Sampling discretization and entropy

Vladimir Temlyakov

Workshop, Moscow, March 10, 2021

Two settings

We are interested in discretizing the L_q, $1 \le q \le \infty$, norm of functions defined on a compact subset of $\mathbb{R}^d$. We distinguish right a way two settings:

Two settings

We are interested in discretizing the L_q, $1 \leq q \leq \infty$, norm of functions defined on a compact subset of $\mathbb{R}^d$. We distinguish right a way two settings:

- Functions belong to an N-dimensional subspace X_N. We call such results the Marcinkiewicz-type discretization theorems.

Two settings

We are interested in discretizing the L_q, $1 \le q \le \infty$, norm of functions defined on a compact subset of $\mathbb{R}^d$. We distinguish right a way two settings:

- Functions belong to an N-dimensional subspace X_N. We call such results the Marcinkiewicz-type discretization theorems.
- Functions belong to a given function class.

Two settings

We are interested in discretizing the L_q, $1 \le q \le \infty$, norm of functions defined on a compact subset of $\mathbb{R}^d$. We distinguish right a way two settings:

- Functions belong to an N-dimensional subspace X_N. We call such results the Marcinkiewicz-type discretization theorems.
- Functions belong to a given function class.
  There are different settings and different ingredients, which play important role in this problem.

Sampling discretization with absolute error

Let $W \subset L_q(\Omega, \mu)$, $1 \leq q < \infty$, be a class of continuous on Ω functions. We are interested in estimating the following optimal errors of discretization of the L_q norm of functions from W

Let $W \subset L_q(\Omega,\mu)$, $1 \le q < \infty$, be a class of continuous on Ω functions. We are interested in estimating the following optimal errors of discretization of the L_q norm of functions from W

$$er_m(W, L_q) := \inf_{\xi^1,\dots,\xi^m} \sup_{f\in W} \left| \|f\|_q^q - \frac{1}{m}\sum_{j=1}^m |f(\xi^j)|^q \right|,$$

Sampling discretization with absolute error

Let $W \subset L_q(\Omega, \mu)$, $1 \le q < \infty$, be a class of continuous on Ω functions. We are interested in estimating the following optimal errors of discretization of the L_q norm of functions from W

$$er_m(W, L_q) := \inf_{\xi^1,\dots,\xi^m} \sup_{f \in W} \left| \|f\|_q^q - \frac{1}{m} \sum_{j=1}^m |f(\xi^j)|^q \right|,$$

$$er_m^o(W, L_q) := \inf_{\xi^1,\dots,\xi^m;\lambda_1,\dots,\lambda_m} \sup_{f \in W} \left| \|f\|_q^q - \sum_{j=1}^m \lambda_j |f(\xi^j)|^q \right|.$$

Entropy numbers

For a compact subset Θ of a Banach space B we define the entropy numbers as follows

$$\varepsilon_n(\Theta, B) := \inf\{\varepsilon : \exists f_1, \dots, f_{2^n} \in \Theta : \Theta \subset \cup_{j=1}^{2^n} (f_j + \varepsilon U(B))\}$$

where $U(B)$ is the unit ball of a Banach space B.

Theorem (T1; VT, 2018)

Assume that a class of real functions W is such that for all $f \in W$ we have $\|f\|_\infty \le M$ with some constant M. Also assume that the entropy numbers of W in the uniform norm L_∞ satisfy the condition

$$\varepsilon_n(W, L_\infty) \le Cn^{-r}, \qquad r \in (0, 1/2).$$

Then

$$er_m(W) := er_m(W, L_2) \le Km^{-r}.$$

Comments

Theorem T1 is a rather general theorem, which connects the behavior of absolute errors of discretization with the rate of decay of the entropy numbers. This theorem is derived from known results in supervised learning theory. It is well understood in learning theory that the entropy numbers of the class of priors (regression functions) is the right characteristic in studying the regression problem.

Comments

Theorem T1 is a rather general theorem, which connects the behavior of absolute errors of discretization with the rate of decay of the entropy numbers. This theorem is derived from known results in supervised learning theory. It is well understood in learning theory that the entropy numbers of the class of priors (regression functions) is the right characteristic in studying the regression problem.

