SIMPLIFIED KALMAN ESTIMATION OF FADING MOBILE RADIO CHANNELS: HIGH PERFORMANCE AT LMS COMPUTATIONAL LOAD

Lars Lindbom
Systems and Control Group, Uppsala University
P.O. Box 27, S–751 03 Uppsala, Sweden

ABSTRACT

Low complexity algorithms for channel estimation in Rayleigh fading environments are presented. The channel estimators are presumed to operate in conjunction with a Viterbi detector, or an equalizer. The algorithms are based on simplified internal modelling of time-variant channel coefficients and approximation of a Kalman estimator. A novel averaging approach is used to replace the on-line update of the Riccati equation with a constant matrix. The associated Kalman gain is expressed in an analytical form. Compared to RLS tracking, both a significantly lower bit error rate and a much lower computational complexity is attained.

1. INTRODUCTION

When digital data are transmitted over mobile radio channels, time-dispersion and multipath fading may occur. To preserve high-quality transmission in such channels, it is necessary to utilize adaptive channel equalization.

An inherent difficulty associated with adaptive equalization is that unknown transmitted data are needed for channel or filter adaptation. They can be replaced by decisioned data in "decision directed mode". With severe fading there is a potential risk of losing tracking ability when (possibly incorrect) decisions are used in the adaptation. Characteristic of adaptive detection schemes based on decisioned data is that small improvements in tracking accuracy result in significantly lower bit error rates.

Tracking of parameters in rapidly time-varying systems with "classical" LMS and windowed RLS algorithms require large adaptation gains, i.e. a small memory. Algorithms with small memories become sensitive to noise and to incorrect decisions. They are therefore inappropriate for tracking rapidly time-varying parameters from noisy data. When prior knowledge about the behaviour of the time-varying parameters exists, algorithms with larger memories can be designed without sacrificing tracking capability.

In fading environments, the coefficients of a FIR channel model typically exhibit trend and quasi-periodic behaviour. One way to incorporate this a priori information in the adaptation is discussed here. It is related to design of recursive algorithms based on hypermodels (internal models), i.e. dynamical models of time-varying parameters, see e.g. [1,2,4].

Most of the current approaches, for tracking of time-varying parameters in fading environments, utilize difference approximations of the derivatives of the parameters in the estimation scheme. This is sometimes called fading-memory prediction, see e.g. [3,5]. These coefficient predictions filters are combined with LMS or RLS algorithms. A brief review of the use of coefficient prediction filters in the adaptation can be found in [6]. In this paper, we suggest new algorithms with simpler and less ad hoc setting of parameters in the adaptation. The internal models used here might be viewed as coefficient prediction filters embedded in a Kalman filter formulation.

The ideal estimator would be one with almost the same performance as the Kalman predictor, but with a computational complexity comparable to that of the LMS algorithm. It will be possible to design estimators with these properties, if transmitted sequences of symbols are *white* and the symbols have *constant modulus.* In this paper, derivation of such channel estimators is outlined. For more details, see [7].

2. STATEMENT OF THE PROBLEM

Consider a received sampled sequence $\{y(n)\}_1^N$. It is generated by transmission of one data burst $\{d(n)\}_1^N$ over an urban mobile radio channel, represented by an equivalent discrete-time baseband (FIR) channel model

$$\begin{aligned} y(n) &= \varphi^H(n)\theta(n) + v(n) \qquad (1)\\ \varphi^H(n) &= (d(n)\ \dots\ d(n-m))\\ \theta^T(n) &= (h_0(n)\ \dots\ h_m(n)). \end{aligned}$$

The noise, $v(n)$, is assumed to be white with zero mean, variance σ_v^2 and independent of the symbol sequence $\{d(n)\}$. The channel coefficients, $\{h_k(n)\}_{k=0}^m$, are Gaussian processes with zero means, variances $\sigma^2_{h_k}$ and subject to *independent* Rayleigh fading. All signals are complex-valued. Here, the objective is to reconstruct $\{d(n)\}_{N_{tr}+1}^{N}$ from $\{y(n)\}_1^N$ and a known training sequence $\{d(n)\}_1^{N_{tr}}$, by estimating the channel response.

