Differentiability of the value function in continuous-time economic models[☆]

Juan Pablo Rincón-Zapatero [a,*], Manuel S. Santos [b]

[a] *Departamento de Economía, Universidad Carlos III de Madrid, Spain*
[b] *Department of Economics, University of Miami, United States*

A B S T R A C T

In this paper we provide some sufficient conditions for the differentiability of the value function in a class of infinite-horizon continuous-time models of convex optimization arising in economics. We dispense with the assumption of interior optimal paths. This assumption is quite unnatural in constrained optimization, and is usually hard to check in applications. The differentiability of the value function is used to prove Bellman's equation as well as the existence and continuity of the optimal feedback policy. We also establish the uniqueness of the vector of dual variables. These results become useful for the characterization and computation of optimal solutions.

1. Introduction

In this paper we study the differentiability of the value function for a class of concave infinite-horizon continuous-time problems with wide applications in economics. We extend the envelope theorem of Benveniste and Scheinkman [1] to optimization problems with constraints. We dispense with an interiority condition for the state and control variables that is usually quite restrictive in economic applications. This interiority condition may rule out periods of zero consumption, irreversibility of investment, bounded capacity, binding monetary constraints, and various financial market restrictions such as short-sale constraints and collateral requirements. Indeed, in his well-known introduction of control theory to economic growth, Arrow [2] formulated an economic problem with inequality constraints to account for feasibility, irreversibility, market clearing, and non-negative restrictions. There are usually no primitive assumptions that may prevent these constraints from being saturated, and hence one cannot generally invoke the envelope theorem of Benveniste and Scheinkman [1].

The differentiability of the value function is essential for the characterization and computation of optimal solutions. Indeed, in continuous time models the differentiability of the value function allows for a simple proof of Bellman's equation and the maximum principle. Hence, from the differentiability of the value function we obtain that the feedback control or policy is a continuous function. For finite-horizon problems, it is known (cf. [3]) that if the value function is differentiable then the path of dual variables or supporting prices is unique. We shall extend this uniqueness result for the infinite-horizon case.

Several papers deal with the existence of dual variables that belong to the superdifferential of the value function (e.g., [4–6]). Our focus here is on the uniqueness of these dual variables. Benveniste and Scheinkman [1] seems to be the

first paper to substantiate the differentiability of the value function for infinite-horizon continuous-time optimization. As in static models, the value function can be characterized as the envelope of short-run, concave and smooth functions. This argument relies on concavity of the objective and interiority of optimal solutions—see Assumption (IN) below. The envelope construction breaks down for boundary solutions. Indeed, in this latter case the derivative of the value function is computed as an infinite integral of derivatives over the optimal path whereas for interior solutions the derivative only depends on the marginal value at time zero. Therefore, for boundary solutions the differentiability of the value function cannot longer be addressed by methods of the kind found in purely static problems. For finite-horizon optimization, Goebel [3] proves that the value function is differentiable after assuming that the terminal, bequest function is differentiable. Of course, this proof cannot be extended to infinite-horizon problems: The dynamic programming method (see Lemma 3.1) implies that for every future terminal time the bequest function corresponds to the true value function. Hence, we still need to establish that this latter function is differentiable.

Viscosity solutions for the Hamilton–Jacobi–Bellman equation are usually quite helpful to study regularity properties of the value function. This elegant method can readily be extended to constrained optimization problems, but it imposes strict concavity of the Hamiltonian function with respect to the dual variables; this is a rather strong restriction for constrained optimization (cf. [7, Proposition 5.7] and specially Remark 5.8).

Let us also mention some other contributions in the economics literature for discrete-time optimization that seem to be of interest for potential extensions of our work to non-convexities (e.g., [8–11]). All these papers relax concavity of the optimization problem, but still demand interiority of optimal solutions. Amir et al. [8] and Amir [9] postulate some monotonicity and supermodularity conditions on the primitive functions. Askri and Le Van [10] extend the general theory of Clarke's gradients to the value function of a non-classical growth model, whereas Cotter and Park [11] consider one-dimensional optimization problems and develop a version of Danskin's theorem as introduced by Milgrom and Segal [12].

The starting point of our analysis is our earlier paper [13] on the differentiability of the value function in discrete-time optimization. The continuous-time formulation, however, is technically more involved and requires to make use of infinite-dimensional calculus. In both cases, we face the problem of the asymptotic behavior of an infinite sequence of derivatives. In our earlier paper [13], we mapped our optimization problem into a competitive economy that precludes existence of asset pricing bubbles (e.g., [14]). Here, we offer a more direct proof based on primitive assumptions. In spite of all technicalities associated with infinite-dimensional optimization, the continuous-time formulation offers more structure because the dynamical system that generates optimal trajectories is a flow: An optimal orbit is conformed by a continuous arc rather than by a countable number of points. This continuity property will be manifested in various stronger results. Theorem 3.2 shows that differentiability of the value function at the initial point x_0 implies differentiability of the function along the whole optimal trajectory, whereas this result is not guaranteed in the discrete-time formulation. Also, in the scalar case the value function is always differentiable at non-stationary points for the continuous-time case, but this is not generally true for discrete-time optimization.

In Section 2 we lay out the continuous-time optimization problem. Section 3 contains our main results on the differentiability of the value function. In Section 4 we apply these results to derive Bellman's equation and the uniqueness of the dual variables. Some examples follow in Section 5. A more technical review of our findings will be offered in Section 6. Various mathematical definitions can be found in the Appendix, as well as additional proofs.

2. The dynamic optimization problem

We consider an infinite-horizon optimization problem. We shall approximate this problem by a sequence of finite-horizon objectives. For finite horizons – rather than for the original optimization problem – we shall make use of a Banach space framework which will be analytically convenient for differentiability. The proof of differentiability of the value function will follow from a limit argument over finite horizons.

2.1. Mathematical setting

Let $t \geq 0$ be the initial date of the optimization problem. Let $I_t = [t, T]$, with $T = \infty$ or $T < \infty$. Let $\beta(s, t) = \exp\left(-\int_t^s \delta(r)\, dr\right)$ be a discount factor over the time interval $[t, s]$, $0 \leq t \leq s$. Function $\delta \geq 0$ is bounded with $\int_t^\infty \delta(r)\, dr = \infty$. Hence, $\beta(\infty, t) = 0$ for all t, and $\beta(t, t) = 1$. Assume that for each $r \in I_t$, there exists a constant $\rho > 0$ such that $\int_r^\infty \beta(s, t)\, ds \leq \rho \beta(r, t)$ for all $r > t$. If δ is a constant discount rate, then $\rho = (1/\delta)$ as $\int_r^\infty e^{-\delta(s-t)}\, ds = (1/\delta)e^{-\delta(r-t)}$.

Let μ_t be a measure on I_t with density $d\mu_t(s) = \beta(s, t)\, ds$. Then, $\mu_t(I_t) < \infty$ for all t. Let $L_n^1(I_t; \mu_t)$ be the set of equivalence classes of Lebesgue-measurable functions x_t in $\mathbb{R}^n$ such that $\int_{I_t} |x_t(s)|\, d\mu_t(s) < \infty$, where $|x_t(s)|$ is a given norm for $x_t(s)$. It follows that $L_n^1(I_t; \mu_t)$ is a Banach space with norm

$$\|x_t\|_{1, \mu_t} = \int_{I_t} |x_t(s)| \beta(s, t)\, ds.$$

Let $\mu_t^\top$ be a measure on I_t with density $d\mu_t^\top(s) = \beta(s, t)^{-1}\, ds = ds/\beta(s, t)$. The space $L_n^\infty(I_t; \mu_t^\top)$ consists of measurable functions p_t on I_t such that $|p_t(s)| \beta(s, t)^{-1}$ is bounded, except possibly on a set of measure zero. It is also a Banach space

with the norm

$$\|p_t\|_{\infty,\mu_t^\top} = \operatorname{ess\,sup}_{s \in I_t} |p_t(s)|\beta(s,t)^{-1} = \inf_{\substack{y(s)=p_t(s) \\ \text{Lebesgue-a.e.}}} \sup_{s \in I_t} |y(s)|\beta(s,t)^{-1}.$$

These two spaces conform a dual pair under the bilinear form

$$\langle x_t, p_t \rangle = \int_{I_t} x_t(s)p_t(s)\,ds, \quad x_t \in L_n^1(I_t; \mu_t), \ p_t \in L_n^\infty(I_t; \mu_t^\top).$$

In what follows, $\dot{x}_t(s)$ is the time derivative of function x_t at time s. Let $W^{1,1}(I_t)$ be the set of functions $x_t \in L_n^1(I_t; \mu_t)$ such that $\dot{x}_t$ exists μ_t a.e. and belong to $L_n^1(I_t; \mu_t)$ and let $W^{1,1}_{\text{loc}}([t, \infty))$ be the set of functions x_t that belong to $W^{1,1}(I_t)$ for every $T < \infty$.

2.2. Continuous-time optimization

The continuous-time optimization problem can now be posed as follows. Given an initial state x_0 and the initial date $t \geq 0$, find a path $x_t^* \in W^{1,1}_{\text{loc}}([t, \infty))$ solving the optimization program

$$V(t, x_0) = \sup \int_t^\infty \ell(x_t(s), \dot{x}_t(s))\beta(s,t)\,ds \tag{1}$$

subject to $(x_t(s), \dot{x}_t(s)) \in \Omega$ for all $s \in [t, \infty)$ and $x_t(t) = x_0$.

(A1) $X \subseteq \mathbb{R}^n$ and $\Omega \subseteq \mathbb{R}^{2n}$ are convex sets with nonempty interior. For each $x \in X$ the set $\Omega_x = \{u : (x, u) \in \Omega\}$ is nonempty.
(A2) Function $\ell : \Omega \longrightarrow \mathbb{R}$ is concave and differentiable of class C^1 in a neighborhood of Ω.

In some economic models, like those studied in this paper, function ℓ may have an unbounded gradient at some portions of the boundary of Ω. This is not a problem if the optimal solution never hits those boundary points, as it happens in the models we consider. Consequently, for the economic examples below we can include the following weak reformulation of Assumption (A2) which can readily be integrated into our main results. This is a standard assumption in economic theory [15].

(A2′) (i) $\ell : \Omega \longrightarrow \mathbb{R}$ is a concave and continuous mapping, and differentiable of class C^1 on the interior of Ω; (ii) Let $\mathcal{X} \subseteq \Omega$ be the set of boundary points in which the derivative of ℓ is not well defined. Assume that $(x_t^*, \dot{x}_t^*)$ is an optimal solution path. Then $\text{measure}\big(\{s \in I_t : (x_t^*(s), \dot{x}_t^*(s)) \cap \mathcal{X} \neq \emptyset\}\big) = 0$.

For instance, in our first example in Section 5, consumption and capital will never equal to zero if the marginal utility of consumption becomes unbounded at zero consumption.

(A3) Pick any $x_0 \in \text{int}\,X$ and $t \geq 0$. Then, there exists an optimal solution x_t^* to Problem (1) over the set $W^{1,1}_{\text{loc}}([t, \infty))$ with $x_t^*(t) = x_0$.

Existence of an optimal solution is guaranteed under various standard assumptions (cf. [16] and the Appendix below). We then have that the value function $V(t, \cdot)$ in (1) is well defined on $\text{int}\,X$. By Bellman's optimality principle, our strategy of proof is to consider the integral functional above over finite intervals $I_t = [t, T]$.

2.3. Some regularity conditions for differentiability of the value function

The following conditions will allow us to dispense with the interiority assumption of Benveniste and Scheinkman [1]. First, if x_t reaches the boundary of X then the value function V may not be differentiable. By backward induction, this lack of differentiability may extend over the optimal path. We therefore assume

(IS) An optimal path $x_t^*(s) \in \text{int}\,X$ for every $s \in I_t$.

