РЕШЕНИЕ СИСТЕМ ЛИНЕЙНЫХ ДИОФАНТОВЫХ УРАВНЕНИЙ [*]

© Г.И. Малашонок

Malaschonok G.I. Solving systems of linear Diophantine equations. New methods to solve linear systems of Diophantine equations are suggested. These problems are considered from two points of view – modular and p-adic. Both methods allow to get solutions of a linear system with the size $n \times m$ with the complexity $O(n^{\beta+1}m)$. For quasi-square system, the p-adic method allows to get solutions with the complexity $O(n^3)$. Both estimations have the accuracy up to the logarithmic multipliers, β – is a power in the estimation of the time of matrix multiplication.

[*]Вестник ТГУ, Т.5, N. 5, 2000, 620–627

ВВЕДЕНИЕ

Задачи, в которых ищутся целочисленные решения алгебраических уравнений, относятся к одним из самых старых задач в математике, которые неизменно привлекали к себе внимание.

Основной метод для решения линейных систем диофантовых уравнений основан на вычислении формы Смита матрицы коэффициентов системы [1, 2]. В недавних работах были получены *вероятностные* методы для нахождения решений линейной системы диофантовых уравнений [3, 4]. Вероятностные методы являются самыми быстрыми из известных сегодня методов. Однако они не гарантируют нахождение решения. Если решение не найдено вероятностным методом, то всегда существует вероятность того, что система совместна, но ее решение не найдено случайно.

В настоящей статье предлагаются детерминистские p-адический и модулярный методы, которые не связаны с вычислением формы Смита матрицы коэффициентов. Оба метода сводят задачу к решению системы в кольце вычетов.

Первый из них, с пощью p-адического подьема, вычисляет базисное множество точек на гиперплоскости всех решений системы. А затем вычисляет все целые точки на этой плоскости.

Второй метод основан на преобразовании к эквивалентной системе с левым квадратным блоком, имеющим диагональный вид. Его диагональные элементы равны определителю соответствующего блока исходной матрицы. Вычисления выполняются с помощью китайской теоремы об остатках. А затем решается система по модулю этого определителя. Кроме того удается объединить детерминистские и вероятностные методы. Когда решение не получено вероятностным методом, то процесс решения продолжается до установления несовместности системы или построения всех решений системы детерминистским p-адическим методом.

Показано, что для квазиквадратных систем сложность p-адического метода $O(n^3)$, а модулярного $O(n^{\beta+1})$. Для существенно прямоугольных систем, когда число неизвестных m более чем вдвое превышает число уравнений n, оба метода имеют одинаковый порядок сложности – $O(n^{\beta+1}m)$. Здесь оценки приводятся с точностью до логарифмического множителя и предполагается применение стандартного умножения. β – показатель степени в оценке сложности матричного умножения.

Приводится алгоритм со сложностью $O(n^{\beta-1}m)$ для решения систем над кольцами вычетов.

Как обычно, все результаты переносятся с кольца целых чисел на кольцо полиномов над полем.

КВАЗИЕВКЛИДОВЫ КОЛЬЦА

Пусть $\mathbf{R}$ – коммутативное кольцо с единицей. Будем называть *диофантовой матрицей* для ненулевой пары α из $\mathbf{R}^2$ унимодулярную матрицу $E_\alpha \in \mathbf{R}^{2\times 2}$ такую, что

$$E_\alpha \alpha = \begin{pmatrix} g \\ 0 \end{pmatrix},\ g \neq 0,\ \det E_\alpha = 1.$$

Для нулевой пары *диофантовой матрицей* будем называть единичную.

Определение 1 *Кольцо* $\mathbf{R}$ *назовем квазиевклидовым, если в кольце определен алгоритм, вычисляющий для любой пары диофантову матрицу.*

Из обратимости диофантовой матрицы следует, что идеал (a, b) является главным идеалом, порожденным g. Поэтому всякое квазиевклидово кольцо является кольцом главных идеалов.

Наоборот. Пусть $\mathbf{R}$ кольцо главных идеалов, в которой существует алгоритм GI для вычисления генератора g идеала (p, q), порожденного двумя элементами $p, q \in \mathbf{R}$:

$$\{g, a, b\} = GI(p, q),\ (g) = (p, q),\ ap + bq = g$$

Такое кольцо будем называть *GI кольцом.* Диофантовой матрицей для ненулевой пары $\alpha = \begin{pmatrix} p \\ q \end{pmatrix}$ будет $E_\alpha = \begin{pmatrix} a & b \\ -q/g & p/g \end{pmatrix}$.

Заметим, что числа a, b и g определены неоднозначно. Поэтому и диофантова матрица определена неоднозначно. Если в $\mathbf{R}$ ввести порядок, то этот выбор можно сделать однозначным.

Если $\mathbf{R}$ евклидово кольцо, то в качестве алгоритма GI можно взять расширенный алгоритм Евклида.

Предложение 1 *Псевдоевклидовость кольца с отношением порядка наследуется его факторкольцом.*

Доказательство.

Пусть $\mathbf{R}'$ – факторкольцо псевдоевклидова кольца $\mathbf{R}$ с некоторым отношением порядка и $\phi : \mathbf{R} \to \mathbf{R}'$ канонический гомоморфизм. Определим функцию $\psi : \mathbf{R}' \to \mathbf{R}$, выбирающую представителя класса смежности для $p' \in \mathbf{R}'$ в прообразе $\phi^{-1}(p')$:

$$\psi(p') = min(\phi^{-1}(p')).$$

Пусть $\psi(p') = p$, $\psi(q') = q$, $\alpha = \begin{pmatrix} p \\ q \end{pmatrix}$, E_α – диофантова матрица для α и $E_\alpha \alpha = \begin{pmatrix} g \\ 0 \end{pmatrix}$. Определим для ненулевой пары $\alpha' = \begin{pmatrix} p' \\ q' \end{pmatrix} \in \mathbf{R}'^2$ диофантову матрицу $E_{\alpha'}$ следующим образом $E_{\alpha'} = \phi(E_\alpha)$. Очевидно, такое определение корректно, так как

$$E_{\alpha'}\alpha' = \begin{pmatrix} g' \\ 0 \end{pmatrix},\ \det E_{\alpha'} = 1,\ g' \neq 0.$$

Покажем, что $g' = \phi(g) \neq 0$.

