Optimization via the Hamilton-Jacobi-Bellman Method: Theory and Applications

Navin Khaneja
lecture notes taken by Christiane Koch

June 24, 2009

1 Variation yields a classical Hamiltonian system

Suppose that we have a system that is described by the following equations of motion,

$$\dot{x} = f(x,u,t) \qquad x \in \mathbb{R}^n\,,\; u \in \Omega \subset \mathbb{R}^m\,, \tag{1}$$

where x denotes the state vector, u the controls, and t time. We can include t in the state vector with

$$\dot{t} = 1\,.$$

Eq. (1) is then reduced to

$$\dot{x} = f(x,u) \qquad x \in \mathbb{R}^{n+1}\,,\; u \in \Omega \subset \mathbb{R}^m\,. \tag{2}$$

The cost function for the optimization problem is typically written as a sum of two terms,

$$J = \phi(x_f) + \int_0^T L(x,u)dt\,, \tag{3}$$

where $\phi(x_f)$ denotes the terminal cost, with $x_f = x(T)$ the state vector at final time T. The second term in Eq. (3) corresponds to the running cost. Adding zero, J can be rewritten

$$\begin{aligned} J &= \phi(x_f) + \int_0^T L(x,u)dt + \int_0^T \lambda^T\big(f(x,u) - \dot{x}\big)dt \\ &= \phi(x_f) + \int_0^T \Big\{L(x,u) + \lambda^T f(x,u) + (\dot{\lambda}^T)x\Big\}dt - \lambda^T(T)x(T) + \lambda^T(0)x(0)\,. \end{aligned}$$

The second line is obtained by integrating -$\int_0^T \lambda^T \dot{x}\, dt$ by parts.

In order to search for the optimum, we vary J,

$$\begin{aligned} \delta J &= \left.\frac{\partial \phi}{\partial x}\right|_{x_f} \delta x_f \\ &\quad + \int_0^T \left\{ \frac{\partial L}{\partial x}\delta x(t) + \frac{\partial L}{\partial u}\delta u(t) + \lambda^T \left[\frac{\partial f}{\partial x}\delta x(t) + \frac{\partial f}{\partial u}\delta u(t)\right] + \dot{\lambda}^T \delta x(t) \right\} dt \\ &\quad -\lambda^T(T)\delta x(T)\,, \end{aligned} \tag{4}$$

where we have assumed that the initial state is fixed, $\delta x(0) = 0$, and $\delta x(T) = \delta x_f$. Regrouping terms, we obtain

$$\begin{aligned}\delta J &= \left[\left.\frac{\partial \phi}{\partial x}\right|_{x_f} - \lambda^T \right] \delta x_f \\ &+ \int_0^T \left\{ \left(\frac{\partial L}{\partial x} + \lambda^T \frac{\partial f}{\partial x} \right) \delta x(t) + \left(\frac{\partial L}{\partial u} + \lambda^T \frac{\partial f}{\partial u} \right) \delta u(t) + (\dot{\lambda}^T)\delta x(t) \right\} dt\end{aligned} \tag{5}$$

We are looking for an optimum of J, i.e. the variations should vanish. The variation with respect to $x(t)$ and x_f, respectively, determine the equation of motion for the adjoint state, $\lambda^T(t)$, and its 'initial' condition: The term inside the square brackets in Eq. (5) is zero if

$$\lambda(T) = \left.\left(\frac{\partial \phi}{\partial x} \right)^T \right|_{x_f},$$

and the terms inside the integral in Eq. (5) that are multiplied with $\delta x(t)$ vanish if

$$\dot{\lambda} = -\left(\frac{\partial L}{\partial x} + \lambda^T \frac{\partial f}{\partial x} \right)^T .$$

We are then left with

$$\delta J = \int_0^T \left(\frac{\partial L}{\partial u} + \lambda^T \frac{\partial f}{\partial u} \right) \delta u(t)\, dt\,.$$

