Solving High-Dimensional Asset Pricing Models via Newton–Kantorovich Iteration

Chase Coleman[a], Pablo Levi[b], John Stachurski[c], Ole Wilms[d] and Junnan Zhang[e]

[a] Research School of Economics, Australian National University
[b] Other affiliations to be added

May 4, 2023

ABSTRACT. To be written

JEL Classifications: D81, G11
Keywords: Asset pricing, wealth-consumption ratio, automatic differentiation

1. INTRODUCTION

As with many sub-fields in economics and finance, researchers working on asset pricing face the need to handle a growing list of state variables when trying to match existing theory to the data. For example, while the seminal paper of Bansal and Yaron (2004) used just two state variables to model the wealth-consumption ratio, subsequent work has added many features to the consumption side of this model. For example, Schorfheide et al. (2018) added preference shocks and an additional stochastic volatility term, leading to four state variables. Gomez-Cram and Yaron (2021) added two more state variables to handle expectations of inflation.

To date, the standard method for handling these kinds of models has been log-linearization. However, the underlying models are highly nonlinear and, moreover, it has been shown that these nonlinearities matter for endogenous quantities

To be added

of interest, such as the risk premium (see, e.g., Pohl et al. (2018)). Hence researchers are seeking better global solution methods that capture the full impact of underlying nonlinearities.

At the same time, for a model of this type, even a moderate number of state variables leads to challenging computational problems. One reason is the nonlinearities discussed above. Another is that, in standard calibrations, discount factors are very close to unity (e.g., 0.998 in Bansal and Yaron (2004)), implying slow convergence of some solution methods. Moreover, the consumption problem is typically embedded in a larger model with still more state variables, so fast and accurate solutions are essential.

In this paper we show how to rapidly approximate global solutions to the kinds of asset pricing models listed above using Newton–Kantorovich iteration backed by three features: automatic differentiation, just-in-time compilation and execution on hardware accelerators such as GPUs.

To implement these feastures we use the software library JAX.

Short history of JAX, discussion of its capabilities.

Some explanation of what we achieve. Review of related literature.

Some benefits of our method.

- We can potentially differentiate the fixed point with respect to the parameters. These gradients help us how the SDF and WC ratio respond to shifts in underlying parameters. They can also provide Jacobians for gradient decent.
- Autodiff means it's possible to plug other fixed point operators, associated with other specifications of recursive utility, directly into the code. There's no need to compute gradients in each case.
- By switching from successive approximation to Newton's method, using autodifferentiation to compute the Jacobian, we change the problem from many small iterations to a small number of computationally expensive ones, which offer better opportunities for parallelization.
- Autodifferentiation is important here, as compared to numerical derivatives, since convergence is fragile for this operator (very slow rate of convergence).

- JAX allows us to invert the Jacobian without actually instantiating the full Jacobian matrix. This is a key feature for successfully solving high dimensional problems.

2. Asset Pricing Background

In this section we introduce the model and then progress to stating the functional equation for the wealth-consumption ration.

In discrete-time no-arbitrage environments, the equilibrium price process $\{P_t\}_{t\geqslant 0}$ associated with a cash flow $\{G_t\}_{t\geqslant 1}$ obeys the fundamental recursion

$$P_t = \mathbb{E}_t \, M_{t+1}(P_{t+1} + G_{t+1}) \quad \text{for all } t \geqslant 0, \tag{1}$$

where $\{M_t\}$ is the sequence of single period stochastic discount factors (see, for example, Kreps (1981), Hansen and Richard (1987) or Duffie (2001)). Most of the well-known "puzzles" in asset pricing theory relate to the difficulty of matching (1) to the data across a diverse range of asset classes. Attempts to resolve these puzzles typically involve relatively sophisticated models for the stochastic discount factor (SDF) process $\{M_t\}$.

In the model, the wealth-consumption ratio obeys

$$\beta^\theta \mathbb{E}_t \left[\left(\frac{\lambda_{t+1}}{\lambda_t} \right)^\theta \left(\frac{C_{t+1}}{C_t} \right)^{1-\gamma} \left(\frac{w(X_{t+1})}{w(X_t) - 1} \right)^\theta \right] = 1,$$

where $\{X_t\}_{t\geqslant 0}$ is a stationary time-homogeneous Markov process on $\mathsf{X} \subset \mathbb{R}^d$. Rearranging the previous expression gives

