# Harmonic Analysis and approximation by ridge functions

Vitaly Maiorov

Technion

August 30, 2015

# The ridge function

- Ridge function (plane wave) is a multivariate function $R: \ \mathbb{R}^d \to \mathbb{R}$ of the form

$$R(x) = r(a \cdot x) = r(a_1 x_d + \cdots + a_d x_d),$$

# The ridge function

- Ridge function (plane wave) is a multivariate function $R: \mathbb{R}^d \to \mathbb{R}$ of the form

$$R(x) = r(a \cdot x) = r(a_1 x_d + \cdots + a_d x_d),$$

- where $r: \mathbb{R} \to \mathbb{R}$ is an univariate function, and $a = (a_1, ..., a_d)$ is a vector from $\mathbb{R}^d$.

# Examples

- Ridge functions appear in the Fourier transform

$$(\mathcal{F}f)(w) = \int_{\mathbb{R}^d} e^{-i(w\cdot x)} f(x)\, dx,$$

# Examples

- Ridge functions appear in the Fourier transform

$$(\mathcal{F}f)(w) = \int_{\mathbb{R}^d} e^{-i(w \cdot x)} f(x)\, dx,$$

- Radon transform

$$(\mathcal{R}f)(t) = \int_{a \cdot x = t} f(x)\, d\sigma(x),$$

# Examples

- Ridge functions appear in the Fourier transform

$$(\mathcal{F}f)(w) = \int_{\mathbb{R}^d} e^{-i(w\cdot x)} f(x)\, dx,$$

- Radon transform

$$(\mathcal{R}f)(t) = \int_{a\cdot x = t} f(x)\, d\sigma(x),$$

- The Hermitte-Genocchi formula for divided differences

$$f[x_0, x_1, \dots, x_n] = \int_{\Sigma_n} f^{(n)}(t \cdot x)\, dx$$

where $\Sigma_n$ is the standard $n$-simplex in $\mathbb{R}^d$,

# Applications

- Diffusion (superposition) of plane waves

# Applications

- Diffusion (superposition) of plane waves
- Ridgelets (Neural Networks): functions of the form $g(a \cdot x + b)$, where $g : \mathbb{R} \to \mathbb{R}$ is fixed function

# Applications

- Diffusion (superposition) of plane waves
- Ridgelets (Neural Networks): functions of the form $g(a \cdot x + b)$, where $g : \mathbb{R} \to \mathbb{R}$ is fixed function
- Learning Theory

# Applications

- Diffusion (superposition) of plane waves
- Ridgelets (Neural Networks): functions of the form $g(a \cdot x + b)$, where $g : \mathbb{R} \to \mathbb{R}$ is fixed function
- Learning Theory
- Tomography

# Applications

- Diffusion (superposition) of plane waves
- Ridgelets (Neural Networks): functions of the form $g(a \cdot x + b)$, where $g : \mathbb{R} \to \mathbb{R}$ is fixed function
- Learning Theory
- Tomography
- Different aspects of ridge functions including applications and also a problems of density and representations see in the book of Allan Pinkus "Ridge Functions", Cambridge tracts in Mathematics

# Finite linear combination of ridges

- Let $n$ be a natural number. We consider the sets of ridge functions with $n$ fixed directions $a_1, \ldots, a_n$

  $$\mathcal{R}_n(a_1, \ldots, a_n) = \{R(x) = r_1(a_1 \cdot x) + \cdots + r_n(a_n \cdot x)\}$$

  of all possible linear combinations of $n$ ridge functions, where $r_i$ run over arbitrary univariate functions $r_1, \ldots, r_n$.

# Finite linear combination of ridges

- Let $n$ be a natural number. We consider the sets of ridge functions with $n$ fixed directions $a_1, \ldots, a_n$

  $$\mathcal{R}_n(a_1, \ldots, a_n) = \{R(x) = r_1(a_1 \cdot x) + \cdots + r_n(a_n \cdot x)\}$$

  of all possible linear combinations of $n$ ridge functions, where $r_i$ run over arbitrary univariate functions $r_1, \ldots, r_n$.
- and with $n$ arbitrary directions

  $$\mathcal{R}_n = \{R(x) = r_1(a_1 \cdot x) + \cdots + r_n(a_n \cdot x)\}$$

  of all possible linear combinations of $n$ ridge functions $r_1(a_1 \cdot x), \ldots, r_n(a_n \cdot x)$ with arbitrary $a_i$.

