High dimensional approximation of parametric PDE's
Theory and Algorithms

Albert Cohen

Laboratoire Jacques-Louis Lions
Université Pierre et Marie Curie
Paris

IHP, October 2016

Overview

Part 1. Theory : high dimensions, parametric PDEs, sparse polynomial approximation

Part 2. Algorithms : Galerkin, power series, sparse interpolation and least-squares

Part 3. Reduced modeling/bases, data assimilation and parameter estimation

References

R. DeVore, “Nonlinear approximation”, Acta Numerica, 1998.

A. Cohen, R. DeVore and C. Schwab, “Analytic regularity and polynomial approximation of parametric and stochastic PDEs”, Analysis and Application, 2011.

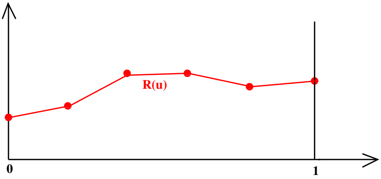
A. Cohen and R. DeVore, “High dimensional approximation of parametric PDEs”, Acta Numerica, 2015.

Part 1

Theory : high dimensions, parametric PDEs, sparse polynomial approximation

The curse of dimensionality

Consider a continuous function $y \mapsto u(y)$ with $y \in [0,1]$.
Sample at equispaced points.
Reconstruct, for example by piecewise linear interpolation.



Error in terms of point spacing $h > 0$: if u has C^2 smoothness

$$\|u - R(u)\|_{L^\infty} \leq C\|u''\|_{L^\infty} h^2.$$

Using piecewise polynomials of higher order, if u has C^m smoothness

$$\|u - R(u)\|_{L^\infty} \leq C\|u^{(m)}\|_{L^\infty} h^m.$$

In terms of the number of samples $n \sim h^{-1}$, the error is estimated by n^{-m}.

In d dimensions : $u(y) = u(y_1, \cdots, y_d)$ with $y \in [0,1]^d$. With a uniform sampling, we still have

$$\|u - R(u)\|_{L^\infty} \leq C\Big(\sup_{|\alpha|=m} \|\partial^\alpha u\|_{L^\infty}\Big) h^m,$$

but the number of samples is now $n \sim h^{-d}$, and the error estimate is in $n^{-m/d}$.

Other sampling/reconstruction methods cannot do better

Can be explained by *n*-width

Let X be a normed space and $\mathcal{K} \subset X$ a compact set.

Linear n-width (Kolmogorov) :

$$d_N(\mathcal{K})_X := \inf_{\dim(E)=n} \max_{u\in\mathcal{K}} \min_{v\in E} \|u - v\|_X.$$

Benchmark for linear approximation methods applied to the elements from $\mathcal{K}$.

If $X = L^\infty([0,1]^d)$ and $\mathcal{K}$ is the unit ball of $C^m([0,1]^d)$ it is known that

$$cn^{-m/d} \leq d_n(\mathcal{K})_X \leq Cn^{-m/d}.$$

Upper bound : approximation by a specific method.

Lower bound : diversity in $\mathcal{K}$.

Exponential growth in d of the needed complexity to reach a given accuracy.

Lower bound : idea of proof

For simplicity work in dimension $d = 1$, and with $\mathcal{K}$ the unit ball of C^1 functions.

Pick $\varphi \in C^\infty$ compactly supported in $]0,1[$ with $\varphi(\frac{1}{2}) = c > 0$ and $\|\varphi\|_{C^1} = 1$.

We build a collection of functions in $\mathcal{K}$ by rescaling by a factor m and scrambling

$$\varphi_\varepsilon = \sum_{k=0}^{m-1} \varepsilon_k m^{-1} \varphi(mx - k), \quad \varepsilon = (\varepsilon_0, \dots, \varepsilon_{m-1}), \quad \varepsilon_j = \pm 1.$$

These functions have values $cm^{-1}\varepsilon_k$ at the m points $x_k = m^{-1}(k + \frac{1}{2})$.

Now we take $m = n + 1$. For any space such that $\dim(E) = n$, we consider

$$F := \{(v(x_0), \dots, v(x_n)) \; : \; v \in E\} \subset \mathbb{R}^{n+1}.$$

Since $\dim(F) \leq n$, there exists a vector $g = (g_0, \dots, g_n) \in F^\perp$. Thus, for any $v \in E$, $\sum_{k=0}^{n} v(x_k) g_k = 0$, which means that $v(x_k)$ has opposite sign to g_k for at least one k.

Thus if we take ε such that $\varepsilon_k = \mathrm{sign}(g_k)$, it follows that for any $v \in E$, we have

$$\|v - \varphi_\varepsilon\|_{L^\infty} \geq \sup_k |v(x_k) - c(n+1)^{-1}\varepsilon_k| \geq c(n+1)^{-1}.$$

which shows that $d_n(\mathcal{K})_{L^\infty} \geq cn^{-1}$.

Same construction in dimension d and for C^m functions gives lower bound $cn^{-m/d}$.

Non-linear methods cannot do better

Use a notion of nonlinear *n*-width (Alexandrov, DeVore-Howard-Micchelli).

Consider maps $E : \mathcal{K} \mapsto \mathbb{R}^n$ (encoding) and $R : \mathbb{R}^n \mapsto X$ (reconstruction).

Introducing the distorsion of the pair (E, R) over $\mathcal{K}$

$$\max_{u\in\mathcal{K}} \|u - R(E(u))\|_X,$$

we define the nonlinear *n*-width of $\mathcal{K}$ as

$$\delta_n(\mathcal{K})_X := \inf_{E,R} \max_{u\in\mathcal{K}} \|u - R(E(u))\|_X,$$

where the infimum is taken over all continuous maps (E, R). Comparison with the Kolmorgorov *n*-width : $\delta_n \leq d_n$ and sometimes substantially smaller.

If $X = L^\infty([0,1]^d)$ and $\mathcal{K}$ is the unit ball of $C^m([0,1]^d)$ it is known that

$$cn^{-m/d} \leq \delta_n(\mathcal{K})_X \leq Cn^{-m/d}.$$

Many other variants of *n*-widths exist (book by A. Pinkus).

Infinitely smooth functions

Nowak and Wozniakowski : if $X = L^\infty([0,1]^d)$ and

$$\mathcal{K} := \{u \in C^\infty([0,1]^d) \; : \; \|\partial^\nu u\|_{L^\infty} \leq 1 \;\text{ for all }\; \nu\}.$$

then, for the linear width,

$$\min\{n \; : \; d_n(\mathcal{K})_X \leq 1/2\} \geq c2^{d/2}.$$

High dimensional problems occur frequently :

PDE's with solutions $u(x,v,t)$ defined in phase space : $d = 7$.

Post-processing of numerical codes : u solver with imput parameters $(y_1, \cdots, y_d)$.

Learning theory : u regression function of imput parameters $(y_1, \cdots, y_d)$

In these applications d may be of the order up to 10^3.

Approximation of stochastic-parametric PDEs : $d = +\infty$.

Smoothness properties of functions should be revisited by other means than C^m classes, and appropriate approximation tools should be used.

Parametric/Stochastic PDEs

We are interested in PDE's of the general form

$$\mathcal{D}(u, y) = 0,$$

where $\mathcal{D}$ is a partial differential operator, u is the unknown and $y = (y_j)_{j=1,\dots,d}$ is a parameter vector of dimension $d >> 1$ or $d = \infty$ ranging in some domain U.

We assume well-posedness of the solution in some Banach space V for every $y \in U$,

$$y \mapsto u(y)$$

is the solution map from U to V.

Solution manifold $\mathcal{M} := \{u(y) \ : \ y \in U\} \subset V$.

The parameters may be deterministic (control, optimization, inverse problems) or random (uncertainty modeling and quantification, risk assessment). In the second case the solution $u(y)$ is a V-valued random variable.

These applications often requires many queries of $u(y)$, and therefore in principle running many times a numerical solver.

Objective : economical numerical approximation of the map $y \mapsto u(y)$.

Related objectives : numerical approximation of scalar quantities of interest $y \mapsto Q(y) = Q(u(y))$, or of averaged quantities $\overline{u} = \mathbb{E}(u(y))$ or $\overline{Q} = \mathbb{E}(Q(y))$.

Guiding example : elliptic PDEs

We consider the steady state diffusion equation

$$-\mathrm{div}(a\nabla u) = f \quad \mathrm{on} \quad D \subset \mathbf{R}^{\mathrm{m}} \quad \mathrm{and} \quad \mathrm{u}_{|\partial \mathrm{D}} = 0,$$

set on a domain $D \subset \mathbb{R}^m$, where $f = f(x) \in L^2(D)$ and $a \in L^\infty(D)$

Lax-Milgram lemma : assuming $a_{\min} := \min_{x\in D} a(x) > 0$, unique solution $u \in V = H_0^1(D)$ with

$$\|u\|_V := \|\nabla u\|_{L^2(D)} \leq \frac{1}{a_{\min}} \|f\|_{V'}.$$

Proof of the estimate : multiply equation by u and integrate

$$a_{\min}\|u\|_V^2 \leq \int_D a\nabla u \cdot \nabla u = -\int_D u\,\mathrm{div}(a\nabla u) = \int_D uf \leq \|u\|_V \|f\|_{V'}.$$

We may extend this theory to the solution of the weak (or variational) formulation

$$\int_D a\nabla u \cdot \nabla v = \langle f, v\rangle, \quad v \in V = H_0^1(D),$$

if $f \in V' = H^{-1}(D)$

Parametrization

Assume diffusion coefficients in the form of an expansion

$$a = a(y) = \overline{a} + \sum_{j\geq 1} y_j\psi_j, \quad y = (y_j)_{j\geq 1} \in U,$$

with $d >> 1$ or $d = \infty$ terms, where $\overline{a}$ and $(\psi_j)_{j\geq 1}$ are functions from L^∞,

Note that $a(y)$ is a function for each given y. We may also write

$$a = a(x,y) = \overline{a}(x) + \sum_{j\geq 1} y_j\psi_j(x), \quad x \in D, y \in U,$$

where x and y are the spatial and parametric variable, respectively. Likewise, the corresponding solution $u(y)$ is a function $x \mapsto u(y,x)$ for each given y. We often ommit the reference to the spatial variable.

Up to a change of variable, we assume that all y_j range in $[-1,1]$, therefore

$$y \in U = [-1,1]^d \text{ or } [-1,1]^{\mathbb{N}}.$$

Uniform ellipticity assumption :

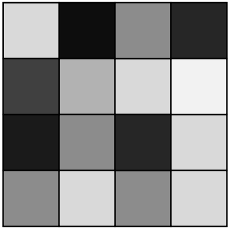
$$(UEA) \qquad 0 < r \leq a(x,y) \leq R, \;\; x \in D, \; y \in U$$

Then the solution map is bounded from U to $V := H_0^1(D)$, that is, $u \in L^\infty(U,V)$:

$$\|u(y)\|_V \leq C_r := \frac{\|f\|_{V'}}{r}, \;\; y \in U,$$

Example of parametrization : piecewise constant coefficients

Assume that a is piecewise constant over a partition $\{D_1, \dots, D_d\}$ of D, and such that on each D_j the value of a varies on $[c - c_j, c + c_j]$ for some $c > 0$ and $0 < c_j < c$.



Then a natural parametrization is

$$a(y) = \overline{a} + \sum_{j=1}^{d} y_j \psi_j, \quad \overline{a} = c, \quad \psi_j = c_j \chi_{D_j},$$

with $y = (y_j)_{j=1,\dots,d} \in U = [-1, 1]^d$.

Example of parametrization : Karhunen-Loeve representation

Assume $a = (a(x))_{x \in D}$ is a random process with average

$$\overline{a}(x) = \mathbb{E}(a(x)),$$

and covariance function

$$C_a(x,z) = \mathbb{E}\Big(\tilde{a}(x)\tilde{a}(z)\Big), \quad \tilde{a} := a - \overline{a}, \quad x, z \in D.$$

Define the integral operator by

$$Tv(x) = \int_D C_a(x,z)v(z)dz,$$

self-adjoint, positive and compact in $L^2(D)$. Therefore it admits an L^2 orthonormal basis $(\varphi_j)_{j\geq 1}$ of eigenfunctions, associated to eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq 0$, such that $\lambda_n \longrightarrow 0$ as $n \longrightarrow +\infty$.

Karhunen-Loeve (KL) decomposition (a.k.a. principal component analysis) :

$$a = \overline{a} + \sum_{j\geq 1} \xi_j \varphi_j, \quad \xi_j := \int_D a(x)\varphi_j(x)dx.$$

Analogy : PCA for finite dimensional random vector

Let $z = (z_1, \dots, z_d) \in \mathbb{R}^d$ be a d-dimentional random vector with average $\overline{z} = \mathbb{E}(z)$ and let $x = z - \overline{x}$ be its centered version. The covariance matrix

$$K = (\mathbb{E}(x_i, x_j))_{i,j=1,\dots,d},$$

is symmetric and positive, since $\langle Kv, w \rangle = \mathbb{E}(\langle x, v \rangle \langle x, w \rangle)$. Its eigenvectors form an orthonormal basis $(\varphi_1, \dots, \varphi_d)$, associated to eigenvalues $\lambda_1 \geq \cdots \geq \lambda_d \geq 0$.

So we write $x = \sum_{j=1}^{d} \xi_j \varphi_j$ with $\xi_j := \langle x, \varphi_j \rangle$, where

$$\mathbb{E}(\xi_j) = \mathbb{E}(\xi_i \xi_j) = 0 \quad \text{if } i \neq j, \quad \mathbb{E}(\xi_j^2) = \lambda_j.$$

This representation has an optimal property : with $V_n = \mathrm{span}\{\varphi_1, \dots, \varphi_n\}$, one has

$$E(\|P_{V_n} x\|^2) = \lambda_1 + \cdots + \lambda_n \geq E(\|P_{W_n} x\|^2),$$

for all W_n of dimension n, or equivalently

$$E(\|x - P_{V_n} x\|^2) = \lambda_{n+1} + \cdots + \lambda_d \leq E(\|x - P_{W_n} x\|^2).$$

Proof by induction : use $W_n = W_{n-1} \oplus^{\perp} \mathbb{R}\psi$ with $\psi \in W_n \cap V_{n-1}^{\perp}$ of norm 1, so that

$$\begin{aligned} E(\|P_{W_n} x\|^2) = E(\|P_{W_{n-1}} x\|^2) + E(|\langle x, \psi \rangle|^2) \quad &\leq \lambda_1 + \cdots + \lambda_{n-1} + \textstyle\sum_{j \geq n} \lambda_j |\langle \psi, \varphi_j \rangle|^2 \\ &\leq \lambda_1 + \cdots + \lambda_n. \end{aligned}$$

Properties of KL representation

The ξ_j are centered and decorelated scalar random variables, with

$$\mathbb{E}(\xi_j) = 0, \quad \mathbb{E}(\xi_i \xi_j) = 0 \quad \text{if} \quad j \neq i, \quad \mathbb{E}(|\xi_j|^2) = \lambda_j.$$

If the random process a is bounded, then the variables ξ_j have bounded range $|\xi_j| \leq c_j$, so that with $y_j := \xi_j / c_j$ and $\psi_j := c_j \varphi_j$ we may also write

$$a = \overline{a} + \sum_{j \geq 1} y_j \psi_j, \quad y = (y_j)_{j \geq 1} \in U = [-1, 1]^{\mathbb{N}}.$$

The KL representation is optimal for trunctation in mean-square $L^2(D)$-error :

$$\inf_{\dim(E)=J} \mathbb{E}(\|\tilde{a} - P_E \tilde{a}\|_{L^2}^2),$$

is attained by $E = E_J := \operatorname{span}\{\psi_1, \dots, \psi_J\}$ with

$$\mathbb{E}(\|\tilde{a} - P_{E_J} \tilde{a}\|_{L^2}^2) = \mathbb{E}\Big(\| \sum_{j > J} y_j \psi_j \|_{L^2}^2\Big) = \sum_{j > J} \lambda_j.$$

Case of a stationary process : $C_a(x, z) = \kappa(x - z)$, that is T is a convolution operator. If D is the m-dimensional 2π-periodic torus, the KL basis is of Fourier type

$$x \mapsto \varphi_k(x) := (2\pi)^{-m/2} e^{ik\dot{x}}, \quad k \in \mathbb{Z}^m.$$

Model reduction

Objective : fast approximate computation of $y \mapsto u(y)$ for many queries of y.

Vehicle : separable (low rank) approximations of the form

$$u(x,y) \approx u_n(x,y) := \sum_{k=1}^{n} v_k(x)\phi_k(y),$$

where $v_k : D \to \mathbb{R}$ with $v_k \in V$ and $\phi_k : U \to \mathbb{R}$. Equivalently

$$u_n(y) := \sum_{k=1}^{n} v_k \phi_k(y) = \sum_{k=1}^{n} \phi_k(y) v_k \in V_n := \mathrm{span}\{v_1, \dots, v_n\} \subset V, \quad y \in U.$$

Thus we approximate simultaneously all solutions $u(y)$ in the same n-dimensional space $V_n \subset V$.

By the way, this is what we do when we use a finite element solver :

$$y \mapsto u_h(y) \in V_h \subset V.$$

So what's new here ?

Accurate solutions may require V_h of very large dimension $n_h = \dim(V_h) >> 1$ and each query $y \mapsto u_h(y)$ is expensive.

We hope to achieve same order of accuracy $n << n_h$ by a choice of V_n adapted to the parametric problem. In practice the functions $v_1, \dots, v_n$ are typically picked from such a finite element space V_h, so that $u_n(y) \in V_h$ for all y but actually belongs to the much smaller space $V_n \subset V_h$.

Measure of performance

1. Uniform sense

$$\|u-u_n\|_{L^\infty(U,V)} := \sup_{y\in U} \|u(y)-u_n(y)\|_V,$$

2. Mean-square sense, for some measure μ on U,

$$\|u-u_n\|^2_{L^2(U,V,\mu)} := \int_U \|u(y)-u_n(y)\|_V^2 d\mu(y).$$

If μ is a probability measure, and y randomly distributed according to this measure, we have

$$\|u-u_n\|^2_{L^2(U,V,\mu)} = \mathbb{E}(\|u(y)-u_n(y)\|_V^2).$$

Note that we always have

$$\mathbb{E}(\|u(y)-u_n(y)\|_V^2) \leq \|u-u_n\|^2_{L^\infty(U,V)}.$$

A “worst case” estimate is always above an “average” estimate.

Optimal spaces ?

Best n-dimensional space for approximation in the uniform sense : the space F_n one that reaches the Kolmogorov n-width of the solution manifold in the V norm

$$d_n = d_n(\mathcal{M}) := \inf_{\dim(E)\leq n} \sup_{v\in\mathcal{M}} \min_{w\in E} \|v-w\|_V = \inf_{\dim(E)\leq n} \sup_{y\in U} \min_{w\in E} \|u(y)-w\|_V.$$

Best n-dimensional space for approximation in the mean-square sense : principal component analysis in V (instead of L^2 with KL basis). Consider an orthonormal basis $(e_k)_{k\geq 1}$ of V and decompose

$$u(y) := \sum_{k\geq 1} u_k(y) e_k, \quad u_k(y) := \langle u(y), e_k\rangle_V.$$

Introduce the infinite correlation matrix $M = (\mathbb{E}(u_k u_l))_{k,l\geq 1}$. It has eigenvalues $(\lambda_k)_{k\geq 1}$ and associated eigenvectors $g_k = (g_{k,l})_{l\in\mathbb{N}}$ which form an orthonormal basis of $\ell^2(\mathbb{N})$. The best space is

$$G_n := \mathrm{span}\{v_1,\dots,v_n\}, \quad v_k := \sum_{l\geq 1} g_{k,l} e_l,$$

and has performance

$$\varepsilon_n^2 := \inf_{\dim(E)\leq n} \mathbb{E}\Big(\min_{w\in E} \|u(y)-w\|_V^2\Big) = \sum_{k>n} \lambda_k \leq d_n^2.$$

Realistic strategies

The optimal spaces F_n and G_n are usually out of reach. There are two main computational approaches to realistically design the approximation $u_n = \sum_{k=1}^{n} v_k \phi_k$.

1. Expand formally the solution map $y \mapsto u(y)$ in a given "basis" $(\phi_k)_{k\geq 1}$ of high dimensional functions

$$u(y) = \sum_{k\geq 1} v_k \phi_k(y),$$

where $v_k \in V$ are viewed as the coefficients in this expansion.

Compute these coefficients for $k = 1, \dots, n$ approximately by some numerical procedure.

Main representative (this lecture) : Polynomial methods (the ϕ_k are multivariate polynomials).

2. Compute first a "good" basis $\{v_1, \dots, v_n\}$ and define V_n as their span. Then, for any given instance y, compute $u_n(y) \in V_n$ by a numerical method.

