Escape Dynamics: A Continuous Time Approach

Dmitri Kolyuzhnov[1] **Anna Bogomolova**[1] **Sergey Slobodyan**[1]

[1]CERGE-EI, Prague

Abstract

We use a continuous-time approximation approach to analyze dynamics of a model where government adaptively learns the Phillips curve while running monetary policy (Phelps problem). This approach is based on approximating the discrete-time dynamics with learning by a limiting continuous-time diffusion and subsequent characterization of the escape dynamics (recurrent excursions from the neighborhood of equilibrium) for this limit process. We characterize escape dynamics by analytically deriving dominant escape path. We discuss the average behavior of the learning process (the mean dynamics) and its relationship to the escape dynamics. Finally, we discuss the appropriateness of our approximation for the parameterizations of the learning which are commonly used in the literature.

Keywords: Large deviations, Stochastic Approximation, Escape Dynamics, Adaptive Learning

1. Introduction

Recent macroeconomic literature attempts to explain the fluctuations in macroeconomic data as a result of the process of endogenous equilibrium search of boundedly rational agents. One of the methods of endogenous equilibrium search that has received wide application in macroeconomic literature is the adaptive learning approach summarized in a book by Evans and Honkapohja (2001) (hereafter, EH). Adaptive learning of agents generates the fluctuating behavior of macroeconomic indicators, such as inflation and unemployment. These fluctuations can be analytically characterized using two forces that exist in such a dynamic system: mean dynamics, which moves the system to the point of weak convergence, and escape dynamics, which moves the system away from it.

Major efforts of researchers were made towards analytical characterization of the second force, escape dynamics. The approach that received a lot of attention in the literature (Cho, Williams and Sargent (2002) (hereafter, CWS), Williams (2001)) is based on the analytical derivation of escape dynamics for the original discrete-time learning process. This approach has theoretical and practical problems. The theoretical problem is that for cases where a shock for the state process is unbounded (e.g. Gaussian) theoretical results allowing full description of escape dynamics are not available. The theory in this case does not allow one to say what is the most probable direction of deviations from the point of convergence and what is the expected time until such an escape.

The practical problem is that the analytical solution of escape dynamic characteristics for the discrete-time process proposed by Williams (2001) implies numerical solution of a system of non-linear differential equations with functions given only numerically. With many lags in the structure of the state process this problem becomes intractable.

The continuous time approach proposed here partially resolves these problems. This approach is first to use a continuous-time approximation of the recursive dynamic system under consideration and then to apply the analytical tools developed by Freidlin and Wentzell (1998) (hereafter, FW) to analyze analytically escape dynamics. As diffusion is a natural approximation for a difference equation with Gaussian noise and FW developed the theory of large deviations for diffusions, the problem of insufficient theoretical results is removed. The second problem is partially alleviated because the diffusion, derived by approximation around the stationary point, is linear. Similar approach was used by Kasa (2004) who considered a simple one-dimensional model.

2. The model

The proposed approach is applied to the Phelps model of the government controlling inflation using the approximate Phillips curve considered originally by CWS. The economy consists of the government and the private sector. Government uses monetary policy instrument x_n to control inflation rate π_n attempting to minimize losses from inflation and unemployment $_n$. It believes (in general, incorrectly) that there exists an exploitable tradeoff between π_n and $_n$, or the Phillips curve, Eq. (1d). The true Phillips curve (1a) is subject to random shifts and contains this tradeoff only for unexpected inflation shocks. The private sector possesses rational

---

*This paper is an abridged and compressed version of Kolyuzhnov *et l* (forthcoming).

expectations $x_n = x_n$ about the inflation rate, and thus unexpected inflation shocks come only from monetary policy errors.

