II.5 Norms

In this lecture we discuss matrix and vector norms.

1. Vector norms: we discuss the standard p-norm for vectors in $\mathbb{R}^n$.
2. Matrix norms: we discuss how two vector norms can be used to induce a norm on matrices. These

satisfy an additional multiplicative inequality.

1. Vector norms

Recall the definition of a (vector-)norm:

Definition 1 (vector-norm) A norm $\|\cdot\|$ on a vector space V (e.g. $\mathbb{R}^n$ or $\mathbb{C}^n$) over a field $\mathbb{F}$ (e.g. $\mathbb{R}$ or $\mathbb{C}$)
is a function that satisfies the following, for $\mathbf{x}, \mathbf{y} \in V$ and $c \in \mathbb{F}$:

1. Triangle inequality: $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$
2. Homogeneity: $\|c\mathbf{x}\| = |c|\|\mathbf{x}\|$
3. Positive-definiteness: $\|\mathbf{x}\| = 0$ implies that $\mathbf{x} = 0$.

Consider the following example:

Definition 2 (p-norm) For $1 \leq p < \infty$ and $\mathbf{x} \in \mathbb{C}^n$, define the p-norm:

$$\|\mathbf{x}\|_p := \left(\sum_{k=1}^n |x_k|^p\right)^{1/p}$$

where x_k is the k-th entry of $\mathbf{x}$. For $p = \infty$ we define

$$\|\mathbf{x}\|_\infty := \max_k |x_k|$$

Theorem 1 (p-norm) $\|\cdot\|_p$ is a norm for $1 \leq p \leq \infty$.

Proof

We will only prove the case $p = 1, 2, \infty$ as general p is more involved.

Homogeneity and positive-definiteness are straightforward: e.g.,

$$\|c\mathbf{x}\|_p = (\sum_{k=1}^n |cx_k|^p)^{1/p} = (|c|^p \sum_{k=1}^n |x_k|^p)^{1/p} = |c|\|\mathbf{x}\|$$

and if $\|\mathbf{x}\|_p = 0$ then all $|x_k|^p$ are have to be zero.

For $p = 1, \infty$ the triangle inequality is also straightforward:

$$\|\mathbf{x}+\mathbf{y}\|_\infty = \max_k(|x_k + y_k|) \leq \max_k(|x_k| + |y_k|) \leq \|\mathbf{x}\|_\infty + \|\mathbf{y}\|_\infty$$

and

$$\|\mathbf{x}+\mathbf{y}\|_1 = \sum_{k=1}^n |x_k + y_k| \leq \sum_{k=1}^n (|x_k| + |y_k|) = \|\mathbf{x}\|_1 + \|\mathbf{y}\|_1$$

For $p = 2$ it can be proved using the Cauchy–Schwartz inequality:

$$|\mathbf{x}^\star \mathbf{y}| \leq \|\mathbf{x}\|_2 \|\mathbf{y}\|_2$$

That is, we have

$$\|\mathbf{x}+\mathbf{y}\|^2 = \|\mathbf{x}\|^2 + 2\mathbf{x}^\top \mathbf{y} + \|\mathbf{y}\|^2 \leq \|\mathbf{x}\|^2 + 2\|\mathbf{x}\|\|\mathbf{y}\| + \|\mathbf{y}\|^2 = (\|\mathbf{x}\| + \|\mathbf{y}\|)$$

∎

In Julia, one can use the inbuilt `norm` function to calculate norms:

```
norm([1,-2,3]) == norm([1,-2,3], 2) == sqrt(1^2 + 2^2 + 3^2);
norm([1,-2,3], 1) == sqrt(1 + 2 + 3);
norm([1,-2,3], Inf) == 3;
```

2. Matrix norms

Just like vectors, matrices have norms that measure their "length". The simplest example is the Fröbenius norm:

Definition 3 (Fröbenius norm) For $A \in \mathbb{C}^{m \times n}$ define

$$\|A\|_F := \sqrt{\sum_{k=1}^m \sum_{j=1}^n A_{kj}^2}$$

This is available as `norm` in Julia:

In [1]:
```
A = randn(5,3)
norm(A) == norm(vec(A))
```

Out[1]:
```
true
```

While this is the simplest norm, it is not the most useful. Instead, we will build a matrix norm from a vector norm:

Definition 4 (matrix-norm) Suppose $A \in \mathbb{C}^{m \times n}$ and consider two norms $\| \cdot \|_X$ on $\mathbb{C}^n$ and $\| \cdot \|_Y$ on $\mathbb{C}^n$. Define the *(induced) matrix norm* as:

$$\|A\|_{X\to Y} := \sup_{\mathbf{v}:\|\mathbf{v}\|_X=1} \|A\mathbf{v}\|_Y$$

Also define

$$\|A\|_X := \|A\|_{X\to X}$$

For the induced p-norm we use the notation $\|A\|_p$.