3. INTERNAL MODELLING AND THE KALMAN PREDICTOR

It is well known that the autocovariance function of a channel coefficient subject to Rayleigh fading is given by $\sigma^2_{h_k} J_o(2\pi f_m/f_s\, \ell)$, where $J_o(\cdot)$ is the Bessel function of the first kind and of order zero, f_m is the maximum Doppler frequency, f_s is the sampling rate ($\gg f_m$) and ℓ is a time lag. Channel coefficients, which approximate the statistical properties of the Rayleigh fading model, can be generated by feeding white complex-valued Gaussian noise through a linear filter, here expressed in a state space form

$$\begin{aligned} x_k(n+1) &= F x_k(n) + G e_k(n) \\ h_k(n) &= H x_k(n) \qquad k = 0, 1, \ldots, m. \end{aligned} \tag{2}$$

The sequences $\{e_k(n)\}_{k=0}^{m}$ are mutually independent white complex-valued Gaussian noises (independent fading), with zero means and variances $\sigma^2_{e_k}$. $J_o(\cdot)$ is symmetrical around $\ell = 0$, so the matrices F, G and H are real-valued. The cut-off frequency of the filter is controlled by the maximum Doppler frequency (f_m).

By introducing the total state vector

$$X(n) = (x_0^T(n), \ldots, x_m^T(n))^T,$$

a state space description of the time-variant model (1) can be expressed as

$$\begin{aligned} X(n+1) &= \mathcal{F} X(n) + \mathcal{G} e(n) \qquad \theta(n) = \mathcal{H} X(n) \\ y(n) &= \Phi^H(n) X(n) + v(n), \end{aligned}$$

where $e(n) = (e_0(n), \ldots, e_m(n))^T$ and

$$\begin{aligned} \mathcal{F} &\triangleq \operatorname{diag}(F, \ldots, F) \qquad \mathcal{G} \triangleq \operatorname{diag}(G, \ldots, G) \\ \mathcal{H} &\triangleq \operatorname{diag}(H, \ldots, H) \\ \Phi^H(n) &\triangleq (d(n)H, \ldots, d(n-m)H). \end{aligned}$$

Above, diag$(\cdot)$ indicates forming a diagonal matrix with the arguments on the diagonal. Given the a priori information (2), the optimal adaptive algorithm for Gaussian noise would be the Kalman predictor

$$\begin{aligned} \varepsilon(n) &= y(n) - \varphi^H(n)\hat{\theta}(n) \\ K(n) &= P(n)\Phi(n)[\sigma_v^2 + \Phi^H(n)P(n)\Phi(n)]^{-1} \\ \widehat{X}(n+1) &= \mathcal{F}\widehat{X}(n) + K(n)\varepsilon(n) \qquad \hat{\theta}(n) = \mathcal{H}\widehat{X}(n) \end{aligned}$$

$$\begin{aligned} P(n+1) &= \mathcal{F}\{P(n) - P(n)Q(n)P(n)\}\mathcal{F}^T \\ &\quad + \mathcal{G}R_e\mathcal{G}^T \\ Q(n) &\triangleq \Phi(n)\Phi^H(n)[\sigma_v^2 + \Phi^H(n)P(n)\Phi(n)]^{-1}, \end{aligned} \tag{3}$$

where $R_e = \operatorname{diag}(\sigma^2_{e_0}, \ldots, \sigma^2_{e_m})$. From a practical point of view, a Kalman predictor based on a very accurate model description of the channel coefficients dynamics is not attractive. This model depends on f_m, which has to be accurately estimated since the power spectrum of a channel coefficient, subject to Rayleigh fading, has a pronounced peak at f_m. The quality of the estimate would then be sensitive to errors in estimates of f_m. Furthermore, the computational load would be high, since the matrices in the state space model (2) must be of rather high order. Fortunately, the high order model based estimator can be replaced by simplified estimators with much lower complexity and with less sensitivity to exact location of the peak at f_m, without losing too much in performance. In the sequel, we will regard the matrices $\{F, G, H\}$ as design variables, rather than very accurate models of the channel coefficient dynamics.

4. STEPS TOWARDS A LOW COMPLEXITY ESTIMATOR

The dominating computational load of a Kalman predictor is the recursive update of the Riccati equation (3). One conceivable way to avoid this update would be to replace $P(n)$ by the solution to an algebraic Riccati equation. However, since $P(n)$ depends on $\Phi^H(n)$, a stationary solution does *not* exist. Another way would be to compute $\{P(n)\}_1^N$ in advance, and store this sequence for subsequent use. However, this is not possible, since the sequence $\{\Phi^H(n)\}_1^N$ will be different from data batch to data batch.

