Владимирова Ю.С., Рамиль Альварес Х.

Алгоритм деления по модулю в симметричных системах счисления

В ряде публикаций (например, [1]-[3]) приведены алгоритмы деления по модулю в троичной симметричной системе счисления. В настоящей статье рассматривается алгоритм деления целых чисел по модулю, известный как деление «в столбик» или «цифра за цифрой», в симметричных системах счисления с произвольным основанием.

Различие алгоритмов деления по модулю в симметричных системах счисления и в системах счисления с неотрицательными цифрами обусловлено тем, что операция деления по модулю определяется в этих системах счисления по-разному. В системах счисления с неотрицательными цифрами делением числа A на число D по модулю с остатком называется представление A в виде:

$$A = B \cdot D + R, \qquad \text{где } 0 \le R < D. \tag{1}$$

Число B называется неполным частным, R – остатком. В симметричных системах счисления делением числа a на число d по модулю с остатком также называется представление a в виде:

$$a = b \cdot d + r, \tag{2}$$

но остаток от деления r может быть как положительным, так и отрицательным числом, удовлетворяющим неравенству

$$-\frac{|d|}{2} \le r \le \frac{|d|}{2} \tag{3}$$

Если d – нечетное число, то $\frac{|d|}{2}$ не является целым, и (3) можно записать как:

$$-\frac{|d|-1}{2} \le r \le \frac{|d|-1}{2};$$

количество целых чисел, удовлетворяющих этому условию равно d. При d четном, $\frac{|d|}{2}$ – целое число, поэтому количество целых чисел, удовлетворяющих (3) составляет $d+1$. Избыточность возникает вследствие того, что при $|r| = \frac{|d|}{2}$ допустимы два представления числа в виде (2):

$$a = b \cdot d + \frac{|d|}{2},$$

при этом можно считать, что r – остатки деления на d во всех случаях удовлетворяют условию:

$$-\frac{|d|}{2} < r \le \frac{|d|}{2}, \tag{4}$$

либо

$$a = (b+1)\cdot d - \frac{|d|}{2},$$

при этом можно считать, что все r удовлетворяют условию:

$$-\frac{|d|}{2} \le r < \frac{|d|}{2}. \tag{4'}$$

Например, деление по модулю числа 22 на 4 в системах счисления с неотрицательными цифрами означает представление 22 в виде:

$$22 = 5\cdot 4 + 2.$$

В симметричных системах счисления допустимо еще одно представление:

$$22 = 6\cdot 4 - 2,$$

в котором абсолютная величина остатка равна половине делителя, а знак противоположен знаку произведения делителя и неполного частного.

Приведем алгоритм деления по модулю для систем счисления с неотрицательными цифрами по основанию q. Предположим, что делимое

$$A = A_n A_{n-1} \ldots A_0$$

и делитель

$$D = D_k D_{k-1} \ldots D_0, \ \ D \ne 0$$

целые числа. Результат обозначим как

$$B = B_n B_{n-1} \ldots B_0.$$

Алгоритм деления «в столбик» состоит в последовательном выполнении, начиная со старшей цифры A_n числа A следующих действий:

$$\begin{aligned} &G_n = A_n, \\ &R_i = G_i - B_i\cdot D, \qquad i = n, \ldots, 0. \\ &G_{i-1} = R_i\cdot q + A_{i-1}, \end{aligned} \tag{5}$$

где B_i – цифра в системе счисления по основанию q, такая, что $B_i\cdot D$ не превышает G_i, R_i – остаток от деления G_i на D, т.е.

$$0 \le R_i < D.$$

Действия производятся до тех пор, пока не будут исчерпаны все разряды A. Последнее значение R_0 – остаток от деления A на D.

Алгоритм деления в симметричных системах счисления существенно отличается от алгоритма (5) тем, что в некоторых случаях очередная цифра частного и остаток определяются не только соотношением очередного делимого и делителя, но и знаком необработанной части делимого. Это отличие возникает в связи с рядом особенностей симметричных систем счисления, в частности, с тем, что необработанная часть делимого в симметричных системах счисления может быть положительной или отрицательной, и тем самым увеличивать или уменьшать обработанную его часть, в то время как в системах счисления с неотрицательными числами необработанная и обработанная части числа могут только суммироваться.

