Lecture 5: Vector and matrix norms, Classical iterative methods

Jamie Haddock

Table of contents

1 Vector and Matrix Norms

Measuring magnitude and length is fundamental to vector and matrix analysis, and will be fundamental to our analysis of numerical methods. Recall that a **norm** $\|\cdot\|$ is a real-valued function defined over a vector space with the following properties for all vectors and scalars α:

- $\|\mathbf{x}\| \geq 0$
- $\|\mathbf{x}\| = 0 \iff \mathbf{x} = \mathbf{0}$
- $\|\alpha\mathbf{x}\| = |\alpha|\|\mathbf{x}\|$
- $\|\mathbf{x}+\mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$

1.1 Vector Norms

Perhaps the most commonly encountered vector norms on $\mathbb{R}^n$ are these three:

- ℓ_2 norm: $\|\mathbf{x}\|_2 = \sqrt{\sum_{i=1}^n x_i^2}$
- ℓ_∞ norm: $\|\mathbf{x}\|_\infty = \max_{i=1,\cdots,n} |x_i|$
- ℓ_1 norm: $\|\mathbf{x}\|_1 = \sum_{i=1}^n |x_i|$

```
using LinearAlgebra

x = [2, -3, 1, -1]
twonorm = norm(x)
```

```
3.872983346207417
```

```
infnorm = norm(x,Inf)
```

```
3.0
```

```
onenorm = norm(x,1)
```

```
7.0
```

We say that a sequence of vectors $\mathbf{x}_1, \mathbf{x}_2, \cdots$ **converges** to $\mathbf{x}$ if

$$\lim_{k\to\infty} \|\mathbf{x}_k - \mathbf{x}\| = 0.$$

This will be very important as we begin to study iterative methods!

Theorem: Norm equivalence

In a finite-dimensional space, convergence in any norm implies convergence in all norms.

1.2 Matrix norms

As you may recall from Linear Algebra, the space of real-valued matrices of a given size define a vector space. There are many norms for this vector space. One such norm is quite interesting since it has a nice interpretation.

This norm is the **Frobenius norm** and it is defined as

$$\|\mathbf{A}\|_F = \sqrt{\sum_{i,j} A_{ij}^2}.$$

This norm is what is by default computed by the `norm` function in Julia.

One of the most interesting aspects of this norm is that we can view it as the ℓ_2 norm on a *vectorization* of the matrix. If you imagine stacking columns of $\mathbf{A}$ to form a vector, then the ℓ_2 norm of this vector is equal to the Frobenius norm of the matrix.

Matrices are actually column-stacked when stored in memory in Julia – this is known as *column-major order*. `MATLAB` is also column-major, while `C` and `Python` are row-major.

However, note that this norm does not inherently involve the *action* of the matrix as an *operator*. There are other matrix norms which do, and these are sometimes more useful.

Definition: Induced (natural) matrix norms

Given a vector norm $\|\cdot\|$, the **induced** or **natural matrix norm** for any $m \times n$ matrix $\mathbf{A}$ is

$$\|\mathbf{A}\| = \max_{\|\mathbf{x}\|=1} \|\mathbf{A}\mathbf{x}\| = \max_{\mathbf{x}\neq 0} \frac{\|\mathbf{A}\mathbf{x}\|}{\|\mathbf{x}\|}.$$

The induced norm definition causes these norms to satisfy some useful inequalities:

Theorem: Norm inequalities

Let $\|\cdot\|$ designate a matrix norm and the vector norm that induced it. Then for all matrices and vectors of compatible sizes,

$$\|\mathbf{A}\mathbf{x}\| \leq \|\mathbf{A}\|\|\mathbf{x}\|.$$

For all matrices of compatible sizes,

$$\|\mathbf{A}\mathbf{B}\| \leq \|\mathbf{A}\|\|\mathbf{B}\|.$$

Exercise: Prove these inequalities!

Answer:

For the first, if $\mathbf{x} = \mathbf{0}$, then $\|\mathbf{A}\mathbf{x}\| = 0 = \|\mathbf{A}\|\|\mathbf{x}\|$. Now, we can deal only with the case $\mathbf{x} \neq 0$,

$$\frac{\|\mathbf{A}\mathbf{x}\|}{\|\mathbf{x}\|} \leq \max_{\mathbf{x}\neq\mathbf{0}} \frac{\|\mathbf{A}\mathbf{x}\|}{\|\mathbf{x}\|} = \|\mathbf{A}\|.$$

Answer:

