STAT 309: MATHEMATICAL COMPUTATIONS I
FALL 2013
PROBLEM SET 2

1. You are not allowed to use the SVD for this problem, i.e. no arguments should depend on the SVD of A or A^*. Let W be a subspace of $\mathbb{C}^n$. The subspace $W^\perp$ below is called the *orthogonal complement* of W.

$$W^\perp = \{\mathbf{v} \in \mathbb{C}^n \mid \mathbf{v}^*\mathbf{w} = 0 \text{ for all } \mathbf{w} \in W\}.$$

For any subspace $W \subseteq \mathbb{C}^n$, we write $P_W \in \mathbb{C}^{n\times n}$ for an orthogonal projection onto W.

(a) Show that $\mathbb{C}^n = W \oplus W^\perp$ and that $W = (W^\perp)^\perp$.

(b) Let $A \in \mathbb{C}^{m\times n}$. Show that

$$\ker(A^*) = \operatorname{im}(A)^\perp \quad \text{and} \quad \operatorname{im}(A^*) = \ker(A)^\perp.$$

(c) Deduce the Fredholm alternative:

$$\mathbb{C}^m = \ker(A^*) \oplus \operatorname{im}(A) \quad \text{and} \quad \mathbb{C}^n = \operatorname{im}(A^*) \oplus \ker(A).$$

In other words any $\mathbf{x} \in \mathbb{C}^n$ and $\mathbf{y} \in \mathbb{C}^m$ can be written uniquely as

$$\begin{aligned}
\mathbf{x} &= \mathbf{x}_0 + \mathbf{x}_1, \quad \mathbf{x}_0 \in \ker(A),\ \mathbf{x}_1 \in \operatorname{im}(A^*),\ \mathbf{x}_0^*\mathbf{x}_1 = 0, \\
\mathbf{y} &= \mathbf{y}_0 + \mathbf{y}_1, \quad \mathbf{y}_0 \in \ker(A^*),\ \mathbf{y}_1 \in \operatorname{im}(A),\ \mathbf{y}_0^*\mathbf{y}_1 = 0.
\end{aligned}$$

(d) Show that

$$\mathbf{x}_0 = P_{\ker(A)}\mathbf{x}, \quad \mathbf{x}_1 = P_{\operatorname{im}(A^*)}\mathbf{x}, \quad \mathbf{y}_0 = P_{\ker(A^*)}\mathbf{y}, \quad \mathbf{y}_1 = P_{\operatorname{im}(A)}\mathbf{y}.$$

(e) Consider the least squares problem for some $\mathbf{b} \in \mathbb{C}^m$,

$$\min_{\mathbf{x}\in\mathbb{C}^n} \|\mathbf{b} - A\mathbf{x}\|_2. \tag{1.1}$$

Show that for any $\mathbf{x} \in \mathbb{C}^n$,

$$\|\mathbf{b} - A\mathbf{x}\|_2 \geq \|\mathbf{b}_0\|_2$$

where $\mathbf{b}_0 = P_{\ker(A^*)}\mathbf{b}$. Deduce that $\mathbf{x} \in \mathbb{C}^n$ is a solution to (1.1) if and only if

$$A\mathbf{x} = \mathbf{b}_1 \quad \text{or, equivalently,} \quad \mathbf{b} - A\mathbf{x} = \mathbf{b}_0. \tag{1.2}$$

Why is $A\mathbf{x} = \mathbf{b}_1$ consistent?

(f) Show that (1.2) is equivalent (i.e., if and only if) to the normal equation

$$A^*A\mathbf{x} = A^*\mathbf{b}. \tag{1.3}$$

Caveat: In numerical analysis, it is in general a terrible idea to solve a least squares problem via its normal equation. Nonetheless (1.3) can be useful in mathematical arguments. We discussed in the lectures the very limited number of scenarios when it makes sense to solve (1.3) via Cholesky decomposition.

Date: October 22, 2013 (Version 1.5); due: November 5, 2013.

