Chapter 4

Dynamic optimization

In the Solow model presented in Chapter 3, the evolution of the capital stock, along with technological change, is the driving force of output growth. In the model, because agents are assumed to save a constant proportion s of income, saving—and thus investment—are exogenously determined. Cass (1965) and Koopmans (1963) developed the first optimizing models of growth by adding *microeconomic foundations*, i.e., by describing how saving comes about as a result of (rational) choice. These foundations were based on the inter-temporal trade-off between consumption today and consumption in the future, also known as the consumption-saving model. Augmenting the Solow model to incorporate endogenous investment decisions by individuals gave rise to the "optimizing neoclassical growth model," which is nowadays at the core of modern macroeconomic theory. This model delivers the same long-run implications as the original Solow model, but because individual utility is explicit, the model can also be used for analyzing normative issues (individual welfare outcomes). In addition, because the decisions are forward-looking, we can obtain richer policy implications even on the positive side. With the introduction of stochastic elements ("shocks"), it also serves as the foundation of the real business cycle theory, which we will discuss in more detail in Chapter 14.

In this chapter, we introduce a simple dynamic optimization model and discuss the main characteristics of this class of models. A key objective is to describe how to characterize solutions to dynamic optimization problems, and to introduce discrete dynamic programming methods. We work with a representative agent who faces a dynamic trade-off, meaning that sacrifices today yield gains in the future. For example, in the standard consumption-saving model, the agent can save today (reducing current consumption) in exchange for increasing consumption possibilities in the future. In such a model, the saving rate s is endogenously determined. When extending this model to incorporate production, saving and investment are optimally determined. Because saving depends on income in the neoclassical growth model, the level of capital and the productivity level will now affect the saving rate.

A key underlying assumption, first introduced by Milton Friedman in 1957, is that individuals are *forward-looking*. This means that they do not make decisions based just on their current income (as in the textbook Keynesian model) but also considering their expected future income. In other words, consumption and saving decisions depend on expectations. Subsequent research on consumption also incorporated the idea of *rational expectations*, which states that individuals use all the information available to them at each point in time

to make the best possible forecast about the future. While this assumption may seem extreme, it is arguably a good first approximation to the average behavior of individuals for important decisions in their lives. Deviations from rational expectations are studied in behavioral macroeconomics, a subfield of our subject that has important contributions but we regard as second-year material and therefore do not cover in this book. In the present chapter, we assume that agents are fully rational and forward-looking. Because their future income is taken into account when making dynamic decisions, it is important to determine the time horizon of their decision-making process. There are two common approaches followed in the literature: (i) agents live a finite number of periods, or (ii) agents live forever. The latter is interpreted as a *dynastic* structure in which individuals alive today care about the welfare of their descendants, as discussed later in this chapter. Because the infinite-horizon models require more mathematical sophistication, we start with a finite-horizon model. We also discuss two alternative ways of solving dynamic optimization problems: using sequential methods and using recursive methods. Sequential methods involve maximizing over sequences. Recursive methods, also labeled dynamic programming methods, involve functional equations and characterize choices as *decision rules*. We begin with sequential methods and then move to recursive methods.

4.1 A dynamic optimization problem

Economic decisions are made by agents. These could be: (i) individuals (or households) deciding how much of a good or service to consume, (ii) firms choosing how much to produce, or (iii) a government deciding on policy. In this chapter, we focus mainly on the decisions of individuals. The problem of the firm is described in detail in Chapter 5, whereas the government policy decisions are deferred to Chapter 15.

Agents live for T periods, and this time horizon can, in principle, be finite or infinite. With reference to the next chapter, which deals with market economies inhabited by many consumers, we assume that consumers are all identical, so we now study a **representative agent**. This agent makes decisions over sequences of **allocations** in order to maximize a lifetime objective. Allocations are quantities of goods or services, such as consumption, hours worked, investment, etc. The typical dynamic optimization problem in sequential form studied in macroeconomics takes the following form:

$$\max_{\{y_t, x_{t+1}\}_{t=0}^{T}} \sum_{t=0}^{T} \beta^t \hat{\mathcal{F}}(y_t) \tag{P1}$$

subject to

$$x_{t+1} = h(x_t, y_t)$$

and

$$x_{t+1} \in \Gamma(x_t).$$

In this problem, $\sum_{t=0}^{T} \beta^t \hat{\mathcal{F}}(y_t)$ represents the **objective function**. The summation between 0 and T indicates that decisions must be made for each and all of those time periods. The function $\hat{\mathcal{F}}$ is the instantaneous objective, representing, for example, a per-period utility function (for individuals) or profit function (for firms). The constant β is

referred to as the stationary discounting weights: it is our discount factor; $1/\beta - 1$ is the discount rate. They are called stationary because the ratio between the weights of any two different dates $t = i$ and $t = j > i$ only depends on the number of periods elapsed between i and j, and not on the values of i or j. In other words, $\beta^{t+k}/\beta^t = \beta^k$.

The sequence $\{y_t\}_{t=0}^T$ represents the choice variables, sometimes referred to as **control variables**. Examples of these are consumption, leisure, saving, and investment levels at each point in time. The sequence $\{x_t\}_{t=0}^T$ represents the **state variables**. Examples are the stocks of capital, debt, or housing. States and controls are related through the constraint $x_{t+1} = h(x_t, y_t)$. For example, h can reflect a budget constraint or a production technology. The value of the initial state, x_0, is exogenously given. To differentiate states from controls, notice that if an agent chooses y_0 in period 0, the value of the state next period, x_1, is automatically determined from $x_1 = h(x_0, y_0)$ because x_0 is exogenous.[1] In period 1, the choice of the control y_1 determines the state in the following period, x_2, and so on. In other words, choosing the control variable optimally at t determines the value of the state variable at $t+1$. Finally, $\Gamma(x_t)$ represents the feasible set which, given the value of the current state x_t, restricts the values that x_{t+1} can take; we will be specific on standard forms Γ can take later.

For dynamic optimization problems to be well defined (i.e., for solutions to exist), we need to make assumptions about the primitives of the model. Quite generally, we know from basic math—the Weierstrass theorem—that a continuous function attains both a maximum and a minimum when evaluated over a non-empty and compact set. A compact set (of values for a finite vector) means that the set is closed and bounded.[2] Sufficient conditions for this theorem are that (i) $\hat{\mathcal{F}}(y_t)$ is continuous for all y_t, (ii) $h(x_t, y_t)$ is continuous and, in its second argument, strictly monotone for all (x_t, y_t), and (iii) $\Gamma(x_t)$ is non-empty, closed, and bounded for all x_t. Assumption (ii) ensures that we can express y_t as a (continuous) function of (x_t, x_{t+1}), which allows us to write the period objective as a continuous function of this vector. Hence, we have an overall continuous function (of the sequence $\{x_1, x_2, \ldots, x_T, x_{T+1}\}$) to be maximized over a non-empty, compact set.

The two most widely used models in macroeconomics are the consumption-saving model and the neoclassical growth model (NGM). The agent's objective in both models is to maximize lifetime utility choosing the optimal path of consumption. They differ in the production structure and assets available to consumers. In the consumption-saving model, the agent has a time-varying endowment, and can save or borrow at market prices. The NGM, instead, considers the production structure from the Solow model but extends it by endogenizing consumption and investment decisions. In order to fix ideas, it is useful to discuss the main assumptions underlying these models and to map them into our generic formulation.

[1] Often, the problem explicitly says "with x_0 given," to emphasize that it is not a choice variable. However, the fact that it is not a choice variable is already clear since it is not listed among them (under the "max").

[2] By drawing simple graphical examples with non-continuous functions, open sets, and unbounded sets, you can illustrate what can go wrong and why a supremum or infimum may exist but neither a maximum nor a minimum exist.

4.1.1 The consumption-saving model

In the consumption-saving model, there is a representative agent who lives for T periods and must choose the optimal stream of consumption $\{c_t\}_{t=0}^T$, where c_t denotes consumption at time t. We can think of consumption at t as a different good from consumption at $t+1$. Preferences are represented by a utility function $U\left(\{c_t\}_{t=0}^T\right)$. A standard assumption is that this function exhibits "additive separability":

$$U\left(\{c_t\}_{t=0}^T\right) = \sum_{t=0}^T \beta^t u\left(c_t\right).$$

Additive separability implies that the marginal utility of consumption at t does not depend on consumption at other times. Notice that the per-period (or instantaneous) utility index $u(\cdot)$ does not depend on time either. The stationary discounting weights satisfy $0 < \beta < 1$, which is consistent with the observation that individuals seem to deem consumption at an early time more valuable than consumption further off in the future. Formally, if consumption were constant over time, i.e., $c_t = c$ for all t, the marginal utility of c_t decreases in t because $u'(c)$ is multiplied by β^t. Of course, consumption in the future can be more valued on the margin if there is less of it, i.e., if it is sufficiently low relative to consumption today.

We assume that the instantaneous utility function $u(c)$ satisfies the following properties.

1. $u(c)$ is strictly increasing.
2. $u(c)$ is strictly concave.
3. $\lim_{c\to 0} u'(c) = \infty$.

The first property states that individuals have positive marginal utility, $u'(c) > 0$. We also assume that marginal utility of consumption is diminishing, or $u''(c) < 0$. The last property is an Inada condition, stating that agents have infinite marginal utility of consumption as c approaches zero.

We abstract from preferences over leisure and, hence, the determination of the number of hours worked (e.g., labor). These will be studied briefly in the next chapter, and more in depth in Chapter 11. For now, we simply assume that individuals have an endowment of one unit of time, and supply it inelastically to production. They are paid a wage w_t, which may vary over time. In addition, they have access to borrowing and lending. We denote their level of *assets* with a_t, with the understanding that $a_t < 0$ indicates that the agent has *debt*. For example, $a_t = 10$ means that another individual owes him or her 10 units of the consumption good, so the agent is a lender, whereas if $a_t = -10$ then the agent owes someone else that amount, so he or she is a borrower. The initial amount of assets is given by a_0. The net interest rate earned on savings is r_t. The budget constraint of the agent can be written as

$$c_t = w_t + (1 + r_t)a_t - a_{t+1}.$$

Borrowing allows the agent to consume more today (in the amount $-a_{t+1}$), but lowers consumption in the future because he or she would need to pay the principal and interest on

the loan, $(1+r_t)a_t$. A standard assumption is that agents cannot borrow more than a certain amount, which is incorporated as the borrowing limit $a_{t+1} \geq \underline{a}$, with $\underline{a}$ being an exogenous constant. In addition, we impose that $a_{T+1} \geq 0$ in order to eliminate the possibility that the agent ends with a positive level of debt in the last period. Absent this constraint, the agent would choose to borrow up to the limit and never pay back, which clearly could not occur in equilibrium (since no lender would be willing to lend the funds in the last period). We will further discuss the issue of the borrowing limit and the terminal condition later in the chapter.

The dynamic optimization problem of the agent can now be written as

$$\max_{\{c_t, a_{t+1}\}_{t=0}^{T}} \sum_{t=0}^{T} \beta^t u(c_t) \tag{P2}$$

subject to

$$c_t = w_t + (1+r_t)a_t - a_{t+1} \quad \forall t \in \{0, ..., T\},$$
$$c_t \geq 0 \quad \forall t \in \{0, ..., T\},$$
$$a_{t+1} \geq \underline{a} \quad \forall t \in \{0, ..., T\},$$

and

$$a_{T+1} \geq 0.$$

The key sequences to be optimally determined are consumption and asset holdings $\{c_t, a_{t+1}\}_{t=0}^{T}$, to maximize the lifetime utility of the agent. We assume that agents know the sequences of prices (wages and interest rates) when making decisions, and take them as given. We defer the discussion of how these are determined in equilibrium to Chapter 5. The main trade-off faced by the agent, then, is whether to consume today or to save (borrow) a unit of consumption in exchange for $1 + r_{t+1}$ additional (less) units in the future.

This model can be mapped to the generic specification as follows: the per-period objective $\hat{\mathcal{F}}(\cdot)$ corresponds to the instantaneous utility function $u(\cdot)$, the control variables are given by the consumption sequences $\{c_t\}_{t=0}^{T}$, and the states correspond to the level of assets $x_t = a_t$. The equation relating controls to states is the resource constraint, $h(a_t, c_t) = w_t + (1+r_t)a_t - c_t$. The feasible set is given by $\Gamma(a_t) = [\underline{a}, w_t + (1+r_t)a_t]$. The lower bound ensures that $a_{t+1} \geq \underline{a}$, whereas the upper bound ensures that savings do not exceed current income so that $c_t \geq 0$ for all t.

