Contiune with the stylized model of Krusell and Smith (1998), where agents have value

$$V(a,e) = \max_{c,a'} \left\{ u(c) + \beta \mathbb{E}\left[V(a',e') | e \right] \right\}$$
$$\text{s.t.} \quad c + a' = we + (1+r)a$$

and $e \in \{0,1\}$ follows a simple, 2×2 transition matrix:

$$\Pi = \begin{pmatrix} \pi_{00} & \pi_{10} \\ \pi_{01} & \pi_{11} \end{pmatrix}, \qquad \pi_{e0} + \pi_{e1} = 1.$$

The natural borrowing constraint is 0, and aggregate production technology is standard neoclassical:

$$F(K,N) = AK^{\alpha}N^{1-\alpha}.$$

Today we consider changes in A:

1. Assume that initially we are in a steady state. The economy is hit by a once and for all shock and transitions to a new steady state.
2. Assume that A is stochastic. Compute the stationary evolution of the aggregate distribution over time.

Both of these are much more time-consuming that just solving for the stationary distribution, and not only because you have to first solve for the stationary distribution(s) anyway. But the first case is perhaps easier to understand than the second.

1. Steady State to Steady State Transitions

In many models we want to know the impact of a big change in the environment, be it technological or policy-related, e.g. the introduction of a different tax policy. Note that in such scenarios,

- Once the shock is realized, everything is deterministic. However, in principle this shock can be anything.
- For a good example, Krueger and Perri (2006) assume that the economy was in a steady state in 1980, and the next year individuals suddenly wake up to new individual income processes that changed over the entire post 1980s period, ending in the mid-1990s.
- **This is still a once and for all shock** since even though the underlying environment is changing every year, the agents learn about the entire future process of changes at the beginning of the change—i.e. *aggregate dynamics is deterministic*.

- The critical assumption is that individuals learn at once how things will change in the future, so according to our RE models agents are able to perfectly forecast the future time path of prices, a.k.a. **perfect foresight**.

- This is because there is *no aggregate uncertainty following the initial aggregate shock*.

Such analyses are interesting for

1. Analyzing how fast we will transition from one steady state to another. This is also related to welfare (transition) costs.

2. Analyzing how the distribution changes along the time path. For instance, suppose it takes 200 years for the transition to be complete. We may only be interested in what happens in the first 20 years, especially if we know there is another big change after 20 years. To know this though, we still have to compute the entire transition.

Guessing a Finite Time Path for Prices The assumption is that we are in a steady state in time 0. At $t = 1$, agents suddenly wake up to a change in environment that they know will last forever (or can perfectly forecast how things will change into the indefinite future). Solving the model is mechanically similar to the stationary case but conceptually different, since we cannot *first* solve for the distribution given prices, *then* check whether prices clear the market.

EXERCISE 1 *1. Compute the stationary distributions for before and after the shock. Call these F_0 and F_∞, and the terminal value function V_∞.*

2. *Assume it takes some arbitrary number of periods T to transition from F_0 to F_∞, usually 200-300 periods. Our hope is that $F_T \approx F_\infty$.*

3. *Guess a sequence of prices $\{R_t\}_{t=2}^{T-1}$. Note that we already know R_1, and impose $R_T = R_\infty$, since that is what we ultimately want.*

4. *Starting from V_T and imposing $V_T = V_\infty$, compute the time sequence of **value functions**, $\{V_{T-1}, V_{T-2}, \ldots, V_1\}$, by backward induction.*

5. *Starting from $t = 1$, compute the implied path of prices $\{r_t\}_{t=2}^{T-1}$. This involves, of course, solving for the time sequence of **distributions**, $\{F_t\}_{t=1}^{T}$.*

6. *Check if $\{r_t\} \approx \{R_t\} - \delta$. If not, update $\{R_t\}$ and repeat from step 3.*

7. *Check if the $F_T \approx F_\infty$. If not, increase T.*

There are three things to note:

1. By now you should be used to iterating on the functional equation V assuming an infinite horizon. Now in step 4, you are *not* solving for convergence by iterating on the value function, but literally get the period-to-period values by backward induction, so each iterated value must be stored.

2. In step 5, you could use either Monte Carlo simulation or forward iteration on an approximated density function (like the p.m.f. in the stationary case). For everyone period, you compute the price to generate a sequence $\{r_t\}_{t=2}^{T}$.

