The Sequence Space

Juan Herreño   Johannes Wieland

UCSD, Spring 2024

Outline

1. Introduction
2. Krussel-Smith (1998) Economy
3. Fake News Algorithm

OUTLINE

1. INTRODUCTION
2. KRUSSEL-SMITH (1998) ECONOMY
3. FAKE NEWS ALGORITHM

Solving Models

- Hard when there is rich heterogeneity.
- Infinitely dimensional state space from the distribution of agent's endogenous state variables—how to handle it?
- In macro, we are often interested what happens when aggregate shock X hits the economy (monetary, fiscal, etc).
- State space: must forecast infinite-dimensional distribution to estimate prices / quantities relevant to agents' problem.
- Sequence space: must find price / quantity sequences that solve market clearing conditions.

$\Rightarrow$ Sequence space is of dimension $T \times$ number of variables, versus infinite dimensional state space.

- Cost: linearization / perfect foresight.

Sequence Space Methods

- Sequence Space methods mean we use linear algebra to solve models with rich heterogeneity.
- But how do we get matrices like these:

$$\nabla_{\boldsymbol{Y}} \mathbf{C} = \begin{pmatrix} \frac{\partial C_0}{\partial Y_0} & \frac{\partial C_0}{\partial Y_1} & \cdots \\ \frac{\partial C_1}{\partial Y_0} & \frac{\partial C_1}{\partial Y_1} & \cdots \\ \frac{\partial C_2}{\partial Y_0} & \frac{\partial C_2}{\partial Y_1} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}$$

- Need to solve the household consumption problem: hard.
- Auclert, Bardoczy, Rognlie, Straub (2021): efficient algorithm that exploits model structure and linearity to get the Jacobians of the model really fast!

OUTLINE

1 INTRODUCTION

2 KRUSSEL-SMITH (1998) ECONOMY

3 FAKE NEWS ALGORITHM

Household

- Preferences:

$$V_t(e, a_-) = \max_{c,k} u(c) + \beta \sum_{e'} V_{t+1}(e', k) P(e, e')$$

- Budget constraint

$$c + k = (1 + r_t) k_- + w_t e n$$

- Borrowing constraint

$$k \geq 0$$

- Optimal policy:

$$c_t^*(e, k_-),\ k_t^*(e, k_-)$$

- Functions of $\{r_t, w_t\}_{t \geq 0}$.

Aggregating Consumer Problem

- Distribution of capital and productivity:

$$D_{t+1}(e', K) = \sum_e D_t(e, k_t^{*-1}(e, K))P(e, e')$$

- Given D_0, $D_{t+1}(e', K)$ are also functions of $\{r_t, w_t\}_{t\geq 0}$.
- Aggregate capital holdings are therefore also a function of $\{r_t, w_t\}_{t\geq 0}$.:

$$\mathscr{K}_t(\{r_s, w_s\}_{s\geq 0}) = \sum_e \int_{k_-} k_t^*(e, k_-)D_t(e, dk_-)$$

- The function $\mathscr{K}_t$ maps an aggregate sequence $\{r_t, w_t\}_{t\geq 0}$ into another aggregate sequence $\{K_t\}_{t\geq 0}$.
- The dimensionality of this problem is the length of the sequence T.

Firm and Market Clearing

- Keep this simple:

$$\begin{aligned} Y_t &= Z_t K_{t-1}^{\alpha} N_t^{1-\alpha} \\ r_t &= \alpha Z_t K_{t-1}^{\alpha-1} N_t^{1-\alpha} - \delta \\ w_t &= (1-\alpha) Z_t K_{t-1}^{\alpha} N_t^{-\alpha} \end{aligned}$$

- Labor supply fixed in this problem so

$$N_t = \sum \pi(e) en$$

- Market for capital has to clear

$$\begin{aligned} &H_t(\boldsymbol{K}, \boldsymbol{Z}) \\ &\equiv \mathscr{K}_t(\mathbf{r}, \mathbf{w}) \\ &\qquad - K_t = 0 \end{aligned}$$

DAG Representation

IRFs

- You know how to compute $\nabla_{\mathbf{K}}\mathbf{r}, \nabla_{\mathbf{U}}\mathbf{r}, \nabla_{K}\mathbf{w}, \nabla_{U}\mathbf{w}$ from firm block.
- You know how to compute $\nabla_{\mathbf{K}}\mathbf{H}, \nabla_{\mathbf{U}}\mathbf{H}$ from the market clearing block.
- How do we compute $\nabla_{\mathbf{r}}\mathscr{K}$ and $\nabla_{\mathbf{w}}\mathscr{K}$?