4.1.2 The neoclassical growth model

The utility function in the NGM is the same as in the consumption-saving model: an agent wants to maximize his or her discounted lifetime welfare. The assumptions underlying the instantaneous utility function u are the same as in the previous section: the marginal utility of consumption is positive but diminishing, and the Inada condition must hold.

What changes relative to the consumption-saving model is that this is not an endowment economy, but a production economy instead. Rather than focusing on assets delivering an exogenous return, we consider capital accumulation, where returns are determined by the productivity of capital. The production structure is identical to the one in the Solow model.

Following what we learned in Chapter 3, we assume that technology can be represented by the production function $y_t = f(k)$, where we have used the fact that aggregate labor $L = 1$. From the properties of $F(K, L)$, it is easy to show that f is a strictly increasing and strictly concave function in k. In some applications, we relax the assumption to weak concavity in order to accommodate linear production functions. Capital evolves as in the Solow model, i.e.,

$$k_{t+1} = (1 - \delta)k_t + i_t \tag{4.1}$$

with k_0 given. We abstract from population growth and technological progress to simplify the exposition. The resource constraint is $c_t + i_t \leq y_t$, assuming that this is a closed economy where savings are equal to investment. We can combine it with equation (4.1), substituting away investment, to write the resource constraint in terms of consumption and capital:

$$c_t + k_{t+1} \leq f(k_t) + (1 - \delta)k_t. \tag{4.2}$$

There are alternative ways to specify how the markets for labor, consumption, and capital are organized. For example, we could consider firms who own the capital stock and accumulate it over time. They produce output and sell it to consumers, who work in firms. Alternatively, we could assume that agents own capital, make investment decisions, and rent it to firms every period in exchange for a rental rate. These decentralized production structures will be discussed at length in Chapter 5, where we also explain how prices are determined. In Chapter 6, we show that as long as markets are perfectly competitive, the allocations (i.e., quantities of consumption, capital, and investment) in a competitive equilibrium are Pareto optimal and solve the following *planning problem*.

$$\max_{\{c_t, k_{t+1}\}_{t=0}^{T}} \sum_{t=0}^{T} \beta^t u(c_t) \tag{4.3}$$

subject to

$$c_t + k_{t+1} \leq f(k_t) + (1 - \delta) k_t \quad \forall t \in \{0, ..., T\}$$

and

$$c_t, k_{t+1} \geq 0 \quad \forall t \in \{0, ..., T\}.$$

The planner chooses allocations directly (note that there are no prices in the equations above) to maximize the lifetime welfare of all agents in the economy. Because all agents are identical, this objective is equivalent to maximizing the welfare of the representative agent. The key sequences to be optimally determined are consumption and capital $\{c_t, k_{t+1}\}_{t=0}^{T}$. The main trade-off faced by consumers is whether to consume today or in the future. Any amount of production that is not consumed can be saved, and therefore invested in new capital, yielding future consumption through the additional amount of goods produced, $f'(k_t) + 1 - \delta$. The rationale is analogous to that in the consumption-saving model. Because we assume strictly positive marginal utility of consumption, the resource constraint always holds with equality.

We can map this model to the generic specification as follows: the per-period objective $\hat{\mathcal{F}}(\cdot)$ corresponds to the instantaneous utility function $u(\cdot)$, the control variables are given by the consumption sequences $\{c_t\}_{t=0}^{T}$, and capital represents the states $x_t = k_t$. The equation

relating controls to states is the resource constraint, $h(k_t, c_t) = f(k_t) + (1-\delta) k_t - c_t$. Finally, the feasible set is given by $\Gamma(k_t) = [0, f(k_t) + (1-\delta) k_t]$. The lower bound ensures that $k_{t+1} \geq 0$, whereas the upper bound ensures that $c_t \geq 0$. The next step is to discuss how to solve for the optimal sequences.

4.2 Sequential methods: finite horizon

When the horizon is finite, $T < \infty$, it is possible to use the Kuhn-Tucker Theorem to solve a sequential maximization problem (4.3). The associated Kuhn-Tucker conditions are necessary and sufficient for an optimum if the objective function is strictly concave in the choice vector and the constraint set is closed, bounded, and convex. Let's start with the simplest case, in a two-period economy.

4.2.1 A two-period consumption-saving model

Consider the consumption-saving model when $T = 2$, that the agent is borrowing constrained $\underline{a} = 0$, and that initial assets holdings are zero, $a_0 = 0$. Moreover, let us consider a specific example where, for illustration, the endowment profile is decreasing over time, $w_0 \geq w_1$ and the interest rate is time invariant, $r_t = r$. The budget constraints can be written as

$$c_0 = w_0 - a_1 \quad \text{and} \quad c_1 = w_1 + (1+r)a_1,$$

where we already used the fact that the lower bound on assets holdings binds in the last period, $a_2 = 0$. Intuitively, the representative agent will never choose to save in period 1 if the economy ends in period 1. These constraints can be combined by replacing a_1 from the second period constraint into the first period constraint, and re-organizing terms, as follows

$$c_0 + \frac{c_1}{1+r} = w_0 + \frac{w_1}{1+r}.$$

This equation is known as the lifetime budget constraint. It states that the discounted value of lifetime consumption (left-hand side) must be equal to the lifetime value of income (right-hand side). Future periods are discounted by $1+r$, the relative price of consumption between periods 0 and 1. In addition, here we impose the constraint that $a_1 \geq 0$, which can alternatively be written as $w_0 - c_0 \geq 0$. The Lagrangian can be written as

$$\begin{aligned} \mathcal{L} &= u(c_0) + \beta u(c_1) + \mu \left\{ w_0 + \frac{w_1}{1+r} - c_0 - \frac{c_1}{1+r} \right\} \\ &+ \lambda \{ w_0 - c_0 \}, \end{aligned}$$

where we introduced the Lagrange multiplier μ on the lifetime budget constraint and λ on the non-negativity constraint. We do not consider the non-negativity constraints on consumption here, because the Inada condition $\lim_{c_t \to 0} u'(c_t) = \infty$ ensures that the agent always chooses $c_t > 0$. The first-order conditions with respect to c_0 and c_1 are

$$\frac{\partial \mathcal{L}}{\partial c_0} \quad : \quad u'(c_0) - \mu - \lambda = 0 \tag{4.4}$$

and

$$\frac{\partial \mathcal{L}}{\partial c_1} \quad : \quad \beta u'(c_1) - \mu \frac{1}{1+r} = 0. \tag{4.5}$$

The additional Kuhn-Tucker conditions are

$$\lambda[w_0 - c_0] = 0, \quad w_0 - c_0 \geq 0, \quad \text{and} \quad \lambda \geq 0. \tag{4.6}$$

The first condition in (4.6) is usually referred to as the **complementary slackness condition**. Let us consider an interior solution with $w_0 - c_0 > 0$, implying $\lambda = 0$ from (4.6). It turns out that a decreasing wage path $w_0 > w_1$ and strictly decreasing marginal utility of consumption can guarantee the solution to be interior when $\beta(1+r) \geq 1$. Combining the first-order conditions (4.4) and (4.5) with $\lambda = 0$, we obtain the first-order condition known as the **Euler equation**

$$u'(c_0) = \beta(1+r)u'(c_1).$$

The left-hand side captures the marginal cost of saving an additional unit (which causes a decrease in consumption in the initial period). The right-hand side captures the marginal benefit of saving, which is given by the discounted value of consumption in period $t = 1$ obtained through the returns to savings, $1 + r$. Note that when $\beta(1+r) = 1$, the agent chooses to consume a constant proportion of lifetime earnings every period

$$c_0 = c_1 = \frac{w_0 + \beta w_1}{1+\beta}.$$

Using the budget constraints, this in turn implies that $w_0 - c_0 = a_1 = (w_0 - w_1)/(2+r) > 0$, that is, the solution is indeed interior.[3]

4.2.2 Generic T-period model

When the budget constraint is non-linear in the state variables, it is not possible to construct the lifetime budget constraint as we did in the example above. However, it is still possible to use the Kuhn-Tucker theorem to solve the finite-horizon model, as the next result shows.

Result 1 *Consider a finite-horizon problem (P1), with $\Gamma(x_t) = [\underline{x}, \gamma(x_t)]$. Now use the constraint $x_{t+1} = h(x_t, y_t)$ to solve for y_t as $y_t = \hat{h}(x_t, x_{t+1})$, replace it in the instantaneous objective function, with $\mathcal{F}(x_t, x_{t+1}) \equiv \hat{\mathcal{F}}\left(\hat{h}(x_t, x_{t+1})\right)$, so that we obtain*

$$\max_{\{x_{t+1} \in \Gamma(x_t)\}_{t=0}^{T}} \sum_{t=0}^{T} \beta^t \mathcal{F}(x_t, x_{t+1}).$$

Suppose that $\mathcal{F}(x_t, x_{t+1})$ is increasing in its first argument, decreasing in its second argument, continuously differentiable, and jointly concave in (x_t, x_{t+1}). If

[3] In the general case, we can proceed by first finding an interior *candidate* solution and then verifying whether or not a_1 satisfies its constraint (in this case that a_1 be non-negative); if this works, it is a solution—given the strictly concave utility function and convex constraint set. If a_1 violates its constraint, set it equal to its boundary value and solve for λ, in order to verify that it is non-negative.

(i) $\mathcal{F}_2\left(x_t^*,\ x_{t+1}^*\right) + \beta\mathcal{F}_1\left(x_{t+1}^*,\ x_{t+2}^*\right) = 0,\ \forall t < T$ *with* $\left\{x_{t+1}^*\right\}_{t=0}^{T-1} \in int\Gamma(x_t)$*, where* $\mathcal{F}_i(x_t, x_{t+1})$ *represents the partial derivative with respect to ith argument.*

(ii) $x_{T+1}^* = \underline{x}$*,*

then the sequence $\left\{x_{t+1}^*\right\}_{t=0}^{T}$ *maximizes the objective.*

We omit the proof of this result here. In our section on infinite-horizon optimization below, we state the corresponding result for that setting and, in our appendix, prove it from first principles. That proof can easily be adapted to the present, finite-horizon case.

The result states that it is optimal to choose the lower bound of the feasible set in the last period, $\underline{x}$, and that an interior solution for $t < T$ satisfies the first-order condition (i), which represents the Euler equation for the general model. This equation is sometimes referred to as a "variational" condition (as part of "calculus of variation"): given two boundary conditions x_t and x_{t+2}, we vary the intermediate value x_{t+1} so as to achieve the best outcome. Combining these variational conditions, we notice that there is a total of $T + 2$ equations and $T+2$ unknowns, namely the sequence of states, plus the initial and terminal conditions. This is called a *difference equation* in the sequence of state variables. It is a *second-order* difference equation because there are two lags of x in the equation. Since the number of unknowns is equal to the number of equations, the difference equation system will typically have a solution.

4.2.3 The finite-horizon neoclassical growth model

In the NGM, we have that $\underline{x} = 0$. Solving for consumption from the resource constraint and inserting it into the objective, the problem can be written more compactly as

$$\max_{\{k_{t+1}\}_{t=0}^{T}} \sum_{t=0}^{T} \beta^t u\left(f(k_t) + (1-\delta)\, k_t - k_{t+1}\right) \tag{P3}$$

subject to

$$k_{t+1} \geq 0 \quad \forall t \leq T.$$

Our assumptions on the utility and production functions ensure concavity of the objective function and a closed, bounded, and convex constraint set, so that the assumptions guaranteeing that there is a unique maximizer are fulfilled.[4] As in the two-period consumption-saving model, we are omitting the constraint $c_t \geq 0$, given the Inada condition in the utility function.