3. How to update a **sequence** of prices is not straightforward. In this simple case, any method will work (e.g. tatonnement or Newton's method on the entire sequence, etc.) but for more complicated models you should be more careful when choosing an updating scheme.

Approximating the Law of Motion Up to now we have focused on guessing/iterating on prices, we could have done the same with the mean level of aggregate capital, K. This is what you will find in many papers and textbooks. This is technically more correct but as long as you only use the mean, equivalent to guessing the prices. This is because the relationship between R and K is fixed by

$$R_t = A\alpha \left(\frac{N}{K_t}\right)^{1-\alpha} - \delta.$$

Note that in the current example, N is fixed by assumption. What is the reason that guessing the mean of K is technically more correct?

1. Period-to-period, agents don't need to know the distribution. They only need to know K (again, since this is sufficient for knowing R).

2. So in the stationary case, R and K are sufficient statistics for each other. However, when the distribution is changing, agents need to know the entire current distribution to forecast *tomorrow*'s (R, K).

3. By recursion, they also need to forecast tomorrow's distribution!

Therefore, the problem faced by the agents can be written

$$\begin{aligned} V(a, e; [m_1, \cdots, m_\infty]) &= \max_{c,a'} \left\{ u(c) + \beta \mathbb{E}\left[V(a', e'; [m_1', \cdots, m_\infty']) | e; [m_1, \cdots, m_\infty] \right] \right\} \\ \text{s.t.} \qquad c + a' &= we + (1 + r(m_1))a \\ [m_1', \cdots, m_\infty'] &= G([m_1, \cdots, m_\infty]), \end{aligned}$$

where $[m_1, \cdots, m_\infty]$ are the moments of the current period distribution F, $m_1 = K$, and r is a function of K. The last line just means that, the agents need to know the evolution of the distribution, i.e. $F' = G(F)$ (rational expectations).

1. Since F is continuous, forecasting a distribution amounts to carrying around a distribution as a state variable.

2. This can be summarized by its infinite number of moments.

3. In practice, we focus only on the first n moments, i.e.

$$V(a, e; [m_1, \ldots, m_n]) = \max_{c,a'} \left\{ u(c) + \beta \mathbb{E} \left[V(a', e'; [m_1', \ldots, m_n']) | e; [m_1, \ldots, m_n] \right] \right\}$$
$$\text{s.t.} \quad c + a' = we + (1 + r(m_1))a.$$

4. We could approximate F differently, e.g. by functional approximation. We won't study this in class though. (Plus honestly, I don't have expertise in it.)

What we have been doing up to now is cutting at $n = 1$. In fact, we will continue to keep $n = 1$ for now, so that it wouldn't matter if we kept focusing only on R. But I will switch to $m_1 = K$ as the state now, for continuity with the next subsection. But why is this a valid approximation?

1. As long as $u(\cdot)$ is CRRA, savings propensities are constant regardless of current period income/wealth...

2. ...if borrowing constraints are not binding. Since aggregate K is mainly determined by agents who are not constrained, the law of motion for K can be well approximated by

$$\log K' = \kappa_0 + \kappa_1 \log K$$

and has been verified by previous papers, most prominently Krusell and Smith (1998).

EXERCISE 2 1. *Compute the stationary distributions for before and after the shock. Call these F_0 and F_∞, the levels of capital (K_0, K_∞), and the terminal value function V_∞.*

2. *Guess (κ_0, κ_1). If the transition is complete at time $T - 1$, we would have*

$$\log K_T = \log K_{T-1} \qquad \textit{and}$$
$$\log K_T = \kappa_0 + \kappa_1 \log K_{T-1}.$$

Hence it makes sense to guess only κ_1 while fixing κ_0 at

$$\kappa_0 = (1 - \kappa_1) \cdot \log K_\infty,$$

while a good initial guess for κ_1 would be 1.

