Mathematical Refresher

Macroeconomics: Economic Cycles, Frictions and Policy

Alejandro Riaño

University of Nottingham, GEP, CFCM, and CESifo

September 2, 2019

Outline

- Functions and their properties
- 2 Derivatives
 - The implicit function theorem
- **3** Unconstrained optimization
 - Concave and convex functions
- Onstrained optimization
 - The envelope theorem
- 6 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 $f : A \to B$. 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 $f:A\to B$ is one-to-one or injective on a subset $C\subset A$ if and only if for every $x,y\in C$,

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

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 → B is one-to-one on C ⊂ A, f⁻¹ (the inverse function of f) assigns to each b ∈ f(C) the unique point in C which mapped to it

Continuity and differentiability

Definition (Continuity)

Let $f : \mathbb{R}^k \to \mathbb{R}^m$ and let $x_0 \in \mathbb{R}^k$ and $y = f(x_0)$ be its image. The function f is continuous at x_0 if whenever $\{x_n\}_{n=1}^{\infty}$ is a sequence in \mathbb{R}^k which converges to x_0 , then the sequence $\{f(x_n)\}_{n=1}^{\infty}$ in \mathbb{R}^m converges to $f(x_0)$. 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 $\lim_{x\to 0^+} f(x) \to +\infty$ and $\lim_{x\to 0^-} f(x) \to -\infty$

Definition (Univariate differentiability)

The derivative of a function $f:\mathbb{R}\to\mathbb{R}$ at x_0 is

$$\frac{df}{dx}(x_0) = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}$$
(1)

The function f is called differentiable if the above limit exists. The derivative is also denoted $f^\prime(x_0)$

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

Rules for computing derivatives

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

$$\frac{d}{dx}(f(x) + g(x)) = f'(x) + g'(x)$$
(2)

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(kf(x)\right) = kf'(x) \tag{3}$$

$$\frac{\mathsf{d}}{\mathsf{d}x}\left(f(x)\cdot g(x)\right) = f'(x)g(x) + f(x)g'(x) \tag{4}$$

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{f(x)}{g(x)}\right) = \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2} \tag{5}$$

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(x^{k}\right) = kx^{k-1} \tag{6}$$

$$\frac{\mathsf{d}}{\mathsf{d}x}\left(\ln x\right) = 1/x\tag{7}$$

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\exp(x)\right) = \exp(x) \tag{8}$$

Multivariate differentiation

Definition (Partial derivative)

Let $f : \mathbb{R}^n \to \mathbb{R}$. Then for each variable x_i at each point $x^0 = (x_1^0, \dots, x_n^0)$, define the partial derivative of f with respect to x_i as

$$\frac{\partial f}{\partial x_i}(x_1^0, \dots, x_n^0) = \lim_{h \to 0} \frac{f(x_1^0, \dots, x_i^0 + h, \dots, x_n^0) - f(x_1^0, \dots, x_i^0, \dots, x_n^0)}{h}$$
(9)

if this limit exists. The partial derivative of F with respect to x_i is also denoted ${\cal F}_{x_i}$ or ${\cal F}_i.$

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

$$\frac{\partial f}{\partial x} = 6xy^3$$
$$\frac{\partial f}{\partial y} = 9x^2y^2$$

Implicit functions

- An implicit function of the endogenous variable y as a function of the exogenous variables (x_1, \ldots, x_n) is written as $G(x_1, \ldots, x_n, y) = 0$
- 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 $G(x_1, \ldots, x_n, y)$ be a continuously differentiable (\mathcal{C}^1) function around $(x_1^*, \ldots, x_n^*, y^*)$. Assume further that $(x_1^*, \ldots, x_n^*, y^*)$ satisfies:

$$G(x_1^*, \dots, x_n^*, y^*) = c$$
$$\frac{\partial G}{\partial y}(x_1^*, \dots, x_n^*, y^*) \neq 0$$

Then there is a C^1 function $y = y(x_1, \ldots, x_n$ defined on an open ball around (x_1^*, \ldots, x_n^*) such that for each $i \in \{1, \ldots, n\}$:

$$\frac{\partial y}{\partial x_i}(x_1^*,\dots,x_n^*) = -\frac{\frac{\partial G}{\partial x_i}(x_1^*,\dots,x_n^*,y^*)}{\frac{\partial G}{\partial y}(x_1^*,\dots,x_n^*,y^*)}$$
(10)