$$_n = \quad -\theta(\pi_n - x_n) + \sigma_1 \quad _{1n}, \quad > 0, \theta > 0, \tag{1a}$$

$$\pi_n = x_n + \sigma_2 \quad _{2n}, \tag{1b}$$

$$x_n = x_n, \tag{1c}$$

$$_n = \gamma_1 \pi_n + \gamma^T{}_{-1} X_{n-1} + \eta_n. \tag{1d}$$

Vector $\gamma = (\gamma_1, \gamma^T_{-1})^T$ represents governments beliefs about the Phillips curve. $_{1n}$ and $_{2n}$ are two uncorrelated Gaussian shocks with zero mean and unit variance. η_n is the Phillips curve shock as perceived by the government, believed to be a white noise uncorrelated with regressors π_n and X_n, where $X_{n-1} = (\quad _{n-1}, \quad _{n-2}, \pi_{n-1}, \pi_n, 1)$

The governments monetary policy rule is determined as follows. Given beliefs γ, it solves $\min_{\{x_n\}_{n=0}^{\infty}} \sum_{n=0}^{\infty} \beta^n (\quad ^2_n + \pi_n^2)$ subject to (1b) and (1d). This Linear-Quadratic problem produces a linear monetary policy rule $x_n = \quad (\gamma)^T X_{n-1}$.

CWS identify three beliefs consistent with the model. Belief 1, $\gamma = (-\theta, 0, 0, 0, 0, \quad (1+\theta^2))^T$, generates policy function $x_n = \theta$. In a model where the government knows the true Phillips curve (1a), this is the Nash, or discretionary equilibrium. Beliefs 2 of the form $\gamma = (0,0,0,0,0, \quad ^*)^T$ lead to $x_n = 0$ and zero average inflation for any : Ramsey, or optimal time-inconsistent equilibrium. Finally, Beliefs 3 where $\gamma_1 + \gamma_4 + \gamma_5$ (sum of coefficients on current and lagged inflation is zero) asymptotically lead to $x_n = 0$: this is an induction hypothesis belief, see Sargent (1999).

In the model with learning, equilibrium is defined as a vector of beliefs at which governments assumptions about orthogonality of η_n to the space of regressors are consistent with observations $\left[\eta_n (\pi_n, X_{n-1})^T\right] = 0.$

CWS call this point a self-confirming equilibrium, or SCE. The only SCE in the model is Belief 1, see Williams (2001).

Learning of the government is introduced as the recursive least squares algorithm with a constant gain[1]:

$$\gamma_{n+1} = \gamma_n + \varepsilon R_n^{-1} (\quad _n - \gamma_{n_1} \pi_n - \gamma_{n_{-1}} X_{n-1}) \cdot [\pi_n X_{n-1}]^T \tag{2a}$$

$$R_{n+1} = R_n + \varepsilon \cdot ([\pi_n X_{n-1}]^T [\pi_n X_{n-1}] - R_n) \quad . \tag{2b}$$

[1] Constant gain learning is used in learning models to reflect discounting of the past experience.

3. Continuous Time Approach

The economic system described above can be written as a stochastic recursive algorithm (SRA)

$$\theta^{\varepsilon}_{n+1} = \theta^{\varepsilon}_n + \varepsilon \quad (\theta^{\varepsilon}_n, \xi_n), \tag{3a}$$

$$\xi_{n+1} = A(\gamma_n)\xi_n + \quad _{n+1}, \tag{3b}$$

where ξ_n is a vector of states, θ_n is composed of a vector of beliefs and elements of covariance matrix. Its properties can be derived using the mathematical results compiled in EH. This process can be approximated by a diffusion

$$\dot{\varphi}(t) = A\varphi(t) + \sqrt{\varepsilon} \quad \dot{} \quad , \tag{4a}$$

where $A = p'(\bar{\theta})$, $\quad = \Sigma(\bar{\theta})^{1/2}$, $\varphi(t) = \theta(t) - \bar{\theta}$, $p(\theta) = \quad [\quad (\theta, \xi)]$, $p(\theta) = \quad [\quad (\theta, \xi)]$, $\Sigma^{ij}(\theta) = \sum_{k=0}^{\infty} \mathrm{cov}[(\quad _i(\theta, X_k(\theta)))(\quad _j(\theta, X_k(\theta)))]$, $\bar{\theta}$ is both the SCE in the CWS model, see Williams (2001, Theorem 3.2) and .a stable point of the ODE

$$\partial\theta / \partial\tau = p(\theta(\tau)) \tag{4b}$$

Action functional for this diffusion is defined as

$$_{0T}(\varphi) = \inf \frac{1}{2} \int_0^T | \quad \dot{} (t)|^2 \, dt \tag{5a}$$

$$\text{s.t. } \dot{\varphi}(t) = A\varphi(t) + \quad \dot{}(t), \ \varphi(0) = 0 \,. \tag{5b}$$

see Dembo and Zeitouni (1998, p. 214).

