AN IMPROVED STEP-SIZE CONTROL FOR LMS FILTERS WITH CORRELATED INPUT SIGNALS

Dani Lippuner and August N. Kaelin

Signal and Information Processing Laboratory
Swiss Federal Institute of Technology, 8092 Zürich, Switzerland
E-Mail contact: lippuner@isi.ee.ethz.ch

ABSTRACT

In this paper, different step-size controls for LMS (Least-Mean-Squares) filters are proposed, which explicitly consider the input-signal correlation. The discussed algorithms are derived from the optimum Kalman filter and modified according to the constraint of a scalar step size, which itself must be a function of only scalar update parameters.

1. INTRODUCTION

Various step-size controls (SSC) for LMS filters have been published in the literature, e.g. [6], [3], which are, however, based on white input signals. In contrast to the optimum Kalman filter, a scalar SSC is not able to decorrelate a correlated input signal. It is nonetheless worthwhile to consider this correlation in the SSC design to at least avoid a correlation-dependent error propagation in the recursive updates. We propose three different versions of such SSCs, which allow to consider the input-signal correlation by an adjustment of one single parameter.

Under the constraint of a scalar step size the Kalman filter is reformulated in Section 2. The resulting optimum *scalarized* Kalman filter (SKF), however, is still dependent on the recursively updated system covariance matrix. By means of an auxiliary parameter, the input-signal correlation dependent *fading factor*, the matrix-updates are approximated by scalar updates in Section 3. A detailed analysis of the fading factor as well as a comparative simulation of the obtained SSCs are given in Section 4.

2. OPTIMUM STEP SIZE

The unknown system, which is assumed to be *time invariant*, is characterized by the state vector $\mathbf{x} = [x_1 \dots x_N]^T$ containing the N impulse-response coefficients. For each input-signal vector $\mathbf{u}_n = [u_n \dots u_{n-N+1}]^T$, a noisy output signal

$$d_n = \mathbf{u}_n^T \mathbf{x} + v_n \,, \quad v_n \sim \mathcal{N}(0, s_n) \,, \tag{1}$$

can be observed, where v_n is the random Gaussian *measurement noise* with variance s_n. We consider the common LMS algorithm, i.e.

$$e_n = d_n - \mathbf{u}_n^T \hat{\mathbf{x}}_n \tag{2}$$

$$\hat{\mathbf{x}}_{n+1} = \hat{\mathbf{x}}_n + \mu_n \mathbf{u}_n e_n \,, \tag{3}$$

which performs a real-time estimate $\hat{\mathbf{x}}_n$ of $\mathbf{x}$. It is intended to find a step size μ_n minimizing the average system distance p_n, i.e.

$$p_n \triangleq \frac{1}{N} \mathrm{tr}\,(\mathbf{P}_n) \,, \quad \mathbf{P}_n \triangleq \mathsf{E}\left\{\tilde{\mathbf{x}}_n \tilde{\mathbf{x}}_n^T\right\} \,, \tag{4}$$

at each instant of time. $\mathbf{P}_n$ is the covariance matrix of the coefficient-error vector $\tilde{\mathbf{x}}_n \triangleq \mathbf{x} - \hat{\mathbf{x}}_n$. In our discussion it is important to distinguish between the *ensemble* variance of e_n, which is defined by

$$r_n = \mathsf{E}\left\{e_n^2\right\} = \mathbf{u}_n^T \mathbf{P}_n \mathbf{u}_n + s_n = \ell_n + s_n \,, \tag{5}$$

and its *temporal* variance, which corresponds to

$$\bar{r}_n = \mathsf{E}_\mathsf{t}\left\{e_n^2\right\} = \mathrm{tr}\,(\mathbf{R}_{\mathbf{u}\,n} \mathbf{P}_n) + s_n = \bar{\ell}_n + s_n \,. \tag{6}$$

$\mathbf{R}_{\mathbf{u}\,n}$ is the input-signal autocorrelation matrix, and ℓ_n ($\bar{\ell}_n$) stands for the ensemble (temporal) systematic error variance. (The measurement noise v_n is assumed to be ergodic, hence $\mathsf{E}_\mathsf{t}\left\{v_n^2\right\} = \mathsf{E}\left\{v_n^2\right\} = s_n$.)

