COL726 Assignment 2

27 January – 15 February, 2022

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. Let the tall matrix $\mathbf{A} \in \mathbb{C}^{m\times n}$ with $m > n$ have singular value decomposition $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^*$ with distinct singular values $\sigma_1 > \sigma_2 > \cdots > \sigma_n$.

   (a) Show that $\inf_{\mathbf{x}\neq\mathbf{0}} \|\mathbf{A}\mathbf{x}\|_2/\|\mathbf{x}\|_2 = \sigma_n$, and it is attained exactly when $\mathbf{x}$ is a nonzero multiple of $\mathbf{v}_n$.

   The above result can be interpreted as saying that $\langle \mathbf{v}_n \rangle$ is the unique one-dimensional subspace within which $\|\mathbf{A}\mathbf{x}\|_2/\|\mathbf{x}\|_2 \leq \sigma_n$ for all nonzero $\mathbf{x}$. Generalize this result to k-dimensional subspaces as follows:

   (b) Show that $\|\mathbf{A}\mathbf{x}\|_2/\|\mathbf{x}\|_2 \leq \sigma_{n-k+1}$ for all nonzero $\mathbf{x} \in \langle \mathbf{v}_{n-k+1}, \ldots, \mathbf{v}_n \rangle$.

   (c) For any k orthonormal vectors $\mathbf{q}_1, \ldots, \mathbf{q}_k$, show that there exists a nonzero $\mathbf{x} \in \langle \mathbf{q}_1, \ldots, \mathbf{q}_k \rangle$ such that $\|\mathbf{A}\mathbf{x}\|_2/\|\mathbf{x}\|_2 \geq \sigma_{n-k+1}$.

   Hint: Show that there is a nonzero $\mathbf{x} \in \langle \mathbf{q}_1, \ldots, \mathbf{q}_k \rangle$ which is orthogonal to $\langle \mathbf{v}_{n-k+2}, \ldots, \mathbf{v}_n \rangle$.

2. Suppose $\mathbf{P}$ and $\mathbf{Q}$ are two Hermitian matrices such that $2\mathbf{P} = \mathbf{I} + \mathbf{Q}$. Show that the following are equivalent (i.e. one is true if and only if all are true):

   - $\mathbf{P}$ is a projector.
   - $\mathbf{Q}$ is a unitary matrix.
   - $\mathbf{P}$ has an SVD with $\sigma_j \in \{0, 1\}$ and $\mathbf{u}_j = \mathbf{v}_j$ for all j.
   - $\mathbf{Q}$ has an SVD with $\sigma_j = 1$ for all j.

3. (a) For any Hermitian matrix $\mathbf{A} \in \mathbb{C}^{m\times m}$, there exists an orthonormal set of eigenvectors $\mathbf{x}_1, \ldots, \mathbf{x}_m$ and real eigenvalues $\lambda_1, \ldots, \lambda_m \in \mathbb{R}$ (not necessarily nonnegative) such that $\mathbf{A}\mathbf{x}_j = \lambda_j\mathbf{x}_j$ for all $j = 1, \ldots, m$. Using this fact, construct a singular value decomposition of $\mathbf{A}$ in terms of the $\mathbf{x}_j$'s and λ_j's.

   (b) Show that any unitary matrix $\mathbf{Q} \in \mathbb{C}^{m\times m}$ can be expressed as the product of Hermitian matrices which each have $\lambda_1 = -1$ and $\lambda_2 = \cdots = \lambda_m = 1$.

   Hint: One of the factorization algorithms we have already studied gets you most of the way there. Can you prove that the steps have the desired properties?

4. Consider the 4×3 matrix $\mathbf{A}$ given below, with $\mu = O(\sqrt{\epsilon_\mathrm{m}})$.

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 \\ \mu & & \\ & \mu & \\ & & \mu \end{bmatrix}$$

Suppose its QR factorization is computed in floating-point arithmetic with (a) the classical Gram-Schmidt method, and (b) the modified Gram-Schmidt method. For both cases, work out on paper the columns $\tilde{\mathbf{q}}_1, \tilde{\mathbf{q}}_2, \tilde{\mathbf{q}}_3$ and $\tilde{\mathbf{r}}_1, \tilde{\mathbf{r}}_2, \tilde{\mathbf{r}}_3$ of the computed factorization, along with bounds on the errors of all the entries. For example,

$$\tilde{\mathbf{q}}_1 = \begin{bmatrix} 1 + O(\epsilon_\mathrm{m}) \\ \mu + O(\mu\epsilon_\mathrm{m}) \\ 0 \\ 0 \end{bmatrix}.$$

Assume that addition of 0 and multiplication by 0 or 1 do not incur any rounding error. What can you conclude about the orthogonality of the computed $\mathbf{Q}$ in both cases?