Let us instead regard $\{P(n)\}$ as a matrix-valued stochastic process, and $\{P(n)\}_1^N$ in each data burst as independent realizations. Assume the ensemble mean $E[P(n)]$ to exist. Taking expectation of (3) gives

$$\begin{aligned} E[P(n+1)] &= \mathcal{F}\{E[P(n)] - E[P(n)Q(n) \\ &\qquad P(n)]\}\mathcal{F}^T + \mathcal{G}R_e\mathcal{G}^T. \end{aligned}$$

After the initial transient phase ($n > n_0$), the time-variations of $P(n)$ are caused by $\Phi^H(n)$. If the deviation of the mean-value sequence $\{E[P(n)]\}_{n>n_0}$ from $E[P(n_0)]$ is not too large, a constant approximation $\bar{P}$ of the sequence $\{E[P(n)]\}_{n\geq n_0}$ can be computed by solving the algebraic equation

$$\begin{aligned} \bar{P} &= \mathcal{F}[\bar{P} - \bar{P}\bar{Q}\bar{P}]\mathcal{F}^T + \mathcal{G}R_e\mathcal{G}^T \\ \bar{Q} &= E[\bar{Q}(n)] \\ \bar{Q}(n) &\triangleq \Phi(n)\Phi^H(n)[\sigma_v^2 + \Phi^H(n)\bar{P}\Phi(n)]^{-1}. \end{aligned} \tag{4}$$

An approximation of the asymptotic value of $E[P(n)]$, which can be calculated at the beginning of each data burst and utilized in the filtering thereafter, is thus sought. Now, $\bar{Q}(n)$ is a random matrix with a finite number of possible values. The evaluation of the expectation $E[\bar{Q}(n)]$ is then relatively straightforward. If the symbol sequence is white, $d(n)$ has constant modulus and the channel coefficients are subject to independent fading, then there exist a positive definite symmetrical block diagonal solution $\bar{P} = \text{diag}(\bar{P}_{00}, \ldots, \bar{P}_{mm})$ to (4). Equation (4) can then be written as a system of *coupled* equations

$$\begin{aligned} \bar{P}_{kk} &= F\left\{\bar{P}_{kk} - \bar{P}_{kk}\frac{H^T H}{\varrho_k + H\bar{P}_{kk}H^T}\bar{P}_{kk}\right\}F^T \\ &+ GG^T\sigma_{e_k}^2 \qquad k = 0, \ldots, m \qquad (5) \\ \varrho_k &\triangleq \frac{\sigma_v^2}{\sigma_d^2} + \sum_{i\neq k}^{m} H\bar{P}_{ii}H^T, \end{aligned}$$

where $E|d(n)|^2 = \sigma_d^2$. We can expect the coupling, here isolated in ϱ_k, to be quite loose when $\sigma_v^2/\sigma_d^2 \gg \sum_{i\neq k}^{m} H\bar{P}_{ii}H^T$. Let us regard ϱ_k as design variables. The system of equations then becomes *decoupled* and approximative solutions to $\bar{P}_{kk}$ can be calculated separately. The equations (5) are then ordinary algebraic Riccati equations, solvable under mild conditions. We can expect the approximations above to be reasonable when $||Q(n)||$ is nearly constant. This is the case when the symbols have constant modulus ($||\Phi(n)\Phi^H(n)||$ is then constant) and $P(n)$ is nearly block diagonal.

5. THE LOW COMPLEXITY ALGORITHM

The on-line computions required are given by

$$\begin{aligned} \varepsilon(n) &= y(n) - \varphi^H(n)\hat{\theta}(n) \\ \hat{x}_k(n+1) &= F\hat{x}_k(n) + L_k d^*(n-k)\varepsilon(n) \\ \hat{h}_k(n) &= H\hat{x}_k(n) \qquad k = 0, 1, \ldots, m. \end{aligned}$$

The gains L_k can be pre-computed from the solutions to

$$\begin{aligned} P_{kk} &= F\left\{P_{kk} - \frac{P_{kk}H^T H P_{kk}}{1 + HP_{kk}H^T}\right\}F^T + \gamma_k GG^T \\ L_k &= \frac{FP_{kk}H^T}{\sigma_d^2(1 + HP_{kk}H^T)} \quad P_{kk} \triangleq \frac{\bar{P}_{kk}}{\varrho_k} \quad \gamma_k \triangleq \frac{\sigma_{e_k}^2}{\varrho_k}. \end{aligned}$$

Here, γ_k can be used to tune the gain to a specific signal-to-noise ratio. Either, γ_k can be changed for each burst, or be kept constant over several bursts. Tuning of γ_k can also be used to account for possible nonstationary behaviour of R_e. The recursions are asymptotically stable for all choices of $\gamma_k \geq 0$. Note that this modelling approach provides extrapolation in time, $\hat{x}_k(n+\tau) = (F)^\tau \hat{x}_k(n)$ $\tau = 0, 1, 2, \ldots$.