- We impose a restriction $r < 1/2$ in Theorem T1 because the probabilistic technique from the supervised learning theory has a natural limitation to $r \leq 1/2$.

Comments

Theorem T1 is a rather general theorem, which connects the behavior of absolute errors of discretization with the rate of decay of the entropy numbers. This theorem is derived from known results in supervised learning theory. It is well understood in learning theory that the entropy numbers of the class of priors (regression functions) is the right characteristic in studying the regression problem.

- We impose a restriction $r < 1/2$ in Theorem T1 because the probabilistic technique from the supervised learning theory has a natural limitation to $r \leq 1/2$.
- It would be interesting to understand if Theorem T1 holds for $r \geq 1/2$.

Let $X \subset \mathbb{R}^d$, $Y \subset \mathbb{R}$ be Borel sets, ρ be a Borel probability measure on a Borel set $Z \subset X \times Y$. For $f : X \to Y$ define *the error*

$$\mathcal{E}(f) := \int_Z (f(\mathbf{x}) - y)^2 d\rho.$$

Connection to learning theory

Let $X \subset \mathbb{R}^d$, $Y \subset \mathbb{R}$ be Borel sets, ρ be a Borel probability measure on a Borel set $Z \subset X \times Y$. For $f : X \to Y$ define *the error*

$$\mathcal{E}(f) := \int_Z (f(\mathbf{x}) - y)^2 d\rho.$$

Let ρ_X be the marginal probability measure of ρ on X, i.e., $\rho_X(S) = \rho(S \times Y)$ for Borel sets $S \subset X$. Define

$$f_\rho(\mathbf{x}) := \mathbb{E}(y|\mathbf{x})$$

to be a conditional expectation of y.

Setting

- The function f_ρ is known in statistics as the *regression function* of ρ. In the sense of error $\mathcal{E}(\cdot)$ the regression function f_ρ is the best to describe the relation between inputs $\mathbf{x} \in X$ and outputs $y \in Y$.

Setting

- The function f_ρ is known in statistics as the *regression function* of ρ. In the sense of error $\mathcal{E}(\cdot)$ the regression function f_ρ is the best to describe the relation between inputs $\mathbf{x} \in X$ and outputs $y \in Y$.
- The goal is to find an estimator $f_{\mathbf{z}}$, on the base of given data $\mathbf{z} := ((\mathbf{x}^1, y_1), \dots, (\mathbf{x}^m, y_m))$ that approximates f_ρ well with high probability.

Setting

- The function f_ρ is known in statistics as the *regression function* of ρ. In the sense of error $\mathcal{E}(\cdot)$ the regression function f_ρ is the best to describe the relation between inputs $\mathbf{x} \in X$ and outputs $y \in Y$.
- The goal is to find an estimator $f_{\mathbf{z}}$, on the base of given data $\mathbf{z} := ((\mathbf{x}^1, y_1), \dots, (\mathbf{x}^m, y_m))$ that approximates f_ρ well with high probability.
- We assume that $(\mathbf{x}^i, y_i)$, $i = 1, \dots, m$ are independent and distributed according to ρ.

Setting

- The function f_ρ is known in statistics as the *regression function* of ρ. In the sense of error $\mathcal{E}(\cdot)$ the regression function f_ρ is the best to describe the relation between inputs $\mathbf{x} \in X$ and outputs $y \in Y$.
- The goal is to find an estimator $f_{\mathbf{z}}$, on the base of given data $\mathbf{z} := ((\mathbf{x}^1, y_1), \dots, (\mathbf{x}^m, y_m))$ that approximates f_ρ well with high probability.
- We assume that $(\mathbf{x}^i, y_i)$, $i = 1, \dots, m$ are independent and distributed according to ρ.
- We measure the error between $f_{\mathbf{z}}$ and f_ρ in the $L_2(\rho_X)$ norm.

