Cold war era: Pontryagin vs. Bellman.

5.1 Dynamic programming and the HJB equation

"In place of determining the optimal sequence of decisions from the *fixed* state of the system, we wish to determine the optimal decision to be made at *any* state of the system. Only if we know the latter, do we understand the intrinsic structure of the solution." —Richard Bellman, "Dynamic Programming", 1957. [Vinter, p. 435]

Reading assignment: read and recall the derivation of HJB equation done in ECE 515 notes.

Other references: Yong-Zhou, A-F, Vinter, Bressan (slides and tutorial paper), Sontag.

5.1.1 Motivation: the discrete problem

Bellman's *dynamic programming* approach is quite general, but it's probably the easiest to understand in the case of purely discrete system (see, e.g., [Sontag, Sect. 8.1]).

$$x_{k+1} = f(x_k, u_k), \qquad k = 0, 1, \ldots, T-1$$

$x_k \in \mathcal{X}$—a finite set of cardinality N, $u_k \in \mathcal{U}$—a finite set of cardinality M (T, N, M are positive integers).

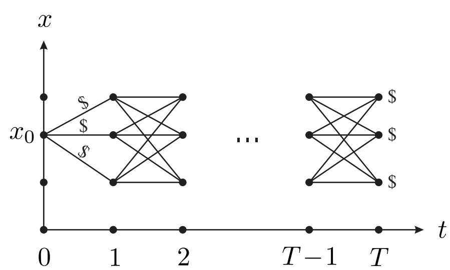

Figure 5.1: Discrete time: going forward

Each transition has a cost, plus possibly a terminal cost. *Goal:* minimize total cost.

Most naive approach: starting at x_0, enumerate all possible trajectories, calculate the cost for each, compare and select the optimal one.

Computational effort – ? (done offline)

M^T possible sequences, need to add T terms to find cost for each $\Rightarrow$ roughly $M^T T$ operations (or $O(M^T T)$).

Alternative approach: let's go backwards!

At $k = T$, terminal costs are known for each x.

At $k = T-1$: for each x, find to which x we should jump at $k = T$ so as to have the smallest "cost-to-go" (1-step running cost plus terminal cost). Write this cost next to x, and mark the selected path (in the picture).

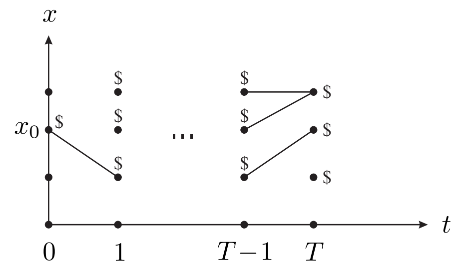

Figure 5.2: Discrete time: going backward

Repeat for $k = T - 2, \ldots, 0$.

Claim: when we're done, we have an optimal path from each x_0 to some x_T (this path is unique unless there exist more than one path giving the same cost—then can select a unique one by coin toss).

Is this obvious?

The above result relies on (already saw in the Maximum Principle)

Principle of optimality: a final portion of an optimal trajectory is itself optimal (with respect to its starting point as initial condition). The reason for this is obvious: if there is another choice of the final portion with lower cost, we could choose it and get lower total cost, which contradicts optimality.

What principle of optimality does for us here is guarantee that the steps we discard (going backwards) cannot be parts of optimal trajectories. (But: cannot discard anything going forward!)

Computational effort – ? (again, offline)

At each time, for each x, need to compare all M controls and add up cost-to-go for each $\Rightarrow$ roughly, $O(T \cdot N \cdot M)$ operations. Recall: in forward approach we had $O(TM^T)$.

Advantages of the backward approach ("dynamic programming" approach):

- Computational effort is clearly smaller if N and M are fixed and T is large (a better organized calculation). But still large if N and M are large ("curse of dimensionality");

  $\longrightarrow$ Note that the comparision is not really fair because the backward approach gives more: optimal policy for any initial condition (and $\forall\, k$). For forward approach, to handle all initial conditions would require $O(TNM^T)$ operations—and still wouldn't cover some states for $k > 0$.

- Even more importantly, the backward approach gives the optimal control policy in the form of *feedback:* given x, we know what to do (for any k).

  $\longrightarrow$ In forward approach, optimal policy depends not on x but on the result of the entire forward pass computation.

5.1.2 The HJB equation and sufficient conditions for optimality

Back to continuous case [ECE 515, Yong-Zhou]

$$\dot{x} = f(t, x, u), \qquad x(t_0) = x_0$$

$$J = \int_{t_0}^{t_1} L(t, x, u)dt + K(x(t_1))$$

t_1 – fixed, $x(t_1)$ – free (other variants also possible—target sets [A-F]).

The above cost can be written as $J(t_0, x_0, u(\cdot))$.

"Dynamic Programming:" the idea is to consider, instead of the problem of minimizing $J(t_0, x_0, u(\cdot))$ for given (t_0, x_0), the *family* of problems

$$\min_u J(t, x, u)$$

where t ranges over $[t_0, t_1)$ and x ranges over $\mathbb{R}^n$.

Goal: derive *dynamic relationships* among these problems, and ultimately solve *all of them* (see Bellman's quote at the beginning).

So, we consider the cost functionals

$$J(t, x, u) = \int_t^{t_1} L(\tau, x(\tau), u(\tau))d\tau + K(x(t_1))$$

where $x(\cdot)$ has initial condition $x(t) = x$. (Yong-Zhou use (s, y) instead of (t, x_0).)

Value function (optimal cost-to-go):

$$V(t, x) := \inf_{u_{[t,t_1]}} J(t, x, u(\cdot))$$

By definition of the control problem, we have $V(t_1, x) = K(x)$.

Note: this is the feature of our specific problem. If there is no terminal cost ($K = 0$), then $V(t_1, x) = 0$. More generally, we could have a target set $S \subset \mathbb{R} \times \mathbb{R}^n$, and then we'd have $V(t, x) = 0$ when $(t, x) \in S$ [A-F].

$\longrightarrow$ By defining V via inf (rather than min), we're not even assuming that an optimal control exists (inf doesn't need to be achieved).

$\longrightarrow$ Also, we can have $V = -\infty$ for some (t, x).

Principle of Optimality (or *Principle of Dynamic Programming*):

For every (t, x) and every $\Delta t > 0$, we have

$$V(t, x) = \inf_{u_{[t,t+\Delta t]}} \left\{ \int_t^{t+\Delta t} L(\tau, x, u)d\tau + V(t + \Delta t, x(t + \Delta t, u_{[t,t+\Delta t]})) \right\}$$

($x(t + \Delta t, u_{[t,t+\Delta t]})$ is the state at $t + \Delta t$ corresponding to the control $u_{[t,t+\Delta t]}$).

This is familiar:

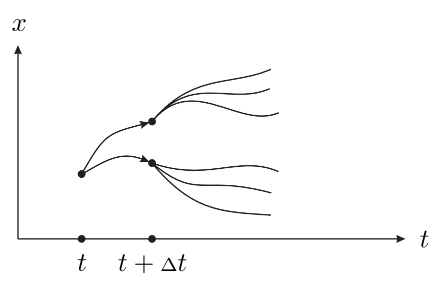

Figure 5.3: Continuous time: principle of optimality

(this picture is in [Bressan]). Continuous version of what we had before.

The statement of principle of optimality seems obvious: to search for an optimal control, we can search for a control over a small time interval that minimizes the cost over this interval plus the cost-to-go from there. However, it's useful to prove it formally—especially since we're using inf and *not assuming existence of optimal control.*

Exercise 17 (due Apr 11) Provide a formal proof of the principle of optimality (using definition of inf).

