Analysis and Applications, Vol. 11, No. 6 (2013) 1350026 (5 pages)

DOI: 10.1142/S0219530513500267

AN INEQUALITY INVOLVING THE ℓ_1, ℓ_2, AND ℓ_∞ NORMS

ARYEH KONTOROVICH

Department of Computer Science
Ben-Gurion University
Beer Sheva, Israel
karyeh@cs.bgu.ac.il

Received 25 September 2012
Revised 12 December 2012
Published 11 October 2013

For $\boldsymbol{x} \in [0,1]^n$ with $\|\boldsymbol{x}\|_1 = 1$ and $\boldsymbol{y} \in [1,\infty)^n$, we prove that

$$\frac{\|\boldsymbol{xy}\|_\infty}{\|\boldsymbol{xy}\|_2} \leq \|\boldsymbol{x}\|_\infty \|\boldsymbol{y}\|_1 \frac{\|\boldsymbol{y}\|_\infty}{\|\boldsymbol{y}\|_2},$$

where $\boldsymbol{xy} \in \mathbb{R}^n$ is the vector with components $x_i y_i$. This bound does not seem to easily follow from known inequalities, and the proof technique may be of independent interest.

Keywords: Vector; norm; inequalities.

Mathematics Subject Classification 2010: 26D20, 47A30

1. Statement of Main Result

Our objective[a] is to prove

Theorem 1.1. *For $n \in \mathbb{N}$, $\boldsymbol{x} \in [0,1]^n$ with $\|\boldsymbol{x}\|_1 = 1$ and $\boldsymbol{y} \in [1,\infty)^n$, we have*

$$\frac{\|\boldsymbol{xy}\|_\infty}{\|\boldsymbol{xy}\|_2} \leq \|\boldsymbol{x}\|_\infty \|\boldsymbol{y}\|_1 \frac{\|\boldsymbol{y}\|_\infty}{\|\boldsymbol{y}\|_2}.$$

This bound does not seem to easily follow from known inequalities.

2. Notation

We use the conventions $[n] = \{1,\ldots,n\}$ and $\mathbb{N} = \{1,2,\ldots\}$ and write $\mathbf{1}_{\{\cdot\}}$ to denote the 0-1 truth value of the subscripted predicate. For $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$, we denote

[a] This problem arose in the course of trying to efficiently learn random finite-state automata. Although that approach had to be abandoned [1], perhaps our result will still be of some interest — if only as a mathematical curiosity.

their componentwise product by $\boldsymbol{xy} \in \mathbb{R}^n$. The ℓ_p norm of a vector $\boldsymbol{x} \in \mathbb{R}^n$ is given by

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

for $1 \le p < \infty$ and $\|\boldsymbol{x}\|_\infty = \max_{i\in[n]} |x_i|$. The $(n-1)$-dimensional simplex Δ_{n-1} is defined by

$$\Delta_{n-1} = \{\mathbf{z} \in [0,1]^n : \|\mathbf{z}\|_1 = 1\}.$$

Denote the all 1s vector by $\bar{1} = (1, 1, \ldots, 1)$.

3. Proof of Theorem 1.1

As $\|\boldsymbol{xy}\|_\infty \le \|\boldsymbol{x}\|_\infty \|\boldsymbol{y}\|_\infty$, we may break up our analysis into two cases: (I) $\|\boldsymbol{xy}\|_\infty = \|\boldsymbol{x}\|_\infty \|\boldsymbol{y}\|_\infty$ and (II) $\|\boldsymbol{xy}\|_\infty < \|\boldsymbol{x}\|_\infty \|\boldsymbol{y}\|_\infty$.

Case (I): $\|\boldsymbol{xy}\|_\infty = \|\boldsymbol{x}\|_\infty \|\boldsymbol{y}\|_\infty$.

