Estimation Theory

Fredrik Rusek

Chapter 12

Chapter 12 – Linear Bayesian Estimation

Summary of chapters 10 and 11

- Bayesian estimators are injecting prior information into the estimation
- Concepts from classical estimation breaks down
  - MVU
  - Efficient estimator
  - unbiasedness
- Performance measure change: **variance -> Bayesian MSE**
- Optimal estimator for Bmse: $E(\theta|x)$. This is the MMSE estimator
- MMSE is difficult since
  - Posterior is hard to find $p(\theta|x)$
  - If we can find $p(\theta|x)$, then $E(\theta|x)$ is still difficult due to integral
- Conjugate priors simplify finding $p(\theta|x)$. Posterior has same distribution as prior (with other parameters). Useful when the posterior acts as prior in a sequential estimation process.
- Other risk functions than the Bmse exists.
  - MAP estimation is solution to hit-and-miss risk
  - Conditional Median is solution to a linear risk function
- Invariance does not hold for MAP
- Bayesian estimators can be used for deterministic parameters, but work well only for parameter values that are close to the prior mean

Chapter 12 – Linear Bayesian Estimation

When an optimal Bayesian estimator is hard to find, we can resort to a linear estimator

$$\hat{\theta} = \sum_{n=0}^{N-1} a_n x[n] + a_N$$

Chapter 12 – Linear Bayesian Estimation

When an optimal Bayesian estimator is hard to find, we can resort to a linear estimator

$$\hat{\theta} = \sum_{n=0}^{N-1} a_n x[n] + a_N$$

This is the same method as the BLUE, but we now make the following changes:

Chapter 12 – Linear Bayesian Estimation

When an optimal Bayesian estimator is hard to find, we can resort to a linear estimator

$$\hat{\theta} = \sum_{n=0}^{N-1} a_n x[n] + a_N$$

This is the same method as the BLUE, but we now make the following changes:

Remove Unbiasedness constraint

Chapter 12 – Linear Bayesian Estimation

When an optimal Bayesian estimator is hard to find, we can resort to a linear estimator

$$\hat{\theta} = \sum_{n=0}^{N-1} a_n x[n] + a_N$$

This is the same method as the BLUE, but we now make the following changes:

Remove Unbiasedness constraint

Change cost function from variance to Bmse $\mathrm{Bmse}(\hat{\theta}) = E\left[(\theta - \hat{\theta})^2\right]$

Chapter 12 – Linear Bayesian Estimation

When an optimal Bayesian estimator is hard to find, we can resort to a linear estimator

$$\hat{\theta} = \sum_{n=0}^{N-1} a_n x[n] + a_N$$

This is the same method as the BLUE, but we now make the following changes:

Remove Unbiasedness constraint

Change cost function from variance to Bmse $\mathrm{Bmse}(\hat{\theta}) = E\left[(\theta - \hat{\theta})^2\right]$

An optimal estimator within this class is termed the

linear minimum mean square error (LMMSE) estimator

Chapter 12 – Linear Bayesian Estimation

Finding the LMMSE estimator $\hat{\theta} = \sum_{n=0}^{N-1} a_n x[n] + a_N$

Cost function

$$\text{Bmse}(\hat{\theta}) = E\left[(\theta - \hat{\theta})^2\right] = E\left[\left(\theta - \sum_{n=0}^{N-1} a_n x[n] - a_N\right)^2\right]$$

Take differentials with respect to a_N

$$\frac{\partial}{\partial a_N} E\left[\left(\theta - \sum_{n=0}^{N-1} a_n x[n] - a_N\right)^2\right] =$$

Chapter 12 – Linear Bayesian Estimation

Finding the LMMSE estimator $\hat{\theta} = \sum_{n=0}^{N-1} a_n x[n] + a_N$

Cost function

$$\mathrm{Bmse}(\hat{\theta}) = E\left[(\theta - \hat{\theta})^2\right] = E\left[\left(\theta - \sum_{n=0}^{N-1} a_n x[n] - a_N\right)^2\right]$$

Take differentials with respect to a_N

$$\frac{\partial}{\partial a_N} E\left[\left(\theta - \sum_{n=0}^{N-1} a_n x[n] - a_N\right)^2\right] = -2E\left[\theta - \sum_{n=0}^{N-1} a_n x[n] - a_N\right] = 0$$

Chapter 12 – Linear Bayesian Estimation

Finding the LMMSE estimator $$\hat{\theta} = \sum_{n=0}^{N-1} a_n x[n] + a_N$$

Cost function

$$\mathrm{Bmse}(\hat{\theta}) = E\left[(\theta - \hat{\theta})^2\right] = E\left[\left(\theta - \sum_{n=0}^{N-1} a_n x[n] - a_N\right)^2\right]$$

Take differentials with respect to a_N

$$\frac{\partial}{\partial a_N} E\left[\left(\theta - \sum_{n=0}^{N-1} a_n x[n] - a_N\right)^2\right] = -2E\left[\theta - \sum_{n=0}^{N-1} a_n x[n] - a_N\right] = 0$$

$$a_N = E(\theta) - \sum_{n=0}^{N-1} a_n E(x[n])$$

Chapter 12 – Linear Bayesian Estimation

Finding the LMMSE estimator $$\hat{\theta} = \sum_{n=0}^{N-1} a_n x[n] + a_N$$

$$\mathrm{Bmse}(\hat{\theta}) = E\left[(\theta - \hat{\theta})^2\right] = E\left[\left(\theta - \sum_{n=0}^{N-1} a_n x[n] - a_N\right)^2\right]$$

Plug in a_N into the cost function $$a_N = E(\theta) - \sum_{n=0}^{N-1} a_n E(x[n])$$

Chapter 12 – Linear Bayesian Estimation

Finding the LMMSE estimator $\hat{\theta} = \sum_{n=0}^{N-1} a_n x[n] + a_N$

$$\mathrm{Bmse}(\hat{\theta}) = E\left[(\theta - \hat{\theta})^2\right] = E\left[\left(\theta - \sum_{n=0}^{N-1} a_n x[n] - a_N\right)^2\right]$$

$$= E\left\{\left[\sum_{n=0}^{N-1} a_n(x[n] - E(x[n])) - (\theta - E(\theta))\right]^2\right\}$$

Plug in a_N into the cost function

$$a_N = E(\theta) - \sum_{n=0}^{N-1} a_n E(x[n])$$

Chapter 12 – Linear Bayesian Estimation

Finding the LMMSE estimator $\hat{\theta} = \sum_{n=0}^{N-1} a_n x[n] + a_N$

$$\mathrm{Bmse}(\hat{\theta}) = E\left[(\theta - \hat{\theta})^2\right] = E\left[\left(\theta - \sum_{n=0}^{N-1} a_n x[n] - a_N\right)^2\right]$$

$$= E\left\{\left[\sum_{n=0}^{N-1} a_n (x[n] - E(x[n])) - (\theta - E(\theta))\right]^2\right\}$$

Assembly into vector notation

$$= E\left\{\left[\mathbf{a}^T(\mathbf{x} - E(\mathbf{x})) - (\theta - E(\theta))\right]^2\right\}$$

Chapter 12 – Linear Bayesian Estimation

Finding the LMMSE estimator $$\hat{\theta} = \sum_{n=0}^{N-1} a_n x[n] + a_N$$

$$\mathrm{Bmse}(\hat{\theta}) = E\left[(\theta - \hat{\theta})^2\right] = E\left[\left(\theta - \sum_{n=0}^{N-1} a_n x[n] - a_N\right)^2\right]$$

$$= E\left\{\left[\sum_{n=0}^{N-1} a_n(x[n] - E(x[n])) - (\theta - E(\theta))\right]^2\right\}$$

$$= E\left\{\left[\mathbf{a}^T(\mathbf{x} - E(\mathbf{x})) - (\theta - E(\theta))\right]^2\right\}$$

Generates 4 terms

Chapter 12 – Linear Bayesian Estimation

Finding the LMMSE estimator $$\hat{\theta} = \sum_{n=0}^{N-1} a_n x[n] + a_N$$

$$\text{Bmse}(\hat{\theta}) = E\left[(\theta - \hat{\theta})^2\right] = E\left[\left(\theta - \sum_{n=0}^{N-1} a_n x[n] - a_N\right)^2\right]$$

$$= E\left\{\left[\sum_{n=0}^{N-1} a_n(x[n] - E(x[n])) - (\theta - E(\theta))\right]^2\right\}$$

$$= E\left\{\left[\mathbf{a}^T(\mathbf{x} - E(\mathbf{x})) - (\theta - E(\theta))\right]^2\right\}$$

$$= E\left[\mathbf{a}^T(\mathbf{x} - E(\mathbf{x}))(\mathbf{x} - E(\mathbf{x}))^T\mathbf{a}\right] - E\left[\mathbf{a}^T(\mathbf{x} - E(\mathbf{x}))(\theta - E(\theta))\right]$$
$$- E\left[(\theta - E(\theta))(\mathbf{x} - E(\mathbf{x}))^T\mathbf{a}\right] + E\left[(\theta - E(\theta))^2\right]$$

Chapter 12 – Linear Bayesian Estimation

Finding the LMMSE estimator $\hat{\theta} = \sum_{n=0}^{N-1} a_n x[n] + a_N$

$$\mathrm{Bmse}(\hat{\theta}) = E\left[(\theta - \hat{\theta})^2\right] = E\left[\left(\theta - \sum_{n=0}^{N-1} a_n x[n] - a_N\right)^2\right]$$

$$= E\left\{\left[\sum_{n=0}^{N-1} a_n (x[n] - E(x[n])) - (\theta - E(\theta))\right]^2\right\}$$

$$= E\left\{\left[\mathbf{a}^T(\mathbf{x} - E(\mathbf{x})) - (\theta - E(\theta))\right]^2\right\}$$

$$= E\left[\mathbf{a}^T(\mathbf{x} - E(\mathbf{x}))(\mathbf{x} - E(\mathbf{x}))^T\mathbf{a}\right] - E\left[\mathbf{a}^T(\mathbf{x} - E(\mathbf{x}))(\theta - E(\theta))\right] - E\left[(\theta - E(\theta))(\mathbf{x} - E(\mathbf{x}))^T\mathbf{a}\right] + E\left[(\theta - E(\theta))^2\right]$$

Observe: a is not random, can be moved outside from expectation operator

Chapter 12 – Linear Bayesian Estimation

Finding the LMMSE estimator $\hat{\theta} = \sum_{n=0}^{N-1} a_n x[n] + a_N$

$$\mathrm{Bmse}(\hat{\theta}) = E\left[(\theta - \hat{\theta})^2\right] = E\left[\left(\theta - \sum_{n=0}^{N-1} a_n x[n] - a_N\right)^2\right]$$

$$= E\left\{\left[\sum_{n=0}^{N-1} a_n (x[n] - E(x[n])) - (\theta - E(\theta))\right]^2\right\}$$

$$= E\left\{\left[\mathbf{a}^T(\mathbf{x} - E(\mathbf{x})) - (\theta - E(\theta))\right]^2\right\}$$

$$= \boxed{E\left[\mathbf{a}^T(\mathbf{x} - E(\mathbf{x}))(\mathbf{x} - E(\mathbf{x}))^T\mathbf{a}\right]} - E\left[\mathbf{a}^T(\mathbf{x} - E(\mathbf{x}))(\theta - E(\theta))\right] - E\left[(\theta - E(\theta))(\mathbf{x} - E(\mathbf{x}))^T\mathbf{a}\right] + E\left[(\theta - E(\theta))^2\right]$$

$$= \mathbf{a}^T\mathbf{C}_{xx}\mathbf{a}$$

Chapter 12 – Linear Bayesian Estimation

Finding the LMMSE estimator $\hat{\theta} = \sum_{n=0}^{N-1} a_n x[n] + a_N$

$$\text{Bmse}(\hat{\theta}) = E\left[(\theta - \hat{\theta})^2\right] = E\left[\left(\theta - \sum_{n=0}^{N-1} a_n x[n] - a_N\right)^2\right]$$

$$= E\left\{\left[\sum_{n=0}^{N-1} a_n(x[n] - E(x[n])) - (\theta - E(\theta))\right]^2\right\}$$

$$= E\left\{\left[\mathbf{a}^T(\mathbf{x} - E(\mathbf{x})) - (\theta - E(\theta))\right]^2\right\}$$

$$= E\left[\mathbf{a}^T(\mathbf{x} - E(\mathbf{x}))(\mathbf{x} - E(\mathbf{x}))^T\mathbf{a}\right] - E\left[\mathbf{a}^T(\mathbf{x} - E(\mathbf{x}))(\theta - E(\theta))\right]$$
$$- E\left[(\theta - E(\theta))(\mathbf{x} - E(\mathbf{x}))^T\mathbf{a}\right] + E\left[(\theta - E(\theta))^2\right]$$

$$= \mathbf{a}^T\mathbf{C}_{xx}\mathbf{a} - \mathbf{a}^T\mathbf{C}_{x\theta}$$

Chapter 12 – Linear Bayesian Estimation

Finding the LMMSE estimator $\hat{\theta} = \sum_{n=0}^{N-1} a_n x[n] + a_N$

$$\mathrm{Bmse}(\hat{\theta}) = E\left[(\theta - \hat{\theta})^2\right] = E\left[\left(\theta - \sum_{n=0}^{N-1} a_n x[n] - a_N\right)^2\right]$$

$$= E\left\{\left[\sum_{n=0}^{N-1} a_n(x[n] - E(x[n])) - (\theta - E(\theta))\right]^2\right\}$$

$$= E\left\{\left[\mathbf{a}^T(\mathbf{x} - E(\mathbf{x})) - (\theta - E(\theta))\right]^2\right\}$$

$$= E\left[\mathbf{a}^T(\mathbf{x} - E(\mathbf{x}))(\mathbf{x} - E(\mathbf{x}))^T\mathbf{a}\right] - E\left[\mathbf{a}^T(\mathbf{x} - E(\mathbf{x}))(\theta - E(\theta))\right]$$
$$- E\left[(\theta - E(\theta))(\mathbf{x} - E(\mathbf{x}))^T\mathbf{a}\right] + E\left[(\theta - E(\theta))^2\right]$$

$$= \mathbf{a}^T\mathbf{C}_{xx}\mathbf{a} - \mathbf{a}^T\mathbf{C}_{x\theta} - \mathbf{C}_{\theta x}\mathbf{a}$$

