Optimisation convexe

Référentiel cours

Bashar Dudin

February 18, 2020

EPITA

Les problèmes d’optimisation

Introduction

L'optimisation fait partie des missions historiques de l'ingénierie. Elle naît avec l'ère industrielle: une fois un concept élaboré il s'agit de réduire les coups, minimiser les risques de défauts de livraisons ou étendre le scope d'action ...

[1] Ou non ...

Introduction

L'optimisation fait partie des missions historiques de l'ingénierie. Elle naît avec l'ère industrielle: une fois un concept élaboré il s'agit de réduire les coups, minimiser les risques de défauts de livraisons ou étendre le scope d'action ...

Les techniques mathématiques qui permettent de résoudre une partie de ces problèmes d'optimisation balayent un large spectre des thématiques mathématiques que vous avez pu aborder jusque là[1]; l'algèbre linéaire, le calcul différentielle et un peu de géométrie. Le cours d'OCVX a pour objectif de vous donner le bon degré de confort pour manipuler ces techniques.

[1] Ou non ...

Problèmes d'optimisation – De quoi on parle?

Voici quelques exemples qu'on pourrait croiser lorsqu'on s'intéresse à l'optimisation.

- Chercher le plus court/rapide chemin entre deux coordonnées GPS.

Problèmes d’optimisation – De quoi on parle?

Voici quelques exemples qu’on pourrait croiser lorsqu’on s’intéresse à l’optimisation.

- Chercher le plus court/rapide chemin entre deux coordonnées GPS.
- Décider des meilleures routes aériennes qui minimisent le prix d’approvisionnement en kérosène.

Problèmes d'optimisation – De quoi on parle?

Voici quelques exemples qu'on pourrait croiser lorsqu'on s'intéresse à l'optimisation.

- Chercher le plus court/rapide chemin entre deux coordonnées GPS.
- Décider des meilleures routes aériennes qui minimisent le prix d'approvisionnement en kérosène.
- Identifier des images d'IRM qui correspondent à des malformations du cerveau.

Problèmes d’optimisation – De quoi on parle?

Voici quelques exemples qu’on pourrait croiser lorsqu’on s’intéresse à l’optimisation.

- Chercher le plus court/rapide chemin entre deux coordonnées GPS.
- Décider des meilleures routes aériennes qui minimisent le prix d’approvisionnement en kérosène.
- Identifier des images d’IRM qui correspondent à des malformations du cerveau.
- Chercher des patterns dans la population d’étudiants intégrants Epita.

Problèmes d'optimisation – De quoi on parle?

Voici quelques exemples qu'on pourrait croiser lorsqu'on s'intéresse à l'optimisation.

- Chercher le plus court/rapide chemin entre deux coordonnées GPS.
- Décider des meilleures routes aériennes qui minimisent le prix d'approvisionnement en kérosène.
- Identifier des images d'IRM qui correspondent à des malformations du cerveau.
- Chercher des patterns dans la population d'étudiants intégrants Epita.
- Décider d'achat/vente d'assets prenant en compte l'historique disponible.

Formulation mathématique

Problèmes d'optimisation – définition formelle

On écrit en général un problème d'optimisation (P) sous la forme standard

$$
\begin{aligned}
&\text{minimiser} && f_0(x) \\
&\text{sujet à} && \\
&& f_i(x) \leq 0, &\quad \forall i \in \{1, \ldots, p\} \\
&& h_j(x) = 0, &\quad \forall j \in \{1, \ldots, m\}
\end{aligned}
$$

où f_0, les f_i et les h_j sont des applications de $\mathbb{R}^n$ vers $\mathbb{R}$. La fonction f_0 est dite ***fonction objectif*** ; suivant le contexte ce sera une fonction de ***coût*** ou d'***erreur***. Les inégalités sont qualifiées de ***contraintes d'inégalités*** et les égalités de ***contraintes d'inégalités***.

Question

Vous pouvez chercher à formuler les problèmes précédents sous forme d'un problème d'optimisation, ce n'est pas toujours évident.

Problèmes d'optimisation – définition formelle

Un problème d'optimisation du type de (P) est dit

- différentiable si toutes les fonctions en jeux le sont ;

Problèmes d'optimisation – définition formelle

Un problème d'optimisation du type de (P) est dit

- différentiable si toutes les fonctions en jeux le sont ;
- ***non-contraint*** s'il n'a aucune contraintes d'inégalités ou égalités ;

Problèmes d'optimisation – définition formelle

Un problème d'optimisation du type de (P) est dit

- différentiable si toutes les fonctions en jeux le sont ;
- ***non-contraint*** s'il n'a aucune contraintes d'inégalités ou égalités ;
- ***convexe*** si l'ensemble des fonctions en jeu sont convexes, les contraintes d'égalités étant de plus affines.

