Data Science

TD 4

1. Optimal Filtering and the Wiener Filter

A geophysicist measures a seismic signal

$$y(t) = x(t) + n(t)$$

where:

- $x(t)$ is the stationary seismic signal of interest. It is centered ($\mu_x = 0$) with autocorrelation function $\gamma_x(\tau) = \sigma_x^2\, e^{-\alpha|\tau|}$
- $n(t)$ is stationary measurement noise, it is centered and independent of $x(t)$
- the signal and the noise are uncorrelated: $\mathbb{E}\{x(t)\, n(t')\} = 0$

The goal is to design an **optimal linear filter** $h(t)$ that estimates $x(t)$ from the noisy measurements $y(t)$. We seek an estimate $x^+(t)$ that minimizes the **mean square error (MSE)**:

$$\text{MSE} = \int_{\mathbb{R}} \mathbb{E}\left\{|x(t) - x^+(t)|^2\right\} \mathrm{d}t$$

1.1. Computing power spectral densities

Q. 1 The noise is assumed independent, white, and identically distributed with variance σ_n^2.

1-a) Write its autocorrelation function.

> **Solution:**
>
> $\gamma_{n(\tau)} = \sigma_n^2\, \delta(\tau)$

1-b) Compute the autocorrelation $\gamma_{y(\tau)}$ of the measured signal y.

> **Solution:**
>
> Since $x(t)$ and $n(t)$ are uncorrelated, we have:
>
> $$\gamma_{y(\tau)} = \gamma_{x(\tau)} + \gamma_n(\tau) = \sigma_x^2\, e^{-\alpha|\tau|} + \sigma_n^2\, \delta(\tau)$$

1-c) Compute the cross-correlation $\gamma_{x,y}(\tau) = \int_{\mathbb{R}} E\{x(t)\, y(t+\tau)\}\, \mathrm{d}t$ between the true signal and the measurements.

> **Solution:**
>
> Since $x(t)$ and $n(t)$ are uncorrelated, we have:
>
> $$\gamma_{x,y}(\tau) = \gamma_x(\tau) + 0 = \sigma_x^2\, e^{-\alpha|\tau|}$$

Q. 2 Compute the power spectral densities of:

- the true signal $S_x(\nu)$
- the noise $S_n(\nu)$
- the measurements $S_y(\nu)$
- the cross power spectral density between x and y: $S_{x,y}(\nu)$

Solution:

Using the Fourier transform of the autocorrelation functions, we have:

- $$S_x(\nu) = \mathcal{F}\{\gamma_x\} = \int_{\mathbb{R}} \sigma_x^2\, e^{-\alpha|\tau|}\, e^{-j2\pi\nu\tau}\,\mathrm{d}\tau = \sigma_x^2 \frac{2\alpha}{\alpha^2 + (2\pi\nu)^2}$$
- $$S_n(\nu) = \mathcal{F}\{\gamma_n\} = \int_{\mathbb{R}} \sigma_n^2\, \delta(\tau)\, e^{-j2\pi\nu\tau}\,\mathrm{d}\tau = \sigma_n^2$$
- $$S_y(\nu) = S_x(\nu) + S_n(\nu) = \sigma_x^2 \frac{2\alpha}{\alpha^2 + (2\pi\nu)^2} + \sigma_n^2$$
- $$S_{x,y}(\nu) = S_x(\nu) = \sigma_x^2 \frac{2\alpha}{\alpha^2 + (2\pi\nu)^2}$$

1.2. Derivation of the Wiener filter

The Wiener filter is the filter with transfer function $\hat{h}$ that minimizes the Mean Square Error. From Parseval-Plancherel we can write:

$$\mathrm{MSE} = \int_{\mathbb{R}} \mathbb{E}\left\{\left|\hat{x}(\nu) - \hat{h}(\nu)\,\hat{y}(\nu)\right|^2\right\}\mathrm{d}\nu$$