We can introduce a (Hamiltonian) function

$$H(x, \lambda, u) = L(x, u) + \lambda^T f(x, u)\,. \tag{6}$$

We then see that

$$\dot{\lambda} = -\left(\frac{\partial H}{\partial x} \right)^T ,$$

and the variation of J becomes

$$\delta J = \int_0^T \frac{\partial H}{\partial u} \delta u(t)\, dt\,.$$

A small variation of u can be written as

$$\delta u = -\varepsilon \left(\frac{\partial H}{\partial u} \right)^T .$$

If u is optimal, then at least

$$\frac{\partial H}{\partial u} = 0$$

such that $\delta J = 0$. We then obtain the following ***set of 1st order necessary conditions***,

$$\dot{x} = f(x, u) = \left(\frac{\partial H}{\partial \lambda} \right)^T \quad \text{with} \quad x(0) \tag{7}$$

$$\dot{\lambda} = -\left(\frac{\partial H}{\partial x} \right)^T \quad \text{with} \quad \lambda(T) = \left.\frac{\partial \phi}{\partial x}\right|_{x_f} \tag{8}$$

$$\left.\frac{\partial H}{\partial u}\right|_{(x(t),\lambda(t),u(t))} = 0 \tag{9}$$

This set of equations looks like a classical Hamiltonian system, where the state vector x takes the role of the coordinates and the adjoint state λ that of the conjugate momenta.

2 The maximum principle

Let us first state Pontryagin's maximum principle [1] before proving it: Along the optimal trajectory, we have

$$u^*(t) = \arg\min_u H\big(x^*(t), \lambda^*(t), u\big) \,. \tag{10}$$

Note that in our context it is rather a minimum than a maximum principle since we seek to minimize the cost. The claim is that if $u^*(t)$ is an *optimal* control, then it minimizes H along the optimal trajectory *globally* at each instant of time. This is a very strong statement since we have only used first order variation.

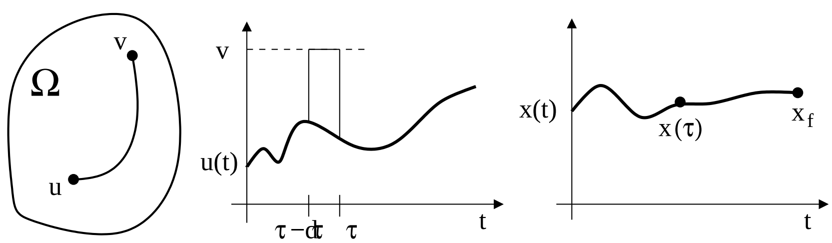
To prove this[1], let us first recast J such that it represents a final-time cost only. This can be done by defining, in Eq. (3), $L(x, u)$ in terms of a state variable. That is, if our state vector is $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, we add x_{n+1} with

$$\dot{x}_{n+1} = L(x, u) \,. \tag{11}$$

We can then write a cost functional that depends on the final time only,

$$J = \phi(x_f) + x_{n+1}(T) \,. \tag{12}$$

Now suppose that we know the optimal control $u^*(t)$ which determines the optimal $x^*(t)$. During a very short time interval, $[\tau - d\tau, \tau]$, let us change the control drastically from u^* to v:



This leads to the following change in x,

$$\delta x(\tau) = \Big[f\big(x^*(\tau), v\big) - f\big(x^*(\tau), u^*(\tau)\big) \Big] d\tau \,. \tag{13}$$

We thus obtain a change of initial condition at time $t = \tau$ for the evolution of $x(t)$ at times $t > \tau$. We can write

$$(x \dot{+} \delta x) = f(x + \delta x, u + \delta u) \,,$$

[1] Note that we will show the proof only for regular extrema. Pontryagin's proof covers also abnormal extrema [1] but is much more difficult to follow. Roughly speaking, abnormal extrema are those where the final state x_f can be reached only for a single T.

which leads to

$$\begin{aligned}
\delta x &= \underbrace{\left.\frac{\partial f}{\partial x}\right|_{(x^*(t),u^*(t))}}_{A(t)} \delta x + \underbrace{\left.\frac{\partial f}{\partial u}\right|_{(x^*(t),u^*(t))}}_{B(t)} \delta u \\
\delta x &= A(t)\delta x(t) + B(t)\delta u(t)\,. \qquad (14)
\end{aligned}$$

This can be expressed in terms of a propagator (or Green's function), Φ,

$$\delta x_f = \delta x(T) = \Phi(T,\tau)\ \delta x(\tau)\,.$$

Our claim is now

$$\left.\frac{\partial \phi}{\partial x}\right|_{x_f} \delta x_f \geq 0\,,$$

or, equivalently,

$$\left.\frac{\partial \phi}{\partial x}\right|_{x_f} \Phi(T,\tau)\ \delta x(\tau) \geq 0\,. \qquad (15)$$

