Michigan State University

CMSE 823 – Numerical Linear Algebra **Spring 2020**

Final Exam

Alexander Harnisch

Problem 1

(a)

For $A \in \mathbb{C}^{m\times n}$:

$$\|A\|_\infty \leq \sqrt{n}\|A\|_2. \tag{1}$$

Proof. First, using the definition [2, (3.2)] of the 2-norm and infinity norm for any $x \in \mathbb{C}^n$:

$$\|x\|_2^2 = \sum_{i=1}^{n} x_i^2 \leq n \max_{1\leq i\leq m} x_i^2 = n\|x\|_\infty^2.$$

And thus

$$\|x\|_2 \leq \sqrt{n}\|x\|_\infty \Leftrightarrow \frac{1}{\|x\|_\infty} \leq \frac{\sqrt{n}}{\|x\|_2}, \quad x \neq 0. \tag{2}$$

Additionally, we have

$$\|x\|_\infty = \max_{1\leq i\leq n} |x_i| \leq \sqrt{\sum_{i=1}^{n} x_i^2} = \|x\|_2 \Leftrightarrow \frac{1}{\|x\|_2} \leq \frac{1}{\|x\|_\infty}, \quad x \neq 0. \tag{3}$$

Combining these results with the definition of the induced matrix norm [2, (3.6)], we find the inequality:

$$\|A\|_\infty = \sup_{x\in\mathbb{C}^n,\, x\neq 0} \frac{\|Ax\|_\infty}{\|x\|_\infty} \overset{(3)}{\leq} \sup_{x\in\mathbb{C}^n,\, x\neq 0} \frac{\|Ax\|_2}{\|x\|_\infty} \overset{(2)}{\leq} \sup_{x\in\mathbb{C}^n,\, x\neq 0} \frac{\sqrt{n}\|Ax\|_2}{\|x\|_2} = \sqrt{n}\|A\|_\infty$$

□

We also find the inequality

$$\|A\|_2 \leq \sqrt{m}\|A\|_\infty.$$

Proof.

$$\|A\|_2 = \sup_{x\in\mathbb{C}^n,\, x\neq 0} \frac{\|Ax\|_2}{\|x\|_2} \overset{(2)}{\leq} \sup_{x\in\mathbb{C}^n,\, x\neq 0} \frac{\sqrt{m}\|Ax\|_\infty}{\|x\|_2} \overset{(3)}{\leq} \sup_{x\in\mathbb{C}^n,\, x\neq 0} \frac{\sqrt{m}\|Ax\|_\infty}{\|x\|_\infty} = \sqrt{m}\|A\|_\infty$$

□

(b)

The inequality (1) is sharp.

Proof. To show that the inequality is sharp, we only need to find an example for which equality holds. Such an example is:

$$M = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}.$$

$\|M\|_\infty = 2$ and $\|M\|_2 = \sqrt{2}$ so

$$\|M\|_\infty = \sqrt{2}\|M\|_2.$$

□

Problem 2

Let $A \in \mathbb{C}^{m\times n}$ with $m \geq n$. A^*A is non-singular if and only if A has full rank.

Proof. The rank of A is the number of non-zero singular values [2, Theorem 5.1]. Let the SVD of A be $A = U\Sigma V^*$. Then it follows, that A is of full rank if and only if Σ is of full rank and

$$A^*A = V\Sigma^*U^*U\Sigma V^* = V\Sigma^*\Sigma V^* = V\Sigma^2V^*.$$

This is an eigenvalue decomposition (similarity transformation). Therefore, A^*A is of full rank if and only if Σ^2 is of full rank, which is of full rank if and only if Σ is of full rank. Σ is diagonal and its diagonal entries are the singular values of A, so Σ is of full rank if and only if A is of full rank. A^*A is a square matrix, so it is non-singular if and only if it has full rank. We conclude that A^*A is non-singular if and only if A is of full rank. □

Problem 3

Let $A \in \mathbb{C}^{m\times m}$, and let a_j be its j-th column. Then:

$$|\det(A)| \leq \prod_{j=1}^{m} \|a_j\|_2. \tag{4}$$

Proof. Let $A = QR$ be the QR decomposition of A. Let q_j be the j-th column of Q and r_{ij} be the (i, j) entry of R.

