ON THE LINEAR LEAST-SQUARE PREDICTION PROBLEM

R. M. Fernández-Alcalá, J. Navarro-Moreno, J.C. Ruiz-Molina and M. D. Estudillo
Department of Statistic and Operations Research. University of Jaén
Campus Las Lagunillas, s/n. 23071, Jaén (Spain)

Keywords: Correlated signal and noise, linear least-square prediction problems.

Abstract: An efficient algorithm is derived for the recursive computation of the filtering and all types of linear least-square prediction estimates (fixed-point, fixed-interval, and fixed-lead predictors) of a nonstationary signal vector. It is assumed that the signal is observed in the presence of an additive white noise which can be correlated with the signal. The methodology employed only requires that the covariance functions involved are factorizable kernels and then it is applicable without the assumption that the signal verifies a state-space model.

1 INTRODUCTION

The estimation of a signal in the presence of additive white noise has been found to be among the central problems of statistical communication theory.

Application of the linear least mean-square error criterion leads to a linear integral equation, called Wiener-Hopf equation, whose solution is the impulse response of the optimal estimate. Although the linear least mean-square estimation problem is completely characterized by the solution to the Wiener-Hopf equation, a great effort has been made in the searching of efficient procedures for the computation of the desired estimator. Roughly speaking two approaches have been applied.

A first via of solution consists in using integral-equations approaches which provide the solution to the Wiener-Hopf integral equation for the impulse response function of the optimal estimator, from the knowledge of the covariance functions of the signal and noise [see, e.g. (Van Trees, 1968), (Kailath et al., 2000), (Fortmann and Anderson, 1973), (Gardner, 1974), (Gardner, 1975), (Navarro-Moreno et al., 2003)]. This technique is closely connected to series representation for stochastic processes and, in general, a series representation for the optimal estimate is provided instead of a recursive computational algorithm. The use of series representation for stochastic processes only allow to derive recursive procedures for the computation of suboptimum estimates.

On the other hand, a conventional approach to estimate a signal observed through a linear mechanism lies in imposing structural assumptions on the covariance functions involved. In this framework, the most representative algorithm is the Kalman-Bucy filter [see. e.g., (Kalman and Bucy, 1961), (Gelb, 1989)] which requires that the signal verifies a state-space model. However, although the Kalman-Bucy filter has been widely applied, there are a great number of physical phenomena that cannot be modelled by a state-space system. For problems with covariance information, linear least mean-square estimation algorithms have been designed under less restrictive structural conditions on the processes involved [(Sugisaka, 1983), (Fernández-Alcalá et al., 2005)]. Specifically, the only hypothesis imposed is that the covariance functions of the signal and noise are expressed in the factorized functional form.

Therefore, under the assumption that the covariance functions of the signal and noise are factorizable kernels, we aim to derived a recursive solution to the linear least-square estimation problem involving correlation between the signal and the observation noise. Specifically, using covariance information, an imbedding method is employed in order to design recursive algorithms for the filter and all kinds of predictors (fixed-point, fixed-interval, and fixed-lead predictors). Moreover, recursive formulas are designed for the error covariances associated with the above estimates.

M. Fernández-Alcalá R., Navarro-Moreno J., C. Ruiz-Molina J. and D. Estudillo M. (2005).
ON THE LINEAR LEAST-SQUARE PREDICTION PROBLEM.
In *Proceedings of the Second International Conference on Informatics in Control, Automation and Robotics - Signal Processing, Systems Modeling and Control*, pages 332-335
DOI: 10.5220/0001160303320335

Then, the paper is structured as follows. In the next section, a general formulation of the linear least-squares filtering and prediction problem is considered. Finally, in Section 3, the recursive algorithms for the filter and all types of predictors as well as their error covariances are derived.