Chapter 12 – Linear Bayesian Estimation

Finding the LMMSE estimator $\hat{\theta} = \sum_{n=0}^{N-1} a_n x[n] + a_N$

$$\text{Bmse}(\hat{\theta}) = E\left[(\theta - \hat{\theta})^2\right] = E\left[\left(\theta - \sum_{n=0}^{N-1} a_n x[n] - a_N\right)^2\right]$$

$$= E\left\{\left[\sum_{n=0}^{N-1} a_n(x[n] - E(x[n])) - (\theta - E(\theta))\right]^2\right\}$$

$$= E\left\{\left[\mathbf{a}^T(\mathbf{x} - E(\mathbf{x})) - (\theta - E(\theta))\right]^2\right\}$$

$$= E\left[\mathbf{a}^T(\mathbf{x} - E(\mathbf{x}))(\mathbf{x} - E(\mathbf{x}))^T\mathbf{a}\right] - E\left[\mathbf{a}^T(\mathbf{x} - E(\mathbf{x}))(\theta - E(\theta))\right] - E\left[(\theta - E(\theta))(\mathbf{x} - E(\mathbf{x}))^T\mathbf{a}\right] + E\left[(\theta - E(\theta))^2\right]$$

$$= \mathbf{a}^T\mathbf{C}_{xx}\mathbf{a} - \mathbf{a}^T\mathbf{C}_{x\theta} - \mathbf{C}_{\theta x}\mathbf{a} + C_{\theta\theta}$$

Chapter 12 – Linear Bayesian Estimation

Finding the LMMSE estimator $\hat{\theta} = \sum_{n=0}^{N-1} a_n x[n] + a_N$

$$\mathrm{Bmse}(\hat{\theta}) = E\left[(\theta - \hat{\theta})^2\right] = E\left[\left(\theta - \sum_{n=0}^{N-1} a_n x[n] - a_N\right)^2\right]$$

$$= E\left\{\left[\sum_{n=0}^{N-1} a_n (x[n] - E(x[n])) - (\theta - E(\theta))\right]^2\right\}$$

$$= E\left\{\left[\mathbf{a}^T(\mathbf{x} - E(\mathbf{x})) - (\theta - E(\theta))\right]^2\right\}$$

$$= E\left[\mathbf{a}^T(\mathbf{x} - E(\mathbf{x}))(\mathbf{x} - E(\mathbf{x}))^T\mathbf{a}\right] - E\left[\mathbf{a}^T(\mathbf{x} - E(\mathbf{x}))(\theta - E(\theta))\right] - E\left[(\theta - E(\theta))(\mathbf{x} - E(\mathbf{x}))^T\mathbf{a}\right] + E\left[(\theta - E(\theta))^2\right]$$

$$= \mathbf{a}^T\mathbf{C}_{xx}\mathbf{a} - \mathbf{a}^T\mathbf{C}_{x\theta} - \mathbf{C}_{\theta x}\mathbf{a} + C_{\theta\theta}$$

$$= \mathbf{a}^T\mathbf{C}_{xx}\mathbf{a} - 2\,\mathbf{a}^T\mathbf{C}_{x\theta} + C_{\theta\theta}$$

Chapter 12 – Linear Bayesian Estimation

Finding the LMMSE estimator $\hat{\theta} = \sum_{n=0}^{N-1} a_n x[n] + a_N$

$$\text{Bmse}(\hat{\theta}) = E\left[(\theta - \hat{\theta})^2\right] = E\left[\left(\theta - \sum_{n=0}^{N-1} a_n x[n] - a_N\right)^2\right]$$

$$= E\left\{\left[\sum_{n=0}^{N-1} a_n(x[n] - E(x[n])) - (\theta - E(\theta))\right]^2\right\}$$

$$= E\left\{\left[\mathbf{a}^T(\mathbf{x} - E(\mathbf{x})) - (\theta - E(\theta))\right]^2\right\}$$

$$= E\left[\mathbf{a}^T(\mathbf{x} - E(\mathbf{x}))(\mathbf{x} - E(\mathbf{x}))^T\mathbf{a}\right] - E\left[\mathbf{a}^T(\mathbf{x} - E(\mathbf{x}))(\theta - E(\theta))\right] - E\left[(\theta - E(\theta))(\mathbf{x} - E(\mathbf{x}))^T\mathbf{a}\right] + E\left[(\theta - E(\theta))^2\right]$$

$$= \mathbf{a}^T\mathbf{C}_{xx}\mathbf{a} - \mathbf{a}^T\mathbf{C}_{x\theta} - \mathbf{C}_{\theta x}\mathbf{a} + C_{\theta\theta}$$

$$= \mathbf{a}^T\mathbf{C}_{xx}\mathbf{a} - 2\,\mathbf{a}^T\mathbf{C}_{x\theta} + C_{\theta\theta}$$

$$\frac{\partial \text{Bmse}(\hat{\theta})}{\partial \mathbf{a}} = 2\mathbf{C}_{xx}\mathbf{a} - 2\mathbf{C}_{x\theta}$$

$$\mathbf{a} = \mathbf{C}_{xx}^{-1}\mathbf{C}_{x\theta}$$

Chapter 12 – Linear Bayesian Estimation

Collect the results using vector notation

$$\frac{\partial \text{Bmse}(\hat{\theta})}{\partial \mathbf{a}} = 2\mathbf{C}_{xx}\mathbf{a} - 2\mathbf{C}_{x\theta}$$

$$\mathbf{a} = \mathbf{C}_{xx}^{-1}\mathbf{C}_{x\theta}$$

$$a_N = E(\theta) - \sum_{n=0}^{N-1} a_n E(x[n])$$

$$= E(\theta) - \mathbf{a}^T E(\mathbf{x})$$

$$\hat{\theta} = \sum_{n=0}^{N-1} a_n x[n] + a_N$$

$$= \mathbf{a}^T \mathbf{x} + a_N$$

Chapter 12 – Linear Bayesian Estimation

Collect the results using vector notation

$$\frac{\partial \text{Bmse}(\hat{\theta})}{\partial \mathbf{a}} = 2\mathbf{C}_{xx}\mathbf{a} - 2\mathbf{C}_{x\theta}$$

$$\mathbf{a} = \mathbf{C}_{xx}^{-1}\mathbf{C}_{x\theta}$$

$$a_N = E(\theta) - \sum_{n=0}^{N-1} a_n E(x[n]) = E(\theta) - \mathbf{a}^T E(\mathbf{x})$$

$$\hat{\theta} = \sum_{n=0}^{N-1} a_n x[n] + a_N = \mathbf{a}^T\mathbf{x} + a_N$$

$$\hat{\theta} = \mathbf{C}_{x\theta}^T\mathbf{C}_{xx}^{-1}\mathbf{x} + E(\theta) - \mathbf{C}_{x\theta}^T\mathbf{C}_{xx}^{-1}E(\mathbf{x}) = E(\theta) + \mathbf{C}_{\theta x}\mathbf{C}_{xx}^{-1}(\mathbf{x} - E(\mathbf{x}))$$

Chapter 12 – Linear Bayesian Estimation

Computing the Bmse cost

$$\mathrm{Bmse}(\hat{\theta}) = \mathbf{a}^T\mathbf{C}_{xx}\mathbf{a} - \mathbf{a}^T\mathbf{C}_{x\theta} - \mathbf{C}_{\theta x}\mathbf{a} + C_{\theta\theta}$$

$$\mathbf{a} = \mathbf{C}_{xx}^{-1}\mathbf{C}_{x\theta}$$

$$\mathrm{Bmse}(\hat{\theta}) = C_{\theta\theta} - \mathbf{C}_{\theta x}\mathbf{C}_{xx}^{-1}\mathbf{C}_{x\theta}$$

Schur complement

Chapter 12 – Linear Bayesian Estimation

Connections

$$\mathrm{Bmse}(\hat{\theta}) = C_{\theta\theta} - \mathbf{C}_{\theta x}\mathbf{C}_{xx}^{-1}\mathbf{C}_{x\theta}$$

We have seen the expression for the BMSE before

Theorem 10.2 (Conditional PDF of Multivariate Gaussian) *If* $\mathbf{x}$ *and* $\mathbf{y}$ *are jointly Gaussian, where* $\mathbf{x}$ *is* $k \times 1$ *and* $\mathbf{y}$ *is* $l \times 1$, *with mean vector* $[E(\mathbf{x})^T \; E(\mathbf{y})^T]^T$ *and partitioned covariance matrix*

$$\mathbf{C} = \begin{bmatrix} \mathbf{C}_{xx} & \mathbf{C}_{xy} \\ \mathbf{C}_{yx} & \mathbf{C}_{yy} \end{bmatrix} = \begin{bmatrix} k \times k & k \times l \\ l \times k & l \times l \end{bmatrix} \tag{10.23}$$

so that

$$p(\mathbf{x}, \mathbf{y}) = \frac{1}{(2\pi)^{\frac{k+l}{2}} \det^{\frac{1}{2}}(\mathbf{C})} \exp\left[-\frac{1}{2}\left(\begin{bmatrix} \mathbf{x} - E(\mathbf{x}) \\ \mathbf{y} - E(\mathbf{y}) \end{bmatrix}\right)^T \mathbf{C}^{-1}\left(\begin{bmatrix} \mathbf{x} - E(\mathbf{x}) \\ \mathbf{y} - E(\mathbf{y}) \end{bmatrix}\right)\right],$$

then the conditional PDF $p(\mathbf{y}|\mathbf{x})$ *is also Gaussian and*

$$E(\mathbf{y}|\mathbf{x}) = E(\mathbf{y}) + \mathbf{C}_{yx}\mathbf{C}_{xx}^{-1}(\mathbf{x} - E(\mathbf{x})) \tag{10.24}$$

$$\mathbf{C}_{y|x} = \mathbf{C}_{yy} - \mathbf{C}_{yx}\mathbf{C}_{xx}^{-1}\mathbf{C}_{xy}. \tag{10.25}$$

Chapter 12 – Linear Bayesian Estimation

Connections

	X and θ jointly Gaussian	X and θ **not** jointly Gaussian

Chapter 12 – Linear Bayesian Estimation

Connections

	X and θ jointly Gaussian	X and θ **not** jointly Gaussian
LMMSE estimator	$\hat{\theta} = E(\theta) + \mathbf{C}_{\theta x}\mathbf{C}_{xx}^{-1}(\mathbf{x} - E(\mathbf{x}))$	$\hat{\theta} = E(\theta) + \mathbf{C}_{\theta x}\mathbf{C}_{xx}^{-1}(\mathbf{x} - E(\mathbf{x}))$
Bmse		

Chapter 12 – Linear Bayesian Estimation

Connections

	X and θ jointly Gaussian	X and θ **not** jointly Gaussian
LMMSE estimator	$\hat{\theta} = E(\theta) + \mathbf{C}_{\theta x}\mathbf{C}_{xx}^{-1}(\mathbf{x} - E(\mathbf{x}))$	$\hat{\theta} = E(\theta) + \mathbf{C}_{\theta x}\mathbf{C}_{xx}^{-1}(\mathbf{x} - E(\mathbf{x}))$
Bmse	$C_{\theta\theta} - \mathbf{C}_{\theta x}\mathbf{C}_{xx}^{-1}\mathbf{C}_{x\theta}$	$C_{\theta\theta} - \mathbf{C}_{\theta x}\mathbf{C}_{xx}^{-1}\mathbf{C}_{x\theta}$

Chapter 12 – Linear Bayesian Estimation

Connections

	X and θ jointly Gaussian	X and θ **not** jointly Gaussian
LMMSE estimator	$\hat{\theta} = E(\theta) + \mathbf{C}_{\theta x}\mathbf{C}_{xx}^{-1}(\mathbf{x} - E(\mathbf{x}))$	$\hat{\theta} = E(\theta) + \mathbf{C}_{\theta x}\mathbf{C}_{xx}^{-1}(\mathbf{x} - E(\mathbf{x}))$
Bmse	$C_{\theta\theta} - \mathbf{C}_{\theta x}\mathbf{C}_{xx}^{-1}\mathbf{C}_{x\theta}$	$C_{\theta\theta} - \mathbf{C}_{\theta x}\mathbf{C}_{xx}^{-1}\mathbf{C}_{x\theta}$
MMSE estimator	$E(\boldsymbol{\theta}\|\mathbf{x}) = E(\theta) + \mathbf{C}_{\theta x}\mathbf{C}_{xx}^{-1}(\mathbf{x} - E(\mathbf{x}))$ **LMMSE = MMSE**	$E(\boldsymbol{\theta}\|\mathbf{x}) \neq E(\theta) + \mathbf{C}_{\theta x}\mathbf{C}_{xx}^{-1}(\mathbf{x} - E(\mathbf{x}))$

