Breaking the Curse of Dimensionality*
Preliminary and Incomplete.

Yuliy Sannikov
Department of Economics
Stanford GSB
sannikov@stanford.edu

April 8, 2022

Abstract

The paper proposes several ideas, as well as two specific computational procedures, to solve economic models with high or infinite-dimensional state spaces. The first idea is to identify the relevant portion of the state space that is *likely* to arise in equilibrium. Another idea is to identify the *major* dimensions of the "likely" state space, such that the variations in the remaining dimensions are small and can be approximated linearly. Computational cost is exponential in major directions, but only quadratic in minor directions. The first proposed computational procedure is based on a randomly generated history of exogenous shocks to the system. This procedure approximates relevant off-path values using just major dimensions of the state space, which are updated in real time. The second iterative procedure is more systematic, taking major and minor dimensions explicitly.

1 Introduction.

The curse of dimensionality is a term coined by Richard E. Bellman to describe the phenomenon that the size of the state space grows exponentially with the number of states. This problem poses tractability challenges in economics in problems where the full description of an economic system at any point of time is multidimensional. The problem arises because economic agents anticipate the future when making decisions. They do not know what will happen, and have to anticipate all the possibilities. With many outcomes that can play out in different ways, the number of combinations becomes exponential.

*****.

This paper proposes several ideas to help analyze economies with high-dimensional or infinite-dimensional state spaces, "breaking the curse of dimensionality." The ideas are based on several of observations about the behavior of complex systems, in order to determine certain approximations that can simplify the problem. The goal is two-fold. First, it is to expand the range of models that an economist can write, with the hope of solving it numerically. Second, it is to obtain a sufficiently readable and intuitive representation of the solution of the complex system. That is, we want to understand what are the key driving forces behind the solution. That is, we want to not only solve complex models, but also to gain clarity about how the model works.

The seminal paper of Krusell and Smith (1998) (hereafter KS) provides inspiration for these two ideas. KS propose a model of an economy, in which individuals save in capital to insure themselves against idiosyncratic unemployment shocks. In addition, aggregate shocks shift the economy between booms and recessions. In booms, say, 4% of individuals are unemployed and in recessions, 10%. Different individuals have different savings levels, which depend on their personal unemployment histories. The map from savings to consumption is nonlinear, hence distribution of savings among employed and unemployed play the role of a state variable. Even though the state, the entire savings distribution, is infinite-dimensional, KS show that average savings are essentially all individuals need to know to make individual decisions. Moreover, the drift of savings is essentially a linear function of savings in booms and recessions along the equilibrium path. It is a beautiful illustration of what we economists needs to understand the complex system.

More generally, KS propose the idea of describing the system by a few set of moments, and testing how many moments are sufficient. Of course, a number of questions remain. Economists have been wondering about how to pick the moments. How large is the limitation of a small number of moments, needed for computational tractability? Is it possible to handle, say, twenty moments? In general, is there a recipe for identifying the right simplification of the high-dimensional space?

This paper proposes two procedures to handle complex models. These procedures give us a lot of hope about high-dimensional models. They identify certain properties of solutions, which may be quite universal, and which offer the possibility that a sufficiently simple description exists. We illustrate these procedures using a variation of the Krusell-Smith model, in which one moment is not sufficient to describe system dynamics on the equilibrium path.

Broadly speaking, why is there hope for tractability? There are two reasons. First, while the state space in terms of all possible combinations of the descriptors of the system is huge, the portion of the state space that occurs along the equilibrium path is typically much smaller. From our knowledge of the model, we expect wealth distributions to be bell-shaped, with certain properties of tails, and with higher mean in employment than in unemployment. Even so, the set is potentially infinite dimensional. The second reason is that the dimensions can be ranked in importance, hence the numerical cost of approximating more dimensions can decay.

Our first procedure solves for the model for a long randomly generated aggregate history. For the model we consider, it is a history of regime switches between recessions and booms. For each moment t of the history, we ask the question about the distribution of savings g_t and the value function of individuals f_t at that moment. The value functions imply consumption-savings policies, which imply how the distribution of savings evolves going forward and give us g_{t+1}. To sum up, next-period wealth distribution g_{t+1} depends on g_t, individual's behavior and the state (boom or recession) in period t. What about value functions? Imagine, for concreteness that period $t+1$ for the history we generated is a boom period. Then, through the Bellman equation, the value function f_t depends on the next-period value function f_{t+1} in the event of boom, as well as value function $\hat{f}_{t+1}$ that individuals would have if it were recession. Thus, along our randomly generated history, we solve for g_t going forward, we solve for f_t going backward, but the piece that is most challenging is this: how do we determine $\hat{f}_{t+1}$? That is what value function do individuals have in recession at the computed distribution of wealth g_{t+1}? We can estimate $\hat{f}_{t+1}$ using data points in recession. For example, we could search for recession periods at which the distribution is closest to g_{t+1}, and interpolate/extrapolate to estimate $\hat{f}_{t+1}$.

The procedure provides remarkably useful information about the solution structure of a complex model. In particular, we learn about the structure of the state space of distributions d_t. Principal component analysis of the set of distributions tells us that, indeed, the state space is high dimensional. Many components matter. However, there is great news: the standard deviations of components decay quite fast.

This observation motivates the second procedure, which solves the system directly using the information about the state space we just learned. In this procedure, we divide principal components into major and minor. We place grid on the major principal components, and approximate the effects of the minor principal components *linearly*. Linearization breaks the curse of dimensionality. While a general d-dimensional system with n grid points along each dimension requires a total of n^d points, a d-dimensional *linear* system requires a $d \times d$ matrix to describe dynamics. The computational cost is no longer exponential along the minor components.

The paper is organized as follows. Section 2 presents the model, a variation of KS, which serves as a laboratory for the proposed computational methods. Section 3 proposes our first computational procedure, based on a randomly generated aggregate history that drives the system. The computed example motivates the identification and ranking in importance of dimensions of the state space to get around the curse of dimensionality. Section 4 describes the second computational procedure, which explicitly incorporates information about the likely portion of the state space as well as major and minor dimensions. Section 5 concludes.

2 The Model.

We consider a variation the model of Krusell and Smith (1998) to illustrate the proposed computational methods. The model draws from the tradition of Bewley, Huggett and Aiyagari, and we build heavily on the contributions of Moll and Kaplan who have developed the continuous-time theory of these models.

