Estimating the n-width of solution manifolds of parametric PDE's

Albert Cohen

Laboratoire Jacques-Louis Lions
Université Pierre et Marie Curie, Paris
Institut Universitaire de France

joint work with Ronald DeVore

Chicheley Hall, 30-06-2014

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

If E is a finite dimensional space, it approximate K with accuracy

$$\mathrm{dist}(K,E)_X := \max_{u\in K}\min_{v\in E}\|u-v\|_X$$

The n-width of K in the metric X is

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

Benchmark for linear approximation methods applied to the elements from K.

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

If $X = L^p(D)$ for some bounded Lipschitz domain $D \subset \mathbb{R}^d$ and K is the unit ball of $W^{s,p}(D)$ it is known that

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

Upper bound : approximation by a specific method.

Lower bound : diversity in K.

Curse of dimensionality : exponential growth in d of the needed n to reach accuracy ε.

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

If E is a finite dimensional space, it approximate K with accuracy

$$\mathrm{dist}(K, E)_X := \max_{u \in K} \min_{v \in E} \|u - v\|_X$$

The n-width of K in the metric X is

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

Benchmark for linear approximation methods applied to the elements from K.

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

If $X = L^p(D)$ for some bounded Lipschitz domain $D \subset \mathbb{R}^d$ and K is the unit ball of $W^{s,p}(D)$ it is known that

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

Upper bound : approximation by a specific method.

Lower bound : diversity in K.

Curse of dimensionality : exponential growth in d of the needed n to reach accuracy ε.

Parametric and stochastic PDE's

We are interested in PDE's of the general form

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

where u is the unkown and a is a parameter which is either finite or infinite dimensional. Typically

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

where V, X, W are Banach spaces and a ranges in some compact set $K \subset X$.

Model 1 : steady state linear diffusion equation.

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

where $f \in L^2(D)$ is fixed.

In this example $\mathcal{P}(u, a) = f + \mathrm{div}(a\nabla u)$ and the spaces are

$$X = L^\infty(D), \quad V = H^1_0(D), \quad W = V' = H^{-1}(D).$$

Parametric and stochastic PDE's

We are interested in PDE's of the general form

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

where u is the unkown and a is a parameter which is either finite or infinite dimensional. Typically

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

where V, X, W are Banach spaces and a ranges in some compact set $K \subset X$.

Model 1 : steady state linear diffusion equation.

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

where $f \in L^2(D)$ is fixed.

In this example $\mathcal{P}(u,a) = f + \mathrm{div}(a\nabla u)$ and the spaces are

$$X = L^\infty(D), \quad V = H_0^1(D), \quad W = V' = H^{-1}(D).$$

Assuming well-posedness of the problem in the Banach space V for every $a \in K$ allows us to define the solution map from K to V

$$a \mapsto u(a)$$

For Model 1, this is done by assuming that

$$0 < r \leq a \leq R, \qquad a \in K \subset X = L^\infty(D).$$

Then Lax-Milgram theory ensures existence in $V = H_0^1(D)$.

A priori bound : the solution map is bounded from K to V. :

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

Note that $a \mapsto u(a)$ is nonlinear.

We also define the solution manifold

$$\mathcal{M} := u(K) = \{u(a) \; : \; a \in K\} \subset V,$$

that is, the set of all solutions as we vary the parameter.

Problem : estimate $d_n(\mathcal{M})_V$.

The solution map and solution manifold

Assuming well-posedness of the problem in the Banach space V for every $a \in K$ allows us to define the solution map from K to V

$$a \mapsto u(a)$$

For Model 1, this is done by assuming that

$$0 < r \leq a \leq R, \qquad a \in K \subset X = L^\infty(D).$$

Then Lax-Milgram theory ensures existence in $V = H_0^1(D)$.

A priori bound : the solution map is bounded from K to V. :

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

Note that $a \mapsto u(a)$ is nonlinear.

We also define the solution manifold

$$M := u(K) = \{u(a) \; : \; a \in K\} \subset V,$$

that is, the set of all solutions as we vary the parameter.

Problem : estimate $d_n(M)_V$.

