

# IFT 3515

## Fonctions à plusieurs variables

### Optimisation sans contraintes

Fabian Bastin  
DIRO  
Université de Montréal

# Optimisation sans contrainte

Nous considérons le problème

$$\min_x f(x),$$

où  $f : Y \rightarrow \mathbb{R}$ ,  $f \in C^1$  sur  $Y$  et  $Y \subseteq \mathbb{R}^n$  est un ensemble ouvert.

Gradient:

$$\nabla_x f(x) = \begin{pmatrix} \frac{df(x)}{dx_1} \\ \frac{df(x)}{dx_2} \\ \vdots \\ \frac{df(x)}{dx_n} \end{pmatrix}$$

## Matrice hessienne

Si  $f$  est deux fois différentiable en  $x$ , la matrice hessienne  $\nabla_{xx}^2 f(x)$  s'écrit

$$\nabla_{xx}^2 f(x) = \begin{pmatrix} \frac{d^2 f(x)}{dx_1^2} & \frac{d^2 f(x)}{dx_1 dx_2} & \dots & \frac{d^2 f(x)}{dx_1 dx_n} \\ \frac{d^2 f(x)}{dx_2 dx_1} & \frac{d^2 f(x)}{dx_2^2} & \dots & \frac{d^2 f(x)}{dx_2 dx_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{d^2 f(x)}{dx_n dx_1} & \frac{d^2 f(x)}{dx_n dx_2} & \dots & \frac{d^2 f(x)}{dx_n^2} \end{pmatrix}.$$

Si  $f$  est deux fois continûment différentiable en  $x$ ,

$$\frac{d^2 f(x)}{dx_i dx_j} = \frac{d^2 f(x)}{dx_j dx_i}.$$

Autrement dit, dans ce cas  $\nabla_{xx}^2 f(x)$  est une matrice symétrique.

## Jacobien et matrice hessienne

Soit  $F : \mathcal{X} \rightarrow \mathbb{R}^m$ , où  $\mathcal{X} \subset \mathbb{R}^n$  est ouvert. Avec  $x = (x_1, \dots, x_n)$ , on peut définir  $F$  à partir de  $m$  fonctions composantes  $f_i(x) : \mathcal{X} \rightarrow \mathbb{R}$ ,  $i = 1, \dots, m$ :

$$F(x) = \begin{pmatrix} f_1(x) \\ \vdots \\ f_m(x) \end{pmatrix}.$$

La matrice jacobienne de  $F$  en  $x$  est

$$J(F(x)) = \begin{pmatrix} \nabla_x^T f_1(x) \\ \vdots \\ \nabla_x^T f_m(x) \end{pmatrix}.$$

Ceci permet d'écrire  $\nabla_{xx}^2 f(x)$  comme la matrice jacobienne du gradient de  $f$ :

$$\nabla_{xx}^2 f(x) = J(\nabla_x f(x)).$$

## Développement de Taylor

Si  $f \in C^1$  sur  $Y$ , alors  $\forall x, y \in Y$ ,

$$f(y) = f(x) + \nabla_x f(x)^T (y - x) + o(\|y - x\|).$$

Note: si  $y \neq x$ ,  $h(y) = o(\|y - x\|)$  comme  $y \rightarrow x$  signifie que

$$\lim_{y \rightarrow x} \frac{h(y)}{\|y - x\|} \rightarrow 0.$$

En d'autres termes,  $h(y)$  converge "plus vite" vers 0 que  $\|y - x\|$ .

Le développement linéaire de  $f$  autour de  $x$  est

$$\tilde{f}(y) = f(x) + \nabla_x f(x)^T (y - x)$$

## Développement de Taylor – version sans résidu

$\exists z \in \{t \in \mathbb{R}^n \mid t = \lambda x + (1 - \lambda)y, \lambda \in [0, 1]\}$  tel que

$$f(y) = f(x) + \nabla_x f(z)^T (y - x)$$

## Développement de Taylor au deuxième ordre

Si  $f \in C^1$  sur  $Y$ , alors  $\forall x, y \in Y$ ,

$$f(y) = f(x) + \nabla_x f(x)^T (y - x) + \frac{1}{2} (y - x)^T (y - x) \nabla_{xx}^2 f(x) o(\|y - x\|^2).$$

Forme explicite du reste (formule de Taylor avec reste intégral) quand  $f \in C^3$  sur un voisinage du segment joignant  $x$  à  $y$ :

$$f(x) = f(x) + \nabla_x f(x)^T (y - x) + \frac{1}{2} (y - x)^T \nabla_{xx}^2 f(x) (y - x) + R_2(y)$$

avec

$$R_2(y) = \frac{1}{2} \int_0^1 (1-t)^2 D^3 f(x + t(y-x)) [y-x, y-x, y-x] dt,$$

où  $D^3 f(z)$  est le tenseur des dérivées troisièmes de  $f$  en  $z$ , et  $D^3 f(z)[u, v, w]$  désigne son application trilinéaire à  $(u, v, w)$ .

