Solving Models of Economic Dynamics with Ridgeless Kernel Regressions

Mahdi Ebrahimi Kahou[1] Jesse Perla[2] Geoff Pleiss[3,4]

ASSA 2026

[1]Bowdoin College, Econ Dept

[2]University of British Columbia, Vancouver School of Economics

[3]University of British Columbia, Stats Dept

[4]Vector Institute

Motivation

- Numerical solutions to dynamical systems are central to many quantitative fields in economics.
- Dynamical systems in economics are **boundary value** problems:
  1. The boundary is at **infinity**.
  2. The values at the boundary are potentially **unknown**.
- Resulting from **forward looking** behavior of agents.
- Examples include the transversality and the no-bubble condition.
- Without them, the problems are **ill-posed** and have infinitely many solutions:
  - These forward-looking boundary conditions are a key limitation on increasing dimensionality.

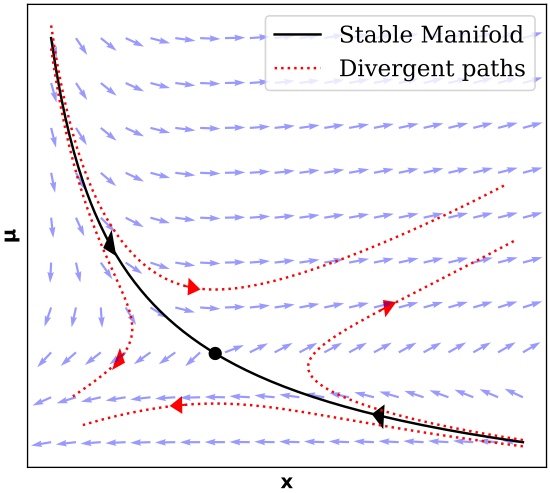
Contribution

Using kernel method to solve a broad class of infinite-horizon, deterministic, continuous-time model

1. **Minimum-norm alignment**:
   - The minimum-norm kernel method aligns with asymptotic boundary conditions.
2. **Learning the right set of steady-states**:
   - Kernel machines learn the boundary values, thereby extrapolating outside the training data.
3. **Robustness and speed**:
   - Competitive in speed and more stable than traditional methods.
4. **Consistency of ML estimates**.

Intuition

- **Violation of the boundary conditions**:
  - Sub-optimal solutions explode over time.
  - They have large derivatives.
  - This behavior is due to the **saddle-path** nature of the problem.
- **Minimum-norm solution** :
  - Penalizing large derivatives rules out explosive paths.
  - The remaining solution is the optimal solution.



The Problem

The class of problems

A differential-algebraic system of equations, coming from an economic optimization problem:

$$\dot{\mathbf{x}}(t) = \mathbf{F}(\mathbf{x}(t), \boldsymbol{\mu}(t), \mathbf{y}(t)) \tag{1}$$

$$\dot{\boldsymbol{\mu}}(t) = r\boldsymbol{\mu}(t) - \boldsymbol{\mu}(t) \odot \mathbf{G}(\mathbf{x}(t), \boldsymbol{\mu}(t), \mathbf{y}(t)) \tag{2}$$

$$\mathbf{0} = \mathbf{H}(\mathbf{x}(t), \boldsymbol{\mu}(t), \mathbf{y}(t)), \tag{3}$$

boundary conditions (at infinity)

$$\mathbf{0} = \lim_{t\to\infty} e^{-rt}\mathbf{x}(t) \odot \boldsymbol{\mu}(t), \tag{4}$$

initial value $\mathbf{x}(0) = \mathbf{x}_0$.

- $\mathbf{x} \in \mathbb{R}^M$: state variables.
- $\boldsymbol{\mu} \in \mathbb{R}^M$: co-state variables.
- $\mathbf{y} \in \mathbb{R}^P$: jump variables.

Goal: finding an approximation for $\mathbf{x}(t)$ and $\mathbf{y}(t)$.

What is the problem?

