Machine Learning Projection Methods for Macro-Finance Models

Alessandro T. Villa and Vytautas Valaitis

Working Paper

2019

Alejandro Rojas-Bernal

March 2020

Motivation

- The parameterized expectations algorithm (PEA) introduced by Den Haan and Marcet (1990) does not deal well with multicollinearity (Condensed PEA) and over-identification (Forward-States PEA).
- For this reason the paper develops an Artificial-Neural-Network-based Expectations Algorithm (ANN) that endogenously reduces the space state and internalizes the multicollinearity.
- Two implementations to illustrate application:
  1. A Ramsey taxation problem with incomplete markets and optimal maturity structure that suffers from: i) exploding state space; ii) forward-looking constraints; and iii) lagged state variables - high multicollinearity.
  2. An extension to a three country international business cycle model for the Kehoe and Perri (2002) solution to the Bachus et a. (1992) puzzle with endogenous incompleteness and enforcement constraint that shows that perfect risk sharing cannot be implemented.

Model - Second Stage Ramsey Primal Approach

Neoclassical incomplete market stochastic model from Aiyagari, Marcet, Sargent, and Seppälä (2002, JPE), and Faraglia, Marcet, Oikonomou, and Scott (Forthcoming, RES) and adding Epstein-Zin preferences:

$$V_t = \left[(1-\beta)\, U(c_t, l_t)^{1-\rho} + \beta \left(\mathbb{E}_t V_{t+1}^{1-\gamma} \right)^{\frac{1-\rho}{1-\gamma}} \right]^{\frac{1}{1-\rho}}$$

s.t.

$$c_t + \sum_{i=1}^{N} p_{i,t} b_{t+1}^{i} = \sum_{i=1}^{N} p_{i-1,t} b_t^{i} + (1-\tau_t)(1-l_t)$$

with buyback assumption, $c_t + g_t = 1 - l_t$, and

$$p_{i,t} = \mathbb{E}_t \beta^i \underbrace{\left(\frac{V_{t+i}}{\left(\mathbb{E}_t V_{t+i}^{1-\gamma} \right)^{\frac{1}{1-\gamma}}} \right)^{\rho-\gamma}}_{\mathcal{M}_t(V_{t+i})} \left(\frac{U_{t+i}}{U_t} \right)^{-\rho} \frac{U_{c,t+i}}{U_{c,t}}$$

Model - First Stage Ramsey Primal Approach

Given $\{g_t\}_{t=0}^{\infty}$ set τ_t and $\left\{\left\{b_t^i\right\}_{i=1}^{N}\right\}_{t=0}^{\infty}$ to maximize welfare over the competitive equilibrium outcome subject to:

$$\sum_{i=1} p_{i-1,t} b_t^i = \tau_t \left(1 - l_t\right) - g_t + \sum_{i=1}^{N} p_{i,t} b_{t+1}^i$$

$$1 - \tau_t = \frac{U_{l,t}}{U_{c,t}}$$

a sequence of measurability constraints

$$b_t^N \beta^N \beta^N u_{c,t+N} - b_{t+1}^N \beta^{N-1} u_{c,t-1+N} - g_t u_{c,t} + \left(u_{c,t} - v_{l,t}\right)\left(g_t + c_t\right) = 0$$

$$\frac{\bar{M}_i}{\beta^i} \geq b_t^i; \; \frac{\underline{\mathrm{M}}_i}{\beta^i} \leq b_t^i$$

Recursive Lagrangian Solution

Computational Nightmare: non-convexities in the implementability constraint (Karantounias, 2018) and complex recursive structure that is solved analytically in Aiyagari, Marcet, Sargent, and Seppälä (2002, JPE)

With $\rho = \gamma$ optimality conditions are given by:

$$c_t : u_{c,t} - v_{l,t} + \mu \left[u_{c,t} - v_{l,t} + u_{cc,t} c + v_{ll,t} \left(c_t + g_t \right) \right] + \sum_{i=1}^{N} \left(\mu_{t-i} - \mu_{t-i+1} \right) b^i_{t-i} u_{cc,t} = 0$$

$$b^i_{t+1} : \mu_t = \left[\mathbb{E}_t u_{c,t+i} \right]^{-1} \left[\mathbb{E}_t \mu_{t+1} u_{c,t+i} + \frac{\xi^i_{U,t}}{\beta^i} - \frac{\xi^i_{L,t}}{\beta^i} \right] \quad \forall i$$

$$\mu_t : \sum_{i=1}^{N} b^i_t \mathbb{E}_t \beta^{i-1} \frac{u_{c,t+i-1}}{u_{c,t}} = c_t - \frac{u_{l,t}}{u_{c,t}} \left(c_t + g_t \right) + \sum_{i=1}^{N} b^i_{t+i} \mathbb{E}_t \beta^i \frac{u_{c,t+i}}{u_{c,t}}$$

