Solution of Systems of Linear Diophantine Equations ⋆

Gennadi I. Malaschonok

Tambov State University,
33 Internatsionalnaya, 392622 Tambov, Russia
malaschonok@math-iu.tambov.su

Abstract. Two new methods to solve linear systems of Diophantine equations are proposed – modular (CRT) and p-adic (Hensel). Each of them allows to obtain solutions of a system with the size $n \times m$ with the complexity $O(n^{\beta} m)$. For quasi-square systems, the p-adic method allows to get solution with the complexity $O(n^3)$, and the modular method – $O(n^{\beta+1})$. Both estimations have the accuracy up to the logarithmic multipliers, β being the power in the estimation of matrix multiplication time.

1 Introduction

The usual method to solve systems of linear Diophantine equations is based on the computation of Smith form of the coefficient matrix [6], [5]. In the recent papers [11], [10] the *probabilistic* methods for finding the solutions of Diophantine systems were obtained. The main lack of these methods is the probabilistic character of the solution and an absence of a capability to prove the incompatibility of a system. These methods do not allow to distinguish a compatible system, in which the solution is not obtained in the process of solving, from an incompatible system.

The main result of the article is the following.

There are offered deterministic p-adic and modular methods which are independent of the Smith form computation. Both methods reduce the problem to the solving of systems of linear equations in the residue ring.

The p-adic method is based on the linear p-adic Hensel lifting. This is a way to obtain the basic set of system solutions on the plane of all solutions. Then we compute all integer points on this plane.

The second method is based on the Chinese remainder theorem. The result of its application is the diagonal form of a non-singular submatrix of maximal rank. Then we solve the system in modulus which equals the determinant of this submatrix.

The complexity of the solution is estimated for systems in the ring of integers $\mathbf{Z}$ and in the ring of polynomials $F[x]$ over a field F.

⋆ in: Computer Algebra in Scientific Computing - CASC'01, Springer, Berlin, 2001

For essentially rectangular systems, when $m > 2n$, both methods have the same order of the complexity – $O(n^{\beta+1}m)$ within a logarithmic factor. For quasi-square systems the complexity of a p-adic method is $O(n^3)$ and the complexity of a modular method is $O(n^{\beta+1})$. Here β is the power in the estimation of complexity of matrix multiplication. Known for today value of β is less than 2.356 [3].

Next section has an auxiliary character, it is devoted to the problem of solving systems of linear equations in a residue ring. Here is presented a method with complexity $O(n^{\beta})$ ring operations. For a system of the order n the asymptotic evaluation of this method is less than $6.54M(n)$, where $M(n)$ is the asymptotic evaluation of Strassen multiplication of matrices of order n.

2 QuasiEuclidean rings

Let $\mathbf{R}$ be a commutative ring with identity, $\alpha \in \mathbf{R}^2$ – be a non-zero pair. An unimodular matrix $E_\alpha \in \mathbf{R}^{2\times 2}$ is called the *Diophantine matrix* of α if

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

The *unit matrix* is called the Diophantine matrix of a *zero* pair.

We must select a special class of the rings, in which an algorithm for computing the Diophantine matrix exists.

Definition 1 . *A ring $\mathbf{R}$ is called quasiEuclidean, if an algorithm for computing the Diophantine matrix for any $\alpha \in \mathbf{R}^2$ exists.*

If $\alpha = \begin{pmatrix} p \\ q \end{pmatrix}$, then $\begin{pmatrix} p \\ q \end{pmatrix} = E_\alpha^{-1} \begin{pmatrix} g \\ 0 \end{pmatrix}$. So the ideal (p, q) is a principal ideal generated by g. Therefore each of the quasiEuclidean rings is a ring of principal ideals, where the algorithm for computing of ideal generator exists.

To the contrary. Let $\mathbf{R}$ be the ring of principal ideals, in which an algorithm for calculation of the generator g of an ideal (p, q), $p, q \in \mathbf{R}$ exists, denote it GI algorithm:

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

For a non-zero pair $\alpha = \begin{pmatrix} p \\ q \end{pmatrix}$, a Diophantine matrix is $E_\alpha = \begin{pmatrix} a & b \\ -q/g & p/g \end{pmatrix}$.

Therefore *a commutative ring is quasiEuclidean if and only if it is a ring of principal ideals, where an algorithm for computing of ideal generator exists.*

Note that the numbers a, b and g are not uniquely defined. Therefore the Diophantine matrix is not uniquely defined. The Diophantine matrix can be determined uniquely if an order is introduced into $\mathbf{R}$.

If $\mathbf{R}$ is the Euclidean ring, then the extended Euclidean algorithm may be taken as GI algorithm.

Proposal 1 . *QuasiEuclidivity of a ring is inherited by its quotient-ring.*

A quatient-ring of the ring of principal ideal is a ring of principal ideal. So to prove this proposal we need to show how to construct the Diophantine matrices in quatient-ring.

Let $\mathbf{R}'$ be a quotient-ring of a quasiEuclidean ring $\mathbf{R}$ with some order relation, and

$$\phi : \mathbf{R} \to \mathbf{R}'$$

be a canonical homomorphism. Let us introduce an order relation in R, and define a function

$$\psi : \mathbf{R}' \to \mathbf{R},$$

selecting a representative of a coset for $p' \in \mathbf{R}'$ in a pre-image $\phi^{-1}(p')$ by a standard way: $\psi(p') = min(\phi^{-1}(p'))$. Let $\psi(p') = p$, $\psi(q') = q$, $\alpha = \begin{pmatrix} p \\ q \end{pmatrix}$, E_α – Diophantine matrix for α and $E_\alpha \alpha = \begin{pmatrix} g \\ 0 \end{pmatrix}$.