Рассмотрим еще одну особенность симметричных систем счисления, являющуюся причиной указанного различия алгоритмов деления. Пусть имеется симметричная система счисления по основанию $p = 2s + 1$, где s – целое положительное число. Покажем, что в этой системе счисления максимальное вещественное число, целая часть которого имеет разрядность n, равно $\frac{p^n}{2}$.

Цифры в данной системе счисления – целые числа, принадлежащие $[\bar{s}, s]$, s – наибольшая цифра в рассматриваемой системе счисления, $\bar{s} = -s$. Тогда максимальное вещественное число, с n-разрядной целой частью имеет вид $s\ldots s.s(s)$, где количество цифр s в целой части числа равно n. Покажем, что в симметричных системах счисления

$$s\ldots s.s(s) = \frac{p^n}{2} \qquad (6)$$

Число в левой части (6) – бесконечно убывающая геометрическая прогрессия с знаменателем p^{-1} и первым членом, равным $s{\cdot}p^{n-1}$. Тогда, получим значение числа $s\ldots s.s(s)$:

$$s \cdot \frac{p^{n-1}}{1-p^{-1}} = \frac{(p-1)}{2} \cdot \frac{p \cdot p^{n-1}}{(p-1)} = \frac{p^n}{2}.$$

Аналогично можно показать, что минимальное вещественное число, с n-разрядной целой частью

$$\bar{s}\ldots\bar{s}.\bar{s}(\bar{s}) = -\frac{p^n}{2}$$

Вследствие этого число $p^n = 10...0$, где количество нулей равно n представимо в виде

$$\frac{p^n}{2} = p^n - \frac{p^n}{2} = 10...0 + \bar{s}\ldots\bar{s}.\bar{s}(\bar{s}) = 1\bar{s}\ldots\bar{s}.\bar{s}(\bar{s}) \qquad (6')$$

где количество цифр $\bar{s}$ в целой части числа составляет n. Таким образом, числа вида $\frac{p^n}{2}$ в симметричных системах счисления представимы двумя способами (6) и (6′).

Следствием (6) является то, что знак числа в симметричных системах счисления определяется знаком его старшего разряда. Из (6) также следует непредставимость в этих системах чисел вида $\frac{p^n}{2}$ конечными дробями. Округление бесконечных дробей из (6) и (6′) дают два представления числа $\frac{p^n}{2}$ в виде конечных дробей – с недостатком: $P' = s\ldots s\,.\,s\ldots s$, т. к. $P' < \frac{p^n}{2}$ и с избытком: $P'' = 1\,\bar{s}\ldots\bar{s}.\,\bar{s}\ldots\bar{s}$, поскольку $P'' > \frac{p^n}{2}$.

Рассмотрим алгоритм деления «в столбик» в симметричной системе счисления по основанию $p = 2s+1$. Обозначим делимое

$$a = a_n a_{n-1}\ldots a_0,$$

делитель

$$d = d_k d_{k-1}\ldots d_0,\ d \neq 0,$$

неполное частное от деления a на d

$$b = b_n b_{n-1}\ldots b_0.$$

В симметричных системах счисления на каждом шаге применяется операция деления по модулю в соответствии с (2) и (3). Поэтому алгоритм аналогичен (5):

$$\begin{aligned} &g_n = a_n,\\ &r_i = g_i - b_i\cdot d, \qquad i = n, \ldots, 0,\\ &g_{i-1} = r_i\cdot p + a_{i-1}, \end{aligned} \tag{7}$$

но здесь r_i – остаток от деления g_i на d в симметричной системе счисления, т.е. b_i выбирается таким образом, чтобы остаток от деления удовлетворял условию

$$-\frac{|d|}{2} \le r_i \le \frac{|d|}{2} \tag{8}$$

Действия (7) производятся до тех пор, пока не будут исчерпаны все разряды числа a. Последнее значение r_0 – остаток от деления a на d в симметричной системе счисления.

В последующих рассуждениях будем считать, что алгоритм деления применяется к произвольному целому a и модулю делителя $|d|$, а затем неполное частное и остаток умножаются на знак d. Таким образом, будем считать, что $d > 0$.

Определим, как зависит очередное частное b_i от соотношения очередного делимого g_i и делителя d. Рассмотрим отдельно случаи, когда делитель нечетный, и когда он четный.

