Sampling recovery. Lecture 2. Recovery and discretization.

Vladimir Temlyakov

May, 2021

For a fixed m and a set of points $\xi := \{\xi^j\}_{j=1}^m \subset \Omega$, let Φ_ξ be a linear operator from $\mathbb{C}^m$ into $L_p(\Omega,\mu)$. Denote for a class $\mathbf{F}$ (usually, centrally symmetric and compact subset of $L_p(\Omega,\mu)$)

$$\varrho_m(\mathbf{F}, L_p) := \inf_{\text{linear}\,\Phi_\xi;\,\xi} \sup_{f\in\mathbf{F}} \|f - \Phi_\xi(f(\xi^1),\dots,f(\xi^m))\|_p.$$

The above described recovery procedure is a linear procedure.

Optimal recovery

For a fixed m and a set of points $\xi := \{\xi^j\}_{j=1}^m \subset \Omega$, let Φ_ξ be a linear operator from $\mathbb{C}^m$ into $L_p(\Omega,\mu)$. Denote for a class $\mathbf{F}$ (usually, centrally symmetric and compact subset of $L_p(\Omega,\mu)$)

$$\varrho_m(\mathbf{F}, L_p) := \inf_{\text{linear}\,\Phi_\xi;\,\xi} \sup_{f\in\mathbf{F}} \|f - \Phi_\xi(f(\xi^1),\dots,f(\xi^m))\|_p.$$

The above described recovery procedure is a linear procedure. The following modification of the above recovery procedure is also of interest. We now allow any mapping $\Phi_\xi : \mathbb{C}^m \to X_N \subset L_p(\Omega,\mu)$ where X_N is a linear subspace of dimension $N \le m$ and define

$$\varrho_m^*(\mathbf{F}, L_p) := \inf_{\Phi_\xi;\xi;X_N,N\le m} \sup_{f\in\mathbf{F}} \|f - \Phi(f(\xi^1),\dots,f(\xi^m))\|_p.$$

In both of the above cases we build an approximant, which comes from a linear subspace of dimension at most m.

Kolmogorov width

It is natural to compare quantities $\varrho_m(\mathbf{F}, L_p)$ and $\varrho_m^*(\mathbf{F}, L_p)$ with the Kolmogorov widths. Let $\mathbf{F} \subset L_p$ be a centrally symmetric compact. The quantities

$$d_n(\mathbf{F}, L_p) := \inf_{\{u_i\}_{i=1}^n \subset L_p} \sup_{f \in \mathbf{F}} \inf_{c_i} \left\| f - \sum_{i=1}^n c_i u_i \right\|_p, \quad n = 1, 2, \dots,$$

are called the Kolmogorov widths of $\mathbf{F}$ in L_p.

Kolmogorov width

It is natural to compare quantities $\varrho_m(\mathbf{F}, L_p)$ and $\varrho_m^*(\mathbf{F}, L_p)$ with the Kolmogorov widths. Let $\mathbf{F} \subset L_p$ be a centrally symmetric compact. The quantities

$$d_n(\mathbf{F}, L_p) := \inf_{\{u_i\}_{i=1}^n \subset L_p} \sup_{f \in \mathbf{F}} \inf_{c_i} \left\| f - \sum_{i=1}^n c_i u_i \right\|_p, \quad n = 1, 2, \dots,$$

are called the Kolmogorov widths of $\mathbf{F}$ in L_p.
We have the following obvious inequalities

$$d_m(\mathbf{F}, L_p) \le \varrho_m^*(\mathbf{F}, L_p) \le \varrho_m(\mathbf{F}, L_p).$$

Kolmogorov width

It is natural to compare quantities $\varrho_m(\mathbf{F}, L_p)$ and $\varrho^*_m(\mathbf{F}, L_p)$ with the Kolmogorov widths. Let $\mathbf{F} \subset L_p$ be a centrally symmetric compact. The quantities

$$d_n(\mathbf{F}, L_p) := \inf_{\{u_i\}_{i=1}^n \subset L_p} \sup_{f \in \mathbf{F}} \inf_{c_i} \left\| f - \sum_{i=1}^n c_i u_i \right\|_p, \quad n = 1, 2, \dots,$$

are called the Kolmogorov widths of $\mathbf{F}$ in L_p.
We have the following obvious inequalities

$$d_m(\mathbf{F}, L_p) \le \varrho^*_m(\mathbf{F}, L_p) \le \varrho_m(\mathbf{F}, L_p).$$

We consider the case $p = 2$, i.e. recovery takes place in the Hilbert space L_2.

