THE PARALLEL APPROXIMATION PROBLEM, AND SUBSET SUMS

Gábor Pataki

Department of Statistics and Operations Research

UNC Chapel Hill

joint work with **Mustafa Tural**

Parallel Approximation: The Basic Problem

Given

$$\boldsymbol{a} \in \mathbb{Z}^n, \ \|\boldsymbol{a}\| \text{ large}$$

find

$$\boldsymbol{p} \in \mathbb{Z}^n, \ \|\boldsymbol{p}\| \text{ small, near parallel to } \boldsymbol{a}.$$

"Near parallel" means: $\exists \lambda \in \mathbb{Q}_{++}, \ r \in \mathbb{Q}^n :$

$$\begin{aligned} a &= \lambda p + r, \\ \lambda &> 0, \quad \text{and} \end{aligned}$$

Parallel Approximation: The Basic Problem

Given

$$\boldsymbol{a} \in \mathbb{Z}^n, \ \| \boldsymbol{a} \| \ \textbf{large}$$

find

$$\boldsymbol{p} \in \mathbb{Z}^n, \ \| \boldsymbol{p} \| \ \textbf{small, near parallel to} \ \boldsymbol{a}.$$

"Near parallel" means: $\exists \lambda \in \mathbb{Q}_{++}, \, r \in \mathbb{Q}^n :$

$$\begin{aligned} a &= \lambda p + r, \\ \lambda &> 0, \quad \text{and} \end{aligned}$$

$$\begin{aligned} \lambda &\to +\infty, \ \text{as} \ \| a \| \to +\infty \\ \| r \| / \lambda &\to 0, \qquad \text{as} \ \| a \| \to +\infty. \end{aligned}$$

Remark:

$$|\sin(a, p)| \le \| r \| / \lambda,$$

so the last condition is quite natural.

Parallel Approximation: The Basic Problem

Remark 2: If

$$\| r \| / \lambda < 1/2,$$

then

$$a = \lambda p + r \Rightarrow (1/\lambda)a = p + (1/\lambda)r \Rightarrow \boldsymbol{p} = \mathbf{round}((\mathbf{1}/\boldsymbol{\lambda})\boldsymbol{a})$$

Example 1 (p found using the U^{-1}-method presented later)

$$
\begin{aligned}
\boldsymbol{a} &= (\mathbf{7621, 4565, 185, 8214, 4447}) \\
\boldsymbol{p} &= (\mathbf{374, 224, 9, 403, 218})
\end{aligned}
$$

$$
\begin{aligned}
\lambda &= 20.3821 \\
r &= (-1.8884, -0.5802, 1.5615, 0.0321, 3.7121)
\end{aligned}
$$

Example 1 (p found using the U^{-1}-method presented later)

$$\boldsymbol{a} = (\mathbf{7621}, \mathbf{4565}, \mathbf{185}, \mathbf{8214}, \mathbf{4447})$$
$$\boldsymbol{p} = (\mathbf{374}, \mathbf{224}, \mathbf{9}, \mathbf{403}, \mathbf{218})$$

$$\lambda = 20.3821$$
$$r = (-1.8884, -0.5802, 1.5615, 0.0321, 3.7121)$$

This a was randomly generated.

Example 2

$$\begin{aligned} \boldsymbol{a} &= \mathbf{(6765, 10946, 17711, 28657, 46368)} \\ \boldsymbol{p} &= \mathbf{(34, 55, 89, 144, 233)} \end{aligned}$$

$$\begin{aligned} \lambda &= 199.0046 \\ r &= (-1.1553, 0.7488, -0.4066, 0.3422, -0.0644) \end{aligned}$$

Example 2

$$
\begin{aligned}
\boldsymbol{a} &= (\mathbf{6765, 10946, 17711, 28657, 46368}) \\
\boldsymbol{p} &= (\mathbf{34, 55, 89, 144, 233})
\end{aligned}
$$

$$
\begin{aligned}
\lambda &= 199.0046 \\
r &= (-1.1553, 0.7488, -0.4066, 0.3422, -0.0644)
\end{aligned}
$$

This a was made up of consecutive **Fibonacci-numbers**, and p comes out to be like that too.

How to find PA

(1) Reducing it to Diophantine Aproximation – less interesting.

(2) The U^{-1}-method – more interesting.

Background: Basis Reduction

Given integral matrix A, basis reduction (BR) computes a unimodular $U (\Leftrightarrow \det U = \pm 1)$ st. the columns of AU are "short" and "nearly" orthogonal.
Best known BR variant: LLL - Lenstra, Lenstra, Lovász; 1982.