<u>Утверждение 1.</u> При нечетном положительном делителе d неполное частное b_i от деления по модулю очередного делимого g_i на d с остатком r_i таково, что выполняется следующее условие:

$$b_i \cdot d - \frac{d-1}{2} \le g_i \le b_i \cdot d + \frac{d-1}{2}. \qquad (9)$$

Доказательство. В силу того, $d > 0$ и нечетно, неравенство (8) принимает вид

$$-\frac{d-1}{2} \le r_i \le \frac{d-1}{2}.$$

Поскольку $g_i = b_i \cdot d + r_i$, то справедливо (9).Утверждение доказано.

Рассмотрим теперь случай, когда делитель d – четное положительное число.

<u>Утверждение 1′.</u> При четном положительном делителе d существуют числа b_i' и r_i' такие, что для очередного делимого g_i выполняются условия:

$$b_i' \cdot d - \frac{d}{2} \le g_i \le b_i' \cdot d + \frac{d}{2}, \qquad (10)$$

$$g_i = b_i' \cdot d + r_i'. \qquad (11)$$

Возможны два варианта определения очередной цифры неполного частного b_i и остатка r_i от деления по модулю очередного делимого g_i на d:

$$b_i = b_i', \;\; r_i = r_i' \qquad (12)$$

и

$$b_i = b_i' + sign\,(r_i'), \; r_i = -\,r_i'. \qquad (12')$$

1) Если $|r_i'| = \frac{d}{2}$, необработанная часть делимого $(a_{i-1}\ldots a_0) \ne 0$ и $sign\,(a_{i-1}\ldots a_0) = sign\,(r_i')$, то b_i и r_i определяются в соответствии с (12′).

2) Если $|r_i'| = \frac{d}{2}$, необработанная часть делимого $(a_{i-1}\ldots a_0) = 0$, $|b_i'| \ne s$, то b_i и r_i определяются в соответствии с (12) или с (12′), оба определения допустимы.

3) В остальных случаях b_i и r_i определяются соответствии с (12).

Доказательство. Справедливость (10) доказывается аналогично утверждению 1. При четном d остаток удовлетворяет (8):

$$-\frac{d}{2} \le r_i' \le \frac{d}{2},$$

откуда следует:

$$b_i' \cdot d - \frac{d}{2} \le g_i' \le b_i' \cdot d + \frac{d}{2}.$$

Покажем, что справедлив п. 1. Неполное частное b после i-ого шага алгоритма (7) представляется в виде:

$$b = (b_n \ldots b_{i+1}) \cdot p^{i+1} + b_i' \cdot p^i + \frac{r_i' \cdot p^i + (a_{i-1} \ldots a_0)}{d}, \qquad (14)$$

где $(b_n \ldots b_{i+1})$ – построенная часть неполного частного.

Так как $|r_i'| = \dfrac{d}{2}$, абсолютное значение $\dfrac{r_i'\cdot p^i + (a_{i-1}...a_0)}{d}$ можно преобразовать:

$$\left|\frac{r_i'\cdot p^i + (a_{i-1}...a_0)}{d}\right| = \left|sign(r_i')\frac{p^i}{2} + \frac{(a_{i-1}...a_0)}{d}\right|. \qquad (15)$$

При $sign\ (a_{i-1}\ldots a_0) = sign\ (r_i')$

$$\left|sign(r_i')\frac{p^i}{2} + \frac{(a_{i-1}...a_0)}{d}\right| > \frac{p^i}{2}.$$

Заменяя в (14) выражение

$$\frac{r_i'\cdot p^i + (a_{i-1}...a_0)}{d} = sign(r_i')\frac{p^i}{2} + \frac{(a_{i-1}...a_0)}{d}$$

на равноценное

$$sign\ (r_i')\cdot p^i - sign\ (r_i')\cdot\frac{p^i}{2} + \frac{(a_{i-1}...a_0)}{d}.$$

Получаем

$$b = (b_n\ldots b_{i+1})\cdot p^{i+1} + (b_i' + sign\ (r_i'))\cdot p^i - sign\ (r_i')\cdot\frac{p^i}{2} + \frac{(a_{i-1}...a_0)}{d}, \qquad (16)$$

т.е. $b_i = b_i' + sign\ (r_i')$, $r_i = -\ r_i'$. Число $b_i' + sign\ (r_i')$ является очередной цифрой неполного частного b, если выполняется

$$b_i' + sign\ (r_i') \le s$$

или

$$|b_i'| < s. \qquad (17)$$

Покажем, что (17) выполняется только если на предыдущем шаге остаток r_{i+1} от деления g_{i+1} на d удовлетворяет условию $|r_{i+1}| = \dfrac{d}{2}$.

