Modeling Spr 23, Take Home Exam 1

INSTRUCTIONS

On the class website, you'll find a starting script for each of the three questions below, and a data set for the first two questions. You should download all of these to the same directory where you'll be working (either to your home directory for Matlab or upload to Octave-Online).

Be sure you start early enough so that if you get stuck, or have technology issues, you still have enough time to ask me.

You may use our text, class notes, or anything on our current class page. Do not use anything else on the internet (ask me first if you need something and you're not sure). You may work with one other person, and both of you should upload your solutions.

Solutions should be uploaded before 11:59PM on Wednesday, Feb 22 to Canvas.

Questions

1. On the class website, there is a Matlab data file called `Exam1Q1.mat`. It contains one set of 100 data triples, (x_i, y_i, z_i) in the array `X`. The data is organized so that `X` is 100×3.

   It also contains one set of 10 ordered pairs, (x_i, y_i) in the array `Xtest`, so that matrix will be 10×2.

   We want to find parameters a, b, c, d, e so that

   $$z_i \approx ax_i^2 + bx_iy_i + cy_i + d\sin(y_i) + e$$

   where the parameters minimize the squared error between the output and the z_i.

   Use these parameters to predict the output for the data in `Xtest`, and your script file should print these 10 values to the screen (or onto the published text).

   See the $m-$file: `Exam1THQ1.m` to help you get started.

   HINT: Try to express the problem as $A\mathbf{x} = \mathbf{b}$ for an appropriate matrix A and vectors $\mathbf{x}$ and $\mathbf{b}$ **before** you sit down with Matlab. Similarly, once a, b, c, d, e have been computed, how should you compute the value of of the function for the new data? (Construct a new matrix "A").

2. This question asks you to verify some of the computations we made in the text. To get the matrix X, download the matlab file `Exam1Q2.mat` from the class website, and in your script file, type: `load Exam1Q2`

   See the $m-$file: `Exam1THQ2.m` to help you get started.

   (You should see that a matrix X that is 5×100 was loaded into memory- We'll treat this as 100 points in $\mathbb{R}^5$)

(a) In the script file, type in the command that will double center the matrix. Call the new matrix Xm (so that Xm is 5×100).

(b) Compute the covariance matrix C using matrix multiplication (not the Matlab command `cov`). Be sure the size is right- Xm is a matrix formed from 100 points in $\mathbb{R}^5$.

(c) Find the value of the largest eigenvalue of C (You might use the Matlab command `eig` and the command `sort`).

(d) Find the value of the largest singular value of Xm.

(e) Using the SVD of Xm, we want to compute the coordinates of the projection of the 100 points to the first column $U(:, 1)$. (Your result should be a vector, either 1×100 or 100×1).

(f) Find the variance of the data in the previous part (100 points in $\mathbb{R}^1$) using vector computations (not the variance command).

(HINT: You should once again sit down and work these out using linear algebra notation before you sit down at Matlab!)

3. Let X be a matrix of 300 random points from $\mathbb{R}^{30}$.

- Mean subtract the data.
- Compute the SVD of the mean subtracted data.
- Project the data to the first two columns of U (from the SVD), and plot the data (in $\mathbb{R}^2$). For example, if we have 25 points in $\mathbb{R}^2$ in a matrix A that is 2×25, then the command would be

```
plot(A(1,:),A(2,:),'k*')
```