Арифметический минимум квадратичной формы и сферические коды

Н. Н. Андреев В. А. Юдин

Посвящается 150-летию со дня рождения Е. И. Золотарева (1847–1878).

В этой статье мы расскажем об одной старой задаче геометрии чисел: *Как много точек с целыми координатами могут располагаться на поверхности эллипсоида в d-мерном евклидовом пространстве, если внутри него нет точек с целыми координатами, за исключением его центра — нуля?* Сформулируем задачу более точно.

Пусть $\mathbb{Z}^d$ — решётка целых чисел в $\mathbb{R}^d$, т. е. множество точек, у которых все координаты — целые числа; точку $x \in \mathbb{Z}^d$ будем называть целой точкой. Обозначим через $xy = x_1y_1 + \dots + x_dy_d$ скалярное произведение векторов $x = (x_1, \dots, x_d)$ и $y = (y_1, \dots, y_d)$, а через $|x| = \sqrt{xx}$ — норму вектора x. Пусть $Axx = \sum_{i,j=1}^{d} a_{ij}x_ix_j$ — положительно определённая квадратичная форма (т. е. $Axx > 0$ для всех $0 \neq x \in \mathbb{R}^d$), порожденная матрицей $A = (a_{ij})_{i,j=1}^{d}$ с определителем $D(A)$. Число

$$\gamma(A) = \inf_{x \in \mathbb{Z}^d \setminus \{0\}} Axx$$

называется арифметическим минимумом квадратичной формы, а величина

$$N(d, A) = \sum_{x \in \mathbb{Z}^d \,:\, Axx = \gamma(A)} 1$$

называется числом представлений ее минимума. Можно показать, что положительная определённость формы Axx гарантирует достижение инфимума $\gamma(A)$.

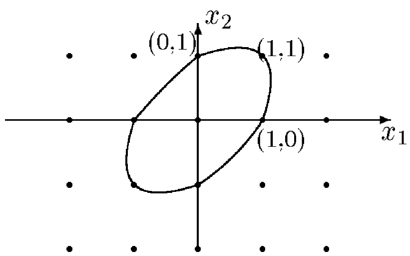
Используя введенные обозначения, интересующую нас задачу можно сформулировать следующим образом: требуется найти числа

$$N_d = \sup_{A} N(d, A),$$

где супремум берётся по множеству всех положительно определённых матриц A.

Математическое просвещение, сер. 3, вып. 2, 1998(133–140)

Так, при $d = 2$ требуется расположить эллипс с центром в начале координат так, чтобы внутри него целых точек не было, а на границе их количество стало максимальным. Докажем, что $N_2 = 6$.



Это значение достигается на квадратичной форме

$$Axx = x_1^2 - x_1x_2 + x_2^2.$$

Арифметический минимум $\gamma(A) = 1$ достигается в шести точках $\pm(1,0)$, $\pm(0,1)$, $\pm(1,1)$.

Покажем теперь, что $N_2 \leqslant 6$. Ввиду симметрии эллипса относительно начала координат для каждой точки эллипса противоположная ей также принадлежит эллипсу. Пусть $\pm x^1, \ldots, \pm x^q$ — все целые точки, на которых достигается минимальное значение формы A (*минимальные целые точки*). Выберем из каждой пары противоположных минимальных точек по одной, обозначим их $x^1, \ldots, x^q$ и разобьем на 4 класса: к классу A_0 отнесем точки, у которых обе координаты — чётные; к классу A_1 — точки, у которых первая координата нечётная, а вторая чётная; класс A_2 будет состоять из точек, у которых первая координата чётная, а вторая нечётная; класс A_3 — из точек, у которых обе координаты нечётные. Каждый класс A_i, $i = 0, 1, 2, 3$ состоит не более, чем из одной минимальной точки. Действительно, допустим, что какой-то из классов содержит не меньше двух точек, например $x, y \in A_i$, тогда их полусумма $\frac{x+y}{2}$ также будет отличной от нуля целой точкой. Ввиду строгой выпуклости эллипса она будет расположена внутри него, что невозможно, так как по условию внутри эллипса нет целых точек, кроме нуля. В классе A_0 вообще нет ни одной минимальной точки (если $x \in \mathbb{Z}^2$, $x \neq 0$, $x \in A_0$, то $\frac{1}{2}x \in \mathbb{Z}^2$, чего не может быть). Таким образом, $q \leqslant 3$ или $N_2 \leqslant 6$. Приведенный выше пример показывает достижимость полученной оценки.

