Balanced Solutions of Linear Diophantine Equations
ACAC 2012

K.A. Draziotis

27-28 August, 2012
ACAC
Athens

Solutions of linear diophantine equations

- Let $a_j \in \mathbb{Z} - \{0\},\ (0 \leq j \leq n)$. We consider the linear diophantine equation

$$f(x_1, \dots, x_n) = \sum_{j=1}^{n} a_j x_j = a_0, \text{ with } |x_j| \leq X_j. \tag{1}$$

Solutions of linear diophantine equations

- Let $a_j \in \mathbb{Z} - \{0\}, \ (0 \leq j \leq n)$. We consider the linear diophantine equation

$$f(x_1, ..., x_n) = \sum_{j=1}^{n} a_j x_j = a_0, \text{ with } |x_j| \leq X_j. \qquad (1)$$

- Without the bound constraints can be solved in polynomial time. For instance see

  •*H.Esmaeili, Lecturas Matemáticas Volumen 27 (2006).*

  • *James Bond. Calculating the general solution of a linear Diophantine equation. American Mathematical Monthly, Vol. 74, p. 955-957, 1967.*

  • *Stanley Kertzner. The linear Diophantine equation. American Mathematical Monthly, Vol. 88, p.200-203, 1981.*

Some Applications.

- Several problems are related with the integer solutions of a linear equation, under the previous bound constraints. Assume that the coefficients a_i $(1 \leq i \leq n)$ are positive and $a_0 = \gcd(a_1, ..., a_n)$. Then the problem of finding some multipliers x_i for the gcd, is called extended gcd problem.

Some Applications.

- Several problems are related with the integer solutions of a linear equation, under the previous bound constraints. Assume that the coefficients a_i $(1 \leq i \leq n)$ are positive and $a_0 = \gcd(a_1, ..., a_n)$. Then the problem of finding some multipliers x_i for the gcd, is called extended gcd problem.
- The decidability problem of the existence of a solution of Egcd under the bound $|x_j| < X_j$ is proved to be NP-complete.

Some Applications.

- Several problems are related with the integer solutions of a linear equation, under the previous bound constraints. Assume that the coefficients a_i $(1 \leq i \leq n)$ are positive and $a_0 = \gcd(a_1, ..., a_n)$. Then the problem of finding some multipliers x_i for the gcd, is called extended gcd problem.
- The decidability problem of the existence of a solution of Egcd under the bound $|x_j| < X_j$ is proved to be NP-complete.
- We shall study this problem and we develop an algorithm which give us a solution for the Egcd problem (under some assumptions)

Some Applications.

- If we restrict the solutions $x_j \in \{0, 1\}$, then we have the $0 - 1$ Knapsack or subset sum problem. Also, if $x_j \in \mathbb{N},\ a_0 = 0$ then the problem of deciding if there is any integer solution is NP-complete.

Some Applications.

- The Frobenius problem seeks for the largest integer M such that $f(x_1, \ldots, x_n) = M$ fails to have a solution.

Some Applications.

- The Frobenius problem seeks for the largest integer M such that $f(x_1, ..., x_n) = M$ fails to have a solution.
- This problem is in the general case NP-hard and is solved, in polynomial time, for $n = 3$

LLL reduction algorithm.

- In 1982, A.Lenstra, H.Lenstra and Lovasz published in their landmark paper the LLL-algorithm, which is a basis reduction algorithm for lattices, based on Hermite's inequality. The aim of LLL algorithm is to find a short vector that approximates the shortest nonzero vector of the lattice. The LLL algorithm runs in polynomial time as a function of the rank of the lattice.

LLL reduction algorithm.

- Let a basis $\{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ of a lattice $L = \mathbf{Z}\mathbf{b}_1 + \cdots + \mathbf{Z}\mathbf{b}_n \subset \mathbf{R}^k$. We associate the Gramm-Schmidt orthogonalization vectors $\{\mathbf{g}_1, \ldots, \mathbf{g}_n\}$ defined by the relations

$$\mathbf{g}_1 = \mathbf{b}_1, \ \mathbf{g}_i = \mathbf{b}_i - \sum_{j=1}^{i-1} \mu_{ij}\mathbf{b}_j,$$

  where the GSO coefficients

$$\mu_{ij} = \frac{\mathbf{b}_i \cdot \mathbf{g}_j}{B_j^2}, \quad B_j = ||\mathbf{g}_j||.$$

  The GSO process produce vectors that form an orthogonal basis of $\mathbf{R}^k$.