Chapter 12 – Linear Bayesian Estimation

Connections

	X and θ jointly Gaussian	X and θ **not** jointly Gaussian
LMMSE estimator	$\hat{\theta} = E(\theta) + \mathbf{C}_{\theta x}\mathbf{C}_{xx}^{-1}(\mathbf{x} - E(\mathbf{x}))$	$\hat{\theta} = E(\theta) + \mathbf{C}_{\theta x}\mathbf{C}_{xx}^{-1}(\mathbf{x} - E(\mathbf{x}))$
Bmse	$C_{\theta\theta} - \mathbf{C}_{\theta x}\mathbf{C}_{xx}^{-1}\mathbf{C}_{x\theta}$	$C_{\theta\theta} - \mathbf{C}_{\theta x}\mathbf{C}_{xx}^{-1}\mathbf{C}_{x\theta}$
MMSE estimator	$E(\boldsymbol{\theta}\vert\mathbf{x}) = E(\theta) + \mathbf{C}_{\theta x}\mathbf{C}_{xx}^{-1}(\mathbf{x} - E(\mathbf{x}))$ **LMMSE = MMSE**	$E(\boldsymbol{\theta}\vert\mathbf{x}) \neq E(\theta) + \mathbf{C}_{\theta x}\mathbf{C}_{xx}^{-1}(\mathbf{x} - E(\mathbf{x}))$
Bmse	$C_{\theta\theta} - \mathbf{C}_{\theta x}\mathbf{C}_{xx}^{-1}\mathbf{C}_{x\theta}$	**Better than** $C_{\theta\theta} - \mathbf{C}_{\theta x}\mathbf{C}_{xx}^{-1}\mathbf{C}_{x\theta}$

Chapter 12 – Linear Bayesian Estimation

Connections

	X and θ jointly Gaussian	X and θ **not** jointly Gaussian
LMMSE estimator	$\hat{\theta} = E(\theta) + \mathbf{C}_{\theta x}\mathbf{C}_{xx}^{-1}(\mathbf{x} - E(\mathbf{x}))$	$\hat{\theta} = E(\theta) + \mathbf{C}_{\theta x}\mathbf{C}_{xx}^{-1}(\mathbf{x} - E(\mathbf{x}))$
Bmse	$C_{\theta\theta} - \mathbf{C}_{\theta x}\mathbf{C}_{xx}^{-1}\mathbf{C}_{x\theta}$	$C_{\theta\theta} - \mathbf{C}_{\theta x}\mathbf{C}_{xx}^{-1}\mathbf{C}_{x\theta}$
MMSE estimator	$E(\boldsymbol{\theta}\vert\mathbf{x}) = E(\theta) + \mathbf{C}_{\theta x}\mathbf{C}_{xx}^{-1}(\mathbf{x} - E(\mathbf{x}))$ **LMMSE = MMSE**	$E(\boldsymbol{\theta}\vert\mathbf{x}) \neq E(\theta) + \mathbf{C}_{\theta x}\mathbf{C}_{xx}^{-1}(\mathbf{x} - E(\mathbf{x}))$
Bmse	$C_{\theta\theta} - \mathbf{C}_{\theta x}\mathbf{C}_{xx}^{-1}\mathbf{C}_{x\theta}$	**Better than** $C_{\theta\theta} - \mathbf{C}_{\theta x}\mathbf{C}_{xx}^{-1}\mathbf{C}_{x\theta}$

The LMMSE yields the same Bmse as if the variables are jointly Gaussian

Chapter 12 – Linear Bayesian Estimation

Example 12.1 $x[n] = A + w[n] \qquad n = 0, 1, \ldots, N-1 \qquad A \sim \mathcal{U}[-A_0, A_0]$

Bayesian options:

1. MMSE
2. MAP
3. LMMSE

Chapter 12 – Linear Bayesian Estimation

Example 12.1 $x[n] = A + w[n] \qquad n = 0, 1, \ldots, N-1 \qquad A \sim \mathcal{U}[-A_0, A_0]$

Bayesian options:

1. MMSE. **Not possible in closed form**
2. MAP. **Possible: truncated sample mean**
3. LMMSE

Chapter 12 – Linear Bayesian Estimation

Example 12.1 $x[n] = A + w[n] \qquad n = 0, 1, \ldots, N-1 \qquad A \sim \mathcal{U}[-A_0, A_0]$

Bayesian options:

1. MMSE. **Not possible in closed form**
2. MAP. **Possible: truncated sample mean**
3. LMMSE

$$\hat{A} = E(\theta) + \mathbf{C}_{\theta x}\mathbf{C}_{xx}^{-1}(\mathbf{x} - E(\mathbf{x}))$$

All means are zero

$$\hat{A} = \mathbf{C}_{\theta x}\mathbf{C}_{xx}^{-1}\mathbf{x}$$

Chapter 12 – Linear Bayesian Estimation

Example 12.1 $x[n] = A + w[n] \qquad n = 0, 1, \ldots, N-1 \qquad A \sim \mathcal{U}[-A_0, A_0]$

Bayesian options:

1. MMSE. **Not possible in closed form**
2. MAP. **Possible: truncated sample mean**
3. LMMSE

$$\hat{A} = E(\theta) + \mathbf{C}_{\theta x}\mathbf{C}_{xx}^{-1}(\mathbf{x} - E(\mathbf{x}))$$

$$\hat{A} = \mathbf{C}_{\theta x}\mathbf{C}_{xx}^{-1}\mathbf{x}$$

$$\begin{aligned}
\mathbf{C}_{xx} &= E(\mathbf{x}\mathbf{x}^T) \\
&= E\left[(A\mathbf{1} + \mathbf{w})(A\mathbf{1} + \mathbf{w})^T\right] \\
&= E(A^2)\mathbf{1}\mathbf{1}^T + \sigma^2\mathbf{I} \\
\mathbf{C}_{\theta x} &= E(A\mathbf{x}^T) \\
&= E\left[A(A\mathbf{1} + \mathbf{w})^T\right] \\
&= E(A^2)\mathbf{1}^T
\end{aligned}$$

Chapter 12 – Linear Bayesian Estimation

Example 12.1 $x[n] = A + w[n] \qquad n = 0, 1, \ldots, N-1 \qquad A \sim \mathcal{U}[-A_0, A_0]$

Bayesian options:

1. MMSE. **Not possible in closed form**
2. MAP. **Possible: truncated sample mean**
3. LMMSE

$$\hat{A} = E(\theta) + \mathbf{C}_{\theta x}\mathbf{C}_{xx}^{-1}(\mathbf{x} - E(\mathbf{x}))$$

$$\hat{A} = \mathbf{C}_{\theta x}\mathbf{C}_{xx}^{-1}\mathbf{x}$$

$$\begin{aligned}\hat{A} &= \mathbf{C}_{\theta x}\mathbf{C}_{xx}^{-1}\mathbf{x} \\ &= \sigma_A^2\mathbf{1}^T(\sigma_A^2\mathbf{1}\mathbf{1}^T + \sigma^2\mathbf{I})^{-1}\mathbf{x}\end{aligned}$$

$$\begin{aligned}\mathbf{C}_{xx} &= E(\mathbf{x}\mathbf{x}^T) \\ &= E\left[(A\mathbf{1} + \mathbf{w})(A\mathbf{1} + \mathbf{w})^T\right] \\ &= E(A^2)\mathbf{1}\mathbf{1}^T + \sigma^2\mathbf{I} \\ \mathbf{C}_{\theta x} &= E(A\mathbf{x}^T) \\ &= E\left[A(A\mathbf{1} + \mathbf{w})^T\right] \\ &= E(A^2)\mathbf{1}^T\end{aligned}$$

$$\sigma_A^2 = E(A^2)$$

Chapter 12 – Linear Bayesian Estimation

Example 12.1 $x[n] = A + w[n] \qquad n = 0, 1, \ldots, N-1 \qquad A \sim \mathcal{U}[-A_0, A_0]$

Bayesian options:

1. MMSE. **Not possible in closed form**
2. MAP. **Possible: truncated sample mean**
3. LMMSE. **Doable**

$$\hat{A} = E(\theta) + \mathbf{C}_{\theta x}\mathbf{C}_{xx}^{-1}(\mathbf{x} - E(\mathbf{x}))$$

$$\hat{A} = \mathbf{C}_{\theta x}\mathbf{C}_{xx}^{-1}\mathbf{x}$$

$$\begin{aligned} \mathbf{C}_{xx} &= E(\mathbf{x}\mathbf{x}^T) \\ &= E\left[(A\mathbf{1} + \mathbf{w})(A\mathbf{1} + \mathbf{w})^T\right] \\ &= E(A^2)\mathbf{1}\mathbf{1}^T + \sigma^2\mathbf{I} \\ \mathbf{C}_{\theta x} &= E(A\mathbf{x}^T) \\ &= E\left[A(A\mathbf{1} + \mathbf{w})^T\right] \\ &= E(A^2)\mathbf{1}^T \end{aligned}$$

$$\sigma_A^2 = E(A^2)$$

$$\begin{aligned} \hat{A} &= \mathbf{C}_{\theta x}\mathbf{C}_{xx}^{-1}\mathbf{x} \\ &= \sigma_A^2\mathbf{1}^T(\sigma_A^2\mathbf{1}\mathbf{1}^T + \sigma^2\mathbf{I})^{-1}\mathbf{x} \end{aligned}$$

$$= \ldots\ldots = \boxed{\hat{A} = \frac{\frac{A_0^2}{3}}{\frac{A_0^2}{3} + \frac{\sigma^2}{N}}\bar{x}}$$

Chapter 12 – Linear Bayesian Estimation

Example 12.1 $x[n] = A + w[n] \qquad n = 0, 1, \ldots, N-1 \qquad A \sim \mathcal{U}[-A_0, A_0]$

Observations

1. With no prior, the sample mean is MVU
2. The LMMSE is a tradeoff between the MVU and the sample mean
3. We did not use the fact that A is uniform, only its mean and variance comes in
4. We do not need A and w to be independent, only uncorrelated
5. No integration is needed

$$\hat{A} = \frac{\frac{A_0^2}{3}}{\frac{A_0^2}{3} + \frac{\sigma^2}{N}} \bar{x}$$

Chapter 12 – Linear Bayesian Estimation

Extension to vector parameter

We can work with each parameter individually

$$\hat{\theta}_i = \sum_{n=0}^{N-1} a_{in}x[n] + a_{iN} \qquad \mathrm{Bmse}(\hat{\theta}_i) = E\left[(\theta_i - \hat{\theta}_i)^2\right] \qquad i = 1, 2, \ldots, p$$

Chapter 12 – Linear Bayesian Estimation

Extension to vector parameter

We can work with each parameter individually

$$\hat{\theta}_i = \sum_{n=0}^{N-1} a_{in} x[n] + a_{iN} \qquad \text{Bmse}(\hat{\theta}_i) = E\left[(\theta_i - \hat{\theta}_i)^2\right] \qquad i = 1, 2, \ldots, p$$

From before, we have that

$$\hat{\theta}_i = E(\theta_i) + \mathbf{C}_{\theta_i x} \mathbf{C}_{xx}^{-1} (\mathbf{x} - E(\mathbf{x})) \qquad i = 1, 2, \ldots, p$$

$$\text{Bmse}(\hat{\theta}_i) = C_{\theta_i \theta_i} - \mathbf{C}_{\theta_i x} \mathbf{C}_{xx}^{-1} \mathbf{C}_{x \theta_i} \qquad i = 1, 2, \ldots, p$$

Chapter 12 – Linear Bayesian Estimation

Extension to vector parameter

We can work with each parameter individually

$$\hat{\theta}_i = \sum_{n=0}^{N-1} a_{in} x[n] + a_{iN} \qquad \mathrm{Bmse}(\hat{\theta}_i) = E\left[(\theta_i - \hat{\theta}_i)^2\right] \qquad i = 1, 2, \ldots, p$$

From before, we have that

$$\hat{\theta}_i = E(\theta_i) + \mathbf{C}_{\theta_i x}\mathbf{C}_{xx}^{-1}(\mathbf{x} - E(\mathbf{x}))$$

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

collect in vector notation

$$\begin{aligned}
\hat{\boldsymbol{\theta}} &= \begin{bmatrix} E(\theta_1) \\ E(\theta_2) \\ \vdots \\ E(\theta_p) \end{bmatrix} + \begin{bmatrix} \mathbf{C}_{\theta_1 x}\mathbf{C}_{xx}^{-1}(\mathbf{x} - E(\mathbf{x})) \\ \mathbf{C}_{\theta_2 x}\mathbf{C}_{xx}^{-1}(\mathbf{x} - E(\mathbf{x})) \\ \vdots \\ \mathbf{C}_{\theta_p x}\mathbf{C}_{xx}^{-1}(\mathbf{x} - E(\mathbf{x})) \end{bmatrix} \\
&= \begin{bmatrix} E(\theta_1) \\ E(\theta_2) \\ \vdots \\ E(\theta_p) \end{bmatrix} + \begin{bmatrix} \mathbf{C}_{\theta_1 x} \\ \mathbf{C}_{\theta_2 x} \\ \vdots \\ \mathbf{C}_{\theta_p x} \end{bmatrix} \mathbf{C}_{xx}^{-1}(\mathbf{x} - E(\mathbf{x})) \\
&= E(\boldsymbol{\theta}) + \mathbf{C}_{\theta x}\mathbf{C}_{xx}^{-1}(\mathbf{x} - E(\mathbf{x}))
\end{aligned}$$