Rincón-Zapatero and Santos [13] provide some examples of non-differentiability when (IS) fails. As shown below for continuous-time one-dimensional optimization this mild interiority requirement is generally not needed.

(LI) Ω can be defined by a finite set of inequalities

$$\Omega =: \{(x, u) : g^i(x, u) \geq 0 \text{ for } i = 1, \ldots, m\},$$

where functions g^i are C^1 in a neighborhood of Ω. Let $g_\sigma = \{g^i : g^i(x, u) = 0\}$. Then, matrix $D_2 g_\sigma(x_t^*(s), \dot{x}_t^*(s))$ has full rank over the optimal path $\{x_t^*(s), \dot{x}_t^*(s)\}$ for almost all $s \geq t$.

The notation is as follows: $D_1 g$ and $D_2 g$ are the Jacobian matrices of $g = (g^1, \ldots, g^m)$ with respect to x and $u = \dot{x}$, respectively. As is well-known, linear independence (LI) implies that matrix $(D_2 g_\sigma)^\top$ has a generalized right-inverse $D_2 g_\sigma^+$, and guarantees uniqueness of the Kuhn–Tucker multipliers in static differentiable programs. It is important to note that (LI)

3

entails that at least one control variable appears in every saturated constraint; for if not, one of the rows of matrix $D_2 g_\sigma$ is made up of zeros, violating the rank condition.

Let the $n \times n$-matrix $G(\sigma; x, u) = -(D_1 g_\sigma^\top D_2 g_\sigma^+)(x, u)$, with the convention that if no constraint is saturated at time s, then G is the null matrix. To shorten the notation we will write

$$G_t^*(s) = G_t(\sigma; x_t^*(s), \dot{x}_t^*(s)) = -(D_1 g_\sigma^\top D_2 g_\sigma^+)(x_t^*(s), \dot{x}_t^*(s)).$$

Assumption (LI) guarantees that matrix G_t^* is locally integrable.

In view of Assumption (IS), the smoothness of functions g^i is only necessary in a neighborhood of $\bigcup_{x \in \text{int} X}(\{x\} \times \Omega_x)$ rather than over the whole set Ω.

Under our strategy of proof, for boundary solutions we will need to rule out some explosive behavior of the derivatives of the value function. These derivatives will grow according to the linear homogeneous system of differential equations $\dot{z}(s) = z(s) G_t^*(s)$, see Theorem 3.2. Hence, we shall consider the associated fundamental matrix $\Phi_t(s)$ with $\Phi_t(t) = I_n$, where I_n is the identity matrix. That is, $\Phi_t(s)$ is the unique matrix satisfying $\dot{\Phi}_t(s) = \Phi_t(s) G_t^*(s)$ for every $s \geq t$ (a.e.). Moreover, the inverse $\Phi_t^{-1}(s)$ exists and $\dot{\Phi}_t^{-1}(s) = -G_t^*(s) \Phi_t^{-1}(s)$ (a.e.). As shown later, the existence of an optimal path $\{(x_t^*(s), \dot{x}_t^*(s))\}$ imposes certain restrictions on the discounted value of $\Phi_t(s)$. We consider below some regularity properties under which this discounted value goes to zero as s goes to ∞.

For the sake of comparison, we include the interiority assumption postulated by Benveniste and Scheinkman [1]. Let $\mathbb{B}$ denote the unit ball of $\mathbb{R}^n$.

(IN) *There exist an open and convex set $U \subset X$, an $\varepsilon > 0$, and a time $h > 0$, such that $\{(x_t^*(s), \dot{x}_t^*(s))\} + \varepsilon\mathbb{B} \subset \Omega$ for all $x_0 \in U$ and almost all $s \in [t, t+h]$.*

In other words, over some initial phase there exists an ε-neighborhood of the optimal path $\{(x_t^*(s), \dot{x}_t^*(s))\}$ that belongs to Ω.

3. Results

3.1. Mathematical preliminaries

We start with the following property for concave optimization problems (cf. [17, Proposition 4.3]). Here, E and F are Banach spaces, and $\partial v(x)$ is the superdifferential of a concave function v.

Proposition 3.1. *Let f be a proper concave function from $E \times F$ to $\mathbb{R} \cup \{-\infty\}$. Consider function $v : E \longrightarrow \mathbb{R} \cup \{-\infty\}$ defined by*

$$v(x) = \sup_{u \in F} f(x, u).$$

Assume that $\bar{u} \in F$ satisfies $v(x) = f(x, \bar{u})$. Then, the following conditions are equivalent:

$$q \in \partial v(x)$$
$$(q, 0) \in \partial f(x, \bar{u}).$$

Remark 3.1. Observe that in our model the value function v is a concave mapping on $X \subseteq \mathbb{R}^n$, and so the superdifferential is well defined at every $x \in \text{int} X$. Then, $\partial v(x) \neq \emptyset$ entails that $\partial f(x, \bar{u}) \neq \emptyset$ at every optimal $\bar{u}$. Therefore, existence of an interior optimal path $\{(x_t^*(s), \dot{x}_t^*(s))\}$ implies that the superdifferential of functional $J_{t,T}$ defined in Lemma 3.2 is nonempty. This is specially important in infinite-dimensional optimization problems, where the superdifferential of a concave functional may not be well defined. Second, $(q, 0) \in \partial f(x, \bar{u})$ if and only if $\bar{u} \in \arg\max f(x, u)$. Therefore, $q \in \partial v(x)$ is independent of the maximizer chosen as f is a concave function.

We now transform a problem with constraints into one of unrestricted maximization by incorporating the indicator function of the feasible set Ω into the integrand of problem (1). Let

$$\mathscr{L}(x, u) = \ell(x, u) - I_\Omega(x, u),$$

where $I_\Omega(x, u) = 0$ if $(x, u) \in \Omega$ and $+\infty$ otherwise.

Assumptions (A1)–(A3) imply that $\mathscr{L}$ is a proper, upper semicontinuous and concave function. Then, problem (1) can now be stated as

$$V(t, x_0) = \max \int_t^\infty \mathscr{L}(x_t(s), \dot{x}_t(s)) \beta(s, t) \, ds$$

subject to $x(t) = x_0$.

Let us rewrite the model in recursive form. This formulation is made possible by the semigroup property of the discount factor $\beta(T, s)\beta(s, t) = \beta(T, t)$ for every $t \leq s \leq T$, and the intertemporal separability of the objective and constraints.

Lemma 3.1 (*Bellman's Principle of Optimality*). *For every $t \leq T < \infty$, the value function can be written as*

$$V(t, x_0) = \max \left\{ \int_t^T \mathcal{L}(x_t(s), \dot{x}_t(s)) \beta(s, t) \, ds + \beta(T, t) V(T, x(T)) \right\}. \tag{2}$$

Moreover, the optimal solution of this finite-horizon problem is given by the optimal pair $\{(x_t^*(s), \dot{x}_t^*(s))\}$ to problem (1) over $[t, T]$.

Our first step is to compute the superdifferential of the integrand in (2) for $T < \infty$. Then, we provide a characterization of the superdifferential of the value function. Let $J_t : I_t \times [L_n^1(I_t; \mu_t)]^2 \longrightarrow \mathbb{R} \cup \{-\infty\}$ be given by

$$J_{t,T}(x_t, u_t) = \begin{cases} \int_t^T \mathcal{L}(x_t(s), u_t(s)) \beta(s, t) \, ds & \text{if } \mathcal{L}(x_t(s), u_t(s)) \in L_n^1(I_t; \mu_t), \\ -\infty & \text{otherwise.} \end{cases} \tag{3}$$

Lemma 3.2. *Function $J_{t,T}$ is proper, upper semicontinuous, and concave. Moreover,*

$$\partial J_{t,T}(x_t, u_t) = \left\{ (p_t, q_t) \in \left[L_n^\infty(I_t; \mu_t^\top) \right]^2 : -(p_t(s), q_t(s)) \in \beta(s, t) \partial \mathcal{L}(x_t(s), u_t(s)) \text{ a.e.} \right\}.$$

Proof. By (A1)–(A3) it is clear that function $J_{t,T}$ is proper, upper semicontinuous, and concave. The superdifferential of function $J_{t,T}$ follows from the characterization of the subdifferential of functionals defined by means of integrals provided in [18,19] and the established duality pairing; see the Appendix for further details. $\square$

The next lemma provides a characterization of the superdifferential $\partial V(t, \cdot)$ of function $x \mapsto V(t, x)$.

Lemma 3.3. *Let $x_0 \in \text{int} X$. Then, $q_0 \in \partial V(t, x_0)$ if and only if there exists $(p_t, q_t) \in L_n^\infty(I_t; \mu_t^\top) \times L_n^\infty(I_t; \mu_t^\top)$ and $\xi_{t,T} \in \partial V(T, x_t^*(T))$ such that*

$$q_0 = - \int_t^T p_t(s) \, ds + \beta(T, t) \xi_{t,T}$$

$$q_t(s) = - \int_s^T p_t(r) \, dr + \beta(T, t) \xi_{t,T}$$

$$-(p_t(s), q_t(s)) \in \beta(s, t) \partial \mathcal{L}(x_t^*(s), \dot{x}_t^*(s)) \quad \text{a.e. } t \leq s \leq T.$$

An immediate consequence of this lemma is the envelope theorem of Benveniste and Scheinkman [1], where for the above indicator function we have $I_\Omega(x, u) = 0$ over an ε-tube of the optimal path.

Theorem 3.1 (*Benveniste and Scheinkman [1]*). *Let (A1)–(A3) be satisfied. Assume that (IN) holds for some optimal solution $\{(x_t^*(s), \dot{x}_t^*(s))\}$. Then, the value function is differentiable at x_0 and the derivative*

$$DV(t, x_0) = -D_2 \ell(x_0, \dot{x}_t^*(t)).$$

Proof. By condition (IN) we get $\partial \mathcal{L}(x_t^*(s), \dot{x}_t^*(s)) = \partial \ell(x_t^*(s), \dot{x}_t^*(s))$ for $s \in [t, t + h]$. Then, by Lemma 3.3 the path $q_t(s)$ is absolutely continuous with $q_t(s) = -D_2 \ell(x_t^*(s), \dot{x}_t^*(s))$ a.e., $s \in [t, t + h]$. Hence,

$$q_0 = q_t(t) = \lim_{s \to t^+} \frac{1}{s - t} \int_t^s -D_2 \ell(x_t^*(r), \dot{x}_t^*(r)) \beta(r, t) \, dr$$

is unique. It follows from Proposition 3.1 that $\partial V(t, x_0)$ is singled-valued. Consequently, $V(t, \cdot)$ is differentiable at x_0. Moreover, (A2) implies that $q_t(t) = -D_2 \ell(x_t^*(t), \dot{x}_t^*(t))$. $\square$

3.2. Differentiability of the value function in constrained optimization

As in Assumption (LI), let $G_t^*(s) = G_t(\sigma(s); x_t^*(s), \dot{x}_t^*(s)) = -(D_1 g_\sigma^\top D_2 g_\sigma^+)(x_t^*(s), \dot{x}_t^*(s))$. Note that Assumptions (A1)–(A3), (IS) and (LI) will be in force for all our main results in this section. We begin with the following characterization of the superdifferential of value function V.