Empirical error

We define the *empirical error* of f as

$$\mathcal{E}_{\mathbf{z}}(f) := \frac{1}{m}\sum_{i=1}^{m}(f(\mathbf{x}^i) - y_i)^2.$$

Empirical error

We define the *empirical error* of f as

$$\mathcal{E}_{\mathbf{z}}(f) := \frac{1}{m}\sum_{i=1}^{m}(f(\mathbf{x}^i) - y_i)^2.$$

Let $f \in L_2(\rho_X)$. The *defect function* of f is

$$L_{\mathbf{z}}(f) := L_{\mathbf{z},\rho}(f) := \mathcal{E}(f) - \mathcal{E}_{\mathbf{z}}(f); \quad \mathbf{z} = (z_1, \dots, z_m), \quad z_i = (\mathbf{x}^i, y_i).$$

Empirical error

We define the *empirical error* of f as

$$\mathcal{E}_{\mathbf{z}}(f) := \frac{1}{m}\sum_{i=1}^{m}(f(\mathbf{x}^i) - y_i)^2.$$

Let $f \in L_2(\rho_X)$. The *defect function* of f is

$$L_{\mathbf{z}}(f) := L_{\mathbf{z},\rho}(f) := \mathcal{E}(f) - \mathcal{E}_{\mathbf{z}}(f); \quad \mathbf{z} = (z_1, \dots, z_m), \quad z_i = (\mathbf{x}^i, y_i).$$

We are interested in estimating $L_{\mathbf{z}}(f)$ for functions f coming from a given class W. We assume that ρ and W satisfy the following condition: for all $f \in W$ and any $(\mathbf{x}, y) \in Z$

$$|f(\mathbf{x}) - y| \leq M. \tag{1}$$

Estimate for the defect function

Theorem (T2, S. Konyagin and VT, 2004)

Assume ρ, W *satisfy (1) and*

$$\varepsilon_n(W, L_\infty) \le Dn^{-r}, \quad r \in (0, 1/2).$$

Then for m, η *satisfying* $m\eta^{1/r} \ge C_1(M, D, r)$ *we have*

$$\rho^m\{\mathbf{z} : \sup_{f \in W} |L_{\mathbf{z}}(f)| \ge \eta\} \le C(M, D, r) \exp(-c(M, D, r) m\eta^{1/r}).$$

Marcinkiewicz problem

Let Ω be a compact subset of $\mathbb{R}^d$ with the probability measure μ. We say that a linear subspace X_N of the $L_q(\Omega)$, $1 \le q < \infty$, admits the Marcinkiewicz-type discretization theorem with parameters m, q and constants C_1, C_2 if there exists a set $\{\xi^\nu \in \Omega, \nu = 1, \dots, m\}$ such that for any $f \in X_N$ we have

$$C_1\|f\|_q^q \le \frac{1}{m}\sum_{\nu=1}^{m} |f(\xi^\nu)|^q \le C_2\|f\|_q^q. \qquad (2)$$

Marcinkiewicz problem

Let Ω be a compact subset of $\mathbb{R}^d$ with the probability measure μ. We say that a linear subspace X_N of the $L_q(\Omega)$, $1 \leq q < \infty$, admits the Marcinkiewicz-type discretization theorem with parameters m, q and constants C_1, C_2 if there exists a set $\{\xi^\nu \in \Omega, \nu = 1, \dots, m\}$ such that for any $f \in X_N$ we have

$$C_1\|f\|_q^q \leq \frac{1}{m}\sum_{\nu=1}^{m}|f(\xi^\nu)|^q \leq C_2\|f\|_q^q. \tag{2}$$

In the case $q = \infty$ we define L_∞ as the space of continuous on Ω functions and ask for

$$C_1\|f\|_\infty \leq \max_{1\leq\nu\leq m}|f(\xi^\nu)| \leq \|f\|_\infty. \tag{3}$$

Marcinkiewicz problem

Let Ω be a compact subset of $\mathbb{R}^d$ with the probability measure μ. We say that a linear subspace X_N of the $L_q(\Omega)$, $1 \le q < \infty$, admits the Marcinkiewicz-type discretization theorem with parameters m, q and constants C_1, C_2 if there exists a set $\{\xi^\nu \in \Omega, \nu = 1, \dots, m\}$ such that for any $f \in X_N$ we have

$$C_1\|f\|_q^q \le \frac{1}{m}\sum_{\nu=1}^{m}|f(\xi^\nu)|^q \le C_2\|f\|_q^q. \tag{2}$$

In the case $q = \infty$ we define L_∞ as the space of continuous on Ω functions and ask for

$$C_1\|f\|_\infty \le \max_{1\le\nu\le m}|f(\xi^\nu)| \le \|f\|_\infty. \tag{3}$$

We will also use a brief way to express the above property: $X_N \in \mathcal{M}(m,q)$ or $X_N \in \mathcal{M}(m,q,C_1,C_2)$.

