A New Proof for the Stability of Equation-Error Models

Roberto López-Valcarce, *Student Member, IEEE,* and Soura Dasgupta, *Fellow, IEEE*

***Abstract*—Some recent works have shown that under an autoregressive constraint on the input signal, least-squares equation-error methods provide stable models of the estimated transfer function. Here we present an alternative proof of this fact which allows to increase the order of the autoregressive input by one, for both the monic and unit-norm approaches.**

***Index Terms*—Equation error, stability, system modeling.**

I. Introduction

THE EQUATION-ERROR (or least-squares) modeling scheme assumes an input/output description of the system under study as the following:

$$\begin{aligned} d(n) &= \sum_{k=0}^{\infty} h_k u(n-k) + v(n) \\ &= \left(\sum_{k=0}^{\infty} h_k z^{-k}\right) u(n) + v(n) \\ &= \underbrace{H(z)u(n)}_{=y(n)} + v(n) \end{aligned} \tag{1}$$

where the processes $u(\cdot)$ and $u(\cdot)$ are jointly stationary and the noise term $v(\cdot)$ is uncorrelated with the input $u(\cdot)$. The model

$$\hat{H}(z) = \frac{B(z)}{A(z)} = \frac{b_0 + b_1 z^{-1} + \cdots + b_N z^{-N}}{a_0 + a_1 z^{-1} + \cdots + a_M z^{-M}} \tag{2}$$

is constructed by minimizing the variance

$$\min_{A(z),\, B(z)} E[|A(z)d(n) - B(z)u(n)|^2] \tag{3}$$

where some constraint must be imposed on $A(z)$ to avoid the zero solution. Usually one fixes $a_0 = 1$ (monic approach) or $\sum_{k=0}^{M} a_k^2 = 1$ (unit norm approach). Let us denote the solution of (3) by $A_*(z)$ and $B_*(z)$. It is desirable then that $A_*(z)$ be minimum phase (i.e., all of its roots lie inside the unit circle) so that $\hat{H}(z)$ is a stable transfer function. It is known that in certain cases $\hat{H}(z)$ may be unstable [5]. Regalia [3] has recently shown that for both monic and unit norm approaches in the case $N = M$, $\hat{H}(z)$ is stable if $u(\cdot)$ is an autoregressive (AR) process of degree less than or equal to M. In particular, no assumptions are made about the spectral characteristics of the noise or the degree of the transfer function $H(z)$. This note provides an alternative proof of this result, generalizing it to the case $N \neq M$ and showing that it suffices to have $u(\cdot)$ AR of degree less than or equal to $N + 1$.

The following notation is used. Let $\mathbf{M}$ be a matrix of size $m \times m$; we denote by $\mathbf{M}\rfloor$, $\lceil\mathbf{M}$, $\lfloor\mathbf{M}$ and $\mathbf{M}\rceil$, respectively, the northwest, southeast, northeast, and southwest matrices of size $(m-1) \times (m-1)$ extracted from $\mathbf{M}$. The inner product of two transfer functions is defined as

$$\langle f(z),\, g(z)\rangle = \frac{1}{2\pi j} \oint S_u(z) f^*(1/z^*) g(z) \frac{dz}{z}$$

where the path of integration is the unit circle and $S_u(z)$ is the power spectral density of $u(\cdot)$. The Szegö polynomials $\{p_k(z)\}_{k=0}^{\infty}$ associated to $u(\cdot)$ form an orthonormal basis for the space of stable, causal transfer functions with respect to this inner product [3]. Each $p_k(z)$ has degree k and all roots outside the unit circle. If $u(\cdot)$ is AR of degree m then for all $k \geq m$, $p_k(z) = z^{-(k-m)} p_m(z)$, and $S_u(z) = 1/[p(z)p^*(1/z^*)]$ with $p(z) = z^{-m} p_m(z^{-1})$ [3].

