100 Notions de maths en IA, LLM et Data Science

Dr. Clotilde Djuikem

10 Notions d'Algèbre

1. Produit Matriciel

Définition : Le produit matriciel est une opération entre deux matrices A et B qui donne une nouvelle matrice C définie par :

$$C = AB, \quad C_{ij} = \sum_k A_{ik} B_{kj}$$

Applications

- Utilisé dans la propagation avant des réseaux neuronaux.
- Permet les transformations linéaires des données.
- Modélisation des systèmes dynamiques en ingénierie et en physique.

2. Décomposition en Valeurs Singulières (SVD)

Définition : La SVD est une factorisation de matrice exprimée comme suit :

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

Applications

- Réduction de dimension, notamment en analyse en composantes principales (PCA).
- Compression des données en apprentissage automatique.
- Filtrage collaboratif dans les systèmes de recommandation.

3. Valeurs Propres et Vecteurs Propres

Définition : Une valeur propre λ et un vecteur propre v d'une matrice A satisfont :

$$Av = \lambda v$$

Applications

- Utilisé pour les méthodes spectrales en clustering.
- Permet le partitionnement de graphes et la segmentation d'images.
- Analyse de la stabilité des systèmes dynamiques en ingénierie et en physique.

4. Multiplication de Tenseurs

Définition : Généralisation du produit matriciel aux tenseurs de dimensions supérieures.

$$C_{ijk} = \sum_{l} A_{ijl} B_{lk}$$

Applications

- Traitement des données en deep learning.
- Utilisé pour les modèles de réseaux de neurones convolutifs.
- Compression et factorisation des tenseurs pour réduire la complexité des modèles.

Définition : La matrice de covariance mesure la relation entre deux variables aléatoires :

$$\mathrm{Cov}(X, Y) = \mathbb{E}[(X - \mu_X)(Y - \mu_Y)]$$

Applications

- Analyse des relations entre variables en statistiques.
- Réduction de dimension et filtrage des données.
- Optimisation des portefeuilles financiers et gestion des risques.

6. Transformation Linéaire

Définition : Une transformation linéaire est une application qui conserve l'addition et la multiplication par un scalaire.

$$T(\mathbf{v}) = A\mathbf{v}$$

Applications

- Manipulation des espaces de représentation en NLP.
- Analyse des transformations des données en vision par ordinateur.
- Propagation des activations dans les réseaux neuronaux profonds.

Exemple : Norme Euclidienne

$$\|\mathbf{x}\|_2 = \sqrt{\sum_i x_i^2}$$

Applications

- Mesures de similarité entre données en machine learning et NLP.
- Utilisé pour l'optimisation et la régularisation des modèles.
- Convergence des algorithmes d'optimisation en deep learning.

Définition : Un espace vectoriel est un ensemble de vecteurs qui peuvent être additionnés et multipliés par un scalaire.

Applications

- Fondamental pour les embeddings de mots en NLP.
- Utilisé pour la manipulation des données en IA et en apprentissage automatique.
- Employé en réduction de dimensionnalité en data science.

Définition : Le rang d'une matrice A est le nombre de colonnes linéairement indépendantes.

Applications

- Vérification de redondance dans les données.
- Analyse de la complexité et de la stabilité des modèles IA.
- Détection des dimensions utiles pour la réduction de données en data science.

Formules :

$$A = LU, \quad A = QR$$

Applications

- Optimisation numérique et résolution efficace de systèmes linéaires.
- Utilisé pour l'implémentation efficace des modèles en machine learning.
- Appliqué en analyse de données pour la factorisation de matrices.

10 Notions de Probabilités et Statistiques

11. Théorème de Bayes

Formule :

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

où $P(A|B)$ est la probabilité de A sachant B et $P(A)$, $P(B)$ sont les probabilités a priori.

Applications

- Modèles probabilistes et inférence en NLP.
- Détection de spams et classification bayésienne.
- Systèmes de recommandation et filtres collaboratifs.

Formule :

$$H(X) = -\sum_i p(x_i) \log p(x_i)$$

où $H(X)$ mesure l'incertitude d'une variable aléatoire et $p(x_i)$ est la probabilité d'occurrence de x_i.

Applications

- Compression de texte et codage de Huffman.
- Modélisation des incertitudes en IA.
- Sélection de variables en machine learning.

Formules :

$$\mathrm{Var}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2], \quad \sigma_X = \sqrt{\mathrm{Var}(X)}$$

où $\mathrm{Var}(X)$ mesure la dispersion des valeurs de X, $\mathbb{E}[X]$ est l'espérance et σ_X est l'écart-type.

Applications

- Normalisation des données en machine learning.
- Détection des anomalies dans les données.
- Analyse du bruit dans les modèles statistiques.

