WA3 - 10:45

On Scattering, Time Reversal and Information Forms

A. M. Bruckstein and T. Kailath

Information Systems Laboratory
STANFORD UNIVERSITY
Stanford, CA94305

ABSTRACT

Redheffer scattering theory provides a natural setting for the solution of various state-space estimation and control problems. In this short note we discuss the algorithms for filtering "in reverse" a process for which a forward model is given and the information form filter that is used when no priors are given on the model state. In the language of scattering theory the updates for both the backwards algorithm and the information form filter are shown to correspond to cascading rotated generator layers associated with a Hamiltonian two-point boundary value problem (TPBVP).

1. Introduction

In fixed interval smoothing problems we have the possibility to process the given observations in any order we wish, this degree of freedom being exploited for computational efficiency and in theoretical derivations. A well-known and appealing solution for the smoothed estimate has the form of a suitably weighted combination of purely causal and anticausal estimates[1],[2]. These are obtained by running two filters, with forwards and a backwards recursions over the data interval. The problem of giving a proper interpretation to the backwards running recursions led to the introduction and study of backwards Markovian models and their corresponding Kalman filters. These concepts provided a natural interpretation to the Mayne-Frazer two filter solution [3]-[6]. In this context, scattering theory proved to be a powerful tool in the derivation of the results and here we review these issues rather briefly, referring to the original works of Kailath, Friedlander, Ljung and Verghese for details.

The usual forwards Kalman filter recursions incorporate the initial conditions information to yield the best Bayesian state estimates. If no a-priori information is supplied problems arise however with the start-up of the Riccati recursions. In order to obtain the Fisher estimates it is formally required to set the initial condition of the Riccati differential equation to infinity. This issue is practically dealt with either starting the recursions with a very "large" diagonal matrix, or passing the equations to the so-called information forms [7]-[9].

The derivation of information forms is not difficult, particularly in the continuous time case and by looking at the resulting recursions one immediately recognizes a striking resemblance to the backwards equations, providing the anticausal Fisher estimates in two filter smoothing algorithms [10]-[11]. Using the scattering framework for a theoretical derivation of the information forms clearly displays why the backwards equations are formally so similar to them.

2. Hamiltonians and Redheffer scattering theory

This work was supported in part by the U.S. Army Research Office, under Contract DAAG29-79-C-0215 and by the National Science Foundation, under Grant ECS78-10003.

Suppose we have a state-space model of a continuous or discrete time signal $z(\cdot)$

$$\frac{d}{dt}x(t) = A_t x(t) + B_t w(t) \qquad (1)$$

$$z(t) = C_t x(t)$$

and we are given its noisy observations

$$y(t) = z(t) + v(t) \qquad (2)$$

The driving and observation noises will be assumed, as usual, uncorrelated white processes with covariances given by Q_t and R_t, however all the results are easily extended to the correlated case.

Given the observations over an interval $\Delta = [\tau_i, \tau_f]$ it is a well known result that the smoothed state estimates $\hat{x}(t \mid \Delta)$ are provided by the solution of a linear Hamiltonian TPBVP as follows (see [12])

$$\frac{d}{dt}\begin{bmatrix} \hat{x}(t \mid \Delta) \\ \lambda(t \mid \Delta) \end{bmatrix} = \begin{bmatrix} A_t & B_t Q_t B_t^* \\ C_t^* R_t^{-1} C_t & -A_t^* \end{bmatrix} \begin{bmatrix} \hat{x}(t \mid \Delta) \\ \lambda(t \mid \Delta) \end{bmatrix} + \begin{bmatrix} 0 \\ -C_t^* R_t^{-1} y(t) \end{bmatrix}$$

with boundary conditions

$$\lambda(\tau_f \mid \Delta) = 0 \quad \text{and} \quad \hat{x}(\tau_i \mid \Delta) = x_i + P_i \lambda(\tau_i \mid \Delta) \qquad (3)$$

In the above boundary conditions x_i and P_i summarize the prior knowledge on the state at the moment τ_i, i.e. the mean and variance of the initial state estimate with no observations.