Chapter 12 – Linear Bayesian Estimation

Extension to vector parameter

We can work with each parameter individually

$$\hat{\theta}_i = \sum_{n=0}^{N-1} a_{in} x[n] + a_{iN} \qquad \mathrm{Bmse}(\hat{\theta}_i) = E\left[(\theta_i - \hat{\theta}_i)^2\right] \qquad i = 1, 2, \ldots, p$$

From before, we have that

$$\hat{\theta}_i = E(\theta_i) + \mathbf{C}_{\theta_i x}\mathbf{C}_{xx}^{-1}(\mathbf{x} - E(\mathbf{x}))$$

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

BMSE

$$\mathbf{M}_{\hat{\theta}} = E\left[(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})^T\right] = \mathbf{C}_{\theta\theta} - \mathbf{C}_{\theta x}\mathbf{C}_{xx}^{-1}\mathbf{C}_{x\theta}$$

$$\mathrm{Bmse}(\hat{\theta}_i) = [\mathbf{M}_{\hat{\theta}}]_{ii}$$

collect in vector notation

$$\begin{aligned}
\hat{\boldsymbol{\theta}} &= \begin{bmatrix} E(\theta_1) \\ E(\theta_2) \\ \vdots \\ E(\theta_p) \end{bmatrix} + \begin{bmatrix} \mathbf{C}_{\theta_1 x}\mathbf{C}_{xx}^{-1}(\mathbf{x} - E(\mathbf{x})) \\ \mathbf{C}_{\theta_2 x}\mathbf{C}_{xx}^{-1}(\mathbf{x} - E(\mathbf{x})) \\ \vdots \\ \mathbf{C}_{\theta_p x}\mathbf{C}_{xx}^{-1}(\mathbf{x} - E(\mathbf{x})) \end{bmatrix} \\
&= \begin{bmatrix} E(\theta_1) \\ E(\theta_2) \\ \vdots \\ E(\theta_p) \end{bmatrix} + \begin{bmatrix} \mathbf{C}_{\theta_1 x} \\ \mathbf{C}_{\theta_2 x} \\ \vdots \\ \mathbf{C}_{\theta_p x} \end{bmatrix} \mathbf{C}_{xx}^{-1}(\mathbf{x} - E(\mathbf{x})) \\
&= E(\boldsymbol{\theta}) + \mathbf{C}_{\theta x}\mathbf{C}_{xx}^{-1}(\mathbf{x} - E(\mathbf{x}))
\end{aligned}$$

Chapter 12 – Linear Bayesian Estimation

Three properties

1. **Invariance holds for affine transformations**

$$\boldsymbol{\alpha} = \mathbf{A}\boldsymbol{\theta} + \mathbf{b} \longrightarrow \hat{\boldsymbol{\alpha}} = \mathbf{A}\hat{\boldsymbol{\theta}} + \mathbf{b}$$

Chapter 12 – Linear Bayesian Estimation

Three properties

1. **Invariance holds for affine transformations**

$$\boldsymbol{\alpha} = \mathbf{A}\boldsymbol{\theta} + \mathbf{b} \longrightarrow \hat{\boldsymbol{\alpha}} = \mathbf{A}\hat{\boldsymbol{\theta}} + \mathbf{b}$$

2. **LMMSE of a sum is a sum of the LMMSE**

$$\boldsymbol{\alpha} = \boldsymbol{\theta}_1 + \boldsymbol{\theta}_2 \longrightarrow \hat{\boldsymbol{\alpha}} = \hat{\boldsymbol{\theta}}_1 + \hat{\boldsymbol{\theta}}_2$$

Chapter 12 – Linear Bayesian Estimation

Three properties

1. **Invariance holds for affine transformations**

$$\boldsymbol{\alpha} = \mathbf{A}\boldsymbol{\theta} + \mathbf{b} \longrightarrow \hat{\boldsymbol{\alpha}} = \mathbf{A}\hat{\boldsymbol{\theta}} + \mathbf{b}$$

2. **LMMSE of a sum is a sum of the LMMSE**

$$\boldsymbol{\alpha} = \boldsymbol{\theta}_1 + \boldsymbol{\theta}_2 \longrightarrow \hat{\boldsymbol{\alpha}} = \hat{\boldsymbol{\theta}}_1 + \hat{\boldsymbol{\theta}}_2$$

3. **Number of observations can be less than parameters to estimate *a significant difference from linear classical estimation***

Chapter 12 – Linear Bayesian Estimation

With $\mathbf{x} = \mathbf{H}\boldsymbol{\theta} + \mathbf{w}$ **we have**

$$\begin{aligned} E(\mathbf{x}) &= \mathbf{H}E(\boldsymbol{\theta}) \\ \mathbf{C}_{xx} &= \mathbf{H}\mathbf{C}_{\theta\theta}\mathbf{H}^T + \mathbf{C}_w \\ \mathbf{C}_{\theta x} &= \mathbf{C}_{\theta\theta}\mathbf{H}^T. \end{aligned}$$

Chapter 12 – Linear Bayesian Estimation

Theorem 12.1 (Bayesian Gauss-Markov Theorem) *If the data are described by the Bayesian linear model form*

$$\mathbf{x} = \mathbf{H}\boldsymbol{\theta} + \mathbf{w} \tag{12.25}$$

where $\mathbf{x}$ *is an* $N \times 1$ *data vector,* $\mathbf{H}$ *is a known* $N \times p$ *observation matrix,* $\boldsymbol{\theta}$ *is a* $p \times 1$ *random vector of parameters whose realization is to be estimated and has mean* $E(\boldsymbol{\theta})$ *and covariance matrix* $\mathbf{C}_{\theta\theta}$*, and* $\mathbf{w}$ *is an* $N \times 1$ *random vector with zero mean and covariance matrix* $\mathbf{C}_w$ *and is uncorrelated with* $\boldsymbol{\theta}$ *(the joint PDF* $p(\mathbf{w}, \boldsymbol{\theta})$ *is otherwise arbitrary), then the LMMSE estimator of* $\boldsymbol{\theta}$ *is*

$$\begin{aligned}\hat{\boldsymbol{\theta}} &= E(\boldsymbol{\theta}) + \mathbf{C}_{\theta\theta}\mathbf{H}^T(\mathbf{H}\mathbf{C}_{\theta\theta}\mathbf{H}^T + \mathbf{C}_w)^{-1}(\mathbf{x} - \mathbf{H}E(\boldsymbol{\theta})) && (12.26)\\ &= E(\boldsymbol{\theta}) + (\mathbf{C}_{\theta\theta}^{-1} + \mathbf{H}^T\mathbf{C}_w^{-1}\mathbf{H})^{-1}\mathbf{H}^T\mathbf{C}_w^{-1}(\mathbf{x} - \mathbf{H}E(\boldsymbol{\theta})). && (12.27)\end{aligned}$$

The performance of the estimator is measured by the error $\boldsymbol{\epsilon} = \boldsymbol{\theta} - \hat{\boldsymbol{\theta}}$ *whose mean is zero and whose covariance matrix is*

$$\begin{aligned}\mathbf{C}_\epsilon &= E_{x,\theta}(\boldsymbol{\epsilon}\boldsymbol{\epsilon}^T)\\ &= \mathbf{C}_{\theta\theta} - \mathbf{C}_{\theta\theta}\mathbf{H}^T(\mathbf{H}\mathbf{C}_{\theta\theta}\mathbf{H}^T + \mathbf{C}_w)^{-1}\mathbf{H}\mathbf{C}_{\theta\theta} && (12.28)\\ &= (\mathbf{C}_{\theta\theta}^{-1} + \mathbf{H}^T\mathbf{C}_w^{-1}\mathbf{H})^{-1}. && (12.29)\end{aligned}$$

The error covariance matrix is also the minimum MSE matrix $\mathbf{M}_{\hat{\theta}}$ *whose diagonal elements yield the minimum Bayesian MSE*

$$\begin{aligned}[\mathbf{M}_{\hat{\theta}}]_{ii} &= [\mathbf{C}_\epsilon]_{ii}\\ &= \mathrm{Bmse}(\hat{\theta}_i). && (12.30)\end{aligned}$$

Chapter 12 – Linear Bayesian Estimation

Wiener filtering

Data: x[0], x[1],x[2],… WSS with zero mean

Covariance matrix = ??

Chapter 12 – Linear Bayesian Estimation

Wiener filtering

Data: x[0], x[1],x[2],… WSS with zero mean

Covariance matrix

$$
\begin{aligned}
\mathbf{C}_{xx} &= \begin{bmatrix} r_{xx}[0] & r_{xx}[1] & \ldots & r_{xx}[N-1] \\ r_{xx}[1] & r_{xx}[0] & \ldots & r_{xx}[N-2] \\ \vdots & \vdots & \ddots & \vdots \\ r_{xx}[N-1] & r_{xx}[N-2] & \ldots & r_{xx}[0] \end{bmatrix} \\
&= \mathbf{R}_{xx}
\end{aligned}
$$

Autocorrelation matrix

(Toeplitz structure)

Chapter 12 – Linear Bayesian Estimation

Wiener filtering

Data: x[0], x[1],x[2],… WSS with zero mean

Covariance matrix

$$
\begin{aligned}
\mathbf{C}_{xx} &= \begin{bmatrix} r_{xx}[0] & r_{xx}[1] & \dots & r_{xx}[N-1] \\ r_{xx}[1] & r_{xx}[0] & \dots & r_{xx}[N-2] \\ \vdots & \vdots & \ddots & \vdots \\ r_{xx}[N-1] & r_{xx}[N-2] & \dots & r_{xx}[0] \end{bmatrix} \\
&= \mathbf{R}_{xx}
\end{aligned}
$$

Parameter to be estimated: s[n], zero mean

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

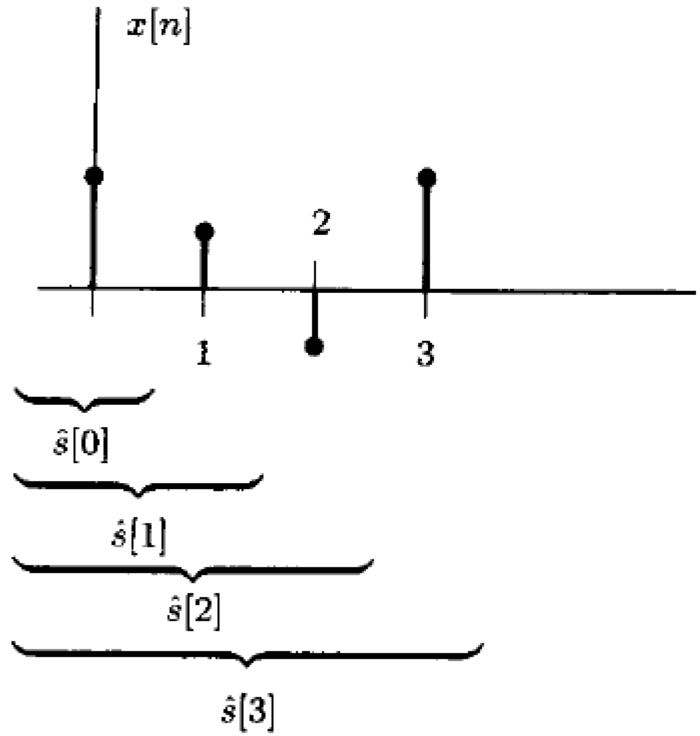
Chapter 12 – Linear Bayesian Estimation

Wiener filtering

Four cases

Filtering

Find s[n] given x[0],…,x[n]



(a) Filtering

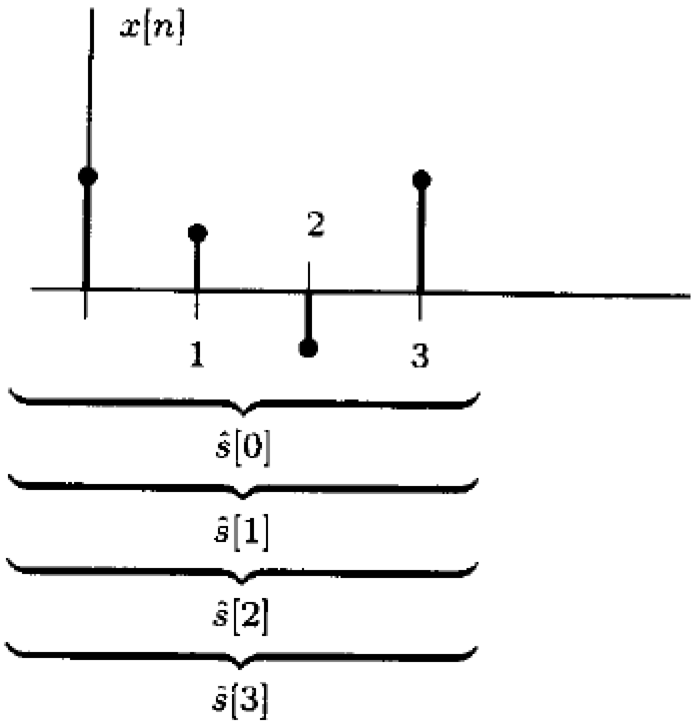
Chapter 12 – Linear Bayesian Estimation

Wiener filtering

Four cases

Smoothing

Find s[n] given x[0],…,x[N]



