LUND UNIVERSITY

Smoothing for Discrete Time Systems Using Operator Factorization

Hagander, Per

1972

Document Version:
Publisher's PDF, also known as Version of record

Link to publication

Citation for published version (APA):
Hagander, P. (1972). *Smoothing for Discrete Time Systems Using Operator Factorization*. (Technical Reports TFRT-7020). Department of Automatic Control, Lund Institute of Technology (LTH).

Total number of authors:
1

General rights
Unless other specific re-use rights are stated the following general rights apply:
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.
• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: https://creativecommons.org/licenses/

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

Download date: 20. Dec. 2025

SMOOTHING FOR DISCRETE TIME SYSTEMS USING OPERATOR FACTORIZATION

PER HAGANDER

REPORT 7214(B) JULY 1972
LUND INSTITUTE OF TECHNOLOGY
DIVISON OF AUTOMATIC CONTROL

Smoothing for discrete time systems using operator factorization

Per Hagander

Abstract

Linear discrete time systems can be described using operators instead of difference equations. For finite time interval the operators could also be interpreted as large three dimensional matrices.

The covariances for a stochastic system are easily expressed in this notation giving a neat solution to the smoothing problem by use of the projection theorem:

$$\hat{x} = R_{xy}R_y^{-1}y$$

It is then shown how R_y is triangularized by the Riccati equation so that the estimate can be obtained from adjoint forward and backward difference equations.

1. Introduction

The two approaches to linear estimation problems, the Wiener filter using covariance functions and the Kalman filter directly using difference or differential equations can be unified by use of the Riccati equation.

In [2] the continuous time linear control and estimation problems were analyzed using operators in function spaces. The same technique is applicable in the discrete time case. This is here demonstrated on the smoothing problem. The projection theorem gives an equation in covariance operators from which the difference equations are obtained by operator factorization using the discrete Riccati equation. The resolvent identity searched for by Kailath and Frost [3] is thus presented.

Although the discrete time problems are conceptually simpler than for continuous time, the algebra is somewhat more involved. The one step ahead predictor plays the role of the filter estimate and the operator factorization contains a special direct term as well as the forward and backward terms.

2. Operator notation

Consider the discrete time system

$$\begin{cases} x(t+1) = \phi(t+1,\ t)x(t) + v(t) \\ x(t_o) = x_o \end{cases} \tag{2.1}$$

with the solution

$$x(t) = \phi(t,t_o)x_o + \sum_{s=t_o}^{t-1} \phi(t,s+1)v(s)$$

where

$$\phi(t,s) = \begin{cases} \prod_{k=s}^{t-1} \phi(k+1,k) & s < t \\ I & s = t \\ 0 & s > t \end{cases}$$

The state of (2.1) at time t can be regarded as the value x(t) of a function x on $\{t_o,t_o+1,\ldots,t_1\}$. The difference equation (2.1) can now be formulated using operators in a space X of such functions:

$$x = gx_o + Lv$$

with

$g: R^n \to X$

$L: X \to X$

Define in X the scalar product

$$x_1 \cdot x_2 = \sum_{t=t_o}^{t_1} x_1^T(t)x_2(t)$$

giving the adjoint of L:

$$L^*: X \to X \; ; \; z = L^*x, \quad z(t) = \sum_{s=t+1}^{t_1} \phi^T(s,t+1)x(s)$$

Define also the restrictions

$$T_o : X \to R^n \quad : \quad T_o x = x(t_o)$$

$$T_1 : X \to R^n \quad : \quad T_1 x = x(t_1)$$

Some useful properties of these operators are now listed:

- $T_o L = 0$

  $T_o g = L$ (the identity)

- $g^* = T_o(\phi^T L^* + I)$ (2.2)

  where ϕ^T is regarded as a diagonal operator in X:

