COL 726 Assignment 2

27 February – 13 March, 2021

Note: All answers should be accompanied by a rigorous justification, unless the question explicitly states that a justification is not necessary.

Updated text is highlighted in blue.

1. Suppose $\mathbf{A}, \mathbf{B} \in \mathbb{C}^{m\times m}$ are two square matrices, and $\mathbf{C} = \mathbf{AB}$. Let the singular values of the three matrices be $a_1 \geq \cdots \geq a_m$, $b_1 \geq \cdots \geq b_m$, and $c_1 \geq \cdots \geq c_m$ respectively. Prove that $a_1 b_1 \geq c_1 \geq \max(a_1 b_m, a_m b_1)$, and similarly, $\min(a_1 b_m, a_m b_1) \geq c_m \geq a_m b_m$.

2. Let S be a subspace of $\mathbb{C}^m$. In the lectures, we defined its orthogonal complement as a subspace T such that $S \cap T = \{0\}$, $S + T = \mathbb{C}^m$, and $S \perp T$ (S is orthogonal to T). Show that this is precisely the set $\{\mathbf{v} \in \mathbb{C}^m : \mathbf{u}^*\mathbf{v} = 0 \text{ for all } \mathbf{u} \in S\}$.

3. Consider a linearly independent set of n real vectors $\mathbf{x}_1, \ldots, \mathbf{x}_n \in \mathbb{R}^m$. Suppose another set of vectors $\mathbf{y}_1, \ldots, \mathbf{y}_n \in \mathbb{R}^m$ is "congruent" to it, in the sense that all lengths and distances are equal: $\|\mathbf{x}_i\|_2 = \|\mathbf{y}_i\|_2$ for all i, and $\|\mathbf{x}_i - \mathbf{x}_j\|_2 = \|\mathbf{y}_i - \mathbf{y}_j\|_2$ for all $i \neq j$. Define the matrices $\mathbf{X} = [\mathbf{x}_1, \ldots, \mathbf{x}_n]$ and $\mathbf{Y} = [\mathbf{y}_1, \ldots, \mathbf{y}_n]$.

   (a) Prove that the reduced QR factorizations of $\mathbf{X}$ and $\mathbf{Y}$ have the same $\hat{\mathbf{R}}$.

   Hint: What can you say about the relationship between $\mathbf{X}^T\mathbf{X}$ and $\mathbf{Y}^T\mathbf{Y}$?

   (b) Give an algorithm to find an orthogonal matrix $\mathbf{Q}$ such that $\mathbf{Q}\mathbf{x}_i = \mathbf{y}_i$ for all i.

4. Consider a matrix $\mathbf{A} \in \mathbb{C}^{m\times n}$ and a vector $\mathbf{v} \in \mathbb{C}^m$. Let $\mathbf{F} = \mathbf{I} - 2\dfrac{\mathbf{v}\mathbf{v}^*}{\mathbf{v}^*\mathbf{v}}$.

   (a) Show that $\mathbf{FA} = \mathbf{A} + \mathbf{v}\mathbf{w}^*$ for some vector $\mathbf{w}$. Find the asymptotic operation count for both ways of computing $\mathbf{FA}$: (i) first computing $\mathbf{F}$ and then performing matrix multiplication, vs. (ii) first computing $\mathbf{w}$, then $\mathbf{v}\mathbf{w}^*$, and then matrix addition.

   (b) Suppose we use an approximate vector $\tilde{\mathbf{v}}$ and obtain $\tilde{\mathbf{F}} = \mathbf{I} - 2\dfrac{\tilde{\mathbf{v}}\tilde{\mathbf{v}}^*}{\tilde{\mathbf{v}}^*\tilde{\mathbf{v}}}$ instead. Show that if $\|\tilde{\mathbf{v}} - \mathbf{v}\|_2 / \|\mathbf{v}\|_2 = O(\epsilon_{\text{machine}})$, then $\|\tilde{\mathbf{F}} - \mathbf{F}\|_2 = O(\epsilon_{\text{machine}})$, and $\text{fl}(\tilde{\mathbf{F}}\mathbf{A}) = \mathbf{F}(\mathbf{A} + \delta\mathbf{A})$ for some $\delta\mathbf{A}$ with $\|\delta\mathbf{A}\|_2 / \|\mathbf{A}\| = O(\epsilon_{\text{machine}})$.

   Note: $\text{fl}(\tilde{\mathbf{F}}\mathbf{A})$ means the matrix product is evaluated exactly and then rounded.

5. Suppose we want to find the general least-squares solution to the linear system $\mathbf{Ax} = \mathbf{b}$, where $\mathbf{A} \in \mathbb{C}^{m\times n}$ has $m > n$ and $\text{rank}(\mathbf{A}) = r < n$. Let the full SVD of $\mathbf{A}$ be $\mathbf{A} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^*$.

   (a) Give an explicit formula for the unique vector $\mathbf{y} \in \text{range}(\mathbf{A})$ which minimizes $\|\mathbf{b} - \mathbf{y}\|_2$.

   (b) Find the general solution of the least-squares problem, i.e. all vectors $\mathbf{x}$ for which $\|\mathbf{b} - \mathbf{Ax}\|_2$ is minimal. Which of these vectors minimizes $\|\mathbf{x}\|_2$?

6. Consider the set P_n of all polynomials of degree $\leq n$ with complex coefficients. Any such polynomial p can be represented as a coefficient vector $[p] \in \mathbb{C}^{n+1}$ via

$$p(x) = a_0 + a_1 x + \cdots + a_n x^n \qquad \Longleftrightarrow \qquad [p] = \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{bmatrix}.$$

Suppose we define an inner product on polynomials as $(p, q) = \int_{-1}^{1} \overline{p(x)} q(x)\, \mathrm{d}x$. Thus two polynomials p, q are orthogonal if $\int_{-1}^{1} \overline{p(x)} q(x)\, \mathrm{d}x = 0$.

(a) Show that there exists a Hermitian matrix $\mathbf{G}$ such that $(p, q) = [p]^* \mathbf{G} [q]$ for all polynomials $p, q \in P_n$.

(b) If $p, q \in P_n$ are two nonzero polynomials, what does it mean to project q orthogonally onto the subspace $\langle p \rangle$ with respect to this inner product? Give an algebraic definition of orthogonal projection, and a formula for the corresponding matrix that acts on the coefficient vector $[q]$.

(c) Given a set of polynomials $p_1, \ldots, p_k$, we can now apply a Gram-Schmidt procedure to obtain a set of orthogonal polynomials (cf. Trefethen and Bau 7). Design and implement such an algorithm as a function $\mathbf{Q} =$ `orthogonalizePolynomials`$(\mathbf{P})$, which takes a matrix $\mathbf{P}$ containing the coefficients of the polynomials $[p_j]$ as columns, and returns an analogous matrix $\mathbf{Q}$ for the orthogonal polynomials.

Apply your function to an identity matrix (representing the polynomials $1, x, x^2, \ldots$), and verify that the coefficients you obtain represent multiples of the Legendre polynomials.

Collaboration policy: Refer to the policy on the course webpage.

If you collaborated with others to solve any question(s) of this assignment, you must give their names in your submission. If you found part of a solution using some online resource, give its URL.

Submission: Submit a PDF of your answers for all questions to Gradescope. Submit the code for Question 6 to Moodle. Both submissions must be uploaded before the assignment deadline.

Code submissions should contain a single .py file which contains all the functions. Functions are permitted but not required to produce any side-effects like printing out values or drawing plots. Any results you are asked to show should go in the PDF.