A Theory of Complexity, Condition and Roundoff

Felipe Cucker

Berkeley 2014

Background

L. Blum, M. Shub, and S. Smale [1989]

On a theory of computation and complexity over the real numbers: NP-completeness, recursive functions and universal machines.

$\mathrm{P}_{\mathbb{R}}$ $\mathrm{NP}_{\mathbb{R}}$

> "Finally, to bring machines over $\mathbb{R}$ closer to the subject of numerical analysis, it would be useful to incorporate round-off error, condition numbers and approximate solutions into our development."

Finite-precision computations

Floating-point number system: $\mathbb{F} \subset \mathbb{R}$.

$$y = \pm\, 0.b_1 b_2 \ldots b_t \times 2^e \qquad |e| \le \mathsf{e}_{\max}$$

$$\mathsf{Range}(\mathbb{F}) := \left[-2^{\mathsf{e}_{\max}}(1-2^{-t}), -2^{-\mathsf{e}_{\max}-1}\right] \cup \{0\} \cup \left[2^{-\mathsf{e}_{\max}-1}, 2^{\mathsf{e}_{\max}}(1-2^{-t})\right].$$

Unit roundoff: $\mathsf{u}_{\mathsf{mach}} := 2^{-t}$.

Rounding function: $\mathsf{fl} : \mathsf{Range}(\mathbb{F}) \to \mathbb{F}$

for all $x \in \mathsf{Range}(\mathbb{F})$, $\mathsf{fl}(x) = x(1+\delta)$ for some δ with $|\delta| < \mathsf{u}_{\mathsf{mach}}$.

Floating-point arithmetic: to $\circ \in \{+, -, \times, /\}$ we associate

$$\widetilde{\circ} : \mathbb{F} \times \mathbb{F} \to \mathbb{F}$$

$$x \,\widetilde{\circ}\, y = (x \circ y)(1+\delta) \text{ for some } \delta \text{ with } |\delta| < \mathsf{u}_{\mathsf{mach}}.$$

Unrestricted exponents: $\mathsf{e}_{\max} = \infty$ $\quad(\mathsf{Range}(\mathbb{F}) = \mathbb{R})$.

Most analyses in the literature assume unrestricted exponents.

Stability and condition

Two factors in the accumulation of errors in a computation:

(1) How sensitive is the result of the computed function φ to perturbations of the data d?

Condition number $\quad \mathrm{cond}^{\varphi}(d)$

it depends only on φ and d

(2) How badly does the algorithm at hand accumulate errors?

Stability analysis

it depends on the algorithm and the dimension of d

Example Linear equation solving: $(A,b) \stackrel{\varphi}{\mapsto} x = A^{-1}b$. Under the assumption of unrestricted exponents, we have

$$\mathsf{cond}^{\varphi}(A) = \kappa(A) := \|A\|\|A^{-1}\|.$$

The computed (using Householder QR decomposition) solution $\widetilde{x}$ satisfies, for some constant C,

$$\frac{\|\widetilde{x} - x\|}{\|x\|} \leq Cn^3 \mathsf{u_{mach}}\, \kappa(A) + o(\mathsf{u_{mach}}). \tag{1}$$

Remark Hestenes and Stiefel showed that $\kappa(A)$ also plays a role in complexity analyses.

Important remark: A wide variety of computational problems:

- decisional
- functional
- set-valued

. . .

results in a variety of condition numbers.

Condition numbers are defined “ad hoc”.

The Theory

Decision problems

Data has discrete and continuous components:

$$\mathcal{I} := \{0,1\}^\infty \times \mathbb{R}^\infty.$$

Here

$$\mathbb{R}^\infty := \bigsqcup_{i=0}^{\infty} \mathbb{R}^i \qquad\qquad \{0,1\}^\infty := \bigsqcup_{i=0}^{\infty} \{0,1\}^i.$$

Definition A *decision problem* is a pair $(A, \boldsymbol{\mu})$ where $A \subset \mathcal{I}$ and $\boldsymbol{\mu} : \mathcal{I} \to [1, \infty]$. Here $\boldsymbol{\mu}$ is the *condition number*.
We denote by Σ the set $\{(u,x) \in \mathcal{I} \mid \boldsymbol{\mu}(u,x) = \infty\}$ and we say that elements in Σ are *ill-posed*.

