Appl. Math. Lett. Vol. 4, No. 2, pp. 43–44, 1991
0893-9659/91 $3.00 + 0.00

A Direct Derivation of the Interpolation Smoother

[1]M. R. Osborne and [2]Tania Prvan

[1]Statistics Research Section
School of Mathematical Sciences
Australian National University
[2]Department of Statistics
Glasgow University

(Received July 1990)

Abstract. It is shown that the fixed interval smoothing algorithm can be derived as a direct and simple consequence of the projection theorem. The derivation permits interpolation of the smoothed state vector values at other than observation points.

Introduction

The Kalman filter gives the best estimate $\mathbf{x}_{k|k}$ of the state vector of random variables $\mathbf{x}_k$ (or $\mathbf{x}(t_k)$) given observations on the system at $t = t_j,\ j = 1, 2, \cdots, k$ when the evolution of the state vector $\mathbf{x}_k$ is determined by the stochastic process

$$\mathbf{x}_{k+1} = T(t_{k+1}, t_k)\mathbf{x}_k + \mathbf{u}(t_{k+1}, t_k), \tag{1.1}$$

$$\mathbf{y}_{k+1} = H_{k+1}\mathbf{x}_{k+1} + \epsilon_{k+1}, \tag{1.2}$$

where $T(u, v)$ is invertible and prescribes the evolution of the deterministic part of the process from $t = u$ to $t = v$, H_k defines an observation on the process made at $t = t_k$, $\mathbf{y}_k$ is the corresponding vector of observed values, and $\mathbf{u}(t_{k+1}, t_k)$ and ϵ_{k+1} are vectors of random variables, uncorrelated with the past and with each other, and possessing zero means and known covariance matrices. The problem of determining $\mathbf{x}_{k|n}$, the best estimate of $\mathbf{x}_k$ given all currently available data, is known as the smoothing problem. It is less commonly addressed (books making no reference to this problem have been written on the Kalman filter), and even comprehensive treatments—for example, the definitive presentation given by Anderson and Moore [1]—develop the result by a quite lengthy argument which makes explicit use of the apparatus of the Kalman filter. That a direct derivation is possible is pointed out by Ansley and Kohn [2]. Their method makes clever use of conditional expectations which are orthogonal projections when the underlying distributions are multivariate normal. The importance of an effective derivation has been further emphasised in [3]. Here an even more direct derivation is given which uses the properties of orthogonal projections in the Hilbert space of random variables generated by the data. It is thus directly compatible with one of the best methods of deriving the filter equations (for example, the derivation given in Luenberger [4]). Interest in the smoothing algorithm has been stimulated in recent years by the exploitation of the Kalman filter in nonparametric and semi-parametric estimation problems (for example, Wecker and Ansley [5]).

2. The smoothing algorithm

Let $t_i \le t \le t_{i+1}$. The best estimate of $\mathbf{x}(t)$ given all the data (written $\mathbf{x}(t|n)$) will be derived under the assumption that the dynamics equation can be expressed as a composition of steps from t_i to t and from t to t_{i+1}. In this case

$$\mathbf{x}_{i+1} = T(t_{i+1}, t)\mathbf{x}(t) + \mathbf{u}(t_{i+1}, t) \tag{2.1}$$

where $\mathbf{u}(t_{i+1}, t)$ possesses the independence properties already assumed. This represents no new assumptions when $t = t_i$ and provides a means for interpolating the $\mathbf{x}_{k|n}$ otherwise.

Consider the identity

$$\begin{aligned}\mathbf{x}(t) &= \mathbf{x}(t|i) + \mathbf{x}(t) - \mathbf{x}(t|i)\\ &= \mathbf{x}(t|i) + T(t_{i+1}, t)^{-1}(\mathbf{x}_{i+1} - \mathbf{x}_{i+1|i}) - \mathbf{r}_{i+1}\end{aligned} \tag{2.2}$$

where $\mathbf{x}(t|i)$ is the best estimate of $\mathbf{x}(t)$ given data up to t_i, and

