Linearizing Heterogeneous Agent Models in Aggregates

Methods by
Boppart et al. (2018), Auclert et al. (2021), and Bayer and Luetticke (2020)

Presented by Marlon Azinović

Advanced Topics in Macroeconomics
Class by Professor José Víctor Ríos Rull
University of Pennsylvania

12.02.2023

Plan

1. Numerical solution methods.
2. What does it mean to “linearize in aggregates”?
3. Reminder: MIT shocks.
4. Reminder: what did Boppart et al. (2018) figure out?
5. What did Auclert et al. (2021) do?
6. What did Bayer and Luetticke (2020) do?

Numerical solution methods

- if we don't solve the model precisely, our results will always depend on both: the model, and the solution method.
- we care about economics implied by our model, so need to know to what extent the solution method influences results.

What does it mean to "linearize in aggregates"?

E.g. Krusell and Smith (1998):

Households solve:

$\max \mathrm{E}_0 \left[\sum_t \beta^t u(c_{i,t})\right]$

$c_{it} = w_t e_{it} + (1 + r_t)k_{it-1} - k_{it}$

$k_{it} \geq 0$

assume e follows a first order Markov process.

Firm problem gives:

$r_t = \alpha Z_t K_{t-1}^{\alpha-1} - \delta$

$w_t = (1 - \alpha) Z_t K_{t-1}^{\alpha}$

The idiosyncratic state:

$x = (e, k_-)$.

The aggregate state:

$X = (Z, D(e, k_-))$,

Z: exog. shock (AR(1))

$D(e, k_-)$: wealth distribution.

Policy functions:

$f((e, k_-), (Z, D(e, k_-))) = f(x, X)$.

First order expansion around X^{ss}:

$f(x, X) \approx f(x, X^{ss}) + (X - X^{ss}) f_X(x, X^{ss})$

still hard if X high dimensional and many or complicated f's.

We could make further approximations: decompose

$$X := (Z, \underbrace{D^1(e)}_{\text{marg. dist.}}, \underbrace{D^2(k)}_{\text{marg. dist.}}, \underbrace{\Theta(D^1(e), D^2(k))}_{\text{Copula}}):$$

$f(x, X) \approx f(x, X^{ss}) + (Z - Z^{ss}) f_Z(x, (Z, D^1, D^2, \Theta)^{ss})$

or

$f(x, X) \approx f(x, X^{ss})$
$+(Z - Z^{ss}) f_Z(x, (Z, D^1, D^2, \Theta)^{ss})$
$+(D^1 - D^{1,ss}) f_{D^1}(x, (Z, D^1, D^2, \Theta)^{ss})$
$+(D^2 - D^{2,ss}) f_{D^2}(x, (Z, D^1, D^2, \Theta)^{ss})$

$\Rightarrow$ One of the two dimensionality reductions in Bayer and Luetticke (2020)

Reminder: MIT shock

- Assume all aggregates are at steady state value at $t = 0$ and $t = T$: $X_{-1} = X_T = X^{ss}$, $D_{-1} = D_T = D^{ss}$.
- Take an exogenous sequence of shocks, *e.g.* $\log(Z_0) = 0.01$, $\log(Z_{0<t<T}) = \rho \log(Z_{t-1})$.
- Task: find the consistent endogenous variables.
  1. guess sequence of aggregate capital $\{K_t\}_{t=0}^T$.
  2. compute sequence of wages and returns $\{w_t\}_{t=0}^T$, $\{r_t\}_{t=0}^T$.
  3. iterate backwards from $t = T$ to get a sequence of Hh policy functions $\{c_t(e, k_-)\}_{t=0}^T$, $\{k_t(e, k_-)\}_{t=0}^T$.
  4. iterate forward to get a sequence of distributions $\{D_t(e, k_-)\}_{t=0}^T$.
  5. compute the sequence of household sides aggregates, $\{\mathcal{C}_t, \mathcal{K}_t\}_{t=0}^T$
  6. compute sequence of market clearing errors $\{K_t - \mathcal{K}_t\}_{t=0}^T$.
  7. improve guess.
- Tension: agents first don't know about the shock, but then know exactly how things will go back to normal.
- Beauty: everything just depends on sequences of aggregates! Het. household side extremely tractable.

