ESTIMATING THE AUTOCORRELATION FUNCTION OF AN ARBITRARILY TIME-VARIANT SYSTEM RESPONSE

Lang Peng

Northeastern University
Boston, Massachusetts 02115
peng.l@husky.neu.edu

Hanoch Lev-Ari

Northeastern University
Boston, Massachusetts 02115
levari@ece.neu.edu

ABSTRACT

We present a technique for determining the autocorrelation of an arbitrarily time-variant system response. Our approach relies on a key relation, introduced in [1], between the system response autocorrelation function and certain 2nd and 4th order moments of the system input and (noisy) output signals, with no other prior information about the dynamics of the system response required. We introduce a "Wiener problem" interpretation of this key relation, which enables us to benefit from the wealth of existing results about the dynamics and performance of standard adaptive filters. In particular, we propose time-recursive estimates for the system response autocorrelation, with significantly reduced computational cost, as compared to previously proposed (non-recursive) estimates. Moreover, our procedure can also be customized to track (slow) variations in the system response autocorrelation when such variations are present. We use an example to demonstrate the advantage of applying standard adaptive algorithms such as LMS, NLMS or RLS to obtain an estimate of the desired system response autocorrelation.

***Index Terms*—** Autocorrelation, time-variant system

1. INTRODUCTION

Time-variant systems are often encountered in engineering applications, ranging from underwater acoustic communications to array processing. The inability of conventional (i.e., stationary-based) system identification techniques to cope with rapid time variation has motivated the development of novel approaches to identify linear, arbitrarily time-variant systems.

Several approaches were suggested for solving this non-stationary identification problem. Some, such as [2,3] involve a statistical characterization of time-variant impulse response in terms of its average power spectrum or, equivalently, its average autocorrelation. Others, such as in [4, 5], involve a dynamic characterization of the time-variant system response in terms of a state-space model. In either case, one is forced to rely on substantial prior information about the nature of the time-variation of the system response.

Our objective is to estimate the necessary prior information, such as the average autocorrelation of the time-variant system response, from measurements of the input and output signals. We assume, as in [1], that the input-output relation of the discrete-time time-variant system is FIR, viz.,

$$d(t) = W(t)U(t) + v(t) \tag{1}$$

where $W(t) = [w_0(t)\ w_1(t) \cdots w_{M-1}(t)], U(t) = [u(t)\ u(t-1) \cdots u(t-M+1)]^T$, $u(t)$ is the scalar input signal and $v(t)$ is additive noise.

A prior-information-free approach for estimation of the autocorrelation of the time variant system response $W(t)$ was introduced in [1,6], based on a novel explicit relation between the autocorrelation of $W(t)$ and certain 2nd and 4th order moments of the signals $u(t)$ and $d(t)$. Such moments can be estimated via long-term averaging directly from measurements of the system input and output signals. This results in a non-recursive (off-line) solution that uses a given signal record of suitable length. The same (non-recursive) solution was also obtained in [5] via an approximate modeling argument that led to a deterministic least squares problems, whose solution provides an estimate for the autocorrelation of $W(t)$.

In this paper we introduce a Wiener problem (i.e., a minimum mean-square error linear estimation) interpretation of the key relation in [1], which enables us to exploit the well-developed theory of linear adaptive filters [7]. In particular we propose time-recursive procedures for estimation of $C_W(\cdot)$, the autocorrelation of $W(t)$, with significantly improved performance, as compared with the static procedures described in [1, 5] (see Sec. 4). To be specific, our procedure:

- offers a dramatic reduction in computational cost
- can efficiently incorporate new measurements as they become available
- can be customized to track (slow) variations in $C_W(\cdot)$ when such variations are present.

In addition we are able to benefit from the wealth of existing results about the dynamics and performance of standard adaptive filters (see Sec. 5). We use an example to demonstrate in Sec. 6 that the normalized LMS algorithm with optimized varying step size (see, e.g. [8]) may offer the best tradeoff between performance and computational cost.

