Instructions:

- The exam is open book. Please finish the exam by your own.
- All the answers should be typed in so that your work will be fully evaluated.
- Please email your finished exam to Prof. Jianliang Qian (email: jqian@msu.edu) by **5PM, Friday, May 1st, 2020.**
- No late exam will be accepted.

1. (a) (5 points) For $A \in \mathbf{C}^{m\times n}$, show that HW #2 3.3 (c)

$$\|A\|_\infty \le \sqrt{n}\|A\|_2.$$

(b) (5 points) Show that the inequality in part a) is sharp.

(a) Using the definition

$$\|A\|_{(m,n)} = \sup_{\substack{x\in\mathbb{C}^n \\ x\neq 0}} \frac{\|Ax\|_{(m)}}{\|x\|_{(n)}} = \sup_{\substack{x\in\mathbb{C}^n \\ \|x\|_{(n)}=1}} \|Ax\|_{(m)}. \qquad (3.6)$$

First, for any $x \in \mathbb{C}^n$:

$$\|x\|_2^2 = \sum_{i=1}^{n} x_i^2 \le n \cdot \max_{1\le i\le m} x_i^2 = n\,\|x\|_\infty^2$$

Definition of 2-norm [3.2] Definition of ∞-norm [3.2]

And thus: $\|x\|_2 \le \sqrt{n}\,\|x\|_\infty \iff \frac{1}{\sqrt{n}\|x\|_\infty} \le \frac{1}{\|x\|_2}$ (1)

Additionally: $\|x\|_\infty = \max_{1\le i\le n} |x_i| \le \sqrt{\sum_i x_i^2} = \|x\|_2$ (2)

With that: $(Ax) \in \mathbb{C}^m$

$$\|A\|_2 \overset{[3.6]}{=} \sup_{\substack{x\in\mathbb{C}^n \\ x\neq 0}} \frac{\|Ax\|_2}{\|x\|_2} \overset{(1)}{\le} \sup_{\substack{x\in\mathbb{C}^n \\ x\neq 0}} \frac{\sqrt{m}\,\|Ax\|_\infty}{\|x\|_2} \overset{(2)}{\le} \sup_{\substack{x\in\mathbb{C}^n \\ x\neq 0}} \frac{\sqrt{m}\,\|Ax\|_\infty}{\|x\|_\infty}$$

$$= \sqrt{m}\,\|A\|_\infty$$

$\iff \|A\|_2 \le \sqrt{m}\,\|A\|_\infty$ (3)

With that:

$$\|A\|_\infty = \sup_{\vec{x}\in\mathbb{C}^n\setminus\{0\}} \frac{\|A\vec{x}\|_\infty}{\|\vec{x}\|_\infty} \overset{(2)}{\leqslant} \sup_{\vec{x}\in\mathbb{C}^n\setminus\{0\}} \frac{\|A\vec{x}\|_2}{\|\vec{x}\|_\infty}$$

$$\overset{(1)}{\leqslant} \sup_{\vec{x}\in\mathbb{C}^n\setminus\{0\}} \frac{\|A\vec{x}\|_2}{\sqrt{n}\|\vec{x}\|_2} = \sup_{\vec{x}\in\mathbb{C}^n\setminus\{0\}} \sqrt{n}\frac{\|A\vec{x}\|_2}{\|\vec{x}\|_2} = \sqrt{n}\|A\|_2 \qquad \square$$

(b) We just need an example for which we get equality:

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$$

$$\|A\|_\infty = 2, \quad \|A\|_2 = \sqrt{2}$$

$$\Rightarrow \|A\|_\infty = \sqrt{n}\|A\|_2 \qquad \square$$

Proof: $\|A\|_2 = \sup_{\substack{x\in\mathbb{C}^n \\ \|x\|_2=1}} \|Ax\|_2$

obv. for $\sqrt{2}x = (1,1)^T$

$\Rightarrow \|A\|_2 = \|Ax\|_2$

$= \left\|\begin{pmatrix} \frac{2}{\sqrt{2}} \\ 0 \end{pmatrix}\right\|_2$

$= \sqrt{\frac{4}{2}}$

$= \sqrt{2}$

2. (10 points) Let $A \in C^{m \times n}$ with $m \geq n$. Show that A^*A is nonsingular if and only if A has full rank.

Let the SVD of A be $A = U\Sigma V^*$

Using

Theorem 5.1. *The rank of A is r, the number of nonzero singular values.*

Proof. The rank of a diagonal matrix is equal to the number of its nonzero entries, and in the decomposition $A = U\Sigma V^*$, U and V are of full rank. Therefore $\text{rank}(A) = \text{rank}(\Sigma) = r$. □

it follows that A is of full rank if and only if Σ is of full rank.

$A^*A = V\Sigma^*\Sigma V^*$ (Eigenvalue decomposition)

$\Rightarrow$ A^*A is non-singular if and only if Σ is of full rank.