Based on the update equation of the coefficient-error vector, i.e.

$$\tilde{\mathbf{x}}_{n+1} = \left(\mathbf{I} - \mu_n \mathbf{u}_n \mathbf{u}_n^T\right) \tilde{\mathbf{x}}_n - \mu_n \mathbf{u}_n v_n \,, \tag{7}$$

the corresponding update of $\mathbf{P}_n$ evaluates to

$$\mathbf{P}_{n+1} = \mathbf{P}_n - \mu_n \left(\mathbf{u}_n \mathbf{u}_n^T \mathbf{P}_n + \mathbf{P}_n \mathbf{u}_n \mathbf{u}_n^T\right) + \mu_n^2 \mathbf{u}_n \mathbf{u}_n^T r_n \,. \tag{8}$$

After computing the trace of (8) and dividing the result by N, we obtain

$$p_{n+1} = p_n - 2\mu_n \frac{\ell_n}{N} + \mu_n^2 \frac{\|\mathbf{u}_n\|^2}{N} (\ell_n + s_n) \,. \tag{9}$$

Although (9) only consists of scalar quantities, the matrix $\mathbf{P}_n$ must be available to obtain ℓ_n. In order to circumvent this computationally expensive burden, we introduce a so-called *fading factor* η_n defined by

$$\eta_n = \frac{\ell_n}{p_n \|\mathbf{u}_n\|^2} \,, \tag{10}$$

such that ℓ_n can be substituted by a product of scalar values. Using (10) in (9), setting the derivative of (9) with respect to μ_n equal to zero, and solving for μ_n yields the optimum (superscript "opt") step size, i.e.

$$\mu_n^{\mathsf{opt}} = \frac{\eta_n p_n}{\eta_n \|\mathbf{u}_n\|^2 p_n + s_n} \tag{11}$$

with the corresponding p_n-update, i.e.

$$p_{n+1} = \left(1 - \eta_n \mu_n^{\mathsf{opt}} \frac{\|\mathbf{u}_n\|^2}{N}\right) p_n \,. \tag{12}$$

Equations (2)–(3), and (10)–(12) will be referred to as the *optimum scalarized Kalman filter* (opt SKF) for a time invariant system.

3. APPROXIMATIONS

In most cases there is no information available to accurately determine the fading factor η_n. However, simulations of the opt SKF (see Figure 1) show that η_n almost retains a constant value. Moreover, we proof in Appendix A that, for a stationary input signal, the fading factor is (in the mean) smaller than (or equal to) one. Equality holds only in the case of an uncorrelated input signal. Referring to Figure 1, we will approximate the fading factor by a constant value, i.e $\eta_n \cong \eta$, and emphasize that, depending on the input signal correlation, *this value may considerably deviate from one*. In addition, to avoid the appearance of η in the step-size definition (11), the quantity

$$p'_n = \eta \cdot p_n \tag{13}$$

instead of p_n will recursively be updated henceforth. Note that this has no effect on the p_n-update equation (12) except for the choice of the initial conditions, i.e. $p'_0 = \eta p_0$.

3.1. Ensemble-Variance Approximation (EVA SKF)

As already announced, we set the fading factor to a constant value, i.e.

$$\eta_n = \frac{\ell_n}{\|\mathbf{u}_n\|^2 p_n} \cong \eta \,. \tag{14}$$

Using (14) and (13), the step-size control (11) and (12) can directly be modified to

$$\mu_n = \frac{p'_n}{\|\mathbf{u}_n\|^2 p'_n + s_n} \,; \qquad p'_{n+1} = \left(1 - \eta \frac{\|\mathbf{u}_n\|^2}{N} \mu_n\right) p'_n \,. \tag{15}$$

This algorithm, which will be referred to as the *scalarized Kalman filter* with *ensemble-variance approximation* (EVA SKF), requires the knowledge of the measurement-noise variance s_n. If $\eta = 1$, the EVA SKF is very similar to the step-size control for uncorrelated input signals (referred to as UIS SSC) published in [3].