Outline

1. Introduction
2. Krussel-Smith (1998) Economy
3. Fake News Algorithm

Fake News Algorithm

- Goal: efficient computation of Jacobians.
- Framework:

$$\begin{aligned} \boldsymbol{v}_t &= v(\boldsymbol{v}_{t+1}, \boldsymbol{X}_t) \\ \boldsymbol{D}_{t+1} &= \Lambda(\boldsymbol{v}_{t+1}, \boldsymbol{X}_t)' \boldsymbol{D}_t \\ \boldsymbol{Y}_t &= y(\boldsymbol{v}_{t+1}, \boldsymbol{X}_t)' \boldsymbol{D}_t \end{aligned}$$

  - $\boldsymbol{X}_t$ are inputs (e.g., Z_t).
  - $\boldsymbol{Y}_t$ are outputs (e.g., C_t, K_t).
  - $\boldsymbol{D}_t$ is the discretized distribution.
- In sequence space:

$$\boldsymbol{Y} = h(\boldsymbol{X})$$

- Want: $\mathscr{J} = \nabla h_{\boldsymbol{Y}}(\boldsymbol{X}^{SS})$

Brute Force

- Column s of $\mathscr{J}^Y_X$ is IRF of outcome (e.g., capital) to a one-time shock (e.g. productivity) at s, $\boldsymbol{X}^s = \boldsymbol{X}_{ss} + \boldsymbol{e}^s dx \equiv \boldsymbol{X}_{ss} + d\boldsymbol{X}^s$.
- Solve problem backward given the known sequence of $\boldsymbol{X}$ to get value functions $\boldsymbol{v}_t$, policy functions $\boldsymbol{y}^s_t$, and transition matrices $\Lambda(\boldsymbol{v}_{t+1}, \boldsymbol{X}_t)'$.
- Given optimal policy functions iterate distribution of agents forward starting from steady state,

$$\boldsymbol{D}^s_{t+1} = (\Lambda^s_t)\boldsymbol{D}^s_t$$

- Aggregate policy functions using the distribution to get outcome

$$\mathscr{Y}^s_t = (\boldsymbol{y}^s_t)'\boldsymbol{D}^s_t$$

- Repeat for $s = 0, \ldots, T$ to get all columns of $\mathscr{J}$.
- This "brute force" method works but is very slow.

Tasks

- Get the first column of $\mathscr{J}_{\mathbf{r}}^{\mathscr{K}}$.

- Get the second column of $\mathscr{J}_{\mathbf{r}}^{\mathscr{K}}$.

Algorithm: Efficient backward step

- Lemma 1: for any $s \geq 1$, $t \geq 1$:

$$\boldsymbol{y}_t^s = \begin{cases} \boldsymbol{y}_{ss} & s < t \\ \boldsymbol{y}_{T-1-(s-t)}^{T-1} & s \geq t \end{cases}, \qquad \Lambda_t^s = \begin{cases} \Lambda_{ss} & s < t \\ \Lambda_{T-1-(s-t)}^{T-1} & s \geq t \end{cases}$$

- Policy functions at t for shock at $t+s$ the same as policy function at 0 for shock at s.

$\Rightarrow$ Policy functions need to be solved backwards only once starting with a shock at $T-1$.

- For any time after the shock, the policy functions are the same as in steady state.

- Key: policy function cannot directly depend on distribution of agents

Fake News

- Denote the sequence $\boldsymbol{\varepsilon}^s$ in which $\varepsilon_s = 1$ and zero otherwise:

$$\varepsilon_t = \begin{cases} 1 & t = s \\ 0 & t \geq 0, t \neq s \end{cases}$$

- Let η_t^s be a fake news shock for ε_t:
  - At t learn that $\varepsilon_s = 1$ with certainty.
  - At $t+1$ lean that $\varepsilon_s = 0$.
  - $\mathbb{E}_k \eta_t^s = 0$ for all $k < t$.
- Lemma: The sequence $\boldsymbol{v}^s$

$$v_t = \begin{cases} \eta_t^{s-t} & t < s \\ 1 & t = s \\ 0 & t > s \end{cases}$$

takes on the same expected values and realized values as $\boldsymbol{\varepsilon}^s$.

