Un peu de calcul différentiel

Résumé

On souhaite généraliser la démarche adoptée pour résoudre géométriquement certains programmes linéaires au cas des problèmes d'optimsation non-linéaires. L'objectif est de trouver des hyperplans d'appui aux sous-niveaux d'une fonction objectif ; ils indiquent une direction de recherche pour minimiser celle-ci. La définition de ces hyperplans d'appui s'effectue par une étude locale des fonctions objectifs, une étude qui généralise l'apport de la dérivée d'une fonction numérique à la compréhension du comportement de celle-ci en un point.

Table des matières

1 Normes sur $\mathbb{R}^n$

Une norme sur un espace vectoriel apporte une manière de mesurer la *longueur* d'un vecteur tout en respectant un minimum la structure vectorielle. Elle permet en particulier de définir une notion de distance entre deux point de $\mathbb{R}^n$ par la *longueur* du vecteur qui les relie. Pouvoir changer de mesure de longueur suivant les problèmes qu'on attaque est crucial lorsque l'on s'attaque à des problèmes d'optimisation. À la fois pour pouvoir modéliser les problèmes en jeu ; comment mesure la différence entre deux mots ? Ou pour accélérer la convergence de certains algorithmes d'apprentissages.

Définition 1.1. Une norme sur $\mathbb{R}^n$ est une application $\|\cdot\| : \mathbb{R}^n \to \mathbb{R}_+$ telle que :

1. $\|x\| = 0 \Leftrightarrow x = 0$;
2. $\forall \lambda \in \mathbb{R}$, $\forall x \in \mathbb{R}^n$, $\|\lambda x\| = |\lambda|\|x\|$ (relation d'*homogéinité*) ;
3. $\forall x \in \mathbb{R}^n$, $\forall y \in \mathbb{R}^n$, $\|x + y\| \leq \|x\| + \|y\|$ (*Inégalité triangulaire*).

Question 1-1 Interpréter chacune des propriétés suivantes avec vos propres mots.

L'inégalité triangulaire donne lieu à une autre inégalité, appelée *inégalité triangulaire inversée* qui peut parfois être utile [1] :

$$\forall x, y \in \mathbb{R}^n, \qquad \big|\|x\| - \|y\|\big| \leq \|x - y\|.$$

1. Il est très probable qu'elle ne fasse que partie de votre culture ...

Question 1-2 Représenter cette inégalité géométriquement. Essayer de la déduire de l'inégalité triangulaire.

Les trois normes les plus fréquentes d'utilisation sont les trois suivantes, elles sont respectivement qualifiées de normes 1, 2 et infinie.

1. $\forall x \in \mathbb{R}^n$, $\|x\|_1 = \sum_{i=1}^n |x_i|$;
2. $\forall x \in \mathbb{R}^n$, $\|x\|_2 = \sqrt{\sum_{i=1}^n x_i^2} = \sqrt{x^T x}$;
3. $\forall x \in \mathbb{R}^n$, $\|x\|_\infty$, $\max\{|x_i| 1 \leq i \leq n\}$.

Les définitions des normes 1 et 2 vont pouvoir se généraliser pour tout $p \geq 1$ [2], par

$$\forall x \in \mathbb{R}^n, \qquad \|x\|_p = \left(\sum |x_i|^p\right)^{\frac{1}{p}}.$$

La norme $\|\cdot\|_p$ est communément appelée la norme p. Montrer le fait que c'est une norme est uniquement délicat pour ce qui est de l'inégalité triangulaire, la preuve de ce fait se base sur des inégalités de convexités dites de Hölder, on n'abordera pas cette preuve dans ce cours.Les normes p ne seront que très localement utilisées en dehors des cas standards des normes 1, 2 et $+\infty$, il est cela dit usuel d'en connaître la définition.
À partir d'une norme on va être en mesure de définir :

Une notion de *distance* entre deux points de $\mathbb{R}^n$ par norme du vecteur de qui relie les deux points en question. Plus formellement étant donné une norme $\|\cdot\|$ on note

$$\forall x, y, \qquad d_{\|\cdot\|}(x, y) = \|x - y\|.$$

Dans le cas des normes p on se contente d'indiquer p en indice.

Question 1-3 Représenter graphiquement les distances 1, 2 et $+\infty$ entre deux vecteurs de $\mathbb{R}^2$.

