

Mathématiques pour l'intelligence artificielle

UE MIA, M2 IAAA, AMU, 2022-2023

Thomas Schatz
Mardi 6 septembre 2022

Introductions

- A propos de moi: maître de conférences à AMU, recherche à l'intersection entre apprentissage non-supervisé de représentations en IA et études du développement de la perception chez le bébé en sciences cognitives
- A propos de vous: nom, prénom, formation en mathématique après le bac, ce que vous envisagez de faire après le M2?

Questions pratiques

- Calendrier des cours sur le site IAAA ([lien](#))
- Questions, etc. mattermost M2 IAAA, canal MIA
- Matériel de cours sur mon site web (<https://thomas.schatz.cogserver.net/teaching/>)
- Evaluation:
 - 2 points: scribe pour une séance
 - 3*6 points: 3 DMs notés (calendrier sur mon site, premier DM en ligne aujourd'hui)
- Questions ?

Pourquoi êtes-vous ici?

Pourquoi êtes-vous ici?

- Quelles mathématiques en apprentissage automatique et pourquoi
 - Apprentissage automatique:
 - apprendre à partir de l'expérience: statistique
 - dans le cas de problèmes requérant des calculs intensifs: informatique
 - Objectif central en apprentissage automatique: **développer des procédures efficaces d'un point de vue à la fois statistique et computationnel** (mais aussi: interprétables, équitables...)
- Points à travailler
 - Statistique (algèbre linéaire, probabilités)
 - Côté informatique: optimisation continue (algèbre linéaire)
 - Interactions statistique/informatique

Plan du cours (prévisionnel)

9 séances de 3h

Partie 1 (DM1) : algèbre linéaire et probabilités

1. Notions de bases sur les preuves (+ Algèbre linéaire?)
2. Algèbre linéaire (+ Probabilités?)
3. Probabilités

Partie 2 (DM2): statistique et optimisation

4. Statistiques
5. Optimisation

Partie 3 (DM3):

6. Optimisation sous contraintes
7. Optimisation stochastique
8. Théorie de l'apprentissage

9. Putting it all together

Plan du cours (prévisionnel)

9 séances de 3h

Partie 1 (DM1) : algèbre linéaire et probabilités

1. Notions de bases sur les preuves (+ Algèbre linéaire?)

2. Algèbre linéaire (+ Probabilités?)

3. Probabilités

Partie 2 (DM2): statistique et optimisation

4. Statistiques

5. Optimisation

Partie 3 (DM3):

6. Optimisation sous contraintes

7. Optimisation stochastique

8. Théorie de l'apprentissage

9. Putting it all together

Séance 1: Notions de bases sur les preuves (+ Algèbre linéaire?)

Preuves mathématiques

Méthode générale:

1. Formalisation des hypothèses et du résultat à prouver
 - On utilise les définitions formelles des notions mathématiques impliquées
 - On donne des noms explicites aux variables impliquées
2. Montrer que les hypothèses impliquent (logiquement) le résultat
 - On utilise uniquement les théorèmes connus (et les axiomes de la théorie dans laquelle on travaille, par exemple les axiomes usuels de la théorie des ensembles) et les règles du raisonnement logique

Exemple: prouver que la somme de deux nombres impairs est paire

Différents types de preuves

1. Prouver que la somme de deux nombre entiers impairs est paire.
2. Prouver que pour tout nombre entier n , $n^2 + n$ est pair.
3. Montrer que pour tout entier n , $2^{3n+1} + 5$ est un multiple de 7.
4. Soit a un nombre rationnel et b un nombre irrationnel.
Prouver que $a + b$ est irrationnel.
5. Montrer que si n est un nombre entier et n^2 est impair, alors n est impair.

Différents types de preuves

1. Prouver que la somme de deux nombre entiers impairs est paire.

Preuve simple

2. Prouver que pour tout nombre entier n , $n^2 + n$ est pair.

Preuve par disjonction des cas

3. Montrer que pour tout entier n , $2^{3n+1} + 5$ est un multiple de 7.

Preuve par récurrence

4. Soit a un nombre rationnel et b un nombre irrationnel.

Prouver que $a + b$ est irrationnel.

Preuve par l'absurde

5. Montrer que si n est un nombre entier et n^2 est impair, alors n est impair.

Preuve par contraposition

Différents types de preuves

- Nombre limité de types de raisonnement
- Chaînés pour construire des preuves complexes
- Existence de motifs de raisonnement typiques en fonction du domaine (entiers, ensembles, algèbre linéaire, analyse, probabilités, etc.)

