22 Hamilton Jacobi Bellman Equation

You must unlearn what you have learned
—Frank Oz

22.1 The Value Function

Our study of optimal control up to this point has been devoted to necessary conditions for the optimal solution that come entirely from Pontryagin's maximum principle. In this chapter we will take a different approach that leads to sufficiency conditions for the optimal solution, but leads to the solution of a (typically nonlinear and non-smooth) partial differential equation rather than a system of ordinary differential equations.

Consider the standard autonomous minimization problem for the cost functional

$$J[\mathbf{u}] = \int_s^{t_f} L(\mathbf{x}, \mathbf{u}) dt + \phi(\mathbf{x}(t_f)),$$

with

$$\mathbf{x}' = \mathbf{f}(\mathbf{x}, \mathbf{u}), \quad \mathbf{x}(s) = \mathbf{y},$$

and $\mathbf{x}(t_f)$ free. To simplify everything that follows we will assume that the solution $\mathbf{x}(t)$ and control $\mathbf{u}(t)$ are sufficiently bounded and that the evolution equation is well behaved as well. In addition we will most commonly write out the functional J with the following dependence on its initial condition, i.e. $J = J[s, \mathbf{y}, \mathbf{u}]$. Hence we assume that for some generic constant C,

$$|\mathbf{f}(\mathbf{x}, \mathbf{u})| \leq C_f, \quad |\mathbf{f}(\mathbf{x}, \mathbf{u}) - \mathbf{f}(\mathbf{y}, \mathbf{u})| \leq C_{fL}|\mathbf{x} - \mathbf{y}|, \tag{22.1}$$

for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and $\mathbf{u} \in U$ where $U \subset \mathbb{R}^m$ is a compact admissible control set. We also assume that the cost function is nicely bounded, i.e.

$$|L(\mathbf{x}, \mathbf{u})| \leq C_L, \quad |L(\mathbf{x}, \mathbf{u}) - L(\mathbf{y}, \mathbf{u})| \leq C_{LL}|\mathbf{x} - \mathbf{y}|, \tag{22.2}$$

$$|\phi(\mathbf{x})| \leq C_p, \quad |\phi(\mathbf{x}) - \phi(\mathbf{y})| \leq C_{pL}|\mathbf{x} - \mathbf{y}|. \tag{22.3}$$

Remark 22.1.1. We note here that it may seem at first glance that these conditions are overly restrictive. However if the solution starts from some bounded set then for some finite time $t < t_f$ all of the solutions $\mathbf{x}(t)$ and $\mathbf{y}(t)$ will be bounded, and so on a finite time interval these bounds are easily verified.

Using this form of the cost functional we are ready to define the value function.

Definition 22.1.2. *The value function for the optimal control problem defined above is given by*

$$V(s, \mathbf{y}) = \min_{\mathbf{u}\in U} J(s, \mathbf{y}, \mathbf{u}). \tag{22.4}$$

Note that the value function is the optimal cost depending on the initial time s and initial condition $\mathbf{y}$ of the state, i.e. the value function is not dependent on the exact solution itself but only on its initial condition.

We note that under the assumptions provided above, the value function is Lipschitz continuous in both s and $\mathbf{y}$.

Lemma 22.1.3. *For the optimal control problem defined above, satisfying the bounds described, there exists a constant C' such that*

$$|V(s, \mathbf{y})| \leq C', \quad |V(s, \mathbf{y}) - V(s, \mathbf{y}')| \leq C'_L \|\mathbf{y} - \mathbf{y}'\|, \tag{22.5}$$

i.e. the value function $V(s, \mathbf{y})$ is Lipschitz continuous. A similar Lipschitz condition holds for two different initial times s and s', i.e. $V(s, \mathbf{y})$ is Lipschitz in s as well as $\mathbf{y}$.

Proof. The boundedness of $V(s, \mathbf{y})$ is fairly straightforward. As L and ϕ are both bounded (and we have assumed that so is the solution $\mathbf{x}$ then it follows that $J[s, \mathbf{y}, \mathbf{u}]$ will be bounded as well and hence it's infimum over $\mathbf{u}$ will still resulted in a bounded quantity.

