Control
and Cybernetics

VOL. 17 (1988) No. 2–3

Multistage convex programming and discrete-time optimal control

by

R. T. ROCKAFELLAR*

Department of Mathematics
University of Washington, GN-50
Seattle
Washington 98195, USA

Optimization problems of general convex type over finitely many time periods are given a new formulation that fits the framework of discrete-time optimal control but also brings out features of importance to applications in mathematical programming. Duality is explored and used to derive optimality conditions. Because of the nature of the problem formulation, these conditions are able to take the form of a "minimaximum" principle that decomposes both the primal and the dual with respect to time.

Keywords: discrete-time optimal control, multistage convex programming, dual problems of optimization, minimaximum principle, decomposition principle, decomposition of temporal constraints.

1. Problem Models with Convexity

This paper is concerned with the interface between discrete-time optimal control and convex programming, which encompasses linear and quadratic programming in particular. Despite the frequent presence of convexity, problems in optimal control have seldom in the past been viewed in a framework of convex programming. This is partly because the control literature, with its traditional emphasis on engineering applications, has not focused particularly on convexity and its consequences. Another reason has been attractiveness of working with concepts specific to control, like the maximum principle.

Problems in convex programming, on the other hand, even when they involve the management of discrete-time dynamical systems through a multi-

---

* This research was supported in part by a grant from the National Science Foundation at the University of Washington, Seattle.

stage decision process, have seldom been seen as instances of optimal control. To some extent this could be due to unfamiliarity of the users of multistage models in operations research with the mathematics of optimal control, typically thought of as infinite-dimensional. In any case, multistage convex programming and discrete-time optimal control have been developed along separate lines by rather separate communities of researchers. A potential now exists for a useful exchange of ideas.

The goal we set here is the introduction of new problem models in discrete-time optimal control that exhibit convexity and promote its role. These models are designed to appeal to mathematical programmers and to open the way to solution techniques in optimal control like some of those in the literature on large-scale convex programming. Our main results are duality relations and the characterization of optimality in terms of a "minimaximum principle".

The guidelines we follow are those of general duality theory [1], [2], and the piecewise linear-quadratic programming models in optimal control that we have developed in [3] for continuous time and in [4] for discrete time. A connection with the discrete-time Bolza problems in [5] may also be noted, although these do not explicitly involve controls and appear rather as analogues of problems in the calculus of variations.

The basic problem we propose to investigate has $N+1$ stages represented by state vectors

$$x_\tau \in R^{n_\tau} \quad \text{for} \quad \tau = 0, 1, \dots, N; \quad x = (x_0, x_1, \dots, x_N). \tag{1.1}$$

The dynamical system is taken to be linear, as a prerequisite to convexity in the problem (actually this is not as restrictive as it may seem), and is placed in the pattern of

$$x_\tau = A_\tau x_{\tau-1} + B_\tau u_\tau + b_\tau \quad \text{for} \quad \tau = 1, \dots, N, \tag{1.2}$$

$$x_0 = B_e u_e + B_e, \tag{1.3}$$

which involves "temporal" control vectors

$$u_\tau \in R^{k_\tau} \quad \text{for} \quad \tau = 1, \dots, N; \quad u = (u_1, \dots, u_N), \tag{1.4}$$

and a "terminal" control vector $u_e \in R^{k_e}$. The vector u_e represents supplementary parameters which may be adjusted in the problem in connection with endopoints. (The subscript e will consistently be used to mark endopoint elements). The nonstandard condition (1.3) allows of course for simple cases like $x_0 = b_e$ (fixed initial state). One can always trivialize u_e, if it is not needed in the model, by taking it to be 0-dimensional.

The reader should note well that the dimensions n_τ in (1.1) and k_τ in (1.4) are allowed to depend on τ. In a typical problem arising from the discretization of a continuous-time problem in optimal control, one would not have such variability: the vectors $x_0, x_1, \dots, x_N$ would all be in a certain

R^n, and $u_1, \dots, u_N$ in R^k. The equation (1.2) would arise from a difference equation

$$x_\tau - x_{\tau-1} = \bar{A}_\tau x_{\tau-1} + B_\tau u_\tau + b_\tau, \tag{1.5}$$

by setting

$$A_\tau = \bar{A}_\tau + I. \tag{1.6}$$

Such cases are obviously covered by our formulation in particular, but the provision for varying dimensionality enlarges the scope of the model quite significantly. In fact it enables the model in principle to encompass the dynamical structures of all multistage decision processes that can be expressed deterministically in terms of finitely many real variables. For if such a process requires the choice of a vector $u_\tau \in R^{k_\tau}$ for $\tau = 0, 1, \dots, N$ (subject presumably to constraints, which for the moment need not concern us), it is possible always to define the "history" $(u_0, u_1, \dots, u_\tau)$ of the process as the state x_τ at time τ, so that

$$x_\tau = (x_{\tau-1}, u_\tau), \qquad x_0 = u_0.$$

These relations can be written in the form of (1.2), (1.3), with

$$A_\tau = \begin{bmatrix} I & & & & \\ & I & & & \\ & & \ddots & & \\ & & & I & \\ & & & & 0 \end{bmatrix}, \quad B_\tau = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ I \end{bmatrix}, \quad B_e = I,$$

for identity matrices I and zero matrices 0 of appropriate sizes, and with b_τ and b_e taken to be zero vectors; u_0 is interpreted in this case as u_e.

This observation makes clear at the same time that the assumption of linearity in the dynamical system is, in itself, no real restriction but merely a convenient normalization for the purposes at hand. The true restrictions enter the model separately in the specifiation of what additional constraints one is allowed to impose on the relationship between $x_{\tau-1}$ and u_τ.

We are now ready to state our optimization problem in its general form, where any additional constraints on $x_{\tau-1}$ and u_τ beyond the dynamical relations are notationally supressed from view through the use of infinite penalities. The problem is

($\mathscr{P}$) minimize the expression

$$\mathscr{F}(u, u_e) = \sum_{\tau=1}^{N} [f_\tau(C_\tau x_{\tau-1}, u_\tau) - c_\tau \cdot x_{\tau-1}] + [f_e(C_e x_N, u_e) - c_e \cdot x_N],$$

over all $u = (u_1, u_N) \in R^{k_1} \times \dots \times R^{k_N}$ and $u_e \in R^{k_e}$,

where $x = (x_0, x_1, \dots, x_N)$ is given by (1.2)–(1.3).

Here $f_\tau : R^{l_\tau} \times R^{k_\tau} \to \bar{R}$ and $f_e : R^{l_e} \times R^{k_e} \to \bar{R}$ are extended-real-valued functions which are assumed to be convex, proper and lower semicontinuous, while C_τ and C_e are matrices of appropriate size, and c_τ and c_e are vectors.

Implicit in $(\mathscr{P})$ are general constraints of the form

$$(C_\tau x_{\tau-1}, u_\tau)\in F_\tau \quad \text{for} \quad r = 1, ..., N, \quad \text{and} \quad (C_e x_N, u_e)\in F_e, \tag{1.7}$$

where

$$F_\tau = \{(s_\tau, u_\tau)\in R^{l_\tau}\times R^{k_\tau} \mid f_\tau(s_\tau, u_\tau) < \infty\}, \tag{1.8}$$

$$F_e = \{(s_e, u_e)\in R^{l_e}\times R^{k_e} \mid f_e(s_e, u_e) < \infty\}. \tag{1.9}$$

Indeed, $\mathscr{F}(u, u_e) < \infty$ if and only if (1.7) is satisfied. The sets F_τ and F_e are nonempty and convex by virtue of the assumptions of convexity and properness placed on f_τ and f_e. They do not have to be closed, however; in some cases one could have $f_\tau(s_\tau, u_\tau)$ approach ∞ as (s_τ, u_τ) nears certain boundary points of F_τ, for instance, and similarly with f_e and F_e. The matrices C_τ and C_e can be identity matrices in particular, but more generally they allow us to deal in a convenient, specific way with the fact that in some models the constraints and objective terms may not fully depend on all the state components.

Represented in (1.7) are a great many possible cases involving restrictions on control vectors and/or state vectors. For the sake of illustration we shall focus here on the following case, which corresponds to an "ordinary" approach to convex programming.

