DS-GA 1014: Midterm Solutions

Optimization and Computational Linear Algebra for Data Science (NYU, Fall 2018)

Instructor: Afonso S. Bandeira (`bandeira@cims.nyu.edu`)

November 6, 2018

The exam ends at the end of the time of the lecture. Please justify your answers, proving the statements you make. You are allowed to refer to results shown in lecture (or that are in the syllabus) as long as you state them precisely, meaning that you should say exactly which hypothesis are needed in the result you use.

This exam is open book/notes. You are allowed to consult notes and books you bring, but not allowed to use electronic devices.

If you need to impose extra conditions on a problem to make it easier (or consider specific cases of the question, like taking n to be 2, for example), state explicitly that you have done so. Solutions where extra conditions were assumed, or where only special cases where treated, will also be graded (probably scored as a partial answer).

The exam has 2 pages. It has 8 question groups that together total 100 points plus extra credit. Extra credit points will be added to your total score. If you have questions (or find a typo) let me know. Any eventual typo will be announced on the board.

1. (*16 points*) Let $A \in \mathbb{R}^{n\times n}$ be a matrix. For each of the subsets of $\mathbb{R}^n$ below, say whether it is necessarily a subspace or may not be one. For each set either prove it must be a subspace, or find a counter-example (a matrix A for which it is not).

   (a) $\{x \in \mathbb{R}^n : x = Ay \text{ for some } y \in \mathbb{R}^n\}$
   (b) $\{x \in \mathbb{R}^n : \sum_{i=1}^n (Ax)_i = 0\}$
   (c) $\{x \in \mathbb{R}^n : \sum_{i=1}^n (Ax)_i = 1\}$
   (d) $\{x \in \mathbb{R}^n : Ax = A^2x\}$
   (e) **Extra Credit (4 points):** $\{x \in \mathbb{R}^n : \|Ax\| \le \|A^2x\|\}$

   Solution.

   (a) Yes, this is $\mathrm{Im}(A)$, which is a subspace.
   (b) Yes, this is $\ker(\mathbf{1}^T A)$, which is a subspace (here $\mathbf{1}$ is the vector with all ones).
   (c) No, for any A we have $\sum_{i=1}^n (A0)_i \neq 1$.
   (d) Yes, this is $\ker(A - A^2)$, which is a subspace.

(e) No. Define A by

$$A = \begin{bmatrix} 2 & 0 \\ 0 & 1/2 \end{bmatrix}.$$

Note that

$$\|A(e_1 + \alpha e_2)\| \leq \|Ae_1\| + \alpha\|Ae_2\| = 2 + \alpha/2$$

and

$$\|A^2(e_1 + \alpha e_2)\| \geq \|A^2 e_1\| - \alpha\|A^2 e_2\| = 4 - \alpha/4.$$

Thus $e_1, e_1 + e_2/2$ are both in the set (which together span all of $\mathbb{R}^2$) but e_2 is not.

2. (*15 points*) Let U be a subspace of $\mathbb{R}^n$ and P_U the (orthogonal) projection onto U.

(a) Show that for all $x \in \mathbb{R}^n$, $\|P_U x\| \leq \|x\|$.

(b) Determine the set of all $x \in \mathbb{R}^n$ for which $\|P_U x\| = \|x\|$, and prove your set is correct.

(c) Determine the set of all $x \in \mathbb{R}^n$ for which $\|P_U x\| = 0$, and prove your set is correct.

Solution. Let $u_1, \ldots, u_k$ be an orthonormal basis for U which we extend to an orthonormal basis $u_{k+1}, \ldots, u_n$ for $U^\perp$.

(a) Note that

$$\begin{aligned}
\|P_U x\|^2 &= \left\|\sum_{i=1}^{k} \langle x, u_i\rangle u_i\right\|^2 && \text{(Equiv defn of } P_U \text{ from Lab)} \\
&= \sum_{i=1}^{k} \langle x, u_i\rangle^2 && \text{(Homework, or calculation)} \\
&\leq \sum_{i=1}^{n} \langle x, u_i\rangle^2 \\
&= \left\|\sum_{i=1}^{n} \langle x, u_i\rangle u_i\right\|^2 && \text{(Homework, or calculation)} \\
&= \|x\|^2.
\end{aligned}$$

(b) The set is U. The above proof is an equality exactly when $\langle x, u_j\rangle = 0$ for $j > k$, i.e., when

$$x = \sum_{i=1}^{k} \langle x, u_i\rangle u_i.$$

Stated differently, this occurs when x is in the span of $u_1, \ldots, u_k$, which is the same as saying $x \in U$.

