Advanced Mathematical Economics

Paulo B. Brito
PhD in Economics: 2021-2022
ISEG
Universidade de Lisboa
pbrito@iseg.ulisboa.pt

Lecture 4
3.11.2021

Contents

Chapter 6

Introduction to functional calculus

6.1 Introduction

In several ares in economics functionals over functions $y : \mathrm{X} \subseteq \mathbb{R} \to \mathbb{R}$ of the following types

$$\mathsf{F}[y] \;= \int_{\mathrm{X}} f(x, y(x))\, dx \tag{6.1}$$

and

$$\mathsf{F}[y] \;= g\Big(\int_{\mathrm{X}} f(x, y(x))\, dx \Big) \tag{6.2}$$

where g are both functions with appropriate properties. Those functional are infinite-dimensional analogs to functionals over $y : \mathbb{N} \to \mathbb{R}$

$$F(y) \;= \sum_{i=1}^{n} f(y_i, i) \text{ or } F(y) = g\Big(\sum_{i=1}^{n} f(y_i, i) \Big).$$

As economic problems involve optimizing functionals (**??**) constrained or not by side conditions which could also be functionals of the same types, we need to use use results from generalized calculus, or, more generally calculus an optimization over normed vector spaces.

In addition, we will see in the next chapter that the simplest calculus of variation problem is

$$\max_{y(\cdot)} \mathsf{F}[y] \;= \int_{\mathrm{X}} f(x, y(x), y'(x))\, dx,$$

whose solution involves generalized calculus, in analogy with the classical optimization problem in which the necessary conditions consist in setting a derivative to zero.

Next we present the main results from generalized calculus and use them to solve some infinite-dimensional optimization problems.

6.2 Some results in generalized calculus

6.2.1 Functionals and Gâteaux differential

Consider two normed vector spaces: can be spaces of numbers, continuous functions, continuous differentiable functions, bounded functions or distributions together with a norm

An operator as a mapping between two normed vector spaces.

A **functional** is a mapping between a normed vector space and the space of real numbers.

Let y belong to the space of functions $\mathcal{Y}$ which map $\mathrm{X} \subseteq \mathbb{R} \mapsto \mathbb{R}$.

Consider in particular a functional over the space of functions $y : \mathrm{X} \subseteq \mathbb{R} \to \mathbb{R}$. The function is continuous at point x_0 if $\lim_{x\to x_0} y(x) = y(x_0)$ for x in any neighborhood of x_0. In particular, this means that $y(x_0^+) = y(x_0^-) = y(x_0)$ where $y(x_0^-) = \lim_{x\uparrow x_0}$ and $y(x_0^+) = \lim_{x\downarrow x_0}$, that the limits to the right and to the left are equal. The function is continuous if it is continuous to all points $x \in \mathrm{X}$.

The variation of the functional is denoted as

$$\Delta \mathsf{F}[y] \ = \mathsf{F}[y + dy] - \mathsf{F}[y]$$

In particular, the variation of the functional in the direction $h(x) \in \mathcal{Y}$ is

$$D\,\mathsf{F}[y] \ = \mathsf{F}[y + \epsilon h] - \mathsf{F}[y].$$

The Gâteaux differential is defined as the variation of the functional in the direction $h \in \mathcal{Y}$ when the constant ε is infinitesimal

$$D_{\eta(\cdot)}\mathsf{F}[y] \ = \lim_{\varepsilon\to 0} \frac{\mathsf{F}[y + \epsilon\,\eta] - \mathsf{F}[y]}{\varepsilon}. \tag{6.3}$$

Equivalently we have [1]

$$F[y + \varepsilon\,\eta] \ = F[y] \ + D_{\eta(\cdot)}\mathsf{F}[y]\varepsilon + o(\varepsilon)$$

In analogy with the case of finite-dimensional analisys[2] it can be shown that the Gâteaux differential is a linear functional of $\eta(x)$. If y is a linear operator in a normed vector space, it is bounded and we can invoke the Riesz-Frechet theorem (Riesz and Sz.-Nagy, 1955, p. 61) which states that if $\mathsf{G}[\eta]$ is a bounded operator in a inner-product space, then there is a ζ such that $\mathsf{G}[\eta]$ can be

[1]To see this observe that equation (6.3) is equivalent to

$$\lim_{\varepsilon\to 0} \frac{\mathsf{F}[y + \epsilon\,\eta] - \mathsf{F}[y]}{\varepsilon} - D_{\eta(\cdot)}\mathsf{F}[y] = 0$$

and therefore, to

$$\lim_{\varepsilon\to 0} \frac{\mathsf{F}[y + \epsilon\,\eta] - \mathsf{F}[y] - D_{\eta(\cdot)}\mathsf{F}[y]\,\varepsilon}{\varepsilon} = 0.$$

Writing the numerator as $g(\varepsilon)$ which is possible because a functional has its range in set $\mathbb{R}$, we can use the the "little-o" notation, such that $o(\varepsilon) = \lim_{\varepsilon\to 0} \frac{g(\varepsilon)}{\varepsilon} \ = 0$.

[2]This generalizes the concept of directional derivative in elementary calculus. Let $f(\mathbf{y}) = f(y_1, \dots, y_n)$ the directional derivative in the direction given by the vector $\mathbf{h} \ = (h_1, \dots, h_n)^\top$ is

$$D_h f(\mathbf{y}) \equiv \lim_{\varepsilon\to 0} \ \frac{f(\mathbf{y} + \varepsilon\mathbf{h}) - f(\mathbf{y})}{\varepsilon} = \sum_i \ \frac{\partial f(\mathbf{y})}{\partial y_i}\, h_i$$

is a linear functional of $\mathbf{h} \ = (h_1, \dots, h_n)^\top$.

represented by $\mathsf{G}[\eta] =< \zeta, \eta >$ and $||\mathsf{G}|| = ||\zeta||$. This means that the Gâteux differential of a linear functional admits the representation

$$D_{\eta(\cdot)}\mathsf{F}[y] \;= \int_{\mathrm{X}} d(x)\,\eta(x)\,dx$$

If we consider $\mathcal{Y}$ as a space of distributions [3] we call **Gâteaux derivative** at point x, to the Gâteaux differential in which the perturbation is a Dirac-δ generalized function at a singular at the point x,

$$d(x) = \frac{\delta \mathsf{F}[y]}{\delta y(x)} = \int_{\mathrm{X}} d(s)\,\delta(s-x)\,ds.$$

Therefore, we can represent the Gâteaux diferential a linear functional as regards any perturbation $\eta(\cdot)$ by

$$D_{\eta(\cdot)}\mathsf{F}[y] = \int_{\mathrm{X}} \frac{\delta \mathsf{F}[y]}{\delta y(x)}\,\eta(x)\,dx. \tag{6.4}$$

We can extend the definition to second order variations

$$\Delta^2\mathsf{F}[y] \;= \Delta\,\Big(\Delta\mathsf{F}[y+d_1y] - \Delta\mathsf{F}[y]\Big) = \mathsf{F}[y+d_1y+d_2y] - \mathsf{F}[y+d_2y] - \mathsf{F}[y+d_1y] + \mathsf{F}[y].$$

The second-order Gâteaux differential of funcional $\mathsf{F}[y]$ is defined as

$$D_{\eta_1(\cdot),\eta_2(\cdot)}\mathsf{F}[y] = \lim_{\varepsilon_2\to 0}\;\lim_{\varepsilon_1\to 0}\frac{\mathsf{F}[y+\varepsilon_1\,\eta_1+\varepsilon_2\,\eta_2] - \mathsf{F}[y+\varepsilon_2\,\eta_2] - \mathsf{F}[y+\varepsilon_1\,\eta 1] - \mathsf{F}[y]}{\varepsilon_1\,\varepsilon_2}.$$

Using our previous results we can write

$$D_{\eta_1(\cdot),\eta_2(\cdot)}\mathsf{F}[y] = \int_{\mathrm{X}}\int_{\mathrm{X}} \frac{\delta^2\mathsf{F}[y]}{\delta y(x)\,\delta y(x')}\,\eta_1(x)\,\eta_2(x')\,dx\,dx'.$$

and the second-order Gâteux derivative, associated to perturbarions $\eta(x)$ and $\zeta(x)$ [4]is

$$\frac{\delta^2\mathsf{F}[y]}{\delta y(x)\,\delta y(x')}.$$

We conjecture that there is symmetry

$$D_{\eta_1(\cdot),\eta_2(\cdot)}\mathsf{F}[y] = D_{\eta_2(\cdot),\eta_1(\cdot)}\mathsf{F}[y]$$

Therefore a generalization of the second order Taylor expansion is

$$\mathsf{F}[y+\varepsilon\,\eta] \;= \mathsf{F}[y] + D_{\eta(\cdot)}\mathsf{F}[y]\,\varepsilon + \frac{1}{2}\,D^2_{\eta(\cdot)}\mathsf{F}[y]\,\varepsilon^2 + o(\varepsilon^2), \tag{6.5}$$

where $D^2_{\eta(\cdot)}\mathsf{F}[y] \equiv D_{\eta(\cdot),\eta(\cdot)}\mathsf{F}[y]$.

There are general properties of functionals over one-dimensional functions of one variable, that is $y \in \mathcal{Y}$, $y : \mathrm{X} \subseteq \mathbb{R} \to \mathbb{R}$, where $\mathsf{F} : \mathcal{Y} \to \mathbb{R}$.

[3]This is the analog to a partial derivative in classical calculus. This is a type of possibly ad-hoc concept which is used in mathematical physics and which is useful for our purposes. In particular it allows to determine elasticities of substitution in a continuum setting. References...

[4]We do not use the notation because $\eta(x)$ is a function of x, and reserve it to the derivative of $\eta(x)$.

1. multiplication by a constant: let $a \in \mathbb{R}$ be a number, then $D_{\eta(\cdot)}\{\, a\,\mathsf{F}[y]\} \;= aD_{\eta(\cdot)}\mathsf{F}[y]$
2. sum of functionals: let $\mathsf{F}_1[y]$ and $\mathsf{F}_1[y]$ be two functionals, then $D_{\eta(\cdot)}\{\, \mathsf{F}_1[y] + \mathsf{F}_2[y]\} \;= D_{\eta(\cdot)}\mathsf{F}_1[y] + D_{\eta(\cdot)}\mathsf{F}_2[y]$
3. product rule: let $\mathsf{F}_1[y]$ and $\mathsf{F}_1[y]$ be two functionals, then $D_{\eta(\cdot)}\{\, \mathsf{F}_1[y]\cdot \mathsf{F}_2[y]\} \;= \{D_{\eta(\cdot)}\mathsf{F}_1[y]\} \cdot \mathsf{F}_2[y] + \{D_{\eta(\cdot)}\mathsf{F}_2[y]\} \;\cdot \mathsf{F}_1[y]$
4. chain rule: let $f : \mathbb{R} \to \mathbb{R}$ be a monotonic function and $\mathsf{F}_1[y]$ be a functional, then $D_{\eta(\cdot)}\{f(\mathsf{F}_1[y])\Big\} \;= f'(\mathsf{F}_1[y])\,\{\, D_{\eta(\cdot)}\mathsf{F}[y]\}$

6.2.2 Linear functionals

From now on we will consider the following two types of functionals which are common in economics:

$$\mathsf{F}[y] \;= \int_{\mathrm{X}} f(y(x))\, dx, \tag{6.6}$$

where we assume that function $f(\cdot)$ is a smooth function, i.e. $f \in C^2(\mathbb{R})$, and the integral exists, and

$$\mathsf{G}[y] \;= g(\mathsf{F}[y]) \equiv g\Big(\int_{\mathrm{X}} f(y(x))\, dx\Big) \tag{6.7}$$

where $g(\cdot)$ is also a $C^2(\mathbb{R})$ function.

From now on we call generalized differential to the differential (in the Gâteaux sense) of functional (**??**) to

$$D_{\eta(\cdot)}\mathsf{F}[y] \;= \int_{\mathrm{X}} \frac{\delta \mathsf{F}[y]}{\delta y(x)}\, \eta(x)\, dx \tag{6.8}$$

where

$$\frac{\delta \mathsf{F}[y]}{\partial y(x)} = \frac{\partial f(y(x))}{\partial y}, \text{for each } \; x \in \mathrm{X},$$

is the generalized derivative of functional (6.6).

The generalized differential of functional (6.7) is

$$D_{\eta(\cdot)}\mathsf{G}[y] \;= \int_{\mathrm{X}} \frac{\delta \mathsf{G}[y]}{\delta y(x)}\, \eta(x)\, dx \tag{6.9}$$

where

$$\frac{\delta \mathsf{G}[y]}{\partial y(x)} = g'(\mathsf{F}[y])\, \frac{\partial f(y(x))}{\partial y}, \text{for each } \; x \in \mathrm{X}.$$

The second-order generalized differential of (6.6) is

$$D^2_{\eta(\cdot)}\mathsf{F}[y] \;= \int_{\mathrm{X}} \frac{\delta^2 \mathsf{F}[y]}{\delta y(x)^2}\, \eta(x)^2\, dx \tag{6.10}$$

where the second-order generalized derivative is

$$\frac{\delta^2 \mathsf{F}[y]}{\partial y(x)^2} = \frac{\partial^2 f(y)}{\partial y^2}, \text{for each } \; x \in \mathrm{X}.$$

The second-order generalized differential of (6.7) is

$$D^2_{\eta(\cdot)}\mathsf{G}[y] \;=\int_{\mathrm{X}} \frac{\delta^2 \mathsf{G}[y]}{\delta y(x)^2}\,\eta(x)^2\,dx \tag{6.11}$$

where the second-order generalized derivative is

$$\frac{\delta^2 \mathsf{G}[y]}{\partial y(x)^2} = g''(\mathsf{F}[y])\left(\frac{\partial f(y(x))}{\partial y}\right)^2 + g'(\mathsf{F}[y])\,\frac{\partial^2 f(y(x))}{\partial y^2}, \text{for each } x \in \mathrm{X}., \text{for each } x \in \mathrm{X}.$$

Examples: Next we present the generalized derivatives for several functionals:

1. for $\mathsf{F}[y] = \int_{\mathrm{X}} a\,dx$, we have $\dfrac{\delta \mathsf{F}[y]}{\partial y(x)} = \dfrac{\delta^2 \mathsf{F}[y]}{\partial y(x)^2} = 0$;
2. for $F[y] = \int_{\mathrm{X}} a\,y(x)\,dx$, we have $\dfrac{\delta \mathsf{F}[y]}{\partial y(x)} = a$ and $\dfrac{\delta^2 \mathsf{F}[y]}{\partial y(x)^2} = 0$;
3. for $F[y] = \int_{\mathrm{X}} y(x)^2\,dx$, we have $\frac{\delta \mathsf{F}[y]}{\partial y(x)} = 2y(x)$ and $\dfrac{\delta^2 \mathsf{F}[y]}{\partial y(x)^2} = 2$;
4. for $F[y] = \int_{\mathrm{X}} e^{y(x)}\,dx$, we have $\frac{\delta \mathsf{F}[y]}{\partial y(x)} = \dfrac{\delta^2 \mathsf{F}[y]}{\partial y(x)^2} = e^{y(x)}$;
5. for $F[y] = \int_{\mathrm{X}} y(x)^{\theta}\,dx$, we have $\frac{\delta \mathsf{F}[y]}{\partial y(x)} = \theta\,y(x)^{\theta-1}$ and $\dfrac{\delta^2 \mathsf{F}[y]}{\partial y(x)^2} = \theta\,(\theta-1)\,y(x)^{\theta-2}$;
6. and for $F[y] = \int_{\mathrm{X}} e^{g(y(x))}\,dx$, we have $\frac{\delta \mathsf{F}[y]}{\partial y(x)} = g'(y(x))\,e^{g(y(x))}$ and $\dfrac{\delta^2 \mathsf{F}[y]}{\partial y(x)^2} = g''(y(x))\,e^{g(y(x))} + (g'(y(x)))^2\,e^{g(y(x))}$.