The principle of optimality is that dynamic relation among family of problems we talked about earlier. But it's difficult to use, so we want to pass to its infinitesimal version (*will get a PDE*). Namely, let's do something similar to the proof of the Maximum Principle : take Δt to be small, write $x(t+\Delta t) = x(t) + \Delta x$ where Δx is then also small, and do a first-order Taylor expansion:

$$V(t+\Delta t, x+\Delta x) \approx V(t,x) + V_t(t,x)\Delta t + V_x(t,x)\cdot \Delta x$$

(V_x is the gradient, can use $\langle \cdot, \cdot \rangle$ notation). Also,

$$\int_t^{t+\Delta t} L(\tau, x, u)d\tau \approx L(t,x,u(t))\cdot \Delta t$$

($u(t)$ is the value at the left endpoint). Plug this into the principle of optimality: $V(t,x)$ cancels out and we get

$$0 \approx \inf_{u_{[t,t+\Delta t]}} \{L(t,x,u(t))\Delta t + V_t(t,x)\Delta t + V_x(t,x)\Delta x\}$$

Divide by Δt:

$$0 \approx \inf_{u_{[t,t+\Delta t]}} \left\{L(t,x,u(t)) + V_t(t,x) + V_x(t,x)\cdot \frac{\Delta x}{\Delta t}\right\}$$

Finally, let $\Delta t \to 0$. Then $\frac{\Delta x}{\Delta t} \to f(t,x,u(t))$, and h.o.t. disappear. Then, inf is over the instantaneous values of u at time t. Also, V_t doesn't depend on u so we can take it out:

$$-V_t(t,x) = \inf_u \{L(t,x,u(t)) + \langle V_x(t,x), f(t,x,u)\rangle\} \tag{5.1}$$

This PDE is called the *HJB equation.*

Recall the boundary condition (where did K go?)

$$V(t_1, x) = K(x)$$

For different problem formulations, with target sets, this boundary condition will change but the HJB equation is the same.

Note that the expression inside the inf is minus the Hamiltonian, so we have

$$\boxed{V_t = \sup_u H(t,x,-V_x,u)}$$

where V_x plays the role of $-p$.

If u^* is the optimal control, then can repeat everything (starting with principle of optimality) with u^* plugged in. Then sup becomes max, i.e., it's achieved for some control u^* along its corresponding trajectory x^*:

$$u^*(t) = \arg\max_u H(t, x^*(t), -V_x(t,x^*(t)), u), \qquad t\in[t_0,t_1] \tag{5.2}$$

(this is state feedback, more on this later). Plugging this into HJB, we get

$$V_t(t, x^*(t)) = H(t, x^*(t), -V_x(t, x^*(t)), u^*(t))$$

The above derivation only shows the *necessity* of the conditions—the HJB and the H-maximization condition on u^*. (If $u^* \neq \arg\max$ then it can't be optimal, but sufficiency is not proved yet.) The latter condition is very similar to the Maximum Principle , and the former is a new condition complementing it.

However, it turns out that these conditions are also *sufficient* for optimality.

It should be clear that we didn't prove sufficiency yet: $V =$ optimal cost, $u^* =$ optimal control $\Rightarrow$ HJB and (5.2), that's what we did so far.

Theorem 10 (sufficient conditions for optimality) *Suppose that there exists a function V that satisfies the HJB equation on $\mathbb{R} \times \mathbb{R}^n$ (or $[t_0, t_1] \times \mathbb{R}^n$) and the boundary condition. (Assuming the solution exists and is $\mathcal{C}^1$.)*

Suppose that a given control u^* and the corresponding trajectory x^* satisfy the Hamiltonian maximization condition (5.2). *(Assuming* min *is achieved.)*

Then u^* is an optimal control, and $V(t_0, x_0)$ is the optimal cost. (Uniqueness of u^* not claimed, can have multiple controls giving the same cost.)

Proof. Let's use (5.1). First, plug in x^* and u^*, which achieve the min:

$$-V_t(t, x^*) = L(t, x^*, u^*) + V_x|_* \cdot f(t, x^*, u^*)$$

Can rewrite as

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

(the total derivative of V along x^*). Integrate from t_0 to t_1 ($x(t_0)$ is given and fixed):

$$0 = \int_{t_0}^{t_1} L(t, x^*, u^*)dt + \underbrace{V(t_1, x^*(t_1))}_{K(x^*(t_1))} - V(t_0, \underbrace{x^*(t_0)}_{x_0})$$

so we have

$$V(t_0, x_0) = \int_{t_0}^{t_1} L(t, x^*, u^*)dt + K(x^*(t_1)) = J(t_0, x_0, u^*)$$

Now let's consider another arbitrary control $\bar{u}$ and the corresponding trajectory $\bar{x}$, again with initial condition (t_0, x_0). Because of the minimization condition in HJB,

$$-V_t(t, \bar{x}) \leq L(t, \bar{x}, \bar{u}) + V_x(t, \bar{x}) \cdot f(t, \bar{x}, \bar{u})$$

or, as before,

$$0 \leq L(t, \bar{x}, \bar{u}) + \left.\frac{dV}{dt}\right|_-$$

(total derivative along $\bar{x}$). Integrating over $[t_0, t_1]$:

$$0 \leq \int_{t_0}^{t_1} L(t, \bar{x}, \bar{u})dt + \underbrace{V(t_1, \bar{x}(t_1))}_{K(\bar{x}(t_1))} - V(t_0, x_0)$$

hence

$$V(t_0, x_0) \leq \int_{t_0}^{t_1} L(t, \bar{x}, \bar{u})dt + K(\bar{x}(t_1)) = J(t_0, x_0, \bar{u})$$

We conclude that $V(t_0, x_0)$ is the cost for u^* and the cost for an arbitrary other control $\bar{u}$ cannot be smaller $\Rightarrow V(t_0, x_0)$ is the optimal cost and u^* is an optimal control. □

Example 13 Consider the parking problem from before: $\ddot{x} = u$, $u \in [-1, 1]$, goal: transfer from given x_0 to rest at the origin ($x = \dot{x} = 0$) in minimum time. The HJB equation is

$$-V_t = \inf_u \{1 + V_{x_1} x_2 + V_{x_2} u\}$$

with boundary condition $V(t, 0) = 0$. The control law that achieves the inf (so it's a min) is

$$u^* = -\text{sgn}(V_{x_2})$$

(as long as $V_{x_2} \neq 0$). Plugging this into the HJB, we have

$$-V_t = 1 + V_{x_1} x_2 - |V_{x_2}|$$

Can you solve the HJB equation? Can you use it to derive the bang-bang principle? Can you verify that the optimal control law we derived using the Maximum Principle is indeed optimal?

(This is a free-time, fixed-endpoint problem. But the difference between different target sets is only in boundary conditions.) □

Saw other examples in ECE 515. Know that HJB is typically very hard to solve.

Important special case for which the HJB simplifies:

System is time-invariant: $\dot{x} = f(x, u)$; infinite horizon: $t_1 = \infty$ (and no terminal cost); Lagrangian is time-independent: $L = L(x, u)$.

(Another option is to work with free terminal time and target set for the terminal state, e.g., time-optimal problems.)

Then it is clear that the cost doesn't depend on the initial time, just on the initial state ⇒ value function depends on x only: $V(t, x) \mapsto V(x)$. Hence $V_t = 0$ and HJB simplifies to

$$0 = \inf_u \{L(x, u) + V_x(x) \cdot f(x, u)\} \tag{5.3}$$

$\longrightarrow$ This is of course still a PDE unless $\dim x = 1$. (But for scalar systems becomes an ODE and can be solved .)

This case has a simpler HJB equation, but there is a complication compared to the finite-horizon case: the infinite integral $\int_{t_0}^{\infty} L(x, u)dt$ may be infinite ($+\infty$). So, we need to restrict attention to trajectories for which this integral is finite.