Information set at every t is $\mathcal{I}_t = \left\{ g_t, \left\{ \left\{ b^i_{t-k} \right\}_{k=0}^{N-1} \right\}_{i=1}^{N}, \left\{ \mu_{t-k} \right\}_{k=1}^{N} \right\}$

Solution - Terms to approximate

To solve the model the terms that need to be approximated are:

$$\mathcal{ANN}_1^i = \mathbb{E}_t u_{c,t+i} \text{ for } i = \{1, \dots, N\}$$

$$\mathcal{ANN}_2^i = \mathbb{E}_t \mu_{t+i} u_{c,t+i} \text{ for } i = \{1, \dots, N\}$$

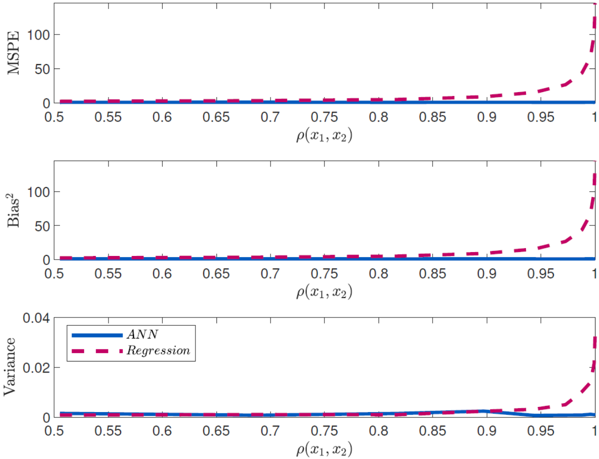
$$\mathcal{ANN}_3^i = \mathbb{E}_t u_{c,t+i-1} \text{ for } i = \{1, \dots, N-1\}$$

Highly multicollinear due to high correlation between contemporaneous variables and the presence of many lags of the same variable. In particular μ_t follows a random walk.

Traditional PEA does not work. Modified Condensed PEA develop by Faraglia, Marcet, Oikonomou, and Scott (2014, WP). PCA, Ridge or Lasso regressions do not converge in this problem.

Solution - Failure of PEA

Take $y_i = 2x_{i1} + 3x_{i2} + x_{i2}^3$ and lets study $MSE_i = \underbrace{\left[y_i - \mathbb{E}\left(\hat{y}_i\right)\right]^2}_{Bias_i^2} + \underbrace{\mathbb{E}\left[\hat{y}_i - \mathbb{E}\left(\hat{y}_i\right)\right]^2}_{Variance_i}$



Solution - Condensed PEA

Objective: Approximate the following function with state variables:

$$\mathbb{E}_t u_{c,t+i} \simeq f_i \left(g_t, \left\{ \left\{ b^i_{t-k} \right\}_{k=0}^{N-1} \right\}_{i=1}^{N}, \{\mu_{t-k}\}_{k=1}^{N} \right) \text{ for } i = \{1, \ldots, N\}$$

$$\mathbb{E}_t \mu_{t+i} u_{c,t+i} \simeq f_{N+i} \left(g_t, \left\{ \left\{ b^i_{t-k} \right\}_{k=0}^{N-1} \right\}_{i=1}^{N}, \{\mu_{t-k}\}_{k=1}^{N} \right) \text{ for } i = \{1, \ldots, N\}$$

Steps of PEA for a given polynomial family, e.g. Hermite:

1. Divide state variables into Core and Non-Core.
2. Parameterize f_i only with Core variables: i) set Maliar and Maliar (2003) bound; ii) simulate dynamics with initial guess; iii) run regression to estimate parameters; and iv) iterate this step until convergence.
3. Regress Non-core on Core and keep residuals (r_1). Then regress expectations on Core ans keep residuals (r_2). Add r_1 and r_2 to the Core set and go back to step 2 until convergence on debt level.
4. Go back to step 1 and randomly change core set to generate MSE reduction.

