Essentials of optimal control theory in ECON 4140 (2018)

Things you need to know (and a few things you need not care about).

A few words about dynamic optimization in general. Dynamic optimization can be thought of as finding the best *function* rather than the best *point.* We have two tools:

- Dynamic programming. In ECON4140, that is used for discrete-time dynamic optimization. The method involves the optimal *value.* If value depends on state $x \in \mathbb{R}$ and time t and the optimal value is $V_t(x_t)$, then one trades off immediate payoff f (direct utility) against future optimal value (indirect utility) $V_{t+1}(x_{t+1})$. If our control at time t is u_t, $f = f(t, x, u)$ depends on time, state and control, and so does $x_{t+1} = g(t, x, u)$, then the best we can do with state $x_t = x$ is to maximize $f(t, x, u) + V_{t+1}(g(t, x, u))$ wrt. our control u. If V_{t+1} is a known function, that gives us the optimal u_t^* in «feedback form», as a function[1] of time and state.
  - In finite horizon T, we can recurse backwards with known V_T, then V_{T-1}, ...
  - Infinite horizon models has some appealing properties, one of which is that if there is no explicit time in the dynamics and only exponential discounting then the time-dimension vanishes. Using a *current-value* formulation $\beta^t f^{\mathsf{cv}} = f$ and assuming f^{cv} a function of state and control only (no «t») as well as $x_{t+1} = g(x_t, u_t)$ (also without explicit t), we get the Bellman equation

    $$V(x) = \max_u \left\{ f^{\mathsf{cv}}(x, u) + \beta V(g(x, u)) \right\}$$

    with the same V on the LHS and the RHS (there are infinitely many steps left both today and tomorrow). The optimal u is given implicitly in terms of V.

- Calculus of variations or the Pontryagin maximum principle. These methods work by varying the strategy, and do not require the value function. There is no «V» in the Hamiltonian nor in the Euler equation, there is only state and control (and in the calculus of variations method, the control is $\dot{x}$).
  - The discrete-time Euler equation (you have seen it in dynamic macro?) does in a way the same thing: Consider a time-homogeneous problem with current-value formulation $\max \sum_{t=0}^{\infty} \beta^t f^{\mathsf{cv}}(x_t, x_{t+1})$. The first-order condition for optimal state x_τ at a certain time $\tau \in \mathbb{N}$ is found by taking the two terms that involve it (namely $\beta^{\tau-1} f^{\mathsf{cv}}(x_{\tau-1}, x_\tau) + \beta^\tau f^{\mathsf{cv}}(x_\tau, x_{\tau+1})$), differentiating wrt. x_τ and putting equal to zero. Notice: no «V» in that condition.

- ECON4140 uses dynamic programming in discrete time and the maximum principle in continuous time. There exist a continuous-time Bellman equation (often used in stochastic systems) and a discrete-time maximum principle, but those are not at all curriculum.

[1] if the maximizer is not unique, is it then a function? Then, it does not matter which one we choose, so we can pick a function.

The maximum principle. Necessary conditions. Let the timeframe $[t_0, t_1]$ be given[2]. Consider the problem to maximize wrt. $u(t) \in U$ the functional $\int_{t_0}^{t_1} f(t, x(t), u(t))\,dt$ where x starts at $x(t_0) = x_0$ (given) and evolves as $\dot{x}(t) = g(t, x(t), u(t))$; we shall consider the following three possible terminal conditions (a) $x(t_1) = x_1$ (given), (b) $x(t_1) \geq x_1$, or (c) $x(t_1)$ free.

Imagine a trade-off between immediate payoff (or, direct utility) today $f(t, x, u)$ and growth $\dot{x}$ of the *state*. With $\dot{x}(t) = g(t, x, u)$, we weigh immediate payoff at one[3] and weigh growth at $p = p(t)$. Our control is then set to maximize the Hamiltonian

$$H(t, x, u, p) = f(t, x, u) + pg(t, x, u)$$

(over u in the control region U which we are allowed to choose from – it need not be interior). The rest of the maximum principle is about determining a weight p such that this gives us a solution to the dynamic problem. p is often referred to as the «adjoint variable» or «costate» or sometimes «shadow price». The following gives necessary conditions:

0: Form the Hamiltonian $H(t, x, u, p) = f(t, x, u) + pg(t, x, u)$.

1: The optimal u^* maximizes H.

2: The adjoint p satisfies $\dot{p} = -\frac{\partial H}{\partial x}$ (evaluated at optimum), with the so-called *transversality conditions* on $p(t_1)$:
(a') no condition on $p(t_1)$ if the problem has $x(t_1) = x_1$; (b') if the problem imposes $x(t_1) \geq x_1$, then $p(t_1) \geq 0$ with equality if $x^*(t_1) > x_1$ in optimum; (c') if there is no restriction on $x(t_1)$, then $p(t_1)$ must be $= 0$.

