The p-norm of circulant matrices

Ludovick Bouthat

Université Laval

Joint Mathematics Meetings; April 6, 2022

Acknowledgement

This research is a collaborate effort with Pr. Apoorva Khare, Pr. Javad Mashreghi and Pr. Frédéric Morneau-Guérin.

It was done with the financial help of the NSERC Graduate Scholarships.

Doubly stochastic matrices

Definition

A square matrix is *doubly stochastic* if:

- nonnegative coefficients;
- row sums $= 1$;
- column sums $= 1$.

The set of doubly stochastic matrices of order n is denoted by Ω_n.

Example

$$D = \begin{bmatrix} 0.1 & 0.3 & 0.6 \\ 0.4 & 0.2 & 0.4 \\ 0.5 & 0.5 & 0 \end{bmatrix}.$$

Geometric interpretation

Theorem (Birkhoff; 1946)

Ω_n is a convex polytope in $\mathbb{R}^{n^2}$ of dimension $(n-1)^2$ with the $n!$ permutation matrices as its vertices.

Geometric interpretation

Theorem (Birkhoff; 1946)

Ω_n is a convex polytope in $\mathbb{R}^{n^2}$ of dimension $(n-1)^2$ with the $n!$ permutation matrices as its vertices.

- Ω_n is *vertex-symmetric* and *edges-symmetric*.

Geometric interpretation

Theorem (Birkhoff; 1946)

Ω_n is a convex polytope in $\mathbb{R}^{n^2}$ of dimension $(n-1)^2$ with the $n!$ permutation matrices as its vertices.

- Ω_n is *vertex-symmetric* and *edges-symmetric*.
- The matrix $J_n := \frac{1}{n}\begin{bmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{bmatrix}$ is the *geometrical centroid* of Ω_n.

The Chebyshev radius

Definition

Let $\mathcal{E}$ be a bounded set of a norm-induced metric space $(\|\cdot\|, \mathcal{U})$. The *Chebyshev radius* $R(\mathcal{E})$ of $\mathcal{E}$ is the smallest radius of a bounding sphere of $\mathcal{E}$. The *Chebyshev center* $C(\mathcal{E})$ of $\mathcal{E}$ is the center of the bounding sphere for which the minimum is attained.

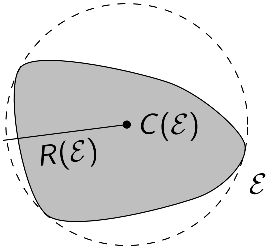

Figure: The Chebyshev radius and the Chebyshev center of $\mathcal{E}$.

Permutation-invariant matrix norm

Definition

A matrix norm $\|\cdot\|$ is *permutation-invariant* if

$$\|QAP\| = \|A\|$$

for every permutation matrix P and Q.

Example

Every matrix norm induced by a permutation-invariant *vector* norm is a permutation-invariant matrix norm.

The Chebyshev radius and center of Ω_n

Proposition

If $\|\cdot\|$ is a permutation-invariant matrix norm, then the Chebyshev center of Ω_n is the matrix J_n and the Chebyshev radius is given by

$$R(\Omega_n) \;=\; \|I_n - J_n\|.$$

The operator norm from $\ell^p \to \ell^p$

Lemma

If $\| \cdot \|_{\ell^p \to \ell^p}$ *denote the operator norm from* $\ell^p \to \ell^p$*, then*

$$\|I_n - J_n\|_{\ell^1 \to \ell^1} = \|I_n - J_n\|_{\ell^\infty \to \ell^\infty} = 2\left(1 - \frac{1}{n}\right)$$

$$\&$$

$$\|I_n - J_n\|_{\ell^2 \to \ell^2} = 1.$$

A generalization

Open question: Determine the operator norm from $\ell^p \to \ell^p$ of

$$A(n, a, b) \;:=\; (a-b)I_n + bnJ_n \;=\; \begin{pmatrix} a & b & b & \cdots & b \\ b & a & b & \cdots & b \\ b & b & a & \cdots & b \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b & b & b & \cdots & a \end{pmatrix}.$$

A generalization

Open question: Determine the operator norm from $\ell^p \to \ell^p$ of

$$A(n, a, b) \;:=\; (a-b)I_n + bnJ_n \;=\; \begin{pmatrix} a & b & b & \cdots & b \\ b & a & b & \cdots & b \\ b & b & a & \cdots & b \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b & b & b & \cdots & a \end{pmatrix}.$$

- In particular, we have $I_n - J_n = A(n, \frac{n-1}{n}, -\frac{1}{n})$.

