16 Pontryagin's maximum principle

This is a powerful method for the computation of optimal controls, which has the crucial advantage that it does not require prior evaluation of the infimal cost function. We describe the method and illustrate its use in three examples. We also give two derivations of the principle, one in a special case under impractically strong conditions, and the other, at a heuristic level only, as an analogue of the method of Lagrange multipliers for constrained optimization.

We continue with the set-up of the preceding section but assume from now on that b, c and C are differentiable in t and x with continuous derivatives, and that the stopping set D is a hyperplane, thus $D = \{y\} + \Sigma$ for some $y \in \mathbb{R}^d$ and some vector subspace Σ of $\mathbb{R}^d$. Define for $\lambda \in \mathbb{R}^d$ the *Hamiltonian*

$$H(t, x, u, \lambda) = \lambda^T b(t, x, u) - c(t, x, u).$$

Pontryagin's maximum principle states that, if $(x_t, u_t)_{t \leqslant \tau}$ is optimal, then there exist *adjoint paths* $(\lambda_t)_{t \leqslant \tau}$ in $\mathbb{R}^d$ and $(\mu_t)_{t \leqslant \tau}$ in $\mathbb{R}$ with the following properties: for all $t \leqslant \tau$,

(i) $H(t, x_t, u, \lambda_t) + \mu_t$ has maximum value 0, achieved at $u = u_t$,

(ii) $\dot{\lambda}_t^T = -\lambda_t^T \nabla b(t, x_t, u_t) + \nabla c(t, x_t, u_t)$,

(iii) $\dot{\mu}_t = -\lambda_t^T \dot{b}(t, x_t, u_t) + \dot{c}(t, x_t, u_t)$,

(iv) $\dot{x}_t = b(t, x_t, u_t)$.

Moreover the following *transversality conditions* hold[28]:

(v) $(\lambda_\tau^T + \nabla C(\tau, x_\tau))\sigma = 0$ for all $\sigma \in \Sigma$,

and, in the time-unconstrained case,

(vi) $\mu_\tau + \dot{C}(\tau, x_\tau) = 0$.

Note that, in the time-unconstrained case, if b, c and C are time-independent, then $\mu_t = 0$ for all t.

The Hamiltonian serves as a way of remembering the first four statements, which could be expressed alternatively as

$$\text{(i) } 0 = \partial H / \partial u, \quad \text{(ii) } \dot{\lambda} = -\partial H / \partial x, \quad \text{(iii) } \dot{\mu} = -\partial H / \partial t, \quad \text{(iv) } \dot{x} = \partial H / \partial \lambda.$$

Beware that the reformulation of (i) is not always correct, for example in cases where the set of actions is an interval and where the maximum is achieved at an endpoint.

Example (Bringing a particle to rest in minimal time). Suppose we can apply a force to a particle, moving on a line, which imparts to it an acceleration a with $|a| \leqslant 1$ in the chosen units. For a given initial position q_0 and velocity p_0, how can we bring the particle to rest at the origin in the shortest time?

[28] Subject to the avoidance of certain pathological behaviour.

Take state $x = (q, p)$ and adjoint variable $\lambda = (\alpha, \beta)$. The problem is time-independent, so there is no need to consider μ. We have $\dot{q}_t = p_t$ and $\dot{p}_t = u_t$, with $|u_t| \leqslant 1$. We seek to minimize $\tau = \inf\{t \geqslant 0 : q_t = p_t = 0\} = \int_0^\tau 1dt$. So take $c = 1, C = 0$ and $D = \{(0, 0)\}$. The Hamiltonian is

$$H = \alpha p + \beta u - 1$$

so $u_t^* = \mathrm{sgn}(\beta_t)$ and $\alpha_t p_t + |\beta_t| = 1$. The adjoint equations are

$$\dot{\alpha}_t = -\partial H/\partial q = 0, \quad \dot{\beta}_t = -\partial H/\partial p = -\alpha_t.$$

So α is a constant and $\beta_t = \beta_\tau + \alpha s$, where $s = \tau - t$ is the time-to-go. Since $p_\tau = 0$, we must have $\beta_\tau = \pm 1$. There remains the problem of determining the values of α and β_τ as a function of (q_0, p_0). We do this backwards.

Suppose $\beta_\tau = 1$ and $\alpha \geqslant 0$, then $\beta_t \geqslant 0$ for all $t \leqslant \tau$, so $u_t = 1$, $p_t = -s$ and $q_t = s^2/2 = p_t^2/2$. On the other hand, if $\beta_\tau = 1$ and $\alpha < 0$, then the preceding calculation applies only for $s \leqslant s_0 = 1/|\alpha|$; once $s > s_0$, we have $\beta_t < 0$, so $u_t = -1$, and integrating the equations of motion back from s_0, we get $p_t = s - 2s_0$ and $q_t = 2s_0 s - s^2/2 - s_0^2$. Similar calculations apply for $\beta_\tau = -1$.