Г. Ф. Вороной провёл аналогичное рассуждение в $\mathbb{R}^d$ и получил оценку $N_d \leqslant 2(2^d - 1)$, однако при $d \geqslant 3$ она оказывается грубой. Так, при $d = 3$ получается оценка $N_3 \leqslant 14$, а на самом деле $N_3 = 12$, как будет показано ниже. Из работы А. Н. Коркина и Е. И. Золотарева [2] непосредственно вытекает оценка снизу $N_d \geqslant d(d+1)$. (Результаты Г. Ф. Вороного, А. Н. Коркина и Е. И. Золотарева приведены в [3].)

Что касается точных значений числа N_d, то при $d \leqslant 6$ их можно получить из работы [2] и из работы Барнса 1957 года. Было доказано, что квадратичные формы, на которых достигается максимум N_d, следует искать среди так называемых предельных квадратичных форм, т.е. таких, для которых постоянная Эрмита

$$\gamma_d = \sup_A \frac{\gamma(A)}{\sqrt[d]{D(A)}}$$

принимает наибольшее возможное значение. Задача вычисления чисел γ_d непосредственно связана с вопросом наиплотнейшей решётчатой упаковки шаров в евклидовом пространстве и с задачей, рассматриваемой в этой статье, как мы увидим позже. Исследования начинаются работами Гаусса и Лагранжа, которые показали, что $\gamma_2 = 2/\sqrt{3}$, $\gamma_3 = \sqrt[3]{2}$. В 1872 году А. Н. Коркин и Е. И. Золотарев в их первой совместной работе [1] доказали, что $\gamma_4 = \sqrt{2}$, а в работе 1877 года вычислили $\gamma_5 = \sqrt[5]{8}$. К настояшему времени найдены точные значения γ_d до $d = 8$ (в двух работах 1925 и 1935 годов Блихфельд показал, что $\gamma_6 = \sqrt[6]{64/3}$, $\gamma_7 = \sqrt[7]{64}$, $\gamma_8 = 2$). Крайне интересным является вопрос о поведении γ_d при $d \to \infty$. Подробно полученные результаты изложены в [4].

В [5] Ватсон упростил способы нахождения чисел N_d для $d \leqslant 6$ и нашёл точные значения для $d = 7, 8, 9$, проводя достаточно длинные вычисления с квадратичными формами. В конце 70-х годов В. И. Левенштейном и независимо Н. Слоэном и А. Одлыжко было показано, что $N_8 = 240$ и $N_{24} = 196560$. После этих работ таблица известных точных значений N_d приняла вид

d	2	3	4	5	6	7	8	9	24
N_d	6	12	24	40	72	126	240	272	196560

В остальных размерностях точные значения не известны.

Приведем аналитический способ получения оценок сверху чисел N_d. Вначале дадим несколько определений.

Пусть $e_1, \ldots, e_d$ — линейно-независимая система векторов из $\mathbb{R}^d$. Решеткой называется множество

$$L = \{k_1 e_1 + k_2 e_2 + \cdots + k_d e_d\}_{k=(k_1,k_2,\ldots,k_d)\in\mathbb{Z}^d}.$$

Вектора, для которых

$$|x| = \inf_{x\in L\setminus\{0\}} |x| \stackrel{\mathrm{def}}{=} \mu_1(L),$$

называются минимальными векторами решётки, совокупность этих векторов обозначим V. Очевидно, что для каждого вектора $x \in V$ противоположный вектор $-x$ также лежит в V. Теперь нетрудно заметить, что для любых двух различных векторов $x, y \in V$ выполнено неравенство $xy \leqslant 1/2\mu_1^2(L)$, т. е. угол между ними не менее $60°$. Без ограничения общности будем считать $\mu_1(L) = 1$.

Возвращаясь к нашей задаче, сделаем линейную замену переменных $x = Ty$, приводящую исходную квадратичную форму к виду $y_1^2 + \dots + y_d^2$. Тем самым вопрос свёлся к поиску в d-мерном евклидовом пространстве решётки L, у которой максимальное количество минимальных векторов.

