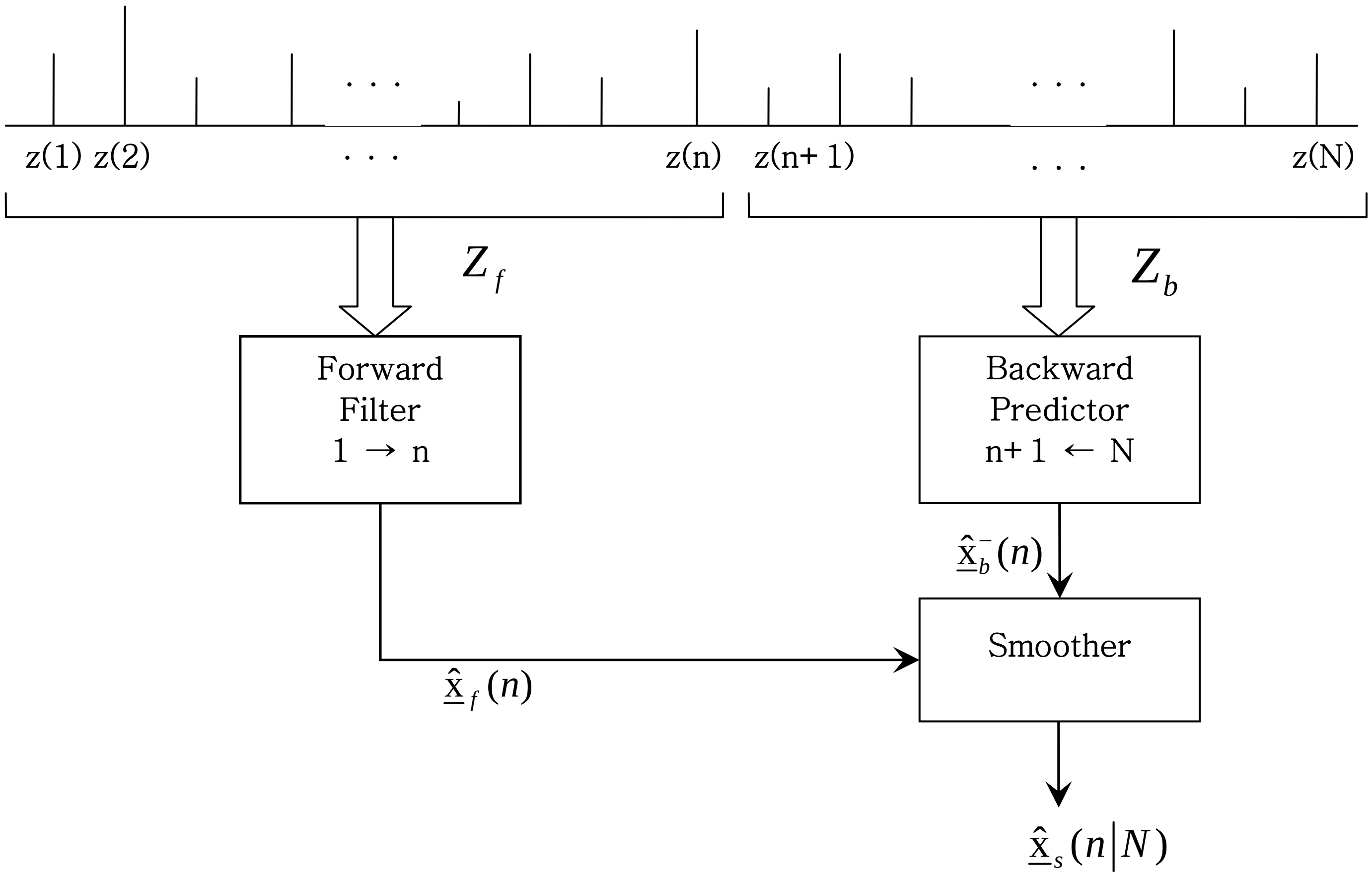
6. Optimal Smoothing

6.1 Concept



Given z(N), sought $\underline{\hat{x}}(n|N)$

(1). Fixed-interval smoothing: N-fixed, $0 < n < N$

(2). Fixed-point smoothing: n-fixed, N-$n>0$

(3). Fixed-lag smoothing: $n = N - k$, $k -$ fixed, $N -$ varies

6.2 Formulation of an Optimal Smoother

Start with a recursive form

$$\underline{\hat{\mathrm{x}}}_{\mathrm{s}}(n|N) = A_f(n)\underline{\hat{\mathrm{x}}}_{\mathrm{f}}(n) + A_b(n)\underline{\hat{\mathrm{x}}}_{\mathrm{b}}^{-}(n). \tag{6.2-1}$$

