# Stochastic Processes

## Midterm 2

9:00 a.m. – 11:30 a.m., 12/01/08

**IMPORTANT:**

- Remember to write down your i.d. number and your name.
- There are 4 problem sets with 110 points in total.
- Show all your work with detailed explanations. Correct answers ***without*** any explanations will carry **NO** credits. However, wrong answers with correct reasonings will get partial credits.

You will need the following definitions/formulas:

- For any arbitrary random variable $X$ with mean $E[X]$ and finite variance $\mathrm{Var}(X)$, we have
$$P\left[\,\left|X - E[X]\right| > k\right] \leq \frac{\mathrm{Var}(X)}{k^2}, \quad \text{for any} \quad k > 0.$$

- Consider a jointly Gaussian random vector $[\mathbf{x}^T \; \mathbf{y}^T]^T$ with $\mathbf{x} \sim \mathcal{N}(\mathbf{m_x}, \mathbf{K_x})$ and $\mathbf{y} \sim \mathcal{N}(\mathbf{m_y}, \mathbf{K_y})$. Then, the conditional density $f_{\mathbf{x}|\mathbf{y}}(\mathbf{x}|\mathbf{y})$ also follows a joint Gaussian density with
$$\mathbf{x}|\mathbf{y} \sim \mathcal{N}\Big(\mathbf{m_x} + \mathbf{K_{xy}}\mathbf{K_y}^{-1}(\mathbf{y} - \mathbf{m_y}), \mathbf{K_x} - \mathbf{K_{xy}}\mathbf{K_y}^{-1}\mathbf{K_{yx}}\Big),$$
where $\mathbf{K_{xy}} = E\left[(\mathbf{x} - \mathbf{m_x})(\mathbf{y} - \mathbf{m_y})^H\right]$ and $\mathbf{K_{yx}} = \mathbf{K}_{\mathbf{xy}}^H$

- A matrix $\mathbf{K}$ is ***positive semi-definite*** if
$$\mathbf{x}^T\mathbf{K}\mathbf{x} \geq 0$$
for any real-valued vector $\mathbf{x}$ with appropriate dimension.

- An unbiased estimator $\hat{\boldsymbol{\theta}}$ of a vector deterministic parameter $\boldsymbol{\theta}$ is said to be more ***efficient*** than any other vector unbiased estimator $\hat{\boldsymbol{\theta}}'$ if
$$\mathbf{K}_{\hat{\theta}} \leq \mathbf{K}_{\hat{\theta}'}$$
where the inequality for the matrix means $\mathbf{K}_{\hat{\theta}'} - \mathbf{K}_{\hat{\theta}}$ is a ***positive semi-definite*** matrix, where $\mathbf{K}_{\hat{\theta}}$ and $\mathbf{K}_{\hat{\theta}'}$ are the covariance matrices of $\hat{\boldsymbol{\theta}}$ and $\hat{\boldsymbol{\theta}}'$, respectively.

1. **Bayes Detection** (10+10=20 points)

   (a) Describe the maximum *a posteriori* (MAP) detection rule, and explain why MAP gives minimum probability of decision error.

   (b) Consider the problem of detecting a known signal $\mathbf{s}$, with dimension $m \times 1$, in the presence of additive noise $\mathbf{n}$. Assume the noise has a Gaussian distribution with $\mathbf{n} \sim \mathcal{N}(\mathbf{0}, \mathbf{K}_n)$ where $\mathbf{K}_n$ is nonsingular. The detector is to determine whether the received signal $\mathbf{y}$ consists of signal plus noise or noise alone. That is, the two hypotheses to be tested for are

$$\begin{aligned} \mathrm{H}_1 : &\quad \mathbf{y} = \mathbf{s} + \mathbf{n} \\ \mathrm{H}_2 : &\quad \mathbf{y} = \mathbf{n}. \end{aligned}$$

   Determine the maximum likelihood decision rule for the two hypotheses.

2. **Estimation** $(10 + 5 + 5 = 20$ points)

Consider the linear model

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w},$$

where $\mathbf{H}$ is a known $m \times n$ observation matrix, $\mathbf{x}$ is an $n \times 1$ unknown parameter, and $\mathbf{w}$ is a Gaussian noise vector with $\mathbf{w} \sim \mathcal{N}(0, \sigma^2\mathbf{I})$. Assume $\mathbf{H}$ is full rank.

