DS-GA 1014: Final Review Questions

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

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

($\star$) denotes more difficult problems.

1. Let $A \in \mathbb{R}^{n \times n}$ and $B \in \mathbb{R}^{n \times n}$ be two matrices such that there exists an orthonormal basis of $\mathbb{R}^n$ for which every vector in this basis is an eigenvector of both A and B. Show that $AB = BA$.

2. Let $A \in \mathbb{R}^{n \times p}$ and $B \in \mathbb{R}^{p \times m}$ be two matrices.

   (a) Show that $\|AB\|_s \leq \|A\|_s \|B\|_s$.

   (b) Recall that the condition number of an invertible matrix A is defined as $\kappa(A) = \|A\|_s \|A^{-1}\|_s$. Show that, for all invertible matrices $A \in \mathbb{R}^{n \times n}$, one has $\kappa(A) \geq 1$, and give an example of a matrix for which $\kappa(A) = 1$.

3. Let $A \in \mathbb{R}^{n \times n}$ be a matrix. If v_1 and v_2 are eigenvectors of A with the same eigenvalue, show that any vector in $\mathrm{span}(v_1, v_2)$ is also an eigenvector of A.

4. Let $A \in \mathbb{R}^{n \times n}$ be symmetric with eigenvalues $\lambda_1, \dots, \lambda_n$. Prove that $\|Ax\| \leq \max_i |\lambda_i| \|x\|$ for any $x \in \mathbb{R}^n$.

5. Let $(t_1, y_1), \dots, (t_n, y_n) \in \mathbb{R}^2$ be a given data set. Let α^*, β^* be minimizers for the least squares error

$$\min_{\alpha,\beta} \sum_{i=1}^{n} (\alpha + \beta t_i - y_i)^2.$$

   Prove that

$$\sum_{i=1}^{n} t_i(\alpha^* + \beta^* t_i - y_i) = 0.$$

6. For $x \in \mathbb{R}^n$ let $\|x\|_\infty = \max_i |x_i|$, i.e., the largest absolute coordinate. Prove that

$$\|x\|_\infty \leq \|x\| \leq \|x\|_\infty \sqrt{n}.$$

7. Let $A \in \mathbb{R}^{n \times n}$ be symmetric with eigenvalues $\lambda_1, \dots, \lambda_n$ which are all non-negative. Prove that $x^T A x \geq 0$ for all $x \in \mathbb{R}^n$.

8. Suppose $A \in \mathbb{R}^{m \times n}$ has rank m. Prove that AA^T is invertible.

9. ($\star$) Consider the optimization problem

$$\begin{array}{ll} \text{minimize}_x & \|x\|^2 \\ \text{subject to} & Ax = b \end{array}$$

where $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$ are fixed and $b \in \text{Im}(A)$.

   (a) Prove that any minimizer x^* must belong to $\text{Im}(A^T)$.

   (b) Give a formula for the minimizer x^*, and show it is unique.

10. Let $A \in \mathbb{R}^{n \times m}$ and $B \in \mathbb{R}^{n \times k}$ and define the block matrix $C \in \mathbb{R}^{n \times (m+k)}$ by

$$C = [A\ B].$$

Either prove the following statement or give a counterexample:

$$\text{Rank}(C) = \text{Rank}(A) + \text{Rank}(B).$$

11. Prove the following converse to the Pythagorean theorem: For any $x, y \in \mathbb{R}^n$, if $\|x + y\|^2 = \|x\|^2 + \|y\|^2$ then $\langle x, y \rangle = 0$.

12. Let $u_1, \ldots, u_k \in \mathbb{R}^n$ be orthonormal. For any $x \in \mathbb{R}^n$ prove that

$$\sum_{i=1}^{k} \langle x, u_k \rangle^2 \le \|x\|^2.$$

13. Suppose $A \in \mathbb{R}^{n \times n}$ is a square matrix whose singular values are all equal to 1. Prove that A is orthogonal.

14. Suppose $A, B \in \mathbb{R}^{n \times n}$ are symmetric and there is an orthogonal matrix $Q \in \mathbb{R}^{n \times n}$ with $QAQ^T = B$. Prove that A and B have the same eigenvalues.

15. Let $v_1, \ldots, v_m \in \mathbb{R}^n$ be orthonormal. Prove that $x \in \text{Span}(v_1, \ldots, v_m)$ if and only if

$$\|x\|^2 = \langle x, v_1 \rangle^2 + \cdots + \langle x, v_m \rangle^2.$$

16. Prove that any rank k matrix $A \in \mathbb{R}^{n \times n}$ can be written as $A = UV^T$ where $U, V \in \mathbb{R}^{n \times k}$ and $\text{Rank}(U) = \text{Rank}(V) = k$.

17. Let $A, B \in \mathbb{R}^{n \times n}$. Show that

   (a) $\text{Tr}(AB) = \text{Tr}(BA)$.

   (b) For any orthonormal basis $v_1, ..., v_n$, $\sum_{i=1}^{n} v_i^T A v_i = \text{Tr}(A)$.

   (c) If A is symmetric and its eigenvalues are $\lambda_1, .., \lambda_n$, then $\sum_{i=1}^{n} \lambda_i = \text{Tr}(A)$.

18. Let $A = U \Sigma V^T$ denote the Singular Value Decomposition of $A \in \mathbb{R}^{m \times n}$. Using this decomposition, construct the orthogonal projection on the image of A ($\text{Im}(A)$). What about the orthogonal projection on its orthogonal complement?