"Spike" perturbations

In particular, for a "spike" perturbation at point $x = x_0$, represented by a delta [5] we have a functional generalization to the partial derivative[6]

$$D_{\delta(x_0)}\mathsf{F}[y] = \int_{\mathrm{X}} \frac{\delta \mathsf{F}[y]}{\delta y(x)}\,\delta(x-x_0)\,dx = \frac{\delta \mathsf{F}[y]}{\delta y(x_0)} = f'(y(x_0)). \tag{6.12}$$

6.2.3 Functionals over two-dimensional spaces

We can In this section definitions to functionals over two-dimensional functions of one variable, that is $\mathbf{y} \in \mathcal{Y}$, $\mathbf{y} : \mathrm{X} \subseteq \mathbb{R} \to \mathbb{R}^2$, where $\mathsf{F} : \mathcal{Y} \to \mathbb{R}$, where $\mathbf{y}(x) = (y_1(x), y_2(x))^\top$.

[5] Dirac-δ is not a function but a distribution. It has the following properties

$$\delta(x-x_0) = \begin{cases} 0 & \text{if } x \neq x_0 \\ \infty & \text{if } x = x_0 \end{cases},$$

$\int_{-\infty}^{\infty} \delta(x)dx = 1$ and $\delta_{-\infty}^{\infty}\,\delta(x-y_0)\,y(x)\,dx = y(x_0)$.

[6] In mathematical physics this is sometimes called as a Volterra derivative.

In particular, let

$$\mathsf{F}[\mathbf{y}] \;= \int_{\mathrm{X}} f(x, y_1(x), y_2(x))\,dx. \tag{6.13}$$

In this case, for a perturbation $\boldsymbol{\eta}(x) = (\eta_1(x), \eta_2(x))^\top$, we definition

$$D_{\boldsymbol{\eta}(\cdot)}\mathsf{F}[\mathbf{y}] \;= \lim_{\varepsilon\to 0} \frac{\mathsf{F}[\mathbf{y}+\varepsilon\,\boldsymbol{\eta}] - \mathsf{F}[y]}{\varepsilon},$$

which implies that the functional (6.13) has the differential

$$\begin{aligned} D_{\boldsymbol{\eta}(\cdot)}\mathsf{F}[\mathbf{y}] &= \int_{\mathrm{X}} \Big\langle \frac{\delta \mathsf{F}[\mathbf{y}]}{\delta \mathbf{y}(x)}, \boldsymbol{\eta}\Big\rangle (x)\,dx \\ &= \int_{\mathrm{X}} \Big(\frac{\delta \mathsf{F}[\mathbf{y}]}{\delta y_1(x)}\,\eta_1(x) + \frac{\delta \mathsf{F}[\mathbf{y}]}{\delta y_2(x)}\,\eta_2(x)\Big)\,dx, \end{aligned} \tag{6.14}$$

where the generaiized gradient is

$$\frac{\delta \mathsf{F}[\mathbf{y}]}{\delta \mathbf{y}(x)} = \begin{pmatrix} \dfrac{\delta \mathsf{F}[\mathbf{y}]}{\delta y_1(x)} \\ \dfrac{\delta \mathsf{F}[\mathbf{y}]}{\delta y_2(x)} \end{pmatrix}.$$

Example:
The functional $\mathsf{F}[\mathbf{y}] = \int_{\mathrm{X}} y_1(x)^\alpha\, y_2(x)^{1-\alpha}\;dx$, for a number α, then the generalized gradient is

$$\frac{\delta \mathsf{F}[\mathbf{y}]}{\delta \mathbf{y}(x)} = \begin{pmatrix} \alpha y_1(x)^{\alpha-1}\, y_2(x)^{1-\alpha} \\ (1-\alpha) y_1(x)^{\alpha}\, y_2(x)^{-\alpha} \end{pmatrix}.$$

6.2.4 Functionals involving first-order derivatives

The previous result is useful to address the derivation of the next functional which is used in calculus of variations

$$\mathsf{F}[y] \;= \int_{\mathrm{X}} f(x, y(x), y'(x))\,dx, \tag{6.15}$$

where we assume function $f(\cdot)$ is continuous and continuous differentiable in (y, y'). The (first-order) generalized differential is

$$D_{\eta(\cdot)}\mathsf{F}[y] \;= \int_{\mathrm{X}} \frac{\delta \mathsf{F}[y]}{\delta y(x)}\;\eta(x)\,dx + \int_{\mathrm{X}} \frac{\partial f}{\partial y'}(x)\,\eta(x) \tag{6.16}$$

where the generalized derivative is

$$\frac{\delta \mathsf{F}[y]}{\delta y(x)} = \frac{\partial f(x, y(x), y'(x))}{\partial y} - \frac{d}{dx}\left(\frac{\partial f(x, y(x), y'(x))}{\partial y'(x)}\right) \tag{6.17}$$

and

$$\int_{\mathrm{X}} \frac{\partial f(x)}{\partial y'}\,\eta(x) = \frac{\partial f(x_1, y(x_1), y'(x_1))}{\partial y'}\;\eta(x_1) - \frac{\partial f(x_0, y(x_0), y'(x_0))}{\partial y'}\,\eta(x_0)$$

if $\mathrm{X} = [x_0, x_1]$ is a closed set, or

$$\int_{\mathrm{X}} f(x)\,\eta(x) = \lim_{x \uparrow x_1} \frac{\partial f(x)}{\partial y'}\,\eta(x) - \lim_{x \downarrow x_0} \frac{\partial f(x)}{\partial y'}\,\eta(x)$$

if $\mathrm{X} = (x_0, x_1)$ is an open set.

To prove this, observe that our perturbations are $y(x) \to y(x) + \varepsilon\eta(x)$ and $y'(x) \to y'(x) + \varepsilon\eta'(x)$, which using equation (6.14) yields

$$\begin{aligned}
D_{\eta(\cdot)}\mathsf{F}[y] &= \int_{\mathrm{X}} \Big(\frac{\delta \mathsf{F}[y]}{\delta y(x)}\;\eta(x) + \frac{\delta \mathsf{F}[y]}{\delta y'(x)}\;\eta'(x) \Big)\,dx \\
&= \int_{\mathrm{X}} \frac{\delta \mathsf{F}[y]}{\delta y(x)}\;\eta(x)\,dx + \int_{\mathrm{X}} \frac{\delta \mathsf{F}[y]}{\delta y'(x)}\;\eta'(x)\,dx \\
&= \int_{\mathrm{X}} \frac{\partial f(x, y(x), y'(x))}{\partial y}\;\eta(x)\,dx + \int_{\mathrm{X}} \frac{\partial f(x, y(x), y'(x))}{\partial y'}\;\eta'(x)\,dx
\end{aligned}$$

Integrating by parts the second integral, we obtain

$$\int_{\mathrm{X}} \frac{\partial f(x, y(x), y'(x))}{\partial y'}\;\eta'(x)\,dx = \int_{\mathrm{X}} \frac{\partial f(x, y(x), y'(x))}{\partial y'}\,\eta(x) - \int_{\mathrm{X}} \frac{d}{dx}\,\Big(\frac{\partial f(x, y(x), y'(x))}{\partial y'}\Big)\,\eta(x)\,dx,$$

that, upon substitution, yields the differential (6.16).

6.2.5 Functionals involving second-order derivatives

Now consider the integral where function f depends on the first and second derivatives of y,

$$\mathsf{F}[y] \;= \int_{\mathrm{X}} f(x, y(x), y'(x), y''(x))\,dx. \tag{6.18}$$

where we assume function $f(\cdot)$ is continuous and continuous differentiable at least up to the second order in (y, y', y'').

The Gâteaux differential of functional (6.18) is

$$D_{\eta(\cdot)}\mathsf{F}[y] \;= \int_{\mathrm{X}} \frac{\delta \mathsf{F}[y]}{\delta y(x)}\;\eta(x)\,dx + \int_{\mathrm{X}} \frac{\partial f(x)}{\partial y'}\,\eta(x) + \int_{\mathrm{X}} \frac{\partial f(x)}{\partial y''}\,\eta'(x) - \int_{\mathrm{X}} \frac{d}{dx}\Big(\frac{\partial f(x)}{\partial y'}\Big)\,\eta(x) \tag{6.19}$$

where we denote $f(x) = f(x, y(x), y'(x), y''(x))$, the Gâteaux derivative is

$$\frac{\delta \mathsf{F}[y]}{\delta y(x)} = \frac{\partial f(x)}{\partial y} - \frac{d}{dx}\,\Big(\frac{\partial f(x)}{\partial y'}\Big) + \frac{d^2}{dx^2}\,\Big(\frac{\partial f(x)}{\partial y''}\Big), \tag{6.20}$$

and, in the case in which set X is closed

$$\begin{aligned}
&\int_{\mathrm{X}} \frac{\partial f(x)}{\partial y'}\,\eta(x) = \frac{\partial f(x_1)}{\partial y'}\,\eta(x_1) - \frac{\partial f(x_0)}{\partial y'}\,\eta(x_0) \\
&\int_{\mathrm{X}} \frac{\partial f(x)}{\partial y''}\,\eta'(x) = \frac{\partial f(x_1)}{\partial y''}\;\eta'(x_1) - \frac{\partial f(x_0)}{\partial y''}\;\eta'(x_0) \\
&\int_{\mathrm{X}} \frac{d}{dx}\Big(\frac{\partial f(x)}{\partial y''}\Big)\,\eta(x) = \frac{d}{dx}\Big(\frac{\partial f(x_1)}{\partial y''}\Big)\,\eta(x_1) - \frac{d}{dx}\Big(\frac{\partial f(x_0)}{\partial y''}\Big)\,\eta(x_0).
\end{aligned}$$

To prove this, we have

$$\begin{aligned} D_{\eta(\cdot)}\mathsf{F}[y] &= \int_{\mathrm{X}} \Big(\frac{\delta \mathsf{F}[y]}{\delta y(x)}\ \eta(x) + \frac{\delta \mathsf{F}[y]}{\delta y'(x)}\ \eta'(x) + \frac{\delta \mathsf{F}[y]}{\delta y''(x)}\ \eta''(x) \Big)\, dx \\ &= \int_{\mathrm{X}} \frac{\delta \mathsf{F}[y]}{\delta y(x)}\ \eta(x)\, dx + \int_{\mathrm{X}} \frac{\delta \mathsf{F}[y]}{\delta y'(x)}\ \eta'(x)\, dx + \int_{\mathrm{X}} \frac{\delta \mathsf{F}[y]}{\delta y''(x)}\ \eta''(x)\, dx \\ &= \int_{\mathrm{X}} \frac{\partial f(x)}{\partial y}\ \eta(x)\, dx + \int_{\mathrm{X}} \frac{\partial f(x)}{\partial y'}\ \eta'(x)\, dx + \int_{\mathrm{X}} \frac{\partial f(x)}{\partial y''}\ \eta''(x)\, dx. \end{aligned}$$

Integrating by parts the second integral, we have

$$\int_{\mathrm{X}} \frac{\partial f(x)}{\partial y'}\ \eta'(x)\, dx = \int_{\mathrm{X}} \frac{\partial f(x)}{\partial y'}\, \eta(x) - \int_{\mathrm{X}} \frac{d}{dx}\ \Big(\frac{\partial f(x)}{\partial y'}\Big)\, \eta(x)\, dx.$$

Integrating by parts the third integral, [7] we have

$$\begin{aligned} \int_{\mathrm{X}} \frac{\partial f(x)}{\partial y''}\ \eta''(x)\, dx = & \int_{\mathrm{X}} \frac{\partial f(x)}{\partial y''}\, \eta'(x) - \int_{\mathrm{X}} \frac{d}{dx}\ \frac{\partial f(x)}{\partial y''}\, \eta'(x)\, dx \\ & + \int_{\mathrm{X}} \frac{d^2}{dx^2}\ \frac{\partial f(x)}{\partial y''}\, \eta(x)\, dx. \end{aligned}$$

6.2.6 Functionals over higher-dimensional independent variables

We can extend those definitions for functionals $\mathsf{F}[y]$ over functions over higher dimensional $y \in \mathcal{Y}$ which is the space of functions $\mathbf{x} \subseteq \mathbb{R}^n \mapsto \mathbb{R}$. The simplest case involves functionals over functions $y(\mathbf{x}) = y(x_1, x_2)$.

The simplest functional is

$$\mathsf{F}[y] = \int_{\mathrm{X}} f(\mathbf{x}, y(\mathbf{x}))\, d\mathbf{x} = \int_{\mathrm{X}_1} \int_{\mathrm{X}_2} f(x_1, x_2, y(x_1, x_2))\, dx_2\, dx_1 \tag{6.21}$$

where $f(\cdot, y)$ is continuously diffferentiable in y and $\mathrm{X} = \mathrm{X}_1 \times \mathrm{X}_2 \subseteq \mathbb{R}^2$. Introducing the perturbation $y(\mathbf{x}) \to y(\mathbf{x}) + \varepsilon\, \eta(\mathbf{x})$ is

The Gâteaux differential is

$$D_{\eta(\cdot)}\ \mathsf{F}[y] = \int_{\mathrm{X}} \frac{\delta \mathsf{F}[y]}{\delta y(\mathbf{x})}\ \eta(\mathbf{x})\, d\mathbf{x} = \int_{\mathrm{X}_1} \int_{\mathrm{X}_2} \frac{\delta \mathsf{F}[y]}{\delta y(x_1, x_2)}\, \eta(x_1, x_2)\, dx_2\, dx_1 \tag{6.22}$$

where the Gâteaux derivative is

$$\frac{\delta \mathsf{F}[y]}{\delta y(\mathbf{x})} = \frac{\partial f(\mathbf{x}, y(\mathbf{x}))}{\partial y} \tag{6.23}$$

For the functional with first derivatives,

$$\begin{aligned} \mathsf{F}[y] = \int_{\mathrm{X}} f(\mathbf{x}, y(\mathbf{x}), \nabla y(\mathbf{x}))\, d\mathbf{x} = \\ \int_{\mathrm{X}_1} \int_{\mathrm{X}_2} f(x_1, x_2, y(x_1, x_2), y_{x_1}(x_1, x_2), y_{x_2}(x_1, x_2))\, dx_2\, dx_1 \end{aligned} \tag{6.24}$$

---

[7] This is because

$$\int u v''\, dx = u v' - \int u' v' dx = u v' - \Big(u' v - \int u'' v dx \Big).$$

where $f(\cdot, y)$ is continuously diffferentiable in $(y, \nabla y)$, $\mathrm{X} = \mathrm{X}_1 \times \mathrm{X}_2 \subseteq \mathbb{R}^2$ and $y_{x_i}(\mathbf{x})$ denote the partial derivatives for $i = 1, 2$. Introducing, the perturbations $y(\mathbf{x}) \to y(\mathbf{x}) + \varepsilon\, \eta(\mathbf{x})$ and $\nabla y(\mathbf{x}) \to \nabla y(\mathbf{x}) + \varepsilon\, \nabla \eta(\mathbf{x})$

The Gâteaux differential is

$$D_{\eta(\cdot)}\ \mathsf{F}[y] = \int_{\mathrm{X}} \frac{\delta \mathsf{F}[y]}{\delta y(\mathbf{x})}\ \eta(\mathbf{x})\, d\mathbf{x} = \int_{\mathrm{X}_1} \int_{\mathrm{X}_2} \frac{\delta \mathsf{F}[y]}{\delta y(x_1, x_2)}\, \eta(x_1, x_2)\, dx_2\, dx_1 \tag{6.25}$$

where the Gâteaux derivative is

$$\frac{\delta \mathsf{F}[y]}{\delta y(x_1, x_2)} = \frac{\partial f(\mathbf{x}, y(\mathbf{x}))}{\partial y} \tag{6.26}$$

6.3 Applications

Next we present some infinite-dimensional functionals which are relatively common in economic models.

6.3.1 Generalized means

The **generalized mean** is very common functional which is used in economics. It is a functional on $x : \mathrm{I} \to \mathbb{R}$, where $\mathrm{I} \subset \mathbb{R}_+$,

$$\mathsf{M}_\rho[x] = \left(\int_{\mathrm{I}} w(i)\, x(i)^\rho\, di \right)^{\frac{1}{\rho}},\ \text{for } \rho \in [-\infty, \infty] \tag{6.27}$$

where $w : \mathrm{I} \to (0, 1)$ is a weighting function such that

$$\int_{\mathrm{I}} w(i)\, di = 1.$$

This functional can be seen as an infinite-dimensional generalization of the CES aggregator.

