Kruna Ratković[1]

Limitations in Direct and Indirect Methods for Solving Optimal Control Problems in Growth Theory

Abstract: *The focus of this paper is on a comprehensive analysis of different methods and mathematical techniques used for solving optimal control problems (OCP) in growth theory. Most important methods for solving dynamic non-linear infinite-horizon growth models using optimal control theory are presented and a critical view of the limitations of different methods is given. The main problem is to determine the optimal rate of growth over time in a way that maximizes the welfare function over an infinite horizon. The welfare function depends on capital-labor ratio, the state variable, and the per-capita consumption, the control variable. Numerical methods for solving OCP are divided into two classes: direct and indirect approach. How the indirect approach can be used is given in the example of the neo-classical growth model. In order to present the indirect and the direct approach simultaneously, two endogenous growth models, one written by Romer and another by Lucas and Uzawa, are studied. Advantages and efficiency of these different approaches will be discussed. Although the indirect methods for solving OCP are still the most expanded in growth theory, it will be seen that using direct methods can also be very efficient and help to overcome problems that can occur by using the indirect approach.*

Key words: *optimal control, direct and indirect methods, growth theory.*

---

[1] University of Donja Gorica, Faculty of Applied Sciences, Montenegro, kruna.ratkovic@gmail.com

Ograničenja u primeni direktnih i indirektnih metoda za rešavanje problema optimalne kontrole u teoriji rasta

Apstrakt: *Fokus ovog članka je na detaljnoj i sveobuhvatnoj analizi glavnih metoda i matematičkih tehnika koje se koriste za rešavanje problema optimalne kontrole u teoriji rasta. Dat je pregled najvažnijih metoda za rešavanje dinamičkih nelinearnih modela rasta koristeći optimalnu kontrolu, kao i kritički osvrt na njihova ograničenja. Osnovni problem koji treba rešiti ovim pristupom je određivanje optimalne stope rasta tokom vremena na način koji maksimizira funkciju blagostanja u beskonačnom vremenskom periodu. Funkcija blagostanja zavisi od koeficijenta kapitalne opremljenosti rada (promjenljive stanja) i od potrošnje po glavi stanovnika (kontrolne promjenljive). Numeričke metode za rešavanje problema optimalne kontrole su podeljene u dve klase: direktni i indirektni pristup. Na primjeru neoklasičnog modela rasta dat je prikaz indirektnog pristupa. Kako bi se predstavio istovremeno indirektni i direktni pristup, u radu će biti data i primena ovih metoda kod dva endogena modela: Romerov i Lucas-Uzawa model. Biće date prednosti i efikasnost jedne metode u odnosu na drugu. Iako se indirektne metode za rešavanje problema optimalne kontrole u ovoj oblasti i dalje najviše upotrebljavaju u praksi, biće viđeno da primjena direktnih metoda može biti vrlo efikasna i korisna u prevazilaženju problema koji se mogu javiti kod indirektnog pristupa.*

Ključne reči: *optimalna kontrola, direktne i indirektne metode, teorija rasta*

1. Introduction

Optimal control theory, as an extension of the calculus of variations, represents a modern approach to dynamic optimization. Because of the complexity of most applications, optimal control problems are most often solved numerically. Numerical methods for solving optimal control problems date back to the 1950s and the work of Bellman (1957). Since the complexity and variety of applications has increased over the last decades, the complexity of methods in optimal control increased as well, such that optimal control theory has become a discipline that is important to many branches of economics. Its principal applications in economics are in financial economics, dynamic macroeconomic theory and resource economics. In particular, optimal control theory has been applied in growth theory since the article of Arrow (1968) and later in both exogenous and endogenous growth models such as Ramsey (1928), Solow (1956), Uzawa (1965), Lucas (1988) and

Romer (1994). Optimal control techniques and the economic interpretation have been discussed by authors such as Arrow and Kurz (1970), Aghion and Howitt (1998), Barro and Sala-i-Martin (2003), Cass (1965), Chiang (1992), Dixit (1990), Dorfman (1969), Intrilligator (1971), Kamien and Schwartz (1991), Koopmans (1963) and others. Since the same principles are applied to the new growth models as to the neoclassical growth model, there is a great literature as an introduction to optimal control techniques (cf. Acemoglu (2007), Aseev (2009)).

The study of these dynamic growth models follows a usual procedure, which is to apply Pontryagin's Maximum Principle and to obtain the necessary optimality conditions, together with the transversality condition. If the initial values for the state variables are defined, it will give us a complete description of the system, enough to explain the economy around its steady state. This formulation gives rise to a system of nonlinear differential equations describing the economy. Nonlinearity usually arises from the diminishing marginal utility of consumption and from the diminishing marginal productivity of the factors of production, but can also be the result of incorporating R&D (Williams and Jones, 1995) or government spending (Barro, 1990). Nonlinearity in the production and utility functions and saddle path stability, can influence the analysis of the transitional dynamics after a structural change or a policy shock, as (Atolia, Chatterjee, & Turnovsky, 2008) point out. One way to overcome this problem would be to *linearize* the dynamic system around its (post-shock) steady state and then to study this linearized, hence simplified, version of the dynamic system as an analogue to the original nonlinear one. But linearization can potentially be fallacious as Wolman and Couper (2003) noticed. Nevertheless, linearization is still the principal method used in the growth literature. Consequently, there were some numerical methods that tried to overcome this problem of linearization such as the projection method (Judd, 1992), the discretization method (Mercenier & Michel, 1994), the shooting method (Judd, 1998), the time elimination method (Mulligan & Sala-i Martin, 1991), the backward integration procedure (Brunner & Strulik, 2002) and the relaxation procedure (Trimborn, Koch, & Steger, 2004).

All of these numerical methods, together with the linearization, rely on indirect methods to solve the OCP of the growth model. According to Betts (2001), indirect methods can sometimes be difficult to apply and the necessary optimality conditions might be hard to determine explicitly. Economic models

are usually formulated with this in mind and sometimes simplicity is imposed onto them so that an analytical solution can be found, which is not always the case.

As opposed to indirect methods, the *direct methods* solve the OCP in a way that the OCP is first discretized and then optimized. More precisely, the direct approach uses the following procedure: the first step is to transcribe the infinite horizon problem into a finite dimensional problem, the second is to prove that this nonlinear programming problem (NLP) is equivalent representation of the original and finally an advanced NLP solver will be used to find the optimal trajectories.