From the vectors

$$\begin{aligned} \mathbf{u}(n) &= [u(n) \quad \cdots \quad u(n-N)]^t, \\ \mathbf{v}(n) &= [v(n) \quad \cdots \quad v(n-M)]^t, \\ \mathbf{y}(n) &= [y(n) \quad \cdots \quad y(n-M)]^t \end{aligned}$$

define $\mathbf{R}_{vv} = E[\mathbf{v}(n)\mathbf{v}(n)^t]$, $\mathbf{R}_{uu} = E[\mathbf{u}(n)\mathbf{u}(n)^t]$, $\mathbf{R}_{yy} = E[\mathbf{y}(n)\mathbf{y}(n)^t]$, $\mathbf{R}_{uy} = E[\mathbf{u}(n)\mathbf{y}(n)^t]$, and $\mathbf{R}_{y/u} = \mathbf{R}_{yy} - \mathbf{R}_{uy}^t \mathbf{R}_{uu}^{-1} \mathbf{R}_{uy}$. The variance (3) can be minimized with respect to $B(z)$, yielding the following reduced cost function [3]

$$K(\mathbf{a}) = \mathbf{a}^t(\mathbf{R}_{vv} + \mathbf{R}_{y/u})\mathbf{a} \tag{4}$$

which is a function of $A(z)$ only, through its coefficient vector $\mathbf{a} = [a_0 \quad \cdots \quad a_M]^t$.

II. The Monic Case

Here we consider the monic constraint $a_0 = 1$. The main tool will be the following result from [1].

Lemma 1: Let $\mathbf{R} > 0$, and let $\mathbf{a}_*$ be the monic vector that minimizes the quadratic cost $J(\mathbf{a}) = \mathbf{a}^t\mathbf{R}\mathbf{a}$. Then the polynomial $A_*(z)$ constructed from the coefficients of $\mathbf{a}_*$ has all roots strictly inside the unit circle if $\Delta \geq 0$, where $\Delta = \lceil\mathbf{R} - \mathbf{R}\rfloor$ is the displacement matrix of $\mathbf{R}$.

Our main result is as follows.

Manuscript received November 30, 1998. This work was supported by the National Science Foundation under Grant ECS-9350346. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. G. B. Giannakis.

The authors are with the Department of Electrical and Computer Engineering, University of Iowa, Iowa City, IA 52240 USA (e-mail: valcarce@icaen.uiowa.edu; dasgupta@icaen.uiowa.edu).

Publisher Item Identifier S 1070-9908(99)04083-3.

Theorem 1: Assume that $u(\cdot)$ is autoregressive of degree not exceeding $N+1$. Then the polynomial constructed from the monic vector that minimizes (4) has all roots strictly inside the unit circle.

Proof: First note from (4) that if $\mathbf{R}_{vv}+\mathbf{R}_{y/u}$ is singular, then the smallest achievable error variance is zero. This means that perfect modeling is possible: $v(\cdot)$ must be identically zero and the optimum model is $\hat{H}_*(z) = H(z)$, which is assumed stable. Therefore, assume that $\mathbf{R}_{vv} + \mathbf{R}_{y/u} > 0$. In view of Lemma 1, it suffices to show that the displacement matrix of $\mathbf{R}_{vv} + \mathbf{R}_{y/u}$ is positive semidefinite. Note that since $\mathbf{R}_{vv}$ is Toeplitz, this displacement matrix is

$$\Delta = \underbrace{\lceil \mathbf{R}_{vv} - \mathbf{R}_{vv} \rfloor}_{=0} + \lceil \mathbf{R}_{y/u} - \mathbf{R}_{y/u} \rfloor.$$

As in [3], consider shifted versions of $H(z)$ expanded over the Szegö polynomials as

$$z^{-l}H(z) = \sum_{k=0}^{\infty} h_k^{(l)} p_k(z), \qquad h_k^{(l)} = \langle z^{-l}H(z),\, p_k(z)\rangle$$

with $l \geq 0$. It is shown in [3] that with $0 \leq i,\, j \leq M$,

$$(\mathbf{R}_{y/u})_{i,\,j} = \sum_{k=N+1}^{\infty} h_k^{(i)} h_k^{(j)}. \tag{5}$$

If $u(\cdot)$ is AR of degree $N+1$ or less, then $p_{k+1}(z) = z^{-1}p_k(z)$ for $k \geq N+1$. As a consequence, for $k \geq N+1$

$$\begin{aligned} h_k^{(l-1)} &= \langle z^{-l+1}H(z),\, p_k(z)\rangle \\ &= \langle z^{-l}H(z),\, z^{-1}p_k(z)\rangle \\ &= \langle z^{-l}H(z),\, p_{k+1}(z)\rangle \\ &= h_{k+1}^{(l)}. \end{aligned}$$

