# Mathematical Refresher 

# Macroeconomics: Economic Cycles, Frictions and Policy 

Alejandro Riaño

University of Nottingham, GEP, CFCM, and CESifo
September 2, 2019

## Outline

(1) Functions and their properties
(2) Derivatives

- The implicit function theorem
(3) Unconstrained optimization
- Concave and convex functions
(4) Constrained optimization
- The envelope theorem
(5) Markov chains
(6 Dynamic programming
- Stochastic dynamic programming


## Functions

## Definition (Function)

A function from a set $A$ to a set $B$ is a rule that assigns to each object in $A$, one and only one object in B . We write this $f: A \rightarrow B$. $A$ is called the domain; $B$ is called the range or image

Some examples:

- Production function: $Y=\left[K^{\theta}+L^{\theta}\right]^{1 / \theta} . A=\mathbb{R}_{+}^{2}$ and $B=\mathbb{R}_{+}$
- Budget constraint: $m=\sum_{i=1}^{n} p_{i} x_{i} . A=\mathbb{R}_{++}^{n} \times \mathbb{R}_{+}^{n}$ and $B=\mathbb{R}_{+}$


## Definition (One-to-one function)

A function $f: A \rightarrow B$ is one-to-one or injective on a subset $C \subset A$ if and only if for every $x, y \in C$,

$$
f(x)=f(y) \quad \Rightarrow x=y
$$

i.e. each $b \in C$ is the image of precisely one element of $C$. E.g. $y=x^{3}$ is one-to-one while $y=x^{2}$ is not

- When $f: A \rightarrow B$ is one-to-one on $C \subset A, f^{-1}$ (the inverse function of $f$ ) assigns to each $b \in f(C)$ the unique point in $C$ which mapped to it


## Continuity and differentiability

## Definition (Continuity)

Let $f: \mathbb{R}^{k} \rightarrow \mathbb{R}^{m}$ and let $x_{0} \in \mathbb{R}^{k}$ and $y=f\left(x_{0}\right)$ be its image. The function $f$ is continuous at $x_{0}$ if whenever $\left\{x_{n}\right\}_{n=1}^{\infty}$ is a sequence in $\mathbb{R}^{k}$ which converges to $x_{0}$, then the sequence $\left\{f\left(x_{n}\right)\right\}_{n=1}^{\infty}$ in $\mathbb{R}^{m}$ converges to $f\left(x_{0}\right)$. The function $f$ is said to be continuous if it is continuous at every point in its domain
E.g. $y=1 / x$ is continuous everywhere except at $x=0$ because $\lim _{x \rightarrow 0^{+}} f(x) \rightarrow+\infty$ and $\lim _{x \rightarrow 0^{-}} f(x) \rightarrow-\infty$

## Definition (Univariate differentiability)

The derivative of a function $f: \mathbb{R} \rightarrow \mathbb{R}$ at $x_{0}$ is

$$
\begin{equation*}
\frac{\mathrm{d} f}{\mathrm{~d} x}\left(x_{0}\right)=\lim _{h \rightarrow 0} \frac{f\left(x_{0}+h\right)-f\left(x_{0}\right)}{h} \tag{1}
\end{equation*}
$$

The function $f$ is called differentiable if the above limit exists. The derivative is also denoted $f^{\prime}\left(x_{0}\right)$

All differentiable functions are continuous but the converse is not true: $y=|x|$ is not differentiable at $x=0$

## Rules for computing derivatives

Let $f(x), g(x): \mathbb{R} \rightarrow \mathbb{R}$ and $k$ be an arbitrary constant. Then,

$$
\begin{align*}
& \frac{\mathrm{d}}{\mathrm{~d} x}(f(x)+g(x))=f^{\prime}(x)+g^{\prime}(x)  \tag{2}\\
& \frac{\mathrm{d}}{\mathrm{~d} x}(k f(x))=k f^{\prime}(x)  \tag{3}\\
& \frac{\mathrm{d}}{\mathrm{~d} x}(f(x) \cdot g(x))=f^{\prime}(x) g(x)+f(x) g^{\prime}(x)  \tag{4}\\
& \frac{\mathrm{d}}{\mathrm{~d} x}\left(\frac{f(x)}{g(x)}\right)=\frac{f^{\prime}(x) g(x)-f(x) g^{\prime}(x)}{g(x)^{2}}  \tag{5}\\
& \frac{\mathrm{~d}}{\mathrm{~d} x}\left(x^{k}\right)=k x^{k-1}  \tag{6}\\
& \frac{\mathrm{~d}}{\mathrm{~d} x}(\ln x)=1 / x  \tag{7}\\
& \frac{\mathrm{~d}}{\mathrm{~d} x}(\exp (x))=\exp (x) \tag{8}
\end{align*}
$$