(g) Show that the pseudoinverse solution

$$\min\left\{\|\mathbf{x}\|_2 : \mathbf{x} \in \operatorname*{argmin}_{\mathbf{x}\in\mathbb{C}^n}\|\mathbf{b} - A\mathbf{x}\|_2\right\}$$

is given by

$$\mathbf{x}_1 = P_{\operatorname{im}(A^*)}\mathbf{x}$$

where $\mathbf{x} \in \mathbb{C}^n$ satisfies (1.2).

(h) Let $A \in \mathbb{C}^{n\times n}$ be normal, i.e., $A^*A = AA^*$. Show that

$$\ker(A^*) = \ker(A) \quad \text{and} \quad \operatorname{im}(A^*) = \operatorname{im}(A)$$

and deduce that for a normal matrix,

$$\mathbb{C}^n = \ker(A) \oplus \operatorname{im}(A).$$

2. Let $A, B \in \mathbb{C}^{m\times n}$ with $n \le m$. In the lectures, we claim that the solution $X \in \mathrm{U}(n)$ to

$$\min_{X^*X=I}\|A - BX\|_F$$

is given by $X = UV^*$ where $B^*A = U\Sigma V^*$ is its singular value decomposition. Here we will prove it and consider some variants.

(a) Show that

$$\|A - BX\|_F^2 = \operatorname{tr}(A^*A) + \operatorname{tr}(B^*B) - 2\operatorname{Re}\operatorname{tr}(X^*B^*A)$$

and deduce that the minimization problem is equivalent to

$$\max_{X^*X=I}\operatorname{Re}\operatorname{tr}(X^*B^*A).$$

(b) Show that

$$\operatorname{Re}\operatorname{tr}(X^*B^*A) \le \sum_{i=1}^{n}\sigma_i(B^*A)$$

for any $X \in \mathrm{U}(n)$. When is the upper bound attained?

(c) Show that

$$\min_{X^*X=I}\|A - BX\|_F^2 = \sum_{i=1}^{n}(\sigma_i(A)^2 - 2\sigma_i(B^*A) + \sigma_i(B)^2).$$

(d) Suppose A has full column rank. Show that the following method produces a Hermitian matrix $X \in \mathbb{C}^{n\times n}$ that solves

$$\min_{X^*=X}\|AX - B\|_F. \tag{2.4}$$

(i) Show that the SVD of A takes the form

$$A = U\begin{bmatrix}\Sigma\\ O\end{bmatrix}V^*$$

where $U \in \mathrm{U}(m)$, $V \in \mathrm{U}(n)$, and $\Sigma = \operatorname{diag}(\sigma_1, \dots, \sigma_n) \in \mathbb{C}^{n\times n}$ is a diagonal matrix.

(ii) Show that

$$\|AX - B\|_F^2 = \|\Sigma Y - C_1\|_F^2 + \|C_2\|_F^2$$

where $Y = V^*XV$ and $C = \begin{bmatrix}C_1\\ C_2\end{bmatrix} = U^*BV$.

(iii) Note that Y must be Hermitian if X is. Show that

$$\|\Sigma Y - C_1\|_F^2 = \sum_{i=1}^{n} |\sigma_i y_{ii} - c_{ii}|^2 + \sum_{j>i} |\sigma_i y_{ij} - c_{ij}|^2 + |\sigma_j y_{ij} - \bar{c}_{ji}|^2$$

and deduce that the minimum value of (2.4) is attained when

$$y_{ij} = \frac{\sigma_i c_{ij} + \sigma_j \bar{c}_{ji}}{\sigma_i^2 + \sigma_j^2}$$

for all $i, j = 1, \dots, n$.

(e) Given $A \in \mathbb{C}^{n\times n}$. Describe how you would find $X \in \mathbb{C}^{n\times n}$ that solves

$$\min_{\det(X) = |\det(A)|} \|A - X\|_F.$$

(*Hint*: Consider the SVD of A).

3. Let $\mathbf{x} \in \mathbb{C}^m$, $\mathbf{y} \in \mathbb{C}^n$, and $A = \mathbf{x}\mathbf{y}^* \in \mathbb{C}^{m\times n}$.

(a) Show that $\operatorname{rank}(A) = 1$ iff $\mathbf{x}$ and $\mathbf{y}$ are both non-zero. Such a matrix is usually called a rank-1 matrix.

