

Dynamic Programming under Certainty

Sergio Feijoo-Moreira*

(based on Matthias Kredler's lectures)

Universidad Carlos III de Madrid

This version: February 11, 2020

[Latest version](#)

Abstract

These are notes that I took from the course Macroeconomics II at UC3M, taught by Matthias Kredler during the Spring semester of 2016. These notes closely follow [Stokey et al. \(1989\)](#). Typos and errors are possible, and are my sole responsibility and not that of the instructor.¹

Contents

| | | |
|----------|---|-----------|
| 1 | Finite Horizon Problems | 2 |
| 1.1 | Introduction | 2 |
| 1.2 | The Lagrangian approach | 3 |
| 1.3 | The Dynamic Programming approach | 5 |
| 1.3.1 | Example of Dynamic Programming: McCall Search Model | 11 |
| 2 | Infinite Horizon Problems | 14 |
| 2.1 | One-Sector Neo-Classical Growth Model | 14 |
| 2.1.1 | The Lagrangian approach and transversality conditions | 16 |
| 2.1.2 | The Dynamic Programming approach | 19 |

*Address: Universidad Carlos III de Madrid. Department of Economics, Calle Madrid 126, 28903 Getafe, Spain. E-mail: sfeijoo@eco.uc3m.es. Web: <https://sergiofeijoo.github.io>.

¹This also applies to Fabrizio Leone, very good economist, and even much better friend, who has also contributed to some parts of these notes.

| | | |
|-------|--|-----------|
| 2.2 | Mathematical Part: Dynamic Programming | 21 |
| 2.2.1 | A function space for V | 22 |
| 2.2.2 | T is a Contraction Mapping | 24 |
| 2.2.3 | T is a Self-Mapping | 26 |
| 2.3 | Principle of Optimality | 27 |
| 2.4 | Bellman equations | 34 |
| | References | 39 |
| | A Appendix: Envelope Theorem | 40 |

1 Finite Horizon Problems

1.1 Introduction

We start with the standard life-cycle consumption-savings problem. There is a unique agent who is born with some wealth, and in each period can decide to consume some of her wealth or save it to consume it in the future. If she decides to save it, we assume that there exists a risk-free mechanism to save assets at a gross return R . We also assume that this consumer receives an exogenous sequence of earnings in each period. Besides, the consumer may be borrowing constrained in some periods, according to an exogenous sequence of borrowing limits. The main components of this model are:

- Time is discrete and finite: $t \in \{0, 1, \dots, T\}$.
- $\{\omega_t\}_{t=0}^T$: exogenous deterministic sequence of earnings,
- c_t : consumption at time t ,
- a_t : assets (or wealth) coming into period t ,
 - $a_0 \geq 0$ given,
 - $a_{t+1} \geq \underline{a}_{t+1}$ for $t = 0, 1, \dots, T - 1$, where $\{\underline{a}_t\}_{t=1}^{T+1}$, $\underline{a}_t \leq 0$, $\forall t$, is an exogenous deterministic sequence of borrowing limits. Besides $\underline{a}_{T+1} = 0$, thus $a_{T+1} \geq 0$ implies that the agents must die with zero debt.
 - Savings yields gross return R from period t to $t + 1$.
- Budget constraint at time t :

$$a_{t+1} \leq R(a_t + \omega_t - c_t), \quad \text{for } t = 0, \dots, T,$$

or, equivalently,

$$\frac{a_{t+1}}{R} \leq a_t + \omega_t - c_t, \quad \text{for } t = 0, \dots, T,$$

- Preferences:

$$\sum_{t=0}^T \beta^t u(c_t),$$

where

- $\beta > 0$ is the discount rate,
- the instantaneous utility function, $u_t(\cdot)$, is
 - * twice continuously differentiable,
 - * $u'(c) > 0$, $\forall c$, i.e. it is strictly increasing in consumption,
 - * $u''(c) < 0$, $\forall c$, i.e. it is strictly concave,
 - * and satisfies the Inada condition

$$\lim_{c \rightarrow 0} u'(c) = \infty.$$

We are interested in solving the following planner's problem²

$$\begin{aligned} \max_{\{c_t, a_{t+1}\}_{t=0}^T} \quad & \sum_{t=0}^T \beta^t u(c_t) \\ \text{s.t.} \quad & a_t + \omega_t - c_t - \frac{a_{t+1}}{R} \geq 0, \quad t = 0, \dots, T, \\ & a_{t+1} - \underline{a}_{t+1} \geq 0, \quad t = 1, \dots, T, \\ & c_t \geq 0, \quad t = 0, \dots, T, \\ & a_0, R, \{\underline{a}_t\}_{t=1}^{T+1} \text{ and } \{\omega_t\}_{t=0}^T \text{ given.} \end{aligned} \tag{1}$$

1.2 The Lagrangian approach

The traditional approach consists on solving the constrained maximization problem, by setting up the following Lagrangian programme

$$\mathcal{L} \left(\{c_t, a_{t+1}, \lambda_t, \mu_t\}_{t=0}^T \right) = \sum_{t=0}^T \left[\beta^t u(c_t) + \lambda_t \left(a_t + \omega_t - c_t - \frac{a_{t+1}}{R} \right) + \mu_t (a_{t+1} - \underline{a}_{t+1}) \right]$$

The first order (necessary) conditions are given by

$$\frac{\partial \mathcal{L}(\cdot)}{\partial c_t} = \beta^t u'(c_t) - \lambda_t \stackrel{!}{=} 0, \quad \text{for } t = 0, 1, \dots, T, \tag{2}$$

$$\frac{\partial \mathcal{L}(\cdot)}{\partial a_{t+1}} = -\frac{\lambda_t}{R} + \mu_t + \lambda_{t+1} \stackrel{!}{=} 0, \quad \text{for } t = 0, 1, \dots, T-1, \tag{3}$$

$$\frac{\partial \mathcal{L}(\cdot)}{\partial a_{T+1}} = -\frac{\lambda_T}{R} + \mu_T \stackrel{!}{=} 0. \tag{4}$$

²Note $c_t > 0$ (by Inada), thus the third constraint will never bind.

From (2) we obtain

$$\lambda_t = \beta^t u'(c_t), \quad (5)$$

and since marginal utility is always positive (by assumption), then we have $\lambda_t > 0$ for $t = 0, \dots, T$. As a consequence, in equilibrium, the full sequence of budget constraints hold with equality. Moreover, (4) and (5) also imply

$$\mu_T = \frac{\lambda_T}{R} = \frac{\beta^T}{R} u'(c_T) > 0,$$

thus the no-borrowing constraint binds at $a_{T+1}^* = \underline{a}_{T+1} = 0$, which implies that the agents won't save in the last period.

Finally, rewriting (3) yields

$$\lambda_t = R(\lambda_{t+1} + \mu_t), \quad (6)$$

where substituting (5) and re-arranging we obtain

$$u'(c_t) = R\beta u'(c_{t+1}) + R\beta^{-t}\mu_t, \text{ for } t = 0, 1, \dots, T-1. \quad (7)$$

Therefore we have:

- If $\mu_t = 0$ for some t , then the agent is not constrained by the borrowing limit, i.e. $a_{t+1}^* > \underline{a}_{t+1}$, and thus we obtain the usual (consumption) Euler equation given by

$$u'(c_t) = R\beta u'(c_{t+1}). \quad (8)$$

This equation tells us that if the consumer is not constrained by the borrowing limit, to be optimizing the marginal cost of saving (i.e., not eating today one unit of the consumption good) given by the LHS of the previous equation and measured by the marginal utility of consuming that unit today, must be equal to the marginal benefit of saving (i.e., eating tomorrow R units of the consumption good) given by the RHS of the previous equation and measured by the discounted marginal utility of consuming R tomorrow. Re-arranging this expression we can obtain a more 'micro' interpretation, given by

$$\frac{u'(c_t)}{\beta u'(c_{t+1})} = R,$$

which tells us that in an interior solution, the *MRS* between c_t and c_{t+1} must be equal to their price ratio (relative price between c_t and c_{t+1}).

- If $\mu_t > 0$ for some t , then the agent is constrained by the borrowing limit, i.e. $a_{t+1}^* = \underline{a}_{t+1}$, and thus we have

$$u'(c_t) > R\beta u'(c_{t+1}),$$

which tells us that the consumer would like to increase c_t even further, but she can't as she is borrowing constrained.

Special case: No borrowing limit Assume that $a_t = -\infty$ for $t = 1, \dots, T$, and $a_{T+1} = 0$. In this case, the borrowing constraint does not bind at any t , which implies $\mu_t = 0$, for $t = 0, 1, \dots, T - 1$. Given that the budget constraint holds with equality, substituting it into (8) yields

$$u' \left(a_t^* + w_t - \frac{a_{t+1}^*}{R} \right) = u' \left(a_{t+1}^* + w_{t+1} - \frac{a_{t+2}^*}{R} \right), \text{ for } t = 0, \dots, T - 1,$$

which is a second order difference equation for $\{a_t^*\}_{t=0}^{T+1}$ with boundary conditions $a_0^* = a_0$ and $a_{T+1}^* = 0$.

Digression: Power utility Consider the life-cycle consumption-savings model without borrowing limits and instantaneous utility function

$$u(c_t) = \frac{c_t^{1-\gamma}}{1-\gamma}.$$

Then, from (8) we obtain the Euler equation

$$c_t^{-\gamma} = \beta R c_{t+1}^{-\gamma} \implies \left(\frac{c_t}{c_{t+1}} \right)^{-\gamma} = \beta R,$$

Taking logs we obtain

$$\ln \left(\frac{c_{t+1}}{c_t} \right) = \frac{\ln \beta + \ln R}{\gamma},$$

where $\ln c_{t+1} - \ln c_t = \ln \left(\frac{c_{t+1}}{c_t} \right)$ is the percentage growth rate of consumption. The inter-temporal rate of substitution (IRS) is defined as

$$IRS = \frac{d \ln \left(\frac{c_{t+1}}{c_t} \right)}{d \ln R} = \frac{1}{\gamma}.$$

1.3 The Dynamic Programming approach

The aim of Dynamic Programming is decomposing a complex problem into many sub-problems and then solving them with a recursive algorithm. In economics, it is particularly useful to break the returns of an optimal plan into two parts: the current return and the discounted continuation return. Such a solution strategy for sequential problems brings us a mayor advantage, it simplifies the computation required to solve the problem. Before starting to analyse somehow deeper the features of the dynamic programming approach, let us state a general notation for this kind of problem. A standard finite sequential

maximization problem can be written as

$$\begin{aligned}
V^*(x_0) = & \sup_{\{x_{t+1}\}_{t=0}^T} \sum_{t=0}^T \beta^t F_t(x_t, x_{t+1}) \\
\text{s.t. } & x_{t+1} \in \Gamma_t(x_t), \quad \text{for } t = 0, 1, \dots, T-1 \\
& x_0 \text{ given}
\end{aligned}$$

Hereafter, we will refer to this formulation as the sequence problem (SP). We call x_t the state (to be defined below), X is the space such that $x_t \in X, \forall t$; $\Gamma_t(x_t) : X \rightrightarrows X$ is the set of feasible actions, which we call the feasible set correspondence, and $F_t(x_t, x_{t+1}) : X \times X \rightarrow \mathbb{R}$ is the return function. The problem above can be read as follows: given an exogenous initial value x_0 , in each time period t from 0 to T we need to choose a set of control variables x_{t+1} that maximize the return function among all the possible choices given by the feasible set correspondence. Finally, $V^*(x_0)$ is called the value function and it specifies the highest possible value that the return (objective) function can reach starting with some x_t at time t . Note that we have used the operator ‘sup’ instead of the operator ‘max’ since, until now, nothing ensures us that the maximum value is attained by any feasible plan that we can choose. Nevertheless, in almost all the possible economic applications we can use the operator ‘max’ considering that we do not allow for infinite returns.