From $a_j = \sum_{i=1}^{j} r_{ij}q_i$ [2, (7.3)] and $r_{ij} = q_i^*a_j$ for $i \neq j$ [2, (7.7)] and the fact that the columns of Q are orthonormal, it follows directly, that:

$$\|a_j\|_2^2 = \left(\sum_{i=1}^{j} r_{ij}q_i\right)^2 = \sum_{i=1}^{j} r_{ij}^2 = \sum_{i=1}^{j} |q_i^*a_j|^2. \tag{5}$$

With that, we find:

$$\begin{aligned}\det(A)^2 = \det(A^*A) = \det(R^*Q^*QR) = \det(R^*R) = \det(R)^2 \\ = \prod_{j=1}^{m} |r_{jj}|^2 \overset{[2,\,(7.7)]}{=} \prod_{j=1}^{m} |q_j^* a_j|^2 \\ \leq \prod_{j=1}^{m} \sum_{i=1}^{j} |q_i^* a_j|^2 \overset{(5)}{=} \prod_{j=1}^{m} \|a_j\|_2^2\end{aligned}$$

Taking the square root gives (4). □

Problem 4

(a)

The primary textbook does not mention an SVD theorem. It does, however, give a formal definition of the Singular Value decomposition [2, pp. 28-29]. The existence and uniqueness is stated and proven [2, Theorem 4.1]. A slide titled “SVD: Theorem” can be found in [1]. It states, that

1. Every matrix $A \in \mathbb{C}^{m\times n}$ has an SVD.
2. The singular values $\{\sigma_i\}$ are uniquely determined.
3. If $A \in \mathbb{C}^{m\times n}$ and the singular values σ_i are distinct, then the left and right singular vectors $\{u_i\}$ and $\{v_i\}$ are uniquely determined up to complex signs (i.e. complex scalar factors of modulus 1).

This is the same as Theorem 4.1 from [2].

(b)

Let $A \in \mathbb{C}^{m\times n}$. Set $\sigma = \|A\|_2$. There are vectors $v \in \mathbb{C}^n$ and $u \in \mathbb{C}^m$ with $\|v\|_2 = \|u\|_2 = 1$ such that $Av = \sigma u$.

Proof. Let $A = U\Sigma V^*$ be the SVD of A. Σ is diagonal and its entries are the singular values of A. $\sigma = \|A\|_2$ is the largest singular value of A [2, Theorem 5.3]. Thus:

$$AV = U\Sigma \Rightarrow Av = \sigma u, \tag{6}$$

where v is the right and u the left normalized singular vector corresponding to the largest singular value σ of A. □

This proof might be circular, depending on how the SVD theorem stated above has been proven. An alternative proof is the compactness argument of the induced 2 norm used in the proof of Theorem 4.1 in [2]. This proof is also given in [1]:

Existence and uniqueness of SVD

Lemma. Let $\sigma_1 = \|A\|_2$. Then there exist $\mathbf{u}_1 \in \mathcal{C}^m$ and $\mathbf{v}_1 \in \mathcal{C}^n$ such that $A\mathbf{v}_1 = \sigma_1 \mathbf{u}_1$, where $\|\mathbf{u}_1\|_2 = 1$ and $\|\mathbf{v}_1\|_2 = 1$.