$$\begin{aligned} (w(X_t) - 1)^\theta &= \beta^\theta \mathbb{E}_t \left[\left(\frac{\lambda_{t+1}}{\lambda_t} \right)^\theta \left(\frac{C_{t+1}}{C_t} \right)^{1-\gamma} w(X_{t+1})^\theta \right] \\ &= \beta^\theta \mathbb{E}_t \left[\exp\left\{ \theta g_{\lambda,t+1} + (1-\gamma) g_{c,t+1} \right\} w(X_{t+1})^\theta \right], \end{aligned}$$

where

$$g_{c,t+1} := \ln \frac{C_{t+1}}{C_t} \quad \text{and} \quad g_{\lambda,t+1} := \ln \frac{\lambda_{t+1}}{\lambda_t}. \tag{2}$$

Let $\mathbb{H}$ be the linear operator defined by

$$(\mathbb{H} f)(x) = \mathbb{E}_x \, f(X_{t+1}) \, \exp\left\{ \theta g_{\lambda,t+1} + (1-\gamma) g_{c,t+1} \right\} \tag{3}$$

at each $x \in \mathsf{X}$, where $\mathbb{E}_x$ conditions on $X_t = x$. With this notation we can now write the equation for the wealth-consumption ratio as

$$(w(x) - 1)^\theta = \beta^\theta \left[(\mathbb{H} w)(x)^\theta \right]. \tag{4}$$

Rearranging once more gives the functional equation

$$w = 1 + \beta(\mathbb{H} w^\theta)^{1/\theta}. \tag{5}$$

A function w solves this equation if and only if w is a fixed point of the operator $\mathbb{T}$ defined by

$$\mathbb{T} w = 1 + \beta \, (\mathbb{H} w^\theta)^{1/\theta}. \tag{6}$$

2.1. **Example: The SSY Case.** In the long run risk model of Schorfheide et al. (2018), the state process takes the form

$$X_t := (h_{\lambda,t}, h_{c,t}, h_{z,t}, z_t) \tag{7}$$

where, for $i \in \{z, c, \lambda\}$,

$$\begin{aligned} h_{i,t+1} &= \rho_i \, h_{i,t} + s_i \, \eta_{i,t+1} \\ \sigma_{i,t} &= \varphi_i \, \bar{\sigma} \, \exp(h_{i,t}), \\ z_{t+1} &= \rho \, z_t + (1 - \rho^2)^{1/2} \, \sigma_{z,t} \, \varepsilon_{t+1} \end{aligned}$$

The consumption and preference shock growth rates are

$$g_{c,t+1} = \mu_c + z_t + \sigma_{c,t} \, \xi_{c,t+1} \quad \text{and} \quad g_{\lambda,t+1} = h_{\lambda,t+1}. \tag{8}$$

The innovations are all independent and standard normal.

In this setting, the operator $\mathbb{H}$ takes the form

$$(\mathbb{H} f)(x) = \mathbb{E}_x \, f(X_{t+1}) \exp \left\{ \theta h_{\lambda,t+1} + (1-\gamma)(\mu_c + z_t + \sigma_{c,t} \xi_{c,t+1}) \right\}.$$

We can reduce dimensionality in the conditional expectation by integrating out the independent innovation $\xi_{c,t+1}$, which leads to

$$(\mathbb{H} f)(x) = \exp \left\{ (1-\gamma)(\mu_c + z) + \frac{1}{2}(1-\gamma)^2 \sigma_c^2 \right\} \mathbb{E}_x \, f(X_{t+1}) \exp \left\{ \theta h_{\lambda,t+1} \right\}, \tag{9}$$

where the conditioning is on $X_t = x$ as given in (7). We can write the operator more explicitly as

$$(\mathbb{H} f)(x) = \kappa(z, h_c) \mathbb{E}_x \, f(X_{t+1}) \exp \left\{ \theta h_{\lambda,t+1} \right\}, \tag{10}$$

where

$$\kappa(z, h_c) := \exp\left\{(1-\gamma)(\mu_c + z) + \frac{1}{2}(1-\gamma)^2[\varphi_c\, \bar{\sigma}\, \exp(h_c)]^2\right\}. \tag{11}$$

The baseline parameter values can be found in the code repository.

