Advanced Macroeconomics

Olaf Posch (oposch@econ.au.dk)

Spring 2010

Contents

Chapter 1

Basic mathematical tools

1.1 Topics in integration

Literature: Sydsæter et al. (2005, chap. 4), Wälde (2009, chap. 4.3)

The objective of this chapter is mainly to recall basic concepts on integration and differential equations and to serve as a reference for later applications.

1.1.1 Definitions

Definition 1.1.1 (Partial derivative) *Let $f = f(x_1, ..., x_n) = f(x)$ where $x \in \mathbb{R}^n$, then*

$$\frac{\partial}{\partial x_i} f = f_{x_i} \tag{1.1}$$

denotes the partial derivative, i.e., the derivative of $f(x)$ with respect to x_i if all the other variables are held constant.

Definition 1.1.2 (Total derivative) *Let $f = f(x_1, ..., x_n) = f(x)$ where $x \in \mathbb{R}^n$, then*

$$df = f_{x_1} dx_1 + f_{x_2} dx_2 + ... + f_{x_n} dx_n = \sum_{i=1}^{n} f_{x_i} dx_i \tag{1.2}$$

denotes the total derivative of $f(x)$.

Example 1.1.3 *Let $f(x_1, ..., x_n) = 0$. Suppose that x_3 to x_n are held constant. Collecting terms in (1.2), we obtain*

$$\frac{dx_1}{dx_2} = -\frac{f_{x_2}}{f_{x_1}}. \tag{1.3}$$

This is an example of the *implicit function theorem* (Sydsæter et al. 2005, Theorem 2.8.1), as the function $f(x_1, x_2, ..., x_n) = 0$ implicitly defines $x_2 = g(x_1, x_3, x_4, ..., x_n)$, and dx_2/dx_1 is the partial derivative of this implicit function with respect to x_1.

Definition 1.1.4 (Indefinite integral) *Let $f(x)$ be a continuous function. The indefinite integral of $f(x)$ is defined as any function $F(x)$ satisfying*

$$\int f(x)dx = F(x) \quad \text{where} \quad F'(x) \equiv \frac{d}{dx}F(x) = f(x). \tag{1.4}$$

The term d/dx often is referred to as the *differential operator*. This definition implies that there is a infinite number of integrals (or solutions). If $F(x)$ is an integral, then $F(x) + C$, where C is a constant, is an integral as well.

Definition 1.1.5 (Definite integral) *Let $f(x)$ be a continuous function. The definite integral of $f(x)$ is defined as any function $F(x)$ satisfying*

$$\int_a^b f(x)dx = \Big|_a^b F(x) = F(b) - F(a) \quad \text{where} \quad F'(x) = f(x) \quad \text{for all } x \text{ in } (a, b). \tag{1.5}$$

If $f(x) \geq 0$ in the interval $[a, b]$, then $\int_a^b f(x)dx$ is the area under the graph of f over $[a, b]$. Note the following implications of (1.5),

$$\frac{d}{dx}\int_a^x f(t)dt = f(x), \quad \frac{d}{dx}\int_x^b f(t)dt = -f(x), \tag{1.6}$$

because $\int_a^x f(t)dt = F(x) - F(a)$.

Definition 1.1.6 *Let $f(x)$ be a continuous function. If f is integrable over an infinite interval, and if the limit of the following expressions exists,*

$$\int_{-\infty}^b f(x)dx \equiv \lim_{a\to-\infty}\int_a^b f(x)dx, \tag{1.7}$$

$$\int_a^\infty f(x)dx \equiv \lim_{b\to\infty}\int_a^b f(x)dx, \tag{1.8}$$

$$\int_{-\infty}^\infty f(x)dx \equiv \lim_{a\to-\infty}\int_a^c f(x)dx + \lim_{b\to\infty}\int_c^b f(x)dx, \quad c \in \mathbb{R}, \tag{1.9}$$

we refer to $F(x)$ as defined in either (1.7) to (1.9) as the improper integral of $f(x)$.

Definition 1.1.7 *A function $f : \mathbb{R} \to \mathbb{R}$ is said to be of class C^k ($k = 1, 2, ...$) if all of its partial derivatives of order up to and including k exist and are continuous. Similarly, a* *transformation $f = (f_1, ..., f_m)$ from (a subset of) $\mathbb{R}^n \to \mathbb{R}^m$ is said to be of class C^k if each of its component functions $f_1, ..., f_m$ is C^k.*

1.1.2 Rules of transformation

Two useful ways to transform an integral involve integration by parts (1.10) and integration by substitution, or change of variable formula (1.11).

Proposition 1.1.8 *For two differentiable functions $f(x)$ and $g(x)$,*

$$\int f(x)g'(x)dx = f(x)g(x) - \int f'(x)g(x)dx. \tag{1.10}$$

Proof. Use the product rule and integrate. ■

Proposition 1.1.9 *For two differentiable functions $f(x)$ and $g(u)$ where $x = g(u)$,*

$$\int f(x)dx = \int f(g(u))g'(u)du. \tag{1.11}$$

Proof. Define $H(u)$ as the integral of $h(u) = f(g(u))g'(u)$ and apply the chain rule. ■

Exercise 1.1.10 *Show that for definite integrals integration by parts is*

$$\int_a^b f(x)g'(x)dx = \Big[f(x)g(x)\Big]_a^b - \int_a^b f'(x)g(x)dx, \tag{1.12}$$

and the formula for integration by substitution where $x = g(u)$ is

$$\int_a^b f(x)dx = \int_{u_1}^{u_2} f(g(u))g'(u)du, \quad g(u_1) = a, \; g(u_2) = b, \tag{1.13}$$

where $u_2 = g^{-1}(b)$ and $u_1 = g^{-1}(a)$.

1.1.3 Differentiation under the integral sign

Integrals appearing in economics often depend on parameters. In comparative static analysis, we compute the change of the value of the integral with respect to a change in the parameter. An important rule for computing the derivative of an integral is Leibniz's formula (Sydsæter et al. 2005, Theorem 4.2.1). Let $F(x)$ be a continuous function defined by

$$F(x) = \int_{a(x)}^{b(x)} f(x,t)dt, \tag{1.14}$$

where $a(x)$, $b(x)$ and $f(x,t)$ are differentiable. Then, the Leibniz formula gives the derivative of this function with respect to x as

$$\frac{d}{dx}F(x) = f(x,b(x))b'(x) - f(x,a(x))a'(x) + \int_{a(x)}^{b(x)} \frac{\partial f(x,t)}{\partial x} dt. \tag{1.15}$$

To prove this result, observe that F can be interpreted as a function of three variables,

$$F(x) = H(x, a(x), b(x)).$$

According to the chain rule, we obtain

$$F'(x) = H_x + H_a a'(x) + H_b b'(x),$$

where H_x is the partial derivative of h w.r.t. x with a and b as constants, $H_x = \int_a^b f_x(x,t)dt$ (Sydsæter et al. 2005, p.154). Moreover, according to (1.6), $H_b = f(x,b)$ and $H_a = -f(x,a)$. Inserting these results again gives the Leibniz formula.

Exercise 1.1.11 *The present discounted value of a continuous flow $f(t)$, $t \in [s,T]$, given the constant rate r is*

$$V(s,r) = \int_s^T f(t)e^{-(t-s)r}dt, \quad r \in \mathbb{R}_+.$$

Find $V_s(s,r)$ and $V_r(s,r)$ by means of Leibnitz's rule and interpret your results.

Exercise 1.1.12 *In a growth model, the total labor force reads*

$$N(t) = \int_{t-T(t)}^{t} n(u)e^{-(t-u)\delta}du, \quad \delta \in \mathbb{R}_+.$$

where $n(u)$ is the number of workers available for operating new equipment, δ its a constant depreciation rate, and $T(t)$ denotes the lifetime of equipment as governed by obsolescence. Compute the growth in working population, $\dot{N}(t)$, and interpret the result.

Exercise 1.1.13 *Find the integrals of the following problems*

1.

$$\int \frac{4x^3+3x^2}{x^4+x^3+1}dx \qquad \textit{(hint: integration by substitution)},$$

2.

$$\int_e^{e^2} \frac{1}{x \ln x}dx \qquad \textit{(hint: integration by substitution)},$$

3.

$$\int x \ln x dx \qquad \textit{(hint: integration by parts)},$$

4.

$$\int \frac{x-1}{x^2-x-2} dx \qquad \textit{(hint: integration by partial fractions)},$$

5.

$$\int \frac{x^3-2x}{x^2+2} dx \qquad \textit{(hint: polynomial long division)}.$$

1.2 Recap differential equations

Literature: Sydsæter et al. (2005, chap. 5, 6), Wälde (2009, chap. 4.1)

1.2.1 Definitions

Unlike ordinary algebraic equations, in a differential equation we are looking for a path or a function instead of a number. The equation includes one or more derivatives of the function. The following definitions give a more formal description.

Definition 1.2.1 (ODE) *An ordinary differential equation is an equation of a function and its derivatives $x', x'', ..., x^k$ and the exogenous variable t,*

$$F(t, x(t), \dot{x}(t), \ddot{x}(t), ...) = 0, \tag{1.16}$$

where k determines the order as long as we can explicitly solve for this variable.

Definition 1.2.2 (ODE system) *A system of first-order ODEs is of the type*

$$\frac{dx}{dt} \equiv \dot{x} = f(t, x(t)), \tag{1.17}$$

where $t \in [t_0, \infty)$ and the vector $x \in \mathbb{R}^n$. The function $f(\cdot)$ maps from $\mathbb{R}^{n+1}$ into $\mathbb{R}^n$.

Definition 1.2.3 (Linear ODE) *Suppose $f(\cdot)$ is a linear mapping, then for $n = 1$ equation (1.17) is a linear ODE,*

$$\dot{x} + a(t)x = b(t), \tag{1.18}$$

where $a(t)$ and $b(t)$ denote continuous functions of t and $x(t)$ is the unknown function.

Definition 1.2.4 (Separability) *Suppose $\dot{x} = F(t,x)$, where $F(t,x)$ can be written as*

$$\dot{x} = f(t)g(x). \tag{1.19}$$

We then refer to $F(t,x)$ as separable in t and x. If t is not explicitly present, the equation is called autonomous. Any autonomous equations is also separable.

It is important to learn to distinguish between separable and nonseparable equations, because separable equations are among those that can be solved in terms of integrals of known functions (for some examples see Sydsæter et al. 2005, chap. 5.3).

1.2.2 Separable and first-order differential equations

The following techniques are useful for many applications in economics. After obtaining an intuition of a *solution*, we quickly recap methods for solving separable equations, first-order linear differential equations and solution techniques via transformations.

Slope fields

Slope fields are especially useful to obtain a feeling for a solution. Consider a one dimensional first-order differential equation of the type $\dot{x} = f(t,x)$ as in (1.18). Drawing straight-line segments or vectors with slopes $f(t,x)$ through several points in the (t,x)-plane gives us a so-called *directional diagram* (or *slope field*).

Exercise 1.2.5 *Draw a direction diagram for the differential equation $\dot{x} = x + t$ and draw the integral curve through $(0,0)$ in the (t,x)-plane.*

Slope fields intuitively suggest that a solution of a differential equation in general is not unique, that means its solution are integral curves that can be made unique if further restrictions are applied. A restriction then forces a unique solution.

Example 1.2.6 (Isoquants of perfect substitutes) *Let $Y = aK + bL$ be a production function where $a, b > 0$. Y denotes output, and K, and L are inputs of capital stock and labor, respectively. An isoquant is defined by $\bar{Y} \equiv Y \in \mathbb{R}_+$. Observe that*

$$K = \bar{Y} - \frac{b}{a}L. \tag{1.20}$$

Differentiating with respect to L gives a differential equation of the form $dK/dL = -b/a$. The solution to this differential equation is given by the integral curves (1.20).

Figure 1.1: Isoquants of perfect substitutes

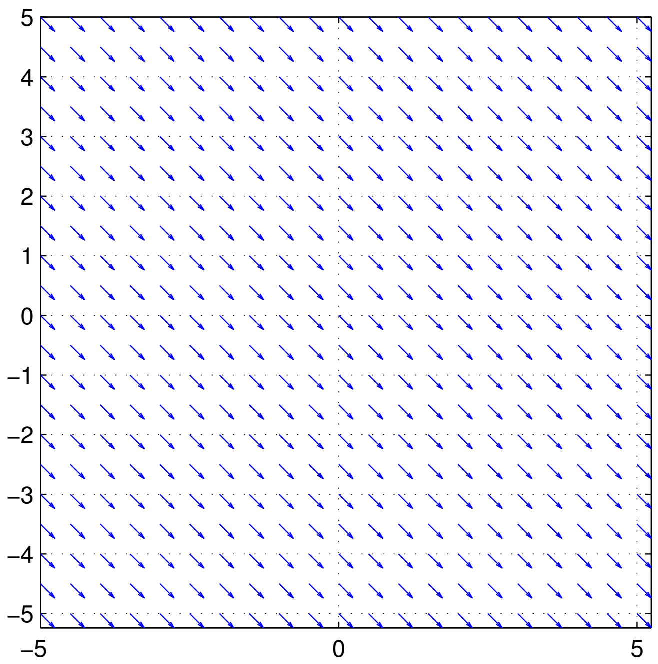
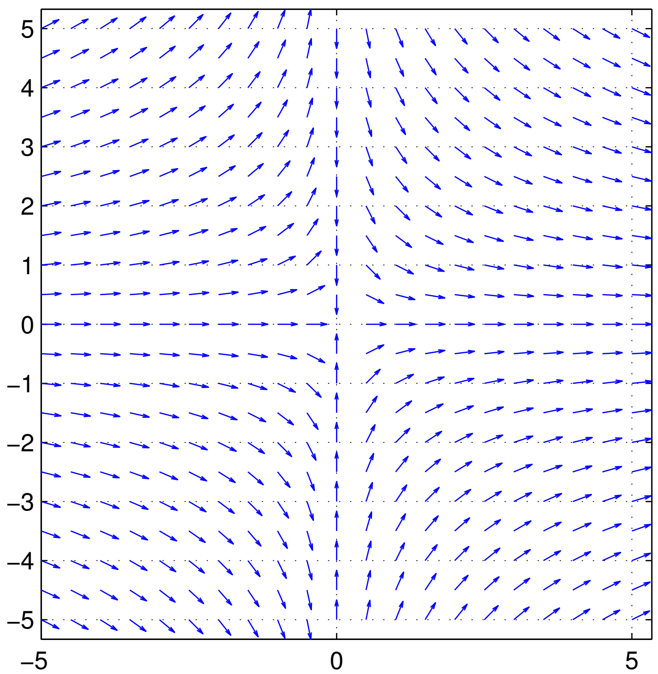


Figure 1.2: Isoquants in the Cobb-Douglas case



Example 1.2.7 (Isoquants in the Cobb-Douglas case) *Let $Y = K^\alpha L^{1-\alpha}$ describe a production function where $0 < \alpha < 1$. Y denotes output, and K, and L are inputs of capital stock and labor, respectively. An isoquant is defined by $\bar{Y} \equiv Y \in \mathbb{R}_+$. Observe that*

$$K = \bar{Y}^{\frac{1}{\alpha}} L^{\frac{\alpha-1}{\alpha}}. \tag{1.21}$$

Differentiating with respect to L gives a differential equation of the form

$$\frac{dK}{dL} = \frac{\alpha-1}{\alpha}(\bar{Y}/L)^{\frac{1}{\alpha}} = \frac{\alpha-1}{\alpha}K/L \quad \Leftrightarrow \quad K' = c_1 K/L, \quad c_1 \equiv \frac{\alpha-1}{\alpha}. \tag{1.22}$$

The solution to this differential equation is given by the integral curves (1.21).

Separable equations

Assume in the following a differential equation of the type (1.19), that is $\dot{x} = f(t)g(x)$. Note that this equation has to be homogeneous to be separable. The first solution technique simply is an application of the *integration by substitution*.

1. For $g(x) \neq 0$ divide by $g(x)$, multiply (1.19) by dt and integrate using $h(x) = 1/g(x)$

$$\int h(x)\dot{x}dt = \int f(t)dt =: F(t) + c_1.$$

2. According to the change of variable formula (1.11)

$$\int h(x)\dot{x}dt = \int h(x)dx =: H(x) + c_2.$$

3. The general solution to (1.19) is $H(x) = F(t) + C$, where $H'(x) = 1/g(x)$, $F'(t) = f(t)$, and C is a constant. If H is invertible, one can explicitly solve for x,

$$x(t) = H^{-1}(F(t) + C).$$

4. For $g(x) = 0$, we obtain the constant solution $x(t) = a$.

The second method is by *separating the variables* (Sydsæter et al. 2005, p.191),

1. For $g(x) \neq 0$ write (1.19) as

$$\frac{dx}{dt} = f(t)g(x).$$

2. Separate the variables
$$\frac{dx}{g(x)} = f(t)dt.$$

3. Integrate each side
$$\int \frac{dx}{g(x)} = \int f(t)dt.$$
Evaluate the two integrals and obtain a solution of (1.19), possibly in implicit form.

4. For $g(x) = 0$, we obtain the constant solution $x(t) = a$.

For illustration, let $F(t)$ be the integral of $f(t)$, and $g(x) = x$ then

$$H(x) = \int h(x)dx = \int \frac{dx}{g(x)} = \ln|x| + C = F(t) + C$$
$$\Leftrightarrow\ x = e^{F(t)+C} = c_1 e^{F(t)}, \quad c_1 = e^C$$

is the general solution to the differential equation of the type (1.19). If $H(x)$ is invertible,

$$x(t) = H^{-1}\left(F(t) + C\right), \quad C \in \mathbb{R}. \tag{1.23}$$

For the simplest case where $g(x) = x$, the general solution reads $x(t) = c_1 e^{F(t)}$.

Exercise 1.2.8 *Find the general solution to the differential equation $\dot{x} = 1/x$.*

Exercise 1.2.9 (Economic growth) *Let $Y_t \equiv Y(t)$ denote aggregate output, $K_t \equiv K(t)$ the capital stock, and $L_t \equiv L(t)$ the number of workers at time t. Suppose for all $t \geq 0$*

$$\begin{aligned} Y_t &= K_t^\alpha L_t^{1-\alpha}, \quad 0 < \alpha < 1, \\ \dot{K}_t &= sY_t, \quad K_0 > 0, \\ L_t &= L_0 e^{nt}, \end{aligned}$$

where α, s, K_0, L_0 and n are all positive constants. Determine the evolution of the capital stock given an initial level of capital stock K_0 and workers L_0.

First-order linear differential equations

Assume in the following a differential equation of the type (1.18). This equation is called linear because the left-hand side is a linear function of x and $\dot{x}$.

In the simplest case, we consider (1.18) with a and b as constants, where $a \neq 0$,

$$\dot{x} + ax = b. \tag{1.24}$$

To solve this equation, we multiply by the positive factor e^{at}, called an *integrating factor.* We then get the equation,

$$\dot{x}e^{at} + axe^{at} = be^{at}.$$

This turns out to be a good idea, since the left-hand side happens to be the derivative of the product xe^{at}. Thus (1.24) is equivalent to

$$\frac{d}{dt}(xe^{at}) = be^{at}.$$

Multiplying by dt and integrating both sides yields

$$xe^{at} = \int be^{at}dt + C = (b/a)e^{at} + C,$$

where C is a constant. Multiplying by e^{-at} gives the solution to (1.24) as,

$$x(t) = b/a + e^{-at}C. \tag{1.25}$$

Comparing this result to the general solution (1.23), it is notable that the solution to the inhomogeneous equation is the *general solution* of the associated homogeneous equation and a *particular solution* of the non-homogeneous equation.

Remark 1.2.10 *The set of solutions of a differential equation is called its general solution, while any specific function that satisfies the equation is called a particular solution.*

This solution technique using the integrating factor can be applied immediately also to the case where a is a constant, and $b(t)$ is time varying,

$$\dot{x} + ax = b(t) \quad \Rightarrow \quad x(t) = Ce^{-at} + e^{-at}\int e^{at}b(t)dt. \tag{1.26}$$

For the general case ($a \neq 0$) as in (1.18), the trick used for solving the equation has to be modified as follows. Multiply (1.18) by the integrating factor $e^{A(t)}$, to obtain

$$\dot{x}e^{A(t)} + a(t)xe^{A(t)} = b(t)e^{A(t)}.$$

We need to find an $A(t)$ such that the left-hand side of this equation equals the derivative of $xe^{A(t)}$. Note that the derivative is $\dot{x}e^{A(t)} + \dot{A}(t)xe^{A(t)}$. We therefore make $A(t)$ satisfy $\dot{A}(t) = a(t)$ by choosing $A(t) = \int a(t)dt$. This makes (1.18) equivalent to

$$\frac{d}{dt}\left(xe^{A(t)}\right) = b(t)e^{A(t)}.$$

Multiplying by dt and integrating gives

$$xe^{A(t)} = \int b(t)e^{A(t)}dt + C, \quad A(t) = \int a(t)dt.$$

Collecting terms gives the solution as

$$x(t) = e^{-\int a(t)dt}\left(C + \int e^{\int a(t)dt}b(t)dt\right). \tag{1.27}$$

A general approach of determining the solution of (1.18), especially useful for higher-order differential equations, is the method of *variation of parameters*. This method makes it possible always to find a particular solution provided the general solution of the associated homogeneous differential equation is known. Recall that any solution of (1.18) satisfies

$$x(t) = x^*(t) + z(t), \tag{1.28}$$

where $x^*(t)$ is a particular solution, and $z(t)$ is the general solution of the homogeneous differential equation associated with (1.18), $\dot{z} = -a(t)z$. Note that this equation clearly is time separable and we can use (1.23) to obtain the general solution as

$$z(t) = Ce^{-\int a(t)dt} \equiv Cv(t), \quad v(t) \equiv e^{-\int a(t)dt}.$$

The key step in using variation of parameters is to suppose that the particular solution reads

$$x^*(t) = u(t)v(t) = u(t)e^{-\int a(t)dt},$$

where $u(t)$ is an yet to be determined function.

Since this solution should be a particular solution of (1.18), we substitute $x^*(t)$ to obtain

$$\begin{aligned} \dot{u}e^{-\int a(t)dt} - a(t)u(t)e^{-\int a(t)dt} + a(t)u(t)e^{-\int a(t)dt} &= b(t) \\ \Leftrightarrow \dot{u} &= b(t)e^{\int a(t)dt}. \end{aligned}$$

We reduced the problem to a simple differential equation which has the solution

$$u(t) = \int b(t)e^{\int a(t)dt}dt + c_1 \quad \Rightarrow \quad x^*(t) = e^{-\int a(t)dt}\int b(t)e^{\int a(t)dt}dt,$$

where c_1 is a constant which can be neglected as we need only one particular solution. Thus, the general solution is given as derived using the integrating factor in (1.27), again

$$x(t) = e^{-\int a(t)dt}\left(C + \int e^{\int a(t)dt}b(t)dt\right). \tag{1.29}$$

Example 1.2.11 (Integral equation) *A differential equation always can be written as an integral equation. For illustration, consider the differential equation $\dot{x} = f(t,x)$. Separating terms and integrating both sides yields,*

$$\begin{aligned} dx &= f(t,x)dt \quad \Rightarrow \quad \int dx = \int f(t,x)dt \\ \Rightarrow \quad x(t) &= x_0 + \int f(t,x)dt. \end{aligned}$$

Differentiating with respect to time again using (1.15) yields $\dot{x} = f(t,x)$. Note this is yet another form of representation, but not necessarily a solution of the differential equation.

Example 1.2.12 (Intertemporal budget constraint) *Consider the dynamic budget constraint, $\dot{a}_t \equiv \dot{a}(t) = f(t,a_t) = r_t a_t + w_t - c_t$ given initial wealth $a_0 = a(0) \in \mathbb{R}$. Solving the differential equation using the integrating factor in (1.27) gives*

$$a_t = e^{\int_0^t r_s ds}\left(a_0 + \int_0^t e^{-\int_0^u r_s ds}(w_u - c_u)du\right).$$

We refer to $e^{\int_0^t r_s ds}$ as the discount factor. Collecting terms, we obtain an intuitive economic interpretation of the intertemporal budget constraint (backward solution),

$$e^{-\int_0^t r_s ds}a_t + \int_0^t e^{-\int_0^u r_s ds}c_u du \; = \; a_0 + \int_0^t e^{-\int_0^u r_s ds}w_u du.$$

The sum of the present value of individual wealth and the present value of future consumption expenditures at time t equal initial wealth and the present value of future income.

Exercise 1.2.13 (Growth at constant rate) *Let $P_t \equiv P(t)$ denote the size of population at time t which grows at constant growth rate n. Describe the law of motion for the population size and solve the associated differential equation. When does the population size double?*

Example 1.2.14 (Endogenous growth) *Let $Y_t = AK_t$ denote aggregate output where $A \in \mathbb{R}_+$ is total factor productivity, K_t is the capital stock. The market clearing condition demands $sY_t = \dot{K}_t$, thus $\dot{K}_t = sAK_t$. Separating the variables gives the solution as*

$$K_t = K_0 e^{sAt} \quad \Rightarrow \quad Y_t = K_0 A e^{sAt}.$$

Both variables the capital stock and aggregate output are growing exponentially at the same constant rate without imposing an exogenous source of growth.

Reducible differential equations

Only in very special cases, differential equation have solutions given by explicit formulas. However, transformation sometimes may convert an apparently complicated differential equation into one of a familiar type. A well known example is *Bernoulli's equation.*

An equation of the type

$$\dot{x} + a(t)x = b(t)x^r, \tag{1.30}$$

where the exponent r is a fixed real number, and where $a(t)$ and $b(t)$ are given continuous functions is called *Bernoulli's equation* or reducible differential equation. Note that if $r = 0$, the equation is linear, and if $r = 1$, it is separable, since $\dot{x} = (b(t) - a(t))x$. We now introduce the following solution technique. Let $x(t) > 0$ for all t, so that the power x^r is always well defined. Now divide by x^r and introduce the transformation

$$z = x^{1-r}. \tag{1.31}$$

Observe that $\dot{z} = (1 - r)x^{-r}\dot{x}$, and substituting into $\dot{x}x^{-r} + a(t)x^{1-r} = b(t)$ yields

$$\frac{1}{1-r}\dot{z} + a(t)z \;=\; b(t) \;\Leftrightarrow\; \dot{z} + (1-r)a(t)z \;=\; (1-r)b(t),$$

which is a linear differential equation for $z = z(t)$. Once $z(t)$ has been found, we simply use (1.31) to determine $x(t)$, which then is the solution of (1.30).

Exercise 1.2.15 *Solve the reducible differential equation $dP_t = cP_t(1 - \lambda P_t)dt$, known as Verhulst equation (growth with carrying capacity $1/\lambda$, logistic growth) where λc denotes the speed of reversion measuring how much the growth rate of P_t declines as P_t increases.*

Remark 1.2.16 *For an economic model to be consistent, the equations in that model must have a solution. If a solution does exist that satisfies the relevant boundary conditions, we are interested whether the solution is unique. Answers to such questions are provided by existence and uniqueness theorems (see Sydsæter et al. 2005, chap. 5.8).*

1.2.3 Second-order differential equations and systems in the plane

Many economic models are based on differential equations in which second- or higher-order derivatives appear. The following sections recap second-order differential equations and systems in the plane before we proceed with linear approximations of nonlinear systems.

Second-order linear differential equations

The general second-order linear differential equation is

$$\ddot{x} + a(t)\dot{x} + b(t)x = f(t), \tag{1.32}$$

where $a(t)$, $b(t)$, and $f(t)$ are all continuous functions of t. Recall that

$$\ddot{z} + a(t)\dot{z} + b(t)z = 0, \tag{1.33}$$

is the associated *homogeneous* equation of (1.32). The following theorem suggest that in order to find the paths $x(t)$ that solve (1.32), i.e. to find its general solution, we have to find the general solution to (1.33) and a particular solution to (1.32).

Theorem 1.2.17 (cf. Sydsæter et al. 2005, Theorem 6.2.1) *The general second-order linear differential equation (1.32),*

$$\ddot{x} + a(t)\dot{x} + b(t)x = f(t)$$

has the general solution

$$x(t) = x^*(t) + z(t),$$

where $x^*(t)$ is any particular solution of the nonhomogeneous equation. Further, the function $z(t)$ solves the associated homogeneous equation with the general solution,

$$z(t) = c_1 v_1(t) + c_2 v_2(t),$$

where $v_1(t)$ and $v_2(t)$ are two solutions that are not proportional, and c_1 and c_2 are constants.

For illustration, we consider finding a general solution to the homogeneous equation with constant coefficients,

$$\ddot{x} + a\dot{x} + bx = 0, \tag{1.34}$$

where a and b are arbitrary constants, and $x(t)$ is the unknown function. According to the theorem, finding the general solution of (1.34) requires to discover two solutions $v_1(t)$

and $v_2(t)$ that are not proportional. Recall that for first-order differential equations with constant coefficients, the general solution is $x(t) = e^{-at}C$. A possible solution therefore is

$$x = e^{\lambda t}, \quad \dot{x} = \lambda e^{\lambda t}, \quad \ddot{x} = \lambda^2 e^{\lambda t},$$

and try adjusting the constant λ in order that $x = e^{\lambda t}$ satisfies (1.34). Inserting gives

$$e^{\lambda t}\lambda^2 + ae^{\lambda t}\lambda + be^{\lambda t} = 0 \quad \Rightarrow \quad \lambda^2 + a\lambda + b = 0, \tag{1.35}$$

which is the *characteristic equation* of the differential equation (1.34). This is a quadratic equation whose two *characteristic roots* are obtained by solving the quadratic formula,

$$\lambda_{1,2} = -\tfrac{1}{2}a \pm \tfrac{1}{2}\sqrt{a^2 - 4b}.$$

There are three different cases to consider that are replicated in the following theorem.

Theorem 1.2.18 (cf. Sydsæter et al. 2005, Theorem 6.3.1) *The general solution of*

$$\ddot{x} + a\dot{x} + bx = 0 \quad \textit{is as follows,}$$

(i) if $a^2 - 4b > 0$*, there are two distinct real roots,*

$$x(t) = c_1 e^{\lambda_1 t} + c_2 e^{\lambda_2 t}, \quad \textit{where} \quad \lambda_{1,2} = -\tfrac{1}{2}a \pm \tfrac{1}{2}\sqrt{a^2 - 4b},$$

(ii) if $a^2 - 4b = 0$*, there is one real double root,*

$$x(t) = c_1 e^{\lambda_1 t} + tc_2 e^{\lambda_2 t}, \quad \textit{where} \quad \lambda_1 = \lambda_2 = -\tfrac{1}{2}a,$$

(iii) if $a^2 - 4b < 0$*, there are two conjugate complex roots,*

$$x(t) = e^{\alpha t}(c_1 \cos\beta t + c_2 \sin\beta t), \quad \textit{where} \quad \alpha = -\tfrac{1}{2}a, \quad \beta = \tfrac{1}{2}\sqrt{4b - a^2},$$

for any arbitrary constants $c_1, c_2 \in \mathbb{R}$*.*

Note that although the roots of the characteristic equation are complex for $a^2 < 4b$, we can obtain real-valued solutions in all three cases (cf. Sydsæter et al. 2005). Moreover, we only have explained how to obtain the general solution to (1.33) which solves the associated homogeneous equation. But how do we find a particular solution $x^*(t)$? In fact the *method of undetermined coefficients* works in many cases (Sydsæter et al. 2005, p.229).

- $f(t) = A$ is a constant

  In this case we check to see if (1.32) has a solution that is constant, $x^* = c$, $\dot{x}^* = \ddot{x}^* = 0$, so the equation reduces to $bc = A$. Hence, $c = A/b$ for $b \neq 0$.

- $f(t)$ is a polynomial of degree n

  A reasonable guess is that $x^*(t) = A_n t^n + A_{n-1} t^{n-1} + ... + A_1 t + A_0$ is also a polynomial of degree n. We then determine the undetermined coefficients $A_n, A_{n-1}, ..., A_0$ by requiring $x^*(t)$ to satisfy (1.32) and equating coefficients of like powers of t.

- $f(t) = pe^{qt}$

  A natural choice is $x^*(t) = Ae^{qt}$. Insert the guess and find that if $q^2 + aq + b \neq 0$, the particular solution is $x^*(t) = p/(q^2 + aq + b)e^{qt}$. If $q^2 + aq + b = 0$, and we either look for constants B or C such that Bte^{qt} or $Ct^2 e^{qt}$ is a solution.

- $f(t) = p \sin rt + q \cos rt$

  Let $x^*(t) = A \sin rt + B \cos rt$ and adjust constants A and B such that the coefficients of $\sin rt$ and $\cos rt$ match. If $f(t)$ is itself a solution of the homogeneous equation, then $x^*(t) = At \sin rt + Bt \cos rt$ is a particular solution for suitable choices of A and B.

Simultaneous equations in the plane

Many dynamic economic models, especially in macroeconomics, involve several unknown functions that satisfy a number of simultaneous differential equations. Consider the following system as a special case of the system in vector notation (1.17),

$$\begin{aligned} \dot{x} &= f(t, x, y), \\ \dot{y} &= f(t, x, y). \end{aligned} \tag{1.36}$$

A solution of (1.36) is a pair of differentiable functions $(x(t), y(t))$ satisfying both equations.

For illustration, consider the following system of linear differential equations,

$$\dot{x} = a_{11}x + a_{12}y + b_1(t), \tag{1.37}$$

$$\dot{y} = a_{21}x + a_{22}y + b_2(t). \tag{1.38}$$

Note that this two-dimensional system of first-order differential equations can be written as a one-dimensional second-order differential equation (and vice versa) as follows. Without loss of generality let $a_{12} \neq 0$ (note that either a_{12} or a_{21} has to be different from zero, otherwise

x and y are two isolated systems). Observe that using (1.37),

$$\begin{aligned} y &= \frac{1}{a_{12}}\dot{x} - \frac{a_{11}}{a_{12}}x - \frac{1}{a_{12}}b_1(t), \\ \dot{y} &= \frac{1}{a_{12}}\ddot{x} - \frac{a_{11}}{a_{12}}\dot{x} - \frac{1}{a_{12}}\dot{b}_1(t). \end{aligned}$$

Inserting into (1.38) and collecting terms yields

$$\begin{aligned} \frac{1}{a_{12}}\ddot{x} - \frac{a_{11}}{a_{12}}\dot{x} - \frac{1}{a_{12}}\dot{b}_1(t) &= a_{21}x + a_{22}\left(\frac{1}{a_{12}}\dot{x} - \frac{a_{11}}{a_{12}}x - \frac{1}{a_{12}}g(t)\right) + b_2(t), \\ \Leftrightarrow\ \ddot{x} - \underbrace{(a_{11}+a_{22})}_{\mathrm{tr}(A)}\dot{x} + \underbrace{(a_{11}a_{22} - a_{12}a_{21})}_{\det(A)}x &= \underbrace{-a_{22}b_1(t) + a_{12}b_2(t) + \dot{b}_1}_{b(t)}, \end{aligned}$$

where we could define a matrix A,

$$A \equiv \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix},$$

such that

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = A\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} b_1(t) \\ b_2(t) \end{pmatrix}$$

is written as the two-dimensional system (1.37) to (1.38) in matrix notation. Thus, we could solve the second-order differential equation in the usual way. Note that for recursive systems, where one of the two variables varies independently of the other, the solution techniques can simply be replaced by a step-by-step procedure.

