Estimation of Stochastic Processes

$\{s_i\} \to$ signal process (unknown) $\}$ discrete time processes
$\{y_i\} \to$ measurement process (known) $\}$ zero mean

known $\leftarrow$ $\boxed{\begin{aligned} R_{sy}(i,l) &= E s_i y_l^T \\ R_y(i,l) &= E y_i y_l^T \end{aligned}}$

Three problems:

1) Smoothing: For each i, & for fixed $N \geq i$, estimate s_i given $\{y_j, 0 \leq j \leq N\}$

$$\hat{s}_{i/N} = \sum_{j=0}^{N} K_{s,ij} y_j \quad \text{(non-causal)}$$

$\longrightarrow$ for smoothing.

i.e. $\min \|s_i - \hat{s}_{i/N}\|^2$

2) Causal Filtering: Estimate s_i given only past & present observations: $\hat{s}_{i/i} = \sum_{j=0}^{i} K_{f,ij} y_j$

i.e. $\min \|s_i - \hat{s}_{i/i}\|^2$ $\boxed{\text{with } K_{f,ij} = 0 \text{ for } j > i}$

3) Prediction: For each i, & for fixed λ, estimate $s_{i+\lambda}$ given $\{y_i\}_{j=0}^{i}$: $\hat{s}_{i+\lambda/i} = \sum_{j=0}^{i} K_{\lambda,ij} y_j$

i.e. $\min \|s_{i+\lambda} - \hat{s}_{i+\lambda/i}\|^2$ (For $\lambda > 0 \to$ prediction)

The Smoothing problem is easy:

$\hat{s}_{i/N} = \sum_{j=0}^{N} K_{s,ij} y_j$. By our orthogonality principle: $\left(s_i - \sum_{j=0}^{N} K_{s,ij} y_j\right) \perp y_l,\ l = 0, \cdots, N$

i.e. $\left\langle \left(s_i - \sum K_{s,ij} y_j\right), y_l \right\rangle = 0, \ \forall\, l = 0, \cdots, N$

or $R_{sy}(i,l) = \sum_{j=0}^{N} K_{s,ij} R_y(j,l) \ \forall\, l = 0, \cdots, N$

i.e. $\left[R_{sy}(i,0) \;\cdots\; R_{sy}(i,N)\right] = \left[K_{s,i0} \cdots K_{s,iN}\right] R_y$ (1)

This is true for $\forall\, i = 0, \dots, N$.

Define $R_{sy} = \begin{bmatrix} R_{sy}(0,0) & \cdots & R_{sy}(0,N) \\ \vdots & \ddots & \vdots \\ R_{sy}(N,0) & \cdots & R_{sy}(N,N) \end{bmatrix}$; $K_s = \begin{bmatrix} K_{s,00} & \cdots & K_{s,0N} \\ \vdots & \ddots & \vdots \\ K_{s,N0} & \cdots & K_{s,NN} \end{bmatrix}$ $\quad R_y = \begin{bmatrix} R_y(0,0) & \cdots & R_y(0,N) \\ \vdots & \ddots & \vdots \\ R_y(N,0) & \cdots & R_y(N,N) \end{bmatrix}$

Define $s = \begin{bmatrix} s_0 \\ \vdots \\ s_N \end{bmatrix}$ $\quad \hat{s}_s = \begin{bmatrix} \hat{s}_{0/N} \\ \vdots \\ \hat{s}_{N/N} \end{bmatrix}$ $\quad y = \begin{bmatrix} y_0 \\ \vdots \\ y_N \end{bmatrix}$

Then $R_y = \langle y, y \rangle$, $R_{sy} = \langle s, y \rangle$

The eqn (1) for all i can be written:

$$\boxed{\hat{s}_s = K_s y \quad \& \quad K_s R_y = R_{sy}}$$

If $R_y > 0$, $\hat{s}_s = R_{sy} R_y^{-1} y = \langle s, y \rangle \langle y, y \rangle^{-1} y$

(Solution is same as the regular llmse)

Causal Filtering (much harder): Two conditions

1) $\left[K_{f,ij} = 0 \text{ for } j > i\right]$

2) $\left(s_i - \sum_{j=0}^{i} K_{f,ij} y_j\right) \perp y_l \quad \forall\, l = 0, \dots, i$