3.2. Partial Temporal-Variance Approximation

A temporal average of the innovation variance as well as the input-signal power can easily be computed as

$$\bar{r}_n = \bar{\ell}_n + s_n = \gamma_{\mathsf{r}} \bar{r}_{n-1} + (1 - \gamma_{\mathsf{r}})\, e_n^2 \tag{16}$$

$$\bar{\lambda}_n = \gamma_{\mathsf{r}} \bar{\lambda}_{n-1} + (1 - \gamma_{\mathsf{r}})\, u_n^2 \,. \tag{17}$$

Based on (10) and the assumption of a constant η, a temporal average of the systematic error variance evaluates to

$$\bar{\ell}_n = \eta N \bar{\lambda}_n p_n \,. \tag{18}$$

Hence, a near-at-hand approximation is carried out by replacing ℓ_n by $\bar{\ell}_n$, such that (9) turns into

$$p_{n+1} = (1 - 2\eta \bar{\lambda}_n \mu_n) p_n + \mu_n^2 \frac{\|\mathbf{u}_n\|^2}{N} \bar{r}_n \,. \tag{19}$$

After computing the optimum step size for the recursion (19) and applying the substitution (13), we obtain the following equations for the step size and the corresponding update of p_n, i.e.

$$\mu_n = \frac{N \bar{\lambda}_n p'_n}{\bar{r}_n} \frac{1}{\|\mathbf{u}_n\|^2} \,; \qquad p'_{n+1} = \left(1 - \eta \bar{\lambda}_n \mu_n\right) p'_n \,. \tag{20}$$

In this algorithm, only $\bar{\ell}_n$ (and consequently also $\bar{r}_n$) is computed temporally; the norm of the input-data vector $\|\mathbf{u}_n\|^2$ itself still represents an instantaneous rather than a temporally averaged quantity. This fact explains the expression "partial" in the title of this subsection.

It is interesting to note that the step size in (20) coincides with that of the *Normalized Least-Mean-Square* algorithm (NLMS) derived in [5]. However, different procedures, such as the "delay-coefficient"-method [7], are proposed to estimate the average system distance. Here, the SSC (20) is derived from the scalarized Kalman filter, which is itself based on a state-space model of the unknown system. We will thus refer to this SSC as the *model-based* NLMS algorithm (MB NLMS).

3.3. Temporal-Variance Approximation (TVA SKF)

Starting from the MB NLMS algorithm, the third and most simple version of the scalarized Kalman filter can be found by invoking the following additional approximation:

$$\|\mathbf{u}_n\|^2 \longrightarrow N \bar{\lambda}_n \,. \tag{21}$$

In comparison with the MB NLMS algorithm, (21) has no influence on the update equation of p'_n and the specification of the fading factor. However, thanks to the cancellation of $\|\mathbf{u}_n\|^2$ and $N\bar{\lambda}_n$, the step size of (20) turns into

$$\mu_n = \frac{p'_n}{\bar{r}_n} \,. \tag{22}$$

This SSC is referred to as the *SKF with temporal-variance approximation* (TVA SKF).

In Table 1, the update equations of the opt SKF of Section 2 and the three approximations of the opt SKF discussed in this section are summarized.

4. DISCUSSION

The fading factor will now be closer investigated. In the simulations of Figure 1 the fading factors of a 128-tap optimum SKF are monitored for three different first-order AR (autoregressive) input signals and two different measurement-noise variances. According to these numerical results, we recognize that the more u_n is correlated the smaller η is, because the input-signal correlation decelerates the convergence of p'_n (and consequently also μ_n).

However, practical rules for the adjustment of η are still missing. Fortunately, the *optimum Kalman filter* (opt KF) gives us an elegant thumb rule to approximately determine η. The simulations of Figure 1 show that the monitored η_n of the opt SKF come surprisingly close to the corresponding (straight-lined) virtual (steady-state) fading factors of the opt KF (superscript "KF"), i.e.

$$\eta \cong \eta_n^{\mathsf{KF}}\Big|_{n\to\infty} = \frac{\mathrm{tr}\left(\mathbf{R_u}\, \mathbf{P}_\infty^{\mathsf{KF}}\right)}{\bar{\lambda}\, \mathrm{tr}\left(\mathbf{P}_\infty^{\mathsf{KF}}\right)} \,. \tag{23}$$

We show in Appendix B that, in case of a *stationary* input signal, (23) approaches the value

$$\eta \cong \left(\frac{\bar{\lambda}}{N} \sum_{i=1}^{N} \frac{1}{\lambda_i}\right)^{-1} \,, \tag{24}$$

where λ_i is the ith eigenvalue of $\mathbf{R_u}$, and $\bar{\lambda}$ is the input-signal power. For a first-order AR process (AR-pole at c,

