Proof. It suffices to verify that (c_∂) and (d_∂) are equivalent to (c) and (d). By ψ_C we denote the indicator functions of a set C which is defined by

$$\psi_C(x) = \begin{cases} 0, & \text{if } x \in C, \\ +\infty, & \text{otherwise.} \end{cases}$$

Now $\bar{x}_1$ minimizes $f_{10}(x_1) + \sum_1^{m_1} \bar{y}_{01} f_{1i}(x_1) + x_1 \cdot \int \rho(s)\sigma(\mathrm{d}s)$ on C_1 if and only if

$$0 \in \partial[f_{10}(\bar{x}_1) + \sum_1^{m_1} \bar{y}_{1i} f_{1i}(\bar{x}_1) + \bar{x}_1 \cdot \int \rho(s)\sigma(\mathrm{d}s) + \psi_{C_1}(\bar{x}_1)]. \tag{31}$$

For a function $g: R^N \to [-\infty, +\infty]$, its effective domain is by definition $\{x \mid g(x) < +\infty\}$. Obviously, the intersection of the relative interiors of the effective domains of the functions appearing in (31) is non-empty since by assumption the f_{1i}, $i = 0, 1, \ldots, m_1$, are finite on R^{n_1}. Thus, we can express the subdifferential in (31) as the sum of the subdifferentials of the individual functions [Rockafellar (1970, Theorem 23.8)], and (31) is equivalent to

$$0 \in \partial f_{10}(\bar{x}_1) + \sum_1^{m_1} \bar{y}_{1i} \partial f_{1i}(\bar{x}_1) + \int \rho(s)\sigma(\mathrm{d}s) + \partial\psi_{C_1}(\bar{x}_1),$$

where we have also used the fact that the $\bar{y}_{1i}$ are non-negative. The preceding relation yields directly (c_∂), if in addition we simply note that a vector $-v$ is a subgradient of ψ_{C_1} at $\bar{x}_1$ if and only if $-v$ is normal to C_1 at $\bar{x}_1$.

The equivalence of (d) and (d_∂) is derived in a similar manner.

Of practical interest are the special cases when the functions f_{1i} and f_{2i} are differentiable and the sets C_1, C_2 are either the positive orthant or the whole spaces. The basic Kuhn–Tucker conditions can then be specialized. This is done in the two corollaries below. By $\nabla_z h(\bar{z}, \bar{v})$ we denote the gradient of $h(z, v)$ with respect to z evaluated at $(\bar{z}, \bar{v})$.

Corollary B. Suppose that $C_1 = R^{n_1}_\oplus$ *is the positive orthant of* R^{n_1}, $C_2 = R^{n_2}_\oplus$ *is the positive orthant of* R^{n_2}, *the functions* $x_1 \mapsto f_{1i}(x_1)$, $i = 0, 1, \ldots, m_1$, *are differentiable on* R^{n_1} *and for each s the functions* $(x_1, x_2) \mapsto f_{2i}(s, x_1, x_2)$, $i = 0, 1, \ldots, m_2$, *are differentiable on* $R^{n_1} \times R^{n_2}$. *Then a pair* $(\bar{x}, \bar{y}) \in X \times Y$ *is a saddle point of the Lagrangian L if and only if* $(\bar{x}, \bar{y})$ *satisfies conditions (a) and (b) of the basic Kuhn–Tucker conditions and there exists a function* $\rho \in \mathscr{L}^1_{n_1}$ *such that*

($c_\oplus$) $\quad \nabla f_{10}(\bar{x}_1) + \sum_1^{m_1} \bar{y}_{1i} \nabla f_{1i}(\bar{x}_1) \geqq -\int_S \rho(s)\sigma(\mathrm{d}s)$ *and*
$\bar{x}_1 \cdot [\nabla f_{10}(\bar{x}_1) + \sum_1^{m_1} \bar{y}_{1i} \nabla f_{1i}(\bar{x}_1)] = -\bar{x}_1 \cdot \int_S \rho(s)\sigma(\mathrm{d}s)$;

