Average best m-term approximation

Jan Vybíral
Austrian Academy of Sciences
RICAM, Linz, Austria

October 2010
SAFS '10
Oppurg, Germany

Outline

- Introduction, motivation and notation
  - Best m-term approximation
  - Models of noise and signals
  - $p = 2$
  - Definition of average best m-term appr.

Outline

- Introduction, motivation and notation
  - Best m-term approximation
  - Models of noise and signals
  - $p = 2$
  - Definition of average best m-term appr.
- Cone measure
- Surface measure

Outline

- Introduction, motivation and notation
  - Best m-term approximation
  - Models of noise and signals
  - $p = 2$
  - Definition of average best m-term appr.
- Cone measure
- Surface measure
- Tensor product measures
- Numerics, graphs, literature

Best m-term approximation

$m \in \mathbb{N}_0$:

$$\Sigma_m := \Big\{ x = \{x_j\}_{j=1}^{\infty} : \#\{j \in \mathbb{N} : x_j \neq 0\} \leq m \Big\}$$

Best m-term approximation

$m \in \mathbb{N}_0$:

$$\Sigma_m := \Big\{ x = \{x_j\}_{j=1}^{\infty} : \#\{j \in \mathbb{N} : x_j \neq 0\} \leq m \Big\}$$

$x \in \ell_q$: *best m-term approximation* error

$$\sigma_m(x)_q := \inf_{y \in \Sigma_m} \|x - y\|_q$$

Best m-term approximation

$m \in \mathbb{N}_0$:

$$\Sigma_m := \left\{ x = \{x_j\}_{j=1}^{\infty} : \#\{j \in \mathbb{N} : x_j \neq 0\} \leq m \right\}$$

$x \in \ell_q$: *best m-term approximation* error

$$\sigma_m(x)_q := \inf_{y \in \Sigma_m} \|x - y\|_q$$

$0 < p \leq q \leq \infty$: *best m-term approximation widths*

$$\sigma_m^{p,q} := \sup_{x : \|x\|_p \leq 1} \sigma_m(x)_q$$

$$2^{-1/p}(m+1)^{1/q-1/p} \leq \sigma_m^{p,q} \leq (m+1)^{1/q-1/p}, \quad m = 0, 1, 2, \dots$$

Applications of best m-term appr.

Applications of best *m*-term appr.

Signal processing: lossy approximation techniques, especially MP3, JPEG, JPEG-2000 and some video formats

Applications of best *m*-term appr.

Signal processing: lossy approximation techniques, especially MP3, JPEG, JPEG-2000 and some video formats

Numerics of PDE's: Dahlke, Cohen, DeVore, Schwab, Hackbusch, Schneider, . . .

Applications of best m-term appr.

Signal processing: lossy approximation techniques, especially MP3, JPEG, JPEG-2000 and some video formats

Numerics of PDE's: Dahlke, Cohen, DeVore, Schwab, Hackbusch, Schneider, . . .

Benchmark for algorithms and a theoretical tool in approximation theory:
Temlyakov, Kashin, Oswald and many others

Use of random structures in discrete mathematics:
Probabilistic method - Erdös, Alon, Spencer, Matoušek, . . .

Use of random structures in discrete mathematics:
Probabilistic method - Erdös, Alon, Spencer, Matoušek, ...

Use of random structures in functional analysis:
(Khintchine), Gluskin, Kashin, Johnson & Lindenstrauss,...

Use of random structures in discrete mathematics:
Probabilistic method - Erdös, Alon, Spencer, Matoušek, ...

Use of random structures in functional analysis:
(Khintchine), Gluskin, Kashin, Johnson & Lindenstrauss,...

Use of random structures in applications (numerical analysis, informatics):
Given (fixed) input is processed with randomized algorithm and gives (with *high probability*) the right output

Compressed Sensing, Monte Carlo,
Approximate Nearest Neighbor

Use of random structures in discrete mathematics:
Probabilistic method - Erdös, Alon, Spencer, Matoušek, ...

Use of random structures in functional analysis:
(Khintchine), Gluskin, Kashin, Johnson & Lindenstrauss,...