For a non-zero pair $\alpha' = \begin{pmatrix} p' \\ q' \end{pmatrix} \in \mathbf{R}'^2$ we define a Diophantine matrix $E_{\alpha'}$ as follows: $E_{\alpha'} = \phi(E_\alpha)$. Such definition is correct since

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

Let's show, that $g' = \phi(g) \neq 0$. Due to the invertibility of Diophantine matrices, p and q are divided by g. Therefore if $\phi(g) = 0$, then $\phi(p) = \phi(q) = 0$. But by the condition, at least one of the numbers $p' = \phi(p)$, $q' = \phi(q)$ is not a zero.

Therefore a set of quasiEuclidean rings contains all Euclidean rings and their quotient-rings.

Let now $\mathbf{R}$ be a quasiEuclidean ring, $A \in \mathbf{R}^{n \times m}$. A matrix A is called *upper triangular*, if all elements under the main diagonal are zeros: $(i, j) = 0$ for all $n \geq j > i \geq 1$. (We say nothing about other matrix elements.)

We say that the matrix $\mathcal{A}$ is *decomposable*, if it is a product of an unimodular matrix $\tilde{\mathcal{A}}$ and an upper triangular matrix $\bar{\mathcal{A}}$: $\mathcal{A} = \tilde{\mathcal{A}}\bar{\mathcal{A}}$.

Proposal 2 . *In a quasiEuclidean ring any matrix is decomposable.*

Proof. We compute such decomposition for a matrix $A = (a_{ij})$ of the size $n \times m$. Let $\alpha_{ij} = \begin{pmatrix} a_{jj} \\ a_{ij} \end{pmatrix}$, $i > j$, $E_{\alpha_{i,j}}$ be a Diophantine matrix for α_{ij}. Unimodular matrix $E^n_{\alpha_{i,j}}$ is a matrix that is obtained from the unit matrix I_n of the order n after replacement of its four elements (jj), (ji), (ij), (ii) by the four elements (18), (19), (21), (22) of matrix $E_{\alpha_{i,j}}$, correspondingly.

Let $E_j = E_{\alpha_{n,j}} \cdots E_{\alpha_{j+1,j}}$, $U = E_{n-1} \cdots E_1$. Then U is a unimodular matrix and UA is an upper triangular matrix. Such computation leads to the elimination of non-zero elements under the diagonal element in the first column, then in the second column, and so on up to the $n-1$ column.

Denote by C_D an asymptotic estimation of the complexity of a Diophantine matrix calculation, $M(n)$ – an asymptotic estimation of the complexity of multiplying matrices of the order n, $M(n) = \alpha n^\beta$. The best known for today value β is less than 2.356 [3].

Lemma 1. *Let* $\mathbf{R}$ *be a quasiEuclidean ring* $A, C \in \mathbf{R}^{n\times n}$, *A is an upper square matrix. Then the matrix* $\begin{pmatrix} A \\ C \end{pmatrix}$ *may be decomposed* $\begin{pmatrix} A \\ C \end{pmatrix} = Q \begin{pmatrix} \bar{A} \\ 0 \end{pmatrix}$ *with an asymptotic estimation of complexity which is not superior than*

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

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

Proof. The decomposition of the 2×1 matrix needs the multiplying by its Diophantine matrix. Without loss of generality, we suppose, that $n = 2^p$ and decompose the matrices A and C into four square blocks:

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

Let us have four decompositions.

$Q_1 \begin{pmatrix} a \\ e \end{pmatrix} = \begin{pmatrix} a_1 \\ 0 \end{pmatrix}$, denote $\begin{pmatrix} b_1 \\ f_1 \end{pmatrix} = Q_1 \begin{pmatrix} b \\ f \end{pmatrix}$.

$Q_2 \begin{pmatrix} a_1 \\ g \end{pmatrix} = \begin{pmatrix} a_2 \\ 0 \end{pmatrix}$, denote $\begin{pmatrix} b_2 \\ h_1 \end{pmatrix} = Q_2 \begin{pmatrix} b_1 \\ h \end{pmatrix}$.

$Q_3 \begin{pmatrix} d \\ f_1 \end{pmatrix} = \begin{pmatrix} d_1 \\ 0 \end{pmatrix}$.

$Q_4 \begin{pmatrix} d_1 \\ h_1 \end{pmatrix} = \begin{pmatrix} d_2 \\ 0 \end{pmatrix}$.