## Formes alternatives

Version en coordonnées:

$$R_2(y) = \frac{1}{2} \sum_{i,j,k=1}^n (y_i - x_i)(y_j - x_j)(y_k - x_k) \int_0^1 (1-t)^2 \frac{\partial^3 f}{\partial x_i \partial x_j \partial x_k} (x + t(y-x)) dt.$$

Forme intégrale avec la matrice Hessienne:

$$R_2(y) = \int_0^1 (1-t) (y-x)^T [\nabla_{xx}^2 f(x + t(y-x)) - \nabla^2 f(x)] (y-x) dt$$

## Dérivée directionnelle

La dérivée de  $f$  au point  $x$  suivant le vecteur  $h$  est, si elle existe, la dérivée en 0 de la fonction à une variable définie par  $t \rightarrow f(x + th)$ , c'est-à-dire

$$D_h f(x) = \lim_{t \rightarrow 0} \frac{f(x + th) - f(x)}{t}$$

### Lemme

Soit  $f : Y \rightarrow \mathbb{R}$ , où  $Y$  est un sous-ensemble ouvert de  $\mathbb{R}^n$ ,  $f \in C^1$ ,  $x \in Y$  et  $h \in \mathbb{R}^n$ .

Alors

$$D_h f(x) = \nabla f(x)^T h.$$

## Preuve du lemme

Preuve.

Le développement de Taylor nous donne

$$f(x + th) = f(x) + \nabla f(x)^T th + o(\|th\|).$$

Dès lors, pour  $t \neq 0$ ,

$$\frac{f(x + th) - f(x)}{t} - \frac{o(\|th\|)}{t} = \nabla f(x)^T h$$

Le résultat suit en faisant tendre  $t$  vers 0. □

## Direction de descente

### Définition (Direction réalisable)

Soient  $X \subseteq \mathbb{R}^n$  et  $f : X \rightarrow \mathbb{R} \in C^1$ . Étant donné  $x \in X$ ,  $d \in \mathbb{R}^n$  est une direction réalisable en  $x$  s'il existe un scalaire  $\alpha_{\max}$  tel que  $x + \alpha d \in X$  pour tout  $\alpha \in [0, \alpha_{\max}]$ .

### Lemme (Direction de descente)

Soient  $X \subseteq \mathbb{R}^n$  et  $f : X \rightarrow \mathbb{R} \in C^1$ . Si  $d \in \mathbb{R}^n$  est une direction réalisable en  $x$  et  $\nabla f(x)^T d < 0$ , alors  $\exists \kappa > 0$  tel pour tout  $\tau \in (0, \kappa]$

$$f(x + \tau d) < f(x)$$

(i.e.,  $d$  est une direction de descente en  $x$ ).

## Preuve du lemme

Preuve.

Puisque  $D_d f(x) = \nabla f(x)^T d < 0$ , il existe un scalaire  $\kappa > 0$  tel que pour tout  $\tau \neq 0$ ,  $-\kappa \leq \tau \leq \kappa$ ,

$$\frac{f(x + \tau d) - f(x)}{\tau} < 0$$

Il suffit de restreindre  $\tau$  à des valeurs strictement positives pour avoir

$$f(x + \tau d) < f(x).$$



## Condition nécessaire d'optimalité

### Théorème (Optimalité au premier ordre)

Soient  $X \subseteq \mathbb{R}^n$  et  $f : X \rightarrow \mathbb{R} \in C^1$ . Si  $x^*$  est un minimum local de  $f$  sur  $X$ , alors pour toute direction  $d \in \mathbb{R}^n$  qui est une direction réalisable en  $x^*$ , nous avons  $\nabla f(x^*)^T d \geq 0$ .

### Preuve.

Comme  $d$  est une direction réalisable, il existe  $\alpha_{\max}$  tel que pour tout  $\alpha \in [0, \alpha_{\max}]$ , le point  $x(\alpha) = x^* + \alpha d \in X$ .

Définissons  $g(\alpha) = f(x(\alpha))$ ,  $0 \leq \alpha \leq \alpha_{\max}$ .  $g$  a un minimum local en  $\alpha = 0$ , et en vertu du développement de Taylor autour de 0,

$$g(\alpha) - g(0) = g'(0)\alpha + o(\alpha).$$

Si  $g'(0) < 0$ , pour  $\alpha$  suffisamment petit,  $g(\alpha) < g(0)$ , aussi 0 ne serait pas minimum local.

## Condition nécessaire d'optimalité (suite)

Preuve.

Dès lors,  $g'(0) \geq 0$ .

Or,

$$g'(\alpha) = \nabla f(x(\alpha))^T \frac{d}{d\alpha} x(\alpha) = \nabla f(x(\alpha))^T d.$$

Ainsi,  $\nabla f(x^*)^T d \geq 0$ . □

Corollaire

Soient  $X \subseteq \mathbb{R}^n$  un ensemble ouvert et  $f : X \rightarrow \mathbb{R} \in C^1$ . Si  $x^*$  est un minimum local de  $f$  sur  $X$ , alors  $\nabla f(x^*) = 0$ .

## Condition nécessaire d'optimalité

Preuve.

Du théorème précédent,  $\forall d \in \mathbb{R}^n$ ,  $\nabla f(x^*)^T d \geq 0$ . □

Preuve alternative.

Par contradiction, supposons que  $x$  est un minimum local et que  $\nabla f(x) \neq 0$ .

Considérons la direction  $d = -\nabla f(x)$ . Puisque  $X$  est ouvert, alors  $d$  est une direction réalisable en  $x$ , et

$$\nabla f(x)^T d = -\|\nabla f(x)\|^2 < 0,$$

puisque  $\nabla f(x) \neq 0$ . Dès lors,  $d$  est une direction de descente en  $x$  et par conséquent il est possible de déterminer un scalaire  $\tau > 0$  tel que  $x + \tau d \in \mathcal{B}(x, \epsilon)$  et  $f(x + \tau d) < f(x)$ , et donc  $x$  n'est pas un minimum local. □

## Condition d'optimalité au deuxième ordre

Comme dans le cas unidimensionnel, les conditions au premier ordre sont rencontrées en un point stationnaire (minimum, maximum, point selle). Nous voulons donc avoir plus pour garantir qu'un point stationnaire est bien un minimum.

