100 Applications Concrètes

Des mathématiques dans notre quotidien

La base des IA, de la finance, de l'écologie et de nos décisions

Dr. Clotilde Djuikem

Avril 2025

À propos de l'autrice

Dr. Clotilde Djuikem est docteure en mathématiques appliquées, spécialiste des modèles mathématiques en épidémiologie, en écologie et dans les systèmes complexes. Elle est également engagée dans la vulgarisation scientifique, la promotion des sciences auprès des jeunes – notamment des filles et des personnes issues des communautés africaines – et l'enseignement des mathématiques appliquées à travers le monde.

Elle partage ses savoirs, son expérience et son énergie sur plusieurs plateformes :

- **YouTube** – `TiohAcademy`
- **LinkedIn** – `Clotilde Djuikem`
- **Site Web officiel** – `www.clotildedjuikem.com`
- **Facebook** – `Clotilde Djuikem`
- **TikTok** – : `Tioh Academy`

Découvrez aussi mon premier ouvrage !

25 Applications fascinantes

Ce livre présente, de façon simple et accessible, 25 situations issues de la vie réelle où les mathématiques jouent un rôle déterminant.

Vous y découvrirez comment les mathématiques interviennent dans :

— le traitement des maladies infectieuses,

— la finance et les décisions économiques,

— les réseaux sociaux,

— les transports et la logistique,

— la météo, l'alimentation, les sondages, et bien plus encore.

Chaque chapitre part d'une question concrète, et montre comment un outil mathématique peut y répondre, avec pédagogie, exemples et illustrations.

Où le trouver ? Sur `Amazon`, ou en me contactant directement via LinkedIn.

Un bon complément à ce livre si vous aimez relier les maths à la vraie vie.

Pourquoi ce livre ?

Reconcilier les adultes avec les mathématiques

Ce livre n'a pas seulement pour but de vulgariser les mathématiques.

Il est aussi une invitation : **une main tendue aux adultes** qui, un jour, ont été convaincus qu'ils n'étaient "pas faits pour les maths".

Beaucoup d'adultes gardent un souvenir douloureux ou frustrant de leurs années scolaires, notamment en mathématiques. Et sans le vouloir, ils transmettent ce vécu à leurs enfants — par leurs mots, leurs soupirs, leurs regards.

"Moi, j'ai toujours été nul en maths."

"Les maths, c'est trop compliqué."

"Tu verras, c'est l'enfer au lycée..."

Ces petites phrases, répétées sans méchanceté, peuvent semer des graines de peur ou de rejet chez les enfants. Et c'est souvent avant même qu'ils aient eu l'occasion de découvrir la beauté et la logique des mathématiques.

Et si on changeait cela ? Et si, en tant qu'adultes, nous changions notre regard sur les maths ? Non pas pour devenir experts, mais pour ouvrir un espace plus doux, plus accueillant, plus curieux.

Ce livre est là pour ça.

À travers des exemples concrets, accessibles, parfois ludiques, il vise à **reconnecter chacun à l'utilité et à la beauté des mathématiques dans la vie quotidienne**.

Car en réconciliant les adultes avec les maths, nous changeons aussi le regard que les enfants porteront demain sur cette discipline.

Ce n'est pas de la magie. C'est des maths. Et c'est pour tout le monde.

Ce livre présente les **50 premières applications concrètes des mathématiques** dans notre quotidien et dans des domaines variés : santé, informatique, économie, écologie, intelligence artificielle, et bien plus encore.

Mais ce n'est que la moitié du voyage. Le second volume, contenant **50 autres applications**, sera publié très prochainement.

Note à la lectrice, au lecteur

Ce livre est un ouvrage de vulgarisation scientifique. Il a pour objectif de montrer, à travers des exemples variés, comment les mathématiques sont présentes dans notre quotidien, nos décisions, nos environnements techniques ou naturels.

Pour rendre les concepts accessibles et vivants, certaines démonstrations sont volontairement simplifiées, et de nombreuses hypothèses implicites ou classiques sont parfois posées sans être détaillées. Cela ne remet pas en cause la rigueur de fond, mais permet de mieux saisir l'intuition derrière les idées.

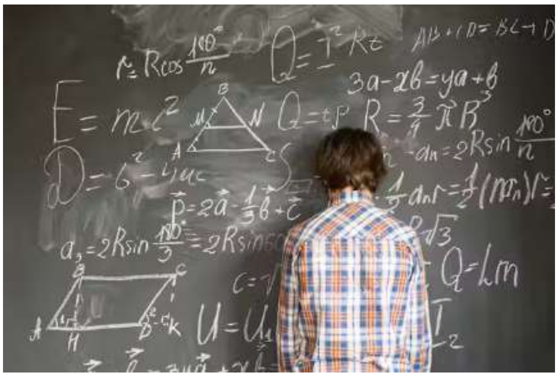
Certaines modélisations mathématiques présentées dans les chapitres reposent également sur des choix d'hypothèses (idéalisations, linéarités, données constantes, etc.), dans le but de montrer la mécanique d'un raisonnement ou la pertinence d'un outil.

Ce livre est une invitation à explorer en profondeur pour bien comprendre. Si vous souhaitez aller plus loin, chaque chapitre peut être un point de départ vers des lectures plus spécialisées.

Table des matières

Histoire des mathématiques

Les mathématiques : entre fascination et crainte



Peur des Maths : Pour certains, elles sont une abstraction complexe, une langue réservée aux initiés.

Joie des Maths : Pour d'autres, elles sont un outil puissant, une logique universelle qui structure le monde.

Pourquoi vulgariser les mathématiques ?

Un constat que j'entends souvent

"Encore un jour où le cosinus ne m'a pas servi !"

Pourquoi je fais de la vulgarisation ?

– Pour montrer que les maths ne sont pas réservées aux spécialistes.

– Pour aider chacun à mieux comprendre le monde qui nous entoure.

– Pour redonner confiance à ceux qui pensent que "les maths, ce n'est pas pour eux".

Les maths c'est la vie !

– Présentes partout dans notre quotidien.

– Une discipline universelle et intemporelle.

– Une base pour l'innovation, la recherche et la technologie.

0.0.1 Un voyage à travers l'histoire des mathématiques

Repères historiques essentiels

– **Antiquité :** calculs babyloniens, géométrie égyptienne.

– **Grèce antique :** Pythagore, Euclide, Archimède.

– **Moyen Âge :** développement de l'algèbre arabe, mathématiciennes africaines.

– **Renaissance :** Descartes, Newton, la révolution analytique.

– **Époque moderne :** intelligence artificielle, finance, écologie, cryptographie.

1. Les maths derrière la génération de texte en IA

Introduction

Problématique

Comment une intelligence artificielle (LLM) génère-t-elle des réponses à l'aide des mathématiques ?

En résumé, le modèle suit six étapes :

1. Encodage du texte via l'algèbre linéaire.
2. Analyse contextuelle avec les probabilités.
3. Passage dans un réseau neuronal de type Transformer.
4. Optimisation des paramètres par descente de gradient.
5. Évaluation via la théorie de l'information.
6. Génération finale du texte.

Encodage du texte

Les mots sont représentés par des vecteurs dans un espace de grande dimension :

$$\mathbf{w} = (w_1, w_2, ..., w_d) \in \mathbb{R}^d,$$

avec R^d représentant un espace mathématique à d dimensions où chaque dimension peut capturer une nuance de sens, et les w_i (avec $i = 1, 2, 3...$) les composantes du vecteurs w dans R_d .

Exemple pour le mot "Fourier" :

$$\mathbf{w}_{\text{Fourier}} = (-0.12, 0.54, ..., -0.89).$$

Ici, chaque chiffre est appris par le modèle et représente la signification du mot dans l'espace correspondant. Néanmoins, dans le cadre de cet exemple, ces valeurs sont purement illustrative.

L'ensemble du texte devient une matrice :

$$\mathbf{X} = \begin{bmatrix} w_{1,1} & w_{1,2} & \dots & w_{1,d} \\ w_{2,1} & w_{2,2} & \dots & w_{2,d} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n,1} & w_{n,2} & \dots & w_{n,d} \end{bmatrix} \in \mathbb{R}^{n \times d},$$

où n est le nombre de mots.

Prédiction par les probabilités

La prédiction des mots ce font de façon progressive. La probabilité du mot suivant $(w_i + 1)$, sachant les mots qui précèdent $(w_1 ... w_i)$, est donc calculée comme suit

$$P(w_{i+1} | w_1, ..., w_i) = \frac{\exp(s(w_i, w_{i+1}))}{\Sigma_j \exp(s(w_i, w_j))},$$

où $s(w_i, w_j)$ mesure la similarité (cosinus) :

$$s(w_i, w_j) = \frac{\mathbf{w}_i \cdot \mathbf{w}_j}{\|\mathbf{w}_i\| \|\mathbf{w}_j\|}.$$

On mesure la similarité s entre le contexte (w_i) et chaque mot possible $(w_j$ ou $w_i + 1)$. Plus l'angle est petit, plus ils sont proches en signification/contexte. La fonction exp transforme ces scores (pour les rendre positifs et accentuer les meilleurs), puis on normalise pour obtenir des résultats probabiliste.

Propagation dans un réseau Transformer

Les Transformers utilisent un mécanisme clé appelé *"attention"*. Il permet au modèle, pour prédire le mot suivant, de *"faire attention"* et de peser l'importance de différents mots dans le contexte, même s'ils sont éloignés. Les Transformers utilisent l'attention comme suit :

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right) V$$

où Q, K, V sont les matrices requêtes, clés et valeurs, et d_k la dimension des vecteurs clés. En effet, la requête demande de l'info, les Clés disent ce que chaque mot offre, les Valeurs donnent le contenu et l'apprentissage est stabilisé par $\sqrt{d_k}$

Optimisation par descente de gradient

Les paramètres sont ajustés à chaque étape d'apprentissage :

$$W^{(t+1)} = W^{(t)} - \eta \nabla_W \mathcal{L}(W)$$

où :

— W représente l'ensemble des paramètres du modèle (incluant les vecteurs mots et autres poids du réseau) qu'on veut améliorer.

— $L(W)$ est la fonction de perte ou d'erreur qui mesure à quel point le modèle se trompe.

— $\nabla_W \mathcal{L}(W)$ (le gradient) indique la direction de la pente de l'erreur.

— η est le taux d'apprentissage, un petit pas qu'on fait pour corriger les paramètres.

Théorie de l'information : évaluer la diversité

L'entropie $H(X)$ permet de quantifier la diversité de la sortie :

$$H(X) = -\sum_i P(x_i) \log P(x_i),$$

où X est l'ensemble des mots possibles en sortie, et $P(x_i)$ est la probabilité de chaque mot x_i prédite par le modèle. En pratique, on ajuste la *"température"* pour contrôler $H(X)$.

— Faible $H(X)$: réponse prévisible.

— Élevé $H(X)$: réponse aléatoire.

Génération finale du texte

Une fois les probabilités calculées et optimisées, l'IA génère un texte. Par exemple :

Exemple de réponse générée

La transformation de Fourier est une technique permettant de convertir un signal du domaine temporel vers le domaine fréquentiel, ce qui est crucial en IA pour analyser des patterns cachés dans les données.

Note : Cette réponse optimisée contraste avec des résultats moins structurés qu'on pourrait obtenir avec un code Python brut. L'amélioration continue des modèles permet d'affiner la qualité et ainsi, la qualité du text généré vient avec l'apprentissage.

Conclusion

Résumé

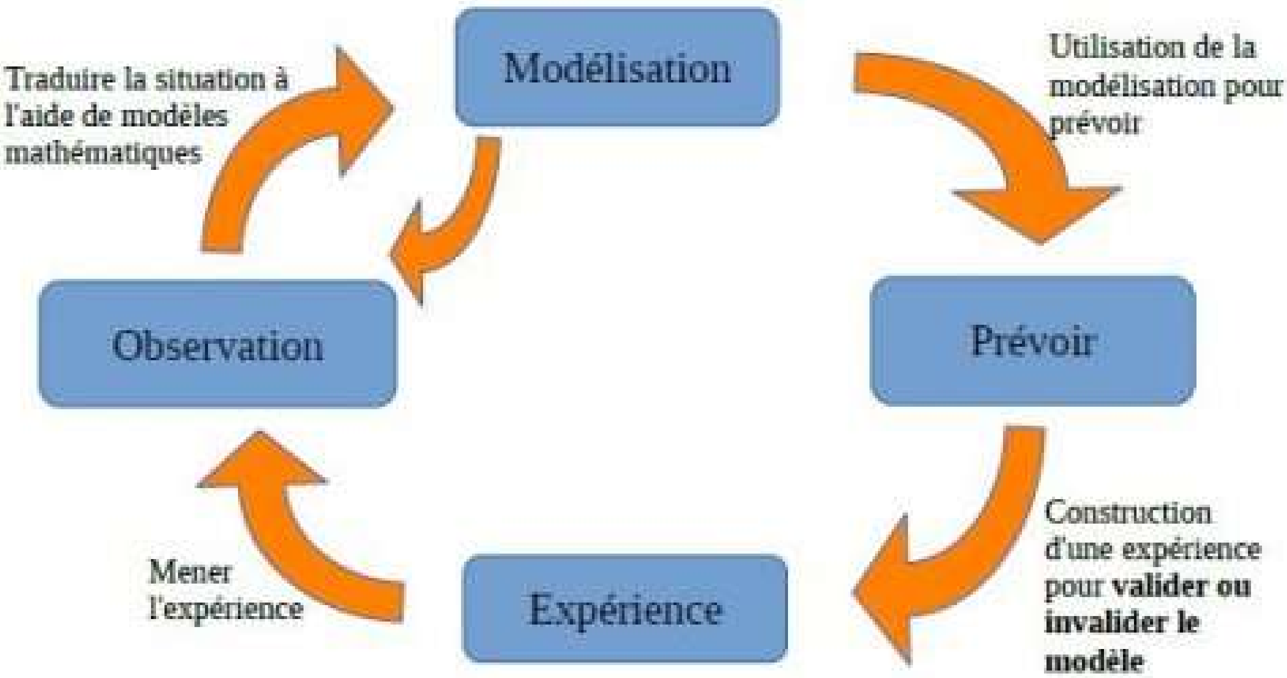
Grâce à l'algèbre linéaire, aux probabilités, aux réseaux Transformers, à l'optimisation et à la théorie de l'information, les IA génératives peuvent produire des textes cohérents et pertinents.

Les maths, c'est la vie.

Pour aller plus loin

Pensez à explorer davantage ce domaine passionnant, et si ce chapitre vous a plu, n'oubliez pas de suivre **Clotilde Djuikem** sur LinkedIn et **Tioh Academy** sur YouTube.

2. C'est quoi la modélisation mathématique ?



Qu'est-ce que la modélisation mathématique ?

Définition

La modélisation mathématique consiste à représenter un phénomène réel à l'aide de mathématiques pour mieux comprendre et résoudre des problèmes. On construit des équations ou des systèmes qui reproduisent le comportement observé dans le monde réel.

Applications :

— Prédire l'évolution d'une épidémie.

— Optimiser la production d'une entreprise.

— Simuler des phénomènes physiques comme la météo ou les fluides.

Les étapes de la modélisation mathématique

1. Identification du problème

Comprendre la situation réelle et choisir les éléments clés à modéliser.

2. Formulation mathématique

Traduire le problème en termes mathématiques avec des variables, des paramètres et des équations.

3. Résolution du modèle

Utiliser des techniques mathématiques (résolution d'équations, simulations) pour trouver des solutions.

4. Interprétation des résultats

Analyser les résultats obtenus et les comparer aux aux données actuelles ou réel.

5. Validation et ajustement

Si le modèle est inexact, ajuster les paramètres ou les hypothèses pour obtenir une meilleure correspondance.

Exemple : La croissance d'une population

Problème : Comment la population d'une espèce évolue-t-elle dans le temps ?

Hypothèse : La population croît proportionnellement à sa taille :

$$\frac{dP}{dt} = rP$$

où $P(t)$ est la population au temps t, et r est le taux de croissance.

Solution :

$$P(t) = P_0 e^{rt}$$

avec P_0 la population initiale.

Interprétation : La population croît de façon exponentielle si $r > 0$.

Types de modèles mathématiques

- **Modèles algébriques :** utilisent des équations simples pour décrire des relations.
- **Modèles différentiels :** décrivent comment un système évolue dans le temps.
- **Modèles statistiques :** basés sur les probabilités pour modéliser des phénomènes incertains.
- **Modèles de simulation :** utilisent des algorithmes pour simuler des systèmes complexes.

Application : Le modèle SIR

Problème : Modéliser la propagation d'une maladie.
Équations :

$$\frac{dS}{dt} = -\beta SI, \quad \frac{dI}{dt} = \beta SI - \gamma I, \quad \frac{dR}{dt} = \gamma I$$

où $S(t)$, $I(t)$, et $R(t)$ sont les populations susceptibles, infectées et rétablies.

Utilité : Ce modèle permet de prédire la dynamique d'une épidémie.

Conclusion

Résumé

La modélisation mathématique est une approche pour comprendre, prédire, et optimiser des phénomènes réels. Elle est présente dans de nombreux domaines comme les sciences, l'économie, ou la biologie.

3. Comment les banques calculent les intérêts ?

Situation du quotidien

Tu ouvres un compte d'épargne. On t'annonce un taux d'intérêt annuel de 5 %. Mais que signifie vraiment ce 5 % ? Combien vas-tu gagner dans un an ? Et dans 5 ans ? Pourquoi certaines personnes gagnent-elles plus, alors qu'elles ont mis la même somme au départ ? **Les mathématiques des intérêts peuvent tout changer dans ta façon d'épargner.**

Les outils mathématiques : intérêts simples et intérêts composés

Il existe deux grandes manières de calculer les intérêts sur ton argent :

- **Intérêt simple** : les intérêts sont calculés uniquement sur le capital de départ.
- **Intérêt composé** : les intérêts générés sont eux-mêmes réinvestis, et produisent à leur tour des intérêts.

Voici les formules utilisées :

- **Intérêt simple** : $A = P(1 + rt)$
- **Intérêt composé** : $A = P(1 + r)^t$

Où :

- A : le montant total après t années
- P : le capital initial (la somme déposée)
- r : le taux d'intérêt annuel (en décimal, donc $5\,\% = 0{,}05$)
- t : la durée en années

Exemple concret

Imaginons que tu déposes 1 000 $ sur un compte épargne à 5 % d'intérêt :

– **Avec intérêt simple**, après 3 ans :

$$A = 1000 \times (1 + 0{,}05 \times 3) = 1000 \times 1{,}15 = 1150\,\$$$

– **Avec intérêt composé**, après 3 ans :

$$A = 1000 \times (1 + 0{,}05)^3 = 1000 \times 1{,}157625 = 1157{,}63\,\$$$

La différence peut sembler petite au début... mais plus le temps passe, plus elle devient importante !

Pourquoi c'est utile ?

Réflexion

Savoir comment les intérêts fonctionnent, c'est :

- éviter de se faire piéger par des prêts coûteux,
- mieux choisir ses produits financiers (compte épargne, assurance vie, etc.),
- comprendre la dynamique de l'endettement ou de l'investissement,
- transmettre à ses enfants des bases solides pour mieux gérer leur argent.

Et dans la vraie vie ?

Les banques utilisent généralement l'intérêt composé. Cela signifie que plus tu laisses ton argent longtemps, plus il travaille pour toi. C'est ce qu'on appelle « l'effet boule de neige ».

Mais attention, c'est exactement le même mécanisme qui fait qu'une dette non remboursée peut exploser avec le temps ! Les mathématiques ne mentent pas : elles montrent comment les décisions d'aujourd'hui façonnent notre avenir financier.

4. Comment les mathématiques nous permettent d'optimiser la répartition des budgets publicitaire ?

Problème posé

Pourquoi les publicités que vous voyez en ligne semblent-elles si bien vous correspondre ? Et comment les entreprises choisissent-elles les bons canaux publicitaires tout en maîtrisant leurs budgets ? Les mathématiques, et en particulier la programmation linéaire (encore appelée optimisation linéaire, est une technique de programmation mathématique, utilisée pour la résolution de problèmes d'optimisation dans lesquels les contraintes et les objectifs peuvent être exprimés sous forme linéaire.), offrent des outils puissants pour répondre à ces questions.

Un objectif clair : optimiser les ressources

L'objectif d'une entreprise est souvent de maximiser son impact publicitaire (nombre de personnes atteintes, taux de clic, conversions, etc.) tout en respectant des contraintes de budget, de disponibilité ou de performance. Cette situation se prête parfaitement à une modélisation en programmation linéaire. De plus, les publicités en ligne semblent souvent bien vous correspondre parce qu'elles sont personnalisées grâce à une collecte massive de données (Cookies et traceurs) sur votre comportement numérique (profil,âge, lieu, préférences, historique d'achat...) et ceci n'est que le résultat d'un algorithm modeliser grâce aux mathématiques.

Formulation générale du modèle

On cherche à maximiser une fonction objectif de la forme :

$$\text{Maximiser} \quad Z = c_1x_1 + c_2x_2 + \cdots + c_nx_n$$

Sous contraintes :

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \leq b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \leq b_2 \\ \vdots \\ x_i \geq 0 \quad \text{pour tout } i \end{cases}$$

où :

- x_i est le montant alloué au canal publicitaire i,
- c_i représente l'efficacité attendue de ce canal (portée, clics, etc.),
- les coefficients a_{ij} traduisent les contraintes de budget ou de stratégie.

Exemple concret

Une entreprise souhaite répartir un budget de 10 000 $ entre deux canaux : - x_1 : Google Ads, - x_2 : Facebook Ads.

Elle évalue que : - Chaque 1 000 $ investis sur Google Ads rapportent 30 points d'impact. - Chaque 1 000 $ sur Facebook Ads en rapportent 50.

Mais elle ne peut pas dépenser : - plus de 10 000 $ au total, - plus de 7 000 $ sur Google, - plus de 5 000 $ sur Facebook.

Le problème devient :

$$\text{Maximiser } Z = 30x_1 + 50x_2$$

avec :

$$\begin{cases} x_1 + x_2 \leq 10 \\ x_1 \leq 7 \\ x_2 \leq 5 \\ x_1, x_2 \geq 0 \end{cases}$$

Résolution

Ce problème peut être résolu graphiquement ou par la méthode du simplexe qui est une procédure itérative permettant d'effectuer une exploration dirigée de l'ensemble des solutions réalisables de base. Ainsi, la solution optimale est :

$$x_1 = 7 \quad \text{et} \quad x_2 = 3$$

Soit : - 7 000 $ sur Google Ads, - 3 000 $ sur Facebook Ads.

Le gain maximal est :

$$Z = 30 \times 7 + 50 \times 3 = 210 + 150 = 360$$

Conclusion

La programmation linéaire permet de prendre des décisions optimales dans un contexte contraint. Dans le domaine publicitaire, elle aide à répartir un budget entre plusieurs canaux pour maximiser l'efficacité des campagnes. Cette méthode, simple mais puissante, est largement utilisée dans le marketing, la logistique, et la finance.

5. Pourquoi étudier les nombres premiers ?

Définition

Un nombre premier est un entier naturel strictement supérieur à 1 qui n'a que deux diviseurs : 1 et lui-même.

Par exemple : 2, 3, 5, 7 et 11 sont des nombres premiers. Ils ne sont divisibles que par 1 et eux-mêmes.

Les briques de l'arithmétique

Les nombres premiers sont considérés comme les "atomes" des entiers. Selon le théorème fondamental de l'arithmétique, tout entier naturel peut s'écrire comme un produit unique de nombres premiers (à l'ordre près).

— Exemple : $12 = 2^2 \times 3$ (identiques à 3×2^2)

— Leur étude permet de mieux comprendre la structure interne des nombres entiers.

Application en cryptographie

Les nombres premiers jouent un rôle central dans la cryptographie moderne, notamment dans les algorithmes RSA (Rivest-Shamir-Adleman).

— **Clés de chiffrement :** On choisit deux grands nombres premiers p et q, puis on calcule $n = p \times q$.

— **Sécurité :** Le produit n est utilisé comme partie de la clé publique. La difficulté à retrouver p et q à partir de n (la factorisation) garantit la sécurité.

Exemple : Envoi sécurisé d'une image

Imaginons que vous voulez envoyer une photo privée en toute sécurité.

1. Le destinataire choisit deux grands nombres premiers p et q, calcule $n = p \times q$ et publie n et un exposant e comme clé publique.
2. Vous convertissez la photo en données numériques, et pour chaque nombre m vous appliquez le chiffrement :
$$c = m^e \mod n$$
3. Vous envoyez les valeurs chiffrées c.
4. Le destinataire utilise sa clé privée d pour retrouver chaque donnée :
$$m = c^d \mod n$$

Ce système est fondé sur les propriétés uniques des nombres premiers.

Application en biologie : les cigales périodiques

Certaines espèces de cigales, notamment en Amérique du Nord, émergent tous les 13 ou 17 ans — deux nombres premiers.

Pourquoi ces cycles ?

— Ils réduisent les chances de coïncider avec les cycles de reproduction de leurs prédateurs (souvent de 2 à 5 ans).

— Cela maximise leurs chances de survie.

Ainsi, les nombres premiers deviennent un outil de survie dans la nature.

Théorie : l'infinité des nombres premiers

Le mathématicien grec Euclide a démontré il y a plus de 2 000 ans que les nombres premiers sont infinis. Utilisons le raisonnement par l'absurde que nous avons étudié précédement.

Supposons qu'il n'en existe qu'un nombre fini, notés $p_1, p_2, \ldots, p_n$. Construisons le nombre :
$$N = p_1 \cdot p_2 \cdot \cdots \cdot p_n + 1$$

Ce nombre n'est divisible par aucun des p_i car il reste 1 dans chaque division. Il est donc soit premier, soit divisible par un autre nombre premier non listé, ce qui contredit l'hypothèse.

Conclusion : il existe une infinité de nombres premiers.

Le plus grand nombre premier connu

Le plus grand nombre premier connu est le **nombre de Mersenne**, de la forme :

$$2^{136\,279\,841} - 1$$

Ce nombre a été découvert le 11 octobre 2024 et contient plus de 41 millions de chiffres. Il a été identifié grâce au projet GIMPS (Great Internet Mersenne Prime Search), utilisant la puissance de milliers d'ordinateurs.

Un nombre de Mersenne est de la forme $2^p - 1$, avec p premier. Ces nombres sont recherchés pour leur beauté mathématique, mais aussi pour leurs applications pratiques comme la cryptographie.

Conclusion

Les nombres premiers sont essentiels en mathématiques, en cryptographie, en biologie, en informatique et au-delà. Ils relient l'abstraction pure à des usages concrets, et continuent de fasciner chercheurs et curieux du monde entier.

6. Pourquoi utilise-t-on le raisonnement par l'absurde ?

Qu'est-ce que le raisonnement par l'absurde ?

Le raisonnement par l'absurde consiste à démontrer la vérité d'une affirmation en supposant qu'elle est fausse, puis en montrant que cette hypothèse conduit à une contradiction.

Principe général

Le principe est simple :

— On suppose que l'énoncé à démontrer est faux.

— On développe cette hypothèse logiquement.

— Si on aboutit à une contradiction, on conclut que notre hypothèse de départ était fausse.

— Par conséquent, l'énoncé initial est vrai.

Ce type de raisonnement est très utilisé en mathématiques, logique, mais aussi en philosophie, informatique ou prise de décision.

Exemple mathématique : l'irrationalité de $\sqrt{2}$

On souhaite prouver que $\sqrt{2}$ est irrationnel.

— Supposons que $\sqrt{2} = \frac{a}{b}$ avec a et b entiers premiers entre eux.

— En élevant au carré, on obtient $2 = \frac{a^2}{b^2} \Rightarrow a^2 = 2b^2$.

— Donc a^2 est pair, donc a est pair.

— On peut écrire $a = 2k$, donc $a^2 = 4k^2 = 2b^2 \Rightarrow b^2 = 2k^2$.

— Donc b est aussi pair, ce qui contredit le fait que a et b soient premiers entre eux.

Conclusion : notre hypothèse est fausse, donc $\sqrt{2}$ est irrationnel.

Exemple philosophique : le paradoxe de l'omnipotence

Peut-on concevoir un être tout-puissant qui crée une pierre qu'il ne peut pas soulever ? En analysant cette question par l'absurde :

— Si une telle pierre existe, l'être n'est plus tout-puissant (il ne peut pas la soulever).

— S'il ne peut pas la créer, il n'est pas tout-puissant non plus.

Cette contradiction permet d'illustrer les limites conceptuelles du concept d'omnipotence.

Exemple en informatique : validation d'un algorithme

Supposons qu'un algorithme de tri laisse une liste non triée.
En suivant cette hypothèse :

— On analyse le fonctionnement de chaque étape.

— On identifie une incohérence : par exemple, une comparaison est contredite par un échange précédent.

Conclusion : si on arrive à une contradiction, on en déduit que l'algorithme fonctionne correctement dans le cadre défini.

Exemple de la vie quotidienne

Vous hésitez à changer de route pour éviter un embouteillage.
Raisonnement par l'absurde :

— Supposons que tout le monde prenne la nouvelle route.

— Elle devient alors plus bouchée que la route initiale.

— Cette contradiction suggère que changer de route ne garantit pas un meilleur choix.

Exemple mathématique : Peut-on diviser par zéro ?

Supposons que l'on puisse diviser par zéro :

— Soit $a = b \neq 0$. Alors $a^2 = ab$.

— On soustrait b^2 : $a^2 - b^2 = ab - b^2$.

— Ce qui donne $(a-b)(a+b) = b(a-b)$.
— En divisant par $a - b = 0$, on obtient $a + b = b$.
— Or $a = b$, donc $2b = b$, donc $2 = 1$.

Cette absurdité montre que l'on ne peut pas diviser par zéro.

Conclusion

Le raisonnement par l'absurde est un outil fondamental pour démontrer, tester ou même simplement réfléchir. Il permet d'aborder des problèmes indirectement et de façon souvent plus accessible. On le retrouve aussi bien dans les démonstrations formelles que dans les raisonnements quotidiens.

