Lecture 18 (hrs. 34,35) - November 5, 2025, 11:30 - 13:30 A13

(2.25) <u>Definition</u> (norm in a vector space).

Let V be a vector space over R. A function $N:V \to R$ is a *norm* in V if it satisfies the following conditions:

(1) for every $v \in V$ it is: $N(v) \geqslant 0$ and $N(v) = 0 \Leftrightarrow v = 0$;
(2) for every $v \in V$ and every $\alpha \in R$ it is: $N(\alpha v) = |\alpha| N(v)$;
(3) for every $v,w \in V$ it is: $N(v + w) \leqslant N(v) + N(w)$.

The pair V,N is called *normed space*.

(2.26) <u>Example</u> (usual norms in $R^n$).

Let $V = R^n$ and $v = [v_1,...,v_n] \in V$. The functions:

- $N_1:R^n \to R$ defined by $N_1(v) = |v_1| + ... + |v_n|$
- $N_2:R^n \to R$ defined by $N_2(v) = \text{sqrt}( v_1^2 + ... + v_n^2 )$
- $N_\infty:R^n \to R$ defined by $N_\infty(v) = \max\{ |v_1|,...,|v_n| \}$

are norms in $R^n$.

(2.27) <u>Homework</u>.

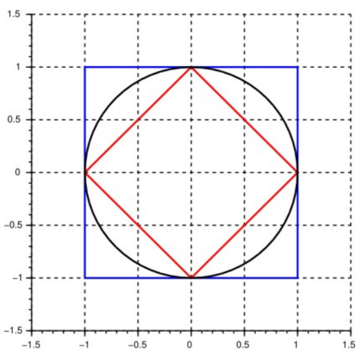
Prove that the functions $N_1$ and $N_\infty$ satisfy the properties of Definition (2.25).

(2.28) <u>Definition</u> (ball in $R^n$).

Let $R^n$,N be a normed space, $v \in R^n$ and $r \in R$. The set:

$$I_N(v,r) = \{ x \in R^n \text{ such that } N(x - v) \leqslant r \}$$

is called *ball of center* v *and radius* r. In the following figure the ball $I_2(0,1)$ is represented in black, $I_\infty(0,1)$ in blue, and $I_1(0,1)$ in red, in the case n = 2.



(2.29) <u>Definition</u> (norm of a matrix).

Let $R^n, N$ be a normed space and $A \in R^{n \times n}$. The quantity:

$$\| A \|_N = \max\{ N(A v),\ N(v) = 1 \}$$

is called *norm of* A *induced by* N.

(2.30) <u>Properties</u> (of the norm of a matrix).

(I) Note that the norm of A induced by N is well-defined: the subset S of vectors v of $R^n$ defined by N(v) = 1 is *closed and bounded*, and the function $v \to N(A v)$ is *continuous*. By the Weierstrass Theorem, the latter *attains a maximum and a minimum* on S. In particular:

$$\text{there exists } y \in R^n \text{ such that } N(y) = 1 \text{ and } \| A \|_N = N(A y)$$

(IIa) For every $A \in R^{n \times n}$ and every $v \in R^n$ it is:

$$N(A v) \leqslant \| A \|_N \, N(v)$$

Indeed: The relation is certainly true if v = 0. If $v \neq 0$ we have:

$$N(A v) = N( A\, N(v)\, \mathrm{unit}(v) )^1 = N( N(v)\, A\, \mathrm{unit}(v) ) = N(v)\, N(A\, \mathrm{unit}(v))$$

Furthermore, by the Definition of norm of A induced by N it is: $N(A\, \mathrm{unit}(v)) \leqslant \| A \|_N$, hence:

$$N(A v) \leqslant \| A \|_N \, N(v)$$

(IIb) There exists $w \in R^n$ such that:

$$N(A w) = \| A \|_N \, N(w)$$

By Properties (I), there exists $y \in R^n$ such that N(y) = 1 and $\| A \|_N = N(A y)$. The statement follows by unit(w) = y.

(III) For every $A, B \in R^{n \times n}$ it is:

$$\| A B \|_N \leqslant \| A \|_N \| B \|_N$$

---

1 Let $R^n, N$ be a normed space and $v \in R^n$ be a non-zero vector. Then:

$$\mathrm{unit}(v) = \frac{1}{N(v)} \, v$$

is the *unit vector* of v. Obviously it is N(unit(v)) = 1.

Indeed: by Properties (I) there exists $y \in R^n$ such that $N(y) = 1$ and $\| A B \|_N = N(A B y)$. Then, by Properties (II):

$$\| A B \|_N = N(A B y) \leqslant \| A \|_N N(B y) \leqslant \| A \|_N \| B \|_N N(y) = \| A \|_N \| B \|_N$$

(2.31) <u>Remark</u>.