Matrices $Q_i = \begin{pmatrix} \alpha_i & \beta_i \\ \gamma_i & \delta_i \end{pmatrix}$, $i = 1, \ldots, 4$, are unimodular, a, a_1, a_2, d, d_1, d_2 are upper triangular matrices. Then we get

$$\bar{A} = \begin{pmatrix} a_2 & b_2 \\ 0 & d_2 \end{pmatrix} \quad \text{and } Q = \begin{pmatrix} \alpha_1\alpha_2 & 0 & \beta_1\alpha_2 & \beta_2 \\ \alpha_4\beta_3\gamma_1 + \beta_4\gamma_2\alpha_1 & \alpha_4\alpha_3 & \alpha_4\beta_3\delta_1 + \beta_4\gamma_2\beta_1 & \beta_4\delta_2 \\ \delta_3\gamma_1 & \gamma_3 & \delta_3\delta_1 & 0 \\ \gamma_4\beta_3\gamma_1 + \delta_4\gamma_2\alpha_1 & \gamma_4\alpha_3 & \gamma_4\beta_3\delta_1 + \delta_4\gamma_2\beta_1 & \delta_4\delta_2 \end{pmatrix}.$$

We obtain matrix Q as a product

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \alpha_4 & 0 & \beta_4 \\ 0 & 0 & 1 & 0 \\ 0 & \gamma_4 & 0 & \delta_4 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \alpha_3 & \beta_2 & 0 \\ 0 & \gamma_2 & \delta_2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \alpha_2 & 0 & 0 & \beta_2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \gamma_3 & 0 & 0 & \delta_3 \end{pmatrix} \begin{pmatrix} \alpha_1 & 0 & \beta_1 & 0 \\ 0 & 1 & 0 & 0 \\ \gamma_1 & 0 & \delta_1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Therefore, these calculations need 28 multiplications of blocks of the order $n/2$ and four decompositions of blocks of the size $n \times n/2$.

We reduce step by step the calculations to the blocks of size 2×1. Finally we obtain the following asymptotic estimation of decomposition

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

Here the number $\lambda_\beta = 7(2^{\beta-2}-1)^{-1}$ lays between $\lambda_3 = 7$ and $\lambda_{\log 6} = 14$.

Corollary 1 . *Let* $\mathbf{R}$ *be a quasiEuclidean ring,* A *be an upper square matrix,* $A \in \mathbf{R}^{m\times m}$, $C \in \mathbf{R}^{(n-m)\times m}$. *Then the matrix* $\begin{pmatrix} A \\ C \end{pmatrix}$ *may be decomposed with an asymptotic estimation of complexity which is not superior*

$$C^1_s(n,m) = (n-m)mC_D + \alpha m^{\beta-1}(\lambda_\beta(n-m) + 4n - 8m),$$

for $n \geq 2m$ *and*

$$C^2_s(n,m) = (n-m)mC_D + \alpha(n-m)^{\beta-2}(\lambda_\beta m(n-m) + 10mn - 4n^2 - 4m^2),$$

for $m \leq n \leq 2m$.

Proof.

a. Let $n \geq 2m$ and m be a divisor of n. Let C_i, $i = 1, \dots, (n-m)/m$, be an m order block of the matrix C. We count the matrix C blocks from top to bottom. And then we decompose consequently the matrices:

$\begin{pmatrix} A_{i+1} \\ 0 \end{pmatrix} = Q_i \begin{pmatrix} A_i \\ C_i \end{pmatrix}$, $i = 1, \dots, (n-2m)/m$, where $A_1 = A$.

We need four multiplications of matrices of the size $m \times m$ at each step.

So the total number of operations is

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

b. Let $m \leq n \leq 2m$ and $n-m$ be a divisor of m. Let us break matrices A and C into blocks of the size $(n-m)\times(n-m)$. The matrix C contains one block row. We denote $C = (C_1, \bar{C}_1)$, here C_1 is a left block of the matrix C. The matrix A contains $k = m/(n-m)$ block rows and columns.

At first we decompose the matrix that is formed by the upper left diagonal block A_1 of the matrix A and the left block C_1 of the matrix C:

$\begin{pmatrix} A_1 \\ C_1 \end{pmatrix} = Q_1 \begin{pmatrix} \bar{A}_1 \\ 0 \end{pmatrix}$, Q_1 is a unimodular matrix.

We compute the product $\begin{pmatrix} A'_1 \\ C'_2 \end{pmatrix} = Q_1 \begin{pmatrix} \bar{A}_1 \\ \bar{C}_1 \end{pmatrix}$, here $\bar{A}_1$ is a part of the first block row of the matrix A situated on the left from the diagonal block A_1. Denote $C'_2{=}(C_2, \bar{C}_2)$, here C_2 is a left square block. We need to do $4(k-1)$ block multiplications.

Then we sequentially decompose matrices that is formed by the i-th diagonal block A_i of the matrix A and the left block C_i of the matrix C'_i:

$\begin{pmatrix} A_i \\ C_i \end{pmatrix} = Q_i \begin{pmatrix} \bar{A}_i \\ 0 \end{pmatrix}$, and compute the product $\begin{pmatrix} A'_i \\ C_{i+1} \end{pmatrix} = Q_i \begin{pmatrix} \bar{A}_i \\ \bar{C}_i \end{pmatrix}$, $i = 2, \dots, k$.

Here Q_i is a unimodular matrix, $\bar{A}_i$ is the part of the i-th block row of the matrix A situated on the left from the diagonal block A_i, $C'_i{=}(C_i, \bar{C}_i)$.

Substitute the blocks (1,1), (1,2), (2,1), (2,2) of the matrix Q_i instead of the blocks (i,i), (i,k+1), (k+1,i), (k+1,k+1) of the unit matrix I_n correspondingly.

As the result we get the matrix Q'_i. To obtain the unimodular matrix Q there is required to multiply the unimodular matrices Q'_i:

$$Q = Q'_k \cdots Q'_2 Q'_1.$$

One such matrix multiplication needs four block multiplications and the i-th step needs $4(k-i)+4$ block multiplications, $i = 2, \ldots, k$.

