ECE 417 Identification & Adaptive Control Fall 2002

Handout: ***Convergence of Extended Least Squares***

We again consider the noisy linear system

$$y(k+1) = \varphi_\circ(k)^T\theta_\circ + w(k+1), \tag{1}$$

where now $\varphi_\circ(k)$ contains noise variables, as well as input-output variables.

The Extended Least Squares algorithm (ELS, or AML) is analyzed in Lai & Wei "Extended least squares and their application to adaptive control and prediction in linear systems," IEEE T.A.C. vol AC-31, no. 10, October 1986, pp. 898–906. We give here a development of the main ideas.

The ELS algorithm The aposteriori error $\widehat{w}$, and the apriori error e are defined, respectively, by

$$\begin{aligned} \widehat{w}_n &= y_n - \varphi_{n-1}^T\hat{\theta}_n \\ e_n &= y_n - \varphi_{n-1}^T\hat{\theta}_{n-1}, \end{aligned}$$

where the pseudo-regression vector φ is given by

$$\varphi_{n-1} = (y_{n-1}, \dots, y_{n-p}, u_{n-1}, \dots, u_{n-p}, \widehat{w}_{n-1}, \dots, \widehat{w}_{n-p})^T.$$

The ELS algorithm is then given by

$$\hat{\theta}_n = \hat{\theta}_{n-1} + P_{n-1}\varphi_{n-1}e_n \tag{2}$$

$$P_n = P_{n-1} - \frac{P_{n-1}\varphi_n\varphi_n^T P_{n-1}}{1+\varphi_n^T P_{n-1}\varphi_n} \tag{3}$$

Just as in the RLS estimator, the recursion for $\hat{\theta}$ can be rewritten using the recursion for P. First expand $P_n\varphi_n$ as follows:

$$\begin{aligned} P_n\varphi_n &= P_{n-1}\varphi_n - \frac{(\varphi_n^T P_{n-1}\varphi_n)P_{n-1}\varphi_n}{1+\varphi_n^T P_{n-1}\varphi_n} \\ &= \frac{P_{n-1}\varphi_n}{1+\varphi_n^T P_{n-1}\varphi_n} \end{aligned}$$

Applying this to the estimate update equation gives,

$$\hat{\theta}_n = \hat{\theta}_{n-1} + \frac{P_{n-2}\varphi_{n-1}e_n}{1+\varphi_{n-1}^T P_{n-2}\varphi_{n-1}} \tag{4}$$

The apriori and aposteriori prediction errors are closely related: Using the recursion for $\hat{\theta}$, we have

$$\varphi_{n-1}^T\hat{\theta}_n = \varphi_{n-1}^T\hat{\theta}_{n-1} + \frac{\varphi_{n-1}^T P_{n-2}\varphi_{n-1}e_n}{1+\varphi_{n-1}^T P_{n-2}\varphi_{n-1}}$$

It follows from subtraction that

$$\widehat{w}_n = y_n - \varphi_{n-1}^T\hat{\theta}_n = y_n - \varphi_{n-1}^T\hat{\theta}_{n-1} - \left(\frac{\varphi_{n-1}^T P_{n-2}\varphi_{n-1}e_n}{1+\varphi_{n-1}^T P_{n-2}\varphi_{n-1}}\right)$$

which shows that

$$\widehat{w}_n = \frac{e_n}{1+\varphi_{n-1}^T P_{n-2}\varphi_{n-1}}. \tag{5}$$

This also gives the recursion $\hat{\theta}_n = \hat{\theta}_{n-1} + P_{n-2}\varphi_{n-1}\widehat{w}_n$.

A key identity The following result is what allows us to mimic the analysis of the RLS algorithm

$$\{C(z)(\widehat{w}(z) - w(z))\}_n \;=\; \varphi_{n-1}^T \tilde{\theta}_n \qquad (6)$$

where $\tilde{\theta}_n$ is defined here as $\theta_\circ - \hat{\theta}_n$, and $\theta_\circ = (-a_1, \ldots, b_1, \ldots, c_1, \ldots)$.