($d_\oplus$) *for almost all s,*
$\nabla_{x_1} f_{20}(s, \bar{x}_1, \bar{x}_2(s)) + \sum_1^{m_2} \bar{y}_{2i}(s) \nabla_{x_1} f_{2i}(s, \bar{x}_1, \bar{x}_2(s)) = \rho(s)$,
$\nabla_{x_2} f_{20}(s, \bar{x}_1, \bar{x}_2(s)) + \sum_1^{m_2} \bar{y}_{2i}(s) \nabla_{x_2} f_{2i}(s, \bar{x}_1, \bar{x}_2(s)) \geqq 0$,
and
$\bar{x}_2(s) \cdot [\nabla_{x_2} f_{20}(s, \bar{x}_1, \bar{x}_2(s)) + \sum_1^{m_2} \bar{y}_{2i}(s) \nabla_{x_2} f_{21}(s, \bar{x}_1, \bar{x}_2(s))] = 0.$

Proof. Naturally it suffices to show that (c_∂) and (d_∂) in the basic Kuhn–Tucker conditions (subdifferential form) take on the special form ($c_\oplus$) and ($d_\oplus$) when the hypotheses of the theorem are satisfied. First observe that when a convex function f is differentiable, then at every point $\hat{x}$ of its domain R^N the subdifferential $\partial f(\hat{x})$ contains only the gradient $\nabla f(\hat{x})$ of f at $\hat{x}$. Moreover, $\hat{x} \in R^N_\oplus$ minimizes the bounded convex differentiable function f on $R^N_\oplus$ if and only if

$$\nabla f(\hat{x}) \geqq 0 \quad \text{and} \quad \hat{x} \cdot \nabla f(\hat{x}) = 0. \tag{32}$$

This follows from the two following observations: (i) $\hat{x}$ minimizes f on $R^N_\oplus$ if and only if

$$0 \in \nabla f(\hat{x}) + \partial \psi_{R_\oplus^N}(\hat{x}),$$

and (ii) $v \in \partial \psi_{R_\oplus^N}(\hat{x})$ if and only if it is a normal to $R^N_\oplus$ at $\hat{x}$, i.e., $v \leqq 0$ and $\hat{x} \cdot v = 0$. Straightforward application of conditions (32) to the expressions to be minimized in (c) and (d) of the basic Kuhn–Tucker conditions yield readily ($c_\oplus$) and ($d_\ominus$).

Corollary C. Suppose that $C_1 = R^{n_1}$, $C_2 = R^{n_2}$, *the functions* $x_1 \mapsto f_{1i}(x_1)$ *for* $i = 0, 1, \ldots, m_1$, *are differentiable on* R^{n_1} *and the functions* $(x_1, x_2) \mapsto f_{2i}(s, x_1, x_2)$ *for* $i = 0, 1, \ldots, m_2$, *are differentiable on* $R^{n_1} \times R^{n_2}$. *Then a pair* $(\bar{x}, \bar{y}) \in X \times Y$ *is a saddle point of the Lagrangian L if and only if in addition to conditions (a) and (b) of the basic Kuhn–Tucker conditions* $(\bar{x}, \bar{y})$ *one has the following: there exists a function* $\rho \in \mathscr{L}^1_{n_1}$ *such that*

(c_∇) $\quad \nabla f_{10}(\bar{x}_1) + \sum_1^{m_1} \bar{y}_{1i} f_{1i}(\bar{x}_1) = -\int \rho(s)\sigma(\mathrm{d}s);$

(d_∇) *for almost all s,*

$$\nabla_{x_1} f_{20}(s, \bar{x}_1, \bar{x}_2(s)) + \textstyle\sum_1^{m_2} \bar{y}_{2i}(s) \nabla_{x_1} f_{2i}(s, \bar{x}_1, \bar{x}_2(s)) = \rho(s),$$
$$\nabla_{x_2} f_{20}(s, \bar{x}_1, \bar{x}_2(s)) + \textstyle\sum_1^{m_2} \bar{y}_{2i}(s) \nabla_{x_2} f_{2i}(s, \bar{x}_1, \bar{x}_2(s)) = 0.$$

Proof. Again this corollary is a direct consequence of Corollary A, once we have observed that the 0 vector is the only normal to R^{n_1} (or R^{n_2}) and that the differentiability assumption implies that the subdifferentials are singletons.

The arguments used in the proof of the theorem can be adapted directly to the special case considered in Corollary C; this is done in Wets (1975, Theorem 2). This allows for significant simplifications in the proof since one can utilize well-known results of classical analysis rather than having to rely on the stronger results of convex analysis and the theory of normal convex integrands.