The solution map and solution manifold

Assuming well-posedness of the problem in the Banach space V for every $a \in K$ allows us to define the solution map from K to V

$$a \mapsto u(a)$$

For Model 1, this is done by assuming that

$$0 < r \leq a \leq R, \qquad a \in K \subset X = L^\infty(D).$$

Then Lax-Milgram theory ensures existence in $V = H_0^1(D)$.

A priori bound : the solution map is bounded from K to V. :

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

Note that $a \mapsto u(a)$ is nonlinear.

We also define the solution manifold

$$M := u(K) = \{u(a) \ : \ a \in K\} \subset V,$$

that is, the set of all solutions as we vary the parameter.

Problem : estimate $d_n(M)_V$.

Motivation : reduced modeling

The parameter may be deterministic (control, optimization, inverse problems) or random (uncertainty modeling and propagation, risk assessment).

These applications often requires many queries of $u(a)$, and therefore in principle running many times a numerical solver. We want to avoid this.

Reduced modeling : find low dimension spaces that simultaneously approximate well all solutions to the parametric PDE.

The Kolmogorov n-width $d_n(M)_X$ is thus a benchmark for reduced modeling method.

Similar benchmark for approximation in an mean square sense : PCA.

The fact that "$d_n(M)_X$ is small for moderate n" is often taken as a assumption to justify the use of reduced modeling method, but rarely proved.

Reduced bases (Maday-Patera) : define a reduced modeling space $E_n = \mathrm{span}\{u_1, \dots, u_n\}$, where the u_i are particular instances (snapshots) from the solution manifold

$$u_i = u(a_i)$$

for some $a_1, \dots, a_n \in K$.

Motivation : reduced modeling

The parameter may be deterministic (control, optimization, inverse problems) or random (uncertainty modeling and propagation, risk assessment).

These applications often requires many queries of $u(a)$, and therefore in principle running many times a numerical solver. We want to avoid this.

Reduced modeling : find low dimension spaces that simultaneously approximate well all solutions to the parametric PDE.

The Kolmogorov n-width $d_n(M)_X$ is thus a benchmark for reduced modeling method.

Similar benchmark for approximation in an mean square sense : PCA.

The fact that "$d_n(M)_X$ is small for moderate n" is often taken as a assumption to justify the use of reduced modeling method, but rarely proved.

Reduced bases (Maday-Patera) : define a reduced modeling space $E_n = \mathrm{span}\{u_1, \dots, u_n\}$, where the u_i are particular instances (snapshots) from the solution manifold

$$u_i = u(a_i)$$

for some $a_1, \dots, a_n \in K$.

Motivation : reduced modeling

The parameter may be deterministic (control, optimization, inverse problems) or random (uncertainty modeling and propagation, risk assessment).

These applications often requires many queries of $u(a)$, and therefore in principle running many times a numerical solver. We want to avoid this.

Reduced modeling : find low dimension spaces that simultaneously approximate well all solutions to the parametric PDE.

The Kolmogorov n-width $d_n(M)_X$ is thus a benchmark for reduced modeling method.

Similar benchmark for approximation in an mean square sense : PCA.

The fact that "$d_n(M)_X$ is small for moderate n" is often taken as a assumption to justify the use of reduced modeling method, but rarely proved.

Reduced bases (Maday-Patera) : define a reduced modeling space $E_n = \mathrm{span}\{u_1, \dots, u_n\}$, where the u_i are particular instances (snapshots) from the solution manifold

$$u_i = u(a_i)$$

for some $a_1, \dots, a_n \in K$.

Greedy selection : having selected $u_1, \dots, u_{n-1} \in M$, choose the next instance by

$$\|u_n - P_{E_{n-1}} u_n\|_V = \max_{v \in M} \|v - P_{E_{n-1}} v\|_V,$$

where P_E is the orthogonal projector onto E

This algorithm is not realistic : one cannot compute $\|v - P_{E_{n-1}} v\|_V$ for all $v \in M$, however can be estimate at moderate cost by a-posteriori error analysis. Therefore, one rather applies a weak-greedy algorithm : u_n such that

$$\|u_n - P_{E_{n-1}} u_n\|_V \geq \gamma \max_{v \in M} \|v - P_{E_{n-1}} v\|_V,$$

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

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

Comparison with n-width ? $\sigma_n(M)_V$ can be much larger than $d_n(M)_V$ for an individual n and M : There exists M and n such that $\sigma_n(M)_V \geq 2^n d_n(M)_V$.

