4 Norms (Part 1)

(Largely based on [TB97, Lecture 2]).

4.1 Transpose and Conjugate Transpose

If $A \in \mathbb{R}^{m\times n}$, then its transpose, $B = A^T \in \mathbb{R}^{n\times m}$, is such that $b_{ij} = a_{ji}$. If $A \in \mathbb{C}^{m\times n}$, then its "*conjugate transpose*" a.k.a. "*hermitian conjugate*" is $B = A^\star \in \mathbb{C}^{n\times m}$ such that $b_{ij} = \bar{a}_{ji}$, where $\bar{x}$ denotes the usual complex conjugate.

Note that $(AB)^\star = A^\star B^\star$. Also, if $A \in \mathbb{R}^{m\times n}$, then $A^T = A^\star$.

If $A \in \mathbb{C}^{m\times m}$ is such that $A = A^*$, then we say that A is *Hermitian.* If $A \in \mathbb{R}^{m\times m}$ and $A = A^T$, we say that A is symmetric.

4.2 Inner products

Recall...

$$(x, y) = \mathbf{x}^T \mathbf{y} = \sum_{i=1}^{m} x_i y_i \text{ for } \mathbf{x}, \mathbf{y} \in \mathbb{R}^n,$$

$$(x, y) = \mathbf{x}^* \mathbf{y} = \sum_{i=1}^{m} \bar{x}_i y_i \text{ for for } \mathbf{x}, \mathbf{y} \in \mathbb{C}^n,$$

where $x^* = \bar{x}^T$ is the "*conjugate transpose*", some times called the "*hermitian conjugate*". There are other Inner Products, of course....

We say $\mathbf{x}$ and $\mathbf{y}$ are *orthogonal* if $\mathbf{x}^\star \mathbf{y} = 0$, but we'll postpone most of the discussion on this until after we've introduced the idea of a *norm.*

4.3 Norms

Much of this section is done in the context of vectors of real numbers. As an exercise you'll determine the corresponding results over $\mathbb{C}^n$.

Definition 4.1. Let $\mathbb{R}^n$ be all the vectors of length n of real numbers. The function $\|\cdot\|$ is called a **norm** on $\mathbb{R}^n$ if, for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$

(i) $\|x\| = 0$ if and only if $x = 0$.

(ii) $\|\lambda \mathbf{x}\| = |\lambda| \|\mathbf{x}\|$ for any $\lambda \in \mathbb{R}$,

(iii) $\|\mathbf{x} + \mathbf{y}\| \leqslant \|\mathbf{x}\| + \|\mathbf{y}\|$ (triangle inequality).

The norm of a vector gives us some information about the "*size*" of the vector. But there are different ways of measuring the size: you could take the absolute value of the largest entry, you cold look at the "distance" for the origin, etc...

Definition 4.2. (i) The 1-*norm* is

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

(ii) The 2-*norm* (a.k.a. the *Euclidean norm*)

$$\|x\|_2 = \left(\sum_{i=1}^{n} x_i^2 \right)^{1/2}.$$

Note, if $\mathbf{x}$ is a vector in $\mathbb{R}^n$, then

$$\mathbf{x}^T \mathbf{x} = x_1^2 + x_2^2 + \cdots + x_n^2 = \|\mathbf{x}\|_2^2.$$

(iii) The ∞-*norm* (also known as the max-*norm*)

$$\|\mathbf{x}\|_\infty = \max_{i=1}^{n} |x_i|.$$

(iv) One can consider all the above to be an example of a p-*norm*

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

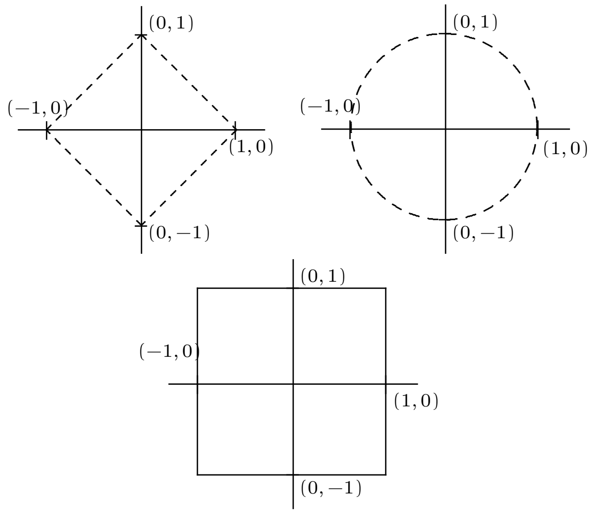

Figure 4.1: The unit vectors in $\mathbb{R}^2$: $\|\mathbf{x}\|_1 = 1$, $\|\mathbf{x}\|_2 = 1$, $\|\mathbf{x}\|_\infty = 1$,

In Figure 4.1, the top left diagram shows the unit ball in $\mathbb{R}^2$ given by the 1-norm: the vectors $\mathbf{x} = (x_1, x_2)$ in R^2 are such that $\|\mathbf{x}\|_1 = |x_1| + |x_2| = 1$ are all found on the diamond. In the top right, the vectors have $\sqrt{x_1^2 + x_2^2} = 1$ and so are arranged in a circle. The bottom diagram gives the unit ball in $\|\cdot\|_\infty$ – the largest component of each vector is 1.

It takes a little bit of effort to show that $\|\cdot\|_2$ satisfies the triangle inequality. First we need

Lemma 4.3 (Cauchy-Schwarz)**.**

$$|x^T y| \leqslant \|\mathbf{x}\|_2 \|\mathbf{y}\|_2, \qquad \forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n.$$

Lemma 4.3 is actaully a special case of *Hölder's Inequality*:

$$|\mathbf{x}^T \mathbf{y}| \leqslant \|\mathbf{x}\|_p \|\mathbf{y}\|_q.$$

This can then be applied to show that

Lemma 4.4. *For all* $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$*,*

$$\|\mathbf{x} + \mathbf{y}\| \leqslant \|\mathbf{x}\|_2 + \|\mathbf{y}\|_2.$$

It follows directly that $\|\cdot\|_2$ is a norm.