ECE 7251: Signal Detection and Estimation

Spring 2002

Prof. Aaron Lanterman

Georgia Institute of Technology

Lecture 18, 2/18/02:

Wiener Filtering

The Setup

- Context: Bayesian linear MMSE estimation for random sequences
- Parameter sequence $\{\Theta_k, k \in \mathbb{Z}\}$
- Data sequence $\{Y_k, k \in I \subset \mathbb{Z}\}$
- Goal: Estimate $\{q_k\}$ as a linear function of the observations:

$$\hat{q}_k(y) = \sum_{j \in I} h(k, j) y_j$$

- Find *h* to minimize mean square error

O.P. to the Rescue

- By the orthogonality principle,

$$E[(\hat{q}_k(Y) - \Theta_k)Y_i^*] = 0 \text{ for } i \in I$$

$$E[(\sum_{j \in I} h(k, j)Y_j - \Theta_k)Y_i^*] = 0$$

$$\sum_{j \in I} h(k, j)E[Y_j Y_i^*] = E[\Theta_k Y_i^*]$$

$$\sum_{j \in I} h(k, j) r_Y(j, i) = r_{\Theta Y}(k, i)$$

Wiener-Hopf Equation

- If processes are stationary, we can write

$$\sum_{j \in I} h(k, j) r_Y \underbrace{(j - i)}_{\text{slight abuse of notation}} = r_{\Theta Y}(k - i)$$

Spectral Representation

- If $I = \mathbb{Z}$, it turns out the filter is LTI:

$$\sum_{j \in I} h(k - j) r_Y(j - i) = r_{\Theta Y}(k - i)$$

WLOG, consider *i*=0: $\sum_{j \in I} h(k - j) r_Y(j) = r_{\Theta Y}(k)$

- Can solve W-H in the Z-transform domain:

$$H(z) S_Y(z) = S_{\Theta Y}(z)$$

$$H(z) = \frac{S_{\Theta Y}(z)}{S_Y(z)}$$

Mean Square Error

- $MSE = E[|(\Theta_k - \hat{q}(Y))|^2]$

$$= E[(\Theta_k - \hat{q}(Y))(\Theta_k^* - \hat{q}^*(Y))]$$

$$= E[(\Theta_k - \hat{q}(Y))\Theta_k^*] + E[(\Theta_k - \hat{q}(Y))\hat{q}^*(Y)]$$

(by O.P.) [second term crossed out]

$$= E[\Theta_k \Theta_k^*] - E[\hat{q}_k(Y)\Theta_k^*]]$$

$$= E[\Theta_k \Theta_k^*] - E[\sum_{j \in I} h(k - j) Y_j \Theta_k^*]]$$

$$= E[\Theta_k \Theta_k^*] - \sum_{j \in I} h(k - j) E[Y_j \Theta_k^*]$$

$$= r_\Theta(0) - \sum_{j \in I} h(k - j) r_{Y\Theta}(j - k)$$

Mean Square Error

$$MSE = r_\Theta(0) - \sum_{j \in I} h(k - j) r_{Y\Theta}(j - k)$$

- Since everything is stationary, can just take *k=0*

$$MSE = r_\Theta(0) - (h * r_{Y\Theta})(0)$$

$$= \int_{-p}^{p} S_\Theta(w) - H(w) S_{Y\Theta}(w) dw$$

$$= \int_{-p}^{p} S_\Theta(w) - \frac{S_{\Theta Y}(w)}{S_Y(w)} S_{Y\Theta}(w) dw$$

$$= \int_{-p}^{p} S_\Theta(w) - \frac{|S_{\Theta Y}(w)|^2}{S_Y(w)} dw$$

Example: Deblurring

- Suppose object is observed through a blurring point spread function *f* and additive noise *W*

$$Y_k = (f * \Theta)_k + W_k$$

- Suppose Θ and *W* are uncorrelated zero-mean
- Recall from ECE6601:

$$S_Y = |F|^2 S_q + S_W \text{ and } S_{\Theta Y} = F^* S_Y$$

- So the Wiener filter is

$$H(z) = \frac{S_{\Theta Y}(z)}{S_Y(z)} = \frac{F^*(z) S_\Theta(z)}{|F(z)|^2 S_q(z) + S_W(z)}$$

Interpretation of Deblurring Filter

- If noise is negligible, i.e. $S_W(w) \approx 0$

$$H(w) = \frac{F^* S_q}{|F|^2 S_q + S_W} \approx \frac{F^* S_q}{F F^* S_q} = \frac{1}{F}$$

- Even if there is no noise, in implementation, straight division by $F(\omega)$ is often ill-posed and not a good idea (round off errors, etc.)