5. Consider two full rank matrices $\mathbf{A} \in \mathbb{C}^{m \times n_1}, \mathbf{B} \in \mathbb{C}^{m \times n_2}$ with $n_1, n_2 < m < n_1 + n_2$, and let $S = \mathrm{range}(\mathbf{A}), T = \mathrm{range}(\mathbf{B})$. Suppose we want to find an orthonormal basis for $S \cap T$.

   (a) Show that $(S \cap T)^\perp = S^\perp + T^\perp$. Use this fact to find an orthonormal basis for $S \cap T$ via three QR factorizations. What is the computational complexity of this method?

   You may use either (or both) definitions: (i) $S_1^\perp = S_2$ if and only if $S_1 \cap S_2 = \{0\}$, $S_1 + S_2 = \mathbb{C}^m$, and $S_1 \perp S_2$, or simply (ii) $S_1^\perp = \{\mathbf{y} \in \mathbb{C}^m : \mathbf{y} \perp \mathbf{x} \text{ for all } \mathbf{x} \in S_1\}$.

   (b) Implement the method as a Python function `intBasis(A, B)`, using `scipy.linalg.qr` for the QR factorizations. The function should return a matrix $\mathbf{Q}$ whose columns are orthonormal and form a basis of $S \cap T$.

   (c) Also define a procedure to validate the algorithm's results. That is, given a matrix $\mathbf{Q}$, how can one verify that (i) its columns are orthonormal, (ii) $\mathrm{range}(\mathbf{Q}) \subseteq \mathrm{range}(\mathbf{A}) \cap \mathrm{range}(\mathbf{B})$, and (iii) $\mathrm{range}(\mathbf{Q})^\perp \subseteq \mathrm{range}(\mathbf{A})^\perp + \mathrm{range}(\mathbf{B})^\perp$, up to some tolerance ϵ? Describe the procedure in your PDF, and implement it as another function `checkIntBasis(Q, A, B, eps)`.

   Note: For subspaces S_1, S_2, let us say $S_1 \subseteq S_2$ to within tolerance ϵ if for all nonzero $\mathbf{x} \in S_1$, the nearest $\mathbf{y} \in S_2$ satisfies $\|\mathbf{y} - \mathbf{x}\|_2 / \|\mathbf{x}\|_2 \leq \epsilon$.

If $S = \{\mathbf{x} \in \mathbb{C}^3 : x_1 = x_2\}$ and $T = \{\mathbf{x} \in \mathbb{C}^3 : x_2 = x_3\}$, then clearly $S \cap T = \{\mathbf{x} \in \mathbb{C}^3 : x_1 = x_2 = x_3\}$. Define appropriate matrices $\mathbf{A}$ and $\mathbf{B}$ and verify that `intBasis` gives the right result, both by inspection and via `checkIntBasis`. You don't have to submit this part.

6. Suppose you have a set of reference points $\mathbf{p}_1, \ldots, \mathbf{p}_n \in \mathbb{R}^2$, and a corresponding set of observed points $\mathbf{q}_1, \ldots, \mathbf{q}_n \in \mathbb{R}^2$. You want to fit an affine transformation to these observations, $\mathbf{q}_i \approx \mathbf{S}\mathbf{p}_i + \mathbf{t}$ for some unknown $\mathbf{S}$ and $\mathbf{t}$. This is a common task in vision and graphics applications.

   (a) Formulate this problem in the form $\min_{\mathbf{x}} \|\mathbf{b} - \mathbf{A}\mathbf{x}\|_2$. State precisely the sizes and entries of the matrix $\mathbf{A}$ and the vector $\mathbf{b}$, and what the entries of $\mathbf{x}$ correspond to.

   (b) Show that $\mathbf{A}$ is rank-deficient if and only if the points $\mathbf{p}_1, \ldots, \mathbf{p}_n$ are collinear.

   (c) Implement a Python function `affineFit(p, q)` to solve this problem using QR factorization. It should take as input $[\mathbf{p}_1, \ldots, \mathbf{p}_n], [\mathbf{q}_1, \ldots, \mathbf{q}_n] \in \mathbb{R}^{2\times n}$ and return $\mathbf{S} \in \mathbb{R}^{2\times 2}$ and $\mathbf{t} \in \mathbb{R}^2$.

   To test your implementation, try choosing random $\mathbf{p}_1, \ldots, \mathbf{p}_n$ and $\mathbf{S}, \mathbf{t}$, then computing $\mathbf{q}_i = \mathbf{S}\mathbf{p}_i + \mathbf{t}$ (maybe with some added noise) and seeing if `affineFit` recovers $\mathbf{S}, \mathbf{t}$ (to within $O(\epsilon_{\mathrm{m}})$ without noise, only approximately with noise). Again, don't submit this part.

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

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

Submission: This assignment has two submission forms on Gradescope. In one form, you have to submit a PDF of your answers for all questions, and in the other you have to submit your code for Questions 5 and 6. Both submissions must be uploaded before the assignment deadline.

Code submissions should contain two files a2q5.py and a2q6.py which contain the requested functions and any helper functions.