ECE 6540, Lecture 14

Bayesian Estimation:

LMMSE Estimators and Scalar Kalman Filters

Last Time

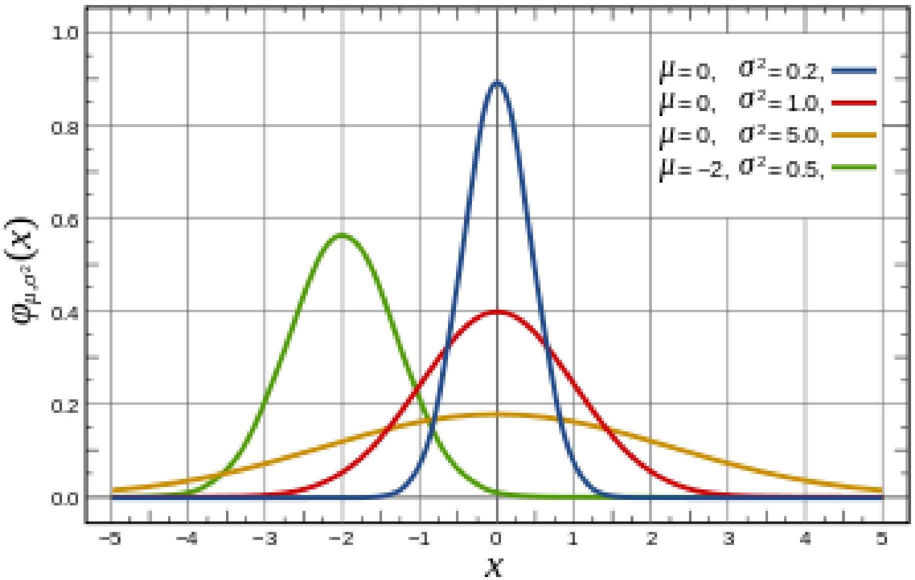
- **Bayes Risk (Average Loss)**
- **Generalized Bayesian Estimator**
- **MMSE Estimator**
- **Maximum A Posterior (MAP) Estimator**
- **Linear MMSE Estimator**

Example:

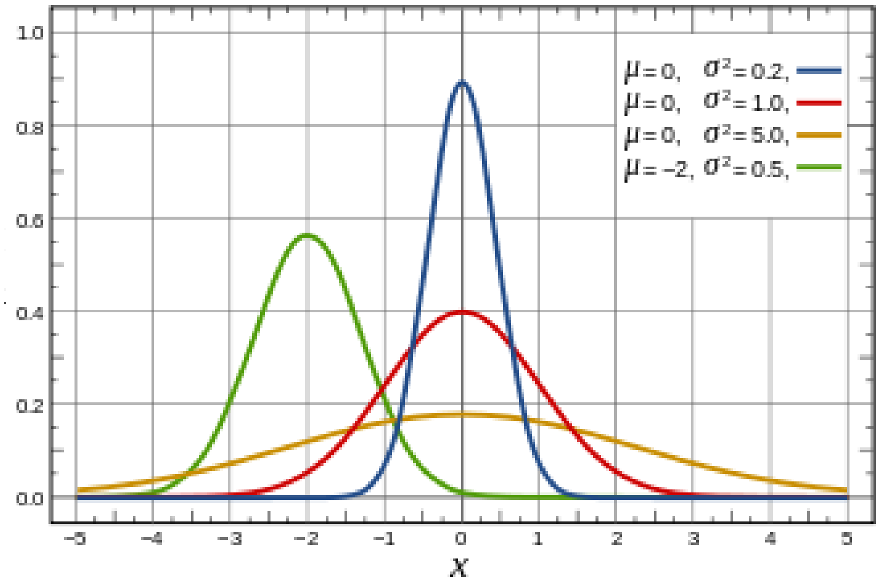
MAP Estimator

Last Time

- **MAP Estimator Example**
  - $\boldsymbol{x} = \boldsymbol{H\theta} + \boldsymbol{w}$
  - Gaussian likelihood $\boldsymbol{x}|\,\boldsymbol{\theta} \sim \mathcal{N}(\theta, \sigma^2\boldsymbol{I})$
  - Assume Gaussian prior $\boldsymbol{\theta} \sim \mathcal{N}\left(0, \sigma_\theta^2\boldsymbol{I}\right)$
  - **Why might we choose this prior?**



Likelihood



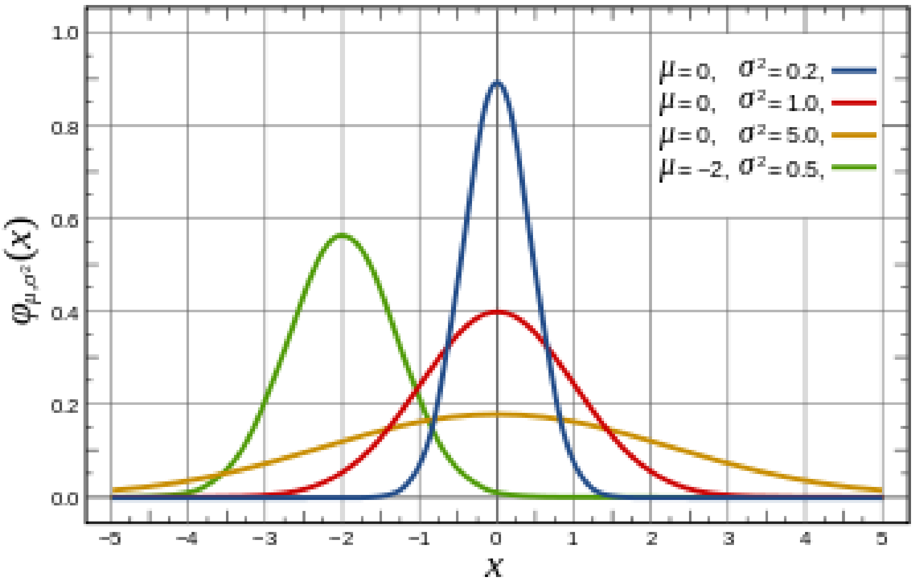
Prior

Source: https://en.wikipedia.org/wiki/Normal_distribution#/media/File:Normal_Distribution_PDF.svg