Inequality

Theorem (VT, 2020)

There exist two positive absolute constants b and B such that for any compact subset Ω of $\mathbb{R}^d$, any probability measure μ on it, and any compact subset $\mathbf{F}$ of $\mathcal{C}(\Omega)$ we have

$$\varrho_{bn}(\mathbf{F}, L_2(\Omega,\mu)) \le B d_n(\mathbf{F}, L_\infty).$$

Theorem (VT, 2020)

There exist two positive absolute constants b and B such that for any compact subset Ω of $\mathbb{R}^d$, any probability measure μ on it, and any compact subset $\mathbf{F}$ of $\mathcal{C}(\Omega)$ we have

$$\varrho_{bn}(\mathbf{F}, L_2(\Omega,\mu)) \le B d_n(\mathbf{F}, L_\infty).$$

Note that for special sets $\mathbf{F}$ (in the reproducing kernel Hilbert space setting) the following inequality is known (see N. Nagel, M. Schäfer, T. Ullrich and D. Krieg, M. Ullrich, 2020)

$$\varrho_n(\mathbf{F}, L_2) \le C \left(\frac{\log n}{n} \sum_{k\ge cn} d_k(\mathbf{F}, L_2)^2 \right)^{1/2}$$

with absolute constants $C, c > 0$.

Numerical integration inequality

We now formulate a known result, which relates other discretization characteristic of a class with its Kolmogorov's width. For a compact subset $\mathbf{F} \subset \mathcal{C}(\Omega)$ define the best error of numerical integration with m knots as follows

$$\kappa_m(\mathbf{F}) := \inf_{\xi^1,\dots,\xi^m;\lambda_1,\dots,\lambda_m} \sup_{f\in\mathbf{F}} \left| \int_\Omega f d\mu - \sum_{j=1}^{m} \lambda_j f(\xi^j) \right|.$$

The following inequality was proved in E. Novak, 1986

$$\kappa_m(\mathbf{F}) \le 2d_m(\mathbf{F}, L_\infty).$$

Let us make a simple well known observation on a relation between recovery and numerical integration. Associate with the recovery operator

$$\Psi(f,\xi) := \sum_{j=1}^{m} f(\xi^j)\psi_j(\mathbf{x})$$

the cubature formula

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

Then

$$\left|\Lambda_m(f,\xi) - \int_\Omega f(\mathbf{x})d\mu\right| = \left|\int_\Omega (\Psi(f,\xi) - f)d\mu\right|$$
$$\leq \|\Psi(f,\xi) - f\|_1 \leq \|\Psi(f,\xi) - f\|_p, \qquad p \geq 1.$$

Therefore, for any function class $\mathbf{F}$ and each $p \geq 1$ we have

$$\kappa_m(\mathbf{F}) \leq \varrho_m(\mathbf{F}, L_p). \tag{1}$$

Theorem (VT, 2020) is proved with a help of a classical type of algorithms – weighted least squares algorithms. Let X_N be an N-dimensional subspace of the space of continuous functions $\mathcal{C}(\Omega)$ and let $\mathbf{w} := (w_1, \dots, w_m) \in \mathbb{R}^m$ be a positive weight, i.e. $w_i > 0$, $i = 1, \dots, m$. Consider the following classical weighted least squares recovery operator (algorithm)

$$\ell 2\mathbf{w}(\xi, X_N)(f) := \arg\min_{u \in X_N} \|S(f-u,\xi)\|_{2,\mathbf{w}}, \quad \xi = \{\xi^j\}_{j=1}^m \subset \Omega,$$

$$\|S(g,\xi)\|_{2,\mathbf{w}} := \left(\sum_{\nu=1}^m w_\nu |g(\xi^\nu)|^2\right)^{1/2}.$$

Theorem (VT, 2020) is proved with a help of a classical type of algorithms – weighted least squares algorithms. Let X_N be an N-dimensional subspace of the space of continuous functions $\mathcal{C}(\Omega)$ and let $\mathbf{w} := (w_1, \dots, w_m) \in \mathbb{R}^m$ be a positive weight, i.e. $w_i > 0$, $i = 1, \dots, m$. Consider the following classical weighted least squares recovery operator (algorithm)

$$\ell 2\mathbf{w}(\xi, X_N)(f) := \arg\min_{u \in X_N} \|S(f-u,\xi)\|_{2,\mathbf{w}}, \quad \xi = \{\xi^j\}_{j=1}^m \subset \Omega,$$

$$\|S(g,\xi)\|_{2,\mathbf{w}} := \left(\sum_{\nu=1}^m w_\nu |g(\xi^\nu)|^2\right)^{1/2}.$$

We note that the least squares operator (empirical risk minimization) is a classical algorithm from statistics and learning theory. It is also well known in learning theory that performance of this algorithm can be controlled by asymptotic characteristics measured in the uniform norm.