We will next consider the Lipschitz behavior of $V(s, \mathbf{y})$ in the second variable. Recall from much earlier in this text TODO: get reference that because $\mathbf{f}$ is Lipschitz in $\mathbf{x}$ then the solution $\mathbf{x}(t)$ is continuously dependent on its initial conditions, that is

$$\|\mathbf{x}(t; s, \mathbf{y}, \mathbf{u}) - \mathbf{x}(t; s, \mathbf{y}', \mathbf{u})\| \leq \exp\left(C_{fL}(t-s)\right) \|\mathbf{y} - \mathbf{y}'\| = \tilde{C}_{ts}\|\mathbf{y} - \mathbf{y}'\|.$$

It follows that

$$\begin{aligned}
&|J[s, \mathbf{y}, \mathbf{u}] - J[s, \mathbf{y}', \mathbf{u}]| \\
&= \left| \int_s^{t_f} \left[L(\mathbf{x}(t; s, \mathbf{y}, \mathbf{u}), \mathbf{u}) - L(\mathbf{x}(t; s, \mathbf{y}', \mathbf{u}), \mathbf{u}) \right] dt + \phi(\mathbf{x}(t_f; s, \mathbf{y}, \mathbf{u})) - \phi(\mathbf{x}(t_f; s, \mathbf{y}', \mathbf{u})) \right| \\
&\leq \int_s^{t_f} C_{LL} \|\mathbf{x}(t; s, \mathbf{y}, \mathbf{u}) - \mathbf{x}(t; s, \mathbf{y}', \mathbf{u})\| dt + C_{pL} \|\mathbf{x}(t_f; s, \mathbf{y}, \mathbf{u}) - \mathbf{x}(t_f; s, \mathbf{y}', \mathbf{u})\| \\
&\leq \int_s^{t_f} C_{LL} e^{C_{fL}(t-s)} \|\mathbf{y} - \mathbf{y}'\| dt + C_{pL} e^{C_{fL}(t_f - s)} \|\mathbf{y} - \mathbf{y}'\| \\
&= \left(\frac{C_{LL}}{C_{fL}} e^{C_{fL}(t_f - 2s)} + C_{pL} e^{C_{fL}(t_f - s)} \right) \|\mathbf{y} - \mathbf{y}'\| \\
&= C'_L \|\mathbf{y} - \mathbf{y}'\|,
\end{aligned}$$

where

$$C'_L = \frac{C_{LL}}{C_{fL}} e^{C_{fL}(t_f - 2s)} + C_{pL} e^{C_{fL}(t_f - s)}.$$

□

The point in developing the value function is that it will give us sufficient conditions for the solution of the optimal control problem, but before we can get there we first consider Bellman's dynamic programming technique.

Theorem 22.1.4 (Dynamic Programming Principle). *For every $\tau \in [s, t_f]$ and $\mathbf{y} \in \mathbb{R}^n$ then*

$$V(s, \mathbf{y}) = \inf_{\mathbf{u} \in U} \left\{ \int_s^\tau L(\mathbf{x}, \mathbf{u}) dt + V(\tau, \mathbf{x}) \right\} = J_\tau. \tag{22.6}$$

Before we develop the proof of this theorem we note that it means that we can break the optimization problem on the full interval $[s, t_f]$ into two parts:

(i) Solve the same problem on the sub-interval $[\tau, t_f]$ with running cost L and terminal cost ϕ which will give the value function $V(\tau, \cdot)$ at the intermediate time τ.

(ii) In the second step we solve the same problem but now on the sub-interval $[s, \tau]$ with running cost L and terminal cost $V(\tau, \cdot)$ which was determined on the previous step.

TODO: need a figure

Proof. The proof here will be done in two parts. First we note that the solution $\mathbf{x}(t)$ with initial condition $\mathbf{x}(s) = \mathbf{y}$ will be denoted as $\mathbf{x} = \mathbf{x}(t; s, \mathbf{y}, \mathbf{u})$ where $\mathbf{u}$ is the selected optimal control i.e. $x(t)$ evolves from initial condition $x(s) = \mathbf{y}$ under the control variable $\mathbf{u}$.

