

ÉCOLE INTERNATIONALE DES SCIENCES DU TRAITEMENT DE L'INFORMATION

COMPRESSIVE SENSING

Mini projet : ***Compressive sensing et données manquantes***



Groupe :
Thibault LEDOUX
Pierre SOULIER
Arthur SAVIO

Prof référent :
Nisrine FORTIN

ING2 MI - Pau

Le 22 mars 2019

Table des matières

1	Introduction	2
2	Codage parcimonieux	3
3	Procédé du compressive sensing	5
4	Données manquantes	8
5	Conclusion	9

1 Introduction

Dans le cadre du module de *Compressive Sensing*, nous avons réalisé ce mini projet concernant les données manquantes. Nous nous sommes intéressés aux données de production des puits en gaz, la production d'huile et la production en eau. Ainsi, nous avons travaillé sur une matrice $N * l$, avec $N = 99$ et $l = 108$.

Ce mini projet possède ainsi trois phases :

- Développer des méthodes de codages parcimoniaux : MP, OMP, StOMP, CoSaMP, IRLS.
- Générer des matrices de mesures et valider vos choix de nombre de mesures
- Appliquer le procédé de compressive sensing afin de compléter des données manquantes

Ainsi les objectifs de ce projet sont d'une première part la validation des méthodes de codage parcimonieux et dans un second temps la reconstruction d'un signal (à partir du dictionnaire appris) dont des données sont manquantes à cause d'une défaillance des capteurs de mesure des données dans un puits pétrolier en utilisant la méthode de codage parcimonieux validée.

Dans ce rapport, nous suivrons donc les trois phases évoquées précédemment dans l'ordre établi des questions.

2 Codage parcimonieux

Question 1

La première question reposait sur deux points. Le premier point était de déduire une valeur pour l'ordre de parcimonie des vecteurs d'apprentissage à partir de la méthode d'apprentissage du dictionnaire par KSVD. Le second point était d'implémenter la méthode CoSaMP. Ainsi on retrouve ces deux aspects dans le code Scilab fourni en complément de ce rapport. A noter que nous avons trouvé un ordre de parcimonie s égale à 10.

Question 2

Montrons que les solutions des problèmes suivants sont équivalentes :

$$\begin{cases} (P_p) : \min \|\alpha\|_p & x = D\alpha \\ (P_2) : \min \|W\alpha\|_2 & x = D\alpha \end{cases}$$

On considère d'abord que W est une matrice quelconque c'est-à-dire :

$$W = (w_{i,j})_{i,j \in \{1,2,\dots,K\}} \text{ où } w_{i,j} \text{ est réel.}$$

En développant $\|W\alpha\|_2$ on obtient :

$$\|W\alpha\|_2 = (\sum_{i=1}^K (\sum_{j=1}^K w_{ij} \times \alpha_j)^2)^{1/2}$$

$$\text{On pose } w_{i,j} = 0 \text{ si } i \neq j \text{ et } w_{i,j} = |\alpha_i|^{p/2-1} \text{ si } i=j$$

$$\|W\alpha\|_2 = (\sum_{i=1}^K (|\alpha_i|^{p/2-1} \times \alpha_i)^2)^{1/2}$$

$$\|W\alpha\|_2 = (\sum_{i=1}^K (|\alpha_i|^{p/2})^2)^{1/2}$$

$$\|W\alpha\|_2 = (\sum_{i=1}^K |\alpha_i|^p)^{1/2}$$

$$\text{Or on sait que } \|\alpha\|_p = (\sum_{i=1}^K |\alpha_i|^p)^{1/p}$$

Pour résoudre (P_p) on cherche le point qui minimise $(\sum_{i=1}^K |\alpha_i|^p)^{1/p}$ or par bijection de $x \rightarrow x^y$ (y réel strictement positif fixé), il s'agit du même point qui minimise $(\sum_{i=1}^K |\alpha_i|^p)^{1/2}$. Donc (P_p) et (P_2) ont la même solution.

Question 3

Il s'agissait ici d'établir les conditions d'existence de solution du problème précédent (Question 2). Au moment de poser $w_i = |\alpha_i|^{p/2-1}$ il faut s'assurer que $\alpha_i \neq 0$ car $p/2 - 1 < 0$.

Question 4

On souhaite approcher x de la manière suivante : $x \simeq D\alpha$. On pose $\beta = W\alpha$, d'où $\alpha = W^{-1}\beta$. On obtient $x \simeq DW^{-1}\beta$ donc on cherche à minimiser $\|x - DW^{-1}\beta\|$:

$$\begin{aligned} \|x - DW^{-1}\beta\|^2 &= \|x\|_2^2 - 2\langle x, DW^{-1}\beta \rangle + \|DW^{-1}\beta\|_2^2 \\ \|x - DW^{-1}\beta\|^2 &= \|x\|_2^2 - 2x^T DW^{-1}\beta + \beta^T (DW^{-1})^T DW^{-1}\beta \end{aligned}$$

On identifie une forme quadratique de la forme $c + b^T\beta + \frac{1}{2}\beta^T A\beta$ avec :

- $c = \|x\|_2^2$
- $b^T = -2x^T DW^{-1}$
- $A = 2(DW^{-1})^T DW^{-1}$

Pour identifier β on résout $A\beta + b = 0$:

$$(DW^{-1})^T DW^{-1}\beta + (-x^T DW^{-1})^T = 0$$

On pose $\Gamma = DW^{-1}$ On a donc :

$$(\Gamma)^T \Gamma \beta + (-x^T \Gamma)^T = 0$$

Par construction, $\Gamma^T \Gamma \in S_n(R)$. Montrons que $\Gamma^T \Gamma$ est définie positive.