So the total number of operations is $(\sum_{i=1}^{k} 4(k-i+1) - 4)M(n-m) + kC_s(n-m) + < kC_s(n-m) + 2(k^2+k-2)M(n-m) = m(n-m)C_D + M(n-m)(\lambda_\beta m/(n-m) + 2(5mn - 2n^2 - 2m^2)/(n-m)^2) = C_s^2(n,m)$.

Remark. *It is easy to show that we can change the upper triangular matrix A to a lower triangular matrix in the condition of Lemma 1. We also can change the order of multipliers in the decomposition formula, the evaluations in the lemma and in the corollary will be the same.*

Theorem 1. *In a quasiEuclidean ring any quadratic matrix $\mathcal{A}$ of the order n may be decomposed with an asymptotic estimation of complexity which is not superior*

$$C_t(n) = \mu_\beta M(n) + \frac{n^2 - n}{2} C_D, \tag{2}$$

where $\mu_\beta = \frac{9 \cdot 2^{\beta-2} + 19}{(2^\beta - 2)(2^{\beta-2} - 1)}$.

Proof. The decomposition of a second order matrix needs the multiplying of its first column by the Diophantine matrix.

We suppose that $n = 2^p$ and decompose the matrix $\mathcal{A}$ into four square blocks:

$$\mathcal{A} = \begin{pmatrix} A & B \\ C & D \end{pmatrix}.$$

1. Let the block A be decomposable:

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

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

2. Let the block $\begin{pmatrix} A_1 \\ C \end{pmatrix}$ be decomposable:

$$\begin{pmatrix} A_1 \\ C \end{pmatrix} = Q^{-1} \begin{pmatrix} A_2 \\ 0 \end{pmatrix}, \tag{3}$$

Q – unimodular, A_2 – upper triangular. Denote $\begin{pmatrix} B_2 \\ D_1 \end{pmatrix} = Q \begin{pmatrix} B_1 \\ D \end{pmatrix}$.

3. Let the block D_1 be decomposable: $D_1 = \tilde{D}_1^{-1} \bar{D}_1$. Denote $D_2 = \bar{D}_1$.

Then the matrix A is decomposable:

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

Such decomposition requires 9 block multiplications, two block decompositions and one decomposition of the rectangular $n \times n/2$ block.

We denote by $C_t(n)$ the complexity of decomposition of a matrix of the order n. Then we obtain according to (1)

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

According to the lemma 1 we receive

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

We shall break matrices A and D into the four square blocks, and continue such procedure up to the second order blocks.

The asymptotic estimation of the complexity is

$$\sum_{i=1}^{p} 2^{i-1}\big(\big(\frac{n}{2^i}\big)^2 C_D + \alpha(9+\lambda_\beta)\big(n/2^i\big)^\beta\big) < \frac{n^2-n}{2} C_D + \alpha\mu_\beta n^\beta = C_t(n)$$

where the number $\mu_\beta = \frac{1}{2^\beta - 2}\big(9 + \frac{7}{2^{\beta-2}-1}\big)$ lays between $\mu_3 = 8/3$ and $\mu_{\log 6} = 23/4$.

Corollary 2 *In a quasiEuclidean ring any rectangular matrix $\mathcal{A}$ of the size $n \times m$ is decomposable. Such decomposition is possible to execute with an asymptotic estimation of complexity which is not higher, than*
$C_t^0(n,m) = C_D(n^2-n)/2 + \alpha n^\beta(\mu_\beta + (m-n)/n)$, *for* $n < m$,
$C_t^1(n,m) = C_D(nm - (m^2+m)/2) + \alpha\mu_\beta m^\beta + \alpha(n-m)^{\beta-1}(\lambda_\beta m(n-m) + 10mn - 4n^2 - 4m^2)$, *for* $m \le n \le 2m$ *and*
$C_t^2(n,m) = C_D(nm - (m^2+m)/2) + \alpha m^{\beta-1}(\mu_\beta m + \lambda_\beta(n-m) + 4n - 8m)$, *for* $n \ge 2m$.

Proof. The proof follows from the estimations for $C_t(n)$ and $C_s(n)$ in (1) and (2).

For $n < m$ it is enough to decompose the left square block and then to multiply the left block from $m-n$ columns of a matrix $\mathcal{A}$ by $\tilde{\mathcal{A}}$.

For $n \ge m$ it is enough to decompose the upper square block and to use the corollary of Lemma 1.

Theorem 2. *In a quasiEuclidean ring any system of linear equations $\mathcal{A}y = c$ with the upper triangular coefficient matrix $\mathcal{A} \in \mathbf{R}^{n\times n}$ may be solved with an asymptotic estimation of complexity*

$$C_n = C_D(n^2-n)/2 + \alpha\phi_\beta n^\beta, \quad \phi_\beta = \frac{1}{2^\beta - 2}\big(5 + \frac{7}{2^{\beta-2}-1}\big). \tag{4}$$

Proof. Solving the system $\mathcal{A}y = c$ is reduced to solving two systems of twice less size and one matrix decomposing.

Let $\begin{pmatrix} A & B \\ 0 & D \end{pmatrix} \begin{pmatrix} y_1 \\ y_0 \end{pmatrix} = \begin{pmatrix} f \\ h \end{pmatrix}$ be the block form of the system $\mathcal{A}y = c$. All the matrix $\mathcal{A}$ blocks are square and A, D are the upper triangular blocks.

