AD viewed from Algebraic Complexity

Sebastian Heinz
Institut für Mathematik
Humboldt–Universität zu Berlin

Outline

- object of interest
- programs in algebraic complexity theory
- families of linear mappings
- connections to AD

Object of interest

What is given?

$F : R^n \longrightarrow \mathbb{R}^m$ continuously differentiable function

- $\mathscr{U}_{x_0} \subseteq \mathbb{R}^n$ open neighborhood of a fixed point $x_0 \in \mathbb{R}^n$

What is to compute?

directional derivatives of $F \simeq$ linearized model $\tilde{F}$ of F

- $\tilde{F} : \mathscr{U}_{x_0} \longrightarrow \mathbb{R}^m$ affine function
- $\tilde{F}(x) = F(x_0) + DF(x_0)(x - x_0)$ for $x \in \mathscr{U}_{x_0}$

What are "best" representations of $\tilde{F}$ or $DF(x_0)$, respectively?

Example (1)

Given problem

$\boxed{A \cdot y = b \text{ linear system}}$

- $\left.\begin{array}{l} A : \mathscr{U}_{x_0} \longrightarrow \mathbb{R}^{m \times m} \\ b : \mathscr{U}_{x_0} \longrightarrow \mathbb{R}^m \end{array}\right\}$ continuously differentiable functions
- $A(\mathscr{U}_{x_0}) \subseteq \mathbb{R}^{m \times m}$ family of invertable matrices

Making F fit to the problem

- for $x \in \mathscr{U}_{x_0}$ set $F(x) := y$ such that $A(x) \cdot y = b(x)$

Example (1)

Evaluation of F at the point x_0

- $A(x_0)$ and $b(x_0)$ Ops: ???
- L-U-factorization of $A(x_0)$ Ops: $O(m^3)$
- solution vector y Ops: $O(m^2)$

Remarks

- F given by a program of elementary operations/functions (Ops)
- no use of fast algorithms

Example (1)

Using AD (forward mode)

- directional derivatives of F Ops: $O(m^3)+???$

Using AD + preaccumulation

- directional derivatives of F Ops: $O(m \cdot n)$

Cases

- preaccumulation pays off if $n \ll m^2$
- preaccumulation (probably) useless if $n \gg m^2$

What can we do if $n \approx m^2$ or $n \gg m^2$?

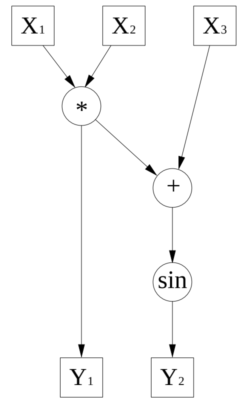
Programs in Algebraic Complexity Theory

Arithmetic circuit $\stackrel{def}{=}$
(DAG) of input-vertices, inner vertices ($+$, $-$, $*$), output-vertices

Remarks

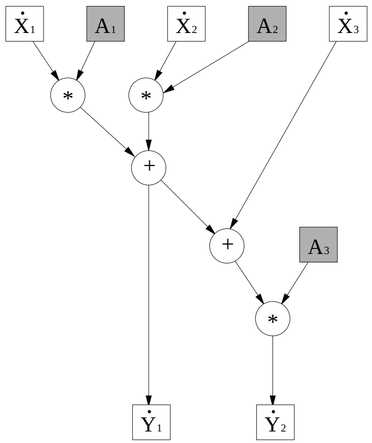
- no branching
- implementation as a straight-line program
- encoding length $\approx$ computation time $\approx$ number of vertices
- univariate functions (sin, cos, inverse,...) allowed

Example (2)



representation of $F(x_1,x_2,x_3) = (x_1 \cdot x_2, \sin(x_1 \cdot x_2 + x_3))$

Example (2)



”generalized” representation of *DF*

”Structure-preserving” elimination

Edge-, vertex-, face-elimination

- $J(\mathbb{R}^p) \subseteq_* \tilde{J}(\mathbb{R}^q)$

”Structure-preserving” elimination

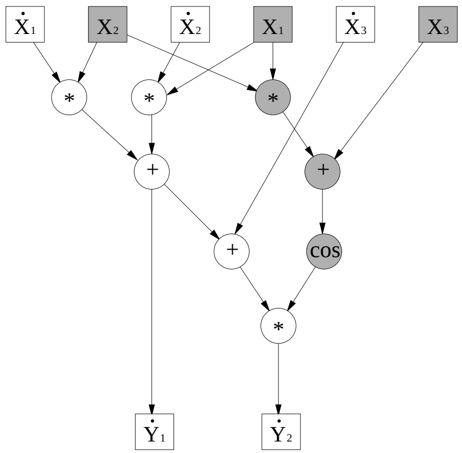
- $J(\mathbb{R}^p) =_* \tilde{J}(\mathbb{R}^q)$

Remark

- ”$\subseteq_*$” and ”$=_*$”: outside a hypersurface in $\mathbb{R}^{m\times m}$

A. Griewank (2003); A. Griewank, O. Vogel(2003?)

Example (2)



representation of $DF(x_1, x_2, x_3)$

Families of linear mappings

Construction

- new variables $a_1, \ldots, a_p$ for the partial derivatives
- $a_1, \ldots, a_p$ algebraically independent
- $J : \mathbb{R}^p \longrightarrow \mathbb{R}^{m\times n}$ polynomial mapping

- **What can we say about the set $J(\mathbb{R}^p) \subseteq \mathbb{R}^{m\times n}$?**
- **Which properties make a ”short” representation of *DF* possible?**

A. Griewank, J. Riehme, T. Steihaug, . . .

References

P. Bürgisser, M. Clausen, M. A. Shokrollahi. *Algebraic complexity theory*, volume 315 of *Grundlehren der mathematischen Wissenschaften*. Springer-Verlag, 1997. With the Collaboration of Thomas Lickteig.

J. von zur Gathen, J. Gerhard. *Modern computer algebra*. Cambridge University Press, 1999.

Connections to AD

Minimal representation of $J(\mathbb{R}^p)$

- ordering of DAGs that represent $J(\mathbb{R}^p)$ nearly everywhere
- How can we compute it?

Dimension of $J(\mathbb{R}^p)$ (in a generic point)

- ”scarcity”
- lower bound of the minimal encoding length of *DF*
- Can we reach that bound?

Idea

- study the set $J^{-1}(Y)$ for a generic point $Y \in J(\mathbb{R}^p)$
- properties of $J^{-1}(Y) \longleftrightarrow$ minimal encoding length of *DF*