We write the model for the traditional continuous timeline $t \in [0, \infty)$, but in many cases we will consider timeline $(-\infty, \infty)$ to discuss an economy that has been running for a long time already. There is a unit mass of individuals indexed by $i \in [0, 1]$. Each individual has labor endowment l_1 when employed and $l_0 \in [0, l_1)$ when unemployed. Given the total supply of labor L_t and capital K_t in the economy, the aggregate production function is

$$z(j_t) K_t^\alpha L_t^{1-\alpha},$$

where $\alpha \in [0, 1)$. The productivity parameter z is indexed by the aggregate economic state, boom if $j = B$ and recession if $j = R$. Given this production function, each individual gets the wage of

$$w_t = (1 - \alpha) z(j_t) \frac{K_t^\alpha}{L_t^\alpha} \tag{1}$$

per unit of labor supply. Each unit of capital receives the rental rate of

$$\alpha z(j_t) \frac{L_t^{1-\alpha}}{K_t^{1-\alpha}}.$$

There are no investment frictions: capital can be freely converted into consumption goods and vice versa. Capital depreciates at rate δ, so given aggregate consumption C_t, the law of motion of aggregate capital is

$$dK_t = z(j_t) K_t^\alpha L_t^{1-\alpha} \, dt - \delta K_t \, dt - C_t \, dt.$$

The risk-free rate in the economy is

$$r_t = \alpha z(j_t) \frac{L_t^{1-\alpha}}{K_t^{1-\alpha}} - \delta. \tag{2}$$

increases in aggregate labor supply and declines in aggregate capital supply.

Hence, the wealth of individual i follows

$$dk_t = r_t k_t \, dt + l_{s(i,t)} w_t \, dt - c_t \, dt,$$

where $s(i, t)$ is the employment indicator of individual i at time t. Individuals have increasing and concave utility $u(c)$ and common discount rate ρ.

In aggregate state of the economy j, an individual transitions from employment to unemployment according to a Poisson process with intensity $\lambda_{10}(j)$ and transitions back to employment with intensity $\lambda_{01}(j)$. The whole economy transitions between aggregate states $j = B$ and R with intensity $\theta \geq 0$.

The full descriptor of the state of the economy at time t includes the aggregate state j_t as well as the density $g_t(k,s)$ over capital holdings and employment states $s \in \{0,1\}$. The state space is infinite dimensional. Individual value functions $f_t(k,s)$ depend on the aggregate state variables, so we can write them as $f(k,s|g_t,j_t)$.

Total labor supply is

$$L_t = l_0 \int_0^\infty g_t(k,0)\,dk + l_1 \int_0^\infty g_t(k,1)\,dk.$$

The market for capital clears if aggregate capital satisfies

$$K_t = \int_0^\infty k(g_t(k,0) + g_t(k,1))\,dk.$$

There is a subtle distinction between this model and that of KS. In KS, aggregate and individual transitions are determined jointly, so that the mass of unemployed is always the same in each state of the economy (for example, unemployment is 10% in recession and 4% in booms). In contrast here, unemployment would *gradually* converge to 10% in recessions, through individual Poisson transitions, and gradually converge to 4% in booms. Due to this distinction, the result of KS that aggregate capital describes the future interest rate process almost completely is no longer true in our version. Unemployment also matters, because the wealth of employed and unemployed evolves differently. Recall from KS that the interest rate process is what individuals need to make optimal consumption-savings decisions.

It is clear how the equilibrium is defined in this standard model, but for completeness let me give a brief description in words. For any initial distribution $g_0(k,s)$, equilibrium consists of full description of how the economy evolves, and individual consumption-savings decisions, after any history of aggregate shocks $\{j_s, s \leq t\}$ and private employment states. K_t determines individual wages w_t as well as the rental rate of capital. Individuals solve their optimization problems to maximize their expected future utilities, and all market have to clear.

2.1 Stationary Solutions.

We can get a lot of intuition behind this model by considering its stationary version, drawing from the beautiful analysis of Achdou et. al (2022). In the dynamic version of the model, individuals take into account their individual employment state, aggregate state j_t, as well as their expectation of the entire future process of interest rates. Future interest rates depend on the law of motion of the entire wealth distribution. In a static version, we can fix the interest rate r exogenously. How do individual behaviors depend on the fixed interest rates as well as transition between employment and unemployment? What does this behavior imply about the distribution of wealth?

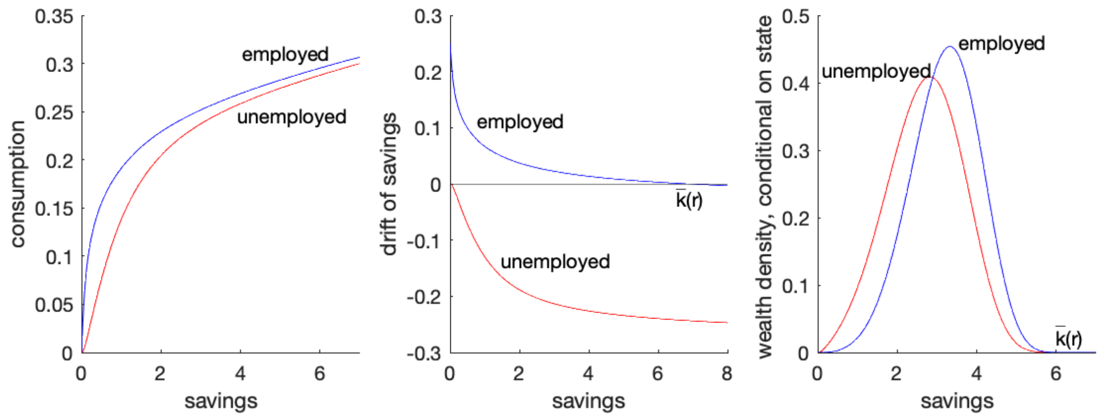

Figure 1: Stationary solution for a fixed rate r.

Figure 1 illustrates the optimal policy. Consumption is increasing in wealth, and higher in the state of employment than in unemployment for any wealth level. The drift of savings is negative in the unemployment state. That is, the unemployed consume more than their total labor and capital income. In the employment state the drift wealth is positive until it reaches 0 at critical level $\bar{k}(r)$, which is finite as long as $r < \rho$.[1] Wealth level can never reach levels where the drift is negative in the state of employment in the stationary equilibrium. Therefore, we can say that in the stationary solution, the drift of wealth of the employed is nonnegative. The stationary distribution of wealth $g(k, s)$ belongs to the interval $[0, \bar{k}(r)]$. Due to the sign of the drifts, conditional on employment the wealth distribution first order stochastically dominates the portion conditional on unemployment.

Given transition intensities the mass of the unemployed is

$$\int_0^\infty g(k,0)\,dk = \frac{\lambda_{10}}{\lambda_{01}+\lambda_{10}}$$

and r is the stationary endogenous equilibrium rate if

$$r = \alpha z \frac{L^{1-\alpha}}{K^{1-\alpha}} - \delta \quad \text{for} \quad K = \int_0^\infty k(g(k,0)+g(k,1))\,dk, \quad L = \frac{\lambda_{10}l_0 + \lambda_{10}l_1}{\lambda_{01}+\lambda_{10}}.$$

The equilibrium rate is determined by demand for capital for production (where rate r is decreasing in K) and supply of capital from the stationary distribution of savings. The drift, as well as the stationary distribution of savings, increase in r.