LLL reduction algorithm.

- An LLL-basis has two characteristics.
  i. Is size reduced, that is $|\mu_{ij}| < 1/2,\ 1 \leq j < i < n$ and
  ii. The vectors of the basis are almost orthogonal to each other. This, is translated to the following (Lovasz) relation $\delta||\mathbf{g}_i||^2 \leq ||\mathbf{g}_{i+1} + \mu_{i+1,i}\mathbf{g}_i||^2$ for some $\delta \in (1/4, 1)$.

LLL reduction algorithm.

- An LLL-basis has two characteristics.
  i. Is size reduced, that is $|\mu_{ij}| < 1/2,\ 1 \leq j < i < n$ and
  ii. The vectors of the basis are almost orthogonal to each other. This, is translated to the following (Lovasz) relation $\delta||\mathbf{g}_i||^2 \leq ||\mathbf{g}_{i+1} + \mu_{i+1,i}\mathbf{g}_i||^2$ for some $\delta \in (1/4, 1)$.
- The LLL-algorithm achieve to give us a small length vector, more specific

$$||\mathbf{b}_1|| \leq 2^{n-1} d(L)^{1/n},$$

where $d(L)$ is the discriminant of the latttice.

LLL reduction algorithm.

- An LLL-basis has two characteristics.
  i. Is size reduced, that is $|\mu_{ij}| < 1/2,\ 1 \leq j < i < n$ and
  ii. The vectors of the basis are almost orthogonal to each other. This, is translated to the following (Lovasz) relation $\delta||\mathbf{g}_i||^2 \leq ||\mathbf{g}_{i+1} + \mu_{i+1,i}\mathbf{g}_i||^2$ for some $\delta \in (1/4, 1)$.
- The LLL-algorithm achieve to give us a small length vector, more specific

$$||\mathbf{b}_1|| \leq 2^{n-1} d(L)^{1/n},$$

  where $d(L)$ is the discriminant of the latttice.
- Some apllications are
  • to the factorization of polynomials over $\mathbf{Z}[x]$
  • to find integer relations between some real numbers $k_1, ..., k_n$.

Solutions of linear diophantine equations.

- We set up some notation.

Solutions of linear diophantine equations.

- We set up some notation.
- Let $X_j,\ j = 1, 2, ..., n$ be positive integers and B be the lattice generated by the following vectors.

$$\{\mathbf{b}_j : \mathbf{b}_j = (0, ..., \frac{1}{X_j}, 0, ..., 0, a_j) \in \mathbf{R}^{n+2}, j = 1, 2, ..., n\}$$

Solutions of linear diophantine equations.

- We set up some notation.
- Let $X_j,\ j = 1, 2, ..., n$ be positive integers and B be the lattice generated by the following vectors.

$$\{\mathbf{b}_j : \mathbf{b}_j = (0, ..., \frac{1}{X_j}, 0, ..., 0, a_j) \in \mathbf{R}^{n+2}, j = 1, 2, ..., n\}$$

- We apply LLL and we get the following vectors

$$B' = \{\mathbf{b}_1^{'}, \mathbf{b}_2^{'}, ..., \mathbf{b}_n^{'}\}.$$

Solutions of linear diophantine equations.

- We set up some notation.
- Let $X_j,\ j = 1, 2, ..., n$ be positive integers and B be the lattice generated by the following vectors.

$$\{\mathbf{b}_j : \mathbf{b}_j = (0, ..., \frac{1}{X_j}, 0, ..., 0, a_j) \in \mathbf{R}^{n+2}, j = 1, 2, ..., n\}$$

- We apply LLL and we get the following vectors

$$B' = \{\mathbf{b}_1^{'}, \mathbf{b}_2^{'}, ..., \mathbf{b}_n^{'}\}.$$

- Finally, using the Gramm-Schimdt orthogonalization process to the LLL reduced basis B', we get the set $G = \{\mathbf{g}_1, \mathbf{g}_2, ..., \mathbf{g}_n\}$. We define $B_j = ||\mathbf{g}_j||$.

Solutions of linear diophantine equations.