Example

$$A = \begin{pmatrix} 289 & 18 \\ 466 & 29 \\ 273 & 17 \end{pmatrix}, \; U = \begin{pmatrix} 1 & -15 \\ -16 & 241 \end{pmatrix}, \; AU = \begin{pmatrix} 1 & 3 \\ 2 & -1 \\ 1 & 2 \end{pmatrix}.$$

Multiplying A by U $\Leftrightarrow$ doing *elementary column operations* on A:

- adding an integer multiple of a column to another; multiplying a column by -1; swapping columns.

Main result

Theorem 1. *Suppose* $a \in \mathbb{Z}^n$, $\| a \| \geq 2^{n(n+1)/4}$. *Let* U *be a unimodular matrix such that*

$$\begin{pmatrix} a \\ I \end{pmatrix} U$$

is LLL-reduced. Let

$$\boldsymbol{p := \pm \ \textbf{the last row of} \ U^{-1}},$$

λp *the projection of* a *onto the line spanned by* $p, r = a - \lambda p$.

Then p, r, λ *satisfy the PA requirements.*

□

In more detail

Theorem 2. *If* a, p, r, λ *are as above, and*

$$f(a) = \frac{2^{n/4}}{\| a \|^{1/n}}.$$

Then

(1) $\| p \| \sqrt{1 + \| r \|^2} \le f(a) \| a \|$.

(2) $\lambda \ge \dfrac{1}{f(a)}$.

(3) $\dfrac{\| r \|}{\lambda} \le 2f(a)$.

$\square$

Notation

We write

$$p = \mathbf{PA}(a)$$

to denote that "p was computed from a using the U^{-1}-method."

Intuition: why we compute p like this

Suppose we **know in advance** that $\boldsymbol{a}$ **decomposes** ...

Theorem (Krishnamoorthy, Pataki 2005) Suppose that

$$a = \lambda p + r$$

with

$$\lambda > 2^{(n+1)/2} \, \| p \| \, (\| r \| + 1)^2,$$

and U is unimodular, s.t.

$$\begin{pmatrix} a \\ I \end{pmatrix} U = \begin{pmatrix} \lambda p + r \\ I \end{pmatrix} U \quad \text{is LLL reduced.}$$

Then $pU = (\overbrace{0 \ \ldots \ 0}^{n-1} \ \alpha \ )$ for some $\alpha \in \mathcal{Z} \setminus \{ 0 \}$.

So it is reasonable to think (and Theorem 1 proves this) that the **last row of $\boldsymbol{U^{-1}}$ gives a good approximation for any $\boldsymbol{a}$ with large norm**.

How will we use PA in IP

Given polyhedron $Q \subseteq \mathbb{R}^n$, vector $p \in \mathbb{Z}^n$,

- $\text{iwidth}(p, Q) = \lfloor \max \{ px \mid x \in Q \} \rfloor - \lceil \min \{ px \mid x \in Q \} \rceil + 1.$
- branching on px means creating the branches

$$px = \lceil \min \rceil, px = \lceil \min \rceil + 1, \ldots, cx = \lfloor \max \rfloor$$

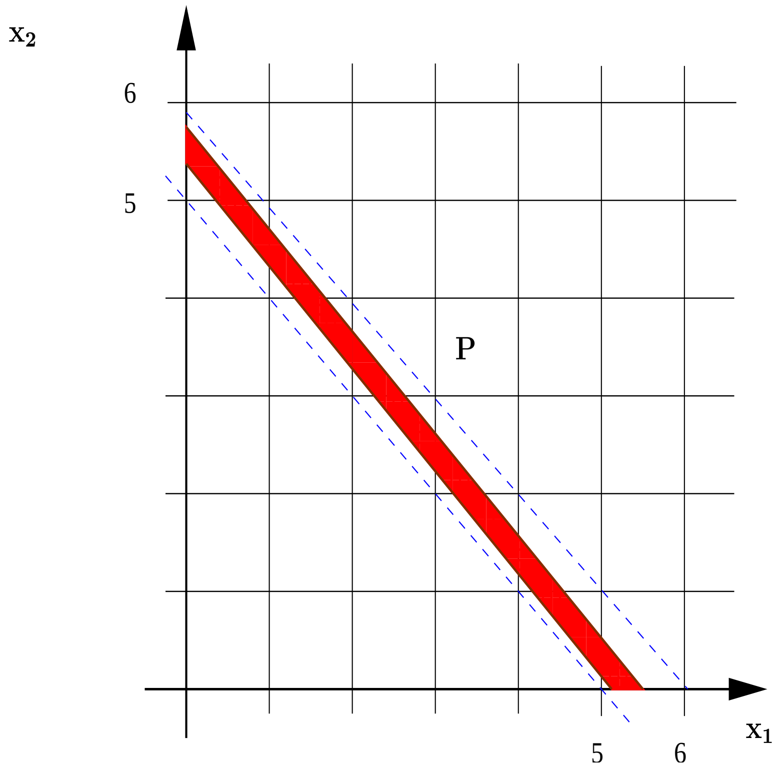
  (adding them to LP relaxation).

