ON ADAPTIVE INVERSE CONTROL

B. Widrow, D. Shur, and S. Shaffer

Information Systems Lab
Stanford University
Stanford, CA 94305

Abstract

An uncertain plant will track an input command signal if the plant is preceded by a controller which approximates the inverse of the plant's transfer function. The controller parameters can be obtained by an adaptive inverse modeling process applied to the plant. If realized as a transversal filter, the controller will be stable whether the plant is minimum or non-minimum phase. Instabilities in the adaptive process which could arise when an adaptive filter is succeeded by a plant in cascade are overcome by the "filtered X" LMS algorithm. This new form of the LMS algorithm converges to a Wiener solution which is unbiased by plant noise and/or drift. The "filtered-X" LMS algorithm offers a simple and economical solution for a variety of practical problems.

I. Introduction

The LMS algorithm of Widrow and Hoff has proved to be a useful tool in a variety of fields for adjusting the parameters of adaptive systems [1,2]. This paper focuses on the servo problem in the field of adaptive control. The parameters of the controller are adapted so as to make the overall transfer function of the controller-plant best match the transfer function of a given reference model.

The approach taken in this paper differs from the classic stochastic control approach [3]. It was shown [4] that controllers based on stochastic control do not present an appropriate solution for the control problem when the plant has a non-minimum phase characteristic. Self tuning regulators based on a linear quadratic Gaussian criterion [5] can cope with non-minimum phase, but they increase the amount of computation, thus making the solution impractical for real time implementation with small computers.

The "overall transfer function" approach was used in [6]. The algorithm proposed in that paper constructs a controller whose transfer function is the inverse of the plant transfer function. Thus the output of the plant will follow a desired command signal. This controller will fail to converge with non-minimum phase plants.

Another "overall transfer function" approach is the self-tuning controller based on pole zero placement [7]. It can be shown that this controller is an adaptive version of the combination of observer and state feedback [8]. The self-tuning controller based on pole zero placement circumvents the non-minimum phase problem by allowing the non-minimum phase zeroes to remain as part of the overall transfer function. Thus when we have a non-minimum phase plant, the output of the plant cannot be forced to follow the output of the reference model.

An adaptive inverse control scheme that can cope with a non-minimum phase plant without increasing the amount of computation was introduced in [2]. The controller converges to a delayed plant inverse, thus yielding an overall transfer function of pure delay. Though the inverse controller was not designed from a regulator point of view, it was shown that it can eliminate drift at the output of the plant.

In this paper a further extension of [2] is presented. It will be shown that the inverse controller of [2] fails to converge to the true inverse in presence of plant noise. A new noise-immune algorithm called the "filtered-X" LMS algorithm is presented here. The parameters of the controller are adapted so as to achieve an unbiased inverse controller for minimum and non-minimum phase unknown plants. The only information which is needed about the unknown plant is an upper bound on the transport delay of the plant. The performance of the "filtered-X" LMS algorithm in the presence of additive plant noise is discussed.

II. Adaptive Inverse Control with Model Reference

The adaptive control problem can be stated as follows: Take as given a plant with transfer function $H_p(z)$, whose parameters are known or time varying. Then adaptively construct an FIR controller such that the output of the plant c_j will best follow (in the MSE sense) the output of a reference model y_{m_j}. The time index is j.

This work was supported by NSF under grant No. ENG 78-08526 and by ONR under contract No. N00014-76-C-0929.

Reprinted from FIFTEENTH ASILOMAR CONFERENCE ON CIRCUITS, SYSTEMS AND COMPUTERS, November 1981

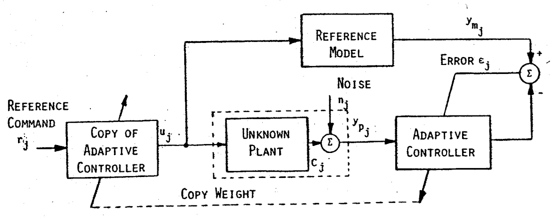

Figure 1. Adaptive reference model controller system.