An implication of the process by which the equilibrium rate is formed is that if individuals want to save more for a fixed rate r (e.g. due to the precautionary motive),

[1] The assumption that $r < \rho$ is important because with $r > \rho$ average savings converge to infinity with or without idiosyncratic risk.

then the amount of capital K in the economy will go up. Then, to reach equilibrium, rate r must go down.

Therefore, let us see how changes in parameters λ_{10} and λ_{01}, which guide transitions between employment and unemployment, affect the optimal savings behavior *for a fixed rate* r. Generally speaking, if the intensity of unemployment shocks λ_{10} increases, then average savings increase. This is not a universally true statement, because obviously in the limit as $\lambda_{10} \to \infty$, everybody is unemployed and average savings equal zero (even if $l_0 > 0$, as long as $r < \rho$). However, at $\lambda_{10} = 0$, average savings are also 0, and they rise as λ_{10} increases away from 0. As recovery rate from unemployment λ_{01} increases, average savings obviously decrease. Moreover, an interesting comparative static arises if both λ_{10} and λ_{01} increase, so that the mass of unemployed stays the same. Then savings decrease because less persistent unemployment spells imply less risk to individuals and lower precautionary motive.

What implications do these observations have on the equilibrium in a nonstationary model with aggregate shocks? As a thought experiment, consider a stationary model permanently in the boom state, and suppose that an unexpected shock puts the economy in permanently recession, with higher λ_{10} and lower λ_{01}. Then unemployment starts to increase, so L_t decreases. These mechanical processes put downward pressure on the rates r_t. The increased precautionary motive also puts downward pressure on r_t, as at any fixed rate r individuals would like to save more. However, r_t is not fixed, and individuals take into account their expectation of all future rates to adjust savings behavior. In typical examples, adjustment in the labor market occurs at much higher speed $\lambda_{10} + \lambda_{01}$ than the speed of wealth adjustment, which is roughly of order r_t. From the perspective of individuals, increased precautionary motive competes with the expectation of lower rates due to contraction in the labor market.

Of course, we have to take this thought experiment of a one-time unexpected with a grain of salt. It is a useful thought experiment, but we have to recognize that an economy permanently on one state is different from one in which individuals anticipate that the regime may switch at any moment. For example, the precautionary motive in an economy permanently in recession is very different from one in which individuals are waiting to return to a boom phase. That is why it is important, and exciting, to learn what actually happens in a fully stochastic economy with aggregate regime switches.

From this discussion of the stationary model, we get some intuition about the dimensionality of our problem. One dimension captures the adjustment of the labor market, and other dimensions captures the slower adjustments of savings due to the different behavior of employed and unemployed, and the precautionary motive. In addition, savings respond to rates, which depend on all of the labor market as well as aggregate wealth, and which create effects that ripple back into behavior. This intuition suggests that an ideal computational procedure for this model would identify the major forces as well as the ripples, and take all of them into account appropriately.

3 Procedure 1: Computation on a random history.

We would like to assess what happens in this economy when the aggregate state j switches between regimes R and B at Poisson rate θ. In this economy with aggregate risk, the state variable takes the form $(g_t, j_t) \in \mathcal{G} \times \{R, B\}$, where $\mathcal{G}$ is the set of wealth distributions, i.e. functions $g : [0, \infty) \times \{0, 1\} \to [0, \infty)$ that satisfy

$$\int_0^\infty (g(k,0) + g(k,1))\, dk = 1.$$

The equilibrium is fully characterized by the value functions $f(k, s|g, j)$. Hence, we would like to find the map from $\mathcal{G} \times \{R, B\}$ to $\mathcal{F}$, where $\mathcal{F}$ is the space of functions $f : [0, \infty) \times \{0, 1\} \to \mathbb{R}$. The values of g_t, j_t and f_t determine all remaining equilibrium objects, i.e. policy $c(k, s|g_t, j_t)$ from the first-order condition

$$u'(c) = \frac{\partial f(k, s|g_t, j_t)}{\partial k},$$

aggregate capital and labor (in both regimes R and B) from

$$K = \int_0^\infty k(g(k,0) + g(k,1))\, dk, \quad L = l_0 \int_0^\infty g(k,0)\, dk + l_1 \int_0^\infty g(k,0)\, dk \tag{3}$$

and wage as well as the risk-free rate from (1) and (2). To sum up, we would like to compute the map from (g, j) to f.

One of our guiding principles, so address the curse of dimensionality, is to focus not on the entire state space but rather on states that the model is likely to reach. Suppose the "state space" contains all pairs (g, j), the "theoretically possible" state space contains pairs (g, j) that can arise in equilibrium in an economy that has been running since time $t = -\infty$, and the "likely" state space contains pairs (g, j) that exclude outliers that are extremely rare. To keep computation manageable, our idea is to focus on "likely" state space. With that goal, algorithm 1 starts by generating a long history of regimes for a given number $2N$ of regime switches. Effectively, this amounts to drawing $2N$ exponential random variables for the rate parameter θ that give the duration of each regime. We thus obtain a regime indicator function $j : [0, T] \to \{R, B\}$, which starts with a block of R on $[0, t_1)$ followed by a block of B over $[t_1, t_2)$ and so no. The last N-th block of B over $[t_{2N-1}, t_{2N})$ ends at $t_{2N} = T$ which is the beginning of the next block (i.e. $j(T) = R$). We then ask the question, what distribution of wealth g_t and what value functions f_t will individuals have at any time $t \in [0, T]$ along this history? For this question to be precise, we also have to specify the starting distribution g_0 at time 0, even though the starting value g_0 will not matter a whole lot when t is large. For convenience, we shall assume that the distribution g_0 at time 0 is one that obtains after the economy has been running continuously over $(-\infty, 0]$ with cycles $j : [0, T] \to \{R, B\}$ repeating over time periods

$[-T, 0]$, $[-2T, -T]$, and so on. This assumption is convenient numerically. It implies that we can connect time 0 to time T when iterating until convergence.