Definition 1.1 (State). *The state at time t is the smallest set of variables at t that allows to*

- *determine the feasible set for the controls (i.e. $\Gamma_t(x)$),*
- *determine the current return at t (i.e. $F_t(x, x')$) given an $x' \in \Gamma_t(x)$,*
- *determine the value tomorrow given an $x' \in \Gamma(x)$.*

We can rewrite this problem following a dynamic programming approach as follows. First note that since time is finite, $V_{T+1}(x) = 0, \forall x \in X$, i.e., since the agent dies at time T , the value perceived by this agent because of saving from T to $T+1$ is 0. For a generic t , we can write the following functional equation (Bellman equation)

$$V_t(x) = \sup_{x' \in \Gamma_t(x)} \{F_t(x, x') + \beta V_{t+1}(x')\}, \quad (9)$$

with the following associated policy function

$$g_t(x) = \arg \sup_{x' \in \Gamma_t(x)} \{F_t(x, x') + \beta V_{t+1}(x')\},$$

One immediate question to address is the following, from here, how can we obtain the Euler equations? For this, we need to use the Envelope theorem! (See [Appendix: Envelope Theorem](#)). Assuming an interior solution, the first order condition of (9) yields

$$\frac{\partial V_t(x)}{\partial x'} = 0 \Leftrightarrow \frac{\partial F_t(x, x')}{\partial x'} + \beta \frac{\partial V_{t+1}(x')}{\partial x'} = 0, \quad (10)$$

where by the Envelope theorem we have that

$$\frac{\partial V_{t+1}(x)}{\partial x} = \frac{\partial F_{t+1}(x, x')}{\partial x},$$

which implies

$$\frac{\partial V_{t+1}(x')}{\partial x'} = \frac{\partial F_{t+1}(x', x'')}{\partial x'}.$$

Then, substituting in (10) we obtain

$$\frac{\partial F_t(x, x')}{\partial x'} + \beta \frac{\partial F_{t+1}(x', x'')}{\partial x'} = 0.$$

Turning to the problem given by (1), we have a slight complication which is the presence of a borrowing constraint. We'll see how to deal with this later. In any case, as time is finite we can solve this problem backwards:

- At period $t = T$ (the last period), we know that the agent will not want to die with positive assets, therefore she will 'eat' all her capital un period T . Let us define the following policy function

$$a' = g_T(a) = 0.$$

This equation tells us that the optimal action of the agent at time T is not saving, i.e. setting $a_{T+1} = 0$. As we can see this function g has a subscript T which denotes the time period and it is a function of a , the state variable of our problem. Furthermore, since the optimal choice for $T + 1$ is setting $a' = 0$, we can define the value of the agent of entering period T with assets level a by the following value function

$$V_T(a) = u(c) = u\left(a + \omega_T - \frac{a'}{R}\right) = u(a + \omega_T) \equiv F_T(a, a' = 0), \quad (11)$$

which, loosely speaking, is a function that tell us the level of utility attained by the agent for each asset level a that she might enter period T with. Note that we are already imposing that $a_{T+1} = 0$, therefore in period T the agent consumption is given by both her asset level and her endowment in period T .

- At the remaining periods $t = T - 1, T - 2, \dots, 1, 0$, let us define the return function as

$$F_t(a, a') = u(c) = u\left(a + \omega_t - \frac{a'}{R}\right),$$

where $a \equiv a_t$ are the assets at the beginning of period $T - 1$ and $a' \equiv a_{t+1}$ are savings for tomorrow. The function $F_t(\cdot)$ is a function of both assets today and tomorrow. Let us also define the feasible set correspondence as

$$\Gamma_t(a) = (\underline{a}_{t+1}, R(a + \omega_{T-1})],$$

which, again, is a function of the assets today. To be more precise, from now onwards we define

- State variables: $a, (t)$, (recall Definition 1.1),
- Control variables: a' .

Loosely speaking, the state variables summarize the state of the economy and the control variables are the choice variables of each agent at each moment of time. Once we have defined this we can easily find the value function of entering period t with assets level a as

$$V_t(a) = \max_{a' \in \Gamma_t(a)} \{F_t(a, a') + \beta V_{t+1}(a')\},$$

with associated policy function

$$(\equiv a') \quad g_t(a) = \arg \max_{a' \in \Gamma_t(a)} \{F_t(a, a') + \beta V_{t+1}(a')\}.$$

First, to simplify the exposition, suppose that there are no borrowing constraints (you may assume $\underline{a}_t = -\infty$ for $t = 1, \dots, T$, and $\underline{a}_{T+1} = 0$). In this case the first order (necessary) conditions are given by

$$\frac{\partial V_t(a)}{\partial a'} = 0 \iff \frac{\partial F_t(a, a')}{\partial a'} + \beta \frac{\partial V_{t+1}(a')}{\partial a'} = 0, \quad (12)$$

where by the Envelope theorem we know that

$$\frac{\partial V_{t+1}(a)}{\partial a} = \frac{\partial F_{t+1}(a, a')}{\partial a},$$

which implies

$$\frac{\partial V_{t+1}(a')}{\partial a'} = \frac{\partial F_{t+1}(a', a'')}{\partial a'},$$

therefore (12) can be rewritten as

$$\frac{\partial F_t(a, a')}{\partial a'} + \beta \frac{\partial F_{t+1}(a', a'')}{\partial a'} = 0.$$

Substituting the specific expression for the return function we obtain

$$-\frac{1}{R} u' \left(a + \omega_t - \frac{a'}{R} \right) + \beta \left[u' \left(a' + \omega_{t+1} - \frac{a''}{R} \right) \right] = 0,$$

or equally

$$u' \left(a + \omega_t - \frac{a'}{R} \right) = R\beta u' \left(a' + \omega_{t+1} - \frac{a''}{R} \right),$$

which can also be expressed as

$$u'(c) = R\beta u'(c').$$

This is exactly the same result that we obtained with the Lagrangian approach.

However, given that in our setting the consumer can be borrowing constrained, we have to be careful when using the Envelope theorem. The next proposition deals with this issue.

Proposition 1.1. *In the life-cycle consumption-savings problem, if $V_{t+1}(a)$ is differentiable, increasing and concave:*

1. $V_t(\cdot)$ is weakly increasing, continuously differentiable and concave,
2. $g_t(\cdot)$ is continuous and

$$g_t(a) \begin{cases} = \underline{a}_{t+1} & \text{if } a \leq a_t^{thr} \\ > \underline{a}_{t+1} & \text{if } a > a_t^{thr} \end{cases}$$

for some $a_t^{thr} \in \mathbb{R}$,

3. $V'_{t+1}(a) = u'(a + w_t - g_t(a))$ (Envelope theorem).

Sketch of proof. Consider the agent's problem at t given $a_t \geq \underline{a}_t$ given by

$$\max_{a' \in [\underline{a}_{t+1}, R(a+w_t)]} \left\{ \underbrace{u \left(a + w_t - \frac{a'}{R} \right) + \beta V_{t+1}(a')}_{\equiv U(a'|a)} \right\}.$$

Define marginal cost and marginal benefit of savings as

$$U'(a'|a) = - \underbrace{\frac{1}{R} u' \left(a + w_t - \frac{a'}{R} \right)}_{\equiv MC(a'|a)} + \underbrace{\beta V'_{t+1}(a')}_{\equiv MB(a')}. \quad (13)$$

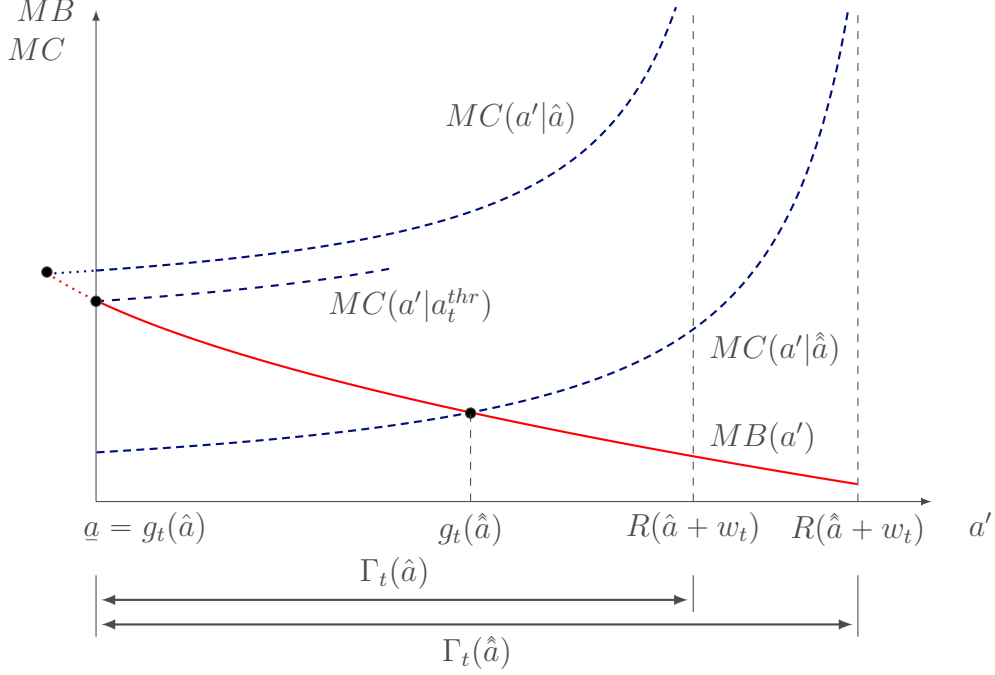
Note:

- (13) is equal to zero at an interior solution.
- For fixed a , $MC(\cdot|a)$ is continuous and strictly increasing in a' (as $u'' < 0$), also

$$\lim_{a' \rightarrow R(a+w_t)} MC(a'|a) = +\infty.$$

- For fixed a' , $MC(a'|\cdot)$ is decreasing in a .
- As $V_{t+1}(\cdot)$ is continuously differentiable and concave, $MB(\cdot)$ is continuous and decreasing.

We will now plot $MC(\cdot|\cdot)$ and $MB(\cdot)$ for different levels of a . Let $\hat{a} < \hat{\hat{a}}$, then



Then, for **Proposition 1.1** :

- Point 2 (form of $g_t(\cdot)$) follows from the previous figure,
- Points 1 and 3:
 - * if $a > a_t^{thr}$: follows from regular Envelope theorem,
 - * if $a < a_t^{thr}$:

$$V'_t(a) = \frac{d}{da}U(g_t(a)|a) = \frac{d}{da}U(\underline{a}|a) = u' \left(a + w_t - \frac{a}{R} \right),$$

- * if $a = a_t^{thr}$: as $g_t(a) \searrow \underline{a}$ as $a \searrow a_t^{thr}$, left and right derivative of $V_t(a_t^{thr})$ coincide and then

$$V'_t(a_t^{thr}) = u' \left(a_t^{thr} + w_t - \frac{a}{R} \right).$$

■

Remark. For differentiability of the value function without interiority assumptions see [Rincón-Zapatero and Santos \(2009\)](#).

1.3.1 Example of Dynamic Programming: McCall Search Model

- Real-options problem.
- Dynamic programming also works for solving discrete choice problem (dynamic choice).

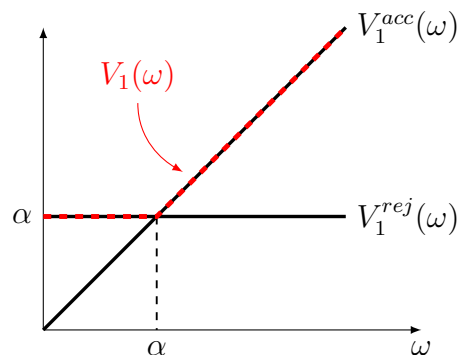
Model:

- Time is finite and discrete: $t = 0, 1$.
- There is one agent (a worker). A $t = 0, 1$ draws i.i.d wage ω_t from a c.d.f. $F(\omega)$ with support $[\underline{\omega}, \bar{\omega}]$. Decision:
 - Accept: stay with ω_t until the end,
 - Reject: get α in the current period, and have a new draw in the next period.
- y_t earnings of the worker.
- Assumption: $\alpha \in (\underline{\omega}, \bar{\omega})$.
- Worker maximizes $\mathbb{E}[y_0 + \beta y_1]$, where $\beta \in (0, 1)$.

To solve this problem, we start by studying the decision of the worker at time $t = 1$. Let $V_1(\omega)$ be the value of having drawn offer ω at $t = 1$. Moreover, let $V_1^{rej}(\omega)$ be the value of rejecting offer ω at $t = 1$ and let $V_1^{acc}(\omega)$ be the value of accepting offer ω at $t = 1$. In this case we have

$$\begin{aligned} V_1^{rej}(\omega) &= \alpha, \\ V_1^{acc}(\omega) &= \omega, \\ V_1(\omega) &= \max\{V_1^{rej}(\omega), V_1^{acc}(\omega)\}. \end{aligned}$$

We can represent this value function as



and the policy function at $t = 1$ is given by

$$g_1(\omega) = \begin{cases} 1 \text{ (Accept)} & \text{if } \omega \geq \alpha \\ 0 \text{ (Reject)} & \text{if } \omega < \alpha \end{cases} \quad (14)$$

As a consequence, conditional on having rejected the wage drawn in period $t = 0$, the reservation wage in period $t = 1$ is just the outside option α .