An alternative approach would be the *method of undetermined coefficients* based on eigenvalues. With $b_1 = b_2 = 0$, system (1.37) and (1.38) reduces to the homogeneous system

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = A\begin{pmatrix} x \\ y \end{pmatrix}. \tag{1.39}$$

Using the approach $x = c_1e^{\lambda t}$ and $y = c_2e^{\lambda t}$, we obtain

$$\begin{pmatrix} \lambda c_1e^{\lambda t} \\ \lambda c_2e^{\lambda t} \end{pmatrix} = A\begin{pmatrix} c_1e^{\lambda t} \\ c_2e^{\lambda t} \end{pmatrix} \quad\Rightarrow\quad A\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \lambda\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \ \Leftrightarrow\ (A - \lambda I)\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = 0.$$

Observe that $(c_1, c_2)^\top$ is the associated eigenvector of the matrix A with eigenvalue λ. The eigenvalues are the solution of the equation

$$\begin{vmatrix} a_{11}-\lambda & a_{12} \\ a_{21} & a_{22}-\lambda \end{vmatrix} = \lambda^2 - (a_{11}+a_{22})\lambda + (a_{11}a_{22} - a_{12}a_{21}) = 0.$$

Recall that for eigenvalues λ_1 and λ_2 of a two-dimensional matrix the following holds,

$$\begin{array}{rrcl} & \lambda_1\lambda_2 & = & \det(A), \quad \lambda_1+\lambda_2 \;=\; \mathrm{tr}(A) \\ \Rightarrow & \lambda_{1,2} & = & \frac{1}{2}\mathrm{tr}(A) \pm \frac{1}{2}\sqrt{(\mathrm{tr}(A))^2 - 4\det(A)}, \end{array}$$

similar to the roots of the characteristic equation of second order. For the case in which A has different real eigenvalues λ_1 and λ_2 (for $(\mathrm{tr}(A))^2 > 4\det(A)$), then A has *two* linearly independent eigenvectors $(v_1, v_2)^\top$ and $(u_1, u_2)^\top$, and the general solution of (1.39) is

$$\begin{pmatrix} x \\ y \end{pmatrix} = d_1 e^{\lambda_1 t} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} + d_2 e^{\lambda_2 t} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \tag{1.40}$$

where d_1 and d_2 are arbitrary constants.

Exercise 1.2.19 *Solve the following system of equations,*

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = A \begin{pmatrix} x \\ y \end{pmatrix}, \quad A = \begin{pmatrix} 0 & 2 \\ 1 & 1 \end{pmatrix}.$$

Linear approximation of systems of differential equations

Suppose we have a nonlinear autonomous system of differential equations of the type

$$\begin{aligned} \dot{x} &= f(x,y), \\ \dot{y} &= g(x,y). \end{aligned} \tag{1.41}$$

Let (x^*, y^*) an *equilibrium point* (or an *equilibrium state*) for the system (1.41),

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} f(x^*, y^*) \\ g(x^*, y^*) \end{pmatrix}.$$

If (x, y) is sufficiently close to (x^*, y^*), then Taylor's formula gives as a linear approximations,

$$\begin{aligned} f(x,y) &\approx f(x^*,y^*) + f_x(x^*,y^*)(x-x^*) + f_y(x^*,y^*)(y-y^*), \\ g(x,y) &\approx g(x^*,y^*) + g_x(x^*,y^*)(x-x^*) + g_y(x^*,y^*)(y-y^*). \end{aligned}$$

Because $f(x^*, y^*) = g(x^*, y^*) = 0$,

$$\begin{aligned} f(x,y) &\approx f_x(x^*,y^*)x + f_y(x^*,y^*)y - f_x x^* - f_y y^* \\ &= a_{11}x + a_{12}y - b_1, \\ g(x,y) &\approx g_x(x^*,y^*)(x - x^*) + g_y(x^*,y^*)(y - y^*) \\ &= a_{21}x + a_{22}y + b_2. \end{aligned}$$

It is therefore reasonable to expect that in a neighborhood of (x^*, y^*), the nonlinear system (1.41) behaves approximately like the linear system

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = A \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}. \tag{1.42}$$

A is the Jacobian matrix of the nonlinear system around the equilibrium state. Note that b_1 and b_2 imply that (x^*, y^*) is also an equilibrium point in the linearized system.

Remark 1.2.20 *Note that the local dynamics of the nonlinear system can be analyzed using the linear approximation as long as A does contain eigenvalues with their real parts different from zero (Theorem of Hartman-Grobman).*

1.3 Qualitative theory

Literature: Sydsæter et al. (2005, chap. 5.7, 6.4 to 6.9), Wälde (2009, chap. 4.2)

Most kinds of economic models involve differential equations do not have the nice property that their solutions can be expressed in terms of elementary functions. Nevertheless, often it is desirable to analyze at least the qualitative behavior of the economic model. The following sections recap the elementary tools to analyze qualitative properties of differential equations.

1.3.1 Definitions

To shed light on the structure of solutions to differential equations that are not explicitly available, we shall introduce the following definitions. For simplicity, we restrict our attention to the two-dimensional case. However, extending the methods for multi-dimensional systems is straightforward, but involves notationally cumbersome derivations.

Definition 1.3.1 *Consider an autonomous nonlinear system of differential equations,*

$$\begin{aligned} \dot{x} &= f(x,y), \\ \dot{y} &= g(x,y). \end{aligned} \tag{1.43}$$

An equilibrium point (x^*, y^*) of the system (1.43) is called locally asymptotically stable, if any path starting near (x^*, y^*) tends to (x^*, y^*) as $t \to \infty$, or simply

$$\lim_{t\to\infty} x = x^*, \quad \lim_{t\to\infty} y = y^*.$$

An equilibrium point (x^*, y^*) of this system is called globally asymptotically stable, if any path (wherever it starts) converges to (x^*, y^*) as $t \to \infty$.

For simplicity, we consider the case of an autonomous linear system as in (1.42) which can actually be thought of as an linearized system around an equilibrium point.

Theorem 1.3.2 (cf. Sydsæter et al. (2005), Theorem 6.6.1) *Suppose that $|A| \neq 0$. Then the equilibrium point (x^*, y^*) for the linear system*

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = A \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

is globally asymptotically stable if and only if

$$\mathrm{tr}(A) = a_{11} + a_{22} < 0, \quad \det(A) = a_{11}a_{22} - a_{12}a_{21} > 0,$$

or equivalently, if and only if all eigenvalues of A have negative real parts.

An intuitive explanation can be obtained from the solution (1.40), where each of the eigenvalues has to be negative in order to ensure *global asymptotic stability*. If the equilibrium point is not necessarily globally asymptotically stable, disregarding the case where one or both eigenvalues are 0, the dynamic behavior quickly can be categorized as follows.

1. If both eigenvalues of A have negative real parts, then (x^*, y^*) is globally asymptotically stable (a *sink*). All solutions converge to the equilibrium point as $t \to \infty$.

2. If both eigenvalues of A have positive real parts, then all solutions starting away from (x^*, y^*) explode as t increases, and the equilibrium point is a *source*.

3. If the eigenvalues of A are real with opposite signs, in other words if the determinant is negative, $\lambda_1 < 0$ and $\lambda_2 > 0$, then (x^*, y^*) is a *saddle point*. Solutions are either diverging from or converging to the equilibrium point as $t \to \infty$.

4. If the eigenvalues are purely imaginary, i.e. complex eigenvalues with zero real parts but nonzero imaginary parts, then (x^*, y^*) is a *centre*. All solution curves are periodic with the same period in the form of ellipses or circles.

Example 1.3.3 *Consider the system* $\dot{x} = 2y$ *and* $\dot{y} = 3x - y$ *with equilibrium point* $(0, 0)$. *Using* $f(x, y) = 2y$ *and* $g(x, y) = 3x - y$, *the coefficient matrix* A *is*

$$A = \begin{pmatrix} 0 & 2 \\ 3 & -1 \end{pmatrix}.$$

Because $\det(A) = -6$, *i.e. the eigenvalues are real with opposite signs, the equilibrium point* (x^*, y^*) *is a* saddle point. *The associated eigenvector from the negative eigenvalue is* $(u_1, u_2)^\top = (-2, 3)^\top$ *which points in the direction of the stable path.*

1.3.2 Autonomous equations, phase diagrams and stability

Many differential equations are of the type or can be expressed in the autonomous form

$$\dot{x} = F(x(t)), \tag{1.44}$$

which is a special case of the equation $F(t, x(t))$ where $F_t = 0$. We refer to equation (1.44) as *autonomous*. To examine the properties of the solution to (1.44), it is useful to study its *phase diagram*. This is obtained by plotting $\dot{x}$ against x in the $x\dot{x}$-plane.

An important property of a differential equation is whether it has any *equilibrium* or *steady states*. Moreover, it is also very useful to know whether an equilibrium state is *stable*. Once we have obtained an understanding about the dynamics in a phase diagram, often the dynamics of the corresponding directional diagram (in the tx-plane) are straightforward.

Suppose that a is an equilibrium state for $\dot{x} = F(x)$, so that $F(a) = 0$. If the slope of $\dot{x}$ at the equilibrium state is negative, $F'(a) < 0$, then $F(x)$ is positive to the left of $x = a$ and negative to the right in a close neighborhood. Hence, we can derive the following results,

$$\begin{aligned} F(a) &= 0 \quad \text{and} \quad F'(a) < 0 \;\Rightarrow\; a \text{ is locally asymptotically stable}, \\ F(a) &= 0 \quad \text{and} \quad F'(a) > 0 \;\Rightarrow\; a \text{ is unstable}. \end{aligned}$$

Example 1.3.4 (Price adjustment mechanism) *Suppose price changes are determined by a function of excess demand,* $D(p_t) - S(p_t)$ *where* $D(p_t)$ *and* $S(p_t)$ *are aggregate demand and supply, respectively, satisfying the nonlinear equation*

$$\dot{p}_t = F(p_t) = H(D(p_t) - S(p_t)).$$

Assume that the function H *satisfies* $H(0) = 0$ *and* $H' > 0$. *If demand exceeds supply at price* p_t, *then* $D(p_t) - S(p_t) > 0$, *so* $\dot{p}_t > 0$, *and the price increases (and vice versa). The* *equilibrium price* p^* *equalizes supply and demand such that* $H(D(p^*) - S(p^*)) = 0$. *Observe that* $F'(p_t) = H'([D'(p_t) - S'(p_t)])$, *and* p^* *is locally asymptotically stable, if* $D'(p_t) < S'(p_t)$, *which holds for the plausible assumptions* $D'(p_t) < 0$ *and* $S'(p_t) > 0$.

Exercise 1.3.5 *Illustrate the phase diagram of the reduced form for Solow's growth model,*

$$\dot{k} = sf(k) - \delta k, \quad f(0) = 0, \; f'(k) > 0, \; f''(k) < 0 \; \forall \; k > 0,$$

where $s > 0$ *and* $\delta > 0$ *denote a constant rate of saving and depreciation respectively, in the* $k\dot{k}$*-plane. Analyze the stability properties in this model.*

1.3.3 Phase plane analysis and stability for nonlinear systems

Even when explicit solutions are unavailable, geometric arguments can still shed light on the structure of the solutions of autonomous systems of differential equations in the plane,

$$\dot{x} \;=\; f(x, y), \quad \dot{y} \;=\; g(x, y). \tag{1.45}$$

A solution $(x(t), y(t))$ describes a curve or path in the xy-plane. For autonomous problems, if $(x(t), y(t))$ is a solution, then $(x(t + a), y(t + a))$ is a solution and both solutions have the same path. Note that $(\dot{x}, \dot{y})$ is uniquely determined at the point (x, y) and two paths in the xy-plane cannot intersect. The *phase plane analysis* is concerned with the technique of studying the behavior of paths in the phase plane.

Vector fields

It follows from (1.45) that the rates of change of $x(t)$ and $y(t)$ are given by $f(x, y)$ and $g(x, y)$, respectively. In particular if $f(x, y) > 0$ and $g(x, y) > 0$, then as t increases, the system will move from a point P in the xy-plane up and to the right. In fact, the direction of motion is given by the tangent vector $(\dot{x}(t), \dot{y}(t))$ at P, while the speed of motion is given by the length of that vector. A family of such vectors, which in practice is only a small representative sample, is called a *vector field*. On the basis of the vector field one can draw paths for the system and thereby the phase diagram of the system.

In general, a point (a, b) where $f(a, b) = g(a, b) = 0$ is called an *equilibrium* for system (1.45). The equilibrium points are the points of intersection of the two curves $f(x, y) = 0$ and $g(x, y) = 0$, which are called the *nullclines* of the system. To draw a phase diagram, we begin by drawing the two nullclines. At each point on the nullcline $f(x, y) = 0$ the $\dot{x}$ component is 0, and the velocity vector is vertical. It points up if $\dot{y} > 0$, and down if $\dot{y} < 0$.

Figure 1.3: Equilibrium point is a *sink*

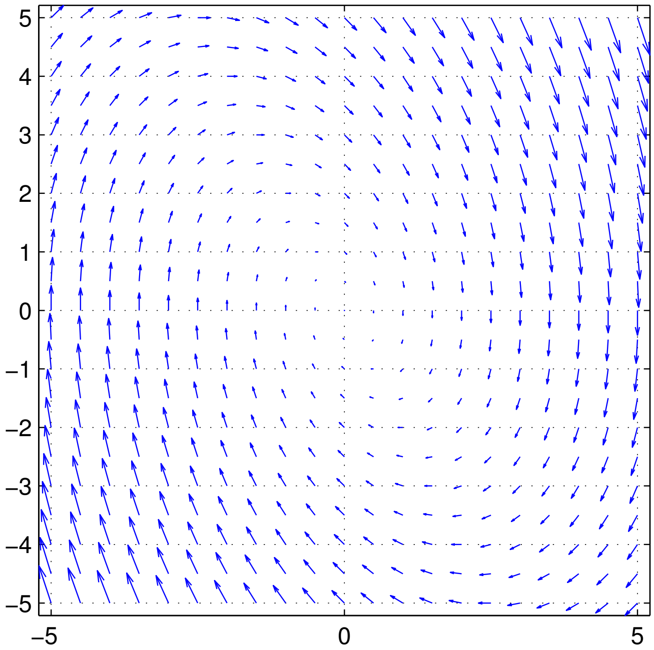


Figure 1.4: Equilibrium point is a *source*

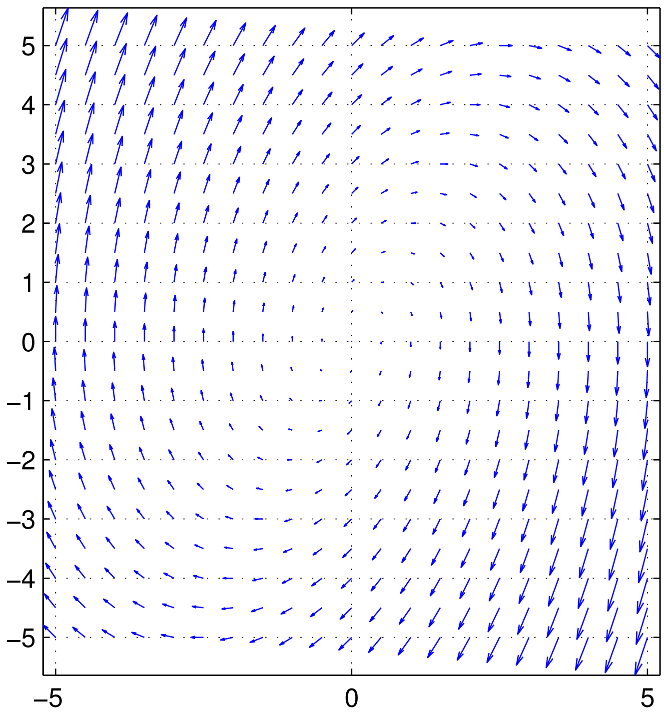


Figure 1.5: Equilibrium point is a *saddle point*

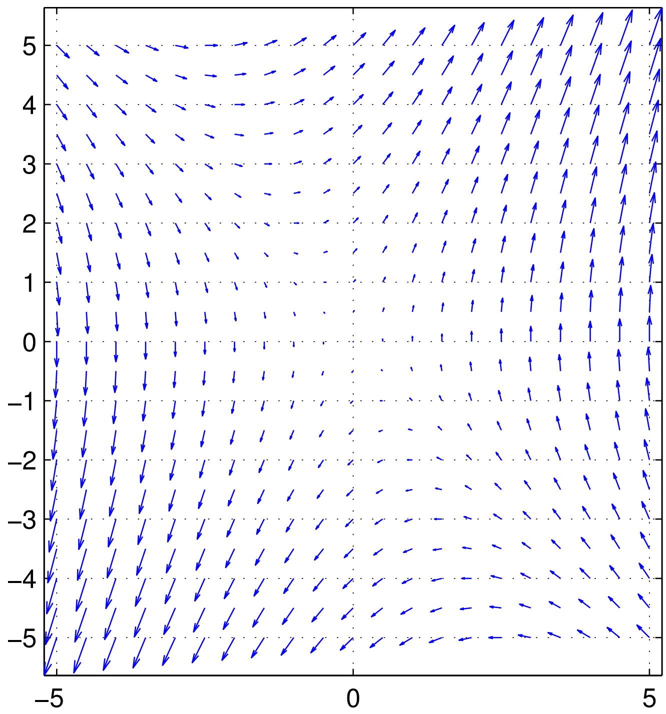
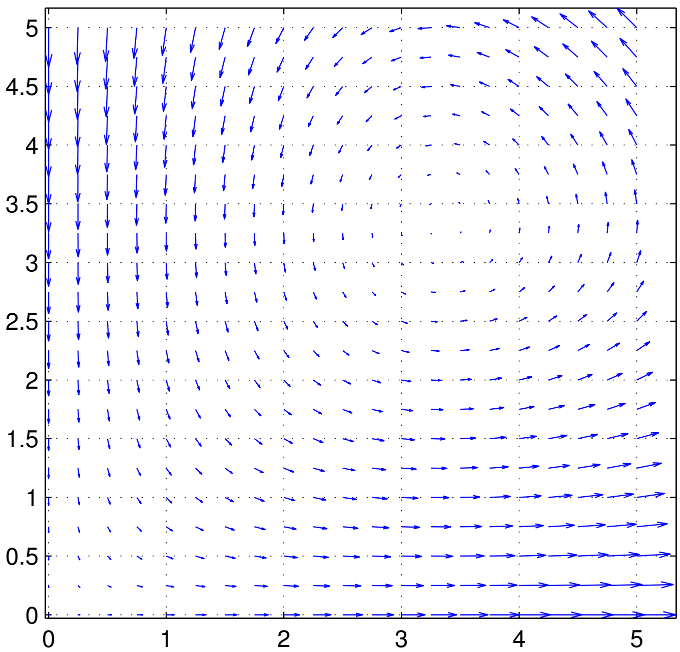


Figure 1.6: Equilibrium point is a *centre*



Similarly, at each point on the nullcline $g(x, y) = 0$. the $\dot{y}$ component is 0, and the velocity vector is horizontal. It points to the right if $\dot{x} > 0$, to the left if $\dot{x} < 0$.

Exercise 1.3.6 *Draw a vector field for the following system describing a model of economic growth in the (k, c)-plane assuming that $k \geq 0$ and $c \geq 0$. Suppose the capital stock, $k(t)$, and consumption, $c(t)$, satisfy the pair of differential equations*

$$\begin{aligned} \dot{k} &= 2.5k - 0.5k^2 - c, \\ \dot{c} &= (0.625 - 0.25k)c. \end{aligned}$$

Infer a phase diagram with nullclines and divide the phase diagram into appropriate regions.

Important properties about the solutions are obtained by partitioning the phase plane into regions where we know the direction of increase or decrease of each variable. In particular, the partition will often indicate whether or not a certain equilibrium point is stable, in the sense that paths starting near the equilibrium point tend to that point as $t \to \infty$.

The Lyapunov theorem (Theorem 1.3.7) now provides us with a tool to analyze *local* stability for a nonlinear system of equations. The intuition behind this theorem is again Hartman-Grobman which can be applied to systems where A has eigenvalues with their real parts *different* from zero. If we impose stronger conditions, one may be able to prove *global* stability using Olech's Theorem (Sydsæter et al. 2005, Theorems 6.8.1 and 6.8.2).

Theorem 1.3.7 (Lyapunov) *Suppose that f and g are C^1 functions (all partial derivatives up to $k = 1$ exist and are continuous) and let (a, b) be an equilibrium point for the system,*

$$\dot{x} = f(x, y), \quad \dot{y} = g(x, y).$$

Let A be the Jacobian matrix,

$$A = \begin{pmatrix} f_x(a,b) & f_y(a,b) \\ g_x(a,b) & g_y(a,b) \end{pmatrix}.$$

If $\mathrm{tr}(A) = f_x(a, b) + g_y(a, b) < 0$, and $\det(A) = f_x(a, b)g_y(a, b) - f_y(a, b)g_x(a, b) > 0$, or if both eigenvalues of A have negative real parts, then (a, b) is locally asymptotic stable.

Theorem 1.3.8 (Olech) *Consider a system, where f and g are C^1 functions in $\mathbb{R}^2$,*

$$\dot{x} = f(x, y), \quad \dot{y} = g(x, y),$$

and let (a, b) be an equilibrium point. Let A be the Jacobian matrix,

$$A(x, y) = \begin{pmatrix} f_x(x, y) & f_y(x, y) \\ g_x(x, y) & g_y(x, y) \end{pmatrix}.$$

Assume that the following three conditions are all satisfied:

$$\begin{array}{ll} (a) & \mathrm{tr}(A(x, y)) = f_x(x, y) + g_y(x, y) < 0 \quad \text{in all of } \mathbb{R}^2, \\ (b) & \det(A(x, y)) = f_x(x, y)g_y(x, y) - f_y(x, y)g_x(x, y) > 0 \quad \text{in all of } \mathbb{R}^2, \\ (c) & f_x(x, y)g_y(x, y) \neq 0 \quad \text{in all of } \mathbb{R}^2 \quad \text{or} \quad f_y(x, y)g_x(x, y) \neq 0 \quad \text{in all of } \mathbb{R}^2. \end{array}$$

Then (a, b) is globally asymptotic stable.

1.4 Calculus of variation

Literature: Kamien and Schwartz (1991, part 1), Sydsæter et al. (2005, chap. 8)
The calculus of variation has a long history (Euler, Lagrange in the 18th century). In economics, some first applications were by Ramsey (1928) to an optimal savings problem, and by Hotelling (1931) to a problem of how to extract a natural resource.

1.4.1 Euler equation

The simplest problem in the calculus of variation takes the form

$$\max \int_{t_0}^{t_1} F(t, x, \dot{x})dt \quad \text{subject to} \quad x(t_0) = x_0, \quad x(t_1) = x_1. \tag{1.46}$$

Here, F is a given *well behaved* function of three variables, whereas t_0 and t_1, as well as x_0 and x_1 are given numbers. Among all well behaved functions $x(t)$ that satisfy $x(t_0) = x_0$ and $x(t_1) = x_1$, find one making the integral (1.46) as large as possible.

Already in 1744, Euler proved that a function $x(t)$ can only solve problem (1.46), if $x(t)$ satisfies the differential equation,

$$\frac{\partial F}{\partial x} - \frac{d}{dt}\left(\frac{\partial F}{\partial \dot{x}}\right) = 0, \tag{1.47}$$

called the *Euler equation*. Replacing F with $-F$ does not change the condition. Hence, the equation is a necessary condition also for solving the corresponding minimization problem.

Note that the term $(d/dt)(\partial F(t, x, \dot{x})/\partial x)$ denotes the *total derivative* of $\partial F(t, x, \dot{x})$ with respect to t. According to the chain rule,

$$\frac{d}{dt}\left(\frac{\partial F}{\partial \dot{x}}\right) = \frac{\partial^2 F(t,x,\dot{x})}{\partial t \partial \dot{x}} + \frac{\partial^2 F(t,x,\dot{x})}{\partial x \partial \dot{x}}\dot{x} + \frac{\partial^2 F(t,x,\dot{x})}{(\partial \dot{x})^2}\ddot{x}.$$

Inserting this into the Euler equation (1.47) and rearranging yields

$$\begin{aligned} \frac{\partial^2 F(t,x,\dot{x})}{\partial t \partial \dot{x}} + \frac{\partial^2 F(t,x,\dot{x})}{\partial x \partial \dot{x}}\dot{x} + \frac{\partial^2 F(t,x,\dot{x})}{(\partial \dot{x})^2}\ddot{x} - \frac{\partial F}{\partial x} = 0 \\ \Leftrightarrow \ F_{\dot{x}\dot{x}}\ddot{x} + F_{x\dot{x}}\dot{x} + F_{t\dot{x}} - F_x = 0, \end{aligned}$$

where $F_{\dot{x}\dot{x}} = \partial^2 F(t,x,\dot{x})/(\partial \dot{x})^2$, $F_{x\dot{x}} = \partial^2 F(t,x,\dot{x})/(\partial x \partial \dot{x})$, $F_{t\dot{x}} = \partial^2 F(t,x,\dot{x})/(\partial t \partial \dot{x})$, and $F_x = \partial F/(\partial x)$. Thus, for $F_{\dot{x}\dot{x}} \neq 0$, the Euler equation is a differential equation of second order, typically does not have an explicit solution. It gives a necessary condition for optimality, but in general is not sufficient. By analogy with static optimization problems, if $F(t,x,\dot{x})$ is concave (convex) in $(x,\dot{x})$, an admissible solution that satisfies the Euler equation solves the maximization (minimization) problem and ensures optimality.

The first known application of the calculus of variation was the *brachistochrone* problem (*brachistos* - shortest, *chronos* - time). Given two points A and B in a vertical plane, the time required for a particle to slide along a curve under the sole influence of gravity will depend on the shape of the curve. Along which curve does the particle go from A to B as quick as possible? The following formulation is from Padra 2006, The Beginnings of variational calculus, and its early relation with numerical methods, *Variational Formulations in Mechanics: Theory and Applications.*

Example 1.4.1 (Brachistochrone) *Consider the following variational problem where a particle slides from x_0 to x_1 under the sole influence of gravity (at gravitational constant g).*

$$I(x) = \min \int_{t_0}^{t_1} \sqrt{\frac{1+\dot{x}^2}{2gx}}dt \quad s.t. \quad x(t_0) = x_0, \quad x(t_1) = x_1.$$

What is the curve traced out by a particle that reaches x_1 in the shortest time? Observe that

$$\begin{aligned} \frac{\partial F}{\partial x} &= -\frac{1}{2}\sqrt{\frac{1+\dot{x}^2}{2gx^3}}, \quad \frac{\partial F}{\partial \dot{x}} = \frac{\dot{x}}{\sqrt{(1+\dot{x}^2)2gx}}, \\ \frac{\partial^2 F}{\partial t \partial \dot{x}} &= 0, \quad \frac{\partial^2 F}{\partial x \partial \dot{x}} = -\frac{1}{2}\frac{\dot{x}}{\sqrt{(1+\dot{x}^2)2gx^3}}, \quad \frac{\partial^2 F}{(\partial \dot{x})^2} = \frac{\sqrt{(1+\dot{x}^2)} - \dot{x}^2(1+\dot{x}^2)^{-1/2}}{(1+\dot{x}^2)\sqrt{2gx}}. \end{aligned}$$

Using the Euler equation (1.47), a necessary condition is

$$\begin{aligned}
\frac{\sqrt{(1+\dot{x}^2)} - \dot{x}^2(1+\dot{x}^2)^{-1/2}}{(1+\dot{x}^2)\sqrt{2gx}}\ddot{x} - \frac{1}{2}\frac{\dot{x}}{\sqrt{(1+\dot{x}^2)2gx^3}}\dot{x} + \frac{1}{2}\frac{\sqrt{1+\dot{x}^2}}{\sqrt{2gx^3}} &= 0\\
\Leftrightarrow \quad \frac{\sqrt{(1+\dot{x}^2)} - \dot{x}^2(1+\dot{x}^2)^{-1/2}}{(1+\dot{x}^2)}\ddot{x} + \frac{1}{2}\frac{1}{\sqrt{(1+\dot{x}^2)x}} &= 0\\
\Leftrightarrow \quad \frac{1-\dot{x}^2(1+\dot{x}^2)^{-1}}{\sqrt{(1+\dot{x}^2)}}\ddot{x} + \frac{1}{2}\frac{1}{\sqrt{(1+\dot{x}^2)x}} &= 0\\
\Leftrightarrow \quad \ddot{x}x + \frac{1}{2}(1+\dot{x}^2) &= 0.
\end{aligned}$$

Any curve that follows the Euler equation is a solution candidate. With two conditions we force the solution to go between exactly through points (t_0, x_0) *and* (t_1, x_1)*. After some steps, the solution turns out to be a cycloid equation.*

Obtaining the Euler equation

The Euler equation plays a similar role in the calculus of variation as the familiar first-order condition in static optimization. Its derivation is very instructive and provides some insights into dynamic optimization. To this end, consider the variational problem (1.46) assuming that admissible functions are C^2. Suppose that $x^* = x^*(t)$ is an optimal solution to the problem and let $h(t)$ be any C^2 function that satisfies the boundary conditions $h(t_0) = 0$, $h(t_1) = 0$. For each real number $a \in \mathbb{R}$, define a *perturbed* function $y(t) = x^*(t) + ah(t)$.

Clearly, $y(t)$ is admissible because it is C^2, and satisfies $y(t_0) = x_0$ and $y(t_1) = x_1$. Let $J(x) \equiv \int_{t_0}^{t_1} F(t, x, \dot{x})dt$ be the objective function. Because of the hypothesis that $x^*(t)$ is optimal, $J(x^*) \geq J(x^* + ah(t))$ for all $a \in \mathbb{R}$. If the function $h(t)$ is kept *fixed*, then $J(x^* + ah(t))$ is a function $g(a)$ of only the single scalar a, given by

$$g(a) \;=\; \int_{t_0}^{t_1} F(t, y(t), \dot{y}(t))dt = \int_{t_0}^{t_1} F(t, x^*(t) + ah(t), \dot{x}^*(t) + a\dot{h}(t))dt.$$

Obviously, $g(0) = J(x^*)$ and $g(a) \leq g(0)$ for all $a \in \mathbb{R}$. Hence the function g has a maximum at $a = 0$, and $d(g(0))/da = 0$. This condition allows us to deduce the Euler equation.

Observe that to calculate $g'(a)$ requires differentiating under the integral sign. We apply Leibnitz's formula (1.15) to obtain

$$g'(a) = \int_{t_0}^{t_1} \frac{\partial}{\partial a} F(t, x^*(t) + ah(t), \dot{x}^*(t) + a\dot{h}(t))dt.$$

According to the chain rule,

$$\frac{\partial}{\partial a}F(t,x^*(t)+ah(t),\dot{x}^*(t)+a\dot{h}(t))=F_y h(t)+F_{\dot{y}}\dot{h}(t),$$

where F_y and $F_{\dot{y}}$ are both evaluated at $(t,x^*+ah(t),\dot{x}^*+a\dot{h}(t))$. For $a=0$, we obtain

$$g'(0)=\int_{t_0}^{t_1}\left[F_x h(t)+F_{\dot{x}}\dot{h}(t)\right]dt.$$

To proceed, integrate the second term of the integrand by parts (1.10) to get

$$\begin{aligned}\int_{t_0}^{t_1}F_{\dot{x}}\dot{h}(t)dt &= \Big|_{t_0}^{t_1}F_{\dot{x}}h(t)-\int_{t_0}^{t_1}\frac{d}{dt}F_{\dot{x}}h(t)dt \\ &= -\int_{t_0}^{t_1}\frac{d}{dt}F_{\dot{x}}h(t)dt,\end{aligned} \tag{1.48}$$

where we used $h(t_0)=h(t_1)=0$ for the last equality. Hence the foc $g'(0)=0$ reduces to

$$g'(0)=\int_{t_0}^{t_1}\left[F_x-\frac{d}{dt}F_{\dot{x}}\right]h(t)dt=0.$$

Because this equation must be valid for *all* functions $h(t)$ that are C^2 on $[t_0,t_1]$ and that are zero at t_0 and t_1, it follows that

$$F_x-\frac{d}{dt}F_{\dot{x}}=0,$$

which is a necessary condition for a maximum (minimum) for the variational problem (1.46). Note that if $dF_{\dot{x}}/dt=0$, we have the familiar first-order condition $F_x=0$. Alternatively, integrating over t the Euler equation can be written as

$$\Big|_{t_0}^{t_1}F_{\dot{x}}(t,x^*,\dot{x}^*)=\int_{t_0}^{t_1}F_x(t,x^*,\dot{x}^*)dt.$$

Optimal savings

The question Ramsey (1928) has addressed is how much investment would be desirable. High consumption today is in itself preferable, but leads to a low rate of investment which in turn results in a lower capital stock in the future, thus reducing the possibilities for future consumption. One must somehow find a way to reconcile the conflict between present and future consumption. To this end, consider the following example.

Example 1.4.2 (Ramsey problem) *Consider an economy where* $K_t\equiv K(t)$ *denotes the* *capital stock, $C_t \equiv C(t)$ consumption, and $Y_t \equiv Y(t)$ aggregate output. Suppose that output*

$$Y_t = f(K_t) = C_t + \dot{K}_t, \quad f'(K_t) > 0,\ f''(K_t) < 0,$$

is a strictly increasing, concave function of the capital stock, divided between consumption C_t and investment $I_t = \dot{K}_t$. Let K_0 be a historically given capital stock, and suppose there is a fixed planning period $[0, T]$. For each choice of investment, capital is fully determined by $K_t = K_0 + \int_0^T \dot{K}_s ds$ which in turn determines C_t. Assume that the society has a utility function $U = U(C_t)$. Suppose that $U'(C_t) > 0$, and $U''(C_t) < 0$, U is strictly increasing and strictly concave. Further we introduce a measure of impatience, discounting U by the discount factor e^{-rt}. The variational problem then reads

$$\max \int_0^T e^{-rt} U(f(K_t) - \dot{K}_t) dt \quad s.t. \quad K(0) = K_0,\ K(T) = K_T,$$

where some terminal condition $K(T)$ is imposed.

Example 1.4.3 (Optimal saving) *Suppose that an individual instantaneous utility is $u(c_t)$ where $u'(c_t) > 0$, and $u''(c_t) < 0$. The household maximizes discounted utility,*

$$\max \int_{t_0}^{t_1} e^{-\rho t} u(c_t) dt \quad s.t. \quad r a_t + w_t = c_t + \dot{a}_t, \quad a(t_0) = a_0,\ a(t_1) = a_1,$$

$\rho > 0$ is the subjective rate of time preference, a_t denotes individual wealth rewarded at the constant interest rate r, and the constant labor supply rewarded at exogenous wage rate w_t. Note that a_t can also be negative (then the household borrows at the interest rate r).

Substituting the budget constraint into the objective function gives

$$\max \int_{t_0}^{t_1} e^{-\rho t} u(r a_t + w_t - \dot{a}_t) dt, \quad a(t_0) = a_0,\ a(t_1) = a_1.$$

To use the Euler equation, we use partial derivatives, $F_a = e^{-\rho t} u'(c_t) r$, and $F_{\dot{a}} = -e^{-\rho t} u'(c_t)$. Hence we obtain from (1.47),

$$\begin{aligned} F_a = dF_{\dot{a}}/dt \ \Leftrightarrow\ e^{-\rho t} u'(c_t) r &= d(-e^{-\rho t} u'(c_t))/dt \\ &= \rho e^{-\rho t} u'(c_t) - e^{-\rho t} u''(c_t) \dot{c}_t \\ \Leftrightarrow\ r - \rho &= -\frac{u''(c_t)}{u'(c_t)} \dot{c}_t, \quad \text{where} \quad -\frac{u''(c_t)}{u'(c_t)} > 0. \end{aligned}$$

Note that $u''(c)/u'(c)$ can be interpreted as the instantaneous growth rate of $u'(c)$. Because $F(t, a, \dot{a})$ is concave, the Euler equation is a sufficient condition for optimality.

Special cases

Euler equations in general are difficult to solve. There are important special cases, however, where the problem reduces substantially.

