NEUVIEME COLLOQUE SUR LE TRAITEMENT DU SIGNAL ET SES APPLICATIONS

NICE du 16 au 20 MAI 1983

Simplified models for perturbed coefficient AR systems

Un modèle simplifié pour les systèmes autoregressifs avec coefficients stochastiques

J.R. Deller, Jr. and Z. Gulboy

Illinois Institute of Technology
Dept. of Electrical Engineering: Speech Processing Laboratory
Chicago, IL 60616 USA

RESUME

Un moyen de rendre compte de l'incertitude du comportement dynamique d' un système autorégressif modèle par l'équation

(1) $$x(n) = \sum_{i=1}^{M} a_i(n)x(n-i) + G(n)u(n)$$

est de rendre les coefficients ai(n) et G(n) stochastiques. Le but de cet article est de présenter un résumé des résultats concernant ce modèle. En particulier, il est supposé que les $a_i(n)$ et G(n) sont independants, et constituent des séquences aléatoires stationnaires du premier ordre, et que des conditions telles que (1) puissent etre approximées par une équation dont les coefficients constants sont les espérances mathématiques:

(2) $$\bar{x}(n) = \sum_{i=1}^{M} \bar{a}_i\bar{x}(n-i) + \bar{G}u(n)$$

Les limites du signal erreur $[x(n) - \bar{x}(n)]$ sont presentées. Elles servent a evaluer la performance du modèle simplifié dans le cas d'un segment fini, et presentent des conditions intrinsèques necessaires pour une bonne approximation. A partir de ces premiers résultats, il n' est possible d'obtenir qu' une condition de convergence asymptotique de l'erreur qui est très restrictive. Une étude additionelle, dans le cas d'un segment infini, permet d'obtenir une condition moins contraignante. Cette condition est exprimée en fonction d'une mesure de perturbation, et des poles du système simplifié.

Les données obtenues par simulation confirment les résultats théoriques.

SUMMARY

One way to account for uncertainty in the dynamics of an autoregressive (AR) system is to allow the coefficients of the model equation to be stochastic:

(1) $$x(n) = \sum_{i=1}^{M} a_i(n)x(n-i) + G(n)u(n)$$

It is the purpose of this paper to present a summary of results concerning this model. In particular, it is assumed that the $a_i(n)$ and G(n) are independent, first order stationary random processes and conditions are sought under which (1) can be approximated by a time invariant equation in which the coefficients and gain are the stochastic means:

(2) $$\bar{x}(n) = \sum_{i=1}^{M} \bar{a}_i\bar{x}(n-i) + \bar{G}u(n)$$

Bounds on the error signal $[x(n) - \bar{x}(n)]$ are presented which are useful in evaluating the performance of the simplified model on the finite time line, and which contain inherent conditions for good approximation. From these initial findings, however, only a very restrictive condition for asymptotic convergence of the error can be inferred. Further study of the long term case yields a significantly relaxed condition for convergence. This condition is formulated in terms of a measure of perturbation and the poles of the simplified system.

Data from simulation studies confirm the theoretical findings.

Acknowledgement

This work was supported in part by the National Science Foundation of the United States under Grant No. ECS-8006866.

1. MOTIVATING PROBLEM

The autoregressive (AR) model, which is widely employed in many interesting problems is often used with the foreknowledge that the time invariant AR equation is only an ad hoc approximation to the true system dynamics. One way to account for uncertainty in the model is to allow the coefficients of the AR equation to be stochastic:

$$x(n) = \sum_{i=1}^{M} a_i(n)x(n-i) + G(n)u(n) \tag{1}$$

where, $a_i(n)$ and $G(n)$ are independent, first order stationary processes and $u(n)$ is an uncorrelated driving sequence. While studying methods for modelling pathologic speech production (speech produced by a person with a functional or organic disease of the larynx), such a model became interesting to the authors. It is the purpose of this paper to present a summary of the results of the study of this stochastic AR model. In particular, we seek conditions under which the stochastic parameter system of (1) can be well approximated by a time invariant model in which the random coefficients and gains are replaced by their mean values:

$$\bar{x}(n) = \sum_{i=1}^{M} \bar{a}_i \bar{x}(n-i) + \bar{G}u(n) \tag{2}$$

in which $\bar{a}_i=E[a_i(n)]$, $\bar{G}=E[G(n)]$, and $\bar{x}(n)$ is the output of the simplified system. The use of the simplified dynamical equation (2) rather than (1), reduces the model to the domain of analysis by well known systems analytic techniques (AR identification, stability analysis, etc.). In Section 2 we present results which are primarily useful in evaluating the performance of the approximation on the finite time interval. In Section 3 the behavior on the infinite time line is studied.