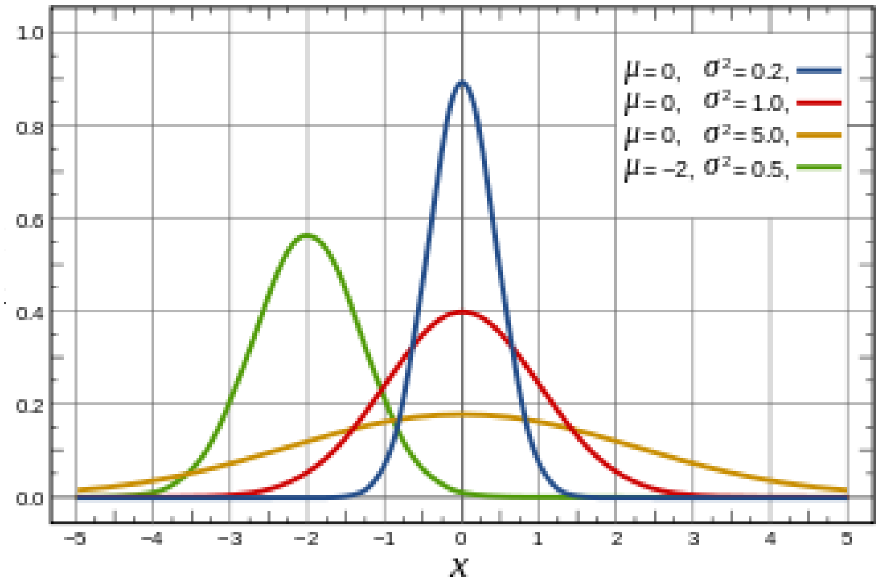
Last Time

- **MAP Estimator Example**
  - $\boldsymbol{x} = \boldsymbol{H\theta} + \boldsymbol{w}$
  - Gaussian likelihood $\boldsymbol{x}|\,\boldsymbol{\theta} \sim \mathcal{N}(\theta, \sigma^2 \boldsymbol{I})$
  - Assume Gaussian prior $\boldsymbol{\theta} \sim \mathcal{N}\left(0, \sigma_\theta^2 \boldsymbol{D}\right)$
  - **Why might we choose this prior?**

Difference Operator



Likelihood

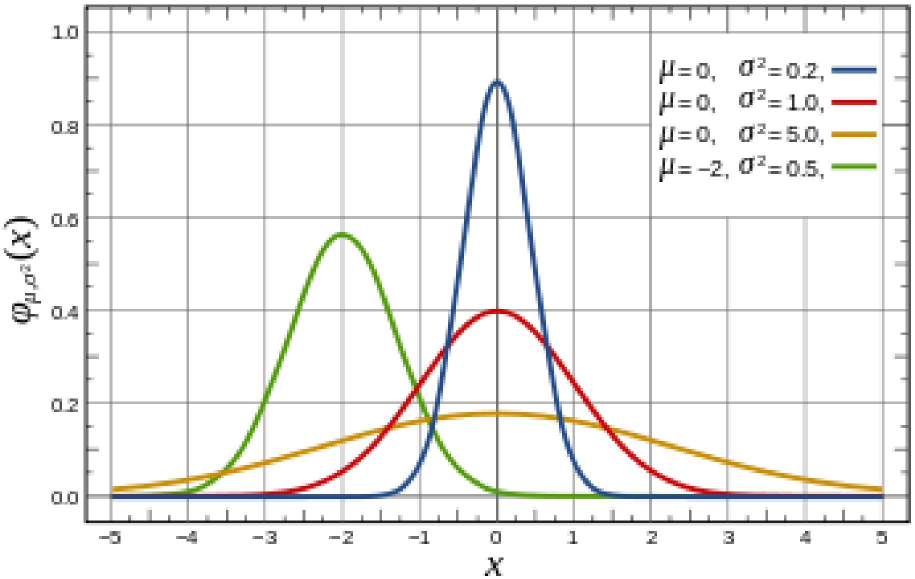


Prior

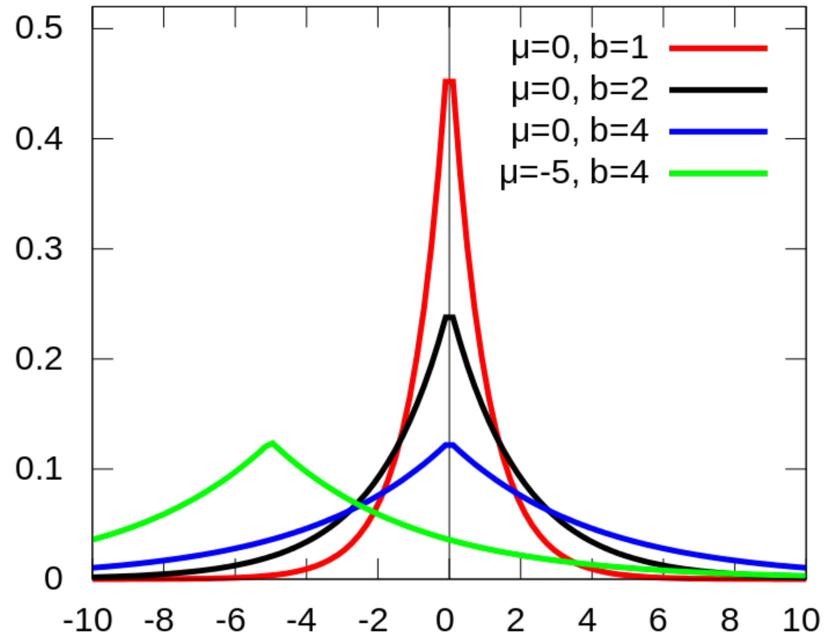
Source: https://en.wikipedia.org/wiki/Normal_distribution#/media/File:Normal_Distribution_PDF.svg

