Podróże w $\mathbb{R}^d$

Wojciech CZERWIŃSKI

Rozważmy następujący problem. Mamy danych wiele d-wymiarowych wektorów: $u_1, \ldots, u_n \in \mathbb{R}^d$, takich że sumują się one do wektora zerowego, czyli $\sum_{i=1}^{n} u_i = \vec{0}$. Podkreślmy, że liczba wektorów n może być dużo większa niż d. Załóżmy dodatkowo, że długości wszystkich wektorów u_i są nie większe niż 1. Dla danego ciągu wektorów $u_1, \ldots, u_n$ rozważmy podróż tym ciągiem w przestrzeni $\mathbb{R}^d$, w której startujemy z zera $\vec{0}$, a potem kolejno przesuwamy się o u_1, o u_2, o u_3 itd., a na końcu o u_n. Oczywiście na koniec całej podróży wrócimy do zera, ale w międzyczasie możemy od tego zera odsunąć się bardzo daleko. Pytanie brzmi, czy dla dowolnych wektorów $u_1, \ldots, u_n$ istnieje takie ich poprzestawianie, inaczej permutacja $u'_1, \ldots, u'_n$, żeby podróż tym ciągiem wektorów nigdy nie oddalała się znacząco od zera. Mówiąc bardziej precyzyjnie, czy dla dowolnego wymiaru d istnieje taka stała C_d, że dla dowolnego ciągu wektorów $u_1, \ldots, u_n \in \mathbb{R}^d$ nie dłuższych niż 1 i sumujących się do $\vec{0}$ istnieje ich permutacja $u'_1, \ldots, u'_n$ taka, że dla dowolnego $k \in \{1, \ldots, n\}$ zachodzi $|\sum_{i=1}^{k} u'_i| \leqslant C_d$. Nim przejdziemy do rozwiązania, zachęcam Ambitnego Czytelnika do samodzielnej próby odpowiedzenia na to pytanie albo przynajmniej postawienia hipotezy i obstawienia ewentualnej wartości C_d.

Dla wymiaru $d = 1$ stosunkowo łatwo jest wykazać, że $C_d = 1$ wystarczy. Konstruujemy ciąg wektorów u'_i tak, żeby wartość bezwzględna $\sum_{i=1}^{k} u'_i$ nigdy nie przekroczyła 1. Postępujemy w następujący sposób. Zaczynamy od ciągu pustego. Powiedzmy, że skonstruowaliśmy już ciąg $u'_1, \ldots, u'_k$, który spełnia nasze warunki dla początkowych wyrazów. Załóżmy bez straty ogólności, że $\sum_{i=1}^{k} u'_i \geqslant 0$. Skoro suma wszystkich wektorów u_i wynosi 0, a suma wektorów już wybranych do ciągu jest nieujemna, to suma wektorów niewybranych jest niedodatnia. A więc istnieje tam jakiś wektor niedodatni, tego właśnie wybieramy jako u'_{k+1}. Postępując w ten sposób do końca, otrzymamy ciąg o wymaganych własnościach. Łatwo też zauważyć, że C_1 jest wybrana optymalnie, ciąg $u_1 = 1$, $u_2 = -1$ nie da się ułożyć lepiej.

Dla wyższych wymiarów sprawa nie jest tak oczywista i mimo bardzo prostego sformułowania ma za sobą długą historię badań. Już w roku 1914 niemiecki matematyk Ernst Steinitz (znany m.in. z twierdzenia o dopełnianiu zbioru liniowo niezależnych wektorów do bazy) udowodnił, że dla dowolnego $d \in \mathbb{N}$ stała $C_d = 2d$ spełnia zadane warunki. Dlatego wspomniany fakt zwany jest lematem Steinitza. Sytuacja staje się jednak ciekawsza, gdy rozważymy inne, nieco dziwniejsze sposoby mierzenia wielkości wektora. Takie funkcje, przyporządkowujące wektorowi z $\mathbb{R}^d$ liczbę mierzącą w pewien rozsądny sposób jego wielkość, zwane są *normami*. Popularne przykłady norm to: długość wektora (zwana normą euklidesową), suma wartości bezwzględnych jego współrzędnych czy też maksymalna wartość bezwzględna jego współrzędnych (zwana normą maksimum), ale istnieją też inne, bardziej wymyślne normy. W ogólności norma to funkcja przyporządkowująca wektorowi $u \in \mathbb{R}^d$ wartość $\|u\| \in \mathbb{R}$, spełniająca trzy proste warunki: 1) $\|u\| = 0$ wtedy i tylko wtedy, gdy u to wektor zerowy, 2) norma skaluje się liniowo, czyli $\|au\| = |a| \cdot \|u\|$ dla dowolnego $a \in \mathbb{R}$ oraz $u \in \mathbb{R}^d$,
3) $\|u + v\| \leqslant \|u\| + \|v\|$ dla dowolnych $u, v \in \mathbb{R}^d$.