Example

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

$$\frac{\partial U}{\partial x} dx + \frac{\partial U}{\partial y} dy = 0$$
$$\frac{dy}{dx} = -\frac{\frac{\partial U}{\partial x}}{\frac{\partial U}{\partial y}}$$

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

Unconstrained optimization

• Let $U \subset \mathbb{R}^n$ and $F: U \to \mathbb{R}$. A point $x^* \in U$ is a max of F on U if $F(x^*) \ge F(x) \ \forall x \in U$

Theorem (First-order necessary conditions)

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

$$\frac{\partial F}{\partial x_i}(x^*) = 0 \quad \forall i = 1, \dots, n \tag{11}$$

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

$$F_x = 2x - 6y + 10 = 0$$

$$F_y = -6x + 4y + 2 = 0,$$

which yields: $(x^*, y^*) = (13/7, 6/7)$

Unconstrained optimization

Theorem (Second-order sufficient conditions)

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

$$\mathcal{H} \equiv D^2 F(x^*) = \begin{pmatrix} \frac{\partial^2 F}{\partial x_1^2}(x^*) & \cdots & \frac{\partial^2 F}{\partial x_n x_1}(x^*) \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 F}{\partial x_1 x_n}(x^*) & \cdots & \frac{\partial^2 F}{\partial x_n^2}(x^*) \end{pmatrix}$$
(12)

1. If \mathcal{H} is negative definite, i.e. if $x'\mathcal{H}x < 0 \ \forall x \in U$, then x^* is a strict local max of F

2. If \mathcal{H} is positive definite, i.e. if $x'\mathcal{H}x > 0 \ \forall x \in U$, 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 $f : \mathbb{R}^n \to \mathbb{R}$ be C^2 and let x^* be a critical point of f (i.e. $f_{x_i} = 0 \quad \forall i = 1, ..., n$) and that the n leading principal minors of \mathcal{H} alternate in sign:

$$|f_{x_1x_1}| < 0, \quad \left| \begin{array}{c} f_{x_1x_1} & f_{x_2x_1} \\ f_{x_1x_2} & f_{x_2x_2} \end{array} \right| > 0, \quad \left| \begin{array}{c} f_{x_1x_1} & f_{x_2x_1} & f_{x_3x_1} \\ f_{x_1x_2} & f_{x_2x_2} & f_{x_3x_2} \\ f_{x_1x_3} & f_{x_2x_3} & f_{x_3x_3} \end{array} \right| < 0, \dots$$
(13)

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

$$f(\lambda x + (1 - \lambda)y) \ge \lambda f(x) + (1 - \lambda)f(y)$$
(14)

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

$$f(\lambda x + (1 - \lambda)y) \le \lambda f(x) + (1 - \lambda)f(y)$$
(15)

Theorem

Let $f : \mathbb{R}^n \to \mathbb{R}$ be 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:

$$\max_{x_1,\ldots,x_n} f(x_1,\ldots,x_n) \tag{16}$$

subject to:

$$g_1(x_1,\ldots,x_n) \le b_1,\ldots,g_k(x_1,\ldots,x_n) \le b_k \tag{17}$$

$$h_1(x_1, \dots, x_n) = c_1, \dots, h_k(x_1, \dots, x_n) = c_m$$
 (18)

- *f* is the objective function
- {g_i} are inequality constraints
- $\{h_j\}$ are equality constraints
- $x_i \ge 0$ non-negativity constraints

Equality constraints

Theorem (Lagrangian)