There is a formal resemblance between condition (d) of the basic Kuhn–Tucker conditions and the Pontryagin maximum principle for optimal control problems [Fleming–Rishel (1975)]. In particular, it asserts that if the optimal $\bar{x}_1$ is known, then an optimal recourse function $\bar{x}_2 \in \mathscr{L}^\infty_{n_2}$ must minimize 'pointwise' $L_2(s, \bar{x}_1, x_2, \bar{y}_2(s))$ for almost all s, i.e., for almost all s, $\bar{x}_2(s)$ is an optimal solution of the minimization problem: find x_2 minimizing

$$f_{20}(s, \bar{x}_1, x_2) + \sum_1^{m_2} \bar{y}_{2i}(s) f_{2i}(s, \bar{x}_1, x_2)$$

on C_2. Assuming that the original problem (1), ..., (3) is solvable, i.e., there actually exist $\bar{x}_1 \in R^{n_1}$, $\bar{x}_2 \in \mathscr{L}^\infty_{n_2}$ satisfying the constraints and for which the minimum is attained, and assuming moreover that the basic Kuhn–Tucker conditions are not only sufficient but also necessary, then, for almost all s, $\bar{x}_2(s)$ must minimize

$$f_{20}(s, \bar{x}_1, x_2)$$

subject to

$$x_2 \in C_2 \quad \text{and} \quad f_{2i}(s, \bar{x}_1, x_2) \leqq 0, \qquad \text{for} \quad i = 1, \ldots, m_2.$$

For fixed s, the optimal solution to this minimization problem must satisfy the appropriate Kuhn–Tucker conditions; see (33) and (34) below. That for almost all s, $(\bar{x}_2(s), \bar{y}_2(s))$ determines an optimal solution and a corresponding vector of multipliers to be associated with this problem, is the content of the following proposition.

Proposition 2. Suppose $(\bar{x}, \bar{y}) \in X \times Y$ is a saddle point for L. Then for almost all s, $(\bar{x}_2(s), \bar{y}_2(s)) \in R^{n_2} \times R^{m_2}$ is a saddle point of the Lagrangian l_2 of the recourse problem, with

$$\begin{aligned} l_2(s, x_2, y_2) &= f_{20}(s, \bar{x}_1, x_2) + \sum_{1}^{m_2} y_{2i} f_{2i}(s, \bar{x}_1, x_2), && \text{if} \quad x_2 \in C_2, \quad y_2 \geqq 0, \\ &= -\infty, && \text{if} \quad x_2 \in C_2, \quad y_2 \ngeqq 0, \\ &= +\infty, && \text{if} \quad x_2 \notin C_2. \end{aligned}$$

Proof. For each s, the recourse problem is a convex minimization problem of finite dimensions. We can thus apply standard results for convex programming problems [Rockafellar (1970, section 28)] to conclude that for almost all s, $(\bar{x}_2(s), \bar{y}_2(s)) \in R^{n_1} \times R^{n_2}$ is a saddle point of l_2 if and only if

$$\begin{aligned} &\bar{x}_2(s) \in C_2, \quad \bar{y}_2(s) \geqq 0, \qquad \text{and for} \quad i = 1, \ldots, m_2, \\ &f_{2i}(s, \bar{x}_1, \bar{x}_2(s)) \leqq 0, \quad \bar{y}_{2i}(s) f_{2i}(s, \bar{x}_1, \bar{x}_2(s)) = 0; \end{aligned} \tag{33}$$

$$\text{the expression } f_{20}(s, \bar{x}_1, x_2) + \sum_{1}^{m_2} \bar{y}_{2i}(s) f_{2i}(s, \bar{x}_1, x_2) \text{ attains its minimum in } x_2 \text{ over the set } C_2 \text{ at } \bar{x}_2(s). \tag{34}$$

But these are exactly conditions (b) and (d) of the basic Kuhn–Tucker conditions if it is known that $\bar{x}_1$ is the first component of an optimal solution for the minimization problem appearing in (d).