(b) Smoothing

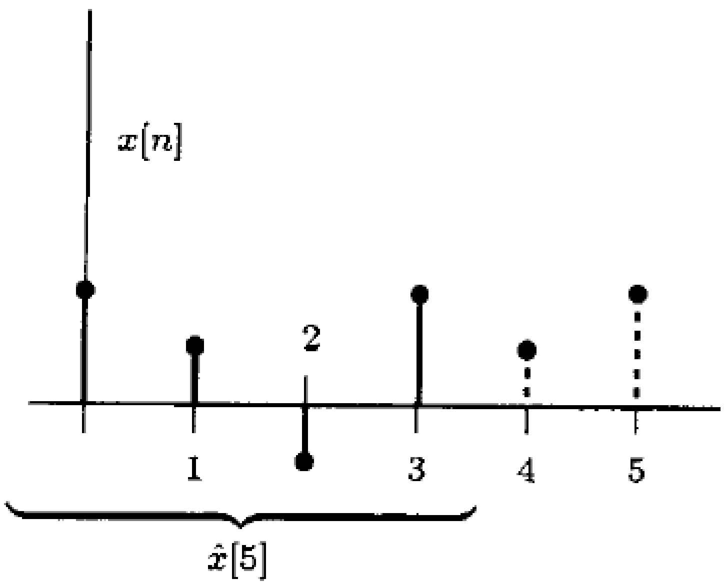
Chapter 12 – Linear Bayesian Estimation

Wiener filtering

Four cases

Prediction

Find s[n-1+L] given x[0],…,x[n-1]



(c) Prediction

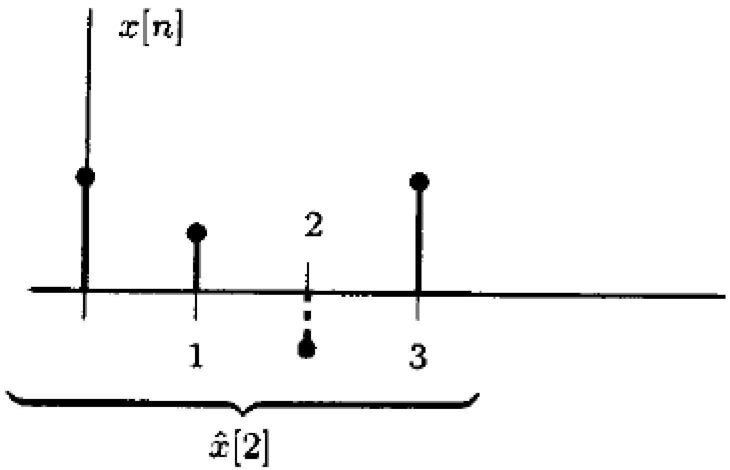
Chapter 12 – Linear Bayesian Estimation

Wiener filtering

Four cases

Interpolation

Find x[n] given x[0],…,x[n-1],x[n+1],…



(d) Interpolation

Chapter 12 – Linear Bayesian Estimation

Wiener filtering

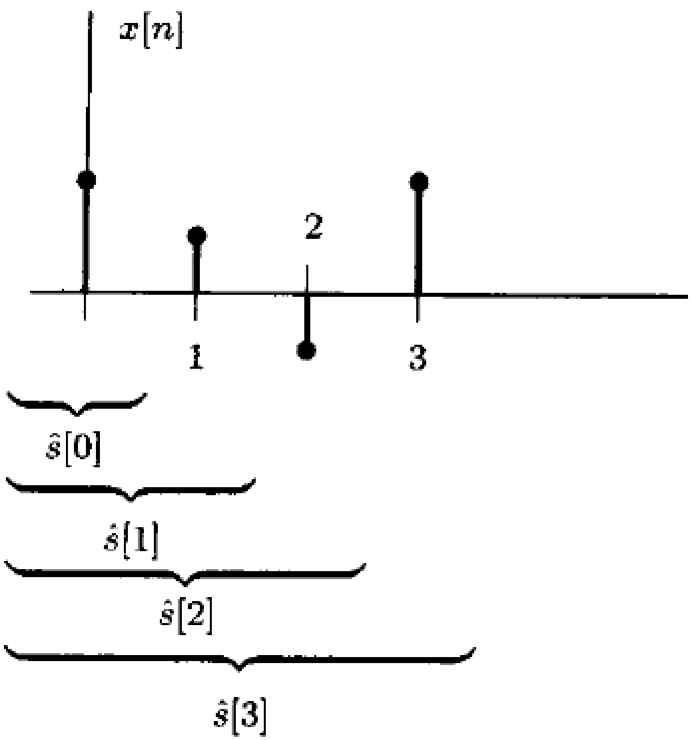
Filtering problem

Signal and noise are uncorrelated

$$\mathbf{C}_{xx} = \mathbf{R}_{ss} + \mathbf{R}_{ww}$$

Since the means are zero, we know that

$$\hat{\boldsymbol{\theta}} = \mathbf{C}_{\theta x}\mathbf{C}_{xx}^{-1}\mathbf{x}$$



(a) Filtering

Chapter 12 – Linear Bayesian Estimation

Wiener filtering

Filtering problem

Signal and noise are uncorrelated

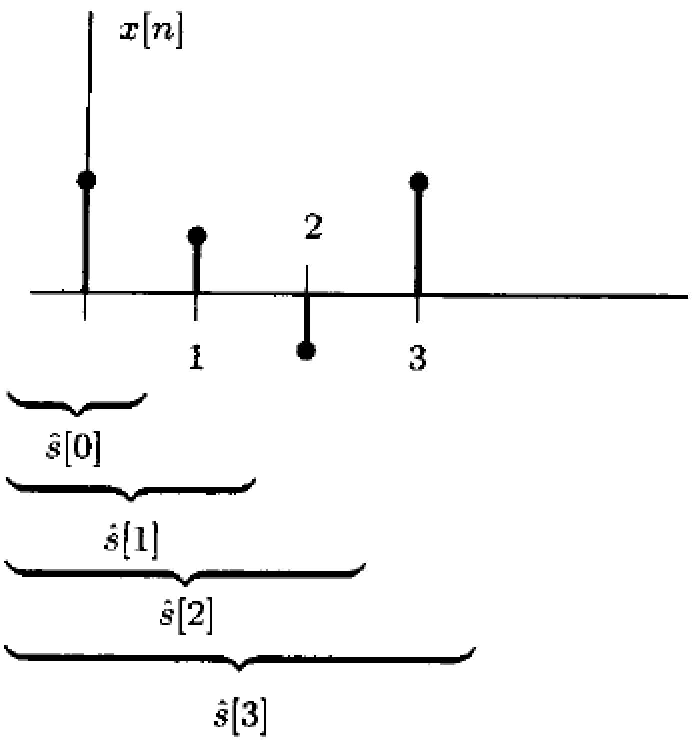
$$\mathbf{C}_{xx} = \mathbf{R}_{ss} + \mathbf{R}_{ww}$$

Since the means are zero, we know that

$$\hat{\boldsymbol{\theta}} = \mathbf{C}_{\theta x}\mathbf{C}_{xx}^{-1}\mathbf{x}$$

We have

$$\mathbf{C}_{\theta x} = E\left(s[n]\begin{bmatrix} x[0] & x[1] & \dots & x[n] \end{bmatrix}\right)$$



(a) Filtering

Chapter 12 – Linear Bayesian Estimation

Wiener filtering

Filtering problem

Signal and noise are uncorrelated

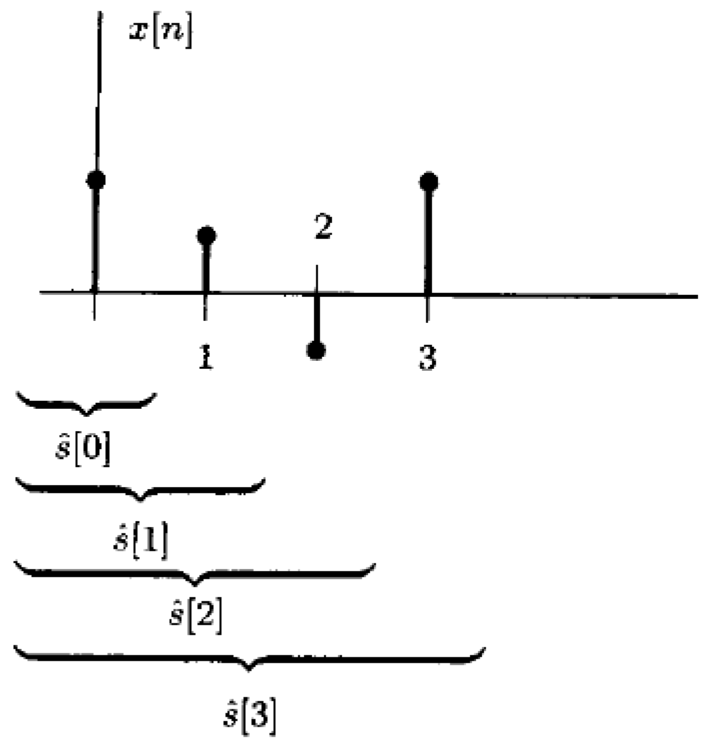
$$\mathbf{C}_{xx} = \mathbf{R}_{ss} + \mathbf{R}_{ww}$$

Since the means are zero, we know that

$$\hat{\boldsymbol{\theta}} = \mathbf{C}_{\theta x}\mathbf{C}_{xx}^{-1}\mathbf{x}$$

We have

$$\begin{aligned}\mathbf{C}_{\theta x} &= E\left(s[n]\begin{bmatrix} x[0] & x[1] & \ldots & x[n] \end{bmatrix}\right)\\ &= E\left(s[n]\begin{bmatrix} s[0] & s[1] & \ldots & s[n] \end{bmatrix}\right)\end{aligned}$$



(a) Filtering

Chapter 12 – Linear Bayesian Estimation

Wiener filtering

Filtering problem

Signal and noise are uncorrelated

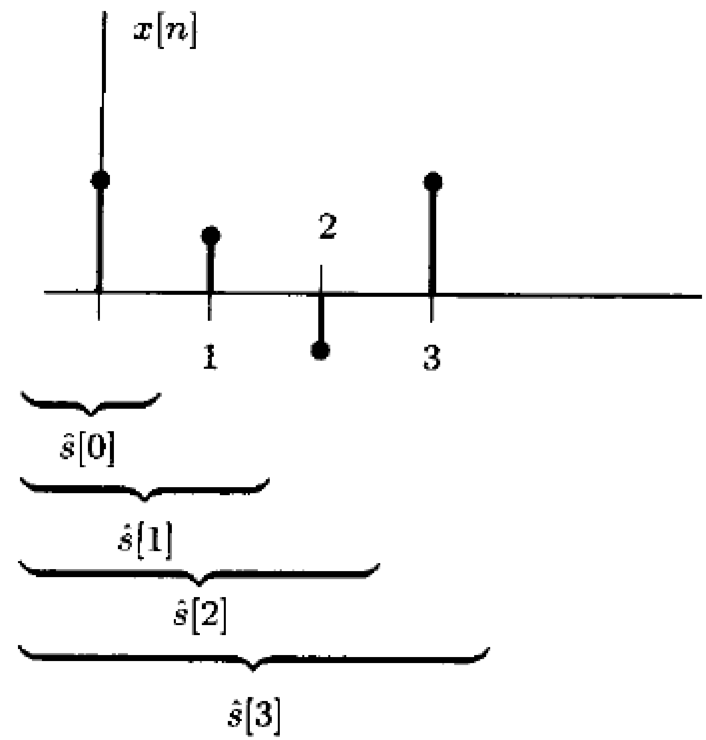
$$\mathbf{C}_{xx} = \mathbf{R}_{ss} + \mathbf{R}_{ww}$$

Since the means are zero, we know that

$$\hat{\boldsymbol{\theta}} = \mathbf{C}_{\theta x}\mathbf{C}_{xx}^{-1}\mathbf{x}$$

We have

$$\begin{aligned}
\mathbf{C}_{\theta x} &= E\left(s[n]\begin{bmatrix} x[0] & x[1] & \dots & x[n] \end{bmatrix}\right) \\
&= E\left(s[n]\begin{bmatrix} s[0] & s[1] & \dots & s[n] \end{bmatrix}\right) \\
&= [r_{ss}[n]\, r_{ss}[n-1] \dots r_{ss}[0]] = \mathbf{r}'^{T}_{ss}
\end{aligned}$$



(a) Filtering

Chapter 12 – Linear Bayesian Estimation

Wiener filtering

Filtering problem

Signal and noise are uncorrelated

$$\mathbf{C}_{xx} = \mathbf{R}_{ss} + \mathbf{R}_{ww}$$

Since the means are zero, we know that

$$\hat{\theta} = \mathbf{C}_{\theta x}\mathbf{C}_{xx}^{-1}\mathbf{x}$$

We have

$$\begin{aligned} \mathbf{C}_{\theta x} &= E\left(s[n]\begin{bmatrix} x[0] & x[1] & \ldots & x[n] \end{bmatrix}\right) \\ &= E\left(s[n]\begin{bmatrix} s[0] & s[1] & \ldots & s[n] \end{bmatrix}\right) \\ &= [r_{ss}[n]\, r_{ss}[n-1] \ldots r_{ss}[0]] = \mathbf{r}_{ss}'^{T} \end{aligned}$$