Вследствие перечисленных выше свойств множества V эта задача оказывается тесно связанной с двумя другими интересными задачами.

Какое максимальное количество точек B_d может иметь симметричный сферический $1/2$-код в размерности d? Другими словами, *какое максимальное количество точек можно разместить на единичной сфере S^{d-1} с условиями, что каждая точка имеет противоположную и что модуль скалярного произведения любых двух различных из них и не противоположных не превосходит $1/2$?*

Найти контактное число M_d шаров в размерности d, т. е. максимальное количество шаров одинакового радиуса, которые могут касаться одного данного шара. Иначе говоря — *какое максимальное количество точек можно расположить на единичной сфере так, чтобы скалярное произведение любых двух из них не превосходило $1/2$?*

Точное значение M_d до недавнего времени было известно только при $d = 2$ $(M_2 = 6)$ и $d = 3$ $(M_3 = 12)$. Наилучшие известные оценки этой величины в зависимости от размерности приведены в [6, т.1, с.42].

Понятно, что

$$N_d \leqslant B_d \leqslant M_d. \tag{1}$$

Мы будем оценивать сверху величину B_d и тем самым получать оценку сверху для количества минимальных векторов в решетке размерности d. При этом мы воспользуемся идеями, предложенными в 1968 году П. Дельсартом. Именно он стал использовать в геометрических задачах положительную определённость. Итак, через $\{P_k^d(t)\}_{k=1}^{\infty}$ обозначим систему многочленов Гегенбауэра с нормировкой $P_k^d(1) = 1$:

$$P_0^d(t) = 1, \quad P_1^d(t) = t, \quad P_2^d(t) = \frac{dt^2 - 1}{d - 1}, \quad P_3^d(t) = \frac{(d+2)t^3 - 3t}{d - 1}, \dots$$

$$(k + d - 2)P_{k+1}^d(t) = (2k + d - 2)tP_k^d(t) - kP_{k-1}^d(t).$$

На интервале $(-1;1)$ они образуют ортогональную систему многочленов с весом $(1-t^2)^{\frac{d-3}{2}}$, т. е.

$$\int_{-1}^{1} P_k^d(t)P_l^d(t)(1-t^2)^{\frac{d-3}{2}} = 0, \quad k \neq l.$$

В геометрических задачах используется их положительная определённость: для любого конечного множества точек $x^{(1)}, \ldots, x^{(N)}$ из S^{d-1}, любого $s \in \mathbb{N}$ и любых p_k, $p_l \in \mathbb{R}$ справедливо неравенство

$$\sum_{k,l=1}^{N} P_s^d(x^{(k)}x^{(l)})p_k p_l \geqslant 0. \tag{2}$$

Пусть в d-мерном евклидовом пространстве нам дан симметричный сферический $1/2$-код, состоящий из N точек, т. е. множество$\{x^{(i)}\}_{i=1}^N$ точек в S^{d-1} таких, что $-1/2 \leqslant x^{(i)}x^{(j)} \leqslant 1/2$ при $x^{(i)} \neq \pm x^{(j)}$. Рассмотрим непрерывную на отрезке $[-1;1]$ функцию $h(t)$ такую, что $h(t) \leqslant 0$ при $t \in [-1/2;1/2]$ и все ее коэффициенты Фурье по системе многочленов Гегенбауэра неотрицательны:

$$h(t) = \sum_{k=0}^{\infty} \widehat{h}_k P_k^d(t), \quad \widehat{h}_k \geqslant 0,\ \widehat{h}_0 > 0.$$

В такой ситуации имеем:

$$I = \sum_{k,l=1}^{N} h(x^{(k)}x^{(l)}) = \sum_{s=0}^{\infty} \widehat{h}_s \sum_{k,l=1}^{N} P_s^d(x^{(k)}x^{(l)}) \geqslant \widehat{h}_0 \sum_{k,l=1}^{N} 1 = N^2 \widehat{h}_0,$$