Now let $V_0(\omega)$ be the value of having drawn offer ω at $t = 0$. Moreover, let $V_0^{rej}(\omega)$ be the value of rejecting offer ω at $t = 0$ and let $V_0^{acc}(\omega)$ be the value of accepting offer ω at $t = 0$. In this case we have

$$\begin{aligned} V_0^{rej}(\omega) &= \alpha + \mathbb{E}[V_1(\omega)], \\ V_0^{acc}(\omega) &= \omega + \beta\omega, \end{aligned}$$

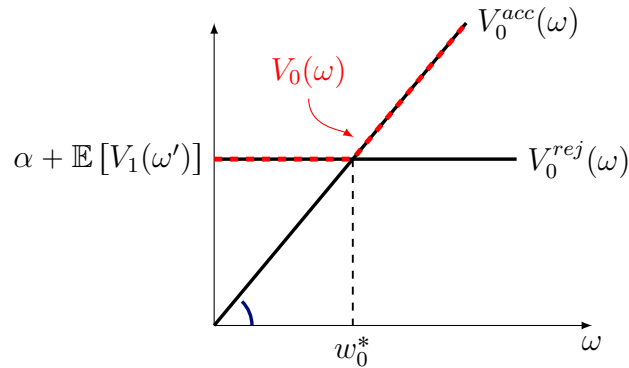
where

$$\begin{aligned} \mathbb{E}[V_1(\omega)] &= \int_{\omega}^{\bar{\omega}} V_1(\omega') dF(\omega') \\ &= \int_{\omega}^{\alpha} \alpha dF(\omega') + \int_{\alpha}^{\bar{\omega}} \omega' dF(\omega') \\ &= F(\alpha)\alpha + \int_{\alpha}^{\bar{\omega}} \omega' dF(\omega'). \end{aligned}$$

Therefore

$$\begin{aligned} V_0(\omega) &= \max \{(1 + \beta)\omega, \alpha + \mathbb{E}[V_1(\omega)]\} \\ &= \max_{a \in \{0,1\}} \{a(1 + \beta)\omega + (1 - a)(\alpha + \mathbb{E}[V_1(\omega)])\}. \end{aligned}$$

We can represent this value function as



The interesting question now is what is the reservation wage in period $t = 0$. To find it, note that the reservation wage in $t = 0$, ω_0^* must leave the worker indifferent between

accepting this wage and sticking with it until the end of period $t = 1$ or rejecting it, obtaining today α and having a new drawn in period $t = 1$. This indifference is given by the equality $V_0^{rej}(\omega_0^*) = V_0^{rej}(\omega_0^*)$, which implies

$$(1 + \beta)\omega_0^* = \alpha + \beta \left[F(\alpha)\alpha + \int_{\alpha}^{\bar{\omega}} \omega' dF(\omega') \right].$$

Therefore

$$\omega_0^* = \frac{\alpha + \beta \left[F(\alpha)\alpha + \int_{\alpha}^{\bar{\omega}} \omega' dF(\omega') \right]}{1 + \beta}.$$

Intuitively, $w_0^* > \alpha$, as α is always available for the worker as an outside option. However, let us prove it. First note that

$$\mathbb{E}[V_1(\omega)] = \underbrace{F(\alpha)\alpha}_{\text{Obtain } \alpha} + \underbrace{\int_{\alpha}^{\bar{\omega}} \omega' dF(\omega')}_{\text{Obtain } \omega > \alpha} > \alpha,$$

thus

$$\alpha + \mathbb{E}[V_1(\omega)] > \alpha + \beta\alpha,$$

and finally this implies that $w_0^* > \alpha = w_1^*$. Consequently, the worker gets less picky over time (option value of waiting).

2 Infinite Horizon Problems

2.1 One-Sector Neo-Classical Growth Model

This is one of the workhorse models of modern macroeconomics. The environment is given by:

- Large number of identical households:
 - Each household starts with k_0 units of capital.
 - Labour endowment: for every t , the household chooses labour $n_t \in [0, 1]$.
- Single consumption good which is produced with only two inputs, capital and labour. The production function is

$$y_t = F(k_t, n_t),$$

where $F(\cdot)$ is a neo-classical production function with the following properties:

- $F(k, n)$ is continuously differentiable.
- Strictly increasing in both arguments, i.e. $F_k(k, n) > 0$, $F_n(k, n) > 0$, $\forall k, n$.
- Strictly concave in both arguments, i.e. $F_{kk}(k, n) < 0$, $F_{nn}(k, n) < 0$, $\forall k, n$. This means that the production function exhibits decreasing marginal returns.
- Constant returns to scale (the production function is homogeneous of degree 1), i.e.,

$$F(\lambda k, \lambda n) = \lambda F(k, n), \quad \forall \lambda > 0.$$

With this assumption the size of the firm is indeterminate (the model cannot tell us anything about this).

- Inada conditions:

$$\begin{aligned} \lim_{k \rightarrow 0} F_k(k, 1) &= \infty & \lim_{n \rightarrow 0} F_n(1, n) &= \infty \\ \lim_{k \rightarrow \infty} F_k(k, 1) &= 0 & \lim_{n \rightarrow \infty} F_n(1, n) &= 0 \end{aligned}$$

- Inputs essential, i.e., $F(0, n) = 0$, $F(k, 0) = 0$, $\forall k, n$.

- Examples:

- * Cobb Douglas:

$$F(k, n) = Ak^\alpha n^{(1-\alpha)},$$

where $\alpha \in (0, 1)$, and A is the total factor productivity.

* CES (Constant elasticity of substitution):

$$F(k, n) = A (\mu k^\rho + (1 - \mu)n^\rho)^{1/\rho},$$

where typically $\mu \in (0, 1)$, $\rho < 1$, and A is the total factor productivity.

• Investment:

$$c_t + k_{t+1} \leq F(k, n) + (1 - \delta)k_t.$$

Note that this could be different due to, for example, irreversibilities or the presence adjustment costs.

• Preferences of the representative household:

$$U(\{c_t\}_{t=0}^{\infty}) = \sum_{t=0}^{\infty} \beta^t u(c_t),$$

where $u'(c) > 0$, $u''(c) < 0$ and

$$\lim_{c \rightarrow \infty} u'(c) = \infty.$$

Note that in this environment, leisure is not valued. Therefore each household will allocate all their time to work, i.e. $n_t = 1, \forall t$.

We are interested in solving the planner's problem. In this environment, a planner will solve

$$\begin{aligned} \max_{\{c_t, k_{t+1}\}_{t=0}^{\infty}} \quad & \sum_{t=0}^{\infty} \beta^t u(c_t) & (15) \\ \text{s.t.} \quad & c_t + k_{t+1} \leq F(k_t, 1) + (1 - \delta)k_t \quad t = 0, 1, \dots, \\ & c_t \geq 0, \quad t = 0, 1, \dots, \\ & k_{t+1} \geq 0, \quad t = 0, 1, \dots, \\ & k_0 \text{ given.} \end{aligned}$$

where we define $f(k_t)$ as the feasible resources at date t , i.e. $f(k_t) = F(k_t, 1) + (1 - \delta)k_t$. Note the following:

1. Since $u'(c_t) > 0$, then the feasibility constraint will always hold with equality (we do not want to leave goods on the table).
2. By the Inada conditions, $c_t = 0$ can't be optimal, therefore we must have $c_t > 0, \forall t$.
3. If $k_{t+1} = 0$ for some t , then $f(k_{t+1}) = 0$. Then we would have that $c_{t+1} = 0$. This possibility is ruled out, again, by the Inada conditions. Then we must have that $k_{t+1} > 0, \forall t$.

2.1.1 The Lagrangian approach and transversality conditions

As we did in the finite horizon case, we can solve this problem setting up the following Lagrangian

$$\mathcal{L}(\{c_t, k_{t+1}, \lambda_t\}_{t=0}^{\infty}) = \sum_{t=0}^{\infty} [\beta^t u(c_t) + \lambda_t (f(k_t) - c_t - k_{t+1})].$$

The first order necessary conditions for having a maximum are given by

$$\frac{\partial \mathcal{L}(\cdot)}{\partial c_t} = 0 \Leftrightarrow \beta^t u'(c_t) - \lambda_t = 0, \quad \forall t, \quad (16)$$

$$\frac{\partial \mathcal{L}(\cdot)}{\partial k_{t+1}} = 0 \Leftrightarrow -\lambda_t + \lambda_{t+1} f'(k_{t+1}) = 0, \quad \forall t, \quad (17)$$

Combining (16) and (17) we obtain

$$u'(c_t) = \beta u'(c_{t+1}) f'(k_{t+1}), \quad \forall t,$$

which is the Euler equation with the usual interpretation, in the optimum, the marginal cost of saving one unit must be equal to the discounted marginal utility of consuming the return of capital. The Euler equations give us necessary conditions for a solution $\{c_t^*, k_{t+1}^*\}_{t=0}^{\infty}$. Note that by substituting the feasibility constraint in the Euler equation we obtain

$$u'(f(k_t) - k_{t+1}) = \beta u'(f(k_{t+1}) - k_{t+2}) f'(k_{t+1}), \quad \forall t,$$

which is a second order difference equation for $\{k_{t+1}^*\}_{t=0}^{\infty}$. This equation together with the initial condition $k_0 = k_0^*$ characterize the optimal path of $\{k_{t+1}^*\}_{t=0}^{\infty}$, but we are still lacking a terminal condition, this is where the transversality condition will come in handy.

Up until now, we have found necessary conditions for a solution, but we are interested in finding sufficient conditions for a solution.

Fact 2.1. *Let $g(\cdot)$ and $f(\cdot)$ be concave functions, and furthermore, let $g(\cdot)$ be an increasing function. Then $h(\cdot) = g(f(\cdot))$ is also a concave function.*

Fact 2.2. *Let $x \in \mathbb{R}^n$ and let $f(x)$ be a function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ that is differentiable and concave. Then*

$$f(x) + Df(x)'[\tilde{x} - x] \geq f(\tilde{x}),$$

where

$$Df(x) = \begin{bmatrix} f_1(x) \\ \vdots \\ f_n(x) \end{bmatrix},$$

is the gradient of f in the point x .

Proposition 2.1. (*Guner, 2008, Proposition 77, pp.136-137*). Consider

$$\begin{aligned} \max_{\{k_{t+1}\}_{t=0}^{\infty}} \quad & \sum_{t=0}^{\infty} \beta^t F(k_t, k_{t+1}) \\ \text{s.t.} \quad & k_{t+1} \geq 0, \quad \forall t, \\ & k_0 \text{ given.} \end{aligned}$$

Let $F(\cdot)$ be continuously differentiable, and let $F(x, y)$ be concave in (x, y) and strictly increasing in x . If the sequence $\{k_{t+1}^*\}_{t=0}^{\infty}$ satisfies:

- $k_{t+1}^* \geq 0, \forall t$;
- (Euler equation): $F_2(k_t^*, k_{t+1}^*) + \beta F_1(k_{t+1}^*, k_{t+2}^*) = 0, \forall t$;
- (Transversality Condition): $\lim_{t \rightarrow \infty} \beta^t F_1(k_t^*, k_{t+1}^*) k_t^* = 0$;

Then $\{k_{t+1}^*\}_{t=0}^{\infty}$ maximizes the objective function, that is, the Euler equations are sufficient for $\{k_0^*, k_1^*, \dots\}$ to maximize

$$U(k_0^*, k_1^*, \dots) = \sum_{t=0}^{\infty} \beta^t F(k_t^*, k_{t+1}^*). \quad (18)$$

