Math 160 (De Loera) **Homework 1**
Due date: April 13, 2014

INSTRUCTIONS

All homeworks will have many problems, both theoretical and practical. Programming exercises need to be submitted via SMARTSITE using the assignment boxes.

Write legibly preferably using word processing if your hand-writing is unclear. Be organized and use the notation appropriately. Show your work on every problem. Correct answers with no support work will not receive full credit.

- **Modeling an Advertising problem**. A company wants to promote its newly developed product by launching an advertising campaign. There are four advertising options to choose from: TV Spot, Newspaper, Radio (prime time), and Radio (afternoon); these options are labelled T, N, P, A respectively. The table below provides, for each type of advertising, the audience reached, the cost and maximum number of ads per week. The company has a budget of 8000 dolars per week and seeks to maximize audience reached. However, the company also wants 5 or more radio spots per week and cannot spend more than 1800 dollars on radio per week. Let T, N, P, A be the decision variables corresponding to the numbers of ads chosen weekly by the company. Formulate this as a linear integer programming problem in SCIP making sure to incorporate all the constraints in the formulation! Solve the problem using SCIP.

Advertising options	TV Spot (T)	Newspaper (N)	Radio (P) (prime time)	Radio (A) (afternoon)
Audience Reached (per ad)	5000	8500	2400	2800
Cost (per ad)	\$ 800	\$ 925	\$ 290	\$ 380
Max Ads (per week)	12	5	25	20

- This exercise is to help you understand the idea of Support Vector Machines (SVM) and the power of linear optimization.

  In this project, we will use linear programming for breast cancer diagnosis. The project will use the Wisconsin Diagnosis Breast Cancer Database (WDBC). The idea is to come up with a discriminant function (a separating plane in this case) to determine if an unknown sample is benign or malignant. In order to do this, you will use part of the data in the above database as a training set to generate your separating plane and the remaining part as a testing set to test your separating plane. Attributes 3 to 32 form a 30-dimensional vector representing each case as a point in 30-dimensional real space R^{30}. To generate the separating plane, a training set, consisting of two disjoint point sets B and M in R^{30} representing confirmed benign and malignant cases. The separating plane, to be determined by solving a single linear program in MATLAB or SCIP is based on the formulation proposed in class.

  1. Formulate the problem as a linear program. Solve the problem using the M and B as a training set from the first 500 cases of the wdbc.data file available from

https://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+(Diagnostic)
The last 69 points should be used as a testing set. Solve the linear program and print out the separating hyperplane, and the minimum value of the LP.

2. Test the separating plane on the 69 cases of the testing set. Report the number of misclassified points on the testing set. It is probably a good idea if you create an MATLAB file to do this.
3. Suppose that the oncologist wants to use only 2 of the 30 attributes in her diagnosis. Determine which pair of attributes is most effective in determining a correct diagnosis as follows. Use each of the possible pairs of possible attributes and for each pair determine from the training set a separating plane in R^2 (you are solving several linear programs to do this!) For each plane use the testing set with the corresponding pair of attributes to determine the number of misclassified cases. Print out the number of misclassified points in the testing set for each pair of attributes. What is the best pair of attributes to make a diagnosis?
4. We saw in class that there is a a *quadratic programming* formulation of the separation problem. Formulate the quadratic program in MATLAB or SCIP and try to solve it that way. Is the separating hyperplane better?

- This exercise will help you practice SVD (singular value decomposition)

Problem 1. Determine SVDs of the following matrices by hand calculation:

$$\text{(a)} \begin{bmatrix} 3 & 0 \\ 0 & -2 \end{bmatrix}, \quad \text{(b)} \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}, \quad \text{(c)} \begin{bmatrix} 0 & 2 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad \text{(d)} \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \quad \text{(e)} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}.$$

[Hint: Let $A = U\Sigma V^\top$. Then, if you express $A^\top A$ using the SVD of A, then you see that σ_j^2, $\boldsymbol{v}_j$ are the jth eigenvalue and the corresponding eigenvector of $A^\top A$.]