Use of random structures in applications (numerical analysis, informatics):
Given (fixed) input is processed with randomized algorithm and gives (with *high probability*) the right output
Compressed Sensing, Monte Carlo,
Approximate Nearest Neighbor

Random signals could be used to test the speed of algorithms

Models of noise

White noise in signal processing (sound, music, video, ...): $\varepsilon\omega$

$\omega = (\omega_1, \dots, \omega_n) \in \mathbb{R}^n$ - independent (normal) Gaussian variables
$\varepsilon \cdot \|\omega\|_2$ - energy level of the noise

Models of noise

White noise in signal processing (sound, music, video, ...): $\varepsilon\omega$

$\omega = (\omega_1, \ldots, \omega_n) \in \mathbb{R}^n$ - independent (normal) Gaussian variables
$\varepsilon \cdot \|\omega\|_2$ - energy level of the noise

The name comes from *white light*
spectral density is uniformly spread over the visible spectrum

Used extensively in signal processing, basic model for "denoising" and other image and video filters

Models of noise

White noise in signal processing (sound, music, video, ...): $\varepsilon\omega$

$\omega = (\omega_1, \dots, \omega_n) \in \mathbb{R}^n$ - independent (normal) Gaussian variables
$\varepsilon \cdot \|\omega\|_2$ - energy level of the noise

The name comes from *white light*
spectral density is uniformly spread over the visible spectrum

Used extensively in signal processing, basic model for "denoising" and other image and video filters

To test an algorithm: generate white noise many times...

Infinite-dimensional models of white noise?

Theoretically, one may consider $\omega = (\omega_1, \dots, \omega_n, \dots)$

Infinite-dimensional models of white noise?

Theoretically, one may consider $\omega = (\omega_1, \dots, \omega_n, \dots)$

But it has almost surely $\|\omega\|_2 = \infty$!

Therefore, one usually chooses n large, but finite

Is there a model for signals? No!?

Is there a model for signals? No!?

It is unlikely, that one model could be good for all signals (music, images, tomography, astronomy, ...)

Is there a model for signals? No!?

It is unlikely, that one model could be good for all signals (music, images, tomography, astronomy, ...)

All signals have something in common - they are usually **sparse** (or nearly sparse) in a basis adapted to the specific area: Fourier series, Gabor frames, wavelets, discrete cosine transform, curvelets and other —lets, ('95 - Donoho, Candés, Johnstone, . . .)

Is there a model for signals? No!?

It is unlikely, that one model could be good for all signals (music, images, tomography, astronomy, ...)

All signals have something in common - they are usually **sparse** (or nearly sparse) in a basis adapted to the specific area: Fourier series, Gabor frames, wavelets, discrete cosine transform, curvelets and other —lets, ('95 - Donoho, Candés, Johnstone, . . .)

We are looking for

- a random model in $\mathbb{R}^n$
- the *typical vectors* are nearly sparse
- should depend on several parameters

Bernoulli-Gaussian model for sparse vectors

Let $x \in \mathbb{R}^n$ with

$$x_i = \varepsilon \varrho_i \omega_i, \quad i = 1, \dots, n,$$

where $\varepsilon > 0$,

$$\varrho_i = \begin{cases} 1 & \text{with prob. } p, \\ 0 & \text{with prob. } 1-p, \end{cases}$$

$0 < p \ll 1$ and $(\omega_1, \dots, \omega_n)$ are again Gaussian variables

Bernoulli-Gaussian model for sparse vectors

Let $x \in \mathbb{R}^n$ with

$$x_i = \varepsilon \varrho_i \omega_i, \quad i = 1, \ldots, n,$$

where $\varepsilon > 0$,

$$\varrho_i = \begin{cases} 1 & \text{with prob. } p, \\ 0 & \text{with prob. } 1-p, \end{cases}$$

$0 < p \ll 1$ and $(\omega_1, \ldots, \omega_n)$ are again Gaussian variables

The number of non-zero components is concentrated around $k := pn \ll n$
Badly concentrated for small k's, two "noisy" for large k's