Define the error terms

$$\underline{\tilde{\mathrm{x}}}_{\mathrm{s}}(n|N) = \underline{\hat{\mathrm{x}}}_{\mathrm{s}}(n|N) - \underline{\mathrm{x}}(n)$$
$$\underline{\tilde{\mathrm{x}}}_{\mathrm{f}}(n) = \underline{\hat{\mathrm{x}}}_{\mathrm{f}}(n) - \underline{\mathrm{x}}(n)$$
$$\underline{\tilde{\mathrm{x}}}_{\mathrm{b}}^{-}(n) = \underline{\hat{\mathrm{x}}}_{\mathrm{b}}^{-}(n) - \underline{\mathrm{x}}(n)$$

Rewrite Eq. (6.2-1) using the error terms

$$\underline{\tilde{\mathrm{x}}}_{\mathrm{s}}(n|N) = [A_f(n) + A_b(n) - I]\underline{\mathrm{x}}(n) + A_f(n)\underline{\tilde{\mathrm{x}}}_{\mathrm{f}}(n) + A_b(n)\underline{\tilde{\mathrm{x}}}_{\mathrm{b}}^{-}(n). \tag{6.2-2}$$

To obtain an unbiased filter, we should have

$$A_f(n) + A_b(n) - I = 0. \tag{6.2-3}$$

Insert Eq. (6.2-3) to Eq. (6.2-2)

$$\underline{\tilde{\mathrm{x}}}_{\mathrm{s}}(n|N) = A_f(n)\underline{\tilde{\mathrm{x}}}_{\mathrm{f}}(n) + (I - A_f(n))\underline{\tilde{\mathrm{x}}}_{\mathrm{b}}^{-}(n)\,.$$

Error covariance of the smoother is

$$\begin{aligned} P_s(n) &= E\{\underline{\tilde{\mathrm{x}}}_{\mathrm{s}}(n|N)\underline{\tilde{\mathrm{x}}}_{\mathrm{s}}^{T}(n|N)\} \\ &= A_f(n)P_f(n)A_f^{T}(n) + (I - A_f(n))P_b^{-}(n)(I - A_f(n))^{T} \end{aligned} \tag{6.2-4}$$

To obtain a minimum variance filter, it should hold that

$$\frac{\partial}{\partial A_f(n)}[tr\,P_s(n)] = 2A_f(n)P_f(n) + 2(I - A_f(n))P_b^{-}(n)(-I) = 0,$$