Circulant matrices

Definition

Definition

Let $\alpha_1, \ldots, \alpha_n$ be real numbers. A *circulant matrix* is an $n \times n$ matrix of the form

$$\text{Circ}(\alpha_1, \ldots, \alpha_n) = \begin{pmatrix} \alpha_1 & \alpha_2 & \alpha_3 & \cdots & \alpha_n \\ \alpha_n & \alpha_1 & \alpha_2 & \cdots & \alpha_{n-1} \\ \alpha_{n-1} & \alpha_n & \alpha_1 & \cdots & \alpha_{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \alpha_2 & \alpha_3 & \alpha_4 & \cdots & \alpha_1 \end{pmatrix}.$$

The norm of $A(n, a, b)$

Positive case

Theorem

If $\alpha_1, \ldots, \alpha_n \geq 0$, *then*

$$\| \operatorname{Circ}(\alpha_1, \ldots, \alpha_n) \|_{\ell^p \to \ell^p} = \alpha_1 + \cdots + \alpha_n .$$

The norm of $A(n, a, b)$

Positive case

Theorem

If $\alpha_1, \ldots, \alpha_n \geq 0$, *then*

$$\|\operatorname{Circ}(\alpha_1, \ldots, \alpha_n)\|_{\ell^p \to \ell^p} = \alpha_1 + \cdots + \alpha_n .$$

Corollary

If $a, b \geq 0$, *then*

$$\|A(n, a, b)\|_{\ell^p \to \ell^p} = (n-1)b + a .$$

The spectral norm of $A(n, -a, b)$

Theorem (B., Khare, Mashreghi, Morneau-Guérin; 2021)

If $a, b \geq 0$, *then*

$$\|A(n, -a, b)\|_{\ell^2 \to \ell^2} = \begin{cases} a + b & \text{if } (n-2)b \leq 2a, \\ (n-1)b - a & \text{if } (n-2)b \geq 2a. \end{cases}$$

The p-norm of $A(n, -a, b)$

Theorem (B., Khare, Mashreghi, Morneau-Guérin; 2021)

If $a, b \geq 0$ and $A = A(n, -a, b)$, then for all $p \geq 2$,

$$\|A\|_{\ell^2 \to \ell^2} \leq \|A\|_{\ell^p \to \ell^p} \leq \|A\|_{\ell^2 \to \ell^2}^{2/p} \|A\|_{\ell^\infty \to \ell^\infty}^{1-2/p}.$$

Moreover, the upper bound becomes an equality if $p = 2, \infty$.

The p-norm of $A(n, -a, b)$

Theorem (B., Khare, Mashreghi, Morneau-Guérin; 2021)

If $a, b \geq 0$ and $A = A(n, -a, b)$, then for all $p \geq 2$,

$$\|A\|_{\ell^2 \to \ell^2} \;\leq\; \|A\|_{\ell^p \to \ell^p} \;\leq\; \|A\|_{\ell^2 \to \ell^2}^{2/p} \|A\|_{\ell^\infty \to \ell^\infty}^{1-2/p}.$$

Moreover, the upper bound becomes an equality if $p = 2, \infty$.

Theorem (Sahasranand; 2022)

If $a, b \geq 0$ and $A = A(n, -a, b)$, then for all $p \geq 2$, $\|A\|_{\ell^p \to \ell^p}$ is monotonically non-decreasing in p.

More general norms

Definition

Let $\|\cdot\|$ be a matrix norm and U be an $n \times n$ matrix. Then

$$\|U\|_\ell := \sup\left\{\frac{\|UX\|}{\|X\|} : X \in \mathbb{C}^{n\times n}, X \neq \mathbf{0}\right\}$$

is the *left operator norm induced by* $\|\cdot\|$ of U and

$$\|U\|_r := \sup\left\{\frac{\|XU\|}{\|X\|} : X \in \mathbb{C}^{n\times n}, X \neq \mathbf{0}\right\}.$$

is the *right operator norm induced by* $\|\cdot\|$ of U.

A general result

Proposition

Let $\| \cdot \|$ be a norm on $\mathbb{C}^{n\times n}$ which satisfies

$$\| \operatorname{diag}(a_1, \ldots, a_n) \| \ = \ \max\{|a_1|, \ldots, |a_n|\} \tag{1}$$

and let $A \in \mathbb{C}^{n\times n}$ be a normal matrix which admits the decomposition $U\operatorname{diag}\big(\lambda_1, \ldots, \lambda_n\big)\, U^*$. Then for each polynomial p in two variables,

$$\|p(A, A^*)\| \ \leq \ \|U\|_\ell \|U^*\|_r \max\left\{|p(\lambda_1, \overline{\lambda_1})|, \ldots, |p(\lambda_n, \overline{\lambda_n})|\right\}$$

and

$$\|p(A, A^*)\| \ \geq \ \frac{1}{\|U\|_r \|U^*\|_\ell} \max\left\{|p(\lambda_1, \overline{\lambda_1})|, \ldots, |p(\lambda_n, \overline{\lambda_n})|\right\}.$$

