MATRIX $p$-NORMS ARE NP-HARD TO APPROXIMATE IF $p$ not equal to 1, 2, infinity

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.
MATRIX $p$-NORMS ARE NP-HARD TO APPROXIMATE IF $p \neq 1, 2, \infty$*

JULIEN M. HENDRICKX† AND ALEX OLSHEVSKY‡

Abstract. We show that, for any rational $p \in [1, \infty)$ except $p \neq 1, 2$, unless $P = NP$, there is no polynomial time algorithm which approximates the matrix $p$-norm to arbitrary relative precision. We also show that, for any rational $p \in [1, \infty)$ including $p = 1, 2$, unless $P = NP$, there is no polynomial-time algorithm which approximates the $\infty, p$ mixed norm to some fixed relative precision.

Key words. matrix norms, complexity, NP-hardness

AMS subject classifications. 68Q17, 15A60

DOI. 10.1137/09076773X

1. Introduction. The $p$-norm of a matrix $A$ is defined as

$$||A||_p = \max_{||x||_p=1} ||Ax||_p.$$  

We consider the problem of computing the matrix $p$-norm to relative error $\epsilon$, defined as follows: given the inputs (i) a matrix $A \in \mathbb{R}^{n \times n}$ with rational entries and (ii) an error tolerance $\epsilon$ which is a positive rational number, output a rational number $r$ satisfying

$$|r - ||A||_p| \leq \epsilon ||A||_p.$$  

We will use the standard bit model of computation. When $p = \infty$ or $p = 1$, the $p$-matrix norm is the largest of the row/column sums and thus may be easily computed exactly. When $p = 2$, this problem reduces to computing an eigenvalue of $A^T A$ and thus can be solved in polynomial time in $n, \log \frac{1}{\epsilon}$ and the bit size of the entries of $A$. Our main result suggests that the case of $p \notin \{1, 2, \infty\}$ may be different.

**Theorem 1.1.** For any rational $p \in [1, \infty)$ except $p = 1, 2$, unless $P = NP$, there is no algorithm which computes the $p$-norm of a matrix with entries in $\{-1, 0, 1\}$ to relative error $\epsilon$ with running time polynomial in $n, \frac{1}{\epsilon}$.

On the way to our result, we also slightly improve the NP-hardness result for the mixed norm $||A||_{\infty, p} = \max_{||x||_p \leq 1} ||Ax||_p$ from [5]. Specifically, we show that, for every rational $p \geq 1$, there exists an error tolerance $\epsilon(p)$ such that, unless $P = NP$, there is no polynomial time algorithm approximating $||A||_{\infty, p}$ with a relative error smaller than $\epsilon(p)$.

*Received by the editors August 11, 2009; accepted for publication (in revised form) by D. Boley July 23, 2010; published electronically October 28, 2010. This research was supported by the National Science Foundation under grant ECCS-0701623. The first author holds postdoctoral fellowships from the F.R.S.-FNRS (Belgian National Fund for Scientific Research) and the Belgian American Education Foundation and is on leave from the Department of Mathematical Engineering of the Université Catholique de Louvain, Louvain-la-Neuve, Belgium.

http://www.siam.org/journals/simax/31-5/76773.html

†Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA. Current address: Institute of Information and Communication Technologies, Electronics and Applied Mathematics, Université catholique de Louvain, 4 avenue George Lemaître, B-1348 Louvain-la-Neuve, Belgium (julien.hendrickx@uclouvain.be).

‡Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA. Current address: Department of Mechanical and Aerospace Engineering, Princeton University, Princeton, NJ (aolshevs@princeton.edu).
1.1. Previous work. When \( p \) is an integer, computing the matrix norm can be recast as solving a polynomial optimization problem. These are known to be hard to solve in general [3]; however, because the matrix norm problem has a special structure, one cannot immediately rule out the possibility of a polynomial time solution. A few hardness results are available in the literature for mixed matrix norms \( \| A \|_{p,q} = \max_{\| x \|_p \leq 1} \| A x \|_q \). Rohn has shown in [4] that computing the \( \| A \|_{\infty,1} \) norm is NP-hard. In her thesis, Steinberg [5] proved more generally that computing \( \| A \|_{p,q} \) is NP-hard when \( 1 \leq q < p \leq \infty \). We refer the reader to [5] for a discussion of applications of the mixed matrix norm problems to robust optimization.

It is conjectured in [5] that there are only three cases in which mixed norms are computable in polynomial time: First, \( p = 1 \), and \( q \) is any rational number larger than or equal to 1. Second, \( q = \infty \), and \( p \) is any rational number larger than or equal to 1. Third, \( p = q = 2 \). Our work makes progress on this question by settling the “diagonal” case of \( p = q \); however, the case of \( p < q \), as far as the authors are aware, is open.

