Quadratic optimization subject to k quadratic has time complexity $n^{O(k)}$

joint work with
D. Grigoriev (IRMAR, Univ. Rennes I, FRANCE)
E. De Klerk (ISA/ITS, TU Delft, Netherlands)

10.10.2002. MSRI

Dima Pasechnik
University of Frankfurt

$$(*)\quad \inf f_0(X), \quad X \in V := \{Y \mid f_j(Y) \geq 0,\ 1 \leq j \leq K\}$$

semialgebraic

$$f_i \in \mathbb{R}[X] := \mathbb{R}[X_1, \dots, X_n], \quad 0 \leq i \leq K$$

polynomials

$$d := \max_i \deg f_i$$ – degree

What is the complexity of solving $(*)$, i.e. # of arithmetic operations needed w.r.t. size of $(*)$

size: n, K, d – BSS* model

— // — + bitlengths of the coefs of f_i's

bit model

*) Blum, Shub, Smale, BAMS (21) 1-46

Representing numbers and vectors

as radicals ($\sqrt{\;\;}$ ..) do <u>not</u> suffice...

$f(z) \in \mathbb{K}[t][r]$ parametrized (univariate) polynomial

(e.g. $\mathbb{K} = \mathbb{Q}$)

$\theta \in \mathbb{K}(\theta)(r)$: $f(\theta) = 0$, $d = \deg f$

root of f

$$\operatorname{sign}\left(\frac{df}{dt}, \frac{d^2 f}{dt}, \dots, \frac{d^{d-1} f}{dt}\right) = (\sigma_1, \dots \sigma_{d-1})$$

$\sigma_i \in \{-1, 1\}$ Thom encoding

$u \in \mathbb{K}(\theta)^m_{(r)}$: $u = \left(\frac{g_1(\theta)}{g_0(\theta)}, \dots, \frac{g_m(\theta)}{g_0(\theta)}\right)$

$g_i \in \mathbb{K}[z][r]$

univariate representation of u

$(f, g_0, g_1 \dots g_m) \in \mathbb{K}[z]^{m+2}_{[r]}$

+ Thom encoding

... that is how we represent the output

General situation for (*)

\# of operations $L\cdot(kd)^{O(n^2)}$
d maximal degree
n # of variables
k # of constraints
L - bitlength of the input ($L=1$ in BSS model)

Canny, Renegar, Basu, Pollack, Roy ← quantifier elimination from 1st order theory formulae over reals.

J. ACM 43 (1996), 1002-1045

\# of operations $L\cdot(kd)^{O(n)}$
de Klerk, Grigoriev, Pasechnik (2002)

used in an essential way for the $L\cdot n^{O(k)}$ - bound for $d=2$ (parametrized version)

---

makes difference only when minimizers are found along with the minimal value.

Particular complexity of cases of (*)

$d=1$ (linear programming)
polynomial in bit model (Khachiyan)
ε-approximation in BSS model
in polynomial of $\log\frac{1}{\varepsilon}$

K
n
o
w
n
u
p
to
n
o
w

$d \geq 4$ exponential (unless $P=NP$)
even for $K=0$
feasibility: $f_0(X) = \left(\sum_{(i,j)\in E\Gamma} (1 - X_i X_j) - C\right)^2 + \sum_i (X_i^2 - 1)^2 = 0$

encodes "does Γ have a cut of size $C/2$" MAXCUT

$d=2$, $K \leq 1$ (trust region)
polynomial in bit model
(folklore since 19??)
for $f_1 \succ 0$

$d=2$ feasibility tested in
$n^{O(K^2)}$ (A. Barvinok, Disc. Comp. Geom. 10 (1993), 1-13)
for homogeneous f_i's
(no feasible X computed)

Main results for $(*)$

$d=2$: the problem solved in $L \cdot n^{O(k)}$ operations (L - bit length of the input), minimizer X_{opt} found.

