OPTIMAL FILTERING

LECTURE 2

1. Geometric Interpretation for Random Variables
2. Wiener filtering, causal and continuous time
3. Some examples

Reading instructions: Kailath, Sect. 3.3, 6.A and 7.A

GEOMETRIC INTERPRETATION FOR RANDOM VARIABLES

View zero mean random variables, $X, Y, Z, \ldots$ as vectors in some abstract vector space (in fact a Hilbert space, i.e., complete, normed linear vector space).

Recall that a random variable X is defined as a function $X(\omega)$ on some probability sample space Ω with points ω. Each $\omega \in \Omega$ defines a coordinate axis and along this axis X will have component $X(\omega)$.

Two elements in this space have an inner (scalar) product

$$\langle X, Y \rangle = \mathrm{E}\{XY\}$$

INNER PRODUCT

Usual properties of inner product, such as

$$\langle X, Y \rangle = \mathrm{E}\{XY\}$$

1. Linearity $\langle \alpha X + \beta Y, Z \rangle = \alpha \langle X, Z \rangle + \beta \langle Y, Z \rangle$
2. Symmetry (reflexivity) $\langle X, Y \rangle = \langle Y, X \rangle^*$
3. Non-degeneracy $\langle X, X \rangle \triangleq \|X\|^2 \geq 0$.
   $\langle X, X \rangle = 0$ iff $X = 0$.

COMPLEX VECTOR RANDOM VARIABLES

Let the inner product be defined by

$$\langle X, Y \rangle = \mathrm{E}\{XY^*\}$$

where

$(\cdot)^*$	–	transpose of a *real* valued vector
$(\cdot)^*$	–	complex conjugate transpose (Hermitian transpose) for *complex* valued vectors

1. Sesquilinearity: $\langle \alpha X + \beta Y, Z \rangle = \alpha \langle X, Z \rangle + \beta \langle Y, Z \rangle$
   $\langle X, \alpha Y + \beta Z \rangle = \alpha^* \langle X, Y \rangle + \beta^* \langle X, Z \rangle$
2. Symmetry (reflexivity) $\langle X, Y \rangle = \langle Y, X \rangle^*$
3. Non-degeneracy $\langle X, X \rangle > 0$, i.e., positive definite,
   $\langle X, X \rangle = 0$, iff $X = 0$.

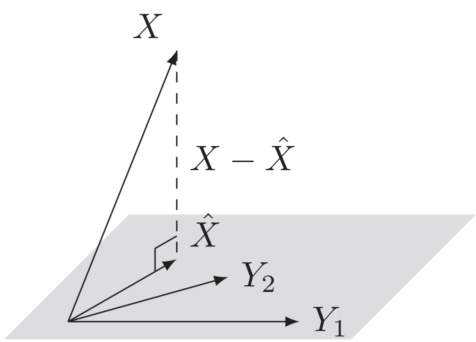
BASIC ORTHOGONALITY PROPERTY

Geometric framework of the LLSE problem:

> Find a vector $\hat{X}$ in the linear subspace spanned by $Y_1, \ldots, Y_N$ such that the squared length of the error vector $X - \hat{X}$ is as small as possible.

Geometrically, $X - \hat{X}$ is clearly smallest when $\hat{X}$ is the foot of the perpendicular from X to the linear subspace of the observations, i.e., the projection of X on to the subspace spanned by the observations $\{Y_i\}$.

BASIC ORTHOGONALITY PROPERTY



$$Y = \begin{bmatrix} Y_1 \\ \vdots \\ Y_N \end{bmatrix}$$

$$\hat{X} = hY$$

$$X - hY \perp Y_i$$

$$\implies \langle X - hY, Y \rangle = \mathrm{E}\{(X - hY)Y^*\} = 0$$

$$\Sigma_{xy} = h\Sigma_{yy} \quad \implies \quad h = \Sigma_{xy}\Sigma_{yy}^{-1}$$

$$\boxed{\hat{X} = \Sigma_{xy}\Sigma_{yy}^{-1}Y}$$

As before!

CONTINUOUS TIME RANDOM PROCESSES

The geometric interpretation is very helpful especially for *continuous* time random processes.