1.2. Outline. We begin in section 2 by providing a proof of the NP-hardness of approximating the mixed norm \( \| \cdot \|_{\infty,p} \) within some fixed relative error for any rational \( p \geq 1 \). The proof may be summarized as follows: observe that, for any matrix \( M \),

\[
\max_{\| x \|_p = 1} \| M x \|_p
\]

is always attained at one of the \( 2^n \) points of \( \{-1,1\}^n \). So by appropriately choosing \( M \), one can encode an NP-hard problem of maximization over the latter set. This argument will prove that computing the \( \| \cdot \|_{\infty,p} \) norm is NP-hard.

Next, in section 3, we exhibit a class of matrices \( A \) such that \( \max_{\| x \|_p = 1} \| A x \|_p \) is attained at each of the \( 2^n \) points of \( \{-1,1\}^n \) (up to scaling) and nowhere else. These two elements are combined in section 4 to prove Theorem 1.1. More precisely, we define the matrix \( Z = (M^T \alpha A^T)^T \), where we will pick \( \alpha \) to be a large number depending on \( n, p \) ensuring that the maximum of \( \| Z x \|_p / \| x \|_p \) occurs very close to vectors \( x \in \{-1,1\}^n \). As mentioned several sentences ago, the value of \( \| A x \|_p \) is the same for every vector \( x \in \{-1,1\}^n \); as a result, the maximum of \( \| Z x \|_p / \| x \|_p \) is determined by the maximum of \( \| M x \|_p \) on \( \{-1,1\}^n \), which is proved in section 2 to be hard to compute. We conclude with some remarks on the proof in section 5.

2. The \( \| \cdot \|_{\infty,p} \) norm. We now describe a simple construction which relates the \( \infty,p \) norm to the maximum cut in a graph.

Suppose \( G = (\{1, \ldots, n\}, E) \) is an undirected, connected graph. We will use \( M(G) \) to denote the edge-vertex incidence matrix of \( G \); that is, \( M(G) \in \mathbb{R}^{E \times n} \). We will think of columns of \( M(G) \) as corresponding to nodes of \( G \) and of rows of \( M(G) \) as corresponding to the edges of \( G \). The entries of \( M(G) \) are as follows: orient the edges of \( G \) arbitrarily, and let the \( i \)-th row of \( M(G) \) have \(+1\) in the column corresponding to the origin of the \( i \)-th edge, \(-1\) in the column corresponding to the endpoint of the \( i \)-th edge, and \( 0 \) at all other columns.

Given any partition of \( \{1, \ldots, n\} = S \cup S^c \), we define \( \text{cut}(G,S) \) to be the number of edges with exactly one endpoint in \( S \). Furthermore, we define \( \max \text{cut}(G) = \max_{S \subset \{1, \ldots, n\}} \text{cut}(G,S) \). The indicator vector of a cut \((S, S^c)\) is the vector \( x \) with \( x_i = 1 \) when \( i \in S \) and \( x_i = -1 \) when \( i \in S^c \). We will use \( \text{cut}(x) \) for vectors \( x \in \{-1,1\}^n \) to denote the value of the cut whose indicator vector is \( x \).

Proposition 2.1. For any \( p \geq 1 \),

\[
\max_{\| x \|_\infty \leq 1} \| M(G) x \|_p = 2 \text{maxcut}(G)^{1/p}.
\]
Proof. Observe that $\|M(G)x\|_p$ is a convex function of $x$, so that the maximum is achieved at the extreme points of the set $\|x\|_\infty \leq 1$, i.e., vectors $x$ satisfying $x_i = \pm 1$. Suppose we are given such a vector $x$; define $S = \{i \mid x_i = 1\}$. Clearly, $\|M(G)x\|_p^p = 2^p \text{cut}(G, S)$. From this the proposition immediately follows.

Next, we introduce an error term into this proposition. Define $f^*$ to be the optimal value $f^* = \max_{\|x\|_\infty \leq 1} \|M(G)x\|_p$; the above proposition implies that $(f^*/2)^p = \text{maxcut}(G)$. We want to argue that if $f_{\text{approx}}$ is close enough to $f^*$, then $(f_{\text{approx}}/2)^p$ is close to $\text{maxcut}(G)$.

