Lecture Notes 8: Dynamic Optimization
Part 1: Calculus of Variations

Peter J. Hammond

Revised 2023 September 20th
typeset from `calcVar23.tex`

Outline

Vickrey–Mirrlees Model

Problem: how much to pay workers of different skills.

Goal: achieve fairness while preserving incentives.

References: William S. Vickrey (1945)
"Measuring Marginal Utility by Reactions to Risk"
Econometrica 13: 319–333.

James A. Mirrlees (1971)
"An Exploration in the Theory of Optimal Income Taxation"
Review of Economic Studies 38: 175–208.

Let $n \in \mathbb{R}_+$ denote a person's skill level, defined to mean that there is a constant rate of marginal substitution of n_1/n_2 between hours of work supplied by workers of the two skill levels n_1 and n_2.

Thus, a worker's productivity is proportional to n, personal skill.

Assume that the distribution of workers' skills can be described by a continuous density function $\mathbb{R}_+ \ni n \mapsto f(n) \in \mathbb{R}_+$ which, like a probability density function, satisfies $\int_0^\infty f(n)dn = 1$.

Objective and Constraints

We consider a "macro" model
with a "representative consumer/worker",
whose preferences for consumption/labour supply pairs $(c, \ell) \in \mathbb{R}^2_+$
are represented by the utility function $u(c) - v(\ell)$,
where $u' > 0$, $v' > 0$, $u'' < 0$, and $v'' > 0$.

The problem is to maximize a social objective by choosing
the pair of functions $\mathbb{R}_+ \ni n \mapsto (c(n), \ell(n)) \in \mathbb{R}^2_+$.

The maximand is "mean utility" which is specified
by the utility integral $\int_0^\infty [u(c(n)) - v(\ell(n))] f(n) \, \mathrm{d}\, n$.

The resource balance constraint takes the form $C \leq F(L)$ where

- $C := \int_0^\infty c(n) f(n) dn$ is mean consumption;
- $L := \int_0^\infty n \, \ell(n) f(n) dn$ is mean effective labour supply.

The aggregate production function $\mathbb{R}_+ \ni L \mapsto F(L) \in \mathbb{R}_+$
is assumed to satisfy $F'(L) > 0$ and $F''(L) \leq 0$ for all $L \geq 0$.

Pseudo First-Order Conditions

Consider the Lagrangian

$$\mathcal{L}(c(\cdot), \ell(\cdot)) := \int_0^\infty [u(c(n)) - v(\ell(n))] f(n) dn \\ - \lambda \left[\int_0^\infty c(n) f(n) dn - F\left(\int_0^\infty n\ell(n) f(n) dn \right) \right]$$

as a functional (rather than a mere function)

of the functions $\mathbb{R}_+ \ni n \mapsto (c(n), \ell(n)) \in \mathbb{R}_+^2$.

We derive "pseudo" first-order conditions by pretending

that the derivatives $\dfrac{\partial \mathcal{L}}{\partial c(n)}$ and $\dfrac{\partial \mathcal{L}}{\partial \ell(n)}$ both exist, for all $n \geq 0$.

This gives the pseudo first-order conditions

$$0 = \frac{\partial \mathcal{L}}{\partial c(n)} = [u'(c(n)) - \lambda] f(n)$$
$$0 = \frac{\partial \mathcal{L}}{\partial \ell(n)} = -v'(\ell(n)) f(n) + \lambda F'(L) n f(n)$$

The Marxist First Best

Karl Marx (1875) *Critique of the Gotha Programme*

For any skill level n such that $f(n) > 0$, we have the two equations

$$0 = [u'(c(n)) - \lambda]f(n) = -v'(\ell(n))f(n) + \lambda F'(L)nf(n)$$

These imply that, for any skill level n with $f(n) > 0$, we want:

- $u'(c(n)) = \lambda$ and so $c(n) = c^*$,
  where the constant c^* uniquely solves $u'(c^*) = \lambda$
  ("to each according to their need");
- $v'(\ell(n)) = \lambda F'(L)n$, implying that $v''(\ell(n)) \cdot \dfrac{d\ell}{dn} = \lambda F' > 0$,
  so $\dfrac{d\ell}{dn} > 0$ ("from each according to their ability")

Exercise

Use concavity and convexity arguments to prove that this is the (essentially unique) solution.

What makes this solution practically infeasible?