Proposition 3.2. *Let $x_0 \in \operatorname{int} X$, and $T < \infty$. Then, $q_0 \in \partial V(t, x_0)$ if and only if there exist $q_t \in L_n^\infty(I_t; \mu_t^\top)$, $-(\bar{p}_t(s), \bar{q}_t(s)) \in \beta(s, t)\partial\ell(x_t^*(s), \dot{x}_t^*(s))$ a.e., and $\xi_{t,T} \in \partial V(T, x_t^*(T))$ such that q_t is the unique absolutely continuous solution in $L_n^\infty(I_t; \mu_t^\top)$ of the linear differential system*

$$\dot{q}_t(s) = \bar{p}_t(s) + G_t^*(s)(\bar{q}_t(s) - q_t(s)), \tag{4}$$

with initial condition

$$q_0 = q_t(t) = -\int_{I_t} \bar{p}_t(s) + G_t^*(s)(\bar{q}_t(s) - q_t(s)) \, ds + \beta(T, t)\xi_{t,T}.$$

Proof. Observe that

$$\partial\mathscr{L}(x, u) = \partial\ell(x, u) - \partial I_\Omega(x, u) = \partial\ell(x, u) - N_\Omega(x, u), \tag{5}$$

where N_Ω is the normal cone of the convex set Ω [20]. By concavity, the normal cone to Ω at (x, u) is given by

$$-N_\Omega(x, u) = \left\{ \sum_{i \in \sigma(x, u)} \lambda^i(D_1 g^i(x, u), D_2 g^i(x, u)) + (z, 0) : \lambda^i \geq 0, \ z \in N_X(x) \right\},$$

where $i = 1, 2, \ldots, \sigma$ refers to those constraints which are saturated at (x, u), and $N_X(x)$ is the normal cone to X at $x \in X$. Note that $N_X(x_t^*(s)) = \{0\}$ because $x_t^*(s)$ is an interior point of X as asserted in Assumption (IS) above.[1]

By Lemma 3.3, we have that $q_0 \in \partial V(t, x_0)$ if and only if there exists $(p_t, q_t) \in [L_n^\infty(I_t; \mu_t^\top)]^2$ such that

$$q_0 = -\int_{I_t} p_t(s) \, ds + \beta(T, t)\xi_{t,T} \tag{6}$$

$$q_t(s) = -\int_{I_s} p_t(r) \, dr + \beta(T, t)\xi_{t,T} \tag{7}$$

$$-(p_t(s), q_t(s)) \in \beta(s, t)\partial\mathscr{L}(x_t^*(s), \dot{x}_t^*(s)) \quad \text{a.e.} \tag{8}$$

By (5) and (8), we can write $p_t = \bar{p}_t + \hat{p}_t$ and $q_t = \bar{q}_t + \hat{q}_t$, where $-(\bar{p}_t, \bar{q}_t) \in \beta(s, t)\partial\ell(x_t^*, \dot{x}_t^*)$ a.e., and $-(\hat{p}_t, \hat{q}_t) \in \beta(s, t)N_\Omega(x_t^*, \dot{x}_t^*)$ (a.e.). Thus, combining these equalities with the characterization of the normal cone $N_\Omega(x_t^*, \dot{x}_t^*)$, we obtain

$$\hat{p}_t(s) = \beta(s, t) \sum_{i \in \sigma(s)} \lambda_t^i(s) D_1 g^i(x_t^*(s), \dot{x}_t^*(s)),$$

$$\hat{q}_t(s) = \beta(s, t) \sum_{i \in \sigma(s)} \lambda_t^i(s) D_2 g^i(x_t^*(s), \dot{x}_t^*(s))$$

a.e., for some $\lambda_t^i(s) \geq 0$, $i = 1, \ldots, \sigma$. By (LI), we can then substitute out

$$\lambda_t(s) = \beta(s, t)^{-1} D_2 g_\sigma^+(x_t^*(s), \dot{x}_t^*(s)) \hat{q}_t(s),$$

so that

$$\hat{p}_t(s) = -G_t^*(s)\hat{q}_t(s) = G_t^*(s)(\bar{q}_t(s) - q_t(s)).$$

Plugging $\hat{p}_t(s)$ into (7) we obtain

$$q_t(s) = -\int_{I_s} \left(\bar{p}_t(r) + G_t^*(r)(\bar{q}_t(r) - q_t(r)) \right) dr + \beta(T, t)\xi_{t,T}. \tag{9}$$

Observe that $\dot{q}_t(s)$ exists a.e., and

$$\dot{q}_t(s) = \bar{p}_t(s) + G_t^*(s)(\bar{q}_t(r) - q_t(r)).$$

Obviously, q_t is absolutely continuous. $\square$

Remark 3.2. From Proposition 3.2 we observe that there is a diffeomorphism between the superdifferentials $\partial V(t, x_0)$ and $\partial V(T, x_t^*(T))$. That is, there exists only one function, $q_t(s, \cdot)$, joining q_0 with $\beta(T, t)\xi_{t,T}$. This is because ℓ is smooth and the saturated constraints satisfy (LI). The flow mapping linking points $q_0 \in \partial V(t, x_0)$ with points $\xi_{t,T} \in \partial V(T, x_t^*(T))$ is illustrated in Fig. 1.

---

[1] This makes clear the need for (IS). If x is not an interior point of X, then there could be infinitely many vectors in the normal cone $N_X(x)$. Uniqueness of q_t will ultimately lead to differentiability of the value function V as shown in Theorem 3.2.
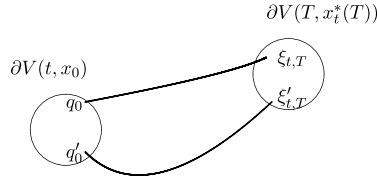
Fig. 1. The flow mapping between $\partial V(t, x_0)$ and $\partial V(T, x_t^*(T))$.

We are now ready to present our basic result on differentiability of the value function V under an additional asymptotic condition to be explained below. Let $\Delta_t(T)$ denote the diameter of compact set $\partial V(T, x_t^*(T))$, and $\|A\|$ some given norm for matrix A.

Theorem 3.2. *Let* $x_0 \in \text{int } X$. *Assume that*

$$\limsup_{T \to \infty} \beta(T, t)\|\Phi_t(T)\|\Delta_t(T) = 0. \tag{10}$$

Then, $V(t, \cdot)$ *is differentiable at* x_0, *and* $V(s, \cdot)$ *is also differentiable along the optimal trajectory* $\{x_t^*\}$ *from* x_0, *for every* $s \geq t$. *Furthermore, if*

$$\lim_{T \to \infty} \beta(T, t)\Phi_t(T)\xi_t(T) = 0 \tag{11}$$

for all $\xi_t(T) \in \partial V(T, x^*(T))$, *then the derivative* $DV(t, x_0)$ *is given by the expression*

$$DV(t, x_0) = \int_t^\infty \Phi_t(s)\big(D_1\ell(x_t^*(s), \dot{x}_t^*(s)) + G^*(s)D_2\ell(x_t^*(s), \dot{x}_t^*(s))\big)\beta(s, t)\, ds. \tag{12}$$

Proof. Let $q_t(s, q_0)$ be a solution of (4) with initial condition $q_t(t) = q_0 \in \partial V(t, x_0)$. Then, $q_t(s, q_0)$ is unique by Proposition 3.2 and Remark 3.2. As is well known from the theory of linear ODEs,

$$q_t(s, q_0) = \Phi_t^{-1}(s)q_0 + \Phi_t^{-1}(s)\int_t^s \Phi_t(r)(\bar{p}_t(r) + G^*(r)\bar{q}_t(r))\, dr, \tag{13}$$

where Φ_t is the associated fundamental matrix defined at the end of Section 2. Letting $s = T$ we can write q_0 as

$$q_0 = \Phi_t(T)q_t(T, q_0) - \int_t^T \Phi_t(r)(\bar{p}_t(r) + G^*(r)\bar{q}_t(r))\, dr.$$

It follows from Proposition 3.2 that $\beta(T, t)^{-1}q_t(T, q_0) = \xi_{t,T} \in \partial V(T, x_t^*(T))$. A similar representation is obtained for some other $q_0' \in \partial V(t, x_0)$ and $\xi_{t,T}' \in \partial V(T, x_t^*(T))$. Hence,

$$\begin{aligned}
|q_0 - q_0'| &\leq \|\Phi_t(T)\|\, |q_t(T, q_0) - q_t(T, q_0')| \\
&= \beta(T, t)\|\Phi_t(T)\|\, |\xi_{t,T} - \xi_{t,T}'| \\
&\leq \beta(T, t)\|\Phi_t(T)\|\Delta_t(T) \to 0, \quad \text{as } T \to \infty,
\end{aligned}$$

where convergence of this last term comes from (10). Therefore, $q_0 = q_0'$, and $\partial V(t, x_0)$ is a singleton, which implies that $V(t, \cdot)$ is differentiable at x_0. To show that $V(s, \cdot)$ is differentiable at $x_t^*(s)$, $s > t$, note that every element in $\partial V(s, x_t^*(s))$ is the image of some $q_t(s; q_0)\beta^{-1}(s, t)$. Then,

$$\bigcup_{q_0 \in \partial V(t, x_0)} \{q_t(s, q_0)\} = \beta(s, t)\partial V(s, x_t^*(s)) \tag{14}$$

for every $s \geq t$. By uniqueness of solutions to linear ODEs, $q_t(s, q_0)$ is unique since $\partial V(t, x_0)$ is a singleton. Therefore, $V(s, \cdot)$ is differentiable at $x_t^*(s)$.

The expression for the derivative (12) obtains from (13). More specifically, letting $s = T$, $q_t(T) = \beta(T, t)\xi_{t,T}$, and using (11), we get

$$q_0 = DV(t, x_0) = -\int_t^\infty \Phi_t(s)\big(\bar{p}_t(s) + G^*(s)\bar{q}_t(s)\big)\, ds,$$

as $T \to \infty$. Now, recall that $\bar{p}_t(s) = -\beta(s, t)D_1\ell(x_t^*(s), \dot{x}_t^*(s))$ and $\bar{q}_t(s) = -\beta(s, t)D_2\ell(x_t^*(s), \dot{x}_t^*(s))$. □

7

3.3. Condition (10)

This asymptotic condition only involves optimal solutions. We are now going to consider several regularity assumptions which guarantee that condition (10) is actually satisfied. We will also show that a slightly weaker version of (10) must be satisfied along an optimal path. Furthermore, condition (10) is not needed for one-dimensional optimization, and it holds vacuously for stationary solutions, and for solutions that eventually lie in the interior.

Our next two propositions apply to all admissible solutions. Hence, the maximal rank condition for matrix $G(x, u)$ in Assumption (LI) should be understood to apply for every $(x, u) \in \text{bd}(\Omega)$.

Proposition 3.3. *Assume that ℓ is a globally Lipschitz function on Ω. Assume that* (LI) *holds at every $(x, u) \in \text{bd}\,\Omega$. Let the following two conditions be satisfied for every admissible solution $\{(x_t(s), \dot{x}_t(s))\}$:*

1. $\lim_{T \to \infty} \beta(T, t) V(T, x_t(T)) = 0$;
2. *There exists an integrable function γ such that $\|G(x_t(s), \dot{x}_t(s))\| \leq \gamma(s)$ for $s \geq t$, and*

$$\int_t^\infty e^{\int_t^s (\gamma(r) - \delta(r))\, dr}\, ds < +\infty, \qquad \int_t^\infty \gamma(s) e^{\int_t^s (\gamma(r) - \delta(r))\, dr}\, ds < +\infty.$$

Then, condition (10) *must hold true.*

Condition 1 of this proposition is a familiar transversality condition which holds for a bounded solution V of the stationary Hamilton–Jacobi–Bellman equation. It also holds in more general environments, e.g., see Lemma A.4 in the Appendix. Condition 2 is closely connected with existence of an optimal solution in the infinite-horizon control problem. Consider for instance the following growth model (for a detailed exposition of the general model, see Section 5). Let $X = \mathbb{R}_+, \ell(x, u) = x - u, \Omega = \{(x, u)\ x \in X, 0 \leq u \leq x\}$, and a constant discount rate $\delta(s) = \delta \leq 1 = |G(x, u)|$. Let $x_0 \geq 1, t = 0$ and $x(s) = x_0 e^{\alpha s}$, where $0 < \alpha < \delta \leq 1$ is constant. Pick a solution $0 < \dot{x}(s) = \alpha x(s) < x(s)$. Hence, the pair $(x_0 e^{\alpha s}, \alpha x_0 e^{\alpha s}) \in \Omega$ for every $s \geq 0$, and the objective attains the following value

$$x_0 (1 - \alpha) \int_0^\infty e^{-\delta s} e^{\alpha s}\, ds = x_0 \frac{1 - \alpha}{\delta - \alpha}.$$

This value gets arbitrarily large as $\alpha \to \delta$. Therefore, the problem has no solution for any $\delta < 1$.

