Lecture XII

Approximating the Invariant Distribution

Gianluca Violante

New York University

Quantitative Macroeconomics

SS Equilibrium in the Aiyagari model

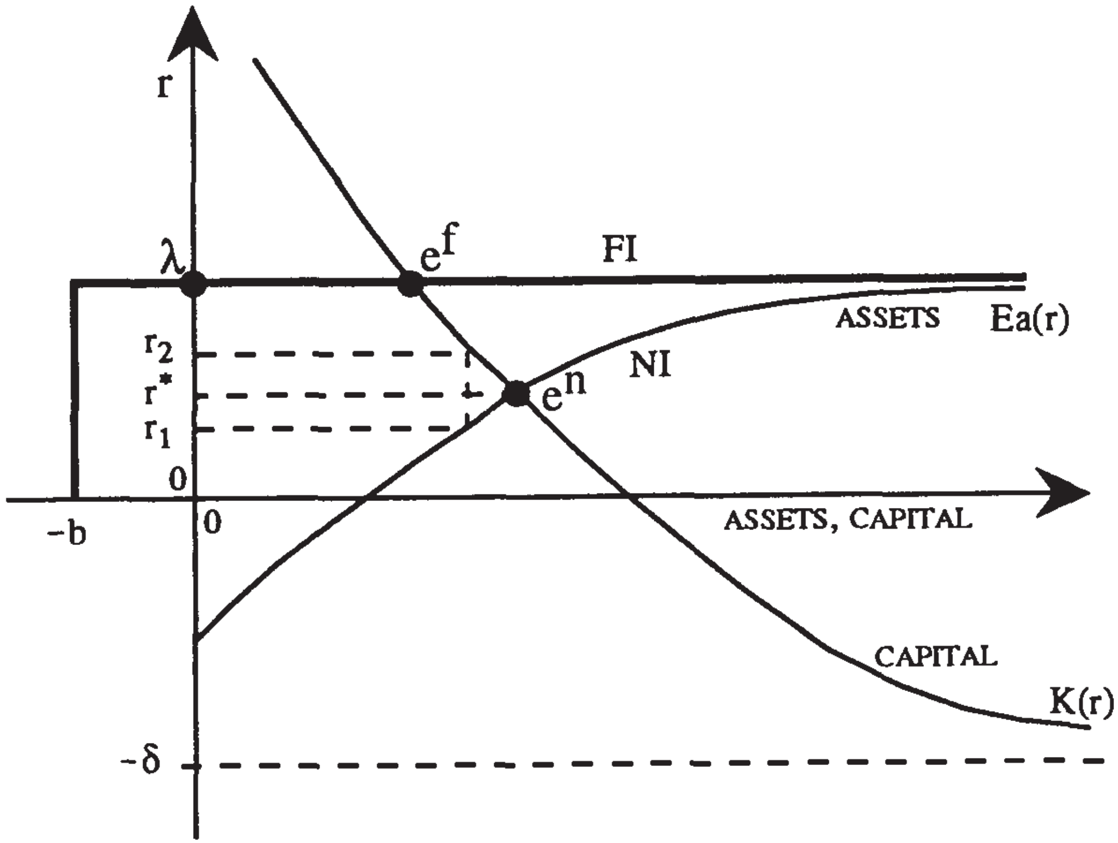

FIGURE IIb
Steady-State Determination

SS Equilibrium of the Aiyagari model

Fixed point algorithm over the interest rate (or K).

1. Fix an initial guess for the interest rate $r^0 \in \left(-\delta, \frac{1}{\beta} - 1\right)$. r^0 is our first candidate for the equilibrium (the superscript denotes the iteration number).

2. Given r^0, solve the dynamic programming problem of the agent to obtain $a' = g(a, y; r^0)$. We described several solution methods.

3. Given the decision rule for assets next period $g(a, y; r^0)$ and the Markov transition over productivity shocks $\Gamma\left(y', y\right)$, we can construct the transition function $Q\left(r^0\right)$ and, by successive iterations, we obtain the fixed point distribution $\Lambda\left(r^0\right)$, conditional on the candidate interest rate r^0.

SS Equilibrium of the Aiyagari model

4. Compute the aggregate demand of capital $K\left(r^0\right)$ from the optimal choice of the firm who takes r^0 as given, i.e.

$$K(r^0) = F_k^{-1}(r^0 + \delta)$$

5. Compute the integral:

$$A\left(r^0\right) = \int_{A \times Y} g(a, y; r^0) d\Lambda(a, y; r^0)$$

which gives the aggregate supply of assets. This can be easily done by using the invariant distribution obtained in step 3.