For a fixed m denote for a class $\mathbf{F}$

$$\varrho_m^{wls}(\mathbf{F}, L_2) := \inf_{\mathbf{w},\, \xi\subset\Omega,\, X_N} \sup_{f\in\mathbf{F}} \|f - \ell 2\mathbf{w}(\xi, X_N)(f)\|_2.$$

One more inequality

For a fixed m denote for a class $\mathbf{F}$

$$\varrho_m^{wls}(\mathbf{F}, L_2) := \inf_{\mathbf{w},\, \xi\subset\Omega,\, X_N} \sup_{f\in\mathbf{F}} \|f - \ell 2\mathbf{w}(\xi, X_N)(f)\|_2.$$

In VT, 2020 we proved the following result, which implies Theorem (VT, 2020).

Theorem (VT, 2020, LS)

There exist two positive absolute constants b and B such that for any compact subset Ω of $\mathbb{R}^d$, any probability measure μ on it, and any compact subset $\mathbf{F}$ of $\mathcal{C}(\Omega)$ we have

$$\varrho_{bn}^{wls}(\mathbf{F}, L_2(\Omega, \mu)) \le B d_n(\mathbf{F}, L_\infty).$$

Notations 1

Let X_N be an N-dimensional subspace of the space of continuous functions $\mathcal{C}(\Omega)$. For a fixed m and a set of points $\xi := \{\xi^\nu\}_{\nu=1}^m \subset \Omega$ we associate with a function $f \in \mathcal{C}(\Omega)$ a vector

$$S(f,\xi) := (f(\xi^1),\dots,f(\xi^m)) \in \mathbb{C}^m.$$

Denote

$$\|S(f,\xi)\|_p := \left(\frac{1}{m}\sum_{\nu=1}^m |f(\xi^\nu)|^p\right)^{1/p}, \quad 1 \le p < \infty,$$

and

$$\|S(f,\xi)\|_\infty := \max_\nu |f(\xi^\nu)|.$$

Notations 2

For a positive weight $\mathbf{w} := (w_1, \dots, w_m) \in \mathbb{R}^m$ consider the following norm

$$\|S(f,\xi)\|_{p,\mathbf{w}} := \left(\sum_{\nu=1}^{m} w_\nu |f(\xi^\nu)|^p\right)^{1/p}, \quad 1 \le p < \infty.$$

Define the best approximation of $f \in L_p(\Omega, \mu)$, $1 \le p \le \infty$ by elements of X_N as follows

$$d(f, X_N)_p := \inf_{u \in X_N} \|f - u\|_p.$$

It is well known that there exists an element, which we denote $P_{X_N,p}(f) \in X_N$, such that

$$\|f - P_{X_N,p}(f)\|_p = d(f, X_N)_p.$$

The operator $P_{X_N,p} : L_p(\Omega, \mu) \to X_N$ is called the Chebyshev projection.

A1. Discretization. Let $1 \le p \le \infty$. Suppose that $\xi := \{\xi^j\}_{j=1}^m \subset \Omega$ is such that for any $u \in X_N$ in the case $1 \le p < \infty$ we have

$$C_1\|u\|_p \le \|S(u,\xi)\|_{p,\mathbf{w}}$$

and in the case $p = \infty$ we have

$$C_1\|u\|_\infty \le \|S(u,\xi)\|_\infty$$

with a positive constant C_1 which may depend on d and p.

Assumptions

A1. Discretization. Let $1 \le p \le \infty$. Suppose that $\xi := \{\xi^j\}_{j=1}^m \subset \Omega$ is such that for any $u \in X_N$ in the case $1 \le p < \infty$ we have

$$C_1\|u\|_p \le \|S(u,\xi)\|_{p,\mathbf{w}}$$

and in the case $p = \infty$ we have

$$C_1\|u\|_\infty \le \|S(u,\xi)\|_\infty$$

with a positive constant C_1 which may depend on d and p.
A2. Weights. Suppose that there is a positive constant $C_2 = C_2(d,p)$ such that $\sum_{\nu=1}^m w_\nu \le C_2$.

Algorithm

Consider the following well known recovery operator (algorithm)

$$\ell p\mathbf{w}(\xi)(f) := \ell p\mathbf{w}(\xi, X_N)(f) := \arg\min_{u\in X_N} \|S(f-u,\xi)\|_{p,\mathbf{w}}.$$

Consider the following well known recovery operator (algorithm)

$$\ell p\mathbf{w}(\xi)(f) := \ell p\mathbf{w}(\xi, X_N)(f) := \arg\min_{u\in X_N} \|S(f-u,\xi)\|_{p,\mathbf{w}}.$$

Note that the above algorithm $\ell p\mathbf{w}(\xi)$ only uses the function values $f(\xi^\nu)$, $\nu = 1,\dots,m$. In the case $p=2$ it is a linear algorithm – orthogonal projection with respect to the norm $\|\cdot\|_{2,\mathbf{w}}$. Therefore, in the case $p=2$ approximation error by the algorithm $\ell 2\mathbf{w}(\xi)$ gives an upper bound for the recovery characteristic $\varrho_m(\cdot, L_2)$. In the case $p\neq 2$ approximation error by the algorithm $\ell p\mathbf{w}(\xi)$ gives an upper bound for the recovery characteristic $\varrho^*_m(\cdot, L_p)$.