Une notion de *voisinage* d'un point de $\mathbb{R}^n$ au sens de la norme utilisée. Cette notion éminemment liée à la première, on l'isole ici parce qu'elle apporte un point de vue auquel vous n'avez pas encore été confrontés. Étant donné un nombre réel $\varepsilon > 0$ on va qualifier ε-voisinage d'un point $x \in \mathbb{R}^n$ tous les points à distance (au sens de la norme utilisée) au plus ε de x. Dans le jargon, on appelle boule ouverte de rayon ε et centrée en x cette notion, pour une norme ambiante $\|\cdot\|$ on note

$$B_{\|\cdot\|}(x, \varepsilon) = \{y \in \mathbb{R}^n \mid d_{\|\cdot\|}(x, y) < \varepsilon\}$$

La boule fermée centrée en x et de rayon ε est la notion correspondante avec les points à distance ε incluses,

$$\overline{B}_{\|\cdot\|}(x, \varepsilon) = \{y \in \mathbb{R}^n \mid d_{\|\cdot\|}(x, y) \leq \varepsilon\}.$$

Dans le cas des normes p on se contente d'indexer les boules par p.
La notion de ε-voisinage permet d'exprimer des phénomènes de passage à la limite et d'études locales. Dire qu'une suite $(u_k)_{k\in\mathbb{N}}$ de points de $\mathbb{R}^n$ converge vers $\ell \in \mathbb{R}^n$, au sens d'une norme $\|\cdot\|$, correspond au fait de dire que pour tout ε-voisinage $B(\ell, \varepsilon)$ de ℓ, il existe un rang N à partir duquel tous les éléments de la suite u_k sont dans $B(\ell, \varepsilon)$.

Question 1-4
- Dessiner les boules unités $\overline{B}_p(\underline{0}, 1)$ pour $p \in \{1, 2, \infty\}$.
- Montrer que les ε-voisinages d'une norme $\|\cdot\|$ sur $\mathbb{R}^n$ sont convexes.
- La définition d'un équivalent à une norme p pour $p < 1$ définit-il une norme ?

2. Il est tout à fait naturel de se poser la question de savoir pourquoi $p \geq 1$! :)

On reprend plus en détails quelques extensions des notions liées aux études locales de fonctions en suites dans $\mathbb{R}$.

Exemple 1.1 (Convergence d'une suite.)**.** Une suite $(u_k)_{k\in\mathbb{N}}$ dans $\mathbb{R}^n$ converge vers un point $\ell \in \mathbb{R}^n$, au sens d'une norme $\|\cdot\|$ si

$$\forall \varepsilon > 0, \exists N \in \mathbb{N}, \quad k \geq N \Rightarrow \|u_k - \ell\| < \varepsilon.$$

Cette défintion se traduit telle quelle par :

$$\forall \varepsilon > 0, \exists N \in N \quad k \geq N \Rightarrow u_k \in B(\ell, \varepsilon).$$

Ou encore : pour tout $\varepsilon > 0$ il existe un rang N à partir duquel tous les éléments de la suite u_k sont dans le ε-voisinage de ℓ pour la norme $\|\cdot\|$.
Dans le cas de la norme infinie, la définition précédente s'écrit explicitement comme

$$\forall \varepsilon > 0, \exists N \in \mathbb{N}, \quad k \geq N \Rightarrow \max_{1\leq i\leq n} \|u_{k,i} - \ell_i\| < \varepsilon.$$

où $u_{k,i}$ (resp. ℓ) est la composante le long de la coordonnée i du vecteur u_k (resp. ℓ_i). Cette définition peut encore être réécrite

$$\forall \varepsilon > 0, \exists N \in \mathbb{N}, \quad k \geq N \Rightarrow \forall i \in \{1, \ldots, n\}, \|u_{k,i} - \ell_i\| < \varepsilon.$$

Un peu de réflexion permet de voir que cette dernière est équivalente

$$\forall i \in \{1, \ldots, n\}, \forall \varepsilon > 0, \exists N \in \mathbb{N}, \quad k \geq N \Rightarrow \|u_{k,i} - \ell_i\| < \varepsilon.$$

Autrement dit, la suite (u_k) converge vers ℓ, au sens de la norme infinie, si et seulement si, chacune des suites **numériques** $(u_{k,i})$ converge vers ℓ_i.

Exemple 1.2 (Limite d'une fonction en un point.)**.** On considère deux normes $\|\cdot\|_\alpha$ et $\|\cdot\|_\beta$ respectivement sur $\mathbb{R}^n$ et $\mathbb{R}^m$. Une fonction $f : (\mathbb{R}^n, \|\cdot\|_\alpha) \to (\mathbb{R}^m, \|\cdot\|_\beta)$ a une limite $\ell \in \mathbb{R}^m$ en un point $a \in \mathbb{R}^n$ si

$$\forall \varepsilon > 0, \exists \eta > 0, \quad x \neq a \text{et} \|x - a\|_\alpha \eta \Rightarrow \|f(x) - \ell\|_\beta < \varepsilon.$$

Chose qu'on peut encore exprimer par

$$\forall \varepsilon > 0, \exists \eta > 0, \quad x \in B_{\|\cdot\|_\alpha}(a, \eta) \setminus \{a\} \Rightarrow f(x) \in B_{\|\cdot\|_\beta}(\ell, \varepsilon),$$

ou encore pour tout ε-voisinage B de ℓ pour la norme $\|\cdot\|_\alpha$ il existe un η-voisinage de a, a étant exclu dont tout élément a une image dans B.