(c) The set is $U^\perp$. If $x \in U^\perp$ then $\langle x, u_i\rangle = 0$ for $i = 1, \ldots, k$ so the above calculation shows $P_U(x) = 0$. Conversely, if $P_U(x) = 0$ then by the above calculation

$$x = \sum_{i=k+1}^{n} \langle u_i, x\rangle u_i,$$

where $u_{k+1}, \dots, u_n \in U^\perp$. Thus $x \in U^\perp$.

3. (*12 points*) For each of the following give a proof or counterexample.

   (a) For all $A, B \in \mathbb{R}^{n\times n}$, if A, B are orthogonal then AB is orthogonal.

   (b) For all $A, B \in \mathbb{R}^{n\times n}$, if A, B are invertible then AB is invertible.

   (c) For all $A, B \in \mathbb{R}^{n\times n}$, $AB = BA$.

   Solution.

   (a) True. Note that $(AB)^T(AB) = B^TA^TAB = B^TIB = I$.

   (b) True. Note that $AB(B^{-1}A^{-1}) = AIA^{-1} = I$, thus $B^{-1}A^{-1}$ is an inverse for AB.

   As an alternative proof, we use the fact that for a square matrix invertibility is equivalent to having full rank, which is equivalent to having trivial kernel by the fundamental theorem. Note that $ABx = 0$ implies $A(Bx) = 0$ which implies $Bx = 0$ (since $\ker(A) = \{0\}$) which implies $x = 0$ (since $\ker(B) = \{0\}$). Thus $\ker(AB) = \{0\}$ proving invertibility.

   (c) False. Define A, B by
   $$A = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$
   so that
   $$AB = \begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix}, \quad BA = \begin{bmatrix} 2 & 1 \\ 2 & 1 \end{bmatrix}.$$

4. (*20 points*) Let $A \in \mathbb{R}^{m\times n}$ with $m > n$ be full rank (meaning that $\operatorname{rank}(A) = n$).

   (a) Does $Ax = b$ have a solution for all $b \in \mathbb{R}^m$? (prove or give a counterexample)

   (b) Can there exist a vector $b \in \mathbb{R}^m$ for which there exist two distinct solutions (meaning that there exist $x \neq y$ such that $Ax = Ay = b$)? Give an example of a matrix A for which this can happen, or prove it cannot happen.

   (c) Is $A^TA \in \mathbb{R}^{n\times n}$ invertible? (prove or give a counterexample)

   (d) Is $AA^T \in \mathbb{R}^{m\times m}$ invertible? (prove or give a counterexample)

   (e) **Extra Credit (5 points):** Prove the largest entry of A^TA is on the diagonal. In other words, prove that
   $$\max_{i,j}\ (A^TA)_{ij} = \max_k\ (A^TA)_{kk}.$$

   Solution.

   (a) False. Let A, b be defined by
   $$A = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad b = e_2.$$
   Then A has rank 1, but $e_2 \notin \operatorname{Im}(A) = \operatorname{Span}(e_1)$, so $Ax = b$ has no solution.

(b) No. Suppose $Ax = Ay = b$ for $x \neq y$. Then $A(x - y) = 0$ showing $0 \neq x - y \in \ker(A)$. But by the fundamental theorem $\dim\ker(A) = 0$, a contradiction.

As an alternative proof, since the columns of A are linearly independent we have that $Ax = b$ implies $x = A^+b$, so if the equation is solvable it must have solution A^+b.

(c) Yes. We show A^TA has trivial kernel. Note that $A^TAx = 0$ implies $\|Ax\|^2 = x^TA^TAx = 0$. This shows $Ax = 0$ which implies $x = 0$ since the columns of A are linearly independent (either by the definition of linear independence and matrix multiplication, or by using the fundamental theorem to see $\dim\ker(A) = 0$).

(d) No. Define $A = e_1$ as above. Then

$$AA^T = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

which has rank 1.

(e) Let $a_1, \dots, a_n$ denote the columns of A. Then we have

$$(A^TA)_{ij} = \langle a_i, a_j \rangle \leq \|a_i\|\|a_j\| \leq \max_k \|a_k\|^2 = \max_k \ (A^TA)_{kk}.$$

5. (*16 points*) Let $A, B \in \mathbb{R}^{m\times n}$.