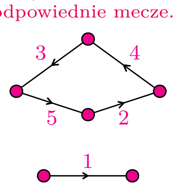
Z faktu, że dla normy euklidesowej $C_d = 2d$ wystarcza, wynika łatwo, że dla dowolnej normy taka stała C_d istnieje, nie wynika jednak w żaden sposób, jak duża jest ta stała. W roku 1931 Borgström wykazał, że dla dowolnej normy optymalna stała C_d jest nie większa niż $\sqrt{(4^d - 1)/3}$. Aż do roku 1978 najlepsza znana stała C_d wciąż była wykładnicza względem d. Dopiero wtedy, w 1978 roku, Sergey Sevastianov opublikował w rosyjskim czasopiśmie w Nowosybirsku dowód uzasadniający, że dla dowolnej normy $C_d = d$ wystarczy. Praca miała dwie strony, przy czym dowód głównego wyniku zajmował w zasadzie jedną stronę, co jest oczywiście nadzwyczajne w wypadku rozwiązania znanego problemu tak długo otwartego. Przetłumaczoną na angielski wersję można znaleźć, wpisując w wyszukiwarce Google frazę „Value of the Steinitz constant”. Mniej więcej rok lub dwa lata temu artykuł ten wpadł mi

Rozwiązanie zadania M 1642.
Odpowiedź: Nie.

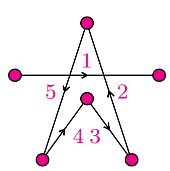
Przypuśćmy, że się udało, i rozważmy sytuację po rozegraniu pierwszej rundy. Nie jest możliwe, że obydwie osoby, które zagrały na ustalonym korcie, na nim zostaną, bo wtedy w drugiej rundzie musiałyby zagrać ze sobą jeszcze raz. Nie jest również możliwe, że obydwaj zawodnicy zmienią kort, bo wówczas w drugiej rundzie znów zagraliby przeciwko sobie (tyle że na innym korcie). Wobec tego na każdym korcie po pierwszej rundzie jeden gracz zostaje (i zmienia połówkę), a jeden zmienia kort. Stąd wniosek, że po drugiej rundzie istnieje zawodnik A, który dotąd grał tylko na pierwszym korcie, i zawodnik B, który dotąd grał tylko na drugim korcie. Ci zawodnicy nie grali jeszcze ze sobą, więc powinni zagrać w trzeciej rundzie – ale nie mają gdzie (gdyby zagrali na pierwszym korcie, A powtórzyłby połówkę, a gdyby na drugim – B powtórzyłby połówkę). Uzyskana sprzeczność oznacza, że nie jest możliwy taki układ rozgrywek.

Rozwiązanie zadania M 1643.
Odpowiedź: Tak.

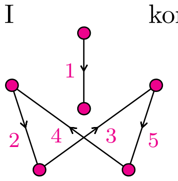
Poniższe obrazki ilustrują przykładowy układ rozgrywek spełniających warunki zadania. Punkty oznaczają zawodników, strzałki – mecze (kierunek strzałki odróżnia na każdym korcie połówkę „wskazującą” od „wskazywanej”), a liczby przy strzałkach to numery rund (od 1 do 5), w których rozgrywane są odpowiednie mecze.



