

CPSC 540 Notes on Norms

Mark Schmidt

Fall 2014

1 General Norms

In class we've used the ℓ_2 -norm and the ℓ_1 -norm as a measure of the length of a vector, and the concept of a norm generalizes this idea. In particular, we say that a function f is a *norm* if it satisfies the following three properties:

- 1 $f(x) = 0$ implies $x = 0$ (separate points)
- 2 $f(\alpha x) = |\alpha|f(x)$, for scalar α (absolute homogeneity)
- 3 $f(x + y) \leq f(x) + f(y)$ (triangle inequality)

Important implication of these properties are that norms are non-negative,

$$f(x) \geq 0,$$

and norms are convex. All possible norms are said to be *equivalent*, in that sense that if f and g are norms then there exist constants β_1 and β_2 such that

$$\beta_1 f(x) \leq g(x) \leq \beta_2 f(x).$$

2 General ℓ_p -norms

The most important class of norms on vectors are the ℓ_p -norms, defined by

$$\|x\|_p = \left(\sum_{i=1}^d |x_i|^p \right)^{1/p},$$

for some $p \geq 1$. The three most important special case are the ℓ_1 -norm, ℓ_2 -norm, and ℓ_∞ norm,

$$\|x\|_1 = \sum_{i=1}^d |x_i|, \quad \|x\|_2 = \sqrt{\sum_{i=1}^d x_i^2}, \quad \|x\|_\infty = \max_i \{|x_i|\}.$$

We recognize the ℓ_2 -norm as the standard straight-line distance in Euclidean-space, and it is often simply denoted $\|x\|$. The constants that relate these norms are

$$\begin{aligned} \|x\|_\infty &\leq \|x\|_1 \leq d\|x\|_\infty \\ \|x\|_\infty &\leq \|x\|_2 \leq \sqrt{d}\|x\|_\infty \\ \|x\|_2 &\leq \|x\|_1 \leq \sqrt{d}\|x\|_2. \end{aligned}$$

Some works use $p < 1$ (e.g., as a regularizer that encourages a higher degree of sparsity than the ℓ_1 -norm), but this does not satisfy the properties of a norm (e.g., $\|x\|_p$ for $p < 1$ is not convex). Some works also define the zero ‘norm’ to be the number of non-zero elements, $\|x\|_0 = \sum_{i=1}^d I(x_i \neq 0)$, and again this is not an actual norm.

3 Cauchy-Schwartz and Hölder Inequalities, Dual Norm

The Cauchy-Schwartz inequality bounds inner products by the product of Euclidean norms,

$$\sum_{i=1}^d |x_i| |y_i| \leq \|x\| \|y\|,$$

which implies

$$x^T y \leq \|x\| \|y\|,$$

and (according to Wikipedia) is one of the most important inequalities in mathematics. A generalization is Hölder’s inequality,

$$\sum_{i=1}^d |x_i| |y_i| \leq \|x\|_p \|y\|_q,$$

where

$$\frac{1}{p} + \frac{1}{q} = 1.$$

Cauchy-Schwartz is the special case where $p = q = 2$, but we could also have $p = 1$ and $q = \infty$. The norm $\|\cdot\|_q$ is said to be the *dual norm* of $\|\cdot\|_p$. The general definition of the dual norm $\|\cdot\|_*$ for a general norm $\|\cdot\|$ is

$$\|x\|_* = \sup\{w^T x \mid \|x\| \leq 1\}.$$