  $\phi^T : X \to X : z = \phi^T x, \; z(t) = \phi^T(t+1,t)x(t)$

- L has no inverse, but introduce the forward shift operator

  $q : X \to X : z = qx, \; z(t) = x(t+1)$

  giving

$$(q-\phi)L = I$$

and

$$L(q-\phi) = I - gT_o \tag{2.3}$$

o Correspondingly

$$(q^{-1} - \phi^T)L^* = I$$

and

$$L^*(q^{-1} - \phi^T) = I - hT_1 \tag{2.4}$$

with

$$h: R^n \to X; \; z = ha, \; z(t) = \phi^T(t_1+1,t+1)a$$

o $h^* = T_1(\phi L + I)$

o $(q-\phi)g = 0$

$(q^{-1}-\phi^T)h = 0$

o $\phi L = qL - I$

$\phi^T L^* = q^{-1}L^* - I$ (2.5)

o $L\phi = Lq + gT_o - I$

$L^*\phi^T = L^*q^{-1} + hT_1 - I$

Remark: Note that X could be regarded as a space of matrices with a trace norm and the operators as three dimensional matrices .
L is then lower triangular, L^* upper triangular and ϕ diagonal. L^* is the transpose of L using this interpretation.

3. Linear, stochastic, time discrete systems

The space X can be extended to contain stochastic processes generated by linear systems driven by white noise. Let v and e of

$$\begin{cases} x(t+1) = \phi x(t) + v(t), \quad x(t_o) = x_o \\ y(t) = \theta x(t) + e(t) \end{cases} \tag{3.1}$$

be zero mean, independent white noise with covariances R_1 and R_2 (>0), and let x_o have zero mean value, covariance R_o and be independent of v and e. The operators L and g are directly generalized. A new scalar product

$$x_1 \cdot x_2 = E \sum_{t=t_o}^{t_1} x_1^T(t) x_2(t)$$

gives the same adjoints. Notice that the deterministic functions constitute a subspace.

The covariance function, $r_x(t,s)$ of x is wellknown

$$r_x(t,s) = \begin{cases} \phi(t,s)R(s) & s < t \\ R(t) & s = t \\ R(t)\phi^T(s,t) & s > t \end{cases}$$

with R(t) from

$$\begin{cases} R(t+1) = \phi R(t) \phi^T + R_1 \\ R(t_o) = R_o \end{cases} \tag{3.2}$$

Let R_x be the operator with $r_x(t,s)$ as kernel. Thus

$$R_x = L\phi R + R + R\phi^T L^*$$

Using the shift operator q (3.2) gives

$$\begin{cases} R_1 = qRq^{-1} - \phi R\phi^T = (q-\phi)R(q^{-1}-\phi^T) + \phi R(q^{-1}-\phi^T) + (q-\phi)R\phi^T \\ R_o T_o = T_o R \end{cases} \tag{3.3}$$

Now (2.2) - (2.4) and (3.3) give

$$R_x = L\phi R + R + R\phi^T L = (I-gT_o)R + L\phi R + (I-gT_o)R\phi^T L^* + gR_o g^* =$$

$$= LR_1 L^* + gR_o g^* \tag{3.4}$$

Using the matrix point of view this is obvious from

$$x = Lv + gx_o$$

4. Smoothing estimate

Now find the best linear estimate $\hat{x}$ of x using $\{y(t_o),\ldots,y(t_1)\}$, best meaning minimal error variance in each component $\hat{x}_i(t)$. This means find the best linear operator F with

$$\hat{x} = Fy$$

Using the projection theorem F must satisfy

$$R_{xy} = FR_y$$

where R_{xy} and R_y have kernels $r_{xy}(t,s)$ and $r_y(t,s)$.
But (3.4):

$$R_{xy} = R_x\theta^T = (LR_1L^* + gR_og^*)\theta^T$$

$$R_y = \theta R_x\theta^T + R_2 = \theta(LR_1L^* + gR_og^*)\theta^T + R_2$$

Introduce P by

$$\begin{cases} P(t_o) = R_o \\ P(t+1) = \phi P\phi^T + R_1 - \phi P\theta^T(\theta P\theta^T + R_2)^{-1}\theta P\phi^T \end{cases} \tag{4.1}$$

or

$$\begin{cases} T_oP = R_oT_o \\ R_1 = qPq^{-1} - \phi P\phi^T + \phi P\theta^T(\theta P\theta^T+R_2)^{-1}\theta P\phi^T = \\ = (q-\phi)P(q^{-1}-\phi^T) + \phi P(q^{-1}-\phi^T) + (q-\phi)P\phi^T + \phi P\theta^T(\theta P\theta^T+R_2)^{-1}\theta P\phi^T \end{cases}$$

