Automatic Differentiation Methods for Analyzing Wealth-Consumption Ratios and Stochastic Discount Factors

August 23, 2022

ABSTRACT. To be written

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

1. INTRODUCTION

For now a place to put some notes about ideas.

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

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

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

$$x_{k+1} = g(x_k) + J(x_k)^{-1} g(x_k) \tag{2}$$

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

In the EZ fixed point problem, it's not too hard to compute an expression for the Jacobian of the fixed point operator, so why use autodiff? This is a fair question. Here are some thoughts?

To be added

- We can also 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.
- 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.
- We might be able to use second order Newton methods, since autodiff can provide derivatives of any order.

General points:

- 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).

2. A Long Run Risk Model

Roadmap to be added.

Consumption growth and the growth rate of the preference shock are given by the generic formulas

$$g_{c,t+1} = g_c(X_t, X_{t+1}, \xi_{t+1}) \quad \text{and} \quad g_{\lambda,t+1} = g_\lambda(X_t, X_{t+1}, \xi_{t+1}), \tag{3}$$

where $\{X_t\}_{t \geqslant 0}$ is a discrete time Markov process on $\mathsf{X} \subset \mathbb{R}^d$, $\{\xi_t\}_{t \geqslant 1}$ is an IID process supported on $\mathbb{Y} \subset \mathbb{R}^k$, and $g_i \colon \mathsf{X} \times \mathsf{X} \times \mathbb{Y} \to \mathbb{R}$ is continuous for each $i \in \{c, \lambda\}$. The processes $\{X_t\}$ and $\{\xi_t\}$ are assumed to be independent.

In the model, the wealth-consumption ratio obeys

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

Rearranging the previous expression gives

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

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

$$(\mathbb{H}g)(x) = \mathbb{E}_x \, g(X_{t+1}) \, \exp\left[\theta g_\lambda(X_t, X_{t+1}, \xi_{t+1}) + (1-\gamma) g_c(X_t, X_{t+1}, \xi_{t+1}))\right] \quad (4)$$

at each $x \in \mathsf{X}$, where $\mathbb{E}_x$ conditions on $X_t = x$. Conditioning on $X_t = x$, writing pointwise on X and using the definition of $\mathbb{H}$ yields $w = 1 + \beta \left(\mathbb{H} w^\theta \right)^{1/\theta}$. A function w solves this equation if and only if w is a fixed point of the operator $\mathbb{T}$ defined by

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

2.1. **Discretization.** In the long run risk model of Schorfheide et al. (2018), the state process takes the form

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

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

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

Consumption growth is given by

$$g_{c,t+1} = \ln \frac{C_{t+1}}{C_t} = \mu_c + z_t + \sigma_{c,t} \, \xi_{c,t+1}. \quad (6)$$

The preference shock λ_t grows as

$$g_{\lambda,t+1} = \ln \frac{\lambda_{t+1}}{\lambda_t} = h_{\lambda,t+1}.$$

The innovations

$$\xi_{c,t}, \quad \varepsilon_t, \text{ and } (\eta_{i,t})_{i \in \{z,c,\lambda\}}$$

are all independent and standard normal.

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

$$\mathbf{H}(n, n') = \sum_{n'=1}^{N} \exp\left\{\theta g_\lambda(x_n, x_{n'}, \xi) + (1-\gamma) g_c(x_n, x_{n'}, \xi)\right\} \mathbf{P}(n, n'), \tag{7}$$

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

$$(\mathbf{T}w) = 1 + \beta \left(\mathbf{H} w^\theta\right)^{1/\theta} \tag{8}$$

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

References

Schorfheide, F., D. Song, and A. Yaron (2018): "Identifying long-run risks: A Bayesian mixed-frequency approach," *Econometrica*, 86, 617–654.