Although x is sparse, it is rather a *randomly filtered noise* then a signal

The linear algebraic equation

$$Ax = y, \quad A \in \mathbb{R}^{n \times n}, \quad y \in \mathbb{R}^n$$

has (if A is regular) exactly one solution x for every y

The linear algebraic equation

$$Ax = y, \quad A \in \mathbb{R}^{n \times n}, \quad y \in \mathbb{R}^n$$

has (if A is regular) exactly one solution x for every y

Compressed sensing describes, how to solve the linear equation

$$Ax = y, \quad A \in \mathbb{R}^{k \times n}, \quad k \ll n, \quad y \in \mathbb{R}^n$$

for special (usually randomly generated) matrices A using the a-priori knowledge, that x is sparse

The linear algebraic equation

$$Ax = y, \quad A \in \mathbb{R}^{n\times n}, \quad y \in \mathbb{R}^n$$

has (if A is regular) exactly one solution x for every y

Compressed sensing describes, how to solve the linear equation

$$Ax = y, \quad A \in \mathbb{R}^{k\times n}, \quad k \ll n, \quad y \in \mathbb{R}^n$$

for special (usually randomly generated) matrices A using the a-priori knowledge, that x is sparse

The methods of compressed sensing extend naturally to $\ell_p^n, p < 1$!

Are vectors from the unit ball of ℓ_p^n really *nearly sparse*?

How does a typical vector of the ℓ_p^n unit ball look like?

or:

Let μ be a probability measure on the unit ball of ℓ_p^n. What is the mean value of $\sigma_m(x)_q$ with respect to this measure?

How does a typical vector of the ℓ_p^n unit ball look like?

or:

Let μ be a probability measure on the unit ball of ℓ_p^n. What is the mean value of $\sigma_m(x)_q$ with respect to this measure?

Choice of μ?

(i) the normalized Lebesgue measure,

(ii) the $n-1$ dimensional Hausdorff measure restricted to the surface of the unit ball of ℓ_p^n and correspondingly normalized,

(iii) the so-called normalized cone measure

Well known case: $p = 2$, $q = \infty$, $m = 0$
μ . . . normalized surface measure on $\mathbb{S}^{n-1}$

....i.e. we want to calculate

$$\int_{\mathbb{S}^{n-1}} \max_{j=1,\dots,n} |x_j| d\mu(x) = \int_{\mathbb{S}^{n-1}} \|x\|_\infty d\mu(x) = \int_{\mathbb{S}^{n-1}} \sigma_0(x)_\infty d\mu(x)$$

Well known case: $p = 2,\ q = \infty,\ m = 0$
μ ... normalized surface measure on $\mathbb{S}^{n-1}$

....i.e. we want to calculate

$$\int_{\mathbb{S}^{n-1}} \max_{j=1,\dots,n} |x_j| d\mu(x) = \int_{\mathbb{S}^{n-1}} \|x\|_\infty d\mu(x) = \int_{\mathbb{S}^{n-1}} \sigma_0(x)_\infty d\mu(x)$$

γ_n ... the standard Gaussian measure on $\mathbb{R}^n$ with the density

$$\frac{1}{(2\pi)^{n/2}} e^{-\|x\|_2^2/2}, \quad x \in \mathbb{R}^n$$

Polar coordinates: ($\Omega_n =$ the area of $\mathbb{S}^{n-1}$)

$$\int_{\mathbb{R}^n} \max_{j=1,\dots,n} |x_j| \, d\gamma_n(x) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \max_{j=1,\dots,n} |x_j| \cdot e^{-\|x\|_2^2/2} dx$$

Polar coordinates: ($\Omega_n =$ the area of $\mathbb{S}^{n-1}$)

$$\int_{\mathbb{R}^n} \max_{j=1,\dots,n} |x_j| \, d\gamma_n(x) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \max_{j=1,\dots,n} |x_j| \cdot e^{-\|x\|_2^2/2} dx$$

$$= \frac{\Omega_n}{(2\pi)^{n/2}} \int_0^\infty r^{n-1} \int_{\mathbb{S}^{n-1}} \max_{j=1,\dots,n} |rx_j| e^{-\|rx\|_2^2/2} d\mu(x) \, dr$$