When discussing our model, we can take the holistic perspective of the state space and focus on the entire function $f(k, s|g, j)$, or we can take the dynamic perspective to focus on a particular realized history of aggregate shocks j. When using the dynamic perspective,[2]. it is helpful to use the notation of $g_t : [0, \infty) \times \{0, 1\} \to [0, \infty)$ to denote the distribution at time t and $f_t : [0, \infty) \times \{0, 1\} \to \mathbb{R}$ to denote the value functions at time t. It is also useful to denote by $\hat{f}_t$ the individuals' value functions in the event that the regime is opposite at time t. That is $\hat{f}_t$ are the value functions for history $\{j_s, s < t, \hat{j}_t\}$ where $\hat{j}_t$ is the opposite regime of j_t. We do not need to use notation for $\hat{g}_t$, the distribution in the event the regime is opposite at time t, since $\hat{g}_t = g_t$ under our assumptions. Formally,

$$f_t = f(\cdot, \cdot|g_t, j_t) \quad \text{and} \quad \hat{f}_t = f(\cdot, \cdot|g_t, \hat{j}_t).$$

Using the dynamic perspective, we can write the HJB equation of an individual in either regime as

$$\rho f_t = \max_c \; u(c) + \underbrace{\frac{\partial f_t}{\partial k}(r_t k + l(s) w_t - c) + \lambda_{s,1-s}(f_t(k, 1-s) - f_t)}_{\mathcal{L}_t f_t} \tag{4}$$

$$+ \frac{\partial f_t}{\partial t} + \theta(\hat{f}_t - f_t).$$

If we take $\hat{f}_t$ as given, this equation can be solved using standard methods as in deterministic models. For a particular random history, algorithm 1 solves equation (4) backwards over each block $[t_n, t_{n+1}]$ from terminal condition $f_{t_{n+1}} = \hat{f}_{t_{n+1}-}$.

Notice that equation (4) defines the differential-jump operator $\mathcal{L}_t$. Then the density function g_t satisfies the Kolmogorov forward equation (KFE)

$$\frac{\partial g_t}{\partial t} = \mathcal{L}_t^* g_t, \tag{5}$$

where $\mathcal{L}_t^*$ is the adjoint of the operator $\mathcal{L}_t$. The fact that the differential operator in KFE is the adjoint of the operator in the HJB equation is convenient numerically. For details, see Achdou et. al. (2022), for example.

For history j, Procedure 1 starts with a guess iteratively solves (4) to update value functions and policies (i.e. the operator $\mathcal{L}_t$), (5) to update the density function g_t and intermittently uses this data to update the counterfactuals, i.e. the value functions $\hat{f}_t$ in the event that the regime at time t had been different from what it is at history j. Appendix A provides details of the particular numerical implementation of Procedure 1 that I use. Here, I would like to discuss a bit the question of estimating $\hat{f}_t$ from data containing f_t and g_t.

[2]Of course, the notation from the dynamic perspective slightly abuses notation.

Specifically, consider the question of estimating $\hat{f}_t$. Suppose the state at time t is $j_t = R$, so $\hat{f}_t$ is a value function for the opposite state $j_t = B$. Consider the set of time points $S_B \subset [0, T)$ generated numerically where the state is B. We have the data points $\{g_{t'}, f_{t'} \mid t' \in S_B\}$, which give us computed value functions in state B. We would like to use this data to estimate the value function at distribution g_t, i.e. the value function $\hat{f}_t$.

There are different approaches to this question, but my numerical implementation matches distributions based on moments given by the principal components of the set of distributions $\{g_{t'} \mid t' \in S_B\}$. Denote the main, say, n principal components by d_1, $d_2 \ldots$ and d_n.[3] For example, consider $n = 2$. Let us represent the distribution $g_{t'}$ by its coordinates $x_{t',i} = \langle g_{t'}, d_i \rangle$ denotes the inner product, $i = 1, 2$. The set of points $X_B = \{x_{t'} \mid t' \in S_B\}$ can be triangulated, using Delaunay triangulation. If point x_t that corresponds to distribution g_t belongs to a particular triangle $x_{t_1}, x_{t_2}, x_{t_3}$, then we take the convex combination

$$x_t = w^1 x_{t_1} + w^2 x_{t_2} + w^3 x_{t_3},$$

and use it to approximate

$$\hat{f}_t = w^1 f_{t_1} + w^2 f_{t_2} + w^3 f_{t_3}$$

as the convex combination of value functions that correspond to the approximating distributions. If $x_{t'}$ is not in the convex hull of X_B, we approximate $\hat{f}_t$ by matching x_t to its closest neighbor in X_B.

A couple of remarks are in order. First, the implementation approximates $\hat{f}_t$ based on only a small number of n moments of g_t in the spirit of KS, but these moments are chosen endogenously as those in which the distributions that the model actually generates exhibit the most variation. This approximation technique is based on the presumption that other moments matter a lot less.[4] Second, the approximation is based on a randomly generated grid rather than a systematic grid, but the grid is based on the actual distribution. This has advantages and drawbacks. As an advantage, we get greater precision, in virtue of having more grid points, in areas where it really matters. Certain regions through which the system moves fast have few grid points. Among drawbacks, random grids lead to less clarity in representing the solution. Also, we sacrifice some precision near the boundary of the "likely" state space.

[3] My numerical implementation of principal component analysis is based on the norm that measures the distance between cumulative distribution functions, instead of pointwise difference between densities.

[4] Other principal components, or moments are ignored in the present algorithm. However, it is possible to take them into account by a more sophisticated procedure, e.g. using a linear regression that includes other neighbors of g_t.

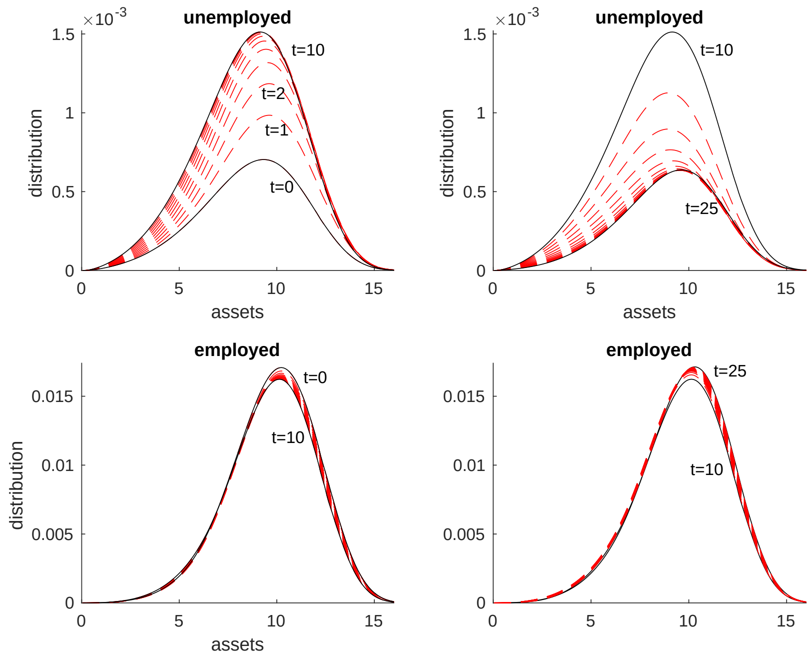

Figure 2: The evolution of distributions.