Minimizing this functional is thus a standard linear control theory problem, complicated by the fact that system (A) is not controllable. The solution to this problem is well known: $(\quad _{1,} \quad) = (1/2) \quad ^T_{1,} \quad . \quad \bar{}^{-1} \cdot \quad _{1,}$, where $\bar{}$ is Gramian of the reachable subsystem, and $T_1 \quad _{1,} \quad \in \partial$, where T is the basis of the reachable subspace. For our choice of boundary ∂ see the next section.

4. Results

For fixed γ, expected value of error term η_n (in stationary state) is given by $[\eta_n] = \quad -(\gamma_1 + \gamma_4 + \gamma_5)\pi \quad -(\gamma_2 + \gamma_3) \quad -\gamma_6$, where π is some stationary inflation rate. This means that from the governments point of view, linear combinations $\gamma_1 + \gamma_4 + \gamma_5$ and $(\gamma_2 + \gamma_3) + \gamma_6$ rather than the whole vector γ matter. As it is exactly perceived error η_n which matters for the adjustment of θ in (2), one presumes that coordinates $(\gamma_1, \gamma_2) \quad \gamma_1 + \gamma_4 + \gamma_5 \quad (\gamma_2 + \gamma_3) + \gamma_6$ are useful in thinking about the model. The simulations support the presumption: Fig.1 plots a whole simulation run started at SCE with ε=0.001, including an escape towards Belief 3 of CWS and very low inflation[2].

[2] We use $\sigma_1 = \sigma_2 = 0.3 \quad = 5, \theta = 1 \ \beta = 0.98$ as in CWS. All the figures are for simulations with $\varepsilon = 0.001$.

In the (γ_1,γ_2) plane all simulation points are very close to a straight line. Very similar graphs are obtained in all 1000 runs, which strongly suggests that we could use coordinates (γ_1,γ_2) to effectively reduce the dimensionality of the system.

Consider equation (2a). At the SCE, the largest eigenvalue of $\overline{R}^{-1}$ is λ_1=3083.8 and the next largest λ_2=29.1, less that 1% of λ_1.Therefore, if one writes (γ_n,ξ_n) as a linear combination of eigenvectors of $\overline{R}^{-1}$, then the projection of (γ_n,ξ_n) onto the $_1$, the eigenvector corresponding to λ_1, is magnified 100 times as strongly as the projection onto $_2$. The dynamics described by $\overline{R}^{-1}$ (γ_n,ξ_n) is, thus, almost 1-dimensional.

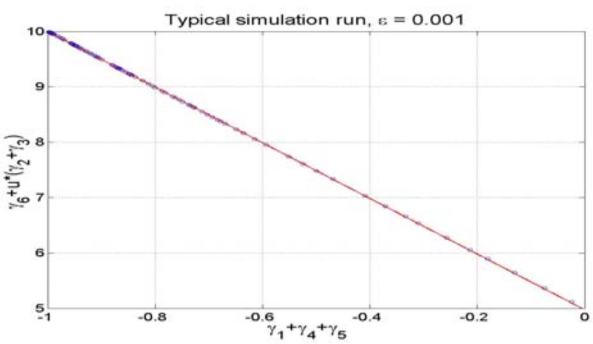

Fig.1.

In coordinates (γ_1,γ_2), $_1$ is approximately proportional to (1, -5). Fig. 1 shows that indeed, all the simulation run points are aligned along this vector.