This together with (5) implies that for $1 \leq i,\, j \leq M$,

$$(\mathbf{R}_{y/u})_{i,\,j} - (\mathbf{R}_{y/u})_{i-1,\,j-1} = h_{N+1}^{(i)} h_{N+1}^{(j)}$$

and therefore $\Delta = \mathbf{h}\mathbf{h}^t \geq 0$ where $\mathbf{h} = [h_{N+1}^{(1)} \quad \cdots \quad h_{N+1}^{(M)}]^t$. ■

III. The Unit Norm Case

Here we consider the case where the constraint $\mathbf{a}^t\mathbf{a} = 1$ is used. First we need to recast the result from Lemma 1 into this setting.

Lemma 2: Let $\mathbf{R} \geq 0$ be a matrix of size $m \times m$; let $J(\mathbf{a}) = \mathbf{a}^t\mathbf{R}\mathbf{a}$. If $\Delta = \lceil \mathbf{R} - \mathbf{R} \rfloor \geq 0$, then there exists a vector $\mathbf{a}_*$ solving

$$\min_{\mathbf{a}} J(\mathbf{a}) \qquad \text{subject to } \mathbf{a}^t\mathbf{a} = 1 \tag{6}$$

such that the polynomial $A_*(z)$ constructed from $\mathbf{a}_*$ has no zeros outside the unit circle. If $\Delta > 0$, then $A_*(z)$ has all zeros strictly inside the unit circle.

Proof: First, we shall prove that by reflecting an unstable zero with respect to the unit circle, the cost $J(\mathbf{a})$ cannot increase if $\Delta \geq 0$. Let $\mathbf{a}_*$ solve (6). Suppose that the unit-norm polynomial $A_*(z)$ has a root at $z = z_0$ with $|z_0| > 1$:

$$A_*(z) = (1 - z_0 z^{-1})C(z),$$

and let

$$\overline{A}_*(z) = (z_0^* - z^{-1})C(z)$$

be the polynomial obtained by reflecting z_0 with respect to the unit circle. It can be checked that $\overline{A}_*(z)$ remains unit norm [3]. Let $\overline{\mathbf{a}}_*$ and $\mathbf{c}$ be the coefficient vectors of $\overline{A}_*(z)$ and $C(z)$, respectively. Since

$$\mathbf{a}_* = \begin{bmatrix}\mathbf{c}\\0\end{bmatrix} - z_0\begin{bmatrix}0\\\mathbf{c}\end{bmatrix}, \quad \overline{\mathbf{a}}_* = z_0^*\left(\begin{bmatrix}\mathbf{c}\\0\end{bmatrix} - \frac{1}{z_0^*}\begin{bmatrix}0\\\mathbf{c}\end{bmatrix}\right)$$

one has

$$\begin{aligned} J(\mathbf{a}_*) &= \mathbf{c}^\dagger(\mathbf{R}\rfloor + |z_0|^2\lceil\mathbf{R} - z_0\lfloor\mathbf{R} - z_0^*\mathbf{R}\rceil)\mathbf{c}, \\ J(\overline{\mathbf{a}}_*) &= |z_0|^2\mathbf{c}^\dagger\left(\mathbf{R}\rfloor + \frac{1}{|z_0|^2}\lceil\mathbf{R} - \frac{1}{z_0^*}\lfloor\mathbf{R} - \frac{1}{z_0}\mathbf{R}\rceil\right)\mathbf{c} \end{aligned}$$

which give $J(\overline{\mathbf{a}}_*) = J(\mathbf{a}_*) + (1 - |z_0|^2)\mathbf{c}^\dagger\Delta\mathbf{c}$.