(1). Let $y_0 = Vs_0 + v$ be a solution of the system $Dy_0 = h$, $V \in \mathbf{R}^{n/2 \times n/2}$, $v, s_0 \in \mathbf{R}^{n/2}$, s_0 is a vector of parameters. Substituting it into the system $Ay_1 + By_0 = f$ we obtain $Ay_1 + Gs_0 = g$, here $G = BV$, $g = f - Bv$.

(2). Decompose the matrix (A, G): $(A, G)P = (\bar{A}, 0)$. P is a unimodular matrix, $\bar{A}$ is an upper triangular matrix. The complexity of such decomposition is $C_s(n/2)$ according to the lemma 1 and the remark.

(3). Let $z_1 = Wt_1 + w$ be a solution of the system $\bar{A}z_1 = g$, $W \in \mathbf{R}^{n/2 \times n/2}$, $w, t_1 \in \mathbf{R}^{n/2}$, t_1 is a vector of parameters.

Then

$$\begin{pmatrix} z_1 \\ z_0 \end{pmatrix} = \operatorname{diag}(W, I) \begin{pmatrix} t_1 \\ t_0 \end{pmatrix} + \begin{pmatrix} w \\ 0 \end{pmatrix}$$

is the solution of the system

$$(\bar{A}, 0) \begin{pmatrix} z_1 \\ z_0 \end{pmatrix} = g.$$

Let $P = \begin{pmatrix} P_a & P_b \\ P_c & P_d \end{pmatrix}$ be the block form of the matrix P. Since $\begin{pmatrix} y_1 \\ s_0 \end{pmatrix} = P \begin{pmatrix} z_1 \\ z_0 \end{pmatrix}$ and $y_0 = Vs_0 + v$, the solution of the initial system $\mathcal{A}y = c$ is obtained in the form

$$y = \begin{pmatrix} P_a W & P_b \\ VP_c W & VP_d \end{pmatrix} \begin{pmatrix} t_1 \\ t_0 \end{pmatrix} + \begin{pmatrix} P_a w \\ VP_c w + v \end{pmatrix}.$$

In the whole we need five block multiplications. Denote by C_n the complexity of solving a system with square coefficient matrix of the order n. Then

$$C_n = 2C_{n/2} + 5M(n/2) + C_s(n/2) = 2C_{n/2} + \alpha(5 + \lambda_\beta)n^{n/2} + (n/2)^2 C_D.$$

Hence we obtain the asymptotic evaluation

$$C_n = C_D(n^2 - n)/2 + \alpha\phi_\beta n^\beta, \quad \text{where} \quad \phi_\beta = \frac{1}{2^\beta - 2}(5 + \frac{7}{2^{\beta-2} - 1}).$$

Theorem 3. *In a quasiEuclidean ring any system of linear equations $\mathcal{A}x = c$, $\mathcal{A} \in \mathbf{R}^{n \times m}$ may be solved with an asymptotic estimation of the complexity C_{nm}:*

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

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

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

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

Proof. The decomposition $(\mathcal{A}, c) = P(\mathcal{A}', c')$ with the upper triangular matrix $(\mathcal{A}', c')$ may be obtained with the complexity $C_t^*(n,m)$ according to the corollary 2.

Therefore it is enough to solve the system

$$\mathcal{A}'x = c' \tag{5}$$

with upper triangular coefficient matrix.

a. Let $m > n$, $\mathcal{A}' = (A, B)$, A be a square block with an upper triangular form.

According to the corollary 1 and the remark we can obtain the unimodular matrix Q such that $(A,B)Q = (\bar{A}, 0)$, with the complexity $C_s^*(n,m)$. Here $\bar{A}$ is an upper triangular matrix.

The solutions of the system (5) and of the system $\bar{A}y = c'$ are connected according to the equation: $x = Q\begin{pmatrix} y \\ t_0 \end{pmatrix}$, where t_0 is a vector of parameters.

Let $Ut_1 + u$ be a solution of the system $\bar{A}y = c'$, $U \in \mathbf{R}^{n\times n}$, $t_1, u, y \in \mathbf{R}^n$, t_1 be a vector of parameters, $t_1 \in \mathbf{R}^n$.

Then the solution of initial system will be $x = Q\mathrm{diag}(U, I)t + Q\begin{pmatrix} u \\ 0 \end{pmatrix}$, where $t \in \mathbf{R}^m$ is a vector of parameters. The computing of matrix product $Q\mathrm{diag}(U, I)$ needs $\alpha n^{\beta-1}m$ operations. According to the theorem 2 the solution of the system $\bar{A}y = c'$ may be obtained with the complexity C_n.

In the whole we obtain the evaluation of the complexity of the system $\mathcal{A}x = c$ solving

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

$i = 1, 2$, where $i = 1$ in the case $m \geq 2n$ and $i = 2$ in the case $2n \geq m > n$.

b. Let $n \geq m$, $(\mathcal{A}', c') = \begin{pmatrix} A & c_A \\ 0 & c_B \end{pmatrix}$, A be a square block with upper triangular form. The system is incompatible if the vector c_B is non-zero.

According to the theorem 2 we can obtain the solution of the system $Ax = c_A$ with the complexity C_m.

