О среднем числе шагов в алгоритме Евклида с выбором минимального по модулю остатка

А. В. Устинов*

Классический алгоритм Евклида, в котором при делении выбирается наименьший неотрицательный остаток

$$a = bq + r, \quad q = [a/b], \quad 0 \le r < q,$$

соответствует разложению числа в стандартную цепную дробь:

$$\frac{a}{b} = t_0 + \cfrac{1}{t_1 + \ddots + \cfrac{1}{t_s}}$$

длины $s = s(a/b)$, в которой t_0 — целое, $t_1, \ldots, t_s$ — натуральные и $t_s \ge 2$ при $s \ge 1$. Алгоритм Евклида с выбором минимального по модулю остатка

$$a = bq + r, \quad q = \left[\frac{a}{b} + \frac{1}{2}\right], \quad -\frac{q}{2} \le r < \frac{q}{2},$$

приводит к разложению в дробь

$$\frac{a}{b} = t_0 + \cfrac{\varepsilon_1}{t_1 + \cfrac{\varepsilon_2}{t_2 + \ddots + \cfrac{\varepsilon_l}{t_l}}}, \tag{1}$$

длины $l = l(a/b)$, где t_0 — целое, $t_1, \ldots, t_l$ — натуральные,

$$\varepsilon_k = \pm 1, \quad t_k \ge 2 \quad (k = 1, \ldots, l), \quad t_k + \varepsilon_{k+1} \ge 2 \quad (k = 1, \ldots, l-1).$$

Существует простой алгоритм (см. [4, § 39]), который превращает обычную цепную дробь в дробь вида (1). К первому неполному частному t_j ($j \ge 1$), равному единице, нужно применить тождество

$$t_{j-1} + \cfrac{1}{1 + \cfrac{1}{t_{j+1} + \ldots}} = t_{j-1} + 1 - \frac{1}{t_{j+1} + 1 + \ldots}.$$

То есть первая единица в разложении вычеркивается, соседние неполные частные увеличиваются на единицу, и между ними ставится знак “минус”. Затем

*Работа выполнена при поддержке фонда РФФИ, грант № 07-01-00306 и проекта ДВО РАН 06-III-A-01-017

находится следующая единица и процедура повторяется. Если в разложении имеется цепочка из подряд идущих единиц, то преобразование применяется к единицам, стоящим на нечетных местах в этой цепочке. Например,

$$[0; 2, 1, 3, 1, 1, 6] = \cfrac{1}{3 - \cfrac{1}{5 - \cfrac{1}{2 + 1/6}}}.$$

Для среднего числа шагов в алгоритме Евклида известны следующие результаты (см. [1, 5]):

$$\frac{2}{R(R+1)} \sum_{b \le R} \sum_{a \le b} s(a/b) = \frac{2 \log 2}{\zeta(2)} \cdot \log R + C_1 + O(R^{-1} \log^4 R), \tag{2}$$

$$\frac{1}{\varphi(b)} \sum_{\substack{1 \le a \le b \\ (a,b)=1}} s(a/b) = \frac{2 \log 2}{\zeta(2)} \cdot \log b + C_2 + O_\varepsilon(b^{-1/6+\varepsilon}), \tag{3}$$

с абсолютными константами

$$C_1 = \frac{2 \log 2}{\zeta(2)} \left(\frac{3 \log 2}{2} + 2\gamma - \frac{\zeta'(2)}{\zeta(2)} - \frac{3}{2} \right) - \frac{3}{2},$$

$$C_2 = C_1 + \frac{2 \log 2}{\zeta(2)} \left(\frac{1}{2} - \frac{\zeta'(2)}{\zeta(2)} \right),$$

где γ — постоянная Эйлера. В то же время для среднего числа шагов в алгоритме Евклида с выбором минимального по модулю остатка известна лишь формула

$$\frac{2}{R(R+1)} \sum_{b \le R} \sum_{a \le b} l(a/b) = \frac{2 \log \varphi}{\zeta(2)} \cdot \log R + C_3 + O(R^{-\beta}),$$

где $\varphi = (1 + \sqrt{5})/2$ — золотое сечение, C_3 — абсолютная постоянная и $\beta > 0$. (см. [3]). Оказывается, что равенства аналогичные (2) и (3) справедливы и для величины $l(a/b)$.

Для фиксированного $x \in (0, 1]$ и рационального $r = [t_0; t_1, \ldots, t_s]$ через $s_x(r)$ будем обозначать статистики Гаусса-Кузьмина

$$s_x(r) = \#\{j = 1, \ldots, s : [0; t_j, \ldots, t_s] \le x\}.$$

В частности, $s_1(r) = s(r)$ — длина цепной дроби для числа r.

Лемма . *Для любого рационального числа a/b*

$$l(a/b) = s_{\varphi-1}(a/b).$$

Доказательство. Пусть $a/b = [t_0; t_1, \ldots, t_s]$. Обозначим через $s'(a/b)$ количество остатков $r_j = [0; t_j, \ldots, t_s]$ $(1 \le j \le s)$, разложение которых начинается с нечетного числа единиц. Соответственно $s''(a/b)$ будет обозначать количество остатков, начинающихся с четного (возможно нулевого) числа единиц. Очевидно, выполняется равенство $s(a/b) = s'(a/b) + s''(a/b)$.

