5374 Fall '22

Numerical Linear Algebra Lecture 4

A rapid linear algebra review

(See Chapter 1 in Solomon's book, and Chapter 4 's section "Sensitivity analysis")

Let's review a few basic concepts.

Vector Spaces

Is a set V with two operations

1. Sum : $v_1, v_2 \in V$ $\quad v_1 + v_2 \in V$
   $(v_1 + v_2 = v_2 + v_1)$
2. Multiplication by a scalar
   $\alpha \in \mathbb{R}$, $v \in V$, $\quad \alpha v \in V$
   $(\alpha \in \mathbb{C})$
3. There is a zero vector
   $v + 0 = 0 + v = v$ for all $v \in V$
4. For every $v \in V$ $\exists!$ v' s.t.
   $v + v' = 0$
5. Given $\alpha_1, \alpha_2 \in \mathbb{R}$ (or $\mathbb{C}$) then
   $\alpha_1 v + \alpha_2 v = (\alpha_1 + \alpha_2) v$

If the multiplication by scalars is defined just for $\mathbb{R}$, the vector space is called real, and for $\mathbb{C}$, the vector space is called complex.

Example 1. Given $n \in \mathbb{N}$, $\mathbb{R}^n = \left\{ x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \,\middle|\, x_k \in \mathbb{R} \right\}$

Example 2. Given $n \in \mathbb{N}$, $\mathbb{C}^n = \left\{ x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \,\middle|\, x_k \in \mathbb{C} \right\}$

Example 3. ("Little L^2")

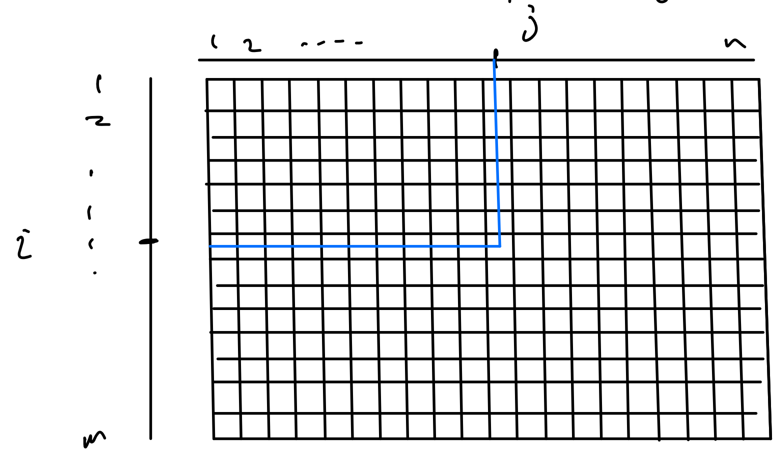
$$\ell_2(\mathbb{Z}) = \left\{ x = (x_n)_{n \in \mathbb{Z}} \,\middle|\, x_n \in \mathbb{C} \ \forall n \in \mathbb{Z} \text{ and } \sum_{n \in \mathbb{Z}} |x_n|^2 < \infty \right\}$$

Example 4. ("Big L^2")

$$L^2(\mathbb{R}) = \left\{ f : \mathbb{R} \to \mathbb{C} \,\middle|\, f \text{ is Lebesgue measurable and } \int_{\mathbb{R}} |f(x)|^2 \, dx < \infty \right\}$$

$$L^2(\mathbb{T}) = \left\{ f : \mathbb{R} \to \mathbb{C} \,\middle|\, f \text{ is Lebesgue measurable and } f(x+1) = f(x) \ \forall x, \ \int_0^1 |f(x)|^2 \, dx < \infty \right\}$$

Example 5 Consider the following $m \times n$ grid:



x_{ij} = real number indicating how bright or dark we make the ij-cell.

We have a set

$$G = \{ (i,j) \mid 1 \leq i \leq m, \ 1 \leq j \leq n \}$$

Functions $f : G \longrightarrow \mathbb{R}$ represent grayscale images, the set of all such functions is a real vector space.