Из-за обратимости диофантовых матриц p и q делятся на g. Поэтому если $\phi(g) = 0$, то $\phi(p) = \phi(q) = 0$, но по условию хоть одно из чисел $p' = \phi(p), q' = \phi(q)$ не должно быть нулем.

Следовательно, множество квазиевклидовых колец содержит все GI кольца и их факторкольца, в том числе евклидовы кольца и их факторкольца.

Будем называть матрицу $A = (a_{ij})$, размера $n \times m$, *верхней треугольной*, если все элементы под главной диагональю нулевые: $(i, j) = 0$ для всех $n \geq j > i \geq 1$. Заметим, что элементы на главной диагонали и над ней могут тоже быть нулевыми.

Будем говорить, что матрица $\mathcal{A}$ *разложима*, если она является произведением унимодулярной матрицы $\tilde{\mathcal{A}}$ на верхнюю треугольную матрицу $\bar{\mathcal{A}}$: $\mathcal{A} = \tilde{\mathcal{A}}\bar{\mathcal{A}}$.

Предложение 2 *В квазиевклидовом кольце все матрицы разложимы.*

Доказательство.

Построим такое разложение для матрицы $A = (a_{ij})$, размера $n \times m$. Пусть $\alpha_{ij} = \begin{pmatrix} a_{jj} \\ a_{ij} \end{pmatrix}$, $i > j$, $E_{\alpha_{i,j}}$ – диофантова матрица для α_{ij}. Обозначим через $E^n_{\alpha_{i,j}}$ унимодулярную матрицу, которая получается после замены в единичной матрице порядка n четырех элементов (jj), (ji), (ij), (ii) соответственно на четыре элемента (11), (12), (21), (22) матрицы $E_{\alpha_{i,j}}$. Пусть $E_j = E_{\alpha_{n,j}} \cdots E_{\alpha_{j+1,j}}$, $U = E_{n-1} \cdots E_1$. Тогда U – унимодулярная матрица, UA – верхняя треугольная матрица. Эти вычисления приводят к последовательному исключению ненулевых элементов сначала в первом столбце, затем во втором столбце, и так далее до $n-1$-го столбца.

Обозначим C_D – асимптотическую оценку сложности вычисления диофантовой матрицы, $M(n)$ – асимптотическую оценку сложности умножения матриц порядка n, $M(n) = \alpha n^\beta$. Известное сегодня значение для β меньше 2,356 [5].

Лемма 1 *Пусть $\mathbf{R}$ – квазиевклидово кольцо, $A, C \in \mathbf{R}^{n\times n}$, A – верхняя треугольная матрица. Тогда матрицу $\begin{pmatrix} A \\ C \end{pmatrix}$ можно разложить $\begin{pmatrix} A \\ C \end{pmatrix} = Q\begin{pmatrix} \bar{A} \\ 0 \end{pmatrix}$ с асимптотической оценкой сложности, не превосходящей*

$$C_s(n) = n^2 C_D + \alpha\lambda_\beta n^\beta, \qquad (1)$$

где $\lambda_\beta = 7(2^{\beta-2} - 1)^{-1}$.

Доказательство. Для разложения матрицы размера 2×1 достаточно умножить на ее диофантовую матрицу. Пусть теперь $n = 2^p$, $p > 0$.

Разобьем A и C на квадратные блоки

$$A = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix},\quad C = \begin{pmatrix} e & f \\ g & h \end{pmatrix}.$$

Пусть имеются четыре разложения:

$$\begin{pmatrix} a \\ e \end{pmatrix} = Q_1^{-1} \begin{pmatrix} a_1 \\ 0 \end{pmatrix}, \ \begin{pmatrix} b_1 \\ f_1 \end{pmatrix} = Q_1 \begin{pmatrix} b \\ f \end{pmatrix},$$
$$\begin{pmatrix} a_1 \\ g \end{pmatrix} = Q_2^{-1} \begin{pmatrix} a_2 \\ 0 \end{pmatrix}, \ \begin{pmatrix} b_2 \\ h_1 \end{pmatrix} = Q_2 \begin{pmatrix} b_1 \\ h \end{pmatrix},$$
$$\begin{pmatrix} d \\ f_1 \end{pmatrix} = Q_3^{-1} \begin{pmatrix} d_1 \\ 0 \end{pmatrix}, \begin{pmatrix} d_1 \\ h_1 \end{pmatrix} = Q_4^{-1} \begin{pmatrix} d_2 \\ 0 \end{pmatrix}.$$