2. RESULTS FOR THE FINITE TIME LINE

General Theory. We condider here the approximation of (1) by (2) on $n\in[1,N]$. In earlier work (Deller, 1981) it was proven, starting with a theorem for a more general class of systems (Meerkov, 1972), that, for the restricted class of systems for which $a_i(n)\cong\bar{a}_i$, and $G(n)\cong0$ $\forall i,n$ the simplification of the model is valid. (In fact, the simplification was shown to be good $\forall$ $n\in[0,\infty]$ if these approximations are sufficiently close.) It is clear that such conditions on $a_i(n)$ and $G(n)$ require these parameters to be small in both mean and variance. Intuitively, it seems reasonable that a small variance condition alone should be sufficient to permit the use of the approximate model. This intuition is formalized in the following theory:

THEOREM 2.1: A general Markov system of the form

$$\underline{y}(n+1) = \underline{\phi}(\underline{\xi}(n), \underline{y}(n), n), \quad \underline{y}(0) = \underline{y}_0 \tag{3}$$

where, $\underline{y}(n)$ is an M-vector, $\underline{\xi}(n)$ is a first order stationary Q-vector, and $\underline{\Phi}$ is a general vector function of $\underline{\xi}$, $\underline{y}$, and n, is well approximated by the equation

$$\bar{\underline{y}}(n+1) = \underline{\phi}(\bar{\underline{y}}(n), n), \quad \bar{\underline{y}}(0) = \underline{y}_0 \tag{4}$$

in the sense that

$$||\underline{y}(n) - \bar{\underline{y}}(n)|| \le \varepsilon\{1 + \sum_{i=1}^{N-1} K^i\}, \qquad \text{w.p.1} \tag{5}$$

if $\exists\,\varepsilon,K$ such that

$$||\underline{\Phi}(\underline{\xi}(n), \underline{y}(n), n) - \underline{\phi}(\underline{y}(n), n)|| \le \varepsilon \tag{6}$$

$$||\underline{\phi}(\underline{y}'(n), n) - \underline{\phi}(\underline{y}''(n), n)|| \le K\,||\underline{y}'(n) - \underline{y}''(n)|| \tag{7}$$

$\forall$ $\underline{y}',\underline{y}''$ in the domain of $\underline{\phi}$.

where, $||\underline{y}|| \triangleq [\sum_{i=1}^{M} y_i^2]^{1/2}$ for an M-vector

and, $||\underline{Y}|| \triangleq \{\text{max eigenvalue of } \underline{Y}\,\underline{Y}^T\}^{1/2}$ for a matrix.

PROOF: See (Deller and Gulboy, 1983).

The following corollary gives the desired results for AR systems:

COROLLARY 2.1: An AR system of form (1) is well approximated by (2) if the following conditions hold:

(8) 1. $|a_i(n) - \bar{a}_i| < \mu$, $|G(n) - \bar{G}| < \mu$, $\forall i,n$
  2. the input and initial conditions are bounded,

in the sense that

$$|x(n) - \bar{x}(n)| \le \mu \cdot f(N), \; n\in[1,N] \tag{9}$$

where f(N) is a bounded function of N.

For a fixed N, this corollary asserts the arbitrary goodness of the approximation by choice of small enough μ, or (small enough bound) in accordance with the intuition which motivated the theorem.

PROOF: The proof is sketched in the Appendix using notation developed in Section 3. A complete proof for the autoregressive moving average (ARMA) model is given in (Deller and Gulboy, 1983).

Simulation studies. Computer simulation studies to verify the results of Corollary 2.1 as well as the earlier work cited at the beginning of this section are reported in (Deller, 1981 (2)).