Last Time

- **MAP Estimator Example**
  - $\boldsymbol{x} = \boldsymbol{H\theta} + \boldsymbol{w}$
  - Gaussian likelihood $\boldsymbol{x}|\,\boldsymbol{\theta} \sim \mathcal{N}(\theta, \sigma^2 \boldsymbol{I})$
  - Assume Laplace prior $\theta_n \sim \text{Laplace}(0, b)$
  - **Why might we choose this prior?**



Likelihood



Prior

Source: https://en.wikipedia.org/wiki/Laplace_distribution#/media/File:Laplace_pdf_mod.svg

Source: https://en.wikipedia.org/wiki/Normal_distribution#/media/File:Normal_Distribution_PDF.svg

Last Time

- **MAP Estimator Example**
  - Gaussian likelihood $x|\theta \sim \mathcal{N}(0,\theta)$
  - Assume inverse gamma prior $\mathrm{p}(\theta) = \frac{\beta^{\alpha}}{\Gamma(\alpha)} \theta^{-\alpha-1} \exp\left(-\frac{\beta}{\theta}\right)$
  - **Determine the MAP Estimator for θ**

Last Time

- **MAP Estimator Example**
  - Gaussian likelihood $x|\,\theta \sim \mathcal{N}(0,\theta)$
  - Assume inverse gamma prior $\mathrm{p}(\theta) = \frac{\beta^\alpha}{\Gamma(\alpha)}\theta^{-\alpha-1}\exp\left(-\frac{\beta}{\theta}\right)$
  - **Determine the MAP Estimator for θ**

$$p(x|\theta)p(\theta) = \frac{1}{\sqrt{2\pi\theta}}\exp\left(\frac{-1}{2\theta}x^2\right)\frac{\beta^\alpha}{\Gamma(\alpha)}\theta^{-\alpha-1}\exp\left(-\frac{\beta}{\theta}\right)$$

$$= \frac{1}{\sqrt{2\pi\theta}}\frac{\beta^\alpha}{\Gamma(\alpha)}\theta^{-\alpha-1}\exp\left(-\frac{1}{\theta}\left(\beta + \frac{x^2}{2}\right)\right)$$

$$\ln\big(p(x|\theta)p(\theta)\big) = \frac{-1}{2}\ln(2\pi\theta) - (\alpha+1)\ln(\theta) - \frac{1}{\theta}\left(\beta + \frac{x^2}{2}\right) + \ln\left(\frac{\beta^\alpha}{\Gamma(\alpha)}\right)$$

$$\frac{\partial \ln\big(p(x|\theta)p(\theta)\big)}{\partial\theta} = \frac{-1}{2\theta} - (\alpha+1)\frac{1}{\theta} + \frac{1}{\theta^2}\left(\beta + \frac{x^2}{2}\right) = 0$$

Last Time

MAP Estimator Example

- Gaussian likelihood $x|\,\theta \sim \mathcal{N}(0,\theta)$
- Assume inverse gamma prior $\mathrm{p}(\theta) = \frac{\beta^{\alpha}}{\Gamma(\alpha)}\theta^{-\alpha-1}\exp\left(-\frac{\beta}{\theta}\right)$
- **Determine the MAP Estimator for θ**

$$\frac{1}{\theta^2}\left(\beta + \frac{x^2}{2}\right) = \frac{1}{\theta}\left((\alpha+1) + \frac{1}{2}\right)$$

$$\frac{1}{\theta^2}\left(\beta + \frac{x^2}{2}\right) = \frac{1}{\theta}\left(\alpha + \frac{3}{2}\right)$$

$$\hat{\theta} = \frac{\left(\beta + \frac{x^2}{2}\right)}{\left(\alpha + \frac{3}{2}\right)}$$

Bayesian Estimation:

Linear MMSE Estimators

Linear MMSE Estimator

- **A simplification**
  - Assume the estimator is a linear system
  - For a **single parameter** estimator

$$\hat{\theta} = \sum_{n=1}^{N} h_n x_n + a = \boldsymbol{h}^T \boldsymbol{x} + a$$

  - For a **multi-parameter** estimator

$$\widehat{\boldsymbol{\theta}} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{a}$$

Linear MMSE Estimator

- **A simplification**
  - Assume the estimator is a linear system
  - For a **single parameter** estimator

$$\hat{\theta} = \sum_{n=1}^{N} h_n x_n + a = \boldsymbol{h}^T \boldsymbol{x} + a$$

  - For a **multi-parameter** estimator

$$\widehat{\boldsymbol{\theta}} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{a}$$

Linear MMSE Estimator

- **<u>Goal:</u> To minimize the Bayesian Mean Square Error**
  - **Step one:** Find the optimal estimate for a

Linear MMSE Estimator

- **<u>Goal:</u> To minimize the Bayesian Mean Square Error**
  - **Step one:** Find the optimal estimate for a

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

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

$$\frac{\partial(BMSE)}{\partial a} = -2E[\theta - (\boldsymbol{h}^T\boldsymbol{x} + a)] = 0$$

$$a^* = E[\theta] - \boldsymbol{h}^T E[\boldsymbol{x}]$$

Linear MMSE Estimator

- **Goal: To minimize the Bayesian Mean Square Error**
  - **Step two:** Find the optimal estimate for $\boldsymbol{h}$

Linear MMSE Estimator

- **Goal: To minimize the Bayesian Mean Square Error**
  - **Step two:** Find the optimal estimate for $\boldsymbol{h}$

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

$$BMSE = E\left[\left(\theta - (\boldsymbol{h}^T\boldsymbol{x} + E[\theta] - \boldsymbol{h}^T E[\boldsymbol{x}])\right)^2\right]$$

$$BMSE = E\left[\left((\theta - E[\theta]) - (\boldsymbol{h}^T\boldsymbol{x} - \boldsymbol{h}^T E[\boldsymbol{x}])\right)^2\right]$$

$$BMSE = E\left[\left((\theta - E[\theta]) - \boldsymbol{h}^T(\boldsymbol{x} - E[\boldsymbol{x}])\right)^2\right]$$

$$BMSE = E\left[(\theta - E[\theta])^2 - \boldsymbol{h}^T(\boldsymbol{x} - E[\boldsymbol{x}])(\theta - E[\theta])\right.$$