4. Equilibrium prices

We know that the 'multipliers' $\bar{y}$ appearing in the saddle point conditions for L can be interpreted as 'equilibrium prices' associated with the constraints of the problem. However, in addition to the standard price system $(\bar{y}_1, \bar{y}_2)$ associated with the constraints, we also have a non-standard 'price' vector $\rho \in \mathscr{L}^1_{n_1}$ which appears in conditions (c) and (d) of the basic Kuhn–Tucker conditions and, at first, does not seem to correspond to any specific constraint. Let us consider the following variant of problem (1), ..., (3):

find $x_1 \in R^{n_1}$, $\chi_1 \in \mathscr{L}^\infty_{n_1}$ and $x_2 \in \mathscr{L}^\infty_{n_2}$ satisfying

$$x_1 \in C_1 \quad \text{and} \quad f_{1i}(x_1) \leqq 0, \quad \text{for} \quad i = 1, \ldots, m_1, \tag{35}$$

satisfying almost surely

$$x_1 - \chi_1(s) = 0, \tag{36}$$

and

$$x_2(s) \in C_2 \quad \text{and} \quad f_{2i}(s, \chi_1(s), x_2(s)) \leqq 0, \quad \text{for} \quad i = 1, \ldots, m_2, \tag{37}$$

and minimizing

$$f_{10}(x_1) + \textstyle\int_S f_{20}(s, \chi_1(s), x_2(s))\sigma(\mathrm{d}s). \tag{38}$$

Naturally this latter problem and the original problem (1), . . ., (3) are equivalent in the sense that the infimum of (38) on the set of admissible solutions determined by (35)–(37) is equal to the infimum of (3) over the feasible region determined by (1) and (2). Moreover, to each pair $(x_1, x_2) \in R^{n_1} \times \mathscr{L}^\infty_{n_2}$ satisfying (1) and (2) corresponds in an obvious way, a family of triples $(x_1, \chi_1, x_2) \in R^{n_1} \times \mathscr{L}^\infty_{n_1} \times \mathscr{L}^\infty_{n_2}$ satisfying (35)–(37) with $\chi_1(s)$ almost surely constant, in fact $\chi_1(s) = x_1$ a.s. Such functions χ_1 are clearly members of $\mathscr{L}^\infty_{n_1}$.

The Lagrangian function $\mathbf{L}$ on $X \times \mathscr{L}^\infty_{n_1} \times Y \times \mathscr{L}^1_{n_1}$ generated by the minimization problem (35), . . ., (38) is

$$\begin{aligned}
\mathbf{L}(x, \chi_1, y, \rho) &= L_1(x_1, y_1) + \textstyle\int_S (x_1 - \chi_1(s)) \cdot \rho(s)\sigma(\mathrm{d}s) \\
&\quad + \textstyle\int_S L_2(s, \chi_1(s), x_2(s), y_2(s))\sigma(\mathrm{d}s) \\
&\qquad \text{if} \quad x \in X_0 \quad \text{and} \quad y \in Y_0, \\
&= -\infty, \qquad \text{if} \quad x \in X_0 \quad \text{and} \quad y \notin Y_0, \\
&= +\infty, \qquad \text{if} \quad x \notin X_0,
\end{aligned} \tag{39}$$

where X_0, Y_0, L_1 and L_2 are as defined by (7), (8) and (9).

The proof of the theorem in the previous section shows that one can also interpret the theorem as a statement about saddle points of $\mathbf{L}$.

Corollary D. The point $(\bar{x}, \chi_1, \bar{y}, \rho)$ is a saddle point of $\mathbf{L}$ on $X \times \mathscr{L}^\infty_{n_1} \times Y \times \mathscr{L}^1_{n_1}$ if and only if it satisfies the following modified form of the basic Kuhn–Tucker conditions:

(a) *$\bar{x}_1 \in C_1$ and for $i = 1, \ldots, m_1$, one has $\bar{y}_{1i} \geqq 0, f_{1i}(\bar{x}_1) \leqq 0$ and $\bar{y}_{1i} f_{1i}(\bar{x}_1) = 0$;*

(b′) *for almost all* s, $\bar{x}_2(s) \in C_2$ *and for* $i = 1, \ldots, m_2$, *one has* $\bar{y}_{2i}(s) \geqq 0$, $f_{2i}(s, \bar{\chi}_1(s), \bar{x}_2(s)) \leqq 0$ *and* $\bar{y}_{2i}(s) f_{2i}(s, \bar{\chi}_1(s), \bar{x}_2(s)) = 0$;