Sufficiency Theorem: Statement

Theorem

Suppose that there exists $\lambda > 0$
such that c^* *and the function* $\mathbb{R}_+ \ni n \mapsto \ell^*(n)$
jointly satisfy the first-order conditions:

$$u'(c^*) = \lambda \quad \text{and} \quad v'(\ell^*(n)) = \lambda F'(L^*) n \quad \text{for all } n \in \mathbb{R}_+$$

where $c^* = F(L^*)$ *and* $L^* = \int_0^\infty n \, \ell^*(n) f(n) \, dn$.

Let $\mathbb{R}_+ \ni n \mapsto (c(n), \ell(n)) \in \mathbb{R}_+^2$
be any other policy satisfying $C = F(L)$
where $C = \int_0^\infty c(n) f(n) \, dn$ *and* $L = \int_0^\infty n \ell(n) f(n) \, dn$.

Then

$$\int_0^\infty [u(c(n)) - v(\ell(n))] f(n) dn \leq u(c^*) - \int_0^\infty v(\ell^*(n)) f(n) dn$$

with strict inequality unless $c(n) = c^*$ *wherever* $f(n) > 0$.

Sufficiency Theorem: Proof, I

Because $u'' < 0$ and so u is strictly concave,
the supergradient property of concave functions implies that

$$u(c(n)) - u(c^*) \leq u'(c^*)[c(n) - c^*] = \lambda[c(n) - c^*]$$

for all n, with strict inequality unless $c(n) = c^*$.

Integrating this inequality gives the first integral inequality

$$\int_0^\infty [u(c(n)) - u(c^*)]\, f(n)\, \mathrm{d}\, n \leq \lambda(C - c^*)$$

with strict inequality unless $c(n) = c^*$ wherever $f(n) > 0$.

Similarly, because $v'' \geq 0$ and so v is convex, for all n
the subgradient property of convex functions implies that

$$v(\ell(n)) - v(\ell^*(n)) \geq v'(\ell^*(n))[\ell(n) - \ell^*(n)] = \lambda F'(L^*)[\ell(n) - \ell^*(n)]$$

Integrating this inequality gives the second integral inequality

$$\int_0^\infty [v(\ell(n)) - v(\ell^*(n))]\, f(n)\, \mathrm{d}\, n \geq \lambda F'(L^*)(L - L^*)$$

Sufficiency Theorem: Proof, II

Subtracting the second integral inequality from the first, then rearranging, one has

$$D := \int_0^\infty \{[u(c(n)) - v(\ell(n))] - [u(c^*) - v(\ell^*(n))]\} f(n) \, \mathrm{d}\, n$$
$$\leq \lambda[(C - c^*) - F'(L^*)(L - L^*)]$$

Note that: (i) $C \leq F(L)$, by feasibility; (ii) $c^* = F(L^*)$;
(iii) because $F'' \leq 0$ and so F is concave,
one has $F(L) - F(L^*) \leq F'(L^*)(L - L^*)$.

It follows that $C - c^* \leq F(L) - F(L^*) \leq F'(L^*)(L - L^*)$ and so

$$C - c^* - F'(L^*)(L - L^*) \leq 0$$

Then, because $\lambda > 0$, the above definition of D implies that $D \leq 0$.

This proves that no feasible policy $\mathbb{R}_+ \ni n \mapsto (c(n), \ell(n)) \in \mathbb{R}^2_+$ can yield more more mean utility $\int_0^\infty \{[u(c(n)) - v(\ell(n))] f(n) \, \mathrm{d}\, n$ than the policy $\mathbb{R}_+ \ni n \mapsto (c^*, \ell^*(n)) \in \mathbb{R}^2_+$ does.

The latter policy is therefore optimal. □

Outline

Problem Formulation

The calculus of variations is used to optimize a functional that maps functions into real numbers.

A typical problem is to choose a path $\mathbf{x}$,
in the form of a function $[t_0, t_1] \ni t \mapsto x(t) \in \mathbb{R}$,
in order to maximize the integral objective functional

$$J(\mathbf{x}) = \int_{t_0}^{t_1} F(t, x(t), \dot{x}(t)) \, \mathrm{d}\, t$$

subject to the fixed end point conditions $x(t_0) = x_0$, $x(t_1) = x_1$.

A variation involves moving away from the first path $\mathbf{x}$
to the variant path $\mathbf{x} + \epsilon \mathbf{u}$,
where $\mathbf{u}$ denotes the differentiable function $[t_0, t_1] \ni t \mapsto u(t) \in \mathbb{R}$,
and $\epsilon \in \mathbb{R}$ is a (small) scalar.