где неравенство верно вследствие неотрицательности коэффициентов $\widehat{h}_s$ и положительной определённости многочленов Гегенбауэра. С другой стороны, так как $h(x^{(k)}x^{(l)}) \leqslant 0$ при $x^{(k)} \neq \pm x^{(l)}$, то

$$\begin{aligned} I &= \sum_{k,l=1}^{N} h(x^{(k)}x^{(l)}) = \\ &= \sum_{x^{(k)}=x^{(l)}} h(x^{(k)}x^{(l)}) + \sum_{x^{(k)}=-x^{(l)}} h(x^{(k)}x^{(l)}) + \sum_{x^{(k)}\neq \pm x^{(l)}} h(x^{(k)}x^{(l)}) \leqslant \\ &\leqslant N(h(1)+h(-1)). \end{aligned}$$

Таким образом

$$N \leqslant \frac{h(1)+h(-1)}{\widehat{h}_0}. \tag{3}$$

Подбирая нужным образом функцию $h(t)$, получаем оценки сверху. Заметим, что число N натуральное, поэтому оно на самом деле оценивается целой частью выражения, стоящего в правой части неравенства (3). Подобным способом получены точные оценки сверху для чисел M_8 и M_{24} и тем самым для чисел N_8 и N_{24}, а решётки с соответствующим количеством минимальных векторов были уже известны: в случае $d = 8$ экстремальной оказалась решётка Коркина-Золотарева E_8, а в случае $d = 24$ — решётка Лича.

Используем неравенство (3) для оценки мощности симметрического 1/2-кода. Рассмотрим многочлен

$$h(t) = t^2(t^2 - \frac{1}{4}).$$

Напишем его разложение в ряд Фурье:

$$h(t) = \frac{d^2-1}{(d+2)(d+4)} P_4^d(t) + \frac{(20-d)(d-1)}{4d(d+4)} P_2^d(t) + \frac{10-d}{4d(d+2)} P_0^d(t). \quad (4)$$

Все его коэффициенты положительны при $2 \leqslant d \leqslant 9$. Следовательно, из неравенства (3) найдем

$$N_d \leqslant \frac{6d(d+2)}{10-d}.$$

Эта оценка оказывается точной в размерностях $d = 3, 4, 6, 7, 8$. Аналогичные (4) многочлены можно строить и в высших размерностях. Рассматривая многочлены

$$h(t) = t^4(t^2 - \frac{1}{4}), \quad h(t) = t^6(t^2 - \frac{1}{4}), \quad h(t) = t^2(t^2 - \frac{1}{16})^2(t^2 - \frac{1}{4}),$$

получаем следующие оценки сверху мощности симметрического 1/2-кода и вследствие неравенства (1) — числа минимальных векторов в решётке:

d	3	4	5	6	7	8	9	10	11	12	13
$N_d \leqslant$	12	24	42	72	126	240	367	560	858	1344	2210

d	17	18	19	20	21	22	23	24
$N_d \leqslant$	11683	16298	22866	32445	46947	70200	111136	196560

Приведем примеры наилучших квадратичных форм и решёток, для которых оценки, представленные в таблице, при $d \leqslant 8$ достигаются. Для

$d=2$ ранее уже был дан очевидный пример квадратичной формы $x_1^2 \pm \pm x_1x_2 + x_2^2$. Для $d=3$ квадратичная форма имеет вид $Axx = x_1^2 + x_2^2 + x_3^2 + x_1x_2 + x_1x_3 + x_2x_3$. Легко устанавливается (выделением полных квадратов), что она положительно определена.

Кроме того, она целочисленна: $Axx \in \mathbb{Z}$ для любого $x \in \mathbb{Z}^3$. Значит, ее арифметический минимум не меньше 1, т. е. $\min_{x\in\mathbb{Z}^3\setminus\{0\}} Axx \geqslant 1$. На самом деле он равен 1 и достигается в 12 точках $\pm(1,0,0)$, $\pm(0,1,0)$, $\pm(0,0,1)$, $\pm(1,-1,0)$, $\pm(1,0,-1)$, $\pm(0,1,-1)$.

Эту последовательность можно продолжить, но нам удобнее иметь дело с решётками (что по смыслу одно и то же). Это даст нам возможность заметить одно весьма любопытное явление: «сечения» минимальных векторов решётки Коркина-Золотарева из $\mathbb{R}^8$ дают экстремальные конструкции для всех меньших размерностей.