Note an equivalent definition of the induced norm:

$$\|A\|_{X\to Y} = \sup_{\mathbf{x}\in\mathbb{R}^n,\mathbf{x}\neq 0} \frac{\|A\mathbf{x}\|_Y}{\|\mathbf{x}\|_X}$$

This follows since we can scale $\mathbf{x}$ by its norm so that it has unit norm, that is, $\frac{\mathbf{x}}{\|\mathbf{x}\|_X}$ has unit norm.

Lemma 1 (matrix norms are norms) Induced matrix norms are norms, that is for $\|\cdot\| = \|\cdot\|_{X\to Y}$ we have:

1. Triangle inequality: $\|A+B\| \leq \|A\| + \|B\|$
2. Homogeneity: $\|cA\| = |c|\|A\|$
3. Positive-definiteness: $\|A\| = 0 \Rightarrow A = 0$

In addition, they satisfy the following additional properties:

1. $\|A\mathbf{x}\|_Y \leq \|A\|_{X\to Y}\|\mathbf{x}\|_X$
2. Multiplicative inequality: $\|AB\|_{X\to Z} \leq \|A\|_{Y\to Z}\|B\|_{X\to Y}$

Proof

First we show the *triangle inequality*:

$$\|A+B\| \leq \sup_{\mathbf{v}:\|\mathbf{v}\|_X=1} (\|A\mathbf{v}\|_Y + \|B\mathbf{v}\|_Y) \leq \|A\| + \|B\|.$$

Homogeneity is also immediate. Positive-definiteness follows from the fact that if $\|A\| = 0$ then $A\mathbf{x} = 0$ for all $\mathbf{x} \in \mathbb{R}^n$. The property $\|A\mathbf{x}\|_Y \leq \|A\|_{X\to Y}\|\mathbf{x}\|_X$ follows from the definition. Finally, the multiplicative inequality follows from

$$\|AB\| = \sup_{\mathbf{v}:\|\mathbf{v}\|_X=1} \|AB\mathbf{v}|_Z \leq \sup_{\mathbf{v}:\|\mathbf{v}\|_X=1} \|A\|_{Y\to Z}\|B\mathbf{v}| = \|A\|_{Y\to Z}\|B\|_{X\to Y}$$

∎

We have some simple examples of induced norms:

Example 1 (1-norm) We claim

$$\|A\|_1 = \max_j \|\mathbf{a}_j\|_1$$

that is, the maximum 1-norm of the columns. To see this use the triangle inequality to find for $\|\mathbf{x}\|_1 = 1$

$$\|A\mathbf{x}\|_1 \leq \sum_{j=1}^n |x_j| \|\mathbf{a}_j\|_1 \leq \max_j \|\mathbf{a}_j\| \sum_{j=1}^n |x_j| = \max_j \|\mathbf{a}_j\|_1.$$

But the bound is also attained since if j is the column that maximises the norms then

$$\|A\mathbf{e}_j\|_1 = \|\mathbf{a}_j\|_1 = \max_j \|\mathbf{a}_j\|_1.$$

In the problem sheet we see that

$$\|A\|_\infty = \max_k \|A[k,:]\|_1$$

that is, the maximum 1-norm of the rows.

Matrix norms are available via `opnorm`:

In [2]:
```
m,n = 5,3
A = randn(m,n)
opnorm(A,1) == maximum(norm(A[:,j],1) for j = 1:n)
opnorm(A,Inf) == maximum(norm(A[k,:],1) for k = 1:m)
opnorm(A) # the 2-norm
```

Out[2]:
```
2.4801255978419205
```

An example that does not have a simple formula is $\|A\|_2$, but we do have two simple cases:

Proposition 1 (diagonal/orthogonal 2-norms) If Λ is diagonal with entries λ_k then $\|\Lambda\|_2 = \max_k |\lambda_k|$. If Q is orthogonal then $\|Q\| = 1$.

In the next chapter we see how the 2-norm for a matrix can be defined in terms of the *Singular Value Decomposition*.