(i) First we prove that $J_\tau \leq V(s, \mathbf{y})$. Let $\varepsilon > 0$ be fixed, then we can choose a control $\mathbf{u}(t)$ defined on the interval $[s, t_f]$ so that

$$J(s, \mathbf{y}, \mathbf{u}) \leq V(s, \mathbf{y}) + \varepsilon.$$

Note that such a choice is certainly possible because we have defined $V(s, \mathbf{y})$ as the optimal such cost for the most optimal choice of the control variable. Now we note that

$$V(\tau, \mathbf{x}(\tau; s, \mathbf{y}, \mathbf{u})) \leq \int_\tau^{t_f} L(\mathbf{x}(t; s, \mathbf{y}, \mathbf{u}), \mathbf{u}) dt + \phi(\mathbf{x}(t_f; s, \mathbf{y}, \mathbf{u}))).$$

It follows that

$$\begin{aligned} J_\tau &\leq \int_s^\tau L(\mathbf{x}(t; s, \mathbf{y}, \mathbf{u}), \mathbf{u}) dt + V(\tau, \mathbf{x}(\tau; s, \mathbf{y}, \mathbf{u})) \\ &\leq J(s, \mathbf{y}, \mathbf{u}) \\ &\leq V(s, \mathbf{y}) + \varepsilon. \end{aligned}$$

As this was done for an arbitrary $\varepsilon > 0$ then we have established that $J_\tau \leq V(s, \mathbf{y})$.

(ii) Now we want to show the converse, that is $V(s, \mathbf{y}) \leq J_\tau$. Once again we let $\varepsilon > 0$ be fixed. It follows that there is a control $\hat{\mathbf{u}}$ defined on $[s, \tau]$ so that

$$\int_s^\tau L(\mathbf{x}(t; s, \mathbf{y}, \hat{\mathbf{u}}), \hat{\mathbf{u}}) dt + V(\tau, \mathbf{x}(\tau; s, \mathbf{y}, \hat{\mathbf{u}}) \leq J_\tau + \varepsilon.$$

In addition there is a control $\mathbf{u}'$ defined on $[\tau, t_f]$ so that

$$J(\tau, \mathbf{x}(\tau; s, \mathbf{y}, \hat{\mathbf{u}}), \mathbf{u}') \leq V(\tau, \mathbf{x}(\tau; s, \mathbf{y}, \hat{\mathbf{u}})) + \varepsilon.$$

Define a control $\mathbf{u}(t)$ as the concatenation/continuation of these two controls to the entire interval $[s, t_f]$, i.e.

$$\mathbf{u}(t) = \begin{cases} \hat{\mathbf{u}}(t) & t \in [s, \tau], \\ \mathbf{u}'(t) & t \in [\tau, t_f]. \end{cases}$$

Combining these estimates together we find that

$$V(s, \mathbf{y}) \leq J(s, \mathbf{y}, \mathbf{u}) \leq J_\tau + 2\varepsilon.$$

As this is done for an arbitrary $\varepsilon > 0$ then we have that $V(s, \mathbf{y}) \leq J_\tau$.

□

The principle of dynamic programming allows us to compute the solution of the optimal control problem in stages, i.e. discretization of the full problem is actually a very good idea. This provides us with a numerical approach we can use for computing solutions, as well as a theoretical tool we will use below to establish sufficient conditions for such solutions to indeed be optimal.

22.2 Hamilton-Jacobi-Bellman Equations

Although we would love to establish sufficient conditions in full generality, we are of course restricted by spatial and temporal restrictions, so we will have to resort to the following derivation which only provides sufficient conditions so long as everything is sufficiently smooth already to allow for all derivative calculations that follow to be computed without concern.

We first consider the optimal control problem described in the previous section with initial condition $\mathbf{x}(s) = \mathbf{y}$, and let $\mathbf{u}(t)$ be any admissible control (not necessarily the optimal one). If we define

$$\Phi_{\mathbf{u}}(t) = \int_s^t L(\mathbf{x}, \mathbf{u}) d\tau + V(t, \mathbf{x}), \quad t \in [s, t_f],$$

then we can note that for $t_1 < t_2$,

$$\Phi_{\mathbf{u}}(t_1) - \Phi_{\mathbf{u}}(t_2) = V(t_1, \mathbf{x}(t_1)) - \left[\int_{t_1}^{t_2} L(\mathbf{x}, \mathbf{u}) dt + V(t_2, \mathbf{x}(t_2)) \right] \leq 0,$$

or in other words $\Phi_{\mathbf{u}}(t)$ is non-decreasing as a function of time t.

We now make the key observation that the selected control $\mathbf{u}$ is optimal if and only if it achieves the minimum cost, i.e.

$$\Phi_{\mathbf{u}}(s) = V(s, \mathbf{y}) = \int_s^{t_f} L(\mathbf{x}, \mathbf{u}) dt + \phi(\mathbf{x}(t_f)) = \Phi_{\mathbf{u}}(t_f).$$

Combining this with the monotonicity of the function $\Phi_{\mathbf{u}}(t)$ we have the following proposition.