2. PROBLEM FORMULATION

We extend somewhat the conceptual frameworks used in [1] and [5]. Thus we assume that $u(\cdot)$, $W(\cdot)$, $v(\cdot)$ are all random and mutually independent. The input signal $u(\cdot)$ and the additive noise $v(\cdot)$ are zero-mean stationary (in the narrow sense). The system response $W(\cdot)$ can be stationary, asymptotically-mean-stationary [9] or slowly-varying non-stationary, and its elements can be correlated with each other.

In this paper we consider two distinct scenarios of estimating the autocorrelation of the system response $W(t)$:

- a *static* scenario in which $W(t)$ is stationary or, more generally, asymptotically-mean-stationary, and we wish to estimate the *average autocorrelation*

  $$C_W(m) \triangleq \langle\, E\{W^*(t)W(t+m)\}\,\rangle_t$$

  where $\langle\ \rangle_t$ denotes a (long-term) time average over $[0\ \infty)$. Notice that when $W(t)$ is stationary, the *time-averaging operation* $\langle\ \rangle_t$ can be dropped.

- a dynamic scenario in which $W(t)$ is non-stationary with slowly-varying statistics, and we wish to estimate and track the *time-varying* autocorrelation

  $$C_W(m;t) \triangleq E\{W^*(t)W(t+m)\}$$

  for every time-instant "*t*".

In both scenarios we make no other prior assumptions about the nature of time-variation of $W(t)$, and we use only measurements of the input signal $u(\cdot)$ and the output signal $d(\cdot)$ to form our autocorrelation estimates.

The autocorrelation $C_W(\cdot)$ provides sufficient prior statistical information for those approaches mentioned before to identify $W(t)$. The static scenario was addressed in [1] and [5]. In particular, $C_W(m)$ was shown in [1] to satisfy a matrix equation, viz.,

$$\mathcal{M}_{\mathcal{U}} \begin{pmatrix} c_v(m) \\ \mathrm{vec}\{C_W(m)\} \end{pmatrix} = \begin{pmatrix} c_d(m) \\ \mathrm{vec}\{C_\xi(m)\} \end{pmatrix} \tag{2a}$$

$$\mathcal{M}_{\mathcal{U}} \triangleq \begin{pmatrix} 1 & [\mathrm{vec}\{C_U^*(m)\}]^* \\ \mathrm{vec}\{C_U^*(m)\} & \Gamma_U(m) \end{pmatrix} \tag{2b}$$

where the asterisk $(*)$ denotes complex transposition and $\mathrm{vec}\{\cdot\}$ denotes vectorization of a matrix by columns [10]. The linear equation (2) relies on several (average) moments that can all be estimated from measurements of the system input and output. Thus, $C_U(m)$ is the autocorrelation of the stationary vector $U(t)$, viz.,

$$C_U(m) \quad \triangleq \quad E\{U(t+m)U^*(t)\}$$

and

$$\Gamma_U(m) = E\left\{\left[\widetilde{U}(t+m)\otimes U(t)\right]\left[\widetilde{U}(t+m)\otimes U(t)\right]^*\right\}$$

is a fourth-order moment of $U(t)$, where the tilde $(\sim)$ denotes element-wise complex conjugation, and $\otimes$ denotes the Kronecker product [10]. Similarly, $c_d(m) \triangleq \langle\, E\{d(t+m)d^*(t)\}\,\rangle_t$ is the average autocorrelation of the non-stationary output signal $d(\cdot)$, and

$$C_\xi(m) \triangleq \langle\, E\{\xi^*(t)\xi(t+m)\}\,\rangle_t$$

is the average autocorrelation of the non-stationary *composite signal* $\xi(t) \triangleq d(t)U^*(t)$.

The linear equation (2) was originally derived under the assumption that $W(t)$ is deterministic (see, e.g. [1]). However, the same relations hold in the more general case when $W(t)$ is random, asymptotically-mean-stationary, possibly with a non-vanishing mean. This more general characterization of $W(t)$ is needed, for instance, in the discussion of fading communication channels [11].

