DeepHAM: A Global Solution Method for Heterogeneous Agent Models with Aggregate Shocks

Jiequn Han[1,2], Yucheng Yang[2], and Weinan E[3,2]

February 5, 2022

[1]Flatiron Institute
[2]Princeton University
[3]Peking University

Link to the paper: https://www.ssrn.com/abstract=3990409

Introduction

- Recent research highlights importance of heterogeneity in macroeconomics.
- Heterogeneous agent (HA) models with aggregate shocks are solved with global Krusell-Smith (KS) method or local perturbation method.

	KS method	Perturbation method
Multiple shocks	No	Yes
Multiple endogenous states	No	Yes
Estimation/Calibration	No	Yes
Large shocks	Yes	No
Risky steady state	Yes	No
Nonlinearity e.g. ZLB	Yes	No

This paper: a new efficient, reliable, and interpretable global solution method for high dimensional HA models with aggregate shocks using deep learning.

Deep Learning for High Dimensional Models

- Deep learning has achieved success in high dimensional problems.
- **Our idea**: use deep learning to "learn" policy function, value function, and distribution representation in high dimensional HA models.
- Three key elements to "learn" high-dim functions:

  1. Deep neural networks to represent function:

  $$f(x) = \mathcal{L}^{out} \circ \mathcal{L}^{N_h} \circ \mathcal{L}^{N_h - 1} \circ \cdots \circ \mathcal{L}^1(x),$$
  $$h_p = \mathcal{L}^p(h_{p-1}) = \sigma\big(W_p h_{p-1} + b_p\big),$$

  σ : element-wise nonlinear activation function: e.g. $\max(0, x)$.

  2. Cast high-dim function into an objective function:

  $$\min_{\theta} \frac{1}{N} \sum_{i=1}^{N} L(f(x_i; \theta)).$$

  3. Efficient optimization: stochastic gradient descent (SGD).

This Paper: DeepHAM Method for N-agent HA Model

1. Use neural networks (NN) to represent value & policy functions.
2. Nest sub-NN of permutation invariant *generalized moments* to represent state distribution.
3. Iteratively update value & policy functions, and *generalized moments*.

Apply DeepHAM to three economies:

1. Krusell-Smith problem: competitive equilibrium.
2. Krusell-Smith with a financial sector (Fernandez-Villaverde et al., 2020).
3. Constrained efficiency problem in HA models with aggregate shocks.

Main features:

1. High accuracy compared to other global solution methods.
2. Efficient computational speed (no curse of dimensionality).
3. Interpretability of distribution representation and function mappings.

Literature

Methodology

Illustration: Krusell and Smith (1998)

- Production economy with a continuum of households: each HH i solves

$$\max_{c_{i,t}\geq 0, a_{i,t+1}\geq \underline{a}} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t u\left(c_{i,t}\right)$$

  subject to budget constraint

$$a_{i,t+1} = w_t \bar{\ell} y_{i,t} + R_t a_{i,t} - c_{i,t}$$

- Idiosyncratic shocks on employment status $y_{i,t}$.
- Representative firm produces $Y_t = Z_t F(K_t, \bar{L})$.
- Aggregate shock $Z_t \sim$ two-state Markov, and enters HH's problem through competitive factor prices:

$$R_t = Z_t \partial_K F(K_t, \bar{L}) - \delta, w_t = Z_t \partial_L F(K_t, \bar{L})$$

Computational Setup: Krusell and Smith (1998)

Curse of dimensionality shows up in recursive form of HH i's problem:

$$V(a_i, y_i, Z, \mathbf{\Gamma}) = \max_{c_i, a_i'} \{u(c_i) + \beta \mathbb{E} V(a_i', y_i', Z', \mathbf{\Gamma}' | y_i, Z)\}$$

subject to budget and borrowing constraints. $\mathbf{\Gamma}$: distribution of all HHs' states.

Krusell-Smith method (KS, 1998; Maliar et al., 2010):

1. Approximate state vector: $\hat{s}_i = (a_i, y_i, Z, m_1)$, where m_1 is first moment of individual asset distribution.
2. Log linear law of motion for m_1:

$$\log(m_{1,t+1}) = A(Z) + B(Z) \log(m_{1t}).$$

Very costly in complex HA models with multiple assets or multiple shocks.