Proof. First consider a finite horizon T and rewrite (18) as

$$U(k_0^*, k_1^*, \dots, k_{T+1}^*) = F(k_0^*, k_1^*) + \beta F(k_1^*, k_2^*) + \beta^2 F(k_2^*, k_3^*) + \dots + \beta^T F(k_T^*, k_{T+1}^*),$$

then we have that

$$D[U(k_0^*, k_1^*, \dots, k_{T+1}^*)] = \begin{bmatrix} \frac{\partial U(k_0^*, k_1^*, \dots, k_{T+1}^*)}{\partial k_0^*} \\ \frac{\partial U(k_0^*, k_1^*, \dots, k_{T+1}^*)}{\partial k_1^*} \\ \vdots \\ \frac{\partial U(k_0^*, k_1^*, \dots, k_{T+1}^*)}{\partial k_{T+1}^*} \end{bmatrix} = \begin{bmatrix} F_1(k_0^*, k_1^*) \\ F_2(k_0^*, k_1^*) + \beta F_1(k_1^*, k_2^*) \\ \vdots \\ \beta^{T-1} F_2(k_{T-1}^*, k_T^*) + \beta^T F_1(k_T^*, k_{T+1}^*) \\ \beta^T F_2(k_T^*, k_{T+1}^*) \end{bmatrix},$$

and thus by **Fact 2.2** we have

$$U(k_1^*, \dots, k_{T+1}^*) + D[U(k_0^*, k_1^*, \dots, k_{T+1}^*)]' \begin{bmatrix} \tilde{k}_0 - k_0^* \\ \tilde{k}_1 - k_1^* \\ \vdots \\ \tilde{k}_T - k_T^* \\ \tilde{k}_{T+1} - k_{T+1}^* \end{bmatrix} \geq U(\tilde{k}_1, \dots, \tilde{k}_{T+1}). \quad (19)$$

Define

$$\begin{aligned}
D &= \lim_{T \rightarrow \infty} \left[U(k_1^*, \dots, k_{T+1}^*) - U(\tilde{k}_1, \dots, \tilde{k}_{T+1}) \right] \\
&= \lim_{T \rightarrow \infty} \left[\sum_{t=0}^T \beta^t F(k_t^*, k_{t+1}^*) - \sum_{t=0}^T \beta^t F(\tilde{k}_t, \tilde{k}_{t+1}) \right], \tag{20}
\end{aligned}$$

and let us rewrite (19) as

$$U(k_1^*, \dots, k_T^*) - U(\tilde{k}_1, \dots, \tilde{k}_{T+1}) \geq -D \left[U(k_0^*, k_1^*, \dots, k_{T+1}^*) \right]' \begin{bmatrix} \tilde{k}_0 - k_0^* \\ \tilde{k}_1 - k_1^* \\ \vdots \\ \tilde{k}_T - k_T^* \\ \tilde{k}_{T+1} - k_{T+1}^* \end{bmatrix},$$

where further rewriting yields

$$U(k_1^*, \dots, k_{T+1}^*) - U(\tilde{k}_1, \dots, \tilde{k}_{T+1}) \geq D \left[U(k_0^*, k_1^*, \dots, k_{T+1}^*) \right]' \begin{bmatrix} k_0^* - \tilde{k}_0 \\ k_1^* - \tilde{k}_1 \\ \vdots \\ k_T^* - \tilde{k}_T \\ k_{T+1}^* - \tilde{k}_{T+1} \end{bmatrix}. \tag{21}$$

Now, substituting (21) in (20) yields

$$\begin{aligned}
D &\geq \lim_{T \rightarrow \infty} \left[F_1(k_0^*, k_1^*)(k_0^* - \tilde{k}_0) + [F_2(k_0^*, k_1^*) + \beta F_1(k_1^*, k_2^*)](k_1^* - \tilde{k}_1) + \dots \right. \\
&\quad \left. \dots + [\beta^{T-1} F_2(k_{T-1}^*, k_T^*) + \beta^T F_1(k_T^*, k_{T+1}^*)](k_T^* - \tilde{k}_T) + \beta^T F_2(k_T^*, k_{T+1}^*)(k_{T+1}^* - \tilde{k}_{T+1}) \right] \\
&= \lim_{T \rightarrow \infty} \left[F_1(k_0^*, k_1^*)(k_0^* - \tilde{k}_0) + \sum_{t=0}^{T-1} \left(\beta^t [F_2(k_t^*, k_{t+1}^*) + \beta F_1(k_{t+1}^*, k_{t+2}^*)](k_{t+1}^* - \tilde{k}_{t+1}) \right) + \dots \right. \\
&\quad \left. \dots + \beta^T F_2(k_T^*, k_{T+1}^*)(k_{T+1}^* - \tilde{k}_{T+1}) \right].
\end{aligned}$$

By assumption, $k_0^* = \tilde{k}_0$, and by the Euler equation, $F_2(k_t^*, k_{t+1}^*) = \beta F_1(k_{t+1}^*, k_{t+2}^*)$, $\forall t$. As a consequence, the previous equation simplifies to

$$\begin{aligned}
D &\geq \lim_{T \rightarrow \infty} \left[\beta^T F_2(k_T^*, k_{T+1}^*)(k_{T+1}^* - \tilde{k}_{T+1}) \right] \\
&= - \lim_{T \rightarrow \infty} \left[\beta^{T+1} F_1(k_{T+1}^*, k_{T+2}^*)(k_{T+1}^* - \tilde{k}_{T+1}) \right] \\
&= - \lim_{T \rightarrow \infty} \left[\beta^{T+1} F_1(k_{T+1}^*, k_{T+2}^*) k_{T+1}^* \right] = 0,
\end{aligned}$$

i.e., $D \geq 0$. The second equality comes from the Euler equation and the third is straightforward considering that $F(\cdot)$ is bounded and $\tilde{k}_{T+1} \geq 0$, $\forall t$. Therefore if the limit of

the previous expression is 0, or in other words, if the transversality condition is satisfied, then $\{k_{t+1}^*\}_{t=0}^\infty$ maximizes the objective function.

Note that the transversality condition can be also be written by going one period backwards, obtaining

$$\lim_{T \rightarrow \infty} [\beta^T F_1(k_T^*, k_{T+1}^*) k_T^*] = 0. \quad (22)$$

■

2.1.2 The Dynamic Programming approach

The problem (15) can be rewritten as the following sequential problem

$$\begin{aligned} \max_{\{k_{t+1}\}_{t=0}^\infty} \quad & \sum_{t=0}^{\infty} \beta^t u(f(k_t) - k_{t+1}) \\ \text{s.t.} \quad & k_{t+1} \in \Gamma(k_t) = [0, f(k_t)], \quad t = 0, 1, \dots \\ & k_0 \text{ given.} \end{aligned} \quad (23)$$

Note that as we have moved to infinite time, we can get rid of the time subscript of the feasibility correspondence. Why we can do this? This follows from the idea of recursion that we will exploit to solve this problem. As time is infinite, the world must look exactly the same today and tomorrow. This idea of recursion allows us to split problem into two parts, today and the entire future. To better understand this approach, note that we can rewrite (23) as

$$\begin{aligned} \max_{\substack{\{k_{t+1}\}_{t=0}^\infty \\ k_0 \text{ given}}} \quad & \sum_{t=0}^{\infty} \beta^t u(f(k_t) - k_{t+1}) = \underbrace{\max_{\substack{k_1 \in \Gamma(k_0) \\ k_0 \text{ given}}} u(f(k_0) - k_1)}_{\text{Today}} + \underbrace{\max_{\substack{\{k_{t+1}\}_{t=1}^\infty \\ k_1 \text{ given}}} \sum_{t=1}^{\infty} \beta^t u(f(k_t) - k_{t+1})}_{\text{Future}} \\ & \text{s.t. } k_{t+1} \in \Gamma(k_t) \\ & = \max_{\substack{k_1 \in \Gamma(k_0) \\ k_0 \text{ given}}} u(f(k_0) - k_1) + \beta \max_{\substack{\{k_{t+1}\}_{t=1}^\infty \\ k_1 \text{ given}}} \sum_{t=1}^{\infty} \beta^{t-1} u(f(k_t) - k_{t+1}) \\ & \quad \text{s.t. } k_{t+1} \in \Gamma(k_t) \\ & = \max_{\substack{k_1 \in \Gamma(k_0) \\ k_0 \text{ given}}} u(f(k_0) - k_1) + \beta \max_{\substack{\{k_{t+1}\}_{t=1}^\infty \\ k_1 \text{ given}}} \sum_{t=0}^{\infty} \beta^t u(f(k_{t+1}) - k_{t+2}). \\ & \quad \text{s.t. } k_{t+2} \in \Gamma(k_{t+1}) \end{aligned}$$

Let us define the maximized value of the problem given k_0 as

$$V(k_0) = \max_{\substack{\{k_{t+1}\}_{t=0}^\infty \\ k_0 \text{ given}}} \sum_{t=0}^{\infty} \beta^t u(f(k_t) - k_{t+1}), \quad \text{s.t. } k_{t+1} \in \Gamma(k_t)$$

then we can write

$$V(k_0) = \max_{k_1 \in \Gamma(k_0)} \{u(f(k_0) - k_1) + \beta V(k_1)\}.$$

As time is infinite, we can write the value function for a generic t as

$$V(k) = \max_{k' \in \Gamma(k)} \{u(f(k) - k') + \beta V(k')\}, \quad (24)$$

where it is important to remark that the time subscript is no longer relevant. Besides, we can define the optimal policy function as

$$g(k) = \arg \max_{k' \in \Gamma(k)} \{u(f(k) - k') + \beta V(k')\} \quad (25)$$

In the dynamic programming approach, instead of looking for the optimal sequence $\{k_{t+1}\}_{t=0}^{\infty}$ like we do in the lagrangian (or sequential) approach, we will look for a pair of functions $V(\cdot)$ and $g(\cdot)$.

There are some open questions that we would like to answer:

- Is $V(k)$ well defined? The answer is yes as long as we replace the max operator for the sup operator (the max may not exist, but the supremum is always well defined).

Let us redefine the problem with a sup operator

$$\begin{aligned} V(k_0) \equiv \sup_{\{k_{t+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t u(f(k_t) - k_{t+1}) \quad (\text{SP}) \\ \text{s.t. } k_{t+1} \in \Gamma(k_t) = [0, f(k_t)], \quad t = 0, 1, \dots, \\ k_0 \text{ given.} \end{aligned}$$

- Is any solution $\tilde{V}(\cdot)$ to the functional equation

$$\tilde{V}(k) = \sup_{0 \leq k' \leq f(k)} \left\{ u(f(k) - k') + \beta \tilde{V}(k') \right\} \quad (\text{FE})$$

equal to $V(\cdot)$, i.e. $\tilde{V}(k) = V(k)$, $\forall k \geq 0$ and gives us the correct maximizing sequence to the sequence problem (SP)? In other words, will a solution to the Functional equation (Bellman equation) be unique? Will it be the same solution as the one obtained from the Sequential Problem?

- If the answer to the previous question is yes, how we can find a solution to the Functional equation (Bellman equation) (FE)? In functional equations the unknowns are precisely functions, therefore we need to define the space in which we want to look for the function that solves our functional equation. In our case, we will look for a solution in the space of continuous and bounded functions.

Once we have defined this, the idea to find the solution to the functional equation will be based on a series of successive approximations. Our aim is to approximate the true value of the value function by performing a sort of backwards induction like what we have to do in finite time. To this end, we take $V_0(\cdot)$ as an initial guess (usually $V_0(\cdot) = 0$), and set up the following algorithm

$$\begin{aligned} V_1(k) &= \max_{k' \in \Gamma(k)} \{u(f(k) - k') + \beta V_0(k')\}, \\ V_2(k) &= \max_{k' \in \Gamma(k)} \{u(f(k) - k') + \beta V_1(k')\}, \\ V_3(k) &= \max_{k' \in \Gamma(k)} \{u(f(k) - k') + \beta V_2(k')\}, \\ &\vdots \\ V_{n+1}(k) &= \max_{k' \in \Gamma(k)} \{u(f(k) - k') + \beta V_n(k')\}. \end{aligned}$$

where the subscript denotes the step in the algorithm. With this procedure, we hope that the value function $V_n(\cdot)$ converges to some $\tilde{V}(\cdot)$ and that the initial guess $V_0(\cdot)$ gets unimportant over time. Note: implicitly we have defined an operator on functions $T : X \rightarrow X$, where X is some space of functions with some properties (we need to define which properties). Then we have that

$$V_{n+1}(\cdot) = TV_n(\cdot),$$

i.e.

$$V_{n+1}(k) = TV_n(k) = \sup_{k' \in \Gamma(k)} \{u(f(k) - k') + \beta V_n(k')\}.$$

To be able to provide an appropriate answer to all these questions, we need to use some mathematical results.