Entropy. Conditional theorem for L_1

Theorem (T3, VT, 2017)

Suppose that a real N-dimensional subspace X_N satisfies the following condition on the entropy numbers of the unit ball $X_N^1 := \{f \in X_N : \|f\|_1 \le 1\}$ with $B \ge 1$

$$\varepsilon_k(X_N^1, L_\infty) \le B \begin{cases} N/k, & k \le N, \\ 2^{-k/N}, & k \ge N. \end{cases}$$

Then there exists a set of $m \le C_1 NB(\log_2(2N\log_2(8B)))^2$ points $\xi^j \in \Omega$, $j = 1, \dots, m$, with large enough absolute constant C_1, such that for any $f \in X_N$ we have

$$\frac{1}{2}\|f\|_1 \le \frac{1}{m}\sum_{j=1}^m |f(\xi^j)| \le \frac{3}{2}\|f\|_1.$$

Entropy. Conditional theorem for L_q

Theorem (T4, DPSTT, 2020)

Let $1 \le q < \infty$. Suppose that a real N-dimensional subspace X_N satisfies the following condition on the entropy numbers of the unit ball $X_N^q := \{f \in X_N : \|f\|_q \le 1\}$

$$\varepsilon_k(X_N^q, L_\infty) \le B \begin{cases} (N/k)^{1/q}, & k \le N, \\ 2^{-k/N}, & k \ge N, \end{cases}$$

with B satisfying the conditions $B \ge 1$ and $\log_2(2B) \le K_1(q)N$. Then there exists a set of $m \le C(q)NB^q(\log_2(2N))^2$ points $\xi^j \in \Omega$, $j = 1, \dots, m$, with large enough constant $C(q)$, such that for any $f \in X_N$ we have

$$\frac{1}{2}\|f\|_q^q \le \frac{1}{m}\sum_{j=1}^{m} |f(\xi^j)|^q \le \frac{3}{2}\|f\|_q^q.$$

Very recently E. Kosov, 2020, proved a theorem with the conditions imposed on the entropy numbers in a weaker metric than the uniform norm. We now formulate his result.
Let $Y_s := \{y_j\}_{j=1}^s \subset \Omega$ be a set of sample points from the domain Ω. Introduce a semi-norm

$$\|f\|_{Y_s} := \|f\|_{L_\infty(Y_s)} := \max_{1\le j\le s} |f(y_j)|.$$

Clearly, for any Y_s we have $\|f\|_{Y_s} \le \|f\|_\infty$.

Entropy. Conditional theorem for L_q continues

Theorem (T5, K, 2020)

Let $1 \le q < \infty$. *There exists a number* $C_1(q) > 0$ *such that for* m *and* B *satisfying* $m \ge C_1(q) N B^q (\log N)^{w(q)}$,

$$w(1) := 2, \quad w(q) := \max(q, 2) - 1, \quad 1 < q < \infty,$$

and for a subspace X_N *satisfying the condition: for any set* $Y_m \subset \Omega$

$$\varepsilon_k(X_N^q, L_\infty(Y_m)) \le B(N/k)^{1/q}, \quad 1 \le k \le N$$

there are points $\xi^j \in \Omega$, $j = 1, \dots, m$, *such that for any* $f \in X_N$ *we have*

$$\frac{1}{2}\|f\|_q^q \le \frac{1}{m}\sum_{j=1}^m |f(\xi^j)|^q \le \frac{3}{2}\|f\|_q^q.$$

(i) There exists a constant $K_1 > 1$ such that

$$\|f\|_\infty \le (K_1 N)^{1/2} \|f\|_2, \quad \forall f \in X_N. \tag{4}$$

(ii) There exists a constant $K_2 > 1$ such that

$$\|f\|_\infty \le K_2 \|f\|_{\log N}, \quad \forall f \in X_N. \tag{5}$$

Bounds for the entropy numbers, $1 \le q \le 2$

Theorem (T6, DPSTT, 2020)