Condition (2) leads to (as before)

$$(*) \text{ — } R_{sy}(i,l) = \sum_{j=0}^{i} K_{f,ij} R_y(j,l), \quad l = 0, \dots, i$$

But because of (1), K_f must be lower (block)? triangular

$$\begin{bmatrix} \hat{s}_{0/0} \\ \vdots \\ \hat{s}_{N/N} \end{bmatrix} = \begin{bmatrix} K_{f,00} & & 0 \\ \vdots & K_{f,11} \ddots & \\ K_{f,N0} & \cdots & K_{f,NN} \end{bmatrix} \begin{bmatrix} y_0 \\ \vdots \\ y_N \end{bmatrix} = K_f y$$

Define A_{lower} by $\left[\{A\}_{lower}\right]_{ij} = \begin{cases} 0 & i<j \\ A_{ij} & i\geq j \end{cases}$

So (*) can be written as

$$\{R_{sy} - K_f R_y\}_{lower} = 0 \quad \text{———} \quad (1)$$

\# It is not immediately clear, how to solve (1).

Weiner-Hopf Technique:

We want $\left[R_{sy} - K_f R_y\right]_{lower} = 0$ i.e. $\left[R_{sy} - K_f R_y\right] = U^+$

[where $[U^+]_{lower} = 0$

Factorize $R_y = L R_e L^T$ ⟶ Exercise: when does such a factorization exist?

L: lower triangular, R_e: diagonal

Then $R_{sy} - K_f L R_e L^T = U^+$ (we want)

$$\underbrace{K_f L}_{\text{lower}} = \underbrace{R_{sy}(L^T)^{-1} R_e^{-1}}_{\text{mixed}} - \underbrace{U^+ (L^T)^{-1} R_e^{-1}}_{\text{strictly upper}}$$

$(L^T)^{-1} R_e^{-1}$ ⟶ upper triangular

Then it must be true that

$$K_f L = \{R_{sy}(L^T)^{-1} R_e^{-1}\}_{lower}$$

or

$$K_f = \{R_{sy}(L^T)^{-1} R_e^{-1}\}_{lower} L^{-1}$$

Lemma: The optimal K_f that solves the causal filtering problem is given by

$$K_f = \{R_{sy}(L^T)^{-1} R_e^{-1}\}_{lower} L^{-1}$$

Ex: Let $y = s + v$ with $\left\langle \begin{bmatrix} s \\ v \end{bmatrix}, \begin{bmatrix} s \\ v \end{bmatrix} \right\rangle = \begin{bmatrix} R_s & 0 \\ 0 & R_v \end{bmatrix}$

Further R_v is diagonal & non-singular.

Then $R_{sy} = R_s$ & $R_y = R_s + R_V$

$$K_f = \left\{ [R_y - R_V](L^T)^{-1} R_e^{-1} \right\}_{lower} L^{-1}$$

$$= \left\{ L - R_V (L^T)^{-1} R_e^{-1} \right\}_{lower} L^{-1}$$

$$= I - \left\{ R_V (L^T)^{-1} R_e^{-1} \right\}_{lower} L^{-1}$$

[since $R_y (L^T)^{-1} R_e^{-1} = L R_e L^T (L^T)^{-1} R_e^{-1} = L$]

Now $\{ R_V \underbrace{(L^T)^{-1}}_{\text{upper triangular}} R_e^{-1} \}_{lower} = R_V R_e^{-1}$ (since R_V, R_e are diagonal)

$$\Rightarrow \quad K_f = I - R_V R_e^{-1} L^{-1}$$

<u>Innovations Process</u> (Recursive solution of the normal eqns, using Gram-Schmidt orthogonalization)

Given $\{y_0, \cdots, y_N\}$ find a orthogonal set of vectors $\{e_0, \cdots, e_n\}$ s.t.

$$\mathcal{L}\{e_0, \cdots, e_n\} = \mathcal{L}\{y_0, \cdots, y_N\} =: \mathcal{L}_N \text{ (say)}$$