Filter-Update Equation		$\hat{\mathbf{x}}_{n+1} = \hat{\mathbf{x}}_n + \mu_n(d_n - \mathbf{u}_n^T\hat{\mathbf{x}}_n)$
Method	*Approximations*	*Equations* $(p'_n = \eta p_n)$
opt SKF		$\mu_n = \frac{\ell_n}{\ell_n + s_n}\frac{1}{\|\mathbf{u}_n\|^2}$ $p_{n+1} = p_n - \frac{1}{N}\mu_n\ell_n$
EVA SKF	$\eta_n = \frac{\ell_n}{\|\mathbf{u}_n\|^2 p_n} \cong \eta$	$\mu_n = \frac{p'_n}{p'_n\|\mathbf{u}_n\|^2 + s_n}$ $p'_{n+1} = \left(1 - \frac{\eta\|\mathbf{u}_n\|^2\mu_n}{N}\right)p'_n$
MB NLMS	$\ell_n \to \bar{\ell}_n$ $\eta_n = \frac{\ell_n}{N\bar{\lambda}_n p_n} \cong \eta$	$\mu_n = \frac{N\bar{\lambda}_n p'_n}{\bar{r}_n}\frac{1}{\|\mathbf{u}_n\|^2}$ $p'_{n+1} = \left(1 - \eta\bar{\lambda}_n\mu_n\right)p'_n$
TVA SKF	(see MB NLMS) $\|\mathbf{u}_n\|^2 \cong N\bar{\lambda}_n$	$\mu_n = \frac{p'_n}{\bar{r}_n}$ $p'_{n+1} = \left(1 - \eta\bar{\lambda}_n\mu_n\right)p'_n$

Table 1: Summary of the proposed step-size controls

$|c| < 1$), (24) is given by [1, 2]

$$\eta = \frac{1-c^2}{1+c^2(N-2)/N} \overset{N \text{ large}}{=} \frac{1-c^2}{1+c^2}\,. \qquad (25)$$

Figure 2 shows the system distances $\|\tilde{\mathbf{x}}_n\|^2$ of the optimum 128-tap Kalman filter (opt KF), the optimum scalarized Kalman filter (opt SKF), its three approximations (EVA SKF, MB NLMS, TVA SKF), and the step-size control for uncorrelated input signals (UIS SSC) [3]. The measurement noise is white with a variance of -10 dB, and the input signal is a first-order AR process (AR-pole at 0.9) with unit power. Note that the performances of the three approximations almost reach that of the opt SKF, whereas the UIS SSC shows a rather poor convergence.

5. CONCLUSIONS

We introduced three different simple step-size controls explicitly considering correlated input signals using a so-called fading factor η. It is analytically shown that $\eta \leq 1$ in the mean. Moreover, it is verified by means of simulations that η can be approximated by the analytically computable η^{KF} of the opt KF, which is only dependent on the input-signal statistics. For a first-order AR input process the nearly optimal performance of the proposed step-size controls has been demonstrated.

A. RANGE DETERMINATION OF THE FADING FACTOR

In this appendix we prove that the fading factor is (in the mean) a value, which is smaller than (or equal to) one. Equality holds only in the case of an uncorrelated input signal.

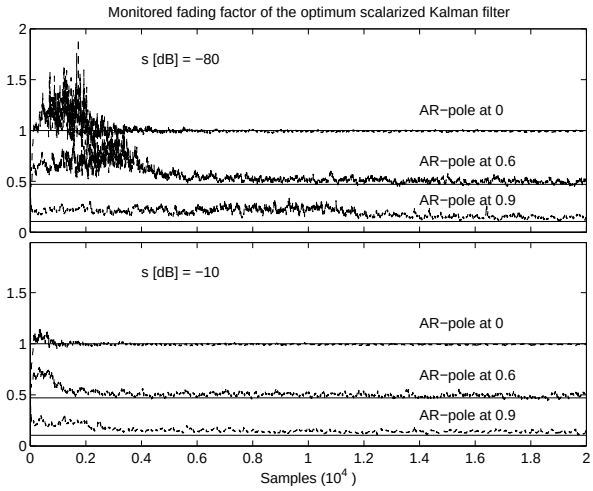

Figure 1: Monitored fading factor of the opt SKF.

