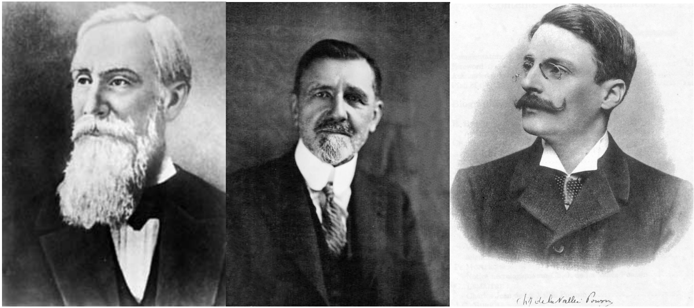
Approximation through Convex Analysis

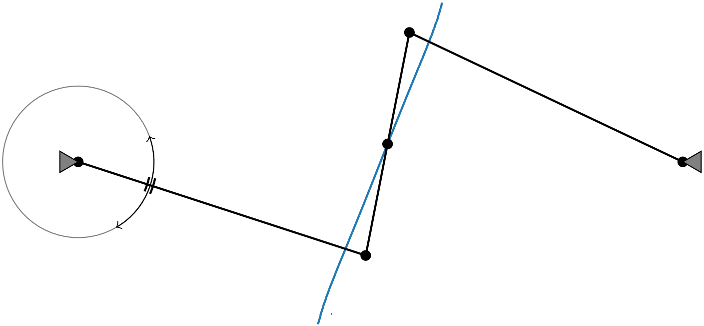
Julien Ugon   Nadia Sukhorukova

Introduction

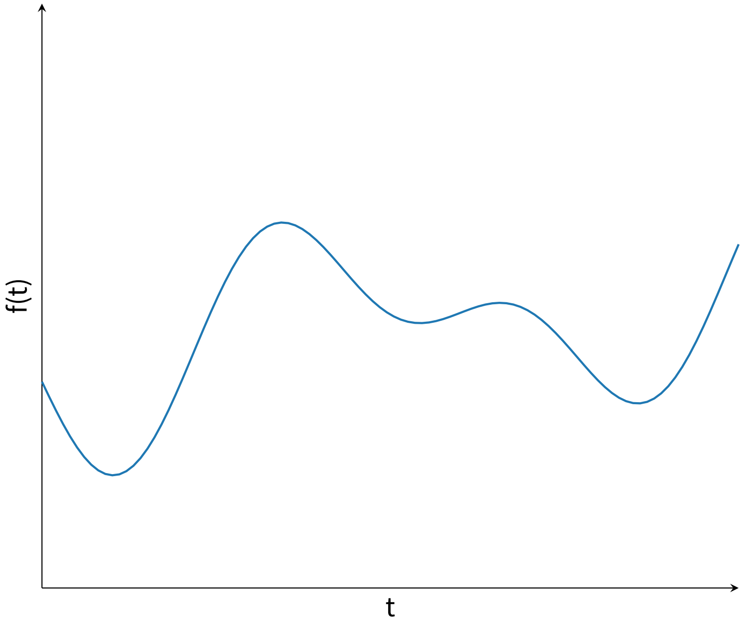


Left to right:
Pafnuty Chebyshev, Emile Borel, Charles Jean de la Vallée-Poussin
(source: Wikipedia)