The derivation of (6) is a simple consequence of the definitions:

$$\begin{aligned}
C(z)(\widehat{w}(z) - w(z))_n &= (C-1)\widehat{w}_n - Cw_n + \widehat{w}_n \\
&= (c_1, \ldots, c_p) \begin{pmatrix} \widehat{w}_{n-1} \\ \vdots \\ \widehat{w}_{n-p} \end{pmatrix} - Cw_n + \widehat{w}_n \\
&= -(a_1, \ldots, a_p) \begin{pmatrix} y_{n-1} \\ \vdots \\ y_{n-p} \end{pmatrix} + (b_1, \ldots, b_p) \begin{pmatrix} u_{n-1} \\ \vdots \\ u_{n-p} \end{pmatrix} + (c_1, \ldots, c_p) \begin{pmatrix} \widehat{w}_{n-1} \\ \vdots \\ \widehat{w}_{n-p} \end{pmatrix} \\
&\quad -y_n + \widehat{w}_n \\
&= \varphi_{n-1}^T \theta_\circ - y_n + (y_n - \phi_{n-1}^T \hat{\theta}_n) \\
&= \varphi_{n-1}^T \tilde{\theta}_n
\end{aligned}$$

Lyapunov Recursion We can now examine the "Lyapunov function" Q which was used in Handout # 2 in the analysis of RLS. The following recursion will be derived below.

$$\begin{aligned}
Q_{k+1} &:= \tilde{\theta}_{k+1}^T P_k^{-1} \tilde{\theta}_{k+1} \\
&= Q_k - 2(\varphi_k^T \tilde{\theta}_{k+1}) \left[\left(\frac{1}{C(z)} - \frac{1}{2} \right) \varphi_k^T \tilde{\theta}_{k+1} \right] \\
&\quad -\varphi_k^T P_{k-1} \phi_k \widehat{w}_{k+1}^2 + 2w_{k+1}^2 \frac{\varphi_k^T P_{k-1} \varphi_k}{1 + \varphi_k^T P_{k-1} \phi_k} \\
&\quad 2w_{k+1} \left(\varphi_k^T \tilde{\theta}_k + \frac{\varphi_k^T P_{k-1} \phi_k}{1 + \varphi_k^T P_{k-1} \varphi_k} (e_{k+1} - w_{k+1}) \right).
\end{aligned}$$

To exploit this recursion, we apply the following lemma from linear systems theory: Suppose that the filter $\frac{1}{C(z)} - \frac{1}{2}$ is strictly positive real. Then there exists $\delta > \infty$ and $K \in \mathbb{R}$ such that

$$\sum_{k=0}^{n} 2 \left(\varphi_k^T \tilde{\theta}_{k+1} \right) \left[\left(\frac{1}{C(z)} - \frac{1}{2} \right) \varphi_k^T \tilde{\theta}_{k+1} \right] \geq \delta \sum_{k=0}^{n} (\varphi_k^T \tilde{\theta}_{k+1})^2 + K, \qquad n \geq 1.$$

The SPR property gives the bound

$$\begin{aligned}
Q_{n+1} &+ \delta \sum_{k=0}^{n} \mid \varphi_k^T \tilde{\theta}_{k+1} \mid^2 + K \\
&+ \sum_{k=0}^{n} \varphi_k^T P_{k-1} \varphi_k \widehat{w}_{k+1}^2
\end{aligned}$$

$$
\begin{aligned}
&+ \quad 2\sum_{k=0}^{n} w_{k+1}\left(\varphi_k^T\tilde\theta_k + \frac{\varphi_k^T P_{k-1}\varphi_k}{1+\varphi_k^T P_{k-1}\varphi_k}(e_{k+1}-w_{k+1})\right)\\
&\le \quad Q_0 + 2\sum_{k=0}^{n}\frac{\varphi_k^T P_{k-1}\varphi_k}{1+\varphi_k^T P_{k-1}\varphi_k} w_{k+1}^2 .
\end{aligned}
$$

Consider the "cross term" in this bound, which depends linearly on w_{k+1}. From property (P1) of white noise, from Handout # 2,

$$
\begin{aligned}
&2\sum_{k=0}^{n} w_{k+1}\left(\varphi_k^T\tilde\theta_k + \frac{\varphi_k^T P_{k-1}\varphi_k}{1+\varphi_k^T P_{k-1}\varphi_k}(e_{k+1}-w_{k+1})\right)\\
&\quad = o\sum_{k=0}^{n}\left(\varphi_k^T\tilde\theta_k + \frac{\varphi_k^T P_{k-1}\varphi_k}{1+\varphi_k^T P_{k-1}\varphi_k}(e_{k+1}-w_{k+1})\right)^2
\end{aligned}
$$

