Vector norms – why are they needed?

When dealing e.g with the solution of a nonlinear equation $f(x)=0$, the error $e = x_{approx} - x_{exact}$ is a single number, thus the absolute value $|e|$ gives us a good idea of the "extent" of error.

When solving a system of linear equations $A\underline{x} = \underline{b}$, the exact solution $\underline{x}_{exact}$ as well as any approximation $\underline{x}_{approx}$ are vectors, and the error:

$$\underline{e} = \underline{x}_{approx} - \underline{x}_{exact}$$

is a vector, too. It is not as straightforward to assess the "magnitude" of such a vector-valued error.

e.g. Consider $\underline{e}_1, \underline{e}_2 \in \mathbb{R}^{1000}$, and

$$\underline{e}_1 = \begin{pmatrix} 0.1 \\ 0.1 \\ 0.1 \\ \vdots \\ \vdots \\ 0.1 \end{pmatrix} \qquad \underline{e}_2 = \begin{pmatrix} 100 \\ 0 \\ 0 \\ \vdots \\ \vdots \\ 0 \end{pmatrix}$$

Which one is worse? e_1 has a modest amount of error, distributed over all components. In $\underline{e}_2$, all but one component are exact, but one of them has a huge discrepancy.

Exactly how we quantify and assess the extent of error is application-dependent. Vector norms are alternative ways to measure this magnitude, and different norms would be appropriate for different tasks.

Def. A vector norm is a function from $\mathbb{R}^n \to \mathbb{R}^+$, which maps a vector $\underline{v}$ to the real number $\|\underline{v}\|$. This symbol must satisfy the properties:

(i) $\|\underline{x}\| \geq 0$ for all $\underline{x} \in \mathbb{R}^n$. Also $\|\underline{x}\| = 0$ iff $\underline{x} = 0$

(ii) $\|\alpha \underline{x}\| = |\alpha| \cdot \|\underline{x}\| \quad \forall\, \underline{x} \in \mathbb{R}^n,\ \alpha \in \mathbb{R}$

(iii) $\|x+y\| \leq \|x\| + \|y\| \quad \forall x, y \in \mathbb{R}^n$ (triangle inequality).

Some norms which can be proven to satisfy these properties, are: (Here $\underline{x} = (x_1, x_2, \dots, x_n)$)

1. The L_1-norm (or 1-norm)

$$\|\underline{x}\|_1 = \sum_{i=1}^{n} |x_i|$$

2. The L_2-norm (or 2-norm, or Euclidean norm)

$$\|\underline{x}\|_2 = \sqrt{\sum_{i=1}^{n} x_i^2}$$

3. The infinity norm (or max-norm)

$$\|\underline{x}\|_\infty = \max_{1 \le i \le n} |x_i|$$

4. (Less common) L_p-norm

$$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p}$$

It is relatively easy to show that these satisfy the defining properties of a norm. e.g. for $\| \ \|_1$:

- $\|\underline{x}\|_1 = \sum_{i=1}^{n} |x_i| \geq 0$

- if $\underline{x} = \underline{0}$, then $\|x\|_1 = 0$

  if $\|x\|_1 = 0 \Rightarrow \sum_{i=1}^{n} |x_i| = 0 \Rightarrow |x_i| = 0 \ \forall i \Rightarrow \underline{x} = 0$

- $\|\alpha \underline{x}\| = \sum_{i=1}^{n} |\alpha x_i| = |\alpha| \sum_{i=1}^{n} |x_i| = |\alpha| \, \|\underline{x}\|_1$

- $\|\underline{x}+\underline{y}\| = \sum_{i=1}^{n} |x_i + y_i| \leq \sum_{i=1}^{n} |x_i| + \sum_{i=1}^{n} |y_i| = \|x\|_1 + \|y\|_1$

Similar proofs can be given for $\|\cdot\|_\infty$ (just as easy), $\|\cdot\|_2$ (a bit more difficult) and $\|\cdot\|_p$ (rather complicated).