Linear MMSE Estimator

- **Goal: To minimize the Bayesian Mean Square Error**
  - **Step two:** Find the optimal estimate for $\boldsymbol{h}$

$$BMSE = C_{\theta\theta} - \boldsymbol{h}^T \boldsymbol{C}_{\boldsymbol{x}\boldsymbol{\theta}} - \boldsymbol{C}_{\boldsymbol{\theta}\boldsymbol{x}} \boldsymbol{h} + \boldsymbol{h}^T \boldsymbol{C}_{\boldsymbol{x}\boldsymbol{x}} \boldsymbol{h}$$

$$\frac{\partial(BMSE)}{\partial \boldsymbol{h}} = -2\boldsymbol{C}_{\boldsymbol{x}\boldsymbol{\theta}} + 2\boldsymbol{C}_{\boldsymbol{x}\boldsymbol{x}} \boldsymbol{h} = 0$$

$$\boldsymbol{C}_{\boldsymbol{x}\boldsymbol{x}} \boldsymbol{h} = \boldsymbol{C}_{\boldsymbol{x}\boldsymbol{\theta}}$$

$$\boldsymbol{h}^* = \boldsymbol{C}_{\boldsymbol{x}\boldsymbol{x}}^{-1} \boldsymbol{C}_{\boldsymbol{x}\boldsymbol{\theta}}$$

Linear MMSE Estimator

- **<u>Goal:</u> To minimize the Bayesian Mean Square Error**
  - **Step three:** Find the optimal estimator

$$\hat{\theta} = \boldsymbol{h}^T\boldsymbol{x} + a$$
$$a^* = E[\theta] - \boldsymbol{h}^T E[\boldsymbol{x}]$$
$$\boldsymbol{h}^* = \boldsymbol{C}_{\boldsymbol{xx}}^{-1}\boldsymbol{C}_{\boldsymbol{x\theta}}$$

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

$$BMSE = C_{\theta\theta} - \boldsymbol{h}^T\boldsymbol{C}_{\boldsymbol{x\theta}} - \boldsymbol{C}_{\boldsymbol{\theta x}}\boldsymbol{h} + \boldsymbol{h}^T\boldsymbol{C}_{\boldsymbol{xx}}\boldsymbol{h}$$
$$BMSE = C_{\theta\theta} - 2\boldsymbol{C}_{\boldsymbol{x\theta}}\boldsymbol{C}_{\boldsymbol{xx}}^{-1}\boldsymbol{C}_{\boldsymbol{x\theta}} + \boldsymbol{C}_{\boldsymbol{x\theta}}\boldsymbol{C}_{\boldsymbol{xx}}^{-1}\boldsymbol{C}_{\boldsymbol{xx}}\boldsymbol{C}_{\boldsymbol{xx}}^{-1}\boldsymbol{C}_{\boldsymbol{x\theta}}$$
$$BMSE = C_{\hat{\theta}} = C_{\theta\theta} - \boldsymbol{C}_{\boldsymbol{x\theta}}\boldsymbol{C}_{\boldsymbol{xx}}^{-1}\boldsymbol{C}_{\boldsymbol{x\theta}}$$

Have we seen these results before?

Linear MMSE Estimator

- **Goal: To minimize the Bayesian Mean Square Error**
  - **Step three:** Find the optimal estimator

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

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

Have we seen these results before?

Linear MMSE Estimator

- **A simplification**
  - Assume the estimator is a linear system
  - For a **single parameter** estimator

$$\hat{\theta} = \sum_{n=1}^{N} h_n x_n + a = \boldsymbol{h}^T \boldsymbol{x} + a$$

  - For a **multi-parameter** estimator

$$\widehat{\boldsymbol{\theta}} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{a}$$

Linear MMSE Estimator

- **<u>Goal:</u> To minimize the Bayesian Mean Square Error**
  - The LMMSE Estimator is

$$\hat{\theta} = \boldsymbol{H}\boldsymbol{x} + a$$
$$\boldsymbol{a}^* = E[\boldsymbol{\theta}] - \boldsymbol{H}E[\boldsymbol{x}]$$
$$\boldsymbol{H}^* = \boldsymbol{C}_{\boldsymbol{xx}}^{-1}\boldsymbol{C}_{\boldsymbol{x\theta}}$$

$$\hat{\boldsymbol{\theta}} = E[\boldsymbol{\theta}] + \boldsymbol{C}_{\boldsymbol{xx}}^{-1}\boldsymbol{C}_{\boldsymbol{x\theta}}{}^{T}(\boldsymbol{x} - E[\boldsymbol{x}])$$
$$C_{\hat{\boldsymbol{\theta}}} = C_{\boldsymbol{\theta\theta}} - \boldsymbol{C}_{\boldsymbol{x\theta}}\boldsymbol{C}_{\boldsymbol{xx}}^{-1}\boldsymbol{C}_{\boldsymbol{x\theta}}$$

Have we seen these results before?

Linear MMSE Estimator

- **Goal: To minimize the Bayesian Mean Square Error**
  - When we have a linear statistical model: $\boldsymbol{x} = \boldsymbol{G\theta} + \boldsymbol{w}$, for noise $\boldsymbol{w}$
  - We assume the expected value and covariance of $\boldsymbol{\theta}$ and $\boldsymbol{w}$ are both known.

$$\widehat{\boldsymbol{\theta}} = E[\boldsymbol{\theta}] + \left(\boldsymbol{C}_{\boldsymbol{\theta\theta}} + \boldsymbol{G}^T \boldsymbol{C}_{\boldsymbol{w}}^{-1} \boldsymbol{G}\right)^{-1} \boldsymbol{G}^T \boldsymbol{C}_{\boldsymbol{w}}^{-1} (\boldsymbol{x} - \boldsymbol{G}E[\boldsymbol{x}])$$
$$\boldsymbol{C}_{\widehat{\boldsymbol{\theta}}} = \boldsymbol{C}_{\boldsymbol{\theta\theta}} - \boldsymbol{G}^T \boldsymbol{C}_{\boldsymbol{xx}}^{-1} \boldsymbol{G}$$

Have we seen these results before?