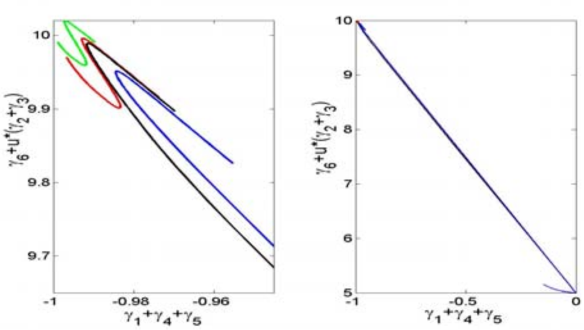

Fig. 2

Several trajectories of the mean dynamics of the model, given by (4b), are presented in Fig. 2 in (γ_1,γ_2) coordinates. The region around the SCE where the mean dynamics points back towards it is very small; if initial deviation from the SCE is relatively large, the mean dynamics trajectory treks towards Belief 3, or induction hypothesis plane, where γ_1 0. After spending some time in the neighborhood of γ_1=0, the trajectory slowly returns back to the SCE. Right panel of Fig.2 tracks several mean dynamics trajectories as they travel to the induction hypothesis plane. The paths are almost indistinguishable at this scale: away from the immediate neighborhood of the SCE, mean dynamics trajectories are rapidly converging to the line connecting the SCE with (0,5).

To understand the relation between the mean and stochastic parts of the dynamics, consider Fig.3. It plots a ratio of relative magnitude of the stochastic dynamics, given by $R_n^{-1}\{ (\gamma_{nn},\xi_n) - [(\gamma_{nn},\xi_n)]\}$, and of mean dynamics $R_n^{-1}[(\gamma_{nn},\xi_n)]$, evaluated at different points along the eigenvector $_1$. For large deviations from the SCE the mean dynamics dominates the stochastic part. In the small region around SCE where the mean dynamics points back towards it, stochastic dynamics is on average hundreds and thousands times larger than the mean dynamics.

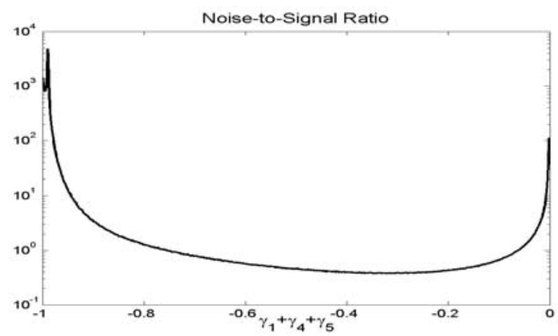

Fig. 3

Using the theoretical derivations and the knowledge of the system obtained through simulations, we think about escapes in the following way. Consider a small neighborhood of the SCE. After the trajectory crosses the boundary ∂ , we assume that the stochastic dynamics does not play any role, and the model s behavior is determined exclusively by its mean dynamics (4b). A process of excursion towards γ_1=0 is, therefore, split into two parts: first, stochastic escape from , and second, almost deterministic movement to γ_1=0 and back to the SCE.

We select the set and its boundary ∂ as follows. The first way is to minimize the action functional on the boundary of a cylinder: a sphere in six-dimensional γ space, and no binding restrictions in 21-dimensional space of components of *R*. This approach is similar to the road taken by CWS. No matter what is the cylinder s radius, this calculation predicts that the escape from the SCE has to occur in the approximate direction (1,-6.5) in (γ_1,γ_2) coordinates. When one selects the cylinder s radius so that it intersects ${}_1$ at γ_1=-0.985, distance between the mean of observed escape points and the theoretically derived escape point equals 0.16, or 135% of the distance from the SCE to the theoretical escape point. We select this cylinder radius because mean dynamics trajectories which intersect vector ${}_1$ beyond γ_1=-0.985 first travel to the induction hypothesis plane, while those intersecting ${}_1$ before this point travel directly to SCE. Therefore, if we observe escape to this point, we can be reasonably sure that further evolution takes the system to induction hypothesis plane before coming back to SCE.

