

LETTER

A New Method of Noise Variance Estimation from Low-Order Yule-Walker Equations

Jonah GAMBA^{†a)}, *Student Member and* Tetsuya SHIMAMURA[†], *Member*

SUMMARY The processing of noise-corrupted signals is a common problem in signal processing applications. In most of the cases, it is assumed that the additive noise is white Gaussian and that the constant noise variance is either available or can be easily measured. However, this may not be the case in practical situations. We present a new approach to additive white Gaussian noise variance estimation. The observations are assumed to be from an autoregressive process. The method presented here is iterative, and uses low-order Yule-Walker equations (LOYWEs). The noise variance is obtained by minimizing the difference in the second norms of the noisy Yule-Walker solution and the estimated noise-free Yule-Walker solution. The noise-free solution is constrained to match the observed autocorrelation sequence. In the iterative noise variance estimation method, a variable step-size update scheme for the noise variance parameter is utilized. Simulation results are given to confirm the effectiveness of the proposed method.

key words: noise variance, Yule-Walker equations, autoregressive process, subspace method

1. Introduction

Autoregressive (AR) parameter estimation has been extensively studied in the literature [1]–[3]. It finds application in various fields including forecasting, economics, and speech processing. Parameter estimation of AR processes corrupted by white Gaussian noise requires *a priori* knowledge of the noise variance or it has to be estimated from an available data sequence. In practical situations, however, the noise variance cannot be known *a priori*.

The method proposed in [5] assumes that the noise variance is available before parameter estimation. Many methods have been proposed to address the problem of noise variance estimation from an observed data sequence [6]–[8]. These methods rely on different strategies such as the use of prefiltering, higher-order Yule-Walker equations (HOYWEs), modified least-squares methods and eigendecomposition. The best performance is commonly achieved at high signal-to-noise ratios (SNRs).

One recent method of noise variance estimation is subspace (SS) approach [8]. In [4], the uniqueness

of the solution via the SS method has been discussed, and in [8] the superiority of that to the other effective methods has been demonstrated. However, in addition to requiring the use of both LOYWEs and HOYWEs, the SS method further have to deal with a computationally intensive quadratic eigenvalue problem to obtain an estimate of the noise variance. Therefore, we herein propose a simpler alternative, which relies *only* on LOYWEs. The proposed method is iterative and utilizes the noisy Yule-Walker equation. At each iteration an estimate of the noise-free Yule-Walker solution is obtained by adjusting the noise variance parameter and constraining the solution to match the autocorrelation sequence of the observed data sequence. The adjustment is achieved through a step-size parameter which depends on the current value of the noise variance parameter. The true noise variance minimizes the difference between the second norms of the noisy Yule-Walker solution and the estimated noise-free Yule-Walker solution.

This paper is organized as follows. In Sect.2 we outline the noisy AR model. The effect of additive noise on AR parameters is highlighted. Section 3 presents the proposed iterative noise variance estimation (INVE) method. In Sect.4 simulation results are presented. Concluding remarks in Sect.5 end this paper.

2. Noisy AR Model

A stationary AR process $\{x(n)\}$ of order p is defined by $\sum_{i=0}^p a(i)x(n-i) = e(n)$, where $\{e(n)\}$ is an uncorrelated driving white noise sequence of variance σ_e^2 and $a(i)$'s are the noise-free AR parameters ($a(0) = 1$). The autocorrelation function (ACF) at lag k for $\{x(n)\}$ is defined by $r_{xx}(k) = E[x(n)x(n+k)]$ where E is the expectation operator. The $r_{xx}(k)$ is given in the above case by $r_{xx}(k) = -\sum_{i=1}^p a(i)r_{xx}(k-i) + \delta(k)\sigma_e^2$, $k \geq 0$, where $\delta(k)$ is the Kronecker delta function. In the presence of noise, the observed data sequence becomes $y(n) = x(n) + w(n)$, where $w(n)$ is assumed to be zero-mean additive white Gaussian noise of variance σ_w^2 . The ACF for $\{y(n)\}$ is similarly defined like that for $\{x(n)\}$ and denoted by $r_{yy}(k)$.

The commonly used Yule-Walker equations (YWEs) [1] are

$$\mathbf{R}_{xx}\mathbf{a} = -\mathbf{r}_x \quad (1)$$

Manuscript received April 4, 2003.

Manuscript revised August 1, 2003.

Final manuscript received September 22, 2003.