However, a more favorable comparison is possible in terms of convergence rates :

Theorem (Binev-Cohen-Dahmen-DeVore-Petrova-Wojtaszczyk, 2011) : For any $s > 0$ one has

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

Greedy selection : having selected $u_1, \dots, u_{n-1} \in M$, choose the next instance by

$$\|u_n - P_{E_{n-1}} u_n\|_V = \max_{v \in M} \|v - P_{E_{n-1}} v\|_V,$$

where P_E is the orthogonal projector onto E

This algorithm is not realistic : one cannot compute $\|v - P_{E_{n-1}} v\|_V$ for all $v \in M$, however can be estimate at moderate cost by a-posteriori error analysis. Therefore, one rather applies a weak-greedy algorithm : u_n such that

$$\|u_n - P_{E_{n-1}} u_n\|_V \geq \gamma \max_{v \in M} \|v - P_{E_{n-1}} v\|_V,$$

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

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

Comparison with n-width ? $\sigma_n(M)_V$ can be much larger than $d_n(M)_V$ for an individual n and M : There exists M and n such that $\sigma_n(M)_V \geq 2^n d_n(M)_V$.

However, a more favorable comparison is possible in terms of convergence rates :

Theorem (Binev-Cohen-Dahmen-DeVore-Petrova-Wojtaszczyk, 2011) : For any $s > 0$ one has

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

Our main result

For a compact set $K \subset X$ and a continuous mapping $u : K \longrightarrow V$, we would like to control the decay $d_n(M)_V$ from $d_n(K)_X$, where $M = u(K)$.

Note that if u was a linear mapping, we would simply have

$$d_n(M)_V \leq C d_n(K)_X, \qquad C := \|u\|_{\mathcal{L}(X,V)}.$$

The following result shows that nonlinear holomorphic maps behave almost like linear maps with respect to the asymptotic decay of N-widths.

Theorem (Cohen-DeVore, 2014) : Let X, V be complex Banach spaces and let

$$K \subset O \subset X,$$

with K compact and O open sets. Assume that there is an extension

$$u : O \longrightarrow V$$

is uniformly bounded and holomorphic (Frechet differentiable in the sense of complex Banach space). Then, for all $t > 0$,

$$\sup_{n \geq 1} n^t d_n(K)_X < \infty \Rightarrow \sup_{n \geq 1} n^s d_n(M)_X < \infty, \qquad s < t - 1.$$

Loss of 1 in the rate : $s < t - 1$. This may be a defect from our method of proof.

"Cartesian proof" : uses scalar parametrization of K and polynomial approximations of the resulting map. A more direct geometrical proof would be welcome.

This result gives upper bounds on $d_n(M)_V$. Lower bounds are an open problem.

In this result, the compact set K may consist only of real valued a, however the open set $O \subset X$ need to contain complex valued a.

Applies to Model 1, with $O := \{a \in X = L^\infty(D) \; : \; \mathrm{Re}(a(x)) > r > 0, \; x \in D\}$. Indeed, Lax-Milgram still applies and ensures that the solution map is uniformly bounded on O. In addition

$$a \to A : v \to \mathrm{div}(a\nabla v) \to A^{-1} \to u(a) = A^{-1}f,$$

is a chain of holomorphic maps between $O \to \mathcal{L}(V, V') \to \mathcal{L}(V', V) \to V$.

For other models, in particular nonlinear equations, the bounded holomorphic extension may be more problematic. Example :

$$u^3 - \mathrm{div}(\exp(a)\nabla u) = f,$$

with $A \leq a \leq B$ for all $a \in K \subset X = L^\infty(D)$.

One way to construct it is by using the implicit function theorem : if $\mathcal{P}$ is holomorphic from $X \times V$ to W and if $\partial_u \mathcal{P}(u(a), a) : V \to W$ is invertible for all $a \in K$, then the IFT in complex Banach space allows to extend $a \mapsto u(a)$ on open balls around each $a \in K$, and by compactness this gives O.

Loss of 1 in the rate : $s < t - 1$. This may be a defect from our method of proof.

"Cartesian proof" : uses scalar parametrization of K and polynomial approximations of the resulting map. A more direct geometrical proof would be welcome.

This result gives upper bounds on $d_n(M)_V$. Lower bounds are an open problem.

