

Algèbre linéaire

Mathématiques pour le TAL 1

Damien Nouvel

Inalco

Écriture des nombres entiers

Écritures historiques (peu régulières)

- **langue naturelle** « trois », « dix-huit », « quatre-vingt-quinze »
- **chiffres romains** « III », « XVIII », « XCV »

Écriture régulière avec des **chiffres**

- **chiffre arabes** en base 10 avec les symboles $\{0\dots9\}$
- **nombres** comme écriture de chiffres $c_1c_2c_3$
- **valeur** $c_1 * 100 + c_2 * 10 + c_3 * 1$
- exemple $142 = 1 * 100 + 4 * 10 + 2 * 1$

Généralisation aux **symboles** avec une **base**

- **base** comme ensemble de b **symboles**
- écriture comme **séquence** de k symboles $s_1s_2\dots s_k$
- valeur $\sum_{i=1\dots k} c_i * b^{(k-i)}$
- bases souvent utilisées
 - **binaires** en base 2 $\{0,1\}$
 - **hexadécimaux** en base 16 de 0 à 9 puis $\{A,B,C,D,E,F\}$

Nombres **négatifs** par ajout du tiret –

Écriture des nombres réels

Écriture d'un **nombre décimal** $n = e, d$

- **partie entière** comme **entier relatif** e
- séparateur de la partie décimale avec « , » en français
 - les anglophones utilisent
 - le point comme séparateur
 - la virgule pour les milliers
- **partie décimale** comme **entier naturel** d
- **valeur** $e + d * 10^{-|d|}$ avec $|d|$ le nombre de chiffres dans d

Écriture d'une **fraction** n/d

- numérateur comme **entier relatif** n
- dénominateur comme **entier naturel** d
- écriture en nombre décimal avec **arrondi** $2/3 = 0,67$

Notation scientifique

Difficultés avec les nombres très grands ou très petits

- difficulté d'**écriture** (beaucoup de symboles)
- appréciation difficile des **ordres de grandeur**

Utilisation de la **notation scientifique** c,dEn

- il n'y a qu'**un seul** chiffre c avant la virgule
- la partie décimale d est un entier naturel
- le symbole E pour la **multiplication par une puissance de 10**
- l'**exposant** n est la puissance de 10 signée (+ ou -)
- **valeur** $c, d * 10^n$
- correspond à la notation scientifique courante $c, d * 10^n$
- souvent utilisés dans les logiciels (Excel, Python, etc.)

Exemples

- $4,5E5 = 450000$
- $3,7E-3 = 0,0037$

Quelques notations mathématiques courantes

Fonctions mathématiques courantes

- **valeur absolue** : $|x|$
- **racine** (carrée) : $\sqrt{x}$
- **factorielle** : $n! = \prod_{x \in \llbracket 1, n \rrbracket} x$
- **exponentielle** : $e^x = \exp(x)$
- **logarithme** : $\log(x)$

Calculs sur des séries de nombres indexés $X = (x_1, x_2, x_3 \dots x_n)$

- **somme** des éléments $\sum_i x_i = x_1 + x_2 + x_3 \dots + x_n$
- **produit** des éléments $\prod_i x_i = x_1 * x_2 * x_3 \dots * x_n$

Exemples

- $4! = 1 * 2 * 3 * 4 = 24$
- $\sum_{k \in \{1, 2, 3\}} k^3 = 1^3 + 2^3 + 3^3 = 1 + 8 + 27 = 36$

Ensembles de nombres

Quelques ensembles de nombres

- $\mathbb{N}$ **entiers naturels** (positifs)
- $\mathbb{Z}$ **entiers relatifs** (entiers positifs et négatifs)
- $\mathbb{D}$ **décimaux**, nombres avec virgules
- $\mathbb{Q}$ **rationnels**, nombres qui peuvent s'écrire comme fractions
- $\mathbb{R}$ **réels** (comme $\sqrt{2}$ ou π)
- $\mathbb{C}$ **complexes** (partie réelle et partie « imaginaire »)

Ensembles de nombres comme hiérarchie $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{D} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$