Exploiting MIT shocks in heterogeneous-agent economies: the impulse response as a numerical derivative

Boppart, Krusell, Mitman (2018)

1. Impulse response to a small MIT shock $\Leftrightarrow$ solution when linearizing in aggregates.

$\Rightarrow$ do not need to derive the first order Taylor terms *i.e.* the huge state-space Jacobian!

$\Rightarrow$ do not need to bother with a potentially super-high-dimensional representation of the aggregate state when solving the Hh problem!

$\Rightarrow$ linear scaling in number of shocks.

2. **Check the accuracy**:
   The impulse response function for different shock sizes (and signs) need to be scaled versions of each other.

Questions?

Using the Sequence-Space Jacobian to Solve and Estimate Heterogeneous-Agent Models

Auclert, Bardóczy, Rognlie, Straub (2021), slides follow slides by Ludwig Straub.

Q: How should we solve heterogeneous-agent general equilibrium models with aggregate shocks in discrete time (if we are willing to linearize in aggregates) ?

1. Write your model as a DAG
2. Compute sequence-space Jacobians

New : provide fast and convenient way to solve a linear system (as Reiter, 2009) but in sequence space (as Boppart et al., 2018)

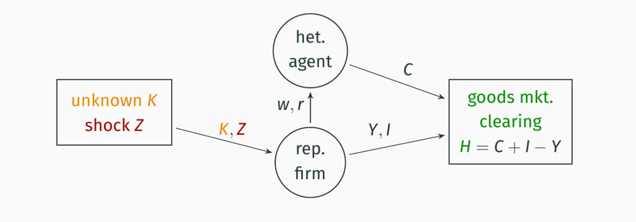
3. Use them for applications:
   - Estimation
   - Local determinacy
   - Nonlinear MIT shocks

Also helpful:

- Provide very clear and user friendly code! See `https://github.com/shade-econ/sequence-jacobian`
- Directed acyclic graphs (DAGs) are very helpful concept to think about heterogeneous agent macro models

Steps

1. Write the model as a collection of blocks along a DAG



2. Compute the Jacobian of each block

$$\mathcal{J}_{t,s}^{C,w} = \frac{\partial C_t}{\partial w_s}, \; \mathcal{J}_{t,s}^{C,r} = \frac{\partial C_t}{\partial r_s}, \; \mathcal{J}_{t,s}^{Y,K} = \frac{\partial Y_t}{\partial K_s}, \dots \tag{1}$$

3. The Jacobians are almost all you need for many applications (like MIT shock IRFs)

Step 1, DAG: Heterogeneous household block

Consider a Krusell and Smith (1998) model with perfect foresight for aggregates as example.

A block maps a sequence of inputs to a sequence of outputs.

The heterogeneous household block maps $\{r_t, w_t\}_{t=0}^{\infty} \to \{C_t\}_{t=0}^{\infty}$.

- Let $\Pi(e'|e)$ be an exogenous Markov chain for skills, such that $L = 1$,
- Households solve

$$\max \mathrm{E}_0 \left[\sum_t \beta^t u(c_{i,t}) \right]$$

$$c_{it} = w_t e_{it} + (1 + r_t) k_{it-1} - k_{it}$$

$$k_{it} \geq 0$$

$\to$ Given initial distribution $D_0(e, k_-)$, the path of aggregate consumption $C_t := \sum_{e,k_-} c_t(e, k_-) D_t(e, k_-)$ only depend on $\{r_t, w_t\}_{t=0}^{\infty} \to \{C_t\}_{t=0}^{\infty}$

Step 1, DAG: Representative firm block

The representative firm block maps $\{K_t, Z_t\}_{t=0}^{\infty} \to \{Y_t, I_t, r_t, w_t\}_{t=0}^{\infty}$.