# Connection with polynomials

- Let $\mathcal{P}_n^d$ be the space of all polynomials on $\mathbb{R}^d$ of degree at most $n$, and

  $$\mathcal{R}_n^{\mathrm{pol}} = \{p_1(a_1 \cdot x) + \cdots + p_n(a_n \cdot x) : \ p_i \in \mathcal{P}_n^1, \ a_i \in \mathbb{R}^d\}.$$

  be the set of combinations of ridge functions with polynomial profiles.

# Connection with polynomials

- Let $\mathcal{P}_n^d$ be the space of all polynomials on $\mathbb{R}^d$ of degree at most $n$, and

  $$\mathcal{R}_n^{\mathrm{pol}} = \{p_1(a_1 \cdot x) + \dots + p_n(a_n \cdot x) : \ p_i \in \mathcal{P}_n^1, \ a_i \in \mathbb{R}^d\}.$$

  be the set of combinations of ridge functions with polynomial profiles.

- Then

$$\mathcal{P}_n^d = \mathcal{R}_{n^{d-1}}^{\mathrm{pol}}$$

# The orthogonal basis of polynomial

- We construct the basis $\mathbf{P} = \mathbf{P}(\mathbb{B}^d)$ on the unit ball $\mathbb{B}^d$ by two following bases: the Gegenbauer polynomial basis on segment $\mathbb{I} = [-1, 1]$

$$\mathbf{C}(\mathbb{I}) = \{C_i^{\frac{d-1}{2}}(t)\}, \quad i = 0, 1, \ldots$$

# The orthogonal basis of polynomial

- We construct the basis $\mathbf{P} = \mathbf{P}(\mathbb{B}^d)$ on the unit ball $\mathbb{B}^d$ by two following bases: the Gegenbauer polynomial basis on segment $\mathbb{I} = [-1, 1]$

$$\mathbf{C}(\mathbb{I}) = \{C_i^{\frac{d-1}{2}}(t)\}, \quad i = 0, 1, \ldots$$

- and the spherical harmonic basis on unit sphere $\mathbb{S}^{d-1}$

$$\mathbf{Y}(\mathbb{S}^{d-1}) = \{Y_j(\xi)\}, \quad j = 0, 1, \ldots$$

# The orthogonal basis of polynomial

- Construct the basis $\mathbf{P}$ by the following convolution of the basis $\mathbf{C}(\mathbb{I})$ and the basis $\mathbf{Y}(\mathbb{S}^{d-1})$

$$\mathbf{P}(\mathbb{B}^d) := \mathbf{C}(\mathbb{I}) \otimes \mathbf{Y}(\mathbb{S}^{d-1})$$

$$:= \{P_{ij} = C_i^{\frac{d-1}{2}} \otimes Y_j\},$$

# The orthogonal basis of polynomial

- Construct the basis $\mathbf{P}$ by the following convolution of the basis $\mathbf{C}(\mathbb{I})$ and the basis $\mathbf{Y}(\mathbb{S}^{d-1})$

$$\mathbf{P}(\mathbb{B}^d) := \mathbf{C}(\mathbb{I}) \otimes \mathbf{Y}(\mathbb{S}^{d-1})$$

$$:= \{P_{ij} = C_i^{\frac{d-1}{2}} \otimes Y_j\},$$

- where

$$C_i^{\frac{d-1}{2}} \otimes Y_j := \int_{\mathbb{S}^{d-1}} C_i^{\frac{d-1}{2}}(x \cdot \xi) Y_j(\xi)\, d\xi$$

# Properties of the basis $\mathbf{P}$

- The system

$$\mathbf{P} = \{P_{i,j}\}$$

  is the complete orthogonal system of polynomials in the Hilbert space $L_2$ of functions defined on the ball $\mathbb{B}^d$.

