UTAH STATE UNIVERSITY
ELECTRICAL and COMPUTER ENGINEERING DEPARTMENT

ECE 6030
Final Examination — 2023

Instructions This is a take-home exam. It is handed out on Wednesday, April 19, 2023 and is due at **5:00 pm on Wednesday April 26, 2023.** No credit can be given if it is late.

This is a take-home exam. The book is allowed, class notes are allowed, the class web page, calculators/computers, homework assignments you have done, and code written for those assignments are allowed. You can use programming environments such as Matlab or Python to assist in your computations. Recorded lectures for this class and lecture notes for this class are allowed. **No other material should be used. Other texts or papers in the literature should not be used.** Use what is in your head!

No comments regarding this test should be directed to anybody at any time except Dr. Moon. This includes other people in the class, other professors, and people who have had the class previously. Even seemingly innocuous comments or questions (such as "Have you finished the first problem yet?") are to be avoided.

Any indication that these instructions have not been followed will result in a score of 0 on this exam and failure in the class.

Your signature below indicates that you have read and complied with the letter and spirit of these instructions.

____

name

____

signature

1. (10 pts) Some interesting matrix identities:

   (a) (5 pts) Let M be a Hermitian symmetric matrix. Show that

   $$U = e^{jM}$$

   is unitary.

   (b) (5 pts) A matrix A is *antihermitian* if $A^H = -A$. Show that if A is antihermitian then

   $$e^A$$

   is unitary.

2. (5 pts) (Inner products; adjoints) Let X be the vector space of polynomials in the real variable t with real coefficients of degree 1 or less. That is, an element $x(t) \in X$ is of the form $x_0 + x_1 t$. Let $Y = \mathbb{R}^2$. Define the linear transformation $L : X \to Y$ by $L(c_0 + c_1 t) = [c_0 + c_1, c_0 - c_1]^T$. Define the inner product associated with the space X by

   $$\langle x(t), y(t) \rangle_X = \int_0^1 x(t)y(t)\, dt.$$

   Define the inner product associated with the space Y by

   $$\langle \mathbf{x}, \mathbf{y} \rangle_Y = x_0 y_0 + x_1 y_1.$$

   For these vector spaces, this linear operator L, and these inner products, explicitly determine: the adjoint L^*, and $\mathcal{R}(L)$ and $\mathcal{R}(L^*)$. Justify your answers.

3. (10 pts) (constrained optimization) Let A be a real $n \times n$ symmetric positive definite matrix with eigendecomposition $A = U\Lambda U^T$, with eigenvalues in Λ ordered from largest to smallest down the diagonal and $U = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_n \end{bmatrix}$. We desire to find an optimum rank-1 approximation to A, which we denote as $\mathbf{b}\mathbf{b}^T$. Then we want to choose $\mathbf{b}$ to minimize

   $$\min_{\mathbf{b}} \|A - \mathbf{b}\mathbf{b}^T\|_F^2$$

   (note the use of the Frobenius norm here). Equivalently, if it helps, you could let $\mathbf{b} = \gamma^{1/2}\mathbf{c}$, where $\mathbf{c}$ has unit length. In this case you would formulate the minimization problem

   $$\min_{\gamma, \mathbf{c}} \|A - \gamma \mathbf{c}\mathbf{c}^T\|_F^2 \text{ subject to } \mathbf{c}^T\mathbf{c} = 1.$$

   You may use either interpretation of this problem.

   Using techniques of vector calculus (e.g., gradients), determine the optimum value of $\mathbf{b}$.

4. (30 pts) (least squares) This problem explores relationships between several least-squares type problems.

   Let A and $\mathbf{b} \neq 0$ be a complex $m \times n$ matrix and an $m \times 1$ vector, respectively, where $m > n+1$. Assume $\operatorname{rank}(A) = n$ and $\mathbf{b} \notin \mathcal{R}(A)$, where $\mathcal{R}(A)$ is the column space (range) of

A. Let $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ and $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ and $\{\sigma_1, \ldots \sigma_n\}$ be the left and right singular vector and singular values of A, ordered such that $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n$. Then $A = \sum_{i=1}^{n} \sigma_i \mathbf{u}_i \mathbf{v}_i^H$.

Let $\{\overline{\mathbf{u}}_1, \ldots, \overline{\mathbf{u}}_m\}$ and $\{\overline{\mathbf{v}}_1, \ldots, \overline{\mathbf{v}}_{n+1}\}$ and $\{\overline{\sigma}_1, \ldots, \overline{\sigma}_{n+1}\}$ be the left and right singular vectors and singular values of the stacked matrix $\begin{bmatrix} A & \mathbf{b} \end{bmatrix}$, so that $\begin{bmatrix} A & \mathbf{b} \end{bmatrix} = \sum_{i=1}^{n+1} \overline{\sigma}_i \overline{\mathbf{u}}_i \overline{\mathbf{v}}_i^H$. (Here the bars, as in $\overline{\mathbf{u}}$, simply mean that these are different values, so that $\mathbf{u}$ is not necessarily the same as $\overline{\mathbf{u}}$.)

