On learnability of E–stable equilibria

Sergey Slobodyan and Atanas Christev

This version: Feb 15, 2006

Authors Affiliation: CERGE-EI, Prague, Czech Republic and Heriot Watt University, Edinburgh, UK

Very Preliminary. Please do not quote!

*Abstract.*While under recursive least squares learning the dynamics of the economy converges to rational expectations equilibria (REE) which are E–stable, some recent examples propose that E–stability is not a sufficient condition for learnability. In this paper, we provide some further evidence on the conditions under which E–stability of a particular equilibrium might fail to imply its stochastic gradient (SG) or generalized SG learnability. We also claim that the requirement on the speed of convergence of the learning process imposed by [4] also implies that E–stable equilibria are likely to be GSG learnable. We show this in a simple "New Keneysian" model of optimal monetary policy design in which the stability of REE under SG learning. In this case, the paper gives the conditions which are necessary for reversal of learnability.

JEL Classification:C620, D830, E170

Keywords: Adaptive learning, E–stability, stochastic gradient, learnability of REE and D-stability

1 Introduction

The concepts of adaptive learning and expectational stability (E–stability) in macroeconomics have received much deserved attention recently, see, for example, [2] for an extensive discussion. The authors provide the methodology and list the conditions under which recursive learning dynamics obtain stability. Rational expectations equilibria which E–stable (REE) may be attainable and can be learned.[1] Admittedly, however, their analysis also points to the lack of general results under a wide variety of learning rules for which E–stability holds,i.e., not all REE are indeed learnable. If recursive least squares (RLS) learning is used by economic agents to update their expectations of the future (or learn adaptively), then it is shown that the concept of *E–stability* plays a crucial role: the RLS learning process converges only to rational expectations

[1] The possible convergence of learning processes and the E–stability criterion of REE dates back to DeCanio (1979) and Evans (1985).

equilibria which are E–stable. Equilibria which are stable under a particular form of adaptive learning are also called *learnable*.

At the same time, earlier work by Bullard (1994) suggests that RLS learning, under certain conditions, may cycle about REE and never obtain them. In addition, [1] and [6] explain that E–stability may not be a sufficient condition for learnability. [1] show that an alternative learning mechanism, namely, stochastic gradient (SG) converges to REE but under different conditions than RLS learning. In a recent paper, furthermore, [5] provides some examples of E–stable equilibria which are not learnable under SG learning. [3] discuss the sufficient conditions under which E–stable equilibria are learnable under SG and Generalized Stochastic Gradient (GSG) learning. In this paper, we provide some further evidence and discussion on the conditions under which E–stability (equivalent to learnability under RLS learning) of a particular equilibrium might fail to imply its SG or GSG learnability. We also claim that the requirement on the speed of convergence of the learning process imposed by [4] also implies that E–stable equilibria are likely to be GSG learnable.

The next section develops a New Keynesian model of optimal monetary policy and examines its REE under SG learning. Section 3 discusses the condition under which these equilibria are E-stable and consequently learnable. Section 4 gives and analyzes a geometric example of the disappearance of D-stability and the conditions under which we observe learnability reversal. Section 5 relates these results to the speed of convergence of SG learning. The last sections provides some concluding remarks.

2 The Model

To illustrate the main result of this paper we make use of a standard theoretical framework on monetary policy design developed in Clarida, Gali and Gertler (1999) and Woodford (2003). This is a dynamic general equilibrium model with staggered sticky price setting and money. In reduced form the model of the private sector yields the following two equations:

$$x_t = -\varphi(i_t - E_t\pi_{t+1}) + E_t x_{t+1} + g_t \tag{1}$$

$$\pi_t = \alpha x_t + \beta E_t \pi_{t+1} + u_t. \tag{2}$$

The intertemporal "IS" curve, equation (1), expresses the output gap, x_t, as a function of the nominal interest rate, i_t, and is influenced by the expectations of future inflation $E_t\pi_{t+1}$, the future output gap $E_t x_{t+1}$ and a demand shock, g_t. Equation (2) is modeled as an expectations-augmented Phillips curve, where current inflation, π_t depends on the current output gap, expected future inflation, $E_t\pi_{t+1}$ and a supply shock, u_t. The model is completely determined with the addition of a monetary policy rule set by the monetary authority. Following Ferrero (2004), we specify a general set of *expectations based* reaction functions that subsume previous results in the literature, e.g. Evans and Honkapohja (2003). Thus the nominal interest rate follows:

$$i_t = \gamma + \gamma_\pi E_t \pi_{t+1} + \gamma_x E_t x_{t+1} + \gamma_g g_t. \tag{3}$$

As a result, the economy evolves according to:

$$Y_t = Q + F E_t Y_{t+1} + S\psi_t, \tag{4}$$

We study this model under SG learning and show the tenuousness of achieving E–stability of REE.

The rational expectations equilibrium (REE) of this model take the form:

$$Y_t = \overline{\mathit{\Phi}} + S\psi_t. \tag{5}$$

Under learning, the expectations are given by:

$$E_t Y_{t+1} = \mathit{\Phi}_t. \tag{6}$$

TO BE COMPLETED

3 E–Stability and Learnability

Denote the rational expectations equilibrium of the above model by $\overline{\mathit{\Phi}}$. This equilibrium is called E–stable if $\overline{\mathit{\Phi}}$ is stable under the dynamics defined by the following ordinary differential equation:

$$\frac{d\mathit{\Phi}}{d\tau} = T(\mathit{\Phi}) - \mathit{\Phi}. \tag{7}$$

$\overline{\mathit{\Phi}}$ is stationary point of (7). It is stable *iff* the Jacobian of (7) evaluated at $\overline{\mathit{\Phi}}$, $J = DT(\mathit{\Phi})\,|_{\mathit{\Phi}=\overline{\mathit{\Phi}}} - I$, has only eigenvalues with negative real parts.

If, instead of using RLS as an adaptive learning algorithm, one relies on SG learning, convergence of the learning process is governed instead by the following differential equation,

$$\frac{d\mathit{\Phi}}{d\tau} = M(\mathit{\Phi}) \cdot (T(\mathit{\Phi}) - \mathit{\Phi})\,. \tag{8}$$

The equilibrium $\overline{\Phi}$ is still a stationary point of (8). It is learnable iff $\overline{\Phi}$ is stable under the flow (8), which means that all eigenvalues of $M(\overline{\Phi}) \cdot J$ have negative real parts, see [1] for a proof. $M(\mathit{\Phi})$ is a matrix of second moments of state variables; it is symmetric and positively definite. Additionally, one could consider Generalized SG learning (GSG), in which case learnability is equivalent to negativity of all eigenvalues of the matrix

$$M \cdot J, \tag{9}$$

where M is arbitrary positive definite matrix.

The problem of a correspondence between E–stability and learnability under GSG learning is, therefore, equivalent to the following linear algebraic problem: given a matrix J with all the eigenvalues to the left of imaginary axis, could one guarantee that no eigenvalue of $M \cdot J$ becomes positive? This problem is well known in economics and is called a D–stability problem, see Arrow (1974) and Johnson (1974). Many sufficient conditions for D–stability are known, but they are usually hard to interpret from economic point of view. We will provide a geometric interpretation of a case when D–stability does not obtain in 2 dimensions.