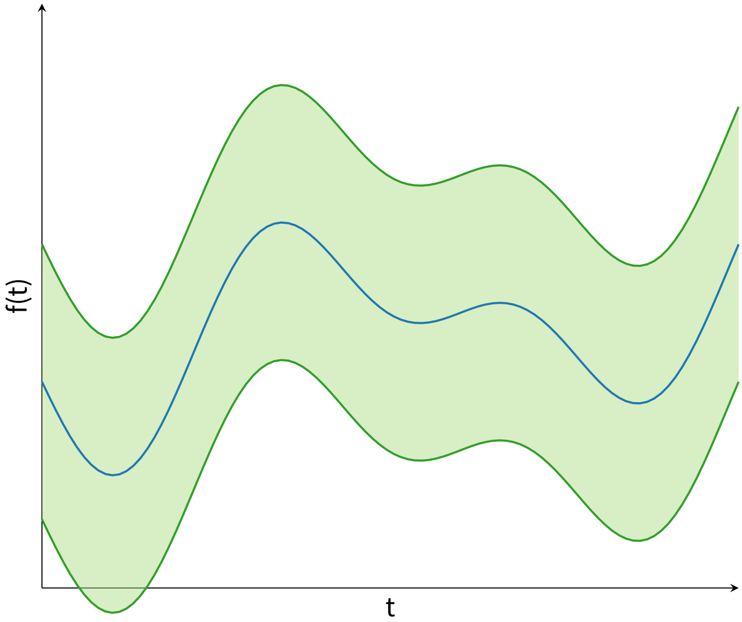
Watt’s linkage



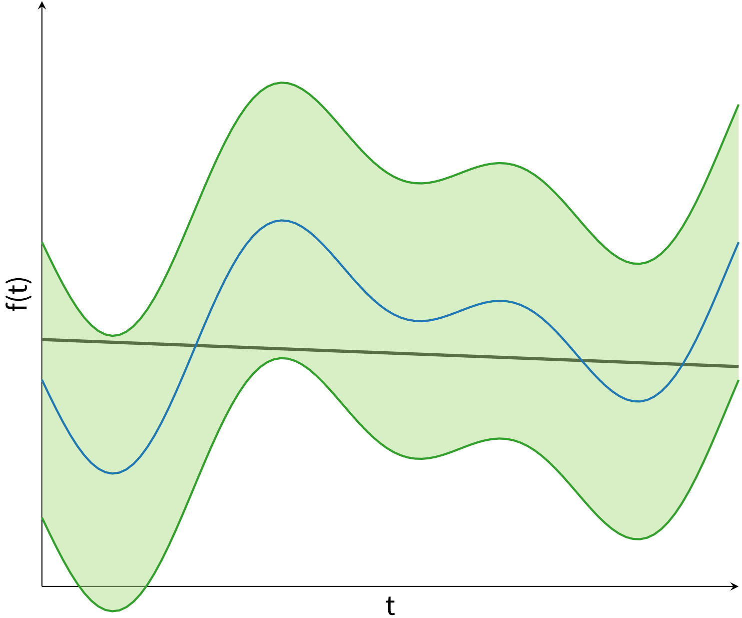
Chebyshev approximation



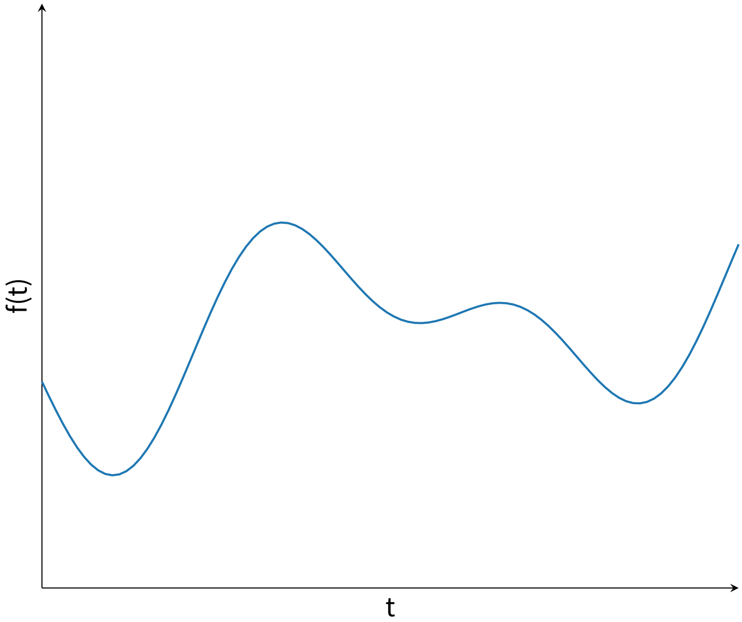
Chebyshev approximation



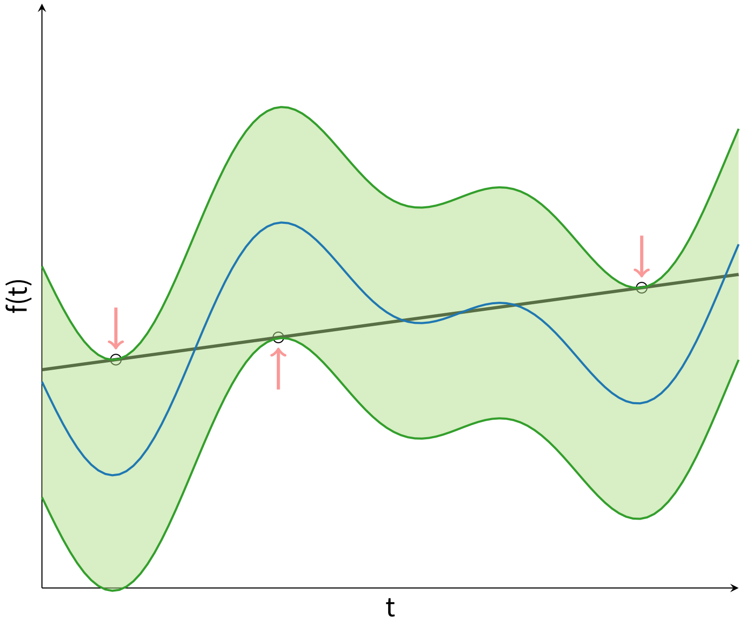
Chebyshev approximation



Chebyshev approximation



Chebyshev approximation



Theorem (Polynomial approximation[1])

$p \in \mathcal{P}_n$ is a best polynomial approximation of a function f if and only if there exists a sequence of at least $n + 2$ points of maximal deviation and the sign of the deviation at these points alternate.

Definition

We call such a sequence a sequence of alternating extreme points.

[1] Chebyshev.

Ways to prove

Chebyshev's Theorem can be proven in different ways. We outline three approaches.

- **Algebraic proofs (Borel, 1905)** rely on the fundamental theorem of algebra (a nontrivial polynomial of degree d vanishes at most d times). This is the classical approach.
- **Convex analytic proofs (Laurent, 1972)** use the fact that any polynomial can be written as a linear combination of monomials to formulate the problem of best polynomial approximation:
  minimise $\Phi(\boldsymbol{a}) = \sup_{t\in[a,b]} \left| f(t) - \sum_{i=0}^{n} a_i t^i \right|$, s.t. $\boldsymbol{a} \in \mathbb{R}^{n+1}$.
  Chebyshev's Theorem is derived from the necessary and sufficient condition for optimality $0 \in \partial\Phi(a^*)$.
- **Semi-infinite programming proofs (Glashoff and Gustafson, 1983)** use a linear semi-infinite programming reformulation:

$$\text{minimise } u \text{ subject to } \left| f(t) - \sum_{i=0}^{n} a_i t^i \right| \leq u, \forall t \in [a, b] \qquad (1)$$

Different proofs: discussion

- Despite superficial differences between these three proofs, it is possible to break them up into building blocks and establish a correspondence between these building blocks (Interconnection Diagram).
- At the same time, each of these proofs highlights a different perspective on the same approximation problem by drawing from different mathematical toolsets (functional analysis, convex analysis).

Examples

Many different generalisations of Chebyshev's Theorem have been obtained, and for a large majority their proofs rely on the generalisation of one of the proofs given above. For instance,

- generalisations of the convex subdifferential enable us to study nonconvex versions of the approximation problem (free knots polynomial splines);
- the reverse approximation problem (maximising the length of an interval where the deviation is below a certain value) can be modelled through a generalised SIP (Still, 1999);
- through function analytical considerations we are no longer restricted to polynomial approximation but can consider other families of functions.

The methodology of this proposal consists in building upon the diversity of these generalisations to propagate results between them and construct more powerful tools in approximation theory, variational analysis and semi-infinite programming. Exploiting the interconnections between these different methodologies is a completely new way of approaching approximation problems.

Open problems

Our ambition is to attack notoriously difficult problems in approximation theory that generalise classic Chebyshev approximation results.
The two problems that we propose to investigate generalise Chebyshev's approximation problem in two different directions.