Example 14 LQR problem:

$$\dot{x} = Ax + Bu$$

$$J = \int_{t_0}^{t_1} (x^T Q x + u^T R u) dt \qquad (t_1 \leq \infty)$$

We'll study this case in detail later and find that optimal controls take the form of *linear* state feedbacks. When $x(t)$ is given by exponentials, it's clear that for $t_1 = \infty$ the integral is finite *if and only if* these exponentials are decaying. So, the infinite-horizon case is closely related to the issue of *closed-loop asymptotic stability*.

□

We will investigate this connection for LQR in detail later.

Back in the nonlinear case, let's for now assume the following:

- $\int_{t_0}^{\infty} L(x,u)dt < \infty$ is possible only when $x(t) \to 0$ as $t \to \infty$.
- $V(0) = 0$ (true, e.g., if $f(0,0) = 0$, $L(0,0) = 0$, and $L(x,u) > 0\ \forall\,(x,u) \neq (0,0)$. Then the best thing to do is not to move—gives zero cost.)

Then we have the following

Sufficient conditions for optimality:

(Here we are assuming that $V(x) < \infty\ \forall\, x \in \mathbb{R}^n$.) Suppose that $V(x)$ satisfies the infinite-horizon HJB (5.3), with boundary condition $V(0) = 0$. Suppose that a given control u^* and the corresponding trajectory x^* satisfy the H-maximization condition (5.2) and $x^*(t) \to 0$ as $t \to \infty$. Then $V(x_0)$ is the optimal cost and u^* is an optimal control.

Exercise 18 (due Apr 18) Prove this.

5.1.3 Some historical remarks

Reference: [Y-Z].

The PDE was derived by Hamilton in 1830s in the context of calculus of variations . Later studied by Jacobi who improved Hamilton's results.

$\longrightarrow$ Read [G-F, Sect. 23] where it's derived using variation of a functional (as a *necessary* condition).

Using it as a *sufficient* condition for optimality was proposed (again in calculus of variations) by Carathéodory in 1920s ("royal road"; he worked with V locally near test trajectory).

Principle of optimality seems almost a trivial observation, was made by Isaacs in early 1950s (slightly before Bellman).

Bellman's contribution was to recognize the power of the approach to study value functions globally. He coined the term "dynamic programming". Applied to a wide variety of problems, wrote a book (all in 1950s).

[A-F, Y-Z] In 1960s, Kalman made explicit connection between Bellman's work and Hamilton-Jacobi equation (not clear if Bellman knew this). He coined the term "HJB equation" and used it for control problems.

A most remarkable fact is that the Maximum Principle was developed in Soviet Union independently around the same time (1950s-1960s)! So it's natural to discuss the connection.

5.2 Relationships between the Maximum Principle and the HJB equation

- There is a notable difference between the form of the optimal control law provided by the Maximum Principle and HJB equation. the Maximum Principle says:

$$\dot{x}^* = H_p|_*, \qquad \dot{p}^* = -H_x|_*$$

u^* pointwise maximizes $H(x^*(t), u, p^*(t))$. This is an *open-loop* specification. HJB says:

$$u^*(t) = \arg\max_u H(t, x^*(t), -V_x|_*, u)$$

This is a *closed-loop (feedback)* specification: assuming we know V everywhere, $u^*(t)$ is determined by current state $x^*(t)$. We discussed this feature at the beginning of this chapter, when we first described Bellman's approach—it determines optimal strategy for *all* states, so we automatically get a feedback.

But to know $V(t, x)$ everywhere, we need to solve HJB which is hard—so there's no free lunch.

- HJB:

$$V_t = \max_u H(t, x, -V_x, u)|_*$$

or sup if max is not achieved—but assume it is and we have u^*. So HJB, like the Maximum Principle , involves a Hamiltonian maximization condition. In fact, let's see if we can prove the Maximum Principle starting from HJB equation. the Maximum Principle doesn't involve any PDE, but instead involves a system of ODEs (canonical equations). We can convert HJB equation into an ODE (having $\frac{\partial}{\partial t}$ only) if we introduce a new vector variable:

$$p^*(t) := -V_x(t, x^*(t))$$

The boundary condition for p^* is

$$p^*(t_1) = -V_x(t_1, x^*(t_1))$$

But recall the boundary condition for HJB: $V(t_1, x) = K(x)$, so we get $p^*(t_1) = -K_x(x^*(t_1))$ and this matches the boundary condition we had for p^* in the Maximum Principle for the case of terminal cost.

Seems OK (we didn't check this carefully for target sets and transversality, though), but let's forget the boundary condition and derive a differential equation for p^*.

$$\begin{aligned}
\dot{p}^* &= -\frac{d}{dt}(V_x|_*) = -V_{tx}|_* - \underbrace{V_{xx}|_*}_{\text{matrix}} \cdot \dot{x}^* \\
&= \text{(swapping partials)} \;\; -V_{xt}|_* - V_{xx}|_* \cdot f|_* \\
&= -\frac{\partial}{\partial x}\left(\underbrace{V_t + \langle V_x, f \rangle}_{=-L \text{ by HJB}} \right)\Bigg|_* + (f_x)^T \underbrace{V_x|_*}_{=-p^*} \\
&= L_x|_* - (f_x)^T\Big|_* p^*
\end{aligned}$$

and this is the correct adjoint equation.

$\longrightarrow$ We saw in the Maximum Principle that p^* has interpretations as the normal to supporting hyperplanes, also as Lagrange multiplier. Now we see another interpretation: it's the gradient of the value function. Can view it as *sensitivity* of the optimal cost with respect to x. Economic meaning: "marginal value", or "shadow price"—it tells us by how much we can increase the benefits by increasing resources/spending, or how much we'd be willing to pay someone else for this resource to still make profit. See [Y-Z, p. 231], also [Luenberger (blue), pp. 95–96].

The above calculation is easy enough. Why didn't we derive HJB and then the Maximum Principle from it, why did we have to spend so much time on the proof of the Maximum Principle ?

Answer: the above assumes $V \in \mathcal{C}^1$, what if this is not the case?!

In fact, all our previous developments (e.g., sufficient conditions) rely on HJB equation having a $\mathcal{C}^1$ solution V. Let's see if this is a reasonable expectation.

Example 15 (Y-Z, p. 163) , [Vinter, p. 36]

$$\dot{x} = xu, \qquad x \in \mathbb{R},\ u \in [-1, 1],\ x(t_0) = x_0$$

$J = x(t_1) \to \min$, t_1 fixed. Optimal solution can be found by inspection: $x_0 > 0 \Rightarrow u = -1 \Rightarrow \dot{x} = -x \Rightarrow x(t_1) = e^{-t_1+t_0}x_0$; $x_0 < 0 \Rightarrow u = 1 \Rightarrow \dot{x} = x \Rightarrow x(t_1) = e^{t_1-t_0}x_0$; $x_0 = 0 \Rightarrow x(t) \equiv 0\ \forall u$. From this we get the value function

$$V(t_0, x_0) = \begin{cases} e^{-t_1+t_0}x_0 & \text{if } x_0 > 0 \\ e^{t_1-t_0}x_0 & \text{if } x_0 < 0 \\ 0 & \text{if } x_0 = 0 \end{cases}$$

Let's fix t_0 and plot this as a function of x_0. In fact, for simplicity let $t_0 = 0$, $t_1 = 1$.