For the second,

$$\|\mathbf{A}\mathbf{B}\| = \max_{\mathbf{x}\neq\mathbf{0}} \frac{\|\mathbf{A}\mathbf{B}\mathbf{x}\|}{\|\mathbf{x}\|} \leq \max_{\mathbf{x}\neq\mathbf{0}} \frac{\|\mathbf{A}\|\|\mathbf{B}\mathbf{x}\|}{\|\mathbf{x}\|} \leq \max_{\mathbf{x}\neq\mathbf{0}} \frac{\|\mathbf{A}\|\|\mathbf{B}\|\|\mathbf{x}\|}{\|\mathbf{x}\|} = \|\mathbf{A}\|\|\mathbf{B}\|.$$

The induced matrix ∞- and 1-norms can be equivalently defined in terms of the entries of the matrix.

Theorem: Matrix ∞- and 1-norms

$$\|\mathbf{A}\|_\infty = \max_{1\leq i\leq n} \sum_{j=1}^{n} |A_{ij}|$$

$$\|\mathbf{A}\|_1 = \max_{1\leq j\leq n} \sum_{i=1}^{n} |A_{ij}|$$

```
A = [2 0; 1 -1]
```

```
2×2 Matrix{Int64}:
 2   0
 1  -1
```

```
Fronorm = norm(A)
```

```
2.449489742783178
```

```
twonorm = opnorm(A)
```

```
2.2882456112707374
```

We can also see that the entry-wise definitions of the ∞- and 1-norms are equivalent to their *induced norm* definition.

```
onenorm = opnorm(A,1)
```

```
3.0
```

```
maximum( sum(abs.(A),dims=1) )
```

```
3
```

```
infnorm = opnorm(A,Inf)
```

```
2.0
```

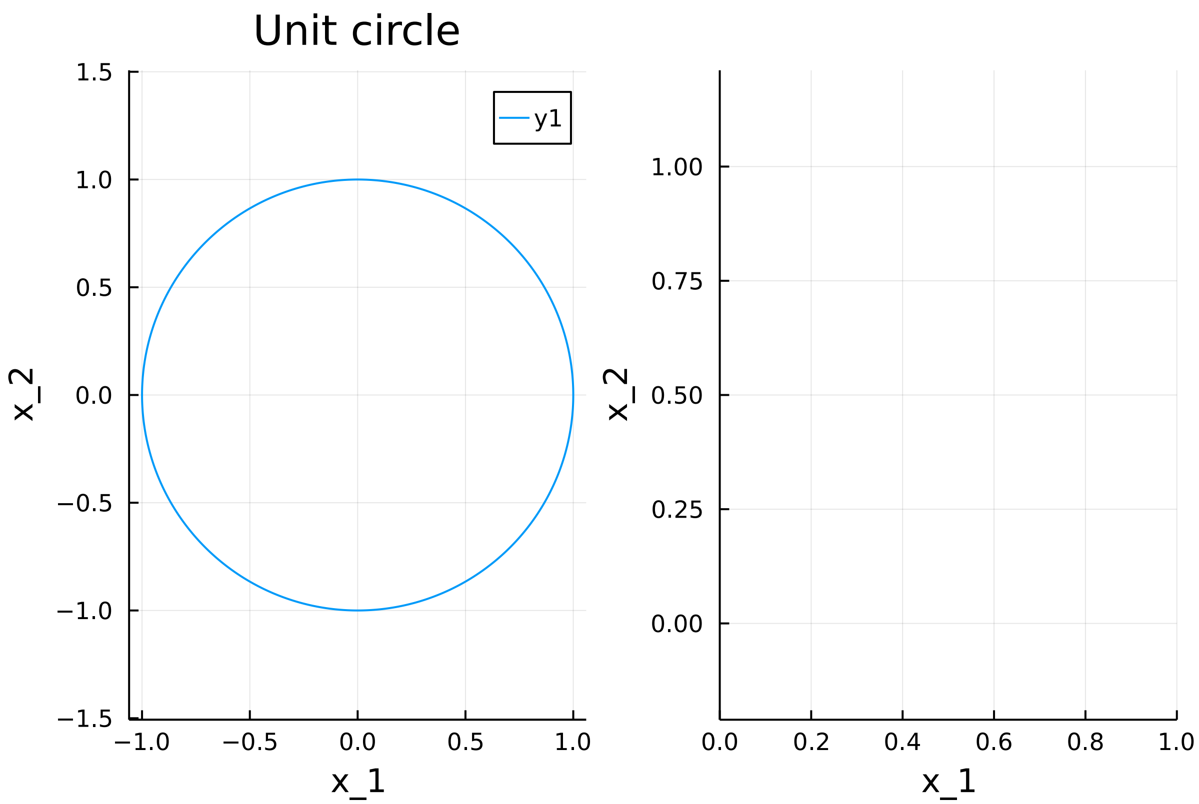
```
maximum( sum(abs.(A),dims=2) )
```

```
2
```

