

SEHYOUN AHN Norges Bank

BI 2019/11/5

Standard disclaimers apply.



$\textbf{Distribution} \Rightarrow \textbf{Macroeconomy}$

- So far, we have focused on how to solve the optimization problems of individuals.
- For the heterogeneous agent model to be "closed," we need their behaviors to affect the macroeconomy.



$\textbf{Distribution} \Rightarrow \textbf{Macroeconomy}$

This requires computation of the aggregate quantities from individual behaviors.

- For example, how do the total amount of savings change if you change interest rate.
- i.e., we need things like

$$\sum f(x_i) \qquad \qquad i \in \mathsf{population}$$

• where x_i changes consistently with individual behaviors



Computation of the Distribution

• Wish: g(x) a distribution function such that

$$\frac{1}{N}\sum f(x_i) = \int f(x)g(d) \,\mathrm{d}x$$

- \blacksquare We proceed assume that such a $g(\cdot)$ exists, and get the conditions it need to satisfy
- This results in the Fokker-Planck Equation



Computation of the Distribution

For the dynamic decision of an individual, we model it as an Ito process

 $\mathrm{d}x_i = \mu(x_i)\,\mathrm{d}t + \sigma\,\mathrm{d}W_t$

This is the optimal behavior determined by the HJB equation, but at this state.



Since we want to approximate

$$\frac{1}{N}\sum f(x_i)$$

we can in fact, just simulate this directly.

- Hence, we take sample of $\{x_i\}$ of the population.
- At each time step, simulate via

$$x_{i,t+\Delta t} = \mu(x_{i,t})\Delta t + \sigma\varepsilon \qquad \qquad \varepsilon \sim N(0, var = \Delta t)$$

- This is called the Euler-Maruyama scheme.
- As $N \to \infty$, it goes to the correct value. But...

- Size of *N*?
- Burn-in?
- Step-size: Δt



- If one has an experience with the Bayesian estimation, one knows how "annoying" the hyperparameter tuning is.
- Also, based on some experiments, the size of N can be quite large.





 One need quite a lot of simulation households for the computation of the steady-state distribution of the Bewley-Huggett-Aiyagari model.
⁰Preliminary plot: need to be checked further.

9 / 27

- Some of these problems are gone with the partial differential equations formulation.
- Things scale differently between different methods
 - We might have to come back to the Monte Carlo simulations in higher dimensions, but for low dimensional problems, PDE methods work better.



Now, to get the partial differential equation formulation, recall we want

$$\frac{1}{N}\sum f(x_i) = \int f(x)g(x)\,\mathrm{d}x$$



We will consider how g(·) evolves over time as people follow their optimal decisions. Hence, let g_t(x) be the time dependent distribution, i.e.,

$$\frac{1}{N}\sum f(x_{i,t}) = \int f(x)g_t(x) \,\mathrm{d}x$$

• Again, we will do Δt approximation.



$$\frac{1}{N}\sum f(x_{i,t+\Delta t}) = \frac{1}{N}\sum \left[f(x_{i,t}) + f'(x_{i,t})(x_{i,t+\Delta t} - x_{i,t}) + \frac{1}{2}f''(x_{i,t})(x_{i,t+\Delta t} - x_{i,t})^2\right]$$

The detail of $(x_{i,t+\Delta t} - x_{i,t})$ is the same as that with HJB equation, so we get

$$\frac{1}{N}\sum f(x_{i,t+\Delta t}) - \frac{1}{N}\sum f(x_{i,t}) = \Delta t \frac{1}{N}\sum f'(x_{i,t})\mu(x_{i,t}) + \Delta t \frac{1}{N}\sum \frac{1}{2}f''(x_{i,t})\sigma^2$$



• Now, we translate the expression, with all our "wishful" $g_t(\cdot)$

$$\int f(x)g_{t+\Delta t}(x) \, \mathrm{d}x - \int f(x)g_t(x) \, \mathrm{d}x$$
$$= \Delta t \int f'(x)\mu(x)g_t(x) \, \mathrm{d}x + \Delta t \int \frac{1}{2}\sigma^2 f''(x)g_t(x) \, \mathrm{d}x$$



$$\int f'(x)\mu(x)g_t(x) \, \mathrm{d}x = f(x)\mu(x)g_t(x)|_{\mathsf{boundary}}$$
$$-\int f(x)\frac{d}{dx}\left(\mu(x)g_t(x)\right) \, \mathrm{d}x$$

Boundary conditions matter, but we leave it for later, then we have it is zero for now.

$$\int f'(x)\mu(x)g_t(x)\,\mathrm{d}x = -\int f(x)\frac{d}{dx}\left(\mu(x)g_t(x)\right)\,\mathrm{d}x$$



Similar application of the integration-by-parts results in

$$\int f''(x)g_t(x)\,\mathrm{d}x = \int f(x)\frac{\mathrm{d}^2}{\mathrm{d}x^2}g_t(x)\,\mathrm{d}x$$



Collecting everything together, we get

$$\int f(x) \frac{g_{t+\Delta t}(x) - g_t(x)}{\Delta t} = -\int f(x) \frac{\mathrm{d}}{\mathrm{d}x} \left(\mu(x)g(x)\right) \mathrm{d}x$$
$$+ \int f(x) \frac{\sigma^2}{2} \frac{\mathrm{d}^2}{\mathrm{d}x^2} g_t(x) \mathrm{d}x$$

- Note that we never defined what $f(\cdot)$ is.
- \blacksquare This means that $g(\cdot)$ need to satisfy these conditions at all "points" ^1

¹Refer to the calculus of variations for technical details. Also, this is a solution concept that's relevant for the Fokker-Planck equation in fact.