Bayesian Estimation:
Linear MMSE Example

Last Time

- **LMMSE Estimator Example**
  - Gaussian likelihood $x|\theta \sim \mathcal{N}(0, \sigma^2)$
  - **Determine the LMMSE Estimator for** $\theta = x^2$ (Chi-squared prior)

Last Time

- **LMMSE Estimator Example**
  - Gaussian likelihood $x|\theta \sim \mathcal{N}(0, \sigma^2)$
  - **Determine the LMMSE Estimator for** $\theta = x^2$ (Chi-squared prior)

$$E\left[\left(\theta - \hat{\theta}\right)^2\right] = E[(\theta - a_0 x - a_1)^2]$$

$$\frac{\partial E\left[\left(\theta - \hat{\theta}\right)^2\right]}{\partial a_1} = E[2(\theta - a_0 x - a_1)] = 0$$

$$E[\theta - a_0 x - a_1] = 0$$

$$E[\theta] - a_1 = 0$$

$$a_1 = E[\theta] = \sigma^2$$

Last Time

- **LMMSE Estimator Example**
  - Gaussian likelihood $x|\theta \sim \mathcal{N}(0, \sigma^2)$
  - **Determine the LMMSE Estimator for** $\theta = x^2$ (Chi-squared prior)

$$E\left[\left(\theta - \hat{\theta}\right)^2\right] = E[(\theta - a_0 x - a_1)^2]$$

$$\frac{\partial E\left[\left(\theta - \hat{\theta}\right)^2\right]}{\partial a_0} = E[(\theta - a_0 x - a_1)x] = 0$$

$$E[\theta x - a_0 x^2 - a_1 x] = 0$$

$$E[\theta x] - a_0 E[x^2] = 0$$

$$E[x^3] - a_0 E[x^2] = 0$$

$$a_0 = 0$$

Last Time

- **LMMSE Estimator Example**
  - Gaussian likelihood $x|\theta \sim \mathcal{N}(0, \sigma^2)$
  - **Determine the LMMSE Estimator for** $\theta = x^2$ (Chi-squared prior)

$$\hat{\theta} = \sigma^2$$

$$BMSE = E\left[\left(\theta - \hat{\theta}\right)^2\right] = E[(x^2 - \sigma^2)^2]$$

$$= E[x^4 - 2x^2\sigma^2 + \sigma^4]$$

$$= 3\sigma^4 - 2\sigma^4 + \sigma^4$$

$$= 2\sigma^4$$

Last Time

- **LMMSE Estimator Example**
  - Gaussian likelihood $x|\theta \sim \mathcal{N}(0, \sigma^2)$
  - **In contrast, the estimator:**

$$\hat{\theta} = x^2$$

$$BMSE = E\left[\left(\theta - \hat{\theta}\right)^2\right] = E[(x^2 - x^2)^2]$$

$$= 0$$

Bayesian Estimation:
Multiple data example

Last Time

- **MMSE / MAP / LMMSE Estimator Example**
  - Consider jointly Gaussian random variables $\boldsymbol{\theta}, \boldsymbol{x}_1, \boldsymbol{x}_2$. $\boldsymbol{x} = \begin{bmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \end{bmatrix}$
  - Assume $\boldsymbol{x}_1, \boldsymbol{x}_2$ are independent.
  - **Determine the MMSE / MAP / LMMSE Estimator for $\boldsymbol{\theta}$.**

Last Time

- **MMSE / MAP / LMMSE Estimator Example**
  - Consider jointly Gaussian random variables $\boldsymbol{\theta}, \boldsymbol{x}_1, \boldsymbol{x}_2$. $\boldsymbol{x} = \begin{bmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \end{bmatrix}$
  - Assume $\boldsymbol{x}_1, \boldsymbol{x}_2$ are independent.
  - **Determine the MMSE / MAP / LMMSE Estimator for $\boldsymbol{\theta}$.**

- **MMSE Estimator:** $\hat{\boldsymbol{\theta}} = E[\boldsymbol{\theta}|\boldsymbol{x}_1, \boldsymbol{x}_2] = E[\boldsymbol{\theta}] + \mathbf{C}_{\boldsymbol{\theta x}} \boldsymbol{C}_{\boldsymbol{xx}}^{-1}(\boldsymbol{x} - E[\boldsymbol{x}])$

$$\mathbf{C}_{\boldsymbol{\theta x}} = E[\boldsymbol{\theta x}^T] = [\mathbf{C}_{\boldsymbol{\theta x_1}} \quad \mathbf{C}_{\boldsymbol{\theta x_2}}]$$

$$\mathbf{C}_{\boldsymbol{xx}} = E[\boldsymbol{xx}^T] = \begin{bmatrix} E[\boldsymbol{x}_1\boldsymbol{x}_1^T] & E[\boldsymbol{x}_1\boldsymbol{x}_2^T] \\ E[\boldsymbol{x}_1\boldsymbol{x}_2^T] & E[\boldsymbol{x}_2\boldsymbol{x}_2^T] \end{bmatrix} = \begin{bmatrix} C_{x_1x_1} & 0 \\ 0 & C_{x_2x_2} \end{bmatrix}$$

Last Time

- **MMSE / MAP / LMMSE Estimator Example**
  - Consider jointly Gaussian random variables $\boldsymbol{\theta}, \boldsymbol{x}_1, \boldsymbol{x}_2$. $\boldsymbol{x} = \begin{bmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \end{bmatrix}$
  - Assume $\boldsymbol{x}_1, \boldsymbol{x}_2$ are independent.
  - **Determine the MMSE / MAP / LMMSE Estimator for $\boldsymbol{\theta}$.**