Given a new observation y_{N+1}, we can find a new vector e_{N+1} s.t. $\{e_0, \cdots, e_{N+1}\}$ is a orthogonal set and $\mathcal{L}\{e_0, \cdots, e_{N+1}\} = \mathcal{L}\{y_0, \cdots y_{N+1}\}$ by defining

$$e_{N+1} = y_{N+1} - \text{Proj}(y_{N+1} / \mathcal{L}_N\} \quad \text{—(1)}$$

Because of orthogonality of $\{e_0, \cdots, e_N\}$:

$$\text{Proj}\{y_{N+1} / \mathcal{L}_N\} = \sum_{j=0}^{N} \langle y_{N+1}, e_j \rangle \|e_j\|^{-2} e_j$$

Then from ① $e_{N+1} = y_{N+1} - \sum \langle y_{N+1}, e_j \rangle \|e_j\|^{-2} e_j$ ②

Recursive formula, use $e_0 = y_0$.

(Identical to Gram - Schmidt Orthogonalization)

Why innovations?

$\text{Proj}\{y_{N+1} | \mathcal{L}_N\}$ = llmse of y_{N+1} given $\mathcal{L}\{y_0, \cdots, y_N\}$

$=: \hat{y}_{N+1}$ (say)

$\hat{y}_{N+1}$ was already known from $\{y_0, \cdots, y_N\}$

New information : $e_{N+1} = y_{N+1} - \hat{y}_{N+1}$

then $\{e_i\} \to$ innovations process.

Two imp properties:

1) e_i is uncorrelated to e_j for all $i \neq j$.

2) $e_i \in \mathcal{L}\{y_0, \cdots, y_i\}$ → causal

AND $y_i \in \mathcal{L}\{e_0, \cdots, e_i\}$ → causally invertible

Iterative Formula (

$\hat{s}^{N+1}_{N+1}$ = llmse of s_{N+1} given $\{y_0, \cdots, y_N, y_{N+1}\}$

= llmse of s_{N+1} given $\{e_0, \cdots, e_N, e_{N+1}\}$

due to orthogonality.

$= \sum_{j=0}^{N} \langle s_{N+1}, e_j \rangle \|e_j\|^{-2} e_j + \{$ llmse of s_{N+1} given $e_{N+1}\}$

i.e. $\hat{s}_{N+1|N+1} = \hat{s}_{N+1|N} + \langle s_{N+1}, e_{N+1} \rangle \|e_{N+1}\|^{-2} e_{N+1}$

with $\hat{s}_{|-1} = 0$

$$e_{N+1} = y_{N+1} - \sum_{j=0}^{N} \langle y_{N+1}, e_j \rangle \|e_j\|^{-2} e_j, \quad e_0 = y_0$$

→ This formula not directly usable since R_{se} is not known & the calculation of R_{se} from R_{sy} depends on inversion of L at each step.

A purely iterative formula would require more structure on R_y. (stationarity)

Algebraic Derivation:

Consider eqn ②: $y_i = e_i + \sum_{j=0}^{i-1} \langle y_i, e_j \rangle \|e_j\|^{-2} e_j$

Collect these eqns: $(i = 0, \dots, N)$

$$\mathbf{y} = L_1 \mathbf{e}$$

where L_1 is lower (block) triangular:

In general block diag but our discussion is scalar diag

$$\begin{bmatrix} y_0 \\ \vdots \\ \\ \\ y_N \end{bmatrix} = \begin{bmatrix} I & & & & \\ \langle y_1, e_0 \rangle \|e_0\|^{-2} & I & & & \\ \langle y_2, e_0 \rangle \|e_0\|^{-2} & \langle y_2, e_1 \rangle \|e_1\|^{-2} & I & & \\ \vdots & & & \ddots & \\ \langle y_N, e_0 \rangle \|e_0\|^{-2} & \langle y_N, e_1 \rangle \|e_1\|^{-2} & \cdots & & I \end{bmatrix} \begin{bmatrix} e_0 \\ \vdots \\ \\ \\ e_N \end{bmatrix}$$

Since $y = L_1 e$, $R_y = E\bar{y}y^T = E\, L_1 e e^T L_1^T = L_1 R_e L_1^T$

Since e_i are uncorrelated by construction,

$R_e = \text{diag}\{E\|e_i\|^2\}$

But recall that the factorization $R_y = LDL^T$ is unique. So the innovations process defined this way is unique. $(L = L_1,\ D = R_e)$
(Some scaling is possible)