Approximation inequality

Theorem (VT, 2020, AI)

Under assumptions **A1** *and* **A2** *for any* $f \in \mathcal{C}(\Omega)$ *we have for* $1 \le p < \infty$

$$\|f - \ell p\mathbf{w}(\xi)(f)\|_p \le (2C_1^{-1}C_2^{1/p} + 1)d(f, X_N)_\infty.$$

Under assumption **A1** *for* $p = \infty$ *for any* $f \in \mathcal{C}(\Omega)$ *we have*

$$\|f - \ell\infty(\xi)(f)\|_\infty \le (2C_1^{-1} + 1)d(f, X_N)_\infty.$$

Proof. We give a detailed proof for the case $1 \le p < \infty$. The case $p = \infty$ is similar and even simpler. From the definition of the operator $P_{X_N,\infty}$ we obtain

$$\|f - P_{X_N,\infty}(f)\|_p \le \|f - P_{X_N,\infty}(f)\|_\infty = d(f, X_N)_\infty. \qquad (2)$$

Proof of the approximation inequality 1

Proof. We give a detailed proof for the case $1 \le p < \infty$. The case $p = \infty$ is similar and even simpler. From the definition of the operator $P_{X_N,\infty}$ we obtain

$$\|f - P_{X_N,\infty}(f)\|_p \le \|f - P_{X_N,\infty}(f)\|_\infty = d(f, X_N)_\infty. \tag{2}$$

Clearly,

$$\|S(f - P_{X_N,\infty}(f), \xi)\|_\infty \le \|f - P_{X_N,\infty}(f)\|_\infty = d(f, X_N)_\infty.$$

Proof of the approximation inequality 2

Therefore, by **A2** we get

$$\|S(f - P_{X_N,\infty}(f), \xi)\|_{p,\mathbf{w}} \le C_2^{1/p} \|S(f - P_{X_N,\infty}(f), \xi)\|_\infty$$
$$\le C_2^{1/p} d(f, X_N)_\infty. \quad (3)$$

Therefore, by **A2** we get

$$\|S(f - P_{X_N,\infty}(f), \xi)\|_{p,\mathbf{w}} \le C_2^{1/p}\|S(f - P_{X_N,\infty}(f), \xi)\|_\infty$$

$$\le C_2^{1/p} d(f, X_N)_\infty. \quad (3)$$

Next, by the definition of the algorithm $\ell p\mathbf{w}(\xi)$ and by **A2** we obtain

$$\|S(f - \ell p\mathbf{w}(\xi)(f), \xi)\|_{p,\mathbf{w}} \le \|S(f - P_{X_N,\infty}(f), \xi)\|_{p,\mathbf{w}}$$

$$\le C_2^{1/p} d(f, X_N)_\infty. \quad (4)$$

Proof of the approximation inequality 3

Bounds (3) and (4) imply

$$\|S(P_{X_N,\infty}(f) - \ell p\mathbf{w}(\xi)(f), \xi)\|_{p,\mathbf{w}} \le 2C_2^{1/p} d(f, X_N)_\infty. \qquad (5)$$

Proof of the approximation inequality 3

Bounds (3) and (4) imply

$$\|S(P_{X_N,\infty}(f) - \ell p\mathbf{w}(\xi)(f), \xi)\|_{p,\mathbf{w}} \le 2C_2^{1/p} d(f, X_N)_\infty. \qquad (5)$$

Then, the discretization assumption **A1** implies

$$\|P_{X_N,\infty}(f) - \ell p\mathbf{w}(\xi)(f)\|_p \le C_1^{-1} 2C_2^{1/p} d(f, X_N)_\infty. \qquad (6)$$

Proof of the approximation inequality 3

Bounds (3) and (4) imply

$$\|S(P_{X_N,\infty}(f) - \ell p\mathbf{w}(\xi)(f),\xi)\|_{p,\mathbf{w}} \le 2C_2^{1/p} d(f,X_N)_\infty. \qquad (5)$$

Then, the discretization assumption **A1** implies

$$\|P_{X_N,\infty}(f) - \ell p\mathbf{w}(\xi)(f)\|_p \le C_1^{-1} 2C_2^{1/p} d(f,X_N)_\infty. \qquad (6)$$

Combining bounds (2) and (6) we conclude

$$\|f - \ell p\mathbf{w}(\xi)(f)\|_p \le (1 + 2C_1^{-1}C_2^{1/p}) d(f,X_N)_\infty,$$

which completes the proof of Theorem (VT, 2020, AI).

The main point is the use of a result on discretization in L_2 from I. Limonova and VT, 2020, which is a generalization to the complex case of an earlier result from F. Dai, A. Prymak, A. Shadrin, V. Temlyakov, and S. Tikhonov, 2020, established for the real case.