Intervalles

- notations des bornes entre crochets $x \in [0; 2]$ pour les réels
 - les anglophones utilisent la virgule au lieu du point virgule
- valeur réelle entre les bornes $[0, 2] = \{x \in \mathbb{R} | x \geq 0, x \leq 2\}$
- exclusion des bornes $]0, 2[= \{x \in \mathbb{R} | x > 0, x < 2\}$
- doubles crochets pour les nombres entiers $\llbracket 0, 2 \rrbracket = \{0, 1, 2\}$

Variables

Une **variable** est prend une valeur dans un **domaine**

Domaine des variables

- **continu** (ensembles infinis, souvent intervalles de réels)
- dans des ensembles **dénombrable**
 - **discrètes** (nombres entiers, ordonnés et comparables)
 - **catégories** (ensembles pas forcément ordonné mais fini)
- **nominales** (chaînes de caractères, langages algébriques)

Exemples

- $x \in [0, 20]$
- $k \in \mathbb{N}$
- $c \in \{\textit{politique}, \textit{sport}, \textit{culture}\}$
- $n \in \{\textit{Damien}, \textit{Marie}, \textit{Pierre}, \textit{Paul}, \dots\}$

Les ensembles

Notion formalisée fin XIX^{ème} (Cantor, Hilbert) : « tout rassemblement en une totalité d'objets [...] déterminés et bien différenciés qui seront appelés les éléments »

Principes généraux

- énumération des éléments avec les **accolades** $\{1, 2, 3\}$
- l'ordre n'importe pas et sans répétitions $\{1, 0, 1\} = \{0, 1\}$
- **ensemble vide** $\emptyset$
- **cardinalité** par barres verticales $|\{1, 3, 7\}| = 3$
- peuvent être nommés par des **variables** $E = \{1, 2, 3\}$
- éléments **quelconques**
 - nombres $E = \{1, 2, 3\}$
 - caractères $E = \{a, b, c\}$
 - mots $E = \{\text{chat}, \text{chien}\}$
 - mélange hétérogène $E = \{1, \text{chat}, \text{abs}, \text{£}\}$
- un élément **appartient** à un ensemble (epsilon) $3 \in \{1, 2, 3\}$
- **sous-ensemble** avec $\subset$ (strict) $\{1, 3\} \subset \{1, 2, 3\}$ (avec égalité $\subseteq$)

Définition d'un ensemble

Définition par **extension**

- énumération des éléments $E = \{1, 2, 3\}$
- utilisation (peu formelle) des trois petits points $E = \{1, 2, 3, \dots\}$

Définition par **compréhension** (ou intention)

- permet de **décrire** les éléments sans les énumérer
- description en **langue naturelle** (peu formelle)
 - $E = \{\text{nombre pairs}\}$
 - $E = \{\text{mots de plus de cinq lettres}\}$
- utilisation **critères** avec la barre verticale ou la virgule
 - $E = \{x | x/2 \text{ est entier}\}$
 - $E = \{m, \text{taille}(m) > 5\}$

Opérations sur les ensembles

Union $\cup$ (en logique disjonction « ou » notée $\vee$)

- réunir deux ensembles (toujours sans répétitions)
- $\{3, 1, 4\} \cup \{2, 3, 4, 5\} = \{1, 2, 3, 4, 5\}$

Intersection $\cap$ (en logique conjonction « et » notée $\wedge$)

- éléments communs aux deux ensembles
- $\{3, 1, 4\} \cap \{2, 3, 4, 5\} = \{3, 4\}$

Complémentaire $\bar{E}$ (en logique négation notée $\neg$)

- éléments qui ne sont pas dans un ensemble
- $\overline{\{3, 1, 4\}} = \{0, 2, 5, 6, 7, 8, \dots\}$

Différence $\setminus$

- éléments qui sont dans le premier mais pas dans le second
- $\{3, 1, 4\} \setminus \{2, 3, 4, 5\} = \{1\}$