[†]The authors are with the Department of Information and Computer Sciences, Saitama University, Saitama-shi, 338-8570 Japan.

a) E-mail: jonah@sie.ics.saitama-u.ac.jp

$$\mathbf{R}_{yy}\hat{\mathbf{a}} = -\mathbf{r}_y \quad (2)$$

where \mathbf{R}_{xx} and \mathbf{R}_{yy} are $p \times p$ autocorrelation matrices (ACMs) of the sequences $x(n)$ and $y(n)$ respectively. The column vectors on the right hand sides of Eqs. (1) and (2) are $\mathbf{r}_x^T = [r_{xx}(1) \dots r_{xx}(p)]$ and $\mathbf{r}_y^T = [r_{yy}(1) \dots r_{yy}(p)]$. The T denotes the transposition operation. The $p \times 1$ vectors \mathbf{a} and $\hat{\mathbf{a}}$ are the noise-free and noisy solutions to the YWEs, respectively. The AR parameter estimates from Eq. (2) are biased since $r_{yy}(k) = r_{xx}(k) + \delta(k)\sigma_w^2$. Using Eqs. (1) and (2), the following relationship between ACMs is validated for noise compensation,

$$\mathbf{R}_{xx} = \mathbf{R}_{yy} - \sigma_w^2 \mathbf{I} \quad (3)$$

where \mathbf{I} is a $p \times p$ identity matrix.

3. Noise Variance Estimation

3.1 Proposed Method

The above relationships for a noisy AR model can be used to obtain an estimate of the noise variance σ_w^2 as follows. From Eq. (1)–Eq. (3) and the fact that $\mathbf{r}_x = \mathbf{r}_y$, we have

$$\mathbf{R}_{yy}(\hat{\mathbf{a}} - \mathbf{a}) = -\sigma_w^2 \mathbf{a}. \quad (4)$$

Taking the norms (any norm) of both sides of Eq. (4) results in the relationship

$$\sigma_w^2 \|\mathbf{a}\| = \|\mathbf{R}_{yy}(\hat{\mathbf{a}} - \mathbf{a})\|. \quad (5)$$

Equation (5) shows that the noise variance is a function of the Euclidean distance between $\hat{\mathbf{a}}$ and \mathbf{a} .

Letting $\Delta \mathbf{a} = \hat{\mathbf{a}} - \mathbf{a}$, Eq. (5) can be expressed more compactly as

$$\sigma_w^2 \|\mathbf{a}\| \leq \|\mathbf{R}_{yy}\| \|\Delta \mathbf{a}\|. \quad (6)$$

Working with the square norm, it is well known [9] that $\|\mathbf{R}_{yy}\| = \lambda_{y_{max}}$, where $\lambda_{y_{max}}$ is the maximum eigenvalue of \mathbf{R}_{yy} . Equation (6) can therefore be written as

$$\sigma_w^2 \|\mathbf{a}\| \leq \lambda_{y_{max}} \|\Delta \mathbf{a}\| \leq \lambda_{y_{max}} \|\hat{\mathbf{a}}\| - \|\mathbf{a}\|. \quad (7)$$

From Eq. (7), it is possible to obtain an estimate of σ_w^2 . To achieve this, a function

$$f(\alpha) = \|\hat{\mathbf{a}}\| - \|\tilde{\mathbf{a}}(\alpha)\| \quad (8)$$

is defined where α is a parameter that gives an estimate of σ_w^2 and $\tilde{\mathbf{a}}(\alpha)$ corresponds to the solution of noise-compensated Yule-Walker equation obtained by combining Eq. (1) with Eq. (3). The value of α that gives the minimum of $f(\alpha)$ results in the noise variance estimate. This property of the function $f(\alpha)$ is described analytically in Appendix. In the proposed method, by adjusting the α in an iterative fashion, the minimum of $f(\alpha)$ is sought. The full INVE method is described below.

The INVE Method

1. Obtain an estimate of the biased noisy autocorrelation sequence $\{\hat{r}_{yy}(k)\}$ as

$$\hat{r}_{yy}(k) = \frac{1}{N} \sum_{n=0}^{N-1-|k|} y(n)y(n+|k|) \quad (9)$$

where N is the length of the observed noisy data sequence, and solve Eq. (2) by the Levinson-Durbin algorithm to obtain an $\hat{\mathbf{a}}$.

2. Set M as a large real number and initialize the noise variance parameter and the step-size parameter as

$$\alpha = \frac{1}{M} \quad (10)$$

$$s = \frac{\alpha}{M} \quad (11)$$

respectively.