4 Geometric Interpretation of Disappearance of Stability

Suppose that the 2×2 matrix J has only eigenvalues with negative real parts. The eigenvalue problem can be written as

$$J \cdot V = V \cdot \Lambda, \tag{10}$$

where V is the matrix whose columns are eigenvectors of J and Λ is diagonal with corresponding eigenvalues λ_i on the main diagonal. For expositional simplicity we assume that both λ_i are real and negative. In case eigenvalues are linearly independent (which is a generic case for non–singular matrices), matrix J could be diagonalized as $J = V \Lambda V^{-1}$. Learnability of the equilibrium under GSG adaptive learning is determined by eigenvalues of $M \cdot J$, where M is symmetric and positively definite and thus could be written as $M = PDP^T$.[2] We are thus interested in the following eigenvalue problem:

$$PDP^T \cdot V\Lambda V^{-1} \cdot \widetilde{V} = \widetilde{V} \cdot \widetilde{\Lambda}, \tag{11}$$

where $\widetilde{V}$ consists of eigenvectors of $M \cdot J$ and $\widetilde{\Lambda}$ is a diagonal matrix with eigenvalues of $M \cdot J$ as entries.

Pre–multiply (11) by P^{-1} and define $\overline{V} = P^{-1}\widetilde{V}$ to get

$$D \cdot P^T V \Lambda V^{-1} P \cdot \overline{V} = D\widetilde{J} \cdot \overline{V} = \overline{V} \cdot \widetilde{\Lambda}. \tag{12}$$

It is obvious that the matrix $\widetilde{J} = P^T V \Lambda V^{-1} P$ has the same eigenvalues as J, namely the values on the main diagonal of Λ. Geometrically, if J represents a linear map in a 2–dimensional space, then $\widetilde{J}$ represents the same map in new coordinates. These new coordinates are given by two orthogonal eigenvalues of M. In the new coordinate system, any vector v is transformed into $P^{-1}v$.

To fix notation, let us order d_1 and d_2, the eigenvalues of M, so that $\frac{d_1}{d_2} < 1$. Eigenvalues of J and $\widetilde{J}$ are $-\lambda_1$ and $-\lambda_2$, ordered so that $\frac{|\lambda_1|}{|\lambda_2|} < 1$. Denote the eigenvectors of $\widetilde{J}$ corresponding to $-\lambda_1$ and $-\lambda_2$ are v_1 and v_2. We are

[2] Eigenvalues of a symmetric matrix are orthogonal, and so $P^{-1} = P^T$.

interested in a case when $D \cdot \widetilde{J}$ has positive eigenvalue $\widetilde{\lambda}$.[3] The eigenvector corresponding to $\widetilde{\lambda}$, let call it $\widetilde{v}$, can be represented as a weighted average of v_1 and v_2, $\widetilde{v} = \alpha v_1 + \beta v_2$. Without loss of generality, set α equal to 1.[4] Write

$$\widetilde{\lambda}\widetilde{v} = D\widetilde{J}\widetilde{v} = D \cdot \widetilde{J} \cdot (v_1 + \beta v_2) = D \cdot \left(\widetilde{J}v_1 + \beta \widetilde{J} v_2\right) = \tag{13}$$

$$D \cdot (-\lambda_1 v_1 - \beta \lambda_2 v_1) \,. \tag{14}$$

This vector equation in the coordinate form could be written as

$$-\lambda_1 \begin{bmatrix} d_1 v_{11} \\ d_2 v_{21} \end{bmatrix} - \beta\lambda_2 \begin{bmatrix} d_1 v_{12} \\ d_2 v_{22} \end{bmatrix} = \widetilde{\lambda} \begin{bmatrix} v_{11} + \beta v_{12} \\ v_{21} + \beta v_{22} \end{bmatrix}, \tag{15}$$

or

$$v_{11}\left(\widetilde{\lambda} + \lambda_1 d_1\right) = -\beta v_{12}\left(\widetilde{\lambda} + \lambda_2 d_1\right), \tag{16a}$$

$$v_{21}\left(\widetilde{\lambda} + \lambda_1 d_2\right) = -\beta v_{22}\left(\widetilde{\lambda} + \lambda_2 d_2\right). \tag{16b}$$