In this result, the compact set K may consist only of real valued a, however the open set $O \subset X$ need to contain complex valued a.

Applies to Model 1, with $O := \{a \in X = L^\infty(D) \ : \ \mathrm{Re}(a(x)) > r > 0, \ x \in D\}$. Indeed, Lax-Milgram still applies and ensures that the solution map is uniformly bounded on O. In addition

$$a \to A : v \to \mathrm{div}(a\nabla v) \to A^{-1} \to u(a) = A^{-1}f,$$

is a chain of holomorphic maps between $O \to \mathcal{L}(V, V') \to \mathcal{L}(V', V) \to V$.

For other models, in particular nonlinear equations, the bounded holomorphic extension may be more problematic. Example :

$$u^3 - \mathrm{div}(\exp(a)\nabla u) = f,$$

with $A \le a \le B$ for all $a \in K \subset X = L^\infty(D)$.

One way to construct it is by using the implicit function theorem : if $\mathcal{P}$ is holomorphic from $X \times V$ to W and if $\partial_u \mathcal{P}(u(a), a) : V \to W$ is invertible for all $a \in K$, then the IFT in complex Banach space allows to extend $a \mapsto u(a)$ on open balls around each $a \in K$, and by compactness this gives O.

Loss of 1 in the rate : $s < t - 1$. This may be a defect from our method of proof.

"Cartesian proof" : uses scalar parametrization of K and polynomial approximations of the resulting map. A more direct geometrical proof would be welcome.

This result gives upper bounds on $d_n(M)_V$. Lower bounds are an open problem.

In this result, the compact set K may consist only of real valued a, however the open set $O \subset X$ need to contain complex valued a.

Applies to Model 1, with $O := \{a \in X = L^\infty(D) \;:\; \mathrm{Re}(a(x)) > r > 0,\; x \in D\}$. Indeed, Lax-Milgram still applies and ensures that the solution map is uniformly bounded on O. In addition

$$a \to A : v \to \mathrm{div}(a\nabla v) \to A^{-1} \to u(a) = A^{-1}f,$$

is a chain of holomorphic maps between $O \to \mathcal{L}(V, V') \to \mathcal{L}(V', V) \to V$.

For other models, in particular nonlinear equations, the bounded holomorphic extension may be more problematic. Example :

$$u^3 - \mathrm{div}(\exp(a)\nabla u) = f,$$

with $A \leq a \leq B$ for all $a \in K \subset X = L^\infty(D)$.

One way to construct it is by using the implicit function theorem : if $\mathcal{P}$ is holomorphic from $X \times V$ to W and if $\partial_u \mathcal{P}(u(a), a) : V \to W$ is invertible for all $a \in K$, then the IFT in complex Banach space allows to extend $a \mapsto u(a)$ on open balls around each $a \in K$, and by compactness this gives O.

Remarks and applications.

Loss of 1 in the rate : $s < t - 1$. This may be a defect from our method of proof.

"Cartesian proof" : uses scalar parametrization of K and polynomial approximations of the resulting map. A more direct geometrical proof would be welcome.

This result gives upper bounds on $d_n(M)_V$. Lower bounds are an open problem.

In this result, the compact set K may consist only of real valued a, however the open set $O \subset X$ need to contain complex valued a.

Applies to Model 1, with $O := \{a \in X = L^\infty(D) \ : \ \mathrm{Re}(a(x)) > r > 0, \ x \in D\}$. Indeed, Lax-Milgram still applies and ensures that the solution map is uniformly bounded on O. In addition

$$a \to A : v \to \mathrm{div}(a\nabla v) \to A^{-1} \to u(a) = A^{-1}f,$$

is a chain of holomorphic maps between $O \to \mathcal{L}(V, V') \to \mathcal{L}(V', V) \to V$.

For other models, in particular nonlinear equations, the bounded holomorphic extension may be more problematic. Example :

$$u^3 - \mathrm{div}(\exp(a)\nabla u) = f,$$

with $A \le a \le B$ for all $a \in K \subset X = L^\infty(D)$.

One way to construct it is by using the implicit function theorem : if $\mathcal{P}$ is holomorphic from $X \times V$ to W and if $\partial_u\mathcal{P}(u(a), a) : V \to W$ is invertible for all $a \in K$, then the IFT in complex Banach space allows to extend $a \mapsto u(a)$ on open balls around each $a \in K$, and by compactness this gives O.

