Optimal Control and Myopic Solution

René Caldentey

March 26, 2003

Consider the following deterministic control problem in Bolza form:

$$\begin{aligned} &\min_{u\in\mathcal{U}}\ J(x(T)) + \int_0^T L(x_t,u_t)\,\mathrm{d}t \\ \text{subject to}\quad &\dot{x}(t) = f(x_t,u_t), \qquad x(0)=x_0. \end{aligned}$$

The functions f, J, and L are assumed to be "sufficiently" smooth.

The solution to this problem can be found solving the associated Hamilton-Jacobi-Bellman equation

$$V_t(t,x) + \min_{u\in U}\{f(x,u)\,V(t,x) + L(x,u)\} = 0$$

with boundary condition $V(T,x)=J(x)$. The value function $V(t,x)$ represents the optimal cost-to-go starting at time t in state x.

Suppose, we fix the control $u\in U$ and solve the first-order PDE

$$W_t(t,x;u) + f(x,u)\,W_x(t,x;u) + L(x,u) = 0, \qquad W(T,x;u) = J(x) \tag{1}$$

using the methods of characteristics. That is, we solve the characteristic ODE $\dot{x}(t)=f(x,u)$ and let $x(t)=H(t;s,y,u)$ the solution passing through the point (s,y), *i.e.*, $x(s)=H(s;s,y,u)=y$.

By construction, along a characteristic curve $(t,x(t))$ the function $W(t,x(t);u)$ satisfies $\dot{W}(t,x(t);u)+L(x(t),u)=0$. Therefore, after integration we have that

$$W(s,x(s);u) = W(T,x(T);u) + \int_s^T L(x(t),u)\,\mathrm{d}t = J(x(T)) + \int_s^T L(x(t),u),$$

where the second equality follows from the boundary condition for W. We can rewrite this last identity for the particular characteristic curve passing through (t,x) as follows

$$W(t,x;u) = J(H(T;t,x,u)) + \int_t^T L(H(s;t,x;u),u)\,\mathrm{d}s.$$

Since the control u has been fixed so far, we call $W(t,x;u)$ the *static value function* associated to control u. Now, if we view $W(t,x;u)$ as a function of u, we can minimize this static value function. We define

$$u^*(t,x) = \arg\min_{u\in U} W(t,x;u) \qquad \text{and} \qquad \mathcal{V}(t,x) = W(t,x;u^*(t,x)).$$

Proposition 1 *Suppose that $u^*(t,x)$ is an interior solution and that $W(t,x;u)$ is sufficiently smooth so that $u^*(t,x)$ satisfies*

$$\frac{dW(t,x;u)}{du}\bigg|_{u=u^*(t,x)} = 0. \tag{2}$$

Then the function $\mathcal{V}(t,x)$ satisfies the PDE

$$\mathcal{V}_t(t,x) + f(x,u^*(t,x))\,\mathcal{V}_x(t,x) + L(x,u^*(t,x)) = 0 \tag{3}$$

with boundary condition $\mathcal{V}(t,x) = J(x)$.

Proof: Let us rewrite the PDE in terms of $W(t,x,u^*)$ to get

$$\underbrace{\frac{\partial W(t,x,u^*)}{\partial t} + f(x,u^*)\,\frac{\partial W(t,x,u^*)}{\partial x} + L(x,u^*)}_{\text{(a)}} + \left[\frac{\partial u^*}{\partial t} + f(x,u^*)\,\frac{\partial u^*}{\partial x}\right]\underbrace{\frac{\partial W(t,x,u^*)}{\partial u^*}}_{\text{(b)}}.$$

We note that by construction of the function W on equation (1) the expression denoted by (a) is equal to zero. In addition, the optimality condition (2) implies that (b) is also equal to zero. Therefore, $\mathcal{V}(t,x)$ satisfies the PDE (3). The border condition follows again from the definition of the value function W. ∎

Given this result, the question that naturally arises is whether $\mathcal{V}(t,x)$ is in fact the value function (that is $\mathcal{V}(t,x) = V(t,x)$) and $u^*(t,x)$ is the corresponding optimal feedback control.