# Matrice (semi-)définie positive

## Définition (Matrice semi-définie positive)

*Une matrice symétrique  $A \in \mathbb{R}^{n \times n}$  est semi-définie positive si  $\forall x \in \mathbb{R}^n$ ,  $x^T Ax \geq 0$ .*

## Définition (Matrice définie positive)

*Une matrice symétrique  $A \in \mathbb{R}^{n \times n}$  est définie positive si  $\forall x \in \mathbb{R}^n$ ,  $x \neq 0$ ,  $x^T Ax > 0$ .*

## Définition (Matrice semi-définie négative)

*Une matrice symétrique  $A \in \mathbb{R}^{n \times n}$  est semi-définie négative si  $\forall x \in \mathbb{R}^n$ ,  $x^T Ax \leq 0$ .*

## Définition (Matrice définie négative)

*Une matrice symétrique  $A \in \mathbb{R}^{n \times n}$  est définie négative si  $\forall x \in \mathbb{R}^n$ ,  $x \neq 0$ ,  $x^T Ax < 0$ .*

# Matrice (semi-)définie positive

## Théorème

*Une matrice symétrique  $A$  est semi-définie positive (définie positive) si et seulement si une des conditions suivantes tient*

1.  $\forall x \in \mathbb{R}^n$  (et  $x \neq 0$ ),  $x^T Ax \geq 0$  ( $> 0$ )
2. toutes les valeurs propres sont non-négatives (positives)
3. les déterminants de tous les mineurs principaux sont non-négatifs (positifs)

## Mineurs principaux

Considérons une matrice  $A \in \mathbb{R}^{n \times n}$ .  $B_i$  est le  $i^e$  mineur principal de  $A$  si  $B$  est une matrice de dimension  $i \times i$ , et  $B(p, q) = A(p, q)$ ,  $p = 1, \dots, i$ ,  $q = 1, \dots, i$ .

$$\left( \begin{array}{ccccc} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \end{array} \right)$$

## Notations

Soit la matrice symétrique  $A \in \mathbb{R}^{n \times n}$ .

- $A \succ 0$ :  $A$  est définie positive.
- $A \succeq 0$ :  $A$  est semi-définie positive.
- $A \prec 0$ :  $A$  est définie négative.
- $A \preceq 0$ :  $A$  est semi-définie négative.

## Conditions nécessaires de deuxième ordre

### Lemme

Soient  $X \subseteq \mathbb{R}^n$  un ensemble ouvert et  $f : X \rightarrow \mathbb{R} \in C^2$ . Si  $x \in X$  est un minimum local de  $f$  sur  $X$ , alors  $\nabla f(x) = 0$  et  $\nabla^2 f(x)$  est une matrice semi-définie positive.

### Preuve.

Nous avons déjà prouvé que  $\nabla f(x) = 0$  si  $x$  est un minimum local. Concentrons-nous dès lors sur le caractère semi-défini positif. Le développement de Taylor au deuxième ordre donne

$$f(x + \tau d) = f(x) + \tau \nabla f(x)^T d + \frac{\tau^2}{2} d^T \nabla^2 f(x) d + o(\|\tau d\|^2)$$

Comme  $\nabla f(x) = 0$ , nous avons

$$f(x + \tau d) - f(x) = \tau^2 d^T \nabla^2 f(x) d + o(\|\tau d\|^2)$$

## Conditions nécessaires de deuxième ordre

Preuve.

Supposons par contradiction que  $\nabla^2 f(x)$  n'est pas semi-définie positive, autrement dit  $\exists y \in \mathbb{R}^n$  tel que  $y^T \nabla^2 f(x) y < 0$ .

Pour  $\tau > 0$  suffisamment petit,  $x + \tau y \in \mathcal{B}(x, \epsilon)$  et

$$\tau^2 y^T \nabla^2 f(x) y + o(\|\tau y\|^2) < 0$$

Dès lors,

$$f(x + \tau y) - f(x) < 0$$

et  $x$  n'est pas un minimum local.



## Conditions non suffisantes

Les conditions que  $\nabla f(x) = 0$  et que  $\nabla^2 f(x)$  est une matrice semi-définie positive ne sont pas suffisantes pour assurer que  $x$  est un minimum local.

Exemple:  $f(x, y) = x^3 + y^3$ . Nous avons

$$\nabla f(x, y) = \begin{pmatrix} 3x^2 \\ 3y^2 \end{pmatrix} \quad \nabla^2 f(x, y) = \begin{pmatrix} 6x & 0 \\ 0 & 6y \end{pmatrix}$$

En  $(0,0)$ , la fonction, son gradient et la matrice hessienne s'annulent, et il est évident qu'une matrice carrée nulle est semi-définie positive.

Pour  $\epsilon > 0$  suffisamment petit,  $(-\epsilon/2, -\epsilon/2) \in \mathcal{B}(x, \epsilon)$  et  $f(-\epsilon/2, -\epsilon/2) < 0$ .

## Conditions suffisantes de second ordre

### Théorème

Soient  $X \subseteq \mathbb{R}^n$  un ensemble ouvert et  $f \in C^2$  sur  $X$ . Si  $\nabla f(x^*) = 0$  et  $\nabla^2 f(x^*)$  est une matrice définie positive, alors il existe un  $\epsilon > 0$  suffisamment petit tel que  $f(x^*) < f(x)$  pour tout  $x \in \mathcal{B}(x^*, \epsilon)$ .