Hypothesis that will be verified in this paper is the following: indirect methods for solving OCP provide better insight into the core of the optimization process in the theory of economic growth, but are sometimes very difficult to solve. The following will be verified as well: direct methods seem to be more efficient, but the process of discretization (transcription of the infinite horizon problem into a finite dimensional problem) of the OCP into a NLP, can be quite artificial and sometimes hard to prove.

The plan will be as follows. Section 2 gives a brief introduction on optimal control theory. The study of the neoclassical growth model (with and without technological progress) using the indirect approach will be surveyed in Section 3. Section 4 deals with the endogenous growth model by Romer but still using the indirect approach. The final section is devoted to the direct approach, which first discretizes and then optimizes the OCP. This method is used for the endogenous growth model by Lucas and Uzawa. The focus of this survey is on the main methods and mathematical techniques used for solving OCP in growth theory and not to give an extensive list to relevant work in the open literature.

2. On optimal control

Optimal control theory is a modern approach to dynamic optimization, with the initial work of Lev Pontryagin and his collaborators (Pontryagin, Boltyanskii, Gamkrelidze, & Mishchenko, 1962) as well as Richard Bellman (Bellman, 1957). In optimal control theory the aim is to find the optimal path for control variables without being constrained to interior solutions, even if it still relies on differentiability of functions that enter in the problem. The approach differs from calculus of variations in the fact that it uses control variables to optimize

the functional, and not state variables. Once the optimal path or value of the control variables is found, the solution to the state variables, or the optimal paths for the state variables are derived. There is a large choice in the literature on optimal control theory (Kirk (2004), Sethi and Thomson (2003), Bryson and Ho (1975)).

In order to solve optimal control problems, numerical methods that are used are divided into two major classes: direct and indirect methods (cf. Rao (2009), Sargent (2000)). In an indirect method, the calculus of variations is used to determine the first-order optimality conditions of the optimal control problem. Unlike ordinary calculus (where the objective is to determine points that optimize a function), the calculus of variations is the subject of determining functions that optimize a function of a function (also known as functional optimization). Applying the calculus of variations to the OCP leads to the first-order necessary conditions for the original optimal control problem. This approach will lead to a two-point (or even multiple-point) boundary value problem. Then, this problem will be solved in order to determine possible optimal trajectories and examine each of them to see if it is a local minimum, maximum, or a saddle point. From the locally optimizing solutions, the particular trajectory with the lowest cost is chosen. In a direct method, the state and/or control of the optimal control problem will be discretized in some manner in order to rewrite the problem to a nonlinear optimization problem or nonlinear programming problem (NLP). The NLP is then solved using well known optimization techniques (cf. (Gill, Murray, & Saunders, 2005), (Gill, Murray, Saunders & Wong, 2015), Betts (2001)).

Direct and indirect methods come from two different philosophies. On the one hand, the indirect approach solves the problem indirectly by converting the optimal control problem to a boundary value problem. Therefore, the optimal solution is indeed the solution of a system of differential equations that satisfies endpoint and/or interior point conditions. On the other hand, in the direct approach the optimal solution is found by rewriting an infinite optimization problem to a finite optimization problem. Even though these two approaches seem unrelated, they have much more in common. In particular, in recent years, researchers have discovered that the optimality conditions from many direct methods have a well-defined meaningful relationship. Therefore, it seems that these two classes of methods are merging as time goes by (cf. von Stryk & Bulirsch (1992)).

The study of dynamic growth models generally follows the indirect approach, which consists in applying Pontryagin's Maximum Principle or the dynamic programming[2], in order to obtain the first order necessary optimality conditions, together with the transversality condition. The Pontryagin's maximum principle and dynamic programming can be seen as equivalently alternative methods. Generally, Pontryagin's maximum principle will be used in continuous time and when there is no uncertainty and dynamic programming in discrete time and when there is uncertainty. Although it seems that the study of growth models in continuous time mainly uses Pontryagin's Maximum Principle, the dynamic programming approach could be used as well (cf. (Stokey, Lucas & Prescott, 1989), (Ljungqvist & Sargent, 2012)), such as for example, Fabbri and Gozzi (2008) apply for the endogenous growth model with vintage capital. Dynamic optimization techniques in growth models together with numerical methods are studied by authors such as Sanderson, Tarasyev & Usova (2011), Tarasyev & Watanabe (2001), Kryazhimskii & Watanabe (2004).

3. Neoclassical growth model

The description of the economic model relies on David Cass (Cass, 1965). The economy evolves over time through the interaction of the households and firms constituting it. There is a single produced commodity (Y), which can be consumed or accumulated as capital (K). The households interact with producers and engage in an intertemporal allocation exercise. It is assumed that there is a single aggregative or representative household and a single representative business firm in the economy. The household increases in size over time and its rate of growth is given by

$$L(t) = L_0 e^{nt} \tag{1}$$

where n is the exponential growth rate of the household size over time and $L(t)$ is the size of the household at t. In what follows, $L(t)$ will be referred to as population at t. The household decides about an optimal consumption path

---

[2] The departure point of dynamic programming method is the idea of embedding a given OCP into a family of optimal control problems, with the consequence that in solving the given problem, we are actually solving the entire family of problems. The core of dynamic programming is the Bellman's principle of optimality and Hamilton-Jacobi-Bellman (HJB) equation. This equation represents a necessary condition for the OCP and it is in general a partial differential equation.

$C(t)$, $t \in [0, +\infty)$, where $C(t)$ is the aggregate consumption enjoyed by $L(t)$. Optimality of the path is judged with reference to an intertemporal welfare or utility function of the form

$$U = \int_0^{+\infty} u(c)L(t)e^{-\rho t}dt = \int_0^{+\infty} u(c)L_0(t)e^{-(\rho-n)t}dt$$

$$= L_0(t)\int_0^{+\infty} u(c)e^{-(\rho-n)t}dt$$

which is a weighted sum of instantaneous utilities derived from c, the per-capita consumption C/L at t. The weights reflect two facts. First, e^{nt} shows that utilities from per-capita consumption receive exponentially higher weights with time to take account of the fact that the household size increases at the rate n. Secondly, $\rho > 0$ stands for the rate of time preference of the representative household. Utilities further down in time are valued less than utilities enjoyed earlier on. This leads to an exponentially decaying weight $e^{-\rho t}$ with the passage of time. Generally, ρ will be referred as the discount rate. In order to assume convergence, it will be assumed that $\rho - n > 0$. This is equivalent to suppose having one positive discount rate r, with $r = \rho - n$. If $L_0 = 1$, then the functional will reduce to

$$U = \int_0^{+\infty} u(c)e^{-rt}dt \ \ with \ r > 0$$

The social utility index function will be assumed to satisfy the following

$$u'(c) > 0 \qquad u''(c) < 0$$
$$\lim_{c \to 0} u'(c) = \infty \quad \lim_{c \to \infty} u'(c) = 0$$