Fig. 3 presents a histogram of observed escape directions tangents. As is easy to see that the theoretical prediction of approximately 6.5 is way off the mode of the empirical distribution, which is -5.

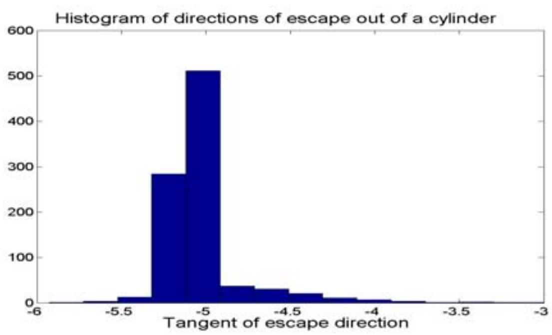

Fig.4

To address these discrepancies we use the information about low signal-to-noise ratio near the SCE mentioned above. Fig.3 plots logarithm of this ratio (averaged over 4000 realizations) for points along ${}_1$. In this situation one expects (4a) to be a good approximation of (2) only if the system stays in the neighborhood of every point θ long enough to allow the average of R^{-1} (γ,ξ_n) to approach R^{-1} $[$ $(\gamma,\xi_n)]$. With signal-to-noise ratios from 0.001 to 0.1, this means hundreds and thousands iterations near every point θ. However, for the values of ε used in CWS and commonly applied elsewhere in the literature (ε=0.001 0.01), the expected escape time is measured in hundreds of iterations: this is the time spent by the system near all points around the SCE.

Given such small signal-to-noise ratio, one could simply disregard the mean dynamics (set $p'(\theta)$=0) and repeat minimization of the action functional. This (second) way provides a much better agreement between the theory and the simulations results: for the same radius of the cylinder, mean observed escape point lies at a distance of 0.005 from the theoretically predicted point, which is only 4.4% of the distance between the SCE and the theoretical escape point. Most runs escape along the direction (1,-5), which is the theoretically predicted one for this approach.

5. Conclusion

Using the theoretical derivations and the knowledge of the system obtained through simulations, we derive the escape dynamics characteristics in two ways. The first way is to minimize the action functional on the spherical set of deviations of beliefs (this is the road taken by CWS). This calculation, though mathematically appealing, does not replicate the behavior in simulations. The second way is based on the diffusion approximation that disregards mean dynamics. It provides a much better agreement between the theoretical and the simulations results.

We suggest two changes to help the proposed approach work better for the model considered: to set lower mean unemployment rate, in order to construct a more balanced covariance matrix and to use learning specifications resulting in lower dimensional problem.

The proposed approach can be used to analyze escape dynamics in more complicated models, where it is not possible to derive analytical characteristics of escape dynamics in discrete time. For example, the model with a dynamic Phillips curve can be studied as a possible extension of the approach proposed here. This is the focus of our current research.

6. References

[1] I.-K. Cho, N. Williams and T. J. Sargent, Escaping Nash Inflation , *Re iew n i St die* , 69, pp. 1-40, 2002.

[2] A. Dembo and O. Zeitouni, *r e e i ti n te ni e nd Appli ti n* , Second Edition (New York: Springer-Verlag), 1998.

[3] G. W. Evans and S. Hohkapohja, *e rnin nd xpe t ti n in M r e n i* (Princeton University Press), 2001.

[4] M. I.Freidlin and A. D. Wentzell, *R nd Pert r ti n yn i l Sy te* , Second Edition (New York: Springer-Verlag), 1998.

[5] K. Kasa, Learning, Large Deviations, And Recurrent Currency Crises , *ntern ti n l n i Re iew*, 45, pp. 141-173, 2004.

[6] D. Kolyuzhnov, A. Bogomolova, S. Slobodyan, Escape Dynamics: A Continuous Time Approximation , *rt in*

[7] N. Williams, *pe yn i in e rnin M del* (Ph.D. thesis, U. of Chicago), 2001.

[8] T.J. Sargent. *T e C n e t A eri n n lti n* (Princeton University Press), 1999.