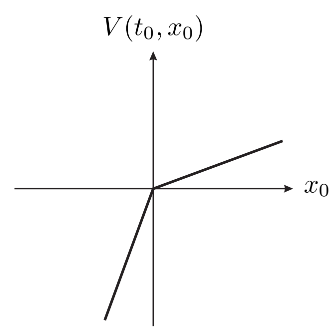

Figure 5.4: $V(t_0, x_0)$ as a function of x_0

So V is not $\mathcal{C}^1$ when $x_0 = 0$, it's only Lipschitz continuous. □

The HJB equation in the above example is

$$-V_t = \inf_u \{V_x \cdot xu\} = -|V_x \cdot x|$$

(we're working with scalars here). Looks pretty simple. So, maybe it has another solution which is $\mathcal{C}^1$, and the value function is just not the right solution? It turns out that:

- The above HJB equation does not admit any $\mathcal{C}^1$ solution [Y-Z];
- The above simple example is not an exception—this situation is quite typical for problems involving control bounds and terminal cost [Vinter];
- The value function can be shown to be Lipschitz for a reasonable class of control problems [Bressan, Lemma 4].

This has implications not just for connecting HJB and the Maximum Principle but for the HJB theory itself: we need to reconsider the assumption that $V \in \mathcal{C}^1$ and work with some more general solution concept that allows V to be nondifferentiable (just Lipschitz). (Cf. the situation with ODEs: $x \in \mathcal{C}^1 \to x \in \mathcal{C}^1$ a.e. (absolutely continuous) $\to$ Filippov's solutions, etc.)

The above difficulty has been causing problems for a long time. The theory of dynamic programming remained non-rigorous until 1980s when, after a series of related developments, the concept of *viscosity solutions* was introduced by Crandall and Lions (see [Y-Z, p. 213]). (This completes the historical timeline given earlier.)

$\longrightarrow$ This is in contrast with the Maximum Principle theory, which was pretty solid already in 1960s.

5.3 Viscosity solutions of the HJB equation

References: [Bressan's tutorial], also [Y-Z], [Vinter].

The concept of viscosity solution relies on some notions from *nonsmooth analysis*, which we now need to review.

5.3.1 One-sided differentials

Reference: [Bressan]

Consider a *continuous* (nothing else) function $v : \mathbb{R}^n \to \mathbb{R}$. A vector p is called a *super-differential* of v at a given point x if

$$v(y) \leq v(x) + \langle p, y - x \rangle + o(|y - x|)$$

along any path $y \to x$.

$\longrightarrow$ See [Bressan] for a different, somewhat more precise definition.

Geometrically:

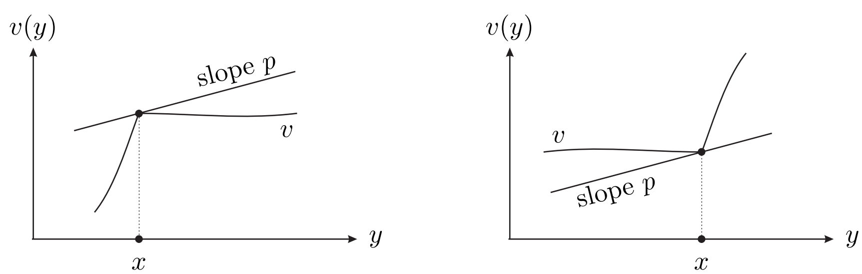

Figure 5.5: (a) Super-differential, (b) Sub-differential

In higher dimensions, this says that the plane $y \mapsto v(x) + \langle p, y - x \rangle$ is tangent *from above* to the graph of v at x. p is the gradient of this linear function. $o(|y - x|)$ means that this is an *infinitesimal*, or local, requirement.

$\longrightarrow$ Such a p in general is not unique, so we have a *set* of super-differentials of v at x, denoted by $D^+v(x)$.

Similarly, we say that $p \in \mathbb{R}^n$ is a *sub-differential* for v at x if

$$v(y) \geq v(x) + \langle p, y - x \rangle - o(|y - x|)$$

as $y \to x$ along any path. Geometrically, everything is reversed: line with slope p (for $n = 1$) or, in general, linear function with gradient p is tangent *from below*, locally up to higher-order terms:

The set of subdifferentials of v at x is denoted by $D^-v(x)$.

Example 16 [Bressan]

$$v(x) = \begin{cases} 0 & \text{if } x < 0 \\ \sqrt{x} & \text{if } 0 \leq x \leq 1 \\ 1 & \text{if } x > 1 \end{cases}$$

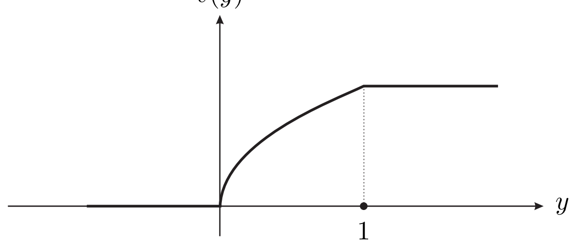

Figure 5.6: The function in Example 16

$D^+v(0) = \emptyset$, $D^-v(0) = [0, \infty)$, $D^+v(1) = [0, 1/2]$, $D^-v(1) = \emptyset$. These are the only interesting points, because of the following result. □

Lemma 11 (some useful properties of $D^{\pm}v$)

- *$p \in D^{\pm}v(x) \Leftrightarrow \exists$ a $\mathcal{C}^1$ function φ such that $\nabla\varphi(x) = p$, $\varphi(x) = v(x)$, and, for all y near x, $\varphi(y) \geq v(y)$ in case of D^+, $\leq$ in case of D^-. (Or can have strict inequality—no loss of generality.) I.e., $\varphi - v$ has a local minimum at x in case of D^+, and local maximum in case of D^-.*

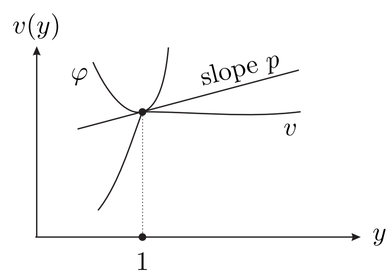

Figure 5.7: Interpretation of super-differential via test function

Will not prove this (see [Bressan]), but will need this for all other statements.

$\longrightarrow$ φ is sometimes called a test function [Y-Z].

- *If v is differentiable at x, then $D^+v(x) = D^-v(x) = \{\nabla v(x)\}$—gradient of v at x. I.e., super/sub-differentials are extensions of the usual differentials. (In contrast, Clarke's generalized gradient extends the idea of continuous differential [Y-Z, bookmark].)*

  PROOF. *Use the first property. $\nabla v(x) \in D^{\pm}v(x)$—clear[1]. If $\varphi \in \mathcal{C}^1$ and $\varphi \geq v$ with equality at x, then $\varphi - v$ has a local minimum at $x \Rightarrow \nabla(\varphi - v) = 0 \Rightarrow \nabla\varphi(x) = \nabla v(x)$ so there are no other p in $D^+v(x)$. Same for $D^-v(x)$.*

[1] The gradient provides, up to higher-order terms, both an over- and under-approximation for a differentiable function.

- *If $D^+v(x)$ and $D^-v(x)$ are both non-empty, then v is differentiable at x and the previous property holds.*

  PROOF. *There exist $\varphi_1, \varphi_2 \in \mathcal{C}^1$ such that $\varphi_1(x) = v(x) = \varphi_2(x)$, $\varphi_1(y) \le v(y) \le \varphi_2(y)$ $\forall\, y$ near x. Thus $\varphi_1 - \varphi_2$ has a local maximum at $x \Rightarrow \nabla(\varphi_1 - \varphi_2)(x) = 0 \Rightarrow \nabla\varphi_1(x) = \nabla\varphi_2(x)$. Now write (assuming $y > x$, otherwise the inequalities are flipped but the conclusion is the same)*

  $$\frac{\varphi_1(y) - \varphi_1(x)}{y - x} \le \frac{v(y) - v(x)}{y - x} \le \frac{\varphi_2(y) - \varphi_2(x)}{y - x}$$

  As $y \to x$, the first fraction approaches $\nabla\varphi_1(x)$, the last one approaches $\nabla\varphi_2(x)$. The two gradients are equal. Hence, by the "Theorem about Two Cops" (or the "Sandwich Theorem"), the limit of the middle fraction exists and equals the other two. This limit must be $\nabla v(x)$, and everything is proved.