In other words, utility is a strictly concave function of c, the marginal utility is unboundedly high for small values of c, while it is as close to zero as possible for large c. A representative firm has access to a technology for producing Y. This is represented by an aggregate production function

$$Y(t) = F\big(K(t), L(t)\big) \tag{2}$$

where $Y(t)$ stands for the flow of output at t and $K(t)$ and $L(t)$ for the flows of capital and labor services entering the production process at t. Note that the

use of the same notations for capital stock and services as well as for population size and labour services implies that the stock-flow ratios for both factors are assumed to be constant (normalized to unity). It will be assumed that function $F(K, L)$ is positively homogenous of degree 1 i.e. $F(\alpha K, \alpha L) = \alpha F(K, L)$, for any $\alpha, K, L > 0$. This implies that, at any unit of time, the production volume is proportional to the factors of production available at this moment. Such production function can be rewritten in per-capita terms and we can define $y = Y/L$ to be the average product of labor and $k = K/L$ to be the capital-labor ratio. The production function can also be expressed by

$$y = \phi(k) \quad with \quad \phi'(k) > 0 \quad and \quad \phi''(k) < 0 \quad for\ all \quad k > 0$$

It is assumed that $\lim_{k \to 0} \phi'(k) = \infty$ and $\lim_{k \to \infty} \phi'(k) = 0$. These conditions are called Inada conditions[3] and they ensure the existence of interior equilibrium. Indeed, they imply that the first units of capital and labor are highly productive and that when capital or labor are sufficiently big, their marginal products are close to zero.

The total output Y is allocated either to consumption C or gross investment I i.e. $Y = C + I$. Therefore, net investment $\dot{K}$ can be expressed as

$$\dot{K} = Y - C - \delta K$$

where δ is the depreciation rate. Dividing this equation by L:

$$\frac{\dot{K}}{L} = y - c - \delta k = \phi(k) - c - \delta K \tag{3}$$

Since

$$\dot{K} \stackrel{\text{def}}{=} \frac{dK}{dt} = k\frac{dL}{dt} + L\frac{dK}{dt}$$

$$= knL + L\dot{k} = L\left(kn + \dot{k}\right) \quad with \quad n = \frac{dL}{dt} \times \frac{1}{L}$$

---

[3] More precisely, Inada conditions require that the first partial derivatives of the production function $F(K, L)$ with respect to the variables K (resp. L) are positive and the second partial derivatives negative. Moreover, limit of the partial derivative of $F(K, L)$, as K (resp. L) approaches ∞, is 0, and as K (resp. L) approaches 0, is ∞ (cf. (Inada, 1963)).

Therefore, equation (3) becomes

$$\dot{k} = \phi(k) - c - (n + \delta)k \qquad (4)$$

This equation which includes only per-capita variables, describes how the capital-labor ratio varies over time. It is the fundamental differential equation of neoclassical growth theory. The OCP is given as follows:

$$Maximize \qquad U = \int_0^{+\infty} u\big(c(t)\big)e^{-rt}dt$$

$$subject \ to \ \ \dot{k} = \phi(k) - c - (n + \delta)k$$

$$k(0) = k_0$$

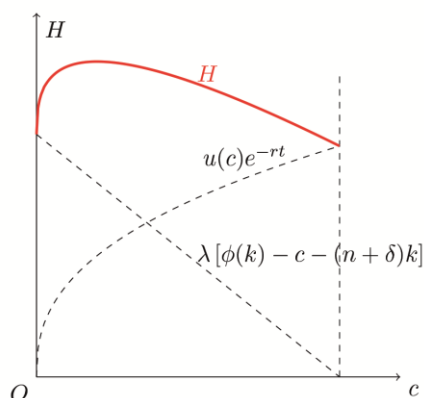$$and \ 0 \leq c(t) \leq \ \phi\big(k(t)\big) \qquad (5)$$

where k is the state variable and c the control variable.

3.1 The maximum principle

The Hamiltonian for this problem is

$$H = u(c)e^{-rt} + \lambda[\phi(k) - c - (n + \delta)k] \qquad (6)$$

and it is not linear in c. The function H and the two additive components $u(c)e^{-rt}$ and $\lambda[\phi(k) - c - (n + \delta)k]$ are represented in the following figure:

The maximum of H corresponds to the value of c inside the control region $[0, \phi(k)]$. By the necessary condition for the existence of a maximum

$$\frac{\partial H}{\partial c} = u'(c)e^{-rt} - \lambda = 0 \tag{7}$$

the following is obtained $u'(c) = \lambda e^{rt}$. The last equation states that optimally, the marginal utility of per-capita consumption should be equal to the shadow price of capital multiplied by the exponential term e^{rt}. The second partials

$$\frac{\partial^2 H}{\partial c^2} = u''(c)e^{-rt} < 0$$

since it is assumed that $u''(c) < 0$ and the function H is indeed maximized.

The maximum principle demands two equations of motion:

$$\dot{\lambda} = -\frac{\partial H}{\partial k} = -\lambda\big(\phi'(k) - (n + \delta)\big) \tag{8}$$

$$\dot{k} = \frac{\partial H}{\partial \lambda} = \phi(k) - c - (n + \delta)k \tag{9}$$

The first equation is a differential equation, and the second retrieves the constraint for the problem of optimal growth. The three equations (7), (8) and (9) should allow us to solve for the three variables c, λ and k. Without given

expressions of the functions $u(c)$ and $\phi(k)$, only qualitative analysis of the model can be done.