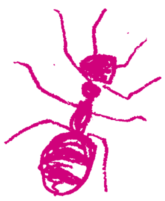
kort I



kort II



kort III



w ręce, ponieważ wynik jest związany z moimi zainteresowaniami naukowymi. Uważam, że dowód jest wyjątkowy. Pál Erdős, jeden z najwybitniejszych matematyków XX wieku często odwoływał się do Księgi, w której Bóg trzyma wszystkie najelegantsze dowody twierdzeń matematycznych. Zainspirowani tym powiedzeniem dwaj matematycy, Eigner i Ziegler, wydali znakomitą książkę „Dowody z Księgi”, którą szczerze polecam każdemu Czytelnikowi. Dowód, o którym mówię, być może również mógłby trafić do takiej Księgi. Mimo że ja i moi koledzy rozumiemy każdy krok tego dowodu z osobna, to nie wiemy, skąd bierze się taki sposób rozumowania, niespotykany nigdzie indziej w naszej dziedzinie. Wydaje się, że za tym rozumowaniem stoi pewna intuicja geometryczna, ale nie wiemy, jaka to jest intuicja. Wierzę, że dogłębne zrozumienie idei ukrytych w tym dowodzie może przyczynić się do kolejnych ciekawych wyników. Kto wie, może ktoś z Czytelników pomoże?

Przedstawiam dowód z oryginalnej pracy Sevastianova nieco przeze mnie zmodyfikowany. Poniżej ustalamy dowolnie wybraną normę wektora $u \in \mathbb{R}^d$, którą oznaczamy $\|u\|$. Przypomnijmy, że naszym celem jest udowodnienie następującego twierdzenia.

Twierdzenie 1. *Jeśli $u_1, \ldots, u_n \in \mathbb{R}^d$, dla każdego $i \in \{1, \ldots, n\}$ zachodzi $\|u_i\| \leqslant 1$ oraz $\sum_{i=1}^n u_i = \vec{0}$, to istnieje permutacja $u'_1, \ldots, u'_n$ ciągu $u_1, \ldots, u_n$ taka, że dla każdego $k \in \{1, \ldots, n\}$ mamy $\|\sum_{i=1}^k u'_i\| \leqslant d$.*

Załóżmy, że $d \leqslant n$, inaczej wniosek jest trywialny. Udowodnimy następujący lemat.

Lemat 1. *Istnieją zbiory $A_d \subseteq A_{d+1} \subseteq \ldots \subseteq A_n = \{1, \ldots, n\}$ oraz wagi $\lambda_k^i \in [0, 1]$ dla $k \in \{d, \ldots, n\}$, $i \in \{1, \ldots, n\}$ takie, że dla dowolnego k mamy $|A_k| = k$, $\sum_{i \in A_k} \lambda_k^i = d$ oraz $\sum_{i \in A_k} \lambda_k^i u_i = \sum_{i \in A_k} u_i$.*

Zobaczmy najpierw, jak z lematu wynika twierdzenie 1. Dla dowolnego $d \leqslant i \leqslant n-1$ zbiór $A_{i+1} \setminus A_i$ ma dokładnie jeden element, nazywamy go u'_{i+1}. Pozostałych d elementów ciągu $u_1, \ldots, u_n$ dowolnie przypisujemy na $u'_1, \ldots, u'_d$. Dla $k \leqslant d$ nierówność z twierdzenia jest oczywista, załóżmy $k > d$. Mamy

$$\Big\|\sum_{i=1}^{k} u'_i\Big\| = \Big\|\sum_{i \in A_k} u_i\Big\| = \Big\|\sum_{i \in A_k} \lambda_k^i u_i\Big\| \leqslant \sum_{i \in A_k} \lambda_k^i \|u_i\| \leqslant \sum_{i \in A_k} \lambda_k^i = d,$$

gdzie druga i ostatnia równość wynikają wprost z lematu. Wystarczy więc udowodnić lemat.

Pokażemy istnienie zbiorów A_i oraz wag λ_k^i spełniających warunki lematu przez indukcję po k. Zacznijmy od bazy indukcji dla $k = n$. Wówczas ustalamy $A_n = \{1, \ldots, n\}$ oraz $\lambda_n^i = \frac{d}{n}$ dla każdego $i \in \{1, \ldots, n\}$, które, jak łatwo sprawdzić, spełniają warunki. Aby wykonać krok indukcyjny z $k+1$ do k, załóżmy, że mamy zdefiniowany zbiór A_{k+1} oraz wagi λ_{k+1}^i, a chcemy zdefiniować zbiór A_k oraz wagi λ_k^i. Niech $A_{k+1} = \{v_1, \ldots, v_{k+1}\} \subseteq \{u_1, \ldots, u_n\}$.