This space has dimension mn, so it is isomorphic to $\mathbb{R}^{mn}$

Matrices and Matrix-vector multiplication

From here on we will mostly talk about the real vector spaces $\mathbb{R}^n$, $n \in \mathbb{N}$ but every once in a while we will refer to other spaces.

A $\underset{\text{\# of columns}}{\overset{\text{\# of rows}}{m \times n}}$ matrix looks like this

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

Any $m \times n$ matrix A can multiply a vector $x \in \mathbb{R}^n$ (from the left) and produce a vector in $\mathbb{R}^m$

Vectors we typically write as columns

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

and the columns of a matrix A can be themselves seen as vectors,

a_{1k}

$$\bar{a}_k = \begin{pmatrix} a_{1k} \\ a_{2k} \\ \vdots \\ a_{mk} \end{pmatrix} \in \mathbb{R}^m$$

Ax is the vector whose i-th entry is given by

$$(Ax)_i = \sum_{j=1}^{n} a_{ij} x_j \quad , \; i = 1, \dots, m$$

the i-th row of A

In particular, if we consider the canonical basis vectors of $\mathbb{R}^n$:

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \; e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, \; e_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

$$Ae_k = \begin{pmatrix} a_{1k} \\ a_{2k} \\ \vdots \\ a_{mk} \end{pmatrix} = \bar{a}_k$$

That is, the columns of the matrix A are simply what we get when we multiply

A by the canonical basis vectors of $\mathbb{R}^n$.

Now that we know this about matrix-vector multiplication and the columns of A it is not difficult to see the following:

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

$$= x_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ \vdots \\ a_{m2} \end{pmatrix} + \cdots + x_n \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix}$$

$$= x_1 \bar{a}_1 + x_2 \bar{a}_2 + \cdots + x_n \bar{a}_n$$

In other words, Ax is a linear combination of the columns of A given by the coefficients of x.

With this is easy to relate statements about families of vectors and matrices.

Exercise: Let A be a $n \times n$ matrix. Show that A is invertible if and only if the columns of A are linearly independent.

Observe that A can be associated with a function from $\mathbb{R}^n$ ($n = \#$ of columns) to $\mathbb{R}^m$ ($m = \#$ of rows):

$$x \in \mathbb{R}^n \longmapsto Ax \in \mathbb{R}^m$$

If we denote this function by L, then it has the following properties:

$$L(\alpha x) = \alpha L(x) \quad \forall \; x \in \mathbb{R}^n, \alpha \in \mathbb{R}$$
$$L(\bar{x}_1 + \bar{x}_2) = L(\bar{x}_1) + L(\bar{x}_2) \quad \forall \; x_1, x_2 \in \mathbb{R}^n$$
$$L(0) = 0$$

A function L from $\mathbb{R}^n$ to $\mathbb{R}^m$ with these properties is called a **linear transformation**.

Exercise: Think about how given a linear transformation $L: \mathbb{R}^n \to \mathbb{R}^m$, then must be a $m \times n$ matrix A s.t. $L(x) = Ax$.

Next time we will say a few things about matrix-matrix multiplication, inner and outer products, and about bases.

Measuring and estimating error

We will be concerned with algorithms for the equation

$$Ax = b$$

The algorithms often will be iterative and don't give an exact answer in a finite number of steps. This raises the need of estimating the distance from our approximate solution to the "analytic" solution.

Moreover, arithmetic operation with "real numbers" cannot always be represented exactly in a computer, so even the back-substitution algorithm will produce errors from the arithmetic operations ("round-off" or "machine error")

This brings us to the way we estimate the sizes of vectors and matrices, norms.

