LECURE 22

- **Readings:** Sect. 8.3–8.4; reread Sect. 4.2 and pp. 225–226

**Lecture outline**

- End of semester
- Review
- Performance criteria for estimators
- (Bayesian) Least mean squares estimation
- (Bayesian) Linear least mean squares estimation
Semester end game

- Chapter 8 all covered
- Chapter 9
  - Sect. 9.1: covered through middle of p. 470
  - Sect. 9.2: covered through middle of p. 482
  - Sect. 9.3: all covered
  - Sect. 9.4: not covered
- Last problem set is for practice, not to be turned in
- Final exam: Tuesday, May 17, 9am–noon
- Many office hours between last lecture and final exam
Review: Bayesian inference

- Posterior computation is use of Bayes’ rule, for example

\[
p_{\theta|x}(\theta|x) = \frac{p_{\Theta}(\theta)p_{X|\Theta}(x|\theta)}{\sum_k p_{\Theta}(k)p_{X|\Theta}(x|k)}
\]

- Estimate \( \hat{\theta} \) is number computed from posterior

- Maximum a posteriori probability (MAP) rule

\[
\hat{\theta}_{\text{MAP}} = \arg \max_{\theta} p_{\Theta|x}(\theta|x) \quad \text{or} \quad \hat{\theta}_{\text{MAP}} = \arg \max_{\theta} f_{\Theta|x}(\theta|x)
\]
Hypothesis testing

- Estimation with discrete $\Theta$ called **hypothesis testing**

- Common formulation:
  - $\theta$ and $\hat{\theta}$ in $\{1, 2, \ldots, m\}$
  - nonnegative cost $c_{ij}$ for choosing $\hat{\theta} = j$ when $\theta = i$
  - $c_{ii} \leq c_{ij}$ for each $j \neq i$; may as well have $c_{ii} = 0$ for each $i$
  - minimize expected cost:
    \[
    \sum_{i=1}^{m} \sum_{j=1}^{m} c_{ij} P(\Theta = i, \hat{\Theta} = j) = \sum_{i=1}^{m} \sum_{j=1}^{m} c_{ij} P(\hat{\Theta} = j | \Theta = i) P(\Theta = i)
    \]
    - good to make $P(\hat{\Theta} \neq \Theta)$ small, but errors not equally important (costly)
    - equal costs makes MAP rule optimal
Binary hypothesis testing example

- Prior given: \( P(\Theta = 1) = p, \quad P(\Theta = 2) = 1 - p \)
- Likelihoods given: \( f_{X|\Theta}(x \mid 1), \quad f_{X|\Theta}(x \mid 2) \)
- Costs given: \( c_{12} \) (mistake 1 for 2), \( c_{21} \) (mistake 2 for 1)
- Minimize expected cost
Bayesian least mean squares (LMS) estimation

- Any estimator is function of observations: \( \hat{\Theta} = g(X) \)
- LMS estimator \( \hat{\Theta}_{\text{LMS}} \) minimizes \( \mathbb{E}[(\Theta - \hat{\Theta})^2] \)
- LMS estimator is \( g_{\text{LMS}}(X) = \mathbb{E}[\Theta | X] \)

- Recall from L12: For random variable \( Y \) and number \( c \)
  \[
  \mathbb{E}[(Y - c)^2] = \text{var}(Y - c) + (\mathbb{E}[Y - c])^2 = \text{var}(Y) + (\mathbb{E}[Y - c])^2
  \]
LMS estimation example

\[ f_\Theta(\theta) \]

\[ f_{X|\Theta}(x | \theta) \]

\[ \theta \]

\[ x \]
Conditional mean squared error

- \( \mathbb{E}[(\Theta - \mathbb{E}[\Theta | X])^2 | X = x] \)
  same as \( \text{var}(\Theta | X = x) \):
  variance of the conditional distribution of \( \Theta \)
Some properties of LMS estimation

- Estimator: \( \hat{\Theta} = \mathbb{E}[\Theta \mid X] \)

- Estimation error: \( \tilde{\Theta} = \hat{\Theta} - \Theta \)

- \( \mathbb{E}[\tilde{\Theta}] = 0 \)

- \( \mathbb{E}[\tilde{\Theta} \mid X = x] = 0 \)

- \( \mathbb{E}[\tilde{\Theta} h(X)] = 0 \), for any function \( h \)

- \( \text{cov}(\tilde{\Theta}, \hat{\Theta}) = 0 \)

- Since \( \Theta = \hat{\Theta} + \tilde{\Theta} \):
  \( \text{var}(\Theta) = \text{var}(\hat{\Theta}) + \text{var}(\tilde{\Theta}) \)
Linear LMS

- Consider estimators of the form $\hat{\Theta}_{\text{LLMS}} = aX + b$
- Minimize $E[(\Theta - aX - b)^2]$
- Best choice of $a, b$; best linear estimator:

\[
\hat{\Theta}_{\text{LLMS}} = \mathbb{E}[\Theta] + \frac{\text{cov}(X, \Theta)}{\text{var}(X)} (X - \mathbb{E}[X])
\]

\[
\mathbb{E}[(\hat{\Theta}_{\text{LLMS}} - \Theta)^2] = (1 - \rho^2)\sigma_\Theta^2
\]
Linear LMS: Example
Linear LMS with more data

- Consider estimators of the form:
  \[ \hat{\Theta} = a_1 X_1 + \cdots + a_n X_n + b \]

- Find best choices of \( a_1, \ldots, a_n, b \)

- Minimize:
  \[ \mathbb{E}[(a_1 X_1 + \cdots + a_n X_n + b - \Theta)^2] \]

- Set derivatives to zero
  linear system in \( b \) and the \( a_i \)

- Only means, variances, covariances matter