Deterministic Dynamic Programming

1 Value Function

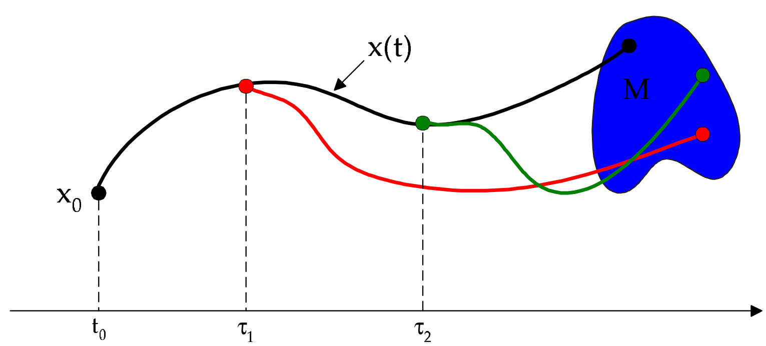
Consider the following optimal control problem in Mayer's form:

$$
\begin{aligned}
& V(t_0, x_0) = \inf_{u \in \mathcal{U}} J(t_1, x(t_1)) && (1)\\
\text{subject to} \quad & \dot{x}(t) = f(t, x(t), u(t)), \quad x(t_0) = x_0 \quad \text{(state dynamics)} && (2)\\
& (t_1, x(t_1)) \in M \quad \text{(boundary conditions).} && (3)
\end{aligned}
$$

The *terminal set* M is a closed subset of $\mathbb{R}^{n+1}$. The admissible control set $\mathcal{U}$ is assumed to be the set of piecewise continuous function on $[t_0, t_1]$. The performance function J is assumed to be C^1. The function $V(\cdot, \cdot)$ is called the *value function* and we shall use the convention $V(t_0, x_0) = \infty$ if the control problem above admits no feasible solution. We will denote by $\mathcal{U}(x_0, t_0)$, the set of feasible controls with initial condition (x_0, t_0), that is, the set of control u such that the corresponding trajectory x satisfies $x(t_1) \in M$.

Proposition 1 *Let $u(t) \in \mathcal{U}(x_0, t_0)$ be a feasible control and $x(t)$ the corresponding trajectory. Then, for any $t_0 \leq \tau_1 \leq \tau_2 \leq t_1$, $V(\tau_1, x(\tau_1)) \leq V(\tau_2, x(\tau_2))$. That is, the value function is a nondecreasing function along any feasible trajectory.*

Proof:



Corollary 1 *The value function evaluated along any optimal trajectory is constant.*

Proof: Let u^* be an optimal control with corresponding trajectory x^*. Then $V(t_0, x_0) = J(t_1, x^*(t_1))$. In addition, for any $t \in [t_0, t_1]$ u^* is a feasible control starting at $(t, x^*(t))$ and so $V(t, x^*(t)) \leq J(t_1, x^*(t_1))$. Finally by proposition (1) $V(t_0, x_0) \leq V(t, x^*(t))$ so we conclude $V(t, x^*(t)) = V(t_0, x_0)$ for all $t \in [t_0, t_1]$. ∎

According to the previous results a necessary condition for optimality is that the value function is constant along the optimal trajectory. The following result provides a sufficient condition.

Theorem 1 *Let $W(s, y)$ be an extended real valued function defined on $\mathbb{R}^{n+1}$ such that $W(s, y) = J(s, y)$ for all $(s, y) \in M$. Given an initial condition (t_0, x_0), suppose that for any feasible trajectory $x(t)$, the function $W(t, x(t))$ is finite and nondecreasing on $[t_0, t_1]$. If u^* is a feasible control with corresponding trajectory x^* such that $W(t, x^*(t))$ is constant then u^* is optimal and $V(t_0, x_0) = W(t_0, x_0)$.*

Proof: For any feasible trajectory x, $W(t_0, x_0) \leq W(t_1, x(t_1)) = J(t_1, x(t_1)$. On the other hand, for x^*,$W(t_0, x_0) = W(t_1, x^*(t_1)) = J(t_1, x^*(t_1)$. ∎