SS Equilibrium of the Aiyagari model

6. Compare $K\left(r^0\right)$ with $A\left(r^0\right)$ to verify whether the asset market clearing condition holds.

7. Use a nonlinear equation solver in your outer loop to update from r^0 to r^1 and go back to step 1.

8. Keep iterating until you reach convergence of the interest rate, i.e., until, at iteration n:

$$\left|r^{n+1} - r^n\right| < \varepsilon$$

Calculating the invariant distribution

- Continuum of households makes the wealth distribution a continuous function and therefore an infinite-dimensional object in the state space. Need approximation.

- Small errors in computing the invariant distribution, particularly in models with aggregate shocks, can accumulate and lead to significant errors in the equilibrium values of aggregate variables.

- Five approaches:

  1. piecewise linear approximation of invariant distribution function
  2. discretization of the invariant density function
  3. eigenvector method
  4. simulation of an artificial panel
  5. approximating the cdf by an exponential function [not today]

Preliminaries

- Let $\mathcal{Y}$ be the grid of y and Γ be its J-state Markov chain with $\Gamma_{ij} = \Pr\{y' = y_j | y = y_i\}$

- We think of separate asset distributions conditional on each value of $y \in \mathcal{Y}$, i.e.

$$\Lambda(a^*, y^*) = \Pr\{a \leq a^*, y = y^*\}$$

- The invariant distribution function satisfies the law of motion:

$$\Lambda(a', y') = \sum_{y \in \mathcal{Y}} \Gamma(y', y) \Lambda\left(g^{-1}(a', y), y\right) \quad \text{for all } (a', y')$$

where $g^{-1}(a', y)$ is the inverse of the saving decision rule with respect to its first argument a. Clearly, we are using the monotonicity of g.

Piecewise linear approximation of invariant distribution

- This method involves approximating Λ by a weighted sum of piecewise-linear functions and iterating on the law of motion above

- The approximation of Λ must be shape-preserving since it would not make sense if Λ decreased over some range of a. This makes piecewise linear basis functions a better choice than say Chebyshev polynomials or cubic splines.

- Other approximations that preserve shape, such as shape-preserving splines could potentially also be appropriate

- First we need a grid $\mathcal{A}$ of interpolation nodes in the interval $[\underline{a}, \bar{a}]$. This grid needs to be a lot finer than the grid used for computing the decision rule. Does not have to be an evenly spaced grid. Let N be the size of this grid.

Piecewise linear approximation of invariant distribution

- Choose an initial distribution Λ^0 over the grid $\mathcal{A} \times \mathcal{Y}$. One choice is, for example, for every pair $(a_k, y_j) \in \mathcal{A} \times \mathcal{Y}$

$$\Lambda^0 (a_k, y_j) = \frac{a_k - \underline{a}}{\bar{a} - \underline{a}} \Gamma_j^*$$

as if the two variables were independent and the distribution over assets a were uniform.

- Update the distribution on grid points as follows. For every pair $(a_k, y_j) \in \mathcal{A} \times \mathcal{Y}$

$$\Lambda^1 (a_k, y_j) = \sum_{y_i \in \mathcal{Y}} \Gamma_{ij} \Lambda^0 \left(g^{-1} (a_k, y_i), y_i \right)$$

Note: (i) need to compute g^{-1} and (ii) g^{-1} not on $\mathcal{A}$

Piecewise linear approximation of invariant distribution

- How do we solve for $a = g^{-1}(a_k, y_i)$? Recall: you also have the consumption decision rule on the original grid used for policy functions (obtained residually from the budget constraint). Then, since $a_k = g(a, y_i)$

  $$c(a, y_i) + a_k = Ra + y_i$$

  and we can calculate a through a nonlinear solver given a piece-wise linear representation of $c(a, y_i)$.

- Given $a = g^{-1}(a_k, y_i)$, we need a linear interpolation of Λ^0 since Λ^0 is only defined on the grid.