so using (2.2) - (2.4) as in section 3

$$R_x = L\phi P + P + P\phi^T L^* + L\phi P\theta^T(\theta P\theta^T + R_2)^{-1}\theta P\phi^T L^*$$

giving

$$R_y = [(\theta P\theta^T + R_2) + \theta L\phi P\theta^T]\ (\theta P\theta^T + R_2)^{-1}\ [(\theta P\theta^T + R_2) + \theta P\phi^T L^*\theta^T]$$

$$R_{xy} = P\theta^T + P\phi^T L^*\theta^T + L\phi P\theta^T(\theta P\theta^T + R_2)^{-1}\ [(\theta P\theta^T + R_2) + \theta P\phi^T L^*\theta^T]$$

Note that

$$y = [(\theta P\theta^T + R_2) + \theta L\phi P\theta^T]\ u$$

means the invertible dynamical system

$$\begin{cases} x(t+1) = \phi x(t) + \phi P\theta^T u(t), \quad x(t_o) = 0 \\ y(t) = \theta x(t) + (\theta P\theta^T + R_2) u(t) \end{cases}$$

Thus R_y invertible and

$$\hat{x} = R_{xy} R_y^{-1} y = \{P(I + \phi^T L^*)[\theta P\theta^T + R_2 + \theta P\phi^T L^*\theta^T]^{-1}\ (\theta P\theta^T + R_2) + L\phi P\theta^T\}$$

$$[\theta P\theta^T + R_2 + \theta P\phi^T L^*\theta^T]^{-1} y$$

Now use an identity for inverse dynamical systems analogous to the one in [1] for the continuous time case:

$$L\phi P\theta^T(\theta P\theta^T+R_2+\theta L\phi P\theta^T)^{-1} = [q-\phi+\phi P\theta^T(\theta P\theta^T+R_2)^{-1}\theta]^{-1}[\phi P\theta^T(\theta P\theta^T+R_2)^{-1}]$$

$$= (q-\phi+K\theta)^{-1}K$$

with

$$K = \phi P\theta^T(\theta P\theta^T+R_2)^{-1} \tag{4.2}$$

Define $\hat{x}_p = (q-\phi+K\theta)^{-1}Ky$ or

$$\hat{x}_p(t+1) = \phi\hat{x}_p(t)+K(y(t)-\theta\hat{x}_p(t)) = (\phi-K\theta)\hat{x}_p(t) + Ky(t),\ \hat{x}_p(t_o) = 0$$

and $\psi(t,s)$ by $\psi(t+1,t) = \phi(t+1,t) - K(t)\theta(t)$

Thus using (2.5)

$$\hat{x} = \hat{x}_p + Pq^{-1}L^*\theta^T[\theta P\theta^T+R_2+\theta P\phi^T L^*\theta^T]^{-1}(\theta P\theta^T+R_2)[\theta P\theta^T+R_2+\theta L\phi P\theta^T]^{-1}y$$

but

$$L^*\theta^T[\theta P\theta^T+R_2+\theta P\phi^T L^*\theta^T]^{-1}(\theta P\theta^T+R_2) = [q^{-1}-\phi^T+\theta^T(\theta P\theta^T+R_2)^{-1}\theta P\phi^T]^{-1}\theta^T$$

and

$$\hat{x} = \hat{x}_p+Pq^{-1}[q^{-1}-\phi^T+\theta^T K^T]^{-1}\theta^T[\theta P\theta^T+R_2+\theta L\phi P\theta^T]^{-1}y =$$

$$= \hat{x}_p + Pq^{-1}[q^{-1} - \psi^T]^{-1}\theta^T(\theta P\theta^T + R_2)^{-1}[I - \theta(q-\psi)^{-1}K]\, y$$

$$= \hat{x}_p + Pq^{-1}[q^{-1} - \psi^T]^{-1}\theta^T(\theta P\theta^T + R_2)^{-1}(y - \theta\hat{x}_p)$$

It is now easily shown that $\hat{x}_p(t)$ is the best one step ahead prediction $\hat{x}(t|t-1)$ of $x(t)$