3.1 Numerical Example.

Following KS, let us take the intensity of transitions between aggregate economic states $j = B, R$ to be $\theta = 1/8$. The average duration of each boom and recession phase is then 8 periods/quarters. In each state, transition intensities between employment and unemployment are

$$R: \quad \lambda_{01} = 0.4, \quad \lambda_{10} = 4/90$$

$$B: \quad \lambda_{01} = 2/3, \ \lambda_{10} = 1/36.$$

Given these parameters, the unemployment rate converge to 4% in booms and to 10% in recessions. Labor endowments are $l_0 = 0$, $l_1 = 0.25$ and individuals have CRRA utility, $u(c) = c^{1-\nu}/(1-\nu)$ with $\nu = 2$, and quarterly discount rate $\rho = 0.01$. Furthermore, take $z(B) = 1.01$, $z(R) = 0.99$, $\delta = 0.025$ and $\alpha = 0.36$.

We start by drawing a sequence of 200 boom and recession intervals, to generate a random history of length $[0, T]$. We then solve forward for the distributions and backward for the value functions along this history. To ensure that the initial distribution g_0 is typical for an economy that has been running for a long time, we set $g_0 = g_T$. We proceed until convergence. If $[t_n, t_{n+1})$ is a recession interval, then estimate the value function $\hat{f}_{t_{n+1}}$ at $(g_{t_{n+1}}, R)$ and solve (4) over the interval $[t_n, t_{n+1}]$ backward from terminal condition $f_{t_{n+1}} = \hat{f}_{t_{n+1}}$. In (4), for $t \in [t_n, t_{n+1})$ we take $\hat{f}_t$ to be our estimate of the value function at (g_t, B).

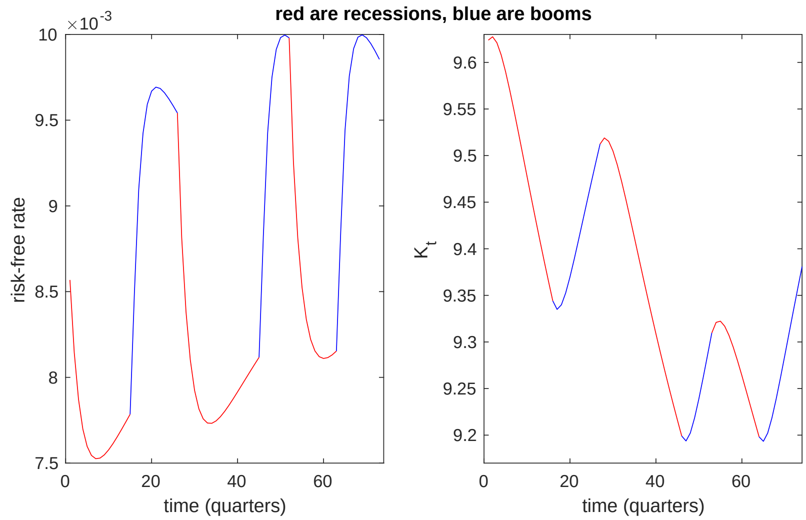

Figure 3: Evolution of the risk-free rate and capital.

Figure 2 shows the evolution of wealth distribution for a history that starts with 10 recession periods (left panels), followed by 15 boom periods (right panels). Top panels are the unemployed individuals and bottom panels are the employed. We see that the typical shape of distributions remains unchanged: they are bell shaped with thicker tails at low wealth levels, and with higher mean for the employed than for unemployed. Important variations occur in the mass of the unemployed, and in mean wealth. The similarity among distribution suggests that, even though the set of all distributions is infinite dimensional, perhaps the set of distributions which are actually likely to arise not that large. The similarity in distributions suggests that we may be able to break the curse of dimensionality.

The two main dimensions, in which the distributions change, are as follows. First, obviously, the proportion of unemployed changes, converging quickly to 10% in recessions and 4% in booms. Second, the unemployed, who spend their savings and have no income, pull the average wealth of both groups to the left, while the employed pull the distribution in the opposite direction. In recessions, the mass of unemployed is greater, so the long-run mean of savings is lower. Changes in employment and savings do not occur at the same rate: saving and spending are slower than finding or losing jobs for these parameters.

This has implications for interest rates. When recession hits, unemployment rises quickly, leading to a drop in interest rates as return on capital decreases. As recession sets in, the increased mass of the unemployed who spend their savings leads to a decrease in the amount of capital in the economy and a rise in interest rates. Hence, in each regime, interest rate as a function of time is non-monotonic. In recessions, the interest rate drops sharply, and then rises slowly. The left panel of Figure 3 shows how the risk-free rate evolves over time. The right panel shows capital, which tends

Components	1	2	3	4	5	6	7	8
Our Version	.4354	.2469	.0862	.0528	.0311	.0242	.0134	.0087
KS: B	.3380	.0342	.0291	.0121	.0054	.0035	.0022	.0013
KS: R	.3160	.0331	.0300	.0236	.0107	.0049	.0038	.0019

Table 1: Standard deviations of the first principal components.

to drop in recessions and rise in booms. In recessions, for example, the rate of drop accelerates to a fairly steady rate, as unemployment rises to 10%. The initial rise in capital at the very onset of the recession is due to the precautionary motive.

Interestingly, here we see one way in which the stationary model we discussed in section 2.1 could possibly give us wrong intuition. In the stationary model, average savings are higher for labor market parameters in recession, due to much greater precautionary motive. In contrast, in the dynamic model capital holdings fall in recessions and rise in booms. This pattern is opposite of what happens in the stationary model. The reason for this discrepancy is that the precautionary motive in the dynamic model is much weaker, as individuals take into account the impeding return to the boom regime. This observation highlights the importance of the dynamic perspective, where individuals take into account the probabilities of switching regimes.

The major forces that affect the evolution of the wealth distribution in the infinite-dimensional space are employment shocks, savings by the employed and dissavings by the unemployed. All these forces affect the interest rate and wage processes, and feed back into savings behavior. These observations suggest a mental model of the major forces and smaller feedback effects. These ripples make the set of distributions infinite dimensional, but the set of distributions is "thin" in the ripple dimensions, so it is reasonable to linearize the ripple effects. To explore how this works in an actual computed solution, we look at the principal components of the set of distributions generated by this algorithm. Table 1 gives the standard deviations of the first eight principal components.

The first row corresponds to our version of the model. The first two large principal components are most important: they correspond to the major effects seen in Figure 2, i.e. changes in unemployment, as well as savings by the employed and spending by the unemployed. The remaining components can be thought of as ripple effects. For comparison rows 2 and 3 illustrate boom and recession regimes in a continuous-time version of KS, where unemployment changes only once when regime switches, and then stays constant in both recession and boom. In those rows, one principal component is clearly dominant, and all remaining "ripple effects" are a lot smaller. This observation suggests that a single moment is the major determinant of equilibrium behavior in KS.