EXAMPLE 1.1. Problem $(\mathscr{P})$ contains as a special case the problem of minimizing

$$\sum_{\tau=1}^{N} [f_{\tau 0}(u_\tau) - c_{\tau 0}\cdot x_{\tau-1}] + [f_{e0}(u_e) - c_{e0}\cdot x_N],$$

in the context of (1.2)–(1.3), subject to

$$f_{\tau i}(u_\tau) \leqslant c_{\tau i}\cdot x_{\tau-1} \quad \text{for } i = 1, ..., l_\tau, \text{ and } u_\tau\in U_\tau, \tag{1.11}$$

$$f_{ei}(u_e) \leqslant c_{ei}\cdot x_N, \quad \text{for } i = 1, ..., l_e, \text{ and } u_e\in U_e, \tag{1.12}$$

where the functions $f_{\tau i}: R^{k_\tau}\to R$ and $f_{ei}: R^{k_e}\to R$ are convex and finite, and the sets $U_\tau\subset R^{k_\tau}$ and $U_e\subset R^{k_e}$ are nonempty, convex and closed. This corresponds notationally to $c_\tau = c_{\tau 0}$, $c_e = c_{e0}$,

$$C_\tau = [l_\tau\times k_\tau \text{ matrix with rows } c_{\tau i}], \tag{1.13}$$

$$C_e = [l_e\times k_e \text{ matrix with rows } c_{ei}], \tag{1.14}$$

and in terms of

$$s_\tau = (..., s_{\tau i}, ...)\in R^{l_\tau} \text{ and } s_e = (..., s_{ei}, ...)\in R^{l_e}, \tag{1.15}$$

the definitions

$$f_\tau(s_\tau, u_\tau) = \begin{cases} \varphi_{\tau 0}(u_\tau) & \text{if } u_\tau \in U_\tau \text{ and } \varphi_{\tau i}(u_\tau) \leqslant s_{\tau i} \quad \text{for } i = 1, \ldots, l_\tau, \\ \infty & \text{for all other } u_\tau, \end{cases} \tag{1.16}$$

$$f_e(s_e, u_e) = \begin{cases} \varphi_{e0}(u_e) & \text{if } u_e \in U_e \text{ and } \varphi_{ei}(u_e) \leqslant s_{ei} \quad \text{for } i = 1, \ldots, l_e \\ \infty & \text{for all other } u_e. \end{cases} \tag{1.17}$$

The functions f_τ and f_e are indeed convex, proper and lower semicontinuous in this case.

EXAMPLE 1.2. Problem $(\mathscr{P})$ contains as a special case the problem of minimizing

$$\sum_{\tau=1}^{N} [p_\tau \cdot u_\tau - c_\tau \cdot x_{\tau-1}] + [p_e \cdot u_e - c_e \cdot x_N],$$

subject to (1.2), (1.3) and

$$C_\tau x_{\tau-1} + D_\tau u_\tau \geqslant q_\tau \text{ and } u_\tau \geqslant 0 \text{ for } \tau = 1, \ldots, N,$$
$$C_e x_N + D_e u_e \geqslant p_e \text{ and } u_e \geqslant 0.$$

All one has to do is to specialize Example 1.1 to

$$f_{\tau 0}(u_\tau) = p_\tau \cdot u_\tau \text{ and } f_{e0}(u_e) = p_e \cdot u_e,$$
$$f_{\tau i}(u_\tau) = q_{\tau i} - d_{\tau i} \cdot u_\tau \text{ for } i = 1, \ldots, l_\tau, \text{ and } U_\tau = R_+^{k_\tau},$$
$$f_{ei}(u_e) = q_{ei} - d_{ei} \cdot u_e \quad \text{for } i = 1, \ldots, l_e, \text{ and } U_e = R_+^{k_e},$$

where $d_{\tau i}$ is the i^{th} row of D_τ and d_{ei} is the i^{th} row of D_e. Other forms of the constraints involving equalities as well as inequalities, or even piecewise linear penalties can be set up in this way; also quadratic and piecewise quadratic programming models. For this we refer the reader to [3], [4].

THEOREM 1.3. *The essential objective function $\mathscr{F}$ being minimized in $(\mathscr{P})$ is convex and lower-semicontinuous (nowhere $-\infty$). Thus $(\mathscr{P})$ is a convex programming problem in the general sense, and its optimal solutions (u, u_e), if any, form a closed convex set.*

Proof. These properties are elementary consequences of our assumptions of f_τ and f_e. ∎

2. Minimax Representation and the Dual Problem

Problem $(\mathscr{P})$ can be given a minimax representation in terms of multiplier vectors

$$v_\tau = R^{l_\tau} \text{ for } \tau = 1, \ldots, N; \qquad v = (v_1, \ldots, v_N), \tag{2.1}$$

and $v_e \in R^{l_e}$. These will turn out later to be the control vectors in a dual dynamical system. To achieve such a representation we must introduce

$$J_\tau(u_\tau, v_\tau) = \inf_{s_\tau \in R^{l_\tau}} \{f_\tau(s_\tau, u_\tau) + s_\tau \cdot v_\tau\}, \tag{2.2}$$

$$J_e(u_e, v_e) = \inf_{s_e \in R^{l_\tau}} \{f_e(s_e, u_e) + s_e \cdot v_e\}. \tag{2.3}$$

Both formulas merely involve taking conjugates of the lower semicontinuous convex functions $f_\tau(\cdot, u_\tau)$ and $f_e(\cdot, u_e)$ along with certain changes of sign. They are invertible as

$$f_\tau(s_\tau, u_\tau) = \sup_{v_\tau \in R^{l_\tau}} \{J_\tau(u_\tau, v_\tau) - s_\tau \cdot v_\tau\}, \tag{2.4}$$

$$f_e(s_e, u_e) = \sup_{v_e \in R^{l_\tau}} \{J_e(u_e, v_e) - s_e \cdot v_e\}, \tag{2.5}$$

by the rules of convex analysis [1, §12].

PROPOSITION 2.1. *The function J_τ is convex-concave, proper and "upper-closed" on $R^{l_\tau} \times R^{l_\tau}$; likewise for J_e on $R^{k_e} \times R^{l_e}$.*

Proof. These are fundamental facts about the correspondence between convex functions and convex-concave functions. For the definitions and details, we refer to [1, §§34–35]. ∎

EXAMPLE 2.2. In the case described in Example 1.1 one has

$$J_\tau(u_\tau, v_\tau) = \begin{cases} f_{\tau 0}(u_\tau) + \sum_{i=1}^{l_\tau} v_{\tau i} f_{\tau i}(u_\tau) & \text{if } u_\tau \in U_\tau,\ v_\tau \in R_+^{l_\tau}, \\ -\infty & \text{if } u_\tau \in U_\tau,\ v_\tau \notin R_+^{l_\tau}, \\ \infty & \text{if } u_\tau \notin U_\tau. \end{cases} \tag{2.6}$$

and analogously

$$J_e(u_e, v_e) = \begin{cases} f_{e0}(u_e) + \sum_{i=1}^{l_e} v_{ei} f_{ei}(u_e) & \text{if } u_e \in U_e,\ v_e \in R_+^{l_e}, \\ -\infty & \text{if } u_e \in U_e,\ v_e \notin R_+^{l_e}, \\ \infty & \text{if } u_e \notin U_e. \end{cases} \tag{2.7}$$

Incidentally, something like the structure in this example can be shown to hold for J_τ and J_e in general. There always exist nonempty convex sets U_τ and V_τ (uniquely determined) such that J_τ is finite on $U_\tau \times V_\tau$ and

$$J_\tau(u_\tau, v_\tau) = -\infty \text{ when } u_\tau \in U_r,\ v_\tau \notin \operatorname{cl} V_\tau, \text{ or when } u_\tau \in \operatorname{ri} U_\tau,\ v_\tau \notin V_\tau, \tag{2.8}$$

$$J_\tau(u_\tau, v_\tau) = \infty \text{ when } v_\tau \in V_\tau,\ u_\tau \notin \operatorname{cl} U_\tau, \text{ or when } v_\tau \in \operatorname{ri} V_\tau,\ u_\tau \notin U_\tau. \tag{2.9}$$

(Here "ri" denotes the relative interior of a convex set [1, §6]). The set $U_\tau \times V_\tau$ is called the effective domain of J_τ; see [1, §34]. Similarly, J_e has an effective domain $U_e \times V_e$.

DEFINITION 2.3. The Lagrangian function associated with problem $(\mathscr{P})$ is

$$\mathscr{J}(u, u_e; v, v_e) = \sum_{\tau=1}^{N} J_\tau(u_\tau, v_\tau) + J_e(u_e, v_e) - [(u, u_e), (v, v_e)], \tag{2.10}$$

where

$$[(u, u_e), (v, v_e)] = \sum_{\tau=1}^{N} x_{\tau-1} \cdot [C_\tau^* v_\tau + c_\tau] + x_N \cdot [C_e^* v_e + c_e]. \tag{2.11}$$

(The convention $\infty - \infty = \infty$ is used to resolve conflicts in the extended arithmetic in (2.10). The asterisk $*$ in (2.11) marks the transpose of a matrix).

PROPOSITION 2.4. *The Lagrangian function $\mathcal{J}$ is convex-concave, proper and "upper closed".*

Proof. This is immediate from the corresponding properties in Proposition 2.1, since the term $[(u, u_e), (v, v_e)]$ is merely affine separately in (u, u_e) and in (v, v_e). ■

THEOREM 2.5. *The essential objective $\mathcal{F}$ in $(\mathcal{P})$ can be expressed by*

$$\mathcal{F}(u, u_e) = \sup_{(v, v_e)} \mathcal{J}(u, u_e; v, v_e).$$

Thus $(\mathcal{P})$ is the primal problem associated with $\mathcal{J}$.