$\Rightarrow$ A^*A is non-singular if and only if A is of full rank. □

Also: A^*A is a square matrix

$\Rightarrow$ A^*A non-singular $\Leftrightarrow$ A^*A is of full rank!

3. (10 points) Let $A \in C^{m\times m}$, and let $\mathbf{a}_j$ be its jth column. Prove the following inequality:

$$|\det(A)| \leq \prod_{j=1}^{m} \|\mathbf{a}_j\|_2.$$

Hadamard's inequality

Let $A = QR$ be the QR factorization of A (which allways exists [Theorem 7.1]).

Then:

$$(\det A)^2 = \det A^*A = \det R^* \underbrace{Q^*Q}_{I} R$$

$$= \det R^* R = (\det R)^2$$

$$= \prod_{j=1}^{m} |r_{jj}|^2 \overset{[7.7]}{=} \prod_{j=1}^{m} |q_j^* a_j|^2 \leq \prod_{j=1}^{m} \sum_{i=1}^{j} |q_i^* a_j|^2$$

$$\overset{\text{Lemma 1}}{=} \prod_{j=1}^{m} \|a_j\|_2^2$$

And thus $|\det A| \leq \prod_{j=1}^{m} \|a_j\|_2$. $\square$

Lemma 1

[7.3]: $a_j = \sum_{i=1}^{j} r_{ij} q_i$

$$\Rightarrow \|a_j\|_2^2 \underset{q_i^* q_k = \delta_{ik}}{=} \sum_{i=1}^{j} r_{ij}^2 \overset{[7.7]}{=} \sum_{i=1}^{j} |q_i^* a_j|^2$$

(Norm is conserved under base change).

4. (a) (5 points) State the Singular Value Decomposition (SVD) theorem for $A \in \mathbf{C}^{m \times n}$.

   (b) (5 points) Let $A \in \mathbf{C}^{m \times n}$. Set $\sigma = \|A\|_2$. Show that there are vectors $\mathbf{v} \in \mathbf{C}^n$ and $\mathbf{u} \in \mathbf{C}^m$ with $\|\mathbf{v}\|_2 = \|\mathbf{u}\|_2 = 1$ such that $A\mathbf{v} = \sigma\mathbf{u}$.

   (c) (5 points) Find an SVD of

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

(a) [P. 28-29: Formal Definition]
[and in SVD slides: SVD: Theorem]

1. SVD allways exists SVD: $U \Sigma V^*$
2. SVD is unique (except complex signs of singular vectors).

(b) [Theorem 5.3]: $\|A\|_2 = \sigma_{max} = \sigma_1$

with $A = U \Sigma V^* \Rightarrow AV = U\Sigma$

$\Rightarrow A v_1 = \sigma_1 u_1$ $\square$

In words: the vector v is the right and u the left singular vector corresponding to the largest singular value $\sigma_1 = \|A\|_2$.

(c)

$$A = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \\ -1 & 0 \end{pmatrix}\begin{matrix} 0 \\ -1 \\ 0 \end{matrix} \begin{matrix} \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} \\ 0 \quad 0 \end{matrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

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

$$= \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 2 \end{pmatrix} \checkmark$$

5. (15 points) Let

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

Use the QR factorization of A to solve the least-squares problem

$$\min \|A\mathbf{x} - \mathbf{b}\|_2,$$

where $\mathbf{b} = [0, 0, 3, 2]^T$ with T indicating transpose.

Algorithm 11.2 | Check with 11.12

$$QR \approx \begin{pmatrix} 0.4472 & -0.4983 \\ 0 & 0 \\ 0.8944 & 0.2491 \\ 0 & 0.8305 \end{pmatrix} \begin{pmatrix} 2.2361 & 0.4472 \\ 0 & 2.4083 \end{pmatrix}$$

$$y = Q^* b \approx (2.68328157, 2.40831552)^T$$

$$\Rightarrow \quad Rx = y \quad \Rightarrow \quad x_2 \approx 1$$

$$x_1 \approx \frac{2.68328157 - 0.4472}{2.2361}$$

$$\approx 1$$

$$x^T = (1, 1) \qquad \text{Check: } A^+ = (\ldots)$$

6. (15 points) Let $A \in R^{n\times n}$ be symmetric and positive definite and $n = j + k$. Partition A into the following 2 by 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}: $A_{11} = R_{11}^T R_{11}$, where R_{11} is upper triangular with positive main-diagonal entries. Let $R_{12} = \left(R_{11}^{-1}\right)^T A_{12}$ and let $\tilde{A}_{22} = A_{22} - R_{12}^T R_{12}$.

(a) Prove that A_{11} is positive definite.

(b) Prove that

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

HW #8

(c) Prove that $\tilde{A}_{22}$ is positive definite.