A basic scheme for constructing an adaptive controller is introduced in Figure 1. This scheme is based on the methodology taught in [2]. Using the LMS algorithm we adapt the parameters of the adaptive controller so as to minimize the mean square of error ε_j. Because the adaptive controller is by definition time varying, it does not have a transfer function. For ease of explanation, let us freeze the parameters of the adaptive controller after they have converged.

At that time

$$H_{\text{unknown plant}} \cdot H_{\text{adaptive controller}} \twoheadrightarrow H_{\text{reference model}} \quad (2.1)$$

where by $\twoheadrightarrow$ we mean "converged to the best match in the MSE sense." From (1) we get

$$H_{\text{adaptive controller}} \twoheadrightarrow \frac{H_{\text{reference model}}}{H_{\text{unknown plant}}} \quad . \quad (2.2)$$

In the special case when the reference model is a pure delay, the transfer function of the adaptive controller converges to a delayed inverse of the unknown plant.

The parameters of the controller are copied and used to construct a series controller. Since the overall transfer function of the controller and the unknown plant matches the transfer function of the reference model, the output of the unknown plant c_j will follow the signal at the output of the reference model y_{m_j}.

III. Effect of Plant Noise on the Convergence of the Adaptive Controller

When the plant is noise-free, the parameters of the adaptive controller converge in accord with (2.2). Once we introduce the possibility of plant noise additive at the output of the plant, the adaptive solution for the parameters of the controller becomes biased from solution (2.2). Thus the overall transfer function of the controller-plant does not match the transfer function of the reference model. See Figure 1.

The optimal noise free Wiener solution for the controller problem is given by

$$W_{opt} = R_c^{-1} \cdot P_c \quad (3.1)$$

where

$$R_c = E[C_j \cdot C_j^T] \quad (3.1a)$$

$$P_c = E[C_j \cdot y_{m_j}] \quad (3.1b)$$

$$C_j = \{c_j, c_{j-1}, c_{j-2}, \ldots, c_{j+1-N}\}^T$$

Due to additive noise at the output of the unknown plant, the parameters of the adaptive controller converge to

$$W_{\text{LMS}} = R_{yp}^{-1} \cdot P_{yp} \quad (3.2)$$

where

$$R_{yp} = E[(C_j+N_j) \cdot (C_j+N_j)^T] \quad (3.2a)$$

$$P_{yp} = E[(C_j+N_j) \cdot y_{m_j}] \quad (3.2b)$$

$$N_j = \{n_j, n_{j-1}, n_{j-2}, \ldots, n_{j+1-N}^T\} \quad .$$

Under the assumption that the input is uncorrelated with the noise, we get

$$R_{yp} = E[C_j C_j^T] + E[N_j N_j^T] \quad (3.3a)$$

$$= R_c + R_n$$

$$R_{yp} = E[(C_j+N_j)\, y_{m_j}] = P_c \quad (3.3b)$$

Using (3.2) and (3.3) we get

$$W_{\text{LMS}} = [R_c+R_n]^{-1} \cdot P_c \quad . \quad (3.4)$$

This solution is biased from (2.2) as a result of plant noise. Using the "ABCD Lemma" of linear algebra,

$$[A+BCD]^{-1} = A^{-1} - A^{-1}B[DA^{-1}B + C^{-1}]^{-1} DA^{-1} \quad (3.5)$$

we get

$$W_{\text{LMS}} = R_c^{-1}P_c - R_c^{-1}R_N[R_c^{-1}R_N + I]^{-1} R_c^{-1}P_c \quad (3.6a)$$

$$W_{\text{LMS}} = W_{opt} - R_c^{-1}R_N[R_c^{-1}R_N + I]^{-1} \cdot W_{opt} \quad (3.6b)$$

$$W_{\text{LMS}} = W_{opt} - bias \quad (3.6c)$$

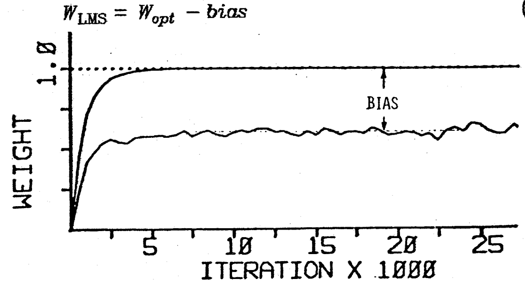

Figure 2. Learning graph of the adaptive delayed inverse model.

