Envelope representations in Hamilton-Jacobi theory for fully convex problems of control

R.Tyrrell Rockafellar and Peter R. Wolenski

Abstract

This paper is a sequel to the one in this same session which surveys recent results on the role of convexity in Hamilton-Jacobi theory. We describe here how value functions in optimal control can be represented as upper and lower envelopes involving so-called kernel functions. Particularly noteworthy is a lower envelope formula given in terms of the dualizing kernel, which is a value function in its own right with many surprising and attractive properties.

1 Introduction

We continue our survey and development of *Fully Convex* (FC) optimal control problems that was begun in our companion paper. We assume the reader is familiar with the notation and assumptions introduced there, and those assumptions are in effect throughout this paper as well.

Recall that the basic model has the form

$$\inf\left\{ g\big(x(0)\big) + \int_0^\tau L\big(x(t),\dot{x}(t)\big)dt \,\middle|\, x(\tau)=\xi \right\}, \qquad (1)$$

where the minimization takes place over the arc space $\mathcal{A}^1_n[0,\tau]$ of absolutely continuous functions, and the initial cost function $g : I\!R^n \to \overline{I\!R} := (-\infty,+\infty]$ and Lagrangian function $L : I\!R^n \times I\!R^n \to \overline{I\!R}$ are both assumed to be convex, lower semicontinuous, and proper. The convexity in L is in both variables (x,v) jointly, which on the one hand is quite special, but on the other hand, leads to a harmonious and symmetric theory while still covering many important applications. The extended-real-valuedness of g and L serve as a modelling tool and assure that problems in optimal control are covered.

Fundamental to many issues in optimal control is the value function $V(\tau,\xi)$, which is the infimum value in (1) with the time τ and terminal state ξ viewed as parameters. Our companion paper dealt with V directly as a solution to the Hamilton-Jacobi equation. Here we describe ways of representing V in terms of better behaved functions, called kernel functions. The terminology comes from a far-reaching analogy between minimizing a sum of functions and integrating a product of functions.

The kernel functions act in regard to the solution of the Hamilton-Jacobi equation as the fundamental solution in linear ODE theory acts in the variation of constants formula, or as the Green's function acts in the representation of solutions to linear elliptic PDE's. The kernels contain two types of parameters, one that in effect trivializes the boundary conditions, and the second over which the differential equation is satisfied inside the domain. The sought-after solution of the original differential equation is represented through employing an operation to the first set of parameters "matching up" the boundary condition, while at the same time, preserving the solution to the differential equation property in the second parameter. In linear ODE the operation is just matrix multiplication, and in elliptic PDE theory it is integration. In FC theory the operation is sup or inf convolution. The use of sup or inf depends on whether the representation is an upper (inf) or lower (sup) envelope, and also dictates the nature of the kernel. We shall soon see there are many advantages to the lower representation since the kernel associated with that envelope has many agreeable properties.

2 Kernel functions and envelopes

2.1 The upper formula

The upper envelope formula relies on a kernel function $E : [0,\infty) \times I\!R^n \times I\!R^n \to \overline{I\!R}$ that treats both endpoints as parameters. Specifically, $E(\tau,\xi',\xi)$ is the infimum value of the optimization problem

$$\inf\left\{ \int_0^\tau L\big(x(t),\dot{x}(t)\big)dt \,\middle|\, x(0)=\xi',\ x(\tau)=\xi \right\} \qquad (2)$$

when $\tau > 0$, and at $\tau = 0$ is 0 if $\xi = \xi'$ and $+\infty$ otherwise. Note that the optimization problems (1) and (2) coincide if g in (1) is taken as

$$g(\xi'') := \begin{cases} 0 & \text{if } \xi'' = \xi', \\ \infty & \text{otherwise.} \end{cases}$$

We call E the *fundamental kernel* in analogy with fundamental matrices of ODE theory. The following theorem is elementary and evident from the definitions.

Theorem 2.1 (upper envelope representation) *For each $(\tau,\xi) \in [0,\infty) \times I\!R^n$, the value function V*

is expressed in terms of E by the formula

$$V(\tau,\xi) = \inf_{\xi' \in \mathbb{R}^n} \Big\{ g(\xi') + E(\tau,\xi',\xi) \Big\}.$$

In fact the upper formula holds for any g. It is called an *upper* envelope because V sits below a collection of functions indexed by ξ' and is the pointwise infimum over this collection.