Application to circulant matrices

- Let $A = \text{Circ}(\alpha_1, \dots, \alpha_\text{n})$ be a circulant matrix and let $\mathcal{P}_A(z) = \alpha_1 + \alpha_2 z + \cdots + \alpha_n z^{n-1}$ be its *associated polynomial*. Let $\omega_n := e^{2\pi i/n}$ and define

$$C_n \;:=\; \text{Circ}(0,1,0,\dots,0) \;=\; \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & \cdots & 0 \end{pmatrix}.$$

Then A admits the decomposition

$$A \;=\; \mathcal{P}_A(W_n C_n W_n^*) \;=\; W_n \text{diag}\Big(\mathcal{P}_A(1), \mathcal{P}_A(\omega_n), \cdots, \mathcal{P}_A(\omega_n^{n-1})\Big) W_n^*.$$

Application to circulant matrices

Corollary

Let $A = \mathrm{Circ}(\alpha_1, \ldots, \alpha_n)$ *and let* $\|\cdot\|$ *be a matrix norm satisfying* (1). *Then*

$$\|A\| \le \|W_n\|_\ell \|W_n^*\|_r \max\{|\mathcal{P}_A(1)|, \ldots, |\mathcal{P}_A(\omega_n^{n-1})|\},$$

and

$$\|A\| \ge \frac{1}{\|W_n\|_r \|W_n^*\|_\ell} \max\{|\mathcal{P}_A(1)|, \ldots, |\mathcal{P}_A(\omega_n^{n-1})|\}.$$

Application to circulant matrices

Corollary

Let $A = \mathrm{Circ}(\alpha_1, \ldots, \alpha_n)$ *and let* $\|\cdot\|$ *be a matrix norm satisfying* (1)*. Then*

$$\|A\| \leq \|W_n\|_\ell \|W_n^*\|_r \max\left\{|\mathcal{P}_A(1)|, \ldots, |\mathcal{P}_A(\omega_n^{n-1})|\right\},$$

and

$$\|A\| \geq \frac{1}{\|W_n\|_r \|W_n^*\|_\ell} \max\left\{|\mathcal{P}_A(1)|, \ldots, |\mathcal{P}_A(\omega_n^{n-1})|\right\}.$$

Unfortunately, this does not yield better results than what was previously known for the particular cases $A(n, -a, b)$ and $I_n - J_n$.

An interesting case

Proposition

If $\| \cdot \|$ is a matrix norm on $\mathbb{C}^{n \times n}$ satisfying $\|I\| = 1$, then the left operator norm and the right operator norm coincide with the norm itself.

An interesting case

Proposition

If $\|\cdot\|$ is a matrix norm on $\mathbb{C}^{n\times n}$ satisfying $\|I\| = 1$, then the left operator norm and the right operator norm coincide with the norm itself.

Remark

The left operator norm and the right operator norm does not always coincide with the norm itself. For instance, for the Frobenius norm $\|\cdot\|_F$, we have

$$\|I\|_\ell = \|I\|_r = 1 \ < \ \sqrt{n} = \|I\|_F.$$

Definition

Definition

Let $A \in M_n(\mathbb{R})$. The *maximizing vectors* of $\frac{\|Ax\|_p}{\|x\|_p}$ are the vectors $x \in \mathbb{R}^n$ for which

$$\frac{\|Ax\|_p}{\|x\|_p} = \|A\|_{\ell^p \to \ell^p}.$$

Example

Let

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

Then the maximizing vectors of $\frac{\|Ax\|_p}{\|x\|_p}$ are those of the form $x = (a, 0)^\intercal$, where $a \in \mathbb{R}$ and $a \neq 0$.

The maximizing vectors of $A(n, -a, b)$

Proposition (B., Mashreghi, Morneau-Guérin; 2022)

Let $a, b \geq 0$, with at least one of them non-zero and let $A = A(n, -a, b)$. Suppose that $n > 2$ and that $1 \leq p \leq \infty$, with $p \neq 2$. If $x \in \mathbb{R}^n$ is a maximizing vector of $\frac{\|Ax\|_p}{\|x\|_p}$, then the entries of x form a set of cardinality at most three.

The maximizing vectors of $A(n, -a, b)$

Proposition (B., Mashreghi, Morneau-Guérin; 2022)

Let $a, b \geq 0$, with at least one of them non-zero and let $A = A(n, -a, b)$. Suppose that $n > 2$ and that $1 \leq p \leq \infty$, with $p \neq 2$. If $x \in \mathbb{R}^n$ is a maximizing vector of $\frac{\|Ax\|_p}{\|x\|_p}$, then the entries of x form a set of cardinality at most three.