Assume that X_N is an N-dimensional subspace of $L_\infty(\Omega)$ satisfying conditions (4) and (5). Then for each $1 \le q \le 2$, there exists a constant $C(q) > 0$ depending only on q such that

$$\varepsilon_k(X_N^q, L_\infty) \le C(q)(K_1 K_2^2 \log N)^{1/q} \begin{cases} (N/k)^{1/q}, & \text{if } 1 \le k \le N, \\ 2^{-k/N}, & \text{if } k > N. \end{cases}$$

Bounds for the entropy numbers, $2 < q < \infty$

Lemma (L1, K, 2020)

Let $q \in (2, \infty)$. Assume that for any $f \in X_N$ we have

$$\|f\|_\infty \leq M\|f\|_q$$

with some constant M. Then for $k \in [1, N]$ we have for any Y_s

$$\varepsilon_k(X_N^q, L_\infty(Y_s)) \leq C(q)M\left(\frac{\log s}{k}\right)^{1/q}.$$

Lemma (L2, VT, 2020)

Let $q \in (2, \infty)$. *Assume that for any* $f \in X_N$ *we have*

$$\|f\|_\infty \le M\|f\|_q$$

with some constant M. *Also, assume that* $X_N \in \mathcal{M}(s, \infty)$ *with* $s \le aN^c$. *Then for* $k \in [1, N]$ *we have*

$$\varepsilon_k(X_N^q, L_\infty) \le C(q, a, c)M \left(\frac{\log s}{k}\right)^{1/q}.$$

Comments

We now formulate Talagrand's inequality for the entropy numbers. For a Banach space X we define the modulus of smoothness

$$\rho(u) := \rho(X,u) := \sup_{\|x\|=\|y\|=1} \left(\frac{1}{2}(\|x+uy\| + \|x-uy\|) - 1\right).$$

Let $\mathcal{D}_n = \{g_j\}_{j=1}^n$ be a system of elements of cardinality $|\mathcal{D}_n| = n$ in a Banach space X. We equip the linear space $W_n := [\mathcal{D}_n] := \operatorname{span}\{\mathcal{D}_n\}$ with the norm

$$\|f\|_A := \|f\|_{A_1(\mathcal{D}_n)} := \inf\left\{\sum_{j=1}^n |c_j| \,:\, f = \sum_{j=1}^n c_j g_j\right\}.$$

Notations

Denote by $W_{n,A}$ the W_n equipped with the norm $\|\cdot\|_A$. We are interested in the dual norm to the norm $\|\cdot\|_A$, which we denote $\|\cdot\|_U$:

$$\|F\|_U := \|F\|_{U(\mathcal{D}_n)} := \sup_{f\in W_n;\|f\|_A\le 1} |F(f)|.$$

Denote $W^*_{n,U}$ the W^*_n equipped with the norm $\|\cdot\|_U$. Note that $\|\cdot\|_U$ is a semi-norm on the dual to X, space X^*.

Talagrand's inequality

Theorem (T7, Tal., 1995)

Let X be q-smooth: $\rho(X,u) \le \gamma u^q$, $1 < q \le 2$ and let $\mathcal{D}_n$ be a normalized system in X of cardinality $|\mathcal{D}_n| = n$. Then for the unit ball $B(X^*)$ of X^* we have

$$\varepsilon_k(B(X^*), \|\cdot\|_{U(\mathcal{D}_n)}) \le C(X)\left(\frac{\log n}{k}\right)^{1-1/q}, \quad k = 1,\dots,n.$$

We note that Talagrand's inequality was slightly improved in V. Temlyakov, 2020: The factor $\log n$ was replaced by $\log(2n/k)$.

Numerical integration seeks good ways of approximating an integral

$$\int_\Omega f(\mathbf{x})d\mu$$

by an expression of the form

$$\Lambda_m(f,\xi) := \sum_{j=1}^{m} \lambda_j f(\xi^j), \tag{6}$$

Numerical integration

Numerical integration seeks good ways of approximating an integral

$$\int_\Omega f(\mathbf{x})d\mu$$

by an expression of the form

$$\Lambda_m(f,\xi) := \sum_{j=1}^{m} \lambda_j f(\xi^j), \tag{6}$$

where $\xi = (\xi^1,\dots,\xi^m), \quad \xi^j \in \Omega, \quad j = 1,\dots,m$. It is clear that we must assume that f is integrable and defined at the points $\xi^1,\dots,\xi^m$. Expression (6) is called a *cubature formula* (ξ,Λ) (if $\Omega \subset \mathbb{R}^d$, $d \geq 2$) or a *quadrature formula* (ξ,Λ) (if $\Omega \subset \mathbb{R}$) with knots $\xi = (\xi^1,\dots,\xi^m)$ and weights $\Lambda := (\lambda_1,\dots,\lambda_m)$.