## Multivariate differentiation

## Definition (Partial derivative)

Let $f: R^{n} \rightarrow R$. Then for each variable $x_{i}$ at each point $x^{0}=\left(x_{1}^{0}, \ldots, x_{n}^{0}\right)$, define the partial derivative of $f$ with respect to $x_{i}$ as

$$
\begin{equation*}
\frac{\partial f}{\partial x_{i}}\left(x_{1}^{0}, \ldots, x_{n}^{0}\right)=\lim _{h \rightarrow 0} \frac{f\left(x_{1}^{0}, \ldots, x_{i}^{0}+h, \ldots, x_{n}^{0}\right)-f\left(x_{1}^{0}, \ldots, x_{i}^{0}, \ldots, x_{n}^{0}\right)}{h} \tag{9}
\end{equation*}
$$

if this limit exists. The partial derivative of $F$ with respect to $x_{i}$ is also denoted $F_{x_{i}}$ or $F_{i}$.

Example: let $f(x, y)=3 x^{2} y^{3}$, then:

$$
\begin{aligned}
& \frac{\partial f}{\partial x}=6 x y^{3} \\
& \frac{\partial f}{\partial y}=9 x^{2} y^{2}
\end{aligned}
$$

## Implicit functions

- An implicit function of the endogenous variable $y$ as a function of the exogenous variables $\left(x_{1}, \ldots, x_{n}\right)$ is written as $G\left(x_{1}, \ldots, x_{n}, y\right)=0$
- The implicit function theorem (IFT) allows us to determine how the endogenous variable changes in response to a change in one of the exogenous variables


## Definition (Implicit function theorem)

Let $G\left(x_{1}, \ldots, x_{n}, y\right)$ be a continuously differentiable ( $\mathcal{C}^{1}$ ) function around $\left(x_{1}^{*}, \ldots, x_{n}^{*}, y^{*}\right)$. Assume further that $\left(x_{1}^{*}, \ldots, x_{n}^{*}, y^{*}\right)$ satisfies:

$$
\begin{aligned}
G\left(x_{1}^{*}, \ldots, x_{n}^{*}, y^{*}\right) & =c \\
\frac{\partial G}{\partial y}\left(x_{1}^{*}, \ldots, x_{n}^{*}, y^{*}\right) & \neq 0
\end{aligned}
$$

Then there is a $\mathcal{C}^{1}$ function $y=y\left(x_{1}, \ldots, x_{n}\right.$ defined on an open ball around $\left(x_{1}^{*}, \ldots, x_{n}^{*}\right)$ such that for each $i \in\{1, \ldots, n\}$ :

$$
\begin{equation*}
\frac{\partial y}{\partial x_{i}}\left(x_{1}^{*}, \ldots, x_{n}^{*}\right)=-\frac{\frac{\partial G}{\partial x_{i}}\left(x_{1}^{*}, \ldots, x_{n}^{*}, y^{*}\right)}{\frac{\partial G}{\partial y}\left(x_{1}^{*}, \ldots, x_{n}^{*}, y^{*}\right)} \tag{10}
\end{equation*}
$$

## Example

- Consider an individual with utility function defined over two goods $x$ and $y$ : $U(x, y)$
- An indifference curve is defined as $\{(x, y): U(x, y)=\bar{U}\}$
- We can use the IFT to calculate the slope of the indifference curve (the marginal rate of substitution):

$$
\begin{gathered}
\frac{\partial U}{\partial x} \mathrm{~d} x+\frac{\partial U}{\partial y} \mathrm{~d} y=0 \\
\frac{\mathrm{~d} y}{\mathrm{~d} x}=-\frac{\frac{\partial U}{\partial x}}{\frac{\partial U}{\partial y}}
\end{gathered}
$$

- Assuming that utility is strictly increasing in both goods $\rightarrow$ MRS is decreasing


## Unconstrained optimization

- Let $U \subset \mathbb{R}^{n}$ and $F: U \rightarrow \mathbb{R}$. A point $x^{*} \in U$ is a max of $F$ on $U$ if $F\left(x^{*}\right) \geq F(x) \forall x \in U$


## Theorem (First-order necessary conditions)

If $x^{*}$ is a local critical point of $F$ and $x^{*}$ is an interior point of $U$ then

$$
\begin{equation*}
\frac{\partial F}{\partial x_{i}}\left(x^{*}\right)=0 \quad \forall i=1, \ldots, n \tag{11}
\end{equation*}
$$