Corollary 2 *Let u^* be an optimal control with corresponding feasible trajectory x^*. Then the restriction of u^* to $[t, t_1]$ is an optimal for the control problem with initial condition $(t, x^*(t))$.*

In many applications, the control problem is given in its Lagrange form

$$V(t_0, x_0) = \inf_{u \in \mathcal{U}(x_0, t_0)} \int_{t_0}^{t_1} L(t, x(t), u(t))\, \mathrm{d}t \tag{4}$$

$$\text{subject to} \qquad \dot{x}(t) = f(t, x(t), u(t)), \qquad x(t_0) = x_0. \tag{5}$$

In this case, the following result is the analogue to proposition (1)

Theorem 2 (Bellman's Principle of Optimality). *Consider an optimal control problem in Lagrange form. For any $u \in \mathcal{U}(s, y)$ and its corresponding trajectory x*

$$V(s, y) \leq \int_s^\tau L(t, x(t), u(t))\, dt + V(\tau, x(\tau)).$$

Proof: Given $u \in \mathcal{U}(s, y)$, let $\tilde{u} \in \mathcal{U}(\tau, x(\tau))$ be arbitrary. Define

$$\bar{u}(t) = \begin{cases} u(t) & s \leq t \leq \tau \\ \tilde{u}(t) & \tau \leq t \leq t_1. \end{cases}$$

Thus, $\bar{u} \in \mathcal{U}(s, y)$ so that

$$V(s, y) \leq \int_s^{t_1} L(t, \bar{x}(t), \bar{u}(t))\, \mathrm{d}t = \int_s^\tau L(t, x(t), u(t))\, \mathrm{d}t + \int_\tau^{t_1} L(t, \tilde{x}(t), \tilde{u}(t))\, \mathrm{d}t. \tag{6}$$

Since the inequality holds for any $\tilde{u} \in \mathcal{U}(\tau, x(\tau))$ we conclude

$$V(s,y) \leq \int_s^\tau L(t, x(t), u(t))\, \mathrm{d}t + V(\tau, x(\tau)). \quad \blacksquare$$

Although the conditions given by theorem (1) are sufficient, they do not provide a concrete way to construct an optimal solution. In the next section, we will provide a direct method to compute the value function.

2 DP's Partial Differential Equations

Define $\mathcal{Q}_0$ the *reachable set* as

$$\mathcal{Q}_0 = \{(s,y) \in \mathbb{R}^{n+1} \, : \, \mathcal{U}(s,y) \neq \emptyset\}.$$

This set define the collection of initial conditions for which the optimal control problem is feasible.

Theorem 3 *Let (s,y) be any interior point of $\mathcal{Q}_0$ at which $V(s,y)$ is differentiable. Then $V(s,y)$ satisfies*

$$V_s(s,y) + V_y(s,y)\, f(s,y,v) \geq 0 \quad \textit{for all } v \in U.$$

If there is an optimal $u^* \in \mathcal{U}(s,y)$, then the PDE

$$\min_{v \in U} \{V_s(s,y) + V_y(s,y)\, f(s,y,v)\} = 0$$

is satisfied and the minimum is achieved by the right limit $u^*(s)^+$ of the optimal control at s.

Proof: Pick any $v \in U$ and let $x_v(t)$ be the corresponding trajectory for $s \leq t \leq s+\epsilon$, $\epsilon > 0$ small. Given the initial condition (s,y), we define the feasible control u_ϵ as follows

$$u_\epsilon(t) = \begin{cases} v & s \leq t \leq s+\epsilon \\ \tilde{u}(t) & s+\epsilon \leq t \leq t_1. \end{cases}$$

Where $\tilde{u} \in \mathcal{U}(s+\epsilon, x_v(s+\epsilon))$ is arbitrary. Note that for ϵ small $(s+\epsilon, x_v(s+\epsilon)) \in \mathcal{Q}_0$ and so $u_\epsilon \in \mathcal{U}(s,y)$. We denote by $x_\epsilon(t)$ the corresponding trajectory. By proposition (1), $V(t, x_\epsilon(t))$ is nondecreasing, hence,

$$D^+V(t, x_\epsilon(t)) := \lim_{h \downarrow 0} \frac{V(t+h, x_\epsilon(t+h)) - V(t, x_\epsilon(t))}{h} \geq 0$$

for any t at which the limit exists, in particular $t = s$. Thus, from the chain rule we get

$$D^+V(s, x_\epsilon(s)) = V_s(s,y) + V_y(s,y)\, D^+x_\epsilon(s) = V_s(s,y) + V_y(s,y)\, f(s,y,v).$$

The equalities use the indentity $x_\epsilon(s) = y$ and the system dynamic equation $D^+ x_\epsilon(t) = f(t, x_\epsilon, u_\epsilon(t)^+)$.