Let f, h_1, \ldots, h_m be C^1 functions. Suppose that \mathbf{x}^* is a local max of f in the constraint set $\{\mathbf{x}: h_1(\mathbf{x}) = c_1, \ldots, h_m(\mathbf{x}) = c_m\}$ and assume that rank of the Jacobian matrix of first-derivatives of the constraints with respect to \mathbf{x} ,

$$\mathcal{J} \equiv Dh(\mathbf{x}^*) = \begin{pmatrix} \partial h_1 / \partial x_1 & \cdots & \partial h_1 / \partial x_n \\ \vdots & \ddots & \vdots \\ \partial h_m / \partial x_1 & \cdots & \partial h_m / \partial x_n \end{pmatrix}$$
(19)

is *m*. Then there exist $\lambda_1^*, \ldots, \lambda_m^*$ such that $(x_1^*, \ldots, x_n^*, \lambda_1^*, \ldots, \lambda_m^*)$ is a critical point of the Lagrangian

$$\mathcal{L}(\mathbf{x},\lambda) \equiv f(\mathbf{x}) - \sum_{j=1}^{m} \lambda_j \left[h_j(\mathbf{x}) - c_j \right]$$
(20)

That is,

$$\frac{\partial \mathcal{L}}{\partial x_i}(\mathbf{x}^*, \lambda^*) = 0 \quad \forall i = 1, \dots, n$$
$$\frac{\partial \mathcal{L}}{\partial \lambda_j}(\mathbf{x}^*, \lambda^*) = 0 \quad \forall j = 1, \dots, m$$

Example

$$\max_{x_1, x_2} U = \alpha \ln x_1 + (1 - \alpha) \ln x_2$$

s.t.:
$$p_1 x_1 + p_2 x_2 = m$$

$$\begin{aligned} & [x_1]: \ \alpha/x_1 - \lambda p_1 = 0 \\ & [x_2]: \ (1 - \alpha)/x_2 - \lambda p_2 = 0 \\ & [\lambda]: \ p_1 x_1 + p_2 x_2 - m = 0 \end{aligned}$$

• Using the first two FOC: $p_2x_2 = \frac{1-\alpha}{\alpha}p_1x_1$. Plugging this back into the the third FOC yields:

$$(x_1^*, x_2^*, \lambda^*) = \left(\frac{\alpha m}{p_1}, \frac{(1-\alpha)m}{p_2}, \frac{1}{m}\right)$$

• Notice that $\mathcal{J}(\mathbf{x}^*) = (p_1 \ p_2)$ has rank 1

Inequality constraints

• We want to solve the following problem:

$$\max_{x_1,\ldots,x_n} f(x_1,\ldots,x_n) \tag{21}$$

subject to:

$$g_1(x_1, \dots, x_n) \le b_1, \dots, g_K(x_1, \dots, x_n) \le b_K$$
 (22)
 $x_1 \ge 0, \dots, x_n \ge 0$ (23)

The Lagrangian is given by:

$$\mathcal{L} = f(\mathbf{x}) - \sum_{k=1}^{K} \lambda_k [g_k(\mathbf{x}) - b_k] + \sum_{i=1}^{n} \mu_i x_i$$
(24)

• The FOC are:

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial x_i}(x^*) &\leq 0, \quad \forall i = 1, \dots, n\\ x_i \frac{\partial \mathcal{L}}{\partial x_i}(x^*) &= 0, \quad \forall i = 1, \dots, n\\ \frac{\partial \mathcal{L}}{\partial \lambda_k}(x^*) &\geq 0, \quad \forall k = 1, \dots, K\\ \lambda_k \frac{\partial \mathcal{L}}{\partial \lambda_k}(x^*) &= 0, \quad \forall k = 1, \dots, K\\ \lambda_k, \mu_i &\geq 0, \quad \forall k = 1, \dots, K, \quad i = 1, \dots, n \end{aligned}$$

