

**Université Lumière Lyon 2**

Institut de la communication

## **Librairie MIMOSA**

# **Mixed Input Multinomial Optimization for Statistical Analysis**

*Par*

Linh Nhi Le Dinh  
Antoine Oruezabala  
Béranger Thomas

Supervisé par :

Ricco Rakotomalala

*Rapport présenté dans le cadre du*

Master 2 SISE - 2024/2025

# Table des matières

<b>1</b>	<b>Éléments introductifs</b>	<b>2</b>
1.1	Contexte . . . . .	2
<b>2</b>	<b>Étude du problème</b>	<b>3</b>
<b>3</b>	<b>Librairie développée</b>	<b>4</b>
<b>4</b>	<b>Lexique</b>	<b>5</b>

# Chapitre 1

## Éléments introductifs

### 1.1 Contexte

Ce document présente la régression logistique multinomiale, par le biais d'une descente de gradient, pour des données mixtes.

## **Chapitre 2**

### **Étude du problème**

## Chapitre 3

# Librairie développée

La librairie développée porte le doux nom de MIMOSA (Mixed Input Multinomial Optimization for Statistical Analysis).

# Chapitre 4

## Lexique

**RÉGRESSION LOGISTIQUE** ■ La régression logistique vise à prédire la probabilité qu'un événement binaire se produise (oui/non, 1/0) en fonction de variables explicatives. Elle utilise une fonction logistique (ou sigmoïde) pour transformer la combinaison linéaire des variables indépendantes en une probabilité comprise entre 0 et 1.

**RÉGRESSION LOGISTIQUE MULTINOMIALE** ■ Extension de la régression logistique permettant de modéliser une variable qualitative à plus de deux modalités.

Cette méthode permet d'estimer la probabilité d'appartenance à chacune des catégories en fonction de variables explicatives quantitatives et/ou qualitatives. Contrairement à la régression logistique binaire qui modélise une seule probabilité  $p$ , la régression logistique multinomiale estime simultanément les probabilités d'appartenance à toutes les catégories  $(p_1, p_2, \dots, p_K)$ , avec la contrainte que leur somme soit égale à 1.

Le modèle utilise une transformation logistique généralisée qui garantit que les probabilités estimées restent comprises entre 0 et 1. Une catégorie est choisie comme référence, et le modèle estime les logarithmes des rapports de probabilités (log-odds) entre chaque catégorie et cette référence.

Limites : - Sensible à la multicollinéarité des variables explicatives - Nécessite un échantillon suffisamment grand, particulièrement quand le nombre de catégories augmente - Suppose l'indépendance des alternatives non pertinentes (IIA)

**DESCENTE DE GRADIENT** ■ La descente de gradient est un algorithme d'optimisation couramment utilisé pour entraîner les modèles de machine learning et les réseaux neuronaux. Ce type d'algorithme entraîne les modèles de machine learning par réduction des erreurs entre les résultats prédits et les résultats réels. ■ Le point de départ n'est qu'un point arbitraire qui nous permet d'évaluer les performances. À partir de ce point de départ, nous allons trouver la dérivée (ou la pente) et, à partir de là, nous pourrons utiliser une ligne tangente pour observer l'inclinaison de la pente. La pente renseigne sur les mises à jour des paramètres, c'est-à-dire les poids et les biais. La pente au point de départ est plus forte, mais au fur et à mesure que de

nouveaux paramètres sont générés, elle devrait progressivement diminuer jusqu'à atteindre le point le plus bas de la courbe, dénommé point de convergence. ■ Comme pour trouver la ligne de meilleur ajustement dans la régression linéaire, l'objectif de la descente de gradient est de minimiser la fonction de coût, ou l'erreur entre  $y$  prédit et  $y$  réel. Pour ce faire, deux points de données sont nécessaires : une orientation et un taux d'apprentissage. Ces facteurs déterminent les calculs de dérivée partielle des itérations futures, ce qui lui permet d'atteindre progressivement le minimum local ou global (c'est-à-dire le point de convergence).. La fonction de perte dans la descente de gradient agit spécifiquement comme un baromètre, évaluant sa précision à chaque itération des mises à jour de paramètres. Jusqu'à ce que la fonction soit proche ou égale à zéro, le modèle continue à ajuster ses paramètres pour obtenir l'erreur la plus faible possible. ■ Il existe trois types d'algorithmes d'apprentissage par descente de gradient : la descente de gradient par lots, la descente de gradient stochastique et la descente de gradient par mini-lots.