The Lagrangian associated to our maximization problem (P3) is

$$\mathcal{L} = \sum_{t=0}^{T} \beta^t \left\{u(f(k_t) + (1-\delta)\, k_t - k_{t+1}) + \mu_t k_{t+1}\right\},$$

[4]To show that the period objective is concave in (k_t, k_{t+1}), it is necessary to go beyond concavity in k_t and k_{t+1} separately: one needs to check that the Hessian is globally negative-definite.

where we introduced the Lagrange/Kuhn-Tucker multipliers $\beta^t \mu_t$ corresponding to the non-negativity constraints on capital for each period t. The first-order conditions are

$$\frac{\partial \mathcal{L}}{\partial k_{t+1}} \quad : \quad -u'(c_t) + \beta u'(c_{t+1})[f'(k_{t+1}) + 1 - \delta] + \mu_t = 0, \quad t \in \{0, ..., T-1\} \tag{4.7}$$

and

$$\frac{\partial \mathcal{L}}{\partial k_{T+1}} \quad : \quad -u'(c_T) + \mu_T = 0. \tag{4.8}$$

The first-order condition in the final period is different from the ones in earlier periods because the economy ends at that time. Finally, the Kuhn-Tucker conditions also include

$$\mu_t k_{t+1} = 0, \ t \in \{0, ..., T\}, \tag{4.9}$$

$$k_{t+1} \geq 0, \ t \in \{0, ..., T\},$$

and

$$\mu_t \geq 0, \ t \in \{0, ..., T\}.$$

Equation (4.9) parallels the first condition of (4.6) in the two-period consumption-savings problem (complementary slackness condition). Because we assume that $u(c_t)$ is strictly increasing, equation (4.8) implies that $\mu_T > 0$. From (4.9) evaluated at $t = T$, we find that $k_{T+1} = 0$. This result establishes the terminal condition for our maximization problem: consumers leave no capital for production after the final period, since they receive no utility from that capital and would rather use it for consumption during their lifetime. The insight is trivial here, but will be relevant also when we consider an infinite-horizon economy.

Given the Inada conditions, in particular $\lim_{k \to 0} f'(k) = \infty$ and $\lim_{c \to 0} u'(c) = \infty$, it is optimal to set $k_{t+1} > 0$ for all $t < T$. Hence, the non-negativity constraint on capital will never be binding, implying that $\mu_t = 0$ for $t < T$. Replacing this result in equation (4.7) delivers the Euler equation,

$$u'\left[f(k_t) + (1-\delta)k_t - k_{t+1}\right] = \beta u'\left[f(k_{t+1}) + (1-\delta)k_{t+1} - k_{t+2}\right]\left[f'(k_{t+1}) + 1 - \delta\right], \tag{4.10}$$

which holds for all $t \in \{0, ..., T-1\}$. This equation, together with the initial condition k_0 and the terminal condition derived before, $k_{T+1} = 0$, determines the capital sequence $\{k_{t+1}\}_{t=0}^{T}$ (e.g., $T+2$ equations and $T+2$ unknowns). Because the first-order conditions are sufficient in the example, there is a unique solution to the difference equation (4.10) describing the evolution of capital over time.

To interpret the key equation for optimization, it is useful to break the Euler equation down into three components:

$$\underbrace{u'(c_t)}_{\substack{\text{marginal cost} \\ \text{of investment}}} = \underbrace{\beta u'(c_{t+1})}_{\substack{\text{utility increase} \\ \text{per unit return}}} \cdot \underbrace{\left[f'(k_{t+1}) + 1 - \delta\right]}_{\text{return on investment}}.$$

The left-hand side represents the marginal cost of investing one unit of consumption today, generating a disutility loss of $u'(c_t)$. The right-hand side represents the marginal benefit of this investment. Higher investment increases capital next period, which produces additional output (plus un-depreciated capital) $f'(k_{t+1}) + 1 - \delta$. This return increases consumption next period, with a discounted per-unit utility gain of $\beta u'(c_{t+1})$.

4.2.4 Solving a finite-horizon model

We have seen how to derive the set of equations that will determine the solution to the maximization problem. How does one solve these equations, however? Several methods are available. The first one uses "backward induction": starting at the final period, T, we can use the resource constraint and the Euler equation iteratively moving backwards. This method is illustrated in the following example, which is designed to yield closed-form solutions at each stage (it is the only known example with strictly concave utility and production functions that can be solved analytically).

Example 4.1 *Consider a T-period economy where utility is logarithmic, $u(c) = \log c$, the production function is Cobb-Douglas $f(k) = Ak^{\alpha}$, and there is full depreciation $\delta = 1$. The Euler equation (4.10) for $t < T$ becomes*

$$\frac{1}{Ak_t^{\alpha} - k_{t+1}} = \beta \frac{1}{Ak_{t+1}^{\alpha} - k_{t+2}} \alpha A k_{t+1}^{\alpha-1}. \tag{4.11}$$

The last term is the marginal product of capital $f'(k) = \alpha A k^{\alpha-1}$. Evaluating equation (4.11) at period $t = T - 1$, replacing the terminal condition $k_{T+1} = 0$ in its right-hand side, and simplifying delivers

$$k_T = \frac{\alpha\beta}{1 + \alpha\beta} A k_{T-1}^{\alpha}.$$

We can now use this result, together with equation (4.11) evaluated at period $T-2$, to obtain

$$k_{T-1} = \frac{\alpha\beta(1 + \alpha\beta)}{1 + \alpha\beta + (\alpha\beta)^2} A k_{T-2}^{\alpha}.$$

Going backwards in time, it is possible to see that

$$k_{T-t} = \frac{\alpha\beta(1 + \alpha\beta + ... + (\alpha\beta)^t)}{1 + \alpha\beta + ... + (\alpha\beta)^{t+1}} A k_{T-t-1}^{\alpha}.$$

Using the fact that $\alpha\beta < 1$, we can use the properties of geometric series in the numerator and the denominator to simplify this expression. That, together with a change of variables, allows us to obtain a formula that describes the evolution of capital in closed form:

$$k_{t+1} = \alpha\beta \frac{1 - (\alpha\beta)^{T-t}}{1 - (\alpha\beta)^{T-t+1}} A k_t^{\alpha}.$$

Consumption becomes

$$c_t = \frac{1 - \alpha\beta}{1 - (\alpha\beta)^{T-t+1}} A k_t^{\alpha}.$$

There are a few characteristics of the solution that we would like to highlight. First, we obtained an expression that, given the initial condition k_0, fully describes the evolution of capital, output, and consumption for the whole time horizon: the outcomes depend explicitly on parameters and on k_0. Second, we see that it is optimal to save (and invest) a proportion

$$s_t = \alpha\beta \frac{1 - (\alpha\beta)^{T-t}}{1 - (\alpha\beta)^{T-t+1}}$$

of output every period. In contrast to the assumption of the Solow model, the optimal saving rate depends on time (the time left to the final date T). But, like in the Solow model, the saving rate does not depend on the level of the capital stock. Two key parameters determine the saving rate: the discount factor and the degree of concavity of the production function. Third, although the utility function is strictly concave, consumption is not fully smoothed: it will, in general, vary over time. The reason for the lack of full smoothing is that the level of the capital stock influences the marginal return on saving when the production function is neoclassical: the higher the capital stock, the lower is this return. The Euler equation tells us that the higher the marginal return on saving, the higher the consumption growth should be, implying that for capital stocks below (above) steady state, consumption rises (falls) over time.[5]

It is possible to obtain analytical solutions in a few other cases; one is where the production function is linear and preferences are represented by a u as given in (2.5); this function is often called the Constant Relative Risk Aversion (CRRA) utility function.

Constant Relative Risk Aversion (CRRA) utility function

The CRRA function is one of the most commonly used additively separable utility functions in macroeconomics[a]

$$u(c) = \frac{c^{1-\sigma} - 1}{1-\sigma} \quad \text{where } \sigma \geq 0 \text{ and } \sigma \neq 1$$

It has, as special cases,

$$\begin{array}{ll} \sigma = 0 & \text{linear utility,} \\ \sigma > 0 & \text{strictly concave utility,} \\ \sigma \to 1 & \text{logarithmic utility.} \end{array}$$

The limit case where $\sigma \to \infty$ utility converges to zero for all c, i.e., the agent is indifferent across paths. However, the relevant limit should be seen as that obtained by raising $\sum_{t=0}^{\infty} \beta^t c_t^{1-\sigma}$ to a power $1/(1-\sigma)$, in which case the limit becomes a "Leontief function": $\min_t \{c_t\}_{t=0}^{\infty}$.[b]

[a] We do not consider uncertainty in this chapter, so the concept of the "relative risk aversion" is not directly relevant. Here this term serves only as a label for a class of utility functions.

[b] The stated utility function here is a monotone transformation of the original function and we know that the behavior given by a utility function is preserved under monotone transformations.

We define the elasticity of intertemporal substitution (EIS) as the percentage change in consumption between periods t and $t+s$ in response to a percentage change in the returns to investment between the same two periods, assuming optimizing behavior subject to a standard budget constraint, taking other returns and the agent's income as given.[6] Thus,

[5] The special case $\alpha = 1$ means that the marginal return on saving is always A. Then, unless $\beta A = 1$, the economy will grow or shrink over time at a constant rate, as will consumption. Consumption, along with capital, will only be constant if $\beta A = 1$.

[6] That is, we use the notion of an *uncompensated* elasticity. Also, our notion presumes the absence of

we have

$$EIS \equiv \frac{\frac{\partial(c_{t+s}/c_t)}{c_{t+s}/c_t}}{\frac{\partial R_{t,t+s}}{R_{t,t+s}}}.$$

In the next example, we show that the CRRA function has a constant intertemporal elasticity of substitution, which is equal to $1/\sigma$. For this reason, a CRRA utility function is also sometimes referred to as a Constant Elasticity of Intertemporal Substitution (CEIS) utility function.

Example 4.2 *Consider a T-period economy where utility is CRRA, $u(c) = \dfrac{c^{1-\sigma}-1}{1-\sigma}$ and the production function is linear $f_t(k_t) = R_t k_t$. In this case, the Euler equation reads*

$$u'(c_t) = \beta u'(c_{t+1}) R_{t+1}.$$

Repeated substitution delivers

$$\begin{aligned} u'(c_t) &= \beta^s u'(c_{t+s}) \underbrace{R_{t+1}R_{t+2}...R_{t+s}}_{\equiv R_{t,t+s}} \\ u'(c) &= c^{-\sigma} \Rightarrow c_t^{-\sigma} = \beta^s c_{t+s}^{-\sigma} R_{t,t+s} \\ \frac{c_{t+s}}{c_t} &= (\beta^s)^{\frac{1}{\sigma}} (R_{t,t+s})^{\frac{1}{\sigma}}. \end{aligned}$$

This means that the EIS becomes

$$\frac{\frac{\partial(c_{t+s}/c_t)}{c_{t+s}/c_t}}{\frac{\partial R_{t,t+s}}{R_{t,t+s}}} = \frac{\partial \log(c_{t+s}/c_t)}{\partial \log R_{t,t+s}} = \frac{1}{\sigma}.$$

When $\sigma \to 1$, the relative expenditure shares $c_t/(c_{t+s}/R_{t,t+s})$ do not change: this corresponds to the logarithmic case. When $\sigma > 1$, an increase in $R_{t,t+s}$ would lead c_t to increase and investment to decrease: the income effect, leading to smoothing across all goods, is larger than the substitution effect. Finally, when $\sigma < 1$, the substitution effect is stronger than the income effect: investment rises whenever $R_{t,t+s}$ increases. When $\sigma = 0$, the elasticity is infinite and investment responds discontinuously to $R_{t,t+s}$.

Another, and absolutely central, reason that the CRRA utility function plays an important role for us is that it is the only one that is consistent with balanced growth, as described in Section 4.3.3.

4.3 Sequential methods: infinite horizon

In this section, we extend our model to infinite periods. The main advantage of an infinite horizon is that the household problem becomes stationary: the maximization problem at

other goods in the utility function, in which case one needs to specify how the choices of the other goods are made as the returns are changed.

date t is exactly the same as in period $t+1$ (for a given starting level of capital). This property is in contrast to that in the previous section, where decisions were significantly affected by how many periods the individual had left (see Example 4.1). A large number of macroeconomic applications, particularly those studying the long-run evolution of aggregate economic variables, use infinite-lived agents as their main building block.

For the typical models that macroeconomists use, the infinite-horizon version behaves very similarly to the finite-horizon version when the latter's remaining time horizon is long enough.[7] However, let us also discuss whether an infinite time horizon is a sensible assumption: after all, people do not live forever. However, to the extent that individuals are altruistic, they care about their descendants. Let $u(c_t)$ denote the utility flow to generation t. We can then interpret β^t as the weight individuals attach to the utility enjoyed by their descendants, t generations down the family tree. Their total welfare is given by $\sum_{t=0}^{\infty} \beta^t u(c_t)$. As long as $\beta < 1$, agents care more about themselves than about their offspring.[8]

4.3.1 Mathematical considerations

Because agents are now choosing infinite sequences of consumption and investment, models with an infinite time horizon demand more advanced mathematical tools.

