

ECE 712 Take-home final exam 2008  
Due: Thursday Dec. 11, 2008 at 4:00pm.

James P. Reilly

Please hand in completed exam to Cheryl Gies in ITB A112 by Thursday Dec. 11, 2008 by 4:00pm.  
This exam is based on the honour system. As such, please sign the following statement:

*This exam is entirely my own work. I may consult any written material or text, but I have NOT received nor offered any help from any other person.*

Name:

Date:

Signature:

Given that individuals may be sitting closely together in the same lab, doing this exam at the same time, it may be difficult to avoid discussion of this exam. I therefore urge you avoid this situation and work e.g., in the library. It is your responsibility to work in an environment where you can honour the above oath. Students can feel free to contact me by email for clarification of any questions.

**For each question, explain your method fully, but as precisely and concisely as possible. Include any matlab code if appropriate.**

1. We have observations of data  $\mathbf{z}(t) \in \Re^m, t = 1, \dots, L$  which are generated according to

$$\mathbf{z}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{n}(t) \tag{1}$$

where  $\mathbf{x}(t) \in \Re^n$  are known vectors, and  $\mathbf{A} \in \Re^{m \times n}$  is unknown and to be estimated. The noise vector  $\mathbf{n}(t) \in \Re^m$  is a Gaussian-distributed random vector with zero mean and covariance matrix  $\mathbf{Q}$ , which is also to be estimated.

The matrices  $\mathbf{A}$  and  $\mathbf{Q}$  can be estimated by maximizing the joint probability distribution  $p(\mathbf{z}(t) | \mathbf{A}, \mathbf{Q}, \mathbf{x}(t))$  for  $t = 1, \dots, L$  with respect to  $\mathbf{A}$  and  $\mathbf{Q}$ . This is equivalent to minimizing the negative log of this distribution, which is given through (1) as

$$-\log p(\mathbf{z} | \mathbf{A}, \mathbf{Q}, \mathbf{x}) = \sum_{t=1}^L (\mathbf{z}(t) - \mathbf{A}\mathbf{x}(t))^T \mathbf{Q}^{-1} (\mathbf{z}(t) - \mathbf{A}\mathbf{x}(t)) + \log \det \mathbf{Q} \tag{2}$$

Show that the minimization of (2) with respect to  $\mathbf{A}$  and  $\mathbf{Q}$  yields

$$\hat{\mathbf{A}} = \left( \sum_{t=1}^L \mathbf{z}(t)\mathbf{x}^T(t) \right) \left( \sum_{t=1}^L \mathbf{x}(t)\mathbf{x}^T(t) \right)^{-1} \quad (3)$$

$$\hat{\mathbf{Q}} = \left( \sum_{t=1}^L \mathbf{z}(t)\mathbf{z}^T(t) \right) - \left( \sum_{t=1}^L \hat{\mathbf{A}}\mathbf{x}(t)\mathbf{z}^T(t) \right) \quad (4)$$

respectively. The distribution  $p(\mathbf{z}|\mathbf{A}, \mathbf{Q}, \mathbf{x})$  is referred to as the *likelihood function* and the values given by (3) and (4) are *maximum likelihood estimates*.

- Download the file q1.mat from the course website. The first column of the variable  $z$  is an array of  $x$ -variables and the second column contains the corresponding  $y$ -variables. The  $(x, y)$  pairs are points taken from a circle. Using methods taken in class, calculate the centre and radius of the circle. Explain your method fully. Include your matlab code.
- (From *Applied Numerical Linear Algebra*, by James W. Demmel.)