Solution - ANN-based expectation algorithm 1

1. Given $\mu_{t-1} = 0$, initial weights on $\mathcal{ANN}$, and imposing Maliar moving bounds (2003) with penalty functions to avoid out-bounded solutions. Then:

$$\forall i: \quad \mu_t = \mathcal{ANN}_1^i\left(\mathcal{I}_t\right)^{-1}\left[\mathcal{ANN}_2^i\left(\mathcal{I}_t\right) + \frac{\xi_{U,t}^i}{\beta^i} - \frac{\xi_{L,t}^i}{\beta^i}\right]$$

Now μ_t is over-identified but can be solved using Forward States approach as described in Faraglia et al. (2014), i.e. use relevant state variables at $t+1$, invoke *LIE*, and solve

$$\forall i: \quad \mu_t = \left[\mathbb{E}_t \mathcal{ANN}_1^i\left(\mathcal{I}_{t+1}\right)\right]^{-1}\left[\left[\mathbb{E}_t \mathcal{ANN}_2^i\left(\mathcal{I}_{t+1}\right)\right] + \frac{\xi_{U,t}^i}{\beta^i} - \frac{\xi_{L,t}^i}{\beta^i}\right]$$

Solution - ANN-based expectation algorithm 2

2. Choose T enough and find $\{c_t\}$ and $\left\{b^i_{t+1}\right\}$ that solve:

$$u_{c,t} - v_{l,t} + \mu \left[u_{c,\,t} - v_{l,t} + u_{cc,t} c + v_{ll,t} \left(c_t + g_t \right) \right] + \sum_{i=1}^{N} \left(\mu_{t-i} - \mu_{t-i+1} \right) b^i_{t-i} u_{cc,t} = 0$$

$$\sum_{i=1}^{N} b^i_t \mathbb{E}_t \beta^{i-1} \frac{u_{c,t+i-1}}{u_{c,t}} = c_t - \frac{u_{l,t}}{u_{c,t}} \left(c_t + g_t \right) + \sum_{i=1}^{N} b^i_{t+i} \mathbb{E}_t \beta^i \frac{u_{c,t+i}}{u_{c,t}}$$

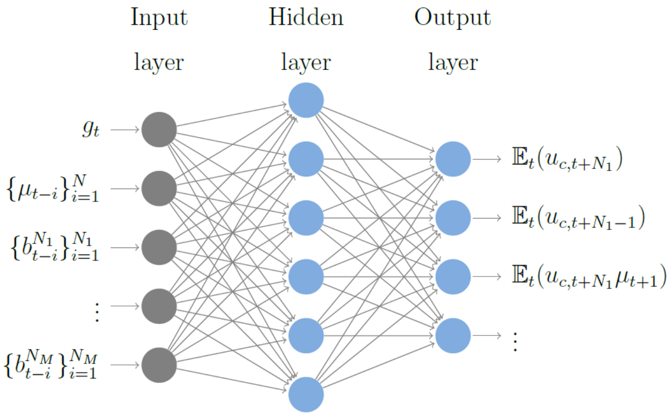
3. If solution error large or no reliable solution, do not actualize and proceed with previous $\mathcal{ANN}$ with a reduced Maliar bound

4. If solution by shrinking bound is not satisfactory keep shrinking bounds keep shrinking bound from iteration $i-1$ to bound from iteration $i-2$ with same $\mathcal{ANN}$

5. If solution is satisfactory after i iterations, the $\mathcal{ANN}$ enters the learning phase supervised by model dynamics. Maliar bounds are increase and the process starts agains until i) prediction error below threshold, and ii) simulated sequences of $\left\{b^i_t\right\}$ and $\{c_t\}$ do not change.

Solution - Graphical Representation ANN

One hidden layers, 10 Neurons and hyperbolic tangent sigmoid as activation function $f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$



Solution - ANN vs Condensed PEA one maturity

One bond with 10 year maturity

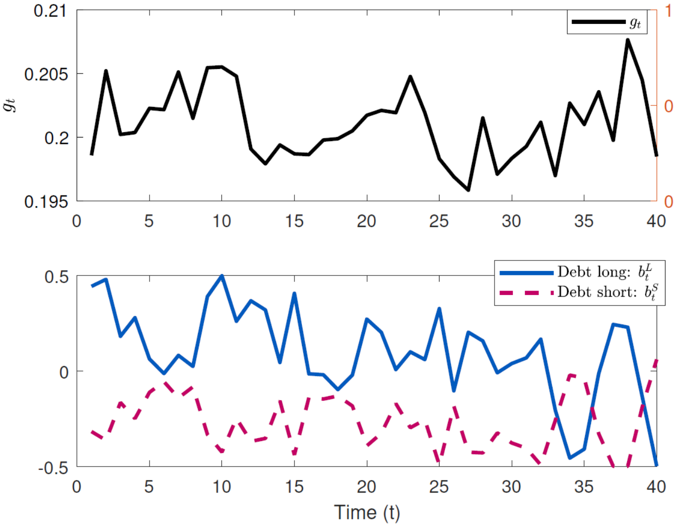
	ANN		C. PEA	
Projected term	*Residual*	$Residual_{\%}$	*Residual*	$Residual_{\%}$
$\mathbb{E}_t(u_{c,t+N}\mu_{t+1})$	0.02	2.00%	0.01	1.25%
$\mathbb{E}_t(u_{c,t+N})$	0.08	1.41%	0.07	1.16%
$\mathbb{E}_t(u_{c,t+N-1})$	0.08	1.42%	0.07	1.18%
Time	69s		695s	

