On best C-approximation

Shmuel Friedland
Univ. Illinois at Chicago

SIAM 2013 - Fort Collins, Co
August 1, 2013
Mostly joint results with G. Ottaviani and M. Stawiska in progress

Overview

- Statement of the problem
- Uniqueness of best approximation
- Number of critical points
- Best border rank k-approximation to symmetric tensors
- Best approximation of vectors fixed under the action of a finite group

Statement of the problem

$\nu : \mathbb{R}^n \to [0, \infty)$ a norm on $\mathbb{R}^n$

$C \subset \mathbb{R}^n$ a closed subset

Problem: approximate a given vector $\mathbf{x} \in \mathbb{R}^n$ by a point $\mathbf{y} \in C$:

$\mathrm{dist}_\nu(\mathbf{x}, C) := \min\{\nu(\mathbf{x} - \mathbf{y}),\ \mathbf{y} \in C\}$

$\mathbf{y}^\star \in C$ is called a best ν-(C)approximation of $\mathbf{x}$ if
$\nu(\mathbf{x} - \mathbf{y}^\star) = \mathrm{dist}_\nu(\mathbf{x}, C)$

$\|\cdot\|$ the Euclidean norm on $\mathbb{R}^n$, $\mathrm{dist}(\mathbf{x}, C) = \mathrm{dist}_{\|\cdot\|}(\mathbf{x}, C)$.

We call a best $\|\cdot\|$-approximation briefly a best (C)-approximation

Example: $\mathbb{R}^n = \mathbb{R}^{m_1 \times \ldots \times m_d}$ - d-mode tensors and C:

1. Tensors of border rank k-at most, denoted as C_k

2. Tensors whose unfolded matrix in mode i has rank $r_i (\le m_i)$ at most

for $i = 1, \ldots, d$ - Best $\mathbf{r} := (r_1, \ldots, r_d)$-approximation, denoted as $C(\mathbf{r})$

Uniqueness of best ν-approximation

Dual norm $\nu^*(\mathbf{x}) := \max\{\mathbf{y}^\top\mathbf{x}, \nu(\mathbf{y}) = 1\}$

Thm: Assume that ν and ν^* are differentiable, (i.e. differ. on $\mathbb{R}^n \setminus \{\mathbf{0}\}$)

Then for a.a. $\mathbf{x} \in \mathbb{R}^n \setminus C$ a best ν-approximation is unique

Outline of prf: $|\mathrm{dist}_\nu(\mathbf{x}, C) - \mathrm{dist}_\nu t(\mathbf{z}, C)| \le \nu(\mathbf{x} - \mathbf{z})$ (triangle inequality)

$\nu(\cdot)$ is Lipschitz $\Rightarrow \mathrm{dist}(\cdot, C)$ Lipschitz

Rademacher's thm: $\mathrm{dist}_\nu(\cdot, C)$ is differentiable a.a.

Assume that $\mathrm{dist}_\nu(\cdot, C)$ is differentiable at $\mathbf{x} \in \mathbb{R}^n \setminus C$

$\mathrm{dist}_\nu(\mathbf{x}, C) = \nu(\mathbf{x} - \mathbf{y}) \Rightarrow \mathbf{u}_1 := D\mathrm{dist}(\mathbf{x}, C)$ is supporting hyperplane

$\{\mathbf{z} \in \mathbb{R}^n,\ \nu(\mathbf{z}) \le \nu(\mathbf{x} - \mathbf{y})\}$ at $\mathbf{x} - \mathbf{y} \Rightarrow \nu^*(\mathbf{u}) = 1, \mathbf{u} := \frac{1}{\nu(\mathbf{x}-\mathbf{y})}\mathbf{u}_1$

$\mathbf{v} := \frac{1}{\nu(\mathbf{x}-\mathbf{y})}(\mathbf{x} - \mathbf{y})$ supporting hyperplane of $\nu^*(\mathbf{z}) \le 1$ at $\mathbf{u}$

ν^*differentiable $\Rightarrow \mathbf{v}$ unique. Hence $\mathbf{x} - \mathbf{y}$ is unique, i.e. $\mathbf{y}$ unique

Semi-algebraic norms

$S \subset \mathbb{R}^n$ is semi-algebraic if it is a finite union of basic semi-algebraic sets given by a finite number of polynomial equalities and inequalities

$p_i(\mathbf{x}) = 0,\ i \in \{1, ..., \lambda\}, q_j(\mathbf{x}) > 0,\ j \in \{1, ..., \lambda'\}$

$\dim S$ is the maximal dimension of the basic set

Semi-algebraic set $S \subset \mathbb{R}^n$ can be described by a quantifier-free first order formula. The class of semi-algebraic sets is closed under finite unions, finite intersections, complements and projections.

