ENGG 5781: Matrix Analysis and Computations	**2018-19 First Term**
Assignment 1	
Instructor: Wing-Kin Ma	Due: October 5, 2018

Answer Problems 1–3, and **either** Problem 4 or Problem 5.

Note:

1. By submitting this assignment, we are assumed to have read the homework guideline `http://www.ee.cuhk.edu.hk/~wkma/engg5781/hw/hw_guidelines.pdf` thereby understanding and respecting the guideline mentioned there.

2. You are allowed to use properties and theorems up to Lecture 2; unless specified, that includes all the results and proofs in the additional notes.

Problem 1 (30%) Are the following sets subspaces? Provide your answer with a proof.

(a) $\mathcal{S} = \{\mathbf{X} \in \mathbb{R}^{m\times n} \mid \mathbf{X}^T\mathbf{Y} + \mathbf{Y}^T\mathbf{X} = \mathbf{0}\}$, where $\mathbf{Y} \in \mathbb{R}^{m\times n}$ is given.

(b) $\mathcal{S} = \{\mathbf{X} \in \mathbb{R}^{m\times n} \mid \mathbf{X} = \mathbf{A}\mathbf{B}^T, \mathbf{A} \in \mathbb{R}^{m\times r}, \mathbf{B} \in \mathbb{R}^{n\times r}\}$, where $r < \min\{m, n\}$.

(c) $\mathcal{S} = \{\boldsymbol{\mathcal{X}} \in \mathbb{R}^{m\times n\times p} \mid x_{ijk} = \sum_{\ell=1}^{r} a_{i\ell}b_{j\ell}c_{k\ell},\ \forall i, j, k,\ c_{k\ell} \in \mathbb{R}\ \forall k, \ell\}$, where $a_{i\ell}, b_{j\ell} \in \mathbb{R}$ are given.

(d) $\mathcal{S} = \{\mathbf{y} \in \mathbb{C}^m \mid |\sum_{i=1}^{m} y_i(\alpha_j)^i| = 0,\ j = 1, \ldots, n\}$, where $\alpha_1, \ldots, \alpha_n \in \mathbb{C}$ are given.

Problem 2 (20%) A non-empty subset $\mathcal{S}$ of $\mathbb{R}^m$ is said to be an affine set if

$$\alpha \in \mathbb{R}, \mathbf{x}, \mathbf{y} \in \mathcal{S} \quad \Longrightarrow \quad \alpha\mathbf{x} + (1-\alpha)\mathbf{y} \in \mathcal{S}.$$

(a) Show that if $\mathcal{S}$ is affine, then any affine combination of $\mathbf{a}_1, \ldots, \mathbf{a}_n \in \mathcal{S}$, i.e.,

$$\mathbf{y} = \sum_{i=1}^{n} \alpha_i \mathbf{a}_i, \quad \alpha_1, \ldots, \alpha_n \in \mathbb{R}, \quad \sum_{i=1}^{n} \alpha_i = 1,$$

lies in $\mathcal{S}$.

(b) Show that an affine set $\mathcal{S}$ can always be represented by $\mathcal{S} = \mathcal{V} + \mathbf{b}$, where $\mathbf{b} \in \mathcal{S}$ and $\mathcal{V}$ is a subspace[1].

Problem 3 (20%) Let $\mathcal{S} = \{\mathbf{y} \in \mathbb{R}^m \mid \mathbf{y} = \mathbf{A}\mathbf{x} - \mathbf{b}, \mathbf{x} \in \mathbb{R}^n, \|\mathbf{x}\|_2 \leq 1\}$. Suppose $\mathbf{b} \notin \mathcal{R}(\mathbf{A})$.

(a) Is $\mathcal{S} \cap \mathcal{S}^\perp = \{\mathbf{0}\}$, or $\mathcal{S} \cap \mathcal{S}^\perp = \emptyset$?

(b) Show that $\mathcal{S}^\perp = \mathcal{N}([\ \mathbf{A}\ \mathbf{b}\ ]^T)$.

[1] The notation $\mathbf{b} + \mathcal{V}$ means that $\mathbf{b} + \mathcal{V} = \{\mathbf{y} = \mathbf{b} + \mathbf{v} \mid \mathbf{v} \in \mathcal{V}\}$.

Problem 4 (30%) Let $\{\mathbf{a}_1, \ldots, \mathbf{a}_n\} \subset \mathbb{R}^m$ be a given set of linearly independent vectors. Let $\mathbf{q}_1, \ldots, \mathbf{q}_n \in \mathbb{R}^m$ whose construction will be specified, and let $\mathcal{S}_i = \text{span}\{\mathbf{q}_1, \ldots, \mathbf{q}_i\}$ for the sake of notational convenience. Consider the following procedure.