**Proposition 2.2.** If $p \geq 1$, $|f^* - f_{\text{approx}}| < \epsilon f^*$ with $\epsilon < 1$, then

$$\left| \left( \frac{f_{\text{approx}}}{2} \right)^p - \text{maxcut}(G) \right| \leq \frac{1}{2} |f^* - f_{\text{approx}}| p \max \left( \frac{f^*}{2}, \frac{f_{\text{approx}}}{2} \right)^{p-1}.$$

It follows from $\epsilon < 1$ that $f_{\text{approx}} \leq 2f^*$. We have therefore

$$\left| \left( \frac{f_{\text{approx}}}{2} \right)^p - \text{maxcut}(G) \right| \leq \frac{1}{2} |f^* - f_{\text{approx}}| \cdot p \cdot (f^*)^{p-1} \leq \frac{\epsilon}{2} p (f^*)^p,$$

where we have used the assumption that $|f^* - f_{\text{approx}}| \leq \epsilon f^*$. The result follows then from $\text{maxcut}(G) = (f^*/2)^p$.

We now put together the previous two propositions to prove that approximating the $\| \cdot \|_{\infty, p}$ norm within some fixed relative error is NP-hard.

**Theorem 2.3.** For any rational $p \geq 1$ and $\delta > 0$, unless $P = \text{NP}$, there is no algorithm which, given a matrix with entries in $\{-1, 0, 1\}$, computes its $p$-norm to relative error $\epsilon = ((33 + \delta)p^{2p-1})^{-1}$ with running time polynomial in the dimensions of the matrix.

**Proof.** Suppose there was such an algorithm. Call $f^*$ its output on the $|E| \times n$ matrix $M(G)$ for a given connected graph $G$ on $n$ vertices. It follows from Proposition 2.2 that

$$\left| \left( \frac{f_{\text{approx}}}{2} \right)^p - \text{maxcut}(G) \right| \leq \frac{2^{p-1}p}{(33 + \delta)p2^{p-1}} \text{maxcut}(G) = \frac{1}{33 + \delta} \text{maxcut}(G).$$

Observing that

$$\frac{32 + \delta}{34 + \delta} \text{maxcut}(G) = \frac{33 + \delta}{34 + \delta} \left( \text{maxcut}(G) - \frac{1}{33 + \delta} \text{maxcut}(G) \right),$$

the former inequality implies

$$\frac{32 + \delta}{34 + \delta} \text{maxcut}(G) \leq \frac{33 + \delta}{34 + \delta} \left( \frac{f_{\text{approx}}}{2} \right)^p \leq \text{maxcut}(G).$$

Since $p$ is rational, one can compute in polynomial time a lower bound $V$ for $\frac{33+\delta}{34+\delta} (f_{\text{approx}}/2)^p$ sufficiently accurate so that $V > \frac{32+\delta/2}{34+\delta} \text{maxcut}(G) > \frac{\epsilon}{16} \text{maxcut}(G)$. 

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.
However, it has been established in [2] that, unless $P = NP$, for any $\delta' > 0$, there is no algorithm producing a quantity $V$ in polynomial time in $n$ such that
\[
\left(\frac{16}{17} + \delta'\right) \maxcut(G) \leq V \leq \maxcut(G). \quad \Box
\]

**Remark.** Observe that the matrix $M(G)$ is not square. If one desires to prove

hardness of computing the zeros to every row of $G$ and is square, and its dimensions are at most $n^2 \times n^2$.

3. A discrete set of exponential size. Let us now fix $n$ and a rational $p > 2$. We denote by $X$ the set $\{-1,1\}^n$ and use $S(a, r) = \{x \in \mathbb{R}^n \mid \|x - a\|_p = r\}$ to stand for the sphere of radius $r$ around $a$ in the $p$-norm. We consider the following matrix in $\mathbb{R}^{2n \times n}$:

\[
A = \begin{pmatrix}
1 & -1 \\
1 & 1 \\
1 & 1 \\
& \ddots & \ddots \\
& \ddots & \ddots \\
1 & -1 \\
1 & 1 \\
\end{pmatrix}
\]

and show that the maximum of $\|Ax\|_p$ for $x \in S(0, n^{1/p})$ is attained at the $2^n$ vectors in $X$ and at no other points. For this, we will need the following lemma.

**Lemma 3.1.** For any real numbers $x$, $y$ and $p \geq 2$

\[
|x + y|^p + |x - y|^p \leq 2^{p-1} (|x|^p + |y|^p).
\]

In fact, $|x + y|^p + |x - y|^p$ is upper bounded by

\[
2^{p-1} (|x|^p + |y|^p) - \frac{(|x| - |y|)^2}{4} \left(p(p-1)||x||^p - 2||x||^{p-2} - 2||y||^{p-2}\right),
\]

where the last term on the right is always nonnegative.