- **MMSE Estimator:** $\widehat{\boldsymbol{\theta}} = E[\boldsymbol{\theta}|\boldsymbol{x}_1, \boldsymbol{x}_2] = E[\boldsymbol{\theta}] + \mathbf{C}_{\boldsymbol{\theta x}} \boldsymbol{C}_{\boldsymbol{xx}}^{-1} (\boldsymbol{x} - E[\boldsymbol{x}])$

$$\widehat{\boldsymbol{\theta}} = \begin{bmatrix} \mathbf{C}_{\boldsymbol{\theta x}_1} & \mathbf{C}_{\boldsymbol{\theta x}_2} \end{bmatrix} \begin{bmatrix} C_{x_1 x_1} & 0 \\ 0 & C_{x_2 x_2} \end{bmatrix} \begin{bmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \end{bmatrix}$$

$$\widehat{\boldsymbol{\theta}} = E[\boldsymbol{\theta}] + \mathbf{C}_{\boldsymbol{\theta x}_1} \boldsymbol{C}_{\boldsymbol{x}_1 \boldsymbol{x}_1}^{-1} (\boldsymbol{x} - E[\boldsymbol{x}]) + \mathbf{C}_{\boldsymbol{\theta x}_2} \boldsymbol{C}_{\boldsymbol{x}_2 \boldsymbol{x}_2}^{-1} (\boldsymbol{x} - E[\boldsymbol{x}])$$

Last Time

- **MMSE / MAP / LMMSE Estimator Example**
  - Consider jointly Gaussian random variables $\boldsymbol{\theta}, \boldsymbol{x}_1, \boldsymbol{x}_2$. $\boldsymbol{x} = \begin{bmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \end{bmatrix}$
  - Assume $\boldsymbol{x}_1, \boldsymbol{x}_2$ are independent.
  - Assume $\boldsymbol{\theta}$ is zero mean.
  - **Determine the MMSE / MAP / LMMSE Estimator for $\boldsymbol{\theta}$.**

- **MMSE Estimator:** $\widehat{\boldsymbol{\theta}} = E[\boldsymbol{\theta}|\boldsymbol{x}_1, \boldsymbol{x}_2] = E[\boldsymbol{\theta}] + \mathbf{C}_{\boldsymbol{\theta x}} \boldsymbol{C}_{\boldsymbol{xx}}^{-1}(\boldsymbol{x} - E[\boldsymbol{x}])$

$$\widehat{\boldsymbol{\theta}} = [\mathbf{C}_{\boldsymbol{\theta}\boldsymbol{x}_1} \quad \mathbf{C}_{\boldsymbol{\theta}\boldsymbol{x}_2}] \begin{bmatrix} C_{x_1x_1} & 0 \\ 0 & C_{x_2x_2} \end{bmatrix} \begin{bmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \end{bmatrix}$$

$$\widehat{\boldsymbol{\theta}} = \mathbf{C}_{\boldsymbol{\theta}\boldsymbol{x}_1} \boldsymbol{C}_{\boldsymbol{x}_1\boldsymbol{x}_1}^{-1}(\boldsymbol{x} - E[\boldsymbol{x}]) + \mathbf{C}_{\boldsymbol{\theta}\boldsymbol{x}_2} \boldsymbol{C}_{\boldsymbol{x}_2\boldsymbol{x}_2}^{-1}(\boldsymbol{x} - E[\boldsymbol{x}]) = E[\boldsymbol{\theta}|\boldsymbol{x}_1] + E[\boldsymbol{\theta}|\boldsymbol{x}_2]$$

Additive property of independent data sets

Bayesian Estimation:
Dynamic / state models (scalar case)

Kalman Filters

- **Consider the signal problem**

$$\boldsymbol{x} = \boldsymbol{s} + \boldsymbol{w}$$
$$x[n] = s[n] + w[n]$$

where $\boldsymbol{w}$ is Gaussian noise with covariance $\boldsymbol{I}\sigma^2$.

What is the MVU Estimator of $\boldsymbol{s}$?

Why is this a pretty terrible estimator?

How could we improve it?

Kalman Filters

- **The Idea:**
  - The signal has:
    - Nonstationary statistics (i.e., the mean changes over time)
    - Correlated values (x[n] is related to x[n-1])
  - Design an estimator such that
    - The current estimated value based on previous measurements
    - This is the idea of a ***filter***.

Kalman Filters

- **First-order Gauss-Markov *Process*:**
  - Assume our signal can be represented by a *dynamical* or *state model*

Measurement

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

Measurement noise

$$s[n] = as[n-1] + u[n]$$

Current state of the system

Previous state of the system

excitation noise, or the *plant noise*

This is our current model.

Kalman Filters

- **First-order Gauss-Markov *Process*:**
  - Assume our signal can be represented by a *dynamical* or *state model*

Measurement

$$x[n] = s[n] + w[n]$$
$$s[n] = as[n-1] + u[n]$$

Measurement noise

Current state of the system

Previous state of the system

excitation noise, or the *plant noise*

This is our has an additional state model.