Throughout this problem, the norms are l_2 norms, unless otherwise specified.

(a) (5 pts) Under what conditions on α is the matrix $A^H A - \alpha I$ guaranteed to be positive definite? Express your answer in terms of the singular values of A.

(b) (5 pts) Solve the standard least-squares problem for the $n \times 1$ vector $\mathbf{x}$:

$$\min_{\mathbf{x}} \|\mathbf{b} - A\mathbf{x}\|^2.$$

Derive the "normal equations" for this problem, that is, equations of the form $M\mathbf{x} = \mathbf{c}$. What are M and $\mathbf{c}$?. Let $\hat{\mathbf{x}}_{\text{ls}}$ denote the solution to this least-squares problem.

(c) (5 pts) Solve the modified least-squares problem for $\mathbf{x}$:

$$\min_{\mathbf{x}} -\rho\|\mathbf{x}\|^2 + \|\mathbf{b} - A\mathbf{x}\|^2, \tag{1}$$

where ρ is a scalar constant. First derive the "normal equations" for this problem of the form $M\mathbf{x} = \mathbf{c}$. What are M and $\mathbf{c}$ in this case? Denote the solution to the modified least-squared problem as $\hat{\mathbf{x}}_{\text{mls}}$.

(d) (5 pts) You will now show that the total least-squares solution is obtained by solving (1) with the appropriate choice for ρ. We have learned that the total least-squares solution is constructed from $\overline{\mathbf{v}}_{n+1}$, the right singular vector of $\begin{bmatrix} A & \mathbf{b} \end{bmatrix}$, with the solution given by

$$\begin{bmatrix} \hat{\mathbf{x}}_{\text{tls}} \\ -1 \end{bmatrix} = \frac{-1}{\overline{v}_{n+1,n+1}} \overline{\mathbf{v}}_{n+1},$$

where $\overline{v}_{n+1,n+1}$ is the last element of the vector $\overline{\mathbf{v}}_{n+1}$ (assumed to be nonzero). Since $\left[\begin{smallmatrix} \hat{\mathbf{x}}_{\text{TLS}} \\ -1 \end{smallmatrix}\right]$ is a scaled version of a right singular vector, it is an eigenvector of what matrix? What is the corresponding eigenvalue? Write down the appropriate eigenequation to show that $\hat{\mathbf{x}}_{\text{TLS}}$ can be computed directly by

$$\hat{\mathbf{x}}_{\text{TLS}} = (A^H A - \overline{\sigma}_{n+1}^2 I)^{-1} A^H \mathbf{b}.$$

(e) (5 pts) Explain why the total least-squares solution can be calculated from the optimization problem in (1). What value of ρ should be used in (1)?

5. (15 pts) (Generalized MUSIC) Let A be a $m \times n$ full-rank matrix (with $m > n$) which depends upon a parameter vector $\boldsymbol{\phi}$; that is, we write $A = A(\boldsymbol{\phi})$. A signal $\mathbf{x}_t$ has the model

$$\mathbf{x}_t = A(\boldsymbol{\phi})\mathbf{s}_t + \mathbf{n}_t$$

where the vector $\mathbf{s}_t$ has uncorrelated elements, so that

$$E[\mathbf{s}_t\mathbf{s}_t^H] = \operatorname{diag}(p_1, p_2, \dots, p_n) = P.$$

The quantity $A(\boldsymbol{\phi})\mathbf{s}_t$ is the signal component of the measured data, and $\mathbf{n}_t$ is a vector of zero-mean white noise with covariance $\sigma_n^2 I$. The overall goal is to estimate $\boldsymbol{\phi}$.

(a) (5 pts) Show that $R_x = E[\mathbf{x}_t\mathbf{x}_t^H] = APA^H + \sigma_n^2 I$.

(b) (5 pts) Let $R_x = U\Lambda U^H$ be an eigendecomposition, and let the eigenvalues be ordered on the diagonal of Λ in non-increasing order: $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m$. Partition the eigenvalues and eigenvectors into two sets: (1) the signal subspace and (2) the noise subspace. How many eigenvalues are in each set? What are the noise eigenvalues equal to?

(c) (5 pts) Let U_s be the matrix with columns formed from the signal eigenvectors. Show that $\mathcal{R}(U_s) = \mathcal{R}(A(\boldsymbol{\phi}))$.