With second order models $\{F, G, H\}$ simple *analytical expressions* for the gains L_k exist. Two second order models will be discussed below.

Lightly damped AR(2) models

In Rayleigh fading environments, the channel coefficients behave as narrow band noise with a spectral peak (cf. $J_o(\cdot)$). The simplest models, which describe such oscillatory behaviour, are lightly damped second order AR models with real-valued coefficients

$$\begin{aligned} h_k(n) &= \frac{1}{1 + a_1 q^{-1} + a_2 q^{-2}} e_k(n) \qquad (6) \\ a_1 &= -2r_d \cos\omega_d \quad a_2 = r_d^2. \end{aligned}$$

The pole locations are $r_d e^{\pm j\omega_d}$. The pole radius (r_d) reflects the damping and ω_d is the dominating frequency of the coefficient variation. Note that $\omega_d \sim 2\pi f_m/f_s$. If f_m is unknown, the spectral peak should be well damped to obtain a robust model.

The model (8) can be represented in an observable canonical state space form by the matrices

$$F = \begin{pmatrix} -a_1 & 1 \\ -a_2 & 0 \end{pmatrix} \quad G = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad H = (1\,0).$$

Analytical expressions for the gains L_k are given by

$$L_k = \frac{FP_{kk}H^T}{\sigma_d^2(1 + p_{11})} = \frac{1}{\sigma_d^2(1 + p_{11})}\begin{pmatrix} -a_1 p_{11} + p_{12} \\ -a_2 p_{11} \end{pmatrix},$$

where p_{11} and p_{12} are elements of P_{kk}, given by

$$\begin{aligned} \alpha_0 &\triangleq 1 + a_1^2 + a_2^2 + \gamma_k \\ \zeta &\triangleq \frac{\alpha_0 - 2a_2 + \sqrt{(\alpha_0 + 2a_2)^2 - 4a_1^2(1 + a_2)^2}}{2} \\ p_{11} &= \frac{\zeta + \sqrt{\zeta^2 - 4a_2^2}}{2} - 1, \quad p_{12} = \frac{a_1 a_2}{1 + a_2 + p_{11}} p_{11}. \end{aligned}$$

The expressions for p_{11} and p_{12} are derived in [7]. This estimator will be recognized as "KLMS" below. The number of real-valued arithmetic operations per iteration in the KLMS algorithm is summarized in the table below. It is compared to LMS tracking of $h_k(n)$.

	Add.	Mult.
KLMS	8(m+1)	12(m+1)
LMS	4(m+1)	6(m+1)

An integrated random walk model

In fading environments, the channel coefficients typically exhibit trend behaviour, i.e. they continue in some direction for a while. A simple way to incorporate such behaviour is to model the coefficients as integrated random walks

$$h_k(n) = h_k(n-1) + \frac{1}{1 - q^{-1}} e_k(n). \qquad (7)$$

This model can be represented in a state space form by the matrices

$$F = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad G = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad H = (1\ 0),$$

corresponding to the state vectors

$$x_k^T(n) = (x_k^{(1)}(n)\ x_k^{(2)}(n)) = (h_k(n)\ \ h_k(n+1) - h_k(n)).$$

Analytical expressions for the gains L_k are given by

$$L_k = \frac{FP_{kk}H^T}{\sigma_d^2(1+p_{11})} = \frac{1}{\sigma_d^2(1+p_{11})}\begin{pmatrix} p_{11}+p_{12} \\ p_{12} \end{pmatrix}$$

where p_{11} and p_{12} are elements of P_{kk}, given by

$$\zeta \triangleq \frac{4+\gamma_k+\sqrt{\gamma_k(16+\gamma_k)}}{2}$$

$$p_{11} = \frac{\zeta+\sqrt{\zeta^2-4}}{2} - 1 \quad p_{12} = \sqrt{\gamma_k(1+p_{11})}.$$

Note that $x_k(n)$ is formally assumed to be a nonstationary process when defining the "nominal model" (9). However, the actual estimates $\hat{x}_k(n)$ will be stationary, since stationary data are assumed. Note also that the forward difference approximations of the derivatives are estimated via $\hat{x}_k^{(2)}(n)$.