Theorem (LT, 2020)

If X_N is an N-dimensional subspace of the complex $L_2(\Omega,\mu)$, then there exist three absolute positive constants C_1', c_0', C_0', a set of $m\le C_1'N$ points $\xi^1,\dots,\xi^m\in\Omega$, and a set of nonnegative weights λ_j, $j=1,\dots,m$, such that

$$c_0'\|f\|_2^2\le \sum_{j=1}^m \lambda_j|f(\xi^j)|^2\le C_0'\|f\|_2^2,\quad \forall f\in X_N.$$

For our application we need to satisfy the assumption **A2** on weights.

Remark

Considering a new subspace $X'_N := \{f : f = g + c,\ g \in X_N,\ c \in \mathbb{C}\}$ *and applying Theorem (LT, 2020) to the* X'_N *with* $f = 1$ $(g = 0,\ c = 1)$ *we conclude that a version of Theorem (LT, 2020) holds with* $m \le C'_1 N$ *replaced by* $m \le C'_1(N+1)$ *and with weights satisfying*

$$\sum_{j=1}^{m} \lambda_j \le C'_0.$$

Proof continues

Let X_n be a subspace of dimension n satisfying: for all $f \in \mathbf{F}$

$$d(f, X_n)_\infty \le 2d_n(\mathbf{F}, L_\infty).$$

Let now $\xi = \{\xi^\nu\}_{\nu=1}^m$ be the set of points from Theorem (LT, 2020) and Remark with $X_N = X_n$. Then $m \le bn$ and assumptions **A1** and **A2** are satisfied with absolute constants C_i, $i = 1, 2$. Applying Theorem (VT, 2020, AI) we complete the proof of Theorem (VT, 2020, LS).

1. Recovery of $\mathbf{W}^r_2$. The following upper bound is known (see Trigub and Belinskii, 2004)

$$d_n(\mathbf{W}^r_2, L_\infty) \le C(r,d)n^{-r}(\log n)^{(d-1)r+1/2}, \quad r > 1/2.$$

By Theorems (VT, 2020) and (VT, 2020, LS) we obtain from the above bound the estimate for $r > 1/2$

$$\varrho_n(\mathbf{W}^r_2, L_2) \le \varrho_n^{wls}(\mathbf{W}^r_2, L_2) \le C'(r,d)n^{-r}(\log n)^{(d-1)r+1/2}.$$

Comment

Very recently the above bound for the $\varrho_n(\mathbf{W}^r_2, L_2)$ was obtained in N. Nagel, M. Schäfer, T. Ullrich, 2020. This is the best known upper bound. For the previous breakthrough result see D. Krieg, M. Ullrich, 2020. Thus, we demonstrate here that Theorem (VT, 2020) is a rather powerful tool in estimation of the recovery numbers. The right order of the quantity $\varrho_n(\mathbf{W}^r_2, L_2)$ is not known. Also, the above bound shows that the best known upper bound for the $\varrho_n(\mathbf{W}^r_2, L_2)$ can be obtained by the weighted least squares algorithm.

2. Recovery of $\mathbf{W}^r_1$. We define the best n-term approximation with respect to the trigonometric system $\mathcal{T}^d := \{e^{i(\mathbf{k},\mathbf{x})}\}_{\mathbf{k}\in\mathbb{Z}^d}$ as follows

$$\sigma_n(f)_p := \inf_{\mathbf{k}^1,\dots,\mathbf{k}^n} \inf_{c_1,\dots,c_n} \left\| f - \sum_{j=1}^{n} c_j e^{i(\mathbf{k}^j,\mathbf{x})} \right\|_p .$$

The following result is known (see, for instance, VT book MA, 2018, p.466)

$$\sigma_n(F_r)_\infty \le C(r,d) n^{-r+1/2}(\log n)^{r(d-1)+1/2}, \quad r>1. \tag{7}$$

$\mathbf{W}^r_1$ classes continue

Bound (7) and the definition of the class $\mathbf{W}^r_1$ imply that there exists a set of frequencies $\Lambda_n = \{\mathbf{k}^j\}_{j=1}^n$ such that for any $f \in \mathbf{W}^r_1$ we have

$$d(f, \mathcal{T}(\Lambda_n))_\infty \le C(r,d) n^{-r+1/2} (\log n)^{r(d-1)+1/2}, \qquad (8)$$

where we use the notation

$$\mathcal{T}(\Lambda) := \{f \,:\, f = \sum_{\mathbf{k}\in\Lambda} c_{\mathbf{k}} e^{i(\mathbf{k},\mathbf{x})}\}.$$

Discretization result

Theorem (VT, 2017)

There are three positive absolute constants C_1, C_2, *and* C_3 *with the following properties: For any* $d \in \mathbb{N}$ *and any* $Q \subset \mathbb{Z}^d$ *there exists a set of* $m \le C_1|Q|$ *points* $\xi^j \in \mathbb{T}^d$, $j = 1, \dots, m$, *such that for any* $f \in \mathcal{T}(Q)$ *we have*

$$C_2\|f\|_2^2 \le \frac{1}{m}\sum_{j=1}^{m}|f(\xi^j)|^2 \le C_3\|f\|_2^2.$$

Discretization result

Theorem (VT, 2017)