(b) Show that

$$\|A\|_F = \|A\|_2 = \|\mathbf{x}\|_2 \|\mathbf{y}\|_2 \tag{3.5}$$

and that

$$\|A\|_\infty \le \|\mathbf{x}\|_\infty \|\mathbf{y}\|_1.$$

What can you say about $\|A\|_1$?

(c) Let $\mathbf{x}_1, \dots, \mathbf{x}_r \in \mathbb{C}^m$ be linearly independent and $\mathbf{y}_1, \dots, \mathbf{y}_r \in \mathbb{C}^n$ be linearly independent. Let

$$A = \mathbf{x}_1\mathbf{y}_1^* + \cdots + \mathbf{x}_r\mathbf{y}_r^*.$$

Show that $\operatorname{rank}(A) = r$. Show that this is not necessarily true if we drop either of the linear independence conditions.

(d) Given any $0 \ne A \in \mathbb{C}^{m\times n}$, show that

$$\operatorname{rank}(A) = \min\{r \in \mathbb{N} \mid A = \textstyle\sum_{i=1}^r \mathbf{x}_i\mathbf{y}_i^*\}.$$

In other words, the rank of a matrix is the smallest r so that it may be expressed as a sum of r rank-1 matrices.

(e) Show the following generalization of (3.5),

$$\|A\|_F \le \sqrt{\operatorname{rank}(A)}\|A\|_2.$$

Note that $\nu\operatorname{rank}(A) = \|A\|_F^2/\|A\|_2^2$ is one of the three notions of numerical ranks that we discussed. It is often used as a continuous surrogate for matrix rank.

(f) Show that with the nuclear norm we get instead

$$\|A\|_* \le \operatorname{rank}(A)\|A\|_2. \tag{3.6}$$

In other words we could also use $\|A\|_*/\|A\|_2$ as a continuous surrogate for matrix rank. In fact, this has been quite popular recently (cf. next problem).

4. Let $A \in \mathbb{C}^{m\times n}$ be a matrix with missing entries. More precisely we let $\Omega \subseteq \{1, \dots, m\} \times \{1, \dots, n\}$ be a subset of the row and column indices. We know the value of a_{ij} if $(i, j) \in \Omega$ but not otherwise. Now one way to perform *matrix completion*, i.e., recovering the missing entries in A, is to find an $X \in \mathbb{C}^{m\times n}$ whereby some loss function f is minimized, subjected to the constraint that x_{ij} agrees with all known entries of A:

$$\begin{aligned} \text{minimize} \quad & f(X) \\ \text{subject to} \quad & x_{ij} = a_{ij} \text{ for } (i,j) \in \Omega. \end{aligned}$$

One could argue that the most natural candidate for f is

$$f(X) = \operatorname{rank}(X), \tag{4.7}$$

but matrix rank is a discrete valued function and techniques of continuous optimization cannot be applied. A popular alternative is to instead use

$$f(X) = \|X\|_*$$

because nuclear norm is the largest convex function that satisfies (3.6) but this won't give the solution to (4.8) in general. Here we will see that an alternative way to solve (4.8) (in principle) as a continuous problem is via SVD.

$$\begin{array}{ll} \text{minimize} & \operatorname{rank}(X) \\ \text{subject to} & x_{ij} = a_{ij} \text{ for } (i,j) \in \Omega, \end{array} \tag{4.8}$$

For $1 \le r \le \min(m,n)$, let $f_r : \mathbb{C}^{m\times n} \to [0,\infty)$ be the function

$$f_r(X) = \sum_{i=r+1}^{\min(m,n)} \sigma_i(X)^2.$$

and consider the minimization problem

$$\begin{array}{ll} \text{minimize} & f_r(X) \\ \text{subject to} & x_{ij} = a_{ij} \text{ for } (i,j) \in \Omega. \end{array} \tag{4.9}$$

Let X_r be a minimizer of (4.9). Show that

$$f_r(X_r) = 0 \quad \text{if and only if} \quad r \ge \operatorname{rank}(X_*)$$

where X_* is a minimizer of (4.8). Discuss how this can be used to obtain a solution to (4.8).