One of the main examples covered by FC theory is linear control problems with hard control control constraints. There, in particular, the fundamental kernel E may not be finite for a given (τ,ξ',ξ). In fact, $E(\tau,\xi',\xi)$ is finite precisely when there is an arc $x(\cdot)$ that goes from ξ' to ξ over the interval $[0,\tau]$ having the integral $\int_0^\tau L\big(x(t),\dot{x}(t)\big)dt4$ finite. An assumption requiring E finite everywhere would amount to global controllability that would preclude many applications. It could happen that $V(\tau,\xi)$ takes on $+\infty$ as well, which is the case when $E(\tau,\xi',\xi)$ is $+\infty$ for all ξ' satisfying $g(\xi') < \infty$.

The appearance of $+\infty$ in V and E is well-handled in the general FC theory and is not really an impediment there, but it clearly becomes troublesome in any practical situation. This motivates an alternative description for V that approximates from below with finite values.

2.2 The lower envelope formula

The lower formula relies on a kernel which introduces parameters in a different manner than the fundamental kernel. The *dualizing* kernel $K : [0,\infty) \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is another instance of the type of value function V defined through (1), in this case with g taken as the linear function $g(\xi') = \langle \eta, \xi' \rangle$ where η is introduced as a new parameter. That is, $K(\tau,\xi,\eta)$ is the infimum value of

$$\inf_{x(\cdot)} \left\{ \langle \eta, x(0) \rangle + \int_0^\tau L\big(x(t),\dot{x}(t)\big)dt \,\middle|\, x(\tau) = \xi \right\}.$$

In view of Theorem 2.1, we could just as well take the formula

$$K(\tau,\xi,\eta) = \inf_{\xi' \in \mathbb{R}^n} \Big\{ \langle \eta, \xi' \rangle + E(\tau,\xi',\xi) \Big\} \qquad (3)$$

as the definition of K.

The convexity assumptions on L imply that $E(\tau,\xi',\xi)$ is convex in (ξ',ξ), and since (3) can be rewritten as

$$-K(\tau,\xi,-\eta) = \sup_{\xi' \in \mathbb{R}^n} \Big\{ \langle \eta, \xi' \rangle - E(\tau,\xi',\xi) \Big\},$$

it follows that $\eta \mapsto K(\tau,\xi,\eta)$ is concave and is closely related to the Legendre-Fenchel conjugate of $\xi' \mapsto E(\tau,\xi',\xi)$. This suggests further that the upper formula (2) can be dualized, and that is in fact the case, which is the substance of the following theorem.

Theorem 2.2 (lower envelope representation) *For each $(\tau,\xi) \in [0,\infty) \times \mathbb{R}^n$, the value function V is expressed in terms of K by the formula*

$$V(\tau,\xi) = \sup_{\eta \in \mathbb{R}^n} \Big\{ K(\tau,\xi,\eta) - g^*(\eta) \Big\}. \qquad (4)$$

It is a *lower* envelope because V sits above a collection of functions indexed by η and is the pointwise supremum over this collection. Since (4) involves the conjugate g^*, we call K the dualizing kernel.

Unlike the fundamental kernel E, the dualizing kernel is always finite-valued (under our assumptions) — see Theorem 3.2 below. The next section shows that the dualizing kernel has other appealing properties as well.

3 The dualizing kernel

3.1 Double Hamilton-Jacobi equation

By definition, the dualizing kernel is a value function in the variables (τ,ξ) for each fixed η, but in fact it is also a value function in (τ,η) of yet another variational problem for each parameter ξ. A consequence of duality theory is that $-K(\tau,\xi,\eta)$ is the infimum value of

$$\inf_{y(\cdot)} \left\{ \int_0^\tau \tilde{L}\big(y(t),\dot{y}(t)\big)dt - \langle \xi, y(\tau) \rangle \,\middle|\, y(0) = \eta \right\},$$

where $\tilde{L}$ is the dual Lagrangian. Thus K satisfies a double Hamilton-Jacobi equation, one associated to ξ as the state variable and η as a parameter, and another with the dual Hamiltonian and the roles of ξ and η reversed. Moreover, K also has a surprising differentiable property in t.

Theorem 3.1 *The kernel $K(\tau,\xi,\eta)$ is continuously differentiable with respect to τ and satisfies, for $\tau \geq 0$,*

$$\frac{\partial K}{\partial \tau}(\tau,\xi,\eta) = \begin{cases} -H(\xi,\eta') & \text{for all } \eta' \in \partial_\xi K(\tau,\xi,\eta), \\ -H(\xi',\eta) & \text{for all } \xi' \in \tilde{\partial}_\eta K(\tau,\xi,\eta), \end{cases}$$

$$K(0,\xi,\eta) = \langle \xi, \eta \rangle,$$

where $\partial K/\partial\tau$ is interpreted as the right partial derivative when $\tau = 0$.