Proposition 1. *The geometrical mean has the special cases:*

1. *if* $\rho = 0$ *then* $\mathsf{M}_0[x] = \exp\left(\int_{\mathrm{I}} \ln\left(x(i)^{w(i)}\ \right) di \right)$ *(geometric mean, in particular a generalized Cobb-Douglas function)*
2. *if* $\rho = 1$ *then* $\mathsf{M}_1[x] = \int_{\mathrm{I}} w(i)\, x(i)\, di$ *(arithmetic mean)*
3. *if* $\rho = -1$ *then* $\mathsf{M}_{-1}[x] = \dfrac{1}{\int_{\mathrm{I}} \frac{w(i)}{x(i)}\, di}$ *(harmonic mean)*
4. *if* $\rho = -\infty$ *then* $\mathsf{M}_{-\infty}[x] = \min\{x\}$ *(in particular a generalized Leontieff production function)*
5. *if* $\rho = \infty$ *then* $\mathsf{M}_\infty[x] = \max\{\ x\}$

In general $\min_x[x] \ \leq \mathsf{M}_\rho\ [x]\ \leq\ \max_x[x]$

Proof. We can write equation (6.27) equivalently as

$$M_\rho[x] = \exp\left\{ \ln\left[\left(\int_{\mathrm{I}} w(i)\, x(i)^\rho\, di \right)^{\frac{1}{\rho}} \right] \right\} = \exp\left\{ \ \frac{1}{\rho}\ \ \ln\left(\int_{\mathrm{I}} w(i)\, x(i)^\rho\, di \ \right) \right\}.$$

But

$$\begin{aligned} \lim_{\rho\to 0}\ \frac{\ln\left(\int_{\mathrm{I}} w(i)\, x(i)^\rho\, di \right)}{\rho} &= \lim_{\rho\to 0}\ \frac{\int_{\mathrm{I}} w(i)\ \ln\left(x(i)\right] x(i)^\rho\, di}{1} \\ &= \int_{\mathrm{I}} w(i)\ \ln\left(x(i)\right) di = \int_{\mathrm{I}} \ln\left(x(i)^{w(i)}\right) di \end{aligned}$$

Then $M_0[x] = \exp\left(\int_{\mathrm{I}} \ln\left(x(i)^{w(i)}\right) \ \right) di$. The generalized mean for $\rho = 1$ and $\rho = -1$ are obtained by direct substitution.

If function $x : \mathbb{R} \to\ \mathrm{X}$ and $x \in L_\infty(\mathrm{X})$ then there is a maximum $x^* = ||x||_\infty$ and we can write

$$M_\infty[x] = \lim_{\rho\to\infty} x^* \left(\int_{\mathrm{I}} w(i) \left(\frac{x(i)}{x^*}\right)^\rho di \ \right)^{\frac{1}{\rho}} = \lim_{\psi\to 0} x^* \left(\int_{\mathrm{I}} w(i) \left(\frac{x(i)}{x^*}\right)^{\frac{1}{\psi}} di \ \right)^{\psi} = x^*$$

where $\psi = 1/\rho$. For any ρ we have $M_\rho[x] = M_{-\rho}\left[\frac{1}{x}\right]$. If $x \in L_\infty(\mathrm{X})$ as $\max\left[\frac{1}{x}\right]\ = \min[x]$ then

$$M_{-\infty}[x] = \lim_{\rho\to\infty} \frac{1}{M_\rho\left[\frac{1}{x}\right]} = \min[x]$$

□

The generalized derivative of equation (6.27), for a finite ρ, is

$$\frac{\delta \mathsf{M}_\rho[x]}{\delta x(i)} = w(i) \left(\frac{\mathsf{M}_\rho[x]}{x(i)}\right)^{1-\rho}, \text{ for each }\ i \in \mathrm{I}.$$

6.3.2 Varieties in consumption theory

Let $c(i)$ be the quantity of good of variety $i \in \mathrm{I} \subseteq [0, \infty)$. We denote by $c(\iota)$ the quantity of a specific variety $i = \iota \in \mathrm{I}$. We can see $c(\iota) = \int_{\mathrm{I}} \delta(i - j\iota)\, c(i) di$. We denote C a variety distribution $C = (c(i))_{i\in\mathrm{I}}$. The set of variety distributions C is the function space $\mathcal{C}$.

A utility functional $\mathsf{U}[c]$ maps a variety distribution into value, which is a number, i.e. $\mathsf{U} : \mathcal{C} \to \mathbb{R}$.

The most common utility functional in the literature is the constant elasticity of substitution utility functional

$$\mathsf{U}[c] = \left(\int_{\mathrm{I}} c(i)^{1-\gamma} di \right)^{\frac{1}{1-\gamma}}$$

Using our previous definitions a marginal change the variety distribution by ζ will lead to a change in utility given by Gâteaux differential

$$D_{\zeta(\cdot)}\ \mathsf{U}[c]\ =\int_{\mathrm{I}}\frac{\delta \mathsf{U}[c]}{\delta c(i)}\,\zeta(i)\ di,\ \text{for any}\ \ i\in \mathrm{I}$$

where the marginal utility of variety i is given by the Gâteaux derivative

$$\begin{aligned}\frac{\delta \mathsf{U}[c]}{\delta c(i)} &= \frac{1}{1-\gamma}\left(\int_{\mathrm{I}} c(i)^{1-\gamma} di\right)^{\frac{\gamma}{1-\gamma}}(1-\gamma)c(i)^{-\gamma}\\ &= \mathsf{U}[c]^{\gamma} c(i)^{-\gamma}\end{aligned}$$

For a specific variety $i=\iota$, the marginal utility can be represented by the Gâteaux derivative for a perturbation $\delta(i-\iota)$, that is

$$\frac{\delta \mathsf{U}[c]}{\delta c(\iota)} = \mathsf{U}[c]^{\gamma} c(\iota)^{-\gamma},\ \text{for}\ \ i=\iota\in \mathrm{I},$$

which allows us to determine the marginal rate of substitution between varieties i_1 and i_2 by

$$MRS_{i_0,i_1} = \left(\frac{c(i_0)}{c(i_1)}\right)^{-\gamma}, \text{for any}\ \ i_0, i_1\in \mathrm{I}.$$

6.3.3 Infinite-dimensional production functions

Let $x(i)$ be the quantity of good of varieties $i\in \mathrm{I}\subseteq[0,\infty)$. We denote x_j the quantity of variety $j\in\mathrm{I}$. We can see $x(j)=\int_{\mathrm{I}}\delta(i-j)q(i)di$. We denote x a variety distribution $x=(x(i))_{\mathrm{I}}$. The set of variety distributions belong to the function space $\mathcal{X}$.

A production function is a functional $\mathsf{F}[x]$ which maps the variety distribution into real production, a non-negative number, i.e. $\mathsf{F}:\mathcal{X}\to\mathbb{R}_+$. We can write it as

$$y=\mathsf{F}[x].$$

There are several production functions used in the literature (see, for instance, Parenti et al. (2017) and Bucci and Ushchev (2016)):

- the generalized AK production function

  $$\mathsf{F}[x]\ =\int_{\mathrm{I}} A(i)\,x(i)\,di;$$

  in which $A(i)$ is the specific marginal productivity of input $i\in\mathrm{I}$,

- the generalized constant elasticity of substitution production function

  $$\mathsf{F}[x]\ =\left(\int_{\mathrm{I}} A(i)\,x(i)^{\frac{\epsilon-1}{\epsilon}}\,di\right)^{\frac{\epsilon}{\epsilon-1}}$$

which is the infinite-dimensional analog of

$$F(\{x\}\ ) = \Big(\sum_{i\in \mathrm{I}} A_i\, x_i^{\frac{\epsilon-1}{\epsilon}}\Big)^{\frac{\epsilon}{\epsilon-1}},$$

where $\mathrm{I} = \{i_1, \dots, i_n\}$;

- the O-ring production function (see Kremer (1993))

$$\mathsf{F}[x]\ = \exp\Big(\int_{\mathrm{I}} \gamma(i) \log\left(x(i)\right) di\Big)$$

  where $\int_{\mathrm{I}} \gamma(i) = 1$. This function is the infinite-dimensional of the finite-dimensional analog

$$F(\{x\}\ ) = \prod_{i\in \mathrm{I}} x_i^{\gamma_i}$$

  for $\sum_{\mathrm{I}} \gamma_i = 1$;

- the translog production function

$$F[x]\ = \exp\left(\int_{\mathrm{I}} \gamma(i) \log\left(x(i)\right) + \gamma(i) \log\left(x(i)^2\right) di\right).$$

6.3.4 In information theory and econometrics

In, according to some authors, one of the most influential paper in the XX century Shannon (1948) introduces a measure of information by entropy. Entropy is a the functional

$$H[f]\ = -\int_{\mathrm{X}} f(x) \ln(f(x)) dx,$$

which is used in an information theoretic approach to econometrics (see Judge and Mittelhammer (2012)).

6.4 Problems involving functionals

6.4.1 Extremes of functionals

The function $y^* \in \mathcal{Y}$ is an extreme of functional $\mathsf{F}[y]$ if the value of functional has the value $\mathsf{F}[y^*]$ and any small arbitrary perturbation $\eta \in \mathcal{Y}$, will not deviate the functional from that value. That is locally we have

$$D_{\eta(\cdot)}\ \mathsf{F}[y^*]\ = 0.$$

An extreme $y^* \in \mathcal{Y}$ is a maximum, only if, for an arbitrary perturbation η, such that $y = y^* + \varepsilon\,\eta$ we have $\mathsf{F}[y^*] \geq \mathsf{F}[y]$, and it is a minimum only if $\mathsf{F}[y^*] \leq \mathsf{F}[y]$.

From the generalized Taylor expansion in equation (6.5) a necessary condition of second order for a maximum is

$$D^2_{\eta(\cdot)}\ \mathsf{F}[y^*]\ \leq 0.$$

6.4.2 Maximum of functionals

Consider the problem: find $y \in \mathcal{Y}$ that solves the problem

$$\max_{y(\cdot)}\ \mathsf{F}[y]$$

A necessary condition for a maximum is that the Gâteaux derivative of $F[y]$ satisfies

$$\frac{\delta \mathsf{F}[y^*]}{\delta y(x)} = \frac{\delta \mathsf{F}[y]}{\delta y(x)}\Big|_{y^*(x)} = 0$$

Proof: Assume we know the optimum $y^*(x)$. The objective functional evaluated at the optimum is $\mathsf{F}[y^*]$, which is a number. Introducing an arbitrary perturbation $\eta(x) \in \mathcal{Y}$ at the optimum we obtain $y(x) = y^*(x) + \varepsilon\,\eta(x)$ which has the value $\mathsf{F}[y] = \mathsf{F}[y^* + \varepsilon\,\eta]$. If $y^*(x)$ is an optimum then we should have $\mathsf{F}[y^*] \geq \mathsf{F}[y]$. Expanding $\mathsf{F}[y]$ in a neighborhood of the optimum we have

$$\mathsf{F}[y]\ = \mathsf{F}[y^*] + D_{\eta(\cdot)}\mathsf{F}[y^*]\,\varepsilon + \frac{1}{2}\ D_{\eta(\cdot),\eta(\cdot)}\mathsf{F}[y^*]\,\varepsilon^2 + o(\varepsilon^2)$$

If y^* is a maximum it satisfies $D_{\eta(\cdot),\eta(\cdot)}\mathsf{F}[y^*] \leq 0$. Therefore we can have $\mathsf{F}[y^*] \geq \mathsf{F}[y]$ only if $D_{\eta(\cdot)}\mathsf{F}[y^*] = 0$.

The Taylor expansion also allows us to find a sufficient condition. Assume that: first, the functional is concave in the sense that it satisfies the condition:

$$D^2_{\eta(\cdot)}\ \mathsf{F}[y]\ \leq 0,\ \text{for any}\ \ y \in \mathcal{Y}$$

and, second, there is an element of $\mathcal{Y}$, $y^*(x)$ satisfying the condition $D_{\eta(\cdot)}\mathsf{F}[y^*] = 0$. Then $y^*(x) \in \mathcal{Y}$ is a maximum.

6.4.3 Constrained maximum of functionals

In economics we are generally interested in problems define by the maximization of a functional depending on constraints. We next consider two types of problems which differ as regards the nature of their constraints: problems with functional constraints and problems with local constraints.

Problems with functional constraints

Consider the two functionals over function $y : \mathrm{X} \to \mathbb{R}$ belonging to the set $y \in \mathcal{Y}$ of bounded functions:

$$\mathsf{W}[y] = \int_{\mathrm{X}} F(y(x))dx$$

and

$$\mathsf{T}[y] = \int_{\mathrm{X}} G(y(x))dx$$

The first problem is: find function $y(x) \in \mathcal{Y}$ that solves

$$\begin{aligned} &\max_{y(\cdot)} \int_{\mathrm{X}} F(y(x))\, dx \\ &\text{subject to} \\ &\int_{\mathrm{X}} G(y(x))dx = 0 \end{aligned} \qquad \text{((P1))}$$

We can define a generalized Lagrangean functional

$$\mathsf{L}[y;\lambda] = \mathsf{W}[y] + \lambda\, \mathsf{T}[y]$$

where $\lambda \in \mathbb{R}$ is the Lagrange multiplier. WE can define a Lagrangean (function)

$$L(y(x),\lambda) \equiv F(y(x)) + \lambda G(y(x)).$$

Therefore the Lagrangean functional can be written as a parameterized (by λ) functional

$$\mathsf{L}[y;\lambda] = \int_{\mathrm{X}} L(y(x),\lambda)\, dx.$$

The necessary conditions for an optimum $y^*(x)$ are

$$\begin{aligned} \frac{\delta \mathsf{L}[y;\lambda]}{\delta y(x)}\Big|_{y^*(x)} &= F'(y^*(x)) + \lambda\, G(y^*(x)) = 0 \\ \frac{\partial \mathsf{L}[y;\lambda]}{\partial \lambda}\Big|_{y^*(x)} &= \int_{\mathrm{X}} G(y^*(x))dx = 0, \end{aligned}$$

where the first condition features a Gâteaux derivative and the second a classical derivative.

Problems with local constraints

Now consider the problem

$$\begin{aligned} &\max_{y(\cdot),z(\cdot)} \int_{\mathrm{X}} F(y(x),z(x))\, dx \\ &\text{subject to} \\ &G(y(x),z(x)) = 0, \text{ for all } x \in \mathrm{X}. \end{aligned} \qquad \text{((P2))}$$

While in problem (P1) we had one constraint, in this case we have an infinity of constraints. Therefore, we have to introduce a Lagrangean function $\lambda : \mathrm{X} \to \mathbb{R}$ (instead of a Lagrange multiplier $\lambda \in \mathbb{R}$ a in the previous problem).

The Lagrangean functional is now

$$\mathsf{L}[y, z; \lambda] = \int_{\mathrm{X}} F(y(x), z(x)) + \lambda(x) G(y(x), z(x)) dx$$

The necessary condition for an extremum $(y^*(x), z^*(x))$ are

$$\begin{aligned}
\frac{\delta \mathsf{L}^*}{\delta y(x)} &= F_y\ (y^*(x), z^*(x)) + \lambda(x)\, G_y\ (y^*(x), z^*(x)) = 0, \text{ for each } x \in \mathrm{X} \\
\frac{\delta \mathsf{L}^*}{\delta z(x)} &= F_z\ (y^*(x), z^*(x)) + \lambda(x)\, G_z\ (y^*(x), z^*(x)) = 0, \text{ for each } x \in \mathrm{X} \\
\frac{\delta \mathsf{L}^*}{\delta \lambda(x)} &= G\ (y^*(x), z^*(x)) = 0, \text{ for each } x \in \mathrm{X}.
\end{aligned}$$

6.4.4 Applications

Problem 1 find the continuous maximum entropy distribution with support $[a, b]$. That is, solve

$$\begin{aligned}
\max_{f(\cdot)}\ \mathsf{H}[f] &= -\int_a^b f(x) \ln(f(x))\, dx \text{ (entropy functional)} \\
&\text{subject to} \\
\mathsf{G}[f] &= \int_a^b f(x) dx = 1 \text{ (functional constraint)}
\end{aligned}$$

The Lagrangean functional is

$$\begin{aligned}
\mathsf{L}[f] &= \int_a^b L(f(x), \lambda) dx = \\
&= \int_a^b -f(x) \ln(f(x)) - \lambda f(x)\, dx
\end{aligned}$$

where λ is a Lagrange multiplier. The previously obtained the first-order conditions yield

$$\frac{\delta \mathsf{L}[f^*]}{\delta f(x)} = \frac{\partial L(f(x))}{\partial f} = -\ln(f(x)) - \lambda - 1 = 0$$

if and only if $f(x) = e^{-(1+\lambda)}$; and substituting in the constraint

$$\mathsf{G}[f] = \int_a^b e^{-(1+\lambda)}\, dx = 1$$

if and only if $e^{-(1+\lambda)} = \frac{1}{b-a}$. Therefore the maximum entropy distribution $f^*(x)$

$$f^*(x) = \frac{1}{b-a}$$

is a uniform distribution.