$$\mathbf{r}_{i+1} = T(t_{i+1}, t)^{-1}\mathbf{u}(t_{i+1}, t),$$

Now $\mathbf{r}_{i+1}$ is uncorrelated with $\mathbf{x}(t) - \mathbf{x}(t|i)$ by assumption, and hence must lie in the Hilbert space of random variables generated by $\mathbf{x}_{i+1} - \mathbf{x}_{i+1|i}$. It follows that $\mathbf{x}(t) - \mathbf{x}(t|i)$ is given by its projection onto this Hilbert space *identically*. The projection theorem (Luenberger [4], p. 92) gives

$$\mathbf{x}(t) - \mathbf{x}(t|i) = \mathcal{E}\{(\mathbf{x}(t) - \mathbf{x}(t|i))(\mathbf{x}_{i+1} - \mathbf{x}_{i+1|i})^T\}S_{i+1|i}^{-1}(\mathbf{x}_{i+1} - \mathbf{x}_{i+1|i}) \tag{2.3}$$

where

$$S_{i+1|i} = \mathcal{E}\{(\mathbf{x}_{i+1} - \mathbf{x}_{i+1|i})(\mathbf{x}_{i+1} - \mathbf{x}_{i+1|i})^T\}. \tag{2.4}$$

Now, by (2.1),

$$\mathcal{E}\{(\mathbf{x}(t) - \mathbf{x}(t|i))(\mathbf{x}_{i+1} - \mathbf{x}_{i+1|i})^T\} = \mathcal{E}\{(\mathbf{x}(t) - \mathbf{x}(t|i))(\mathbf{x}(t) - \mathbf{x}(t|i))^T\}T(t_{i+1}, t)^T \tag{2.5}$$

because $\mathbf{u}(t_{i+1}, t)$ is uncorrelated with $\mathbf{x}(t) - \mathbf{x}(t|i)$. Thus

$$\mathbf{x}(t) - \mathbf{x}(t|i) = S(t|i)T(t_{i+1}, t)^T S_{i+1|i}^{-1}(\mathbf{x}_{i+1} - \mathbf{x}_{i+1|i}). \tag{2.6}$$

The desired result is now obtained by projecting both sides of (2.6) onto the Hilbert space generated by $\mathbf{y}_1, \mathbf{y}_2, \cdots, \mathbf{y}_n$ and the initial conditions. This gives

$$\mathbf{x}(t|n) - \mathbf{x}(t|i) = S(t|i)T(t_{i+1}, t)^T S_{i+1|i}^{-1}(\mathbf{x}_{i+1|n} - \mathbf{x}_{i+1|i}). \tag{2.7}$$

That this projection commutes with matrix multiplication is *Theorem 4.6.1* in Luenberger [3].

Let

$$A(t_{i+1}, t) = S(t|i)T(t_{i+1}, t)^T S_{i+1|i}^{-1}.$$

Then it follows immediately on subtracting (2.7) from (2.6) that

$$S(t|n) = A(t_{i+1}, t)S_{i+1|n}A(t_{i+1}, t)^T. \tag{2.8}$$

This expression for $S(t|n)$ is an agreeable simplification of the form usually quoted.

3. Conclusion

The smoothing equation (2.7) is shown to be a direct and simple consequence of the projection theorem.

References

1. B.D.O. Anderson and J.B. Moore, *Optimal Filtering*, Prentice-Hall, Englewood Cliffs, N.J., (1979).
2. C.F. Ansley and Robert Kohn, A geometrical derivation of the fixed interval smoothing algorithm, *Biometrica* **69**, 486-487 (1982).
4. P. de Jong, Smoothing and interpolation with the state space model, *J. Am. St. Assoc.* **84**, 1085-1088 (1989).
4. D.G. Luenberger, *Optimization by Vector Space Methods*, Wiley, New York, (1969).
5. W. Wecker and C.F Ansley, The signal extraction approach to nonlinear regression and spline smoothing, *J. Am. Stat. Assoc.* **78**, 81-89 (1983).

[1]GPO Box 4, Canberra, ACT 2601, Australia
[2]Glasgow G12 8QW, Scotland