Summing up:

<u>Theorem</u>: The smoothing estimate of the system (3.1) is given by

$$\hat{x}(t|t_1) = \hat{x}(t|t-1) + P(t)\lambda(t-1) \qquad t_o \le t \le t_1$$

where

$$\begin{cases} \hat{x}(t+1|t) = \phi\hat{x}(t|t-1) + K(t)(y(t) - \theta\hat{x}(t|t-1)) \\ \hat{x}(t_o|t_o-1) = 0 \end{cases}$$

$$\begin{cases} \lambda(t-1) = (\phi - K\theta)^T\lambda(t) + \theta^T(\theta P\theta^T + R_2)^{-1}[y(t) - \theta\hat{x}(t|t-1)] \\ \lambda(t_1) = 0 \end{cases}$$

and $P(t)$ and $K(t)$ from (4.1) and (4.2) $\square$

<u>Remarks</u>: Introduce $\tilde{x}_p = x - \hat{x}_p$ and $\tilde{y}_p = y - \theta\hat{x}_p$!

(i) $\hat{x}_p(t+1) = \psi\hat{x}_p(t) + Ky(t) = \phi\hat{x}_p(t) + K\tilde{y}_p(t)$

Notice that

$$K = \phi P\theta^T(\theta P\theta^T+R_2)^{-1} = (\phi-\phi P\theta^T(\theta P\theta^T+R_2)^{-1}\theta)P\theta^T R_2^{-1} = (\phi-K\theta)P\theta^T R_2^{-1}$$

(ii) $\hat{x}_p(t+1)=\psi\hat{x}_p(t) + \psi P\theta^T R_2^{-1}y(t)$

Let $P(t,s)$ be the covariance function of $\tilde{x}_p$

(iii) $$P(t,s) = \begin{cases} \psi(t,s)P(s) & t > s \\ P(t) & t = s \\ P(t)\psi^T(s,t) & t < s \end{cases}$$

Now

$$\hat{x}_p(t) = \sum_{s=t_o}^{t-1} \psi(t,s)P(s)\theta^T R_2^{-1}y(s) = \sum_{s=t_o}^{t-1} \phi(t,s)P(s)\theta^T(\theta P\theta^T+R_2)^{-1}\tilde{y}_p(s)$$

and

$$\lambda(t-1) = \sum_{s=t}^{t_1} \psi^T(s,t)\theta^T(\theta P\theta^T+R_2)^{-1}\tilde{y}_p(s)$$

Note that

$$\mathrm{Cov}[y(t)|x(t)] = R_2$$
$$\mathrm{Cov}[\tilde{y}_p(t)|x(t)] = \theta P\theta^T + R_2$$

Define

$$\nu_m = \theta^T(\theta P\theta^T + R_2)^{-1}\tilde{y}_p$$

$$y_m = \theta^T R_2^{-1} y$$

Hence

$$\text{(iv)}\ \hat{x}(t) = \hat{x}(t|t_1) = \hat{x}_p(t) + P(t)\lambda(t-1) =$$

$$= \sum_{s=t_o}^{t-1} \phi(t,s)P(s)\nu_m(s) + \sum_{s=t}^{t_1} P(t)\psi^T(s,t)\nu_m(s) =$$

$$= \sum_{s=t_o}^{t-1} \psi(t,s)P(s)y_m(s) + \sum_{s=t}^{t_1} P(t)\psi^T(s,t)\nu_m(s) =$$

$$= \sum_{s=t_o}^{t-1} P(t,s)y_m(s) + \sum_{s=t}^{t_1} P(t,s)\nu_m(s)$$

This last formula is the discrete time analog of the adjoint interpretation by Kailath and Frost [3 eq 9].

References:

[1] Hagander, P: The use of Operator Factorization for Linear Control and Estimation, Submitted for publication in Automatica.

[2] Hagander, P: Inversion of a Dynamical System by an Operator Identity, Automatica Vol. 8. pp 361 - 362, May 72.

[3] Kailath, T and P Frost: An Innovations Approach to Least-Squares Estimation - Part II:"Linear Smoothing in Additive White Noise, IEEE AC-16 pp 655 - 660, Dec 68.