The term $R_c^{-1}R_N$ has the interpretation of a matrix version of the noise to signal ratio. This ratio, which is a matrix rather than a simple scalar, is a function of the spectrum of the output c_j and the noise n_j.

Figure 2 shows the convergence of one normalized weight of a simulated adaptive inverse model controller. As can be seen from Figure 2, for a noise-free plant the weight converges to the true inverse solution. However, when additive noise is present at the output of the plant, the weight converges to a biased solution. The bias in the weight can be calculated from (3.6b). Increasing the noise to signal ratio (without changing either the spectrum of the noise or the spectrum of the signal) increases the bias in the parameters of the controller. Another effect of the additive noise at the output of the plant is an increase in the noise component of the controller parameters. This effect can be reduced by reducing the adaption step size μ, consequently making the adaption process slower.

IV. The Filtered-X LMS Algorithm.

The problem of plant noise has been attacked through a fresh approach to inverse modeling. This has motivated the development of a new algorithm: the "filtered-X" LMS algorithm. This method allows adaptation of the inverse filter placed forward of the plant in the cascade sequence. The adaption input is first prefiltered by some $H(z)$. Plant noise does not appear in the adaptive filter input when this algorithm is practiced. But plant noise clearly is a component of the error ε_j, as seen in Fig. 3. In accord with Wiener theory, this noise will have no effect on the converged solution for the inverse filter.

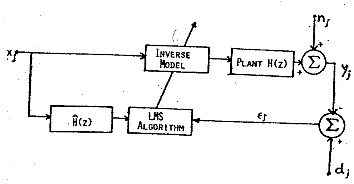

FIGURE 3. THE FILTERED-X LMS ALGORITHM.

The mean square of the error ε_j in Fig. 3 will be a quadratic function of the adaptive filter weights. Therefore, adaptation has the potential for smooth convergence with a unimodal mean square error function.

We now show that the filtered-X LMS algorithm causes convergence in the mean to the inverse filter.

The weight vector update equation is

$$W_{j+1} = W_j + 2\mu\varepsilon_j(\hat{h}_j * X_j^T) \tag{4.1}$$

$$= W_j + 2\mu(d_j - h_j * W_j^T X_j)(\hat{h}_j * X_j^T)$$

$$= W_j + 2\mu d_j \hat{h}_j * X_j^T - 2\mu h_j * W_j^T X_j X_j^T * \hat{h}_j \tag{4.2}$$

where $*$ denotes the convolution operator and it is understood that the Z -transform of h_j is $H(z)$.

From (4.1) we get

$$W_j = W_o + 2\mu\sum_{i=0}^{j-1} \varepsilon_i (X_i^T * \hat{h}_i) \tag{4.3}$$

Thus W_j depends on the values $X_{j-1}, X_{j-2}, \cdots, X_o$.

Also

$$W_j X_i = W_o X_i + 2\mu\sum_{i=0}^{j-1} \varepsilon_i (X_i^T * \hat{h}_i) X_i$$

$$E(W_j X_i) = E(W_o X_i) + 2\mu E\left(\sum_{i=0}^{j-1} \varepsilon_i (X_i^T * \hat{h}_i) X_i\right) \tag{4.4}$$

If X_j is a white sequence with zero mean, then X_j is uncorrelated with X_{j-1} etc. and therefore W_j is uncorrelated with X_j. But since X_j has been filtered by $\hat{H}(z)$ we have essentially the situation of correlated inputs for which $E(W_j X_i)$ is not zero. From (4) $E(W_j X_i)$ is proportional to μ; also $E(W_j X_i)$ decays exponentially with $|i-j|$ for large $|i-j|$. Since X_j has zero mean we can make the correlation of $W_j X_i$ as small as we like by a suitable choice of μ. That is, for μ small enough W_j and X_i are essentially uncorrelated. With this assumption taking expectations in Eq. (4.2) yields,

$$E(W_{j+1}) = E(W_j) + 2\mu E(d_j \hat{h}_j * X_j^T) - 2\mu E(W_j) E(h_j * X_j X_j^T * \hat{h}_j)$$

$$E(W_{j+1}) = E(W_j) + 2\mu E(d_j \hat{Y}_j^T) - 2\mu E(W_j) E(Y_j \hat{Y}_j^T) \tag{4.5}$$

where $\hat{Y}_j = \hat{h}_j * X_j$.