$f : \mathbb{R}^n \to \mathbb{R}$ semi-algebraic if $G(f) = \{(\mathbf{x}, f(\mathbf{x})) : x \in \mathbb{R}^n\}$ semi-algebraic

ν semi-algebraic if $\nu(\cdot)$ semi-algebraic

$\|(x_1, \ldots, x_n)^\top\|_a := (\sum_{i=1}^n |x_i|^a)^{\frac{1}{a}}, a \geq 1$

is semi-algebraic if $a = \frac{b}{c}$ is rational as $G(\|\cdot\|_a) =$

$\{(x_1, \ldots, x_n, t)^\top : x_i = \pm y_i^c, y_i \geq 0, t = s^c, s \geq 0, \sum_{i=1}^n y_i^b - s^b = 0\}$

Uniqueness of ν-approxim. in semi-algebraic setting

Lemma: $C \subset \mathbb{R}^n$ closed semi-algebraic set, ν a semi-algebraic

Then the function $\mathrm{dist}_\nu(\cdot, C)$ is semi-algebraic.

Thm : Let $C \subset \mathbb{R}^n$ semi-algebraic, ν semi-algebraic norm, ν and ν^* are differentiable. Then the set of all points $\mathbf{x} \in \mathbb{R}^n \setminus C$, denoted by $S(C)$, where ν-approximation to $\mathbf{x}$ in C is not unique is a semi-algebraic set which does not contain an open set. In particular $S(C)$ is contained in some hypersurface $H \subset \mathbb{R}^n$.

Prf: $f(\mathbf{x}) = \mathrm{dist}_\nu(\mathbf{x}, C)$, $G(f) \times C$

$T(f) := \{(\mathbf{x}, t, \mathbf{y}, \mathbf{x}, t, \mathbf{z}) \in \mathbb{R}^{2(2n+1)}, (\mathbf{x}, t, \mathbf{y},), (\mathbf{x}, t, \mathbf{z},) \in G(f) \times C,$

$\nu(\mathbf{x} - \mathbf{y}) = \nu(\mathbf{x} - \mathbf{z}) = t, \ \|\mathbf{y} - \mathbf{z}\|^2 > 0\}$

$\pi_n(T(f)) = S(C)$ - hence semi-algebraic

Rademacher's or Durfee's theorem yield $\dim S(C) < n$.

C an irreducible variety

C is the zero set of polynomials $p_i(\mathbf{x}), \mathbf{x} \in \mathbb{R}^n, i = 1, \ldots, m$.

$V := C_{\mathbb{C}}$ is the zero set of $p_i(\mathbf{z}), \mathbf{z} \in \mathbb{C}^n, i = 1, \ldots, m$.

$D(V)(\mathbf{y}) = ([\frac{\partial p_i}{\partial x_j}(\mathbf{y})]_{i=j=1}^{m,n})^\top \in \mathbb{C}^{n \times m}$, $\mathbf{x} = (x_1, \ldots, x_n)^\top \in \mathbb{C}^n$.

For $\mathbf{y} \in V$, $D(V)(\mathbf{y})$-Jacobian of V at $\mathbf{y}$, $D(V)(\mathbf{y})$ real for $\mathbf{y} \in C$.

C, V is of real and complex dimension d respectively if

$\mathrm{rank}\, D(V)(\mathbf{y}) = n - d$ for most $\mathbf{y} \in V$. (Such $\mathbf{y}$ is a smooth point of V.)

For a smooth $\mathbf{y} \in V$ the null space of $D(V)(\mathbf{y})^\top$ is tangent hyperplane

and the column space of $D(V)(\mathbf{y})$ is cotangent hyperplane

Fix $\mathbf{x} \in \mathbb{R}^n$ and let $g_{\mathbf{x}}(\mathbf{u}) := \|\mathbf{x} - \mathbf{u}\|^2$ be a function on C.

smooth $\mathbf{y} \in C$ is a critical point of $g_{\mathbf{x}}$ iff $\mathbf{x} - \mathbf{y} \in \mathrm{Range}(D(V)(\mathbf{y}))$

i.e. all $n - d + 1$ minors of $[D(V)(\mathbf{y})\ (\mathbf{x} - \mathbf{y})] \in \mathbb{R}^{n \times (m+1)}$ vanish

The critical variety

$\Sigma_1(V) \subset \mathbb{C}^n \times V$ all $(\mathbf{x}, \mathbf{y})$ such that all $n - d + 1$ minors of

$[D(V)(\mathbf{y})\ (\mathbf{x} - \mathbf{y})]$ vanish

$\Sigma_1(V)$ a variety that contains $\mathbb{C}^n \times \mathrm{Sing}\ V$, ($\dim C^n \times \mathrm{Sing}\ V \geq n$)