- *The set $\{x : D^+v(x) \ne \emptyset\}$ (respectively, $\{x : D^-v(x) \ne \emptyset\}$) is non-empty, and actually dense in the domain of v.*

  IDEA OF PROOF.

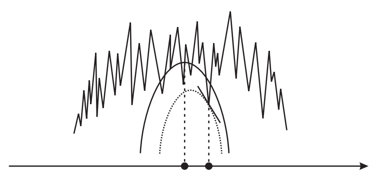

Figure 5.8: Idea behind the denseness proof

Take an arbitrary x_0. For φ steep enough (but $\mathcal{C}^1$), $\varphi - v$ will have a local maximum at a nearby x, as close to x_0 as we want. Just move φ to cross that point $\Rightarrow \nabla\varphi(x) \in D^-v(x)$. (Same for D^+.)

5.3.2 Viscosity solutions of PDEs

$$F(x, v, v_x) = 0 \tag{5.4}$$

(t can be part of x, as before).

Definition 3 A continuous function $v(x)$ is a *viscosity subsolution* of the PDE (5.4) if

$$F(x, v, p) \le 0 \qquad \forall\, p \in D^+v(x),\ \forall\, x$$

Equivalently, $F(x, \varphi, \varphi_x) \le 0$ for every $\mathcal{C}^1$ function φ such that $\varphi - v$ has a local minimum at x, and this is true for every x. Similarly, $v(x)$ is a *viscosity supersolution* of the PDE (5.4) if

$$F(x, v, p) \ge 0 \qquad \forall\, p \in D^-v(x),\ \forall\, x$$

or, equivalently, $F(x, \varphi, \varphi_x) \ge 0$ for every $\mathcal{C}^1$ function φ such that $\varphi - v$ has a local maximum at x, and this is true for every x. v is a *viscosity solution* if it is both viscosity super- and sub-solution.

⟶ Note that the above definitions impose conditions on v only at points where D^+v and D^-v are non-empty. We know that the set of these points is dense.

⟶ At all points where v is differentiable, the PDE must hold in the usual sense. (If v is Lipschitz then this is a.e. by Rademacher's Theorem.)

Example 17 $F(x, v, v_x) = 1 - |v_x|$, so the PDE is $1 - |v_x| = 0$.

(Is this the HJB for some control problem?)

$v = \pm x$ (plus constant) are classical solutions.

Claim: $v(x) = |x|$ is a viscosity solution.

At points of differentiability—OK. At $x = 0$, we have:

$D^+v(0) = \emptyset \Rightarrow$ nothing to check. $D^-v(0) = [-1, 1]$, $1 - |p| \geq 0 \; \forall \, p \in [-1, 1]$—OK. □

⟶ Note the lack of symmetry: if we rewrite the PDE as $|v_x| - 1 = 0$ then $|x|$ is *not* a viscosity solution! So we need to be careful with the sign.

The term "viscosity solution" comes from the fact that v can be obtained from smooth solutions to the PDE

$$F(x, v_\varepsilon, \nabla v_\varepsilon) = \varepsilon \Delta v_\varepsilon$$

in the limit as $\varepsilon \to 0$, and the Laplacian on the right-hand side is a term used to model viscosity of a fluid (cf. [Feynman, Eq. (41.17), vol. II]). See [Bressan] for a proof of this convergence result.

Begin optional:

The basic idea behind that proof is to consider a test function $\varphi \in \mathcal{C}^2$ (or approximate it by some $\psi \in \mathcal{C}^2$). Let's say we're in supersolution case. Then for each ε, $\varphi - v_\varepsilon$ has a local maximum at some x_ε near a given x, hence $\nabla \varphi(x_\varepsilon) = \nabla v_\varepsilon(x_\varepsilon)$ and $\Delta \varphi(x_\varepsilon) \leq \Delta v_\varepsilon(x_\varepsilon)$. This gives $F(x, v_\varepsilon, \nabla \varphi_\varepsilon) \geq \varepsilon \Delta \varphi$. Taking the limit as $\varepsilon \to 0$, we have $F(x, v, \nabla \varphi) \geq 0$ as needed.

End optional

5.3.3 The HJB partial differential equation and the value function

Now go back to the optimal control problem and consider the HJB equation

$$-v_t - \min_u \{L(x, u) + v_x \cdot f(x, u)\} = 0$$

(or inf), with boundary condition $v(t_1, x) = K(x)$. Now we have (t, x) in place of what was just x earlier. Note that u is minimized over, so it doesn't explicitly enter the PDE.

Theorem 12 *Under suitable technical assumptions (rather strong: f, L, K are bounded uniformly over x and globally Lipschitz as functions of x; see [Bressan] for details), the value function $V(t, x)$ is a unique viscosity solution of the HJB equation. It is also Lipschitz (but not necessarily $\mathcal{C}^1$).*

We won't prove everything, but let's prove one claim: that V is a viscosity subsolution. Need to show: $\forall \, \varphi(t, x) \in \mathcal{C}^1$ such that $\varphi - V$ attains a local minimum at (t_0, x_0)—arbitrary time/space—we have

$$\varphi_t(t_0, x_0) + \inf_u \{L(x_0, u) + \varphi_x(t_0, x_0) \cdot f(x_0, u)\} \geq 0$$

Assume the contrary: $\exists \, \varphi(t, x)$ and $u_0 \in \mathcal{U}$ such that

1) $\varphi(t_0, x_0) = V(t_0, x_0)$

2) $\varphi(t, x) \geq V(t, x) \ \forall \, (t, x)$

3) $\varphi_t(t_0, x_0) + L(x_0, u_0) + \varphi_x(t_0, x_0) \cdot f(x_0, u_0) < -\Theta < 0$

Goal: Show that this control u_0 is "too good to be true," i.e., it provides a cost decrease inconsistent with the principle of optimality (too fast).

Using control u_0 on $[t_0, t_0 + \Delta t]$, from the above three items we have

$$
\begin{aligned}
V(t_0 + \Delta t, x(t_0 + \Delta t)) - V(t_0, x_0) &\leq \varphi(t_0 + \Delta t, x(t_0 + \Delta t)) - \varphi(t_0, x_0) \\
&= \int_{t_0}^{t_0+\Delta t} \frac{d\varphi}{dt}(t, x(t))dt = \int_{t_0}^{t_0+\Delta t} \varphi_t(t, x(t)) + \varphi_x(t, x(t)) \cdot f(x(t), u_0)dt \\
\text{(by continuity, if } \Delta t \text{ is small, using item 3)} \ &\leq -\int_{t_0}^{t_0+\Delta t} L(t, x(t), u_0)dt - \Delta t \cdot \Theta
\end{aligned}
$$

On the other hand, recall the principle of optimality:

$$
V(t_0, x_0) = \inf_u \left\{ \int_{t_0}^{t_0+\Delta t} Ldt + V(t_0 + \Delta t, x(t_0 + \Delta t)) \right\} \leq \int_{t_0}^{t_0+\Delta t} L(t, x(t), u_0)dt + V(t_0 + \Delta t, x(t_0 + \Delta t))
$$

(the second integral is along $x(t)$ generated by $u(t) \equiv u_0$, $t \in [t_0, t_0 + \Delta t]$). This implies

$$
V(t_0 + \Delta t, x(t_0 + \Delta t)) - V(t_0, x_0) \geq -\int_{t_0}^{t_0+\Delta t} Ldt \tag{5.5}
$$

(again along $u(t) \equiv u_0$). We have arrived at a contradiction. □

Try proving that V is a super-solution. The idea is the same, but the contradiction will be that the cost now doesn't decrease fast enough to satisfy the inf condition. See [Bressan] for help, and for all other claims too.