It is instructive to see how the principal components change over time. Figure 4 shows a simulated path of the economy over 200 periods. The four panels show the

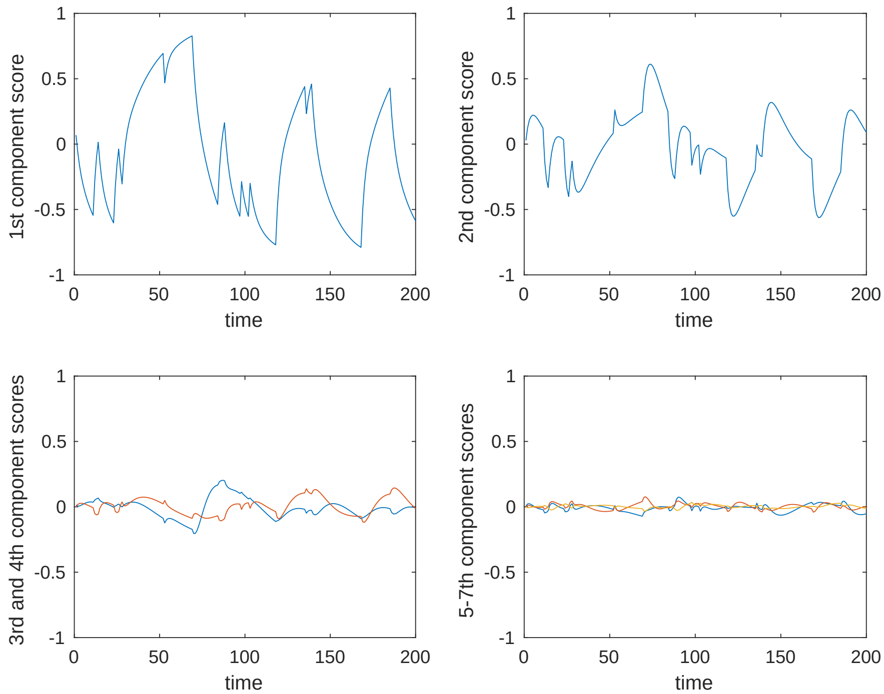

Figure 4: Sample paths of component scores.

paths of scores on the top 7 components of the distribution of wealth. The scale of the vertical axis is intentionally kept identical across the four panels. We see that the major components follow a clear pattern, while the paths of more minor components look like ripples. The major components are the main drivers of the system. It looks as if we should be able to forecast how the system will evolve, at least conditional on regimes, from the first components quite accurately. The movement of higher components looks more ripple-like and erratic. Because the variation of higher order component scores are small, we can guess that, perhaps, it is reasonable to estimate the effect of higher-order components to be linear.

Indeed our second algorithm uses the output of algorithm 1 and these ideas. Specifically, we can use the output of algorithm 1 to place a more systematic grid on the state space along the major principal components. We can take into account many dimensions of the state space by approximating the effect of higher-order components to be linear.

4 Computing on the state space of distributions.

The second algorithm operates directly in the state space of distributions. In order to run, we need to have a fairly good idea of the part of "likely" region of state space.

Before describing the numeric procedure, it is useful to develop the equilibrium equations in the infinite-dimensional space a bit further. Recall the notation $f(k, s|g, j)$, which highlights how the entire value function depends on the distribution of wealth g as well as the aggregate regime j. Denote by ∇f the Jacobian of the entire function f in the space of distributions. If $\dot{g}$ is rate of change of the distribution of wealth with time, conditional on the current aggregate regime remaining unchanged, then the time change of f is $\nabla f\, \dot{g}$.

$$\rho f = \max_c \; u(c) + \underbrace{\frac{\partial f}{\partial k}(r(g,j)k + l(s)w(g,j) - c) + \lambda_{s\hat{s}}(f(k,\hat{s}|g,j) - f)}_{\mathcal{L}f} \tag{6}$$

$$+\nabla f\, \dot{g} + \theta(f(k, s \mid g, \hat{j}) - f).$$

Here $\hat{j}$ denotes the regime opposite from j, and likewise $\hat{s} = 1-s$. The current interest and wage rates $r(g, j)$ and $w(g, j)$ are completely determined by the pair (g, j) via (1) and (2) with capital and labor determined by (3). Recall that the rate of change of the wealth distribution with time is given by the Kolmogorov equation (5), which we can write as

$$\dot{g} = \mathcal{L}^* g,$$

where $\mathcal{L}^*$ is the adjoint of the operator $\mathcal{L}$ in (7). Procedure 2 aims to solve equation (7) on the space of distributions directly.

4.1 Procedure 2.

The procedure is based on a selection of the *major* principal components d_1 through d_n which are taken into account explicitly, and *minor components* d_{n+1} through d_{n+m} which are taken into account via linear approximations. For example, we can take $n = 3$ and $m = 10$. The complexity of our procedure is exponential in n but only quadratic in m.

Denote by $\bar{g}$ the center of the "likely" region of the state space, i.e. the average of the distributions generated by procedure 1. We create a grid around $\bar{g}$ as follows. The size of the grid in major directions is

$$\{-N_1, \ldots - 1, 0, 1, 2 \ldots N_1\} \times \{-N_2, \ldots N_2\} \ldots \times \{-N_n, \ldots N_n\}.$$

Denote the size of the grid in dimension $k = 1, 2, \ldots n$ by Δ_k, and let us choose Δ_k so that $\Delta_k N_k$ corresponds to several standard deviations of the principal component k.

Point $x = (x_1, x_2 \ldots x_n)$, $x_i \in [-N_i \ldots N_i]$ of the grid corresponds to the distribution

$$g(x) = \bar{g} + \sum_{k=1}^{n} \Delta_k x_k d_k.$$

In addition, at each x for each minor component $n+k$, $k = 1 \dots m$, we place an additional grid point for the distribution $g(x,k) \equiv g(x) + \Delta_{n+k} d_{n+k}$. In the end, the size of our grid is

$$2(m+1)\prod_{k=1}^{n}(2N_k+1),$$

where the factor of 2 accounts for regimes $j \in \{R, B\}$.

Notice that procedure 1 fixed neither the set of distributions nor the principal component dimensions that were used for approximation. Those were updated dynamically. In contrast, procedure 2 fixes the set of distributions. Hence, procedure 2 is more similar to traditional computation on a fixed finite-dimensional grid. Notice also that not all our grid points $g(x)$ may be true densities (i.e. they may not all be nonnegative). That is okay, because what they aim to measure is the effect of change of the wealth distribution in a particular direction.

To start procedure 2, we need a guess of value functions $f(\cdot,\cdot \mid g, j)$ on the grid. Those can be obtained using the method we employed in procedure 1, when finding value function $\hat{f}$ for a particular regime $\hat{j}$ at a particular distribution, using data generated from a particular random history. We then update value functions by iterating equation (7).

