EE263 Homework 5

Summer 2022

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.

8.1231. Invertibility of the KKT matrix. Recall the linearly-constrained norm minimization problem we saw in lecture: Given $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, $C \in \mathbb{R}^{k\times n}$, and $d \in \mathbb{R}^k$, find $x \in \mathbb{R}^n$ to

$$\begin{array}{ll} \text{minimize} & \|Ax - b\| \\ \text{subject to} & Cx = d. \end{array}$$

We assume that $k \leq n$, so there are at least as many variables as equality constraints. As discussed in lecture, using the theory of constrained optimization, a necessary condition for a point $x \in \mathbb{R}^n$ to be a minimizer of this problem is for there to exist $\lambda \in \mathbb{R}^k$ (the Lagrange multiplier) so that x, λ solve:

$$\begin{bmatrix} A^{\mathsf{T}}A & C^{\mathsf{T}} \\ C & 0 \end{bmatrix}\begin{bmatrix} x \\ \lambda \end{bmatrix} = \begin{bmatrix} A^{\mathsf{T}}b \\ 0 \end{bmatrix}.$$

This system is called the Karush-Kuhn-Tucker (KKT) system, and the matrix

$$K = \begin{bmatrix} A^{\mathsf{T}}A & C^{\mathsf{T}} \\ C & 0 \end{bmatrix} \in \mathbb{R}^{(n+k)\times(n+k)}$$

is called the KKT matrix. In this problem you will fill in the proof that K is invertible if and only if C has linearly independent rows, and

$$S = \begin{bmatrix} A \\ C \end{bmatrix}$$

has linearly independent columns. Notice that if $A = I$, then S always has linearly independent columns and K is invertible if and only if C has linearly independent rows (which covers the case of the least norm problem studied in lecture).

a) Show that $\text{null}(A) \cap \text{null}(C) = \{0\}$ if and only S has linearly independent columns.

b) Show that if C is does not have linearly independent rows, then K is singular (i.e., not invertible).

c) Show that if there exists a nonzero vector $u \in \text{null}(A) \cap \text{null}(C)$, then K is singular.

d) Show that if K is singular, then either C does not have linearly independent rows, or there exists a nonzero vector $u \in \text{null}(A) \cap \text{null}(C)$.

8.1320. Portfolio selection with sector neutrality constraints. We consider the problem of selecting a portfolio composed of n assets. We let $x_i \in \mathbb{R}$ denote the investment (say, in dollars) in asset i, with $x_i < 0$ meaning that we hold a short position in asset i. We normalize our total portfolio as $\mathbf{1}^\mathsf{T} x = 1$, where $\mathbf{1}$ is the vector with all entries 1. (With normalization, the x_i are sometimes called *portfolio weights*.)

The portfolio (mean) return is given by $r = \mu^\mathsf{T} x$, where $\mu \in \mathbb{R}^n$ is a vector of asset (mean) returns. We want to choose x so that r is large, while avoiding risk exposure, which we explain next.

First we explain the idea of *sector exposure*. We have a list of k economic sectors (such as manufacturing, energy, transportation, defense, ...). A matrix $F \in \mathbb{R}^{k\times n}$, called the *factor loading matrix*, relates the portfolio x to the *factor exposures*, given as $R^{\text{fact}} = Fx \in \mathbb{R}^k$. The number R_i^{fact} is the portfolio risk exposure to the ith economic sector. If R_i^{fact} is large (in magnitude) our portfolio is exposed to risk from changes in that sector; if it is small, we are less exposed to risk from that sector. If $R_i^{\text{fact}} = 0$, we say that the portfolio is *neutral* with respect to sector i.

Another type of risk exposure is due to fluctations in the returns of the individual assets. The *idiosyncratic risk* is given by

$$R^{\text{id}} = \sum_{i=1}^n \sigma_i^2 x_i^2,$$

where $\sigma_i > 0$ are the standard deviations of the asset returns. (You can take the formula above as a definition; you do not need to understand the statistical interpretation.)