**Proof.** By symmetry we can assume that $x \geq y \geq 0$. In that case, we need to prove

\[
(x + y)^p + (x - y)^p \leq 2^{p-1}(x^p + y^p) - \frac{(x - y)^2}{4} \left(p(p-1)(x + y)^{p-2} - 2(x - y)^{p-2}\right).
\]

Divide both sides by $(x + y)^p$, and change the variables to $z = (x - y)/(x + y)$:

\[
1 + z^p \leq \frac{(1 + z)^p + (1 - z)^p}{2} - \frac{p(p-1)}{4}z^2 - \frac{1}{2}z^p.
\]

The original inequality holds if this inequality holds for $z \in [0, 1]$. Let’s simplify:

\[
2 + z^p \leq (1 + z)^p + (1 - z)^p - \frac{p(p-1)}{2}z^2.
\]
Observe that we have equality when \( z = 0 \), so it suffices to show that the right-hand side grows faster than the left-hand side, namely,
\[
z^{p-1} \leq (1 + z)^{p-1} - (1 - z)^{p-1} - (p - 1)z,
\]
and this follows from
\[
(1 + z)^{p-1} \geq 1 + (p - 1)z \geq (1 - z)^{p-1} + z^{p-1} + (p - 1)z,
\]
where we have used the convexity of \( f(a) = a^{p-1} \).

Now we prove that every vector of \( X \) optimizes \( \|Ax\|_p/\|x\|_p \) or, equivalently, optimizes \( \|Ax\|_p^p \) over the sphere \( S(0, n^{1/p}) \).

**Lemma 3.2.** For any \( p \geq 2 \), the supremum of \( \|Ax\|_p^p \) over \( S(0, n^{1/p}) \) is achieved by any vector in \( X \).

**Proof.** Observe that \( \|Ax\|_p^p = n2^p \) for any \( x \in X \). To prove that this is the largest possible value, we write
\[
\|Ax\|_p^p = \sum_{i=1}^{n} |x_i - x_{i+1}|^p + |x_i + x_{i+1}|^p,
\]
using the convention \( n + 1 = 1 \) for the indices. Lemma 3.1 implies that
\[
|x_i - x_{i+1}|^p + |x_i + x_{i+1}|^p \leq 2^{p-1} (|x_i|^p + |x_{i+1}|^p).
\]
By applying this inequality to each term of (3.1) and by using \( \|x\|_p = n \), we obtain
\[
\|Ax\|_p^p \leq \sum_{i=1}^{n} 2^{p-1} (|x_i|^p + |x_{i+1}|^p) = 2^p \sum_{i=1}^{n} |x_i|^p = 2^pn. \]

Next we refine the previous lemma by including a bound on how fast \( \|Ax\|_p^p \) decreases as we move a little bit away from the set \( X \) while staying on \( S(0, n^{1/p}) \).

**Lemma 3.3.** Let \( p \geq 2, c \in (0, 1/2] \), and suppose \( y \in S(0, n^{1/p}) \) has the property that
\[
\min_{x \in X} \|y - x\|_\infty \geq c.
\]

Then
\[
\|Ay\|_p^p \leq n2^p - \frac{3(p - 2)}{2pn^2} c^2.
\]

**Proof.** We proceed as before in the proof of Lemma 3.2 until the time comes to apply Lemma 3.1, when we include the error term which we had previously ignored:
\[
\|Ay\|_p^p \leq n2^p - \frac{1}{4} \sum_{i=1}^{n} (|y_i| - |y_{i+1}|)^2 \left( p(p - 1) \|y_i| + |y_{i+1}|\|^{p-2} - 2 \|y_i| - |y_{i+1}|\|^{p-2} \right),
\]
Note that on the right-hand side, we are subtracting a sum of nonnegative terms. The upper bound will still hold if we subtract only one of these terms, so we conclude that, for each \( k \),
\[
\|Ay\|_p^p \leq n2^p - \frac{1}{4} (|y_k| - |y_{k+1}|)^2 \left( p(p - 1) \|y_k| + |y_{k+1}|\|^{p-2} - 2 \|y_i| - |y_{i+1}|\|^{p-2} \right).
\]
By assumption, there is at least one \( y_k \) with \( |y_k| - 1 \geq c \). Suppose first that \( |y_k| > 1 \). Then we have \( |y_k| > 1 + c \), and there must be a \( y_j \) with \( |y_j| < 1 \), for otherwise \( y \) would not be in \( S(0,n^{1/p}) \). Similarly, if \( |y_k| < 1 \), then \( |y_k| < 1 - c \) and there is a \( j \) for which \( |y_j| > 1 \). In both cases, this implies the existence of an index \( m \) with \( |y_m| \) and \( |y_{m+1}| \) differing by at least \( c/n \) and such that at least one of \( |y_m| \) and \( |y_{m+1}| \) is larger than or equal to \( 1 - c \). Therefore,

\[
||Ay||^p \leq n2^p - \frac{1}{4} \frac{c^2}{n^2} \left[ p(p-1) ||y_m| + |y_{m+1}| - 2 ||y_m| - |y_{m+1}| \right].
\]