Proof : 1. scalar parametrization

Assume that $d_n(K)_X \lesssim n^{-s}$. Then, there exists a family $(\psi_j)_{j\geq 1}$ of functions in X such that any $a \in K$ is of the form

$$a = a(z) := \sum_{j\geq 1} z_j\psi_j, \quad z := (z_j)_{j\geq 1} \in \mathcal{U} := \otimes_{j\geq 1}\{|z_j| \leq 1\}.$$

and such that

$$\|\psi_j\|_X \lesssim j^{-s}$$

In other words, K is contained in the simpler box shaped domain

$$Q := \{a(z) \; : \; z \in \mathcal{U}\}.$$

Hint : use Auerbach lemma (any finite dimensional space $X_n \subset X$ has a basis $(e_i)_{i=1,\dots,n}$ with dual basis $(\tilde{e}_i)_{i=1,\dots,n}$ such that $\|e_i\|_X = \|\tilde{e}_i\|_{X'} = 1$).

Strong geometrical simplification !

This is at the origin of the loss of 1 in the rate : $\|\psi_j\|_X \lesssim j^{-s}$ does not allow us to retrieve better than $d_n(Q)_X \lesssim n^{-(s-1)}$.

We are not ensured that $Q \subset O$ and therefore that u is defined on Q. We first assume that this holds and fix this problem later. Therefore $d_n(M)_V \leq d_n(u(Q))_V$

Proof : 1. scalar parametrization

Assume that $d_n(K)_X \lesssim n^{-s}$. Then, there exists a family $(\psi_j)_{j\geq 1}$ of functions in X such that any $a \in K$ is of the form

$$a = a(z) := \sum_{j\geq 1} z_j \psi_j, \quad z := (z_j)_{j\geq 1} \in \mathcal{U} := \otimes_{j\geq 1}\{|z_j| \leq 1\}.$$

and such that

$$\|\psi_j\|_X \lesssim j^{-s}$$

In other words, K is contained in the simpler box shaped domain

$$Q := \{a(z) \ : \ z \in \mathcal{U}\}.$$

Hint : use Auerbach lemma (any finite dimensional space $X_n \subset X$ has a basis $(e_i)_{i=1,\dots,n}$ with dual basis $(\tilde{e}_i)_{i=1,\dots,n}$ such that $\|e_i\|_X = \|\tilde{e}_i\|_{X'} = 1$).

Strong geometrical simplification !

This is at the origin of the loss of 1 in the rate : $\|\psi_j\|_X \lesssim j^{-s}$ does not allow us to retrieve better than $d_n(Q)_X \lesssim n^{-(s-1)}$.

We are not ensured that $Q \subset O$ and therefore that u is defined on Q. We first assume that this holds and fix this problem later. Therefore $d_n(M)_V \leq d_n(u(Q))_V$

Proof : 2. polynomial approximation

We show that $d_n(u(Q))_V \lesssim n^{-t}$ for $t < s-1$, using the parametrization of $u(Q)$ by the scalar solution map

$$z \in \mathcal{U} \mapsto u(z) = u(a(z)) \in V.$$

Strategy : build a numerical approximation of this map by an optimal truncation of power series $u(z) = \sum_{\nu \in \mathcal{F}} t_\nu z^\nu$, where

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

where $\mathcal{F}$ is the set of all finitely supported sequences of integers (finitely many $\nu_j \neq 0$). This series is indexed by countably many integers.

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

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

Resulting upper bound on n-width : with $E_n := \text{span}\{t_\nu \; : \; \nu \in \Lambda\}$, we have

$$d_n(u(Q))_V \leq \sup_{z \in \mathcal{U}} \min_{w \in E_n} \|u(z) - w\|_V \leq \sup_{z \in \mathcal{U}} \Big\| u(z) - \sum_{\nu \in \Lambda} t_\nu z^\nu \Big\|_V.$$

We show that $d_n(u(Q))_V \lesssim n^{-t}$ for $t < s-1$, using the parametrization of $u(Q)$ by the scalar solution map

$$z \in \mathcal{U} \mapsto u(z) = u(a(z)) \in V.$$

Strategy : build a numerical approximation of this map by an optimal truncation of power series $u(z) = \sum_{\nu \in \mathcal{F}} t_\nu z^\nu$, where

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

where $\mathcal{F}$ is the set of all finitely supported sequences of integers (finitely many $\nu_j \neq 0$). This series is indexed by countably many integers.