We will choose the portfolio weights x so as to maximize $r - \lambda R^{\text{id}}$, which is called the *risk-adjusted return*, subject to neutrality with respect to all sectors, *i.e.*, $R^{\text{fact}} = 0$. Of course we also have the normalization constraint $\mathbf{1}^\mathsf{T} x = 1$. The parameter λ, which is positive, is called the *risk aversion parameter*. The (known) data in this problem are $\mu \in \mathbb{R}^n$, $F \in \mathbb{R}^{k\times n}$, $\sigma = (\sigma_1, \ldots, \sigma_n) \in \mathbb{R}^n$, and $\lambda \in \mathbb{R}$.

a) Explain how to find x, using methods from the course. You are welcome (even encouraged) to express your solution in terms of block matrices, formed from the given data.

b) Using the data given in `sector_neutral_portfolio_data.json`, find the optimal portfolio. Report the associated values of r (the return), and R^{id} (the idiosyncratic risk). Verify that $\mathbf{1}^\mathsf{T} x = 1$ (or very close) and $R^{\text{fact}} = 0$ (or very small).

8.1460. Filling-in missing data. In this problem we have a signal, $y_i \in \mathbb{R}$ for $i = 1, \ldots, n$, which we view as $y \in \mathbb{R}^n$. We will have $n = 100$. The signal y comes from measurements of a physical system, and so y_{i+1} is measured a short time interval after y_i. Unfortunately, during the data acquisition process some of the data was lost and so the signal we have has gaps in it. Specifically, we have a known set $K \subset \mathbb{Z}$ and we know y_i only for values $i \in K$.

The data for this problem is in the file `missing.json`. The supplied vector `known` contains the list of known points K, and the vector `yknown` is the list of values of y at the points in K. The length of `yknown` is therefore $|K|$.

a) For a signal $z \in \mathbb{R}^n$, we define the discrete derivative $z^{\text{der}} \in \mathbb{R}^{n-1}$ by

$$z_i^{\text{der}} = z_{i+1} - z_i \quad \text{for } i = 1, \ldots, n-1$$

Find the matrix G such that $z^{\mathrm{der}} = Gz$

b) Our first approach will be to find the signal z which minimizes $\|z^{\mathrm{der}}\|$ and satisfies

$$z_i = y_i \quad \text{if } i \in K$$

Give a method finding the optimal z.

c) Find the optimal z in the previous part and plot z_i against i. Be sure to plot the points (i, z_i), not just a line joining them.

d) One way to do a better job at filling in the missing data is to put additional criteria on our estimate. Here we will do this by additionally penalizing the second derivative of z. Define the discrete second derivative $z^{\mathrm{hes}} \in \mathbb{R}^{n-2}$ by

$$z_i^{\mathrm{hes}} = z_{i+2} - 2z_{i+1} + z_i \quad \text{for } i = 1, \ldots, n-2$$

Find the matrix H such that $z^{\mathrm{hes}} = Hz$

e) Define the two objective functions

$$J_1 = \|Gz\|^2 \qquad J_2 = \|Hz\|^2$$

We would like to find the signal z that minimizes

$$J_1 + \mu J_2$$

and satisfies

$$z_i = y_i \quad \text{if } i \in K$$

Give a method for finding the optimal z.

f) Plot the trade-off curve of J_2 (on the vertical axis) versus J_1 (on the horizontal axis). Give the interpretation of the endpoints of this curve.

g) Find the optimal z for the three different cases $\mu = 5, 20, 100$.

11.1900. Some basic properties of eigenvalues. Show the following:

a) The eigenvalues of A and A^{T} are the same.

b) A is invertible if and only if 0 is not an eigenvalue of A.

c) If A is invertible, and its eigenvalues are $\lambda_1, \ldots, \lambda_n$, then the eigenvalues of A^{-1} are $1/\lambda_1, \ldots, 1/\lambda_n$.

d) The eigenvalues of A and $T^{-1}AT$ are the same.

e) Suppose $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{n \times m}$. Show that if $\lambda \in \mathbb{C}$ is a nonzero eigenvalue of AB,then λ is also an eigenvalue of BA. Conclude that the nonzero eigenvalues of AB and BA are the same.

You may use the following without proof: $\det A = \det(A^{\mathsf{T}})$, $\det A = \det(-A)$, $\det(AB) = \det A \det B$, and, if A is invertible, $\det A^{-1} = 1/\det A$.