5. Let $A \in \mathbb{C}^{m\times n}$ and $\mathbf{b} \in \mathbb{C}^m$. We will discuss a variant of $A\mathbf{x} \approx \mathbf{b}$ where the error occurs only in A. Note that in ordinary least squares we assume that the error occurs only in $\mathbf{b}$ while in total least squares we assume that it occurs in both A and $\mathbf{b}$.

(a) Show that if $0 \ne \mathbf{x} \in \mathbb{C}^m$, then

$$\left\| A\left(I - \frac{\mathbf{x}\mathbf{x}^*}{\mathbf{x}^*\mathbf{x}}\right)\right\|_F^2 = \|A\|_F^2 - \frac{\|A\mathbf{x}\|_2^2}{\mathbf{x}^*\mathbf{x}}.$$

(b) Show that the matrix

$$E = \frac{(\mathbf{b} - A\mathbf{x})\mathbf{x}^*}{\mathbf{x}^*\mathbf{x}} \in \mathbb{C}^{m\times n}$$

has the smallest 2-norm of all $m \times n$ matrices E that satisfy

$$(A + E)\mathbf{x} = \mathbf{b}.$$

(c) What are the solutions of

$$\min_{(A+E)\mathbf{x}=\mathbf{b}} \|E\|_2 \qquad \text{and} \qquad \min_{(A+E)\mathbf{x}=\mathbf{b}} \|E\|_F?$$

The minimum is taken over all $E \in \mathbb{R}^{m\times n}$ such that $(A+E)\mathbf{x} = \mathbf{b}$ is consistent (i.e., has a solution).

(d) Given $\mathbf{a} \in \mathbb{C}^n$, $\mathbf{b} \in \mathbb{C}^m$, and $\delta > 0$. Show how to solve the problems

$$\min_{\|E\|_F \le \delta} \|E\mathbf{a} - \mathbf{b}\|_2 \qquad \text{and} \qquad \max_{\|E\|_F \le \delta} \|E\mathbf{a} - \mathbf{b}\|_2$$

over all $E \in \mathbb{C}^{m\times n}$.

6. The files required for this problem are in `http://www.stat.uchicago.edu/~lekheng/courses/309/stat309-hw2/`. The matrix in `processed.mat` (Matlab format) or `processed.txt` (comma separated, plain text) is a 49×7 matrix where each row is indexed by a country in `row.txt` and each column is indexed by a demographic variable in `column.txt`, ordered as in the respective files. So for example, if we denote the matrix by $A = [a_{ij}] \in \mathbb{R}^{49\times 7}$, then $a_{23} = -0.2743$ is Austria's population per square kilometers (row index 2 = Austria, column index 3 = population per square kilometers). As you probably notice, this matrix has been slightly preprocessed. If you want to see the raw data, you can find them in `raw.txt` (e.g. the actual value for Austria's population per square kilometers is 84) but you don't need the raw data for this problem.

(a) Find the first two right singular vectors of A, $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^7$. Project the data onto the two-dimensional space spanned by $\mathbf{v}_1$ and $\mathbf{v}_2$. Plot this in a graph where the x- and y-axes correspond to $\mathbf{v}_1$ and $\mathbf{v}_2$ respectively and where the points correspond to the countries — label each point by the country it corresponds to. Identify the two obvious outliers.

(b) Now do the same with the two left singular vectors of A, $\mathbf{u}_1, \mathbf{u}_2 \in \mathbb{R}^{49}$, i.e., project the data onto the two-dimensional space spanned by $\mathbf{u}_1$ and $\mathbf{u}_2$ and plot this in a graph as before. Note that in this case, the points correspond to the demographic variables — label them accordingly.

(c) Overlay the two graphs in (a) and (b). Identify the two demographic variables near the two outlier countries — these explain why the two countries are outliers.

(d) Remove the two outlier countries and redo (a) with this 47×7 matrix. This allows you to see features that were earlier obscured by the outliers. Which two European countries are most alike Japan?

The graphs in (a) and (b) are called *scatter plots* and the overlayed one in (c) is called a *biplot*. See `http://en.wikipedia.org/wiki/Biplot` for more information. The reason we didn't need to adjust the scale of the axes using the singular values of A like in the Wikipedia description is because the preprocessing has taken care of the scaling; if we had started from the raw data, then we would need to deal with this complication.