Compressive Sensing 2022
Claude Petit, Aline Roumy

Quiz & oral questions -2- Corrigé

---

1. **Lister et classer les algorithmes de reconstruction parcimonieuses.**

Pour résoudre le problème P_0, on a trois catégories d'algorithmes:

- Algorithmes de relaxation convexe, qui transforment le problème P_0 en problème convexe P_1 et le résolvent avec des algorithmes d'optimisation convexe, en s'assurant d'avoir des garanties pour que les solutions de P_0 et P_1 soient identiques. Parmi ces algorithmes on trouve Basis Pursuit (BP), Lasso, BPDN et QCBP (ils sont tous trois équivalents).
- Algorithmes de seuillage, qui s'intéressent à $y = Mx$: Basic Tresholding (en une passe), IHT.
- Algorithmes gloutons, qui s'intéressent à la parcimonie et reconstruisent le support itérativement : OMP, CoSaMP et variantes.

2. **Quels sont les principes et points importants de l'acquisition comprimée (CS) ?**

- le CS permet d'échantillonner et comprimer des données de façon simultanée, contrairement aux procédés classiques qui le font séquentiellement et acquièrent des données de façon inutile.
- il s'agit de reconstruire de façon efficace (avec un minimum de mesures) des vecteurs parcimonieux, en exploitant cette parcimonie.
- l'encodage est linéaire (c'est juste une multiplication matricielle), mais le décodage efficace est non linéaire, effectué par des algorithmes de minimisation convexe ou des algorithmes gloutons.
- résoudre directement le système linéaire $y = Mx$ est NP-difficile car il est sous-dimensionné.
- le nombre de mesures m pour reconstruire efficacement tous les vecteurs k-parcimonieux de taille n est de l'ordre de $k \ln n$ en utilisant des algorithmes bien choisis. Le facteur $\ln n$ est le coût qu'il faut payer car on ne connaît pas le support du vecteur parcimonieux.

3. **Quelles sont les applications de l'acquisition comprimée ?**

- Compression de données, représentations parcimonieuses, échantillonnage de données.
- Acquisition de données à moindre coût (décrire le principe de la "single pixel camera" : reconstruire une image grossière à partir de très peu de mesures).
- Imagerie à résonance magnétique (MRI) ; décrire brièvement le principe : interaction entre un champ magnétique en trois dimensions et les protons contenus dans les molécules d'eau présentes dans le corps. Le signal mesuré par le processus MRI est directement la transformée de Fourier en 3D des variations du champ (les données sont donc directement récupérées en fréquence). On utilise alors le fait que certains types d'images médicales sont naturellement parcimonieuses.
- Radar : la fonction transformant le signal émis en signal reçu après réflexions sur les objets peut être modélisée par un opérateur linéaire. La réponse radio est souvent parcimonieuse car la majorité du signal émis n'est pas réfléchi.
- Codes correcteurs d'erreurs.
- Débruitage de signaux.

- Statistique et Machine Learning : modèles de régression pour prédire un phénomène, apprentissage de modèles.

4°. Énoncer les problèmes de reconstruction parcimonieuse P_0, P_1, P_2. Quelles sont leurs différences ?

De façon générale, pour $p \geq 0$,

$$P_p : \ \hat{x} = \underset{z:Mx=Mz}{\operatorname{argmin}} ||z||_p \tag{1}$$

- Si $p = 0$, on cherche la solution la plus parcimonieuse au système $y = Mx = Mz$ (d'inconnue z). C'est un problème NP-difficile et non convexe.
- Si $p = 1$, la solution reste parcimonieuse et le problème devient convexe; on peut le résoudre avec des algorithmes efficaces.
- Si $p = 2$, le problème est convexe et l'on a à disposition toute la machinerie des espaces de Hilbert et des projections orthogonales, mais la norme 2 ne préserve plus la parcimonie de la solution.

5°. Décrire ou bien l'algorithme OMP ou bien l'algorithme IHT en pseudo code et préciser quel en est le principe. Quelles sont les conditions pour qu'il converge ? Quel est le critère d'arrêt ?

IHT est un algorithme de seuillage itératif. À chaque itération, il estime le vecteur x à partir de y et de M (en effectuant des produits scalaires), puis dans le vecteur obtenu, ne garde que les k coordonnées les plus élevées afin de conserver une solution parcimonieuse. Si la matrice possède la propriété RIP (précisez k et ϵ), on peut montrer qu'à chaque itération, la norme mesurant l'écart entre le vecteur initial x et son estimation décroit strictement. Le théorème du point fixe assure alors que l'algorithme converge vers la bonne solution. Le nombre d'itérations n'est pas fixe. Il doit être estimé en fonction de l'erreur que l'on tolère entre le vecteur initial et le vecteur estimé à t.

OMP est un algorithme itératif qui reconstruit le support du vecteur coordonnée par coordonée et s'arrête après k itérations. À la première itération, l'algorithme sélectionne l'indice de la colonne la plus corrélée (en termes de produits scalaires) au vecteur observé y et décide que cet indice appartient au bon support. Il estime alors le vecteur x en projetant y sur l'espace engendré par les colonnes déjà sélectionnés. Ce projeté constitue le vecteur estimé, et le projeté sur l'orthogonal de cet espace constitue le reste r. À chaque suivante, on recommence la même opération, mais sur le reste r au lieu de y.