Since the maximum principle demands the differentiation of H and the discount factor further adds complexity to the derivatives, it is useful to define a new Hamiltonian without the discount factor. First, a new Lagrange multiplier $m = \lambda e^{rt}$ can be defined. The new Hamiltonian $H_c = He^{rt}$, called the current-value Hamiltonian is:

$$H_c = u(c) + m[\phi(k) - c - (n + \delta)k] \tag{10}$$

In this case, the maximum principle demands that

$$\frac{\partial H_c}{\partial c} = u'(c) - m = 0 \tag{11}$$

The function H_c is indeed maximized since the second partials $\partial^2 H_c / \partial c^2 = u''(c) < 0$. The equations of motion for the state variable k and the current-value multiplier m are:

$$\dot{k} = \frac{\partial H_c}{\partial m} = \phi(k) - c - (n + \delta)k \tag{12}$$

$$\dot{m} = -\frac{\partial H_c}{\partial k} + rm = -m[\phi'(k) - (n + \delta + r)] \tag{13}$$

3.2 Phase diagram analysis

A phase diagram analysis can be undertaken for the current-value maximum conditions (11), (12) and (13). The normal phase diagram should be in the km space, since the two equations of motion are in variables k and m. In order to construct such a diagram, the variable c should first be eliminated. But, since the condition (11) contains a function $u'(c)$ of c, it becomes easier to eliminate the variable m and to construct the phase diagram in kc space. First, condition (11) is differentiated, with respect to t: $\dot{m} = u''(c)\dot{c}$. By rearranging condition (13), it becomes

$$\dot{c} = -\frac{u'(c)}{u''(c)}[\phi'(k) - (n + \delta + r)] \tag{14}$$

which is a differential equation in the variable c. Therefore, this gives us the following system of differential equations:

$$\dot{k} = \phi(k) - c - (n + \delta)k$$

$$\dot{c} = -\frac{u'(c)}{u''(c)}[\phi'(k) - (n + \delta + r)] \qquad (15)$$

To construct the phase diagram, the curves $\dot{k} = 0$ and $\dot{c} = 0$ need to be plotted. Therefore, two equations are obtained:

$$c = \phi(k) - (n + \delta)k$$
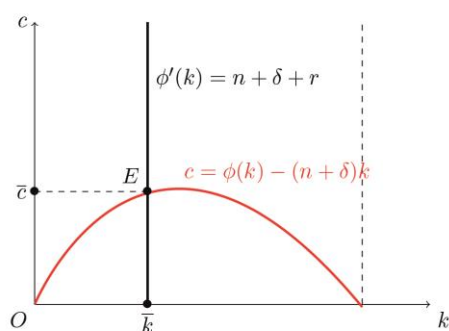
$$\phi'(k) = n + \delta + r$$

Figure 1: Curves $\dot{k} = 0$ and $\dot{c} = 0$

The first curve is the difference between the curve of the function $\phi(k)$ and the upward-sloping line $(n + \delta)k$. The second curve requires that the slope of the function $\phi(k)$ takes the value $n + \delta + r$. Since the function $\phi(k)$ is monotonic and increasing, there is only one point in which this condition is satisfied. Hence, the second curve must be plotted as a vertical straight line, with horizontal intercept $\bar{k}$. The intersection of the two curves determines the steady-state values of k and c. These values, denoted by $\bar{k}$ and $\bar{c}$ are known in literature as the modified-golden-rule values of capital-labor ratio and per-capita consumption.
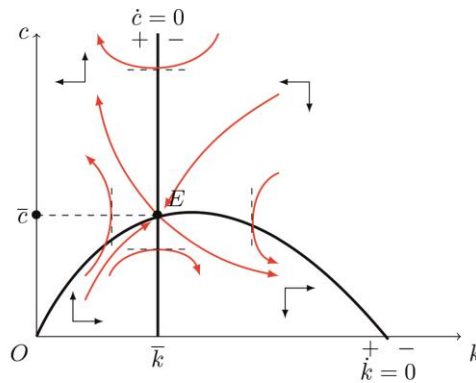
Figure 2: Phase diagram

If one looks at the pattern of the streamlines in Figure 2, it seems that the equilibrium $E(\overline{k}, \overline{c})$ is a saddle point. This can be verified by examining the characteristic roots of the linearization of the system of differential equations (15) of the model. In order to do it, one needs to form the Jacobian matrix from (15) and to evaluate it at the steady-state point E.

$$J_{E(\overline{k},\overline{c})} = \begin{pmatrix} \dfrac{\partial \dot{k}}{\partial k} & \dfrac{\partial \dot{k}}{\partial c} \\ \dfrac{\partial \dot{c}}{\partial k} & \dfrac{\partial \dot{c}}{\partial c} \end{pmatrix} = \begin{pmatrix} \phi'(\overline{k}) - (n+\delta) & -1 \\ -\dfrac{u'(\overline{c})}{u''(\overline{c})} \phi''(\overline{k}) & \left(-\dfrac{u'(\overline{c})}{u''(\overline{c})}\right)'_c [\phi'(\overline{k}) - (n+\delta+r)] \end{pmatrix}$$

$$= \begin{pmatrix} r & -1 \\ -\dfrac{u'(\overline{c})}{u''(\overline{c})} \phi''(\overline{k}) & 0 \end{pmatrix}$$

Since $r_1 r_2 = |J_E| = -\dfrac{u'(\overline{c})}{u''(\overline{c})} \phi''(\overline{k}) < 0$, two roots r_1 and r_2 have opposite signs and therefore the steady state is locally a saddle point.

3.3 Transversality conditions

Selecting a stable branch in the family of streamlines is almost the same as choosing a particular solution from a family of general solutions by finding a

particular constant. In order to do it, one will use the boundary conditions that can be: either a specific initial c_0 such that (k_0, c_0) is on the stable branch or an appropriate transversality condition. Since one has general functions and therefore does not have quantitative solutions of differential equations, it is not possible to describe a transversality condition. However, the phase-diagram analysis can be used in order to verify that the steady-state solution satisfies the expected transversality conditions. First transversality condition that the steady-state solution is expected to satisfy is

$$\lim_{t \to \infty} \lambda = 0 \qquad (16)$$

Indeed, the solution path for λ is $\lambda^* = u'(c)e^{-rt}$. It is clear that $\lim_{t \to \infty} e^{-rt} = 0$. On the other hand $\lim_{t \to \infty} u'(c)$ is finite, since it was assumed earlier that $\lim_{c \to 0} u'(c) = \infty$ and in this case c^* doos not tend to zero when $t \to \infty$. Another transversality condition in the solution is

$$\lim_{t \to \infty} H = 0 \qquad (17)$$

One knows from (6) that the solution path for H is

$$H^* = u(c^*)e^{-rt} + \lambda^*[\phi(k^*) - c^* - (n + \delta)k^*]$$

By (16), $\lim_{t \to \infty} \lambda^* = 0$ and $\lim_{t \to \infty} u(c^*)e^{-rt} = 0$, since $\lim_{t \to \infty} u(c^*)$ is finite and $\lim_{t \to \infty} e^{-rt} = 0$. The expression in the brackets $\phi(k^*) - c^* - (n + \delta)k^*$ is exactly $\dot{k}$ which is equal to zero in the steady state. Therefore, this transversality condition is also satisfied.