In fact, the equation (2) holds even in the dynamic case, but with time-dependent versions of $C_W(\cdot)$, $c_d(\cdot)$ and $C_\xi(\cdot)$. This observation allows us to apply adaptive (time-recursive) algorithms to track the time-variation of $C_W(m;t)$ (see Sec. 4). Moreover, the same algorithms can also be used to provide a computationally-efficient solution in the static case, offering a significant cost reduction as compared to the non-recursive solutions described in [1] and [5].

Our approach to constructing time-recursive solutions for the linear equation (2) relies on a minimum-mean-square-error (MMSE) linear estimation interpretation of this equation which we introduce in Sec. 3. We show that (2) is a so-called discrete-time Wiener-Hopf equation for a suitably defined "Wiener problem". This observation puts at our disposal a large body of techniques and results from the theory of adaptive linear filters [7]. In particular it offers a rich selection of computationally efficient implementations and some closed-form results about the steady state and tracking performance of our $C_W(\cdot)$ estimate (see Sec. 5).

3. LEAST-SQUARES INTERPRETATION OF (2)

We observe that $\mathcal{M}_{\mathcal{U}}$ can be interpreted as

$$\mathcal{M}_{\mathcal{U}} = E\left\{\begin{pmatrix} 1 \\ \Psi_m(t) \end{pmatrix}\begin{pmatrix} 1 \\ \Psi_m(t) \end{pmatrix}^*\right\}$$

where $\Psi_m(t) \triangleq \widetilde{U}(t+m)\otimes U(t)$ is the Kronecker product of two column vectors. Thus $\mathcal{M}_{\mathcal{U}}$ is the autocorrelation matrix of $\mathcal{U}(t) \triangleq \begin{pmatrix} 1 \\ \Psi_m(t) \end{pmatrix}$, as pointed out in [1]. We now also observe that the right-hand-side of (2a) is in fact equal to

$$E\left\{\begin{pmatrix} 1 \\ \Psi_m(t) \end{pmatrix} \mathcal{D}_m^*(t)\right\} \triangleq \mathcal{R}_{\mathcal{U}\mathcal{D}}(m)$$

where $\mathcal{D}_m(t) \triangleq d^*(t+m)d(t)$. Hence (2) is the Wiener-Hopf equation for the MMSE problem

$$\min_{H,b} E\left|\mathcal{D}_m(t) - \left[\begin{array}{cc} b & H \end{array}\right]\begin{pmatrix} 1 \\ \Psi_m(t) \end{pmatrix}\right|^2$$

whose (optimal) solution is (recall that H is a row vector)

$$b_{opt} = c_v^*(m), \qquad H_{opt} = [\text{vec}\{C_W(m)\}]^*$$

Alternatively, we can write this as

$$\min_{H,b} E\left|\mathcal{D}_m(t) - \Big(H\Psi_m(t) + b\Big)\right|^2 \tag{3}$$

which is a linear MMSE estimation problem (a.k.a, "Wiener problem") for non-zero-mean random variables.

Since (3) involves non-zero-mean random variables, the corresponding Wiener-Hopf equation determines both the additive constant "b" and the coefficient vector "H" (equivalently both $c_v(m)$ and $\text{vec}\{C_W(m)\}$). We shall refer to this format as the *non-centered* Wiener-Hopf equation. It can be replaced by two separate equations, one for "b" alone, and the other one for "H" alone. The resulting *centered* Wiener-Hopf equation is

$$E\{\overline{\Psi_m(t)} \quad \overline{\Psi_m(t)}^*\}H^* = E\{\overline{\Psi_m(t)} \quad \overline{\mathcal{D}_m(t)}^*\} \tag{4}$$

where $\overline{\Psi_m(t)}$, $\overline{\mathcal{D}_m(t)}$ are centered versions of these random variables, namely,

$$\begin{aligned} \overline{\Psi_m(t)} &\triangleq \Psi_m(t) - E\Psi_m(t) \\ \overline{\mathcal{D}_m(t)} &\triangleq \mathcal{D}_m(t) - E\mathcal{D}_m(t) \end{aligned}$$

This is the standard version of the Wiener-Hopf equation presented in all textbooks (see, e.g. [7]) and it corresponds to the centered Wiener problem

$$\min_{H} E\left|\overline{\mathcal{D}_m(t)} - H\,\overline{\Psi_m(t)}\right|^2 \tag{5}$$

4. ADAPTIVE SOLUTIONS

Based on our Wiener problem interpretation (3), we propose to use standard adaptive algorithms to provide a (time-recursive) solution for the linear equation (2). We present here the recursions for normalized LMS and for RLS, which we then use in Sec. 6 in our numerical example.