- $\mathbf{y}_0$ is unknown.
- The optimal solutions is a **saddle-path**: unstable nature

Method

Method

- Pick a set of points $\mathcal{D} \equiv \{t_1, \cdots, t_N\}$ for some fixed interval $[0, T]$
- Large machine learning models to learn $\hat{\mathbf{x}}(t)$ and $\hat{\mathbf{y}}(t)$

$$\min_{\hat{\mathbf{x}},\hat{\mathbf{y}}} \sum_{t_i \in \mathcal{D}} \Bigg[\eta_1 \underbrace{\left\| \dot{\hat{\mathbf{x}}}(t_i) - \mathbf{F}(\hat{\mathbf{x}}(t_i), \hat{\mathbf{y}}(t_i)(t_i)) \right\|_2^2}_{\text{Residuals}^2\text{: state variables}} + \eta_2 \underbrace{\left\| \dot{\hat{\mathbf{y}}}(t_i) - \mathbf{G}(\hat{\mathbf{x}}(t_i), \hat{\mathbf{y}}(t_i)) \right\|_2^2}_{\text{Residuals}^2\text{: jump variables}}$$

$$+ \eta_3 \underbrace{\left\| \mathbf{H}(\hat{\mathbf{x}}(t_i), \hat{\mathbf{y}}(t_i)) \right\|_2^2}_{\text{Residuals}^2\text{: algebraic constraint}} \Bigg] + \eta_4 \underbrace{\left\| \hat{\mathbf{x}}(0) - \mathbf{x}_0 \right\|_2^2}_{\text{Residuals}^2\text{: initial conditions}} .$$

- This optimization **ignores** the boundary conditions.
- The implicit bias automatically satisfy the boundary conditions.
- Recent works suggest the implicit bias is toward smallest Sobolev semi-norms.

Ridgeless kernel regression

$$\hat{\dot{\mathbf{x}}}(t) = \sum_{j=1}^{N} \boldsymbol{\alpha}_j^x K(t, t_j), \quad \hat{\dot{\mathbf{y}}}(t) = \sum_{j=1}^{N} \boldsymbol{\alpha}_j^y K(t, t_j)$$

$$\hat{\mathbf{x}}(t) = \mathbf{x}_0 + \int_0^t \hat{\dot{\mathbf{x}}}(\tau) d\tau, \quad \hat{\mathbf{y}}(t) = \hat{\mathbf{y}}_0 + \int_0^t \hat{\dot{\mathbf{y}}}(\tau) d\tau$$

- $\mathbf{x}_0$ is given.
- $\hat{\mathbf{y}}_0$, $\boldsymbol{\alpha}_j^x$, and $\boldsymbol{\alpha}_j^y$ are learnable parameters.
- $K(\cdot,\cdot)$: Matérn Kernel, with smoothness parameter ν and length scale ℓ.

Ridgeless kernel regression: minimum Sobolev seminorm solutions

We also solve the ridgeless kernel regression

$$\lim_{\lambda \to 0} \min_{\hat{\mathbf{x}}, \hat{\mathbf{y}}} \sum_{t_i \in \mathcal{D}} \left[\eta_1 \left\| \dot{\hat{\mathbf{x}}}(t_i) - \mathbf{F}(\hat{\mathbf{x}}(t_i), \hat{\mathbf{y}}(t_i)(t_i)) \right\|_2^2 + \eta_2 \left\| \dot{\hat{\mathbf{y}}}(t_i) - \mathbf{G}(\hat{\mathbf{x}}(t_i), \hat{\mathbf{y}}(t_i)) \right\|_2^2 \right.$$
$$\left. + \eta_3 \left\| \mathbf{H}(\hat{\mathbf{x}}(t_i), \hat{\mathbf{y}}(t_i)) \right\|_2^2 \right] + \eta_4 \left\| \hat{\mathbf{x}}(0) - \hat{\mathbf{x}}_0 \right\|_2^2 + \lambda \underbrace{\left[\sum_{m=1}^{N_x} \| \hat{\mathbf{x}}^{(m)} \|_{\mathcal{H}}^2 + \sum_{m=1}^{N_y} \| \hat{\mathbf{y}}^{(m)} \|_{\mathcal{H}}^2 \right]}_{\text{The Sobolev semi-norm}}$$

- Targeting Sobolev semi-norm.
- This choice is very natural: it solves the instability issues of the classical algorithm.

