The unifom-convexity inequality for ℓ_p-norms

Jiří Matoušek

Rev. 30/I/2009

We want to prove the inequality

$$\frac{\|\mathbf{x}+\mathbf{y}\|_p^2+\|\mathbf{x}-\mathbf{y}\|_p^2}{2} \geq \|\mathbf{x}\|_p^2+(p-1)\|\mathbf{y}\|_p^2, \quad 1<p<2, \tag{1}$$

where $\mathbf{x},\mathbf{y} \in \mathbb{R}^d$ are arbitrary vectors. (The proof can also be extended for vectors in ℓ_p or functions in L_p, but some things come out slightly simpler in finite dimension.) This exposition is based on a sketch given as the first proof of Proposition 3 in Ball, Carlen, and Lieb [1]. The second proof from that paper has been worked out by Assaf Naor; see `http://www.cims.nyu.edu/~naor/homepage/files/inequality.pdf`. I consider the first proof somewhat more conceptual and accessible for a non-expert.

First we pass to an inequality formally stronger than (1), with the same right-hand side:

$$\left(\frac{\|\mathbf{x}+\mathbf{y}\|_p^p+\|\mathbf{x}-\mathbf{y}\|_p^p}{2}\right)^{2/p} \geq \|\mathbf{x}\|_p^2+(p-1)\|\mathbf{y}\|_p^2. \tag{2}$$

To see that the l.h.s. of (2) is never smaller than the l.h.s. of (1), we use the following well-known fact: The qth degree average $\left(\frac{a^q+b^q}{2}\right)^{1/q}$ is a nondecreasing function of q for a,b fixed. We apply this with $a=\|\mathbf{x}+\mathbf{y}\|_p^2$, $b=\|\mathbf{x}-\mathbf{y}\|_p^2$, $q=1$ and $q=p/2<1$, and we see that the new inequality indeed implies the old one. The computation with the new inequality is more manageable.

It is instructive to see what (2) asserts if the vectors $\mathbf{x},\mathbf{y}$ are replaced by real numbers x,y. For simplicity, let us re-scale so that $x=1$, and suppose that y is very small. Then the l.h.s. becomes $\left(\frac{(1+y)^p+(1-y)^p}{2}\right)^{2/p}$, and a Taylor expansion of this gives $(1+p(p-1)y^2/2+O(y^3))^{2/p} = 1+(p-1)y^2+O(y^3)$, while the r.h.s. equals $1+(p-1)y^2$. So both sides agree up to the quadratic term, and in particular, we see that the coefficient $p-1$ in (2) cannot be improved. (This kind of argument doesn't show a similar kind of tighthess for the previous inequality (1), though, but as is discussed in [1], the apparently stronger inequality (2) can be deduced from the validity of the seemingly weaker (1) for all L_p spaces.)

The basic idea of the proof of (2) is this: With $\mathbf{x}$ and $\mathbf{y}$ fixed, we introduce an auxiliary real parameter $t\in[0,1]$, and we consider the functions $L(t)$ and $R(t)$ obtained by substituting $t\mathbf{y}$ for $\mathbf{y}$ in the left-hand and right-hand sides of (2), respectively. That is,

$$\begin{aligned} L(t) &:= \left(\frac{\|\mathbf{x}+t\mathbf{y}\|_p^p+\|\mathbf{x}-t\mathbf{y}\|_p^p}{2}\right)^{2/p} \\ R(t) &:= \|\mathbf{x}\|_p^2+(p-1)t^2\|\mathbf{y}\|_p^2. \end{aligned}$$

Evidently $L(0) = R(0) = \|\mathbf{x}\|_p^2$. We would like to verify that the first derivatives $L'(t)$ and $R'(t)$ both vanish at $t = 0$ (this is easy), and that for the second derivatives we have $L''(t) \geq R''(t)$ for all $t \in [0, 1]$, which will imply $L(1) \geq R(1)$ by double integration.