Kalman Filters

- **First-order Gauss-Markov *Process*:**
  - Assume our signal can be represented by a *dynamical* or *state model*

Measurement

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

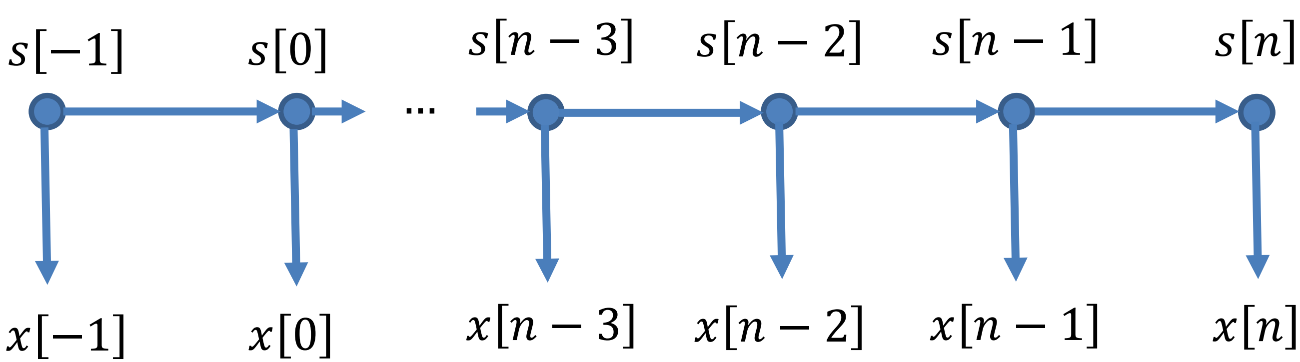
$$s[n] = as[n-1] + u[n]$$

Measurement noise

Current state of the system

Previous state of the system

excitation noise, or the *plant noise*



Kalman Filters

- **First-order Gauss-Markov *Process*:**
  - Assume our signal can be represented by a *dynamical* or *state model*

Measurement

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

Measurement noise

$$s[n] = as[n-1] + u[n]$$

Current state of the system

Previous state of the system

excitation noise, or the *plant noise*

What the closed form solution for $s[n]$?

Kalman Filters

- **First-order Gauss-Markov *Process*:**
  - Assume our signal can be represented by a *dynamical* or *state model*

$$s[n] = as[n-1] + u[n]$$

$$s[0] = as[-1] + u[0]$$
$$s[1] = as[0] + u[1] = a^2 s[-1] + au[0] + u[1]$$
$$s[2] = a^3 s[-1] + a^2 u[0] + au[1] + u[2]$$
$$s[3] = a^4 s[-1] + a^3 u[0] + a^2 u[1] + au[2] + u[3]$$

Kalman Filters

- **First-order Gauss-Markov *Process*:**
  - Assume our signal can be represented by a *dynamical* or *state model*

$$s[n] = as[n-1] + u[n]$$

$$\begin{aligned}
s[0] &= as[-1] + u[0] \\
s[1] &= as[0] + u[1] = a^2 s[-1] + au[0] + u[1] \\
s[2] &= a^3 s[-1] + a^2 u[0] + au[1] + u[2] \\
s[3] &= a^4 s[-1] + a^3 u[0] + a^2 u[1] + au[2] + u[3]
\end{aligned}$$

$$s[n] = a^{n+1} s[-1] + \sum_{k=0}^{n} a^k u[n-k]$$

Kalman Filters

- **First-order Gauss-Markov *Process*:**
  - Assume our signal can be represented by a *dynamical* or *state model*

$$s[n] = a^{n+1}s[-1] + \sum_{k=0}^{n} a^k u[n-k]$$

  - Assume (for now)
    - $s[-1] \sim \mathcal{N}(\mu_s, \sigma_s^2)$
    - $u[n] \rightarrow \boldsymbol{u} \sim \mathcal{N}(\boldsymbol{0}, \sigma_u^2 \boldsymbol{I})$
    - $s[-1]$ and $u[n]$ are independent

What is the expected value of $s[n]$?

Kalman Filters

- **First-order Gauss-Markov *Process*:**
  - Assume our signal can be represented by a *dynamical* or *state model*

$$s[n] = a^{n+1}s[-1] + \sum_{k=0}^{n} a^k u[n-k]$$

  - Assume (for now)
    - $s[-1] \sim \mathcal{N}(\mu_s, \sigma_s^2)$
    - $u[n] \rightarrow \boldsymbol{u} \sim \mathcal{N}(\mathbf{0}, \sigma_u^2 \boldsymbol{I})$
    - $s[-1]$ and $u[n]$ are independent

What is the expected value of $s[n]$?

$$E\big[s[n]\big] = a^{n+1}\mu_s$$

Kalman Filters

- **First-order Gauss-Markov *Process*:**
  - Assume our signal can be represented by a *dynamical* or *state model*

$$s[n] = a^{n+1}s[-1] + \sum_{k=0}^{n} a^k u[n-k]$$

  - Assume (for now)
    - $s[-1] \sim \mathcal{N}(\mu_s, \sigma_s^2)$
    - $u[n] \rightarrow \boldsymbol{u} \sim \mathcal{N}(\boldsymbol{0}, \sigma_u^2 \boldsymbol{I})$
    - $s[-1]$ and $u[n]$ are independent

What is the expected value of $s[n]$ (in terms of $s[n-1]$)?

Kalman Filters

- **First-order Gauss-Markov *Process*:**
  - Assume our signal can be represented by a *dynamical* or *state model*

$$s[n] = a^{n+1}s[-1] + \sum_{k=0}^{n} a^k u[n-k]$$

  - Assume (for now)
    - $s[-1] \sim \mathcal{N}(\mu_s, \sigma_s^2)$
    - $u[n] \rightarrow \boldsymbol{u} \sim \mathcal{N}(\boldsymbol{0}, \sigma_u^2 \boldsymbol{I})$
    - $s[-1]$ and $u[n]$ are independent

What is the expected value of $s[n]$ (in terms of $s[n-1]$)?

$$E\big[s[n]\big] = aE\big[s[n-1]\big]$$

Kalman Filters

- **First-order Gauss-Markov *Process*:**
  - Assume our signal can be represented by a *dynamical* or *state model*

$$s[n] = a^{n+1}s[-1] + \sum_{k=0}^{n} a^k u[n-k]$$

  - Assume (for now)
    - $s[-1] \sim \mathcal{N}(\mu_s, \sigma_s^2)$
    - $u[n] \to \boldsymbol{u} \sim \mathcal{N}(\boldsymbol{0}, \sigma_u^2 \boldsymbol{I})$
    - $s[-1]$ and $u[n]$ are independent

What is the covariance between of $s[n], s[m]$?