  No branches created $\longrightarrow Q$ has **no integral point**.
- If $p = e_i$, this is branching on a single variable x_i.
- If the only dense constraint is $\beta' \leq ax \leq \beta$, then **branching on px** with p near parallel to a may be **much better** than branching on any **single x_i**.

Example: an integer infeasible polyhedron

$$\begin{array}{rcl} 106 \leq & 21x_1 + 19x_2 & \leq 113 \\ 0 \leq & x_1, x_2 & \leq 6 \end{array}$$

Proving infeasibility is



- **Hard for x_is**;
- **Easy for $x_1 + x_2$**: max = 5.94, min = 5.04.

How will we use PA in IP

- In a knapsack problem with weight vector a, find

$$\boldsymbol{p = \mathrm{PA}(a),}$$

  and branch on px.

- In fact, we will use a natural reformulation, in which branching on the last single variable is equivalent to branching on px in the original.
- Motivation: Krishnamoorthy-Pataki (2005-2006): **analysis** of such problems with a having a **decomposition known in advance**.

Knapsack problem and its reformulation

$$\textbf{(KP)} \quad \begin{pmatrix} \beta' \\ 0 \end{pmatrix} \leq \begin{pmatrix} a \\ I \end{pmatrix} x \leq \begin{pmatrix} \beta \\ u \end{pmatrix} \longrightarrow \begin{pmatrix} \beta' \\ 0 \end{pmatrix} \leq \begin{pmatrix} a \\ I \end{pmatrix} U y \leq \begin{pmatrix} \beta \\ u \end{pmatrix} \quad \textbf{(KP-R)}$$

where U is a unimodular matrix, computed to make the columns of

$$\begin{pmatrix} a \\ I \end{pmatrix} U \text{ LLL-reduced.}$$

(KP-R) is called the *rangespace reformulation* of (KP).

Story of reformulations: Aardal-Hurkens-Lenstra '98; Aardal-Bixby-Hurkens-Lenstra-Smeltink '00; Louveaux-Wolsey '02; Aardal-Lenstra '04 for equality constrained problems; Krishnamoorthy-P. 2005-06: general framework (stay in same space, or not), and analysis.

Fact: $\mathbf{p} = \mathbf{PA}(\mathbf{a}) \Rightarrow$
branching on px in (KP) $\equiv$ branching on y_n in (KP-R).

Why? y feasible in (KP-R) $\Leftrightarrow x = Uy$ feasible in (KP). Also,

$$px = p(Uy) = (pU)y = y_n.$$

(see in detail Krishnamoorthy-P.)

First Testcase IP: the Subset Sum Problem

Given positive integers $\boldsymbol{a = (a_1, \ldots, a_n)}$ and a target $\boldsymbol{\beta}$, find if there is a subset of the a_i that sums to β.

That is, find out if there is a feasible solution to

$$\boldsymbol{ax = \beta, \quad x \in \{0,1\}^n. \quad (*)}$$

Let

$$\boldsymbol{G_n(M) \ = \ \{\, a \,|\, a_i \in \{1, 2, \ldots, M\}\}.}$$

Known results:

- Lagarias and Odlyzko: Let $\boldsymbol{M \geq 2^{cn^2}}$. For **almost all** $\boldsymbol{a \in G_n(M)}$ their polytime algorithm **finds** the solution, if it exists. Almost all: for all but a fraction of $\boldsymbol{1/2^n}$.
- Frieze: simplifies and strengthens their argument.
- Furst, Kannan: With similar bound on M, they **solve** the subset sum problem for almost all a. (If there is no solution, they prove this).