1. objective function (the integrand) does *not* depend on x explicitly, that is $F(t, \dot{x})$

2. the integrand does *not* depend on t explicitly, that is $F(x, \dot{x})$

In both cases, the problem reduces to solving a first-order differential equation which, in general, is easier to handle than the usual second-order Euler equation.

Exercise 1.4.4 *Characterize the possible solutions to the variational problem*

$$\max \int_1^T (3\dot{x} - t\dot{x}^2)dt \quad s.t. \quad x(1) = x_1,\ x(T) = x_T.$$

Exercise 1.4.5 *Characterize the possible solutions to the variational problem*

$$\min \int_0^T x\sqrt{1+\dot{x}^2}dt \quad s.t. \quad x(0) = x_0,\ x(T) = x_T.$$

1.4.2 More general terminal conditions

So far, boundary values of the unknown function have been fixed. In economic applications the initial point is usually fixed because it represents a historically given situation. However, the terminal value can be free, or subject to more general restrictions. In what follows we review two most common terminal conditions that appear in economic models.

The two problems can be formulated as

$$\max \int_{t_0}^{t_1} F(t, x, \dot{x})dt \quad s.t. \quad x(t_0) = x_0,\ x(t_1) \text{ free}, \tag{1.49}$$

and

$$\max \int_{t_0}^{t_1} F(t, x, \dot{x})dt \quad s.t. \quad x(t_0) = x_0,\ x(t_1) \geq x_1. \tag{1.50}$$

Again, F is a given *well behaved* function of three variables, whereas t_0 and t_1, as well as x_0 are given numbers. Among all well behaved functions $x(t)$ that satisfy $x(t_0) = x_0$ and $x(t_1)$ satisfying either terminal condition, find one making the integral (1.46) as large as possible.

An important observation is that an optimal solution to either of the two problems must satisfy the Euler equation. Suppose x^* solves either problem. The condition $x^*(t_0) = x_0$ places one restriction on the constants in the general solution of the Euler equation. A so called *transversality condition* is needed to determine the other constant. The relevant condition is given in the following theorem (Sydsæter et al. 2005, Theorem 8.5.1).

Theorem 1.4.6 (Transversality conditions) *If $x^*(t)$ solves the variational problem with either (1.49) or (1.50) as the terminal condition then $x^*(t)$ must satisfy the Euler equation (1.47). For $x(t_1)$ free the transversality condition is*

$$(F_{\dot{x}})_{t=t_1} = 0.$$

With the terminal condition $x(t_1) \geq 0$, the transversality condition is

$$(F_{\dot{x}})_{t=t_1} \leq 0, \quad (F_{\dot{x}})_{t=t_1} = 0 \quad \text{if} \quad x^*(t_1) > x_1.$$

If $F(t, x, \dot{x})$ is concave in $(x, \dot{x})$, then any admissible $x^*(t)$ satisfying both the Euler equation and the appropriate transversality condition will solve the problem (1.49) or (1.50).

Obtaining the transversality condition

Recall that an optimality condition for deriving the Euler equation from (1.48) was

$$\int_{t_0}^{t_1} F_{\dot{x}} \dot{h}(t) dt \;=\; \Big|_{t_0}^{t_1} F_{\dot{x}} h(t) - \int_{t_0}^{t_1} \frac{d}{dt} F_{\dot{x}} h(t) dt.$$

The first term of the right-hand side vanishes only if we demand that $h(t_1) = 0$. However, we now allow for solutions where $\big|_{t_0}^{t_1} F_{\dot{x}} h \neq 0$, because $h(t_1) \neq 0$. Hence, we obtain

$$\int_{t_0}^{t_1} F_{\dot{x}} \dot{h}(t) dt \;=\; F_{\dot{x}} h(t_1) - \int_{t_0}^{t_1} \frac{d}{dt} F_{\dot{x}} h(t) dt.$$

As a result, the first-order condition becomes

$$g'(0) = \int_{t_0}^{t_1} \left[F_x - \frac{d}{dt} F_{\dot{x}} \right] h(t) dt + F_{\dot{x}} h(t_1) = 0,$$

which must be valid for *all* functions $h(t)$ that are C^2 on $[t_0, t_1]$. In that the Euler equation has to be augmented by the *transversality condition* $F_{\dot{x}}(t_1, x, \dot{x}) = 0$. Intuitively, a change in the variable $\dot{x}$ at the time t_1 should not lead to an increase in the objective function.

Exercise 1.4.7 (Atkinson's pensioner) *Let $a(t)$ denote a pensioner's wealth at time t, and let w be the (constant) pension income. Suppose that the person can borrow and save at the same constant interest rate r. Consumption at time t is $c_t = ra_t + w - \dot{a}_t$. Suppose the* *pensioner plans consumption from $t = 0$ until terminal date T such as to maximize*

$$\max \int_0^T e^{-\rho t} u(c_t) dt \quad s.t. \quad a(0) = a_0, \; a(T) \geq a_T,$$

where u is a utility function with $u' > 0$, $u'' < 0$, and ρ is a discount rate (see Atkinson 1971). Characterize the possible solutions.

Other extensions such as problems with variable final time, infinite horizon, or several unknown functions will be dealt using the more general *control theory*.

1.5 Control theory

Literature: Kamien and Schwartz (1991, part 2), Sydsæter et al. (2005, chap. 9,10)
Optimal control theory often is able to solve complicated structured problems. It is a modern extension of the classical calculus of variation and goes back to Pontryagin et al. (1962).

1.5.1 Maximum principle

Consider a system whose state at time t is characterized by a number $x(t)$, the *state variable*. The process that controls $x(t)$ at least partially is called a *control function* $u(t)$. In what follows, we assume that the rate of change of $x(t)$ depends on t, $x(t)$, and $u(t)$. The state at some initial point t_0 is typically known, $x(t_0) = x_0$. Hence, the basic problem reads,

$$\max \int_{t_0}^{t_1} f(t, x(t), u(t)) dt \quad s.t. \quad \dot{x} = g(t, x(t), u(t)), \quad x(t_0) = x_0. \tag{1.51}$$

Note that the variational problem is given for $\dot{x} = g(t, x(t), u(t)) = u(t)$. By choosing different control functions $u(t)$, the system can be steered along many different paths, not all of which are equally desirable. As usual we therefore define an *objective function*, which is the integral $J = \int_{t_0}^{t_1} f(t, x(t), u(t)) dt$. Certain restrictions are often placed on the final state $x(t_1)$. Moreover, the time t_1 at which the process stops is not necessarily fixed. Among all admissible pairs $(x(t), u(t))$ that obey the differential equation in (1.51) with $x(t_0) = x_0$ and that satisfy the constraints imposed on $x(t_1)$, find one that maximizes the objective function in (1.51), i.e. find the optimal pair.

We introduce a function $\lambda = \lambda(t)$ associated with the constraint (or *transition equation*) for each t in $[t_0, t_1]$. We refer to this function as the *adjoint function* (or *costate* variable) associated with the differential equation. Corresponding to the Lagrangian function is the

Hamiltonian H. For each time t in $[t_0, t_1]$ and each possible triple (x, u, λ), of the state, control, and adjoint variables, the Hamiltonian is defined by

$$H(t, x, u, \lambda) = f(t, x, u) + \lambda(t)g(t, x, u). \tag{1.52}$$

The *maximum principle* gives necessary conditions for optimality, similar to the Euler equation (including all necessary conditions emerging from the classical theory), for a wide range of dynamic optimization problems. Suppose that $(x^*(t), u^*(t))$ is an optimal pair for the problem (1.51). Then there exists a continuous function and piecewise differentiable function $\lambda(t)$ such that, for all t in $[t_0, t_1]$ (Sydsæter et al. 2005, Theorem 9.2.1),

$$\begin{aligned} u &= u^*(t) \quad \text{maximizes} \quad H(t, x^*(t), u, \lambda(t)) \quad \text{for} \quad u \in (-\infty, \infty), & (1.53)\\ \dot{\lambda}(t) &= -H_x(t, x^*(t), u^*(t), \lambda(t)), \quad \lambda(t_1) = 0. & (1.54) \end{aligned}$$

The requirement that $\lambda(t_1) = 0$ is called *transversality condition.* It tells us that in the case where $x(t_1)$ is free, the adjoint variable vanishes at t_1. If the requirement

$$H(t, x, u, \lambda(t)) \quad \text{is concave in} \quad (x, u) \quad \text{for each} \quad t \in [t_0, t_1]$$

is added we obtain sufficient conditions (Sydsæter et al. 2005, Theorem 9.2.2). In a way, changing $u(t)$ on a small interval causes $f(t, x, u)$ to change immediately. Moreover, at the end of this interval $x(t)$ has changed and this change is transmitted throughout the remaining time interval. In order to steer the process optimally, the choice of $u(t)$ at each instant of time must *anticipate* the future changes in $x(t)$. In short, we have to plan ahead or have to be *forward looking.* In a certain sense, the adjoint equation takes care of this need for forward planning, $\lambda(t_0) = \int_{t_0}^{t_1} H_x(s, x^*(s), u^*(s), \lambda(s))ds$.

Since the control region is $(-\infty, \infty)$, a *necessary* condition for (1.53) is that

$$H_u(t, x^*(t), u^*(t), \lambda(t)) = 0.$$

If $H(t, x(t), u, \lambda(t))$ is concave in u, it is also sufficient for the maximum condition (1.53) to hold, because an interior stationary point for a concave function is (globally) optimal.

To summarize, necessary conditions for an optimal solution to the control problem (1.51) are obtained defining the Hamiltonian $H(t,x,u,\lambda) = f(t,x,u) + \lambda(t)g(t,x,u)$ which requires

$$
\begin{aligned}
H_u &= 0, &(1.55)\\
H_x &= -\dot{\lambda}, &(1.56)\\
H_\lambda &= \dot{x}. &(1.57)
\end{aligned}
$$

For a maximum, the necessary conditions are sufficient for optimality if f and g are concave in (x,u), respectively and $\lambda(t) \geq 0$. Similarly for a minimum, if f is concave, g is convex in (x,u) and $\lambda \leq 0$, the necessary conditions are sufficient.

Example 1.5.1 *Consider the variational problem*

$$\max \int_{t_0}^{t_1} f(t,x,\dot{x})dt \quad s.t. \quad x(0) = x_0.$$

Using $u = \dot{x}$, it becomes a control problem. The Hamiltonian is $H = f(t,x,u) + \lambda u$, and

$$H_u = f_u + \lambda = 0, \quad H_x = f_x = -\dot{\lambda},$$

are the first-order conditions. Hence, $f_u = f_{\dot{x}}$ and therefore $f_{\dot{x}} = -\lambda$. As a result,

$$f_x = -\dot{\lambda} = -\frac{d}{dt}\lambda = \frac{d}{dt}f_{\dot{x}},$$

which is the Euler equation of the variational problem. Moreover, $0 = \lambda(t_1) = -f_{\dot{x}}$ is the transversality condition. Assuming that f is a C^2 function, if the Hamiltonian attains its maximum at $u^*(t)$, not only is $H_u = 0$, but also $H_{uu} \leq 0$, implying that $f_{\dot{x}\dot{x}} \leq 0$ which is the Legendre condition in the calculus of variation.

Obtaining the maximum principle

Consider the control problem (1.51) assuming that admissible functions are C^2. Because the constraint is a differential equation on the interval $[t_0, t_1]$, it can be regarded as an infinite number of equality constraints, one for each time t. Economists usually incorporate equality constraints by forming a Lagrangian function, with a Lagrange multiplier corresponding to each constraint. Thus, the problem can be written as

$$\max \int_{t_0}^{t_1} \big(f(t,x(t),u(t)) - \lambda(t)\left[\dot{x} - g(t,x(t),u(t))\right]\big)dt \quad s.t. \quad x(t_0) = x_0.$$

Observing that integration by parts (1.12) gives,

$$\int_{t_0}^{t_1} \lambda(t)\dot{x}(t)dt = \lambda(t_1)x(t_1) - \lambda(t_0)x(t_0) - \int_{t_0}^{t_1} \dot{\lambda}(t)x(t)dt,$$

we obtain

$$\max \int_{t_0}^{t_1} \Big(f(t,x,u) + \lambda g(t,x,u) + \dot{\lambda}x\Big)\, dt + \lambda(t_0)x(t_0) - \lambda(t_1)x(t_1) \quad s.t. \quad x(t_0) = x_0.$$

Suppose that $u^* = u^*(t)$ is an optimal control, and let $h(t)$ be any C^2 function. For each real number $a \in \mathbb{R}$, let $y(t) \equiv u^*(t) + ah(t)$ be an admissible control and $x(t,a)$ be the admissible state variable that satisfies the constraint $\dot{x} = g(t,x,y)$ and initial condition, $x(t_0) = x_0$. Note that for any given control, $u(t)$, $\dot{x} = g(t,x,u)$ is a first-order differential equation which solution depends on an arbitrary constant. This constant is given by the initial condition $x(t_0) = x_0$, which in general forces a unique solution $x(t,a)$ for a given control $y(t)$. Keeping u^* and h fixed, we solve our problem by maximizing $J(a)$ with respect to a,

$$\begin{aligned} J(a) \quad = \quad & \int_{t_0}^{t_1} \Big(f(t,x(t,a),u^*(t)+ah(t)) + \lambda(t)g(t,x(t,a),u^*(t)+ah(t)) + \dot{\lambda}x(t,a)\Big)dt \\ & +\lambda(t_0)x(t_0,a) - \lambda(t_1)x(t_1,a). \end{aligned}$$

The initial condition is given by $\lambda(t_0)x(t_0,a) = \lambda(t_0)x_0 \ \forall a$. The first-order condition reads

$$J'(a) \quad = \quad \int_{t_0}^{t_1} \Big((f_x + \lambda(t)g_x + \dot{\lambda})x_a + (f_u + \lambda(t)g_u)h(t)\Big)dt - \lambda(t_1)x_a(t_1,0) = 0.$$

Observe that the condition is satisfied if

(a) $\lambda(t)$ solves the first-order differential equation

$$\dot{\lambda} + \lambda(t)g_x + f_x = 0$$

with terminal condition is $\lambda(t_1) = 0$,

(b) and $f_u + \lambda g_u = 0$.

Of course, by construction the constraint $\dot{x} = g(t,x,u)$ must hold. Finally, defining the Hamiltonian $H(t,x,u,\lambda) = f(t,x,u) + \lambda(t)g(t,x,u)$ gives the necessary conditions of the maximum principle (1.55) to (1.57), together with conditions $x(t_0) = x_0$ and $\lambda(t_1) = 0$. For a maximum (minimum) we have that $H_{uu} \leq 0$ ($H_{uu} \geq 0$).

1.5.2 Regularity conditions

In many applications of control theory to economics, the control function are explicitly or implicitly restricted in various ways. In general, assume that $u(t)$ takes values in a fixed subset U of the reals, called *control region.* An important aspect of control theory is that $u(t)$ can take values at the boundary of U. One example are *bang-bang* controls

$$u(t) = \begin{cases} 1 & \text{for} \quad t \in [t_0, t'] \\ 0 & \text{for} \quad t \in (t', t_1] \end{cases},$$

which involve a single shift at time t'. In this case $u(t)$ is *piecewise continuous*, with a jump discontinuity at $t = t'$. A function is piecewise continuous if it has at most a finite number of discontinuities on each finite interval, with finite jumps at each point of discontinuity. The value of a control $u(t)$ at a point of discontinuity will not be of any importance, but let us agree to choose the value of $u(t)$ at a point of discontinuity t as the left-hand limit of $u(t)$ at t'. Then $u(t)$ will be *left-continuous* which will be implicitly assumed.

A solution to a control problem where $u = u(t)$ has discontinuities is a *continuous* function that has a derivative that satisfies the equation, except at points where $u(t)$ is discontinuous. The graph of $x(t)$ will, in general, have kinks at the points of discontinuity of $u(t)$, and will usually not be differentiable at these kinks. It is, however, still continuous at the kinks.

So far no restrictions have been placed on the functions $g(t, x, u)$ and $f(t, x, u)$. In general, we implicitly assume that both functions are of class C^1, that is f, g, and their first-order partial derivatives with respect to x and u are continuous in (t, x, u).

1.5.3 Standard end constraint problems

The *standard end constrained problem* imposes one of the following terminal conditions,

$$(a)\ x(t_1) = x_1, \quad (b)\ x(t_1) \geq x_1, \quad (c)\ x(t_1) \text{ free.}$$

It can be shown that the necessary conditions for optimality are the same as for the basic control problem, but the transversality condition is either

$$(a')\ \lambda(t_1) \text{ free} \quad (b')\ \lambda(t_1) \geq 0,\ (\text{with } \lambda(t_1) = 0 \text{ if } x^*(t_1) > x_1) \quad (c')\ \lambda(t_1) = 0. \tag{1.58}$$

Exercise 1.5.2 (Optimal consumption) *Solve the following control problem*

$$\int_0^T u(c_t)dt \quad s.t. \quad \dot{a}_t = ra_t - c_t, \quad a(0) = a_0,\ a(T) \geq 0,$$

where r denotes the constant rental rate of capital. Suppose that $u' > 0$, $u'' < 0$ is strictly concave and assume that $c(t) > 0$ so that the control region is $(0, \infty)$.

Exercise 1.5.3 (Bang-bang) *Solve the following control problem*

$$\max \int_0^1 (2x - x^2)dt, \quad s.t. \quad \dot{x} = u, \quad x(0) = 0, \ x(1) = 0,$$

for the control region $u \in [-1, 1]$.

1.5.4 Variable final time

In the optimal control problems studied so far the time interval has been fixed. Yet for some control problems in economics, the final time is a variable to be chosen optimally, along the path $u(t)$, $t \in [t_0, t_1]$. A variable final time problem is for example

$$\max_{\{u,t_1\}} \int_{t_0}^{t_1} f(t, x(t), u(t))dt \quad s.t. \quad \dot{x} = g(t, x(t), u(t)), \quad x(t_0) = x_0. \tag{1.59}$$

The maximum principle with variable final time then states that all necessary conditions hold, and, in addition (Sydsæter et al. 2005, Theorem 9.8.1)

$$H(t_1^*, x^*(t_1), u^*(t), \lambda(t_1^*)) = 0.$$

Basically, one additional unknown is determined by one extra condition. Hence, the method for solving variable final time problems is first to solve the problem with fixed t_1 for every $t_1 > t_0$. The optimal final time t_1^* must then satisfy the additional restriction. Note that concavity of the Hamiltonian in (x, u) is *not* sufficient for optimality when t_1 is free.

Exercise 1.5.4 (Hotelling's rule) *Consider $x(t)$ as the amount of an exhaustible resource in a reservoir at time t, where $x(0) = x_0$. Let $u(t)$ denote the rate of extraction such that $x(t) = x_0 - \int_0^t u(s)ds$. Suppose the price of the resource at time t is $q(t)$, and the sales revenue per unit of time at t is $q(t)u(t)$. Assume further that the cost of extraction is $C = C(t, x, u)$, thus the instantaneous rate of profit at time t is*

$$\pi(t, x, u) = q(t)u(t) - C(t, x, u).$$

Let the discount rate be r, so that the control problem reads

$$\max \int_0^T e^{-rt}\pi(t, x, u)dt, \quad s.t. \quad \dot{x} = -u(t), \quad x(0) = x_0.$$

It is natural to assume that $u(t) \geq 0$, and that $x(T) \geq 0$. Consider the following problems.

(a) Characterize the optimal rate of extraction $u^* = u^*(t)$ which solves the control problem over a fixed extraction period $[0, T]$. If necessary, impose further restrictions on C.

(b) Find the optimal stopping time T that solves the control problem for $C = C(t, u)$.

This example builds on Hotelling (1931).

1.5.5 Current value formulations

Many control problems in economics have the following structure

$$\max \int_{t_0}^{t_1} e^{-(t-t_0)r} f(t, x(t), u(t)) dt \quad s.t. \quad \dot{x} = g(t, x(t), u(t)), \quad x(t_0) = x_0. \tag{1.60}$$

The new feature is the explicit appearance of the discount factor $e^{-(t-t_0)r}$. For such problems it is often convenient to formulate the maximum principle in a slightly different form. The usual Hamiltonian is $H = e^{-(t-t_0)r} f(t, x, u) + \lambda(t) g(t, x, u)$. Multiply by $e^{(t-t_0)r}$ to obtain the *current value* Hamiltonian, $H^c = He^{(t-t_0)r} = f(t, x, u) + e^{(t-t_0)r}\lambda(t) g(t, x, u)$. Introducing $m(t) = e^{(t-t_0)r}\lambda(t)$ as the *current value* (not discounted) shadow price for the problem,

$$H^c(t, x, u, m) = f(t, x, u) + m(t) g(t, x, u).$$

In fact, the maximum principle comprises (Sydsæter et al. 2005, Theorem 9.9.1)

$$\begin{aligned} u &= u^*(t) \quad \text{maximizes} \quad H^c(t, x^*(t), u, m(t)) \quad \text{for} \quad u \in U, \\ \dot{m}(t) &= rm(t) - H^c_x, \end{aligned}$$

with transversality conditions

$$(a') \; m(t_1) \text{ free} \quad (b') \; m(t_1) \geq 0 \text{ (with } m(t_1) = 0 \text{ if } x^*(t_1) > x_1) \quad (c') \; m(t_1) = 0, \tag{1.61}$$

similar to the standard end constrained as in (1.58).

Exercise 1.5.5 *Solve the following control problem*

$$\max \int_0^{20} e^{-0.25t}(4K - u^2) dt \quad s.t. \quad \dot{K} = -0.25K + u, \quad K(0) = K_0, \; K(20) \text{ free},$$

where $u \geq 0$ denotes the repair effort and $K(t)$ the value of a machine, $4K - u^2$ is the instantaneous net profit at time t, and $e^{-0.25t}$ is the discount factor.

1.5.6 Infinite horizon

Most of the optimal growth models appearing in literature have an infinite time horizon. This assumption often does simplify formulas and conclusions, though at the expense of some new mathematical problems that need to be sorted out. A typical infinite horizon optimal control problem takes the form

$$\max \int_{t_0}^{\infty} e^{-(t-t_0)r} f(t, x(t), u(t)) dt \quad s.t. \quad \dot{x} = g(t, x(t), u(t)), \quad x(t_0) = x_0. \tag{1.62}$$

Often no condition is placed on $x(t)$ as $t \to \infty$, but many problems do impose the constraint

$$\lim_{t\to\infty} x(t) \geq x_1, \quad x_1 \in \mathbb{R}.$$

Because of the presence of the discount factor, it is convenient to use the current value formulation with the current value Hamiltonian,

$$H^c(t, x, u, m) = f(t, x, u) + m(t)g(t, x, u),$$

and $m(t)$ as the current value shadow price. From the maximum principle, it can be shown that sufficient conditions are as follows (Sydsæter et al. 2005, Theorem 9.11.1)

$$\begin{array}{ll} (a) & u = u^*(t) \quad \text{maximizes} \quad H^c(t, x^*(t), u, m(t)) \quad \text{for} \quad u \in U, \\ (b) & \dot{m}(t) = rm(t) - H^c_x, \\ (c) & H^c(t, x, u, m(t)) \quad \text{is concave with respect to } (x, u), \\ (d) & \lim_{t\to\infty} m(t)e^{-rt}[x(t) - x^*(t)] \geq 0 \quad \text{for all admissible } x(t). \end{array}$$

Other necessary conditions where a certain growth condition replaces the transversality condition are in Sydsæter et al. (2005, Theorem 9.11.2).

Remark 1.5.6 (Malinvaud) *Note that the inequality (d) must be shown for all admissible $x(t)$, which often is problematic. The following conditions are equivalent to (d) for the case where the terminal condition is $\lim_{t\to\infty} x(t) \geq x_1$ (Michel 1982, Sydsæter et al. 2005),*

$$\begin{array}{ll} (A) & \lim_{t\to\infty} m(t)e^{-rt}[x_1 - x^*(t)] \geq 0, \\ (B) & \textit{there is a number } M \textit{ such that } |m(t)e^{-rt}| \leq M \textit{ for all } t \geq t_0, \\ (C) & \textit{there is a number } s \textit{ such that } m(t) \geq 0 \textit{ for all } t \geq s. \end{array}$$

Suppose that $x(t) \geq x_1$ for all t. Then it suffices to check conditions (A) and (C). This result *is referred to as the Malinvaud transversality condition.*

Exercise 1.5.7 (Infinite horizon) *Consider the control problem*

$$\max \int_0^\infty -u^2 e^{-rt} dt \quad s.t. \quad \dot{x} = ue^{-at}, \quad x(0) = 0, \ \lim_{t\to\infty} x(t) \geq K, \quad u \in \mathbb{R}.$$

The constants r, a, and K are positive, with $a > r/2$. Find the optimal solution.

1.5.7 Several control and state variables

As we show below, most of the results obtained by studying control problems with only one state and one control variable can be generalized to control problems with an arbitrary number of state and control variables.

The *standard problem* is to find for fixed values of t_0 and t_1 a pair of vector functions $(x(t), u(t)) = ((x_1(t), ..., x_n(t))^\top, (u_1(t), ..., u_r(t))^\top)$ on $[t_0, t_1]$, which maximizes the objective function

$$\max \int_{t_0}^{t_1} f(t, x(t), u(t)) \quad s.t. \quad \dot{x} = g(t, x(t), u(t)), \quad x_i(t_0) = x_i^0, \quad i = 1, ..., n,$$

where $\dot{x} = g(t, x(t), u(t)) = (g_1(t, x(t), u(t)), ..., g_n(t, x(t), u(t)))^\top$, satisfying initial conditions $x^0 = (x_1^0, ..., x_n^0)^\top \in \mathbb{R}^n$, the terminal conditions

$$\begin{array}{lll} (a) & x_i(t_1) = x_i^1, & i = 1, ..., l \\ (b) & x_i(t_1) \geq x_i^1, & i = l+1, ..., m \\ (c) & x_i(t_1) \text{ free}, & i = m+1, ..., n \end{array} \tag{1.63}$$

and the control region, $u(t) = (u_1(t), ..., u_r(t))^\top \in U \subseteq \mathbb{R}^r$ where U is a given set in $\mathbb{R}^r$. Any pair $(x(t), u(t))$ is admissible if $u_1(t), ..., u_r(t)$ are all piecewise continuous, $u(t)$ takes values in U and $x(t)$ is the corresponding continuous and piecewise differentiable vector function that satisfies the dynamic constraints as well as initial and terminal conditions. The functions f and $g = (g_1, ..., g_n)^\top$ are C^1 with respect to the $n + r + 1$ variables.

The Hamiltonian $H(t, x, u, \lambda)$, with $\lambda = (\lambda_1, ..., \lambda_n)^\top$, is then defined by

$$H(t, x, u, \lambda) = f(t, x, u) + \lambda^\top g(t, x, u) = f(t, x, u) + \sum_{i=1}^n \lambda_i g_i(t, x, u).$$

The *maximum principle* then reads as follows (Sydsæter et al. 2005, Theorem 10.1.1). Suppose that $(x^*(t), u^*(t))$ is an optimal pair for the standard end constrained problem.

Then there exists a continuous and piecewise differentiable function $\lambda(t) = (\lambda_1(t), ..., \lambda_n(t))^\top$ such that for all t in $[t_0, t_1]$

$$
\begin{aligned}
u &= u^*(t) \quad \text{maximizes} \quad H(t, x^*(t), u^*(t)) \quad \text{for} \quad u \in U, && (1.64)\\
\dot{\lambda}_i &= -H_{x_i}(t, x^*(t), u^*(t), \lambda(t)), \quad i = 1, ..., n. && (1.65)
\end{aligned}
$$

Corresponding to the terminal conditions (1.63), one has the transversality conditions,

$$
\begin{array}{lll}
(a) & \lambda_i(t_1) \text{ free} & i = 1, ..., l \\
(b) & \lambda_i(t_1) \geq 0 \quad (\lambda_i(t_1) = 0 \text{ if } x_i^*(t_1) > x_1^i) & i = l+1, ..., m \\
(c) & \lambda_i(t_1) = 0 & i = m+1, ..., n
\end{array}.
$$

For sufficient conditions see e.g. Sydsæter et al. (2005, Theorems 10.1.2 and 10.1.3).

Exercise 1.5.8 (Optimal resource depletion) *Consider an economy using an exhaustible resource, $R_t \equiv R(t)$, as an input factor to produce output,*

$$
Y_t = R_t^\alpha K_t^{1-\alpha}, \quad 0 < \alpha < 1, \quad K(0) = K_0,
$$

where $K_t \equiv K(t)$ is the aggregate capital stock. Capital is accumulated if net investment is positive, that is total output exceeds aggregate consumption, $C_t \equiv C(t)$,

$$
I_t \equiv I(t) = \dot{K}_t = Y_t - C_t.
$$

Let $X_t \equiv X(t)$ be the amount of the resource in a reservoir at time t, and $X(0) = X_0$. Suppose the planner intends to consume all stocks completely, $X(T) = K(T) = 0$, to maximize the utility $U = \int_0^T \ln C_s ds$. Find the optimal paths for consumption and resource depletion.

This exercise builds on Dasgupta and Heal (1974).

1.6 Dynamic programming

Literature: Kamien and Schwartz (1991, chap. 2.21), Wälde (2009, chap. 3.3,6)
This chapter gives a brief introduction to *continuous-time* dynamic programming, showing how to solve optimization problems using dynamic programming methods. A typical problem to be tackled by dynamic programming takes the form of a control problem,

$$
\max \int_{t_0}^{\infty} e^{-\int_{t_0}^t \rho(s)ds} f(t, x(t), u(t))dt \quad s.t. \quad \dot{x} = g(t, x(t), u(t)), \quad x(t_0) = x_0, \tag{1.66}
$$

where we are focusing on infinite horizon models throughout the chapter. Suppose that $(x^*(t), u^*(t))$ is an optimal pair among the admissible pairs for the problem (1.66).

1.6.1 Bellman's principle

We define the (optimal) value function at time t_0 by

$$V(t_0, x(t_0)) = \int_{t_0}^{\infty} e^{-\int_{t_0}^{t} \rho(s)ds} f(t, x^*(t), u^*(t))dt. \tag{1.67}$$

Note that the value function does not depend on the control. The reason is, as it will become clear below, that the optimal controls $u^*(t)$ will depend on $x(t)$. In the optimum the controls are a function of the state variables.

Solving the control problem (1.67) using dynamic programming essentially requires a *three-step* procedure (Wälde 2009, chap. 6). As a first step, similar to the Euler equation or the maximum principle, a necessary condition for optimality is

$$\rho(t_0)V(t_0, x(t_0)) = \max_{u \in U} \left\{ f(t_0, x(t_0), u(t_0)) + \frac{d}{dt}V(t_0, x(t_0)) \right\}, \tag{1.68}$$

to which we refer as the *Bellman equation* or sometimes called the fundamental equation of dynamic programming. As a corollary, the *first-order condition* is

$$f_u(t_0, x(t_0), u(t_0)) + \frac{\partial}{\partial u}\left(\frac{d}{dt}V(t_0, x(t_0))\right) = 0,$$

where in both equations using $dV(t_0, x(t_0)) = V_t dt + V_x dx$,

$$\frac{d}{dt}V(t_0, x(t_0)) = V_t + V_x\dot{x} \;\; = \;\; V_t + g(t_0, x(t_0), u(t_0))V_x.$$

As the value function does not depend on the control, the first-order condition simplifies to

$$f_u(t_0, x(t_0), u(t_0)) + g_u(t_0, x(t_0), u(t_0))V_x = 0. \tag{1.69}$$

In a second step, we determine the *evolution of the costate variable*, defined as the law of motion of the partial derivative of the value function with respect to the state variable. Using the maximized Bellman equation we obtain

$$\begin{aligned} \rho(t_0)V_x &= f_x(t_0, x_0, u(x(t_0))) + V_{tx} + g_x(t_0, x(t_0), u(t_0))V_x + V_{xx}\dot{x} \\ \Leftrightarrow \;\; (\rho - g_x)V_x &= f_x + V_{tx} + V_{xx}\dot{x}. \end{aligned} \tag{1.70}$$

Observing that the time-derivative of the costate is

$$\frac{d}{dt}V_x(t_0, x(t_0)) = V_{xt} + V_{xx}\dot{x},$$

we insert this into (1.70) to obtain

$$\dot{V}_x \quad = \quad (\rho - g_x)V_x - f_x, \tag{1.71}$$

which describes the evolution of the costate variable, the shadow price of the state variable.

As the final step we use the time-derivative of the first-order condition (1.69),

$$f_{ut} + f_{ux}\dot{x} + f_{uu}\dot{u} + (g_{ut} + g_{ux}\dot{x} + g_{uu}\dot{u})V_x + \dot{V}_x g_u = 0,$$

substituting $\dot{V}_x$ by the expression in (1.71), and the costate V_x using the first-order condition (1.69) to obtain a generalized *Euler equation*,

$$f_{ut} + f_{ux}\dot{x} + f_{uu}\dot{u} - (g_{ut} + g_{ux}\dot{x} + g_{uu}\dot{u})f_u/g_u - (\rho - g_x)f_u - f_x g_u = 0. \tag{1.72}$$

Example 1.6.1 *Using the standard variational problem, $u = \dot{x}$, and $\rho = 0$ it simplifies to*

$$f_{\dot{x}t} + f_{\dot{x}x}\dot{x} + f_{\dot{x}\dot{x}}\ddot{x} - f_x = 0,$$

which is the familiar Euler equation in (1.47).

Example 1.6.2 *Consider a typical control problem,*

$$\max \int_0^\infty e^{-\rho t}u(c(t))dt \quad s.t. \quad \dot{a} = ra - c, \quad a(0) = a_0,$$

where ρ, i are positive constants. Suppose the control is $u = c$, and the state is $x = a$,

$$\begin{aligned} f(t, a(t), c(t)) &= f(c(t)) = u(c) \quad \Rightarrow f_c = u'(c),\ f_{cc} = u''(c),\ f_a = f_{ct} = f_{ca} = 0, \\ g(t, a(t), c(t)) &= g(x(t), c(t)) = ra - c \quad \Rightarrow g_a = r,\ g_c = -1, g_{ct} = g_{ca} = g_{cc} = 0. \end{aligned}$$

Going step-by-step through the suggested procedure or just plugging the partial derivatives in the generalized Euler equation (1.72) gives the necessary condition, $u''(c)\dot{c} = (\rho - r)u'(c)$.