Proof. By definition, we have

$$\|A\|_2 = \sup_{\|\mathbf{y}\|_{2(n)}=1} \|A\mathbf{y}\|_{2(m)}.$$

Since the function $\|A\mathbf{y}\|_{2(m)}$ is continuous on the compact set $\{\mathbf{y} \in \mathcal{C}^n : \|\mathbf{y}\|_{2(n)} = 1\}$, there exists $\mathbf{x} \in \mathcal{C}^n$, $\|\mathbf{x}\|_{2(n)} = 1$, so that

$$\|A\|_2 = \sup_{\|\mathbf{y}\|_{2(n)}=1} \|A\mathbf{y}\|_{2(m)} = \max_{\|\mathbf{y}\|_{2(n)}=1} \|A\mathbf{y}\|_{2(m)} = \|A\mathbf{x}\|_{2(m)}.$$

Let $\mathbf{v}_1 = \mathbf{x}$ so that $\|\mathbf{v}_1\|_2 = 1$. Let $\mathbf{z} = A\mathbf{x}$. Define $\mathbf{u}_1 = \frac{\mathbf{z}}{\|\mathbf{z}\|_2}$. Then

$$\|\mathbf{z}\|_{2(m)} = \sigma_1 = \|A\|_2, \quad \mathbf{z} = \|A\|_2 \mathbf{u}_1 = \sigma_1 \mathbf{u}_1.$$

Thus $A\mathbf{v}_1 = \sigma_1 \mathbf{u}_1$, where $\|\mathbf{u}_1\|_{2(m)} = 1$ and $\|\mathbf{v}_1\|_{2(n)} = 1$.
In particular, $\mathbf{u}_1^* A \mathbf{v}_1 = \sigma_1$.

Figure 1: Proof of (b) given in [1].

(c)

A reduced SVD of

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

is

$$A = U\Sigma V^* = \begin{bmatrix} 0 & 1 \\ 0 & 0 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

To make it a full SVD, we only need to make the columns of U a full basis of $\mathbb{R}^3$ by adding the column $(0,1,0)^\top$ and adding a row of zeros to Σ.

Just as a short justification of the result: For this simple matrix I could come up with it in my head by knowing that Σ is a diagonal matrix with the square roots of the eigenvalues of $A^\top A$ on its diagonal [2, Theorem 5.4].

Problem 5

We want to solve the least-squares problem $\min_{x\in\mathbb{R}^2} \|Ax-b\|_2$, where $b=(0,0,3,2)^\top$ and

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

using Algorithm 11.2 defined in [2]. The result can be cross-checked numerically and by using [2, (11.12)].

1. The first step is to compute the QR decomposition of A. This can most conveniently be achieved by Gram-Schmidt Orthogonalization or Householder Triangularization. Let's choose the former, because it requires less typing:

$$q_1' = a_1 = \begin{bmatrix}1\\0\\2\\0\end{bmatrix}$$

$$r_{11} = \|q_1'\| = \sqrt{1^2+2^2} = \sqrt{5}$$

$$q_1 = \frac{q_1'}{\|q_1'\|} = \frac{1}{\sqrt{5}}\begin{bmatrix}1\\0\\2\\0\end{bmatrix}$$

$$r_{12} = q_1^\top a_2 = \frac{1}{\sqrt{5}}(-1+2) = \frac{1}{\sqrt{5}}$$

$$q_2' = a_2 - r_{12}q_1 = \begin{bmatrix}-1\\0\\1\\2\end{bmatrix} - \frac{1}{5}\begin{bmatrix}1\\0\\2\\0\end{bmatrix} = \frac{1}{5}\begin{bmatrix}-6\\0\\3\\10\end{bmatrix}$$

$$r_{22} = \|q_2'\| = \sqrt{5.8}$$

$$q_2 = \frac{q_2'}{\|q_2'\|} = \frac{1}{5\sqrt{5.8}}\begin{bmatrix}-6\\0\\3\\10\end{bmatrix}$$

And thus:

$$A = QR = \begin{bmatrix}\frac{1}{\sqrt{5}} & -\frac{6}{5\sqrt{5.8}}\\ 0 & 0\\ \frac{2}{\sqrt{5}} & \frac{3}{5\sqrt{5.8}}\\ 0 & \frac{10}{5\sqrt{5.8}}\end{bmatrix}\begin{bmatrix}\sqrt{5} & \frac{1}{\sqrt{5}}\\ 0 & \sqrt{5.8}\end{bmatrix}.$$