Now, we'll try to construct a geometric interpretation of the ℓ_2 norm.

```
# sample a lot of vectors on the unit circle in R^2
theta = 2pi*(0:1/600:1)
x = [ fun(t) for fun in [cos,sin], t in theta ]; #what a cool comprehension!
```

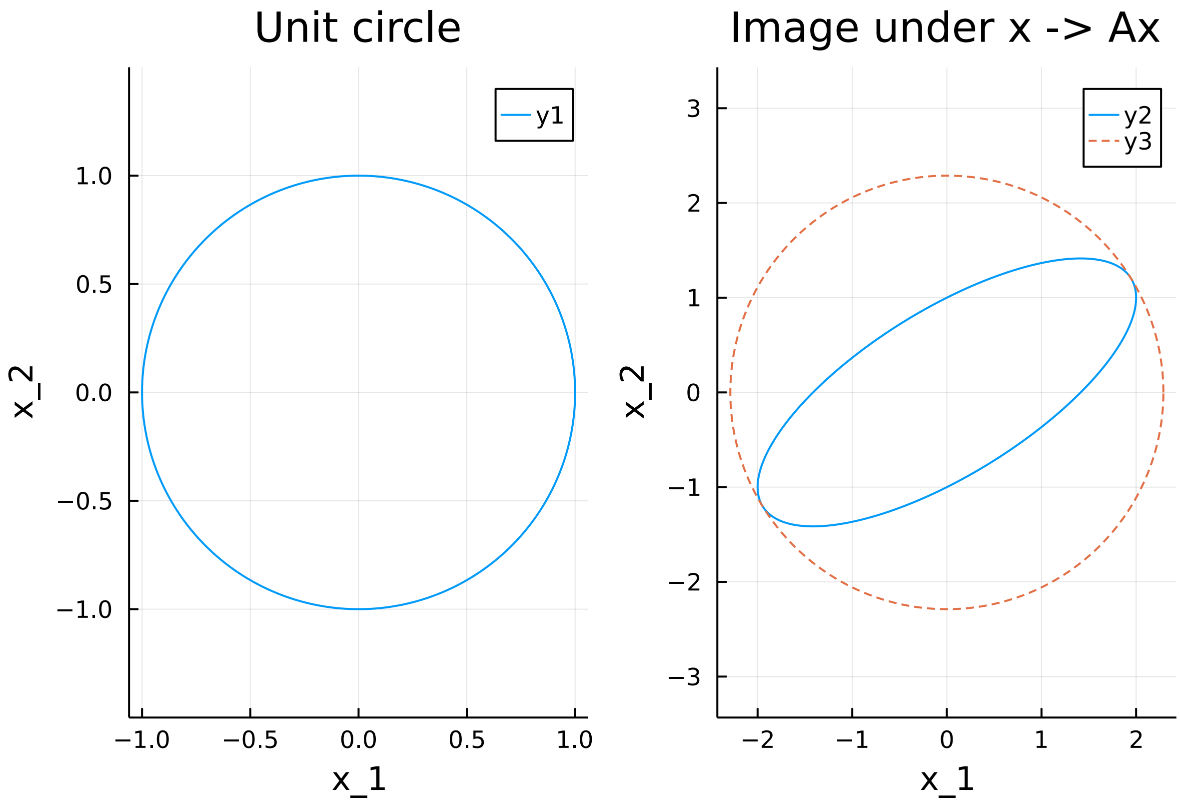
```
using Plots

plot(aspect_ratio=1, layout=(1,2), xlabel="x_1", ylabel="x_2") #creates a "layout" -- subsequent plot! c
plot!(x[1,:],x[2,:], subplot=1,title="Unit circle")
```



Now, the function $\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x}$ defines a mapping from $\mathbb{R}^2$ to $\mathbb{R}^2$. Let's see what this does to the vectors in the unit circle!

```
Ax = A*x;
```

```
plot!(Ax[1,:],Ax[2,:],subplot=2,title="Image under x -> Ax")
plot!(twonorm*x[1,:],twonorm*x[2,:], subplot=2,l=:dash)
```



2 Classical Iterative Methods for Solving Linear Systems

We've talked a bit last week about *direct* methods for solving linear systems of equations. There is another class of methods known as *iterative methods* which use an iterative sequence of steps to make incremental improvement of an approximate solution to the system.