14. Loi Normale

Formule :

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

où μ est la moyenne et σ l'écart-type, définissant une distribution de données centrée autour de μ.

Applications

- Hypothèses sur la distribution des erreurs en régression.
- Génération de données synthétiques.
- Modélisation des phénomènes naturels et financiers.

Exemple : Test de Student

$$t = \frac{\bar{X} - \mu}{s/\sqrt{n}}$$

où $\bar{X}$ est la moyenne de l'échantillon, μ la moyenne théorique, s l'écart-type et n la taille de l'échantillon.

Applications

- Validation des modèles en machine learning.
- Comparaison des performances d'algorithmes.
- Analyse A/B en marketing et expérimentation.

16. Espérance Mathématique

Formule :

$$\mathbb{E}[X] = \sum_i x_i P(X = x_i)$$

où $\mathbb{E}[X]$ est la valeur moyenne attendue d'une variable aléatoire discrète.

Applications

- Prédiction des sorties des modèles IA.
- Stratégies optimales en apprentissage par renforcement.
- Évaluation des risques en finance et assurance.

Formule :

$$D_{KL}(P||Q) = \sum_i P(x_i) \log \frac{P(x_i)}{Q(x_i)}$$

où $P(x_i)$ est la distribution réelle et $Q(x_i)$ une distribution approximée.

Applications

- Comparaison des distributions en apprentissage automatique.
- Optimisation des modèles génératifs (GANs, VAE).
- Sélection de modèles statistiques et inférence bayésienne.

Propriété : Processus où l'état futur dépend uniquement de l'état présent.

$$P(X_{t+1}|X_t, X_{t-1}, \ldots) = P(X_{t+1}|X_t)$$

Applications

- Modélisation du langage en NLP (HMM).
- Systèmes de reconnaissance vocale et traitement du signal.
- Modélisation des séries temporelles en finance.

Formule :

$$P(N(t) = k) = \frac{(\lambda t)^k e^{-\lambda t}}{k!}$$

où λ est le taux moyen d'occurrence et t la période d'observation.

Applications

- Modélisation d'événements rares (pannes, appels clients).
- Analyse du trafic réseau et gestion des files d'attente.
- Détection d'anomalies en cybersécurité et finance.

20. Bootstrap et Estimation Bayésienne

Principe : Méthode d'échantillonnage avec remise pour estimer la distribution d'une statistique.

Applications

- Estimation des intervalles de confiance en ML.
- Validation de la robustesse des modèles.
- Approche bayésienne pour la prédiction et la prise de décision.

10 Notions d'Optimisation

Formule :

$$\theta_{t+1} = \theta_t - \eta \nabla J(\theta_t)$$

où θ_t représente les paramètres du modèle à l'itération t, η est le taux d'apprentissage et $\nabla J(\theta_t)$ est le gradient de la fonction de coût.

Applications

- Apprentissage des réseaux neuronaux.
- Optimisation des paramètres en machine learning.
- Ajustement des coefficients en régression.

22. Descente de Gradient Stochastique (SGD)

Formule :

$$\theta_{t+1} = \theta_t - \eta \nabla J_i(\theta_t)$$

où $\nabla J_i(\theta_t)$ est le gradient calculé sur un échantillon unique de données.

Applications

- Optimisation des modèles en mini-batchs.
- Réduction du coût de calcul en deep learning.
- Accélération de l'apprentissage des grands modèles.

Formules :

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2$$

où m_t et v_t sont les moyennes mobiles des gradients et β_1, β_2 sont des coefficients d'atténuation.

Applications

- Accélération de la convergence en deep learning.
- Optimisation des modèles complexes.
- Stabilisation de l'entraînement des réseaux neuronaux profonds.

Formule :

$$x_{t+1} = x_t - H^{-1}\nabla f(x_t)$$

où H est la matrice hessienne de la fonction $f(x)$.

Applications

- Optimisation des fonctions convexes.
- Calcul des points critiques en optimisation.
- Convergence rapide pour les fonctions bien conditionnées.

25. Méthode de Lagrange

Formule :

$$\mathcal{L}(x, \lambda) = f(x) + \lambda g(x)$$

où λ est un multiplicateur de Lagrange, utilisé pour prendre en compte la contrainte $g(x) = 0$.

Applications

- Régularisation des réseaux de neurones.
- Optimisation sous contrainte en IA.
- Calcul des optima en économie et en physique.

Définition : Ensemble de variables liées par des contraintes formalisées mathématiquement.

Applications

- Résolution de problèmes d'optimisation sous contrainte.
- Planification et allocation de ressources.
- Intelligence artificielle et systèmes experts.