(c) *the expression*

$$f_{10}(x_1) + \sum_1^{m_1} \bar{y}_{1i} f_{1i}(x_1) + x_1 \cdot \int \rho(s)\sigma(\mathrm{d}s)$$

attains its minimum in x_1 *over the set* C_1 *at* $\bar{x}_1$;

(d′) *for almost all* s, *the expression*

$$f_{20}(s, \chi_1, x_2) + \sum_1^{m_2} \bar{y}_{2i}(s) f_{2i}(s, \chi_1, x_2) - \chi_1 \cdot \rho(s)$$

attains its minimum in (χ_1, x_2) *over the set* $R^{n_1} \times C_2$ *at* $(\bar{\chi}_1(s), \bar{x}_2(s))$.

(e′) *for almost all* s, $\bar{\chi}_1(s) = \bar{x}_1$.

The new condition (e′) shows that for saddle points, the term

$$\int_S (\bar{x}_1 - \bar{\chi}_1(s)) \cdot \rho(s)\sigma(\mathrm{d}s),$$

distinguishing **L** from L, must be equal to zero. In fact, it is easy to verify that $(\bar{x}, \bar{y})$ is a saddle point for L on $X \times Y$ if and only if there exists $\bar{\chi}_1 \in \mathscr{L}^\infty_{n_1}$ and $\rho \in \mathscr{L}^1_{n_1}$ such that $(\bar{x}, \bar{\chi}_1, \bar{y}, \rho)$ is a saddle point for **L** on $X \times \mathscr{L}^\infty_{n_1} \times Y \times \mathscr{L}^1_{n_1}$. Moreover we have that for saddle points $(\bar{x}, \bar{y})$ and $(\bar{x}, \bar{\chi}_1, \bar{y}, \rho)$,

$$L(\bar{x}, \bar{y}) = \mathbf{L}(\bar{x}, \bar{\chi}_1, \bar{y}, \rho).$$

From this it follows, among other things, that the saddle values of L and **L** are equal.

The $(\bar{y}, \rho)$ components of a saddle point of the modified Lagrangian **L** can heuristically be interpreted as 'equilibrium prices'. The interpretation to attach to the $\bar{y}$ variables follows from the standard economic analysis which yields 'equilibrium prices' for problems modelling allocation of scarce resources. This is however not so for the variable ρ.

The perturbation function

$$\phi : U \to [-\infty, +\infty]$$

as defined by (21) yields for every possible perturbation of the constraints, $f_{1i} \leqq 0$, $i = 1, \ldots, m_1$, and $f_{2i} \leqq 0$ a.1., $i = 1, \ldots, m_2$, the corresponding value of the infimum of the program. If these perturbations became available at a price, it might become advantageous to modify the constraints by 'buying' some of the available perturbations. A set of prices is an *equilibrium price* system if at these prices no perturbation is worth buying. If equilibrium prices y^* exist (elements of the space Y paired with U) they must satisfy the relation

$$\phi(u) + \langle u, y^* \rangle \geqq \phi(0) \qquad \text{for all} \quad u,$$

where $\langle u, y^* \rangle$ represents the additional cost of purchasing perturbations u at prices y^*. From the definition of the perturbation function ϕ and Proposition 1, it follows that the inequality above can also be expressed as

$$\inf_x L(x, y^*) \geqq \inf \mathbf{P},$$

observing that $\phi(0)$ is the optimal value of the program $\mathbf{P}$. This infimum is $-\infty$ unless $y^* \geqq 0$. Explicitly, this means that

$$y_1^* \geqq 0, \qquad \text{and for almost all } s, \quad y_2^*(s) \geqq 0.$$

We can interpret a saddle point of L, if one exists, as a pair $(\bar{x}, \bar{y})$ determining an optimal solution $\bar{x}$ to $\mathbf{P}$ and an 'equilibrium price system' $\bar{y}$. The main result of this paper implies that