# Properties of basis $\mathbf{P}$

- $\mathbf{P}$ is invariante with respect to rotation operator in $\mathbb{R}^d$, i.e. if $A \in \mathbb{SO}(d)$, and

$$T(A)f = f(Ax)$$

is rotation operator in $\mathbb{R}^d$ with the matrix of operator $T(A)$ in the basis $\mathbf{P}$

$$\mathbf{T}(A) = (T(A)P_{i,j}, P_{i',j'}), \qquad P_{i,j}, P_{i',j'} \in \mathbf{P},$$

# Properties of basis **P**

- **P** is invariante with respect to rotation operator in $\mathbb{R}^d$, i.e. if $A \in \mathbb{SO}(d)$, and

$$T(A)f = f(Ax)$$

is rotation operator in $\mathbb{R}^d$ with the matrix of operator $T(A)$ in the basis **P**

$$\mathbf{T}(A) = (T(A)P_{i,j}, P_{i',j'}), \qquad P_{i,j}, P_{i',j'} \in \mathbf{P},$$

- then

$$\mathbf{P}(A\,\cdot) = \mathbf{T}(A)\mathbf{P}(\cdot)$$

# Approximation by ridge functions

- Let $L_q$, $1 \le q \le \infty$, be the space of $q$-integrable functions on the unite ball $\mathbb{B}^d = \{x : |x| \le 1\}$.

# Approximation by ridge functions

- Let $L_q$, $1 \leq q \leq \infty$, be the space of $q$-integrable functions on the unite ball $\mathbb{B}^d = \{x : \ |x| \leq 1\}$.
- Given natural $n$ consider in the space $L_p$ the set

$$\mathcal{R}_n = \{R(x) = r_1(a_1 \cdot x) + \cdots + r_n(a_n \cdot x)\}$$

of all possible linear combinations of $n$ ridge functions $r_1(a_1 \cdot x), \ldots, r_n(a_n \cdot x)$.

# Approximation by ridge functions

- Let $W = \{f\}$ be a function class. Define the best approximation of a function class $W$ by the set of ridge functions $\mathcal{R}_n$:

$$e(W, \mathcal{R}_n, L_q) = \sup_{f \in W} e(f, \mathcal{R}_n, L_q)$$

$$= \sup_{f \in W} \inf_{R \in \mathcal{R}_n} \|f - R\|_{L_q}$$

# Connection with alg. polynomials

- We have

$$\mathcal{R}^{\mathrm{pol}}_{n^{d-1}} = \mathcal{P}^d_n,$$

# Connection with alg. polynomials

- We have

$$\mathcal{R}^{\mathrm{pol}}_{n^{d-1}} = \mathcal{P}^d_n,$$

- and hence

$$e(f, \mathcal{R}_{n^{d-1}}) \leq e(f, \mathcal{R}^{\mathrm{pol}}_{n^{d-1}}).$$

# Connection with alg. polynomials

- We have

$$\mathcal{R}^{\rm pol}_{n^{d-1}} = \mathcal{P}^d_n,$$

- and hence

$$e(f, \mathcal{R}_{n^{d-1}}) \le e(f, \mathcal{R}^{\rm pol}_{n^{d-1}}).$$

- Therefore,

$$e(f, \mathcal{R}_{n^{d-1}}) \le e(f, \mathcal{P}^d_n).$$

# Connection with trig. polynomials

- Let $f \in L_2$. Consider the $i$-moment vector $\hat{f}_i = (\langle f, P_{i,j} \rangle)_{j=1,\ldots,m_i}$ of the function $f$ where $m_i = \mathrm{dim}\mathcal{H}_i$.