(a) Prove that if $\operatorname{Im}(B) \subseteq \operatorname{Im}(A)$ then for any $y \in \mathbb{R}^m$ the equation $Ax = BB^Ty$ has a solution $x \in \mathbb{R}^n$.

(b) Suppose $\operatorname{Im}(B) = \operatorname{Im}(A)$ and the columns of B are orthonormal. Prove that for any $y \in \mathbb{R}^m$ the solution x to $Ax = BB^Ty$ minimizes

$$\min_v \|Av - y\|.$$

Solution.

(a) Note that $BB^Ty = B(B^Ty) \in \operatorname{Im}(B) \subseteq \operatorname{Im}(A)$. Thus $BB^Ty \in \operatorname{Im}(A)$ and the equation is solvable.

(b) By the assumption $BB^T = P_{\operatorname{Im}(A)}$ so the solution x satisfies

$$Ax = P_{\operatorname{Im}(A)}(y) = \operatorname{argmin}_{u\in\operatorname{Im}(A)} \|u - y\|.$$

But since $Av \in \operatorname{Im}(A)$ for all v we see that x is the required minimizer.

6. (*9 points*) Determine all solutions to $\operatorname{Tr}(AA^T) = 0$ for $A \in \mathbb{R}^{m\times n}$ and prove that your answer is correct. Here the trace of a matrix $B \in \mathbb{R}^{m\times m}$ is defined by

$$\operatorname{Tr}(B) = \sum_{i=1}^m B_{ii}.$$

Solution. Note that $0 = \mathrm{Tr}(AA^T) = \mathrm{Tr}(A^TA) = \sum_{i=1}^m \sum_{j=1}^n A_{ij}^2$, so every entry of A must be zero.

As an alternative solution, we know that $AA^T \succeq 0$ from the homework, so $\mathrm{Tr}(AA^T) = 0$ implies all of the eigenvalues of AA^T are 0, which implies $AA^T = 0$. But $\ker(A^T) = \ker(AA^T)$ so $A^T = 0$ giving $A = 0$.

7. (*12 points*)

   (a) Prove that if $B \in \mathbb{R}^{n\times n}$ is symmetric with strictly positive eigenvalues (i.e., $\lambda_i > 0$ for $i = 1, \ldots, n$) then B is invertible.

   (b) Prove that $A^TA + \alpha I$ is invertible for any $A \in \mathbb{R}^{m\times n}$ and any $\alpha > 0$.

   Solution.

   (a) By the spectral theorem we have $B = V\Lambda V^T$. Since the diagonal entries of Λ are all positive, and the rank of B is equal to the number of non-zero eigenvalues, we see that B has rank n and is thus invertible (with inverse $V\Lambda^{-1}V^T$).

   (b) First note that A^TA is symmetric so by the spectral theorem we have $A^TA = V\Lambda V^T$. By the homework we know that $A^TA \succeq 0$, so that all entries of Λ are non-negative. Then we have
   $$A^TA + \alpha I = V(\Lambda + \alpha I)V^T$$
   where $\Lambda + \alpha I$ is diagonal with positive entries. This is a spectral decomposition of A^TA showing that A^TA has positive eigenvalues, and thus is invertible by part a.

8. **Extra Credit (9 points):** Let $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$. Prove that $Ax = b$ has a solution if there is no $y \in \mathbb{R}^m$ such that $A^Ty = 0$ and $y^Tb > 0$.

   Solution. By assumption, we have that $A^Ty = 0$ implies $y^Tb = 0$ (since if $y^Tb < 0$ then $A^T(-y) = 0$ and $(-y)^Tb > 0$). Suppose, for contradiction, that $b \notin \mathrm{Im}(A)$. Then we can write $b = u + v$ where $u \in \mathrm{Im}(A)$, $v \in \ker(A^T)$, and $v \neq 0$. But then, letting $y = v$, we have $A^Tv = 0$ and
   $$v^Tb = v^T(u + v) = v^Tv > 0,$$
   a contradiction.

9. **Extra Extra Credit:** Let $a_1, \ldots, a_k \in \mathbb{R}^n$ where $n \geq 2$ and $\|a_i\| = 1$ for $i = 1, \ldots, k$. Assume there is a $\mu \in (0, 1)$ with $\langle a_i, a_j\rangle^2 = \mu$ for all $i \neq j$. Prove that $k \leq n^2$.