2 PROBLEM STATEMENT

Let $\{x(t), 0 \leq t < \infty\}$ be a zero-mean signal vector of dimension n which is observed through the following equation:

$$y(t) = x(t) + v(t), \qquad 0 \leq t < \infty$$

where $y(t)$ represents the n-dimensional observation vector and $v(t)$ is a centered white observation noise with covariance function $E[v(t)v'(s)] = r\delta(t-s)$, with r a positive definite covariance matrix of dimension $n \times n$, and correlated with the signal.

We assume that the autocovariance function of the signal and the cross-covariance function between the signal and the observation noise are factorizable kernels which can be expressed in the following form:

$$\begin{aligned} R_x(t,s) &= \begin{cases} A(t)B'(s), & 0 \leq s \leq t \\ B(t)A'(s), & 0 \leq t \leq s \end{cases} \\ R_{xv}(t,s) &= \begin{cases} \alpha(t)\beta'(s), & 0 \leq s \leq t \\ \gamma(t)\lambda'(s), & 0 \leq t \leq s \end{cases} \end{aligned} \quad (1)$$

where $A(t)$, $B(t)$, $\alpha(t)$, $\beta(t)$, $\gamma(t)$, and $\lambda(t)$ are bounded matrices of dimensions $n \times k$, $n \times k$, $n \times l$, $n \times l$, $n \times l'$, and $n \times l'$, respectively.

We consider the problem of finding the linear least mean-square error estimator, $\hat{x}(t/T)$, with $t \geq T$, of the signal $x(t)$ based on the observations $\{y(s), s \in [0,T]\}$. It is known that such an estimate is the orthogonal projection of $x(t)$ onto $H(y,t)$ (the Hilbert space spanned by the process $\{y(s), s \in [0,T]\}$). Hence, $\hat{x}(t/T)$ can be expressed as a linear function of all the observed data of the form

$$\hat{x}(t/T) = \int_0^T h(t,s,T)y(s)ds, \;\; 0 \leq s \leq T \leq t \quad (2)$$

As a consequence of the orthogonal projection theorem, we obtain that the impulse response function $h(t,s,T)$ must satisfy the Wiener-Hopf equation

$$R_{xy}(t,s) = \int_0^T h(t,\sigma,T)R(\sigma,s)d\sigma + h(t,s,T)r \quad (3)$$

for $0 \leq s \leq T \leq t$, where $R_{xy}(t,s) = R_x(t,s) + R_{xv}(t,s)$, and $R(t,s) = R_x(t,s) + R_{xv}(t,s) + R_{vx}(t,s)$.

From (1), it is easy to check that $R_{xy}(t,s)$ and $R(t,s)$ can be written as follows:

$$\begin{aligned} R_{xy}(t,s) &= \begin{cases} F(t)\Gamma'(s), & 0 \leq s \leq t \\ G(t)\Lambda'(s), & 0 \leq t \leq s \end{cases} \\ R(t,s) &= \begin{cases} \Lambda(t)\Gamma'(s), & 0 \leq s \leq t \\ \Gamma(t)\Lambda'(s), & 0 \leq t \leq s \end{cases} \end{aligned} \quad (4)$$

where $F(t) = [A(t), \alpha(t), 0_{n\times l'}]$, $G(t) = [B(t), 0_{n\times l}, \gamma(t)]$, $\Lambda(t) = [A(t), \alpha(t), \lambda(t)]$, and $\Gamma(t) = [B(t), \beta(t), \gamma(t)]$ are matrices of dimensions $n \times m$ with $m = k + l + l'$, and $0_{p\times q}$ denotes the $(p \times q)$-dimensional matrix whose elements are zero.

Note that, we can expressed the optimal linear filter and all kinds of predictors through the equations (2) and (3). Specifically, by considering $T = t$ we have the filtering estimate $\hat{x}(t/t)$, the fixed-point predictor $\hat{x}(t_d/T)$ is derived by taking a fixed instant $t = t_d > T$, for the fixed-interval predictor, we consider a fixed observation interval $[0, T_d]$, with $T_d < t$, and finally the fixed-lead prediction estimate $\hat{x}(T+d/T)$, is given by (2) and (3) with $t = T + d$, for any $d > 0$.

