

Optimisation et RO

Certificat Chef de Projet IA

Quizz - 2023/2024

Instructions

- Ce court questionnaire permet de valider le module Optimisation et recherche opérationnelle réalisé lors de la semaine du 9 au 13 octobre 2023.
- Chaque question ne possède qu'une réponse correcte.
- Vos réponses individuelles à ce questionnaire sont à transmettre par mail à clement.royer@lamsade.dauphine.fr.

Questions sur la partie 1 - Optimisation convexe

Question 1 *Quel est l'intérêt principal de formuler un problème de sciences des données comme un programme convexe (type LP, QP, SDP, CP) ?*

- a) *Il existe des solveurs numériques bien établis dédiés à ce type de problèmes.*
- b) *Il existe des algorithmes de points intérieurs avec garanties théoriques pour résoudre ce type de problèmes.*
- c) *On peut toujours trouver une solution au problème.*
- d) *On peut toujours résoudre le problème avec plusieurs millions de variables.*

Question 2 *On cherche à coller à un jeu de données via un modèle linéaire. Laquelle des affirmations suivantes est alors fausse ?*

- a) *Calculer le modèle par les moindres carrés est adapté au cas d'erreurs gaussiennes.*
- b) *On peut calculer l'approximation en norme ℓ_1 en résolvant un programme linéaire.*
- c) *On ne peut jamais calculer un modèle linéaire qui interpole parfaitement les données.*
- d) *Il existe toujours un modèle qui donne la meilleure approximation de Chebyshev.*

Questions sur la partie 2 - Optimisation non convexe

Question 3 Parmi les affirmations suivantes, laquelle est **fausse** ?

- a) Il existe des fonctions non convexes pour lesquelles tout minimum local est global.
- b) Il existe des fonctions convexes qui ont un unique minimum global.
- c) Il existe des fonctions convexes qui possèdent des maxima locaux non globaux.
- d) Il existe des fonctions non convexes qui possèdent des maxima locaux et des minima locaux.

Question 4 Supposons que l'on applique une descente de gradient avec une longueur de pas bien choisie à un problème convexe. Laquelle des affirmations ci-dessous est alors **fausse** ?

- a) L'algorithme converge en vitesse $\mathcal{O}(1/K)$.
- b) L'algorithme converge vers la valeur optimale du problème.
- c) L'algorithme peut être accéléré pour obtenir une vitesse de convergence optimale.
- d) L'algorithme n'est pas applicable à une fonction fortement convexe.

Questions sur la partie 3 - Gradient stochastique

Question 5 Parmi les propositions ci-dessous, laquelle **n'est pas une raison** pour utiliser les méthodes de gradient stochastique ?

- a) La taille du jeu de données peut être très grande.
- b) L'accès aux données peut être coûteux.
- c) Le gradient par rapport à une partie des données peut être une bonne direction.
- d) La descente de gradient ne converge pas sur ces problèmes.

Question 6 Parmi ces approches, laquelle est considérée comme la moins coûteuse dans un contexte de somme finie ?

1. Appliquer une descente de gradient.
2. Appliquer l'algorithme du gradient stochastique avec mini-batch.
3. Appliquer l'algorithme du gradient accéléré.
4. Appliquer l'algorithme du gradient stochastique de base.

Questions sur la partie 4 - Optimisation sans(?) dérivées

Question 7 Parmi les approches suivantes, laquelle est équivalente au processus dit de backpropagation ?

- a) Différences finies + itération gradient proximal;
- b) Calcul du sous-différentiel complet + itération de sous-gradient;
- c) Différentiation automatique+ itération de gradient stochastique;
- d) Codage à la main des dérivées + itération de descente de gradient.

Question 8 Laquelle de ces affirmations sur l'optimisation sans dérivées est **fausse** ?

- a) Les algorithmes génétiques peuvent être performants sur des problèmes de petite dimension (2-3).
- b) Les méthodes basées sur des modèles exploitent plus l'information des valeurs de fonctions déjà obtenues que celles de recherche directe.
- c) Un algorithme d'optimisation sans dérivées ne peut utiliser aucune information sur la fonction, et la traite forcément comme une "boîte noire".
- d) Les dernières avancées en optimisation sans dérivées sont basées sur l'interaction avec d'autres domaines comme les approches de bandits.

Questions sur la partie 5 - Optimisation à grande échelle

Question 9 Quel est l'intérêt principal de la méthode SVRG ?

- a) Elle fonctionne en parallèle;
- b) Elle utilise le momentum;
- c) Elle est moins coûteuse que le gradient stochastique;
- d) Elle permet de réduire la variance du gradient stochastique.

Question 10 Dans quel contexte **n'est-il pas** pertinent d'utiliser une approche de descente par coordonnées ?

- a) Lorsque le problème est séparable;
- b) Lorsque le problème est partiellement séparable;
- c) Lorsque le problème est non lisse;
- d) Lorsque la dimension du problème est trop élevée pour stocker un gradient en mémoire.