Assume that the above linear Hamiltonian system is solved by simple forward propagation of the extended state from some arbitrary initial condition. It is immediate that the corresponding extended final state will be given by the following formula

$$\langle ext\ state \rangle_{\tau_f} = M^T \langle ext\ state \rangle_{\tau_i} + \Sigma^T \qquad (4)$$

where M^T is a transition matrix and the vector Σ^T summarizes the effect of nonzero input (both being obviously independent of the assumed initial extended state) It is clear further that if we have the pair $[M^T, \Sigma^T]$ the values of the solution of the Hamiltonian TPBVP at the boundaries are easily obtained, by simply solving for the unknowns in (4). Thus in order to solve the Hamiltonian TPBVP's over arbitrarily varying intervals Δ_t we have to obtain recursively in time the pairs $[M^T(\Delta_t), \Sigma^T(\Delta_t)]$ or, as we shall see, some related quantities that are obtained by a Mason exchange rule from them.

Scattering theory deals with the recursive computation of such $[M, \Sigma]$ pairs, given the so called generator sequence, which corresponds to the state-space model and noise parameters [10],[13]. Associated to the solution of the linear Hamiltonian which is characterized by the pair $[M^T, \Sigma^T]$ we have a pair $[M^S, \Sigma^S]$ relating the final state and the initial adjoint vector to the initial state and the final (and always zero) adjoint vector. This pair is given by the "Mason exchange rule" **X** as follows

$$\mathbf{X}\left[\begin{array}{cc|c} \alpha & \beta & \sigma_u^T \\ b & \alpha & \sigma_d^T \end{array}\right] = \left[\begin{array}{cc|c} \alpha - \beta\alpha^{-1}b & \beta\alpha^{-1} & \sigma_u^T - \beta\alpha^{-1}\sigma_d^T \\ -\alpha^{-1}b & \alpha^{-1} & -\alpha^{-1}\sigma_d^T \end{array}\right] \qquad (5)$$

Scattering theory, we recall, also provides direct recursions, involving Riccati equations, for these, so called *scattering* representations. The scattering generator sequence corresponding to the Hamiltonian system (3) is

$$[g^S(t),\gamma^S(t)] = \left[\begin{array}{cc|c} A_t & B_t Q_t B_t^* & 0 \\ -C_t^* R_t^{-1} C_t & A_t^* & C_t^* R_t^{-1} y(t) \end{array}\right] \quad (6)$$

3. Time reversal and backwards generators

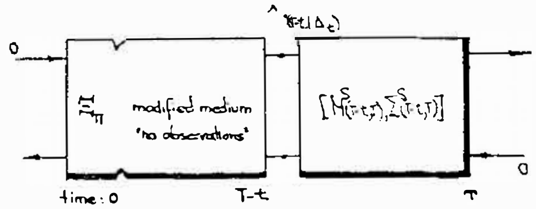
Suppose we are given a Markovian signal model, with initial condition information and observations over an interval $\Delta = [0,T]$ and we wish to process the observations sequentially, but *backwards* in time. Thus we are interested in recursions providing $\hat{x}(T-t\,|\,\Delta_t)$ for increasing t, were the estimates are based on both the given initial conditions and the observations over the interval $\Delta_t = [T-t, T]$.

The problem as posed is connected with the problem of missing observations, since $\hat{x}\,(t\,|\,\Delta_t)$ is nothing but the smoothed state estimate when the data points over $[0,t)$ are for some reason unavailable. If for some period of time no observations are given, we can account for this in the scattering picture by a simple modification of the medium generator. The modification amounts to setting either R_s^{-1} to zero (i.e. assuming infinite observation noise variance rendering the observations useless) or or taking $C_s = 0$ for that period of time (i.e. assuming the observations are of noise alone). The so-modified S-domain generators are

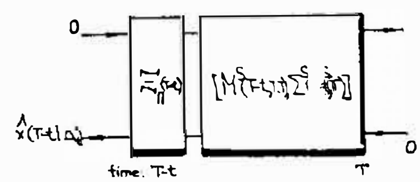
$$\left[\begin{array}{cc|c} A_s & B_s Q_s B_s^* & 0 \\ 0 & A_s^* & 0 \end{array}\right] \quad (7)$$

and these provide the Lyapunov recursions for the evolution of the second-order statistics of the state.