3. RESULTS FOR THE INFINITE TIME LINE

Preliminary discussion. It is clear that, for a fixed μ, the assymptotic convergence of the bound ((9)) on the approximation error depends on the convergence of f(N) as $N\to\infty$. In turn, examination of the proof of COROLLARY 2.1 (Eqn. A - 2) reveals that a sufficient condition for this convergence is that:

$$(p + \alpha\mu) < 1 \tag{10}$$

where, p=max pole magnitude of the averaged model (2) and α is a small positive number. This condition requires p<1 which, in turn, assures the assymptotic convergence of the other potentially explosive term in (A- 5) i.e.,

$$\sum_{i=1}^{N-1} p^i \tag{11}$$

The implication of this later condition is that the system of (2) need be bounded input-bounded output (BIBO) stable (all of its poles must be inside the unit circle in the z-plane) while that of the former is that the perturbed coefficient system be 'pointwise stable', i.e. that its poles be inside the unit circle for each n. This condition is unnecessarily restrictive as we demonstrate presently:

General Theory. Consider rewriting (1) in the state space formulation:

$$\underline{x}(n+1) = \underline{A}(n)\,\underline{x}(n) + G(n)\underline{u}(n) \tag{12}$$

$$x(n) = \underline{c}^T \underline{x}(n+1)$$

where, $\underline{x}(n)=[x(n-1)\ldots x(n-M)]^T$
$\underline{u}(n)=[u(n-1)\ldots u(n-M)]^T$
$\underline{c}=[1\;0\;0\;\ldots\;0]$

$$\underline{A}(n) = \begin{bmatrix} a_1(n)\; a_2(n) \ldots a_M(n) \\ \underline{I}_{(M-1)\times(M-1)} \quad \underline{0} \end{bmatrix}$$

and similarly rewriting (2) as

(13) $\bar{\underline{x}}(n+1) = \bar{\underline{A}}\bar{\underline{x}}(n) + \bar{\underline{G}}u(n)$

$\bar{x}(n) = \underline{c}^T\bar{\underline{x}}(n+1)$

where $\underline{x}, \underline{u}$ $\underline{c}$ and $\bar{\underline{A}}$ are defined in the obviously similar way. We define the error matrix,

(14) $\tilde{\underline{A}}(n) = \underline{A}(n) - \bar{\underline{A}}$

in which all entries are zero except for the first row which contains the coefficient perturbations $\tilde{a}_i(n)=a_i(n)-\bar{a}_i$.

The following are useful facts about the matrix $\bar{\underline{A}}$:

(15.1) The eigenvalues of $\bar{\underline{A}}$, $\bar{\lambda}_i$, are the poles of the AR system (2) (Kailath, 1980).

(15.2) It follows that $\|\bar{\underline{A}}\|$ = max pole magnitude of the system (2)

We define $p = \max_i |\bar{\lambda}_i|$

(15.3) Further, it follows that $\bar{\underline{A}}^T\bar{\underline{A}}$ has eigenvalues $|\bar{\lambda}_i|^2$ and norm $\max_i |\bar{\lambda}_i|^2 = p^2$.

The ultimate goal in this section is to examine the behavior of the approximation of (1) by (2) (or, equivalently, (12) by (13)) as $n \to \infty$. We make the assumption that the approximate system (13), is stable (either in the sense of Lyapunov or in the BIBO sense see below) which implies that the signal $\bar{x}(n)$ will remain bounded in response to bounded inputs.

It remains to determine conditions on the stochastic parameter system to assure that it likewise does not "blow up" in response to reasonable inputs and hence diverge away from the approximation system. Thus far we have only the very restrictive conditions noted above.

