Optimisation non-linéaire - Projet

10 Décembre 2018

Partie 1 : étude du problème NNLS (non-negative least squares).
En optimisation, le problème des moindres carrés non-négatifs (NNLS) est un problème de type *moindres carrés* où les variables ne sont pas autorisées à devenir négatives. Formellement, étant donné une matrice $A \in \mathbb{R}^{m \times n}$ et un vecteur colonne $b \in \mathbb{R}^m$, l'objectif est de trouver le vecteur $x \in \mathbb{R}^n$ minimisant la quantité suivante :

$$\min_x \|Ax - b\|_2^2 \quad \text{tel que } x \geq 0.$$

Le but de cette première partie est d'étudier ce problème à travers les questions qui suivent.

(a) Ecrivez les conditions d'optimalité du premier ordre. Sont-elles nécessaires, suffisantes, nécessaires et suffisantes ? Illustrez ces conditions à l'aide d'un exemple de petite taille ($m = 6$ et $n = 4$).

(b) Etant donné un vecteur $x \in \mathbb{R}^n$, quelle valeur faut-il donner au paramètre $\gamma \in \mathbb{R}$ afin de minimiser la quantité $\|A(\gamma x) - b\|_2^2$?

(c) Implémentez une fonction `function [x,e,t]=nnls_nomequipe(A,b,x0,timelimit,choix)` permettant d'appliquer la méthode du gradient (lorsque `choix=1`), la méthode du gradient accéléré 1 (lorsque `choix=2`) et la méthode de Coordinate Descent (lorsque `choix=3`).

- Implémentez soigneusement les différentes méthodes afin de pouvoir les comparer
- La méthode du gradient accéléré et la méthode Coordinate Descent possèdent plusieurs degrés de liberté (paramètres, ordre des variables, etc). A vous d'effectuer un choix afin de construire la méthode la plus rapide possible.
- Testez et commentez les résultats obtenus sur les jeux de données `exemplennls1.mat` et `exemplennls2.mat`.

Partie 2 : factorisation nonnégative de matrices (NMF)
Etant donné une matrice $X \in \mathbb{R}^{m \times n}$ et un naturel $r < \min(m, n)$, le problème NMF consiste à trouver des matrices nonnégatives $W \in \mathbb{R}^{m \times r}$ et $H \in \mathbb{R}^{r \times n}$ minimisant la quantité suivante :

$$\min_{W,H} \|X - WH\|_F^2 \quad \text{tel que } W \geq 0 \text{ et } H \geq 0.$$

(a) Ce problème est-il convexe ? Et si H est fixée ? Et si W est fixée ?

(b) Implémentez une fonction `function [W,H,e,t]= nmf_nomequipe(X,W0,H0,timelimit)` effectuant une optimisation alternée de W et H.

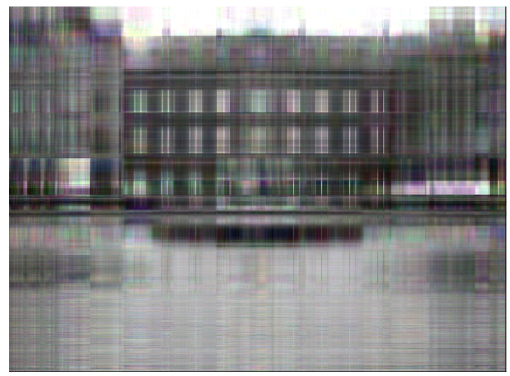
- Si H est fixée, l'optimisation de W consiste en l'optimisation de m problèmes NNLS. Si W est fixée, l'optimisation de H consiste en l'optimisation de n problèmes NNLS.
- A vous d'effectuer les meilleurs choix possibles dans l'implémentation afin de construire la méthode la plus rapide possible.

(c) Pour la matrice X ci-dessous, trouvez à l'aide de votre code le plus petit naturel r tel qu'il existe une factorisation exacte $X = WH$ avec des facteurs $W \in \mathbb{R}^{6\times r}$ et $H \in \mathbb{R}^{r\times 6}$.

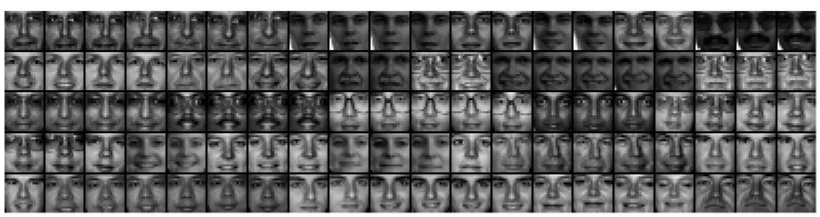
$$X = \begin{pmatrix} 0 & 1 & 2 & 2 & 1 & 0 \\ 0 & 0 & 1 & 2 & 2 & 1 \\ 1 & 0 & 0 & 1 & 2 & 2 \\ 2 & 1 & 0 & 0 & 1 & 2 \\ 2 & 2 & 1 & 0 & 0 & 1 \\ 1 & 2 & 2 & 1 & 0 & 0 \end{pmatrix}$$

(cette matrice est enregistrée dans `exemplenmf1.mat`, utilisez le fichier `RunMeNMF.m`)

(d) Dans le dossier du projet se trouve une image `pl-originale.png`. A l'aide de votre code et en utilisant le fichier `RunMeNMF2.m`, on vous demande de factoriser les trois bandes de l'image en testant différentes valeurs de r afin de réduire la qualité de l'image.



(e) Dans le fichier de données `exemplenmf2.mat`, une série de 2429 d'images de visages (19×19 pixels) sont vectorisées dans une matrice 361×2429. A l'aide de votre code et en utilisant le fichier `RunMeNMF3.m`, on vous demande d'effectuer une factorisation de cette matrice et d'afficher le facteur W pour différentes valeurs de r. A quoi correspond ce facteur W ? Et H ?



(à renvoyer pour le 22 décembre 2018)