Approximate Commutative Algebra – an Impossible Concept ?

Hans J. Stetter , Vienna

stetter@aurora.anum.tuwien.ac.at

Workshop on Approximate Commutative Algebra, 20 - 24 Feb. 2006
in: Special Semester on Gröbner Bases, RICAM, Linz, Austria

Algebra and Approximation
a Fundamental Antithesis

Algebra		Analysis
discrete world		continuous world
symbols, structures		spaces, functions
no neighborhoods		neighborhoods
equal <=> not equal	elements	equal . . . close . . . distant
satisfied <=> not satisfied	equations	sat. . . . nearly sat. . . . not satisfied
true <=> false	relations	nearly true . . . false
discrete	mappings	continuous

Approximation can only be considered in an analytic context, hence

(Commutative) Algebra and Approximation

are **antithetic** concepts

But many problems in Commutative Algebra
are posed over number fields which carry a

natural topology

e.g. real numbers or complex numbers

Here, $1.41421 \quad \approx \quad \sqrt{2}$ may be

an obvious approximation

This suggests an

embedding

of commutative algebra into analysis

via the coefficient field

Successful Prototype for this Embedding :

Linear Algebra over $\mathbb{R}^n$ or $\mathbb{C}^n$

becomes

Analytic Geometry

becomes

Numerical Linear Algebra

Embedding into Analysis permits

- introduction of approximation
- handling of data with limited accuracy
- use of analytic tools for problem solving
- use of floating-point in computation

Embedding requires

- reconsideration of all concepts
- redefinition of many concepts
- introduction of various new concepts
- disposal of some classical concepts

Embedding provides

- new insights
- means to avoid ill-conditioning
- faster algorithms
- safer algorithms

These benefits arise even for exact computation with exact data !

Analytic treatment of an **algebraic** problem
has been introduced by

C. F. Gauss

Solution of systems of linear equations
by iterative approximation

A Revolutionary Idea :

Have an algebraic problem
whose (exact) solution may be obtained by
a *finite number* of (exact) *rational operations*

Instead take an approximate solution and
improve it iteratively and

Stop when accuracy is sufficient
e.g. relative to *data accuracy*

Developed and used by Gauss
in a large-scale surveying project
Communicated (by letter) to another surveyor
on Dec. 26, 1823

Numerous details of the numerical computation
are touched upon in this communication :

- computation of a good initial approximation
- work with *corrections* not with values
- meaningful accuracy during iteration
- details of choosing the iteration sequence
- stopping criterion

Today , in Scientific Computing ,

- "all" linear systems solved in floating-point
- "almost all" linear systems solved iteratively

where "solved" $\equiv$ solved approximately

Is this meaningful and achievable
also for polynomial problems ?

Fundamental Idea for the Analytic Treatment of Quantitative Algebraic Problems

"quantitative": Besides by their structure,
problem and results are characterized by numbers

Problem : "data" $\quad a \in \mathcal{A} \subseteqq \mathbb{C}^M$

Result : "solution" $\quad z \in \mathcal{Z} \subseteqq \mathbb{C}^m$

Model of Quantitative Problem :

Given: a mapping $\boldsymbol{F}: \ \mathcal{A} \to \mathcal{Z}$ (implicitly!)

Sought: $F(a)$ for specified $a \in \mathrm{dom} F \subset \mathcal{A}$

F is the exact "data→result mapping"

structural aspects of problem are contained in F

definition of F may include values like 1, 0, etc.

must include *structural sparsity*

Note: a coefficient 0 may signify

- an absent term (*sparsity*)
- a tiny value (*approximate data*)

Analytic Properties of $F : \mathcal{A} \to \mathcal{Z}$

$$\dim \mathcal{A} =: M \ , \quad \dim \mathcal{Z} =: m$$

F is **well-posed** in a neighborhood $N(a)$ of a if

domain $F \supset N(a)$, *open* in $\mathcal{A}$

$m \le M$ and F is *surjective* in $N(a)$

F is *continuous* in $N(a)$

otherwise it is **ill-posed**

F is **well-conditioned** in $N(a)$ if

F is *well-posed* in $N(a)$

F is *differentiable* in $N(a)$

$\| F' \|$ is "small"