- $Y_t = Z_t K_t^{\alpha}$
- $I_t = K_t - (1-\delta)K_{t-1}$
- $r_t = \alpha Z_t K_t^{\alpha-1} - \delta$
- $w_t = (1-\alpha) Z_t K_{t-1}^{\alpha}$

$\rightarrow$ We can anlytically map the sequences $\{K_t, Z_t\}_{t=0}^{\infty}$ into sequences $\{Y_t, I_t, r_t, w_t\}_{t=0}^{\infty}$.

Step 1, DAG: Good market clearing block

The good market clearing block maps $\{Y_t, I_t, C_t\}_{t=0}^{\infty} \to \{H_t\}_{t=0}^{\infty}$.

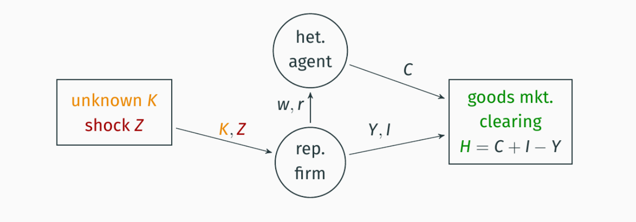
- $$H_t = C_t + I_t - Y_t$$

$\to$ We can anlytically map the sequences $\{C_t, I_t, Y_t\}_{t=0}^{\infty}$ into sequences $\{H_t\}_{t=0}^{\infty}$.

Step 1, DAG: Model

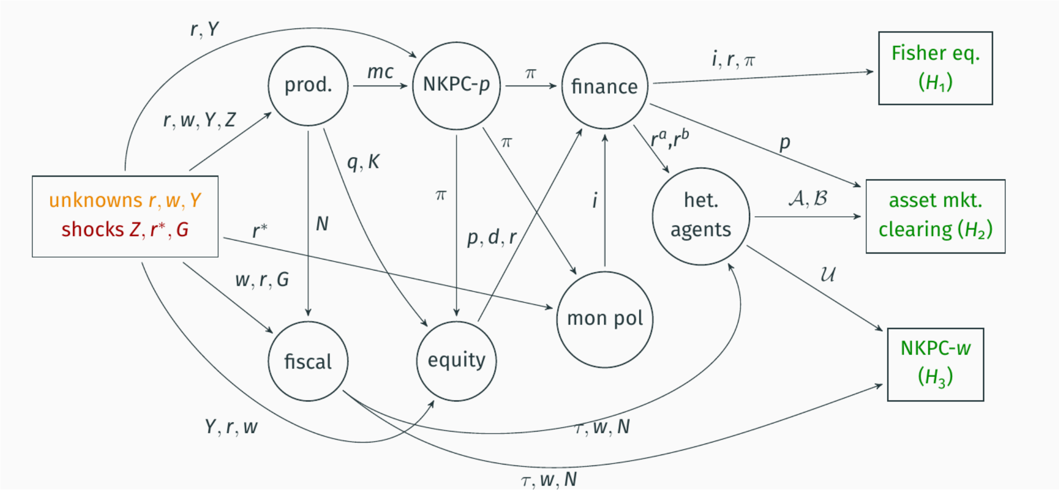
We can view the model as a set of blocks, arranged along a DAG:

- some inputs are **exogenous shocks** (like $\{Z_t\}$)
- some inputs are endogenous **unknowns** (like $\{K_t\}$)
- some outputs are **target** sequences that must equal 0 in GE (like $\{H_t\}$)



- We can collapse the DAG into a mapping $\{\{Z_t\}, \{K_t\}\} \to \{H_t\}$, where the GE path of $\{K_t\}$ satisfies $\{H_t\} = 0$.