2. Compute $y = Q^*b = Q^\top b$:

$$y^\top = b^\top Q = \begin{bmatrix}0 & 0 & 3 & 2\end{bmatrix}\begin{bmatrix}\frac{1}{\sqrt{5}} & -\frac{6}{5\sqrt{5.8}}\\ 0 & 0\\ \frac{2}{\sqrt{5}} & \frac{3}{5\sqrt{5.8}}\\ 0 & \frac{10}{5\sqrt{5.8}}\end{bmatrix} = \begin{bmatrix}\frac{6}{\sqrt{5}} & \frac{29}{5\sqrt{5.8}}\end{bmatrix}$$

3. Solve $Rx = y$ for x, which solves $\min_{x\in\mathbb{R}^2} \|Ax - b\|_2$.

$$\begin{bmatrix} \sqrt{5} & \frac{1}{\sqrt{5}} \\ 0 & \sqrt{5.8} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \frac{6}{\sqrt{5}} \\ \frac{29}{5\sqrt{5.8}} \end{bmatrix}$$

Note that $\sqrt{5.8} = \frac{\sqrt{145}}{5} = \frac{29}{5\sqrt{5.8}}$, thus the solution is

$$x = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

.

Problem 6

Let $A \in \mathbb{R}^{n\times n}$ be symmetric and positive definite with $n = j + k$. Partition A into the following 2 b 2 blocks:

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

where A_{11} is $j \times j$ and A_{22} is $k \times k$. Let R_{11} be the Cholesky factor of A_{11}: $A_{11} = R_{11}^\top R_{11}$, where R_{11} is upper-triangular with positive main-diagonal entries. Let $R_{12} = (R_{11}^{-1})^\top A_{12}$ and let $\tilde{A}_{22} = A_{22} - R_{12}^\top R_{12}$.

(a)

A_{11} is positive definite.

Proof. We know that A is positive definite:

$$x^\top Ax > 0 \quad \forall x \in \mathbb{R}^n \setminus \{0\}.$$

From that it follows, that for

$$\tilde{x} \in \left\{(y, 0)^\top | y \in \mathbb{R}^j\right\} \subset \mathbb{R}^n$$

we get

$$\tilde{x}^\top A\tilde{x} = y^\top A_{11}y > 0.$$

We conclude that if A is positive definite, A_{11} is also positive definite. □

(b)

$\tilde{A}_{22} = A_{22} - A_{21}A_{11}^{-1}A_{12}$

Proof. We know, that $A_{11} = R_{11}^\top R_{11}$, $R_{12} = (R_{11}^{-1})^\top A_{12}$ and $\tilde{A}_{22} = A_{22} - R_{12}^\top R_{12}$. We also know, that $A_{21} = A_{12}^\top$, because A is symmetric.

With that: $R_{12}^\top = A_{12}^\top R_{11}^{-1} = A_{21}R_{11}^{-1}$ and it follows, that

$$R_{12}^\top R_{12} = A_{21}R_{11}^{-1}(R_{11}^{-1})^\top A_{12} = A_{21}(R_{11}^\top R_{11})^{-1}A_{12} = A_{21}A_{11}^{-1}A_{12}$$

and thus $\tilde{A}_{22} = A_{22} - A_{21}A_{11}^{-1}A_{12}$. □

(c)

$\tilde{A}_{22}$ is positive definite. An informal proof is given by [2, p. 175]. However, I'll still write out a formal proof:

Proof. We can factorize A as follows:

$$A = UDU^\top \begin{bmatrix} \mathbb{I}_j & 0 \\ A_{21}A_{11}^{-1} & \mathbb{I}_k \end{bmatrix} \begin{bmatrix} A_{11} & 0 \\ 0 & \tilde{A}_{22} \end{bmatrix} \begin{bmatrix} \mathbb{I}_j & A_{11}^{-1}A_{12} \\ 0 & \mathbb{I}_k \end{bmatrix}$$