NLMS with time varying step size [8]

The NLMS algorithm with time varying step size for the Wiener problem (3) is

$$\begin{aligned} e_m(t) &= \mathcal{D}_m(t) - H_m(t-1)\Psi_m(t) \\ H_m(t) &= H_m(t-1) + \mu(t)e_m(t)\left[\Psi_m^*(t)\Psi_m(t)\right]^{-1}\Psi_m^*(t) \\ \mu(t) &= \mu(t-1)\,\frac{1-\mu(t-1)/M^2}{1-\mu^2(t-1)/M^2} \end{aligned} \tag{6}$$

where M^2 is the length of the vector $\Psi_m(t)$ and the initial value of μ can be chosen as $\mu(0) = 1 - \frac{J_{min}}{\sigma_d^2}$ where J_{min} is the variance of the residual $e_{opt}(t)$ when estimation is assumed to be perfect i.e., $\mathcal{D}_m(t) = H_{m,opt}\Psi_m(t) + e_{m,opt}(t)$ and σ_d^2 is the initial estimation variance.

RLS

The RLS algorithm for the Wiener problem (3) is ($0 < \lambda < 1$)

$$\begin{aligned} \pi_m(t) &= P_m(t-1)\Psi_m(t) \\ k_m(t) &= \frac{\pi_m(t)}{\lambda + \Psi_m^*(t)\pi_m(t)} \\ e_m(t) &= \mathcal{D}_m(t) - H_m(t-1)\Psi_m(t) \\ H_m(t) &= H_m(t-1) + k_m(t)e_m^*(t) \\ P_m(t) &= \lambda^{-1}P_m(t-1) - \lambda^{-1}k_m(t)\Psi_m^*(t)P_m(t-1) \end{aligned}$$

with initial values $H_m(0) = 0$ and $P_m(0) = \delta^{-1}I$.

Using time-recursive solutions for (2) results in a significant reduction in computational cost (see Sec. 5). It also makes it possible to track variations of $C_W(m;t)$ in the dynamic version of our problem.

5. ACCURACY ANALYSIS

In the static case the variance of each element of the estimate $\widehat{C}_W(m)$ obtained by a non-recursive approach is inversely proportional to N, the length of our data record (for large N values). In fact, it is known [12] that the probability density function of the scaled estimation error $\sqrt{N}\text{vec}\{\widehat{C}_W(m) - C_W(m)\}$ converges, as $N \to \infty$, to a Gaussian zero-mean distribution with covariance equal to $J_{min}\mathcal{R}_U^{-1}(m)$, where $\mathcal{R}_U(m)$ is the centered covariance of $\Psi_m(t)$ and J_{min} is the minimal achievable value of the Wiener problem cost function (3). This provides an explicit characterization for the accuracy of $\widehat{C}_W(m)$, namely for large N we have the explicit error covariance expression $E\Big\{\Big[\text{vec}\{\widehat{C}_W(m) - C_W(m)\}\Big]\Big[\text{vec}\{\widehat{C}_W(m) - C_W(m)\}\Big]^*\Big\} \approx \frac{1}{N}\,J_{min}\mathcal{R}_U^{-1}(m)$.