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{aligned} \max_{c,n,l} \log c + al, & a > 0\\ \text{s.t.:} \\ c \leq wn\\ l+n \leq 1\\ c,n,l \geq 0 \end{aligned}
```

- 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 $f, h_1, \ldots, h_k : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$ be C^1 functions. Let $\mathbf{x}^*(a)$ denote the solution of the problem of maximizing $f(\mathbf{x}; a)$ on the constraint set $\{h_1(\mathbf{x}; a) = 0, \ldots, h_k(\mathbf{x}; a) = 0\}$ for any fixed parameter a. Then,

$$\frac{d}{da}f(\mathbf{x}^{*}(a);a) = \frac{\partial L}{\partial a}(\mathbf{x}^{*}(a),\mu(a);a)$$
(25)

- 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 (x_1^*, x_2^*) into the objective function, you obtain the 'value function':

$$V(p_1, p_2, m) = \alpha \ln\left(\frac{\alpha m}{p_1}\right) + (1 - \alpha) \ln\left(\frac{(1 - \alpha)m}{p_2}\right)$$
$$\rightarrow \frac{\partial V}{\partial m} = \frac{1}{m}$$

• 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 $n \times 1$, $\overline{x} \in \mathbb{R}^n$, 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,

$$P_{ij} = \mathsf{Prob}(x_{t+1} = \overline{x}_j | x_t = \overline{x}_i)$$

(iii) A $n \times 1$ vector π_0 recording the probabilities of being in each state *i* at time 0,

 $\pi_0 = \mathsf{Prob}(x_0 = \overline{x}_i)$

We need P and π_0 to satisfy the following conditions:

- (i) For all i = 1, ..., n, $\sum_{j=1}^{n} P_{ij} = 1$
- (ii) $\sum_{i=1}^{n} \pi_{0i} = 1$

 \boldsymbol{P} is called a stochastic matrix

Example

- Suppose an individual can be in either of two states: employed, $e_{\rm r}$ 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:

$$\begin{split} \overline{x} &= \{u, e\}, & \text{(state space)} \\ \pi_0 &= [0.5 \ 0.5]', & \text{(initial state probability)} \\ P &= \begin{bmatrix} p_{uu} & p_{ue} \\ p_{eu} & p_{ee} \end{bmatrix} = \begin{bmatrix} 0.5 & 0.5 \\ 0.043 & 0.957 \end{bmatrix} & \text{(stochastic matrix)} \end{split}$$

An example — a realization of $\{x_t\}$



$$\overline{x} = \{u, e\},\$$

$$\pi_0 = \begin{bmatrix} 0.5 & 0.5 \end{bmatrix}',\$$

$$P = \begin{bmatrix} p_{uu} & p_{ue} \\ p_{eu} & p_{ee} \end{bmatrix} = \begin{bmatrix} 0.5 & 0.5 \\ 0.043 & 0.957 \end{bmatrix}.$$

Asymptotic distribution of \boldsymbol{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) β is the discount factor

Dynamic programming

• We are interested in solving problems of the form:

$$\max_{\{x_{t+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^{t} F(x_{t}, x_{t+1})$$
s.t.:
$$x_{t+1} \in \Gamma(x_{t}), \quad t = 0, 1, 2, \dots$$

$$x_{0} \in X \text{ given.}$$
(26)
(27)
(27)
(28)

- 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 $x_{t+1} \in \Gamma(x_0)$
- After choosing x_1 , the decision-maker obtains a reward $F(x_0, x_1)$
- In t = 2, the available options for the decision-maker are $x_{t+2} \in \Gamma(x_1)$
- 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:

$$\max_{\{c_t, i_t\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t \ln(c_t)$$
(29)

$$y_t = k_t^{\alpha} \tag{30}$$

$$c_t + i_t = y_t \tag{31}$$

$$k_{t+1} = (1 - \delta)k_t + i_t \tag{32}$$

$$k_0 > 0$$
 given (33)

• Simplifying:

$$\max_{\{k_{t+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t \ln \left[k_t^{\alpha} + (1-\delta)k_t - k_{t+1} \right]$$
(34)

- 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 $v: X \to \mathbb{R}$ as follows:

$$v(x_0) = \max_{x \in \Gamma(x_0)} \sum_{t=0}^{\infty} \beta^t F(x_t, x_{t+1})$$
(35)

- We call $v(x_0)$ a value function. It tells us the value of the maximized objective function for a given value of the state variable
- Similarly, $v(x_1)$ would give us the maximum utility value that the decision-maker can achieve from period 1 afterwards given his/her state x_1 Then,