Now observe that \( ||y_m| - |y_{m+1}| \leq |y_m| + |y_{m+1}| \) and that \( |y_m| + |y_{m+1}| \geq (1-c) \geq 1/2 \) because \( c \in (0,1/2) \). These two inequalities suffice to establish that the term in square brackets is at least \( (1/2)^{p-2}(p(p-1) - 2) \geq (3/2^p)(p-2) \) so that

\[
||Ay||^p \leq n2^p - \frac{3(p-2)}{2^p n^2}. \]

**4. Proof of Theorem 1.1.** We now relate the results of the last two sections to the problem of the \( p \)-norm. For a suitably defined matrix \( Z \) combining \( A \) and \( M(G) \), we want to argue that the optimizer of \( ||Zx||_p/||x||_p \) is very close to satisfying \( \max x \in S(0,n^{1/p}) ||Zx||_p \) achieves its supremum, then

\[
\min_{x \in \Lambda} ||x^* - x||_p \leq \frac{1}{4^p n^{6/32}}.
\]

**Proof.** Suppose the conclusion is false. Then using Lemma 3.3 with \( c = 1/4^p n^6 \), we obtain

\[
||Ax^*||_p \leq n2^p - \frac{3(p-2)}{2^p 4^p n^{14}} = n2^p - \frac{3(p-2)}{32^p n^{14}}.
\]

It follows from Proposition 2.1 that

\[
||Mx^*||_p \leq 2^p \maxcut(G) \leq 2^p n^2
\]

so that

\[
||Zx^*||_p = ||Ax^*||_p + \left( \frac{p-2}{64^p n^8} \right)^p ||Mx^*||_p \leq 2^p - \frac{3(p-2)}{32^p n^{14}} + \frac{2^p(2p-2)n^2}{64^p p^p n^{8p}}.
\]

Observe that the last term in this inequality is smaller than the previous one (in absolute value). Indeed, for \( p > 2 \), we have that \( 3/32^p > (2/64)^p, p-2 > (p-2)/p^p \), and \( 1/n^{14} > n^2/n^{8p} \). We therefore have \( ||Zx^*||_p < 2^p n^2 \). By contrast, let \( x \) be any vector in \( \{-1,1\}^n \). Then \( x \in S(0,n^{1/p}) \) and

\[
||Zx||_p \geq ||Ax||_p \geq 2^p n,
\]
which contradicts the optimality of \( x^\ast \).

Next we seek to translate the fact that the optimizer \( x^\ast \) is close to \( X \) to the fact that the objective value \(||\mathbf{z} x||_p/||x||_p|\) is close to the largest objective value at \( X \).

**Proposition 4.2.** Let \( p > 2 \), \( G \) be a graph on \( n \) vertices, and

\[
\mathbf{z} = \begin{pmatrix} \frac{64pn^8}{p^2 - 2} A \\ M(G) \end{pmatrix}.
\]

If \( x^\ast \) is the vector at which the optimization problem

\[
\max_{x \in S(0,n^{1/p})} ||\mathbf{z} x||_p
\]

achieves its supremum and \( x_\ast \) is the rounded version of \( x^\ast \) in which every component is rounded to the closest of \(-1\) and \(1\), then

\[
\left| ||\mathbf{z} x^\ast||_p^p - ||\mathbf{z} x_\ast||_p^p \right| \leq \frac{1}{n^2}.
\]

**Proof.** Observe that \( x^\ast \) is the same as the extremizer of the corresponding problem with \( \tilde{z} \) instead of \( Z \) so that \( x \) satisfies the conclusion of Proposition 4.1. Consequently every component of \( x^\ast \) is closer to one of \( \pm 1 \) than to the other, and so \( x_\ast \) is well defined. We have,

\[
||\mathbf{z} x^\ast||_p^p - ||\mathbf{z} x_\ast||_p^p = \left( 64 \frac{p}{p - 2} n^8 \right)^p (||Ax^\ast||_p^p - ||Ax_\ast||_p^p) + (||Mx^\ast||_p^p - ||Mx_\ast||_p^p).
\]

