VFI, PFI and DP in Stochastic Environments

Macroeconomics 3: TA class #3

Andrea Pasqualini

Bocconi University

15 February 2021

Plan for Today

Objective: **Solve numerically for $V(K, A)$ and $K'(K, A)$**

Operating example: Neoclassical Growth Model (stochastic version)

$$V(K, A) \equiv \max_{C, K'} \frac{C^{1-\gamma}}{1-\gamma} + \beta \mathbf{E}\left(V(K', A') | A\right)$$
$$\text{s.t.} \begin{cases} C + K' \leq A K^{\alpha} + (1-\delta) K \\ C, K' > 0 \\ \log(A') = (1-\rho)\log(\mu) + \rho \log(A) + \epsilon \\ \epsilon \overset{iid}{\sim} \mathcal{N}(0, \sigma^2) \end{cases}$$

- Same methods as last time
- Same objects of interest
- Adding discretization methods for time series processes

Shocks are operationally useful

- Simulations
- Impulse-Response Functions
- Forecast Error-Variance Decomposition

The Discretization Problem

Objective: **approximate a continuous stochastic process with a discrete one**

Same problem faced in our last class: the computer has no concept of set density

$$\log(A') = (1-\rho)\log(\mu) + \rho\log(A) + \epsilon \qquad\qquad \epsilon \overset{iid}{\sim} \mathcal{N}(0,\sigma^2)$$

$$\underbrace{\begin{bmatrix} \log(A_1) \\ \log(A_2) \\ \vdots \\ \log(A_m) \end{bmatrix}}_{\text{grid for } \log(A)} = \underbrace{\begin{bmatrix} \Pi_{1,1} & \Pi_{1,2} & \cdots & \Pi_{1,m} \\ \Pi_{2,1} & \Pi_{2,2} & \cdots & \Pi_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ \Pi_{m,1} & \Pi_{m,2} & \cdots & \Pi_{m,m} \end{bmatrix}}_{\text{transition probabilities, } \Pi} \cdot \underbrace{\begin{bmatrix} \log(A_1) \\ \log(A_2) \\ \vdots \\ \log(A_m) \end{bmatrix}}_{\text{grid for } \log(A)}$$

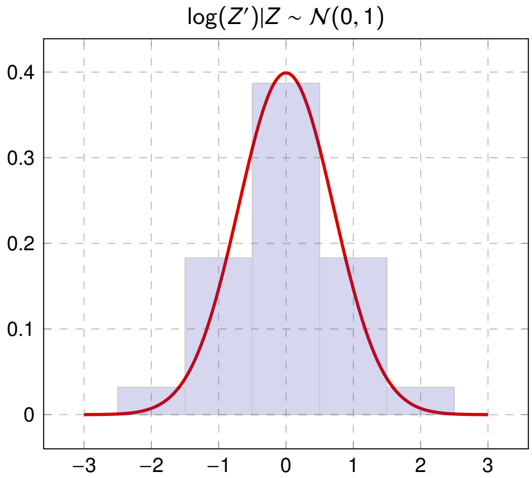
- Need to set up a grid for A (same as before)
- Need to figure out the transition probabilities (new!)
  - Important to compute the conditional expected continuation value $\mathbf{E}\left(V(K', A')|A\right)$

The Discretization Problem (cont'd)

Objective: **approximate a continuous stochastic process with a discrete one**

Must match

- Unconditional exp. value
- Conditional exp. value
- Unconditional variance
- Conditional variance
- (Optional) skewness
- (Optional) kurtosis
- (Optional) higher-order moments



Note: the parameters in this figure are chosen exclusively for illustration purposes

Overview of Methods

Tauchen

- Constructs a histogram for a conditional distribution function
- Can control the grid directly
- Easy to code, easy intuition
- Approximation errors with high-persistence processes

Tauchen-Hussey

- Constructs a histogram for a conditional distribution function
- Imposes a fancy grid, no control over it (except for no. of points)
- Approximation errors with high-persistence processes

Rouwenhorst

- Recursively approximates a conditional distribution function
- No control on grid (except for no. of points)
- Robust to high-persistence processes
- "It just works," non-obvious intuition

The Tauchen Algorithm

1. Forget about the unconditional average of the process (will recover it later)
2. Create a grid for the support S of the probability distribution function (this is a vector)
3. Compute all possible transitions $S' - \rho S$ (this is a matrix)
4. Evaluate the relevant CDF (e.g., Gaussian) at the possible transitions
5. Make the resulting matrix such that each row sums to one
6. Shift the grid by the unconditional average, if needed

The Tauchen-Hussey Algorithm

1. Forget about the unconditional average of the process (will recover it later)
2. Obtain the grid for S by computing the zeros of a Gauss-Hermite polynomial of degree m
3. Rescale the grid points by $\sqrt{2\sigma^2}$ (this is a vector)
4. Compute the relevant *conditional* PDF at the possible transitions (this is a matrix)
5. Rescale the computed conditional PDF to account for discrete points
6. Normalize the matrix so that rows sum to one
7. Shift the grid by the unconditional average, if needed