19. Show that for any $x, y \in \mathbb{R}^n$ we have $x^T y = \dfrac{\|x+y\|^2 - \|x-y\|^2}{4}$

20. Let $A \in \mathbb{R}^{n\times n}$ be a square matrix with the (unusual) property that the image space (or column space) $\mathrm{Im}(A)$ of A is equal to its kernel (or nullspace) $\ker(A)$.

    (a) What can you say about A^2?

    (b) Give an example of such an A.

21. Which of the following functions $f : \mathbb{R}^n \to \mathbb{R}$ are convex? (justify your answer)

    (a) $f(x) = \|x\|^2$,

    (b) $f(x) = -\|x\|^2$,

    (c) $f(x) = -(e_1^T x)^3$

    (d) $f(x) = -e_1^T x$,

22. Let A and b be given by:

$$A = \begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix} \text{ and } b = \begin{bmatrix} 6 \\ 2 \end{bmatrix}.$$

    (a) Write the quadratic $f(x) = x^T A x - b^T x$ associated with the above matrix in terms of x_1, x_2, the entries of x.

    (b) Determine the points for which the gradient vanishes. Are they local/global maxima, local/global minima, or neither (saddle points)?.

23. Find a basis for the orthogonal complement of the subspace

$$V = \left\{ \begin{bmatrix} x+y \\ x+y \\ x-y \\ x-y \end{bmatrix} : x, y \in \mathbb{R} \right\}$$

24. ($\star$) Let $\mathcal{P}_d$ denote the space of univariate polynomials $p(x)$ of degree smaller or equal to d, $\deg(p(x)) \le d$.

    (a) Is $\mathcal{P}_d$ a vector space? (with the normal adition of polynomials and scalar multiplication)

    (b) Let $\mathcal{D}$ denote the differentiaion operator, meaning that $\mathcal{D}(p(x)) = p'(x)$. For example, $\mathcal{D}\left(x^2\right) = 2x$. Prove this operator is linear.

    (c) What is the kernel $\ker(\mathcal{D})$ of $\mathcal{D}$?

    (d) What is the image space $\mathrm{Im}(\mathcal{D})$ of $\mathcal{D}$?

    (e) For fixed $s \in \mathbb{R}$ consider the operator $T_s : \mathcal{P}_d \to \mathcal{P}_d$ that adds s to the argument: $T_s(p) = p_s$ where $p_s(x) = p(x+s)$.

        i. Show this map is linear.

        ii. ($\star$) Represent this operator in terms of $\mathcal{D}$.

iii. ($\star$) Suppose we choose the basis $1, x, \ldots, x^d$ for the space $\mathcal{P}_d$. Show that the matrix corresponding to T_1 has a very nice form. [Hint: Pascal's Triangle.]

25. ($\star$) Let $A = U\Sigma V^T$ denote the Singular Value Decomposition of $A \in \mathbb{R}^{m\times n}$. Let $A' = U\Sigma' V^T$ where Σ' is obtained from Σ by replacing every entry by zero except for the entry corresponding to the largest singular value.

    (a) Show that A' is the best rank 1 approximation of A in the Frobenius norm, meaning that A' is the solution to $\min_{B:\ \mathrm{rank}(B)=1} \|B - A\|_F$.

    (b) Show that A' is the best rank 1 approximation of A in the spectral norm, meaning that A' is the solution to $\min_{B:\ \mathrm{rank}(B)=1} \|B - A\|$.

26. Let $x \in \mathbb{R}^n$ such that $\|x\| = 1$ and let $H \in \mathbb{R}^{n\times n}$ be given by $H = I - 2xx^T$ (H is a so-called Householder matrix). Show that all eigenvalues of H are either $\lambda = 1$ or $\lambda = -1$.

27. ($\star$) Let G denote a d-regular graph on n nodes (a simple graph where every node has degree d). Show that the largest eigenvalue of the adjacency matrix of a d-regular graph is d.

28. A matrix $A \in \mathbb{R}^{n\times n}$ is lower triangular if $A_{ij} = 0$ for all $j > i$. Show that if $A, B \in \mathbb{R}^{n\times n}$ are both lower triangular then AB is also lower triangular.

29. Given a matrix $A \in \mathbb{R}^{m\times n}$, which of the following sets $S \subset \mathbb{R}^n$ are a subspace, a convex set that is not necessarily a subspace, or not necessarily a convex set (justify your answer, either with a proof or with a counter example)

    (a) $S = \{x \in \mathbb{R}^n :\ \|x\|^2 < 1\}$

    (b) $S = \{x \in \mathbb{R}^n :\ \|x\|^2 > 1\}$

    (c) $S = \mathrm{Im}(A)$

30. Which of the following functions $f : \mathbb{R}^n \to \mathbb{R}$ are nonconvex, convex, or strictly convex? (justify your answer)

    (a) $f(x) = e_1^T x$,

    (b) $f(x) = \|x\|$

    (c) $f(x) = (e_1^T x)^2$

    (d) $f(x) = (e_1^T x)^5$

    (e) $f(x) = -\|x\|$,

31. Show that $\sum_{k=1}^n \sqrt{k} x_k \le n\|x\|$ for all $x \in \mathbb{R}^n$.

32. Show that if $A \succ 0$ there exists $T \succ 0$ such that $A = T^2$. Recall that $A \succ 0$ means that A is symmetric and all its eigenvalues are positive.

33. Let's say we are given 6 data points (x, y) in 2 dimensions: $(-1, -2), (1, -2), (-1, 0), (1, 0), (-1, 2)$, and $(1, 2)$.

(a) Describe the result of one-dimensional PCA, attempting to approximate the data by a line.

(b) Describe the result of linear regression, attempting to write y a an affine function of x.