Advantage of Fake News

- Given linearization only care about expected values and realized values.
- Rather than compute IRFs to $\boldsymbol{\varepsilon}^s$ for $s \geq 0$, we need IRFs to fake news shocks η^s for $s \geq 0$ and contemporaneous shock $\boldsymbol{\varepsilon}^0$.
- This is more efficient because η^s and $\boldsymbol{\varepsilon}^0$ share a common feature: the policy function only changes in the period the shock becomes known and then reverts. From then on we only need to solve the distribution forward using steady state policy functions.

Algorithm: Fake News

- Define the difference in outcomes at t for shock at s versus outcomes at $t-1$ for shock at $s-1$.

$$\mathscr{F}_{t,s} dx \equiv d\mathscr{Y}_t^s - d\mathscr{Y}_{t-1}^{s-1}$$

- Not zero: policy function the same, but distribution different.
- Lemma:

$$\mathscr{F}_{t,s} dx = \boldsymbol{y}_{ss}' (\Lambda_{ss}')^{t-1} d\boldsymbol{D}_1^s$$

- Policy functions the same. The only difference is the initial distribution.
- To a first order, the distribution affects outcomes as if all agents followed their steady state policy functions.

Fake News Interpretation

$$\mathscr{F}_{t,s}dx = \boldsymbol{y}_{ss}'(\Lambda_{ss}')^{t-1}d\boldsymbol{D}_1^s$$

- Agents at $t = 0$ were told a shock will happen at $t = s$.
- They change their optimal policy rules resulting in a new distribution of outcomes at $t = 1$, $d\boldsymbol{D}_1^s$.
- At $t = 1$ agents learn (to their surprise) that the shock does not happen.
- So all policy rules revert to steady state.
- But the distribution has not reverted to steady state and will affect economic outcomes.
- $\boldsymbol{y}_{ss}'(\Lambda_{ss}')^{t-1}d\boldsymbol{D}_1^s$ traces out IRF from $t = 1$ onwards to a first order.
- Requires only one updating step in solving $d\boldsymbol{D}_1^s$

Fake News Matrix

- Define the fake news matrix as

$$\mathscr{F}_{t,s}dx \equiv \begin{cases} d\mathscr{Y}_0^s & t = 0 \\ \boldsymbol{y}_{ss}'(\Lambda_{ss}')^{t-1}d\boldsymbol{D}_1^s & t \geq 1 \end{cases}$$

- Then the Jacobian of h is given by

$$\mathscr{J}_{t,s} = \sum_{k=0}^{\min\{s,t\}} \mathscr{F}_{t-k,s-k}$$

- $t = 0$ follows by definition.
- Why does this make sense?

Fake News Matrix Examples

- $t > 0, s = 0$:

$$\mathscr{J}_{t,0} = \boldsymbol{y}'_{ss}(\Lambda'_{ss})^{t-1} d\boldsymbol{D}_1^0$$

  - Shock at 0 affected distribution at 1. Now trace out this effect.

- $t = 1, s = 1$:

$$\mathscr{J}_{1,1} = d\mathscr{Y}_0^0 + \boldsymbol{y}'_{ss} d\boldsymbol{D}_1^1$$

  - Sum of surprise contemporaneous shock and fake news shock.

- $t = 1, s = 2$:

$$\mathscr{J}_{1,2} = d\mathscr{Y}_0^1 + \boldsymbol{y}'_{ss} d\boldsymbol{D}_1^2$$

  - As if we got news today about shock tomorrow, but taking into account change in distribution given that news was known earlier.

$$\mathscr{J}_{3,4} = d\mathscr{Y}_0^1 + \boldsymbol{y}'_{ss} d\boldsymbol{D}_1^2 + \boldsymbol{y}'_{ss}(\Lambda'_{ss}) d\boldsymbol{D}_1^3 + \boldsymbol{y}'_{ss}(\Lambda'_{ss})^2 d\boldsymbol{D}_1^4$$