A basic question is whether the solution to the planner's problem exists once choices are no longer elements of the Euclidean space $\Re^T$. In more general notation, suppose we are seeking to maximize a function $U(x)$, where $x \in S$ and S is a set that includes infinite sequences. If U is continuous, we can invoke the Weierstrass theorem, provided that the set S is nonempty and compact. For finite sequences, continuity and compactness are defined in standard ways. But for infinite-dimensional sequences, several issues arise. How do we define continuity in this setup? What is an open set? What does compactness mean? Answering these questions in detail is beyond the scope of this book; we refer you, for example, to Stokey and Lucas (1989). For illustration, we will, however, provide some specific examples where the maximization problem may be ill-defined (i.e., have no solution) unless a set of necessary conditions holds.

Unbounded utility Continuity of the objective requires boundedness. A necessary condition for the lifetime utility U to be bounded is that consumption streams do not yield "infinite" utility. If two consumption streams do so, they cannot be compared and the maximization problem is ill-defined. For example, consider a plan specifying equal amounts of consumption each period, $\{c_t\}_{t=0}^{\infty} = \{\bar{c}\}_{t=0}^{\infty}$, delivering $U = \sum_{t=0}^{\infty} \beta^t u(\bar{c})$. Clearly, this function is unbounded (e.g., does not have a finite limit) unless $\beta < 1$.

This may not be sufficient, however. Suppose that the constraints allow for a constantly increasing consumption stream $\{c_t\}_{t=0}^{\infty} = \left\{c_0 (1+\gamma)^t\right\}_{t=0}^{\infty}$. The lifetime utility is now $U = \sum_{t=0}^{\infty} \beta^t u\left(c_0 (1+\gamma)^t\right)$. Even if $\beta < 1$ (so β^t is decreasing to 0), the argument inside the

[7] Game theory is a sharp contrast here: we know that an infinite horizon can then open up to the existence of many equilibria, e.g., the trigger-strategy outcomes in repeated games.

[8] In this simple example we attach only one consumption level to each generation. The example could be extended to the case where people of generation t consume in multiple periods; though more complicated to formulate, this extension would be straightforward and the main insights would carry over.

utility function is growing at rate γ. The shape of the utility function, hence, is key to determining whether the maximization problem is well defined. In the case of a CRRA utility function $u(c) = (c^{1-\sigma} - 1)/(1-\sigma)$, we obtain a level of utility that is a geometric sum. Hence, boundedness requires $\beta(1+\gamma)^{1-\sigma} < 1$. If $\sigma < 1$, so that there is less than logarithmic curvature, this requirement involves an upper bound on γ for utility to be a positive, finite number. If $\sigma > 1$, positive growth ($\gamma > 0$) implies that the boundedness condition is met; with negative growth at a sufficiently high rate, it will not be.[9]

Constraint sets that are "too large" The problem can be ill-defined if the constraint sets are "too large". In the consumption-saving model discussed at the outset of the chapter, we imposed the constraint that borrowing could not exceed the exogenous amount $\underline{a}$. This condition ensures that the constraint set is bounded, and hence that the problem is well defined.

In some applications, however, imposing the constraint $a_{t+1} \geq \underline{a}$ for each t can be "too restrictive"; it will rule out feasible borrowing. Instead, for example, we could only impose that $a_{T+1} \geq 0$, i.e., that the agent cannot borrow in the very last period. This terminal constraint is important: if this constraint was not imposed, any individual would have incentives to accumulate an infinite amount of debt going into the final period. Such a restriction also makes sense: knowing that the consumer would engage in such a scheme, no lender would be willing to lend at any positive rate, as the loan would be defaulted with probability 1.

With an infinite horizon, a final-period constraint loses meaning in a literal sense. Instead, we impose a constraint that is its appropriate infinite-period extension, known as the **no Ponzi game** (nPg) condition. In words, this condition rules out "borrowing at infinity, measured in present value." This requirement represents a restriction on the agent's constraint set, preventing it from being so large as to allow the agent to attain arbitrarily high utility. To see how the condition comes about, let us look at a simple example.

Suppose we endow a consumer with a given initial amount of net assets, a_0, representing claims against other agents. Additionally, suppose that the agent has no other sources of income, so the budget constraint is

$$c_t + a_{t+1} = (1+r)a_t, \forall t \geq 0,$$

where we assume that $r > 0$.[10] Consider a candidate solution to consumer's maximization problem $\{c_t^*\}_{t=0}^{\infty}$. Absent further constraints, the agent could improve on $\{c_t^*\}_{t=0}^{\infty}$ as follows:

1. Let $\tilde{c}_0 = c_0^* + \epsilon$, with $\epsilon > 0$, making $\tilde{a}_1 = a_1^* - \epsilon$.

2. For every $t \geq 1$ leave $\tilde{c}_t = c_t^*$ by setting $\tilde{a}_{t+1} = a_{t+1}^* - \epsilon(1+r)^t$.

Given a strictly increasing utility function, the agent is clearly better off under this alternative consumption allocation, which satisfies the budget constraint period by period. Because this

[9]Negative growth, i.e., $\gamma < 0$, does not occur in the context of our standard balanced-growth model, but it can hypothetically occur if natural resources are assumed to be finite and essential for production. In this case, utility could hence become unboundedly negative—recall that u is negative in this case—and there may even be no allocation of resources with finite utility. For more, see Chapter 25.

[10]Cases with $r \leq 0$ can sometimes be relevant but we do not consider them here.

sort of improvement is possible for *any* candidate solution, there cannot be a maximum for lifetime utility.[11] Note that with this alternative allocation, the agent's debt is growing without bound at rate $1+r$, and it is never repaid. This type of scheme, borrowing $(1+r)a_t$ every period t to keep rolling the debt, is often called a "Ponzi scheme" or a "Ponzi game". It is crucial in the infinite-horizon model to impose a constraint that rules out the Ponzi scheme. A condition that works is the nPg condition in the following form:

$$\lim_{t\to\infty} \frac{a_{t+1}}{(1+r)^t} \geq 0. \tag{4.12}$$

Intuitively, the agent cannot engage in borrowing and lending so that their "terminal asset holdings" (in present-value terms) are negative, because this means that they would borrow and not pay back.

We can use the nPg condition to simplify, or *consolidate*, the sequence of budget constraints. To do this, solve for a_1 from the budget constraint in period 1 and substitute the expression into the budget constraint in period 0: we obtain a budget containing c_0, c_1, a_0, and a_2. Next, replace a_2 in this expression by solving for it in the next budget constraint, and proceed this way forward. After T substitutions, we obtain

$$\sum_{t=0}^{T} c_t \frac{1}{(1+r)^t} = a_0(1+r) - \frac{a_{T+1}}{(1+r)^T}.$$

Taking limits, we arrive at

$$\sum_{t=0}^{\infty} c_t \frac{1}{(1+r)^t} = a_0(1+r) - \lim_{T\to\infty} \frac{a_{T+1}}{(1+r)^T} \leq a_0(1+r), \tag{4.13}$$

where the inequality comes from the nPg condition. This is the lifetime budget constraint in the infinite-horizon model.

We motivated ruling out Ponzi schemes as a natural infinite-horizon extension of the constraint $a_{T+1} \geq 0$ in a finite-horizon model. There is also a constraint often labeled the "natural borrowing limit." It can be used both in the finite-horizon and the infinite-horizon cases and it captures, precisely, the notion of a loosest possible constraint on borrowing: the only restriction is that you are able to pay back, in a present-value sense, if you set consumption to zero at all future times. In the particular model here, the natural borrowing limit is zero at all times, but in general the constraint depends on the future stream of non-asset income. This case is analyzed in Appendix 4.A.1, which shows that imposing the natural borrowing limit, imposing the lifetime budget constraint, and imposing the nPg condition are all equivalent.

Note that the lifetime budget constraint is often written with equality, i.e.,

$$\sum_{t=0}^{\infty} c_t \frac{1}{(1+r)^t} = a_0(1+r), \tag{4.14}$$

[11] One could imagine a maximum if an arbitrary upper bound is placed on consumption at each date, but such ad-hoc assumptions are undesirable.

because, as we shall see below, $\lim_{T\to\infty} a_{T+1}/(1+r)^T$ will never be chosen to be positive, so long as the utility function is strictly increasing. The reason is that assets do not themselves contribute to utility, so it is always better to increase consumption as long as it is positive. The inequality in the opposite direction of (4.12),

$$\lim_{T\to\infty} \frac{a_{T+1}}{(1+r)^T} \leq 0, \tag{4.15}$$

therefore, is a part of the optimality condition. The condition (4.15) will be referred to as the **transversality condition** (TVC). The equality (4.14) is the combination of the inequalities (4.13) (which comes from the nPg condition (4.12)) and the TVC (4.15). The nPg and the TVC are different in nature: the former is a restriction on what paths the consumer is allowed to choose whereas the latter is a self-imposed condition—it is chosen by the consumer.

Let us now work out a full solution to the problem above.

Example 4.3 *Consider the infinite-horizon version of the consumption-saving model (without borrowing constraints) given a logarithmic utility function* $u(c) = \log c$. *The optimization problem is:*

$$\max_{\{c_t, a_{t+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t \log c_t$$

subject to

$$c_t + a_{t+1} = a_t (1+r) \quad \forall t \geq 0$$

and the nPg condition.

To solve this problem, replace the period budget constraints with a consolidated one evaluated at equality (hence, already using the TVC, i.e., the fact that it is not optimal to save a positive amount, in present-value terms, at infinity):

$$\sum_{t=0}^{\infty} c_t \left(\frac{1}{1+r}\right)^t = a_0 (1+r).$$

With this simplification, the first-order conditions are

$$\beta^t \frac{1}{c_t} = \lambda \left(\frac{1}{1+r}\right)^t, \forall t \geq 0,$$

where λ *is the Lagrange multiplier associated with the consolidated budget constraint. From the first-order conditions it follows that*

$$c_t = [\beta (1+r)]^t c_0, \forall t \geq 1.$$

Substituting this expression into the consolidated budget constraint, we obtain

$$\begin{aligned} \sum_{t=0}^{\infty} \beta^t (1+r)^t \frac{1}{(1+r)^t} c_0 &= a_0 (1+r) \\ c_0 \sum_{t=0}^{\infty} \beta^t &= a_0 (1+r). \end{aligned}$$

From here, $c_0 = a_0 (1-\beta)(1+r)$, *and consumption in the periods* $t \geq 1$ *can be recovered from* $c_t = [\beta (1+r)]^t c_0$.

Sufficient conditions: the transversality condition (TVC) In general, the infinite-horizon maximization problems involve the same mathematical techniques as the finite-horizon ones. In particular, we make use of (Kuhn-Tucker) first-order conditions. In the neoclassical growth model, these lead to a second-order difference equation, the Euler equation, defining a path for the state variable given the initial condition k_0. But unlike in the finite horizon case, where it was optimal to set $k_{T+1} = 0$, there is no final condition that allows us to pin down the sequence of capital. Therefore, the difference equation that characterizes the first-order condition may have an infinite number of solutions. To determine the solution, we need to make use of an additional optimality condition that we already briefly discussed: the *transversality condition*. This condition, which we will now discuss in more general terms, captures the principle that it cannot be optimal for an agent to choose a sequence of capital involving, in present-value utility terms, a positive shadow value as $t \to \infty$. In the consumption-saving problem such behavior is clearly sub-optimal: lower savings would be feasible and yield higher lifetime utility.

We will not prove the necessity of the TVC here. We will, however, provide a sufficiency condition for a generic optimization problem. We will, moreover, offer a proof strategy that also allows us to derive what form the TVC must take (it is not always obvious what its precise form should be). The message of the following proposition is that if we have a convex maximization problem (utility is concave and the constraint set convex) and a sequence $\{x_{t+1}\}_{t=0}^{\infty}$ that satisfies the Kuhn-Tucker first-order conditions and the transversality condition, then indeed we have a maximum. Formally, we have the following.