First Testcase IP: The Subset Sum Problem

$$ax = \beta, \quad x \in \{0,1\}^n. \quad (*)$$

- Furst-Kannan proof technique: they find a short x candidate solution using BR. For almost all instances, if this x is not a solution, then there isn't one.
- These results don't deal with a small portion of the a vectors, so some a vectors could lead to hard instances
- Some instances should be hard from a theoretical viewpoint.
- Obviously: There are at most 2^n feasible right hand sides. So, if $\| a \|$ is large, then most of the possible $\sum_i a_i$ right hand sides are infeasible; but *proving* the infeasibility may be hard.
- We will look at **all** instances with a having large norm, and **almost all** β rhs values.

Main result on Subset Sums

Theorem 3. *Suppose* $a \in \mathbb{Z}^n$, $\| a \| \geq 2^{3n^2}$.

(1) Let $p = \mathrm{PA}(a)$. *Then for almost all* $\beta \in \{1, \ldots, \sum a_i\}$ *the infeasibility of* (*) *is proven by branching on* px.

(2) For almost all $\beta \in \{1, \ldots, \sum a_i\}$ *the infeasibility of* (*) *is proven by branching on* y_n *in the rangespace reformulation.*

Main result on Subset Sums

Theorem 4. *Suppose* $\boldsymbol{a} \in \mathbb{Z}^n, \|\boldsymbol{a}\| \geq \mathbf{2}^{3n^2}$.

(1) Let $\boldsymbol{p} = \mathbf{PA}(\boldsymbol{a})$. *Then for almost all* $\boldsymbol{\beta} \in \{\mathbf{1}, \ldots, \sum \boldsymbol{a}_i\}$ *the infeasibility of* **(*)** *is proven by branching on* $\boldsymbol{px}$.

(2) For almost all $\boldsymbol{\beta} \in \{\mathbf{1}, \ldots, \sum \boldsymbol{a}_i\}$ *the infeasibility of* **(*)** *is proven by branching on* $\boldsymbol{y_n}$ *in the rangespace reformulation.*

Theorem 5. *Suppose* $\boldsymbol{a} \in \mathbb{Z}^n, \|\boldsymbol{a}\| \geq \mathbf{2}^{3n^2}$.

(1) Let $\boldsymbol{p} = \mathbf{PA}(\boldsymbol{a})$. *Then for almost all* **infeasible** $\boldsymbol{\beta} \in \{\mathbf{1}, \ldots, \sum \boldsymbol{a}_i\}$ *the infeasibility of* **(*)** *is proven by branching on* $\boldsymbol{px}$.

(2) For almost all **infeasible** $\boldsymbol{\beta} \in \{\mathbf{1}, \ldots, \sum \boldsymbol{a}_i\}$ *the infeasibility of* **(*)** *is proven by branching on* $\boldsymbol{y_n}$ *in the rangespace reformulation.*

(Part of) Main result on PA, general version

Theorem 6. *Suppose a and U are as before, P_k the submatrix of U^{-1} made up of the last k rows, $a(k)$ the projection of a onto the rowspace of P_k, $r = a - a(k)$. Then*

$$\text{(1)} \quad \sqrt{\det P_k P_k^T}\sqrt{(1 + \| r \|^2)} \leq 2^{k(n+k-1)/4} \| a \|^{1-k/n}.$$

Example

$$a = (7621, 4565, 185, 8214, 4447)$$

$$U = \begin{pmatrix} -4 & 0 & -3 & 0 & 5 \\ 5 & 2 & -3 & 1 & -2 \\ -3 & -5 & -4 & -8 & 8 \\ 1 & -1 & 4 & -2 & -1 \\ 0 & 0 & 1 & 3 & -5 \end{pmatrix},$$

$$\mathbf{U}^{-1} = \begin{pmatrix} 372 & 223 & 9 & 401 & 217 \\ -656 & -393 & -16 & -707 & -383 \\ 127 & 76 & 3 & 137 & 74 \\ 581 & 348 & 14 & 626 & 339 \\ 374 & 224 & 9 & 403 & 218 \end{pmatrix},$$

$$\| r \| = 4.4857,\ 2.3100,\ 2.0809,\ 7 \times 10^{-5} \ \textbf{for}\ k = 1, 2, 3, 4;$$

Recap

(1) Parallel Approximation – a natural and basic problem.

(2) Precise definition of **“near parallel”**, and construction by Basis Reduction.

(3) Variant: successive approximation.

(4) Use of the near parallel vector: branching direction in Subset Sum problems.