Quasi-algebra property.

We begin with a very simple general observation on a connection between norm discretization and numerical integration.

Quasi-algebra property. We say that a function class W has the quasi-algebra property if there exists a constant a such that for any $f, g \in W$ we have $fg/a \in W$.

Quasi-algebra property.

We begin with a very simple general observation on a connection between norm discretization and numerical integration.

Quasi-algebra property. We say that a function class W has the quasi-algebra property if there exists a constant a such that for any $f, g \in W$ we have $fg/a \in W$.

The above property was introduced and studied in detail by H. Triebel. He introduced this property under the name *multiplication algebra*. Normally, the term *algebra* refers to the corresponding property with parameter $a = 1$. To avoid any possible confusions we call it *quasi-algebra*. We refer the reader to the very resent book of Triebel, 2018, which contains results on the multiplication algebra (quasi-algebra) property for a broad range of function spaces.

Upper bound

Proposition (P1; VT, 2018)

Suppose that a function class W has the quasi-algebra property and for any $f \in W$ we have for the complex conjugate function $\bar{f} \in W$. Then for a cubature formula $\Lambda_m(\cdot, \xi)$ we have: for any $f \in W$

$$| \|f\|_2^2 - \Lambda_m(|f|^2, \xi)| \le a \sup_{g \in W} \left| \int_\Omega g d\mu - \Lambda_m(g, \xi) \right|.$$

Function classes

We discuss some classical classes of smooth periodic functions. We begin with a general scheme and then give a concrete example. Let $F \in L_1(\mathbb{T}^d)$ be such that $\hat{F}(\mathbf{k}) \neq 0$ for all $\mathbf{k} \in \mathbb{Z}^d$, where

$$\hat{F}(\mathbf{k}) := \mathcal{F}(F, \mathbf{k}) := (2\pi)^{-d} \int_{\mathbb{T}^d} F(\mathbf{x}) e^{-i(\mathbf{k},\mathbf{x})} d\mathbf{x}.$$

Function classes

We discuss some classical classes of smooth periodic functions. We begin with a general scheme and then give a concrete example. Let $F \in L_1(\mathbb{T}^d)$ be such that $\hat{F}(\mathbf{k}) \neq 0$ for all $\mathbf{k} \in \mathbb{Z}^d$, where

$$\hat{F}(\mathbf{k}) := \mathcal{F}(F, \mathbf{k}) := (2\pi)^{-d} \int_{\mathbb{T}^d} F(\mathbf{x}) e^{-i(\mathbf{k},\mathbf{x})} d\mathbf{x}.$$

Consider the space

$$W_2^F := \{f : f(\mathbf{x}) = J_F(\varphi)(\mathbf{x}) := (2\pi)^{-d} \int_{\mathbb{T}^d} F(\mathbf{x} - \mathbf{y}) \varphi(\mathbf{y}) d\mathbf{y},$$

$$\|\varphi\|_2 < \infty\}.$$

For $f \in W_2^F$ we have $\hat{f}(\mathbf{k}) = \hat{F}(\mathbf{k})\hat{\varphi}(\mathbf{k})$ and, therefore, our assumption $\hat{F}(\mathbf{k}) \neq 0$ for all $\mathbf{k} \in \mathbb{Z}^d$ implies that function φ is uniquely defined by f. Introduce a norm on W_2^F by

$$\|f\|_{W_2^F} := \|\varphi\|_2, \qquad f = J_F(\varphi).$$

Quasi-algebra property for function classes

For $f \in W_2^F$ we have $\hat{f}(\mathbf{k}) = \hat{F}(\mathbf{k})\hat{\varphi}(\mathbf{k})$ and, therefore, our assumption $\hat{F}(\mathbf{k}) \neq 0$ for all $\mathbf{k} \in \mathbb{Z}^d$ implies that function φ is uniquely defined by f. Introduce a norm on W_2^F by

