RAMAROVAHOAKA Jean Angelo
Introduction à l'Apprentissage Automatique
Mathématiques Informatique et Statistique appliquées
Mention Informatiques et Technologie

Ajustement Linéaire : Méthodes des Moindres Carrés et Descente de Gradient

L'ajustement linéaire est une technique utilisée pour modéliser la relation entre une variable indépendante x et une variable dépendante y à l'aide d'une droite de la forme $y = ax + b$. Voici comment on peut déterminer les coefficients a et b à l'aide des deux méthodes principales : la méthode des moindres carrés et la descente de gradient.

1. Méthode des Moindres Carrés

Principe La méthode des moindres carrés vise à minimiser l'erreur quadratique totale entre les valeurs observées y_i et les valeurs prédites $\hat{y}_i$. Cela se traduit par la minimisation de la somme des carrés des différences verticales entre les points de données et la droite de régression.

Formulation Mathématique Soit un ensemble de données $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$. La fonction de coût $J(a, b)$ est définie par :

$$J(a,b) = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = \sum_{i=1}^{n} (y_i - (ax_i + b))^2$$

Pour minimiser cette fonction, nous dérivons J par rapport à a et b et égalons ces dérivées à zéro :

$$\frac{\partial J}{\partial a} = -2 \sum_{i=1}^{n} x_i (y_i - (ax_i + b)) = 0$$

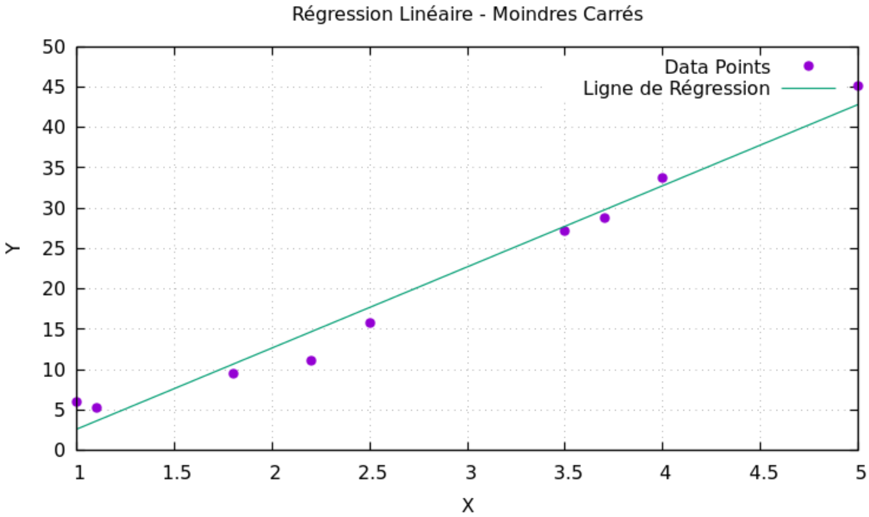
$$\frac{\partial J}{\partial b} = -2 \sum_{i=1}^{n} (y_i - (ax_i + b)) = 0$$

En résolvant ces équations simultanément, nous obtenons :

$$a = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{n \sum x_i^2 - (\sum x_i)^2}$$

$$b = \frac{\sum y_i - a \sum x_i}{n}$$

Illustration de la méthode moindre carrée



Avantages - Offre une solution analytique exacte. - Facile à implémenter pour des problèmes linéaires simples.

2. Méthode de Descente de Gradient

Principe La descente de gradient est une méthode itérative qui optimise la fonction de coût en ajustant les coefficients a et b dans la direction des gradients négatifs.

Algorithme 1. **Initialisation** : Choisissez des valeurs initiales pour a et b (souvent zéro). 2. **Calcul des Prédictions** : Pour chaque point de données (x_i, y_i), calculez la prédiction :

$$\hat{y}_i = ax_i + b$$

3. **Calcul des Gradients** :

$$\frac{\partial J}{\partial a} = -\frac{2}{n}\sum_{i=1}^{n} x_i(y_i - \hat{y}_i)$$

$$\frac{\partial J}{\partial b} = -\frac{2}{n}\sum_{i=1}^{n} (y_i - \hat{y}_i)$$

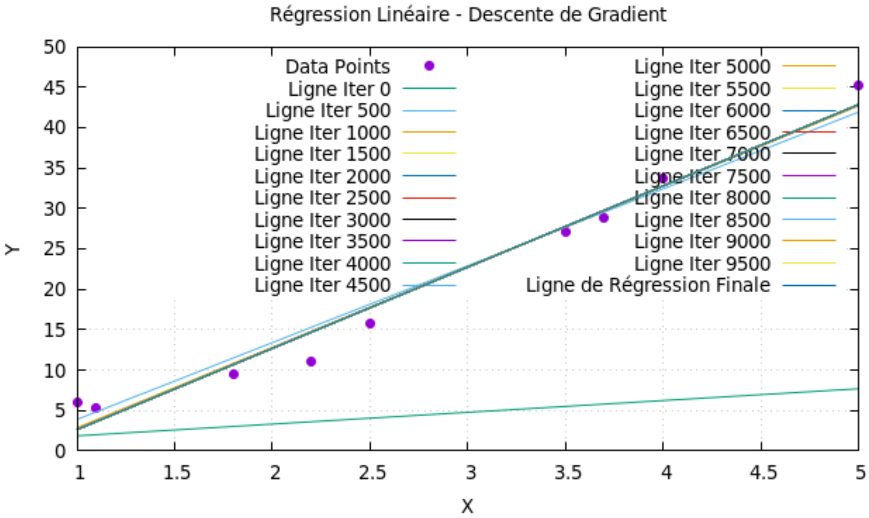
4. **Mise à Jour des Coefficients** :

$$a = a - \alpha\frac{\partial J}{\partial a}$$

$$b = b - \alpha\frac{\partial J}{\partial b}$$

5. **Répétez** jusqu'à convergence.

Illustration de la méthode déscente de gradient



Avantages - Appliqué aux problèmes non linéaires et à grande échelle. - Fonctionne lorsque les dérivées analytiques sont difficiles à calculer.

Conclusion

- **Méthode des Moindres Carrés** : Fournit une solution analytique pour les modèles linéaires simples. - **Descente de Gradient** : Approche itérative flexible pour modèles complexes.

Les deux méthodes visent à minimiser l'erreur entre les valeurs observées et prédites, mais diffèrent dans leur approche et leur applicabilité.