- We shall prove the following Theorem.
  Let $\gcd(a_1, \dots, a_n) = 1$. If the following two assumptions hold
  A_1. $(a_n X_n)^2 + (a_j X_j)^2 < \dfrac{1}{2^{n+1}} (X_n X_j)^2, \; j = 1, 2, \dots, n-1,$
  A_2. $\left\lceil \dfrac{a_0}{B_n^2} \right\rfloor = a_0,$
  then we can find in polynomial time, an integer solution $(x_1, \dots, x_n)$ of the equation $\sum_{j=1}^{n} a_j x_j = a_0$, such that $|x_j| < c(n) X_j \prod_{i=1}^{n} X_i, \;\; j = 1, 2, \dots, n, \; c(n) = \sqrt{3}(1.25)^{(n-1)/2}$

Solutions of linear diophantine equations.

- We shall prove the following Theorem.
  Let $\gcd(a_1, \ldots, a_n) = 1$. If the following two assumptions hold
  A_1. $(a_n X_n)^2 + (a_j X_j)^2 < \dfrac{1}{2^{n+1}} (X_n X_j)^2, \; j = 1, 2, \ldots, n-1,$
  A_2. $\left\lceil \dfrac{a_0}{B_n^2} \right\rfloor = a_0,$
  then we can find in polynomial time, an integer solution $(x_1, \ldots, x_n)$ of the equation $\sum_{j=1}^{n} a_j x_j = a_0$, such that $|x_j| < c(n) X_j \prod_{i=1}^{n} X_i, \;\; j = 1, 2, \ldots, n, \; c(n) = \sqrt{3}(1.25)^{(n-1)/2}$
- Also, the proof of the Theorem provide us with an algorithm.

The Algorithm

- There are two basic steps, first LLL to the lattice B, and second size reduction to a new lattice of the form $M = B + \mathbf{Zb}$.

The Algorithm

- There are two basic steps, first LLL to the lattice B, and second size reduction to a new lattice of the form $M = B + \mathbf{Zb}$.
- Say that we have the linear diophantine equation

$$a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 + a_5x_5 = a_0.$$

The Algorithm

- There are two basic steps, first LLL to the lattice B, and second size reduction to a new lattice of the form $M = B + \mathbf{Zb}$.
- Say that we have the linear diophantine equation

$$a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 + a_5x_5 = a_0.$$

- We apply LLL to the rows of the Lattice given by the matrix

$$M = \begin{bmatrix} \frac{1}{X_1} & 0 & 0 & 0 & 0 & 0 & a_1 \\ 0 & \frac{1}{X_2} & 0 & 0 & 0 & 0 & a_2 \\ 0 & 0 & \frac{1}{X_3} & 0 & 0 & 0 & a_3 \\ 0 & 0 & 0 & \frac{1}{X_4} & 0 & 0 & a_4 \\ 0 & 0 & 0 & 0 & \frac{1}{X_5} & 0 & a_5 \end{bmatrix}$$

The Algorithm

- Then to the new reduced basis we add the row $(0, 0, 0, 0, 0, 1, -a_0)$.

The Algorithm

- Then to the new reduced basis we add the row $(0,0,0,0,0,1,-a_0)$.
- Then apply size reduction to the rows, that is (here $n=5$)

$$row(n+1) \leftarrow row(n+1) - \lceil \mu_{n+1,j} \rfloor row(j)$$

$$\mu_{ij} = \frac{\mathbf{b}_i^{'} \cdot \mathbf{g}_j}{B_j^2}, \quad B_j = ||\mathbf{g}_j||.$$

first for $j = n$ and then $j = 1, 2, ..., n-1$.

The Algorithm

- Finally, we multiply the $i-$ entry of the last vector with X_i.

The Algorithm

- Finally, we multiply the $i-$ entry of the last vector with X_i.
- The resulting vector gives a solution of the diophantine equation which satisfies the bound of our theorem.

The Algorithm

- Finally, we multiply the $i-$ entry of the last vector with X_i.
- The resulting vector gives a solution of the diophantine equation which satisfies the bound of our theorem.
- But, at least experimentally satisfies the better bound $|x_j| < X_j$.