Main representative : Reduced Bases (RB) methods emulate the n-width spaces F_n for uniform, or $L^\infty(U, V)$, approximation. Proper Orthogonal Decompositions (POD) methods emulate the principal component spaces G_n for mean-square, or $L^2(U, V, \mu)$, approximation.

Remarks

In the second approach, the functions v_k are typically computed in an heavy offline stage, then for any given y, the computation of $u_n(y)$ is done in a cheap online stage.

The first approach gives immediate access to the approximation u_n for all values of y since the functions v_k and ϕ_k are both precomputed offline, the online stage is then a trivial recombination.

Other important distinction : intrusive versus non-intrusive methods. The latter are based on post-processing individual solution instances

$$u(y^i), \quad y^i \in U, \quad i = 1, \dots, m.$$

They may benefit of a pre-existing numerical solver viewed as a blackbox and do not necessarily require full knowledge of PDE model.

In practice, the v_k are typically chosen in a discrete (finite element) space $V_h \subset V$, with $n_h = \dim(V_h) >> n$. Equivalently, we apply the above technique to the discrete solution map $y \mapsto u_h(y) \in V_h$. The error may thus be decomposed into the finite element discretization error and the model reduction error.

How to defeat the curse of dimensionality ?

The map $y \mapsto u(y)$ is high dimensional, or even infinite dimensional $y = (y_j)_{j \geq 1}$.

We are thus facing the curse of dimensionality when trying to approximate it with conventional discretization tools in the y variable (Fourier series, finite elements).

A general function of d variable with m bounded derivatives cannot be approximated in L^∞ with rate better than $n^{-m/d}$ where n is the number of degrees of freedom.

A possible way out : exploit anisotropic features in the function $y \mapsto u(y)$.

The PDE is parametrized by a function a (diffusion coefficient, velocity, domain boundary) and y_j are the coordinates of a in a certain basis representation $a = \overline{a} + \sum_{j \geq 1} y_j \psi_j$.

If the ψ_j decays as $j \to +\infty$ (for instance if a has some smoothness) then the variable y_j are less active for large j.

We shall see that in certain relevant instances, this mechanism allows to break the curse of dimensionality by using suitable expansions : we obtain approximation rates $\mathcal{O}(n^{-s})$ that are independent of d in the sense that they hold when $d = \infty$.

One key tool for obtaining such result is the concept of sparse approximation.

Sparsity

Small dimensional phenomenon in high dimensional context

Simple example : vector $x = (x_1, \cdots, x_N) \in \mathbf{R}^{\mathrm{N}}$ representing a signal, image or function, discretized with $N >> 1$.

The vector x is sparse if only few of its coordinates are non-zero.

How to quantify this ?

The set of *n*-sparse vectors

$$\Sigma_n := \{x \in \mathbb{R}^{\mathrm{N}} \; ; \; \#\{\mathrm{i} \; ; \; \mathrm{x_i} \neq 0\} \leq \mathrm{n}\}$$

As n gets smaller, $x \in \Sigma_n$ gets sparser.

More realistic : a vector is quasi-sparse if only a few numerically significant coordinates concentrate most of the information. How to measure this notion of concentration ?

Remarks :

A vector in Σ_n is characterized by n non-zero values and their n positions.

Intrinsically nonlinear concepts : $x, y \in \Sigma_n$ does not imply $x + y \in \Sigma_n$.

Sparsity is often hidden, and revealed through an appropriate representation (change of basis).

Example : representations of natural images in wavelet bases are quasi-sparse, and therefore used in image compression standards (JPEG 2000)

Sparse approximation in ℓ^q spaces : fundamental lemma (Stechkin)

Consider sequences $\mathbf{d} = (d_\nu)_{\nu\in\mathcal{F}}$ in $\ell^q(\mathcal{F})$ where $\mathcal{F}$ is a countable index set.

Best n-term approximation : we seek to approximate $\mathbf{d}$ by a sequence supported on a set of size n.

Best choice : $\mathbf{d}_n$ defined by leaving d_ν unchanged for the n largest $|d_\nu|$ and setting the others to 0.

Lemma : for $0 < p < q \leq \infty$, one has

$$\mathbf{d} \in \ell^p(\mathcal{F}) \implies \|\mathbf{d} - \mathbf{d}_n\|_{\ell^q} \leq C(n+1)^{-s}, \quad s = \frac{1}{p} - \frac{1}{q}, \quad C := \|\mathbf{d}\|_{\ell^p}.$$

Proof : introduce $(d_k^*)_{k\geq 1}$ the decreasing rearrangement of $(|d_\nu|)_{\nu\in\mathcal{F}}$, and combine

$$\|\mathbf{d} - \mathbf{d}_n\|_{\ell^q}^q = \sum_{k>n} |d_k^*|^q = \sum_{k>n} |d_k^*|^{q-p}|d_k^*|^p \leq C^p |d_{n+1}^*|^{q-p}$$

with

$$(n+1)|d_{n+1}^*|^p \leq \sum_{k=1}^{n+1} |d_k^*|^p \leq C^p.$$

Note that a large value of s corresponds to a value $p < 1$ (non-convex spaces).

A sharp result : weak ℓ^p spaces

A sequence $\mathbf{d} = (d_\nu)_{\nu\in\mathcal{F}}$ belongs to $w\ell^p(\mathcal{F})$ if and only if

$$\#\{\nu \text{ s.t. } |d_\nu| > \eta\} \leq C\eta^{-p},$$

or equivalently, the decreasing rearrangement $(d_k^*)_{k\geq 1}$ of $(|d_\nu|)$ satisfies

$$d_k^* \leq Ck^{-1/p}.$$

The $w\ell^p$ quasi-norm can be defined by

$$\|\mathbf{d}\|_{w\ell^p} := \sup_{k\geq 1} k^{1/p} d_k^*.$$

Obviously $\ell^p \subset w\ell^p \subset \ell^{p+\varepsilon}$ with strict inclusions.

Lemma : for $0 < p < q \leq \infty$, one has

$$\mathbf{d} \in w\ell^p(\mathcal{F}) \iff \|\mathbf{d} - \mathbf{d}_n\|_{\ell^q} \leq C(n+1)^{-s}, \quad s = \frac{1}{p} - \frac{1}{q}, \quad C := \|\mathbf{d}\|_{\ell^p}.$$

Proof : we now write

$$\|\mathbf{d} - \mathbf{d}_n\|_{\ell^q}^q = \sum_{k>n} |d_k^*|^q \leq \|\mathbf{d}\|_{w\ell^p}^q \sum_{k>n} k^{-q/p} \leq C\|\mathbf{d}\|_{w\ell^p}^q n^{1-q/p}$$

which gives the forward implication (converse is left as an exercise).

From sequence approximation to Banach space valued function approximation

If a V-valued u has an expansion of the form $u(y) = \sum_{\nu\in\mathcal{F}} u_\nu \phi_\nu(y)$, in a given basis $(\phi_\nu)_{\nu\in\mathcal{F}}$, we use Stechkin's lemma to study the approximation of u by

$$u_n := \sum_{\nu\in\Lambda_n} u_\nu \phi_\nu,$$

where $\Lambda_n \subset \mathcal{F}$ corresponds to the n-largest $\|u_\nu\|_V$.

If $\sup_{y\in U} |\phi_\nu(y)| = 1$, then by triangle inequality

$$\|u - u_n\|_{L^\infty(U,V)} \leq \sum_{\nu\notin\Lambda_n} \|u_\nu\phi_\nu\|_{L^\infty(U,V)} = \sum_{\nu\notin\Lambda_n} \|u_\nu\|_V,$$

If $(\phi_\nu)_{\nu\in\mathcal{F}}$ is an orthonormal basis of $L^2(U,\mu)$, then by Parseval equality

$$\|u - u_n\|^2_{L^2(U,V,\mu)} = \sum_{\nu\notin\Lambda_n} \|u_\nu\|^2_V,$$

For concrete choices of bases a relevant question is thus : what smoothness properties of a function ensure that its coefficient sequence belongs to ℓ^p for small values of p ?

In the case of wavelet bases, such properties are characterized by Besov spaces.

In our present setting of high-dimensional functions $y \mapsto u(y)$ we shall rather use tensor-product polynomial bases instead of wavelet bases. Sparsity properties will follow to the anisotropic features of these functions.

Return to the main guiding example

Steady state diffusion equation

$$-\mathrm{div}(a\nabla u) = f \quad \text{on} \quad D \subset \mathbb{R}^m \quad \text{and} \quad \mathrm{u}_{|\partial \mathrm{D}} = 0,$$

where $f = f(x) \in L^2(D)$ and the diffusion coefficients are given by

$$a = a(x,y) = \bar{a}(x) + \sum_{j\geq 1} y_j \psi_j(x),$$

where $\bar{a}$ and the $(\psi_j)_{j\geq 1}$ are given functions and $y \in U := [-1,1]^{\mathbb{N}}$. Uniform ellipticity assumption :

$$(UEA) \qquad 0 < r \leq a(x,y) \leq R, \quad x \in D, \; y \in U.$$

Equivalent expression of (UEA) : $\bar{a} \in L^\infty(D)$ and

$$\sum_{j\geq 1} |\psi_j(x)| \leq \bar{a}(x) - r, \quad x \in D,$$

or

$$\left\| \frac{\sum_{j\geq 1} |\psi_j|}{\bar{a}} \right\|_{L^\infty(D)} \leq \theta < 1.$$

Lax-Milgram : solution map is well-defined from U to $V := H_0^1(D)$ with uniform bound

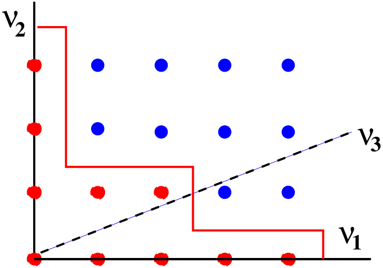
$$\|u(y)\|_V \leq C_r := \frac{\|f\|_{V'}}{r}, \quad y \in U, \quad \text{where} \quad \|v\|_V := \|\nabla v\|_{L^2}.$$

Sparse polynomial approximations using Taylor series

We consider the expansion of $u(y) = \sum_{\nu\in\mathcal{F}} t_\nu y^\nu$, where

$$y^\nu := \prod_{j\geq 1} y_j^{\nu_j} \quad \text{and} \quad t_\nu := \frac{1}{\nu!}\partial^\nu u_{|y=0} \in V \quad \text{with} \quad \nu! := \prod_{j\geq 1} \nu_j! \quad \text{and} \quad 0! := 1.$$

where $\mathcal{F}$ is the set of all finitely supported sequences of integers (finitely many $\nu_j \neq 0$). The sequence $(t_\nu)_{\nu\in\mathcal{F}}$ is indexed by countably many integers.



Objective : identify a set $\Lambda \subset \mathcal{F}$ with $\#(\Lambda) = n$ such that u is well approximated by the partial expansion

$$u_\Lambda(y) := \sum_{\nu\in\Lambda} t_\nu y^\nu.$$

Best n-term approximation

A-priori choices for Λ have been proposed, e.g. (anisotropic) sparse grid defined by restrictions of the type $\sum_j \alpha_j \nu_j \leq A(n)$ or $\prod_j (1+\beta_j \nu_j) \leq B(n)$.

Instead we want to choose Λ optimally adapted to u. By triangle inequality we have

$$\|u-u_\Lambda\|_{L^\infty(U,V)} = \sup_{y\in U} \|u(y)-u_\Lambda(y)\|_V \leq \sup_{y\in U} \sum_{\nu\notin\Lambda} \|t_\nu y^\nu\|_V = \sum_{\nu\notin\Lambda} \|t_\nu\|_V$$

Best n-term approximation in $\ell^1(\mathcal{F})$ norm : use $\Lambda = \Lambda_n$ index set of n largest $\|t_\nu\|_V$.

Stechkin lemma : if $(\|t_\nu\|_V)_{\nu\in\mathcal{F}} \in \ell^p(\mathcal{F})$ for some $p<1$, then for this Λ_n,

$$\sum_{\nu\notin\Lambda_n} \|t_\nu\|_V \leq Cn^{-s}, \quad s := \frac{1}{p}-1, \quad C := \|(\|t_\nu\|_V)\|_{\ell^p}.$$

Question : do we have $(\|t_\nu\|_V)_{\nu\in\mathcal{F}} \in \ell^p(\mathcal{F})$ for some $p<1$?

One main result

Theorem (Cohen-DeVore-Schwab, 2011) : under the uniform ellipticity assumption (UAE), then for any $p < 1$,

$$(\|\psi_j\|_{L^\infty})_{j>0} \in \ell^p(\mathbb{N}) \implies (\|t_\nu\|_V)_{\nu\in\mathcal{F}} \in \ell^p(\mathcal{F}).$$

Interpretations :

(i) The Taylor expansion of $u(y)$ inherits the sparsity properties of the expansion of $a(y)$ into the ψ_j.

(ii) We approximate $u(y)$ in $L^\infty(U, V)$ with algebraic rate $\mathcal{O}(n^{-s})$ despite the curse of (infinite) dimensionality, due to the fact that y_j is less influencial as j gets large.

(iii) The solution manifold $\mathcal{M} := \{u(y)\ ;\ y \in U\}$ is uniformly well approximated by the n-dimensional space $V_n := \mathrm{span}\{t_\nu\ :\ \nu \in \Lambda_n\}$. Its n-width satisfies the bound

$$d_n(\mathcal{M})_V \le \max_{y\in U} \mathrm{dist}(u(y), V_n)_V \le \max_{y\in U} \|u(y) - u_{\Lambda_n}(y)\|_V \le Cn^{-s}.$$

Such approximation rates cannot be proved for the usual a-priori choices of Λ.

Same result for more general linear equations $Au = f$ with affine operator dependance : $A = \overline{A} + \sum_{j\ge1} y_j A_j$ uniformly invertible over $y \in U$, and $(\|A_j\|_{V\to W})_{j\ge1} \in \ell^p(\mathbb{N})$, as well as other models.

Idea of proof : extension to complex variable

Estimates on $\|t_\nu\|_V$ by complex analysis : extend $u(y)$ to $u(z)$ with $z = (z_j) \in \mathbb{C}^{\mathbb{N}}$.

Uniform ellipticity $\sum_{j\geq 1} |\psi_j| \leq \overline{a} - r$ implies that with $a(z) = \overline{a} + \sum_{j\geq 1} z_j \psi_j$,

$$0 < r \leq \Re(a(x,z)) \leq |a(x,z)| \leq 2R, \quad x \in D,$$

for all $z \in \mathcal{U} := \{|z| \leq 1\}^{\mathbb{N}} = \otimes_{j\geq 1}\{|z_j| \leq 1\}$.

Lax-Milgram theory applies : $\|u(z)\|_V \leq C_0 = \frac{\|f\|_{V^*}}{r}$ for all $z \in \mathcal{U}$.

The function $u \mapsto u(z)$ is holomorphic in each variable z_j at any $z \in \mathcal{U}$: its first derivative $\partial_{z_j} u(z)$ is the unique solution to

$$\int_D a(z) \nabla \partial_{z_j} u(z) \cdot \nabla v = -\int_D \psi_j \nabla u(z) \cdot \nabla v, \quad v \in V.$$

Note that ∇ is with respect to spatial variable $x \in D$.

Extended domains of holomorphy : if $\rho = (\rho_j)_{j\geq 0}$ is any positive sequence such that for some $\delta > 0$

$$\sum_{j\geq 1} \rho_j |\psi_j(x)| \leq \overline{a}(x) - \delta, \quad x \in D,$$

then u is holomorphic with uniform bound $\|u(z)\| \leq C_\delta = \frac{\|f\|_{V^*}}{\delta}$ in the polydisc

$$\mathcal{U}_\rho := \otimes_{j\geq 1}\{|z_j| \leq \rho_j\},$$

If $\delta < r$, we can take $\rho_j > 1$.

Estimate on the Taylor coefficients

Use Cauchy formula. In 1 complex variable if $z \mapsto u(z)$ is holomorphic and bounded in a neighbourhood of disc $\{|z| \leq b\}$, then for all z in this disc

$$u(z) = \frac{1}{2i\pi} \int_{|z'|=b} \frac{u(z')}{z - z'} dz',$$

which leads by n differentiation at $z = 0$ to $|u^{(n)}(0)| \leq n! b^{-n} \max_{|z| \leq b} |u(z)|$.

This yields exponential convergence rate $b^{-n} = \exp(-cn)$ of Taylor series for 1-d holomorphic functions. Curse of dimensionality : in d dimension, this yields sub-exponential rate $\exp(-cn^{1/d})$ where n is the number of retained terms.

Recursive application of this to all variables z_j such that $\nu_j \neq 0$, with $b = \rho_j$ gives

$$\|\partial^\nu u_{|z=0}\|_V \leq C_\delta \nu! \prod_{j \geq 1} \rho_j^{-\nu_j},$$

and thus

$$\|t_\nu\|_V \leq C_\delta \prod_{j \geq 1} \rho_j^{-\nu_j} = C_\delta \rho^{-\nu},$$

for any sequence $\rho = (\rho_j)_{j \geq 1}$ such that

$$\sum_{j \geq 1} \rho_j |\psi_j(x)| \leq \bar{a}(x) - \delta.$$

Optimization

Since ρ is not fixed we have

$$\|t_\nu\|_V \leq C_\delta \inf\{\rho^{-\nu} \ : \ \rho \ \text{s.t.} \ \sum_{j\geq 1} \rho_j |\psi_j(x)| \leq \overline{a}(x) - \delta, \ \ x \in D\}.$$

We do not know the general solution to this problem, except in particular case, for example when the ψ_j have disjoint supports.

Instead design a particular choice $\rho = \rho(\nu)$ satisfying the constraint with $\delta = r/2$, for which we prove that

$$(\|\psi_j\|_{L^\infty})_{j\geq 1} \in \ell^p(\mathbb{N}) \implies (\rho(\nu)^{-\nu})_{\nu\in\mathcal{F}} \in \ell^p(\mathcal{F}),$$

therefore proving the main theorem.

A simple case

Assume that the ψ_j have disjoint supports. Then we maximize separately the ρ_j so that

$$\sum_{j\geq 1} \rho_j |\psi_j(x)| \leq \overline{a}(x) - \frac{r}{2}, \quad x \in D,$$

which leads to

$$\rho_j := \min_{x\in D} \frac{\overline{a}(x) - \frac{r}{2}}{|\psi_j(x)|}.$$

We have, with $\delta = \frac{r}{2}$,

$$\|t_\nu\|_V \leq C_\delta \rho^{-\nu} = C_\delta b^\nu,$$

where $b = (b_j)$ and

$$b_j := \rho_j^{-1} = \max_{x\in D} \frac{|\psi_j(x)|}{\overline{a}(x) - \frac{r}{2}} \leq \frac{\|\psi_j\|_{L^\infty}}{R - \frac{r}{2}}.$$

Therefore $b \in \ell^p(\mathbb{N})$. From (UEA), we have $|\psi_j(x)| \leq \overline{a}(x) - r$ and thus $\|b\|_{\ell^\infty} < 1$.

We finally observe that

$$b \in \ell^p(\mathbb{N}) \;\; \text{and} \;\; \|b\|_{\ell^\infty} < 1 \iff (b^\nu)_{\nu\in\mathcal{F}} \in \ell^p(\mathcal{F}).$$

Proof : factorize

$$\sum_{\nu\in\mathcal{F}} b^{p\nu} = \prod_{j\geq 1} \sum_{n\geq 0} b_j^{pn} = \prod_{j\geq 1} \frac{1}{1 - b_j^p}.$$

What about weak ℓ^p-spaces ?

Do we have a result of the type

$$b \in w\ell^p(\mathbb{N}) \text{ and } \|b\|_{\ell^\infty} < 1 \iff (b^\nu)_{\nu\in\mathcal{F}} \in w\ell^p(\mathcal{F}) \quad ?$$

Awnser : no related to results in number theory.

Take the typical sequence $(b_j) = (j+1)^{-1/p}$. so that $b \in w\ell^p(\mathbb{N})$ and $\|b\|_{\ell^\infty} < 1$. We want to know if $\#\{\nu \in \mathcal{F} \ : \ b^\nu \geq \eta\} \leq C\eta^{-p}$ for all $\eta > 0$, or equivalently if

$$t(A) := \#\Big\{\nu \in \mathcal{F} \ : \ \prod_{j\geq 2} j^{\nu_j} \leq A\Big\} \leq CA, \quad A > 0.$$

The left side can be rewritten as $t(A) = \sum_{n=2}^{\lfloor A \rfloor} f(n)$, where $f(n)$ is the number of possible multiplicative partitions (factorisatio numerorum) of n. Estimating $f(n)$ and $t(A)$, is a problem in number theory (Canfield-Erdos-Pomerance). It is known that

$$\frac{t(A)}{A} \sim \exp\left\{\frac{4\sqrt{\log(A)}}{\sqrt{2e}\log(\log(A))}(1+o(1))\right\} \to +\infty$$

Thus the result valid for ℓ^p spaces cannot be true for $w\ell^p$ spaces.