Polar coordinates: ($\Omega_n =$ the area of $\mathbb{S}^{n-1}$)

$$\int_{\mathbb{R}^n} \max_{j=1,\dots,n} |x_j| \, d\gamma_n(x) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \max_{j=1,\dots,n} |x_j| \cdot e^{-\|x\|_2^2/2} dx$$

$$= \frac{\Omega_n}{(2\pi)^{n/2}} \int_0^\infty r^{n-1} \int_{\mathbb{S}^{n-1}} \max_{j=1,\dots,n} |rx_j| e^{-\|rx\|_2^2/2} d\mu(x) \, dr$$

$$= \frac{\Omega_n}{(2\pi)^{n/2}} \int_0^\infty r^n e^{-r^2/2} dr \cdot \int_{\mathbb{S}^{n-1}} \max_{j=1,\dots,n} |x_j| d\mu(x)$$

Maximum of n indep. Gaussian variables: $\approx \sqrt{\log(n+1)}$,

Maximum of n indep. Gaussian variables: $\approx \sqrt{\log(n+1)}$,

$$\int_0^\infty r^n e^{-r^2/2} dr = 2^{(n-1)/2}\Gamma((n+1)/2)$$

Maximum of n indep. Gaussian variables: $\approx \sqrt{\log(n+1)}$,

$$\int_0^\infty r^n e^{-r^2/2} dr = 2^{(n-1)/2}\Gamma((n+1)/2)$$

and

$$\Omega_n = \frac{2\pi^{n/2}}{\Gamma(n/2)}:$$

Maximum of n indep. Gaussian variables: $\approx \sqrt{\log(n+1)}$,

$$\int_0^\infty r^n e^{-r^2/2} dr = 2^{(n-1)/2}\Gamma((n+1)/2)$$

and

$$\Omega_n = \frac{2\pi^{n/2}}{\Gamma(n/2)}:$$

$$\int_{\mathbb{S}^{n-1}} \sigma_0(x)_\infty d\mu(x) \approx \sqrt{\frac{\log(n+1)}{n}}, \quad n \in \mathbb{N}$$

Interpretation:

The average coordinate of $x \in \mathbb{S}^{n-1}$: $= n^{-1/2}$
The average of the maximal coordinates: $\approx \sqrt{log(n+1)} \cdot n^{-1/2}$

Interpretation:

The average coordinate of $x \in \mathbb{S}^{n-1}$: $= n^{-1/2}$
The average of the maximal coordinates: $\approx \sqrt{log(n+1)} \cdot n^{-1/2}$

A random point of $x \in \mathbb{S}^{n-1}$ is a good model of *noise* not of *signals*

Interpretation:

The average coordinate of $x \in \mathbb{S}^{n-1}$: $= n^{-1/2}$
The average of the maximal coordinates: $\approx \sqrt{log(n+1)} \cdot n^{-1/2}$

A random point of $x \in \mathbb{S}^{n-1}$ is a good model of *noise* not of *signals*

This, and other aspects of the geometry of $\mathbb{S}^{n-1}$ is well known as the *concentration of measure* (V. Milman, M. Talagrand, M. Ledoux)

Interpretation:

The average coordinate of $x \in \mathbb{S}^{n-1}$: $= n^{-1/2}$
The average of the maximal coordinates: $\approx \sqrt{log(n+1)} \cdot n^{-1/2}$

A random point of $x \in \mathbb{S}^{n-1}$ is a good model of *noise* not of *signals*

This, and other aspects of the geometry of $\mathbb{S}^{n-1}$ is well known as the *concentration of measure* (V. Milman, M. Talagrand, M. Ledoux)

Surprisingly, the same holds true for all $0 < p$!