Proof. We first prove that function $V(t, \cdot)$ is globally Lipschitz continuous on X. For $x_0 \in X$, let $x_t(s, x_0)$ be an admissible trajectory satisfying $x_t(t, x_0) = x_0$. Let $x_1, x_2 \in X$ and $T \geq t$. Let $x_t^*(s, x_1)$ be an optimal trajectory from x_1, and let $x_t(s, x_2)$ refer to an admissible trajectory from x_2. Then, by Lemma A.2 in the Appendix we can pick $x_t(s, x_2)$ so that $|\dot{x}_t^*(s, x_1) - \dot{x}_t(s, x_2)| \leq \gamma(s) |x_t^*(s, x_1) - x_t(s, x_2)|$. Also, by the asserted Lipschitz condition on ℓ there exists a constant K such that

$$
\begin{aligned}
V(t, x_1) - V(t, x_2) &\leq \int_t^T \ell(x_t^*(s, x_1), \dot{x}_t^*(s, x_1)) - \ell(x_t(s, x_2), \dot{x}_t(s, x_2)) \beta(s, t)\, ds \\
&\quad + \beta(T, t)(V(T, x_t^*(T, x_1)) - V(T, x_t(T, x_2))) \\
&\leq K \int_t^T |x_t^*(s, x_1) - x_t(s, x_2)| + |\dot{x}_t^*(s, x_1) - \dot{x}_t(s, x_2)| \beta(s, t)\, ds \\
&\quad + \beta(T, t)(V(T, x_t^*(T, x_1)) - V(T, x_t(T, x_2))).
\end{aligned}
$$

Moreover, by Lemma A.3 in the Appendix we get

$$|x_t^*(s, x_1) - x_t(s, x_2)| \leq k e^{\int_t^s \gamma(r)\, dr},$$

for some constant k. Now, combining these inequalities it follows that

$$
\begin{aligned}
V(t, x_1) - V(t, x_2) &\leq K \left(\int_t^T (1 + k\gamma(s)) e^{\int_t^s (\gamma(r) - \delta(r))\, dr}\, ds \right) |x_1 - x_2| \\
&\quad + \beta(T, t)(V(T, x_t^*(T, x_1)) - V(T, x_t(T, x_2))).
\end{aligned}
$$

Exchanging the roles of x_1 and x_2, and letting $T \to \infty$, by condition 1 we obtain

$$|V(t, x_1) - V(t, x_2)| \leq K \left(\int_t^\infty (1 + k\gamma(s)) e^{\int_t^s (\gamma(r) - \delta(r))\, dr}\, ds \right) |x_1 - x_2|.$$

Hence, $V(T, \cdot)$ is a globally Lipschitz function, and so the diameter $\Delta_t(T)$ of the superdifferential is always bounded. Finally, observe that $\|\Phi_t(s)\| \leq e^{\int_t^s \gamma(r)\, dr}$ as

$$\|\dot{\Phi}_t(s)\| = \|G_t^*(s)\|\, \|\Phi_t(s)\| \leq \gamma(s) \|\Phi_t(s)\| \Rightarrow \|\Phi_t(s)\| \leq e^{\int_t^s \gamma(r)\, dr}.$$

Therefore, by condition 2 we then have that our asymptotic condition (10) holds true. $\square$

We can also establish a similar result without demanding Lipschitzianity of ℓ. In contrast to the previous approach, we incorporate some uniformity conditions and a standard monotonicity condition. Again, let $\mathbb{B}$ be the unit ball in $\mathbb{R}^n$.

(NB) (i) *The state space is* $X = \mathbb{R}^n_+$. *For* $x_0 \in \operatorname{int} X$ *the optimal trajectory* $\{x^*_t(s)\}$ *from* x_0 *belongs to a set* X' *such that* $X' + \varepsilon \mathbb{B} \subseteq \mathbb{R}^n_+$ *at a distance* $\geq \varepsilon$ *for some* $\varepsilon > 0$; (ii) *For all* $s \geq t$ *let* $\partial V(s, \cdot) \geq 0$; *let* $D_1 \ell(x^*_t(s), \dot{x}^*_t(s)) + G^*_t(s) D_2 \ell(x^*_t(s), \dot{x}^*_t(s)) \geq 0$ *over the optimal solution* $\{(x^*_t(s), \dot{x}^*_t(s))\}$.

Observe that the interiority requirement of Assumption (NB)(i) is a strengthening of Assumption (IS) since the orbit $\{x^*_t(s)\}$ must be uniformly separated from the boundary of $\mathbb{R}^n_+$.

Proposition 3.4. *Assume that the discount rate is a constant* $\delta > 0$ *so that* $\beta(s, t) = e^{-\delta(s-t)}$ *for all* $s > 0$. *Let* (NB)(i) *be satisfied. Assume that there are constants* $a, b \geq 0$ *such that*

$$|V(x)| \leq a|x| + b, \quad \forall x \in X. \tag{15}$$

Finally, let the following condition be satisfied: There exists an integrable function γ *such that for every admissible arc* $(x_t, \dot{x}_t) \in \operatorname{bd} X' \times \Gamma(X')$, *where* $X' + \varepsilon \mathbb{B} \subseteq \mathbb{R}^n_+$, *we have*

$$\|G(x_t(s), \dot{x}_t(s))\| \leq \gamma(s) \quad \forall s \geq t \tag{16}$$

with

$$\int_t^\infty e^{\int_t^s (\gamma(r) - \delta)\, dr} < +\infty. \tag{17}$$

Then, condition (10) *must hold true.*

Proof. Under the asserted conditions, it follows from Corollary A.1 in the Appendix that the time-homogeneous value function V is globally Lipschitz on X', and so the diameter $\Delta_t(T)$ of the superdifferential is always bounded on X'. Also, as in the proof above we have $\|\Phi_t(s)\| \leq e^{\int_t^s \gamma(r)\, dr}$. Moreover, since the optimal trajectory $\{x^*_t(s)\}$ belongs to X', by (16)–(17) it follows that condition (10) holds true. $\square$

As discussed in the Appendix, condition (15) can be obtained under very general assumptions. Finally, we can also show that a slightly weaker version of (10) is necessary for optimality.

Proposition 3.5. *Let Assumption* (NB)(ii) *hold. Then*

$$\limsup_{T \to \infty} \|\beta(T, t) \Phi_t(T) \xi_t(T)\| < \infty.$$

Proof. Observe that (NB)(ii) implies that every vector of dual variables $q_T \in \partial V(T, x_t(T, x_0))$ is non-negative. Then, by Proposition 3.2 for any $\xi_t(T) \in \partial V(T, x_t(T, x_0))$ there is some $q_0 \in \partial V(t, x_0)$ such that

$$q_0 \geq \beta(T, t) \Phi_t(T) \xi_t(T) \geq 0. \quad \square$$

3.4. Differentiability for the scalar case

In the one-dimensional case with a constant discount factor we have that differentiability is attained without Assumption (IS) and condition (10). In higher dimensions our argument below does not work, since an absolutely continuous curve has zero Lebesgue measure.

Corollary 3.1. *Let* $n = 1$ *and suppose that the discount rate* δ *is constant. Consider that* $x_0 \in \operatorname{int} X$ *is such that the optimal path* $x(s)$ *from* x_0 *satisfies* $\dot{x}^*_t(s) \neq 0$ *on some interval* $t \leq s \leq T$. *Then,* V *is differentiable at* x_0.

Proof. We argue by contradiction. First, note that the value function V is time-homogeneous, since the discount rate δ is constant. If V is not differentiable at x_0, then by Proposition 3.2 we get that V is not differentiable at $x(s)$ for any $s \geq t$ either. Hence, V is not differentiable in a set of positive Lebesgue measure, by assumption. This is in contradiction with the concavity of V, since a real concave function has at most countably many points of non-differentiability. $\square$

Actually, since the optimal trajectory x^*_t is absolutely continuous, it must be that the set $\{x(s) : t \leq s \leq T\}$ is a singleton if and only if $\dot{x}^*_t$ is zero over the interval $[t, T]$. Therefore, in the one-dimensional case with a constant discount rate, the value function is differentiable at all interior points of the state space, with the possible exception of stationary points. We study now the differentiability of the value function at stationary points for a general state space $X \subset \mathbb{R}^n$.

3.5. Differentiability at stationary points

By an optimal stationary point we mean a constant optimal solution $x^* = x^*_t(s)$ for almost all s, so that $\dot{x}^*_t(s) = 0$ for all s.

Corollary 3.2. *Assume that the discount rate δ is constant. Let $x^* \in \text{int}\, X$ be an optimal stationary point. Suppose that all coordinates of vector $D_1\ell(x^*, 0) + G(x^*, 0)D_2\ell(x^*, 0)$ are positive. Then, V is differentiable at x^*.*

Proof. Using Eq. (13) in Theorem 3.2 and the identity $q(T) = \beta(T, 0)\xi_{t,T}$ we know that $q_0 \in \partial V(x_0)$ if and only if for every T there exists $\xi_T \in \partial V(x^*(T))$ such that

$$q_0 = \int_0^T (D_1\ell(x^*(s), \dot{x}^*(s)) + G^*(s)D_2\ell(x^*(s), \dot{x}^*(s)))\beta(s, 0)\Phi(s)\, ds + \beta(T, 0)\Phi(T)\xi_T.$$

As x^* is a stationary point this equality reads

$$q_0 = \int_0^T (D_1\ell(x^*, 0) + G(x^*, 0)D_2\ell(x^*, 0))e^{(G(x^*,0)-\delta I_n)s}\, ds + e^{(G(x^*,0)-\delta I_n)T}\xi_T. \tag{18}$$

Note that now the fundamental matrix is $\Phi(s) = e^{G(x^*,0)s}$; moreover, both q_0, ξ_T belong to $\partial V(x^*)$ for any T, and by assumption, each component of vector $D_1\ell(x^*, 0) + G(x^*, 0)D_2\ell(x^*, 0)$ is strictly positive. Hence, $e^{(G(x^*,0)-\delta I_n)T}$ tends to the null matrix as $T \to \infty$. Therefore, V is differentiable at x^* because q_0 is univocally defined as

$$q_0 = \left(\int_0^\infty e^{(G(x^*,0)-\delta I_n)s}\, ds \right) (D_1\ell(x^*, 0) + G(x^*, 0)D_2\ell(x^*, 0)). \quad \square$$

Therefore, under strict monotonicity (cf. (NB)(ii)) this method of proof shows that condition (10) is vacuously satisfied for stationary solutions.