Consider forming a LLSE of X from observations of a random process

$$\{Y(\tau),\ a \leq \tau \leq b\}$$

The optimal estimator is uniquely defined by the orthogonality conditions,

$$\hat{X} = \int_a^b h(\tau)Y(\tau)d\tau \quad \text{linear estimate}$$

$$\left(X - \int_a^b h(\tau)Y(\tau)d\tau\right) \perp Y(t) \quad a \leq t \leq b \quad \text{orth. cond.}$$

CONTINUOUS TIME RANDOM PROCESSES

$$\Longrightarrow \ \mathrm{E}\{XY^*(t)\} = \int_a^b h(\tau)\mathrm{E}\{Y(\tau)Y^*(t)\}d\tau$$

$$\Longrightarrow \ \Sigma_{xy}(t) = \int_a^b h(\tau)\Sigma_{yy}(\tau, t)d\tau$$

Results in an integral equation.

This equation is *much* more difficult to obtain with other techniques.

WIENER FILTERS

Original problem considered by Norbert Wiener (1894 – 1964):

Prediction of airplane trajectories for the control of anti-aircraft guns.

Let $x(\cdot)$ and $y(\cdot)$ be two zero mean jointly wide sense stationary (wss) random processes with known covariance functions, $r_y(t)$ and $r_{xy}(t)$.

Given observations $\{y(\tau), \infty < \tau \leq t\}$ find the LLSE of $x(t+\lambda)$.

$x(t)$ – position at time t and $y(t)$ may be noisy range and velocity measurements at time t.

$$\hat{x}(t+\lambda) = \int_{-\infty}^{t} h(t,\tau)y(\tau)d\tau \quad (1)$$

minimize

$$\mathrm{E}\{(x(t+\lambda) - \hat{x}(t+\lambda))^2\} \quad \text{(Scalar case for simplicity)}$$

CAUSAL WIENER FILTERING

Orthogonality condition

$$(x(t+\lambda) - \hat{x}(t+\lambda)) \perp y(\sigma) \quad -\infty < \sigma \leq t$$

which yields

$$\mathrm{E}\{x(t+\lambda)y(\sigma)\} = \int_{-\infty}^{t} h(t,\tau)\mathrm{E}\{y(\tau)y(\sigma)\}d\tau \quad -\infty < \sigma \leq t$$

$$\Longrightarrow \ldots \Longrightarrow$$

$h(t,\tau)$ of the form $h(t,\tau) = h(t-\tau)$ and

$$r_{xy}(t+\lambda) = \int_{0}^{\infty} h(\tau)r_y(t-\tau)d\tau \qquad t \geq 0$$

CAUSAL WIENER FILTERING

Let $h(t)$ be causal, $h(t) = 0$, $t < 0$.

$$\hat{x}(t+\lambda) = \int_{-\infty}^{\infty} h(t-\tau)y(\tau)d\tau = h(t) * y(t) \qquad (1')$$

Where $h(t)$ fulfills the, so-called, *Wiener-Hopf equation*

$$r_{xy}(t+\lambda) = \int_{-\infty}^{\infty} h(\tau)r_y(t-\tau)d\tau \quad t \geq 0$$

Looks like it is solvable by Laplace Transforms but this *does not work* because of the constraint $t \geq 0$!

SPECIAL CASES

Example 1. *White noise*

Let $y(\cdot)$ be white noise, i.e., $r_y(t) = \delta(t)$. We have

$$r_{xy}(t+\lambda) = \int_{-\infty}^{\infty} h(\tau)r_y(t-\tau)d\tau = h(t) \quad t \geq 0$$

If $x(t+\lambda) = y(t+\lambda)$, i.e., a prediction problem we have

$$r_{xy}(t+\lambda) = r_y(t+\lambda) = \delta(t+\lambda)$$

we get

$$h(t) = \delta(t+\lambda) \quad t \geq 0$$

But $\lambda > 0 \quad \Longrightarrow \quad h(t) = 0$.

White noise is of course predicted by its mean, $\hat{y}(t+\lambda) = 0$. □

SPECIAL CASES

Example 2. *Smoothing (or non-causal Wiener filter)*

Observe $y(\tau)$, $-\infty < \tau < \infty$ and estimate $x(t)$. This gives

$$r_{xy}(t) = \int_{-\infty}^{\infty} h(\tau) r_y(t-\tau) d\tau$$

$\mathcal{L}$-transform gives

$$S_{xy}(s) = H(s) S_y(s) \quad \Longrightarrow \quad h(t) = \mathcal{L}^{-1}\left\{\frac{S_{xy}(s)}{S_y(s)}\right\}$$

$$\hat{x}(t) = \int_{-\infty}^{\infty} h(t-\tau) y(\tau) d\tau$$

Not realizable as a causal filter but it may be approximated by truncating the impulse response. □

SPECTRAL FACTORIZATION

In order to proceed with the solution to the Wiener-Hopf equation we need to introduce the concept of *Spectral Factorization*.