Definitions: $0 < p \leq q \leq \infty$, $m \leq n$:

$$\Delta_p^n = \{(t_1, \ldots, t_n) \in \mathbb{R}_+^n : \sum_{j=1}^{n} t_j^p = 1\}$$

Definitions: $0 < p \le q \le \infty$, $m \le n$:

$\Delta_p^n = \{(t_1, \ldots, t_n) \in \mathbb{R}_+^n : \sum_{j=1}^n t_j^p = 1\}$

$\mu \ldots$ a Borel probability measure on Δ_p^n

$$\sigma_m^{p,q}(\mu) = \int_{\Delta_p^n} \sigma_m(x)_q \, d\mu(x)$$

average surface best m-term width with respect to μ

Definitions: $0 < p \leq q \leq \infty$, $m \leq n$:

$\Delta_p^n = \{(t_1, \ldots, t_n) \in \mathbb{R}_+^n : \sum_{j=1}^n t_j^p = 1\}$

$\mu \ldots$ a Borel probability measure on Δ_p^n

$$\sigma_m^{p,q}(\mu) = \int_{\Delta_p^n} \sigma_m(x)_q \, d\mu(x)$$

average surface best m-term width with respect to μ

$\nu \ldots$ a Borel probability measure on $[0,1] \cdot \Delta_p^n$

$$\sigma_m^{p,q}(\nu) = \int_{\Delta_p^n} \sigma_m(x)_q \, d\nu(x)$$

average volume best m-term width with respect to ν

Cone measure

$n \geq 2$:

$$\mu_p(\mathcal{A}) = \frac{\lambda([0,1] \cdot \mathcal{A})}{\lambda([0,1] \cdot \Delta_p^n)}, \quad \mathcal{A} \subset \Delta_p^n$$

is the normalized *cone measure* on Δ_p^n

Cone measure

$n \geq 2$:

$$\mu_p(\mathcal{A}) = \frac{\lambda([0,1]\cdot\mathcal{A})}{\lambda([0,1]\cdot\Delta_p^n)}, \quad \mathcal{A} \subset \Delta_p^n$$

is the normalized *cone measure* on Δ_p^n

polar decomposition identity:

$$\frac{\int_{\mathbb{R}_+^n} f(x)\,d\lambda(x)}{\lambda([0,1]\cdot\Delta_p^n)} = n\int_0^\infty r^{n-1}\int_{\Delta_p^n} f(rx)\,d\mu_p(x)\,dr, \quad f \in L_1(\mathbb{R}_+^n)$$

. . . substitute for *polar coordinates*

Cone measure

$n \geq 2$:

$$\mu_p(\mathcal{A}) = \frac{\lambda([0,1]\cdot\mathcal{A})}{\lambda([0,1]\cdot\Delta_p^n)}, \quad \mathcal{A} \subset \Delta_p^n$$

is the normalized *cone measure* on Δ_p^n

polar decomposition identity:

$$\frac{\int_{\mathbb{R}_+^n} f(x)d\lambda(x)}{\lambda([0,1]\cdot\Delta_p^n)} = n\int_0^\infty r^{n-1}\int_{\Delta_p^n} f(rx)d\mu_p(x)dr, \quad f \in L_1(\mathbb{R}_+^n)$$

...substitute for *polar coordinates*

$p = 1, 2, \infty$ ...coincides with the surface measure

Let $p = 2$ and let $\omega = (\omega_1, \ldots, \omega_n)$ be indep. Gaussian variables. Then

$$\mathbb{P}\left(\frac{(\omega_1, \ldots, \omega_n)}{\|\omega\|_2} \in \mathcal{A}\right) = \mu(A)$$

Proof: both the measures are rotational invariant on $\mathbb{S}^{n-1}$

Let $p = 2$ and let $\omega = (\omega_1, \dots, \omega_n)$ be indep. Gaussian variables. Then

$$\mathbb{P}\left(\frac{(\omega_1, \dots, \omega_n)}{\|\omega\|_2} \in \mathcal{A}\right) = \mu(A)$$

Proof: both the measures are rotational invariant on $\mathbb{S}^{n-1}$

Variables with density $c_p e^{-t^p}$, $t > 0$, may be used if $p \neq 2$ (Schechtman and Zinn, 1990)

$$\mathbb{P}\left(\frac{(\omega_1, \dots, \omega_n)}{\|\omega\|_p} \in \mathcal{A}\right) = \mu_p(A)$$