27. Relaxation Convexe

Principe : Transformation d'un problème d'optimisation combinatoire en un problème convexe plus facile à résoudre.

Applications

- Formulation de problèmes complexes en optimisation.
- Approximations en machine learning.
- Simplification des problèmes NP-difficiles.

Formule Générale :

$$\max_{x} \quad c^T x \quad \text{sous contraintes} \quad Ax \leq b$$

où c est le vecteur de coefficients, x les variables à optimiser, et A, b définissent les contraintes.

Applications

- Allocation des ressources en apprentissage.
- Optimisation des chaînes logistiques et industrielles.
- Planification des tâches et gestion de production.

29. Régularisation L1 et L2

Formules :

$$\text{L1: } \lambda \sum |w_i|, \quad \text{L2: } \lambda \sum w_i^2$$

où λ est un coefficient de régularisation et w_i les poids du modèle.

Applications

- Réduction du surajustement (overfitting).
- Sélection de variables en machine learning.
- Amélioration de la généralisation des modèles.

Exemple : Problème du Voyageur de Commerce (TSP)

$$\min \sum d(i,j) x_{ij}$$

où $x_{ij} = 1$ si la ville i est visitée avant la ville j, et $d(i,j)$ est la distance entre elles.

Applications

- Recherche de chemins optimaux en graphes.
- Planification de tournées et logistique.
- Séquençage génétique et réseaux de transport.

10 Notions de Graphe

31. Algorithme de Dijkstra

Principe : Trouver le chemin le plus court entre un nœud source et tous les autres nœuds d'un graphe pondéré.

Formule de mise à jour :

$$d(v) = \min(d(v), d(u) + w(u, v))$$

où $d(v)$ est la distance minimale au sommet v, et $w(u, v)$ est le poids de l'arête entre u et v.

Applications

- Optimisation des trajets en logistique et transport.
- Recherche d'itinéraires dans les graphes de connaissances en IA.
- Planification des parcours en robotique et jeux vidéo.

Formule de mise à jour :

$$PR(v) = (1 - d) + d \sum_{u \in L(v)} \frac{PR(u)}{C(u)}$$

où $PR(v)$ est le score de PageRank du nœud v, d est un facteur de téléportation, $L(v)$ l'ensemble des pages pointant vers v, et $C(u)$ le nombre de liens sortants de u.

Applications

- Algorithme clé de Google Search.
- Modélisation des graphes neuronaux (GNN) en deep learning.
- Recommandation de contenu en data science.

Formule de centralité de betweenness :

$$C_B(v) = \sum_{s \neq v \neq t} \frac{\sigma(s, t|v)}{\sigma(s, t)}$$

où $\sigma(s, t)$ est le nombre total de chemins entre s et t, et $\sigma(s, t|v)$ est le nombre de ces chemins passant par v.

Applications

- Analyse des réseaux sociaux (ex: influenceurs sur Twitter).
- Détection des points critiques dans les infrastructures réseaux.
- Études des interactions entre neurones dans les modèles de deep learning.

34. Graphes Orientés

Définition : Graphe $G = (V, E)$ avec un ensemble de sommets V et un ensemble d'arêtes E telles que chaque arête $(u, v) \in E$ est dirigée de u vers v.

Applications

- Modélisation des relations causales en IA.
- Représentation des transitions d'état en apprentissage par renforcement.
- Organisation des dépendances en programmation et bases de données.

Algorithme récursif :

$$DFS(v) = \text{Marquer } v; \quad \forall w \in \text{Adj}(v), .$$

$$\text{si } w \text{ non marqué, alors } DFS(w)$$

où $\text{Adj}(v)$ est l'ensemble des voisins du sommet v.

Applications

- Exploration des graphes de connaissances en LLM et NLP.
- Analyse des relations entre entités en data science.
- Résolution de labyrinthes et de jeux vidéo.

Algorithme : Exploration couche par couche d'un graphe. **Complexité :** $O(V + E)$ avec V sommets et E arêtes.

Applications

- Optimisation des parcours dans les réseaux (ex: calcul du plus court chemin sans poids).
- Détection de communautés dans les réseaux sociaux.
- Recherche de correspondances dans les graphes d'entités en IA.

37. Coloration de Graphes

Définition : Trouver une fonction $c : V \to \{1, ..., k\}$ telle que $c(u) \neq c(v)$ pour toute arête (u, v). **Formule d'optimisation :**

$$\min k \quad \text{sous contrainte } c(u) \neq c(v) \text{ pour } (u, v) \in E.$$

Applications

- Allocation des fréquences radio en télécommunications.
- Optimisation des emplois du temps et planifications.
- Segmentation d'images et classification en IA.