So,

$$\hat{s}[n] = \mathbf{r}_{ss}'^{T}\left(\mathbf{R}_{ss} + \mathbf{R}_{ww}\right)^{-1}\mathbf{x}.$$

Chapter 12 – Linear Bayesian Estimation

Wiener filtering

Filtering problem

Signal and noise are uncorrelated

$$\mathbf{C}_{xx} = \mathbf{R}_{ss} + \mathbf{R}_{ww}$$

Since the means are zero, we know that

$$\hat{\theta} = \mathbf{C}_{\theta x}\mathbf{C}_{xx}^{-1}\mathbf{x}$$

We have

$$\begin{aligned}
\mathbf{C}_{\theta x} &= E\left(s[n]\begin{bmatrix} x[0] & x[1] & \dots & x[n] \end{bmatrix}\right) \\
&= E\left(s[n]\begin{bmatrix} s[0] & s[1] & \dots & s[n] \end{bmatrix}\right) \\
&= [r_{ss}[n]\, r_{ss}[n-1] \dots r_{ss}[0]] = \mathbf{r}'^{T}_{ss}
\end{aligned}$$

So,

$$\hat{s}[n] = \mathbf{r}'^{T}_{ss}\left(\mathbf{R}_{ss} + \mathbf{R}_{ww}\right)^{-1}\mathbf{x}.$$

With

$$\mathbf{a} = \left(\mathbf{R}_{ss} + \mathbf{R}_{ww}\right)^{-1}\mathbf{r}'_{ss}$$

We get

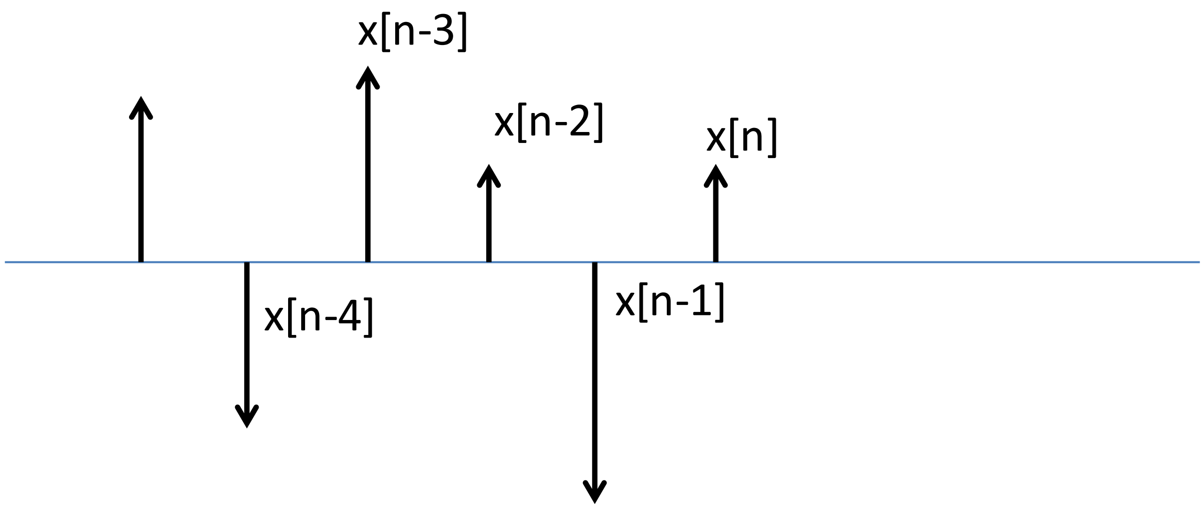
$$\hat{s}[n] = \mathbf{a}^{T}\mathbf{x}$$

Chapter 12 – Linear Bayesian Estimation

Wiener filtering

Why is it a filtering?

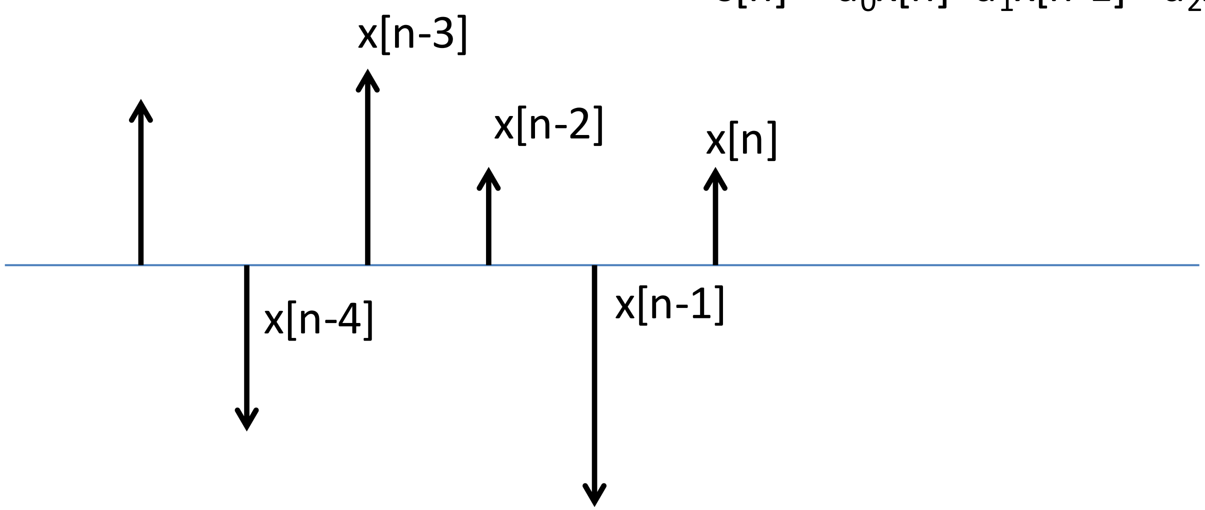
Find s[n]



Chapter 12 – Linear Bayesian Estimation

Wiener filtering

Why is it a filtering?



Find s[n]

We do this as a weighted sum

$s[n] = a_0x[n]+a_1x[n-1]+ a_2x[n-2]$

Chapter 12 – Linear Bayesian Estimation

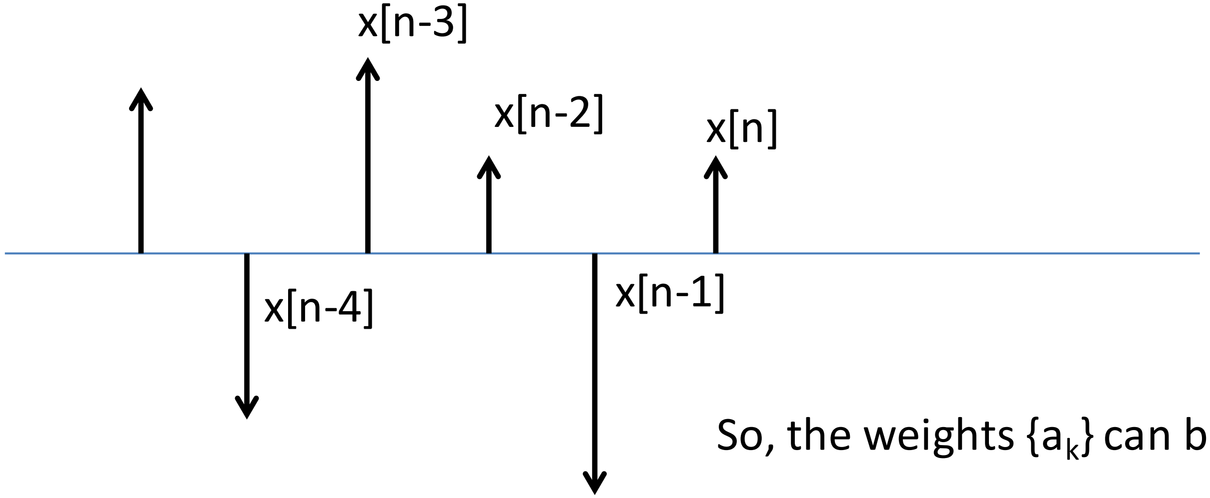
Wiener filtering

Why is it a filtering?

Find s[n]

We do this as a weighted sum

$s[n] = a_0x[n]+a_1x[n-1]+ a_2x[n-2]$



So, the weights $\{a_k\}$ can be seen as a FIR filter

Chapter 12 – Linear Bayesian Estimation

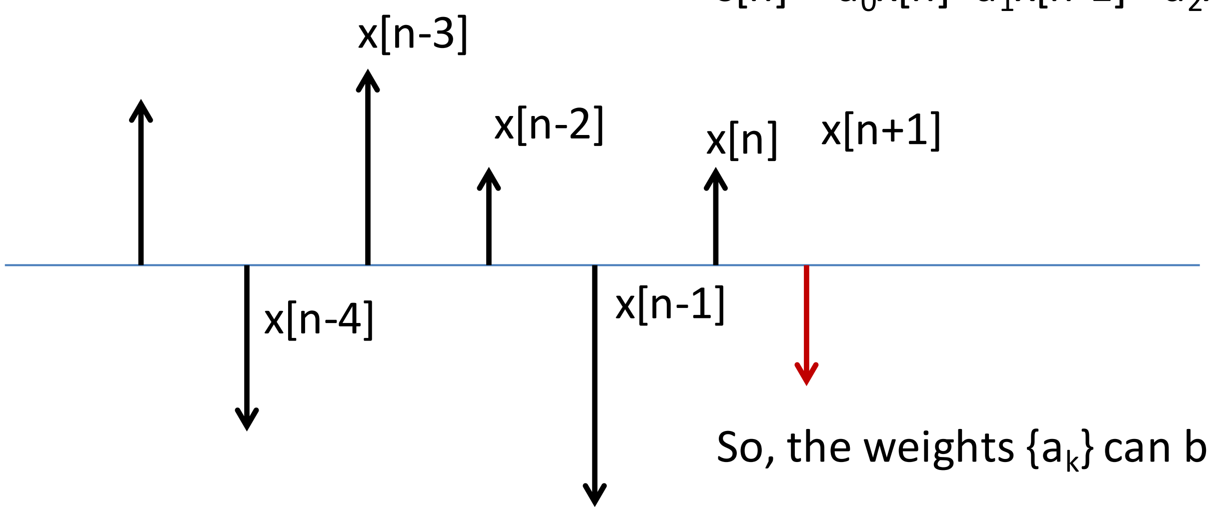
Wiener filtering

Why is it a filtering?

Find s[n]

We do this as a weighted sum

$s[n] = a_0x[n]+a_1x[n-1]+ a_2x[n-2]$



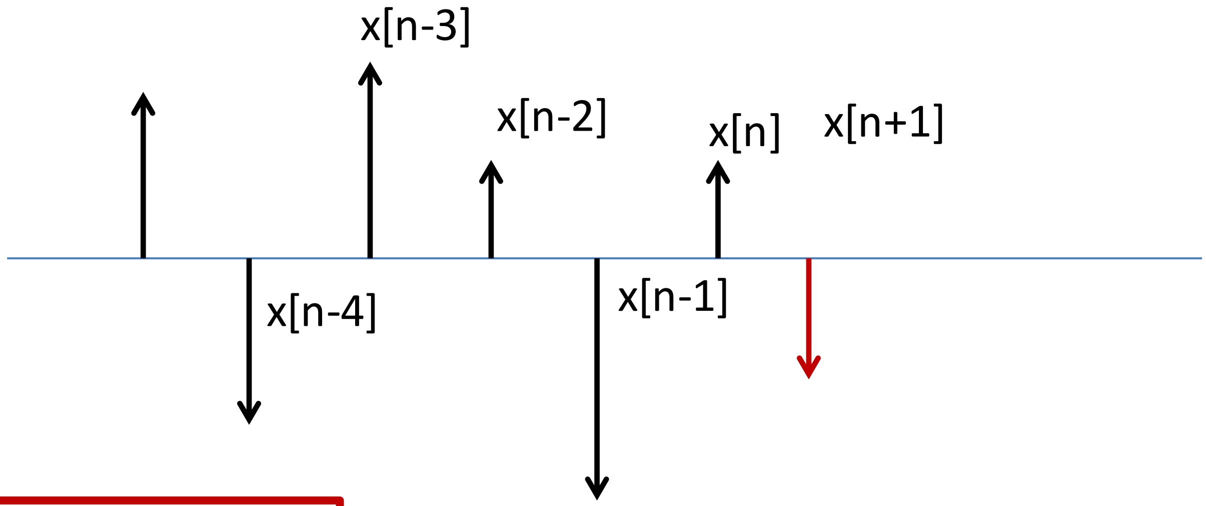
So, the weights $\{a_k\}$ can be seen as a FIR filter

However, at the next time, the weights $\{a_k\}$ are not the same (edge effect)

Chapter 12 – Linear Bayesian Estimation

Wiener filtering

Why is it a filtering?