that is,

$$A_f(n) = P_b^{-}(n)(P_f(n) + P_b^{-}(n))^{-1} \tag{6.2-5}$$

or

$$(I - A_f(n)) = P_f(n)(P_f(n) + P_b^{-}(n))^{-1}. \tag{6.2-6}$$

Inserting Eqs. (6.2-5) and (6.2-6) to Eq. (6.2-4) and arranging gives

$$
\begin{aligned}
P_s(n) &= A_f(n)P_f(n)A_f^{\,T}(n) + (I - A_f(n))P_b^-(n)(I - A_f(n))^T \\
&= P_b^-(n)\left[P_f(n) + P_b^-(n)\right]^{-1} P_f(n)\left[P_f(n) + P_b^-(n)\right]^{-T}\left[P_b^-(n)\right]^T \\
&\quad + P_f(n)\left[P_f(n) + P_b^-(n)\right]^{-1} P_b^-(n)\left[P_f(n) + P_b^-(n)\right]^{-T}\left[P_f(n)\right]^T \\
&= P_b^-(n)\left[P_f(n) + P_b^-(n)\right]^{-1} P_f(n)\left[I + \left(P_b^-(n)\right)^{-1} P_f(n)\right]^{-1} \\
&\quad + P_f(n)\left[P_f(n) + P_b^-(n)\right]^{-1} P_b^-(n)\left[P_f^{-1}(n)P_b^-(n) + I\right]^{-1} \\
&= P_b^-(n)\left[P_f(n) + P_b^-(n)\right]^{-1}\left[P_f^{-1}(n) + \left(P_b^-(n)\right)^{-1}\right]^{-1} \\
&\quad + P_f(n)\left[P_f(n) + P_b^-(n)\right]^{-1}\left[P_f^{-1}(n) + \left(P_b^-(n)\right)^{-1}\right]^{-1} \\
&= \left[P_b^-(n) + P_f(n)\right]\left[P_f(n) + P_b^-(n)\right]^{-1}\left[P_f^{-1}(n) + \left(P_b^-(n)\right)^{-1}\right]^{-1} \\
&= \left[P_f^{-1}(n) + \left(P_b^-(n)\right)^{-1}\right]^{-1}
\end{aligned}
$$

or

$$[P_s(n)]^{-1} = [P_f(n)]^{-1} + [P_b^-(n)]^{-1}. \tag{6.2-7}$$

Since $P_f(n), P_b^-(n), P_s(n)$ are all positive semidefinite, we see

$$P_s(n) \le P_f(n).$$

Inserting Eqs. (6.2-5) and (6.2-6) to Eq. (6.2-1) gives

$$\begin{aligned}
\underline{\hat{\mathrm{x}}}_{\mathrm{s}}(n|N) &= A_f(n)\underline{\hat{\mathrm{x}}}_{\mathrm{f}}(n) + (I - A_f(n))\underline{\hat{\mathrm{x}}}_{\mathrm{b}}^-(n) \\
&= P_b^-(n)\left[P_f(n) + P_b^-(n)\right]^{-1}\underline{\hat{\mathrm{x}}}_{\mathrm{f}}(n) + P_f(n)\left[P_f(n) + P_b^-(n)\right]^{-1}\underline{\hat{\mathrm{x}}}_{\mathrm{b}}^-(n) \\
&= \left[\left(P_b^-(n)\right)^{-1}P_f(n) + I\right]^{-1}\underline{\hat{\mathrm{x}}}_{\mathrm{f}}(n) + \left[I + P_f^{-1}(n)P_b^-(n)\right]^{-1}\underline{\hat{\mathrm{x}}}_{\mathrm{b}}^-(n) \\
&= \left[\left(P_b^-(n)\right)^{-1}P_f(n) + P_f^{-1}(n)P_f(n)\right]^{-1}\underline{\hat{\mathrm{x}}}_{\mathrm{f}}(n) + \left[\left(P_b^-(n)\right)^{-1}P_b^-(n) + P_f^{-1}(n)P_b^-(n)\right]^{-1}\underline{\hat{\mathrm{x}}}_{\mathrm{b}}^-(n) \\
&= \left[\left(P_b^-(n)\right)^{-1} + P_f^{-1}(n)\right]^{-1}P_f^{-1}(n)\underline{\hat{\mathrm{x}}}_{\mathrm{f}}(n) + \left[\left(P_b^-(n)\right)^{-1} + P_f^{-1}(n)\right]^{-1}\left(P_b^-(n)\right)^{-1}\underline{\hat{\mathrm{x}}}_{\mathrm{b}}^-(n) \\
&= P_s(n)\left[P_f^{-1}(n)\underline{\hat{\mathrm{x}}}_{\mathrm{f}}(n) + \left(P_b^-(n)\right)^{-1}\underline{\hat{\mathrm{x}}}_{\mathrm{b}}^-(n)\right].
\end{aligned} \tag{6.2-8}$$

