Alternation conditions for multivariate approximation

Julien Ugon

WoMBaT

24 November 2016

Introduction

We are interested in the problem of *finding the best uniform approximation of a continuous function f by a polynomial of degree at most n*:

$$\text{minimise } \|p - f\|_\infty \text{ subject to } p \in \Pi_n[x_1, \ldots, x_d],$$

where

- Π_n is the set of polynomials of the variables $x_1, \ldots, x_d$ of degree at most n.
- $f : \mathbb{R}^d \to \mathbb{R}$ and bounded.

Preliminary considerations: the good

- Π_n is a linear vector space: $\forall p \in \Pi_n$

$$p(x_1, \ldots, x_d) = \sum_{i \in I} a_{\mathbf{i}} x^{\mathbf{i}}$$

where $\mathbf{i} = (i_1, \ldots, i_d)$, $I = \{\mathbf{i} : i_1 + \ldots + i_d \leq n\|$ and $x^{\mathbf{i}} = x_1^{i_1} x_2^{i_2} \ldots x_d^{i_d}$.

- The objective function of the problem is proper convex.

Preliminary considerations: the bad

- dimension of the problem can be very large: $\binom{d+n}{d}$
- objective function is not differentiable
- problem is not separable

It is not possible to solve this problem except for very low values of n or d.

Question

How can we address the curse of dimensionality for this problem?

Question

How can we address the curse of dimensionality for this problem?

Can we replace that question with another question?

Question

How can we address the curse of dimensionality for this problem?

Can we replace that question with another question? Yes!

Theorem (Chebyshev)

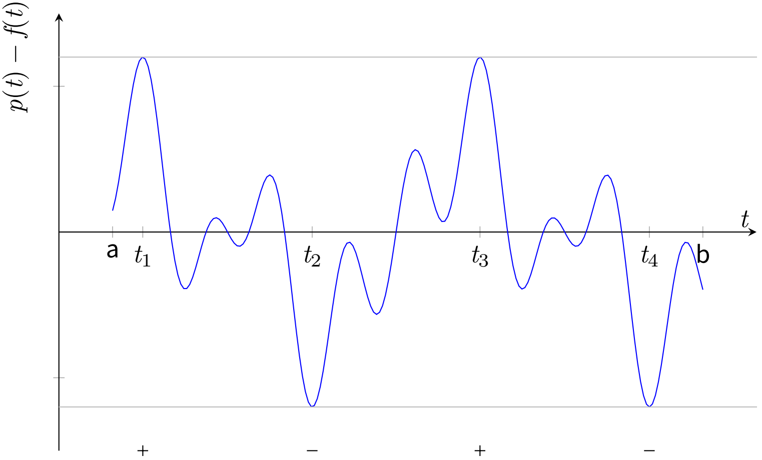
A polynomial $p \in \Pi_n[x]$ is a best uniform approximation of a continuous function $f : \mathbb{R} \to \mathbb{R}$ on an interval $[a, b]$ if and only if there exists $n + 2$ points $a \le t_1 < \ldots < t_{n+2} \le b$ and a number $\sigma \in \{-1, 1\}$ such that $f(t_i) - p(t_i) = \sigma^i \|f - p\|_\infty$.

Definition (Extreme point)

- Points where the maximal deviation is attained are extreme points. We denote them $E^+(p)$ and $E^-(p)$.
- Points $t_1, \ldots, t_n$ for an alternating sequence of extreme points.

Existing results

Case $d = 1$



Theorem

A linear function is a best linear approximation of a function $f\colon \mathbb{R}^d \to \mathbb{R}$ *if and only if* $\operatorname{co} E^+(p) \cap \operatorname{co} E^-(p) \neq \emptyset$.

