

## MODULE B6

## Éclatement de variables

Sauf mention contraire,  $\mathcal{X}$  (resp.  $\mathcal{Y}$ ) est un espace de HILBERT, muni du produit scalaire noté  $\langle \cdot, \cdot \rangle$  et on note  $\| \cdot \|$  la norme qui découle du produit scalaire.

Dans ce module, on s'intéresse au problème de minimisation suivant

$$\min_{x=(x_{(1)}, \dots, x_{(n)}) \in \mathcal{X}_1 \times \dots \times \mathcal{X}_n} J(x_{(1)}, \dots, x_{(n)}) \quad (\mathcal{P})$$

où  $J : \mathcal{X}_1 \times \dots \times \mathcal{X}_n \rightarrow \mathbb{R} \cup \{+\infty\}$  est une fonction s.c.i. de domaine non vide. On supposera que ce problème admet au moins une solution.

## 1 Principe et observations

Dans les applications actuelles de l'optimisation continue, on est amené à gérer des problèmes de dimension toujours plus grande. La dimension du problème s'entend ici comme la dimension de l'espace  $\mathcal{X}$  sur lequel la fonction objectif est optimisée<sup>1</sup>. Dans ces conditions, minimiser directement sur  $\mathcal{X}$  peut s'avérer techniquement difficile, voire impossible, ou très coûteux en calculs. L'idée est donc, si le problème le permet, de réduire la dimension du problème en le scindant en plusieurs problèmes de dimensions plus petites.

Une stratégie classique et intuitive, que l'on va explorer dans ce module, est de considérer une stratégie itérative, où à chaque itération, on "gèle" une partie de l'espace  $\mathcal{X}$  pour considérer un problème qui est mécaniquement de dimension plus petite. Une fois ce sous-problème "traité" (plusieurs traitements vont être considérés dans les sections qui suivent), on met à jour la variable sur laquelle le sous-problème est défini, puis on change de sous-problème (et de variable associée). Formellement, cela revient à décomposer l'espace  $\mathcal{X}$  en  $n$  sous-espaces :

$$\mathcal{X} = \mathcal{X}_1 \times \dots \times \mathcal{X}_n$$

puis à considérer successivement les problèmes suivants :

$$\min_{x_{(\ell)} \in \mathcal{X}_\ell} J((x_{(1)})_{k+1}, \dots, (x_{(\ell-1)})_{k+1}, x_{(\ell)}, (x_{(\ell+1)})_k, \dots, (x_{(n)})_k)$$

où  $(x_{(i)})_k$  correspond à un état donné d'une des sous-variables du problème. Dans cette configuration, la mise-à-jour est cyclique et déterministe (on choisit un ordre et on s'y tient). D'autres modes de mises-à-jour peuvent être envisagés, comme le choix d'un ordre aléatoire pour l'ordre des mises-à-jour. Toutefois, on s'astreint généralement à parcourir toutes les variables avant de mettre à jour à nouveau une variable. En pratique, l'ordre aléatoire s'avère plus efficace ; toutefois, pour des raisons de lisibilité, on ne considère ici que l'ordre cyclique déterministe.

Toujours pour des raisons de lisibilité, on considèrera le problème à deux blocs ( $n = 2$ )

$$\min_{X=(x,z) \in \mathcal{X} \times \mathcal{Z}} J(x, z) \quad (\mathcal{P}_{2 \text{ blocs}})$$

1. Un exemple simple est le traitement d'une image, pour lequel la dimension peut être le nombre de pixels. L'avènement des appareils photos haute résolution et des écrans hautes résolutions conduisent à travailler avec des nombres de pixels de plus en plus grands.

## 2 Minimisation alternée ou *block-coordinate descent*

### 2.1 Position du problème

On considère le problème  $(\mathcal{P}_{2 \text{ blocs}})$ , dans le cas où chaque fonction partielle

$$x \mapsto J(x, z) \quad \text{et} \quad z \mapsto J(x, z)$$

admet un minimiseur pour tout  $z \in \mathcal{Z}$  et  $x \in \mathcal{X}$  respectivement (si son domaine n'est pas vide) et que ceux-ci sont calculables de manière exacte.

Discutons des conditions d'existence de minimiseurs. Si  $J$  est coercive et s.c.i, alors il est aisé de vérifier que c'est également le cas de ses fonctions partielles. Dans ce cas, l'existence des minimiseurs est assurée. Une condition plus faible que la coercivité, mais qui suffira pour cette section, est l'existence d'un point  $(x^0, z^0) \in \mathcal{X} \times \mathcal{Z}$  tel que l'ensemble  $\text{niv}_{\leq J(x^0, z^0)} J$  soit borné. Dans ce cas, pour tout  $X = (x, z) \in \text{niv}_{\leq J(x^0, z^0)} J$ , les ensembles de sous-niveau partiels

$$\left\{ x' \in \mathcal{X} \mid J(x', z) \leq J(x, z) \right\} \quad \text{et} \quad \left\{ z' \in \mathcal{Z} \mid J(x, z') \leq J(x, z) \right\}$$

sont bornés, et les fonctions partielles (s.c.i.)  $x \mapsto J(x, z)$  et  $z \mapsto J(x, z)$  admettent un minimiseur  $z^*$  et  $x^*$  respectivement.