Proof. Formulas (2.10) and (2.11) allow us to write

$$\mathcal{J}(u, u_e; v, v_e) = \sum_{\tau=1}^{N} [J_\tau(u_\tau, v_\tau) - (C_\tau x_{\tau-1}) \cdot v_\tau - c_\tau \cdot x_{\tau-1}] + $$
$$+ [J_e(u_e, v_e) - (C_e x_N) \cdot v_e - c_e \cdot x_N].$$

It follows that

$$\sup_{(v, v_e)} \mathcal{J}(u, u_e; v, v_e) = \sum_{\tau=1}^{N} \left[\sup_{v_\tau \in R^{l_\tau}} \{J_\tau(u_\tau, v_\tau) - (C_\tau x_{\tau-1}) \cdot v_\tau\} - C_\tau \cdot x_{\tau-1}\right] +$$
$$+ \left[\sup_{v_e \in \mathbf{R}^{l_e}} \{J_e(u_e, v_e) - (C_e x_N) \cdot v_e\} - c_e \cdot x_N\right],$$

and this reduces by (2.4) and (2.5) to the given definition of $\mathcal{F}(u, u_e)$. ■

Theorem 2.5 points the way towards setting up as dual to $(\mathcal{P})$ the problem of maximizing in (v, v_e) the infimum of $\mathcal{J}(u, u_e; v, v_e)$ with respect to (u, u_e). As the first step in that direction we show that the form $[(u, u_e), (v, v_e)]$ has an alternative expression in terms of the dual dynamical system

$$y_\tau = V_\tau^* y_{\tau+1} + C_\tau^* v_\tau + c_\tau \text{ for } \tau = 1, \ldots, N, \tag{2.12}$$

$$y_{N+1} = C_e^* v_e + c_e, \tag{2.13}$$

which involves the state vectors

$$y_\tau \in R^{k_\tau} \text{ for } \tau = 1, \ldots, N, N+1; \quad y = (y_1, \ldots, y_N, y_{N+1}). \tag{2.14}$$

This system can be integrated backward in time. The transformation $(v, v_e) \mapsto y$ is affine.

PROPOSITION 2.6. *The expression* (2.11) *can be written equivalently in terms of the dual dynamical system as*

$$[(u, u_e), (v, v_e)] = \sum_{\tau=1}^{N} y_{\tau+1} \cdot [B_\tau u_\tau + b_\tau] + y_1 \cdot [B_e u_e + b_e]. \tag{2.15}$$

Proof. The right side of (2.11) can be written by way of (2.12)–(2.13) as

$$\sum_{\tau=1}^{N} x_{\tau-1} \cdot [y_\tau - A_\tau^* y_{\tau+1}] + x_N \cdot y_{N+1},$$

while the right side of (2.15) can be written by way of (1.2)–(1.3) as

$$\sum_{\tau=1}^{N} y_{\tau+1} \cdot [x_\tau - A_\tau x_{\tau-1}] + y_1 \cdot [x_0].$$

These expressions both reduce to

$$x_0 \cdot y_1 + x_1 \cdot y_2 + \dots + x_N \cdot y_{N+1} - \sum_{\tau=1}^{N} y_{\tau+1} \cdot A_\tau x_{\tau-1},$$

and are therefore equal. ∎

Next in constructing the dual of $(\mathscr{P})$ we need functions g_τ and g_e whose relationship to J_τ and J_e is dual to that of f_τ and f_e in (2.4)–(2.5):

$$g_\tau(r_\tau, v_\tau) = \inf_{u_\tau \in R^{k_\tau}} \{J_\tau(u_\tau, v_\tau) - r_\tau \cdot u_\tau\}, \tag{2.16}$$

$$g_e(r_e, v_e) = \inf_{u_e \in R^{k_e}} \{J_e(u_e, v_e) - r_e \cdot u_e\}. \tag{2.17}$$

PROPOSITION 2.7. *The function* $g_\tau : R^{k_\tau} \times R^{l_\tau} \to \bar{R}$ for $\tau = 1, \dots, N$ *and* $g_e : R^{k_e} \times \times R^{l_e} \to \bar{R}$ *are concave, proper and upper semicontinuous. They are paired directly with the functions* f_τ *and* f_e *by the formulas*

$$g_\tau(r_\tau, v_\tau) = \inf_{s_\tau, u_\tau} \{f_\tau(s_\tau, u_\tau) - r_\tau \cdot u_\tau + v_\tau \cdot s_\tau\}, \tag{2.18}$$

$$g_e(r_e, v_e) = \inf_{s_e, u_e} \{f_e(s_e, u_e) - r_e \cdot u_e + v_e \cdot s_e\}, \tag{2.19}$$

and

$$f_\tau(s_\tau, u_\tau) = \sup_{r_\tau, v_\tau} \{g_\tau(r_\tau, v_\tau) - s_\tau \cdot v_\tau + u_\tau \cdot r_\tau\}, \tag{2.20}$$

$$f_e(s_e, u_e) = \sup_{r_e, v_e} \{g_e(r_e, v_e) - s_e \cdot v_e + r_e \cdot u_e\}. \tag{2.21}$$

Proof. The first pair of formulas is obtained by substituting (2.2) and (2.3) into (2.16) and (2.17). In terms of the conjugate functions f_τ^* and f_e^*, these say that

$$g_\tau(r_\tau, v_\tau) = -f_\tau^*(-v_\tau, r_\tau) \text{ and } g_e(r_e, v_e) = -f_e^*(-v_e, r_e).$$

Inasmuch as $(f_\tau^*)^* = f_\tau$ and $(f_e^*)^* = f_e$ (because f_τ and f_e are convex, proper and lower semicontinuous), we then have (2.20) and (2.21). ∎

We are able now to formulate the problem that in relation to the Lagrangian function $\mathscr{J}$ will be shown to be dual to $(\mathscr{P})$, namely

$(\mathscr{D})$ maximize the expression

$$\mathscr{G}(v, v_e) = \sum_{\tau=1}^{N} [g_\tau(B_\tau^* y_{\tau+1}, v_\tau) - b_\tau \cdot y_{\tau+1}] + [g_e(B_e^* y_1, u_e) - b_e \cdot y_1]$$

over all $v = (v_1, \ldots, v_N) \in R^{l_1} \times \ldots \times R^{l_N}$ and $v_e \in R^{l_e}$,

where $y = (y_1, \ldots, y_N, y_{N+1})$ is given by (2.12)–(2.13).

The nature of this problem will be elucidated in a moment, but first we record a crucial fact.

THEOREM 2.8. *The essential objective $\mathscr{G}$ in $(\mathscr{D})$ can be expressed by*

$$\mathscr{G}(v, v_e) = \inf_{(u, u_e)} \mathscr{J}(u, u_e; v, v_e).$$

Thus $(\mathscr{D})$ is the dual problem associated with $\mathscr{J}$.

Proof. Using (2.15) as the alternative expression for (2.11) in the definition (2.10) for $\mathscr{J}$, we obtain

$$\mathscr{J}(u, u_e; v, v_e) = \sum_{\tau=1}^{N} [J_\tau(u_\tau, v_\tau) - (B_\tau^* y_{\tau+1}) \cdot u_\tau - b_\tau \cdot y_{\tau+1}] + \\ + [J_e(u_e, v_e) - (B_e^* y_1) \cdot u_e - b_e \cdot y_1].$$

This yields

$$\inf_{(u, u_e)} \mathscr{J}(u, u_e; v, v_e) = \sum_{\tau=1}^{N} \big[\sup_{u_\tau \in R^{k_\tau}} \{J_\tau(u_\tau, v_\tau) - (B_\tau^* y_{\tau+1}) \cdot u_\tau\} - b_\tau \cdot y_{\tau+1}\big] + \\ + \big[\sup_{u_e \in R^{k_e}} \{J_e(u_e, v_e) - (B_e^* y_1) \cdot u_e\} - b_e \cdot y_1\big].$$

The definitions (2.16) and (2.17) of g_τ and g_e turn this into $\mathscr{G}(v, v_e)$. ∎

THEOREM 2.9. *The objective function $\mathscr{G}$ being maximized in $(\mathscr{D})$ is concave and upper semicontinuous (nowhere $+\infty$). Thus $(\mathscr{D})$ is a convex programming problem in the general sense, and its optimal solutions (v, v_e), if any, form a closed convex set.*

Proof. This follows at once from the properties of g_τ and g_e in Proposition 2.7. ∎

Problem $(\mathscr{D})$, like $(\mathscr{P})$, implicitly involves constraints of the form

$$(B_\tau^* y_{\tau+1}, v_\tau) \in G_\tau \text{ for } \tau = 1, \ldots, N, \text{ and } (B_e^* y_1, v_e) \in G_e, \tag{2.22}$$

where G_τ and G_e are the effective domains of the concave functions g_τ and g_e:

$$G_\tau = \{(r_\tau, v_\tau) \in R^{k_\tau} \times R^{l_\tau} | g_\tau(r_\tau, v_\tau) > -\infty\}, \tag{2.23}$$

$$G_e = \{(r_e, v_e) \in R^{k_e} \times R^{l_e} | g_e(r_e, v_e) > -\infty\}. \tag{2.24}$$

One has $\mathscr{G}(v, v_e) > -\infty$ if and only if (2.22) is satisfied. The sets G_τ and G_e are convex and nonempty, but they need not be closed (even though g_τ and g_e are upper semicontinuous).