$\Sigma_0(V) := \Sigma_1(V) \setminus \mathbb{C}^n \times \mathrm{Sing}\ V$ all points

$(\mathbf{y} + \mathbf{z}, \mathbf{y}), \mathbf{y} \in V \setminus \mathrm{Sing}\ (V), \mathbf{z} \in \mathrm{Range} D(V)(\mathbf{y})$

So $\Sigma_0(V)$ vector bundle over $\mathbf{V} \setminus \mathrm{Sing}\ V$ and quasi-variety

Critical variety:$\Sigma(V) = \mathrm{Clos}(\Sigma_0(V))$,an irreducible component of $\Sigma_1(V)$

$\dim \Sigma(V) = n$ and $\pi_n : \Sigma(V) \to \mathbb{C}^n$ dominating of degree $\delta(V)$:

$\exists$ subvariety $U \subset \mathbb{C}^n$ s.t. $|\pi_n^{-1}(\mathbf{x})| = \delta(V)$ for each $\mathbf{x} \in \mathbb{C}^n \setminus U$

i.e. each $\mathbf{x} \in \mathbb{C}^n \setminus U$ has $\delta(V)$ smooth critical points in V

Friedland-Ottaviani computed $\delta(V)$ for

$\mathbb{R}^n = \mathbb{R}^{m_1 \times \ldots \times m_d}$ and C: tensors of rank 1 and zero tensor

Smooth and singular C

Assume that C smooth, ($V = C_{\mathbb{C}}$ may be singular). Then

$\mathrm{dist}(\mathbf{x}, C) = \min\{\|\mathbf{x}-\mathbf{y}\|, (\mathbf{x},\mathbf{y}) \in \Sigma(V) \cap (\mathbb{R}^n \times C)\}$ (1)

Suppose C contains singular points

$\mathrm{Sing}\,(C) = \cup_{i=1}^{p} C_i$, each $C_i \subset \mathbb{R}^n$ is irreducible

$\mathrm{dist}(\mathbf{x}, C) \le \mathrm{dist}(\mathbf{x}, C_i)$ and $\mathrm{dist}(\mathbf{x}, C) = \mathrm{dist}(\mathbf{x}, C_i)$ on semi-algebraic Q_i

Suppose that (2): $\dim Q_i < n$, $i = 1, \ldots, p$ holds. Then (1) holds

(2) holds for $\mathbb{R}^n = \mathbb{R}^{m_1 \times m_2}$ and C matrices of rank k at most

Probably (2) holds for tensors?

Assume (2) does not hold.

$C = \cup_{i=0}^{q}(C_i \setminus \mathrm{Sing}\ C_i)$ a smooth stratification of C ($C_0 := C$)

$V_i := (C_i)_{\mathbb{C}}$ Then (1) holds if we replace $\Sigma(V)$ by $\cup_{i=0}^{q}\Sigma(V_i)$,

I.e., for $\mathbf{x} \in \mathbb{R}^n \setminus (\cup U_i)$ $\mathrm{dist}(\mathbf{x}, C)$ is the minimal distance of

$\mathbf{x}$ to k ($\le \sum \delta(V_i)$) points on C

Approximation of symmetric tensors: rank one at most

$m^{\times d} := (\underbrace{m, \ldots, m}_{d})$, $\mathbb{R}^n = \mathbb{R}^{m^{\times d}}$, $\mathrm{S}(d, \mathbb{R}^n)$ symmetric tensors

$C \subset \mathbb{R}^n$ either C_k or $C(\mathbf{r})$

Thm There exists a semi-algebraic set $Q \subset \mathrm{S}(d, \mathbb{R}^m), \dim Q < \binom{m+d-1}{d}$

for $\mathcal{T} \in \mathrm{S}(d, \mathbb{R}^n) \setminus Q$ best rank 1-approximation unique, and symmetric

Prf. 1. Banach 1939, Chen-He-Li-Zhang 2012, Friedland 2013:

best rank 1-approxim. of symmetric tensor can be chosen symmetric

2. Friedland-Ottaviani: $f := \mathrm{dist}(\cdot, C_1)|\mathrm{S}(d, \mathbb{R}^m)$. If f differentiable at $\mathcal{T}$

then best rank 1-approximation unique up to permutation of factors in

$\mathcal{X} = \mathbf{x}_1 \otimes \ldots \otimes \mathbf{x}_d$. Use 1. to deduce $\mathcal{X}$ symmetric

3. Friedland-Stawiska: the set Q of symmetric tensor with

not unique best rank approximation is semi-algebraic

Approximation of symmetric tensors: b. rank k at most

$N(m,d) = \frac{1}{2}(\binom{m+d-3}{d-2} + 2m - 2)$ for $d \geq 3$, $\quad (N(m,3) = \frac{3m-2}{2})$

Thm For $d \geq 3, 2 \leq k \leq N(m,d)$ the semi-algebraic set of all symmetric tensors for which best border rank k approximation is unique, (denoted as $P_k \subset \mathbb{R}^n$), has dimension $\binom{m+d-1}{d}$.