# Connection with trig. polynomials

- Let $f \in L_2$. Consider the $i$-moment vector $\hat{f}_i = (\langle f, P_{i,j} \rangle)_{j=1,\dots,m_i}$ of the function $f$ where $m_i = \mathrm{dim} \mathcal{H}_i$.
- Let $a_1, \dots, a_n$ be any points on the sphere $\mathbb{S}^{d-1}$. Introduce in the space $\mathbb{R}^{m_i}$ the subspace of trigonometric polynomials

$$\mathcal{T}_i(a_1, \dots, a_n) = \mathrm{span}\{(Y_{ij}(a_1))_{j=1}^{m_i}, \dots, (Y_{ij}(a_n))_{j=1}^{m_i}\}$$

with harmonics $a_1, \dots, a_n$.

# Connection with trig. polynomials

- Let $f \in L_2$. Consider the $i$-moment vector $\hat{f}_i = (\langle f, P_{i,j} \rangle)_{j=1,\ldots,m_i}$ of the function $f$ where $m_i = \mathrm{dim}\mathcal{H}_i$.
- Let $a_1, \ldots, a_n$ be any points on the sphere $\mathbb{S}^{d-1}$. Introduce in the space $\mathbb{R}^{m_i}$ the subspace of trigonometric polynomials

$$\mathcal{T}_i(a_1, \ldots, a_n) = \mathrm{span}\{(Y_{ij}(a_1))_{j=1}^{m_i}, \ldots, (Y_{ij}(a_n))_{j=1}^{m_i}\}$$

  with harmonics $a_1, \ldots, a_n$.
- Then

$$e(f, \mathcal{R}_n) = \inf_{a_1,\ldots,a_n} \sum_{i=0}^{\infty} e(\hat{f}_i, \mathcal{T}_i(a_1, \ldots, a_n), l_2^{m_i})$$

# Connection with trig. polynomials

- Let $f \in L_2$. Consider the $i$-moment vector $\hat{f}_i = (\langle f, P_{i,j} \rangle)_{j=1,\dots,m_i}$ of the function $f$ where $m_i = \mathrm{dim}\mathcal{H}_i$.

# Connection with trig. polynomials

- Let $f \in L_2$. Consider the $i$-moment vector $\hat{f}_i = (\langle f, P_{i,j} \rangle)_{j=1,\dots,m_i}$ of the function $f$ where $m_i = \mathrm{dim}\mathcal{H}_i$.
- Let $a_1, \dots, a_n$ be any points on the sphere $\mathbb{S}^{d-1}$. Introduce in the space $\mathbb{R}^{m_i}$ the subspace of trigonometric polynomials

$$\mathcal{T}_i(a_1, \dots, a_n) = \mathrm{span}\{(Y_{ij}(a_1))_{j=1}^{m_i}, \dots, (Y_{ij}(a_n))_{j=1}^{m_i}\}$$

with harmonics $a_1, \dots, a_n$.

# Connection with trig. polynomials

- Let $f \in L_2$. Consider the $i$-moment vector $\hat{f}_i = (\langle f, P_{i,j} \rangle)_{j=1,\ldots,m_i}$ of the function $f$ where $m_i = \mathrm{dim}\mathcal{H}_i$.
- Let $a_1, \ldots, a_n$ be any points on the sphere $\mathbb{S}^{d-1}$. Introduce in the space $\mathbb{R}^{m_i}$ the subspace of trigonometric polynomials

$$\mathcal{T}_i(a_1, \ldots, a_n) = \mathrm{span}\{(Y_{ij}(a_1))_{j=1}^{m_i}, \ldots, (Y_{ij}(a_n))_{j=1}^{m_i}\}$$

with harmonics $a_1, \ldots, a_n$.
- Then

$$e(f, \mathcal{R}_n) = \inf_{a_1,\ldots,a_n} \sum_{i=0}^{\infty} e(\hat{f}_i, \mathcal{T}_i(a_1, \ldots, a_n), l_2^{m_i})$$

# Approximation of smoothness functions

- Let $L_q(\mathbb{B}^d)$ be the space of $q$-integrable functions on $\mathbb{B}^d$. Consider the Sobolev class of functions

$$W_p^r(\mathbb{B}^d) = \{f \in L_p : \ \|f\|_{L_p} + \sum_{|s|=r} \|\mathcal{D}^s f\|_{L_p} \leq 1\}.$$

in the space $L_q(\mathbb{B}^d)$ (the case of compact embedding $\frac{r}{d} > (\frac{1}{p} - \frac{1}{q})_+$).