Remark Different condition numbers for the same subset $A \subset \mathcal{I}$ define different decision problems. This is akin to the situation in classical (i.e., both discrete and infinite-precision BSS) complexity theory where different encodings of the intended input data define (sometimes radically) different problems.

Finite-precision machines, input size, and cost

Definition A *finite-precision BSS machine* is a BSS machine performing finite-precision computations. To define the latter, we fix a number $\mathsf{u}_{\mathsf{mach}} \in (0,1)$ (the *unit roundoff*) and let

$$\mathsf{k}_{\mathsf{mach}} := \left\lceil \log_2 \frac{1}{\mathsf{u}_{\mathsf{mach}}} \right\rceil.$$

In a $\mathsf{u}_{\mathsf{mach}}$*-computation* , built-in constants, input values, and the result of arithmetic operations, call any such number z, are systematically replaced by $\mathsf{fl}(z)$ satisfying

$$\mathsf{fl}(z) = z(1+\delta) \quad \text{for some } |\delta| < \mathsf{u}_{\mathsf{mach}}. \tag{2}$$

We will refer to $\mathsf{k}_{\mathsf{mach}} \in \mathbb{N}$ as the *precision* of M.

Complexity = dependence of cost on size.

For $(u,x) \in \{0,1\}^s \times \mathbb{R}^n \subset \mathcal{I}$, we let length$(u,x)$ to be $s+n$ and

$$\text{size}(u,x) := \text{length}(u,x) + \lceil \log_2 \boldsymbol{\mu}(u,x) \rceil.$$

Note that if (u,x) is ill-posed then size$(u,x) = \infty$ and that otherwise size$(u,x) \in \mathbb{N}$.

Arithmetic cost: number of steps performed before halting. We denote it by ar_cost$_M(u,x)$.

Accuracy cost: smallest value of $\mathsf{k}_{\mathsf{mach}}$ guaranteeing a correct answer.

Close to the cost in practice of operating with floating-point numbers since, assuming the exponents of such numbers are moderately bounded, this cost is at most quadratic on $\mathsf{k}_{\mathsf{mach}}$.

Clocked computations

Definition Let Arith : $\mathbb{N}\times\mathbb{N}\to\mathbb{N}$ and Prec : $\mathbb{N}\to\mathbb{N}$. A decision problem $(S,\boldsymbol{\mu})$ is *solved with cost* (Arith, Prec) when there exists a machine M satisfying the following. For every $(u,x)\in\mathcal{I}$ with $\boldsymbol{\mu}(u,x)<\infty$ the computation of M with input (u,x) satisfies

$$\mathsf{ar\_cost}_M(u,x)\le \mathsf{Arith}(\mathsf{length}(u,x),\mathsf{k_{mach}}).$$

Furthermore, if

$$\mathsf{k_{mach}}\ge \mathsf{Prec}(\mathsf{size}(u,x))$$

then all computations of M correctly decide whether $(u,x)\in S$.

(i) Computations are clocked, i.e., their arithmetic cost is bounded by a function on two parameters immediately available: length of the input data and machine precision.

(ii) Computations are unreliable: there is no guarantee that the precision used is enough to ensure a correct output. Even for exact computations correctness is not guaranteed.

A hierarchy theorem

Proposition **(Precision Hierarchy Theorem)** *Let* $T : \mathbb{N} \to \mathbb{N}$ *be time constructible and* $P_1, P_2 : \mathbb{R}_+ \to \mathbb{R}_+$ *such that* P_2 *is continuous and increasing and* $P_1 < \frac{P_2}{2}$. *There exists a decision problem* $(B, \boldsymbol{\mu})$ *which can be decided with* $\mathsf{ar\_cost}(u, x) \leq \mathcal{O}(T(\mathsf{length}(u, x)))$ *and* $\mathsf{k_{mach}} = P_2(\mathsf{size}(u, x)) + 3$, *but cannot be decided with* $\mathsf{k_{mach}} = P_1(\mathsf{size}(u, x))$ *(no matter the arithmetic cost).*

