ECEN 5244 - Stochastic Environmental Signal Processing

Homework #4

Issued 11/16/23
Due 12/12/23

4.1 In Bayesian random parameter estimation show that if the measurements $\overline{y}$ depend on only one component of $\overline{x}$ (say, x_j), *and* if the components of $\overline{x}$ are all statistically independent, then the lower bound on the error in estimating x_j is:

$$\left\langle e_j^2 \right\rangle \geq \frac{1}{\left\langle \left(\frac{\partial \ln p(x_j, \overline{y})}{\partial x_j} \right)^2 \right\rangle}$$

Start your proof using the multiparameter error covariance bound based on the Fisher information matrix:

$$\left\langle \overline{e}\,\overline{e}^t \right\rangle \triangleq \overline{\overline{R}}_{ee} \geq \overline{\overline{J}}^{-1}$$

where

$$\overline{\overline{J}} = \left\langle \left[\nabla_x \ln p(\overline{x}, \overline{y}) \right]^t \nabla_x \ln p(\overline{x}, \overline{y}) \right\rangle$$

4.2 Prove the alternate form of the Cramér-Rao bound for a single random parameter:

$$\left\langle e_j^2 \right\rangle \geq \frac{1}{\left\langle \left[\frac{\partial \ln p(x_j, \overline{y})}{\partial x_j} \right]^2 \right\rangle} = \frac{-1}{\left\langle \frac{\partial^2 \ln p(x_j, \overline{y})}{\partial x_j^2} \right\rangle}$$

That is, show that the second equality in the above expression holds.

4.3 Show that for M independent identically distributed continuous random variables x_j, $j = 1 \ldots M$, that the entropy H_M for the M-vector:

$$\overline{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_M \end{bmatrix}$$

is equal to M times the entropy for any single component of $\overline{x}$.

4.4 (after Van Trees, Problem 2.2.2) The conditional probability density functions for measurements y under each of two hypotheses H_0 and H_1 are:

$$\begin{aligned} p\left(y|H_0\right) &= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2} \\ p\left(y|H_1\right) &= \frac{1}{2} e^{-|y|} \end{aligned}$$

(i) Find an expression for the likelihood ratio $\Lambda\left(y\right)$.

(ii) Assuming that the likelihood ratio test is of the form:

$$\Lambda\left(y\right) \underset{H_0}{\overset{H_1}{\gtrless}} \eta$$

decribe the decision regions as a function of η.

4.5 (after Van Trees, Problem 2.3.3) A measured random variable y is Gaussian for each of five hypotheses:

$$
\begin{aligned}
p\left(y|H_k\right) &= \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y-m_k)^2}{2\sigma^2}}, \qquad k = 1\ldots 5 \\
m_1 &= -2m \\
m_2 &= -m \\
m_3 &= 0 \\
m_4 &= m \\
m_5 &= 2m
\end{aligned}
$$

The hypotheses are equally likely and the criterion for hypothesis detection is minimum probability of error.

(i) Determine the decision regions on the y-axis.

(ii) Compute the error probabality.

4.6 The Matlab file "hw4_6.m" and associated SOHO (Solar and Heliospheric Observatory) .jpg image files on Canvas can be used to read and analyze the pdf of four-band imagery of the sun. These data were observed using the SOHO Extreme ultraviolet Imaging Telescope (EIT) on November 20, 2011 at ~0100 UTC. Run this code to load the images and perform a preliminary rendering and scatter analysis of the data recorded by EIT at the following UV wavelengths:

Plane 1: 171Å
Plane 2: 195Å
Plane 3: 284Å
Plane 4: 495Å

For this homework build on this Matlab code to study i) the entropy of the imagery, ii) develop a principle components analysis of the imagery, iii) estimate magnetic field anomolies and detect sunspots using statistical LMMSE estimation applied to EIT data, and iv) study the estimate error for orthogonality to the data.

Once developed, your algorithm should be tested on real (up-to-date) SOHO data that you can download at:

http://sohowww.nascom.nasa.gov/data/realtime-images.html

Note that the test (unknown) data will be assumed to be data from another instrument on SOHO, the SDO Helioseismic and Magnetic Imager (SDO/HMI), which observes at a longer wavelength (6173Å) than does EIT.

(1) From our discussion of the entropy of an image, find the information content in each of the image planes by calculating the expected number of bits that would be required to code each pixel. (Note that this would be the entropy of each bit plane obtained using the natural logarithm divided by ln(2).) How does this compare to the number of bits required to encode random images (i.e., with values uniformly distributed over all levels of each image plane)?

(2) Using the multispectral pixels as a set of four-dimensional vectors, compute the covariance matrix. How does the black background affect the covariance matrix calculation? Can you avoid using pixels with intensity value 3 or less?

(3) Using the covariance matrix from (2), compute the principal component imagery. Render this data in both image form and using scatter plots as was rendered the original data. Comment on the correlation between the principal component amplitudes and imagery. What are the standard deviations of the principal component amplitudes?

(4) Render a single false-colored (RGB) image of the first three (most dominant) principal components? Comment on the utility of this image.

(5) Render the HMI continuum and magnetic anomaly images (planes 1 and 2 of data x, respectively) and compare visually with the rendered data from EIT and its principal components.

(6) Compute the D-matrix to estimate the HMI continuum and magnetic anomaly data from the EIT data.

(7) Render a scatter plot of the continuum and magnetic anomaly data. Is the strong orthogonality condition achieved? The weak condition?

(8) Test your LMMSE algorithm on up-to-date EIT spectral data from the SOHO data site. This can be obtained by downloading the .jpg images and reading them into Matlab using simple variants on the m-file provided. How do the estimates appear compared to the observed SDO/HMI data?