**2. Partie I – Fondations mathématiques et pipelines de données**

La première partie jette les bases nécessaires à toute approche de deep learning. Elle se subdivise en deux volets : la théorie (algèbre linéaire, calcul différentiel et probabilités) et la pratique (gestion et transformation des données).

* **Algèbre linéaire et structures de données** : on y présente les notions d’espace vectoriel, de matrices et de tenseurs. On y explique l’importance des opérations de base (addition, multiplication, transposition) dans la manipulation efficace de grandes quantités de données, ainsi que la notion de projection pour la réduction de dimension.
* **Calcul différentiel appliqué** : cette section détaille comment les dérivées permettent de mesurer la sensibilité d’une fonction de coût par rapport à ses paramètres. On y introduit la règle de la chaîne, essentielle pour comprendre comment le gradient se propage dans un réseau de neurones.
* **Probabilités et fonctions de perte** : on établit le lien entre les distributions probabilistes et les critères d’apprentissage. La notion de vraisemblance et de cross-entropy y est abordée pour justifier les fonctions de coût utilisées en classification et en régression.
* **Préparation des données** : le livre propose des méthodes pour charger, nettoyer et transformer des ensembles variés (images, textes, séries temporelles). On découvre la création de pipelines efficaces, incluant le batching, le mélange aléatoire et les techniques d’augmentation pour enrichir artificiellement les données d’entraînement.