Le Compressed Sensing et l'algorithme optimal

CUVILLIER Thomas

24 juin 2010

Définition 1

Soit X un espace de Banach

Soit K compact $K \subset X$

On appelle et on note épaisseur de Gelfand

$$d^n(K)_X = inf_{codim(Y)=n} sup_{x \in K \cap Y} |x|_X$$

Les meilleurs Y dans cette définition sont ceux qui ”traversent“ K sur une “longueur” minimale.

Définition 2

Application aux espaces l_p.
Rappel : On est en dimension finie donc

$$||x||_{l_p} = (\sum_{j=1}^{N} |x_j|^p)^{1/p}$$

On note

$$U(L_p^N) = [xEL_p^N, ||x||_{L_p} \leq 1]$$

C'est la boule unité dans L_p^N

Théorême 1 : Inégalité de Gluskin et Garnaev

$$d^n(U(L_1^N))_{l_2^N} \leq C_0\sqrt{(log(N/n)+1)/n}$$

$\Rightarrow$ l'épaisseur minimale de la boule L_1 en dimension N, par rapport à un sous espace de codimension n varie en $\sqrt{log(N/n)/n}$.

Les épaisseurs de Gelfrand permettent d'estimer les performances d'un algorithme, en les comparant aux performances théoriques maximales.

Ainsi, en algorithmique, on utilise une paire encodeur / décodeur.
Soit x un signal.

Definition encodeur

Un encodeur Φ est une matrice (N, n) qui va extraire des informations (n informations) de x.

$$y = \Phi x$$

Definition décodeur

Le décodeur est un algorithme Δ de $\mathbb{R}^n$ dans $\mathbb{R}^N$ qui va chercher à “retrouver” x, à partir de y
c’est à dire on cherche Δ tel que

$$||\Delta\Phi x - x||$$

Soit minimal.

On mesure l'efficacité de l'algorithme par la donnée suivante

Performance d'un algorithme

Soit K un compact de X,
X espace de Banach
$E(K, \Phi, \Delta)_X = sup_{x \in K} ||\Delta\Phi x - x||_X$

Algorithme optimal

L'algorithme optimal qui utilise n de N informations initiales est
$E_{n,N}(K)_X = inf_{(\Phi,\Delta) \in A_{n,N}}) E(K, \Psi, \Phi)_X$

Algorithme optimal et épaisseurs de Gelfand

Soit K compact, $K \subset \mathbb{R}^N$ tq
$K = -K$ et
$\exists C$ tq $K + K \subset C.K$ alors on a

$$d^n(K)_X \leq E_{n,N}(K)_X < Cd^n(K)_X$$

C'est à dire, la meilleure performance algorithmique possible dépend linéairement des épaisseurs de Gelfand

Premier théorème de Candès et Tao

f signal, $f \in \mathbb{R}^n$

Soit T le support de f

Soit $K \subset [0, 1, ..., n]$ l'ensemble des mesures.

Alors si $K > C.|T|.log(N)$ on peut retrouver parfaitement f à partir de la donnée de f̂(k), avec probabilité $\geq 1 - O(N^{-p.C})$

On résoud $argmin_{x \in \mathbb{R}^N}(||x||_{L_1}, \Phi x = \Phi f)$

Si le support de f, n'est pas “petit”.
Alors on pose $f = f_T + f_{T^c}$ T ensemble des |T| plus grandes valeurs de f. Alors on a $||f - f_{T^c}||_{l_2} \leq 1/\sqrt{|T|}$
On a alors

Second théorème

f signal, $f \in \mathbb{R}^n$
Soit $K \subset [0, 1, ..., n]$ l'ensemble des mesures.
Alors si $K > C.|T|.log(N)$ on peut trouver une approximation x de f qui vérifie

$$||f - x||_{L_2} \leq C/\sqrt{|T|}$$

avec probabilité $\geq 1 - O(N^{-p.C})$

Rappelons les deux inégalités fondamentales.

Rappel

$K > C.|T|.log(N)$ ou K est le nombre de mesures.
$||f - x||_{L_2} \leq C/\sqrt{|T|}$ ou x est le résultat de l'algorithme

On a donc

Efficacité

$||f - x||_{L_2} \leq C\sqrt{log(N)}/|K|$

On rappelle que

efficacité optimale

$E_{n,N}(K)_X < Cd^n(K)_X$

et qu'on a

Inégalité de Gluskin et Garnaev

$$d^n(U(L_1^N))_{l_2^N} \leq C_0\sqrt{(log(N/n)+1)/n}$$

On appelle n (à la place de K) le nombre de mesures. On compare alors l'efficacité de l'algorithme de Candès et Tao avec l'optimale.

comparaison

$||f - x||_{L_2} \leq C\sqrt{log(N)}/|n| \Rightarrow$ Candès et Tao

$E_{n,N}(K)_X \leq C_0\sqrt{(log(N/n)+1)/n} \Rightarrow$ Optimale

L'efficacité de l'algorithme de Candès et Tao est donc très proche de celle optimale, alors qu'elle consiste à effectuer des mesures aléatoires.

L'essentiel

[1] *R. Baraniuk, M.Davenport, R.Devore, et M.Wakin* **The Johnson-Linderstrauss Lemma meets Compressed Sensing**
[2] *E.Candès et T.Tao* **Near Optimal Signal Recovery From Random Projections : Universal Encoding Strategies**
[3] *Yves Meyer* **Compressed sensing**