3. *Solve the value function assuming an infinite horizon, including K as a state. The grid for K can be chosen coarsely in an interval that contains $[K_0, K_\infty]$. Individual forecasts for K' are used to project next period price R' (and hence (r', w')). That is,*

$$V(a, e; K) = \max_{c,a'} \left\{ u(c) + \beta \mathbb{E} \left[V(a', e'; K') | e; K \right] \right\}$$
$$\textit{s.t.} \quad c + a' = we + (1 + r(m_1))a$$
$$\log K' = \kappa_0 + \kappa_1 \log K.$$

4. *Starting from $t = 1$ and $K_1 = K_0$, compute the implied path of capital $\{K_t\}_{t=2}^T$. This involves, of course, solving for the time sequence of* ***distributions***, *$\{F_t\}_{t=2}^T$.*

5. *Check if $K_T \approx K_{T-1}$. If not, increase T.*

6. *Regress K_t on K_{t-1} to find the implied $\tilde{\kappa}_1$. If $\kappa_1 \approx \tilde{\kappa}_1$, terminate. Otherwise update and repeat from 3. When updating, you can assume $\kappa_1 \in [0, 1]$.*

Again, the time sequence of distributions in step 4 can be solved for using Monte Carlo simulations or forward iterating on the distribution. In step 3, the law of motion for the individual state a' is determined by individual optimization, while the law of motion for aggregate capital K' is given exogenously by the guess. Few last things to note:

1. As long as you assume that the only information in agents' aggregate state space is the mean, m_1, the algorithm would be equivalent to guessing a law of motion for R.

2. Compared to guessing the entire sequence of K or R in Exercise 1, we are instead guessing the aggregate law of motion. We replace convergence on the entire sequence to convergence on the constants assumed in the law of motion.

3. Having said that, it is not always the case that K and (R, w) are isomorphic, e.g. if there is a labor-leisure choice or occupational choice. In those cases, further approximations would be needed to approximate the relationship between K and L, and/or K and R. Typically these will all take a log-linear form, as long as we maintain the isoelastic utility/technology assumption.

4. This could be interpreted in two ways. The first is the way I have been explaining it, that it is an approximation to the agents' information set. It could also be seen as assuming that agents are not perfectly rational, or that they only have limited information, i.e. "bounded rationality." Note the difference—depending on your model, you will want to motivate it differently.

2. Aggregate Shocks

Up to now, the aggregate state was always deterministic—even when it was changing over time, the future path was assumed to be (nearly) perfectly known to all agents. Now we assume that the aggregate state is also stochastic—for our example we will assume that now A evolves according to the 2×2 transition matrix

$$\Gamma = \begin{pmatrix} p_{bb} & p_{gb} \\ p_{bg} & p_{gg} \end{pmatrix}, \qquad p_{Ab} + p_{Ag} = 1,$$

and $A \in \{g, b\}$, which denotes a good and bad aggregate state, respectively. Further assume that the transition matrix for e is also different in good and bad states, i.e. there are two transition matrices (Π^g, Π^b) with elements $\pi^A_{ee'}$. While we can make whatever restrictions we want, Krusell and Smith (1998) assume that aggregate employment can take on only two values $\{u_g, u_b\}$. So the parameters $\pi^A_{ee'}$ and $p_{AA'}$ are restricted as

$$p_{AA'}\left[\pi^A_{00}u_A + \pi^A_{10}(1-u_A)\right] = u_{A'}$$

for all 4 pairs of (A, A'), where u_A corresponds to the steady state unemployment rate under Π_A. Since the transition matrix is collinear, more restrictions need to be made, which are chosen to match aggregate unemployment fluctuations, individaul unemployment spell durations, etc., in the data.

But note that this **does not** imply that aggregate capital takes on only two values—since capital is determined by endogenous savings, it depends on the entire distribution. Since the A-shock is a perfect predictor of u_A, employment is no longer an aggregate state. An agent's problem is in this environment is

$$\begin{aligned} V(a,e;F,A) &= \max_{c,a'}\left\{u(c) + \beta\mathbb{E}\left[V(a',e';F',A')|e;F,A\right]\right\} \\ \text{s.t.} \qquad c + a' &= we + (1+r(F,A))a \\ F' &= G(F,A) \end{aligned}$$

and the shocks (e, A) follow the assumed Markov processes. Note that since F is a state for the agents, F' is deterministic given today's state—i.e. I not only know my own states and decisions but also that of everyone else (even though I might not know *who* they are). In other words, K' is not affected by A'. Hence the expectation operator is taken over only (e, A). Note also that the aggregate law of motion now depends on the aggregate state, since it is stochastic!

Monte Carlo Simulation (a.k.a. "Krusell and Smith (1998)") This method is similar to the second method in the previous section. As there, we approximate F only by the first moment of aggregate capital $m_1 = K$.

