Nierówności macierzowe

Tomasz Tkocz

Streszczenie. Jest to referat wygłoszony na Toruńskiej Letniej Szkole Matematyki 2009.

Profesor Steinhaus podobno mawiał, że na każdym referacie trzeba coś pokazać — udowodnić. Żeby mi na to nie brakło czasu, rozwiążmy sobie, tak na rozgrzewkę, dobrze wszystkim znane zadanko.

Zadanie 1 (Hadamard). *Niech $A = [a_{ij}]_{i,j=1,\dots n}$ będzie macierzą $n \times n$ o rzeczywistych współczynnikach. Udowodnić, że*

$$|\det A| \leqslant \sqrt{\prod_{j=1}^{n} \sum_{i=1}^{n} a_{ij}^2}.$$

W szczególności dla $n = 2, A = \left[\begin{smallmatrix} r & s \\ t & u \end{smallmatrix}\right]$:

$$|ru - st| \leqslant \sqrt{r^2 + t^2}\sqrt{s^2 + u^2} \quad \textit{(odtworzona nierówność Schwarza)},$$

dla $n = 3, A = \left[\begin{smallmatrix} r & s & t \\ u & v & w \\ x & y & z \end{smallmatrix}\right]$:

$$|rvz + swx + tuy - tvx - wyr - zsu|$$
$$\leqslant \sqrt{r^2 + u^2 + x^2}\sqrt{s^2 + v^2 + y^2}\sqrt{t^2 + w^2 + z^2}.$$

Rozwiązanie. Mamy

$$|\det A| = \text{objętość równoległościanu rozpiętego przez wektory kolumnowe macierzy A} \leqslant \text{iloczyn ich długości}.$$

$\square$

Morał może być na przykład taki, że (ładne) nierówności można szukać w macierzach.

No to szukamy. W fizyce statystycznej kluczowe znaczenie ma takie zwierzę jak suma statystyczna, która jest postaci $\operatorname{tr} e^A$, gdzie A to przekształcenie liniowe samosprzężone. Bardzo ważne są jej różnorakie oszacowania, co jest świetną motywacją dla naszych poszukiwań.

Dalej zakładamy, że w przestrzeni liniowej $\mathbb{C}^n$ mamy zadany iloczyn skalarny $<,>$.

Przykładowo, zachodzi

Twierdzenie 1 (Peierls–Bogolubow). *Dla operatorów liniowych samosprzężonych $A, B\colon \mathbb{C}^n \longrightarrow \mathbb{C}^n$ zachodzi*

$$\frac{\operatorname{tr} e^{A+B}}{\operatorname{tr} e^B} \geqslant \exp\left(\frac{\operatorname{tr} Ae^B}{\operatorname{tr} e^B}\right).$$

W szczególności ($B = I$)

$$\frac{1}{n}\operatorname{tr} e^A \geqslant e^{\frac{1}{n}\operatorname{tr} A}.$$

Nie można nie wspomnieć o

Twierdzenie 2 (Golden–Thompson). *Dla operatorów liniowych samosprzężonych $A, B\colon \mathbb{C}^n \longrightarrow \mathbb{C}^n$ zachodzi*

$$\operatorname{tr} e^{A+B} \leqslant \operatorname{tr}\left(e^A e^B\right),$$

a nawet więcej

$$\operatorname{tr} e^{A+B} \leqslant \operatorname{tr}\left(e^{A/m} e^{B/m}\right)^m \leqslant \operatorname{tr}\left(e^A e^B\right), \qquad 0 < m \in \mathbb{Z}.$$

Spróbujmy teraz prześledzić szkic dowodu tej ślicznej nierówności. Dzięki temu poznamy podstawowe metody teorii nierówności macierzowych i zobaczymy jeszcze kilka innych pomocniczych, ale ładnych oszacowań.

Po kolei; dlaczego operatory samosprzężone (tzn. takie, które zachowują iloczyn sklarny) są przyjemne? Wyjaśnia to małe przypomnienie z algebry liniowej

Twierdzenie 3 (spektralne). *Jeśli operator liniowy $A\colon \mathbb{C}^n \longrightarrow \mathbb{C}^n$ jest samosprzężony[1] ($A^\dagger = A$), to istnieje baza ortonormalna $u_1, \dots, u_n$ oraz liczby rzeczywiste $\lambda_1, \dots \lambda_n$ takie, że $Au_k = \lambda_k u_k, k = 1, \dots, n$.*

Definicja 1. Niech A — operator samosprzężony. Dla $f\colon \mathbb{R} \longrightarrow \mathbb{R}$ definiujemy operator $f(A)\colon \mathbb{C}^n \longrightarrow \mathbb{C}^n$ określony na bazie $u_1, \dots, u_n$ z twierdzenia spektralnego wzorem $f(A)u_k = f(\lambda_k)u_k$.