To ensure that the end point conditions $x(t_0) + \epsilon u(t_0) = x_0$
and $x(t_1) + \epsilon u(t_1) = x_1$ remain satisfied by $\mathbf{x} + \epsilon \mathbf{u}$,
one imposes the conditions $u(t_0) = u(t_1) = 0$.

Toward a Necessary First-Order Condition

A maximum of $J(\mathbf{x}) = \int_{t_0}^{t_1} F(t, x(t), \dot{x}(t))\, \mathrm{d}\, t$ is a path $\mathbf{x}^*$
or function $[t_0, t_1] \ni t \mapsto x^*(t) \in \mathbb{R}$:
(i) that satisfies the end point conditions $x^*(t_0) = x_0$, $x^*(t_1) = x_1$;
(ii) with the property that $J(\mathbf{x}^*) \geq J(\mathbf{x})$
for any alternative path $\mathbf{x} = (x(t))_{t \in [t_0, t_1]}$
that also satisfies the end point conditions $x(t_0) = x_0$, $x(t_1) = x_1$.

A necessary condition for $\mathbf{x}^*$ to maximize $J(\mathbf{x})$ w.r.t. $\mathbf{x}$
is that $J(\mathbf{x}^*) \geq J(\mathbf{x}^* + \epsilon \mathbf{u})$ for all small ϵ.

Alternatively, the function

$$\mathbb{R} \ni \epsilon \mapsto f_{\mathbf{x}^*, \mathbf{u}}(\epsilon) := J(\mathbf{x}^* + \epsilon \mathbf{u}) \in \mathbb{R}$$

must satisfy, for all small ϵ, the inequality

$$f_{\mathbf{x}^*, \mathbf{u}}(0) = J(\mathbf{x}^*) \geq J(\mathbf{x}^* + \epsilon \mathbf{u}) = f_{\mathbf{x}^*, \mathbf{u}}(\epsilon)$$

In case the function $\epsilon \mapsto f_{\mathbf{x}^*, \mathbf{u}}(\epsilon)$ is differentiable at $\epsilon = 0$,
a necessary first-order condition is therefore $f'_{\mathbf{x}^*, \mathbf{u}}(0) = 0$.

Evaluating the Derivative

Our definitions of the functions J and $f_{\mathbf{x}^*,\mathbf{u}}$ imply that

$$f_{\mathbf{x}^*,\mathbf{u}}(\epsilon) = J(\mathbf{x}^* + \epsilon\mathbf{u}) = \int_{t_0}^{t_1} F(t, x^*(t) + \epsilon u(t), \dot{x}^*(t) + \epsilon \dot{u}(t))\,\mathrm{d}\,t$$

By Leibniz's formula, the derivative $f'_{\mathbf{x}^*,\mathbf{u}}(0)$ w.r.t. ϵ at $\epsilon = 0$ equals the derivative of the integrand.

It follows that $f'_{\mathbf{x}^*,\mathbf{u}}(0) = \int_{t_0}^{t_1} [F'_x(t)u(t) + F'_{\dot{x}}(t)\dot{u}(t)]\,\mathrm{d}\,t$ where,
for each $t \in [t_0, t_1]$,
the partial derivatives $F'_x(t)$ and $F'_{\dot{x}}(t)$ of $F(t, x, \dot{x})$
are evaluated at the triple $(t, x^*(t), \dot{x}^*(t))$.

Integrating by Parts

The product rule for differentiation implies that

$$\frac{d}{dt}[F'_{\dot{x}}(t)u(t)] = \left[\frac{d}{dt}F'_{\dot{x}}(t)\right] u(t) + F'_{\dot{x}}(t)\dot{u}(t)$$

and so, integrating by parts, one has

$$\int_{t_0}^{t_1} F'_{\dot{x}}(t)\dot{u}(t)dt = |_{t_0}^{t_1} F'_{\dot{x}}(t)u(t) - \int_{t_0}^{t_1} \left[\frac{d}{dt}F'_{\dot{x}}(t)\right] u(t)\,\mathrm{d}\,t$$

But the end point conditions imply that $u(t_0) = u(t_1) = 0$, so the first term on the right-hand side vanishes.