Let  $\mathbf{A}, \mathbf{B}$  and  $\mathbf{C}$  be matrices with dimensions such that the product  $\mathbf{A}^T\mathbf{C}\mathbf{B}^T$  is defined. Let  $\chi$  be the set of matrices  $\mathbf{X}$  minimizing  $\|\mathbf{A}\mathbf{X}\mathbf{B} - \mathbf{C}\|_F$ , and let  $\mathbf{X}_o$  be the unique member of  $\chi$  minimizing  $\|\mathbf{X}\|_F$ . Show that  $\mathbf{X}_o = \mathbf{A}^+\mathbf{C}\mathbf{B}^+$ . *Hint:* use the SVDs of  $\mathbf{A}$  and  $\mathbf{B}$ .

- Consider the configuration shown in Fig. 1a. The sequence  $x[n]$  is of length  $m$  and  $f_1[n]$  and  $f_2[n]$  are sequences of length  $n \ll m$ . The outputs  $y_1[n]$  and  $y_2[n]$  are the convolution of  $x[n]$  with  $f_1[n]$  and  $f_2[n]$  respectively; i.e.,

$$y_i[n] = \sum_k f_i[k]x[n-k], \quad i \in [1, 2].$$

We observe only the sequences  $y_1[n]$  and  $y_2[n]$ .

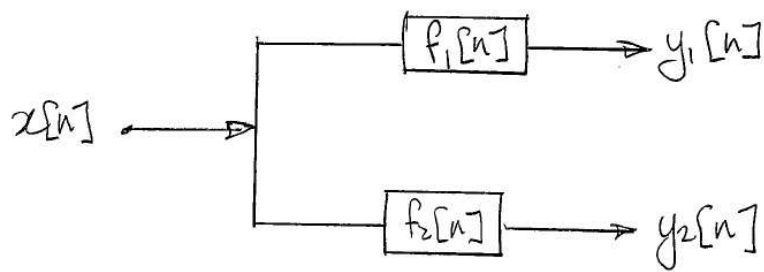
- Show that the convolution operation can be expressed as a matrix–vector multiplication.
  - Show how  $f_1[n]$ ,  $f_2[n]$  and  $x[n]$  can all be determined, based ONLY on the observations  $y_1[n]$  and  $y_2[n]$ . *Hint:* Consider the configuration of Fig. 1b. What are  $w_1$  and  $w_2$  so that the output  $z[n] = 0$ ?
  - Are there any required conditions on  $x[n]$ , or on  $f_1[n], f_2[n]$ ?
- Let the matrix  $\mathbf{X}$  be partitioned in the following blocks:

$$\mathbf{X} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} \quad (5)$$

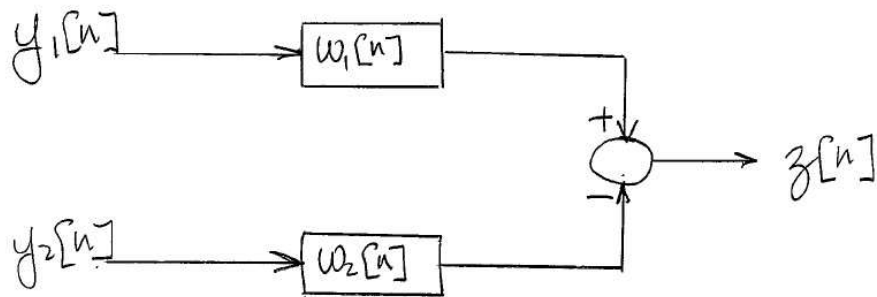
Find a block lower triangular matrix  $\mathbf{L}$  and a block upper triangular matrix  $\mathbf{U}$  so that

$$\mathbf{X} = \mathbf{L} \begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{S}_A \end{bmatrix} \mathbf{U}, \quad (6)$$

where the diagonal blocks of  $\mathbf{L}$  and  $\mathbf{U}$  are the identity, and  $\mathbf{S}_A$  is the *Schur complement* of  $\mathbf{A}$ , given as  $\mathbf{S}_A = \mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B}$ . Then, using this result, find the upper left block of  $\mathbf{X}^{-1}$ .



(a)



(b)

Figure 1: Configuration for Question 4.