Problèmes d'optimisation – définition formelle

Un problème d'optimisation du type de (P) est dit

- différentiable si toutes les fonctions en jeux le sont ;
- ***non-contraint*** s'il n'a aucune contraintes d'inégalités ou égalités ;
- ***convexe*** si l'ensemble des fonctions en jeu sont convexes, les contraintes d'égalités étant de plus affines.

Sous la première hypothèse on a une série d'outils mathématiques qui nous permettront d'apporter un éclairage riche sur (P).

Problèmes d’optimisation – définition formelle

Un problème d’optimisation du type de (P) est dit

- différentiable si toutes les fonctions en jeux le sont ;
- ***non-contraint*** s’il n’a aucune contraintes d’inégalités ou égalités ;
- ***convexe*** si l’ensemble des fonctions en jeu sont convexes, les contraintes d’égalités étant de plus affines.

Sous la première hypothèse on a une série d’outils mathématiques qui nous permettront d’apporter un éclairage riche sur (P). Si l’on rajoute la seconde on est en mesure de construire des procédés itératifs efficaces en état de *résoudre* ces problèmes.

Problèmes d'optimisation – définition formelle

Un problème d'optimisation du type de (P) est dit

- différentiable si toutes les fonctions en jeux le sont ;
- ***non-contraint*** s'il n'a aucune contraintes d'inégalités ou égalités ;
- ***convexe*** si l'ensemble des fonctions en jeu sont convexes, les contraintes d'égalités étant de plus affines.

Sous la première hypothèse on a une série d'outils mathématiques qui nous permettront d'apporter un éclairage riche sur (P). Si l'on rajoute la seconde on est en mesure de construire des procédés itératifs efficaces en état de *résoudre* ces problèmes. La dernière nous garantie de trouver *la* solution optimale.

Problèmes d'optimisation – définition formelle

Un problème d'optimisation du type de (P) est dit

- différentiable si toutes les fonctions en jeux le sont ;
- ***non-contraint*** s'il n'a aucune contraintes d'inégalités ou égalités ;
- ***convexe*** si l'ensemble des fonctions en jeu sont convexes, les contraintes d'égalités étant de plus affines.

Sous la première hypothèse on a une série d'outils mathématiques qui nous permettront d'apporter un éclairage riche sur (P). Si l'on rajoute la seconde on est en mesure de construire des procédés itératifs efficaces en état de *résoudre* ces problèmes. La dernière nous garantie de trouver *la* solution optimale.

Fake News

Les éléments en italiques sont là pour marquer le fait que nos assertions à ce stade sont encore un peu fausses. L'image est un peu moins idyllique.

Problèmes d'optimisation – lexique

Étant donné un problème d'optimisation (P) on appelle:

- ***point admissible*** de (P) tout point de $\mathbb{R}^n$ satisfaisant toutes les contraintes. L'ensemble de tous les points admissibles est appelé ***lieu admissible*** de (P).

Problèmes d'optimisation – lexique

Étant donné un problème d'optimisation (P) on appelle:

- ***point admissible*** de (P) tout point de $\mathbb{R}^n$ satisfaisant toutes les contraintes. L'ensemble de tous les points admissibles est appelé ***lieu admissible*** de (P).
- ***valeur objectif*** d'un point admissible la valeur que prend la fonction objectif en celui-ci.

Problèmes d'optimisation – lexique

Étant donné un problème d'optimisation (P) on appelle:

- ***point admissible*** de (P) tout point de $\mathbb{R}^n$ satisfaisant toutes les contraintes. L'ensemble de tous les points admissibles est appelé ***lieu admissible*** de (P).
- ***valeur objectif*** d'un point admissible la valeur que prend la fonction objectif en celui-ci.
- ***valeur optimale*** de (P) la meilleure borne inférieure sur la fonction objectif.

Problèmes d'optimisation – lexique

Étant donné un problème d'optimisation (P) on appelle:

- ***point admissible*** de (P) tout point de $\mathbb{R}^n$ satisfaisant toutes les contraintes. L'ensemble de tous les points admissibles est appelé ***lieu admissible*** de (P).
- ***valeur objectif*** d'un point admissible la valeur que prend la fonction objectif en celui-ci.
- ***valeur optimale*** de (P) la meilleure borne inférieure sur la fonction objectif.
- ***point optimale*** de (P) tout point admissible dont la valeur objectif est la valeur optimale.