Moreover, the range of f_0 on V, $f_0(V) \subseteq \mathbb{R}$ found (same complexity).

$d=2$ Cor.[1] Decision version of $(*)$ is in P for fixed k.

Cor.[2] $f_0(V)$ has at most $n^{O(k)}$ components.

$d=2$ Cor.[3]

$$\max \langle C, X \rangle$$

$$X \in \{Y \mid \langle A_i, Y \rangle = b_i,\ 1 \le i \le k\}$$

$$\mathbb{R}^{n \times n} \ni X \succcurlyeq 0$$

$$\operatorname{rk} X \le \text{[redacted]} (\le n)$$

solvable in $n^{\min(\Theta(k), O(n^2))}$ operations (cf. announced in 1997 J. Glob. Opt. 10, 351-65 by Porkolab & Khachiyan)

Proof of Cor. 3

Let $G=(g_{ij})$, $1\le i\le n$; $1\le j\le r$

set $X = GG^T$. Then

$$f_0(G) = -\mathrm{Tr}(CGG^T)$$

$$f_i(G) = \mathrm{Tr}(A_i GG^T) - b_i \quad \square$$

Some reductions...

$$f_i \geq 0 \overset{1\le i\le k}{} \Rightarrow f_i - s_i = 0$$

new extra variables OK, as k is fixed

$$(*) \Rightarrow \min X_0$$

$$(X_0, X_1, \ldots, X_n) := X \in V$$

$$V := \{Y \mid F(Y) := \sum_{i=1}^{k} f_i(Y)^2 + (f_0(Y) - Y_0)^2 = 0\}$$

so V is a real algebraic set

Outline of the algorithm

1. Deform $F(X)$ to make it "generic"

$$\tilde{F}(X) = \sum_{i=1}^{k} \Big(\overbrace{X^T(H_i + \varepsilon_2 D_i)X + b_i^T X + c_i}^{\tilde{f}_i}\Big)^2 +$$

$$+ \Big(\overbrace{X^T(H_0 + \varepsilon_2 D_0)X + b_0^T X}^{\tilde{f}_0} - X_0\Big)^2 = \varepsilon_1$$

$$\mathbb{R} \Rightarrow \mathbb{R}((\varepsilon_1^{1/\infty}))((\varepsilon_2^{1/\infty})), \quad 0 \ll \varepsilon_2 \leq \varepsilon_1$$

Puiseux series

$$\tilde{V} = \bigcup_{\substack{U,W \subseteq \{1,\dots,n\} \\ r = |U| = |W| \geq n-k}} \varphi_{UW}(y, t)$$

$y = (y_1 \dots y_k)$

$t = (t_a, \dots, t_k)$

φ_{UW} - semialgebraic

2. Solve subproblems for $\varphi = \varphi_{UW}$

$$\inf_{X \in \varphi(y,t)} X_0$$

obtain candidates for the minimizer $\tilde{X} \in \mathbb{R}((e^{1/\infty}))^n$

3. Compute the "standard part" of each $\tilde{X}$:

$$st(\tilde{X}) = \lim_{\substack{\varepsilon_2 \to 0 \\ \varepsilon_1 \to 0}} \tilde{X}$$

and select the best one.

Puiseux series fields

$\mathbb{K}$ - a field (a real field), i.e. equipped with the ordering "$\leq$"$_{x^2 \geq 0}$; $-1 \neq \sum_j x_j^2$

$\mathbb{K}$ is real closed when $\mathbb{K}[i]$ is an algebraically closed field (here $i = \sqrt{-1}$)

$\mathbb{K}(z)$ - field of rational functions in z with coefficients in $\mathbb{K}$

$\mathbb{K}((z))$ — –// – formal power series (meromorphic) —//— $\left(\sum_{j \geq j_0} \alpha_j z^j \right)$

$\mathbb{K}((z^{1/\infty})) = \bigcup_{m=1}^{\infty} \mathbb{K}((z^{1/m}))$ —//— —//— Puiseux series —//—

It is real closed if $\mathbb{K}$ is.