Obtaining the Bellman equation

The heuristic derivation of the Bellman equation is very instructive and provides insights into dynamic optimization. It shows Bellman's trick to simplify the multi-dimensional problem of choosing a complete path of optimal controls, to a one-dimensional problem of choosing the optimal control in the initial period (Chang 1988, Sennewald and Wälde 2006). Consider the control problem (1.66) assuming that admissible functions are C^1. Suppose further that an optimal process $u^* = u^*(t)$ exists. For small $h > 0$, and $\rho(t) \geq 0$ we may write

$$\begin{aligned} V(t_0, x(t_0)) &= \int_{t_0}^{t_0+h} e^{-\int_{t_0}^{t} \rho(s)ds} f(t, x^*(t), u^*(t))dt \\ &\quad + e^{-\int_{t_0}^{t_0+h} \rho(s)ds} \int_{t_0+h}^{\infty} e^{-\int_{t_0+h}^{t} \rho(s)ds} f(t, x^*(t), u^*(t))dt. \end{aligned}$$

The term $\int_{t_0+h}^{\infty} e^{-\int_{t_0+h}^{t} \rho(s)ds} f(t, x^*(t), u^*(t))dt$ simply denotes the value of the optimal program at $t = t_0 + h$. Hence, for any control $u(t)$ with $t \geq t_0 + h$,

$$\int_{t_0+h}^{\infty} e^{-\int_{t_0+h}^{t} \rho(s)ds} f(t, x(t), u(t))dt \leq V(t_0 + h, x(t_0 + h)),$$

with equality for the optimal pair $(x^*(t), u^*(t))$. Therefore,

$$0 = \int_{t_0}^{t_0+h} e^{-\int_{t_0}^{t} \rho(s)ds} f(t, x^*(t), u^*(t))dt + e^{-\int_{t_0}^{t_0+h} \rho(s)ds} V(t_0 + h, x(t_0 + h)) - V(t_0, x(t_0)).$$

Dividing by h and let $h \to 0$ (from above), the equation becomes

$$\begin{aligned} 0 &= \lim_{h\to 0} \frac{1}{h} \int_{t_0}^{t_0+h} e^{-\int_{t_0}^{t} \rho(s)ds} f(t, x^*(t), u^*(t))dt \\ &\quad + \lim_{h\to 0} \frac{1}{h} \left(e^{-\int_{t_0}^{t_0+h} \rho(s)ds} V(t_0 + h, x(t_0 + h)) - V(t_0, x(t_0)) \right). \end{aligned} \tag{1.73}$$

The last term is the derivative of $e^{-\int_{t_0}^{t_0+h} \rho(s)ds} V(t_0 + h, x(t_0 + h))$ with respect to h,

$$\begin{aligned} \frac{d}{dh} e^{-\int_{t_0}^{t_0+h} \rho(s)ds} V(t_0 + h, x(t_0 + h)) &= -\rho(t_0 + h) e^{-\int_{t_0}^{t_0+h} \rho(s)ds} V(t_0 + h, x(t_0 + h)) \\ &\quad + e^{-\int_{t_0}^{t_0+h} \rho(s)ds} \frac{d}{dh} V(t_0 + h, x(t_0 + h)), \end{aligned}$$

where $\frac{d}{dh} V(t_0 + h, x(t_0 + h))$ for $h = 0$ is equal to $\frac{d}{dt} V(t_0, x(t_0))$. Similarly, the first term is the derivative of $\int_{t_0}^{t_0+h} e^{-\int_{t_0}^{t} \rho(s)ds} f(t, x^*(t), u^*(t))dt$ with respect to h, which for $h = 0$ is

$$\begin{aligned}\frac{d}{dh}\int_{t_0}^{t_0+h} e^{-\int_{t_0}^{t} \rho(s)ds} f(t, x^*(t), u^*(t))dt &= e^{-\int_{t_0}^{t_0+h} \rho(s)ds} f(t_0 + h, x^*(t_0 + h), u^*(t_0 + h)) \\ &= f(t_0, x^*(t_0), u^*(t_0)).\end{aligned}$$

Therefore, we may rewrite (1.73) as

$$\rho(t_0)V(t_0, x(t_0)) = f(t_0, x^*(t_0), u^*(t_0)) + \frac{d}{dt}V(t_0, x(t_0)),$$

which is the maximized Bellman equation (1.68).

Obtaining the Bellman equation II

The second heuristic derivation is an application of the Leibnitz formula (1.15) which sheds light on the economic content of the Bellman equation (as taken from Wälde 2009, chap. 6). Given the control problem in (1.66), we may define a criterion function,

$$U(t, x(t), u(t)) = \int_t^\infty e^{-\int_t^\tau \rho(s)ds} f(\tau, x(\tau), u(\tau))d\tau,$$

simply denoting the value of a given program at time t. Its time-derivative reads,

$$\begin{aligned}\frac{d}{dt}U(t, x(t), u(t)) &= -e^{-\int_t^t \rho(s)ds} f(t, x(t), u(t)) + \int_t^\infty \frac{\partial}{\partial t}\left(e^{-\int_t^\tau \rho(s)ds} f(\tau, x(\tau), u(\tau))\right) d\tau \\ &= -f(t, x(t), u(t)) + \rho(t)U(t, x(t), u(t)).\end{aligned}$$

Observe that

$$U(t, x(t), u(t)) = \frac{f(t, x(t), u(t)) + \dot{U}(t, x(t), u(t))}{\rho(t)}$$

denotes the present value of an perpetuity (periodic payment continuing indefinitely). These periodic payments consists of the instantaneous payment $f(t, x(t), u(t))$, say instantaneous utility, plus the present value of a perpetuity that reflects the change in the periodic payment, where future payments are discounted at the rate $\rho(t)$. Considering the optimal control $u^*(t)$, $U(t)$ denotes the value of the optimal program, $V(t, x^*(t))$. Collecting terms we obtain

$$\rho V(t, x^*(t)) = f(t, x^*(t), u^*(t)) + \frac{d}{dt}V(t, x^*(t)),$$

which corresponds to the Bellman equation (1.68).

Remark 1.6.3 (Costate variables) *From (1.71), the evolution of the costate is*

$$\dot{V}_x - (\rho - g_x)V_x + f_x = 0.$$

Consider the present value Hamiltonian without explicit discounting ($\rho = 0$). The shadow price $\lambda(t)$ solves

$$\dot{\lambda} + \lambda(t)g_x + f_x = 0,$$

which is the evolution of the costate in the dynamic programming approach where $V_x = \lambda(t)$. For the current value Hamiltonian ($\rho \geq 0$), the shadow price $m(t)$ solves

$$\dot{m} + m(t)g_x + f_x = \rho m(t),$$

which gives the costate as the current value (not discounted) shadow price $V_x = m(t)$.

Remark 1.6.4 (Infinite horizon) *Solving the differential equation*

$$\dot{U}(t, x(t), u(t)) - \rho(t)U(t, x(t), u(t)) = -f(t, x(t), u(t))$$

requires $\lim_{T\to\infty} e^{-\int_t^T \rho(s)ds} U(T, x(T), u(T)) = 0$ to obtain the criterion function,

$$U(t, x(t), u(t)) = \int_t^\infty e^{-\int_t^\tau \rho(s)ds} f(\tau, x(\tau), u(\tau))d\tau.$$

Remark 1.6.5 (Transversality condition) *Often the limiting inequality in the dynamic programming approach is written as (Sennewald 2007, Theorem 4)*

$$\lim_{t\to\infty} e^{-\rho t}V(t, x(t)) \geq 0 \quad \text{and} \quad \lim_{t\to\infty} e^{-\rho t}V(t, x^*(t)) = 0, \quad \rho > 0,$$

for all admissible $x(t)$, which replaces the transversality condition as a sufficient condition.

Remark 1.6.6 (Boundedness condition) *By considering infinite horizon problems as in (1.62) or (1.66), we implicitly assume that the integral*

$$U(t, x(t), u(t)) = \int_t^\infty e^{-\int_t^\tau \rho(s)ds} f(\tau, x(\tau), u(\tau))d\tau$$

converges for all admissible pairs $(x(t), u(t))$. This assumption has to be checked after having found an optimal control $u^*(t)$. Typically certain growth restrictions emerge ensuring that the integral indeed is bounded, to which we refer as the boundedness conditions.

1.6.2 The envelope theorem

In order to understand the independence of the Bellman equation to the control variable, it is instructive to consider the following theorem (Wälde 2009, Theorem 3.2.1).

Theorem 1.6.7 (Envelope theorem) *Suppose $g(x,u)$ is a C^1 function. Choose u such that $g(x,u)$ is maximized for a given x, assuming that an interior solution exists. Let $f(x)$ be the resulting function of x,*

$$f(x) = \max_{u \in U} g(x,u).$$

Then, the derivative of f with respect to x equals the partial derivative of g with respect to x, if g is evaluated at $u = u(x)$ that maximizes $g(x,u)$,

$$\frac{d}{dx} f(x) = \frac{\partial}{\partial x} g(x,u)\bigg|_{u=u(x)}.$$

Proof. Consider the function $f(x) = f(x, u(x))$. If $u = u(x)$ is a maximum point of $g(x,u)$,

$$df(x,u(x)) \quad = \quad \frac{\partial}{\partial x} g(x,u) dx\bigg|_{u=u(x)} + \frac{\partial}{\partial u} g(x,u) du\bigg|_{u=u(x)} = \frac{\partial}{\partial x} g(x,u) dx\bigg|_{u=u(x)},$$

because $g_u = 0$ at $u = u(x)$ is a necessary condition for a maximium. ■

Exercise 1.6.8 (Envelope theorem) *Let a benevolent planner maximize the social welfare function $U(A,B)$, where A and B are consumption goods. The technologies are $A = A(cL_A)$, $B = B(L_B)$, and the economy's resource constraint is $L_A + L_B = L$. Solve the problem of the optimal allocation of labor to the sectors,*

$$\max_{L_A} \{U(A(cL_A), B(1-L_A))\} \quad s.t. \quad L_B = 1 - L_A.$$

Study the effects of an increase in the technology parameter c on social welfare

(a) without using the envelope theorem,

(b) using the envelope theorem.

Exercise 1.6.9 (Capital adjustment costs) *Solve the optimal control problem of a firm with capital adjustment costs,*

$$\max \int_0^\infty e^{-rt} (F(K_t) - \Phi(I_t))\, dt \quad s.t. \quad \dot{K}_t = I_t - \delta K_t, \quad K(0) = K_0.$$

Assume the production function to be strictly concave, $F' > 0$ and $F'' < 0$. If the firm accumulates capital, it faces quadratic adjustment costs of $\Phi(I_t) = vI_t + I_t^2/2$, where $v > 0$.

1.6.3 Several control and state variables

Because using dynamic programming tackles control problems, extensions such as several control and state variables do not pose new conceptional challenges, however, involve more cumbersome notation. Let us briefly consider the infinite horizon problem,

$$\max \int_{t_0}^{\infty} e^{-(t-t_0)\rho} f(t, x(t), u(t))dt \quad s.t. \quad \dot{x} = g(t, x(t), u(t)), \quad x(t_0) = x_0, \tag{1.74}$$

where $(x(t), u(t)) = ((x_1(t), ..., x_n(t))^\top, (u_1(t), ..., u_r(t))^\top)$ is a pair of vector functions defined on $[t_0, \infty)$, satisfying $\dot{x} = g(t, x(t), u(t)) = (g_1(t, x(t), u(t)), ..., g_n(t, x(t), u(t)))^\top$, and initial conditions $x^0 = (x_1^0, ..., x_n^0) \in \mathbb{R}^n$, and the control region $u(t) = (u_1(t), ..., u_r(t))^\top \in U \subseteq \mathbb{R}^r$.

As a first step, the *Bellman equation* reads

$$\rho V(t_0, x(t_0)) = \max_{u \in U} \left\{ f(t_0, x(t_0), u(t_0)) + \frac{d}{dt} V(t_0, x(t_0)) \right\}, \tag{1.75}$$

where

$$\frac{d}{dt} V(t_0, x(t_0)) = V_t + V_{x_1} g_1 + ... + V_{x_n} g_n.$$

Observe that we have r *first-order conditions*, for $i = 1, ..., r$

$$0 \quad = \quad \frac{\partial}{\partial u_i} f(t_0, x(t_0), u(t_0)) + V_{x_1} \frac{\partial}{\partial u_i} g_1 + ... + V_{x_n} \frac{\partial}{\partial u_i} g_n.$$

The second step is to obtain the *evolution of n costate variables.* For this we use the maximized Bellman equation, for $j = 1, ..., n$

$$\rho V_{x_j} = f_{x_j}(t_0, x(t_0), u(x(t_0))) + V_{t x_j} + V_{x_1 x_j} \dot{x}_1 + ... + V_{x_n x_j} \dot{x}_n + V_{x_1} \frac{\partial}{\partial x_j} g_1 + ... + V_{x_n} \frac{\partial}{\partial x_j} g_n.$$

Observing that the total derivative of the costate of variable $j = 1, ..., n$ is

$$\frac{d}{dt} V_{x_j}(t_0, x(t_0)) = V_{t x_j} + V_{x_1 x_j} \dot{x}_1 + ... + V_{x_n x_j} \dot{x}_n,$$

we obtain

$$\dot{V}_{x_j} = \rho V_{x_j} - f_{x_j} - V_{x_1} \frac{\partial}{\partial x_j} g_1 - ... - V_{x_n} \frac{\partial}{\partial x_j} g_n,$$

describing the evolution of the costate variable j, the shadow price of the state variable j.

As the final step we use the time-derivatives of r first-order conditions, substituting $\dot{V}_{x_j}$ and again the r first-order conditions to substitute costates V_{x_j} to obtain *Euler equations* for the r control variables. Unfortunately, however, it is not always possible to fully eliminate shadow prices from the resulting equations. Appropriate assumptions on f and g may help. The general solution to the problem (1.74) is a system of Euler equations.

1.6.4 An example: Lucas' model of endogenous growth

Consider a closed economy with competitive markets, with identical, rational agents and a constant returns technology, $Y(t) = F(K, N^e)$. At date t there are N workers in total with skill level $h(t)$. Suppose a worker devotes a fraction $u(t)$ of his non-leisure time to current production, and the remaining $1 - u(t)$ to human capital accumulation. Then the effective workforce (that is effective hours) devoted to production is $N^e(t) = u(t)h(t)N$. Suppose that preferences over per-capita consumption streams are given by

$$U \equiv \int_0^\infty e^{-\rho t} \frac{c^{1-\sigma}}{1-\sigma} dt, \tag{1.76}$$

where the subjective discount rate $\rho > 0$, and $\sigma > 0$ (Lucas 1988, Benhabib and Perli 1994).

Production is divided into consumption and capital accumulation. Let $k(t) \equiv K(t)/N$ denote individual physical capital, where $K(t)$ is the total stock of capital,

$$\dot{K} = AF(K, N^e) - Nc(t), \quad A \in \mathbb{R}_+. \tag{1.77}$$

The hourly wage rate per unit of effective labor is $w(t)$, that is the individual's labor income at skill h is $w(t)h(t)u(t)$. Further, the rental rate of physical capital is r_t. To complete the model, the effort $1 - u(t)$ devoted to the accumulation of human capital must be linked to the rate of change in its level, $h(t)$. Suppose the technology relating the growth of human capital $\dot{h}$ to the level already attained and the effort devoted to acquiring more is

$$\dot{h} = (1 - u(t))\delta h(t), \quad \delta \in \mathbb{R}_+. \tag{1.78}$$

According to (1.78), if no effort is devoted to human capital accumulation, $u(t) = 1$, then non accumulates. If all effort is devoted to this purpose, $u(t) = 0$, $h(t)$ grows at rate δ. In between these extremes, there are no diminishing returns to the stock $h(t)$.

The resource allocation problem faced by the representative individual is to choose a time path for $c(t)$ and for $u(t)$ in $U \subseteq \mathbb{R}_+ \times [0,1]$ such as to maximize life-time utility,

$$\max_{\{c(t),u(t)\}_{t=0}^{\infty}} \int_0^{\infty} e^{-\rho t} \frac{c^{1-\sigma}}{1-\sigma} dt \quad s.t. \quad \dot{x} = g(x(t), c(t), u(t)), \quad x(0) = x_0 \in \mathbb{R}_+^2, \tag{1.79}$$

where $x \equiv (k,h)^\top$,

$$\dot{x} \equiv \left(AF(K,N^e)/N - c(t), (1-u(t))\delta h(t)\right)^\top, \tag{1.80}$$

and individual income flow at date t is $AF(K,N^e)/N = r(t)k(t) + w(t)u(t)h(t)$.

As the first step, the *Bellman equation* reads

$$\rho V(x(0)) = \max_{(c(0),u(0))\in U} \left\{ \frac{c^{1-\sigma}}{1-\sigma} + \frac{d}{dt} V(x(0)) \right\}, \tag{1.81}$$

where

$$\frac{d}{dt} V(x(0)) = (r(t)k(t) + w(t)u(t)h(t) - c(t))V_k + (1-u(t))\delta h(t) V_h.$$

Observe that we have two *first-order conditions*,

$$\begin{aligned} c^{-\sigma} - V_k &= 0, & (1.82)\\ w(t)h(t)V_k - \delta h(t) V_h &= 0. & (1.83) \end{aligned}$$

The second step is to obtain the *evolution of the costate variables*. For this we use the maximized Bellman equation and the envelope theorem to obtain for physical capital

$$\begin{aligned} \rho V_k &= (r(t)k(t) + w(t)u(x)h(t) - c(x))V_{kk} + r(t)V_k + (1-u(x))\delta h(t) V_{hk} \\ &= \dot{k} V_{kk} + r(t) V_k + \dot{h} V_{hk}, \end{aligned}$$

and for human capital

$$\begin{aligned} \rho V_h &= (r(t)k(t) + w(t)u(x)h(t) - c(x))V_{hk} + w(t)u(x)V_k + (1-u(x))\delta V_h \\ &\quad +(1-u(x))\delta h(t) V_{hh} \\ &= \dot{k} V_{hk} + w(t)u(x)V_k + (1-u(x))\delta V_h + \dot{h} V_{hh}. \end{aligned}$$

Observing that the costate variables obey

$$\dot{V}_k = \dot{k} V_{kk} + \dot{h} V_{kh}, \quad \text{and} \quad \dot{V}_h = \dot{k} V_{kh} + \dot{h} V_{hh},$$

we may write the evolution of the two costate variables as

$$\dot{V}_k = (\rho - r(t))V_k, \quad \dot{V}_h = (\rho - (1 - u(x))\delta)V_h - w(t)u(x)V_k.$$

As the final step we use the first-order conditions (1.82) and (1.83) to substitute costates V_h and V_k to obtain *Euler equations* for optimal consumption,

$$-\sigma c^{-\sigma-1}\dot{c} = (\rho - r(t))c^{-\sigma} \quad \Rightarrow \quad \dot{c} \;=\; \frac{r(t) - \rho}{\sigma}c(t), \tag{1.84}$$

and for the optimal time allocated to production

$$\begin{aligned} \dot{w}(t)V_k + w(t)(\rho - r(t))V_k &= (\rho - (1 - u(x))\delta)V_h\delta - w(t)u(x)V_k\delta \\ \Leftrightarrow \;\; \dot{w}(t)/w(t) - r(t) &= -(1 - u(x))\delta - u(x)\delta \\ \Leftrightarrow \;\; \dot{u}/u(t) + \dot{h}/h(t) &= \delta/\alpha - c(t)/k(t) \\ \Rightarrow \quad \dot{u} &= \left(\frac{1 - \alpha}{\alpha}\delta - c(t)/k(t) + u(t)\delta\right) u(t). \end{aligned} \tag{1.85}$$

Together with appropriate transversality conditions, initial conditions, and the constraints in (1.80), the Euler equations describe the equilibrium dynamics.

We may summarize the reduced form system as

$$\begin{aligned} \dot{k} &= r(t)k(t) + w(t)u(t)h(t) - c(t), \\ \dot{h} &= (1 - u(t))\delta h(t), \\ \dot{c} &= (r(t) - \rho)c(t)/\sigma, \\ \dot{u} &= \left(\tfrac{1-\alpha}{\alpha}\delta - c(t)/k(t)\right) u(t) + u^2(t)\delta. \end{aligned}$$

and the transversality condition reads (Benhabib and Perli 1994, p.117)

$$\lim_{t\to\infty} \left[V_k e^{-\rho t}[k(t) - k^*(t)] + V_h e^{-\rho t}[h(t) - h^*(t)]\right] \geq 0,$$

for all admissible $k(t)$ and $h(t)$.

1.7 Basic concepts of probability theory

Literature: Karlin and Taylor (1975, chap. 1,2), Spanos (1999, chap. 3,4,8), Ljungqvist and Sargent (2004, chap. 2.1 to 2.3),

This section contains a brief review of the basic elementary notions and terminology of probability theory for later reference. The following concepts will be assumed familiar to the reader. More detailed treatments of these topics can be found in any good standard text for a course in probability theory (our notation closely follows Spanos 1999).

1.7.1 The notion of a probability model

Definition 1.7.1 (Probability space) *The trinity $(\Omega, \mathfrak{F}, P)$ where Ω is the sample space (outcomes set), $\mathfrak{F}$ is an event space associated with Ω, and P is a probability function from $\mathfrak{F}$ to the real numbers between 0 and 1 satisfying axioms*

1. $P(\Omega) = 1$,
2. $P(A) \geq 0$ *for any event* $A \in \mathfrak{F}$,
3. *for a countable sequence of mutually exclusive events* $A_1, A_2, ... \in \mathfrak{F}$, *countable additivity* $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$, *holds,*

is referred to as a probability space.

Definition 1.7.2 (Random variable) *A random variable on the probability space $(\Omega, \mathfrak{F}, P)$ is a function $X : \Omega \to \mathbb{R}$ that satisfies the restriction $X \leq x := \{\omega : X(\omega) \leq x\} \in \mathfrak{F}$ for all $x \in \mathbb{R}$. A random variable X is said to be continuous if its range is any uncountable subset of $\mathbb{R}$. If the subset is countable, the random variable X is said to be discrete.*

Definition 1.7.3 (Cumulative distribution function) *We refer to*

$$F_X : \mathbb{R} \to [0, 1], \quad F_X(x) = P(X \leq x)$$

as the cumulative distribution function (cdf) of the random variable X.

Remark 1.7.4 *The properties of the cdf $F_X(x)$ of the random variable X are*

1. $F_X(x) \leq F_X(y)$, *for* $x \leq y$, $x, y \in \mathbb{R}$,
2. $\lim_{x \searrow x_0} F_X(x) = F_X(x_0)$, *for any* $x \in \mathbb{R}$,
3. $\lim_{x \to \infty} F_X(x) := F_X(\infty) = 1$, $\lim_{x \to -\infty} := F_X(-\infty) = 0$,

i.e., F_X is a non-decreasing, right-continuous function with $F_X(\infty) = 1$, $F_X(-\infty) = 0$.

Definition 1.7.5 (Density function) *Assuming that there exists a function of the form*

$$f_X : \mathbb{R} \to [0, \infty) \quad \textit{such that} \quad F_X(x) = \int_{-\infty}^{x} f_X(u)du, \quad \textit{where} \quad f_X(u) \geq 0,$$

f_X is said to be a density function of the random variable X which corresponds to F_X.

Remark 1.7.6 *The density function, for a continuous random variable, satisfies*

1. *$f_X(x) \geq 0$ for all $x \in \mathbb{R}$,*
2. *$\int_{-\infty}^{\infty} f_X(x)dx = 1$,*
3. *$F_X(b) - F_X(a) = \int_a^b f_X(x)dx$ for $a < b$, $a, b \in \mathbb{R}$,*
4. *$P(X = x) = 0$ for all $x \in \mathbb{R}$.*

Definition 1.7.7 (Probability mass function) *Assuming that there exist a function*

$$f_X : \mathbb{R} \to [0, 1] \quad \textit{such that} \quad F_X(x) = \sum_{u : u \leq x} f_X(u) \quad \textit{where} \quad 1 \geq f_X(u) \geq 0.$$

f_X is said to be a probability mass function of the random variable X corresponding to F_X.

Remark 1.7.8 *In that the cdf for a discrete random variable is a step function with the jumps defined by f_X. The probability mass function, for a discrete random variable, satisfies*

1. *$f_X(x) \geq 0$ for all $x \in \mathbb{R}$,*
2. *$\sum_{x_i \in \mathbb{R}} f_X(x_i) = 1$,*
3. *$F_X(b) - F_X(a) = \sum_{a < x_i \leq b} f_X(x_i)$ for $a < b$, $a, b \in \mathbb{R}$,*
4. *$P(X = x) = f_X(x)$ for all $x \in \mathbb{R}$.*

In the literature, the probability mass function is also referred to as the density function for discrete random variables.

Example 1.7.9 *The Normal distribution has the density function*

$$f_X(x) = \frac{1}{\sqrt{2\sigma^2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad \mu \in \mathbb{R}, \ \sigma^2 \in \mathbb{R}_+, \quad x \in \mathbb{R},$$

with the cdf

$$F_X(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\sigma^2\pi}} e^{-\frac{(s-\mu)^2}{2\sigma^2}} ds, \quad \mu \in \mathbb{R}, \ \sigma^2 \in \mathbb{R}_+, \quad x \in \mathbb{R}.$$

Example 1.7.10 *The continuous Uniform distribution has the density function*

$$f_X(x) = \frac{1}{b-a}, \quad a, b \in \mathbb{R}, \quad x \in [a, b],$$

and $f_X(x) = 0$ for $x \notin [a, b]$. The cdf is explicitly available and reads

$$F_X(x) = \int_a^x \frac{1}{b-a} du \;=\; \frac{x-a}{b-a}, \quad a, b \in \mathbb{R}, \quad x \in [a, b],$$

and $F_X(x) = 0$ for $x < a$, $F_X(x) = 1$ for $x > b$.

Example 1.7.11 *The discrete Poisson distribution has the probability mass function*

$$f_X(x) = \frac{e^{-\lambda}\lambda^x}{x!}, \quad \lambda \in \mathbb{R}_+, \quad x = 1, 2, ...,$$

and $f_X(x) = 0$ for $x \notin 1, 2, ...$. The cdf reads

$$F_X(x) = \sum_{k=0}^{x} \frac{e^{-\lambda}\lambda^x}{k!} \quad \lambda \in \mathbb{R}_+, \quad x = 1, 2, ... \, .$$

Remark 1.7.12 *The support of the density f_X (or of the probability mass function) is the range of values of the random variable X for which the density function is positive,*

$$\mathbb{R}_X := \{x \in \mathbb{R} : f_X(x) > 0\},$$

The subscript X for f_X will usually be omitted unless there is a possible ambiguity.

Definition 1.7.13 (Probability model) *A collection of density functions or cumulative distribution functions indexed by a set of unknown parameters θ, one density for each possible value of θ in the d-dimensional parameter space $\Theta \subset \mathbb{R}^d$,*

$$\{f(x; \theta), \theta \in \Theta, x \in \mathbb{R}_X\} \quad \text{or} \quad \{F(x; \theta), \theta \in \Theta, x \in \mathbb{R}_X\}$$

is referred to as a probability model.

Example 1.7.14 *The probability model of a Binomial distribution is*

$$f(x; \theta) = \binom{n}{x} \theta^x (1-\theta)^{n-x}, \quad 0 < \theta < 1, \quad 0 \le x \le n, \; n = 1, 2, ... \, .$$

Definition 1.7.15 (Expectation operator) *Let X be a random variable and $f(x;\theta),\ \theta \in \Theta$ an associated parametric family of densities, then $E(\cdot)$ is the expectation operator*

$$\begin{aligned} E(X) &= \int_{-\infty}^{\infty} x f(x;\theta)dx, \quad \textit{for continuous random variables,} \\ E(X) &= \sum_{x_i \in \mathbb{R}_X} x_i f(x_i;\theta)dx, \quad \textit{for discrete random variables.} \end{aligned}$$

Remark 1.7.16 *For random variables X_1 and X_2 and the constants a, b, and c, $E(\cdot)$ satisfies the following properties of a linear operator,*

1. $E(c) = c$,
2. $E(aX_1 + bX_2) = aE(X_1) + bE(X_2)$.

Definition 1.7.17 (Variance operator) *Let X be a random variable and $f(x;\theta),\ \theta \in \Theta$ an associated parametric family of densities, then $Var(\cdot)$ is the variance operator*

$$\begin{aligned} Var(X) &= \int_{-\infty}^{\infty} (x - E(X))^2 f(x;\theta)dx, \quad \textit{for continuous random variables,} \\ Var(X) &= \sum_{x_i \in \mathbb{R}_X} (x_i - E(X))^2 f(x;\theta)dx, \quad \textit{for discrete random variables.} \end{aligned}$$

Remark 1.7.18 *For stochastically independent random variables X_1, X_2 and the constants a, b, and c, $Var(\cdot)$ satisfies the following properties*

1. $Var(c) = 0$,
2. $Var(aX_1 + bX_2) = a^2 Var(X_1) + b^2 Var(X_2)$.

Definition 1.7.19 (Raw moments) *A generalization of the mean is the definition of raw moments,*

$$\mu_r'(\theta) \equiv E(X^r) = \int_{-\infty}^{\infty} x^r f(x;\theta)dx, \quad r = 1, 2, ... \ .$$

Definition 1.7.20 (Central moments) *A direct generalization of the variance is central moments,*

$$\mu_r(\theta) \equiv E((X - \mu)^r) = \int_{-\infty}^{\infty} (x - \mu)^r f(x;\theta)dx, \quad r = 1, 2, ...,$$

where $\mu \equiv E(X) = \mu_1'$ denotes the mean.

Definition 1.7.21 (Joint distribution function) *Given a pair (X, Y) of random variables, their joint distribution function is the function F_{XY} of two real variables,*

$$F_{XY} : \mathbb{R}_X \times \mathbb{R}_Y \to [0, 1], \quad F_{XY}(x, y) = P(X \leq x, Y \leq y).$$

The function $F(x, \infty) \equiv \lim_{y\to\infty} F(x, y)$ is called the marginal distribution function of X. Similarly, the function $F(\infty, y) \equiv \lim_{x\to\infty} F(x, y)$ is called the marginal distribution of Y.

Remark 1.7.22 *In the case where a subset $T \in \mathbb{R}_X \times \mathbb{R}_Y$ is countable and the probability $P((X, Y) \in T) = 1$, the joint probability mass function (joint density function) is*

$$f_{XY}(x, y) = P(X = x, Y = y).$$

In the continuous case there is a joint density function $f_{XY} : \mathbb{R}_X \times \mathbb{R}_Y \to \mathbb{R}_+$ with

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x, y)dxdy = 1 \quad \text{and} \quad F_{XY}(x, y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f_{XY}(u, v)dudv,$$

where $F_{XY}(x, y)$ denotes the joint cumulative distribution function.

Definition 1.7.23 (Covariance and correlation) *If X and Y are jointly distributed random variables, their covariance is the product moment*

$$Cov(X, Y) = E\left[(X - E(X))(Y - E(Y))\right].$$

For the case where $Cov(X, Y) = 0$ the random variables X and Y are said to be uncorrelated. The correlation between X and Y is

$$Corr(X, Y) = \frac{Cov(X, Y)}{\sqrt{Var(X)Var(Y)}}.$$

Remark 1.7.24 *For random variables X, Y and Z and the constants a, b, and c, $Cov(\cdot)$ satisfies the following properties*

1. $Cov(X, Y) = E(XY) - E(X)E(Y)$,
2. $Cov(X, Y) = Cov(Y, X)$,
3. $Cov(aX + bY, Z) = aCov(X, Z) + bCov(Y, Z)$,
4. $Var(aX + bY) = a^2Var(X) + b^2Var(Y) + 2abCov(X, Y)$.

Definition 1.7.25 (Conditional probability) *The formula for the conditional probability of event A given event B takes the form*

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \quad P(B) > 0.$$

Definition 1.7.26 (Conditional distribution functions) *Let X and Y be random variables. If X and Y can attain only countably many different values and a joint probability function exists, the conditional probability gives rise to the formula*

$$f_{Y|X}(y|x) = \frac{f_{XY}(x,y)}{f_X(x)}, \quad f_X(x) > 0, \ y \in \mathbb{R}_Y,$$

where $f_{Y|X}$ denotes the conditional density of Y given that $X = x$ and

$$F_{Y|X}(y|x) = \frac{P(Y \leq y, X = x)}{P(X = x)}, \quad P(X = x) > 0,$$

denotes the conditional cdf. For uncountable many different values we define

$$f_{Y|X}(y|x) = \frac{f_{XY}(x,y)}{f_X(x)} \quad \text{and} \quad F_{Y|X}(y|x) = \int_{-\infty}^{y} f_{Y|X}(v|x)dv$$

as the conditional density and conditional cdf, respectively.

Definition 1.7.27 (Stochastic independence) *The random variables X and Y are said to be stochastically independent if for any events A and B,*

$$P(X \in A, Y \in B) = P(X \in A)P(Y \in B)$$

Remark 1.7.28 *The stochastic independence of X and Y is equivalent to the factorization of the joint distribution function,*

$$F_{XY}(x,y) = F_X(x)F_Y(y), \quad x \in \mathbb{R}_X, \ y \in \mathbb{R}_Y.$$

In particular, $E(XY) = E(X)E(Y)$, or $Cov(X,Y) = E(XY) - E(X)E(Y) = 0$. Similarly, stochastic independence implies that for all $x \in \mathbb{R}_X, \ y \in \mathbb{R}_Y$,

$$f_{XY}(x,y) = f_X(x)f_Y(y), \quad f_{X|Y}(x|y) = f_X(x), \quad f_Y(y) > 0.$$

Remark 1.7.29 *From $Cov(X,Y) = Corr(X,Y) = 0$ does not follow that X and Y are stochastically independent. For example, suppose $X \sim N(0,1)$ and $Y = h(X) = X^2$. Observe* *that $E(XY) = X^3 = 0 = E(X)E(Y)$ and therefore $Cov(X,Y) = Corr(X,Y) = 0$.*

Definition 1.7.30 (Random sample) *The sample $X_{(n)}^{iid} = (X_1, X_2, ..., X_n)$ is called a random sample if the random variables $(X_1, ..., X_n)$ are independent and identically distributed.*

Definition 1.7.31 (Sampling model) *A sampling model is a set of random variables $(X_1, X_2, ..., X_n)$, a sample, with a certain probabilistic structure. The primary objective of the sampling model is to relate the observed data to the probability model.*

Remark 1.7.32 *A statistical model therefore consists of both a probability model and a sampling model. Particular examples are given below.*

- Bernoulli model

$$\text{Probability model:} \qquad \left\{f(x;\theta) = \theta^x (1-\theta)^{1-x}, 0 \leq \theta \leq 1, x = \{0,1\}\right\}$$

- Normal model

$$\text{Probability model:} \qquad \left\{f(x;\theta) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \theta = (\mu, \sigma^2) \in \mathbb{R} \times \mathbb{R}_+, x \in \mathbb{R}\right\}$$

In each case the sampling model considers $(X_1, X_2, ..., X_n)$ as a random sample.

1.7.2 Functions of random variables

Often, we find ourselves faced with functions of one or several random variables whose distribution we need but we only know the distribution of the original variables. Below we replicate techniques how to obtain the distributions (cf. Spanos 1999, chap. 11.7.1).

Theorem 1.7.33 (Change of variables for densities) *Let X be a continuous random variable defined on $\mathbb{R}_X = (a,b)$ where $a, b \in \mathbb{R}$, $a < b$ with density $f_X(x)$. Let Y be a random variable defined $Y = h(X)$ where $h(\cdot)$ is strictly monotonic and $h(\cdot)$ has a differentiable inverse. Then the density function of Y is*

$$f_Y(y) = f_X(h^{-1}(y)) \left|\frac{dh^{-1}(y)}{dy}\right|, \quad \mathbb{R}_Y = (h(a), h(b)),$$

which follows from $F_Y(y) = F_X(h^{-1}(y))$ using the chain rule.

Example 1.7.34 *Consider the random variable* $Y = \exp(X)$ *with* X *on* $\mathbb{R}_X = (-\infty, \infty)$ *being normally distributed. From Theorem 1.7.33 the density of* Y *is given by,*

$$f_Y(y) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(\ln y-\mu)^2}{2\sigma^2}} \frac{d\ln(y)}{dy} = \frac{1}{y}\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(\ln y-\mu)^2}{2\sigma^2}},$$

where $\mathbb{R}_Y = (e^{\infty}, e^{-\infty}) = (\infty, 0)$*, which is the density of the Log-Normal distribution.*

Remark 1.7.35 *If* X *is a continuous random variable with the density function* $f_X(x)$ *and* g *is a function, then* $Y = g(X)$ *is also a random variable and the expectation of* $g(X)$ *is computed from*

$$E(Y) = E(g(X)) = \int_{-\infty}^{\infty} g(x) f_X(x) dx.$$

Theorem 1.7.36 (Change of variables for joint densities) *Let* X *and* Y *be continuous random variables defined on* $\mathbb{R}_X$ *and* $\mathbb{R}_Y$*, respectively. Let* $Z = h(X, Y)$ *be a function of* X *and* Y*. Its cdf can be derived via*

$$F_Z(z) = P(Z \leq z) = P(h(X,Y) \leq z) = \int\int_{\{(x,y): h(x,y)\leq z\}} f(x,y) dx dy.$$

Example 1.7.37 *Consider the convolution of two independent random variables* X *and* Y*, that is* $Z = X + Y$ *where the density functions take the form*

$$f_X(x) = e^{-x},\ x > 0, \quad f_Y(y) = e^{-y},\ y > 0, \quad \mathbb{R}_Z = (0, \infty).$$

Using the general result of Theorem 1.7.36 as follows

$$\begin{aligned} F_Z(z) &= \int_0^z \int_0^{z-x} e^{-x-y} dx dy = -\int_0^z e^{-x-y} dx\Big|_0^{z-x} = \int_0^z e^{-x}\left(1 - e^{x-z)}\right) dx \\ &= \int_0^z e^{-x} dx - e^{-z} z = -\left(e^{-z} - 1\right) - e^{-z} z = 1 - e^{-z} - z e^{-z}, \end{aligned}$$

and the density function is $f_Z(z) = e^{-z} + ze^{-z} - e^{-z} = ze^{-z},\ z > 0.$

1.7.3 Stochastic processes

The basic concept required when working with models under uncertainty is that of stochastic processes, which extends the notion of a random variable.