3.6. Some counterexamples

3.6.1. Necessity of Assumption (IS)
We will show the necessity of (IS) in a simple specification of the optimal growth model that will be studied in detail in Section 5. Consider $X = [0, \infty)$, a linear utility $U(c) = c$, a constant discount rate $\delta > 0$, and a linear production function $f(k) = \alpha k$ for some $\alpha > 0$ and k in $[0, 1]$. For $k \geq 1$, suppose that f is increasingly monotone, smooth, concave, and $\lim_{k \to \infty} f'(k) = 0$. According to Dmitruk and Kuz'kina [16, Theorem 1], the problem admits a solution for any discount rate $\delta > 0$; moreover, every trajectory is bounded.

For $0 < k_0 < 1$, consider the family of admissible trajectories $\dot{k}(s) = \alpha k(s)$ for $0 \leq s \leq T$, and $\dot{k}(s) = 0$ for $s \geq T$. Pick $T = -\frac{1}{\alpha} \ln k_0$; that is, $k(T) = k_0 e^{\alpha T} = 1$. By Lemma 3.1,

$$V(k_0) = \sup_{0 \leq \dot{k}_t \leq f(k_t)} \left\{ \int_0^T (f(k(s)) - \dot{k}(s))e^{-\delta s}\, ds + e^{-\delta T}V(k(T)) \right\}$$
$$\geq e^{-\delta T}V(k(T)) = e^{-\delta T}V(1) = V(1)k_0^{\delta/\alpha}.$$

The value function is continuous on X, with $V(k) > 0$ for any $k > 0$ and $V(0) = 0$. Hence, the above inequality determines that $\partial V(0) = \emptyset$ if $\alpha > \delta$.

3.6.2. Necessity of Assumption (LI)
Even in the scalar case, Assumption (LI) cannot be weakened. Consider the following problem

$$V(x_0) = -\max \int_0^\infty -x(t)e^{-\delta t}\, dt, \quad \delta > 0,$$

subject to the constraints: $\dot{x} \geq -2x$ and $\dot{x} \geq -\frac{1}{2}x$. This set of feasible choices Ω is depicted in Fig. 2. At point $x_0 = 0$ both constraints are saturated, thus (LI) does not hold since the problem is one-dimensional. In the region where $x > 0$ the smallest admissible derivative is $\dot{x} = -\frac{1}{2}x$. Hence, for $x_0 > 0$ the optimal path is $x(t) = x_0 e^{-t/2}$. It follows that $x(t) > 0$ for every t, since the stationary point $x_0 = 0$ is never reached in finite time. In the region where $x < 0$ we also need to pick the smallest admissible derivative because it is now positive. More precisely, $\dot{x} = -2x$. Hence, for $x_0 < 0$ the optimal path is $x(t) = x_0 e^{-2t} < 0$ for every t, which again converges to $x = 0$. Clearly, $x_0 = 0$ is an optimal stationary point.

Therefore, the value function

$$V(x_0) = \begin{cases} \dfrac{x_0}{2 + \delta}, & \text{if } x_0 < 0; \\ \dfrac{x_0}{\frac{1}{2} + \delta}, & \text{if } x_0 \geq 0. \end{cases}$$
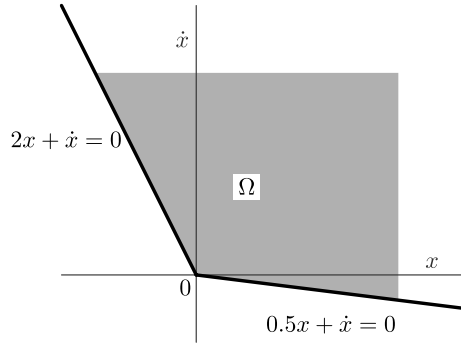
This function is not differentiable at $x_0 = 0$.

Fig. 2. A feasible set where (LI) does not hold.

4. Duality theory and Bellman's equation

We first show uniqueness of dual arcs satisfying a transversality condition. This uniqueness result easily follows from the differentiability of the value function and some properties of partial superdifferentials of saddle functions discussed in the Appendix. We also derive Bellman's equation and show the continuity of the optimal feedback control or policy function. Of course, if the policy function is continuous then the optimal solution $x_t^*(s)$ is a C^1 function of s.

Let us begin with the Hamiltonian associated with the optimization problem

$$H(x, q) = \sup_u \{\mathscr{L}(x, u) + qu\}. \tag{19}$$

Combining Lemma 3.3 with Proposition A.2 in the Appendix, an optimal solution $u = x_t^*$ must satisfy the Hamiltonian inclusions

$$\begin{aligned} -\dot{q}_t(s) &\in \beta(s, t)\partial_x H(x_t^*(s), q_t(s)), \\ \dot{x}_t^*(s) &\in \beta(s, t)\partial_q H(x_t^*(s), q_t(s)), \end{aligned} \tag{20}$$

for almost all $s \in [t, T]$. Here, $\partial_x H$ denotes the superdifferential of the concave function $x \mapsto H(x, q)$ for a fixed q, and $\partial_q H$ denotes the subdifferential of the convex function $q \mapsto H(x, q)$ for a fixed[2] x. If a pair (x_t^*, q_t) satisfies the Hamiltonian inclusions at all times, then we say that q_t is the dual variable. It has the interpretation of a shadow price.

Theorem 4.1. *Let the pair (x_t^*, q_t) satisfy the Hamiltonian inclusions* (20) *with $x_t^*(t) = x_0$. Assume that the following transversality condition holds[3]*

$$\lim_{T \to \infty} q_t(T)x_t^*(T) = 0. \tag{21}$$

Then, the path of dual variables $q_t(s)$ is unique.

Bellman's equation is a fundamental tool in solving dynamic programming problems. As is well known, Bellman's equation requires some smoothness of the value function; moreover, the optimal policy correspondence is obtained as the arg max of this equation. Therefore, the differentiability of the value function is helpful for the existence and numerical solution of Bellman's equation. Let us rewrite (19) as

$$H(x, q) = \sup_{u \in \Omega_x} \{\ell(x, u) + qu\}.$$

Assuming a constant discount rate: $\delta(s) = \delta$ for every s, we get Bellman's equation as

$$-\delta V(x) + H(x, DV(x)) = 0 \quad \text{for all } x \in \text{int } X.$$

That is,

$$-\delta V(x) + H(x, DV(x)) = -\delta V(x) + \sup_{u \in \Omega_x} \{\ell(x, u) + DV(x)u\} = 0 \quad \text{for all } x \in \text{int } X.$$

Let us define the optimal policy correspondence $u \in h(x) = \partial_q H(x, DV(x))$. This is the set of admissible values of $u \in \Omega_x$ that solves $\max_{u \in \Omega_x} \{\ell(x, u) + qu\}$.

---

[2] For the concavity of the function $x \mapsto H(x, q)$ and the convexity of the function $q \mapsto H(x, q)$, see [20, Chapter VII].

[3] As is well known (cf., [6]) Assumption (NB)(ii) implies (21) along an optimal solution.
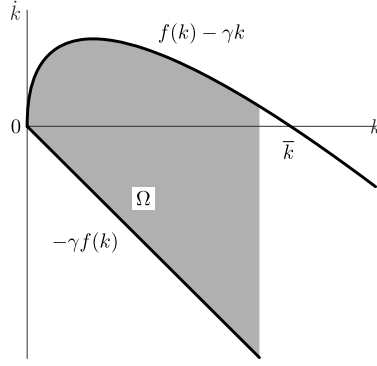
Fig. 3. Feasible set Ω in the optimal growth model.

Proposition 4.1. *Assume that the multivalued mapping $x \rightrightarrows \Omega_x$ is continuous and that Ω_x is a compact set for every $x \in X$. Assume that ℓ is strictly concave with respect to u. Then, the optimal $\dot{x}_t^*$ is given by a continuous function $\dot{x}_t^* = h(x_t)$ for all $x_t \in \text{int } X$, where $h(x) = \partial_q H(x, DV(x))$.*

Proof. Since V is differentiable on int X, function $(x, u) \mapsto \ell(x, u) + DV(x)u$ is continuous. Hence, by Berge's Theorem, h is upper hemicontinuous. Moreover, by the strict concavity of ℓ in u, the maximizer $h(x)$ is unique, and thus h is a continuous function. Finally, the expression $h(x) = \partial_q H(x, DV(x))$ follows from the first-order condition. $\square$

5. Examples

5.1. The one-sector growth model with irreversible investment

Consider the following version of the neoclassical growth model:

$$\max_{c_t(s), i_t(s)} \int_t^\infty U(c_t(s))\beta(s, t) \, ds \text{ subject to}$$

$$\dot{k}_t(s) = i_t(s) - \gamma k_t(s),$$

$$c_t(s) + i_t(s) = f(k_t(s)),$$

$$k_t(s) \geq 0, \qquad c_t(s) \geq 0, \qquad i_t(s) \geq 0, \qquad k_t(t) = k_0.$$

The notation is as follows: $k_t(s)$ is capital at time s, $c_t(s)$ is consumption, and $i_t(s)$ is investment. The utility function, $U : \mathbb{R}_+ \longrightarrow \mathbb{R}$, is increasing, concave, differentiable over $[0, \infty)$ with $U'(0^+) < +\infty$ or $U'(0^+) = +\infty$. The production function, $f : \mathbb{R}_+ \longrightarrow \mathbb{R}_+$, is bounded, increasing, concave, and differentiable in $[0, \infty)$ with $f'(0^+) = +\infty$.

As is well understood, the problem can be mapped into variables $(k_t, \dot{k}_t)$ corresponding to our original framework:

$$\max_{k_t(s), \dot{k}_t(s)} \int_t^\infty U(f(k_t(s)) - \gamma k_t(s) - \dot{k}_t(s))\beta(s, t) \, ds \text{ subject to}$$

$$-\gamma k_t(s) \leq \dot{k}_t(s) \leq f(k_t(s)) - \gamma k_t(s), \quad k_t(s) \geq 0.$$

Then, the instantaneous utility function is $\ell(k, u) = U(f(k) - \gamma k - u)$ with derivatives

$$D_1\ell(x, u) = U'(c)(f'(k) - \gamma), \qquad D_2\ell(k, u) = -U'(c).$$

The constraints are $g^1(k, u) = u + \gamma k$, $g^2(k, u) = f(k) - \gamma k - u$. The feasible set is depicted in Fig. 3.

It follows that

$$G(\{1\}; k, u) = -D_1 g^1(k, u) D_2 g^1(k, u) = -\gamma,$$

$$G(\{2\}; k, u) = -D_1 g^2(k, u) D_2 g^2(k, u) = f'(k) - \gamma.$$

Note that both constraints cannot be binding at the same time. Therefore,

$$G_t^*(s) = \begin{cases} -\gamma, & \text{if } \sigma = \{1\}; \\ f'(k_t^*(s)) - \gamma, & \text{if } \sigma = \{2\}, \end{cases}$$

and $\Phi_t(s) = e^{\int_t^s G_t^*(r) \, dr}$.

We are now ready to check that all our regularity conditions are generally satisfied. First, let us show that Assumption (IS) holds. If the optimal solution is at the boundary with $\sigma = \{1\}$, then it decreases at a constant rate so that $k_t^*(s) =$

$k_0 e^{-\gamma(s-t)} > 0$ for every $s \geq t$, and never hits 0. If the optimal solution is at the boundary with $\sigma = \{2\}$, then $\dot{k}_t^*$ is positive around $k = 0$. Moreover, an optimal solution with $\sigma = \{2\}$ for every $s \geq t$ cannot be possible, since it implies zero consumption for the optimal solution at all times. Hence, (10) is vacuously satisfied. For $\sigma = \{1\}$, we should note that $e^{\int_t^T (-\delta + G_t^*(s)) \, ds} = e^{(T-t)(-\delta-\gamma)} \to 0$ as $T \to \infty$. Again, (10) is vacuously satisfied in the case of $\sigma = \{1\}$ because the restriction $f'(0^+) = +\infty$ implies that the optimal capital stock $k_t^*(s)$ will never be arbitrarily close to zero. Actually, this model satisfies all the conditions postulated in (NB) above. Finally, as seen above (LI) holds trivially since an optimal solution cannot be at both extremes of the boundary at the same time.