Using the 'initial' condition, $\lambda(T)$, cf. Eq. (8), we can write for λ at time τ,

$$\lambda(\tau) = \left[\left.\frac{\partial \phi}{\partial x}\right|_{x_f} \Phi(T,\tau)\right]^T = \big(\Phi(T,\tau)\big)^T\ \lambda(T) \qquad (16)$$

Inserting Eq. (16) and Eq. (13) into Eq. (15), we obtain

$$\lambda^T(\tau)\Big[f\big(x^*(\tau),v\big) - f\big(x^*(\tau),u^*(t)\big)\Big]\delta\tau\ \geq\ 0$$

or, equivalently,

$$\lambda^T(\tau)\, f\big(x^*(\tau),u^*(t)\big)\ \leq\ \lambda^T(\tau)\, f\big(x^*(\tau),v\big)$$

Since Eq. (6) for a final-time only cost becomes

$$H(x,\lambda,u) = \lambda^T\ f(x,u)\,, \qquad (17)$$

we obtain for all $v \in \Omega$

$$H(x^*,\lambda^*,u^*)\ \leq\ H(x^*,\lambda^*,v)\,, \qquad (18)$$

which proves that indeed $u^*(t)$ minimizes H globally along the optimal trajectory, i.e. Eq. (10). This necessary condition is stronger than just the first order variation being zero. In fact, it is so restrictive that often it allows to determine the optimal solution. ***In conclusion, if*** $u^*(t)$ ***is the global optimum, then there is a canonical way to define*** H ***with the necessary condition, Eq. (18) and generate the state*** $x(t)$ ***and co-state*** $\lambda(t)$***.***

Eq. (17) implies

$$\frac{\partial H}{\partial x} = \lambda^T \frac{\partial f}{\partial x}\,,$$

which, together with the definition of $A(t)$, cf. Eq. (14), leads to

$$\dot{\lambda} = -\left(\frac{\partial H}{\partial x}\right)^T = -A(t)\,\lambda\,. \qquad (19)$$

An example: Controlling the force on a particle [1]

The equation of motion is given by

$$\ddot{x} = u \quad \text{with} \quad |u| \leq 1$$

which can be rewritten

$$\begin{aligned} \dot{x}_1 &= x_2\,, \\ \dot{x}_2 &= u\,. \end{aligned}$$

The goal is to drive the system from any initial point in phase space, $(x_1(0), x_2(0))$, to the origin in minimal time. We therefore have to formulate time as a cost,

$$\dot{t} = \dot{x}_3 = 1\,.$$

Evaluating Eq. (17), the equations of motion lead to the Hamiltonian function for this optimization problem,

$$H = \lambda_1 x_2 + \lambda_2 u + \lambda_3\,.$$

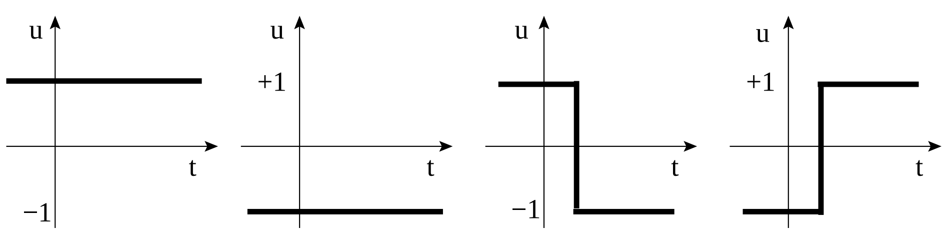
The λ_j are found from the Hamiltonian equations, cf. Eq. (8),

$$\begin{aligned} \dot{\lambda}_1 &= -\frac{\partial H}{\partial x_1} = 0 \quad \rightarrow \quad \lambda_1 = c = const \\ \dot{\lambda}_2 &= -\lambda_1 \quad \rightarrow \quad \lambda_2 = c\,t + d \\ \dot{\lambda}_3 &= 0 \end{aligned}$$

We can now determine the control from Eq. (9), or, more specifically from the condition to maximize H. Taking the restriction, $|u| \leq 1$, into account, H takes its maximum value for

$$u = -\mathrm{sgn}(\lambda_2)\,.$$

This corresponds to using the maximum allowed force, i.e. bang-bang control. Since λ_2 is a linear function, it changes sign only once. We therefore obtain four possibilities for u:



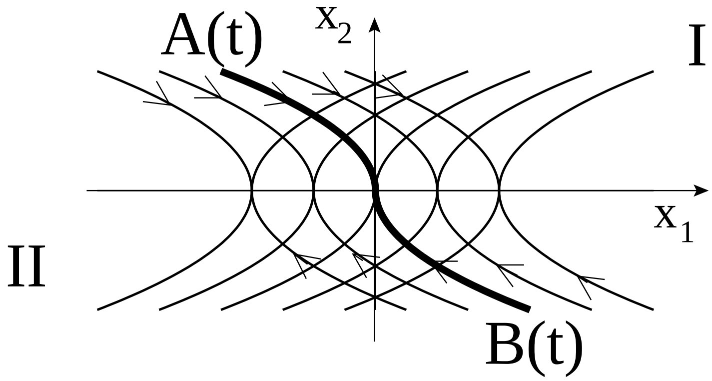
Evaluating the Hamiltonian function, Eq. (7), we can also obtain the evolution of the state variables. For $u = 1$, $\dot{x}_2 = u$, yields

$$\begin{aligned} x_2(t) &= t + c \\ x_1(t) &= \frac{t^2}{2} + c\,t + d = \frac{x_2^2}{2} + d_1\,, \end{aligned}$$

and for $u = -1$, we get correspondingly

$$x_1(t) = -\frac{x_2^2}{2} + d_1\,.$$

That is, we obtain parabolas in phase space:



If the controls are not switched, our initial point must be on either curve $A(t)$ or $B(t)$, cf. Eq. (14), since these are the only parabolas that take us to the origin. If the initial point is $(x_1(0), x_2(0)) \in \mathrm{I}$, then the control is switched from -1 to 1 when the parabola that contains $(x_1(0), x_2(0))$ and describes the flow $(x_1(t), x_2(t)$ crosses $B(t)$. The phase space flow then continues on $B(t)$ to the origin. If $(x_1(0), x_2(0)) \in \mathrm{II}$, then the control is switched from + 1 to 1 when the parabola that contains $(x_1(0), x_2(0))$ and describes $(x_1(t), x_2(t)$ crosses $A(t)$. The phase space flow then continues on $A(t)$ to the origin. We have solved this example by constructing the Hamiltonian function. Alternatively, it could also be solved by considering the final-time cost, $\phi(x_f)$, which in the example above corresponds to

$$\phi(x_{1,f}, x_{2,f}, x_{3,f}) = x_{3,f} = T\,,$$

i.e. the total time to reach the target.[2]

3 Hamilton-Jacobi-Bellman principle

When we transform all running costs into final costs $\phi(x_f)$, the optimization problem consists in designing the trajectory from initial point x_0 to final point x_f that minimizes the final cost,

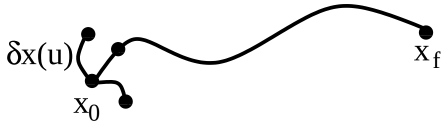
$$\min\, \phi(x_f)\,.$$

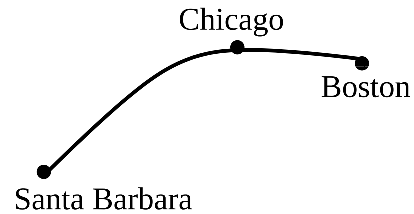
We seek to solve this problem globally for all initial points x_0 and introduce the optimal return function,

$$V(x_0) = \min\, \phi(x_f)\,, \tag{20}$$

that is $V(x)$ corresponds to the best or minimum cost starting from initial value x. Suppose that we have solved the optimization problem and know $V(x)$. This defines the optimal trajectory from x_0 to x_f:

[2]Note that $\phi(x_f)$ doesn't include the target state because we assume the target state to be fixed and the zero variations $\delta x_{1,f}$, $\delta x_{2,f}$ would leave λ_1 and λ_2 undefined.