Problem 2[8] find the continuous maximum entropy distribution with support $(-\infty, \infty)$ which such that average satisfies $\mathbb{E}[x] = \mu$ and the variance satisfies $\mathbb{E}[(x-\mu)^2] = \sigma^2$. Recall that

$$\mathbb{E}[x] = \int_{-\infty}^{\infty} x f(x) dx$$
$$\mathbb{V}[x] = \int_{-\infty}^{\infty} (x-\mu)^2 f(x) dx$$

The problem is: find $f(x)$ which solves the problem

$$\max_{f(\cdot)} \mathsf{H}[f] = -\int_a^b f(x) \ln(f(x))\, dx \text{ (entropy functional)}$$

subject to

$$\int_{-\infty}^{\infty} f(x) dx = 1 \text{ (f is a density function)}$$
$$\int_{-\infty}^{\infty} x f(x) dx = \mu \text{ (with average equal to μ)}$$
$$\int_{-\infty}^{\infty} x^2 f(x) dx = \sigma^2 + \mu^2 \text{ (with variance equal to σ^2).}$$

The last restriction is equivalent to $\mathbb{V}[x] = \sigma^2$, because

$$\mathbb{V}[x] = \int_{-\infty}^{\infty} (x-\mu)^2 f(x) dx = \int_{-\infty}^{\infty} x^2 f(x) dx - 2\int_{-\infty}^{\infty} \mu x f(x) dx + \int_{-\infty}^{\infty} \mu^2 f(x) dx = \int_{-\infty}^{\infty} x^2 f(x) dx - \mu^2.$$

The Lagrangean functional is

$$\begin{aligned} \mathsf{L}\ [f] &= \int_a^b L(x, f(x))\, dx = \\ &= \int_a^b -f(x)\ln(f(x)) - \lambda_0 f(x) - \lambda_1 x f(x) - \lambda_2 x^2 f(x) dx \end{aligned}$$

where λ_0, λ_1 and λ_2 are Lagrange multipliers (they are all numbers).

The first order condition is

$$\frac{\delta \mathsf{L}[f^*]}{\delta f(x)} = -\ln(f(x)) - 1 - \lambda_0 - \lambda_1 x - \lambda_2 x^2 = 0$$

Therefore

$$f(x) = e^{-(1+\lambda_0+\lambda_1 x+\lambda_2 x^2)}$$

There are two ways of finding the solution:

[8] See (Shannon, 1948, p.36), or (Cover and Thomas, 2006, ch. 12)

1. First approach: conjecture that the solution is a Gaussian integral

$$g(x) = ae^{-\frac{(x-b)^2}{2c^2}}$$

where a, b, and c are undetermined coefficients, and should be c a real and positive number. If the conjecture on the general form of the solution is correct, then we can obtain the three parameters by substituting this function in the three constraints. Fortunately, this is the case because we obtain a system of three equations in the three unknowns a, b, and c:

$$\int_{-\infty}^{\infty} g(x)dx = a\sqrt{2\pi c}$$
$$\int_{-\infty}^{\infty} xg(x)dx = ba\sqrt{2\pi c}$$
$$\int_{-\infty}^{\infty} x^2 g(x)dx = a\sqrt{2\pi c}\,(b^2 + c)$$

Solving the system yields

$$a = \frac{1}{2\pi\sigma^2},\ b = \mu,\ \text{and}\ \ c = \sigma^2,$$

which implies that function $g(x)$ becomes

$$g(x) = e^{\log(a) - \frac{1}{2c}(x^2 - 2bx + b^2)}.$$

Matching the exponent with $f(x)$, we find the Lagrange multipliers

$$\lambda_0 = -1 - \ln(2\pi\sigma^2)^{-\frac{1}{2}} + \frac{\mu^2}{2\sigma^2},\ \lambda_1 = -\frac{\mu}{\sigma^2},\ \lambda_2 = \frac{1}{2\sigma^2}.$$

2. Second approach: alternatively, we can substitute our candidate solution in the constraints and try to determine the Lagrange multipliers. Assuming that $\mathrm{Re}(\lambda_2) > 0$, we find

$$\int_{-\infty}^{\infty} f(x)dx = \sqrt{\frac{\pi}{\lambda_2}}e^{-1-\lambda_0+\frac{\lambda_1^2}{4\lambda_2}} = 1$$
$$\int_{-\infty}^{\infty} xf(x)dx = -\frac{\lambda_1}{2\lambda_2}\int_{-\infty}^{\infty} f(x)dx = \mu$$
$$\int_{-\infty}^{\infty} x^2 f(x)dx = \left(\frac{\lambda_1^2 + 2\lambda_2}{(2\lambda_2)^2}\right)\int_{-\infty}^{\infty} f(x)dx = \sigma^2 + \mu^2.$$

If we solve for the Lagrange multipliers we obtain the same result

The optimal density function is the Normal distribution:

$$f^*(x) = \frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

Intuition: the problem is assume we have an unknown distribution which is required to have an average equal to μ and a variance equal to σ^2, two known numbers, what is this distribution if we require the principle of maximum entropy to apply ? The answer is: the normal distribution.

6.4.5 Applications to the consumer problem

The optimal choice of varieties

Assume we have a continuum of varieties $\mathrm{I} \in [0, \infty)$ and a basket of consumption containing different varieties is $C = \left(c(i)\right)_{i \in \mathrm{I}}$ where $c(i) \geq$ is the quantity of variety $i \in \mathrm{I}$ in basket C. The composition of the basket can be seen as a function mapping between the space of varieties and a real number, the quantity consumed: $c : \mathrm{I} \rightarrow \mathbb{R}_{+}$, belonging to a space of positively-valued bounded functions, $c \in \mathcal{C}$.

The value of the basket is measured by the utility functional $\mathsf{U} : \mathcal{C} \rightarrow \mathbb{R}$:

$$\mathsf{U}[c] = \left(\int_{\mathrm{I}} c(i)^{1-\gamma} di\right)^{\frac{1}{1-\gamma}}.$$

The consumer has income $y > 0$ which can be spent on the purchase of baskets. The total expenditure is the functional

$$\mathsf{E}[c] = \int_{\mathrm{I}} p(i)\, c(i)\, di$$

where $p(.)$ is the relative price of variety relative to the income deflator. The budget constraint is $\mathsf{E}[p, c] = y$, assuming that the consumer has no savings.

The consumer problem is to find the basket $\Big(c(i)\Big)_{i \in \mathrm{I}}$ that solves the problem

$$\begin{aligned} &\max_{c(\cdot)} \mathrm{U}[c] = \left(\int_{\mathrm{I}} c(i)^{1-\gamma} di\right)^{\frac{1}{1-\gamma}} \\ &\text{subject to} \\ &\mathsf{E}[c] = \int_{\mathrm{I}} p(i)\, c(i)\, di = y. \end{aligned}$$

The Lagrangean functional can be interpreted as the indirect utility functional

$$\begin{aligned} \mathsf{L}[c; \lambda] &= \mathsf{U}[c] - \lambda \mathsf{E}[c] \\ &= \left(\int_{\mathrm{I}} c(i)^{1-\gamma} di\right)^{\frac{1}{1-\gamma}} - \lambda \int_{\mathrm{I}} p(i)\, c(i)\, di \end{aligned}$$

The first order conditions are

$$\begin{aligned} &\frac{\delta \mathsf{L}[c^*]}{\delta c(i)} = \frac{\delta \mathsf{U}[c^*]}{\delta c(i)} - \lambda \frac{\delta \mathsf{E}[c^*]}{\delta c(i)} = 0 \\ &\mathsf{E}[c^*] = y. \end{aligned}$$

The first condition is equivalent to

$$\left(\frac{c(i)}{\mathsf{U}[c]}\right)^{-\gamma} = \lambda\, p(i) \iff c(i) = \mathsf{U}[c]\Big(\lambda\, p(i)\Big)^{-\frac{1}{\gamma}}.$$

Substituting in the utility functional

$$\mathsf{U}[c] \; = \Big(\int_{\mathrm{I}} \Big(\mathsf{U}[c] \Big(\lambda\, p(i) \Big)^{-\frac{1}{\gamma}} \Big)^{1-\gamma} \; di \Big)^{\frac{1}{1-\gamma}}$$

allows us to find λ as

$$\frac{1}{\lambda} \; = P \equiv \Big(\int_{\mathrm{I}} p(i)^{\frac{\gamma-1}{\gamma}} \Big)^{\frac{\gamma}{\gamma-1}},$$

that we interpret as the consumer price deflactor (or the true index of cost of living) using as numéraire the income deflactor. Substuting in the constraint and simplifying yields $P\,\mathsf{U}[c] = y$. Substituting again in the expression for $c(i)$ yields the solution to the consumer problem as the basket $\big(c^*(i)\big)_{i\in\mathrm{I}}$ where

$$c^* = \frac{y}{P} \Big(\frac{P}{p(i)} \Big)^{\frac{1}{\gamma}}, \text{ for each } \; i \in \mathrm{I}.$$

Therefore, the demand for every variety is proportional to income, and the is negatively related to the price-elasticity of demand, which is homogenous among varieties and is equal to $1/\gamma$, because

$$\frac{\partial \ln c^*(i)}{\partial \ln p(i)} = -\frac{1}{\gamma} < 0.$$

6.4.6 Application to production theory

This problem is a component of most neo-Keynesian models, as the problem for a final producer using a continuum of inputs $X = (x(i))_{i\in\mathrm{I}}$. When bundles of inputs are Infinite-dimensional, as our X, models usually consider they refer to intermediary goods not factors of production.

Total sales in a competitive sales market $S(p, y) = py$ where p is the final good price and y is the output. Assume that the firm has a generalized constant elasticity of substitution (CES) production function

$$y = \mathsf{F}[x] \equiv \Big(\int_{\mathrm{I}} A(i)\, x(i)^{\frac{\epsilon-1}{\epsilon}} \, di \Big)^{\frac{\epsilon}{\epsilon-1}}$$

with specific factor productivity $A(i)$. Therefore, nominal sales are a functional

$$\mathsf{S}[x] \; = p\,\mathsf{F}[x],$$

where p is the price of the output. The total cost is also a functional

$$\mathsf{C}[x] \; = \int_{\mathrm{I}} w(i) x(i) di,$$

where $w(i)$ denotes the price of input $i \in \mathrm{I}$.

Therefore, firms profits are also a functional over the x

$$\Pi[x] = p\,\mathsf{F}[x] - \mathsf{C}[x].$$

Assuming that the firm is price-taker in all the markets, both the in the product market and the I markets for inputs, the firm's problem is to find the optimal input bundle $X^* = \left(x^*\right)_{i\in\mathrm{I}}$ that solves the problem

$$\begin{aligned}
&\max_{x(\cdot)} \Pi[x] = p\left(\int_{\mathrm{I}} A(i)\, x(i)^{\frac{\epsilon-1}{\epsilon}}\, di\right)^{\frac{\epsilon}{\epsilon-1}} - \int_{\mathrm{I}} w(i)x(i)di \\
&\quad\text{subject to} \\
&\quad\left(\int_{\mathrm{I}} A(i)\, x(i)^{\frac{\epsilon-1}{\epsilon}}\, di\right)^{\frac{\epsilon}{\epsilon-1}} = y
\end{aligned}$$

6.4.7 Optimal taxation

Let $\theta \in \Theta \subseteq \mathbb{R}_+$ denote the skill level of the population and let $f(\theta)$ be their density. Therefore $\int_\Theta f(\theta)\, d\theta = 1$. Let $\ell(\theta)$, $y(\theta) = \theta\,\ell(\theta)$, and $c(\theta)$ the work effort, the income and consumption by people of skill level θ. The utility for a household with skill level θ is $u(\theta) = U(c(\theta), \ell(\theta))$.

Assume the government has an exogenous expenditure level G and wants to find an optimal tax schedule $T(\theta) = \tau(\theta)\, y(\theta)$ which implements a social optimum. Assume that the social optimal criterium is the average social welfare

$$\mathsf{W}[u] = \int_\Theta W\big(u(\theta)\big)\, f(\theta)\, d\theta.$$

Assume that the central planner not only observes the distribution of income $y(\theta) = \theta\,\ell(\theta)$ but has also complete information on the work effort, that is it is able to separate the productivity θ from the work effort $\ell(\theta)$.

The problem is simpler if we use the implicit function theorem to solve $u(\theta) = U(c(\theta), \ell(\theta))$ for $c(\theta) = C\big(u(\theta), \ell(\theta)\big)$.

We can find the optimal allocation of utility and work effort by solving the functional problem

$$\begin{aligned}
&\max_{u(\cdot),\ell(\cdot)} \int_\Theta W\big(u(\theta)\big)\, f(\theta)\, d\theta \\
&\quad\text{subject to} \\
&\quad\int_\Theta \Big(\theta\,\ell(\theta) - C\big(u(\theta), \ell(\theta)\big)\Big)\, f(\theta)\, d\theta = G
\end{aligned}$$

If we can find c^* and ℓ^* then the optimal optimal tax-transfer policy that implements the social optimum, and finances the government expenditure G, has the following property

$$T^*\,(\theta) = y^*\, - c^*.$$

6.5 References

Generalized calculus in economics: Ok (2007), which observes that in most textbooks in functional analysis there is not much applications to differential and integral calculus and applications to optimizations.

Generalized calculus in applied mathematics: (Siddiqi, 2018, ch 5), (Drabek and Milota, 2013, ch 3)

Application of functional analysis to optimization, however with an emphasis in non-differentiable functions Clarke (2013)

Chapter 7

Calculus of variations

Calculus of variations provided the first approach for dealing with optimal variational problems since the early XVIII century (see Goldstine (1980)).

Explicit applications of calculus of variations to economics started, apparently, by a mathematician paper on the optimal pricing of a monopolistic firm Evans (1924). It was the main tools for solving intertemporal optimization problems at least until the interwar period (see Pomini (2018)). However, it is still useful, because it provides a better intuition on the optimization of functionals.

7.1 Calculus of variations: introduction

Calculus of variations problems consist in finding an extreme of a functional over a function $y : \mathrm{X} \to \mathcal{Y}$, which can be subject to additional requirements. Solving a calculus of variations problem means finding function $y^*(.)$ belonging to an an admissible set of continuous and differentiable functions $\mathcal{Y}$ (not necessarily everywhere).

The **objective functional** takes the form

$$\mathsf{J}[y] \equiv \int_{x_0}^{x_1} F(x, y(x), y'(x))\, dx \tag{7.1}$$

in which we function $F(.)$ is known.

We assume that $F_{y'}(x, y, y') \equiv \dfrac{\partial F(x, y, y')}{\partial y'} \neq 0$, except maybe on a subset of measure zero (i.e., it may be piecewise-differentiable). This is the characteristic of function $F(.)$ which makes the problem a dynamic optimization problem (if time is the independent variable), in the sense that the optimization involves a trade-off between the current state $y(x)$ and the change in the current state, $y'(x)$. If $F_{y'}(x, y(x), y'(x)) = 0$ globally, for any $x \in \mathrm{X}$ the problem will degenerate to a static functional optimization.