- Example: Let $F(x, y)=x^{2}-6 x y+2 y^{2}+10 x+2 y-5$; its critical points are found by solving:

$$
\begin{aligned}
& F_{x}=2 x-6 y+10=0 \\
& F_{y}=-6 x+4 y+2=0
\end{aligned}
$$

which yields: $\left(x^{*}, y^{*}\right)=(13 / 7,6 / 7)$

## Unconstrained optimization

## Theorem (Second-order sufficient conditions)

Let $F: U \rightarrow \mathbb{R}$ be twice continuously differentiable ( $\mathcal{C}^{2}$ ) in $U$, and let $x^{*}$ be a critical point of $F$. Define the Hessian of $F$ as

$$
\mathcal{H} \equiv D^{2} F\left(x^{*}\right)=\left(\begin{array}{ccc}
\frac{\partial^{2} F}{\partial x_{1}^{2}}\left(x^{*}\right) & \cdots & \frac{\partial^{2} F}{\partial x_{n} x_{1}}\left(x^{*}\right)  \tag{12}\\
\vdots & \ddots & \vdots \\
\frac{\partial^{2} F}{\partial x_{1} x_{n}}\left(x^{*}\right) & \cdots & \frac{\partial^{2} F}{\partial x_{n}^{2}}\left(x^{*}\right)
\end{array}\right)
$$

1. If $\mathcal{H}$ is negative definite, i.e. if $x^{\prime} \mathcal{H} x<0 \forall x \in U$, then $x^{*}$ is a strict local max of $F$
2. If $\mathcal{H}$ is positive definite, i.e. if $x^{\prime} \mathcal{H} x>0 \forall x \in U$, then $x^{*}$ is a strict local min of $F$
3. If $\mathcal{H}$ is indefinite, the $x^{*}$ is neither a local max nor a local min of $F$

## Unconstrained optimization

## Theorem

Let $f: \mathbb{R}^{n} \rightarrow \mathbb{R}$ be $\mathcal{C}^{2}$ and let $x^{*}$ be a critical point of $f$ (i.e. $f_{x_{i}}=0 \quad \forall i=1, \ldots, n$ ) and that the $n$ leading principal minors of $\mathcal{H}$ alternate in sign:

$$
\left|f_{x_{1} x_{1}}\right|<0, \quad\left|\begin{array}{ll}
f_{x_{1} x_{1}} & f_{x_{2} x_{1}}  \tag{13}\\
f_{x_{1} x_{2}} & f_{x_{2} x_{2}}
\end{array}\right|>0,\left|\begin{array}{ccc}
f_{x_{1} x_{1}} & f_{x_{2} x_{1}} & f_{x_{3} x_{1}} \\
f_{x_{1} x_{2}} & f_{x_{2} x_{2}} & f_{x_{3} x_{2}} \\
f_{x_{1} x_{3}} & f_{x_{2} x_{3}} & f_{x_{3} x_{3}}
\end{array}\right|<0, \ldots
$$

at $x^{*}$. Then $x^{*}$ is a strict local max of $f$. If all the leading principal minors of $\mathcal{H}$ are all positive then $x^{*}$ is a strict local min of $f$

## Concave and convex functions

## Definition (Concave and Convex functions)

A real-valued function $f: \mathbb{R}^{n} \rightarrow \mathbb{R}$ is concave if $\forall x, y \in \mathbb{R}^{n}$ and $\lambda \in[0,1]$

$$
\begin{equation*}
f(\lambda x+(1-\lambda) y) \geq \lambda f(x)+(1-\lambda) f(y) \tag{14}
\end{equation*}
$$

$f$ is convex if $\forall x, y \in \mathbb{R}^{n}$ and $\lambda \in[0,1]$

$$
\begin{equation*}
f(\lambda x+(1-\lambda) y) \leq \lambda f(x)+(1-\lambda) f(y) \tag{15}
\end{equation*}
$$

## Theorem <br> Let $f: \mathbb{R}^{n} \rightarrow \mathbb{R}$ be $\mathcal{C}^{2}$. Then $f$ is concave if and only if its Hessian matrix is negative semidefinite $\forall x \in \mathbb{R}^{n}$. $f$ is convex if and only if its Hessian matrix is positive semidefinite $\forall x \in \mathbb{R}^{n}$

## Constrained optimization

- We are interested in solving problems of the type:

$$
\begin{gather*}
\max _{x_{1}, \ldots, x_{n}} f\left(x_{1}, \ldots, x_{n}\right)  \tag{16}\\
\quad \text { subject to: } \\
g_{1}\left(x_{1}, \ldots, x_{n}\right) \leq b_{1}, \ldots, g_{k}\left(x_{1}, \ldots, x_{n}\right) \leq b_{k}  \tag{17}\\
h_{1}\left(x_{1}, \ldots, x_{n}\right)=c_{1}, \ldots, h_{k}\left(x_{1}, \ldots, x_{n}\right)=c_{m} \tag{18}
\end{gather*}
$$