Causal & causally invertible :

$$y_i \longrightarrow \boxed{L^{-1}} \xrightarrow{\ e_i\ } \boxed{L} \xrightarrow{\ y_i\ }$$

Filtering using Innovations Process

$$\hat{s}_{i|i} := \text{llmse of } s_i \text{ given } \{y_0, \cdots, y_i\}$$
$$= \text{llmse of } s_i \text{ given } \{e_0, \cdots, e_i\}$$
$$= \sum_{j=0}^{i} \langle s_i, e_j \rangle \|e_j\|^{-2} e_j \quad \text{— } (1)$$

Recall $R_{se}(i,j) = E s_i e_j^T = \langle s_i, e_j \rangle$

From (1), $\hat{s}_{i|i} = \sum_{j=0}^{i} \underbrace{\langle s_i, e_j \rangle \|e_j\|^{-2}}_{g_{f,ij}} e_j$ for $j < i$

ie. $g_{f,ij} = \begin{cases} R_{se}(i,j)\, \|e_j\|^{-2}, & j \leq i \\ 0 & \text{otherwise} \end{cases}$

But we know R_{sy} & not R_{se}. But recall,

$e = L^{-1} y$. So, $R_{se} = \langle s, e \rangle = \langle s, y \rangle L^{-T}$

$= R_{sy} L^{-T} \longrightarrow (2)$

$G_f = [g_{f,ij}]$

So $G_f = \{R_{sy} L^{-T} R_e^{-1}\}_{lower}$

So (1) can be written for all $i = 0, \dots, N$ together as:

$$\hat{s}_f = G_f e = G_f L^{-1} y = \underbrace{\left[R_{sy} L^{-T} R_e^{-1}\right]_{lower}}_{K_f} L^{-1} y$$

Examples from Section 4.4 (Kailath)

Note: $y = Le$ or $e = L^{-1} y = Wy$

$e_i = w_{i0} y_0 + w_{i1} y_1 + w_{i2} y_2 + \dots + w_{i,i-1} y_{i-1} + y_i$

depends on i $\longrightarrow$ so time varying

This map becomes time-invariant for Stationary Processes observed from $i = -\infty$.

Stationary Process $\rightarrow$ Computation of corr innovations process.

Let $\{y_i\}$ $-\infty < i < \infty$ be scalar stationary, zero mean

$R_y(i) = \langle y_j, y_{j-i} \rangle$

Compute $e_i = y_i - \hat{y}_{i/i-1}$ for $-\infty < i < \infty$

$$R_y = \begin{bmatrix} \ddots & & & & & \\ & R_y(1) & R_y(0) & R_y^T(1) & & \\ & & R_y(1) & \boxed{R_y(0)} & R_y^T(1) & \\ & & & R_y(1) & R_y(0) & R_y^T(1) \\ & & & & & \ddots \end{bmatrix}$$

Doubly Infinite Toeplitz

We want to factor $R_y = LDL^T$, where

$$D = \begin{bmatrix} \ddots & & & & 0 \\ & r_e & & & \\ & & \boxed{r_e} & & \\ & & & r_e & \\ 0 & & & & \ddots \end{bmatrix} \quad L = \begin{bmatrix} \ddots & & & & 0 \\ \ddots & l_1 & 1 & & \\ & \ddots & l_1 & \boxed{1} & \\ & & \ddots & l_1 & 1 \\ & & & \ddots & \ddots \end{bmatrix}$$

We will require:

1) L & D to be doubly infinite
2) L to be Toeplitz (equivalently time invariant
3) L to be lower triangular causal map between $e_i \to y_i$)
4) L^{-1} to be Toeplitz (time inv. map $y_i \to e_i$)
5) L^{-1} to be lower triangular $\to$ causal map betw. $y_i \to e_i$
6) The maps L & L^{-1} should be stable (i.e. WSS processes should remain WSS)

$$y_k = e_k + \sum_{i=1}^{\infty} l_i e_{k-i} \quad \& \quad e_k = y_k + \sum_{i=1}^{\infty} w_i y_{k-i}$$

i.e. $\sum_{i=1}^{\infty} |l_i|^2 < \infty \quad \& \quad \sum_{i=1}^{\infty} |w_i|^2 < \infty$