EXERCISE 3 *The following is a simple version of Krusell and Smith (1998), but not so much simpler.*

1. *Compute the stationary distribution F_0 under the assumption the long run mean of A. Use this value function and F_0 as initial guesses.*

2. *Guess $\{\kappa^A_0, \kappa^A_1\}_{A=g,b}$ as in the previous section. We cannot impose any general restrictions on the parameters anymore, but reasonable guesses would be $\kappa_0 \approx 0$ and $\kappa_1 \approx 1$. Note that we need a set of $\{\kappa^A_0, \kappa^A_1\}$'s for each state $A \in \{g, b\}$, so this is a set of two equations.*

3. *Solve the value function assuming an infinite horizon, including K as a state. The grid for K can be chosen coarsely in an interval that contains K_{ss}. In each state A, individual forecasts for K' are projected according to (κ^A_0, κ^A_1), from which values agents compute the next period price R' (and hence (r', w')). The aggregate transition matrix is used to compute the next period probabilities of (g, b).*

4. *Starting from* $K_1 = K_{ss}$*, compute the implied path of capital* $\{K_t\}_{t=2}^T$ *for N individuals; say* $N = 100,000$ *and* $T = 5500$*. Throw out the first 500 periods.*

5. *Regress* K' *on* K *for the* $T - 500$ *periods to find the implied* $(\tilde{\kappa}_0^A, \tilde{\kappa}_1^A)$*. If* $\kappa \approx \tilde{\kappa}$*, terminate. Otherwise update and repeat from 3. Note that we cannot numerically ensure whether T is large enough. One method may be to increase T* ***after*** *this step, and checking whether the set of* κ*'s change significantly.*

As you can see this method is rather complicated and non-intuitive.

- The original Krusell and Smith (1998) paper is still heralded as the first paper to solve an incomplete markets model with aggregate uncertainty.

- Their name has become synonymous with almost any incomplete markets model with aggregate uncertainty.

- But there is no analytical foundation *a priori* why such an approximation method should work (but it does), or even if an equilibrium exists.

- Recent papers such as Cao (2011) have developed existence proofs with some additional assumptions (but generalizing in others).

Explicit Aggregation (Den Haan and Rendahl, 2010) Den Haan and Rendahl (2010) develop a simpler solution technique that is based on the simple observation that

$$K = \int a dF$$
$$K' = \int a^*(a, e; F, A) dF \tag{1}$$

of which approximation is theoretically sound. (Doesn't tell us anything about equilibrium existence though.) The basic idea is as follows.[1] First split the capital holdings by employment status, so intead of $s = [F, A]$, the true aggregate state,

1. define $s = [A, K_0, K_1]$ as the aggregate state, where (K_0, K_1) are the capital stock per capita held by unemployed and employed agents ($e \in \{0, 1\}$). Let $(\tilde{K}_0, \tilde{K}_1)$ denote the capital stocks *in this period* but *after agents make their decisions*. Since tomorrow they are hit by a shock,

$$u_{A'} K_0' = p_{AA'} \left[\pi_{00}^A u_A \tilde{K}_0 + \pi_{10}^A (1 - u_A) \tilde{K}_1 \right] \tag{2a}$$
$$(1 - u_{A'}) K_1' = p_{AA'} \left[\pi_{01}^A u_A \tilde{K}_0 + \pi_{11}^A (1 - u_A) \tilde{K}_1 \right]. \tag{2b}$$

This interim transformation is not entirely necessary, but reduces the number of approximations made and hence conducive for accuracy, since by deinition

$$K = u_A K_0 + (1 - u_A) K_1 \tag{3}$$

[1]My notes are not exactly what Den Haan and Rendahl (2010) do; they propose it more formally with more exact approximations. I just propose the idea since it has to be modified anyway to fit your needs.

With this, we don't have to worry about the exogenous law of motion induced by the transition matrices, but only solve for the law of motion between beginning and end of period capital holdings, that is the $\tilde{K}_e$'s as a function of the K_e's.