In the dynamic (slowly-varying) case we can use standard tracking algorithms such as LMS or RLS. Again, there are known results about the accuracy of the resulting $\widehat{C}_W(m;t)$

	$\mathfrak{D}_S$	$\mathfrak{D}_L$
LMS	$\frac{1}{2}\mu J_{min} M^2$	$\frac{1}{2\mu} tr\left[\mathcal{R}_U^{-1} R_{\mathfrak{w}}\right]$
RLS	$\frac{1-\lambda}{2} J_{min} tr\left[\mathcal{R}_U^{-1}\right]$	$\frac{1}{2\mu} tr\left[\mathcal{R}_U^{-1} R_{\mathfrak{w}}\right]$

Table 1. Performance summary of LMS and RLS (small step size)

estimates (see, e.g. [7]). In particular, the steady-state variance of elements of $\widehat{C}_W(m;t)$ is proportional to the step size μ for LMS, and to $(1-\lambda)$ for exponentially weighted RLS. Table 1 gives a summary of the performance of LMS and RLS, where $\mathfrak{D}_s$ is the steady state performance , $\mathfrak{D}_L$ is the tracking performance and $R_{\mathfrak{w}}$ is the autocorrelation function of the process $\mathfrak{w}(t)$ that generates $W(t)$ via the random-walk model $W(t) = W(t-1) + \mathfrak{w}(t)$.

6. NUMERICAL EXAMPLES

Fig. 1 and Fig. 2 both use the same set of data: the input signal $u(t)$ and the additive noise $v(t)$ are both zero-mean Gaussian white noise signals, with $c_u(0) = 1$. The memoryless system gain $w(t)$ is generated by passing a white Gaussian signal through a narrow band (linear phase, FIR) lowpass filter, with a cutoff frequency of 0.035. The level of the additive output noise $v(t)$ is adjusted to achieve $SNR = 5$dB. Since the length of the system response is $M = 1$, the autocorrelation $c_w(m)$ is a scalar in our simulation. Estimator variance is calculated by averaging over 200 independent realizations. All estimator variances are normalized by $c_w^2(0)$.

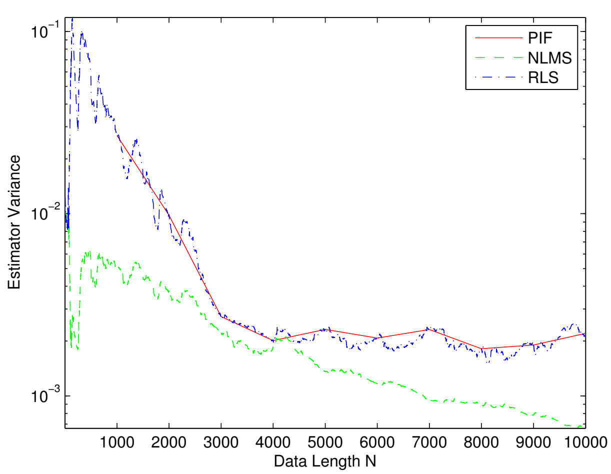

Fig. 1. Convergence speed of three algorithms: non-recursive , NLMS, and RLS for estimating $c_w(2)$

Those figures compare three estimators: (i) non-recursive solution of (2), as proposed in [1,5], (ii) corresponding NLSM with time varying steps [8] and (iii) exponentially-weighted

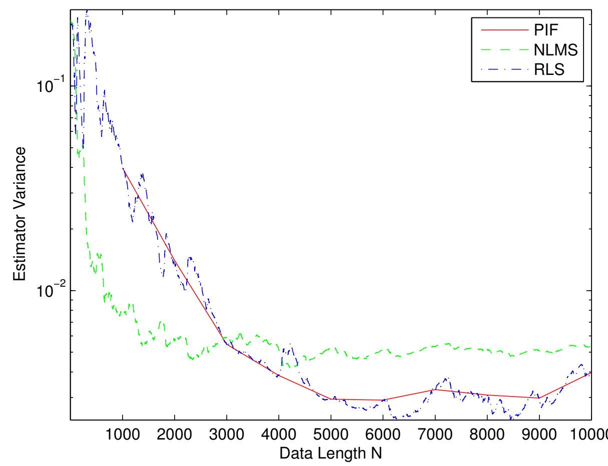

Fig. 2. Convergence speed of three algorithms: non-recursive, NLMS and RLS for estimating $c_w(3)$

RLS, with $\lambda = 0.999$. The results demonstrate the superiority of NLMS over both the non-recursive solution and exponentially-weighted RLS for *short data records*. Lower estimator variances can be sometimes achieved with RLS (see Fig. 2) with sufficiently long data records. Given these mild differences in performance, NLMS appears to be the preferred choice, due to its significantly lower computational cost.