Dividing (16a) by (16b), get

$$\frac{v_{11}}{v_{21}} \frac{\widetilde{\lambda} + \lambda_1 d_1}{\widetilde{\lambda} + \lambda_1 d_2} = \frac{v_{12}}{v_{22}} \frac{\widetilde{\lambda} + \lambda_2 d_1}{\widetilde{\lambda} + \lambda_2 d_2}. \tag{17}$$

For the above equation to have a solution $\widetilde{\lambda} > 0$, one needs $sign(\frac{v_{11}}{v_{21}}) = sign(\frac{v_{12}}{v_{22}})$. In other words, both eigenvectors of J after rotation into the coordinates defined by eigenvectors of M should be located in the same quadrant of the plane. In this case, $T = \frac{v_{22}}{v_{21}} \frac{v_{11}}{v_{12}} > 0$. This is the first condition necessary to generate learnability reversal.[5]

---

[3] As $\det(D\widetilde{J}) = \det(D) * \det(\widetilde{J})$, there are either two or zero eigenvalues with positive real part.

[4] If $\widetilde{v}$ is eigenvector of $M\widetilde{J}$, then $\alpha\widetilde{v}$ is also an eigenvector for any real α.

[5] Suppose v_1 and v_2 are two eigenvectors of J, and p_1 and p_2 are two orthogonal eigenvectors of M. In the new coordinates the eigenvectors of MJ are given as

$$\begin{bmatrix} p_1^T \\ p_2^T \end{bmatrix} \begin{bmatrix} v_1 & v_2 \end{bmatrix} = \begin{bmatrix} p_1^T v_1 & p_1^T v_2 \\ p_2^T v_1 & p_2^T v_2 \end{bmatrix}.$$

If we change the direction of p_1, first coordinates of *both* eigenvectors of MJ change sign. Possible moves are from the quadrant I to II and back, and from the quadrant III to IV and back. If one changes the direction of p_1, both eigenvectors of MJ move $I \leftrightarrow IV$ or $II \leftrightarrow III$. Finally, changing the direction of both p_1 and p_2 makes both eigenvectors move in $I \leftrightarrow III$ or $II \leftrightarrow IV$ direction. If instead one switches the direction of v_1 or v_2 to the opposite, the corresponding eigenvector of MJ moves between I and III or II and IV.

Thus, there are just two possibilities regarding the mutual location of the two eigenvalues of MJ: they either could both be put into the first quadrant of the plane by some combination of directional changes described above, or one of them is in the first quadrant and the other is in the second.

Rewrite (17) as

$$T\left(\widetilde{\lambda}+\lambda_2 d_2\right)\left(\widetilde{\lambda}+\lambda_1 d_1\right)=\left(\widetilde{\lambda}+\lambda_2 d_1\right)\left(\widetilde{\lambda}+\lambda_1 d_2\right), \tag{18}$$

$$\widetilde{\lambda}^2+\widetilde{\lambda}\frac{\lambda_2 d_1+\lambda_1 d_2-T\left(\lambda_2 d_2+\lambda_1 d_1\right)}{1-T}+\lambda_1\lambda_{21}d_1 d_2=0. \tag{19}$$

Note that T cannot equal one; in this case, two eigenvectors of J are collinear, which is not generic situation.[6]

The quadratic equation $\widetilde{\lambda}^2+b\widetilde{\lambda}+c=0$, where $c>0$, has at least one solution with positive real part only if $b<0$. Define $T=1-\epsilon$, $\epsilon>0$ and rewrite the condition $b<0$ as

$$1+\frac{\lambda_2}{\lambda_1}\frac{d_2}{d_1}+\frac{\left(\frac{\lambda_2}{\lambda_1}-1\right)\left(1-\frac{d_2}{d_1}\right)}{\epsilon}<0, \tag{20}$$

$$\epsilon\cdot(1+xy)<(1-x)(1-y), \tag{21}$$

where we defined $x=\frac{\lambda_2}{\lambda_1}$ and $y=\frac{d_2}{d_1}$. Finally, resolving the above inequality with respect to y, one gets

$$y>\frac{x-(1-\epsilon)}{(1-\epsilon)\,x-1},\ x>\frac{1}{1-\epsilon}, \tag{22}$$

$$y<\frac{x-(1-\epsilon)}{(1-\epsilon)\,x-1},\ x<\frac{1}{1-\epsilon}. \tag{23}$$