General polynomial time: the class P_{ro}

Definition A decision problem $(S, \boldsymbol{\mu})$ belongs to P_{ro} (*roundoff polynomial cost*) when there exists a finite-precision BSS machine M solving S with cost $(\mathsf{Arith}, \mathsf{Prec})$ and such that

(i) Prec is bounded by a polynomial function, and

(ii) the function $\mathsf{Arith}(\mathsf{length}(u,x), \mathsf{Prec}(\mathsf{size}(u,x)))$ is bounded by a polynomial in $\mathsf{size}(u,x)$, for all $(u,x) \in \mathcal{I}$.

Direct algorithms: the class $\mathrm{P}_{\mathsf{dir}}$

Definition A decision problem $(S, \boldsymbol{\mu})$ belongs to $\mathrm{P}_{\mathsf{dir}}$ (*direct polynomial cost*) when there exists a machine M satisfying the following. For every $(u, x) \in \mathcal{I}$ the computation of M with input (u, x) satisfies

$$\mathsf{ar\_cost}_M(u, x) \leq (\mathsf{length}(u, x))^{\mathcal{O}(1)}.$$

Furthermore, if

$$\mathsf{k_{mach}} \geq (\mathsf{size}(u, x))^{\mathcal{O}(1)}$$

then all computations of M correctly decide whether $(u, x) \in S$. If correctness is ensured as soon as $\mathsf{k_{mach}} \geq (\log \mathsf{size}(u, x))^{\mathcal{O}(1)}$ we say that $(S, \boldsymbol{\mu})$ can be solved with *logarithmic precision*.

Examples. Deciding whether $\det(A) > 0$, whether S is p.s.d., ...

Proposition *We have* $\mathrm{P}_{\mathsf{dir}} \subsetneq \mathrm{P}_{\mathsf{ro}}$.

Notation:

$\mathcal{C}$ an algebraic circuit (n input gates, 1 output gate)

$f_{\mathcal{C}} : \mathbb{R}^n \to \mathbb{R}$ function computed by the circuit

$$S_{\mathcal{C}} := \{x \in \mathbb{R}^n \mid f_{\mathcal{C}}(x) \geq 0\}.$$

Example Instances for CircEval are algebraic circuits $\mathcal{C}$ together with a point $x \in \mathbb{R}^n$. The problem is to decide whether $x \in S_{\mathcal{C}}$. To specify a condition number we first define

$$\varrho_{\mathsf{eval}}(\mathcal{C}, x) := \begin{cases} \sup\{\varepsilon < 1 \mid \text{all } \varepsilon\text{-evaluations of } \mathcal{C} \text{ at } x \text{ yield } x \in S_{\mathcal{C}}\} & \text{if } x \in S_{\mathcal{C}} \\ \sup\{\varepsilon < 1 \mid \text{all } \varepsilon\text{-evaluations of } \mathcal{C} \text{ at } x \text{ yield } x \notin S_{\mathcal{C}}\} & \text{otherwise.} \end{cases}$$

We then take as condition number

$$\mu_{\mathsf{eval}}(\mathcal{C}, x) := \frac{1}{\varrho_{\mathsf{eval}}(\mathcal{C}, x)}.$$

Nondeterministic Polynomial Cost

Problems in (all versions of) NP are sets S for which membership of an element x to S can be established through a "short" proof y.

$\mathsf{NP}, \mathrm{NP}_{\mathbb{R}}$ short $=$ small length

$\mathrm{NP}^{\mathsf{U}}_{\mathsf{ro}}$ short $=$ small length $+$ small condition

$\mathrm{NP}^{\mathsf{B}}_{\mathsf{ro}}$ short $=$ small length $+$ small condition $+$ small magnitude

The class $\mathrm{NP}^{\mathsf{U}}_{\mathsf{ro}}$

Definition A decision problem $(W, \boldsymbol{\mu}_W)$ belongs to $\mathrm{NP}^{\mathsf{U}}_{\mathsf{ro}}$ (*non-deterministic roundoff polynomial cost*) when there exist a decision problem $(B, \boldsymbol{\mu}_B)$, a machine M deciding $(B, \boldsymbol{\mu}_B)$ in P_{ro}, and polynomials p, Q, such that for $(u, x) \in \mathcal{I}$,