Unfortunately, this is not generally true. In fact, to prove that $V(t,x) = \mathcal{V}(t,x)$ we would need to show that

$$u^*(t,x) = \arg\min_{u\in U}\left\{f(x,u)\,\mathcal{V}_x(t,x) + L(x,u)\right\}.$$

Since we have assumed that $u^*(t,x)$ is an interior solution then the first order optimality condition for the minimization problem above is given by

$$f_u(x,u^*(t,x))\,\mathcal{V}_x(t,x) + L_u(x,u^*(t,x)) = 0.$$

Using the optimality condition (2) we have that

$$\mathcal{V}_x(t,x) = W_x(t,x;u^*) = J'(H)\,H_x + \int_t^T L_x(H,u^*)\,H_x\,\mathrm{d}t,$$

where $H = H(T;t;x,u^*)$ and $H_x = H_x(T;t;x,u^*)$ the partial derivative of $H(T;t;x,u^*)$ with respect to x keeping $u^* = u^*(t,x)$ fixed. Thus, the first order optimality condition that needs to be verified is

$$f_u(x,u^*(t,x))\left(J'(H)\,H_x + \int_t^T L_x(H,u^*)\,H_x\,\mathrm{d}t\right) + L_u(x,u^*(t,x)) = 0. \tag{4}$$

On the other hand, the optimality condition (2) that $u^*(t,x)$ satisfies is

$$J'(H)\,H_u + \int_t^T \left[L_x(H,u^*)\,H_u + L_u(H,u^*)\right]\,\mathrm{d}t = 0. \tag{5}$$

It should be clear that condition (5) does not necessarily imply condition (4) and so $\mathcal{V}(t,x)$ and $u^*(t,x)$ are not guaranteed to be the value function and the optimal feedback control, respectively. The following example shows the suboptimality of $u^*(t,x)$.

Example 1: Consider the traditional linear-quadratic control problem

$$\begin{aligned} &\min_u \left\{ x^2(T) + \int_0^T (x^2(t) + u^2(t))\,\mathrm{d}t \right\} \\ \text{subject to}\quad &\dot{x}(t) = \alpha x(t) + \beta u(t), \qquad x(0) = x_0. \end{aligned}$$

- **Exact solution to the HJB equation:** This problem is traditionally tackled solving an associated Riccati differential equation. We suppose that the optimal control satisfies

$$u(t,x) = -\beta\,k(t)\,x,$$

where the function $k(t)$ satisfies the Riccati ODE

$$\dot{k}(t) + 2\alpha\,k(t) = \beta^2\,k^2(t) - 1, \qquad k(T) = 1.$$

We can get a particular solution assuming $k(t) = \bar{k} =$ constant. In this case,

$$\beta^2\,\bar{k}^2 - 2\alpha\,\bar{k} - 1 = 0 \qquad \Longrightarrow \qquad \bar{k}^{\pm} = \frac{\alpha \pm \sqrt{\alpha^2+\beta^2}}{\beta^2}.$$

Now, let us define $k(t) = z(t) + \bar{k}^+$ then the Riccati becomes

$$\dot{z}(t) + 2(\alpha - \beta^2\,\bar{k}^+)\,z(t) = \beta^2\,z^2(t) \;\Longrightarrow\; \frac{\dot{z}(t)}{z^2(t)} + \frac{2(\alpha - \beta^2\,\bar{k}^+)}{z(t)} = \beta^2.$$

If we set $w(t) = z^{-1}(t)$ then the last ODE is equivalent to

$$\dot{w}(t) + 2(\alpha - \beta^2\,\bar{k}^+)\,w(t) = \beta^2.$$

This a simple linear differential equation that can be solved using the integrating factor $\exp(2(\alpha - \beta^2\,\bar{k}^+)\,t$, that is,

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big(\exp\Big(2(\alpha - \beta^2\bar{k}^+)\,t\Big)\,w(t)\Big) = \exp\Big(2(\alpha - \beta^2\,\bar{k}^+)\,t\Big)\,\beta^2.$$

The solution to this ODE is

$$w(t) = \widetilde{k}\,\exp\Big(-2(\alpha - \beta^2\,\bar{k}^+)\,t\Big) + \frac{\beta^2}{2(\alpha - \beta^2\,\bar{k}^+)},$$

where $\widetilde{k}$ is a constant of integration. Using the fact that $\alpha - \beta^2\,\bar{k}^+ = -\sqrt{\alpha^2+\beta^2}$ and $k(t) = \bar{k}^+ + 1/w(t)$ we get

$$k(t) = \frac{\alpha + \sqrt{\alpha^2+\beta^2}}{\beta^2} + \frac{2\,\sqrt{\alpha^2+\beta^2}}{2\,\widetilde{k}\,\sqrt{\alpha^2+\beta^2}\,\exp\Big(-2(\alpha - \beta^2\,\bar{k}^+)\,t\Big) - \beta^2}.$$

The value of $\widetilde{k}$ is obtained from the border condition $k(T) = 1$.