Or on a : $x^T \Gamma^T \Gamma x = \|\Gamma x\|_2^2$.

Ainsi $\Gamma^T \Gamma$ est définie positive.

De plus, la matrice $\Gamma^T \Gamma \in M_n(R)$ et par densité de $GL_n(R)$ dans $M_n(R)$, la probabilité qu'elle ne soit pas inversible est "très" faible.

Donc on peut exprimer β de la manière suivante : $\beta = (\Gamma^T \Gamma)^{-1} \Gamma^T x$.

α s'exprime donc de la manière suivante : $\alpha^{(k)} = W^{-1}\beta = W^{-1}(\Gamma^T \Gamma)^{-1} \Gamma^T x$.

3 Procédé du compressive sensing

Question 6

Pour rappel, nous utilisons les cinq matrices de mesures suivantes :

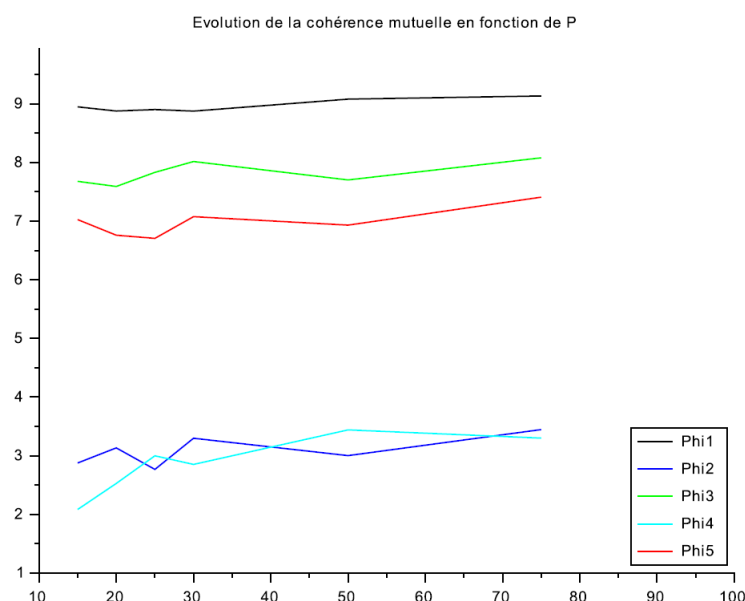
- ϕ_1 : la matrice aléatoire générée à partir d'un processus uniformément distribué.
- ϕ_2 : la matrice aléatoire générée à partir d'un processus bernoullien $\{-1, 1\}$.
- ϕ_3 : la matrice aléatoire générée à partir d'un processus bernoullien $\{0, 1\}$.
- ϕ_4 : la matrice aléatoire générée à partir d'un processus gaussien identique et indépendamment distribué (i.i.d) avec une moyenne nulle et une variance $\frac{1}{M} : \mathcal{N}(0, \frac{1}{M})$.
- ϕ_5 : la matrice creuse ou parcimonieuse générée de façon aléatoire.

Le calcul des vecteurs de mesures en utilisant ces 5 matrices de mesure est sur le script Scilab.

Par la suite, nous avons calculé la cohérence mutuelle entre le dictionnaire D et ces différentes matrices de mesures. Nos résultats sont présentés dans le tableau ci-dessous.

Cohérence mutuelle	$P = 15$	$P = 20$	$P = 25$	$P = 30$	$P = 50$	$P = 75$
ϕ_1	8.93	8.94	8.95	8.97	9.00	9.02
ϕ_2	2.85	2.97	3.03	3.08	3.22	3.35
ϕ_3	7.71	7.76	7.79	7.81	7.88	7.93
ϕ_4	2.89	2.99	3.07	3.12	3.25	3.37
ϕ_5	6.87	6.97	6.94	6.97	7.03	7.08

Pour être plus précis, ces valeurs ont été moyennées à partir d'un échantillon de 1000 valeurs à chaque calcul. Par ailleurs, nous avons réalisé un affichage graphique de cette cohérence mutuelle en fonction des différents pourcentages de mesures P , sur un autre cas que précédemment, pour illustrer le tableau.



Question 7 - Analyse des résultats

La cohérence mutuelle mesure la corrélation maximale entre les vecteurs lignes des différentes ϕ et les vecteurs colonnes de D . Ainsi, ici, on remarque que ce sont les matrices ϕ_2 et ϕ_4 qui présentent la meilleure corrélation. Donc, pour la suite du projet, nous retiendrons ces deux matrices.