6.3 Optimal Fixed-Interval Smoother

Given the system

$$\begin{aligned} &\underline{\mathrm{x}}(n+1) = \Phi\underline{\mathrm{x}}(n) + \Gamma\underline{\mathrm{w}}(n), \quad \underline{\mathrm{w}}(n) \sim N(\underline{0}, Q(n)) \\ &\underline{\mathrm{z}}(n) = H\underline{\mathrm{x}}(n) + \underline{\mathrm{v}}(n), \quad \underline{\mathrm{v}}(n) \sim N(\underline{0}, R(n)) \end{aligned} \tag{6.3-1}$$

the forward (Kalman) filter is written

$$\hat{\underline{\mathrm{x}}}_{\mathrm{f}}(n) = \hat{\underline{\mathrm{x}}}_{\mathrm{f}}^{-}(n) + K_f(n)[\underline{\mathrm{z}}(n) - H\hat{\underline{\mathrm{x}}}_{\mathrm{f}}^{-}(n)] \tag{6.3-2}$$

$$P_f(n) = P_f^{-}(n) - K_f(n)HP_f^{-}(n) \tag{6.3-3}$$

$$K_f(n) = P_f^{-}(n)H^T\left[HP_f^{-}(n)H^T + R(n)\right]^{-1}. \tag{6.3-4}$$

$$\hat{\underline{\mathrm{x}}}_f^{-}(n+1) = \Phi\hat{\underline{\mathrm{x}}}_f(n) \tag{6.3-5}$$

$$P_f^{-}(n+1) = \Phi P_f(n)\Phi^T + \Gamma Q(n)\Gamma^T \tag{6.3-6}$$

For the backward filter (predictor), define

$$S(n) \triangleq P_b^{-1}(n) = \left(Cov\left[\underline{\mathrm{x}}(n) - \hat{\underline{\mathrm{x}}}_{\mathrm{b}}(n)\right]\right)^{-1} \tag{6.3-7}$$

and

$$S^{-}(n) \triangleq \left(P_b^{-}(n)\right)^{-1} = \left(Cov\left[\underline{x}(n) - \hat{\underline{x}}_{\mathrm{b}}^{-}(n)\right]\right)^{-1}. \tag{6.3-8}$$

It will also be convenient to define

$$\hat{\underline{y}}(n) \triangleq P_b^{-1}(n)\hat{\underline{x}}_{\mathrm{b}}(n) = S(n)\hat{\underline{x}}_{\mathrm{b}}(n), \tag{6.3-9}$$

and

$$\hat{\underline{y}}^{-}(n) \triangleq \left(P_b^{-}(n)\right)^{-1}\hat{\underline{x}}_{\mathrm{b}}^{-}(n) = S^{-}(n)\hat{\underline{x}}_{\mathrm{b}}^{-}(n). \tag{6.3-10}$$

Now, rewrite Eq. (6.3-1) for $\underline{x}(n)$

$$\underline{x}(n) = \Phi^{-1}\underline{x}(n+1) - \Phi^{-1}\Gamma\underline{w}(n).$$

Since $\hat{\underline{w}}(n) = \underline{0}$, the optimum a priori backward estimate is

$$\hat{\underline{x}}_{\mathrm{b}}^{-}(n) = \Phi^{-1}\hat{\underline{x}}_{\mathrm{b}}(n+1). \tag{6.3-11}$$

The equivalent form for the Kalman filter in Section 4.1 provides two useful results. First, Eq. (4.1-28) gives

$$P_b^{-1}(n) = \left(P_b^-(n)\right)^{-1} + H^T R^{-1}(n) H \,.$$

As a result of Definitions Eqs. (6.3-7) and (6.3-8), this equation becomes

$$S(n) = S^-(n) + H^T R^{-1}(n) H \,. \tag{6.3-12}$$

Second, from Eq. (4.1-27) we have

$$P_b^{-1}(n)\underline{\hat{\mathrm{x}}}_{\mathrm{b}}(n) = \left(P_b^-(n)\right)^{-1} \underline{\hat{\mathrm{x}}}_{\mathrm{b}}^-(n) + H^T R^{-1}(n)\underline{\mathrm{z}}(\mathrm{n}) \,,$$

or equivalently,

$$\underline{\hat{\mathrm{y}}}(n) = \underline{\hat{\mathrm{y}}}^-(n) + H^T R^{-1}(n)\underline{\mathrm{z}}(\mathrm{n}) \,. \tag{6.3-13}$$