2.2 Mathematical Part: Dynamic Programming

Let us redefine the problem given by (15) with a more general notation. Consider the sequence problem (SP) approach given by

$$\begin{aligned} V^*(x_0) &= \sup_{\{x_{t+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t F(x_t, x_{t+1}) & (\text{SP}) \\ \text{s.t. } & x_{t+1} \in \Gamma(x_t), \quad \text{for } t = 0, 1, \dots, \\ & x_0 \text{ given.} \end{aligned}$$

The dynamic programming approach to this problem is given by the Functional equation (Bellman equation)

$$V(x) = \sup_{x' \in \Gamma(x)} F(x, x') + \beta V(x'). \quad (\text{FE})$$

which implicitly defines an operator T given by

$$T : (TV)(x) = \sup_{x' \in \Gamma(x)} \{F(x, x') + \beta V(x')\},$$

$$\text{read as : } V_{n+1}(x) = \sup_{x' \in \Gamma(x)} \{F(x, x') + \beta V_n(x')\}.$$

We will now focus on the properties of the Bellman equation in an infinite time setting. The first difference to be noted relies on the absence of an explicit final condition of the form $V_{T+1} = 0$. We are interested in studying under which conditions this recursive problem is well defined and actually admits a (unique) solution. To this end, we will follow the next steps:

1. Define a function space for V ,
2. Show that T is a contraction mappings,
3. Show that T is a self-mapping,
4. The Principle of Optimality (SP \Leftrightarrow FE).

2.2.1 A function space for V

Let us introduce some mathematical definitions that we will need.³

Definition 2.1 (Metric). *A metric or distance function on X is a function $d : X \times X \rightarrow \mathbb{R}^+$ such that for every $x, y, z \in X$ we have*

$$(i) \quad d(x, y) \geq 0,$$

$$(ii) \quad d(x, y) = 0 \Leftrightarrow x = y,$$

$$(iii) \quad d(x, y) = d(y, x),$$

$$(iv) \quad \text{Triangle inequality: } d(x, z) \leq d(x, y) + d(y, z).$$

Definition 2.2 (Metric Space). *A Metric Space is pair (X, d) , where X is a set and d is a metric defined on it.*

Some sets X have an algebraic structure that allows us to operate with their elements. In particular we will consider vector spaces. We are interested not only in measuring the distance between points, but also in giving a meaning to the length of a vector.

³Unless stated otherwise, all the definitions and theorems that you'll find until Section 2.3 have been obtained from the lecture notes of the Mathematics course taught by [Juan Pablo Rincón-Zapatero](#) (n.d.). You may go back to his notes for a detailed and clearer explanation, proofs and/or examples.

Definition 2.3 (Vector Space). A vector space X is a set that is closed under finite vector addition and scalar multiplication. Let $f, g \in X$, then

- Addition: $(f + g)(x) = f(x) + g(x)$,
- Scalar Multiplication: $(\lambda f)(x) = \lambda f(x)$.

Definition 2.4 (Norm). Let X be a vector space over the reals (hence, in particular, $0 \in X$). A norm on X is a function $\|\cdot\| : X \rightarrow \mathbb{R}^+$ such that for every $x, y \in X$ and for every scalar $\lambda \in \mathbb{R}$ we have

- (i) $\|x\| \geq 0$,
- (ii) $\|x\| = 0 \Leftrightarrow x = 0$,
- (iii) Triangle inequality: $\|x + y\| \leq \|x\| + \|y\|$,
- (iv) Homogeneity: $\|\lambda x\| = |\lambda| \|x\|$.

Definition 2.5 (Normed Vector Space). A normed space is a pair $(X, \|\cdot\|)$ where X is a vector space and $\|\cdot\|$ is a norm.

Definition 2.6. Given a set X , the vector space of real bounded and continuous functions f on X , is denoted $\mathcal{C}(X)$

$$\mathcal{C}(X) = \{f : X \rightarrow \mathbb{R} : f \text{ bounded and continuous}\}.$$

Definition 2.7 (Supremum Norm). Let X be a set and let $\mathcal{C}(X)$ be the space of all bounded and continuous functions defined on X . The supremum norm is the norm defined on $\mathcal{C}(X)$ by

$$\|f\|_\infty = \sup_{(x \in X)} |f(x)|.$$

Remark. $(\mathcal{C}(X), \|\cdot\|_\infty)$ is a normed vector space.

Definition 2.8 (Bolzano-Weierstrass Property). A subset A of a metric space X has the Bolzano-Weierstrass Property if every sequence in A has a convergent subsequence, i.e. has a subsequence that converges to a point in A .

Theorem 2.1 (Compact Metric Space). A subset of a metric space is compact if and only if it enjoys the Bolzano-Weierstrass Property.

Definition 2.9 (Cauchy Sequence). Let (X, d) be a metric space. A sequence $x_n \in X$ is Cauchy if given any $\varepsilon > 0$, there exists $n_0(\varepsilon)$ such that

$$d(x_p, x_q) < \varepsilon \text{ whenever } p, q \geq n_0(\varepsilon).$$

Remark. Every convergent sequence is a Cauchy sequence, but the reciprocal is not true.

Definition 2.10 (Complete Metric Space). A metric space (X, d) is complete if every Cauchy sequence converges.

Definition 2.11 (Banach Space). A normed space $(X, \|\cdot\|)$ is a Banach Space if it is complete.

Lemma 2.1. The space $\mathcal{C}(X)$ is complete, therefore it is a Banach Space.

We will look for the function V in $\mathcal{C}(X)$, the space of continuous and bounded functions with the supremum norm. This vector space is a complete normed vector space, and therefore it is a Banach space.

2.2.2 T is a Contraction Mapping

It turns out that under some regularity conditions, the operator T is a contraction mapping. Since we are looking for our function in a Banach space, then if we can show that T is a contraction mapping we can apply the Banach Fixed Point theorem, that ensures the existence and uniqueness of the limit function V and also provides an algorithm to find it. To be able to prove that T is a contraction mapping, we will use Blackwell's sufficient conditions.

Definition 2.12 (Mapping). A map is a way of associating unique objects to every element in a given set. So a map $f : A \rightarrow B$ from A to B is a function f such that for every $a \in A$, there is a unique object $b \in B$.

Definition 2.13 (Fixed Point). Let $T : X \rightarrow X$ be a mapping. We say that $x \in X$ is a fixed point of T if $T(x) = x$.

Definition 2.14 (Contraction Mapping). Let (X, d) be a metric space. The operator $T : (X, d) \rightarrow (X, d)$ is a contraction mapping with parameter (modulus) β if for some $0 < \beta < 1$ we have that

$$d(Tx, Ty) \leq \beta d(x, y), \quad \forall x, y \in X.$$

Theorem 2.2 (Banach Fixed Point Theorem - Contraction Mapping Theorem). Let (X, d) be a complete metric space and let T a contraction operator. Then T admits a unique fixed point, which can be approached from successive iterations of the operator T from any initial element $x \in X$. In particular, if the contraction parameter is β , then we have that

- T has exactly one fixed point $V \in X$, i.e. $T(V) = V$,

- and for any $V_0 \in X$,

$$d(T^n V_0, V) \leq \beta^n d(V_0, V), \quad n = 0, 1, \dots$$

Corollary 2.1. *Let $(X, \|\cdot\|)$ be a Banach Space and $T : (X, \|\cdot\|) \rightarrow (X, \|\cdot\|)$ be a contraction mapping with fixed point $V \in X$. If X' is a closed subset of X and $T(X') \subseteq X'$, then $V \in X'$. If, in addition, $T(E') \subseteq E'' \subseteq E'$, then $V \in E''$*

Proof. Choose $V_0 \in X'$. Then the sequence $\{T^n V\}_{n=0}^{\infty} \rightarrow V$, $T^n V \in X'$. Since E' is closed, $V \in X'$ (by the definition of closeness). If also $T(E) \subseteq X''$, then since $V = T(V)$ and $V \in X'$, then $T(V) \in X''$ and thus $V \in X''$. ■

How can we identify that some operator T is a contraction? We can show it either by first principles (i.e., applying the previous mathematical artillery) or we can use Blackwell's sufficient conditions.

Theorem 2.3 (Blackwell's Sufficient Conditions). *Let $X \subset \mathbb{R}^\ell$ and $\mathcal{B}(X)$ be the space of bounded functions $f : X \rightarrow \mathbb{R}$ with the supremum norm. Let $T : \mathcal{B}(X) \rightarrow \mathcal{B}(X)$ be an operator satisfying*

1. *Monotonicity: if $f, g \in \mathcal{B}(X)$ and $f(x) \leq g(x)$, $\forall x \in X$, then this implies that $(Tf)(x) \leq (Tg)(x)$, $\forall x \in X$.*
2. *Discounting: \exists some $\beta \in (0, 1)$ such that*

$$[(T(f + a))](x) \leq (Tf)(x) + \beta a, \quad \forall f \in \mathcal{B}(X), \forall a \in \mathbb{R}^+, \forall x \in X.$$

Then T is a contraction operator with contraction parameter (modulus) β .

Example Let us apply this procedure to see whether our operator in the growth model is a contraction mapping. From (24), write

$$(TV)(k) = \max_{k' \in \Gamma(k)} \{u(f(k) - k') + \beta V(k')\},$$

First, let $W(k) \geq V(k)$, $\forall k$, then

$$\begin{aligned} (TW)(k) &= \max_{0 \leq k' \leq f(k)} \{u(f(k) - k') + \beta W(k')\} \\ &\geq \max_{0 \leq k' \leq f(k)} \{u(f(k) - k') + \beta V(k')\} = (TV)(k), \end{aligned}$$

as for each $0 \leq k' \leq f(k)$ (for fixed k), and $W(k') \geq V(k')$ by assumption. Note that fixing k implies that the feasibility constraint does not change. Therefore, monotonicity holds in the neoclassical growth model.

Furthermore

$$\begin{aligned}
[T(V + a)](k) &= \max_{0 \leq k' \leq f(k)} \{u(f(k) - k') + \beta[V(k') + a]\} \\
&= \max_{0 \leq k' \leq f(k)} \{u(f(k) - k') + \beta V(k') + \beta a\} \\
&= \max_{0 \leq k' \leq f(k)} \{u(f(k) - k') + \beta V(k')\} + \beta a, \\
&= (TV)(k) + \beta a
\end{aligned}$$

and as $=$ implies \leq , then discounting also holds in the neoclassical growth model.

Therefore, T is a contraction operator in the neoclassical growth model.

2.2.3 T is a Self-Mapping

We want to show that T is a self mapping because if this holds and our initial guess V_0 is, for example, a continuous and bounded function, then T will map continuous and bounded functions into continuous and bounded functions, and therefore V , our main interest, will also be a continuous and bounded function. To be able to show that T is a self-mapping, we will use Berge's Maximun Theorem. Let us first introduce two relevant theorems.

Theorem 2.4 (Weierstrass' Theorem). *If $f : X \rightarrow G$ is a continuous function that maps a compact metric space X into a metric space G , then $f(X)$ is compact.*

Theorem 2.5 (Extreme Value Theorem). *Let K be a compact subset of the metric space X and $f : K \rightarrow \mathbb{R}$ a continuous function. Then there exists $x_1, x_2 \in K$ such that*

$$f(x_2) \leq f(x) \leq f(x_1), \quad \forall x \in K. \quad (26)$$

Now, consider the general problem

$$v(x) = \sup_{y \in \Gamma(x)} f(x, y),$$

where $x \in X$, $X \subset \mathbb{R}^\ell$, $y \in Y$, $Y \subset \mathbb{R}^m$, $f : X \times Y \rightarrow \mathbb{R}$ and $\Gamma : X \rightrightarrows Y$. Note that if $f(x, \cdot)$ is continuous in y (for fixed x) and Γ is a non-empty and compact valued correspondence (i.e. $\Gamma(x)$ is a compact set $\forall x \in X$), then by fixing an x we are maximizing a continuous function on a compact set. Then, by Weierstrass' Theorem (Extreme Value Theorem), the maximum exists. Let us write the value function $h(x)$ as

$$h(x) = \max_{y \in \Gamma(x)} f(x, y),$$

and the optimal correspondence (policy function) $G(x)$ as

$$G(x) = \arg \max_{y \in \Gamma(x)} f(x, y) = \{y \in \Gamma(x) : f(x, y) = h(x)\}.$$

Theorem 2.6 (Theorem of the Maximum of Berge). *Assume that $v(x) < +\infty$, $\forall x \in X$. Then*

1. *If f is lower semi-continuous and Γ is non-empty and lower hemi-continuous, then $h(x)$ is lower-semi-continuous.*
2. *If f is upper semi-continuous and Γ is non-empty, compact-valued and upper hemi-continuous, then $h(x)$ is upper-semi-continuous.*
3. *If f is continuous and Γ is non-empty, compact-valued and continuous, then $h(x)$ is continuous and $G(x)$ is non-empty, compact-valued and upper hemi-continuous (we lose lower hemi-continuity).*

Corollary 2.2. *Suppose that $X \subset \mathbb{R}^m$ is convex, $f : X \times Y \rightarrow \mathbb{R}$ is continuous and concave in y (for fixed x) and that $\Gamma : X \rightrightarrows Y$ is a continuous set-valued mapping, compact valued and with convex graph. Then*

1. *The value function $h(x)$ is concave and the optimal correspondence $G(x)$ is convex-valued.*
2. *If $f(x, \cdot)$ is strictly concave in y for every $x \in X$, then $G(x)$ is single-valued and it is continuous as a function.*

Remark. *We can use as a fact that any correspondence $\Gamma(x) = [f(x), g(x)]$ in which the bounds $f(\cdot)$ and $g(\cdot)$ are continuous functions is continuous.*

Example Let us apply this to the one-sector neo-classical growth model. Define

- $f(x, y) \equiv u(f(k) - k') + \beta V(k')$. Since both $u(\cdot)$ and $f(\cdot)$ are continuous functions, as long as $V(\cdot)$ is a continuous function then $u(f(k) - k') + \beta V(k')$ is also a continuous function.
- $\Gamma(k) = [0, f(k)]$. Since 0 and $f(x)$ are both continuous functions of x , then $\Gamma(x)$ is a non-empty, compact and continuous correspondence.