Improved summability results

One defect of the previous result is that it depends on the ψ_j only through $\|\psi_j\|_{L^\infty}$, without taking their support into account.Improved results can be obtained, without relying on complex variable, by better exploiting the specific structure of PDE.

Recursive formula for the Taylor coefficients : with $e_j = (0,\dots,0,1,0,\dots)$ the Kroeneker sequence of index j, the coefficient t_ν is solution to

$$\int_D \bar{a}\nabla t_\nu \nabla v = -\sum_{j:\ \nu_j\neq 0}\int_D \psi_j \nabla t_{\nu-e_j}\nabla v, \quad v\in V.$$

We introduce the quantities

$$d_\nu := \int_D \bar{a}|\nabla t_\nu|^2 \quad \text{and} \quad d_{\nu,j} := \int_D |\psi_j|\,|\nabla t_\nu|^2.$$

Recall that (UEA) implies that $\left\|\frac{\sum_{j\geq 1}|\psi_j|}{\bar{a}}\right\|_{L^\infty(D)} \leq \theta < 1$. In particular

$$\sum_{j\geq 1} d_{\nu,j} \leq \theta d_\nu.$$

We use here the equivalent norm $\|v\|_V^2 := \int_D \bar{a}|\nabla v|^2$.

Lemma : under (UEA), one has $\sum_{\nu\in\mathcal{F}} d_\nu = \sum_{\nu\in\mathcal{F}}\|t_\nu\|_V^2 < \infty$.

Proof

Taking $v = t_\nu$ in the recursion gives

$$d_\nu = \int_D \overline{a}|\nabla t_\nu|^2 = -\sum_{j:\ \nu_j \neq 0} \int_D \psi_j \nabla t_{\nu - e_j} \nabla t_\nu .$$

Apply Young's inequality on the right side gives

$$d_\nu \leq \sum_{j:\ \nu_j \neq 0} \Big(\frac{1}{2} \int_D |\psi_j|\,|\nabla t_\nu|^2 + \frac{1}{2} \int_D |\psi_j|\,|\nabla t_{\nu - e_j}|^2 \Big) = \frac{1}{2} \sum_{j:\ \nu_j \neq 0} d_{\nu,j} + \frac{1}{2} \sum_{j:\ \nu_j \neq 0} d_{\nu - e_j, j} .$$

The first sum is bounded by θd_ν, therefore

$$\Big(1 - \frac{\theta}{2}\Big) d_\nu \leq \frac{1}{2} \sum_{j:\ \nu_j \neq 0} d_{\nu - e_j, j} .$$

Now summing over all $|\nu| = k$ gives

$$\Big(1 - \frac{\theta}{2}\Big) \sum_{|\nu| = k} d_\nu \leq \frac{1}{2} \sum_{|\nu| = k} \sum_{j:\ \nu_j \neq 0} d_{\nu - e_j, j} = \frac{1}{2} \sum_{|\nu| = k-1} \sum_{j \geq 1} d_{\nu, j} \leq \frac{\theta}{2} \sum_{|\nu| = k-1} d_\nu .$$

Therefore $\sum_{|\nu|=k} d_\nu \leq \kappa \sum_{|\nu|=k-1} d_\nu$ with $\kappa := \frac{\theta}{2-\theta} < 1$, and thus $\sum_{\nu \in \mathcal{F}} d_\nu < \infty$.

Rescaling

Now let $\rho = (\rho_j)_{j\geq 1}$ be any sequence with $\rho_j > 1$ such that $\sum_{j\geq 1} \rho_j |\psi_j| \leq \bar{a} - \delta$ for some $\delta > 0$, or equivalently such that $\left\| \frac{\sum_{j\geq 1} \rho_j |\psi_j|}{\bar{a}} \right\|_{L^\infty(D)} \leq \theta < 1$.

Consider the rescaled solution map $\tilde{u}(y) = u(\rho y)$ where $\rho y := (\rho_j y_j)_{j\geq 1}$ which is the solution of the same problem as u with ψ_j replaced by $\rho_j \psi_j$.

Since (UEA) holds for for these rescaled functions, the previous lemma shows that

$$\sum_{\nu\in\mathcal{F}} \|\tilde{t}_\nu\|_V^2 < \infty,$$

where

$$\tilde{t}_\nu := \frac{1}{\nu!} \partial^\nu \tilde{u}(0) = \frac{1}{\nu!} \rho^\nu \partial^\nu u(0) = \rho^\nu t_\nu.$$

This therefore gives the weighted ℓ^2 estimate

$$\sum_{\nu\in\mathcal{F}} (\rho^\nu \|t_\nu\|_V)^2 \leq C < \infty.$$

In particular, we retrieve the estimate $\|t_\nu\|_V \leq C\rho^{-\nu}$ that was obtained by the complex variable approach, however the above estimate is stronger.

An alternate summability result

Applying Hölder's inequality gives

$$\sum_{\nu\in\mathcal{F}} \|t_\nu\|_V^p \leq \Big(\sum_{\nu\in\mathcal{F}} (\rho^\nu \|t_\nu\|_V)^2\Big)^{p/2} \Big(\sum_{\nu\in\mathcal{F}} \rho^{-q\nu}\Big)^{1-p/2},$$

with $q = \frac{2p}{2-p} > p$, or equivalently $\frac{1}{q} = \frac{1}{p} - \frac{1}{2}$.
The sum in second factor is finite provided that $(\rho_j^{-1})_{j\geq 1} \in \ell^q$. Therefore, the following result holds.

Theorem (Bachmayr-Cohen-Migliorati, 2015) : Let p and q be such that $\frac{1}{q} = \frac{1}{p} - \frac{1}{2}$. Assume that there exists a sequence $\rho = (\rho_j)_{j\geq 1}$ of numbers larger than 1 such that

$$\sum_{j\geq 1} \rho_j |\psi_j| \leq \bar{a} - \delta,$$

for some $\delta > 0$ and

$$(\rho_j^{-1})_{j\geq 1} \in \ell^q.$$

Then $(\|t_\nu\|_V)_{\nu\in\mathcal{F}} \in \ell^p(\mathcal{F})$.

The above conditions ensuring ℓ^p summability of $(\|t_\nu\|_V)_{\nu\in\mathcal{F}}$ are significantly weaker than those in the first summability theorem especially for locally supported ψ_j.

Disjoint supports

Assume that the ψ_j have disjoint supports.

Then with $\delta = \frac{r}{2}$, we choose

$$\rho_j := \min_{x \in D} \frac{\overline{a}(x) - \frac{r}{2}}{|\psi_j(x)|} > 1.$$

so that $\sum_{j \geq 1} \rho_j |\psi_j| \leq \overline{a} - \delta$ holds.

We have

$$b_j := \rho_j^{-1} = \frac{|\psi_j(x)|}{\overline{a}(x) - \frac{r}{2}} \leq \frac{\|\psi_j\|_{L^\infty}}{R - \frac{r}{2}}.$$

Thus in this case, our result gives for any $0 < q < \infty$,

$$(\|\psi_j\|_{L^\infty})_{j \geq 1} \in \ell^q(\mathbb{N}) \implies (\|t_\nu\|_V)_{\nu \in \mathcal{F}} \in \ell^p(\mathcal{F}),$$

with $\frac{1}{q} = \frac{1}{p} - \frac{1}{2}$.

Similar improved results if the ψ_j have supports with limited overlap, such as wavelets.

No improvement in the case of globally supported functions, such as typical KL bases.

Other models

Model 1 : same PDE but no affine dependence, e.g. $a(x,y) = \overline{a}(x) + (\sum_{j\geq 0} y_j \psi_j(x))^2$. Assuming that $\overline{a}(x) \geq r > 0$ guarantees ellipticity uniformly over $y \in U$.

Model 2 : similar problems + non-linearities, e.g.

$$g(u) - \mathrm{div}(a\nabla u) = f \;\; \text{on} \;\; D = D(y) \;\; u_{|\partial D} = 0,$$

with same assumptions on a and f. Well-posedness in $V = H^1_0(D)$ for all $f \in V'$ is ensured for certain nonlinearities, e.g. $g(u) = u^3$ of u^5 in dimension $m = 3$ ($V \subset L^6$).

Model 3 : PDE's on domains with parametrized boundaries, e.g.

$$-\Delta v = f \;\; \text{on} \;\; D = D_y \;\; u_{|\partial D} = 0.$$

where the boundary of D_y is parametrized by y, e.g.

$$D_y := \{(x_1, x_2) \in \mathbb{R}^2 \;\; : \;\; 0 < x_1 < 1 \;\; \text{and} \;\; 0 < x_2 < b(x_1, y)\},$$

where $b = b(x,y) = \overline{b}(x) + \sum_j y_j \psi_j(x)$ satisfies $0 < r < b(x,y) < R$. We transport this problem on the reference domain $[0,1]^2$ and study

$$u(y) := v(y) \circ \phi_y, \quad \phi_y : [0,1]^2 \longrightarrow D_y, \quad \phi_y(x_1, x_2) := (x_1, x_2 b(x_1, y)).$$

which satisfies a diffusion equation with coefficient $a = a(x,y)$ non-affine in y.

Polynomial approximation for these models

In contrast to our guiding example (which we refer to as model 0), bounded holomorphic extension is generally not feasible in a complex domain containing the polydisc $\mathcal{U} = \otimes_{j\geq 1}\{|z_j| \leq 1\}$. For this reason,Taylor series are not expected to converge.

Instead we consider the tensorized Legendre expansion

$$u(y) = \sum_{\nu\in\mathcal{F}} v_\nu L_\nu(y),$$

where $L_\nu(y) := \prod_{j\geq 1} L_{\nu_j}(y_j)$ and $(L_k)_{k\geq 0}$ are the Legendre polynomials normalized in $L^2\left([-1,1], \frac{dt}{2}\right)$.

Thus $(L_\nu)_{\nu\in\mathcal{F}}$ is an orthonormal basis for $L^2(U, V, \mu)$ where $\mu := \otimes_{j\geq 1}\frac{dy_j}{2}$ is the uniform probability measure and we have

$$v_\nu = \int_U u(y)L_\nu(y)d\mu(y).$$

We also consider the L^∞-normalized Legendre polynomials $P_k = (1+2k)^{-1/2}L_k$ and their tensorized version P_ν, so

$$u(y) = \sum_{\nu\in\mathcal{F}} w_\nu P_\nu(y),$$

where $w_\nu := \left(\prod_{j\geq 1}(1+\nu_j)^{1/2}\right)v_\nu$.

Main result

Theorem (Chkifa-Cohen-Schwab, 2013) : For models 0, 1, 2 and 3, and for any $p < 1$,

$$(\|\psi_j\|_X)_{j>0} \in \ell^p(\mathbb{N}) \implies (\|v_\nu\|_V)_{\nu\in\mathcal{F}} \text{ and } (\|w_\nu\|_V)_{\nu\in\mathcal{F}} \in \ell^p(\mathcal{F}).$$

with $X = L^\infty$ for models 0, 1, 2, and $X = W^{1,\infty}$ for model 3.

By the same application of Stechkin's argument as for Taylor expansions, best n-term truncations for the L^∞ normalized expansion converge rate $\mathcal{O}(n^{-s})$ in $L^\infty(U, V)$ where $s = \frac{1}{p} - 1$.

Best n-term truncations for the L^2 normalized expansion converge with rate $\mathcal{O}(n^{-r})$ in $L^2U, V, \mu)$ where $r = \frac{1}{p} - \frac{1}{2}$.

In the particular case of our guiding example, model 0, we can obtain improved summability results for Legendre expansions, similar to Taylor expansions.

Key ingredient in the proof of the above theorem : estimates of Legendre coefficients for holomorphic functions in a "small" complex neighbourhood of U.

Taylor vs Legendre expansions

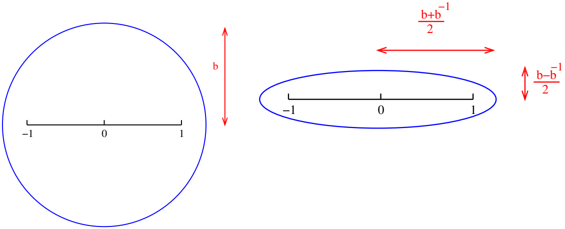
In one variable :

- If u is holomorphic in an open neighbourhood of the disc $\{|z| \leq b\}$ and bounded by M on this disc, then the n-th Taylor coefficient of u is bounded by

$$|t_n| := \left| \frac{u^{(n)}(0)}{n!} \right| \leq Mb^{-n}$$

- If u is holomorphic in an open neighbourhood of the domain $\mathcal{E}_b$ limited by the ellipse of semi axes of length $(b+b^{-1})/2$ and $(b-b^{-1})/2$, for some $b>1$, and bounded by M on this domain, then the n-th Legendre coefficent w_n of u is bounded by

$$|w_n| \leq Mb^{-n}(1+2n)\phi(b), \qquad \phi(b) := \frac{\pi b}{b-1}$$



A general assumption for sparsity of Legendre expansions

We say that the solution to a parametric PDE $\mathcal{D}(u,y)=0$ satisfies the (p,ε)-holomorphy property if and only if there exist a sequence $(c_j)_{j\geq 1}\in\ell^p(\mathbb{N})$, a constant $\varepsilon>0$ and $C_0>0$, such that : for any sequence $\rho=(\rho_j)_{j\geq 1}$ such that $\rho_j>1$ and

$$\sum_{j\geq 1}(\rho_j-1)c_j\leq\varepsilon,$$

the solution map has a complex extension

$$z\mapsto u(z),$$

of the solution map that is holomorphic with respect to each variable on a domain of the form $\mathcal{O}_\rho=\otimes_{j\geq 1}\mathcal{O}_{\rho_j}$ where $\mathcal{O}_{\rho_j}$ is an open neigbourhood of the elliptical domain $\mathcal{E}_{\rho_j}$, with bound

$$\sup_{z\in\mathcal{E}_\rho}\|u(z)\|_V\leq C_0,$$

where $\mathcal{E}_\rho=\otimes_{j\geq 1}\mathcal{E}_{\rho_j}$.

Under such an assumption, one has (up to additional harmless factors) an estimate of the form

$$\|w_\nu\|_V\leq C_0\inf\{\rho^{-\nu}\ \ ;\ \ \rho\ \text{ s.t. }\ \sum_{j\geq 1}(\rho_j-1)c_j\leq\varepsilon\},$$

allowing us to prove that $(\|w_\nu\|_V)_{\nu\in\mathcal{F}}\in\ell^p(\mathcal{F})$.

A general framework for establishing the (p,ε)-holomorphy assumption

Assume a general problem of the form

$$\mathcal{P}(u,a)=0,$$

with $a=a(y)=\overline{a}+\sum_{j\geq 1} y_j\psi_j$, where

$$\mathcal{P}: V\times X\to W,$$

with V,X,W a triplet of complex Banach spaces, and $\overline{a}$ and ψ_j are functions in X.

Theorem (Chkifa-Cohen-Schwab, 2013) : assume that

(i) The problem is well posed for all $a\in Q=a(U)$ with solution $u(y)=u(a(y))\in V$.

(ii) The map $\mathcal{P}$ is differentiable (holomorphic) from $X\times V$ to W.

(iii) For any $a\in Q$, the differential $\partial_u\mathcal{P}(u(a),a)$ is an isomorphism from V to W

(iv) One has $(\|\psi_j\|_X)_{j\geq 1}$ in $\ell^p(\mathbb{N})$ for some $0<p<1$,

Then, for $\varepsilon>0$ small enough, the (p,ε)-holomorphy property holds.

Idea of proof

Based on the holomorphic Banach valued version of the implicit function theorem (see e.g. Dieudonné).

1. For any $a \in Q = \{a(y) \; : \; y \in U\}$ we can find a $\varepsilon_a > 0$ such that the map $a \longrightarrow u(a)$ has an holomorphic extension on the ball $B(a, \varepsilon_a) := \{\tilde{a} \in X \; : \; \|\tilde{a} - a\|_X < \varepsilon_a\}$.

2. Using the decay properties of the $\|\psi_j\|_X$, we find that Q is compact in X. It can be covered by a finite union of balls $B(a_i, \varepsilon_{a_i})$, for $i = 1, \ldots, M$.

3. Thus $a \longrightarrow u(a)$ has an holomorphic extension on a complex neighbourhood $\mathcal{N}$ of Q of the form

$$\mathcal{N} = \cup_{i=1}^{M} B(a_i, \varepsilon_{a_i}).$$

4. For ε small enough, one proves that if $\sum_{j \geq 1} (\rho_j - 1) c_j \leq \varepsilon$ with $c_j := \|\psi_j\|_L$ then with $\mathcal{O}_\rho = \otimes_{j \geq 1} \mathcal{O}_{\rho_j}$ where $\mathcal{O}_b := \{z \in \mathbb{C} \; : \; \mathrm{dist}(z, [-1, 1])_{\mathbb{C}} \leq b - 1\}$ is a neighborhood of $\mathcal{E}_b$, one has

$$z \in \mathcal{O}_\rho \implies a(z) \in \mathcal{N}.$$

This gives holomorphy of $z \mapsto a(z) \mapsto u(z) = u(a(z))$ in each variable for $z \in \mathcal{O}_\rho$.

Lognormal coefficients

We assume diffusion coefficients are given by

$$a = \exp(b),$$

with b a random function defined by an affine expansion of the form

$$b = b(y) = \sum_{j\geq 1} y_j \psi_j,$$

where (ψ_j) is a given family of functions from $L^\infty(D)$ and $y = (y_j)_{j\geq 1}$ a sequence of i.i.d. standard Gaussians $\mathcal{N}(0,1)$ variables.

Thus y ranges in $U = \mathbb{R}^{\mathbb{N}}$ equipped with the probabilistic structure $(U, \mathcal{B}(U), \gamma)$ where $\mathcal{B}(U)$ is the cylindrical Borel Σ-algebra and γ the tensorized Gaussian measure.

Commonly used stochastic model for diffusion in porous media.

The solution $u(y)$ is well defined in V for those $y \in U$ such that $b(y) \in L^\infty(D)$, with

$$\|u(y)\|_V \leq \frac{1}{a_{\min}(y)} \|f\|_{V'} \leq \exp(\|b(y)\|_{L^\infty}) \|f\|_{V'}.$$

Affine Gaussian representations

Given a centered Gaussian process $(b(x))_{x\in D}$ with covariance function $C_b(x,z) = \mathbb{E}(b(x)b(z))$, one frequently consider the Karhunen-Loeve expansion,

$$b = \sum_{j\geq 1} \xi_j \varphi_j,$$

where ξ_j are i.i.d. $\mathcal{N}(0,\sigma_j^2)$ and $(\varphi_j)_{j\geq 1}$ are $L^2(D)$-orthonormal, and normalize

$$\psi_j = \sigma_j \varphi_j \quad \text{and} \quad y_j = \sigma_j^{-1}\xi_j,$$

so that $b = \sum_{j\geq 1} y_j \psi_j$. However, other representations may be relevant.

Example : b the Brownian bridge on $D=[0,1]$ defined by $C_b(x,z) := \min\{x,z\} - xz$.

1. Normalized KL : $\psi_j(x) = \frac{\sqrt{2}}{\pi j}\sin(\pi j x)$.

2. Levy-Ciesielski representation : uses Schauder basis (primitives of Haar system)

$$\psi_{l,k}(x) := 2^{-l/2}\psi(2^l x - k), \quad k = 0,\dots,2^l-1, \quad l \geq 0, \quad \psi(x) := \frac{1}{2}(1-|2x-1|)_+.$$

Then with coarse to fine ordering $\psi_j = \psi_{l,k}$ for $j = 2^l + k$, one has $b = \sum_{j\geq 1} y_j \psi_j$.

Main theoretical questions

1. Integrability : under which conditions is $y \mapsto u(y)$ Bochner measurable with values in V and satifies for $0 \leq k < \infty$.

$$\|u\|_{L^k(U,V,\gamma)}^k = \mathbb{E}(\|u(y)\|_V^k) < \infty,$$

In view of $\|u(y)\|_V \leq \exp(\|b(y)\|_{L^\infty})\|f\|_{V'}$, this holds if $\mathbb{E}(\exp(k\|b(y)\|_{L^\infty}) < \infty$.

2. Approximability : if $u \in L^2(U,V,\gamma)$, consider the multivariate Hermite expansion

$$u = \sum_{\nu\in\mathcal{F}} u_\nu H_\nu, \quad H_\nu(y) := \prod_{j\geq 1} H_{\nu_j}(y_j) \quad \text{and} \quad u_\nu := \int_U u(y)H_\nu(y)d\gamma(y)$$

where $\mathcal{F}$ is the set of finitely supported integer sequences $\nu = (\nu_j)_{j\geq 1}$.