Algorithm 1:

$\mathbf{q}_1 = \mathbf{a}_1 / \|\mathbf{a}_1\|_2$;
for $i = 2, \ldots, n$ **do**
    $\tilde{\mathbf{q}}_i = \Pi_{\mathcal{S}_{i-1}^{\perp}}(\mathbf{a}_i)$;
    $\mathbf{q}_i = \tilde{\mathbf{q}}_i / \|\tilde{\mathbf{q}}_i\|_2$;
end

(a) Show that, for any $i \in \{1, \ldots, n\}$,

$$\text{span}\{\mathbf{a}_1, \ldots, \mathbf{a}_i\} = \mathcal{S}_i.$$

DO NOT use the proof in Lecture 1, page 47–50. Consider the problem as if you did not know what is Gram-Schmidt, which you can easily find in textbooks or in the world-wide web. Use ONLY the basic notions of subspace, with projection onto subspaces included, to rediscover the result.

(b) Use the basic notions of subspaces, as well as those of orthogonality and LS, to show that $\tilde{\mathbf{q}}_i = \mathbf{a}_i - \mathbf{Q}_{i-1}\mathbf{Q}_{i-1}^T \mathbf{a}_i$, where $\mathbf{Q}_i = [\ \mathbf{q}_1, \ldots, \mathbf{q}_i\ ]$.

(c) Suggest how we may modify the algorithm when $\{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$ is linearly dependent. Note that the ultimate goal is to find an orthogonal basis for $\text{span}\{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$.

Note: Since this problem requires some explanation, we consider it as a semi-essay problem. English fluency, clarity of presentation, originality of the presentation (relative to others), etc., will be heavily taken into account.

Problem 5 (30%) This is a MATLAB problem. The problem we deal with is a face recognition problem. Your main reference is [1], Chapter 9. I also recommend [2] as an additional reference for you to get better understanding of the context. Download from the course website the following three files: `X.mat`, `y_part_a.mat`, and `y_part_b.mat`. Loading `X.mat` on MATLAB, you will see five matrices, namely, `X_1`, `X_2`, `X_3`, `X_4`, `X_5`. I will sometimes use $\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3, \mathbf{X}_4, \mathbf{X}_5$, our standard matrix notations, to describe them. Each $\mathbf{X}_i$ is a collection of 63 images, with size 192×168, taken from the same person. Specifically, each column of $\mathbf{X}_i$ is an image stored in the vectorized form. You can see them, say, for $\mathbf{X}_1$, by calling

```
>> for i=1:63, subplot(8,8,i); imshow(reshape(X_1(:,i),192,168)); end;
```

I should mention that the data come from "Yale Face Database B."

(a) Load `y_part_a.mat`. You will find a vector `y`, or $\mathbf{y}$. You can see it by calling

```
>> imshow(reshape(y,192,168));
```

You will see that it is a noisy image. Form a multi-person data matrix $\mathbf{X} = [\ \mathbf{X}_1, \ldots, \mathbf{X}_5\ ]$, apply LS

$$\min_{\mathbf{w} \in \mathbb{R}^n} \ \|\mathbf{y} - \mathbf{X}\mathbf{w}\|_2^2,$$

and use the LS solution to construct an estimated, and hopefully noise-cleaned, image. Show us the image you recover. Then, from the LS solution, identify the person $\mathbf{y}$ is associated with. Note: You can certainly watch the images and *subjectively* identify who that person is, but I would like to see a quantitative way.

(b) Load `y_part_b.mat`, which contains a vector $\mathbf{y}$. Using `imshow`, you will see that a small part of the image is severely corrupted. Try LS and show the recovered image. Then, try Algorithm 11 in [1], AltMin for Robust Regression (AM-RR), with parameter $k = 4775$. You should also write a short description concerning what is the rationale of AM-RR, and how it works. Show the recovered image of AM-RR.

Note: You will need to submit your MATLAB code online via Blackboard. You also need to provide a description on your assignment. We do not do reverse engineering tasks such as guessing how a MATLAB code works in the absence of any description. If you are in doubt, talk to us to understand more.

References

[1] P. Jain and P. Kar. Non-convex optimization for machine learning. *Foundations and Trends® in Machine Learning*, 10(3–4):142–336, 2017.

[2] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma. Robust face recognition via sparse representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(2):210–227, 2009.