3: Also, the differential equation for x must hold: an optimal x^* must satisfy $\dot{x}^*(t) = g(t, x^*(t), u^*(t))$ with initial condition $x^*(t_0) = x_0$ and if applicable, the terminal condition.

These conditions may be regarded as a solution steps recipe although in practice it may not be so straightforward as to call it a «cookbook». Next page:

[2] On page 7: there are problems where time can be optimized too.

[3] Here there is a theoretical catch which is not exam relevant, except see the second bullet below in order not to be confused by any «p_0»:
Suppose that there is no «optimization», and that there is only one control $u^*(t)$ such that the terminal condition holds. If your control has to be reserved to fulfill that condition, then you cannot optimize for utility. Then the weight on f has to be zero. That is the «p_0» constant in the book, which looks a bit akin to the *Fritz John* type conditions covering the constraint qualification in nonlinear programming.

- You can disregard the p_0 (i.e., put it equal to one) for exam purposes.
- *But:* Do *not* put $p(t_0)$ equal to one, because the constant p_0 is *not* the same thing as $p(t_0)$! (Nor the same as $p(0)$. In case you wonder what the notation is about: it is from the case with several states $\mathbf{x} \in \mathbf{R}^n$. Then we have an n-dimensional $\mathbf{p}(t)$, and the p_0 is then the «zeroeth» dimension.)

More «cookbook»-alike solution steps:

step 0: Form the Hamiltonian $H(t, x, u, p) = f(t, x, u) + pg(t, x, u)$.

step 1: The optimal u^* maximizes H.

- Whatever state x and costate p might be, then that gives us a relation between u^* and (t, x, p). With the possible reservation that the maximizer may not be unique[4], this gives us u^* as a function

$$\hat{u} \quad \text{of} \quad (t, x, p)$$

where $x = x^*$ is the optimal state, and p is the adjoint satisfying the next step. (Note that in practice you may have to split between cases.)

step 2: We have a differential equation for p:
$\dot{p} = -\frac{\partial H}{\partial x}$ (evaluated at optimum),
with the so-called *transversality conditions* on $p(t_1)$:

(i) In case the terminal value $x(t_1)$ is fixed, there is no condition on $p(t_1)$.

(ii) In case the problem imposes $x(t_1) \geq x_1$, then we get a complementary slackness condition on $p(t_1)$: it is ≥ 0, with equality if $x(t_1) > x_1$ (the latter corresponds to the next item).

(iii) If there is no restriction on $x(t_1)$, then $p(t_1)$ must be $= 0$.

If we have a function $\hat{u}(t, x, p)$ for the optimal control, then plugging this into $-\frac{\partial H}{\partial x}$ will give $\dot{p}$ as a function of (t, x^*, p).

step 3: Then we have the differential equation for the state. Inserting $\hat{u}$ there as well, gives a differential equation system

$$\dot{x}^* = \phi(t, x^*, p), \qquad \dot{p} = \psi(t, x^*, p)$$

and the conditions on $x(t_0)$, $x(t_1)$ and $p(t_1)$ determine the integration constants.

Sufficient conditions. We have two sets of sufficient conditions. Suppose we have found a pair (x^*, u^*) which satisfies the necessary conditions. This pair is a candidate for optimality. We can conclude that it is indeed optimal if it satisfies one of the following:

- The *Mangasarian sufficiency condition:* With the $p = p(t)$ that the maximum principle produces, then H is concave wrt. (x, u) for all $t \in (t_0, t_1)$.
- The *Arrow sufficiency condition:* Insert the function $\hat{u}(t, x, p)$ for u in the Hamiltonian to get the function $\hat{H}(t, x, p) = H(t, x, \hat{u}(t, x, p), p)$. With the $p = p(t)$ that the maximum principle produces, then $\hat{H}$ is concave wrt. x for all $t \in (t_0, t_1)$.

The Arrow condition is more powerful: if Mangasarian applies, then Arrow will always apply. However, the Mangasarian condition could be easier to verify.

[4] in which case several functions are possible. For necessary conditions, you should consider them all. For sufficient conditions, you may «guess one and verify that it solves». See the «modified example 2».