Use Kruskal's theorem to show that a symmetric tensor of the form $\mathcal{T} = \sum_{i=1}^{k} \otimes^d \mathbf{u}_i$, k-as above

has rank k if any $\min(m,k)$ vectors from $\mathbf{u}_1, \ldots, \mathbf{u}_k$ and $\min(k, \binom{m+d-3}{d-2})$ vectors from $\otimes^{d-2}\mathbf{u}_1, \ldots, \otimes^{d-2}\mathbf{u}_k$ linearly independent

Problem : Is $\dim(\mathbb{R}^n \setminus P_k) < \binom{m+d-1}{d}$?

Weaker problem: Is the best border rank k-approximation to a symmetric tensor can be chosen symmetric?

Approximation of vectors in an invariant subspace

$C \subset \mathbb{R}^n$ irreducible variety, $\mathcal{G} \subset \mathbb{R}^{n \times n}$ a finite nontrivial group of matrices acting on $\mathbb{R}^n$: $\mathbf{x} \mapsto A\mathbf{x}$,

Assume:

1. $\mathbf{U} \subset \mathbb{C}^n$ the subspace of all fixed points of $\mathcal{G}$ $\dim \mathbf{U} = d \in [1, n-1]$.
2. C is fixed by $\mathcal{G}$. $C_{\mathbf{U}} := C_{\mathbb{C}} \cap \mathbf{U}$ contains a smooth point of $V = C_{\mathbb{C}}$.

Let $\mathbf{y} \in C_{\mathbf{U}}$ is a smooth point of V. $\exists O(\mathbf{y}) \subset \mathbb{C}^n$ neighborhood of $\mathbf{y}$ s.t. for $\mathbf{x} \in O(\mathbf{y}) \cap \mathbf{U}$, the unique critical point $\mathbf{y}(\mathbf{x}) \in O(\mathbf{y})$ is in $\mathbf{U}$.

If $\mathbf{y}$ a smooth point of C then for each $\mathbf{x} \in O(\mathbf{y}) \cap \mathbb{R}^n$ the best C approximation is in $\mathbf{U}_{\mathbb{R}} := \mathbf{U} \cap \mathbb{R}^n$.

A sufficient condition for best approx. to be in $\mathbf{U}_{\mathbb{R}}$

Define $\Phi := (\mathbf{U} \times V) \cap \Sigma(V)$.

Φ contains irreducible components $\Phi_1, \ldots \Phi_l$, $\dim \Phi_i = d$,

s.t. $(\mathbf{x}, \mathbf{y}(x))$, $\mathbf{x} \in O(\mathbf{y})$ as above, is in some Φ_i.

$\Phi_i \subset \mathbf{U} \times C_{\mathbf{U}}$, $i = 1, \ldots, l$

Let $\Phi = \Phi' \cup (\cup_{i=1}^{l} \Phi_i)$.

Suppose that $\pi_n(\Phi' \cap (\mathbb{R}^n \times C))$ is not dense in $\mathbf{U}_{\mathbb{R}}$.

Then the best approximation to $\mathbf{x} \in \mathbf{U}_{\mathbb{R}}$ can be chosen in $\mathbf{U}_{\mathbb{R}}$.

This is true for symmetric matrices where $C = C_k$.

References 1

- S. Banach, Über homogene polynome in (L^2), *Studia Math.* 7 (1938), 36–44.
- B. Chen, S. He, Z. Li, and S, Zhang, Maximum block improvement and polynomial optimization, *SIAM J. Optimization*, 22 (2012), 87–107
- P. Comon, G. Golub, L.-H. Lim, and B. Mourrain, Symmetric tensors and symmetric tensor rank, *SIAM Journal on Matrix Analysis and Applications*, 30 (2008), 1254-1279.
- A. H. Durfee, Neighborhoods of algebraic sets, *Transactions of the American Mathematical Society*, 276 (1983), 517–530
- S. Friedland, Best rank one approximation of real symmetric tensors can be chosen symmetric, *Front. Math. China*, 8 (1) (2013), 19–40.

References 2

- S. Friedland and G. Ottaviani, The number of singular vector tuples and uniqueness of best rank one approximation of tensors, arXiv:1210.8316 .
- L.-H. Lim. Singular values and eigenvalues of tensors: a variational approach. *Proc. IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing* (CAMSAP '05), 1 (2005), 129-132.