Q. 3 3-a) Show that

$$\mathrm{MSE} = \int_{\mathbb{R}} S_x(\nu) - 2\,\mathfrak{R}\left(\hat{h}(\nu)\,S_{x,y}(\nu)\right) + \left|\hat{h}(\nu)\right|^2 S_y(\nu)\,\mathrm{d}\nu$$

Solution:

By expanding the square and using the definition of spectral density, we have:

$$\mathrm{MSE} = \int_{\mathbb{R}} \mathbb{E}\{|\hat{x}(\nu)|^2\} - 2\,\mathfrak{R}\left(\mathbb{E}\left\{\hat{x}(\nu)\,\overline{\hat{h}(\nu)\,\hat{y}(\nu)}\right\}\right) + \mathbb{E}\left\{\left|\hat{h}(\nu)\,\hat{y}(\nu)\right|^2\right\}\mathrm{d}\nu$$

Using the definition of spectral density, we obtain:

$$\mathrm{MSE} = \int_{\mathbb{R}} S_x(\nu) - 2\,\mathfrak{R}\left(\hat{h}(\nu)\,S_{x,y}(\nu)\right) + \left|\hat{h}(\nu)\right|^2 S_y(\nu)\,\mathrm{d}\nu$$

3-b) To minimize the MSE, take the derivative with respect to $\hat{h}(\nu)$ and set it to zero. Show that the optimal filter (Wiener filter) satisfies:

$$\hat{h}(\nu) = \frac{S_x(\nu)}{S_x(\nu) + S_n(\nu)}, \quad \forall \nu \in \mathbb{R}$$

Solution:

Taking the derivative of the MSE with respect to $\hat{h}(\nu)$ and setting it to zero, we have:

$$\frac{\mathrm{d}\ \mathrm{MSE}}{\mathrm{d}\hat{h}(\nu)} = \int_{\mathbb{R}} -2\, S_{x,y}(\nu) + 2\hat{h}(\nu)\, S_y(\nu)\, \mathrm{d}\nu = 0$$

Solving for $\hat{h}(\nu)$, we obtain:

$$\hat{h}(\nu) = \frac{S_{x,y}(\nu)}{S_y(\nu)}$$

Replacing $S_{x,y}(\nu)$ and $S_y(\nu)$ by their expressions found in the previous question, we get:

$$\hat{h}(\nu) = \frac{S_x(\nu)}{S_x(\nu) + S_n(\nu)}$$

1.3. Application

Q. 4 4-a) Write the Wiener filter expression using the densities computed in question 2:

Solution:

Replacing the expressions of $S_x(\nu)$ and $S_n(\nu)$ found in question 2, we have:

$$\begin{aligned}\hat{h}(\nu) &= \frac{\sigma_x^2 \frac{2\alpha}{\alpha^2+(2\pi\nu)^2}}{\sigma_x^2 \frac{2\alpha}{\alpha^2+(2\pi\nu)^2} + \sigma_n^2} \\ &= \frac{\sigma_x^2\,(2\alpha)}{\sigma_x^2\,(2\alpha) + \sigma_n^2\,(\alpha^2 + (2\pi\nu)^2)} \\ &= \frac{1}{1 + \frac{\sigma_n^2}{\sigma_x^2}\frac{\alpha^2+(2\pi\nu)^2}{2\alpha}}\end{aligned}$$

4-b) Analyze the behavior of the Wiener filter:
- What happens at low frequencies?
- What happens at high frequencies?
- How does the filter depend on the signal-to-noise ratio: $\mathrm{SNR} = \frac{\sigma_x^2}{\sigma_n^2}$?

4-c) Sketch (qualitatively) the magnitude $\left|\hat{h}(\nu)\right|^2$

2. Spatial Interpolation and Kriging

A geophysicist has measured a physical property (e.g., porosity, seismic velocity) at N discrete locations $\boldsymbol{x} = [x_1, x_2, ..., x_N]^T$ in a geological formation. The measurements are $Z(x_i) = z_i$ (*i.e.* the vector $\boldsymbol{z} = [z_1, z_2, ..., z_N]^T$).