1. The first one, approximation by polynomial splines with free knots, leads to a nonconvex optimisation problem, which we can study using nonconvex analysis.
2. The second one is multivariate approximation. It is still a convex problem, but the functions under consideration are much more complex.

Polynomial splines

Definition (Polynomial Spline)

A polynomial spline is a piecewise polynomial function. Each polynomial piece lies on an interval $[\xi_i, \xi_{i+1}],\ i = 0, \dots, N-1$. The points ξ_0 and ξ_N are the external knots, and the points ξ_i, ($i = 1, \dots, N-1$) are the internal knots of the polynomial spline.

A spline is generally not infinitely differentiable at its knots. Its degree is defined as the maximum degree of all its polynomial pieces. We let $\mathcal{S}_n$ be the set of polynomial splines of degree at most n.

Problem formulation

Consider the problem of finding a best approximation by a continuous spline $s(t, \boldsymbol{\xi}, \boldsymbol{a}) \in \mathcal{S}_n$:

$$\text{minimise } \Psi(s) = \sup_{t \in [\xi_0, \xi_N]} \left| s(t, \boldsymbol{\xi}, \boldsymbol{a}) - f(t) \right|, \tag{2}$$

where $\boldsymbol{\xi}$ is the vector of knots and $\boldsymbol{a}$ is the vector of coefficients of the polynomial pieces. Any spline is a linear function of the coefficients $\boldsymbol{a}$ and so if the knots are fixed the problem of minimising the function Ψ from (2) with respect to $\boldsymbol{a}$ is convex.If the knots are also variable, this problem is nonconvex.

Polynomial splines formulation

Splines are formulated as:

$$s(t) = a_{00} + \sum_{i=0}^{N-1} \sum_{j=1}^{n} a_{ij}(t - \xi_i)_+^j, \tag{3}$$

where $t_+ = \max(0, t)$ is the truncated power function, $\xi_i, i = 0, \ldots, N-1$ are knots and a_{ij} is the j-th coefficient for the i-th polynomial piece.

Necessary optimality conditions

The characterisation theorem is given as:

Theorem

For a polynomial spline $s \in \mathcal{S}_n$ with $N+1$ knots $\xi_0, \ldots, \xi_N$ to be a best Chebyshev approximation to a continuous function f over the interval $[\xi_0, \xi_N]$, the following two conditions need to be satisfied:

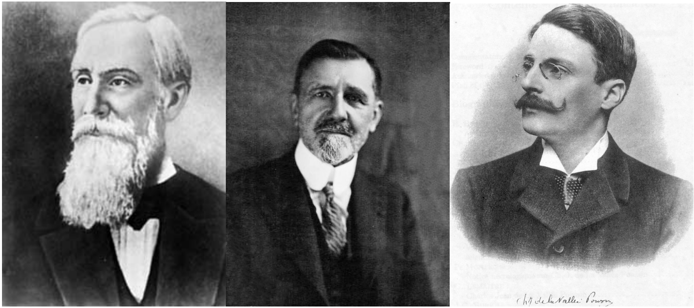
1. there exists an interval $[\xi_p, \xi_q]$ containing an alternating sequence of at least $n(q-p)+2+l$ points, where l is the number of points where the spline is not differentiable inside (ξ_p, ξ_q);
2. if the spline is non-differentiable at an end point of this sequence, then at this point either its right derivative is strictly greater than its left derivative and $(s-f)$ is negative, or its right derivative is strictly less than its left derivative and $(s-f)$ is positive.

Discussions

This result is interesting from several points of view.

1. It corresponds to a necessary and sufficient condition for stationarity in the sense of Demyanov-Rubinov, that is, there is no negative directional derivative.
2. It improves other best approximation conditions, at least when only continuity is required.
3. Since many optimisation methods verify necessary optimality conditions as their termination conditions, the improvement in necessary optimality conditions leads to an improvement in the approximation quality.

Introduction



Left to right:
Pafnuty Chebyshev, Emil Borel, Charles Jean de la Vallée-Poussin
(source: Wikipedia)