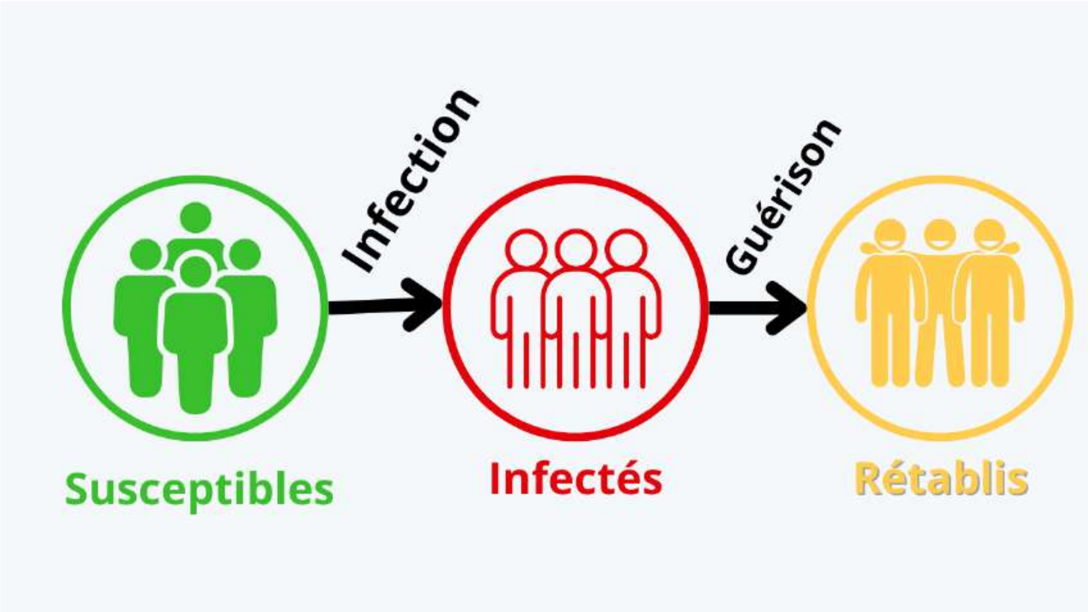
7. Maths pour comprendre les maladies

Cette situation est-elle vraiment derrière nous ?

Cette image reste gravée dans nos mémoires. La pandémie de COVID-19 a bouleversé le monde entier. Mais comment les scientifiques prennent-ils des décisions en période de crise sanitaire ? **Les mathématiques sont l'un des outils les plus puissants pour comprendre et anticiper la propagation des maladies.**

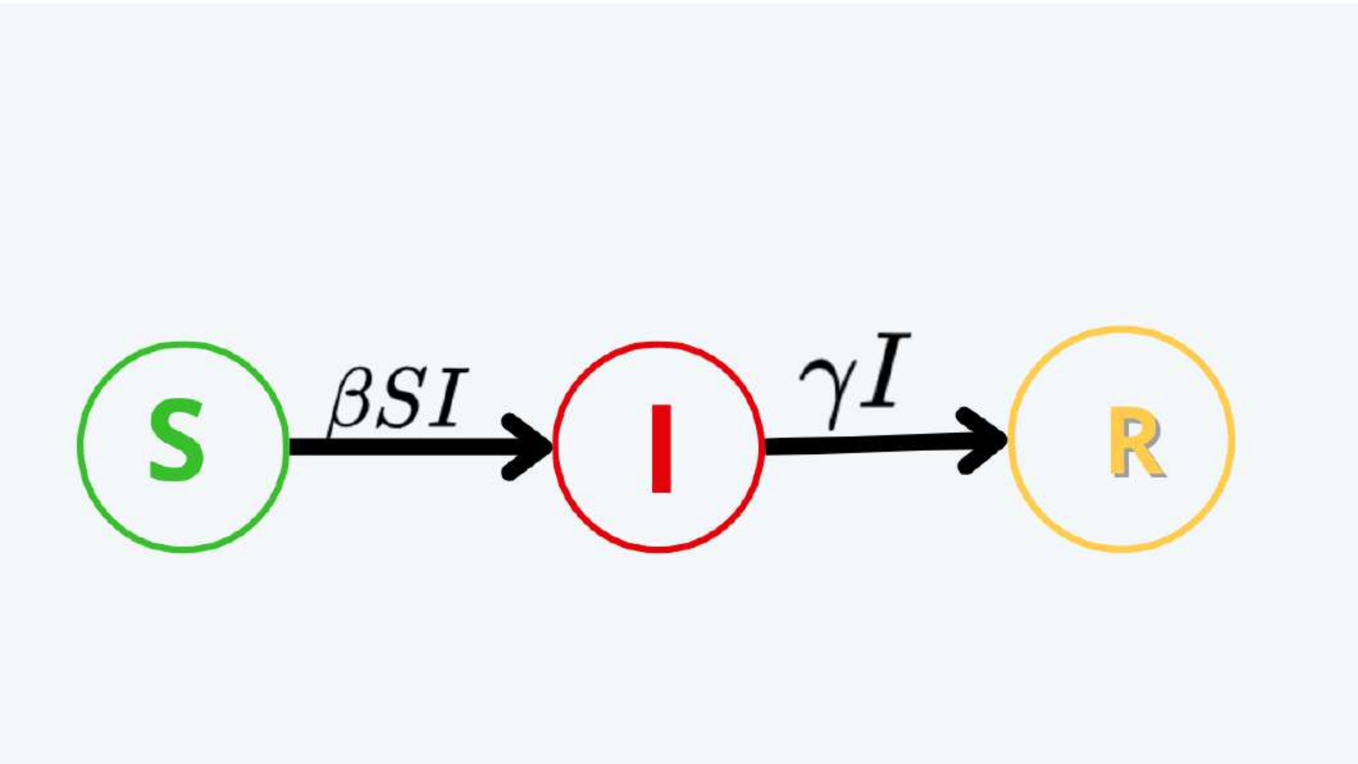
L'un des modèles les plus utilisés : le SIR

Le modèle SIR est l'un des modèles épidémiologiques les plus simples. Il divise la population en trois catégories :



Les variables du modèle SIR

- **S :** Susceptibles – les personnes pouvant être infectées.
- **I :** Infectés – les personnes qui propagent la maladie.
- **R :** Rétablis – guéris, et qui ne peuvent plus être infectés.



Les équations du modèle SIR

Les équations différentielles qui décrivent ce modèle sont :

$$\frac{dS}{dt} = -\beta SI$$
$$\frac{dI}{dt} = \beta SI - \gamma I$$
$$\frac{dR}{dt} = \gamma I$$

– β : taux de transmission de la maladie.

– γ : taux de guérison (en général, $1/\gamma$ est la durée moyenne de guérison).

La première équation étant la variation dans le temps des personnes succeptibles dêtre infectées, la seconde la variation de ceux propageant la maladie et la dernière la variation des personnes guéris et ne pouvant plus être infecté.

Le nombre de base de reproduction $\mathcal{R}_0$

Le fameux $\mathcal{R}_0$ est un indicateur clé :

$$\mathcal{R}_0 = \frac{\beta}{\gamma}$$

Interprétation

– Si $\mathcal{R}_0 > 1$: l'épidémie se propage.

– Si $\mathcal{R}_0 < 1$: l'épidémie régresse.

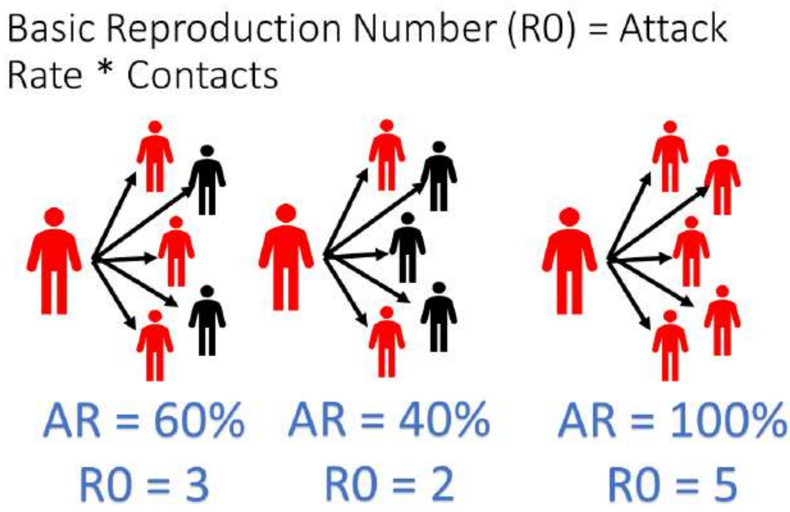
Exemple concret

Imaginons :

$$\beta = 0.3 \quad \text{et} \quad \gamma = 0.1 \Rightarrow \mathcal{R}_0 = \frac{0.3}{0.1} = 3$$

Cela signifie que, dans cette situation, une personne infectée va en moyenne contaminer 3 autres personnes si personne n'est immunisé et qu'aucune mesure n'est mise en place.

C'est une dynamique exponentielle : ces 3 personnes en contamineront chacune 3 autres, et ainsi de suite. Ce type de propagation rapide explique pourquoi certaines épidémies peuvent se répandre très vite si on n'agit pas rapidement. Pour bien illustrer voilà 3 cas :



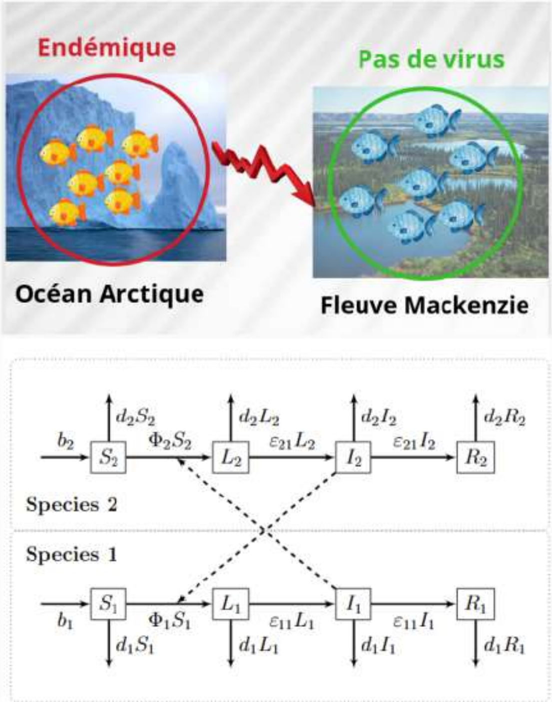
Pourquoi c'est important ?

Manipuler $\mathcal{R}_0$ pour sauver des vies

La compréhension de $\mathcal{R}_0$ permet d'adapter les politiques publiques :

- Le confinement
- La quarantaine
- La vaccination ciblée
- La gestion des contacts

Pour aller plus loin... avec les poissons !



équations differentielles

$$\dot{S}_1 = b_1 - \beta_{111} S_1 I_1 - \beta_{121} S_1 I_2 - d_1 S_1$$
$$\dot{L}_1 = \beta_{111} S_1 I_1 + \beta_{121} S_1 I_2 - (\varepsilon_{11} + d_1) L_1$$
$$\dot{I}_1 = \varepsilon_{11} L_1 - (\gamma_{11} + d_1) I_1$$
$$\dot{R}_1 = \gamma_{11} I_1 - d_1 R_1$$
$$\dot{S}_2 = b_2 - \beta_{211} S_2 I_1 - \beta_{221} S_2 I_2 - d_2 S_2$$
$$\dot{L}_2 = \beta_{211} S_2 I_1 + \beta_{221} S_2 I_2 - (\varepsilon_{21} + d_2) L_2$$
$$\dot{I}_2 = \varepsilon_{21} L_2 - (\gamma_{21} + d_2) I_2$$
$$\dot{R}_2 = \gamma_{21} I_2 - d_2 R_2.$$

Les modèles SIR ne servent pas qu'à comprendre les virus humains ! Ils s'appliquent aussi aux maladies animales, comme celles que j'étudie actuellement sur les poissons dans le cadre de mon postdoctorat au Canada.

8. La transmission et la propagation du VIH/SIDA

FIGURE 8.1 – Illustration de la propagation du VIH/SIDA

Problème

Comment les mathématiques peuvent-elles modéliser la transmission du VIH, la progression vers le SIDA, ainsi que les mortalités naturelles et liées à la maladie, tout en calculant le nombre de reproduction de base R_0 ?

Réponse : Nous allons utiliser un système d'équations différentielles pour représenter l'évolution des différentes catégories de la population, et analyser comment les paramètres influencent la propagation.

Modèle mathématique avec mortalité naturelle et liée à la maladie

Nous considérons trois catégories :

- S : individus susceptibles,
- I_1 : individus infectés par le VIH (asymptomatiques),
- I_2 : individus atteints du SIDA (symptomatiques).

Le système différentiel représentant la variation temporelle de chacun des individus ci-dessus est :

$$\frac{dS}{dt} = \Lambda_n(S + I_1 + I_2) - \beta S(I_1 + \varepsilon I_2) - \mu_n S$$
$$\frac{dI_1}{dt} = \beta S(I_1 + \varepsilon I_2) - \sigma I_1 - \mu_n I_1$$
$$\frac{dI_2}{dt} = \sigma I_1 - (\mu_2 + \mu_n)I_2$$

où :

- β : taux de transmission du VIH,
- $\varepsilon < 1$: sidéens transmettent moins que les porteurs asymptomatiques,
- σ : taux de progression vers le SIDA,
- Λ_n : taux de natalité,
- μ_n : mortalité naturelle,
- μ_2 : mortalité liée au SIDA.

Ces variations sont fortement affecté par les taux de natalités et de mortalité (naturelle et liée au VIH) d'où la proportionalité.

Calcul du nombre de reproduction de base R_0

Le R_0 représente le nombre moyen de nouvelles infections provoquées par un seul infecté au début de l'épidémie.

$$R_0 = \frac{\beta\Lambda_n(\varepsilon + \mu_n + \mu_2)}{\mu_n(\mu_n + \sigma)(\mu_n + \mu_2)}$$

- Si $R_0 > 1$, l'épidémie se propage,
- Si $R_0 < 1$, l'infection s'éteint (hypothèse théorique en l'absence de traitement).

Influence des paramètres sur R_0

- β : plus le taux de transmission est élevé, plus R_0 augmente.
- μ_n : une mortalité naturelle plus forte diminue R_0, car les individus vivent moins longtemps pour transmettre.
- σ : une progression rapide vers le SIDA diminue aussi R_0, car les personnes symptomatiques sont moins actives sexuellement et vivent moins longtemps.

Applications et utilité du modèle

Ce modèle permet de :

- Simuler l'évolution de l'épidémie en fonction de différents paramètres,
- Estimer le seuil d'intervention nécessaire pour faire descendre R_0 en dessous de 1,
- Guider les politiques de prévention et de traitement (ex. : campagnes ciblées, accès aux antirétroviraux).

Limites du modèle

Critiques

- La progression vers le SIDA est supposée constante (σ), sans prise en compte des traitements qui la ralentissent.
- Le modèle ne tient pas compte des campagnes de prévention, ni de l'effet des médicaments sur β.
- Il suppose une population homogène en matière de comportement, ce qui ne reflète pas la réalité sociale (groupes à risque, etc.).

Conclusion

Les mathématiques permettent de modéliser l'évolution du VIH/SIDA, de calculer le R_0, et d'évaluer l'impact des paramètres épidémiologiques. Ces modèles sont des outils essentiels pour orienter les décisions de santé publique, en particulier dans les contextes à ressources limitées.

continue from here!!!

9. Décrypter les articles sans les lire

Le Contexte

Question

Comment peut-on découvrir les thèmes sous-jacents dans un corpus d'articles sans les lire, en utilisant les mathématiques ?

Les maths

La modélisation par Latent Dirichlet Allocation (LDA) repose sur un modèle probabiliste pour identifier les thèmes dominants dans un ensemble de documents.

Trois éléments clés sont utilisés :

— La distribution des thèmes dans les documents ;

— La distribution des mots dans les thèmes ;

— Les lois de Dirichlet pour modéliser l'incertitude.

Application : Politique, Économie et Sport

Exemple : trois articles sont analysés avec LDA.

— **Article 1** (élections) : $\theta_1 = [0.7, 0.2, 0.1]$
$\rightarrow$ 70% politique, 20% économie, 10% sport.

— **Article 2** (croissance) : $\theta_2 = [0.1, 0.8, 0.1]$

— **Article 3** (football) : $\theta_3 = [0.05, 0.05, 0.9]$

Les mots caractéristiques pour chaque thème :

$$\phi_{\text{politique}} = [\text{"élection"}, \text{"gouvernement"}, \text{"parti"}]$$
$$\phi_{\text{économie}} = [\text{"croissance"}, \text{"inflation"}, \text{"marché"}]$$
$$\phi_{\text{sport}} = [\text{"match"}, \text{"équipe"}, \text{"score"}]$$

Question : Que signifient θ et ϕ ? Comment les calcule-t-on ? C'est là qu'interviennent les maths.

Fondements mathématiques du LDA

Une distribution, c'est la façon dont quelque chose est réparti entre plusieurs groupes. Le symbole $\sim$ signifie "suivant la loi de".

1. Distribution des thèmes θ_d :

$$\theta_d \sim \text{Dirichlet}(\alpha)$$

- θ_d : vecteur représentant la proportion des thèmes dans le document.
- α : paramètre contrôlant la concentration thématique.

2. Distribution des mots ϕ_k :

$$\phi_k \sim \text{Dirichlet}(\beta)$$

- ϕ_k : vecteur des probabilités des mots pour le thème k.
- β : paramètre contrôlant la diversité des mots dans un thème.

Processus de génération des documents

LDA suppose que chaque mot d'un document provient d'un thème latent :

- On tire $\theta_d \sim \text{Dirichlet}(\alpha)$ pour chaque document.
- Pour chaque mot :
  - on tire un thème $z_{dn} \sim \theta_d$
  - on tire un mot $w_{dn} \sim \phi_{z_{dn}}$

Vraisemblance jointe :

$$\mathrm{P}(w, z, \theta, \phi | \alpha, \beta) = \prod_{d=1}^{D} \mathrm{P}(\theta_d|\alpha) \prod_{n=1}^{N_d} \mathrm{P}(z_{dn}|\theta_d)\mathrm{P}(w_{dn}|z_{dn}, \phi)$$

Estimation des paramètres LDA

Pour estimer θ_d et ϕ_k, on utilise :

- l'échantillonnage de Gibbs,

— ou l'inférence variationnelle.

Maximisation de la vraisemblance :

$$\max_{\theta,\phi} \sum_{d=1}^{D} \log \mathrm{P}(w_d|\theta_d, \phi)$$

Résultats :

— θ_d indique les thèmes dominants d'un document ;

— ϕ_k donne les mots les plus représentatifs de chaque thème.

Critique du modèle LDA

Forces :

— Simple, efficace, interprétable.

— Utilisable dans de nombreux domaines (politique, santé, sport...).

Limites :

— Nécessité de fixer le nombre de thèmes à l'avance.

— Perte du contexte : l'ordre des mots n'est pas pris en compte.

— Hypothèse d'indépendance des documents.

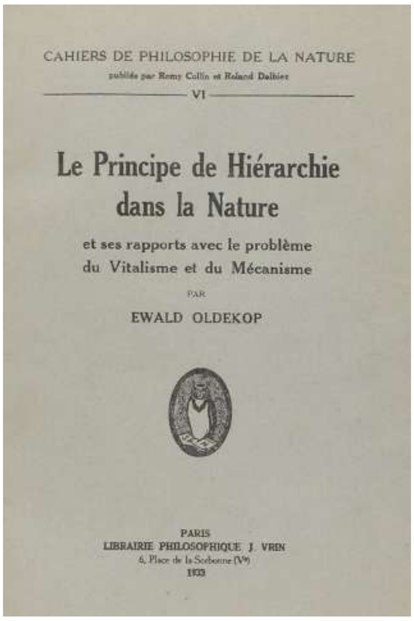
Recommandation : combiner LDA avec d'autres méthodes (réseaux de neurones, embeddings...) pour une analyse plus riche.

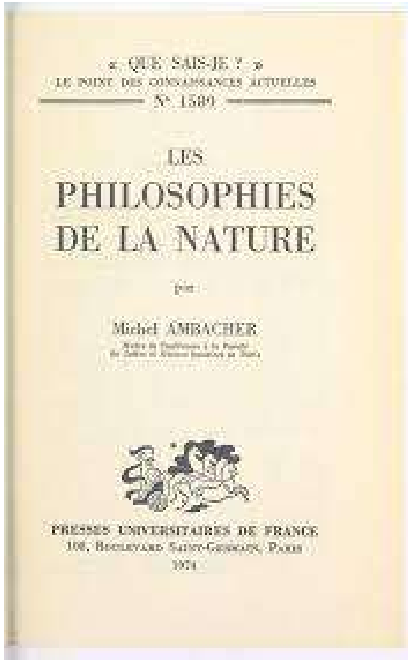
Conclusion

La Latent Dirichlet Allocation (LDA) offre une approche mathématique puissante pour découvrir automatiquement les thèmes cachés dans un corpus. Elle permet de structurer l'information, d'organiser des archives, ou de mieux comprendre les tendances d'un domaine.

10. Comparaison des textes philosophiques

Comment structurer une réflexion sans tout lire ?





Question

Lorsqu'on compare des textes philosophiques, comment peut-on structurer et analyser leurs contenus de façon quantitative, sans tout lire ligne par ligne ? Existe-t-il une manière de repérer rapidement les idées dominantes ?

Une réponse mathématique : TF-IDF

L'un des outils mathématiques utilisés en analyse de texte est la méthode TF-IDF. C'est une technique issue du traitement automatique du langage, mais qui s'applique aussi très bien à des textes philosophiques.

Elle permet de quantifier l'importance relative d'un mot dans un ensemble de textes.

Trois définitions essentielles

- Le **corpus** : ensemble de textes analysés.
- Le **TF (Term Frequency)** : fréquence du mot dans un texte donné.
- L'**IDF (Inverse Document Frequency)** : rareté du mot dans le corpus global.

Qu'est-ce qu'un corpus ?

Un **corpus**, en analyse textuelle, désigne simplement un ensemble de textes. Par exemple, si nous analysons des œuvres de Platon, Aristote et Descartes, ces trois textes formeront notre corpus philosophique.

Term Frequency (TF)

La TF d'un mot mesure combien de fois ce mot apparaît dans un texte, divisé par le nombre total de mots du texte.

Exemple : Dans "La République" de Platon, si le mot "**justice**" apparaît 50 fois sur 10000 mots :

$$\mathrm{TF}(\text{“justice”}) = \frac{50}{10\,000} = 0{,}005$$

Inverse Document Frequency (IDF)

L'IDF indique si un mot est rare dans l'ensemble du corpus. La formule est :

$$\mathrm{IDF}(t) = \log\left(\frac{N}{n_t}\right)$$

Où :

- N : nombre total de textes du corpus,
- n_t : nombre de textes contenant le mot t.

Exemple simplifié avec 3 textes

1. Texte 1 : *"La justice est essentielle à l'harmonie sociale."*

2. Texte 2 : *"L'éthique et la politique sont des notions centrales en philosophie."*
3. Texte 3 : *"La justice et la morale guident les décisions humaines."*

On souhaite calculer la TF-IDF du mot "justice".

Calcul de la TF (par texte)

– Texte 1 : $\text{TF}(\text{“justice”}) = \frac{1}{7} \approx 0{,}142$

– Texte 2 : $\text{TF}(\text{“justice”}) = 0$

– Texte 3 : $\text{TF}(\text{“justice”}) = \frac{1}{8} \approx 0{,}111$

Calcul de l'IDF du mot "justice"

Le mot "justice" apparaît dans 2 des 3 textes :

$$\text{IDF}(\text{“justice”}) = \log\left(\frac{3}{2}\right) \approx 0{,}176$$

Calcul final : TF-IDF = TF × IDF

– Texte 1 : $\text{TF-IDF} = 0{,}142 \times 0{,}176 \approx 0{,}025$

– Texte 2 : $\text{TF-IDF} = 0$

– Texte 3 : $\text{TF-IDF} = 0{,}111 \times 0{,}176 \approx 0{,}019$

Conclusion mathématique

Le mot "justice" est le plus représentatif du **Texte 1**. Il est présent dans le Texte 3, mais de manière moins dominante. Dans le Texte 2, il est absent.

Dans la pratique : visualisation en Python

```
from sklearn.feature_extraction.text import TfidfVectorizer
import pandas as pd

# Corpus de textes philosophiques
corpus = [
    "La justice est essentielle à l'harmonie sociale.",
    "L'éthique et la politique sont des notions centrales en philosophie.",
    "La justice et la morale guident les décisions humaines."
]

# Initialisation du vecteur TF-IDF
vectorizer = TfidfVectorizer()
tfidf_matrix = vectorizer.fit_transform(corpus)

# Conversion en DataFrame pour une meilleure lisibilité
df_tfidf = pd.DataFrame(tfidf_matrix.toarray(), columns=vectorizer.get_feature_name

# Affichage des scores TF-IDF
df_tfidf
```

Interprétation

Pertinence

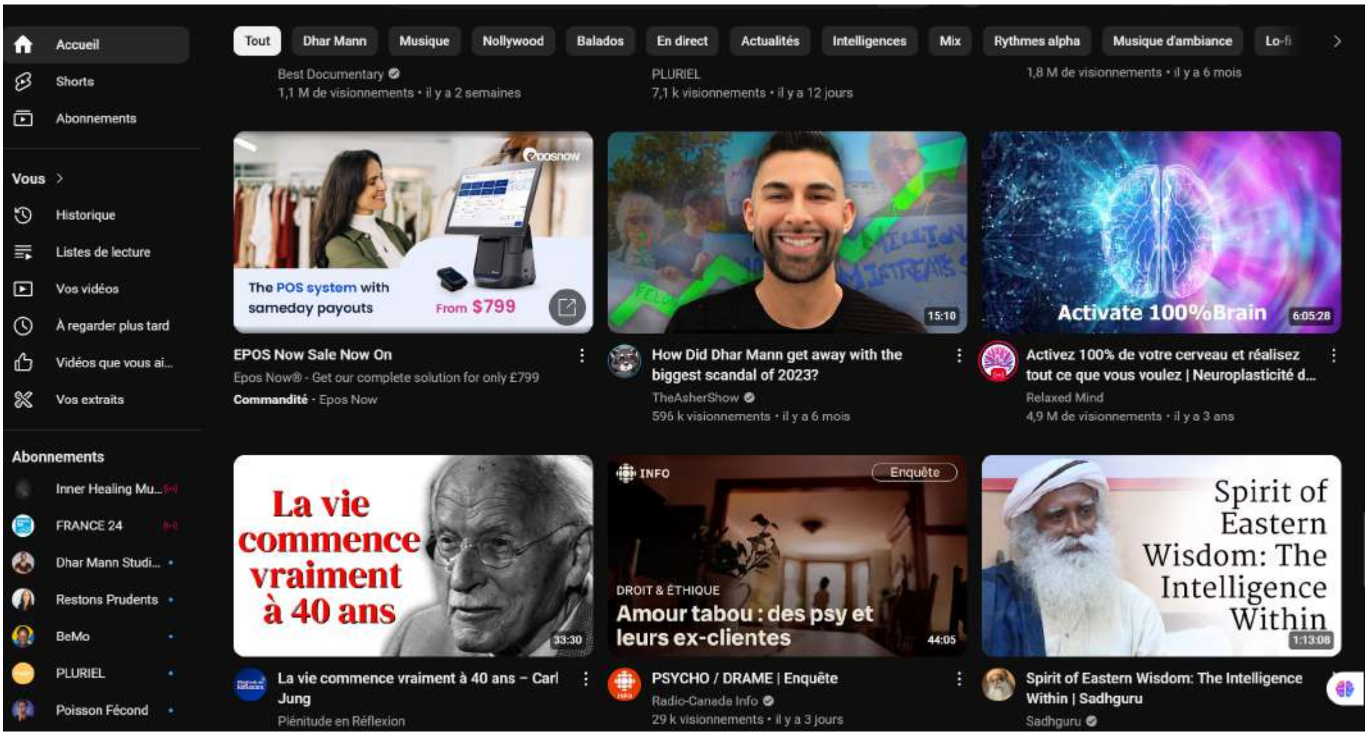
Si tu cherches à étudier la notion de justice dans ces trois textes, alors le Texte 1 est celui qui semble le plus pertinent, suivi par le Texte 3. Cela montre comment les mathématiques peuvent aider à orienter une lecture, même dans un domaine aussi subjectif que la philosophie.

Remarque : Dans la réalité, l'analyse textuelle va plus loin : elle tient compte du contexte, des synonymes, des relations sémantiques et du style de l'auteur. Mais TF-IDF est déjà un bon point de départ.

11. Maths dans la recommandation des vidéos

Comment Netflix ou YouTube devinent ce qu'on aime ?

Tu as sûrement remarqué que Netflix ou YouTube te proposent des vidéos que tu as très envie de regarder, souvent sans même que tu les cherches. Mais comment font-ils ? Est-ce qu'ils lisent dans nos pensées ? Non... ce sont des mathématiques qui travaillent discrètement en arrière-plan.



Question

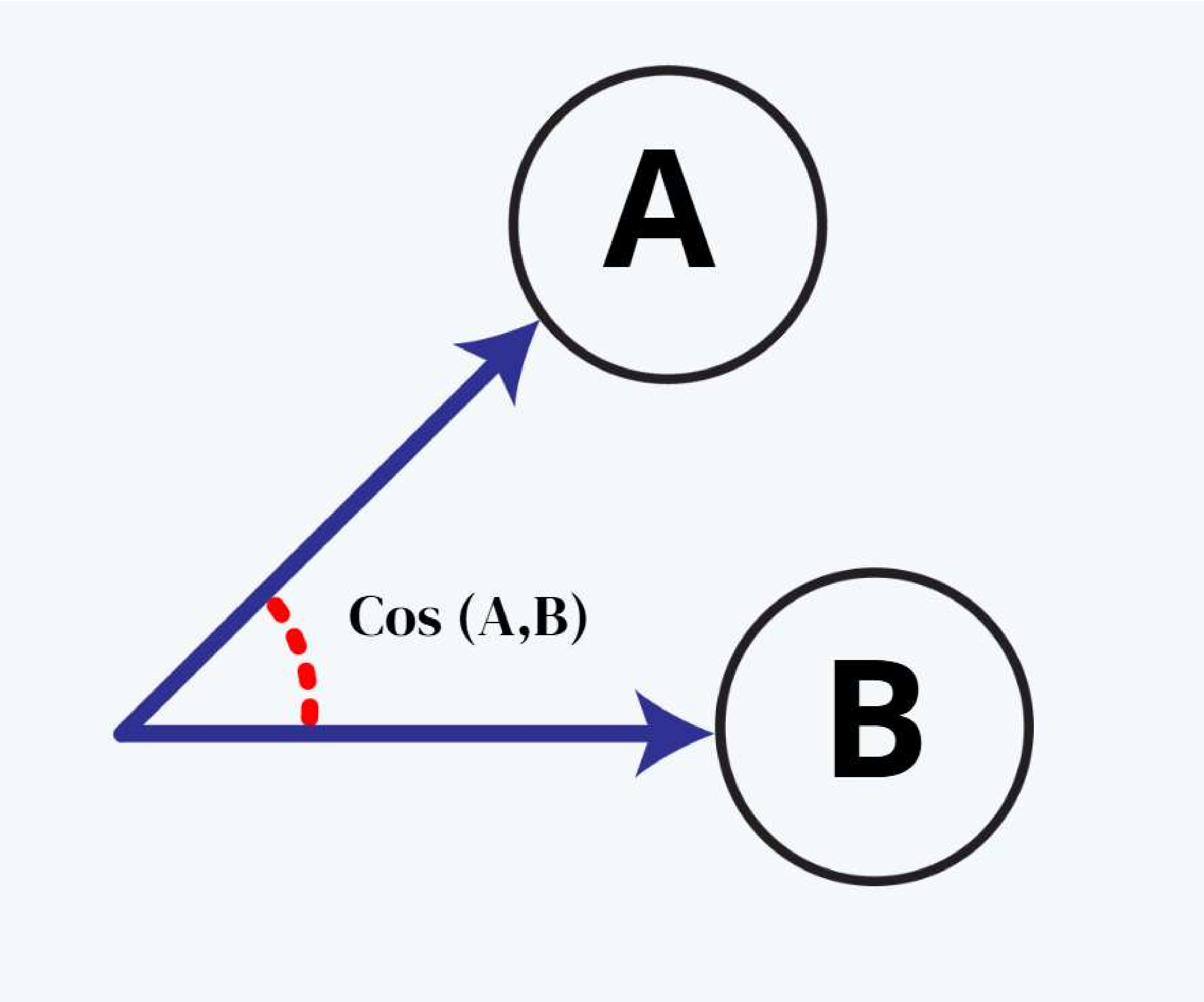
Comment ces plateformes savent-elles exactement quelles vidéos te recommander ?