Using this probabilistic bound, we can then proceed using simple algebraic manipulations. Observe that from (5) we have

$$
\varphi_k^T\tilde\theta_k + \frac{\varphi_k^T P_{k-1}\varphi_k}{1+\varphi_k^T P_{k-1}\varphi_k} e_{k+1} = \varphi_k^T\tilde\theta_{k+1}.
$$

This and (P1) thus gives

$$
\begin{aligned}
&2\sum_{k=0}^{n} w_{k+1}\left(\varphi_k^T\tilde\theta_k + \frac{\varphi_k^T P_{k-1}\varphi_k}{1+\varphi_k^T P_{k-1}\varphi_k}(e_{k+1}-w_{k+1})\right)\\
&\quad = \quad o\left(\sum_{k=0}^{n}\left(\varphi_k^T\tilde\theta_{k+1} + \frac{\varphi_k^T P_{k-1}\varphi_k}{1+\varphi_k^T P_{k-1}\varphi_k} w_{k+1}\right)^2\right)\\
&\quad = \quad o\left(\sum_{k=0}^{n}\left((\varphi_k^T\tilde\theta_{k+1})^2 + \left(\frac{\varphi_k^T P_{k-1}\varphi_k}{1+\varphi_k^T P_{k-1}\varphi_k}\right)^2 w_{k+1}^2\right)\right)\\
&\quad = \quad o\left(\sum_{k=0}^{n}(\varphi_k^T\tilde\theta_{k+1})^2 + \frac{\varphi_k^T P_{k-1}\varphi_k}{1+\varphi_k^T P_{k-1}\varphi_k} w_{k+1}^2\right)
\end{aligned}
$$

In words, the cross term is insignificant when compared with other quantities in the Lyapunov recursion.

The bound on Q can therefore be written as

$$
\begin{aligned}
&Q_{n+1} + \sum_{k=0}^{n}(\varphi_k^T\tilde\theta_{k+1})^2 + \sum_{k=0}^{n}\varphi_k^T P_{k-1}\varphi_k \widehat{w}_{k+1}^2\\
&\quad = \quad O\left(\sum_{k=0}^{n}\frac{\varphi_k^T P_{k-1}\varphi_k}{1+\varphi_k^T P_{k-1}\varphi_k} w_{k+1}^2\right)\\
&\quad = \quad O\left(\sum_{k=0}^{n}\frac{\varphi_k^T P_{k-1}\varphi_k}{1+\varphi_k^T P_{k-1}\varphi_k}\right) = O(\log\lambda_{\max} P_n^{-1})
\end{aligned}
$$

where the last bound is identical to that used in the analysis of RLS.

Conclusions: Just as with RLS, suppose that the disturbance $\mathbf{w}$ is a white noise sequence, and that for each n, the noise variable w_n is statistically independent of $\{\phi_i, y_i, v_i, i \le n-1\}$. Assume moreover that $\frac{1}{C(z)} - \frac{1}{2}$ is strictly positive real. Then, the ELS algorithm has the following properties:

1. Since P is positive definite, we have the bound

$$(\lambda_{\min} P_n^{-1})|\tilde{\theta}_{n+1}|^2 \le Q_{n+1} \le O(\log \lambda_{\max} P_n^{-1})$$

   Rearranging, this gives

$$|\tilde{\theta}_{n+1}|^2 = O\left(\frac{\log \lambda_{\max} P_n^{-1}}{\lambda_{\min} P_n^{-1}}\right)$$

2. We have from the bound on Q,

$$\sum_{k=0}^{n} (\varphi_k^T \tilde{\theta}_{k+1})^2 = O(\log \lambda_{\max} P_n^{-1})$$