- $f$ is the objective function
- $\left\{g_{i}\right\}$ are inequality constraints
- $\left\{h_{j}\right\}$ are equality constraints
- $x_{i} \geq 0$ non-negativity constraints


## Equality constraints

## Theorem (Lagrangian)

Let $f, h_{1}, \ldots, h_{m}$ be $\mathcal{C}^{1}$ functions. Suppose that $\mathbf{x}^{*}$ is a local max of $f$ in the constraint set $\left\{\mathbf{x}: \quad h_{1}(\mathbf{x})=c_{1}, \ldots, h_{m}(\mathbf{x})=c_{m}\right\}$ and assume that rank of the Jacobian matrix of first-derivatives of the constraints with respect to $\mathbf{x}$,

$$
\mathcal{J} \equiv D h\left(\mathbf{x}^{*}\right)=\left(\begin{array}{ccc}
\partial h_{1} / \partial x_{1} & \cdots & \partial h_{1} / \partial x_{n}  \tag{19}\\
\vdots & \ddots & \vdots \\
\partial h_{m} / \partial x_{1} & \cdots & \partial h_{m} / \partial x_{n}
\end{array}\right)
$$

is $m$. Then there exist $\lambda_{1}^{*}, \ldots, \lambda_{m}^{*}$ such that $\left(x_{1}^{*}, \ldots, x_{n}^{*}, \lambda_{1}^{*}, \ldots, \lambda_{m}^{*}\right)$ is a critical point of the Lagrangian

$$
\begin{equation*}
\mathcal{L}(\mathbf{x}, \lambda) \equiv f(\mathbf{x})-\sum_{j=1}^{m} \lambda_{j}\left[h_{j}(\mathbf{x})-c_{j}\right] \tag{20}
\end{equation*}
$$

That is,

$$
\begin{aligned}
\frac{\partial \mathcal{L}}{\partial x_{i}}\left(\mathbf{x}^{*}, \lambda^{*}\right) & =0 \quad \forall i=1, \ldots, n \\
\frac{\partial \mathcal{L}}{\partial \lambda_{j}}\left(\mathbf{x}^{*}, \lambda^{*}\right) & =0 \quad \forall j=1, \ldots, m
\end{aligned}
$$

## Example

$$
\begin{gathered}
\max _{x_{1}, x_{2}} U=\alpha \ln x_{1}+(1-\alpha) \ln x_{2} \\
\text { s.t.: } \\
p_{1} x_{1}+p_{2} x_{2}=m
\end{gathered}
$$

- Lagrangian: $\mathcal{L}=\alpha \ln x_{1}+(1-\alpha) \ln x_{2}-\lambda\left[p_{1} x_{1}+p_{2} x_{2}-m\right]$
- FOC:

$$
\begin{aligned}
{\left[x_{1}\right]: } & \alpha / x_{1}-\lambda p_{1}=0 \\
{\left[x_{2}\right]: } & (1-\alpha) / x_{2}-\lambda p_{2}=0 \\
{[\lambda]: } & p_{1} x_{1}+p_{2} x_{2}-m=0
\end{aligned}
$$

- Using the first two FOC: $p_{2} x_{2}=\frac{1-\alpha}{\alpha} p_{1} x_{1}$. Plugging this back into the the third FOC yields:

$$
\left(x_{1}^{*}, x_{2}^{*}, \lambda^{*}\right)=\left(\frac{\alpha m}{p_{1}}, \frac{(1-\alpha) m}{p_{2}}, \frac{1}{m}\right)
$$

- Notice that $\mathcal{J}\left(\mathbf{x}^{*}\right)=\left(\begin{array}{ll}p_{1} & p_{2}\end{array}\right)$ has rank 1


## Inequality constraints

- We want to solve the following problem:

$$
\begin{gather*}
\max _{x_{1}, \ldots, x_{n}} f\left(x_{1}, \ldots, x_{n}\right)  \tag{21}\\
\quad \text { subject to: } \\
g_{1}\left(x_{1}, \ldots, x_{n}\right) \leq b_{1}, \ldots, g_{K}\left(x_{1}, \ldots, x_{n}\right) \leq b_{K}  \tag{22}\\
x_{1} \geq 0, \ldots, x_{n} \geq 0 \tag{23}
\end{gather*}
$$

The Lagrangian is given by:

$$
\begin{equation*}
\mathcal{L}=f(\mathbf{x})-\sum_{k=1}^{K} \lambda_{k}\left[g_{k}(\mathbf{x})-b_{k}\right]+\sum_{i=1}^{n} \mu_{i} x_{i} \tag{24}
\end{equation*}
$$