Ce qui se cache derrière : des algorithmes de recommandation

Réponse

Derrière chaque suggestion de vidéo se cache un algorithme mathématique. Il se base sur tes préférences... mais aussi sur celles d'autres utilisateurs qui te ressemblent dans leurs goûts.

L'un des outils les plus simples pour modéliser cela s'appelle la similarité cosinus. C'est une manière de mesurer à quel point deux personnes partagent des goûts similaires.



L'algorithme : la similarité cosinus

Imaginons que chaque utilisateur est représenté par un vecteur, qui contient ses notes (ou temps de visionnage) pour différentes vidéos.

On calcule alors :

$$\text{Similitude}(A, B) = \frac{A \cdot B}{\|A\| \|B\|} = \cos(A, B)$$

Où :

– $A \cdot B$ est le produit scalaire entre les deux vecteurs

– $\|A\|$ et $\|B\|$ sont les longueurs (ou normes) des vecteurs

– $\cos(A, B)$ mesure l'angle entre ces deux vecteurs et leur alignement. Plus l'angle est petit, plus les vecteurs sont alignés dans la même direction, et plus le cosinus est proche de 1.

Exemple concret

Deux utilisateurs ont noté trois vidéos :

$$A = [4, 5, 3], \quad B = [5, 4, 3]$$

On calcule :

$$\text{Similitude}(A, B) = \frac{(4 \times 5) + (5 \times 4) + (3 \times 3)}{\sqrt{4^2 + 5^2 + 3^2} \times \sqrt{5^2 + 4^2 + 3^2}} = \frac{49}{\sqrt{50} \times \sqrt{50}} = 0{,}98$$

Que signifie ce 0,98 ?

Interprétation

Une similarité de 0,98 signifie que les préférences de A et B sont très proches.
L'algorithme va donc recommander à A les vidéos que B a bien notées — même si A ne les a jamais vues.

Ce principe permet à Netflix, YouTube ou Spotify de proposer un contenu qui semble "deviner" ce que tu aimes.

Et au-delà des vidéos ?

Autres applications

Ce type de raisonnement par similarité ne sert pas que pour les vidéos :

– Sur les sites e-commerce, il permet de recommander des produits.

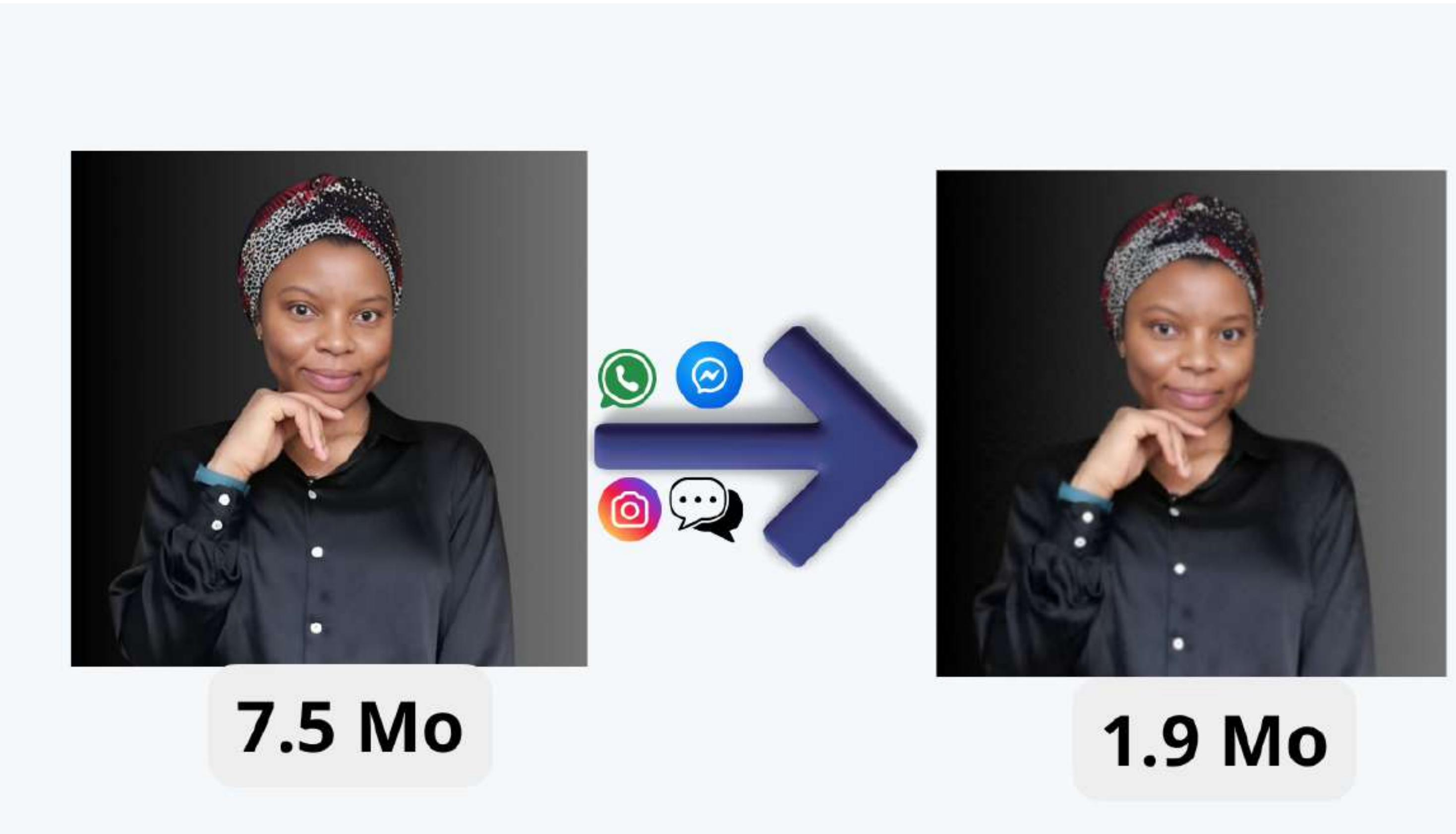
– Il est aussi utilisé pour comparer des articles ou suggérer du contenu sur les sites d'actualité.

– Et même dans les moteurs de recherche pour classer les résultats.

Conclusion : Tu ne le vois pas, mais chaque clic, chaque pause ou chaque like devient une information mathématiquement traitée pour affiner ce qu'on te montre. *La personnalisation numérique, c'est des maths au service de tes habitudes.*

12. Comment les applications compressent les images ?

Contexte : Deux images à comparer



Comprendre le mécanisme

Question

Comment les applications compressent-elles vos images pour un envoi rapide tout en préservant une qualité acceptable ?

Les maths à l'œuvre : La décomposition en valeurs singulières (SVD) est une technique mathématique qui permet de réduire la quantité de données tout en conservant l'essentiel des détails visuels. Cette méthode permet d'envoyer des images plus rapidement tout en minimisant la perte de qualité.

Trois éléments clés sont utilisés :

– La matrice d'image

– La décomposition SVD

– La réduction du rang

Qu'est-ce qu'une matrice d'image ?

Une image peut être représentée comme une matrice, où chaque élément correspond à un pixel, et sa valeur indique son intensité lumineuse.

10	70	20
60	80	40
30	90	50

Cette petite matrice 3×3 représente une image très simplifiée. Les chiffres indiquent ici l'intensité lumineuse de chaque pixel (niveaux de gris).

La décomposition en valeurs singulières (SVD)

La SVD est une méthode de factorisation matricielle qui décompose la matrice A représentant l'image en trois matrices :

$$A = U\Sigma V^T$$

– U : contient les vecteurs singuliers à gauche (structure par lignes)

– Σ : matrice diagonale avec les valeurs singulières

– V^T : vecteurs singuliers à droite (structure par colonnes)

Que sont les valeurs singulières ? Ce sont les racines carrées des valeurs propres de $A^T A$. Les matrices U et V contiennent les vecteurs propres correspondants.

Exemple simple : matrice 3×3

Soit la matrice :

$$A = \begin{pmatrix} 10 & 70 & 20 \\ 60 & 80 & 40 \\ 30 & 90 & 50 \end{pmatrix}$$

Sa décomposition SVD donne :

$$U = \begin{pmatrix} -0.4249 & 0.6406 & -0.6396 \\ -0.6365 & -0.7138 & -0.2921 \\ -0.6436 & 0.2830 & 0.7111 \end{pmatrix}, \quad \Sigma = \begin{pmatrix} 165.52 & 0 & 0 \\ 0 & 30.94 & 0 \\ 0 & 0 & 12.11 \end{pmatrix}$$

$$V^T = \begin{pmatrix} -0.3731 & -0.8373 & -0.3996 \\ -0.9029 & 0.4268 & -0.0514 \\ -0.2136 & -0.3416 & 0.9152 \end{pmatrix}$$

Comprendre les valeurs singulières

– 165.52 : composante la plus importante

– 30.94 : significative

– 12.11 : plus faible, donc moins d'impact visuel

Interprétation : La majeure partie de l'information de l'image est contenue dans les premières valeurs singulières.

Réduction du rang et compression

Pour comprimer l'image, on peut ne conserver que les valeurs singulières les plus grandes :

$$\Sigma_k = \begin{pmatrix} 165.52 & 0 & 0 \\ 0 & 30.94 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Et on reconstruit une image approximative :

$$A_k = U\Sigma_k V^T = \begin{pmatrix} 8.35 & 67.36 & 27.09 \\ 59.25 & 78.79 & 43.24 \\ 31.84 & 92.94 & 42.12 \end{pmatrix}$$

Cette matrice A_k est une approximation de l'image d'origine, plus légère.

Pourquoi ça fonctionne ?

La compression par SVD marche bien parce que les images comportent des redondances visuelles.

En supprimant les composantes les moins importantes (valeurs singulières faibles), on réduit le fichier sans trop dégrader la qualité.

Remarque : Plus on garde de valeurs singulières, meilleure sera la qualité — mais la taille du fichier sera aussi plus grande.

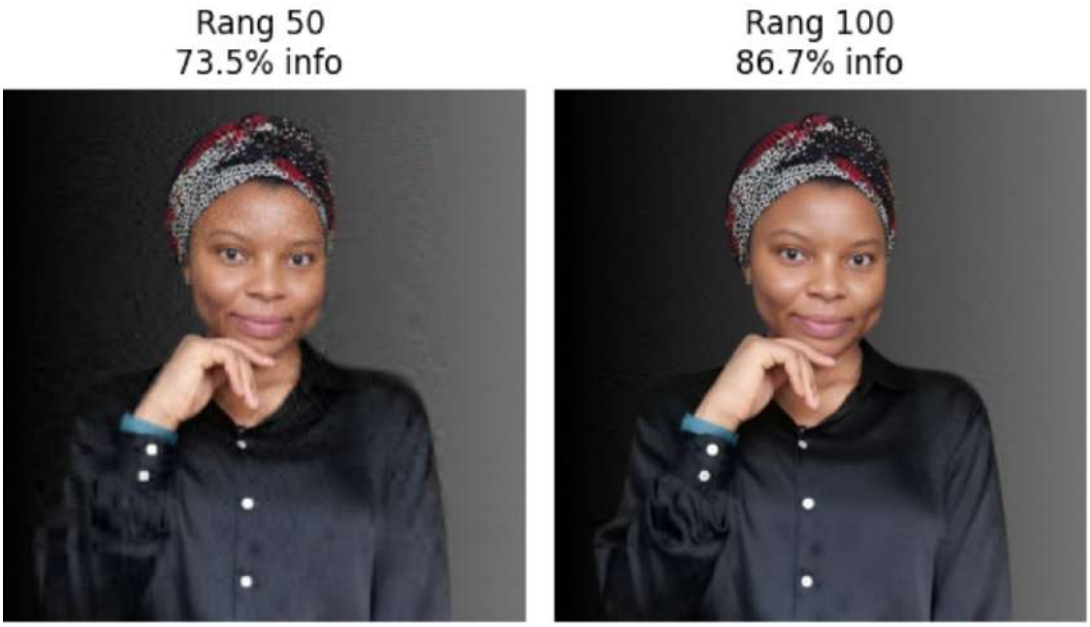
Code Python

```
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
from PIL import Image

# Charger l'image en couleur (RGB)
image_path = "Clo.jpeg"  # Remplace par ton fichier
image = Image.open(image_path).convert("RGB")  # Charger en couleur
image = np.array(image, dtype=np.float32)  # Convertir en matrice NumPy

# Dimensions de l'image
rows, cols, channels = image.shape

# Taille réelle de l'image originale en mémoire (en Ko)
original_size = image.nbytes / 1024
print(f"Taille de l'image originale : {original_size:.2f} Ko")

# Fonction pour appliquer la SVD et reconstruire chaque canal couleur
def compress_svd(image, k):
    compressed_channels = []
    for i in range(3):  # R, G, B
        U, S, Vt = np.linalg.svd(image[:, :, i], full_matrices=False)  # SVD sur chaque canal
        Sk = np.zeros((k, k))  # Matrice diagonale réduite
        np.fill_diagonal(Sk, S[:k])  # Conserver les k premières valeurs singulières
        Uk = U[:, :k]  # Conserver les k premiers vecteurs singuliers
        Vtk = Vt[:k, :]  # Conserver les k premiers vecteurs singuliers
        compressed_channel = np.dot(Uk, np.dot(Sk, Vtk))  # Reconstruction
        compressed_channels.append(compressed_channel)

    # Recombiner les 3 canaux couleur
    compressed_image = np.stack(compressed_channels, axis=2)
```

Exemple visuel avec compression

Taille de l'image originale : 7500.00 Ko

Nombre de valeurs singulières (k)	Taille estimée (Ko)	Réduction (%)	Info conservée (%)
Image originale	7500.000000	0.000000	100.000000
5	93.808594	98.749219	40.612946
20	375.234375	94.996875	57.357067
50	938.085938	87.492188	73.486572
100	1876.171875	74.984375	86.667274



Conclusion : L'importance des maths dans la technologie

Pourquoi c'est utile ?

– Permet d'envoyer des images plus rapidement

– Réduit l'espace de stockage sur les serveurs

– Maintient une bonne qualité d'image

Autres applications :

– Compression de vidéos (YouTube, Netflix)

– Réduction de dimension en intelligence artificielle

– Analyse d'images médicales et astronomiques

13. Détection des anomalies du cycle menstruel

Le contexte

Question

Comment utiliser les mathématiques pour détecter les anomalies dans le cycle menstruel sur plusieurs mois ?

Les maths à l'œuvre : Les mathématiques permettent d'analyser les écarts par rapport à une durée de cycle typique en utilisant deux outils statistiques fondamentaux : **la moyenne** et **l'écart type**.

Deux éléments clés sont utilisés :

– Le calcul de la moyenne

– Le calcul de l'écart type sur plusieurs mois

Calcul de la durée moyenne et de l'écart type

Pour détecter les anomalies dans le cycle menstruel, on calcule la durée moyenne du cycle, notée μ, et l'écart type, noté σ, sur plusieurs mois.

1. Durée moyenne du cycle :

$$\mu = \frac{1}{n}\sum_{i=1}^{n} D_i = \frac{D_1 + D_2 + \cdots + D_n}{n}$$

– D_i : durée du i-ème cycle

– n : nombre de cycles observés (par exemple, 6 mois)

2. Écart type :

$$\sigma = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(D_i - \mu)^2}$$

– σ mesure la variabilité des cycles autour de la moyenne

Exemple de calcul

Durées de cycle sur 6 mois : 28, 30, 27, 29, 31, 26 jours.

$$\mu = \frac{28+30+27+29+31+26}{6} = 28{,}5 \text{ jours}$$

Si on se limite à la moyenne, on pourrait croire que tout va bien dans le cycle, n'est-ce pas ?

$$\sigma = \sqrt{\frac{(28-28{,}5)^2+(30-28{,}5)^2+(27-28{,}5)^2+(29-28{,}5)^2+(31-28{,}5)^2+(26-28{,}5)^2}{5}} \approx 2 \text{ jours}$$

Cela signifie que la durée des cycles varie, en moyenne, de 2 jours autour de 28,5 jours.

Détection d'une anomalie

Une anomalie peut être détectée si un cycle sort de l'intervalle suivant :

$$[\mu - 2\sigma, \mu + 2\sigma]$$

Critères de détection :

– Cycle trop court : $D < \mu - 2\sigma$

– Cycle trop long : $D > \mu + 2\sigma$

Exemple : Avec $\mu = 28{,}5$ et $\sigma = 2$:

$$[\mu - 2\sigma, \mu + 2\sigma] = [24{,}5 \text{ jours}, 32{,}5 \text{ jours}]$$

– Un cycle de 24 jours serait considéré comme anormalement court

– Un cycle de 33 jours serait considéré comme anormalement long

Limites des méthodes statistiques

Bien que ces outils soient utiles, ils ont aussi leurs limites.

1. Limitations des modèles basés sur la moyenne et l'écart type :

– Les cycles varient naturellement, ce qui peut créer de fausses alertes

– Des cycles très courts ou très longs faussent les calculs

2. Absence de facteurs contextuels :

– Stress, alimentation, traitements médicaux peuvent modifier les cycles sans qu'il y ait d'anomalie clinique
– Ces modèles ne s'adaptent pas à la variabilité propre à chaque personne

Conclusion

Les méthodes statistiques comme la moyenne et l'écart type sont des outils utiles pour surveiller la régularité du cycle menstruel. Mais elles ont leurs limites, et pour aller plus loin, il faut croiser ces indicateurs avec des données contextuelles ou explorer des modèles plus personnalisés.

Les maths peuvent aider à mieux comprendre son corps, mais ne remplacent pas l'écoute, l'observation personnelle et l'accompagnement médical.

14. Fluctuations hormonales dans le cycle menstruel

Le contexte

Question

Comment les mathématiques, en particulier les fonctions sinus, peuvent-elles modéliser les fluctuations hormonales au cours du cycle menstruel ?

Les maths à l'œuvre : Les niveaux d'hormones comme l'œstrogène et la progestérone varient de manière cyclique. Ces variations peuvent être modélisées par des fonctions sinus pour représenter les hausses et les baisses au fil du temps.

Deux éléments clés sont utilisés :

– Les fonctions sinus pour modéliser les variations périodiques

– L'amplitude pour quantifier les niveaux hormonaux en pg/mL et ng/mL

Modélisation mathématique par une fonction sinus

Le cycle menstruel a une durée moyenne de 28 jours. On peut modéliser le niveau hormonal à l'aide de la fonction suivante :

$$H(t) = A \cdot \sin\left(\frac{2\pi}{T}t + \phi\right) + B$$

Signification des termes :

– $H(t)$: niveau de l'hormone au jour t

– A : amplitude (variation maximale)

– $T = 28$: période du cycle (en jours)

– ϕ : phase (moment où le cycle commence)

– B : niveau moyen autour duquel l'hormone varie

Exemple : Pour modéliser l'œstrogène avec :

$$A = 50 \text{ pg/mL}, \quad B = 150 \text{ pg/mL}, \quad \phi = 0$$

On obtient :

$$H(t) = 50 \cdot \sin\left(\frac{2\pi}{28}t\right) + 150$$

Cela signifie que l'œstrogène varie entre 100 et 200 pg/mL au cours du cycle.

Quantification jour après jour

1. Début du cycle (menstruation) – $t = 0$

$$H(0) = 50 \cdot \sin(0) + 150 = 150 \text{ pg/mL}$$

2. Jour 14 (ovulation) – $t = 14$

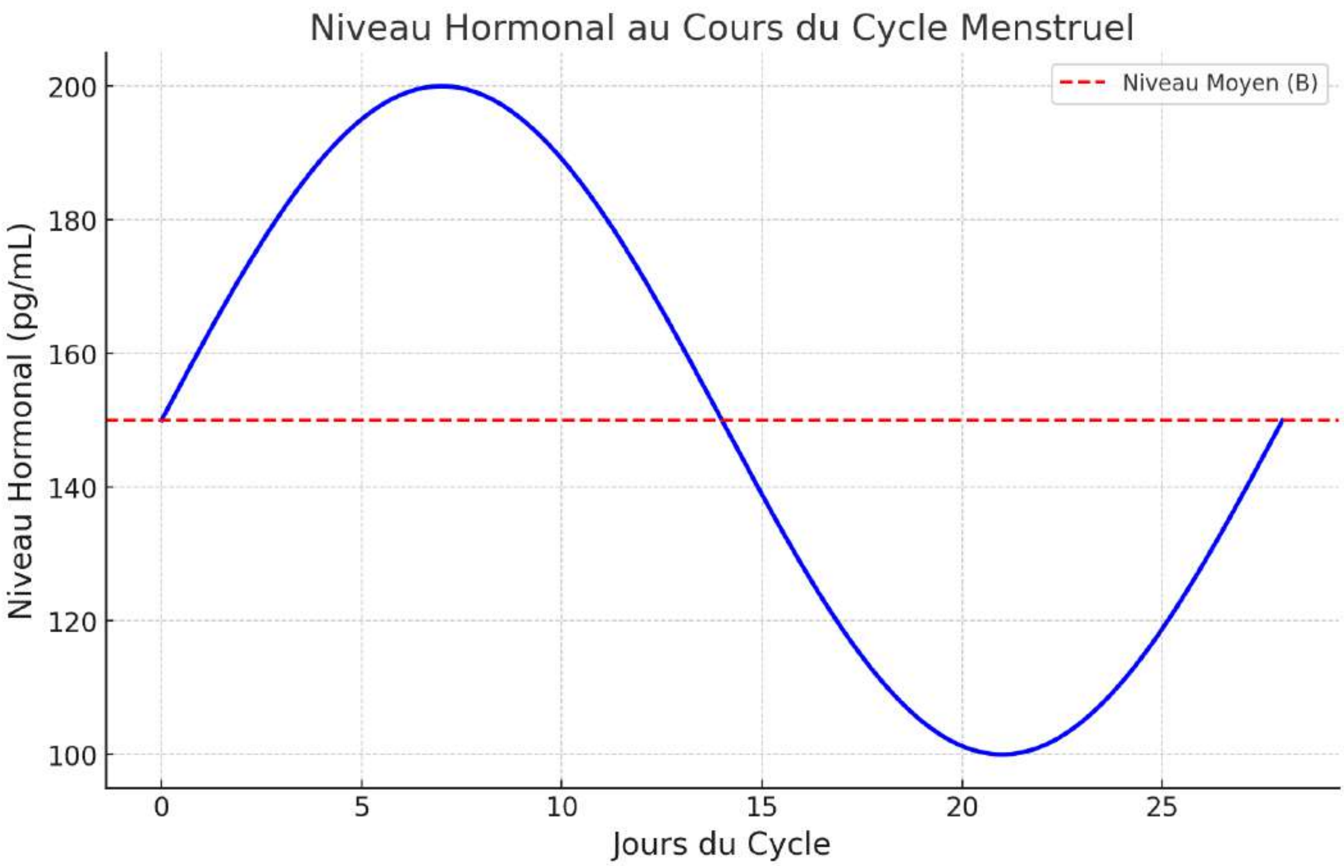
$$H(14) = 50 \cdot \sin(\pi) + 150 = 150 \text{ pg/mL}$$

3. Fin du cycle – $t = 28$

$$H(28) = 50 \cdot \sin(2\pi) + 150 = 150 \text{ pg/mL}$$

L'œstrogène monte, redescend, puis revient à son niveau moyen.

Visualisation du modèle



Cette courbe sinus montre clairement comment les niveaux hormonaux varient tout au long du cycle.

Applications pratiques

- Suivre le cycle menstruel avec précision
- Repérer les irrégularités hormonales
- Adapter les traitements ou interventions médicales

Critique du modèle sinusoïdal

1. Simplification excessive : Les hormones sont influencées par de nombreux facteurs comme le stress, l'alimentation ou la santé générale.

2. Variabilité individuelle : Chaque personne a un cycle unique. Ce modèle ne prend pas en compte les irrégularités, ni des cas comme le SOPK.

3. Modélisation statique : Le sinus suppose un cycle identique chaque mois, ce qui n'est pas toujours réaliste.

Conclusion

Les fonctions sinus permettent de visualiser et quantifier les fluctuations hormonales du cycle menstruel. Mais pour aller plus loin, il faut intégrer la variabilité individuelle, les événements contextuels, et éventuellement des modèles plus adaptatifs.

Une belle preuve que les maths peuvent éclairer notre compréhension du corps.

15. Comment l'IRM utilise les maths ?

Qu'est-ce que l'IRM ?

Question

Comment l'imagerie par résonance magnétique (IRM) permet-elle de voir l'intérieur du corps humain sans chirurgie ?

L'IRM est une technologie d'imagerie médicale qui utilise un champ magnétique puissant, des ondes radio, et des mathématiques pour créer des images détaillées des organes et des tissus internes.

Le principe de l'IRM : les bases physiques

Comment ça marche ?

– L'IRM utilise un puissant champ magnétique pour aligner les protons (les noyaux d'hydrogène dans l'eau) dans le corps humain.
– Une onde radio est ensuite envoyée, perturbant l'alignement des protons.
– Lorsque le champ radio est éteint, les protons reviennent à leur alignement initial, émettant des signaux.

Ces signaux sont captés par l'appareil IRM et transformés en images grâce à des mathématiques avancées.

Transformée de Fourier : le cœur de l'IRM

La transformée de Fourier est un outil mathématique essentiel en IRM, qui permet de convertir les signaux captés (dans le domaine fréquentiel) en images (dans le domaine spatial).

– *Domaine fréquentiel* : les signaux reçus par l'IRM sont d'abord dans ce domaine.

– *Domaine spatial* : la transformée de Fourier convertit ces signaux en images montrant les structures internes.

Formule de la Transformée de Fourier

Formule

$$F(f) = \int_{-\infty}^{\infty} f(t)e^{-2i\pi ft}\, dt$$

– $f(t)$: le signal dans le domaine temporel ou spatial
– $F(f)$: le signal transformé dans le domaine fréquentiel
– $e^{-2i\pi ft}$: la base des ondes sinusoïdales utilisées pour décomposer le signal

Comment l'IRM utilise ces transformations

Processus :

1. Excitation des protons : alignement et perturbation par onde radio.
2. Émission de signaux à différentes fréquences.
3. Captation des signaux numériques.
4. Transformée de Fourier pour construire une image.

Exemple visuel : construction d'une image IRM

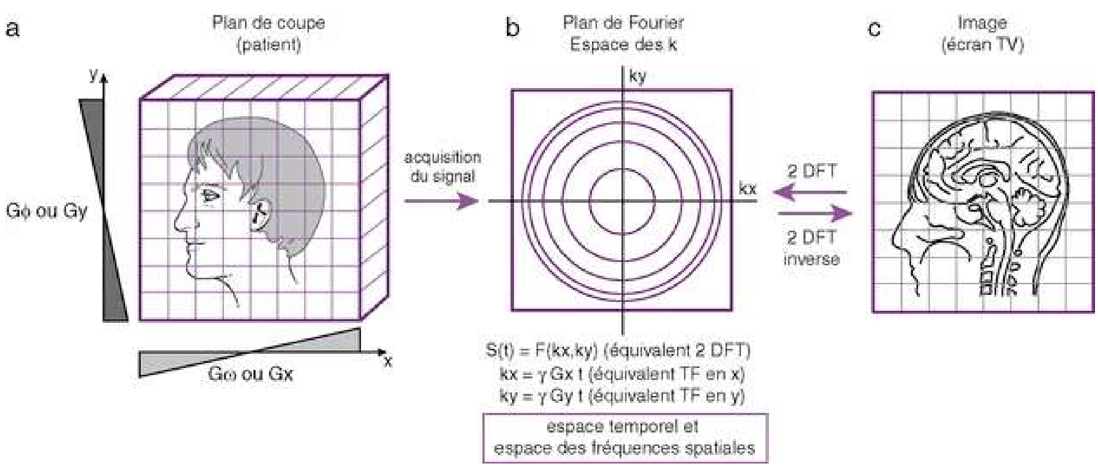


FIGURE 15.1 – *
Source : @Medecine Key

Interprétation : dans le plan de Fourier, les signaux complexes sont décomposés en composantes simples. Cela permet de reconstruire une image claire des tissus internes.

Applications pratiques de l'IRM

Pourquoi l'IRM est-elle si utile ?

– *Détection des tumeurs* : même invisibles par d'autres techniques.

– *Observation du cerveau* : détection d'anomalies neurologiques.

– *Analyse des tissus mous* : ligaments, muscles, organes...

Conclusion

L'IRM montre comment les mathématiques, notamment la transformée de Fourier, permettent de voir à l'intérieur du corps humain de manière non invasive.

Grâce à ces outils, les spécialistes peuvent diagnostiquer des maladies, planifier des traitements et mieux comprendre la physiologie humaine – sans avoir besoin de chirurgie.

16. Optimiser la chimiothérapie

Figure 16.1 – Illustration du traitement du cancer et de la chimiothérapie

Problème

Comment les mathématiques peuvent-elles nous aider à modéliser l'efficacité de la chimiothérapie pour détruire les cellules cancéreuses tout en minimisant les effets secondaires ?

Pour répondre à cette question, les mathématiques proposent des modèles basés sur les **équations différentielles**. Ces modèles permettent de décrire la dynamique des cellules cancéreuses et des cellules saines sous l'effet de la chimiothérapie, afin d'optimiser le dosage et améliorer le traitement.

Modélisation des cellules cancéreuses

La croissance des cellules cancéreuses $C(t)$ est souvent modélisée par une équation logistique modifiée :

$$\frac{dC(t)}{dt} = rC(t)\left(1 - \frac{C(t)}{K}\right) - d_C C(t) - e_C(t)C(t)$$

où :

— $C(t)$: nombre de cellules cancéreuses au temps t,

— r : taux de croissance des cellules cancéreuses,

— K : capacité maximale (carrying capacity),

— d_C : taux de mortalité naturelle des cellules cancéreuses,

— $e_C(t)$: effet du traitement (efficacité de la chimiothérapie au temps t).