3.4 Exogenous growth model with technological progress

The neoclassical growth model that has been discussed in the previous section provides a steady state in which per-capita consumption c, stays constant at $\bar{c}$, with no further improvement in the average standard of living, which comes from the static nature of the production function $Y = F(K, L)$. Once technological progress is allowed, one can easily remove the cap on per-capita consumption. There are three types of ``neutral'' technological progresses (the term neutral is because it leaves a certain economic variable unaffected under certain condition): Hicks-neutral, Harrod-neutral and Solow-neutral. The production function with Harrod-neutral technological progress is

$$Y = F(K, A(t)L) \qquad (18)$$

where F is linearly homogenous in K and $A(t)L$. Harrod-neutral technological progress leaves the output-capital ratio (Y/K) unchanged at the same MPPK (marginal product of capital). In exogenous growth models, Harrod neutrality is frequently assumed since it is perfectly consistent with the notion of a steady state. If we define efficiency labor $\eta = AL$, then the production function (18) becomes

$$Y = F(K, \eta) \qquad (19)$$

If one now consider efficiency labor η to be the relevant labor-input variable, then (19) can be seen as a static production function. Using the homogeneity, one can rewrite:

$$y_\eta = \phi(k_\eta) \quad where \quad y_\eta = \frac{Y}{\eta} \quad and \quad k_\eta = \frac{K}{\eta}$$

This function is analogue to the production function $y = \phi(k)$ in the neoclassical growth model. Following the same steps as at the beginning of this section, while maintaining all the assumptions on ϕ, one can find the equation of motion for the new state variable k_η.
One obtains

$$\dot{k}_\eta = \phi(k_\eta) - c_\eta - (a + n + \delta)k_\eta \qquad (20)$$

$$where \quad c_\eta = \frac{C}{\eta}, \quad a = \frac{\dot{A}}{A} \quad and \quad n = \frac{\dot{L}}{L} \qquad (21)$$

Hence, one has the following OCP:

$$Maximize \qquad U = \int_0^{+\infty} u\left(c_\eta(t)\right) e^{-rt} dt$$

$$subject \;\; to \;\; \dot{k}_\eta = \phi(k_\eta) - c_\eta - (a + n + \delta)k_\eta$$

$$k_\eta(0) = k_{\eta 0}$$

$$and \;\; 0 \leq c_\eta \leq \phi(k_\eta) \qquad (22)$$

The current-value Hamiltonian in the new problem is

$$H'_c = u(c_\eta) + m'[\phi(k_\eta) - c_\eta - (a + n + \delta)k_\eta] \qquad (23)$$

where m' is the new current-value costate variable. By analogy, the maximum principle demands for the following conditions:

$$m' = u'(c_\eta) \qquad (24)$$

$$\dot{k}_\eta = \phi(k_\eta) - c_\eta - (a + n + \delta)k_\eta \qquad (25)$$

$$\dot{m}' = -m'[\phi'(k_\eta) - (a + n + \delta + r)] \qquad (26)$$

By eliminating the variable m', the following system of differential equations is obtained:

$$\dot{k}_\eta = \phi(k_\eta) - c_\eta - (a + n + \delta)k_\eta$$

$$\dot{c}_\eta = -\frac{u'(c_\eta)}{u''(c_\eta)}[\phi'(k_\eta) - (a + n + \delta + r)]$$

The new phase diagram is constructed:

$$\dot{k}_\eta = 0 \quad \Leftrightarrow \quad c_\eta = \phi(k_\eta) - (a + n + \delta)k_\eta$$
$$\dot{c}_\eta = 0 \quad \Leftrightarrow \quad \phi'(k_\eta) = a + n + \delta + r$$

The new phase diagram is qualitatively the same as the one in the neoclassical growth model. However, there is one major difference which concerns the new steady state. Before, the per-capita consumption $c = C/L$ was constant in a steady state. Now, one has that $c_\eta = C/\eta = C/AL$ is constant. Therefore, per-capita consumption $c = c_\eta A$ will rise over time, as long as A increases as a result of technological progress.

4. Endogenous growth model by Romer: indirect approach

Even though the exogenous technological progress remains quite simple, it does not explain the origin of progress. In order to do it, it is necessary to endogenize the technological progress. Endogenous growth models were conceived as an attempt to overcome this theoretical nuance and to give a consistent report to what causes economies to keep on growing (Romer (1990), Lucas (1988), Uzawa (1965), Jones and Manuelli (1990)).

Knowledge in the Romer model (Romer, 1990), can be classified into two components: human capital and technology. Human capital is person-specific and it is a ``rival good″ in the sense that its use by one firm excludes its use by another. Technology is available to the public and it is a ``nonrival good″ in the sense that its use by one firm does not limit its use by others. Human capital and technology are created by conscious action. In order to reduce the number of state variables, human capital is assumed to be fixed and inelastically supplied. Human capital will be denoted by S, and by S_0 its fixed total. Since S can be used for the production of the final good Y or for the improvement of technology A, one has:

$$S_0 = S_Y + S_A$$

On the other hand, technology A is not fixed and can be created by engaging human capital S_A in research and applying the existing technology A:

$$\dot{A} = \sigma S_A A \tag{27}$$

where σ denotes the research success parameter. One sees that $\dot{A}/A = \sigma S_A > 0$ if σ and S_A are positive. Therefore, technology can grow without bound. In the production of the final good Y variables L and K are inputs with human capital S_Y and technology A. The production function is assumed to be of Cobb-Douglas type:

$$Y = S_y^\alpha L_0^\beta A \bar{x}^{1-\alpha-\beta} \tag{28}$$

where L_0 denotes fixed and inelastically supplied amount of ordinary labor and x is the common level of use of ``design″. Since the amount of capital actually used will be

$$K = \gamma A \bar{x}$$

Therefore, the equation (28) becomes:

$$Y = S_y^\alpha L_0^\beta A \left(\frac{K}{\gamma A} \right)^{1-\alpha-\beta}$$

$$= \left(S_y A \right)^\alpha (L_0 A)^\beta K^{1-\alpha-\beta} \gamma^{\alpha+\beta-1} \tag{29}$$

One can see from the last expression that technology can be considered as human capital increasing $(S_y A)$ and as labor-increasing $(L_0 A)$, but it is detached from K. It is in a certain way endogenously characterized by Harrod-neutrality.