To find $S^-(n)$, write the covariance matrix $P_b^-(n)$ as

$$
\begin{aligned}
P_b^-(n) &= Cov\left[\underline{\mathrm{x}}(n) - \hat{\underline{\mathrm{x}}}_\mathrm{b}^-(n)\right] \\
&= Cov\left[\Phi^{-1}\underline{\mathrm{x}}(n+1) - \Phi^{-1}\Gamma\underline{\mathrm{w}}(n) - \Phi^{-1}\hat{\underline{\mathrm{x}}}_\mathrm{b}(n+1)\right] \\
&= \Phi^{-1}\left[P_b(n+1) + \Gamma Q(n)\Gamma^T\right]\Phi^{-T}.
\end{aligned}
$$

Then apply Eq. (6.3-7) to obtain

$$S^-(n) = \left[\Phi^{-1}\left[S^{-1}(n+1) + \Gamma Q(n)\Gamma^T\right]\Phi^{-T}\right]^{-1}.$$

By the matrix inversion lemma, this expression becomes

$$S^-(n) = \Phi^T\left[S(n+1) - S(n+1)\Gamma\left[\Gamma^T S(n+1)\Gamma + Q^{-1}(n)\right]^{-1}\Gamma^T S(n+1)\right]\Phi.$$

Define a backward Kalman gain by

$$K_b(n) = S(n+1)\Gamma\left[\Gamma^T S(n+1)\Gamma + Q^{-1}(n)\right]^{-1}, \qquad (6.3\text{-}14)$$

so that our expresion for $S^-(n)$ is

$$S^-(n) = \Phi^T \left[I - K_b(n)\Gamma^T \right] S(n+1)\Phi\,. \tag{6.3-15}$$

Finally, we need an equation for computing $\hat{\underline{\mathrm{y}}}^-(n)$ from $\hat{\underline{\mathrm{y}}}(n+1)$. Using Eqs. (6.3-10), (6.3-11), and (6.3-9) we find

$$\begin{aligned}\hat{\underline{\mathrm{y}}}^-(n) &= S^-(n)\hat{\underline{\mathrm{x}}}_{\mathrm{b}}^-(n) \\ &= S^-(n)\Phi^{-1}\hat{\underline{\mathrm{x}}}_{\mathrm{b}}(n+1) \qquad . \\ &= S^-(n)\Phi^{-1}S^{-1}(n+1)\hat{\underline{\mathrm{y}}}(n+1)\end{aligned}$$

Substitute Eq. (6.3-15) for $S^-(n)$ to obtain

$$\hat{\underline{\mathrm{y}}}^-(n) = \Phi^T \left[I - K_b(n)\Gamma^T \right] \hat{\underline{\mathrm{y}}}(n+1)\,. \tag{6.3-16}$$

Now, we need to formulate $P_s^{-1}(n)$ and $\hat{\underline{\mathrm{x}}}_{\mathrm{s}}(n|N)$ in terms of $S^-(n)$ and $\hat{\underline{\mathrm{y}}}^-(n)$ so that we can directly connect the backward filter formulation with the formulation of the optimal smoother. Eq. (6.2-7) may be rewritten by employing $S^-(n)$ and then applying the matrix inversion lemma

$$
\begin{aligned}
P_s(n) &= ([P_f(n)]^{-1} + [P_b^-(n)]^{-1})^{-1} \\
&= ([P_f(n)]^{-1} + S^-(n)II)^{-1} \\
&= P_f(n) - P_f(n)S^-(n)\left[IP_f(n)S^-(n) + I\right]^{-1} IP_f(n)
\end{aligned}
$$