Which is a similarity transformation of A, because U is invertible:

$$U^{-1} = \begin{bmatrix} \mathbb{I}_j & 0 \\ -A_{21}A_{11}^{-1} & \mathbb{I}_k \end{bmatrix}.$$

So D can be written as $D = U^{-1}A(U^{-1})^T$ and it follows that

$$y^\top Dy = y^\top U^{-1}A(U^{-1})^T y = x^\top Ax > 0, \quad y \neq 0$$

with $x = (U^{-1})^\top y$. So D is positive definite if and only if A is positive definite. The lower right 2 by 2 block of D is $\tilde{A}_{22}$ and it is uncoupled from the upper left block of D, because the upper right and lower left blocks are zero. Therefore, by the same argument used in **(a)**, A is positive definite if and only if A_{11} and $\tilde{A}_{22}$ are both positive definite. □

Problem 7

If $A \in \mathbb{R}^{m\times m}$ is symmetric and positive definite, then solving the linear system $Ax = b$ amounts to computing

$$x = \sum_{i=1}^{m} \frac{c_i}{\lambda_i} v_i,$$

where λ_i are the eigenvalues of A and v_i are the corresponding eigenvectors, and c_i are some constants determined by b and v_i.

Proof. We know that A is symmetric and therefore unitarily diagonalizable [2, Theorem 24.7] and all its eigenvalues are positive real numbers (if $Ax = \lambda x$ for $x \neq 0$, we have $x^\top Ax = \lambda x^\top x > 0$). So we can write A as

$$A = Q\Lambda Q^\top, \quad QQ^\top = \mathbb{I}_m.$$

The columns of Q are the normalized eigenvectors of A v_i and Λ is diagonal with A's eigenvalues λ_i on its diagonal. Thus:

$$Ax = b \Leftrightarrow Q\Lambda Q^\top x = b$$

and it follows

$$x = \Lambda^{-1}(Q^\top b)Q = \Lambda^{-1}cQ = \sum_{i=1}^{m} \frac{c_i}{\lambda_i} v_i, \quad c_i = v_i^\top b.$$

□

Problem 8

Let $A \in \mathbb{C}^{m \times n}$, $m \geq n$, with linearly independent columns:

$$A = [a_1, \ldots, a_n].$$

We want to find eigenvalues and eigenvectors of the projection matrix

$$P = \mathbb{I} - A(A^*A)^{-1}A^* = \mathbb{I} - AA^+.$$

Let the SVD of A be $A = U\Sigma V^*$. So $A^*A = V\Sigma U^* U\Sigma V^* = V\Sigma^2 V^*$. And the pseudoinverse of A becomes

$$A^+ = (A^*A)^{-1}A^* = V\Sigma^{-1}U^*.$$

So AA^+ is the orthogonal projector onto the range of A. That means P is the orthogonal projector onto the nullspace of A.

Proof.

$$PA = (\mathbb{I} - A(A^*A)^{-1}A^*)A = A - AV\Sigma^{-1}U^*U\Sigma V^* = A - AV\Sigma^{-1}\Sigma V^* = 0$$

□

So we conclude that any vector, that lies in the nullspace of A, is an eigenvector of P with eigenvalue 1. Conversely, any vector inside the range of A is an eigenvector of P with eigenvalue 0. So all of A's columns a_i are eigenvectors of P with eigenvalue 0 (as proven above). And all multiples of the $m - n$ linearly independent vectors which, together with the n a_i vectors, form a full basis of $\mathbb{R}^m$, are eigenvectors of P with eigenvalue 1.

References

[1] Jianliang Qian. *Lecture Slides for CMSE 823 - Numerical Linear Algebra.* 2020.

[2] Lloyd N. Trefethen and David Bau. *Numerical Linear Algebra.* SIAM, 1997. ISBN 0-89871-361-7.