1 Preliminary background

1.1 Linear systems of equations

Linear systems are one of the most fundamental mathematical tools in science and engineering. The solution to a lot of computational mathematical problems boils down to solution of linear systems. Knowing how to treat such systems opens the door to treat sophisticated problems that arise in different applications. There arise in optimization problems, solution of non-linear equations, solution of partial differential equations and more.

In introduction to linear algebra we saw how to solve linear systems. However, in practice, when solving these systems using a computer program, some huge errors may occur due to the finite precision of our processor. In addition, some of the standard methods may be inefficient or impractical in some cases, leading the need to more sophisticated algorithms.

In "numerical linear algebra" we will consider linear systems such as:

$$\begin{cases} 3x_1 & + & 4x_2 & - & 2x_3 & = & 5 \\ 10x_1 & + & 2x_2 & + & x_3 & = & 15 \\ 1x_1 & + & x_2 & + & x_3 & = & 1 \end{cases} \tag{1}$$

We will use matrix notation: this system is written as $A\mathbf{x} = \mathbf{b}$ where

$$A = \begin{bmatrix} 3 & 4 & -2 \\ 10 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} 5 \\ 15 \\ 1 \end{bmatrix}$$

In this course we will consider matrices $A \in \mathbb{C}^{m\times n}$, or $A \in \mathbb{R}^{m\times n}$, and will denote their entries by

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{m1} & a_{m2} & & a_{mn} \end{bmatrix}$$

A vector $\mathbf{x} \in \mathbb{C}^n$ or $\mathbf{x} \in \mathbb{R}^n$

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

For simplicity we will consider real matrices throughout the course, but initially we will consider also complex valued matrices for some fundamental definitions. Most of the algorithms that we will learn later also have a complex valued version.

There are two types of approaches for solving these linear systems:

1. **Direct Methods:** Gaussian elimination and LU decomposition.
   In these methods we perform an algorithm, that would have got the exact solution if there were no round-off errors. Since there are round-off errors, we may get significant errors in the solution if we are not careful.

2. **Iterative Methods**
   These methods will work using iterations and will provide an approximate solution that gets improved as we do more iterations. The solution will be approximate regardless of round-off errors. In many applications, direct solution is much more expensive than iterative solution.

The existence of a solution. We know that the system $A\mathbf{x} = \mathbf{b}$ above has either of the following:

1. *A unique solution.* In this case the columns or rows of the matrix A are linearly independent, and the matrix is non-singular. In this case the matrix is also said to be of "full-rank".

2. *No solution*–the vector $\mathbf{b}$ cannot be expressed as a linear combination of the columns of A. In this case there must be at least one column which is linearly dependent with the others.

3. *Infinite number of solutions.* In this case, the vector $\mathbf{b}$ can be expressed as a linear combination of the columns of A, but there must be at least one column which is

linearly dependent with the others. That means that there is a non-trivial vector $\mathbf{e}$ such that $A\mathbf{e} = 0$, and the matrix is called singular.

We will now learn methods for getting the solution $\mathbf{x}$ of the system $A\mathbf{x} = \mathbf{b}$.

Transposition The transpose of a matrix A (or conjugate transpose) will be denoted by A^T (or A^*).

$$(A^T)_{ij} = (A)_{ji} = a_{ji} \quad (A^*)_{ij} = (\bar{A})_{ji} = \bar{a}_{ji},$$

where $\bar{a}$ denotes the conjugate of a scalar a.
A matrix will be called Symmetric (Hermitian) if $A^T = A$ ($A^* = A$).

1.2 Errors in the solution of linear systems

We need a distance measures for matrices and vectors - norms.
Examples:

- Suppose we have a linear system $A\mathbf{x} = \mathbf{b}$ that we wish to solve. From practical reasons, we get a slightly perturbed system

  $$\tilde{A}\mathbf{x} = \tilde{\mathbf{b}},$$

  represented in our system. How this will affect the answer $\mathbf{x}$?

- We have a model that corresponds to data pairs $\{(x_i, y_i)\}$:

  $$Model(x_i, \theta) \approx y_i.$$

  How do we wrap the differences between prediction and observed data into a single number to measure how good our model is?

To answer such questions, and many others, we will first have to define the distance measure of vectors and matrices. Such measures will be called “norms”.

1.3 Vector norms

Definition 1 (Vector norms)**.** *The term* $\|\cdot\|$ *is called a norm over the vector space* $\mathbb{C}^n$ *(or* $\mathbb{R}^n$*) if the norm is*

- *Non-negative:* $\forall \mathbf{v} \in \mathbb{C}^n : \quad \|\mathbf{v}\| \geq 0$*, and* $\|\mathbf{v}\| = 0 \Leftrightarrow \mathbf{v} = 0$*.*
- *Homogenous:* $\forall \mathbf{v} \in \mathbb{C}^n, \alpha \in \mathbb{C} : \|\alpha \mathbf{v}\| = |\alpha| \|\mathbf{v}\|$*.*
- *Triangle inequality:* $\forall \mathbf{v}, \mathbf{u} \in \mathbb{C}^n : \|\mathbf{v} + \mathbf{u}\| \leq \|\mathbf{v}\| + \|\mathbf{u}\|$*.*