7. CONCLUDING REMARKS AND RELATION TO PRIOR WORK

The only prior work on explicit estimation of $C_W(m)$ from the signal $u(\cdot)$ and $d(\cdot)$, and without using any prior information about the dynamics of $W(t)$, was reported in [1, 5]. Both provided a non-recursive (off-line) estimate of $C_W(m)$, albeit using very different approaches to derive their results. In contrast, in this paper we introduce a Wiener problem interpretation for the linear equation (2), involving the composite signals $\mathcal{D}_m(t)$ and $\Psi_m(t)$, which makes it possible to use standard linear adaptive filtering algorithms to obtain an estimate for $C_W(\cdot)$, in both the static and the dynamic versions of our problem. In particular, the LMS family of adaptive algorithms offers a significant reduction in computational cost — from $\mathcal{O}(M^6) + \mathcal{O}(M^4N)$ to $\mathcal{O}(M^2N)$ for a single lag value — as compared with the offline methods of [1, 5]. In addition, we have relaxed the assumptions made in previous work about the variation of $W(t)$, allowing it to be asymptotically-mean-stationary, or even arbitrarily non-stationary with slowly-varying statistics. Thus our approach achieves a drastic cost reduction, and applies to a wider family of $W(t)$ dynamics, as compared with the methods of [1, 5].

8. REFERENCES

[1] Y. Ehrlich-Steinsaltz and H. Lev-Ari, “Identification of arbitrarily time-variant systems,” in *Proc. of the IEEE International Conference on Acoustics, Speech and Signal Processing*, (Seattle, WA), pp. 2181–2184, May 1998.

[2] J. Lin, J. G. Proakis, F. Ling, and H. Lev-Ari, “Optimal tracking of time-varying channels: a frequency domain approach for known and new algorithms,” *IEEE Transactions on Selected Areas in Communications*, vol. 13, pp. 141–154, Jan 1995.

[3] P. W. Wu and H. Lev-Ari, “Optimized estimation of moments for nonstationary signals,” *IEEE Trans. on Signal Processing*, vol. 45, pp. 1210–1221, May 1997.

[4] S. Haykin, A. H. Sayed, *et al.*, “Adaptive tracking of linear time-variant systems by extended rls algorithms,” *IEEE Transactions on Signal Processing*, vol. 45, pp. 1118–1128, May 1997.

[5] M. K. Tsatsanis, G. B. Giannakis, and G. Zhou, “Estimation and equalization of fading channels with random coefficients,” *Signal Processing*, vol. 53, pp. 211–229, September 1996.

[6] L. Peng and H. Lev-Ari, “Lag-recursive estimation of the average autocorrelation of an arbitrarily time-variant system response,” *Statistical Signal Processing Workshop (SSP)*, pp. 353–356, Aug 2012.

[7] S. Haykin, *Adaptive Filter Theory*. Upper Saddle River, N.J.: Prentice Hall, fourth ed., 2002.

[8] D. T. M. Slock, “On the convergence behavior of the lms and the normalized lms algorithms,” *IEEE Transactions on Signal Processing*, vol. 41, pp. 2811–2825, September 1993.

[9] M. J. Genossar, H. Lev-Ari, and T. Kailath, “Consistent estimation of the cyclic autocorrelation,” *IEEE Trans. Signal Processing*, vol. 42, pp. 595–603, March 1994.

[10] J. W. Brewer, “Kronecker products and matrix calculus in system theory,” *IEEE Transactions on Circuits and Systems*, vol. 25, pp. 722–781, September 1978. issue 9.

[11] P. A. Bello, “Characterization of randomly time-variant linear channels,” *IEEE Transactions on Communicatioins Systems*, vol. CS-11, pp. 360–393, December 1963.

[12] P. J. Brockwell and R. A. Davis, *Time Series: Theory and Methods*. New York, NJ: Springer-Verlag, second ed., 1991.