Then by Berge's maximum theorem, we have that $TV(\cdot)$ is a continuous function.

2.3 Principle of Optimality

The value function $V^*(x_0)$ given by (SP) tells us the infinite discounted value of following the best sequence $\{x_{t+1}^*\}_{t=0}^{\infty}$. Our hope is that, rather than finding the best sequence $\{x_{t+1}^*\}_{t=0}^{\infty}$, we can try to find the function $V^*(x_0)$ as a solution to (FE). If our conjecture

is correct, then the function V that solves (FE) would give us the value function, i.e. $V(x_0) = V^*(x_0)$. In this section we will cover some theorems on the relationship between the sequence problem and the functional equation approach. Let us recover the definition for the value function $V^*(\cdot)$ for the sequence problem, which is given by

$$V^*(x_0) = \sup_{\{x_{t+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t F(x_t, x_{t+1}) \quad (\text{SP})$$

s.t. $x_{t+1} \in \Gamma(x_t), \forall t,$
 x_0 given.

and the dynamic programming approach to this problem is given by the Functional equation (Bellman equation)

$$V(x) = \sup_{x' \in \Gamma(x)} F(x, x') + \beta V(x'). \quad (\text{FE})$$

In this section we will show that

- The value function $V^*(\cdot)$ satisfies (FE).
- If there exists a solution to (FE), then it is the value function $V^*(\cdot)$.
- There is indeed a solution to (FE).
- A sequence $\{x_{t+1}^*\}_{t=0}^{\infty}$ attains the maximum in (SP) if it satisfies

$$V(x_t^*) = F(x_t^*, x_{t+1}^*) + \beta V(x_{t+1}^*), \forall t.$$

Notation and definitions Let us set some notation and useful definitions:

- ‘Plan’: any sequence $\{x_{t+1}\}_{t=0}^{\infty}$,
- ‘Feasible plan’: plan $\{x_{t+1}\}_{t=0}^{\infty}$ that satisfies $x_{t+1} \in \Gamma(x_t)$ for $t = 0, 1, \dots$
- $\Pi(x_0)$: the set of feasible plans given a particular x_0 , i.e.

$$\Pi(x_0) = \left\{ \{x_{t+1}\}_{t=0}^{\infty} : x_{t+1} \in \Gamma(x_t) \text{ for } t = 0, 1, \dots \right\}$$

- $\tilde{x} = \{x_1, x_2, \dots\} \in \Pi(x_0)$ is a typical feasible plan.
- For each $n = 0, 1, \dots$ we will define

$$u_n(\tilde{x}) = \sum_{t=0}^n \beta^t F(x_t, x_{t+1}),$$

as the discounted return from following a feasible plan \tilde{x} from date 0 to date n.

- Let $u : \Pi(x_0) \rightarrow \mathbb{R}$

$$u(\tilde{x}) = \lim_{n \rightarrow \infty} u_n(\tilde{x}) = \lim_{n \rightarrow \infty} \sum_{t=0}^n \beta^t F(x_t, x_{t+1}) = \sum_{t=0}^{\infty} \beta^t F(x_t, x_{t+1}),$$

be the infinite discounted sum of returns from following feasible plan \tilde{x} .

Then, the value function is defined as

$$V^*(x_0) = \sup_{\tilde{x} \in \Pi(x_0)} u(\tilde{x}).$$

As $V^*(x_0)$ is a supremum, it is the only function that satisfies the following two requirements:

$$V^*(x_0) \geq u(\tilde{x}), \quad \forall \tilde{x} \in \Pi(x_0), \quad (\text{SP1})$$

and for any $\varepsilon > 0$

$$V^*(x_0) \leq u(\tilde{x}) + \varepsilon, \quad \text{for some } \tilde{x} \in \Pi(x_0). \quad (\text{SP2})$$

This implies that no other feasible sequence can do better than $V^*(\cdot)$, but we can get ε close, as close as we want. Moreover, $V(x)$ being a supremum means

$$V(x) \geq F(x, y) + \beta V(y), \quad \forall y \in \Gamma(x), \forall x \in X, \quad (\text{FE1})$$

and for any $\varepsilon > 0$

$$V(x) \leq F(x, y) + \beta V(y) + \varepsilon \quad \text{for some } y \in \Gamma(x), \forall x \in X, \quad (\text{FE2})$$

which has the same interpretation as before.

Assumption (A1). $\Gamma(x) \neq \emptyset, \forall x \in X$, i.e., there is always at least one feasible sequence.

Assumption (A2). $F : A \rightarrow \mathbb{R}$ is bounded, where $A = \{(x, y) : x \in X, y \in \Gamma(x)\}$ is the graph of $\Gamma(\cdot)$.

Remark. Under **A1** and **A2**, $V^*(x_0)$ is uniquely defined.

Proof. By **A2**, $F(x, x')$ is bounded. Let B be a bound on $|F(x, x')|$ where $(x, x') \in A$, then

$$|V^*(x_0)| \leq \sup_{\tilde{x} \in \Pi(x_0)} \sum_{t=0}^{\infty} \beta^t |F(x_t, x_{t+1})| \leq \sup_{\tilde{x} \in \Pi(x_0)} \sum_{t=0}^{\infty} \beta^t B = \frac{B}{1 - \beta},$$

thus $V^*(x_0)$ is also bounded.

Furthermore, $V^*(x_0)$ is the unique function that satisfies **(SP1)** and **(SP2)**. Hence, if $|V^*(x_0)| < \infty$, then it must give more utility than any other feasible plan. In order to prove **(SP2)**, by contradiction suppose there is no such \tilde{x} , then for all $\tilde{x} \in \Pi(x_0)$

$$V^*(x_0) - \varepsilon \geq u(\tilde{x}),$$

which also implies that

$$V^*(x_0) - \varepsilon \geq \max_{\tilde{x} \in \Pi(x_0)} u(\tilde{x}) = V^*(x_0),$$

which is a contradiction. ■

We will first prove a Lemma that will be key to the following proofs. This Lemma simply tells us that we can separate discounted infinite sum of returns from any feasible plan into current and future returns. This separation is key in dynamic programming.

Lemma 2.2 (Lemma 4.1 in (Stokey et al., 1989)). *Let X, Γ, F and β satisfy **A2**. Then for any $x_0 \in X$ and any $\tilde{x} = (x_0, x_1, \dots) \in \Pi(x_0)$*

$$u(\tilde{x}) = F(x_0, x_1) + \beta u(\tilde{x}'),$$

where $\tilde{x}' = (x_1, \dots)$ is the continuation plan given x_0 .

Proof. Under **A2**, for any $x_0 \in X$ and any $\tilde{x} \in \Pi(x_0)$,

$$u(\tilde{x}) = \lim_{n \rightarrow \infty} \sum_{t=0}^n \beta^t F(x_t, x_{t+1}),$$

which can be expressed as

$$\begin{aligned} u(\tilde{x}) &= F(x_0, x_1) + \lim_{n \rightarrow \infty} \sum_{t=1}^n \beta^t F(x_t, x_{t+1}) \\ &= F(x_0, x_1) + \beta \lim_{n \rightarrow \infty} \sum_{t=0}^n \beta^t F(x_{t+1}, x_{t+2}) \\ &= F(x_0, x_1) + \beta u(\tilde{x}'). \end{aligned}$$

■

Fact 2.3. *Weak inequalities are preserved in the limit. Let $x_n \in \mathbb{R}$, $n = 0, 1, \dots$, then if $x_n \rightarrow x$ and $x_n \geq y$, $\forall n$, then $x \geq y$.*

Theorem 2.7 (**SP** \Rightarrow **FE**, Theorem 4.2 in (Stokey et al., 1989)). *Let X, Γ, F and β satisfy (A1) and (A2). Then the value function V^* satisfies the Functional equation.*

Proof. Suppose $\beta > 0$ (otherwise the result is trivial) and choose x_0 . We know that (SP1) and (SP2) hold since V^* is the value function. We will divide the proof in two steps, first we will show that (SP1) and (SP2) imply (FE1) and then we will show that they also imply (FE2).

(a) Show (FE1):

Fix some choice $x_1 \in \Gamma(x_0)$. Take a decreasing sequence that converges to 0, i.e. $\varepsilon_n/\beta \rightarrow 0$. Then, $\forall \varepsilon_n$ by (SP2) there exists some feasible continuation plan $\tilde{x}' = (x_1, x_2, \dots) \in \Pi(x_1)$ such that

$$V^*(x_1) \leq u(\tilde{x}') + \frac{\varepsilon_n}{\beta},$$

which can be rewritten as

$$u(\tilde{x}') \geq V^*(x_1) - \frac{\varepsilon_n}{\beta}. \quad (27)$$

This is true as $V^*(x_1)$ is the supremum function. Since $(x_0, \tilde{x}') \in \Pi(x_0)$, by (SP1) and the previous Lemma we have that

$$\begin{aligned} V^*(x_0) &\geq u(\tilde{x}) = F(x_0, x_1) + \beta u(\tilde{x}') \\ &\geq F(x_0, x_1) + \beta \left[V^*(x_1) - \frac{\varepsilon_n}{\beta} \right] \\ &= F(x_0, x_1) + \beta V^*(x_1) - \varepsilon_n, \end{aligned}$$

where the second equality follows from substituting (27). Since weak inequalities are preserved in the limit, we have

$$V^*(x_0) \geq F(x_0, x_1) + \beta V^*(x_1),$$

which is precisely (FE1). Since this holds for any $x_1 \in \Gamma(x_0)$, this establishes (FE1).

(b) Show (FE2):

Choose $x_0 \in X$ and fix $\varepsilon > 0$. By (SP2) and the previous Lemma, for some $\tilde{x} \in \Pi(x_0)$

$$\begin{aligned} V^*(x_0) &\leq u(\tilde{x}) + \varepsilon = F(x_0, x_1) + \beta u(\tilde{x}') + \varepsilon \\ &\leq F(x_0, x_1) + \beta V^*(x_1) + \varepsilon, \end{aligned}$$

where the second inequality follows from (SP1). This establishes (FE2). ■

This theorem shows that the supremum function V^* obtained from (SP) satisfies the (FE). The following theorem provides a converse: If V satisfies (FE) (i.e. if V is a solution to (FE)) and if its bounded, then V is the supremum function (i.e. $V = V^*$) that solves (SP).

Theorem 2.8 (FE \Rightarrow SP, Theorem 4.3 in (Stokey et al., 1989)). *Let X, Γ, F and β satisfy (A1) and (A2). If V is a solution to (FE) and satisfies*

$$\lim_{n \rightarrow \infty} \beta^n V(x_n) = 0, \quad \forall \tilde{x} \in \Pi(x_0), \quad \forall x_0. \quad (28)$$

Then $V = V^$.*

Proof. We will divide the proof in two steps, first we will show that (FE1) and (FE2) imply (SP1) and (SP2).