So the total complexity is

$$C_{nm} = C_t^i(n,m) + C_m \quad i = 1, 2, \tag{6}$$

where $i = 1$ in the case $2m \geq n \geq m$ and $i = 2$ in the case $n \geq 2m$.

Therefore the theorem assertion follows.

In particular for the system with square coefficient matrix we obtain

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

We have $\psi_3 < 4,67$, $\psi_{log7} < 6,54$, $\psi_{log6} = 10,5$.

3 Hensel lift

Let $\mathbf{R}$ be a commutative domain with identity, $\mathbf{K}$ be a quotient field of $\mathbf{R}$, $A \in \mathbf{R}^{n\times m}$, $c \in \mathbf{R}^n$. Let

$$Ax = c \tag{7}$$

be a system of linear equations in $\mathbf{R}$.

Suppose, that $n < m$, $\text{rank}A = n$, $k = m - n + 1$ is a number of the linearly independent solutions of the system (7). Let matrix $A = (A_0, A_1)$ have a square block A_0, $\det A_0 \neq 0$, and the last row of a the matrix $A_0^{-1}(A_1, c)$ not contain nulls. The system in general may be reduced to such case (it may be done for example as in [10]).

Let $\mathbf{e}_j$ be the j column of the unit matrix $I_k = (\mathbf{e}_1, \ldots, \mathbf{e}_k)$ of the order k. We seek the solution of the system (7) in the form $\begin{pmatrix} x_j \\ \xi_j \mathbf{e}_j \end{pmatrix}$, $x_j \in \mathbf{K}^{n-1}$, $\xi_j \in \mathbf{K}\setminus 0$, $(j = 1, \ldots, k)$. Then we obtain a determined system of the order n.

It is known, that such system may be solved by p-adic lift [4]. If R is a ring of integers, the solution may be obtained with

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

bits operations. If $\mathbf{R}$ is a ring of polynomials over a field F, then solving the system needs

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

operations in the field F [11], [4].

To solve k such systems we need

$$k\mathcal{D}_Z, \quad \text{and} \quad k\mathcal{D}_{F[x]} \tag{8}$$

operations correspondingly.

As $\xi_j \neq 0$, $j = 1, \ldots, k$, the solutions of these k systems are linearly independent. Reduce the components of each solution to their common denominator (it is a divider of some minor of the matrix A), and write the solutions in the following way: $x_j = \begin{pmatrix} \mathbf{x}_j \\ \alpha_j \mathbf{e}_j \end{pmatrix} \chi_j^{-1}$, $j = 1, \ldots, k$, $\mathbf{x}_j \in \mathbf{R}^{n-1}$, $\alpha_j, \chi_j \in \mathbf{R}\setminus 0$.

The **probabilistic** approach to finding Diophantine solutions consists in the following.

Denote $\begin{pmatrix} \mathbf{x}_j \\ \alpha_j \mathbf{e}_j \end{pmatrix}$ by $\mathbf{w}_j$. It is easy to show that if $\{\mathbf{w}_j {\chi_j}^{-1} | i = 1, 2, \ldots, k\}$ is a basis set of solutions of the system (7), then

$$\left\{ \frac{\sum_{j=1}^k \mathbf{w}_j \eta_j}{\sum_{j=1}^k \chi_j \eta_j} \middle| \ \eta_j \in \mathbf{R}, \sum_{j=1}^k \chi_j \eta_j \neq 0 \right\}$$

is the set of all solutions of this system [10].

So the system has Diophantine solutions if the ideal generated by the denominators of a basis set of solutions is unit. In this case we can find such

$\eta_j \in \mathbf{R}$ that $\sum_{j=1}^{k} \chi_j \eta_j = 1$ and $\sum_{j=1}^{k} \mathbf{w}_j \eta_j$ is a Diophantine solution. Moreover if $x_i = \mathbf{w}_i \chi_i^{-1}$, $i = 1, 2, \ldots, h$ is a basis set of solutions and $\chi_1 = 1$, then the set $\mathbf{w}_1$, $\mathbf{w}_i - \mathbf{w}_1(\chi_i - 1)$, $i = 2, 3, \ldots, h$, is an integer basis [10].

If the ideal generated by the denominators of a basis set of solutions is not unit then we can rearrange columns in the matrix A and repeat the whole computation process. It might be necessary to repeat such procedure several times.

An expected number of such computations for the ring $\mathbf{Z}$ is $O(k^{-1}(\log n + \log\log \|(A, c)\|))$ [10], [11].

The **deterministic** approach to finding Diophantine solutions consists in the following.

The number $\mathrm{LCM}(\alpha_1, \ldots, \alpha_k)$ is a divider of the minor of a matrix (A, c), constructed from the first $n - 1$ columns and the last column, according to Crammer's rule. Denote $\delta = \mathrm{LCM}(\alpha_1, \ldots, \alpha_k)$.

Multiply the numerator and the denominator of the solution x_j by $\delta \alpha_j^{-1}$ and denote $\mathbf{x}_j \delta \alpha_j^{-1} = \mathbf{v}_j$, $\chi_j \delta \alpha_j^{-1} = \delta_j$ for each j. We obtain the basic set of the solutions of the system (7) in the form $x_j = \begin{pmatrix} \mathbf{v}_j \\ \delta \mathbf{e}_j \end{pmatrix} \delta_j^{-1}$.