We have then proved the following

Proposition 5.1. *In the one-sector growth model with irreversible investment the value function is differentiable at interior points. Moreover, the derivative is*

$$DV(t, k_0) = \int_t^\infty e^{\int_t^s (G_t^*(r) - \delta(r)) \, dr} U'(c_t^*(s))(f'(k_t^*(s)) - \gamma - G_t^*(s)) \, ds,$$

where $G_t^* = 0$ if the optimal arc lies in the interior of correspondence Ω.

Note that the envelope theorem of Benveniste and Scheinkman [1] cannot be invoked for cases in which some constraint could be binding. The irreversibility assumption may bind if capital is high enough, and zero consumption may be obtained if capital is low enough. Clearly, for a constant discount rate $\delta > 0$, differentiability of the value function follows from our above results for the scalar case.

5.2. A monetary economy

Consider the following cash-in-advance model

$$\max_{(c_t(s), m_t(s), k_t(s), \dot{k}_t(s))} \int_t^\infty U(c_t(s))\beta(s, t) \, ds \text{ subject to}$$

$$\dot{k}_t(s) + \dot{m}_t(s) = f(k_t(s)) - \gamma k_t(s) - c_t(s) + x_t(s) - \pi_t(s)m_t(s),$$

$$m_t(s) \geq c_t(s) + \dot{k}_t(s) + \gamma k_t(s),$$

$$k_t(s) \geq 0, \qquad c_t(s) \geq 0.$$

Here, c_t is consumption, m_t is a stock of real monetary holdings, k_t is capital, x_t is the value of government transfers rebated to the consumer as a consequence of the inflation tax, and π_t is the rate of inflation. Both U and f satisfy the same properties as in the previous example. For simplicity, the cash-in-advance constraint $m_t \geq c_t + \dot{k}_t(s) + \gamma k_t(s)$ applies to purchases of both the consumption good and gross investment.

Let us rewrite this problem in terms of the state variables (k, m). Then, the instantaneous objective is rewritten as:

$$\ell((k, m), (\dot{k}, \dot{m})) = U(f(k) - \gamma k + x - \pi m - \dot{k} - \dot{m}),$$

and the constraints:

$$g^1((k, m), (\dot{k}, \dot{m})) = f(k) - \gamma k + x - \pi m - \dot{k} - \dot{m} \geq 0, \quad \text{(non-negative consumption);}$$

$$g^2((k, m), (\dot{k}, \dot{m})) = \gamma k + \dot{k} \geq 0, \quad \text{(irreversible investment);}$$

$$g^3((k, m), (\dot{k}, \dot{m})) = m + \dot{m} - f(k) - x + \pi m \geq 0, \quad \text{(cash-in-advance).}$$

We are therefore confronted with a two-dimensional problem. As in the growth model, the pure state constraint $k \geq 0$ is not binding, as $f'(0^+) = \infty$. Thus, optimal trajectories (k_t^*, m_t^*) lie in the interior of the state space $X = \mathbb{R}_+^2$, and (IS) is satisfied. In order to check (LI) we consider Jacobian matrices $D_2(g^1, g^2)$, $D_2(g^1, g^3)$, $D_2(g^2, g^3)$ and $D_2(g^1, g^2, g^3)$ and verify the full-rank assumption. Of course, if only one constraint is saturated, then (LI) follows trivially. All matrices

$$D_2(g^1, g^2) = \begin{pmatrix} -1 & -1 \\ 1 & 0 \end{pmatrix}, \qquad D_2(g^1, g^3) = \begin{pmatrix} -1 & -1 \\ 0 & 1 \end{pmatrix}, \qquad D_2(g^2, g^3) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

have maximal rank. The three constraints (g^1, g^2, g^3) can only be binding for zero money holdings, $m = 0$. This case has been ruled out. Therefore, (LI) is always satisfied.

In order to check asymptotic condition (10), from our arguments in the previous example we know that there are periods in which constraints g^1 (zero consumption) and g^2 (irreversible investment) will not be saturated. Hence, let us focus on the simple case in which only g^3 (cash-in-advance) is binding for all $s \geq t$. Then, $G_t(\{3\}; ((m), (\dot{m}))) = -D_1(g^3)^\top D_2(g^3)^+ = -(1 + \pi)$. Therefore, $\Phi_t(T)e^{-\delta T} = e^{\int_t^T (G_t^*(r) - \delta) \, dr} = e^{\int_t^T (-1 - \pi - \delta) \, dr}$. Of course, this expression goes to zero, and hence (10) will always hold whenever the set of optimal solutions (k, m) remains in a compact set separated from the boundary of $\mathbb{R}_+^2$.

Observe that our asymptotic condition (10) should not be confused with transversality condition (21). The transversality condition is about asymptotic values (i.e., price times quantity), whereas (10) is about asymptotic shadow prices for constraints that are always binding. For instance, in the literature of the optimum quantity of money, it is well known that there are no steady states for $\pi > -\delta$. For our asymptotic condition (10) the requirement is simply $\pi > -1 - \delta$. Further, (10) is vacuously satisfied for time intervals in which none of the constraints is saturated.

6. Concluding remarks

This paper contains several results on the differentiability of the value function for a class of infinite-horizon continuous-time optimization problems with saturated constraints. One main goal of our exercise is to dispense with the interiority condition of Benveniste and Scheinkman [1]. We additionally show that the path of dual variables is unique, and derive a version of Bellman's equation for constrained optimization so that the feedback control or policy function is a continuous mapping. Therefore, the differentiability of the value function is essential for the characterization and computation of optimal solutions.

As illustrated in our examples above, there are many economic models with saturated constraints that violate the interiority condition of Benveniste and Scheinkman [1]. To circumvent this interiority condition, we postulate three additional assumptions which seem indispensable. First, the path of state variables must lie in the interior of the domain; for if not, the superdifferential of the value function may be multivalued or may be undefined. Second, as in the static case we require a linear independence assumption on the saturated constraints. And third, we rule out explosive behavior of the derivatives of the value function. The existence of an optimal path already imposes some restrictions on the dynamics, since the derivatives cannot grow faster than the discount factor. Moreover, we provide some mild regularity conditions on the optimization problem that imply our asymptotic condition.

The analysis presents several differences with respect to the discrete-time case considered in our previous paper [13]. In discrete-time, Bellman's equation is guaranteed under general assumptions. (For instance, this equation holds for bounded, non-continuous objective functions.) In continuous-time, we need certain smoothness conditions to write down Bellman's equation. Furthermore, iterations must proceed over time intervals rather than over simple dates as every time t has measure zero. Hence, the continuous-time problem requires the use of infinite-dimensional calculus. We transform a problem with constraints into one of unconstrained optimization, and build the analysis over finite-horizon optimization problems in a Banach-space setting. We characterize the superdifferential of the value function at time $t = 0$ as a sum of the superdifferential of the value function at every time $T > 0$ and an integral of derivatives of the return function and constrains over the interval $[0, T]$. Then, our asymptotic condition implies that the discounted value of the superdifferential of the value function at time T converges to zero as T goes to infinity.

As already remarked, the continuous-time formulation offers more structure than the discrete-time counterpart, since an optimal trajectory is conformed by a continuous arc rather than by a sequence of countable points. This continuity property is actually manifested in stronger results and sharper examples. For instance, for one-dimensional optimization the value function is differentiable under general conditions in the continuous-time case. Also, as illustrated in several examples above the assumptions are usually easier to check in applications: For continuous arcs it is simpler to track down points of switching binding constraints.

Appendix

For a given Banach space E and its dual $E^\top$, let $\langle \cdot, \cdot \rangle$ be the associated bilinear form over $E \times E^\top$. That is, for fixed $x \in E$ mapping $\langle x, \cdot \rangle$ defines a continuous linear functional on $E^\top$ and for fixed $p \in E^\top$ mapping $\langle \cdot, p \rangle$ defines a continuous linear functional on E.

For a bounded linear mapping $A : E \longrightarrow F$ between Banach spaces E and F, with dual spaces $E^\top$ and $F^\top$, respectively, the adjoint is the unique linear mapping $A^\top : F^\top \longrightarrow E^\top$ satisfying

$$\langle x, A^\top p \rangle = \langle Ax, p \rangle, \quad \forall x \in E, \ \forall p \in F^\top.$$

Let us now recall some basic definitions from convex analysis. Assume that $f : F \longrightarrow \mathbb{R} \cup \{\infty\}$ is an upper semicontinuous, concave function. Then, the effective domain of f is $\text{dom} f = \{x \in F : f(x) < \infty\}$. Function f is called proper if $\text{dom} f \neq \emptyset$. The set

$$\partial f(x) = \{p \in F^\top : \langle x - x', p \rangle \leq f(x) - f(x') \ \forall x' \in F\}$$

is the superdifferential of function f at x. An element $p \in \partial f(x)$ is called a supergradient of f at x. Let $\text{dom} \, \partial f = \{x \in F : \partial f(x) \neq \emptyset\}$. The superdifferential of f is always well defined at interior points of $\text{dom} f$, that is, $\text{int} \, \text{dom} f \subseteq \text{dom} \, \partial f$.

Let $A : E \longrightarrow F$ be a continuous linear operator. Assume that there is $x \in E$ such that $A(x) \in \text{int} \, \text{dom} f$. Then, the following equality holds, see [22, Proposition 5.7]:

$$\partial (f \circ A)(x) = (A^\top \circ \partial f)(A(x)). \tag{22}$$

Let us then introduce the families of linear mappings $A_t : \mathbb{R}^n \times L^1_n(I_t; \mu_t) \longrightarrow \left[L^1_n(I_t; \mu_t)\right]^2$:

$$A_t(x_0, u_t) = (x_t, u_t), \quad \text{where } x_t(s) = x_0 + \int_t^s u_t(r) \, dr \tag{23}$$

and $B_t : \mathbb{R}^n \times L^1_n(I_t; \mu_t) \longrightarrow \mathbb{R}^n$:

$$B_t(x_0, u_t) = x_0 + \int_t^T u_t(r) \, dr. \tag{24}$$

Proposition A.1. 1. *Operator A_t is linear and continuous. Its adjoint*

$$A_t^\top : \left[L_n^\infty(I_t; \mu_t^\top) \right]^2 \longrightarrow \mathbb{R}^n \times L_n^\infty(I_t; \mu_t^\top)$$

is defined as

$$A_t^\top(p_t, q_t) = \left(\int_{I_t} p_t(s)\, ds, \int_{I_s} p_t(r)\, dr + q_t \right).$$

2. *Operator B_t is linear and continuous. Its adjoint*

$$B_t^\top : \mathbb{R}^n \longrightarrow \mathbb{R}^n \times L_n^\infty(I_t; \mu_t^\top)$$

is defined as

$$B_t^\top(y_0) = (y_0, y_0).$$

Proof. 1. Obviously, A_t is linear. Let us show that it is well defined and continuous. We have

$$\int_{I_t} |x_t(s)| \beta(s,t)\, ds \le |x_0| \mu_t(I_t) + \int_{I_t} \beta(s,t) \int_t^s |u_t(r)|\, dr.$$

By an application of Fubini's theorem to the second term in the right-hand side we get

$$\int_{I_t} \beta(s,t) \int_t^s |u_t(r)|\, dr ds = \int_{I_t} |u_t(r)| \int_{I_r} \beta(s,t)\, ds\, dr$$

$$\le \rho \int_{I_t} |u_t(r)| \beta(r,t)\, dr < \infty,$$

since $u \in L_n^1(I_t; \mu_t)$, and by assumption $\int_r^\infty \beta(s,t)\, ds \le \rho\beta(r,t)$. It is easy to prove from these inequalities that the mapping is continuous.