To understand the effect of the derivative on the optimum, consider instead the objective functional

$$\mathsf{J}_0[y] \equiv \int_{x_0}^{x_1} F(x, y(x))\, dx.$$

If there are no other conditions, if $y^*(x)$ is the optimum, a necessary condition is

$$\frac{\delta \mathsf{J}_0[y^*]}{\delta y(x)} = F_y(x, y^*(x)) = 0, \text{ for every } x \in \mathrm{X}$$

where $F_y(x,y) = \dfrac{\partial F(x,y)}{\partial y}$. This condition is a point-wise optimality criterium: the optimum $y^*(x)$ is found by finding an extremum for every point in $x \in \mathrm{X}$ independent of any other point $\in \mathrm{X}$. If the objective function depends on the derivative of function $y(\cdot)$, $y'(\cdot)$, this means that the local interaction influences the value of the problem. This has two consequences: first, the optimum cannot be just determined by a point-wise extremum, and, second, any constraint on the value of y will influence the solution.

This also means that we should look for solutions $y \in C^1(\mathrm{X};\mathcal{Y})$, where $C^1(\mathrm{X};\mathcal{Y})$ is the set of continuously differentiable functions mapping X into $\mathcal{Y}$.

Two observations are important referring to the nature of the independent variable, x, and to its domain X.

First, in most economic applications, x is a non-negative real number referring to time , i.e. $x = t$ and $\mathrm{X} = \mathrm{T} \subseteq \mathbb{R}_+$. However in some microeconomic problems or static macroeconomic problems with heterogeneity among agents, in which there are, for example, information or searching frictions, we need to solve optimal control problems in which the independent variable is not time and has a support belonging to a continuum, for instance $\mathrm{X} = [x_0, x_1]$. In time-dependent problems we call $x_0 = t_0$ the initial time and $x_1 = t_1$ the terminal time, or horizon, while for non-time-dependent models the designation depends on the context. For example in models in which x refers to the skill level x_0 refers to the lowest skill in the distribution and x_1 to the highest skill. Therefore, from now on we call x_0 the lower bound and x_1 the upper bound of X.

Second, in time-dependent problems we usually assume that $x_0 = t_0$ and $x_1 = t_1$ may be fixed (v.g., in macroeconomic models) or free (v.g., in microeconomic problems). If x refers to other type of variables x_0 and x_1 may refer to cutoff points which can be free and optimally determined.

At last, another important point to be made, which is particularly important in macroeconomics is related to the boundedness of X. We can consider x_1 to be bounded or unbounded $x_1 = \infty$. In the case in which x refers to time we have to distinguish between **finite or infinite horizon** cases.

7.2 Bounded domains and equality constraints

In this section, we start with the simplest case, in subsection 7.2.1 the case in which the boundary of X and the values of the state variables at that boundary are also known, i.e, x_0, x_1, $y(x_0)$ and $y(x_1)$ are known. In section 7.2.2 we consider the cases in which x_0 are x_1 known but $y(x_0)$ and $y(x_1)$ are free. In section 7.2.3, the cases in which known $y(x_0)$ and $y(x_1)$ are known but x_0 are x_1 are free and the cases in which x_0, x_1, $y(x_0)$ and $y(x_1)$ are all free. At last, in section 7.2.4, we deal with two cases which are common to time-dependent models: the existence of terminal constraints and the infinite horizon problem.

7.2.1 The simplest CV problem

The **simplest problem of calculus of variations** is the following: find a function $y : \mathrm{X} \to \mathrm{Y} \subseteq \mathbb{R}$, that maximizes the **objective functional** (7.1) such that: (i) the set of independent variables is closed and bounded $\mathrm{X} = [x_0, x_1]$, with fixed limits x_0 and x_1 satisfying $x_0 < x_1$, (ii) the functions $y(\cdot)$ is continuous and continuously differentiable (except at a finite number of points) and satisfies $y(x_0) = y_0$ and $y(x_1) = y_1$.

The **admissibility set**,

$$\mathcal{Y} \equiv \Big\{ y(x) \in \mathrm{X} : y(x_0) = y_0,\ y(x_1) = y_1 \Big\} \subseteq \mathcal{C}^1(\mathbb{R})$$

is the set of all functions continuous and differentiable functions that satisfy the lower and upper boundary data and $\mathrm{X} = [x_0, x_1]$.

Therefore, the problem is to find a function $y^* \in \mathcal{Y}$, which maximizes the functional (7.1).

Formally, the simplest problem is:

$$\begin{aligned} &\max_{y(\cdot)\in\mathcal{Y}} \int_{x_0}^{x_1} F(x, y(x), y'(x))\, dx \\ &\text{subject to} \\ &x_0 \text{ and } x_1 \text{ fixed} \\ &y(x_0) = y_0 \text{ and } y(x_1) = y_1 \text{ fixed} \end{aligned} \tag{P1}$$

We denote by $\varphi = (x_0, x_1, y_0, y_1, .)$ be vector of the data of the problem containing the lower and upper values of the independent variable, the associated values of the state function, and other parameters that might exist in function $F(.)$.

The **value function**, $V(\varphi) = \mathsf{J}[y^*]$, is a real-valued function depending on the data of the problem, that is

$$V(x_0, x_1, y_0, y_1, .) = \mathsf{J}[y^*] \equiv \max_{y\in\mathcal{Y}} \int_{x_0}^{x_1} F(x, y^*(x), y^{*'}(x))\, dx,$$

where $\mathcal{Y}$ is the admissibility set.

Proposition 1. ***First order necessary conditions for the simplest problem, (P1)*** *: $y^* : [x_0, x_1] \to \mathrm{Y}$ is a solution of the simplest CV problem only if it satisfies the* ***Euler-Lagrange equation*** [1]*:*

$$F_y(x, y^*(x), y^{*'}(x)) = \frac{d}{dx}\Big(F_{y'}(x, y^*(x), y^{*'}(x)) \Big),\ \textit{for}\ x \in (x_0, x_1) \tag{7.2}$$

together with the boundary conditions

$$y^*(x_0) = y_0,\ \textit{and}\ y^*(x_1) = y_1 \tag{7.3}$$

[1] We use the notation $F_y\ (x, y, y') = \dfrac{\partial F(x, y, y')}{\partial y}$ and $F_{y'}(x, y, y') = \dfrac{\partial F(x, y, y')}{\partial y'}$.

Proof. (Heuristic) Assume we know y^*. Then the maximum value for the functional is

$$\mathsf{J}[y^*] = \int_{x_0}^{x_1} F(x, y^*(x), y^{*'}(x))\, dx.$$

Function y^* is an optimum only if $\mathsf{J}[y^*] \geq \mathsf{J}[y]$ for any other admissible function $y : \mathrm{X} \to \mathrm{Y}$. Take an admissible variation over y^*, $y = y^* + \delta y$ such that the variation is a *parameterized perturbation* of y^*, that is $\delta y(x) = \varepsilon\, \eta(x)$ where $\eta \in \mathcal{Y}$ and ε is a number. A variation to be admissible has to satisfy $y(x_1) = y^*(x_1) = y_1$ and $y(x_0) = y^*(x_0) = y_0$. Therefore, an admissible perturbation has to satisfy $\eta(x_0) = \eta(x_1) = 0$ and it can take arbitrary values $\eta \in \mathrm{Y}$ for x in the interior of the domain X.

The value functional for the perturbed function y is

$$\mathsf{J}[y] = \mathsf{J}[y^* + \varepsilon\eta] \; = \int_{x_0}^{x_1} F(x, y^*(x) + \varepsilon\eta(x), y^{*'}(x) + \varepsilon\eta'(x))dx.$$

The variation of the functional is

$$\begin{aligned}\Delta J(\varepsilon) = &\; \mathsf{J}[y^* + \varepsilon\, \eta] - \; \mathsf{J}[y^*] \\ = &\int_{x_0}^{x_1} \Big(F(x, y^*(x) + \varepsilon\eta(x), y^{*'}(x) + \varepsilon\eta'(x)) - F(x, y^*(x), y^{*'}(x)) \Big)\, dx.\end{aligned}$$

Defining the Gâteaux derivative of a functional evaluated at y^* for the perturbation η

$$\delta\mathsf{J}[y^*](\eta) = \lim_{\varepsilon \to 0} \; \frac{\Delta J(\varepsilon)}{\varepsilon},$$

a first-order expansion of the functional $\mathsf{J}[y]$ in a neighbourhood of y^*,

$$\mathsf{J}[y] \; = \mathsf{J}[y^*] + \delta\mathsf{J}[y^*](\eta)\varepsilon + o(\varepsilon)$$

Then, at the optimum, $\mathsf{J}[y^*] \geq \mathsf{J}[y]$ only if the first integral derivative of J is zero: $\delta\mathsf{J}[y^*](\eta) = 0$. We can find the Gâteaux derivative by using the formula

$$\delta\mathsf{J}[y^*](\eta) = \frac{d}{d\varepsilon}\mathsf{J}[y^* + \varepsilon\eta]\Big|_{\varepsilon=0}.$$

In this case, we have

$$\begin{aligned}\delta\mathsf{J}[y^*](\eta) = &\int_{x_0}^{x_1} \Big(F_y(x, y^*(x), y^{*'}(x))\, \eta(x) + F_{y'}(x, y^*(x), y^{*'}(x)\Big)\, \eta'(x)\Big)\, dx \\ = &\int_{x_0}^{x_1} F_y(x, y^*(x), y^{*'}(x))\, \eta(x)\, dx + \int_{x_0}^{x_1} F_{y'}(x, y^*(x), y^{*'}(x)\Big)\, \eta'(x)\, dx.\end{aligned}$$

Integrating by parts the second integral yields

$$\begin{aligned}\int_{x_0}^{x_1} F_{y'}(x, y^*(x), y^{*'}(x))\, \eta'(x)dx = &\int_{x_0}^{x_1} F_{y'}(x, y^*(x), y^{*'}(x))\, \eta(x) - \int_{x_0}^{x_1} dF_{y'}(x, y^*(x), y^{*'}(x))\, \eta(x) = \\ = &\; F_{y'}(x_1, y^*(x_1), y^{*'}(x_1))\, \eta(x_1) - F_{y'}(x_0, y^*(x_0), y^{*'}(x_0))\, \eta(x_0) \\ &- \int_{x_0}^{x_1} \frac{d}{dx} F_{y'}(x, y^*(x), y^{*'}(x))\, \eta(x)\, dx.\end{aligned}$$

Therefore,

$$\begin{aligned}\delta \mathsf{J}[y^*](\eta) = \int_{x_0}^{x_1} \Big(F_y(x, y^*(x), y^{*'}(x)) - \frac{d}{dx} F_{y'}(x, y^*(x), y^{*'}(x)) \Big)\, \eta(x)\, dx + \\ + F_{y'}(x_1, y^*(x_1), y^{*'}(x_1))\, \eta(x_1) - F_{y'}(x_0, y^*(x_0), y^{*'}(x_0))\, \eta(x_0).\end{aligned} \tag{7.4}$$

As, in this case with fixed boundary values for the variable y, the admissible perturbation satisfies $\eta(x_1) = \eta(x_0) = 0$ equation (7.4) reduces to

$$\delta \mathsf{J}[y^*](\eta) = \int_{x_0}^{x_1} \Big(F_y(x, y^*(x), y^{*'}(x)) - \frac{d}{dx} F_{y'}(x, y^*(x), y^{*'}(x)) \Big)\, \eta(x)\, dx$$

If $F(.)$ is a continuous function we can use the following result (see (Gel'fand and Fomin, 1963, p.9)): if $h := [x_0, x_1] \to \mathbb{R}$ is a continuous function and $\int_{x_0}^{x_1} h(x)\eta(x)dx$ for all C^1 functions η and if $\eta(x_0) = \eta(x_1) = 0$ then $\int_{x_0}^{x_1} h(x)\eta(x)dx = 0$ if and only if $h(x) = 0$ for all $x \in (x_0, x_1)$.

Therefore $\delta \mathsf{J}[y^*](\eta) = 0$ if an only if $F_y(x, y^*(x), y^{*'}(x)) - \frac{d}{dx} F_{y'}(x, y^*(x), y^{*'}(x)) = 0$ for every $x \in \mathrm{X} = [x_0, x_1]$. □

Writing $F^*\ (x) \equiv F(x, y^*(x), y^{*'}(x))$ and $F_j^*\ (x) \equiv F_j(x, y^*(x), y^{*'}(x))$ for $j \in y, y'$, its derivatives evaluated at the optimum, the Euler-Lagrange equation is a 2nd order ODE (ordinary differential equation) if $F_{y'y'} \neq 0$: if we expand the right-hand-side we find

$$F^*_{y'y'}(x)\, y''(x) + F^*_{y'y}(x)\, y'\ (x) + F^*_{y'x}(x) - F^*_y(x) = 0, \text{ for each } x \in [x_0, x_1].$$

We can transform it into a system of first order ODE's if we define $y_1 = y$ and $y_2 = y'$ then

$$\begin{aligned} y_1' &= y_2 \\ F_{y_2 y_2}(x, y_1, y_2)\, y_2' &= F_{y_1}(x, y_1, y_2) - F_x(x, y_1, y_2) - F_{y_2 y_1}(x, y_1, y_2) y_2.\end{aligned}$$

The first order necessary condition only allows for the determination of an extremum. In order to get the a necessary condition for a maximand we need a second order condition:

Proposition 2. ***Second order necessary conditions****: the solution to the CV problem $y^* : \mathrm{X} \to \mathrm{Y}$ is a maximand only if it satisfies the Legendre-Clebsch condition*

$$F_{y'y'}(x, y^*(x), y^{*'}(x)) \leq 0 \tag{7.5}$$

Proof. (Heuristic but more complicated). Performing a second -order expansion of the functional $\mathsf{J}[x]$ in a neighbourhood of y^*, we obtain

$$\mathsf{J}[y]\ = \mathsf{J}[y^*] + \delta \mathsf{J}[y^*](\eta)\varepsilon + \frac{1}{2}\ \delta^2 \mathsf{J}[y^*](\eta)\varepsilon^2 + o(\varepsilon^2),$$

where

$$\delta^2 \mathsf{J}[y^*](\eta) = \frac{d^2}{d\varepsilon^2}\Big|_{\varepsilon=0} \mathsf{J}[y^*\ + \varepsilon\eta].$$

Because at the optimum for any admissible perturbation η we have $\delta \mathsf{J}[y^*](\eta) = 0$, and at a have a maximum $\mathsf{J}[y] \ \leq \mathsf{J}[y^*]$, a necessary condition is $\delta^2 \mathsf{J}[y^*](\eta) \leq 0$.

The second-order functional derivative is

$$\delta^2 \mathsf{J}[y^*](\eta) = \int_{x_0}^{x_1} \Big(F_{yy}(x, y^*(x), {y^*}'(x))\, \eta(x)^2 + \\ + 2F_{yy'}(x, y^*(x), {y^*}'(x))\, \eta(x)\, \eta'(x) + F_{y'y'}(x, y^*(x), {y^*}'(x))\, (\eta'(x))^2 \Big)\, dx.$$

As

$$\begin{aligned} \int_{x_0}^{x_1} 2F^*_{yy'}(x)\, \eta(x)\, \eta'(x) dx &= \int_{x_0}^{x_1} F^*_{yy'}(x) \frac{d}{dx}\, (\eta(x)^2) dx = \\ &= F^*_{yy'}(x)\ \eta(x)^2 \Big|_{x_0}^{x_1} - \int_{x_0}^{x_1} \frac{d}{dx} \left(F^*_{yy'} \right) (\eta(x)^2) dx \\ &= - \int_{x_0}^{x_1} \frac{d}{dx} \left(F^*_{yy'} \right) (\eta(x)^2) dx \end{aligned}$$

because of the admissibility conditions $\eta(x_0) = \eta(x_1) = 0$. Then

$$\delta^2 \mathsf{J}[y^*](\eta) = \int_{x_0}^{x_1} \left(\left(F^*_{yy}(x) - \frac{d}{dx} F^*_{yy'}(x) \right) \eta(x)^2 + F^*_{y'y'}(x) (\eta'(x))^2 \right) dx.$$

Following (Liberzon, 2012, p.59-60)), it can be shown that $\delta^2 \mathsf{J}[y^*](\eta) \leq 0$ only if condition (7.5) holds. □

Proposition 3. ***Sufficient conditions****: let $y^* \in \mathcal{Y}$ verify*

$$F_y\ (t, y^*, {y^*}') = \frac{d}{dx}\ F_{y'}(t, y^*, {y^*}') \ \text{ and } F_{y'y'}(t, y^*, {y^*}') \leq 0$$

then (under some additional conditions on the trajectory of y) y^* is an optimiser to $\mathsf{J}[y]$.