Since the net investment is just the output that is not consumed, by substituting:

$$\dot{K} = Y - C$$

$$= \gamma^{\alpha+\beta-1} A^{\alpha+\beta} (S_0 - S_A)^\alpha L_0^\beta K^{1-\alpha-\beta} - C \qquad (30)$$

One can now consider an OCP with two state variables A and K, and two control variables C and S_A, while the equations of motion are (27) and (30). Assuming the constant elasticity of substitution (CES) utility function:

$$u(C) = \frac{C^{1-\theta}}{1-\theta} \quad with \quad 0 < \theta < 1$$

the following OCP can be defined:

$$Maximize \qquad \int_0^{+\infty} \frac{C^{1-\theta}}{1-\theta} e^{-\rho t} dt$$

$$subject \ to \ \ \dot{A} = \sigma S_A A$$

$$\dot{K} = \gamma^{\alpha+\beta-1} A^{\alpha+\beta} (S_0 - S_A)^\alpha L_0^\beta K^{1-\alpha-\beta} - C$$

$$and \ A(0) = A_0 \quad K(0) = K_0 \qquad (31)$$

In order to simplify the expression (30), one defines $\Delta = \gamma^{\alpha+\beta-1} A^{\alpha+\beta} (S_0 - S_A)^\alpha L_0^\beta K^{1-\alpha-\beta}$. Then the current-value Hamiltonian is

$$H_c = \frac{C^{1-\theta}}{1-\theta} + \lambda_A (\sigma S_A A) + \lambda_K (\Delta - C) \qquad (32)$$

where λ_A and λ_K denote the shadow prices of A and K, respectively. By the necessary condition for the existence of a maximum

$$\frac{\partial H_c}{\partial C} = C^{-\theta} - \lambda_K = 0 \qquad \Leftrightarrow \qquad \lambda_K = C^{-\theta} \qquad (33)$$

$$\frac{\partial H_c}{\partial S_A} = \lambda_A \sigma A - \lambda_K \alpha (S_0 - S_A)^{-1} \Delta = 0 \quad \Leftrightarrow \quad \Delta = \frac{\lambda_A \sigma A}{\lambda_K \alpha}(S_0 - S_A) \qquad (34)$$

The maximum principle demands in addition to the two equations of motion $\dot{A}$ and $\dot{K}$, two equations of motion for the costate variables:

$$\dot{\lambda}_A = -\frac{\partial H_c}{\partial A} + \rho \lambda_A = -\lambda_A \sigma S_A - \lambda_K (\alpha + \beta) A^{-1} \Delta + \rho \lambda_A$$

$$\dot{\lambda}_K = -\frac{\partial H_c}{\partial K} + \rho \lambda_K = -\lambda_K (1 - \alpha - \beta) K^{-1} \Delta + \rho \lambda_K \qquad (35)$$

Having four differential equations, it is not possible to construct a phase diagram and it is difficult to solve the system explicitly. However, some questions can be answered in the steady state. Variables Y, K, A and C grow at the same rate. Therefore, by (27), one has:

$$\frac{\dot{A}}{A} = \frac{\dot{Y}}{Y} = \frac{\dot{K}}{K} = \frac{\dot{C}}{C} = \sigma S_A \qquad (36)$$

Since the growth rate should be expressed only in terms of parameters, it is necessary to find the expression for S_A. Indeed, the expressions for $\dot{\lambda}_A / \lambda_A$ and $\dot{\lambda}_K / \lambda_K$ can be found from (35) and then use the fact that in the steady state: $\dot{\lambda}_A / \lambda_A = \dot{\lambda}_K / \lambda_K$. Therefore, S_A reaches in the steady state the constant value:

$$S_A = \frac{\sigma(\alpha+\beta)S_0 - \alpha\rho}{\sigma(\alpha\theta+\beta)} \qquad (37)$$

The growth rate is

$$\frac{\dot{A}}{A} = \frac{\dot{Y}}{Y} = \frac{\dot{K}}{K} = \frac{\dot{C}}{C} = \frac{\sigma(\alpha+\beta)S_0 - \alpha\rho}{\alpha\theta+\beta} \qquad (38)$$

The expression (38) tells us how the various parameters affect the growth rate. One can see that the research success parameter σ and the human capital S_0 have a positive effect on the growth rate, and that the discount rate ρ has the negative effect.

5. Endogenous growth model by Lucas and Uzawa: direct approach

The Uzawa-Lucas model introduces human capital h, as another productive input of the economy that is produced by a different technology than that of physical capital K. Also, labor L can either be employed in final output production, μ, with the rest $1 - \mu$ that is dedicated to formal education. This model exposes steady-state growth, in the sense that consumption and capital (physical and human) is unbounded. The introduction of human capital and specialized labor adds another state and control variable respectively to the system, and adds complexity as well. In fact, the transition process of this model is still vague, since the indirect methods for solving the underlying OCP give origin to rigid ordinary differential equations, which gives an extra difficulty, as Trimborn et al. (2004) noticed. This model uses the framework specified in Section 3. Households try to maximize their utility by consuming according to a standard CES function $u(C)$. One has the following production function:

$$Y = AK^{\alpha}(uhL)^{1-\alpha} \tag{39}$$

Physical capital K and human capital h follow the laws of motion:

$$\dot{K} = Y - C - \delta K \tag{40}$$

$$\dot{h} = B(1-u)h - \delta_h h \tag{41}$$

where B is a constant reflecting productivity of quality adjusted effort in education and δ_h $(0 \leq \delta_h < B)$ is the depreciation rate of human capital, which is set $\delta_h = 0$.

Using per-capita variables k, y and c and without population growth $(n = 0)$, the following OCP can be defined:

$$Maximize \qquad \int_0^{+\infty} \frac{C^{1-\theta}}{1-\theta} e^{-\rho t} dt$$

$$subject \ \ to \ \ \dot{k} = AK^{\alpha}(uh)^{1-\alpha} - c - \delta k$$

$$\dot{h} = B(1-u)h$$

$$and \ \ k(0) = k_0 \quad h(0) = h_0 \tag{42}$$

The outline that will be presented results from the work of Fontes (2001) and follows the approach used by Lopes et al. (2013). The first step is to transcribe the infinite horizon problem into a finite dimensional problem. The second is to prove that this nonlinear programming problem (NLP) is equivalent to the original one. Finally, an advanced NLP solver could be used in order to find the optimal trajectories. As opposed to indirect methods, the OCP is first discretized and then optimized.

The first step is to present the following theorem that enables us to discretize the infinite horizon problem.

Theorem 1. (Lopes et al., 2013)

Consider the following generic optimal control problem:

$$P_\infty: \quad Minimize \quad \int_0^{+\infty} L\big(x(t), u(t)\big) dt$$

$$subject \ \ to \ \ \dot{x} = f(x, u)$$

$$x(0) = x_0$$

$$and \ \ x(t) \in \Gamma(t) \ \ u(t) \in \Omega(t) \tag{43}$$

where we assume that there is a finite solution. Furthermore, we assume that after some time T, the state is within some invariant set S (i.e. $x(t) \in S$, $S \subset \Gamma(t)$, $\forall t \geq T$) for which the problem still has a finite solution.