Problèmes d’optimisation – premières remarques qualitatives

Comme est le cas de tout système d’équations, il est utile de se poser le type de questions suivantes:

Problèmes d’optimisation – premières remarques qualitatives

Comme est le cas de tout système d’équations, il est utile de se poser le type de questions suivantes:

- y a-t-il au moins une solution?

Problèmes d'optimisation – premières remarques qualitatives

Comme est le cas de tout système d'équations, il est utile de se poser le type de questions suivantes:

- y a-t-il au moins une solution?
- s'il y a au moins une solution combien?

Problèmes d’optimisation – premières remarques qualitatives

Comme est le cas de tout système d’équations, il est utile de se poser le type de questions suivantes:

- y a-t-il au moins une solution?
- s’il y a au moins une solution combien?
- peut-on toujours décrire l’ensemble des solutions?

Problèmes d'optimisation – premières remarques qualitatives

Comme est le cas de tout système d'équations, il est utile de se poser le type de questions suivantes:

- y a-t-il au moins une solution?
- s'il y a au moins une solution combien?
- peut-on toujours décrire l'ensemble des solutions?
- y a-t-il moyen d'approcher des solutions?

Question

Chercher un problème d'optimisation qui:

- a un lieu admissible est vide ;
- a plus d'un seul point optimal ;
- n'a pas de valeur optimale mais a un lieu admissible non vide ;
- a une valeur optimale mais pas de point optimal.

Cadre de la première UE d’OCVX

Contours du cours

On se limite au cours du premier semestre de majeure au cas des problèmes d'optimisations ***sans contraintes***.

[2] Support Vector Machines.

Contours du cours

On se limite au cours du premier semestre de majeure au cas des problèmes d'optimisations ***sans contraintes***.

- C'est un cadre suffisant pour les premières applications des techniques d'optimisations à un premier niveau de ML ; il recouvre le cas des différentes régressions, de l'entraînement d'un réseau de neurones ainsi que les cas des algos de classification standards.

[2] Support Vector Machines.

Contours du cours

On se limite au cours du premier semestre de majeure au cas des problèmes d'optimisations ***sans contraintes***.

- C'est un cadre suffisant pour les premières applications des techniques d'optimisations à un premier niveau de ML ; il recouvre le cas des différentes régressions, de l'entraînement d'un réseau de neurones ainsi que les cas des algos de classification standards.
- Il représente un premier niveau à atteindre qui permet de fixer votre attitude vis-à-vis d'un problème d'optimisation, sans s'encombrer de concepts plus abstraits à concevoir.

[2] Support Vector Machines.

Contours du cours

On se limite au cours du premier semestre de majeure au cas des problèmes d'optimisations ***sans contraintes***.

- C'est un cadre suffisant pour les premières applications des techniques d'optimisations à un premier niveau de ML ; il recouvre le cas des différentes régressions, de l'entraînement d'un réseau de neurones ainsi que les cas des algos de classification standards.
- Il représente un premier niveau à atteindre qui permet de fixer votre attitude vis-à-vis d'un problème d'optimisation, sans s'encombrer de concepts plus abstraits à concevoir.
- Il ne recouvre pas le cas des SVM[2].

[2] Support Vector Machines.

Régression mon amie

Map Fitting

Définition

Une famille différentiable d'applications $f_\alpha : \mathbb{R}^n \to \mathbb{R}$ indexées par $\alpha \in \mathbb{R}^k$ est une famille de fonctions pour laquelle l'application $\phi : \mathbb{R}^k \times \mathbb{R}^n \to \mathbb{R}$ qui envoie (α, x) sur $f_\alpha(x)$ est différentiable.

Map Fitting

On considère un ensemble de couples $(X_i, y_i) \in \mathbb{R}^n \times \mathbb{R}$ pour $i \in \{1, \ldots, p\}$ et une famille différentiable d'applications $\{f_\alpha\}_{\alpha \in \mathbb{R}^k}$. Le problème de *map fitting* relatif aux données précédentes consiste à trouver les meilleurs paramètres $\alpha^\star$ tels que $f_{\alpha^\star}$ approche *au mieux*[a] les (X_i, y_i).

[a]Pour une métrique pré-choisie.