Proposition 22.2.1. *The control $\mathbf{u}(t)$ is optimal if and only if $\Phi_{\mathbf{u}}(t)$ is constant in t.*

This gives us a route to identifying a sufficient condition for the control $\mathbf{u}(t)$ to be optimal. Rather than stating the final result here and proceeding as if proving a theorem, we will constructively build the result in the following steps in the hopes that this will be more intuitive to the reader.

(i) To begin, we fix some initial data $(s, \mathbf{y})$ and assume that the control $\mathbf{u}(t)$ defined on the interval $[s, t_f]$ is admissible for the problem described in the previous section. If we additionally assume that $\mathbf{u}(t)$ is optimal for this problem, and therefore $\Phi_{\mathbf{u}}(t)$ is constant then assuming that the value function is differentiable and $\mathbf{u}(t)$ is continuous everywhere (as mentioned at the beginning of this section, this is not always true, but would require some extra caution to deal with) then we can see that

$$\begin{aligned} 0 &= \frac{d}{d\tau}\left[\int_s^\tau L(\mathbf{x}(t), \mathbf{u}(t))dt + V(\tau, \mathbf{x}(\tau)\right] \\ &= L(\mathbf{x}(\tau), \mathbf{u}(\tau)) + V_t(\tau, \mathbf{x}(\tau)) + \nabla V(\tau, \mathbf{x}(\tau)) \cdot \mathbf{f}(\mathbf{x}(\tau), \mathbf{u}(\tau)), \end{aligned}$$

where $\mathbf{x}' = \mathbf{f}$, and the ∇ operator here refers to a gradient along the $\mathbf{x}$ variable. If we consider the particular case of $\tau = s$ then we arrive at

$$L(\mathbf{y}, \mathbf{u}(s)) + V_t(s, \mathbf{y}) + \nabla V(s, \mathbf{y}) \cdot \mathbf{f}(\mathbf{y}, \mathbf{u}(s)) = 0.$$

(ii) If instead we chose an arbitrary admissible (but constant in this case) control $\mathbf{v}$ then $\Phi_{\mathbf{v}}(t)$ would be non-decreasing so that

$$\begin{aligned} 0 &\le \frac{d}{d\tau}\left[\int_s^\tau L(\mathbf{x}(t), \mathbf{v})dt + V(\tau, \mathbf{x}(\tau))\right] \\ &= L(\mathbf{x}(\tau), \mathbf{v}) + V_t(\tau, \mathbf{x}(\tau)) + \nabla V(\tau, \mathbf{x}(\tau)) \cdot \mathbf{f}(\mathbf{x}(\tau), \mathbf{v}). \end{aligned}$$

In particular if we consider the specific value $\tau = s$ then we see that

$$L(\mathbf{y}, \mathbf{v}) + V_t(s, \mathbf{y}) + \nabla V(s, \mathbf{y}) \cdot \mathbf{f}(\mathbf{y}, \mathbf{v}) \ge 0. \tag{22.7}$$

(iii) Combining the previous two steps, we note that

$$\min_{\mathbf{v}\in U}\{L(\mathbf{y}, \mathbf{v}) + V_t(s, \mathbf{y}) + \nabla V(s, \mathbf{y}) \cdot \mathbf{f}(\mathbf{y}, \mathbf{v})\} = 0, \tag{22.8}$$

and

$$\mathbf{u}(s) = \mathbf{u}^*(s, \mathbf{y}) \in \operatorname*{argmin}_{\mathbf{v}\in U}\{L(\mathbf{y}, \mathbf{v}) + \nabla V(s, \mathbf{y}) \cdot \mathbf{f}(\mathbf{y}, \mathbf{v})\}. \tag{22.9}$$

Note that this minimum is achieved at exactly the same point as the maximum of the Hamiltonian $H = \mathbf{p} \cdot \mathbf{f} - L$ if we allow the substitution $p \to -\nabla V$.

In addition if the minimum above is achieved uniquely then this procedure will uniquely determine the optimal control $\mathbf{u}(s, \mathbf{y})$.

