EE263 Homework 6

Fall 2023

3.660. Some true/false questions. Determine if the following statements are true or false. No justification or discussion is needed for your answers. What we mean by "true" is that the statement is true for all values of the matrices and vectors given. You can't assume anything about the dimensions of the matrices (unless it's explicitly stated), but you can assume that the dimensions are such that all expressions make sense. For example, the statement "$A+B = B+A$" is true, because no matter what the dimensions of A and B (which must, however, be the same), and no matter what values A and B have, the statement holds. As another example, the statement $A^2 = A$ is false, because there are (square) matrices for which this doesn't hold. (There are also matrices for which it does hold, *e.g.*, an identity matrix. But that doesn't make the statement true.)

a) If $x^T Ax = x^T Bx$ for all x, then $A = B$.

b) If $x^T Ay = x^T By$ for all x and y, then $A = B$.

c) If $\|Ax\| = \|Bx\|$ for all x, then $A = B$.

d) If A and B are both stable, then $A + B$ is also stable.

e) The matrix $\begin{bmatrix} 2a & 3b \\ 4c & 5d \end{bmatrix}$ is equal to $A \begin{bmatrix} a & b \\ c & d \end{bmatrix} B$ for some matrices A and B.

f) If R is upper triangular and orthogonal, then R is diagonal.

g) If A is square, then there always exists a matrix C such that $AC = CA^T$.

h) If $x, y \in \mathbf{R}^n$ then the $n \times n$ matrix xy^T is diagonalizable.

8.160. Designing an equalizer for backwards-compatible wireless transceivers. You want to design the equalizer for a new line of wireless handheld transceivers (more commonly called walkie-talkies). The transmitter for the new line of transceivers has already been designed (and cannot be changed) – if the input signal is $x \in \mathbb{R}^n$, then the transmitted signal is $y = A_{\text{new}}x \in \mathbb{R}^m$, where $A_{\text{new}} \in \mathbb{R}^{m\times n}$ is known. An equalizer for A_{new} is a matrix $B \in \mathbb{R}^{n\times m}$ such that $By = x$ for every $x \in \mathbb{R}^n$.

The new line of transceivers will replace an older model. Given an input signal $x \in \mathbb{R}^n$, the old line of transceivers transmit a signal $y_{\text{old}} = A_{\text{old}}x \in \mathbb{R}^m$, where $A_{\text{old}} \in \mathbb{R}^{m\times n}$ is known. In addition to providing exact equalization for the new line of transceivers, you want your equalizer to be able to at least partially equalize signals transmitted using the old line of transceivers. In other words, to the extent that it is possible, you want the new line of transceivers to be backwards compatible with the old line of transceivers.

a) Explain how to find an equalizer B that minimizes

$$J = \|BA_{\text{old}} - I\|_F^2 = \sum_{i=1}^{n}\sum_{j=1}^{n}(BA_{\text{old}} - I)_{ij}^2$$

among all B that exactly equalize A_{new}. Such a B is an exact equalizer for A_{new}, and an approximate equalizer for A_{old}. State any assumptions that are needed for your method to work.

b) The file `backwards_compatible_transceiver_data.json` defines the following variables.

- `Anew`, the $m \times n$ matrix that describes the transmitter used in the new line of transceivers
- `Aold`, the $m \times n$ matrix that describes the transmitter used in the old line of transceivers
- `x`, a vector of length n that serves as an example input signal

Apply your method to this example data. Report the optimal value of J. The pseudoinverse $A_{\text{new}}^{\dagger}$ is another exact equalizer for A_{new}. Compare the optimal value of J, and the value of J achieved by $A_{\text{new}}^{\dagger}$.

c) The example signal x defined in the data file is a binary signal. Form the signal $y_{\text{old}} = A_{\text{old}}x$ transmitted by the old line of transceivers, and construct an estimate of x by equalizing y_{old} using B, and then rounding the result to a binary signal. More concretely, compute the estimate $\hat{x} \in \mathbb{R}^n$, where

$$\hat{x}_i = \begin{cases} 1 & (By_{\text{old}})_i > \frac{1}{2}, \\ 0 & \text{otherwise.} \end{cases}$$

Report the bit error rate of your estimate, which is defined as

$$\frac{1}{n}\sum_{i=1}^{n} \mathbf{I}(x_i \neq \hat{x}_i),$$

where $\mathbf{I}(x_i \neq \hat{x}_i)$ is an indicator function:

$$\mathbf{I}(x_i \neq \hat{x}_i) = \begin{cases} 1 & x_i \neq \hat{x}_i, \\ 0 & \text{otherwise.} \end{cases}$$

Similarly, report the bit error rate if $A_{\text{new}}^{\dagger}$ is used as the equalizer.

15.2150. Norm expressions for quadratic forms. Let $f(x) = x^{\mathsf{T}}Ax$ (with $A = A^{\mathsf{T}} \in \mathbb{R}^{n\times n}$) be a quadratic form.

a) Show that f is positive semidefinite (*i.e.*, $A \geq 0$) if and only if it can be expressed as $f(x) = \|Fx\|^2$ for some matrix $F \in \mathbb{R}^{k\times n}$. Explain how to find such an F (when $A \geq 0$). What is the size of the smallest such F (*i.e.*, how small can k be)?

b) Show that f can be expressed as a difference of squared norms, in the form $f(x) = \|Fx\|^2 - \|Gx\|^2$, for some appropriate matrices F and G. How small can the sizes of F and G be?

