



Université d'Abomey-Calavi
Institut de Formation et de Recherche Informatique(IFRI)

Exposé sur l'Algorithme de Gradient Descent

Nom: ZINZINDOHOUE

Prénoms: Babatoundé César Boutoile

Filière: Système d'information et Réseau Informatique

06 Mars 2023

PLAN

I. Introduction

II. Historique

III. Présentation de la méthodologie adoptée

**IV. Domaine d'utilisation de l'agorithme du Gradient
dessent**

V. Comparaisons avec les algorithmes similaires

VI. Avantages et limites

VII. Conclusion

Introduction

Le domaine de la Data Science est un domaine en plein essor porteur de grandes opportunités et qui a donné naissance à de nombreuses disciplines. L'une d'entre elle est le Machine Learning qui est lié à la technologie de l'I.A. . Cette discipline consiste pour le data scientist à permettre aux ordinateurs d'apprendre d'eux-mêmes et d'améliorer leurs performances. Pour ce faire, le machine learning fait appel à des notions mathématiques assez complexes, dont l'algorithme de gradient descent. Dans la suite de notre développement, nous allons d'une part nous intéresser aux généralités liées à l'algorithme de gradient descent, son domaine d'expertise, nous parlerons d'autre part des avantages ainsi que les limites liés à cet algorithme.

A. Historique

Le Gradient Descent est une méthode d'optimisation qui a été développée dans les années 1800 par les mathématiciens Augustin-Louis Cauchy et Leonhard Euler. Cependant, la version moderne du Gradient Descent que nous connaissons aujourd'hui a été introduite dans les années 1950 par le mathématicien Richard Bellman et le statisticien Leonard Savage. Cette notion a été initialement utilisée pour résoudre les problèmes d'optimisation dans les domaines de la physique et de l'ingénierie.

Cependant, dans les années 1960, il a commencé à être appliqué en apprentissage automatique et en intelligence artificielle, où il est devenu une méthode essentielle pour l'optimisation des modèles en Machine Learning.

Au fil des décennies, différentes variantes du Gradient Descent ont été développées pour optimiser son rendement. Par exemple, le stochastic Gradient Descent a été introduit dans les années 90 pour traiter des Big-Data, tandis que le Gradient Descent par mini-lots est devenu une méthode courante pour améliorer la stabilité des modèles Deep Learning. De nos jours, le Gradient Descent reste une méthode de base pour l'optimisation de modèles de machine learning et est largement utilisé dans la conception de nombreuses applications telles que : la reconnaissance de la parole, la reconnaissance d'images, la traduction automatique et la prédiction des séries chronologiques.

B-Présentation de la méthodologie adoptée

i. Énoncé de l'algorithme de Gradient Descent

Si une fonction f est différentiable et convexe, et que son gradient est borné par une constante M sur l'ensemble de ces variables, alors l'algorithme de descente de gradient converge vers un minimum global de $f(x)$ à un taux de convergence de l'ordre $1/l$ avec l le nombre d'itérations.

En d'autres termes, si l'algorithme de descente de gradient est appliqué à une fonction convexe et différentiable, et que le gradient de cette fonction ne varie pas trop rapidement, alors l'algorithme convergera vers le minimum global de cette fonction avec une certaine rapidité.

ii. Méthodologie adoptée

1-Initialisation des paramètres: Les paramètres du modèle sont initialisés avec des valeurs aléatoires ou prédéfinies.

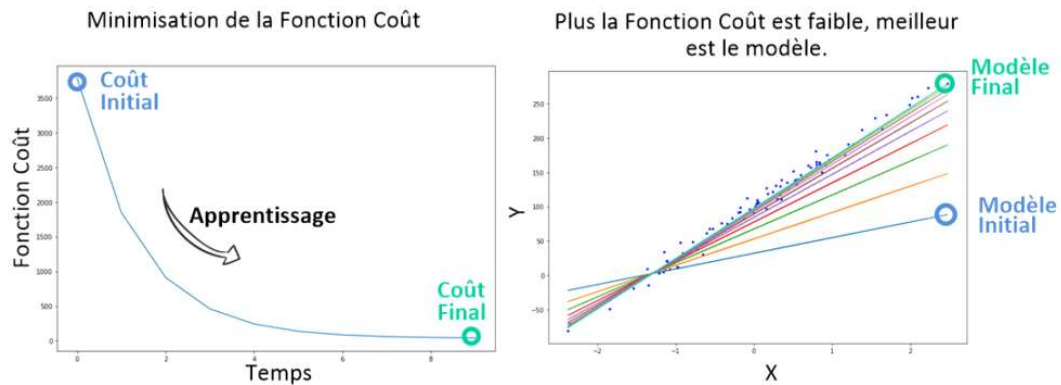
2-Le calcul de la fonction cout: en utilisant les paramètres du modèle et l'ensemble de données d'entraînement.