A factorization is hard to do directly. So we take z-transforms: (z-Spectra)

Define 1) $[\cdots z^2 \; z \; \boxed{1} \; z^{-1} \; z^{-2} \cdots] =: \Lambda(z) \to$ vector

2) $S_y(z) = \sum_{i=-\infty}^{+\infty} R_y(i) z^{-i}$ Note: scalars
(z-spectra of y)

3) $L(z) = 1 + \sum_{i=1}^{\infty} l_i z^{-i}$, $L\left(\frac{1}{z}\right)$ accordingly.

4) $D(z) = r_e$

First check: $\Lambda(z) R_y = S_y(z) \Lambda(z)$

Eigenvector

Eigenvalue

$$\overbrace{[\cdots\ z\ \boxed{1}\ z^{-1}\ \cdots]}^{\Lambda(z)}\ \overbrace{\begin{bmatrix} R_y(0) & R_y(1) & R_y(2) & \cdots \\ R_y(1) & R_y(0) & R_y(1) & \cdots \\ R_y(2) & R_y(1) & R_y(0) & \cdots \end{bmatrix}}^{R_y} = [\cdots zR_y(0)\cdots \mid \cdots 1\cdot R_y(0)\cdots \mid \cdots z^{-1}R_y(0)\cdots]$$

$$= \overbrace{(\cdots + R_y(0) + \cdots)}^{S_y(z)}[\cdots\ z\ \boxed{1}\ z^{-1}\ \cdots] = S_y(z)\Lambda(z)$$

Next check.

$$\Lambda(z)\, L D L^T = L(z)\Lambda(z)\, D\, L^T = L(z)\, r_e\, L\left(\tfrac{1}{z}\right)\Lambda(z)$$

$$= L(z)\, r_e\, L^*(1/z^*)$$

Check:

$$[\cdots\ z^2\ z\ \boxed{1}\ z^{-1}\ z^{-2}] \begin{bmatrix} 1 & l_1 & l_2 & \cdots \\ 0 & \boxed{1} & l_1 & l_2 \\ 0 & 0 & 1 & l_1 \end{bmatrix} = [\cancel{z}\quad zl_1 + 1\quad zl_2 + l_1 + z^{-1}]$$

$$L(z) = 1 + l_1 z^{-1} + l_2 z^{-2} \qquad L\left(\tfrac{1}{z}\right) \doteq (1 + l_1 z + l_2 z^2)[z\ 1\ z^{-1}\cdots]$$

Similarly, $e(z) = L^{-1}(z)\, y(z)$ where $e(z) = \sum_{i=-\infty}^{+\infty} e_i z^{-i}$

$y(z) = \sum_{i=-\infty}^{+\infty} y_i z^{-i}$

\# Note: $S_y(e^{j\omega})$ is power spectral density of y_i.

\# We want the series: $S_y(z) = \sum_{i=-\infty}^{+\infty} R_y(i) z^{-i}$ — (*) to be well defined. For this we assume (Exponentially bounded)

(A1) — $|R_y(i)| < K\alpha^{|i|}$ for some $K > 0$ & $0 < \alpha < 1$.

Then ROC of (*) is $\alpha < |z| < \frac{1}{\alpha}$ (Note this includes unit circle $|z| = 1$)

(absolute convergence) $\boxed{\Rightarrow \text{Finite power (Ensure)}}$

Q) How do requirements (2) to (6) above translate to requirements on $L(z)$? How do we find $L(z)$?

Review std. signals & systems def^{n}s.
: Causal, Anticausal, Minimum phase, stable

<u>Answers</u>: Assume further: <u>$S_y(z)$ is rational.</u> (A2)

Under (A1) + (A2):
If $L(z)$ and $L^{-1}(z)$ both are BIBO stable
$\Leftrightarrow$ <u>$L(z)$ is BIBO stable & min. phase</u>

$\Rightarrow$ [$L(z)$ is z-transform of a causal seq. (lower triangular)
Exercise
$L^{-1}(z)$ is z-tr. of a causal seq (L^{-1} is lower tr.)
$S_y(e^{j\omega}) > 0$, $-\pi \leq \omega \leq \pi$ (Since no poles or zeros on unit circle)

$\hookrightarrow$ Note one directional implication: Assumption of BIBO stability of $L^{-1}(z)$ slightly stronger than req.