We have $R'(t) = 2(p-1)t\|\mathbf{y}\|_p^2$ (so $L(0) = 0$) and $R''(t) = 2(p-1)\|\mathbf{y}\|_p^2$.

For dealing with $L(t)$, it is convenient to write $f(t) := (\|\mathbf{x} + t\mathbf{y}\|_p^p + \|\mathbf{x} - t\mathbf{y}\|_p^p)/2$. Then

$$\begin{aligned} L'(t) &= \frac{2}{p} f(t)^{\frac{2}{p}-1} f'(t) \\ &= \frac{2}{p} f(t)^{\frac{2}{p}-1} \frac{p}{2} \sum_i \left(|x_i + ty_i|^{p-1} \mathrm{sgn}(x_i + ty_i) y_i - |x_i - ty_i|^{p-1} \mathrm{sgn}(x_i - ty_i) y_i \right) \end{aligned}$$

(we note that the function $z \mapsto |z|^p$ has a continuous first derivative, namely, $p|z|^{p-1}\mathrm{sgn}(z)$, provided that $p > 1$). The above formula for $L'(t)$ shows $L'(0) = 0$.

For the second derivative we have to be careful, since the graph of the function $z \mapsto |z|^{p-1}$ has a sharp corner at $z = 0$, and thus the function isn't differentiable there for our range of p. We thus proceed with the calculation of $L''(t)$ only for those t with $x_i \pm ty_i \neq 0$ for all i, which excludes finitely many values. Then

$$\begin{aligned} L''(t) &= \frac{2}{p} \left(\frac{2}{p} - 1 \right) f(t)^{\frac{2}{p}-2} f'(t)^2 + \frac{2}{p} f(t)^{\frac{2}{p}-1} f''(t) \\ &\geq \frac{2}{p} f(t)^{\frac{2}{p}-1} f''(t) \\ &= f(t)^{\frac{2}{p}-1} (p-1) \left(\sum_i |x_i + ty_i|^{p-2} y_i^2 + \sum_i |x_i - ty_i|^{p-2} y_i^2 \right). \end{aligned}$$

Next, we would like to bound the sums in the last formula using $\|\mathbf{x}\|_p$ and $\|\mathbf{y}\|_p$. We use the so-called *reverse Hölder inequality*, which asserts, for nonnegative a_i's and strictly positive b_i's, $\sum_i a_i b_i \geq \left(\sum_i a_i^r\right)^{1/r} \left(\sum_i b_i^s\right)^{1/s}$, where $0 < r < 1$ and $\frac{1}{s} = 1 - \frac{1}{r} < 0$. This inequality is not hard to derive from the "usual" Hölder inequality $\sum_i a_i b_i \leq \|\mathbf{a}\|_p \|\mathbf{b}\|_q$, $1 < p < \infty$, $\frac{1}{p} + \frac{1}{q} = 1$. In our case we use the reverse Hölder inequality with $r = p/2$, $s = p/(p-2)$, $a_i = y_i^2$, and $b_i = |x_i + ty_i|^{p-2}$ or $b_i = |x_i - ty_i|^{p-2}$, and we arrive at

$$L''(t) \geq (p-1) f(t)^{\frac{2}{p}-1} \|\mathbf{y}\|_p^2 \left(\|\mathbf{x} + t\mathbf{y}\|_p^p + \|\mathbf{x} - t\mathbf{y}\|_p^p \right) = 2(p-1)\|\mathbf{y}\|_p^2.$$

We have thus proved $L''(t) \geq R''(t)$ for all but finitely many t. The function $L'(t) - R'(t)$ is thus continuous in $(0, 1)$ and nondecreasing on each of the open intervals between the excluded values of t (by the Mean Value Theorem), and so $L'(t) \geq R'(t)$ for all t. The desired conclusion $L(1) \geq R(1)$ follows, again by the Mean Value Theorem.

References

[1] K. Ball, E. A. Carlen, and E. H. Lieb. Sharp uniform convexity and smoothness inequalities for trace norms. *Invent. Math.* 115,1(1994) 463–482.