3. Substitute Eq. (3) into Eq. (1) with the setting of $\sigma_w^2 = \alpha$ and solve Eq. (1) to obtain an estimate of \mathbf{a} denoted by $\tilde{\mathbf{a}}(\alpha)$.
4. Calculate Eq. (8).
5. If $\alpha \geq \hat{r}_{yy}(0)$, terminate and obtain the value of α for which $f(\alpha)$ is minimum. Otherwise, go to Step 6.
6. Calculate s according to

$$s = \begin{cases} \hat{r}_{yy}(0)/M, & \text{if } \alpha \geq 1 \\ \alpha/M, & \text{otherwise} \end{cases} \quad (12)$$

and increase α by the value of s as

$$\alpha = \alpha + s. \quad (13)$$

7. Go to Step 3.

3.2 Step-size Adjustment

The INVE method requires the step-size parameter s to increase the value of α at each iteration. The use of a constant step-size parameter will result in high variance in the estimated values of noise variance. This phenomenon is especially pronounced at high SNRs. To understand the increase in variance we observe that if the step value is kept constant at a value for example β , then only noise variance with at most a value with some order of magnitude greater than β can be possibly estimated. Therefore the step-size parameter must be adjusted according to the SNR of the observed sequence. For this purpose the adjustment of Eq. (12) is used. For SNRs less than 0 dB, the step-size parameter s is kept constant. For SNRs more than 0 dB, however, s is decreased at each iteration.

4. Simulation Examples

Computer simulations were carried out to evaluate the

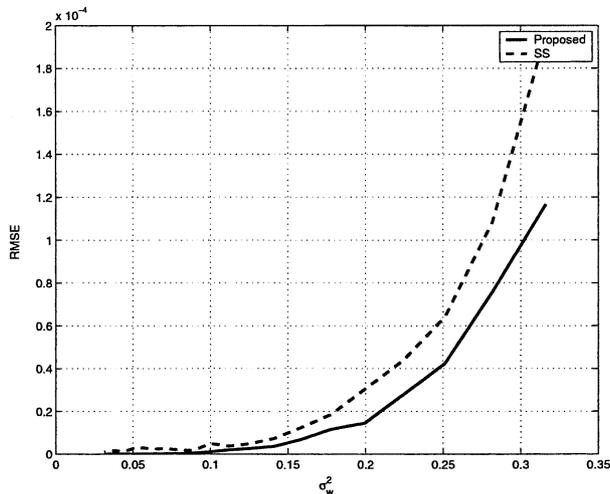


Fig. 1 RMSE of noise variance estimation for the SS and INVE methods at SNRs varying from 5 dB to 15 dB.

performance of the INVE method. The aim of the simulations was to compare the proposed method with other established methods for noise variance estimation. The INVE method was compared to the SS method for the reasons already stated in Sect. 1. In order to make this performance assessment, an AR process with the following transfer function was generated.

$$H(z) = (1 - 1.7466z^{-1} + 1.9112z^{-2} - 1.7030z^{-3} + 1.3450z^{-4} - 0.7839z^{-5} + 0.2266z^{-6})^{-1}. \quad (14)$$

The transfer function $H(z)$ was obtained from conjugate poles with amplitudes $\{0.70, 0.80, 0.85\}$ and respective frequencies of $\{0.10\pi, 0.30\pi, 0.60\pi\}$ radians per sample. The amplitude of the sequence was scaled to have zero mean and unit variance. From the scaled sequence, the last 40 data points were taken to represent the noise-free data sequence. A noisy AR data sequence was obtained by adding a white Gaussian noise sequence of the same length. The SNR was varied from 5 dB to 15 dB, where $SNR = 10 \log(1/\sigma_w^2)$. The σ_w^2 is the variance of the additive white Gaussian noise. The AR order p for parameter estimation was set to 17. The order was obtained by using the approximation $p = (N/3 + N/2)/2$ [10], where N is the length of the data sequence. The constant M for the INVE method was set to 1000. The noise variance was estimated at various SNRs in the above mentioned range. In all the simulations, 100 independent runs were averaged. The 100 independent runs were considered to be sufficient to give consistently accurate estimates of the noise variance.

Figure 1 illustrates the root mean square error (RMSE) of noise variance estimation where the INVE method is compared with the SS method. The RMSE at each noise variance σ_w^2 has been calculated as

Table 1 The standard deviation of noise variance estimates in the simulations.