Матрицы $Q_i = \begin{pmatrix} \alpha_i & \beta_i \\ \gamma_i & \delta_i \end{pmatrix}$, $i = 1,..,4$, унимодулярные, a, a_1, a_2, d, d_1, d_2 – верхние треугольные. Тогда получим $\bar{A} = \begin{pmatrix} a_2 & b_2 \\ 0 & d_2 \end{pmatrix}$ и

$$Q = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \alpha_4 & 0 & \beta_4 \\ 0 & 0 & 1 & 0 \\ 0 & \gamma_4 & 0 & \delta_4 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \alpha_3 & \beta_2 & 0 \\ 0 & \gamma_2 & \delta_2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \times$$
$$\times \begin{pmatrix} \alpha_2 & 0 & 0 & \beta_2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \gamma_3 & 0 & 0 & \delta_3 \end{pmatrix} \begin{pmatrix} \alpha_1 & 0 & \beta_1 & 0 \\ 0 & 1 & 0 & 0 \\ \gamma_1 & 0 & \delta_1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} =$$
$$\begin{pmatrix} \alpha_1\alpha_2 & 0 & \beta_1\alpha_2 & \beta_2 \\ \eta_1 & \alpha_4\alpha_3 & \eta_2 & \beta_4\delta_2 \\ \delta_3\gamma_1 & \gamma_3 & \delta_3\delta_1 & 0 \\ \eta_3 & \gamma_4\alpha_3 & \eta_4 & \delta_4\delta_2 \end{pmatrix}.$$

Используются обозначения $\eta_1 = \alpha_4\beta_3\gamma_1 + \beta_4\gamma_2\alpha_1$, $\eta_2 = \alpha_4\beta_3\delta_1 + \beta_4\gamma_2\beta_1$, $\eta_3 = \gamma_4\beta_3\gamma_1 + \delta_4\gamma_2\alpha_1$, $\eta_4 = \gamma_4\beta_3\delta_1 + \delta_4\gamma_2\beta_1$.

Следовательно, достаточно выполнить 28 умножений блоков порядка $n/2$ и четыре разложения для блоков размера $n \times n/2$.

Продолжая выполнять разбиения на все более мелкие блоки, вплоть до блоков 2×1, получим асимптотическую оценку сложности вычислений

$$C_s(n) = n^2C_D + 7\sum_{i=1}^{p} 4^i M(\frac{n}{2^i}) < n^2C_D + \alpha\lambda_\beta n^\beta.$$

Здесь число $\lambda_\beta = 7(2^{\beta-2}-1)^{-1}$ находится в пределах от $\lambda_3 = 7$ до $\lambda_{\log 6} = 14$.

Следствие 1 *Пусть* $\mathbf{R}$ *– квазиевклидово кольцо. Тогда матрицу* $\begin{pmatrix} A \\ C \end{pmatrix}$ *над кольцом* $\mathbf{R}$*, размера* $n \times m$*, с верхним треугольным блоком* $A \in \mathbf{R}^{m\times m}$ *можно разложить с асимптотической оценкой сложности, не превосходящей* $C_s^1(n,m) =$
$= (n-m)mC_D + \alpha m^{\beta-1}(\lambda_\beta(n-m) + 4n - 8m)$, *при* $n \geq 2m$ *и* $C_s^2(n,m) = (n-m)mC_D +$
$+\alpha(n-m)^{\beta-2}(\lambda_\beta m(n-m) + 10mn - 4n^2 - 4m^2)$, *при* $m \leq n \leq 2m$.

Доказательство.

a. Пусть $n \geq 2m$ и m делит n. Сначала выполним разложение для матрицы, образованной блоком A и верхним квадратным блоком C_1 матрицы C: $\begin{pmatrix} A \\ C_1 \end{pmatrix} = Q_1 \begin{pmatrix} \bar{A} \\ 0 \end{pmatrix}$. Затем $(n-2m)/m$ раз проведем разложение матрицы, образованной блоком $\bar{A}$ на предыдущем шаге и i-тым сверху квадратным блоком C_i матрицы C, $i = 2, \dots k$. Перемножение унимодулярных матриц, получаемых на каждом шаге, требует еще четырех умножений матриц размера $m \times m$. Всего получим

$$C_s^1(n,m) = C_s(m)\frac{(n-m)}{m} + 4M(m)\frac{(n-2m)}{m}.$$

b. Пусть $m \leq n \leq 2m$ и $n - m$ делит m. Разобьем матрицы A и C на блоки размера $(n-m) \times (n-m)$. Матрица C содержит одну блочную строку. Обозначим $C = (C_1, \bar{C}_1)$, где C_1 – левый блок матрицы C. Матрица A содержит $k = m/(n-m)$ блочных строк и столбцов.

Сначала разложим матрицу, образованную верхним левым диагональным блоком A_1 матрицы A и левым блоком C_1 матрицы C: $\begin{pmatrix} A_1 \\ C_1 \end{pmatrix} = Q_1 \begin{pmatrix} \bar{A}_1 \\ 0 \end{pmatrix}$, Q_1 – унимодулярная матрица. И вычислим произведение $\begin{pmatrix} \bar{A'}_1 \\ \bar{C'}_2 \end{pmatrix} = Q_1 \begin{pmatrix} \bar{A}_1 \\ \bar{C}_1 \end{pmatrix}$, где $\bar{A}_1$ - первая блочная строка матрицы A начиная со второго блока. Обозначим $\bar{C'}_1 = (C_1, \bar{C}_1)$, где C_1 – левый квадратный блок. Всего необходимо $4(k-1)$ блочных умножений.

Затем последовательно раскладываем матрицы, образованные i-тым (i=2,..., k) диагональным блоком A_i матрицы A и левым блоком C_i матрицы $\bar{C'}_i$: $\begin{pmatrix} A_i \\ C_i \end{pmatrix} = Q_i \begin{pmatrix} \bar{A}_i \\ 0 \end{pmatrix}$. И вычисляем произведение $\begin{pmatrix} \bar{A'}_i \\ \bar{C}_{i+1} \end{pmatrix} = Q_1 \begin{pmatrix} \bar{A}_i \\ \bar{C}_i \end{pmatrix}$, где $\bar{A}_i$ - i-я блочная строка матрицы A, начиная с $i+1$-го блока, $\bar{C'}_i = (C_i, \bar{C}_i)$.