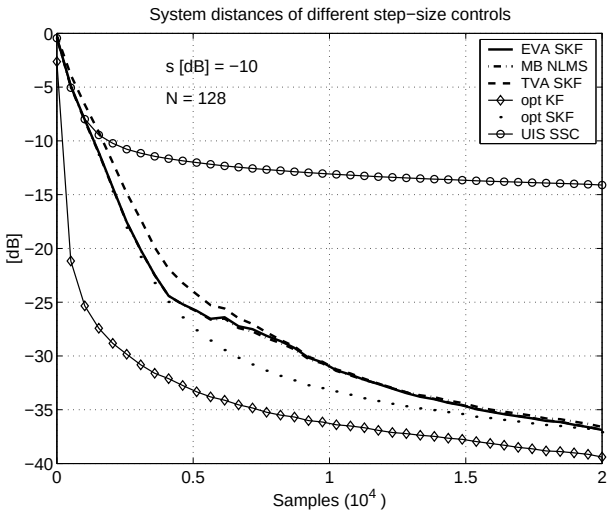

Figure 2: Comparison of system distances.

We rewrite the temporally averaged fading factor of the MB NLMS in Table 1 as

$$\eta_n = \frac{\bar{\ell}_n}{N\bar{\lambda}p_n} = \frac{\mathrm{tr}\,(\mathbf{R_u}\,\mathbf{P}_n)}{\bar{\lambda}\mathrm{tr}\,(\mathbf{P}_n)}\,, \qquad (26)$$

where we have assumed a stationary input-signal, hence $\mathbf{R_u}_{\,n} = \mathbf{R_u}$. The temporal (short-term) average of the $\mathbf{P}_n$-update equation (8) is given by

$$\mathbf{P}_{n+1} = \mathbf{P}_n - \mu_n\,(\mathbf{R_u}\,\mathbf{P}_n + \mathbf{P}_n\mathbf{R_u}\,) + \\ +\ \mu_n^2\mathbf{R_u}\ (\mathrm{tr}\,(\mathbf{R_u}\,\mathbf{P}_n) + s_n)\ . \qquad (27)$$

According to [4] the recursion (27) describes the behavior of (8) reasonably well for small step sizes. Since μ_n is approximately proportional to the number of filter coefficients, i.e. $\mu_n \propto 1/N$, this assumption is truly justified for high-order FIR filters.

Next, we normalize $\mathbf{R_u}$ by the input-signal power $\bar{\lambda}$ and use the following unitary similarity transformation:

$$\mathbf{\Lambda} = \mathbf{T}^T\frac{1}{\bar{\lambda}}\mathbf{R_u}\,\mathbf{T}\,. \qquad (28)$$

The diagonal matrix $\mathbf{\Lambda}$ contains the N eigenvalues λ_i' of $\mathbf{R_u}/\bar{\lambda}$, such that $\mathrm{tr}\left(\mathbf{\Lambda}\right) = \sum_{i=1}^{N} \lambda_i' = N$ holds. The transformation matrix $\mathbf{T}$ satisfies $\mathbf{T}^T\mathbf{T} = \mathbf{I}$. We also apply $\mathbf{T}$ to $\mathbf{P}_n$ to obtain

$$\mathbf{\Psi}_n = \mathbf{T}^T\mathbf{P}_n\mathbf{T}\,, \tag{29}$$

where, in contrast to $\mathbf{\Lambda}$, $\mathbf{\Psi}_n$ is not diagonal in general. Making use of (29) in the update equation (27) leads to the following recursion:

$$\mathbf{\Psi}_{n+1} = \left(\mathbf{I} - 2\bar{\lambda}\mu_n\mathbf{\Lambda}\right)\mathbf{\Psi}_n + \bar{\lambda}\mu_n^2\mathbf{\Lambda}\left(\mathrm{tr}\left(\mathbf{\Lambda}\mathbf{\Psi}_n\right) + s_n\right)\,. \tag{30}$$

Due to our assumption of a sufficiently small step size, we will neglect higher-order terms of μ_n, such that (30) can be approximated as:

$$\mathbf{\Psi}_n \cong \prod_{j=0}^{n}\left(\mathbf{I} - 2\bar{\lambda}\mu_j\mathbf{\Lambda}\right)\mathbf{\Psi}_0 \cong \left(\mathbf{I} - 2\bar{\lambda}\sum_{j=0}^{n}\mu_j\mathbf{\Lambda}\right)\mathbf{\Psi}_0\,. \tag{31}$$