Computational Setup: DeepHAM

- In N-agent Krusell-Smith problem (N finite but large), define cumulative utility for HH i up to t:

$$\widetilde{U}_{i,t} = \sum_{\tau=0}^{t} \beta^{\tau} u\left(c_{i,\tau}\right)$$

- Solve Markov Nash equilibrium: optimal policy $\mathcal{C}(s)$ solve HH i's problem:

$$V(s_{i,0}) = \max_{c_{i,t} \geq 0, a_{i,t+1} \geq \underline{a}} \mathbb{E}_0 \widetilde{U}_{i,\infty} = \max_{\{c_{i,t} \geq 0, a_{i,t+1} \geq \underline{a}\}_{t=0}^{T}} \mathbb{E}_0 \left(\widetilde{U}_{i,T} + \beta^T V(s_{i,T}) \right)$$

subject to other HHs also follow $\mathcal{C}(s)$. $s_i = (a_i, y_i, Z, \mathbf{\Gamma})$, where $\mathbf{\Gamma}$ is distribution of all HHs' states.

- Value function $V(s)$ and policy function $\mathcal{C}(s)$ are parameterized by deep neural networks.

DeepHAM: Outline of Algorithm

Initialization: initial policy $\mathcal{C}^{(0)}$, initial value and policy neural networks with parameters Θ^V and Θ^C. In the k-th iteration:

1. Prepare the stationary distribution $\mu(\mathcal{C}^{(k-1)})$ according to the policy $\mathcal{C}^{(k-1)}$
2. Given Θ^C, update the value function by solving

$$\min_{\Theta^V} \mathbb{E}_{\mu(\mathcal{C}^{(k-1)})} \left[V_{\mathsf{NN}}(s_i; \Theta^V) - \widehat{V}_i \right]^2,$$

   where $\widehat{V}_i = \sum_{\tau=0}^{T_v} \beta^\tau u(c_\tau^i)$ is truncated lifetime utility with large T_v.
3. Given Θ^V, optimize the policy function by solving

$$\max_{\Theta^C} \mathbb{E}_{\mu(\mathcal{C}^{(k-1)})} \left[\sum_{t=0}^{T} \beta^t u(c_{i,t}) + \beta^T V_{\mathsf{NN}}(s_{i,T}; \Theta^V) \right],$$

   following the spirit of fictitious play.

DeepHAM: Value Function Learning

In iteration k, given policy function $\mathcal{C}^{(k-1)}(s)$:

1. Simulate the economy many times for T_v periods ($T_v \gg T$), sample states s from simulations. Then the value of each state s can be approximately calculated as cumulative utility in the following T periods:

$$\widetilde{V}^{(k)}(s) \approx \mathbb{E}\widetilde{U}_T = \mathbb{E}\sum_{\tau=0}^{T} \beta^{\tau} u\left(c_{i,\tau}\right)$$

2. Learn value function $V^{(k)}(s)$ parameterized by deep neural networks with regression.

DeepHAM: Policy Function Learning

In iteration k, given value function $V^{(k)}(s)$, optimize policy function $\mathcal{C}^{(k)}(s)$ in a fictitious play.

- Fictitious play: separate N-agent problem into N individual problems where other agents' strategies are fixed and follow the policy in the past "play".

DeepHAM: Policy Function Learning

In iteration k, given value function $V^{(k)}(s)$, optimize policy function $\mathcal{C}^{(k)}(s)$ in a fictitious play.

- Fictitious play: separate N-agent problem into N individual problems where other agents' strategies are fixed and follow the policy in the past "play".
- Solve the following problem iteratively:
  1. At round $\ell + 1$, everybody's policy $\mathcal{C}^{(k,\ell)}(s)$ is known.
  2. For agent $i = 1$, solve for her optimal policy $\mathcal{C}^{(k,\ell+1)}(s)$:

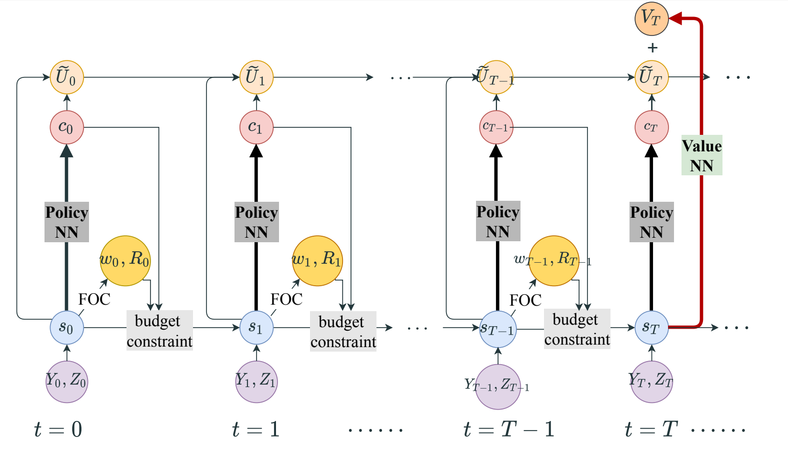
$$\max_{\mathcal{C}^{(k,\ell+1)}(s)} \mathbb{E}_{\mu(\mathcal{C}^{(k-1)})} \left(\widetilde{U}_{i,T} + \beta^T V^{(k)}(s_{i,T}) \right)$$

  subject to others all following $\mathcal{C}^{(k,\ell)}(s)$ in the first T periods (T large).