If $u^* \in \mathcal{U}(s, y)$ is an optimal control with trajectory x^* then corollary 1 implies $V(t, x^*(t)) = J(t_1, x^*(t_1))$ for all $t \in [s, t_1]$, so differentiating (from the right) this equality at $t = 2$ we conclude

$$V_s(s, y) + V_y(s, y)\, f(s, y, u^*(s)^+) = 0. \quad \blacksquare$$

Corollary 3 (Hamilton-Jacobi-Bellman equation (HJB)) *For a control problem given in Lagrange form (4)-(5), the value function at a point $(s, y) \in int(\mathcal{Q}_0)$ satisfies*

$$V_s(y, s) + V_y(s, y)\, f(s, y, v) + L(s, y, v) \geq 0 \quad \text{for all } v \in U.$$

If there exists an optimal control u^* then the PDE

$$\min_{v \in U} \{V_s(y, s) + V_y(s, y)\, f(s, y, v) + L(s, y, v)\} = 0$$

is satisfied and the minimum is achieved by the right limit $u^*(s)^+$ of the optimal control at s.

In many applications, instead of solving the HJB equation a candidate for the value function is identified, say by inspection. It is important to be able to decide whether or not the proposed solution is in fact optimal.

Theorem 4 (Verification Theorem) *Let $W(s, y)$ be a C^1 solution to the partial differential equation*

$$\min_{v \in U} \{V_s(s, y) + V_y(s, y)\, f(s, y, v)\} = 0$$

with boundary condition $W(s, y) = J(s, y)$ for all $(s, y) \in M$. Let $(t_0, x_0) \in \mathcal{Q}_0$, $u \in \mathcal{U}(t_0, x_0)$ and x the corresponding trajectory. Then, $W(t, x(t))$ is nondecreasing on t. If u^* is a control in $\mathcal{U}(t_0, x_0)$ defined on $[t_0, t_1^*]$ with corresponding trajectory x^* such that for any $t \in [t_0, t_1^*]$

$$W_s(t, x^*(t)) + W_y(t, x^*(t))\, f(t, x^*(t), u^*(t)) = 0$$

then u^* is an optimal control in $calU(t_0, x_0)$ and $V(s, y) = W(s, y)$.

Example 1:

$$\min_{\|u\| \leq 1} \; J(t_0, x_0, u) = \frac{1}{2}(x(\tau))^2$$
$$\text{subject to} \qquad \dot{x}(t) = u(t), \qquad x(t_0) = x_0$$

where $\|u\| = \max_{0 \leq t \leq \tau}\{|u(t)|\}$. The HJB equation is $\min_{|u| \leq 1}\{V_t(t, x) + V_x(t, x)\, u\} = 0$ with boundary condition $V(\tau, x) = \frac{1}{2}x^2$. We can solve this problem by inspection. Since the only cost is associated to the terminal state $x(\tau)$, and optimal control will try to make $x(\tau)$ as close to zero as possible, *i.e.*,

$$u^*(t, x) = -\text{sgn}(x) = \begin{cases} 1 & x < 0 \\ 0 & x = 0 \\ -1 & x > 0. \end{cases} \qquad \text{(Bang-Bang policy)}$$

We should now verify that u^* is in fact an optimal control. Let $J^*(t,x) = J(t,x,u^*)$. Then, it is not hard to show that

$$J^*(t,x) = \frac{1}{2}(\max\{0\,;\,|x| - (\tau - t)\})^2$$

which satisfies the boundary condition $J^*(\tau, x) = \frac{1}{2}x^2$. In addition,

$$J_t^*(t,x) = (|x| - (\tau - t))^+ \quad \text{and} \quad J_x^*(t,x) = \mathsf{sgn}(x)\,(|x| - (\tau - t))^+.$$

Therefore, for any u such that $|u| \le 1$ it follows that

$$J_t^*(t,x) + J_x^*(t,x)\,u = (1 + \mathsf{sgn}(x)\,u)\,(|x| - (\tau - t))^+ \ge 0$$

with the equality holding for $u = u^*(t,x)$. Thus, $J^*(t,x)$ is the value function and u^* is optimal.