Предположим, что $|b_i'| = |s|$. Рассмотрим случай $b_i' = s$. Тогда из (7) получим

$$g_i = s\cdot d + \frac{d}{2},$$

и, учитывая то, что $s = \dfrac{(p-1)}{2}$,

$$g_i = \frac{(p-1)\cdot d}{2} + \frac{d}{2} = \frac{p\cdot d}{2}.$$

Поскольку в силу (7)

$$g_i = r_{i+1}\cdot p + a_i,$$

то должно выполняться:

$$\left(\frac{d}{2} - r_{i+1}\right)\cdot p = a_i. \qquad (18)$$

Так как числа r_{i+1} и $\frac{d}{2}$ – целые, т.е. $\left(\frac{d}{2} - r_{i+1}\right)\cdot p$ кратно p, $|a_i| \le s$ – цифра в симметричной системе счисления по основанию p, то (18) выполняется только при $a_i = 0$ и $r_{i+1} = \frac{d}{2}$.

Аналогично доказывается, что $b_i' = \bar{s}$ только при $r_{i+1} = -\frac{d}{2}$.

Таким образом, если $|b_i'| = s$, то на предыдущем шаге $|r_{i+1}| = \frac{d}{2}$, и если $sign\,(a_i\ldots a_0) = sign\,(r_{i+1})$, то b_{i+1} и r_{i+1} определялись на $(i+1)$-ом шаге в соответствии с (12′). При этом если $|b_{i+1}| < s$, то на i-ом шаге

$$g_i = -\,s\cdot d - \frac{d}{2},$$

и $sign\,(a_{i-1}\ldots a_0) \ne sign\,(r_i)$. Если $|b_{i+1}| = s$, то $|r_{i+2}| = \frac{d}{2}$, и т.д. пока на одном из предыдущих шагов j не окажется $|b_j| < s$, тогда на этом шаге b_j и r_j определялись в соответствии с (12′).

Если на всех предыдущих шагах $|b_j| = s$, то при определении старшей цифры b_k неполного частного b на шаге $k \le n$

$$|g_k| = \frac{d}{2}.$$

Тогда значение $|b_k|$ равно 0 или 1, и если $sign\,(a_k\ldots a_0) = sign\,(r_k)$, то b_k и r_k определялись на k-ом шаге в соответствии с (12′), после чего $sign\,(a_k\ldots a_0) \ne sign\,(r_k)$

2) Покажем, что справедлив п. 2. При $(a_{i-1}\ldots a_0) = 0$ представление (14) неполного частного b при $|r_i'| = \frac{d}{2}$ принимает вид

$$b = (b_n\ldots b_{i+1})\cdot p^{i+1} + b_i'\cdot p^i + sign\,(r_i')\,\frac{p^i}{2},$$

представление (16) принимает вид:

$$b = (b_n\ldots b_{i+1})\cdot p^{i+1} + (b_i' + sign\,(r_i'))\cdot p^i - sign\,(r_i')\cdot \frac{p^i}{2}.$$

Так как $|b_i'| \ne s$, оба представления допустимы. Утверждение доказано.

В симметричных системах счисления для определения очередного b_i, на каждом шаге алгоритма производится сравнение $|r_i|$ и $\frac{|d|}{2}$, требующее вычисления $\frac{|d|}{2}$, которое при нечетном d не является целым числом. В [1] предложен способ избежать вычисления $\frac{|d|}{2}$ в алгоритме деления чисел с плавающей запятой в троичной симметричной системе счисления, который применим к алгоритму деления по модулю в любых симметричных системах счисления. Указанный способ заключается предварительном вычислении $2a$ и $2|d|$, сравнении на каждом шаге $2g_i$ и $|d|$, с последующем вычитанием из $2g_i$ удвоенного делителя, т.е. $2|d|$.

Окончательно, алгоритм деления a на d с остатком в симметричной системе счисления по основанию p описывается формулами, аналогичными (7), но применяемыми к делимому $A = (A_n A_{n-1} \ldots A_0) = 2a$ и делителю $2|d|$:

$$\begin{aligned} & g_n = A_n, \\ & r_i = g_i - 2b_i \cdot |d|, \qquad i = n, \ldots, 0, \\ & g_{i-1} = r_i \cdot p + A_{i-1}, \end{aligned} \tag{19}$$

b_i выбирается таким образом, чтобы остаток от деления удовлетворял условию

$$-|d| \le 2r_i \le |d|.$$

Остаток от деления a и d равен половине остатка от деления A на $2|d|$, поэтому после выполнения алгоритма (19) остаток должен быть скорректирован. Также после вычисления неполного частного от деления a на d и остатка, корректируются их знаки: если a и d – числа разного знака, то знаки неполного частного и остатка меняются на противоположные.