SNR dB	Standard Deviation	
	INVE Method	SS Method
5	0.0738	0.1086
6	0.0553	0.1057
7	0.0538	0.0790
8	0.0076	0.0713
9	0.0369	0.0440
10	0.0038	0.0439
11	0.0077	0.0339
12	0.0068	0.0286
13	0.0006	0.0117
14	0.0087	0.0267
15	0.0040	0.0160

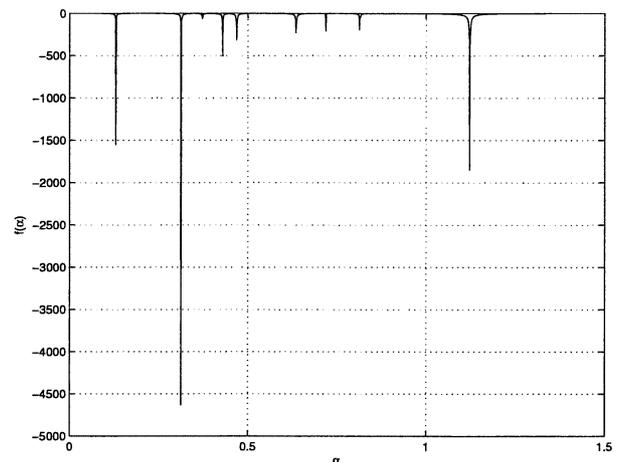


Fig. 2 $f(\alpha)$ at SNR=5 dB ($\sigma_w^2 = 0.3162$).

$$RMSE_{\sigma_w^2} = \sqrt{\left(\frac{1}{100} \sum_{i=1}^{100} (\sigma_w^2 - \hat{\sigma}_{wi}^2)^2 \right)} \quad (15)$$

where $\hat{\sigma}_{wi}^2$ is the i -th independent estimate of σ_w^2 . Figure 1 shows that while the RMSE performance is competitive at comparatively high SNRs (with small σ_w^2), the proposed INVE method provides lower RMSE than the SS method as the SNR is decreased (as σ_w^2 is increased). The standard deviation of noise variances estimates for the same range of SNRs was also compared for the two methods. As shown in Table 1, the proposed method shows lower standard deviation over the SNR range studied in the simulations.

Figure 2 illustrates the application of the INVE method at SNR=5 dB. By seeking the minimum point of $f(\alpha)$ the INVE method can accurately give a noise variance estimate at a low SNR. The function $f(\alpha)$ has negative values and if the minimum point is found, then the location of that point, α , corresponds to the estimate of noise variance. Figure 2 shows that α becomes very close to the true one 0.3162.

5. Concluding Remarks

In this paper we have proposed a noise variance estimation method. The noise is assumed to be white. The proposed INVE method is capable of accurately estimating the noise variance. Simulation results have shown that the INVE method performs better than the SS method, especially at low SNRs. Future work would aim to analyze the statistical properties of the proposed method and also to investigate its performance in the case where the noise is colored.

Acknowledgement

The authors are grateful to the reviewer of the journal for kind suggestions and advice.

References

- [1] S.M. Kay, Modern Spectral Estimation: Theory and Application, Prentice Hall, Englewood Cliffs, NJ, 1988.
- [2] S.V. Vaseghi, Advanced Signal Processing and Digital Noise Reduction, John Wiley and Teubner, England, 1996.
- [3] G.E.P. Box, G.M. Jenkins, and G.C. Reinsel, Time Series Analysis: Forecasting and Control, Prentice Hall, Englewood Cliffs, NJ, 1994.
- [4] C.E. Davila, "On the noise-compensated Yule-Walker equations," IEEE Trans. Signal Process., vol.49, no.6, pp.1119–1121, June 2001.
- [5] S.M. Kay, "Noise compensation for autoregressive spectral estimates," IEEE Trans. Acoust., Speech Signal Process., vol.ASSP-28, no.3, pp.292–303, June 1980.
- [6] W.X. Zheng, "Estimation of autoregressive signals from noisy measurements," IEE Proc., Vis. Image Signal Process., vol.144, no.1, pp.39–45, 1997.
- [7] H.-T. Hu, "Linear prediction analysis of speech signals in the presence of white Gaussian noise with unknown variance," IEE Proc., Vis. Image Signal Process., vol.145, no.4, pp.303–308, 1998.
- [8] C.E. Davila, "A subspace approach to estimation of autoregressive parameters from noisy measurements," IEEE Trans. Signal Process., vol.46, no.2, pp.531–534, 1998.
- [9] G.H. Golub and C.F. Van Loan, Matrix Computations, The Johns Hopkins University Press, Baltimore and London, 1989.
- [10] T.E. Landers and R.T. Lacoss, "Some geophysical applications of autoregressive spectral estimates," IEEE Trans. Geosci. Electron., vol.GE-15, no.1, pp.26–32, Jan. 1977.