$$\inf_x L(x, \bar{y}) = L(\bar{x}, \bar{y}) = \inf \mathbf{P},$$

with $\bar{y} \geqq 0$. The price vector $\bar{y}$ decomposes into two parts $\bar{y}_{1i}$ associated with the first stage constraints $f_{1i}(x_1) \leqq 0$, $i = 1, \ldots, m_1$, and $\bar{y}_{2i}(s)$ associated with the second stage constraints $f_{2i}(s, x_1, x_2) \leqq 0$, $i = 1, \ldots, m_2$. The components of $\bar{y}_{1i}$ can be interpreted as the marginal prices associated with the constraints $f_{1i} \leqq 0$, i.e., the maximal prices one would be willing to pay to have the constraints relaxed by one unit. The same interpretation is valid for $\bar{y}_{2i}(s)$, except that this time the price depends on the variable s. Knowing $\bar{y}_{2i}(s)$ for $i = 1, \ldots, m_2$, is as if we could predict the price system which would determine equilibrium prices, whatever be the outcome of the random events, up to a set of measure zero.

The variable $\chi_1(s)$ which appears in the formulation of problem (35), ..., (38) can be viewed as a possible alteration of the variable x_1 after the random elements have been observed. Naturally condition (36) requires that this alteration be in fact 0 with probability 1. Let us assume that equilibrium prices $(\bar{y}_1, \bar{y}_2)$ for problem (1), ..., (3), or equivalently problem (35), ..., (38), are known and that $(\bar{x}, \bar{y})$ is a saddle point for L with $\bar{x}$ an optimal solution for the original problem. Finding an optimal solution for (35), ..., (38) can be achieved by solving the following problem: find $x_1 \in R^{n_1}$, $\chi_1 \in \mathscr{L}^\infty_{n_1}$, $x_2 \in \mathscr{L}^\infty_{n_2}$ satisfying

$$x_1 \in C_1, \qquad x_2(s) \in C_2 \quad \text{a.s.}, \tag{40}$$

and satisfying almost surely

$$x_1 - \chi_1(s) = 0, \tag{41}$$

and minimizing

$$L_1(x_1, \bar{y}_1) + \int_S L_2(s, x_1, \chi_1(s), x_2(s), \bar{y}_2(s))\sigma(\mathrm{d}s). \tag{42}$$

Let h be the function to be minimized, i.e.,

$$h: R^{n_1} \times \mathscr{L}^{\infty}_{n_1} \times \mathscr{L}^{\infty}_{n_2} \to R$$

is the expression appearing in (42). We now embed this convex minimization problem in a class of perturbed problems, by replacing relation (41) by

$$x_1 - \chi_1(s) = v(s), \tag{43}$$

with $v(s)$ an element of $\mathscr{L}^{\infty}_{n_1}$. Let

$$G: (R^{n_1} \times \mathscr{L}^{\infty}_{n_1} \times \mathscr{L}^{\infty}_{n_2}) \times \mathscr{L}^{\infty}_{n_1} \to]-\infty, +\infty]$$

be defined as follows. If $(x_1, \chi_1, x_2) \in R^{n_1} \times \mathscr{L}^{\infty}_{n_1} \times \mathscr{L}^{\infty}_{n_2}$ and $v \in \mathscr{L}^{\infty}_{n_1}$ satisfy

$$x_1 \in C_1,$$

and almost surely

$$x_1 - \chi_1(s) = v(s), \qquad x_2(s) \in C_2,$$

then $G((x_1, \chi_1, x_2), v) = h(x_1, \chi_1, x_2)$; otherwise $G((x_1, \chi_1, x_2), v) = +\infty$. Problem (40), (41) and (42) is thus equivalent to finding (x_1, χ_1, x_2) such that $G((x_1, \chi_1, x_2), 0)$ is minimized.

The perturbation space $\mathscr{L}^{\infty}_{n_1}$ is put in duality with the (weak dual) space $\mathscr{L}^{1}_{n_1}$ by the bilinear form

$$\langle v, \rho \rangle = \int v(s) \cdot p(s) \sigma(\mathrm{d}s).$$

The resulting Lagrangian $K((x_1, \chi_1, x_2), \rho)$ is the function

$$K: (R^{n_1} \times \mathscr{L}^{\infty}_{n_1} \times \mathscr{L}^{\infty}_{n_2}) \times \mathscr{L}^{1}_{n_1} \to]-\infty, +\infty],$$

defined by

$$\begin{aligned} K((x_1, \chi_1, x_2), \rho) &= h(x_1, \chi_1, x_2) + \textstyle\int_S (x_1 - \chi_1(s)) \cdot \rho(s) \sigma(\mathrm{d}s), \\ &\qquad\qquad\qquad \text{if } \; x \in X_0, \\ &= +\infty, \qquad\qquad \text{if } \; x \notin X_0, \end{aligned}$$

and obtained by the formula

$$K((x_1, \chi_1, x_2), \rho) = \inf_{v \in \mathscr{L}^{\infty}{}_{n_1}} \{ G((x_1, \chi_1, x_2), v) + \langle v, \rho \rangle \}. \tag{44}$$

