

E.I.S.T.I - Département Mathématiques
2^{ème} année Ingénieurs - Mathématiques & Informatique

Compressive Sensing
Apprentissage d'un dictionnaire par K-SVD – TP noté

TD & TP - 2018/2019 – Prof : Nisrine Fortin

Les données d'apprentissage que vous utiliserez pour l'apprentissage d'un dictionnaire sont déposées sur Arel. Elles sont présentées sous le format ".dat" : data.dat pour la matrice de vecteurs d'apprentissage. Pour utiliser ces données sur Scilab, vous pourrez utiliser 'load("chemin/fichier.dat")'.

Exercice 1 (APPRENTISSAGE D'UN DICTIONNAIRE). Les données¹ représentées dans la matrice "data.dat" correspondent à la production des puits en gaz, la production d'huile et la production en eau. Cette matrice est de taille $\mathbf{N} \times \ell$, $\mathbf{N} = 99$ et $\ell = 108$.

Les colonnes \mathbf{X}_i , $i = 1, \dots, \ell$ représentent des vecteurs d'apprentissage et nous permettrons d'apprendre un dictionnaire $\mathbf{D} \in \mathcal{M}_{\mathbf{N} \times \mathbf{K}}(\mathbb{R})$ dans lequel les signaux \mathbf{X}_i sont parcimonieux.

Utiliser la méthode K-SVD pour apprendre un dictionnaire $\mathbf{D} \in \mathcal{M}_{\mathbf{N} \times \mathbf{K}}(\mathbb{R})$ concernant nos données "data.sci". ($\mathbf{N} = 99$, $\mathbf{K} = 100$, $\mathbf{L} = 10$ et $\epsilon = 10^{-6}$).

Exercice 2 (STAGEWISE ORTHOGONAL MATCHING PURSUIT). On se propose d'améliorer la méthode OMP en utilisant une stratégie de sélection de plusieurs atomes. La méthode de Stagewise Orthogonal Matching Pursuit (StOMP) sélectionne plusieurs atomes (au lieu d'un seul atome pour l'OMP – annexe TD-TP Semaine 2-Codage parcimonieux) dépassant un seuil calculé à chaque itération.

L'algorithme StOMP (ci-dessous) diffère de l'algorithme OMP à l'étape de la sélection d'atomes (étape 2) et à l'étape de la mise à jour (étape 3), le StOMP sélectionne plusieurs atomes alors que l'OMP sélectionne un seul atome (étape 3), de plus le StOMP remet à jour l'ensemble des indices actifs en ignorant les atomes sélectionnés pour lesquels une composante de α est déjà nulle (étape 7).

L'algorithme StOMP procède selon les étapes suivantes :

1. On calcule la contribution de tous les atomes \mathbf{d}_j , $j = 1, \dots, \mathbf{K}$ au résiduel courant,

$$\mathbf{C}_j = \frac{|\mathbf{d}_j^t \mathbf{R}^{(k-1)}|}{\|\mathbf{d}_j\|_2}.$$

1. Pour des raisons de confidentialité, les données ne sont pas présentées avec plus de détails

2. La sélection :

- (a) On calcule le seuillage $\mathbf{S}^{(k)}$: $\mathbf{S}^{(k)} = \mathbf{t} \frac{\|R^{(k-1)}\|_2}{\sqrt{\mathbf{K}}}$, $2 \leq \mathbf{t} \leq 3$.
- (b) On sélectionne l'ensemble Λ_k des indices des atomes dont la contribution est supérieure au seuillage $\mathbf{S}^{(k)}$

$$\Lambda_k = \left\{ j \in \{1, \dots, \mathbf{K}\}, \quad \mathbf{C}_j > \mathbf{S}^{(k)} \right\}.$$

3. **La mise à jour :** On met à jour l'ensemble des indices $\mathbf{P}_k = \mathbf{P}_{k-1} \cup \Lambda_k$.

4. On construit la matrice Φ_k en considérant les colonnes \mathbf{d}_p , $p \in \mathbf{P}_k$.

$$\Phi_k = [\mathbf{d}_p, p \in \mathbf{P}_k].$$

5. On résout le problème d'optimisation $\alpha_{\mathbf{P}_k}^{(k)} = \arg \min \|x - \Phi_k \alpha_{\mathbf{P}_k}\|_2$ par la méthode des moindres carrées.
6. On met à jour le résiduel, $R^{(k)} = x - \mathbf{D}\alpha$.
7. On remet à jour l'ensemble des indices :

$$\mathbf{P}_k = \left\{ \text{des indices } j \text{ pour lesquels } \alpha^{(k)}(j) \neq 0 \right\}.$$

Le StOMP s'initialise et s'arrête suivant les mêmes critères que l'OMP.

1. **Coder l'algorithme du Stagewise Orthogonal Matching Pursuit.**
2. Apprendre un dictionnaire adapté à l'exercice 1, en utilisant le StOMP comme méthode de codage parcimonieux.