To estimate s[n], we filter the recent observations with a filter that is dependent on n

The filter is time-variant

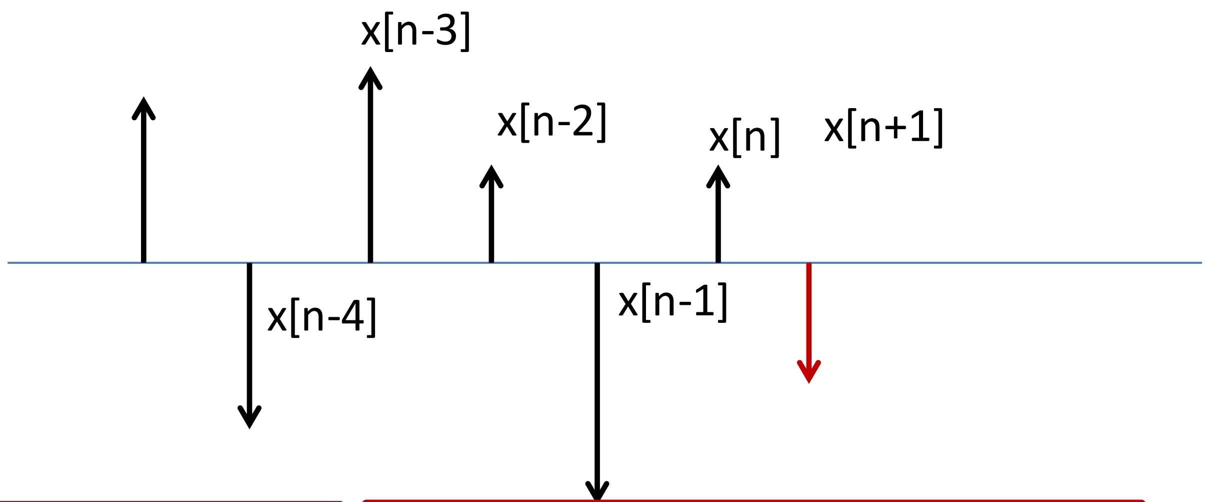
$$h^{(n)}[k] = a_{n-k} \qquad k = 0, 1, \ldots, n$$

a is computed for a given n (not shown explicitly)

Chapter 12 – Linear Bayesian Estimation

Wiener filtering

Why is it a filtering?



To estimate s[n], we filter the recent observations with a filter that is dependent on n

The filter is time-variant

$$h^{(n)}[k] = a_{n-k} \qquad k = 0, 1, \ldots, n$$

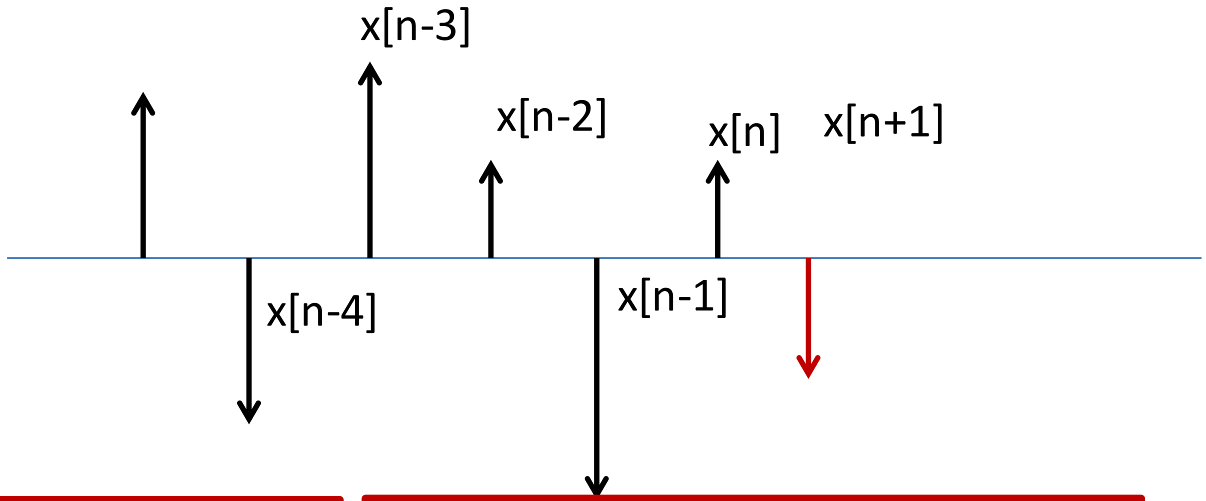
We get,

$$\hat{s}[n] = \sum_{k=0}^{n} a_k x[k]$$

Chapter 12 – Linear Bayesian Estimation

Wiener filtering

Why is it a filtering?



To estimate s[n], we filter the recent observations with a filter that is dependent on n

The filter is time-variant

$$h^{(n)}[k] = a_{n-k} \qquad k = 0, 1, \ldots, n$$

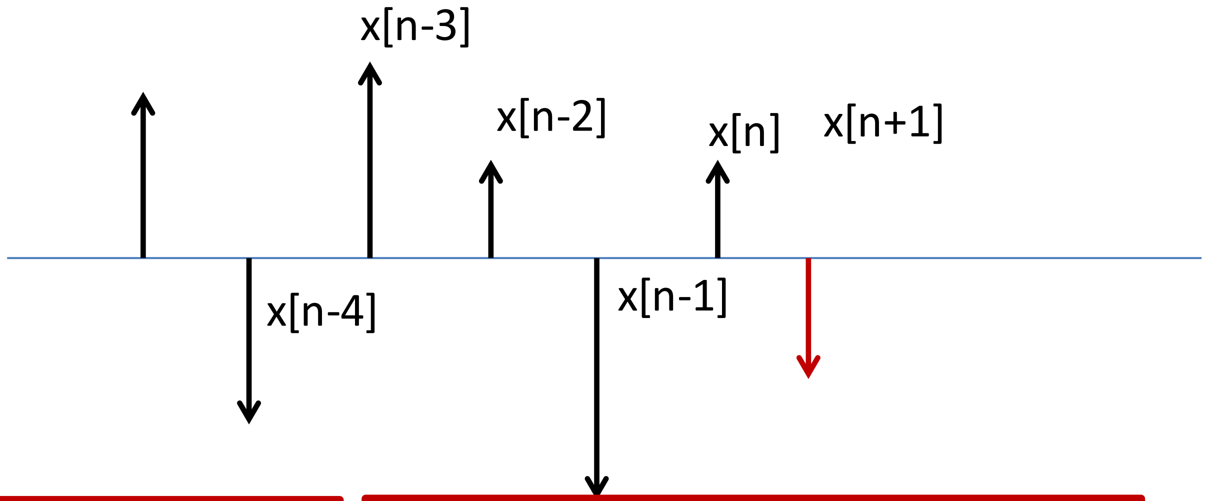
We get,

$$\begin{aligned}\hat{s}[n] &= \sum_{k=0}^{n} a_k x[k] \\ &= \sum_{k=0}^{n} h^{(n)}[n-k]x[k]\end{aligned}$$

Chapter 12 – Linear Bayesian Estimation

Wiener filtering

Why is it a filtering?



To estimate s[n], we filter the recent observations with a filter that is dependent on n

The filter is time-variant

$$h^{(n)}[k] = a_{n-k} \qquad k = 0, 1, \ldots, n$$

We get,

$$\begin{aligned} \hat{s}[n] &= \sum_{k=0}^{n} a_k x[k] \\ &= \sum_{k=0}^{n} h^{(n)}[n-k] x[k] \end{aligned}$$

$$\hat{s}[n] = \sum_{k=0}^{n} h^{(n)}[k] x[n-k]$$

Chapter 12 – Linear Bayesian Estimation

Wiener filtering

Observe

$\mathbf{h} = [h^{(n)}[0]\, h^{(n)}[1] \ldots h^{(n)}[n]]^T$ is **a** but flipped upside-down

$\mathbf{a} = [a_0\, a_1 \ldots a_n]^T$

To estimate s[n], we filter the recent observations with a filter that is dependent on n

The filter is time-variant

$h^{(n)}[k] = a_{n-k} \qquad k = 0, 1, \ldots, n$

We get,

$$\begin{aligned} \hat{s}[n] &= \sum_{k=0}^{n} a_k x[k] \\ &= \sum_{k=0}^{n} h^{(n)}[n-k] x[k] \end{aligned}$$

$$\hat{s}[n] = \sum_{k=0}^{n} h^{(n)}[k] x[n-k]$$

Chapter 12 – Linear Bayesian Estimation

Wiener filtering

Observe

$$\mathbf{h} = [h^{(n)}[0]\, h^{(n)}[1] \ldots h^{(n)}[n]]^T$$

is **a** but flipped upside-down

$$\mathbf{a} = [a_0\, a_1 \ldots a_n]^T$$

Recall

$$\mathbf{a} = (\mathbf{R}_{ss} + \mathbf{R}_{ww})^{-1}\, \mathbf{r}'_{ss}$$

$$\mathbf{r}'^{T}_{ss} = [r_{ss}[n]\, r_{ss}[n-1] \ldots r_{ss}[0]]$$

We get,

$$\begin{aligned} \hat{s}[n] &= \sum_{k=0}^{n} a_k x[k] \\ &= \sum_{k=0}^{n} h^{(n)}[n-k] x[k] \end{aligned}$$

$$\hat{s}[n] = \sum_{k=0}^{n} h^{(n)}[k] x[n-k]$$

Chapter 12 – Linear Bayesian Estimation

Wiener filtering

Observe

$$\mathbf{h} = [h^{(n)}[0]\, h^{(n)}[1] \ldots h^{(n)}[n]]^T$$

is **a** but flipped upside-down

$$\mathbf{a} = [a_0\, a_1 \ldots a_n]^T$$

Recall

$$\mathbf{a} = (\mathbf{R}_{ss} + \mathbf{R}_{ww})^{-1}\, \mathbf{r}'_{ss}$$

$$\mathbf{r}'^T_{ss} = [r_{ss}[n]\, r_{ss}[n-1] \ldots r_{ss}[0]]$$

$$(\mathbf{R}_{ss} + \mathbf{R}_{ww})\,\mathbf{a} = \mathbf{r}'_{ss}$$

We get,

$$\begin{aligned} \hat{s}[n] &= \sum_{k=0}^{n} a_k x[k] \\ &= \sum_{k=0}^{n} h^{(n)}[n-k] x[k] \end{aligned}$$

$$\hat{s}[n] = \sum_{k=0}^{n} h^{(n)}[k] x[n-k]$$

Chapter 12 – Linear Bayesian Estimation

Wiener filtering

Observe

$\mathbf{h} = [h^{(n)}[0]\, h^{(n)}[1] \ldots h^{(n)}[n]]^T$ is **a** but flipped upside-down

$\mathbf{a} = [a_0\, a_1 \ldots a_n]^T$

Recall

$$\mathbf{a} = (\mathbf{R}_{ss} + \mathbf{R}_{ww})^{-1}\, \mathbf{r}'_{ss}$$

$$\mathbf{r}'^{T}_{ss} = [r_{ss}[n]\, r_{ss}[n-1] \ldots r_{ss}[0]]$$

$$(\mathbf{R}_{ss} + \mathbf{R}_{ww})\, \mathbf{a} = \mathbf{r}'_{ss}$$

$$(\mathbf{R}_{ss} + \mathbf{R}_{ww})\, \mathbf{h} = \mathbf{r}_{ss}$$

$$\mathbf{r}_{ss} = [r_{ss}[0]\, r_{ss}[1] \ldots r_{ss}[n]]^T$$

We get,

$$\hat{s}[n] = \sum_{k=0}^{n} a_k x[k] = \sum_{k=0}^{n} h^{(n)}[n-k]x[k]$$

$$\hat{s}[n] = \sum_{k=0}^{n} h^{(n)}[k]x[n-k]$$

Chapter 12 – Linear Bayesian Estimation

Wiener filtering

$$(\mathbf{R}_{ss} + \mathbf{R}_{ww})\,\mathbf{h} = \mathbf{r}_{ss}$$

$$\mathbf{r}_{ss} = [r_{ss}[0]\; r_{ss}[1] \ldots r_{ss}[n]]^T$$

Chapter 12 – Linear Bayesian Estimation

Wiener filtering

Wiener-Hopf equations

$$\begin{bmatrix} r_{xx}[0] & r_{xx}[1] & \dots & r_{xx}[n] \\ r_{xx}[1] & r_{xx}[0] & \dots & r_{xx}[n-1] \\ \vdots & \vdots & \ddots & \vdots \\ r_{xx}[n] & r_{xx}[n-1] & \dots & r_{xx}[0] \end{bmatrix} \begin{bmatrix} h^{(n)}[0] \\ h^{(n)}[1] \\ \vdots \\ h^{(n)}[n] \end{bmatrix} = \begin{bmatrix} r_{ss}[0] \\ r_{ss}[1] \\ \vdots \\ r_{ss}[n] \end{bmatrix}$$

$$(\mathbf{R}_{ss} + \mathbf{R}_{ww})\,\mathbf{h} = \mathbf{r}_{ss}$$

$$\mathbf{r}_{ss} = [r_{ss}[0]\; r_{ss}[1] \dots r_{ss}[n]]^T$$

Chapter 12 – Linear Bayesian Estimation

Wiener filtering

Wiener-Hopf equations

$$\begin{bmatrix} r_{xx}[0] & r_{xx}[1] & \ldots & r_{xx}[n] \\ r_{xx}[1] & r_{xx}[0] & \ldots & r_{xx}[n-1] \\ \vdots & \vdots & \ddots & \vdots \\ r_{xx}[n] & r_{xx}[n-1] & \ldots & r_{xx}[0] \end{bmatrix} \begin{bmatrix} h^{(n)}[0] \\ h^{(n)}[1] \\ \vdots \\ h^{(n)}[n] \end{bmatrix} = \begin{bmatrix} r_{ss}[0] \\ r_{ss}[1] \\ \vdots \\ r_{ss}[n] \end{bmatrix}$$

These equations can be solved recursively by the Levinson algorithm

Observe: The matrix is Toeplitz, but cannot be approximated as circulant as n grows. Therefore, Szegö theory does not apply.