To find the adjoint $A_t^\top$, consider $(x_0, u) \in \mathbb{R}^n \times L_n^1(I_t; \mu_t)$ and $(p_t, q_t) \in [L_n^\infty(I_t; \mu_t^\top)]^2$. Then, using the duality pairings

$$\langle A_t(x_0, u), (p_t, q_t) \rangle = \left\langle x_0 + \int_t^s u(r)\, dr, p_t \right\rangle + \langle u, q_t \rangle$$

$$= x_0 \int_{I_t} p_t(s)\, ds + \int_{I_t} \left(\int_t^s u(r)\, dr \right) p_t(s)\, ds + \langle u, q_t \rangle.$$

Changing the order of integration in the second summand and applying Fubini's Theorem, we find

$$\langle A_t(x_0, u), (p_t, q_t) \rangle = x_0 \int_{I_t} p_t(s)\, ds + \int_t^T u(s) \int_{I_s} p_t(r)\, dr\, ds + \langle u, q_t \rangle$$

$$= \left\langle x_0, \int_{I_t} p_t(s)\, ds \right\rangle + \left\langle u, \int_{I_s} p_t(r)\, dr \right\rangle + \langle u, q_t \rangle$$

$$= \langle (x_0, u), A_t^\top(p_t, q_t) \rangle.$$

The result for $A_t^\top$ is thus established.

2. Linearity and continuity of B_t is proved similarly. Moreover, by related computations we get

$$\langle B_t(x_0, u), y_0 \rangle = \langle x_0, y_0 \rangle + \left\langle \int_t^T u(s)\, ds, y_0 \right\rangle = \langle (x_0, u), B_t^\top(y_0) \rangle.$$

Hence, $B_t^\top(y_0) = (y_0, y_0)$. □

Now, for $T > t$ let $\mathcal{J}_{t,T} : \mathbb{R}^n \times L_n^1(I_t; \mu_t) \longrightarrow \mathbb{R} \cup \{-\infty\}$ be defined as in (3). That is,

$$\mathcal{J}_{t,T}(x_0, u) = J_{t,T}(A_t(x_0, u)) + \beta(T,t)V(T, B_t(x_0, u)).$$

It follows from Lemma 3.1 that the value function

$$V(t, x_0) = \sup_{u \in L_n^1(I_t; \mu_t)} \mathcal{J}_{t,T}(x_0, u). \tag{25}$$

By assumptions (A1)–(A3), mapping $V(t, \cdot)$ is well defined and concave over $\mathrm{int}\,X$ at each t, and $\partial V(T, x_t^*(T))$ is not empty for every T.

The following lemma characterizes the superdifferential of $\mathcal{J}_{t,T}$. In the sequel, $p_t(I_s)$ will denote $\int_s^T p_t(r)\,dr$.

Lemma A.1. *Assume that J_t is well-defined in a neighborhood of a feasible solution $A_t(x_0, u)$ with $x_t(s) \in \mathrm{int}\,X$ for all $s \geq t$, and V is well defined in a neighborhood of $x_t(T)$. Then,*

$$\partial \mathcal{J}_{t,T}(x_0, u) = \Big\{ \big(-p_t(I_t) + \beta(s,t)\xi_{t,T}, \, -p_t(I_s) - q_t + \beta(s,t)\xi_{t,T}\big) :$$
$$-(p_t(s), q_t(s)) \in \beta(s,t)\partial(\mathcal{L} \circ A_t)(x_0, u), \; \xi_{t,T} \in \partial V(T, x(T)) \; a.e. \Big\}.$$

Proof. By the concavity of these functions, we must have

$$\partial \mathcal{J}_{t,T} = \partial(J_{t,T} \circ A_t) + \beta(T,t)\partial(V(T, \cdot) \circ B_t).$$

Also, by (22)

$$\partial(J_{t,T} \circ A_t) = A_t^\top \circ \partial J_{t,T} \circ A_t$$

and

$$\partial(V(T, \cdot) \circ B_t) = B_t^\top \circ \partial V(T, \cdot) \circ B_t.$$

Combining Lemma 3.2 and Proposition A.1, an element of $A_t^\top(\partial J_{t,T}(A_t(x_0, u)))$ must be of the form $(-p_t(I_t), -p_t(I_s) - q_t)$, with $-(p_t(s), q_t(s)) \in \beta(s,t)\partial\mathcal{L}\big(A_t(x_0, u)\big)$, as well as a typical element of the set $\beta(T,t)B_t^\top(\partial V(T, B_t(x_0, u)))$ must be of the form $\beta(T,t)(\xi_{t,T}, \xi_{t,T})$ with $\xi_{t,T} \in \partial V(T, x(T))$. □

Proof of Lemma 3.3. Note that at the optimal solution $A_t(x_0, \dot{x}_t^*)$ all the conditions of Lemma A.1 are satisfied. By Proposition 3.1 we then have $q_0 \in \partial V(t, x_0)$ if and only if $(q_0, 0) \in \partial \mathcal{J}_{t,T}(x_0, \dot{x}_t^*)$. Now, the proof follows as a straightforward consequence of the above characterizations of the subdifferential of $\mathcal{J}_{t,T}$ at $(x_0, \dot{x}_t^*)$.

More precisely, by Lemma A.1 we must have

$$q_t(s) = -\int_s^T p_t(r)\,dr + \beta(T,s)\xi_{t,T}$$

with

$$-(p_t(r), q_t(r)) \in \beta(r,t)\partial\mathcal{L}(x_t^*(r), \dot{x}_t^*(r)) \quad \text{a.e. } t \leq r \leq T. \quad □$$

The following result is used in the proof of Proposition 3.3. It states that $\Gamma(x) = \Omega_x$ enjoys a kind of Lipschitz property.

Lemma A.2. *For all $x, x' \in X$ and $u \in \Gamma(x)$, there exists $u' \in \Gamma(x')$ such that*

$$|u - u'| \leq \|G(x', u')\| \, |x - x'|.$$

Proof. Consider function $\eta(x') = \min\{|u - u'| : u' \in \Gamma(x')\}$. There is no restriction of generality to assume that $u \notin \Gamma(x')$. The set $\Gamma(x')$ is closed and convex; thus, the minimum is attained at a unique point, $u' \in \mathrm{bd}\,\Gamma(x')$. Since $u \notin \Gamma(x')$, function $|u - u'|$ is differentiable with respect to u', with gradient $-(u - u')/|u - u'|$. By Assumption (LI) and the convexity of η, one can show[4] that η is differentiable and the derivative $D\eta(x') = -\lambda D_1 g(x', u')$, with $\lambda = \big((u - u')/|u - u'|\big)D_2 g_\sigma^+(x', u')$. Hence, $D\eta(x') = -\big((u - u')/|u - u'|\big)G(x', u')$. By the convexity of η and the fact that $\eta(x) = 0$, we must have

$$\eta(x') - \eta(x) = \eta(x') \leq D\eta(x') \cdot (x - x') = -\big((u - u')/|u - u'|\big)G(x', u') \cdot (x' - x).$$

Therefore, taking norms, we obtain $|u - u'| \leq \|G(x', u')\| \, |x' - x|$ for some $u' \in \Gamma(x')$. □

Note that in Proposition 3.3 we assume that $\|G\|$ is bounded by an integrable function $\gamma(s)$. Then, the correspondence Γ is globally Lipschitz in the sense defined in [24]. The following result is actually a simple consequence of Theorem 3.11 in [24] on the Lipschitz dependence of solutions with respect to initial conditions for the differential inclusion $\dot{x} \in \Gamma(x)$.

Lemma A.3. *Let Γ be globally Lipschitz in $x \in X$. Then, for any fixed $T \geq t$, the correspondence $x \rightrightarrows \{x_t(T, x) : \dot{x}_t(s, x) \in \Gamma(x_t(s, x)), s \in [t, T]\}$ is globally Lipschitz on X.*

---

[4] E.g., [23].

As is well known (cf. [16]), the existence of an optimal control can be ensured under the following additional assumptions: (i) Correspondence $x \rightrightarrows \Gamma(x)$ is upper semicontinuous and compact valued; and (ii) For any $t \leq T' < T''$, the negative part of the functional $\int_{T'}^{T''} \ell(x_t(s), \dot{x}_t(s)) \beta(s, t) \, ds$ converges to zero uniformly over all admissible trajectories as $T', T'' \to \infty$. For completeness, we provide here sufficient conditions for existence of optimal paths, which will allow us to establish condition 1 in Proposition 3.3. Let $\delta = \inf_{s \geq t} \delta(s)$.

Lemma A.4. *Let* (A1)–(A2) *hold. Assume that correspondence $x \rightrightarrows \Gamma(x)$ is upper semicontinuous and compact valued. Moreover, for all pairs $(x, u) \in \Omega$ the following conditions are satisfied:*

1. *For some constants α, η with $\alpha < \delta$*

$$\langle x, u \rangle \leq \alpha(|x|^2 + \eta). \tag{26}$$

2. *For some constants C and K*

$$|\ell(x, u)| \leq C(|x| + K). \tag{27}$$

Then, there exists an optimal solution $x_t^* \in W_{loc}^{1,1}([0, \infty))$ for problem (1), *and the value function satisfies $|V(t, x)| \leq a|x| + b$ for suitable constants a, b. Furthermore,*

$$\lim_{T \to \infty} \beta(T, t) V(T, x_t(T)) = 0$$

for any feasible trajectory.

Proof. We only prove the linear growth condition on the value function and the property of the limit. For any admissible x_t, let $y(s) = |x_t(s)|^2 + \eta$. It follows from (26) that $\dot{z} = 2\langle x_t, \dot{x}_t \rangle \leq 2|x_t \parallel \dot{x}_t| \leq 2\rho y$. Hence $y(s) \leq (|x_0|^2 + \eta)e^{2\alpha(s-t)}$, and thus $|x_t(s)| \leq \left(\sqrt{|x_0|^2 + \eta} \right) e^{\alpha(s-t)}$. Then,

$$V(t, x_0) = \int_t^\infty \ell(x_t^*(s), \dot{x}_t^*(s)) \beta(s, t) \, ds \leq C \frac{1}{\delta - \alpha} \left(\sqrt{|x_0|^2 + \eta} + K \right) \leq a|x_0| + b,$$

for suitable constants a, b. The last claim of the lemma follows from the fact that V has linear growth in x and the admissible trajectories are at most of exponential growth ρ. $\square$

Now, we present some preparatory results for the proof of Proposition 3.4. From bound (15) we establish a global Lipschitz property over a restricted domain. Rockafellar and Wets [25, Example 9.14], provides a more limited result on global Lipschitzianity for *bounded* convex functions.