Using (28), (29), and (31), the *numerator* of the fading factor (26) evaluates to:

$$\begin{aligned}\mathrm{tr}\left(\mathbf{R_u}\,\mathbf{P}_n\right) &= \bar{\lambda}\mathrm{tr}\left(\mathbf{T}\mathbf{\Lambda}\mathbf{T}^T\mathbf{T}\mathbf{\Psi}_n\mathbf{T}^T\right)\\ &= \bar{\lambda}\mathrm{tr}\left(\mathbf{\Lambda}\mathbf{\Psi}_n\right)\\ &\cong \bar{\lambda}\mathrm{tr}\left(\left(\mathbf{\Lambda} - 2\bar{\lambda}\sum_{j=0}^{n}\mu_j\mathbf{\Lambda}\mathbf{\Lambda}\right)\mathbf{\Psi}_0\right)\\ &\cong \bar{\lambda}\psi_0 N\left(1 - 2\bar{\lambda}\sum_{j=0}^{n}\mu_j\frac{1}{N}\mathrm{tr}\left(\mathbf{\Lambda}\mathbf{\Lambda}\right)\right)\,.\end{aligned} \tag{32}$$

In the last line of (32), the initial matrix $\mathbf{\Psi}_0$ was simplified to $\psi_0\mathbf{I}$, because we want to focus on the most common case, where $\mathbf{P}_0$ (and consequently also $\mathbf{\Psi}_0$) is initialized as a scaled identity matrix. Similarly, the *denominator* of the fading factor (26) results in:

$$\begin{aligned}\bar{\lambda}\mathrm{tr}\left(\mathbf{P}_n\right) &\cong \bar{\lambda}\mathrm{tr}\left(\left(\mathbf{I} - 2\bar{\lambda}\sum_{j=0}^{n}\mu_j\mathbf{\Lambda}\right)\mathbf{\Psi}_0\right)\\ &\cong \bar{\lambda}\psi_0 N\left(1 - 2\bar{\lambda}\sum_{j=0}^{n}\mu_j\right)\,.\end{aligned} \tag{33}$$

A comparison between (32) and (33) reveals that the presence of the term $\mathrm{tr}\left(\mathbf{\Lambda}\mathbf{\Lambda}\right)/N$ distinguishes the numerator from the denominator. We now have to show that this term is greater than (or equal to) one. Note that

$$\mathrm{tr}\left(\mathbf{\Lambda}\mathbf{\Lambda}\right) = \mathrm{tr}\left(\frac{1}{\bar{\lambda}}\mathbf{R_u}\,\frac{1}{\bar{\lambda}}\mathbf{R_u}\right) = \sum_{i=1}^{N}{\mathbf{r}_i}^T\mathbf{r}_i\,, \tag{34}$$

where $\mathbf{r}_i$ is the ith column vector of $\mathbf{R_u}/\bar{\lambda}$. Because of the normalization by $\bar{\lambda}$, the ith element of $\mathbf{r}_i$ is equal to one, such that the norm of $\mathbf{r}_i$ is greater than (or equal to) one, i.e. ${\mathbf{r}_i}^T\mathbf{r}_i = 1 + \epsilon_i^2$. This immediately allows to deduce that $\mathrm{tr}\left(\mathbf{\Lambda}\mathbf{\Lambda}\right)/N$ is greater than (or equal to) one, i.e.

$$\frac{\mathrm{tr}\left(\mathbf{\Lambda}\mathbf{\Lambda}\right)}{N} = \frac{1}{N}\sum_{i=1}^{N}{\mathbf{r}_i}^T\mathbf{r}_i = 1 + \frac{1}{N}\sum_{i=1}^{N}\epsilon_i^2 \geq 1\,. \tag{35}$$

As a consequence, the numerator (32) of the fading factor is (in the mean) smaller than (or equal to) its denominator (33), such that

$$\mathsf{E}_\mathsf{t}\left\{\eta_n\right\} \leq 1 \tag{36}$$

must hold.