F is **ill-conditioned** if $\| F' \|$ "large" in $N(a)$

Ill-conditioning implies that the solution z is *highly sensitive* to certain changes in a

Ill-posed Algebraic Problems

Type 1 : dim (domain of F) $=: d < M$

Results only defined for *data on a manifold* $\mathcal{S}$ of dimension d in $\mathcal{A}$ (i.e. only for special data)

Type 2 : dim (image of F) $=: \bar{d} < m$

All exact *results* lie on a *manifold* $\bar{\mathcal{S}}$ of dimension $\bar{d}$ in $\mathcal{Z}$
(i.e. approximate solution is not a proper result)

Type 3 : F is *discontinuous* in $N(a)$

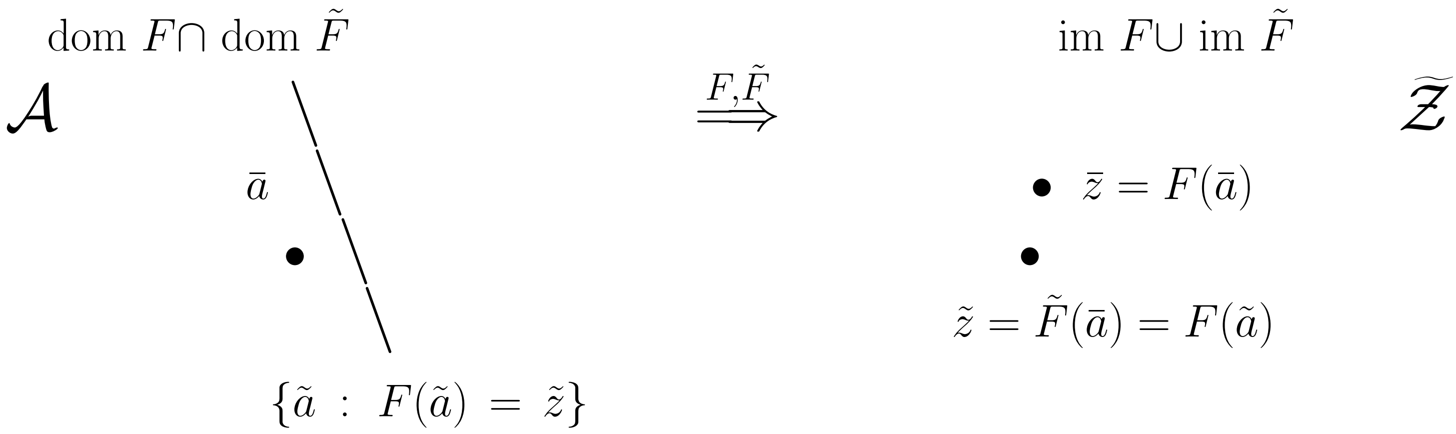
i.e. results "jump" at nearby data

Many classes of interesting quantitative algebraic problems are ill-posed

Meaningful *approaches* to ill-posed problems are an important task of Approximate Algebra

==================================

Approximate Commutative Algebra



Tasks :

- design approximate map $\tilde{F}$
  (algorithm, implementation)
- evaluate $\tilde{F}$ for $\bar{a} \;\Rightarrow\; \tilde{z}$
  (run code)
- find $\tilde{a}$ closest to $\bar{a}$, with $F(\tilde{a}) = \tilde{z}$
  (backward error)
- correct $\tilde{z}$ towards $\bar{z}$
  (successive approximation)
- find analytic properties of F around $\bar{a}$
  (wellposedness, condition, nearsingularity)
- re-define F and/or $\tilde{F}$ if appropriate
  (respect analytic properties of F)

Topologies in $\mathcal{A}$ and $\mathcal{Z}$

Formally, we have

$$\text{data space } \mathcal{A} \equiv \mathbb{C}^M$$

$$\text{result space } \mathcal{Z} \equiv \mathbb{C}^m$$

Note: Structural data are not components of $\mathcal{A}$

May use some **weighted** norm to define distance:

2-norm : $\| \tilde{a} - a \|_{e,2} := \left(\Sigma_\nu \left(\frac{|\tilde{\alpha}_\nu - \alpha_\nu|}{\varepsilon_\nu} \right)^2 \right)^{1/2}$

max-norm : $\| \tilde{a} - a \|_{e,\infty} := \max_\nu \frac{|\tilde{\alpha}_\nu - \alpha_\nu|}{\varepsilon_\nu}$

