Vector LMMSE estimator

Like MMSE scalar to vector was a trivial extension so is the LMMSE scalar to vector extension

$$\hat{\theta}_i = \sum_{n=0}^{N-1} a_{in}\, x[n] + a_{iN} \quad \text{, choose } a \text{ to minimize}$$

$$Bmse(\hat{\theta}_i) = E\left[(\theta_i - \hat{\theta}_i)^2\right]$$

$$\hat{\theta}_i = E[\theta_i] + C_{\theta_i x} C_{xx}^{-1}(x - E(x)) \quad i = 1 \text{ to } p$$

$$Bmse(\hat{\theta}_i) = C_{\theta_i\theta_i} - C_{\theta_i x} C_{xx}^{-1} C_{x\theta_i}$$

<u>Vector:</u> $$\underline{\hat{\theta}} = \begin{bmatrix} E(\theta_1) \\ \vdots \\ E(\theta_p) \end{bmatrix} + \begin{bmatrix} C_{\theta_1 x} \\ \vdots \\ C_{\theta_p x} \end{bmatrix} C_{xx}^{-1}(x - E(x))$$

$$\underline{\hat{\theta}} = E(\underline{\theta}) + C_{\theta x} C_{xx}^{-1}(x - E(x))$$

Let $$M_{\hat{\theta}} = E\left[(\underline{\theta} - \underline{\hat{\theta}})(\underline{\theta} - \underline{\hat{\theta}})^T\right]$$
$$= C_{\theta\theta} - C_{\theta x} C_{xx}^{-1} C_{x\theta}$$ (no longer scalar quantity)

Then $$Bmse(\theta_i) = \left[M_{\hat{\theta}}\right]_{ii}$$ → i'th diagonal element

<u>Note</u>: Gauss-Markov theorem specifies the vector LMMSE estimator for the Bayesian Linear model $x = H\theta + w$ without knowing the PDF for w.

$$\hat{\theta} = E(\theta) + \left(C_\theta^{-1} + H^T C_w^{-1} H\right)^{-1} H^T C_w^{-1}(x - HE(x))$$

Same as MMSE estimator when θ, w jointly Gaussian.

Wiener Filtering

$x[n] = s[n] + w[n]$

$\theta = \begin{pmatrix} s[0] \\ \vdots \\ s[N-1] \end{pmatrix}$ $s[n]$ 0-mean

Sometimes referred to as Wiener smoothing

Classical estimate was useless $\hat{\theta} = x$ $\left(\hat{s}[n] = x[n]\right.$

Lets assume $x[0], \dots x[N-1]$ is a wide-sense stationary signal, i.e. mean & variance independent of n. Then C_{xx} takes the symmetric Toeplitz form

$$C_{xx} = \begin{bmatrix} r_{xx}[0] & r_{xx}[1] & r_{xx}[2] & \cdots & r_{xx}[N-1] \\ r_{xx}[1] & r_{xx}[0] & r_{xx}[1] & & \vdots \\ r_{xx}[2] & r_{xx}[1] & r_{xx}[0] & & \vdots \\ \vdots & & & & \\ r_{xx}[N-1] & - & - & - & r_{xx}[0] \end{bmatrix}$$

$$r_{xx}[k] = \frac{1}{N} \sum_{n=0}^{N-1} x[n]\, x[n+k] \quad \text{auto-correlation}$$

If we assume $\qquad = E\left[x[n]\, x[n+k]\right]$

that noise w is independent of s then

$$r_{xx}[k] = r_{ss}[k] + r_{ww}[k]$$

$$\Rightarrow C_{xx} = C_{ss} + C_{ww}$$

$$C_{\theta x} = \underset{\substack{\text{using 0-mean} \\ \text{for } s \text{ and } w}}{E\left[s x^T\right]} = C_{ss}$$

Therefore $\hat{s} = \underbrace{C_{ss}\left(C_{ss} + C_{ww}\right)^{-1}}_{\text{Wiener smoothing matrix}} x$

Notice $\hat{s}$ is not simply x.

But we need to know C_{ss}.