The results to follow are based in internal stability considerations since they are easier for formulate than those centering on BIBO stability. It should be noted, however, that since the transfer function of an all pole system is irreducible, internal stability and BIBO stability are equivalent (Kailath, 1980). This fact is also intuitively obvious from the autoregressive nature of the computation: a bounded output will require all previous (internal) values to be bounded as well. The main stability condition is contained in the following:

THEOREM 3.1: Let σ_i^2 denote the variance of $a_i(n)$. Then system of (12) and, hence, (1), is internally stable with probability one (wp1) if the condition

(16) $\|\bar{\underline{A}}^T\bar{\underline{A}} + \underline{S}\| < 1$

is satisfied, in which $\underline{S} = \text{diag}\,(\sigma_1^2, \sigma_2^2 \ldots \sigma_M^2)$. To prove the theorem we need the following based on (Serfling, 1980):

DEFINITION: Let X(n) be a sequence of random variables which is convergent in the rth mean ie.,

(17) $\lim_n E[|X(n) - X|^r] = 0$

X(n) is said to converge sufficiently fast in the rth mean if

(18) $\sum_{n=1}^{\infty} E[|X(n) - X|^r] < \infty$

LEMMA: Convergence sufficiently fast in the rth mean implies convergence wp1.

PROOF: The proof is given in (Serfling, 1980).

PROOF OF THEOREM 3.1: We seek a condition which causes the homogeneous solution of (12) to tend to $\underline{0}$ as $n \to \infty$, for arbitrary bounded $\underline{x}(0)$. If $\underline{u}(n)=\underline{0}$, it is easy to show that

(19) $E\|\underline{x}(n+1)\|^2 = E\|\underline{x}^T(n)\underline{A}^T(n)\underline{A}(n)\underline{x}(n)\|$

Note that

(20) $\underline{A}^T(n)\underline{A}(n) = E((\bar{\underline{A}}^T + \tilde{\underline{A}}(n))(\bar{\underline{A}} + \tilde{\underline{A}}(n))$

and it is not difficult to show that the cross terms in the quadratic are zero ie.,

(21) $E[\tilde{\underline{A}}^T(n)\bar{\underline{A}}] = E[\bar{\underline{A}}^T\tilde{\underline{A}}(n)] = \underline{0}$

and that the final term in the quadratic is equal to the matrix, $\underline{S}$, so that (19) becomes

(22) $E\|\underline{x}(n+1)\|^2 = E\|\underline{x}^T(n)\,\{\bar{\underline{A}}^T\bar{\underline{A}}+\underline{S}\}\underline{x}(n)\|$

$\leqslant \|\bar{\underline{A}}^T\bar{\underline{A}}+\underline{S}\|\; E\|\underline{x}(n)\|^2$

which can also be written

(23) $E\|\underline{x}(n+1)\|^2 \leqslant \|\bar{\underline{A}}^T\bar{\underline{A}}+\underline{S}\|^n\; E\|\underline{x}(0)\|^2$

Eqn. (23) implies that $E\|\underline{x}(n+1)\|^2$ and hence $E[x^2(n)]$ converges sufficiently fast in the mean if condition (16) holds. According to the LEMMA, therefore, (16) is a sufficient condition for internal stability of (12), and, hence, of (1). Q.E.D.

It is of interest to interpret the stability condition in terms of the poles of the systems (1) and (2) instead of the coefficients. We assume for simplicity that the variance of the random coefficients is the same for all i, so that $\underline{S} = \sigma^2\underline{I}$. The following result obtains:

COROLLARY 3.1: A sufficient condition for the internal stability of the system of form (12) and, hence, of (1), wp1 is that

(24) $p^2 + \sigma^2 < 1$

where p is the max pole magnitude of the system (13) or (2).

PROOF: From (15.2) it is clear that $\bar{\underline{A}}^T\bar{\underline{A}} + \underline{S}$ has eigenvalues $\rho_i = |\bar{\lambda}_i|^2 + \sigma^2$. By definition, $\|\bar{\underline{A}}^T\bar{\underline{A}} + \underline{S}\| = \max \rho_i = p^2 + \sigma^2$. Therefore, (24) and (16) are equivalent under the equal variance assumption.

It is interesting to note the similar condition which results from a different approach in the small perturbation case. Again we assume equal variances.

THEOREM 3.2: A sufficient condition for the internal stability of (12), and, hence, (1), is

(25) $p^2 + c^2\sigma^2 < 1$

where $c \stackrel{\Delta}{=} \sum_{i=1}^{M} \dfrac{p^{M-i}}{\prod_{|\bar{\lambda}_i| \neq p} (p - |\bar{\lambda}_i|)}$

and where the $\bar{\lambda}_i$ are as defined in (15.1).

