# Numerical Linear Algebra

## Sheet 1 — MT24

## Norms and SVD, up to lecture 4 (solutions for Sections A,C)

Questions are split into three sections: Section A (basic, not marked, solutions provided): 1–3. Section B (will be marked): 4–8. Section C (new, solutions provided): 9.

1. Show that $\|x\|_\infty = \max_i |x_i|$ satisfies the axioms for a vector norm.

Solution: Verify the three properties required satisfied by a vector norm:
1) positivity: $\max_i |x_i| \ge 0, \max_i |x_i| = 0 \iff x = 0$ by properties of $|\cdot|$;
2) scaling: $|\alpha x_i| = |\alpha||x_i| \implies \max_i |\alpha x_i| = |\alpha| \max_i |x_i|$;
3) triangle inequality: $|x_i + y_i| \le |x_i| + |y_i| \implies \max_i |x_i + y_i| \le \max_i |x_i| + \max_i |y_i|$.

Thus $\|\cdot\|_\infty$ is a vector norm.

2. Show that if $\|x\|$ is a vector norm then $\sup_x \frac{\|Ax\|}{\|x\|}$ satisfies the axioms for a matrix norm. Further show that
$$\|A\,B\| \le \|A\|\,\|B\|.$$

(Solution:) Suppose $\|\cdot\|$ is some vector norm. Verify the three requirements for matrix norms:
1) positivity: $\|x\| \ge 0\,\forall x \implies \|Ax\| \ge 0\,\forall x$, so $\|Ax\|/\|x\| \ge 0$ for $x \ne 0$;
If $\|A\| = 0$ then $\|Ax\| = 0\,\forall x \implies Ax = 0\,\forall x \implies A = 0$. Clearly $\|0\| = 0$.
2) scaling: $\|\alpha x\| = |\alpha|\,\|x\| \implies \|(\alpha A)x\| = |\alpha|\,\|Ax\|$, so $\|(\alpha A)x\|/\|x\| = |\alpha|\,\|Ax\|/\|x\|\,\forall x$;
3) triangle inequality: $\|x + y\| \le \|x\| + \|y\| \implies \|A(x + y)\| \le \|Ax\| + \|Ay\|$,
so $\sup_x \|(A + B)x\|/\|x\| \le \sup_x \|Ax\|/\|x\| + \sup_x \|Bx\|/\|x\|$.

Now, to show $\|AB\| \le \|A\|\,\|B\|$. The result is trivial if $B = 0$, thus

$$\|AB\| = \sup_{x\ne 0} \frac{\|ABx\|}{\|x\|} = \sup_{Bx\ne 0} \frac{\|ABx\|}{\|x\|} = \sup_{Bx\ne 0} \frac{\|ABx\|}{\|Bx\|}\frac{\|Bx\|}{\|x\|}.$$


Yuji Nakatsukasa: nakatsukasa@maths.ox.ac.uk

3. By considering the individual columns $a_j$ of $A$ and $b_j$ of $B = Q\,A$, show that

$$\|Q\,A\|_{\mathrm{F}} = \|A\|_{\mathrm{F}}$$

if $Q$ is an orthogonal matrix.

(Solution:) Suppose $A \in \mathbb{R}^{m\times n}$ and let $Q \in \mathbb{R}^{m\times m}$ be orthogonal. Recall that $\|Qx\|_2 = \|x\|_2$ since $\|Qx\|_2^2 = (Qx)^{\mathrm{T}}(Qx) = x^{\mathrm{T}}Q^{\mathrm{T}}Qx = x^{\mathrm{T}}x = \|x\|_2^2$.
Partition $A \in \mathbb{R}^{m\times n}$ by columns, $A = [a_1\ a_2\ \cdots\ a_n]$, where $a_j \in \mathbb{R}^m$. Then we can write the Frobenius norm of $A$ using dot products:

$$\|A\|_{\mathrm{F}}^2 = \sum_{i=1}^{m}\sum_{j=1}^{n}|\alpha_{ij}|^2 = \sum_{j=1}^{n} a_j^{\mathrm{T}} a_j.$$