Results (V., 2010):

(i) $n \geq 2$, $0 \leq m \leq n-1$, $0 < p < \infty$:

$$\sigma_m^{p,\infty}(\mu_p) \lesssim \left[\frac{\log\left(\frac{en}{m+1}\right)}{n}\right]^{1/p};$$

Results (V., 2010):
(i) $n \geq 2$, $0 \leq m \leq n-1$, $0 < p < \infty$:

$$\sigma_m^{p,\infty}(\mu_p) \lesssim \left[\frac{\log\left(\frac{en}{m+1}\right)}{n} \right]^{1/p};$$

(ii) $n \geq 2$, $0 < p < \infty$:

$$\sigma_0^{p,\infty}(\mu_p) \gtrsim \left[\frac{\log(n+1)}{n} \right]^{1/p};$$

Results (V., 2010):

(i) $n \geq 2$, $0 \leq m \leq n-1$, $0 < p < \infty$:

$$\sigma_m^{p,\infty}(\mu_p) \lesssim \left[\frac{\log\left(\frac{en}{m+1}\right)}{n}\right]^{1/p};$$

(ii) $n \geq 2$, $0 < p < \infty$:

$$\sigma_0^{p,\infty}(\mu_p) \gtrsim \left[\frac{\log(n+1)}{n}\right]^{1/p};$$

(iii) $n \geq 2$, $0 < p \leq q < \infty$:

$$\sigma_0^{p,q}(\mu_p) \approx n^{1/q-1/p}$$

Lebesgue measure

ν_p : p-normalized Lebesgue measure

$$\nu_p(A) = \frac{\lambda(A)}{\lambda([0,1]\cdot\Delta_p^n)}, \quad A \subset \mathbb{R}_+^n$$

Lebesgue measure

ν_p : p-normalized Lebesgue measure

$$\nu_p(A) = \frac{\lambda(A)}{\lambda([0,1] \cdot \Delta_p^n)}, \quad A \subset \mathbb{R}_+^n$$

The identity

$$\sigma_m^{p,q}(\nu_p) = \sigma_m^{p,q}(\mu_p) \cdot \frac{n}{n+1}$$

holds for all $0 < p \leq q \leq \infty$, all $n \geq 2$ and all $0 \leq m \leq n-1$

Surface measure

$n \geq 2$:

$$\varrho_p(\mathcal{A}) = \frac{\mathcal{H}(\mathcal{A})}{\mathcal{H}(\Delta_p^n)}, \qquad \mathcal{A} \subset \Delta_p^n$$

the normalized $n-1$ dimensional Hausdorff measure on Δ_p^n

Methods: (i) cone measure (!)

Methods: (i) cone measure (!)

(ii) Density of ϱ_p w.r.t. μ_p
($p \geq 1$: Naor and Romik, 2003; $p < 1$, V., 2010)

$$\frac{d\varrho_p}{d\mu_p}(x) = \frac{n\lambda([0,1]\cdot\Delta_p^n)}{\mathcal{H}(\Delta_p^n)}\Big\|\nabla(\|\cdot\|_p)(x)\Big\|_2 = C_{p,n}^{-1}\left(\sum_{i=1}^n x_i^{2p-2}\right)^{1/2},$$

where $C_{p,n}$ is the normalizing constant

For $p < 1$, ϱ_p has a singularity if $x_i \to 0$

Methods: (i) cone measure (!)