SKETCH OF PROOF: (A complete proof is given in (Gulboy, 1982).) It is argued by Gulboy (using the results of perturbation theory (Kato, 1980)) that the eigenvalues of the perturbed matrix, $\underline{A}(n)$, are given approximately by

(26) $\lambda_i(n) \simeq \bar{\lambda}_i + c\,\tilde{a}_i(n)$

for small σ^2. Since the homogeneous solution,

$$(27)\qquad ||\underline{x}(n+1)|| = ||\prod_{r=1}^{n-1} \underline{A}(r)\ \underline{x}(0)|| \le \prod_{r=1}^{n-1} p^2(r)\ ||\underline{x}(0)||$$

where, p(r) = max eigenvalue ("pole") of $\underline{A}(r)$. Using (26) in (27) and taking the expectation leads to the conclusion that

$$(28)\qquad \lim_{n} ||\underline{x}(n+1)|| = 0,\ \text{wp1}$$

Q.E.D.

These results imply that the perturbed AR system of (1) need not be 'pointwise stable' in order for the system to produce a bounded output in the long term. In fact, it is quite clear from conditions (24) and (25) that one or more of the "poles" of $\underline{A}(n)$ may be outside the unit circle at any time value with nonzero probability. It is therefore possible to rationally approximate systems of form (1) which do not meet the restrictive conditions deduced from Corollary 2.1 by an average coefficient system of form (2). The usefulness of the approximation in the long term can be assessed by computation of the bound in (9) for large N, although it is clear that this bound may diverge even if the error in approximation does not.

We note finally, that the sufficient conditions proved in this section can also be shown to be necessary for the first order case, M=1 (Gulboy, 1982, Ch.3). Because the minimum norm is used in computing bounds above, one may conjecture that the conditions cited are at least close to being necessary in the higher order cases. This conjecture is supported by the following experimental data.

Simulation study. The identification of a pertubed sixth order AR process is considered. The location of the three pairs of complex poles are those of the formants of the vowel /i/, as in "BEET" when the sampling frequency is 10kHz.

Table 1. Pole Loc

Magnitude	Angle (rd)
0.9488	0.1700
0.9688	1.4391
0.9444	1.8910

The signals x(n) and $\bar{x}(n)$ were generated using a cascade form implementation where random sequences were added to the coefficients in order to simulate the time varying system. All random sequences were chosed to have the same variance, which corresponds to the case where all σ_i's are equal to σ.

In a first experiment we show the verification of the theoretical maximum admissible value of the variance σ^2. It was shown in Corollary 3.1 that a sufficient condition for convergence of x(n) with probability one is,

$$(29)\qquad p^2 + \sigma^2 < 1$$

which gives in this case:

$$(30)\qquad \sigma^2 < 0.6134$$

since p = 0.9688. This theoretical value was verified by generating 100 different realizations of x(n) for each value of σ which was progressively increased. The divergence of x(n) (or instability) was detected as an overflow error in the subroutine generating x(n).

It was observed that:

a) If $\sigma^2 < 0.060$ no overflow occurred.

b) If $\sigma^2 = 0.061$ x(n) diverged at least once among the 100 realizations.

c) For $\sigma^2 = 0.070$ overflow occurred once out of two realizations.

These results confirm theoretical expectations. The experimental value of the maximum admissible variance on the coefficients was slightly smaller than the value predicted in (30), but this discrepancy can be explained by the effect of the quantization on the coefficients $a_i(n)$ which is equivalent to an extra internal perturbation, and hence, causes the effective variance to be slightly greater than σ^2.

4. CONCLUSIONS

Results concerning the approximation of a stochastic coefficient AR system of form (1) by the time invariant model of form (2) have been presented. Under the condition that the coefficients are bounded, Corollary 2.1 can be employed to bound the approximation error over the finite time interval. As expected this bound vanishes as as the bound on the coefficient perturbations becomes small.