<u>Def^{n}</u>: (Canonical Spectral Factorization): Let $S_y(z)$ be a rational z-spectrum of a finite power process (and assume $S_y(e^{j\omega})$ is strictly positive) Then the canonical spectral factorization of $S_y(z)$ is

$\equiv$ $L^{-1}(z)$ BIBO stable

$$S_y(z) = L(z) r_e L^{*}\left(\frac{1}{z^{*}}\right) \left(= L(z) r_e L\left(\frac{1}{z}\right)\right)$$

where $L(z)$ is stable (BIBO) & minimum phase
$r_e > 0$ & $L(\infty) = 1$.

$\hookrightarrow$ Normalizing condition
\# guarantees uniqueness (How? Exercise) $\leftarrow$ [equivalent to diagonal elements being 1 for finite horizon

$$y_i \rightarrow \boxed{L^{-1}(z)} \xrightarrow[\text{white noise}]{e_i} \boxed{L(z)} \rightarrow y_i$$

Whitening Filter   Modelling filter

$E x_1 x_2^* = R_1$

$E x_2 x_1^* = R_1^*$

Scalar Rational Z-Spectra

Clearly: $S_y(z) = \sum_{i=-\infty}^{\infty} R_y(i) z^{-i} = S_y(1/z)$ $\;= S_y^*\left(\frac{1}{z^*}\right)$

$\Leftrightarrow$ For every pole (or zero) $z = \alpha$, $\exists$ a pole (resp. zero) at $z = \frac{1}{\alpha^*}$

So
$$S_y(z) = r_e \frac{\prod_{i=1}^{m} (z - \alpha_i)\left(\frac{1}{z} - \alpha_i^*\right)}{\prod_{i=1}^{n} (z - \beta_i)\left(\frac{1}{z} - \beta_i^*\right)}$$

with $|\alpha_i| < 1$, $|\beta_i| < 1$, $r_e > 0$

So canonical factorization is:

$$L(z) = z^{n-m} \frac{\prod_{i=1}^{m} (z - \alpha_i)}{\prod_{i=1}^{n} (z - \beta_i)} \qquad |\alpha_i| < 1, \; |\beta_i| < 1$$

(z^{n-m}: required to satisfy $L(\infty) = 1$)

Example: $S_y(z) = \sum_{i=-\infty}^{+\infty} a^{|i|} z^{-i} = \frac{1 - a^2}{(1 - a z^{-1})(1 - a z)}$

$0 < a < 1$

Then $L(z) = \frac{1}{1 - a z^{-1}}$

A useful result from Signals & Systems:

FACT: Filtering of Stationary Processes:

$$u_i \longrightarrow \boxed{H(z)} \longrightarrow y_i \longrightarrow \text{stationary}$$

zero mean stationary

LTI

$$S_y(z) = H(z)\, S_u(z)\, H^*\left(\frac{1}{z^*}\right) \qquad S_y(e^{j\omega}) = H(e^{j\omega})\, S_u(e^{j\omega})\, H^*(e^{j\omega})$$

$$S_{yu}(z) = H(z) \qquad S_{yu}(e^{j\omega}) = H(e^{j\omega})\, S_u(e^{j\omega})$$

Example: ARMA: $y_{i+1} = a_0 y_i + a_1 y_{i-1} + u_i + b u_{i-1}$, $i > -\infty$

$$H(z) = \frac{Y(z)}{U(z)} = \frac{z+b}{z^2 - a_0 z - a_1}, \quad |b| < 1$$

1) u_i is stationary white noise $\langle u_i, u_j \rangle = Q\delta_{ij}$

2) $\langle u_i, y_j \rangle = 0$ for $j \leq i$

3) stable

$$S_y(z) = H(z) \cdot Q \cdot H^*\left(\frac{1}{z^*}\right) = H(z)\, Q\, H\left(\frac{1}{z}\right)$$

$$= Q\left(\frac{z+b}{z^2 - a_0 z - a_1}\right)\left(\frac{z^{-1}+b}{z^{-2} - a_0 z^{-1} - a_1}\right)$$

Then, $L(z) = \dfrac{z(z+b)}{(z^2 - a_0 z - a_1)}$ if $|b| < 1$, $r_e = Q$.