EXAMPLE 2.10. In the case of the convex programming model in Examples 1.1 and 2.2, one has

$$g_\tau(r_\tau, v_\tau) = \begin{cases} \min_{u_\tau \in U_\tau} \{f_{\tau 0}(u_\tau) + \sum_{i=1}^{l_\tau} v_{\tau i} f_{\tau i}(u_\tau) - r_{\tau i} \cdot u_\tau\} & \text{if } v_\tau \geqslant 0, \\ -\infty \text{ if } v_\tau \ngeqslant 0. \end{cases}$$

$$g_e(r_e, v_e) = \begin{cases} \min_{u_e \in U_e} \{f_{e0}(u_e) + \sum_{i=1}^{l_e} v_{ei}[f_{ei}(u_e) - r_{ei}]\} & \text{if } v_e \geqslant 0, \\ -\infty & \text{if } v_e \ngeqslant 0. \end{cases}$$

Thus in (D) one seeks to maximize the expression

$$\sum_{\tau=1}^{N} \inf_{u_\tau \in U_\tau} \{(f_{\tau 0} + \sum_{i=1}^{l_\tau} v_{\tau i} f_{\tau i})(u_\tau) - y_{\tau+1} \cdot B_\tau u_\tau\} +$$

$$+ \inf_{u_e \in U_e} \{(f_{e0} + \sum_{i=1}^{l_e} v_{ei} f_{ei})(u_e) - y_1 \cdot B_e u_e\},$$

subject to $v_\tau \geqslant 0$ for $\tau = 1, \ldots, N$ and $v_e \geqslant 0$. This corresponds to the ordinary Lagrangian duality scheme in convex programming and suffers from the drawback that unless further assumptions are made about the functions involved, one cannot proceed to a level where the "inf" terms can be made more explicit. The linear case is an exception, as demonstrated in the example that follows. Other cases can be worked out too, but the real point is that the duality scheme plugs in at this stage to everything in the convex programming literature on ordinary duality.

EXAMPLE 2.11. In the linear programming case in Example 1.2, the dual problem consists of maximizing

$$\sum_{\tau=1}^{N} [q_\tau \cdot v_\tau - b_\tau \cdot y_{\tau+1}] + [q_e \cdot v_e - b_e \cdot y_1],$$

subject to (2.12), (2.13) and

$$B_\tau^* y_{\tau+1} + D_\tau^* v_\tau \leqslant p_\tau \text{ and } v_\tau \geqslant 0 \text{ for } \tau = 1, \ldots, N,$$

$$B_e^* y_1 + D_e^* v_e \leqslant p_e \text{ and } v_e \geqslant 0.$$

This can be seen by first calculating the Lagrangian terms

$$J_\tau(u_\tau, v_\tau) = \begin{cases} p_\tau \cdot u_\tau + q_\tau \cdot v_\tau - v_\tau \cdot D_\tau u_\tau & \text{if } u_\tau \geqslant 0,\ v_\tau \geqslant 0, \\ -\infty & \text{if } u_\tau \geqslant 0,\ v_\tau \ngeqslant 0, \\ \infty & \text{if } u_\tau \ngeqslant 0, \end{cases}$$

$$J_e(u_e, v_e) = \begin{cases} p_e \cdot u_e + q_e \cdot v_e - v_e \cdot D_e u_e & \text{if } u_e \geqslant 0,\ v_e \geqslant 0, \\ -\infty & \text{if } u_e \geqslant 0,\ v_e \ngeqslant 0, \\ \infty & \text{if } u_e \ngeqslant 0, \end{cases}$$

from (2.2), (2.3), and then using the definitions (2.16), (2.17) to obtain

$$g_\tau(r_\tau, v_\tau) = \begin{cases} q_\tau \cdot v_\tau & \text{if } r_\tau + D_\tau^* v_\tau \leqslant p_\tau, \\ -\infty & \text{if } r_\tau + D_\tau^* v_\tau \nleqslant p_\tau, \end{cases}$$

$$g_e(r_e, v_e) = \begin{cases} q_e \cdot v_e & \text{if } r_e + D_e^* v_e \leqslant p_e, \\ -\infty & \text{if } r_e + D_e^* v_e \leqslant p_e. \end{cases}$$

See [4] for extensions of this pattern to piecewise linear and quadratic programming.

3. Duality Relations

Theorems relating the optimal values in $(\mathscr{P})$ and $(\mathscr{D})$, namely the quantities

$$\inf(\mathscr{P}) = \inf_{(u, u_e)} \mathscr{F}(u, u_e), \qquad \sup(\mathscr{D}) = \sup_{(v, v_e)} \mathscr{G}(v, v_e),$$

are the key to deriving optimality conditions for these problems, because of convexity. They also furnish criteria for the existence of optimal solutions.

The inequality

$$\inf(\mathscr{P}) \geqslant \sup(\mathscr{D}),$$

always holds by virtue of the formulas for $\mathscr{F}$ and $\mathscr{G}$ in terms of the Lagrangian $\mathscr{J}$ as demonstrated in Theorems 2.5 and 2.9:

$$\inf(\mathscr{P}) = \inf_{(u, u_e)} \sup_{(v, v_e)} \mathscr{J}(u, u_e; v, v_e) \geqslant \sup_{(v, v_e)} \inf_{(u, u_e)} \mathscr{J}(u, u_e; v, v_e) = \sup(\mathscr{D}).$$

Our interest lies in the circumstances under which $\inf(\mathscr{P}) = \sup(\mathscr{D})$ holds and one or both of these extrema is attained. We use the convention of writing $\min(\mathscr{P})$ in place of $\inf(\mathscr{P})$, or $\max(\mathscr{D})$ in place of $\sup(\mathscr{D})$, to indicate attainment. In the general convex case we are dealing with, additional assumptions in the form of "constraint qualifications" are needed for the results we want.

DEFINITION 3.1. We shall say that the *primal constraint qualification* holds if for some choice of (u, u_e) and the corresponding primal trajectory x determined from (1.2)–(1.3) one has

$$(C_\tau x_{\tau+1}, u_\tau) \in \operatorname{ri} F_\tau \quad \text{for} \quad \tau = 1, \ldots, N, \quad \text{and} \quad (C_e x_N, u_e) \in \operatorname{ri} F_e, \tag{3.1}$$

where F_τ and F_e are the convex sets in (1.8)–(1.9) and "ri" denotes relative interior. (See [1, §6] for a discussion of relative interiors and how they can be calculated in various situations). Similarly, we shall say that the *dual constraint qualification* holds if for some choice of (v, v_e) and the corresponding dual trajectory y determined from (2.12)–(2.13) one has

$$(B_\tau^* y_{\tau+1}, v_\tau) \in \operatorname{ri} G_\tau \quad \text{for} \quad \tau = 1, \ldots, N, \quad \text{and} \quad (B_e^* y_1, v_e) \in \operatorname{ri} G_e. \tag{3.2}$$

THEOREM 3.2. *If the primal constraint qualification holds, one has*

$$\infty > \inf(\mathscr{P}) = \max(\mathscr{D}) \geqslant -\infty, \tag{3.3}$$

while if the dual constraint qualification holds one has

$$\infty \geqslant \min(\mathscr{P}) = \sup(\mathscr{D}) > -\infty. \tag{3.4}$$

If both hold, one therefore has

$$\infty > \min(\mathscr{P}) = \max(\mathscr{D}) > -\infty. \tag{3.5}$$

Proof. The general duality theory for optimization problems of convex type will be applied as in [2] and more specifically [1, §30]. This requires the introduction of primal perturbations $w = (w_1, \ldots, w_N)$ and w_e with

$$w_\tau \in R^{l_\tau} \quad \text{for} \quad \tau = 1, \ldots, N, \quad \text{and} \quad w_e \in R^{l_e},$$

and the function

$$\Phi(u, u_e; w, w_e) = \sup_{(v, v_e)} \{\mathscr{J}(u, u_e; v, v_e) - (w, w_e) \cdot (v, v_e)\}, \tag{3.6}$$

as well as dual perturbations $z = (z_1, \ldots, z_N)$ and z_e with

$$z_\tau \in R^{k_\tau} \quad \text{for} \quad \tau = 1, \ldots, N, \quad \text{and} \quad z_e \in R^{k_e},$$

and the function

$$\Psi(v, v_e; z, z_e) = \inf_{(u, u_e)} \{\mathscr{J}(u, u_e; v, v_e) - (z, z_e) \cdot (u, u_e)\}. \tag{3.7}$$

Clearly from Theorems 2.5 and 2.9 one has

$$\mathscr{F}(u, u_e) = \Phi(u, u_e; 0, 0) \quad \text{and} \quad \mathscr{G}(v, v_e) = \Psi(v, v_e; 0, 0). \tag{3.8}$$

In fact the calculations in these theorems give

$$\Phi(u, u_e; w, w_e) = \sum_{\tau=1}^{N} [f_\tau(C_\tau x_{\tau-1} + w_\tau, u_\tau) - c_\tau \cdot x_{\tau-1}] + \\ + [f_e(C_e x_N + w_e, u_e) - c_e \cdot x_N], \tag{3.9}$$

$$\Psi(v, v_e; z, z_e) = \sum_{\tau=1}^{N} [g_\tau(B_\tau^* x_{\tau+1} + z_\tau, v_\tau) - b_\tau \cdot y_{\tau+1}] + \\ + [g_e(B_e^* y_N + z_e, v_e) - b_e \cdot y_1], \tag{3.10}$$

where as always, x is determined from (u, u_e) by the primal dynamics (1.2)–(1.3) and y from (v, v_e) by the dual dynamical (2.16)–(2.17). The functions Φ and φ, the latter defined by

$$\varphi(w, w_e) = \inf_{(u, u_e)} \Phi(u, u_e; w, w_e) \quad [\varphi(0, 0) = \inf(\mathscr{P})], \tag{3.11}$$

are convex, while the functions Ψ and ψ, the latter defined by

$$\psi(z, z_e) = \sup_{(v, v_e)} \Psi(v, v_e; z, z_e) \quad [\psi(0,0) = \sup(\mathscr{D})], \tag{3.12}$$

are concave.