The issue of whether K is uniquely determined as the solution to *one* of the HJ equations in the last theorem has only been resolved recently by Galbraith [3]; the growth assumptions we have placed on the Hamiltonian is more general than earlier viscosity theory allowed. The uniqueness of K satisfying *both* of the equations and some further regularity was proved in our original paper [2].

3.2 Continuity and differentiability properties

As already mentioned, the fundamental kernel E may take on $+\infty$'s, but K does not. This section suggests that K's further structural makeup could make it more practically useful for numerical computation than E. Continuity and differentiability properties of K are stated in the next two theorems.

Theorem 3.2 *The dualizing kernel K is locally Lipschitz (and in particular, is finite) on $[0,\infty)\times \mathbb{R}^n \times \mathbb{R}^n$. For each $\tau \geq 0$, the map $\xi \mapsto K(\tau,\xi,\eta)$ is convex for each $\eta \in \mathbb{R}^n$, and $\eta \mapsto K(\tau,\xi,\eta)$ is concave for each $\xi \in \mathbb{R}^n$.*

Theorem 3.3 *For each $(\tau,\xi,\eta) \in (0,\infty) \times \mathbb{R}^n \times \mathbb{R}^n$, the directional derivative*

$$\begin{aligned} &dK(\tau,\xi,\eta)(\theta,\omega,\zeta) \\ &\quad := \lim_{h\downarrow 0} \tfrac{1}{h}\big\{K(\tau+h\theta,\xi+h\omega,\eta+h\zeta) - K(\tau,\xi,\eta)\big\} \end{aligned}$$

exists in all directions $(\theta,,\omega,\zeta)$, and is equal to

$$\begin{aligned} \theta k(\tau,\xi,\eta) \; &+ \max\big\{\langle \eta',\omega\rangle \,\big|\, \eta' \in \partial_\xi K(\tau,\xi,\eta)\big\} \\ &+ \min\big\{\langle \xi',\zeta\rangle \,\big|\, \xi' \in \partial_\eta K(\tau,\xi,\eta)\big\}. \end{aligned}$$

The quantity $k(\tau,\xi,\eta)$ is the common value of $H(\xi,\eta')$ with $\eta' \in \partial_\xi K(\tau,\xi,\eta)$, and $H(\xi',\eta)$ with $\xi' \in \tilde{\partial}_\eta K(\tau,\xi,\eta)$.

3.3 Propagation of the Legendre-Fenchel conjugate formula

Theorem 2.2 has an interesting interpretation that sheds light on the nature of the evolution of solutions to the Hamilton-Jacobi equation. Recall that the boundary conditions are $V(0,\xi) = g(\xi)$ and $K(0,\xi,\eta) = \langle \xi,\eta\rangle$. Thus the lower formula at $\tau = 0$ reduces to

$$g(\xi) = \sup_{\eta\in\mathbb{R}^n} \Big\{\langle \xi,\eta\rangle - g^*(\eta)\Big\},$$

which of course is just the recovery formula of a convex function from its Legendre-Fenchel conjugate. Theorem 2.2 says that this formula is propagated forward in τ by replacing the bilinear function $(\xi,\eta) \mapsto \langle \xi,\eta\rangle$ by the convex/concave function $(\xi,\eta) \mapsto K(\tau,\xi,\eta)$.

4 Subgradient formula

In control applications, the subgradients of V are perhaps more important than the values of V itself, and so it is highly desirable if these can be obtained in a straightforward manner. The dualizing kernel shows itself again as an important entity because the subgradients of V can be obtained from K through solving the finite-dimensional optimization problem contained in the lower envelope formula (4). It is not clear how one could obtain a similar type of characterization through the subgradients of E.

Theorem 4.1 *For every $(\tau,\xi) \in [0,\infty)\times\mathbb{R}^n$, one has*

$$\partial V(\tau,\xi) \; = \bigcup \Big\{\partial_{\tau,\xi} K(\tau,\xi,\eta) \,\Big|\, \eta \in M(\tau,\xi)\Big\},$$

where

$$M(\tau,\xi) := \text{argmax}_\eta \Big\{K(\tau,\xi,\eta) - g^*(\eta)\Big\}$$

is the set of η achieving the maximum in (4). In other words, we have $(\sigma,\eta') \in \partial V(\tau,\xi)$ if and only if

$$\exists\, \eta \in M(\tau,\xi) \;\; \text{with} \;\; \begin{cases} \eta' \in \partial_\xi K(\tau,\xi,\eta), \\ \sigma = -H(\xi,\eta'). \end{cases}$$

The previous theorem raises further questions about the nature of the optimization problem in (4), and in particular whether optimal solutions exist. The following result addresses this issue.