Formule du Théorème de Ford-Fulkerson :

$$\max \sum_{u} f(s,u) = \sum_{v} f(v,t)$$

où $f(s,u)$ est le flot partant de la source s vers u, et $f(v,t)$ est le flot arrivant au puits t.

Applications

- Optimisation des flux d'information dans les réseaux.
- Planification des chaînes d'approvisionnement.
- Distribution optimale des charges dans les serveurs cloud.

Définitions :

- **Un graphe eulérien** contient un cycle passant par chaque arête exactement une fois.
- **Un graphe hamiltonien** contient un cycle passant par chaque sommet une fois.

Critère d'Euler : Un graphe est eulérien s'il contient 0 ou 2 sommets de degré impair.

Applications

- Planification des trajets en logistique et transport.
- Optimisation des circuits dans les microprocesseurs.
- Séquençage de l'ADN en bio-informatique.

Exemple : Réseau Bayésien

$$P(X_1, \ldots, X_n) = \prod_i P(X_i | \text{Parents}(X_i))$$

où chaque variable X_i dépend d'un sous-ensemble de variables appelées parents.

Applications

- Modélisation des dépendances en IA et apprentissage probabiliste.
- Détection des anomalies en data science.
- Traitement des données incertaines en finance et médecine.

10 Notions de la Théorie de l'Information et Compression

41. Entropie et divergence de Kullback-Leibler (KL)

Définition : Mesure de l'information et de la distance entre distributions de probabilité.

$$D_{KL}(P\|Q) = \sum_x P(x) \log \frac{P(x)}{Q(x)}$$

Applications

- Modélisation du langage naturel et NLP.
- Optimisation des modèles génératifs (GANs, VAEs).
- Sélection de modèles statistiques.

Définition : Algorithme de compression sans perte basé sur des arbres binaires.

$$L = \sum_i p_i \cdot l_i$$

Applications

- Compression de texte et d'images (JPEG, PNG).
- Réduction de taille des modèles NLP.
- Codage de données en télécommunications.

43. Transformée de Fourier discrète (DFT) et FFT

Définition : Transformation d'un signal du domaine temporel au domaine fréquentiel.

$$X_k = \sum_{n=0}^{N-1} x_n e^{-i2\pi kn/N}$$

Applications

- Analyse des signaux audio et traitement du son.
- Compression d'images et de vidéos.
- Détection de motifs en séries temporelles.

Définition : Décomposition des signaux en fonctions élémentaires localisées en temps et en fréquence.

$$W_{j,k} = \sum_{n} x_n \psi_{j,k}(n)$$

Applications

- Compression de données audio et d'images (JPEG 2000).
- Analyse multi-échelle en vision par ordinateur.
- Détection des anomalies dans les séries temporelles.

45. Théorie du signal et filtrage

Définition : Traitement des signaux pour améliorer leur qualité ou extraire des informations utiles.

$$y[n] = \sum_{k=0}^{M} b_k x[n-k] - \sum_{j=1}^{N} a_j y[n-j]$$

Applications

- Reconnaissance vocale et NLP.
- Suppression du bruit dans les signaux audio.
- Analyse des données EEG et ECG en médecine.

Définition : Réseaux neuronaux pour la compression et l'apprentissage des représentations latentes.

$$z = \mu + \sigma \odot \epsilon$$

Applications

- Réduction de dimension en machine learning.
- Génération de nouvelles images et textes.
- Dénoising autoencoders pour le traitement du signal.

47. Théorème de l'information de Shannon

Définition : Décrit la capacité maximale d'un canal de communication.

$$C = B \log_2(1 + S/N)$$

Applications

- Optimisation des algorithmes de compression.
- Codage des données en télécommunications.
- Analyse des réseaux neuronaux et information theory.

Définition : Définit la limite fondamentale de la compression avec perte.

$$R(D) = \min_{p(\hat{X}|X):E[d(X,\hat{X})]\leq D} I(X;\hat{X})$$

$$D(R) = \min_{p(\hat{X}|X):I(X;\hat{X})\leq R} E[d(X,\hat{X})]$$

Applications

- Compression vidéo et audio (MP3, MPEG).
- Réduction de modèles pour le deep learning.
- Transmission efficace des données.

Définition : Représentation des données avec un nombre minimal de coefficients non nuls.

$$x = D\alpha, \quad \text{avec} \quad \|\alpha\|_0 \ll N$$

$$\min_{\alpha} \|\alpha\|_0 \quad \text{sous} \quad \|x - D\alpha\|_2 \leq \epsilon$$

Applications

- Compression et stockage efficace des modèles IA.
- Reconstruction d'images et de signaux.
- Détection d'anomalies en machine learning.