«Example» 1 (not yet worked out completely). A concave problem. In matrix notation, let $f(t,x,u) = \Big[(k_1,\ k_2)\binom{x}{u} - \frac{1}{2}(x,u)\mathbf{Q}\binom{x}{u}\Big]e^{-rt}$ where $\mathbf{Q} = \begin{pmatrix} q_{11} & q_{12} \\ q_{12} & q_{22} \end{pmatrix}$ is symmetric and positive semidefinite with $q_{22} > 0$, and let $g(t,x,u) = (m_1,\ m_2)\binom{x}{u}$. Suppose u can take any real value. No matter what the terminal condition on x is, we will have $H(t,x,u,p) = f + pg$ is concave in (x,u) regardless of the sign of p, so Mangasarian will apply to the solution we find as follows[5]:

step 0: We have

$$\begin{aligned} H(t,x,u,p) &= \Big[(k_1,\ k_2)\binom{x}{u} - \tfrac{1}{2}(x,u)\mathbf{Q}\binom{x}{u}\Big]e^{-rt} + p(m_1,\ m_2)\binom{x}{u} \\ &= \big[k_1 x - \tfrac{1}{2}q_{11}x^2 + (k_2 - q_{12}x)u - \tfrac{1}{2}q_{22}u^2\big]e^{-rt} + m_1 px + m_2 pu \end{aligned}$$

step 1: The optimal u^* maximizes H, and can be written as $\hat{u}(t,x^*,p)$ where (from the first-order condition for u)

$$\hat{u}(t,x,p) = \frac{(k_2 - q_{12}x)e^{-rt} + m_2 p}{q_{22}}$$

step 2: The differential equation for p is $\dot{p} = (q_{11}x - k_1 + q_{12}u)e^{-rt} - m_1 p$ to be evaluated at optimum; that is, we insert x^* for x and $\hat{u}(t,x^*,p)$ for u:

$$\dot{p} = \big(q_{11}x^* - k_1 + q_{12}\frac{(k_2 - q_{12}x^*)e^{-rt} + m_2 p}{q_{22}}\big)e^{-rt} - m_1 p$$

Notice this is linear in (x^*, p) (we should reorder coefficients, but unfortunately it is not homogeneous). Then we have the transversality conditions corresponding to whatever the terminal conditions were for x.

step 3: Inserting $\hat{u}$ in the differential equation for x^*, we find

$$\dot{x}^* = m_1 x^* + m_2\hat{u}(t,x^*,p) = m_1 x^* + m_2\frac{(k_2 - q_{12}x^*)e^{-rt} + m_2 p}{q_{22}}$$

Again, this is linear in (x^*, p).

The problem then reduces to solving a linear equation system, which unfortunately has time-dependent coefficients. The current-value formulation below resolves that!

[5] Indeed, we can assume U to be an interval as well, and concavity will still hold – but then the problem becomes harder when u hits and endpoint.

Example 2 (H not concave wrt. (x, u), but Arrow's condition applies) This looks like one given in class, but modified so that Mangasarian does not apply. Suppose u can take values in $[0, K]$ where $K > 0$ is a constant, and consider the problem $\max \int_0^T e^{-\delta t} u^2 dt$ subject to $\dot{x} = -u$, $x(0) = x_0$ (a constant > 0) and $x(T)$ required to be ≥ 0. We therefore assume $x_0 < KT$ (otherwise, we would have $x(t) \geq 0$ automatically). Let in this problem $\delta > 0$. Notice that the Hamiltonian $H(t, x, u, p) = e^{-\delta t} u^2 - pu$ is convex in u. That means (a) that the maximizing u^* is either 0 or K, and (b) we cannot use Mangasarian. But we can use Arrow: since x does not enter H, then inserting for $\hat{u}$ (which does not depend on x, as the maximization does not) we get no x in $\hat{H}$. Considering $\hat{H}$ a function of x only, it is constant, and that is concave. (Not strictly, but we do not need that.) So whatever we get out of the following, will indeed be optimal. Let us work out the steps.

step 0: Define $H(t, x, u, p) = e^{-\delta t} u^2 - pu$.

step 1: The optimal u^* maximizes H. By convexity, we must have either the endpoint 0 or the endpoint K, and we just compare the two, $e^{-\delta t} K^2 - pK$ vs. zero. We have

$$\hat{u} = \begin{cases} K & \text{if} \quad K > pe^{\delta t} \\ 0 & \text{if} \quad K < pe^{\delta t} \\ 0 \text{ or } K & \text{if} \quad K = pe^{\delta t} \quad \text{(the maximization cannot tell which one)} \end{cases}$$

This condition can not determine $\hat{u}$ in the case $pe^{\delta t} = K$. If that happens at only one point in time (or never), then it is not a problem, as changing an integrand at a single point does not change any state. If we were to have $p(t) = Ke^{-\delta t}$ on an entire positive interval, we would be in trouble (although, by sufficient conditions, we could try to guess and verify!).

step 2: Because H does not depend on x, then $\dot{p} = 0$, so p is a constant P. By the transversality condition, $P \geq 0$ with equality if $x^*(T) > 0$.