3-Le calcul du gradient: qui permet l'indication de la direction de la pente la plus raide de la fonction de cout.

4-La mise à jour des paramètres: en suivant la direction opposée du gradient et en utilisant une certaine taille de pas ou taux d'apprentissage pour garantir une convergence rapide.

5-La répétition du processus: Les étapes 2 à 4 sont répétées jusqu'à ce que la fonction de cout soit minimisée ou que le nombre maximum d'itérations soit atteint.

En machine learning, l'algorithme de gradient descent est fondamental puisque la machine apprend grâce à lui.



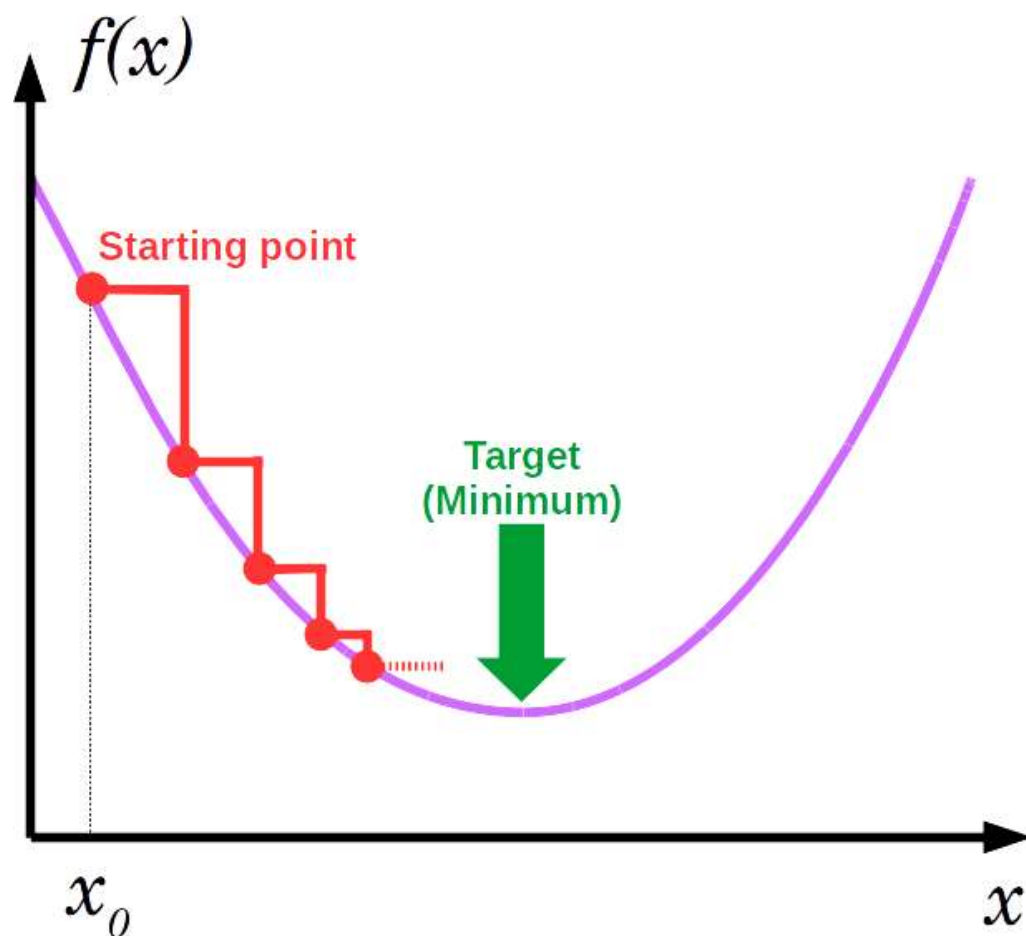
C-Domains d'utilisation du Gradient Descent