Via uniqueness in the above theorem [Bressan]: Sufficient conditions for optimality now generalize: if V is a viscosity solution of HJB and it is the cost for u^*, then V and u^* are the optimal cost and optimal control.

Chapter 6

The linear quadratic regulator

Will specialize to the case of linear systems and quadratic cost. Things will be simpler, but we'll be able to dig deeper.

Another approach would have been to do this first and then generalize to nonlinear systems. Historically, however, things developed the way we cover them.

References: ECE 515 class notes; Kalman's 1960 paper (posted on the web—must read!), which is the original reference on LQR and other things (controllability); there exist many classic texts: Brockett (FDLS), A-F, Anderson-Moore, Kwakernaak-Sivan (all these should be on reserve at Grainger).

6.1 The finite-horizon LQR problem

System:

$$\dot{x} = A(t)x + B(t)u, \qquad x(t_0) = x_0$$

$x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$ (unconstrained). Cost (1/2 is to match [Kalman], [A-F]):

$$J = \int_{t_0}^{t_1} \frac{1}{2}\left(x^T(t)Q(t)x(t) + u^T(t)R(t)u(t)\right)dt + \frac{1}{2}x^T(t_1)Mx(t_1)$$

where t_1 is a specified final time, $x(t_1)$ is free, $Q(\cdot), R(\cdot), M$ are matrices satisfying $M = M^T \geq 0$, $Q(t) = Q^T(t) \geq 0$, $R(t) = R^T(t) > 0\ \forall t$. We will justify/revise these assumptions later ($R > 0$ because we'll need R^{-1}).

Free t_1, target sets—also possible.

$L = x^TQx + u^TRu$—this is another explanation of the acronym "LQR".

Intuitively, this cost is very reasonable: keep x and u small (remember $Q, R \geq 0$).

6.1.1 The optimal feedback law

Let's derive necessary conditions for optimality using the Maximum Principle first. Hamiltonian:

$$H = -\frac{1}{2}(x^TQx + u^TRu) + p^T(Ax + Bu)$$

(*everything depends on* t). Optimal control u^* must maximize H. (Cf. Problem 11 earlier.)

$$H_u = -Ru + B^T p \;\Rightarrow\; \boxed{u^* = R^{-1}B^T p^*}$$

$H_{uu} = -R < 0 \;\Rightarrow$ this is indeed a maximum (we see that we need the strict inequality $R > 0$ to write R^{-1}). We have a formula for u^* in terms of the adjoint p, so let's look at p more closely. Adjoint equation:

$$\dot{p} = -H_x = Qx - A^T p$$

Boundary condition: $p(t_1) = -K_x(x(t_1)) = -Mx(t_1)$. We'll show that the relation $p^*(t) = (\text{matrix}) \cdot x^*(t)$ holds for all t, not just t_1. Canonical system (using $u^* = R^{-1}B^T p^*$):

$$\begin{pmatrix} \dot{x}^* \\ \dot{p}^* \end{pmatrix} = \begin{pmatrix} A & BR^{-1}B^T \\ Q & -A^T \end{pmatrix} \begin{pmatrix} x^* \\ p^* \end{pmatrix}$$

Consider its transition matrix $\Phi(\cdot,\cdot)$, and partition it as

$$\begin{pmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{pmatrix}$$

(all have two arguments), so

$$\begin{pmatrix} x^*(t) \\ p^*(t) \end{pmatrix} = \begin{pmatrix} \Phi_{11}(t,t_1) & \Phi_{12}(t,t_1) \\ \Phi_{21}(t,t_1) & \Phi_{22}(t,t_1) \end{pmatrix} \begin{pmatrix} x^*(t_1) \\ p^*(t_1) \end{pmatrix}$$

(flowing back, $\Phi(t,t_1) = \Phi^{-1}(t_1,t)$, not true for Φ_{ij}). Using the terminal condition $p^*(t_1) = -Mx^*(t_1)$, we have

$$x^*(t) = (\Phi_{11}(t,t_1) - \Phi_{12}(t,t_1)M)x^*(t_1)$$

and

$$p^*(t) = (\Phi_{21}(t,t_1) - \Phi_{22}(t,t_1)M)x^*(t_1) = (\Phi_{21}(t,t_1) - \Phi_{22}(t,t_1)M)(\Phi_{11}(t,t_1) - \Phi_{12}(t,t_1)M)^{-1}x^*(t)$$

This is formally for now, will address the existence of the inverse later. But at least for $t = t_1$, we have

$$\Phi(t_1,t_1) = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} \;\Rightarrow\; \Phi_{11}(t_1,t_1) = I,\; \Phi_{12}(t_1,t_1) = 0$$

so $\Phi_{11}(t_1,t_1) - \Phi_{12}(t_1,t_1)M = I$—invertible, and by continuity it stays invertible for t near t_1.

To summarize, $p^*(t) = -P(t)x^*(t)$ for a suitable matrix $P(t)$ defined by

$$P(t) = -(\Phi_{21}(t,t_1) - \Phi_{22}(t,t_1)M)(\Phi_{11}(t,t_1) - \Phi_{12}(t,t_1)M)^{-1}$$

which exists at least for t near t_1. (The "$-$" sign is to match a common convention and to have $P \geq 0$ which is later related to the cost.) The optimal control is

$$u^*(t) = R^{-1}(t)B^T(t)p^*(t) = \underbrace{-R^{-1}(t)B^T(t)P(t)}_{=:K(t)}x^*(t)$$

which is a *linear state feedback*!

$\longrightarrow$ Note that K is *time-varying* even if the system is LTI.

$u^*(t) = K(t)x^*(t)$ —this shows that *quadratic costs* are very compatible with *linear systems*. (This idea essentially goes back to Gauss: least square problems $\leftrightarrow$ linear gradient descent laws.)

$\longrightarrow$ Note that the LQR problem does not impose the feedback form of the solution. We could just as easily derive an open-loop formula for u^*:

$$x^*(t_0) = x_0 = (\Phi_{11}(t_0, t_1) - \Phi_{12}(t_0, t_1)M)x^*(t_1)$$

hence

$$p^*(t) = (\Phi_{21}(t, t_1) - \Phi_{22}(t, t_1)M)(\Phi_{11}(t_0, t_1) - \Phi_{12}(t_0, t_1)M)^{-1}x_0$$

again modulo the existence of an inverse. But as long as $P(t)$ above exists, the closed-loop feedback system is well-defined $\Rightarrow$ can solve it and convert from feedback to open-loop form.

But feedback form is nicer in many respects.

$\longrightarrow$ We have linear state feedback *independent* of the Riccati equation (which comes later).

We have $p^*(t) = -P(t)x^*(t)$ but the formula is quite messy.

6.1.2 The Riccati differential equation

Idea: Let's derive a differential equation for P by differentiating both sides of $p = -Px$:

$$\dot{p}^* = -\dot{P}x^* - P\dot{x}^*$$

and using the canonical equations

$$\dot{x}^* = Ax^* + BR^{-1}B^T p^*$$

and

$$\dot{p}^* = Qx^* - A^T p^*$$

Plugging these in, we have

$$Qx^* - A^T p^* = -\dot{P}x^* - PAx^* - PBR^{-1}B^T p^*$$

Finally, using $p^* = -Px^*$,

$$Qx^* + A^T Px^* = -\dot{P}x^* - PAx^* + PBR^{-1}B^T Px^*$$

This must hold along the optimal trajectory. But the initial state x_0 was arbitrary, and $P(t)$ doesn't depend on x_0. So we must have the matrix differential equation

$$\dot{P}(t) = -P(t)A(t) - A^T(t)P(t) - Q(t) + P(t)B(t)R^{-1}(t)B^T(t)P(t)$$

This is the *Riccati differential equation* (RDE).