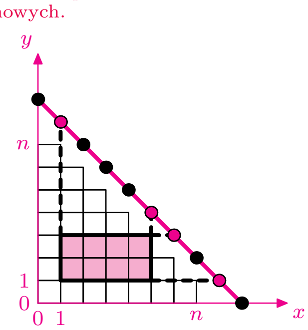
Rozważmy następujący układ równań i nierówności z $k+1$ zmiennymi $\mu_1, \ldots, \mu_{k+1} \in \mathbb{R}$: $0 \leqslant \mu_i \leqslant 1$ dla dowolnego i, $\sum_{i=1}^{k+1} \mu_i = d+1$ oraz $\sum_{i=1}^{k+1} \mu_i v_i = \sum_{i=1}^{k+1} v_i$. Zbiór rozwiązań S tego układu jest niepusty, gdyż, jak nietrudno sprawdzić, rozwiązaniem jest $\mu_i = \lambda_{k+1}^i + (1 - \lambda_{k+1}^i) \cdot \frac{1}{k+1-d}$. W ogólności dla każdego układu m_r równań i m_n nierówności liniowych w $\mathbb{R}^d$ zbiór rozwiązań jest wielościanem, o ile jest ograniczony. Co więcej, okazuje się, że w każdym wierzchołku tego wielościanu dokładnie $d - m_r$ nierówności staje się równościami. Nietrudno w to uwierzyć, bo skoro mamy do czynienia z wierzchołkiem, to liczba równości powinna być równa wymiarowi przestrzeni, a m_r równości mamy już gotowe z równań. Zachęcamy Czytelników do precyzyjnego wykazania tego faktu. Wybierzmy więc dowolny wierzchołek wielościanu S zawierającego rozwiązania naszego układu równań. W układzie mamy $d+1$ równości (równość $\sum_{i=1}^{k+1} \mu_i v_i = \sum_{i=1}^{k+1} v_i$ jest w istocie równością na każdej z d współrzędnych) oraz $2(k+1)$ nierówności. A zatem w wierzchołku S spełnionych jest $k-d$ dodatkowych równości spośród nierówności $0 \leqslant \mu_i \leqslant 1$. Chcemy wykazać, że istnieje j takie, że $\mu_j = 1$. Jedyny przypadek, w którym nie

Rozwiązanie zadania M 1644. Umieśćmy schodkowy trójkąt w układzie współrzędnych w taki sposób, aby środki jego pól były w punktach $(x + \frac{1}{2}, y + \frac{1}{2})$ dla wszystkich par liczb całkowitych nieujemnych x, y takich, że $x + y \leqslant n - 1$.

Zauważmy, że przedłużenia boków dowolnego prostokąta złożonego z całych pól schodkowego trójkąta przecinają prostą $y = n + 2 - x$ w czterech różnych punktach spośród $n+3$ następujących:

$$(0, n+2), \quad (1, n+1), \quad (2, n), \quad \ldots, \quad (n+2, 0).$$

Odwrotnie, każde cztery różne punkty spośród powyższych wyznaczają dokładnie jeden prostokąt – dwa niższe punkty są zawarte w przedłużeniach poziomych boków prostokąta, a dwa wyższe – w przedłużeniach boków pionowych.



Zbudowana wzajemnie jednoznaczna odpowiedniość między szukanymi prostokątami a czwórkami spośród $n+3$ ustalonych punktów świadczy o tym, że odpowiedź na postawione w zadaniu pytanie to