With the above discussion the problem we posed is completely solved, since all we have to do is look at the Hamiltonian medium over $[0,T]$ and assume *no observations* during the interval $[0,t)$. A graphical description of this is



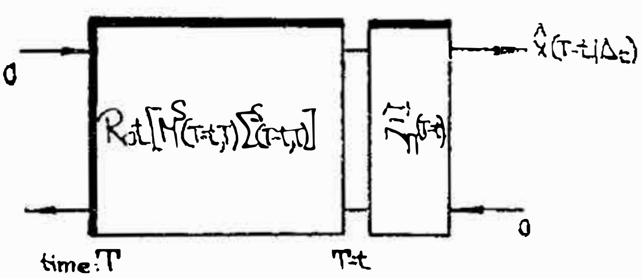
and from it we can simply read out all the required estimates. Therefore we know how to obtain the Bayesian estimate $\hat{x}\,(T-t\,|\,\Delta_t)$, however, if we wish to develop recursions for it (i.e. to obtain the backwards Kalman filter) the following equivalent description is perhaps more relevant



Here the "no observations" medium was deleted and its effect replaced by a suitable initial conditions layer $\Xi_\Pi(T-t)$, displaying the initial conditions propagated to $T-t$ as follows $\Xi_\Pi(s) = \left[\begin{array}{cc|c} P_{ts} & P_{ts} & x_{ts} \\ P_{ts} & P_{ts} & x_{ts} \end{array}\right]$.

In order to obtain recursions for the resulting scattering representation (where the estimates appear as sources) we only have to realize that it is necessary to determine the generators for a suitably modified backwards trajectory. Thus the initial conditions, the given statistics of the state at the origin, when propagated with no observations provide the parameterized family of layers that left-modify the $[.M\Sigma]$ trajectory of interest. (Indeed, it is obvious that the initial conditions together with the given system dynamics and driving-noise statistics provide a-priori information on the entire state-trajectory.)

Although the above discussion solves the problem of determining recursions for $\hat{x}\,(T-t\,|\,\Delta_t)$ we sill have to perform a "cosmetic" step to give an interpretation of the result as a backward Kalman filter. As is clear from the previous diagram the estimate of interest appears as a lower source of the scattering representation and the medium extends to the left, whereas in the usual setting the Kalman filter estimates appear as upper sources with the medium extending to the right. If we wish to "correct" for this we only have to rotate the whole picture 180° and by the nature of the scattering representation everything falls in the right place, as follows



Mathematically, a 180° rotation in the scattering domain is the following operation

$$Rot[M^S,\Sigma^S] = [JM^S J, J\Sigma^S] \quad \textit{where} \quad J = \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix} \quad (8)$$

and we also note the fact that

$$Rot\{\Xi_1 \gg \Xi_2 \gg \cdots \gg \Xi_n\} = Rot\,\Xi_n \gg \cdots \gg Rot\,\Xi_2 \gg Rot\,\Xi_1 \quad (9)$$

where $\gg$ denotes the layer cascading operation, and thus contains all the update equations.

The backwards filtering generators

The processing real-time t will be the index of the generator, and as t goes from zero to T the data will be processed in the reverse direction. We shall have (for infinitesimal δ)

$$[I+g_r^S(t)\delta\,,\ \gamma_r^S(t)\delta] = \quad (10)$$
$$\Xi_\Pi^{-1}(T-t) \gg [I+Jg^S(T-t)J\delta, J\gamma^S(T-t)\delta] \gg \Xi_\Pi(T-t+\delta)$$

and in the above we used the rotation symmetry of the initial condition layers, i.e. $Rot\,\Xi_\Pi(\cdot) = \Xi_\Pi(\cdot)$.