Applications

Linear asset pricing

$$
\begin{aligned}
\dot{\mathbf{x}}(t) &= c + g\mathbf{x}(t) && (5) \\
\dot{\mathbf{y}}(t) &= r\mathbf{y}(t) - \mathbf{x}(t) && (6) \\
0 &= \lim_{t \to \infty} e^{-rt}\mathbf{y}(t) && (7)
\end{aligned}
$$

- $\mathbf{x}(t) \in \mathbb{R}$: dividends, $\mathbf{y}(t) \in \mathbb{R}$: prices, and $\mathbf{x}_0$ given.
- Equation (5): how the dividends evolve in time.
- Equation (6): how the prices evolve in time.
- Equation (7): "no-bubble" condition, the boundary condition at infinity.

Why do we need the boundary condition?

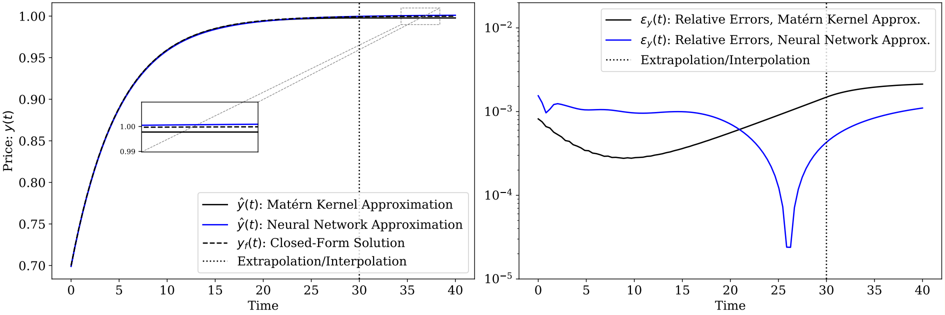
$$\dot{\mathbf{x}}(t) = c + g\mathbf{x}(t)$$
$$\dot{\mathbf{y}}(t) = r\mathbf{y}(t) - \mathbf{x}(t)$$

- The solutions:

$$\mathbf{y}(t) = \mathbf{y}_f(t) + \zeta e^{rt}$$

- $\mathbf{y}_f(t) = \int_0^\infty e^{-r\tau}\mathbf{x}(t+s)ds$: price based on the fundamentals.
- ζe^{rt}: explosive bubble terms, it has to be **ruled out** by the boundary condition.
- Triangle inequality: $\|\mathbf{y}_f\| < \|\mathbf{y}\|$.
- The price based on the fundamentals has the **lowest norm**.

Results



$$\mathcal{D} = \{0, 1, \cdots, 30\}$$

- The explosive solutions are ruled out without directly imposing the boundary condition.
- Very accurate approximations, both in the short- and medium-run.
- Learns the steady-state.

Neoclassical growth model: the agent's problem

$$\max_{\mathbf{y}(t)} \int_0^\infty e^{-rt} \ln(\mathbf{y}(t))dt$$
$$\text{s.t.} \quad \dot{\mathbf{x}}(t) = f(\mathbf{x}(t)) - \mathbf{y}(t) - \delta\mathbf{x}(t)$$

for a given $\mathbf{x}_0$.

- $\mathbf{x}(t) \in \mathbb{R}$: capital, $\mathbf{y}(t) \in \mathbb{R}$: consumption, and a concave production function $f(x) = x^a$.

Constructing the Hamiltonian ...