- The FOC are:

$$
\left.\left.\begin{array}{rl}
\frac{\partial \mathcal{L}}{\partial x_{i}}\left(x^{*}\right) & \leq 0, \quad \forall i=1, \ldots, n \\
x_{i} \frac{\partial \mathcal{L}}{\partial x_{i}}\left(x^{*}\right) & =0, \quad \forall i=1, \ldots, n \\
\frac{\partial \mathcal{L}}{\partial \lambda_{k}}\left(x^{*}\right) & \geq 0, \quad \forall k=1, \ldots, K \\
\lambda_{k} \frac{\partial \mathcal{L}}{\partial \lambda_{k}}\left(x^{*}\right) & =0, \quad \forall k \\
\lambda_{k}, \mu_{i} & \geq 0, \quad \forall k
\end{array}\right)=1, \ldots, K, \quad i=1, \ldots, n\right) ~ l
$$

## Exercise

- An individual enjoys utility from consumption $c$ and leisure $l$; he has one unit of time that he can devote to work $n$ at the wage $w$ or enjoy leisure
- Find the optimal allocation of consumption, leisure and working hours that maximizes the consumer's utility subject to his budget and time allocation constraints, i.e.

$$
\begin{gathered}
\max _{c, n, l} \log c+a l, \quad a>0 \\
\text { s.t.: } \\
c \leq w n \\
l+n \leq 1 \\
c, n, l \geq 0
\end{gathered}
$$

- Write the Lagrangian of the problem
- Show that the non-negativity constraint for consumption cannot be binding
- Under what conditions will the consumer devote all his time to working?


## Envelope theorem

- How does the optimal value of an objective function changes as one parameter changes?


## Theorem (Envelope theorem)

Let $f, h_{1}, \ldots, h_{k}: \mathbb{R}^{n} \times \mathbb{R} \rightarrow \mathbb{R}$ be $\mathcal{C}^{1}$ functions. Let $\mathbf{x}^{*}(a)$ denote the solution of the problem of maximizing $f(\mathbf{x} ; a)$ on the constraint set $\left\{h_{1}(\mathbf{x} ; a)=0, \ldots, h_{k}(\mathbf{x} ; a)=0\right\}$ for any fixed parameter $a$. Then,

$$
\begin{equation*}
\frac{d}{d a} f\left(\mathrm{x}^{*}(a) ; a\right)=\frac{\partial L}{\partial a}\left(\mathrm{x}^{*}(a), \mu(a) ; a\right) \tag{25}
\end{equation*}
$$

- Go back to the example in Slide \# 18. We want to know how does the utility of the individual changes when his income increases
- Plugging $\left(x_{1}^{*}, x_{2}^{*}\right)$ into the objective function, you obtain the 'value function':

$$
\begin{aligned}
V\left(p_{1}, p_{2}, m\right)=\alpha \ln & \left(\frac{\alpha m}{p_{1}}\right)+(1-\alpha) \ln \left(\frac{(1-\alpha) m}{p_{2}}\right) \\
& \rightarrow \frac{\partial V}{\partial m}=\frac{1}{m}
\end{aligned}
$$

- Which is equal to $\frac{\partial \mathcal{L}}{\partial m}=\lambda^{*}=\frac{1}{m}$


## Markov chains

## Definition

A time-invariant Markov chain is defined by a triple of objects:
(i) A vector of dimension $n \times 1, \bar{x} \in \mathbb{R}^{n}$, that records the possible values of the state of the system,
(ii) An $n \times n$ matrix $P$ which records the probabilities of moving from one value of the state to another in one period,

$$
P_{i j}=\operatorname{Prob}\left(x_{t+1}=\bar{x}_{j} \mid x_{t}=\bar{x}_{i}\right)
$$

(iii) A $n \times 1$ vector $\pi_{0}$ recording the probabilities of being in each state $i$ at time 0 ,

$$
\pi_{0}=\operatorname{Prob}\left(x_{0}=\bar{x}_{i}\right)
$$

We need $P$ and $\pi_{0}$ to satisfy the following conditions:
(i) For all $i=1, \ldots, n, \sum_{j=1}^{n} P_{i j}=1$
(ii) $\sum_{i=1}^{n} \pi_{0 i}=1$
$P$ is called a stochastic matrix