If $|b| > 1$, $L(z) = \dfrac{z\left(z + \frac{1}{b}\right)}{z^2 - a_0 z - a_1}$, $r_e = b^2 Q$.

Infinite Horizon Wiener Theory (Discrete time)

- $\{s_i\}, \{y_i\} \to$ zero mean jointly WSS $\to R_s(i), R_y(i), R_{sy}(i) \to$ known

Smoother: l.l.m.s.e. of s_i using $\{y_m\}_{m=-\infty}^{+\infty}$

$$\hat{s}_i = \sum_{m=-\infty}^{+\infty} w_{im}\, y_m$$

Orthogonality condition: $\left(s_i - \sum_{m=-\infty}^{+\infty} w_{im} y_m\right) \perp y_l \quad -\infty < l < \infty$

or $R_{sy}(i-l) = \sum_{m=-\infty}^{+\infty} w_{im} R_y(m-l) \xrightarrow{\forall l}$ infinite set of linear eqns.

Let $m - l = m'$ & $i - l = i'$

then $\underbrace{R_{sy}(i')}_{\text{ind. of } l} = \sum_{m'=-\infty}^{+\infty} w_{i'+l,\, m'+l} R_y(m'), \quad -\infty < l < \infty$

ind. of $l \Longrightarrow \underline{w_{i'+l,\, m'+l} = w_{i',m'} = k_{i'-m} \quad \forall l}$

Here, $R_{sy}(i) = \sum_{m=-\infty} k_{i-m} R_y(m)$ — (*)

& $\hat{s}_i = \sum_{m=-\infty}^{+\infty} k_{i-m} R_y(m)$

(*) can be solved by taking DTFT:

Thm: Given two discrete time zero mean jointly stationary random processes $\{s_i, y_i\}$ the llmse (smoother) of s_i give $\{y_i\}_{i=-\infty}^{+\infty}$ is the LTI filter:

$$K(e^{j\omega}) = \frac{S_{sy}(e^{j\omega})}{S_y(e^{j\omega})}$$

Proof: Exercise.

Discrete Time Wiener-Hopf Filter

$\hat{s}_{i|i} = \sum_{m=-\infty}^{i} w_{im} y_m$ (Determine $\{w_{im}\}$ s.t. $\|s_i - \hat{s}_{i|i}\|$ is minimized)

Using orthogonality: $\left(s_i - \sum_{m=-\infty}^{i} w_{im} y_m\right) \perp y_l$ for $-\infty < l \le i$

$$R_{sy}(i-l) = \sum_{m=-\infty}^{i} w_{im} R_y(m-l) \quad \text{for } -\infty < l \le i$$

$$\Leftrightarrow R_{sy}(i') = \sum_{m=-\infty}^{i'+l} w_{i'+l,m} R_y(m-l) \quad \text{for } i' \geq 0 \quad \Big| \; i' = i-l$$

$$\Leftrightarrow R_{sy}(i') = \sum_{m'=-\infty}^{i'} w_{i'+l,m'+l} R_y(m') \quad \text{for } i' \geq 0 \quad \Big| \; m' = m-l \qquad (*)$$

holds for all l, but LHS is ind. of l, so

$$w_{i'+l,m'+l} = w_{i'm'} = k_{i'-m'}$$

Here $(*) \Leftrightarrow R_{sy}(i) = \sum_{m=-\infty}^{i} k_{i-m} R_y(m) = \sum_{m=0}^{\infty} k_m R_y(i-m)$, $i \geq 0$

Equivalently:
$$\boxed{\begin{array}{l} R_{sy}(i) = \sum_{m=-\infty}^{+\infty} k_m R_y(i-m) \qquad i \geq 0 \\ \& \qquad k_m = 0 \qquad m < 0 \end{array}} \qquad (**)$$