Кроме того, для вычисления унимодулярной матрицы Q требуется перемножить унимодулярные матрицы Q'_i, размера $n \times n$, соответствующие Q_i. Матрица Q'_i получается из единичной матрицы I_n заменой четырех блоков (i,i), (i,k+1), (k+1,i), (k+1,k+1) соответственно на четыре блока (1,1), (1,2), (2,1), (2,2) матрицы Q_i. Умножение на Q'_i требует четырех блочных умножений, а всего на i-том шаге нужно $4(k-i)+4$ блочных умножений, $i = 2, \dots, k$.

Суммируя по всем шагам, получим $C_s^2(n,m)$: $kC_s(n-m) + \left(\sum_{i=1}^{k} 4(k-i+1) - 4\right) M(n-m) =$
$= kC_s(n-m) + 2(k^2+k-2)M(n-m) =$
$= m(n-m)C_D + M(n-m)(\lambda_\beta m/(n-m) +$
$+2(5mn - 2n^2 - 2m^2)/(n-m)^2)$.

Замечание. *Не трудно показать, что в условии леммы можно заменить верхнюю* *треугольную матрицу A на нижнюю треугольную матрицу. А так же можно раскладывать в произведение, когда унимодулярная матрица умножается справа. Оценки при этом сохраняются.*

Теорема 1 *В квазиевклидовом кольце всякую квадратную матрицу $\mathcal{A}$ порядка n можно разложить с асимптотической оценкой сложности $C_t(n)$, не превосходящей*

$$\mu_\beta M(n) + \frac{n^2-n}{2} C_D, \quad \mu_\beta = \frac{9\cdot 2^{\beta-2}-2}{(2^\beta-2)(2^{\beta-2}-1)}.$$

Доказательство. Для разложения матрицы второго порядка достаточно умножить ее на диофантову матрицу ее первого столбца.

Не уменьшая общности, будем полагать, что $n = 2^p$. Разобьем матрицу $\mathcal{A}$ на четыре квадратных блока $\mathcal{A} = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$.

(1). Пусть имеется разложение блока A :

$$A = \tilde{A}^{-1}\bar{A}.$$

Обозначим $A_1 = \tilde{A}A$, $B_1 = \tilde{A}B$.

(2). Пусть разложение блока $\begin{pmatrix} A_1 \\ C \end{pmatrix}$ имеет вид

$$\begin{pmatrix} A_1 \\ C \end{pmatrix} = Q^{-1}\begin{pmatrix} A_2 \\ 0 \end{pmatrix},$$

Q – унимодулярная, A_2 – верхняя треугольная. Обозначим $\begin{pmatrix} B_2 \\ D_1 \end{pmatrix} = Q\begin{pmatrix} B_1 \\ D \end{pmatrix}$.

(3). Пусть имеется разложение блока D_1 :

$$D_1 = \tilde{D}_1^{-1}\bar{D}_1.$$

Обозначим $D_2 = \bar{D}_1$.

Тогда разложима матрица A:

$$\mathcal{A} = \tilde{\mathcal{A}}^{-1}\bar{\mathcal{A}},$$

$\tilde{\mathcal{A}} = \mathrm{diag}(1, \tilde{\mathrm{D}}_1)\ \mathrm{Q}\ \mathrm{diag}(\tilde{\mathrm{A}}, 1)$, $\bar{\mathcal{A}} = \begin{pmatrix} A_2 & B_2 \\ 0 & D_2 \end{pmatrix}$.

Всего, при таком разложении, требуется 9 блочных умножений, два блочных разложения и одно разложение для блока размера $n \times n/2$. Обозначим $C_t(n)$ сложность разложения матрицы $\mathcal{A}$. Получим

$$C_t(n) = 2C_t(n/2) + 9M(n/2) + C_s(n/2). \quad (2)$$

По лемме 1 получим

$$C_t(n) = 2C_t(n/2) + (n/2)^2 C_D + \alpha(9+\lambda_\beta)(n/2)^\beta.$$

Будем разбивать блоки A и D далее на квадратные блоки. И так будем поступать вплоть до блоков второго порядка. При этом асимптотическая оценка сложности вычислений будет $C_t(n) = \sum_{i=1}^{p} 2^{i-1}\left(\left(\frac{n}{2^i}\right)^2 C_D + \alpha(9+\lambda_\beta)\left(n/2^i\right)^\beta\right)$ $< \frac{n^2-n}{2}C_D + \alpha\mu_\beta n^\beta$,
где число $\mu_\beta = \frac{1}{2^\beta-2}\left(9 + \frac{7}{2^{\beta-2}-1}\right)$ находится в пределах от $\mu_3 = 8/3$ до $\mu_{\log 6} = 23/4$.

Следствие 2 *В квазиевклидовом кольце разложение прямоугольной матрицы $\mathcal{A}$ размера $n \times m$ можно выполнить с асимптотической оценкой сложности, не превосходящей*
$C_t^0(n,m) = C_D(n^2-n)/2 + \alpha n^\beta(\mu_\beta + (m-n)/n)$ *при* $n < m$,
$C_t^1(n,m) = C_D(nm - (m^2+m)/2) + \alpha\mu_\beta m^\beta +$
$\alpha(n-m)^{\beta-1}(\lambda_\beta m(n-m) + 10mn - 4n^2 - 4m^2)$, *при* $m \le n \le 2m$ *и*
$C_t^2(n,m) = C_D(nm - (m^2+m)/2) +$
$\alpha m^{\beta-1}(\mu_\beta m + \lambda_\beta(n-m) + 4n - 8m)$, *при* $n \ge 2m$.