Step 1, DAG: two asset HANK as another example



Step 2, Jacobians: What do we want?

How does the model react to a sequence of shocks d$\mathbf{Z}$?
$d\mathbf{K}$ needs to be consistent with $\mathbf{H}(\mathbf{K}, \mathbf{Z}) = 0$.
By the implicit function theorem

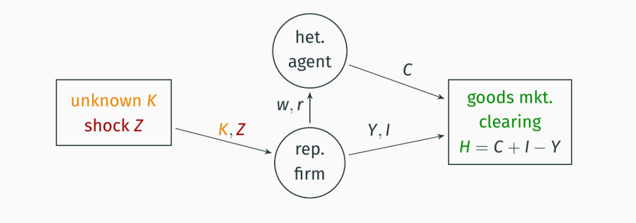
$$d\mathbf{K} = -\left(\underbrace{\frac{\partial \mathbf{H}}{\partial \mathbf{K}}}_{(n_{targets}T)\times(n_{targets}T)} \right)^{-1} \left(\frac{\partial \mathbf{H}}{\partial \mathbf{Z}} \right) d\mathbf{Z} \tag{2}$$

So what's a good way to get $\frac{\partial \mathbf{H}}{\partial \mathbf{K}}$ and $\frac{\partial \mathbf{H}}{\partial \mathbf{Z}}$?

Compute the sequence space Jacobians for each block and chain them with the chain rule!

Step 2, Jacobians: Block Jacobians

A block Jacobian is the derivatives of its outputs wrt its inputs.



- Heterogeneous agent block: two Jacobians: $\mathcal{J}^{C,w}_{t,s} := \frac{\partial C_t}{\partial w_s}$, $\mathcal{J}^{C,r}_{t,s} := \frac{\partial C_t}{\partial r_s}$
- Firm block: eight Jacobians: $\underbrace{\mathcal{J}^{w,\mathbf{K}}}_{T\times T}, \mathcal{J}^{w,\mathbf{Z}}, \mathcal{J}^{r,\mathbf{K}}, \mathcal{J}^{r,\mathbf{Z}}, \mathcal{J}^{Y,\mathbf{K}}, \mathcal{J}^{Y,\mathbf{Z}}, \mathcal{J}^{I,\mathbf{K}}, \mathcal{J}^{I,\mathbf{Z}}$
- Market clearing block: three Jacobians: $\mathcal{J}^{\mathbf{H},C}$, $\mathcal{J}^{\mathbf{H},I}$, $\mathcal{J}^{\mathbf{H},Y}$
- Now we use the chain rule to get

$$\frac{\partial \mathbf{H}}{\partial \mathbf{K}} = \mathcal{J}^{\mathbf{H},C}\mathcal{J}^{C,r}\mathcal{J}^{r,\mathbf{K}} + \mathcal{J}^{\mathbf{H},C}\mathcal{J}^{C,w}\mathcal{J}^{w,\mathbf{K}} + \mathcal{J}^{\mathbf{H},I}\mathcal{J}^{I,\mathbf{K}} + \mathcal{J}^{\mathbf{H},Y}\mathcal{J}^{Y,\mathbf{K}} \tag{3}$$

$$\frac{\partial \mathbf{H}}{\partial \mathbf{Z}} = \mathcal{J}^{\mathbf{H},C}\mathcal{J}^{C,r}\mathcal{J}^{r,\mathbf{Z}} + \mathcal{J}^{\mathbf{H},C}\mathcal{J}^{C,w}\mathcal{J}^{w,\mathbf{Z}} + \mathcal{J}^{\mathbf{H},I}\mathcal{J}^{I,\mathbf{Z}} + \mathcal{J}^{\mathbf{H},Y}\mathcal{J}^{Y,\mathbf{Z}} \tag{4}$$