It is easy to verify that if $(\bar{x}, \bar{y})$ is a saddle point of L, then there exists $\bar{\chi}_1$ and ρ so that $(\bar{x}, \bar{\chi}_1, \bar{y}, \rho)$ is a saddle point of $\mathbf{L}$ and consequently that $((\bar{x}_1, \bar{\chi}_1, \bar{x}_2), \rho)$ is a saddle point of K. The purpose of this analysis has been to show that ρ can be interpreted as an 'equilibrium price' for possible perturbations $v(s)$ of the constancy requirement on x_1. In problem (40), . . ., (42), and also in problem (35), . . ., (38), we must still choose x_1 independent of s (before the random event, represented here by s, is observed); however, when carrying over the decision to the recourse problem we are allowed to modify this decision. If no modification is sought no additional cost is incurred. But if we seek a non-zero perturbation $v(s)$ a price must be paid. The interpretation of the Lagrangian multipliers $\rho(s)$ as equilibrium prices shows that no matter what particular s is actually observed, no perturbations $v(s)$ are actually worth buying at those prices or equivalently with the price system ρ no a postiori alteration of $\bar{x}_1$ will be profitable.

References

Blackwell, D., 1962, Discrete dynamic programming, Annals Math. Stat. 33, 719–726.

Blackwell, D. and M. Girshick, 1954, Theory of games and statistical decisions (Wiley, New York).

Dynkin, E., 1974, Optimal programs and stimulating prices in stochastic models of economic growth, in: J. Los and M. Los, eds., Mathematical models of economics (North-Holland, Amsterdam).

Fleming, W. and R. Rishel, 1975, Deterministic and stochastic optimal control (Springer, New York).

Gartska, S., 1973, An economic interpretation of stochastic programs, Manuscript (University of Chicago).

Gartska, S. and R. Wets, 1974, Decision rules for stochastic programs, Mathematical Programming 7.

Kushner, H., 1972, Necessary conditions for continuous parameter stochastic optimization problems, SIAM J. Control 10, 550–565.

Rockafellar, R.T., 1966, Extension of Fenchel's duality theorem for convex functions, Duke Math. Journal 33, 81–89.

Rockafellar, R.T., 1970, Convex analysis (Princeton University Press, Princeton, N.J.).

Rockafellar, R.T., 1971, Integrals which are convex functionals II, Pacific J. Math. 39, 143–148.

Rockafellar, R.T., 1974a, Conjugate duality and optimization, SIAM Monograph Series (Philadelphia).

Rockafellar, R.T., 1974b, Lagrange multipliers for an N-stage model in stochastic convex programming, in: J.P. Aubin, ed., Analyse convexe et ses applications, Lecture notes in mathematical systems and economics 102 (Springer, New York) 180–187.

Rockafellar, R.T. and R. Wets, forthcoming – a, Stochastic convex programming: Basic duality, Pacific J. Math.

Rockafellar, R.T. and R. Wets, forthcoming – b, Stochastic convex programming: Extended duality and singular multipliers, Pacific J. Math.

Rockafellar, R.T. and R. Wets, forthcoming – c, Stochastic convex programming: Relatively complete recourse and induced feasibility, SIAM J. Control.

Wets, R., 1970, Problèmes duaux en programmation stochastique, C.R. Acad. Sc. Paris 270, 47–50.

Wets, R., 1974, On the relation between stochastic and deterministic optimization, in: A. Bensoussan and J.L. Lions, eds., International Symposium on Control Theory, Numerical Methods and Computer Systems Modelling, I.R.I.A., Springer-Verlag lecture notes 107 (Springer, New York) 350–361.

Wets, R., 1975, Duality relations in stochastic programming, in: Symposia Mathematica (Academic Press, New York).