

# Algèbre Linéaire pour le Data



**Composante**  
École Nationale  
Supérieure  
d'Électrotechnique  
d'Électronique  
d'Informatique  
d'Hydraulique  
et des  
Télécommunications

## En bref

- **Code Ametys:** N8EN03B
- **Ouvert aux étudiants en échange:** Non

## Présentation

### Objectifs

Ce cours vise à fournir aux étudiants les bases essentielles de l'algèbre linéaire pour la modélisation, l'analyse, et la mise en œuvre des méthodes pour le Data Mining. L'objectif est d'être à même de manipuler les transformations linéaires, de comprendre les notions de réduction de dimension, et d'appliquer ces outils à des problèmes divers, qui seront ensuite réalisés dans des séances de travaux pratiques.

### Description

Ce cours commence par rappeler l'ensemble des notions déjà vues en algèbre linéaire numérique, avec notamment les systèmes d'équations linéaires et leur résolution. Certaines extensions sont développées en détail, comme la notion de corrélation canonique entre sous-espaces, les formules de Sherman-Morrisson-Woodbury, ainsi que les notions d'analyse d'erreur a priori et d'analyse d'erreur a posteriori.

Une partie importante du cours est consacrée à la réduction de dimension, en particulier au travers des décompositions matricielles, comme la décomposition en valeurs singulière, ou les méthodes de Krylov basées sur le processus de bidiagonalisation de Lanczos-Golub-Kahan, ou bien encore les méthodes de factorisation non-négative, et leur rôle dans les systèmes de recommandation, la compression de données et le traitement de l'information.

La dernière partie du cours présente diverses notions complémentaires, comme la régularisation de Tikhonov, la validation croisée généralisée, les moindres carrés totaux, les moindres carrés linéaires mixtes, ainsi que l'homogénéisation des données par des techniques de Scaling et leur application à la classification spectrale ou le page-ranking, avec l'exploitation du théorème de Perron-Frobenius en particulier. Les travaux pratiques, qui comptent pour presque la moitié des séances, permettent d'illustrer sur des cas concrets les diverses notions abordées.