Deblurring Error

$$MSE = \int_{-p}^{p} S_\Theta(w) - \frac{|S_{\Theta Y}(w)|^2}{S_Y(w)} dw$$

$$= \int_{-p}^{p} S_\Theta(w) - \frac{|F(w)|^2 |S_\Theta(w)|^2}{|F(w)|^2 S_\Theta(w) + S_W(w)} dw$$

$$= \int_{-p}^{p} \frac{S_\Theta[|F|^2 S_\Theta + S_W] - |F|^2 |S_\Theta|^2}{|F|^2 S_\Theta + S_W} dw$$

$$= \int_{-p}^{p} \frac{S_\Theta S_W}{|F|^2 S_\Theta + S_W} dw$$

Discussion on Deblurring

- Advantage of Wiener approach:
  - LTI filtering implementation
- Disadvantages of Wiener approach:
  - No natural way to incorporate nonnegativity constraints (in image processing, for instance)
  - Only truly optimal for Gaussian statistics
- Competing approaches include iterative methods such as the “Richardson-Lucy” algorithm (an EM-style procedure)
  - Computationally intensive
  - Can naturally incorporate nonnegativity
  - Sometimes better match to real statistics

Real-Time Wiener Filtering

- What if we don’t have “future” measurements?
  - Must restrict *h* to be causal
- Solution:

$$H(z) = \frac{1}{S_Y^-(z)} \left\{ \frac{S_{\Theta Y}(z)}{S_Y^+(z)} \right\}_+$$

where the meaing of the plus and minus superscripts and subscripts will be defined on later slides

Spectral Factorization

- If *Y* has a spectrum satisfying the Paley-Wiener criterion:

$$\int_{-p}^{p} \log S_Y(w) dw > -\infty$$

then the spectrum can be factored as

$$S_Y(w) = S_Y^+(w) S_Y^-(w)$$

where

$\mathcal{F}^{-1}\{S_Y^+\}$ is causal

$\mathcal{F}^{-1}\{S_Y^-\}$ is anticausal

Factoring Rational Spectra

- If the spectrum is a ratio of polynomials, we can factor as

$$S_Y(z) = \underbrace{S_Y^+(z)}_{\text{Poles and zeros inside unit circle}} \underbrace{S_Y^-(z)}_{\text{Poles and zeros outside unit circle}} = S_Y^+(z) S_Y^+(z^{-1})$$

- Aside: spectral factorization into causal and anticausal factors is analogous to Cholesky decomposition of a covariance matrix into lower and upper triangular factors

Causal Part Extraction

- We can split f into its causal and anticausal parts:

$$f(k) = \underbrace{\{f(k)\}_+}_{\text{causal}} + \underbrace{\{f(k)\}_-}_{\text{anticausal}}$$

$$f(k)_+ = f(k)u(k), \ \{f(k)\}_- = f(k)u(-k-1)$$

- Use similar notation for Z-transform domain

$$F(z) = \{F(z)\}_+ + \{F(z)\}_-$$

$$\{F\}_+ = \mathcal{Z}\{\mathcal{Z}^{-1}\{F\}u(k)\}$$

$$\{F\}_- = \mathcal{Z}\{\mathcal{Z}^{-1}\{F\}u(-k-1)\}$$

How to Extract Causal Parts

- If F is a ratio of polynomials can usually do a partial fraction expansion:

$$F(z) = \underbrace{\{F(z)\}_+}_{\text{Poles and zeros inside unit circle}} + \underbrace{\{F(z)\}_-}_{\text{Poles and zeros outside unit circle}}$$

- Can also do polynomial long division (see Ed Kamen's book)
- Almost always a total pain and really annoying