Assume $E(Y_j \hat{Y}_j^T)$

(1) is finite (requires H(z) to be stable)

(2) is non-singular (requires some correlation between ε_j and $\hat{y}_j$). This implies that the model should have at least one weight of delay n where n is the number of pure delays in the plant.

Therefore from [9]

$$\lim_{j\to\infty} E(W_{j+1}) = R_{\hat{y}}^{-1} P_{\hat{y}},$$

where

$$R_{\hat{y}} = E(Y_j \hat{Y}_j^T)$$

$$P_{\hat{y}}^T = E(d_j \hat{Y}_j^T)$$

Assume that the X vector is derived from a double sided, doubly infinite tapped delay line. Then the Z transform of the optimum weight vector sequence is

$$H^{-1}(z)\hat{H}^{-1}(z^{-1})\Phi_{xx}^{-1}(z)\Phi_{xd}(z)\hat{H}(z^{-1})$$

$$= H^{-1}(z)\Phi_{xx}^{-1}(z)\Phi_{xd}(z) \tag{4.6}$$

which is the Wiener solution.

An alternative form of the filtered-X LMS algorithm is shown in figure 4. This structure has similar properties to the configuration of figure 3 since

$$W_{j+1} = W_j + 2\mu\varepsilon_j(\hat{Y}_j)$$

$$= W_j + 2\mu(d_j - W_j^T Y_j)\hat{Y}_j,$$

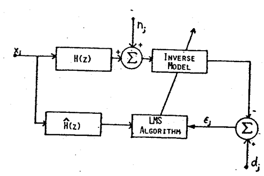

FIGURE 4. ALTERNATIVE FILTERED-X LMS ALGORITHM.

which is the same as Eq. (4.2). Hence similar convergence results hold here. Both structures converge to the same solution, which is unaffected by additive plant noise. It suffices to show that the orthogonality condition of Wiener filters [10],

$E(\varepsilon_j x_j) = 0$, leads to the following Wiener solution .

$$E((d_j - W_j^T Y_j + n_j)\hat{Y}_j) = 0$$

$$\text{i.e. } W^* = R^{-1}P,$$

assuming n_j and X_j are uncorrelated and

$$R = R_{yy}$$

$$P = R_{yd}$$

V. Bounds on Covergence Rate

We stated previously that it was necessary to filter X_j such that ε_j and X_j' (filtered X_j) are correlated. We now reason that a good choice of X_j' is simply to delay X_j by the same amount as the pure delay in the plant.

If the only *a priori* knowledge available is a bound on the plant delay, then it is necessary that the delays of the components of X_j' at least equal that bound. Since the speed of convergence of the adaptation process is limited by the eigenvalue disparity in the covariance matrix [9,13], we investigate the effects of choice of $\hat{H}$ on this matrix.

If the plant transfer function is $z^{-q}H(z)$ where $H(z)$ is any rational transfer function, and $\hat{H}(z) = z^{-q}$,

then

$$Y_j\hat{Y}_j^T = \begin{bmatrix} y_0\hat{y}_0 & y_0\hat{y}_1 & \cdots & y_0\hat{y}_n \\ y_1\hat{y}_0 & y_1\hat{y}_1 & \cdots & y_1\hat{y}_n \\ \vdots & \vdots & & \vdots \\ y_n\hat{y}_0 & y_n\hat{y}_1 & \cdots & y_n\hat{y}_n \end{bmatrix}$$

Taking Expectations gives

$$E(Y_j\hat{Y}_j)^T = \begin{bmatrix} h_0 & h_1 & \cdots & h_n \\ 0 & h_0 & \cdots & h_{n-1} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & h_0 \end{bmatrix}$$

since

$$E(y_j\hat{y}_i) = 0 \qquad \text{for } i < j$$

$$= h_{i-j} \qquad \text{for } i \geq j$$

All the eigenvalues are the same, which implies that, subject to the operating assumptions for a small fixed μ, the speed of convergence will not be improved by another choice of $H(z)$.