In the long term, however, Corollary 2.1 yields only a very restrictive condition for the error to remain bounded, i.e., for the perturbed system to converge. In particular, it is seen to be sufficient that the poles of (1) remain inside the unit circle for all n. This condition is relaxed in the results of the theorems and corollaries of Section 3 in which it is shown that the stochastic system need not be 'pointwise stable' to assure convergence of its output in steady state.

The approximation bound of Corollary 2.1 can be used to estimate the asymptotic error, though, from the results above, it is clear that this bound can diverge even if the error does not.

Simulation studies confirm the theoretical results.

5. REFERENCES

Deller, Jr., J.R., 1981, IEEE Trans. Acoust., Speech, Signal Proc., vol. 29, 312.

Deller, Jr., J.R., 1981 (2), Procedings IEEE Int. Conf. Acoust., Speech, Sig. Proc., vol. 2, 869.

Deller, Jr., J.R., and Z. Gulboy, 1983 (in press), Int. J. Sys. Sci.

Gulboy, Z., 1982, Deterministic Models for Perturbed Coefficient AR Systems (M.Sc. Dissertation), Illinois Institute of Technology, Chicago.

Meerkov, S.M., 1972, Automatika Telemekhanika, vol. 3, 66.

Kailath, T., 1980, Linear Systems, Prentice-Hall, Englewood Cliffs, NJ.

Kato, T., 1980, Perturbation Theory, Springer-Verlag, New York, NY.

Serfling, R.J., 1980, Approximation Theorems of Mathematical Statistics, John Wiley and Sons, New York, NY.

6. APPENDIX: SKETCH OF THE PROOF OF COROLLARY 2.1

This proof was deferred in order to take advantage of the results of Section 3. For a complete version of this proof, see (Deller and Gulboy, 1983).

Consider approximating the system of (12) by that of (13) We first show that condition (8)-1 is tantamount to (6) of Theorem 2.1. It is easy to show from (12), (13), and (8) that

$$\text{(A-1)} \quad ||\underline{\Phi}(\underline{\xi}(n), \underline{x}(n), n) - \underline{\phi}(\underline{x}(n), n)|| \leqslant ||\underline{A}(n) - \overline{\underline{A}}||\ ||\underline{x}(n)|| + \mu U$$

where $\underline{x}(n+1) = \underline{\Phi}(\underline{\xi}, \underline{x}, n)$,

and $\overline{\underline{x}}(n+1) = \underline{\phi}(\overline{\underline{x}}, n)$ and

if $||\underline{u}(n)|| \leqslant U$ as assumed in (8)-2. Now, $||\underline{A}(n) - \overline{\underline{A}}|| \leqslant \mu M^{\frac{1}{2}}$ and a bound on $||\underline{x}(n)||$ can be shown to be

$$\text{(A-2)} \quad ||\underline{x}(n)|| \leq X(p+c\mu)^{n-1} + U(\overline{G}+\mu)\{1 + \sum_{\ell=1}^{n-1} (p+c\mu)^{\ell}\} \triangleq S(n)$$

in which the notation of (24) and (27), and the result of Theorem 3.2 are employed. Clearly, $S(n) \leqslant S(N)$ $\forall\, n \leq N$. Therefore,

$$\text{(A-3)} \quad ||\underline{\phi}(\underline{\xi}(n), \underline{x}(n), n) - \underline{\phi}(\underline{x}(n), \underline{n})|| \leq \mu(M^{1/2}S(N) + U) \triangleq \varepsilon(N) \quad n\varepsilon[1,N] \qquad \text{w.p.1.}$$

Next, we verify that condition (7) holds:

$$\text{(A-4)} \quad ||\underline{\phi}(\underline{x}(n), n) - \underline{\phi}(\overline{\underline{x}}(n), n)|| = ||\overline{\underline{A}}(\underline{x}(n) - \overline{\underline{x}}(n))|| \leq p||(\underline{x}(n) - \overline{\underline{x}}(n))|| \qquad \text{w.p.1.}$$

Examining (A-3) and (A-4) and using Theorem 2.1 it is clear that

$$\text{(A-5)} \quad ||\underline{x}(n) - \overline{\underline{x}}(n)|| \leq \varepsilon(N)\{1 + \sum_{i=1}^{N-1} p^{i}\}$$

and (9) follows immediately. Q.E.D.