Step 3, use Jacobians: Getting IRFs of interest

Suppose we want GE response to shock $d\mathbf{Z}$ (e.g. AR(1)).
$\mathbf{K}$ needs to be consistent with $\mathbf{H}(\mathbf{K}, \mathbf{Z}) = 0$.
Equipped with all the Jacobians, we can first get

$$d\mathbf{K} = -\left(\underbrace{\frac{\partial \mathbf{H}}{\partial \mathbf{K}}}_{(n_{targets}T)\times(n_{targets}T)} \right)^{-1} \left(\frac{\partial \mathbf{H}}{\partial \mathbf{Z}}\right) d\mathbf{Z} \tag{5}$$

and can than get any other GE impulse response we want from plain matrix multiplications. E.g. IRF of output:

$$d\mathbf{Y} = \mathcal{J}^{Y,\mathbf{K}} d\mathbf{K} + \mathcal{J}^{Y,\mathbf{Z}} d\mathbf{Z}. \tag{6}$$

Step 3, use Jacobians: Applications

- Estimation: especially fast if something is estimated where the Jacobians don't have to be recomputed (like estimating the shock process).
- Determinacy: the $n_{targets}T \times n_{targets}T$ matrix $\frac{\partial \mathbf{H}}{\partial \mathbf{K}}$ is invertible if the model is local determinate (can study determinacy)
- MIT shocks: Can use the s.s. Jacobian $\frac{\partial \mathbf{H}}{\partial \mathbf{K}}$ to rapidly solve $\mathbf{H}(\mathbf{K}, \mathbf{Z}) = 0$ for the nonlinear perfect-foresight $\mathbf{K}$, given an MIT-shock $d\mathbf{Z}$. This is basically a Newton method with using the s.s. Jacobian as approximation of the Jacobian.

How to get Jacobians fast?

We saw that once we have the Jacobians, we have all we need. So how to get them?

- For simple blocks (like the firm block) they are simple and sparse. This has to be exploited!
- For heterogeneous agent blocks more complicated.

Jacobians for Simple Blocks

Remember, the representative firm block maps $\{K_t, Z_t\}_{t=0}^{\infty} \to \{Y_t, I_t, r_t, w_t\}_{t=0}^{\infty}$.

- $Y_t = Z_t K_{t-1}^{\alpha}$
- $I_t = K_t - (1-\delta)K_{t-1}$
- $r_t = \alpha Z_t K_{t-1}^{\alpha-1} - \delta$
- $w_t = (1-\alpha) Z_t K_{t-1}^{\alpha}$

$\Rightarrow$ the firm block jacobians are hence simple and sparse.

E.g.

$$\mathcal{J}_{t,s}^{w,K} := \frac{\partial w_t}{\partial K_s} = \begin{cases} (1-\alpha)\alpha Z_t K_{t-1}^{\alpha-1} & \text{if } s = t-1 \\ 0 & \text{else} \end{cases} \tag{7}$$

Jacobians for het. agent blocks

- Assume at $t = 0$ the model is in steady state and we want to know how aggregate consumption at time t responds to an anticipated wage change at s: $\mathcal{J}_{t,s} := \mathcal{J}^{C,w}_{t,s} = \frac{\partial C_t}{\partial w_s} \forall t, s \in \{0, \dots T-1\}$.
- Direct algorithm: perturb $w_s = w + \epsilon$, then
  1. iterate backwards to get the perturbed policies: $\mathbf{c}^s_t(e, k_-), \mathbf{k}^s_t(e, k_-)$.
  2. iterate forward to get the perturbed distribution $D^s_t(e, k_-)$.
  3. get perturbed aggregate consumption: $C^s_t = \sum_{e,k_-} \mathbf{c}_t(e, k_-) D_t(e, k_-)$.
  4. Compute $\mathcal{J}_{t,s} = \frac{C^s_t - C}{\epsilon}$
- Slow! Because 1. - 4. has to be T times. Once for each s.
- Paper proposes fake news algorithm that is T times faster.
  - requires only a single forward and a single backward iteration.
  - key idea: exploit time symmetries around the steady sate.