Then there exists a terminal cost function W, such that the optimal control problem is equivalent to the finite horizon problem:

$$P_T: \quad Minimize \quad \int_0^T L\big(x(t), u(t)\big) dt + W\big(x(t)\big)$$

$$subject \ \ to \ \ \dot{x} = f(x, u)$$

$$x(0) = x_0$$

$$and \ \ x(t) \in \Gamma(t) \ \ u(t) \in \Omega(t) \ \ x(T) \in S \tag{44}$$

Proof. Consider problem P_∞ with $\Gamma(t) = S$ for all $t \geq T$. Then the value for P_∞ is

$$V(0, x_0) = min\left\{\int_0^T L(x, u(t))dt + V(t, x(t))\right\}$$

Now, one can define

$$W(x(t)) = V(T, x(T)) = \min_{x \in S}\left\{\int_T^{+\infty} L(x, u)dt\right\}$$

and discretize P_∞ into P_T.

Theorem 1 can be applied only when the set S is such that the characterization of the solution to the problem is possible

$$W(x(t)) = \min_{x \in S}\left\{\int_T^{+\infty} L(x(t), u(t))dt\right\}$$

and therefore $W(x)$ can explicitly be computed for any $x \in S$.

The following procedure is applied:

1. Transcription of the infinite-horizon problem into an equivalent finite-horizon problem by applying Theorem 1;

2. Adding the necessary boundary conditions that ensure that the set S is invariant (and so $W(x)$ exists for any $x \in S$);

3. Using a NLP solver to determine the trajectories of the control and state variables.

It is known that in balanced growth we have:

$$\frac{\dot{c}}{c} = \frac{\dot{y}}{y} = \frac{\dot{k}}{k} = \frac{\dot{h}}{h} = \gamma \qquad (45)$$

If one defines

$$w = \frac{k}{h} \quad and \quad \chi = \frac{c}{k}$$

and it is known that in the steady state, for $t \geq T$, one has $\dot{w} = 0$ and $\dot{\chi} = 0$ (Barro et al., 2003). Define the control functions

$$c(t) = \chi k(t) \quad and \quad k(t) = wh(t) \tag{46}$$

such that χ and w are equal to a given positive constant.

One has

$$\frac{\dot{k}}{k} = Ak^{\alpha-1}(uh)^{1-\alpha} - \chi - \delta = \left(\frac{uh}{Ak}\right)^{1-\alpha} - \chi - \delta = \left(\frac{u}{Aw}\right)^{1-\alpha} - \chi - \delta$$

but also

$$\frac{\dot{k}}{k} = \frac{w\dot{h}}{wh} = B(1-u) = \gamma$$

Therefore, equation (46) is obtained with $u(t) = u$ constant and one has:

$$\left(\frac{u}{Aw}\right)^{1-\alpha} - \chi - \delta = B(1-u)$$

Moreover, one has

$$\frac{\dot{c}}{c} = \frac{\dot{k}}{k} = \frac{\dot{h}}{h} = B(1-u)$$

which implies that $\dot{c} = \gamma c$ in a balanced growth path. At the time $t = T$ consumption c will continue to grow at rate γ

$$c(t) = c(T)e^{\gamma(t-T)}, \quad t \in [T, +\infty)$$

which enables us to compute W. Utility will be bounded as $\rho > \gamma(1-\theta)$ so that $U(\cdot, \infty) = 0$. Then the boundary cost is given by

$$W = -\int_T^{+\infty} \frac{\left(c(T)e^{\gamma(t-T)}\right)^{1-\theta}}{1-\theta} e^{-\rho t}dt$$

Finding the integral

$$W = -\frac{e^{\gamma(1-\theta)-\rho}}{\rho - \gamma(1-\theta)} \cdot \frac{\left(c(T)e^{-\gamma T}\right)^{1-\theta}}{1-\theta} \tag{47}$$

The boundary condition will be

$$S = \left\{ (k, h) \in R^2 : \frac{\dot{k}}{k} = \frac{\dot{h}}{h} = 0 \right\} \quad s.t. \quad k(T), h(T) \in S \qquad (48)$$

The system defined by (42) together with the boundary cost W (47) and the boundary condition S (48) can now be solved numerically using a NLP solver. Lopes et al. (2013) use the software ICLOCS that solves OCP with general path and boundary constraints and free or fixed final time (actually this software uses another software called IPOPT to solve the transformed NLP problem).

More generally, there are numerous and powerful softwares (NLP solvers) that are used to find numerical solutions of OCP. Some of them can be named: SNOPT, ICLOCS, DIRCOL, SOCS, OTIS, GESOP/ASTOS, DITAN and PyGMO/PyKEP.

These numerical results should be evaluated in order to measure the accuracy of numerical methods. There are different studies of this procedure of evaluation (Aruoba et al. (2006), Ambler and Pelgrin (2006) and Heer and Maubner (2008)).

It is important to notice, that using the direct approach, optimal trajectories can be determined numerically and without demanding the linearization of the differential equations. This can avoid some problems that can occur by a change of base or any other manipulation. It also allows the study of the transition process when it is not necessary for the system to depart from a steady state or to be at a steady state. Furthermore, it is a powerful tool to study some complex phenomena like anticipated or multiple, sequential shocks. Mainly, to study such shocks using indirect methods, authors like Trimborn (2007) suggest a reformulation of the optimization problem, which consists in decomposing the functional form of the objective function from $f^{(1)}$ to $f^{(2)}$ and the state equations from $g^{(1)}$ to $g^{(2)}$ at time $\tilde{t}$ when shock occurs. The necessary optimality conditions would have to be augmented with the conditions derived from the interior boundary condition. Moreover, the adjoint variable functions introduced by the Maximum Principle have to satisfy the condition of continuity, also known as the Weierstrass-Erdmann corner condition. Using the direct approach, it is not required to reformulate the model. Indeed, one does not need to determine or even know any necessary optimality conditions, which can be particularly useful for problems whose adjoint functions are hard to find.

Conclusion

The indirect and direct methods for solving OCP can be used in both, exogenous and endogenous models of growth. The indirect methods for solving OCP provide better insight into the core of the optimization process in the theory of economic growth, but are sometimes very difficult to solve. In particular, the linearization of the dynamic system of differential equations can become very difficult. As a result, simplicity of the model is often required, so that the analytical solutions could be found. On the other side, direct methods seem to be faster and more efficient, even if they do not follow the logical, standard procedure of optimization. Moreover, they allow the study of the transitional dynamics of models that are not at their steady-state, and also to study expected and unexpected shocks. However, the process of discretization (transcription of the infinite horizon problem into a finite dimensional problem) of the OCP into a NLP, can become quite artificial and sometimes hard to prove. In general, the choice of a method for solving OCP will mainly depend on two major points: the importance of the insight of the optimization process and the complexity of the given OCP.

