11 Norms of vectors and matrices. Condition number

11.1 Norms

Let x be a real number. I define its absolute value as

$$|x| = \begin{cases} x, & \text{if } x \geq 0, \\ -x, & \text{otherwise.} \end{cases}$$

Absolute value has several nice properties. In particular I have that $|x| > 0$ for all non-zero x, and is equal to zero only if $x = 0$; for any other real number α I have that $|\alpha x| = |\alpha| \cdot |x|$; and finally, for any $x, y \in \mathbf{R}$ I have the so-called *triangle inequality*

$$|x + y| \leq |x| + |y|. \tag{11.1}$$

Exercise 1. Give a proof for the triangle inequality.

The name *triangle inequality* refer to the fact that absolute value allows to measure the distance d between any two real numbers, vis., $d(x, y) = |x - y|$, and the distance certainly satisfies

$$|x - y| = |x - z + z - y| \leq |x - z| + |y - z|,$$

i.e., distance from x to y is no bigger than the distance from x to some other number z plus that distance from z to y, geometrically, the length of one of the sides of a triangle cannot be bigger then the sum of the lengthes of two other sides. Hence the name.

To reiterate, the absolute value allows me to measure the distance between any two points, and also measure the size of a given real number x, which is the distance from zero. It would be very rewarding to be able to do something similar with the elements of $\mathbf{R}^k$, which is k-dimensional vector space, where the elements are column-vectors with k coordinates:

$$\mathbf{R}^k \ni \boldsymbol{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_k \end{bmatrix}, \quad x_j \in \mathbf{R}.$$

This is done by using *norm*.

Definition 11.1. *A norm on the vector space $\mathbf{R}^k$ is a function*

$$\|\cdot\| \colon \mathbf{R}^k \longrightarrow \mathbf{R},$$

such that

1. $\|\boldsymbol{x}\| \geq 0$*, for any* $\boldsymbol{x} \in \mathbf{R}^k$*, and* $\|\boldsymbol{x}\| = 0$ *if and only if* $\boldsymbol{x} = 0$ *(note that here* 0 *is not a number zero but a zero vector with* k *zero coordinates);*
2. *if* $\alpha \in \mathbf{R}$ *then* $\|\alpha \boldsymbol{x}\| = |\alpha| \cdot \|\boldsymbol{x}\|$ *for any* $\boldsymbol{x} \in \mathbf{R}^k$*;*

Math 488/688: Numerical Analysis
e-mail: artem.novozhilov@ndsu.edu. Fall 2021

3. $\|\boldsymbol{x}+\boldsymbol{y}\| \leq \|\boldsymbol{x}\|+\|\boldsymbol{y}\|$ *for any* $\boldsymbol{x}, \boldsymbol{y} \in \mathbf{R}^k$ *(triangle inequality).*

Example 11.2. Let

$$\|\boldsymbol{x}\|_1=\sum_{i=1}^k |x_i|.$$

Let me check all the norm axioms. First, clearly, $\|\boldsymbol{x}\|_1 \geq 0$ as the sum of nonnegative terms; if $\boldsymbol{x}=0$ then $\|\boldsymbol{x}\|_1=0$, also, if $\|\boldsymbol{x}\|_1=0$ then each coordinate x_i must be zero, and hence $\boldsymbol{x}=0$. To check the second axiom I compute

$$\|\alpha \boldsymbol{x}\|_1=\sum_{i=1}^k |\alpha x_i|=|\alpha| \sum_{i=1}^k |x_i|=|\alpha| \cdot \|\boldsymbol{x}\|_1$$

as required.

For the triangle inequality

$$\|\boldsymbol{x}+\boldsymbol{y}\|_1=\sum_{i=1}^k |x_i+y_i| \leq \sum_{i=1}^k \big(|x_i|+|y_i|\big)=\sum_{i=1}^k |x_i|+\sum_{i=1}^k |y_i|=\|\boldsymbol{x}\|_1+\|\boldsymbol{y}\|_1.$$

So I have my first legitimate example of a norm.

Example 11.3.

$$\|\boldsymbol{x}\|_2=\left(\sum_{i=1}^k x_i^2\right)^{\frac{1}{2}},$$

which is usually called the *Euclidian norm.*

Example 11.4.

$$\|\boldsymbol{x}\|_\infty=\max_i |x_i|.$$

I invite the student to check the norm axioms for this example.