There are three positive absolute constants C_1, C_2, and C_3 with the following properties: For any $d \in \mathbb{N}$ and any $Q \subset \mathbb{Z}^d$ there exists a set of $m \le C_1|Q|$ points $\xi^j \in \mathbb{T}^d$, $j = 1, \dots, m$, such that for any $f \in \mathcal{T}(Q)$ we have

$$C_2\|f\|_2^2 \le \frac{1}{m}\sum_{j=1}^{m} |f(\xi^j)|^2 \le C_3\|f\|_2^2.$$

Therefore, conditions **A1** and **A2** are satisfied for the $\mathcal{T}(\Lambda_n)$ and by Theorem (VT, 2020, AI) we obtain from (8)

$$\varrho_n(\mathbf{W}^r_1, L_2) \le C(r,d)n^{-r+1/2}(\log n)^{r(d-1)+1/2}, \quad r > 1. \tag{9}$$

Discretization result

Theorem (VT, 2017)

There are three positive absolute constants C_1, C_2, and C_3 with the following properties: For any $d \in \mathbb{N}$ and any $Q \subset \mathbb{Z}^d$ there exists a set of $m \leq C_1|Q|$ points $\xi^j \in \mathbb{T}^d$, $j = 1, \dots, m$, such that for any $f \in \mathcal{T}(Q)$ we have

$$C_2\|f\|_2^2 \leq \frac{1}{m}\sum_{j=1}^{m}|f(\xi^j)|^2 \leq C_3\|f\|_2^2.$$

Therefore, conditions **A1** and **A2** are satisfied for the $\mathcal{T}(\Lambda_n)$ and by Theorem (VT, 2020, AI) we obtain from (8)

$$\varrho_n(\mathbf{W}^r_1, L_2) \leq C(r,d)n^{-r+1/2}(\log n)^{r(d-1)+1/2}, \quad r > 1. \tag{9}$$

Moreover, as a recovering algorithm we can take the $\ell 2\mathbf{w}_n(\xi)$ with $\mathbf{w}_n = (1/n, \dots, 1/n)$, which is a standard least squares algorithm.

A sharpness comment

It is known that bound (9) is not sharp. The following relation holds

$$\varrho_n(\mathbf{W}^r_1, L_2) \asymp n^{-r+1/2}(\log n)^{r(d-1)}, \quad r > 1. \tag{10}$$

The lower bound in (10) follows from the known lower bound for $d_n(\mathbf{W}^r_1, L_2)$ (VT, 1982). The matching upper bound follows from the upper bound for recovery by Smolyak's algorithm (sparse grids method) obtained in VT, 1993. Smolyak's algorithm is a simple direct method, which gives an explicit formula for an approximant. It is known that this algorithm is optimal in the sense of order in some cases of recovery functions with mixed smoothness.

Recovery of $\mathbf{H}^r_2$. The following bound for the Kolmogorov width is known (see E. Belinskii, 1991)

$$d_n(\mathbf{H}^r_2, L_\infty) \le C(r,d) n^{-r} (\log n)^{(d-1)(r+1/2)+1/2}, \quad r > 1/2. \quad (11)$$

We obtain from bound (11) by Theorems (VT, 2020) and (VT, 2020, LS) the estimate for $r > 1/2$

$$\varrho_n(\mathbf{H}^r_2, L_2) \le \varrho^{wls}_n(\mathbf{H}^r_2, L_2) \le C'(r,d) n^{-r} (\log n)^{(d-1)(r+1/2)+1/2}.$$

A historical comment

Let us make a brief historical comment on optimal recovery of classes $\mathbf{H}^r_p$ in L_p. For more detailed discussion we refer the reader to Ding Dũng, V.N. Temlyakov, and T. Ullrich, 2018 and VT book MA, 2018. The first result in this direction was established in VT, 1985: for $r > 1/p$ and $1 \le p \le \infty$ we have

$$\varrho_n(\mathbf{H}^r_p, L_p) \le C(r,d,p) n^{-r} (\log n)^{(d-1)(r+1)}.$$

A historical comment

Let us make a brief historical comment on optimal recovery of classes $\mathbf{H}^r_p$ in L_p. For more detailed discussion we refer the reader to Ding Dũng, V.N. Temlyakov, and T. Ullrich, 2018 and VT book MA, 2018. The first result in this direction was established in VT, 1985: for $r > 1/p$ and $1 \le p \le \infty$ we have