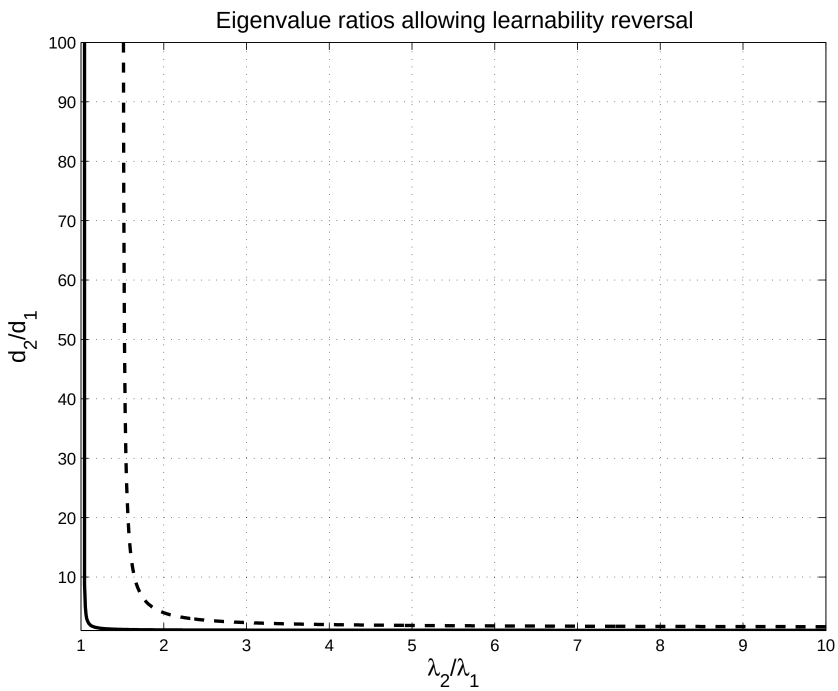
The function $y=\frac{x-(1-\epsilon)}{(1-\epsilon)x-1}$ has a singularity at $x=\frac{1}{1-\epsilon}>1$. For the values of x less than $\frac{1}{1-\epsilon}$, y is less than one. Recalling that we have fixed $y=\frac{d_2}{d_1}<1$, this branch of the solution is not admissible. Thus, the only branch of the solution which might interest us is given in the first line of the above inequality.[7] The solution is illustrated in Figure 1 for two values of ϵ, $\epsilon=0.0333$ (solid) and $\epsilon=0.333$ (dashed). The inequality is satisfied in the area of the Figure located above and to the right of the corresponding line.

One could make the following conclusions. If eigenvalues of J are not too collinear (the value of T not too close to 1), only the very high values of $x=\frac{\lambda_2}{\lambda_1}$ and $y=\frac{d_2}{d_1}$ guarantee presence of a positive $\widetilde{\lambda}$ and, therefore, reversal of learnability.[8] If, on the other hand, v_1 and v_2 are almost collinear, learnability could be reversed for relatively mild ratios of $\frac{\lambda_2}{\lambda_1}$ and $\frac{d_2}{d_1}$. Collinearity of two eigenvectors of a matrix is not a generic property, and thus very low ϵ are unlikely to be observed for a matrix J.

[6]Note, however, that if $T=1$, then (17) has a solution $\widetilde{\lambda}=0$. We would be interested in behavior of $\widetilde{\lambda}$ as a function of T and will use $T=1$ as the benchmark case.

[7]Similar considerations demonstrate that if $\epsilon<0$, then both branches of the inequality do not satisfy conditions $x>1$, $y>1$.