Effet de la chimiothérapie sur les cellules saines

Le traitement affecte aussi les cellules saines $S(t)$, modélisées par une autre équation différentielle :

$$\frac{dS(t)}{dt} = r_S S(t) - d_S S(t) - e_S(t)S(t)$$

où :

— $S(t)$: nombre de cellules saines au temps t,

— r_S : taux de régénération des cellules saines,

— d_S : taux de mortalité naturelle des cellules saines,

— $e_S(t)$: effet indésirable du traitement (toxicité).

Optimisation du dosage

L'objectif est de trouver une fonction de dosage $e_C(t)$ qui maximise la destruction des cellules cancéreuses tout en minimisant les dégâts sur les cellules saines.

Objectif mathématique

$$\min_{e_C(t)} \left(\int_0^T S(t)\,dt - \int_0^T C(t)\,dt \right)$$

$$\text{sous contrainte : } e_C(t) \geq 0, \quad e_S(t) = \alpha e_C(t)$$

où α est un paramètre mesurant la toxicité du traitement et T la durée totale du protocole.

Exemple de calcul

Supposons que l'on applique une dose constante $e_C(t) = e_0$. Le système devient :

$$\frac{dC(t)}{dt} = rC(t) - (d_C + e_0)C(t)$$
$$\frac{dS(t)}{dt} = (r_S - d_S - \alpha e_0)S(t)$$

En résolvant ce système, on peut observer comment les populations de cellules évoluent et ajuster e_0 pour obtenir un compromis entre efficacité et tolérance.

Applications et utilité

Ces modèles permettent :

— d'**optimiser le dosage** de la chimiothérapie,

— de **prédire la réponse du patient** en fonction du protocole choisi,

— de **personnaliser le traitement** selon les caractéristiques biologiques du patient.

Critique des hypothèses

Limites du modèle

— **Croissance logistique** : La croissance tumorale est simplifiée, alors qu'elle dépend de multiples facteurs biologiques.

— **Efficacité constante** : Le modèle suppose un effet constant de la chimiothérapie, ce qui ne reflète pas la réalité (résistance, métabolisme...).

— **Toxicité linéaire** : L'impact sur les cellules saines est modélisé de façon linéaire, alors qu'il est souvent non linéaire et variable d'un patient à l'autre.

Conclusion

Grâce aux équations différentielles, il est possible de modéliser l'effet de la chimiothérapie sur le corps humain et de guider les médecins dans le choix du dosage. Ces approches mathématiques contribuent à sauver des vies, en améliorant les chances de guérison tout en réduisant les effets secondaires.

17. Optimiser la gestion des entrepôts

Introduction : le défi de l'optimisation

Question

Comment organiser un entrepôt pour minimiser le temps passé à récupérer les articles et maximiser l'efficacité des opérations ?

Les mathématiques, en particulier les outils d'optimisation, permettent d'améliorer l'agencement des stocks et les itinéraires des préparateurs de commandes.

Optimisation des stocks : l'analyse ABC

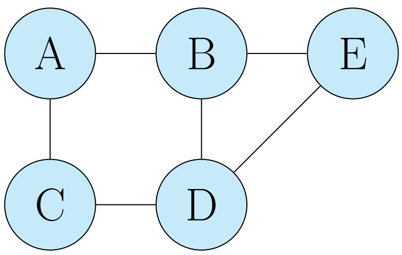
L'analyse ABC classe les articles selon leur importance :

- **Classe A** : Articles les plus demandés (20% des articles mais 80% des mouvements)
- **Classe B** : Articles moyennement demandés
- **Classe C** : Articles les moins demandés

Idée simple : placer les articles de classe A près de la zone de préparation pour réduire le temps de collecte.

Exemple : organisation d'un entrepôt

Considérons un petit entrepôt avec 5 articles : A, B, C, D et E.



Les distances entre les nœuds sont de 1 unité, sauf pour $D \leftrightarrow E$, qui vaut 2 unités.

Optimiser les itinéraires : le problème du voyageur de commerce (TSP)

Une fois les articles bien placés, il reste à optimiser le trajet du préparateur de commandes.

Objectif : minimiser la distance pour collecter tous les articles d'une commande. **Modèle** : problème du voyageur de commerce (TSP).

Exemple : une commande demande les articles A, C et E.

Meilleur itinéraire pour A, C et E

On commence toujours par A. Voici trois options :

Option 1 : $A \to C \to D \to E$

$$\text{Distance} = d(A, C) + d(C, D) + d(D, E) = 1 + 1 + 2 = 4$$

Option 2 : $A \to B \to E \to D \to C$

$$\text{Distance} = 1 + 1 + 2 + 1 = 5$$

Option 3 : $A \to C \to D \to B \to E$

$$\text{Distance} = 1 + 1 + 1 + 1 = 4$$

Conclusion : Les options 1 et 3 sont les plus efficaces (4 unités chacune).

Limites du modèle

Limites pratiques

Même si ce modèle est utile, il a ses faiblesses :

- Il suppose des distances fixes et linéaires, ce qui ignore les obstacles ou la congestion.
- Il se base sur l'historique, mais pas sur les variations saisonnières.
- Il devient vite trop complexe à calculer pour de grands entrepôts.
- Il nécessite des outils numériques (WMS) et de la formation, ce qui peut coûter cher.

Conclusion

Les mathématiques appliquées à la logistique permettent de mieux organiser les entrepôts : moins de déplacements, plus de rapidité, et une meilleure satisfaction client.

Mais comme souvent, **la théorie a besoin de s'adapter aux réalités du terrain** pour être pleinement efficace.

18. Les maths derrière les mouvements des robots !

Introduction : Le défi des mouvements robotiques

Question

Les robots modernes accomplissent des tâches complexes avec précision. Comment leurs bras parviennent-ils à se déplacer de manière aussi fluide et ciblée ?

La réponse réside dans les mathématiques : Plus précisément, dans la **cinématique**, une branche qui permet de calculer avec précision les mouvements et les positions des bras robotiques.

Le pouvoir de la cinématique

Les mouvements des robots sont modélisés grâce à la **cinématique**.

La cinématique étudie les mouvements en fonction du temps, sans tenir compte des causes physiques (comme les forces ou les masses).

En utilisant des notions comme la **cinématique directe** et la **cinématique inverse**, les ingénieurs peuvent prédire et contrôler la position exacte de l'extrémité du bras avec une extrême précision.

Exemple : contrôler un bras robotique

Dans un projet de robotique, on cherche à contrôler un bras articulé composé de deux segments.

Problème : Comment orienter les deux articulations pour que l'extrémité du bras atteigne une position cible dans l'espace ?

Cinématique directe : calculer la position à partir des angles

On considère :

– deux articulations avec angles θ_1 et θ_2

– des segments de longueurs $L_1 = 30\text{cm}$ et $L_2 = 20$ cm

Formules de la position :

$$x = L_1 \cos(\theta_1) + L_2 \cos(\theta_1 + \theta_2)$$
$$y = L_1 \sin(\theta_1) + L_2 \sin(\theta_1 + \theta_2)$$

– x : Position horizontale de l'extrémité du bras

– y : Position verticale de l'extrémité du bras

Exemple numérique :

Si on choisit :

$$\theta_1 = 45^\circ, \quad \theta_2 = 30^\circ$$

Alors on obtient :

$$x \approx 26{,}39 \text{ cm}, \quad y \approx 33{,}46 \text{ cm}$$

Conclusion : Le bras atteint une position précise, calculée à partir des deux angles.

Conclusion : les maths au service des robots

Les bras robotiques peuvent accomplir des tâches très complexes... parce qu'ils obéissent à des calculs mathématiques.

La cinématique permet de prédire et de contrôler les mouvements en temps réel, avec efficacité et fluidité.

19. Maximiser la diffusion de l'information

Introduction : le défi de la diffusion d'information

Question

Dans un réseau social, certaines personnes jouent un rôle central dans la diffusion de l'information. Comment les identifier pour maximiser l'impact d'une publicité ou d'un message ?

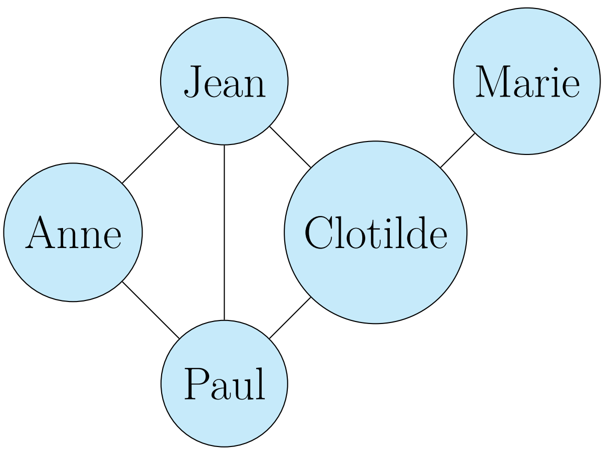
Les mathématiques, via la théorie des graphes, offrent des outils pour analyser et optimiser ces réseaux de communication.

La théorie des graphes : exemple avec 5 personnes

La théorie des graphes est un domaine des mathématiques qui modélise les relations entre des objets. Dans le contexte des réseaux sociaux :

– Chaque personne est un **nœud**

– Chaque relation (amitié, abonnement) est une **arête**



Ce graphe simple montre les connexions entre cinq personnes dans un réseau.

Excentricité et Proximité

- **L'excentricité** ($e(v)$) : Somme des distances les plus courtes entre un nœud v et tous les autres nœuds du réseau.
- **La proximité** ($p(v)$) : Inverse de l'excentricité, elle mesure l'importance du nœud dans le réseau.

Par exemple, pour le nœud Jean, on calcule :

$$e(\text{Jean}) = d(\text{Jean}, \text{Anne}) + d(\text{Jean}, \text{Clotilde}) + d(\text{Jean}, \text{Paul}) + d(\text{Jean}, \text{Marie})$$

où $d(\cdot, \cdot)$ désigne la distance minimale dans le graphe.

$$p(\text{Jean}) = \frac{1}{e(\text{Jean})}$$

Calcul pour Jean

$$\begin{aligned} e(\text{Jean}) &= d(\text{Jean}, \text{Anne}) + d(\text{Jean}, \text{Clotilde}) + d(\text{Jean}, \text{Paul}) + d(\text{Jean}, \text{Marie}) \\ &= 1 + 1 + 1 + 2 = 5 \end{aligned}$$

Donc :

$$p(\text{Jean}) = \frac{1}{5} = 0{,}20$$

On peut maintenant comparer cela aux autres personnes du réseau.

Proximité de chaque nœud

Personne	**Excentricité**	**Proximité**
Anne	6	$\frac{1}{6} \approx 0.17$
Jean	5	$\frac{1}{5} = 0.20$
Clotilde	7	$\frac{1}{7} \approx 0.14$
Paul	6	$\frac{1}{6} \approx 0.17$
Marie	8	$\frac{1}{8} = 0.125$

On observe que **Jean** a la plus grande proximité, ce qui en fait le meilleur candidat pour diffuser une information.

Interprétation des résultats

Le tableau montre que Jean, ayant l'excentricité la plus faible et la plus grande proximité, est le **nœud central du réseau**. Il est donc le candidat idéal pour maximiser la portée d'un message.

Conclusion et applications

La théorie des graphes est un outil puissant pour :

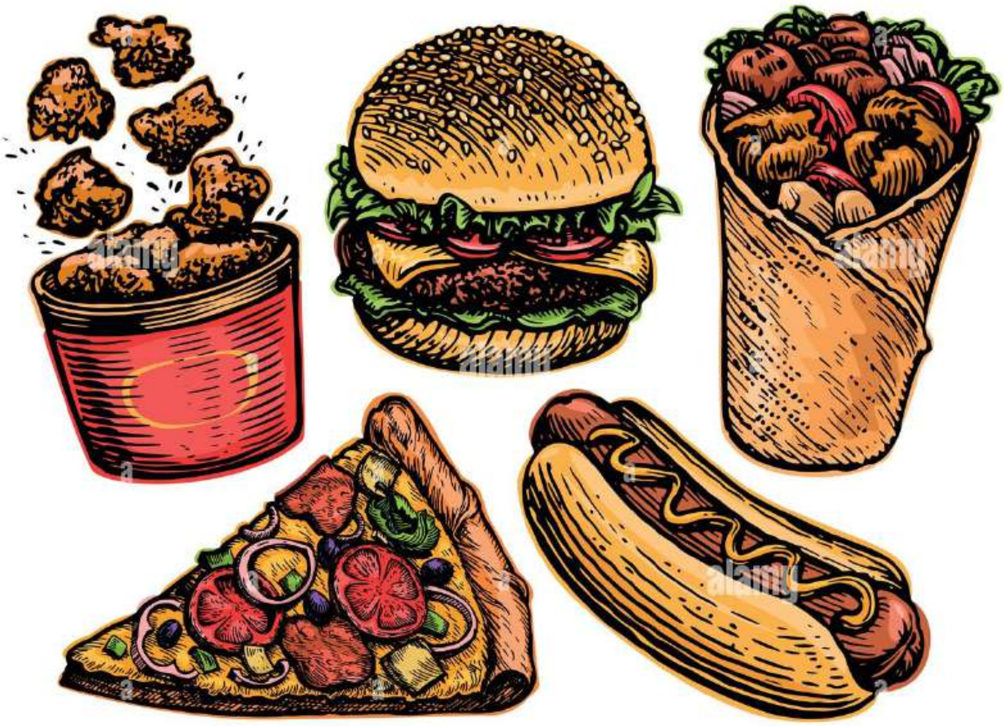
– analyser les structures de réseaux sociaux

– identifier les nœuds clés pour la diffusion de l'information

– cibler les meilleures personnes dans une stratégie de marketing

Challenge : Pouvez-vous calculer l'excentricité et la proximité de vos connexions sur LinkedIn ? Qui sont vos influenceurs invisibles ?

20. Les commandes dans un fast-food



Le contexte

Question

Comment utiliser les chaînes de Markov pour prédire la demande de repas (pizza, kebab, burger) dans un fast-food ?

Les maths à l'œuvre : On utilise une chaîne de Markov pour modéliser les transitions entre différents types de repas commandés. L'objectif : prédire la demande future à partir des données historiques.

États considérés : Pizza, Kebab, Burger

Modélisation de la demande

La chaîne de Markov est modélisée par une matrice de transition P, où chaque élément représente la probabilité de passer d'un type de commande à un autre.

- **État 1 :** Pizza
- **État 2 :** Kebab
- **État 3 :** Burger

Matrice de transition :

$$P = \begin{pmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{pmatrix}$$

Chaque élément p_{ij} représente la probabilité qu'un client passe de la commande i à la commande j.

Calcul des probabilités de transition

Pour déterminer les valeurs de la matrice P, on analyse les données historiques. Par exemple, si 100 clients ont commandé une pizza, et que :

- 40 reprennent une pizza,
- 30 changent pour un kebab,
- 30 pour un burger,

alors :

$$p_{11} = \frac{40}{100}, \quad p_{12} = \frac{30}{100}, \quad p_{13} = \frac{30}{100}$$

Les autres lignes sont calculées de la même manière.

Exemple complet de matrice de transition

Supposons que la matrice P soit :

$$P = \begin{pmatrix} 0.6 & 0.2 & 0.2 \\ 0.3 & 0.4 & 0.3 \\ 0.2 & 0.1 & 0.7 \end{pmatrix}$$

Prévision de la demande future

La demande future est calculée en multipliant la distribution actuelle des commandes par la matrice de transition.

Distribution actuelle :

$$v_0 = \begin{pmatrix} 0.5 \\ 0.3 \\ 0.2 \end{pmatrix}$$

Cela signifie : 50 pizza, 30% kebab, 20% burger.

Demande prédite (le lendemain) :

$$v_1 = P \cdot v_0 = \begin{pmatrix} 0.47 \\ 0.32 \\ 0.21 \end{pmatrix}$$

Interprétation des résultats

– 47% des clients commanderont une pizza

– 32% commanderont un kebab

– 21% un burger

Cette information permet d'ajuster les quantités à préparer pour optimiser les stocks.

Limites du modèle

Points faibles :

– Le modèle suppose que les probabilités sont fixes

– Il dépend entièrement des données passées

– Il ne réagit pas aux changements soudains

– Il ne prend pas en compte des facteurs extérieurs (météo, promo, etc.)

Pistes d'amélioration : Utiliser des modèles plus dynamiques comme les simulations stochastiques ou les modèles bayésiens.

Conclusion

Les chaînes de Markov permettent :

– de prédire la demande,

– de mieux gérer les stocks,

– de limiter le gaspillage alimentaire.

Une preuve concrète que les maths peuvent aider… même à mieux gérer les burgers et les pizzas !

21. Dans la tête d'un PDG

Scénario : vous êtes à la tête d'une entreprise. . .

Scénario

Vous êtes le PDG de l'entreprise A, une grande société. Votre principal concurrent, l'entreprise B, est sur le point de prendre une décision stratégique cruciale. Que faites-vous ?

Deux options :

– Coopérer avec votre rival (en maintenant les prix ou en sortant un produit au même moment)

– Le défier (en baissant les prix pour gagner des parts de marché)

La matrice des gains : entreprises A et B

La matrice suivante montre les résultats possibles selon les décisions prises par les deux entreprises :

	B : Coopérer	**B : Défier**
A : Coopérer	(500, 500)	(0, 750)
A : Défier	(750, 0)	(250, 250)

Interprétation :

– (500, 500) : coopération mutuelle, marché partagé

– (0, 750) : A coopère, B défie → B gagne gros

– (750, 0) : A défie, B coopère → A gagne gros

– (250, 250) : les deux défient → pertes d'opportunité

Le dilemme stratégique

Le choix stratégique

Quelle stratégie allez-vous choisir pour maximiser les gains de votre entreprise ?

– Si vous coopérez, vous avez des gains stables (500).
– Si vous défiez, vous prenez un risque : vous pouvez gagner 750, mais aussi tomber à 250.

Remarque : La meilleure stratégie dépend souvent de ce que l'autre décide de faire. C'est tout le cœur de la théorie des jeux.

L'équilibre de Nash

Définition : L'équilibre de Nash est atteint lorsqu'aucun joueur n'a intérêt à changer sa stratégie, tant que l'autre ne change pas la sienne.

Exemple dans notre cas :

– Si A pense que B va coopérer, A peut avoir intérêt à défier.
– Mais si B anticipe cela, B défiera aussi.
– Alors A aussi devra défier, pour ne pas perdre de parts.

La meilleure stratégie

L'équilibre de Nash se produit quand les deux défient (250, 250). Aucune entreprise ne peut faire mieux en changeant seule de stratégie. C'est un équilibre, même s'il n'est pas optimal globalement.

Exemple concret :

– Si A augmente ses prix, B en profite pour gagner plus → A perd.
– Si A baisse ses prix alors que B aussi, tout le monde perd un peu, mais reste stable.

Conclusion : L'art de la stratégie

La théorie des jeux vous aide à :

– anticiper les décisions des autres,

– prendre des décisions plus rationnelles,

– éviter les pertes en coopérant de manière stratégique.

Dans les affaires, en politique, ou dans la vie quotidienne, la stratégie est une affaire… de maths.

22. Les maths les inondations ?

FIGURE 22.1 – L'eau s'écoule... mais parfois déborde. Et les maths nous aident à comprendre pourquoi.

Question de départ

Problème

Comment les mathématiques peuvent-elles nous aider à modéliser l'écoulement de l'eau et à anticiper les inondations à l'aide d'équations aux dérivées partielles ?

Réponse : Nous utilisons les **équations de Saint-Venant**, qui modélisent l'écoulement de l'eau en fonction de la vitesse, du débit et de la hauteur d'eau.

Modélisation de l'écoulement : équations de Saint-Venant

Les équations de Saint-Venant sont des équations aux dérivées partielles (EDP) qui modélisent l'écoulement de l'eau dans les rivières, les canaux ou les surfaces urbaines.

Équation de continuité :

$$\frac{\partial H}{\partial t} + \frac{\partial (Hv)}{\partial x} = 0$$

Équation de quantité de mouvement :

$$\frac{\partial v}{\partial t} + v\frac{\partial v}{\partial x} + g\frac{\partial H}{\partial x} = 0$$

Où :

- $H(x,t)$: hauteur de l'eau à un point x au temps t
- $v(x,t)$: vitesse de l'écoulement
- g : accélération de la gravité
- $\partial H/\partial t$, $\partial v/\partial x$: dérivées partielles temporelles et spatiales

Lien avec les inondations

Quand $H(x,t)$ devient trop élevé (fortes pluies, obstacles, mauvaise évacuation), une inondation peut se produire.

Facteurs aggravants

- Précipitations intenses $\rightarrow$ augmentation rapide de $H(x,t)$
- Obstacles urbains (routes, murs, bâtiments)
- Réseaux de drainage insuffisants

Exemple de calcul simple

Supposons qu'une pluie continue ajoute progressivement de l'eau :

$$H(x,t) = H_0 + \int_0^t P(\tau)\, d\tau$$

où :

- H_0 : hauteur initiale de l'eau
- $P(t)$: précipitations à l'instant t

Quand $H(x,t)$ dépasse un certain seuil, une inondation est probable.

Utilité des équations de Saint-Venant

Ces équations permettent de :

- **Prévoir les crues** dans les zones à risque
- **Tester des scénarios** en cas d'orage ou de saturation
- **Évaluer l'impact des infrastructures** sur le comportement de l'eau

Conclusion

Les équations de Saint-Venant permettent de modéliser l'écoulement de l'eau, de prévoir les crues et de proposer des mesures préventives. Les mathématiques, ici, deviennent un outil concret pour protéger les populations et adapter nos villes au changement climatique.

23. Réduire l'empreinte carbone

Le problème : serveurs énergivores

GreenTech, une entreprise spécialisée en technologies vertes, est confrontée à un problème majeur : ses serveurs consomment une grande quantité d'électricité, contribuant ainsi fortement à son empreinte carbone.

Question

Comment optimiser cette consommation d'énergie pour réduire les coûts et l'impact environnemental, tout en maintenant les performances des serveurs ?

Les mathématiques, en particulier l'optimisation, peuvent offrir une solution efficace.

Modélisation mathématique de la consommation d'énergie

Pour résoudre ce problème, GreenTech modélise la consommation d'énergie de ses serveurs à l'aide de la fonction suivante :

$$E(x) = 2x^2 - 20x + 100 \quad \text{(en kWh)}$$

où x représente le nombre d'opérations traitées par unité de temps.

Objectif : trouver la valeur de x qui minimise la fonction $E(x)$.

Résolution du problème : trouver le minimum

On dérive la fonction :

$$E'(x) = 4x - 20$$

En résolvant $E'(x) = 0$, on obtient :

$$4x - 20 = 0 \quad \Rightarrow \quad x_0 = 5$$

Analyse du point critique

On vérifie la nature du point $x_0 = 5$:

$$E''(x) = 4$$

Puisque $E''(x_0) > 0$, il s'agit bien d'un minimum local.

GreenTech devrait donc configurer ses serveurs pour traiter 5 opérations par unité de temps afin d'optimiser la consommation d'énergie.

Tableau de variation

x	$-\infty$	5	$+\infty$
$E'(x)$	$-$	0	$+$
$E(x)$	$\searrow$	50	$\nearrow$

Interprétation : le minimum de la fonction est atteint pour $x = 5$, avec une consommation minimale de :

$$E(5) = 2(5)^2 - 20(5) + 100 = 50\,\text{kWh}$$

Critique du modèle utilisé

Bien que simple et efficace, ce modèle comporte des limites :

– La relation entre x et $E(x)$ est supposée quadratique, ce qui peut être trop simplifié.

– Les coefficients peuvent varier selon les conditions réelles (température, charge, maintenance...).

– Le modèle ne prend pas en compte les incertitudes (pannes, pics d'activité...).

Suggestion : une modélisation plus robuste intégrerait des simulations stochastiques et une analyse de sensibilité.

Conclusion

Cet exemple montre comment l'optimisation mathématique peut contribuer à :

- Réduire l'empreinte carbone
- Améliorer l'efficacité énergétique
- Réaliser des économies substantielles

Une application simple, mais concrète, des maths dans la transition écologique.

24. Qui est responsable ? Maths Pour juges ?

Question de départ

Problème

Quand une entreprise cause des dommages, comment peut-on mesurer sa responsabilité ? Peut-on utiliser les mathématiques pour modéliser ce processus et éviter des conflits juridiques complexes ?

Réponse : Oui. Grâce à des modèles mathématiques, il est possible de prédire l'ampleur des dommages et la responsabilité associée à une entreprise à chaque instant.

Modèles de responsabilité

Prenons une entreprise qui cause des dommages environnementaux au fil du temps. Voici une modélisation possible :

- Les dommages $D(t)$ augmentent avec l'activité de l'entreprise $A(t)$
- La responsabilité $R(t)$ dépend des dommages cumulés et des lois en vigueur

$$D(t) = \int_0^t f(A(\tau))\, d\tau \quad \text{et} \quad R(t) = g(D(t), P)$$

où P représente les facteurs juridiques et les éléments de preuve disponibles.

Exemple de situation

Imaginons qu'une entreprise dépasse un seuil critique de pollution. Sa responsabilité, $R(t)$, augmente avec le temps, et si elle dépasse un seuil S, elle devient légalement responsable.

$$\text{Responsabilité} = \begin{cases} 1 & \text{si } R(t) \geq S \\ 0 & \text{sinon} \end{cases}$$

Si $R(t) \geq S$, l'entreprise est considérée comme responsable et devra réparer les dommages causés.

Applications et utilité

Ces modèles permettent de :

- **Anticiper les risques** : Suivi en temps réel de la responsabilité potentielle.
- **Prévenir les litiges** : Prise de décisions avant d'atteindre un seuil critique.
- **Gérer les risques** : Mieux comprendre l'impact de ses activités sur l'environnement et la société.

Challenges dans ce travail

Malgré leur utilité, ces modèles présentent plusieurs difficultés :

- **Définir les fonctions f et g** : Elles doivent être crédibles, représentatives, mais aussi simples à exploiter.
- **Accès aux données** : Il faut pouvoir observer ou estimer précisément l'activité de l'entreprise et ses impacts.
- **Modèle non-linéaire** : Les lois, les dommages et les preuves ne s'ajustent pas toujours de façon simple.
- **Acceptation juridique** : Les juges doivent être convaincus par l'usage de ces outils.

Conclusion

Les mathématiques permettent de mieux comprendre, anticiper et modéliser la responsabilité d'une entreprise. Elles deviennent un outil d'aide à la décision, utile pour prévenir les conflits et favoriser la justice environnementale et sociale.

25. Rentabilité d'une machine

Le Contexte

Question

Comment déterminer la rentabilité d'une machine de production alimentaire en utilisant un modèle déterministe ?

Les maths : On utilise un modèle déterministe pour calculer le temps de rentabilité en fonction :

- des coûts initiaux,
- des coûts opérationnels,
- de la production,
- des prix de vente.

Formulation du modèle déterministe

- C_0 : Coût initial (achat, installation)
- C_{op} : Coût opérationnel par unité de temps
- $P(t)$: Production par unité de temps
- p_{vente} : Prix de vente unitaire
- $M(t)$: Coût de maintenance par unité de temps

NB : Ces données sont supposées connues. En réalité, le coût de maintenance est souvent imprévisible (pannes aléatoires), d'où l'intérêt futur d'un modèle stochastique.

Modélisation mathématique

Revenus cumulés jusqu'au temps t :

$$R(t) = P(t) \times p_{\text{vente}} \times t$$

Coûts cumulés jusqu'au temps t :

$$C(t) = C_0 + C_{\text{op}} \times t + M(t) \times t$$

Condition de rentabilité :

$$R(t_r) \geq C(t_r) \quad \Rightarrow \quad t_r \geq \frac{C_0}{P(t) \cdot p_{\text{vente}} - (C_{\text{op}} + M(t))}$$

Le temps t_r est le moment où la machine devient rentable.

Exemple concret

Supposons les valeurs suivantes :

— $C_0 = 100\,000$

— $C_{\text{op}} = 2\,000$ par mois

— $P(t) = 500$ unités/mois

— $p_{\text{vente}} = 100$

— $M(t) = 500$ par mois

$$t_r \geq \frac{100\,000}{(500 \times 100) - (2\,000 + 500)} = \frac{100\,000}{47\,500} \approx 2{,}1 \text{ mois}$$

Interprétation

— La machine devient rentable après environ 2,1 mois d'utilisation.

— Ce modèle est simple, rapide à appliquer, et utile lorsque les paramètres sont bien connus.

Limites du modèle

Critiques :

— Il suppose des coûts et revenus fixes.

— Il ne tient pas compte des pannes imprévues ou fluctuations de la demande.

— Il ignore la valeur temporelle de l'argent ou les effets des réinvestissements.

Recommandation : Compléter cette approche avec :

— une analyse de sensibilité ;

— des simulations stochastiques pour intégrer les incertitudes réelles.

Conclusion

Le modèle déterministe de rentabilité est un outil pratique pour estimer à quel moment une machine devient un investissement rentable. En partant d'hypothèses simples, il donne une première estimation du retour sur investissement.

26. L'évolution de la température dans une pièce

Problème

Comment les mathématiques peuvent-elles nous aider à comprendre et prédire l'évolution de la température dans une pièce ?

Pour modéliser la manière dont la chaleur se propage dans une pièce, on utilise une équation mathématique bien connue : **l'équation de la chaleur**. Il s'agit d'une équation aux dérivées partielles (EDP) qui décrit la diffusion de la température dans l'espace au cours du temps.

Le principe de l'équation de la chaleur

L'équation de la chaleur s'écrit généralement sous la forme suivante :

$$\frac{\partial T}{\partial t} = \alpha \nabla^2 T + Q(x, t)$$

où :

— $T(x, t)$ représente la température à un point x et au temps t,

— α est la diffusivité thermique du matériau (elle dépend du matériau étudié),

— $\nabla^2 T$ est le *Laplacien* de T, qui mesure la variation spatiale de la température,

— $Q(x, t)$ est une fonction qui modélise les sources de chaleur (comme un radiateur ou un climatiseur).

Modélisation d'une pièce simple

Imaginons une pièce rectangulaire dont les murs sont maintenus à une température constante T_0, et dans laquelle un radiateur fournit une chaleur constante Q_0 à un endroit donné. Initialement, toute la pièce est à température T_0. L'objectif est de savoir comment la température évolue dans le temps à partir de cette configuration.

Solution générale de l'équation

Dans un cadre plus théorique, la solution générale peut s'écrire sous la forme suivante :

$$T(x,t) = T_0 + \int_0^t \int_{\text{Domaine}} G(x,\xi,t-\tau) Q(\xi,\tau)\, d\xi\, d\tau$$

où $G(x,\xi,t)$ est une fonction appelée *fonction de Green*, qui décrit comment une source ponctuelle influence la température autour d'elle.

Cependant, dans la pratique, il est difficile de calculer cette intégrale exactement. On utilise donc des méthodes numériques.

Méthodes numériques utilisées

Parmi les méthodes numériques utilisées pour résoudre l'équation de la chaleur, on trouve :

— La méthode des différences finies,

— La méthode des éléments finis,

— Des outils de simulation comme Python ou MATLAB.