Доказательство.

Доказательство сразу следует из оценок для $C_t(n)$ и $C_s(n)$.

При $n < m$ достаточно разложить левый квадратный блок, а затем умножить правый блок из $m-n$ столбцов матрицы $\mathcal{A}$ на $\tilde{\mathcal{A}}$.

При $n > m$ достаточно разложить верхний квадратный блок и воспользоваться следствием Леммы 1.

Теорема 2 *В квазиевклидовом кольце всякая система линейных уравнений $\mathcal{A}y = c$ с верхней треугольной матрицей коэффициентов $\mathcal{A} \in \mathbf{R}^{n\times n}$ может быть решена с асимптотической оценкой сложности*

$$C_n = C_D(n^2-n)/2 + \alpha\phi_\beta n^\beta,$$

$$\phi_\beta = \frac{1}{2^\beta-2}\left(5 + \frac{7}{2^{\beta-2}-1}\right).$$

Доказательство.

Решение системы $\mathcal{A}y = c$ можно свести к решению двух систем вдвое меньшего размера и одному матричному разложению.

Пусть $\begin{pmatrix} A & B \\ 0 & D \end{pmatrix}\begin{pmatrix} y_1 \\ y_0 \end{pmatrix} = \begin{pmatrix} f \\ h \end{pmatrix}$ – блочная запись системы $\mathcal{A}y = c$. Все блоки матрицы $\mathcal{A}$ квадратные, блоки A, D – верхние треугольные.

(1). Пусть решением системы $Dy_0 = h$ будет $y_0 = Vs_0 + v$, $V \in \mathbf{R}^{n/2\times n/2}$, $v, s_0 \in \mathbf{R}^{n/2}$, s_0 – вектор параметров. Подставим его в систему $Ay_1 + By_0 = f$. Получим $Ay_1 + Gs_0 = g$, где $G = BV$, $g = f - Bv$.

(2). Разложим матрицу (A, G):
$(A,G)P = (\bar{A}, 0)$. P - унимодулярная матрица, $\bar{A}$ - верхняя треугольная. По лемме 1 и замечанию к лемме оценка сложности такого разложения матрицы (A, G) составляет $C_s(n/2)$.

(3). Пусть решением системы $\bar{A}z_1 = g$ будет $z_1 = Wt_1 + w$, $W \in \mathbf{R}^{n/2 \times n/2}$, $w, t_1 \in \mathbf{R}^{n/2}$, t_1 – вектор параметров.

Тогда $\begin{pmatrix} z_1 \\ z_0 \end{pmatrix} = \mathrm{diag(W,I)}\begin{pmatrix} t_1 \\ t_0 \end{pmatrix} + \begin{pmatrix} w \\ 0 \end{pmatrix}$ будет решением системы $(\bar{A}, 0)\begin{pmatrix} z_1 \\ z_0 \end{pmatrix} = g$. Пусть $P = \begin{pmatrix} P_a & P_b \\ P_c & P_d \end{pmatrix}$ – блочная запись матрицы P. Так как $\begin{pmatrix} y_1 \\ s_0 \end{pmatrix} = P\begin{pmatrix} z_1 \\ z_0 \end{pmatrix}$ и $y_0 = Vs_0 + v$, то решение исходной системы $\bar{A}y = c$ получим в виде

$$y = \begin{pmatrix} P_aW & P_b \\ VP_cW & VP_d \end{pmatrix}\begin{pmatrix} t_0 \\ t_1 \end{pmatrix} + \begin{pmatrix} P_aw \\ VP_cw + v \end{pmatrix}.$$

Всего требуется выполнить пять блочных умножений. Обозначим C_n сложность решения системы порядка n. Тогда
$C_n = 2C_{n/2} + 5M(n/2) + C_s(n/2) = 2C_{n/2} + \alpha(5 + \lambda_\beta)n^{n/2} + (n/2)^2C_D$. Отсюда получаем асимптотическую оценку

$$C_n = C_D(n^2 - n)/2 + \alpha\phi_\beta n^\beta,$$

$$\phi_\beta = \frac{1}{2^\beta - 2}\big(5 + \frac{7}{2^{\beta-2} - 1}\big).$$

Теорема 3 *В квазиевклидовом кольце всякая система линейных уравнений $\mathcal{A}x = c$, $\mathcal{A} \in \mathbf{R}^{n\times m}$ может быть решена со следующими асимптотическими оценками сложности*

$$C_{nm} = (nm - l)C_D + \alpha\psi_\beta l^\beta,$$

$\psi_\beta = \mu_\beta + \phi_\beta + \lambda_\beta\frac{n-m}{m} + \frac{4n}{m} - 8$ *при* $2m \le n$;
$\psi_\beta = \mu_\beta + \phi_\beta + \lambda_\beta\frac{m-n}{n} + \frac{6m}{n} - 9$ *при* $2n \le m$;

$$C_{nm} = (nm - l)C_D + \alpha\psi_\beta l^\beta +$$

$$+\alpha|n - m|^{\beta-2}(\lambda_\beta l|m - n| + 10nm - 4m^2 - 4n^2),$$

$\psi_\beta = \mu_\beta + \phi_\beta$ *при* $m \le n \le 2m$,
$\psi_\beta = \mu_\beta + \phi_\beta + 2\frac{m}{n} - 1$ *при* $n < m \le 2n$,
где $l = \min(m, n)$.