(ii) Density of ϱ_p w.r.t. μ_p
($p \geq 1$: Naor and Romik, 2003; $p < 1$, V., 2010)

$$\frac{d\varrho_p}{d\mu_p}(x) = \frac{n\lambda([0,1] \cdot \Delta_p^n)}{\mathcal{H}(\Delta_p^n)} \left\| \nabla(\| \cdot \|_p)(x) \right\|_2 = C_{p,n}^{-1} \left(\sum_{i=1}^{n} x_i^{2p-2} \right)^{1/2},$$

where $C_{p,n}$ is the normalizing constant

For $p < 1$, ϱ_p has a singularity if $x_i \to 0$

Result (V. 2010): $n \geq 2$, $0 < p \leq \infty$:

$$\sigma_0^{p,\infty}(\varrho_p) \lesssim \left[\frac{\log(n+1)}{n} \right]^{1/p}$$

Tensor products

$0 < p < \infty$, $\beta > -1$: We define $\theta_{p,\beta}$ by

$$\frac{d\theta_{p,\beta}}{d\mu_p}(x) = C_{p,\beta}^{-1} \cdot \prod_{i=1}^{n} x_i^{\beta},$$

Tensor products

$0 < p < \infty$, $\beta > -1$: We define $\theta_{p,\beta}$ by

$$\frac{d\theta_{p,\beta}}{d\mu_p}(x) = C_{p,\beta}^{-1} \cdot \prod_{i=1}^{n} x_i^{\beta},$$

$$C_{p,\beta} = \int_{\Delta_p^n} \prod_{i=1}^{n} x_i^{\beta} d\mu_p(x) \ldots \text{normalizing constant}$$

$\beta > -1$ implies $C_{p,\beta} < \infty$

Result: For every fixed $m \in \mathbb{N}$,

$$\frac{1}{\left(\frac{1}{p}+1\right)^m} \lesssim \liminf_{n\to\infty} \sigma^{p,\infty}_{m-1}(\theta_{p,p/n-1})$$

$$\leq \limsup_{n\to\infty} \sigma^{p,\infty}_{m-1}(\theta_{p,p/n-1}) \lesssim \frac{1}{\left(\frac{1}{p}+1\right)^m} + \frac{e^{-m}}{m!},$$

Interpretation:

The measure $\theta_{p,p/n-1}$ is strongly concentrated on nearly-sparse vectors. A typical vector with respect to $\theta_{p,p/n-1}$ *corresponds* to a structured signal

Can we take the limit $n \to \infty$: extension to function spaces?

Interpretation:

The measure $\theta_{p,p/n-1}$ is strongly concentrated on nearly-sparse vectors. A typical vector with respect to $\theta_{p,p/n-1}$ *corresponds* to a structured signal

Can we take the limit $n \to \infty$: extension to function spaces?

Potentially, one could prove theorems like: *"An algorithm is good for* `almost all` *pictures chosen w.r.t.* $\theta_{p,\beta}$ *..."*

Generating the points w.r.t. $\theta_{p,\beta}$?

Generating the points w.r.t. $\theta_{p,\beta}$?

Generate $\omega_1', \dots, \omega_n'$ w.r.t. the density $c_{p,\beta} t^{\beta} e^{-t^p}, t > 0$, $c_{p,\beta} \dots$ normalizing constant

Generating the points w.r.t. $\theta_{p,\beta}$?

Generate $\omega_1', \ldots, \omega_n'$ w.r.t. the density $c_{p,\beta} t^{\beta} e^{-t^p}, t > 0$,
$c_{p,\beta} \ldots$ normalizing constant

$$\mathbb{P}\left(\frac{(\omega_1', \ldots, \omega_n')}{\|\omega'\|_p} \in \mathcal{A} \right) = \theta_{p,\beta}(\mathcal{A})$$

Generating the points w.r.t. $\theta_{p,\beta}$?

Generate $\omega_1',\dots,\omega_n'$ w.r.t. the density $c_{p,\beta}t^{\beta}e^{-t^p}, t>0$,
$c_{p,\beta}\dots$ normalizing constant

$$\mathbb{P}\left(\frac{(\omega_1',\dots,\omega_n')}{\|\omega'\|_p}\in\mathcal{A}\right)=\theta_{p,\beta}(\mathcal{A})$$

We take

$$\frac{(\omega_1',\dots,\omega_n')}{\left(\sum_{j=1}^n(\omega_j')^p\right)^{1/p}}\in\Delta_p^n.$$

Generating ω_i'?

U ... Uniform random number generator: values in $[0, 1]$

Generating ω_i'?