$$\|f\|_{W_2^F} := \|\varphi\|_2, \qquad f = J_F(\varphi).$$

For convenience, with a little abuse of notation we use notation W_2^F for the unit ball of the space W_2^F. We are interested in the following question. Under what conditions on F the fact that $f, g \in W_2^F$ implies that $fg \in W_2^F$ and

$$\|fg\|_{W_2^F} \leq C_0 \|f\|_{W_2^F} \|g\|_{W_2^F}?$$

Quasi-algebra property for function classes

For $f \in W_2^F$ we have $\hat{f}(\mathbf{k}) = \hat{F}(\mathbf{k})\hat{\varphi}(\mathbf{k})$ and, therefore, our assumption $\hat{F}(\mathbf{k}) \neq 0$ for all $\mathbf{k} \in \mathbb{Z}^d$ implies that function φ is uniquely defined by f. Introduce a norm on W_2^F by

$$\|f\|_{W_2^F} := \|\varphi\|_2, \qquad f = J_F(\varphi).$$

For convenience, with a little abuse of notation we use notation W_2^F for the unit ball of the space W_2^F. We are interested in the following question. Under what conditions on F the fact that $f, g \in W_2^F$ implies that $fg \in W_2^F$ and

$$\|fg\|_{W_2^F} \leq C_0 \|f\|_{W_2^F} \|g\|_{W_2^F}?$$

In other words: Which properties of F guarantee that the class W_2^F has the quasi-algebra property? We give a simple sufficient condition.

Sufficient condition for quasi-algebra prperty

Proposition (P2; VT, 2018)

Suppose that for each $\mathbf{n} \in \mathbb{Z}^d$ *we have*

$$\sum_{\mathbf{k}\in\mathbb{Z}^d} |\hat{F}(\mathbf{k})\hat{F}(\mathbf{n}-\mathbf{k})|^2 \le C_0^2|\hat{F}(\mathbf{n})|^2. \tag{7}$$

Then, for any $f, g \in W_2^F$ *we have* $fg \in W_2^F$ *and*

$$\|fg\|_{W_2^F} \le C_0\|f\|_{W_2^F}\|g\|_{W_2^F}.$$

Classes with mixed smoothness

As an example consider the class $\mathbf{W}^r_2$ of functions with bounded mixed derivative. By the definition $\mathbf{W}^r_2 := W^{F_r}_2$ with function $F_r(\mathbf{x})$ defined as follows. For a number $k \in \mathbb{Z}$ denote $k^* := \max(|k|, 1)$. Then for $r > 0$ we define F_r by its Fourier coefficients

$$\hat{F}_r(\mathbf{k}) = \prod_{j=1}^{d} (k_j^*)^{-r}. \quad (8)$$

Classes with mixed smoothness

As an example consider the class $\mathbf{W}^r_2$ of functions with bounded mixed derivative. By the definition $\mathbf{W}^r_2 := W^{F_r}_2$ with function $F_r(\mathbf{x})$ defined as follows. For a number $k \in \mathbb{Z}$ denote $k^* := \max(|k|, 1)$. Then for $r > 0$ we define F_r by its Fourier coefficients

$$\hat{F}_r(\mathbf{k}) = \prod_{j=1}^{d} (k_j^*)^{-r}. \tag{8}$$

Lemma (L3, VT, 2018)

Function $F = F_r$ with $r > 1/2$ satisfies condition (7) with $C_0 = C(r, d)$.

Classes with mixed smoothness

As an example consider the class $\mathbf{W}_2^r$ of functions with bounded mixed derivative. By the definition $\mathbf{W}_2^r := W_2^{F_r}$ with function $F_r(\mathbf{x})$ defined as follows. For a number $k \in \mathbb{Z}$ denote $k^* := \max(|k|, 1)$. Then for $r > 0$ we define F_r by its Fourier coefficients

$$\hat{F}_r(\mathbf{k}) = \prod_{j=1}^{d} (k_j^*)^{-r}. \tag{8}$$

Lemma (L3, VT, 2018)

Function $F = F_r$ with $r > 1/2$ satisfies condition (7) with $C_0 = C(r, d)$.