We thus see how a simple exercise in "time-travel" provides nice interpretations to the recursions appearing in the Mayne-Frazer two-filter smoothing result. Although in the continuous case the backwards generators automatically display the parameters of a reversed time Markovian model, in the discrete case some more algebra is required. The problem is that in discrete-time the estimate we obtain is $\hat{x}\,(T-t\,|\,\Delta_t)$, which is not a predicted but a filtered estimate. In order to make an identification with a forward case one thus has to either obtain the forwards Kalman recursions for the filtered estimates or to obtain backwards recursions for $\hat{x}\,(T-t-1\,|\,\Delta_t)$. We have to point out, however, that numerically the above calculations are quite straightforward, in fact all one needs is a good computer subroutine to perform the cascade operation.

4. On the derivation of information forms

Assume we used the T-domain recursions to obtain a transfer representation of the Hamiltonian medium over say $[0,t]$ Now using the Mason exchange rule we have that the error covariance and the state estimate are given by

$$P_t = \beta\alpha^{-1} \quad \text{and} \quad \hat{x}_t = \sigma_u^T - \beta\alpha^{-1}\sigma_d^T \tag{11}$$

If instead of $[M^T, \Sigma^T]$ we consider the same representation with the rows interchanged

$$J[M^T, \Sigma^T] = [JM^T, J\Sigma^T] = \left[\begin{array}{cc|c} b & \alpha & \sigma_d^T \\ a & \beta & \sigma_u^T \end{array}\right] \tag{12}$$

then it is easily seen that we have the following result

$$\mathbf{X}(J[M^T, \Sigma^T]) = \left[\begin{array}{cc|c} \times & P_t^{-1} & -P_t^{-1}U_t \\ \times & \times & \times \end{array}\right] \tag{13}$$

This shows that if we wish to obtain evolution equations for the information variables we should find generators that, in the transfer domain directly provide $[JM^T, -J\Sigma^T]$. The corresponding S-domain generators will then give the direct recursions for the information form.

The generators of information forms

If the scattering domain infinitesimal generator sequence $[g^S(\cdot), \gamma^S(\cdot)]$ is given then we have the forwards evolution equation in the T-domain

$$\frac{d}{dt}[M^T, \Sigma^T] = [g^T(t)M^T, g^T(t)\Sigma^T + \gamma^T(t)] \tag{14}$$

and recall that the relation between scattering and transfer generators is

$$[g^T(\cdot), \gamma^T(\cdot)] = \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix} [g^S(\cdot), \gamma^S(\cdot)] \tag{15}$$

Therefore we obtain (using the fact that J is orthogonal) the following evolution equations for the required modified representation

$$\frac{d}{dt}[JM^T, -J\Sigma^T] = [Jg^T(t)J\,(JM^T), Jg^T(t)J(-J\Sigma^T) - J\gamma^T(t)] \tag{16}$$

The above equation shows that the information form S-generator sequence can be obtained as follows

$$[g_I^S(\cdot), \gamma_I^S(\cdot)] = [-Jg^S(\cdot)J, J\gamma^S(\cdot)] \tag{17}$$

where we used repeatedly the relation between transfer domain and scattering domain generators and some easy matrix multiplications. Explicitly the information form generator is therefore also a 180°-rotated generator (up to an additional sign change) and given by

$$\left[\begin{array}{cc|c} -A_t^* & C_t^* R_t^{-1} C_t & C_t^* R_t^{-1} z(t) \\ -B_t Q_t B_t^* & -A_t & 0 \end{array}\right] \tag{18}$$

providing equations for $\eta_t = P_t\hat{x}(t \mid \Delta_t)$ and $W_t = P_t^{-1}$ which are indeed the information form recursions obtained by direct differentiation and using the original Kalman filter equations. From the fact that the corresponding generator sequence is a rotated version of the original generator one also expects the formal similarity to backwards evolution equations. This fact however seems to have no apparent deeper interpretation or meaning.