2.2. **Example: The GCY Case.** In this section, we analyze the stability properties of the model of Gomez-Cram and Yaron (2020). The authors add inflation dynamics to a long-run risk model similar to the one of Schorfheide et al. (2018). In particular, they assume that the expected inflation rate $z_{\pi,t}$ affects the mean growth rate of consumption:

$$g_{c,t+1} = \ln\left(\frac{C_{t+1}}{C_t}\right) = \mu_c + z_t + \sigma_{c,t}\, \xi_{c,t+1}, \tag{12}$$

where

$$\begin{aligned} z_{t+1} &= \rho\, z_t + \rho_\pi\, z_{\pi,t} + \sigma_{z,t}\, \eta_{t+1} \\ z_{\pi,t+1} &= \rho_{\pi\pi}\, z_{\pi,t} + \sigma_{z\pi,t}\, \eta_{\pi,t+1} \end{aligned}$$

and

$$\sigma_{i,t} = \varphi_i\, \bar{\sigma}\, \exp(h_{i,t}) \quad \text{where} \quad h_{i,t+1} = \rho_i\, h_{i,t} + s_i\, \eta_{i,t+1} \quad \text{for } i \in \{z, c, z\pi\}.$$

Note that also the expected inflation rate $z_{\pi,t}$ has stochastic volatility $\sigma_{z\pi,t}$. As in the model of Schorfheide et al. (2018), the process $\{\lambda_t\}$ follows (8) and all shocks are IID and standard normal. Hence, the state vector x contains 6 states and is given by

$$x = (z, z_\pi, h_z, h_c, h_{z\pi}, h_\lambda) \in \mathsf{X} := \mathbb{R}^6.$$

We can apply the same conditioning as in the model of Schorfheide et al. (2018) and we need to compute $\mathcal{S}_c$ numerically again. The baseline parameter values can be found in the code repository.

3. Solution Method

Background on Newton's algorithm in $\mathbb{R}^N$. Let f be a smooth map from $\mathbb{R}^N$ to itself. We want to find the $x \in \mathbb{R}^N$ that solves $f(x) = x$. Ordinary successive approximation uses

$$x_{k+1} = f(x_k) \tag{13}$$

Newton's method first sets $g(x) = f(x) - x$, so that we are seeking a root x satisfying $g(x) = 0$, and then iterates on

$$x_{k+1} = g(x_k) + J(x_k)^{-1} g(x_k) \tag{14}$$

where $J(x)$ is the Jacobian of g at x.

3.1. **Discretization.** For computation, the state process is distcretized onto a state space $S = \{x_1, \ldots, x_N\}$ of size $N \in \mathbb{N}$ and the operator $\mathbb{H}$ is represented by a matrix $\mathbf{H}$, with

$$\mathbf{H}(n, n') = \sum_{n'=1}^{N} \exp\left\{\theta g_\lambda(x_n, x_{n'}, \xi) + (1-\gamma) g_c(x_n, x_{n'}, \xi)\right\} \mathbf{P}(n, n'), \tag{15}$$

where $\mathbf{P}$ is an $N \times N$ matrix with $\mathbf{P}(n, n')$ representing the probability that the discretized state process transitions from state x_n to state $x_{n'}$ in one unit of time.

The discretization of $\mathbb{T}$ is written as $\mathbf{T}$ and the problem is to find the fixed point of

$$(\mathbf{T}w) = 1 + \beta \left(\mathbf{H} w^\theta\right)^{1/\theta} \tag{16}$$

in the set of strictly positive vectors in $\mathbb{R}^N$.

References

BANSAL, R. AND A. YARON (2004): "Risks for the long run: A potential resolution of asset pricing puzzles," *The Journal of Finance*, 59, 1481–1509.

DUFFIE, D. (2001): *Dynamic Asset Pricing Theory*, Princeton University Press.

GOMEZ-CRAM, R. AND A. YARON (2020): "How Important Are Inflation Expectations for the Nominal Yield Curve?" *The Review of Financial Studies*, 34, 985–1045.

——— (2021): "How important are inflation expectations for the nominal yield curve?" *The Review of Financial Studies*, 34, 985–1045.

HANSEN, L. P. AND S. F. RICHARD (1987): "The Role of Conditioning Information in Deducing Testable Restrictions Implied by Dynamic Asset Pricing Models," *Econometrica*, 55, 587–613.

KREPS, D. M. (1981): "Arbitrage and equilibrium in economies with infinitely many commodities," *Journal of Mathematical Economics*, 8, 15–35.

POHL, W., K. SCHMEDDERS, AND O. WILMS (2018): "Higher order effects in asset pricing models with long-run risks," *The Journal of Finance*, 73, 1061–1111.

SCHORFHEIDE, F., D. SONG, AND A. YARON (2018): "Identifying long-run risks: A Bayesian mixed-frequency approach," *Econometrica*, 86, 617–654.