(i) if $(u, x) \in W$ then there exists $y^* \in \mathbb{R}^m$, such that $(u, x, y^*) \in B$, and $\log \boldsymbol{\mu}_B(u, x, y^*) \leq Q(\log \boldsymbol{\mu}_W(u, x))$, and

(ii) if $(u, x) \notin W$ then, for all $y \in \mathbb{R}^m$ we have $(u, x, y) \notin B$ and $\log \boldsymbol{\mu}_B(u, x, y) \leq Q(\log \boldsymbol{\mu}_W(u, x))$.

Here $m = p(\mathsf{length}(u, x))$.

Example Instances for CircFeas are algebraic circuits $\mathcal{C}$ (with input variables $Y_1, \ldots, Y_m$). The problem is to decide whether there exists $y \in \mathbb{R}^m$ such that $y \in S_{\mathcal{C}}$ (in which case, we say that $\mathcal{C}$ is *feasible*). We take as condition number

$$\mu_{\mathsf{feas}}(\mathcal{C}) := \frac{1}{\varrho_{\mathsf{feas}}(\mathcal{C})}$$

where

$$\varrho_{\mathsf{feas}}(\mathcal{C}) := \begin{cases} \sup_{y \in S_{\mathcal{C}}} \varrho_{\mathsf{eval}}(\mathcal{C}, y) & \text{if } \mathcal{C} \text{ is feasible,} \\ \inf_{y \in \mathbb{R}^m} \varrho_{\mathsf{eval}}(\mathcal{C}, y) & \text{otherwise.} \end{cases}$$

Note that in the feasible case, $\mu_{\mathsf{feas}}(\mathcal{C})$ is the condition of its best conditioned solution, and in the infeasible case, it is the condition of the worst conditioned point in $\mathbb{R}^m$.

Proposition $\mathsf{CircFeas} \in \mathrm{NP}^{\mathsf{U}}_{\mathsf{ro}}$.

Proposition $\mathrm{P}_{\mathsf{ro}} \subset \mathrm{NP}^{\mathsf{U}}_{\mathsf{ro}}$.

Definition A P_{ro}-*reduction* from $(W, \boldsymbol{\mu}_W)$ to $(S, \boldsymbol{\mu}_S)$ is a machine $\overline{M}$ which, given a point $(u, x) \in \mathcal{I}$ and a number $k \in \mathbb{N}$, performs a discrete computation and returns a pair $(v, z) \in \mathcal{I}$ with $\mathsf{ar\_cost}_{\overline{M}}(u, x)$ polynomially bounded on $\mathsf{length}(u, x)$ and k.

In addition, we require the existence of some $D, p > 0$ such that for all $k \geq D\,\mathsf{size}(u, x)^p$ one has

(i) $(u, x) \in W \iff (v, z) \in S$, and

(ii) $\log \boldsymbol{\mu}_S(v, z)$ is polynomially bounded in $\mathsf{size}_W(u, x)$.

If all of the above holds, we write $(W, \boldsymbol{\mu}_W) \preceq_{\mathsf{ro}} (S, \boldsymbol{\mu}_S)$.

Proposition *If* $(W, \boldsymbol{\mu}_W) \preceq_{\mathsf{ro}} (S, \boldsymbol{\mu}_S)$ *and* $(S, \boldsymbol{\mu}_S) \in \mathrm{P}_{\mathsf{ro}}$ *then* $(W, \boldsymbol{\mu}_W) \in \mathrm{P}_{\mathsf{ro}}$.

Definition A problem $(S, \boldsymbol{\mu}_S)$ is $\mathrm{NP}^{\mathsf{U}}_{\mathsf{ro}}$-*hard* when for any problem $(W, \boldsymbol{\mu}_W) \in \mathrm{NP}^{\mathsf{U}}_{\mathsf{ro}}$ we have $(W, \boldsymbol{\mu}_W) \preceq_{\mathsf{ro}} (S, \boldsymbol{\mu}_S)$.