Likewise, the error covariances associated with the above estimates can be defined as

$$P(t/T) = E[(x(t) - \hat{x}(t/T))(x(t) - \hat{x}(t/T))'] \quad (5)$$

with a suitable estimation instant, t, and a specific observation interval $[0,T]$.

Therefore, in the next section, the Wiener-Hopf equation (3) will be used, with the aid of invariant imbedding, in order to design recursive procedures for the filter and all kinds of predictors of the signal vector $x(t)$ as well as their associated error covariances. We must note that the only hypothesis assumed is that the covariance functions involved are factorizable kernels of the form (1).

3 RECURSIVE LINEAR ESTIMATION ALGORITHMS

Under the hypotheses established in Section 2, an efficient recursive algorithm for the linear least-square filter, and the fixed-point, fixed-interval and fixed-lead prediction estimates of the signal and their associated error covariance functions is presented in the following theorem.

Theorem 1 *The filter and the fixed-point, fixed-interval and fixed-lead prediction estimates of the signal $x(t)$ are recursively computed as follows:*

$$\begin{aligned} \hat{x}(t/t) &= F(t)L(t) \\ \hat{x}(t_d/T) &= F(t_d)L(T) \\ \hat{x}(t/T_d) &= F(t)L(T_d) \\ \hat{x}(T+d/T) &= F(T+d)L(T) \end{aligned} \quad (6)$$

where the m-dimensional vector $L(T)$ obeys the differential equation

$$\begin{aligned} \frac{\partial}{\partial T}L(T) =& J(T)\left[y(T) - \Lambda(T)L(T)\right] \\ L(0) =& 0_m \end{aligned} \quad (7)$$

with 0_m the m-dimensional vector with zero elements, and where $J(T)$ is given by the expression

$$J(T) = \left[\Gamma'(T) - Q(T)\Lambda'(T)\right] r^{-1} \quad (8)$$

with $Q(T)$ satisfying the differential equation

$$\begin{aligned} \frac{\partial}{\partial T}Q(T) =& J(T)\left[\Gamma(T) - \Lambda(T)Q(T)\right] \\ Q(0) =& 0_{m\times m} \end{aligned} \quad (9)$$

Moreover, the optimal linear estimation error covariance functions associated with the filtering estimate, $P(t/t)$, the fixed-point predictor, $P(t_d/T)$, the fixed-interval predictor, $P(t/T_d)$, and the fixed-lead predictor, $P(T+d/T)$, are formulated as follows:

$$\begin{aligned} P(t/t) =& R_x(t,t) - F(t)Q(t)F'(t) \\ P(t_d/T) =& R_x(t_d,t_d) - F(t_d)Q(T)F'(t_d) \\ P(t/T_d) =& R_x(t,t) - F(t)Q(T_d)F'(t) \\ P(T+d/T) =& R_x(T+d,T+d) \\ & - F(T+d)Q(T)F'(T+d) \end{aligned} \quad (10)$$

proof 1 *From* (4)*, the Wiener-Hopf equation* (3) *can be rewritten as*

$$h(t,s,T)r = F(t)\Gamma'(s) - \int_0^T h(t,\sigma,T)R(\sigma,s)d\sigma$$

Now, we introduce an auxiliary function $J(s,T)$ satisfying the equation

$$J(s,T)r = \Gamma'(s) - \int_0^T J(\sigma,T)R(\sigma,s)d\sigma \quad (11)$$

Then, it is obvious that the impulse response function is given by the expression

$$h(t,s,T) = F(t)J(s,T) \quad (12)$$

Next, differentiating (11) *with respect to T, we obtain that $J(s,T)$ obeys the following partial differential equation:*

$$\frac{\partial}{\partial T}J(s,T) = -J(T)\Lambda(T)J(s,T) \quad (13)$$

where $J(T) = J(T,T)$.