50. Analyse des séries temporelles avec Fourier et ondelettes

Définition : Identification des tendances cachées dans les séries temporelles.

$$W_{j,k} = \sum_n x_n \psi_{j,k}(n)$$

$$S_X(f) = \lim_{T \to \infty} \frac{1}{T} E\left[|X_T(f)|^2\right]$$

Applications

- Prédiction des marchés financiers.
- Détection des tendances dans les signaux biologiques.
- Surveillance des infrastructures critiques.

10 Notions de Géométrie et Topolog

51. Espaces métriques et distances

Définition : Mesure de la similarité entre données.

$$d(x, y) = \|x - y\|$$

$$d(x, y) = \sqrt{\sum_i (x_i - y_i)^2}$$

Applications

- Clustering et classification en apprentissage automatique.
- Recherche d'informations en NLP.
- Détection d'anomalies en Data Science.

52. Analyse de persistance topologique

Définition : Détection des structures cachées dans les données.

$$H_k(X) = \frac{Z_k(X)}{B_k(X)}$$

$$P_k(X) = \{(b_i, d_i)\}$$

Applications

- Analyse de réseaux neuronaux.
- Reconnaissance de formes en vision par ordinateur.
- Étude des structures complexes en biologie computationnelle.

53. Graphe de Voronoï et diagrammes de Delaunay

Définition : Partition d'un espace en régions influencées par un ensemble de points.

$$V(p_i) = \{x \in \mathbb{R}^n \mid d(x, p_i) < d(x, p_j), j \neq i\}$$

$$D(p) = \bigcup_i \text{convex hull}(p_i)$$

Applications

- Clustering et réseaux neuronaux.
- Optimisation en graphes et planification.
- Analyse des données spatiales.

Définition : Représentation des données dans des espaces courbés.

$$T_pM = \{v \in \mathbb{R}^n \mid \text{tangent à } M \text{ en } p\}$$

$$g_{ij} = \left\langle \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \right\rangle$$

Applications

- Réduction de dimension en apprentissage automatique.
- Modélisation des données en NLP et vision.
- Analyse topologique des espaces de représentations.

Définition : Techniques de réduction de dimension et de visualisation.

$$X' = WX$$

$$D_{ij} = \|x_i - x_j\|^2$$

Applications

- Exploration des données en haute dimension.
- Amélioration des performances des algorithmes de ML.
- Visualisation des relations entre variables.

Définition : Utilisé pour la modélisation en GANs et modèles génératifs.

$$d_g(x, y) = \inf \int_0^1 \|\gamma'(t)\| dt$$

$$x' = G(z)$$

Applications

- Génération de données synthétiques.
- Apprentissage des représentations en NLP.
- Clustering et interpolation en espace latent.

57. Topologie de réseaux neuronaux

Définition : Étude de la structure et de l'optimisation des architectures de deep learning.

$$\mathcal{L} = \sum_i y_i \log \hat{y}_i$$

$$A' = f(WA + b)$$

Applications

- Conception de nouvelles architectures neuronales.
- Optimisation de la connectivité des réseaux.
- Analyse des performances des modèles IA.

58. Applications de la géométrie algébrique

Définition : Modélisation des formes complexes en IA à l'aide d'équations polynomiales.

$$f(x, y) = 0$$

$$J(f) = \det\left(\frac{\partial f_i}{\partial x_j}\right)$$

Applications

- Représentation des surfaces et des formes.
- Modélisation d'objets 3D en vision par ordinateur.
- Optimisation des modèles IA en robotique.

59. Optimisation sur variétés riemanniennes

Définition : Apprentissage dans des espaces non euclidiens.

$$\nabla_M f(x) = P_{T_x M} \nabla f(x)$$

$$x_{k+1} = R_{x_k}(\alpha d_k)$$

Applications

- Optimisation des modèles de deep learning.
- Amélioration des algorithmes de descente de gradient.
- Apprentissage des représentations géométriques en IA.

60. Courbure et géométrie différentielle

Définition : Étude de la courbure et des structures différentielles des espaces de représentation.

$$K = \frac{R_{1212}}{g_{11}g_{22} - g_{12}^2}$$

$$\nabla_Y X = \sum_i \Gamma_{ij}^k X^j Y^i \frac{\partial}{\partial x^k}$$

Applications

- Analyse des structures de données non euclidiennes.
- Modélisation des espaces latents en machine learning.
- Optimisation des algorithmes géométriques en vision par ordinateur.

10 Notions de Processus Stochastiq

Définition : Un processus stochastique est une famille de variables aléatoires indexées par un paramètre (temps, espace, etc.).

$$X_t : \Omega \to \mathbb{R}$$

Applications

- Modélisation des séries temporelles en Data Science.
- Prédiction des prix en finance.
- Suivi des phénomènes physiques et biologiques.