The set $R^{n \times n}$ is, with the usual matrix addition and multiple operations, a vector space over R. Introducing a norm N in $R^n$, the function $A \to \| A \|_N$ from $R^{n \times n}$ to R is *a norm* in $R^{n \times n}$ (this explains the name given to the function). Therefore, the Properties of the norm (Definition (2.25)) hold:

(1) for every $A \in R^{n \times n}$ it is: $\| A \|_N \geqslant 0$, and $\| A \|_N = 0 \Leftrightarrow A = 0$;

(2) for every $A \in R^{n \times n}$ and every $\alpha \in R$ it is: $\| \alpha A \|_N = |\alpha| \, \| A \|_N$;

(3) for every $A,B \in R^{n \times n}$ it is: $\| A + B \|_N \leqslant \| A \|_N + \| B \|_N$.

(2.32) <u>Remark</u> (how to compute the norm of a matrix).

Let $A \in R^{n \times n}$ and let $a_1,...,a_n$ be the columns of A. it is:

- $\| A \|_1 = \max\{ N_1(a_1),...,N_1(a_n) \}$
- $\| A \|_2 = \text{sqrt}($ maximum of the eigenvalues of $A^tA$ $)$[2]
- $\| A \|_\infty = \| A^t \|_1$ i.e., denoted by $r_1,...,r_n$ the *rows* of A: $\| A \|_\infty = \max\{ N_1(r_1),...,N_1(r_n) \}$

Note that while the computation of $\| A \|_1$ and $\| A \|_\infty$ is elementary, that of $\| A \|_2$ in general *is not*.

(2.33) <u>Example</u> (conditioning: the case $\delta A = 0$, $\delta b \neq 0$).

Let us return to the conditioning of the solution of the system $A x = b$. Let N be a norm in $R^n$.

Let $\delta A = 0$ and $\delta b \neq 0$. Then the columns $x^*$ and $\hat{x}$ satisfy:

$$A x^* = b \quad , \quad A \hat{x} = b + \delta b$$

hence, by the invertibility of A, it is:

$$\delta x = \hat{x} - x^* = A^{-1} (b + \delta b) - A^{-1} b = A^{-1} \delta b$$

By introducing the *absolute measure* of the deviation $N(\delta x)$ and that of the perturbation $N(\delta b)$, using Property (IIa) we obtain:

$$\forall \; \delta b \; , \; N(\delta x) = N(A^{-1} \delta b) \leqslant \| A^{-1} \|_N N(\delta b)$$

2 The matrix $A^tA$ is *symmetric* and *positive semidefinite,* hence all *i*ts eigenvalues are nonnegative.

The above is the *best possible bound* for the absolute measure of the deviation as a function of the absolute measure of the perturbation. Indeed, Property (IIb) shows that:

$$\exists\ \delta b : N(\delta x) = \| A^{-1} \|_N N(\delta b)$$

If $b \neq 0$ (and therefore $x^* \neq 0$), we can also introduce the *relative measures* of the deviation $\varepsilon_x = N(\delta x)/N(x^*)$ and of the perturbation $\varepsilon_b = N(\delta b)/N(b)$. For these measures we have:

$$\varepsilon_x = \frac{N(\delta x)}{N(x^*)} \leqslant \frac{\| A^{-1} \|_N N(\delta b)}{N(x^*)}$$

<u>But</u>:

$$A x^* = b \Rightarrow N(b) = N(A x^*) \leqslant \| A^{-1} \|_N N(x^*) \Rightarrow \frac{1}{N(x^*)} \leqslant \frac{\| A^{-1} \|_N}{N(b)}$$

hence:

$$\forall\ \delta b\ ,\ \forall\ b \neq 0 : \varepsilon_x \leqslant \| A^{-1} \|_N \| A \|_N\ \varepsilon_b$$

The above is the *best possible bound* for the relative size of the deviation as a function of the relative size of the perturbation. Indeed, Property (IIb) shows that:

$$\exists\ \delta b \text{ and } \exists\ b \neq 0 : \varepsilon_x = \| A^{-1} \|_N \| A \|_N\ \varepsilon_b$$

(2.34) <u>Definition</u> (condition number of a matrix).

Let $A \in R^{n \times n}$ be an *invertible* matrix and N be a norm in $R^n$. The number:

$$c_N(A) = \| A^{-1} \|_N \| A \|_N$$

is the *condition number of* A (using the norm N).

(2.35) <u>Remark</u>.

Since $A^{-1} A = I$, we have (using Property (III) of (2.30)):

$$\| I \|_N = \| A^{-1} A \|_N \leqslant \| A^{-1} \|_N \| A \|_N$$

Moreover, by the definition of norm of I:

$$\| I \|_N = \max\{ N(I v), N(v) = 1 \} = \max\{ N(v), N(v) = 1 \} = 1$$

hence:

$$c_N(A) = \| A^{-1} \|_N \| A \|_N \geqslant 1$$

(2.36) <u>Theorem</u> (conditioning).

Let $A \in R^{n \times n}$ be an *invertible* matrix and let N be a norm in $R^n$. Then: for every $b \neq 0$, every $\delta b$ such that $b + \delta b \neq 0$ and every $\delta A$ such that $c_N(A)\ \varepsilon_A < 1$ it is:

$$\varepsilon_x \leqslant \frac{c_N(A)}{1 - c_N(A)\,\varepsilon_A} (\varepsilon_A + \varepsilon_b)$$