Problem 2. Let $A \in \mathbb{R}^{m\times n}$, and suppose $B \in \mathbb{R}^{n\times m}$ is obtained by rotating A 90 degree clockwise on paper. More precisely,

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \qquad B = \begin{bmatrix} a_{m1} & \cdots & a_{21} & a_{11} \\ \vdots & \vdots & \vdots & \vdots \\ a_{mn} & \cdots & a_{2n} & a_{1n} \end{bmatrix}.$$

Do A and B have the same singular values? Prove that the answer is yes or give a counterexample.

[Hint: Express B as a product of $A^\top$ and a matrix P that permutes the column vectors of $A^\top$.]

Problem 3. Write a MATLAB program which, given a real 2×2 matrix A, plots the right singular vectors $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ in the unit circle and also the left singular vectors $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ in the appropriate ellipse, as in the figure of my Lecture 13. Apply your program to the matrices in Problem 1, and attach the resulting figures to your homework submission.

Problem 4 Two matrices $A, B \in \mathbb{R}^{m\times m}$ are *orthogonally equivalent* if $A = QBQ^\top$ for some orthogonal matrix Q. Is it true or false that A and B are orthogonally equivalent if and only if they have the same singular values?

- Download the image called mandril.mat (available at `https://www.math.ucdavis.edu/~deloera/TEACHING/MATH160/mandril.mat`) using the following MATLAB command (This loads a matrix X containing a face of a cute mandrill, and a map containing a colormap of the image.) *load mandrill;*

Display this matrix on your screen by:

```
>> image(X); colormap(map)
```

Then, attach it in your HW sheets.

(b) Compute the SVD of this mandrill image and plot the distribution of its singular values on your screen (Note that the MATLAB `svd` function returns three matrices U, S, V for a given input matrix. So, the singular values are nicely plotted by:

```
>> stem(diag(S)); grid
```

Then print this figure and attach it in your HW sheets.

(c) Let σ_j, $\boldsymbol{u}_j$, $\boldsymbol{v}_j$ be the jth singular value, the jth left and right singular vectors of the mandrill image, respectively. In other words, they are `S(j,j)`, `U(:,j)`, `V(:,j)` of the SVD of X in MATLAB. Let us define the rank k approximation of the image X as

$$X_k := \sigma_1 \boldsymbol{u}_1 \boldsymbol{v}_1^{\mathsf{T}} + \cdots + \sigma_k \boldsymbol{u}_k \boldsymbol{v}_k^{\mathsf{T}}.$$

Then, for $k = 1, 6, 11, 31$, compute X_k of the mandrill, and display the results. Fit these four images in one page by using subplot function in MATLAB (i.e., use `subplot(2,2,1)` to display the first image, `subplot(2,2,2)` to display the second image, etc.)

(d) For $k = 1, 6, 11, 31$, display the residuals, i.e., $X - X_k$, fit them in one page, print them, and attach that page in your HW sheets.

(e) For $k = 1, 6, 11, 31$, compute $\|X - X_k\|_2$ by the `norm` function of MATLAB. Then, compare the results with σ_{k+1}. More precisely, compute the relative error and report the results:

$$\frac{|\sigma_{k+1} - \|X - X_k\|_2\,|}{\sigma_{k+1}}.$$

- Consider the matrix $A = \begin{pmatrix} -2 & 11 \\ -10 & 5 \end{pmatrix}$

  (a) Determine an SVD of A by hand calculation. (b) List the singular values, left singular vectors, and right singular vectors of A. Draw a labeled picture of the unit circle in R^2 and its image under A, together with the singular vectors, with the coordinates of their vertices marked. (c) Verify that $det(A) = r_1 r_2$ and $|det(A)| = \sigma_1\sigma_2$ the singular values.

- Suppose A is a 3×3 square non-singular matrix and let $E = \{y : y = Ax,\ \|x\| = 1\}$ be the image of the unit sphere. (a) Explain why E is an ellipsoid and write its equation down. (b) What are their principal axis and their lengths? What is the volume of the solid ellipsoid E?

- Suppose A is an $m \times m$ matrix with SVD $A = U\Sigma V^T$. Use this to find the **singular value** decomposition of the $2m \times 2m$ matrix:

$$\begin{pmatrix} 0 & A^T \\ A & 0 \end{pmatrix}$$

- What are the singular values of a $1 \times n$ matrix? Write down its singular value decomposition.
- Prove that if A is a square non-singular matrix then the singular values of the A^{-1} are the reciprocals of the singular values of A.