Доказательство. Матрица $(\mathcal{A}, c)$ раскладывается в произведение $(\mathcal{A}, c) = P(\mathcal{A}', c')$ унимодулярной и верхней треугольной за $C_t^*(n, m)$ операций по следствию 2 теоремы 1. Следовательно, достаточно решить систему

$$\mathcal{A}'x = c' \tag{3}$$

с верхней треугольной матрицей коэффициентов.

a. Пусть $m > n$, $\mathcal{A}' = (A, B)$, A – квадратный блок, имеющий верхний треугольный вид. По замечанию к лемме 1 можно построить унимодулярную матрицу Q, такую, что $(A, B)Q = (\bar{A}, 0)$, где $\bar{A}$ – верхняя треугольная, за $C_s^*(n, m)$ операций. Решения системы (3) и системы $\bar{A}y = c'$ связаны матрицей Q: $x = Q\begin{pmatrix} y \\ t_0 \end{pmatrix}$, где t_0 – вектор параметров. Пусть решением системы $\bar{A}y = c'$ будет $Ut_1 + u$, $U \in \mathbf{R}^{n\times n}$, $t_1, u, y \in \mathbf{R}^n$, t_1 – вектор параметров, $t_1 \in \mathbf{R}^n$. Тогда решением исходной системы будет $x = Q\mathrm{diag(U,I)t + Q}\begin{pmatrix} u \\ 0 \end{pmatrix}$, где $t \in \mathbf{R}^m$ – вектор параметров. Для вычисления матричного произведения $Q\mathrm{diag(U,I)}$ нужно $\alpha n^{\beta-1}m$ операций. По теореме 2 решение системы $\bar{A}y = c'$ можно получить за C_n операций.

Суммируя все операции, получим оценку сложности такого решения

$$C_{nm}^{1i} = C_t^0(n, m) + C_s^i(m, n) + C_n + M(n)(m/n),$$

$i = 1, 2$, где $i = 1$ соответствует случаю $m \ge 2n$, $i = 2$ соответствует случаю $2n \ge m > n$.

b. Пусть $n \ge m$, $\mathcal{A}' = \begin{pmatrix} A \\ B \end{pmatrix}$, A – квадратный блок, имеющий верхний треугольный вид. По теореме 2 решение системы $Ax = c'$ можно получить за C_m операций, согласно теореме 2. Следовательно, общее число операций для этого решения будет

$$C_{nm} = C_t^i(n, m) + C_m \quad i = 1, 2, \tag{4}$$

где $i = 1$ соответствует случаю $2m \ge n \ge m$, $i = 2$ соответствует случаю $n \ge 2m$.

Утверждение теоремы получается подстановкой всех оценок в правую часть.

В частности для случая квадратной матрицы коэффициентов системы получим

$$C_{nn} = n^2C_D + \alpha\psi_\beta n^\beta, \;\; \psi_\beta = \frac{7 \cdot 2^{\beta-1}}{(2^\beta - 2)(2^{\beta-1} - 2)}.$$

Для $\beta = 3$: $\psi_\beta < 4,67$. Для $\beta = \log 7$: $\psi_\beta < 6,54$. Для $\beta = \log 6$: $\psi_\beta = 10,5$.

ПОДЪЕМ ПО ГЕНЗЕЛЮ

Пусть $\mathbf{R}$ – коммутативная область с единицей, $\mathbf{K}$ – поле частных области $\mathbf{R}$, $A \in \mathbf{R}^{n\times m}$, $c \in \mathbf{R}^n$

$$Ax = c \tag{5}$$

– система линейных уравнений над $\mathbf{R}$.

Будем полагать, что $n < m$, $\mathrm{rank}A = n$, $k = m - n + 1$ – число линейно независимых решений системы (5). Пусть $A = (A_0, A_1)$, блок A_0 квадратный, $\det A_0 \ne 0$ и последняя строка матрицы $A_0^{-1}(A_1, c)$ не содержит нулей. Общий случай сводится к данному, например, так, как предлагается в [4]. Пусть $\mathbf{e}_j$– столбцы единичной матрицы $I_k = (\mathbf{e}_1, \dots, \mathbf{e}_k)$ порядка k.

Если искать решение системы (5) в виде $\begin{pmatrix} x_j \\ \xi_j \mathbf{e}_j \end{pmatrix}$, $x_j \in \mathbf{K}^{n-1}$, $\xi_j \in \mathbf{K}\backslash\{0\}$, $(j = 1, \dots, k)$, то придем к определенной системе порядка k.

Решение такой системы в поле $\mathbf{K}$ с помощью p-адического подъема по Гензелю можно получить за

$$\mathcal{D}_Z = O(n^3(\log n + \log \|A\| + \log p)^2)$$

бит-операций, если $\mathbf{R}$ – кольцо целых чисел $\mathbf{Z}$. Если же $\mathbf{R} = F[x]$ – кольцо полиномов над полем F, то достаточно

$$\mathcal{D}_{F[x]} = O(n^3(\deg A + \deg p)^2)$$

операций в поле F [3, 6]. Для получения k решений потребуется соответственно $k\mathcal{D}_Z$ и $k\mathcal{D}_{F[x]}$ операций.

Так как $\xi_j \neq 0, j = 1, \dots, k$, то все k решений будут линейно независимы. Приводя каждое из решений к общему знаменателю, который является делителем некоторого минора матрицы A, запишем эти решения в виде

$$x_j = \begin{pmatrix} \mathbf{x}_j \\ \alpha_j \mathbf{e}_j \end{pmatrix} \chi_j^{-1}, \quad j = 1, \dots, k,$$

$\mathbf{x}_j \in \mathbf{R}^{n-1}$, $\alpha_j, \chi_j \in \mathbf{R}$.

Обозначим $\delta = (\alpha_1, \dots, \alpha_\mathrm{k})$ и отметим, что δ – делитель минора матрицы (A, c), образованного первыми $n-1$ столбцами и последним столбцом.