Consider now the optimal return for small deviations from the optimal trajectory,

$$V\big(x + \delta x(u)\big)\,.$$

The claim is now that

$$\min_u V\big(x + \delta x(u)\big) = V(x)\,. \tag{21}$$

This is the ***principle of dynamic programming*** or ***Hamilton-Jacobi-Bellman (HJB) principle***. It states that if we have found the optimal way to go from Santa Barbara to Boston, and Chicago is located on this way, then also the way from Chicago to Boston is optimal. We can rewrite Eq. (21) by expressing δx in terms of the equation of motion of x, Eq. (1),

$$\min_u \left[V(x) + \frac{\partial V}{\partial x} f(x,u)\,\delta t\,, \right] = V(x)$$

which leads to the ***Bellman PDE***,

$$\min_u \left[\frac{\partial V}{\partial x} f(x,u) \right] = 0\,. \tag{22}$$

The optimal cost must satisfy Eq. (22). If we want to achieve the optimization in mininum time, we need to consider t as a state variable such that our equations of motion become

$$\dot{x} = f(x,u,t) \quad , \quad \dot{t} = 1 \tag{23}$$

Eq. (22) is then rewritten

$$\min_u \left[\frac{\partial V}{\partial t} + \frac{\partial V}{\partial x} f(x,u) \right] = 0\,, \tag{24}$$

which is the ***Hamilton-Jacobi-Bellman equation***. Note that if there is no restriction on time, then V does not depend on t, i.e. the optimum is achieved at infinite time (these problems are called infinite horizon problems).

We now show the connection between the HJB equation, Eq. (24), and the Pontryagin principle, Eq. (10), by considering what happens along the optimal trajectory. Evaluating, along the optimal trajectory, the partial derivative of $V(x)$ with respect to x which also depends on x, we find

$$\frac{\partial V}{\partial x}\Big(x^*(t)\Big) = \lambda^T(t)\,. \tag{25}$$

Inserting this into Eq. (22), we see that

$$\min_u \Big[\lambda^T(t) f\big(x^*(t), u\big) \Big] = 0\,.$$

Since the term inside the square brackets is nothing but the Hamiltonian, cf. Eq. (17), along the optimal trajectory, $x^*(t)$, we see that the optimal control minimizes H at each time. That is, we recover Pontryagin's minimum principle. We can therefore determine the optimal control as the argument that minimizes the Hamiltonian,

$$u^*(x) = \arg\min_u \Big[H\big(x^*(t), \lambda(t), u\big) \Big] . \tag{26}$$

or,

$$u^*(x) = \arg\min_u \left[\frac{\partial V}{\partial x} \, f(x, u) \right] . \tag{27}$$

Eq. (27) implies

$$\frac{\partial V}{\partial x} \, f(x, u^*(x)) = 0 . \tag{28}$$

Eq. (26) implies that the Hamiltonian is zero along the optimal trajectory, *why?*

$$H\big(x^*(t), \lambda(t), u^*(t)\big) = 0 ,$$

and that the optimal cost V is constant along the optimal trajectory,

$$\left. \frac{dV}{dt} \right|_{(x^*(t))} = 0 ,$$

since

$$\frac{dV}{dt} = H .$$

We now show that Eq. (25) indeed satisfies the equation of motion for $\lambda^T(t)$, Eq. (8). Let us denote partial derivatives by subscripts,

$$\frac{\partial V}{\partial x} = V_x .$$

Taking the derivative of Eq. (25), we obtain

$$\dot{\lambda}^T = V_{xx}\big(x^*(t)\big) \; f(x^*, u^*) .$$

On the other hand, we can differentiate Eq. (28) with respect to x and find

$$0 = \frac{\partial}{\partial x} \Big[V_x f(x, u^*) \Big] = V_{xx} \, f\big(x, u(x)\big) + V_x f_x + V_x f_u .$$

For optimal u, $f_u = 0$ and we can express V_{xx} in terms of $V_x f_x$ and obtain

$$\dot{\lambda}^T = -\frac{\partial V}{\partial x} \frac{\partial f}{\partial x} .$$

With Eq. (25), we find

$$\dot{\lambda}(t) = -\frac{\partial f}{\partial x}^T \lambda .$$

From the construction of the Hamiltonian, Eq. (17), it follows that

$$-\frac{\partial H}{\partial x} = \lambda^T \frac{\partial f}{\partial x}\,,$$

and we indeed recover Eq. (8). The connection to Pontryagin's maximum principle is easily made by considering the optimal return at the final point,

$$V(x_f) = \phi(x_f)\,.$$

This is inserted into Eq. (25) for $t = T$ and we obtain

$$\lambda^T(T) = \frac{\partial \phi}{\partial x}(x_f)\,.$$

References

[1] L. S. Pontryagin. *The Mathematical Theory of Optimal Processes.* Gordon and Breach Science Publishers, 1986.