3 Feedback Control

In the previous example, the notion of a *feedback control* policy was introduced. Specifically, a feedback control $\mathbf{u}$ is a mapping from $\mathbb{R}^{n+1}$ to U such that $\mathbf{u} = \mathbf{u}(t,x)$ and the system dynamics $\dot{x} = f(t,x,\mathbf{u}(t,x))$ has a unique solution for each initial condition $(s,y) \in \mathcal{Q}_0$. Given a feedback control $\mathbf{u}$ and an initial condition (s,y), we can define the trajectory $x(t;s,y)$ as the solution to

$$\dot{x} = f(t,x,\mathbf{u}(t,x)) \qquad x(s) = y.$$

The corresponding control policy is $u(t) = \mathbf{u}(t, x(t;s,y))$.

A feedback control $\mathbf{u}^*$ is an *optimal feedback control* if for any $(s,y) \in \mathcal{Q}_0$ the control $u(t) = \mathbf{u}^*(t, x(t;s,y))$ solve the optimization problem (1)-(3) with initial condition (s,y).

Theorem 5 *If there is an optimal feedback control* $\mathbf{u}^*$ *and* $t_1(s,y)$ *and* $x(t_1;s,y)$ *are the terminal time and terminal state for the trajectory*

$$\dot{x} = f(t,x,\mathbf{u}(t,x)) \qquad x(s) = y$$

then the value function $V(s,y)$ *is differentiable at each point at which* $t_1(s,y)$ *and* $x(t_1;s,y)$ *are differentiable with respect to* (s,y).

Proof: From the optimality of $\mathbf{u}^*$ we have that

$$V(s,y) = J(t_1(s,y), x(t_1(s,y);s,y)).$$

The result follows from this identity and the fact that J is C^1. ∎

4 The Linear-Quadratic Problem

Consider the following optimal control problem.

$$\min x(T)'\, Q_T\, x(T) + \int_0^T \big[x(t)'\, Q\, x(t) + u(t)'\, R\, u(t)\big]\, \mathrm{d}t \tag{7}$$

$$\text{subject to} \qquad \dot{x}(t) = A\, x(t) + B\, u(t) \tag{8}$$

where the $n \times n$ matrices Q_T and Q are symmetric positive semidefinite and the $m \times m$ matrix R is symmetric positive definite. The HJB equation for this problem is given by

$$\min_{u \in \mathbb{R}^m} \big\{V_t(t,x) + V_x(t,x)'\,(Ax + Bu) + x'\, Q\, x + u'\, R\, u\big\} = 0$$

with boundary condition $V(T, x) = x'\, Q_T\, x$.

We guess a quadratic solution for the HJB equation. That is, we suppose that $V(t,x) = x'\, K(t)\, x$ for a $n \times n$ symmetric matrix $K(t)$. If this is the case then

$$V_t(t,x) = 2K(t)\, x \quad \text{ and } \quad V_x(t,x) = x'\, \dot{K}(t)\, x.$$

Plugging back these derivatives on the HJB equation we get

$$\min_{u \in \mathbb{R}^m} \Big\{x'\, \dot{K}(t)\, x + 2x'\, K(t) A\, x + 2x'\, K(t) B\, u + x'\, Q\, x + u'\, R\, u\Big\} = 0. \tag{9}$$

Thus, the optimal control satisfies

$$2B'K(t)\, x + 2R\, u = 0 \qquad \Longrightarrow \qquad u^* = -R^{-1}B'K(t)\, x.$$

Substituting the value of u^* in equation (9) we obtain the condition

$$x' \, \Big(\dot{K}(t) + K(t)A + A'K(t) - K(t)BR^{-1}B'K(t) + Q\Big)\, x = 0 \quad \text{ for all } (t,x).$$

Therefore, for this to hold matrix $K(t)$ must satisfy the *continuous-time Ricatti equation* in matrix form

$$\dot{K}(t) = -K(t)A - A'K(t) = K(t)BR^{-1}B'K(t) - Q, \quad \text{ with boundary condition } K(T) = Q_T. \tag{10}$$

Reversing the argument it can be shown that if $K(t)$ solves (10) then $W(t,x) = x'K(t)x$ is a solution of the HJB equation and so nt the verification theorem we conclude that it is equal to the value function. In addition, the optimal feedback control is $\mathbf{u}^*(t,x) = -R^{-1}B'K(t)x$.