Define the smoothing gain by

$$
K_s(n) = P_f(n)S^-(n)\left[I + P_f(n)S^-(n)\right]^{-1},
$$

so the final form for $P_s(n)$ be

$$
P_s(n) = \left[I - K_s(n)\right]P_f(n). \tag{6.3-17}
$$

Similarly, Eq. (6.2-8) may be rewritten by employing $\underline{\hat{\mathrm{y}}}^-(n)$ and the matrix inversion lemma

$$
\begin{aligned}
\underline{\hat{\mathrm{x}}}_{\mathrm{s}}(n|N) &= P_s(n)[[P_f(n)]^{-1}\underline{\hat{\mathrm{x}}}_{\mathrm{f}}(n) + [P_b^-(n)]^{-1}\underline{\hat{\mathrm{x}}}_{\mathrm{b}}^-(n)] \\
&= P_s(n)[P_f(n)]^{-1}\underline{\hat{\mathrm{x}}}_{\mathrm{f}}(n) + P_s(n)\underline{\hat{\mathrm{y}}}^-(n) \qquad (6.3\text{-}18) \\
&= \left[I - K_s(n)\right]\underline{\hat{\mathrm{x}}}_{\mathrm{f}}(n) + P_s(n)\underline{\hat{\mathrm{y}}}^-(n).
\end{aligned}
$$

Now, consider the initial conditions required by the smoothing filter. The initial conditions for the forward predictor are the same as in Section 5.1. For the backward predictor we observe that at time *N*, $P_s(N) = P_f(N)$. Write Eq. (6.2-7) for $n = N$ as

$$[P_s(N)]^{-1} = [P_f(N)]^{-1} + [P_b^-(N)]^{-1} = [P_f(N)]^{-1} + S^-(N),$$

and it follows that

$$S^-(N) = 0. \qquad (6.3\text{-}19)$$

Then Eq. (6.3-10) implies that

$$\underline{\hat{y}}^-(N) = 0. \qquad (6.3\text{-}20)$$

Eqs. (6.3-2), (6.3-3), (6.3-4), (6.3-5), (6.3-6), (6.3-14), (6.3-15), (6.3-16), (6.3-17), (6.3-18), (6.3-19), and (6.3-20) form an optimal smoother.

Summary of the Optimal Fixed-Interval Smoother

Given System	$\underline{x}(n+1) = \Phi \underline{x}(n) + \Gamma \underline{w}(n), \quad \underline{w}(n) \sim N(\underline{0}, Q(n))$ $\underline{z}(n) = H\underline{x}(n) + \underline{v}(n), \quad \underline{v}(n) \sim N(\underline{0}, R(n))$
Forward Filter	$\hat{\underline{x}}_f(n) = \hat{\underline{x}}_f^-(n) + K_f(n)[\underline{z}(n) - H\hat{\underline{x}}_f^-(n)]$, Given $\hat{\underline{x}}_f^-(1)$ $P_f(n) = P_f^-(n) - K_f(n) H P_f^-(n)$, Given $P_f^-(1)$ $K_f(n) = P_f^-(n) H^T \left[H P_f^-(n) H^T + R(n) \right]^{-1}$ $\hat{\underline{x}}_f^-(n+1) = \Phi \hat{\underline{x}}_f(n)$ $P_f^-(n+1) = \Phi P_f(n) \Phi^T + \Gamma Q(n) \Gamma^T$

Backward Filter	$S(n) = S^-(n) + H^T R^{-1}(n) H, \quad S^-(N) = 0$ $\underline{\hat{\mathrm{y}}}(n) = \underline{\hat{\mathrm{y}}}^-(n) + H^T R^{-1}(n) \underline{\mathrm{z}}(\mathrm{n}), \quad \underline{\hat{\mathrm{y}}}^-(N) = 0$ $\underline{\hat{\mathrm{y}}}^-(n) = \Phi^T \left[I - K_b(n) \Gamma^T \right] \underline{\hat{\mathrm{y}}}(n+1)$ $S^-(n) = \Phi^T \left[I - K_b(n) \Gamma^T \right] S(n+1) \Phi$ $K_b(n) = S(n+1) \Gamma \left[\Gamma^T S(n+1) \Gamma + Q^{-1}(n) \right]^{-1}$
Optimal Smoother	$K_s(n) = P_f(n) S^-(n) \left[I + P_f(n) S^-(n) \right]^{-1}$ $P_s(n) = \left[I - K_s(n) \right] P_f(n)$ $\underline{\hat{\mathrm{x}}}_{\mathrm{s}}(n \vert N) = \left[I - K_s(n) \right] \underline{\hat{\mathrm{x}}}_{\mathrm{f}}(n) + P_s(n) \underline{\hat{\mathrm{y}}}^-(n)$