The Rouwenhorst Algorithm

1. Set $p, q \in (0, 1)$
2. For $m = 2$ grid points, construct Π_2 as

$$\Pi_2 = \begin{bmatrix} p & 1-p \\ 1-q & q \end{bmatrix}$$

3. For $m > 2$ grid points
   3.1 Construct Π_m as

$$\Pi_m = p \begin{bmatrix} \Pi_{m-1} & 0 \\ 0 & 0 \end{bmatrix} + (1-p) \begin{bmatrix} 0 & \Pi_{m-1} \\ 0 & 0 \end{bmatrix} + (1-q) \begin{bmatrix} 0 & 0 \\ \Pi_{m-1} & 0 \end{bmatrix} + q \begin{bmatrix} 0 & 0 \\ 0 & \Pi_{m-1} \end{bmatrix}$$

   3.2 Divide by 2 all but the top and bottom rows of Π_m (those rows in the middle sum to 2)
4. Create a grid of linearly spaced points for the support of the PDF
   4.1 Compute $f = \sqrt{N-1} \cdot \sigma / \sqrt{1-\rho^2}$ (it relates to the uncond. variance of the AR(1))
   4.2 Create $A = \{a_1, \dots, a_m\}$ with $a_1 = -f$ and $a_m = f$
   4.3 Shift A by the unconditional average, if necessary

- Setting $p = q$ ensures homoskedasticity in the structure of shocks/innovations
- Setting $p = q = (1+\rho)/2$ matches the variance of the original AR(1) process

Ergodic Distribution

Objective: **Compute the ergodic PDF of a Markov Chain**

The ergodic distribution π of a Markov Chain with transition matrix Π is such that

$$\begin{cases} \pi = \Pi' \pi \\ \pi \iota = 1 \end{cases}$$

where ι is a vector of 1's

The system of equations above says that

- The vector π is one eigenvector of the matrix Π...
- ... in particular, the one whose elements sum to one

There are countless ways to compute the ergodic distribution, but this one works quite well

Calibration

Symbol	Meaning	Value
α	Capital intensity in PF	0.30
β	Discount parameter	0.95
γ	CRRA parameter	1.50
δ	Capital depreciation	0.10
μ	Uncond. avg. of productivity	1.00
ρ	Persistence of productivity	0.70
σ	St.dev. of productivity shocks	0.10

The same disclaimer as in the previous class applies
$\rightarrow$ The calibration presented here is not credible in any meaningful empirical setting

Simulation

Consider the necessary and sufficient conditions for the equilibrium in any model with rational expectations

$$\mathbf{E}_t\left(f\left(X_{t-1}, X_t, X_{t+1}\right)\right) = 0$$

The solution to such model is a "policy function" $g(\cdot)$ such that

$$X_{t+1} = g(X_{t-1}, X_t)$$

What we call here "policy function" $g(\cdot)$ is a vector function containing

- The policy functions (strictly speaking) from the Bellman problem
- The laws of motion (e.g., the one for capital)
- Exogenous stochastic processes (e.g., the one for productivity)

A simulation takes some initial conditions for X_{t-1} and X_t and applies the function $g(\cdot)$ repeatedly for a given series of shocks

Simulation (cont'd)

Steps to simulate from the Stochastic Neoclassical Growth Model

1. Set a number of periods T to simulate
2. Set K_0, that is the initial condition
3. For each $t \in \{0, \ldots, T\}$
   3.1 Draw a state A_t from the relevant CDF
   3.2 Compute current consumption and future capital holdings using the policy functions $C_t = C(K_t, A_t)$ and $K_{t+1} = K'(K_t, A_t)$
   3.3 Compute all other endogenous variables using other equations of the model (e.g., production, investment)

Impulse-Response Functions

Objective: **Marginal effect of an exogenous shock on an endogenous variable**

IRFs are the marginal effects of shocks on endogenous variables predicted by the model

Formally, the response at horizon h of variable X_{t+h} to a shock (impulse) to S_t is

$$IRF_{X,S}(h) \equiv \frac{\partial X_{t+h}}{\partial S_t}$$

IRFs are simple simulations

- The initial condition is typically the steady state of the model
- At time t, a sudden unexpected shock realizes
- At time $t + h$, for all $h > 0$, all shocks are shut down

Exercises

1. Use the code for VFI/PFI I have shown in class #2
   1.1 Write code that, given an initial condition for the state variable, simulates the model
   1.2 In what sense such simulation is uninteresting?

2. Use the code for VFI I have shown in this class
   2.1 How would the discretization of the stochastic process work if $\rho = 0$ (i.e., the process of A itself is a sequence of i.i.d. random variables)?
   2.2 Code up the related discretization method and solve for the policy function

3. Consider the code I have shown in this class for simulating the model
   3.1 My code forces the shock to be on the grid for A: how would you modify the numerical policy functions (and those only!) to accomodate for any $A \in \mathbb{R}$?
   3.2 Code up your answer to the previous question
   3.3 Simulate the model for some periods (e.g., $T = 250$)
   3.4 Compute the impulse-response functions of consumption, investment and production to a one-standard deviation shock to productivity
   3.5 Provide the economic intuition behind the the IRFs you have obtained