In short, it is possible to come up with different specific definitions of vector norms.

Example 11.5 (Ball of radius 1 in $\mathbf{R}^2$)**.**

Since we live in $\mathbf{R}^k$ then matrices $\boldsymbol{A}=[a_{ij}]_{k\times k}$ for us are not elements of $\mathbf{R}^k$ (they are elements of $\mathbf{R}^{k^2}$) but maps acting on $\mathbf{R}^k$, i.e.,

$$\boldsymbol{A}\colon \mathbf{R}^k \longrightarrow \mathbf{R}^k,$$

where it simply means that for any vector $\boldsymbol{x} \in \mathbf{R}^k$ I can compute $\boldsymbol{A}\boldsymbol{x} \in \mathbf{R}^k$, which in coordinates a matrix times vector multiplication. My goal is to define "a size" or a norm of a matrix as well.

Of course, I could have generalized the given above examples to $\mathbf{R}^{k^2}$ spaces, but for my purposes I need something that is directly related to the implemented norm of a vector. This is an example of a *subordinate* norm of a matrix, you can read more in any linear algebra book, for my purposes I just will need definitions.

Definition 11.6. *The subordinate infinity norm of matrix $\boldsymbol{A}$ is*

$$\|\boldsymbol{A}\|_\infty = \max_i \sum_{j=1}^{k} |a_{ij}| \quad \textit{(largest row sum)}.$$

The subordinate 1-norm of matrix $\boldsymbol{A}$ is

$$\|\boldsymbol{A}\|_1 = \max_j \sum_{i=1}^{k} |a_{ij}| \quad \textit{(largest column sum)}.$$

Note that $\|\boldsymbol{A}\|_1 = \|\boldsymbol{A}^\top\|$. Similarly, for any other vector norm one can define a corresponding subordinate norm. For the subordinate norms all the usual axioms of norms hold (you can check them for the given two examples). Additionally, for any norm $\|\cdot\|$ I have that

$$\|\boldsymbol{AB}\| \leq \|\boldsymbol{A}\| \cdot \|\boldsymbol{B}\|, \quad \|\boldsymbol{Ax}\| \leq \|\boldsymbol{A}\| \cdot \|\boldsymbol{x}\|.$$

Here and below if I do not specify which norm I use, it means it is true for any vector norm and its subordinate counterpart.

11.2 Condition number of a matrix

Let us go back to the system of linear algebraic equations

$$\boldsymbol{Ax} = \boldsymbol{b}.$$

Here $\boldsymbol{A}$ and $\boldsymbol{b}$ are input data of our problem, and $\boldsymbol{x}$ is the output, our solution. In reality all the data is known only approximately, up the measurement error. Therefore it is natural to ask what kind of error we may expect in our solution if we can estimate the error in the input data.

To be precise, let me assume that instead of the exact vector $\boldsymbol{b}$ I know only its approximate measurement $\boldsymbol{b} + \Delta\boldsymbol{b}$. Here $\|\Delta\boldsymbol{b}\|$ is the *absolute error*, and

$$\frac{\|\Delta\boldsymbol{b}\|}{\|\boldsymbol{b}\|}$$

is the *relative error*. Using $\boldsymbol{b} + \Delta\boldsymbol{b}$ instead of $\boldsymbol{b}$ will cause me to compute something like $\boldsymbol{x} + \Delta\boldsymbol{x}$, where $\Delta\boldsymbol{x}$ is the absolute error in my solution.

Consider the following line of equalities and inequalities, where I use the properties of norms and the fact that $\boldsymbol{Ax} = \boldsymbol{b}$.