This entire quantity is nonnegative since \( x^\ast \) is the maximum of \( ||\mathbf{z} x||_p \) on \( S(0,n^{1/p}) \). Moreover, \( ||Ax^\ast||_p^p - ||Ax_\ast||_p^p \) is nonpositive since, by Proposition 3.2, \( ||Ax||_p \) achieves its maximum over \( S(0,n^{1/p}) \) on all the elements of \( X \). Consequently,

\[
\left| ||\mathbf{z} x^\ast||_p^p - ||\mathbf{z} x_\ast||_p^p \right| \leq ||Mx^\ast||_p^p - ||Mx_\ast||_p^p \leq (||Mx^\ast||_p^p - ||Mx_\ast||_p^p) \max(||Mx^\ast||_p^p, ||Mx_\ast||_p^p)^{p-1}.
\]

We now bound all the terms in the last equation. First

\[
||Mx^\ast||_p^p - ||Mx_\ast||_p^p \leq ||M||_2 ||x^\ast - x_\ast||_2 \leq ||M||_F \sqrt{n} ||x^\ast - x_\ast||_{\infty} = \frac{n \sqrt{n}}{4^p n^6},
\]

where we have used \( ||M(G)||_F = \sqrt{2|E|} < n \) and Proposition 4.1 for the last inequality. Now that we have a bound on the first term in (4.1), we proceed to the last term. It follows from the definition of \( M \) that

\[
||Mx_\ast||_p^p \leq 2^p \cdot \left( \frac{n}{2} \right)^p \leq 2^p n^2.
\]

Next we bound \( ||Mx^\ast||_p^p \). Observe that a particular case of (4.2) is

\[
||Mx^\ast||_p^p < ||Mx_\ast||_p^p + 1.
\]

Moreover, observe that \( ||Mx_\ast||_p > 1 \). (The only way this does not hold is if every entry of \( x_\ast \) is the same, i.e., \(||Mx_\ast||_p = 0\). But then (4.3) implies that \(||Mx^\ast||_p < 1 \), which is
impossible since \( G \) has at least one edge.), So (4.3) implies that \( \|Mx^*\|_p \leq 2\|Mx_1\|_p \), and so

\[
\|Mx^*\|_p^p \leq 4^p n^2.
\]

Thus

\[
\max(\|Mx^*\|_p, \|Mx_1\|_p) \leq 4^p n^2,
\]

and therefore \( \max(\|Mx^*\|_p, \|Mx_1\|_p)^{p-1} \leq 4^p n^2 \). Indeed, this bound is trivially valid if \( \max(\|Mx^*\|_p, \|Mx_1\|_p)^p \leq 1 \) and follows from \( a^{p-1} < a^p \) for \( a \geq 1 \) otherwise. Using this bound and the inequality (4.2), we finally obtain

\[
\|Zx^*\|_p^p - \|Zx_1\|_p^p \leq \frac{n^{1.5}}{4^p n^6} p \cdot 4^p n^2 \leq \frac{1}{n^2}.
\]

Finally let us bring it all together by arguing that if we can approximately compute the \( p \)-norm of \( Z \), we can approximately compute the maximum cut.

**Proposition 4.3.** Let \( p > 2 \). Consider a graph \( G \) on \( n > 2 \) vertices and the matrix

\[
Z = \left( \frac{64}{p-2} n^8 A \right)^p,
\]

and let \( f^* = \|Z\|_p \). If

\[
|f_{\text{approx}} - f^*| \leq \frac{(p-2)p}{132p^p n^{8p+3}}
\]

then

\[
\left| \left( \frac{n}{2p} f_{\text{approx}}^p - n \left( \frac{64pn^8}{p-2} \right)^p \right) - \maxcut(G) \right| \leq \frac{1}{n}.
\]

**Proof.** Observe that \( \frac{n}{2p} f^* = \max_{x \in S(0,n^{1/p})} \|Zx\|_p \). It follows thus from Proposition 4.2 that

\[
\left| n f^* - \max_{x \in X} \|Zx\|_p \right| < \frac{1}{n^2}.
\]

Recall that \( \|Zx\|_p^p = \|Mx\|_p^p + (\frac{64}{p-2} n^8)^p \|Ax\|_p^p \) and that \( \|Ax\|_p^p = n2^p \) for every \( x \in X \). Therefore,

\[
\max_{x \in X} \|Zx\|_p^p = \left( \frac{64pn^8}{p-2} \right)^p n2^p + \max_{x \in X} \|Mx\|_p^p = \left( \frac{64pn^8}{p-2} \right)^p n2^p + 2^p \maxcut(G),
\]

and by combining the last two equations, we have

\[
\left| \left( \frac{n}{2p} f_{\text{approx}}^p - n \left( \frac{64pn^8}{p-2} \right)^p \right) - \maxcut(G) \right| \leq \frac{1}{2pn^2}.
\]