Oczywiście ta definicja jest poprawna (tzn. nie zależy od wyboru bazy $u_1, \dots, u_n$ z twierdzenia spektralnego), bo $f(A)$ w powyższy sposób jest tak naprawdę określony jako homotetie na odpowiednich podprzestrzeniach wyznaczonych przez A.

[1] przez $A^\dagger$ oznaczamy hermitowskie sprzężenie operatora A

Na przykład, dla $f(x) = x^2$ mamy $f(A)u_k = \lambda_k^2 u_k$, czyli macierz operatora $f(A)$ w bazie $u_1, \ldots, u_n$ jest postaci $\begin{bmatrix} \lambda_1^2 & & 0 \\ & \ddots & \\ 0 & & \lambda^2 \end{bmatrix}$, czyli jest kwadratem macierzy operatora A w tej bazie, tak jak powinno byc.

Koniec przypomnienia. Zachodzi kluczowe dla nas

Twierdzenie 4. *Dla operatora samosprzężonego $A\colon \mathbb{C}^n \longrightarrow \mathbb{C}^n$ nieujemnego*[2]*, funkcji wypukłej $f\colon [0,\infty) \longrightarrow \mathbb{R}$ i bazy ortonormalnej $v_1, \ldots, v_n$ zachodzi*

$$\operatorname{tr} f(A) \geqslant \sum_{k=1}^{n} f\left(< Av_k, v_k >\right).$$

Stąd łatwo udowodnimy taki ładny lemacik.

Lemat 1. *Dla operatora liniowego $X\colon \mathbb{C}^n \longrightarrow \mathbb{C}^n$ zachodzi*

$$\left|\operatorname{tr} X^{2m}\right| \leqslant \operatorname{tr}\left(X^\dagger X\right)^m, \qquad 0 \leqslant m \in \mathbb{Z}.$$

Dowód. W powyższym twierdzeniu 4 bierzemy $f(x) = x^m$, $A = X^\dagger X$ i za bazę $v_1, \ldots, v_n$ bierzemy bazę ortonormalną w której macierz $[x_{ij}]_{i,j=1,\ldots,n}$ operatora X jest górnotrójkątna (że taka istnieje jest domowym ćwiczeniem z algebry liniowej). Mamy wówczas

$$\operatorname{tr}\left(X^\dagger X\right)^m \geqslant \sum_{k=1}^{n} < X^\dagger X v_k, v_k >^m$$
$$= \sum_{k=1}^{n} \left(\|k\text{-ta kolumna macierzy } [x_{ij}]\|^2\right)^m \geqslant \sum_{k=1}^{n} |x_{kk}|^{2m},$$

ale $|\operatorname{tr} X^{2m}| = |\sum_{k=1}^n x_{kk}^{2m}| \leqslant \sum_{k=1}^n |x_{kk}|^{2m} \leqslant \operatorname{tr}\left(X^\dagger X\right)^m$. ◻

Dzięki tej nierówności można indukcyjnie udowodnić

Lemat 2. *Dla operatorów liniowych samosprzężonych $X, Y\colon \mathbb{C}^n \longrightarrow \mathbb{C}^n$ zachodzi*

$$\left|\operatorname{tr}(XY)^{2^m}\right| \leqslant \operatorname{tr}\left(X^{2^m} Y^{2^m}\right), \qquad 0 \leqslant m \in \mathbb{Z}.$$

[2] dla każdego $x \in \mathbb{C}^n$ zachodzi $< Ax, x > \geqslant 0$

Kładziemy $X = I + \frac{1}{2^m}A, Y = I + \frac{1}{2^m}B$, gdzie A, B to operatory samosprzężone i ufając, że

$$\begin{aligned} (XY)^{2^m} &= \left(\left(I + \frac{1}{2^m}A\right)\left(I + \frac{1}{2^m}B\right)\right)^{2^m} \\ &= \left(I + \frac{1}{2^m}(A+B) + \left(\frac{1}{2^m}\right)^2 AB\right)^{2^m} \xrightarrow{m \longrightarrow \infty} e^{A+B}, \\ X^{2^m}Y^{2^m} &= \left(I + \frac{1}{2^m}A\right)^{2^m}\left(I + \frac{1}{2^m}B\right)^{2^m} \xrightarrow{m \longrightarrow \infty} e^A e^B, \end{aligned}$$

otrzymujemy $\operatorname{tr} e^{A+B} \leqslant \operatorname{tr}\left(e^A e^B\right)$.

Literatura

[1] S. Golden, *Lower Bounds for the Helmholtz Function*, Physical Review **137**, B1127–1128 (1965)

[2] B. Simon, *The statistical Mechanics of Lattice Gases*, Princeton, New Jersey, Princeton University Press

[3] M. Suzuki, *Quantum Statistical Monte Carlo Methods and Applications to Spin Systems*, Journal of Statistical Physics **43**, 883–909 (1986)

[4] C. J. Thompson, *Inequality with Applications in Statistical Mechanics*, Journal of Mathematical Physics **6**, 1812–1813 (1965)