The Euler Equation

Substituting $-\int_{t_0}^{t_1}\left[\frac{d}{dt}F'_{\dot{x}}(t)\right]u(t)\,\mathrm{d}\,t$ for the term $\int_{t_0}^{t_1}F'_{\dot{x}}(t)\dot{u}(t)\,\mathrm{d}\,t$ in the equation $f'_{\mathbf{x}^*,\mathbf{u}}(0)=\int_{t_0}^{t_1}[F'_x(t)u(t)+F'_{\dot{x}}(t)\dot{u}(t)]\,\mathrm{d}\,t$, then recognizing the common factor $u(t)$, we finally obtain

$$f'_{\mathbf{x}^*,\mathbf{u}}(0)=\int_{t_0}^{t_1}\left[F'_x(t)-\frac{d}{dt}F'_{\dot{x}}(t)\right]u(t)\,\mathrm{d}\,t$$

The first-order condition is $f'_{\mathbf{x}^*,\mathbf{u}}(0)=0$
for every differentiable function $t\mapsto u(t)$
satisfying the two end point conditions $u(t_0)=u(t_1)=0$.

This condition holds if and only if, for (almost) all $t\in[t_0,t_1]$,
the integrand is zero, or equivalently,
if and only if the Euler equation $\frac{d}{dt}F'_{\dot{x}}(t)=F'_x(t)$ holds.

Outline

Are We Saving Too Little?

Kenneth Arrow, Gretchen Daily, Partha Dasgupta, Paul Ehrlich, Lawrence Goulder, Geoffrey Heal, Simon Levin, Karl-Göran Mäler, Stephen Schneider, David Starrett and Brian Walker (2004) "Are We Consuming Too Much?" *Journal of Economic Perspectives* 18: 147–172.

Macroeconomic variation in the Solow–Swan growth model.

Given a capital stock K, output Y is given by the production function $Y = f(K)$, where $f' > 0$, and $f'' \leq 0$.

Net investment = gross investment, without depreciation.

So given capital K and consumption C, investment I is given by

$$I = \dot{K} = f(K) - C$$

The Ramsey Problem and Beyond

The economy's intertemporal objective is taken to be

$$\int_0^T e^{-rt} U(C(t))\,\mathrm{d}\,t = \int_0^T e^{-rt} U(f(K) - \dot{K})\,\mathrm{d}\,t$$

Frank Ramsey (*Economic Journal*, 1928)
assumed $T = \infty$ (infinite horizon) and $r = 0$ (no discounting).

Nicholas Stern (of the *Stern Review* on Climate Change)
and others advocate:

- $T = \infty$;
- r as the hazard rate in an exogenous Poisson process that determines the latest date at which extinction occurs;
  this implies that e^{-rt} is the exogenous maximum probability that the human race has not become extinct by time t.

Chichilnisky, Hammond, and Stern in a special (2020) issue
of *Social Choice and Welfare* honouring Kenneth Arrow.

Applying the Calculus of Variations

We apply the calculus of variations
to the objective $\int_0^T e^{-rt} U(f(K) - \dot{K}) \, \mathrm{d}\, t$
with the end conditions $K(0) = \bar{K}$, which is exogenous,
and $K(T) = 0$ at the finite time horizon T.

Euler's equation takes the form $\frac{d}{dt} F'_{\dot{K}}(t) = F'_K(t)$
where $F(t, K, \dot{K}) = e^{-rt} U(f(K) - \dot{K}) = e^{-rt} U(C)$.

So Euler's equation becomes $\frac{d}{dt}[-e^{-rt} U'(C)] = e^{-rt} U'(C) f'(K)$.

Equivalently, after evaluating the time derivative,

$$-U''(C)\dot{C}e^{-rt} + rU'(C)e^{-rt} = e^{-rt} U'(C) f'(K)$$

Cancelling the common factor e^{-rt} and dividing by $U'(C) > 0$,
then rearranging, one obtains

$$-\frac{U''(C)}{U'(C)} \dot{C} = f'(K) - r$$

Further Interpretation

Define the (negative) elasticity of marginal utility as

$$\eta(C) := -\frac{d \ln U'(C)}{d \ln C} = -\frac{U''(C)C}{U'(C)}$$

This is related to the curvature of the utility function,
and to how quickly marginal utility $U'(C)$ decreases as C increases.

Rearranging the equation $-U''(C)\dot{C}/U'(C) = f'(K) - r$ yet again, one obtains the equation

$$\eta(C)\frac{\dot{C}}{C} = f'(K) - r$$

whose left hand side is the proportional rate of consumption growth multiplied by: (i) the elasticity of marginal utility;
or (ii) the elasticity of what macroeconomists call
"an intertemporal marginal rate of substitution";
or (iii) by analogy with the theories of risk and inequality aversion, the degree of relative fluctuation aversion.

Final Recommendation

Morton I. Kamien and Nancy L. Schwartz (2012)
Dynamic Optimization, Second Edition:
The Calculus of Variations and Optimal Control
in Economics and Management (Dover Publications)