1. Deformation and dimension reduction

- $\tilde{V}$ is smooth i.e.

  $\nabla \tilde{F} = 0$, $\tilde{F} = \varepsilon_1$ has no solutions

- By semialgebraic Sard's thm,*) the set of critical values of $\tilde{F}$ is 0-dim. over $\mathbb{R}((\varepsilon_2^{1/\infty}))$. But ε_1 is transcendental.

- at the extremal points of X_0 there exists Lagrange multiplier $\lambda \neq 0$.

$$1 = \lambda \frac{\partial \tilde{F}}{\partial X_0}$$

$$0 = \lambda \frac{\partial \tilde{F}}{\partial X_1} = \sum_{i=0}^{k} t_i((H_i + \varepsilon_2 D_i)X + b_i)$$

$$\vdots$$

$$0 = \lambda \frac{\partial \tilde{F}}{\partial X_n} \qquad \boxed{\begin{array}{c} t_i = f_i(X) \\ 0 \leq i \leq k \end{array}}$$

- thus we obtain linear conditions on X

$$\Phi(t)X := \sum_{i=0}^{k} t_i(H_i + \varepsilon_2 D_i)X = -\sum_{i=0}^{k} b_i t_i + \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

want $\Phi(t)$ of high rank!

- For any $t \in \mathbb{R}((e^{1/\infty})) - \{0\}$,
  $\operatorname{rk}\Phi(t) \geq n-k$
  when $D_i = \operatorname{diag}(1^i, 2^i, \dots, n^i)$

▶ Let $\mathcal{J}$ be the homogeneous ideal:

$$\mathcal{J} = \left(\det B_{UW} \;\middle|\; \begin{array}{l} U, W \in \{1, \dots n\} \\ |U| = |W| = n-k \end{array}\right)$$

$$\mathcal{J} \subseteq \mathbb{C}[y, z, \mu]$$

where

$$B(y, z, \mu) = \sum_{i=0}^{k} y_i (\mu H_i + z D_i)$$

$$(y^*, z^*, 1) \in Z(\mathcal{J}) \Leftrightarrow \operatorname{rk}(A(y^*, z^*)) < n-k$$

$$\begin{array}{ccc} \mathcal{J} & & Z(\mathcal{J}) \\ \downarrow & & \downarrow \\ \mathcal{J}' := \mathcal{J} \cap \mathbb{C}[z, \mu] & & Z(\mathcal{J}') \\ \Downarrow & & \# \\ 0 \neq p(z, \mu) & & \emptyset \end{array}$$

By "Main thm" of elimination theory, Zariski-closed

$\operatorname{rk}(A(y^*, z)) < n-k \Rightarrow p(z, 1) = 0$

$\Rightarrow z \neq \varepsilon_2$, as ε_2 is transcendental.

2. Solving low-dimensional subproblems for φ_{uw}

For each u, w, solve the parametric problem, w.r.t. to the parameter $r \geq 0$

$$\min X_0 \qquad \boxed{\varphi = \varphi_{uw}}$$

$(y,t) \in \tilde{V}_{uw}(r)$, where

$$\tilde{V}_{uw}(r) = \{(y,t) \mid \|\varphi(y,t)\|^2 \leq r^2,$$
$$\Delta := \det(\Phi_{uw}(t)) \neq 0,\ \tilde{F}(\varphi(y,t)) = \varepsilon_1 \}$$

Represent $\tilde{V}_{uw}(r)$ as an algebraic set given by $Q=0$, for

$$Q(y,t,r,Y) := (Y_2\Delta - 1)^2 + \Delta^8(\tilde{F}(\varphi(y,t)) - \varepsilon_1)^2 + \Delta^4(\|\varphi(y,t)\|^2 + Y_1^2 - r^2)$$

Can be done by quantifier elimination (the worsens the complexity to $n^{O(k^2)}$), or better by a modification of a procedure from Basu, Pollack, Roy (1996).