Piecewise linear approximation of invariant distribution

- For an $a \in [a_n, a_{n+1}]$, we define:

$$\Lambda^0(a, y_i) = \Lambda^0(a_n, y_i) + \frac{\Lambda^0(a_{n+1}, y_i) - \Lambda^0(a_n, y_i)}{a_{n+1} - a_n}(a - a_n)$$

- Compare $\Lambda^1(a_k, y_j)$ to $\Lambda^0(a_k, y_j)$, for example with the $supnorm$

$$\max_{j,k} |\Lambda^1(a_k, y_j) - \Lambda^0(a_k, y_j)| < \varepsilon$$

and stop when your tolerance level ε is reached

- How do we compute the aggregate supply of capital

$$A = \sum_{y_i \in \mathcal{Y}} \int_A a d\Lambda(a, y_i)$$

Piecewise linear approximation of invariant distribution

- We assume the distribution of assets is uniform within the element $[a_n, a_{n+1}]$ and therefore

$$\begin{aligned}\int_{a_n}^{a_{n+1}} a d\Lambda(a, y_i) &= \int_{a_n}^{a_{n+1}} a \left[\frac{\Lambda\left(a_{n+1}, y_i\right) - \Lambda\left(a_n, y_i\right)}{a_{n+1} - a_n}\right] da \\ &= \frac{\Lambda\left(a_{n+1}, y_i\right) - \Lambda\left(a_n, y_i\right)}{a_{n+1} - a_n} \left[\frac{a^2}{2}\right]_{a_n}^{a_{n+1}} \\ &= \frac{\Lambda\left(a_{n+1}, y_i\right) - \Lambda\left(a_n, y_i\right)}{2} \left(a_{n+1} + a_n\right)\end{aligned}$$

- Then:

$$A = \sum_{y_i \in \mathcal{Y}} \sum_{n=1}^{N-1} \frac{\Lambda\left(a_{n+1}, y_i\right) - \Lambda\left(a_n, y_i\right)}{2} \left(a_{n+1} + a_n\right) + \Lambda\left(\underline{a}, y_i\right) \underline{a}$$

where we account for the fact that there may be a mass point at the borrowing constraint.

Discretization of the invariant density function

- A simpler approach involves finding an approximation to the invariant density function $\lambda(a, y)$
- We will approximate the density by a probability distribution function defined over a discretized version of the state space. Once again the grid should be finer than the one used to compute the optimal savings rule.
- Choose initial density functions $\lambda^0(a_k, y_i)$. For example

$$\lambda^0(a_k, y_i) = \begin{cases} 0 & \text{if } k > 1 \text{ and } i > 1 \\ 1 & \text{if } k = 1 \text{ and } i = 1 \end{cases}$$

i.e., all the mass is at the borrowing limit and at the lowest realization of the income shock

Discretization of the invariant density function

- For every (a_k, y_j) on the grid:

$$\lambda^1(a_k, y_j) = \sum_{y_i \in \mathcal{Y}} \Gamma_{ij} \sum_{m \in \mathcal{M}_i} \frac{a_{k+1} - g(a_m, y_i)}{a_{k+1} - a_k} \lambda^0(a_m, y_i)$$

$$\lambda^1(a_{k+1}, y_j) = \sum_{y_i \in \mathcal{Y}} \Gamma_{ij} \sum_{m \in \mathcal{M}_i} \frac{g(a_m, y_i) - a_k}{a_{k+1} - a_k} \lambda^0(a_m, y_i)$$

where

$$\mathcal{M}_i = \{m = 1, ..., N | a_k \leq g(a_m, y_i) \leq a_{k+1}\}$$

Discretization of the invariant density function

- We can think of this way of handling the discrete approximation to the density function as forcing the agents in the economy to play a lottery.

- Lottery: with the probability of going to a_k given that your optimal policy is to save $a' = g\left(a_m, y_i\right) \in \left[a_{k,} a_{k+1}\right]$ is given by $\frac{a_{k+1} - g(a_m, y_i)}{a_{k+1} - a_k}$ and with probability $1 - \frac{a_{k+1} - g(a_m, y_i)}{a_{k+1} - a_k}$ you go to a_{k+1}.

- Then, once you have found λ, the aggregate supply of capital is computed as:

$$A = \sum_{y_i \in \mathcal{Y}} \sum_{a_k \in \mathcal{A}} a_k \lambda\left(a_k, y_j\right)$$

Eigenvector method

- Let A be a square matrix. If there is a vector $v \neq 0$ in R^n such that:

  $$vA = ev$$

  for some scalar e, then e is called the eigenvalue of A with corresponding (left) eigenvector v.

- Recall that the definition of invariant pdf for a Markov transition matrix Q is λ^* that satisfies:

  $$\lambda^* Q = \lambda^*$$

  thus an invariant distribution is the eigenvector of the matrix Q associated to eigenvalue 1.

- How do we guarantee this eigenvector is unique and how do we find it in that case?