B. VIRTUAL FADING FACTOR OF THE OPTIMUM KALMAN FILTER

For a time-invariant unknown system, the covariance matrix of the optimum (superscript "KF") Kalman filter propagates as

$$\mathbf{P}_{n+1}^{\mathsf{KF}} = \mathbf{P}_n^{\mathsf{KF}} - \frac{1}{r_n^{\mathsf{KF}}}\mathbf{P}_n^{\mathsf{KF}}\mathbf{u}_n\mathbf{u}_n^T\mathbf{P}_n^{\mathsf{KF}}\,, \tag{37}$$

where $r_n^{\mathsf{KF}} = \mathbf{u}_n^T\mathbf{P}_n^{\mathsf{KF}}\mathbf{u}_n + s_n$ is the ensemble innovation variance. Using the *matrix-inversion lemma*, the following relationship between the covariance matrix and the *empirical autocorrelation matrix* $\bar{\mathbf{R}}_{\mathbf{u}\,n} = (1/n)\sum_{i=1}^{n}\mathbf{u}_n\mathbf{u}_n^T$ of the input signal can be established:

$$\mathbf{P}_n^{\mathsf{KF}} = \frac{n+1}{s_n}\bar{\mathbf{R}}_{\mathbf{u}\,n}^{-1} = \tau_n\bar{\mathbf{R}}_{\mathbf{u}\,n}^{-1}\,. \tag{38}$$

The virtual fading factor of the opt KF thus evaluates to

$$\eta_n^{\mathsf{KF}} = \frac{\mathrm{tr}\left(\bar{\mathbf{R}}_{\mathbf{u}\,n}\mathbf{P}_n^{\mathsf{KF}}\right)}{\bar{\lambda}_n\mathrm{tr}\left(\mathbf{P}_n^{\mathsf{KF}}\right)} = \frac{\tau_n\mathrm{tr}\left(\mathbf{I}\right)}{\tau_n\bar{\lambda}_n\mathrm{tr}\left(\bar{\mathbf{R}}_{\mathbf{u}\,n}^{-1}\right)} = \frac{N/\bar{\lambda}_n}{\mathrm{tr}\left(\bar{\mathbf{R}}_{\mathbf{u}\,n}^{-1}\right)}\,. \tag{39}$$

Considering that $\lim_{n\to\infty}\bar{\mathbf{R}}_{\mathbf{u}\,n} = \mathbf{R_u}$ the steady-state fading factor is thus given by

$$\eta^{\mathsf{KF}} = \lim_{n\to\infty}\frac{N/\bar{\lambda}_n}{\mathrm{tr}\left(\bar{\mathbf{R}}_{\mathbf{u}\,n}^{-1}\right)} = \left(\frac{\bar{\lambda}}{N}\sum_{i=1}^{N}\frac{1}{\lambda_i}\right)^{-1}\,, \tag{40}$$

where λ_i is the ith eigenvalue of $\mathbf{R_u}$. The reciprocal of η^{KF}, i.e. $\beta = 1/\eta^{\mathsf{KF}}$, also appears in a slightly different context in [1], where β is referred to as the "correlation amplification factor."

C. REFERENCES

[1] T. Adali and S. H. Ardalan. On the effect of input signal correlation on weight misadjustment in the RLS algorithm. *IEEE Trans. on Signal Processing*, 43:988–991, April 1995.

[2] M. G. Belanger. *Adaptive Digital Filters and Signal Analysis*. New York and Basel: Marcel Dekker, Inc., 1987.

[3] S. C. Douglas. Normalized data nonlinearities for LMS adaptation. *IEEE Trans. on Signal Processing*, 42:1352–1365, June 1994.

[4] H. J. Kushner. *Approximation and Weak Convergence Methods for Random Processes with Applications to Stochastic System Theory*. Cambridge, Mass.: MIT Press, 1984.

[5] A. Mader, H. Puder, and G. U. Schmidt. Step-size control for acoustic echo cancellation filters - an overview. *Signal Processing*, 80:1697–1719, September 2000.

[6] D. T. M. Slock. On the convergence behavior of the LMS and the normalized LMS algorithms. *IEEE Trans. on Signal Processing*, 41:2811–2825, September 1993.

[7] S. Yamamoto and S. Kitayama. An adaptive echo canceller with variable step gain method. *Trans. of the IECE of Japan*, E 65:1–8, 1982.