### Preuve.

En utilisant le développement de Taylor à l'ordre 2

$$f(x^* + \tau d) = f(x^*) + \tau \nabla f(x^*)^T d + \frac{\tau^2}{2} d^T \nabla^2 f(x^*) d + o(\|\tau d\|^2),$$

nous pouvons écrire, puisque  $\nabla f(x^*) = 0$ ,

$$f(x^* + \tau d) - f(x^*) = \frac{\tau^2}{2} d^T \nabla^2 f(x^*) d + o(\|\tau d\|^2)$$

## Conditions suffisantes de second ordre

Preuve.

Si  $A$  est symétrique, il est possible de montrer que

$$\lambda_{\min} \leq \frac{d^T A d}{\|d\|^2} \leq \lambda_{\max}$$

où  $\lambda_{\min}$  ( $\lambda_{\max}$ ) est la plus petite (grande) valeur propre de  $A$ . On peut en effet effectuer la décomposition spectrale  $A = Q\Lambda Q^T$  où

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}, \quad Q = (\mathbf{u}_1 \quad \cdots \quad \mathbf{u}_n),$$

tels que  $A\mathbf{u}_i = \lambda_i \mathbf{u}_i$ ,  $\|\mathbf{u}_i\| = 1$ ,  $QQ^T = I$ .

## Conditions suffisantes de second ordre

Preuve.

Avec  $y := Q^T d$ , nous obtenons

$$d^T A d = y^T \Lambda y = \sum_{i=1}^n \lambda_i y_i^2.$$

Dès lors

$$d^T A d \geq \lambda_{\min} \sum_{i=1}^n y_i^2 = \lambda_{\min} \|y\|^2 = \lambda_{\min} \|d\|^2$$

$$d^T A d \leq \lambda_{\max} \sum_{i=1}^n y_i^2 = \lambda_{\max} \|y\|^2 = \lambda_{\max} \|d\|^2$$

Il suit

$$d^T \nabla^2 f(x) d \geq \lambda_{\min} \|d\|^2.$$

## Conditions suffisantes de second ordre

Preuve.

Ainsi,

$$\begin{aligned} f(x^* + \tau d) - f(x^*) &= \frac{\tau^2}{2} d^T \nabla^2 f(x^*) d + o(\|\tau d\|^2) \\ &\geq \frac{\tau^2}{2} \lambda_{\min} \|d\|^2 + o(\|\tau d\|^2) \end{aligned}$$

Il suit que pour  $\tau > 0$  suffisamment petit,  $x^* + \tau d \in \mathcal{B}(x^*, \epsilon)$  et  
 $f(x^* + \tau d) - f(x^*) > 0$  ou  $f(x^* + \tau d) > f(x^*)$ . □

Corollaire

Soient  $X \subseteq \mathbb{R}^n$  un ensemble ouvert et  $f \in C^2$  sur  $X$ . Si  $\nabla f(x^*) = 0$  et  $\nabla^2 f(x^*)$  est une matrice définie positive, alors  $x^*$  est un minimum local de  $f$  sur  $X$ .

## Fonction convexe

Soit  $X \subseteq \mathbb{R}^n$  un ensemble convexe et  $f : X \rightarrow \mathbb{R}$ .

$f$  est dite **convexe** si  $\forall x_1, x_2 \in X, \forall t \in [0, 1]$ :

$$f(tx_1 + (1 - t)x_2) \leq tf(x_1) + (1 - t)f(x_2).$$

$f$  est dite **strictement convexe** si  $\forall x_1, x_2 \in X, x_1 \neq x_2, \forall t \in (0, 1)$ :

$$f(tx_1 + (1 - t)x_2) < tf(x_1) + (1 - t)f(x_2).$$

Une fonction  $f$  est dite **(strictement) concave** si  $-f$  est (strictement) convexe.

## Propriétés

### Théorème

*Une fonction  $f : X \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$  est convexe si et seulement si la fonction  $g : \mathbb{R} \rightarrow \mathbb{R}$  donnée par*

$$g(\alpha) = f(x + \alpha y)$$

*est convexe en tout  $x + \alpha y \in X$ .*

L'intérêt de ce résultat en optimisation est que si une fonction est convexe, elle le sera également le long de toute direction de recherche.

## Caractérisations équivalentes

### Théorème

Soit  $f : X \subseteq \mathbb{R}^n$ ,  $f \in C^2$ ,  $X$  ouvert. Les affirmations suivantes sont équivalentes

- (i)  $f$  est convexe;
- (ii)  $\forall x, y \in X, f(y) \geq f(x) + \nabla f(x)^T(y - x);$
- (iii)  $\forall x \in X, \nabla^2 f(x) \succeq 0.$

## Convexité et fonctions quadratique

Considérons la fonction quadratique

$$f(x) = x^T A x + b^T x + c.$$

$f(\cdot)$  est

- convexe si et seulement si  $A \succeq 0$ ;
- strictement convexe si et seulement si  $A \succ 0$ ;
- concave si et seulement si  $A \preceq 0$ ;
- strictement concave si et seulement si  $A \prec 0$ .

## Fonction convexe – minimisation globale

### Théorème

*Soit le problème d'optimisation sans contrainte*

$$\min_x f(x),$$

*où  $f \in C^1$  est convexe. Alors, tout point  $x^*$  satisfaisant  $\nabla f(x^*) = 0$  est un minimum global.*

### Preuve.

Comme  $f$  est convexe et dans  $C^1$ , nous avons  $\forall x, y$ ,

$$f(y) \geq f(x) + \nabla f(x)^T (y - x).$$

En particulier,  $\forall y$ ,

$$f(y) \geq f(x^*) + \nabla f(x^*)^T (y - x^*).$$

et comme  $\nabla f(x^*) = 0$ ,  $f(y) \geq f(x^*)$ .