Définition : Un processus de Wiener est un processus stochastique continu à accroissements indépendants et stationnaires.

$$W_t \sim \mathcal{N}(0, t)$$

Applications

- Modélisation du mouvement brownien.
- Calcul stochastique et équations différentielles stochastiques.
- Prédiction des marchés financiers.

63. Processus de Markov

Propriété : Un processus de Markov est un processus où l'état futur dépend uniquement de l'état présent.

$$P(X_{t+1}|X_t, X_{t-1}, \ldots) = P(X_{t+1}|X_t)$$

Applications

- Chaînes de Markov en NLP et modèles génératifs.
- Modélisation des transitions d'état en apprentissage par renforcement.
- Analyse des files d'attente et gestion des ressources.

Formule :

$$P(N(t) = k) = \frac{(\lambda t)^k e^{-\lambda t}}{k!}$$

où λ est le taux moyen d'occurrence et t la période d'observation.

Applications

- Modélisation d'événements rares (appels clients, défaillances).
- Analyse du trafic en cybersécurité.
- Gestion des flux dans les systèmes en réseau.

65. Processus de Branching

Définition : Un processus de branching décrit l'évolution d'une population où chaque individu engendre une descendance aléatoire.

$$Z_{n+1} = \sum_{i=1}^{Z_n} X_i$$

Applications

- Modélisation des infections épidémiologiques.
- Croissance des populations et systèmes biologiques.
- Simulation de la propagation des informations.

66. Processus de Martingale

Propriété : Un processus X_t est une martingale si :

$$\mathbb{E}[X_{t+1}|X_t, X_{t-1}, ...] = X_t$$

Applications

- Théorie des jeux et stratégies optimales.
- Modélisation des marchés financiers.
- Algorithmes d'optimisation en apprentissage machine.

Équation de Langevin :

$$dX_t = \mu(X_t, t)dt + \sigma(X_t, t)dW_t$$

Applications

- Modélisation du mouvement des particules en physique.
- Simulation des systèmes financiers et biologiques.
- Détection des anomalies dans les séries temporelles.

68. Processus Stationnaire

Définition : Un processus est stationnaire si sa distribution ne dépend pas du temps.

$$\mathbb{E}[X_t] = \mathbb{E}[X_{t+\tau}], \quad \mathrm{Var}(X_t) = \mathrm{Var}(X_{t+\tau})$$

Applications

- Analyse des signaux en traitement du signal.
- Modélisation des séries temporelles en finance.
- Étude des réseaux neuronaux récurrents en deep learning.

69. Processus Gaussien

Définition : Un processus gaussien est un ensemble de variables aléatoires telles que toute combinaison linéaire suit une distribution gaussienne.

$$X_t \sim \mathcal{N}(\mu_t, K(t, t'))$$

Applications

- Méthodes d'apprentissage non paramétriques.
- Optimisation bayésienne en intelligence artificielle.
- Interpolation et modélisation en statistiques.

Formule :

$$X_t = X_{t-1} + \epsilon_t, \quad \epsilon_t \sim \mathcal{N}(0, \sigma^2)$$

Applications

- Modélisation des trajectoires boursières.
- Génération de bruit en vision par ordinateur.
- Étude des déplacements en biologie et physique.

10 Notions d'Analyse

71. Théorème des Valeurs Intermédiaires

Définition : Si une fonction continue f sur un intervalle $[a, b]$ prend deux valeurs $f(a)$ et $f(b)$, alors elle prend toute valeur intermédiaire :

$$\forall y \in [f(a), f(b)], \quad \exists c \in (a, b) \text{ tel que } f(c) = y.$$

Applications

- Convergence et existence des solutions en apprentissage automatique.
- Recherche des points fixes en IA.
- Interpolation et approximation de données.

72. Théorème de Point Fixe de Banach

Définition : Si f est une contraction sur un espace métrique complet, alors il existe un unique point fixe x^* tel que :

$$f(x^*) = x^*.$$

Applications

- Algorithmes itératifs en optimisation et deep learning.
- Convergence des méthodes numériques en IA.
- Résolution des équations fonctionnelles en machine learning.

Définition : L'analyse asymptotique étudie le comportement des fonctions lorsque $x \to \infty$. On note :

$$f(x) = O(g(x)) \text{ si } \exists C, x_0 \text{ tq } \forall x > x_0, |f(x)| \leq C|g(x)|.$$

Applications

- Analyse de la complexité algorithmique en IA.
- Étude des performances des modèles de deep learning.
- Optimisation des algorithmes d'entraînement.