$$v(x_0) = \max_{x_1 \in \Gamma(x_0)} \left\{ \underbrace{F(x_0, x_1) + \beta v(x_1)}_{=\sum_{t=0}^{\infty} \beta^t F(x_t, x_{t+1})} \right\}$$
(36)

- This is called the recursive representation of a DP
- A solution for the DP problem specified above is a policy rule, $g: X \to \Gamma(x)$ 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 $x \in X$:

$$v(x) = \max_{x' \in \Gamma(x)} \left\{ F(x, x') + \beta v(x') \right\}$$
(*)

where \boldsymbol{x}' denotes the next-period value of \boldsymbol{x}

Note that the solution to (★) 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 ...
- and we can iterate on the following recursion:

$$v_{j+1} = \max_{x \in \Gamma(x)} \left\{ F(x, x') + \beta v_j(x') \right\}$$
(37)

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

• 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 $k, k' \in [0, f(k) + (1 \delta)k) \Rightarrow \Gamma(k) = [0, f(k) + (1 \delta)k)$
- The Bellman equation of this problem is then:

$$v(k) = \max_{k' \in [0, k^{\alpha} + (1-\delta)k)} \left\{ \ln[k^{\alpha} + (1-\delta)k - k'] + \beta v(k') \right\}$$
(38)

Intuition on how to solve this in the computer

- Create a grid of N points for the state variable k, $\mathcal{K}=[k_1,k_2,\ldots,k_N]$ 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 $v(k) = [v_1, v_2, \dots, v_N]$
- 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

$$\ln(k^{\alpha} + (1 - \delta)k - k') + \beta v_0(k')$$
(39)

we're going to create a big $N \times N$ matrix, where future capital k' 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

$$\begin{split} &\ln(k^{\alpha} + (1-\delta)k - k') \text{ looks like this:} \\ & \left[\begin{array}{cccc} \ln[k_{1}^{\alpha} + (1-\delta)k_{1} - k_{1}] & \ln[k_{2}^{\alpha} + (1-\delta)k_{2} - k_{1}] & \cdots & \ln[k_{N}^{\alpha} + (1-\delta)k_{N} - k_{1}] \\ \ln[k_{1}^{\alpha} + (1-\delta)k_{1} - k_{2}] & \ln[k_{2}^{\alpha} + (1-\delta)k_{2} - k_{2}] & \cdots & \ln[k_{N}^{\alpha} + (1-\delta)k_{N} - k_{2}] \\ & \vdots & \vdots & \ddots & \vdots \\ \ln[k_{1}^{\alpha} + (1-\delta)k_{1} - k_{N}] & \ln[k_{2}^{\alpha} + (1-\delta)k_{2} - k_{N}] & \cdots & \ln[k_{N}^{\alpha} + (1-\delta)k_{N} - k_{N}] \end{array} \right] \end{split}$$

For the second component, notice two things: (i) $v_0(k')$ is of size $N \times 1$ and (ii) it only depends on future capital k'

$$\beta v_0(k) \cdot [1 \cdots 1]_{1 \times N} = \beta \begin{bmatrix} v_0(k_1) & v_0(k_1) & \cdots & v_0(k_1) \\ v_0(k_2) & v_0(k_2) & \cdots & v_0(k_2) \\ \vdots & \vdots & \ddots & \vdots \\ v_0(k_N) & v_0(k_N) & \cdots & v_0(k_N) \end{bmatrix}$$

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, ... N you will have a new maximized value. So you end up with a new vector $v_1(k) = [v_1(k_1) \cdots v_1(k_N)]$
- Then, you compare $v_0(k)$ and $v_1(k)$, if they're close enough $||v_1(k) v_0(k)|| < \varepsilon$, 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 $k' = k_5$).
- When you have achieved convergence, you also find the policy rule, that tells you
 what should be the optimal future capital k' 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 AR(1) process:

$$z_t = \rho z_{t-1} + \varepsilon_t, \quad \varepsilon \stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma_{\varepsilon}^2)$$
(40)

- Production now looks like this: $y_t = \exp(z_t)k_t^{\alpha}$
- 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 ε_t drawn from a Normal distribution with mean 0 and variance σ_{ε}^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 , $\mathbb{E}_t[V(k', z')]$

Stochastic dynamic programming

• The Bellman equation of our new problem is:

$$v(k,z) = \max_{k' \in [0,\exp(z)k^{\alpha} + (1-\delta)k)} \left\{ \ln[\exp(z)k^{\alpha} + (1-\delta)k - k'] + \beta \sum_{z'} P(z'|z)V(k',z') \right\}$$
(41)

- P is the stochastic matrix associated to the Markov-chain approximation of the $\{z_t\}$
- 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.