Preuves mathématiques

- Pour aller plus loin:
 - Devlin, K. J. (2012). *Introduction to mathematical thinking*.
 - Lien vers pdf et plus de références sur la page du cours

Algèbre linéaire

Algèbre linéaire

1. Espaces vectoriels et fonctions linéaires
2. Matrices
3. Angles et orthogonalité
4. Structure des applications linéaires et matrices
5. Espaces de matrices et normes
6. Algèbre linéaire numérique

Algèbre linéaire

1. Espaces vectoriels et fonctions linéaires
2. Matrices
3. Angles et orthogonalité
4. Structure des applications linéaires et matrices
5. Espaces de matrices et normes
6. Algèbre linéaire numérique

Objets mathématiques

- **Types d'objets en mathématiques** nombres entiers, fractionnels, réels, complexes, fonctions, ensembles, distributions de probabilité...
- **Comme en informatique mais souvent plus idéalisé** (par exemple nombre réels vs nombre IEEE floating point)
- **Qu'est-ce qu'une fonction en mathématiques ?**

Fonction en mathématiques

- Qu'est-ce qu'une fonction en mathématiques ?

$$f : \left| \begin{array}{ll} E & \rightarrow F \\ x & \mapsto f(x) \end{array} \right.$$

- Exemples

$$+ : \left| \begin{array}{ll} \mathbf{R}^2 & \rightarrow \mathbf{R} \\ x, y & \mapsto x + y \end{array} \right. \quad \exp : \left| \begin{array}{ll} \mathbf{R} & \rightarrow \mathbf{R} \\ x & \mapsto e^x \end{array} \right.$$

$$\nabla : \left| \begin{array}{ll} \mathcal{C}^1(\mathbf{R}^n, \mathbf{R}) & \rightarrow \mathcal{C}^0(\mathbf{R}^n, \mathbf{R}^n) \\ f & \mapsto \nabla f \end{array} \right.$$

Fonction linéaire

- Objet central de l'algèbre linéaire (du point de vue conceptuel)
- Idée générale: fonction qui préserve les combinaisons linéaires
- Définition formelle

$$f : \left\{ \begin{array}{ll} E & \rightarrow F \\ x & \mapsto f(x) \end{array} \right. \quad \text{est une fonction linéaire si et seulement si}$$

E et F sont des espaces vectoriels sur un même corps K et pour toute paire de scalaires $(\alpha, \beta) \in K^2$ et toute paire de vecteurs $(u, v) \in E^2$,

$$f(\alpha u + \beta v) = \alpha f(u) + \beta f(v)$$

Espace vectoriel

Définition

$$(E, K, +, \cdot) \quad + : \left| \begin{array}{ccc} E^2 & \rightarrow & E \\ x, y & \mapsto & x + y \end{array} \right| \cdot : \left| \begin{array}{ccc} K \times E & \rightarrow & E \\ \alpha, x & \mapsto & \alpha \cdot x \end{array} \right|$$

est un espace vectoriel, si et seulement si K est un corps et

Il existe $0_E \in E$, tel que

Pour tout $(u, v, w) \in E^3$,

Pour tout $(\alpha, \beta) \in K^2$

Il existe $u^{-1} \in E$, tel que

$$u + (v + w) = (u + v) + w \quad \text{associativité}$$

$$u + v = v + u \quad \text{commutativité}$$

$$u + 0_E = u \quad \text{élément neutre}$$

$$u + u^{-1} = 0_E \quad \text{inverse}$$

$$\alpha \cdot (\beta \cdot u) = (\alpha \beta) \cdot u \quad \text{compatibilité}$$

$$(\alpha + \beta)u = \alpha \cdot u + \beta \cdot u \quad \text{distributivité}$$

$$\alpha(u + v) = \alpha \cdot u + \alpha \cdot v \quad \text{distributivité}$$

$$1_K \cdot u = u \quad \text{identité multiplicative}$$

Espace vectoriel

Exemples

- espaces Euclidiens $\mathbf{R}^n$
- espaces Hermitiens $\mathbf{C}^n$
- espace des fonctions continues sur $[a, b] \subset \mathbf{R}$
à valeurs réelles $\mathcal{C}^0([a, b], \mathbf{R})$

Fonction linéaire

- Objet central de l'algèbre linéaire (du point de vue conceptuel)
- Idée générale: fonction qui préserve les combinaisons linéaires
- Définition formelle

$$f : \left\{ \begin{array}{ll} E & \rightarrow F \\ x & \mapsto f(x) \end{array} \right. \quad \text{est une fonction linéaire si et seulement si}$$

E et F sont des espaces vectoriels sur un même corps K et pour toute paire de scalaires $(\alpha, \beta) \in K^2$ et toute paire de vecteurs $(u, v) \in E^2$,

$$f(\alpha u + \beta v) = \alpha f(u) + \beta f(v)$$

Notions élémentaires sur les espaces vectoriels

Définition (Vect)

Espace vectoriel engendré par une famille de vecteurs $(v_i)_{i \in I} \in E^I$:

$$\text{Vect}((v_i)_{i \in I}) := \left\{ \sum_{i \in I} \lambda_i v_i \mid (\lambda_i)_{i \in I} \in K^I, \text{ with finitely many } \lambda_i \text{ different from zero} \right\}$$

Définition (indépendance linéaire)

$v \in E$ est *linéairement indépendant* de $(v_i)_{i \in I} \in E^I$ si et seulement si v n'appartient pas à $\text{Vect}((v_i)_{i \in I})$.