Theorem 4.2 *For any $(\tau,\xi) \in (0,\infty,)\times\mathbb{R}^n$, the following properties in the lower envelope formula are equivalent:*

(a) the argmax set $M(\tau,\xi)$ is nonempty and compact;

(b) for every $\beta \in \mathbb{R}$, the upper level set

$$\big\{\eta \,\big|\, K(\tau,\xi,\eta) - g^*(\eta) \geq \beta\big\}$$

is compact;

(c) ξ belongs to the interior of the set

$$\big\{\xi \,\big|\, V(\tau,\xi) < \infty\big\}.$$

5 Example prototype

Of course one should not expect the kernels to easily be calculated in all case, since nonlinear equations generally do not have closed form solutions. However, this section considers a certain type of Lagrangian in which expressions for the kernels E and K can be easily formulated, in which case the envelope formulas become more explicit.

Consider a Lagrangian $L : \mathbb{R}^n \times \mathbb{R}^n \to \overline{\mathbb{R}}$ of the form

$$L(x,v) = L_0(v - Ax),$$

where A is an $n\times n$ matrix and L_0 is a proper convex function on $\mathbb{R}^n$ that is lsc and coercive. For example, L could come from the *dynamics* in a Linear-Quadratic Regulator (LQR) problem with hard constraints (see Section 2 of our companion paper), in which case $L_0(w)$ is the infimum in u of $u^\top Ru$ taken over the set $\{u : u \in U,\, w = Bu\}$ (by convention, the inf over the empty set is $+\infty$).

Define $\Psi : [0,\infty) \times I\!R^n \to I\!R$ by

$$\Psi(\tau,\eta) := \int_0^\tau H_0(e^{-tA^\top}\eta)\,dt,$$

this expression being finite and convex in η. Then the dualizing kernel is given by

$$K(\tau,\xi,\eta) = \langle e^{-\tau A}\xi, \eta\rangle - \Psi(\tau,\eta)$$

and the fundamental kernel is given by

$$E(\tau,\xi',\xi) = \Phi(\tau, e^{-\tau A}\xi - \xi'),$$

where $\Phi : [0,\infty) \times I\!R^n \to \overline{I\!R}$ is the function defined by taking $\Phi(\tau,\cdot)$ to be the (proper, lsc and coercive) convex function conjugate to $\Psi(\tau,\cdot)$ for each τ. Thus, for initial function $g : I\!R^n \to \overline{I\!R}$ one has the upper envelope representation

$$V(\tau,\xi) = \inf_{\xi'}\Big\{g(\xi') + \Phi(\tau, e^{-\tau A}\xi - \xi')\Big\},$$

and the lower envelope representation

$$V(\tau,\xi) = \sup_\eta\Big\{\langle e^{-\tau A}\xi,\eta\rangle - \Psi(\tau,\eta) - g^*(\eta)\Big\}.$$

One may note directly in this prototypical example that $E(\tau,\cdot,\cdot)$, $\tau > 0$ is not necessarily finite everywhere, but that K always is.

6 Conclusion

The FC problem formulation covers a particular type of nonlinear optimal control problem that extends the LQR model. The popularity of the LQR model lies in the fact that explicit solutions are available in feedback form, whereas in general no explicit representation could be expected. The reasoning behind the derivation of these properties in LQR relies on the nature of quadratic functions, of course, but the essential ingredient needed for a global representation is just the joint convexity of the Lagrangian. This weaker assumption allows a control designer much greater freedom in handling control constraints, while at the same time, maintains global features of the problem and offers formulas for the solution.

References

[1] R.T. Rockafellar and P.R. Wolenski, "Convexity in Hamilton-Jacobi theory I: dynamics and duality," SIAM J. Control and Opt. 39 (2001), 1323–1350.

[2] R.T. Rockafellar and P.R. Wolenski, "Convexity in Hamilton-Jacobi theory II: envelope representations," SIAM J. Control Opt. 39 (2001), 1351-1372.

[3] G. Galbraith, "Applications of variational analysis to optimal trajectories and nonsmooth HJ theory", Ph.D. thesis, University of Washington, 1999.