La régression linéaire

Le plus simple des problèmes de *map fitting* est celui de la régression linéaire. Dans le cas de dimension 1 (on cherche à approcher une fonction de $\mathbb{R}$ dans $\mathbb{R}$) il se décline comme ceci:

La régression linéaire

Le plus simple des problèmes de *map fitting* est celui de la régression linéaire. Dans le cas de dimension 1 (on cherche à approcher une fonction de $\mathbb{R}$ dans $\mathbb{R}$) il se décline comme ceci:

- la famille différentiable à laquelle on s'intéresse est indexées par $\mathbb{R}^2$: $f_\alpha(x) = \alpha_1 x + \alpha_0$ pour $\alpha = (\alpha_0, \alpha_1)$;

La régression linéaire

Le plus simple des problèmes de *map fitting* est celui de la régression linéaire. Dans le cas de dimension 1 (on cherche à approcher une fonction de $\mathbb{R}$ dans $\mathbb{R}$) il se décline comme ceci:

- la famille différentiable à laquelle on s'intéresse est indexées par $\mathbb{R}^2$: $f_\alpha(x) = \alpha_1 x + \alpha_0$ pour $\alpha = (\alpha_0, \alpha_1)$;
- la métrique standard utilisée est la ***MSE*** pour ***Mean Square Error*** donnée pour un f_α par

$$\mathcal{E}(\alpha) = \sum_{i=1}^{p} \frac{1}{p}\big(f_\alpha(X_i) - y_i\big)^2$$

  c'est une estimation moyenne de la variance des prédictions de f_α.

La régression linéaire

Le plus simple des problèmes de *map fitting* est celui de la régression linéaire. Dans le cas de dimension 1 (on cherche à approcher une fonction de $\mathbb{R}$ dans $\mathbb{R}$) il se décline comme ceci:

- la famille différentiable à laquelle on s'intéresse est indexées par $\mathbb{R}^2$: $f_\alpha(x) = \alpha_1 x + \alpha_0$ pour $\alpha = (\alpha_0, \alpha_1)$;
- la métrique standard utilisée est la ***MSE*** pour ***Mean Square Error*** donnée pour un f_α par

$$\mathscr{E}(\alpha) = \sum_{i=1}^{p} \frac{1}{p}\big(f_\alpha(X_i) - y_i\big)^2$$

  c'est une estimation moyenne de la variance des prédictions de f_α.

Le but est de trouver un paramètre $\alpha = (\alpha_1, \alpha_0)$ tel que $\mathscr{E}(\alpha)$ est minimal, autrement dit de résoudre le problème d'optimisation sans contraintes

$$\text{minimiser } E(\alpha).$$

La régression linéaire

Le problème de régression linéaire a une solution ***analytique*** ; càd une solution donnée par une expression explicite en fonction des entrées.

Cette solution implique cependant l'inversion d'une matrice de taille équivalent à celle des données en entrée. Chose particulièrement coûteuse.

Empathie machine

Approximation séquentielle

Il est rare qu'un problème d'optimisation ait une solution analytique. Même quand cela est le cas il est souvent plus efficace de chercher une solution approchée.

Un processus itératif qui résout un problème d'optimisation (P) est

Approximation séquentielle

Il est rare qu'un problème d'optimisation ait une solution analytique. Même quand cela est le cas il est souvent plus efficace de chercher une solution approchée.

Un processus itératif qui résout un problème d'optimisation (P) est

- un choix initial d'un point de départ (de préférence) admissible x_0 ;

Approximation séquentielle

Il est rare qu'un problème d'optimisation ait une solution analytique. Même quand cela est le cas il est souvent plus efficace de chercher une solution approchée.

Un processus itératif qui résout un problème d'optimisation (P) est

- un choix initial d'un point de départ (de préférence) admissible x_0 ;
- un processus itératif qui construit un point admissible x_{n+1} à partir de x_n et de données locales ayant une valeur objectif plus petite que celle de x_n.

Approximation séquentielle

Il est rare qu'un problème d'optimisation ait une solution analytique. Même quand cela est le cas il est souvent plus efficace de chercher une solution approchée.

Un processus itératif qui résout un problème d'optimisation (P) est

- un choix initial d'un point de départ (de préférence) admissible x_0 ;
- un processus itératif qui construit un point admissible x_{n+1} à partir de x_n et de données locales ayant une valeur objectif plus petite que celle de x_n.