- **Myopic Solution:** If we solve the problem using the myopic approach described at the beginning of this notes, we get that the characteristic curve is given by

$$\dot{x}(t) = \alpha\, x(t) + \beta\, u \quad \Longrightarrow \quad \ln(\alpha\, x(t) + \beta\, u) = \alpha(t + A),$$

with A a constant of integration. The characteristic passing through the point (t, x) satisfies $A = \ln(\alpha\, x + \beta\, u)/\alpha - t$ and is given by

$$x(\tau) = \frac{(\alpha\, x + \beta\, u)\, \exp(\alpha(\tau - t)) - \beta\, u}{\alpha}.$$

The value of the static value function $W(t, x; u)$ is given by

$$W(t, x; u) = \Big(\frac{(\alpha\, x + \beta\, u)\, \exp(\alpha(T - t)) - \beta\, u}{\alpha}\Big)^2 + \int_t^T \Bigg[\Big(\frac{(\alpha\, x + \beta\, u)\, \exp(\alpha(\tau - t)) - \beta\, u}{\alpha}\Big)^2 + u^2\Bigg]\, \mathsf{d}\tau.$$

If we compute the derivative of $W(t, x; u)$ with respect to u and make it equal to zero we get, after some manipulations, that the optimal myopic solution is

$$u^*(t, x) = \left(\frac{\alpha\, \beta\, (3 \exp(2(T - t)) - 4 \exp(T - t) + 1)}{\beta^2 (3 \exp(2(T - t)) - 8 \exp(T - t) + 5) + 2(T - t)(\alpha^2 + \beta^2)}\right)\, x.$$

Interestingly, this myopic feedback control is also linear on x as in the optimal solution, however, the solution is clearly different and suboptimal.

The previous example shows that in general the use of a myopic policy produces suboptimal solutions. However, a questions remains still open which is under what conditions is the myopic solution optimal? A general solution to this problem can be obtained by looking under what restrictions on the problem's data the optimality condition (4) is implied by condition (5).

In what follows we present one specific case for which the optimal solution is given by the myopic solution. Consider the control problem

$$\begin{aligned} &\min_{u \in \mathcal{U}} \; J(x(T)) + \int_0^T L(u(t))\, \mathrm{d}t \\ \text{subject to} \quad &\dot{x}(t) = f(x(t), u(t)) := g(x(t))\, h(u(t)), \qquad x(0) = x_0. \end{aligned}$$

In this case, it can be shown that the characteristic equation passing through the point (t, x) is given by

$$x(\tau) = G^{-1}(h(u)(\tau - t) + G(x)), \qquad \text{where } G(x) := \int \frac{\mathrm{d}x}{g(x)}.$$

In this case, the static value function is

$$W(t, x; u) = J(G^{-1}(h(u)(T - t) + G(x))) + L(u)(T - t)$$

and the myopic solution satisfies $\frac{\mathrm{d}}{\mathrm{d}u} W(t, x; u) = 0$ or equivalently

$$\begin{aligned} 0 &= J'(G^{-1}(h(u)(T - t) + G(x)))\, h'(u)\, (T - t)\, G^{-1\prime}(h(u)(T - t) + G(x)) + (T - t)\, L'(u) \iff \\ 0 &= J'(G^{-1}(h(u)(T - t) + G(x)))\, f_u(x, u)\, G'(x) G^{-1\prime}(h(u)(T - t) + G(x)) + L'(u) \iff \\ 0 &= f_u(x, u)\, W_x(t, x; u) + L'(u). \end{aligned}$$

The second equality uses the identities $G'(x) = 1/f(x)$ and $f_u(x,u) = f(x)\,h'(u)$. Therefore, the optimal myopic policy $u^*(t,x)$ satisfies

$$0 = f_u(x,u^*)\,\mathcal{V}_x(t,x) + L'(u^*)$$

i.e., the first order optimality condition (4).

<u>**Example 2:**</u> Consider control problem

$$\begin{aligned} &\min_u \left\{ x^2(T) + \int_0^T u^2(t)\,\mathsf{d}t \right\} \\ \text{subject to} \quad &\dot{x}(t) = x(t)u(t), \qquad x(0) = x_0. \end{aligned}$$

In this case, the characteristic passing through (t,x) is given by

$$x(\tau) = x\,\exp(u(\tau - t)).$$

The static value function is

$$W(t,x;u) = x^2\,\exp(2u(T-t)) + u^2\,(T-t).$$

Minimizing W over u implies

$$x^2\,\exp(2u^*(T-t)) + u^* = 0$$

and the corresponding value function

$$V(t,x) = \mathcal{V}(t,x) = u^*(t,x)\Big(u^*(t,x)\,(T-t) - 1\Big).$$