Theorem A.1. *Let $X \subseteq \mathbb{R}_+^n$ be a convex set such that $X + \varepsilon \mathbb{B} \subseteq \mathbb{R}_+^n$ for some $\varepsilon > 0$. Let $f : \mathbb{R}_+^n \longrightarrow \mathbb{R}$ be a convex function such that $f(x) \leq a|x| + b$ for all $x \in \mathbb{R}_+^n$ and suitable constants $a, b \geq 0$. Assume that f is bounded below by some constant m. Then, f is globally Lipschitz on X.*

Proof. There is no loss of generality to let $m = 0$. Consider two arbitrary points $x_1, x_2 \in X$. Let $v = x_2 - x_1$. Then, there are two possibilities for the line $x_1 + \lambda v, \lambda \in \mathbb{R}$. Either the line intercepts the boundary $A = \text{bd } \mathbb{R}_+^n$ at two different points, or the line intercepts A at a single point. In the first case, there exist $y_1, y_2 \in A, \|y_1 - y_2\| \geq 2\varepsilon$, such that

$$y_1 = x_1 + \lambda_1 v,$$
$$y_2 = x_1 + \lambda_2 v,$$

with $\lambda_1, \lambda_2 \in \mathbb{R}$ of opposite sign. Observe that

$$\tau = \frac{|x_1 - x_2|}{|y_1 - y_2|} < 1,$$

since

$$|y_1 - y_2| = |y_1 - x_1| + |x_1 - x_2| + |x_2 - y_2| \geq 2\varepsilon + |x_1 - x_2|.$$

Let us consider convex combination $x_1 = \tau y_1 + (1 - \tau)x_2$. By the convexity of f

$$f(x_1) \leq \tau f(y_1) + (1 - \tau)f(x_2).$$

Hence,

$$f(x_1) - f(x_2) \leq \tau(f(y_1) - f(x_2)) \leq \tau(f(y_1) - m)$$
$$\leq \frac{a|y_1| + b - m}{|y_1 - y_2|} |x_1 - x_2|.$$

Observe that expression

$$\frac{|y_1|}{|y_1 - y_2|}$$

is bounded by 1 by the Pythagorean Theorem as $|y_1 - y_2| \geq 2\varepsilon$. Thus, we have

$$f(x_1) - f(x_2) \leq \left(a + \frac{b - m}{2\varepsilon}\right) |x_1 - x_2|. \tag{28}$$

Now, pick convex combination $x_2 = \tau y_2 + (1 - \tau)x_1$. Using the same arguments, we get inequality

$$f(x_2) - f(x_1) \leq \left(a + \frac{b - m}{2\varepsilon}\right) |x_1 - x_2|.$$

It follows from these two inequalities that

$$|f(x_1) - f(x_2)| \leq \left(a + \frac{b - m}{2\varepsilon}\right) |x_1 - x_2|.$$

If line $x_1 + \lambda v$ intersects A at a single point, there is only one $y_1 \in A$ such that $y_1 = x_1 + \lambda_1 v$ for some $\lambda_1 < 0$ (we can exchange the roles of x_1 and x_2 if needed), whereas $y_\lambda = x_1 + \lambda v \in X$ for any $\lambda > 0$. Again, from $x_1 = \tau y_1 + (1 - \tau)x_2$ we can arrive to inequality (28). Consider now $x_2 = \tau y_\lambda + (1 - \tau)x_1$ for y_λ defined above. Then, as $f(x_2) \leq \tau f(y_\lambda) + (1 - \tau)f(x_1)$ we have

$$f(x_2) - f(x_1) \leq \frac{a|y_\lambda| + b - m}{|y_\lambda - y_1|} |x_1 - x_2|,$$

which holds for every $\lambda > 0$. Letting $\lambda \to \infty$,

$$f(x_2) - f(x_1) \leq a |x_1 - x_2|.$$

Hence,

$$|f(x_1) - f(x_2)| \leq \left(a + \frac{b - m}{2\varepsilon}\right) |x_1 - x_2|.$$

Therefore, in all cases f is globally Lipschitz with constant

$$K = a + \frac{b - m}{2\varepsilon}. \quad \square$$

The following result weakens the uniform lower bound m on f.

Corollary A.1. *Let $X \subseteq \mathbb{R}_+^n$ be a convex set such that $X + \varepsilon \mathbb{B} \subseteq \mathbb{R}_+^n$ for some $\varepsilon > 0$. Let $f : \mathbb{R}_+^n \longrightarrow \mathbb{R}$ be a convex function such that $-a'|x| - b' \leq f(x) \leq a|x| + b$ for all $x \in \mathbb{R}_+^n$ and for some constants $a, b, a', b' \geq 0$. Then, f is globally Lipschitz on X.*

Proof. Let us define function $g(x) = f(x) + a'|x| + b'$. This function fulfills all the hypotheses of our theorem since it is convex, bounded below by $m = 0$, and bounded above by $(a + a')|x| + b + b'$. It follows that g is globally Lipschitz on X with Lipschitz constant K as estimated above. It is then easy to see that $|p| \leq K$ for all $p \in \partial g(x)$ and $x \in X$. Moreover,

$$\partial g = \partial(f + a'| \cdot | + b') = \partial f + a'\partial| \cdot |.$$

Hence, $\partial f = \partial g - a'\partial| \cdot |$. Observe that $\partial| \cdot |(x) = \frac{x}{|x|}$ for every point $x \in X$ with $|x| \neq 0$. Then, we have $|q| \leq K + a'$ for $q \in \partial f(x)$ and $x \in X$. Therefore, f is globally Lipschitz on X. $\quad \square$

The next proposition can be found in [17, Problem 22]. See also [21].

Proposition A.2. *Let H be a proper, concave, upper semicontinuous function from $\mathbb{R}^n \times \mathbb{R}^m$ to $\mathbb{R} \cup \{-\infty\}$. Let*

$$H(x, q) = \sup_{u \in \mathbb{R}^m} \{f(x, u) + qu\}.$$

Then, $x \mapsto H(x, q)$ is a concave mapping for a fixed q, and $q \mapsto H(x, q)$ is a convex mapping for a fixed x. Moreover, the following conditions are equivalent

$$-(p, q) \in \partial f(x, u)$$
$$-p \in \partial_x H(x, q) \quad and \quad u \in \partial_q H(x, q).$$

Proof of Theorem 4.1. Suppose that the pair (x_t^*, q_t) satisfies the Hamiltonian inclusions (20) with $x_t^*(t) = x_0$. It is well known that this condition along with (NB) and (21) constitute a sufficient criterion for optimality of $(x_t^*, \dot{x}_t^*)$ for problem (1). For instance, the proof given in [6] can be easily adapted to our framework; we do not repeat the details here. Let us then assume that $(x_t^*, \dot{x}_t^*)$ is an optimal path with two associated paths of dual variables q_t and q_t' satisfying both the Hamiltonian inclusions (20) and the transversality condition (21). For x_0 fixed, let

$$V_T(t, x_0) = \max \int_t^T \mathscr{L}(x_t(s), \dot{x}_t(s))\beta(s, t)\, ds + q_t(T)x_t(T) \tag{29}$$

subject to $x(t) = x_0$,

and

$$V_T'(t, x_0) = \max \int_t^T \mathscr{L}(x_t(s), \dot{x}_t(s))\beta(s, t)\, ds + q_t'(T)x_t(T) \tag{30}$$

subject to $x(t) = x_0$.

Note that the added linear parts $q_t(T)x_t(T)$ and $q_t'(T)x_t(T)$ are chosen so that $(x_t^*, \dot{x}_t^*)$ with $x_t^*(t) = x_0$ is the optimal solution for both optimization problems. We can readily see that functions $V_T(t, x_0)$ and $V_T'(t, x_0)$ are concave; moreover, by the same arguments as in Lemma 3.3 these functions are of class C^1 in x. By the transversality condition (21), the sequences of functions $\{V_T(t, x_0)\}_{T \geq 0}$ and $\{V_T'(t, x_0)\}_{T \geq 0}$ converge pointwise to function $V(t, x_0)$ as $T \to \infty$. Hence, the sequences of derivative functions $\{DV_T(t, x_0)\}_{T \geq 0}$ and $\{DV_T'(t, x_0)\}_{T \geq 0}$ converge uniformly to function $DV(t, x_0)$ on every compact set $K \subset \mathrm{int}(X)$ (see [20, Theorem 25.7]). By Remark 3.2 the convergence of these derivatives to a unique common value $DV(t, x_0)$ implies that $q_t(T) = q_t'(T)$. Therefore, we get uniqueness of the path of dual variables q_t. $\quad\square$

References

[1] L.M. Benveniste, J.A. Scheinkman, On the differentiability of the value function in dynamic models of economics, Econometrica 47 (1979) 727–732.
[2] K.J. Arrow, Applications of control theory to economic growth, in: B. Dantzig, A.F. Veinott (Eds.), Mathematics of the Decisions Sciences, Part 2. Lectures in Applied Mathematics, Vol. 12, American Mathematical Society, Providence, RI, 1968, pp. 85–119.
[3] R. Goebel, Regularity of the optimal feedback and the value function in convex problems of optimal control, Set-Valued Anal. 12 (2004) 127–145.
[4] A. Araujo, J.A. Scheinkman, Maximum principle and transversality condition for concave infinite horizon economic models, J. Econom. Theory 30 (1983) 1–16.
[5] J.P. Aubin, F.H. Clarke, Shadow prices and duality for a class of optimal control problems, SIAM J. Control Optim. 17 (1979) 567–586.
[6] L.M. Benveniste, J.A. Scheinkman, Duality theory for dynamic optimization models of economics: the continuous time case, J. Econom. Theory 27 (1982) 1–19.
[7] M. Bardi, I. Capuzzo-Dolcetta, Optimal Control and Viscosity Solutions of Hamilton–Jacobi–Bellman Equations, Birkhäuser, Boston, MA, 1997.
[8] R. Amir, L.J. Mirman, W.R. Perkins, One-sector nonclassical optimal growth: optimality conditions and comparative dynamics, Int. Econom. Rev. 32 (1991) 625–644.
[9] R. Amir, Sensitivity analysis of multisector optimal economic dynamics, J. Math. Econom. 25 (1996) 123–141.
[10] K. Askri, C. Le Van, Differentiability of the value function of nonclassical optimal growth models, J. Optim. Theory Appl. 97 (1998) 591–604.
[11] K.D. Cotter, J.-H. Park, Non-concave dynamic programming, Econ. Lett. 90 (2006) 141–146.
[12] P. Milgrom, I. Segal, Envelope theorems for arbitrary choice sets, Econometrica 70 (2002) 583–601.
[13] J.P. Rincón-Zapatero, M.S. Santos, Differentiability of the value function without interiority assumptions, J. Econom. Theory 144 (2009) 1948–1964.
[14] M.S. Santos, M. Woodford, Rational asset pricing bubbles, Econometrica 65 (1997) 19–58.
[15] M.S. Santos, Differentiability and comparative analysis in discrete-time optimization, J. Econom. Theory 57 (1992) 222–229.
[16] A.V. Dmitruk, N.V. Kuz'kina, Existence theorem in the optimal control problem on an infinite time interval, Math. Notes 78 (2005) 466–480.
[17] J.P. Aubin, Optima and Equilibria, Springer-Verlag, New York, 1993.
[18] T. Rockafellar, Integrals which are convex functionals, Pacific J. Math. 24 (1968) 525–539.
[19] T. Rockafellar, Integrals which are convex functionals, II, Pacific J. Math. 39 (1971) 439–469.
[20] T. Rockafellar, Convex Analysis, Princeton University Press, Princeton, NJ, 1970.
[21] T. Rockafellar, Generalized Hamiltonian equations for convex problems of Lagrange, Pacific J. Math. 33 (1970) 411–427.
[22] I. Ekeland, R. Témam, Convex Analysis and Variational Problems, SIAM, New York, 1999.
[23] J. Gauvin, F. Dubeau, Differential properties of the marginal function in mathematical programming, Math. Program. Stud. 19 (1982) 101–119.
[24] F.H. Clarke, Yu.S. Ledyaev, R.J. Stern, P.R. Wolenski, Nonsmooth Analysis and Control Theory, Springer-Verlag, New York, 1998.
[25] T. Rockafellar, R. Wets, Variational Analysis, Springer-Verlag, Berlin, 1998.