Exemple : méthode des différences finies

La méthode des différences finies consiste à discrétiser l'espace en petits points (notés i) et le temps en instants (notés n). Cela permet d'approximer l'équation continue par une équation discrète :

$$\frac{T_i^{n+1} - T_i^n}{\Delta t} = \alpha \frac{T_{i+1}^n - 2T_i^n + T_{i-1}^n}{(\Delta x)^2} + Q_i^n$$

Cette équation permet de calculer la température au point i à l'instant suivant $n+1$, en utilisant les températures aux points voisins au temps n.

Interprétation et utilité

Grâce à cette modélisation, on peut :

- Comprendre comment la chaleur se diffuse dans une pièce,
- Prédire la température à différents endroits dans le futur,
- Optimiser le placement des radiateurs pour un confort thermique maximal.

Ces outils sont très utiles dans la conception thermique des bâtiments, la gestion énergétique et l'amélioration du confort des habitants.

Conclusion

L'équation de la chaleur est un outil puissant pour comprendre et prédire l'évolution de la température dans un espace fermé. En combinant théorie mathématique et simulation numérique, elle permet de concevoir des environnements intérieurs plus confortables et plus efficaces sur le plan énergétique.

27. Le temps d'attente dans un service public

FIGURE 27.1 – Illustration visuelle de la problématique des files d'attente dans un service public

Problème

Comment les mathématiques peuvent-elles nous aider à modéliser et optimiser le temps d'attente des clients dans un service public, en utilisant la théorie des files d'attente ?

Pour répondre à cette question, on peut utiliser la **théorie des files d'attente**, et plus précisément le modèle **M/M/1**. Ce modèle mathématique repose sur deux hypothèses principales :

— Les clients arrivent selon un *processus de Poisson* avec un taux moyen λ (nombre moyen d'arrivées par minute),

— Le temps de service est distribué de façon *exponentielle* avec un taux moyen μ (nombre moyen de clients servis par minute).

Le modèle M/M/1 : définitions et formules

Dans le modèle M/M/1, on définit :

— λ : taux moyen d'arrivée des clients,

— μ : taux moyen de service,

— $\rho = \frac{\lambda}{\mu}$: taux d'utilisation du serveur (proportion du temps où il est occupé).

Les formules clés sont les suivantes :

$$p_0 = 1 - \rho \quad \text{(probabilité qu'il n'y ait aucun client, serveur inactif)}$$

$$\bar{N} = \frac{\rho}{1-\rho} \quad \text{(nombre moyen de clients dans le système)}$$

$$T = \frac{1}{\mu(1-\rho)} \quad \text{(temps moyen passé par un client dans le système)}$$

Le temps T inclut à la fois le temps d'attente dans la file et le temps de service.

Quels facteurs influencent le temps d'attente ?

Facteurs influençant le temps d'attente

— **Taux d'arrivée élevé (λ)** : Plus les clients arrivent rapidement, plus la file d'attente s'allonge.

— **Taux de service insuffisant (μ)** : Si le service est lent, les clients attendent plus longtemps.

— **Optimisation des ressources** : Ajouter des agents ou améliorer la rapidité du service permet de réduire l'attente.

Exemple concret

Prenons un service public où :

$$\lambda = 5 \text{ clients par minute}, \quad \mu = 8 \text{ clients par minute}$$

Alors :

$$\rho = \frac{5}{8} = 0{,}625$$

$$T = \frac{1}{8(1 - 0{,}625)} = \frac{1}{3} = 3{,}33 \text{ minutes}$$

Chaque client passe donc en moyenne 3,33 minutes dans le système (file + service).

Pourquoi ce modèle est utile ?

Le modèle M/M/1 peut être utilisé pour :

— Optimiser l'organisation du personnel dans un service public,

— Réduire les files d'attente dans les guichets administratifs,

— Améliorer la satisfaction des usagers,

— Aider les gestionnaires à anticiper les situations de surcharge.

Conclusion

Grâce à la théorie des files d'attente, il est possible de modéliser les flux de clients dans un service et de prendre des décisions éclairées pour réduire le temps d'attente. Le modèle M/M/1 est simple mais puissant : il offre une première approximation efficace pour améliorer l'organisation des services publics.

28. Traduction directe de la parole

Figure 28.1 – Illustration de la reconnaissance vocale et de la traduction automatique

Problème

Comment les mathématiques peuvent-elles nous aider à modéliser la reconnaissance de la parole et la traduction automatique ?

Pour comprendre comment les machines reconnaissent la parole et traduisent automatiquement, les mathématiques utilisent un outil fondamental : les **chaînes de Markov cachées** (ou HMM, pour Hidden Markov Models). Ces modèles probabilistes permettent de représenter une séquence de sons et de prédire les mots les plus probables dans une phrase.

Modélisation de la reconnaissance vocale avec les HMM

Les HMM modélisent une suite de sons (ou de phonèmes) comme une série d'états cachés, chacun associé à une probabilité d'observer un certain son. La structure du modèle repose sur deux types de probabilités :

- La probabilité de transition entre les états (passage d'un son à un autre),
- La probabilité d'observer un certain son à partir d'un état donné.

L'expression mathématique de la probabilité d'une séquence est :

$$P(\text{séquence}) = P(s_1) \cdot P(o_1|s_1) \cdot P(s_2|s_1) \cdot P(o_2|s_2) \cdot \ldots$$

avec :

- $P(s_1)$: probabilité initiale du premier état (premier son),
- $P(s_j|s_i)$: probabilité de passer de s_i à s_j,
- $P(o_i|s_i)$: probabilité d'entendre le son o_i si l'état réel est s_i.

Lien avec la traduction automatique

Les HMM ne servent pas qu'à la reconnaissance de la parole. Ils jouent aussi un rôle clé dans la traduction automatique. En effet, une phrase peut être vue comme une séquence de mots dont les structures peuvent être analysées par des probabilités de transition.

Applications des HMM

- **Reconnaissance vocale** : Identifier les mots prononcés à partir des sons captés par un micro.
- **Traduction automatique** : Trouver la phrase équivalente dans une autre langue à partir de séquences de mots et de probabilités de correspondance.

Exemple de calcul

Imaginons que quelqu'un prononce deux sons o_1 et o_2, associés aux états cachés s_1 et s_2. La probabilité que cette séquence corresponde à une phrase particulière est :

$$P(\text{phrase}) = P(s_1) \cdot P(o_1|s_1) \cdot P(s_2|s_1) \cdot P(o_2|s_2)$$

C'est cette probabilité que le système de reconnaissance vocale maximise pour déterminer les mots prononcés.

Pourquoi ces modèles sont utiles ?

Les chaînes de Markov cachées permettent de :

- **Reconnaître la parole** avec une grande précision, même dans un environnement bruyant,
- **Traduire automatiquement** des phrases en exploitant la structure des langues,
- **Modéliser les systèmes de sons** dans une langue, en tenant compte des transitions probables entre phonèmes.

Conclusion

Les chaînes de Markov cachées (HMM) sont des outils mathématiques puissants qui permettent de modéliser les sons de la parole humaine et d'en tirer du sens. Grâce à elles, les systèmes de reconnaissance vocale et de traduction automatique deviennent de plus en plus performants, rapprochant un peu plus la machine de la compréhension du langage humain.

29. Pourquoi calculer des dérivées ?

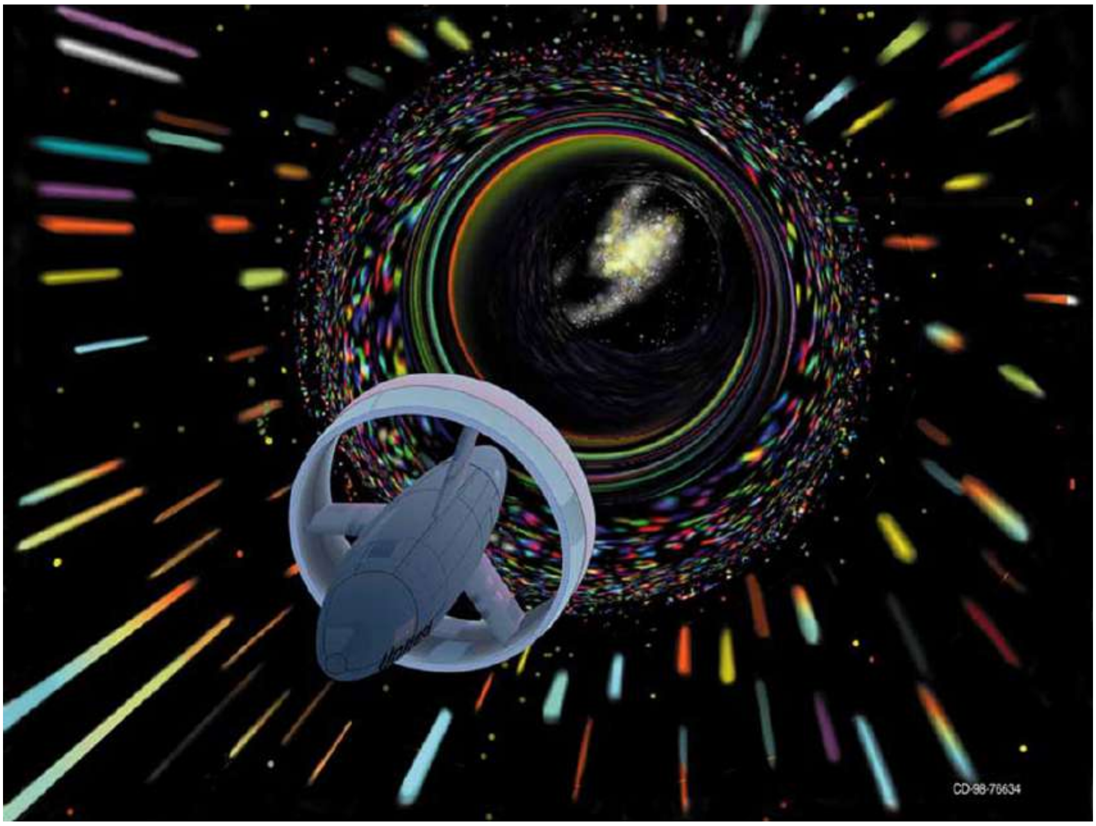


FIGURE 29.1 – Illustration de la vitesse et du changement de position dans le temps

Définition de la vitesse

La vitesse mesure le changement de position d'un objet au cours du temps. C'est une dérivée de la position par rapport au temps.

Exemple : Une voiture parcourt 100 km en 2 heures. Sa vitesse moyenne est donnée par :

$$\text{Vitesse} = \frac{\text{Distance}}{\text{Temps}} = \frac{100\ \text{km}}{2\ \text{h}} = 50\ \text{km/h}$$

Mais qu'en est-il de la vitesse à chaque instant ? C'est là qu'intervient la dérivée.

Dérivée et vitesse instantanée

La dérivée

La dérivée mesure comment une quantité change par rapport à une autre. En physique, la dérivée de la position donne la **vitesse instantanée**, c'est-à-dire la vitesse à un instant précis.

Exemple : Si la position d'une voiture est donnée par $s(t) = 5t^2$, alors :

$$s'(t) = \frac{ds}{dt} = 10t$$

Cela signifie qu'à chaque instant t, la vitesse de la voiture est $10t$ km/h.

Utilisation des dérivées en optimisation

Les dérivées sont très utiles pour optimiser des fonctions. Elles permettent de :

— Trouver les **minima et maxima** d'une fonction (par exemple, un coût minimal ou un profit maximal),

— Résoudre des problèmes d'optimisation dans l'industrie, l'économie, l'énergie, etc.

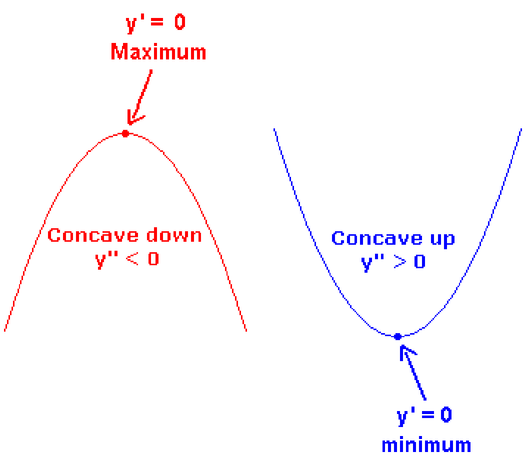


FIGURE 29.2 – Utilisation des dérivées pour trouver un minimum

Rôle des dérivées dans l'intelligence artificielle

Pourquoi les dérivées en IA ?

Les dérivées permettent de corriger les erreurs dans un réseau de neurones, en ajustant les paramètres pour que les prédictions soient meilleures. Ce processus s'appelle la **rétropropagation**.

— La **fonction de perte** mesure l'écart entre la prédiction et la réalité.

— La **descente de gradient** utilise les dérivées de cette fonction pour améliorer le modèle.

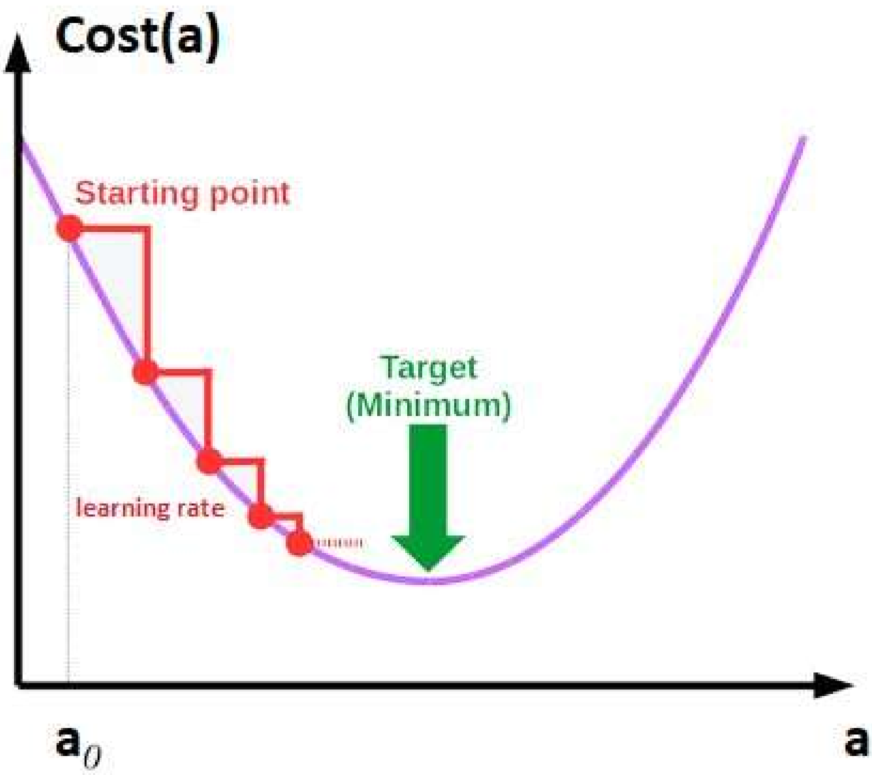


FIGURE 29.3 – La descente de gradient ajuste les paramètres grâce aux dérivées

Dérivées en épidémiologie

Pourquoi les dérivées en santé publique ?

Les dérivées permettent de modéliser la vitesse à laquelle une maladie se propage dans une population.

Exemple : Le modèle SIR est basé sur les dérivées des populations de personnes susceptibles, infectées et rétablies :

$$\frac{dS}{dt} = -\beta SI \quad , \quad \frac{dI}{dt} = \beta SI - \gamma I \quad , \quad \frac{dR}{dt} = \gamma I$$

avec :

— β : taux de transmission,

— γ : taux de guérison.

Ces équations différentielles permettent de prédire le pic de l'épidémie et d'anticiper les besoins sanitaires.

Conclusion

Les dérivées sont partout : dans les sciences, l'économie, la médecine, l'intelligence artificielle... Elles nous permettent de comprendre et de prévoir les changements, d'optimiser nos décisions et de mieux maîtriser les phénomènes qui nous entourent.

30. Modèle ARIMA : mathématiques et applications

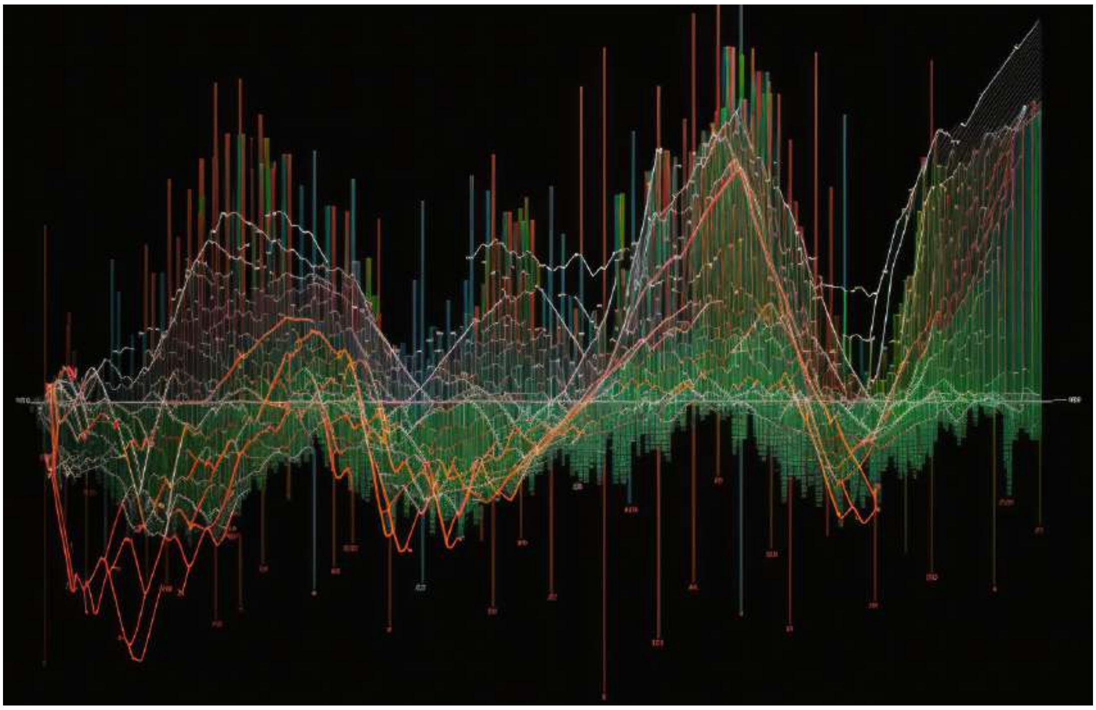


FIGURE 30.1 – Illustration de l'analyse de séries temporelles avec le modèle ARIMA

Problème

Comment les mathématiques permettent-elles de modéliser et prévoir l'évolution d'une série temporelle comme la température, la consommation ou la demande ?

Le modèle ARIMA (AutoRegressive Integrated Moving Average) est un outil mathématique très utilisé pour l'analyse et la prévision de séries temporelles. Il combine trois composantes :

— **AR (auto-régressif)** : la valeur actuelle dépend des valeurs passées,

— **I (intégré)** : on applique des différences pour rendre la série stationnaire,

— **MA (moyenne mobile)** : on ajuste les erreurs passées pour améliorer la prédiction.

Formule générale du modèle ARIMA

Le modèle ARIMA s'écrit :

$$\Delta^d Y_t = \phi_1 \Delta^d Y_{t-1} + \cdots + \phi_p \Delta^d Y_{t-p} + \epsilon_t + \theta_1 \epsilon_{t-1} + \cdots + \theta_q \epsilon_{t-q}$$

où :

- $\Delta^d Y_t$ est la différence d'ordre d de la série Y_t,
- ϕ_i sont les coefficients auto-régressifs (AR),
- θ_j sont les coefficients de moyenne mobile (MA),
- ϵ_t est le bruit aléatoire (erreur).

Étapes de l'algorithme ARIMA

1. **Identification** : choix des paramètres p, d, q à l'aide des graphes ACF et PACF.
2. **Différenciation** : rendre la série stationnaire par différenciation :
$$\Delta^d Y_t = Y_t - Y_{t-d}$$
3. **Estimation** : estimation des paramètres ϕ_i et θ_j en minimisant :
$$\sum_t \epsilon_t^2$$
4. **Validation** : vérification que les résidus sont indépendants et suivent une loi normale.
5. **Prévision** : calcul des valeurs futures à l'aide du modèle ajusté.

Exemple : prédiction de la température mensuelle

On dispose d'une série de températures mensuelles sur 5 ans. En utilisant les courbes ACF et PACF :

- On observe une tendance → on applique une différence première ($d = 1$),

— On choisit $p = 1$ et $q = 1$: modèle ARIMA(1,1,1).

L'équation estimée devient :

$$\Delta Y_t = 0.7\Delta Y_{t-1} + \epsilon_t + 0.5\epsilon_{t-1}$$

Validation du modèle

Pour s'assurer que le modèle est fiable, on effectue :

— un **test de Ljung-Box** pour vérifier l'absence d'autocorrélation des résidus,

— un **test de normalité** pour s'assurer que les résidus sont gaussiens.

Si ces tests sont satisfaisants, le modèle peut être utilisé pour la prévision.

Prévision à court terme

À partir du modèle estimé, on peut prédire :

$$\hat{Y}_{t+1} = \phi_1 \Delta Y_t + \theta_1 \epsilon_t$$

Exemple : Si la dernière température est $15.2°C$, la variation précédente est -0.2, et l'erreur précédente est 0.2, alors :

$$\hat{Y}_{t+1} = 15.2 + 0.7 \times (-0.2) + 0.5 \times 0.2 = 15.5°C$$

Calcul de l'intervalle de confiance

L'incertitude autour de la prévision est donnée par :

$$\text{IC} = \hat{Y}_{t+k} \pm z_{\alpha/2}\sqrt{\text{Var}(\epsilon_t)}$$

Exemple : avec $\alpha = 0.05$ et $\sqrt{\text{Var}(\epsilon_t)} = 0.3$, on a :

$$\text{IC} = 15.5 \pm 1.96 \times 0.3 = [14.92,\ 16.08]°C$$

Conclusion

Le modèle ARIMA permet de modéliser des séries temporelles non stationnaires, en combinant trois mécanismes puissants. Il est largement utilisé en économie, météorologie, finance, etc. En maîtrisant ses étapes, on peut effectuer des prévisions fiables et utiles dans de nombreux domaines.

31. Les maths derrière les K-means

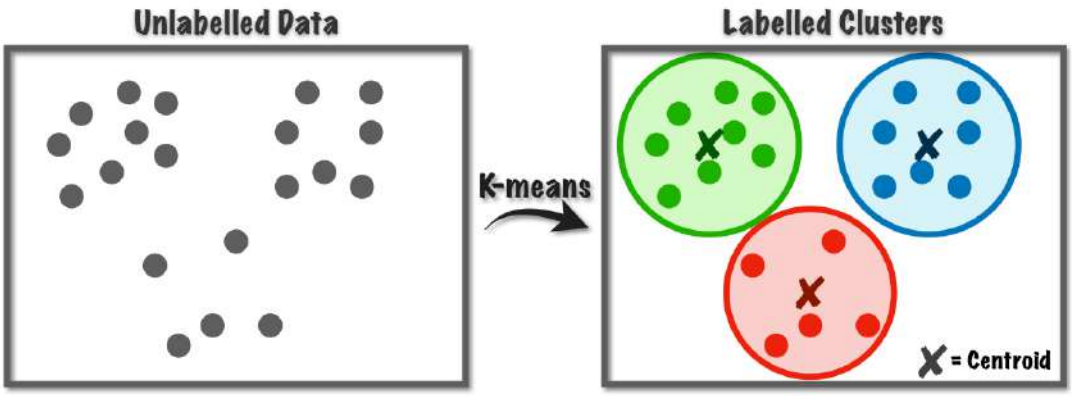


FIGURE 31.1 – Illustration de regroupement de données par K-Means

Question

Comment l'algorithme K-Means utilise-t-il la géométrie euclidienne pour regrouper les points de données en clusters homogènes ?

Réponse : K-Means cherche à regrouper des données proches les unes des autres en minimisant la somme des distances au carré entre chaque point et le centroïde de son cluster. La distance utilisée est la *distance euclidienne*, qui mesure la proximité géométrique entre deux points.

La fonction mathématique derrière K-Means

L'objectif de l'algorithme est de minimiser une fonction de coût définie comme suit :

$$J = \sum_{k=1}^{K} \sum_{\mathbf{x}_i \in C_k} \|\mathbf{x}_i - \mu_k\|^2$$

où :

— C_k est le k-ième cluster,

— μ_k est le centroïde du cluster C_k,

— $\|\mathbf{x}_i - \mu_k\|$ est la distance euclidienne entre un point $\mathbf{x}_i$ et son centroïde.

Distance euclidienne : définition

Distance euclidienne

$$\|\mathbf{x}_1 - \mathbf{x}_2\| = \sqrt{(x_{11} - x_{21})^2 + (x_{12} - x_{22})^2 + \ldots + (x_{1n} - x_{2n})^2}$$

C'est la distance la plus courte entre deux points dans un espace à n dimensions.

Elle permet de mesurer la proximité entre les points. C'est ce critère qui garantit des clusters compacts et bien séparés.

Optimisation géométrique

Pourquoi minimiser la distance ?

En minimisant la distance euclidienne intra-cluster, on crée des groupes de points homogènes. Chaque point est ainsi le plus proche possible de son centre de cluster.

L'objectif final est de minimiser la variance à l'intérieur des groupes.

Les étapes de l'algorithme K-Means

1. **Initialisation** : Choisir aléatoirement K centroïdes.
2. **Assignation** : Affecter chaque point au centroïde le plus proche.
3. **Mise à jour** : Recalculer les centroïdes comme moyenne des points dans chaque cluster.
4. **Répétition** : Répéter assignation et mise à jour jusqu'à convergence.

Exemple de clustering : données initiales

Considérons 6 points à regrouper en 2 clusters :

Point	$\mathbf{x}_1$	$\mathbf{x}_2$	$\mathbf{x}_3$	$\mathbf{x}_4$	$\mathbf{x}_5$	$\mathbf{x}_6$
x	2	5	9	4	8	7
y	3	4	6	7	1	5

Initialisation : On choisit deux centroïdes aléatoires :

$$\mu_1 = (2, 3), \quad \mu_2 = (8, 1)$$

Étapes du clustering

- Calcul des distances euclidiennes entre chaque point et les centroïdes.
- Affectation des points aux clusters selon la plus petite distance.
- Mise à jour des centroïdes en calculant la moyenne des points de chaque cluster.
- Répétition jusqu'à convergence.

Résultat final : Les 6 points sont regroupés en deux groupes homogènes, chacun centré autour d'un centroïde calculé à partir des données.

Applications de K-Means

- **Segmentation de marché** : regrouper des clients selon leurs comportements d'achat.
- **Classification d'images** : regrouper les pixels ou objets similaires.
- **Analyse de données génétiques** : identifier des structures dans des séquences biologiques.

Conclusion

L'algorithme K-Means est un outil simple et puissant qui utilise la géométrie euclidienne pour structurer les données. Il permet de regrouper des éléments similaires et de découvrir des structures cachées dans des jeux de données complexes.

32. Les maths derrière les réseaux neuronaux

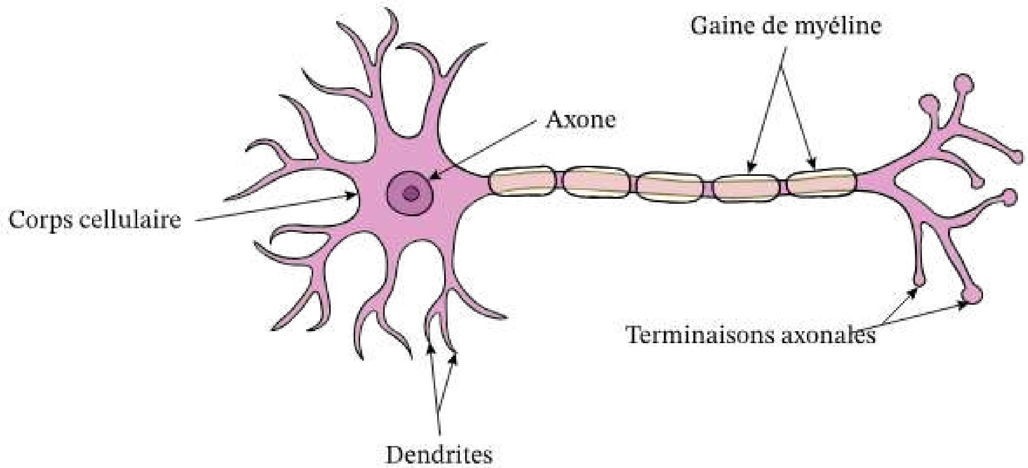


FIGURE 32.1 – Illustration de la structure générale d'un réseau neuronal

Pourquoi les réseaux neuronaux ?

Les réseaux neuronaux sont utilisés pour résoudre des problèmes complexes comme la reconnaissance d'images, la traduction de langues et la prédiction de données. Mais comment fonctionnent-ils ?

Réponse : Ils s'appuient sur des concepts mathématiques comme les fonctions d'activation, les fonctions de perte et la descente de gradient pour apprendre et généraliser à partir de données !

Structure d'un réseau neuronal

Un réseau neuronal est constitué de plusieurs couches :

— **Couche d'entrée** : reçoit les données initiales (ex. : pixels d'une image),

— **Couches cachées** : effectuent des calculs intermédiaires pour extraire des motifs,

— **Couche de sortie** : fournit le résultat final (ex. : prédiction de classe).

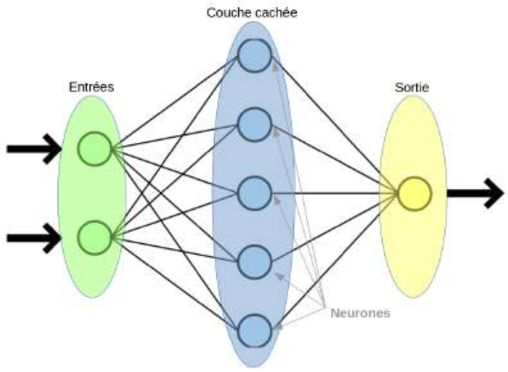


FIGURE 32.2 – Exemple de structure de réseau neuronal

Mathématiques derrière un neurone

Chaque neurone applique une fonction mathématique :

$$a = f\left(\sum_i w_i x_i + b\right)$$

où :

— x_i : entrées du neurone,

— w_i : poids associés à chaque entrée,

— b : biais,

— f : fonction d'activation (ReLU, sigmoïde...),

— a : sortie du neurone (appelée activation).

Propagation avant et rétropropagation

— **Propagation avant** : les données traversent le réseau jusqu'à la prédiction.

— **Rétropropagation** : l'erreur est propagée en sens inverse pour corriger les poids.