Example: Autoregressive random process

$$s[n] = \sum_{i=1}^{p} a_i s[n-i] + \varepsilon[n]$$

a_i known coefficients ε WGN with known σ_ε^2

Observed $x[n] = s[n] + w[n]$

w WGN with known σ_w^2

Estimate $s[n]$ $n = 0, \ldots N-1$ from $x[n]$ $n=0 \ldots N$

use Wiener filter.

For simplicity take an AR-1 process

$$s[n] = a s[n-1] + \varepsilon[n]$$
$$x[n] = s[n] + w[n]$$

$C_{ww} = \sigma_w^2 I$, what is C_{ss}?

For a AR-1 process, we can show that the autocorrelation function is $r_{ss}[k] = \dfrac{\sigma_\varepsilon^2}{1-a^2} a^{|k|}$

This can be used to construct C_{ss}

Then $\hat{s} = C_{ss}\left(C_{ss} + C_{ww}\right)^{-1} \underline{x}$

$$= C_{ss}\left(C_{ss} + \sigma_w^2 I\right)^{-1} \underline{x}$$

MATLAB example: $a = 0.95$, $\sigma_\varepsilon^2 = 1.0$, $\sigma_w^2 = 1.5$

```
[s x] = ARprocess(0.95, 1.0, 1.5, 100)
ŝ = WienerARprocess(x, 0.95, 1.0, 1.5)
```

Wiener Filtering for images

Image degradation model: $F(x,y) = h(x,y) * s(x,y) + n(x,y)$

$s(x,y)$ = true signal, $h(x,y)$ = point spread function of degradation (blur for instance), $n(x,y)$ = WGN.

In the frequency domain: $F(u,v) = \underbrace{H(u,v)}_{\text{known}} S(u,v) + N(u,v)$

We want to find a LMMSE estimator such that

$J = E\left(\| \hat{S} - S \|^2 \right)$ is minimum (in freq-domain)

Since the estimator is linear it can be written as a convolution in time-domain: $\hat{S} = G F$ in freq. domain

Wiener filter

$$J = E\left[|S - GF|^2 \right]$$

$$= E\left[|S - G(HS + N)|^2 \right]$$

$$= E\left[|\underbrace{(1 - GH)}_{\text{call this } Q} S - GN|^2 \right]$$

$$= E\left[(QS - GN)(QS - GN)^* \right]$$

complex conjugate since we are in Fourier-domain

$$= E\left[QQ^* SS^* \right] \rightarrow QQ^* \underbrace{E[SS^*]}_{P_S}$$

P_S: power spect. density for S

$$- E\left[QSG^*N^* \right] \rightarrow QG^* E[SN^*]$$

$$- E\left[GNQ^*S^* \right]$$

0 since S & N independent

$$+ E\left[GNG^*N^* \right] \rightarrow GG^* \underbrace{E[NN^*]}_{P_N}$$

$$= QQ^* P_S + GG^* P_N$$

$$= (1 - GH)(1 - GH)^* P_S + GG^* P_N$$

$$\frac{\partial J}{\partial G} = -H(1-GH)^* P_S + G^* P_N = 0$$

$$-HP_S + HH^* G^* P_S + G^* P_N = 0$$

$$G^* = \frac{HP_S}{HH^* P_S + P_N} = \frac{HP_S}{|H|^2 P_S + P_N}$$

$$G = \frac{H^* P_S}{|H^2| P_S + P_N}$$

→ real #

so $\hat{S}(u,v) = \frac{H^* P_S}{|H^2| P_S + P_N} F(u,v)$

$\hat{S}(x,y) =$ inverse F.T $\left(\hat{S}(u,v) \right)$

rewriting

$$\hat{S}(u,v) = \left(\frac{1 \times |H|^2}{H \times (|H^2| + P_N/P_S)} \right) F(u,v)$$

H: known degradation function

P_N/P_S ratio of noise to signal powerspectrum (function of u,v)

usually we take $P_N/P_S = K$ a constant

$$\hat{S}(u,v) = \frac{|H|^2 F(u,v)}{H(|H|^2 + K)}$$, K acts as a parameter