## Convexité stricte

Soit  $f : X \subseteq \mathbb{R}^n$ ,  $f \in C^2$ ,  $X$  ouvert.  $f$  est strictement convexe ssi

$$\forall x, y \in X \text{ tels que } x \neq y, f(y) > f(x) + \nabla f(x)^T(y - x).$$

Si  $\forall x \in X$ ,  $\nabla^2 f(x) \succ 0$ , alors  $f$  est strictement convexe. Mais l'inverse n'est pas vrai!  
Exemple:  $f(x) = \sum_{i=1}^4 x_i^4$ .

## Convexité stricté et unicité de la solution

### Théorème

*Soit le problème d'optimisation sans contrainte*

$$\min_{x \in X} f(x),$$

*où  $f \in C^1$  est strictement convexe et  $X$  est un ensemble convexe. La solution optimale, si elle existe, est unique.*

## Convexité stricté et unicité de la solution

Preuve.

Supposons par l'absurde qu'il existe deux solutions optimales différentes  $x_1^*$  et  $x_2^*$ .

Prenons

$$z = \frac{x_1^* + x_2^*}{2}$$

En vertu de la convexité stricte,

$$f(z) < \frac{1}{2}f(x_1^*) + \frac{1}{2}f(x_2^*) = f(x_1^*).$$



## Algorithme

Un algorithme défini sur un sous-ensemble  $X \subseteq \mathbb{R}^n$  est un processus itératif qui, partant d'une solution initiale  $x_0 \in X$ , génère une suite de points  $\{x_k\}$  dans  $X$ .

À un algorithme correspond une multi application (multi function ou “mapping”)  $A : X \rightarrow X$  associant à un point  $x_k \in X$  un sous-ensemble  $A(x_k) \subseteq X$ .

Exemple:

$$A(x) = \left\{ \left[0, \frac{x_1}{2}\right], \left[0, \frac{x_2}{2}\right] \right\}.$$

Dénotons par  $X^* \subseteq X$  l'ensemble des solutions recherchées.

## Algorithme de descente

### Définition (Algorithme de descente)

Un algorithme  $A$  est un algorithme de descente par rapport à une fonction  $z : X \rightarrow \mathbb{R}$  continue si

1.  $x \notin X^*$  et  $y \in A(x) \Rightarrow z(y) < z(x)$
2.  $x \in X^*$  et  $y \in A(x) \Rightarrow z(y) \leq z(x)$ .

### Définition

Un algorithme  $A$  est fermé au point  $x \in X$  si

1.  $\{x_k\} \in X$  a la propriété que  $\lim_{k \rightarrow \infty} x_k = x$
2.  $\{y_k \in A(x_k)\}$  a la propriété que  $\lim_{k \rightarrow \infty} y_k = y$

alors  $y \in A(x)$ .

Note: la notion de fermeture pour les multi applications correspond à celle de continuité pour les fonctions.

## Point d'accumulation et ensemble fermé

Soit  $X \subseteq \mathbb{R}^n$ .

### Définition (Point d'accumulation)

Considérons une suite de points  $x_k \in X$ ,  $k = 1, 2, \dots$  telle que  $\lim_{k \rightarrow \infty} x_k = x$ .  $x$  est un point d'accumulation (point limite) de  $X$ .

### Définition (Ensemble fermé.)

$X$  est un ensemble fermé si tout point d'accumulation de  $X$  appartient à  $X$ .

Par exemple, dans  $\mathbb{R}^2$  l'ensemble

$$\Gamma = \{[x, y] \in \mathbb{R}^2 \mid x \in [a, b], y = 0\}$$

est un ensemble fermé.

## Ensemble ouvert, borné, compact

### Définition (Ensemble ouvert)

Soit  $X \subseteq \mathbb{R}^n$ .  $X$  est un ensemble ouvert si  $\forall x \in X$ , alors la boule  $B(x, \epsilon) = \{z \in R^n \mid \|x - z\| < \epsilon\} \subseteq X$ .

(Contre)-exemple: dans  $R^2$  l'ensemble

$$\Gamma = \{[x, y] \in \mathbb{R}^2 \mid x \in (a, b), y = 0\}$$

n'est pas un ensemble ouvert puisque nous ne pouvons pas modifier la valeur de  $y$ .

### Définition (Ensemble borné)

Un ensemble  $X$  est borné si  $\exists M$  t.q.  $\forall x \in X$ ,  $\|x\| \leq M$ .

### Définition (Ensemble compact)

Dans  $\mathbb{R}^n$ , un ensemble  $X$  est compact si et seulement si il est fermé et borné.

## Théorème de convergence globale

Supposons que  $X \subseteq \mathbb{R}^n$  est un ensemble non vide et fermé, et soit  $X^*$  l'ensemble non vide des solutions. Soit  $A$  un algorithme défini sur  $X$  qui, partant d'un point  $x_0 \in X$ , génère une suite de points  $\{x_k\}$  comme suit:

- si  $x_k \in X^*$ , l'algorithme s'arrête,
- sinon, soit  $x_{k+1} \in A(x_k)$ . Poser  $k \leftarrow k + 1$  et recommencer.