Thus we find there is a *switching locus* given by $q = -sgn(p)p^2/2$. Each initial state (q_0, p_0) above the locus lies on a unique parabola $q = -p^2/2 + c$, with $c > 0$. The optimal control is initially to take $a = -1$, thereby moving round the parabola to hit the switching locus. On hitting the locus, the acceleration acceleration changes sign, bringing the particle to rest at the origin by moving along the locus.

Example (Monopolist). Miss Prout holds the entire remaining stock of Cambridge elderberry wine for the vintage year 1959. If she releases it at rate u, then she realises a unit price $p(u) = 1 - u/2$ for $0 \leqslant u \leqslant 2$ and $p(u) = 0$ for $u \geqslant 2$. She holds amount x at time 0. What is her maximal total discounted return

$$\int_0^\infty e^{-\alpha t} u_t p(u_t) dt$$

and how should she achieve it?

The current stock evolves by $\dot{x}_t = -u_t$. Set $\tau = \inf\{t \geqslant 0 : x_t = 0\}$. Note that the rewards from any two controls which agree on $[0, n]$ can differ by at most $\int_n^\infty e^{-\alpha t} dt = e^{-\alpha n}/\alpha$ so it will suffice to find an optimal control among those for which $\tau < \infty$. So let us restrict now to such controls. We take $A = [0, \infty)$, $c = -e^{-\alpha t} up(u)$, $C = 0$ and $D = \{0\}$. The Hamiltonian is

$$H = -\lambda u + e^{-\alpha t} up(u),$$

which is maximized to a positive value at $u = 1 - \lambda e^{\alpha t}$, provided this is positive, and to 0 at 0 otherwise. The adjoint equation $\dot{\lambda}_t = -\partial H/\partial x = 0$ shows that λ is a constant, and the transversality condition $\mu_\tau = \dot{C} = 0$ shows that H is maximized to 0 at τ, so $u_\tau = 0$, and so $\lambda_\tau = e^{-\alpha\tau}$. Now

$$x = \int_0^\tau u_t dt = \int_0^\tau (1 - e^{-\alpha(\tau - t)}) dt = \tau - (1 - e^{-\alpha\tau})/\alpha.$$

This equation is satisfied by a unique $\tau \in (0, \infty)$, though we cannot solve it explicitly, and then the optimal control is $u_t = 1 - e^{-\alpha(\tau - t)}$. Finally, the maximal reward is

$$V(x) = \int_0^\infty e^{-\alpha t} u_t p(u_t) dt = \frac{(1 - e^{-\alpha\tau})^2}{2\alpha}.$$

Example (Insect optimization). A colony of insects consists of workers and queens, numbering w_t and q_t at time t. If a proportion u_t of the workers' effort at time t is devoted to producing more workers, then the numbers evolve according to the differential equations

$$\dot{w}_t = a u_t w_t - b w_t, \quad \dot{q}_t = (1 - u_t) w_t,$$

where a, b are positive constants, with $a > b$. How should the workers behave to maximize the number of queens produced by the end of the season?

Write T for the length of the season and take as state the number of workers. Then $c = -(1-u)w, C = 0$ and $D = \mathbb{R}$. The Hamiltonian is

$$H = \lambda(au - b)w + (1-u)w = \begin{cases} (1 - \lambda b)w, & \text{if } u = 0, \\ (\lambda a - \lambda b)w, & \text{if } u = 1. \end{cases}$$

So $u_t = 0$ if $\lambda_t a < 1$ and $u_t = 1$ if $\lambda_t a \geqslant 1$. The adjoint equations are

$$\dot{\lambda}_t = -\partial H / \partial w = \begin{cases} \lambda_t b - 1, & \text{if } u_t = 0, \\ -\lambda_t(a - b), & \text{if } u_t = 1. \end{cases}$$

Hence, for small time-to-go $s = T - t$, we have $u_t = 0$, so $\lambda_t = (1 - e^{-bs})/b$. We switch to $u_t = 1$ when $a(1 - e^{-bs})/b = 1$, that is, at

$$s_0 = \frac{1}{b} \log\left(\frac{a}{a-b}\right).$$

There is only one switch because $\dot{\lambda}_t$ is always negative. Hence, regardless of the length of the season, the workers should produce only more workers until there is s_0 time to go, when they should all switch to making queens.

A heuristic derivation of Pontryagin's maximum principle can be made by analogy with the method of Lagrange multipliers for constrained optimization problems. Recall that to maximize $f(x)$ subject to a d-dimensional constraint $g(x) = b$, one introduces the *Lagrangian*

$$L(x, \lambda) = f(x) - \lambda^T(g(x) - b),$$

where $\lambda \in \mathbb{R}^d$. For each λ, we seek $x(\lambda)$ to maximize $L(x, \lambda)$ and then seek λ so that $g(x(\lambda)) = b$. Then $x(\lambda)$ is the desired maximizer. Now, suppose we wish to maximize

$$-\int_0^T c(x_t, u_t) dt - C(x_T)$$

subject to $\dot{x}_t = b(x_t, u_t)$. We might try to maximize for each path $(\lambda_t)_{t\leqslant T}$

$$\begin{aligned} L(x,\lambda) &= \int_0^T \{-c(x_t,u_t) - \lambda_t^T(\dot{x}_t - b(x_t,u_t))\}dt - C(x_T) \\ &= -\lambda_T^T x_T + \lambda_0^T x_0 + \int_0^T \{\dot{\lambda}_t^T x_t + \lambda_t^T b(x_t,u_t) - c(x_t,u_t)\}dt - C(x_T). \end{aligned}$$

Then to maximize over x we might set

$$0 = \partial L/\partial x_t = \dot{\lambda}_t^T + \lambda_t^T \nabla b(x_t,u_t) - \nabla c(x_t,u_t),$$

which is the adjoint equation, and, in permitted directions,

$$0 = \partial L/\partial x_T = -\lambda_T^T - \nabla C(x_T),$$

which is the transversality condition.