Propriétés des opérateurs

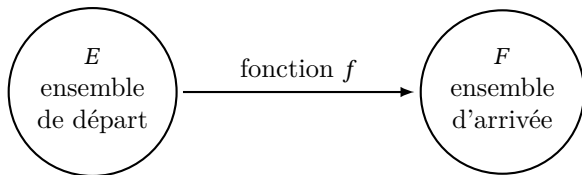
Quelques **propriétés** des opérateurs

- ensembles **disjoints** n'ont pas d'éléments communs $E \cap F = \emptyset$
- **commutativité** $E \cup F = F \cup E$ et $E \cap F = F \cap E$
- **idempotence** $E \cup E = E \cap E = E$
- pour l'ensemble vide
 - $E \cap \emptyset = \emptyset$
 - $E \cup \emptyset = E$
- **distributivité**
 - $E \cap (F \cup G) = (E \cap F) \cup (E \cap G)$
 - $E \cup (F \cap G) = (E \cup F) \cap (E \cup G)$

Cardinalité des ensembles

- $|\emptyset| = 0$
- $|E \cup F| = |E| + |F| - |E \cap F|$

Définition des fonctions



Définition d'une fonction

- **correspondances** entre deux ensembles $f : E \rightarrow F$
 - ensemble de **départ** E (ou domaine)
 - ensemble d'**arrivée** F
- **application** à un **élément** de l'ensemble de départ
 - ensemble de départ en **paramètre** de la fonction $f(x)$
 - résultat **déterminé** par les paramètres $y = f(x)$

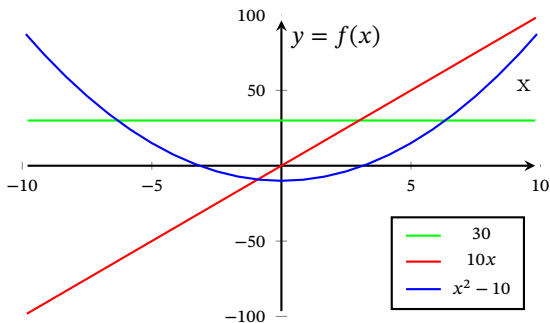
Exemple, multiplication d'un entier naturel par deux

- $f : \mathbb{N} \rightarrow \mathbb{N}$
- $f(x) = 2x$

Tracer une fonction

Déterminer des (x,y) avec

- domaine x en abscisse
- résultat de la fonction $y = f(x)$ en ordonnée



Quelques compléments sur les fonctions

Une fonction peut prendre **plusieurs paramètres**

- $f : \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$
- $f(x, y) = 3x - 5y$

Les fonctions peuvent être

- **injectives** chaque élément a un seul antécédent
- **surjectives** tous les éléments du domaine cible sont atteints
- **bijectives** si elle est injective et surjective

On peut (souvent) calculer avec une fonction

- sa **dérivée**
pour $f(x) = 2x$ la dérivée est $f'(x) = 2$
- son **intégrale**
pour $f(x) = 2x$ son intégrale est $f(x) = x^2$
- sa fonction **inverse**
pour $f(x) = 2x$ est $f^{-1}(x) = \frac{x}{2}$

Produit cartésien de deux ensembles

Énumération des **possibilités** de combinaisons de **deux ensembles**

- notation par le symbole **croix** $E \times F$
- les éléments du produits sont des « paires »
- éléments notés avec des **parenthèses** $(x, y) \in E \times F$
- **cardinalité** du produit cartésien $|E \times F| = |E| * |F|$
- utilisation de la puissance $E \times E = E^2$

Exemples

- positions dans un jeu d'échec
 - colonnes avec des lettres $C = \{A, B, C, D, E, F, G, H\}$
 - lignes avec des chiffres $L = \{1, 2, 3, 4, 5, 6, 7, 8\}$
 - ensemble des positions possibles
 $C \times L = \{(A, 1), (A, 2), (A, 3), \dots (B, 1), (B, 2) \dots (H, 8)\}$
 - une position $(D, 1) \in C \times L$
- position sur une carte
- taille et âge d'une personne
- modèle et couleur d'une voiture

Produit cartésien et vecteurs

Extension du produit cartésien à plus de deux ensembles

- possibilités avec plusieurs ensembles $E \times F \times G$
- élément avec des **composantes** ou **dimensions**
 $(x, y, z) \in E \times F \times G$
- même calcul de **cardinalité** $|E \times F \times G| = |E| * |F| * |G|$
- triplets, quadruplets, etc. selon le nombre de dimensions