Вычисление b_i на каждом шаге осуществляется следующим образом:

если $2|g_i| < |d|$, то b_i выбирается так, чтобы , $-|d| \le 2r_i \le |d|$;

если $2|g_i| = |d|$, т.е. d – четный, то

- если $(A_{n-i} \ldots A_0) > 0$, то b_i выбирается так, чтобы , $2r_i = -|d|$;
- если $(A_{n-i} \ldots A_0) < 0$, то b_i выбирается так, чтобы , $2r_i = |d|$;
- если $(A_{n-i} \ldots A_0) = 0$, то b_i выбирается так, чтобы знак r_i совпадал со знаком g_i.

При выборе очередной цифры b_i можно воспользоваться методом, приведенном в [3] в разделе 4.3.1.

Рассмотрим пример деления в симметричной системе счисления по основанию 9 числа

$$a = 554_{10} = 1\ \bar{2}\ \bar{1}\ \bar{4}_{\bar{9}}$$

на

$$d = 3_{10} = 3_{\bar{9}}.$$

Удвоим делимое и делитель:

$$A = 1108_{10} = 2\ \bar{4}\ \bar{3}\ 1_{\bar{9}}$$

$$2d = 6_{10} = 1\ \bar{3}_{\bar{9}}.$$

$$
\begin{array}{r|l}
-\;2\;\bar{4}\;\bar{3}\;1 & \underline{1\;\bar{3}} \\
\underline{0}\phantom{\;\bar{4}\;\bar{3}\;1} & 0\;2\;3\;\bar{4} \\
-\;2\;\bar{4}\phantom{\;\bar{3}\;1} & \\
\underline{1\;3}\phantom{\;\bar{3}\;1} & \\
-\;2\;\bar{3} & \\
\underline{2\;0} & \\
\bar{3}\;1 & \\
\underline{\bar{3}\;3} & \\
\bar{2} &
\end{array}
$$

Поскольку $A_3 = 2$ и $A_3 < d = 3$, то $b_3 = 0$. Следующее делимое, определяется формулой

$$g_{i\text{-}1} = r_i{\cdot}p + A_{i\text{-}1},$$

из алгоритма (19), которой соответствует приписывание к полученному на предыдущем шаге остатку следующей цифры делимого. Получим $g_2 = 2\ \bar{4}$, из чего определим

$$|b_2| = 2,\ |r_2| = 2.$$

Знак b_2 совпадает со знаком g_2 и d.

На следующем шаге $g_1 = 2\ \bar{3}$. Определим, что b_1 может быть равным 2, и остаток равен 3 или b_1 может быть равным 3, и остаток равен $\bar{3}$. С учетом знака нерассмотренной части числа, равной 1, получим,

$$|b_1| = 3,\ |r_1| = \bar{3},$$

На следующем шаге $g_0 = \bar{3}\ 1$, и $b_0 = \bar{4}$, остаток от деления A на $2|d|$ равен $\bar{2}$, т.е. остаток от деления a на d равен $\bar{1}$. Таким образом, неполное частное от деления $1\ \bar{2}\ \bar{1}\ \bar{4}\ _9$ на 3 равно $2\ 3\ \bar{4}\ _9$, остаток $\bar{1}$.

Литература

Брусенцов Н.П. Алгоритмы деления для троичного кода с цифрами 0, 1, -1. // Вычислительная техника и вопросы кибернетики. Вып. 10. – Л.: Изд–во ЛГУ, 1974. С. 39-44.

Рамиль Альварес Х. Троичное деление целых чисел //Рамиль Альварес Х. Алгоритмы троичной арифметики. – М., Фонд «Новое тысячелетие», 2012. С. 7-13.

Кнут Д. Искусство программирования, т. 2. Получисленные алгоритмы. 3-е изд. — М.: «Вильямс», 2007.

Опубликовано: Программные системы и инструменты № 14. Под ред. Л.Н. Королева. – М.: Издательский отдел ВМиК МГУ, 2013. С. 199 208.