Proposition 4.4 *Consider the infinite-horizon version of the maximization problem* (P1)*, where we use $x_{t+1} = h(x_t, y_t)$ to replace y_t in the objective:*

$$\max_{\{x_{t+1}\in\Gamma(x_t)\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t \mathcal{F}(x_t, x_{t+1}).$$

If $x_{t+1}^* \in int\ \Gamma(x_t)$ for all t,

(i) Euler equation: $\mathcal{F}_2\left(x_t^*,\ x_{t+1}^*\right) + \beta\mathcal{F}_1\left(x_{t+1}^*,\ x_{t+2}^*\right) = 0\ \forall t$

(ii) TVC: $\lim_{t\to\infty} \beta^t \mathcal{F}_1\left(x_t^*,\ x_{t+1}^*\right) x_t^* = 0$,

$\mathcal{F}\left(x_t, x_{t+1}\right)$ is jointly concave in (x_t, x_{t+1}) and increasing in its first argument, and $\Gamma(x)$ is a convex set for all x, then $\left\{x_{t+1}^*\right\}_{t=0}^{\infty}$ maximizes the objective.

Proof. See Appendix 4.A.3 ■

The proof is relatively mechanical: it repeatedly uses the Euler equation, (i), and a definition of concavity, namely, that a concave function is always globally below its tangent hyperplane. It shows that the stated sequence gives a higher value to the objective than any other feasible sequence so long as a remaining inequality is met. That remaining inequality is met when the TVC is satisfied: condition (ii) of the theorem.

In the neoclassical growth model, the partial derivative $\mathcal{F}_1\left(x_t, x_{t+1}\right)$ becomes $u'(c_t)[f'(k_t)+1-\delta]$, which corresponds to the marginal utility of increasing capital in period t. The TVC reads

$$\lim_{t\to\infty} \beta^t u'(c_t)[f'(k_t) + 1 - \delta]k_t = 0.$$

It states that the discounted present-value of each additional unit of capital times the stock of capital has to be zero in the limit. If this requirement is not met, it would be beneficial to modify the path of capital in order to increase consumption, generating higher lifetime utility, without violating feasibility.

In cases where a proposed chosen sequence becomes stationary in the limit—such as in the neoclassical growth model without growth—the TVC is satisfied "automatically" so long as $\beta < 1$. When consumption and capital grow, i.e., due to technical change, the condition is less trivial. Under the form of utility that allows balanced growth, $u(c) = (c^{1-\sigma} - 1)/(1-\sigma)$, $\beta^t u'(c_t) k_t$ will grow at the gross rate $\beta\gamma^{1-\sigma}$, since both c_t and k_t grow at gross rate γ on a balanced path. The transversality condition thus requires $\beta\gamma^{1-\sigma} < 1$, which we recall is also the condition that makes utility bounded.

It is yet again worth emphasizing that the transversality condition and the no-Ponzi game conditions are conceptually distinct. The TVC (jointly with the Euler equation being met at all times) is a sufficient condition for optimization. It states that the value of assets cannot be positive in the limit, because the agent could otherwise be made better off by reducing them. Hence, it is eliminating sequences that cannot be optimal by ruling out over-accumulation of wealth.[12] The nPg condition, on the other hand, is an institutional constraint ensuring that the agent cannot have positive debt (in present value terms) in the limit. An agent would want to choose to increase debt unboundedly if he or she were allowed to violate it. The latter could not reasonably be thought to occur in a market economy and therefore the constraint is imposed. In sum, the TVC is a self-imposed constraint ruling out over-accumulation of assets, whereas the nPg condition is an externally imposed constraint ruling out of extreme under-accumulation (exploding debt).

4.3.2 Solving the infinite-horizon neoclassical growth model

The infinite-horizon maximization problem can be solved taking the limit of the solution we found in the finite horizon case, and making sure that the transversality condition is met. A question, to which we will return later, is whether this model, like the Solow model, will imply global convergence to a steady state. A preliminary inquiry involves finding the set of steady states. This is straightforward: the Euler equation (4.10) implies that for any constant positive level of consumption—which means that $u'(c_t) = u'(c_{t+1})$ can be eliminated from the equation—we obtain

$$1 = \beta(f'(\bar{k}) + 1 - \delta). \tag{4.16}$$

Given that f is strictly concave and satisfies Inada conditions, this implies a unique solution.

We will return to the general formulation later, once we have covered recursive methods, and then establish global convergence to $\bar{k}$ from any positive starting level of capital. For now, we will look at an example where we can solve for dynamics explicitly.

Example 4.5 *Consider an infinite-horizon economy where utility is logarithmic, $u(c) = \log c$, the production function is Cobb-Douglas $f(k) = Ak^\alpha$, and there is full depreciation: $\delta = 1$. Truncating the economy at $t = T$ delivers the same optimality conditions and analytical*

[12] The nature of this argument is that TVC is a necessary condition for optimization, which it is under some conditions; we do not prove this here.

expression for capital next period that we found in Example 4.1. Denote the solution to the truncated problem with k_{t+1}^T. Taking limits, we can find a candidate solution to the infinite horizon problem k_{t+1} as

$$\begin{aligned} \lim_{T\to\infty} k_{t+1}^T &= \lim_{T\to\infty} \alpha\beta \frac{1-(\alpha\beta)^{T-t}}{1-(\alpha\beta)^{T-t+1}} A k_t^\alpha \\ k_{t+1} &= \alpha\beta A k_t^\alpha \end{aligned} \tag{4.17}$$

To check that the TVC holds, as pointed out above, we only need to look at a limit and not at other aspects of the sequence for capital. So, (i), (4.17) implies that the capital sequence converges to a limit (as will, then, consumption) and, (ii) therefore the limit behavior of the expression $\beta^t \mathcal{F}_1(k_t,\ k_{t+1})\, k_t$ will boil down to a constant times β^t. Thus, the TVC will be met. To see the concrete expressions, first note that $\mathcal{F}_1(k_t,\ k_{t+1}) = u'(c)f'(k_t)$, since consumption is simply $c = f(k_t) - k_{t+1}$. The TVC can be written as

$$\lim_{t\to\infty} \beta^t u'(c_t) f'(k_t) k_t = \lim_{t\to\infty} \beta^t \frac{A\alpha k_t^{\alpha-1}}{A k_t^\alpha - k_{t+1}} k_t.$$

Using equation (4.17) and simplifying delivers $\lim_{t\to\infty} \beta^t \alpha/(1-\alpha\beta) = 0$, since $\beta < 1$.

Another common method for solving infinite-horizon problems is known as "guess and verify." The strategy consists of guessing a generic functional form for k_{t+1} as a function of k_t and using the Euler equation to verify that the guess is correct. We illustrate this method in the following example.

Example 4.6 *Consider the economy from Example 4.5. Guess that $k_{t+1} = sAk_t^\alpha$, where s is an unknown parameter (the saving rate). Insert this guess into the Euler equation (4.10) to obtain*

$$\frac{1}{Ak_t^\alpha - sAk_t^\alpha} = \beta\alpha A \frac{k_{t+1}^{\alpha-1}}{Ak_{t+1}^\alpha - sAk_{t+1}^\alpha}.$$

After some manipulations, we can verify that $s = \alpha\beta$ satisfies the equation above no matter what k_t is.

A natural challenge with this approach is of course how to come up with an insightful initial guess. A standard procedure is to use a combination of the truncation-at-T method and the guess and verify method. In other words, it is possible to solve the problem backwards a couple of periods to get a sense of the possible functional form, and then verify whether this guess is correct using the Euler equation.

Note that in the infinite-horizon economy of Example 4.5, the stock of capital evolves as assumed by the Solow model: agents invest a *constant* proportion, $s = \alpha\beta$, of their income every period. Moreover, using what we have learned in Chapter 2, we can show that this economy converges to a steady state $\bar{k}$. To compute it, simply evaluate our solution at the steady state, $k_t = k_{t+1} = \bar{k}$,

$$\bar{k} = \beta\alpha A \bar{k}^\alpha \Rightarrow \bar{k} = (\beta\alpha A)^{\frac{1}{1-\alpha}},$$

which is identical to what we found in Chapter 3, given $\delta = 1$.

In general, when closed-form solutions are not possible to solve for, we need to resort to numerical methods. It is possible to solve for sequences numerically, but seeking numerical solutions using dynamic programming is also possible. Dynamic programming also delivers conceptual insights. The following section turns to this method.

In the case of a relatively simple model like the NGM, yet another method of characterizing the dynamics is to use *phase diagrams*. The solution to a general NGM can be written as a set of two difference equations for two variables (k_t, c_t):

$$k_{t+1} - k_t = f(k_t) - \delta k_t - c_t$$

and

$$u'(c_t) = \beta u'(c_{t+1})(f'(k_{t+1}) + 1 - \delta).$$

The first is the resource constraint for the social planner, and the second is the Euler equation. We can look for the dynamic path of (k_t, c_t) that satisfies these two difference equations, together with the nonnegativity constraints for both variables and the TVC. The graphical analysis using the phase diagram is explained in Appendix 4.A.4.

4.3.3 Balanced growth in the neoclassical growth model

The economies described in the previous sections eventually converge to a steady state, exhibiting no growth in the long run. It is possible to study a version of the NGM in which there is balanced growth, with a production structure similar to the one studied in Chapter 3.2. There, we considered $Y_t = F(K_t, A_t L_t)$, where $A_t L_t$ denotes "efficiency units of labor." An important assumption that allows us to obtain balanced growth is that L_t and A_t grow at the (constant) rates n and γ, respectively. Because of population growth, we now need to take a stand on the welfare criterion by giving weight to the current population and people born in the future. In what follows, we assume that the social planner is "utilitarian" and maximize the sum of utility in the entire economy:

$$\sum_{t=0}^{\infty} \beta^t L_t u\left(c_t\right),$$

where $c_t = C_t/L_t$ is the per capita consumption. An alternative formulation is to assume the social planner cares about the "per capita" utility in each period: $\sum_{t=0}^{\infty} \beta^t u\left(c_t\right)$. As will become clear with the procedure below, this different assumption amounts to a different discount factor by the social planner. In our benchmark formulation, with positive population growth, we give more weight to future utility than in the alternative formulation. Quantitatively, the population growth rates tend to be small in advanced economies, and the impact of the alternative assumptions on the outcome is unlikely to be significant.

The resource constraint with growth becomes

$$C_t = F(K_t, A_t L_t) + (1 - \delta)K_t - K_{t+1}.$$

Defining variables in "per efficiency units of labor", $\tilde{x}_t = X_t/(A_t L_t)$, we can follow similar steps as those in Section 3.2 to write the resource constraint as

$$\tilde{c}_t = f(\tilde{k}_t) + (1 - \delta)\tilde{k}_t - (1 + \gamma)(1 + n)\tilde{k}_{t+1}. \tag{4.18}$$

We further assume that the instantaneous utility is CRRA, $u(c_t) = c_t^{1-\sigma}/(1-\sigma)$, $\sigma > 0$ and $\sigma \neq 1$. Appendix 4.A.2 shows that this function (and the logarithmic function, which should be seen as the case $\sigma \to 1$) is the only form that is consistent with balanced growth. Normalizing $A_0 = 1$ and $L_0 = 1$, we can rewrite the objective function as

$$\begin{aligned}\sum_{t=0}^{\infty} \beta^t L_t \frac{c_t^{1-\sigma}}{1-\sigma} &= \sum_{t=0}^{\infty} \beta^t (1+n)^t \frac{((1+\gamma)^t \tilde{c}_t)^{1-\sigma}}{1-\sigma} \\ &= \sum_{t=0}^{\infty} \tilde{\beta}^t \frac{\tilde{c}_t^{1-\sigma}}{1-\sigma},\end{aligned} \tag{4.19}$$

where now we use the adjusted discount factor $\tilde{\beta}$:

$$\tilde{\beta} \equiv \beta(1+n)(1+\gamma)^{1-\sigma}. \tag{4.20}$$

The maximization problem for the social planner can then be written as maximizing (4.19) subject to the resource constraint (4.18) and the non-negativity constraints $\tilde{c}_t \geq 0$ and $k_{t+1} \geq 0$. The problem can be solved using standard procedures, and the Euler equation can be derived as:

$$(1+n)(1+\gamma)(\tilde{c}_t)^{-\sigma} = \tilde{\beta}(\tilde{c}_{t+1})^{-\sigma}\left[f'(\tilde{k}_{t+1}) + (1-\delta)\right].$$

Note that, given the definition of $\tilde{\beta}$ in (4.20), the term $(1+n)$ drops out when the Euler equation is written with the original discount factor β. Therefore, the social planner's intertemporal allocation is not affected by the population growth under our utilitarian assumption.