Examples

- Let

$$84 \cdot 10^5 x_1 + 4 \cdot 10^6 x_2 + 15688 x_3 + 6720 x_4 + 15 x_5 = 371065262.$$

This is example 1 of Aardal, Hurkens, Lenstra

Ref : Aardal, K.; Hurkens, C.; Lenstra, A..; Solving a linear Diophantine equation with lower and upper bounds on the variables. Integer programming and combinatorial optimization, LNCS, **1412**.

and they get the solution $\mathbf{x} = (36, 17, 39, 8, -22)$, with $||\mathbf{x}|| \simeq 60.44$. Assumption A_1 is fulfilled if

$$\max_{1\leq j\leq 4} |a_j| < \frac{1}{8} X_5, \ |a_5| < \frac{1}{8} \max_{1\leq j\leq 5} X_j.$$

So it is enough to choose

$$X = X_1 = \cdots = X_5 = 8 \cdot 84 \cdot 10^5 + 1.$$

Examples

- We consider the matrix

$$M = \begin{bmatrix} \frac{1}{67200001} & 0 & 0 & 0 & 0 & 0 & 8400000 \\ 0 & \frac{1}{67200001} & 0 & 0 & 0 & 0 & 4000000 \\ 0 & 0 & \frac{1}{67200001} & 0 & 0 & 0 & 15688 \\ 0 & 0 & 0 & \frac{1}{67200001} & 0 & 0 & 6720 \\ 0 & 0 & 0 & 0 & \frac{1}{67200001} & 0 & 15 \end{bmatrix}$$

Examples

- Applying LLL to the rows of M we get
  $M_{LLL} =$

$$
\begin{bmatrix}
-\frac{10}{67200001} & \frac{21}{67200001} & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \frac{15}{67200001} & -\frac{35}{67200001} & -\frac{8}{67200001} & 0 & 0 \\
\frac{1}{67200001} & -\frac{2}{67200001} & -\frac{25}{67200001} & -\frac{1}{67200001} & -\frac{72}{67200001} & 0 & 0 \\
\frac{5}{67200001} & -\frac{10}{67200001} & -\frac{95}{67200001} & -\frac{76}{67200001} & \frac{72}{67200001} & 0 & 0 \\
\frac{2}{67200001} & -\frac{4}{67200001} & -\frac{42}{67200001} & -\frac{21}{67200001} & \frac{1}{67200001} & 0 & -1
\end{bmatrix}
$$

Examples

- We add a new row $\mathbf{b}_6 = (0, 0, 0, 0, 0, 1, -a_0)$.

$$\begin{bmatrix} -\frac{10}{67200001} & \frac{21}{67200001} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{15}{67200001} & -\frac{35}{67200001} & -\frac{8}{67200001} & 0 & 0 \\ \frac{1}{67200001} & -\frac{2}{67200001} & -\frac{25}{67200001} & -\frac{1}{67200001} & -\frac{72}{67200001} & 0 & 0 \\ \frac{5}{67200001} & -\frac{10}{67200001} & -\frac{95}{67200001} & -\frac{76}{67200001} & \frac{72}{67200001} & 0 & 0 \\ \frac{2}{67200001} & -\frac{4}{67200001} & -\frac{42}{67200001} & -\frac{21}{67200001} & \frac{1}{67200001} & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 & -a_0 \end{bmatrix}$$

Examples

- Then applying size reduction to the previous lattice i.e.

$$row(6) \leftarrow row(6) - \lceil \mu_{6,j} \rfloor row(j),$$

first for $j = 6$ and then for $j = 1, 2, 3, 4, 5$, we shall get

Examples

- $\hat{M}_{LLL} =$

$$\begin{bmatrix} -\frac{10}{67200001} & \frac{21}{67200001} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{15}{67200001} & -\frac{35}{67200001} & -\frac{8}{67200001} & 0 & 0 \\ \frac{1}{67200001} & -\frac{2}{67200001} & -\frac{25}{67200001} & -\frac{1}{67200001} & -\frac{72}{67200001} & 0 & 0 \\ \frac{5}{67200001} & -\frac{10}{67200001} & -\frac{95}{67200001} & -\frac{76}{67200001} & \frac{72}{67200001} & 0 & 0 \\ \frac{2}{67200001} & -\frac{4}{67200001} & -\frac{42}{67200001} & -\frac{21}{67200001} & \frac{1}{67200001} & 0 & -1 \\ \frac{36}{67200001} & \frac{17}{67200001} & \frac{39}{67200001} & \frac{8}{67200001} & -\frac{2}{6109091} & 1 & 0 \end{bmatrix}$$