Now write $QA = Q[a_1\ a_2\ \cdots\ a_n] = [Qa_1\ Qa_2\ \cdots\ Qa_n]$, and compute

$$\|QA\|_{\mathrm{F}}^2 = \sum_{j=1}^{n}(Qa_j)^{\mathrm{T}}(Qa_j) = \sum_{j=1}^{n} a_j^{\mathrm{T}}Q^{\mathrm{T}}Qa_j = \sum_{j=1}^{n} a_j^{\mathrm{T}} a_j = \|A\|_{\mathrm{F}}^2.$$

(tutors: perhaps discuss what happens when $Q$ is (i) tall-orthonormal (same identity holds), and (ii) fat-orthonormal rows (identity fails). )

4. From the definition of the vector 1–norm show that

$$\|A\|_1 = \max_j \sum_i |a_{ij}|.$$

5. Full SVD. Prove the existence of $A = U \begin{bmatrix} \Sigma \\ 0_{(m-n)\times n} \end{bmatrix} V^*$, where $U \in \mathbb{C}^{m\times m}$ and $V \in \mathbb{C}^{n\times n}$ are unitary matrices i.e., $U^*U = I_m$ and $V^*V = I_n$, and $\Sigma \in \mathbb{R}^{n\times n}$ is diagonal.

6. What is the SVD of a normal matrix $A$, with respect to the eigenvalues and eigenvectors? What if $A$ is (real) symmetric? And unitary?

7. If $A \in \mathbb{R}^{n\times n}$ is nonsingular, what is the SVD of $A^{-1}$ in terms of that of $A$?

8. Let $B$ be a square $n \times n$ matrix. Bound the $i$th singular values of $AB$ using $\sigma_i(A)$ and $\sigma_i(B)$: Specifically, prove that for each $i$,

$$\sigma_i(A)\sigma_n(B) \le \sigma_i(AB) \le \sigma_i(A)\sigma_1(B).$$


Yuji Nakatsukasa: nakatsukasa@maths.ox.ac.uk

9. (optional; harder) Let $A \in \mathbb{R}^{m\times n}, m \ge n$ and $\sigma_1(A) \ge \sigma_2(A) \ge \cdots \ge \sigma_n(A) \ge 0$ be its singular values. Prove that for $k = 1, 2, \ldots, n$,

$$\sum_{i=1}^{k} \sigma_i(A) = \max_{Q^TQ=I_k, W^TW=I_k} \operatorname{trace}(Q^TAW).$$

   ($Q \in \mathbb{R}^{m\times k}, W \in \mathbb{R}^{n\times k}$ are orthonormal. Recall for an $k \times k$ matrix $B$, $\operatorname{trace}(B) = \sum_{i=1}^{k} B_{ii}$; a useful property is $\operatorname{trace}(CD) = \operatorname{trace}(DC)$ as long as $CD$ is square.)

(Solution:) Equality is seen to be attained when $Q = [u_1, u_2, \ldots, u_k], W = [v_1, \ldots, v_k]$, since then $\operatorname{trace}(Q^TAW) = \operatorname{trace}(\operatorname{diag}(\sigma_1(A), \ldots, \sigma_k(A))) = \sum_{i=1}^{k} \sigma_i(A)$. We need to prove this is an upper bound for $\operatorname{trace}(Q^TAW)$. First note that $\sigma_i(AB) \le \sigma_i(A)\|B\|$ holds for any $A, B$ s.t. $AB$ is defined (e.g. via Courant-Fisher). Now since $Q, W$ are orthonormal, $\sigma_i(Q) = \sigma_i(W) = 1$ for all $i = 1, \ldots, k$. We thus have $\sigma_i(Q^TAW) \le \sigma_i(A)$ for all $i$. We are thus done if we prove $\operatorname{trace}(B) \le \sum_{i=1}^{k} \sigma_i(B)$ for any $k \times k$ matrix $B$. Let $B = U_B\Sigma_BV_B^T$ be the SVD. Then $\operatorname{trace}(B) = \operatorname{trace}(U_B\Sigma_BV_B^T) = \operatorname{trace}(\Sigma_BV_B^TU_B) = \sum_{i=1}^{k} \sigma_i(B)(V_B^TU_B)_{ii} \le \sum_{i=1}^{k} \sigma_i$, because $V_B^TU_B$ is orthogonal and so its entries are bounded by 1 in absolute value.


Yuji Nakatsukasa: nakatsukasa@maths.ox.ac.uk