Definition 1.7.38 (Stochastic process) *A stochastic process is simply an indexed collection* $\{X_t\}_{t\in\mathbb{T}}$ *of random variables defined on the same probability space* $(\Omega, \mathfrak{F}, P)$*, i.e.* X_t *is a random variable relative to* $(\Omega, \mathfrak{F}, P)$*, for each* t *in the index set* $\mathbb{T}$ *(henceforth time).*

Remark 1.7.39 *We refer to the range of the random variable defined by the union of the sets of values of* $\mathbb{R}_{X(t)}$ *for each* t *as the state space of the stochastic process,* $\mathcal{R} = \cup_{t\in\mathbb{T}}\mathbb{R}_{X(t)}$.

Definition 1.7.40 *In the case where index set* $\mathbb{T}$ *is countable, we call* $\{X_t\}_{t\in\mathbb{T}}$ *a discrete-time stochastic process. On the other hand, when* $\mathbb{T}$ *is an uncountable set, such as an interval on the real line, we call* $\{X_t\}_{t\in\mathbb{T}}$ *a continuous-time stochastic process.*

Definition 1.7.41 *In the case where the state space* $\mathcal{R}$ *is a countable space, we call* $\{X_t\}_{t\in\mathbb{T}}$ *a discrete-state stochastic process. On the other hand, when* $\mathcal{R}$ *is an uncountable space we call* $\{X_t\}_{t\in\mathbb{T}}$ *a continuous-state stochastic process.*

We proceed to define some dependence restrictions that will be useful in later applications, because the concept of purely independent stochastic processes or random samples appears to be too restrictive. Below we use a discrete-time notation but with minor modifications, the concepts can be written in the more general notation $0 < t_1 < t_2 < ... < t_n < \infty$.

Definition 1.7.42 (Independence) *The stochastic process* $\{X_t\}_{t\in\mathbb{T}}$ *is said to be independent if the joint density function* $f(x_1, x_2..., x_T) := f_{\{X_t\}_{t\in\mathbb{T}}}(x_1, x_2..., x_T)$ *can be factorized,*

$$f(x_1, x_2, ..., x_T) = \prod_{i=1}^{T} f_i(x_i), \quad (x_1, x_2, ..., x_T) \in \mathcal{R},$$

or similarly, the conditional density equals the unconditional density for $\tau > 0$,

$$f(x_{k+\tau}|x_k, x_{k-1}, ..., x_1) = f(x_{k+\tau}), \quad k = 1, 2, ... \, .$$

Definition 1.7.43 (Asymptotic independence) *The stochastic process* $\{X_t\}_{t\in\mathbb{T}}$ *is said to be asymptotically independent if as* $\tau \to \infty$ $f(x_{k+\tau}|x_k, x_{k-1}, ..., x_1) \simeq f(x_{k+\tau})$, $k = 1, 2, ...$, *i.e., elements become independent as the distance between them increases to infinity.*

Definition 1.7.44 (Markov dependence) *The stochastic process* $\{X_t\}_{t\in\mathbb{T}}$ *is said to be Markov dependent if* $f(x_{k+1}|x_k, x_{k-1}, ..., x_1) = f(x_{k+1}|x_k)$, $k = 1, 2, ...$.

This notion of dependence can be easily extended to higher-orders m. Then the stochastic process is said to be Markov dependent of order m. A similar concept based on the first two moments is the case of *non-correlation* or no linear dependence, which can be extended to *non-correlation of order* m or even to *asymptotic non-correlation* (Spanos 1999, chap. 8.4).

Markov chains

Probably the most well-known stochastic process is the so-called *Markov chain*, which is a Markov process whose distribution (state space) is *discrete* (countable) while the time dimension can either be discrete or continuous.

Definition 1.7.45 (Markov chain) *The stochastic process* $\{X_t\}_{t\in\{0,1,2,...\}}$ *is said to be a Markov chain if for arbitrary times* $0 \le t_1 < t_2 < ... < t_n$

$$P(X_{t_n} = x_n | X_{t_{n-1}} = x_{n-1}, X_{t_{n-2}} = x_{n-2}, ..., X_{t_1} = x_1) = P(X_{t_n} = x_n | X_{t_{n-1}} = x_{n-1}).$$

The joint distribution of the process takes the form

$$P(X_{t_n} = x_n, X_{t_{n-1}} = x_{n-1}, ..., X_{t_1} = x_1) = P(X_{t_1} = x_1) \prod_{k=2}^{n} P(X_{t_k} = x_k | X_{t_{k-1}} = x_{k-1}),$$

where $P(X_{t_1} = x_1)$ *is the initial condition, and* $p_{ij}^{(k)} := P(X_{t_k} = j | X_{t_{k-1}} = i)$, $k = 2, 3, ...$, *the one-step transition probabilities from state* i *to state* j.

Remark 1.7.46 *A particular important case is when the process is homogeneous in time,* $p_{ij}^{(k)} = p_{ij}$, *for all* $k = 2, 3, ...$, *the transition probabilities do not change over time. In this case the* n*-step transition probabilities is obtained from the one-step transition probabilities.*

Definition 1.7.47 (Martingale) *A stochastic process* $\{X_t\}_{t\in\mathbb{N}}$ *with* $E(|X_t|) < \infty$ *for all* $t \in \mathbb{N}$ *is said to be a martingale if* $E(X_t | X_{t-1}, X_{t-2}, ..., X_1) = X_{t-1}$.

Remark 1.7.48 *A stochastic process* $\{X_t\}_{t\in\mathbb{N}}$ *with* $E(|X_t|) < \infty$ *for all* $t \in \mathbb{N}$ *is said to be a martingale difference if* $E(X_t | X_{t-1}, X_{t-2}, ..., X_1) = 0$. *The term martingale difference stems from the fact that this process can always be generated as a difference of a martingale process* $\{Y_t\}_{t\in\mathbb{N}}$, *defining the process* $\{X_t := Y_t - Y_{t-1}\}_{t\in\mathbb{N}}$,

$$E(X_t | X_{t-1}, X_{t-2}, ..., X_1) = E(Y_t | Y_{t-1}, Y_{t-2}, ..., Y_1) - Y_{t-1} = 0.$$

Reversing the argument, $\{Y_t = \sum_{k=1}^{t} X_k\}_{t\in\mathbb{N}}$ *is a martingale.*

Chapter 2

Stochastic models in discrete time

2.1 Topics in difference equations

Literature: Ljungqvist and Sargent (2004, chap. 2), Sydsæter et al. (2005, chap. 11)
The objective of this chapter is mainly to recall basic concepts on difference equations. We start after some preliminary definitions with deterministic difference equation and include stochastics afterwards.

2.1.1 Definitions

Let $t = 0, 1, 2, ...$ denote different discrete-time periods. Usually we refer to $t = 0$ as the *initial period*. If $x(t)$ is a function defined for $t = 0, 1, 2, ...$, we use $x_0, x_1, x_2, ...$ to denote $x(0), x(1), x(2), ...$, and in general we write x_t for $x(t)$.

Definition 2.1.1 *Let denote Δ a linear operator that satisfies the properties*

$$\begin{aligned} \Delta x_t &= x_t - x_{t-1}, \quad \Delta^0 x_t = x_t, \\ \Delta^2 x_t &= \Delta(\Delta x_t) = \Delta x_t - \Delta x_{t-1} = x_t - 2x_{t-1} + x_{t-2}, \\ \Delta^k x_t &= \Delta\left(\Delta^{k-1} x_t\right), \quad k = 2, 3, ... \, . \end{aligned}$$

We refer to Δ as the difference operator.

Remark 2.1.2 *Let $a, b \in \mathbb{R}$ and $t = 0, 1, 2, ...,$*

$$\begin{aligned} \Delta a &= 0, \\ \Delta(ay_t + bx_t) &= a\Delta y_t + b\Delta x_t, \\ \Delta^k t^k &= k, \quad k = 0, 1, 2, ..., \\ \Delta^{k+1} t^k &= 0, \quad k = 0, 1, 2, ... \ . \end{aligned}$$

Definition 2.1.3 *Let denote L a linear operator defined by*

$$L^k x_t \;=\; (1 - \Delta)^k x_t \;=\; x_{t-k}, \quad L^0 = 1, \; k \geq 0.$$

We refer to L as the lag operator or backshift operator.

Remark 2.1.4 *Let $a \in \mathbb{R}$ and $S = \sum_{i=0}^{k} (aL)^i$,*

$$\begin{aligned} (aL)S &= \sum_{i=1}^{k+1} (aL)^i \;\Rightarrow\; (1 - aL)S = 1 - a^{k+1} L^{k+1} \\ \Leftrightarrow \sum_{i=0}^{k} (aL)^i &= \frac{1 - a^{k+1} L^{k+1}}{1 - aL}. \end{aligned} \tag{2.1}$$

Lemma 2.1.5 *For any $a \neq 1$,*

$$\sum_{i=0}^{k} a^i = \frac{1 - a^{k+1}}{1 - a}.$$

Proof. An immediate implication of the geometric series in (2.1). ■

Lemma 2.1.6 *For any $a \neq 1$,*

$$\sum_{i=0}^{k} i a^i = \frac{1}{1 - a} \left(\frac{a - a^{k+1}}{1 - a} - k a^{k+1} \right).$$

Proof. By inserting and collecting terms. ■

Definition 2.1.7 *A difference equation is an equation for $t \in \mathbb{Z}$ of either type*

$$G(t, x_t, \Delta x_t, \Delta^2 x_t, ..., \Delta^k x_t) \;=\; 0, \tag{2.2}$$

$$F(t, L^0 x_t, ..., L^k x_t) = F(t, x_t, ..., x_{t-k}) \;=\; 0, \tag{2.3}$$

where k denotes the order, if k denotes the maximum time difference with respect to x.

Definition 2.1.8 *A linear difference equation of order k for $t \in \mathbb{Z}$ of the type*

$$x_t + c_1 x_{t-1} + ... + c_k x_{t-k} = r_t, \quad c_1, ..., c_k \in \mathbb{R}, \tag{2.4}$$

where r_t is a given series is referred to as the normal form. For $r_t = 0$ the difference equation is called homogeneous, for $r \neq 0$ we refer to (2.4) as inhomogeneous.

2.1.2 Deterministic difference equations

The following results and solution techniques are useful and are reproduced for later reference.

First-order linear difference equations

Assume in the following a first-order difference equation of the type (2.4),

$$x_t = a x_{t-1} + r_t, \quad a \in \mathbb{R}. \tag{2.5}$$

Given an initial value x_0, we can solve (2.5) iteratively,

$$\begin{aligned} x_1 &= a x_0 + r_1, \\ x_2 &= a x_1 + r_2 = a^2 x_0 + a r_1 + r_2, \\ &\vdots \\ x_t &= a^t x_0 + \sum_{i=0}^{t-1} a^i r_{t-i}. \end{aligned}$$

Exercise 2.1.9 *Obtain the general solution to (2.5) using the lag operator.*

For $a \neq 1$ and $r_t = r$ we obtain,

$$x_t = a^t x_0 + r \sum_{i=0}^{t-1} a^i = a^t x_0 + r \frac{1 - a^t}{1 - a} = \left(x_0 - \frac{r}{1-a} \right) a^t + \frac{r}{1-a}.$$

- For $|a| < 1$ the series converges to the *equilibrium* or stationary state

  $$\lim_{t \to \infty} x_t = \lim_{t \to \infty} \left(x_0 - \frac{r}{1-a} \right) a^t + \frac{r}{1-a} = \frac{r}{1-a},$$

  and the equation is called *globally asymptotically stable*. Two kinds of stability can be identified, either x_t converges monotonically to the equilibrium ($a > 0$), or x_t exhibits decreasing fluctuations or *damped oscillations* around the equilibrium state ($a < 0$).

- For $|a| > 1$ the series diverges from the equilibrium state, except when the initial value is the equilibrium state. Either x_t increases monotonically ($a > 0$), or x_t exhibits increasing fluctuations or *explosive oscillations* around the equilibrium state ($a < 0$).

For $a = 1$ and $r_t = r$ we obtain,

$$x_t = x_0 + r \sum_{i=0}^{t-1} 1^i = x_0 + rt,$$

i.e., the solution is a linear function of t.

Example 2.1.10 (Intertemporal budget constraint) *Let a_t denote the value of assets held at the end of period t. Further, let c_t be the amount withdrawn for consumption and w_t the labor income during period t. Suppose $r > 0$ is the constant interest rate,*

$$a_t = (1+r)a_{t-1} + w_t - c_t, \quad a(0) = a_0.$$

Using the property of the lag operator (2.1), we obtain the solution

$$\left(\frac{1}{1+r}\right)^t a_t = a_0 + \sum_{i=0}^{t-1} \left(\frac{1}{1+r}\right)^{t-i} (w_{t-i} - c_{t-i}),$$

which denotes the present discounted value of the assets in the account at time t.

Exercise 2.1.11 (Cobweb model) *Let total demand and supply of a nondurable good be*

$$\begin{aligned} D &= a - bp_t, \quad a, b > 0, \\ S &= c + dp_{t-1}, \quad c, d > 0. \end{aligned}$$

Market clearing demands that $D(p_t) = S(p_{t-1})$. Obtain the equilibrium price.

Second-order linear difference equations

The solution techniques for solving difference equations are similar to differential equations. Compare the following two theorems to second-order differential equations. Consider a linear second-order difference equation in normal form (2.4),

$$x_{t+2} + ax_{t+1} + bx_t = c_t, \quad a, b \in \mathbb{R}, \ b \neq 0. \tag{2.6}$$

Suppose we found the particular solution of (2.6), e.g., for the case where $c_t = c$ we obtain $u^* = c/(1+a+b)$. The general solution of (2.6) requires to solve the homogeneous equation,

$$x_{t+2} + ax_{t+1} + bx_t = 0, \quad a, b \in \mathbb{R},\ b \neq 0. \tag{2.7}$$

We try to find solutions of the form $x_t = \lambda^t$. Inserting these expressions into (2.6) yields the characteristic equation $\lambda^t(\lambda^2 + a\lambda + b) = 0$. The following theorem considers three cases.

Theorem 2.1.12 (cf. Sydsæter et al. (2005), Theorem 11.4.1) *The equation*

$$x_{t+2} + ax_{t+1} + bx_t = 0, \quad a, b \in \mathbb{R},\ b \neq 0 \quad \textit{has the following general solution.}$$

(i) If $a^2 - 4b > 0$, there are two distinct real roots,

$$x(t) = c_1\lambda_1^t + c_2\lambda_2^t, \quad \textit{where} \quad \lambda_{1,2} = -\frac{a}{2} \pm \frac{1}{2}\sqrt{a^2 - 4b}.$$

(ii) If $a^2 - 4b = 0$, there is one real double root,

$$x(t) = c_1\lambda_1^t + c_2 t\lambda_2^t, \quad \textit{where} \quad \lambda_1 = \lambda_2 = -\frac{a}{2}.$$

(iii) If $a^2 - 4b < 0$, there are two conjugate complex roots,

$$x(t) = r^t(c_1 \cos\theta t + c_2 \sin\theta t), \quad \textit{where} \quad r = \sqrt{b}, \quad \cos\theta = -\frac{a}{2\sqrt{b}}, \ \theta \in [0, \pi],$$

where $c_1, c_2 \in \mathbb{R}$.

Suppose an economy evolves according to a system of difference equations. If appropriate initial conditions are imposed, then the linear system has a unique solution. An important question is whether small changes in the initial conditions have any effect on the long-run behavior of the solution. If small changes in the initial conditions lead to significant differences in the long run behavior of the solution, then the system is *unstable.* If the effect dies out as time approaches infinity, the system is called stable.

Consider in particular the second-order difference equation (2.6). If the general solution of the associated homogeneous equation (2.7) tends to 0 as $t \to \infty$, for all values of arbitrarily chosen constants, then the equation is called *globally asymptotically stable.*

Theorem 2.1.13 (cf. Sydsæter et al. (2005), Theorem 11.4.2) *The equation*

$$x_{t+2} + ax_{t+1} + bx_t = c_t$$

is called globally asymptotically stable if and only if the following two equivalent conditions are satisfied

(i) The roots of the characteristic equation $\lambda^2 + a\lambda + b = 0$ *have solutions with real parts strictly less than 1 (moduli strictly less than 1).*

(ii) $|a| < 1 + b$ *and* $b < 1$.

Systems of linear difference equations

We consider a system of first-order linear difference equations in the normal form,

$$\begin{aligned} x_{1t} &= a_{11}x_{1,t-1} + a_{12}x_{2,t-1} + ... + a_{1n}x_{n,t-1} + r_{1t} \\ x_{2t} &= a_{21}x_{1,t-1} + a_{22}x_{2,t-1} + ... + a_{2n}x_{n,t-1} + r_{2t} \\ &\vdots \\ x_{nt} &= a_{n1}x_{1,t-1} + a_{n2}x_{2,t-1} + ... + a_{nn}x_{n,t-1} + r_{nt}, \end{aligned}$$

or in matrix notation,

$$x_t = Ax_{t-1} + r_t, \tag{2.8}$$

where

$$x_t = \begin{bmatrix} x_{1t} \\ \vdots \\ x_{nt} \end{bmatrix}, \quad A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}, \quad r_t = \begin{bmatrix} r_{1t} \\ \vdots \\ r_{nt} \end{bmatrix}.$$

Remark 2.1.14 *A system of* n *first-order linear difference equations can be written as an* n*-th order linear difference equation, and vice versa.*

Example 2.1.15 *Consider the system*

$$x_{1t} = a_{11}x_{1,t-1} + a_{12}x_{2,t-1}, \tag{2.9}$$

$$x_{2t} = a_{21}x_{1,t-1} + a_{22}x_{2,t-1}. \tag{2.10}$$

Using the lag operator we can write (2.10) as

$$(1 - a_{22}L)x_{2t} = a_{21}x_{1,t-1} \;\Leftrightarrow\; (1 - a_{22}L)x_{2,t-1} = a_{21}x_{1,t-2}.$$

Multiplying (2.9) by $(1 - a_{22}L)$ *and inserting the last result yields,*

$$\begin{aligned} (1 - a_{22}L)x_{1t} &= (1 - a_{22}L)a_{11}x_{1,t-1} + a_{12}a_{21}x_{1,t-2} \\ \Leftrightarrow\; x_{1t} &= (a_{11} + a_{22})x_{1,t-1} + (a_{12}a_{21} - a_{22}a_{11})x_{1,t-2}, \end{aligned}$$

which is a second-order difference equation.

Example 2.1.16 *Consider the scalar second-order autoregression,*

$$x_{t+1} = \alpha + \rho_1 x_t + \rho_2 x_{t-1} + r_{t+1}.$$

Observe that we can represent this relationship as the system

$$\begin{bmatrix} x_{t+1} \\ x_t \\ 1 \end{bmatrix} = \begin{bmatrix} \rho_1 & \rho_2 & \alpha \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_t \\ x_{t-1} \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} r_{t+1},$$

which is a system of first-order vector difference equation (2.8).

Example 2.1.17 *Consider the scalar second-order autoregression with moving average terms,*

$$x_{t+1} = \alpha + \rho_1 x_t + \rho_2 x_{t-1} + r_{t+1} + \gamma r_t.$$

Observe that we can represent this relationship as the system

$$\begin{bmatrix} x_{t+1} \\ x_t \\ r_{t+1} \\ 1 \end{bmatrix} = \begin{bmatrix} \rho_1 & \rho_2 & \gamma & \alpha \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_t \\ x_{t-1} \\ r_t \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} r_{t+1},$$

which is a system of first-order vector difference equation (2.8).

The general solution of the associated homogeneous equation system to (2.8),

$$x_t = Ax_{t-1}, \tag{2.11}$$

can be obtained as follows. We try a solution where $x_t = W\lambda^t$, for $\lambda \neq 0$ and the unknown $n \times 1$ vector $W = (w_1, ..., w_n)^\top$. Inserting into (2.11) yields

$$W\lambda^t = AW\lambda^{t-1} \;\Leftrightarrow\; (A - \lambda I)W = 0,$$

where I is the $n \times n$ identity matrix. Thus admissible solutions for λ are the eigenvalues of A and admissible solutions for W are the associated eigenvectors. Suppose rank$(A - \lambda I) < n$, the characteristic equation is a nth-order polynomial in λ,

$$\det(A - \lambda I) = 0.$$

Note that if $\text{rank}(A-\lambda I) = n$, then $\det(A-\lambda I) \neq 0$ and we would have only trivial solutions $W = 0$. The n characteristic roots are the eigenvalues of the coefficient matrix A. For any given λ_i, $i = 1, ..., n$ we obtain a linear homogeneous system in $W^{(i)}$,

$$(A - \lambda_i I)W^{(i)} = 0,$$

where $W^{(i)}$ are the eigenvectors associated with λ_i. If we obtain n distinct roots, the general solution of (2.11) reads

$$x_t = c_1\lambda_1^t W^{(1)} + c_2\lambda_2^t W^{(2)} + ... + c_n\lambda_n^t W^{(n)}, \quad c_1, ..., c_n \in \mathbb{R}.$$

A necessary and sufficient condition for the system (2.8) to be *globally asymptotically stable* is that all the eigenvalues of the matrix A have moduli (strictly) less than 1. In this case the solution of the associated homogeneous system (2.11) converges to 0.

Theorem 2.1.18 (cf. Sydsæter et al. (2005), Theorem 11.6.2) *If all the eigenvalues of $A = (a_{ij})_{n\times n}$ have moduli (strictly) less that 1, the difference equation*

$$x_t = Ax_{t-1} + r,$$

is globally asymptotically stable. Any solution x_t converges to the constant equilibrium state vector $(I - A)^{-1}r$.

The following theorem is useful to show that the coefficient matrix A has only eigenvalues with moduli less than 1.

Theorem 2.1.19 (cf. Sydsæter et al. (2005), Theorem 11.6.3) *Let $A = (a_{ij})$ be an arbitrary $n \times n$ matrix and suppose that*

$$\sum_{j=1}^{n} |a_{ij}| < 1 \quad \textit{for all} \quad i = 1, ..., n.$$

Then all eigenvalues of A have moduli less than 1.

2.1.3 Stochastic difference equations

Many macro economists would refer to first-order stochastic linear vector difference equations together with Markov chains as their workhorses (Ljungqvist and Sargent 2004). Indeed, they are useful because they describe a time series with parsimony.

Consider the system difference equation in the normal form similar to (2.8),

$$x_{t+1} = Ax_t + Cr_{t+1}, \tag{2.12}$$

for $t = 0, 1, 2, ...$ where x_t is an $n \times 1$ state vector, A is an $n \times n$ matrix, C is an $n \times m$ matrix, r_t is an $m \times 1$ vector. We now refer to r_t as a random variable satisfying certain assumptions. Because r_t is random, at least for $t \neq 0$ the state variable x_t will be a random variable as well, of which the properties are now of interest.

Definition 2.1.20 (White noise process) *Suppose that r_{t+1} is an $m \times 1$ random vector of the discrete-time and continuous-state stochastic process $\{r_t\}_{t \in \mathbb{N}}$ with*

$$\begin{aligned} E(r_{t+1}) &= 0 \quad \textit{for all} \quad t \in \mathbb{N}, \\ E(r_t r_{t-j}^\top) &= \begin{cases} I & \textit{if} \quad j = 0 \\ 0 & \textit{if} \quad j \neq 0 \end{cases}, \end{aligned}$$

where I is the $m \times m$ identity matrix. Then r_{t+1} is said to be (vector) white noise.

Definition 2.1.21 (State-space system) *Let $\{r_t\}_{t \in \mathbb{N}}$ be a white noise process. Then*

$$\begin{aligned} x_{t+1} &= Ax_t + Cr_{t+1}, \\ y_t &= Gx_t, \end{aligned}$$

where y_t is a vector of variables observed at t, is said to be a state-space system.

If y_t includes linear combinations of x_t only, it represents a *linear state-space* system.

Example 2.1.22 *Consider the scalar second-order autoregression,*

$$x_{t+1} = \alpha + \rho_1 x_t + \rho_2 x_{t-1} + r_{t+1}.$$

Observe that we can put this relationship in the form of a linear state-space representation,

$$\begin{aligned} \begin{bmatrix} x_{t+1} \\ x_t \\ 1 \end{bmatrix} &= \begin{bmatrix} \rho_1 & \rho_2 & \alpha \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_t \\ x_{t-1} \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} r_{t+1}, \\ x_t &= \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_t \\ x_{t-1} \\ 1 \end{bmatrix}. \end{aligned}$$

In analogy to deterministic difference equations, (2.12) can be solved for x_t iteratively as a function of t and realizations of r_t, provided we have a given initial vector x_0 for $t = 0$,

$$x_{t+1} = A^{t+1} x_0 + \sum_{j=0}^{t} A^j C r_{t+1-j}. \tag{2.13}$$

However, because x_t is a random variable the solution does not provide the actual value or the realization of x_t for $t \geq 0$, but gives the joint distribution of x_t at time t given the available *information set* at $t-1$ and distributional assumptions of the random shocks r_t. Often, the form (2.13) is referred to as the *moving average* representation.

Definition 2.1.23 (Impulse response function) *Suppose that the eigenvalues of A in the solution (2.13) have moduli strictly less than unity (except for the constants). Defined as a function of lag j, $h_j = A^j C$, is referred to as the impulse response function.*

Both the solution and the associated impulse response function show how x_{t+1} is affected by lagged values of the shocks. Thus, the contribution of shock r_{t-j} to x_t is $A^j C$.

Distributional properties and limiting distribution

In order to understand the distributional properties of x_t we start by assuming that $\{r_t\}_{t\in\mathbb{N}}$ is vector white noise, that is $E(r_t) = 0$, $E(r_t r_t^\top) = I$, and $E(r_t r_j^\top) = 0$ for all $j \neq t$. Apply the expectation operator to (2.12), and obtain the *mean* defining $\mu_t \equiv E(x_t)$ as

$$\begin{aligned} E(x_{t+1}) &= AE(x_t) + CE(r_{t+1}) \\ \Leftrightarrow \ \mu_{t+1} &= A\mu_t \ \Rightarrow \ \mu_t = c_1 \lambda_1^t W^{(1)} + \ldots + c_n \lambda_n^t W^{(n)}, \quad c_1, \ldots, c_n \in \mathbb{R} \end{aligned}$$

for the case where we have n distinct roots. If we assume that all of the eigenvalues of A are strictly less than unity in modulus, except possibly for one that is affiliated with the constant term, then x_t possesses a stationary mean satisfying $\mu = \mu_{t+1} = \mu_t$, or equivalently,

$$(I - A)\mu = 0.$$

It characterizes the mean μ as an eigenvector associated with the single unit eigenvalue. Notice that

$$x_{t+1} - \mu_{t+1} = A(x_t - \mu_t) + C r_{t+1}. \tag{2.14}$$

Also, the fact that the remaining eigenvalues of A are less that unity in modulus implies that starting from any μ_0, the expected value converges towards $\mu_t \to \mu$. In that we regard the initial condition x_0 as being drawn from a distribution with mean $\mu_0 = E(x_0)$.

From equation (2.14), postmultiplying both sides with $(x_{t+1}-\mu_{t+1})^\top$ and applying the expectation operator, we can compute that the stationary variance matrix satisfies

$$\begin{aligned}
E\left[(x_{t+1}-\mu_{t+1})(x_{t+1}-\mu_{t+1})^\top\right] &= E\left[(A(x_t-\mu_t)+Cr_{t+1})(A(x_t-\mu_t)+Cr_{t+1})^\top\right] \\
&= E\left[A(x_t-\mu_t)(x_t-\mu_t)^\top A^\top\right]+E\left[Cr_{t+1}r_{t+1}^\top C^\top\right] \\
&= AE\left[(x_t-\mu_t)(x_t-\mu_t)^\top\right]A^\top+CC^\top \\
\Leftrightarrow\ \gamma_{t+1,0} &= A\gamma_{t,0}A^\top+CC^\top,
\end{aligned}$$

defining the variance matrix

$$\gamma_{t,0}\equiv E\left[(x_t-\mu_t)(x_t-\mu_t)^\top\right].$$

The equation is a *discrete Lyapunov* equation in the $n\times n$ matrix $\gamma_{t,0}$. It can be solved using specialized software, or e.g., using Matlab (Ljungqvist and Sargent 2004, p.45).

Similarly, by virtue of $\mu_{t+1}=A\mu_t$ and (2.12), it can be shown that

$$\gamma_{t,j}\equiv E\left[(x_{t+j}-\mu_{t+j})(x_t-\mu_t)^\top\right]=A^j\gamma_{t,0}. \tag{2.15}$$

Once we solved for the autocovariance function $\gamma_{t,0}$, the remaining second moments $\gamma_{t,j}$ can be deduced from the given formula. The sequence of $\gamma_{t,j}$ as a function of j is also called the *autocovariogram.* Defining long-run stationary moments as

$$\mu_\infty \equiv \lim_{t\to\infty}\mu_t, \tag{2.16}$$

$$\gamma_{\infty,0} \equiv \lim_{t\to\infty}\gamma_{t,0}, \tag{2.17}$$

we pinned down two moments of the long-run or *limiting distribution* of x_t. Observe that we obtained this result without making distributional assumptions, but only requiring zero first moment and constant variance for the stochastic shocks. Further note that although the stochastic process $\{r_t\}_{t\in\mathbb{N}}$ by assumption is *white noise*, from (2.15) obviously the stochastic process $\{x_t\}_{t\in\mathbb{N}}$ has *asymptotic non-correlation.*

Exercise 2.1.24 *Consider the following first-order difference equation in matrix notation*

$$\begin{bmatrix} y_t \\ 1 \end{bmatrix}=\begin{bmatrix} a & b \\ 0 & 1 \end{bmatrix}\begin{bmatrix} y_{t-1} \\ 1 \end{bmatrix}+\begin{bmatrix} \sigma_\epsilon \\ 0 \end{bmatrix}\epsilon_t,$$

where $\{\epsilon_t\}_{t\in\mathbb{N}}$ is white noise, $E(\epsilon_t)=0$ and $E(\epsilon_t^2)=1$. Obtain the mean, $\mu_\infty\equiv\lim_{t\to\infty}\mu_t$, and the variance $\gamma_{\infty,0}\equiv\lim_{t\to\infty}\gamma_{0,t}$ of the limiting distribution.

Sunspots and indeterminacy

Let $\{\epsilon_t\}_{t\in\mathbb{T}}$ be a stochastic process and suppose there is the relationship

$$x_t = \frac{1}{a}E_t(x_{t+1}) + \epsilon_t, \quad a \neq 0. \tag{2.18}$$

A solution is an expression such that

$$E_t(x_{t+1}) = (x_t - \epsilon_t)a.$$

In what follows, we restrict our attention only to *linear* processes that fulfill the equation. Then all admissible solutions satisfy the condition,

$$x_{t+1} = (x_t - \epsilon_t)a + \tilde{\eta}_{t+1}, \tag{2.19}$$

with an arbitrary random variable $\tilde{\eta}_t$ satisfying $E_t(\tilde{\eta}_{t+1}) = 0$. Therefore the solution can be *indetermined*, because it depends on arbitrary random extrinsic events. Suppose $\tilde{\eta}_t$ is a function of *fundamental shocks* ϵ_t with $E_t(\epsilon_{t+1}) = 0$ and *sunspots* η_t with $E_t(\eta_{t+1}) = 0$,

$$\tilde{\eta}_t = \eta_t + \delta\epsilon_t, \quad \delta \in \mathbb{R}.$$

Note that for $E_t(\epsilon_{t+1}) \neq 0$, the equations demand $\delta = 0$. Assume that $|Corr(\eta_t, \epsilon_t)| < 1$, i.e., contemporaneous fundamental and sunspot shocks are not perfectly correlated.

Observe that the characteristic equation is $\lambda = a$. For the case where $|a| < 1$ the solution to the associated homogeneous equation to (2.19) is stable. If the infinite geometric series converges, it can be shown that the general solution to the inhomogeneous equation is

$$x_t = \delta\epsilon_t + (\delta - 1)\sum_{i=1}^{\infty} a^i \epsilon_{t-i} + \sum_{i=0}^{\infty} a^i \eta_{t-i}.$$

Hence, there is an infinite number of admissible solutions for different values of δ, and the solution is *indetermined*. These depend on contemporaneous fundamental shocks ($\delta \neq 0$), lagged fundamental shocks ($\delta \neq 1$) and contemporaneous and lagged sunspot shocks η_t.

If $|a| > 1$, we use the inverted lag operator $L^{-k}x_t = x_{t+k}$, solving (2.19) forward to get

$$\left(1 - (1/a)L^{-1}\right) x_t = \epsilon_t - (1/a)\tilde{\eta}_{t+1}.$$

If the infinite geometric series converges, we may write

$$x_t = \sum_{i=0}^{\infty} a^{-i} \left(\epsilon_{t+i} - (1/a)\tilde{\eta}_{t+1+i} \right).$$

Because the solution depends on future realizations (not on expected future realizations), we shift the solution one period to the future and apply the expectation operator to obtain

$$E_t(x_{t+1}) = \sum_{i=0}^{\infty} a^{-i} E_t \left(\epsilon_{t+1+i} \right),$$

where we used the law of iterated expectations. Inserting into (2.18) yields ($a \neq 0$)

$$x_t = \sum_{i=0}^{\infty} a^{-i-1} E_t \left(\epsilon_{t+1+i} \right) + \epsilon_t = \sum_{i=1}^{\infty} a^{-i} E_t \left(\epsilon_{t+i} \right) + \epsilon_t.$$

The solution depends on contemporaneous fundamental shocks, and expected future shocks ($E_t\epsilon_{t+i} \neq 0$), but not on sunspot shocks. The solution therefore is *determined.*

Exercise 2.1.25 (Inflation and indeterminacy) *Consider the following macroeconomic model (a reduced form description of a new-Keynesian model as in Clarida et al. 2000),*

$$\begin{aligned} y_t &= \psi(E_t(\pi_{t+1}) - i_t), \quad \psi > 0, \\ i_t &= \phi \pi_t, \quad \phi > 0. \end{aligned}$$

Output y_t is negatively related to the real interest rate, $i_t - E_t(\pi_{t+1})$, and the central bank responds with a Taylor rule setting the nominal interest rate i_t to the inflation rate π_t. Should the central bank respond actively to inflation ($\phi > 1$)?

2.2 Discrete-time optimization

Literature: Stokey et al. (1989), Ljungqvist and Sargent (2004, chap. 3,9), Sydsæter et al. (2005, chap. 12), (Wälde 2009, chap. 2,3,8,9)
This chapter briefly summarizes solution techniques to *discrete-time* dynamic optimization problems. Most of the chapter is concerned with dynamic programming.