Best n-term approximation : $u_n = \sum_{\nu\in\Lambda_n} u_\nu H_\nu$, with Λ_n indices of n largest $\|u_\nu\|_V$.

Stechkin lemma : if $(\|u_\nu\|_V)_{\nu\in\mathcal{F}} \in \ell^p(\mathcal{F})$ for some $0 < p < 2$ then

$$\|u - u_n\|_{L^2(U,V,\gamma)} \leq Cn^{-s}, \quad s := \frac{1}{p} - \frac{1}{2}, \quad C := \|(\|u_\nu\|_V)_{\nu\in\mathcal{F}}\|_{\ell^p}$$

Existing results

Integrability : sufficient conditions for $u \in L^k(U, V, \gamma)$ for all $0 \leq k < \infty$ are known.

1. Smoothness : $C_b \in C^\alpha(D \times D)$ for some $\alpha > 0$ (Charrier).
2. Summability : $\sum_{j \geq 1} \|\psi_j\|_{L^\infty} < \infty$ (Schwab-Gittelson-Hoang)
3. $\sum_{j \geq 1} \|\psi_j\|_{L^\infty}^{2-\delta} \|\psi_j\|_{C^\alpha}^{\delta} < \infty$ for some $0 < \delta < 1$ (Dashti-Stuart)

Approximability : first available result is as follows.

Theorem (Hoang-Schwab, 2014) : for any $0 < p \leq 1$, if $(j\|\psi_j\|_{L^\infty}) \in \ell^p(\mathbb{N})$ then $(\|u_\nu\|_V) \in \ell^p(\mathcal{F})$.

Remarks :

The condition $(j\|\psi_j\|_{L^\infty}) \in \ell^p(\mathbb{N})$ is strong, compared to L^2-integrability conditions.

It typically imposes high order of smoothness of the covariance function.

For example it is not satified by KL or Schauder representation of Brownian bridge.

Condition based on $\|\psi_j\|_{L^\infty}$. Can we better exploit the support properties ?

Improved summability result

Theorem (Bachmayr-Cohen-DeVore-Migliorati, 2015) :

Let $0 < p < 2$ and define $q := q(p) = \frac{2p}{2-p} > p$ (or equivalently $\frac{1}{q} = \frac{1}{p} - \frac{1}{2}$).

Assume that there exists a positive sequence $\rho = (\rho_j)_{j\geq 1}$ such that

$$(\rho_j^{-1})_{j\geq 1} \in \ell^q(\mathbb{N}) \quad \text{and} \quad \sup_{x\in D} \sum_{j\geq 1} \rho_j |\psi_j(x)| < \infty.$$

Then $y \mapsto u(y)$ is measurable and belongs $L^k(U, V, \gamma)$ for all $0 \leq k < \infty$ and

$$(\|u_\nu\|_V)_{\nu\in\mathcal{F}} \in \ell^p(\mathcal{F}).$$

Remarks :

Similar result for the Taylor and Legendre coefficients for the affine parametric model $a(y) = \overline{a} + \sum_{j\geq 1} y_j\psi_j$ however by different arguments.

Proof is rather specific to the linear diffusion equation (yet extensions possible).

The above conditions for ℓ^p summability of $(\|u_\nu\|_V)_{\nu\in\mathcal{F}}$ are weaker than ℓ^p summability of $(j\|\psi_j\|_{L^\infty})_{j\geq 1}$ especially for locally supported ψ_j.

The case of the Brownian bridge

KL representation :

Globally supported functions $\psi_j(x) = \frac{\sqrt{2}}{\pi j}\sin(\pi j x)$.

The decay of $(\|\psi_j\|_{L^\infty})_{j\geq 1}$ is not sufficient to apply our results.

No provable approximability by best n-term Hermite series.

Schauder representation :

Wavelet type functions with decay in scale $\|\psi_\lambda\|_{L^\infty} \sim 2^{-l/2}$.

This allows to apply our result $\rho_\lambda = 2^{\beta l}$, for any $\beta < \frac{1}{2}$.

Our result imply that $(\|u_\nu\|_V)_{\nu\in\mathcal{F}} \in \ell^p(\mathcal{F})$ for any p such that $\frac{1}{2} > \frac{1}{p} - \frac{1}{2}$.

In particular, best n-term Hermite approximations satisfy

$$\|u - u_{\Lambda_n}\|_{L^2(U,V,\gamma)} \leq Cn^{-s}, \quad s = \frac{1}{p} - \frac{1}{2} < \frac{1}{2}.$$

Representations of gaussian processes

Objective : for general gaussian processes b, identify an "optimal" representation for solving the approximation problem.

By analogy with the Brownian bridge, one expect a wavelet type basis.

Existing work in this direction : Cieceslki-Kerkycharian-Roynette, Benassi-Jaffard-Roux, Kerkyacharian-Ogawa-Petrushev-Picard.

Luschgy-Pages : $b = \sum_{j\geq 1} y_j\psi_j$ with y_j i.i.d. $\mathcal{N}(0,1)$ iff $(\psi_j)_{j\geq 1}$ is a tight frame for the reproducing kernel Hilbert space $\mathcal{H}$ defined by the covariance function of the process.

Matérn processes : covariance given by $K(x,x') = k(x-x')$ with

$$k(x) = \frac{2^{1-\nu}}{\Gamma(\nu)}\left(\frac{\sqrt{2\nu}|x|}{\lambda}\right)^{\nu} K_\nu\left(\frac{\sqrt{2\nu}|x|}{\lambda}\right),$$

where $\nu,\lambda > 0$ and K_ν is the modified Bessel function of the second kind. One has

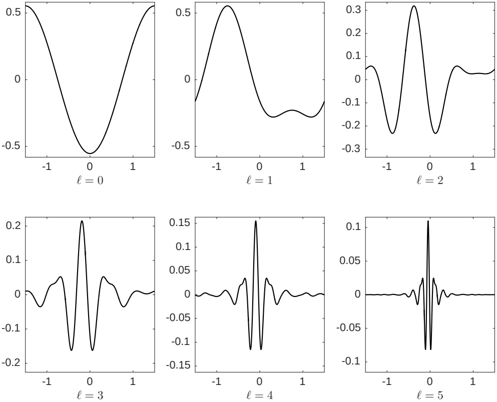
$$\hat{k}(\omega) = c_{\nu,\lambda}\left(\frac{\nu}{2\pi^2\lambda^2} + |\omega|^2\right)^{-(\nu+d/2)}, \quad c_{\nu,\lambda} := \frac{\Gamma(\nu+d/2)(2\nu)^\nu}{\pi^{2\nu+d/2}\Gamma(\nu)\lambda^{2\nu}}.$$

The associated RKHS is $\mathcal{H} \sim H^r(D)$ with $r = \nu + d/2$.

Matern wavelets

Bachmayr-Cohen-Migliorati (2016) : construction of expansions into wavelets type frames satisfying the expected decay $2^{-\nu l}$ with scale level. These expansions lead to better approximations than KL expansions which have decay $j^{-r/d}$ but global support.

Example : $\lambda = 1$, $\nu = \frac{1}{2}$



Main ingredient in the proof of the main result

1. Relate Hermite coefficients u_ν and partial derivatives $\partial^\mu u$. Base on 1-d Rodrigues formula : $H_n(t) = \frac{(-1)^n}{\sqrt{n!}} \frac{g^{(n)}(t)}{g(t)}$, where $g(t) := (2\pi)^{-1/2} \exp(-t^2/2)$. After some computation this leads to weighted ℓ^2 identity for any sequence $\rho := (\rho_j)_{j\geq 1}$.

$$\sum_{\|\mu\|_{\ell^\infty}\leq r} \frac{\rho^{2\mu}}{\mu!} \int_U \|\partial^\mu u(y)\|_V^2 \, d\gamma(y) = \sum_{\nu\in\mathcal{F}} b_\nu \|u_\nu\|_V^2,$$

where $b_\nu := \sum_{\|\mu\|_{\ell^\infty}\leq r} \binom{\nu}{\mu} \rho^{2\mu}$.

2. Prove finiteness of left hand side $\sum_{\|\mu\|_{\ell^\infty}\leq r} \frac{\rho^{2\mu}}{\mu!} \int_U \|\partial^\mu u(y)\|_V^2 \, d\gamma(y)$ when

$$\sup_{x\in D} \sum_{j\geq 1} \rho_j |\psi_j(x)| =: K < C_r := r^{-1/2} \ln 2.$$

Use PDE : $\int_D a(y)\nabla\partial^\mu u(y)\cdot\nabla v = -\sum_{\nu\leq\mu,\, \nu\neq\mu} \binom{\mu}{\nu} \int_D \psi^{\mu-\nu} a(y)\nabla\partial^\nu u(y)\cdot\nabla v$.

3. Derive ℓ^p estimate by mean of Hölder's inequality :

$$\Big(\sum_{\nu\in\mathcal{F}} \|u_\nu\|_V^p\Big)^{1/p} \leq \Big(\sum_{\nu\in\mathcal{F}} b_\nu \|u_\nu\|_V^2\Big)^{1/2} \Big(\sum_{\nu\in\mathcal{F}} b_\nu^{-q/2}\Big)^{1/q}.$$

We prove that the second factor is finite if $(\rho_j^{-1})_{j\geq 1} \in \ell^q(\mathbb{N})$ and r such that $\frac{2}{r+1} < p$.

Conclusions

The curse of dimensionality can be “defeated” by exploiting both smoothness and anisotropy in the different variables.

For certain models, this can be achieved by sparse polynomial approximations.

The way we parametrize the problem, or represent its solution, is crucial.

Part 2

Algorithms : Galerkin, power series, sparse interpolation and least-squares

From approximation results to numerical methods

The results so far are approximation results. They say that for several models of parametric PDEs, the solution map $y \mapsto u(y)$ can be accurately approximate (with rate n^{-s} for some $s > 0$) by multivariate polynomials having n terms.

These polynomials are computed by best n-term truncation of Taylor or Legendre or Hermite series, but this is not feasible in practical numercial methods.

Problem 1 : the best n-term index sets Λ_n are computationally out of reach. Their identification would require the knowledge of all coefficients in the expansion.

Objective : identify non-optimal yet good sets Λ_n.

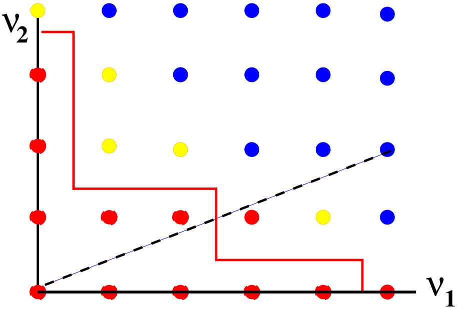
Problem 2 : the exact polynomial coefficients t_ν (or v_ν, w_ν, u_ν) of u for the indices $\nu \in \Lambda_n$ cannot be computed exactly.

Objective : numerical strategy for approximately computing polynomial coefficients.

Numerical methods : strategies to build the sets Λ_n

(i) Non-adaptive, based on the available a-priori estimates for the $\|t_\nu\|_V$ (or $\|v_\nu\|_V$, $\|w_\nu\|_V$, $\|u_\nu\|_V$). Take Λ_n to be the set corresponding to the n largest such estimates.

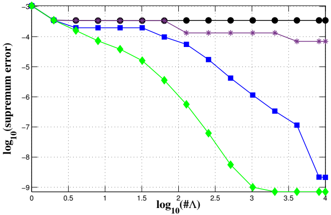
(ii) Adaptive, based on a-posteriori information gained in the computation $\Lambda_1 \subset \Lambda_2 \subset \cdots \subset \Lambda_n \cdots$.



Adaptive vs non-adaptive

Adaptive methods are known to converge better than non-adaptive ones, but their analysis is more difficult.

A test case for linear-affine model in dimension $d = 64$: comparison between the approximation performance with Λ_n given by standard choices $\{\sup \nu_j \leq k\}$ (black) or $\{\sum \nu_j \leq k\}$ (purple) and by anisotropic choices based on a-priori bounds (blue) or adaptively generated (green).



Highest polynomial degree for Λ_{1000} with different choices : 1, 2, 162 and 114.

Downward closed index sets

For adaptive algorithms it is critical that the index chosen sets are downward closed

$$\nu \in \Lambda \text{ and } \mu \leq \nu \implies \mu \in \Lambda,$$

where $\mu \leq \nu$ means that $\mu_j \leq \nu_j$ for all $j \geq 1$.

Such sets are also called lower sets. This property does not generally holds for the sets corresponding to the n largest estimates, however the same convergence rates as proved in the approximation theorems, can be proved when imposing such a structure.

If Λ is downward closed, we consider the polynomial space

$$\mathbb{P}_\Lambda = \mathrm{span}\{y \to y^\nu \,:\, \nu \in \Lambda\} = \mathrm{span}\{L_\nu \,:\, \nu \in \Lambda\} = \mathrm{span}\{H_\nu \,:\, \nu \in \Lambda\}$$

and its V-valued version

$$V_\Lambda := \{\sum_{\nu \in \Lambda} v_\nu y^\nu \,:\, v_\nu \in V\} = V \otimes \mathbb{P}_\Lambda.$$

After having selected Λ_n we search for a computable approximation of u in V_{Λ_n}.

Spatial discretization

Note that $\dim(V_{\Lambda_n}) = \infty$. In practice we use $V_{\Lambda_n,h} = V_h \otimes \mathbb{P}_{\Lambda_n}$ which has dimension

$$n_{tot} = \dim(V_{\Lambda_n,h}) = \dim(V_h)\dim(\mathbb{P}_{\Lambda_n}) = n_h n < \infty.$$

This amount in applying polynomial approximation to the approximate solution map $y \mapsto u_h(y) \in V_h$, defined e.g. by the Galerkin method

$$\int_D \nabla a(y) u_h(y) \nabla v_h = \int_D f v_h, \quad v_h \in V_h,$$

Total approximation error estimate e.g. in L^∞

$$\varepsilon_{tot} = \min_{v \in V_{\Lambda_n,h}} \|u - v\|_{L^\infty} \leq \varepsilon_n + \varepsilon_h,$$

where

$$\varepsilon_n = \min_{v \in V_{\Lambda_n,h}} \|u_h - v\|_{L^\infty(U,V_h)} \quad \text{and} \quad \varepsilon_h = \sup_{y \in U} \|u(y) - u_h(y)\|_V.$$

By the same sparsity analysis as for the exact solution map we obtain estimates of the form $\varepsilon_n \leq C n^{-s}$. The spatial error is controlled provided that $u(y)$ has additional spatial smoothness : $\varepsilon_h \leq C h^t \leq C n_h^{-t/m}$ if $u \in L^\infty(U, H^{1+t}(D))$ and $D \subset \mathbb{R}^m$.

Balancing with $n_h \sim n^{-ms/t}$, this leads to the estimate $\varepsilon_{tot} \leq C n_{tot}^{-\frac{s}{1+ms/t}}$.

Sharper analysis of spatial discretization

For a given truncated series, we could use different spatial resolution in the discretization of each coefficients. For example, for the Taylor series, use

$$\sum_{\nu\in\Lambda_n} t_{\nu,h_\nu} y^\nu, \quad t_{\nu,h_\nu} \in V_{h_\nu}.$$

The total number of degrees of freedom is now $n_{tot} = \sum_{\nu\in\Lambda_n} n_{h_\nu} \sim \sum_{\nu\in\Lambda_n} h_\nu^{-d}$ and the total error is of the order $\varepsilon_{tot} \le \varepsilon_n + \sum_{\nu\in\Lambda_n} \varepsilon_\nu$, where

$$\varepsilon_n = \min_{v\in V_{\Lambda_n}} \|u-v\|_{L^\infty(U,V)} \le Cn^{-s} \quad \text{and} \quad \varepsilon_\nu = \min_{v_h\in V_{h_\nu}} \|t_\nu - v_h\|_V \le Ch_\nu^t \|t_\nu\|_{H^{1+t}}.$$

The error analysis now relates to the sparsity of the sequence $(\|t_\nu\|_{H^{1+t}})_{\nu\in\mathcal{F}}$.

Theorem (Cohen-DeVore-Schwab, 2011) : If $(\|t_\nu\|_{H^{1+t}})_{\nu\in\mathcal{F}} \in \ell^p(\mathcal{F})$ for some $p<1$, then optimal tuning of the h_ν gives the total error $\varepsilon_{tot} \le Cn^{-\min\{s,\frac{t}{m}\}}$ with $s = \frac{1}{p} - 1$.

Theorem (Bachmayr-Cohen-Dinh-Schwab, 2016) : for the affine diffusion model, let p and q be such that $\frac{1}{q} = \frac{1}{p} - \frac{1}{2}$. Assume that the domain D is smooth or convex and that there exists a sequence $\rho = (\rho_j)_{j\ge1}$ of numbers larger than 1 such that

$$\sum_{j\ge1} \rho_j |\psi_j| \le \overline{a} - \delta \quad \text{and} \quad \sum_{j\ge1} \rho_j |\nabla\psi_j| < \infty$$

for some $\delta > 0$ and $(\rho_j^{-1})_{j\ge1} \in \ell^q$. Then $(\|t_\nu\|_{H^2})_{\nu\in\mathcal{F}} \in \ell^p(\mathcal{F})$.

Strategies to build the polynomial approximation : intrusive methods

1. Galerkin method : based on a space-parameter variational form (test the parametric PDE on arbitrary $y \mapsto v(y)$ and integrate both in x and y). Example for model 0 : find $u \in L^2(U, V, \mu)$ such that for all $v \in L^2(U, V, \mu)$,

$$A(u, v) := \int_U \int_D a(x, y) \nabla u(x, y) \nabla v(x, y) dx d\mu(y) = \int_U \langle f, v(y) \rangle d\mu(y) =: L(v),$$

The problem is coercive in $L^2(U, V, \mu)$. Galerkin formulation : find $u_n \in V_{\Lambda_n}$ such that

$$A(u_n, v_n) = L(v_n), \quad v_n \in V_{\Lambda_n}.$$

Cea's lemma gives error estimate

$$\|u - u_n\|_{L^2(U,V,\mu)} \leq (R/r)^{1/2} \min_{v \in V_{\Lambda_n}} \|u - v\|_{L^2(U,V,\mu)}.$$

After space discretization, Galerkin problem in $V_{\Lambda_n,h}$ gives a $(nn_h) \times (nn_h)$ system.

Lognormal case : lack of coervivity, Galerkin method needs some massaging.

2. Exact computation of the Taylor coefficients $\|t_\nu\|_V$, based on the recursive formula.

After space discretization, sequence of n systems of size $n_h \times n_h$.

Adaptive algorithms with optimal theoretical guarantees exist for both method 1 (Gittelson-Schwab) and 2 (Chkifa-Cohen-DeVore-Schwab).

These methods apply to other models, however mainly confined to linear PDEs, with affine parameter dependence.

Exact adaptive computation of the Taylor coefficients

With e_j the Kroenecker sequence of index j,

$$\int_D \bar{a} \nabla t_\nu \nabla v = -\sum_{j:\ \nu_j \neq 0} \int_D \psi_j \nabla t_{\nu - e_j} \nabla v, \quad v \in V.$$

If Λ_n is downward closed, this allows us to compute all t_ν by recursively solving n boundary value problems, or $n_h \times n_h$ systems after space discretization in V_h.

Adaptive method : start with $\Lambda_1 = \{0\}$. Given that we have computed Λ_k and the $(t_\nu)_{\nu \in \Lambda_k}$ we compute the t_ν for ν in the margin

$$\mathcal{M}(\Lambda_k) = \mathcal{M}_k := \{\nu \notin \Lambda_k\,;\, \nu - e_j \in \Lambda_k \ \text{ for some } j\},$$

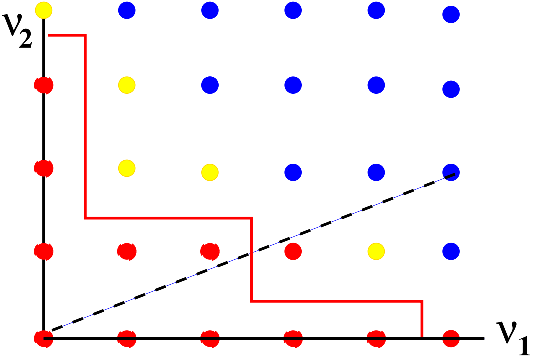
and build the new set by bulk search : choose $\Lambda_{k+1} = \Lambda_k \cup \mathcal{S}_k$, with $\mathcal{S}_k \subset \mathcal{M}_k$ smallest such that $\sum_{\nu \in \mathcal{S}_k} \|t_\nu\|_V^2 \geq \theta \sum_{\nu \in \mathcal{M}_k} \|t_\nu\|_V^2$, for a fixed $\theta \in]0,1[$.