Exemple 1.3 (Continuité d'une fonction en un point.)**.** La continuité d'une fonction $f : (\mathbb{R}^n, \|\cdot\|_\alpha) \to (\mathbb{R}^m, \|\cdot\|_\beta)$ en un point $a \in \mathbb{R}^n$ s'écrit

$$\forall \varepsilon > 0, \exists \eta > 0, \quad \|x - a\|_\alpha \eta \Rightarrow \|f(x) - f(a)\|_\beta < \varepsilon.$$

Autrement dit pour tout ε-voisinage B de $f(a)$ pour la norme $\|\cdot\|_\alpha$ il existe un η-voisinage de a dont tout élément a une image dans B.

Question 1-5

— Donner des exemples de fonctions continues (pour les normes de votre choix) de $\mathbb{R}^2$ dans $\mathbb{R}$.
— Comment peut-on en construire d'autres à partir de celles-ci ? Que dire de l'addition, la multiplication, le quotient ou la composition de fonctions continues ?
— Comment tester la continuité de fonctions à plusieurs variables en se ramenant au cas des fonctions numériques ?
— Est-ce que ça marche tous le temps ?

Exemple 1.4 (Comparaisons en un point.). Une fonction $f : (\mathbb{R}^n, \|\cdot\|_\alpha) \to (\mathbb{R}^m, \|\cdot\|_\beta)$ est un o en un point $a \in \mathbb{R}^n$ d'une fonction g, sur les mêmes espaces et par rapport aux mêmes normes, s'il existe une fonction $\epsilon : \mathbb{R}^n \to \mathbb{R}$ telle que $f = \epsilon g$ avec $\lim_{t \to a} \epsilon(t) = 0$. Quand g n'est pas nulle sur un voisinage de a la définition précédente est équivalente au fait que

$$\lim_{t \to a} \frac{\|f(t)\|_\alpha}{\|g(t)\|_\beta} = 0.$$

On écrit dans ce cas que f est un $o_a(g)$. On laisse tomber le point a quand celui-ci est clair du contexte. Pour obtenir les notions d'équivalence et de $\mathcal{O}$ en un point a, il suffit de remplacer les limites de ϵ ou de quotient respectivement par 1 et une constante positive quelconque.

Question 1-6 Justifier les affirmations suivantes :

1. $\langle h, h \rangle$ est un $o_{\underline{0}}(h)$;
2. $\sin\big(\langle h, h \rangle\big) \sim_{\underline{0}} \langle h, h \rangle$.

À ce stade on est en droit de se demander si le choix des normes avec lesquels on travaille a une incidence sur les limites des suites qu'on étudie (une même suite convergerait pour une norme et pas pour une autre), les fonctions qu'on regarde (une même fonction serait continue pour une norme et pas pour une autre) etc. Le théorème suivant permet de se rassurer sur ce point (du moins en dimension finie).

Définition 1.2. Deux normes $\|\cdot\|_\alpha$ et $\|\cdot\|_\beta$ sur $\mathbb{R}^n$ sont dites ***équivalentes*** s'il existe des constants c, $C \in \mathbb{R}_+^*$ telles que

$$\forall x \in \mathbb{R}^n, \qquad c\|x\|_\alpha \leq \|x\|_\beta \leq C\|x\|_\alpha.$$

En reprenant les exemples précédents, il est relativement simple de se convaincre que deux normes équivalentes donnent des suites de mêmes natures, les mêmes fonctions continues et les mêmes comparaisons.

Question 1-7 Montrer que les normes 1, 2 et ∞ sont équivalentes.

Théorème 1.1. *Toutes les normes sur $\mathbb{R}^n$ sont équivalentes.*

Remarque 1. La preuve de ce théorème est en dehors du périmètre de ce cours. Elle nécessite certaines notions de topologie qui vous manquent ; notamment la notion de compacité. On se contentera du résultat.

Le fait que les normes soient équivalentes en terme de questionnement topologique (les études des phénomènes locaux) ne signifie pas qu'utiliser l'une ou l'autre en modélisation revient au même.

Question 1-8 Quelles normes utiliseriez-vous pour modéliser les problématiques suivantes :

1. Le calcul de la distance que doit parcourir un oiseau entre deux coordonnées GPS pas trop éloignées.
2. Le calcul de la distance que parcours un touriste à Manhattan entre deux musée.
3. Le calcul du nombre de modifications nécessaires (lettre à lettre) pour changer un mot en un autre.

2 Différentiabilité et différentielle en un point

3 Gradient en un point

4 Caractérisation du premier ordre de la convexité

5 Hessienne en un point

6 Caractérisation du second ordre de la convexité