Proving the theorem for this case amounts to showing that

$$\|\boldsymbol{y}\|_2 \le \|\boldsymbol{y}\|_1 \|\boldsymbol{xy}\|_2 \tag{3.1}$$

whenever $\boldsymbol{x} \in \Delta_{n-1}, y \in [1, \infty)^n$ and

$$\|\boldsymbol{xy}\|_\infty = \|\boldsymbol{x}\|_\infty \|\boldsymbol{y}\|_\infty. \tag{3.2}$$

Assume without loss of generality that $y_1 \ge y_2 \ge \cdots \ge y_n$. Also, there is no loss of generality in assuming that $x_1 = \|\boldsymbol{x}\|_\infty$. We will show that (3.1) holds under the condition

$$x_1 \ge n^{-1}, \tag{3.3}$$

which certainly holds whenever (3.2) does. Let $M = x_1$ be a parameter and let us minimize the quadratic form $\sum_{i=2}^n x_i^2 y_i^2$ over $x_2, \ldots, x_n$, subject to $x_i \ge 0, i = 2, \ldots, n$ and $\sum_{2=1}^n x_i = 1 - x_1$. The following Lagrangian incorporates the objective function and the equality constraint (the inequality constraint will be satisfied automatically):

$$\mathcal{L} = \frac{1}{2}\sum_{i=2}^{n} x_i^2 y_i^2 - \lambda \sum_{i=2}^{n} x_i.$$

Setting $\partial\mathcal{L}/\partial x_i = 0$, we get

$$x_i = \lambda / y_i^2$$

for $i = 2, \ldots, n$, while the normalization constraint yields

$$\lambda = \frac{1 - M}{\sum_{i=2}^{n} y_i^{-2}} = (1 - M)W,$$

where $W^{-1} = \sum_{i=2}^{n} y_i^{-2}$. Thus,

$$\min\{\|\boldsymbol{xy}\|_2^2 : \boldsymbol{x} \in \Delta_{n-1}, x_1 = M\} = M^2 y_1^2 + (1-M)^2 W.$$

Proving (3.1) is now reduced to showing that

$$\|\boldsymbol{y}\|_2^2 \le \|\boldsymbol{y}\|_1^2 (M^2 y_1^2 + (1-M)^2 W).$$

Minimizing the quadratic polynomial $f(M) = M^2 y_1^2 + (1-M)^2 W$ with respect to M and recalling condition (3.3), we may take

$$M = \max\left\{\frac{W^2}{W^2 + y_1^2}, \frac{1}{n}\right\}.$$

However, if $W^2/(W^2 + y_1^2) \ge n^{-1}$ then we have $W^2 + y_1^2 \ge nW^2$, which means that

$$\frac{W^2}{W^2 + y_1^2} \le \frac{1}{n},$$

and thus we must have $M = n^{-1}$.

At this point, (3.1) has been reduced to showing that

$$\|\boldsymbol{y}\|_2^2 \le \|\boldsymbol{y}\|_1^2 (n^{-2} y_1^2 + (1-n^{-1})^2 W) = \|\boldsymbol{y}\|_1^2 \left(\frac{y_1^2}{n^2} + \frac{(1-n^{-1})^2}{\sum_{i=2}^{n} y_i^{-2}} \right).$$

Actually, since $y_i \ge 1$ implies

$$\sum_{i=2}^{n} y_i^{-2} \le n - 1,$$

it suffices to prove the stronger claim:

$$\|\boldsymbol{y}\|_2^2 \le \|\boldsymbol{y}\|_1^2 \left(\frac{y_1^2}{n^2} + \frac{(1-n^{-1})^2}{n-1} \right) = \|\boldsymbol{y}\|_1^2 \left(\frac{y_1^2 + n - 1}{n^2} \right).$$

