Introduction

Albert Marcet and Thomas J. Sargent

Moving-Average Perceptions Learning with Autoregressive Recursive Least Squares: Speed of Convergence of

...
\[
\begin{align*}
\frac{\ln p_2 + \ln 2}{\ln p_1} &= \ln 2 \\
\frac{\ln x_2 + \ln 1 + \ln 2}{\ln x_1} &= \ln 2 \\
\end{align*}
\]

Defining the state of the system, 2, and the system noise 2 as...
Definition

A stationary equilibrium

is the fixed point of the mapping between the two parameters.

The process is described by the following two conditions:

\[ \begin{align*}
&\text{Condition 1: } (g)^{2}\mathcal{W}(g)^{2}\mathcal{W}(g)^{T} = (g)^{2}
\end{align*} \]

The expected value of the mapping is equal to the expected value of the parameters.

Next, we describe the relationship between the two points of $g$.

2 Existence and Uniqueness of Stationary Equilibrium

Parameters $g$ is the fixed point of $\mathcal{W}(g)^{2}\mathcal{W}(g)^{T}$, where $g$ is the stationary equilibrium.

The fixed point of $\mathcal{W}(g)^{2}\mathcal{W}(g)^{T}$ is determined by the unique solution of $\mathcal{W}(g)^{2}\mathcal{W}(g)^{T} = 0$.

The equation for $g$ is given by:

\[ (g)^{2}\mathcal{W}(g)^{2}\mathcal{W}(g)^{T} = (g)^{2} \]

where

\[ \mathcal{W}(g) = (g)^{2} \]

The solution to $\mathcal{W}(g)^{2}\mathcal{W}(g)^{T} = 0$ is $g = 0$.

Note that this is the fixed point of $\mathcal{W}(g)^{2}\mathcal{W}(g)^{T}$.

\[ \mathcal{W}(g) = (g)^{2} \]

where

\[ g = (g)^{2} \]

The fixed point is determined by $g = (g)^{2}$.

Therefore, the stationary equilibrium is given by $g = (g)^{2}$.
Lemma 0. Let $\Gamma$ be the noise that satisfies (11.9). We have

\[ \lim_{n \to \infty} \frac{\|\mathbf{y} - \mathbf{x}\|^2}{n} = 0 \]

where $\mathbf{y}$ is the output of the parameter estimator ($p$), and $\mathbf{x}$ is the actual parameter set. This is because the noise term $\mathbf{n}$ is white and has zero mean. Therefore, the error term $\mathbf{e} = \mathbf{y} - \mathbf{x}$ is also white and has zero mean, which implies that the variance of $\mathbf{e}$ is zero. Hence, the above limit statement holds.

Proof: From equation (6.9), we can write

\[ \frac{\partial \sigma^2}{\partial \beta} = \frac{\partial \sigma^2}{\partial \beta} = 0 \]

and

\[ \frac{\partial \sigma^2}{\partial \gamma} = \frac{\partial \sigma^2}{\partial \gamma} = 0 \]

where $\sigma^2$ is the variance of the noise term $\mathbf{n}^2$. If we set $\partial \gamma = 0$, we get

\[ \frac{\partial \sigma^2}{\partial \beta} = \frac{\partial \sigma^2}{\partial \beta} = 0 \]

and

\[ \frac{\partial \sigma^2}{\partial \gamma} = \frac{\partial \sigma^2}{\partial \gamma} = 0 \]

where $\mathbf{n}$ is the noise term and $\beta$ is the parameter set.

Theorem 0. Let $\mathbf{y}$ be the output of the parameter estimator ($p$), and $\mathbf{x}$ be the actual parameter set. Then, the error term $\mathbf{e} = \mathbf{y} - \mathbf{x}$ is white and has zero mean. Therefore, the variance of $\mathbf{e}$ is zero. Hence, the above limit statement holds.

Proof: From equation (6.9), we can write

\[ \frac{\partial \sigma^2}{\partial \beta} = \frac{\partial \sigma^2}{\partial \beta} = 0 \]

and

\[ \frac{\partial \sigma^2}{\partial \gamma} = \frac{\partial \sigma^2}{\partial \gamma} = 0 \]

where $\sigma^2$ is the variance of the noise term $\mathbf{n}^2$. If we set $\partial \gamma = 0$, we get

\[ \frac{\partial \sigma^2}{\partial \beta} = \frac{\partial \sigma^2}{\partial \beta} = 0 \]

and

\[ \frac{\partial \sigma^2}{\partial \gamma} = \frac{\partial \sigma^2}{\partial \gamma} = 0 \]

where $\mathbf{n}$ is the noise term and $\beta$ is the parameter set.
It is possible to find parameter values for which no stationary equilibrium exists.

