Rappels de mathématiques

Introduction à la Data Science - Efreitech

Maxime Jumelle - Taieb Badis

www.blent.ai

Plan

1. Calcul matriciel
2. Statistiques / Probabilités
3. Analyse

Calcul matriciel

Produit matriciel

$$\begin{pmatrix} 3 & 1 \\ -1 & 2 \end{pmatrix} \times \begin{pmatrix} -2 & 4 \\ 1 & 3 \end{pmatrix} = \begin{pmatrix} -5 & 15 \\ 4 & 2 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \times \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 7 \end{pmatrix}$$

Attention

Le produit matriciel n'est pas toujours commutatif ! En règle général, pour deux matrices A et B :

$$AB \neq BA$$

Opérations particulières

Inverse

$$M^{-1} = \frac{1}{\det(M)}\mathrm{Com}(M)^\top$$

Transposée

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}^\top = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}$$

Normes

Considérons $x = (x_1, \ldots, x_n)^\top \in \mathbb{R}^n$.

Norme L^1 :

$$\|x\|_1 = \sum_{i=1}^{n} |x_i|$$

Norme euclidienne (ou L^2) :

$$\|x\|_2 = \sqrt{\sum_{i=1}^{n} x_i^2}$$

Norme infinie (ou L^∞) :

$$\|x\|_\infty = \max_{1 \le i \le n} |x_i|$$

Exercice

Montrer que $\|x\|_2^2 = x^\top x$.

Statistiques / Probabilités

Notions de bases

L'élément de base est **la variable aléatoire** X sur un ensemble Ω appelé **univers**. Une variable aléatoire peut être à plusieurs types :

- **discrète finie** si elle peut prendre un nombre fini de valeurs (jeu de carte, lancer de dés, ...)
- **discrète infinie** est elle peut prendre une infinité de valeurs dénombrable (nombre de piles successifs au pile ou face)
- **continue** lorsque les valeurs possibles sont indénombrables (temps d'attente, distance, ...)

On note $X(\Omega)$ l'ensemble des valeurs que peut prendre une variable aléatoire.
Parmi les lois les plus connus, on retrouve la loi de Bernouilli, binomiale et normale.

Mesure de probabilité

Une mesure de probabilité $\mathbb{P}$ est une mesure à valeurs dans $[0, 1]$ telle que

- $\mathbb{P}(\emptyset) = 0$
- Pour des événements $A_1, A_2, ...$ deux à deux disjoints :

$$\mathbb{P}\left(\bigcup_{n\geq 1} A_n\right) = \sum_{n\geq 1} \mathbb{P}(A_n)$$

Espérance mathématique

Definition (Espérance mathématique)
L'espérance mathématique d'une variable X est définie par :

$$\text{(Discrète)} \quad \mathbb{E}(X) = \sum_{x \in X(\Omega)} x\mathbb{P}(X = x)$$

$$\text{(Continue)} \quad \mathbb{E}(X) = \int xf(x)dx$$

Attention : l'espérance mathématique n'est pas la moyenne, même si elle y ressemble !

Moyenne

Definition (Moyenne)
La moyenne des observations $x_1, ..., x_n$ est définie par :

$$\mu = \frac{1}{n}\sum_{i=1}^{n} x_i$$

[Loi des grands nombres] Si les observations $x_1, ..., x_n$ sont issues d'une variable aléatoire X alors

$$\frac{1}{n}\sum_{i=1}^{n} x_i \longrightarrow \mathbb{E}(X)$$

Variance

Definition (Variance)
La variance d'une variable X est définie par :

$$\text{(Discrète)} \quad \mathbb{V}(X) = \sum_{x \in X(\Omega)} (x - \mu)^2 \mathbb{P}(X = x)$$

$$\text{(Continue)} \quad \mathbb{V}(X) = \int (x - \mu)^2 f(x) dx$$

On peut montrer la formule de König-Huygens :

$$\mathbb{V}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$$

Une variance est toujours positive !

Variance empirique

Definition (Ecart-type)

L'écart-type d'une variable aléatoire est la racine carrée de la variance (si elle existe).

$$\sigma_X = \sqrt{\mathbb{V}(X)}$$

Definition (Variance empirique)

La variance empirique des observations $x_1, ..., x_n$ est définie par :

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

Covariance

Definition (Covariance)

La covariance est une mesure de concentration entre deux variables aléatoires X et Y :

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

Definition (Corrélation linéaire)

La corrélation est une mesure de lien linéaire entre deux variables aléatoires X et Y :

$$\rho(X, Y) = \frac{\mathrm{Cov}(X, Y)}{\sqrt{\mathbb{V}(X)\mathbb{V}(Y)}}$$

On a toujours $\rho(X, Y) \in [-1, 1]$. Attention : corrélation $\neq$ dépendence.

Conclusion

- Bien différencier le cadre probabiliste (espérance, variance, ...) du cadre statistique (moyenne, variance empirique, ...)
- En probabilités, on **prédit** les futures observations
- En statistiques, on **analyse** les observations dont on dispose

Analyse

Gradient

Considérons une fonction $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ différentiable.

Le gradient d'une fonction à valeurs réelles est le vecteur de ses dérivées partielles.

$$\nabla f(x_1, \ldots, x_n) = \begin{pmatrix} \dfrac{\partial f}{\partial x_1} \\ \vdots \\ \dfrac{\partial f}{\partial x_n} \end{pmatrix}$$