(iv) Continuing the analogy introduced above, we re-introduce the Hamiltonian, but now replacing the co-state with the negative gradient of the value function. This leads to what is referred to as the Hamilton-Jacobi-Bellman (HJB) equation which provides a sufficient condition for the optimal control $\mathbf{u}$

$$V_t(s, \mathbf{y}) = \sup_{\mathbf{u}} H(s, \mathbf{x}, \mathbf{u}, -\nabla V). \tag{22.10}$$

with endpoint condition

$$V(t_f, \mathbf{y}) = \phi(\mathbf{y}). \tag{22.11}$$

Remark 22.2.2. The derivation provided above makes two critical assumptions which it turns out are not necessary but are used here for readability and ease of discussion. These two assumptions are:

(i) The partial derivative $\partial_t V$ and the gradient ∇V all exist at all values of $(s, \mathbf{y})$.

(ii) An optimal control exists for the optimal control problem that is continuous at the initial time s.

It turns out that neither of these assumptions are necessary to show that any control yielding a value function that satisfies the HJB, will be optimal. Without guaranteeing the existence of the partial derivatives though, we are restricted to considering what are called viscosity solutions which are an extension of the classical solution of partial differential equations.

Remark 22.2.3. In summary, the value function which is a solution of the HJB equation will give the optimal cost of the formulated optimal control problem. The corresponding optimal solution $\tilde{\mathbf{x}}(t)$ and $\tilde{\mathbf{u}}(t)$, if found locally are necessarily the optimizers, and if found globally, are necessarily and sufficiently the optimizers of the problem. In other words, the HJB equation gives both sufficient and necessary conditions for a global optimizer of the corresponding optimal control problem.

Typically one identifies a potential optimal solution using Pontraygin's Maximum Principle and then this optimal control $\mathbf{u}(t)$ and corresponding state $\mathbf{x}(t)$ are checked against the HJB equation. If HJB is satisfied then we are guaranteed that this is indeed the unique optimal solution.

Example 22.2.4. Returning to the cancer treatment problem, we will consider the optimal control problem with state equation and cost functional given by:

$$x' = \alpha x - u, \quad x(0) = x_0,$$
$$J[u] = x(t_f) + \int_0^{t_f} u^2 dt.$$

We can calculate

$$\begin{aligned} -V_t(t,x) &= \inf_{u \in U} \left\{ L + \frac{DV}{Dx} \cdot f \right\} \\ &= \inf_{u \in U} \{u^2 + V_{x'}(\alpha x - u)\}, \quad V(t_f, x) = x. \end{aligned}$$

The part inside the infimum above is convex in u, and hence has a unique minimum: $2u - V_x = 0$ implying that $\tilde{u} = \frac{1}{2}V_x$. This in turn implies that

$$-V_t(t,x) = \frac{1}{4}V_x^2 + V_x(\alpha x - \frac{1}{2}V_x) = -\frac{1}{4}V_x^2 + \alpha x V_x.$$

Then as our final statement of the HJB, we have

$$-V_t(t,x) = -\frac{1}{4}V_x^2 + \alpha x V_x,$$

with $V(t_f, x) = x$. Solutions of this PDE would yield the unique optimizer $\tilde{u} = \frac{1}{2}V_x$. As one can see though, this is not a simple process, nor is it very tractable.

To work through something that actually is tractable, we consider a much more simplified problem which we can easily solve via Pontryagin if we wanted to. In practice, even interpreting the optimal control solution computed from Pontryagin into the HJB formulation is a complicated process.

Example 22.2.5. Consider the simple example where $x' = u$ with $x(0) = 1$ and we are trying to minimize

$$J[u] = \int_0^{t_f} u^2 dt + \frac{1}{2}x(t_f)^2.$$

This leads to the Hamiltonian defined by $H = up - u^2$ and hence the HJB is given by:

$$-V_t = \inf_u \left\{u^2 + V_x u\right\} = -\frac{1}{4}V_x^2,$$

where the optimal control $\tilde{u} = -\frac{1}{2}V_x$, and we have the additional endpoint condition $V(t_f) = \frac{1}{2}x(t_f)^2$. For this particular problem, we note that the function $V(t,x) = -xx(t_f) + \frac{t}{4}x(t_f)^2 + x(t_f)^2 + \frac{2-t_f}{4}x(t_f)^2$ satisfies the HJB equation. This means that the optimal control is given by the constant $\tilde{u} = \frac{1}{2}x(t_f)^2$. We can use this to find the actual evolution of the state, and the value of $x(t_f)$.