Weights ε_ν should relate to accuracy levels of α_ν

e.g. $\bar{\alpha}_\nu$ specified to 4 decimals $\Rightarrow$ $\varepsilon_\nu = 10^{-4}$

$$e := (\varepsilon_1, \ldots, \varepsilon_M)$$

Different data components α_ν
may have different tolerances ε_ν

Empirical Data

data from measurements, observations etc. have

limited accuracy

specified data $\bar{a} \quad \in \mathcal{A} \subset \mathbb{R}^M$ or $\mathbb{C}^M$

tolerance $e \quad \in \mathbb{R}_+^M$

validity parameter $\delta \in \mathbb{R}_+$

family of sets of **equivalent data** $(\delta > 0)$:

$$N_\delta(\bar{a}, e) := \{\tilde{a} \in \mathcal{A} : \|\tilde{a} - \bar{a}\|_e \le \delta\}$$

validity scale (*continuous gradual* transition!)

δ: 0 ...	1 ...	3 ...	10 ...	30 ...
valid	probably valid	possibly valid	probably invalid	invalid

Empirical polynomial in $\mathbb{C}^s$:

$$p(x;\, a) := \sum_\nu \alpha_\nu \, x^{j_\nu} \quad \text{with} \quad x^{j_\nu} := x_1^{j_{\nu 1}} \ldots x_s^{j_{\nu s}}$$

where *some* (or all) coefficients α_ν are empirical

Valid Instances of Empirical Data

$\tilde{a} \in \mathcal{A}$ is a valid instance for $(\bar{a}, e)$ if

$$\tilde{a} \in N_\delta(\bar{a}, e) \quad \text{with} \quad \delta = \text{“O(1)”}$$

i.e. $\tilde{a}$ is *indistinguishable* from $\bar{a}$

on the specified accuracy level,

Valid Results for Empirical Data

computational result $\tilde{z} = \tilde{F}(\bar{a}) \in \mathcal{Z}$

interpreted as **exact** result for **modified** data $\tilde{a}$

Equivalent-data manifold $\mathcal{M}(\tilde{z})$:

$$\mathcal{M}(\tilde{z}) := \{ \tilde{a} \in \mathcal{A} : \tilde{a} \xrightarrow{F} \tilde{z} \} \subset \mathcal{A}$$

E.g. Polynomial zeros :

$p(x; \bar{a}) := \bar{\alpha}_0 + \bar{\alpha}_1 x + \ldots + \bar{\alpha}_{M-1} x^{M-1} + x^M$

$\mathcal{M}(\tilde{z}) := \{ \tilde{a} \in \mathcal{A} : p(\tilde{z}; \tilde{a}) = 0 \}$ *linear* (M-1)-manifold

$$\tilde{z} \quad \delta\text{-valid} \quad \Leftrightarrow \quad N_\delta(\bar{a}, e) \cap \mathcal{M}(\tilde{z}) \neq \emptyset$$

Find *minimal* δ : $\tilde{z}$ is **valid** if $\delta_{\min} = O(1)$

Validity Checking

Have to **compute**

$$\delta_{\min}(\tilde{z}) \;:=\; \min\;(\|\tilde{a}-\bar{a}\|_e\,,\;\tilde{a}\in\mathcal{M}(\tilde{z}))$$

For a *linear* manifold $\mathcal{M}(\tilde{z})$, this is a

- linear *least squares* problem with a *2-norm*
- linear *optimization* problem with a *max-norm*

E.g. Polynomial zeros :

Find *nearest* coefficient set $\tilde{a}_{\min}$ with $p(\tilde{z};\ \tilde{a}_{\min}) = 0$

$$\delta_{\min}(\tilde{z}) \;:=\; \|\,\tilde{a}_{\min}-\bar{a}\,\|_e$$

Validity checking permits the following

Algorithmic Scheme

- compute approximate result $\tilde{z}$ *somehow*
  (approximate algorithm, floating-point)
- check validity of $\tilde{z}$

if not valid : refine $\tilde{z}$, go to check

else : **accept** $\tilde{z}$ as **valid result**

Valid Result Sets

Problem with empirical data $(\bar{a};\, e)$:

$$Z_\delta(\bar{a};\, e) \;:=\; \{\, \tilde{z} \in \mathcal{Z} \;:\; \tilde{z} \;\text{ a } \delta\text{-valid result}\,\}$$

size of $Z_\delta(\bar{a};\, e)$ indicates condition of result $F(\bar{a})$

"Pseudospectra" introduced for matrix eigenvalues in 90's

Valid result sets for *distinct* results not meaningful when these results are to be used *jointly*.