Минимальные вектора решётки Коркина-Золотарева E_8 состоят из двух групп векторов. В первую входят 128 векторов вида $\frac{1}{\sqrt{8}}(\pm1,\pm1,\ldots,\pm1)$ с чётным числом (0,2,4,6,8) плюсов. Во вторую входят вектора, у которых две произвольные координаты равны $\pm1/\sqrt{2}$, а остальные — 0. Их количество $C_8^2\cdot 4 = 112$. В сумме получим 240 векторов. Их совокупность (обозначим ее через W_8) и есть экстремальный набор векторов при $d=8$.

Замечая, что «сечение» решётки гиперплоскостью $ax=0$ снова есть решётка, а минимальные вектора первоначальной решётки становятся минимальными векторами «сечения», возьмём сечение в $\mathbb{R}^8$ решётки Коркина-Золотарева гиперплоскостью $ax=0$, $a=(1,1,\ldots,1)$ и положим

$$W_7 = \{x \in W_8 \ : \ x_1+x_2+x_3+x_4+x_5+x_6+x_7+x_8 = 0\}.$$

Подсчитаем количество элементов в W_7. Из первой группы минимальных векторов решётки Коркина-Золотарева в W_7 войдут лишь те вектора, у которых 4 координаты равны $+1$ и четыре равны -1. Их количество равно $C_8^4 = 70$. Из второй группы векторов лишь половина (координаты имеют разные знаки) удовлетворяет уравнению $ax=0$. Таким образом, W_7 содержит $70+56=126$ векторов и является экстремальным набором векторов при $d=7$.

Аналогично можно доказать, что множества

$$W_6 = \{x \in W_8 \ : \ x_1+x_2=0,\ x_3+x_4+x_5+x_6+x_7+x_8=0\},$$
$$W_5 = \{x \in W_8 \ : \ x_1+x_2=0,\ x_3+x_4+x_5+x_6=0,\ x_7+x_8=0\},$$
$$W_4 = \{x \in W_8 \ : \ x_1+x_2=0,\ x_3+x_4=0,\ x_5+x_6=0,\ x_7+x_8=0\},$$

являясь наборами минимальных векторов решёток в $\mathbb{R}^6$, $\mathbb{R}^5$, $\mathbb{R}^4$, содержат соответственно 72, 40 и 24 вектора.

В заключение отметим, почему предложенный метод получения оценок сверху чисел N_d «грубит», например, при $d = 5$. Полученная оценка $N_5 \leqslant 42$ не является точной — на самом деле $N_5 = 40$. Чтобы ответить на этот вопрос, надо проанализировать предложенное нами доказательство. При $d = 5$ оценка оказывается точной в тех случаях, когда неравенство (2) обращается в равенство для экстремальной конструкции при $s = 1, 2, 3, 4$. Однако этого не происходит для W_5: $\sum_{x,y \in W_5} P_4^5(xy) > 0$. Аналогичная ситуация и при $d = 3$, но тут «везёт» в вычислениях: $[\frac{90}{7}] = 12$.

Итак, для размерностей $d = 3$, 4, 6, 7, 8, 24 мы привели, как нам кажется, наиболее простой метод нахождения чисел N_d. При $d = 10$, 11, 12, 13, 17, 18, 19, 20, 21, 22, 23 получены новые оценки сверху для количества минимальных векторов решёток в этих размерностях.

Список литературы

[1] *Korkine A., Zolotareff G.* Sur les formes quadratiques positives quaternaires // Math. Ann. V. 5. 1872. P. 581–583.
Русск. пер.: *Золотарев Е. И.* Полное собр. соч., вып. 1. Изд. АН СССР, 1931.

[2] *Korkine A., Zolotareff G.* Sur les formes quadratiques positives // Math. Ann. V.11. 1877. P. 242–292.
Русск. пер.: там же.

[3] *Делоне Б. Н.* Петербургская школа теории чисел. М.-Л.: Изд. АН СССР, 1947.

[4] *Рышков С. С., Барановский Е. П.* Классические методы теории решетчатых упаковок // УМН. Т. 34, вып. 4. 1974. С. 3–63.

[5] *Watson G. L.* The number of minimum points of a positive quadric form // Dissertationes mathematicae. LXXXIV. 1971, pp. 2–42.

[6] *Конвей Дж., Слоэн Н.* Упаковки шаров, решетки и группы. М.: Мир, 1990.