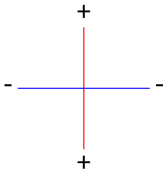
Examples with $d = 2$:

+

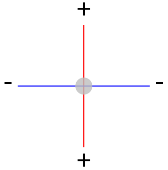
- -

+

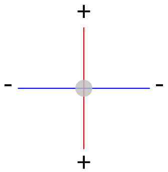
Examples with $d = 2$:

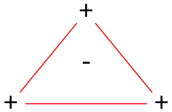


Examples with $d = 2$:

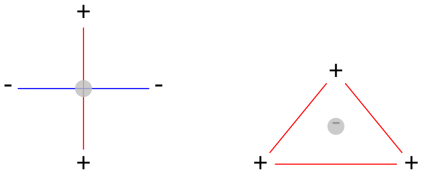


Examples with $d = 2$:

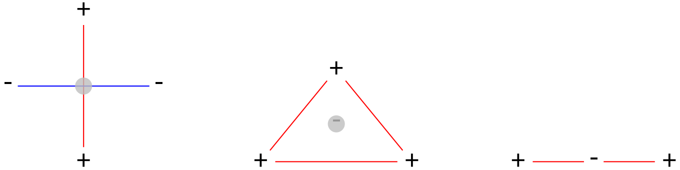




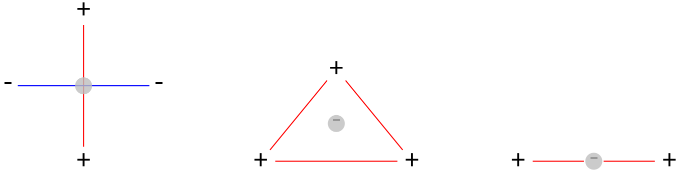
Examples with $d = 2$:



Examples with $d = 2$:



Examples with $d = 2$:



Some observations

- Two different spaces: $\mathbb{R}^d$ and $\mathbb{R}^{\binom{d+n}{d}}$

Some observations

- Two different spaces: $\mathbb{R}^d$ and $\mathbb{R}^{\binom{d+n}{d}} = \begin{cases} \mathbb{R}^{n+1} & \text{if } d = 1 \\ \mathbb{R}^{d+1} & \text{if } n = 1 \end{cases}$

Some observations

- Two different spaces: $\mathbb{R}^d$ and $\mathbb{R}^{\binom{d+n}{d}} = \begin{cases} \mathbb{R}^{n+1} & \text{if } d = 1 \\ \mathbb{R}^{d+1} & \text{if } n = 1 \end{cases}$
- Characterisations are geometrical
- In terms of extreme points

Questions

- Can we provide a geometrical characterisation of best approximants for general degree n and dimension d?
- What is the relationship of these characterisations with the notion of alternation?

Characterising solutions

Define by $M^n(\mathbf{x})$ the set of all monomials of degree at most n of $\mathbf{x}$. If $d = 2$,

$$M^2(\mathbf{x}) = \begin{pmatrix} 1 \\ x_1 \\ x_2 \\ x_1^2 \\ x_1 x_2 \\ x_2^2 \end{pmatrix},$$

Remark

$$M^1(\mathbf{x}) = \mathbf{x}$$

Characterising solutions

Theorem

A linear function is a best linear approximation of a function $f : \mathbb{R}^d \to \mathbb{R}$ *if and only if*

$$\mathrm{co}\{M^n(\mathbf{x}) : \mathbf{x} \in E^+(p)\} \cap \mathrm{co}\{M^n(\mathbf{x}) : \mathbf{x} \in E^-(p)\} \neq \emptyset$$

Proof.

Convex analysis/subdifferential calculus □

Reducing the dimension

Theorem

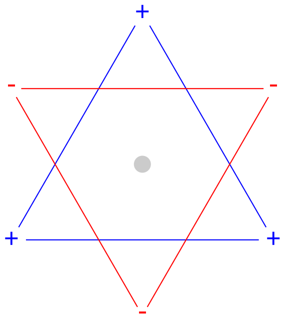
Given a hyperplane $\mathcal{H}$ separating two half-spaces $\mathcal{H}^+$ and $\mathcal{H}^-$. If p is a best uniform approximation in $\Pi^n[\mathbf{x}]$, then

$$\operatorname{co}\{M^{n-1}(\mathbf{x}) : \mathbf{x} \in E^+_{\mathcal{H}}(p)\} \cap \operatorname{co}\{M^{n-1}(\mathbf{x}) : \mathbf{x} \in E^-_{\mathcal{H}}(p)\} \neq \emptyset$$

$$E^+_{\mathcal{H}}(p) = E^+(p) \cap \mathcal{H}^+ \cup E^-(p) \cap \mathcal{H}^-$$
$$E^-_{\mathcal{H}}(p) = E^-(p) \cap \mathcal{H}^+ \cup E^+(p) \cap \mathcal{H}^-$$