Previous papers prove that there exists no other equilibrium. This paper proves that there is no other stationary equilibrium. The statement of the proposition is as follows.

Proposition 2: If the parameters satisfy the conditions
\[ \frac{d_2 + d}{d_2 - d} > \left| \frac{\gamma - 1}{\gamma - 1} \right| \]

then there exists a unique stationary equilibrium.

We prove this proposition by contradiction. Assume there exists a stationary equilibrium \( (d_0, \gamma_0) \) that satisfies
\[ \frac{d_0 + d}{d_0 - d} > \left| \frac{\gamma - 1}{\gamma - 1} \right| \]

Then, we can write
\[ \frac{d_0 + d}{d_0 - d} > \left| \frac{\gamma - 1}{\gamma - 1} \right| \]

This contradicts the condition for the existence of a stationary equilibrium. Therefore, there exists a unique stationary equilibrium.

Proof: Suppose there exists a stationary equilibrium \( (d_0, \gamma_0) \) that satisfies
\[ \frac{d_0 + d}{d_0 - d} > \left| \frac{\gamma - 1}{\gamma - 1} \right| \]

Then, we can write
\[ \frac{d_0 + d}{d_0 - d} > \left| \frac{\gamma - 1}{\gamma - 1} \right| \]

This contradicts the condition for the existence of a stationary equilibrium. Therefore, there exists a unique stationary equilibrium.
Under pseudo-linear regression, the system evolves according to

\[ \frac{dx}{dt} = f(x) \]

where \( f(x) \) is a function of \( x \), and \( x \) is the vector of state variables. The parameters of the model are estimated using the least squares method.

The key result of the paper is that the model is stable if the parameters satisfy certain conditions. The conditions are:

1. \( \gamma > 0 \)
2. \( \rho < 0 \)

These conditions ensure that the system converges to a stable equilibrium point. The proof is based on the analysis of the characteristic equation of the linearized system around the equilibrium point.

The paper also discusses the implications of these results for economic models, particularly those involving recursive least squares estimation.
\[ y - (g)^*W = y/p \]

\[ \left[ g - i\eta \right] (g)\frac{1}{D}\frac{1}{n}(g) = \frac{1}{D}\frac{1}{n}(g) \]

In summary, under pseudo-linear regression, we have the ordinary differential equation

\[ \left( g \right)^{*W} W_{-1} y = \frac{1}{D}\frac{1}{n}(g) \]

where

\[ \left[ g - i\eta \right] (g)\frac{1}{D}\frac{1}{n}(g) = \frac{1}{D}\frac{1}{n}(g) \]

\[ \left[ g - i\eta \right] (g)\frac{1}{D}\frac{1}{n}(g) \]

The associated ordinary differential equations

Instrumental for an ARMA(1,1) model given by (6,24b) can be shown that \( y \) is an optimal form of

\[ \begin{bmatrix} 1-\lambda^2 \lambda \frac{1}{1-\lambda^2} \end{bmatrix} \]

\[ \begin{bmatrix} 1-\lambda^2 \lambda \frac{1}{1-\lambda^2} \end{bmatrix} \]

Geometric distribution of the second error term, the ARMA(1,1) process is

\[ \begin{bmatrix} 1-\lambda^2 \lambda \frac{1}{1-\lambda^2} \end{bmatrix} \]

Linear regression chooses \( y \) to impose the optimal orthogonal approximation on the orthogonal condition. The pseudo-linear regression method is superior for recursive estimation. For recursive prediction error method, both pseudo-linear regression and recursive instrumental variable estimate. Both pseudo-linear regression and recursive instrumental variable estimate the parameters of an ARMA(1,1) process.
\[ \phi(\theta) L^n = (\theta)^d \]  
\[ A(\theta) L^n = (\theta)^S \]  

To determine the coefficients that can be represented by the matrices \( \phi(\theta) \) and \( A(\theta) \), we need to consider the least squares estimate of the regression equation, which is given by:

\[ \hat{\theta} = (\phi(\theta)^T \phi(\theta))^{-1} \phi(\theta)^T y \]

The normal equations for these two regressions are:

\[ 0 = \hat{\theta}^T \phi(\theta) \]  
\[ 0 = \hat{\theta}^T A(\theta) \]

Consider the regressions \( \phi(\theta) \) and \( A(\theta) \).

**Proposition 5**: Suppose that \( \theta \) satisfies \( \hat{\theta} = \theta \). Then, the solution to the normal equations is the least squares solution for the regression equation.