(a) Show (SP1):

By (FE1), we have that $\forall \tilde{x} \in \Pi(x_0)$

$$\begin{aligned} V(x_0) &\geq F(x_0, x_1) + \beta V(x_1) \\ &\geq F(x_0, x_1) + \beta [F(x_1, x_2) + \beta V(x_2)] \\ &\geq F(x_0, x_1) + \beta F(x_1, x_2) + \beta^2 [F(x_2, x_3) + \beta V(x_3)] \\ &\geq \dots \\ &\geq u_n(\tilde{x}) + \beta^{n+1} V(x_{n+1}), \quad \text{for } n = 1, 2, \dots, \end{aligned}$$

as $n \rightarrow \infty$ by (28) and since weak inequalities are preserved in the limit, we have that

$$\begin{aligned} V(x_0) &\geq \lim_{n \rightarrow \infty} (u_n(\tilde{x}) + \beta^{n+1} V(x_{n+1})) \\ &\geq u(\tilde{x}). \end{aligned}$$

Since $\tilde{x} \in \Pi(x_0)$ was arbitrarily chosen, this establishes (SP1).

(b) Show (SP2):

Also, for any $\varepsilon > 0$ note that we can choose $\{\delta_t\}_{t=1}^{\infty}$, $\delta_t > 0$, $\forall t$, such that

$$\sum_{t=1}^{\infty} \beta^{t-1} \delta_t \leq \varepsilon \tag{29}$$

By (FE2) we can choose $x_{t+1} \in \Gamma(x_t)$, $\forall t$ such that

$$V(x_t) \leq F(x_t, x_{t+1}) + \beta V(x_{t+1}) + \delta_{t+1},$$

then we have

$$\begin{aligned} V(x_0) &\leq F(x_0, x_1) + \beta V(x_1) + \delta_1 \\ &\leq F(x_0, x_1) + \beta [F(x_1, x_2) + \beta V(x_2) + \delta_2] + \delta_1 \\ &\leq F(x_0, x_1) + \beta F(x_1, x_2) + \beta V(x_2) + \delta_1 + \beta \delta_2 \\ &\leq \dots \\ &\leq \sum_{t=0}^n \beta^t F(x_t, x_{t+1}) + \beta^{n+1} V(x_{n+1}) + \sum_{t=1}^n \delta_t \beta^{t-1} \\ &\leq u_n(\tilde{x}) + \beta^{n+1} V(x_{n+1}) + \sum_{t=1}^n \delta_t \beta^{t-1}. \end{aligned}$$

As weak inequalities are preserved in the limit, using (28) and (29) we have that

$$\begin{aligned} V(x_0) &\leq \lim_{n \rightarrow \infty} \left(u_n(\tilde{x}) + \beta^{n+1}V(x_{n+1}) + \sum_{t=1}^n \delta_t \beta^{t-1} \right) \\ &\leq u(\tilde{x}) + \varepsilon, \end{aligned}$$

for some $\tilde{x} \in \Pi(x_0)$ given any arbitrary $\varepsilon > 0$. This establishes (SP2). ■

It is important to note that the proof requires that $\lim_{n \rightarrow \infty} \beta^n V(x_n) = 0$ holds for all feasible plans. Obviously this is satisfied if F and as a result V is bounded. If boundedness is not satisfied, and if there is a feasible plan that does not satisfy $\lim_{n \rightarrow \infty} \beta^n V(x_n) = 0$; then we can not conclude that V is the supremum function (even if it satisfies (FE)). If the boundedness fails, then there might be solutions to (FE) that are not supremum function.

Theorem 2.9 (FE \Rightarrow Optimal Policy. Theorem 4.4 in (Stokey et al., 1989)). *Let X, Γ, F and β satisfy (A1) and (A2). Let $\tilde{x}^* \in \Pi(x_0)$ be a feasible plan starting from x_0 satisfying*

$$V^*(x_0) = u(\tilde{x}^*),$$

i.e., it attains the supremum in (SP). Then

$$V^*(x_t^*) = F(x_t^*, x_{t+1}^*) + \beta V^*(x_{t+1}^*), \quad t = 0, 1, \dots \quad (30)$$

Proof. Since \tilde{x}^* attains the supremum,

$$\begin{aligned} V^*(x_0^*) &= u(\tilde{x}^*) = F(x_0, x_1^*) + \beta u(\tilde{x}^*) \\ &\geq u(\tilde{x}) = F(x_0, x_1) + \beta u(\tilde{x}'), \quad \text{for any } \tilde{x} \in \Pi(x_0). \end{aligned} \quad (31)$$

In particular, the inequality follows for all plans with $x_1 = x_1^*$. As $(x_1^*, x_2, x_3, \dots) \in \Pi(x_1^*)$ implies that $(x_0, x_1^*, x_2, x_3, \dots) \in \Pi(x_0)$ it follows that

$$u(\tilde{x}'^*) \geq u(\tilde{x}'), \text{ for any } \tilde{x} \in \Pi(x_1^*).$$

Therefore, $u(\tilde{x}'^*) = V(x_1^*)$. Substituting this into (31) gives (30) for $t = 0$. Continuing by induction establishes (30) for any t . ■

Theorem 2.10 (Optimal Policy \Rightarrow FE. Theorem 4.5 in (Stokey et al., 1989)). Let X, Γ, F and β satisfy **A1** and **A2**. Let $\tilde{x}^* \in \Pi(x_0)$ be a feasible plan starting from x_0 satisfying (30), i.e.

$$V^*(x_t^*) = F(x_t^*, x_{t+1}^*) + \beta V^*(x_{t+1}^*), \quad t = 0, 1, \dots,$$

with

$$\limsup_{t \rightarrow \infty} \beta^t V^*(x_t^*) \leq 0, \tag{32}$$

then \tilde{x}^* attains the supremum in (SP) for initial state x_0 .

Proof. Note that from (30) we can write

$$\begin{aligned} V^*(x_0^*) &= F(x_0^*, x_1^*) + \beta V^*(x_1^*) \\ &= F(x_0^*, x_1^*) + \beta [F(x_1^*, x_2^*) + \beta V^*(x_2^*)] \\ &= \dots \\ &= u_n(\tilde{x}^*) + \beta^{n+1} V^*(\tilde{x}_{n+1}^*), \end{aligned}$$

Taking limits when $n \rightarrow \infty$ we obtain

$$\begin{aligned} V^*(x_0^*) &= u(\tilde{x}^*) + \lim_{n \rightarrow \infty} \beta^{n+1} V^*(\tilde{x}_{n+1}^*) \\ &\leq u(\tilde{x}^*), \end{aligned}$$

where the second inequality follows from (32). Note that, from (SP1) we must have that

$$V^*(x_0^*) \geq u(\tilde{x}^*),$$

therefore both results imply

$$V^*(x_0^*) = u(\tilde{x}^*).$$

■

2.4 Bellman equations

Consider the Bellman equation (functional equation) given by

$$V(x) = \max_{y \in \Gamma(x)} \{F(x, y) + \beta V(y)\}, \tag{BE}$$

where $x \in X$, being X the state space; $\Gamma : X \rightarrow X$ is the correspondence describing the feasibility constraints; $A = \{(x, y) \in X \times X : y \in \Gamma(x)\}$ is the graph of Γ ; $F : A \rightarrow \mathbb{R}$ is the return function and $0 < \beta < 1$ is the discount factor. The following assumptions guarantee that a solution to (BE) exists, is unique and has some desirable properties.

Assumption (B1). X is a convex subset of \mathbb{R}^ℓ . $\Gamma : X \rightarrow X$ is a non-empty, compact-valued and continuous correspondence.

Assumption (B2). $F : A \rightarrow \mathbb{R}$ is bounded and continuous, where the graph of the correspondence A is defined as $A = \{(x, y) \in X \times X : x \in X, y \in \Gamma(x)\}$.

Note that the max operator is well defined in (BE) since Γ is compact-valued and F is continuous. Besides, if F is bounded, the solution to the supremum function V^* will be also bounded. Therefore, we can try to find a solution (BE) in the space of bounded and continuous functions $\mathcal{C}(X)$ with the supremum norm

$$\|f\|_\infty = \sup_{x \in X} |f(x)|.$$

The functional equation in (BE) implicitly defines an operator on the elements of $\mathcal{C}(X)$ given by

$$(Tf)(x) = \max_{y \in \Gamma(x)} \{F(x, y) + \beta f(y)\}. \quad (\text{T})$$

Finally, given a fixed point $T(V) = V \in \mathcal{C}(X)$, we can characterize the policy correspondence $g : X \rightarrow X$ given by

$$g(x) = \{y \in \Gamma(x) : V(x) = F(x, y) + \beta V(y)\}. \quad (\text{G})$$

Theorem 2.11 (Theorem 4.6 in (Stokey et al., 1989)). Under **B1** and **B2** on X , $\Gamma(x)$, F and β :

- (a) $T : \mathcal{C}(X) \rightarrow \mathcal{C}(X)$,
- (b) T has exactly one fixed point V in $\mathcal{C}(X)$,
- (c) For any $V_0 \in \mathcal{C}(X)$

$$\|T^n V_0 - V\|_\infty \leq \beta^n \|V_0 - V\|_\infty, \quad \text{for } n = 0, 1, \dots$$

- (d) $g : X \rightarrow X$ is non-empty, compact-valued and upper hemi-continuous.

Proof. Under **B1** and **B2**, for each $f \in \mathcal{C}(X)$ and $x \in X$, the problem in (T) is to maximize the continuous function $f(x, \cdot) + \beta f(\cdot)$ over the compact set $\Gamma(x)$. Hence the maximum is attained. As both F and f are bounded, clearly Tf is also bounded, and as F and f are continuous, and Γ is compact-valued and continuous, it follows from the **Theorem of the Maximum of Berge** that Tf is continuous. Therefore $T : \mathcal{C}(X) \rightarrow \mathcal{C}(X)$.

It is trivial to show that T satisfies **Blackwell's Sufficient Conditions**. As $\mathcal{C}(X)$ is a Banach space, it then follows from **Banach Fixed Point Theorem - Contraction Mapping Theorem** that T has a unique fixed point $V \in \mathcal{C}(X)$ and that for any $V_0 \in \mathcal{C}(X)$

$$\|T^n V_0 - V\|_\infty \leq \beta^n \|V_0 - V\|_\infty, \quad \text{for } n = 0, 1, \dots$$

The stated properties on g follow from the **Theorem of the Maximum of Berge** applied to (BE). ■

Assumption (B3). *For all y , $F(\cdot, y)$ is strictly increasing (i.e. F is strictly increasing in the current state. Loosely speaking we can say that more capital does not hurt).*

Assumption (B4). *Γ is monotone, i.e. if $x \leq x'$, then $\Gamma(x) \subseteq \Gamma(x')$ (i.e. the feasibility set is expanding in the current state).*

Theorem 2.12 (Theorem 4.7 in (Stokey et al., 1989)). *Under **B1**, **B2**, **B3** and **B4** on X , $\Gamma(x)$, F and β , V is strictly increasing.*

Proof. Let $\mathcal{C}'(X) \subset \mathcal{C}(X)$ be the set of bounded, continuous, nondecreasing functions on X , and let $\mathcal{C}''(X) \subset \mathcal{C}(X)$ be the set of strictly increasing functions. As $\mathcal{C}'(X)$ is a closed subset of the complete metric space $\mathcal{C}(X)$, by **Theorem 2.11** and **Corollary 2.2** it is sufficient to show that $T[\mathcal{C}'(X)] \subseteq \mathcal{C}''(X)$. Assumptions **B3** and **B4** ensure that this is the case. To see this, let $x_0, x_1 \in X$ such that $x_0 < x_1$. Then

$$\begin{aligned} V(x_0) &= \max_{y \in \Gamma(x_0)} \{F(x_0, y) + \beta V(y)\} \\ &= F(x_0, g(x_0)) + \beta V(g(x_0)), \quad \text{for some } g(x_0) \\ &< \underbrace{F(x_1, g(x_0)) + \beta V(g(x_0))}_{\text{by B3}} \\ &\leq \max_{y \in \Gamma(x_1)} \{F(x_1, y) + \beta V(y)\} = V(x_1), \end{aligned}$$

where the last inequality follows from the fact that $\Gamma(x_0) \subseteq \Gamma(x_1)$ by **B4** (and thus, $g(x_0) \in \Gamma(x_1)$). ■

Assumption (B5). *F is strictly concave. For $(x, y) \in A$ and $(x', y') \in A$ such that $(x, y) \neq (x', y')$ and for all $\lambda \in (0, 1)$,*

$$F(\lambda(x, y) + (1 - \lambda)(x', y')) > \lambda F(x, y) + (1 - \lambda)F(x', y').$$

Assumption (B6). *Γ is convex, i.e. for all $\lambda \in (0, 1)$ and for all $x, x' \in X$ and $y \in \Gamma(x)$ and $y' \in \Gamma(x')$,*

$$\lambda y + (1 - \lambda)y' \in \Gamma(\lambda x + (1 - \lambda)x').$$

Theorem 2.13 (Theorem 4.8 in (Stokey et al., 1989)). Under **B1**, **B2**, **B5** and **B6** on X , $\Gamma(x)$, F and β , V is strictly concave and g is single-valued and continuous as a function.