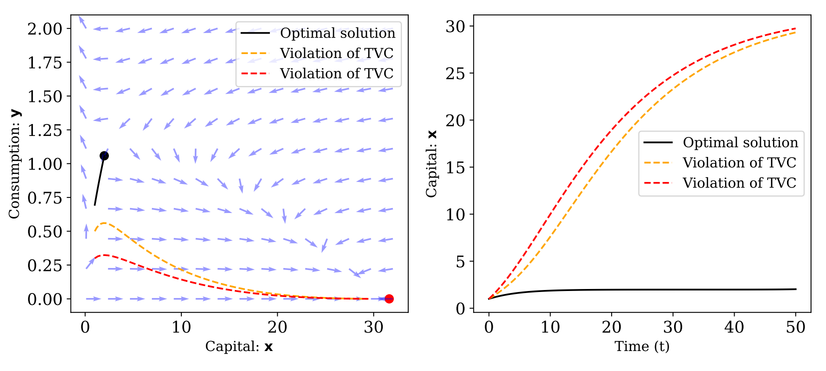
$$\dot{\mathbf{x}}(t) = f(\mathbf{x}(t)) - \mathbf{y}(t) - \delta\mathbf{x}(t) \tag{8}$$
$$\dot{\mathbf{y}}(t) = \mathbf{y}(t)\left[f'(\mathbf{x}(t)) - \delta - r\right] \tag{9}$$
$$0 = \lim_{t\to\infty} e^{-rt}\frac{\mathbf{x}(t)}{\mathbf{y}(t)} \tag{10}$$

- Equation (10) : transversality condition (TVC)

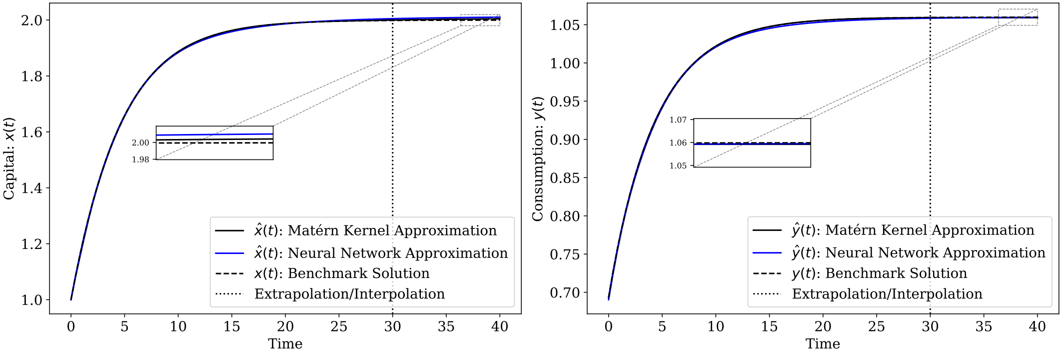
Why do we need the boundary condition?

Ignoring the transversality condition:

$$\dot{\mathbf{x}}(t) = f(\mathbf{x}(t)) - \mathbf{y}(t) - \delta\mathbf{x}(t)$$
$$\dot{\mathbf{y}}(t) = \mathbf{y}(t)\left[f'(\mathbf{x}(t)) - \delta - r\right]$$
$$\mathbf{x}(0) = \mathbf{x}_0 \text{ given.}$$



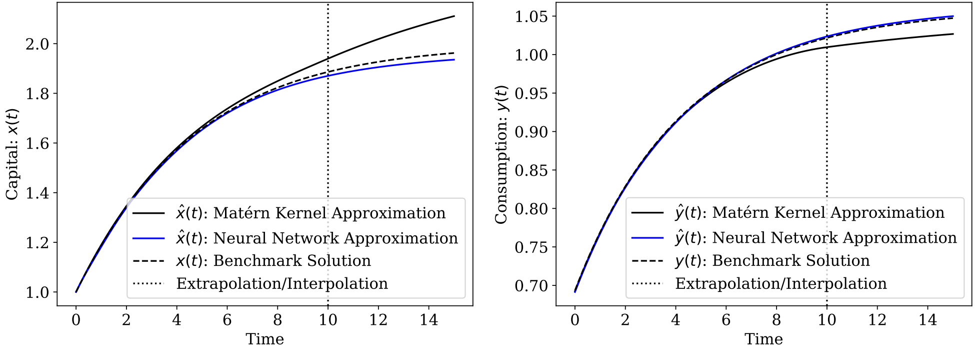
Results



$$\mathcal{D} = \{0, 1, \cdots, 30\}$$

- The explosive solutions are ruled out without directly imposing the boundary condition.
- Very accurate approximations, both in the short- and medium-run.
- Learns the **right steady-state**. ▸▸ Relative errors

Short-run planning: “In the long run, we are all dead”



$$\mathcal{D} = \{0, 1, \cdots, 10\}$$

- The explosive solutions are ruled out without directly imposing the boundary condition.
- Provides a very accurate approximation in the short-run.