## Example

- Suppose an individual can be in either of two states: employed, $e$, and unemployed, $u$
- Also, suppose that at the beginning of her life this individual has an equal chance to be employed or unemployed
- Every period, the probability that an unemployed individual stays unemployed is $50 \%$. The probability that an employed individual moves into unemployment is 4.3\%
- Let's characterize the Markov chain governing this individual's employment status:

$$
\begin{aligned}
\bar{x} & =\{u, e\} \\
\pi_{0} & =\left[\begin{array}{ll}
0.5 & 0.5
\end{array}\right]^{\prime} \\
P & =\left[\begin{array}{ll}
p_{u u} & p_{u e} \\
p_{e u} & p_{e e}
\end{array}\right]=\left[\begin{array}{cc}
0.5 & 0.5 \\
0.043 & 0.957
\end{array}\right]
\end{aligned}
$$

(state space)
(initial state probability)
(stochastic matrix)

## An example - a realization of $\left\{x_{t}\right\}$



## Asymptotic distribution of $P$



- If, at any given point in time, you take a large sample of individuals whose employment status is governed by the the Markov process I showed you before, you'd find that approximately $92 \%$ would be employed, and the remaining $8 \%$ would be unemployed


## Dynamic programming: intuition

- A dynamic programming (DP) problem is an optimization problem in which decisions have to be taken sequentially over several time periods
- Periods are usually 'linked' in the sense that actions taken at a particular point affect the reward possibilities in future periods
- In practice this is achieved by defining a state variable, which restricts the set of actions available to the decision-maker at any point in time
- In turn, the decision-maker's actions affect the value of the state variable(s)


## Dynamic programming: definitions

## Definition

A DP problem is defined by a tuple $\{X, \Gamma(x), F, \beta\}$, where:
(i) $X$ is the state space of the problem,
(ii) $\Gamma(x)$ is the feasible set given a particular state $x$,
(iii) $F$ is the one-period reward function,
(iv) $\beta$ is the discount factor

## Dynamic programming

- We are interested in solving problems of the form:

$$
\begin{gather*}
\max _{\left\{x_{t+1}\right\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^{t} F\left(x_{t}, x_{t+1}\right)  \tag{26}\\
\text { s.t.: } \\
x_{t+1} \in \Gamma\left(x_{t}\right), \quad t=0,1,2, \ldots  \tag{27}\\
x_{0} \in X \text { given. } \tag{28}
\end{gather*}
$$

- This representation of the problem is called a sequential problem
- i.e. the decision-maker begins from some fixed state $x_{0} \in X$ at $t=0$
- The set of actions available at $t=0$ is given by $x_{t+1} \in \Gamma\left(x_{0}\right)$
- After choosing $x_{1}$, the decision-maker obtains a reward $F\left(x_{0}, x_{1}\right)$
- In $t=2$, the available options for the decision-maker are $x_{t+2} \in \Gamma\left(x_{1}\right)$
- The problem in period 2 looks exactly the same, the only difference is that we start with state $x_{1}$ instead of $x_{0}$. And so, we keep going...


## Deterministic neoclassical growth model

- This is how the SP problem looks like:

$$
\begin{gather*}
\max _{\left\{c_{t}, i_{t}\right\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^{t} \ln \left(c_{t}\right)  \tag{29}\\
\text { s.t.: } \\
y_{t}=k_{t}^{\alpha}  \tag{30}\\
c_{t}+i_{t}=y_{t}  \tag{31}\\
k_{t+1}=(1-\delta) k_{t}+i_{t}  \tag{32}\\
k_{0}>0 \text { given } \tag{33}
\end{gather*}
$$

- Simplifying:

$$
\begin{equation*}
\max _{\left\{k_{t+1}\right\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^{t} \ln \left[k_{t}^{\alpha}+(1-\delta) k_{t}-k_{t+1}\right] \tag{34}
\end{equation*}
$$

- Notice that the individual enters every period with a given capital stock $k_{t}$, and then he chooses $k_{t+1}$. Knowing both $k_{t}$ and $k_{t+1}$ determines the level of consumption
- In this representation, $k_{t}$ is the state variable and $k_{t+1}$ is the action/control variable


## Dynamic programming: intuition

- Suppose that the problem above has already been solved for every value of $x_{0}$
- Then we could define a function $v: X \rightarrow \mathbb{R}$ as follows:

$$
\begin{equation*}
v\left(x_{0}\right)=\max _{x \in \Gamma\left(x_{0}\right)} \sum_{t=0}^{\infty} \beta^{t} F\left(x_{t}, x_{t+1}\right) \tag{35}
\end{equation*}
$$