Duality theory centers on the properties of these perturbation functions, in particular their effective domains, which are the convex sets

$$\operatorname{dom} \varphi = \{(w, w_e) | \varphi(w, w_e) < \infty\} = \{(w, w_e) | \exists (u, u_e) \text{ such that}$$
$$(C_\tau x_{\tau-1} + w_\tau, u_\tau) \in F_\tau \text{ for } \tau = 1, \dots, N, \text{ and } (C_e x_N + w_e, u_e) \in F_e\}, \tag{3.13}$$

$$\operatorname{dom} \psi = \{(z, z_e) | \psi(z, z_e) > -\infty\} = \{(z, z_e | \exists (v, v_e) \text{ such that}$$
$$(B_\tau^* y_{\tau+1} + z_\tau, v_\tau) \in G_\tau \text{ for } \tau = 1, \dots, N, \text{ and } (B_e^* y_1 + z_e, v_e) \in G_e\}. \tag{3.14}$$

These are important because of the conjugacy relations

$$\mathscr{G}(v, v_e) = -\varphi^*(v, v_e) = \inf_{(w, w_e)} \{\varphi(w, w_e) - (v, v_e) \cdot (w, w_e)\}, \tag{3.15}$$

$$\mathscr{F}(u, u_e) = -\psi^*(u, u_e) = \sup_{(z, z_e)} \{\psi(z, z_e) - (u, u_e) \cdot (z, z_e)\}. \tag{3.16}$$

These relations can readily be verified directly from our formulas, but they also hold by the general duality scheme being employed [1, Thm. 30.2].

One knows from conjugate function theory that

$$\infty > \varphi(0,0) = \max\{-\varphi^*\} \quad \text{if} \quad (0,0) \in \operatorname{ri} \operatorname{dom} \varphi, \tag{3.17}$$

(because φ is convex), and

$$-\infty < \psi(0,0) = \min\{-\psi^*\} \quad \text{if} \quad (0,0) \in \operatorname{ri} \operatorname{dom} \psi, \tag{3.18}$$

(because ψ is concave) [1, Thm. 27.1]. The equation in (3.17) corresponds to the duality assertion (3.3), and the equation in (3.18) to (3.4). The proof we are faced with reduces then to the verification that

$$(0,0) \in \operatorname{ri} \operatorname{dom} \varphi \Leftrightarrow \text{the primal constraint qualification is satisfied,} \tag{3.19}$$

$$(0,0) \in \operatorname{ri} \operatorname{dom} \psi \Leftrightarrow \text{the dual constraint qualification is satisfied.} \tag{3.20}$$

To calculate the set $\operatorname{ri} \operatorname{dom} \varphi$ we define the convex set

$$F = F_1 \times \dots \times F_N \times F_e,$$

the affine transformation

$$T : (w, w_e; u, u_e) \mapsto (C_1 x_0 + w_1, u_1, \dots, C_N x_{N-1} + w_N, u_N; C_e x_N + w_e, u_e),$$

and the linear transformation

$$S : (w, w_e; u, u_e) \mapsto (w, w_e).$$

These allow us to write (3.13) in the form

$$\operatorname{dom} \varphi = S(T^{-1}(F)).$$

The calculus rules for relative interiors of convex sets in [1, Thms. 6.6, 6.7] then yield

$$\operatorname{ri}\operatorname{dom}\varphi = S\left(T^{-1}(\operatorname{ri} F)\right),$$

where

$$\operatorname{ri} F = (\operatorname{ri} F_1)\times \dots \times(\operatorname{ri} F_N)\times(\operatorname{ri} F_e).$$

This is precisely what (3.19) means. The verification of (3.20) follows the same lines. ∎

REMARK 3.3. The proof of Theorem 3.2 provides a basis for interpreting the optimal solutions to $(\mathscr{D})$ relative to $(\mathscr{P})$. It shows through (3.15) and (3.19) that under the primal constraint qualification one has

$$\partial\varphi\,(0,0) = \arg\max\,(\mathscr{D}), \tag{3.21}$$

$$\varphi'(0,0;w,w_e) = \sup\left\{(w,w_e)\cdot(v,v_e)\,|\,(v,v_e)\in\arg\max\,(\mathscr{D})\right\}, \tag{3.22}$$

[1, Thms. 23.4, 23.5]. Similarly it shows through (3.16) and (3.20) that under the dual constraint qualification one has

$$\partial\psi\,(0,0) = \arg\min\,(\mathscr{P}), \tag{3.23}$$

$$\psi'(0,0;z,z_e) = \inf\left\{(z,z_e)\cdot(u,u_e)\,|\,(u,u_e)\in\arg\min\,(\mathscr{P})\right\}. \tag{3.24}$$

Although the significance of the primal and dual constraint qualifications can be brought into sharper detail in a particular instance of problem $(\mathscr{P})$ by the use of the calculus of relative interiors, both conditions can also be stated in another form that in some situations could be easier to verify. This other form involves the recession functions

$$\hat{f}_\tau(\hat{s}_\tau,\hat{u}_\tau) = \lim_{\lambda\uparrow\infty}\left[f_\tau(s_\tau+\lambda\hat{s},u_\tau+\lambda\hat{u}_\tau) - f_\tau(s_\tau,u_\tau)\right]/\lambda \quad \text{for } (s_\tau,u_\tau)\in F_\tau, \tag{3.25}$$

$$\hat{f}_e(\hat{s}_e,\hat{u}_e) = \lim_{\lambda\uparrow\infty}\left[f_e(s_e+\lambda\hat{s}_e,u_e+\lambda\hat{u}_e) - f_e(s_e,u_e)\right]/\lambda \quad \text{for } (s_e,u_e)\in F_e, \tag{3.26}$$

$$\hat{g}_\tau(\hat{r}_\tau,\hat{v}_\tau) = \lim_{\lambda\uparrow\infty}\left[g_\tau(r_\tau+\lambda\hat{r}_\tau,v_\tau+\lambda\hat{v}_\tau) - g_\tau(r_\tau,v_\tau)\right]/\lambda \quad \text{for } (r_\tau,v_\tau)\in G_\tau, \tag{3.27}$$

$$\hat{g}_e(\hat{r}_e,\hat{v}_e) = \lim_{\lambda\uparrow\infty}\left[g_e(r_e+\lambda\hat{r}_e,v_e+\lambda\hat{v}_e) - g_e(r_e,v_e)\right]/\lambda \quad \text{for } (r_e,v_e)\in G_e. \tag{3.28}$$

(These formulas are instensitive to the choice of the base point, as long as it belongs to the effective domain in question. Thus in (3.25), for instance, one gets the same function $\hat{f}_\tau$ regardless of the particular choice of $(s_\tau,u_\tau)\in F_\tau$; see [1, Thm, 8.5]). In terms of these functions we define

$$\hat{\mathscr{F}}(\hat{u},\hat{u}_e) = \sum_{\tau=1}^{N}\left[\hat{f}_\tau(C_\tau\hat{x}_{\tau-1},\hat{u}_\tau) - c_\tau\cdot\hat{x}_{\tau-1}\right] + \left[\hat{f}_e(C_e\hat{x}_N,\hat{u}_e) - c_e\cdot\hat{x}_N\right], \tag{3.29}$$

where $\hat{x}$ is the state trajectory generated from $(\hat{u},\hat{u}_e)$ by the "homogenized primal dynamics":

$$\hat{x}_\tau = A_\tau \hat{x}_{\tau-1} + B_\tau \hat{u}_\tau \quad \text{for} \quad \tau = 1, \dots, N, \quad \text{with} \quad \hat{x}_0 = B_e \hat{u}_e, \tag{3.30}$$

(where b_τ and b_e have been suppressed). We also define

$$\hat{\mathscr{G}}(\hat{v}, \hat{v}_e) = \sum_{\tau=1}^{N} [\hat{g}_\tau (B_\tau^* \hat{y}_{\tau+1}, \hat{v}_\tau) - b_\tau \cdot \hat{y}_{\tau+1}] + [\hat{g}_e (B_e^* \hat{y}_1, \hat{v}_e) - b_e \cdot \hat{y}_1], \tag{3.31}$$

where $\hat{y}$ is the state trajectory generated from $(\hat{v}, \hat{v}_e)$ by the "homogenized dual dynamics":

$$\hat{y}_\tau = A_\tau^* \hat{y}_{\tau+1} + C_\tau^* \hat{v}_\tau \quad \text{for} \quad \tau = 1, \dots, N, \quad \text{with} \quad \hat{y}_{N+1} = C_e^* v_e. \tag{3.32}$$

PROPOSITION 3.4.