6. A SIMULATION STUDY

Consider transmission of data bursts of 170 symbols over a two taps Rayleigh fading channel ($m = 1$). The first 14 symbols constitute the training sequence. The channel coefficients change independently. We restrict attention to the case where the variance of the coefficients are equal, i.e. $E|h_0(n)|^2 = E|h_1(n)|^2$. The maximum Doppler frequency is 83 Hz and the symbol rate is 25kHz ($= f_s$). The symbols take the values $\{\pm 1 \pm j\}$, equally likely, and correspond to the set of bits $\{00, 01, 11, 10\}$. The symbols are differentially encoded. (This specification is similar to the proposed North American digital mobile radio standard. Here, we disregard the $\pi/4$-shift in $\pi/4$-DQPSK.)

The channel estimator is used in conjunction with a Viterbi detector. As internal (design) model we choose (8), with $r_d = .998$ and $w_d = .015$. (The autocovariance function associated with this model approximates $J_o(.021\,\ell)$ reasonably well for $\ell < 100$.)

In Figure 1, the performance (bit error rate) is compared for the "KLMS" algorithm (dashed), the Kalman predictor (dashed-dotted) based on the second order model, the RLS algorithm with forgetting factor (dotted), and using known channel states in the Viterbi algorithm (solid). With RLS tracking, the best performance, shown in the figure, was achieved with forgetting factor 0.7. (LMS tracking gives similar performance as RLS.) The difference between the Kalman predictor and the simplified algorithm is negligible.

The results were based on simulations of 1000 bursts (312 000 unknown bits), each at 15, 20 and 25 dB SNR. Time-varying coefficients were generated by filtering white complex-valued noise. The filters used were a second order slightly damped filter, followed by a 7'th order Butterworth filter. It was ascertained that the level-crossing statistics corresponded closely to the Rayleigh fading model.

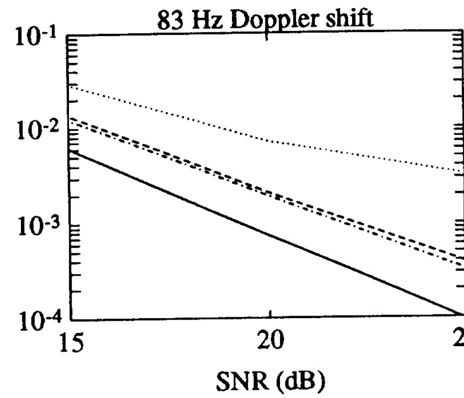

Figure 1: BER versus SNR of adaptive Viterbi based on the KLMS (dashed), the Kalman predictor (dashed-dotted), RLS tracking (dotted) and known channel (solid).

REFERENCES

[1] A Benveniste, M Métivier and P Priouret, *Adaptive Algorithms and Stochastic Approximations.* Springer-Verlag, Berlin Heidelberg, 1990.

[2] A Benveniste, "Design of adaptive algorithms for tracking of time-varying systems," *International Journal of Adaptive Control and Signal Processing,* vol. 2, pp. 3-29, 1987.

[3] A P Clark,*Adaptive Detectors for Digital Modems.* Pentech Press, London, 1989.

[4] G Davidov, A Shavit and Y Koren, "Estimation of dynamical-varying parameters by the internal model principle," *IEEE Transactions on Automatic Control,* vol. 37, pp. 498-503, 1992.

[5] S Hariharan and A P Clark, "HF channel estimation using a fast transversal filter algorithm," *IEEE Transactions on Acoustics, Speech, and Signal Processing,* vol. 38, pp. 1353-1362, 1990.

[6] G Kubin, "Adaptation in rapidly time-varying environments using coefficient filters," *Proceedings of ICASSP,* April 1991, Toronto, Ontario, Canada, vol. 3, pp. 2097-2100.

[7] L. Lindbom, *Adaptive equalization for fading mobile radio channels.* Licentiate Thesis, Department of Technology, Uppsala University, Sweden, Nov. 1992.