   But, since $\widehat{w}_n - w_n = \frac{1}{C} \varphi_{n-1}^T \tilde{\theta}_n$, and the polynomial C is stable, we must also have

$$\sum_{k=0}^{n+1} (\widehat{w}_k - w_k)^2 = O(\log \lambda_{\max} P_n^{-1})$$

3. Finally, using the identity

$$\varphi_k^T \tilde{\theta}_k = \varphi_k^T \tilde{\theta}_{k+1} - \varphi_k^T P_{k-1} \varphi_k \widehat{w}_{k+1}$$

   we get the elementary bound

$$\begin{aligned} \frac{|\varphi_k^T \tilde{\theta}_k|^2}{1 + \varphi_k^T P_{k-1} \varphi_k} &\le 2 \frac{(\varphi_k^T \tilde{\theta}_{k+1})^2}{1 + \varphi_k^T P_{k-1} \varphi_k} + 2 \frac{(\varphi_k^T P_{k-1} \varphi_k)^2 \widehat{w}_{k+1}^2}{1 + \varphi_k^T P_{k-1} \varphi_k} \\ &\le 2(\varphi_k^T \tilde{\theta}_{k+1})^2 + 2 \varphi_k^T P_{k-1} \varphi_k \widehat{w}_{k+1}^2. \end{aligned}$$

   The bound on Q then gives

$$\sum_{k=0}^{n} \frac{|\varphi_k^T \tilde{\theta}_k|^2}{1 + \varphi_k^T P_{k-1} \varphi_k} = O(\log \lambda_{\max} P_n^{-1}).$$

Typically, we will find that $\frac{1}{n+1} \log \lambda_{\max} P_n^{-1} \to 0$. This is a very mild *stability* result. So, the bounds above show that $\widehat{w}_k$ is a good approximation of w_k, and that output predictions will be accurate, even without persistence of excitation.

Proof of Lyapunov recursion:

$$\begin{aligned} Q_{k+1} &= \tilde{\theta}_{k+1}^T (P_{k-1}^{-1} + \varphi_k \varphi_k^T) \tilde{\theta}_{k+1} \\ &= \tilde{\theta}_{k+1}^T P_{k-1}^{-1} \tilde{\theta}_{k+1} + (\varphi_k^T \tilde{\theta}_{k+1})^2 \end{aligned}$$

Also, $\tilde{\theta}_{k+1} = P_{k-1}[P_{k-1}^{-1}\tilde{\theta}_k - \varphi_k \widehat{w}_{k+1}]$. So,

$$\begin{aligned} Q_{k+1} &= (P_{k-1}^{-1}\tilde{\theta}_k - \varphi_k \widehat{w}_{k+1})^T P_{k-1} P_{k-1}^{-1} P_{k-1} (P_{k-1}^{-1}\tilde{\theta}_k - \varphi_k \widehat{w}_{k+1}) + (\varphi_k^T \tilde{\theta}_{k+1})^2 \\ &= \tilde{\theta}_k^T P_{k-1}^{-1} \tilde{\theta}_k - 2\widehat{w}_{k+1} \varphi_k^T \tilde{\theta}_k + \varphi_k^T P_{k-1} \varphi_k \widehat{w}_{k+1}^2 + (\varphi_k^T \tilde{\theta}_{k+1})^2 \end{aligned}$$

Expanding the term $\widehat{w}_{k+1}\varphi_k^T\tilde{\theta}_k$ will give the bound: We have

$$\begin{aligned} \varphi_k^T \tilde{\theta}_k &= \varphi_k^T \tilde{\theta}_{k+1} + \frac{\varphi_k^T P_{k-1} \varphi_k e_{k+1}}{1 + \varphi_k^T P_{k-1} \varphi_k} \\ &= \varphi_k^T \tilde{\theta}_{k+1} + \varphi_k^T P_{k-1} \varphi_k \widehat{w}_{k+1} \end{aligned}$$

Hence,

$$\begin{aligned} Q_{k+1} &= Q_k - 2\varphi_k^T \tilde{\theta}_{k+1} \widehat{w}_{k+1} - 2\varphi_k^T P_{k-1} \varphi_k \widehat{w}_{k+1}^2 + \varphi_k^T P_{k-1} \varphi_k \widehat{w}_{k+1}^2 + (\varphi_k^T \tilde{\theta}_{k+1})^2 \\ &= Q_k - 2\varphi_k^T \tilde{\theta}_{k+1} \widehat{w}_{k+1} - \varphi_k^T P_{k-1} \phi_k \widehat{w}_{k+1}^2 + (\phi_k^T \tilde{\theta}_{k+1})^2 \end{aligned}$$

The “key identity” completes the proof. □