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

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

Resulting upper bound on n-width : with $E_n := \mathrm{span}\{t_\nu \; : \; \nu \in \Lambda\}$, we have

$$d_n(u(Q))_V \leq \sup_{z\in\mathcal{U}} \min_{w \in E_n} \|u(z) - w\|_V \leq \sup_{z \in \mathcal{U}} \Big\| u(z) - \sum_{\nu\in\Lambda} t_\nu z^\nu \Big\|_V.$$

We show that $d_n(u(Q))_V \lesssim n^{-t}$ for $t < s-1$, using the parametrization of $u(Q)$ by the scalar solution map

$$z \in \mathcal{U} \mapsto u(z) = u(a(z)) \in V.$$

Strategy : build a numerical approximation of this map by an optimal truncation of power series $u(z) = \sum_{\nu \in \mathcal{F}} t_\nu z^\nu$, where

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

where $\mathcal{F}$ is the set of all finitely supported sequences of integers (finitely many $\nu_j \neq 0$). This series is indexed by countably many integers.

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

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

Resulting upper bound on n-width : with $E_n := \mathrm{span}\{t_\nu \ : \ \nu \in \Lambda\}$, we have

$$d_n(u(Q))_V \leq \sup_{z \in \mathcal{U}} \min_{w \in E_n} \|u(z) - w\|_V \leq \sup_{z \in \mathcal{U}} \Big\| u(z) - \sum_{\nu \in \Lambda} t_\nu z^\nu \Big\|_V.$$

Proof : 3. best n-term approximation and summability

By triangle inequality we have

$$d_n(u(Q))_V \leq \sup_{z \in \mathcal{U}} \|u(z) - u_\Lambda(z)\|_V \leq \sup_{z \in \mathcal{U}} \|\sum_{\nu \notin \Lambda} t_\nu z^\nu\|_V \leq \sum_{\nu \notin \Lambda} \|t_\nu\|_V.$$

Best n-term approximation in the $\ell^1(\mathcal{F})$ norm : use for Λ the n largest $\|t_\nu\|_V$.

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

$$\sum_{\nu \notin \Lambda} \|t_\nu\|_V \leq C n^{-t}, \quad t := \frac{1}{p} - 1, \quad C := \|(\|t_\nu\|_V)\|_p.$$

Proof : with $(t_k)_{k>0}$ the decreasing rearrangement, we combine

$$\sum_{\nu \notin \Lambda} \|t_\nu\|_V = \sum_{k>n} t_k = \sum_{k>n} t_k^{1-p} t_k^p \leq t_n^{1-p} C^p \quad \text{and} \quad n t_n^p \leq \sum_{k=1}^{n} t_k^p \leq C^p.$$

The ℓ^p summability of $(\|t_\nu\|_V)_{\nu \in \mathcal{F}}$ is based on the following fundamental result : if u is bounded and holomorphic on O and $Q \subset O$, then for any $p < 1$,

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

Here $\|\psi_j\|_X \lesssim j^{-s}$ implies that $(\|\psi_j\|_X)_{j>0} \in \ell^p(\mathbb{N})$ for any p such that $sp > 1$ and thus we obtain that $d_n(u(Q))_V \lesssim n^{-t}$ for any $t = \frac{1}{p} - 1 < s - 1$.

By triangle inequality we have

$$d_n(u(Q))_V \leq \sup_{z\in\mathcal{U}} \|u(z)-u_\Lambda(z)\|_V \leq \sup_{z\in\mathcal{U}} \|\sum_{\nu\notin\Lambda} t_\nu z^\nu\|_V \leq \sum_{\nu\notin\Lambda} \|t_\nu\|_V.$$

Best n-term approximation in the $\ell^1(\mathcal{F})$ norm : use for Λ the n largest $\|t_\nu\|_V$.