Optimization is solved on Monte Carlo simulation with N agents on a large number of sample paths in a computational graph.

Computational Graph Formulation for DeepHAM

$$\max_{\Theta^C} \mathbb{E}_{\mu(\mathcal{C}^{(k-1)})} \left(\widetilde{U}_{i,T} + \beta^T V_{\mathsf{NN}}(s_{i,T}; \Theta^V) \right)$$



Budget constraint $a_{i,t+1} = R_t a_{i,t} + w_t \bar{\ell} y_{i,t} - c_{i,t}$. $s_t = (a_{i,t}, y_{i,t}, Z_t, \mathbf{\Gamma}_t)$.

DL-Based Model Reduction and Generalized Moments

- General form of value & policy functions are like (ignore y):

$$V(a_i; a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_N; Z), \quad c(a_i; a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_N; Z)$$

- For permutational invariance, specify its approximated form with symmetry preserving generalized moments $\frac{1}{N}\sum_i \mathcal{Q}_j(a_i)$, with basis function $\mathcal{Q}_1(\cdot)$, ..., $\mathcal{Q}_J(\cdot)$ parameterized by (sub) neural networks:

$$V(a_i; \frac{1}{N}\sum_i \mathcal{Q}_1(a_i), \ldots, \frac{1}{N}\sum_i \mathcal{Q}_J(a_i); Z)$$
$$c(a_i; \frac{1}{N}\sum_i \tilde{\mathcal{Q}}_1(a_i), \ldots, \frac{1}{N}\sum_i \tilde{\mathcal{Q}}_J(a_i); Z)$$

- Special case: $\mathcal{Q}(a) = a$ yields the first moment.
- The algorithm optimally solve generalized moments that matters most for policy and value functions. ("numerically determined sufficient statistics")
- Generalized moments provide interpretabilitiy on how heterogeneity matters in model.

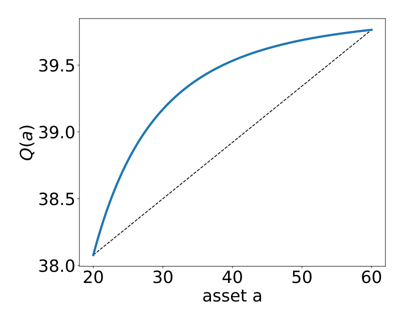
Accuracy Results for Krusell-Smith Problem

Method and Moment Choice	Bellman error	Std of error
KS Method (Maliar et al., 2010)	**0.0253**	0.0002
DeepHAM with 1st moment	0.0184	0.0023
DeepHAM with 1 generalized moments	0.0151	0.0015

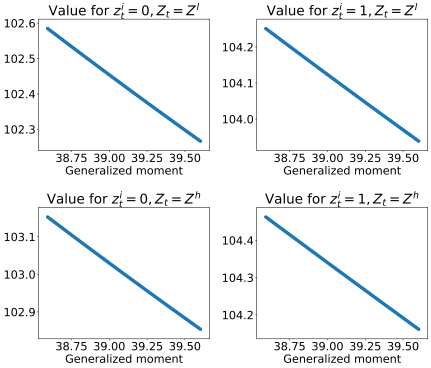
Definition of Bellman Error

- Highly accurate solutions compared to Krusell-Smith (KS) method.
- Even only with first moment as model input, DeepHAM outperform KS method due to better capture of nonlinearity.
- Generalized moment yields more accurate solution than the first moment, as it extract more relevant information.

Interpretation of Generalized Moments



(a) Plot of $\mathcal{Q}_1(a)$



(b) Map $\frac{1}{N}\sum_i \mathcal{Q}_1(a_i)$ to value function

- Basis function is concave in individual asset, while value function is linear wrt generalized moment.
- Heterogeneity matters! Unanticipated redistributive policy shock: asset from rich to poor HH $\Rightarrow$ generalized moment $\uparrow \Rightarrow$ other agents' welfare $\downarrow$. No effect with KS method, as first moment not change.