15.2790. Ellipsoids.

a) Write a function that, given a a 2×2 real, positive definite symmetric matrix $A > 0$, plots the ellipse

$$E = \{ x \in \mathbb{R}^2 \mid x^T A x = 1 \}$$

Make sure that your plot is shown so that horizontal and vertical lengths are the same, that is, with aspect ratio 1. Turn in your code.

Julia hint: use the following to draw a plot with correct aspect ratio.

```
Using Plots;  plot(x, y, aspect_ratio=:equal)
```

b) Use your code to plot the ellipsoid for the matrix

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$$

c) Use your code to plot the ellipsoid for the matrix

$$A = \begin{bmatrix} 0.2 & -0.1 \\ -0.1 & 0.4 \end{bmatrix}$$

On your plot, also show semiaxes.

d) Consider an estimation problem, where we have three sensors, define by $b_i \in \mathbb{R}^2$ for $i = 1, 2, 3$. We measure $y_i = b_i^\mathsf{T} x$. The vectors b_i are

$$b_1 = \begin{bmatrix} 0.89 \\ 0.45 \end{bmatrix} \quad b_2 = \begin{bmatrix} 0.45 \\ 0.89 \end{bmatrix} \qquad b_3 = \begin{bmatrix} -0.71 \\ 0.71 \end{bmatrix}$$

Plot the set of $x \in \mathbb{R}^2$ for which $\|y\| \leq 1$. On your plot, show also the b_i (that is, plot a line from the origin to b_i).

16.2980. Smoothing. We have a discrete-time signal given by $x \in \mathbb{R}^n$. We get to measure $y \in \mathbb{R}^n$, given by

$$y_i = \sum_{k=-h}^{h} c_k x_{i+k} + w_i \qquad \text{for } i = 1, \ldots, n$$

where w_i is noise. Here we use the convention that $x_i = 0$ for $i < 1$ or $i > n$. That is, y is c convolved with x plus noise. In applications, very often the effect of convolution with c is to smooth or blur x, and we would like to undo this.

The file `regl_data.json` contains c, w and x.

a) In Julia, construct the $n \times n$ matrix such that $y = Ax + w$. Plot the singular values σ_k against k.

b) Plot the first 6 right singular vectors of A (i.e. plot V_{ij} against i for $j = 1, \ldots, 6$.) Explain what you see.

c) Find and plot the least-squares estimate of x given y_{meas}, computng y_{meas} using c, x and w given in `regl_data.json`. Explain what happens.

d) Many of the singular values of A are very small; this means that the measurement in the directions of the corresponding right singular vectors is being swamped by the noise.

If we believe these components are small, we can remove them from our estimate of x altogether by *truncating* the SVD of A and using the truncated SVD to compute the estimate. This is called the *truncated SVD regularization* of least-squares.

Suppose we decided only to keep the first r components. Then truncate by letting $\tilde{V}$ and $\tilde{U}$ be the first r columns of V and U, and letting $\tilde{\Sigma}$ be the top-left $r \times r$ submatrix of Σ. Then we can construct an estimator that ignores the noise components by

$$A_{\text{est}} = \tilde{V}\tilde{\Sigma}^{-1}\tilde{U}^T$$

and set

$$x_{\text{est}} = A_{\text{est}} y_{\text{meas}}$$

For values of r in $5, 10, 15, 30, 50$, compute and plot the corresponding estimates of x. Explain what you see.

e) For each r between 1 and 35, compute the norm of the error

$$\|x - x_{\text{est}}\|$$

Plot this against r. Explain what you see.

f) Pick the 'best' r and plot the corresponding estimate.

g) Another approach is to use Tychonov regularization. Find and plot the vector $x_{\text{reg}} \in \mathbb{R}^n$ that minimizes the function

$$\|Ax - y\|^2 + \mu\|x\|^2,$$

where $\mu > 0$ is the regularization parameter. Pick a value of μ that gives a good estimate, in your opinion.

h) The regularized solution is a linear function of y, so it can be expressed as $x_{\text{reg}} = By$ where $B \in \mathbb{R}^{n\times n}$. Express the SVD of B in terms of the SVD of A. To be more specific, let

$$B = \sum_{i=1}^{n} \tilde{\sigma}_i \tilde{u}_i \tilde{v}_i^T$$

denote the SVD of B. Express $\tilde{\sigma}_i$, $\tilde{u}_i$, $\tilde{v}_i$, for $i = 1, \ldots, n$, in terms of σ_i, u_i, v_i, $i = 1, \ldots, n$ (and, possibly, μ). Recall the convention that $\tilde{\sigma}_1 \geq \cdots \geq \tilde{\sigma}_n$.

i) Find the norm of B. Give your answer in terms of the SVD of A (and μ).

j) Find the worst-case relative inversion error, defined as

$$\max_{y\neq 0} \frac{\|ABy - y\|}{\|y\|}.$$

Give your answer in terms of the SVD of A (and μ).