Proof. See (Liberzon, 2012, p.62-68) □

Proposition 4. ***Necessary and sufficient conditions****: consider the simplest calculus of variations problem and assume that $F_{y'y'}(x, y(x), y'(x)) \leq 0$ for every $x \in [x_0, x_1]$ then equations* (7.2) *and* (7.3) *are necessary and sufficient conditions.*

7.2.2 Free boundary values for the state variable

Now we consider the problem: find function y^* among admissible functions $y \in \mathcal{Y}$ having the following properties: $y : \mathrm{X} \to Y \subseteq \mathbb{R}$, where $\mathrm{X} = [x_0, x_1]$ has known boundaries, x_0 and x_1, and such that $y(x_0)$ and/or $y(x_1)$ are free. The objective functional is again (7.1).

Formally, the problem are:

$$\max_{y(\cdot)} \int_{x_0}^{x_1} F(x, y(x), y'(x))\, dx$$

$$\begin{aligned}
&\text{subject to} \\
&x_0 \text{ and } x_1 \text{ fixed} \\
&y(x_0) = y_0 \text{ and } y(x_1) = y_1 \text{ free} && \text{(P2a)} \\
&y(x_0) = y_0 \text{ fixed, } \; y(x_1) = y_1 \text{ free} && \text{(P2b)} \\
&y(x_0) = y_0 \text{ free } y(x_1) = y_1 \text{ fixed} && \text{(P2c)}
\end{aligned}$$

We have the following data, value functions and admissibility sets:

- in the case of the problem (P2a), the parameter set is $\varphi = (x_0, x_1, .)$, the value function is

  $$V(x_0, x_1, .) = \max_{y \in \mathcal{Y}} \int_{x_0}^{x_1} F(x, y(x), y'(x))dx.$$

  and the admissible set is

  $$\mathcal{Y} \equiv \Big\{ y(x) \in \mathrm{Y} : \; x \in \mathrm{X} \Big\} \; \subseteq \mathcal{C}^1(\mathbb{R});$$

- in the case of the problem (P2b), the parameter set is $\varphi = (x_0, x_1, y_0, .)$, the value function is

  $$V(x_0, x_1, y_0, .) = \max_{y \in \mathcal{Y}} \int_{x_0}^{x_1} F(x, y(x), y'(x))dx.$$

  and the admissible set is

  $$\mathcal{Y} \equiv \Big\{ y(x) \in \mathrm{Y} : \; y(x_0) = y_0, x \in \mathrm{X} \Big\} \; \subseteq \mathcal{C}^1(\mathbb{R});$$

- in the case of the problem (P2c), the parameter set is $\varphi = (x_0, x_1, y_1, .)$, the value function is

  $$V(x_0, x_1, y_1 .) = \max_{y \in \mathcal{Y}} \int_{x_0}^{x_1} F(x, y(x), y'(x))dx.$$

  and the admissible set is

  $$\mathcal{Y} \equiv \Big\{ y(x)\mathrm{Y} : \; y(x_1) = y_1, \; x \in \mathrm{X} \Big\} \; \subseteq \mathcal{C}^1(\mathbb{R}).$$

Proposition 5. ***First order necessary conditions for the free terminal state problem:*** *$y^* \in \mathcal{Y}$ is the solution to one of the CV problem with free boundary values for the state variable and known terminal values for the independent variable, x_0 and x_1, problems P2a, P2b, or P2c, only if it satisfies the Euler equation* (7.2) *and the boundary conditions:*

1. *if both boundary values are free (problem* (P2a)*)*

$$F_{y'}(x_0, y^*(x_0), y^{*'}(x_0)) = 0, \; \textit{and} \; F_{y'}(x_1, y^*(x_1), y^{*'}(x_1)) = 0 \tag{7.7}$$

2. *if the lower boundary value is given by* $y(x_0) = y_0$*, and the upper boundary value is free (problem* (P2b)*)*
$$y^*(x_0) = y_0, \ \textit{and}\ F_{y'}(x_1, y^*(x_1), y^{*'}(x_1)) = 0 \tag{7.8}$$

3. *if the upper boundary value is given by* $y(x_1) = y_1$*, and the lower boundary value is free (problem* (P2c)*)*
$$F_{y'}(x_0, y^*(x_0), y^{*'}(x_0)) = 0, \ \textit{and}\ y^*(x_1) = y_1. \tag{7.9}$$

Proof. (Heuristic) Now the boundary values for perturbation are $\eta(x_0)$ and $\eta(x_1)$ can take any value, including zero if the associated boundary value $y(x)$, for $x \in \{x_0, x_1\}$ is fixed. The proof follows the same steps as in the proof of Proposition 1. However, in equation (7.4), in order to get $\delta \mathsf{J}[y^*](\eta) = 0$, and after introducing the Euler-Lagrange condition, we should have

$$F_{y'}(x_j, y^*(x_j), y^{*'}(x_j))\, \eta(x_j) = 0, \ \text{for}\ j = 0, 1. \tag{7.10}$$

Thus we have two cases, concerning the adjoint conditions at boundary x_j, for $j = 0, 1$, for an optimum. First, if the value of the state variable for the boundary x_j is known, i.e., $y(x_j) = y_j$, an admissible perturbation should verify $\eta(x_j) = 0$, implying that condition (7.10) holds automatically. This is the case in Proposition 1. Second, if the value of the state variable for the boundary x_j is free, then the related perturbation value is arbitrary and $\eta(x_j) \neq 0$ in general. The optimally condition (7.10) holds if and only if $F_{y'}(x_j, y^*(x_j), y^{*'}(x_j)) = 0$ which provides one adjoint condition allowing for the determination of the optimal boundary value for the state variable $y^*(x_j)$. This is how we adjoint (7.7) to (7.9) depending on which boundary value for the state variable is free. □

In time-varying models in which the value of the state variable is known at time $t = 0$ and the terminal value of the state variable is endogenous we supplement the Euler-Lagrange with condition (7.8).

However, there are models in which the initial value of the state variable is unknown. This is the case, for instance, in optimal taxation models of the Mirrlees (1971) type in which the independent variable are skill values and the initial condition is related to the cutoff level of skill bellow which taxes should be zero. In this case condition (7.7) can be used.

Observation: as the Euler-Lagrange is a second-order differential equation, in order to fully solve a model we need to have information on the value of y at the two boundaries for $x = x_0$ and $x = x_1$.

7.2.3 Free boundary values for the independent variable

Now we consider the problem: find function $y^* \in \mathcal{Y}$ which is the set of functions $y : \mathrm{X}^* \to \mathbb{R}$, where X^* has at least one unknown boundary, x_0^* and/or x_1^*, but such that the terminal values for the state variable are known. That is $\mathrm{X}^* = [x_0^*, x_1]$ or $\mathrm{X}^* = [x_0, x_1^*]$ or $\mathrm{X}^* = [x_0^*, x_1^*]$ where x_j is known and x_j^* is free. If a boundary value for the independent variable is free the related boundary value for the state variable is known, that is $y(x_j^*) = y_j$. The objective functional is again (7.1).

Formally, the problem are :

$$\begin{aligned}
&\max_{y(\cdot)} \int_{x_0}^{x_1} F(x, y(x), y'(x))\, dx \\
&\text{subject to} \\
&y(x_0) = y_0 \text{ and } y(x_1) = y_1 \text{ fixed} \\
&x_0 \text{ and } x_1 \text{ free} && \text{(P3a)} \\
&x_0 \text{ fixed, } x_1 \text{ free} && \text{(P3b)} \\
&x_0 \text{ free } x_1 \text{ fixed} && \text{(P3c)}
\end{aligned}$$

In this case the data of the problem is $\varphi = (y_0, y_1, .)$. and the value functional is

$$V(y_0, y_1, .) = \max_{y \in \mathcal{Y}, x_0, x_1} \int_{x_0}^{x_1} F(x, y(x), y'(x))dx = \int_{x_0^*}^{x_1^*} F(x, y^*(x), y^{*'}(x))dx$$

where we x_0^* and/or x_1^* are determined endogenously. Now, we have $\mathrm{X}^* = [x_0^*, x_1^*]$ and the admissible set for problem (P3a) is

$$\mathcal{Y} = \{ y(x) \in \mathrm{Y} : y(x_0) = y_0, \ y(x_1) = y_1, \ x \in \mathrm{X}^* \}.$$

The other admissible sets are defined accordingly.

Proposition 6. ***First order necessary conditions for the free boundaries value problem:*** *$y^* \in \mathcal{Y}$ is the solution to the CV problem with known boundary values for the state variable, y_0 and y_1, and free terminal values for the independent variable, problems P3a, P3b, or P3c, only if it satisfies the Euler equation* (7.2) *and the boundary conditions:*

1. *if both boundary values for the indepdendent variable are free (problem P3a)*

$$\begin{aligned}
&F(x_0^*, y_0, y^{*'}(x_0^*)) - F_{y'}(x_0^*, y_0, y^{*'}(x_0^*))\, y^{*'}(x_0^*) = 0 \\
&\qquad \textit{and} \ \ F(x_1^*, y_1, y^{*'}(x_1^*)) - F_{y'}(x_1^*, y_1, y^{*'}(x_1^*))\, y^{*'}(x_1^*) = 0 \qquad (7.12)
\end{aligned}$$

2. *if the lower boundary value for the independent variable is known, $x_0^* = x_0$, and the upper boundary for the independent variable is free (problem P3b)*

$$x_0^* = x_0, \ \textit{and}\, F(x_1^*, y_1, y^{*'}(x_1^*)) - F_{y'}(x_1^*, y_1, y^{*'}(x_1^*))y^{*'}(x_1^*) = 0 \qquad (7.13)$$

3. *if the upper boundary value for the independent variable is known, $x_1^* = x_1$, and the lower boundary for the independent variable is free (problem P3c)*

$$F(x_0^*, y_0, y^{*'}(x_0^*)) - F_{y'}(x_0^*, y_0, y^{*'}(x_0^*))y^{*'}(x_0^*) = 0, \ \textit{and}\, x_1^* = x_1. \qquad (7.14)$$

Proof. (Heuristic) Let us assume that we know the solution $y^*(x)$ for $x \in [x_0^*, x_1^*]$, that is for all values of the independent variable contained between the two optimally chosen boundary values.

In this case we have to introduce two types of perturbations: a perturbation to the state variable $y(x) = y^*(x) + \varepsilon\eta(x)$, by function $\eta(\cdot)$, and to the independent variable $x = x^* + \varepsilon\chi$ by a constant χ. If we denote $y_j^* = y^*(x_j^*)$, for $j = 0, 1$, the two boundary values for the independent and dependent variables are $P_j^* \equiv (x_j^*, y_j^*)$ for $j = 0, 1$ at the optimum. The related terminal points for the perturbed solution are written as $P_j = (x_j^* + \varepsilon\chi_j, y_j^* + \varepsilon\eta_j)$ for $j = 0, 1$.

At the optimum the objective functional is

$$\mathsf{J}[y^*; x^*] = \int_{x_0^*}^{x_1^*} F(x, y^*(x), y^{*'}(x))dx$$

and

$$\mathsf{J}[y^* + \varepsilon\eta; x^* + \varepsilon\chi] = \int_{x_0^*+\varepsilon\chi_0}^{x_1^*+\varepsilon\chi_1} F(x, y^*(x) + \varepsilon\eta(x), y^{*'}(x) + \varepsilon\eta'\ (x))dx.$$

Then, denoting $\Delta\mathsf{J}(\varepsilon) = \mathsf{J}[y^* + \varepsilon\eta; x^* + \varepsilon\chi] - \mathsf{J}[y^*; x^*]$ we have

$$\begin{aligned}
\Delta\mathsf{J}(\varepsilon) &= \int_{x_0^*+\varepsilon\chi_0}^{x_1^*+\varepsilon\chi_1} F\big(x, y^*(x) + \varepsilon\eta(x), y^{*'}(x) + \varepsilon\,\eta'(x)\big)\,dx - \int_{x_0^*}^{x_1^*} F(x, y^*(x), y^{*'}(x))\,dx\\
&= \int_{x_0^*}^{x_1^*} \Big(F(x, y^*(x) + \varepsilon\eta(x), y^{*'}(x) + \varepsilon\eta'\ (x)) - F(x, y^*(x), y^{*'}(x))\Big)dx+\\
&\quad + \int_{x_1^*}^{x_1^*+\varepsilon\chi_1} F(x, y^*(x) + \varepsilon\eta(x), y^{*'}(x) + \varepsilon\eta'\ (x))dx-\\
&\quad - \int_{x_0^*+\varepsilon\chi_0}^{x_0^*} F(x, y^*(x) + \varepsilon\eta(x), y^{*'}(x) + \varepsilon\eta'\ (x))dx
\end{aligned}$$

Denoting $F^*(x) = F(t, y^*(x), \dot{y}^*(x))$ and using the mean-value theorem,

$$\Delta\mathsf{J}(\varepsilon) = \varepsilon\int_{x_0^*}^{x_1^*} \big(F_y^*(x)\eta(x) + F_{y'}^*(x)\eta'(x)\big)\,dx + F(\tilde{x}_1)\,\varepsilon\,\chi_1 - F(\tilde{x}_0)\,\varepsilon\,\chi_0$$

where $\tilde{x}_1 \in (x_1^*, x_1^* + \varepsilon\,\chi_1)$ and $\tilde{x}_0 \in (x_0^*, x_0^* + \varepsilon\,\chi_0)$. Taking $\delta\mathsf{J}[y^*; x^*]\ (\eta, \chi) = \lim_{\varepsilon\to 0} \dfrac{\Delta\mathsf{J}(\varepsilon)}{\varepsilon}$, the functional derivative becomes

$$\delta\mathsf{J}[y^*; x^*]\ (\eta, \chi) = \int_{x_0^*}^{x_1^*} \Big(F_y^*(x)\eta(x) + F_{y'}^*(x)\,\eta'(x)\Big)dx + F^*(x)\Big|_{x=x_1^*}\chi_1 - F^*(x)\Big|_{x=x_0^*}\chi_0,$$

where $\chi = (\chi_0, \chi_1)$. Integration by parts yields

$$\begin{aligned}
\delta\mathsf{J}[y^*; x^*]\ (\eta, \chi) = &\int_{x_0^*}^{x_1^*} \Big(F_y^*(x) - \frac{d}{dx}F_{y'}^*(x)\Big)\,\eta(x)dx+\\
&+ F_{y'}^*(x)\eta(x)\Big|_{x=x_1^*} - F_{y'}^*(x)\eta(x)\Big|_{x=x_0^*} + F^*\,(x)|_{x=x_1^*}\,\chi_1 - F^*(t)|_{x=x_0^*}\,\chi_0.
\end{aligned}$$

We only know the perturbations for the state variables at the perturbed boundaries x_0 and x_1 and not at x_0^* and x_1^*, which inhibits the computation of the integral in the last equation. In order to find $\eta(x_j^*)$, using the approximation $\eta(x_j) \approx \eta(x_j^*) + y'(x_j^*)\,\chi_j$, we introduce

$$\eta(x_j^*) \approx \eta_j - y'(x_j^*)\chi_j,\ \text{for}\ j = 0, 1.$$

Therefore,

$$\delta \mathsf{J}[y^*; x^*]\ (\eta, \chi) = \int_{x_0^*}^{x_1^*} \left(F_y^*(x) - \frac{d}{dx} F_{y'}^*(x) \right) \eta(x) dx + F_{y'}^*(x_1^*)\eta_1 - F_{y'}^*(x_0^*)\eta_0 +$$
$$+ \left(F^*(x) - F_{y'}^*(x) y'(x) \right)\Big|_{x=x_1^*} \chi_1 - \left(F^*(x) - F_{y'}^*(x) y'(x) \right)\Big|_{x=_0^*} \chi_0$$

As the terminal values of the state variables, $y(x_0^*) = y_0$ and $y(x_1^*) = y_1$, are known then the terminal perturbation for the independent variable should satisfy $\eta_0 = \eta_1 = 0$. Therefore, $\delta \mathsf{J}[y^*; x^*]\ (\eta, \chi) = 0$ if and only if the Euler-Lagrange equation holds and $\left(F^*(x) - F_{y'}^*(x) y'(x) \right)\Big|_{x=x_1^*} \chi_1 = 0$ and/or $\left(F^*(x) - F_{y'}^*(x) y'(x) \right)\Big|_{x=_0^*} \chi_0 = 0$. This encompasses the three cases in equations (P5), (7.13) and (7.14). □

7.2.4 Free boundaries for both independent and dependent variables

The most general problem is: find function $y^* \in \mathcal{Y}$ among functions $y : \mathrm{X}^* \to \mathbb{R}$, where X^* has at least one unknown boundary, x_0^* and/or x_1^*, as in the previous subsection, and the terminal values for the state variables, $y(x_0^*)$ and/or $y(x_1^*)$ are also free. The objective functional is again (7.1).