Fake news Algorithm

- Consider: $\mathcal{J}_{t,s}$ corresponds to the response of aggregate consumption at time t to the news that at time s the wage will be higher.
- Notice that, if we had $\mathcal{F}$, we could reconstruct $\mathcal{J}$ easily. Where

$$\mathcal{J} = \begin{bmatrix} \mathcal{J}_{00} & \mathcal{J}_{01} & \mathcal{J}_{02} & \cdots \\ \mathcal{J}_{10} & \mathcal{J}_{11} & \mathcal{J}_{12} & \cdots \\ \mathcal{J}_{20} & \mathcal{J}_{21} & \mathcal{J}_{22} & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{bmatrix}, \mathcal{F} = \begin{bmatrix} \mathcal{J}_{00} & \mathcal{J}_{01} & \mathcal{J}_{02} & \cdots \\ \mathcal{J}_{10} & \mathcal{J}_{11} - \mathcal{J}_{00} & \mathcal{J}_{12} - \mathcal{J}_{01} & \cdots \\ \mathcal{J}_{20} & \mathcal{J}_{21} - \mathcal{J}_{10} & \mathcal{J}_{22} - \mathcal{J}_{11} & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{bmatrix} \tag{8}$$

  We call $\mathcal{F}$ the fake news matrix. Consider consumption and wage.
- $\mathcal{F}_{t,s} = \begin{cases} \frac{\partial C_t}{\partial w_s} & \text{, if } s = 0 \text{ or } t = 0 \\ \frac{\partial C_t}{\partial w_s} - \frac{\partial C_{t-1}}{\partial w_{s-1}} & \text{, else.} \end{cases}$
- $\mathcal{F}_{t,s}$ corresponds to the response of aggregate consumption at time t to:
  1. at $t = 0$: news that at time s the wage will be higher
  2. at $t = 1$: news that you were tricked and wage at time s won't be higher!
- Can get $\mathcal{F}$ with a single forward and backward iteration.

Why single backward and forward iteration to get $\mathcal{F}$?

- Key observation for fake-news shocks:
  - At $t = 0$ distribution is in steady state $\mathbf{D}_0^s(e, k_-) = \mathbf{D}_{ss}(e, k_-)$. Only policies $\mathbf{c}_0^s(e, k_-) \neq \mathbf{c}_{ss}(e, k_-)$, $\mathbf{k}_0^s(e, k_-) \neq \mathbf{k}_{ss}(e, k_-)$ and consequently the transition of the distribution $\Lambda_0^s \neq \Lambda_{ss}$ react.
  - From $t = 1$ on the policies are the steady state policies $\mathbf{c}_t^s(e, k_-) = \mathbf{c}_{ss}(e, k_-)$, $\mathbf{k}_t^s(e, k_-) = \mathbf{k}_{ss}(e, k_-)$ and as a result the transition matrix for the distribution is the steady state transition matrix $\Lambda_t^s = \Lambda_{ss}$. But because the $t = 0$ policy was different from the steady state policy, we have $\mathbf{D}_1^s(e, k_-) \neq \mathbf{D}_{ss}(e, k_-)$ and consequently $\mathbf{D}_t^s(e, k_-) = \Lambda_{ss}^{t-1}\mathbf{D}_1^s(e, k_-)$.

Why single backward iteration?

Single backward iteration is enough to recover $\mathbf{c}_0^s(e, k_-), \mathbf{k}_0^s(e, k_-)$

1. set $s = T - 1$, a single backward iteration gives us $\mathbf{c}_t^{T-1}(e, k_-), \mathbf{k}_t^{T-1}(e, k_-)$ for $t = 0, \ldots, T-1$.
2. notice that only $(s - t)$ matters for the policy reaction.
   Hence

$$\mathbf{c}_0^s(e, k_-) = \mathbf{c}_{T-1-s}^{T-1}(e, k_-), \tag{9}$$

   which we have!
3. From these, we already have the first row of the fake news matrix: $\mathcal{J}_{0s} = \frac{\partial C_0}{\partial w_s} = \int \mathbf{c}_0^s(e, k_-) D_{ss}(e, k_-) = \mathbf{c}_0' \mathbf{D}_{ss} = \mathcal{F}_{0s}$, and the resulting distribution at date 1: $\mathbf{D}_1^s$.