We should discuss how we handle the term $\nabla f \, \mathcal{L}^* g$ when solving (7). In the basis of principal components, $\dot{g}$ has coordinates $\dot{g}_k = \langle \mathcal{L}^* g, d_k \rangle$. We ignore the coordinates beyond the n major and m minor principal components. Then, for $k = 1, \dots n$, we evaluate $D_k f \equiv \nabla f \, d_k$ using our grid in the standard way, i.e. via the upwind scheme.

For $k = n+1 \dots n+m$, we estimate[5]

$$\nabla f \; d_k = \frac{f(x,k) - f(x)}{\Delta_{n+k}}.$$

Hence, the discretized version of equation (7) which we solve backwards is

$$-\frac{\partial f}{\partial t} = \max_c \; u(c) + \underbrace{\frac{\partial f}{\partial k}(r(g,j)k + l(s)w(g,j) - c) + \lambda_{s\hat{s}}(f(k,\hat{s}|g,j) - f)}_{\mathcal{L}f} \qquad (7)$$

$$+ \sum_{k=1}^{n} D_k f \; \dot{g}_k + \sum_{k=1}^{m} \frac{f(x,k) - f(x)}{\Delta_{n+k}} \dot{g}_{n+k} + \theta(f(k,s \mid g,\hat{j}) - f) - \rho f.$$

Our iterative procedure proceeds backward in time until convergence at all grid points x and (x,k).

[5]Technically, this way of estimating the derivative is leads to a numerical scheme that is non-monotone. Hence, this issue can cause problems in numerical stability, and it requires special attention. Stability issues do not arise in examples I computed, and casual analysis shows that stability is connected with the thinness of the set of distributions along minor principal components. Forces keep distributions close to the mean in those directions help stability.

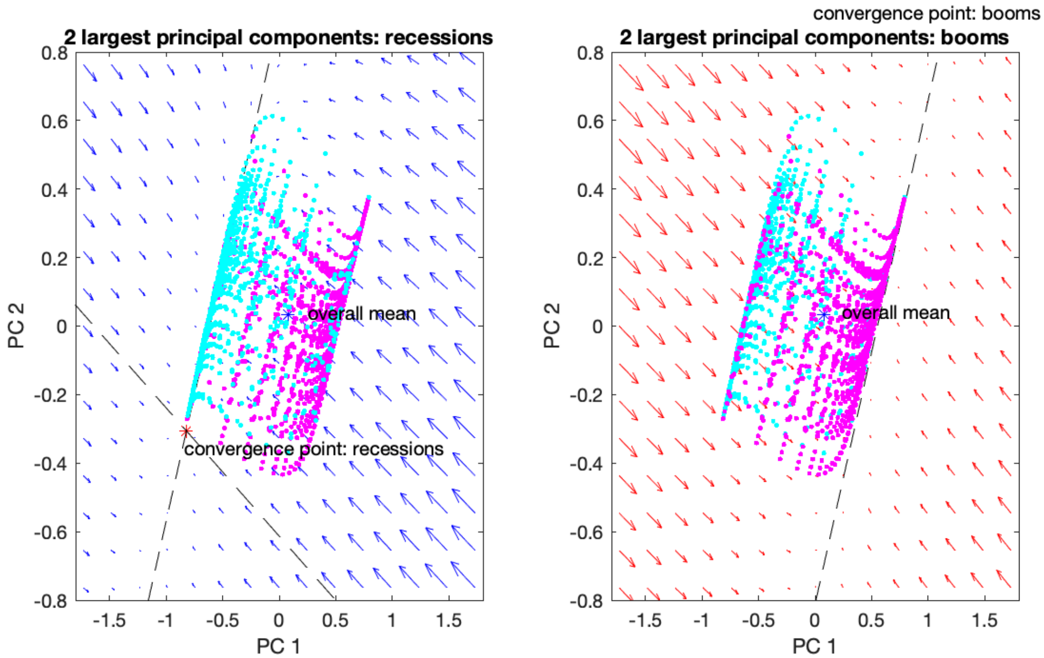

Figure 5: Dynamics of the major principal components (the rest are set to 0).

4.2 Computed Examples.

We start by placing a grid over the top two major principal components of distributions and incorporating 8 minor components linearly. The evolution of distributions can be described through two phase diagrams of the scores on principal components, one for each state. Figure 5 shows the evolution of the first two principal component scores when the coordinates of the rest are set to 0. Recall from Table 1 that the importance of principal components beyond the first two drops dramatically for our version of the model

The scale of each axis corresponds to the distance between the cumulative distribution functions. One unit of distance corresponds to shifting the distribution by one unit of capital within each group. The principal components are normalized to have a positive mass of capital. Each panel indicates the point to which the distribution converges if the economy stays in one state for a long time. For booms, that point is not in the area of the chart but rather above it. The dashed lines *roughly* indicate the eigenvector directions for the movement of the system towards the convergence point. The eigenvector directions can also be read from the directions of the arrows. This is possible to do because the dynamics in this example turns out to be quite linear.

The two panels also display scatter plots for the distributions generated by procedure 1. The scattered points make sense in light of the phase diagrams. They give

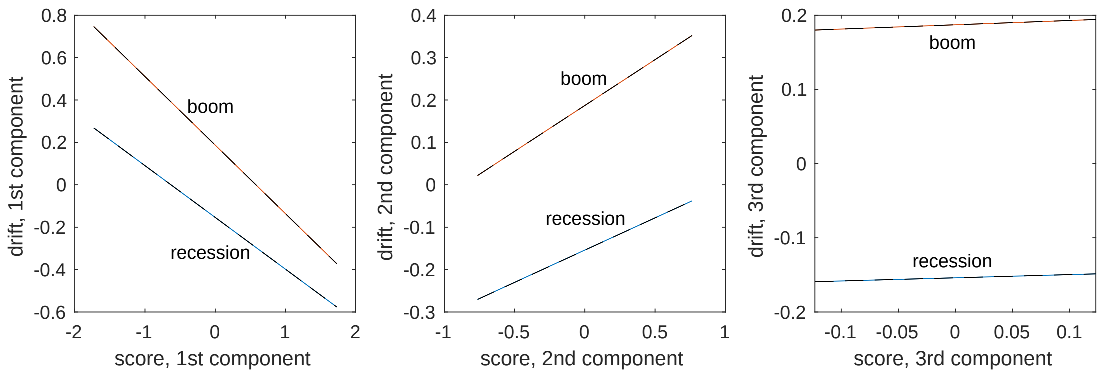

Figure 6: Linearity of component dynamics.

us information about the domain and concentration of wealth distributions.

Computation has been performed on the space of about 4 standard deviations from the mean in the direction of both principal components. We see from the graph that this is a generous range, especially for the first component, although it excludes some distributions which are theoretically possible with a large score on the second component.

