

MATH0500 - Introduction à l'algorithmique numérique

Projet de programmation

Dans ce projet, nous allons mettre en pratique les méthodes vues au cours pour traiter des matrices creuses.

Nous allons travailler avec un fichier correspondant aux choix de cours d'étudiants. Notre première tâche sera de transformer le fichier en une matrice dont les lignes et les colonnes sont les cours et les étudiants et les éléments de la matrice valent 1 lorsque l'étudiant suit le cours correspondant.

Nous calculerons ensuite quelques statistiques sur les cours et sur les étudiants en effectuant des opérations matricielles.

Fichier d'entrée

Le fichier d'entrée qui vous est donné est un fichier qui renseigne les choix de cours. Il s'agit d'un fichier csv contenant, pour chaque ligne, le matricule de l'étudiant concerné, et le cours choisi par cet étudiant et enfin le quadrimestre auquel ce cours est donné.

1. Ecrire une fonction C qui prend en entrée le fichier csv et qui donne en sortie une matrice creuse où chaque ligne correspond à un cours, chaque colonne à un étudiant (représenté par son matricule), et chaque élément (i, j) de la matrice vaut 1 si l'étudiant j a suivi le cours i .
2. Ecrire une fonction C qui calcule la transposée de la matrice.

Statistiques simples sur les cours

Nous allons maintenant calculer quelques statistiques simples sur les cours en effectuant des opérations matricielles sur notre matrice creuse.

3. Ecrire une fonction C qui prend en entrée une matrice creuse A et un vecteur dense x et qui renvoie le produit Ax sous forme dense.
4. Utiliser la fonction précédente avec la matrice créée en 1. et un vecteur de 1 afin de calculer un vecteur contenant le nombre d'étudiants dans chaque cours.
5. Ecrire une fonction C qui prend en entrée une matrice creuse A et un vecteur creux x et qui renvoie le produit Ax sous forme creuse.
6. On suppose que tous les matricules impairs correspondent à des filles. Utiliser la fonction précédente pour calculer un vecteur contenant le nombre de filles de chaque cours. Pour chaque cours, on souhaite également calculer le nombre d'étudiants ayant un matricule commençant par 2008. Utiliser similairement la fonction précédente pour ce faire.

Statistiques plus élaborées

Nous allons maintenant utiliser le calcul de valeurs propres afin de détecter des groupes d'étudiants les plus homogènes possibles.

7. Ecrire une fonction `C` qui calcule le produit de 2 matrices creuses et renvoie une matrice creuse.
 8. Ecrire une fonction `C` qui applique la méthode de la puissance à une matrice A et renvoie la valeur propre de plus grand module ainsi que le vecteur propre (creux) associé.
 9. Appelons A la matrice calculée en 1. Calculer le produit $A^T A$ qui est une matrice dont l'élément i, j renseigne le nombre de cours communs entre l'étudiant i et l'étudiant j . De même, calculer la matrice AA^T dont chaque élément i, j renseigne le nombre d'étudiants en commun entre le cours i et le cours j .
 10. Calculer le vecteur propre associé à la valeur propre de plus grand module de $A^T A$. Le vecteur propre peut être vu comme un cluster d'étudiants le plus homogène possible, correspondant au plus grand groupe. Faites de même pour AA^T . Interpréter les résultats.
- Bonus** Calculer le 2e cluster en calculant le vecteur propre associé à la valeur propre de 2e plus grand module.

Consignes

Le projet se réalise par groupes de 2. Vous ne devez pas écrire de rapport . Le code sera par contre testé sur une machine unique. Le critère le plus important concerne l'efficacité de votre code et le fait que le code a été écrit par vous.

Une présentation orale de 10 minutes où votre code sera testé est prévue le vendredi 20 décembre. Pour ce faire, votre code doit être envoyé par mail à q.louveaux@uliege.be sous forme d'archive **zip**. Un ordre de passage sera déterminé dans le courant du mois de décembre.

Le plagiat de codes trouvés sur internet ou de codes d'autres groupes d'étudiants sera systématiquement testé et, le cas échéant, résultera en une note de 0/20 pour le projet.