Time continuous

Let $S_y(s)$ be a spectral density ($S_y(i\omega) > 0$) which is rational in s, then $S_y(s)$ may be factored

$$S_y(s) = \underbrace{H(s)}_{S_y^+(s)} \underbrace{H^*(-s^*)}_{S_y^-(s)}$$

where $H(s)$ is stable and minimum phase (i.e. zeros and poles strictly in the LHP).

SPECTRAL FACTORIZATION, CONT.

Thus, $S_y(s)$ can be written as

$$S_y(s) = R\frac{\Pi_{i=1}^{m}(s-\beta_i)(-s-\beta_i^*)}{\Pi_{i=1}^{n}(s-\alpha_i)(-s-\alpha_i^*)} \quad \text{real}(\beta_i), \text{real}(\alpha_i) < 0, \quad R > 0$$

The spectral factor is

$$S_y^+(s) = R^{1/2}\frac{\Pi_{i=1}^{m}(s-\beta_i)}{\Pi_{i=1}^{n}(s-\alpha_i)}$$

Remarks: In the text book *Linear Estimation*,

$S_y(s) = L(s)RL^*(-s^*)$, $R > 0$, where $L(\infty) = 1$.

When $L(s)$ is the transfer function of a *real* system, $L^*(-s^*) = L(-s)$.

ADDITIVE DECOMPOSITION

$$H(s) = \mathcal{L}\{h(t)\} = \int_{-\infty}^{\infty} h(t)e^{-st}dt$$

$$= \overbrace{\underbrace{\int_{-\infty}^{0^-} h(t)e^{-st}dt}_{\{H(s)\}_-}}^{\text{anti-causal}} + \overbrace{\underbrace{\int_{0^-}^{\infty} h(t)e^{-st}dt}_{\{H(s)\}_+}}^{\text{causal}}$$

When $H(s)$ is rational in s, it can be shown that

Causal part $\{H(s)\}_+$ has LHP poles

Anti-causal part $\{H(s)\}_-$ has RHP poles

CAUSAL WIENER FILTER

Back to the Wiener solution (of the Wiener-Hopf equation).

Given

$$r_{xy}(t+\lambda) = \int_0^\infty h(\tau) r_y(t-\tau) d\tau \quad t \geq 0$$

$$h(t) = 0 \quad t < 0$$

$$S_{xy}(s) = \int_{-\infty}^{\infty} r_{xy}(t) e^{-st} dt \qquad S_y(s) = \int_{-\infty}^{\infty} r_y(t) e^{-st} dt$$

$$S_{xy}(s) e^{s\lambda} = \int_{-\infty}^{\infty} r_{xy}(t) e^{-s(t-\lambda)} dt = \int_{-\infty}^{\infty} r_{xy}(t+\lambda) e^{-st} dt$$

Then

$$H(s) = \frac{1}{S_y^+(s)} \left\{ \frac{S_{xy}(s) e^{s\lambda}}{S_y^-(s)} \right\}_+$$

CAUSAL WIENER FILTER, PROOF

Let

$$g(t) = r_{xy}(t+\lambda) - \int_0^\infty h(\tau) r_y(t-\tau) d\tau \quad t \geq 0$$

then

$$g(t) = \begin{cases} 0 & t \geq 0 \\ ? & t < 0 \end{cases}$$

Take the Laplace transform

$$\underbrace{G(s)}_{\substack{\text{anti-causal since} \\ g(t) = 0,\ t \geq 0}} = S_{xy}(s) e^{s\lambda} - H(s) S_y(s)$$

CAUSAL WIENER FILTER, PROOF

Factorize

$$S_y(s) = S_y^+(s) \underbrace{S_y^-(s)}_{\text{no zeros in LHP}}$$

$$\underbrace{\frac{G(s)}{S_y^-(s)}}_{\substack{\text{Anticausal!}\\ \text{Anti-causal “signal”}\\ \text{filtered through}\\ \text{anti-causal filter}}} = \frac{S_{xy}(s)e^{s\lambda}}{S_y^-(s)} - \underbrace{H(s)S_y^+(s)}_{\substack{\text{Causal!}\\ \text{Causal “signal” filtered}\\ \text{through causal filter}}}$$

CAUSAL WIENER FILTER, PROOF

$$H(s)S_y^+(s) = \left\{ \frac{S_{xy}(s)e^{s\lambda}}{S_y^-(s)} \right\}_+$$

Causal Wiener filter solution for continuous time

$$\boxed{H(s) = \frac{1}{S_y^+(s)} \left\{ \frac{S_{xy}(s)e^{s\lambda}}{S_y^-(s)} \right\}_+}$$