Norms (see Solomon's chapter 4)

A **norm** in a vector space V is a function

$$N : V \longrightarrow \mathbb{R}$$

Such that
- $N(x) \geq 0 \quad \forall x$
- $N(x) = 0$ if, and only if, $x = 0$
- $N(\alpha x) = |\alpha| N(x) \quad \forall \alpha \in \mathbb{R}$ (or $\mathbb{C}$)
- $N(x+y) \leq N(x) + N(y) \quad \forall x, y \in V$

Typically $N(x)$ is denoted by $\|x\|$, sometimes with a subscript like $\|x\|_2$ or $\|x\|_\infty$ to make explicit which norm we are talking about:

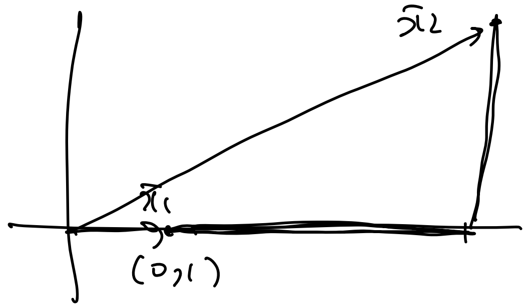
Example 1. The Euclidean norm

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \qquad \|x\|_2 = \left(\sum_{k=1}^{n} x_k^2 \right)^{1/2}$$

When we write $\|x\|$ (without a subscript) we will assume we mean this norm.

Example 2. The Manhattan or Taxi metric (AKA ℓ_1)

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \qquad \|x\|_1 = |x_1| + |x_2| + \cdots + |x_n|$$



$(3,3)$

$\|\bar{x}_1 - \bar{x}_2\|_1$

$= |3-1| + |3-0|$

$= 2 + 3 = 5$

Example 3 The l_p norm $(1 \le p < \infty)$

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \qquad \|x\|_p = \left(\sum_{k=1}^{n} |x_k|^p \right)^{1/p}$$

To check that this satisfies the triangle inequality, see the Minkowski inequality.

Example 4 The l_∞ norm

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \qquad \|x\|_\infty = \max_{1 \le k \le n} |x_k|$$

Exercise: Fix $x \in \mathbb{R}^n$, show that

$$\lim_{p \to \infty} \|x\|_p = \|x\|_\infty$$

Matrix norms

Let A be a $n \times n$ matrix, let's go over some norms we can define for A.

Example 1. The Frobenius norm, $\|A\|_{Fr}$

$$\|A\|_{Fr} = \left(\sum_{i=1}^{n} \sum_{j=1}^{m} a_{ij}^{2} \right)^{1/2}$$

Exercise: Given A, it's trace, denoted $tr(A)$ is the sum of the elements in the diagonal of A. Check that

$$\|A\|_{Fr} = \left(tr\left(A \cdot A^{t}\right) \right)^{1/2}$$

where A^{t} denotes the transpose of A.

Example 2 The $(l_2 - l_2)$ operator norm of A

Let A be a $m \times n$ matrix, we define

$$\|A\|_{op} = \max_{\substack{x \in \mathbb{R}^n, \\ x \neq 0}} \frac{\|Ax\|_2}{\|x\|_2}$$

The number $\|A\|_{op}$ is the smallest number among all $\lambda \geq 0$ s.t.

$$\|Ax\|_2 \leq \lambda \|x\|_2 \quad \forall \; x \in \mathbb{R}^n.$$

In particular,

$$\|Ax\|_2 \leq \|A\|_{op} \|x\|_2.$$

We will say two norms $\|x\|_\alpha$, $\|x\|_\beta$ are **equivalent** with constants c_1, c_2 if

$$c_1 \|x\|_\beta \leq \|x\|_\alpha \leq c_2 \|x\|_\beta \quad \forall \; x \in \mathbb{R}^n$$

Remark. In the space

$$C[0,1] = \{ f : [0,1] \to \mathbb{R} \mid f \text{ is continuous} \}$$

the two norms:

$$\|f\|_2 = \left(\int_0^1 |f|^2 dx \right)^{1/2}$$

$$\|f\|_1 = \int_0^1 |f| \, dx$$

are **NOT** equivalent.