The latter claim is equivalent to the assertion that

$$\|\boldsymbol{y}\|_2^2 \le \|\boldsymbol{y}\|_1^2 \left(\frac{\|\boldsymbol{y}\|_\infty^2 + n - 1}{n^2} \right) \tag{3.4}$$

for all $\boldsymbol{y} \in [1, \infty)^n$. Indeed, we claim that for $\boldsymbol{y} \in [1, L]^n$, the ratio $\|\boldsymbol{y}\|_2/\|\boldsymbol{y}\|_1$ is maximized when $\boldsymbol{y}$ is of the form $\boldsymbol{y} = (L, 1, 1, \ldots, 1)$. To see this, define the function $f : [0, L]^n \to \mathbb{R}$ by $f(\boldsymbol{y}) = (\|\boldsymbol{y}\|_2/\|\boldsymbol{y}\|_1)^2$. It is a simple consequence of Hölder's inequality that f is maximized over $[0, L]^n$ at the standard basis elements $\{\boldsymbol{e}_i\}$, such as $\boldsymbol{e}_1 = (1, 0, 0, \ldots, 0)$. Since $\partial f/\partial y_i|_{\boldsymbol{y}=\boldsymbol{e}_1} < 0$ for $i \ne 1$, it follows that there exist some $\epsilon > 0$ such that $f(\boldsymbol{y})$ is maximized over $[\epsilon, L]^n$ at $(L, \epsilon, \epsilon, \ldots, \epsilon)$. However, since $f(\boldsymbol{y}) = f(\boldsymbol{y}/\min_{i \in [n]} y_i)$, it follows that $f(\boldsymbol{y})$ is maximized over $[1, L/\epsilon]^n$ at $(L/\epsilon, 1, 1, \ldots, 1)$. Since L was arbitrary, the claim follows.

Substituting the values $\|\boldsymbol{y}\|_2^2 = L^2 + n - 1$ and $\|\boldsymbol{y}\|_1^2 = (L + n - 1)^2$ into (3.4) yields (after simplifying) the equivalent inequality $n^2 \le (L + n - 1)^2$, which clearly holds when $L \ge 1$.

Case (II): $\|\boldsymbol{xy}\|_\infty < \|\boldsymbol{x}\|_\infty \|\boldsymbol{y}\|_\infty$.

Proving the theorem for this case is equivalent to showing that the function

$$F(\boldsymbol{x}, \boldsymbol{y}) = \frac{\|\boldsymbol{y}\|_2 \|\boldsymbol{xy}\|_\infty}{\|\boldsymbol{y}\|_1 \|\boldsymbol{y}\|_\infty \|\boldsymbol{x}\|_\infty \|\boldsymbol{xy}\|_2}$$

is upper-bounded by 1 on $\Delta_{n-1} \times [1, \infty)^n$.

Fix an $\boldsymbol{x} \in \Delta_{n-1}$ and pick any $\boldsymbol{y} \in [1, \infty)^n$. Define the quantities

$$\begin{aligned} I &= \{i \in [n] : y_i = \|\boldsymbol{y}\|_\infty\}, \\ w &= \max\{y_i : i \notin I\}, \\ h &= \|\boldsymbol{y}\|_\infty - w, \end{aligned} \tag{3.5}$$

(in words, I is the set of indices of $\boldsymbol{y}$'s maximal coordinates, w is the second largest entry in $\boldsymbol{y}$, and h is the gap between w and $\|\boldsymbol{y}\|_\infty$). Thus, we may write $\boldsymbol{y}$ as $\boldsymbol{y} = \tilde{\boldsymbol{y}}(h)$, where

$$\begin{aligned} \tilde{\boldsymbol{y}}(h)_i &= \min\{y_i, w\} + \mathbf{1}_{\{i \in I\}} h \\ &= \begin{cases} w + h, & i \in I \\ y_i, & i \notin I \end{cases}. \end{aligned}$$