[8]Notice that the angle between eigenvectors of J is preserved under the rotation into the coordinate system determined by the eigenvectors of M. Therefore, we can talk about collinearity for eigenvectors of J and $\widetilde{J}$ interchangeably.



5 Reversal of Learnability and the Speed of Convergence.

[4] studies the speed with which adaptive learning dynamics converges to the REE. The paper considers a one–dimensional case. The author shows that in case the (only) eigenvalue of J is below -0.5, the convergence of the adaptive learning process to the REE occurs with the "root-t" speed, in other words, exponentially fast. If this root is above -0.5 but below 0, so that the REE is still E–stable, the convergence is much slower. Ferrero (2004) further quantifies the welfare losses related to a slow convergence, and shows that they could be very substantial. He advocates adopting monetary policy which assures exponentially fast convergence to the REE, and thus the eigenvalue bounded from above by -0.5. In a multi–dimensional context, the speed of convergence is determined by the eigenvalue of J which is the smallest in absolute value. This policy prescription then reads that monetary policy should be structured in such a way that the real part of closest to the imaginary axis eigenvalue is below -0.5.

Comparing the policy prescription with the result derived in the previous section, it immediately obtains that accepting Ferrero's recommendations makes learnability reversals more difficult. Suppose that we pick values of $\epsilon > 0$, $\lambda_1 > 0$, $\lambda_2 > 0$, and $\frac{d_2}{d_1} > 0$ from some joint distribution. It is immediately obvious that by restricting the support of the distribution to $\lambda_1 > 0.5$, $\lambda_2 > \lambda_1$

will make picking a point with very high $\frac{\lambda_2}{\lambda_1}$ much less likely.[9]

If one is willing to impose the probability distribution described above, the results in this paper further suggest that deriving the probability of picking matrices J and M so that J is stable but MJ is not, would be a relatively simple numerical exercise.

6 Conclusion

While under recursive least squares learning the dynamics of the model converges to rational expectations equilibria (REE) which are E–stable, see Evans and Honkapohja (2003), some recent examples propose that E–stability is not a sufficient condition for learnability. In this paper, we provide some further evidence on the conditions under which E–stability of a particular equilibrium might fail to imply its stochastic gradient (SG) or generalized SG learnability. We also claim that the requirement on the speed of convergence of the learning process imposed by [4] also implies that E–stable equilibria are likely to be GSG learnable. We show this in a simple "New Keneysian" model of optimal monetary policy design in which we investigate the stability of REE under SG learning. Our findings are two fold: we examine the model under an alternative learning scheme SG and derive the conditions in this case which are necessary for the reversal of learnability.

References

[1] Emilio Barucci and Leonardo Landi. Least mean squares learning in self-referential linear stochastic models. *Economics Letters*, 57:313–317, 1997.

[2] George W Evans and Seppo Honkapohja. *Learning and Expectations in Macroeconomics*. Princeton University Press, Pricenton, Oxford, 2001.

[3] George W Evans, Seppo Honkapohja, and Noah Williams. Generalized stochastic gradient learning. NBER Technical Working Paper 317, 2005.

[4] Giuseppe Ferrero. Monetary policy and the transition to rational expectations. Banca'd Italia Termi di discussioni 499, 2004.

[5] Chryssi Giannitsarou. E–stability does not imply learnability. *Macroeconomic Dynamics*, 9:276–287, 2005.

[6] Maik Heinemann. Convergence of adaptive learning and expectational stability: The case of multiple rational-expectations equilibria. *Macroeconomic Dynamics*, 4:263–288, 2000.

[9] By how much less likely depends on the way the assumed probability density changes when we change the support. One natural possibility is to increase probability density at every remaining point proportionally. In this case, it is a straightforward exercise to determine a probability of picking a matrix J with a ratio $\frac{\lambda_2}{\lambda_1}$ above any given number. The probability should drop dramatically for any imaginable distribution .