Supposons que la suite de points  $\{x_k\}$  générés par l'algorithme est contenue dans un sous ensemble compact de  $X$  et qu'il existe une fonction continue  $z$  par rapport à laquelle  $A$  est un algorithme de descente. Sous ces conditions, si la multi application  $A$  est fermée aux points appartenant à  $X^*$ , alors

1. soit l'algorithme s'arrête après un nombre fini d'itérations à un point de  $X^*$ ,
2. soit il génère une suite infinie de points  $\{x_k\}$  telle que tout point d'accumulation de  $\{x_k\}$  (c'est-à-dire tout point limite d'une sous-suite convergente de  $\{x_k\}$ ) appartient à  $X^*$ .

## Descente par coordonnées (coordinate descent)

Soit  $f : Y \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ .  $f$  est minimisée pour chaque coordonnée séquentiellement.  
Ainsi, les directions utilisées en séquence sont

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \dots, \quad e_n = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

À chaque fois que  $n$  itérations sont complétées, la séquence est reprise.

# Méthode de Gauss-Seidel

**Étape 0** Soit  $f$  la fonction à minimiser, et  $\delta > 0$  le niveau de tolérance. Poser  $k = 0$  et choisir une solution initiale  $x^0$ .

**Étape 1** 1. Poser  $y_0 := x_k$ .

2. Pour  $j = 1, \dots, n$ ,

- calculer  $\alpha^* \in \arg \min_{\alpha} f(y_{j-1} + \alpha e_j)$ .
- Poser  $y_j = y_{j-1} + \alpha_j e_j$ .

3. Poser  $x_{k+1} = y_n$ .

**Étape 2** Arrêt si  $\|x_{k+1} - x_k\| < \delta$  et poser  $x^* := x_{k+1}$ .

Sinon, poser  $k := k + 1$  et retourner à l'étape 1.

# Méthode de Jacobi

**Étape 0** Soit  $f$  la fonction à minimiser, et  $\delta > 0$  le niveau de tolérance. Poser  $k = 0$  et choisir une solution initiale  $x_0$ .

**Étape 1**

1. Pour  $j = 1, \dots, n$ , calculer  $\alpha_j^* \in \arg \min_{\alpha} f(x_k + \alpha e_j)$ .
2. Soit  $\alpha^* = (\alpha_1^*, \dots, \alpha_n^*)$ . Poser  $x_{k+1} = x_k + \alpha^*$ .

**Étape 2** Arrêt si  $\|x_{k+1} - x_k\| < \delta$  et poser  $x^* := x_{k+1}$ .  
Sinon, poser  $k := k + 1$  et retourner à l'étape 1.

## Méthode de recherche linéaire

Une méthode de recherche linéaire pour optimiser une fonction  $f : R^n \rightarrow R$  génère une suite de points  $\{x_k\}$  où  $x_0$  est un point initial choisi dans  $\mathbb{R}^n$  et où

$$x_{k+1} = x_k + \alpha_k d_k$$

i.e.  $x_{k+1}$  est généré à partir de  $x_k$  en choisissant une direction  $d_k \in R^n$  et en prenant un pas  $\alpha_k \in \mathbb{R}$  dans cette direction pour s'éloigner de  $x_k$ . Les méthodes diffèrent par leurs choix de direction  $d_k$  et de pas  $\alpha_k$ .

## Méthode de plus forte pente

Supposons qu'à chaque itération  $k$ ,  $d_k$  soit choisi comme une direction de descente. Dès lors, à une itération  $k$ , si  $\nabla f(x_k) \neq 0$ , il est possible de trouver  $\alpha_k$

$$f(x_k + \alpha_k d_k) < f(x_k).$$

Supposons que  $f(x)$  soit bornée inférieurement sur  $Y$  par  $\kappa$ . Dès lors  $f(x_k) > f(x_{k+1}) > \dots \geq \kappa$ .

### Théorème

*Toute fonction monotonement décroissante et minorée converge.*

En d'autres termes,  $f(x_k)$  converge vers une valeur  $\bar{f}$ . Mais il se peut que  $\bar{f} > f^*$ !  $d_k$  et  $\alpha_k$  doivent être choisis adéquatement pour avoir  $\bar{f} = f^*$

## Méthode de plus forte pente

Cherchons la direction qui fait décroître  $f$  le plus vite localement.

Le développement linéaire de  $f$  autour de  $x$  était

$$f(y) = f(x) + \nabla f(x)^T (y - x) + o(\|y - x\|)$$

et donc

$$f(x_k + \alpha d_k) - f(x_k) = \alpha \nabla f(x_k)^T d_k + o(\|\alpha d_k\|)$$

L'inégalité de Cauchy-Schwarz nous indique que

$$|\nabla f(x_k)^T d_k| \leq \|\nabla f(x_k)\| \|d_k\|$$

et l'égalité est atteinte avec  $d_k = \pm \nabla f(x_k)$ . En faisant tendre  $\alpha$  vers 0, on voit que

$$|f(x_k + \alpha d_k) - f(x_k)| \approx \alpha |\nabla f(x_k)^T d_k|$$

et donc localement, la plus forte décroissance est atteinte en prenant  $d_k = -\nabla f(x_k)$ :  
**plus forte pente.**

## Méthode de plus forte pente (méthode du gradient)

Étape 0 Soit  $f \in C^1$  la fonction à minimiser, et  $\delta > 0$  le niveau de tolérance.

Poser  $k = 0$  et choisir une solution initiale  $x_0$ .

Étape 1 Arrêt si  $\|\nabla f(x_k)\| < \delta$  et poser  $x^* := x_k$ .

Étape 1 Calculer  $\alpha^* \in \arg \min_{\alpha \geq 0} f(x_k - \alpha \nabla f(x_k))$ .