Along the balanced growth path, each variable grows at a constant rate. From our normalized model above, the balanced growth path can be found by imposing steady-state conditions in the $(\tilde{c}_t, \tilde{k}_{t+1})$ variables. This procedure reduces the Euler equation to

$$(1+n)(1+\gamma) = \tilde{\beta}\left[f'(\bar{\tilde{k}}) + (1-\delta)\right].$$

By setting $n = \gamma = 0$, we can check that the above equations coincide with the ones derived under no growth. If we assume a Cobb-Douglas production function, replacing $f'(\bar{\tilde{k}}) = \alpha \bar{\tilde{k}}^{\alpha-1}$ in the equation above, we can deliver $\bar{\tilde{k}}$ as a function of the relevant parameters of the model.

The question of whether or not the economy will converge to its balanced growth path from arbitrary initial conditions will be dealt with in the next section.

4.4 Recursive methods

In the previous section, we solved the maximization problem searching for a sequence of real numbers $\{x_{t+1}^*\}_{t=0}^{\infty}$ that achieves the highest value of the objective function. This involved finding a solution to an infinite sequence of equations (e.g., a difference equation). It is conceptually useful to break down this high-dimensional problem into a sequence of similar, but smaller problems, which all are tied to each other: "recursive" refers to this tie. This principle is at the core of the recursive method known as *dynamic programming*, introduced

by Richard E. Bellman in the 1950s. A key difference with sequential methods is that the solution to the optimization problem will now be a *function* rather than a sequence of numbers. In what follows, we present the idea behind recursive methods and how they can be used; we discuss the precise theoretical links between the sequential approach and the functional approach briefly in Section 4.4.3 below.

4.4.1 Dynamic programming and the Bellman equation

An implicit assumption in the sequential formulation was that the whole path of x_{t+1} was chosen in the initial period. The key to dynamic programming is to think of dynamic decisions as being made not once-and-for-all but period by period instead. In other words, the value of x_{t+1} is decided in period t rather than at date 0. A key question is whether the two formulations are identical. They will be, as long as the problem at hand is **stationary**. This is the case whenever the structure of the maximization problem that a decision maker faces is identical in nature at every point in time. To make ideas more concrete, let's revisit the infinite-horizon version of the problem (P1).

$$V(x_0) \equiv \max_{\{x_{t+1}\in\Gamma(x_t)\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t \mathcal{F}(x_t, x_{t+1}), \tag{4.21}$$

where $\Gamma(x_t)$ represents the feasible choice set for x_{t+1} given x_t. In the expression, the "value function" $V(x_0)$ represents the value of the objective function at the optimum given the initial condition x_0. This high-dimensional problem can be broken down and re-written as

$$\begin{aligned}
V(x_0) &= \max_{\{x_{t+1}\in\Gamma(x_t)\}_{t=0}^{\infty}} \left\{ \mathcal{F}(x_0, x_1) + \beta \sum_{t=1}^{\infty} \beta^{t-1} \mathcal{F}(x_t, x_{t+1}) \right\} \\
&= \max_{x_1\in\Gamma(x_0)} \left\{ \mathcal{F}(x_0, x_1) + \beta \left[\max_{\{x_{t+1}\in\Gamma(x_t)\}_{t=1}^{\infty}} \sum_{t=1}^{\infty} \beta^{t-1} \mathcal{F}(x_t, x_{t+1}) \right] \right\} \\
&= \max_{x_1\in\Gamma(x_0)} \left\{ \mathcal{F}(x_0, x_1) + \beta \underbrace{\left[\max_{\{x_{t+2}\in\Gamma(x_{t+1})\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^{t} \mathcal{F}(x_{t+1}, x_{t+2}) \right]}_{\equiv V(x_1)} \right\}
\end{aligned}$$

The simple mathematical idea that $\max_{x,y} f(x,y) = \max_y\{\max_x f(x,y)\}$, provided that each of the max operators is well-defined, allows us to maximize "in steps." To do so, first notice that the problem in squared brackets (in the last row) has the same structure as equation (4.21), but with a different initial condition. In other words, it represents the value of the objective function attained by an agent that chooses the optimal sequence of x_{t+1} given the initial condition x_1. Because the time horizon is the same, and neither the instantaneous objective function nor the feasible set Γ change over time, then it must be the case that the problem in squared brackets is equal to $V(x_1)$. This implies that we can write

$$V(x_0) = \max_{x_1\in\Gamma(x_0)} \left\{ \mathcal{F}(x_0, x_1) + \beta V(x_1) \right\}.$$

The two representations are identical as long as the problem is stationary. Intuitively, a dynamic problem is stationary if the problem at t and at $t+1$ look the same, so one can capture all relevant information for the decision maker in a way that does not involve time. Not all problems are stationary. For example, consider the finite-horizon version of the neoclassical growth model. In it, agents care about how many periods are left when choosing investment. The decision problem changes as the terminal period approaches. One can still describe the problem recursively: break it up into a current choice between consumption and investment and all future choices as implicitly captured by a value function. The difference is that the value function will depend on time, V_t.

With infinitely many periods, the remaining horizon is the same at every t. For different values of the initial capital stock, the choices differ, but do not depend on the time period in which the choice is made. This means that a decision at any point in time does not depend on anything but the level of capital at that point in time.

In general, the predetermined, payoff-relevant information for the consumer is—as was mentioned earlier in this chapter—called a *state variable*. The state variable for the planner in the neoclassical growth model is the current stock of capital k_t. In our generic example, it is simply given by x_t. When transforming a sequential problem into a recursive one, it is important to choose the state variables appropriately so that the resulting problem is indeed recursive.

When the problem is stationary, decisions take a stationary form:

$$x_{t+1} = g(x_t).$$

In particular, the function determining how future capital depends on the current state does not vary with time. The function $g(\cdot)$ is known as the **decision rule** or **policy function**.

Given an arbitrary initial condition x_t, we can write the recursive problem as

$$V(x_t) = \max_{x_{t+1} \in \Gamma(x_t)} \{\mathcal{F}(x_t, x_{t+1}) + \beta V(x_{t+1})\}.$$

This is the dynamic-programming formulation. The derivation was completed for a given value of x_t on the left-hand side of the equation. On the right-hand side, however, we need to know V evaluated at any value for x_{t+1} in order to be able to perform the maximization. Going forward, we now change our notation and use x to denote current values and, adding a prime, x', to denote next period's values. This change of notation is conceptually important: in a stationary dynamic program, "t" actually has no place as time is no longer of essence. Above, it was used to connect the sequence problem to the recursive one; the only remaining need is to distinguish today's x from tomorrow's x'. We write our dynamic-programming equation as the quest for a $V(x)$ satisfying

$$V(x) = \max_{x' \in \Gamma(x)} \{\mathcal{F}(x, x') + \beta V(x')\} \tag{4.22}$$

for all values of x. This equation is called the "Bellman equation." It is a *functional equation*: the unknown is a function $V(x)$; i.e., it needs to satisfy equation (4.22) for all values of the argument of the function, x. It can be intuitively interpreted as follows: the discounted lifetime value of our representative agent's objective function is equal to the instantaneous

value $\mathcal{F}(x, x')$ received today, plus the discounted value enjoyed from tomorrow into the future, $\beta V(x')$, all assuming optimal choices.

We use the function g alluded to above to denote the arg max in the functional equation:

$$g(x) = \arg\max_{x' \in \Gamma(x)} \{\mathcal{F}(x, x') + \beta V(x')\}$$

for all x, or the decision rule for x': $x' = g(x)$. This notation presumes that a maximum exists and it is unique. Otherwise, g would either not exist or not be a function and, rather, a correspondence. Note that the following must hold for all x by definition.

$$V(x) = \mathcal{F}(x, g(x)) + \beta V(g(x)).$$

4.4.2 Writing a problem recursively

Sometimes we are presented with the sequential formulation of a problem, and it is necessary to 'translate it' into a recursive formulation. We can do this for several reasons. Perhaps the most important one is conceptual: finding a recursive formulation offers an added understanding of the problem at hand. That is because when a problem has been formulated recursively, we know what behavior will depend on: the state variable(s) of the recursive formulation, and on nothing else. In practice, this often means that we can obtain insights into how a problem can be solved. Another reason for studying the recursive problem is that it is associated with convenient computational methods as will be described in Chapter 10.

Consider the infinite-horizon version of the neoclassical growth model, which corresponds to (P3) with $T = \infty$. We first need to make sure that the problem is indeed stationary. That is, that the decision-maker faces the same type of problem every period given the same value of the state variables. That brings us to the question of what the relevant state variables ought to be.

The standard NGM In the NGM, the state variable is given by the stock of capital at the outset of the period, k. In other words, $x = k$ in this model. The value function depends on k, implying that we can write $V(k)$. The control variable is $c = f(k) + (1 - \delta)k - k'$, where we already used the notation that next period variables are denoted with primes (e.g. k_{t+1} is written as k'). The recursive formulation of the neoclassical growth model is

$$V(k) = \max_{k' \in \Gamma(k)} \{u(f(k) + (1 - \delta)k - k') + \beta V(k')\}, \tag{P4}$$

with $\Gamma(k) = [0, f(k) + (1 - \delta)k]$, for all k. While this problem is relatively straightforward to write in recursive form, this is not always the case, as we see with the next two examples.

The periodic NGM:

Consider a small deviation from the problem above, where the level of technology oscillates deterministically between two values A_h and A_l, where $A_h > A_l$. In particular, period t output equals $A_h f(k_t)$ if t is even and $A_l f(k_t)$ if t is odd. The resource constraint is $c_t + k_{t+1} = A_t f(k_t) + (1 - \delta)k_t$, as before. At first sight, this problem does not look stationary because the level of productivity is changing over time. However,

it is changing deterministically and predictably. We can write the problem recursively exploiting the periodicity in the evolution of TFP. To do so, we need to incorporate an additional state variable: whether we are in an even or odd period, which in turn corresponds to an h or l value for A_t. We can therefore write

$$V_h(k) = \max_{k' \in \Gamma_h(k)} \{u(A_h f(k) + (1-\delta)k - k') + \beta V_l(k')\}$$

$$V_l(k) = \max_{k' \in \Gamma_l(k)} \{u(A_l f(k) + (1-\delta)k - k') + \beta V_h(k')\}$$

with $\Gamma_i(k) = [0, A_i f(k) + (1-\delta)k - k']$ for $i \in \{h, l\}$, for all k. Clearly, given the same values of these two states, the decision maker would always select the same k'. So we were able to write the problem recursively with the appropriate choice of state variables.

The delayed depreciation NGM: The objective and resource constraints are standard,

$$\begin{aligned} \max_{\{c_t\}_{t=0}^{\infty}} \quad & \sum_{t=0}^{\infty} \beta^t u(c_t) \\ \text{s.t.} \quad & c_t + i_t = F(k_t), \end{aligned}$$

but capital depreciates fully in two periods, and does not depreciate at all before that, so that the law of motion for capital given a sequence of investments $\{i_t\}_{t=0}^{\infty}$ is given by

$$k_t = i_{t-1} + i_{t-2}.$$

In the first period, $k_0 = i_{-1} + i_{-2}$, with two initial conditions i_{-1} and i_{-2}. The recursive formulation for this problem (with "primes" denoting consecutive periods) becomes

$$\begin{aligned} V(i', i) = \quad & \max_{c, i''} \{u(c) + V(i'', i')\} \\ \text{s.t.} \quad & c = f(i' + i) - i''. \end{aligned}$$

Notice that there are two state variables in this problem. That is unavoidable here; there is no way of summarizing what one needs to know at a point in time with only one state variable. Both i_{-1} and i_{-2} are natural state variables: they are predetermined, they affect utility and decision making, and neither is redundant.

4.4.3 Properties of the value function

We have argued that there are two different ways of solving a dynamic maximization problem: (i) the sequential method, where the task is to find a sequence (that maximizes the objective function), and (ii) dynamic programming, where the task is to find a function solving the recursively stated problem, i.e., solving a functional equation. A formal mathematical statement that these two methods are equivalent is beyond the scope of this text; Stokey and Lucas (1989) has all the details. For each of the two approaches, it is important to specify a space within which one looks for a solution—a space of "allowable" sequences and a space of allowable functions, respectively. For dynamic programming, one most conveniently restrict attention to bounded, continuous functions V. For this choice to work out,

one needs assumptions on primitives ($\mathcal{F}$, Γ, and β in our generic formulation); minimum conditions involve (i) continuity of $\mathcal{F}$, (ii) non-emptiness, continuity, and compactness of Γ, and (iii) $\beta < 1$.[13] Under these assumptions, the functional equation defined by the dynamic program has a unique solution (it is the value function what is unique; the policy function can be a correspondence, unless we place further restrictions on $\mathcal{F}$ and Γ.) Moreover, for any given value of the state, it allows us to construct a sequence of choices corresponding to solutions to the sequentially stated problem. Conversely, solutions in the sequence space can be used to construct a value function solving the dynamic programming problem.