- We call $v\left(x_{0}\right)$ a value function. It tells us the value of the maximized objective function for a given value of the state variable
- Similarly, $v\left(x_{1}\right)$ would give us the maximum utility value that the decision-maker can achieve from period 1 afterwards given his/her state $x_{1}$ Then,

$$
\begin{equation*}
v\left(x_{0}\right)=\max _{x_{1} \in \Gamma\left(x_{0}\right)}\{\underbrace{F\left(x_{0}, x_{1}\right)+\beta v\left(x_{1}\right)}_{=\sum_{t=0}^{\infty} \beta^{t} F\left(x_{t}, x_{t+1}\right)}\} \tag{36}
\end{equation*}
$$

- This is called the recursive representation of a DP
- A solution for the DP problem specified above is a policy rule, $g: X \rightarrow \Gamma(x)$ which specifies the optimal action at each stage as a function of the state variable


## The Principle of optimality

## Theorem

The value function $v$ satisfies the Bellman equation at each $x \in X$ :

$$
v(x)=\max _{x^{\prime} \in \Gamma(x)}\left\{F\left(x, x^{\prime}\right)+\beta v\left(x^{\prime}\right)\right\}
$$

where $x^{\prime}$ denotes the next-period value of $x$

- Note that the solution to $(\star)$ is a function, not just a number. This is called a functional equation


## The Principle of optimality

- Here's the cool thing (if you throw in some extra technical conditions), we can take whatever initial guess for $v$, call it $v_{0} \ldots$
- and we can iterate on the following recursion:

$$
\begin{equation*}
v_{j+1}=\max _{x \in \Gamma(x)}\left\{F\left(x, x^{\prime}\right)+\beta v_{j}\left(x^{\prime}\right)\right\} \tag{37}
\end{equation*}
$$

until $\left\|v_{j+1}-v_{j}\right\|<\varepsilon$ and we will converge to the solution of $(\star)$ !

- You can see that these conditions make this problem very amenable to be solved in a computer


## Deterministic neoclassical growth model

- Now let's write the problem in a recursive way (write up the Bellman equation). Let ' denote the future value of a variable
- Notice first that $c>0$ this means that given $k, k^{\prime} \in[0, f(k)+(1-\delta) k) \Rightarrow$ $\Gamma(k)=[0, f(k)+(1-\delta) k)$
- The Bellman equation of this problem is then:

$$
\begin{equation*}
v(k)=\max _{k^{\prime} \in\left[0, k^{\alpha}+(1-\delta) k\right)}\left\{\ln \left[k^{\alpha}+(1-\delta) k-k^{\prime}\right]+\beta v\left(k^{\prime}\right)\right\} \tag{38}
\end{equation*}
$$

## Intuition on how to solve this in the computer

- Create a grid of $N$ points for the state variable $k, \mathcal{K}=\left[k_{1}, k_{2}, \ldots, k_{N}\right]$ around the steady state
- We want to find the value function $v(k)$. This function will provide a value for each point in the grid. Hence $v(k)$ will be a $N \times 1$. vector $v(k)=\left[v_{1}, v_{2}, \ldots, v_{N}\right]$
- Start with an initial guess for $v(k)$ (it can be whatever you want, for instance the zero vector), call it $v_{0}(k)$
- Now we construct the recursion over which we're going to iterate. For every $k$ today, we will choose $k^{\prime}$ that maximizes

$$
\begin{equation*}
\ln \left(k^{\alpha}+(1-\delta) k-k^{\prime}\right)+\beta v_{0}\left(k^{\prime}\right) \tag{39}
\end{equation*}
$$

we're going to create a big $N \times N$ matrix, where future capital $k^{\prime}$ goes in the rows and current capital $k$ goes in the columns. Let me show you how the first component of it looks like.

## Intuition on how to solve this in the computer

$\ln \left(k^{\alpha}+(1-\delta) k-k^{\prime}\right)$ looks like this:

$$
\left[\begin{array}{cccc}
\ln \left[k_{1}^{\alpha}+(1-\delta) k_{1}-k_{1}\right] & \ln \left[k_{2}^{\alpha}+(1-\delta) k_{2}-k_{1}\right] & \cdots & \ln \left[k_{N}^{\alpha}+(1-\delta) k_{N}-k_{1}\right] \\
\ln \left[k_{1}^{\alpha}+(1-\delta) k_{1}-k_{2}\right] & \ln \left[k_{2}^{\alpha}+(1-\delta) k_{2}-k_{2}\right] & \cdots & \ln \left[k_{N}^{\alpha}+(1-\delta) k_{N}-k_{2}\right] \\
\vdots & \vdots & \ddots & \vdots \\
\ln \left[k_{1}^{\alpha}+(1-\delta) k_{1}-k_{N}\right] & \ln \left[k_{2}^{\alpha}+(1-\delta) k_{2}-k_{N}\right] & \cdots & \ln \left[k_{N}^{\alpha}+(1-\delta) k_{N}-k_{N}\right]
\end{array}\right]
$$