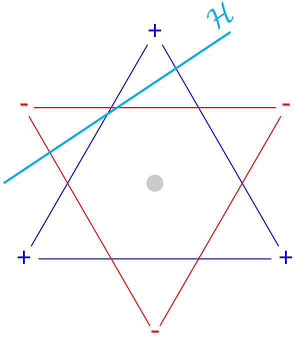
Reducing the dimension

Example for $n = 2$ and $d = 2$



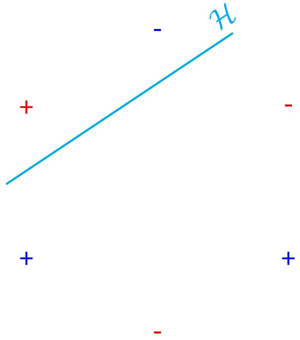
Reducing the dimension

Example for $n = 2$ and $d = 2$



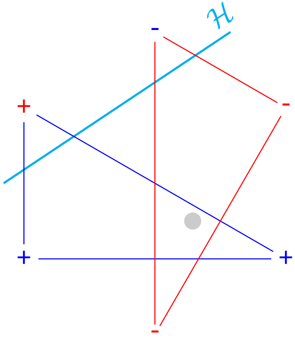
Reducing the dimension

Example for $n = 2$ and $d = 2$



Reducing the dimension

Example for $n = 2$ and $d = 2$



Reducing the dimension: sufficiency

Theorem

If

$$\mathrm{co}\{M^{n-1}(\mathbf{x}) : \mathbf{x} \in E_{\mathcal{H}}^{+}(p)\} \cap \mathrm{co}\{M^{n-1}(\mathbf{x}) : \mathbf{x} \in E_{\mathcal{H}}^{-}(p)\} \neq \emptyset$$

for any hyperplane $\mathcal{H}$ then p is a best approximation.

Relation to alternation

\+ - + - + -

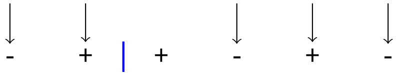
Relation to alternation

+ - | + - + -

Relation to alternation

- + | + - + -

Relation to alternation



Relation to alternation

Reducing the dimension

Theorem

Select d extreme points. If p is a best uniform approximation in $\Pi^n[\mathbf{x}]$, then

$$\mathrm{co}\{M^{n-1}(\mathbf{x}) : \mathbf{x} \in E^{+}_{\mathcal{H}}(p)\} \cap \mathrm{co}\{M^{n-1}(\mathbf{x}) : \mathbf{x} \in E^{-}_{\mathcal{H}}(p)\} \neq \emptyset$$

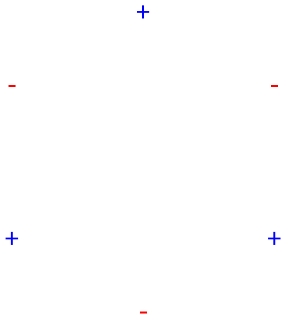
where $\mathcal{H}$ is a hyperplane containing these points. Note that these d points are not in $E^{+}_{\mathcal{H}}(p)$ or in in $E^{+}_{\mathcal{H}}(p)$.

Theorem

If the above result is true for any d extreme points, then p is a best approximation.