L'algorithme de Gradient Descent est adapté aux données de type numérique, il convient aux petites données mais surtout aux données de grandes tailles. Il est couramment utilisé dans de nombreux domaines tels que:

- 1-L'apprentissage automatique: plus particulièrement dans l'apprentissage supervisé, où il est utilisé pour minimiser la fonction de perte lors de l'entraînement des modèles de machine learning comme les réseaux de neurones artificiels.
- 2-La vision par Ordinateur: où il ajuste les paramètres des modèles de vision par ordinateur, tels que les filtres de convolution utilisés pour la détection des caractéristiques dans des images.
- 3-La reconnaissance vocale: où il ajuste les paramètres des modèles de reconnaissance vocale, tels que les réseaux de neurones récurrents et les réseaux de neurones convolutionnels.
- 4-La finance: où il ajuste les paramètres des modèles financiers, tels que les modèles de prévision de la volatilité et les modèles de prévision des prix.

5-L'Optimisation de la conception: utile notamment dans plusieurs domaines comme la mécanique des fluides,l'aéronautique ou encore l'automobile.

6-La recherche opérationnelle:Utilisé dans de nombreux domaines pour optimiser les modèles et les processus ,et pour résoudre des problèmes d'optimisation.



D-Comparaison avec d'autres algorithmes similaires

Il existe plusieurs algorithmes similaires à l'algorithme de Gradient

Descent qui sont également utilisés pour résoudre les problèmes d'optimisation numérique.

1-L'algorithme de descente de gradient stochastique

Contrairement à la descente de gradient classique, cet algorithme calcule le gradient sur un seul exemple d'entraînement à la fois, plutôt que sur l'ensemble des données complet. Cette approche permet de réduire le temps de calcul nécessaire à l'optimisation des modèles de Machine Learning. Cette algorithme est utilisé notamment pour entraîner les réseaux de neurones pour la classification des images

2-L'algorithme de Newton

Contrairement à l'algorithme de descente de gradient qui ne prend en compte que la pente, il utilise également les informations sur la courbure de la fonction. Celui-ci lui permet de converger rapidement vers le minimum global, mais peut-être plus coûteux en temps de calcul.

3-L'algorithme de quasi-Newton

C'est généralisation de l'algorithme de Newton; il permet de conserver les avantages de l'algorithme de Newton tout en évitant les coûts élevés de calcul en temps.

4-L'algorithme de conjugaison de gradient

Cet algorithme prend en compte les fonctions non convexes contrairement au gradient descent. Plus précisément, c'est une méthode itérative qui cherche la direction descente en utilisant des informations sur les directions précédemment trouvées. Cette approche est utile pour les fonctions non convexes, car elle permet de se déplacer dans plusieurs directions à la fois pour éviter de se retrouver coincé dans un minimum local.

E-Avantages et limites de l'algorithme de Gradient Descent

a. Avantages

Les avantages liés à l'algorithme de Gradient Descent sont multiples. On peut citer:

- * La Convergence rapide vers la solution optimale en ajustant les paramètres du modèle.
- * La flexibilité grâce auquel on peut optimiser différents types de fonctions de coût et de modèle de machine learning.
- * L'évolutivité caractérisée par la création du stochastic Gradient et le Mini-Batch Gradient Descent.
- * La facilité des implémentations.

b. Limites

Comme contraintes liées à l'algorithme de Gradient Descent, on peut citer:

- * La convergence vers un minimum local
- * Le risque de surapprentissage
- * Sensible aux points de départ.

En résumé, le Gradient Descent est une méthode puissante pour optimiser les modèles de machine learning, mais elle nécessite une attention particulière pour être mise en œuvre de manière optimale.

Conclusion

L'algorithme de gradient descent est un algorithme d'optimisation numérique couramment utilisé dans de nombreux domaines tels que l'apprentissage automatique, la vision par ordinateur ou encore l'automatisation de la conception. Le but de cet algorithme est de minimiser une fonction de coût en ajustant les paramètres du modèle. L'algorithme calcule le gradient de la fonction coût par rapport aux paramètres du modèle. L'algorithme est itératif et répète le processus jusqu'à l'obtention d'un minimum local ou global. Cet algorithme peut être réalisé de différentes manières, notamment avec l'algorithme de Newton par exemple. Il présente des avantages de même que des inconvénients en fonction du problème spécifique à résoudre.