If $\Delta \geq 0$, then for $|z_0| > 1$ we have $J(\overline{\mathbf{a}}_*) \leq J(\mathbf{a}_*)$. Since $\mathbf{a}_*$ minimizes J equality must hold, i.e., $J(\overline{\mathbf{a}}_*) = J(\mathbf{a}_*)$. Let $\lambda = J(\mathbf{a}_*)$. It can be shown that λ is the smallest eigenvalue of $\frac{1}{2}(\mathbf{R}+\mathbf{R}^t)$ with associated unit norm eigenvector $\mathbf{a}_*$. If λ is simple, then the minimum of the cost J is unique, so we have a contradiction. If λ is multiple then $J(\overline{\mathbf{a}}_*) = J(\mathbf{a}_*)$ is possible. However, by reflecting the zero inside the unit circle the cost does not change, which means that a polynomial without roots outside the unit circle can be found in the solution set.

To see that $\Delta > 0$ gives strict stability, consider the family of unit norm polynomials with a zero at $z = re^{j\omega}$, with ω some fixed angle:

$$A_r(z) = g(r)(1 - re^{j\omega}z^{-1})C(z).$$

The normalization factor $g(r) = [\mathbf{c}^\dagger\mathbf{c}(1-r)^2 + r]^{-1/2}$ with $\mathbf{c}$ again the coefficient vector of $C(z)$, ensures that $A_r(z)$ has unit norm. Note that $g(1) = 1$, i.e., we have assumed that $C(z)$ is scaled to yield $(1 - e^{j\omega}z^{-1})C(z)$ unit norm. With $\mathbf{a}_r$ the coefficient vector of $A_r(z)$, one can write

$$J(\mathbf{a}_r) = g^2(r)[J(\mathbf{a}_1)r + \mathbf{c}^\dagger(\mathbf{R}\rfloor - r(\mathbf{R}\rfloor + \lceil\mathbf{R}) + r^2\lceil\mathbf{R})\mathbf{c}].$$

It is easy to verify that

$$\left.\frac{\partial J(\mathbf{a}_r)}{\partial r}\right|_{r=1} = \mathbf{c}^\dagger\Delta\mathbf{c}$$

which is strictly positive if $\Delta > 0$. This precludes the possibility of a root on the unit circle.

We can state now:

Theorem 2: Assume that $u(\cdot)$ is AR of degree not exceeding $N+1$. Then the polynomial constructed from the unit norm vector that minimizes (4) has no roots outside the unit circle. ■

Proof: The result follows immediatly in view of Lemma 2 by mimicking the proof of Theorem 1.

■

Remark: Note that for the monic constraint $\Delta \geq 0$ implies strict stability, but not for the unit norm constraint (for example, one can always construct settings for which the matrix $\mathbf{R}_{vv} + \mathbf{R}_{y/u}$ is Toeplitz. If its smallest eigenvalue is simple, then the roots of $A_*(z)$ lie *all on the unit circle* [2]). However, it turns out that whenever $A_*(z_0) = 0$ with $|z_0| = 1$, then $B_*(z_0) = 0$ as well, as can be seen from the results in [4]. Therefore, all the poles of the approximant $\hat{H}(z) = B_*(z)/A_*(z)$ on the unit circle are canceled out, and all of the remaining poles lie strictly inside the unit circle.

Acknowledgment

The authors would like to thank the anonymous Reviewer B for pointing out the results in [4] to us.

References

[1] J. Idier and J.-F. Giovannelli, “Structural stability of least squares prediction methods,” *IEEE Trans. Signal Processing,* vol. 46, pp. 3109–3111, Nov. 1998.

[2] J. Makhoul, “On the eigenvectors of symmetric Toeplitz matrices,” *IEEE Trans. Acoust., Speech, Signal Processing,* vol. ASSP-29, pp. 868–872, Aug. 1981.

[3] P. A. Regalia, “An unbiased equation error identifier and reduced order approximations,” *IEEE Trans. Signal Processing,* vol. 42, pp. 1397–1412, June 1994.

[4] P. A. Regalia, M. Mboup, and M. Ashari, “A class of first- and second-order interpolation problems in model reduction,” *Arch. Elektron. Übertragungstech.,* vol. 49, pp. 332–343, Sept. 1995.

[5] T. Söderström and P. Stoica, “On the stability of dynamic models obtained by least squares identification,” *IEEE Trans. Automat. Contr.,* vol. AC-26, pp. 575–577, 1981.