2.2.1 A typical control problem

Consider a system that is observed at times $t = 0, 1, ..., T$. Suppose the *state* of the system at time t is characterized by a real number x_t. Assume that the initial state x_0 is historically given, and the system is steered through time by a sequence of *controls* $u_t \in U$. Briefly formulated, the problem reads

$$\max \sum_{t=0}^{T} f(t, x_t, u_t) \quad s.t. \quad x_{t+1} = g(t, x_t, u_t), \quad x_0 \quad \text{given.}$$

Among all *admissible* sequence pairs $(\{x_t\}, \{u_t\})$ find one, $(\{x_t^*\}, \{u_t^*\})$, that makes the value of the objective function as large as possible. Such an admissible pair is called an *optimal pair*, and the corresponding sequence $\{u_t^*\}_{t=0}^T$ is called an *optimal control.*

2.2.2 Solving using classical calculus methods

Consider the most simple version of a dynamic stochastic general equilibrium model as a straightforward extension of the following deterministic model of overlapping generations.

A simple deterministic overlapping generations model

Let there be an individual living for two periods with utility function, $U_t \equiv U(c_t, c_{t+1})$, where consumption in the first and second period is denoted by c_t and c_{t+1}, respectively. The individual receives labor income w_t in both periods, and allocates income to consumption, c_t and c_{t+1}, or saving, s_t which increases consumption possibilities in the second period generating, $(1 + r_{t+1})s_t$. Summarized the control problem reads

$$\max \ U(c_t, c_{t+1}) \quad s.t. \quad c_{t+1} = w_{t+1} + (1 + r_{t+1})(w_t - c_t).$$

For simplicity, the state variable is zero at the *beginning* of the first and at the *end* of the second period, that is by assumption there is no initial wealth and the individual does not leave any bequest. This points to the fact that *timing* within a period is an important issue in discrete-time models. Here, the budget constraint reveals some insights.

Observe that there are no interest payments in the first period, wages are paid and consumption takes place at the end of a period. Because savings, $s_t = w_t - c_t$, are used in the form of productive capital in the second period, returns r_{t+1} are determined in $t+1$. Without any further restrictions, we implicitly assume perfect capital markets, that is individuals can save and borrow any amount they desire at the rate r_{t+1}.

The problem can be solved simply by defining the *Lagrangian*,

$$\mathcal{L} = U(c_t, c_{t+1}) + \lambda \left((1 + r_{t+1})(w_t - c_t) + w_{t+1} - c_{t+1} \right),$$

where first-order conditions are

$$\begin{aligned} \mathcal{L}_{c_t} &= U_{c_t} - \lambda(1 + r_{t+1}) = 0, \\ \mathcal{L}_{c_{t+1}} &= U_{c_{t+1}} - \lambda = 0. \end{aligned}$$

Combining the conditions gives a necessary condition for optimality,

$$U_{c_t} = U_{c_{t+1}}(1 + r_{t+1}) \;\Leftrightarrow\; \frac{U_{c_t}}{U_{c_{t+1}}} = \frac{1}{(1 + r_{t+1})^{-1}}. \tag{2.20}$$

Obviously, optimal behavior requires that the *marginal rate of substitution* of consumption c_t and c_{t+1} given a consumption bundle (c_t, c_{t+1}) must equal their relative price. Otherwise, individuals could increase their overall utility simply by adjusting their consumption levels,

$$dU_t = U_{c_t} dc_t + U_{c_{t+1}} dc_{t+1} = 0 \;\Leftrightarrow\; \frac{dc_{t+1}}{dc_t} = -\frac{U_{c_t}}{U_{c_{t+1}}}.$$

An important measure is the *intertemporal elasticity of substitution* of consumption at two points in time,

$$\theta_{c_t, c_{t+1}} \equiv -\frac{U_{c_t}/U_{c_{t+1}}}{c_t/c_{t+1}} \frac{d(c_t/c_{t+1})}{d(U_{c_t}/U_{c_{t+1}})} = -\frac{d \ln(c_t/c_{t+1})}{d \ln(U_{c_t}/U_{c_{t+1}})}, \tag{2.21}$$

measuring the percentage change in relative consumption by a percentage change in their relative price. For the Cobb-Douglas case, $U_t = U(c_t, c_{t+1}) = c_t^{\gamma} c_{t+1}^{1-\gamma}$ we obtain

$$\theta_{c_t, c_{t+1}} = -\frac{\frac{\gamma}{1-\gamma} c_{t+1}/c_t}{c_t/c_{t+1}} \frac{d(c_t/c_{t+1})}{\frac{\gamma}{1-\gamma} d(c_{t+1}/c_t)} = -\frac{1}{(c_t/c_{t+1})^2} \frac{d((c_{t+1}/c_t)^{-1})}{d(c_{t+1}/c_t)} = 1.$$

Note that *overall* utility, $U_t = U(c_t, c_{t+1})$, is separable into *instantaneous* utility levels, u,

$$\ln U_t = \gamma \ln c_t + (1 - \gamma) \ln c_{t+1} = \gamma u(c_t) + (1 - \gamma) u(c_{t+1}). \tag{2.22}$$

The *time preference rate* is the rate at which future utility is discounted. Technically, it is the marginal rate of substitution of instantaneous utility $u(c_t)$ and $u(c_{t+1})$ minus one,

$$\rho = \frac{U_{u(c_t)}}{U_{u(c_{t+1})}} - 1 = \frac{\gamma}{1-\gamma} - 1. \tag{2.23}$$

Hence, the rate of time preference is positive for $\gamma > 1/2$, which makes sense as γ points to the relative importance of instantaneous utility in (2.22). Accordingly, for $\gamma > 0$, instantaneous utility in t is preferred to instantaneous utility in $t+1$.

This allows to derive an intuitive condition under which consumption increases over time (or equivalently where savings are positive for invariant labor income). Using the utility function $U_t = U(c_t, c_{t+1}) = c_t^{\gamma} c_{t+1}^{1-\gamma}$ and the first-order condition (2.20),

$$\frac{\gamma}{1-\gamma}\frac{c_{t+1}}{c_t} = 1 + r_{t+1} \;\Leftrightarrow\; c_{t+1} = \frac{1-\gamma}{\gamma}(1+r_{t+1})c_t,$$

we obtain

$$c_{t+1} > c_t \;\Leftrightarrow\; 1 + r_{t+1} > \frac{\gamma}{1-\gamma} \;\Leftrightarrow\; r_{t+1} > \rho.$$

Thus consumption increases if the interest rate is higher than the time preference rate, that is returns to saving are sufficiently high to overcompensate impatience.

A simple stochastic overlapping generations model

Let there be an aggregate technology

$$Y_t = A_t K_t^{\alpha} L^{1-\alpha}. \tag{2.24}$$

Suppose $\{A_t\}_{t\in\mathbb{T}}$ is a stochastic process. By assumption, the fundamental source of uncertainty is exogenous resulting from the technology used by firms,

$$A_t \sim (A, \sigma^2), \quad A_t > 0,$$

where $\{A_t\}_{t\in\mathbb{T}}$ is a continuous-state stochastic process with mean A and variance σ^2. At the beginning of period t, the capital stock K_t is inherited from the previous period, the capital stock therefore is *predetermined*. Then, total factor productivity is revealed and firms choose factor inputs, and households receive factor income and choose their consumption level.

The crucial assumption is that wages and interest payments are known with *certainty* at the end of the period. As a result of the timing, firms do not bear any risk and pay marginal product of labor, $w_t = Y_L$, and capital, $r_t = Y_K$, to workers and capital owners, respectively. Therefore, workers and capital owners bear all risk because their returns are uncertain. Let overall utility be time separable, the control problem reads

$$\max \ E_t \{u(c_t) + \beta u(c_{t+1})\} \quad s.t. \quad c_{t+1} = w_{t+1} + (1 + r_{t+1})(w_t - c_t),$$

where $\beta = 1/(1+\rho)$, $\rho > 0$ denotes the subjective discount factor measuring the individual's impatience to consume. Despite the uncertainty about w_{t+1} and r_{t+1}, the dynamic budget constraint has to hold. In that *contingent claims* have to ensure that negative savings are indeed settled in the second period. Inserting the budget constraint yields,

$$\begin{aligned} &\max \ \{u(c_t) + \beta E_t u((1 + r_{t+1})(w_t - c_t) + w_{t+1})\} \\ \Leftrightarrow \ &\max \ \left\{u(c_t) + \beta \int_0^\infty f(s) u((1 + r_{s,t+1})(w_t - c_t) + w_{s,t+1}) ds\right\}, \end{aligned}$$

where $f(s)$ is the density function of A_t given the information set t. Optimality requires

$$\begin{aligned} u'(c_t) - \beta \int_0^\infty f(s) u'((1 + r_{s,t+1})(w_t - c_t) + w_{s,t+1})(1 + r_{s,t+1}) ds = 0 \\ \Leftrightarrow \ u'(c_t) = \beta E_t u'((1 + r_{t+1})(w_t - c_t) + w_{t+1})(1 + r_{t+1}). \end{aligned} \tag{2.25}$$

Hence, marginal utility in t has to be equal to *expected* discounted instantaneous marginal utility in $t + 1$ corrected by the interest rate effect.

In some cases, where the instantaneous utility function u allows to separate the capital returns from consumption in the first period, an explicit expression or *closed-form* solution for c_t can be obtained. Using instantaneous utility $u(c_t) = \ln c_t$, the first-order condition in (2.25), and the assumption $w_{t+1} = 0$, which turns out to be necessary for a closed form,

$$1/c_t = \beta E_t \frac{1 + r_{t+1}}{(1 + r_{t+1})(w_t - c_t)} \ \Leftrightarrow \ (1 + \beta) c_t = w_t, \tag{2.26}$$

where from the budget constraint $w_t = c_t + (1 + r_{t+1})^{-1} c_{t+1}$. Inserting yields

$$(1 + \beta) c_t = c_t + (1 + r_{t+1})^{-1} c_{t+1} \ \Leftrightarrow \ c_{t+1} = \beta(1 + r_{t+1}) c_t, \tag{2.27}$$

which has the same structure as in the deterministic setup. In contrast to the solution under certainty there is still uncertainty about the consumption level in $t + 1$. Nonetheless, the optimal consumption level in the first period is available in closed form, $c_t = 1/(1 + \beta) w_t$.

We now aggregate over all individuals in order to find the *reduced form* of the overlapping generations model. Consumption of all young individuals in period t is given from (2.26),

$$C_t^y = \frac{w_t N}{1+\beta} = \frac{w_t L}{1+\beta} = \frac{1}{1+\beta}(1-\alpha)Y_t,$$

where we used the fact that under perfect factor markets, using Euler's theorem provided the production function has constant returns to scale, $Y_t = r_t K_t + w_t L$. Consumption of the old individuals from (2.27) is

$$C_t^o = \beta(1+r_t)Nc_{t-1} = \frac{\beta}{1+\beta}(1+r_t)Lw_{t-1} = \frac{\beta}{1+\beta}(1+r_t)(1-\alpha)Y_{t-1}.$$

Aggregate consumption in t therefore reads

$$C_t = C_t^y + C_t^o = \frac{1}{1+\beta}(1-\alpha)\left(Y_t + \beta(1+r_t)Y_{t-1}\right).$$

Market clearing on the goods market demands that supply equals demand,

$$Y_t + K_t = C_t + I_t, \tag{2.28}$$

where the left-hand side gives supply as current production and the capital stock sold by the old generation as it is of no use in the next period. The right-hand side gives demand as total consumption plus investment. Investment is determined by the aggregate savings of the young generation,

$$I_t = L(w_t - c_t) = \left(1 - \frac{1}{1+\beta}\right) Lw_t = \frac{\beta}{1+\beta}(1-\alpha)Y_t.$$

Because the old generation leaves no bequests, the capital stock K_{t+1} is fully determined by investment of the young generation,

$$K_{t+1} = I_t = \frac{\beta}{1+\beta}(1-\alpha)A_t K_t^{\alpha} L^{1-\alpha},$$

which is a non-linear first-order stochastic difference equation in the capital stock. Analyzing the properties of stochastic processes with *stochastic coefficients* in general is fairly complex. For the reduced form equation above, however, a log-transformation yields

$$\begin{aligned}
\ln K_{t+1} &= \ln\left(\frac{\beta}{1+\beta}(1-\alpha)L^{1-\alpha}\right) + \alpha \ln K_t + \ln A_t \\
&\equiv \ln\left(\frac{\beta}{1+\beta}(1-\alpha)L^{1-\alpha}\right) + \mu_\epsilon + \alpha \ln K_t + \sigma_\epsilon \epsilon_t, \quad \epsilon_t \sim (0,1),
\end{aligned}$$

where we defined $\mu_\epsilon = E(\ln A_t)$, and $\sigma_\epsilon \epsilon_t = \ln A_t - E(\ln A_t)$. Because $0 < \alpha < 1$, it is a stable linear first-order stochastic difference equation converging towards a *stochastic steady state*, or more precisely towards some *stationary distribution* of the log capital stock. It can be written as a stochastic difference equation in normal form,

$$\begin{bmatrix} \ln K_{t+1} \\ \epsilon_{t+1} \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha & \sigma_\epsilon & b \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \ln K_t \\ \epsilon_t \\ 1 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \epsilon_{t+1},$$

where ϵ_t is white noise with $E(\epsilon_t) = 0$ and $Var(\epsilon_t) = 1$, and b denotes a constant,

$$b = \ln \left(\frac{\beta}{1+\beta}(1-\alpha)AL^{1-\alpha} \right) + \mu_\epsilon.$$

From this we obtain the first two moments of the stationary distribution as

$$E(\ln K_t) = \frac{b}{1-\alpha} = \frac{1}{1-\alpha} \ln \left(\frac{\beta}{1+\beta}(1-\alpha)L^{1-\alpha} \right) + \frac{\mu_\epsilon}{1-\alpha}, \quad Var(\ln K_t) = \frac{\sigma_\epsilon^2}{1-\alpha^2}.$$

Given a specific distributional assumption for A_t, we could explicitly relate μ_ϵ and σ_ϵ to the moments of the fundamental uncertainty, A_t, that is as function of A and σ^2, respectively.

2.2.3 Solving using dynamic programming

Below, we closely follow Sydsæter et al. (2005, chap. 12). Return to the control problem,

$$\max \sum_{t=0}^{T} f(t, x_t, u_t) \quad s.t. \quad x_{t+1} = g(t, x_t, u_t), \quad x_0 \quad \text{given.}$$

Suppose we choose arbitrary values for $\{u_t\}_{t=0}^T$ with $u_t \in U$ for all t, the states $\{x_t\}_{t=1}^T$ can be computed recursively. Each choice of controls give rise to a sequence or a path which usually have different utility or value $\sum_{t=0}^T f(t, x_t, u_t)$. Often, controls u_t that depend only on time are referred to as *open-loop controls*, while controls that depend on the state of the system are called *closed-loop controls*, *feedback controls*, or *policies*.

Bellman's principle

Suppose that at time $t = s$ the state is $x_s = x \in \mathbb{R}$. The best we can do in the remaining periods is to choose $\{u_t\}_{t=s}^T$ and thereby also $\{x_t\}_{t=s+1}^T$ to maximize $\sum_{t=s}^T f(t, x_t, u_t)$ with $x_s = x$ subject to $x_{t+1} = g(t, x_t, u_t)$ for $t > s$. The *optimal control*, $\{u_t^*\}_{t=0}^T$, will depend on

x, in particular, $u_s^* = u^*(x_s) = u^*(x)$. We define the value function at time s,

$$J_s(x) = \max_{u_s,\ldots,u_T \in U} \sum_{t=s}^{T} f(t, x_t, u_t) = \sum_{t=s}^{T} f(t, x_t^*, u_t^*), \tag{2.29}$$

where

$$x_s = x, \quad x_{s+1} = g(s, x_s, u_s). \tag{2.30}$$

If we choose $u_s = u \in U$, then at time $t = s$ we obtain the reward $f(s, x, u)$ and the state changes to $x_{s+1} = g(s, x, u)$ as from (2.30). The highest reward starting from the state x_{s+1} is $J_{s+1}(x_{s+1}) = J_{s+1}(g(s, x, u))$ according the definition in (2.29). Hence the best choice of $u_s = u$ at time s must be a value of u that maximizes $f(s, x, u) + J_{s+1}(g(s, x, u))$,

$$J_s(x) = \begin{cases} \max_{u\in U} \{f(s, x, u) + J_{s+1}(g(s, x, u))\}, & s = 0, 1, ..., T-1 \\ \max_{u\in U} \{f(T, x, u)\}, & s = T \end{cases}. \tag{2.31}$$

Note that often (2.31) is referred to as the *fundamental equation* of dynamic programming, because it is the basic tool for solving dynamic optimization problems:

1. Find the optimal function $J_T(x) = \max_{u\in U} f(T, x, u)$ for $s = T$, where (usually) the maximizing value of u will depend on x, and was denoted by $u_T^*(x)$ above.

2. Use (2.31) in order to determine $J_{T-1}(x)$ and the corresponding u_{T-1}^* of the preceding period, and work backwards recursively to determine all the value functions and hence the optimal control $\{u_t^*\}_{t=0}^T$.

Example 2.2.1 (Overlapping generations) *Consider the following control problem of overlapping generations, with instantaneous utility $u' > 0$ and $u'' < 0$,*

$$\max \sum_{t=0}^{T} \beta^t u(c_t) \quad s.t. \quad a_{t+1} = (1 + r_t)a_t + w_t - c_t, \quad a_0 = a = 0, \quad T = 1,$$

where we introduced a state variable $a_t \geq 0$.

1. *find the optimal value $J_T(a) = \max_{c\geq 0} \beta u(c_1)$. Note that the maximum of $\beta u(c_1)$ can be obtained by the corner solution $\beta u(c_1) = \beta u(w_1 + r_1 a_1)$ such that $c_1^*(a) = w_1 + (1 + r_1)a_1$. It is reasonable that $a_1 = a_1(a)$ because $a_1 = w_0 - c_0 + (1 + r_0)a_0$, $a_0 = a_0(a)$.*

2. *determine $J_{T-1}(a)$ recursively for the preceding (that is the initial) period from*

$$J_0(a) = \max_{c\geq 0}\{u(c_0) + J_T(a)\} = \max_{c\geq 0}\{u(c_0) + \beta u(w_1 + (1 + r_1)a_1)\},$$

which becomes using $a_0 = a_0(a) = a = 0$,

$$J_0(a) = \max_{c \geq 0} \{u(c_0) + \beta u(w_1 + (1 + r_1)(w_0 - c_0))\}.$$

The optimal $c_0^*(a)$ satisfies the condition

$$\begin{aligned} u'(c_0) + \beta u'(w_1 + (1 + r_1)(w_0 - c_0))(-(1 + r_1)) &= 0 \\ \Leftrightarrow \quad \frac{u'(c_0)}{u'(w_1 + (1 + r_1)(w_0 - c_0))} &= \beta(1 + r_1). \end{aligned}$$

It coincides with the optimality condition $u'(c_0)/u'(c_1) = \beta(1+r_1)$ in the model of overlapping generations in (2.20), where $U(c_1, c_2) = c_1^\gamma c_2^{1-\gamma}$, and the factor $\beta = (1 - \gamma)/\gamma$.

In simple cases, the control problem can be solved quite easily by ordinary calculus methods. In principle, all finite horizon dynamic programming problems can be solved using classical methods, however, the method becomes quite messy if the horizon T is large.

Infinite horizon

Economists often study dynamic optimization problems over an infinite horizon. This avoids specifying what happens after the finite horizon is reached, as well as having the horizon as an extra exogenous variable that features in the solution. Consider the following problem

$$\max \sum_{t=0}^{\infty} \beta^t f(x_t, u_t) \quad s.t. \quad x_{t+1} = g(x_t, u_t), \quad x_0 \quad \text{given}, \tag{2.32}$$

where $\beta \in (0, 1)$ is a constant *discount factor*, and x_0 is a given number in $\mathbb{R}$. Note that neither f nor g depends explicitly on t. For this reason, problem (2.32) is called *autonomous*. Assume that the infinite sum converges, that is f satisfies some boundedness conditions.

Bellman's principle

Suppose that at time $t = s$ the state is $x_s = x \in \mathbb{R}$. The optimal control $\{u_t^*\}_{t=s}^{\infty}$ defines the value function as

$$J_s(x) = \sum_{t=s}^{\infty} \beta^t f(x_t^*, u_t^*), \quad V(x) \equiv J_0(x),$$

Roughly, if we choose the control u, the immediate reward is $\beta^s f(x, u)$ and the state changes to $x_{s+1} = g(x, u)$. Choosing an optimal control sequence from $t = s + 1$ on gives a total reward over all subsequent periods that equals $J_{s+1}(g(x,u)) = \beta J_s(g(x,u))$. For $s = 0$,

$$V(x) = \max_{u \in U} \{f(x,u) + \beta V(g(x,u))\}, \tag{2.33}$$

is a necessary condition for (2.32) which is the *Bellman equation*, i.e., the first step of the *three-steps* procedure for solving discrete infinite horizon problems (Wälde 2009, chap. 3.3). As a corollary, we can compute the *first-order condition* which reads

$$f_u(x,u) + \beta V'(g(x,u)) g_u(x,u) = 0, \tag{2.34}$$

which is the solution to the control problem (2.32) and makes $u^* = u(x)$ a function of the state variable. Note that (2.34) is a first-order difference equation in x_s, for any $s \geq 0$

$$f_u(x_s,u_s) + \beta V'(x_{s+1}) g_u(x_s,u_s) = 0, \tag{2.35}$$

because the future looks exactly the same at time 0 as at time s. As we usually do not know the functional form of $V(x)$, or say $V'(x)$, we need to go through two further steps in order to eliminate the costate variable V', replacing it by known functions of f and g.

In a second step, we determine the *evolution of the costate variable.* Replacing the control variable u_s by the optimal control $u_s^* = u_s(x_s)$ gives the maximized Bellman equation,

$$V(x_s) = f(x_s, u_s(x_s)) + \beta V(g(x_s, u_s(x_s))).$$

Computing the derivative with respect to x_s, we obtain using the envelope theorem

$$V'(x_s) = f_x(x_s, u_s(x_s)) + \beta V'(x_{s+1}) g_x(x_s, u_s). \tag{2.36}$$

Interpreted as a shadow price, it gives the value of a marginal increase in the state variable.

As the final step we insert the first-order condition (2.35),

$$V'(x_{s+1}) = -\frac{1}{\beta}\frac{f_u(x_s,u_s)}{g_u(x_s,u_s)}, \quad V'(x_{s+2}) = -\frac{1}{\beta}\frac{f_u(x_{s+1},u_{s+1})}{g_u(x_{s+1},u_{s+1})},$$

into (2.36) shifted one period ahead to obtain

$$\begin{aligned} V'(x_{s+1}) &= f_x(x_{s+1},u_{s+1}) + \beta V'(x_{s+2}) g_x(x_{s+1},u_{s+1}) \\ \Leftrightarrow \frac{f_u(x_s,u_s)}{g_u(x_s,u_s)} &= \beta \frac{f_u(x_{s+1},u_{s+1})}{g_u(x_{s+1},u_{s+1})} g_x(x_{s+1},u_{s+1}) - \beta f_x(x_{s+1},u_{s+1}), \end{aligned} \tag{2.37}$$

which is a generalized version of the discrete-time *Euler equation.*

Example 2.2.2 (Infinite horizon) *Consider the following infinite horizon control problem with instantaneous utility* $u' > 0$ *and* $u'' < 0$*,*

$$\max \sum_{t=0}^{\infty} \beta^t u(c_t) \quad s.t. \quad a_{t+1} = (1+r_t)a_t + w_t - c_t, \quad a_0 = a = 0,$$

where $a_t \geq 0$ *denotes individual's real wealth. Observe that*

$$\begin{aligned} f(a_t, c_t) &= u(c_t) \quad \Rightarrow \; f_a = 0,\; f_c = u'(c_t), \\ g(a_t, c_t) &= w_t - c_t + (1+r_t)a_t \quad \Rightarrow \; g_a = 1 + r_t,\; g_c = -1. \end{aligned}$$

Going step-by-step through the suggested procedure or just plugging the partial derivatives in the generalized Euler equation (2.37) gives the necessary condition,

$$-u'(c_t) = -\beta u'(c_{t+1})(1 + r_{t+1}) \;\Leftrightarrow\; \frac{u'(c_t)}{u'(c_{t+1})} = \beta(1 + r_{t+1}), \tag{2.38}$$

which is exactly the same condition for the overlapping generations model. Apparently, the change in optimal consumption does not depend on the time horizon. Be aware, however, that the consumption level could (indeed should) depend on the planning horizon.

In the example, the implicit timing is different to models where $a_{t+1} = (1+r_t)(w_t - c_t + a_t)$. Observe, however, it does not matter for the Euler equation.

2.2.4 Stochastic control problems

This section considers how to control a dynamic system subject to random disturbances. Stochastic dynamic programming is a central tool for tackling this problem.

We consider the following infinite horizon stochastic control problem,

$$\max E \sum_{t=0}^{\infty} \beta^t f(X_t, u_t), \quad X_{t+1} = g(X_t, u_t, Z_t), \quad x_0 = x,\; z_0 = z,\; (x, z) \in \mathbb{R}^2, \tag{2.39}$$

where $\{Z_t\}_{t=0}^{\infty}$ is a Markov dependent stochastic process, each random variable defined on the same probability space $(\Omega, \mathfrak{F}, P)$. Note that $P(Z_{t+1} = z_{t+1} | Z_t = z_t)$ as well as functions f and g do *not* explicitly depend on t making the control problem *autonomous* or *stationary.* For $s \geq 0$, each optimal choice $u_s^* = u_s^*(X_s, Z_s)$ will be a function of the current state X_s and the random variable Z_s, to which we refer as *Markov policies* or *Markov controls.*

Bellman's principle

The heuristic argument for the optimality equation works just as well in the stochastic control problem because of the *autonomous* formulation and the *Markov property* of the stochastic process the future looks exactly the same at time $t = s$ as it does at time $t = 0$, however, with the next periods value function discounted at β. Hence the *Bellman equation* reads

$$V(x, z) = \max_{u \in U}\{f(x, u) + \beta E_0(V(X_1, Z_1))\}, \tag{2.40}$$

defining the expectation operator $E_0(V(X_1, Z_1)) \equiv E(V(g(x, u, z), Z_1)|X_0 = x, Z_0 = z)$. Observe that the term z enters the value function, which is considered as an *exogenous* state variable. There are cases, however, where this state variable is *endogenous* as well.

Observe that the Bellman equation again is a *functional equation* which determines the unknown function V that occurs on both sides. Once V is known, the optimal Markov control is obtained from maximizing in the optimality equation. Note that certain boundedness conditions on f are assumed to hold (Sydsæter et al. 2005, chap. 12.7).

We now proceed with the *three-steps* procedure for solving infinite horizon stochastic control problems (Wälde 2009, chap. 9). As a corollary, we obtain first-order conditions,

$$f_u(x, u) + \beta E_0\left(V_x(g(x, u, z), Z_1)g_u(x, u, z)\right) = 0,$$

which basically is a first-order stochastic difference equation,

$$f_u(X_s, u_s) + \beta E_s\left(V_x(X_{s+1}, Z_{s+1})g_u(X_s, u_s, Z_s)\right) = 0, \tag{2.41}$$

indeed providing a functional relationship of the control and the state variables.

In a second step, we determine the *evolution of the costate variable.* Using the maximized Bellman equation,

$$V(X_s, Z_s) = f(X_s, u_s^*) + \beta E_s(V(X_{s+1}, Z_{s+1})),$$

where $u_s^* = u_s(X_s, Z_s)$, and the envelope theorem gives the (evolution of the) costate as

$$V_x(X_s, Z_s) = f_x(X_s, u_s(X_s, Z_s)) + \beta E_s\left(V_x(X_{s+1}, Z_{s+1})g_x(X_s, u_s(X_s, Z_s), Z_s)\right). \tag{2.42}$$

As the final step we use the first-order condition (2.41),

$$\beta E_s\left(V_x(X_{s+1}, Z_{s+1})\right) = -\frac{f_u(X_s, u_s)}{g_u(X_s, u_s, Z_s)},$$

making use of the fact that $E_s(g_u(X_s, u_s, Z_s)) = g_u(X_s, u_s, Z_s)$ is deterministic conditional on the information set available at s. Inserting into (2.42) by using a similar argument for $E_s(g_x(X_s, u_s, Z_s)) = g_x(X_s, u_s, Z_s)$ we obtain

$$\begin{aligned} V_x(X_s, Z_s) &= f_x(X_s, u_s) - g_x(X_s, u_s, Z_s)\frac{f_u(X_s, u_s)}{g_u(X_s, u_s, Z_s)}, \\ V_x(X_{s+1}, Z_{s+1}) &= f_x(X_{s+1}, u_{s+1}) - g_x(X_{s+1}, u_{s+1}, Z_{s+1})\frac{f_u(X_{s+1}, u_{s+1})}{g_u(X_{s+1}, u_{s+1}, Z_{s+1})}, \end{aligned}$$

where we shifted the resulting expression also one period ahead. Inserting both expressions back into (2.42) we managed replacing the costate by terms of known functions f and g,

$$\frac{f_u(X_s, u_s)}{g_u(X_s, u_s, Z_s)} = \beta E_s\left(\frac{f_u(X_{s+1}, u_{s+1})}{g_u(X_{s+1}, u_{s+1}, Z_{s+1})}g_x(X_{s+1}, u_{s+1}, Z_{s+1}) - f_x(X_{s+1}, u_{s+1})\right),$$

and obtained the discrete-time *Euler equation* of the stochastic control problem in (2.39).

Exercise 2.2.3 (Real business cycles) *Consider the prototype real business cycles (RBC) model as in King et al. (1988). Suppose a benevolent planner maximizes*

$$\max E\sum_{t=0}^{\infty}\beta^t\left(\ln C_t + \theta\ln(1-N_t)\right), \quad s.t. \quad K_{t+1} = A_t N_t^{\alpha} K_t^{1-\alpha} - C_t + (1-\delta)K_t, \quad K_0 > 0.$$

C_t is commodity consumption, N_t is the labor input and K_t is the predetermined capital stock in period t. A_t denotes stochastic temporary changes in total factor productivity, δ is the rate of physical depreciation of capital, and $0 < \alpha < 1$ is the output elasticity of labor. Solve the optimization problem and obtain the Euler equation for consumption.

Chapter 3

Stochastic models in continuous time

3.1 Topics in stochastic differential equations and rules for differentials

Literature: Øksendal (1998, chap. 4,5), Kloeden and Platen (1999, chap. 3,4), Spanos (1999, chap. 8), Protter (2004, chap. 1,2), Wälde (2009, chap. 10)
The objective of this chapter is to introduce concepts for stochastic models in continuous time, where usually uncertainty enters in the form of stochastic differential equations to model specific dynamics, e.g., the evolution of prices or technology frontiers.

3.1.1 Definitions

For later reference we consider two fundamental stochastic processes, the *Brownian motion* and the *Poisson process*. While the *Brownian motion* is often used to model the behavior of prices (e.g., returns, exchange rates, interest rates), the *Poisson process* captures rare events.

Definition 3.1.1 (Standard Brownian motion) *The stochastic process $\{B_t\}_{t\in[0,\infty)}$ is said to be a standard Brownian motion process if the following conditions hold,*

(i) $B_{t+h} - B_t \sim N(0,|h|)$, *for* $(t+h) \in [0,\infty)$,

(ii) B_t *has independent increments, that is for* $0 \leq t_1 < t_2 < t_3 < \infty$,

$$\begin{pmatrix} B_{t_1} - B_{t_2} \\ B_{t_2} - B_{t_3} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} t_2 - t_1 & 0 \\ 0 & t_3 - t_2 \end{pmatrix} \right),$$

(iii) $B_0 = 0$.

Remark 3.1.2 *There is confusion in the literature in so far as the process first noticed by Brown is called a Brownian motion or a Wiener process. We closely follow Spanos (1999) and refer to the continuous-time process as a Brownian motion and its discrete counterpart as a Wiener process. This, however, is not standard terminology.*

Remark 3.1.3 *For $\{B_t\}_{t\in[0,\infty)}$ being a Brownian motion, the following properties hold*

1. *$\{B_t\}_{t\in[0,\infty)}$ is a Markov process. This follows directly from the fact that*

$$B_{t_2+t_1} = B_{t_1} + [B_{t_1+t_2} - B_{t_1}],$$

which says that the new state $B_{t_2+t_1}$ is the sum of the old state B_{t_1} and an independent Normal random variable $[B_{t_1+t_2} - B_{t_1}]$.

2. *The sample paths of a Brownian motion process are continuous but almost nowhere differentiable. Think of the zig-zag trace of a particle in a liquid.*

3. *The Brownian motion processes $\left\{\sqrt{c}B\left(\frac{t}{c}\right)\right\}_{t\in[0,\infty)}$, where $c > 0$ and $\{B_t\}_{t\in[0,\infty)}$ have the same joint distribution. This property is referred to as the scaling property.*

Remark 3.1.4 *Related stochastic processes to the standard Brownian motion, $\{B_t\}_{[0,\infty)}$, are*

1. *the process $\{\mu t + \sigma B_t\}_{t\in[0,\infty)}$ is said to be a Brownian motion with drift,*

2. *the process $\{B_t - tB_1\}_{t\in[0,1]}$ is said to be a Brownian bridge,*

3. *the process $\left\{\mu + \frac{\sigma}{\sqrt{2\theta}}e^{-\theta t}B\left(e^{2\theta t}\right)\right\}_{t\in[0,\infty)}$ is said to be an Ornstein-Uhlenbeck process,*

4. *the process $\{\int_0^t B_u du\}_{t\in[0,\infty)}$ is said to be an integrated Brownian motion process,*

5. *the process $\{\exp(B_t)\}_{t\in[0,\infty)}$ is said to be a geometric Brownian motion process.*

Definition 3.1.5 (Poisson process) *The stochastic process $\{N_t\}_{t\in[0,\infty)}$ is said to be a Poisson process if the following conditions hold,*

1. *$N_t = \max\{n : S_n(t) \leq t\}_{t\in[0,\infty)}$ is a point process, where*

2. *$S_n(t) = \sum_{i=1}^{n} X_i$ for $n \geq 1$, $S_0 = 0$ is a partial sum stochastic process, and*

3. *X_n with $n = 1, 2, ...$ is a sequence of independent identically exponentially distributed random variables.*

Remark 3.1.6 *Let N_t denote the number of phone calls up to date t, where random variable $S_n(t) = \min\{t : N(t) = n\}$ then is the date at which the telephone rings for the nth time, and $X_n = S_n - S_{n-1}$, $n = 1, 2, 3, ...$ denotes the time interval between calls n and $n - 1$.*

Remark 3.1.7 *Since the event $N_t \geq k$ is equivalent to $S_k(t) \leq t$, it follows*

$$P(N_t \geq k) = P(S_k(t) \leq t).$$

In view of the fact that $S_n(t)$ is the sum of i.i.d. exponentially distributed random variables, we can deduce the density function of N_t as follows (see Spanos 1999, chap. 8.11),

$$f_N(k) = P(N_t = k) = P(S_k(t) \leq t) - P(S_{k+1}(t) \leq t) = \frac{e^{-\lambda t}(\lambda t)^k}{k!}.$$

The Poisson process $\{N_t\}_{t\in[0,\infty)}$ has mean and variance

$$E(N_t) = \lambda t, \quad Var(N_t) = \lambda t,$$

i.e., the standard Poisson process is not stationary since its first two moments depend on t. It follows that $dN_t = 1$ with probability λdt and $dN_t = 0$ with probability $(1 - \lambda)dt$.