J'ai décrit les algorithmes en bon français et pas en pseuso code (qui est disponible dans les diapos ou les notes de lecture) afin de bien résumer les étapes importantes.

6°. Les matrices gaussiennes possèdent la propriété RIP avec grande probabilité ; donner quelques idées essentielles et les grandes lignes de la démonstration.

Commencer par énoncer la propriété RIP, puis lister les étapes de la démo (en donnant quelques équations) :
1. Pour un vecteur x fixé, une matrice gaussienne aléatoire M (qui dépend de x) vérifie des inégalités de concentration : la norme du vecteur Mx ($||Mx||^2$) ne s'éloigne que très peu de sa moyenne $||x||^2$. Quand m grandit, la probabilité que l'écart soit plus grand qu'une valeur ϵ décroit à vitesse exponentielle avec m et ϵ. On le prouve avec l'inégalité de Markov et celle de Chernoff.
2. Pour un nombre fini de vecteurs, le lemme de Johnson Lindenstrauss (JL) assure ensuite, par un argument de borne de l'union, que la probabilité précédente tend rapidement vers 1, de façon uniforme pour tous les vecteurs de l'ensemble (contrairement à l'inégalité de concentration, une seule matrice M vérifie l'inégalité pour tous les vecteurs x de l'ensemble. Donc, M ne dépend pas de x).

3. Par linéarité, la démonstration peut être faire sur des vecteurs de norme 1 ou des vecteurs quelconques, sans perte de généralité.
4. Par un argument combinatoire de recouvrement sur la sphere de $\mathbb{R}^n$, le lemme de JL reste valable pour un ensemble de vecteurs parcimonieux de la sphère, même si cet ensemble est infini.
5. On retrouve alors exactement la propriété RIP par une dernière utilisation de la borne de l'union en comptant tous les ensembles k parcimonieux de la sphère.

7°. Que signifient les formules $m \geq Ck \ln(n/k)$, $m \geq 2k$, $m \geq k+1$? Dans quels cas s'appliquent-elles ?

- Pour reconstruire un vecteur k-parcimonieux x fixé, c'est à dire donné de façon non uniforme (la matrice M doit être trouvée pour chaque x), il suffit de $m = k+1$ mesures. Les méthodes de résolution sont alors algébriques (on résoud des systèmes), mais leur complexité (le problème est NP-difficile) et leur non robustesse au bruit les rendent rapidement inutiles.
- Pour reconstruire de façon uniforme tout vecteur k-parcimonieux x (c'est à dire que la matrice M ne dépend pas de x), il suffit de $m = 2k$ mesures, mais là encore les méthodes de résolution sont complexes et non robustes au bruit ou à une parcimonie approchée.
- $m \geq Ck \ln(n/k)$ est l'ordre de grandeur du nombre de mesures qu'il faut pour reconstruire tout vecteur k parcimonieux par des méthodes efficaces, de façon stable et robuste au bruit. Cet ordre de grandeur est atteint par des algorithmes comme par exemple BP ou OMP.

8°. Garanties de reconstruction: définir RIP, spark, NSP, ERC et cohérence mutuelle. À quoi servent-ils ? Donner un exemple de garantie sur le nombre de mesures.

- La propriété spark est une condition nécessaire et suffisante de reconstruction exacte de tous les vecteurs k-parcimonieux solution du problème P_0. Elle est complexe à déterminer (le problème est NP-difficile) donc non applicable dès la dimension est grande. Elle permet de reconstruire des vecteurs avec un nombre de mesures de l'ordre de $m \geq 2k$.
- La propriété du noyau d'ordre k (NSP pour null space property) est une condition nécessaire et suffisante de reconstruction exacte de tous les vecteurs k parcimonieux du problème P_1 par des algorithmes de minimisation de la norme $||.||_1$. C'est donc une propriété uniforme, mais difficile à démontrer, donc peu pratique. Elle est vérifiée par une matrice si son noyau ne contient pas de vecteurs parcimonieux (ils doivent avoir un poids d'au moins $2k$).
- La propriété RIP est une condition nécessaire et suffisante de reconstruction, uniforme, qui permet en un nombre de mesures de l'ordre de $m \geq Ck \ln(n/k)$ de reconstruire tout vecteur k parcimonieux à l'aide de nombreux algorithmes (BP, IHT, OMP, etc.). C'est une propriété uniforme et il facile de trouver une matrice (aléatoirement) la vérifiant.
- La cohérence mutuelle μ mesure l'écart angulaire maximal entre toutes les paires de vecteurs colonnes de M (une fois normalisés). Lorsque la cohérence est suffisamment petite, elle donne une condition suffisante pour reconstruire tout vecteur k-parcimonieux avec plusieurs algorithmes (dont OMP). Points positifs : c'est une condition uniforme et relativement facile à vérifier. Points négatifs : c'est une condition juste suffisante (et pas nécessaire) et surtout elle nécessite un nombre de mesures de l'ordre de $m \geq Ck^2$ au lieu de $m \geq Ck \ln(n/k)$.
- La condition ERC est une condition nécessaire et suffisante sur les coefficients de la matrice M, qui assure que l'algorithme OMP reconstruit de façon unique tout vecteur k-parcimonieux: $\max_{i \in \overline{S}} ||M_S^\dagger M_i|| < 1$.