We can actually define norms for (square) matrices, as well. A matrix norm is a function

$\|\cdot\| : \mathbb{R}^{n\times n} \longrightarrow \mathbb{R}$ which satisfies:

(i) $\|M\| \geq 0 \ \forall M \in \mathbb{R}^{n\times n}$. $\|M\| = 0$ iff $M = 0$

(ii) $\|\alpha M\| = |\alpha| \, \|M\|$

(iii) $\|M + N\| \leq \|M\| + \|N\|$

(iv) $\|M \cdot N\| \leq \|M\| \cdot \|N\|$.

(Property (iv) is the one that has slightly different flavor than vector norms).

Although more types of matrix norms exist, one common category is that of matrix norms induced by vector norms.

Def. If $\|\cdot\|_\star$ is a valid vector norm, its induced matrix norm is defined as

$$\|M\|_\star = \max_{x\in\mathbb{R}^n,\ x\neq 0} \left\{ \frac{\|Mx\|_\star}{\|x\|_\star} \right\}$$

or equivalently:

$$\|M\|_* = \max_{\substack{x \in \mathbb{R}^n \\ \|x\|_* = 1}} \{\|Mx\|_*\}$$

Note, again, that not all valid matrix norms are induced by vector norms. One notable example is the very commonly used Frobenius norm:

$$\|M\|_F = \sqrt{\sum_{i,j=1}^{n} M_{ij}^2}$$

We can easily show though that induced norms satisfy properties (i) through (iv). (i)-(iii) are rather trivial, eg:

$$\|M+N\| = \max_{\underline{x} \neq 0} \frac{\|(M+N)\underline{x}\|}{\|\underline{x}\|} \leq \max_{\underline{x} \neq 0} \frac{\|M\underline{x}\| + \|N\underline{x}\|}{\|\underline{x}\|}$$

$$= \max_{\underline{x} \neq 0} \frac{\|Mx\|}{\|x\|} + \max_{\underline{x} \neq 0} \frac{\|Nx\|}{\|x\|} = \|M\| + \|N\|.$$

Property (iv) is slightly trickier to show.

First, a lemma:

<u>Lemma</u> If $\|\cdot\|$ is a matrix norm induced by a vector norm $\|\cdot\|$, then:

$$\|Ax\| \leq \|A\| \cdot \|x\|$$

<u>Proof</u>: Since $\|A\| = \max\limits_{x \neq 0} \dfrac{\|Ax\|}{\|x\|}$, we have that for an arbitrary $y \in \mathbb{R}^n$: $(y \neq 0)$

$$\|A\| = \max_{x \neq 0} \frac{\|Ax\|}{\|x\|} \geq \frac{\|Ay\|}{\|y\|} \Rightarrow$$

$\Rightarrow \|Ay\| \leq \|A\| \, \|y\|$. This holds for $y \neq 0$, but we can see it is also true for $y = 0$

Prop. (iv)

$$\|MN\| = \max_{\|x\|=1} \|MNx\| \leq \max_{\|x\|=1} \|M\| \, \|Nx\| =$$

$$= \|M\| \cdot \max_{\|x\|=1} \|Nx\| = \|M\| \cdot \|N\| \Rightarrow$$

$$\|MN\| \leq \|M\| \cdot \|N\|.$$

Although the definition of an induced norm allowed us to prove certain properties, it does not necessarily provide a convenient formula for evaluating the matrix norm.

Fortunately, such formulas do exist for the L_1 and L_∞ induced matrix norms. Given here without proof:

$$\|A\|_1 = \max_j \sum_{i=1}^{n} |A_{ij}| \quad \text{(max. absolute column sum)}$$

$$\|A\|_\infty = \max_i \sum_{j=1}^{n} |A_{ij}| \quad \text{(max. absolute row sum)}.$$

($\|\cdot\|_2$ is much more complicated!)

Where do these vector/matrix norms come handy?