Denote the vector $(\delta_1, \ldots, \delta_k)$ by $\mathbf{u}$, $\mathbf{u} \in \mathbf{R}^k$, and the matrix $(\mathbf{v}_1, \ldots, \mathbf{v}_k)$ by V, $V \in \mathbf{R}^{(n-1)\times k}$.

Theorem 4. *The system (7) and*

$$\begin{cases} Vy = 0 \bmod \delta \\ \mathbf{u}y = \delta \end{cases} \tag{9}$$

are consistent simultaneously in the ring $\mathbf{R}$.

If Y is the set of the solutions of the system (9), then
$X = \left\{ \begin{pmatrix} y \\ Vy\delta^{-1} \end{pmatrix} \mid y \in Y \right\}$ *is the set of the solutions of the system (7).*

Proof. The set of the solutions of the linear system (7) is a plane of the dimension $m - n$ in the space $\mathbf{K}^m$. It is defined by $m - n + 1$ linearly independent solutions. Therefore, for any solution $z \in \mathbf{K}^m$ of the system (7) the numbers $\sigma_i \in \mathbf{K}$, $i = 1, \ldots, k$, exist, such that

$$\sum_{i=1}^{k} x_i \sigma_i = z \quad \text{and} \quad \sum_{i=1}^{k} \sigma_i = 1.$$

Denote $\sigma_i \delta_i^{-1} = \tau_i \in \mathbf{K}$ and $z = (y, w)$, $y \in \mathbf{R}^k$, $w \in \mathbf{R}^{m-k}$, $t = (\tau_1, \ldots, \tau_k)$.

Then

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

and these equations may be written as the system:

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

Multiplying each part of this system by the matrix

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

invertible in $\mathbf{K}$, we obtain a system

$$\begin{pmatrix} \delta I_k & -I_k & 0 \\ 0 & V & -\delta I_{n-1} \\ 0 & \mathbf{u} & 0 \end{pmatrix} \begin{pmatrix} t \\ y \\ w \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \delta \end{pmatrix}. \tag{10}$$

Let the system (7) have the Diophantine solution $\mathbf{z} = (\mathbf{y}, \mathbf{w}) \in \mathbf{R}^m$. Then the system (10) is consistent. Therefore the system (9) is also consistent.

To the contrary, if the system (9) is consistent and $\mathbf{y} \in \mathbf{R}^k$ is its solution, then $(\delta^{-1}\mathbf{y}, \mathbf{y}, \delta^{-1}V\mathbf{y})$ is the solution of a system (10). Therefore $\mathbf{z} = (\mathbf{y}, \delta^{-1}V\mathbf{y})$ is the Diophantine solution of the system (7).

Now we estimate the complexity of the solution of the system (9) in the ring of integers and in the ring of polynomials. The matrix V has the size $(n-1) \times k$, and the module δ is a divider of a minor of the order n. Therefore, for $\mathbf{R} = \mathbf{Z}$

$$\log \delta \leq n(\log \|A\| + 0.5 \log n), \tag{11}$$

where $\|A\|$ is an absolute value of the greatest element of the matrix A. For $\mathbf{R} = F[z]$

$$\deg \delta \leq n \deg \|A\|, \tag{12}$$

where $\deg \|A\|$ is the greatest degree of elements of the matrix A.

Let us use the standard algorithm of multiplying integers and polynomials. Then the multiplying of N bit numbers requires N^2 bit operations, the multiplying of two polynomials of a degree N requires N^2 multiplications in the field F.

By Theorem 3 we obtain the following asymptotic estimations of operations for $\mathbf{Z}$ and $F[x]$, correspondingly:

$$\eta_\beta knM(n)(\log \|A\| + 0.5 \log n)^2, \quad \eta_\beta knM(n) \deg^2 \|A\|, \tag{13}$$

for $k \geq n$, and for $k < n$ we obtain

$$\delta k^{\beta-1}(\log \|A\| + (\log n)/2)^2(n^3(4 + \lambda_\beta) + n^2k(\nu_\beta + \mu_\beta - \lambda_\beta - 4)), \tag{14}$$

$$\delta k^{\beta-1} \deg^2 \|A\|(n^3(4 + \lambda_\beta) + n^2k(\nu_\beta + \mu_\beta - \lambda_\beta - 4)). \tag{15}$$

In general, these estimations do not depend on the complexity (8) of k solutions of the system (7).

However, for quasi-square systems it is not so, the contribution of the value of (8) is essential. In this case for $\mathbf{Z}$ and $F[x]$ we have the following estimations

$$O(n^3(\log n + \log \|A\| + \log p)^2), \quad O(n^3(\deg \|A\| + \deg p)^2). \tag{16}$$

Fast methods of multiplying integers and polynomials correspondingly improve these estimations. For example, the Karasuba method demands $O(N^{\log 3})$ operations and the FFT method demands $O(N \log N \log\log N)$ operations.

4 Modular Method

Let us solve the system (7), supposing that $A = (A_0, A_1)$, block A_0 is a square block, $\det A_0 = \delta \neq 0$. A recursive method [9] allows to transform the system (7) to the diagonal form

$$(\delta I_n, A'_1)x = c' \tag{17}$$

with $O(n^{\beta-1}m)$ operations in the ring $\mathbf{R}$. The methods described in [7] and [8] can do this transformation with $O(n^2 m)$ operations.