It is common to choose $\hat{H}_j = H_j$. This leads to

$$E(Y_j\hat{Y}_j) = \begin{bmatrix} h_0^2 & h_0h_1 & \cdots & h_0h_n \\ h_1h_0 & h_0^2 & \cdots & h_1h_n \\ \vdots & \vdots & & \vdots \\ h_nh_0 & h_nh_1 & \cdots & h_0^2 \end{bmatrix}$$

where all eigenvalues are in general not the same.

In practice the LMS algorithm will usually be operated with a small μ to maintain a small misadjustment [11]. With small μ the above analysis gives bounds on the convergence rate of the filtered-X LMS algorithm in terms of the eigenvalues of the covariance matrix. The misadjustment of the filtered- X LMS algorithm is currently under study.

VI. Simulation of Adaptive Inverse Control With Model Reference.

The idea of model reference control [12] is to build, design, or to adapt a system in such a way that its overall input-output response characteristic best matches a reference model response or some form of ideal response. It is easy to include a model-reference feature in adaptive inverse control systems of the type shown in Fig. 3. The idea is expressed in Fig. 5.

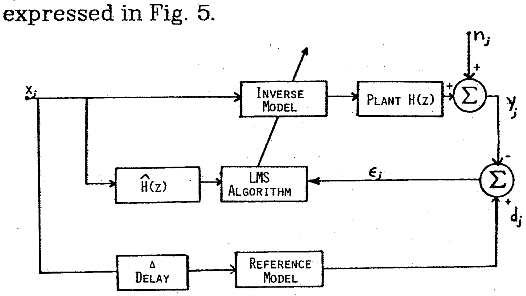

FIGURE 5. ADAPTIVE INVERSE CONTROL WITH MODEL REFERENCE.

It is important that a reference model response be chosen that can be accurately realized by the cascaded adaptive filter and the plant $P(z)$, given that the weights of the adaptive filter are set to minimize mean square error. The system of Fig. 5 will perform remarkably well as long as it is given a feasible task. It should not be asked to respond faster or more intricately than it is able for the given plant $P(z)$ and its FIR adaptive transversal controller. If the plant has transport delay or is non-minimum phase, inclusion of a delay Δ will give more precise model tracking but a delayed response.

An experiment was performed with the model reference adaptive inverse control system of Fig. 5.

$$\text{Plant:} \qquad \frac{2.4z^{-1}(1-.8z^{-1})}{(1+.6z^{-1})(1-.7z^{-1})}$$

$$\text{Reference Model:} \qquad \frac{.25z^{-1}}{(1-.5z^{-1})^2}$$

Thirteen model weights and forty five controller weights were used. The input to the entire system consisted of white noise. The step response of the uncompensated plant is given in Fig. 6. Figure 7 shows the step response of the compensated plant superimposed on the step response of the reference model. The quality of the fit depends on the number of weights one allows for in the controller. As is evident from the figure a very close match between the model output and the entire control system output can be obtained with only a moderate number of weights.

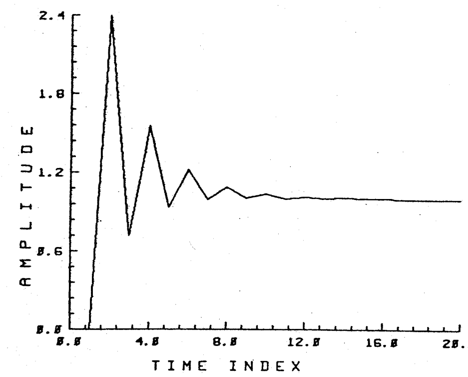

FIGURE 6. STEP RESPONSE OF UNCOMPENSATED PLANT.