5 The Method of Characteristics for First-Order PDEs

5.1 First-Order Homogeneous Case

Consider the following first-order homogeneous PDE

$$u_t(t,x) + a(t,x)u_x(t,x) = 0, \qquad x \in \mathbb{R}, t > 0,$$

with boundary conditions $u(x,0) = \phi(x)$ for all $x \in \mathbb{R}$. We assume that a and ϕ are "smooth enough" functions. A PDE problem in this form is referred to as a Cauchy problem.

We will investigate the solution to this problem using the method of characteristics. The characteristics of this PDE are curves in the $x - t$ plane defined by

$$\dot{x}(t) = a(x(t), t), \qquad x(0) = x_0. \tag{11}$$

Let $\tilde{x} = \tilde{x}(t)$ be a solution with $\tilde{x}(0) = x_0$. Let u be a solution to the PDE, we want to study the evolution of u along $\tilde{x}(t)$.

$$\dot{u}(t, \tilde{x}(t)) = u_t(t, \tilde{x}(t)) + u_x(t, \tilde{x}(t))\, \dot{\tilde{x}}(t) = u_t(t, \tilde{x}(t)) + u_x(t, \tilde{x}(t))\, a(\tilde{x}(t), t) = 0.$$

So, $u(t,x)$ is constant along the characteristic curve $\tilde{x}(t)$, that is,

$$u(t, \tilde{x}(t)) = u(0, \tilde{x}(0)) = \phi(x_0), \qquad \forall t > 0. \tag{12}$$

Thus, if we are able to solve the ODE (13) then we would be able to find the solution to the original PDE.

Example 2: Consider the Cauchy problem

$$\begin{aligned} u_t + x\ u_x &= 0, \qquad x \in \mathbb{R}, t > 0 \\ u(x, o) &= \phi(x), \qquad x \in \mathbb{R}. \end{aligned}$$

The characteristic curves are defined by

$$\dot{x}(t) = x(t),\ x(0) = x_0,$$

so $x(t) = x_0 \exp(t)$. So for a given (t,x) the characteristic passing through this point has initial condition $x_0 = x \exp(-t)$. Since $u(t, x(t)) = \phi(x_0)$ we conclude that $u(t,x) = \phi(x \exp(-t))$.

5.2 First-Order NonHomogeneous Case

Consider the following nonhomogeneous problem.

$$
\begin{aligned}
&u_t(t,x) + a(t,x)\ u_x(t,x) = b(t,x), \qquad x \in \mathbb{R}, t > 0 \\
&u(x,0) = \phi(x), \qquad x \in \mathbb{R}.
\end{aligned}
$$

Again, the characteristic curves are given by

$$
\dot{x}(t) = a(x(t),t), \qquad x(0) = x_0. \tag{13}
$$

Thus, for a solution $u(t,x)$ of the PDE along a characteristic curve $\tilde{x}(t)$ we have that

$$
\dot{u}(t,\tilde{x}(t)) = u_t(t,\tilde{x}(t)) + u_x(t,\tilde{x}(t))\,\dot{\tilde{x}}(t) = u_t(t,\tilde{x}(t)) + u_x(t,\tilde{x}(t))\,a(\tilde{x}(t),t) = b(t,\tilde{x}(t)).
$$

Hence, the solution to the PDE is given by

$$
u(t,\tilde{x}(t)) = \phi(x_0) + \int_0^t b(\tau,\tilde{x}(\tau))\,\mathrm{d}\tau
$$

along the characteristic $(t,\tilde{x}(t))$.