Hence, we have

$$\frac{g_{t+\Delta t}(x) - g_t(x)}{\Delta t} = -\frac{\mathrm{d}}{\mathrm{d}x}(\mu(x)g(x)) + \frac{\sigma^2}{2}\frac{\mathrm{d}^2}{\mathrm{d}x^2}g(x)$$

Hence, in limit

$$\frac{\mathrm{d}g}{\mathrm{d}t} = -\frac{\mathrm{d}}{\mathrm{d}x}(\mu(x)g(x)) + \frac{\sigma^2}{2}\frac{\mathrm{d}^2}{\mathrm{d}x^2}g(x)$$

This is the Fokker-Planck equation



- Again, what did we gain?
- Went from a population of

$$\mathrm{d}x_{i,t} = \mu(x_{i,t})\,\mathrm{d}t + \sigma\,\mathrm{d}W_{i,t}$$

to

$$\frac{\mathrm{d}g}{\mathrm{d}t} = -\frac{\mathrm{d}}{\mathrm{d}x}(\mu(x)g(x)) + \frac{\sigma^2}{2}\frac{\mathrm{d}^2}{\mathrm{d}x^2}g(x)$$



■ Nicer to approximate numerically²



²in low dimension

Discretization Methods

- HJB equations are HARD discretization problems.
- **FPK** equations are nicer, and we have more room for decision making.
- Part of this is because we do not have max... where a* depends on the value function.
- We have a different problems of
 - positivity
 - preservation of mass
- but they are still easier.



• We can do the same finite-difference methods as in the HJB equation case

$$\frac{\mathrm{d}g}{\mathrm{d}t} = 0 = \frac{\mathrm{d}}{\mathrm{d}x}(-\mu(x)g(x)) + \frac{\sigma^2}{2}\frac{\mathrm{d}^2}{\mathrm{d}x^2}g(x)$$

■ i.e.,

$$\left[\frac{\mathrm{d}}{\mathrm{d}x}(-\mu(x)\cdot g(x))\right]_i + \frac{\sigma^2}{2}\left[\frac{\mathrm{d}^2}{\mathrm{d}x^2}g(x)\right]_i = 0$$



For example, if we take forward difference, we get

$$\begin{bmatrix} -\frac{\mu(x_{i+1})}{x_{i+1} - x_i} \cdot g(x_{i+1}) + \frac{\mu(x_i)}{x_{i+1} - x_i} \cdot g(x_i) \end{bmatrix} \\ + \frac{\sigma^2}{2} \begin{bmatrix} 2\frac{g(x_{i-1})}{(x_i - x_{i-1})(x_{i+1} - x_{i-1})} \\ - 2\frac{g(x_i)}{(x_i - x_{i-1})(x_{i+1} - x_i)} \\ + 2\frac{g(x_{i+1})}{(x_{i+1} - x_i)(x_{i+1} - x_{i-1})} \end{bmatrix} = 0$$



This expression looks terrible



- This expression looks terrible
- But... everything is linear in $g(x_i)$
- Hence, we end up with the same matrix form as with HJB equation.

$$Ag = 0$$



For example, the A from above with forward difference is given by

$$A(i, i-1) = \frac{\sigma^2}{(x_i - x_{i-1})(x_{i+1} - x_i)}$$
$$A(i, i) = \frac{\mu(x_i)}{x_{i+1} - x_i} + \frac{\sigma^2}{(x_i - x_{i-1})(x_{i+1} - x_i)}$$
$$A(i, i+1) = -\frac{\mu(x_{i+1})}{x_{i+1} - x_i} + \frac{\sigma^2}{(x_{i+1} - x_i)(x_{i+1} - x_{i-1})}$$

Again, this can be implemented exactly the same as with HJB.



- In fact, if we have use the same grid as that for the HJB equation, A_{FP} is exactly the transform of A_{HJB}.
- My Advisor (Benjamin Moll): "You get distribution for free!"
- This is also good if you use uniformly-spaced grids.
 - Not good for non-uniformly spaced grids, and an adjustment is necessary.³

³Achdou, Yves, et al. Income and wealth distribution in macroeconomics: A continuous-time approach. No. w23732. National Bureau of Economic Research, 2017.

26 / 27

We Need a Cute Cat Picture about Now...



³https://xkcd.com/231/

We Need a Cute Cat Picture about Now...



³https://xkcd.com/231/