Chapter 12 – Linear Bayesian Estimation

Wiener filtering

Wiener-Hopf equations

$$\begin{bmatrix} r_{xx}[0] & r_{xx}[1] & \dots & r_{xx}[n] \\ r_{xx}[1] & r_{xx}[0] & \dots & r_{xx}[n-1] \\ \vdots & \vdots & \ddots & \vdots \\ r_{xx}[n] & r_{xx}[n-1] & \dots & r_{xx}[0] \end{bmatrix} \begin{bmatrix} h^{(n)}[0] \\ h^{(n)}[1] \\ \vdots \\ h^{(n)}[n] \end{bmatrix} = \begin{bmatrix} r_{ss}[0] \\ r_{ss}[1] \\ \vdots \\ r_{ss}[n] \end{bmatrix}$$

$$\sum_{k=0}^{n} h^{(n)}[k] r_{xx}[l-k] = r_{ss}[l] \qquad l = 0, 1, \dots, n$$

$$r_{xx}[-k] = r_{xx}[k]$$

Chapter 12 – Linear Bayesian Estimation

Wiener filtering

Wiener-Hopf equations

$$\begin{bmatrix} r_{xx}[0] & r_{xx}[1] & \cdots & r_{xx}[n] \\ r_{xx}[1] & r_{xx}[0] & \cdots & r_{xx}[n-1] \\ \vdots & \vdots & \ddots & \vdots \\ r_{xx}[n] & r_{xx}[n-1] & \cdots & r_{xx}[0] \end{bmatrix} \begin{bmatrix} h^{(n)}[0] \\ h^{(n)}[1] \\ \vdots \\ h^{(n)}[n] \end{bmatrix} = \begin{bmatrix} r_{ss}[0] \\ r_{ss}[1] \\ \vdots \\ r_{ss}[n] \end{bmatrix}$$

$$\sum_{k=0}^{n} h^{(n)}[k] r_{xx}[l-k] = r_{ss}[l] \qquad l = 0, 1, \ldots, n$$

$$r_{xx}[-k] = r_{xx}[k]$$

As n grows, the filter converges to a stationary solution

$$\sum_{k=0}^{\infty} h[k] r_{xx}[l-k] = r_{ss}[l] \qquad l = 0, 1, \ldots .$$

Chapter 12 – Linear Bayesian Estimation

Wiener filtering

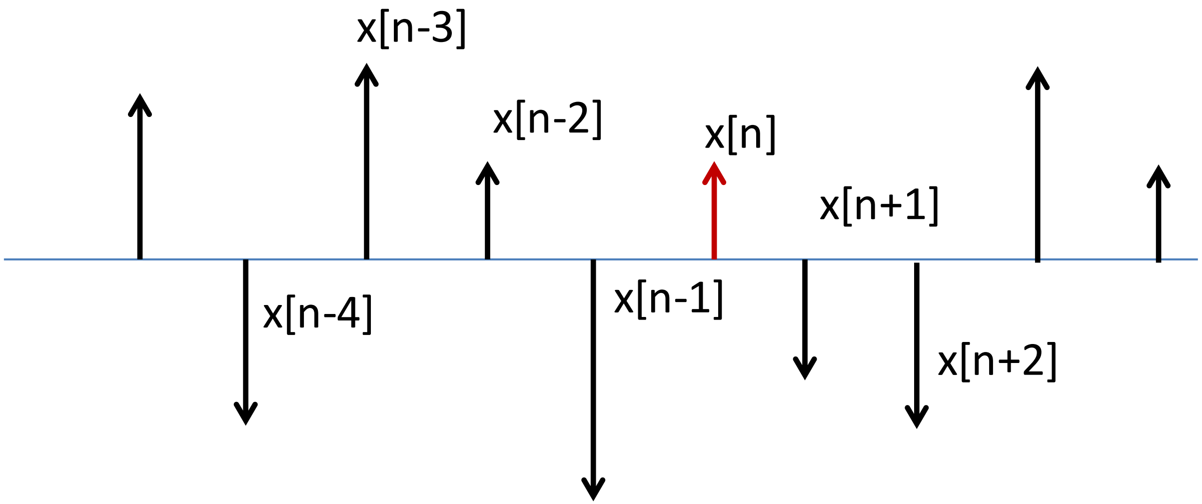
To find h[n], we can apply **spectral factorization** (= same method as is used to find a minimum phase version of a filter)

$$\sum_{k=0}^{\infty} h[k] r_{xx}[l-k] = r_{ss}[l] \qquad l = 0, 1, \ldots \, .$$

Chapter 12 – Linear Bayesian Estimation

Wiener smoothing

Now consider asymptotic Wiener smoothing

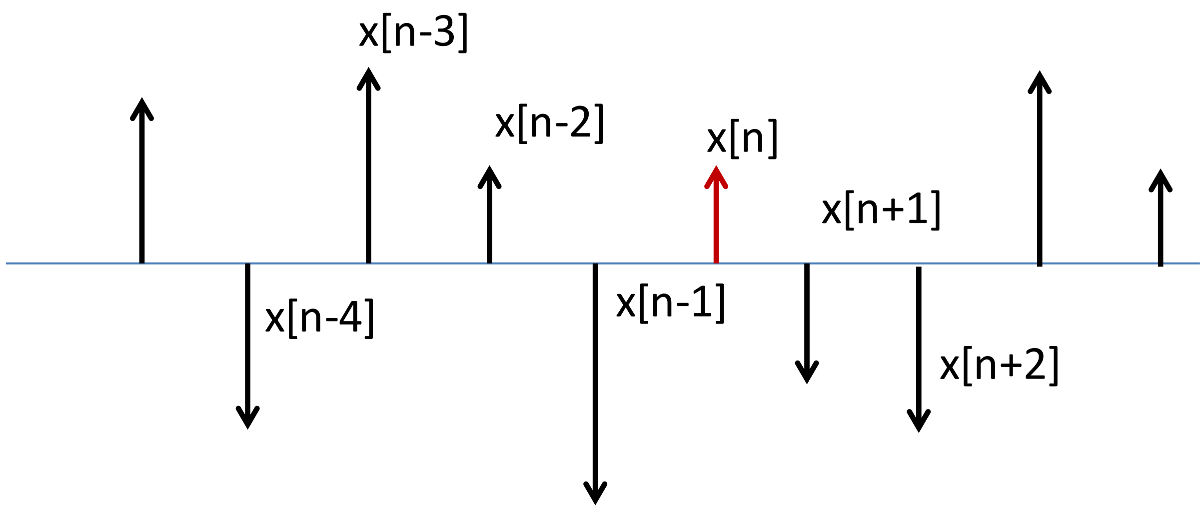


We can still express the estimation of s[n] as a filtering of {x[k]}

Chapter 12 – Linear Bayesian Estimation

Wiener smoothing

Now consider asymptotic Wiener smoothing



We can still express the estimation of s[n] as a filtering of {x[k]}
The filter is ***not*** causal

$$\hat{s}[n] = \sum_{k=-\infty}^{\infty} h[k]x[n-k]$$

Chapter 12 – Linear Bayesian Estimation

Wiener smoothing

Filtering

$$\hat{s}[n] = \sum_{k=0}^{\infty} h[k]x[n-k]$$

$$\sum_{k=0}^{\infty} h[k]r_{xx}[l-k] = r_{ss}[l] \qquad l = 0, 1, \ldots \,.$$

Smoothing

$$\hat{s}[n] = \sum_{k=-\infty}^{\infty} h[k]x[n-k]$$

???????

Chapter 12 – Linear Bayesian Estimation

Wiener smoothing

Filtering

$$\hat{s}[n] = \sum_{k=0}^{\infty} h[k]x[n-k]$$

$$\sum_{k=0}^{\infty} h[k]r_{xx}[l-k] = r_{ss}[l] \qquad l = 0, 1, \ldots .$$

Smoothing

$$\hat{s}[n] = \sum_{k=-\infty}^{\infty} h[k]x[n-k]$$

$$\sum_{k=-\infty}^{\infty} h[k]r_{xx}[l-k] = r_{ss}[k] \qquad -\infty < l < \infty$$

(12.61): Typo in book *l*

Chapter 12 – Linear Bayesian Estimation

Wiener smoothing

$$\sum_{k=-\infty}^{\infty} h[k] r_{xx}[l-k] = r_{ss}[k] \qquad -\infty < l < \infty$$

No edge effect in the smoothing setup!

Can be solved by approximating R_{xx} as a circulant matrix (Szegö theory)

Chapter 12 – Linear Bayesian Estimation

Wiener smoothing

$$\sum_{k=-\infty}^{\infty} h[k]r_{xx}[l-k] = r_{ss}[k] \qquad -\infty < l < \infty$$

No edge effect in the smoothing setup!

Can be solved by approximating R_{xx} as a circulant matrix (Szegö theory)

$$h[n] \star r_{xx}[n] = r_{ss}[n]$$

$$\begin{aligned} H(f) &= \frac{P_{ss}(f)}{P_{xx}(f)} \\ &= \frac{P_{ss}(f)}{P_{ss}(f) + P_{ww}(f)}. \end{aligned}$$

Chapter 12 – Linear Bayesian Estimation

Wiener prediction

Filtering equations

$$\begin{bmatrix} r_{xx}[0] & r_{xx}[1] & \dots & r_{xx}[n] \\ r_{xx}[1] & r_{xx}[0] & \dots & r_{xx}[n-1] \\ \vdots & \vdots & \ddots & \vdots \\ r_{xx}[n] & r_{xx}[n-1] & \dots & r_{xx}[0] \end{bmatrix} \begin{bmatrix} h^{(n)}[0] \\ h^{(n)}[1] \\ \vdots \\ h^{(n)}[n] \end{bmatrix} = \begin{bmatrix} r_{ss}[0] \\ r_{ss}[1] \\ \vdots \\ r_{ss}[n] \end{bmatrix}$$

Chapter 12 – Linear Bayesian Estimation

Wiener prediction

Filtering equations….Filtering "predicts" s[n] given x[n]

$$\begin{bmatrix} r_{xx}[0] & r_{xx}[1] & \dots & r_{xx}[n] \\ r_{xx}[1] & r_{xx}[0] & \dots & r_{xx}[n-1] \\ \vdots & \vdots & \ddots & \vdots \\ r_{xx}[n] & r_{xx}[n-1] & \dots & r_{xx}[0] \end{bmatrix} \begin{bmatrix} h^{(n)}[0] \\ h^{(n)}[1] \\ \vdots \\ h^{(n)}[n] \end{bmatrix} = \begin{bmatrix} r_{ss}[0] \\ r_{ss}[1] \\ \vdots \\ r_{ss}[n] \end{bmatrix}$$

Chapter 12 – Linear Bayesian Estimation

Wiener prediction

Filtering equations….Filtering ”predicts” s[n] given x[n]

$$\begin{bmatrix} r_{xx}[0] & r_{xx}[1] & \dots & r_{xx}[n] \\ r_{xx}[1] & r_{xx}[0] & \dots & r_{xx}[n-1] \\ \vdots & \vdots & \ddots & \vdots \\ r_{xx}[n] & r_{xx}[n-1] & \dots & r_{xx}[0] \end{bmatrix} \begin{bmatrix} h^{(n)}[0] \\ h^{(n)}[1] \\ \vdots \\ h^{(n)}[n] \end{bmatrix} = \begin{bmatrix} r_{ss}[0] \\ r_{ss}[1] \\ \vdots \\ r_{ss}[n] \end{bmatrix}$$

In prediction x[n] is not available, therefore there is no h[0] coefficient.

Chapter 12 – Linear Bayesian Estimation

Wiener prediction

Filtering equations….Filtering ”predicts” s[n] given x[n]

$$\begin{bmatrix} r_{xx}[0] & r_{xx}[1] & \dots & r_{xx}[N-1] \\ r_{xx}[1] & r_{xx}[0] & \dots & r_{xx}[N-2] \\ \vdots & \vdots & \ddots & \vdots \\ r_{xx}[N-1] & r_{xx}[N-2] & \dots & r_{xx}[0] \end{bmatrix} \begin{bmatrix} h[1] \\ h[2] \\ \vdots \\ h[N] \end{bmatrix}$$

In prediction x[n] is not available, therefore there is no h[0] coefficient.

Chapter 12 – Linear Bayesian Estimation

Wiener prediction

Filtering equations....Filtering ”predicts” s[n] given x[n]

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

In prediction x[n] is not available, therefore there is no h[0] coefficient.

We are also predicting l steps into the future

Chapter 12 – Linear Bayesian Estimation

Wiener prediction

Filtering equations….Filtering ”predicts” s[n] given x[n]

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

In prediction x[n] is not available, therefore there is no h[0] coefficient.

We are also predicting l steps into the future

Wiener-Hopf prediction equations. For l=1 we obtain the Yule-Walker equations
Solved by Levinson recursion or spectral factorization (not Szegö Theory)