Structure générale de **n-uplet**

- élément d'un produit cartésien $t = (x, y, z, t) \in E \times F \times G \times H$
- données dans l'espace mathématique considéré
 - pour les nombres, notion de **vecteur**
 - pour les bases de données, des **enregistrements**
- composants numérotés par **indices** $X = (x_1, x_2, x_3, x_4, x_5) \in \mathbb{R}^5$
- opérateurs classiques
 - addition $(x_1, x_2, x_3) + (y_1, y_2, y_3) = (x_1 + y_1, x_2 + y_2, x_3 + y_3)$
 - soustraction $(x_1, x_2, x_3) - (y_1, y_2, y_3) = (x_1 - y_1, x_2 - y_2, x_3 - y_3)$
 - multiplication scalaire $a * (x_1, x_2, x_3) = (a * x_1, a * x_2, a * x_3)$

Introduction aux matrices

Représentation de **tableaux** rectangulaires de nombres

- plusieurs **lignes** (vecteurs) ou **colonnes** (vecteurs)
- exemple de matrice 3×2

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}$$

- notation par **indices** sur une matrice M de dimensions 2×4

$$\begin{bmatrix} M_{11} & M_{12} & M_{13} & M_{14} \\ M_{21} & M_{22} & M_{23} & M_{24} \\ M_{31} & M_{32} & M_{33} & M_{34} \end{bmatrix}$$

Matrice termes par documents

Exemple avec trois phrases

- d1 : « un beau chat »
- d2 : « un chat et un gros chat »
- d3 : « un beau chien et un beau chat et un gros lion »

Représentation de chaque document par un **vecteur**

- vocabulaire comme **colonnes** (*un, beau, chat, et, gros, chien, lion*)
- **fréquence** des mots pour chaque document (ligne) $d_i \in \mathbb{N}^7$

	un	beau	chat	et	gros	chien	lion
d1	1	1	1	0	0	0	0
d2	2	0	2	1	1	0	0
d3	3	2	1	2	1	1	1

Document $d2$ est représenté par le **vecteur** : $d2 = (2, 0, 2, 1, 1, 0, 0)$

Produit vectoriel et matriciel

Produit de deux vecteurs (en anglais, *dot product*)

- seulement possible à **dimensions égales**
- vecteur ligne $(1, m)$ multiplié par un vecteur colonne $(m, 1)$
- somme des multiplications par composantes
 - $(x_1, x_2, x_3) * (y_1, y_2, y_3) = x_1 * y_1 + x_2 * y_2 + x_3 * y_3$
 - $X * Y = \sum_i x_i * y_i$

Produit matriciel

- produit des matrices (m, n) et (n, p) est (m, p)
- chaque élément de $P = M1 * M2$ résulte de la multiplication
 - ligne de $M1$ par la colonne $M2$
 - $MP_{ij} = \sum_k M1_{ik} * M2_{kj}$

$$M1 = \begin{bmatrix} 1 & 0 & 2 \\ 3 & 1 & -1 \end{bmatrix} \quad M2 = \begin{bmatrix} 0 & 3 \\ -1 & 2 \\ 1 & 0 \end{bmatrix}$$

$$P = \begin{bmatrix} 1 * 0 + 0 * -1 + 2 * 1 & 1 * 3 + 0 * 2 + 2 * 0 \\ 3 * 0 + 1 * -1 + -1 * 1 & 3 * 3 + 1 * 2 + -1 * 0 \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ -2 & 11 \end{bmatrix}$$

refaire mieux illustré

Projection par produit matriciel

Projection par produit de deux matrices

- D **données** (lignes)
- W **poids** de la projection (colonnes)
- P **résultat** de $D * W$ (lignes de D et colonnes de W)

La matrice W calcule des **combinaisons linéaires** de D

Quelques projections selon les nombres dans W

- matrice **identité** : carrée avec des 1 sur la diagonale
- matrice **filtre** : colonnes sélectionnées par des 1
- matrice **somme** : une colonne remplie de 1