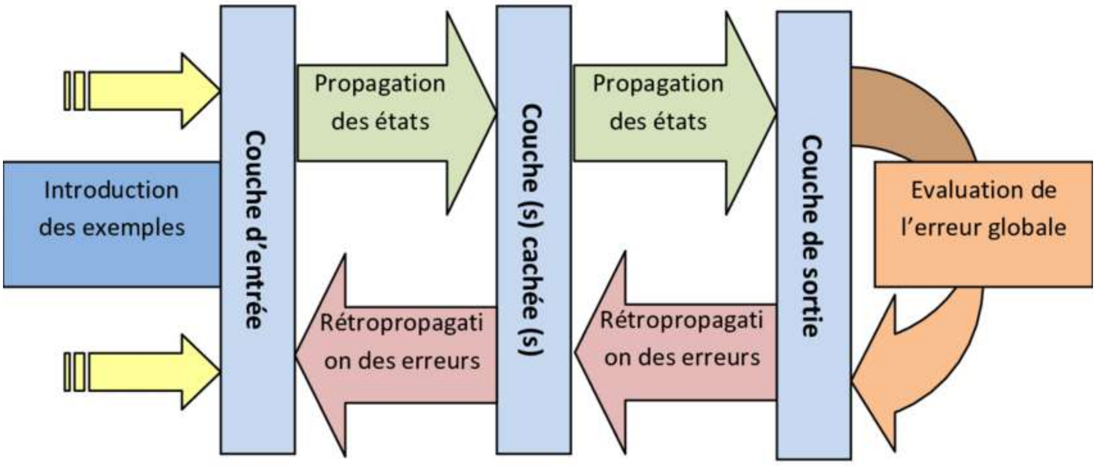


FIGURE 32.3 – La rétropropagation ajuste les poids pour réduire l'erreur

Optimisation : fonction de perte et descente de gradient

Pour apprendre, le réseau cherche à minimiser une **fonction de perte**. Un exemple courant est l'erreur quadratique moyenne (MSE) :

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_{\text{réel}} - y_{\text{prédit}})^2$$

Les poids sont mis à jour grâce à la descente de gradient :

$$w = w - \eta\frac{\partial \text{Perte}}{\partial w}, \quad b = b - \eta\frac{\partial \text{Perte}}{\partial b}$$

où η est le taux d'apprentissage.

Applications concrètes

Les réseaux neuronaux sont utilisés dans de nombreux domaines :

— **Vision par ordinateur** : reconnaissance faciale, détection d'objets,

— **Traitement du langage naturel** : traduction automatique, synthèse vocale,

— **Finance** : prévisions de marché, détection de fraudes,

— **Santé** : analyse d'imagerie médicale, diagnostics assistés.

Conclusion

Les réseaux neuronaux reposent sur des concepts mathématiques puissants pour apprendre à partir de données. Grâce à la propagation avant, la rétropropagation, les dérivées et l'optimisation, ces outils transforment aujourd'hui des secteurs entiers de notre société.

33. Pourquoi changer la norme dans Python ?

Problème

Comment les différentes normes L^1, L^2, et L^∞ sont-elles utilisées pour améliorer la performance des modèles de régression linéaire ?

Réponse : Les normes permettent de régulariser les modèles en ajoutant une pénalité sur les coefficients. Cela permet d'éviter le surajustement (overfitting) et de rendre le modèle plus simple et plus robuste.

Rappel : la régression linéaire classique

Le but est de trouver les coefficients $\beta_0, \beta_1, \ldots, \beta_n$ qui minimisent l'erreur entre les valeurs prédites $\hat{y}_i$ et les valeurs réelles y_i :

$$\hat{y}_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_n x_{in}$$

$$\mathrm{MSE} = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2$$

Mais cette approche peut être sensible au bruit ou au surajustement.

Régularisation par la norme L^1 : Régression Lasso

$$\min_{\beta}\left(\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 + \lambda\sum_{j=1}^{n}|\beta_j|\right)$$

Avantages :

— Réduit certains coefficients à zéro,

— Sélectionne les variables automatiquement.

Régularisation par la norme L^2 : Régression Ridge

$$\min_\beta \left(\frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y}_i)^2 + \lambda \sum_{j=1}^{n} \beta_j^2 \right)$$

Avantages :

- Réduit tous les coefficients mais ne les annule pas,
- Utile quand les variables explicatives sont corrélées entre elles.

Norme L^∞ : Moindres carrés contraints

$$\min_\beta \max_{i=1,\dots,N} |y_i - \hat{y}_i|$$

Avantages :

- Rend le modèle robuste aux valeurs aberrantes (outliers),
- Moins utilisée que Lasso et Ridge, mais intéressante en optimisation robuste.

Optimisation mathématique du problème

Le problème de régression revient à résoudre une équation matricielle :

$$\boldsymbol{\beta} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y}$$

où :

- $\mathbf{X}$ est la matrice des variables explicatives,
- $\mathbf{y}$ est le vecteur des sorties observées,
- $\boldsymbol{\beta}$ est le vecteur des coefficients à estimer.

Exemple visuel : régularisation en Python

Dans les bibliothèques Python comme ‘scikit-learn‘, la commande ‘.fit()‘ effectue cette résolution en utilisant les normes choisies.

```python
import numpy as np
from sklearn.linear_model import LinearRegression, Lasso, Ridge

# Générer des données fictives
X = np.array([[1], [2], [3], [4], [5]])
y = np.array([1.2, 1.9, 3.0, 4.1, 5.1])

# Régression Linéaire Simple
reg = LinearRegression().fit(X, y)
print("Coefficients Régression Simple:", reg.coef_)

# Régression Ridge (norme L^2)
ridge = Ridge(alpha=1.0).fit(X, y)
print("Coefficients Ridge (L^2):", ridge.coef_)

# Régression Lasso (norme L^1)
lasso = Lasso(alpha=0.1).fit(X, y)
print("Coefficients Lasso (L^1):", lasso.coef_)
```

```
Coefficients Régression Simple: [1.]
Coefficients Ridge (L^2): [0.90909091]
Coefficients Lasso (L^1): [0.95]
```

FIGURE 33.1 – Effet de la régularisation sur une régression en Python

Qu'est-ce qu'une norme ?

Définition : Une norme est une fonction $\|\cdot\| : \mathbb{R}^n \to \mathbb{R}$ qui respecte les propriétés suivantes :

1. **Positivité** : $\|\mathbf{x}\| \geq 0$, et $\|\mathbf{x}\| = 0$ si $\mathbf{x} = 0$,
2. **Homogénéité** : $\|\alpha\mathbf{x}\| = |\alpha|\|\mathbf{x}\|$,
3. **Inégalité triangulaire** : $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$.

Exemples courants :

— Norme L^1 : $\|\mathbf{x}\|_1 = \Sigma_{i=1}^n |x_i|$

— Norme L^2 : $\|\mathbf{x}\|_2 = \sqrt{\Sigma_{i=1}^n x_i^2}$

— Norme L^∞ : $\|\mathbf{x}\|_\infty = \max_i |x_i|$

Équivalence des normes : Toutes les normes sont équivalentes dans $\mathbb{R}^n$, ce qui signifie qu'elles définissent la même topologie. On a par exemple :

$$C_1\|\mathbf{x}\|_1 \leq \|\mathbf{x}\|_2 \leq C_2\|\mathbf{x}\|_1 \quad \text{et} \quad \|\mathbf{x}\|_\infty \leq \|\mathbf{x}\|_2 \leq \sqrt{n}\|\mathbf{x}\|_\infty$$

Conclusion

Les normes L^1, L^2 et L^∞ sont des outils mathématiques puissants pour régulariser les modèles, éviter le surajustement, et améliorer la généralisation. Leur choix influence directement la forme du modèle obtenu et sa capacité à bien prédire sur de nouvelles données.

34. Optimisation des modèles avec la méthode de Lagrange

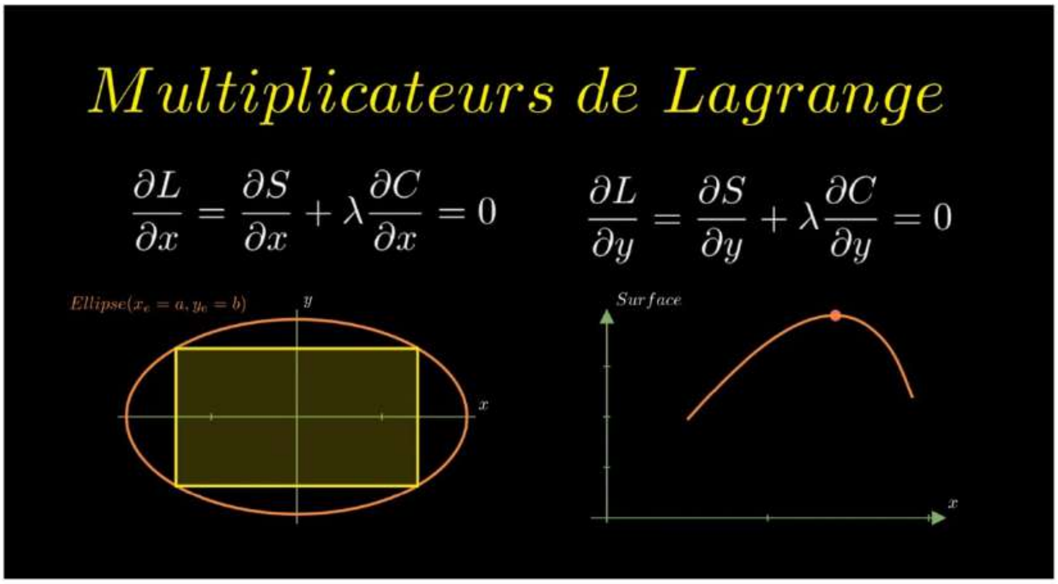


FIGURE 34.1 – Optimisation de modèles sous contraintes avec les multiplicateurs de Lagrange

Question

Comment les multiplicateurs de Lagrange sont-ils utilisés pour optimiser des modèles avec des contraintes dans le cadre d'algorithmes d'apprentissage comme le Support Vector Machine (SVM) ?

Réponse : Les multiplicateurs de Lagrange permettent d'intégrer des contraintes dans une optimisation. Cela facilite la recherche des paramètres optimaux d'un modèle tout en respectant certaines conditions structurelles.

Fondements mathématiques

Soit une fonction objectif $f(\mathbf{x})$ à optimiser, sous contrainte d'égalité $g_i(\mathbf{x}) = 0$. On introduit les multiplicateurs de Lagrange λ_i et on définit la fonction de Lagrange :

$$\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}) = f(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i g_i(\mathbf{x})$$

Pour trouver un optimum, on résout les équations de premier ordre :

$$\frac{\partial \mathcal{L}}{\partial \mathbf{x}} = 0 \quad \text{et} \quad \frac{\partial \mathcal{L}}{\partial \lambda_i} = 0 \quad \text{pour chaque } i$$

Application aux SVM

Les Support Vector Machines cherchent à séparer des données en maximisant la marge entre les classes tout en respectant les étiquettes des données d'entraînement.

La fonction de Lagrange associée est :

$$\mathcal{L}(\mathbf{w}, b, \boldsymbol{\alpha}) = \frac{1}{2}\|\mathbf{w}\|^2 - \sum_{i=1}^{n} \alpha_i \left[y_i(\mathbf{w} \cdot \mathbf{x}_i + b) - 1\right]$$

Conditions de stationnarité :

$$\frac{\partial \mathcal{L}}{\partial \mathbf{w}} = \mathbf{w} - \sum_{i=1}^{n} \alpha_i y_i \mathbf{x}_i = 0$$
$$\frac{\partial \mathcal{L}}{\partial b} = \sum_{i=1}^{n} \alpha_i y_i = 0$$
$$\frac{\partial \mathcal{L}}{\partial \alpha_i} = y_i(\mathbf{w} \cdot \mathbf{x}_i + b) - 1 = 0 \quad \forall i$$

En résolvant ce système, on obtient les valeurs optimales des paramètres du modèle : $\mathbf{w}$, b et α_i.

Pourquoi utiliser les multiplicateurs de Lagrange ?

- Ils permettent de reformuler un problème avec contraintes comme un problème sans contraintes.
- Dans le cas des SVM, cela permet d'identifier les vecteurs supports, qui définissent la frontière de décision.
- L'optimisation consiste à maximiser la marge entre les classes tout en respectant la structure imposée par les étiquettes.

Conclusion

La méthode de Lagrange est un outil fondamental en optimisation sous contraintes. Elle est utilisée dans de nombreux domaines, notamment en machine learning pour entraîner des modèles comme les SVM, qui doivent satisfaire à des contraintes de classification tout en minimisant une fonction de coût.

Elle permet d'obtenir des modèles à la fois rigoureux, interprétables et efficaces.

35. Les tenseurs, c'est juste des matrices ? Pas si vite !

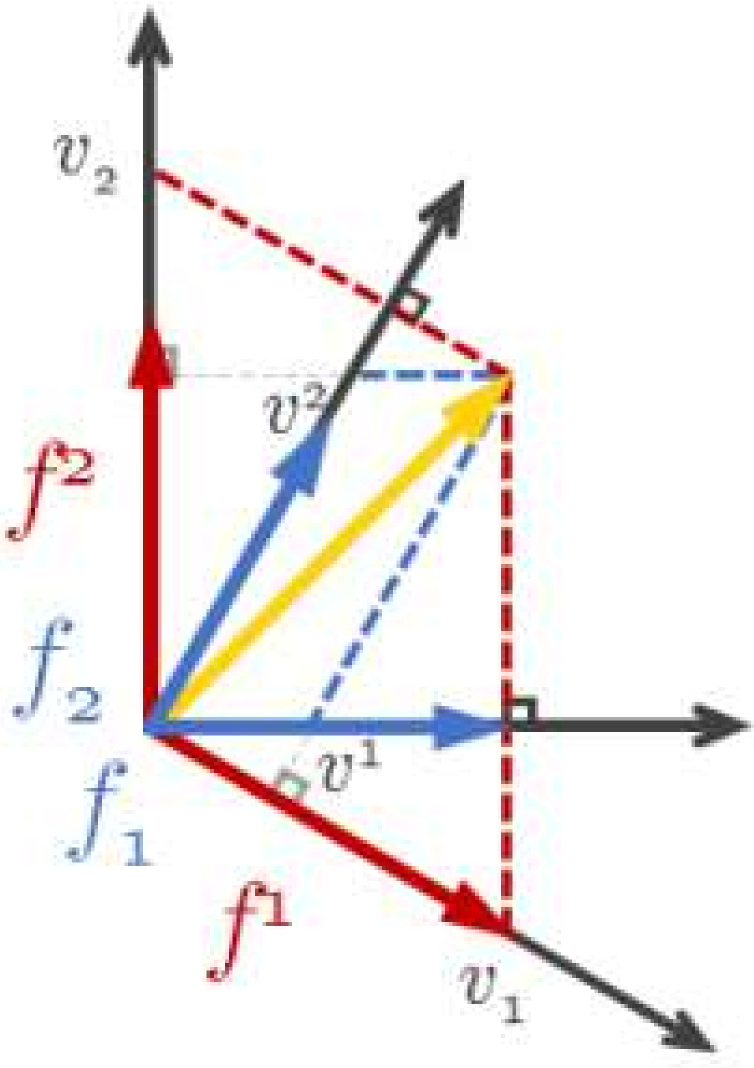


FIGURE 35.1 – Les tenseurs sont au cœur de l'architecture de TensorFlow

Question

Comment TensorFlow utilise-t-il les tenseurs pour effectuer des calculs complexes en apprentissage automatique ?

Réponse : Dans TensorFlow, les tenseurs transportent des données multidimensionnelles (images, séries temporelles, etc.) à travers un graphe de calcul. Cela permet l'exécution d'opérations comme la multiplication de matrices, essentielle pour entraîner des réseaux de neurones.

Définition des tenseurs

Un tenseur est une structure mathématique qui généralise les scalaires, vecteurs et matrices :

- **Scalaire** = tenseur d'ordre 0 : $T = 5$
- **Vecteur** = tenseur d'ordre 1 : $T = (t_1, t_2, t_3)$
- **Matrice** = tenseur d'ordre 2 :
$$T = \begin{pmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{pmatrix}$$
- **Tenseur d'ordre 3 ou plus** : Cubes ou blocs de données, utilisés pour modéliser des images, vidéos, etc.

Multiplication de tenseurs

L'une des opérations clés est la multiplication matricielle :

$$C = A \cdot B \quad \text{où} \quad C_{ij} = \sum_{k=1}^{n} A_{ik} B_{kj}$$

Cette opération est omniprésente dans les couches de réseaux de neurones pour propager l'information d'une couche à l'autre.

Le graphe de calcul dans TensorFlow

TensorFlow repose sur une architecture en graphe :

- Chaque **nœud** est une opération mathématique (addition, multiplication, activation...).
- Chaque **arc** transporte un tenseur entre deux nœuds.

Exemple d'un graphe de calcul :

$$z = x^2 + y^2$$

Ce calcul est représenté par deux nœuds (x^2, y^2), suivis d'un nœud pour la somme.

Rétropropagation et optimisation

Les tenseurs sont aussi utilisés pour optimiser les modèles :

$$\theta = \theta - \alpha \nabla_\theta J(\theta)$$

où :

— θ : paramètres du modèle,

— α : taux d'apprentissage,

— $\nabla_\theta J(\theta)$: gradient de la fonction de coût.

Cette mise à jour permet au modèle de s'améliorer au fil des itérations.

Tenseurs en Python avec TensorFlow

Exemple de création et manipulation de tenseurs :

```
import tensorflow as tf

# Création d'un tenseur 2x2
T = tf.constant([[1, 2], [3, 4]])
print(T)

# Multiplication matricielle
A = tf.constant([[1, 2], [3, 4]])
B = tf.constant([[5, 6], [7, 8]])
C = tf.matmul(A, B)
print(C)
```

Pourquoi les tenseurs sont-ils si importants ?

— **Images** : représentées comme tenseurs 3D (hauteur × largeur × canaux),

— **Séries temporelles** : données séquentielles sur plusieurs pas de temps,

— **Réseaux de neurones** : chaque couche reçoit et retourne un tenseur.

Avantages

— Manipulation efficace de grandes quantités de données,

— Calcul optimisé sur GPU/TPU,

— Flexibilité dans la modélisation,

— Base de l'optimisation par rétropropagation.

Conclusion

Les tenseurs ne sont pas "juste des matrices" : ils sont l'ossature des systèmes d'intelligence artificielle modernes. Dans TensorFlow, ils permettent de représenter, transformer, apprendre et optimiser des modèles complexes sur des données réelles.

36. Regrouper automatiquement des données avec DBSCAN

Problème

Comment regrouper des données sans connaître à l'avance le nombre de groupes (clusters), tout en identifiant automatiquement les points aberrants (outliers) ?

Réponse : L'algorithme DBSCAN (Density-Based Spatial Clustering of Applications with Noise) repose sur des notions mathématiques de densité et de voisinage. Il permet de regrouper les données de manière intelligente sans avoir besoin de spécifier le nombre de groupes.

Les concepts mathématiques derrière DBSCAN

- **Point cœur :** Un point ayant au moins un certain nombre de voisins dans un rayon donné (`minPts` voisins dans un rayon ε).
- **Point bordure :** Un point qui est voisin d'un point cœur, mais qui ne remplit pas lui-même la condition pour être un point cœur.
- **Bruit (outlier) :** Un point qui ne fait partie d'aucun groupe, car il est isolé.

Les paramètres

L'algorithme DBSCAN utilise deux paramètres principaux :

- ε : le rayon de voisinage autour d'un point.
- `minPts` : le nombre minimal de points pour former un groupe (ou cluster).

Comment fonctionne DBSCAN ?

1. On choisit un point non encore visité.
2. On identifie tous les points dans un rayon ε.
3. S'il y a au moins `minPts` voisins, on crée un cluster.
4. On agrandit ce cluster en ajoutant tous les voisins des voisins, tant qu'ils remplissent la condition.
5. Les points qui ne sont dans aucun cluster sont considérés comme du bruit.

Exemple visuel

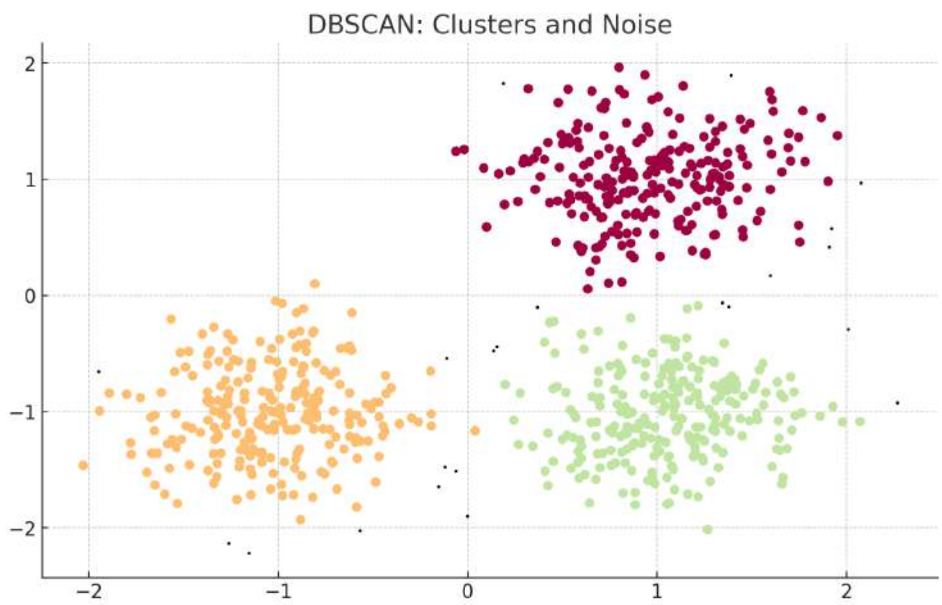


FIGURE 36.1 – Les points colorés représentent des clusters. Les points noirs sont considérés comme du bruit (outliers).

Applications concrètes

— **Détection d'anomalies :** Identifier automatiquement des comportements anormaux dans des données (fraude, panne, etc.).

— **Analyse géographique :** Regrouper des lieux proches (par exemple des magasins, des capteurs, des habitations).

— **Réseaux sociaux :** Identifier des communautés ou des sous-groupes dans des graphes de relations.

Pourquoi c'est utile ?

Contrairement à des algorithmes comme K-Means, DBSCAN ne nécessite pas de connaître le nombre de groupes à l'avance. Il est aussi capable d'identifier les formes irrégulières de groupes et de gérer les données bruitées. C'est donc un outil très utile dans des domaines comme l'intelligence artificielle, la cybersécurité ou encore la géographie.

Conclusion

DBSCAN repose sur des idées simples mais puissantes : la densité locale, la notion de voisinage, et l'extension d'un groupe de manière dynamique. Grâce à ces concepts mathématiques, on peut automatiser le regroupement de données complexes sans avoir à tout paramétrer soi-même.

37. Comment les maths aident à identifier les visages ?

Question centrale

Comment un système de reconnaissance faciale parvient-il à identifier un visage parmi des millions ? La réponse se trouve dans l'utilisation de modèles mathématiques capables de transformer une image en données comparables.

Les systèmes de reconnaissance faciale s'appuient sur un ensemble de caractéristiques physiques mesurables d'un visage : la distance entre les yeux, la largeur du nez, la forme du menton, etc. Ces éléments sont traduits en valeurs numériques et stockés sous forme de vecteurs.

Normalisation des caractéristiques

Chaque caractéristique est souvent normalisée entre 0 et 1 afin de rendre les comparaisons cohérentes. On utilise la formule suivante :

$$x_i' = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}}$$

Cela permet de comparer différents visages sur une même échelle, quelles que soient les unités initiales.

Distance euclidienne

Pour mesurer la proximité entre deux visages, on peut calculer la distance euclidienne entre leurs vecteurs :

$$d = \sqrt{(v_1 - w_1)^2 + (v_2 - w_2)^2 + \cdots + (v_n - w_n)^2}$$

Plus d est faible, plus les visages sont similaires.

Distance euclidienne pondérée

Certaines caractéristiques comptent plus que d'autres dans l'identification. On utilise alors une distance euclidienne pondérée :

$$d_{\text{pondéré}} = \sqrt{a_1(v_1 - w_1)^2 + \cdots + a_n(v_n - w_n)^2}$$

Similarité cosinus

Une autre approche repose sur la comparaison de la direction des vecteurs, plutôt que leur distance. On utilise alors la similarité cosinus :

$$\cos(\mathbf{v}, \mathbf{w}) = \frac{\mathbf{v} \cdot \mathbf{w}}{\|\mathbf{v}\|\|\mathbf{w}\|}$$

Plus le résultat est proche de 1, plus les vecteurs pointent dans la même direction, donc plus les visages sont proches.

Exemple de calcul

Prenons deux vecteurs caractéristiques normalisés :

$$\mathbf{v} = (0.5, 0.7, 0.4, 0.6, 0.8), \quad \mathbf{w} = (0.6, 0.7, 0.5, 0.55, 0.85)$$

La distance euclidienne est :

$$d \approx \sqrt{(0.5 - 0.6)^2 + (0.4 - 0.5)^2 + \dots} \approx 0.15$$

La similarité cosinus est :

$$\cos(\mathbf{v}, \mathbf{w}) \approx 0.997$$

Pertinence des résultats

Lorsque la distance est faible et que la similarité cosinus est proche de 1, le système considère que les visages sont très similaires. Il peut alors identifier un visage inconnu en le comparant à une base de données.

Conclusion

Les mathématiques permettent de convertir des visages en vecteurs numériques et de les comparer avec rigueur. Ces méthodes sont aujourd'hui utilisées dans la sécurité, la recherche d'images, les réseaux sociaux et bien d'autres domaines.

38. Maths pour la régression linéaire

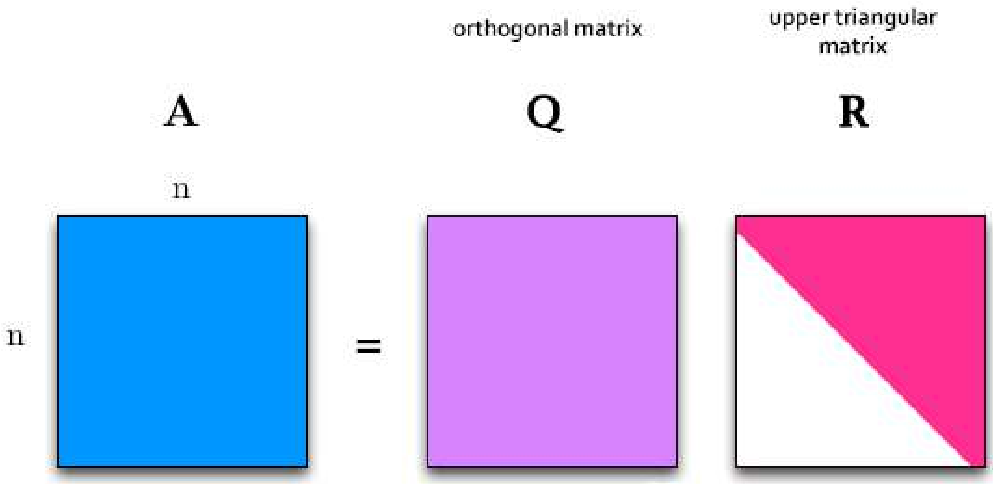


FIGURE 38.1 – QR factorisation : un outil fondamental pour les moindres carrés et la stabilité numérique

Problème

Comment utiliser la factorisation QR pour résoudre des équations linéaires issues de données réelles, et pourquoi cette méthode est-elle aussi efficace ?

De données à une équation linéaire

Prenons un ensemble de données observées $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ qu'on souhaite ajuster par un modèle linéaire de type $y = mx + c$.

Cela nous mène à un système d'équations :

$$A\mathbf{x} = \mathbf{b}, \quad \text{avec} \quad A = \begin{pmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} m \\ c \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

Pourquoi la factorisation QR ?

Lorsque ce système est surdéterminé (plus d'équations que d'inconnues), il n'admet généralement pas de solution exacte. On cherche donc la meilleure approximation possible : c'est l'objectif des **moindres carrés**.

Cette méthode revient à minimiser la norme de l'erreur :

$$\min_{\mathbf{x}} \|A\mathbf{x} - \mathbf{b}\|$$

Et la QR factorisation nous y aide en décomposant :

$$A = QR, \quad \text{où } Q \text{ est orthogonale et } R \text{ est triangulaire supérieure.}$$

On transforme alors le problème en un système plus simple :

$$R\mathbf{x} = Q^T\mathbf{b}$$

Orthogonalisation de Gram-Schmidt

L'un des moyens d'obtenir cette décomposition est l'algorithme de Gram-Schmidt.

- On commence avec les colonnes de A, notées $v_1, v_2, \ldots$.
- On construit une base orthogonale $q_1, q_2, \ldots$ en éliminant les composantes parallèles.

Concrètement :

$$q_1 = \frac{v_1}{\|v_1\|}, \quad q_2 = \frac{v_2 - \text{proj}_{q_1}(v_2)}{\|v_2 - \text{proj}_{q_1}(v_2)\|}, \quad \text{où } \text{proj}_{q_1}(v_2) = \frac{v_2 \cdot q_1}{\|q_1\|^2} q_1$$

Construire la matrice R

La matrice R encode les relations entre les vecteurs d'origine et la nouvelle base orthogonale :

$$R_{ij} = q_i \cdot v_j$$

On obtient alors :

$$A = QR$$

Utilisation en Python

En Python, la QR factorisation est directe :

```
import numpy as np

A = np.array([[1, 1], [1, -1], [1, 1]])
Q, R = np.linalg.qr(A)
```

Cette commande applique en interne une version numérique de Gram-Schmidt (ou une version modifiée plus stable), pour fournir les matrices Q et R.

Exemple numérique

Prenons :

$$A = \begin{pmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \end{pmatrix}$$

On calcule :

$$q_1 = \frac{v_1}{\|v_1\|}, \quad q_2 = \frac{v_2 - \mathrm{proj}_{q_1}(v_2)}{\|v_2 - \mathrm{proj}_{q_1}(v_2)\|}$$

Puis :

$$R = \begin{pmatrix} \|v_1\| & q_1 \cdot v_2 \\ 0 & \|v_2 - \mathrm{proj}_{q_1}(v_2)\| \end{pmatrix}$$

Applications concrètes

— **Résolution de systèmes linéaires** lorsque le système est surdéterminé.

— **Régression linéaire** par moindres carrés.

— **Réduction de dimension** et prétraitement dans des algorithmes de machine learning.

Conclusion

La QR factorisation n'est pas juste une routine Python. Elle repose sur des principes d'algèbre linéaire solides qui garantissent stabilité et efficacité. En comprenant l'orthogonalisation et la structure des matrices Q et R, on exploite tout le potentiel de cet outil pour mieux analyser et modéliser les données.

39. Distance Euclidienne vs Distance de Manhattan : différences et applications

Contexte

En analyse de données et en apprentissage automatique, la manière de mesurer la distance entre deux points influence fortement les performances des algorithmes. Parmi les mesures les plus courantes, on retrouve :

- La distance euclidienne (ou « à vol d'oiseau »), - La distance de Manhattan (ou « à angle droit »).