Cette démarche ne nous offre en général qu'une approximation d'une solution. Elle a cependant l'avantage de pouvoir se dérouler en temps raisonnable. Il faut par ailleurs prendre en compte que l'implémentation des flottants en machines nous contraint déjà à approcher les grandeurs qu'on manipule.

Objectifs du cours

Objectifs du cours

- ***Savoirs***
  - Une compréhension des éléments composants un problème d’optimisation et des éléments nécessaires à son étude qualitative.
  - Une connaissance du panel d’outils à disposition pour résoudre un problème d’optimisation et des hyperparamètres qui déclinent et gouvernent ceux-ci.
  - Une familiarité avec les problèmes numériques qui apparaissent dans toute implémentation.

Objectifs du cours

- ***Savoirs***
  - Une compréhension des éléments composants un problème d'optimisation et des éléments nécessaires à son étude qualitative.
  - Une connaissance du panel d'outils à disposition pour résoudre un problème d'optimisation et des hyperparamètres qui déclinent et gouvernent ceux-ci.
  - Une familiarité avec les problèmes numériques qui apparaissent dans toute implémentation.
- ***Savoir-faire***
  - Une capacité à implémenter des algorithmes d'optimisation sans contraintes.
  - Une capacité à *benchmarker* des implémentations d'algorithmes d'optimisation et à analyser les performances de ces derniers.

Objectifs du cours

- ***Savoirs***
  - Une compréhension des éléments composants un problème d'optimisation et des éléments nécessaires à son étude qualitative.
  - Une connaissance du panel d'outils à disposition pour résoudre un problème d'optimisation et des hyperparamètres qui déclinent et gouvernent ceux-ci.
  - Une familiarité avec les problèmes numériques qui apparaissent dans toute implémentation.
- ***Savoir-faire***
  - Une capacité à implémenter des algorithmes d'optimisation sans contraintes.
  - Une capacité à *benchmarker* des implémentations d'algorithmes d'optimisation et à analyser les performances de ces derniers.
- ***Attitude*** - Analyse de risque
  - Un questionnement fin quant à la validité et limites des algorithmes implémentés.
  - ***Un*** test n'est pas une ***statistique*** et une ***statistique*** ne vient pas sans ***variabilité***.

Contenus notionnels

Tables des matières

1. Pré-requis techniques.
    - Des éléments de géométries
        - Interprétation géométrique du produit scalaire.
        - Courbes de niveau et epigraphes de fonctions.
        - Parties et fonctions convexese en dimension finie.
    - Des éléments de topologie
        - Comment calculer des distances et définir des voisinanges dans $\mathbb{R}^n$
        - Approcher et comparer des fonctions à plusieurs variables.
    - Des éléments de calcul différentiel
        - Approcher localement une fonction multivariées par une fonction affine.
        - Écriture en base ; jacobienne et gradient.
        - Interprétation géométrique du gradient.
        - Approximation locale de second ordre : la hessienne.

Tables des matières

Évaluation

Évaluation

Deux modes d’évaluations vont intéragir dans le cadre de ce cours

1. des évaluations automatisées via ***moodle***, ce qui est suffisant pour garantir le fait que vous consacrez suffisamment de temps à comprendre les éléments de cours et l’intérioriser
2. des évaluations sur soutenances de TP afin d’avoir un regard sur l’ensemble des éléments que vous aurez pu mobiliser pour atteindre les objectifs de cours.

Évaluation

Deux modes d'évaluations vont intéragir dans le cadre de ce cours

1. des évaluations automatisées via ***moodle***, ce qui est suffisant pour garantir le fait que vous consacrez suffisamment de temps à comprendre les éléments de cours et l'intérioriser
2. des évaluations sur soutenances de TP afin d'avoir un regard sur l'ensemble des éléments que vous aurez pu mobiliser pour atteindre les objectifs de cours.

On attend de vous de faire preuve ***d'autonomie*** lors du suivi de ce cours. Les participations actives des étudiants au maintien et à la création de contenus seront des éléments d'appréciation dans la notation finale.

Moodle

Moodle

L'ensemble des contenus de cours, d'annonces, de bibliographies de travail à rendre et des tests d'évaluation seront disponibles sur un cours moodle auquels vous serez inscrit bientôt.

Vous y trouverez également un ***glossaire*** et un ***wiki*** collaboratifs sur lesquels on vous invite à résumer les éléments de cours ou de thématiques autour de l'optimisation convexe que vous souhaitez.

Bon courage pour la suite!