Eigenvector method

- Perron-Frobenius Theorem: Let Q be a transition matrix of an homogeneous ergodic Markov chain. Then Q has a unique dominant eigenvalue $e = 1$ such that:
  - its associated eigenvector has all positive entries
  - all other eigenvalues are smaller than e in absolute value
  - Q has no other eigenvector with all non-negative entries
- This eigenvector (renormalized so that it sums to one) is the unique invariant distribution

Eigenvector methods

- How do we construct Q?

- Let $Q\left(a',y';a,y\right)$ be the $JN \times JN$ transition matrix from state $\left(a,y\right)$ into $\left(a',y'\right)$. Since the evolution of y' is independent of a:

$$Q\left(a',y';a,y\right) = Q_a\left(a';a,y\right) \otimes \Gamma\left(y',y\right)$$

where, if $a' = g\left(a,y\right)$ is such that $a_{k+1} \leq a' < a_k$, then we define:

$$\begin{aligned} Q_a\left(a_k;a,y\right) &= \frac{a_{k+1} - g\left(a,y\right)}{a_{k+1} - a_k} \\ Q_a\left(a_{k+1};a,y\right) &= \frac{g\left(a,y\right) - a_k}{a_{k+1} - a_k} \end{aligned}$$

... same lottery we used before for the density

Eigenvector methods

- In practice: the Q function is very large and very sparse and there are many eigenvalues very close to one
- The corresponding eigenvectors generally have negative components that are inappropriate for the density function.
- Idea: perturb the zero entries of Q by adding a perturbation constant η and renormalizing the rows of Q.
- Then find the unique non-negative eigenvector associated with the unit root of the matrix. The stationary density can be obtained by normalizing this eigenvector to add to one.
- How do we select η? Authors suggest

$$\eta \leq \frac{\min Q}{2N} \quad \text{or even smaller}$$

Monte-Carlo simulation

- One must generate a large sample of households and track them over time.
- Monte Carlo simulation is memory and time consuming: not recommended for low dimensional problems.
- Valuable method however when the dimension of the problem is large since it is not subject to the curse of dimensionality that plagues all other methods.
- Choose a sample size I. Typically, $I \geq 10,000$
- At $t = 0$, initialize states by (i) draws from Γ^* and (ii) some value for $a_i^0 = a^*$, for all agents, e.g., steady-state capital of the representative agent (divided by I)

Monte-Carlo simulation

- Update asset holdings each individual i in the sample by using the decision rule:

$$a_i^1 = g\left(a_i^0, y_i^1\right)$$

  where y_i^1 is drawn from $\Gamma\left(y_i^1, y_i^0\right)$.

- Use a random number generator for the uniform distribution in $[0, 1]$. Suppose the draw from the uniform is u. Then $y_i^1 = y_{j^*}$ where j^* is the smallest index for gridpoints on $\mathcal{Y}$ such that:

$$u \leq \sum_{j=1}^{j^*} \Gamma\left(y_j, y_i^0\right)$$

- Because of randomness, the fraction of households with income values y will never be exactly equal to $\Gamma^*(y)$.

Monte-Carlo simulation

- Correction: you can make an adjustment where you (i) check for which values of y you have an excess relative to the stationary share $\Gamma^* (y)$ and reassign the status of individuals with other realizations of y (for which you have less than the stationary share) appropriately.

- At every t compute the vector M^t of statistics of the wealth distribution (e.g., mean, variance, IQR, etc...).

- Stop if M^t and M^{t-1} are close enough.

- Then A is just the mean of wealth holding in the final sample.

Speed vs stability of the algorithm

- Two tricks that make the algorithm a lot more stable, albeit slower

1. Don't use gradient methods to solve for r^* in the outer loop. Use bisection. Given r^0, to obtain the new candidate r^1 bisect between r^0 and $\left[F_K(A\left(r^0\right)) - \delta\right]$

$$r^1 = \frac{1}{2}\left\{r^0 + \left[F_K(A\left(r^0\right)) - \delta\right]\right\}$$

which are, by construction, on opposite sides of the steady-state interest rate r^*. Even better, use "dampening" in updating:

$$r^1 = \omega r^0 + (1-\omega)\left[F_K(A\left(r^0\right)) - \delta\right]$$

with weight $\omega = 0.8 - 0.9$ for example.

Speed vs stability of the algorithm

2. When you resolve the household problem for a new value of the interest rate r^{n+1}, do not initialize the decision rule from the one corresponding to r^n obtained in the previous loop. Always start from the same guess to avoid propagation error in wealth distribution. Slower, but safer.