(a) The maximum likelihood estimator for $\mathbf{x}$ is given by

$$\begin{aligned} \hat{\mathbf{x}}_{ML} &= \underbrace{(\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T}_{\mathbf{T}_{ML}} \cdot \mathbf{y} \\ &= \mathbf{T}_{ML} \cdot \mathbf{y}, \end{aligned}$$

where $\mathbf{T}_{ML} = (\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T$ is the linear transformation matrix.

Show that $\hat{\mathbf{x}}_{ML}$ is the most efficient estimator in the class $\mathcal{L}$ of ***unbiased*** linear estimators, where $\mathcal{L}$ is the set $\mathcal{L} = \{\hat{\mathbf{x}} | \hat{\mathbf{x}} = \mathbf{T} \cdot \mathbf{y}, \quad \forall\, \mathbf{T}$ such that $E[\hat{\mathbf{x}}] = \mathbf{x}\}$.

(Hint: For $E[\hat{\mathbf{x}}] = \mathbf{x}$, we must have $\mathbf{TH} = \mathbf{I}$)

(b) From part (a), explain that the variance of every component in $\hat{\mathbf{x}}_{ML}$ is no larger than that of $\hat{\mathbf{x}}$.

(c) Geometrically explain that the least square estimator $\hat{\mathbf{x}}_{LS}$ for $\mathbf{x}$ satisfies the normal equation

$$\left(\mathbf{H}^T\mathbf{H}\right)\hat{\mathbf{x}}_{LS} = \mathbf{H}^T\mathbf{y}.$$

3. **Conditional Expectation of Jointly Gaussian** (10+10=20 points)

Let $\mathbf{x} \sim \mathcal{N}(\mathbf{m}_x, \mathbf{K}_x)$, $\mathbf{y} \sim \mathcal{N}(\mathbf{m}_y, \mathbf{K}_y)$, and $\mathbf{z} \sim \mathcal{N}(\mathbf{m}_z, \mathbf{K}_z)$ be collectively jointly Gaussian random vectors, *i.e.* the elements of the random vector $[\mathbf{x}^T, \mathbf{y}^T, \mathbf{z}^T]^T$ are jointly Gaussian. Suppose $\mathbf{y}$ and $\mathbf{z}$ are not statistically independent. Define $\hat{\mathbf{z}} \triangleq \mathbf{z} - E[\mathbf{z}|\mathbf{y}]$.

(a) Find the cross-covariance matrix $\mathbf{K}_{y\hat{z}} = E\Big[(\mathbf{y} - \mathbf{m}_y)(\hat{\mathbf{z}} - E[\hat{\mathbf{z}}])^T\Big]$.

(b) Show that $E[\mathbf{x}|\mathbf{y}, \mathbf{z}] = E[\mathbf{x}|\mathbf{y}, \hat{\mathbf{z}}]$.

4. **Gaussian Sample** $(10 \times 5 = 50$ points)

Let $X_1, \cdots, X_n$ be i.i.d. Gaussian random variables each with ***unknown*** mean $\mu$ and ***unknown*** variance $\sigma^2$. The sample mean of the Gaussian sample is $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$, and sample variance is $S_n^2 \triangleq \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2$.

(a) Show that the sample mean is unbiased and consistent.

(b) Find the maximum likelihood (ML) estimator $\hat{\sigma}^2_{ML}$ for the variance based on $X_1, \cdots, X_n$.

(c) Are the ML estimator $\hat{\sigma}^2_{ML}$ for the variance unbiased? Please explain.

(d) Find the mean squared error (MSE) between $\hat{\sigma}^2_{ML}$ and $\sigma^2$. Then, compare your result with the MSE between $S_n^2$ and $\sigma^2$.
(Hint: $\mathrm{Var}\left(S_n^2\right) = \frac{2\sigma^4}{n-1}$ and $S_n^2$ is unbiased)

(e) In order to find an interval estimator for the unknown mean, we typically use the following probability

$$P\left[-z \leq \frac{\bar{X}_n - \mu}{S_n/\sqrt{n}} \leq z\right] = 1 - \alpha$$

to specify an interval for $\mu$ with a confidence level $1-\alpha$. Suppose we are given that the value $\beta$ satisfies $P\big[T_{n-1} \leq \beta\big] = 0.025$, where $T_{n-1}$ is the Student $T$ random variable with $n-1$ degrees of freedom. Find the interval with a confidence level 0.95.