Wiener-Hopf Equation

Note: $\hat{s}_{i/i} = \sum_{m=-\infty}^{i} k_{i-m} y_m = \sum_{m=0}^{\infty} k_m y_{i-m}$

$$y_i \longrightarrow \boxed{K(z)} \longrightarrow \hat{s}_{i/i} \qquad \Big| \; K(z) = \sum_{i=0}^{\infty} k_i z^{-i}$$

<u>Solution:</u>

Assume: $S_y(z) > 0$ on $|z| = 1$

Define $g_i := R_{sy}(i) - \sum_{m=0}^{\infty} k_m R_y(i-m)$, $-\infty < i < \infty$

By $(**)$, g_i is strictly anti-causal i.e. $g_i = 0$ for $i \geq 0$.

Take z-transform of g_i over $(-\infty, \infty)$.

$$G(z) = S_{sy}(z) - K(z) S_y(z)$$

Use the canonical factorization of the z-spectrum $S_y(z)$.

$$S_y(z) = L(z)\, r_e L^*\left(\frac{1}{z^*}\right) \quad \text{— (1)}$$

\# Recall, $L(z)$ is causal, BIBO stable, min. phase.

Also $L(\infty) = 1$.

\# Divide (1), by $r_e L^*\left(\frac{1}{z^*}\right)$

$$\frac{G(z)}{r_e L^*\left(\frac{1}{z^*}\right)} = \frac{S_{sy}(z)}{r_e L^*\left(\frac{1}{z^*}\right)} - K(z)L(z)$$

FACT: $G(z)\, r_e^{-1}\left[L^{-1}\left(\frac{1}{z^*}\right)\right]^*$ will have strictly anti-causal inverse transform.

Proof: (Outline): $L^{-1}(z)$ corr. to a lower triangular doubly infinite L. So $\left[L^{-1}\left(\frac{1}{z^*}\right)\right]^*$ corr to a upper triangular doubly infinite L.

OR: $L^{-1}(z) = 1 + \sum_{j=1}^{\infty} w_j z^{-j}$ $\quad \left[L^{-1}\left(\frac{1}{z^*}\right)\right]^* = 1 + \sum_{j=1}^{\infty} w_j^* z^j$

Either logic implies $\left[L^{-1}\left(\frac{1}{z^*}\right)\right]^*$ corr. to an anti-causal inv. transform.

$$\underbrace{G(z)}_{\text{strictly anticausal IT}}\; r_e \underbrace{\left[L^{-1}\left(\frac{1}{z^*}\right)\right]^*}_{\text{anti-causal IT}} \longrightarrow \text{strictly anti-causal IT}$$

→ Recall proof from Signals & system (Exercise)

$$\underbrace{\frac{G(z)}{r_e L^*\left(\frac{1}{z^*}\right)}}_{\text{strictly anti-causal}} = \underbrace{\frac{S_{sy}(z)}{r_e L^*\left(\frac{1}{z^*}\right)}}_{\text{mixed}} - \underbrace{K(z)L(z)}_{\text{Causal}}$$

So causal parts must match $\Rightarrow$

$$\left\{\frac{S_{sy}(z)}{r_e L^*\left(\frac{1}{z^*}\right)}\right\}_{\substack{\text{causal}\\\text{part}}} = K(z)L(z)$$

i.e. $$K(z) = \left\{\frac{S_{sy}(z)}{r_e L^*\left(\frac{1}{z^*}\right)}\right\}_{\substack{\text{causal}\\\text{part}}} L^{-1}(z)$$

<u>Thm:</u> (Wiener Filter) Consider two zero-mean jointly WSS scalar random processes $\{s_i\}$ & $\{y_i\}$ with known rational z-spectra & z-cross spectra $S_s(z)$, $S_y(z)$ & $S_{sy}(z)$ resp. Assume further, that $S_y(z)$ has no unit circle zeroes. Then the llms estimates of s_i given $\{y_m, -\infty < m \leq i\}$ is given by the filter $K(z) = \left\{\frac{S_{sy}(z)}{r_e L^*\left(\frac{1}{z^*}\right)}\right\}_+ \frac{1}{L(z)}$.

<u>FACT:</u> $K(z)$ is BIBO stable

Proof: Exercise.