Appendix: Analytical Derivation

Let there be α such that $0 \leq \alpha \leq r_{yy}(0)$. From Eq. (3) there exists an ACM $\tilde{\mathbf{R}}_{xx}$ such that

$$\tilde{\mathbf{R}}_{xx} = \mathbf{R}_{yy} - \alpha \mathbf{I}. \quad (\text{A} \cdot 1)$$

If the solution to the system of Eq. (1) is constrained such that the autocorrelation sequence is preserved, we can write

$$\tilde{\mathbf{R}}_{xx} \tilde{\mathbf{a}} = -\mathbf{r}_x. \quad (\text{A} \cdot 2)$$

By using the fact that $\mathbf{r}_x = \mathbf{r}_y$, it is possible to establish the relationship between the noisy solution and the estimated solution. Replacing \mathbf{r}_y with \mathbf{r}_x in Eq. (2) and inserting the left side of Eq. (A·2) into the resulting equation gives

$$\mathbf{R}_{yy} \hat{\mathbf{a}} = \tilde{\mathbf{R}}_{xx} \tilde{\mathbf{a}}. \quad (\text{A} \cdot 3)$$

Substituting for $\tilde{\mathbf{R}}_{xx}$ (Eq. (A·1)) on the right side of Eq. (A·3) results in

$$\begin{aligned} \mathbf{R}_{yy} \hat{\mathbf{a}} &= (\mathbf{R}_{yy} - \alpha \mathbf{I}) \tilde{\mathbf{a}}, \\ \mathbf{R}_{yy} (\tilde{\mathbf{a}} - \hat{\mathbf{a}}) &= \alpha \mathbf{I} \tilde{\mathbf{a}} = \alpha \tilde{\mathbf{a}}. \end{aligned} \quad (\text{A} \cdot 4)$$

In order to get the relationship between the relative magnitudes (in the norm sense) of \mathbf{a} and $\tilde{\mathbf{a}}$ we consider the following. By changing the sign of both sides of Eq. (4) and subtracting Eq. (A·4) from it we get

$$\mathbf{R}_{yy} (\mathbf{a} - \tilde{\mathbf{a}}) = \sigma_w^2 \mathbf{a} - \alpha \tilde{\mathbf{a}}. \quad (\text{A} \cdot 5)$$

Taking the norms of Eq. (A·5) results in

$$\|\mathbf{R}_{yy} (\mathbf{a} - \tilde{\mathbf{a}})\| = \|\sigma_w^2 \mathbf{a} - \alpha \tilde{\mathbf{a}}\|. \quad (\text{A} \cdot 6)$$

By considering the relationship

$$\|\mathbf{r}_x\| = \|\tilde{\mathbf{R}}_{xx} \tilde{\mathbf{a}}\| \quad (\text{A} \cdot 7)$$

and results from the eigendecomposition theory, we can infer the magnitude of $\tilde{\mathbf{a}}$ in Eq. (A·6). The following equations can be easily verified:

$$\begin{aligned} \mathbf{R}_{xx} \mathbf{a} &= \lambda_x \mathbf{a}, \\ (\mathbf{R}_{yy} - \sigma_w^2 \mathbf{I}) \mathbf{a} &= \lambda_x \mathbf{a}, \\ \mathbf{R}_{yy} \mathbf{a} &= (\lambda_x + \sigma_w^2) \mathbf{a} = \lambda_y \mathbf{a}. \end{aligned} \quad (\text{A} \cdot 8)$$

Equation (A·8) is the well known fact that the effect of additive noise is to increase the eigenvalues by an amount equal to the noise variance. Equation (8) can be found by considering the following cases.