# Approximation of smoothness functions

- Theorem 1 (Uniform distribution of directions)
  If $d = 2$ and $1 \leq p, q \leq \infty$ then

$$c_1 \frac{n^{(\frac{3}{p} - \frac{3}{q})_+}}{n^r} \leq e(W_p^r, \mathcal{R}_n^{\mathrm{unif}}, L_q) \leq c_2 \frac{n^{(\frac{3}{p} - \frac{3}{q})_+}}{n^r}.$$

# Approximation of smoothness functions

- Theorem 1 (Uniform distribution of directions)
  If $d = 2$ and $1 \le p, q \le \infty$ then
$$c_1 \frac{n^{(\frac{3}{p} - \frac{3}{q})_+}}{n^r} \le e(W_p^r, \mathcal{R}_n^{\mathrm{unif}}, L_q) \le c_2 \frac{n^{(\frac{3}{p} - \frac{3}{q})_+}}{n^r}.$$
- Theorem 2 (Arbitrary distribution of directions) If $1 \le q \le p \le \infty$ then
$$\frac{c_1}{n^{\frac{r}{d-1}}} \le e(W_p^r, \mathcal{R}_n^{\mathrm{arb}}, L_q) \le \frac{c_2}{n^{\frac{r}{d-1}}}$$
  (Oskolkov 1998, $d = 2,\ p = q = 2$,
  Maiorov 1999, $d \ge 2,\ q \le p$)

# Approximation of smoothness functions

- Consequence 1 If $p = q$, then the approximations by ridges with uniform and arbitrary distributions of directions are asymptoticaly coincide, that is

$$e(W_p^r, \mathcal{R}_n^{\mathrm{arb}}, L_p) \asymp e(W_p^r, \mathcal{R}_n^{\mathrm{unif}}, L_p) \asymp \frac{1}{n^{\frac{r}{d-1}}}$$

# Sketch of proof

- Upper bound follows from the inclusion $\mathcal{R}_n^{pol} \subseteq \mathcal{P}_n^d$ and known results about approximation multivarite functions by polynomials.

# Sketch of proof

- Upper bound follows from the inclusion $\mathcal{R}_n^{pol} \subseteq \mathcal{P}_n^d$ and known results about approximation multivarite functions by polynomials.

- Lower bound is based on estimation of number of connected components of the polynomial variety of vector-moments

$$\mathcal{M}_n = \{(\langle \mathcal{R}(a_1, \dots, a_n), \mathbf{P} \rangle) : \ a_1, \dots, a_n \in \mathbb{R}^d\}$$

.

# Efficiency of arbitrary directions

- In the case $p > q$ the approximation by ridges with arbitrary directions is better of approximation with uniform distribution

# Efficiency of arbitrary directions

- In the case $p > q$ the approximation by ridges with arbitrary directions is better of approximation with uniform distribution
- Theorem 3 Let $d = 2$, $p = 2$ and $2 \leq q \leq \infty$. Then

$$e(W_2^r, \mathcal{R}_n^{\mathrm{arb}}, L_q) \asymp n^{-r+2}$$

# Efficiency of arbitrary directions

- In the case $p > q$ the approximation by ridges with arbitrary directions is better of approximation with uniform distribution
- Theorem 3 Let $d = 2$, $p = 2$ and $2 \leq q \leq \infty$. Then

$$e(W_2^r, \mathcal{R}_n^{\mathrm{arb}}, L_q) \asymp n^{-r+2}$$

- just as

$$e(W_2^r, \mathcal{R}_n^{\mathrm{unif}}, L_q) \asymp n^{-r+3}$$

# Approximation of harmonic functions

- Let $h$ be a harmonic function on ball $\mathbb{B}^d$ (i.e $\Delta f = 0$).