74. Théorème d'Approximation de Weierstrass

Définition : Toute fonction continue sur un intervalle $[a, b]$ peut être uniformément approchée par des polynômes :

$$\forall \epsilon > 0, \exists P_n(x) \text{ tel que } \sup_{x \in [a,b]} |f(x) - P_n(x)| < \epsilon.$$

Applications

- Approximation des réseaux neuronaux.
- Compression des modèles d'IA.
- Interpolation et régularisation des données.

75. Espaces de Sobolev

Définition : Un espace de Sobolev $W^{k,p}(\Omega)$ généralise les notions de différentiabilité et d'intégrabilité en exigeant que les dérivées faibles soient dans L^p.

$$W^{k,p}(\Omega) = \{f \in L^p(\Omega) \mid D^\alpha f \in L^p(\Omega), |\alpha| \leq k\}.$$

Applications

- Analyse des régularités dans les modèles d'IA.
- Apprentissage des solutions des équations aux dérivées partielles.
- Traitement d'image et filtrage en deep learning.

76. Théorème de la Convergence Dominée

Définition : Si une suite de fonctions f_n est dominée par une fonction intégrable g, alors :

$$\lim_{n \to \infty} \int f_n = \int \lim_{n \to \infty} f_n.$$

Applications

- Convergence des méthodes de Monte Carlo en IA.
- Validation de la convergence des algorithmes probabilistes.
- Théorèmes limites en deep learning.

77. Théorème de Lagrange et Développement de Taylor

Définition : Le théorème de Lagrange garantit qu'une fonction dérivable possède un point c où :

$$f(b) - f(a) = f'(c)(b - a).$$

Le développement de Taylor permet d'approximer $f(x)$ par un polynôme :

$$f(x) \approx f(a) + f'(a)(x - a) + \frac{f''(a)}{2!}(x - a)^2 + \dots$$

Applications

- Approximation des modèles d'IA.
- Analyse de la stabilité des algorithmes.
- Calcul de gradients et descentes de gradient.

78. Théorème Spectral

Définition : Tout opérateur autoadjoint A dans un espace de Hilbert possède une décomposition en valeurs propres :

$$A = \sum \lambda_i P_i.$$

Applications

- Analyse des matrices de covariance en apprentissage automatique.
- Réduction de dimension via PCA.
- Étude des réseaux de neurones et des graphes.

Définition : L'analyse variationnelle étudie l'optimisation des fonctionnelles. Une solution optimale satisfait :

$$\frac{d}{d\epsilon} J(u + \epsilon v)\Big|_{\epsilon=0} = 0.$$

Applications

- Optimisation des modèles d'IA.
- Calcul de gradients dans l'apprentissage automatique.
- Formulation des modèles physiques et statistiques.

Définition : Un système dynamique est défini par :

$$\frac{dx}{dt} = f(x).$$

On parle de chaos si la sensibilité aux conditions initiales est exponentielle :

$$d(x_t, y_t) \approx e^{\lambda t} d(x_0, y_0).$$

Applications

- Modélisation des comportements imprévisibles en IA.
- Dynamique des apprentissages en deep learning.
- Analyse des trajectoires dans les algorithmes évolutionnaires.

10 Notions de Séries et Transformé

Définition : Toute fonction périodique peut être exprimée en une somme infinie de sinusoïdes.

$$f(x) = a_0 + \sum_{n=1}^{\infty} (a_n \cos(nx) + b_n \sin(nx))$$

Applications

- Analyse des signaux et compression.
- Modélisation des séries temporelles.
- Traitement du son et reconnaissance vocale.

82. Transformée de Fourier Discrète (DFT)

Définition : Conversion d'un signal temporel en signal fréquentiel.

$$X_k = \sum_{n=0}^{N-1} x_n e^{-i2\pi kn/N}$$

Applications

- Traitement des signaux en vision par ordinateur.
- Compression des images et sons (JPEG, MP3).
- Détection de motifs dans les séries temporelles.

83. Transformée de Fourier Rapide (FFT)

Définition : Algorithme optimisé pour calculer la DFT en $O(N \log N)$ au lieu de $O(N^2)$.

Applications

- Optimisation du traitement d'images.
- Analyse spectrale des signaux biologiques.
- Compression et filtrage en télécommunications.

84. Transformée en Ondelette (Wavelet)

Définition : Analyse fréquentielle avec localisation temporelle.

$$W_{j,k} = \sum_{n} x_n \psi_{j,k}(n)$$

Applications

- Compression d'images (JPEG 2000).
- Analyse multi-échelle des signaux.
- Détection d'anomalies dans les données temporelles.

Définition : Approximation locale d'une fonction par une somme infinie de termes polynomiaux.

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n$$

Applications

- Approximation des fonctions non linéaires en IA.
- Analyse des erreurs en calcul numérique.
- Optimisation en apprentissage automatique.