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

$$\sum_{\nu\notin\Lambda} \|t_\nu\|_V \leq Cn^{-t}, \quad t := \frac{1}{p}-1, \quad C := \|(\|t_\nu\|_V)\|_p.$$

Proof : with $(t_k)_{k>0}$ the decreasing rearrangement, we combine

$$\sum_{\nu\notin\Lambda} \|t_\nu\|_V = \sum_{k>n} t_k = \sum_{k>n} t_k^{1-p} t_k^p \leq t_n^{1-p} C^p \quad \text{and} \quad nt_n^p \leq \sum_{k=1}^n t_k^p \leq C^p.$$

The ℓ^p summability of $(\|t_\nu\|_V)_{\nu\in\mathcal{F}}$ is based on the following fundamental result : if u is bounded and holomorphic on O and $Q \subset O$, then for any $p<1$,

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

Here $\|\psi_j\|_X \lesssim j^{-s}$ implies that $(\|\psi_j\|_X)_{j>0} \in \ell^p(\mathbb{N})$ for any p such that $sp>1$ and thus we obtain that $d_n(u(Q))_V \lesssim n^{-t}$ for any $t = \frac{1}{p}-1 < s-1$.

By triangle inequality we have

$$d_n(u(Q))_V \le \sup_{z\in\mathcal{U}} \|u(z)-u_\Lambda(z)\|_V \le \sup_{z\in\mathcal{U}} \|\sum_{\nu\notin\Lambda} t_\nu z^\nu\|_V \le \sum_{\nu\notin\Lambda} \|t_\nu\|_V.$$

Best n-term approximation in the $\ell^1(\mathcal{F})$ norm : use for Λ the n largest $\|t_\nu\|_V$.

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

$$\sum_{\nu\notin\Lambda} \|t_\nu\|_V \le Cn^{-t}, \quad t:=\frac{1}{p}-1, \quad C:=\|(\|t_\nu\|_V)\|_p.$$

Proof : with $(t_k)_{k>0}$ the decreasing rearrangement, we combine

$$\sum_{\nu\notin\Lambda} \|t_\nu\|_V = \sum_{k>n} t_k = \sum_{k>n} t_k^{1-p} t_k^p \le t_n^{1-p} C^p \quad \text{and} \quad n t_n^p \le \sum_{k=1}^{n} t_k^p \le C^p.$$

The ℓ^p summability of $(\|t_\nu\|_V)_{\nu\in\mathcal{F}}$ is based on the following fundamental result : if u is bounded and holomorphic on O and $Q\subset O$, then for any $p<1$,

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

Here $\|\psi_j\|_X \lesssim j^{-s}$ implies that $(\|\psi_j\|_X)_{j>0} \in \ell^p(\mathbb{N})$ for any p such that $sp>1$ and thus we obtain that $d_n(u(Q))_V \lesssim n^{-t}$ for any $t=\frac{1}{p}-1<s-1$.

Proof : 4. estimates of Taylor coefficients (Cohen-DeVore-Schwab, 2011)

The map $z \mapsto u(z)$ is bounded on $\mathcal{U}$ and holomorphic in each variable z_j, since we have assumed that $Q = a(\mathcal{U}) \subset \mathcal{O}$. For any sequence $\rho = (\rho_j)_{j\geq 1}$ such that $\rho_j \geq 1$, we consider the polydisc

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

If $\sum_{j\geq 1}(\rho_j - 1)\|\psi_j\|_X \leq \varepsilon$ for $\varepsilon > 0$ sufficiently small, then $a(\mathcal{U}_\rho) \subset \mathcal{O}$, and therefore u is bounded and holomorphic on $\mathcal{U}_\rho$.

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 $|u^{(n)}(0)| \leq n! b^{-n} \max_{|z|\leq b} |u(z)|$.