# Approximation of harmonic functions

- Let $h$ be a harmonic function on ball $\mathbb{B}^d$ (i.e $\Delta f = 0$).
- Assume that best approximation of harmonic function $h$ by the subspace

$$\mathcal{P}_n = \mathrm{span}\{x_1^{s_1} \cdots x_d^{s_d} : \ s_1 + \cdots s_d \le n\}$$

of polynomials of degree $n$ in $L_2$-norm satisfies

# Approximation of harmonic functions

- Let $h$ be a harmonic function on ball $\mathbb{B}^d$ (i.e $\Delta f = 0$).
- Assume that best approximation of harmonic function $h$ by the subspace

$$\mathcal{P}_n = \mathrm{span}\{x_1^{s_1} \cdots x_d^{s_d} : \ s_1 + \cdots s_d \leq n\}$$

of polynomials of degree $n$ in $L_2$-norm satisfies

- $$e(h, \mathcal{P}_n) \asymp \frac{1}{n^{\alpha}}.$$

# Approximation of harmonic functions

- Theorem 4 Let $a_1, \dots, a_n$ be a fixed collection of equidistributed point on the sphere $\mathbb{S}^{d-1}$. Then the best approximation of $h$ by the ridge functions subspace $\mathcal{R}_n^{\mathrm{unif}}$ satisfies

$$e(h, \mathcal{R}_n^{\mathrm{unif}}) \asymp \frac{1}{n^{\frac{\alpha}{d-1}}}.$$

# Approximation of harmonic functions

- **Theorem 4** Let $a_1, \dots, a_n$ be a fixed collection of equidistributed point on the sphere $\mathbb{S}^{d-1}$. Then the best approximation of $h$ by the ridge functions subspace $\mathcal{R}_n^{\mathrm{unif}}$ satisfies

$$e(h, \mathcal{R}_n^{\mathrm{unif}}) \asymp \frac{1}{n^{\frac{\alpha}{d-1}}}.$$

- **Theorem 5** The best approximation of $h$ by the ridge functions subspace $\mathcal{R}_n^{\mathrm{arb}}$

$$e(h, \mathcal{R}_n^{\mathrm{arb}}) \asymp \frac{1}{n^{\frac{\alpha}{d-\frac{3}{2}}}} \quad (\text{with factor } \ln n).$$

(Oskolkov 1998, $d = 2$, Maiorov 2014, $d \geq 2$)

# Approximation of radial functions

- Let $g(x)$ be a radial function of the form $g(x) = q(\|x\|)$.

# Approximation of radial functions

- Let $g(x)$ be a radial function of the form $g(x) = q(\|x\|)$.
- Theorem 6 Let $g$ be a radial $s$-monotone function on $\mathbb{B}^d$, and $e(g, \mathcal{P}_n) \asymp \frac{1}{n^\alpha}$.

# Approximation of radial functions

- Let $g(x)$ be a radial function of the form $g(x) = q(\|x\|)$.
- Theorem 6 Let $g$ be a radial $s$-monotone function on $\mathbb{B}^d$, and $e(g, \mathcal{P}_n) \asymp \frac{1}{n^\alpha}$.
- Then the best approximation of $g$ by the ridge functions set $\mathcal{R}_n$ with arbitrary $n$ directions satisfies

$$e(g, \mathcal{R}_n) \asymp \frac{1}{n^{\frac{\alpha}{d-1}}}$$

(Oskolkov 1998, $d = 2$, Konovalov, Leviatan, Maiorov, 2008))

# Approximation of convex functions

- Theorem 5 Let $\mathcal{F}$ be the class of all convex Lipschitz functions on $\mathbb{B}^d$.

# Approximation of convex functions

- Theorem 5 Let $\mathcal{F}$ be the class of all convex Lipschitz functions on $\mathbb{B}^d$.
- Then the approximation of $\mathcal{F}$ by the ridge functions set $\mathcal{R}_n$ with arbitrary $n$ directions satisfies

$$e(\mathcal{F}, \mathcal{R}_n^{\mathrm{arb}}) \asymp \frac{1}{n^{\frac{3}{2(d-1)}}}$$

(Konovalov, Kopotun, Maiorov, 2010)

THANKS