Lemma L3 and Proposition P1 imply that the class $\mathbf{W}_2^r$ has the quasi-algebra property.

Fibonacci cubature formulas

We now illustrate how a combination of Proposition P1 and known results on numerical integration gives results on discretization. We discuss classes of periodic functions of two variables. Let $\{b_n\}_{n=0}^{\infty}$, $b_0 = b_1 = 1$, $b_n = b_{n-1} + b_{n-2}$, $n \geq 2$, – be the Fibonacci numbers.

We now illustrate how a combination of Proposition P1 and known results on numerical integration gives results on discretization. We discuss classes of periodic functions of two variables. Let $\{b_n\}_{n=0}^{\infty}$, $b_0 = b_1 = 1$, $b_n = b_{n-1} + b_{n-2}$, $n \geq 2$, – be the Fibonacci numbers. For continuous functions of two variables, which are 2π-periodic in each variable, we define cubature formulas

$$\Phi_n(f) := b_n^{-1} \sum_{\mu=1}^{b_n} f\left(2\pi\mu/b_n, 2\pi\{\mu b_{n-1}/b_n\}\right),$$

which are called the Fibonacci cubature formulas. In this definition $\{a\}$ is the fractional part of the number a.

Known result for numerical integration

For a function class W denote

$$\Phi_n(W) := \sup_{f\in W} |\Phi_n(f) - \hat{f}(\mathbf{0})|.$$

The following result is known

$$\Phi_n(\mathbf{W}_2^r) \asymp b_n^{-r}(\log b_n)^{1/2}, \qquad r > 1/2. \tag{9}$$

Combining (9) with Proposition P1 we obtain the following discretization result.

Discretization result

Theorem (T8; VT, 2018)

Let $d = 2$, $r > 1/2$ *and* μ *be the Lebesgue measure on* $[0, 2\pi]^2$. *Then*

$$er_m(\mathbf{W}^r_2, L_2) \le C(r)m^{-r}(\log m)^{1/2}.$$

Smoothness classes

We proved the following bounds for the class $\mathbf{W}^r_2$ of functions on d variables with bounded in L_2 mixed derivative.

We proved the following bounds for the class $\mathbf{W}^r_2$ of functions on d variables with bounded in L_2 mixed derivative.

Theorem (T9; VT, 2018)

Let $r > 1/2$ and μ be the Lebesgue measure on $[0, 2\pi]^d$. Then

$$er^o_m(\mathbf{W}^r_2, L_2) \asymp m^{-r}(\log m)^{(d-1)/2}.$$

Lower bound

Assume that a class of real functions $W \subset \mathcal{C}(\Omega)$ has the following extra property.

Property A. For any $f \in W$ we have $f^+ := (f+1)/2 \in W$ and $f^- := (f-1)/2 \in W$.

Assume that a class of real functions $W \subset \mathcal{C}(\Omega)$ has the following extra property.

Property A. For any $f \in W$ we have $f^+ := (f+1)/2 \in W$ and $f^- := (f-1)/2 \in W$.

In particular, this property is satisfied if W is a convex set containing function 1.

Lower bound

Assume that a class of real functions $W \subset \mathcal{C}(\Omega)$ has the following extra property.

Property A. For any $f \in W$ we have $f^+ := (f+1)/2 \in W$ and $f^- := (f-1)/2 \in W$.

In particular, this property is satisfied if W is a convex set containing function 1.

For a function class $W \subset \mathcal{C}(\Omega)$ consider the best error of numerical integration by cubature formulas with m knots:

$$\kappa_m(W) := \inf_{(\xi,\Lambda)} \sup_{f\in W} |I_\mu(f) - \Lambda_m(f,\xi)|,$$

$$I_\mu(f) := \int_\Omega f d\mu, \quad \Lambda_m(f,\xi) := \sum_{j=1}^m \lambda_j f(\xi^j).$$

Connection to numerical integration

Theorem (T10; VT, 2018)

Suppose $W \subset \mathcal{C}(\Omega)$ *has Property A. Then for any* $m \in \mathbb{N}$ *we have*

$$er^o_m(W, L_2) \geq \frac{1}{2}\kappa_m(W).$$

The End

Thank you!