The norm of a circulant operator

Ludovick Bouthat

Université Laval

January 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.

NSERC
CRSNG

Doubly stochastic matrices

Definition

A square matrix is *doubly stochastic* if its coefficients are nonnegative and if the sum of the coefficients of each of its row and each of its column is equal to 1.

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

Example

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

Important cases

- The matrix $J_n = \frac{1}{n}\begin{bmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{bmatrix}$ acts as the *absorbing element* of Ω_n.

Definition

Let $(S, \cdot)$ be a set S with a closed binary operation $\cdot$ on it. An *absorbing element* is an element z such that for all $s \in S$, $z \cdot s = s \cdot z = z$.

The van der Waerden's conjecture

Definition

The permanent of an $n \times n$ matrix $A = (a_{ij})$ is defined as

$$\text{perm}(A) = \sum_{\sigma \in S_n} \prod_{i=1}^{n} a_{i\sigma(i)}.$$

Theorem (van der Waerden's conjecture, 1926)

The minimum permanent among all $D \in \Omega_n$ is attained uniquely by the matrix J_n.

The van der Waerden's conjecture

Definition

The permanent of an $n \times n$ matrix $A = (a_{ij})$ is defined as

$$\operatorname{perm}(A) = \sum_{\sigma \in S_n} \prod_{i=1}^{n} a_{i\sigma(i)}.$$

Theorem (van der Waerden's conjecture, 1926)

The minimum permanent among all $D \in \Omega_n$ is attained uniquely by the matrix J_n.

• Proofs of this conjecture were published in 1980 by Gyires and in 1981 Egorychev and Falikman.

Infinite products

- Ω_n is closed under the regular matrix multiplication.

Infinite products

- Ω_n is closed under the regular matrix multiplication.

Question: What is the long-term behaviour of a product of d.s. matrices?

Infinite products

- Ω_n is closed under the regular matrix multiplication.

Question: What is the long-term behaviour of a product of d.s. matrices?

Example

$$D = \begin{bmatrix} 0.1 & 0.3 & 0.6 \\ 0.4 & 0.2 & 0.4 \\ 0.5 & 0.5 & 0 \end{bmatrix} \rightsquigarrow D^8 = \begin{bmatrix} 0.334636 & 0.334634 & 0.33073 \\ 0.333332 & 0.333335 & 0.333332 \\ 0.332031 & 0.332031 & 0.335938 \end{bmatrix}$$

Infinite products

- Ω_n is closed under the regular matrix multiplication.

Question: What is the long-term behaviour of a product of d.s. matrices?

Example

$$D = \begin{bmatrix} 0.1 & 0.3 & 0.6 \\ 0.4 & 0.2 & 0.4 \\ 0.5 & 0.5 & 0 \end{bmatrix} \rightsquigarrow D^8 = \begin{bmatrix} 0.334636 & 0.334634 & 0.33073 \\ 0.333332 & 0.333335 & 0.333332 \\ 0.332031 & 0.332031 & 0.335938 \end{bmatrix}$$

In fact, $D^k \to J_3$ when $k \to \infty$.

Infinite products

- Ω_n is closed under the regular matrix multiplication.

Question: What is the long-term behaviour of a product of d.s. matrices?

Example

$$D = \begin{bmatrix} 0.1 & 0.3 & 0.6 \\ 0.4 & 0.2 & 0.4 \\ 0.5 & 0.5 & 0 \end{bmatrix} \rightsquigarrow D^8 = \begin{bmatrix} 0.334636 & 0.334634 & 0.33073 \\ 0.333332 & 0.333335 & 0.333332 \\ 0.332031 & 0.332031 & 0.335938 \end{bmatrix}$$

In fact, $D^k \to J_3$ when $k \to \infty$.

Counterexample: $I^k = I$ for every $k \in \mathbb{N}$.

Approach Idea

Question: When does $A_1 \cdots A_k \xrightarrow{k \to \infty} J_n$ for $(A_i)_{i \geq 1} \subseteq \Omega_n$?

Approach Idea

Question: When does $A_1 \cdots A_k \xrightarrow{k\to\infty} J_n$ for $(A_i)_{i\geq 1} \subseteq \Omega_n$?

Strategy:

- Consider $\|A_1 \cdots A_k - J_n\|$ for different matrix norms.

Approach Idea

Question: When does $A_1 \cdots A_k \xrightarrow{k\to\infty} J_n$ for $(A_i)_{i\geq 1} \subseteq \Omega_n$?

Strategy:

1. Consider $\|A_1 \cdots A_k - J_n\|$ for different matrix norms.
2. Find a norm for which $\sup_{A\in\Omega_n} \|A - J_n\|$ is *small*.

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

A simpler formula

Proposition

If $\|\cdot\|$ is a permutation-invariant matrix norm, then

$$\sup_{A \in \Omega_n} \|A - J_n\| = \|I_n - J_n\|.$$

The Schatten p-norm

Definition

If the singular values of the $m \times n$ matrix A are denoted by σ_i, then the Schatten p-norm is defined by

$$\|A\|_{S_p} = \left(\sum_{i=1}^{\min\{m,n\}} \sigma_i^p \right)^{\frac{1}{p}}.$$

Lemma

If $\|\cdot\|_{S_p}$ denote the Schatten p-norm, then

$$\|I_n - J_n\|_{S_p} = (n-1)^{1/p}.$$

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

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

$$\&$$

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

Main Problem

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

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

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

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

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

Positive case

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

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

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

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

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

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

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

A similar pattern

A simple observation

If $A = \text{Circ}(\alpha_1, \alpha_2, \alpha_3)$ for $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}$ and $\Delta = \alpha_1\alpha_2 + \alpha_2\alpha_3 + \alpha_3\alpha_1$, then

$$\|A\|_{\ell^2 \to \ell^2} = \begin{cases} \sqrt{\alpha_1^2 + \alpha_2^2 + \alpha_3^2 - \Delta} & \text{if } \Delta \leq 0, \\ |\alpha_1 + \alpha_2 + \alpha_3| & \text{if } \Delta \geq 0. \end{cases}$$

A similar pattern

A simple observation

If $A = \mathrm{Circ}(\alpha_1, \alpha_2, \alpha_3)$ for $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}$ and $\Delta = \alpha_1\alpha_2 + \alpha_2\alpha_3 + \alpha_3\alpha_1$, then

$$\|A\|_{\ell^2 \to \ell^2} = \begin{cases} \sqrt{\alpha_1^2 + \alpha_2^2 + \alpha_3^2 - \Delta} & \text{if } \Delta \leq 0, \\ |\alpha_1 + \alpha_2 + \alpha_3| & \text{if } \Delta \geq 0. \end{cases}$$

Open question: Does a similar pattern emerge for every circulant matrix?

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, we have equality in the upper bound for $p = 2, \infty$.

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, we have equality in the upper bound for $p = 2, \infty$.

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.

Concluding remarks

The p-generalized variance

Definition

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

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

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

Concluding remarks

The p-generalized variance

Remark

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

$$n \operatorname{Var}_p^p(x) = \|(I_n - J_n)x\|_p^p.$$

Concluding remarks

The p-generalized variance

Remark

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

$$n \operatorname{Var}_p^p(x) = \|(I_n - J_n)x\|_p^p.$$

$$\text{Determining } \|I_n - J_n\|_{\ell^p \to \ell^p} \iff \text{Maximizing } \frac{\operatorname{Var}_p(x)}{\|x\|_p}.$$

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

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

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

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

where

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

Question: Does the above estimation, or its variations, 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. arXiv preprint arXiv:2111.11389.