A. *Case* $0 \leq \alpha \leq \sigma_w^2$

Since $\mathbf{R}_{xx} = \mathbf{R}_{yy} - \sigma_w^2 \mathbf{I}$, $\tilde{\mathbf{R}}_{xx} = \mathbf{R}_{yy} - \alpha \mathbf{I}$ and also considering the case where $\alpha < \sigma_w^2$, we can evaluate the magnitudes of the eigenvalues of \mathbf{R}_{xx} and $\tilde{\mathbf{R}}_{xx}$ if \mathbf{R}_{yy} is known. Letting the maximum eigenvalue of \mathbf{R}_{yy} be λ_{ymax} , by Eq. (A·8), the maximum eigenvalue of \mathbf{R}_{xx} is $\lambda_{xmax} = \lambda_{ymax} - \sigma_w^2$ and correspondingly the maximum eigenvalue of $\tilde{\mathbf{R}}_{xx}$ is $\tilde{\lambda}_{xmax} = \lambda_{ymax} - \alpha$. Using Eq. (1), $\|\mathbf{r}_x\| = \|\mathbf{R}_{xx} \mathbf{a}\| \leq \|\mathbf{R}_{xx}\| \|\mathbf{a}\| = \lambda_{xmax} \|\mathbf{a}\|$ or in short $\|\mathbf{r}_x\| \leq \lambda_{xmax} \|\mathbf{a}\|$. The Euclidean length of \mathbf{a} is then given by the inequality $\|\mathbf{a}\| \geq \|\mathbf{r}_x\| / \lambda_{xmax}$. Similarly Eq. (2) gives $\|\hat{\mathbf{a}}\| \geq \|\mathbf{r}_x\| / \lambda_{ymax}$ and Eq. (A·7) yields $\|\tilde{\mathbf{a}}\| \geq \|\mathbf{r}_x\| / \tilde{\lambda}_{xmax}$. For $\lambda_{xmax} < \tilde{\lambda}_{xmax}$, there exists a range of values of α in the interval $[0, r_{yy}(0)]$ for which

$$\lambda_{xmax} \leq \tilde{\lambda}_{xmax} \leq \lambda_{ymax},$$

$$\begin{aligned} 1/\lambda_{xmax} &\geq 1/\tilde{\lambda}_{xmax} \geq 1/\lambda_{ymax}, \\ \|\hat{\mathbf{a}}\| &\leq \|\tilde{\mathbf{a}}\| \leq \|\mathbf{a}\|. \end{aligned} \tag{A.9}$$

B. Case $\sigma_w^2 < \alpha \leq r_{yy}(0)$

In this case the equations are similarly defined as in the case A. The final result is that

$$\begin{aligned} \lambda_{ymax} &\geq \lambda_{xmax} > \tilde{\lambda}_{xmax}, \\ 1/\lambda_{ymax} &\leq 1/\lambda_{xmax} < 1/\tilde{\lambda}_{xmax}, \\ \|\hat{\mathbf{a}}\| &\leq \|\mathbf{a}\| < \|\tilde{\mathbf{a}}\|. \end{aligned} \tag{A.10}$$

Equations (A.9) and (A.10) look contradictory

but can be interpreted from two points of view. First, for $0 \leq \alpha \leq \sigma_w^2$, $\|\mathbf{a}\| - \|\tilde{\mathbf{a}}\| \geq 0$ but for $\sigma_w^2 < \alpha \leq r_{yy}(0)$, $\|\mathbf{a}\| - \|\tilde{\mathbf{a}}\| < 0$. On the other hand, $\|\hat{\mathbf{a}}\| - \|\tilde{\mathbf{a}}\| \leq 0$ for $0 \leq \alpha \leq \sigma_w^2$ and decreases and $\|\hat{\mathbf{a}}\| - \|\tilde{\mathbf{a}}\| \leq 0$ for $\sigma_w^2 < \alpha \leq r_{yy}(0)$ but increases towards 0 in the vicinity of σ_w^2 . Therefore the minimum value of $\|\hat{\mathbf{a}}\| - \|\tilde{\mathbf{a}}\|$ occurs at $\alpha = \sigma_w^2$. Furthermore, $\|\hat{\mathbf{a}}\| - \|\tilde{\mathbf{a}}\|$ can be easily evaluated on $[0, r_{yy}(0)]$.

From Eqs. (A.9) and (A.10) and restricting α in the interval $[0, r_{yy}(0)]$, we derive the following. If $\|\tilde{\mathbf{a}}\|$, denoted by $\|\tilde{\mathbf{a}}(\alpha)\|$, is continuous on the interval $[0, r_{yy}(0)]$, then $f(\alpha) = \|\hat{\mathbf{a}}\| - \|\tilde{\mathbf{a}}(\alpha)\| \leq 0$, with a minimum value only when $\alpha = \sigma_w^2$.