- Good news! The «?» case for $\hat{u}$ will not be an issue: there can only be at most t for which $K = Pe^{\delta t}$. If there is one such, then we switch control (and we switch from K to 0 ... exercise: why?).

step 3: We have $\dot{x} = -K$ (if $K > Pe^{\delta t}$) or $= 0$ (if the reverse inequality holds). We need to determine when we have what.

So now we have the conditions, and we can start to «nest» out what could happen.

- Could we have $x(T) > 0$? Then we must have $p(T) = 0$ hence $P = 0$. Then we would always have $K > Pe^{\delta t}$ and always $u^* = K$. But then $x^*(T) = x_0 - KT$ which by assumption is ≤ 0. Contradiction! So $x(T) = 0$.

- Indeed, we cannot have $u^* \equiv K$. Therefore, we must have $K = Pe^{\delta t^*}$ for some (necessarily unique) $t^* \in (0, T)$. Then $u^* = K$ on $(0, t^*)$ and 0 afterwards. Adjust then t^* so that we hit zero there: $t^* = x_0/K$.

So we must choose u^* maximally until we hit zero, and keep x^* constant at zero from then on. By Arrow's condition, this is indeed the optimal solution.

Modified example 2. Now drop the assumption that $\delta > 0$. The case $\delta < 0$ will not add much insight – we will push the $u^* = K$ period to the end instead. But suppose now $\delta = 0$, and let us see what happens.

step 0: Define $H(t, x, u, p) = u^2 - pu$.

step 1: The optimal u^* maximizes H. Again, we have endpoint solution:

$$\hat{u} = \begin{cases} K & \text{if} \quad K > p \\ 0 & \text{if} \quad K < p \\ (0 \text{ or } K & \text{if} \quad K = p\text{; the maximization cannot tell which one)} \end{cases}$$

step 2: Again, p is a constant $P \geq 0$ (equal to zero if $x^*(T) > 0$.)

step 3: The differential equation for the state must hold.

Again we get a contradiction if we assume $x(T) > 0$. So $x(T) = 0$. In particular, that means we cannot have $u^* \equiv 0$. Therefore, we must have $K \geq P$. And we cannot have $u^* \equiv K$, as that yields $x^*(T) < 0$. So $K \leq P$.

- With $K = P$, the necessary conditions cannot tell us whether to use $u^* = 0$ or $u^* = K$.
- But we know that we must have $u^* = K$ for precisely long enough to end up at zero. The necessary conditions cannot tell *when* to run at full throttle.
- Actually, it does not matter. No matter when, we would get the same performance Kx_0. But that is maybe not so easy to see, was it?
- Well let us argue as follows: in the case $\delta > 0$, we had $u^* = K$ up to $t^* = x_0/K$. Let us just make the guess that this is optimal.
  - It does satisfy all the conditions from the maximum principle!
  - By the Arrow condition, it is optimal.

  (It just isn't uniquely optimal. In fact, all the other «$u^* = K$ for a period totalling x_0/K in length» will be optimal – and Arrow's condition will verify that!)

Even more modified example 2. Restrict $u(t)$ to being *either* zero or K. We know already that we have optimal solutions for the previous modification, with that property, so they must be optimal here as well. But take note that Arrow's condition works even then U is not a convex set, and could be used to verify optimality!

Existence/uniqueness of optimal control: no general results on curriculum. The only «exam relevant» concerning uniqueness is if you can show that only one control satisfies necessary conditions. The only «exam relevant» concerning existence is if you have found a solution by sufficient conditions, or – for calculus of variations, mainly, when conditions are the same for min and max – when it is clear that *no* solution exists.

Sensitivity. The optimal value V depends on (t_0, x_0, t_1, x_1) although no «x_1» if $x(t_1)$ is free. With exception for the latter, and to the level of precision of this course, we have the following sensitivity properties:

- $\frac{\partial V}{\partial x_0} = p(t_0)$.
- $\frac{\partial V}{\partial x_1} = -p(t_1)$ except in the free-end case. Note that the «x_1» variable is a *constraint* you have to fulfil, so the interpretation is that $p(t_1)$ is the marginal loss of tightening it by requiring you to leave one more unit at the table in the end.
- $\frac{\partial V}{\partial t_1} = H(t_1, x_1^*, u^*(t_1), p(t_1))$ is the marginal value of having one more time unit in the end. That means, that if you were actually allowed to choose *when* to stop, then the first-order condition would be $H = 0$ at the final time. (But for optimal stopping, the sufficient conditions presented herein are no longer valid!)
- $\frac{\partial V}{\partial t_0} = -H(t_0, x_0, u^*(t_0), p(t_0))$. The minus sign because increasing t_0 gives you one unit less of time.