Poser  $x_{k+1} = x_k - \alpha^* \nabla f(x^k)$ ,  $k \leftarrow k + 1$ .

Retour à l'étape 1.

## Méthode de plus forte pente: convergence

Posons  $\delta = 0$  plutôt que  $\delta > 0$ .

### Théorème

Soit  $f \in C^1$ . Tout point d'accumulation  $x^*$  d'une sous-suite convergente de la suite générée par la méthode du gradient est tel que  $\nabla f(x^*) = 0$ .

### Preuve.

Considérons une sous-suite convergente  $\{x_{k_j}\}$  de  $\{x_k\}$  telle que  $x_{k_j} \rightarrow x^*$  quand  $k_j \rightarrow \infty$ . Dès lors

$$f(x_{k_{j+1}}) \leq f(x_{k_j+1}) = \min_{\alpha \geq 0} f(x_{k_j} - \alpha \nabla f(x_{k_j}))$$

et donc,  $\forall \alpha \geq 0$ ,

$$f(x_{k_{j+1}}) \leq f(x_{k_j} - \alpha \nabla f(x_{k_j}))$$

## Méthode de plus forte pente: convergence

Preuve.

Comme  $f \in C^1$ ,  $f$  et  $\nabla f$  sont continus sur  $\mathbb{R}^n$ . En outre, comme

$$\lim_{j \rightarrow \infty} f(x_{k_j}) = f(x^*),$$

nous avons,  $\forall \alpha \geq 0$ ,

$$\lim_{j \rightarrow \infty} f(x_{k_j} - \alpha \nabla f(x_{k_j})) = f(x^* - \alpha \nabla f(x^*))$$

De là,

$$f(x^*) \leq f(x^* - \alpha \nabla f(x^*)), \quad \forall \alpha \geq 0.$$

Si  $\nabla f(x^*) \neq 0$ ,  $-\nabla f(x^*)$  est une direction de descente, et cette inégalité n'est pas valide. On en déduit  $\nabla f(x^*) = 0$ .



## Directions orthogonales

Lemme (Déplacement en zig-zag)

*Si  $x_{k+1} = x_k - \alpha_k \nabla f(x_k)$  où  $\alpha_k$  est tel que*

$$f(x_k) - \alpha_k \nabla f(x_k) = \min_{\alpha \geq 0} f(x_k - \alpha \nabla f(x_k))$$

*alors  $\nabla f(x_{k+1})^T \nabla f(x_k) = 0$ .*

Preuve.

Si  $\nabla f(x_k) = 0$ , alors le résultat tient immédiatement. Supposons donc  $\nabla f(x_k) \neq 0$ . Considérons la fonction

$$m : \mathbb{R}^+ \rightarrow \mathbb{R}, \quad m(\alpha) = f(x_k - \alpha \nabla f(x_k)).$$

Dès lors,

$$m'(\alpha) = -\nabla f(x_k)^T \nabla f(x_k - \alpha \nabla f(x_k))$$

## Directions orthogonales

Preuve.

Or,  $\alpha_k$  est choisi de sorte à minimiser  $m(\alpha)$ , aussi  $m'(\alpha_k) = 0$ . De là,

$$0 = -\nabla f(x_k)^T \nabla f(x_{k+1})$$



Notes:

1. L'ordre de convergence de la méthode du gradient est linéaire.
2. Au début de son application, la méthode permet de faire diminuer la valeur de la fonction économique relativement rapidement, mais son efficacité à ce chapitre diminue au cours des itérations.

## Calcul du pas

Nous pouvons utiliser les méthodes vues précédemment dans le cadre unidimensionnel pour calculer

$$\min_{\alpha \geq 0} f(x_k - \alpha \nabla f(x_k))$$

mais il est difficile de vérifier a priori si les hypothèses de ces méthodes sont satisfaites.

Mais avons-nous besoin de calculer le minimum exact le long de  $-\nabla f(x_k)$ ?

Nous avons plus généralement à considérer le problème de recherche linéaire qui consiste à déterminer une longueur de pas adaptée étant donnée une direction de recherche  $d_k$ .

## Méthode de recherche linéaire

1. Choisir un point initial  $x_0$  et poser  $k = 0$ .
2. Arrêt si la convergence est atteinte (habituellement en testant si  $\|\nabla f(x_k)\| < \delta$ , avec  $\delta > 0$  petit).
3. Calculer une direction de recherche  $d_k$  à partir de  $x_k$ , telle que  $d_k$  est une direction de descente.
4. Calculer une longueur de pas adéquate  $\alpha_k > 0$  telle que

$$f(x_k + \alpha_k d_k) < f(x_k).$$

Le calcul de  $\alpha_k$  est appelé recherche linéaire, et est habituellement l'objet d'une boucle de calcul itérative interne.

5. Poser

$$x_{k+1} = x_k + \alpha_k d_k.$$

Incrémenter  $k$  et retourner en 2.

## Calcul de la longueur de pas

Le défi dans le calcul de  $\alpha_k$  est d'éviter que le pas soit trop long ou trop court.

**Recherche linéaire exacte:** nous cherchons  $\alpha_k$  solution du programme

$$\min_{\alpha \geq 0} f(x_k + \alpha d_k)$$

L'approche est cependant souvent considérée comme trop coûteuse.

**Recherche linéaire inexacte**

1. formuler un critère qui assure que les pas ne sont ni trop longs ni trop courts
2. choisir une bonne longueur de pas initial
3. construire une séquence de mises à jour qui satisfont les critères précédents après quelques étapes

## Recherche linéaire par “backtracking”

Une première approche consiste à fixer une valeur maximale de la longueur du pas, et de réduire celle-ci jusqu'à obtenir une réduction de la fonction.