The goal is to estimate the value $\tilde{Z}(x_0)$ at an unmeasured location x_0 using a linear combination of the observations:

$$\tilde{Z}(x_0) = \sum_{i=1}^{N} \lambda_i Z(x_i)$$

2.1. Expressing the covariance structure

We model $Z(x)$ as a stationary random field (spatial random signal) with zero mean $\mathbb{E}\{Z(x)\} = 0$ and covariance function

$$\gamma(h) = \mathbb{E}\big\{Z(x)\overline{Z}(x+h)\big\} = \sigma^2 e^{-\frac{|h|}{a}}$$

where a is the correlation length (range) and σ^2 a constant.

Q. 5 5-a) What is the covariance $\gamma(0)$ at zero distance? Interpret this value.

Solution:

The covariance at zero distance is:

$$\gamma(0) = \mathbb{E}\big\{Z(x)\overline{Z}(x)\big\} = \sigma^2$$

This value represents the variance of the random field $Z(x)$, indicating the spread or variability of the measurements around their mean (which is zero in this case).

5-b) How does the covariance drop at a distance a?

Solution:

At a distance equal to the correlation length a, the covariance drops to:

$$\gamma(a) = \mathbb{E}\big\{Z(x)\overline{Z}(x+a)\big\} = \sigma^2 e^{-\frac{|a|}{a}} = \sigma^2 e^{-1} \approx 0.368\,\sigma^2$$

This indicates that at a distance equal to the correlation length, the covariance has decreased to approximately 36.8% of its maximum value at zero distance.

2.2. Optimal weights

We want to find weights $\boldsymbol{\lambda} = [\lambda_1, \lambda_2, ..., \lambda_N]$ that minimize the mean square prediction error:

$$\text{MSE} = \mathbb{E}\Big\{\big|\tilde{Z}(x_0) - Z(x_0)\big|^2\Big\}$$

Q. 6 6-a) Show that the prediction error can be written as:

$$\text{MSE} = \gamma(0) - 2\sum_{i=1}^{N} \lambda_i \gamma(x_0 - x_1) + \sum_{i=1}^{N}\sum_{j=1}^{N} \lambda_i \lambda_j \gamma\big(x_i - x_j\big)$$

Solution:

By expanding the square and using the definition of covariance, we have:

$$\text{MSE} = \mathbb{E}\left\{\left|\sum_{i=1}^{N} \lambda_i Z(x_i) - Z(x_0)\right|^2\right\} = \mathbb{E}\left\{\left(\sum_{i=1}^{N} \lambda_i Z(x_i) - Z(x_0)\right)\overline{\sum_{j=1}^{N} \lambda_j Z\big(x_j\big) - Z(x_0)}\right\}$$

Expanding this expression, we get:

$$\text{MSE} = \mathbb{E}\left\{\sum_{i=1}^{N}\sum_{j=1}^{N} \lambda_i \lambda_j Z(x_i)\overline{Z\big(x_j\big)} - 2\sum_{i=1}^{N} \lambda_i Z(x_i)\overline{Z(x_0)} + Z(x_0)\overline{Z(x_0)}\right\}$$

Using the definition of covariance, we can rewrite this as:

$$\text{MSE} = \sum_{i=1}^{N}\sum_{j=1}^{N}\lambda_i\lambda_j\gamma(x_i - x_j) - 2\sum_{i=1}^{N}\lambda_i\gamma(x_0 - x_i) + \gamma(0)$$

Rearranging the terms, we obtain:

$$\text{MSE} = \gamma(0) - 2\sum_{i=1}^{N}\lambda_i\gamma(x_0 - x_i) + \sum_{i=1}^{N}\sum_{j=1}^{N}\lambda_i\lambda_j\gamma(x_i - x_j)$$

6-b) Find the system of linear equations in λ_i that minimizes the MSE by setting the derivative with respect to λ_i to zero.