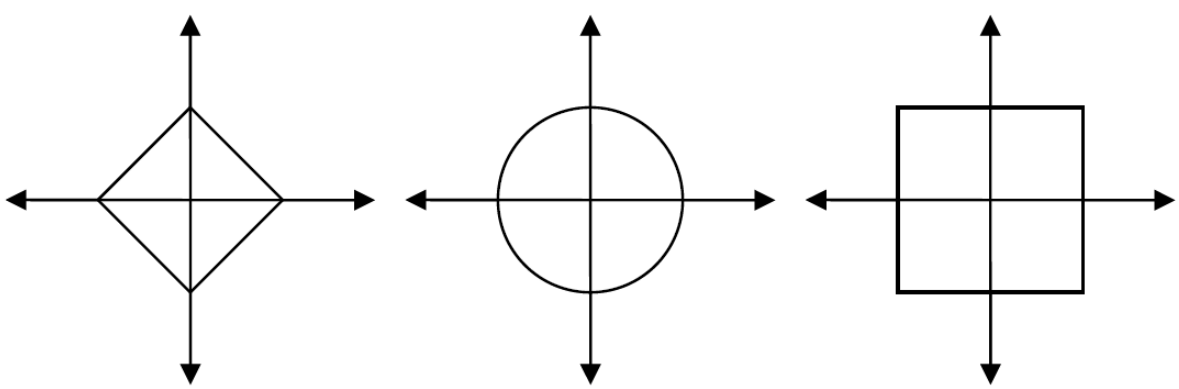
The common vector norm is called the ℓ_p norm

$$\|\mathbf{v}\|_p = \left(\sum_{i=1}^{n} |v_i|^p \right)^{\frac{1}{p}},$$

and usually we use one of the following (more popular) norms :

$$\|\mathbf{v}\|_1 = \sum_{i=1}^{n} |v_i|, \quad \|\mathbf{v}\|_2 = \left(\sum_{i=1}^{n} v_i^2 \right)^{\frac{1}{2}}, \quad \|\mathbf{v}\|_\infty = \max_i |v_i| \quad (**)$$

for which the three unit circles are (respectively for the ℓ_1, ℓ_2, and ℓ_∞)



(**) Explanation: by definition we have that $\max_i |v_i| \leq \|\mathbf{v}\|_p \leq \sqrt[p]{n} \max_i |v_i|$, and taking $p \to \infty$ will result in $\sqrt[p]{n} \to 1$.

Definition 2 (Vector inner product)**.** *The operation* $\langle \cdot, \cdot \rangle$ *is an inner product between two vectors* $\mathbf{u}, \mathbf{v} \in \mathbb{C}^n$ *if the following hold*

- *Conjugate symmetry:* $\langle \mathbf{u}, \mathbf{v} \rangle = \overline{\langle \mathbf{v}, \mathbf{u} \rangle}$
- *Linearity*[1]*:* $\langle \mathbf{u}, \mathbf{v}_1 + \mathbf{v}_2 \rangle = \langle \mathbf{u}, \mathbf{v}_2 \rangle + \langle \mathbf{u}, \mathbf{v}_1 \rangle$ *and* $\langle \alpha \mathbf{u}, \beta \mathbf{v} \rangle = \bar{\alpha}\beta \langle \mathbf{u}, \mathbf{v} \rangle$ *for* $\alpha, \beta \in \mathbb{C}$.
- *Non-negativity:* $\langle \mathbf{u}, \mathbf{u} \rangle \geq 0$ *and* $\langle \mathbf{u}, \mathbf{u} \rangle = 0$ *iff* $\mathbf{u} = 0$.

The main inner product that we will consider in this course is the dot product between two vectors:

$$\langle \mathbf{u}, \mathbf{v} \rangle = \sum_i \bar{u}_i v_i = \mathbf{u}^* \mathbf{v}.$$

It can be shown that $\|\mathbf{u}\| = \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle}$ is a norm for every inner product. For this norm, the Cauchy Schwartz inequality is given by

$$|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\| \|\mathbf{v}\|.$$

Example 1 (Energy norm)**.** *Let* $M \in \mathbb{C}^{n \times n}$ *be Hermitian positive definite. Then*

1. *For every inner product* $\langle \mathbf{u}, \mathbf{v} \rangle$*, the inner product* $\langle \mathbf{u}, M\mathbf{v} \rangle$ *is also a valid inner product, and is often denoted as* $\langle \mathbf{u}, \mathbf{v} \rangle_M$*.*
2. *The inner product* $\sqrt{\langle \mathbf{u}, M\mathbf{u} \rangle}$ *is a norm on* $\mathbb{C}^n$*, often denoted as* $\|\mathbf{u}\|_M$*.*

[1] There is some disagreement on this definition. In some cases the definition states: $\langle \alpha \mathbf{u}, \beta \mathbf{v} \rangle = \alpha \bar{\beta} \langle \mathbf{u}, \mathbf{v} \rangle$ for $\alpha, \beta \in \mathbb{C}$. Here, we choose our form somewhat arbitrarily, since this is the definition used by Matlab/Julia where the Hermitian inner product is calculated by `dot(u,v)`.