We claim that, for a fixed $\boldsymbol{x} \in \Delta_{n-1}$, the function

$$f(h) = F(\boldsymbol{x}, \tilde{\boldsymbol{y}}(h))^2$$

is decreasing for $h \geq 0$. Once this claim is proven, we may let $\boldsymbol{y}' = \tilde{\boldsymbol{y}}(0)$, compute I', w', h' from $\boldsymbol{y}'$ as in (3.5) and repeat this process, obtaining a sequence of $\boldsymbol{y}', \boldsymbol{y}'', \boldsymbol{y}''', \ldots$, until either case (I) becomes applicable or the vector $\bar{1} = (1, 1, \ldots, 1)$ is reached. Since by the claim $f(0) \leq f(h)$, we have

$$F(\boldsymbol{x}, \boldsymbol{y}) \leq F(\boldsymbol{x}, \boldsymbol{y}') \leq F(\boldsymbol{x}, \boldsymbol{y}'') \leq \cdots.$$

Note also that $F(\boldsymbol{x}, \bar{1}) = n^{-1/2} \|\boldsymbol{x}\|_2^{-1} \leq 1$. Thus, showing that $f(h)$ is decreasing is all that remains to finish off case (II).

Let us write

$$\begin{aligned} \|\tilde{\boldsymbol{y}}\|_2^2 &= |I|(w+h)^2 + \sum_{i \notin I} y_i^2 = k(w+h)^2 + \zeta \\ \|\tilde{\boldsymbol{y}}\|_\infty^2 &= (w+h)^2 \\ \|\tilde{\boldsymbol{y}}\|_1^2 &= \left(|I|(w+h) + \sum_{i \notin I} y_i \right)^2 = (k(w+h) + \theta)^2 \\ \|\boldsymbol{x}\tilde{\boldsymbol{y}}\|_2^2 &= \sum_{i \in I} x_i^2 (w+h)^2 + \sum_{i \notin I} x_i^2 y_i^2 = \eta(w+h)^2 + \nu, \end{aligned}$$

where $k, \zeta, \theta, \eta, \nu$ are defined in the obvious way.

Consider a fixed $(\boldsymbol{x}, \boldsymbol{y}) \in \Delta_{n-1} \times [1, \infty)^n$ and let I, w, h be as in (3.5). Putting $j = \operatorname{argmax}\{x_i y_i : i \in [n]\}$, we have $\|\boldsymbol{x}\tilde{\boldsymbol{y}}\|_\infty = x_j y_j$. Hence,

$$\begin{aligned} f(h) = F(\boldsymbol{x}, \tilde{\boldsymbol{y}}(h))^2 &= \frac{\|\tilde{\boldsymbol{y}}\|_2^2 \|\boldsymbol{x}\tilde{\boldsymbol{y}}\|_\infty^2}{\|\tilde{\boldsymbol{y}}\|_1^2 \|\tilde{\boldsymbol{y}}\|_\infty^2 \|\boldsymbol{x}\|_\infty^2 \|\boldsymbol{x}\tilde{\boldsymbol{y}}\|_2^2} \\ &= x_j^2 y_j^2 \|\boldsymbol{x}\|_\infty^{-2} \frac{k(w+h)^2 + \zeta}{(k(w+h)+\theta)^2 (w+h)^2 (\eta(w+h)^2 + \nu)} \\ &= x_j^2 y_j^2 \|\boldsymbol{x}\|_\infty^{-2} \left(k + \frac{\zeta}{(w+h)^2} \right) \\ &\quad \cdot \frac{1}{(k(w+h)+\theta)^2} \cdot \frac{1}{\eta(w+h)^2 + \nu}, \end{aligned}$$

which is obviously decreasing in h. This proves case (II).

Acknowledgments

I thank the two anonymous referees for helpful comments and corrections on the manuscript.

Reference

[1] D. Angluin, D. Eisenstat, L. (Aryeh) Kontorovich and L. Reyzin, Lower bounds on learning random structures with statistical queries, in *ALT*, Lecture Notes in Computer Science, Vol. 6331 (Springer, 2010), pp. 194–208.