Solution:

Taking the derivative of the MSE with respect to λ_k and setting it to zero, we have:

$$\frac{\mathrm{d}\ \text{MSE}}{\mathrm{d}\lambda_k} = -2\gamma(x_0 - x_k) + 2\sum_{j=1}^{N}\lambda_j\gamma(x_k - x_j) = 0$$

Rearranging this equation, we obtain:

$$\sum_{j=1}^{N}\lambda_j\gamma(x_k - x_j) = \gamma(x_0 - x_k), \quad \forall k = 1, 2, \ldots, N$$

This gives us a system of N linear equations that can be solved to find the optimal weights λ_i.

6-c) Rewrite it in a matrix form leading to the simple kriging equations:

$$\boldsymbol{\lambda} = \boldsymbol{C}^{-1} \cdot \boldsymbol{c}$$

$$\tilde{Z}(x_0) = \boldsymbol{z}^T \cdot \boldsymbol{C}^{-1} \cdot \boldsymbol{c} = \sum_{i=1}^{N}\lambda_i Z(x_i)$$

Solution:

Let $\boldsymbol{\lambda} = [\lambda_1, \lambda_2, \ldots, \lambda_N]^T$ be the vector of weights, and let $\boldsymbol{C}$ be the covariance matrix with entries $\gamma(x_i - x_j)$ for $i, j = 1, 2, \ldots, N$. We can rewrite the system of equations in matrix form as:

$$\boldsymbol{C}\boldsymbol{\lambda} = \boldsymbol{c}$$

where:

- $C_{i,j} = \gamma(x_i - x_j)$
- $\boldsymbol{c}_i = \gamma(x_0 - x_i)$

This matrix equation can be solved for the weight vector $\boldsymbol{\lambda}$.

2.3. Properties

Q. 7 Estimate the bias of the estimator $\tilde{Z}(x_0)$, that is $\text{Bias} = \mathbb{E}\{\tilde{Z}(x_0) - Z(x_0)\}$

Solution:

The bias of the estimator $\tilde{Z}(x_0)$ is given by:

$$\text{Bias} = \mathbb{E}\{\tilde{Z}(x_0) - Z(x_0)\} = \mathbb{E}\left\{\sum_{i=1}^{N} \lambda_i Z(x_i) - Z(x_0)\right\}$$

Since $Z(x)$ has zero mean, we have:

$$\mathbb{E}\{Z(x_i)\} = 0, \quad \forall i$$

Therefore, the bias simplifies to:

$$\text{Bias} = \sum_{i=1}^{N} \lambda_i \mathbb{E}\{Z(x_i)\} - \mathbb{E}\{Z(x_0)\} = 0 - 0 = 0$$

This shows that the estimator $\tilde{Z}(x_0)$ is unbiased.

Q. 8 Estimate the variance of the estimator $\text{Var}(\tilde{Z}(x_0)) = \mathbb{E}\left\{\left|\tilde{Z}(x_0) - Z(x_0)\right|^2\right\}$

Solution:

The variance of the estimator $\tilde{Z}(x_0)$ can be computed as:

$$\text{Var}(\tilde{Z}(x_0)) = \mathbb{E}\left\{\left|\tilde{Z}(x_0) - \text{Bias}\right|^2\right\} = \mathbb{E}\left\{\left|\sum_{i=1}^{N} \lambda_i Z(x_i)\right|^2\right\}$$

Expanding this expression, we have:

$$\text{Var}(\tilde{Z}(x_0)) = \mathbb{E}\left\{\sum_{i=1}^{N}\sum_{j=1}^{N} \lambda_i \lambda_j Z(x_i)\overline{Z(x_j)}\right\}$$

Using the definition of covariance, we can rewrite this as:

$$\begin{aligned} \text{Var}(\tilde{Z}(x_0)) &= \sum_{i=1}^{N}\sum_{j=1}^{N} \lambda_i \lambda_j \gamma(x_i - x_j) \\ &= \boldsymbol{\lambda}^T \cdot \boldsymbol{C} \cdot \boldsymbol{\lambda} \\ &= \boldsymbol{\lambda}^T \cdot \boldsymbol{c} \end{aligned}$$

This expression gives the variance of the estimator $\tilde{Z}(x_0)$ in terms of the weights λ_i and the covariance function $\gamma(h)$.