Extensions

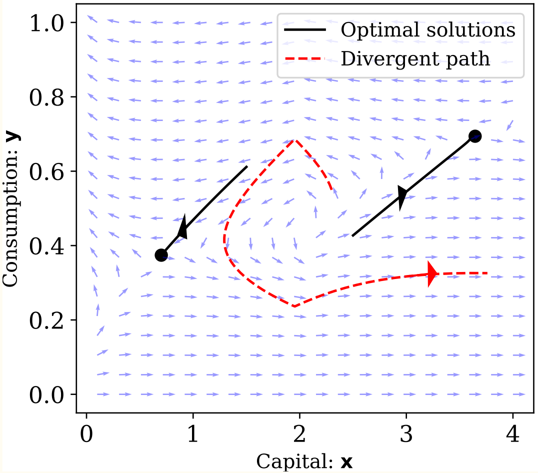
Neoclassical Growth Model: Non-Concave Production Function

- So far we have had a **unique** saddle-path converging to a unique **saddle** steady state.
- What if we have **two** saddle steady states, very close to each other (equilibrium multiplicity)?
- Neoclassical growth model with a non-concave production function (threshold externalities):

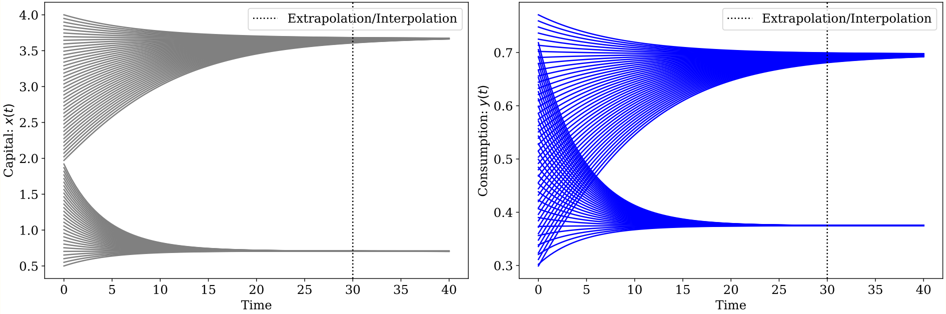
$$f(x) = A \max\{x^a, b_1 x^a - b_2\}$$

Non-concave production function: vector field

$$\dot{\mathbf{x}}(t) = f(\mathbf{x}(t)) - \mathbf{y}(t) - \delta\mathbf{x}(t)$$
$$\dot{\mathbf{y}}(t) = \mathbf{y}(t)\left[f'(\mathbf{x}(t)) - \delta - r\right]$$
$$\mathbf{x}(0) = \mathbf{x}_0 \text{ given.}$$



Results



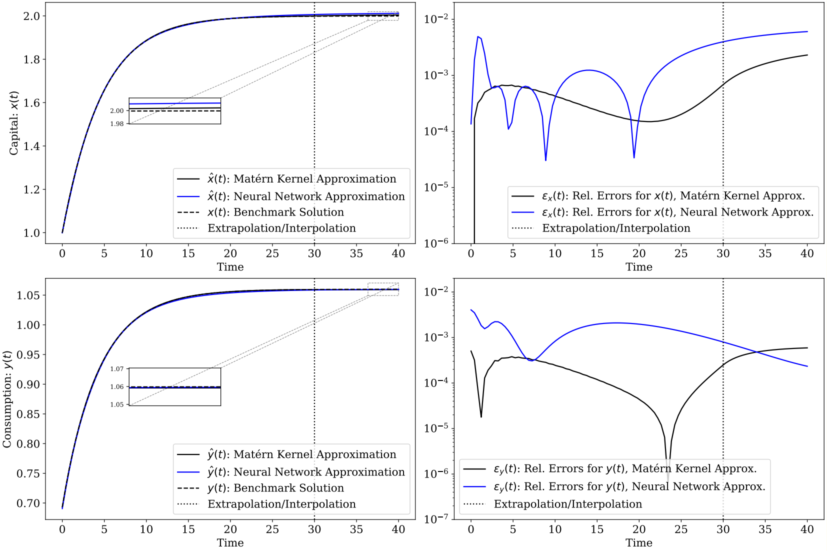
- The approximate solutions approach the right steady states.
- The transversality conditions are satisfied without being directly imposed.
- The steady states are learned. Full DAE