Let us now evaluate the error introduced by the approximation \( f_{\text{approx}} \):

\[
\left| \left( \frac{n}{2p} f_{\text{approx}}^p - n \left( \frac{64pn^8}{p-2} \right)^p \right) - \maxcut(G) \right| \leq \frac{1}{2pn^2} + \frac{n}{2p} |f_{\text{approx}}^p - f^*| \leq \frac{1}{2pn^2} + \frac{n}{2p} |f_{\text{approx}}^p - f^*| p \max(f^*, f_{\text{approx}})^{p-1}.
\]
It remains to bound the last term of this inequality. First we use the fact that \( f^* \geq 1 \) and (4.4) to argue

\[
(4.5) \quad f^{(p-1)} \leq f^p \leq 2^p \left( \frac{64pn^8}{p-2} \right)^p + \frac{2^p}{n} \maxcut(G) + \frac{1}{n^2} \leq 2^p \left( \frac{66pn^8}{p-2} \right)^p,
\]

where we have used \( \maxcut(G) < n^2 \) and \( 1 \leq p/(p-2) \) for the last inequality. By assumption, \(|f_{\text{approx}} - f^*| \leq 1 \), and since \( f^* \geq 1 \),

\[
f_{\text{approx}}^{(p-1)} \leq (2f^*)^{p-1} \leq (2f^*)^p \leq 4^p \left( \frac{66pn^8}{p-2} \right)^p.
\]

Putting it all together and using the bound on \(|f_{\text{approx}} - f^*|\), we obtain (assuming \( n > 1 \))

\[
\left( \frac{n}{2^p f_{\text{approx}}} - n \left( \frac{64pn^8}{p-2} \right)^p \right) - \maxcut(G) \leq \frac{1}{2^p n^2} + \frac{(p-2)p^p n^8 p^{p+3}}{132 p^p n^8 p^{p+3} p} 2^p n^p \left( \frac{66pn^8}{p-2} \right)^p
\]

\[
\leq \frac{1}{2^p n^2} + \frac{1}{n^2}
\]

\[
\leq \frac{1}{n}.
\]

**Proposition 4.4.** Fix a rational \( p \in [1, \infty) \) with \( p \neq 1, 2 \). Unless \( P = NP \), there is no algorithm which, given input \( \epsilon > 0 \) and a matrix \( Z \), computes \( ||Z||_p \) to a relative accuracy \( \epsilon \), in time which is polynomial in \( 1/\epsilon \), the dimensions of \( Z \), and the bit size of the entries of \( Z \).

**Proof.** Suppose first that \( p > 2 \). We show that such an algorithm could be used to build a polynomial time algorithm solving the maximum cut problem. For a graph \( G \) on \( n \) vertices, fix

\[
\epsilon = \left( 132^n \left( \frac{p}{p-2} \right)^{n^8p^3} \right)^{-1} \left( 132^n \left( \frac{p}{p-2} \right)^{n^8} \right)^{-1},
\]

build the matrix \( Z \) as in Proposition 4.3, and compute the norm of \( Z \); let \( f_{\text{approx}} \) be the output of the algorithm. Observe that, by (4.5),

\[
||Z||_p \leq \frac{132pn^8}{p-2},
\]

so

\[
|f_{\text{approx}} - ||Z||_p| \leq \epsilon ||Z||_p \leq \epsilon \left( 132^n \left( \frac{p}{p-2} \right)^{n^8} \right) \leq \left( 132^n \left( \frac{p}{p-2} \right)^{n^8} \right)^{-1}.
\]

It follows then from Proposition 4.3 that

\[
n \left( \frac{f_{\text{approx}}}{2} \right)^p - n \left( 64 \cdot \left( \frac{p}{p-2} \right)^{n^8} \right)^p
\]

is an approximation of the maximum cut with an additive error at most \( 1/n \). Once we have \( f_{\text{approx}} \), we can approximate this number in polynomial time to an additive accuracy of \( 1/4 \). This gives an additive error \( 1/4 + 1/n \) approximation algorithm for
maximum cut, and since the maximum cut is always an integer, this means we can compute it exactly when \( n > 4 \). However, maximum cut is an NP-hard problem \([1]\).

For the case of \( p \in (1, 2) \), NP-hardness follows from the analysis of the case of \( p > 2 \) since, for any matrix \( Z \), \( \|Z\|_p = \|Z^T\|_{p'} \), where \( 1/p + 1/p' = 1 \).

**Remark.** In contrast to Theorem 2.3 which proves the NP-hardness of computing the matrix \( \infty, k \)-norm to relative accuracy \( \epsilon = 1/C(p) \), for some function \( C(p) \), Proposition 4.4 proves the NP-hardness of computing the \( p \)-norm to accuracy \( 1/C'(p)n^{8p+11} \) for some function \( C'(p) \). In the latter case, \( \epsilon \) depends on \( n \).