U . . . Uniform random number generator: values in $[0, 1]$

$p = 2, \beta = 0$, "naive way": $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{X} e^{-t^2/2} dt = U$

Generating ω_i'?

U ... Uniform random number generator: values in $[0,1]$

$p = 2, \beta = 0$, "naive way": $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{X} e^{-t^2/2} dt = U$

Box-Muller method:

$$X = \sqrt{-2 \ln U} \cos(2\pi V), \qquad Y = \sqrt{-2 \ln U} \sin(2\pi V)$$

Generating ω_i'?

U ... Uniform random number generator: values in $[0,1]$

$p=2, \beta=0$, "naive way": $\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{X} e^{-t^2/2}dt = U$

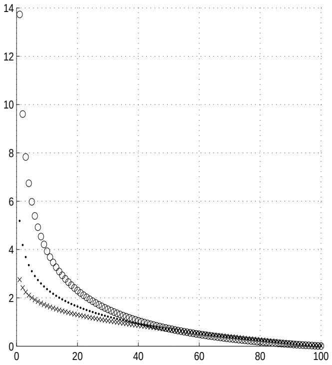
Box-Muller method:

$$X=\sqrt{-2\ln U}\cos(2\pi V), \qquad Y=\sqrt{-2\ln U}\sin(2\pi V)$$

Marsaglia polar method

$$S := U^2+V^2<1, \quad X=U\sqrt{\frac{-2\ln S}{S}}, \quad Y=V\sqrt{\frac{-2\ln S}{S}}$$

...described in detail in Knuth and implemented in *GNU Scientific Library* for general p and β



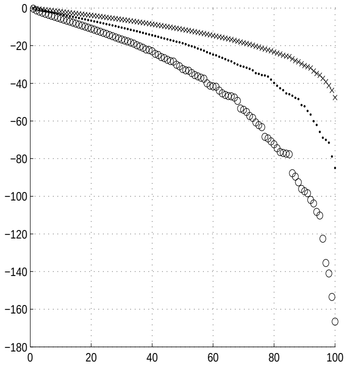


Figure: Approximations of $n^{1/p} \cdot \int_{\Delta_p^n} x_m^* d\mu_p(x)$ (left) and $\log_{10}(\int_{\Delta_p^n} x_m^* d\theta_{p,p/n-1})$ (right) for $n = 100$, $p = 1/2(\circ)$, $p = 1(\bullet)$ and $p = 2(\times)$ based on sampling of 10^8 random points.

Literature:

Geometry of ℓ_p^n balls:

F. Barthe, M. Csörnyei and A. Naor, *A note on simultaneous polar and Cartesian decomposition*, Springer, Berlin, 2003.

F. Barthe, O. Guédon, S. Mendelson and A. Naor, *A probabilistic approach to the geometry of the l_p^n-ball*, Ann. Probab. 33 (2005), no. 2, 480–513.

A. Naor, *The surface measure and cone measure on the sphere of l_p^n.* Trans. Amer. Math. Soc. 359 (2007), no. 3, 1045–1079.

A. Naor and D. Romik, *Projecting the surface measure of the sphere of l_p^n*, Ann. Inst. H. Poincar Probab. Statist. 39 (2003), no. 2, 241–261.

G. Schechtman and J. Zinn, *On the volume of the intersection of two L_p^n balls*, Proc. AMS 110 (1), 217-224, (1990).

Literature:

Non-linear approximation:

K. Oskolkov, *Polygonal approximation of functions of two variables*, Math. USSR Sbornik 35, 851-861, (1979).

R. A. DeVore, *Nonlinear approximation*, Acta Num. 51-150, (1998).

Special functions:

M. Abramowitz and I. A. Stegun, *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, 1964.

Random variables:

M. Ledoux and M. Talagrand, *Probability in Banach spaces*, Springer-Verlag, Berlin, 1991.

Numerics:

GNU Scientific Library, `http://www.gnu.org/software/gsl/`

D. E. Knuth, The Art of Computer Programming, Vol. 2: *Seminumerical Algorithms*, 3rd ed., Addison-Wesley 1998.