Example 3: Consider the Cauchy problem

$$
\begin{aligned}
&u_t +\ u_x = x, \qquad x \in \mathbb{R}, t > 0 \\
&u(x,o) = \phi(x), \qquad x \in \mathbb{R}.
\end{aligned}
$$

The characteristic curves are defined by

$$
\dot{x}(t) = 1,\ x(0) = x_0,
$$

so $x(t) = x_0 + t$. So for a given (t,x) the characteristic passing through this point has initial condition $x_0 = x - t$. In addition, along a characteristic $\tilde{x}(t) = x_0 + t$ starting at x_0, we have

$$
u(t,\tilde{x}(t)) = \phi(x_0) + \int_0^t \tilde{x}(\tau)\,\mathsf{d}\tau = \phi(x_0) + x_0\,t + \frac{1}{2}t^2.
$$

Thus, the solution to the PDE is given by

$$
u(t,x) = \phi(x-t) + \left(x - \frac{t}{2}\right)\ t.
$$

5.3 Applications to Optimal Control

Given that the partial differential equation of dynamic programming is a first-order PDE, we can try to apply the method of characteristic to find the value function. In general, the HJB is not a standard first-order PDE because of the maximization that takes place. So in general, we can not just solve a simple first-order PDE to get the value function of dynamic programming. Nevertheless, in some situations it is possible to obtain good results as the following example shows.

Example 1:(Method of Characteristics) Consider the optimal control problem

$$\min_{\|u\|\leq 1} \; J(t_0, x_0, u) = \frac{1}{2}(x(\tau))^2$$
$$\text{subject to} \qquad \dot{x}(t) = u(t), \qquad x(t_0) = x_0$$

where $\|u\| = \max_{0\leq t\leq \tau}\{|u(t)|\}$.

A candidate for value function $W(t,x)$ should satisfy the HJB equation

$$\min_{|u|\leq 1} \; \{W_t(t,x) + W_x(t,x)\,u\} = 0,$$

with boundary condition $W(\tau, x) = \frac{1}{2}x^2$.

For a given $u \in U$, let solve the PDE

$$W_t(t,x;u) + W_x(t,x;u)\,u = 0, \qquad W(\tau,x;u) = \frac{1}{2}x^2. \tag{14}$$

A characteristic curve $\tilde{x}(t)$ is found solving

$$\dot{x}(t) = u, \qquad x(0) = x_0,$$

so $\tilde{x}(t) = x_0 + ut$. Since the solution to the PDE is constant along the characteristic curve we have

$$W(t,\tilde{x}(t);u) = W(\tau,\tilde{x}(\tau);u) = \frac{1}{2}(x(\tau))^2 = \frac{1}{2}(x_0 + u\tau)^2.$$

The characteristic passing through the point (t,x) has initial condition $x_0 = x - ut$, so the general solution to the PDE (14) is

$$W(t,x;u) = \frac{1}{2}(x + (\tau - t)u)^2.$$

Since our objective is to minimize the terminal cost, we can identify a policy by minimizing $W(t,x;u)$ over u above. It is straightforward to see that the optimal control (in feedback form) satisfies

$$u^*(x,t) = \begin{cases} -1 & \text{if } x > \tau - t \\ \frac{-x}{\tau - t} & \text{if } |x| \leq \tau - t \\ 1 & \text{if } x < t - \tau. \end{cases}$$

The corresponding "candidate" for value function $W^*(t,x) = W(t,x;u^*(t,x))$ satisfies

$$W(t,x) = \frac{1}{2}\Big(\max\{0\,;\,|x| - (\tau - t)\}\Big)^2$$

which we already know satisfies the HJB equation.

6 Extensions

6.1 Connecting the HJB Equation with Pontryagin Principle

We consider the optimal control problem in Lagrange form. In this case, the HJB equation is given by

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

with boundary condition $V(t_1, x(t_1)) = 0$.