2. Whatever numerical scheme you use for the policy function, the trick is to find a way to approximate equation (1) without relying on the distribution function. For example, if the individual policy rule were linear,

$$a^*(a, e; s) = \phi^0_{e,s} + \phi^1_{e,s} a,$$

Then it is immediate that

$$u_A \tilde{K}_0 = \int a^*(a, 0; s) dF = u_A \phi^0_{0,s} + \phi^1_{0,s} \int_{e=0} a dF = u_A \left[\phi^0_{0,s} + \phi^1_{0,s} K_0\right]$$

and likewise for $\tilde{K}_1$. So once we get the individual policy functions, the aggregate law of motion should be just exactly same function. Of course in practice we don't assume linear policy functions, but typically some type of functional approximation:

$$a^*(a, e; s) = \phi_{e,s}(a),$$

but this *includes* linear interpolations (which is just a B-spline of order 1). While not entirely necessary, the computational benefit of this method is really achieved only with linear interpolation. Approximation is achieved by

$$\tilde{K}_0 = \phi_{0,s}(K_0), \quad \tilde{K}_1 = \phi_{1,s}(K_1), \tag{4}$$

i.e. the aggregation is taken *within* the basis functions. In case of linear interpolations, this is equivalent to an approximation using a first-order Taylor expansion. The resulting law of motion is then obtained from (2)-(3) in 1.

Essentially, they are proposing an aggregate law of motion that is consistent with individual decisions, rather than assuming an ad-hoc function.

EXERCISE 4 *The original paper is a bit more complicated and general, since they allow for approximating for higher-order moments as well. We assume again that agents only need to know $m_1 = K$ to make forecasts. In this sense we make the same assumption as before, but the forecasting rule will be different!*

1. *Compute the stationary distribution under the assumption the long run mean of A. Use the value and policy functions in steady state as initial guesses.*

2. *Solve the value function assuming an infinite horizon, including (K_0, K_1) as states. The grids for (K_0, K_1) can be chosen coarsely in an interval that contains K_{ss}. In each state A, individual forecasts for K' are projected according to the policy functions from the prevous iteration according to :*

   (a) *Let $a^*_{n+1}(a, e; s)$ be the policy function you want to solve for in this iteration, and a^*_n the policy function you already have from the previous iteration.*

(b) When solving for a_{n+1}, your law of motion for K' is obtained from (4) using a_n, and then applying (2)-(3). Given this agents can forecast the next period price R' (and hence (r', w')).

3. *If you want to check for accuracy, you can now simulate once the implied path of capital $\{K_t\}_{t=2}^{T}$. You can use either Monte Carlo or approximate the distribution. Compare if*

$$\frac{\int a^*(a,0;s)dF(a,0)}{u_A} \approx a^*(K_u, 0; s)$$
$$\frac{\int a^*(a,1;s)dF(a,1)}{1-u_A} \approx a^*(K_e, 1; s).$$

If the procedure is inaccurate, you need to increase the number of moments in the explicit aggregation.

This paper was purely methodological, in contrast to the Aiyagari (1994); Krusell and Smith (1998) papers which originally were addressing other questions at the time, but historically ended up being cited mainly as numerical innovations anyway.

Note that at no stage does this method rely on simulating the distribution to check whether the forecasting rule is consistent with the aggregating individual policies; this is embedded into the individual's problem already by the explicit aggregation (so we could also use this method for deterministic transitions). That is, there are no parameters like the set of κ's we needed to calibrate in the original Krusell and Smith (1998) method. The last step 3 above is not part of the algorithm, but only a post-computation check (which they have already done for you, so it works!)

If you want to see economics nerdiness at its best, take a look at the Jan 2010 issue of the Journal of Economic Dynamics and Control, where this paper was published, for a competition of ***"who can solve the Krusell and Smith (1998) model the fastest and/or most accurately?"***

References

Aiyagari, S. Rao, "Uninsured Idiosyncratic Risk and Aggregate Saving," *Quarterly Journal of Economics*, August 1994, *109* (3), 659–684.

Cao, Dan, "Collateral Shortages, Asset Price and Investment Volatility with Heterogeneous Beliefs," 2011. Georgetown University working paper.

Haan, Wouter J. Den and Pontus Rendahl, "Solving the incomplete markets model with aggregate uncertainty using explicit aggregation," *Journal of Economic Dynamics and Control*, January 2010, *34* (1), 69–78.

Krueger, Dirk and Fabrizio Perri, "Does Income Inequality Lead to Consumption Inequality? Evidence and Theory," *Review of Economic Studies*, 2006, *73*, 163–193.

Krusell, Per and Anthony Smith, "Income and Wealth Heterogeneity in the Macroeconomy," *Journal of Political Economy*, 1998, *106*, 867–896.