Next, the fundamental premise behind the proposed method of computation is that the minor components, the "ripples" in the system, have essentially linear effects on system dynamics. It turns out that the equilibrium of this model is fairly linear in all components. Figure 6 shows the result of computation with three major components (and over 10 minor components). The three panels show the drift of each major component in each state of the economy, as a function of deviations from the mean along the direction of that component. The colored curve shows the true drift produced by computation, and the dashed straight lines are linear. Visually the dynamics is indistinguishable from linear (and numerically, each curve goes away from the straight line by an amount of less than 0.0001).

For comparison, in Figure 7 the same computational method is applied to a different but related model: a pure endowment model with money. Individuals receive exogenously given endowments in employment and unemployment states, and trade fiat money in fixed nominal supply to self insure against idiosyncratic shocks. Money is an unproductive asset with endogenous value that is used purely as an equilibrium store of value. The details of the model are in Appendix A. In this model, the return on money is a lot more volatile than the return on capital in our main model. In our main model individuals consume capital, hence it is easy to move consumption between periods. In the monetary example, consumption goods are non-storable and the real return on money fluctuates in the range of up to 10% per period. These rigidities give rise to come curvature in the major components, relative to our main model.

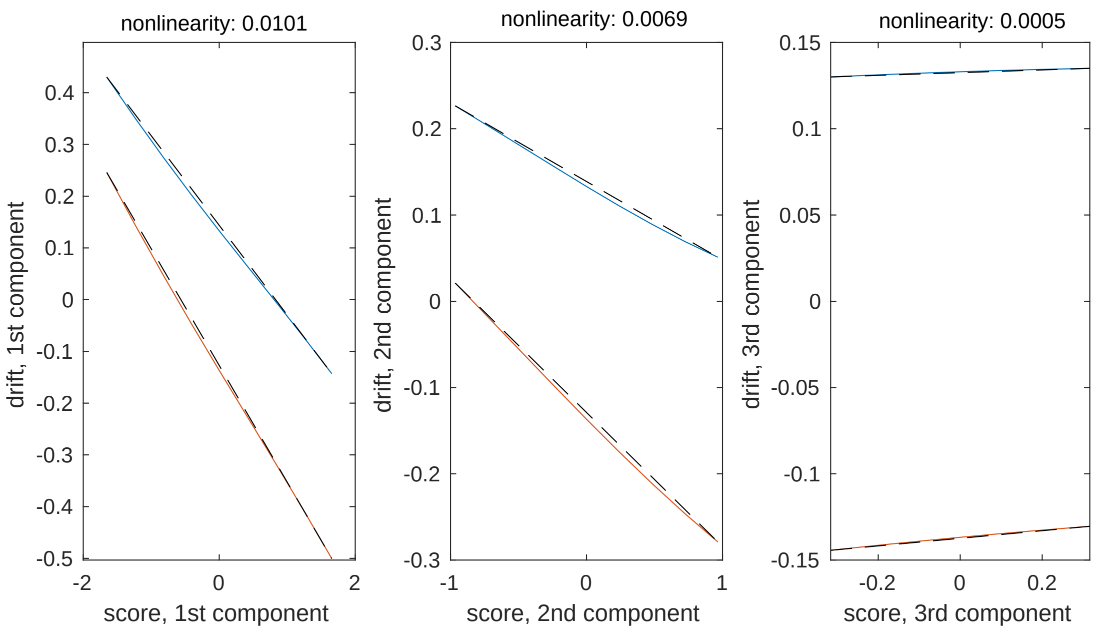

Figure 7: Nonlinearities in a different model: pure endowment economy with money.

Figure 7 shows the drift of each of the first three component scores in both states, as a function of deviation of that score from the mean. The computation is based on three major components and over 10 minor components, as in Figure 6. Here, we clearly see nonlinearities in the first two components. The effect of the third component is, at least visually, linear. The figure also displays a nonlinearity index for each component, defined as the distance between the true drift and its dashed linear approximation. The nonlinearity index depends on the scale of horizontal axis, which reflects the equilibrium variation of each principal component.

In all our examples, the assumption of linearity in minor principal components is adequately satisfied in the range of values typically observed along the equilibrium path.

5 Conclusions

This paper proposes a computational method for problems with many state variables. The method gets around the curse of dimensionality to some extent: it lets us incorporate a large number of dimensions with nontrivial interactions. The tractability of the proposed computational method is based on the assumption that for a system that has been running for a long time, if we know the values of certain key moments, we can guess the values of the remaining moments with high precision. That is, the "likely" portion of the state space is wide in some directions but thin in other directions. This property is intuitive and has theoretical foundations; we also confirm it numerically in all our examples. In addition, the proposed method also relies on the assumption of differentiability in the state space, so that linear approximations are adequate for changes of the state vector in directions designated as minor.

One may wonder how our method compares to, for example, the method of KS that accounts for the effect of the infinite-dimensional state via specific moments, such as average capital holdings. In principle, there are advantages to using moments with a clear economic interpretation, such as average wealth or the unemployment rate. If those moments have appropriate correlation structure with the major principal components, numerically it may not really matter which moments we use.[6] It is important that, given the values of the major moments we choose, we are fairly certain about the value of the state vector or its effect on the system, if we use our first procedure or the method of KS that takes into account only a small set of moments, or at least that the residual uncertainty is small, if we account for it linearly via our second procedure. In general, the higher the dimensionality of the state space, the lower the correlation between a randomly chosen moment and any specific moment. Hence, the systematic procedure we propose of choosing the key moments based on principal components should prove of great value, especially for models where one does not have a strong theoretical prior about which moments are important. Of course, it should prove tremendously useful to be able to account for limitless other minor moments by approximating their effect on the system as linear.

6 Appendix A.

(TO BE COMPLETED)

[6]However, from the point of view of coding, it is certainly convenient to use orthogonal moments, such as moments given by principal component analysis.

Bibliography.

Achdou, Y., J. Han, J-M. Lasry, P-L. Lions and B. Moll (2022) ”Income and wealth distribution in macroeconomics: A continuous-time approach,” *The Review of Economic Studies* 89 (1), 45-86.

Aiyagari, S. R. (1994), “Uninsured Idiosyncratic Risk and Aggregate Saving,” *The Quarterly Journal of Economics*, 109, 659–84.

Bewley, T. (1986), “Stationary Monetary Equilibrium with a Continuum of Independently Fluctuating Consumers,” in Hildenbrand, W. and Mas-Collel, A. (eds) *Contributions to Mathematical Economics in Honor of Gerard Debreu* (Amsterdam: North-Holland).

Huggett, M. (1993), “The Risk-free Rate in Heterogeneous-agent Incomplete-insurance Economies,” *Journal of Economic Dynamics and Control*, 17, 953–969.

Krusell, P. and A. A. Smith (1998), “Income and Wealth Heterogeneity in the Macroeconomy,” *Journal of Political Economy*, 106, 867–896.