All coefficients of the system (17) are minors of the order n of matrix (A, c). Therefore estimations (11) and (12) are suitable to the ring of integers and ring of polynomials, correspondingly.

Calculation needs

$$O(mn^{\beta+1}(\log n + \log \|A\|)^2) \quad \text{and} \quad O(mn^{\beta+1} \deg^2 \|A\|),$$

operations in the rings $\mathbf{Z}$ and $F[x]$, correspondingly.

The Chinese remainder theorem allows to reduce amount of operations up to

$$O(mn^{\beta}(\log n + \log \|A\|)^2) \text{ and } O(mn^{\beta} \deg^2 \ |A\ |). \tag{18}$$

The solving of the system (17) is reduced to the solving of a system

$$A'_1 y = c' \bmod \delta. \tag{19}$$

in the quotient-rings $\mathbf{R}/\delta\mathbf{R}$.

Let Y be a set of the solutions of the system (19), then

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

is the set of the solutions of the system (17).

Let us estimate the complexity of the solution of the system (19). The matrix A' has the size $n \times (m-n)$, and the module δ is a minor of the order n.

By Theorem 3 we obtain the following estimations for $\mathbf{Z}$ and $F[x]$, correspondingly:

$$O(n^{\beta+1}(m-n)(\log \|A\| + \log n)^2), \quad O(n^{\beta+1}(m-n) \deg^2 \|A\|). \tag{20}$$

for $m - n \geq n$, and for $m - n \leq n$

$$O(n^2 m(m-n)^{\beta-1}(\log \|A\| + \log n)^2), \quad O(n^2 m(m-n)^{\beta-1} \deg^2 \|A\|). \tag{21}$$

In general (20) and (21) are the estimations for the modular method. However, for quasi-square systems the value $m-n$ can be neglected and the estimation (18) will give the basic contribution. In this case we obtain the following estimations for the number of operations:

$$O(n^{\beta+1}(\log n + \log \|A\|)^2) \quad \text{and} \quad O(n^{\beta+1} \deg^2 \|A\|) \tag{22}$$

for $\mathbf{Z}$ and $F[x]$, correspondingly.

5 Resume

The evaluations (13)–(16) and (20)–(22) obtained in the last two sections allow to formulate the next theorem about the complexity (number of bit operations) of solving a linear Diophantine system of the size $n \times m$.

Theorem 5. *For essentially rectangular systems, when $m > 2n$, both methods have the identical complexity $O(mn^{\beta+1})$. For quasi-square systems the complexity of the p-adic method is $O(n^3)$ and the complexity of the modular method is $O(n^{\beta+1})$.*

All estimations in the theorem are given within a logarithmic factors. In the case of the polynomial ring $F[x]$ the analogous theorem is true for the evaluation of the number of operations in the field F.

The comparison of the p-adic and the modular methods shows, that for quasi-square systems the p-adic method has the advantage, it is $O(n^{\beta-2})$ times as fast.

The other advantage of the p-adic method for systems of the arbitrary size is the possibility to use the probability approach. The general solution of a system can be obtained with a high probability with $O((m-n+1)n^3)$ operations [10]. An expected amount of attempts will be $O((\log n + \log\log \|A\|)/(m-n+1))$.

These estimations can still be improved, if the fast algorithms of multiplying would be used.

References

1. Abbott, J., Bronstein, M. and Mulders, T.: Fast Deterministic Computation of Determinants of Dense Matrices. In: Proceedings of ISSAC'99. ACM Press, Vancouver (1999) 197–204
2. Bareiss, E. H.: Computational solutions of matrix problems over an integral domain. J. Inst. Maths Applics **10** (1972) 68–104
3. Coppersmith, D. and Winograd, S.: Matrix multiplication via arithmetic progressions. Journal of Symbolic Computation **9** (1990) 251–280
4. Dixon, J.: Exact solution of linear equations using p-adic expansions. Numer. Math. **40** (1982) 137-141
5. Gregory, R. T., Krishnamurthy E. V.: Methods and Applications of Error-Free Computation. Springer-Verlag, Berlin (1984)
6. Hurt, M. F., Waid, C. A.: A generalized inverse which give all the integral solutions to a system of linear equations. SIAM J. Appl. Math. **10**, (1970) 547–550

7. Malaschonok, G. I.: Solution of a system of linear equations in an integral domain. USSR Journal of Computational Mathematics and Mathematical Physics **23** (1983) 1497–1500
8. Malaschonok, G. I.: Algorithms for the solution of systems of linear equations in commutative rings. In: Mora, T. and Traverso, C. (eds.): Effective Methods in Algebraic Geometry. Progress in Mathematics 94, Birkhauser (1991) 289–298
9. Malaschonok, G. I.: Recursive Method for the Solution of Systems of Linear Equations. In: Sydow, A. (ed.): Computational Mathematics (Proceedings of the 15th IMACS World Congress, Vol. I, Berlin, August 1997). Wissenschaft & Technik Verlag, Berlin (1997) 475–480
10. Malaschonok, G. I.: Effective Matrix Methods in Commutative Domains, In: Krob, D., Mikhalev, A. A., Mikhalev, A. V. (eds.): Formal Power Series and Algebraic Combinatorics. Springer (2000) 506–517
11. Mulders, T., Storjohann, A.: Diophantine Linear System Solving. In: Proceedings of ISSAC'99. ACM Press, Vancouver (1999) 181–188