Définition (famille libre)

Une famille de vecteurs $(v_i)_{i \in I} \in E^I$ est dite *libre* si et seulement si pour tout $j \in I$, v_j est linéairement indépendant de $(v_i)_{i \in I \setminus \{j\}}$.

Définition (base)

$(v_i)_{i \in I}$ est une base de E si et seulement si $(v_i)_{i \in I}$ est une famille libre de E telle que $E = \text{Vect}((v_i)_{i \in I})$ (i.e. la famille $(v_i)_{i \in I}$ engendre E).

Définition (dimension finie)

Un espace vectoriel E est dit de dimension finie si et seulement si il existe une base de E de formée d'un nombre fini de vecteurs.

Théorème (dimension)

Si E est un espace vectoriel de dimension finie, il existe un unique entier naturel n , la dimension de E , tel que toute base de E est formée de exactement n vecteurs.

Caractérisation de l'indépendance linéaire

Théorème

Soit $(v_i)_{i \in I} \in E^I$ est une famille de vecteurs de l'espace vectoriel E .

- S'il existe une famille $(\lambda_i)_{i \in I} \in K^I$ de scalaires avec au plus un nombre fini de λ_i différents de 0 et telle que :

1. il existe $i \in I$, tel que $\lambda_i \neq 0$

2. $\sum_{i \in I} \lambda_i v_i = 0$

alors $(v_i)_{i \in I}$ est une famille liée.

- Réciproquement, si pour toute famille $(\lambda_i)_{i \in I} \in K^I$ de scalaires avec au plus un nombre fini de λ_i différents de 0, on a:

$$\left(\sum_{i \in I} \lambda_i v_i = 0 \right) \Rightarrow (\text{pour tout } i \in I, \lambda_i = 0),$$

alors $(v_i)_{i \in I}$ est une famille libre.

Décomposition sur une base

Définition (décomposition sur une base)

Soit une base $V = (v_i)_{i \in I}$ de E et $\Lambda = (\lambda_i)_{i \in I} \in K^I$ une famille de scalaires telle que au plus un nombre fini des λ_i soient différents de 0.

On dit que Λ est une décomposition de $v \in E$ sur la base V si et seulement si $v = \sum_{i \in I} \lambda_i v_i$.

Théorème (existence et unicité de la décomposition)

Etant donné une base $V = (v_i)_{i \in I}$ de E , tout vecteur $v \in E$ peut-être décomposé sur cette base et cette décomposition est unique.

Toute base peut donc servir de système de coordonnées pour E

Algèbre linéaire

1. Espaces vectoriels et fonctions linéaires
2. Matrices
3. Angles et orthogonalité
4. Structure des applications linéaires et matrices
5. Espaces de matrices et normes
6. Algèbre linéaire numérique

Matrice

- Objet central de l'algèbre linéaire (en pratique)
- Idée générale: simplement un tableau de nombres en deux dimensions
- Pourquoi est-ce utile ?

$$M = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1j} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2j} & \dots & a_{2p} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{i1} & a_{i2} & \dots & a_{ij} & \dots & a_{ip} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nj} & \dots & a_{np} \end{pmatrix}$$

Matrice

- Objet central de l'algèbre linéaire (en pratique)
- Idée générale: simplement un tableau de nombres en deux dimensions
- Pourquoi est-ce utile ?
 - Permet de représenter les fonctions linéaires en dimensions finie de manière simple à appréhender et pratique pour les calculs

$$M = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1j} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2j} & \dots & a_{2p} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{i1} & a_{i2} & \dots & a_{ij} & \dots & a_{ip} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nj} & \dots & a_{np} \end{pmatrix}$$

Matrice d'une fonction linéaire

E, F espaces vectoriels de dimension finie

$V = (v_1, \dots, v_p)$ base de E

$W = (w_1, \dots, w_n)$ base de F

$f : E \rightarrow F$, fonction linéaire

$M = M(f, V, W)$ matrice de f de V vers W .

$$M = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1j} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2j} & \dots & a_{2p} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{i1} & a_{i2} & \dots & a_{ij} & \dots & a_{ip} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nj} & \dots & a_{np} \end{pmatrix} \begin{matrix} \leftarrow w_1 \\ \leftarrow w_2 \\ \\ \leftarrow w_i \\ \\ \leftarrow w_n \end{matrix}$$

$$\begin{matrix} \uparrow & \uparrow & & \uparrow & & \uparrow \\ f(v_1) & f(v_2) & & f(v_j) & & f(v_p) \end{matrix}$$