It is $\mathrm{NP}^{\mathsf{U}}_{\mathsf{ro}}$-*complete* when it is $\mathrm{NP}^{\mathsf{U}}_{\mathsf{ro}}$-hard and belongs to $\mathrm{NP}^{\mathsf{U}}_{\mathsf{ro}}$.

Theorem CircFeas *is* $\mathrm{NP}^{\mathsf{U}}_{\mathsf{ro}}$-*complete.*

Corollary $\mathrm{P}_{\mathsf{ro}} = \mathrm{NP}^{\mathsf{U}}_{\mathsf{ro}} \iff \mathsf{CircFeas} \in \mathrm{P}_{\mathsf{ro}}$.

Open Question Does one have $\mathrm{P}_{\mathsf{ro}} = \mathrm{NP}^{\mathsf{U}}_{\mathsf{ro}}$? As usual, we believe this is not the case.

The class $\mathrm{NP}^{\mathsf{B}}_{\mathsf{ro}}$

Given $k \in \mathbb{N}$ we consider the floating-point system F_k with

$$t = k, \quad \text{and} \quad \mathsf{e}_{\max} = 2^k - 1.$$

For $x \in \mathbb{R}$ we define the *magnitude* of x to be

$$\mathsf{mgt}(x) := \min\{k \geq 1 \mid x \in \mathsf{Range}(F_k)\},$$

and for $x \in \mathbb{R}^n$, $\mathsf{mgt}(x) := \max_{i \leq n} \mathsf{mgt}(x_i)$.

Definition A decision problem $(W, \boldsymbol{\mu}_W)$ belongs to $\mathrm{NP}_{\mathsf{ro}}^{\mathsf{B}}$ (*bounded non-deterministic roundoff polynomial cost*) when there exist a decision problem $(B, \boldsymbol{\mu}_B)$, a machine M deciding $(B, \boldsymbol{\mu}_B)$ in P_{ro}, and polynomials p, q, Q, such that for $(u, x) \in \mathcal{I}$,

(i) if $(u, x) \in W$ then there exists $y^* \in \mathbb{R}^m$, such that $(u, x, y^*) \in B$, $\log \boldsymbol{\mu}_B(u, x, y^*) \leq Q(\log \boldsymbol{\mu}_W(u, x))$, and $\mathsf{mgt}(y^*) \leq q(\mathsf{size}_W(u, x))$, and

(ii) if $(u, x) \notin W$ then, for all $y \in \mathbb{R}^m$ we have $(u, x, y) \notin B$ and $\log \boldsymbol{\mu}_B(u, x, y) \leq Q(\log \boldsymbol{\mu}_W(u, x))$.

Here $m = p(\mathsf{length}(u, x))$.

Example Instances for CircBFeas are algebraic circuits $\mathcal{C}$ (with input variables $Y_1, \dots, Y_m$). The problem is to decide whether there exists $y \in \mathbb{R}^m$ such that $y \in S_{\mathcal{C}}$. What makes this problem different from CircFeas is its condition number. Here we take

$$\mu_{\mathsf{Bfeas}}(\mathcal{C}) := \frac{1}{\varrho_{\mathsf{Bfeas}}(\mathcal{C})}$$

where

$$\varrho_{\mathsf{Bfeas}}(\mathcal{C}) := \begin{cases} \sup_{y \in S_{\mathcal{C}}} \varrho_{\mathsf{eval}}(\mathcal{C}, y) 2^{-\mathsf{mgt}(y)} & \text{if } \mathcal{C} \text{ is feasible,} \\ \inf_{y \in \mathbb{R}^m} \varrho_{\mathsf{eval}}(\mathcal{C}, y) & \text{otherwise.} \end{cases}$$

Theorem CircBFeas *is* $\mathrm{NP}^{\mathsf{B}}_{\mathsf{ro}}$*-complete.*

Corollary $\mathrm{P}_{\mathsf{ro}} = \mathrm{NP}^{\mathsf{B}}_{\mathsf{ro}} \iff \mathsf{CircBFeas} \in \mathrm{P}_{\mathsf{ro}}$.