Note what this means: By finding the eigendecomposition of R_x, we are able to find the space where the vectors of $A(\boldsymbol{\phi})$ live, even if (to this point) we don't know exactly what those vector are.

6. (10 pts) Let $\mathbf{x}$ be a 2×1 random vector with mean $E\mathbf{x} = 0$ and covariance matrix $R = E[\mathbf{x}\mathbf{x}^T]$. Let R have the eigendecomposition

$$R = U\Lambda U^T = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\frac{1}{\sqrt{2}}.$$

Suppose that $\mathbf{x}$ is observed when corrupted by noise,

$$\mathbf{y} = \mathbf{x} + \mathbf{v}$$

where $\mathbf{v}$ is a random vector with mean $E[\mathbf{v}] = 0$ and covariance $E[\mathbf{v}\mathbf{v}^T] = \sigma_v^2 I$.

We take the observed vector $\mathbf{y}$ and want to process it to obtain a scale measurement, using a vector $\mathbf{w}$:

$$z = \mathbf{w}^T\mathbf{y} = \mathbf{w}^T\mathbf{x} + \mathbf{w}^T\mathbf{v}.$$

We desire to choose the vector $\mathbf{w}$ to maximize the signal "strength" to the noise "strength", where the signal part is

$$z_{\text{signal}} = \mathbf{w}^T\mathbf{x}$$

and the noise part is

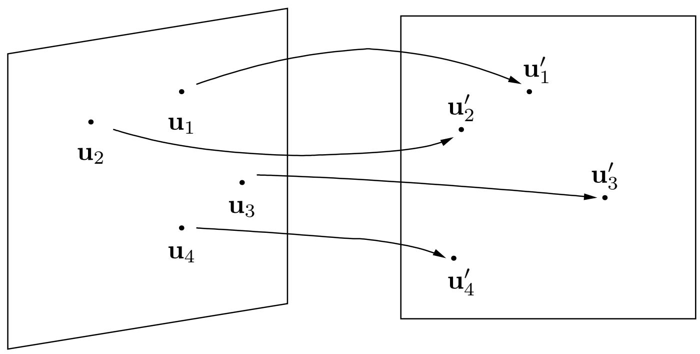
$$z_{\text{noise}} = \mathbf{w}^T\mathbf{v}.$$

More specifically, we desire to choose $\mathbf{w}$ to maximize the signal to noise ratio

$$\text{SNR} = \frac{E[z_{\text{signal}}^2]}{E[z_{\text{noise}}^2]}$$

What is the unit-length $\mathbf{w}$ that maximizes the SNR? What is the corresponding maximum SNR?

7. (15 pts) (Ability to read and apply new concepts). Suppose that there is a collection of m points in space, which when viewed in one coordinate system have coordinates $\mathbf{u}'_i$, $i = 1, 2, \ldots, m$. Suppose that these same m points, when viewed in another coordinate system, have coordinates $\mathbf{u}_i$, $i = 1, 2, \ldots, m$. For example, the coordinates $\mathbf{u}'_i$ might be the positions of certain objects on one map, and the $\mathbf{u}_i$ might be the positions of those same objects, but on a different map that uses a different coordinate system, as suggested by the following figure:



The goal of this problem is to develop a means of mapping between these coordinate systems. This kind of problem arises in quantitative measurements using cameras (photogrammetry), when the internal camera coordinates system are calibrated against the real world.

The sets of coordinates have correspondences, so that $\mathbf{u}'_i$ corresponds to (represents the same point as) $\mathbf{u}_i$, for $i = 1, 2, \ldots, m$.

In this problem, projective coordinates are used. In projective coordinates, a point in d-dimensional space is represented using a vector with $d + 1$ elements. To begin, the point (u, v) (in two dimensions) is represented using the triple $(u, v, 1)$. Then, under the projective coordinate framework, any triple which is a non-zero multiple of this point represents the same point. That is, the triple $\lambda(u, v, 1)$ represents the point (u, v), for any $\lambda \neq 0$.

Since there is no absolute scale relating points in projective coordinates, the relationship between corresponding points has an extra scale factor. The *homography* between the set of points $\mathbf{u}'_i$ and $\mathbf{u}_i$ is the matrix H such that

$$\alpha_i \mathbf{u}'_i = H\mathbf{u}_i, \qquad i = 1, 2, \ldots, m.$$

Here, α_i is a scale parameter α_i, which introduces an extra degree of freedom in to the computations to allow for scale factors.