Remark 22.2.6. If the previous solution for $V(t,x)$ wasn't obvious to you, don't be alarmed. Solutions of nonlinear PDEs typically aren't. Such a solution is more easily engineered from Pontryagin's maximum principle, but even this process is most definitely nontrivial.

Remark 22.2.7. In general the HJB equation is convenient theoretically because it yields both necessary and sufficient conditions for the existence of a unique (if the solution to the corresponding PDE is unique) minimizer, however as indicated by this simple example above these solutions are typically very hard to come by. In practice, the HJB is almost certainly a nonlinear PDE that in practice can at best be approximated numerically. Even then, we have omitted the messier complication that in practice the optimal control $\tilde{\mathbf{u}}$ is most often not continuous, and the optimal state trajectory need not be differentiable, indicating that the PDE form of the HJB need only be understood in the weak sense, i.e. we must consider weak solutions in order to truly determine the optimal solution. These weak solutions are often very hard to come by and frequently the rigorous search for the solution to an optimal control problem invokes the theory of non-smooth partial differential equations, a field unto itself. This is typically done through the theory of viscosity solutions.

Example 22.2.8. Consider the HJB equation for the Linear Quadratic Regulator, i.e.

$$-V_t(t,\mathbf{x}) = \inf_{\mathbf{u}} \left\{\mathbf{x}^T Q\mathbf{x} + \mathbf{u}^T R\mathbf{u} + \frac{DV}{D\mathbf{x}} \cdot (A\mathbf{x} + B\mathbf{u})]\right\},$$

where the infimum is given by

$$\tilde{\mathbf{u}} = -\frac{1}{2}R^{-1}B^T\frac{DV}{D\mathbf{x}},$$

implying that the value function satisfies

$$-V_t = \mathbf{x}^T Q\mathbf{x} + \frac{DV}{D\mathbf{x}} \cdot A\mathbf{x} - \frac{1}{4}\left(\frac{DV}{D\mathbf{x}}\right)^T BR^{-1}B^T\frac{DV}{D\mathbf{x}},$$

with $V(t_f, \mathbf{x}) = \mathbf{x}^T M\mathbf{x}$. It turns out that this will work so long as $\frac{1}{2}\frac{DV}{D\mathbf{x}} = P\mathbf{x}$ where P is the Riccatti matrix as demonstrated previously.

Exercises

Note to the student: Each section of this chapter has several corresponding exercises, all collected here at the end of the chapter. The exercises between the first and second line are for Section 1, the exercises between the second and third lines are for Section 2, and so forth.

You should **work every exercise** (your instructor may choose to let you skip some of the advanced exercises marked with *). We have carefully selected them, and each is important for your ability to understand subsequent material. Many of the examples and results proved in the exercises are used again later in the text. Exercises marked with ⚠ are especially important and are likely to be used later in this book and beyond. Those marked with † are harder than average, but should still be done.

Although they are gathered together at the end of the chapter, we strongly recommend you do the exercises for each section as soon as you have completed the section, rather than saving them until you have finished the entire chapter.

22.1. Prove that the value function is indeed Lipschitz continuous. *HINT: Use an argument similar to the proof of the dynamic programming principle, that is you will want to make use of the existence of a non-optimal control* $\mathbf{u}(t)$.

22.2. Find the value function $V(s, y)$ for the optimal control problem of Example 21.1.5 from the previous Chapter.

22.3. Determine the Hamilton-Jacobi-Bellman equation (with boundary condition) for the car-driving problem, that is

$$x_1' = x_2 \qquad x_2' = u \qquad x_1(0) = -a, \quad x_2(0) = 0$$
$$J = \int_0^{t_f} dt, \quad U_l \le u \le U_u.$$

You will probably need to use the sign function to make this work.

22.4. Determine the HJB equation with endpoint condition for the cash balance problem, i.e.

$$\min_{u(t)} \left\{-x(t_f) - y(t_f)\right\},$$
$$x' = r_1 x - d + u - \alpha|u|, \quad x(0) = x_0$$
$$y' = r_2 y - u, \quad y(0) = y_0.$$

22.5. Show that the proposed value function $V(t, x)$ in Example 22.2.5 is indeed a solution of the HJB equation with the prescribed boundary condition.

22.6. Using the explicit state evolution equation in Example 22.2.5 and the described optimal control, solve for $x(t)$ and determine $x(t_f)$.

Notes