Définitions mathématiques

Distance euclidienne

C'est la distance la plus intuitive : la ligne droite entre deux points dans un espace multidimensionnel.

$$d_E(x, y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}$$

Caractéristiques :

— Sensible aux valeurs extrêmes.

— Adaptée aux données continues et normalement distribuées.

— Utilisée dans des algorithmes comme K-Means, KNN, PCA.

Exemple

Deux points : $A = (1, 2), B = (4, 6)$

$$d_E(A, B) = \sqrt{(4-1)^2 + (6-2)^2} = \sqrt{9+16} = 5$$

Distance de Manhattan

Elle mesure la somme des différences absolues sur chaque dimension — comme un trajet en ville où l'on suit une grille.

$$d_M(x, y) = \sum_{i=1}^{n} |x_i - y_i|$$

Caractéristiques :

— Moins sensible aux valeurs extrêmes.

— Mieux adaptée aux données ordinales ou hétérogènes.

— Souvent utilisée dans le clustering basé sur la densité.

Exemple

Même points : $A = (1, 2), B = (4, 6)$

$$d_M(A, B) = |4 - 1| + |6 - 2| = 3 + 4 = 7$$

La distance de Manhattan est toujours supérieure ou égale à la distance euclidienne (égalité seulement si une seule dimension varie).

Conséquences analytiques

— **Distance euclidienne :** favorise les clusters sphériques, sensible à l'échelle et aux unités.

— **Distance de Manhattan :** adaptée aux variables avec des unités différentes, plus robuste aux valeurs aberrantes.

Choix stratégique : - Pour des données bien normalisées : distance euclidienne. - Pour des données mixtes ou bruitées : distance de Manhattan.

Tableau comparatif

Critère	Euclidienne	Manhattan
Formule	$\sqrt{\Sigma(x_i - y_i)^2}$	$\Sigma \lvert x_i - y_i \rvert$
Sensibilité aux extrêmes	Forte	Faible
Données adaptées	Continues, gaussiennes	Ordinales, hétérogènes
Algorithmes typiques	K-Means, PCA	DBSCAN, arbres de décision

Illustration graphique

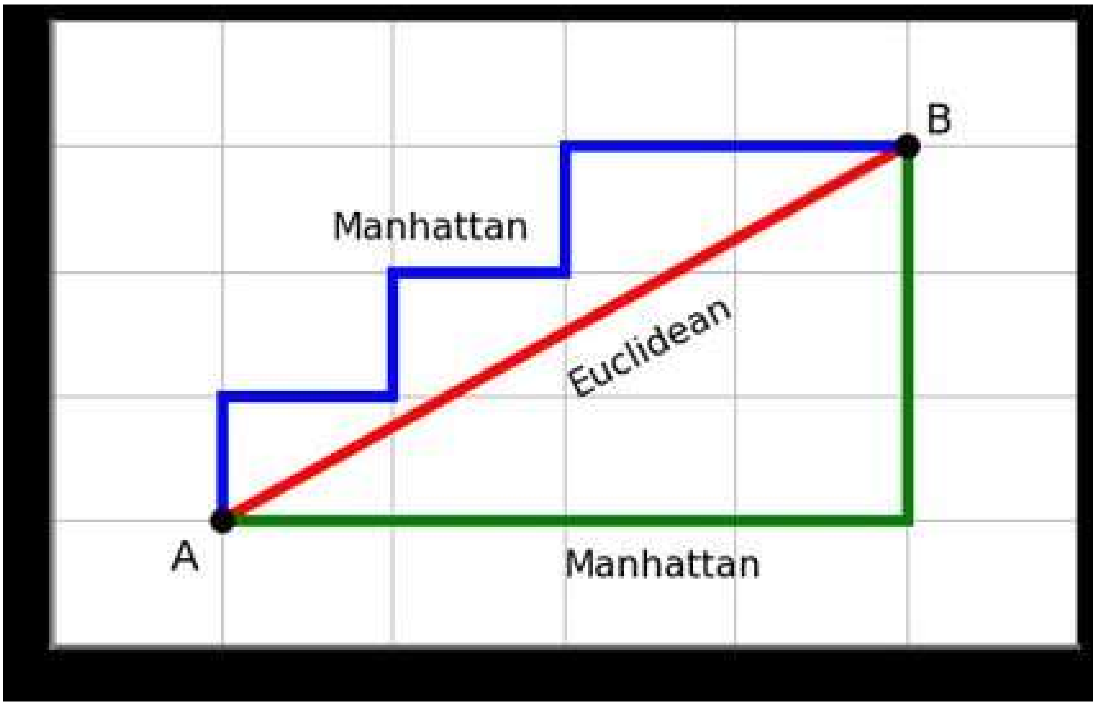


FIGURE 39.1 – Comparaison graphique entre distance euclidienne (ligne droite) et distance de Manhattan (chemin en escalier)

Conclusion

— La distance euclidienne est simple et intuitive, mais sensible aux valeurs extrêmes.

— La distance de Manhattan est plus robuste, idéale pour des données où les dimensions ne sont pas comparables.

— Le bon choix dépend des propriétés des données et de l'objectif d'analyse.

40. Tu connais PCA ? Mais connais-tu les maths derrière ?

Question centrale

Comment l'Analyse en Composantes Principales (PCA) permet-elle de réduire la dimension des données tout en conservant un maximum d'information ?

Objectif de PCA

PCA (Principal Component Analysis) est une méthode utilisée pour transformer un jeu de données en un espace de plus faible dimension, tout en préservant autant que possible la variance présente dans les données.

Autrement dit, PCA cherche à identifier les directions dans lesquelles les données varient le plus, et projette les données sur ces axes pour obtenir une représentation simplifiée.

Les étapes mathématiques de PCA

1. Centrage des données

Soit une matrice de données $X \in \mathbb{R}^{n\times d}$, avec n observations et d variables. On commence par centrer X en retirant la moyenne de chaque colonne :

$$X_{\text{centré}} = X - \bar{X}$$

2. Calcul de la matrice de covariance

$$C = \frac{1}{n} X_{\text{centré}}^T X_{\text{centré}}$$

Cette matrice donne des informations sur la variance et la corrélation entre les variables.

3. Décomposition en valeurs propres

On calcule les valeurs propres λ_i et les vecteurs propres v_i de la matrice C :

$$Cv_i = \lambda_i v_i$$

Les vecteurs propres indiquent les directions principales de variation (axes), et les valeurs propres indiquent la quantité de variance capturée dans chaque direction.

4. Sélection des composantes principales

On trie les valeurs propres par ordre décroissant :

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_d$$

On choisit les k premières composantes, où $k < d$, selon le pourcentage de variance que l'on souhaite conserver.

5. Projection des données

On forme la matrice des composantes principales :

$$W_k = [v_1, v_2, ..., v_k] \in \mathbb{R}^{d \times k}$$

Puis on projette les données :

$$Z = X_{\text{centré}} W_k$$

avec $Z \in \mathbb{R}^{n \times k}$, une représentation réduite du jeu de données.

Pourquoi PCA est optimal ?

PCA a plusieurs propriétés clés :

- Il maximise la variance conservée dans les nouvelles dimensions.
- Les axes principaux sont orthogonaux : les nouvelles variables sont indépendantes.
- La projection minimise l'erreur quadratique moyenne entre les données originales et leur reconstruction.

Liens avec d'autres domaines

Régression linéaire

PCA peut être vu comme une régression dans laquelle la direction de projection minimise la somme des carrés des distances perpendiculaires aux axes.

Probabilités

Dans une distribution normale multivariée, les composantes principales correspondent aux axes qui maximisent la vraisemblance des données.

Applications de PCA

PCA est utilisé dans de nombreux contextes :

— **Compression d'images** : réduction du nombre de pixels sans trop de perte d'information.

— **Détection d'anomalies** : identification de points qui s'éloignent de la structure principale.

— **Réduction de bruit** : suppression des dimensions peu informatives.

— **Visualisation** : transformation de données multidimensionnelles en 2D ou 3D.

Conclusion

PCA est un outil mathématique puissant qui permet de simplifier, visualiser et mieux comprendre des jeux de données complexes. Son efficacité repose sur l'algèbre linéaire (valeurs propres, projections) et il reste un pilier incontournable dans de nombreux domaines liés à l'analyse de données.

41. Comprendre et combattre un nuisible du cacao

Problème

Comment comprendre la croissance de la population du foreur du cacao afin de réduire son impact sur les plantations ?

Contexte : Le foreur du cacao est un insecte nuisible qui cause d'importants dégâts aux plantations. Il passe par plusieurs stades de vie : œuf, larve, adulte mâle et adulte femelle. Chaque stade joue un rôle différent dans la prolifération du nuisible. En modélisant cette dynamique, on peut anticiper et contrôler les infestations.

Cycle de vie du foreur du cacao

Le développement du foreur se fait en quatre stades principaux :

— E : Œufs pondus par les femelles adultes,

— L : Larves (les plus destructrices, car elles se nourrissent de la plante),

— M : Mâles adultes (n'interviennent pas dans la reproduction),

— F : Femelles adultes (assurent la ponte des œufs).

Le cycle suit la trajectoire suivante : Œuf → Larve → Mâle ou Femelle → Nou

Pourquoi modéliser ?

Un modèle mathématique permet de répondre à plusieurs questions cruciales :

— Quels stades cibler pour freiner l'évolution de la population ?

— Quel est l'impact du taux de ponte sur l'invasion ?

— Quelles stratégies seraient les plus efficaces (pesticides, prédateurs naturels, etc.) ?

Cela aide les agriculteurs à intervenir au bon moment, avec les bons moyens.

Les équations du modèle

On modélise l'évolution de chaque catégorie par une équation différentielle :

$$\frac{dE}{dt} = \beta F - \mu_E E$$
$$\frac{dL}{dt} = \mu_E E - \mu_L L$$
$$\frac{dM}{dt} = \frac{1}{2}\mu_L L - \mu_M M$$
$$\frac{dF}{dt} = \frac{1}{2}\mu_L L - \mu_F F$$

Où :

— β est le nombre moyen d'œufs pondus par chaque femelle,

— $\mu_E, \mu_L, \mu_M, \mu_F$ sont les taux de mortalité aux différents stades.

Ces équations décrivent la dynamique globale de la population dans le temps.

Analyse et stratégies de contrôle

L'analyse du modèle révèle plusieurs points clés :

— Les larves sont responsables des dégâts $\rightarrow$ c'est le stade critique à contrôler.

— Une forte valeur de β entraîne une croissance rapide de la population.

— En ciblant les œufs (μ_E) et les larves (μ_L), on réduit efficacement le nombre d'adultes reproducteurs.

Recommandation

L'intervention la plus efficace est d'augmenter la mortalité au stade larvaire (par pesticides ou prédateurs naturels), car c'est à ce moment que le nuisible est le plus vulnérable et destructeur.

Conclusion

Ce modèle mathématique met en lumière les mécanismes clés qui contrôlent la croissance de la population du foreur du cacao. Il sert d'outil d'aide à la décision pour les agriculteurs, en leur permettant d'intervenir de manière ciblée et stratégique afin de protéger leurs cultures.

42. Comprendre la valeur des options

Définition

Une option est un contrat financier qui donne le droit, mais non l'obligation, d'acheter (option d'achat ou call) ou de vendre (option de vente ou put) un actif à un prix fixé (prix d'exercice), à une date donnée ou avant.

Exemple : Si vous avez une option vous donnant le droit d'acheter une action à 100 euros alors que le marché la propose à 120 euros, vous réalisez un gain de 20 euros.

Pourquoi acheter des options ?

— **Se couvrir contre les risques :** Une entreprise peut verrouiller un prix de vente minimum.

— **Spéculer :** Un investisseur peut parier sur la hausse ou la baisse d'un actif avec une mise limitée.

— **Stratégies flexibles :** Les options permettent de construire des combinaisons complexes d'investissement.

Comment fonctionne une option ?

— **Option d'achat (call)** : donne le droit d'acheter un actif à un prix déterminé.

— **Option de vente (put)** : donne le droit de vendre un actif à un prix déterminé.

— **Gain potentiel** : si le prix de l'actif monte, une option d'achat devient plus intéressante.

— **Perte maximale** : limitée au coût de l'option (la prime).

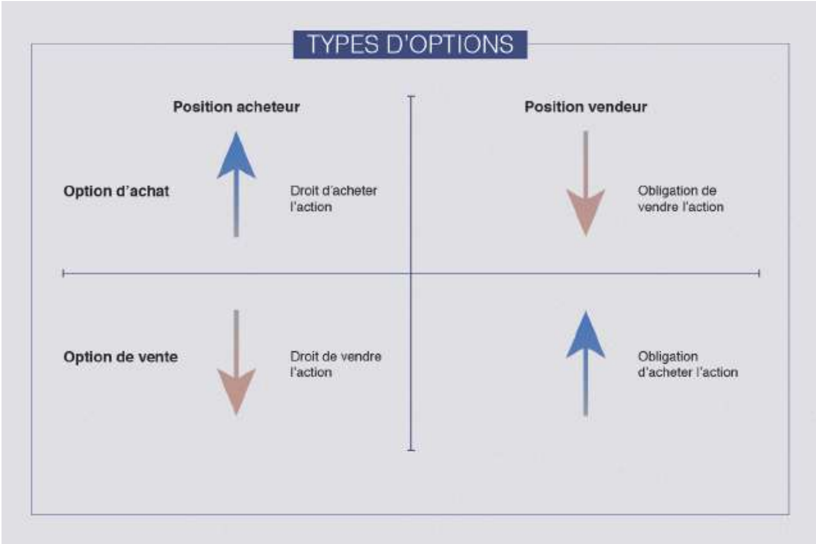


FIGURE 42.1 – Illustration du fonctionnement d'une option

Modélisation mathématique des options

On suppose que le prix $S(t)$ de l'actif suit un mouvement brownien géométrique :

$$dS(t) = \mu S(t)dt + \sigma S(t)dW(t)$$

où :

— μ est le taux de croissance espéré,

— σ est la volatilité (écart-type des variations),

— $dW(t)$ représente un mouvement brownien standard.

Formule de Black-Scholes

La valeur théorique $V(S,t)$ d'une option vérifie l'équation :

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0$$

Pour une option d'achat européenne, la formule explicite est :

$$C(S,t) = SN(d_1) - Ke^{-r(T-t)}N(d_2)$$

avec :

$$d_1 = \frac{\ln\left(\frac{S}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}}, \quad d_2 = d_1 - \sigma\sqrt{T-t}$$

où N est la fonction de répartition de la loi normale.

Exemple numérique

Paramètres :

— $S = 100$ €, $K = 105$ €, $r = 5\%$, $\sigma = 20\%$, $T - t = 0.5$

Calcul :

$$d_1 = 0{,}106, \quad d_2 = -0{,}034$$

$$C = 100 \times N(0{,}106) - 105 \times e^{-0{,}05 \times 0{,}5} \times N(-0{,}034) \approx 5{,}27 \text{ €}$$

Scénarios possibles

Scénario 1 : L'option n'est pas exercée

Le prix reste inférieur à 105 €, l'option expire sans valeur :

$$\text{Perte} = \text{prime payée} = 5{,}27\pounds$$

Scénario 2 : L'option est exercée

Le prix monte à 120 € :

$$\text{Gain} = (120 - 105) - 5{,}27 = 9{,}73\pounds$$

Applications pratiques des mathématiques financières

Les modèles permettent de :

— **Évaluer des produits financiers complexes**,

— **Prendre des décisions éclairées** sur les couvertures ou investissements,

— **Gérer les risques de portefeuille**.

Limites du modèle de Black-Scholes

— **Volatilité constante** : peu réaliste dans la vraie vie.

— **Pas de sauts de prix** : le modèle n'intègre pas les crashs ou pics.

— **Pas de coûts de transaction** : hypothèse simplificatrice.

Conclusion

Grâce aux mathématiques, on peut modéliser, comprendre et anticiper les comportements de produits financiers comme les options. Cela permet d'améliorer la prise de décision en finance et de mieux gérer les incertitudes du marché.

43. Comment les maths gèrent le trafic routier ?

FIGURE 43.1 – La modélisation mathématique au service de la fluidité routière

Pourquoi les bouchons ?

Le trafic routier est souvent difficile à gérer. Comment prévoir et éviter les embouteillages, et optimiser la fluidité de la circulation ?

Réponse : Les mathématiques, à travers des modèles comme l'équation de Lighthill-Whitham-Richards (LWR), nous aident à comprendre et gérer le trafic en ville !

Modélisation mathématique du trafic

Le modèle LWR repose sur une équation de conservation. On modélise :

$$\frac{\partial \rho}{\partial t} + \frac{\partial q}{\partial x} = 0$$

où :

- $\rho(x, t)$ est la densité des véhicules (en nombre de voitures par kilomètre),
- $q(x, t) = \rho(x, t) \cdot v(x, t)$ est le flux de trafic,
- $v(x, t)$ est la vitesse moyenne des véhicules.

Ce modèle permet de prédire la formation et la disparition des embouteillages, de la même façon que les vagues dans un fluide.

Exemple : formation d'un embouteillage

- **Trafic fluide** : À faible densité, les véhicules roulent rapidement et le flux augmente.
- **Trafic dense** : Quand la densité dépasse un certain seuil ρ_{cr}, les voitures ralentissent et des bouchons apparaissent.

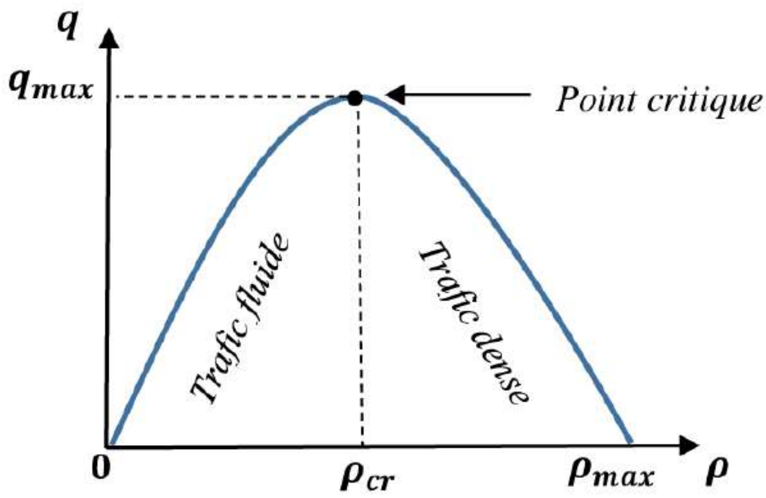


FIGURE 43.2 – Relation entre densité ρ et flux q : le point critique marque la transition vers l'embouteillage

Ce type de graphique montre que le flux de trafic atteint un maximum pour une densité critique ρ_{cr}, au-delà de laquelle le système sature.

Applications concrètes du modèle

Grâce à ces équations, on peut :

— **Optimiser les feux de circulation** pour maximiser la fluidité,

— **Prédire les embouteillages** en amont et recommander des itinéraires alternatifs,

— **Améliorer les GPS intelligents** qui s'appuient sur des modèles prédictifs pour proposer les meilleurs trajets.

Conclusion

Les équations différentielles comme celle de LWR sont au cœur des systèmes modernes de gestion du trafic. En comprenant la dynamique de la circulation, les mathématiques deviennent un outil puissant pour réduire les embouteillages et améliorer la mobilité urbaine.

44. Pourquoi les Logarithmes ?

Problème

Comment effectuer facilement des multiplications ou divisions complexes à une époque sans calculatrice ni ordinateur ?

Contexte historique

Les logarithmes ont été inventés au début du 17e siècle par John Napier. À cette époque, les scientifiques avaient besoin d'outils pour accélérer les calculs astronomiques et éviter les erreurs lors des longues multiplications.

Idée principale

Les logarithmes transforment des multiplications en additions, et des divisions en soustractions.

Exemples d'utilisation historique

— **Astronomie** : déterminer les positions des planètes et des étoiles.

— **Navigation** : calcul des distances et des trajectoires maritimes.

— **Finance** : calculs d'intérêts composés sur le long terme.

Simplification des calculs

Voici deux formules classiques :

$$a \times b = 10^{\log(a)+\log(b)} \qquad \frac{a}{b} = 10^{\log(a)-\log(b)}$$

Ainsi, au lieu de multiplier directement 5000×3200, on utilise les tables logarithmiques, on additionne les logarithmes, puis on reprend l'antilogarithme.

Aujourd'hui : des usages toujours essentiels

Même si on n'utilise plus de tables logarithmiques à la main, le concept reste fondamental dans plusieurs disciplines :

- **Informatique** : complexité algorithmique ($O(\log n)$).
- **Physique** : mesure des décibels, des magnitudes sismiques.
- **Chimie** : calcul du pH, défini comme $-\log[H^+]$.
- **Économie et Data Science** : transformation logarithmique des données.

Les bases et formules essentielles

Définition : Si $a^x = b$, alors $x = \log_a(b)$.
Formules fondamentales :

- $\log_a(xy) = \log_a(x) + \log_a(y)$
- $\log_a(x/y) = \log_a(x) - \log_a(y)$
- $\log_a(x^n) = n \log_a(x)$
- $\log_a(b) = \dfrac{\log_c(b)}{\log_c(a)}$ (changement de base)

Bases fréquentes :

- $\log_{10}(x)$: logarithme décimal.
- $\ln(x)$ ou $\log_e(x)$: logarithme naturel (base $e \approx 2{,}718$), très utilisé en analyse.

Conclusion

Les logarithmes sont un exemple parfait d'outil mathématique né d'un besoin pratique. Ils ont révolutionné les sciences dès leur invention, et continuent aujourd'hui d'être indispensables dans la modélisation, l'analyse de données, et la compréhension des phénomènes exponentiels.

45. Optimiser les Cultures

Problème

Un entrepreneur agricole cultive à la fois du plantain et de la patate douce. Ces deux plantes partagent les mêmes parcelles et se disputent les ressources naturelles : eau, lumière, nutriments du sol. Comment modéliser cette compétition pour aider cet entrepreneur à optimiser sa production ?

Réponse

On peut utiliser un modèle mathématique inspiré de la biologie des populations : le **modèle de compétition de Lotka-Volterra**, adapté ici à l'agriculture.

Contexte agricole

Dans de nombreuses régions tropicales d'Afrique, on cultive ensemble des plantes comme :

— Le **plantain** — grand, feuillu, gourmand en ressources.

— La **patate douce** — rampante, productive, mais vulnérable au manque de lumière.

Ces cultures peuvent se nuire mutuellement si elles sont mal réparties. D'où l'intérêt de modéliser leur croissance conjointe.

Modéliser la croissance du plantain

On note $P(t)$ la production de plantain au temps t. Son évolution est donnée par :

$$\frac{dP}{dt} = r_P P \left(1 - \frac{P}{K_P}\right) - \alpha_{PD} PD$$

Interprétation des paramètres :

— r_P : taux de croissance du plantain.

— K_P : rendement maximal que le sol peut supporter (capacité de charge).

— α_{PD} : effet négatif de la patate douce sur le plantain.

La dernière partie de l'équation représente la perte de rendement due à la compétition : chaque unité de patate douce gêne le développement du plantain.

Modéliser la croissance de la patate douce

De façon similaire, on note $D(t)$ la production de patate douce. On obtient :

$$\frac{dD}{dt} = r_D D \left(1 - \frac{D}{K_D}\right) - \alpha_{DP} DP$$

— r_D : taux de croissance de la patate douce.

— K_D : capacité de charge pour la patate douce.

— α_{DP} : effet du plantain sur la patate douce (souvent via l'ombre portée).

Ces deux équations forment un système dynamique décrivant l'interaction entre les deux cultures.

Difficultés d'application concrète

— **Paramètres incertains** : Les valeurs de r, K, et α dépendent du climat, du type de sol, de la densité de plantation.

— **Aléas climatiques** : Sécheresse, maladies, ou attaques de parasites peuvent perturber les prévisions.

— **Interaction non linéaire** : Les effets de l'ombre, de l'humidité ou de la densité peuvent amplifier ou atténuer les interactions de manière complexe.

Objectif : optimiser les rendements

En simulant ce système d'équations avec différents paramètres (densités de plantation, calendrier des semis, apports d'eau, etc.), on peut :

— Repérer les configurations qui maximisent le rendement total.

— Trouver des compromis durables entre les deux cultures.

— Réduire les pertes liées à la mauvaise concurrence des plantes.

Conclusion

Les équations différentielles offrent aux agriculteurs une vision plus claire des interactions entre leurs cultures. Elles ne remplacent pas l'expérience du terrain, mais elles permettent de tester des scénarios avant de les mettre en œuvre. Une modélisation réussie peut donc conduire à des pratiques plus durables, plus rentables et mieux adaptées aux ressources disponibles.

46. Créer des Mandalas avec la Multiplication

Une idée surprenante

Et si je te disais que tu peux créer des mandalas simplement en... faisant des tables de multiplication ? Oui, grâce aux maths, l'art et la beauté géométrique se rejoignent. Explorons comment les motifs issus de la multiplication modulaire peuvent donner naissance à des figures fascinantes.

Une table de multiplication circulaire ?

Prenons un cercle avec n points également espacés. On les numérote de 0 à $n-1$. Pour chaque point i, on trace un segment vers le point $(i \times k) \mod n$.

Formule utilisée :

$$f(i, k, n) = (i \times k) \mod n$$

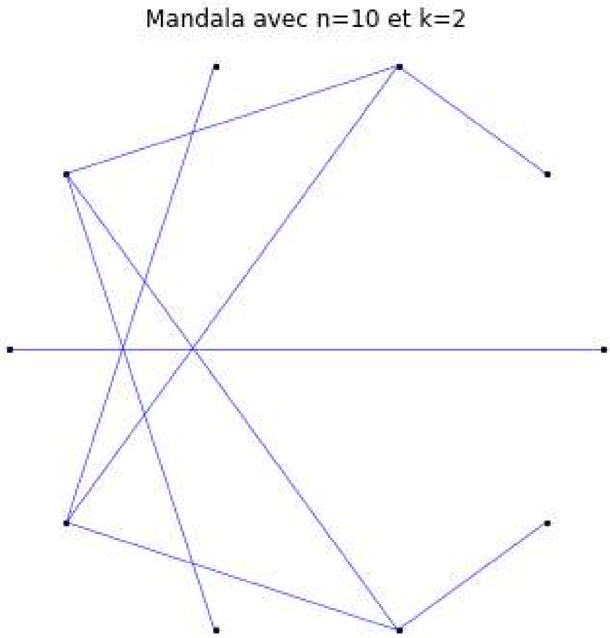
Le paramètre k est un multiplicateur, et n représente le nombre total de points sur le cercle.

Exemple avec 10 points et $k = 2$

Avec $n = 10$ et $k = 2$, on relie :

$$0 \to 0$$
$$1 \to 2$$
$$2 \to 4$$
$$3 \to 6$$
$$4 \to 8$$
$$5 \to 0$$
etc.

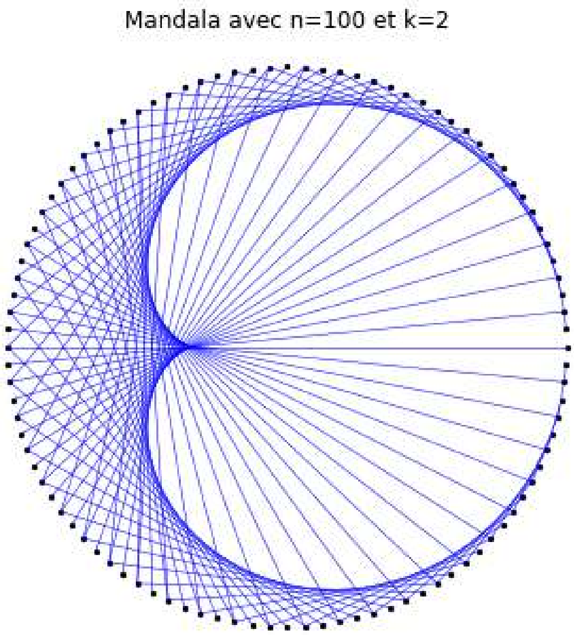
Cela donne un motif géométrique étonnamment structuré, bien que très simple à construire.



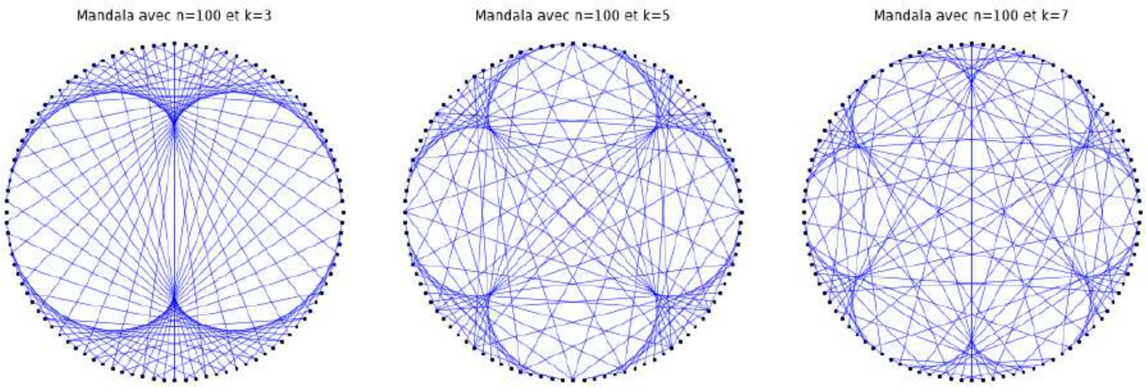
Plus de points, plus de beauté

Si on augmente n, les formes deviennent plus complexes, presque hypnotiques :

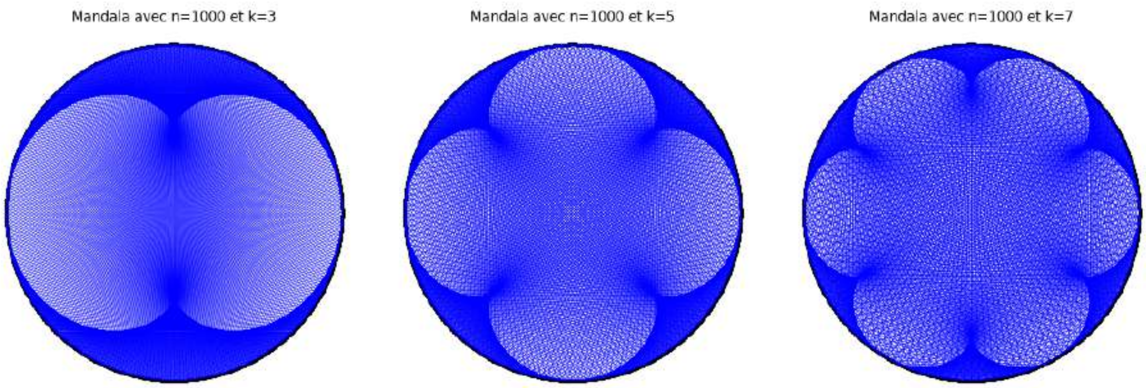
— $n = 100$, $k = 2$:



— $n = 100$, $k = 3, 5, 7$:



— $n = 1000$, $k = 3, 5, 7$:



Ce que nous apprennent ces figures

Ces dessins ne sont pas juste beaux. Ils illustrent plusieurs idées mathématiques :

— La notion de **modulo**, ou reste de division.