(a) A positive definite $\Leftrightarrow x^T A x > 0 \quad \forall x \in \mathbb{R}^n \setminus 0$

From that it follows that A_{11} is also positive definite because we can choose

$\tilde{x} \in \{(y, \vec{0})^T \mid y \in \mathbb{R}^j\} \subset \mathbb{R}^n$

and know that $y^T A_{11} y = \tilde{x}^T A \tilde{x} > 0$

$\square$

(b) Given: $A_{11} = R_{11}^T R_{11}$, $R_{12} = (R_{11}^{-1})^T A_{12}$

and $\tilde{A}_{22} = A_{22} - R_{12}^T R_{12}$

with that: $R_{12}^T = A_{12}^T R_{11}^{-1} = A_{21} R_{11}^{-1}$

$\Rightarrow R_{12}^T R_{12} = A_{21} \underbrace{R_{11}^{-1} (R_{11}^{-1})^T}_{(R_{11}^T R_{11})^{-1} = A_{11}^{-1}} A_{12} = A_{21} A_{11}^{-1} A_{12}$

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

$\square$

(c) An informal proof can be found in the textbook [p. 174]. However, I'll still write out a formal proof:

We can factorize A as follows:

$$A = \begin{pmatrix} I & 0 \\ A_{21}A_{11}^{-1} & I \end{pmatrix} \begin{pmatrix} A_{11} & 0 \\ 0 & \tilde{A}_{22} \end{pmatrix} \begin{pmatrix} I & A_{11}^{-1}A_{12} \\ 0 & I \end{pmatrix}$$

$$=: \quad U \quad \cdot \quad D \quad \cdot \quad U^T$$

Note that U is invertible:

$$U^{-1} = \begin{pmatrix} I & 0 \\ -A_{21}A_{11}^{-1} & I \end{pmatrix}.$$

So D can be written as:

$$D = U^{-1} A (U^{-1})^T.$$

Therefore: $\tilde{x}^T D \tilde{x} = \tilde{x}^T U^{-1} A (U^{-1})^T \tilde{x} = x^T A x > 0$

with $x = (U^{-1})^T \tilde{x}$, $\tilde{x} \neq 0$

So D is a similarity transformation of A and D is positive-definite if and only if A is positive definite, which we know it is. Since $D = \begin{pmatrix} A_{11} & 0 \\ 0 & \tilde{A}_{22} \end{pmatrix}$ and A_{11} is p.d., $\tilde{A}_{22}$ is p.d. $\square$

7. (10 points) Show that if $A \in \mathbf{R}^{m\times m}$ is symmetric and positive definite, then solving the linear system $A\mathbf{x} = \mathbf{b}$ amounts to computing

$$\mathbf{x} = \sum_{i=1}^{m} \frac{c_i}{\lambda_i}\mathbf{v}_i,$$

where λ_i are the eigenvalues of A and $\mathbf{v}_i$ are the corresponding eigenvectors, and c_i are some constants determined by $\mathbf{b}$ and $\mathbf{v}_i$.

Again, the proof is implicit in [Lecture 23].

The eigenvalues of A are all positive real numbers. If $Ax = \lambda x$ for $x \neq 0$, we have

$x^T A x = \lambda x^T x > 0$ and therefore $\lambda > 0$.

And ev that corres to distinct ew's are orthogonal [p. 173]. (hermitian matrices are normal).

So A is unitary diagonalizable.

$$A = Q D Q^T, \quad QQ^T = I$$

The colums of Q are the orthogonal ew of A V_i and the diagonal of D are the ev's λ_i.

Therefore $Ax = b \iff QDQ^T x = b$

$$\iff x = D^{-1}(Q^T b) Q = D^{-1}\vec{c}\, Q$$

$$= \sum_{i=1}^{m} \frac{c_i}{\lambda_i} V_i, \quad c_i = V_i^T b$$

$\square$

8. (15 points) Let $A \in \mathbf{C}^{m\times n}, m \geq n$, with linearly independent columns:

$$A = [\mathbf{a}_1, \cdots, \mathbf{a}_n].$$

Find eigenvalues and eigenvectors of the projection matrix

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

For $m=n$, $P=0$ and there is nothing to show.

SVD: $A = U\Sigma V^*$

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

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

So AA^+ is the orthogonal projector onto Range(A).

Therefore, $P = I - AA^+$ is the orthogonal projector onto the nullspace of A.

Thus, P has eigenvalues 1 for any eV $v \in$ Null(A) and ew 0 for any $w \in$ Range(A) = span$\{a_1, \cdots, a_n\}$.

So $a_1, \ldots, a_n$ are eigenvectors of P with ev 0. And any $x \in$ Null$(A) \subset \mathbb{R}^m$ is an ev with ew 1.

Proof: $PA = \left(I - A(A^*A)^{-1}A^*\right)A$

$= A - AV\Sigma^{-1}U^*U\Sigma V^* = A - AV\Sigma^{-1}\Sigma V^*$

$= 0$