Hölder's inequality is a generalization of Cauchy-Schwarz; it allows an arbitrary number of sequences of variables, as well as different weights. First we need the following:

Theorem 1 (Young). For $a, b > 0$ and $p, q > 0$ such that $\frac{1}{p} + \frac{1}{q} = 1$,

$$ab \le \frac{a^p}{p} + \frac{b^q}{q}.$$

This is a special case of the weighted AM-GM inequality.

Theorem 2 (Hölder). Let $a_1, \dots, a_n; b_1, \dots, b_n; \dots; z_1, \dots, z_n$ be sequences of nonnegative real numbers, and let $\lambda_a, \dots, \lambda_z$ be positive reals summing to 1. Then

$$(a_1 + \cdots + a_n)^{\lambda_a}(b_1 + \cdots + b_n)^{\lambda_b} \cdots (z_1 + \cdots + z_n)^{\lambda_z} \ge a_1^{\lambda_a} b_1^{\lambda_b} \cdots z_1^{\lambda_z} + a_n^{\lambda_a} b_n^{\lambda_b} \cdots z_n^{\lambda_z}.$$

Proof. First we prove that

$$(a_1^p + \cdots + a_n^p)^{\frac{1}{p}}(b_1^q + \cdots + a_n^q)^{\frac{1}{q}} \ge a_1 b_1 + \cdots + a_n b_n, \quad (1)$$

when $p, q > 0$ and $\frac{1}{p} + \frac{1}{q}$, which is equivalent to the theorem statement for 2 variables.

Let $A = (a_1^p + \cdots + a_n^p)^{\frac{1}{p}}$ (also denoted $||a||_p$) and $B = (b_1^q + \cdots + b_n^q)^{\frac{1}{q}}$. Let $a_i' = \frac{a_i}{A}$ and $b_i' = \frac{b_i}{B}$. Now that we've "normalized" $(a_1, \dots, a_n)$ and $(b_1, \dots, b_n)$ so that $(a_1'^p + \cdots + a_n'^p)^{\frac{1}{p}} = 1$ and $(b_1'^q + \cdots + b_n'^q)^{\frac{1}{q}} = 1$, we can apply Minkowski's inequality.

$$a_1' b_1' + \cdots + a_n' b_n' \le \frac{1}{p}(a_1'^p + \cdots + a_n'^p)^{\frac{1}{p}} + \frac{1}{q}(b_1'^q + \cdots + b_n'^q)^{\frac{1}{q}} = \frac{1}{p} + \frac{1}{q} = 1$$

Multiplying by AB gives (1).

The general inequality follows by induction on the number of sequences. For example, passing from 2 to 3 sequences, apply Hölder to with weights $\frac{\lambda_a}{\lambda_a + \lambda_b}, \frac{\lambda_b}{\lambda_a + \lambda_b}$, then with weights $\lambda_a + \lambda_b$, λ_c. □

Theorem 3 (Minkowski). Let $p > 1$ and let $a_1, \dots, a_n, b_1, \dots, b_n, \dots, z_1, \dots, z_n$ be positive numbers. Then

$$\begin{aligned}(|a_1|^p + \cdots + |a_n|^p)^{\frac{1}{p}} + (|b_1|^p + \cdots + |b_n|^p)^{\frac{1}{p}} + \cdots + (|z_1|^p + \cdots + |z_n|^p)^{\frac{1}{p}} \\ \ge [(|a_1 + \cdots + z_1|)^p + (|a_2 + \cdots + z_2|)^p + \cdots + (|a_n + \cdots + z_n|)^p]^{\frac{1}{p}}\end{aligned}$$

Proof. We first prove Minkowski for 2 sequences. We have

$$\begin{aligned}\sum_{k=1}^{n} |a_k + b_k|^p &= \sum_{k=1}^{n} (|a_k| + |b_k|)|a_k + b_k|^{p-1} \\ &= \sum_{k=1}^{n} |a_k||a_k + b_k|^{p-1} + \sum_{k=1}^{n} |b_k||a_k + b_k|^{p-1} \\ &\le \left(\left(\sum_{k=1}^{n} |a_k|^p \right)^{\frac{1}{p}} + \left(\sum_{k=1}^{n} |b_k|^p \right)^{\frac{1}{p}} \right) \left(\sum_{k=1}^{n} |a_k + b_k|^{(p-1)\left(\frac{p}{p-1}\right)} \right)^{\frac{p-1}{p}}\end{aligned}$$

$$\left(\sum_{k=1}^{n} |a_k + b_k|^p \right)^{\frac{1}{p}} \le \left(\sum_{k=1}^{n} |a_k|^p \right)^{\frac{1}{p}} + \left(\sum_{k=1}^{n} |b_k|^p \right)^{\frac{1}{p}}$$

The general case follows by induction. □

As a corollary, $||x||_p = \sqrt[p]{x_1^p + \cdots + x_n^p}$ is a valid norm in n-dimensional space.