The following result establishes the validity of Pontryagin's maximum principle, subject to the existence of a twice continuously differentiable solution to the Hamilton-Jacobi-Bellman equation, with well-behaved minimizing actions. These hypotheses are unnecessarily strong and are too strong for many applications. A proof of the principle under weaker hypotheses lies beyond the scope of this course. We assume that the action space A is an open subset in $\mathbb{R}^p$ and that b and the cost functions c and C are continuously differentiable.

Proposition 16.1. *Suppose that there exists a function $F : \tilde{S} \cup \tilde{D} \to \mathbb{R}$, twice differentiable with continuous derivatives, and a function $u : \tilde{S} \to A$ such that*

$$c(t,x,a) + \dot{F}(t,x) + \nabla F(t,x) b(t,x,a) \geqslant 0$$

for all $a \in A$, with equality when $a = u(t,x)$, for all $(t,x) \in \tilde{S}$. Suppose also that $F = C$ on $\tilde{D}$. Fix a starting point $(0,x)$ and assume that u defines a continuous feasible control and controlled path $(u_t,x_t)_{t\leqslant\tau}$ starting from $(0,x)$. Set $\mu_t = -\dot{F}(t,x_t)$ and $\lambda_t^T = -\nabla F(t,x_t)$, then

$$\begin{aligned} \dot{\lambda}_t^T &= -\lambda_t^T \nabla b(t,x_t,u_t) + \nabla c(t,x_t,u_t), \\ \dot{\mu}_t &= -\lambda_t^T \dot{b}(t,x_t,u_t) + \dot{c}(t,x_t,u_t), \end{aligned}$$

and, for any $\sigma \in \Sigma$, we have

$$(\lambda_\tau^T + \nabla C)(\tau,x_\tau))\sigma = 0,$$

and, in the time-unconstrained case,

$$\mu_\tau + \dot{C}(\tau,x_\tau) = 0.$$

Proof. Define, for $(t,x) \in \tilde{S}$ and $a \in A$,

$$J(t,x,a) = c(t,x,a) + \dot{F}(t,x) + \nabla F(t,x) b(t,x,a).$$

Then $J(t,x,a) \geqslant 0$ and $J(t,x,u(t,x)) = 0$ so, since A is open, we have

$$(\partial J/\partial a)(t,x,u(t,x)) = 0,$$

and hence

$$0 = (\partial/\partial x)J(t,x,u(t,x)) = \nabla J(t,x,u(t,x)), \quad 0 = (\partial/\partial t)J(t,x,u(t,x)) = \dot{J}(t,x,u(t,x)).$$

Write $a = u(t,x)$, then

$$0 = \nabla J(t,x,a) = \nabla c(t,x,a) + \nabla F(t,x)\nabla b(t,x,a) + \{\nabla \dot{F}(t,x) + \nabla^2 F(t,x) b(t,x,a)\}$$

and

$$0 = \dot{J}(t,x,a) = \dot{c}(t,x,a) + \nabla F(t,x)\dot{b}(t,x,a) + \{\nabla \dot{F}(t,x) b(t,x,a) + \ddot{F}(t,x)\}.$$

Hence

$$\begin{aligned}\dot{\lambda}_t^T &= -\nabla \dot{F}(t,x_t) - \nabla^2 F(t,x_t) b(t,x_t,u_t) \\ &= \nabla c(t,x_t,u_t) + \nabla F(t,x_t)\nabla b(t,x_t,u_t) = \nabla c(t,x_t,u_t) - \lambda_t^T \nabla b(t,x_t,u_t)\end{aligned}$$

and

$$\begin{aligned}\dot{\mu}_t &= -\ddot{F}(t,x_t) - \nabla \dot{F}(t,x_t) b(t,x_t,u_t) \\ &= \dot{c}(t,x,u_t) + \nabla F(t,x_t)\dot{b}(t,x_t,u_t) = \dot{c}(t,x,u_t) - \lambda_t^T \dot{b}(t,x_t,u_t).\end{aligned}$$

On differentiating the equality $F = C$ at (τ, x_τ) in the direction σ , we obtain

$$(\lambda_\tau^T + \nabla C)(\tau, x_\tau))\sigma = 0,$$

and, in the time-unconstrained case, we can differentiate at (τ, x_τ) in t to obtain

$$\mu_\tau + \dot{C}(\tau, x_\tau) = 0.$$

$\square$