Bottom Line

- $t = 1, s = 0$:

$$\mathscr{J}_{t,s} = \mathscr{J}_{t-1,s-1} + \mathscr{F}_{t,s}$$

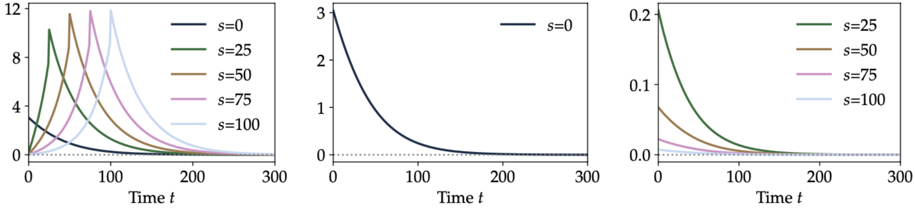
  - A shock at s from time t looks very similar to a shock at $s-1$ from time $t-1$.
  - Policy functions are exactly the same.
  - Only difference is that the distribution at t is different from $t-1$, which is captured by the fake news term.
- The objects we need to calculate this are:
  1. Solve consumer problem backwards once for T periods to get policy functions $\{\boldsymbol{y}_{T-1-s}^{T-1}, \boldsymbol{\Lambda}_{T-1-s}^{T-1}\}_{s=0}^{T-1}$.
  2. Get $\{d\mathscr{Y}_0^s\}_{s=0}^T$ by combining policy functions for $s = 0, ..., T$ with steady state distribution.
  3. Compute $d\boldsymbol{D}_1^s$ using steady state distribution and $s = 0, ..., T$ policy functions.

$\Rightarrow$ Very small number of steps / computations to get all of the Jacobian and therefore the IRFs.

RESULT

Figure 2: Jacobian $\mathcal{J}^{\mathcal{K},r}$ and fake news matrix $\mathcal{F}^{\mathcal{K},r}$ in the Krusell-Smith model.

(a) Columns of the Jacobian $\mathcal{J}^{\mathcal{K},r}$ (b) First column of $\mathcal{F}^{\mathcal{K},r}$ (c) Later column of $\mathcal{F}^{\mathcal{K},r}$

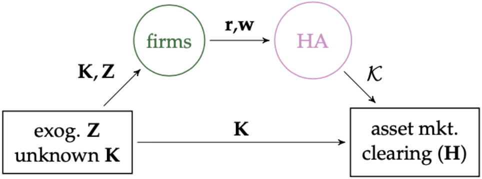


Dimensionality Reduction

- The dimensionality of the problem is $T \times$ number of variables, but with large T and many variables this can get complicated.
- Solution: substitute out for some variables, like we did for r_t, w_t earlier.
- This can be automated by writing the model in separate blocks.
- E.g., firm block $r_t = \alpha Z_t K_{t-1}^{\alpha-1} N_t^{1-\alpha} - \delta$ and $w_t = (1-\alpha) Z_t K_{t-1}^{\alpha} N_t^{-\alpha}$ can be used to solved out for r_t, w_t as a function of Z_t, K_t, N_t.
- Each block takes inputs (e.g., Z_t, K_t, N_t) and computes outputs (r_t, w_t).

Directed Acyclical Graph

Figure 3: DAG representation of Krusell-Smith economy



- The computer can do the substitution if we write the model as a Directed Acyclical Graph (DAG).
- Dimensionality $T \times 2$ even though we have 4 endogenous variables.
- Restriction: cannot have cycle. E.g., if K was not an unknown, the graph would be cylical.
- Always(?) possible to write model as DAG.

More

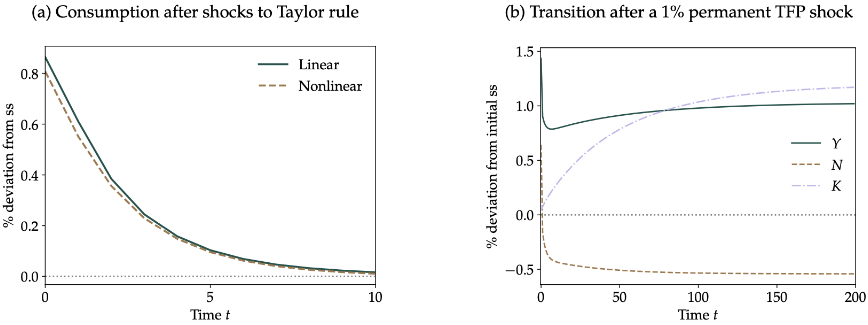
- How to pick T: check if higher T matters for IRFs. Authors recommend $T = 300 - 1000$.

- Estimation: Can compute moments and likelihood quickly from Jacobians. Limit is how quickly you can recompute Jacobians given new parameter values.

- Non-linear perfect foresight dynamics.

More

Figure 7: Nonlinear impulse responses and transitional dynamics for the two-asset HANK model



- See also McKay and Wieland (2022, Econometrica) for how to implement a ZLB.