Note that $-p(t_1)$ is the t_0-present value. For current value formulation, see the next page

Infinite horizon: not curriculum. Conditions for infinite horizon are not curriculum. At worst, you could be asked what happens when t_1 becomes large. (That is, limit of finite horizon problems.)

That means that phase planes for infinite horizon problems are not curriculum per se. But phase planes for differential equation systems *are* curriculum per se, and could also be helpful for solving the optimal control problems you *can* be asked to handle. See next page and examples in lectures.

Variable final time to be optimized. (Not to be stressed in this semester, but an example could be given in class. The Hotelling rule, ...)

In some problems, the upper horizon t_1 is subject to optimization. Suppose that the optimal t_1 is $> t_0$ (so that it is interior); the first-order condition is then $\frac{\partial V}{\partial t_1}$, that is, $H(t_1, x_1^*, u^*(t_1), p(t_1)) = 0$. This and the maximum principle forms *necessary* conditions.

Note: When t_1 is free, we no longer have sufficient conditions. (Not in this course nor textbook! There are some, given in Seierstad and Sydsæter's 1987 textbook on optimal control theory.)

[Current values: fits one page, so pagebreak here.]

Current-value formulation. You will not be asked directly to know it (it has been given at the exam, but then with a hint on what to do to get it) – but it could be very helpful, especially in the following case: Suppose running utility has exponential discounting and there is no other «explicit time-dependence» (and in particular so in infinite horizon, which is *not* curriculum. Here is what happens: Suppose that $g(t, x, u)$ does not depend on t directly, and that $f(t, x, u) = e^{-rt} f^{\mathtt{cv}}(x, u)$, the «cv» for «current-value». Define $\lambda = e^{rt}p$. Then $H(t, x, u, p) = e^{-rt}\big[f^{\mathtt{cv}} + pe^{rt}g\big]$ which equals $e^{-rt}H^{\mathtt{cv}}(x, u, \lambda)$ where $H^{\mathtt{cv}}(x, u, \lambda) = f^{\mathtt{cv}}(x, u) + \lambda g(x, u)$ is called the current-value Hamiltonian. (There is literature where that is just called «Hamiltonian» as well.) We have $\dot{p} = \frac{d}{dt}\lambda e^{-rt} = [\dot{\lambda} - r\lambda]e^{-rt}$ which equals $-e^{-rt}\frac{\partial H^{\mathtt{cv}}}{\partial x}$, so we get

$$\dot{\lambda} - r\lambda = -\frac{\partial H^{\mathtt{cv}}}{\partial x}$$

with the same transversality conditions for λ as for p. The optimal u^* maximizes $H^{\mathtt{cv}}$ and – whenever unique – will be given as a function $\hat{u}^{\mathtt{cv}}(x, \lambda)$.

Example 1 revisited. With current-value formulation, we get

step 0: Current-value Hamiltonian

$$H^{\mathtt{cv}}(x, u, \lambda) = k_1 x - \tfrac{1}{2}q_{11}x^2 + (k_2 - q_{12}x)u - \tfrac{1}{2}q_{22}u^2 + m_1\lambda x + m_2\lambda u$$

step 1: The optimal u^* can be written as $\dfrac{k_2 - q_{12}x + m_2\lambda}{q_{22}}$

step 2: We get the differential equation $\dot{\lambda} - r\lambda = q_{11}x - k_1 + q_{12}u - m_1\lambda$ to be evaluated at optimum:

$$\dot{\lambda} - r\lambda = q_{11}x^* - k_1 + q_{12}\frac{k_2 - q_{12}x^* + m_2\lambda}{q_{22}} - m_1\lambda$$

Linear in (x^*, λ), and now the coefficients are constant.

step 3: Inserting $\hat{u}$ in the differential equation for x^*, we find

$$\dot{x}^* = m_1x^* + m_2\frac{k_2 - q_{12}x^* + m_2\lambda}{q_{22}}$$

Linear in (x^*, λ), and now the coefficients are constant.

Because of the constant coefficients, we can solve this system completely. Align the integration constants to $x(t_0) = x_0$ and the terminal/transversality conditions, and we have solved the maximization problem.

Phase planes. Also «requires» a system without explicit t-dependence («autonomous»), so the current-value formulation is most helpful here too. See example given in class.