Conjecture

Under the previous hypothesis, the entries of x form a set of cardinality at most *two*.

The special case $I_n - J_n$

Proposition

Let $1 \leq p \leq \infty$, with $p \neq 2$ and $\rho := p-1$. Let x_p be the unique root of the function

$$x \longmapsto \rho\left(1+x^{\frac{1}{\rho}}\right)\left(1-x^{\rho-1}\right)+\left(1-x^{\frac{1-\rho}{\rho}}\right)\left(1+x^{\rho}\right)$$

in the interval $[0,1]$, and let $m_1 := \left\lfloor \frac{n}{1+x_p} \right\rfloor$ and $m_2 := \left\lceil \frac{n}{1+x_p} \right\rceil$. Suppose that the previous Conjecture is valid. Then

$$\|I_n - J_n\|_{\ell^p \to \ell^p} = \max_{m \in \{m_1, m_2\}} \frac{\left(\left(\frac{n}{m}-1\right)^{p-1}+1\right)^{\frac{1}{p}}\left(\left(\frac{n}{m}-1\right)^{\frac{1}{p-1}}+1\right)^{1-\frac{1}{p}}}{\frac{n}{m}}.$$

Concluding remarks

Generalized measures of statistical dispersion

Definition

If $x \in \mathbb{R}^n$ and $p \in [1, \infty)$, then the *p-deviation* of x is

$$D_p(x) := \left(\frac{1}{n} \sum_{i=1}^{n} |x_i - \overline{x}|^p \right)^{1/p},$$

where $\overline{x} := \frac{x_1 + x_2 + \cdots + x_n}{n}$.

Concluding remarks

The p-generalized variance

Remark

Let $x \in \mathbb{R}^n$ and $p \in [1, \infty)$. We have the identity

$$n\, D_p^p(x) \;=\; \|(I_n - J_n)x\|_p^p.$$

Concluding remarks

The p-generalized variance

Remark

Let $x \in \mathbb{R}^n$ and $p \in [1, \infty)$. We have the identity

$$n\, D_p^p(x) \;=\; \|(I_n - J_n)x\|_p^p.$$

$$\text{Determining } \|I_n - J_n\|_{\ell^p \to \ell^p} \quad \Longleftrightarrow \quad \text{Maximizing } \frac{D_p(x)}{\|x\|_p}.$$

Concluding remarks

A connection with harmonic analysis

Remark

Let $\mathcal{P}_n$ denote the space of polynomials of degree at most n. We may interpret $A = A(n, -a, b)$ as an operator on $\mathcal{P}_{n-1}$. More explicitly, for each polynomial $f(z) = a_0 + a_1 z + \cdots + a_{n-1} z^{n-1} \in \mathcal{P}_{n-1}$, we have

$$(Af)(z) \;=\; -(a+b)f(z) + b(a_0 + a_1 + \cdots + a_{n-1})\varphi(z),$$

where

$$\varphi(z) \;=\; 1 + z + \cdots + z^{n-1}.$$

Concluding remarks

A connection with harmonic analysis

- Using this interpretation, we get

$$\|A\|_{\ell^p \to \ell^p} \;\leq\; a + b + bn\|\varphi\|_{L^1(\mathbb{T})},$$

where

$$\|\varphi\|_{L^1(\mathbb{T})} \;=\; \int_0^{2\pi} |\varphi(e^{i\theta})| \, \frac{d\theta}{2\pi}.$$

Concluding remarks

A connection with harmonic analysis

- Using this interpretation, we get

$$\|A\|_{\ell^p \to \ell^p} \;\leq\; a + b + bn\|\varphi\|_{L^1(\mathbb{T})} ,$$

where

$$\|\varphi\|_{L^1(\mathbb{T})} \;=\; \int_0^{2\pi} |\varphi(e^{i\theta})| \, \frac{d\theta}{2\pi} .$$

Question: Can variations of the above estimation lead to a precise formula for $\|A\|_{\ell^p \to \ell^p}$?

References

Š. Schwarz (1980) Infinite products of doubly stochastic matrices. *Acta Math. Univ. Comenian.*, 39: 131–150.

L. Bouthat, A. Khare, J. Mashreghi & F. Morneau-Guérin (2021) The p-norm of circulant matrices, *Linear and Multilinear Algebra*, DOI: 10.1080/03081087.2021.1983513

K. R. Sahasranand (2022) The p-norm of circulant matrices via Fourier analysis. *Concr. Oper.*, 9(1):1–5.

L. Bouthat, J. Mashreghi & F. Morneau-Guérin (2022) On the norm of normal matrices, *RIMS Kôkyûroku Bessatsu*, https://r-libre.teluq.ca/2603/, (submitted for publication).