Obtain a set of param. univariate representations: (dropping Y)

$$(f(z,r), g_0(z,r), \dots, g_{k+1+n-r}(z,r))$$

defining points (y, t, X_0)

1. (cont'd) φ_{UW}

Let $U, W \subseteq \{1, 2, \dots, n\}$

$r = |U| = |W| \geq n - k$

$$\varphi := \varphi_{UW} : \mathbb{K}^{n-r} \times \mathbb{K}^{k+1} \to \mathbb{K}^n$$

$$\varphi(y, t) \longmapsto X$$

$$X_W = \Phi_{UW}^{-1}(t) \cdot (b(t)_U - \Phi_{U\overline{W}}(t) y)$$

$$X_{\overline{W}} = y$$

$\Phi_{UW}^{-1}(t)$ is computed by Cramer's rule

$$A_{ij}^{-1} = \frac{(-1)^{i+j} \det(A(i,j))}{\det A}$$

$\mathbb{K}[t] \ni \det A = \det(A(t))$ computed by "interpolation", $\deg(\det A) = n$

$$\text{extr. pts of } \tilde{V} = \bigcup_{\substack{U,W \\ |U|=|W| \geq n-k}} \{\varphi_{UW}(y,t) \mid \tilde{F}(\varphi_{UW}(y,t)) = \varepsilon_1\}$$

//

3. Computing $st(X)$

Recall that $st(X) := \lim_{\varepsilon \to 0} X$

$X \in \tilde{V}$.

- $st(\tilde{V}) = V$

▶ Obviously, $st(\tilde{V}) \subseteq V$.

For $X \in V$, $\tilde{F}(X) < \varepsilon_1$.

On the other hand, there exists

Z satisfying $\tilde{F}(X + \varepsilon_1^{1/4} Z) > \varepsilon_1$

$\Rightarrow$ $Y = \lambda X + (1-\lambda)(X + \varepsilon_1^{1/4} Z)$ satisfies

$\tilde{F}(Y) = \varepsilon_1$ and $st(Y) = X$. ◀

- $X = \varphi_{UW}(y(z,r), t(z,r))$

X is defined by param. univariate representation

$(f(z,r), g_0(z,r) \cdot \Delta(z,r), h_1(z,r), \ldots$

$\ldots, h_{n+1}(z,r))$

$f, g_0, h_i \in \mathbb{K}[z,r][\varepsilon_1, \varepsilon_2]$,

of degrees at most $n^{O(k)}$.

3! Computing $\lim_{\varepsilon_2 \to 0} X(z, r, \varepsilon_1, \varepsilon_2)$

- $f(z, r, \varepsilon_1, \varepsilon_2) = \sum_{i=0}^{m} \varepsilon_2^i f_i(z, r, \varepsilon_1)$

  Compute Thom encodings for the real roots of f_0 (assuming $\frac{1}{r} \gg \varepsilon_1$ an infinitesimal)

  Disregard roots Θ with $\|\Theta\| \notin \mathbb{R}$.

  W.l.o.g.

  $$g_0 \cdot \Delta = \sum_{i=0}^{m'} \varepsilon_2^i q_i(z, r, \varepsilon_1) \text{ and}$$

  $$q_0 \neq 0$$

- If $\gcd(f_0, q_0) \neq 1$, divide q_0 by it

- Form (at most $n^{O(k)}$) representations

  $$(f_0, q_0, \dots, \overline{h}_i, \dots) \in \mathbb{R}[z, r, \varepsilon_1]$$

  $$\operatorname{sign}(f_0', f_0'', \dots f_0^{(N)}) = \sigma_\Theta$$

- Repeat for ε_1 & $\frac{1}{r}$.

- Finally, select the best among $Y = St(X)$ satisfying $F(Y) = 0$