Complexity of optimization

Nilava Metya

October 21, 2023

Goal

Solve polynomial systems of equations.

Example: Sudoku

$$\begin{bmatrix} x_1 & x_2 & x_3 & x_4 \\ x_5 & x_6 & x_7 & x_8 \\ x_9 & x_{10} & x_{11} & x_{12} \\ x_{13} & x_{14} & x_{15} & x_{16} \end{bmatrix}$$

We want[*] to consider the ideal generated by $F_j = \prod_{k=1}^{4}(x_j - k) = x_j^4 - 10x_j^3 + 35x_j^2 - 50x_j + 24$ for each $j = 1, \cdots, 16$ and the polynomials. And also the polynomials

$$G_{ij} = \frac{F_i - F_j}{x_i - x_j} = x_i^3 + x_i^2 x_j + x_i x_j^2 + x_j^3 - 10(x_i^2 + x_i x_j + x_j^2) + 35(x_i + x_j) - 50$$

for $i \neq j$. These polynomials determine the space of solutions to the above sudoku. Additionally we want to input the information given as the starting point of the sudoku.

[*] Jesús Gago-Vargas, María Isabel Hartillo-Hermoso, Jorge Martín-Morales, and José María Ucha-Enríquez. "Sudokus and Gröbner Bases: Not Only a Divertimento". In: *Computer Algebra in Scientific Computing*. 2006. URL: https://api.semanticscholar.org/CorpusID:11562585.

Example of example: Sudoku

$$\begin{bmatrix} 2 & 4 & x_3 & x_4 \\ x_5 & 1 & x_7 & 2 \\ 1 & x_{10} & x_{11} & 4 \\ x_{13} & x_{14} & 1 & 3 \end{bmatrix}$$

I took my ideal to be generated by the relations

$$\begin{aligned} \text{row sum} &= 10 \\ \text{column sum} &= 10 \\ \text{block sum} &= 10 \end{aligned}$$

and the additional things like $x_1 - 2, x_2 - 4, \cdots$.
`M2` gives solution

$$\begin{bmatrix} 2 & 4 & 3 & 1 \\ 3 & 1 & 4 & 2 \\ 1 & 3 & 2 & 4 \\ 4 & 2 & 1 & 3 \end{bmatrix}$$

Example of example: Sudoku

$$\begin{bmatrix} 3 & 4 & x_3 & x_4 \\ x_5 & 1 & x_7 & 2 \\ 1 & x_{10} & x_{11} & 4 \\ x_{13} & x_{14} & 1 & 3 \end{bmatrix}$$

has no solution

But `M2` gives solution

$$\begin{bmatrix} 3 & 4 & 2 & 1 \\ 2 & 1 & 5 & 2 \\ 1 & 3 & 2 & 4 \\ 4 & 2 & 1 & 3 \end{bmatrix}$$

Once I enforce that $F_j = 0 \forall 1 \leq j \leq 16$, `M2` indeed says that there is no solution.

Since I'm a hater of learning by reading, and prefer learning by computing examples ...

I shall not define what a (reduced) Gröbner basis is.

Gaussian elimination

$$
\begin{aligned}
5x + 7y &= 1 \\
3x + 10y &= -3
\end{aligned}
$$

Gaussian elimination

$$
\begin{aligned}
{\color{red}5}x + 7y &= 1 \\
{\color{red}3}x + 10y &= -3
\end{aligned}
$$

Gaussian elimination

$$\begin{aligned} \color{red}{5}x + 7y &= 1 && \times 3 \\ \color{red}{3}x + 10y &= -3 && \times 5 \end{aligned}$$

We get: $-29y = 18$. Then plug back y.

A slight change in perspective

Instead of looking at

$$\begin{aligned} 5x + 7y &= 1 \\ 3x + 10y &= -3 \end{aligned}$$

A slight change in perspective

I urge your to look at

$$\begin{aligned} \color{red}{5x} + 7y &= 1 \\ \color{red}{3x} + 10y &= -3 \end{aligned}$$

Non-linear analog of Gaussian elimination

$$x^2y + 8 = 0$$
$$xy^2 - 4 = 0$$

Non-linear analog of Gaussian elimination

$$x^2y + 8 = 0$$
$$xy^2 - 4 = 0$$

Non-linear analog of Gaussian elimination

$$\begin{aligned} x^2y + 8 &= 0 & \times y \\ xy^2 - 4 &= 0 & \times x \end{aligned}$$

We get $-2y = x$. Plugging into the first equation gives $2x^3 = 8 \implies x = \sqrt[3]{16} \implies y = -\sqrt[3]{2}$.

`M2` gives the reduced Gröbner basis of the ideal $\langle x^2y + 8, xy^2 - 4 \rangle$ as $\{x + 2y, y^3 + 2\}$.