$\longrightarrow$ Boundary condition: $P(t_1) = M$ (since $p^*(t_1) = -Mx^*(t_1)$ and $x^*(t_1)$ was free).

RDE is a *first-order nonlinear* (quadratic) ODE, a *matrix* one. We can propagate its solution backwards from t_1 to obtain $P(t)$. Global existence is an issue, more on this later.

Another option is to use a previous formula by which we defined $P(t)$—that calls for solving *two first-order* ODEs (canonical system) which are *linear*:

$$\begin{pmatrix} \dot{x} \\ \dot{p} \end{pmatrix} = \begin{pmatrix} A & BR^{-1}B^T \\ Q & -A^T \end{pmatrix} \begin{pmatrix} x \\ p \end{pmatrix}$$

But we don't just need to solve these for some initial condition, we need to have the transition matrix Φ, which satisfies a *matrix* ODE. In fact, let $X(t), Y(t)$ be $n \times n$ matrices solving

$$\begin{pmatrix} \dot{X} \\ \dot{Y} \end{pmatrix} = \begin{pmatrix} A & BR^{-1}B^T \\ Q & -A^T \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix}, \qquad X(t_1) = I,\ Y(t_1) = -M$$

Then we can check that $P(t) := -Y(t)X^{-1}(t)$ satisfies the RDE and boundary condition [ECE 515].

$\longrightarrow$ Cf. earlier discussion of solution of RDE in the context of calculus of variations (sufficient optimality conditions).

Summary so far: optimal control u^* (if it exists) is given by

$$u^* = -R^{-1}B^T \underbrace{Px^*}_{=-p^*}$$

where P satisfies RDE + boundary condition.

$\longrightarrow$ *Note:* if u^* exists then it's automatically unique. Indeed, solution of RDE with a given boundary condition is unique, and this uniquely specifies the feedback law $u^* = Kx^*$ and the corresponding optimal trajectory (x_0 being given).

But the above analysis was via the Maximum Principle so it only gives necessary conditions. What about sufficiency?

Road map: sufficiency; global existence of P; infinite horizon ($t_1 \to \infty$).

6.1.3 The optimal cost

Actually, we already proved (*local*) sufficiency for this control for the LQR problem in Problem 11. To get global sufficiency, let's use sufficient conditions for global optimality in terms of the HJB equation:

$$-V_t = \inf_u \left\{ \frac{1}{2}x^T Qx + \frac{1}{2}u^T Ru + \langle V_x, Ax + Bu \rangle \right\}$$

and boundary condition: $V(t_1, x) = \frac{1}{2}x^T Mx$ The $\inf_u$ on the right-hand side is in this case a minimum, and it's achieved for $u^* = -R^{-1}B^T V_x$. Thus we can rewrite the HJB as

$$\begin{aligned} -V_t &= \frac{1}{2}x^T Qx + \frac{1}{2}(V_x)^T BR^{-1}RR^{-1}B^T V_x + (V_x)^T Ax - (V_x)^T BR^{-1}B^T V_x \\ &= \frac{1}{2}x^T Qx - \frac{1}{2}(V_x)^T BR^{-1}B^T V_x + (V_x)^T Ax \end{aligned}$$

We want to verify optimality of the control

$$u^* = -R^{-1}B^T Px, \qquad P(t_1) = M$$

V didn't appear in the the Maximum Principle based analysis, but now we need to find it. The two expressions for u^*—as well as boundary conditions—would match if

$$V(t, x) = \frac{1}{2}x^T P(t)x$$

This is a very natural guess for the value function. Let's verify that it indeed satisfies the HJB:

$$V_t = \frac{1}{2}x^T \dot{P}x$$

(note that this is partial derivative with respect to t, not total derivative),

$$V_x = Px$$

Plugging this in:

$$-\frac{1}{2}x^T\dot{P}x = \frac{1}{2}x^TQx - \frac{1}{2}x^TP^TBR^{-1}B^TPx + \underbrace{x^TPAx}_{=\frac{1}{2}x^T(PA+A^TP)x}$$

We see that RDE implies that this equality indeed holds. Thus our linear feedback is a (unique) optimal control, and V is the cost.

Let's reflect on how we proved optimality:

- We derived a candidate control from the Maximum Principle ;
- We found a suitable V that let us verify its optimality via the HJB equation.

This is a typical path, suggested by Carathéodory (and Kalman).

Exercise 19 (due Apr 25)

Let $t \le t_1$ be an arbitrary time at which the solution $P(t)$ of RDE exists. *Assumptions—as in the lecture.*

- Prove that $P(t)$ is symmetric;
- Prove that $P(t) \ge 0$;
- Can you prove that $P(t) > 0$? If not, can you prove it under some additional assumption(s)?

6.1.4 Existence of solution for RDE

We still have to address the issue of global existence for $P(t)$. We know from the theory of ODEs that $P(t)$ exists for $t < t_1$ sufficiently close to t_1. Two cases are possible:

1) $P(t)$ exists for all $t < t_1$;
2) There exists a $\bar{t} < t_1$ such that for some i, j, $P_{ij}(t) \to \pm\infty$ as $t \searrow \bar{t}$, i.e., we have finite escape for some entry of P.

We know from calculus of variations that in general, global existence of solutions to Riccati-type equations is not assured, and related to existence of conjugate points. In the present notation, $P(t) = -Y(t)X^{-1}(t)$ and zeros of $X(t)$ correspond (roughly) to conjugate points as we defined them in calculus of variations . However, for the LQR problem, we can prove global existence rather easily:

Suppose Case 2 above holds. We know from Problem 19 that $P(t) = P^T(t) \ge 0$, $t \in (\bar{t}, t_1]$. If an off-diagonal element of P goes to $\pm\infty$ as $t \searrow \bar{t}$ while diagonal elements stay bounded, then some 2×2 principal minor becomes negative—impossible.

$$\begin{pmatrix} * & P_{ij} \\ P_{ij} & * \end{pmatrix}$$

So $P_{ii} \to \infty$ for some i. Let $x_0 = e_i = (0, \dots, 1, \dots, 0)$ with 1 in the i-th position. Then $x_0^T P(t) x_0 = P_{ii}(t) \to \infty$ as $t \searrow \bar{t}$. But $\frac{1}{2} x_0^T P(t) x_0$ is the optimal cost for the system with initial condition x_0 at time t. This cost cannot exceed the cost corresponding to, e.g., control $u(\tau) \equiv 0$, $\tau \in [t, t_1]$. The state trajectory for this control is $x(\tau) = \Phi_A(\tau, t) x_0$, and the cost is

$$\int_t^{t_1} \frac{1}{2} \left(x_0^T \Phi_A^T(\tau, t) Q(\tau) \Phi_A(\tau, t) x_0 \right) d\tau + x_0^T \Phi_A^T(t_1, t) M \Phi_A(t_1, t) x_0$$

which clearly remains bounded, as $t \searrow \bar{t}$, by an expression of the form $\alpha(\bar{t}, t_1) \underbrace{|x_0|^2}_{=1}$, a contradiction. □

So everything is valid for arbitrary t.

So, can we actually compute solutions to LQR problems?

Example 18 (simplest possible, from ECE 515)

$$\dot{x} = u, \qquad x, u \in \mathbb{R}$$

$$J = \int_0^{t_1} \frac{1}{2}(x^2 + u^2) dt$$

$A = 0$, $B = 1$, $Q = 1$, $R = 1$, $M = 0$. RDE: $\dot{P} = -1 + P^2$, P–scalar. The ODE $\dot{z} = -1 + z^2$ doesn't have a finite escape time going backwards (remember that we're concerned with *backward* completeness).

(Digression– But: $\dot{z} = -z^2$ does have finite escape backward in time! We would get this RDE if $Q = 0$, $R = -1$. So the assumption $R \geq 0$ is important. Where did we use our standing assumptions in the above completeness proof?)