— L'effet d'une multiplication répétée dans un système cyclique.

— L'émergence de motifs **symétriques** à partir d'une règle simple.

En ajustant juste k et n, on obtient toute une famille de mandalas. On peut même en faire un projet artistique basé sur les mathématiques.

Et en Python ?

Voici un code simple pour générer de tels mandalas :

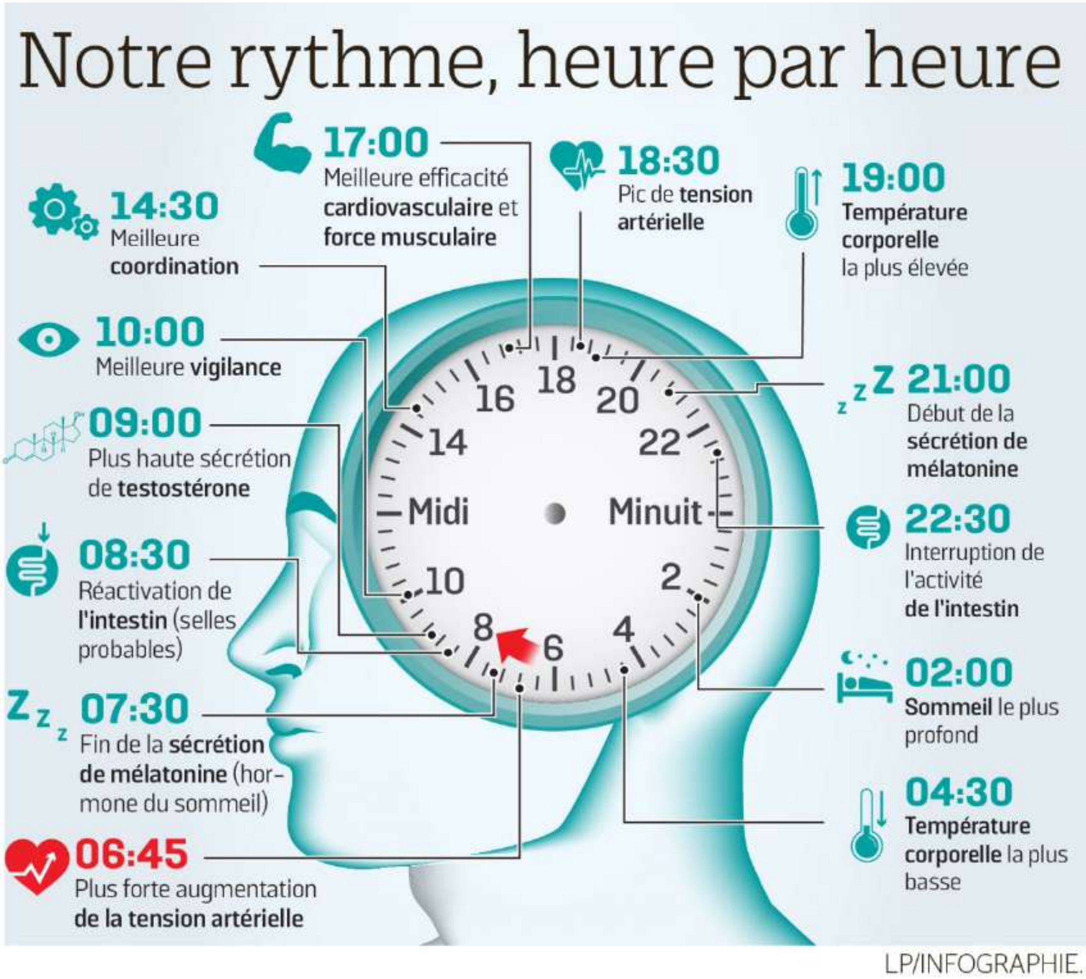
```
import matplotlib.pyplot as plt
import numpy as np

n = 200  # nombre de points
k = 2    # multiplicateur

angles = np.linspace(0, 2*np.pi, n, endpoint=False)
points = np.array([[np.cos(a), np.sin(a)] for a in angles])

fig, ax = plt.subplots(figsize=(6,6))
ax.set_aspect('equal')
ax.axis('off')

for i in range(n):
    j = (i * k) % n
    ax.plot([points[i][0], points[j][0]],
            [points[i][1], points[j][1]],
            color='blue', linewidth=0.5)

plt.show()
```

N'hésite pas à jouer avec les paramètres n et k. Tu verras apparaître des rosaces, des spirales, des figures étoilées. . .

Conclusion

Les mandalas issus des tables de multiplication modulaire montrent que les mathématiques sont partout, même dans l'art. Un simple produit suivi d'un modulo peut donner naissance à un monde visuel riche et harmonieux. Et si tu ne t'es jamais senti · e créatif · ve avec les maths, c'est peut-être le moment d'ouvrir un carnet… ou un terminal Python.

47. Rôle des Maths dans Nos Rythmes Quotidiens



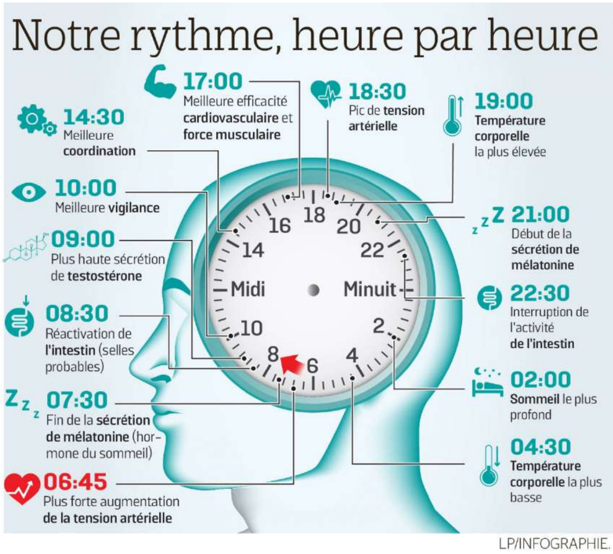
Qu'est-ce qu'une Horloge Circadienne ?

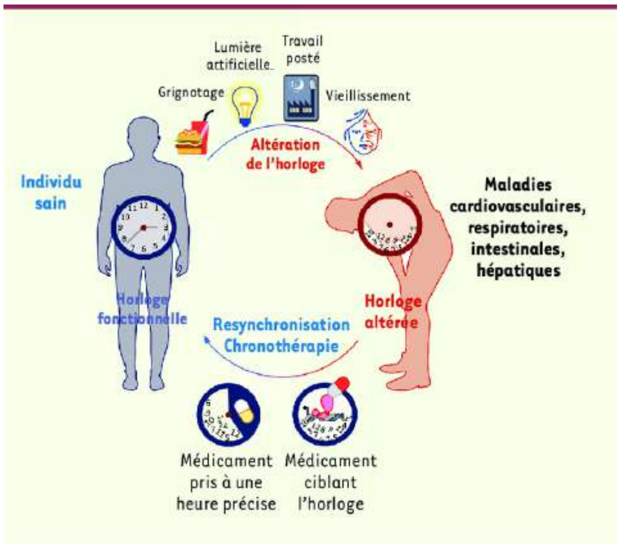
Introduction

L'horloge circadienne est notre "montre interne" qui régule des cycles de 24 heures, influençant notre sommeil, notre énergie, et nos processus biologiques. Elle se synchronise principalement avec la lumière du jour. Comprendre comment cette horloge fonctionne et reste en phase avec l'environnement est crucial pour notre santé et bien-être.

Les Horloges dans le Corps Humain

Voici les principaux lieux dans le corps où se trouvent des horloges circadiennes : noyau suprachiasmatique (SCN) du cerveau, foie, cœur, pancréas, reins, intestins, poumons, peau, muscles squelettiques, tissu adipeux (graisse).





Modélisation Mathématique de l'Horloge Circadienne

Pour modéliser une horloge circadienne, on peut utiliser une équation qui décrit la production et la dégradation d'une protéine P qui contrôle le rythme.

— La protéine P augmente et diminue de façon cyclique.

— Son niveau influence directement notre rythme de veille et de sommeil.

Équation du modèle très basique

$$\frac{dP}{dt} = \frac{\alpha}{1 + P^n} - \gamma P$$

Décomposition de l'Équation

— $\frac{\alpha}{1+P^n}$ modélise la production de P, qui diminue lorsque P est élevé.

— $-\gamma P$ représente la dégradation naturelle de P, nécessaire pour que le cycle recommence.

Pourquoi la synchronisation est-elle importante ?

L'alignement de l'horloge circadienne avec le cycle jour/nuit est essentiel pour que notre corps fonctionne efficacement. La désynchronisation peut causer des troubles du sommeil et autres effets sur la santé.

Le Modèle Mathématique de Goodwin

Le modèle de Goodwin est utilisé pour modéliser les oscillations circadiennes. Il repose sur trois variables :

- x : ARN messager.
- y : Protéine produite à partir de x.
- z : Inhibiteur qui régule la production de x.

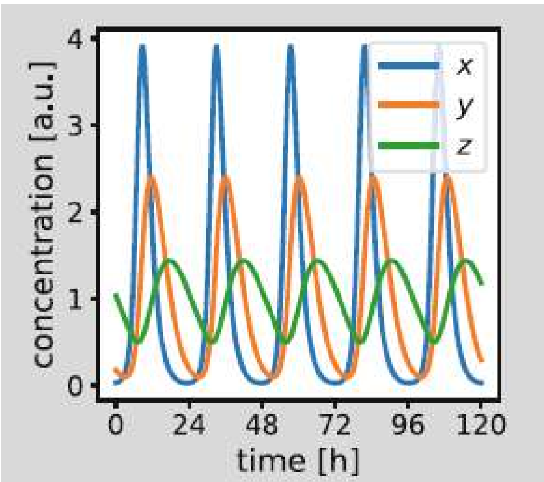
Les équations associées sont :

$$\frac{dx}{dt} = \nu_1 \frac{K_1^n}{K_1^n + z^n} - \nu_2 \frac{x}{K_2 + x}$$

$$\frac{dy}{dt} = \nu_3 x - \nu_4 \frac{y}{K_4 + y}$$

$$\frac{dz}{dt} = \nu_5 y - \nu_6 \frac{z}{K_6 + z}$$

Exemple de Synchronisation Circadienne



Crédit : Marta del Olmo et al

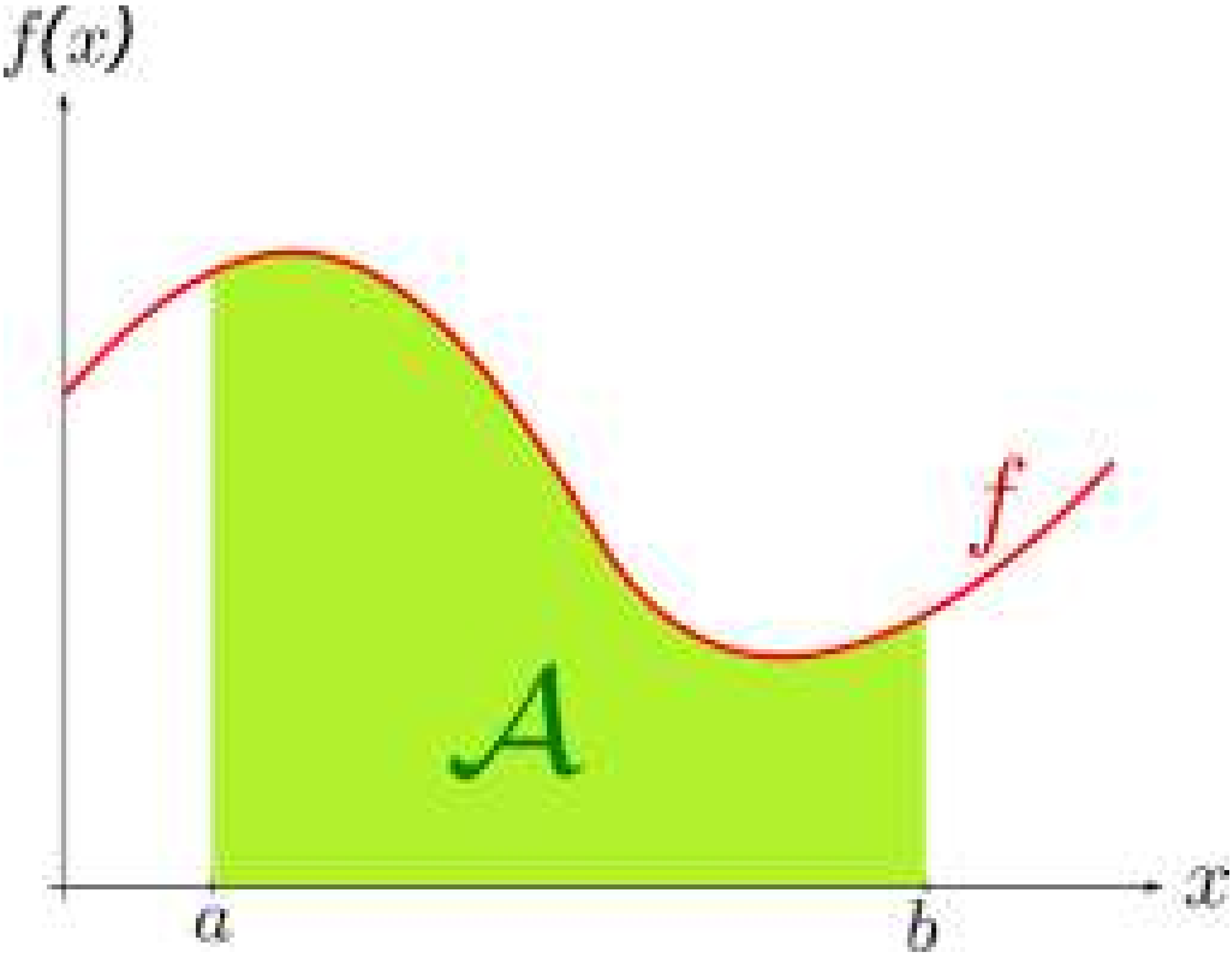
Conclusion

Ce modèle simple montre comment une horloge biologique peut se synchroniser naturellement avec l'environnement, en maintenant un rythme régulier essentiel pour notre santé.

Conclusion Générale

Les mathématiques nous aident à mieux comprendre l'horloge circadienne et son rôle dans notre fonctionnement quotidien. Les modèles présentés montrent comment un phénomène biologique complexe peut être simulé à l'aide d'équations simples mais puissantes.

48. Pourquoi calculer des intégrales ?



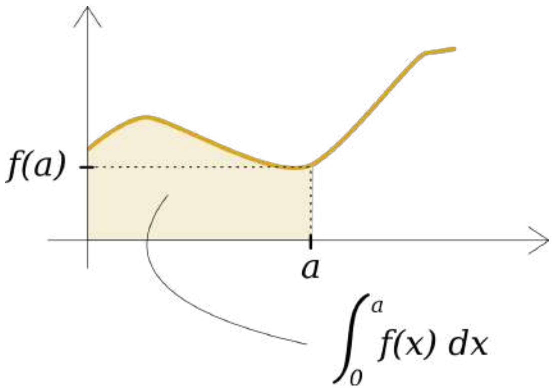
Un Problème par Exemple

Pourquoi calculer des intégrales ?

Pour des formes simples comme les carrés ou rectangles, on calcule facilement l'aire. Cependant, pour des formes complexes comme une portion de plage, les calculs deviennent compliqués. C'est là que les **intégrales** aident.

Définition de l'intégrale

L'intégrale sert à calculer l'aire sous une courbe, le volume d'un solide, ou l'accumulation sur un intervalle.



Application : Vitesse et Distance

Si vous voulez calculer la distance parcourue par une voiture en fonction de sa vitesse variable, l'intégrale permet de trouver la distance totale en additionnant les petits changements de position sur chaque intervalle de temps.

$$\text{Distance} = \int_a^b v(t)\,dt$$

C'est un outil pour additionner une infinité de petites quantités.

Optimisation des Stocks avec les Intégrales

Contexte

Dans une entreprise, la gestion efficace des stocks est essentielle pour réduire les coûts et optimiser l'approvisionnement. Comment peut-on utiliser les intégrales pour prendre de meilleures décisions ?

Supposons qu'une entreprise de distribution reçoive des produits de manière continue. La demande pour ces produits varie au cours du temps. Pour éviter un manque ou un surplus de stock, il faut calculer l'accumulation de produits sur une période donnée :

$$S(t) = \int_0^T \text{Entrées}(t) - \text{Sorties}(t)\,dt$$

Cela permet de suivre l'évolution des stocks à chaque instant.

Intégrales en Économie

Pourquoi les Intégrales en Économie ?

Les intégrales permettent de calculer les bénéfices, les coûts cumulés ou encore l'accumulation de capital.

Exemple : Si $R(t)$ représente le revenu à un instant t et $C(t)$ le coût à ce même instant, alors le bénéfice total sur une période de t_1 à t_2 est :

$$\text{Bénéfice total} = \int_{t_1}^{t_2} (R(t) - C(t))\, dt$$

Cela permet d'avoir une vision précise des profits accumulés.

Croissance du Capital et Intérêts Composés

Intérêts composés

Les intégrales sont également utilisées pour modéliser la croissance d'un capital à intérêts composés, en prenant en compte l'accumulation continue d'intérêts.

Exemple : Si un capital C_0 est placé avec un taux d'intérêt variable $r(t)$, le montant accumulé entre t_1 et t_2 est :

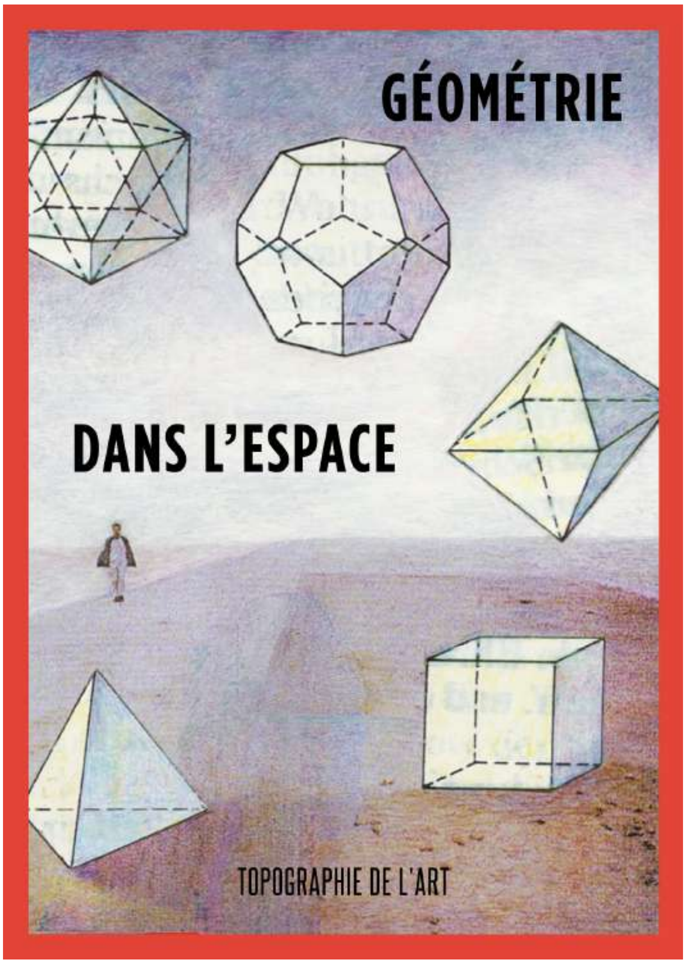
$$C(t) = C_0 \cdot e^{\int_{t_1}^{t_2} r(t)\, dt}$$

Ce modèle permet d'estimer la croissance continue du capital.

Conclusion

Les intégrales sont omniprésentes : elles interviennent dans l'aménagement des plages, en économie pour estimer des bénéfices ou suivre la croissance d'un capital, ou encore pour gérer efficacement des stocks. C'est un outil fondamental pour additionner des quantités variables sur une période continue.

49. Pourquoi fait-on la Géométrie de l’Espace ?



Pourquoi faire la géométrie de l’espace ?

Utilité de la Géométrie de l’Espace

La géométrie de l’espace est utilisée pour modéliser et comprendre des objets en trois dimensions, ce qui est essentiel dans des domaines comme la construction, l’ingénierie, la conception d’objets et même l’aéronautique.

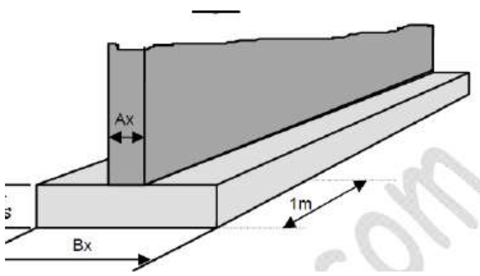
Elle permet de calculer les volumes, les surfaces et les distances entre des objets dans l'espace, facilitant ainsi des projets pratiques comme la construction d'un bâtiment ou le design d'un produit industriel.

Cas 1 : Calcul de volume de béton en construction

Problème

Comment calculer le volume de béton nécessaire pour remplir une fondation trapézoïdale ?

$$V = \frac{1}{2} \times (B_1 + B_2) \times h \times L$$

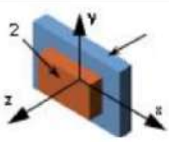
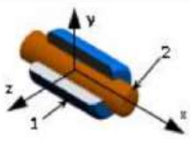
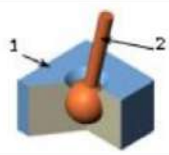


Cas 2 : Conception mécanique et optimisation

Problème

Comment optimiser la surface de contact d'une pièce mécanique pour minimiser les frottements ?

Modélisation de la surface en utilisant des solides de révolution, et calcul de la surface latérale à l'aide de la géométrie 3D pour optimiser les points de contact.

Nom du Contact:	Contact Plan	Contact Cylindrique	Contact Sphérique
Exemple:			
Nature du contact:	Plan	Cylindre	Sphère

Cas 3 : Optimisation du matériau pour une imprimante 3D

Problème

Comment calculer le volume de matériau à utiliser pour un prototype de pièce imprimée en 3D ?

$$V_{\text{total}} = V_{\text{cylindre}} + V_{\text{demi-sphère}} = \pi r^2 h + \frac{4}{3}\pi r^3$$

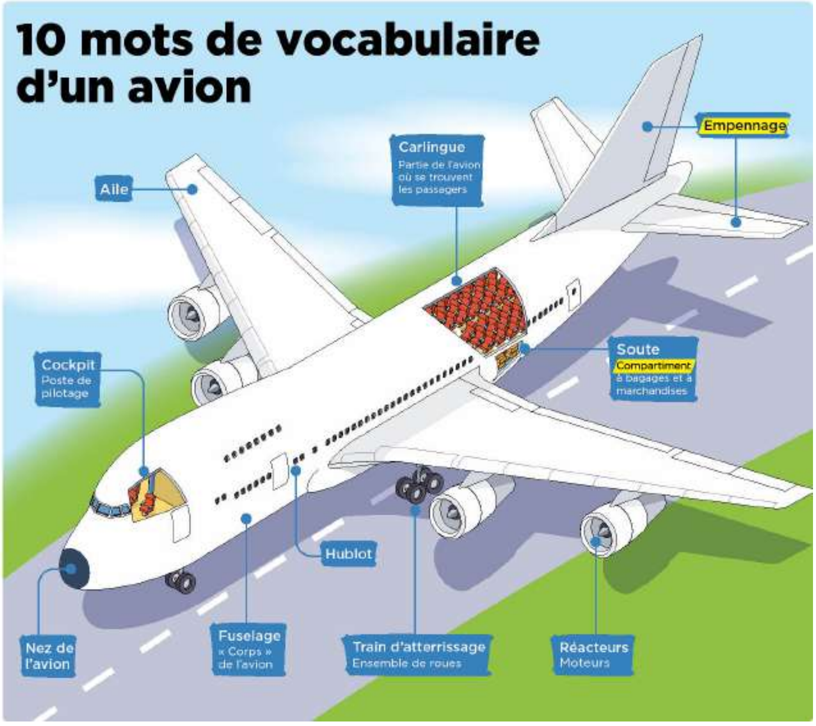
Cas 4 : Calcul de surface pour un revêtement d'avion

Problème

Comment calculer la surface totale d'une carlingue d'avion pour un revêtement précis ?

On modélise la carlingue comme un cylindre :

$$A = 2\pi r h$$

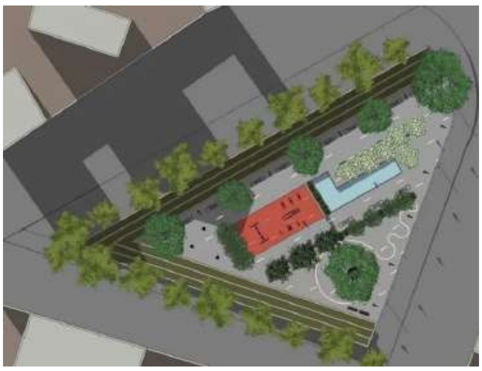


Cas 5 : Aménagement d'un espace urbain

Problème

Comment concevoir un parc public en calculant l'espace optimal à utiliser tout en respectant les contraintes géométriques ?

Utilisation des formules géométriques pour maximiser l'espace vert tout en intégrant des structures (aires de jeux, fontaines) et des chemins.

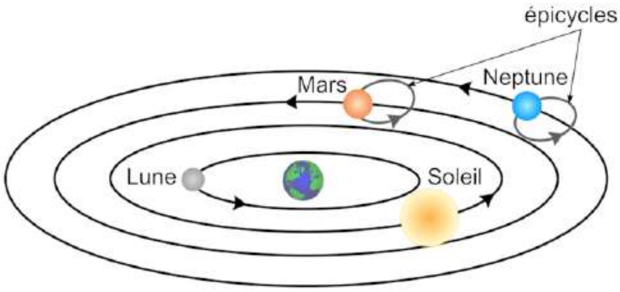


Cas 6 : Observation des planètes

Problème

Comment calculer la distance entre deux planètes en tenant compte de leurs trajectoires ?

Utilisation de la géométrie de l'espace et des orbites elliptiques (sections coniques) pour déterminer la distance minimale lors d'une opposition planétaire.



Cas 7 : Création d'environnements virtuels

Problème

Comment créer des environnements virtuels réalistes pour la réalité virtuelle ?

La géométrie de l'espace est utilisée pour modéliser des objets 3D avec des textures et des proportions réalistes.

Conclusion

La géométrie de l'espace permet de modéliser, concevoir et construire des objets et des structures en trois dimensions. De la construction à l'ingénierie, en passant par la modélisation 3D et la réalité virtuelle, elle constitue une base pour comprendre les formes qui nous entourent.

50. Pourquoi le Barycentre à l'école ?

Contexte

Question

Comment déterminer le centre optimal pour une infrastructure dans une ville ?

Le barycentre est utilisé pour trouver le centre de gravité d'une répartition de population dans une ville, facilitant la localisation optimale pour des infrastructures, comme des hôpitaux ou des centres commerciaux.

Définition du Barycentre

Qu'est-ce qu'un barycentre ?

Le barycentre, aussi appelé centre de gravité, est un point qui représente la position moyenne pondérée d'un ensemble de points en fonction de leurs poids respectifs.

Formule Mathématique :

$$x_G = \frac{\sum_{i=1}^{n} m_i x_i}{\sum_{i=1}^{n} m_i}, \quad y_G = \frac{\sum_{i=1}^{n} m_i y_i}{\sum_{i=1}^{n} m_i}$$

— m_i : Poids ou importance du point i (par exemple, population d'une zone)

— (x_i, y_i) : Coordonnées du point i

Application : Exemple Pratique

Considérons une ville où chaque zone représente un quartier avec sa propre population :

— Zone A : près du parc avec 5000 habitants.

— Zone B : au centre-ville avec 3000 habitants.

— Zone C : au bord de la rivière avec 2000 habitants.

Coordonnées : $(1, 3), (4, 5), (6, 1)$

Question

Où devrions-nous placer une nouvelle école pour minimiser les déplacements des habitants ?

Calcul du Barycentre pour un Emplacement Optimal

Utilisons la formule du barycentre :

$$x_G = \frac{5000 \times 1 + 3000 \times 4 + 2000 \times 6}{5000 + 3000 + 2000} = \frac{33000}{10000} = 3.3$$

$$y_G = \frac{5000 \times 3 + 3000 \times 5 + 2000 \times 1}{5000 + 3000 + 2000} = \frac{29000}{10000} = 2.9$$

Résultat

Le point optimal pour l'infrastructure se situe donc autour de $(3.3, 2.9)$, au centre des trois zones.

Ce barycentre permet de placer l'école à un endroit stratégique, facilitant l'accès pour un maximum d'habitants.

Avantages du Barycentre en Urbanisme

- Optimisation des distances : minimiser les déplacements pour la majorité de la population.
- Accessibilité accrue : faciliter l'accès aux infrastructures.
- Réduction des coûts : planification stratégique des emplacements.

Limites et Améliorations

Limites :

- Hypothèses simplifiées sur la répartition uniforme de la population.
- Ne prend pas en compte les obstacles géographiques.

Améliorations possibles :

- Utiliser des modèles intégrant la topographie.
- Prendre en compte les flux de déplacement.

Conclusion

L'utilisation du barycentre dans la planification urbaine permet d'optimiser l'emplacement des infrastructures, réduisant les coûts et augmentant l'accessibilité.

Ressources et inspirations

Ce livre est né de ma passion pour les mathématiques appliquées à la vie réelle, mais aussi de nombreuses lectures, vidéos, discussions, et recherches.

Voici quelques références, sites, ouvrages et plateformes qui m'ont inspirée dans la rédaction de ces applications.

Sites et plateformes

- https://netflixtechblog.com – Le blog technique de Netflix, pour mieux comprendre leurs systèmes de recommandation.
- https://mathigon.org – Une plateforme interactive pour explorer les maths autrement.
- https://towardsdatascience.com – Articles simples sur la data science et les algorithmes.
- https://ourworldindata.org – Statistiques mondiales, utiles pour les chapitres sur la santé ou les inégalités.

Ouvrages et lectures marquantes

- Ian Stewart – Les mathématiques du quotidien
- Jordan Ellenberg – How Not to Be Wrong : The Power of Mathematical Thinking
- Cathy O'Neil – Weapons of Math Destruction
- Livres de vulgarisation en philosophie, sociologie, et économie, pour enrichir les contextes d'application

Logiciels et outils utilisés

- **Python** – pour illustrer les exemples de TF-IDF, SIR, ou recommandations

- **LaTeX** – pour la mise en forme du livre
- **Canva et Inkscape** – pour certaines illustrations

Remerciements particuliers

Merci à celles et ceux qui, de près ou de loin, m'ont inspirée ou encouragée à écrire ce livre. À ma communauté en ligne, à mes élèves, et à ma fille qui me pousse chaque jour à rendre les mathématiques plus vivantes et accessibles.