On the other hand, from (4) *and* (11)*, it is easy to check that*

$$J(T)r = \Gamma'(T) - \int_0^T J(\sigma,T)\Gamma(\sigma)d\sigma\Lambda'(T)$$

Then, the definition of a function $Q(T)$ as

$$Q(T) = \int_0^T J(\sigma,T)\Gamma(\sigma)d\sigma \quad (14)$$

leads to the equation (8)*.*

The equation (9) *is obtained by differentiating* (14) *with respect to T and using* (13) *in the resultant equation.*

Next, introducing a new auxiliary function

$$L(T) = \int_0^T J(\sigma,T)y(\sigma)d\sigma \quad (15)$$

and substituting (12) *in* (2)*, we have that*

$$\hat{x}(t/T) = F(t)L(T), \;\; \forall t \geq T \quad (16)$$

Then, by considering a suitable estimation instant, t, and a specific observation interval $[0,T]$ in (16)*, the filter and all kinds of predictors are given by the expressions* (6)*.*

Moreover, differentiating (15) *with respect to T and considering* (13) *in the resultant equation, it is easy to check that the above function $L(T)$ satisfies the differential equation* (7)*.*

Finally, in order to derived the expressions (10) *for the error covariances associated with the above estimates, we remark that, from the orthogonal projection lemma, the error covariance function* (5)*, can be rewritten as*

$$P(t/T) = R_x(t,t) - E[\hat{x}(t/T)\hat{x}'(t/T)]$$

Then, substituting (16) *in the above equation and using* (11)*, it is easy to check that*

$$P(t/T) = R_x(t,t) - F(t)Q(T)F'(t)$$

As consequence, the expressions given in (10) *can be obtained.*

ACKNOWLEDGMENT

This work was supported in part by Project MTM2004-04230 of the Plan Nacional de I+D+I, Ministerio de Educación y Ciencia, Spain. This project is financed jointly by the FEDER.

REFERENCES

Fernández-Alcalá, R. M., Navarro-Moreno, J., and Ruiz-Molina, J. C. (2005). Linear Least-Square Estimation Algorithms Involving Correlated Signal and Noise. *IEEE Trans. Signal Processing*. Accepted for publication.

Fortmann, T. E. and Anderson, B. D. O. (1973). On the Approximation of Optimal Realizable Linear Filters Using a Karhunen-Loève Expansion. *IEEE Trans. Inform. Theory*, IT-19:561–564.

Gardner, W. A. (1974). A Simple Solution to Smoothing, Filtering, and Prediction Problems Using Series Representations. *IEEE, Trans. Inform. Theory*, IT-20:271–274.

Gardner, W. A. (1975). A Series Solution to Smoothing, Filtering, and Prediction Problems Involving Correlated Signal and Noise. *IEEE, Trans. Inform. Theory*, IT-21:698–699.

Gelb, A. (1989). *Applied Optimal Estimation*. The Analytic Sciences Corporation.

Kailath, T., Sayed, A., and Hassibi, B. (2000). *Linear Estimation*. Prentice Hall.

Kalman, R. E. and Bucy, R. S. (1961). New Results in Linear Filtering and Prediction Theory. *Trans. ASME, J. Basic Engineering, Ser. D*, 83:95–108. In: Epheremides, A. and Thomas, J.B. (Ed.) 1973. Random Processes. Multiplicity Theory and Canonical Decompositions.

Navarro-Moreno, J., Ruiz-Molina, J., and Fernández, R. M. (2003). Approximate Series Representations of Second-Order Stochastic Processes. Applications to Signal Detection and Estimation. *IEEE, Trans. Inform. Theory*, 49(6):1574–1579.

Sugisaka, M. (1983). The Design of On-line Least-Squares Estimators Given Covariance Specifications Via an Imbedding Method. *Applied Mathematics and Computation*, (13):55–85.

Van Trees, H. L. (1968). *Detection, Estimation, and Modulation Theory-Part I*. Wiley, New York.