1. Soit  $\alpha_{\text{init}} > 0$ . Poser  $\alpha_{(0)} = \alpha_{\text{init}}$  et  $\ell = 1$ .
2. Tant que  $f(x_k + \alpha_{(\ell)} d_k) \geq f(x_k)$ , répéter
  - poser  $\alpha_{(\ell+1)} = \kappa \alpha_{(\ell)}$ , avec  $\kappa \in (0, 1)$  fixé,
  - incrémenter  $\ell$  par 1.
3. Poser  $\alpha_k = \alpha_{(\ell)}$ .

Cette méthode permet d'éviter que le pas ne devienne trop petit, mais pas qu'il soit trop long relativement à la décroissance de  $f$ , et ne garantit pas une décroissance significative de  $f$ . Afin d'améliorer la méthode, nous devons renforcer l'exigence

$$f(x_k + \alpha_{(\ell)} d_k) < f(x_k).$$

## Condition d'Armijo

La **condition d'Armijo** impose que nous obtenions au moins une fraction  $\beta$  de la réduction prédictive par le développement de Taylor au premier ordre de  $f$  en  $x_k$ :

$$f(x_k + \alpha_k d_k) \leq f(x_k) + \alpha_k \beta \nabla f(x_k)^T d_k$$

pour un certain  $\beta \in (0, 1)$  fixé (par exemple  $\beta = 0.1$  ou même  $\beta = 0.0001$ ).

Ceci aura aussi pour effet de prévenir des pas trop grands.

La procédure de backtracking devient

1. Soit  $\alpha_{\text{init}} > 0$ . Poser  $\alpha_{(0)} = \alpha_{\text{init}}$  et  $\ell = 1$ .
2. Tant que  $f(x_k + \alpha_k d_k) > f(x_k) + \alpha_k \beta \nabla f(x_k)^T d_k$ , répéter
  - poser  $\alpha_{(\ell+1)} = \kappa \alpha_{(\ell)}$ , avec  $\kappa \in (0, 1)$  fixé,
  - incrémenter  $\ell$  par 1.
3. Poser  $\alpha_k = \alpha_{(\ell)}$ .

## Convergence du backtracking-Armijo

### Théorème (Terminaison du backtracking-Armijo)

Soit  $f(x) \in C^1$ , et  $\nabla f(x)$  lipschitzienne autour de  $x_k$  avec une constante  $\gamma_k$ . Soit  $d_k$  une direction de descente en  $x_k$ . Alors, pour un  $\beta \in (0, 1)$  fixé,

1. la condition d'Armijo  $f(x_k + \alpha d_k) \leq f(x_k) + \alpha\beta\nabla f(x_k)^T d_k$  est satisfaite pour tout  $\alpha \in [0, a_k^{\max}]$ , où

$$a_k^{\max} = \frac{2(\beta - 1)\nabla f(x_k)^T d_k}{\gamma_k \|d_k\|^2}$$

2. de plus, pour  $\kappa \in (0, 1)$  fixé, la longueur de pas générée par la procédure de recherche linéaire backtracking-Armijo se termine avec

$$\alpha_k \geq \min \left\{ \alpha_{init}, \frac{2\kappa(\beta - 1)\nabla f(x_k)^T d_k}{\gamma_k \|d_k\|^2} \right\}$$

## Remarque

En pratique,  $\gamma_k$  n'est pas connu. Dès lors, nous ne pouvons pas simplement calculer  $\alpha_k^{\max}$  et  $\alpha_k$  à partir des formules explicites données par le théorème.

La condition d'Armijo ne suffit pas à elle seule comme elle ne prévient pas des pas trop petits. Considérons par exemple la fonction  $f(x) = x^2$  et  $x_0 = 2$ . Prenons  $d_k = -1$  et  $\alpha_k = 2^{-k-1}$ . Dans ce cas  $x_k \rightarrow 1$ . Pourtant, la condition d'Armijo est satisfaite pour  $c_1$  choisi suffisamment petit.

# Convergence de la recherche linéaire avec Armijo

## Théorème

*Supposons que  $f(x) \in C^1$  et que  $\nabla f(x)$  est lipschitzienne sur  $\mathbb{R}^n$ . Alors, pour les itérés générés par la méthode de recherche linéaire avec pas d'Armijo par backtracking, une des trois situations suivantes a lieu:*

1.  $\nabla f_k(x) = 0$  après un nombre fini d'itérations
2.  $\lim_{k \rightarrow \infty} f(x_k) = -\infty$
3.  $\lim_{k \rightarrow \infty} \nabla f(x_k) = 0$

## Condition de Wolfe

La condition d'Armijo n'est toutefois pas toujours suffisante pour assurer un bon comportement de l'algorithme de recherche linéaire.

Une longueur de pas  $\alpha_k$  est dite satisfaire les conditions de Wolfe par rapport à la direction de descente  $d_k$  si les inégalités suivantes sont satisfaites:

$$f(x_k + \alpha_k d_k) \leq f(x_k) + \alpha_k \beta_1 \nabla f(x_k)^T d_k \quad (\text{condition d'Armijo})$$

$$\nabla f(x_k + \alpha_k d_k)^T d_k \geq \beta_2 \nabla f(x_k)^T d_k \quad (\text{condition de courbure})$$

avec  $0 < \beta_1 < \beta_2 < 1$ .

La condition d'Armijo assure une réduction suffisante de la fonction objectif, tandis que la condition de courbure assure une réduction suffisante de la pente.