Key property (saturation) : under (UEA), for any lower set Λ there exists a constant C such that

$$\sum_{\nu \notin \Lambda} \|t_\nu\|_V^2 \leq C \sum_{\nu \in \mathcal{M}(\Lambda)} \|t_\nu\|_V^2.$$

This guarantees ℓ^2 error reduction by fixed factor at each step $k \to k+1$.

In addition, can be proved to converge with optimal convergence rate $\#(\Lambda_k)^{-s}$.


v_2
v_1

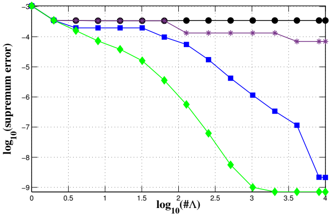
Test case in high dimension $d = 64$

Physical domain $D = [0,1]^2 = \cup_{j=1}^{d} D_j$.

Diffusion coefficients $a(x,y) = 1 + \sum_{j=1}^{d} y_j \left(\frac{0.9}{j^2}\right) \chi_{D_j}$. Thus $U = [-1,1]^{64}$.

Adaptive search of Λ implemented in C++, spatial discretization by FreeFem++.

Comparison between the Λ_k generated by the adaptive algorithm (green) and non-adaptive choices $\{\sup \nu_j \leq k\}$ (black) or $\{\sum \nu_j \leq k\}$ (purple) or k largest a-priori bounds on the $\|t_\nu\|_V$ (blue).



Highest polynomial degree with $\#(\Lambda) = 1000$ coefficients : 1, 2, 162 and 114.

Computation of the average solution

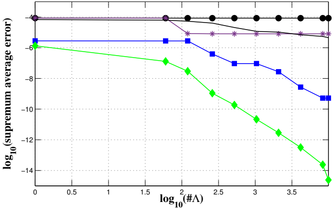
Assuming that y is uniformly distributed on $U = [-1,1]^{64}$, we compute the average solution

$$\bar{u} = \mathbb{E}(u),$$

either by the deterministic approach

$$\bar{u}_\Lambda := \sum_{\nu\in\Lambda} t_\nu \mathbb{E}(y^\nu), \quad \mathbb{E}(y^\nu) = \prod_{j\geq 1}\Big(\int_{-1}^{1} t^{\nu_j}\,\frac{dt}{2}\Big) = \prod_{j\geq 1}\frac{1+(-1)^{\nu_j}}{2+2\nu_j},$$

or by the Monte Carlo approach $\bar{u}_n := \frac{1}{n}\sum_{i=1}^{n} u(y^i)$, where $y^1, \cdots, y^n$ are n independent realization of y.



Error curves in terms of number of solved bvp (MC in full line).

Strategies to build the polynomial approximation : non-intrusive methods

Based on snapshots $u(y^i)$ where $y^i \in U$ for $i = 1, \dots, m$..

1. Pseudo spectral methods : computation of $\sum_{\nu\in\Lambda_n} v_\nu L_\nu$ by quadrature

$$v_\nu = \int_U u(y)L_\nu(y)d\mu(y) \approx \sum_{i=1}^m w_i u(y^i)L_\nu(y^i).$$

2. Interpolation : with $m = n = \dim(\mathbb{P}_{\Lambda_n})$ search for $u_n = I_{\Lambda_n} u \in V_{\Lambda_n}$ such that

$$u_n(y^i) = u(y^i), \quad i = 1, \dots, n.$$

3. Least-squares : with $m \geq n$, search for $u_n \in V_{\Lambda_n}$ minimizing

$$\sum_{i=1}^m \|u(y^i) - u_n(y^i)\|_V^2.$$

4. Underdetermined least-squares : with $m < n$ search for $u_n \in V_{\Lambda_n}$ minimizing

$$\sum_{i=1}^m \|u(y^i) - u_n(y^i)\|_V^2 + \pi(u_n),$$

where π is a penalization functional. Compressed sensing : take for π the (weighted) ℓ^1 sum of V-norms of Legendre coefficients of u_n (promote sparse solutions).

Advantages of non-intrusive methods

Applicable to a broad range of models, in particular non-linear PDEs.

Adaptive algorithms seem to work well for the interpolation and least squares approach, however with no theoretical guarantees.

Additional prescriptions for non-intrusive methods :

(i) Progressive : enrichment $\Lambda_n \longrightarrow \Lambda_{n+1}$ requires only one or a few new snapshots.

(ii) Stable : moderate growth with n of the norm of the reconstruction operator (Lebesgue constant in the case of interpolation).

Main issue : how to best choose the point y^i ?

In the following we focus on interpolation and least-squares, which we present for simplicity for scalar valued functions (extension to V or V_h valued function is trivial).

Concepts from Information Based Complexity

We consider general recovery algorithms of the form

$$A_n(u) := A_n(u(y^1), \dots, u(y^n)),$$

where A_n is a linear or non-linear map from $\mathbb{R}^n$ to a space of functions defined on U.

Non-adaptive algorithms : $y^1, \dots, y^n$ are a-priorily chosen.

Adaptive algorithms : the choice of y^k may depend on $\{u(y^1), \dots, u(y^{k-1})\}$

We measure the error in a given norm : $\|u - A_n(u)\|_X$.

If $X = L^\infty$ no loss in imposing interpolation : $A_n(u)(y^i) = u(y^i)$ for $i = 1, \dots, n$.

Optimal recovery performance over a class $\mathcal{K} \subset X$ (worst case scenario) :

$$e_n = e_n(\mathcal{K})_X := \inf_{A_n, (y^i)} \sup_{u \in \mathcal{K}} \|u - A_n(u)\|_X.$$

Objective : computationally feasible algorithm that perform almost as good as e_n.

Similar concepts developped for the tasks of integration and optimization.

Additional prescriptions

Nestedness : A_{n+1} uses the same points as A_n and one new point y^{n+1}

Stability : a small perturbation of u in X induces a small perturbation of $A_n(u)$ in X

Universality : A_n is simultaneously near-optimal for a large range of model classes.

Curse of dimensionality

Example : $X = L^\infty([-1,1]^d)$ and $\mathcal{K}$ is the unit ball of $C^m([-1,1]^d)$, then

$$cn^{-m/d} \leq e_n(\mathcal{K})_X \leq Cn^{-m/d}, \quad n \geq 0,$$

where $c, C > 0$ depend on (m, d).

Does adaptivity helps ?

Theorem (Bakhvalov, 1971) : If $\mathcal{K}$ is a convex symmetric set of X, there exist a near best algorithm A_n which is non-adaptive and linear.

Yet in many practical applications, adaptivity appears to be very helpful, for instance to detect sharp features or anisotropies in high dimension.

The relevant classes $\mathcal{K}$ are not well understood, especially in high dimensions. They are expected to be non-convex.

A commonly used non-polynomial method : RKHS interpolation

Given a set of point $\{y^1, \dots, y^n\}$, there are infinitely many functions that admit the values $\{u(y^1), \dots, u(y^n)\}$ at these points.

Some a-priori information needs to be injected in order to make a choice. One way to do this is through the minimization of a certain energy among all possible candidates.

Remark : in the univariate case, the piecewise linear and cubic spline interpolants on an interval I minimize the elastic and torsion energies $\int_I |v'|^2$ and $\int_I |v''|^2$, respectively.

Reproducing Kernel Hilbert Space (RKHS) : a Hilbert space of function $\mathcal{H}$ defined on some domain U that is continuously embedded in the space of continuous function $C(U)$ (in our case $U = [-1,1]^d$ or $[-1,1]^{\mathbb{N}}$).

We assume that the space is rich enough such that for all $\{y^1, \dots, y^n\}$ and values $\{v_1, \dots, v_n\}$ there exists $v \in \mathcal{H}$ such that $v(y^i) = v_i$ for $i = 1, \dots, n$.

Example : Sobolev space $H^s(U)$ with $s > d/2$.

RKHS interpolation (Kimmeldorf-Wahba, 1971, Duchon, 1977) : define interpolant as

$$I_n u = I_{\{y^1, \dots, y^n\}} u := \operatorname{argmin}\Big\{ \|v\|_{\mathcal{H}} \ : \ v(y^i) = u(y^i), \quad i = 1, \dots, n \Big\}.$$

This minimizer turns out to be easily computable.

The reproducing kernel

For any $y \in U$, there exists $K_y \in \mathcal{H}$ such that

$$\langle K_y, v \rangle_{\mathcal{H}} = v(y), \quad v \in \mathcal{H}.$$

The functions $(K_y)_{y \in U}$ are complete in $\mathcal{H}$. We define the reproducing kernel (RK) as

$$K(y,z) := \langle K_y, K_z \rangle_{\mathcal{H}} = K_y(z) = K_z(y)$$

The RK satisfies the positive definiteness property

$$\sum_{i=1}^{n} \sum_{j=1}^{n} K(y^i, y^j) c_i c_j > 0, \quad (c_1, \dots, c_n) \neq (0, \dots, 0), \quad y^1, \dots, y^n \in U, \quad n \geq 0.$$

Conversly (Aronszajn, 1950), a function K satisfying this property generates a RKHS $\mathcal{H} = \mathcal{H}_K$ defined as the closure of the linear combinations of the functions $K_y = K(y, \cdot)$ for the norm induced by the inner product $\langle K_y, K_z \rangle_{\mathcal{H}} := K(y,z)$.

Radial basis functions (RBF) : if RK is of the form $K(y,z) = k(|y-z|)$, the functions K_y is the translate at y of the radial function $z \mapsto k(|z|)$ (e.g. Gaussian $e^{-a|z|^2}$).

Computation of the RKHS interpolation

It is then easily seen that the RKHS interpolation is of the form $I_n u = \sum_{j=1}^n c_j K_{y^j}$, where $(c_1, \dots, c_n)$ is the unique solution to the system

$$\sum_{j=1}^n K(y^i, y^j) c_j = u(y^i), \quad i = 1, \dots, n.$$

RKHS interpolation is formally equivalent to gaussian process interpolation which was introduced in geostatistics engineering as Kriging (Matheron, 1978).

For a given positive definite kernel K we consider a centered gaussian process v with covariance $K(y, z)$ (reflects the uncertainty on the unknown u).

Then I_n can be defined as the conditional expectation

$$I_n(y) = \mathbb{E}\Big(v(y) \,\Big|\, v(y^j) = u(y^j),\ j = 1, \dots, n\Big).$$

It is also the best linear estimator $I_n u(y) = \sum_{j=1}^n a_j(y) u(y^j)$ minimizing among all $a_1, \dots, a_n$ the mean square error $\mathbb{E}\Big(|u(y) - \sum_{j=1}^n a_j u(y^j)|^2\Big)$

Adaptive strategies

The gaussian process interpretation leads to natural strategies for adaptive algorithms :

1. Point selection : given $y^1, \dots, y^n$, choose y^{n+1} where $\mathbb{E}(|u(y) - I_n u(y)|^2)$ is largest.

2. Kernel adaptation : using cross-validation, for example with anisotropic gaussians

$$K(y,z) = K_b(y,z) = \exp\Big(-\sum_{j=1}^{d} b_j |y_j - z_j|^2\Big), \quad b = (b_1, \dots, b_d),$$

find b which minimizes $\sum_{i=1}^n \Big| u(y^i) - I_{\{y^1,\dots,y^n\}-\{y^i\}} u(y^i) \Big|^2$

Not much is known on the analysis of these stragegies (need relevant model classes).

Works in arbitrarily high dimension, however costful in moderately large dimension due to the two above non-convex optimization problems.

The system is always solvable but is often ill-conditionned.

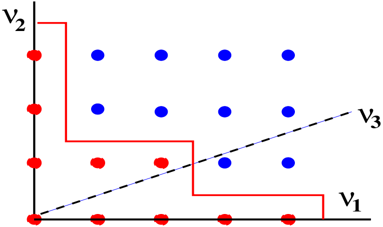
Sparse polynomial interpolation

We want to use general multivariate polynomial spaces of the form

$$\mathbb{P}_\Lambda = \text{span}\{y \mapsto y^\nu \; : \; \nu = (\nu_j)_{j\geq 1} \in \Lambda\}, \quad y^\nu := \prod_{j\geq 1} y_j^{\nu_j}.$$

We assume that Λ is a lower set :

$$\nu \in \Lambda \text{ and } \mu \leq \nu \implies \mu \in \Lambda.$$



Motivation : for relevant classes of functions u arising from parametric PDEs, there exists sequences of lower sets $(\Lambda_n)_{n\geq 0}$ such that for some $s > 0$,

$$\min_{v\in\mathbb{P}_{\Lambda_n}} \|u - v\|_{L^\infty(U)} \leq Cn^{-s}.$$

Univariate nested interpolation

Let $\{t_0, t_1, t_2 \dots\}$, be an infinite sequence of pairwise distinct points in $[-1, 1]$ and let I_k be the univariate interpolation operator on $\mathbb{P}_k$ associated to the section $\{t_0, \dots, t_k\}$.

Hierarchical (Newton) form :

$$I_k = \sum_{l=0}^{k} \Delta_l, \quad \Delta_l := I_l - I_{l-1} \quad \text{and} \quad I_{-1} := 0.$$

Note that $\Delta_k \mathbb{P}_l = 0$ and $\Delta_k u(t_l) = 0$ for all $l < k$. Expansion in a hierarchical basis

$$\Delta_l u = \alpha_l h_l, \quad \alpha_l := u(t_l) - I_{l-1} u(t_l) \ \text{ and } \ h_l(t) = \prod_{j=0}^{l-1} \frac{t - t_j}{t_l - t_j}.$$

The choice of $\{t_0, t_1, t_2 \dots\}$ is important for stability. The usual choices, such as Chebychev or Clemshaw-Curtis, are not section of a single infinite sequence.

Leja points : initialize with arbitrary t_0, usually $t_0 = 1$, then

$$t_l := \text{argmax}_{t \in [-1,1]} \prod_{j=0}^{l-1} |t - t_l|.$$

Note that this choice ensures $\|h_l\|_{L^\infty} \leq 1$. Close to Fekete points $\text{argmax} \prod_{j \neq l} |t_j - t_l|$.

Tensorization

Tensorized grid : for any multi-index ν, we define the point

$$z_\nu := (t_{\nu_1}, t_{\nu_2}, \dots) \in U.$$

Tensorized operators : for any multi-index μ, we define

$$I_\mu = \otimes_{j\geq 1} I_{\mu_j} \quad \text{and} \quad \Delta_\mu := \otimes_{j\geq 1} \Delta_{\mu_j}.$$

I_μ is the interpolation operator on the space of polynomials of degree μ_j in each y_j

$$\mathbb{P}_\mu = \mathbb{P}_{R_\mu}, \quad R_\mu = \{\nu \ : \ \nu \leq \mu\},$$

associated to the grid of point

$$\Gamma_{R_\nu} := \{z_\nu \ : \ \nu \in R_\mu\}.$$

Observe that

$$I_\mu = \otimes_{j\geq 1}\Big(\sum_{l=0}^{\mu_j} \Delta_l\Big) = \sum_{\nu\leq\mu} \Delta_\nu = \sum_{\nu\in R_\mu} \Delta_\nu.$$

Sparsification

Theorem (Cohen-Chkifa-Schwab, 2013, Dyn-Floater, 2013, Kuntzmann, 1959) : if Λ is any downward closed set, the grid

$$\Gamma_\Lambda := \{z_\nu \ : \ \nu \in \Lambda\},$$

is unisolvent for $\mathbb{P}_\Lambda = \mathrm{span}\{y \mapsto y^\nu \ : \ \nu \in \Lambda\}$ and the interpolant is given by

$$I_\Lambda := \sum_{\nu\in\Lambda} \Delta_\nu, \ \ \Delta_\nu := \otimes_{j\geq 1}\Delta_{\nu_j}.$$

Proof : Γ_Λ has the right cardinality, it suffices to prove that $I_\Lambda u(z_\mu) = u(z_\mu)$ for any $\mu \in \Lambda$. This follows from

$$I_\Lambda = I_\mu + \sum_{\nu\in\Lambda \ \nu\not\leq\mu} \Delta_\nu.$$

and observe that $I_\mu u(z_\mu) = u(z_\mu)$ and $\Delta_\nu u(z_\mu) = 0$ if $\nu \not\leq \mu$.

Hierarchical computation

With the tensorized hierarchical basis $H_\nu(y) = \prod_{j\geq 1} h_{\nu_j}(y_j)$, we have

$$\Delta_\nu u(y) = \alpha_\nu H_\nu(y).$$

where the coefficients α_ν can be computed recursively.

Write $\Lambda = \Lambda_n = \{\nu^1, \dots, \nu^n\}$ where the enumeration is such that $\Lambda_k = \{\nu^1, \dots, \nu^k\}$ is downward closed for all $k = 1, \dots, n$. Then

$$\alpha_{\nu^k} = u(z_{\nu^k}) - I_{\Lambda_{k-1}} u(z_{\nu^k}).$$

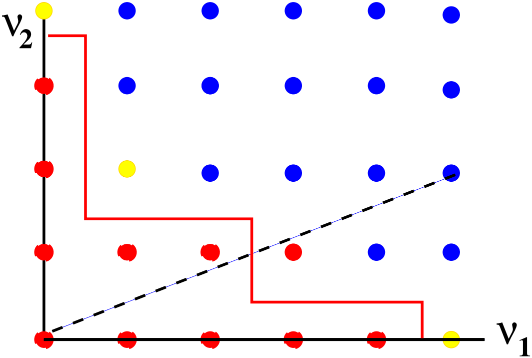
Remark : the same general principles (tensorization, sparsification, hierarchical computation) apply to any other systems such as trigonometric polynomials or hierarchical piecewise linear finite elements.

Adaptive algorithms

Given Λ, we consider its set of neighbors $\mathcal{N}(\Lambda)$ consisting of those $\nu \notin \Lambda$ such that $\Lambda \cup \{\nu\}$ is also downward closed.

Adaptive algorithm : given Λ_n, define $\Lambda_{n+1} := \Lambda_n \cup \{\nu^*\}$ with

$$\nu^* := \mathrm{argmax}\{\|\Delta_\nu u\|_{L^\infty} \ : \ \nu \in \mathcal{N}(\Lambda_n)\}.$$



Theoretical difficulties

The previous adaptive algorithm may fail to converge (in particular $\Delta_\nu u = 0$ for some ν and $\Delta_\mu u \neq 0$ for a $\mu \geq \nu$.

Behaves well in many practical situations.

More conservative variant : Use the above selection rule if n is even, and for odd n choose $\nu^* \in \mathcal{N}(\Lambda_n)$ which was already contained in $N(\Lambda_k)$ for the smallest value of k.

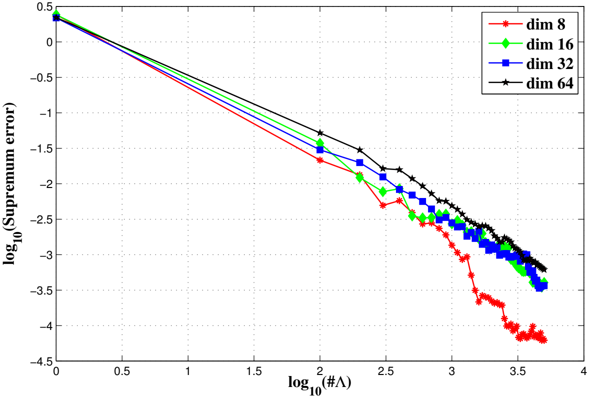
Other variants : measure $\Delta_\nu u$ in L^p norm, use $|\int_U \Delta_\nu u|$ (integration), or $\nu^* \in \mathcal{N}(\Lambda_n)$ minimizing $u(z_\nu)$ (optimization)...

Robustness to dimension growth

We apply the adaptive interpolation algorithm to

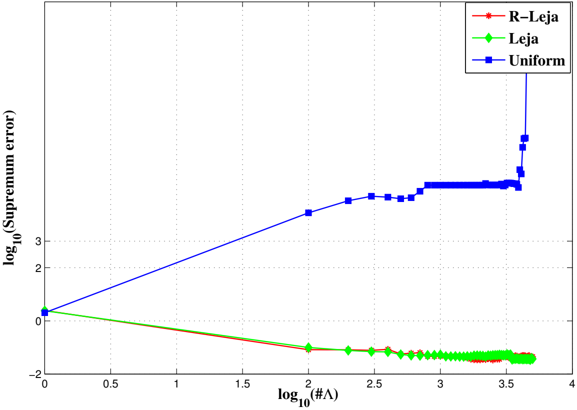
$$u(y) := \Big(1 + \sum_{j=1}^{d} \gamma_j y_j\Big)^{-1}, \quad \gamma_j = \frac{3}{5j^3},$$

for different numbers d of variables.