Согласно приведенному выше алгоритму, при переводе обычной цепной дроби в дробь вида (1) каждый отрезок из k подряд идущих единиц заменяется на $[k/2]$ неполных частных. Поэтому исчезает в точности $s'(a/b)$ неполных частных:

$$l(a/b) = s(a/b) - s'(a/b) = s''(a/b).$$

Но цепная дробь для числа $r \in (0,1)$ начинается с четного числа единиц тогда и только тогда, когда $r \in (0, \varphi - 1)$. Следовательно, $s''(a/b) = s_{\varphi-1}(a/b)$. $\square$

Теорема 1. *При любом* $R \geq 2$

$$\frac{2}{R(R+1)} \sum_{b \leq R} \sum_{a \leq b} l(a/b) = \frac{2 \log \varphi}{\zeta(2)} \cdot \log R + C_3 + O(R^{-1} \log^4 R),$$

Доказательство. Формула (2) обобщается на случай статистик Гаусса-Кузьмина (см. [1])

$$\frac{2}{R(R+1)} \sum_{b \leq R} \sum_{a \leq b} s_x(a/b) = \frac{2}{\zeta(2)} \left(\log(1+x) \log R + C_1(x)\right) + O(R^{-1} \log^4 R), \quad (4)$$

где при $x < 1$

$$C_1(x) = \log(1+x) \left(2\gamma - \frac{\zeta'(2)}{\zeta(2)} - \frac{\log(1+x)}{2} + \log x - \frac{3}{2}\right) +$$
$$+ h_1(x) + h_2(x) + \frac{\zeta(2)}{2} \cdot \frac{x^2}{x+1},$$

а функции $h_1(x)$ и $h_2(x)$ задаются абсолютно сходящимися рядами

$$h_1(x) = \sum_{n=1}^{\infty} \frac{1}{n} \left(\sum_{m=1}^{n} \frac{x}{n+mx} - \log(1+x) \right), \quad h_2(x) = \sum_{n=1}^{\infty} \frac{1}{n} \left(\sum_{\frac{n}{x} \leq m < \frac{n}{x}+n} \frac{1}{m} - \log(1+x) \right).$$

Подставляя в (4) равенство из леммы, получаем утверждение теоремы с константой

$$C_3 = \frac{2}{\zeta(2)} C_1(\varphi) = \frac{2 \log \varphi}{\zeta(2)} \left(2\gamma - \frac{\zeta'(2)}{\zeta(2)} - \frac{3}{2} \log \varphi - \frac{3}{2}\right) + \frac{2}{\zeta(2)} (h_1(\varphi) + h_2(\varphi)) + \frac{1}{\varphi^3}.$$

$\square$

Теорема 2. *При любом* $b \geq 2$ *выполняется равенство*

$$\frac{1}{\varphi(b)} \sum_{\substack{1 \leq a \leq b \\ (a,b)=1}} l(a/b) = \frac{2 \log \varphi}{\zeta(2)} \cdot \log b + C_4 + O_\varepsilon(b^{5/6} \log^{7/6+\varepsilon} b)$$

с константой

$$C_4 = C_3 + \frac{2 \log \varphi}{\zeta(2)} \left(\frac{1}{2} - \frac{\zeta'(2)}{\zeta(2)} \right).$$

Доказательство. Подставляя равенство из леммы в обобщение результата Портера (3) (см. [2]))

$$\frac{1}{\varphi(b)}\sum_{\substack{1\le a\le b\\(a,b)=1}} s_x(a/b)=\frac{2}{\zeta(2)}\left(\log(x+1)\log b+C_2(x)\right)+O_{\varepsilon,x}(b^{5/6}\log^{7/6+\varepsilon}b),$$

где

$$C_2(x)=C_1(x)+\log(1+x)\left(\frac{1}{2}-\frac{\zeta'(2)}{\zeta(2)}\right),$$

получаем утверждение теоремы.

□

Замечание. Рассмотрим дисперсию величины $l(a/b)$

$$D_l(R)=\frac{2}{R(R+1)}\sum_{d\le R}\sum_{c\le d}\left(l(c/d)-E_l(R)\right)^2,$$

где $E_l(R)$ — математическое ожидание, стоящее в левой части равенства (2). Одним из результатов работы [3] является асимптотическая формула

$$D_l(R)=D_1\cdot\log R+D_0+O(R^{-\beta}), \tag{5}$$

с абсолютными постоянными D_1, D_0 и некоторым $\beta>0$. Для аналогично определяемой дисперсии $D_s(R)$ величины $s(a/b)$ известен более точный результат (см. [1])

$$D_s(R)=D_1'\cdot\log R+D_0'+O_\varepsilon(R^{-1/4+\varepsilon}).$$

Его доказательство остается справедливым, если вместо длины цепной дроби $s(a/b)$ рассматривать статистики Гаусса-Кузьмина $s_x(a/b)$. Отсюда, с учетом доказанной леммы, следует, что равенство (5) справедливо при любом $\beta<1/4$.

Список литературы

[1] Устинов А. В. Асимптотическое поведение первого и второго моментов для числа шагов в алгоритме Евклида. — *Изв. РАН*, в печати.

[2] Устинов А. В. О числе решений сравнения $xy\equiv l \pmod q$ под графиком дважды непрерывно дифференцируемой функции. — *Алгебра и анализ*, в печати.

[3] Baladi V., Vallée B. Euclidean algorithms are Gaussian. — *J. Number Theory*, **110** (2005), 331–386.

[4] Perron O. *Die Lehre von den Kettenbruechen (Band 1).* — Stuttgart: B.G. Teubner Verlagsgesellschaft, 1954.

[5] Porter J. W. On a theorem of Heilbronn. — *Mathematika*, 1975, v. 22, № 1, 20–28.