Examples

- $\hat{M}_{LLL} =$

$$\begin{bmatrix} -\frac{10}{67200001} & \frac{21}{67200001} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{15}{67200001} & -\frac{35}{67200001} & -\frac{8}{67200001} & 0 & 0 \\ \frac{1}{67200001} & -\frac{2}{67200001} & -\frac{25}{67200001} & -\frac{1}{67200001} & -\frac{72}{67200001} & 0 & 0 \\ \frac{5}{67200001} & -\frac{10}{67200001} & -\frac{95}{67200001} & -\frac{76}{67200001} & \frac{72}{67200001} & 0 & 0 \\ \frac{2}{67200001} & -\frac{4}{67200001} & -\frac{42}{67200001} & -\frac{21}{67200001} & \frac{1}{67200001} & 0 & -1 \\ \frac{36}{67200001} & \frac{17}{67200001} & \frac{39}{67200001} & \frac{8}{67200001} & -\frac{2}{6109091} & 1 & 0 \end{bmatrix}$$

- We take the 6th row and multiply each entry with X, then we shall get the vector $\mathbf{x} = (36, 17, 39, 8, -22)$. Which is the same as was found in Lenstra's paper.

Examples

- Now if we choose $X_1 = 30, X_2 = X_3 = X_4 = X_5 = 50$ and we repeat the previous procedure we shall get the solution $\mathbf{y} = (26, 38, 39, 8, -22)$. This solution has euclidean length $\simeq 64.72$. Note that is larger than the previous solution $\mathbf{x} = (36, 17, 39, 8, -22)$, with $||x|| \simeq 60.44$. It has the advantage that satisfy our constraints (the solution $\mathbf{x}$ does not).

Sketch of the Proof

- Let $\{\mathbf{b}_1^{'}, \mathbf{b}_2^{'}, ..., \mathbf{b}_n^{'}\}$ be the LLL reduced basis of the vectors

$$B = \{\mathbf{b}_j : \mathbf{b}_j = (0, ..., \frac{1}{X_j}, 0, ..., 0, a_j) \in \mathbf{R}^{n+2}, j = 1, 2, ..., n\}.$$

Sketch of the Proof

- Let $\{\mathbf{b}_1^{'}, \mathbf{b}_2^{'}, ..., \mathbf{b}_n^{'}\}$ be the LLL reduced basis of the vectors

$$B = \{\mathbf{b}_j : \mathbf{b}_j = (0, ..., \frac{1}{X_j}, 0, ..., 0, a_j) \in \mathbf{R}^{n+2}, j = 1, 2, ..., n\}.$$

- We apply the Gramm-Schimdt orthogonalization process to the LLL reduced basis and we get the set $\mathbf{g}_1, \mathbf{g}_2, ..., \mathbf{g}_n$. We set

$$\mu_{ij} = \frac{\mathbf{b}_i^{'} \cdot \mathbf{g}_j}{B_j^2}, \quad B_j = ||\mathbf{g}_j||.$$

Sketch of the Proof

- From relations

$$\mu_{n+1,j} = \frac{\mathbf{b}'_{n+1} \cdot \mathbf{g}_j}{B_j^2}, \quad B_j = ||\mathbf{g}_j||$$

we get the following linear system of $n-$equations with $n-$unknowns

$$\hat{x}_1 \hat{b}^*_{j1} + \cdots + \hat{x}_n \hat{b}^*_{jn} = \varepsilon_j, \ 1 \leq j \leq n,$$

where

$$\hat{b}^*_{ij} = \frac{b^*_{ij}}{X_j}$$

Sketch of the Proof

- From relations

$$\mu_{n+1,j} = \frac{\mathbf{b}'_{n+1} \cdot \mathbf{g}_j}{B_j^2}, \quad B_j = ||\mathbf{g}_j||$$

we get the following linear system of $n-$equations with $n-$unknowns

$$\hat{x}_1 \hat{b}^*_{j1} + \cdots + \hat{x}_n \hat{b}^*_{jn} = \varepsilon_j, \; 1 \leq j \leq n,$$

where

$$\hat{b}^*_{ij} = \frac{b^*_{ij}}{X_j}$$

- From assumptions A_1 and A_2 we can prove that

$$|\varepsilon_j| < \frac{1}{2}(1 \leq j \leq n-1), \varepsilon_n < 1.$$

Sketch of the Proof

- Finally ,using Cramer rule and Hadamard inequality we can prove that

$$|x_j| < c(n)X_j \prod_{i=1}^{n} X_i, \quad j = 1, 2, ..., n.$$

Sketch of the Proof

- Finally ,using Cramer rule and Hadamard inequality we can prove that

$$|x_j| < c(n)X_j \prod_{i=1}^{n} X_i, \quad j = 1, 2, ..., n.$$

- Thank you!