Robustness to noise

Same function u in dimension $d = 16$, with noisy samples (noise level $= 10^{-2}$). using adaptive interpolation based on different univariate sequences.



Stability

We want to study the Lebesgue constant

$$\mathbb{L}_\Lambda := \|I_\Lambda\|_{L^\infty\to L^\infty} = \sup_u \frac{\|I_\Lambda u\|_{L^\infty}}{\|u\|_{L^\infty}}$$

Useful for approximation since

$$\|u - I_\Lambda u\|_{L^\infty} \le \|u - v\|_{L^\infty} + \|I_\Lambda v - I_\Lambda u\|_{L^\infty}, \quad v \in \mathbb{P}_\Lambda,$$

and thus

$$\|u - I_\Lambda u\|_{L^\infty} \le (1 + \mathbb{L}_\Lambda) \min_{v\in\mathbb{P}_\Lambda} \|u - v\|_{L^\infty}$$

The following result relates $\mathbb{L}_\Lambda$ to the univariate Lebesgue constant

$$\mathbb{L}_k := \|I_k\|_{L^\infty\to L^\infty} = \sup_u \frac{\|I_k u\|_{L^\infty}}{\|u\|_{L^\infty}}$$

Theorem (Chkifa-Cohen-Schwab, 2013) : if $\mathbb{L}_k \le (1+k)^a$, then $\mathbb{L}_\Lambda \le \#(\Lambda)^{1+a}$.

Stability of univariate sequences

For Leja point, it is known (Taylor-Totik, 2008) that $\mathbb{L}_k$ is sub-exponential

$$\lim_{k \to +\infty} \mathbb{L}_k^{1/k} = 1.$$

Numerical computation seems to indicate that

$$\mathbb{L}_k \leq 1 + k.$$

Clemshaw-Curtis points $C_k = \{\cos(l\pi/k) \ : \ l = 0, \dots, k\}$ are dyadically nested :

$$C_{2^j+1} \subset C_{2^{j+1}+1}.$$

For the values $k = 2^{j+1}$ we know that $\mathbb{L}_k \sim \log(k)$.

Problem : how to fill in the intermediate values ?

Sequencial enumeration : disastrous behaviour of $\mathbb{L}_k$.

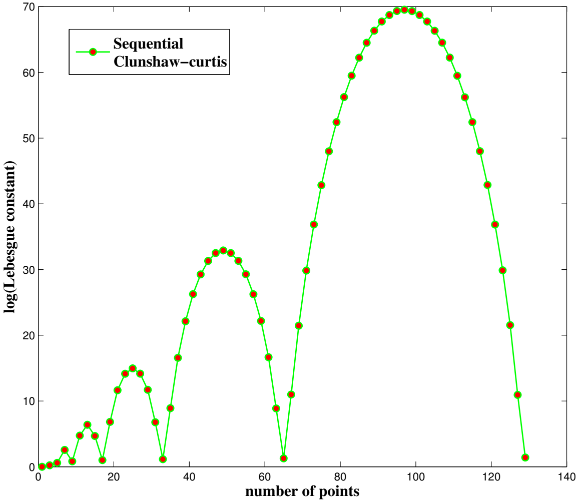
Van der Corput enumeration : it can be proved (Chkifa, 2013) that

$$\mathbb{L}_k \leq (1 + k)^2.$$

This is also the projection of the Leja point for the complex unit disc (R-Leja points).

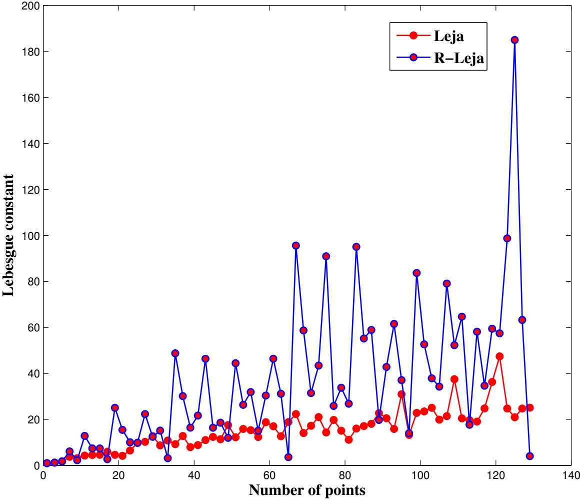
Stability

Lebesgue constant for the Clemshaw-Curtis point with sequencial intermediate filling.



Stability

The Lebesgue constant for the Leja points (red) and the R-Leja points (blue).



Error estimates

So with R-Leja points $\mathbb{L}_k \leq (1+k)^2$ and therefore $\mathbb{L}_\Lambda \leq \#(\Lambda)^3$.

For relevant solution map of parametric PDEs, we know that there exists lower sets $(\Lambda_n)_{n\geq 1}$, such that

$$\min_{v\in V_{\Lambda_n}} \|u-v\|_{L^\infty(U,V)} \leq Cn^{-s}.$$

Therefore, one first interpolation estimate is of the type

$$\|u - I_{\Lambda_n} u\|_{L^\infty(U,V)} \leq Cn^{-s+3},$$

which is uneffective if $s < 3$.

Better : use series, e.g. Taylor $\sum_{\nu\in\mathcal{F}} t_\nu y^\nu$ and write

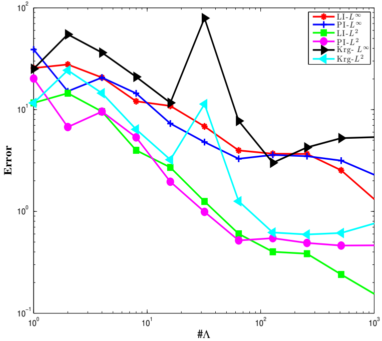
$$\begin{aligned}\|u - I_\Lambda u\|_{L^\infty(U,V)} \quad &= \|\sum_{\nu\in\mathcal{F}} t_\nu y^\nu - I_\Lambda(\sum_{\nu\in\mathcal{F}} t_\nu y^\nu)\|_{L^\infty(U,V)} \\ &= \|\sum_{\nu\notin\Lambda} I_\Lambda(t_\nu y^\nu)\|_{L^\infty(U,V)} \\ &\leq \sum_{\nu\notin\Lambda} \|t_\nu\|_V \|I_\Lambda(y^\nu)\|_{L^\infty(U)} = \sum_{\nu\notin\Lambda} \|t_\nu\|_V \|I_{R_\nu}(y^\nu)\|_{L^\infty(U)} \\ &\leq \sum_{\nu\notin\Lambda} a_\nu \|t_\nu\|_V,\end{aligned}$$

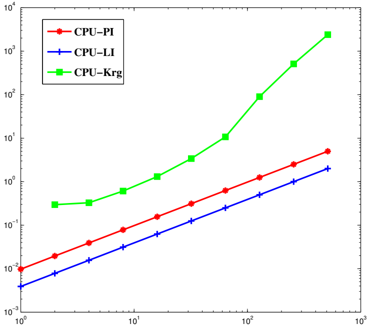
with $a_\nu := \mathbb{L}_{R_\nu} \leq \prod_{j\geq 1}(1+\nu_j)^2$.

If we have estimates of the form $\|t_\nu\|_V \leq C\rho^\nu$ for sequences $(\rho_j)_{j\geq 1}$ satisfying the admissibility constraint, the presence of the algebraic factor can be absorbed in the analysis showing that $(a_\nu \|t_\nu\|_V)_{\nu\in\mathcal{F}} \in \ell^p(\mathcal{F})$ and so we finally obtain the same convergence rate $\|u - I_{\Lambda_n} u\|_{L^\infty(U,V)} \leq Cn^{-s}$ with $s = \frac{1}{p} - 1$ as for the Taylor series.

Comparison with kriging interpolation algorithms

Test case : $y = (y_1, y_2, y_3, y_4, y_5)$ shape parameters in the design of an airfoil and $u(y)$ is the lift to drag ratio (scalar quantity of interest) obtained by ONERA numerical solver.





Error curves in terms of number of points are comparable.

The CPU cost for sparse interpolation scales linearly with the number of points.

This contrasts with kriging methods which require solving ill-conditionned linear systems of growing size + optimization of the parameters of a Gaussian kernel.

Least-squares methods

General context :

- Reconstruction of unknown function $u : U \to \mathbb{R}$ from scattered measurements.
- Measurements $u^i = u(y^i)$ for $i = 1, \dots, m$ with $y^i \in U \subset \mathbb{R}^d$.
- Measurement are costly : one cannot afford to have $m >> 1$.
- Measurements could be noisy : $u^i = u(y^i) + \eta_i$.
- The y^i can be chosen by us or imposed, deterministic or random.
- Questions : how should we sample ? how should we reconstruct ?

Prior : approximability. Analysis of the PDE models shows that in both there exists sequences of m dimensional linear spaces $(V_n)_{n>0}$ such that the unknown function u is well approximated by such spaces

$$e_n(u) := \inf_{v \in V_n} \|u - v\| \leq \varepsilon(n),$$

where $\varepsilon(n)$ is a known bound (such as Cn^{-s}) and where

$$\|v\| := \|v\|_{L^2(U,\rho)},$$

with ρ some probability measure on U.

Least-squares approximation

For a certain value of $n \leq m$ solve

$$\pi = \mathrm{Argmin}_{v \in V_n} \frac{1}{m} \sum_{i=1}^{n} |u^i - v(y^i)|^2.$$

Widely used since its introduction by Gauss.

This is solved by taking a basis $L_1, \dots, L_n$ of V_n and searching π in the form $\pi = \sum_{j=1}^{n} c_j L_j$. The vector $\mathbf{c} = (c_j)^t$ is solution to the normal equations

$$\mathbf{G}\mathbf{c} = \mathbf{a},$$

with $\mathbf{a} = (a_k)^t = (\frac{1}{m} \sum_{i=1}^{m} u^i L_k(y^i))$ and $\mathbf{G} = (G_{k,j}) = (\frac{1}{m} \sum_{i=1}^{m} L_k(y^i) L_j(y^i))$.

The solution always exists and is unique if $\mathbf{G}$ is invertible.

When $u^i = u(x_j)$ this can be viewed as the orthogonal projection of u onto V_n in the sense of the Hilbertian norm

$$\|v\|_m := \left(\frac{1}{m} \sum_{i=1}^{m} |v(y^i)|^2\right)^{1/2}.$$

General questions

1. How accurate is the least square approximation ?
2. Stability with respect to data perturbations ?
3. How large should we take m compared to n ?

The parameter n should typically be thought as describing the level of regularization, and its choice leads to a trade-off :

If n is small : high amount of regularization, which stabilizes the problem (for example, if V_1 is the space of constant function, $n = 1$ gives the average of the data). But the spaces V_n have poor approximation properties.

If n is large : the spaces V_n have better approximation properties, but the least-square approximation may become unstable and therefore unaccurate (the maximal value $m = n$ corresponds to the interpolation problem which might be unstable, e.g. if we use polynomials with uniform distribution of the y^i).

How can we describe the optimal compromise ?

How does this depend on the distribution of the samples y^i ?

A stochastic setting

The y^i are assumed to be i.i.d. according to the probability measure ρ. over U.

We recall the $L^2(U,\rho)$ norm

$$\|v\| := \Big(\int_U |v|^2 d\rho\Big)^{1/2}.$$

The norm $\|v\|_m := \Big(\frac{1}{m}\sum_{i=1}^m |v(y^i)|^2\Big)^{1/2}$ may be thought as its empirical counterpart. It is a stochastic quantity that depends on the draw, and one has $\mathbb{E}(\|v\|_m^2) = \|v\|^2$.

We want to compare the least-square approximation error $\|u-\pi\|$ with the best approximation error

$$e_n(u) := \inf_{v\in V_n} \|u-v\|,$$

We sometimes assume a known uniform bound $\|u\|_{L^\infty} \leq M$ and consider the truncated least-squares estimator

$$\tilde u := T_M\pi, \qquad T_M(t) := \mathrm{sign}(t)\min\{M,|t|\}.$$

An important quantity

Let $L_1, \dots, L_n$ be an orthonormal basis of V_n for the $L^2(U, \rho)$ norm. We introduce

$$k_n(y) = \sum_{j=1}^{n} |L_j(y)|^2,$$

which is the diagonal of the orthogonal projection kernel onto V_n, and also the inverse of the Christoffel function $\phi_n(y) = \inf\{\|v\|^2 \;:\; v \in V_n, v(y) = 1\}$. We define

$$K_n := \|k_n\|_{L^\infty} = \sup_{y \in U} \sum_{j=1}^{n} |L_j(y)|^2.$$

Both are independent on the choice orthonormal basis : they only depend on (V_n, ρ).

One obvious bound is $K_n \leq \sum_{j=1}^{n} \|L_j\|_{L^\infty}^2$. On the other hand, we have

$$K_n \geq \int_U (\sum_{j=1}^{n} |L_j(y)|^2) d\rho = \sum_{j=1}^{n} \|L_j\|^2 = n.$$

Equality holds for spaces with flat orthonormal bases such as trigonometric polynomials on the Torus with uniform measure.

Remark : one can use other bases than (L_j) to solve the least-squares problem.

Main results

Cohen-Davenport-Leviatan (JFoCM, 2013) : let $r > 0$ be arbitrary and let $\kappa = \kappa(r) := \frac{1-\log 2}{2+2r}$. Then, if $K_n \leq \kappa \frac{m}{\log m}$, then the least-squares approximation is

(i) stable : with probability larger than $1 - m^{-r}$, for any data $(u^i)_{i=1,\dots,n}$

$$\|\pi\|^2 \leq 8\Big(\frac{1}{m}\sum_{i=1}^{m} |u^i|^2\Big).$$

(ii) accurate : if $\|u\|_{L^\infty} \leq M$ and $\tilde u := T_M \pi$ with $T_M(t) := \mathrm{sign}(t)\min\{M, |t|\}$,

$$\mathbb{E}(\|u - \tilde u\|^2) \leq (1 + \varepsilon(m))e_n(u)^2 + 8M^2 m^{-r}.$$

where $\varepsilon(m) := \frac{4\kappa}{\log m} \to 0$ as $m \to +\infty$.

Chkifa-Cohen-Migliorati-Nobile-Tempone (M2AN, 2014) : we also have with probability larger than $1 - 2m^{-r}$,

$$\|u - \pi\|^2 \leq (1 + \sqrt{2})e_n(u)_\infty^2, \quad \text{with} \quad e_n(u)_\infty := \inf_{v \in V_n} \|u - v\|_{L^\infty}.$$

Interpretation

The condition $K_n \leq \kappa \frac{m}{\log m}$ describes the amount of regularization that is needed for stabilizing the method.

It suggests to choose the largest value of n such that this condition holds.

Earlier results (Birgé-Massart, Baraud) : stability condition $K_n \leq \kappa \sqrt{\frac{m}{\log m}}$.

Numerical experience suggest that our condition is optimal.

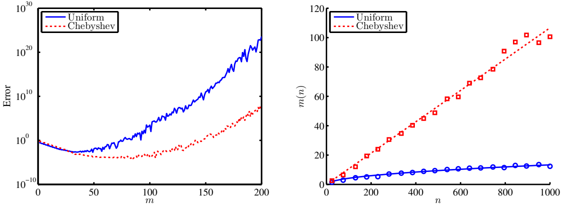
Links with results in Compressed Sensing (Rauhut) : a $N \times m$ matrix $\Phi = (L_j(y^i))$ satisfies the RIP property of order k with high probability for $m \geq cMk(\log m)^3 \log N$ where $M = \max_{j=1,\dots,N} \|L_j\|_{L^\infty}$.

A simple example

$U = [-1, 1]$ and $V_n = \mathbb{P}_{n-1}$.

(i) Uniform distribution $\rho = \frac{dt}{2}$: the L_j are normalized Legendre polynomials and $K_n = \sum_{j=1}^n |L_j(1)|^2 = \sum_{j=1}^n (2j-1) = n^2$. Best choice of n of the order of $m^{1/2}$

(ii) Non-uniform distribution $\rho = \frac{dt}{\pi\sqrt{1-t^2}}$: the L_j are Chebychev polynomials and $K_n = 2n + 1$. Best choice of n of the order of m.



Example : $u(x) = (1 + 25x^2)^{-1}$. Left : approximation error as a function of n for $m = 200$. Right : best value n as a function of m.

In this case, similar stability results can be obtained with deterministic sampling.

Remark : $U = \mathbb{R}$ with gaussian measures is not covered (Hermite polynomials).

Other examples

Local bases : Let V_n be the space of piecewise constant functions over a partition $\mathcal{P}_n$ of U into n cells. An orthonormal basis is given by the functions $\rho(T)^{-1/2}\chi_T$.

If the partition is uniform with respect to ρ, i.e. $\rho(T) = \frac{1}{n}$ for all $T \in \mathcal{P}_n$, then $K_n = n$.

Trigonometric system : with ρ the uniform measure on a torus, since L_j is the complex exponential, one has $K_n = n$.

Spectral spaces on Riemannian manifolds : let $\mathcal{M}$ be a compact Riemannian manifold without boundary and let V_n be spanned by the m first eigenfunctions L_j of the Laplace-Beltrami operator. Then under mild assumptions (doubling properties and Poincaré inequalities), $K_n = \mathcal{O}(n)$ (estimation based on analysis of the Heat kernel in Dirichlet spaces by Kerkyacharian and Petrushev).

Such spaces are therefore well suited for stable least-squares methods. Example : spherical harmonics. Note that individually the eigenfunctions do not satisfy $\|L_j\|_{L^\infty} = \mathcal{O}(1)$.

Key idea in the proof of main result

The stability is controlled by the deviation of the least-square matrix

$$\mathbf{G} := \Big(\frac{1}{m}\sum_{i=1}^{m} L_k(y^i)L_j(y^i)\Big)_{j,k=1,\dots,n},$$

from the identity matrix $\mathbf{I}$ in spectral norm. We want to control, say with $\delta = \frac{1}{2}$,

$$P(\delta) = \Pr\{\|\mathbf{G}-\mathbf{I}\|_2 \geq \delta\},$$

where $\|\cdot\|_2$ is the spectral norm. We show that if $K_n \leq \kappa(r)\frac{m}{\log m}$ then $P(\frac{1}{2}) \leq 2m^{-r}$.

We have $\mathbf{G} = \frac{1}{m}\sum_{i=1}^{m}\mathbf{Y}_i$, with $\mathbf{Y}_i$ i.i.d. copies of the rank one random $n \times n$ matrix $\mathbf{Y} = (L_k(y)L_j(y))_{j,k=1,\dots,n}$, with expectation $\mathbb{E}(\mathbf{Y}) = \mathbf{I}$.

Matrix Chernoff bound (Ahlswede-Winter 2000, Tropp 2011) : if $\|\mathbf{Y}\|_2 \leq K$ a.s.., then

$$\Pr\left\{\left\|\frac{1}{m}\sum_{i=1}^{m}\mathbf{Y}_i - \mathbb{E}(\mathbf{Y})\right\|_2 \geq \delta\right\} \leq 2n\exp\Big(-\frac{mc(\delta)}{K}\Big), \quad c(\delta) := \delta + (1-\delta)\log(1-\delta) > 0.$$

Here $K = \sup_{y\in U}\sum_{j=1}^{n}|L_j(y)|^2 = K_n$, thus $P(\frac{1}{2}) \leq 2m^{-r}$ if $K_n \leq \kappa(r)\frac{m}{\log m}$.

The noisy case

Non parametric bounded regression model : (y^i, u^i), i.i.d. copies of a variable (y, z).

Regression function $u(y) = \mathbb{E}(z|y)$.

We write $u^i = u(y^i) + \eta_i$, with η_i copies of $\eta = y - u(x)$ such that $\mathbb{E}(\eta) = 0$.

We denote by $\sigma^2 := \sup_{y \in U} \mathbb{E}(|\eta|^2|y)$ the noise variance.

Cohen-Davenport-Leviatan (JoFoCM, 2013) : Assume $K_n \leq \kappa \frac{m}{\log m}$ If $\|u\|_{L^\infty} \leq M$ and $\tilde{u} := T_M \pi$, then

$$\mathbb{E}(\|u - \tilde{u}\|^2) \leq (1 + \varepsilon(m)) e_n(u)^2 + 8M^2 m^{-r} + 8\sigma^2 \frac{n}{m}.$$

A typical result on bounded regression with no assumption on K_m (Van de Geer) : assuming $|z| \leq M$ a.s.,

$$\mathbb{E}(\|u - \tilde{u}\|^2) \leq C\Big(e_n(u)^2 + \max\{M^2, \sigma^2\}\frac{n}{m}\Big).$$

Application to parametric PDE's

With $\Lambda \subset \mathcal{F}$, approximation by polynomial space

$$V_\Lambda := \left\{ \sum_{\nu\in\Lambda} v_\nu y^\nu, \ \ v_\nu \in V \right\} = V \otimes \mathbb{P}_\Lambda,$$

with Λ a downward closed index set.