(a) *The dual constraint qualification holds if and only if every* $(\hat{u}, \hat{u}_e)$ *satisfying* $\hat{\mathscr{F}}(\hat{u}, \hat{u}_e) \leqslant 0$ *actually satisfies* $\hat{\mathscr{F}}(\hat{u}, \hat{u}_e) = 0 = \hat{\mathscr{F}}(-\hat{u}, -\hat{u}_e)$.

(b) *The primal constraint qualification holds if and only if every* $(\hat{v}, \hat{v}_e)$ *satisfying* $\hat{\mathscr{G}}(\hat{v}, \hat{v}_e) \geqslant 0$ *actually satisfies* $\hat{\mathscr{G}}(\hat{v}, \hat{v}_e) = 0 = \hat{\mathscr{G}}(-\hat{v}, -\hat{v}_e)$.

Proof. We shall demonstrate that $\hat{\mathscr{F}}$ is the support function of the convex set dom ψ,

$$\hat{\mathscr{F}}(\hat{u}, \hat{u}_e) = \sup \{(\hat{u}, \hat{u}_e) \cdot (z, z_e) | (z, z_e) \in \operatorname{dom} \psi\}, \tag{3.33}$$

and similarly for $\hat{\mathscr{G}}$, with a change of signs for the sake of concavity:

$$\hat{\mathscr{G}}(\hat{v}, \hat{v}_e) = \inf \{(\hat{v}, \hat{v}_e) \cdot (w, w_e) | (w, w_e) \in \operatorname{dom} \varphi\}. \tag{3.34}$$

The desired conclusions will then be immediate from (3.19), (3.20), and the basic theory of support functions [1, Cor. 13.3.4].

The conjugacy between f_τ and g_τ as expressed by (2.18) and (2.20) yields by [1, Thm. 13.5] the support function formulas

$$\hat{f}_\tau (\hat{s}_\tau, \hat{u}_\tau) = \sup \{\hat{u}_\tau \cdot r_\tau - \hat{s}_\tau \cdot v_\tau | (r_\tau, v_\tau) \in G_\tau\}, \tag{3.35}$$

$$\hat{g}_\tau (\hat{r}_\tau, \hat{v}_\tau) = \inf \{\hat{v}_\tau \cdot s_\tau - \hat{r}_\tau \cdot u_\tau | (s_\tau, v_\tau) \in F_\tau\}.$$

Likewise from (2.19) and (2.21):

$$\hat{f}_e (\hat{s}_e, \hat{u}_e) = \sup \{\hat{u}_e \cdot r_e - \hat{s}_e \cdot v_e | (r_e, v_e) \in G_e\}, \tag{3.37}$$

$$\hat{g}_e (\hat{r}_e, \hat{v}_e) = \inf \{\hat{v}_e \cdot s_e - \hat{r}_e \cdot u_e | (s_e, v_e) \in F_e\}. \tag{3.38}$$

Working first toward (3.33), we use the description of dom ψ in (3.14) with

$$r_\tau = B_\tau^* y_{\tau+1} + z_\tau \quad \text{and} \quad r_e = B_e^* y_1 + z_e,$$

to write

$$\sup \{(\hat{u}, \hat{u}_e) \cdot (z, z_e) | (z, z_e) \in \operatorname{dom} \psi\} = \sup \Big\{ \sum_{\tau=1}^{N} \hat{u}_\tau \cdot [r_\tau - B_\tau^* y_{\tau+1}] +$$

$$+ \hat{u}_e \cdot [r_e - B_e^* y_1] \,|\, (r_\tau, v_\tau) \in G_\tau \text{ for some } v_\tau, (r_e, v_e) \in G_e \text{ for some } v_e \Big\}. \tag{3.39}$$

The equation

$$\sum_{\tau=1}^{N} y_{\tau+1} \cdot B_\tau u_\tau + y_1 \cdot B_e u_e = \sum_{\tau=1}^{N} \hat{x}_{\tau-1} \cdot [C_\tau^* v_\tau + c_\tau] + \hat{x}_0 \cdot [C_e^* v_e + c_e],$$

is true as a special case of the one in Proposition 2.6, namely where the b_τ and b_e terms are omitted so that the primal dynamics is given by (3.30). The value in (3.39) can therefore be written also as

$$\sup \Big\{ \sum_{\tau=1}^{N} (\hat{u}_\tau \cdot r_\tau - \hat{x}_{\tau-1} \cdot [C_\tau^* v_\tau + c_\tau]) + (\hat{u}_e \cdot r_e - \hat{x}_N \cdot [C_e^* v_e + c_e]) \Big| (r_\tau, v_\tau) \in G_\tau$$

$$\text{for } \tau = 1, \ldots, N, \text{ and } (r_e, v_e) \in G_e \Big\} = \sum_{\tau=1}^{N} \Big[\sup_{(r_\tau, v_\tau) \in G_\tau} \{ \hat{u}_\tau \cdot r_\tau -$$

$$- (C_\tau \hat{x}_{\tau-1}) \cdot v_\tau \} - c_\tau \cdot \hat{x}_{\tau-1} \Big] \Big[\sup_{(r_e, v_e) \in G_e} \{ \hat{u}_e \cdot r_e - (C_e \hat{x}_N) \cdot v_e \} - c_e \cdot \hat{x}_N \Big], \quad (3.40)$$

which by (3.35) and (3.37) turns out to be $\hat{\mathscr{F}}(\hat{u}, \hat{u}_e)$, as defined in (3.29). Thus (3.33) is true. The verification of (3.34) follows the same pattern. ∎

PROPOSITION 3.5. *If* ($\mathscr{P}$) *has at least one feasible solution, i.e. there exists* (u, u_e) *such that the implicit constraints* (1.7) *are satisfied, then for any such* (u, u_e) *one has the formula*

$$\hat{\mathscr{F}}(\hat{u}, \hat{u}_e) = \lim_{\lambda \uparrow \infty} [\mathscr{F}(u + \lambda \hat{u}, u_e + \lambda \hat{u}_e) - \mathscr{F}(u, u_e)]/\lambda. \quad (3.41)$$

In other words, $\hat{\mathscr{F}}$ *is the recession function associated with* $\mathscr{F}$. *Duality, if* ($\mathscr{D}$) *has at least one* (v, v_e) *satisfying the implicit constraints* (2.22), *then*

$$\hat{\mathscr{G}}(\hat{v}, \hat{v}_e) = \lim_{\lambda \uparrow \infty} [\mathscr{G}(v + \lambda \hat{v}, v_e + \lambda \hat{v}_e) - \mathscr{G}(v, v_e)]/\lambda. \quad (3.42)$$

Proof. These expressions can be calculated directly from the definitions of $\hat{\mathscr{F}}$ and $\hat{\mathscr{G}}$ those of the functions $\hat{f}_\tau$, $\hat{g}_\tau$, $\hat{f}_e$, and $\hat{g}_e$. ∎

COROLLARY 3.6. *The dual constraint qualification is satisfied in particular if for some number* α *the level set* $\{(u, u_e) | \mathscr{F}(u, u_e) \leqslant \alpha\}$ *is nonempty and bounded, as for instance when the set of optimal solutions to* ($\mathscr{P}$) *is nonempty and bounded* (*the case of* $\alpha = \inf(\mathscr{P})$).

Similarly, the primal constraint qualification is satisfied in particular if for some number β *the level set* $\{(v, v_e) | \mathscr{G}(v, v_e) \geqslant \beta\}$ *is nonempty and bounded, as for instance when the set of optimal solutions to* ($\mathscr{D}$) *is nonempty and bounded* (*the case of* $\beta = \sup(\mathscr{D})$).

Proof. The first of the level set properties corresponds to $\mathscr{F}$ being a proper convex function whose recession function is positive except at the origin; cf. [1, Thm. 8.7]. Similarly for the second property. ∎

EXAMPLE 3.7. In the ordinary convex programming case described in Examples 1.1 and 2.2, the primal constraint qualification holds in particular if there exists (u, u_e) satisfying (with the corresponding x):

$$f_{\tau i}(u_\tau) < c_{\tau i} \cdot x_{\tau-1} \quad \text{for} \quad i=1, \ldots, l_\tau, \quad \text{and} \quad u_\tau \in U_\tau, \tag{3.43}$$

$$f_{ei}(u_e) < c_{ei} \cdot x_N \quad \text{for} \quad i=1, \ldots, l_e, \quad \text{and} \quad u_e \in U_e. \tag{3.44}$$