Домножим числитель и знаменатель каждого решения j на $\delta\alpha_j^{-1}$. Обозначим $\mathbf{x}_j\delta\alpha_j^{-1} = \mathbf{v}_j$, $\chi_j\delta\alpha_j^{-1} = \delta_j$. Получим базисное множество решений системы (5) в виде $x_j = \begin{pmatrix} \mathbf{v}_j \\ \delta \mathbf{e}_j \end{pmatrix} \delta_j^{-1}$.

Обозначим $\mathbf{u} = (\delta_1, \dots, \delta_k)$ – вектор и $V = (\mathbf{v}_1, \dots, \mathbf{v}_k)$ – матрицу, $\mathbf{u} \in \mathbf{R}^k$, $V \in \mathbf{R}^{(n-1)\times k}$.

Теорема 4 *Для того чтобы была совместна система (5) в кольце $\mathbf{R}$, необходимо и достаточно, чтобы была совместна в $\mathbf{R}$ система*

$$\left\{ \begin{array}{lcl} Vy & = & 0 \bmod \delta \\ \mathbf{u}y & = & \delta. \end{array} \right. \tag{6}$$

Если Y – множество решений ситемы (6), то $X = \left\{ \begin{pmatrix} y \\ Vy\delta^{-1} \end{pmatrix} \mid y \in Y \right\}$ – множество решений системы (5).

Доказательство. Множество решений системы – это гиперплоскость размерности $m-n$ в пространстве $\mathbf{K}^m$. Она определяется $m-n+1$ линейно независимым решением. Следовательно, для всякого решения $z \in \mathbf{K}^m$ системы (5) найдутся числа $\sigma_i \in \mathbf{K}$, $i = 1, \dots, k$, такие что $\sum_{i=1}^k x_i\sigma_i = z$ и $\sum_{i=1}^k \sigma_i = 1$. Обозначим $\sigma_i\delta_i^{-1} = \tau_i \in \mathbf{K}$, $z = (y, w)$, $y \in \mathbf{R}^k$, $w \in \mathbf{R}^{m-k}$, $t = (\tau_1, \dots, \tau_k)$. Тогда

$$\sum_{i=1}^k \delta_i\tau_i x_i = z, \quad \mathbf{u}t = 1.$$

Если рассматривать (t, y, w) как вектор неизвестных, то получим систему уравнений

$$\begin{pmatrix} \delta I_k & -I_k & 0 \\ V & 0 & -I_{n-1} \\ \mathbf{u} & 0 & 0 \end{pmatrix} \begin{pmatrix} t \\ y \\ w \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Домножая каждую часть на обратимую в $\mathbf{K}$ матрицу

$$\begin{pmatrix} I_k & 0 & 0 \\ -V & \delta I_{n-1} & 0 \\ -\mathbf{u} & 0 & \delta \end{pmatrix},$$

приходим к системе

$$\begin{pmatrix} \delta I_k & -I_k & 0 \\ 0 & V & -\delta I_{n-1} \\ 0 & \mathbf{u} & 0 \end{pmatrix} \begin{pmatrix} t \\ y \\ w \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \delta \end{pmatrix} \tag{7}$$

Пусть система (5) имеет диофантово решение $\mathbf{z} = (\mathbf{y}, \mathbf{w}) \in \mathbf{R}^m$. Тогда совместна система (7). Следовательно совместна и система (6). Наоборот, пусть совместна система (6) и $\mathbf{y} \in \mathbf{R}^k$ – ее решение, тогда $(\delta^{-1}\mathbf{y}, \mathbf{y}, \delta^{-1}V\mathbf{y})$ решение системы (7). Следовательно $\mathbf{z} = (\mathbf{y}, \delta^{-1}V\mathbf{y})$ – диофантово решение системы (5).

Оценим сложность решения системы (6) в кольце целых чисел и кольце многочленов.

Матрица V имеет размеры $(n-1)\times k$, а модуль δ является делителем минора порядка n. Следовательно, для $\mathbf{R} = \mathbf{Z}$

$$\log \delta \leq n(\log \|A\| + 0.5 \log n), \tag{8}$$

где $\|A\|$ – модуль наибольшего элемента матрицы A. Для $\mathbf{R} = F[z]$

$$\deg \delta \leq n \deg \|A\|, \tag{9}$$

где $\deg \|A\|$ – степень наибольшего по степени элемента матрицы A.

Пусть применяется стандартный алгоритм умножения целых чисел и многочленов: для умножения N- битных чисел требуется N^2 бит-операций, а для умножения двух многочленов степени N требуется N^2 операций умножения в поле F.

При $k \geq n$ по формуле (3) получим асимптотические оценки числа операций соответственно для $\mathbf{Z}$ и $F[x]$:

$$\eta_\beta knM(n)(\log\|A\| + 0.5\log n)^2$$

$$\eta_\beta knM(n)\deg^2\|A\|.$$

При $k \leq n$ по формуле (4) получим асимптотические оценки числа операций соответственно для кольца целых чисел $\mathbf{Z}$:
$\delta k^{\beta-1}(\log\|A\| + \frac{\log n}{2})^2(n^3(4+\lambda_\beta)$
$+n^2k(\nu_\beta + \mu_\beta - \lambda_\beta - 4))$ и для кольца многочленов над полем $F[x]$:
$\delta k^{\beta-1}\deg^2\|A\|(n^3(4+\lambda_\beta)$
$+n^2k(\nu_\beta + \mu_\beta - \lambda_\beta - 4))$.

Отметим, что в общем случае сложность вычисления k решений систем в поле K не влияет на эти асимптотические оценки.