Under various conditions, we know that there exists downward closed sets $\Lambda_1 \subset \Lambda_2 \subset \cdots \subset \Lambda_n \ldots$, with $n := \#(\Lambda_n)$ such that

$$\inf_{v\in V_{\Lambda_n}} \|u - v\|_{L^2(U,V,\rho)} \leq Cn^{-s},$$

with s robust with respect to the parametric dimension d.

We use $V_n = V_{\Lambda_n}$ and solve the least square problem

$$\tilde{u} := \mathrm{Argmin}_{v\in V_n} \frac{1}{m} \sum_{i=1}^{m} \|u^i - v(y^i)\|_V^2,$$

with $u^i = u(y^i)$ computed by a numerical solver for each sample $y^i \in U$.

Chkifa-Cohen-Nobile-Tempone (M2AN, 2014) : with $\rho = \otimes^d(\frac{dx}{2})$ the uniform distribution over U, one has $K_n \leq n^2$ for all downward closed sets Λ_n such that $\#(\Lambda_n) = n$. Improvement to $K_n \sim n^\alpha$ with $\alpha := \frac{\log 2}{\log 3}$ when using the tensor-product Chebychev probability distribution.

Towards optimal sampling

For many relevant instances of approximation families $(V_n)_{n>0}$ and measure ρ, the quantity K_n behaves super-linearly in n leading to too much demanding conditions on the number m of samples to guarantee stability and accuracy.

This can be overcome by weighted least-squares (Doostan 2014, Narayan 2015).

Introduce an auxiliary measure σ and weight function $w \geq 0$ such that $d\rho = wd\sigma$.

Take $\{y^1, \dots, y^m\}$ i.i.d. with respect to σ and solve the weighted least-squares problem

$$\pi = \mathrm{Argmin}_{v\in V_n} \frac{1}{m}\sum_{i=1}^m w(y^i)|u^i - v(y^i)|^2,$$

The case $w = 1$ and $\sigma = \rho$ gives the non-weighted least-squares, and the modified discrete norm

$$\|v\|_m^2 := \frac{1}{m}\sum_{i=1}^m w(y^i)|v(y^i)|^2,$$

is again an unbiased approximation of $\|v\|^2 = \int_U |v(y)|^2 d\rho = \int_U w(y)|v(y)|^2 d\sigma$.

We want to pick the pair (σ, w) in order to optimize the approximation. The same idea is used when performing importance sampling in Monte-Carlo method.

Convergence analysis

Based on $k_{n,w}(y) := w(y)\sum_{j=1}^n |L_j(y)|^2 = w(y)k_n(y)$ and $K_{n,w} := \|k_{n,w}\|_{L^\infty} \geq n$.

Cohen-Migliorati (2016) : if $K_{n,w} \leq \kappa \frac{m}{\log m}$, then the weighted least-squares approximation satisifie with probability larger than $1 - {}^{-r}$, for any data $(u^i)_{i=1,\dots,m}$

$$\|\pi\|^2 \leq 8\Big(\frac{1}{m}\sum_{i=1}^m |u^i|^2\Big).$$

and if $\|u\|_{L^\infty} \leq M$ and $\tilde{u} := T_M\pi$ with $T_M(t) := \mathrm{sign}(t)\min\{M, |t|\}$,

$$\mathbb{E}(\|u - \tilde{u}\|^2) \leq (1 + \varepsilon(m))e_n(u)^2 + 8M^2m^{-r}.$$

where $\varepsilon(m) := \frac{4\kappa}{\log m} \to 0$ as $m \to +\infty$. Also, with probability larger than $1 - 2m^{-r}$,

$$\|u - \pi\|^2 \leq (1 + \sqrt{2})e_n(u)_\infty^2, \quad \text{with} \quad e_n(u)_\infty := \inf_{v\in V_n} \|u - v\|_{L^\infty}.$$

An optimal sampling method : take $w(y) = w_n(y) = \frac{n}{k_n(y)}$ so that $d\sigma = d\sigma_n = \frac{k_n}{n}d\rho$ is a probability measure and $k_{n,w} = n$. Then, stability and optimal accuracy is achieved by the weighted least-squares method under the minimal condition $\frac{m}{\ln m} \sim n$.

This approach allows to treat polynomial approximation on unbounded domains, for example $U = \mathbb{R}^d$ or $\mathbb{R}^{\mathbb{N}}$ with Gaussian measure.

Conclusions

Adaptive algorithms with optimal theoretical guarantees are still to be developed, in particular for non-intrusive approaches (interpolation, collocation, least-squares).

Main advantages of sparse interpolation methods over kriging : scalability and stability. Advantages of kriging : points can be completely arbitrary.

General principle : hierarchical 1d system $->$ tensorization $->$ sparsification Open question : $1d$ sequences with logarithmic growth Lebesgue constants

Weighted least-squares : the optimal pair (w_n, σ_n) depends on n. This is a problem for adaptive methods in which we may want to vary the value of n.

In certain simple cases, there exists explicit asymptotics for the Christoffel function $k_n(y)^{-1}$ of the form $\frac{n}{k_n} \sim w$. This yields a near optimal pairs (w, σ) that do not change with n. Example : polynomials $V_n = \mathbb{P}_{n-1}$ on $U = [-1, 1]$ and ρ a Jacobi weight, take σ to be the equilibrium measure $\frac{dt}{2\pi\sqrt{1-t^2}}$.

Producing an i.i.d. sample with respect to the optimal pair (w_n, σ_n) is not always easily feasible, in particular in high dimension.

Deterministic counterpart to these results ? Error estimate in L^∞ ? Is there an optimal set of points ? Related work : magic points for interpolation (Maday-Patera).

Part 3

Reduced modeling/bases, data assimilation, parameter estimation

Reduced order modeling and n-width

Recall the benchmark of Kolmogorov n-width of the solution manifold

$$d_n(\mathcal{M})_V = \inf_{\dim(E)=n} \max_{v\in\mathcal{M}} \min_{w\in E} \|v-w\|_V = \inf_{\dim(E)=n} \max_{y\in U} \min_{w\in E} \|u(y)-w\|_V.$$

Uniform approximation estimates of the solution map $y \mapsto u(y)$ by polynomial (or other separable) expansions give an upper bound on n-width

$$d_n(\mathcal{M})_V \leq \min_{v\in V_{\Lambda_n}} \|u-v\|_{L^\infty(U,V)} \leq Cn^{-s}.$$

We do not know other approaches to estimate the n-width of the solution manifold by above.

These estimates might very pessimistic in the sense that he actual n-width $d_n(\mathcal{M})_V$ is much smaller than the right side.

We do not have results proving lower bounds for the n-widths of solution manifolds.

It is desirable to have numerical reduced modeling methods that can provably perform as good as the n-width benchmark.

Reduced bases (Maday, Patera)

Define a reduced modeling space $V_n = \mathrm{span}\{u_0, \dots, u_{n-1}\}$, where the u_i are particular instances (snapshots) from the solution manifold

$$u_i = u(y^i)$$

for some $y^0, \dots, y^{n-1} \in U$.

Greedy selection : having selected $u_0, \dots, u_{k-1} \in \mathcal{M}$, choose the next instance by

$$u_k = \mathrm{argmax}\{\|u - P_{V_k} u\|_V \ : \ v \in \mathcal{M}\},$$

where P_E is the V-orthogonal projector onto E, or equivalently $u_k = u(y^k)$, with

$$y^k = \mathrm{argmax}\{\|u(y) - P_{V_k} u(y)\|_V \ : \ y \in U\}.$$

This algorithm is not realistic : $\|u(y) - P_{V_k} u(y)\|_V$ is unknown, however can be estimate at moderate cost by a-posteriori error analysis. Therefore, one rather apply a weak-greedy algorithm : u_k such that

$$\|u_k - P_{V_k} u_k\|_V \geq \gamma \max\{\|v - P_{V_k} v\|_V \ : \ v \in \mathcal{M}\},$$

for some fixed $0 < \gamma < 1$.

Comparison with n-width

Performance of reduced bases : $\sigma_n(\mathcal{M})_V := \max\{\|v - P_{V_n}v\|_V \ : \ v \in \mathcal{M}\}$

Comparison with n-width : how does $\sigma_n(\mathcal{M})_V$ compares with $d_n(\mathcal{M})_V$?

Theorem (Buffa-Maday-Patera-Prudhomme-Turinici, 2012) : $\sigma_n(\mathcal{M})_V \leq n2^n d_n(\mathcal{M})_V$.

Theorem (Binev-Cohen-Dahmen-DeVore-Petrova-Wojtaszczyk, 2013) : for any $n > 0$ and $\varepsilon > 0$, there exists $\mathcal{M}$ such that $\sigma_n(\mathcal{M})_V \geq (1-\varepsilon)2^n d_n(\mathcal{M})_V$.

A more favorable comparison is possible in terms of convergence rates :

Theorem (Binev-Cohen-Dahmen-DeVore-Petrova-Wojtaszczyk, 2013) : For any $s > 0$ one has

$$\sup_{n\geq 1} n^s d_n(\mathcal{M})_V < \infty \implies \sup_{n\geq 1} n^s \sigma_n(\mathcal{M})_V < \infty,$$

and for any $a > 0$ there exists $b > 0$ such that

$$\sup_{n\geq 1} e^{an^s} d_n(\mathcal{M})_V < \infty \implies \sup_{n\geq 1} e^{bn^s} \sigma_n(\mathcal{M})_V < \infty.$$

A matrix reformulation

In order to prove the theorem, we introduce the functions $\{u_0^*, u_1^*, \cdots\}$ obtained by applying Gram-Schmidt orthonormalization algorithm to the sequence $\{u_0, u_1, \cdots\}$. We consider the lower triangular matrix $A = (a_{i,j})_{i,j\geq 0}$ defined by

$$u_i = \sum_{j=0}^{i} a_{i,j} u_j^* .$$

This matrix satisfies two fundamental properties. Since $a_{n,n} = \langle u_n, u_n^* \rangle = \|u_n - P_{V_n} u_n\|_V$, we have

$$\gamma \sigma_n \leq |a_{n,n}| \leq \sigma_n \quad (P1),$$

where $\sigma_n := \sigma_n(\mathcal{M})_V$. Since for $m \geq n$ we have $\|u_m - P_{V_n} u_m\|_V \leq \sigma_n$, we have

$$\sum_{j=n}^{m} a_{m,j}^2 \leq \sigma_n^2 \qquad (P2)$$

Conversely, for any matrix satisfying these two properties with $(\sigma_n)_{n\geq 0}$ a non-increasing sequence going to 0, there exists a compact set $\mathcal{M}$ in $\ell^2(\mathbb{N})$ (the lines of the matrix) such that a realization the weak-greedy algorithm exactly leads to this matrix.

A key lemma

Note that since $u_i \in \mathcal{M}$ for all i, there exists a m dimensional space W of $\ell^2(\mathbb{N})$ such that each row of A is approximated by W with accuracy $d_m := d_m(\mathcal{M})_V$ in $\ell^2(\mathcal{N})$.

The same holds for any submatrix of A by restriction of W.

Lemma : let $G = (g_{i,j})$ be a $K \times K$ matrix with rows $\mathbf{g}_1, \dots, \mathbf{g}_K$. If W is any m dimensional subspace of $\mathbb{R}^K$ for some $0 < m \leq K$, and P is the orthogonal projection from $\mathbb{R}^K$ onto W, then

$$\det(G)^2 \leq \left(\frac{1}{m}\sum_{i=1}^{K} \|P\mathbf{g}_i\|_{\ell^2}^2\right)^m \left(\frac{1}{K-m}\sum_{i=1}^{K} \|\mathbf{g}_i - P\mathbf{g}_i\|_{\ell^2}^2\right)^{K-m}.$$

We apply this lemma to $K \times K$ matrix $G = (g_{i,j})$ which is formed by the rows and columns of A with indices $N+1, \dots, N+K$. By Property $(P2)$, we obtain

$$\|P\mathbf{g}_i\|_{\ell^2} \leq \|\mathbf{g}_i\|_{\ell^2} \leq \sigma_{N+1}, \quad i = 1, \dots, K,$$

We also have,

$$\|\mathbf{g}_i - P\mathbf{g}_i\|_{\ell^2} \leq d_m, \quad i = 1, \dots, K.$$

It follows that

$$\gamma^{2K}\prod_{i=1}^{K}\sigma_{N+i}^2 \leq \prod_{i=1}^{K} a_{N+i,N+i}^2 = \det(G)^2 \leq \left(\frac{K}{m}\right)^m \left(\frac{K}{K-m}\right)^{K-m} \sigma_{N+1}^{2m} d_m^{2K-2m}.$$

Application : exponential rates

We take $N = 0$, $K = n$ and any $1 \leq m < n$. Using the monotonicity of $(\sigma_n)_{n\geq 0}$ and $\sigma_1 \leq \sigma_0 \leq d_0$, we obtain

$$\sigma_n^{2n} \leq \prod_{j=1}^{n} \sigma_j^2 \leq \gamma^{-2n} \left(\frac{n}{m}\right)^m \left(\frac{n}{n-m}\right)^{n-m} d_m^{2n-2m} d_0^{2m}.$$

Since $x^{-x}(1-x)^{x-1} \leq 2$ for $0 < x < 1$, it follows that

$$\sigma_n \leq \sqrt{2}\gamma^{-1} d_0^{\frac{m}{n}} \min_{1\leq m<n} d_m^{\frac{n-m}{n}}, \quad n \geq 1,$$

and particular

$$\sigma_{2n} \leq \gamma^{-1}\sqrt{2d_0 d_n}.$$

From this, one easily derive

$$\sup_{n\geq 1} e^{an^s} d_n(\mathcal{M})_V < \infty \implies \sup_{n\geq 1} e^{bn^s} \sigma_n(\mathcal{M})_V < \infty.$$

Application : algebraic rates

We take $N = K = n$ and any $1 \leq m < n$. Using the monotonicity of $(\sigma_n)_{n\geq 0}$, we obtain

$$\sigma_{2n}^{2n} \leq \prod_{j=n+1}^{2n} \sigma_j^2 \leq \gamma^{-2n}\left(\frac{n}{m}\right)^m \left(\frac{n}{n-m}\right)^{n-m} \sigma_n^{2m} d_m^{2n-2m}.$$

In the case $n = 2k$ and $m = k$ we have for any positive integer k,

$$\sigma_{4k} \leq \sqrt{2}\gamma^{-1}\sqrt{\sigma_{2k} d_k}.$$

Assuming that $d_n \leq C_0 n^{-s}$ for all $n \geq 1$ and $d_0 \leq C_0$, we obtain by induction that for all $j \geq 0$ and $n = 2^j$,

$$\sigma_n = \sigma_{2^j} \leq C2^{-sj} \leq n^{-s}, \quad C := \ 2^{3s+1}\gamma^{-2}C_0.$$

Indeed, the above obviously holds for $j = 0$ or 1 since for these values, we have $\sigma_{2^j} \leq \sigma_0 = d_0 \leq C_0 \leq C2^{-sj}$. Assuming its validity for some $j \geq 1$, we find that

$$\begin{aligned} \sigma_{2^{j+1}} &\leq \sqrt{2}\gamma^{-1}\sqrt{\sigma_{2^j} d_{2^{j-1}}} \\ &\leq \gamma^{-1} 2^{\frac{3s}{2}} \sqrt{2CC_0} 2^{-s(j+1)} \\ &= \sqrt{C}\sqrt{2^{3s+1} C_0 \gamma^{-2}} 2^{-s(j+1)} = C2^{-s(j+1)}, \end{aligned}$$

where we have used the definition of C. For values $2^j < n < 2^{j+1}$, we obtain the general result by writing

$$\sigma_n \leq \sigma_{2^j} \leq C2^{-sj} \leq 2^s C n^{-s} = C_1 n^{-s}.$$

Proof of the key lemma

Let $G = (g_{i,j})$ be a $K \times K$ matrix with rows $\mathbf{g}_1, \dots, \mathbf{g}_K$, and let W be any m dimensional subspace of $\mathbb{R}^K$ for some $0 < m \leq K$ with projector P. Take $\varphi_1, \dots, \varphi_m$ any orthonormal basis for the space W and complete it into an orthonormal basis $\varphi_1, \dots, \varphi_K$ for $\mathbb{R}^K$.

We denote by Φ the $K \times K$ orthogonal matrix whose j-th column is φ_j, then the matrix $C := G\Phi$ has entries $c_{i,j} = \langle \mathbf{g}_i, \varphi_j \rangle$. We have

$$\det(G)^2 = \det(C)^2.$$

With $\mathbf{c}_j$ the j-th column of C, the arithmetic-geometric mean inequality yields

$$\prod_{j=1}^{m} \|\mathbf{c}_j\|_{\ell^2}^2 \leq \Big(\frac{1}{m} \sum_{j=1}^{m} \|\mathbf{c}_j\|_{\ell^2}^2 \Big)^m = \Big(\frac{1}{m} \sum_{j=1}^{m} \sum_{i=1}^{K} \langle \mathbf{g}_i, \varphi_j \rangle^2 \Big)^m = \Big(\frac{1}{m} \sum_{i=1}^{K} \|P\mathbf{g}_i\|_{\ell^2}^2 \Big)^m.$$

Likewise, since φ_j is orthogonal to W when $j > m$,

$$\prod_{j=m+1}^{K} \|\mathbf{c}_j\|_{\ell^2}^2 \leq \Big(\frac{1}{K-m} \sum_{j=m+1}^{K} \|\mathbf{c}_j\|_{\ell^2}^2 \Big)^{K-m} = \Big(\frac{1}{K-m} \sum_{i=1}^{K} \|\mathbf{g}_i - P\mathbf{g}_i\|_{\ell^2}^2 \Big)^{K-m}.$$

We conclude by using Hadamard's inequality, which gives

$$\det(C)^2 \leq \prod_{j=1}^{K} \|\mathbf{c}_j\|_{\ell^2}^2 \leq \Big(\frac{1}{m} \sum_{i=1}^{K} \|P\mathbf{g}_i\|_{\ell^2}^2 \Big)^m \Big(\frac{1}{K-m} \sum_{i=1}^{K} \|\mathbf{g}_i - P\mathbf{g}_i\|_{\ell^2}^2 \Big)^{K-m}.$$

Parametric PDEs and data assimilation

Parametric PDEs of the general form

$$\mathcal{P}(u, y) = 0,$$

are used to describe many physical processes.

In some settings, we know the governing PDEs $\mathcal{P}$ but do not know the parameters y of the solution we are trying to capture.

Example : groundwater modeling where the process is governed by

$$-\mathrm{div}(a\nabla u) = f,$$

with suitable boundary conditions and

$$a = a(y) = \overline{a} + \sum_{j=1}^{d} y_j \psi_j, \quad y = (y_j)_{j=1,\dots,d} \in \mathcal{P} = [-1, 1]^d.$$

The parameter vector $y \in \mathcal{P}$ describing the diffusion properties of the underground could be unknown to us. So we make some local measurements by "drilling".

How can we best combine these measurements with the model to reconstruct $u(y)$?

Another example

Reconstruction of acoustic fields from recorded data (ANR project ECHANGE 2012).

Abstraction of the Problem (Maday-Patera-Penn-Yano)

Let $\mathcal{M} = \{u(\cdot, y) \ : \ y \in \mathcal{P}\}$ be the solution manifold of the parametric PDE.

We assume $\mathcal{M}$ to be compact in the solution (Hilbert) space V.

We wish to approximate an element $u \in \mathcal{M}$ based on the knowledge of m observations

$$\ell_i(u), \quad i = 1, \dots, m,$$

where the ℓ_i are linear forms on V.

The ℓ_i are imposed to us (or chosen by us within a given dictionnary $\mathcal{D} \subset V'$).

Questions :

What is the best algorithm for approximating u from this information ?

What is the best error we can achieve ?

Problems of this type are known as optimal recovery.

The model

We know that u lies on the solution manifold $\mathcal{M}$. However, this manifold is complex and its exact description is numerically out of reach.