Let us define the so-called *Hamiltonian*

$$H(t,x,u,\lambda) := \lambda f(x,t,u) - L(t,x,u).$$

Thus, the HJB equation implies that the value function satisfies

$$\max_{u \in U} H(t,x,u,-V_x) = 0,$$

and so the optimal control can be found maximizing the Hamiltonian. Specifically, let $x^*(t)$ be the optimal trajectory and let $P(t) = -V_x(t, x^*(t))$, then the optimal control satisfies the so-called *Maximum Principle*

$$H(t, x^*(t), u^*(t), P(t)) \leq H(t, x^*(t), u, P(t)), \qquad \text{for all } u \in U.$$

In order to complete the connection with Pontryagin principle we need to derive the adjoint equations. Let $x^*(t)$ be the optimal trajectory and consider a small perturbation $x(t)$ such that

$$x(t) = x^*(t) + \delta(t), \quad \text{where } |\delta(t)| < \epsilon.$$

First, we note that the HJB equation together with the optimality of x^* and its corresponding control u^* implies that

$$H(t, x^*(t), u^*(t), -V_x(t, x^*(t))) - V_t(t, x^*(t)) \geq H(t, x(t), u^*(t), -V_x(t, x(t))) - V_t(t, x(t)).$$

Therefore, the derivative of $H(t, x(t), u^*(t), -V_x(t, x(t))) + V_t(t, x(t))$ with respect to x so be equal to zero at $x^*(t)$. Using the definition of H this condition implies that

$$-V_{xx}(t, x^*(t))\, f(t, x^*(t), u^*(t)) - V_x(t, x^*(t))\, f_x(t, x^*(t), u^*(t)) - L_x(t, x^*(t), u^*(t)) - V_{xt}(t, x^*(t)) = 0.$$

In addition, using the dynamics of the system we get that

$$\dot{V}_x(t, x^*(t)) = V_{tx}(t, x^*(t)) + V_{xx}(t, x^*(t))\, f(t, x^*(t), u^*(t)),$$

therefore

$$\dot{V}_x(t, x^*(t)) = V_x(t, x^*(t))\, f(t, x^*(t), u^*(t)) + L(t, x^*(t), u^*(t)).$$

Finally, using the definition of $P(t)$ and H we conclude that $P(t)$ satisfies the adjoint condition

$$\dot{P}(t) = \frac{\partial}{\partial x} H(t, x^*(t), u^*(t), P(t)).$$

The boundary condition for $P(t)$ are obtained from the boundary conditions of the HJB, that is,

$$P(t_1) = -V_x(t_1, x(t_1)) = 0. \quad \text{(transversality condition)}$$

6.2 Economic Interpretation of the Maximum Principle

Let us again consider the control problem in Lagrange form. In this case the performance measure is

$$V(t, x) = \min \int_0^T L(t, x(t), u(t))\, \mathrm{d}t.$$

The function L corresponds to the instantaneous "cost" rate. According to our definition of $P(t) = -V_x(t, x(t))$, we can interpret this quantity as the marginal profit associated to a small change on the state variable x. The economic interpretation of the Hamiltonian is as follows:

$$\begin{aligned} H\, \mathrm{d}t &= P(t)\, f(t, x, u)\, \mathrm{d}t - L(t, x, u)\, \mathrm{d}t \\ &= P(t)\, \dot{x}(t)\, \mathrm{d}t - L(t, x, u)\, \mathrm{d}t \\ &= P(t)\, \mathrm{d}x(t) - L(t, x, u)\, \mathrm{d}t. \end{aligned}$$

The term $-L(t, x, u)\, \mathrm{d}t$ corresponds to the instantaneous profit made at time t at state x is control u is selected. We can look at this profit as a *direct contribution.* The second term $P(t)\, \mathrm{d}x(t)$ represents the instantaneous profit that it is generated by changing the state from $x(t)$ to $x(t) + \mathrm{d}x(t)$. We can look at this profit as an *indirect contribution.* Therefore $H\, \mathrm{d}t$ can be interpreted as the *total contribution* made from time t to $t + \mathrm{d}t$ given the state $x(t)$ and the control u.

With this interpretation, the *Maximum Principle* simply state that an optimal control should try to maximize the total contribution for every time t. In other words, the Maximum Principle *decouples* the dynamic optimization problem in to a series of static optimization problem, one for every time t.

Note also that if we integrate the adjoint equation we get

$$P(t) = \int_t^{t_1} H_x\, \mathrm{d}t.$$

So P(t) is the cumulative gain obtained over $[t, t_1]$ by marginal change of the state space. In this respect, the adjoint variables behave in much the same way as dual variables in LP.