$$\varrho_n(\mathbf{H}^r_p, L_p) \le C(r,d,p) n^{-r} (\log n)^{(d-1)(r+1)}.$$

We note that the problem of the right asymptotic behavior of $\varrho_n(\mathbf{H}^r_p, L_p)$, $1 \le p \le \infty$, is a great open problem. As far as we know it is only solved in the case $d = 2$, $p = \infty$, (see, for instance, VT book MA, 2018, p.308):

$$\varrho_n(\mathbf{H}^r_\infty, L_\infty) \asymp n^{-r} (\log n)^{r+1}.$$

Some further comments

We discussed recovery by weighted least squares algorithms $\ell 2\mathbf{w}(\xi)$. We may want to have the recovery algorithm $\ell 2\mathbf{w}(\xi)$ to be a classical least squares algorithm, i.e. $\mathbf{w} = \mathbf{w}_m := (1/m, \dots, 1/m)$. For that we need an analog of the discretization Theorem (LT, 2020) with the weight $\mathbf{w}_m$. There is such an analog of Theorem (LT, 2020) but under an extra assumption on the subspace X_N.

Some further comments

We discussed recovery by weighted least squares algorithms $\ell 2\mathbf{w}(\xi)$. We may want to have the recovery algorithm $\ell 2\mathbf{w}(\xi)$ to be a classical least squares algorithm, i.e. $\mathbf{w} = \mathbf{w}_m := (1/m, \dots, 1/m)$. For that we need an analog of the discretization Theorem (LT, 2020) with the weight $\mathbf{w}_m$. There is such an analog of Theorem (LT, 2020) but under an extra assumption on the subspace X_N.

Condition E(t). We say that an orthonormal system $\{u_i(\mathbf{x})\}_{i=1}^N$ defined on Ω satisfies Condition E(t) with a constant t if for all $\mathbf{x} \in \Omega$

$$\sum_{i=1}^{N} |u_i(\mathbf{x})|^2 \leq N t^2.$$

Discretization result

Theorem (LT, 2020, EW)

Let $\Omega \subset \mathbb{R}^d$ be a compact set with the probability measure μ. Assume that $\{u_i(\mathbf{x})\}_{i=1}^N$ is a real (or complex) orthonormal system in $L_2(\Omega,\mu)$ satisfying Condition $E(t)$. Then there is an absolute constant C_1 such that there exists a set $\{\xi^j\}_{j=1}^m \subset \Omega$ of $m \le C_1 t^2 N$ points with the property: For any $f = \sum_{i=1}^N c_i u_i$ we have

$$C_2\|f\|_2^2 \le \frac{1}{m}\sum_{j=1}^m |f(\xi^j)|^2 \le C_3 t^2 \|f\|_2^2,$$

where C_2 and C_3 are absolute positive constants.

A variant of the Kolmogorov width

For a fixed m and a set of points $\xi := \{\xi^j\}_{j=1}^m \subset \Omega$ denote for a class $\mathbf{F}$ (usually, a centrally symmetric compact in $L_2(\Omega,\mu)$)

$$\varrho_m^{ls}(\mathbf{F}, L_2) := \inf_{\xi, X_N} \sup_{f\in\mathbf{F}} \|f - \ell 2\mathbf{w}_m(\xi, X_N)(f)\|_2.$$

We now define $E(t)$-conditioned Kolmogorov width

$$d_N^{E(t)}(\mathbf{F}, L_p) := \inf_{\{u_1,\dots,u_N\} \text{ satisfies Condition} E(t)} \sup_{f\in\mathbf{F}} \inf_{c_1,\dots,c_N} \|f - \sum_{i=1}^N c_i u_i\|_p.$$

Theorem (LT, 2020, EW) combined with Theorem (VT, 2020, AI) gives the following analog of Theorem (VT, 2020, LS).

Theorem (VT, 2020, EW)

Let $\mathbf{F}$ be a compact subset of $\mathcal{C}(\Omega)$. There exist two positive constants b and B which may depend on t such that

$$\varrho_{bn}^{ls}(\mathbf{F}, L_2) \leq B d_n^{E(t)}(\mathbf{F}, L_\infty).$$

Comments on techniques

The discretization Theorems (LT, 2020) and (LT, 2020, EW) play a key role in the proof of the recovery inequalities. Proofs of these theorems are based on fundamental results from J. Bourgain, J. Lindenstrauss, and V. Milman, 1989, J. Batson, D.A. Spielman, and N. Srivastava, 2014, and A. Marcus, D.A. Spielman, and N. Srivastava, 2015. In particular, they are based on the following result by A. Marcus, D.A. Spielman, and N. Srivastava, 2015, which solved the famous Kadison-Singer problem.