3.1.2 Stochastic differential equations

These stochastic processes can now be combined in various ways to construct more complex processes, which can nicely be represented by *stochastic differential equations*. An ordinary differential equation,

$$\dot{x} = \frac{dx}{dt} = a(t, x(t)),$$

may be thought of as a degenerated case of a stochastic differential equation in the absence of uncertainty. Using the symbolic *differential* form, we could write

$$dx = a(t, x(t))dt$$

or more accurately,

$$x(t) = x_0 + \int_0^t a(s, x(s))ds.$$

A generalized formulation can be obtained for stochastic processes. The Brownian motion constitutes the principal element for a class called *diffusion processes* based on stochastic differentials. By a *stochastic differential* we mean an expression of the type

$$dX_t = a(t, X_t)dt + b(t, X_t)dB_t, \tag{3.1}$$

which is just a symbolic way of writing

$$X_t = X_s + \int_s^t a(u, X_u)du + \int_s^t b(u, X_u)dB_u, \tag{3.2}$$

for any $0 \leq s \leq t$. The first integral in (3.2) is an ordinary Riemann or Lebesgue integral and the second integral is an Itô integral. In general, X_t inherits the non-differentiability of sample paths from the Brownian motion in the stochastic integral.

Remark 3.1.8 *Stochastic differentials can also be obtained using other processes,*

$$dX_t = a(t, X_t)dt + b(X_{t-})dN_t.$$

N_t is a càdlàg (from the French "continue à droite, limite à gauche") Poisson process, and N_{t-} denotes the left-limit $\lim_{\tau \to t} N_\tau$. *X_t coincides with X_{t-} if X_t has continuous paths. We may define stochastic differentials using combination of stochastic processes,*

$$dX_t = a(t, X_t)dt + b(t, X_t)dB_t + c(X_{t-})dN_t.$$

Moreover, the coefficients can be stochastic, i.e., with direct dependence on the forcing process

$$dX_t = a(t, X_t, N_t)dt + b(X_{t-})dN_t.$$

There is a price for the convenient notation, namely that stochastic differentials, interpreted in terms of stochastic integrals, do not transform according to the rules of classical calculus. Instead an additional term appears and the resulting expression is called the *Itô formula.*

3.1.3 Functions of stochastic processes

An important aspect when working with stochastic processes in continuous time is that rules for computing differentials of functions of those processes are different from classical ones. We start with one stochastic process, in particular with a one-dimensional Brownian motion, and generalize results for higher dimensions and other stochastic processes afterwards.

Computing differentials for Brownian motions

The *Itô formula* as a rule can be found in many texts (e.g. Øksendal 1998, Theorem 4.1.2).

Theorem 3.1.9 (Itô's formula) *Let X_t be a scalar stochastic process given by*

$$dX_t = a(t)dt + b(t)dB_t.$$

Let $g(t,x)$ be a C^2 function on $[0,\infty) \times \mathbb{R}$. Then, $Y_t = g(t, X_t)$ obeys the differential

$$dY_t = \frac{\partial}{\partial t} g(t, X_t)dt + \frac{\partial}{\partial x} g(t, X_t)dX_t + \tfrac{1}{2}\frac{\partial^2}{(\partial x)^2} g(t, X_t)(dX_t)^2,$$

where $(dX_t)^2 = (dX_t)(dX_t)$ is computed according to the rules,

$$dtdt = dtdB_t = 0, \quad dB_t dB_t = dt.$$

Using $(dX_t)^2 = a^2(t)dtdt + 2a(t)b(t)dtdB_t + b^2(t)(dB_t)^2 = b^2(t)dt$, we obtain

$$dY_t = \left(\frac{\partial}{\partial t}Y_t + a(t)\frac{\partial}{\partial x}Y_t + \tfrac{1}{2}b^2(t)\frac{\partial^2}{(\partial x)^2}Y_t\right)dt + b(t)\frac{\partial}{\partial x}Y_t dB_t. \tag{3.3}$$

Note that Itô's formula is also referred to as *change of variables*. A similar rule can be stated and proved for the class of *semimartingales*, which includes *Lévy* processes, in particular the Poisson process and the Brownian motion (cf. Protter 2004, chap. 2.3, 2.7, Theorem 2.32).

Remark 3.1.10 *As a sketch of a proof, consider the stochastic differential $dX_t = adt + bdB_t$, or equivalently, $X_t = X_s + a(t-s) + b(B_t - B_s)$, and the C^2 function $g(t, X_t)$. Using Taylors theorem, a second-order approximation around s and X_s is*

$$\begin{aligned} g(t, X_t) &\approx g(s, X_s) + \frac{\partial}{\partial t}g(s, X_s)(t-s) + \frac{\partial}{\partial x}g(s, X_s)(X_t - X_s) \\ &+\tfrac{1}{2}\left(\frac{\partial^2}{(\partial x)^2}g(s, X_s)(X_t - X_s)^2 + 2\frac{\partial^2}{\partial x \partial t}g(s, X_s)(X_t - X_s)(t-s) + \frac{\partial^2}{(\partial t)^2}g(s, X_s)(t-s)^2\right). \end{aligned}$$

Substitute $\Delta g(t, X_t) = g(t, X_t) - g(s, X_s)$, $\Delta X_t = X_t - X_s$, and $\Delta t = t - s$,

$$\begin{aligned} \Delta g(t, X_t) \approx{}& \frac{\partial}{\partial t}g(s, X_s)\Delta t + \frac{\partial}{\partial x}g(s, X_s)\Delta X_t \\ &+\tfrac{1}{2}\left(\frac{\partial^2}{(\partial x)^2}g(s, X_s)(\Delta X_t)^2 + 2\frac{\partial^2}{\partial x \partial t}g(s, X_s)\Delta X_t \Delta t + \frac{\partial^2}{(\partial t)^2}g(s, X_s)(\Delta t)^2\right). \end{aligned}$$

Here, $\Delta X_t = a\Delta t + b\Delta B_t$, and ΔB_t is of order $\sqrt{\Delta t}$ since $Var(\Delta B_t) = E\left((\Delta B_t)^2\right) = \Delta t$.

As $\Delta t \to dt$, all terms with higher orders, $\Delta t \Delta t, \Delta t \Delta B_t \to 0$, $\Delta B_t \Delta B_t \to dt$, and thus

$$(\Delta X_t)^2 = a^2(\Delta t)^2 + 2ab\Delta t \Delta B_t + b^2(\Delta B_t)^2 \ \to \ b^2 dt.$$

Example 3.1.11 *Consider the simplest case where $X_t = B_t$ and $g(t,x) = \mu t + \sigma x$. Then, $Y_t = g(t, B_t) = \mu t + \sigma B_t$, and using Itô's formula Y_t obeys the differential*

$$dY_t = \mu dt + \sigma dB_t,$$

which is the stochastic differential of a Brownian motion with drift. Observe that if $g(t,x)$ is linear, the Itô formula reduces to the chain rule of classical calculus.

Example 3.1.12 *Consider the case where $dX_t = \sigma(t)dB_t$ and the function $g(t,x) = e^x$. Then $Y_t = g(t, X_t) = \exp(X_t)$, and from Itô's formula Y_t obeys the differential*

$$dY_t = \tfrac{1}{2}\sigma^2(t)Y_t dt + \sigma(t)Y_t dB_t,$$

which is the stochastic differential of a geometric Brownian motion (geometric diffusion). Observe that using $g(t,x) = e^{x - \frac{1}{2}\int_0^t \sigma^2(u)du}$ we obtain

$$dY_t = \left(-\tfrac{1}{2}\sigma^2(t)Y_t + \tfrac{1}{2}\sigma^2(t)Y_t\right)dt + \sigma(t)Y_t dB_t = \sigma(t)Y_t dB_t.$$

This shows that the counterpart of the exponential in the Itô calculus is $\exp(X_t - \frac{1}{2}\int_0^t \sigma^2(u)du)$.

Exercise 3.1.13 *Find the following stochastic integrals in terms of classical calculus*

1.
$$\int_0^t b dB_s,$$

2.
$$\int_0^t s dB_s,$$

3.
$$\int_0^t f(s) dB_s,$$

4.
$$\int_0^t B_s dB_s,$$

where B_t is a standard Brownian motion.

Remark 3.1.14 *An Itô stochastic integral can be thought of as a random variable on the bounded interval* $[t_0, t]$,

$$X_t = X_{t_0} + \int_{t_0}^{t} f(s, B_s) dB_s,$$

satisfying the following properties

1. $E(X_{t_2} - X_{t_1}) = 0 \quad \textit{for} \quad t_0 \le t_1 \le t_2,$
2. $E(X_t^2) = \int_{t_0}^{t} E\left(f(s, B_s)^2\right) ds < \infty,$
3. $\int_{t_0}^{t_2} f(s, B_s)dB_s = \int_{t_0}^{t_1} f(s, B_s)dB_s + \int_{t_1}^{t_2} f(s, B_s)dB_s \quad \textit{for} \quad t_0 \le t_1 \le t_2.$

The Itô formula can easily be generalized to an m-dimensional Brownian motion Z_t, where the $Z_i(t)$ for $i = 1, 2, ..., m$ are scalar processes which are pairwise correlated,

$$E\left[(Z_i(t) - Z_i(s))(Z_j(t) - Z_j(s))\right] = (t - s)\rho_{ij}, \quad 0 \le s \le t, \quad i, j = 1, 2, ..., m,$$

where ρ_{ij} is the *correlation coefficient* between increments of stochastic processes. When the Brownian motion are pairwise independent, then $\rho_{ij} = 0$ for $i \ne j$ and $\rho_{ij} = 1$ for $i = j$. An Itô formula for correlated Brownian motions is below (e.g. Merton 1999, Theorem 5.1).

Theorem 3.1.15 (Itô's formula for systems of Brownian motions) *Let*

$$dX_t = u(t, X_t)dt + \sigma(t, X_t)dZ_t,$$

be an n*-dimensional Itô process in matrix notation where*

$$X_t = \begin{bmatrix} X_1(t) \\ \vdots \\ X_n(t) \end{bmatrix}, \; u(t, X_t) = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}, \; \sigma(t, X_t) = \begin{bmatrix} \sigma_{11} & \dots & \sigma_{1m} \\ \vdots & & \vdots \\ \sigma_{n1} & \dots & \sigma_{nm} \end{bmatrix}, \; dZ_t = \begin{bmatrix} dZ_1(t) \\ \vdots \\ dZ_m(t) \end{bmatrix}.$$

Let $g(t, x) = (g_1(t, x), ..., g_p(t, x))$ *be a* C^2 *map from* $[0, \infty) \times \mathbb{R}^n$ *into* $\mathbb{R}^p$*. Then the process* $Y_t = g(t, X_t)$ *is again an Itô process, whose component* k*, obeys the differential*

$$dY_k(t) = \frac{\partial}{\partial t} g_k(t, X_t) + \sum_{i=1}^{n} \frac{\partial}{\partial x_i} g_k(t, X_t) dX_i + \tfrac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2}{\partial x_i \partial x_j} g_k(t, X_t) dX_i dX_j, \tag{3.4}$$

where

$$dtdt = dtdZ_i = 0, \quad dZ_i dZ_j = \rho_{ij} dt,$$

and ρ_{ij} *is the correlation coefficient between increments of the stochastic processes.*

Example 3.1.16 *Consider two independent stochastic processes* X_1 *and* X_2*,*

$$dX_t = u_t dt + \begin{bmatrix} \gamma_{11} & 0 \\ 0 & \gamma_{22} \end{bmatrix} dB_t \equiv u_t dt + \gamma dB_t,$$

or equivalently, consider a two-dimensional system of stochastic differential equations,

$$\begin{aligned} dX_1(t) &= u_1(t)dt + \gamma_{11} dB_1(t), \\ dX_2(t) &= u_2(t)dt + \gamma_{22} dB_2(t), \end{aligned}$$

and let $g(t,x) = x_1 x_2$ *be a* C^2 *function on* $\mathbb{R}^2 \to \mathbb{R}$*. The stochastic differential for the product scalar process* $Y_t = X_1 X_2$ *reads* $dY_t = X_2 dX_1 + X_1 dX_2 + dX_1 dX_2$*, where* $(dX_1)^2 = \gamma_{11}^2 dt$*,* $(dX_2)^2 = \gamma_{22}^2 dt$*,* $dX_1 dX_2 = 0$*. To summarize, we obtain*

$$dY_t = (u_1(t)X_2 + u_2(t)X_1)dt + \gamma_{11} X_2 dB_1(t) + \gamma_{22} X_1 dB_2(t).$$

Example 3.1.17 *Consider two dependent stochastic processes* X_1 *and* X_2*,*

$$dX_t = u_t dt + \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix} dB_t \equiv u_t dt + \gamma dB_t,$$

or equivalently, consider a two-dimensional system of stochastic differential equations,

$$\begin{aligned} dX_1(t) &= u_1(t)dt + \gamma_{11} dB_1(t) + \gamma_{12} dB_2(t), \\ dX_2(t) &= u_2(t)dt + \gamma_{21} dB_1(t) + \gamma_{22} dB_2(t), \end{aligned}$$

and let $g(t,x) = x_1 x_2$ *be a* C^2 *function on* $\mathbb{R}^2 \to \mathbb{R}$*. The stochastic differential for the product scalar process* $Y_t = X_1 X_2$ *reads* $dY_t = X_2 dX_1 + X_1 dX_2 + dX_1 dX_2$*, where* $(dX_1)^2 = (\gamma_{11}^2 + \gamma_{12}^2)dt$*,* $(dX_2)^2 = (\gamma_{21}^2 + \gamma_{22}^2)dt$*, and* $dX_1 dX_2 = (\gamma_{11}\gamma_{21} + \gamma_{22}\gamma_{12})dt$*. Hence,*

$$dY_t = (u_1(t)X_2 + u_2(t)X_1 + \gamma_{11}\gamma_{21} + \gamma_{22}\gamma_{12})dt + (\gamma_{11}X_2 + \gamma_{21}X_1)dB_1(t) + (\gamma_{22}X_1 + \gamma_{12}X_2)dB_2(t).$$

Example 3.1.18 *Consider two dependent stochastic processes* X_1 *and* X_2*,*

$$dX_t = u_t dt + \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix} dZ_t \equiv u_t dt + \mathrm{diag}(\sigma) dZ_t,$$

or equivalently, consider a two-dimensional system of stochastic differential equations,

$$\begin{aligned} dX_1(t) &= u_1(t)dt + \sigma_1 dZ_1(t), \\ dX_2(t) &= u_2(t)dt + \sigma_2 dZ_2(t), \end{aligned}$$

and let $g(t,x) = x_1 x_2$ *be a* C^2 *function on* $\mathbb{R}^2 \to \mathbb{R}$. *Then, the stochastic differential for the product scalar process* $Y_t = X_1 X_2$ *reads* $dY_t = X_2 dX_1 + X_1 dX_2 + dX_1 dX_2$, *where* $(dX_1)^2 = \sigma_1^2 dt$, $(dX_2)^2 = \sigma_2^2 dt$, $dX_1 dX_2 = (\rho_{12}\sigma_1\sigma_2)dt$. *Hence,*

$$dY_t = (u_1(t)X_2 + u_2(t)X_1)dt + \sigma_1 X_2 dZ_1(t) + \sigma_2 X_1 dZ_2(t) + \rho_{12}\sigma_1\sigma_2 dt.$$

The Itô formula contains an m-dimensional standard Brownian motion B_t, where $B_i(t)$ for $i = 1, 2, ..., m$ are a scalar processes which are pairwise independent,

$$E\left[(B_i(t) - B_i(s))(B_j(t) - B_j(s))\right] = (t-s)\delta_{ij}, \quad 0 \le s \le t, \quad i, j = 1, 2, ..., m,$$

and δ_{ij} is the *Kronecker delta symbol* defined by,

$$\delta_{ij} = \begin{cases} 1 & : \; i = j \\ 0 & : \; i \neq j \end{cases}.$$

Observe that the appropriate Itô formula is contained in Theorem 3.1.15 for $\rho_{ij} = \delta_{ij}$. Note that the stochastic differential is simply the sum of the differentials of each stochastic process independently, plus a term capturing the *correlation* structure.

Computing differentials for Poisson processes

When we consider the stochastic differential of a function of a variable that is driven by a scalar Poisson process, we need to take into account the following version of *Itô's formula*.

Theorem 3.1.19 (cf. Sennewald 2007, Theorem 1) *Let* X_t *be an stochastic process,*

$$dX_t = a(t, X_t)dt + c(t, X_{t-})dN_t.$$

Let $g(t,x)$ *be a* C^1 *function on* $[0,\infty) \times \mathbb{R} \to \mathbb{R}$. *Then,* $Y_t = g(t, X_t)$ *obeys the differential*

$$dY_t = \frac{\partial}{\partial t} g(t, X_t)dt + a(t, X_t)\frac{\partial}{\partial x} g(t, X_t)dt + (g(t, X_{t-} + c(t, X_{t-})) - g(t, X_{t-}))dN_t. \quad (3.5)$$

This rule is intuitive, because the differential of a function is given by the usual chain rule plus an obvious *jump term*. If the point process N_t increases by one, X_t increases by $c(t, X_{t-})$, which translates into a jump in Y_t from $g(t, X_{t-})$ to the new value $g(t, X_{t-} + c(t, X_{t-}))$, or equivalently, it jumps by $g(t, X_{t-} + c(t, X_{t-})) - g(t, X_{t-})$. Remember that N_t simply counts the arrivals up to t and a new arrival lets N_t increase by one, that is $dN_t = 1$.

Theorem 3.1.20 (cf. Sennewald 2007, Theorem 1) *Let*

$$dX_t = u(t, X_t)dt + v(t, X_{t-})dN_t$$

be an n-dimensional Poisson process in matrix notation, where

$$X_t = \begin{bmatrix} X_1(t) \\ \vdots \\ X_n(t) \end{bmatrix}, \; u(t, X_t) = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}, \; v(t, X_{t-}) = \begin{bmatrix} v_{11} & \dots & v_{1m} \\ \vdots & & \vdots \\ v_{n1} & \dots & v_{nm} \end{bmatrix}, \; dN_t = \begin{bmatrix} dN_1(t) \\ \vdots \\ dN_m(t) \end{bmatrix}.$$

Let $g(t, x)$ be a C^1 map from $[0, \infty) \times \mathbb{R}^n$ into $\mathbb{R}$. Then the scalar process $Y_t = g(t, X_t)$ is again a Poisson process, and obeys the differential

$$dY_t = \frac{\partial}{\partial t} g(t, X_t)dt + \sum_{i=1}^{n} \frac{\partial}{\partial x_i} g(t, X) u_i dt + \sum_{j=1}^{m} \left(g(t, X_{t-} + v_j) - g(t, X_{t-})\right) dN_j(t), \quad (3.6)$$

where v_j denotes the jth column of the $n \times m$ matrix $v(t, X_{t-})$.

Remark 3.1.21 *A stochastic integral driven by a Poisson process is a random variable on the bounded interval $[t_0, t]$,*

$$X_t = X_{t_0} + \int_{t_0}^{t} f(s, N_s)dN_s,$$

satisfying the following properties (cf. García and Griego 1994)

1. $\int_{t_0}^{t} f(s, N_s)\, dN_s - \int_{t_0}^{t} f(N_s, s)\, \lambda ds$ *is a martingale,*
2. $\int_{t_0}^{t_2} f(s, N_s)dN_s = \int_{t_0}^{t_1} f(s, N_s)dN_s + \int_{t_1}^{t_2} f(s, N_s)dN_s$ *for* $t_0 \leq t_1 \leq t_2$,
3. $E(X_{t_2} - X_{t_1}) = E\left(\int_{t_1}^{t_2} f(s, N_s)dN_s\right) = E\left(\int_{t_1}^{t_2} f(s, N_s)\lambda ds\right)$ *for* $t_0 \leq t_1 \leq t_2$.

Example 3.1.22 *Consider two dependent stochastic processes*

$$dX_t = u(t)dt + \begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{bmatrix} dN_t,$$

or equivalently, consider a two-dimensional system of stochastic differential equations,

$$\begin{aligned} dX_1(t) &= u_1(t)dt + v_{11}dN_1(t) + v_{12}dN_2(t), \\ dX_2(t) &= u_2(t)dt + v_{21}dN_1(t) + v_{22}dN_2(t), \end{aligned}$$

and let $g(t, x)$ a C^1 function $[0, \infty) \times \mathbb{R}^2 \to \mathbb{R}$. Then, the process $Y_t = g(t, X_t)$ obeys

$$\begin{aligned} dY_t &= \frac{\partial}{\partial t} g(t, X_t)dt + \left[\begin{array}{cc} \frac{\partial}{\partial x_1} g(t, X_t) & \frac{\partial}{\partial x_2} g(t, X_t) \end{array} \right] u(t)dt \\ &\quad + \left[\begin{array}{cc} g(t, X_{t-} + v_1) - g(t, X_{t-}) & g(t, X_{t-} + v_2) - g(t, X_{t-}) \end{array} \right] dN_t, \end{aligned}$$

where

$$v_1 \equiv \left[\begin{array}{c} v_{11} \\ v_{21} \end{array} \right], \quad v_2 \equiv \left[\begin{array}{c} v_{12} \\ v_{22} \end{array} \right].$$

This feature is frequently encountered in economic models when there is one economy-wide source of uncertainty, say new technologies arrive or commodity price shocks occur according to some Poisson process, and many variables in this economy (for example relative prices) are affected simultaneously by this shock (cf. Wälde 2009, Lemma 10.1).

Computing differentials for jump-diffusions

This section replicates a version of Itô's formula (change of variables) for a setting where variables are driven by Brownian motion and Poisson processes, henceforth *jump-diffusion* models (see Sennewald 2007, chap. 6). Similar rules for computing differentials can be found in Wälde (2009, chap. 10), a more general version is in Protter (2004, Theorem 32).

Theorem 3.1.23 (Itô's formula for jump-diffusions) *Let*

$$dX_t = a(t, X_t)dt + b(t, X_t)dB_t + c(t, X_{t-})dN_t$$

be a scalar stochastic process, where B_t is a scalar Brownian motion, N_t is a scalar Poisson process, both forcing processes being stochastically independent. Let $g(t, x)$ be a C^2 function on $[0, \infty) \times \mathbb{R}$. Then, $Y_t = g(t, X_t)$ obeys the differential

$$\begin{aligned} dY_t &= \left(\frac{\partial}{\partial t} g(t, X_t) + a(t, X_t) \frac{\partial}{\partial x} g(t, X_t) + \tfrac{1}{2} b^2(t, X_t) \frac{\partial^2}{(\partial x)^2} g(t, X_t) \right) dt \\ &\quad + b(t, X_t) \frac{\partial}{\partial x} g(t, X_t) dB_t + (g(t, X_{t-} + c(t, X_{t-})) - g(t, X_{t-}))dN_t. \end{aligned} \tag{3.7}$$

Note that as long as no jump occurs, $dN_t = 0$, we simply obtain the case of a scalar Brownian motion. If a jump occurs, $dN_t = 1$, the system jumps by $c(t, X_{t-})$, and continues as a Brownian motion with drift until the next jump occurs. Thus, the rule (3.7) simply extends previous rules by allowing for possible jump terms.

Exercise 3.1.24 (Option pricing) *Let the price of a stock, S_t follow*

$$dS_t = \alpha S_t dt + \sigma S_t dB_t.$$

B_t is a standard Brownian motion, i.e., α denotes the (instantaneous) expected rate of return and σ^2 its variance. Suppose the price of an option is $Y_t = g(t, S_t)$. Let the price of a riskless asset, P_t, follow $dP_t = rP_t dt$. Obtain a pricing formula for Y_t, given the portfolio strategy of holding n_1 units of stocks, n_2 units of options, and n_3 units of the riskless asset.

This exercise employs Itô's formula to derive a pricing formula for options as in Black and Scholes (1973) closely following Merton (1976).

3.1.4 Solutions of stochastic differential equations

As with linear ODEs, the solution of a linear stochastic differential equation (SDE) can be found explicitly. The method of solution also involves an integrating factor, or equivalently, a *fundamental solution* of an associated homogeneous differential equation.

SDEs driven by Brownian motions

Below we will describe solution techniques for scalar SDEs, while extensions to linear systems of SDEs are straightforward (cf. Kloeden and Platen 1999, chap. 4.2, 4.8),

$$dX_t = (a_1(t)X_t + a_2(t))dt + (b_1(t)X_t + b_2(t))dB_t, \tag{3.8}$$

where the coefficients a_1, a_2, b_1, b_2 are specified functions of time t or constants. When all coefficients are constants the SDE is *autonomous* and its solutions are homogeneous Markov processes. When $a_2(t) \equiv 0$ and $b_2(t) \equiv 0$, (3.8) reduces to the *homogeneous* linear SDE

$$dX_t = a_1(t)X_t dt + b_1(t)X_t dB_t, \tag{3.9}$$

with *multiplicative noise*. When $b_1(t) \equiv 0$, in (3.8) the SDE has the form

$$dX_t = (a_1(t)X_t + a_2(t))dt + b_2(t)dB_t, \tag{3.10}$$

with *additive noise*, and the SDE is said to be linear in the narrow-sense.

Consider the homogeneous equation associated with (3.10), $dX_t = a_1(t)X_t dt$, which is an ODE with the fundamental solution

$$\Phi_{t,t_0} \equiv \exp\left(\int_{t_0}^{t} a_1(s)ds\right). \tag{3.11}$$

Applying Itô's formula (3.7) to the transformation $g(t,x) = \Phi_{t,t_0}^{-1}x$, we obtain

$$\begin{aligned} d(\Phi_{t,t_0}^{-1}X_t) &= -a_1(t)\Phi_{t,t_0}^{-1}X_t dt + \Phi_{t,t_0}^{-1}dX_t \\ &= -a_1(t)\Phi_{t,t_0}^{-1}X_t dt + a_1(t)\Phi_{t,t_0}^{-1}X_t dt + a_2(t)\Phi_{t,t_0}^{-1}dt + b_2(t)\Phi_{t,t_0}^{-1}dB_t \\ &= a_2(t)\Phi_{t,t_0}^{-1}dt + b_2(t)\Phi_{t,t_0}^{-1}dB_t. \end{aligned}$$

Observe that Φ_{t,t_0}^{-1} can be interpreted as an integrating factor for (3.10). Integrating both sides gives the solution to the SDE with *additive noise* (3.10),

$$X_t = \Phi_{t,t_0}\left(X_{t_0} + \int_{t_0}^{t} a_2(s)\Phi_{s,t_0}^{-1}ds + \int_{t_0}^{t} b_2(s)\Phi_{s,t_0}^{-1}dB_s\right), \tag{3.12}$$

where Φ_{t,t_0} is given in (3.11).

Exercise 3.1.25 *Use Itô's formula to prove that (3.12) indeed is a solution of (3.10).*

For the linear case with multiplicative noise, we may use the result for the linear SDE with additive noise. It follows by Itô's formula that $d(\ln X_t)$ for the *homogeneous* equation of the SDE with multiplicative noise in (3.9) obeys

$$d(\ln X_t) = a_1(t)dt - \tfrac{1}{2}b_1^2(t)dt + b_1(t)dB_t,$$

which becomes a linear SDE with additive noise in $\ln X_t$. Integrating both sides gives the fundamental solution to the homogeneous SDE with multiplicative noise (3.9),

$$X_t = X_{t_0}\exp\left(\int_{t_0}^{t}\left(a_1(s) - \tfrac{1}{2}b_1^2(s)\right)ds + \int_{t_0}^{t} b_1(s)dB_s\right). \tag{3.13}$$

In fact, the solution in (3.13) coincides with (3.11) for $b_1(t) \equiv 0$. In that we may use the integrating factor for the general linear SDE in (3.8),

$$\Phi_{t,t_0} \equiv \exp\left(\int_{t_0}^{t}\left(a_1(s) - \tfrac{1}{2}b_1^2(s)\right)ds + \int_{t_0}^{t} b_1(s)dB_s\right), \tag{3.14}$$

where $d\left(\Phi_{t,t_0}\right) = a_1(t)\Phi_{t,t_0}dt + b_1(t)\Phi_{t,t_0}dB_t$. Observe that both random variables Φ_{t,t_0} and X_t from (3.8) have stochastic differentials involving the same process B_t, so Itô's formula must be used with the transformation $g(x_1, x_2) = g(\Phi_{t,t_0}^{-1}, X_t)$. After some algebra,

$$d(\Phi_{t,t_0}^{-1} X_t) \quad = \quad (a_2(t) - b_1(t)b_2(t))\,\Phi_{t,t_0}^{-1}dt + b_2(t)\Phi_{t,t_0}^{-1}dB_t.$$

Integrating both sides gives the solution to the *general linear* SDE in (3.8) as

$$X_t = \Phi_{t,t_0}\left(X_{t_0} + \int_{t_0}^{t} (a_2(s) - b_1(s)b_2(s))\,\Phi_{s,t_0}^{-1}ds + \int_{t_0}^{t} b_2(s)\Phi_{s,t_0}^{-1}dB_s\right), \qquad (3.15)$$

where Φ_{t,t_0} is given in (3.14). Observe that for $b_1 \equiv 0$, the general solution in (3.15) indeed reduces to solution of the narrow-sense linear SDE in (3.12).

Exercise 3.1.26 (Langevin equation) *A molecular bombardment of a speck of dust on a water surface results into a Brownian motion. Compute the velocity, X_t, of a particle when the acceleration of the particle obeys $dX_t = -aX_tdt + bdB_t$.*

Example 3.1.27 (Option pricing) *Consider an European call, i.e., an option which can be exercised only at maturity, gives the holder the right to buy the asset at fixed maturity T and price $\bar{S}$ (strike price). Observe that the value of the option, $Y_t = g(t, S_t)$, satisfies*

$$Y_t = g(t, 0) = 0, \quad Y_T = g(T, S_T) = \max\{0, S_T - \bar{S}\}.$$

Using these boundary conditions, the partial differential equation (PDE) for the price of the option, $\frac{\partial Y_t}{\partial t} = rY_t - rS_t\frac{\partial Y_t}{\partial S} - \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 Y_t}{\partial S^2}$ has the solution

$$Y_t = \Phi(d_1(t, S_t))S_t - e^{-(T-t)r}\Phi(d_2(t, S_t))\bar{S} \quad \textit{with}$$

$$d_1(t, S_t) = \frac{\ln(S_t/\bar{S}) + \left(r + \frac{1}{2}\sigma^2(T-t)\right)}{\sigma\sqrt{T-t}}, \quad d_2(t, S_t) = \frac{\ln(S_t/\bar{S}) + \left(r - \frac{1}{2}\sigma^2(T-t)\right)}{\sigma\sqrt{T-t}},$$

where $\Phi(\cdot)$ is the cdf of the standard normal distribution (Black and Scholes 1973, p.644).

Remark 3.1.28 (Computing moments) *If we take the expectation of the integral form of equation (3.8) and use the property of an Itô stochastic integral, $E\int f(s, X_s)dB_s = 0$, we obtain an ODE for the mean, $m_1(t) \equiv E(X_t)$, namely*

$$dm_1(t) = (a_1(t)m_1(t) + a_2(t))dt.$$

Similarly, the second moment, $m_2(t) \equiv E(X_t^2)$, satisfies the ODE,

$$dm_2(t) = \left(\left(2a_1(t) + b_1^2(t)\right) m_2(t) + 2(a_2(t) + b_1(t)b_2(t))m_1(t) + b_2^2(t)\right) dt.$$

The solutions to the ODEs above are the time-dependent moments of the distribution of X_t. Letting $t \to \infty$, we may obtain the moments of some limiting distribution.

Exercise 3.1.29 (Geometric Brownian motion) *Use Itô's formula to solve*

$$dP_t = \mu P_t dt + \sigma P_t dB_t,$$

where P_t denotes the size of population. Obtain the expected rate of population growth.

SDEs driven by Poisson processes

Consider the following scalar SDE (cf. García and Griego 1994),

$$dX_t = (a_1(t)X_t + a_2(t))dt + (b_1(t)X_{t-} + b_2(t))dN_t, \tag{3.16}$$

where the coefficients a_1, a_2, b_1, b_2 are specified functions of time t or constants. Similar to the approach of solving SDEs with driven by Brownian motions, we use an educated guess of the fundamental solution to the homogeneous differential equation of (3.16) as

$$\Phi_{t,t_0} \equiv \exp\left(\int_{t_0}^{t} a_1(s)ds + \int_{t_0}^{t} \ln(1 + b_1(s))dN_s\right), \tag{3.17}$$

where $d\left(\Phi_{t,t_0}\right) = a_1(t)\Phi_{t,t_0}dt + b_1(t)\Phi_{t-,t_0}dN_t$. Observe that both random variables Φ_{t,t_0} and X_t from (3.16) have stochastic differentials involving the same process N_t, so the Itô formula must be used with the transformation $g(x_1, x_2) = g(\Phi_{t,t_0}^{-1}, X_t)$. After some algebra,

$$d(\Phi_{t,t_0}^{-1}X_t) \quad = \quad a_2(t)\Phi_{t,t_0}^{-1}dt + \frac{b_2(t)}{1 + b_1(t)}\Phi_{t-,t_0}^{-1}dN_t.$$

Integrating both sides gives the solution to the general linear SDE in (3.16) as

$$X_t = \Phi_{t,t_0}\left(X_{t_0} + \int_{t_0}^{t} a_2(s)\Phi_{s,t_0}^{-1}ds + \int_{t_0}^{t} \frac{b_2(s)}{1 + b_1(s)}\Phi_{s-,t_0}^{-1}dN_s\right), \tag{3.18}$$

where the fundamental solution Φ_{t,t_0} is given in (3.17).

Exercise 3.1.30 (Geometric Poisson process) *Use Itô's formula to solve*

$$dA_t = \mu A_t dt + \gamma A_{t-} dN_t, \quad \gamma > -1.$$

Obtain the expected growth rate of and illustrate your result graphically.

Reducible stochastic differential equations

With an appropriate transformation $X_t = U(t, Y_t)$ a nonlinear SDE

$$dY_t = a(t, Y_t)dt + b(t, Y_t)dB_t + c(t, Y_{t-})dN_t,$$

may be reduced to a linear SDE in X_t,

$$dX_t = (a_1(t)X_t + a_2(t))\, dt + (b_1(t)X_t + b_2(t))\, dB_t + (c_1(t)X_{t-} + c_2(t))\, dN_t.$$

In principle, all reducible ODEs are candidates for reducible SDEs. A reducible SDE has an explicit solution provided that the integrals exist Kloeden and Platen (1999, chap. 4.4).

For illustration, consider a stochastic *Bernoulli equation*,

$$dX_t = (aX_t^n + bX_t)dt + cX_t dB_t + J_t X_{t-} dN_t,$$

a stochastic process with constants coefficients $a, b, c \in \mathbb{R}$, and polynomial drift of degree $n \neq 1$. Let B_t denote a Brownian motion, N_t is a Poisson process, and J_t is a stochastically independent random variable where the first two moments exist. It can be shown that

$$X_t = \Phi_{t,t_0} \left(X_{t_0}^{1-n} + (1-n)a \int_{t_0}^{t} \Phi_{s,t_0}^{n-1} ds \right)^{\frac{1}{1-n}}, \tag{3.19}$$

where

$$\Phi_{t,t_0} = \exp\left(\left(b - \tfrac{1}{2}c^2\right)(t - t_0) + (B_t - B_{t_0})c + \int_{t_0}^{t} \ln(1 + J_u) dN_u \right).$$

A special case for $n = 2$ is referred to as the *stochastic Verhulst equation.*

3.1.5 An example: Merton's model of growth under uncertainty

This section introduces Merton's asymptotic theory of growth under uncertainty, where the dynamics of the capital-to-labor ratio is a diffusion-type stochastic process. The particular source of uncertainty chosen is the population size although the analysis would be equally applicable to other specifications (Merton 1999, chap. 17).

Consider a one-sector growth model of the Solow type with a constant-returns-to-scale, strictly concave production function, $Y_t = F(K_t, L_t)$, where Y_t denotes output, K_t denotes the capital stock, L_t denotes the labor force. Capital will be accumulated in a deterministic way according to

$$dK_t = (Y_t - \delta K_t - C_t)dt, \tag{3.20}$$

where δ is the rate of physical depreciation and C_t is aggregate consumption.