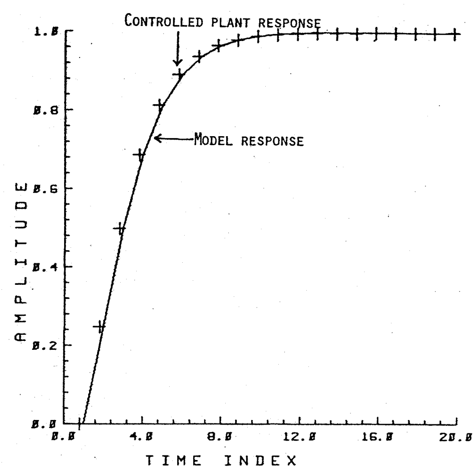

FIGURE 7. CONTROL SYSTEM STEP RESPONSE WITH NOISE FREE PLANT.

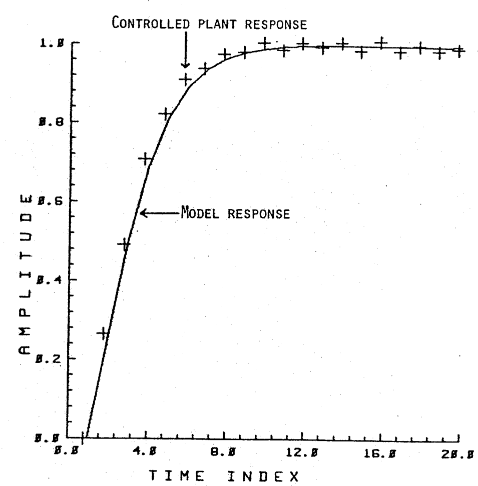

FIGURE 8. CONTROL SYSTEM STEP RESPONSE WITH NOISY PLANT.

The adaption was repeated with additive white noise at the plant output of normalized amplitude 0.1 (normalized with respect to the system input). Figure 8 shows the step response using the "noisy" weights. As expected there is a degradation proportional to the amount of plant noise present.

VII. Conclusions

A method for adaptive inverse control unbiased by additive plant noise has been introduced. The technique is easy to implement and exhibits robust, predictable behaviour. Research is ongoing in this area and further results will be reported in the future.

VIII. References

[1] B. Widrow et al., "Adaptive Noise Cancelling: Principles and Applications," *Proc. IEEE*, Vol. 63, pp. 1692-1716, December 1975.

[2] B. Widrow, J. McCool, and B. Medoff, "Adaptive Control by Inverse Modeling," Conf. Rec. of 12th Asilomar Conference on Circuits, Systems and Computers, pp. 90-94, November 1978.

[3] K. Astrom, *Introduction to Stochastic Control*, New York, Academic Press, 1970.

[4] K. Astrom, U. Borisson, L. Ljung, and B. Wittenmark, "Theory and Application of Self Tuning Regulators," *Automatica*, Vol. 13, pp. 457-476, 1977.

[5] K. Astrom and B. Wittenmark, "Analysis of a Self-Tuning Regulator for Nonminimum Phase Systems," IFAC Symposium on Stochastic Control, Budapest, Hungary, 1974.

[6] J. Martin-Sanchez, "A New Solution to Adaptive Control," *Proc. IEEE*, Vol. 64, No. 8, pp. 1209-1218, August 1976.

[7] K. Astrom and B. Wittenmark, "Self-Tuning Controllers Based on Pole-Zero Placement," *Proc. IEE*, Vol. 127, Pt. D, No. 3, pp. 120-130, May 1980.

[8] T. Kailath, *Linear Systems*, Prentice-Hall, 1980.

[9] B. Widrow, P. Mantey, L. Griffiths and B. Goode, "Adaptive Antenna Systems," *Proc. IEEE*, Vol. 55, No. 12, pp. 2143-2159, December 1967.

[10] T. Kailath, *Lectures on Linear Least-Squares Estimation*, Springer Verlag, Wien-New York, 1976.

[11] B. Widrow, "On the Speed of Convergence of Adaptive Algorithms," Conf. Rec. of 14th Asilomar Conference on Circuits, Systems and Computers, pp. 239-242, November 1980.

[12] I. Landau, "A Survey of Model Reference Adaptive Techniques-Theory and Applications," *Automatica*, Vol. 10, pp. 353-379, 1974.

[13] D. Leuenberger, *Introduction to Linear and Nonlinear Programming*, Addison Wesley, 1973.