MSS theorem

Theorem (MSS, 2015)

Let a system of vectors $\mathbf{v}_1, \dots, \mathbf{v}_M$ *from* $\mathbb{C}^N$ *have the following properties: for all* $\mathbf{w} \in \mathbb{C}^N$

$$\sum_{j=1}^{M} |\langle \mathbf{w}, \mathbf{v}_j \rangle|^2 = \|\mathbf{w}\|_2^2 \tag{12}$$

and for some $\epsilon > 0$ *we have* $\|\mathbf{v}_j\|_2^2 \le \epsilon$, $j = 1, \dots, M$. *Then there is a partition of* $\{1, 2, \dots, M\}$ *into two sets* S_1 *and* S_2 *such that for all* $\mathbf{w} \in \mathbb{C}^N$ *and for each* $i = 1, 2$

$$\sum_{j \in S_i} |\langle \mathbf{w}, \mathbf{v}_j \rangle|^2 \le \frac{(1 + \sqrt{2\epsilon})^2}{2} \|\mathbf{w}\|_2^2.$$

NOU lemma

The following Lemma (NOU, 2016) was derived from Theorem (MSS, 2015) in S. Nitzan, A. Olevskii, and A. Ulanovskii, 2016.

Lemma (NOU, 2016)

Let a system of vectors $\mathbf{v}_1, \dots, \mathbf{v}_M$ *from* $\mathbb{C}^N$ *satisfy* (12) *for all* $\mathbf{w} \in \mathbb{C}^N$ *and*

$$\|\mathbf{v}_j\|_2^2 = N/M, \qquad j = 1, \dots, M.$$

Then there is a subset $J \subset \{1, 2, \dots, M\}$ *such that for all* $\mathbf{w} \in \mathbb{C}^N$

$$c_0\|\mathbf{w}\|_2^2 \le \frac{M}{N} \sum_{j \in J} |\langle \mathbf{w}, \mathbf{v}_j \rangle|^2 \le C_0\|\mathbf{w}\|_2^2,$$

where c_0 *and* C_0 *are some absolute positive constants.*

LT lemma

The following generalization of Lemma (NOU, 2016) was proved in LT, 2020.

Lemma (LT, 2020)

Let a system of vectors $\mathbf{v}_1, \dots, \mathbf{v}_M$ *from* $\mathbb{C}^N$ *satisfy* (12) *for all* $\mathbf{w} \in \mathbb{C}^N$ *and*

$$\|\mathbf{v}_j\|_2^2 \le \theta N/M, \qquad \theta \le M/N, \qquad j = 1, \dots, M. \tag{13}$$

Then there is a subset $J \subset \{1, 2, \dots, M\}$ *such that for all* $\mathbf{w} \in \mathbb{C}^N$

$$c_0\theta\|\mathbf{w}\|_2^2 \le \frac{M}{N}\sum_{j\in J}|\langle \mathbf{w}, \mathbf{v}_j\rangle|^2 \le C_0\theta\|\mathbf{w}\|_2^2, \quad |J| \le C_1\theta N, \tag{14}$$

where c_0, C_0, *and* C_1 *are some absolute positive constants.*

We note that the proof of Lemma (LT, 2020) in LT, 2020 gives a little stronger result than Lemma (LT, 2020) – the tight frame condition (12) can be replaced by a frame condition

$$A\|\mathbf{w}\|_2^2 \le \sum_{j=1}^{M} |\langle \mathbf{w}, \mathbf{v}_j \rangle|^2 \le B\|\mathbf{w}\|_2^2.$$

This stronger version of Lemma (LT, 2020) was used in the followup paper N. Nagel, M. Schäfer, T. Ullrich, 2020, for sampling recovery.

A historical remark

Currently, it is well understood that Theorem (MSS, 2015), Lemma (NOU, 2016), Lemma (LT, 2020), and their further developments are very important in the sampling discretization and sampling recovery. As far as the author knows, the first application of Lemma (NOU, 2016) for discretization of the L_2 norms of functions in subspaces of the multivariate trigonometric polynomials was done in VT, 2017. B. Kashin brought Lemma (NOU, 2016) to the author's attention. Lemma(NOU, 2016) and its generalization Lemma (LT, 2020) allow us to prove discretization results with equal weights in a style of Theorem (LT, 2020, EW). This approach only works for subspaces satisfying extra conditions, for example, Condition E(t).

However, it is not enough for proving Theorem (LT, 2020) because we do not know if there is a subspace with this property, which provides approximation of the order of the corresponding Kolmogorov's width. It turns out that we can treat discretization of the L_2 norm for arbitrary subspaces if we allow the use of appropriate weights instead of the equal weights. Results from J. Batson, D.A. Spielman, and N. Srivastava, 2014, allow us to do that. The first application of the results from J. Batson, D.A. Spielman, and N. Srivastava, 2014, in discretization of the L_2 norm was done in VT, 2017. The other way to go from equal weights to appropriate general weights and to treat general subspaces uses the technique known as change of measure (see, for instance, J. Bourgain, J. Lindenstrauss, and V. Milman, 1989). This technique was applied to discretization of the L_2 norm in F. Dai, A. Prymak, A. Shadrin, V. Temlyakov, and S. Tikhonov, 2020.

Thank you!