Question 8 - Reconstruction

Nous arrivons au moment de la reconstruction des vecteurs d'origine à l'aide du procédé du compressive sensing. Pour cela, nous avons utilisé les différentes méthodes de codage parcimonieux : Matching Pursuit, Orthogonal Matching Pursuit, Stagewise Orthogonal Matching Pursuit, Compressive Sampling du Matching Pursuit et Iteratively Reweighted Least Squares. Pour la reconstruction, nous n'avons considéré que les trois premiers vecteurs d'apprentissage et les deux matrices de mesures retenues précédemment.

Le tableau suivant présente les erreurs relatives de reconstruction pour la matrice de mesures ϕ_2 associée au vecteur d'apprentissage X_3 . Les tableaux pour X_1 et X_2 et avec la matrice de mesures ϕ_4 sont sensiblement les mêmes.

Erreurs relatives	P = 15	P = 20	P = 25	P = 30	P = 50	P = 75
MP	0.027	0.028	0.020	0.018	0.013	0.014
OMP	0.019	0.018	0.013	0.009	0.003	0.003
StOMP	0.026	0.031	0.027	0.032	0.026	0.026
CoSaMP	0.101	0.030	0.045	0.011	0.006	0.004
IRLS	0.076	0.139	0.148	0.075	0.002	0.00001

Par ailleurs, nous avons également tracé un graphique pour les différentes erreurs relatives en fonction de chaque méthode et de chaque pourcentage de mesure P . On peut ainsi le retrouver ci-dessous.



Question 9 - Analyse des résultats

Avant de rentrer dans les détails, il faut noter que les résultats affichés sont très aléatoires en fonction des choix faits. L'erreur relative trouvée pour chaque méthode nous montre qu'elles sont plutôt assez fiables. Les méthodes CoSaMP et IRLS sont les plus aléatoires. StOMP reste la plupart du temps stable quelque soit le pourcentage P . Ainsi, les méthodes les plus optimales sont MP et surtout l'OMP. On retiendra alors cette dernière pour la partie des données manquantes.

4 Données manquantes

Question 10

Dans cette dernière partie, nous allons reconstruire à l'aide de deux méthodes différentes le signal à partir d'un vecteur d'origine présentant des données manquantes par exemple à cause d'une défaillance des capteurs de mesure de données.

La première méthode consiste à prendre en compte les données absentes. Dans notre cas, nous avons sélectionné aléatoirement un vecteur de notre signal d'origine puis nous avons supprimé un certain nombre de valeur pour obtenir un vecteur Zr de taille $N - e$ avec $N = 99$ la taille du signal d'origine et e le nombre d'éléments manquants. Plus précisément, $Zr = \Phi * X^r$, avec Φ une matrice de 0 et de 1 lorsqu'il y a un élément manquant puis X^r un vecteur pris aléatoirement du signal d'origine.

La seconde méthode consiste à ne pas prendre en compte les données absentes. Pour cela, nous avons une nouvelle fois sélectionné aléatoirement un vecteur de notre signal d'origine puis nous avons cette fois-ci mis à zéro un certain nombre de valeur pour obtenir un vecteur $Z0$ de même taille que le signal d'origine. Pour ce cas, nous avons utilisé matrice de mesure ϕ_4 avec $P = 100$ pour ne pas avoir de problèmes de dimension dans les calculs.

Nous avons ainsi fait varier le nombre d'éléments manquants pour chaque méthode. En utilisant les mêmes méthodes que précédemment pour le compressive sensing et en utilisant la méthode de l'OMP, nous avons trouvé les résultats suivants :

Éléments manquants	2	5	10	20	30	50	75
Erreurs pour la méthode 1 : Zr	0.351	0.040	0.125	0.015	0.006	0.117	0.040
Erreurs pour la méthode 2 : $Z0$	0.499	0.315	0.305	0.323	0.319	0.371	0.262

Ainsi, on remarque que les erreurs pour la première méthode sont plus fiables que celles pour la méthode 2. De même, pour la méthode 1, on remarque également que la précision est plus grande moins il y a d'éléments manquants, ce qui est logique. Pour la méthode 2, peu importe le nombre de zéros, l'erreur est assez forte. Il faut savoir que les résultats sont très changeants d'un vecteur à un autre et donc qu'il est tout de même plutôt difficile de faire une conclusion générale pour ces deux méthodes. On peut toutefois dire que ces méthodes permettent quand même de reconstruire le signal correctement.

5 Conclusion

Pour conclure, ce mini projet nous a permis de mettre en application les méthodes de codage parcimonieux vues en cours. Nous avons pu comparer ces différentes méthodes sur l'exemple concret du puits pétrolier. Cela nous a également permis de mettre en corrélation tous les éléments du cours de *Compressive Sensing*. Enfin, dans un dernier temps, nous avons également pu nous confronter au cas où des données peuvent être manquantes à cause de défaillances et ainsi utiliser des méthodes pour contrer le problème et reconstruire le signal avec la plus petite erreur possible.