Formally, the problem are :

$$\max_{y(\cdot)} \int_{x_0}^{x_1} F(x, y(x), y'(x))\, dx$$

subject to

$$y(x_0) = y_0 \text{ and } x_0 \text{ free},\ y(x_1) = y_1 \text{ and } x_1 \text{ fixed} \tag{P4a}$$
$$y(x_0) = y_0 \text{ and } x_0 \text{ fixed},\ y(x_1) = y_1 \text{ and } x_1 \text{ free} \tag{P4b}$$
$$y(x_0) = y_0,\ x_0,\ y(x_1) = y_1 \text{ and } x_1 \text{ free} \tag{P4c}$$

In this case the data of the problem, $\varphi = (.)$, only involves parameters that may be present in function $F(.)$. The value functional is

$$V(.) = \max_{y \in \mathcal{Y}, x_0, x_1} \int_{x_0}^{x_1} F(x, y(x), y'(x)) dx = \int_{x_0^*}^{x_1^*} F(x, y^*(x), y^{*'}(x)) dx$$

where we x_0^* and/or x_1^* and $y^*(x_0^*)$ and/or $y^*(x_1^*)$ are determined endogenously.

Proposition 7. ***First order necessary conditions for the free terminal boundary problem:*** *$y^* \in \mathcal{Y}$ is the solution to the CV problem with free boundary values for the state variable and for the independent variable, only if it satisfies the Euler equation* (7.2) *and the boundary conditions:*

1. *if both values for lower boundary are free (problem P4a)*

$$F_y(x_0^*, y^*(x_0^*), y^{*'}(x_0^*)) = F_{y'}(x_0^*, y^*(x_0^*), y^{*'}(x_0^*)) = 0 \tag{7.16}$$

2. *if both values for upper boundary are free (problem P4b)*

$$F_y(x_1^*, y^*(x_1^*), y^{*'}(x_1^*)) = F_{y'}(x_1^*, y^*(x_1^*), y^{*'}(x_1^*)) = 0 \tag{7.17}$$

3. *if all terminal values for x and $y(x)$ are free (problem P4c)*

$$F_y(x_0^*, y^*(x_0^*), y^{*'}(x_0^*)) = F_{y'}(x_0^*, y^*(x_0^*), y^{*'}(x_0^*)) = 0 \tag{7.18a}$$

$$F_y(x_1^*, y^*(x_1^*), y^{*'}(x_1^*)) = F_{y'}(x_1^*, y^*(x_1^*), y^{*'}(x_1^*)) = 0 \tag{7.18b}$$

Proof. We use the previous proof and, in equation (7.15), we consider $\eta_0 \neq 0$, $\eta_1 \neq 0$, $\chi_0 \neq 0$ and $\chi_1 \neq 0$. □

Table 7.1 assembles all the previous results. Observe that if we consider all the possible combinations of the information on both boundaries we have **16 possible cases**.

Table 7.1: Adjoint conditions for bounded domain CV problems

data		optimum	
x_j	$y(x_j)$	x_j^*	$y^*(x_j^*)$
fixed	fixed	x_j	y_j
fixed	free	x_j	$F_{y'}(x_j, y^*(x_j), y^{*'}(x_j)) = 0$
free	fixed	$F(x_j^*, y_j, y^{*'}(x_j^*)) - y^{*'}(x_j^*)F_{y'}(x_j^*, y_j, y^{*'}(x_j^*)) = 0$	y_j
free	free	$F\left(x_j^*, y^*(x_j^*), y^{*'}(x_j^*)\right) = 0$	$F_{y'}(x_j^*, y^*(x_j^*), y^{*'}(x_j^*)) = 0$

The index refers to the lower boundary when $j = 0$ and to the upper boundary when $j = 1$

7.2.5 Inequality terminal constraint

As for the static optimization problem we can consider inequality constraints, for instance inequality constraints on the value of the variable y for some value of the independent variable.

We consider a problem in which the two limits for independent variable are known, i.e, x_0 and x_1 are known, $y(x_0) = y_0$ is known, but we $y(x_1)$ is constrained by the condition $R(x_1, y(x_1)) \geq 0$.

Formally, the problem are :

$$\begin{aligned} &\max_{y(\cdot)} \int_{x_0}^{x_1} F(x, y(x), y'(x))\, dx \\ &\text{subject to} \\ &x_0 \text{ and } x_1 \text{ fixed} \\ &y(x_0) = y_0 \text{ fixed} \\ &R(x_1, y(x_1)) \geq 0. \end{aligned} \tag{P5}$$

Proposition 8. ***First order necessary conditions for the constrained terminal state problem:*** *$y^* \in \mathcal{Y}$ is the solution to the CV problem (P5) only if it satisfies the Euler-Lagrange equation* (7.2)*, the initial condition $y^*(x_0) = y_0$ and the boundary condition*

$$F_{y'}(x_1, y^*(x_1), y^{*'}(x_1))R(x_1, y^*(x_1)) = 0 \tag{7.19}$$

Proof. In this case we consider the functional we introduce a Lagrange multiplier (a real number) associated to the terminal condition, yielding the Lagrange functional

$$L[y] = \int_{x_0}^{x_1} F(x, y(x), y'(x))dx + \mu R(x_1, y(x_1)).$$

We proceed as previously to get the optimality conditions for a perturbation $\eta \in \mathcal{Y}$ over the optimal function y^*. The first order necessary condition is

$$\delta L[y^*](\eta) = \int_{x_0}^{x_1} \left(F_y^*(x) - \frac{d}{dx} F_{y'}^*(x) \right) \eta(x)dx + \left(F_{y'}^*(x_1) + \mu R_y^*(x_1) \right) \eta(x_1) = 0$$

where $F^*(x) = F(x, y^*(x), y^{*'}(x))$ and $R_y^*(x_1) = \partial_y R(x_1, y^*(x_1))$. Because of the free terminal state, admissible perturbations are such that $\eta(x_1) \neq 0$. Therefore $\delta L[y^*](\eta) = 0$ requires that the adjoint condition $F_{y'}^*(x_1) + \mu R_y^*(x_1) = 0$ holds.

Due to the existence of a static inequality constraint at the boundary x_1, the Karush-Kuhn-Tucker (KKT) complementarity slackness conditions are also necessary:

$$\mu R^*(x_1) = 0, \ \mu \geq 0 \text{ and } \ R^*(x_1) \geq 0$$

where $R^*(x_1) = R(x_1, y^*(x_1))$. Multiplying the adjoint condition by $R^*(x_1)$ we obtain an equivalent condition

$$R^*(x_1)\, F_{y'}^*(x_1) + \mu\, R^*(x_1)\, R_y^*(t_y) = 0,$$

which is equivalent to $F_{y'}^*(x_1) R^*(x_1) = 0$, after considering the KKT condition. Therefore $\delta L[y^*](\eta) = 0$ if the Euler-Lagrange equation (7.2) and adjoint boundary condition (7.19) hold. □

7.2.6 Applications

Spatial pricing by a monopolist

A monopolist firm, located at $x = 0$, can sell a unique product in quantity $y(x)$ at location which are distant from the production location by $x \in X = [0, \overline{x}]$. The output sold until distance x is $Y(x) = \int_0^x y(s)\, ds$, and the total output sold is $Y(\overline{x}) = \int_0^{\overline{x}} y(s)\, ds$. Then $Y'(x) = y(x)$. The problem is time-independent.

The firm faces a demand function $y = f(x, p)$, such that $y(x) = f(x, p(x))$.

We can assume that $y(x) = p(x)^{-\eta(x)}$ where $\eta(x)$ is the elasticity of demand which can be space-independent or space-dependent.

The firm faces two types of costs: (1) a transportation cost $\tau = \tau(x)$, where $\tau(0) = 0$ and $\tau'(x) > 0$, and (2) a production cost $c = C(Y(\overline{x})$ is a function of the total output which is sold.

We are assuming implicitly that the production takes place at site $x = 0$. The objective of the firm is to maximize profit, which is the functional

$$\pi[p] \; = \int_0^{\overline{x}} \Big(p(x) - \tau(x)\Big)\, f(x, p(x))\, dx - C(Y(\overline{x})),$$

where $Y(\overline{x}) = \int_0^{\overline{x}} y(p(x))\, dx$. If we can use the inverse function theorem, the inverse demand function $p(x) = P(x, y(x))$ can be recovered, allowing to write the profit functional as

$$\pi[y] \; = \int_0^{\overline{x}} \Big(P(x, y(x)) - \tau(x)\Big)\, y(x)\, dx - C(Y(\overline{x})).$$

The firm's problem can be cast as a calculus of variations problem

$$\max_{y(\cdot)} \int_0^{\overline{x}} \Big(P(x, Y'(x)) - \tau(x)\Big)\, Y'(x)\, dx - C(Y(\overline{x})).$$

$$\begin{aligned}
&\text{subject to}\\
&\overline{x} \;\; \text{fixed}\\
&Y(0) = 0, \text{for} \;\; x \in 0\\
&Y(\overline{x}) \;\; \text{free}
\end{aligned}$$

Assume that the demand function is $y(x) = p(x)^{-\eta(x)}$ where $0 < \eta(x) < 1$. The first order conditions are

$$\begin{aligned}
p(x) &= \mu(x)\big(\tau(x) + C'(Y(x))\big)\\
Y(0) &= 0\\
C'(Y(\overline{x}) &= C'(Y(x)), \;\text{for all} \;\; x \in \mathrm{X},
\end{aligned}$$

where the spatial markup is $\mu(x) = \frac{\eta(x)}{1-\eta x} > 1$. Then $\mu'(x) = \frac{\eta'(x)}{(1-\eta x)^2}$ which implies that $\mathrm{sign}(\mu'(x)) = \mathrm{sign}(\eta'(x))$.

Writing $C'(Y(\overline{x}) = C'(Y(x)) = -\overline{c}$ for any $x \in X$ then we obtain

$$p(x) = \mu(x)\; (\tau(x) + \overline{c}\;), \;\text{for}\; x \in X$$

that is, the price is determined from a spatial markup over the marginal cost of transportation and production. As $\tau'(x) \geq 0$ we have the possible cases: (1) price discrimination if $\mu'(x) \geq 0$ then $\partial_x p(x) > 0$ because $\mu(x) > 1$; (2) "dumping" that is $\partial_x p(x) < 0$ only if $\mu'(x) < 0$ which only occurs if $\eta'(x) < 0$, that is the demand elasticity decreases with the distance to the location of production.

Furthermore, as $Y'(x) = y(x) = p(x)^{-\eta(x)}$, and $Y(0) = 0$, and $Y(\overline{x}\;) = \Phi(\overline{c}) = C^{-1}(\overline{c})$ then

$$Y(x) \; = \int_0^x \; y(s)\, ds = \int_0^x \; p(s)^{-\eta(s)} \;\; ds = \int_0^x \; \Big(\mu(s)\; (\tau(s) + \overline{c}\;)\Big)^{-\eta(s)} \;\; ds$$

where

$$c^* \;\; = \Big\{\; z > 0 : \; C^{-1}(z) = \int_0^{\overline{x}} \; \Big(\mu(x)\; (\tau(x) + z)\Big)^{-\eta(x)} \;\; dx \Big\}.$$

As $\dfrac{\partial Y(\overline{x})}{\partial z} < 0$ then there is a unique c^* which is a function of the distribution of the transport costs and of the spatial demand schedule.

References: see Norman (1981), Fratrik (1982) and (Takayama, 1994, ch. 11.4).

Optimal taxation with asymmetric information

Mirrlees (1971) can be seen as a model which addresses the optimal taxation problem that we presented in section 6.4.7 but assumes that the government only observes the income distribution $y(\theta)$ but cannot observe the work effort $\ell(\theta)$, that is, he cannot separate productivity from effort in $y(\theta) = \theta\,\ell(\theta)$.

This poses an incentive problem, high skill people (with a high θ) may choose to reduce the work effort $\ell(\theta)$ such that the taxable income $y(\theta)$ is samaller. This, of course, may generate a general shift in the tax schedule to finance the government expenditure G. The solution by Mirrlees generate a new literature on mechanism design.

His solution was: the tax schedule must be drawn in a way such that people with higher skills still have an interest to have a high work effort. This has been called in the literature an **incentive compatible condition**. Using a variational approach (see also Diamond (1998) and Saez (2001)) the incentive compatible condition is a differential equation

$$\frac{du(\theta)}{d\theta} = -\frac{\ell(\theta)\,u_\ell\,(\theta)}{\theta}, \text{ for each } \theta \in \Theta$$

Using the notations in section 6.4.7 the optimal allocation problem becomes

$$\begin{aligned}
\max_{u(\cdot),\ell(\cdot)} \quad & \int_\Theta W\big(u(\theta)\big)\,f(\theta)\,d\theta \\
& \text{subject to} \\
& \frac{du(\theta)}{d\theta} = -\frac{\ell(\theta)\,u_\ell\,(\theta)}{\theta} \\
& \int_\Theta \Big(\theta\,\ell(\theta) - C\big(u(\theta),\ell(\theta)\big)\Big)\,f(\theta)\,d\theta = G \\
& u(\theta) \text{ free for } \theta \in \partial\Theta,
\end{aligned}$$

where $\partial\Theta$ denotes the boundary of Θ. The solution of this problem is easier if we use the Pontriyagin's maximum principle.

7.3 Calculus of variations in time

We can directly apply the previous results for problems in which time is the independent variable. When time is the independent variable the domain of the independent variable it $\mathrm{T} \subseteq \mathbb{R}_+$, if we have a finite interval $\mathrm{T} = [t_0, t_1]$, the dependent variable is $y(t)$, which is a mapping $y : \mathrm{T} \to \mathcal{Y} \subseteq \mathbb{R}$, and we denote the time derivative by $\dot{y} = \dfrac{dy(t)}{dt}$.

A particular important problem is the discounted infinite-horizon problem

7.3.1 Discounted infinite horizon

The most common problem in macroeconomics and growth theory has three main common features. First, time is the independent variable, and assumes that the initial time and values are known, usually $x_0 = 0$ and $y(0) = y_0$, and an unbounded value for the terminal time, $x_1 \to \infty$. Second, the objective function is of type $F(t, y, \dot{y}) = f(y, \dot{y})e^{-\rho t}$, where $e^{-\rho t}$ is a discount factor with a time-independent rate of discount $\rho \geq 0$, and the current value objective function $f(y, \dot{y})$ is time -independent. Third, there are two main versions to the problem depending on the terminal value of the state variable, that can be free or constrained.

Free asymptotic state

Find function $y^* \in \mathcal{Y}$ which is the set of functions $y : [0, \infty) \to \mathbb{R}$ such that $y(0) = y_0$ where y_0 are given that maximizes

$$\mathsf{J}[y] \equiv \int_0^\infty f(y(t), \dot{y}(t))e^{-\rho t}dt, \ \rho \geq 0. \tag{7.20}$$

This can be treated as a problem with a fixed initial time and value for the state variable, a fixed terminal time but a free terminal value for the state variable.