This can be seen right from Definition 3.2 and the fact that $U_\tau \subset cl\,[\text{ri}\, U_\tau]$ and $U_e \subset cl\,[\text{ri}\, U_e]$. (Actually the primal constraint qualification is more subtle than (3.43)–(3.45) and nicely covers cases where linear equations have been represented by pairs of inequalities). The dual constraint qualification, on the other hand, can be analyzed in the form provided by Proposition 3.4(a). Let us introduce the recesssion functions

$$\hat{f}_{\tau i}(\hat{u}_\tau) = \lim_{\lambda \uparrow \infty} [f_{\tau i}(u_\tau + \lambda \hat{u}_\tau) - f_{\tau i}(u_\tau)]/\lambda,$$

and similarly $\hat{f}_{ei}(\hat{u}_e)$. (These functions might be extended-real-valued even though $f_{\tau i}$ and f_{ei} are finite everywhere). Let us also denote by $\hat{U}_\tau$ and $\hat{U}_e$ the recession cones of U_τ and U_e [1, §8]. Then

$$\hat{f}_\tau(C_\tau x_{\tau-1}, u_\tau) = \begin{cases} \hat{f}_{\tau 0}(\hat{u}_\tau) & \text{if } \hat{u}_\tau \in \hat{U}_\tau \text{ and} \\ & \quad \hat{f}_{\tau i}(\hat{u}_\tau) \leqslant c_{\tau i}\, \hat{x}_{\tau-1} \text{ for } i=1, \ldots, l_\tau, \\ \infty & \text{otherwise,} \end{cases}$$

$$\hat{f}_e(C_e x_N, u_e) = \begin{cases} \hat{f}_{e0}(\hat{u}_e) & \text{if } \hat{u}_e \in \hat{U}_e \text{ and} \\ & \quad \hat{f}_{ei}(\hat{u}_e) \leqslant c_{ei}\, \hat{x}_N \text{ for } i=1, \ldots, l_e, \\ \infty & \quad \text{otherwise.} \end{cases}$$

The dual constraint qualification is therefore satisfied in particular if the only choice of $(\hat{u}, \hat{u}_e)$ such that

$$\sum_{\tau=1}^{N} [\hat{f}_{\tau 0}(\hat{u}_\tau) - c_\tau \cdot \hat{x}_{\tau-1}] + [\hat{f}_{e0}(\hat{u}_e) - c_e \cdot \hat{x}_N] \leqslant 0,$$

$$\hat{f}_{\tau i}(\hat{u}_\tau) \leqslant c_{\tau i} \cdot \hat{x}_{\tau-1} \quad \text{for} \quad i=1, \ldots, l_\tau, \quad \text{and} \quad \hat{u}_\tau \in \hat{U}_\tau,$$

$$\hat{f}_{ei}(\hat{u}_e) \leqslant c_{ei} \cdot \hat{x}_N \quad \text{for} \quad i=1, \ldots, l_e, \quad \text{and} \quad \hat{u}_e \in \hat{U}_e,$$

is $(\hat{u}, \hat{u}_e) = (0, 0)$. This is trivially true, for instance, if the sets U_τ and U_e are bounded (because then $U_\tau = \{0\}$ and $U_e = \{0\}$), and this case is also obvious from Corollary 3.6. More generally, however, the functions $\hat{f}_{\tau i}$ and $\hat{f}_{ei}$ express the growth properties of $f_{\tau i}$ and f_{ei}, which can lead to the dual constraint qualification being easily verifiable in cases where U_τ and U_e are not necessarily bounded.

4. Saddle Point Conditions and Decomposition

The Lagrangian $\mathscr{J}$ which we have introduced in representing problem $(\mathscr{P})$ and constructing its dual $(\mathscr{D})$ has the important property that $\mathscr{J}(u, u_e; v, v_e)$ is separable in the u_τ and u_e components for fixed (v, v_e)

but also separable in the v_τ and v_e components for fixed (u, u_e). This is true because of the alternative formulas for the term $[(u, u_e), (v, v_e)]$ provided in (2.11) and (2.15). We shall demonstrate in this section that such separability leads to a primal-dual decomposition scheme in the form of a "minimaximum" principle like the one recently derived for the special case of piecewise linear-quadratic optimal control in [3], [4]. This principle gives a joint decomposition of $(\mathscr{P})$ and $(\mathscr{D})$ with respect to time. A predecessor in "continuous-time programming" can be seen in the work of Grinold [6, p. 46] and another in a context of optimal control and the calculus of variations in Rockafellar [7, Thm. 6].

To set the stage, we begin with a fundamental fact about the relationship between $(\mathscr{P})$ and $(\mathscr{D})$.

THEOREM 4.1. *One has* $\min(\mathscr{P}) = \max(\mathscr{D})$ *if and only if* $\mathscr{J}$ *has a saddle point, in which event the saddle points of* $\mathscr{J}$ *are precisely the elements* $(\bar{u}, \bar{u}_e; \bar{v}, \bar{v}_e)$ *such that* $(\bar{u}, \bar{u}_e)$ *is an optimal solution to* $(\mathscr{P})$ *and* $(\bar{v}, \bar{v}_e)$ *is an optimal solution to* $(\mathscr{D})$. *Then*

$$\min(\mathscr{P}) = \max(\mathscr{D}) = \mathscr{J}(\bar{u}, \bar{u}_e; \bar{v}, \bar{v}_e) \quad [\textit{finite}].$$

Proof. We get this immediately from the representations established in Theorems 2.5 and 2.9. We need only invoke elementary and well known facts of general duality theory (cf. [2, Thm. 2], for instance). ∎

COROLLARY 4.2.

(a) *If the primal constraint qualification holds, then a necessary as well as sufficient condition for* $(\bar{u}, \bar{u}_e)$ *to be an optimal solution to* $(\mathscr{P})$ *is the existence of some* $(\bar{v}, \bar{v}_e)$ *such that* $(\bar{u}, \bar{u}_e; \bar{v}, \bar{v}_e)$ *is a saddle point of* $\mathscr{J}$.

(b) *If the dual constraint qualification holds, then a necessary as well as sufficient condition for* $(\bar{v}, \bar{v}_e)$ *to be an optimal solution to* $(\mathscr{D})$ *is the existence of some* $(\bar{u}, \bar{u}_e)$ *such that* $(\bar{u}, \bar{u}_e; \bar{v}, \bar{v}_e)$ *is a saddle point of* $\mathscr{J}$.

Proof. This combines Theorem 4.1 with Theorem 3.2. ∎

THEOREM 4.3. ("*Minimaximum Principle*"). *In order that* $(\bar{u}, \bar{u}_e; \bar{v}, \bar{v}_e)$ *be a saddle point of* $\mathscr{J}$, *it is necessary and sufficient that the following conditions in terms of the corresponding state trajectories* $\bar{x}$ *and* $\bar{y}$ *hold at each time* τ:

$(\bar{u}_\tau, \bar{v}_\tau)$ is a saddle point of

$$\bar{J}_\tau(u_\tau, v_\tau) = J_\tau(u_\tau, v_\tau) - u_\tau \cdot B_\tau^* \bar{y}_{\tau+1} - v_\tau \cdot C_\tau \bar{x}_{\tau-1}, \tag{4.1}$$

and

$(\bar{u}_e, \bar{v}_e)$ is a saddle point of

$$\bar{J}_e(u_e, v_e) = J_e(u_e, v_e) - u_e \cdot B_e^* \bar{y}_1 - v_e \cdot C_e \bar{x}_N. \tag{4.2}$$

Proof. The saddle point condition for $(\bar{u}, \bar{u}_e; \bar{v}, \bar{v}_e)$ consists of the two relations

$$(\bar{u}, \bar{u}_e) \in \operatorname*{argmin}_{(u, u_e)} \mathscr{J}(u, u_e; \bar{v}, \bar{v}_e), \tag{4.3}$$

$$(\bar{v}, \bar{v}_e) \in \underset{(v, v_e)}{\operatorname{argmax}} \mathscr{J}(\bar{u}, \bar{u}_e; v, v_e). \tag{4.4}$$

Expressing $\mathscr{J}(u, u_e, \bar{v}, \bar{v}_e)$ first in terms of (2.11), we see that (4.3) is equivalent to

$$\bar{u}_\tau \in \underset{u_\tau}{\operatorname{argmin}} \{J_\tau(u_\tau, \bar{v}_\tau) - u_\tau \cdot B_\tau^* \bar{y}_{\tau+1} - \bar{v}_\tau \cdot C_\tau \bar{x}_{\tau-1}\}, \tag{4.5}$$

for $\tau = 1, \ldots, N$, and

$$\bar{u}_e \in \underset{u_e}{\operatorname{argmin}} \{J_e(u_e, \bar{v}_e) - u_e \cdot B_e^* \bar{y}_1 - \bar{v}_e \cdot C_e \bar{x}_N\}. \tag{4.6}$$

Expressing $\mathscr{J}(u, u_e; v, v_e)$ then in terms of (2.15), we see on the other hand that (4.4) is equivalent to

$$\bar{v}_\tau \in \underset{v_\tau}{\operatorname{argmax}} \{J_\tau(\bar{u}_\tau, v_\tau) - \bar{u}_\tau \cdot B_\tau^* \bar{y}_{\tau+1} - v_\tau \cdot C_\tau \bar{x}_{\tau-1}\}, \tag{4.7}$$

for $\tau = 1, \ldots, N$, and

$$\bar{v}_e \in \underset{v_e}{\operatorname{argmin}} \{J_e(\bar{u}_e, v_e) - \bar{u}_e \cdot B_e^* \bar{y}_1 - v_e \cdot C_e \bar{x}_N\}. \tag{4.8}$$

The combination of (4.5) and (4.7) is equivalent to (4.1), while the combination of (4.6) and (4.8) is equivalent to (4.2). ∎

COROLLARY 4.4. *Consider relative to a given $(\bar{u}, \bar{u}_e; \bar{v}, \bar{v}_e)$ and the corresponding trajectories $\bar{x}$ and $\bar{y}$ the following subproblems:*

$(\bar{\mathscr{P}}_\tau)$ *minimize* $f_\tau(C_\tau \bar{x}_{\tau-1}, u_\tau) - u_\tau \cdot B_\tau^* \bar{y}_{\tau-1}$ *in* u_τ,

$(\bar{\mathscr{D}}_\tau)$ *maximize* $g_\tau(B_\tau^* \bar{y}_{\tau+1}, v_\tau) - v_\tau \cdot C_\tau \bar{x}_{\tau-1}$ *in* v_τ,

for $\tau = 1, \ldots, N$ *and also*

$(\bar{\mathscr{P}}_e)$ *minimize* $f_e(C_e \bar{x}_N, u_e) - u_e \cdot B_e^* \bar{y}_1$ *in* u_e,

$(\bar{\mathscr{D}}_e)$ *maximize* $g_e(B_e^* \bar{y}_N, v_e) - v_e \cdot C_e^* \bar{x}_1$ *in* v_e.