$$\begin{aligned}
\boldsymbol{A}(\boldsymbol{x} + \Delta\boldsymbol{x}) &= \boldsymbol{b} + \Delta\boldsymbol{b} \implies \\
\boldsymbol{Ax} + \boldsymbol{A}\Delta\boldsymbol{x} &= \boldsymbol{b} + \Delta\boldsymbol{b} \implies \\
\boldsymbol{A}\Delta\boldsymbol{x} &= \Delta\boldsymbol{b} \implies \\
\boldsymbol{\Delta} x &= \boldsymbol{A}^{-1}\Delta\boldsymbol{b} \implies \\
\|\Delta\boldsymbol{x}\| &= \|\boldsymbol{A}^{-1}\Delta\boldsymbol{b}\| \leq \|\boldsymbol{A}^{-1}\| \cdot \|\Delta\boldsymbol{b}\| = \\
&= \|\boldsymbol{A}^{-1}\| \cdot \|\Delta\boldsymbol{b}\| \frac{\|\boldsymbol{b}\|}{\|\boldsymbol{b}\|} = \\
&= \|\boldsymbol{A}^{-1}\| \cdot \|\boldsymbol{Ax}\| \frac{\|\Delta\boldsymbol{b}\|}{\|\boldsymbol{b}\|} \leq \\
&\leq \|\boldsymbol{A}^{-1}\| \cdot \|\boldsymbol{A}\| \cdot \|\boldsymbol{x}\| \frac{\|\Delta\boldsymbol{b}\|}{\|\boldsymbol{b}\|} \implies \\
\frac{\|\boldsymbol{\Delta} x\|}{\|\boldsymbol{x}\|} &\leq \|\boldsymbol{A}^{-1}\| \cdot \|\boldsymbol{A}\| \cdot \frac{\|\Delta\boldsymbol{b}\|}{\|\boldsymbol{b}\|} .
\end{aligned}$$

In word, the relative error in my solution is bounded by the relative error in vector $\boldsymbol{b}$ times $\|\boldsymbol{A}\| \cdot \|\boldsymbol{A}^{-1}\|$.

Let me consider another situation, when I have measurement error in the matrix $\boldsymbol{A}$ itself, i.e., I know $\boldsymbol{A} + \Delta\boldsymbol{A}$ instead of the exact $\boldsymbol{A}$. I have

$$\begin{aligned}
(\boldsymbol{A} + \Delta\boldsymbol{A})(\boldsymbol{x} + \Delta\boldsymbol{x}) &= \boldsymbol{b} \implies \\
\boldsymbol{Ax} + \Delta\boldsymbol{A}(x + \Delta\boldsymbol{x}) + \boldsymbol{A}\Delta\boldsymbol{x} &= \boldsymbol{b} \implies \\
\boldsymbol{A}\Delta\boldsymbol{x} &= -\Delta\boldsymbol{A}(x + \Delta\boldsymbol{x}) \implies \\
\Delta\boldsymbol{x} &= -\boldsymbol{A}^{-1}\Delta\boldsymbol{A}(x + \Delta\boldsymbol{x}) \implies \\
\frac{\|\Delta\boldsymbol{x}\|}{\|\boldsymbol{x} + \Delta\boldsymbol{x}\|} &\leq \|\boldsymbol{A}\| \cdot \|\boldsymbol{A}^{-1}\| \cdot \frac{\|\Delta\boldsymbol{A}\|}{\|\boldsymbol{A}\|} ,
\end{aligned}$$

which leads to a similar conclusion: my relative error in the solution is bounded by the relative error in the data times the same expression $\|\boldsymbol{A}\| \cdot \|\boldsymbol{A}^{-1}\|$. This prompts me to introduce the following definition.

Definition 11.7. *For a nonsingular matrix* $\boldsymbol{A}$ *its condition number* $\kappa(\boldsymbol{A})$ *is defined as*

$$\kappa(\boldsymbol{A}) = \|\boldsymbol{A}\| \cdot \|\boldsymbol{A}^{-1}\|.$$

Note that

$$1 = \|\boldsymbol{I}\| = \|\boldsymbol{AA}^{-1}\| \leq \|\boldsymbol{A}\| \cdot \|\boldsymbol{A}^{-1}\|.$$

Now I can reformulate my conclusions from the two examples above as follows: For a matrix with large condition number it is possible to have a large relative error in the solution even if the relative error of the initial data is small. And this is exactly what happened with my example of a 2 by 2 system, where even the professional level solve from `numpy` module had a somewhat significant error. Because of this possibility, matrices for which $\kappa(\boldsymbol{A}) \gg 1$ are called *ill-conditioned*. For the ill-conditioned matrices one cannot expect that the exact methods like $\boldsymbol{LU}$-factorization will result in a close to the exact solution because of the round up errors and accumulations of mistakes. A good professional software for solving systems of linear equations should at least issue a warning if the user is dealing with an ill-conditioned matrix.