Optimisation convexe

Contours du cours

Bashar Dudin

February 20, 2019

EPITA

Les problèmes d’optimisation

Formulation mathématique

ML – Un exemple d’opti à garder en tête

Méthodes itératives pour résoudre un problème d’opti

Contours du cours

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’opti – 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’opti – 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'opti – 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'opti – 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’opti – 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.

Problèmes d'opti – définition formelle

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

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

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 en problème d'optimisation, ce n'est pas toujours évident.

Problèmes d'opti – définition formelle

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

- différentiable si toutes les fonctions en jeux le sont ;

Problèmes d'opti – 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'opti – 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'opti – 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'opti – 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'opti – 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'opti – 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'opti – 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'opti – 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'opti – 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'opti – 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'opti – 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'opti – 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’opti – 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’opti – 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'opti – 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.

Map Fitting

On s'intéresse plus en détail à un problème d'opti qu'on qualifie de *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

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

  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

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

Les problèmes d'optimisation

Formulation mathématique

ML – Un exemple d'opti à garder en tête

Méthodes itératives pour résoudre un problème d'opti

Contours du cours

Approximation séquentielle

Il est rare qu'un problème d'optimisation ait une solution analytique[2]. 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

[2] Une solution donnée par une expression explicite en fonction des entrées.

Approximation séquentielle

Il est rare qu'un problème d'optimisation ait une solution analytique[2]. 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 ;

[2] Une solution donnée par une expression explicite en fonction des entrées.

Approximation séquentielle

Il est rare qu'un problème d'optimisation ait une solution analytique[2]. 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.

[2] Une solution donnée par une expression explicite en fonction des entrées.

Approximation séquentielle

Il est rare qu'un problème d'optimisation ait une solution analytique[2]. 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.

Lors de ce cours on va introduire l'ensemble des contenus nécessaires à la compréhension des méthodes itératives standards pour résoudre des problèmes d'optimisation.

[2] Une solution donnée par une expression explicite en fonction des entrées.

Les problèmes d’optimisation

Formulation mathématique

ML – Un exemple d’opti à garder en tête

Méthodes itératives pour résoudre un problème d’opti

Contours du cours

Contours du cours – évaluation

- La première partie de cours, plutôt théorique sera évaluée par un DM ainsi qu’un partiel sur table en avril.
- La seconde, plus applicative sera évaluée par un travail d’analyse par groupe. Il donnera lieu à une soutenance.

Contours du cours – contenus

1. Pré-requis techniques.
   - Des éléments de géométries ; occasionne un retour sur PROL ; la plus simple des familles de problèmes d'optimisation et le point de départ des techniques étudiées.
   - Du calcul différentiel ; introduit les éléments utiles pour généraliser la démarche PROL.

Contours du cours – contenus

1. Pré-requis techniques.
    - Des éléments de géométries ; occasionne un retour sur PROL ; la plus simple des familles de problèmes d'optimisation et le point de départ des techniques étudiées.
    - Du calcul différentiel ; introduit les éléments utiles pour généraliser la démarche PROL.
2. Étude qualitative des problèmes d'optimisation.
    - Cas sans contrainte ; lieu critique.
    - Le langrangien et le problème dual.
    - Les conditions KKT.
    - L'apport de la convexité et applicabilité en général.

Contours du cours – contenus

1. Pré-requis techniques.
    - Des éléments de géométries ; occasionne un retour sur PROL ; la plus simple des familles de problèmes d'optimisation et le point de départ des techniques étudiées.
    - Du calcul différentiel ; introduit les éléments utiles pour généraliser la démarche PROL.
2. Étude qualitative des problèmes d'optimisation.
    - Cas sans contrainte ; lieu critique.
    - Le langrangien et le problème dual.
    - Les conditions KKT.
    - L'apport de la convexité et applicabilité en général.
3. Méthodes de résolutions itératives.
    - Les descentes de gradients et méthodes de Newton.
    - Les Méthodes de point intérieurs.
    - La SMO pour les SVMs.

Contours du cours – contenus

1. Pré-requis techniques.
   - Des éléments de géométries ; occasionne un retour sur PROL ; la plus simple des familles de problèmes d'optimisation et le point de départ des techniques étudiées.
   - Du calcul différentiel ; introduit les éléments utiles pour généraliser la démarche PROL.
2. Étude qualitative des problèmes d'optimisation.
   - Cas sans contrainte ; lieu critique.
   - Le langrangien et le problème dual.
   - Les conditions KKT.
   - L'apport de la convexité et applicabilité en général.
3. Méthodes de résolutions itératives.
   - Les descentes de gradients et méthodes de Newton.
   - Les Méthodes de point intérieurs.
   - La SMO pour les SVMs.
   - Guillaume ? À l'aide.

Bon courage pour la suite!