Then in order that $(\bar{u}, \bar{u}_e; \bar{v}, \bar{v}_e)$ be a saddle point of $\mathscr{J}$ the following conditions are necessary and sufficient:

$\bar{u}_\tau$ *is optimal for* $(\bar{\mathscr{P}}_\tau)$, $\bar{v}_\tau$ *is optimal for* $(\bar{\mathscr{D}}_\tau)$, *and* $\inf(\bar{\mathscr{P}}_\tau) = \sup(\bar{\mathscr{D}}_\tau)$ (4.9)

for $\tau = 1, \ldots, N$, *and*

$\bar{u}_e$ *is optimal for* $(\bar{\mathscr{P}}_e)$, $\bar{v}_e$ *is optimal for* $(\bar{\mathscr{D}}_e)$, *and* $\inf(\bar{\mathscr{P}}_e) = \sup(\bar{\mathscr{D}}_e)$. (4.10)

Proof. For the convex-concave function $\bar{J}_\tau$ in (4.1) we have by (2.4) that

$$f_\tau(C_\tau \bar{x}_{\tau-1}, u_\tau) - u_\tau \cdot B_\tau^* \bar{y}_{\tau+1} = \sup_{v_\tau} \bar{J}_\tau(u_\tau, v_\tau),$$

and by (2.16) that

$$g_\tau(B_\tau^* \bar{y}_{\tau+1}, v_\tau) - v_\tau \cdot C_\tau \bar{x}_{\tau-1} = \inf_{u_\tau} \bar{J}_\tau(u_\tau, v_\tau).$$

Therefore $(\bar{\mathscr{P}}_\tau)$ and $(\bar{\mathscr{D}}_\tau)$ are the primal and dual problems associated with $\bar{J}$. The saddle point condition (4.1) can thus be written equivalently as (4.9).

By a parallel argument, the saddle point condition (4.2) is equivalent to (4.10). ∎

EXAMPLE 4.5. In the ordinary convex programming model in Examples 1.1 and 2.2, the saddle point conditions in Theorem 4.3 reduce to the following. The pair $(\bar{u}_\tau, \bar{v}_\tau)$ is a saddle point of the expression

$$[f_{\tau 0}(u_\tau) - u_\tau \cdot B_\tau^* \bar{y}_{\tau+1}] + \sum_{i=1}^{l_\tau} v_{\tau i} [f_{\tau i}(u_\tau) - c_{\tau i} \cdot \bar{x}_{\tau-1}],$$

relative to $u_\tau \in U_\tau$ and $v_\tau \geqslant 0$, while the pair $(\bar{u}_e, \bar{v}_e)$ is a saddle point of the expression

$$[f_{e0}(u_e) - u_e \cdot B_e^* \bar{y}_1] + \sum_{i=1}^{l_e} v_{ei} [f_{ei}(u_e) - c_{ei} \cdot \bar{x}_N],$$

relative to $u_e \in U_e$ and $v_e \geqslant 0$. These conditions mean that $\bar{u}_\tau$ is an optimal solution to, and $\bar{v}_\tau$ a Kuhn-Tucker vector for, the problem

$(\bar{\mathscr{P}}_\tau)$ minimize $f_{\tau 0}(u_\tau) - u_\tau \cdot B_\tau^* \bar{y}_{\tau+1}$ subject to

$$u_\tau \in U_\tau \text{ and } f_{\tau i}(u_\tau) \leqslant c_{\tau i} \cdot \bar{x}_{\tau-1} \text{ for } i = 1, \ldots, l_\tau,$$

while $\bar{u}_e$ is an optimal solution to, and $\bar{v}_e$ a Kuhn-Tucker vector for, the problem

$(\bar{\mathscr{P}}_e)$ minimize $f_{e0}(u_e) - u_e \cdot B_e^* \bar{y}_1$ subject to

$$u_e \in U_e \text{ and } f_{ei}(u_e) \leqslant c_{ei} \cdot \bar{x}_N \text{ for } i = 1, \ldots, l_e.$$

EXAMPLE 4.6. The linear programming model in Examples 1.2 and 2.11 gives the subproblems

$(\bar{\mathscr{P}}_\tau)$ minimize $[p_\tau - B_\tau^* \bar{y}_{+1}] \cdot u_\tau$ subject to $u_\tau \geqslant 0$, $D_\tau u_\tau \geqslant [q_\tau - C_\tau \bar{x}_{\tau-1}]$,

$(\bar{\mathscr{D}}_\tau)$ maximize $[q_\tau - C_\tau \bar{x}_{\tau-1}] \cdot v_\tau$ subject to $v_\tau \geqslant 0$, $D_\tau^* v_\tau \leqslant [p_\tau - B_\tau^* \bar{y}_{\tau+1}]$,

and

$(\bar{\mathscr{P}}_\tau)$ minimize $[p_e - B_e^* \bar{u}_1] \cdot u_e$ subject to $u_e \geqslant 0$, $D_e u_e \geqslant [q_e - C_e \bar{x}_N]$,

$(\bar{\mathscr{D}}_\tau)$ maximize $[q_e - C_e \bar{x}_N] \cdot v_e$ subject to v_e, $D_e^* u_e \leqslant [p_e - B_e^* \bar{y}_1]$.

For such problems the linear programming duality theorem tells us that the equations $\inf(\bar{\mathscr{P}}_\tau) = \sup(\bar{\mathscr{D}}_\tau)$ in (4.9) and $\inf(\bar{\mathscr{P}}_e) = \sup(\bar{\mathscr{D}}_e)$, in (4.10) are redundant. Corollary 4.4 thus characterizes the optimal solutions $(\mathscr{P})$ in terms of optimal solutions to certain temporal linear programming problems $(\bar{\mathscr{P}}_\tau)$ and $(\bar{\mathscr{P}}_e)$ and their duals.

This pattern extends to problem models in piecewise linear-quadratic programming. See [4] for details.

REMARK 4.7. The patterns developed here show that not just linear-quadratic programming or ordinary convex programming fit this situation. Any primal--dual pair of problems in the literature of finite-dimensional convex optimization can be set up to appear as the subproblems of Corollary 4.4 and yield a corresponding version of ($\mathscr{P}$) and ($\mathscr{D}$).

References

[1] ROCKAFELLAR R. T. Convex Analysis. Princeton Univ. Press, 1970.

[2] ROCKAFELLAR R. T. Conjugate Duality and Optimization. C.B.M.S. Series No 16, SIAM Publications, Philadelphia 1974.

[3] ROCKAFELLAR R. T. Linear-quadratic programming and optimal control. *SIAM J. Control Opt.* (1987).

[4] ROCKAFELLAR R. T., WETS R. J.-B. Generalized linear-quadratic problems of deterministic stochastic optimal control in discrete time. Forthcoming.

[5] ROCKAFELLAR R. T., WETS R. J.-B. Deterministic and stochastic problems of Bolza in discrete time. Stochastics.

[6] GRINOLD R. Symmetric duality for continuous linear programming problem. *SIAM J. Appl. Math.* **18** (1970), 84–97.

[7] ROCKAFELLER R. T. Conjugate convex functions in optimal control and the calculus of variations. *J. Math. Anal. Appl.* **32** (1970), 174–222.

Received, January 1987.

Wieloetapowe programowanie wypukłe a sterowanie optymalne z czasem dyskretnym

W pracy przedstawiono nowe sformułowanie zadania optymalizacji wypukłej na skończonej liczbie etapów. Pozwala ono na potraktowanie zadania jako zadanie sterowania optymalnego a także wprowadza pewne cechy istotne dla programowania matematycznego. Warunki optymalności wyprowadzono z rozważań dotyczących zadania dualnego. Ze względu na właściwości sformułowania warunki optymalności przyjmują postać zasady „minimaksimum", która prowadzi do dekompozycji względem czasu zarówno zadania pierwotnego jak i dualnego.

Многоэтапное выпуклое программирование и дискретное оптимальное управление

В работе представлена новая формулировка задачи выпуклой оптимизации для конечного числа этапов. Она позваляет воспринимать проблему в виде задачи оптимального управления, а также вводит новые существенные факторы для математического программирования. Условия оптимальности выводятся из рассмотрений, касающихся дуальной задачи. Учитывая свойства формулировки условия оптимальности принимают вид принципа „минимакса", который ведет к декомпозиции по времени как первичной так и дуальной задачи.