$$\binom{n+3}{4}.$$

Rozwiązanie zadania F 1004.
Proces zamarzania jest powolny, a więc możemy przyjąć, że górna powierzchnia lodu ma temperaturę powietrza -10°C, natomiast dolna, stykająca się z wodą ma temperaturę 0°C, równą temperaturze zamarzającej wody. Ciepło przepływa od cieplejszej wody pod powierzchnią lodu do zimniejszego powietrza nad jego powierzchnią i podczas całego procesu różnica temperatur $\Delta T = -10$ K pozostaje stała, ale rośnie grubość lodu. Powstanie warstwy lodu o grubości dx i polu powierzchni S wymaga odebrania ciepła $dQ = L\rho S \cdot dx$. Szybkość przepływu ciepła jest proporcjonalna do powierzchni, różnicy temperatur ΔT i odwrotnie proporcjonalna do grubości lodu x (tzn. jest proporcjonalna do szybkości zmian temperatury z grubością) i wynosi:

$$\frac{dQ}{dt} = -kS\frac{\Delta T}{x}.$$

Otrzymujemy więc:

$$L\rho S x\frac{dx}{dt} = -kS\Delta T.$$

Oznacza to stałą szybkość zmiany kwadratu grubości warstwy

$$\frac{d}{dt}\left(\frac{x^2}{2}\right) = x\frac{dx}{dt}.$$

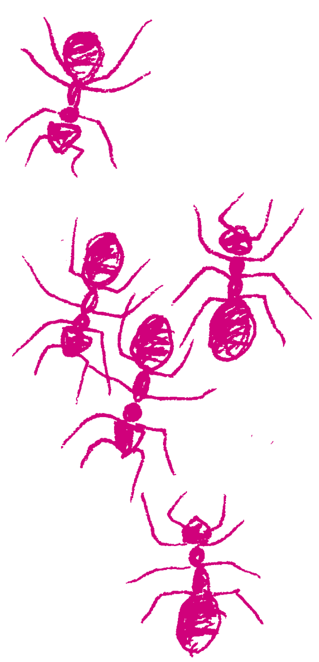
Ostatecznie otrzymujemy dla początkowej grubości $x_p = 5$ cm i końcowej $x_k = 10$ cm:

$$\frac{1}{2}\left(x_k^2 - x_p^2\right) = \frac{-k\Delta T}{L\rho}t$$

i

$$t = \frac{-\left(x_k^2 - x_p^2\right)L\rho}{2k\Delta T}.$$

Po podstawieniu danych liczbowych $t \approx 5{,}22 \cdot 10^4$ s $\approx 14{,}5$ godziny.



jest to natychmiastowe, to gdy każda ze wspomnianych $k-d$ równości jest postaci $\mu_i = 0$. Wiemy jednak, że wówczas suma pozostałych $(k+1)-(k-d) = d+1$ zmiennych μ_i jest równa $d+1$, więc każda z nich musi być równa jeden. A więc tak czy inaczej istnieje j takie, że $\mu_j = 1$. Definiujemy więc $A_k = A_{k+1} \setminus v_j$ oraz $\lambda_k^i = \mu_i$. Nietrudno sprawdzić, że istotnie wszystkie warunki są spełnione, co kończy dowód lematu.

Czy oprócz ładnego dowodu i ciekawej historii oszacowanie na stałą w lemacie Steinitza przydaje się do czegoś? Tak, zdecydowanie, przykładem może być ta praca `https://arxiv.org/abs/1707.00481` opublikowana na konferencji SODA w 2018 roku, jednej z najlepszych światowych konferencji informatycznych. Gwoli ścisłości należy przyznać, że użyta jest tam konkretna norma: *norma maksimum* oznaczana $\|u\|_\infty$, przypomnijmy, że przypisuje ona wektorowi $u \in \mathbb{R}^d$ maksimum z wartości bezwzględnych jego współrzędnych. A więc do tego konkretnego zastosowania wystarczyłby już oryginalny wynik Steinitza z 1914 roku. Faktycznie, z tego, że dla $u \in \mathbb{R}^d$ zachodzi $\|u\|_\infty \leqslant |u| \leqslant \sqrt{d}\|u\|_\infty$ ($|u|$ oznacza długość euklidesową wektora u), oraz tego, że $C_d = 2d$ wystarcza dla normy euklidesowej, wynika, że $C_d = 2d\sqrt{d}$ wystarcza dla normy maksimum. Ja przedstawię pokrótce inne zastosowanie, które wydaje mi się również interesujące, a może być też bardzo użyteczne.