We now summarize several key theoretical properties of the dynamic programming problem and briefly comment on them. For proofs, see Stokey and Lucas (1989). You may have studied the necessary ingredients into the proofs in preparatory math, in particular the fact that the dynamic program defines a *contraction mapping*.

1. It is possible to find $V(x)$ by an iterative process. The procedure is as follows.

   i. Select any initial $V_0(x)$ function. One example is $V_0(x) = 0 \ \forall x$.

   ii. Define a sequence of functions as follows: for all x,

   $$V_{n+1}(x) = \max_{x' \in \Gamma(k)} \mathcal{F}(x, x') + \beta V_n(x')$$

   for $n = 0, 1, 2, \ldots$.

   Then the following is true:

   (i) the resulting sequence $\{V_j(x)\}_{j=0}^{\infty}$ converges to $V(x)$, i.e., to the function that solves the dynamic program;

   (ii) the distance to the solution V gets smaller and smaller at a constant rate: $\|V_{n+1} - V\| \leq \beta \|V_n - V\|$, where $\|f\|$ denotes a distance between functions.[14]

   Notice that the particular initial guess $V_0 = 0$ delivers an interpretation of V_n: it represents present-value utility for an economy consisting of n periods. Clearly, V_n is an addition of "utils" over time which involves optimal choice of x' in all periods, including the last one, V_1.

   Feature (ii) is rather remarkable: it holds no matter what the initial guess is and it is therefore very useful in practical applications: for finding a solution numerically. In particular, if one can bound the initial error by some value $\bar{\epsilon} \equiv \|V_0 - V\|$ and then generate the sequence of functions, one knows that after n iterations the error is at most $\beta^n \bar{\epsilon}$. Finally, (ii) is extremely useful in allowing us to prove properties of the value function, such as those we state next.

[13] Some maximization problems involve discrete choice; we could, for example, restrict the choice of capital to belong to a finite set of values. The conditions on primitives then become correspondingly weaker: $\mathcal{F}$ has to be bounded and Γ nonempty.

[14] The distance is a norm; in practice the sup-norm is used, i.e., the largest difference across all values of the argument of the function, i.e., $\|f\| \equiv \sup_x f(x)$.

2. Assume that $\mathcal{F}$ is strictly increasing in its first argument and that Γ is monotone: if $x \leq \tilde{x}$, then $\Gamma(x) \subseteq \Gamma(\tilde{x})$. Then V is strictly increasing.

3. Assume that $\mathcal{F}$ is strictly concave (in its two arguments jointly) and that $\Gamma(x)$ is convex in the following sense: if $x' \in \Gamma(x)$ and $\tilde{x}' \in \Gamma(\tilde{x})$, then $\theta x' + (1-\theta)\tilde{x}' \in \Gamma(\theta x + (1-\theta)\tilde{x})$. Then V is strictly concave and the policy unique (i.e., g is a well-defined function).

4. Assume that $\mathcal{F}$ and Γ satisfy the properties in the previous statement and, in addition, that $\mathcal{F}$ is continuously differentiable. Then for any x where the choice is interior, V is differentiable, i.e., $V'(x)$ exists.

5. Assume that $\mathcal{F}$ and Γ satisfy all the conditions above. Then, the policy function $g(x)$ is strictly increasing.

Very briefly, property 1 above follows from the Bellman equation being a contraction mapping.[15] Properties 2 and 3 are proved using property 1 as follows: take a V_0 with the desired characteristic (e.g., concavity), show that if V_n has the desired characteristic then so does V_{n+1}, and finally argue that the limit V^* inherits the characteristic at least weakly (e.g., V^* is concave). Strictness of the characteristic (e.g., V^* is strictly concave) follows by applying the argument again, with the limit V^* inserted on the right-hand side of the Bellman equation. Property 4 is also possible to prove relatively straightforwardly but note that it applies only in the case where V is concave. Property 5 follows from the first-order condition

$$-\mathcal{F}_2(x, x') = \beta V'(x').$$

The left-hand side of this equality is clearly increasing in x', since $\mathcal{F}(x, x')$ is strictly concave with respect to its second argument, and the right-hand side is strictly decreasing in x', since $V(x)$ is strictly concave under the stated assumptions. Furthermore, since the right-hand side is independent of x but the left-hand side is decreasing in x, the optimal choice of x' is increasing in x.

4.4.4 Solving for the value function

As for the sequentially formulated maximization problem, analytical solutions are only available in very special cases. When they do exist, there are again a few different methods available to find them. One is a "guess and verify method," also known as the "method of undetermined coefficients." This method, however, requires an insightful initial guess for $V(x)$. When the initial guess is unavailable, we can use the iterative process described above, starting from $V_0(x) = 0$, quite similarly to how we proceeded to solve the finite-horizon model backwards.

We again illustrate these solution methods using the neoclassical growth model, (P4). Further details can be found in Chapter 10.

Example 4.7 *Consider the economy from Example 4.5. The corresponding Bellman equation is*

$$V(k) = \max_{k' \geq 0} \left\{ \log\left(Ak^{\alpha} - k'\right) + \beta V(k') \right\}.$$

[15] A contraction mapping obtains for the Bellman equation so long as $\beta < 1$.

Let the initial guess of the value function be $V_0(k) = 0$. Then,

$$V_1(k) = \max_{k' \geq 0} \{\log [Ak^\alpha - k']\}. \tag{4.23}$$

The right-hand side is maximized by taking $k' = 0$, yielding $V_1(k) = \log A + \alpha \log k$. Using this in the next iteration, we obtain

$$V_2(k) = \max_{k' \geq 0} \{\log [Ak^\alpha - k'] + \beta [\log A + \alpha \log k']\}.$$

The first-order condition delivers

$$\frac{1}{Ak^\alpha - k'} = \frac{\beta\alpha}{k'} \Rightarrow k' = \frac{\alpha\beta Ak^\alpha}{1+\alpha\beta}.$$

We can interpret the resulting expression for k' as the rule that determines how much it would be optimal to save if we were at period $T-1$ in the finite horizon model. We can substitute this into $V_2(k)$ to yield

$$\begin{aligned} V_2(k) &= \log\left[Ak^\alpha - \frac{\alpha\beta Ak^\alpha}{1+\alpha\beta}\right] + \beta\left[\log A + \alpha \log \frac{\alpha\beta Ak^\alpha}{1+\alpha\beta}\right] \\ &= \underbrace{\log\left(A - \frac{\alpha\beta A}{1+\alpha\beta}\right) + \beta \log A + \alpha\beta \log \frac{\alpha\beta A}{1+\alpha\beta}}_{=a_2} + \underbrace{\left(\alpha + \alpha^2\beta\right)}_{=b_2} \log k. \end{aligned}$$

The same procedure can be used to obtain a $V_3(k)$, and so on. This procedure would make the sequence of value functions converge to $V(k)$.

We can also, at this point, guess and verify a functional form for V. Grouping terms in V_2, we can see that it takes the form $V_n(k) = a_n + b_n \log k$ for all n. Therefore, we may already guess that the function to which this sequence is converging has to be of the form:

$$V(k) = a + b \log k.$$

In order to determine the corresponding parameters a, b, we take first-order conditions and find that $k' = \dfrac{\beta b}{1+\beta b} Ak^\alpha$. Inserting this into the right-hand side of the Bellman equation

$$RHS \equiv \max_{k' \geq 0} \{\log (Ak^\alpha - k') + \beta (a + b \log k')\}.$$

and equating the resulting expression to our guess delivers a system of two equations in two unknowns, a and b. The solutions will be

$$b = \frac{\alpha}{1-\alpha\beta} \quad \text{and} \quad a = \frac{1}{1-\beta}\frac{1}{1-\alpha\beta}\left[\log A + \log (1-\alpha\beta)^{1-\alpha\beta} + \log (\alpha\beta)^{\alpha\beta}\right].$$

The resulting decision rule is exactly the same policy rule obtained using sequential methods,

$$k' = \alpha\beta Ak^\alpha.$$

4.4.5 The functional Euler equation

In the sequentially formulated maximization problem, the Euler equation turned out to be a crucial part of characterizing the solution. With the recursive strategy, an Euler equation can be derived as well. Consider the Bellman equation for a general instantaneous objective function $\mathcal{F}(x, x')$:

$$V(x) = \max_{x' \in \Gamma(x)} \{\mathcal{F}(x, x') + \beta V(x')\}.$$

Under suitable assumptions, this problem delivers the policy function $x' = g(x)$. Hence,

$$V(x) = \mathcal{F}(x,\ g(x)) + \beta V(g(x)). \tag{4.24}$$

Assuming an interior solution, $g(x)$ satisfies the first-order condition

$$\mathcal{F}_2(x,\ x') + \beta V'(x') = 0.$$

Evaluated at the optimum, i.e., at $x' = g(x)$,

$$\mathcal{F}_2(x,\ g(x)) + \beta V'(g(x)) = 0.$$

In contrast to how we derived the sequential formulation, we now need to know the derivative of the value function $V'(\cdot)$ in order to solve for the policy rule. Even though it is not possible, in general, to write $V(x)$ in terms of primitives, we can find its derivative. Using the equation (4.24) above, one can differentiate both sides with respect to x (recall that the equation holds for all x and, again under some assumptions stated earlier, is differentiable). We obtain

$$V'(x) = \mathcal{F}_1(x,\ g(x)) + \underbrace{g'(x)\Big[\mathcal{F}_2(x,\ g(x)) + \beta V'(g(x))\Big]}_{\text{indirect effect through optimal choice of } x'},$$

where "$g'(x)$" represents the derivative of the policy function. The differentiability of g was not established above, and it is harder to prove, but it is actually not required in order to derive the result that just will follow. Namely, from the first-order condition, the argument in brackets in the equation just stated is zero and hence

$$V'(x) = \mathcal{F}_1(x,\ g(x)),$$

which again holds for all values of x. The indirect effect disappears: this is an application of a general result known as the **Envelope Theorem**.

Updating, we know that $V'(g(x)) = \mathcal{F}_1\Big(g(x),\ g(g(x))\Big)$ also has to hold. The first-order condition can now be rewritten as follows:

$$\mathcal{F}_2(x,\ g(x)) + \beta \mathcal{F}_1\Big(g(x),\ g(g(x))\Big) = 0\ \forall x. \tag{4.25}$$

This is the Euler equation stated as a functional equation: it does not contain the unknowns x_t, x_{t+1}, and x_{t+2} but, rather, has to hold for all x. Recall our previous Euler equation formulation

$$\mathcal{F}_2(x_t,\ x_{t+1}) + \beta \mathcal{F}_1(x_{t+1},\ x_{t+2}) = 0, \forall t,$$

where the unknown was the sequence $\{x_t\}_{t=1}^{\infty}$. Now instead, the unknown is the function g. That is, under the recursive formulation, the Euler equation turned into a functional equation: the **functional Euler equation**.

Solving directly for policy The previous discussion suggests that a third way of searching for a solution to the dynamic problem is to consider the functional Euler equation, and solve it for the function g. We have previously seen that we can (i) look for sequences solving a nonlinear difference equation plus a transversality condition; or (ii) we can solve a Bellman (functional) equation for a value function.

The functional Euler equation offers another way to solve directly for behavior, thus bypassing the value function. Its key feature is that it expressed an intertemporal tradeoff. Here, the recursive approach provides some extra structure relative to the sequential Euler equation: it tells us that the optimal sequence of capital stocks needs to be connected using a stationary function. Mathematically, unlike the Bellman equation, the functional Euler equation is not a contraction mapping. It is also only a necessary condition on the optimal policy function g and, viewed in isolation, can allow multiple solutions. Only one of these solutions, however, is the policy function that solves the right-hand side of the Bellman equation. It will become clear below how, at least locally, multiple solutions are possible.

The functional Euler equation approach is often used in practice in solving dynamic problems numerically. We now show an example for which an analytical solution exists.