Our final theorem demonstrates that the \( p \)-norm is still hard to compute when restricted to matrices with entries in \( \{-1, 0, 1\} \).

**Theorem 4.5.** Fix a rational \( p \in [1, \infty) \) with \( p \neq 1, 2 \). Unless \( P = NP \), there is no algorithm which, given input \( \epsilon \) and a matrix \( M \) with entries in \( \{-1, 0, 1\} \), computes \( \|M\|_p \) to relative accuracy \( \epsilon \), in time which is polynomial in \( \epsilon^{-1} \) and the dimensions of the matrix.

**Proof.** As before, it suffices to prove the theorem for the case of \( p > 2 \); the case of \( p \in (1, 2) \) follows because \( \|Z\|_p = \|Z^T\|_{p'} \), where \( 1/p + 1/p' = 1 \).

Define

\[
Z^* = \left( \left\lceil \left( \frac{64p}{p-2} n^8 \right)^p \right\rceil M(G) \right) A
\]

where \( \lceil \cdot \rceil \) refers to rounding up to the closest integer. Observe that, by an argument similar to the proof of the previous proposition, computing \( \|Z^*\|_p \) to an accuracy \( \epsilon = (C(p)n^{8p+11})^{-1} \) is NP-hard for some function \( C(p) \). But if we define

\[
Z^{**} = \begin{pmatrix}
A \\
A \\
\vdots \\
A \\
M
\end{pmatrix}
\]

where \( A \) is repeated \( \left\lceil \left( \frac{64p}{p-2} n^8 \right)^p \right\rceil \) times, then

\[
\|Z^{**}\|_p = \|Z^*\|_p.
\]

The matrix \( Z^{**} \) has entries in \( \{-1, 0, 1\} \), and its size is polynomial in \( n \), so it follows that it is NP-hard to compute \( \|Z^{**}\|_p \) within the same \( \epsilon \).

**Remark.** Observe that the argument also suffices to show that computing the \( p \)-norm of square matrices with entries in \( \{-1, 0, 1\} \) is NP-hard: simply pad each row of \( Z^{**} \) with enough zeros to make it square. Note that this trick was also used in section 2.

5. **Concluding remarks.** We have proved the NP-hardness of computing the matrix \( p \)-norm approximately with relative error \( \epsilon = 1/C(p)n^{8p+11} \), where \( C(p) \) is some function of \( p \), and the NP-hardness of computing the matrix \( \infty, p \)-norm to some fixed relative accuracy depending on \( p \). We finish with some technical remarks about various possible extensions of the theorem:

- Due to the linear property of the norm \( \|\alpha A\| = |\alpha| \|A\| \), our results also imply the NP-hardness of approximating the matrix \( p \)-norm with any fixed or polynomially growing additive error.
Our construction also implies the hardness of computing the matrix $p$-norm for any irrational number $p > 1$ for which a polynomial time algorithm to approximate $x^p$ is available.

Our construction may also be used to provide a new proof of the NP-hardness of the $\| \cdot \|_{p,q}$ norm when $p > q$, which has been established in [5]. Indeed, it rests on the matrix $A$ with the property that $\max \|Ax\|_p/\|x\|_p$ occurs at the vectors $x \in \{-1,1\}^n$. We use this matrix $A$ to construct the matrix $Z = (\alpha A M)^T$ for large $\alpha$ and argue that $\max \|Zx\|_p/\|x\|_p$ occurs close to the vectors $x \in \{-1,1\}^n$. At these vectors, it happens $Ax$ is a constant, so we are effectively maximizing $\|Mx\|_p$, which is hard as shown in section 2.

If one could come up with such a matrix for the case of the mixed $\| \cdot \|_{p,q}$ norm, one could prove NP-hardness by following the same argument. However, when $p > q$, actually the same matrix $A$ works. Indeed, one could simply argue that

$$\|A\|_{p,q} = \max_{x \neq 0} \|Ax\|_q/\|x\|_p = \max_{x \neq 0} \|Ax\|_q \|x\|_q/\|x\|_p,$$

and since the maximum of $\|x\|_q/\|x\|_p$ when $1 \leq q < p \leq \infty$ occurs at the vectors $x \in \{-1,1\}^n$, we have that both terms on the right are maximized at $x = \in \{-1,1\}^n$, and that is where $\|Ax\|_q/\|x\|_p$ is maximized.

Finally we note that our goal was only to show the existence of a polynomial time reduction from the maximum cut problem to the problem of matrix $p$-norm computation. It is possible that more economical reductions which scale more gracefully with $n$ and $p$ exist.

REFERENCES


