Young's, Hölder's and Minkowski's Inequalities

In class I derived the *triangle inequality* for the 2-norm (often called the Euclidean norm) on the vector space $\mathbb{R}^2$,

$$||\mathbf{x}||_2 \equiv \sqrt{|x_1|^2 + |x_2|^2} \quad \Rightarrow \quad ||\mathbf{x}+\mathbf{y}||_2 \leq ||\mathbf{x}||_2 + ||\mathbf{y}||_2.$$

I used a somewhat brute force calculation together with the simple fact that

$$0 \leq (a-b)^2 = a^2 - 2ab + b^2 \quad \Rightarrow \quad ab \leq \tfrac{1}{2}a^2 + \tfrac{1}{2}b^2.$$

On your homework you are asked to extend this result from $\mathbb{R}^2$ to $\mathbb{R}^n$.

Below I will show how to generalize the triangle inequality to the p-norm, $p \geq 1$, which on $\mathbb{R}^n$ is defined by

$$||\mathbf{x}||_p \equiv \sqrt[p]{\sum_{i=1}^{n} |x_i|^p}.$$

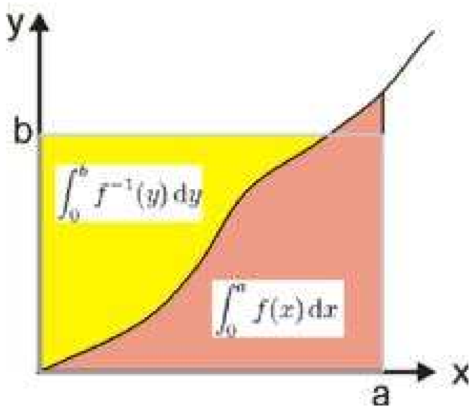
The derivation is much more subtile than was required for the 2-norm. I'll break the problem up into establishing three separate inequalities: (1) Young's Inequality, (2) Hölder's Inequality, and finally (3) Minkowski's Inequality which is the name often used to refer to the p-norm triangle inequality. .

(1) Young's Inequality

For any real numbers $a \geq 0$ and $b \geq 0$ and $p > 1$ we have

$$ab \leq \frac{1}{p}a^p + \frac{1}{q}b^q, \quad \text{where} \quad q = \frac{p}{p-1}.$$

In class we used the special case with $p = q = 2$ to derive Cauchy-Schwarz.



Clearly, the sum of the highlighted areas given by $\int_0^b f^{-1}(y)\,dy + \int_0^a f(x)\,dx$ is greater than or equal to area of the box of height b and width a. (The image above comes from Wikipedia.)

Consider the figure above where $f(x)$ at this stage can be any increasing function of x which satisfies $f(0) = 0$. It is easy to see by comparing the box delineated by the coordinated axes and $x = a$ and $y = b$ to the two shaded regions in the figure we must have

$$ab \le \int_0^b f^{-1}(y)\,dy + \int_0^a f(x)\,dx.$$

The first integral is the area of the shaded region on the left of the graph $y = f(x)$ and the second integral is the area of the shaded region on the right. Now, let's specifically pick f. For $p > 1$ set $f(x) = x^{p-1}$. Compute f's inverse function to get $f^{-1}(y) = y^{q-1}$ where $q = p/(p-1)$. Integrate

$$\int_0^a f(x)\,dx = \frac{1}{p}a^p \quad \text{and} \quad \int_0^b f^{-1}(y)\,dy = \frac{1}{q}b^q,$$

and the desired inequality follows. This clever Calculus I based proof can be found at [1].

(2) Hölder's Inequality

For $p > 1$ and $q = p/(p-1)$, Hölder's Inequality says

$$\sum_{i=1}^{n} a_i b_i \le ||\mathbf{a}||_p ||\mathbf{b}||_q.$$

p and q are said to be *dual exponents* and are related by $1/p + 1/q = 1$. The Cauchy-Schwarz Inequality is a special case of Hölder when $p = q = 2$.

Hölder follows easily from Young's Inequality. Let $\widehat{\mathbf{a}} = \mathbf{a}/||\mathbf{a}||_p$ and $\widehat{\mathbf{b}} = \mathbf{b}/||\mathbf{b}||_q$. Using Young's Inequality n times we get

$$\sum_{i=1}^{n} \left(\widehat{a}_i \widehat{b}_i\right) \le \sum_{i=1}^{n} \left(\frac{1}{p}|\widehat{a}_i|^p + \frac{1}{q}|\widehat{b}_i|^q\right)$$
$$= \frac{1}{p}\frac{1}{||\mathbf{a}||_p^p}\sum_{i=1}^{n}|a_i|^p + \frac{1}{q}\frac{1}{||\mathbf{b}||_q^q}\sum_{i=1}^{n}|b_i|^q$$
$$= \frac{1}{p} + \frac{1}{q} = 1.$$

Since by the "$\widehat{\ }$" normalization above we have

$$\sum_{i=1}^{n} \widehat{a}_i \widehat{b}_i = \frac{1}{||\mathbf{a}||_p ||\mathbf{b}||_q} \sum_{i=1}^{n} a_i b_i,$$

the desired inequality follows.

[1] `https://en.wikipedia.org/wiki/Young's_inequality`

(3) Minkowski's Inequality

For any $p \geq 1$, Minkowski says $||\mathbf{x}+\mathbf{y}||_p \leq ||\mathbf{x}||_p + ||\mathbf{y}||_p$. The case when $p = 1$ is obviously true. To see it's also true for any $p > 1$ write

$$||\mathbf{x} + \mathbf{y}||_p^p = \sum_{i=1}^{n} |x_i + y_i|^p = \sum_{i=1}^{n} |x_i + y_i| \, |x_i + y_i|^{p-1}.$$

We know that for real numbers $|x_i + y_i| \leq |x_i| + |y_i|$, Use this in the inequality above to conclude

$$||\mathbf{x} + \mathbf{y}||_p^p \leq \sum_{i=1}^{n} |x_i| \, |x_i + y_i|^{p-1} + \sum_{i=1}^{n} |y_i| \, |x_i + y_i|^{p-1}.$$

Finally, use Hölder on the first sum on the right hand side above with $a_i = |x_i|$ and $b_i = |x_i + y_i|^{p-1}$, and again on the second sum with $a_i = |y_i|$ and $b_i = |x_i + y_i|^{p-1}$ to arrive at

$$||\mathbf{x} + \mathbf{y}||_p^p \leq (||\mathbf{x}||_p + ||\mathbf{y}||_p) \sqrt[q]{\sum_{i=1}^{n} |x_i + y_i|^{(p-1)q}}.$$

The fact that $q = p/(p-1)$ allows us to rewrite this as

$$||\mathbf{x} + \mathbf{y}||_p^p \leq (||\mathbf{x}||_p + ||\mathbf{y}||_p) \, ||\mathbf{x} + \mathbf{y}||_p^{p-1}$$

from which the Minkowski's Inequality follows.