Kalman Filters

- **First-order Gauss-Markov *Process*:**
  - Assume our signal can be represented by a *dynamical* or *state model*

What is the covariance between of $s[n], s[m]$?

$$E[(s[n] - a^{n+1}\mu_s)(s[m] - a^{m+1}\mu_s)]$$

$$= E\left[\left(a^{n+1}s[-1] + \sum_{k=0}^{n} a^k u[n-k] - a^{n+1}\mu_s\right)\left(a^{n+1}s[-1] + \sum_{\ell=0}^{m} a^k u[n-k] - a^{m+1}\mu_s\right)\right]$$

$$= E\left[\left(a^{n+1}(s[-1] - \mu_s) + \sum_{k=0}^{n} a^k u[n-k]\right)\left(a^{m+1}(s[-1] - \mu_s) + \sum_{\ell=0}^{m} a^\ell u[n-\ell]\right)\right]$$

Due to the independence of $s[-1]$ and $u[n]$

$$= a^{m+n+2}\sigma_s^2 + \sum_{k=0}^{n}\sum_{\ell=0}^{m} a^{k+\ell} E\big[u[m-k]u[n-\ell]\big]$$

Due to the independence of each $u[n]$

$$C(s[n], s[m]) = a^{m+n+2}\sigma_s^2 + \sigma_u^2 a^{m-n} \sum_{k=0}^{n} a^{2k}$$

Kalman Filters

- **First-order Gauss-Markov *Process*:**
  - Assume our signal can be represented by a *dynamical* or *state model*

What is the covariance between of $s[n], s[m]$?

$$C(s[n], s[m]) = a^{m+n+2}\sigma_s^2 + \sigma_u^2 a^{m-n} \sum_{k=0}^{n} a^{2k}$$

$$\mathrm{var}(s[n]) = C(s[n], s[n]) = a^{2n+2}\sigma_s^2 + \sigma_u^2 \sum_{k=0}^{n} a^{2k}$$

If $a = 1$ *?:*

Kalman Filters

- **First-order Gauss-Markov *Process*:**
  - Assume our signal can be represented by a *dynamical* or *state model*

What is the covariance between of $s[n], s[m]$?

$$C(s[n], s[m]) = a^{m+n+2}\sigma_s^2 + \sigma_u^2 a^{m-n} \sum_{k=0}^{n} a^{2k}$$

$$\mathrm{var}(s[n]) = C(s[n], s[n]) = a^{2n+2}\sigma_s^2 + \sigma_u^2 \sum_{k=0}^{n} a^{2k}$$

If $a = 1$*:*

$$C(s[n], s[n]) = \sigma_s^2 + n\,\sigma_u^2$$

Kalman Filters

- **First-order Gauss-Markov *Process*:**
  - Assume our signal can be represented by a *dynamical* or *state model*

What is the covariance between of $s[n], s[m]$?

$$C(s[n], s[m]) = a^{m+n+2}\sigma_s^2 + \sigma_u^2 a^{m-n} \sum_{k=0}^{n} a^{2k}$$

$$\mathrm{var}(s[n]) = C(s[n], s[n]) = a^{2n+2}\sigma_s^2 + \sigma_u^2 \sum_{k=0}^{n} a^{2k}$$

If $|a| < 1, n \to \infty$: ?

Kalman Filters

- **First-order Gauss-Markov *Process*:**
  - Assume our signal can be represented by a *dynamical* or *state model*

What is the covariance between of $s[n], s[m]$?

$$C(s[n], s[m]) = a^{m+n+2}\sigma_s^2 + \sigma_u^2 a^{m-n} \sum_{k=0}^{n} a^{2k}$$

$$\mathrm{var}(s[n]) = C(s[n], s[n]) = a^{2n+2}\sigma_s^2 + \sigma_u^2 \sum_{k=0}^{n} a^{2k}$$

If $|a| < 1, n \to \infty$*:* (use geometric series relationship)

$$C(s[n], s[m]) = \frac{\sigma_u^2\, a^{m-n}}{1 - a^2}$$

$$C(s[n], s[n]) = \frac{\sigma_u^2}{1 - a^2}$$

Kalman Filters

- **First-order Gauss-Markov *Process*:**
  - Assume our signal can be represented by a *dynamical* or *state model*

What is the covariance between of $s[n], s[m]$?

$$C(s[n], s[m]) = a^{m+n+2}\sigma_s^2 + \sigma_u^2 a^{m-n} \sum_{k=0}^{n} a^{2k}$$

$$\mathrm{var}(s[n]) = C(s[n], s[n]) = a^{2n+2}\sigma_s^2 + \sigma_u^2 \sum_{k=0}^{n} a^{2k}$$

In terms of s[n-1]: ?

Kalman Filters

- **First-order Gauss-Markov *Process*:**
  - Assume our signal can be represented by a *dynamical* or *state model*

What is the covariance between of $s[n], s[m]$?

$$C(s[n], s[m]) = a^{m+n+2}\sigma_s^2 + \sigma_u^2 a^{m-n} \sum_{k=0}^{n} a^{2k}$$

$$\operatorname{var}(s[n]) = C(s[n], s[n]) = a^{2n+2}\sigma_s^2 + \sigma_u^2 \sum_{k=0}^{n} a^{2k}$$

In terms of s[n-1]:

$$\operatorname{var}(s[n-1]) = a^{2n}\sigma_s^2 + \sigma_u^2 \sum_{k=0}^{n-1} a^{2k}$$

$$\operatorname{var}(s[n]) = a^2 \operatorname{var}(s[n-1]) + \sigma_u^2$$

Kalman Filters

- **First-order Gauss-Markov *Process*:**
  - Assume our signal can be represented by a *dynamical* or *state model*

$$s[n] = as[n-1] + u[n]$$

$$E\left[s[n]\right] = aE\left[s[n-1]\right]$$
$$\mathrm{var}(s[n]) = a^2\mathrm{var}(s[n-1]) + \sigma_u^2$$