Solution - ANN vs Condensed PEA two bonds

One short term bond with maturity 1 and one long term bond with maturity 10, and an AR(1) process for g_t with mean 0.2 and persistence parameter 0.8

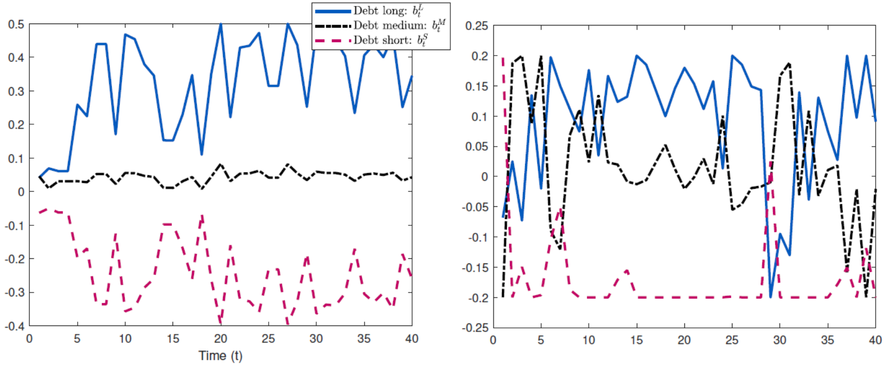
	ANN		C. PEA	
Projected term	*Residual*	$Residual_{\%}$	*Residual*	$Residual_{\%}$
$\mathbb{E}_t(u_{c,t+N}\mu_{t+1})$	0.019	0.32%	0.005	0.45%
$\mathbb{E}_t(u_{c,t+N})$	0.026	0.45%	0.021	0.36%
$\mathbb{E}_t(u_{c,t+N-1})$	0.0063	0.59%	0.021	0.36%
$\mathbb{E}_t(u_{c,t+S})$	0.026	0.44%	0.023	0.40%
$\mathbb{E}_t(u_{c,t+S}\mu_{t+1})$	0.0062	0.58%	0.006	0.60%
Time	20min		450min	

Solution - ANN vs Condensed PEA two bonds



Solution - ANN three bonds

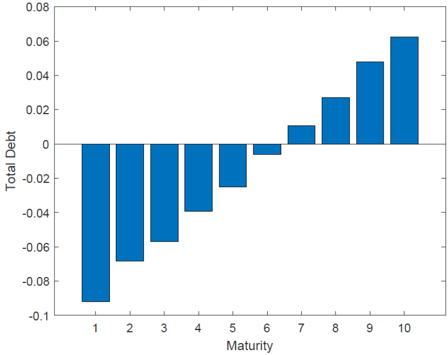
Maturities: 1, 5, and 10. Bounds: 0.5 and 0.2



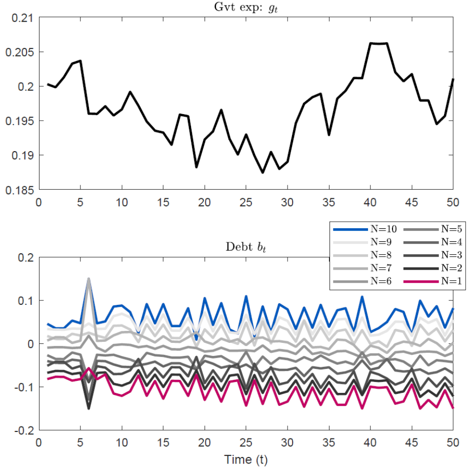
Short in Long Term and Long in Short Term! Long term bond is more responsive to shocks. This combination decreases value of liabilities when g_t is high. Use less correlated instruments.

Solution - ANN ten bonds

Average outstanding debt at each maturity



Simulation



Solution - ANN vs Condensed PEA computational complexity