To eliminate the dependence on α in the equation of the form $\alpha\mathbf{u}' = H\mathbf{u}$, we can multiply on the left by a matrix $G(\mathbf{u}')$ whose rows are orthogonal to $\mathbf{u}'$:

$$G(\mathbf{u}')\alpha\mathbf{u}' = \alpha G(\mathbf{u}')\mathbf{u}' = 0 = G(\mathbf{u}')H\mathbf{u}.$$

One way to obtain this matrix (which works in three dimensions) is to use the matrix representing the cross product. For two vectors $\mathbf{u}$ and $\mathbf{v}$, the vector cross product $\mathbf{u} \times \mathbf{v}$ produces another vector. If $\mathbf{u} = [u, v, w]^T$, then the cross product $\mathbf{u} \times \mathbf{v}$ can be computed by

$$\mathbf{u} \times \mathbf{v} = \begin{bmatrix} 0 & -w & v \\ w & 0 & -u \\ -v & u & 0 \end{bmatrix} \mathbf{v}.$$

It can be shown that for any vector, the cross product with itself is equal to 0. That is, $\mathbf{u} \times \mathbf{u} = \mathbf{0}$. We will denote the cross product matrix by $G(\mathbf{u})$:

$$G(\mathbf{u}) = \begin{bmatrix} 0 & -w & v \\ w & 0 & -u \\ -v & u & 0 \end{bmatrix}. \tag{2}$$

This matrix is formulated to compute the cross product, so that $\mathbf{u} \times \mathbf{v} = G(\mathbf{u})\mathbf{v}$.

Applying this cross product matrix to each correspondence relationship $\alpha_i \mathbf{u}'_i = H\mathbf{u}_i$, we obtain the equations

$$\begin{aligned} G(\mathbf{u}'_1)H\mathbf{u}_1 &= 0 \\ G(\mathbf{u}'_2)H\mathbf{u}_2 &= 0 \\ &\vdots \\ G(\mathbf{u}'_m)H\mathbf{u}_m &= 0 \end{aligned} \tag{3}$$

These equations are linear in the matrix H, but the unknown matrix H appears in the middle of the expression.

We now refer the student to equation (9.19) of the book, which says that for matrices A, Y, and B,

$$\text{vec}(AYB) = (B^T \otimes A)\,\text{vec}\,Y.$$

Here, $\text{vec}(\cdot)$ means to stack its argument up by columns. For example, for the matrix $X = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$,

$$\text{vec}(X) = \begin{bmatrix} a \\ c \\ b \\ d. \end{bmatrix}$$

The notation $\otimes$ indicates the Kronecker product. (See section 9.1).

Let us apply this identity to each of the equations in (3)

$$\begin{aligned} (\mathbf{u}_1^T \otimes G(\mathbf{u}'_1))\,\mathbf{h} &= \mathbf{0} \\ (\mathbf{u}_2^T \otimes G(\mathbf{u}'_2))\,\mathbf{h} &= \mathbf{0} \\ &\vdots \\ (\mathbf{u}_m^T \otimes G(\mathbf{u}'_m))\,\mathbf{h} &= \mathbf{0}, \end{aligned}$$

where $\mathbf{h} = \text{vec}(H)$. Now the matrix of interest appears (reshaped as a vector) on the right. Since $\mathbf{h}$ is common to each equation, this can be written as

$$\begin{bmatrix} \mathbf{u}_1^T \otimes G(\mathbf{u}'_1) \\ \mathbf{u}_2^T \otimes G(\mathbf{u}'_2) \\ \vdots \\ \mathbf{u}_m^T \otimes G(\mathbf{u}'_m) \end{bmatrix} \mathbf{h} = \mathbf{0}.$$

Denote the matrix here by W. We thus have the expression $W\mathbf{h} = 0$. The $\mathbf{h}$ we are seeking is a nonzero vector that lies in the nullspace of the matrix W.

However, in practice the measurements of the coordinates may be noisy, so that W is a full rank matrix (only the trivial nullspace), so that there is no exact solution. In this case, we want to find a nonzero vector $\mathbf{h}$ so that $W\mathbf{h}$ is "small." This leads finally to posing the following problem:

$$\text{minimize} \|W\mathbf{h}\|^2 \text{ subject to } \|\mathbf{h}\|^2 = 1.$$

The constraint $\|\mathbf{h}\|^2 = 1$ is imposed to avoid the trivial solution $\mathbf{h} = 0$. (Since the original equations have the α_i scaling factors, the actual norm of $\mathbf{h}$ is immaterial, as long as it is not equal to 0.)

(a) (5 pts) Show that the matrix $G(\mathbf{u})$ defind in (2) satisfies $G(\mathbf{u})\mathbf{u} = \mathbf{0}$.

(b) (10 pts) Show how to use the SVD to solve the constrained optimization problem

$$\text{minimize} \|W\mathbf{h}\|^2 \text{ subject to } \|\mathbf{h}\|^2 = 1.$$