In the discrete case the corresponding generators are also easily found, however they do not obey the same rotation type of relation to the original generators. The computation of the generators involves, again, several cascadings. The result also shows that to propagate information forms is as easy (or difficult) as to solve the original recursions. Computation of the new generator takes only $O(\mathit{dimension}^2)$ operations, and in the case of a time-invariant model only part of it has to be recomputed at each step, involving one matrix multiplication.

5. Concluding remarks

Using Redheffer scattering results the derivation backwards filters and models, as well as information forms, is a rather simple task. In order to obtain the medium generators, which immediately provide the time update formulas, one only needs to perform some simple algebraic manipulations. In deriving the reverse filtering recursions and the information forms it becomes clear that the similarities between them are due to the fact that both recursions involve rotated generator sequences. The similarity is however formal only and there is no obvious deeper meaning to it.

Since the Redheffer scattering results provide a rather "automatic" way to address all state-space problems the intuition that could be gained through different derivations, for example the ones based on the innovations principle and the orthogonal projection interpretation of estimation, is necessarily lost. Therefore scattering theory should be viewed as an exceptional tool for deriving the algorithms, and later one may gain further interpretations and insights through alternative approaches.

References

[1] D. Q. Mayne: A SOLUTION OF THE SMOOTHING PROBLEM FOR LINEAR DYNAMICAL SYSTEMS, Automatica, vol4, p73, 1966

[2] D. C. Fraser: A NEW TECHNIQUE FOR OPTIMAL SMOOTHING OF DATA, PhD Dissertation, M.I.T., Cambridge, 1967

[3] L. Ljung and T. Kailath: BACKWARDS MARKOVIAN MODELS FOR SECOND ORDER STOCHASTIC PROCESSES, IEEE Trans. I.T., vol IT-22/4, p49 , 1976

[4] G. Verghese and T. Kailath: A FURTHER NOTE ON BACKWARDS MARKOVIAN MODELS, IEEE Trans. I.T., vol IT-25/1, p121, 1979

[5] L. Ljung, T. Kailath, B. Friedlander: SCATTERING THEORY AND LINEAR LEAST SQUARES ESTIMATION, Part I: Continuous Time Problems, Proc. IEEE, vol 64, p1, 1976

[6] B. Friedlander, T. Kailath, L. Ljung: SCATTERING THEORY AND LINEAR LEAST SQUARES ESTIMATION, Part II: Discrete Time Problems, J. Franklin Inst., vol 301, p71, 1976

[7] A. Jazwinski: STOCHASTIC PROCESSES AND FILTERING THEORY, Academic Press, N.Y., 1970

[8] F. C. Schweppe: UNCERTAIN DYNAMIC SYSTEMS, Prentice Hall, N.Y., 1973

[9] A. Gelb (ed): APPLIED OPTIMAL ESTIMATION, MIT Press, Cambridge, 1974

[10] G. Verghese, B. Friedlander, T. Kailath: SCATTERING THEORY AND LINEAR LEAST SQUARES ESTIMATION, Part III: The Estimates, IEEE Trans. on Aut. Contr., vol AC-25, p794, 1980

[11] L. Ljung, T. Kailath: A UNIFIED APPROACH TO SMOOTHING FORMULAS, Automatica, vol 12, p147, 1976

[12] A. E. Bryson, M. Frazier: SMOOTHING FOR LINEAR AND NONLINEAR DYNAMIC SYSTEMS, ASD-TDR-63-119, 1963

[13] T. Kailath: REDHEFFER SCATTERING THEORY AND LINEAR STATE-SPACE ESTIMATION PROBLEMS, Ricerche di Automatica, vol X/2, p136, 1979