Однако, для квазиквадратных систем это не так, ибо величина k мала по сравнению с n, и ей можно пренебречь. В этом случае для $\mathbf{Z}$ и $F[x]$ получим, соответственно, следующие оценки числа операций:

$$O(n^3(\log n + \log\|A\| + \log p)^2)$$

$$O(n^3(\deg\|A\| + \deg p)^2).$$

Если применять быстрые методы умножения целых чисел и многочленов, то эти оценки можно еще улучшить. Так, метод Карацубы требует $O(N^{\log 3})$ операций, а сегодняшний рекорд составляет – $O(N\log N\log\log N)$.

МОДУЛЯРНЫЙ МЕТОД

Будем решать систему (5), полагая что $A = (A_0, A_1)$, блок A_0 квадратный, $\det A_0 = \delta$, $\delta \neq 0$. Рекурсивный метод [7] позволяет привести систему (5) к диагональному виду

$$(\delta I_n, A'_1)x = c' \qquad (10)$$

за $O(n^{\beta-1}m)$ операций в кольце $\mathbf{R}$.

Все коэффициенты системы (10) – это миноры порядка n матрицы (A, c). Следовательно, на них распространяются оценки (8) и (9), соответственно, для кольца целых чисел и кольца многочленов над полем F.

Если применять стандартное умножение, то количество операций, которое потребуется для вычисления коэффициентов системы (10) для колец $\mathbf{Z}$ и $F[x]$, соответственно будет

$$O(mn^{\beta+1}(\log n + \log\|A\|)^2) \text{ и } O(mn^{\beta+1}\deg^2\|A\|).$$

Китайская теорема об остатках, составляющая основу модулярного метода, позволяет уменьшить количество операций до

$$O(mn^\beta(\log n + \log\|A\|)^2) \text{и} O(mn^\beta\deg^2\|A\|). \qquad (11)$$

Решение системы (10) сводится к решению системы

$$A'_1 y = c' \bmod \delta. \qquad (12)$$

в факторкольце $\mathbf{R}/\delta\mathbf{R}$. Пусть Y – множество решений системы (12), тогда

$$X = \{ \begin{pmatrix} (c' - A'_1 y)\delta^{-1} \\ y \end{pmatrix} \mid y \in Y \}$$

– множество решений системы (10).

Оценим сложность решения системы (12). Матрица A' имеет размеры $n \times (m-n)$, а модуль δ является минором порядка n. По теореме 3 получим следующие оценки.

При $m - n \geq n$ получим соответственно для $\mathbf{Z}$ и $F[x]$:

$$O(n^{\beta+1}m(\log\|A\| + \log n)^2),$$

$$O(n^{\beta+1}m\deg^2\|A\|). \qquad (13)$$

При $m - n \leq n$ получим соответственно для $\mathbf{Z}$ и $F[x]$:

$$O(n^3(m-n+1)^{\beta-1}(\log\|A\| + \log n)^2),$$

$$O(n^3(m-n+1)^{\beta-1}\deg^2\|A\|). \qquad (14)$$

В общем случае оценки (13), (14) являются оценками сложности решения системы модулярным методом, так как оценка (11) имеет меньший порядок по n и $m-n$. Однако для квазиквадратных систем величиной $m-n$ можно пренебречь и оценка (11) будет давать основной вклад. В этом случае для $\mathbf{Z}$ и $F[x]$ получим соответственно следующие оценки числа операций:

$$O(n^{\beta+1}(\log n + \log\|A\|)^2) \text{ и } O(n^{\beta+1}\deg^2\|A\|).$$

ЗАКЛЮЧЕНИЕ

Сравнение p-адического и модулярного метода показывает, что для квазиквадратных систем p-адический метод имеет существенное преимущество, он быстрее в $O(n^{\beta-2})$ раз.

Другим достоинством p-адического метода для систем прозвольного вида является тот факт, что общее решение системы может быть получено с большой вероятностью за время $O((m-n+1)n^3)$ [6]. Ожидаемое количество итераций при этом будет

$$O(\frac{\log n + \log\log\|A\|}{m-n+1}).$$

Каждая новая итерация начинается со случайного выбора ведущего минора порядка n.

Для существенно прямоугольных систем, когда $m > 2n$, оба метода имеют одинаковый порядок сложности $O(mn^{\beta+1})$.

Если пользоваться быстрыми алгоритмами умножения целых чисел, то эти оценки можно улучшить пропорционально ускорению алгоритма умножения.

ЛИТЕРАТУРА

1. *Gregory R.T., Krishnamurthy E.V.* Methods and Applications of Error-Free Computation. Berlin, 1984.

2. *Hurt M.F., Waid C.A.* A generalized inverse wich give all the integral solutions to a system of linear equations// SIAM J. Appl. Math. V. 10. 1970. P. 547–550.

3. *Mulders T. and Storjohann A.* Diophantine Linear System Solving// Proceedings of ISSAC'99: ACM International Symposium on Symbolic and Algebraic Computation. July 1999, Vancouver, Canada. P. 181–188.

4. *Malaschonok G.* Effective Matrix Methods in Commutative Domains// Formal Power Series and Algebraic Combinatorics/ Ed. by D. Krob, A.A. Mikhalev, A.V. Mikhalev. Springer, 2000. P. 506–517.

5. *Coppersmith D. and Winograd S.* Matrix multiplication via arithmetic progressions// Journal of Symbolic Computation. 1990. V. 9. P. 251–280.

6. *Dixon J.* Exact solution of linear equations using p-adic expansions// Numer. Math. 1982. V. 40. P. 137–141.

7. *Malaschonok G.* Recursive Method for the Solution of systems of Linear Equations// Computational Mathematics/ Ed. by A. Sydow. Proceedings of the 15th IMACS World Congress. Berlin, 1997. V. 1. P. 475–480.

Поступила в редакцию 1 октября 2000 г.