The operations and are associated with the recursive prediction error method and are given by:

\[ \begin{align*}
R(\theta) L^n &= \theta^d \\
A(\theta) L^n &= \theta^S
\end{align*} \]

Under the recursive prediction error method, the ordinary differential equation is

\[ R(\theta) \frac{dW}{d\theta} = \theta^p \]

\[ \left[ \theta - \hat{\theta}_0(\theta) L \right] \frac{dW}{d\theta} W_{1-\theta} = \theta^p \]

where

\[ \begin{align*}
0 &= \hat{\theta}^T \phi(\theta) \\
0 &= \hat{\theta}^T A(\theta)
\end{align*} \]

The ordinary differential equation is solved by the ordinary differential equation method.
Consider the "late" ordinary differential equations (6.28) for pseudo-linear regression.

Pseudo-linear regression

whether its coefficients are strictly negative in real part.

To check the local stability of the RPEM, we have to compute $\dot{y}$ and check

\[
\dot{y} = \frac{d}{dp} \left[ \frac{-g}{(g)_{1}^p} \right] \Rightarrow (g)_{1}^p \cdot \dot{W}_{1-1} = d \cdot \dot{W}
\]

or

\[
\dot{g} = \left[ I - \frac{d}{dp} \left( \frac{-g}{(g)_{1}^p} \right) \right] \Rightarrow (g)_{1}^p \cdot \dot{W}_{1-1} = d \cdot \dot{W}
\]

and

\[
\dot{g} = \left[ I - \frac{d}{dp} \left( \frac{-g}{(g)_{1}^p} \right) \right] \Rightarrow (g)_{1}^p \cdot \dot{W}_{1-1} = d \cdot \dot{W}
\]

Comparing the individual derivative and evaluating all $\dot{g}$ values

\[
\dot{g} = \left[ I - \frac{d}{dp} \left( \frac{-g}{(g)_{1}^p} \right) \right] \Rightarrow (g)_{1}^p \cdot \dot{W}_{1-1} = d \cdot \dot{W}
\]

we have to study the matrix we have to employ the matrix

In the vicinity of each point of $\dot{g}$ of $\dot{g}$, this system has dynamics that are

\[
\dot{y} = (g)_{1}^p \cdot \dot{W}_{1-1} = \frac{d}{dp}
\]

or

\[
\dot{y} = \frac{d}{dp} \left[ \frac{-g}{(g)_{1}^p} \right] \Rightarrow (g)_{1}^p \cdot \dot{W}_{1-1} = d \cdot \dot{W}
\]

Consider the "late" ordinary differential equation for the RPEM.

Recursive prediction error method

Local analysis of the ordinary differential equations

\[
1 - \left[ \frac{\rho(g)_{1}^p}{W_{1-1}} \right] = (g)_{1}^p
\]

For formulas of moment matrices, we have the following formulas

\[
\dot{g} = (g)_{1}^p
\]

We also have

\[
1 - \left[ \frac{\rho(g)_{1}^p}{W_{1-1}} \right] = (g)_{1}^p
\]

Next, we have from the definition of $\dot{g}$ in the pseudo-linear regression (hexagonal)

\[
1 - \left[ \frac{\rho(g)_{1}^p}{W_{1-1}} \right] = (g)_{1}^p
\]

To obtain these, we first use (6.28) to compute

\[
1 - \left[ \frac{\rho(g)_{1}^p}{W_{1-1}} \right] = (g)_{1}^p
\]

To compute $\dot{g}$, we need formulas for $\dot{(g)}_{1}^p\cdot \dot{W}_{1-1}$ and $(g)_{1}^p\cdot \dot{W}_{1-1}$.
The dynamic of the algorithm in the vicinity of $\beta$ are governed by

\[ \frac{dW}{dt} = \frac{\beta \frac{d}{dt} e}{\sigma^2} \frac{dW}{dt} + \frac{\beta}{\sigma^2} e \]
Figure 6.1: Parameter 0.1 of the recursive prediction error method

Figure 6.2: Parameters 0.1 and 0.1 of the ordinary differential equation determined by the recursive prediction error method

Figure 6.3: Parameter 0.1 of the recursive prediction error method

Differential equation determined by the recursive prediction error method

We have computed solutions of the ordinary differential equation for many other values. Notice how graphically, Figures 6.5 and 6.6 resemble Figures 6.3 and 6.4.