An example

Maximize $f = x + y + z$ subject to $g = x^4 + y^4 + 3z^4 - z - 1 = 0$

$\mathcal{L} = (x + y + z) + \lambda(x^4 + y^4 + 3z^4 - z - 1)$.

$$\partial_x \mathcal{L} = 1 + 4\lambda x^3$$
$$\partial_y \mathcal{L} = 1 + 4\lambda y^3$$
$$\partial_z \mathcal{L} = 1 + \lambda(12z^3 - 1)$$
$$\partial_\lambda \mathcal{L} = g$$

Trying to define an ideal in `SageMath` given by the above generators and finding a Gröbner basis tells us that we need to solve an equation of degree 36.
If we add another **generic** linear constraint, this degree is now 12.
Another **generic** linear constraint makes the degree 4.
Adding another **generic** equation means that there's no solution, which gives degree 0.

Define these numbers to be the algebraic degrees: $d_1 = 36, d_2 = 12, d_3 = 4$.

Polar Variety

Imagine a compact ellipsoid X and a point $V = \boldsymbol{v}$. Imagine that your eyes are at $\boldsymbol{v}$. What do you see?

Picture on blackboard

Suppose $X \subseteq \mathbb{P}^3$ is given by a homogeneous polynomial f of degree d and $\boldsymbol{v} = (v_0 : v_1 : v_2 : v_3)$ is the point where your eyes are. What you see is a curve, name it $P(X, \boldsymbol{v})$, is determined by f and $\partial_{\boldsymbol{v}} f$.

Theorem (Bezout)

Let $f_1, \cdots, f_k$ be general polynomials in n variables of degree $d_1, \cdots, d_k$ respectively. For $I = \langle f_1, \cdots, f_k \rangle$ we have $\dim I = n - k$ and $\deg I = d_1 \cdots d_k$.

So this $P(X, V)$ *typically* has degree $d(d-1)$.

DEFINITION (POLAR VARIETY)

The polar variety of a variety $X \subseteq \mathbb{P}^n$ with respect to a projective subspace $V \subseteq \mathbb{P}^n$ is

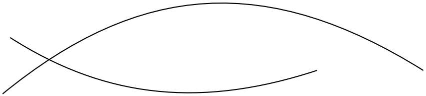
$$P(X,V) = \overline{\{\boldsymbol{p} \in \operatorname{Reg}(X) \setminus V : V + \boldsymbol{p} \text{ intersects } X \text{ at } \boldsymbol{p} \text{ non-transversally}\}}.$$

Let $i \in \{0, 1, \cdots, \dim X\}$. If V is generic with $\dim V = \operatorname{codim}(X) - 2 + i$, then the degree of $P(X,V)$ is independent of V:

$$\mu_i(X) = \deg P(X,V).$$

Careful about transversality

Transversality depends on the ambient space. . .



The above intersection is transversal in $\mathbb{R}^2$, but non-transversal in $\mathbb{R}^3$.

DEFINITION (CONORMAL VARIETY)

The *conormal variery* $N_X \subseteq \mathbb{P}^n \times \mathbb{P}^n$ is the Zariski closure of of the collection of all pairs $(\boldsymbol{x}, \boldsymbol{h}) \in \mathbb{P}^n \times \mathbb{P}^n$ such that $\boldsymbol{x}$ is a non-singular point in X and $\boldsymbol{h}$ represents a hyperplane tangent to X at $\boldsymbol{x}$.

Now $H^*(\mathbb{P}^n \times \mathbb{P}^n, \mathbb{Z}) = \mathbb{Z}[s,t]/\left\langle s^{n+1}, t^{n+1}\right\rangle$. The class of the conormal variety N_X in this cohomology ring is a binary form of degree $n+1 = \mathrm{codim}(N_X)$ whose coefficients are nonnegative integers:

$$[N_X] = \sum_{i=1}^{n} \delta_i(X) s^{n+1-i} t^i$$

.

THEOREM

$\delta_i(X) = \mu_i(X)$.

For a general optimization problem

Given a compact smooth algebraic variety $\mathcal{M}$ in $\mathbb{R}^m$, we consider a linear functional ℓ and an affine-linear space L of codimension r in $\mathbb{R}^m$. It is assumed that the pair (ℓ, L) is in general position[†] relative to $\mathcal{M}$. Our aim is to study the following optimization problem:

$$\text{maximize } \ell \text{ over } L \cap \mathcal{M}.$$

[‡]

Theorem

The algebraic degree of the above problem is $\mu_r(\mathcal{M})$.

[†] this assumption is very important

[‡] Türkü Özlüm Çelik, Asgar Jamneshan, Guido Montúfar, Bernd Sturmfels, and Lorenzo Venturello. "Wasserstein distance to independence models". In: *Journal of symbolic computation* 104 (2021), pp. 855–873.