Proposition 9. ***First order necessary conditions for the discounted infinite horizon problem with free terminal state:*** *$y^* \in \mathcal{Y}$ is the solution to the discounted infinite horizon CV problem with a known initial data, $\varphi = (y_0, \rho, .)$, and with a free terminal state only if it satisfies the Euler-Lagrange equation*

$$\frac{d}{dt}\left(f_{\dot{y}}(y^*(t), \dot{y}^*(t))\right) = f_y(y^*(t), \dot{y}^*(t)) + \rho f_{\dot{y}}(y^*(t), \dot{y}^*(t)), \ \textit{for} \ \ t \in [0, \infty), \tag{7.21}$$

the so-called transversality condition

$$\lim_{t\to\infty} f_{\dot{y}}(y^*(t), \dot{y}^*(t))e^{-\rho t} = 0 \tag{7.22}$$

and the initial condition $y^*(0) = y_0$

Proof. In the proof for the free boundaries value problem we extend $x_1 \to \infty$ and take it as fixed but let $\lim_{t\to\infty} y^*(t)$ be free. In this discounted problem the Euler-Lagrange equation (7.2) $F^*_y(t) = \frac{d}{dt}F^*_{\dot{y}}(t)$ is equivalent to

$$e^{-\rho t}f_y(y^*, \dot{y}^*) = \frac{d}{dt}\left(e^{-\rho t}f_{\dot{y}}(y^*, \dot{y}^*)\right),$$

and the terminal condition (7.22) is obtained fro the boundary condition $\lim_{t\to\infty} F^*_{\dot{y}}(t) = 0$. □

Observe that the Euler-Lagrange is again a 2nd order non-linear autonomous ODE

$$f_y(y^*, \dot{y}^*) + \rho f_{\dot{y}}(y^*, \dot{y}^*) - f_{\dot{y}y}(y^*, \dot{y}^*)\dot{y} - f_{\dot{y}\dot{y}}(y^*, \dot{y}^*)\ddot{y} = 0.$$

The constrained terminal state problem

In several problems in economics the former condition can lead to an asymptotic state which does not make economic sense (v.g, a negative level for a capital stock).

The most common discounted infinite horizon model is usually the following: find function $y^* \in \mathcal{Y}$ which is the set of functions $y : [0, \infty) \to \mathbb{R}$ such that $y(0) = y_0$ and $\lim_{t\to\infty} R(t, y(t)) \geq 0$, where $x_0 = 0$ and $y(x_0) = y_0$ are given, that maximizes the objective functional (7.20)

Proposition 10. ***First order necessary conditions for the discounted infinite horizon problem with constrained terminal state**: $y^* \in \mathcal{Y}$ is the solution to the discounted infinite horizon CV problem with a known initial data, $(x_0, y(x_0)) = (0, y_0)$, and with a terminal state constrained by $\lim_{t\to\infty} R(t, y(t)) \geq 0$ only if it satisfies the Euler-Lagrange equation* (7.21)*, the initial condition $y^*(0) = y_0$, and the (so-called) transversality condition*

$$\lim_{t\to\infty} f_{\dot{y}}(y^*(t), \dot{y}^*(t))R(t, y^*(t))e^{-\rho t} = 0 \tag{7.23}$$

Exercise: prove this. Observe that as the terminal constraint is $\lim_{t\to\infty} y(t) \geq 0$ we have to introduce a Lagrange multiplier associated to the terminal time.

7.3.2 Applications

The resource depletion problem

Assume we have a resource W (v.g., a cake) with initial size W_0 and we want to consume it along period $[0, \bar{t}]$. If $C(t)$ denotes the consumption at time x we evaluate the consumption of the resource by the functional $\int_0^{\bar{t}} \ln(C(t))e^{-\rho t}dt$. Several properties: (1) we are impatient (we discount time at a rate $\rho > 0$); (2) the felicity at every point in time is only a function of the instantaneous consumption (preferences are inter temporally additive); (3) more consumption means more felicity but at a decreasing rate (the increase in utility for big bites is smaller than for small bites); and (4) there is no satiation (there is not a bite with a zero or negative marginal utility): consumption is always good.

Cake eating problem with the terminal state given CE problem: find $C^* = (C^*(t))_{0\leq t\leq \bar{t}}$ that

$$\max_C \int_0^{\bar{t}} \ln(C(t))e^{-\rho t}dt$$

subject to

$$\dot{W}(t) = -C(t), \text{ for } t \in [0, T]$$

given $W(0) = W_0$ and $W(\bar{t}) = 0$.

Formulated as a CV problem: find $W^* = (W^*(t))_{0\leq t\leq \bar{t}}$ such that

$$V(W_0, \bar{t}, \rho) = \max_W \mathsf{J}[W] = \max_W \int_0^{\bar{t}} \ln(-\dot{W}(t))e^{-\rho t}dt$$

given $W(0) = W_0$ and $W(\bar{t}) = 0$. The data of the problem is the vector of constants $\varphi = (0, \bar{t}, W_0, 0, \rho)$

The solution of the problem, $(W^*(t))_{t=0}^{\bar{t}}$, is obtained from

$$\begin{cases} \ddot{W}^* + \rho \dot{W}^* = 0, & 0 < t < \bar{t} \\ W^*(0) = W_0, & t = 0 \\ W^*(T) = 0, & t = T \end{cases}$$

The solution of the Euler equation is [2]

$$W(t) = W(0) - \frac{k}{\rho}\left(1 - e^{-\rho t}\right)$$

where k is an arbitrary constant. Using the adjoint conditions $W^*(\bar{t}) = 0$ and $W^*(0) = W_0$ we find the solution

$$W^*(t) = \frac{e^{-\rho t} - e^{-\rho \bar{t}}}{1 - e^{-\rho \bar{t}}} W_0, \text{ for } t \in [0, \bar{t}].$$

The value of the cake is

$$\begin{aligned} V(\varphi) = & \int_0^{\bar{t}} \ln\left(-\dot{W}^*(t)\right) e^{-\rho t} dt = \\ = & \frac{1}{\rho}\left[\left(1 + \ln\left(\frac{1 - e^{-\rho \bar{t}}}{\rho W_0}\left(e^{-\rho \bar{t}} - 1\right)\right)\right)\right] + \bar{t} e^{-\rho \bar{t}} \end{aligned}$$

if the consumer is rational this should be equal its reservation price for the cake. If $\rho = 0.01$ and the cake lasts for one week and the calorie content is $W_0 = 1000$ then the reservation price for should be $V(10, 0.01, 1/52) \approx 0.12$ per 100 calories.

Cake eating problem: infinite horizon If we assume an infinite horizon and the terminal condition $\lim_{t\to\infty} W(t) \geq 0$, meaning that we cannot have a negative level of resource asymptotically. The first order conditions are:

$$\begin{cases} \rho \dot{W}^*(t) + \ddot{W}^*(t) = 0 \\ W^*(0) = W_0 \\ -\lim_{t\to\infty} e^{-\rho t} \frac{W^*(t)}{\dot{W}^*(t)} = 0 \end{cases}$$

We already found

$$W(t) = W_0 - \frac{k}{\rho}\left(1 - e^{-\rho t}\right)$$

then

$$\dot{W}(t) = -k e^{-\rho t}$$

Solution (as $k = \rho W_0$)

$$W^*(t) = W_0 e^{-\rho t}, \; t \in \mathbb{R}_+$$

Again $\lim_{t\to\infty} W^*(t) = 0$.

[2]Hint: setting $z = \dot{W}$ we get a first-order ODE $\dot{z} = -\rho z$ with solution $\dot{z} = k e^{-\rho t}$. As $dW(t) = z(t)dt$, if we integrate we have $\int_{W(0)}^{W(t)} dW = \int_0^t z(s)ds = \int_0^t k e^{-\rho s} ds$.

The benchmark representative problem

The benchmark representative consumer problem in macroeconomics is to find optimal consumption and asset holdings (C, A) such that $C : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ and $A : \mathbb{R}_+ \rightarrow \mathbb{R}$ that maximize the value functional

$$U[C] = \int_0^\infty u(C(t))e^{-\rho t}dt$$

subject to the instantaneous budget constraint

$$\dot{A} = Y - C + rA$$

given $A(0) = A_0$ and the non-Ponzi game condition $\lim_{t\to} A(t)e^{-rt} \geq 0$. In the above equations Y and r denote, respectively the non-financial income and the interest rate, and are both positive. The following assumptions on utility are standard: $u(0) = 0$, $u^{'}(C) > 0$ and $u^{''}(C) < 0$.

The inverse of the elasticity of intertemporal substitution can be proved to be

$$\theta(C) = -\frac{u^{''}(C)C}{u^{'}(C)} > 0.$$

Assumption: the elasticity of intertemporal substitution $\theta(C) = \theta$ is constant and

$$\theta > \frac{r-\rho}{r} > 0.$$

We can transform it into a CV problem by observing that consumption is a function of the both wealth and savings, $\dot{A}$,

$$C = C(A, \dot{A}) \equiv Y + rA - \dot{A}.$$

Therefore, the problem becomes a CV problem with value functional

$$\mathsf{J}[A] = \int_0^\infty u\left(Y + rA(t) - \dot{A}(t)\right) e^{-\rho t}dt$$

where $f(A(t), \dot{A}(t)) = u\left(Y + rA(t) - \dot{A}(t)\right)$. The optimality conditions (which are necessary and sufficient in this case) are

$$\begin{cases} (r-\rho)u^{'}(C(A,\dot{A})) + (r\dot{A} - \ddot{A})u^{''}(C(A,\dot{A})) = 0 \\ A(0) = A_0 \\ -\lim_{t\to\infty} e^{-\rho t}u^{'}(C(A,\dot{A}))A(t) \ = 0 \end{cases}$$

Observing that $\dot{C} = r\dot{A} - \ddot{A}$ and using the definition of the inverse intertempopral elasticity of substitution we can transform the Euler equation into

$$\dot{C} = \gamma C, \text{ for } \gamma \equiv \frac{r-\rho}{\theta} > 0.$$

This allows us to find a general solution for optimal consumption

$$C(t) = C(0)\, e^{\gamma t},$$

where $C(0)$ is an arbitrary unknown admissible level for consumption, i.e., it should be non-negative. In order to find that value we use the transversality condition. But for this we need to determine admissible values for A. The asset dynamics is then governed by

$$\dot{A} = Y + rA - ke^{\gamma t}, \text{for } t > 0,\ A(0) = A_0,\ \text{for } t = 0$$

The solution to this initial value problem is

$$A(t) = -\frac{Y}{r} + \left(A_0 + \frac{Y}{r}\right)e^{rt} + \frac{C(0)}{r-\gamma}\left(e^{rt} - e^{\gamma t}\right).$$

With the previous assumption we have $r > \gamma$. As $u^{'}(C) = C^{-\theta}$ with an isoelastic utility function we find

$$\begin{aligned}
\lim_{t\to\infty} u^{'}(C(t))A(t)e^{-\rho t} &= \lim_{t\to\infty}\left(C(0)e^{\gamma t}\right)^{-\theta} A(t)e^{-\rho t} = \\
&= \lim_{t\to\infty} C(0)^{-\theta}e^{-rt}\left[-\frac{Y}{r} + \left(A_0 + \frac{Y}{r}\right)e^{rt} + \frac{C(0)}{r-\gamma}\left(e^{rt} - e^{\gamma t}\right)\right] = \\
&= \lim_{t\to\infty} C(0)^{-\theta}\left[A_0 + \frac{Y}{r} + \frac{C(0)}{r-\gamma} - \frac{C(0)}{r-\gamma}e^{(\gamma - r)t}\right] = \\
&= C(0)^{-\theta}\left[A_0 + \frac{Y}{r} + \frac{C(0)}{r-\gamma}\right] = 0
\end{aligned}$$

if and only if $C(0) = (r-\gamma)\left(A_0 + \frac{Y}{r}\right)$. Therefore the optimal consumption and asset holdings are

$$C^*(t) = (r-\gamma)\left(A_0 + \frac{Y}{r}\right)e^{\gamma t},\ t \in [0,\infty) \tag{7.24}$$

$$A^*(t) = -\frac{Y}{r} + \left(A_0 + \frac{Y}{r}\right)e^{\gamma t},\ t \in [0,\infty). \tag{7.25}$$

Observations: First, if we define human capital as the present value, at rate r, of the non-financial income

$$H(t) = \int_t^{\infty} Ye^{r(t-s)}ds$$

we find $H(0) = \frac{Y}{r}$. Therefore the solution is a linear function of the total capital, financial and non-financial

$$C^*(t) = (r-\gamma)(A_0 + H(0))e^{\gamma t},\ A^*(t) = -H(0) + (A_0 + H(0))e^{\gamma t}$$

Second, because $\gamma > 0$ then the asymptotic value of the optimal A becomes unbounded. However, it still satisfies that boundary condition $\lim_{t\to\infty} A^*(t)e^{-rt} = 0$ because, by assumption, $r > \gamma$. What matters is not the absolute level of A but its level in present-value terms.

7.4 Bibliography

- General references: Gel'fand and Fomin (1963)
- With applications to economics: Liberzon (2012)
- Other van Brunt (2010)
- History of the calculus of variations: Goldstine (1980)

Bibliography

Bucci, A. and Ushchev, P. (2016). Specialization vs Competition: An Anatomy of Increasing Returns to Scale. HSE Working papers WP BRP 134/EC/2016, National Research University Higher School of Economics.

Clarke, F. (2013). *Functional Analysis, Calculus of Variations and Optimal Control*. Graduate Texts in Mathematics. Springer, 2013 edition.

Cover, T. M. and Thomas, J. A. (2006). *Elements of Information Theory, Second Edition*. Wiley-Interscience, second edition.

Diamond, P. A. (1998). Optimal income taxation: An example with a U-shaped pattern of optimal marginal tax rates. *The American Economic Review*, 88(1):83–95.

Drabek, P. and Milota, J. (2013). *Methods of nonlinear analysis: applications to differential equations*. Birkhauser advanced texts. Birkhäuser, 1 edition.

Evans, G. C. (1924). The dynamics of monopoly. *The American Mathematical Monthly*, 31(2):77–83.

Fratrik, M. R. (1982). Spatial competition within an optimal control framework. Working Paper 60, Bureau of Economics, Federal Trade Comission, Washington, DC. USA.

Gel'fand, I. M. and Fomin, S. V. (1963). *Calculus of Variations*. Dover.

Goldstine, H. H. (1980). *A History of the Calculus of Variations from the 17th through the 19th Century*. Studies in the History of Mathematics and Physical Sciences 5. Springer-Verlag New York, 1 edition.

Judge, G. G. and Mittelhammer, R. C. (2012). *An Information Theoretic Approach to Econometrics*. Cambridge University Press.

Kremer, M. (1993). The O-ring theory of economic development. *The Quarterly Journal of Economics*, 108(3):551–575.

Liberzon, D. (2012). *Calculus of Variations and Optimal Control Theory: A Concise Introduction*. Princeton UP.

Mirrlees, J. A. (1971). An exploration in the theory of optimum income taxation. *Review of Economic Studies*, 38:175–208.

Norman, G. (1981). Spatial competition and spatial price discrimination. *The Review of Economic Studies*, 48(1):97–111.

Ok, E. A. (2007). *Real Analysis with Economic Applications*. Princeton University Press. tenho.

Parenti, M., Ushchev, P., and Thisse, J.-F. (2017). Toward a theory of monopolistic competition. *Journal of Economic Theory*, 167(C):86–115.

Pomini, M. (2018). Economic dynamics and the calculus of variations in the interwar period. *Journal of the History of Economic Thought*, 40(1):57–79.

Riesz, F. and Sz.-Nagy, B. (1955). *Functional Analysis*. Frederick Ungar Publishing Co., republished by dover (1990) edition.

Saez, E. (2001). Using elasticities to derive optimal income tax rates. *Review of Economic Studies*, 68:205–229.

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, 27:379–423, 623–656.

Siddiqi, A. H. (2018). *Functional analysis and applications*. Industrial and Applied Mathematics. Springer.

Takayama, A. (1994). *Analytical Methods in Economics*. Harvester Wheatsheaf.

van Brunt, B. (2010). *The Calculus of Variations*. Universitext. Springer.