Tunelem pomiędzy punktem $x \in \mathbb{R}^d$ a $y \in \mathbb{R}^d$ o promieniu $s \in \mathbb{R}^+$ nazwijmy zbiór, który zawiera odcinek pomiędzy x a y oraz punkty, które są oddalone od tego odcinka o co najwyżej s, gdzie odległość mierzymy normą maksimum. Precyzyjnie rzecz biorąc, taki tunel to zbiór

$$T = \{z \in \mathbb{R}^d : \exists_{0 \leqslant \alpha \leqslant 1} \|z - \alpha \cdot x - (1-\alpha)\cdot y\|_\infty \leqslant s\}.$$

Po pierwsze zauważmy, że z twierdzenia 1 prosto wynika następujący wniosek.

Wniosek 1. *Jeśli $u_1, \ldots, u_n \in \mathbb{R}^d$ oraz dla każdego $i \in \{1, \ldots, n\}$ zachodzi $\|u_i\|_\infty \leqslant N$, to istnieje permutacja $u_1', \ldots, u_n'$ ciągu $u_1, \ldots, u_n$ taka, że podróż z punktu $x \in \mathbb{R}^d$ do punktu $y = x + \sum_{i=1}^n u_i$ ciągiem $u_1', \ldots, u_n'$ odbywa się wewnątrz tunelu pomiędzy x a y o promieniu $2dN$.*

Zachęcamy Czytelnika do samodzielnego wykazania wniosku, dowód jest nietrudny. Jesteśmy już gotowi do sformułowania twierdzenia.

Twierdzenie 2. *Rozważmy układ n równań liniowych $Mx = y$, gdzie M jest macierzą o współczynnikach całkowitych $n \times d$, a y wektorem z $\mathbb{Z}^d$. Niech N będzie maksimum z wartości bezwzględnych liczb występujących w macierzy M i wektorze y. Wówczas jeśli istnieje pewne rozwiązanie tego układu w liczbach naturalnych, to istnieje również rozwiązanie $x \in \mathbb{N}^n$ takie, że $\|x\|_\infty \leqslant (5dN+1)^d$.*

Zauważmy, że ograniczenie na normę rozwiązania nie zależy od liczby równań n, a jedynie od liczby zmiennych d i wielkości liczb N, i to jest właśnie główna siła twierdzenia 2. Aby udowodnić twierdzenie 2, oznaczmy kolumny macierzy M jako wektory $u_1, \ldots, u_d$ oraz $x = (x_1, \ldots, x_d)$. Wówczas równanie $Mx = y$ przyjmuje postać $\sum_{i=1}^d u_i x_i = y$. Rozważmy pewne rozwiązanie tego układu, które istnieje zgodnie z założeniem twierdzenia 2. Zawiera ono x_1 wektorów u_1, x_2 wektorów u_2 itd., aż w końcu x_n wektorów u_n. Zgodnie z wnioskiem istnieje taka permutacja $x_1', \ldots, x_m'$ tych $m = x_1 + \ldots + x_n$ wektorów, że cała podróż z punktu 0 do punktu y ciągiem $x_1', \ldots, x_m'$ odbywa się wewnątrz tunelu z 0 do y o promieniu $2dN$. Taki tunel zawiera się cały w d-wymiarowej kostce o boku $4dN + N \leqslant 5dN$, na potrzeby dowodu twierdzenia 2 wystarczy nam takie zgrubne oszacowanie. Zauważmy, że jeśli w trakcie naszej podróży odwiedzimy dwa razy ten sam punkt, to możemy tę podróż skrócić, pomijając pętlę wychodzącą i wracającą do tego samego punktu, nie zmieni to oczywiście celu podróży. Postępując w ten sposób, możemy skrócić podróż ciągiem $x_1', \ldots, x_m'$ do takiej podróży, która każdy punkt w kostce o boku $5dN$ odwiedza co najwyżej jeden raz. Punktów o współczynnikach całkowitych w kostce jest nie więcej niż $(5dN+1)^d$, a więc nasza podróż będzie miała co najwyżej tyle kroków. Nietrudno zauważyć, że dowolna taka podróż natychmiast daje rozwiązanie równania $\sum_{i=1}^d u_i x_i = y$, gdzie $\sum_{i=1}^n x_i \leqslant (5dN+1)^d$, co kończy dowód twierdzenia 2.