Exponential cost

Definition A decision problem $(S, \boldsymbol{\mu})$ belongs to $\mathrm{EXP}_{\mathsf{ro}}$ (*roundoff exponential cost*) when there exists a finite-precision BSS machine M deciding S with cost $(\mathsf{Arith}, \mathsf{Prec})$ and such that

(i) Prec is bounded by a exponential function, and

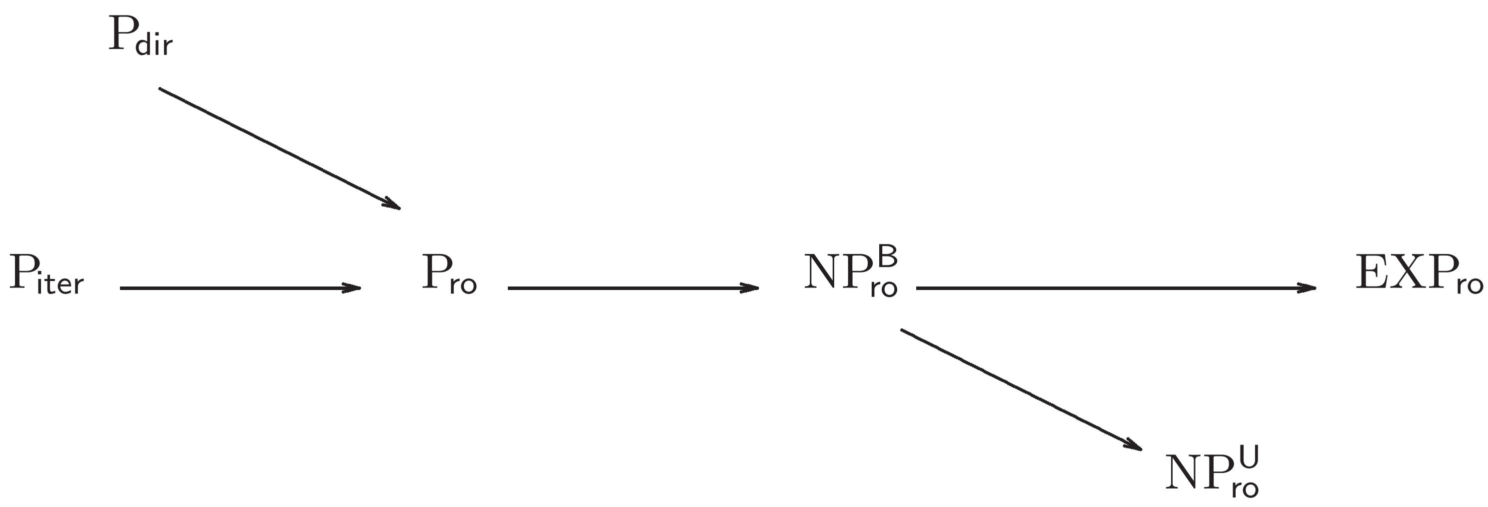
(ii) the function $\mathsf{Arith}(\mathsf{length}(u, x), \mathsf{Prec}(\mathsf{size}(u, x)))$ is bounded by an exponential in $\mathsf{size}(u, x)$, for all $(u, x) \in \mathcal{I}$.

In both (i) and (ii) by exponential we understand a function of the kind $n \mapsto a^{n^d}$ for some $a > 1$ and $d > 0$.

Theorem $\mathsf{CircBFeas} \in \mathrm{EXP}_{\mathsf{ro}}$.

Corollary $\mathrm{NP}^{\mathsf{B}}_{\mathsf{ro}} \subset \mathrm{EXP}^{|}_{\mathsf{ro}}$ *and the inclusion is strict.*

Open Question A major open question in this is whether $\mathrm{NP}^{\mathsf{U}}_{\mathsf{ro}}$ is included in $\mathrm{EXP}_{\mathsf{ro}}$ or, equivalently, whether $\mathsf{CircFeas}$ belongs to $\mathrm{EXP}_{\mathsf{ro}}$. We conjecture that this question has a positive answer.


P_{dir}
P_{iter}
P_{ro}
NP^{B}_{ro}
EXP_{ro}
NP^{U}_{ro}