We set the initial value of \( t \) in the simulation at \( t = 0 \), and simulated

\[
\begin{bmatrix}
4 & 1 \\
2 & 1
\end{bmatrix}
\]

For the second set of parameter values, we calculated the solution of the ordinary differential equation, which is consistent with the other solutions calculated for similar conditions.
Analytic results. Then we consider a full information case.

In this section we describe some results on the rate of convergence that we obtain by

4 Speed of Convergence

In ordinary differential equation

(i) Local stability is governed by the eigenvalues associated with a smaller

According to recent results in numerical methods, there is the possibility of the

(ii) Global convergence is governed by the eigenvalues associated with a larger

The proposition stated in the appendix and in Mareci and Sargent (1989) on

The proposition supports the following conclusions about

The proposition supports the following conclusions about

Local Learning Systems can be discovered by studying their associated ordinary

The proposition supports the following conclusions about

The proposition supports the following conclusions about
based on the assumption that there is a for which the equation for covariance needs to be calculated. The equation for covariance of the rate of convergence is:

\[ \text{Cov}(\tilde{g}, \tilde{t}) \]

In this section we describe a numerical procedure for exploring the behavior of this equation.

**Covariance Based on Simulation**

For example, if we simulate the process of \( \tilde{g} \) and \( \tilde{t} \), we can estimate the covariance by averaging the product of the observed \( \tilde{g} \) and \( \tilde{t} \) values over a large number of simulations. This approach is often used in simulations to estimate the covariance between two random variables. The covariance can be calculated using the following formula:

\[ \text{Cov}(\tilde{g}, \tilde{t}) = \frac{1}{n} \sum_{i=1}^{n} (\tilde{g}_i - \mu_g)(\tilde{t}_i - \mu_t) \]

where \( \mu_g \) and \( \mu_t \) are the means of the \( \tilde{g} \) and \( \tilde{t} \) distributions, respectively, and \( n \) is the number of samples.

This formula represents the covariance as the average of the products of deviations of the \( \tilde{g} \) and \( \tilde{t} \) values from their respective means. A positive covariance indicates that the variables tend to move in the same direction, while a negative covariance indicates that they tend to move in opposite directions. A covariance of zero indicates that there is no linear relationship between the two variables.

In the previous section, we derived a formula for the rate of convergence of \( \tilde{g} \) given \( \tilde{t} \). Now we need to calculate the covariance of \( \tilde{g} \) and \( \tilde{t} \) to understand how these variables jointly contribute to the rate of convergence. The covariance provides insight into the degree of linear dependence between \( \tilde{g} \) and \( \tilde{t} \) and helps to quantify the extent to which changes in one variable are associated with changes in the other.
Table 6.4: Response and Error Data for the Model of Section 1 with Hidden Layer 2.
The texture can stay for a long time close to a point on the board, so the texture of a point is very close to that of its neighbors. The texture of an image (or in this case, the nearest neighbor of the image) is often used as a reference point. For instance, if the image is a photograph, the texture of a pixel is the average of the pixel's neighbors. If the image is a graph, the texture of a pixel is the average of the pixel's neighbors. The texture of an image (or in this case, the nearest neighbor of the image) is often used as a reference point.

The table below shows the speed of convergence when the derivative of a function is zero.
5. Conclusions

The efficient market hypothesis does not imply that errors in expectations will lead to misprices, and the number of mistakes or errors in the model. For very large samples, the beliefs have converged, the mean error is zero, and the model will work as predicted. In this case, the model will work as predicted. The mean error is zero, and the model will work as predicted.
Assumption 1. For \( g \in D \), \( T \) is twice differentiable and has one derivative.

Assumption 2. The operator \( T \) has a unique fixed point in the set \( D \).

The two propositions below pertain to versions of the algorithms (6.14).

\[
\begin{align*}
\text{(A.2)} & \quad (g, g') \\
\text{(A.1)} & \quad (g, g')
\end{align*}
\]

Appendix

The covariance matrix \( M' \) is non-singular.

By hypothesis, the recursive conditional least squares algorithm is identical to the recursive conditional least squares algorithm.
We analyzed a version of the model in [Mercel et al. (1999)] and found that the predictions are consistent with the observed data. Our model is based on a set of differential equations that describe the interactions between different components of the system.

In our model, we considered two main components: A and B. Component A is influenced by component B, while component B is influenced by a separate input, C. The model can be expressed as:

\[ \frac{dA}{dt} = f(A, B, C) \]

\[ \frac{dB}{dt} = g(A, B, C) \]

\[ \frac{dC}{dt} = h(A, B, C) \]

Where \( f, g, \) and \( h \) are functions that describe the interactions between the components. Our analysis focused on the stability of these systems and the conditions under which each component can be sustained.