Example 4.8 *Consider the model used in Example 4.7. With full depreciation* $\mathcal{F}(k,k') = u(f(k)-k')$*. Then, the respective derivatives are:*

$$\begin{aligned}\mathcal{F}_1(k,k') &= u'(f(k)-k')\,f'(k)\\ \mathcal{F}_2(k,k') &= -u'(f(k)-k').\end{aligned}$$

In the particular parametric example, and replacing $k' = g(k)$*, equation (4.25) becomes:*

$$\frac{1}{Ak^{\alpha}-g(k)} - \frac{\beta\alpha A\,(g(k))^{\alpha-1}}{A\,(g(k))^{\alpha} - g\,(g(k))} = 0, \forall k.$$

This is a functional equation in $g(k)$*. Guess that* $g(k) = sAk^{\alpha}$*, i.e., saving is a constant fraction of output. Substituting this guess into functional Euler equation delivers*

$$\frac{1}{(1-s)\,Ak^{\alpha}} = \frac{\alpha\beta A\,(sAk^{\alpha})^{\alpha-1}}{A\,(sAk^{\alpha})^{\alpha} - sA\,(sAk^{\alpha})^{\alpha}}.$$

As can be seen, k *cancels out, and the remaining equation can be solved for* s*. Collecting terms and factoring out* s*, we obtain*

$$s = \alpha\beta.$$

Thus, $\alpha\beta Ak^{\alpha}$ *satisfies the functional Euler equation for all values of* k*. It is, moreover, the same answer that we arrived at in Example 4.7.*

4.4.6 Dynamics in the optimizing neoclassical growth model

We now apply recursive methods to characterize the solution to the planning problem involving the standard neoclassical growth model. We could also apply our methods to the case of consumer saving under a constant wage and interest rate, but we leave that application for the reader. It may be useful to first note that there is a unique maximum attainable

level of the capital stock, k_u. That is, if the capital stock starts out below (or at) k_u, it cannot take a higher value than k_u in the future. This result follows because for any k, $f(k) + (1-\delta)k$ is the highest feasible saving k', because this implies zero consumption. Hence, $f(k_u) + (1-\delta)k_u = k_u$ defines k_u and we can restrict attention to a search for value and policy functions over the closed and bounded set $[0, k_u]$. There are two solutions to this equation: a strictly positive k_u and 0; we discard the latter.

We know from the functional Euler equation (4.25) that, for all $k \in [0, k_u]$, $g(k)$ satisfies

$$u'\left(f(k) + (1-\delta)k - g(k)\right) = \beta u'\left(f(g(k)) + (1-\delta)g(k) - g(g(k))\right)\left(f'(g(k)) + 1 - \delta\right). \tag{4.26}$$

In particular, at a steady-state $\bar{k}$, $g(\bar{k}) = \bar{k}$, which—given that the argument of u' is non-zero—allows us to write the steady-state condition as

$$1 = \beta(f'(\bar{k}) + 1 - \delta), \tag{4.27}$$

which is identical to (4.10) above. Now we can proceed to show global convergence to the steady state.

We use the property that under the standard assumptions the value function $V(k)$ is strictly concave.[16] This property implies that we can write

$$\left[V'(k) - V'(g(k))\right]\left[k - g(k)\right] \leq 0 \quad \forall k \in [0, k_u].$$

To see that the inequality holds, note that whenever $g(k) > k$, the expression contained in the left-most bracket is positive, whereas the right-most bracket contains a negative quantity, and vice versa.

Using the envelope theorem, we know that

$$V'(k) = u'\left(f(k) + (1-\delta)k - g(k)\right)\left(f'(k) + 1 - \delta\right).$$

From the first-order condition in the Bellman equation, we obtain

$$V'(g(k)) = u'\left(f(k) + (1-\delta)k - g(k)\right)\frac{1}{\beta}.$$

Inserting these two expressions into the left-most bracket of the inequality and factorizing u', we obtain

$$u'(f(k) - g(k))\left[f'(k) + 1 - \delta - \frac{1}{\beta}\right]\left[k - g(k)\right] \leq 0 \quad \forall k \in [0, k_u]. \tag{4.28}$$

When $k = \bar{k}$, from (4.27), $f'(k) + 1 - \delta = 1/\beta$ and $g(k) = k$ holds. When $k > \bar{k}$, then $f'(k) + 1 - \delta < 1/\beta$, and (4.28) implies $g(k) < k$, and capital decreases over time. When $k < \bar{k}$, capital increases. This result implies, given that we also know that g is increasing (property 5 above), that the system is globally stable and converges monotonically to $\bar{k}$.

[16] This requires us to show that $u(f(k) + (1-\delta)k - k')$ is concave in (k, k'). This can be accomplished by computing the Hessian and showing that it is negative definite.

Approximating the policy function Most often, it is not possible to solve for the policy function in analytical form. In such cases, we can approximate the solution locally with a linear function,

$$g(k) \sim a_0 + a_1 k$$

where a_0 and a_1 are the coefficients that we need to solve for, around the steady-state value of capital. We need two equations to solve for our two unknowns, but since $g(k) = k = \bar{k}$ in steady state we obtain one condition as

$$a_0 = \bar{k}(1 - a_1),$$

where $\bar{k}$ solves (4.10). To obtain a_1, we use a procedure that parallels linearization techniques discussed in Chapter 3, but is less cumbersome because we exploit the fact that the functional Euler equation must hold for all k. Thus, we can simply differentiate equation (4.26) with respect to k, i.e., take the derivative of the left-hand side, LHS, and set it equal to the derivative of the right-hand side, RHS. We obtain, with the obvious notation, $LHS_k(k) = RHS_k(k)$, which must hold for all k. Evaluating this expression at steady state allows us to obtain an equation that provides a second condition relating a_1 to a_0. Hence,

$$LHS_k(\bar{k}) = u''(\bar{c})[f'(\bar{k}) + 1 - \delta - \underbrace{g'(\bar{k})}_{=a_1}]$$

$$RHS_k(\bar{k}) = \beta u''(\bar{c})\Big[(f'(\bar{k}) + 1 - \delta)a_1 - \underbrace{g'(\bar{k})g'(\bar{k})}_{=a_1^2}\Big]\Big[f'(\bar{k}) + 1 - \delta\Big] + \beta u'(\bar{c})f''(\bar{k})a_1.$$

where we used the fact that, under our guess, $g(g(k)) = a_0 + a_1[a_0 + a_1 k]$; furthermore, $\bar{c}$ is defined as $f(\bar{k}) - \delta\bar{k}$. Setting $LHS_k(\bar{k}) = RHS_k(\bar{k})$ and simplifying (recall $f'(\bar{k}) + 1 - \delta = 1/\beta$) delivers

$$u''(\bar{c})\left[1 - \beta a_1\right] = u''(\bar{c})\left[a_1 - \beta a_1^2\right] + \beta^2 u'(\bar{c})f''(\bar{k})a_1, \tag{4.29}$$

which can be used to solve for a_1 given $\bar{k}$. Clearly, a_1 satisfies a second-order polynomial equation. It is straightforward to show that this equation has two real solutions, of which one is strictly between zero and one, and the other is strictly greater than one. Therefore, locally, $g(k)$ can be approximated by two functions. This was alluded to above and it is not a surprise, given that the functional Euler equation is only a necessary condition. A function g which solves this equation also has to attain the maximum on the right-hand side of the Bellman equation. Of the two functions obtained here, only one has that property: the one with an a_1 between zero and one. It is the one corresponding to a slope less than one in (k, k') space, giving monotone convergence to the steady state. Concretely, by denoting $\Theta \equiv \beta^2 u' f''/u'' > 0$, the solution of (4.29) is

$$a_1 = \frac{1 + \beta + \Theta - \sqrt{(1 + \beta + \Theta)^2 - 4\beta}}{2\beta}. \tag{4.30}$$

It can easily be checked that $a_1 \in (0, 1]$, a_1 is decreasing in Θ, and $a_1 = 1$ when $\Theta = 0$ and $a_1 \to 0$ as $\Theta \to \infty$.

Using this procedure, it is possible to also obtain higher-order approximations to the policy function g around the steady state. To obtain the second-order Taylor approximation, simply differentiate the functional equation once more with respect to k, which delivers g'' at the steady state: it can be solved for as a function of g and g', which were previously solved-out. Further differentiations will, successively, give us any desired higher-order approximations.[17]

Calibration We discussed the concept of calibration in Chapter 3: the adoption of specific functional forms and parameter values with the purpose of generating quantitative predictions. We applied it in that chapter to gauge the Solow model's quantitative predictions for the speed of convergence to a steady state. The NGM developed in this chapter, instead, differs in its convergence properties: the saving rate is not constant away from steady state but is instead endogenous and, as we have seen, a function of the current level of the capital stock. How strong the dependence on capital is depends on the utility function. More precisely, equation (4.30) determines a_1, the slope of the saving function as it crosses the steady-state line: a slope of zero implies convergence in one period (infinite speed) and a slope of 1 implies no (infinitely slow) convergence. The remaining model parameters can be selected as in Chapter 3 so our only question here is: what is a_1, or rather, what is the value of Θ at the steady state? Recall that $\Theta = \beta^2 u' f''/u''$, $\bar{c} = f(\bar{k}) - \delta\bar{k}$, and $f'(\bar{k}) = 1/\beta - (1-\delta)$. Furthermore, suppose that $u(c) = (c^{1-\sigma} - 1)/(1-\sigma)$. Then it can be shown that

$$\Theta = -\frac{\beta^2 f''(\bar{k})(f(\bar{k}) - \delta\bar{k})}{\sigma}.$$

Because $f''(\bar{k})$ is negative, Θ is positive and decreasing in σ. When $\sigma = 0$, the utility function is linear, and in this case $\Theta \to \infty$. The above result implies $a_1 \to 0$ in this case, implying k' does not depend on k. Convergence, in this case, is immediate. The opposite extreme is when $\sigma \to \infty$, $\Theta \to 0$ and $a_1 \to 1$. In this case, the utility function is "infinitely curved" (Leontief utility). When σ is very large, consumers are extremely unwilling to change consumption over time and convergence is very slow.

Therefore, σ is an important parameter in determining the speed of convergence. What do empirical studies of consumption suggest as an appropriate value of σ? Using aggregate consumption data, Hall (1988) estimates that $1/\sigma = 0.1$. Attanasio and Weber (1993) and Attanasio and Weber (1995) use micro data instead and finds that $1/\sigma \in [0.3, 0.8]$. In many applied macroeconomic studies using the CRRA function, however, there are also stochastic components, and under uncertainty, σ is also equal to the coefficient of relative risk aversion (uncertainty is discussed in Chapter 7 below) and this coefficient is often estimated to be above 1. Perhaps for this reason, much of the applied literature focuses on $\sigma = 1$ (log utility) or $\sigma = 2$, though rarely much higher than one.

Further assume that the production function $f(k)$ is a power function: $f(k) = k^\alpha$. This formulation implies the aggregate production function is of the Cobb-Douglas form. In this case, Θ can be solved as

$$\Theta = \frac{\beta^2(1-\alpha)}{\sigma}\left(\frac{1}{\beta} - 1 + \delta\right)\left[\frac{1}{\alpha}\left(\frac{1}{\beta} - 1 + \delta\right) - \delta\right],$$

[17] Taylor-based approximations, such as those used here, naturally require the underlying function to be differentiable. Accordingly, we approximate the endogenous function g under that assumption.

which is decreasing in α and increasing in δ. Convergence is slower when the production function is closer to linear and the depreciation rate is small. This qualitative property is shared with the Solow model.

4.5 Concluding remarks

In this chapter, we showed a set of tools that are useful for solving standard dynamic optimization models used in macroeconomics. We discussed the sequential formulation, where the aim is to choose the best sequence of allocations (quantities) that maximize lifetime utility, starting with a finite horizon economy and then moving to an infinite horizon economy, highlighting the mathematical complications that arise. We then showed how to use dynamic programming methods to write the optimization problem recursively, where the key is to solve for policy functions determining optimal allocations. In practice, all these methods are used. We spent less time on motivating specific functional forms and, throughout, only used two examples: a pure consumption-saving problem under price-taking and an optimizing neoclassical growth model. Neither of these cases allowed valued leisure or other, richer optimization problems. However, the methods we introduce here are straightforwardly extended to other, richer contexts and will indeed be used over and over in the rest of the text.