Reduced modeling methods, e.g. reduced basis and polynomial chaos, allow us to identify nested finite dimensional spaces V_k which approximate $\mathcal{M}$ up to known tolerances ε_k.

$$V_0 \subset V_1 \subset \cdots \subset V_n, \quad \dim(V_k) = k.$$

such that

$$\mathrm{dist}(u, V_k) := \min_{w \in V_k} \|u - w\|_V \leq \varepsilon_k, \quad k = 1, \dots, n, \quad u \in \mathcal{M},$$

or equivalently

$$\sup_{y \in \mathcal{P}} \mathrm{dist}(u(y), V_k) \leq \varepsilon_k, \quad k = 1, \dots, n.$$

In the reduced basis method, the V_k are generated by snapshots $\{u(y^1), \dots, u^(y^k)\}$ selected e.g. by greedy algorithms (Maday-Patera). Performance ε_k "nearly" as good as that of ideal n-width spaces (Binev, AC, Dahmen, DeVore, Petrova, Wojtaszczyk) :

$$\sup_{k \geq 1} k^s \varepsilon_k \leq C_s \sup_{k \geq 1} k^s d_k, \quad d_k := d_k(\mathcal{M}) := \inf_{\dim(V)=k} \sup_{u \in \mathcal{M}} \mathrm{dist}(u, V).$$

So, a more friendly model is to replace the assumption $u \in \mathcal{M}$ by these weaker but better understood assumptions on approximability

Model meets data

We may write

$$\ell_i(u) = \langle u, \omega_i \rangle, \quad i = 1, \dots, m,$$

and define the measurement space

$$W := \mathrm{span}\{\omega_1, \dots, \omega_m\}.$$

The measurement data determine $w = P_W u \in W$.

Model class : $\mathcal{K}$ determined by

$$V_0 \subset V_1 \subset \cdots \subset V_n \quad \text{and} \quad \varepsilon_0 \geq \varepsilon_1 \geq \cdots \geq \varepsilon_n \geq 0.$$

Two settings : one space

$$\mathcal{K} := \mathcal{K}^{\mathrm{one}} := \{u \in V \ : \ \mathrm{dist}(u, V_n) \leq \varepsilon_n\},$$

or multi-space

$$\mathcal{K} := \mathcal{K}^{\mathrm{mult}} := \{u \in V \ : \ \mathrm{dist}(u, V_k) \leq \varepsilon_k, \quad k = 0, 1, \dots, n\}.$$

Our knowledge about the function is thus that it belongs to

$$\mathcal{K}_w := \{u \in V \ : \ P_W u = w\} = \mathcal{K} \cap \mathcal{H}_w, \quad \mathcal{H}_w := \{u \ : \ P_W u = w\}.$$

Algorithms

Given the data $w = P_W u$ we want to reconstruct an approximation $\tilde{u}(w)$ to u.

Algorithm : computable map $A : W \to V$ giving the approximation $\tilde{u}(w) = A(P_W u)$.

Ambiguity : all elements $u \in \mathcal{K}_w$ are assigned the same approximation $A(w)$.

Error of an algorithm A for an individual u :

$$E_A(u) = \|u - A(P_W u)\|.$$

.

Performance of an algorithm A over the class $\mathcal{K}_w$:

$$E_A(\mathcal{K}_w) := \sup_{u \in \mathcal{K}_w} \|u - A(P_W u)\|.$$

Performance of an algorithm A over the class $\mathcal{K}$:

$$E_A(\mathcal{K}) := \sup_{w \in W} \sup_{u \in \mathcal{K}_w} \|u - A(P_W u)\|.$$

Best algorithm

Best algorithm A^* for the given model class $\mathcal{K}$:

$$A^* := \mathrm{argmin}_A E_A(\mathcal{K})$$

This algorithm is simple to describe : for any bounded set $\mathcal{S} \in V$ we define the Chebyshev ball as the ball $B(v^*, R^*)$ of minimal radius which contains $\mathcal{S}$.

We say that v^* is the Chebyshev center of $\mathcal{S}$ and $R^* = \mathrm{rad}(S)$ its Chebyshev radius.

Then the best algorithm takes for $A(w) = u^*(w)$ the Chebyshev center of $\mathcal{K}_w$ and has performances

$$E_{A^*}(\mathcal{K}_w) = \mathrm{rad}(\mathcal{K}_w),$$

and

$$E_{A^*}(\mathcal{K}) = \sup_{w \in W} \mathrm{rad}(\mathcal{K}_w).$$

The Chebychev center and radius of a general set are generally not easy to find.

The one space case

This case was studied by Maday-Patera-Penn-Yano (MPPY).

Their algorithm can be described as follows :

1. For the given data $w \in W$, determine $\tilde{u}(w) \in \mathcal{H}_w = \{u \in V \ : \ P_W u = w\}$ and $\tilde{v}(w) \in V_n$ such that

$$\|\tilde{u}(w) - \tilde{v}(w)\| = \min\{\|u - v\| \ : \ u \in \mathcal{H}_w, v \in V_n\} = \mathrm{dist}(\mathcal{H}_w, V_n).$$

2. Define $A(w) = \tilde{u}(w)$.

The analysis of this algorithm is based on the inf-sup constant

$$\beta(V_n, W) := \inf_{v \in V_n} \sup_{w \in W} \frac{\langle v, w \rangle}{\|v\| \, \|w\|}.$$

We mainly use the notation $\mu(V_n, W) := \beta(V_n, W)^{-1}$.

Notice that $\mu(V_n, W) = +\infty$ means that there are non-trivial $v \in V_n \cap W^{\perp}$. So $\mathcal{K}_{w=0}$ is an unbounded set and therefore $E_A(\mathcal{K}) = E_A(\mathcal{K}_0) = +\infty$ for any algorithm A.

This happens in particular if $\dim(V_n) > \dim(W)$.

Optimality of MPPY algorithm in the one space case

Our results :

The center of Chebyshev ball of $\mathcal{K}_w$ coincides with $A(w)$ computed by the MPPY algorithm.

Its radius is given by

$$R(\mathcal{K}_w)^2 = \mu(V_n, W)^2 \left(\varepsilon_n^2 - \|\tilde{u}(w) - \tilde{v}(w)\|^2\right).$$

Therefore $A = A^*$ is the best possible algorithm and its performance is

$$E_{A^*}(\mathcal{K}) = \mu(V_n, W)\varepsilon_n,$$

corresponding to the case $w = 0$ for which $\tilde{u}(w) = \tilde{v}(w) = 0$.

Remarks : finding $A^*(w)$ does not require the knowledge of ε_n.

$\mathcal{K}_W$ is an ellipsoid : interstection of the cylinder $\mathcal{K}$ with affine space $\mathcal{H}_w$.

MPPY also how to select the measurements in order to make $\mu(V_n, W)$ small.

Optimized measurements ?

Given $(V_n)_{n\geq 0}$ reduced model space we want to select the measurement functions $(\omega_i)_{i\geq 1}$ out of a dictionnary $\mathcal{D}$ (a set of norm 1 functions, complete in V).

Objective : guarantee a lower bound on $\beta(V_n, W)$ with $W = \mathrm{span}\{\omega_1, \dots, \omega_m\}$ and $m = m(n) \geq n$ not too large.

Evaluation of $\beta(V_n, W)$ requires SVD of an $n \times m$ matrix and its maximization over all possible choices of $\{\omega_1, \dots, \omega_m\}$ is computationally intensive.

Observe that

$$\beta(V_n, W) := \inf_{v\in V_n} \sup_{w\in W} \frac{\langle v, w\rangle}{\|v\|\,\|w\|} = \inf_{v\in V_n, \|v\|=1} \|P_W v\|.$$

Therefore $\beta(V_n, W) \geq \gamma > 0$ if and only if

$$\sup_{v\in V_n, \|v\|=1} \|v - P_W v\| \leq \delta := \sqrt{1-\gamma^2} < 1.$$

This leads to consider OMP-type algorithms selecting dictionnary elements for the collective approximation of the elements of V_n (work in progress with Binev and Mula).

The multi-space case

Recall that

$$\mathcal{K}^{\text{mult}} = \bigcap_{j=0}^{n} \mathcal{K}^j, \quad \mathcal{K}^j = \{u \,:\, \operatorname{dist}(u, V_j) \le \varepsilon_j\}.$$

and therefore

$$\mathcal{K}_w = \mathcal{K}_w^{\text{mult}} = \bigcap_{j=0}^{n} \mathcal{K}_w^j, \quad \mathcal{K}_w^j = \mathcal{K}^j \cap \mathcal{H}_w.$$

Thus, $\mathcal{K}_w$ is an intersection of ellipsoids.

Finding Chebyshev center and radius of an intersection of ellipsoids is NP hard.

The multi-space case

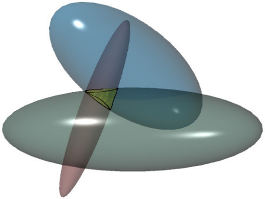
Recall that

$$\mathcal{K}^{\text{mult}} = \bigcap_{j=0}^{n} \mathcal{K}^j, \quad \mathcal{K}^j = \{u \,:\, \operatorname{dist}(u, V_j) \le \varepsilon_j\}.$$

and therefore

$$\mathcal{K}_w = \mathcal{K}_w^{\text{mult}} = \bigcap_{j=0}^{n} \mathcal{K}_w^j, \quad \mathcal{K}_w^j = \mathcal{K}^j \cap \mathcal{H}_w.$$

Thus, $\mathcal{K}_w$ is an intersection of ellipsoids.



Finding Chebyshev center and radius of an intersection of ellipsoids is NP hard.

Our contribution to the multi-space case

A "poor man's algorithm" : choose the best one space, giving thus the performance

$$E_A(\mathcal{K}) = \min_{k=0,\dots,n} \mu(V_k, W)\varepsilon_k.$$

Notice that $\mu(V_k, W)$ increases with k while ε_k decreases.

Simple examples (in 2d or 3d) show that $\mathrm{rad}(\mathcal{K}_w)$ can be arbitrarily smaller than the above poor man's estimate. Better algorithms are thus desirable.

1. Numerical :

We study an algorithm A that finds a point $\tilde{u}(w)$ in the intersection $\mathcal{K}_w = \bigcap_{j=0}^n \mathcal{K}_w^j$.

Therefore, this algorithm is near optimal :

$$E_A(\mathcal{K}_w) \leq 2\mathrm{rad}(\mathcal{K}_w) = 2E_{A^*}(\mathcal{K}_w)$$

2. Theoretical :

We give computable a-priori bounds for the Chebyshev radius in the multi-space case.

These bounds are always smaller than poor man's estimate, sometimes much smaller.

Algorithms for the multi-space case

Finding a point in the intersection of convex sets is a standard problem known as convex feasability (see e.g. book by Patrick Combettes).

Standards methods are available. Here we discuss alternate projections.

Observation : If $\mathcal{K}_w$ is non-empty, it contains at least a point in the space $\tilde{V} := V_n + W$. Therefore we may restrict ourself to this finite dimensional setting.

We use the orthogonal projections $P_{\mathcal{K}^j}$ and $P_{\mathcal{H}_w}$ onto the cylinders $\mathcal{K}_j$ and affine space $\mathcal{H}_w$. Both are trivial to compute using appropriate bases.

With $u^0 = 0$, consider the sequence

$$u^{k+1} := P_{\mathcal{K}_n} P_{\mathcal{K}_{n-1}} \dots P_{\mathcal{K}_0} P_{\mathcal{H}_w} u^k.$$

Remarks :

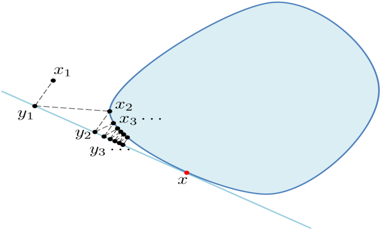
(i) The nestedess $V_0 \subset V_1 \subset \cdots \subset V_n$ implies that $u^k \in \mathcal{K} = \mathcal{K}^{\rm mult} = \bigcap_{j=0}^n \mathcal{K}_w^j$ at each iterations $\mathcal{K}$.

(ii) Projection $P_{\mathcal{S}}$ onto one convex set has the property that $\|P_{\mathcal{S}} v - w\| \leq \|v - w\|$ for all $w \in \mathcal{S}$. This implies that

$$\mathrm{dist}(u^{k+1}, \mathcal{K}_w) \leq \mathrm{dist}(u^k, \mathcal{K}_w).$$

Convergence

Convergence of alternate projection always hold but can be arbitrarily slow.

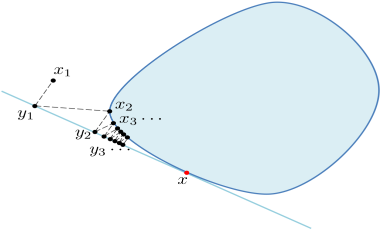


Establishing convergence rate require some geometric properties, such as uniform convexity. We do not have this, however we do have a restricted form :

If $u_1, u_2 \in \mathcal{K}$ with $u_1 - u_2 \in W^\perp$, then the ball centered at $u_0 = \frac{1}{2}(u_1 + u_2)$ with radius $r = c \min\{\delta, \delta^2\}$ where $\delta := \|u_1 - u_2\|$ is contained in $\mathcal{K}$. The constant c depends on the ε_j and $\mu_j := \mu(V_j, W)$ for $j = 0, \dots, n$.

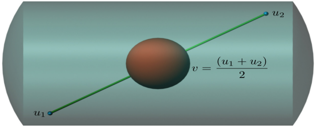
Convergence

Convergence of alternate projection always hold but can be arbitrarily slow.



Establishing convergence rate require some geometric properties, such as uniform convexity. We do not have this, however we do have a restricted form :

If $u_1, u_2 \in \mathcal{K}$ with $u_1 - u_2 \in W^{\perp}$, then the ball centered at $u_0 = \frac{1}{2}(u_1 + u_2)$ with radius $r = c \min\{\delta, \delta^2\}$ where $\delta := \|u_1 - u_2\|$ is contained in $\mathcal{K}$. The constant c depends on the ε_j and $\mu_j := \mu(V_j, W)$ for $j = 0, \ldots, n$.



Applications

We exploit the restricted convexity to prove the following :

Convergence rate : u^k converges to a point in $\mathcal{K}_w$ with rate at worse $\mathcal{O}(k^{-1/2})$.

If $\mathcal{K}_w$ contains a point in the interior of $\mathcal{K}$: exponential rate $\mathcal{O}(\rho^k)$ with $\rho < 1$.

A posteriori estimate : for any $v \in \mathcal{K}$, one has

$$\mathrm{dist}(v, \mathcal{K}_w) \leq C \max\{\mathrm{dist}(v, \mathcal{H}_w), \mathrm{dist}(v, \mathcal{H}_w)^{1/2}\},$$

where C depends on the ε_j and $\mu_j := \mu(V_j, W)$ for $j = 0, \dots, n$.

Can be applied to $v = u^k$ to get a stopping criterion.

Favorable bases

Theoretical analysis of the one-space and multi-space problem is facilitated by the introduction of appropriate bases.

Let $\{\phi_1, \dots, \phi_n\}$ be an orthonormal system such that $V_j = \mathrm{span}\{\phi_1, \dots, \phi_j\}$ for $j \le n$.

This basis is well adapted to describe the model : $u = \sum_{j=1}^n \alpha_j \phi_j + e$ with $e \in V_n^\perp$ belongs to $\mathcal{K}$ if and only if

$$\sum_{k=j+1}^{n} \alpha_k^2 + \|e\|^2 \le \varepsilon_j, \quad j = 0, \dots, n.$$

Let $\{\omega_1, \dots, \omega_m\}$ be an orthonormal basis of W, and introduce the cross-grammian

$$G := \langle \omega_i, \phi_j \rangle.$$

Applying SVD to G leads to new orthonormal bases $\{\omega_1^*, \dots, \omega_m^*\}$ and $\{\phi_1^*, \dots, \phi_n^*\}$ of W and V_n such that :

$$\langle \omega_i^*, \phi_j^* \rangle = s_i \delta_{i,j}, \quad 1 \ge s_1 \ge s_2 \ge \dots \ge s_n \ge 0.$$

$$P_W \phi_j^* = s_j \omega_j^*, \quad j = 1, \dots, n, \quad \text{and} \quad \beta(V_n, W) = s_n.$$

A-priori estimates

Favorable bases allow us to derive a-priori bound on $\mathrm{rad}(\mathcal{K}_w)$ in the multi-space case.

Such bound can be usee in giving a stoping criterion for the convex optimization algorithm.

Let $\Lambda = (\lambda_{i,j})$ be the change of basis matrix from (ϕ_j) to (ϕ_j^*), and compute

$$\theta_i = \sum_{j=1}^{n} |\lambda_{i,j}| \varepsilon_{j-1}.$$

Define k the largest integer such that $\sum_{j=k}^n s_j^2 \theta_j^2 \geq \varepsilon_n^2$.

A priori bound : $\mathrm{rad}(\mathcal{K}_w) \leq 2(\varepsilon_n^2 + \sum_{j=k}^n \theta_j^2)^{1/2}$.

This bound can be much smaller than the poor man's estimate $\min_{j=1,\dots,n} \mu(V_j, W)\varepsilon_j$.

Parameter estimation

The parameter to data map takes the form

$$y \mapsto u(y) \mapsto (\ell_i(u(y))_{i=1,\dots,m}.$$

We have discussed the optimal recovery of $u(y)$ from the measured data $(\ell_i(u(y))_{i=1,\dots,m}$.

Parameter estimation deals with the inverse problem of recovering y from such data.

We can address this problem by using the previous result on the recovery of $u(y)$ and try to identify y from $u(y)$.

Example : diffusion equation

$$-\mathrm{div}(a\nabla u) = f \quad \text{on} \quad D \subset \mathbf{R}^{\mathrm{m}} \quad \text{and} \quad \mathrm{u}_{|\partial \mathrm{D}} = 0,$$

with y uniquely associated to the diffusion coefficient $a = a(y)$.

For a fixed right side f, let us denote by $u_a \in V = H^1_0(D)$ the solution for a given a.

Question : can we identify a from u_a in a stable manner ?

Here, stable identification means an estimate of the form $\|a - b\|_X \leq C\|u_a - u_b\|_V$ in some norm $\|\cdot\|_X$ and for any a and b in some class $\mathcal{A}$.

The most natural class is $\mathcal{A}_0 := \{a \in L^\infty(D) \; : \; r < a < R\}$ for some $0 < r < R < \infty$.

Some basic facts

1. Let T_a be the operator $u \mapsto \mathrm{div}(a\nabla u)$ and S_a be its inverse from V' to V. Then, one has

$$\|T_a - T_b\|_{V\to V'} = \|a-b\|_{L^\infty},$$

and for all $a, b \in \mathcal{A}_0$, one has

$$r^2\|S_a - S_b\|_{V'\to V} \leq \|a-b\|_{L^\infty} \leq R^2\|S_a - S_b\|_{V'\to V}.$$

In particular this means that there exist a right side $f = f(a,b)$ such that

$$\|a-b\|_{L^\infty} \leq R^2\|u_a - u_b\|_V$$

However this is not what we look for, since this f changes with (a,b).

2. There exists right sides f such that $a \mapsto u_a$ is not injective : take a smooth v compactly supported in D, and such that v is constant in a subdomain $\tilde{D} \subset D$. For any a and b that differ only on $\tilde{D}$, we have

$$f := -\mathrm{div}(a\nabla v) = -\mathrm{div}(b\nabla v),$$

and so $v = u_a = u_b$ for this f.

Stable identifiability

The problem of identifying a from u_a has been studied since the 1980's e.g. by Falk, Kohn and Lowe, Hoffman and Sprekel, typically with a boundary condition $a\nabla u \cdot \mathbf{n} = g$ and the additional assumption that ∇u does not vanish.

Bonito-Cohen-DeVore-Petrova-Welper (2016) : stable identifiability results for the Dirichlet problem for certain type of right side f.

In both types of results,

(i) Additional smoothness assumptions required : $a \in \mathcal{A}_s := \{a \in \mathcal{A}_0 \ : \ \|a\|_{H^s} \leq M\}$.

(ii) We cannot control $\|a-b\|_{L^\infty}$, instead we estimate $\|a-b\|_{L^2}$.

(iii) The control is not Lipschitz but Hölder continuous.

One sample result : with D a Lipschitz domain and $f \in L^\infty(D)$ such that $f(x) > c > 0$ on D, then for all $a, b \in \mathcal{A}_1$

$$\|a-b\|_{L^2} \leq C\|u_a - u_b\|_V^\alpha, \quad \text{with} \quad \alpha = \frac{1}{6}.$$

Similar result with $a, b \in \mathcal{A}_s$ for some $s > \frac{1}{2}$, with smaller Hölder exponent α.

Univariate case : stable idendifiability for $a, b \in \mathcal{A}_0$ with (sharp) exponent $\alpha = \frac{1}{3}$.

Conclusions

Reduced bases achieve “almost” the same performance as optimal spaces corresponding to Kolmogorov n-width in the sense of preserving algebraic or exponential convergence rates.

Can be used in the reconstruction of $u(y)$ from uncomplete data : one-space (poor man) and multi-space strategies.

Perspective : replace linear spaces V_n by non-linear (sparse) reduced models.

Open questions in parameter estimation : can we treat $a \in \mathcal{A}_s$ for any fixed $s > 0$? What is the sharpest Hölder exponent in the dependence $u_a \mapsto a$? In several dimension is the map $a \mapsto u_a$ injective on $\mathcal{A}_0$?