References

Acemoglu, D. (2007). *Introduction to modern economic growth (Levine's Bibliography)*. UCLA Department of Economics.

Aghion, P., & Howitt, P. (1998). *Endogenous growth theories*. Cambridge, MA: MIT Press.

Arrow, K.J. (1968). *Applications of control theory to economic growth*. (pp. 85-119). Providence: American Mathematical Society.

Arrow, K.J., & Kurz, M. (1970). *Public investment, the rate of return and optimal fiscal policy*. Baltimore - London: Johns Hopkins Press.

Aseev, S. (2009). *Infinite-horizon optimal control with applications in growth theory*. lecture notes.

Atolia, M., Chatterjee, S., & Turnovsky, S.J. (2008). *How Misleading is Linearization?, Evaluating the Dynamics of the Neoclassical Growth Model*. Florida: Department of Economics, State University. Working Papers.

Barro, R.J., & Sala-i, M.X. (2003). *Economic Growth, 2nd ed*. The MIT Press. Vol. 1.

Bellman, R. (1957). *Dynamic programming, 1st ed*. Princeton, NJ, USA: Princeton University Press.

Betts, J.T. (2001). *Practical methods for optimal control using nonlinear programming*. Philadelphia: SIAM Press - Princeton University Press.

Brunner, M., & Strulik, H. (2002). Solution of perfect foresight saddlepoint problems: A simple method and applications. *Journal of Economic Dynamics and Control,* 26(5), 737-753.

Cass, D. (1965). Optimum growth in an aggregative model of capital accumulation. *Review of Economic Studies,* 32(3), 233-240.

Chiang, A.C. (1992). *Elements of dynamic optimization.* New York: McGraw-Hill..

Dixit, A.K. (1990). *Optimization in Economic Theory.* Oxford University Press.

Dorfman, R. (1969). An Economic Interpretation of Optimal Control Theory. *American Economic Review,* 59(5), 817-831.

Fabbri, G., & Gozzi, F. (2008). Solving optimal growth models with vintage capital: The dynamic programming approach. *Journal of Economic Theory,* 143(1), 331-373.

Fontes, F. (2001). A general framework to design stabilizing nonlinear model predictive controllers. *Systems & Control Letters,*

Gill, P.E., Murray, W., & Saunders, M.A. (2005). SNOPT: An SQP algorithm for large-scale constrained optimization. *SIAM Rev.,* 47, 99-131.

Gill, P.E., Murray, W., Saunders, M.A., & Wong, E. (2015). User's guide for SNOPT 7. 5: Software for large-scale nonlinear programming. *Center for Computational Mathematics Report,* 15-3, CCoM, La Jolla, CA: Department of Mathematics, University of California, San Diego.

Inada, K.I. (1963). On a two-sector model of economic growth: Comments and a generalization. *Review of Economic Studies,* 30(2), 119-127.

Intriligator, M.D. (1971). *Mathematical optimization and economic theory.* Englewood Clis, N. J.: Prentice Hall..

Judd, K.L. (1992). Projection methods for solving aggregate growth models. *Journal of Economic Theory,* 58(2), 410-452.

Judd, K.L. (1998). *Numerical Methods in Economics.* The MIT Press. Vol. 1.

Kamien, M.I., & Schwartz, N.L. (1991). *Dynamic Optimization: The Calculus of Variations and Optimal Control in Economics and Management, 2nd ed.* Elsevier Science.

Koopmans, T. (1963). On the concept of optimal economic growth. *Cowles Foundation Discussion Papers,* 163, Cowles Foundation for Research in Economics, Yale University.

Ljungqvist, L., & Sargent, T.J. (2012). *Recursive Macroeconomic Theory, 3rd ed.* The MIT Press. Vol. 1.

Lopes, M.A., Fontes, F.A.C.C., & Fontes, D.A.C.C. (2013). Optimal Control of Infinite-Horizon Growth Models-A direct approach. *FEP Working Papers,* Universidade do Porto, Faculdade de Economia do Porto.

Lucas, R.J. (1988). On the mechanics of economic development. *Journal of Monetary Economics,* 22(1), 3-42.

Mercenier, J., & Michel, P. (1994). Discrete-Time Finite Horizon Appromixation of Infinite Horizon Optimization Problems with Steady-State Invariance. *Econometrica,* 62(3), 635-656.

Mulligan, C.B., & Sala-i Martin, X. (1991). A Note on the Time-Elimination Method For Solving Recursive Dynamic Economic Models. *NBER Technical Working Papers,* 0116, National Bureau of Economic Research, Inc..

Pontryagin, L.S., Boltyanskii, V.G., Gamkrelidze, R.V., & Mishchenko, E.F. (1962). *The mathematical theory of optimal processes.* New York: Interscience.

Ramsey, F. (1928). A mathematical theory of saving. *Economic Journal,* 38, 543-559.

Rao, A.V. (2009). A survey of numerical methods for optimal control. *Advances in the Astronautical Sciences,* 135(1), 497-528.

Romer, P.M. (1990). Endogenous Technological Change. *Journal of Political Economy,* 98(5), 71-102.

Romer, P.M. (1994). The origins of endogenous growth. *The Journal of Economic Perspectives,* 8(1), 3-22.

Sargent, R. (2000). Optimal control. *Journal of Computational and Applied Mathematics,* 124(12), 361-371.

Solow, R. (1956). A contribution to the theory of economic growth. *The Quarterly Journal of Economics,* 70(1), 65-94.

Stokey, N., Lucas, R.E., & Prescott, E.C. (1989). *Recursive methods in economic dynamics.* Cambridge, Mass: Harvard University Press.

Trimborn, T., Koch, K.J., & Steger, T.M. (2004). Multi-dimensional transitional dynamics: A simple numerical procedure. *CER-ETH Economics working paper series,*

Uzawa, H. (1965). Optimum technical change in an aggregative model of economic growth. *International Economic Review,* 6(1), 18-31.

von Stryk, O., & Bulirsch, R. (1992). Direct and indirect methods for trajectory optimization. *Ann. Oper. Res.,* 37(1-4), 357-373.

Wolman, A.L., & Couper, E.A. (2003). Potential consequences of linear approximation in economics. *Economic Quarterly,* 51-67.