Conclusion

- Long-run (**global**) conditions can be replaced with appropriate regularization (**local**) to achieve the optimal solutions.
- The minimum-norm implicit bias of large ML models aligns with optimality in economic dynamic models.
- Both kernel and neural network approximations accurately learn the right steady state(s).
- Proceeding with **caution**: can regularization be thought of as an equilibrium selection device?

Appendix

Neoclassical growth: relative errors



Back

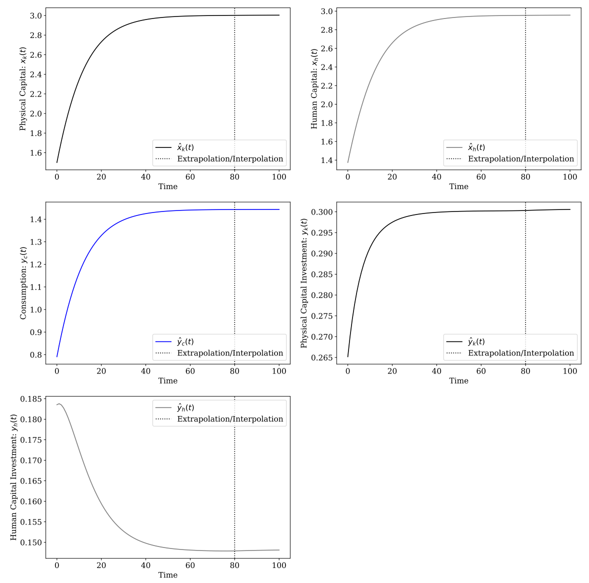
Human capital and growth

$$
\begin{aligned}
\dot{\mathbf{x}}_k(t) &= \mathbf{y}_k(t) - \delta_k \mathbf{x}_k(t), \\
\dot{\mathbf{x}}_h(t) &= \mathbf{y}_h(t) - \delta_k \mathbf{x}_h(t) \\
\dot{\mathbf{y}}_c(t) &= \mathbf{y}_c(t) \left[f_1 \left(\mathbf{x}_k(t), \mathbf{x}_h(t) \right) - \delta_k - r \right], \\
0 &= f \left(\mathbf{x}_k(t), \mathbf{x}_h(t) \right) - \mathbf{y}_c(t) - \mathbf{y}_k(t) - \mathbf{y}_h(t), \\
0 &= f_2 \left(\mathbf{x}_k(t), \mathbf{x}_h(t) \right) - f_1 \left(\mathbf{x}_k(t), \mathbf{x}_h(t) \right) + \delta_k - \delta_h. \\
0 &= \lim_{t \to \infty} e^{-rt} \frac{\mathbf{x}_k(t)}{\mathbf{y}_c(t)}, \quad 0 = \lim_{t \to \infty} e^{-rt} \frac{\mathbf{x}_h(t)}{\mathbf{y}_c(t)}.
\end{aligned}
$$

- $\mathbf{x}_k$: physical capital, $\mathbf{x}_h$: human capital, $\mathbf{y}_c$: consumption, $\mathbf{y}_k$: investment in physical capital, $\mathbf{y}_h$: investment in human capital
- $f\left(x_k, x_h\right) = x_k^{a_k} x_h^{a_h}$

Back

Results


Physical Capital: $x_k(t)$
$\hat{x}_k(t)$
Extrapolation/Interpolation
Time
Human Capital: $x_h(t)$
$\hat{x}_h(t)$
Extrapolation/Interpolation
Time
Consumption: $y_c(t)$
$\hat{y}_c(t)$
Extrapolation/Interpolation
Time
Physical Capital Investment: $y_k(t)$
$\hat{y}_k(t)$
Extrapolation/Interpolation
Time
Human Capital Investment: $y_h(t)$
$\hat{y}_h(t)$
Extrapolation/Interpolation
Time