DeepHAM for Constrained Efficiency Problem

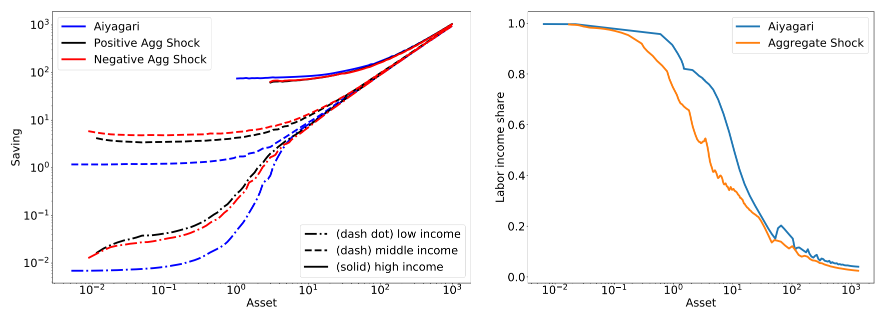
- Constrained efficiency problem is hard to solve in HA models, as planner's state involves agent distribution.
- Literature only solves for HA models without aggregate shocks (Davila et al., 2012; Nuno and Moll, 2018).
- DeepHAM can solve planner's problem as easily as solve competitive equilibrium, only remove the fictitious play procedure.
- We solve constrained efficiency problem of Davila et al. (2012), and same model with aggregate shocks and countercyclical unemployment risk.
- It takes DeepHAM 20 minutes to solve the constrained efficiency problem in Davila et al. (2012) on GPU, which takes conventional methods > 10 hours on CPU.

Constrained Efficiency for HA Models w or w/o Agg Shock

	No aggregate shock		Aggregate shock	
	Market	Constrained Opt.	Market	Constrained Opt.
Average assets	30.635	119.741	34.296	95.811
Wealth Gini	0.864	0.862	0.812	0.878
Consumption Gini	0.615	0.386	0.578	0.388

- Both models: constrained optimal capital $\gg$ capital in competitive equilibrium.
- Why? Overcome pecuniary externality: $K \uparrow \Rightarrow$ wage $\uparrow, R \downarrow$, redistribute from rich HHs to poor HHs (high labor share).
- Constrained optimal capital in model with agg shock $<$ without agg shock.
- Why? Agg shock $\Rightarrow$ precautionary saving $\uparrow$ by poor HHs $\Rightarrow$ labor share lower than model w/o agg shock. So planner raises K less.

Constrained Efficiency for HA Models w or w/o Agg Shock



(c) Policy function **(d)** Labor share distribution

Agg shock ⇒ precautionary saving ↑ by poor HHs ⇒ labor share lower than model w/o agg shock. So planner raises K less in constrained efficient equilibrium.

Conclusion

- We develop DeepHAM, an efficient, reliable, and interpretable deep learning based method to solve HA models with aggregate shocks globally.
- Our method obtains highly accurate solutions, and can be applied to high-dim HA models without curse of dimensionality.
- For the first time, we solve constrained efficiency problem in HA models with aggregate shock.
- Deep learning based model reduction can inform interpretable generalized moments of distribution that matters.
- The method not only solve competitive equilibrium of HA models, but also strategic equilibrium.
- As is happening in other disciplines, such toolkits may become the new generation of model solvers in economics.

Many potential exciting applications!

Thank You!

Comments and questions are welcome!

Contact: jiequnhan@gmail.com

yuchengy@princeton.edu

Link to the paper: https://www.ssrn.com/abstract=3990409

Appendix

Literature

- Solving HA models with aggregate shocks:
    1. Global KS method: Krusell and Smith (1998), Den Haan (2010) project, Fernandez-Villaverde et al. (2019), etc.
    2. Local perturbation method: Reiter (2009), Ahn et al. (2017), Winberry (2018), Bayer and Luetticke (2020); Boppart, Krusell and Mitman (2018), Auclert et al. (2021), etc.
- Deep learning for high dimensional problems:
    1. Stochastic control & PDE: Han and E (2016), Han, Jentzen and E (2018).
    2. Macroeconomics: Duarte (2018), Fernandez-Villaverde et al. (2020, 2021), Maliar et al. (2021), Azinovic et al. (2020), etc.
- How heterogeneity matters in macro: Kaplan and Violante (2018), Kaplan et al. (2018), Auclert (2019), etc.
- Constrained efficiency problem in HA models: Davila et al. (2012), Nuno and Moll (2018), Bhandari et al. (2021), etc.

back

Accuracy Measures: Bellman Equation Errors

For the KS problem, only using solved value function $V(\cdot)$, Bellman equation error is

$$\text{err}_\text{B} = V(a_i, y_i, Z, \mathbf{a}^{-i}, \mathbf{y}^{-i}) - \max_{c_i} \left\{ u(c_i) + \beta \sum_{y', Z', \mathbf{y}'^{-i}} V(a_i', y_i', Z', \widehat{\mathbf{a}'}^{-i}, \mathbf{y}'^{-i}) \times \Pr\left(Z', y'^{i}, \mathbf{y}'^{-i} | Z, y^{i}, \mathbf{y}^{-i} \right) \right\}$$

back