Why single forward iteration?

- We have $\mathbf{D}_t^s = \underbrace{(\Lambda_{ss}')^{t-1}}_{\text{indep. of s.}} \underbrace{\mathbf{D}_1^s}_{\text{indep. of t.}}$.
- In a single forward iteration, we can therefore compute $C_t^s = \mathbf{c}_{ss}'(\Lambda_{ss}')^{t-1}\mathbf{D}_1^s$.
- Now we are done, because $\mathcal{F}_t^s = \frac{C_t^s - C_{ss}}{\epsilon}$.

Runtimes 1

Algorithm	Krusell-Smith	HD Krusell-Smith	one-asset HANK	two-asset HANK
Direct	**26 s**	**1939 s**	**176 s**	**2107 s**
step 1 (backward)	16 s	1338 s	150 s	1291 s
step 2 (forward)	10 s	601 s	27 s	815 s
Fake news	**0.104 s**	**8.429 s**	**0.646 s**	**5.697 s**
step 1 (backward)	0.067 s	5.433 s	0.525 s	5.206 s
step 2 (forward)	0.010 s	1.546 s	0.021 s	0.122 s
step 3	0.023 s	1.445 s	0.092 s	0.346 s
step 4	0.004 s	0.004 s	0.008 s	0.023 s
Gridpoints n_g	3,500	250,000	3,500	10,500

Solving discrete time heterogeneous agent models with aggregate risk and many idiosyncratic states by perturbation

Bayer and Luetticke (2020)

1. Compute steady state.
2. Identify the most relevant parameters in the representation of the policy function to be linearized (*i.e.* reducing the number of "f" (in this case the number of coefficients in the representation of the policy functions)).
3. Only capture the change due to most relevant changes in the aggregate state by assuming a fixed copula (*i.e.* reducing the dimensionality of the "X" we need to expand w.r.t. (in this case the histogram)).
4. Do linearization or second order expansion of the most relevant f w.r.t the most releveant X as in Schmitt-Grohé and Uribe (2004).

Credits also to: Reiter (2009) and Winberry (2018).

Thank **YOU** !

References I

Auclert, A., Bardóczy, B., Rognlie, M., and Straub, L. (2021). Using the sequence-space jacobian to solve and estimate heterogeneous-agent models. *Econometrica*, 89(5):2375–2408.

Bayer, C. and Luetticke, R. (2020). Solving discrete time heterogeneous agent models with aggregate risk and many idiosyncratic states by perturbation. *Quantitative Economics*, 11(4):1253–1288.

Boppart, T., Krusell, P., and Mitman, K. (2018). Exploiting mit shocks in heterogeneous-agent economies: the impulse response as a numerical derivative. *Journal of Economic Dynamics and Control*, 89:68–92.

Krusell, P. and Smith, Jr, A. A. (1998). Income and wealth heterogeneity in the macroeconomy. *Journal of political Economy*, 106(5):867–896.

Reiter, M. (2009). Solving heterogeneous-agent models by projection and perturbation. *Journal of Economic Dynamics and Control*, 33(3):649–665.

Schmitt-Grohé, S. and Uribe, M. (2004). Solving dynamic general equilibrium models using a second-order approximation to the policy function. *Journal of economic dynamics and control*, 28(4):755–775.

Winberry, T. (2018). A method for solving and estimating heterogeneous agent macro models. *Quantitative Economics*, 9(3):1123–1151.