Can we solve the RDE from this example?

$$\int_{P(t)}^{0} \frac{dP}{P^2 - 1} = \int_t^{t_1} dt$$

(recall: $P(t_1) = M = 0$). The integral on the left-hand side is $\tanh^{-1}$, so the solution is $P(t) = \tanh(t_1 - t)$ (recall: $\sinh x = \frac{e^x - e^{-x}}{2}, \dots$).

□

So we see that even in such a simple example, solving RDE analytically is not easy. However, the improvement compared to the general nonlinear problems is very large: HJB (*PDE*) → RDE (*ODE*), can solve efficiently by numerical methods.

But: see a need for further simplification.

6.2 The infinite-horizon LQR problem

As in the general case of HJB, we know that the value function becomes time-independent—and the solution of the problem considerably simplifies—if we make suitable additional assumptions. In this case, they are:

- System is LTI: $\dot{x} = Ax + Bu$, A, B–constant matrices;

- Cost is time-independent: $L = \frac{1}{2}(x^T Qx + u^T Ru)$, Q, R–constant matrices;
- No terminal cost: $M = 0$;
- Infinite horizon: $t_1 = \infty$.

The first 3 assumptions—no problem, certainly simplifying. The last one, as we know, requires care (e.g., need to be sure that the cost is $< \infty$, expect this to be related to closed-loop stability). So, let's make the first 3 assumptions, keep t_1 finite for the moment, and a bit later examine the limit as $t_1 \to \infty$.

$$J = \int_{t_0}^{t_1} \frac{1}{2}(x^T Qx + u^T Ru)dt$$

where t_1 is large but finite, treat as parameter. Let's explicitly include t_1 as a parameter in the value function: $V_{t_1}(t_0, x_0)$, or $V_{t_1}(t, x)$, $t_0, t \leq t_1$. We know that this is given by

$$V_{t_1}(t_0, x_0) = \frac{1}{2}x_0^T P(t_0; 0, t_1)x_0$$

where $P(t_0; 0, t_1)$ (0 is the zero matrix) stands for the solution at time t_0 of the RDE

$$\dot{P} = -PA - A^T P - Q + PBR^{-1}B^T P \tag{6.1}$$

with boundary condition $P(t_1) = 0$ (since there is no terminal cost). The optimal control is

$$u^*(t) = -R^{-1}B^T P(t; 0, t_1)x^*(t)$$

6.2.1 Existence and properties of the limit

We want to study $\lim_{t_1 \to \infty} P(t_0; 0, t_1)$. It doesn't necessarily exist in general, but it does under suitable assumptions on (A, B) which guarantee that the cost remains finite.

In force from now on:

Assumption: (A, B) is a controllable pair.

Reason for this assumption: if not, then we may have unstable uncontrollable modes and then no chance of having $V_{t_1} < \infty$ as $t_1 \to \infty$. We see that stabilizability might also be enough, more on this later.

Claim: the limit $\lim_{t_1 \to \infty} P(t_0; 0, t_1)$ exists, and has some interesting properties.

Proof. [Kalman], [Brockett]

$$\frac{1}{2}x_0^T P(t_0; 0, t_1)x_0 = \int_{t_0}^{t_1} \frac{1}{2}\left((x^*)^T Qx^* + (u^*)^T Ru^*\right) dt$$

where u^* is the optimal control for finite-horizon LQR which we found earlier.

It is not hard to see that $x_0^T P(t_0; 0, t_1)x_0$ is a monotonically nondecreasing function of t_1, since $Q \geq 0$, $R > 0$. Indeed, let now u^* be optimal for $t_1 + \Delta t$. Then

$$V_{t_1+\Delta t}(t_0, x_0) = \int_{t_0}^{t_1+\Delta t} \frac{1}{2}((x^*)^T Qx^* + (u^*)^T Ru^*)dt \geq \int_{t_0}^{t_1} \frac{1}{2}((x^*)^T Qx^* + (u^*)^T Ru^*)dt \geq V_{t_1}(t_0, x_0)$$

We would know that it has a limit if we can show that it's bounded. Can we do that?

Use controllability!

There exists a time $\bar{t} > t_0$ and a control $\bar{u}$ which steers the state from x_0 at $t = t_0$ to 0 at $t = \bar{t}$. After $\bar{t}$, set $\bar{u} \equiv 0$. Then

$$V_{t_1}(t_0, x_0) \leq J(t_0, x_0, \bar{u}) = \int_{t_0}^{\bar{t}} \frac{1}{2}(\bar{x}^T Q \bar{x} + \bar{u}^T R \bar{u}) dt \qquad \forall\, t_1 \geq \bar{t}$$

(where $\bar{x}$ is the corresponding trajectory), and this doesn't depend on t_1 and works for all $t_1 \geq \bar{t}$— uniform bound. So, $\lim_{t_1 \to \infty} x_0^T P(t_0; 0, t_1) x_0$ exists.

Does this imply the existence of the matrix limit $\lim_{t_1 \to \infty} P(t_0; 0, t_1)$?

Yes. Let's play the same kind of game as before (when we were proving global existence for $P(t)$) and plug in different x_0.

Plug in $x_0 = e_i \Rightarrow \lim_{t_1 \to \infty} P_{ii}(t_0; 0, t_1)$ exists.

Plug in $x_0 = e_i + e_j = (0, \dots, 1, \dots, 1, \dots, 0)$ (1 in the i-th and j-th spots) $\Rightarrow \lim P_{ii} + 2P_{ij} + P_{jj}$ exists $\Rightarrow$ since $\lim P_{ii}$ and $\lim P_{jj}$ exist, $\lim P_{ij}$ exists.

What are the "interesting properties"?

Does $\lim_{t_1 \to \infty} P(t_0; 0, t_1)$ depend on t_0?

Remember that $P(t_0; 0, t_1)$ is the solution at time t_0 of the *time-invariant* matrix ODE (6.1), and we're passing to the limit as $t_1 \to \infty \implies$ in the limit, *the solution has flown for infinite time backwards* starting from the zero matrix.

So $\lim_{t_1 \to \infty} P(t_0; 0, t_1)$ is a constant matrix, independent of t_0. Call it P_∞, or just P:

$$P = \lim_{t_1 \to \infty} P(t; 0, t_1) \qquad \forall\, t$$

Passing to the limit on both sides of RDE, we see that $\dot{P}(t; 0, t_1)$ must also approach a constant matrix $\Rightarrow$ this constant matrix must be the zero matrix $\Rightarrow P$ satisfies the *algebraic Riccati equation* (ARE)

$$PA + A^T P + Q - PBR^{-1}B^T P = 0 \tag{6.2}$$

$\longrightarrow$ This simplification corresponds to getting rid of V_t in HJB when passing to the infinite-horizon case. But now we're left with no derivatives at all!

A way to think about this limit:

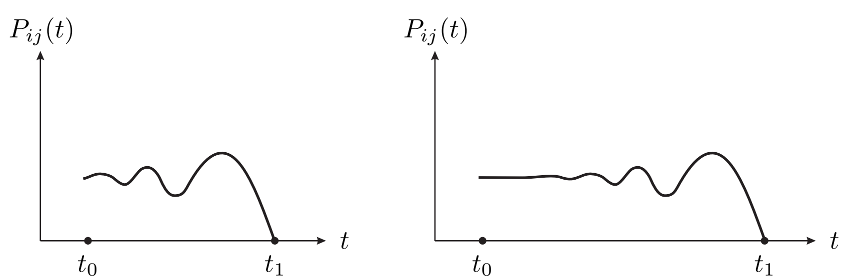

Figure 6.1: Steady-state solution of RDE

As $t_1 \to \infty$, $P(t; 0, t_1)$ approaches *steady state.*