Définition : Conversion des équations différentielles en équations algébriques.

$$F(s) = \int_0^{\infty} e^{-st} f(t) dt$$

Applications

- Analyse des systèmes dynamiques.
- Modélisation des signaux et contrôle automatique.
- Résolution des équations différentielles en IA.

Définition : Extension discrète de la transformée de Laplace.

$$X(z) = \sum_{n=0}^{\infty} x_n z^{-n}$$

Applications

- Analyse des systèmes à temps discret.
- Conception de filtres en traitement du signal.
- Modélisation des processus en apprentissage automatique.

Définition : Multiplication d'un signal par une fonction de fenêtre pour éviter les effets de bord.

Applications

- Amélioration de la précision spectrale.
- Analyse des signaux en audio et électrophysiologie.
- Réduction des artefacts en traitement du signal.

Définition : La convolution d'un signal avec un noyau est utilisée pour le filtrage et la détection de motifs.

$$(f * g)(t) = \int_{-\infty}^{\infty} f(\tau)g(t-\tau)d\tau$$

Applications

- Réseaux de neurones convolutifs (CNN).
- Traitement d'images et reconnaissance faciale.
- Analyse des données en apprentissage automatique.

90. Fonction de Transfert en Traitement du Signal

Définition : Modélisation d'un système linéaire dans le domaine fréquentiel.

$$H(s) = \frac{Y(s)}{X(s)}$$

Applications

- Conception des filtres numériques.
- Modélisation des réseaux neuronaux récurrents.
- Optimisation des signaux en reconnaissance vocale.

10 Notions de Logique Mathématiq

Définition : Système logique basé sur des propositions qui peuvent être vraies ou fausses.

$$P \vee \neg P = \text{Vrai} \quad \text{(principe du tiers exclu)}$$

Applications

- Vérification de programmes informatiques.
- Modélisation des circuits logiques.
- Systèmes de raisonnement automatisé en IA.

Définition : Extension de la logique propositionnelle intégrant des variables et des quantificateurs.

$$\forall x \in D, P(x) \Rightarrow Q(x)$$

Applications

- Intelligence artificielle et systèmes experts.
- Représentation des connaissances en NLP.
- Vérification formelle des logiciels.

93. Théorème de Complétude de Gödel

Définition : Un système formel est complet si toute formule vraie peut être prouvée.

$$\vdash \phi \Rightarrow \text{si } \phi \text{ est vraie, elle est prouvable.}$$

Applications

- Fondements des mathématiques et IA symbolique.
- Preuves automatisées et vérification des théorèmes.
- Modélisation des bases de connaissances.

Définition : Tout système formel suffisamment puissant contient des énoncés indécidables.

$$\exists \phi, \quad \neg(\vdash \phi) \land \neg(\vdash \neg\phi)$$

Applications

- Limites des systèmes d'IA formelle.
- Théories des nombres et formalisation des mathématiques.
- Étude des langages formels et programmation logique.

Définition : Extension de la logique classique permettant d'exprimer des modalités comme la nécessité et la possibilité.

$$\Box P \Rightarrow P, \quad P \Rightarrow \Diamond P$$

Applications

- Raisonnement en intelligence artificielle.
- Vérification de protocoles et systèmes distribués.
- Modélisation de la connaissance et de la croyance.

Définition : Logique sans tiers exclu, utilisée en mathématiques constructives.

$$P \vee \neg P \quad \text{n'est pas toujours vrai.}$$

Applications

- Vérification formelle des programmes.
- Preuves assistées par ordinateur.
- Fondements de la théorie des types.

Définition : Système algébrique basé sur des opérations logiques.

$$A \land (B \lor C) = (A \land B) \lor (A \land C)$$

Applications

- Conception des circuits numériques.
- Optimisation des bases de données.
- Développement des moteurs de recherche.

Définition : Logique généralisant la logique booléenne pour gérer des degrés de vérité.

$$\mu_A(x) \in [0, 1]$$

Applications

- Systèmes de contrôle intelligents.
- Traitement du langage naturel.
- Prise de décision et IA hybride.

Définition : Modèle formel de calcul basé sur les fonctions anonymes et l'application de fonctions.

$$(\lambda x.f(x))(y) = f(y)$$

Applications

- Théorie des langages de programmation.
- Intelligence artificielle et raisonnement automatique.
- Conception des compilateurs et analyse statique.

Définition : Système logique où chaque terme appartient à un type bien défini.

$$x : A, \quad \text{signifie que } x \text{ est de type } A.$$

Applications

- Vérification de la sûreté des programmes.
- Développement des langages fonctionnels.
- Preuves formelles et assistants de preuve.

Envie d'en apprendre plus ?

Suivez **Clotilde Djuikem** sur LinkedIn et **Tioh Academy** sur YouTube !