For the second component, notice two things: (i) $v_{0}\left(k^{\prime}\right)$ is of size $N \times 1$ and (ii) it only depends on future capital $k^{\prime}$

$$
\beta v_{0}(k) \cdot[1 \cdots 1]_{1 \times N}=\beta\left[\begin{array}{cccc}
v_{0}\left(k_{1}\right) & v_{0}\left(k_{1}\right) & \cdots & v_{0}\left(k_{1}\right) \\
v_{0}\left(k_{2}\right) & v_{0}\left(k_{2}\right) & \cdots & v_{0}\left(k_{2}\right) \\
\vdots & \vdots & \ddots & \vdots \\
v_{0}\left(k_{N}\right) & v_{0}\left(k_{N}\right) & \cdots & v_{0}\left(k_{N}\right)
\end{array}\right]
$$

You add these two components, and then for each column (current capital) you choose the row (future capital) that maximizes this objective function

## Intuition on how to solve this in the computer

- For each column $i=1, \ldots N$ you will have a new maximized value. So you end up with a new vector $v_{1}(k)=\left[v_{1}\left(k_{1}\right) \cdots v_{1}\left(k_{N}\right)\right]$
- Then, you compare $v_{0}(k)$ and $v_{1}(k)$, if they're close enough $\left\|v_{1}(k)-v_{0}(k)\right\|<\varepsilon$, stop. You've solved the problem. Otherwise, continue iterating
- You also know the maximizing value of future capital (for example, for current capital $k=k_{3}$, the level of future capital that maximizes the objective is $k^{\prime}=k_{5}$ ).
- When you have achieved convergence, you also find the policy rule, that tells you what should be the optimal future capital $k^{\prime}$ for whatever value of current capital $k$


## Stochastic dynamic programming

- What happens when we add stochastic shocks to our DP problem?
- For instance, in the one-sector growth model that we just saw, we can add productivity shocks that affect production every period
- More specifically, let $z_{t}$ be an $\operatorname{AR}(1)$ process:

$$
\begin{equation*}
z_{t}=\rho z_{t-1}+\varepsilon_{t}, \quad \varepsilon \stackrel{\text { i.i.d. }}{\sim} \mathcal{N}\left(0, \sigma_{\varepsilon}^{2}\right) \tag{40}
\end{equation*}
$$

- Production now looks like this: $y_{t}=\exp \left(z_{t}\right) k_{t}^{\alpha}$
- Now we combine dynamic programming with Markov chains!


## Stochastic dynamic programming

- Timing: assume that in period $t$, before you make a decision of how much capital to choose for $t+1$, you observe a realization $\varepsilon_{t}$ drawn from a Normal distribution with mean 0 and variance $\sigma_{\varepsilon}^{2}$
- Since you already knew $z_{t-1}$, you can compute $z_{t}=\rho z_{t-1}+\varepsilon_{t}$
- $z$ becomes a new state variable in our dynamic programming in addition to $k$
- To optimally choose $k_{t+1}$ all you need to know is your capital in period $t, k_{t}$, and the value of the productivity shock in period $t, z_{t}$
- $z_{t}$ is called an exogenous state, since its not influenced by the actions of the decision-maker
- The other important difference, is that now the decision-maker needs to forecast the future evolution of $z$ since we're now interested in the expected value of our value function conditional on $z_{t}, \mathbb{E}_{t}\left[V\left(k^{\prime}, z^{\prime}\right)\right]$


## Stochastic dynamic programming

- The Bellman equation of our new problem is:

$$
\begin{equation*}
v(k, z)=\max _{k^{\prime} \in\left[0, \exp (z) k^{\alpha}+(1-\delta) k\right)}\left\{\ln \left[\exp (z) k^{\alpha}+(1-\delta) k-k^{\prime}\right]+\beta \sum_{z^{\prime}} P\left(z^{\prime} \mid z\right) V\left(k^{\prime}, z^{\prime}\right)\right\} \tag{41}
\end{equation*}
$$

- $P$ is the stochastic matrix associated to the Markov-chain approximation of the $\left\{z_{t}\right\}$
- We can solve this problem in a similar fashion as we solved the deterministic problem. The main difference is that the state space is bigger (when we are creating the matrices that I showed you before, the dimensions are going to be $N \times N_{z}$, where $N_{z}$ is the number of grid points for the approximation for $z_{t}$
- As you can see, the number of points in the state space grows exponentially. You cannot add too many state variables to your problem, otherwise it will be impossible to solve it! This is called the Curse of Dimensionality


## References

- Simon, C. P. and L. E. Blume Mathematics for Economists, 2010.
- Adda, J. and R. W. Coooper Dynamic Economics, 2003.