The source of uncertainty is the population size L_t. Suppose that

$$dL_t = nL_t dt + \sigma L_t dB_t, \tag{3.21}$$

where $\{B_t\}_{t=0}^{\infty}$ is a standard Brownian motion. By inspection of the random variable L_t will have a log-normal distribution with $E_0(\ln L_t) = \ln L_0 + (n - \frac{1}{2}\sigma^2)t$, and $Var_0(\ln L_t) = \sigma^2 t$.

Exercise 3.1.31 *Show that $E_0(\ln L_t) = \ln L_0 + (n - \frac{1}{2}\sigma^2)t$ and $Var_0(\ln L_t) = \sigma^2 t$, whereas $E_0(L_t) = L_0 e^{nt}$ and $Var_0(L_t) = e^{2nt}\big(e^{\sigma^2 t} - 1\big)$.*

As in the certainty model, the dynamics can be reduced to a one-dimensional process for variables in intensive form. Using Itô's formula, the capital stock per capita, $k_t \equiv K_t/L_t$ follows the stochastic differential equation

$$\begin{aligned} dk_t &= (y_t - \delta k_t - c_t)dt - k_t n dt + k\sigma^2 dt - \sigma k_t dB_t \\ &= (y_t - (\delta + n - \sigma^2)k_t - c_t)dt - \sigma k_t dB_t \\ &\equiv b(k_t)dt - \sqrt{a(k_t)}dB_t, \end{aligned}$$

where $y_t \equiv Y_t/L_t = f(k_t)$ denotes output, and $c_t \equiv C_t/L_t = (1-s)Y_t/L_t$ is consumption both in intensive form (per capita). Hence, the accumulation equation in intensive form is a diffusion process. In particular, the transition probabilities for k_t are completely determined by the functions $b(k_t) \equiv sf(k_t) - (\delta + n - \sigma^2)k_t$ as the (instantaneous) expected change in k_t per unit of time, and $a(k_t) \equiv (\sigma k_t)^2$ as the (instantaneous) variance.

Exercise 3.1.32 *Obtain the non-stochastic steady state values for capital in intensive form and for the rental rate of capital.*

Before going on to analyze the distributional characteristics of k_t, it is important to distinguish between the stochastic process for k_t and the one for K_t. While the sample path for k_t is not differentiable, the sample path for K_t is. Thus, unlike in portfolio models, there is no current uncertainty, but only future uncertainty. The returns to capital (and labor) over the next instant are known with certainty (Merton 1999, p.584),

$$r_t = f'(k_t), \quad w_t \equiv f(k_t) - kf'(k_t). \tag{3.22}$$

Just as in the certainty model the existence and properties of the steady-state economy can be examined, so can they be for the uncertainty model. Instead of there being a unique point k in the steady-state, there is now a unique *distribution* for k which is time and initial condition independent and toward which the stochastic process for k_t tends.

Throughout the analysis we assume that the following set of sufficient conditions for existence are satisfied Merton (1999, p.585): (i) $f(k_t)$ is concave and satisfies the Inada conditions, (ii) $c(k_t) < f(k_t)$ for all $k_t < \bar{k}$ for some positive $\bar{k}$, and (iii) $\delta + n - \sigma^2 > 0$. Then it is possible to deduce a general functional representation for the steady-state probability distribution. Let $\pi_k(\cdot)$ be the *asymptotic* or *limiting* density function (steady-state density) for the capital stock per effective worker. It will satisfy (Merton 1999, Appendix 17B)

$$\pi_k(k) \equiv \lim_{t\to\infty} \pi_k(k_t) = \frac{\mathbb{C}_0}{a(k_t)} \exp\left[\int^{k_t} \frac{2b(x)}{a(x)} dx\right], \tag{3.23}$$

where $\mathbb{C}_0$ is a constant chosen so that $\int_0^\infty \pi_k(x)dx = 1$.

A Cobb-Douglas economy (constant-savings-function)

There is a specific functional form where the steady-state distributions for all variables can be solved for in closed form. If it is assumed that the production function is Cobb-Douglas, $f(k_t) = k_t^\alpha$, and that saving is a constant fraction of output (s is a constant, $0 < s \leq 1$), then by substituting the particular functional form in (3.23) it can be shown that

$$\begin{aligned} \pi_k(k) &= \frac{\mathbb{C}_0}{\sigma^2 k^2} \exp\left[\int^k \frac{2sx^\alpha - 2(\delta + n - \sigma^2)x}{\sigma^2 x^2} dx\right] \\ &= \frac{\mathbb{C}_0}{\sigma^2} k^{\frac{-2(\delta+n)}{\sigma^2}} \exp\left[-\frac{2s}{(1-\alpha)\sigma^2} k^{-(1-\alpha)}\right]. \end{aligned} \tag{3.24}$$

To specify the constant term, we employ the condition $\int_0^\infty \pi_k(k)dk = 1$. It can be obtained indirectly from the Gamma(γ, ω) distribution (see Merton 1999, chap. 17.4).

Remark 3.1.33 (Gamma(γ, ω) distribution) *The probability model, i.e., the collection of the density function indexed by unknown parameters (γ, ω) and the parameter space, of* *the Gamma distribution is*

$$\left\{ f(x;(\gamma,\omega)) = \frac{\omega^{-1}}{\Gamma(\gamma)}\left(\frac{x}{\omega}\right)^{\gamma-1}\exp\left[-\frac{x}{\omega}\right], \quad (\gamma,\omega)\in\mathbb{R}_+^2, \quad x\in\mathbb{R}_+ \right\}.$$

It follows that $E(X)=\gamma\omega$, and $Var(X)=\gamma\omega^2$ (see Spanos 1999, p.140).

Given that capital rewards are $r_t = \alpha k_t^{\alpha-1}$, we may use the change of variable formula for densities in Theorem 1.7.33, $\pi_r(r) = \pi_k(k)/(|dr/dk|)$, to obtain

$$\pi_r(r) \quad = \quad \frac{\mathbb{C}_0}{(1-\alpha)\alpha\sigma^2}\left(\frac{r}{\alpha}\right)^{\frac{\sigma^2-2(\delta+n)}{(\alpha-1)\sigma^2}-1}\exp\left[-\frac{2s}{(1-\alpha)\alpha\sigma^2}r\right] \equiv \frac{\mathbb{C}_0\omega^{\gamma-1}\alpha^{-\gamma}}{(1-\alpha)\sigma^2}\left(\frac{r}{\omega}\right)^{\gamma-1}\exp\left[-\frac{r}{\omega}\right],$$

where we defined $\gamma \equiv \frac{2(\delta+n)-\sigma^2}{(1-\alpha)\sigma^2} > 0$ and $\omega \equiv \frac{(1-\alpha)\alpha\sigma^2}{2s}$. By inspection, r has a Gamma(γ,ω) distribution, where

$$\mathbb{C}_0 = \frac{(1-\alpha)\sigma^2}{\Gamma[\gamma]\omega^\gamma\alpha^{-\gamma}}$$

is needed to satisfy the property of a density function.

Exercise 3.1.34 *Obtain the stochastic differential for the rental rate of capital, $r_t = \alpha k_t^{\alpha-1}$, and interpret the coefficients of the resulting stochastic differential equation. Compute the solution to the equation and derive the mean and the variance of the limiting distribution.*

Similarly, the density functions and moments of the distributions for all the variables can be deduced from (3.24) and using Theorem 1.7.33. The analysis would be identical for other types of consumption functions, in particular, for closed-form policy functions resulting from stochastic control problems.

3.2 Stochastic dynamic programming

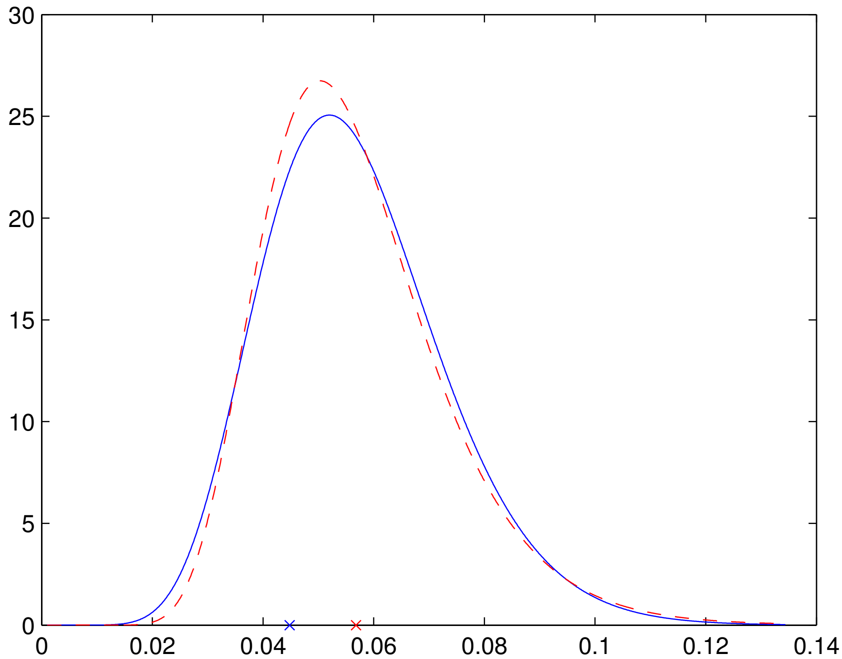
Literature: Dixit and Pindyck (1994, chap. 3,4), Kloeden and Platen (1999, chap. 6.5), Chang (2004, chap. 4), Turnovsky (2000, chap. 15), Wälde (2009, chap. 11)
This section considers how to control a dynamic system subject to random disturbances, studying optimal stochastic control problems under Brownian and Poisson uncertainty.

Consider the following typical infinite horizon stochastic control problem,

$$\max E\int_0^\infty e^{-\rho t}f(t,X_t,u_t)dt \quad s.t. \quad dX_t = a(t,X_t,u_t)dt + b(t,X_t,u_t)dB_t + c(t,X_{t-},u_{t-})dN_t,$$
$$X_0 = x,\ (B_0,N_0)=z,\ (x,z)\in\mathbb{R}\times\mathbb{R}_+^2,$$

where $\{B_t\}_{t=0}^\infty$ is a standard Brownian motion, and $\{N_t\}_{t=0}^\infty$ is a Poisson process.

Figure 3.1: Asymptotic Gamma(γ, ω) distribution for capital rewards r_t (solid), compared to a Log-Normal distribution (dashed) with mean and variance, $E(r) = \gamma\omega$, $Var(r) = \gamma\omega^2$; calibrated parameter values are $(\rho, \alpha, \theta, \delta, n, \sigma) = (.04, .6, .6, .025, .025, .2)$.



3.2.1 Bellman's principle

Closely following Sennewald (2007), we obtain the *Bellman equation* at time s as

$$\rho V(s,x) = \max_{u \in U} \left\{ f(s,x,u) + \frac{1}{dt} E_s dV(s,x) \right\},$$

which is a necessary condition for optimality. Using Itô's formula (change of variables),

$$\begin{aligned} dV(s,x) &= \left(\frac{\partial}{\partial t} V(s,x) + a(s,x,u) \frac{\partial}{\partial x} V(s,x) + \tfrac{1}{2} b^2(s,x,u) \frac{\partial^2}{\partial x^2} V(s,x) \right) dt \\ &\quad + b(s,x,u) \frac{\partial}{\partial x} V(s,x) dB_t + (V(s, x + c(s,x,u)) - V(s,x)) dN_t. \end{aligned}$$

If we take the expectation of the integral form, and use the property of stochastic integrals,

$$\begin{aligned} E_0 \int_0^t b(s,x,u) \frac{\partial}{\partial x} V(s,x) dB_s &= 0, \\ E_0 \int_0^t (V(s, x + c(s,x,u)) - V(s,x)) dN_s &= \int_0^t (V(s, x + c(s,x,u)) - V(s,x)) \lambda ds, \end{aligned}$$

assuming that the above integrals exist, in particular that the reward function satisfies some boundedness condition (Sennewald 2007, Theorem 2), we may write

$$\begin{aligned}E_s dV(s,x) &= \left(\frac{\partial}{\partial t}V(s,x) + a(s,x,u)\frac{\partial}{\partial x}V(s,x) + \tfrac{1}{2}b^2(s,x,u)\frac{\partial^2}{\partial x^2}V(s,x)\right)dt \\ &\quad +(V(s,x+c(s,x,u)) - V(s,x))\lambda dt \\ &\equiv \left(V_t + a(s,x,u)V_x + \tfrac{1}{2}b^2(s,x,u)V_{xx} + (V(s,x+c(s,x,u)) - V(s,x))\lambda\right)dt,\end{aligned}$$

and the Bellman equation becomes (suppressing functional arguments)

$$\rho V(s,x) = \max_{u\in U}\left\{f(\cdot) + V_t + a(\cdot)V_x + \tfrac{1}{2}b^2(\cdot)V_{xx} + (V(s,x+c(\cdot)) - V(s,x))\lambda\right\}.$$

A neat result about the continuous-time formulation under uncertainty is that the Bellman equation is, in effect, a deterministic differential equation because the expectation operator disappears (Chang 2004, p.118). Hence, the *first-order condition* reads

$$f_u(\cdot) + a_u(\cdot)V_x(s,x) + \tfrac{1}{2}b_u^2(\cdot)V_{xx}(s,x) + V_x(s,x+c(s,x,u))c_u(\cdot)\lambda = 0.$$

In contrast to deterministic control problems, we now obtain a second-order effect resulting from the Brownian motion forcing term, and a first-order term linking the utility before and after a jump resulting from the Poisson process.

For the *evolution of the costate* we use the maximized Bellman equation,

$$\rho V(s,x) = f(\cdot) + V_t + a(\cdot)V_x + \tfrac{1}{2}b^2(\cdot)V_{xx} + (V(s,x+c(\cdot)) - V(s,x))\lambda,$$

where the optimal control is a function of the state variables (the dependence on the state vector z has been neglected for notational convenience). We then make use of the envelope theorem to compute the costate,

$$\begin{aligned}\rho V_x(s,x) &= f_x(s,x,u(x)) + V_{tx} + a_x(s,x,u(x))V_x + a(s,x,u(x))V_{xx} \\ &\quad + \tfrac{1}{2}\left(b_x^2(s,x,u(x))V_{xx} + b^2(s,x,u(x))V_{xxx}\right) \\ &\quad +(V_x(s,x+c(t,x,u(x)))(1+c_x(\cdot)) - V_x(s,x))\lambda.\end{aligned}$$

Collecting terms we obtain

$$\begin{aligned}\left(\rho - a_x(\cdot) + \lambda\right)V_x &= f_x(\cdot) + V_{tx} + a(\cdot)V_{xx} + \tfrac{1}{2}b^2(\cdot)V_{xxx} + \tfrac{1}{2}b_x^2(\cdot)V_{xx} \\ &\quad +V_x(s,x+c(\cdot))(1+c_x(\cdot))\lambda. \qquad (3.25)\end{aligned}$$

Using Itô's formula, the costate obeys

$$\begin{aligned} dV_x(s,x) &= \left(V_{tx} + a(\cdot)V_{xx}(s,x) + \tfrac{1}{2}b^2(\cdot)V_{xxx}\right) dt \\ &\quad + V_{xx}b(\cdot)dB_t + \left(V_x(s,x+c(s,x,u)) - V_x(s,x)\right) dN_t. \end{aligned}$$

Inserting (3.25) yields

$$\begin{aligned} dV_x(s,x) &= \left((\rho - a_x(\cdot) + \lambda)\, V_x(s,x) - f_x(\cdot) - \tfrac{1}{2}b_x^2(\cdot)V_{xx} - \lambda(1 + c_x(\cdot))V_x(s,x+c(\cdot))\right) dt \\ &\quad + V_{xx}b(\cdot)dB_t + \left(V_x(s,x+c(s,x,u)) - V_x(s,x)\right) dN_t, \end{aligned}$$

which describes the evolution of the costate variable.

As the final step, to obtain the *Euler equation* we use the first-order condition,

$$f_u(\cdot) = -a_u(\cdot)V_x(s,x) - \tfrac{1}{2}b_u^2(\cdot)V_{xx}(s,x) - V_x(s,x+c(s,x,u))c_u(\cdot)\lambda.$$

to substitute unknown functions by known functions. In general we are not able to eliminate shadow prices form the resulting equation. In the case where $b(\cdot) \equiv 0$ and $c(\cdot) \equiv 0$ we obtain the deterministic version of the Euler equation in (1.72).

Exercise 3.2.1 (Optimal saving under Poisson uncertainty) *Consider the problem,*

$$\begin{aligned} \max E \int_0^\infty e^{-\rho t} u(C_t) dt \quad s.t. \quad & dK_t = \left(AK_t^\alpha L^{1-\alpha} - \delta K_t - C_t\right) dt - \gamma K_{t-} dN_t, \\ & K_0 = x,\ N_0 = z, \quad (x,z) \in \mathbb{R}_+^2, \quad 0 < \gamma < 1, \end{aligned}$$

where N_t denotes the number of disasters up to time t, occasionally destroying γ percent of the capital stock K_t with an arrival rate $\lambda > 0$. Suppose that $u' > 0$ and $u'' < 0$. Solve the planners problem and find the optimal consumption path using the inverse function.

Remark 3.2.2 (Hyperbolic utility and attitudes toward risk) *The class of hyperbolic absolute risk aversion (HARA) include the widely used (isoelastic) power utility or constant relative risk aversion (CRRA), (negative) exponential utility or constant absolute risk aversion (CARA), and quadratic utility (Merton 1999, chap. 5.6),*

$$v(c) = \frac{\theta}{1-\theta}\left(\frac{\eta c}{\theta} + \delta\right)^{1-\theta}, \quad \theta \neq 0, \quad \eta > 0, \quad \frac{\eta c}{\theta} + \delta > 0, \quad \delta = 1 \quad if \quad \theta \to -\infty$$

whose measure of absolute risk aversion is positive and hyperbolic in consumption, i.e.

$$ARA(c) = -\frac{v''(c)}{v'(c)} = \left(\frac{\eta c}{\theta} + \delta\right)^{-1} \eta$$

which implies that $\delta > 0$ for $\theta < 0$. For CRRA utility ($\theta > 0$), use $\delta = 0$, for CARA utility let $\theta \to -\infty$ and use $\delta = 1$ (which gives negative exponential utility in the limit).

Exercise 3.2.3 *Let $u(c)$ be utility generated by consuming, $c > 0$. Find an expression for $u(c)$ of the type constant absolute risk aversion (CARA), and constant relative risk aversion (CRRA), respectively. Distinguish between the cases $\theta = 1$ and $\theta \neq 1$ for the latter.*

Example 3.2.4 (Deriving a budget constraint) *Consider an individual that invests in both a risky asset (bond with default risk) and a riskless asset (bond). Suppose the price of the risky asset obeys*

$$dv_t = \alpha v_t dt + \beta v_{t-} dN_t, \quad \beta > -1, \tag{3.26}$$

while on unit of the riskless asset gives instantaneous returns r, or equivalently, $db_t = rb_t dt$. Let the individual receive fixed income w, and have expenditures for consumption c_t. Consider a portfolio strategy which holds $n_1(t)$ units of the risky asset, and $n_2(t)$ units of riskless bonds. Then the value of this portfolio is $a_t = v_t n_1(t) + b_t n_2(t)$ and its differential obeys

$$da_t = dn_1(t)v_t + dn_2(t)b_t + (\alpha n_1(t)v_t + rn_2(t)b_t)dt + \beta n_1(t)v_{t-}dN_t.$$

Let θ_t denote the share of the risky asset, $\theta_t \equiv n_1(t)v_t/a_t$, such that $1 - \theta_t = n_2(t)b_t/a_t$, and

$$da_t = dn_1(t)v_t + dn_2(t)b_t + (\alpha\theta_t + (1 - \theta_t)r)a_t dt + \beta\theta_{t-}a_{t-}dN_t.$$

Since investors use their savings to accumulate assets,

$$dn_1(t)v_t + dn_2(t)b_t \;=\; (\pi_v v_t n_1(t) + \pi_b b_t n_2(t) + w_t - c_t)\, dt,$$

where π_v and π_b denote percentage dividend payments on the assets, respectively. Thus

$$\begin{aligned} da_t &= (\pi_v v_t n_1(t) + \pi_b b_t n_2(t) + w_t - c_t + (\alpha\theta_t + (1 - \theta_t)r)a_t)\, dt + \beta\theta_{t-}a_{t-}dN_t \\ &= (((\pi_v + \alpha)\theta_t + (1 - \theta_t)(r + \pi_b))\, a_t + w_t - c_t)\, dt + \beta\theta_{t-}a_{t-}dN_t \\ &\equiv (((r_v - r_b)\theta_t + r_b)\, a_t + w_t - c_t)\, dt + \beta\theta_{t-}a_{t-}dN_t, \end{aligned} \tag{3.27}$$

where we defined

$$r_v \equiv \pi_v + \alpha, \quad r_b \equiv \pi_b + r = \pi_b + \dot{b}_t/b_t,$$

as the return on the risky asset conditioned on no jumps, and the riskless asset, respectively. Both consist of dividend payments in terms of the asset price and the deterministic part.

Exercise 3.2.5 (Optimal consumption and portfolio choice) *Consider an individual portfolio decision problem between investing in a risky asset (bond with default risk) and a riskless asset (government bill). There is no dividend payments. Individual debt is bounded by the individual's lifetime labor income, discounted at the riskless rate, $a_t > -w/r \in A_t \subset \mathbb{R}$. Solve the household's problem*

$$\max E \int_0^\infty e^{-\rho t} u(c_t) dt \quad s.t. \quad da_t = (((\alpha - r)\theta_t + r) a_t + w - c_t) dt + \beta\theta_{t-} a_{t-} dN_t,$$
$$a_0 = x,\ N_0 = z, \quad (x, z) \in A_0 \times \mathbb{R}_+.$$

To avoid trivial investment problems (bang-bang), assume $r < \alpha + \beta\lambda < \alpha$, where from (3.26) $\lambda\beta$ is the expected jump in stock returns, and α its instantaneous drift term, $\beta > -1$.

This exercise builds on Sennewald and Wälde (2006).

3.2.2 An example: The matching approach to unemployment

This section introduces a simple model of the labor market that captures the salient features of the theory of unemployment (Pissarides 2000, chap. 1). Suppose entry into unemployment is an exogenous process, resulting from stochastic structural change or from new entry into the labor force. The transition out of unemployment is modeled as a trading process, with unemployed workers and firms with job vacancies wanting to trade labor services.

The central idea is that trade in the labor market is a decentralized economic activity. It is uncoordinated, time-consuming, and costly for both firms and workers. As a modeling device, the *matching function* captures this costly trading process. It gives the number of jobs formed at any moment in time as a function of the number of workers looking for jobs, the number of firms looking for workers, and possibly some other variables. It has its parallel in the neoclassical assumption of the existence of an aggregate production function.

Suppose there are L workers in the labor force. Let u denote the fraction of unmatched workers, i.e., the *unemployment rate*, and v the number of vacant jobs as a fraction of the labor force, i.e., the *vacancy rate*. The number of job matches is

$$mL = m(uL, vL), \quad m_u > 0,\ m_v > 0,\ m_{uu} < 0,\ m_{vv} < 0,$$

defining the *matching function* with constant returns to scale (homogeneous of degree one). The job vacancies and unemployed workers that are matched at any point are randomly selected from the sets vL and uL. Hence the process that changes the state of vacant jobs is Poisson with rate $m(uL, vL)/(vL)$.

By the homogeneity of the matching function, $m(uL, vL)/(vL)$ is a function of the ratio of vacancies to unemployment only. It is convenient to introduce the v/u ratio as a separate variable, denoted by θ, and write the rate at which vacant jobs become filled as

$$q(\theta) \equiv m(u/v, 1) = m(1/\theta, 1), \quad q'(\theta) \leq 0.$$

During an infinitesimal small time interval, a vacancy is matched to an unemployed worker with probability $q(\theta)$, so the mean duration of a vacant job is $1/q(\theta)$. Unemployed workers move into employment according to a related Poisson process with rate

$$m(uL, vL)/(uL) = m(1, v/u) = m(1, \theta) = \theta q(\theta).$$

The mean duration of unemployment is $1/(\theta q(\theta))$. Thus unemployed workers find jobs more easily when there are more jobs relative to available workers, and firms with vacancies find workers more easily when there are more workers relative to the available jobs.

Without growth or turnover in the labor force, the mean number of workers who enter unemployment during an infinitesimal small time interval is $(1-u)L\lambda$ and the mean number who leave unemployment is $m(1, v/u)L = uL\theta q(\theta)$, where $\theta q(\theta)$ is the transition probability of the unemployed. The evolution of mean unemployment is given by the difference

$$d(uL) = (1-u)L\lambda dt - \theta q(\theta)uLdt \quad \Rightarrow \quad \dot{u} = (1-u)\lambda - \theta q(\theta)u.$$

Thus, in the steady state, the mean of unemployment is in terms of the two transition rates,

$$(1-u)\lambda = \theta q(\theta)u \quad \Rightarrow \quad u = \frac{\lambda}{\lambda + \theta q(\theta)}. \tag{3.28}$$

It implies that for given λ and θ, there is a unique equilibrium mean unemployment rate. In that λ is a model parameter whereas θ is yet an unknown. It can be shown that θ can be determined by an equation derived from the assumption of profit maximization and that it is unique and independent of u. Hence, the solution for u is also unique. By the properties of the matching function, (3.28) can be represented in the vacancy-unemployment space by a downward-sloping and convex curve (known as the *Beveridge curve*).

Let us now formulate an individual's budget constraint which incorporates this idea of labor matching (Wälde 2009, chap. 11.2). Suppose that Z_t denotes labor income which has two uncertain states,

$$Z_t = \begin{cases} w & \text{wage income when employed} \\ b & \text{unemployment benefits else} \end{cases},$$

where

$$dZ_t = -(w-b)dN_1(t) + (w-b)dN_2(t), \quad N_t = \begin{bmatrix} N_1(t) \\ N_2(t) \end{bmatrix}. \tag{3.29}$$

N_t is a two-dimensional Poisson process where $N_1(t)$ is counting the numbers of separations, whereas $N_2(t)$ is counting the number of matches, with state dependent arrival rates,

$$\lambda_1(Z_t) = \begin{cases} \lambda & Z_t = w & \text{(individual is currently employed)} \\ 0 & Z_t = b & \text{(is unemployed)} \end{cases},$$

and

$$\lambda_2(Z_t) = \begin{cases} 0 & Z_t = w & \text{(individual is currently employed)} \\ \theta q(\theta) & Z_t = b & \text{(is unemployed)} \end{cases}.$$

Hence, λ_1 and λ_2 are the probabilities of loosing and finding a job (or match), respectively. The exogenous arrival rate λ_2 is related to the matching function, which can be interpreted as the probability of filling a vacancy. Hence, a simple individual budget constraint reads

$$da_t = (ra_t + Z_t - c_t)dt, \tag{3.30}$$

where labor income is stochastic and given by (3.29).

Exercise 3.2.6 (Matching on the labor market) *Consider the control problem*

$$\max E \int_0^\infty e^{-\rho t} u(c_t)dt \qquad s.t. \qquad \begin{aligned} &da_t = (ra_t + Z_t - c_t)dt, \\ &dZ_t = -(w-b)dN_1(t) + (w-b)dN_2(t), \\ &a_0 = x,\ Z_0 = z, \quad (x,z) \in \mathbb{R}_+ \times \{w,b\}, \end{aligned}$$

where Z_t denotes two-states labor income, and the Poisson process N_t counts the number of changing states from unemployment to employment, and vice versa. Suppose that $u' > 0$ and $u'' < 0$. Obtain the expected present value of being unemployed and of being employed.

3.2.3 An example: Wälde's model of endogenous growth cycles

Consider a closed economy with competitive markets. Suppose that technological progress is labor augmenting and embodied in capital. A capital good K_j of vintage j allows workers to produce with a labor productivity A^j, where $A > 1$ (Wälde 2005). Each vintage produces a single output good according to the production function

$$Y_j = K_j^{\alpha}(A^j L_j)^{1-\alpha}, \quad 0 < \alpha < 1,$$

where L_j denotes the amount of labor allocated to vintage j. The sum of labor employment over vintages equals aggregate constant labor supply, $\sum_{j=0}^{q} L_j = L$. Output is used for consumption, C_t, investment into physical capital, I_t, and venture-capital investment, R_t. Market clearing demands

$$Y_t \equiv \sum_{j=0}^{q} Y_j = C_t + R_t + I_t.$$

Allowing labor to be mobile across vintages, wage rates equalize, $w_t^L = Y_L$, and total output can be represented by a simple Cobb-Douglas production function

$$Y_t = K_t^{\alpha}(A^q L)^{1-\alpha}. \tag{3.31}$$

K_t is obtained by aggregating vintage specific capital stocks,

$$K_t \equiv B^{-q}K_0 + B^{1-q}K_1 + ... + K_q = \sum_{j=0}^{q} B^{j-q}K_j, \quad B = A^{\frac{1-\alpha}{\alpha}}, \tag{3.32}$$

defining the capital stock index in units of the consumption good (or the most recent vintage). As long capital goods are traded, the price of an installed unit of the most recent vintage q equals the price of the investment good (normalized to unity). Since the different vintages are perfect substitutes in production, the price of vintage j as from (3.32) is $v_j = B^{j-q}$.

Capital goods of vintage j are subject to physical depreciation at the rate δ. If net investment exceeds depreciation, the capital stock of this vintage accumulates according to

$$dK_j = (I_j - \delta K_j)dt, \quad j = 0, ..., q. \tag{3.33}$$

The objective of research is to develop capital goods that yield a higher labor productivity than existing capital goods, i.e., that of vintage 0 until the most recent vintage q. When research is successful, the capital stock of the next vintage $q + 1$ increases by the size of the first new machine, i.e., assumed to be a constant fraction of the capital stock index K_t,

$$dK_{q+1} = \kappa K_{t-} dN_t, \quad 0 < \kappa \ll 1,$$

where $\{N_t\}_{t=0}^{\infty}$ is a Poisson process at arrival rate

$$\lambda_t = (R_t/K_t)^{1-\gamma}. \tag{3.34}$$

This formulation removes the scale effect in the present model (cf. Wälde 2005). In contrast to quality-ladder models, the output of successful research is not only a blueprint, engineers actually construct a first machine that implies higher labor productivity. This first prototype (or pilot plant) can be regarded as the payoff for investment into research.

Exercise 3.2.7 (Cyclical endogenous growth) *Consider the planner's problem*

$$\max E \int_0^{\infty} e^{-\rho t} u(C_t) dt \qquad s.t. \qquad \begin{aligned} & dK_t = (I_t - \delta K_t)dt + (\kappa - s)K_{t-} dN_t, \\ & K_0 = x, \ N_0 = z, \quad (x,z) \in \mathbb{R}^2, \end{aligned}$$

where $s \equiv 1 - A^{-\frac{1-\alpha}{\alpha}}$ denotes the economic rate of depreciation, while $0 < \kappa \ll 1$ is the size of the first new machine of the next vintage. New capital goods are discovered at the arrival rate $\lambda_t = (R_t/K_t)^{1-\gamma}$, $0 < \gamma < 1$. Total output is produced according to $Y_t = K_t^{\alpha}(A^q L)^{1-\alpha}$, where market clearing demands $Y_t = C_t + R_t + I_t$. Suppose that $u' > 0$ and $u'' < 0$. Obtain the the Euler equation for optimal consumption and illustrate your results.

Bibliography

Atkinson, A. B., "Capital taxes, the redistribution of wealth and individual savings," *Rev. Econ. Stud.*, 1971, *38.*

Benhabib, Jess and Roberto Perli, "Uniqueness and Indeterminacy: On the Dynamics of Endogenous Growth," *J. Econ. Theory*, 1994, *63*, 113–142.

Black, Fischer and Myron Scholes, "The Pricing of Options and Corporate Liabilities," *J. Polit. Economy*, 1973, *81* (3), 637–654.

Chang, Fwu-Ranq, "The Inverse Optimal Problem: A Dynamic Programming Approach," *Econometrica*, 1988, *56* (1), 147–172.

———, *Stochastic optimization in continuous time*, Cambridge Univ. Press, 2004.

Clarida, Richard, Jordi Galí, and Mark Gertler, "Monetary policy rules and macroeconomic stability: Evidence and some theory," *Quart. J. Econ.*, 2000, pp. 147–180.

Dasgupta, Partha and Geoffrey Heal, "The Optimal Depletion of Exhaustible Resources," *Rev. Econ. Stud.*, 1974, *41*, 3–28.

Dixit, A. K. and R. S. Pindyck, *Investment under uncertainty*, Princeton Univ. Press, 1994.

García, M. A. and R. J. Griego, "An elementary theory of stochastic differential equations driven by a poisson process," *Commun. Statist.-Stochastic Models*, 1994, *10*, 335–363.

Hotelling, Harold, "The Economics of Exhaustible Resources," *J. Polit. Economy*, 1931, *39* (2), 137–175.

Kamien, Morton I. and Nancy L. Schwartz, *Dynamic Optimization*, Elsevier Science, 1991.

Karlin, Samuel and Howard M. Taylor, *A First Course in Stochastic Processes*, 2nd ed., Academic Press, 1975.

King, Robert G., Charles I. Plosser, and Sergio T. Rebelo, "Production, growth and business cycles - I. The basic neoclassical model," *J. Monet. Econ.*, 1988, *21*, 195–232.

Kloeden, Peter E. and Eckhard Platen, *Numerical solution of stochastic differential equations*, Springer, 3rd ed., 1999.

Ljungqvist, Lars and Thomas J. Sargent, *Recursive macroeconomic theory*, MIT Press, 2004.

Lucas, Robert E. Jr., "On the Mechanics of Economic Development," *J. Monet. Econ.*, 1988, *22*, 3–42.

Merton, Robert C., "Option Pricing When Underlying Stock Returns Are Discontinuous," *J. Finan. Econ.*, 1976, *3*, 125–144.

——, *Continuous-time finance*, Cambridge, Mass.: Blackwell, 1999.

Michel, Phillippe, "On the Transversality Condition in Infinite Horizon Optimal Problems," *Econometrica*, 1982, *50* (4), 975–985.

Øksendal, B., *Stochastic Differential Equations*, Berlin: Springer, Fifth Edition, 1998.

Pissarides, Christopher A., *Equilibrium Unemployment Theory*, 2nd ed., MIT Press, 2000.

Protter, Philip E., *Stochastic Integration and Differential Equations*, Springer, 2004.

Ramsey, Frank, "A Mathematical Theory of Saving," *Econ. J.*, 1928, *38*, 543–559.

Sennewald, Ken, "Controlled stochastic differential equations under Poisson uncertainty and with unbounded utility," *J. Econ. Dynam. Control*, 2007, *31* (4), 1106–1131.

—— and Klaus Wälde, "Itô's Lemma and the Bellman equation for Poisson processes: An applied view," *Journal of Economics*, 2006, *89* (1), 1–36.

Spanos, A., *Probability Theory and Statistical Inference*, Cambridge Univ. Press, 1999.

Stokey, Nancy L., Robert E. Lucas Jr., and Edward C. Prescott, *Recursive Methods in Economic Dynamics*, Harvard University Press, 1989.

Sydsæter, Knut, Peter Hammond, Atle Seierstad, and Arne Strøm, *Further Mathematics for Economic Analysis*, Prentice Hall, 2005.

Turnovsky, S. J., *Methods of Macroeconomic Dynamics*, Cambridge, Massachusetts: MIT Press, Second Edition, 2000.

Wälde, Klaus, "Endogenous growth cycles," *Int. Econ. Rev.*, 2005, *46* (3), 867–894.

——, *Applied Intertemporal Optimization*, http://www.waelde.com/aio: Lecture Notes, University of Glasgow, 2009.