Must be *jointly validated* if jointly used $(k \le M)$:

$$\mathcal{M}\left(\{\tilde{z}_1, \ldots, \tilde{z}_k\}\right) \;:=\; \mathcal{M}(\tilde{z}_1) \cap \ldots \cap \mathcal{M}(\tilde{z}_k)$$

E.g. *Several* polynomial zeros $\tilde{z}_\kappa,\ \kappa = 1(1)k$:

$$\mathcal{M}\left(\{\tilde{z}_1, \ldots, \tilde{z}_k\}\right) \;:=\; \{\tilde{a} \in \mathcal{A} \;:\; p(\tilde{z}_\kappa;\, \tilde{a}) = 0\,,\ \kappa = 1(1)k\}$$

$\{\tilde{z}_1, \ldots, \tilde{z}_k\}$ *simultaneously* δ-valid iff

$$N_\delta(\bar{a},\, e) \;\cap\; \mathcal{M}\left(\{\tilde{z}_1, \ldots, \tilde{z}_k\}\right) \;\neq\; \emptyset$$

$$Z_\delta[\{\tilde{z}_1, \ldots, \tilde{z}_k\}] \;\overset{\subset}{\neq}\; Z_\delta[\tilde{z}_1] \cup \ldots \cup Z_\delta[\tilde{z}_k]$$

M pol. zeros: $\mathcal{M}(\{\tilde{z}_1, \ldots, \tilde{z}_M\}) = \{\tilde{a}^*\}$, interpolating pol.

Validity for Ill-Posed Problems

Type 1: $\mathcal{S} :=$ domain $(F) \subset \mathcal{A}$, dim $\mathcal{S} =: d < M$

$$\mathcal{M}(\tilde{z}) \;:=\; \{\, \tilde{a} \in \mathcal{S} :\; F(\tilde{a}) \;=\; \tilde{z} \,\} \;\subset\; \mathcal{S} \;\subset\; \mathcal{A}$$

E.g., *Multiple* zero of $p(x;\, \bar{a})$: discriminant $(p, p') \;=\; 0$
now, refinement becomes (linearized) *optimization*

Type 2: $\bar{\mathcal{S}} :=$ image $(F) \subset \mathcal{Z}$, dim $\bar{\mathcal{S}} =: \bar{d} < m$

$\tilde{z} \notin \bar{\mathcal{S}}$ is *not a result*, $\mathcal{M}(\tilde{z})$ is *not defined*

Must check validity of $\tilde{z}$ in further use
Refine $\tilde{z}$ towards $F(\bar{a})$ on $\bar{\mathcal{S}}$

E.g., numerically computed Groebner basis polynomials do *not* satisfy syzygies , may refine towards "exact" syzygies

Type 3: F discontinuous, e.g. discrete

$\tilde{z}$ valid if $\mathcal{M}(\tilde{z}) \;\supset\; N_\delta(\bar{a},\, e)$, $\delta = \mathrm{O}(1)$

else: Check validity of "more degenerate" results

E.g., rank of empirical matrix $=$ *lowest* valid rank

Basic Principles

interpret quantitative problem as a

data $\longrightarrow$ result map $F : \mathcal{A} \rightarrow \mathcal{Z}$

Design approximate map $\tilde{F}$

Evaluate $\tilde{F}$ for specified data

obtain approximate result $\tilde{z}$

Compute backward error of $\tilde{z}$

accept or refine $\tilde{z}$

Note :

Want exact results (except for round-off)

for valid algebraic problems

i.e. for data very close to specified ones

Conclusions

Approximate Commutative Algebra

is meaningful and feasible

in the context of

quantitative problems

with coefficients with a natural topology

it is indispensible for

empirical problems