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

$$\|\partial^\nu u_{|z=0}\|_V \leq C\nu! \prod_{j>0} \rho_j^{-\nu_j} = C\nu!\rho^{-\nu}$$

where C is the uniform bound for u on $\mathcal{O}$, and therefore

$$\|t_\nu\|_V \leq C \inf\Big\{\rho^{-\nu} \;\; ; \;\; \rho \;\text{ s.t. }\; \sum_{j\geq 1}(\rho_j - 1)\|\psi_j\|_X \leq \varepsilon, \;\; x \in D\Big\}.$$

Optimizing on ρ gives a specific $\rho = \rho(\nu)$ for which we prove that

$$(\|\psi_j\|_X)_{j\geq 1} \in \ell^p(\mathbb{N}) \Rightarrow (\rho(\nu)^{-\nu})_{\nu\in\mathcal{F}} \in \ell^p(\mathcal{F}).$$

The map $z \mapsto u(z)$ is bounded on $\mathcal{U}$ and holomorphic in each variable z_j, since we have assumed that $Q = a(\mathcal{U}) \subset O$. For any sequence $\rho = (\rho_j)_{j\geq 1}$ such that $\rho_j \geq 1$, we consider the polydisc

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

If $\sum_{j\geq 1}(\rho_j - 1)\|\psi_j\|_X \leq \varepsilon$ for $\varepsilon > 0$ sufficiently small, then $a(\mathcal{U}_\rho) \subset O$, and therefore u is bounded and holomorphic on $\mathcal{U}_\rho$.

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 $|u^{(n)}(0)| \leq n!b^{-n} \max_{|z|\leq b} |u(z)|$.

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

$$\|\partial^\nu u_{|z=0}\|_V \leq C\nu! \prod_{j>0} \rho_j^{-\nu_j} = C\nu!\rho^{-\nu}$$

where C is the uniform bound for u on O, and therefore

$$\|t_\nu\|_V \leq C \inf\Big\{\rho^{-\nu} \;\; ; \;\; \rho \;\; \text{s.t.} \;\; \sum_{j\geq 1}(\rho_j - 1)\|\psi_j\|_X \leq \varepsilon, \;\; x \in D\Big\}.$$

Optimizing on ρ gives a specific $\rho = \rho(\nu)$ for which we prove that

$$(\|\psi_j\|_X)_{j\geq 1} \in \ell^p(\mathbb{N}) \Rightarrow (\rho(\nu)^{-\nu})_{\nu\in\mathcal{F}} \in \ell^p(\mathcal{F}).$$

In the case where Q is not contained in O, we fix this problem by local parametrizations : using the compactness of K, we build a covering

$$K \subset \cup_{i=1}^{P} Q_i,$$

with

$$Q_i := \left\{ a_i + \sum_{j \geq 1} z_j \psi_j^* \ : \ z \in \mathcal{U} \right\} \subset O,$$

and such that $\|\psi_j^*\|_X \lesssim j^{-s}$.

Therefore

$$M = u(K) \subset \cup_{i=1,\dots,P} M_i, \qquad M_i = u(Q_i).$$

and using the same techniques, we prove

$$d_n(M_i)_V \lesssim n^{-t}, \quad i = 1, \dots, P, \quad t < s - 1.$$

This amounts to using piecewise polynomial approximations of the solution map on a fixed partition of K.

We obtain the same result for $d_n(M)_V$ up to a multiplicative constant, since

$$d_{nP}(M)_V \leq \sup_{i=1,\dots,P} d_n(M_i)_V.$$

Proof : 5. localization

In the case where Q is not contained in O, we fix this problem by local parametrizations : using the compactness of K, we build a covering

$$K \subset \cup_{i=1}^{P} Q_i,$$

with

$$Q_i := \Big\{ a_i + \sum_{j \geq 1} z_j \psi_j^* \ : \ z \in \mathcal{U} \Big\} \subset O,$$

and such that $\|\psi_j^*\|_X \lesssim j^{-s}$.

Therefore

$$M = u(K) \subset \cup_{i=1,\dots,P} M_i, \qquad M_i = u(Q_i).$$

and using the same techniques, we prove

$$d_n(M_i)_V \lesssim n^{-t}, \quad i = 1, \dots, P, \quad t < s - 1.$$

This amounts to using piecewise polynomial approximations of the solution map on a fixed partition of K.

We obtain the same result for $d_n(M)_V$ up to a multiplicative constant, since

$$d_{nP}(M)_V \leq \sup_{i=1,\dots,P} d_n(M_i)_V.$$

Conclusions

The n-width is asymptotically (almost) preserved by holomorphic maps.

Open questions : exponential rates, lower bounds, less smooth u...

The proof is based on a parametrization and a non-linear approximation process.

Other, more direct, approaches to evaluate $d_n(M)_X$?

Similar result for low-rank approximation in the mean square sense ?

THANKS !

QUESTIONS ?