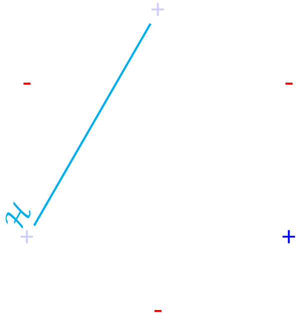
Reducing the dimension

Example for $n = 2$ and $d = 2$



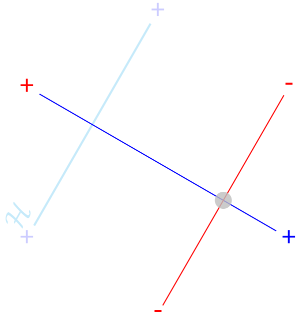
Reducing the dimension

Example for $n = 2$ and $d = 2$

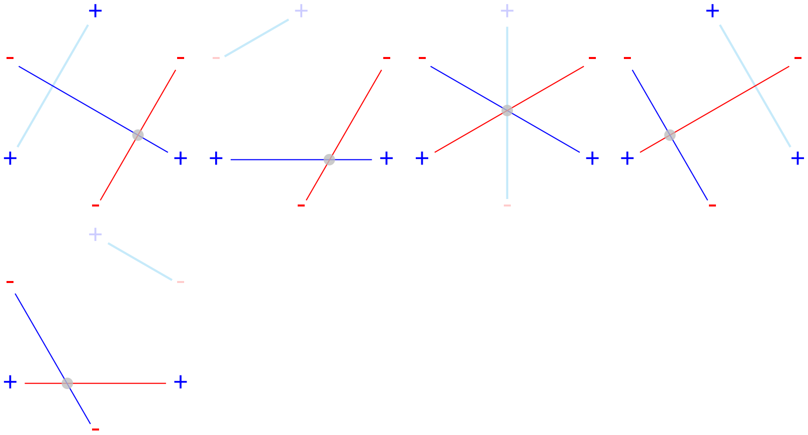


Reducing the dimension

Example for $n = 2$ and $d = 2$

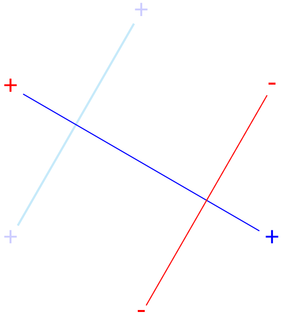


A “geometric” characterisation



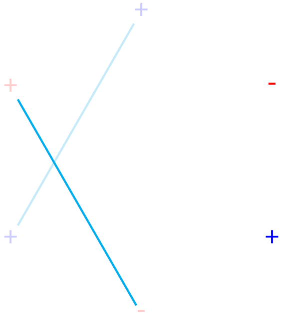
Reducing the dimension

Example for $n = 3$ and $d = 2$



Reducing the dimension

Example for $n = 3$ and $d = 2$



A conceptual algorithm...

Conceptual algorithm

repeat $n-1$ *times*

Step 1. Select d points and a hyperplane $\mathcal{H}$ containing them

Step 2. Update the signs of the remaining points according to $\mathcal{H}$

We need to apply this algorithm $\frac{((n-1)d)!}{(d!)^{n-2}}\binom{N}{(n-1)d}$ times.

Tentative definition of alternation...

Definition (alternating sequence)

A set of k points in $\mathbb{R}$ alternates if by removing any point from the set one obtains an alternating set of $k-1$ points by inverting the signs of the points on one side of the removed point.

Tentative definition of alternation...

Definition (alternating sequence)

A set of k points in $\mathbb{R}$ alternates if by removing any point from the set one obtains an alternating set of $k-1$ points by inverting the signs of the points on one side of the removed point.

Definition (Generalised alternation)

A set of k points in $\mathbb{R}^d$ alternate if by removing any d point from the set one obtains a set of $k-d$ points by inverting the signs of the points on one side of (a hyperplane containing) the removed point.

Future research

- We have an algorithm, but it is not practical.
- Do we need to consider all possible combinations of d points?
- Could it be that we only need to consider facets of the polytope formed by the extreme points? (Or maybe find a counter-example?)
- How does this compare against checking a simple intersection in a (much larger) space?