Proof. Let $\mathcal{C}'(X) \subset \mathcal{C}(X)$ be the set of bounded, continuous, weakly concave functions on X , and let $\mathcal{C}''(X) \subset \mathcal{C}(X)$ be the set of strictly concave functions. As $\mathcal{C}'(X)$ is a closed subset of the complete metric space $\mathcal{C}(X)$, by **Theorem 2.11** and **Corollary 2.2** it is sufficient to show that $T[\mathcal{C}'(X)] \subseteq \mathcal{C}''(X)$. To verify that this is so, let $f \in \mathcal{C}'(X)$, let $x_0, x_1 \in X$ with $x_0 \neq x_1$, and let $\lambda \in (0, 1)$. Define $x_\lambda = \lambda x_0 + (1 - \lambda)x_1$ and let $y_i \in \Gamma(x_i)$ attain $(Tf)(x_i)$ for $i = 0, 1$. Then,

$$\begin{aligned} \lambda(Tf)(x_0) + (1 - \lambda)(Tf)(x_1) &= \lambda[F(x_0, y_0) + \beta V(y_0)] + (1 - \lambda)[F(x_1, y_1) + \beta V(y_1)] \\ &= \underbrace{\lambda F(x_0, y_0) + (1 - \lambda)F(x_1, y_1)}_{< F(\lambda x_0 + (1 - \lambda)x_1, \lambda y_0 + (1 - \lambda)y_1) \text{ by } \mathbf{B5}} + \cdots \\ &\quad \cdots + \beta \underbrace{[\lambda V(y_0) + (1 - \lambda)V(y_1)]}_{\leq f(\lambda y_0 + (1 - \lambda)y_1) \text{ as } f \in \mathcal{C}'(X)} \\ &< (Tf)(\lambda x_0 + (1 - \lambda)x_1) = (Tf)(x_\lambda), \end{aligned}$$

as $\lambda y_0 + (1 - \lambda)y_1 \in \Gamma(x_\lambda)$ by **B6**. Therefore Tf is strictly concave, i.e. $Tf \in \mathcal{C}''(X)$. ■

Theorem 2.14 (Convergence of the policy functions, Theorem 4.9 in (Stokey et al., 1989)). Under **B1**, **B2**, **B5** and **B6** on X , $\Gamma(x)$, F and β , if $V_0 \in \mathcal{C}(X)$ and $\{V_n, g_n\}$ are defined by

$$\begin{aligned} V_{n+1} &= TV_n, \quad n = 0, 1, \dots, \\ g_n(x) &= \arg \max_{y \in \Gamma(x)} \{F(x, y) + \beta V_n(y)\}. \end{aligned}$$

Then $g_n \rightarrow g$ pointwise. If, in addition, X is compact, then the convergence is uniform.

Proof. Let $\mathcal{C}''(X) \subset \mathcal{C}(X)$ be the set of strictly concave functions $f : X \rightarrow \mathbb{R}$. As shown in **Theorem 2.13**, $V \in \mathcal{C}''(X)$. Besides, as shown in the proof of that theorem, $T[\mathcal{C}'(X)] \subseteq \mathcal{C}''(X)$. As $V_0 \in \mathcal{C}'(X)$, it then follows that every function V_n , $n = 1, 2, \dots$ is strictly concave. Define the functions $\{f_n\}$ and f by

$$\begin{aligned} f_n(x, y) &= F(x, y) + \beta V_n(y), \quad n = 1, 2, \dots, \\ f(x, y) &= F(x, y) + \beta V(y). \end{aligned}$$

As F satisfies **B5**, it follows that each function f_n , $n = 1, 2, \dots$ is strictly concave as f is. Therefore **Theorem 2.13** applies and the desired results are proved. ■

Once we have shown that it exists a unique solution $V \in \mathcal{C}(X)$ to the functional equation (BE), we would like to treat the maximum problem in that equation as an ordinary programming problem and use the standard methods of calculus to characterize the policy function g . For example, consider the one-sector neo-classical growth model:

$$V(k) = \max_{k' \in \Gamma(k)} \{u(f(k) - k') + \beta V(k')\}.$$

If we knew that V was differentiable (and the solution to (BE) was always interior), then the policy function g would be given implicitly by the first-order condition

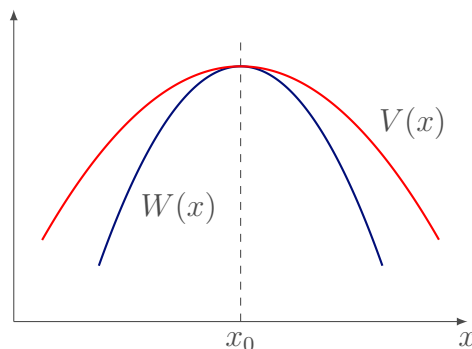
$$-u'(f(k) - g(k)) + \beta V'(g(k)) = 0.$$

Besides, if we knew that V was twice differentiable, the monotonicity of g could be established by differentiating the previous expression with respect to k and examining the resulting expression for g' . However, this ultimately depends upon the differentiability of the functions u , f , V and g . We are free to make any differentiability assumptions on u and f , but the properties of V and g must be established.

Theorem 2.15 (Benveniste and Scheinkman, Theorem 4.10 in (Stokey et al., 1989)). *Let $X \subseteq \mathbb{R}^\ell$ be a convex set, let $V : X \rightarrow \mathbb{R}$ be a concave function, let $x_0 \in \text{int}(X)$ and let D be a neighbourhood of x_0 . If there is a concave, differentiable function $W : D \rightarrow \mathbb{R}$ with $W(x_0) = V(x_0)$ and $W(x) \leq V(x)$, for all $x \in D$, then V is differentiable at x_0 and*

$$V_i(x_0) = W_i(x_0), \quad i = 1, \dots, \ell.$$

Sketch of proof. See next Figure. This Theorem tells us that if we can find a function W for which V is like an Envelope, then if W is differentiable at x_0 so is V with identical derivative.



See Stokey et al. (1989, p.85) for a complete proof. ■

Assumption (B7). F is continuously differentiable in the interior of A .

Theorem 2.16 (Differentiability of the value function, Theorem 4.11 in (Stokey et al., 1989)). *Under Assumptions B1, B2, B5, B6 and B7 on X , $\Gamma(x)$, F and β , if $x_0 \in \text{int}(X)$ and $g(x_0) \in \text{int}[\Gamma(x_0)]$, then V is continuously differentiable at x_0 with derivatives given by*

$$V_i(x_0) = F_i(x_0, g(x_0)), \quad i = 1, \dots, \ell.$$

Proof. The proof relies on **Theorem 2.15**. As $g(x_0) \in \text{int}[\Gamma(x_0)]$ and γ is continuous, it follows that $g(x_0) \in \text{int}[\Gamma(x)]$ for all x in some neighbourhood D of x_0 . Define W on D by

$$W(x) = F(x, g(x_0)) + \beta V(g(x_0)).$$

As F is concave by **B5** and differentiable by **B7**, it follows that W is concave and differentiable. Moreover, as $g(x_0) \in \text{int}[\Gamma(x)]$ for all $x \in D$, it follows that

$$W(x) \leq \max_{y \in \Gamma(x)} \{F(x, y) + \beta V(y)\} = V(x), \quad \text{for all } x \in D,$$

with equality at x_0 . Then V and W satisfy the hypothesis of **Theorem 2.15**, and thus V is differentiable at x_0 with

$$V_i(x_0) = F_i(x_0, g(x_0)), \quad i = 1, \dots, \ell.$$

■

References

- Guner, N. (2008), *Advanced Macroeconomic Theory, Lecture Notes*.
- Juan Pablo Rincón-Zapatero (n.d.), *Mathematics, Lecture Notes*.
- Rincón-Zapatero, J. P. and Santos, M. S. (2009), *Differentiability of the value function without interiority assumptions*, Vol. 144.
- Stokey, N. L., Lucas, R. E. and Prescott, E. (1989), *Recursive Methods in Economic Dynamics*, Harvard University Press.

A Appendix: Envelope Theorem

Let the Bellman equation

$$V_t(X_t) = \sup_{X_{t+1} \in \Gamma_t(X_t)} \{F_t(X_t, X_{t+1}) + \beta V_{t+1}(X_{t+1})\} \quad (33)$$

for $t = 0, \dots, T$, and the policy function (correspondence)

$$g_t(X_t) = \arg \sup_{X_{t+1} \in \Gamma_t(X_t)} \{F_t(X_t, X_{t+1}) + \beta V_{t+1}(X_{t+1})\} \quad (34)$$

Proposition A.1 ((Weak) Envelope Theorem). *Let $V_{t+1}(X_{t+1}) : \mathbb{R} \rightarrow \mathbb{R}$ and $F_t(X_t, X_{t+1}) : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ be differentiable. Then if the optimal policy function (or correspondence) satisfies*

- $g_t(X_t)$ is a differentiable function, and
- $g_t(X_t) \in \text{int}(\Gamma_t(X_t))$

then $V_t(x_t) : \mathbb{R} \rightarrow \mathbb{R}$ is differentiable and

$$\frac{dV_t(X_t)}{dX_t} = \frac{\partial F_t(X_t, X_{t+1})}{\partial X_t} \Bigg|_{\substack{X_t = X^* \\ X_{t+1} = g_t(X^*)}}$$

Proof. Let us assume that everything is differentiable in (33). Then the F.O.C. w.r.t. the choice variable X_{t+1} assuming interior solution at some $X_t = \bar{X}$ is given by

$$\frac{dV_t(X_t)}{dX_{t+1}} \Bigg|_{X_t = \bar{X}} = 0 \Leftrightarrow \frac{\partial F_t(X_t, X_{t+1})}{\partial X_{t+1}} \Bigg|_{\substack{X_t = \bar{X} \\ X_{t+1} = X_{t+1}}} + \frac{dV_{t+1}(X_{t+1})}{dX_{t+1}} \Bigg|_{X_{t+1} = X_{t+1}} = 0 \quad (35)$$

Let us evaluate the Bellman equation (33) in the optimum given by (34), obtaining

$$V_t(X_t) = F_t(X_t, g(X_t)) + \beta V_{t+1}(g(X_t))$$

At any arbitrary point \bar{X}_t , we have that the F.O.C. of this equation w.r.t the state variable X_t is given by

$$\begin{aligned} \frac{dV_t(X_t)}{dX_t} \Bigg|_{X_t = \bar{X}} &= \frac{\partial F_t(X_t, g(X_t))}{\partial X_t} \Bigg|_{\substack{X_t = \bar{X} \\ X_{t+1} = g_t(\bar{X})}} + \frac{\partial F_t(X_t, g(X_t))}{\partial g(X_t)} \Bigg|_{\substack{X_t = \bar{X} \\ X_{t+1} = g_t(\bar{X})}} \cdot \frac{g(X_t)}{X_t} + \\ &+ \frac{dV_{t+1}(g(X_t))}{dg(X_t)} \Bigg|_{\substack{X_t = \bar{X} \\ X_{t+1} = g_t(\bar{X})}} \cdot \frac{g(X_t)}{X_t} \\ &= \left[\frac{\partial F_t(X_t, g(X_t))}{\partial g(X_t)} \Bigg|_{\substack{X_t = \bar{X} \\ X_{t+1} = g_t(\bar{X})}} + \frac{dV_{t+1}(g(X_t))}{dg(X_t)} \Bigg|_{\substack{X_t = \bar{X} \\ X_{t+1} = g_t(\bar{X})}} \right] \frac{g(X_t)}{X_t} + \\ &+ \frac{\partial F_t(X_t, g(X_t))}{\partial X_t} \Bigg|_{\substack{X_t = \bar{X} \\ X_{t+1} = g_t(\bar{X})}} \\ &= \frac{\partial F_t(X_t, g(X_t))}{\partial X_t} \Bigg|_{\substack{X_t = \bar{X} \\ X_{t+1} = g_t(\bar{X})}} \end{aligned}$$

where the last equality comes from (35). Since \bar{X} was arbitrary, it also holds for X^* , and therefore we obtain

$$\frac{dV_t(X_t)}{dX_t} = \frac{\partial F_t(X_t, X_{t+1})}{\partial X_t} \Bigg|_{\substack{X_t = X^* \\ X_{t+1} = g_t(X^*)}}$$

■