### 2.2 Algorithme BCD

L'algorithme BCD (pour *Block-Coordinate Descent*) est la stratégie la plus intuitive pour résoudre le problème  $(\mathcal{P}_{2 \text{ blocs}})$  sous les conditions imposées plus haut. Il s'agit tout simplement d'alterner des minimisations partielles de la fonction  $J$ , c'est-à-dire de considérer les itérations suivantes :

$$z_0 \in \mathcal{Z} \quad \text{et} \quad \forall k \in \mathbb{N}, \quad \begin{cases} x_{k+1} \in \underset{x \in \mathcal{X}}{\operatorname{argmin}} J(x, z_k) \\ z_{k+1} \in \underset{z \in \mathcal{Z}}{\operatorname{argmin}} J(x_{k+1}, z) \end{cases}$$

Notons que les itérées ne sont en général pas définies de manière unique. Pour garantir leur unicité, la continuité et la coercivité des fonctions partielles suffit.

Pour que les itérées soient bien définies, il suffit que  $z_0$  soit choisi de sorte que

$$\operatorname{dom} J(\cdot, z_0) \neq \emptyset \quad \text{soit} \quad \exists x \in \mathcal{X}_1, \quad J(x, z_0) \in \mathbb{R}$$

### 2.3 Propriétés de convergence

#### Proposition 1 (Convergence du critère)

Soit  $J : \mathcal{X} \times \mathcal{Z} \rightarrow \mathbb{R} \cup \{+\infty\}$ . On suppose qu'il existe  $(x^0, z^0) \in \operatorname{dom} J$  tel que l'ensemble de sous-niveau  $\text{niv}_{\leq J(x^0, z^0)} J$  soit borné. On considère la suite  $((x_k, z_k))_{k \in \mathbb{N}}$  générée par l'algorithme

$$\forall k \in \mathbb{N}, \quad \begin{cases} x_{k+1} \in \underset{x \in \mathcal{X}}{\operatorname{argmin}} J(x, z_k) \\ z_{k+1} \in \underset{z \in \mathcal{Z}}{\operatorname{argmin}} J(x_{k+1}, z) \end{cases}$$

Alors les suites  $(J(x_k, z_k))_{k \in \mathbb{N}}$  et  $(J(x_{k+1}, z_k))_{k \in \mathbb{N}}$  sont décroissantes et convergent vers la même limite.

**DÉMONSTRATION :** Il est immédiat que, par optimalité, on a pour tout  $k \in \mathbb{N}$

$$J(x_{k+1}, z_k) \leq J(x_k, z_k)$$

et

$$J(x_{k+1}, z_{k+1}) \leq J(x_{k+1}, z_k)$$

Les deux suites considérées dans la proposition 1 sont donc décroissantes et minorées, et par conséquent convergentes. L'identité de leur limite découle de l'encadrement

$$J(x_{k+1}, z_{k+1}) \leq J(x_{k+1}, z_k) \leq J(x_k, z_k)$$

qui permet de conclure par comparaison. ■

La règle de FERMAT assure que les points générés par l'algorithme BCD vérifient

$$0 \in \partial_x J(x_{k+1}, z_k) \quad \text{et} \quad 0 \in \partial_z J(x_{k+1}, z_{k+1})$$

Lorsque  $J$  satisfait les hypothèses de la proposition 1, les deux suites

$$((x_{k+1}, z_k))_{k \in \mathbb{N}} \quad \text{et} \quad ((x_k, z_k))_{k \in \mathbb{N}}$$

sont bornées, donc admettent chacune une sous-suite qui converge. Notons  $(x^*, z^*)$  et  $(\tilde{x}^*, \tilde{z}^*)$  les deux limites respectives. Si, de plus,  $J$  est continue, alors

$$\lim_{j \rightarrow +\infty} J(x_{k_j+1}, z_{k_j}) = J(x^*, z^*) \quad \text{et} \quad \lim_{j \rightarrow +\infty} J(x_{\tilde{k}_j}, z_{\tilde{k}_j}) = J(\tilde{x}^*, \tilde{z}^*)$$

Ainsi, si les sous-différentiels partiels sont fermés, alors on a

$$0 \in \partial_x J(x^*, z^*) \quad \text{et} \quad 0 \in \partial_z J(\tilde{x}^*, \tilde{z}^*)$$

Notons que, par unicité de la limite, on a  $J(x^*, z^*) = J(\tilde{x}^*, \tilde{z}^*)$ .

Pour aller au-delà de ce premier résultat, il faut faire des hypothèses supplémentaires sur le problème. Si on suppose que

$$\forall (x, z) \in \mathcal{X} \times \mathcal{Z}, \quad J(x, z) = f(x, z) + \chi_{C_1}(x) + \chi_{C_2}(z)$$

avec  $f$  une fonction continûment différentiable, alors les remarques précédentes se traduisent de la manière suivante :

$$\frac{\partial J}{\partial x}(x_{k+1}, z_k) = 0 \quad \text{et} \quad \frac{\partial J}{\partial z}(x_{k+1}, z_{k+1}) = 0$$

### Lemme 1

Soit  $J : \mathcal{X} \times \mathcal{Z} \rightarrow \mathbb{R} \cup \{+\infty\}$  une fonction continûment différentiable sur son domaine. On suppose qu'il existe  $(x^0, z^0) \in \text{dom } J$  tel que l'ensemble de sous-niveau  $\text{niv}_{\leq J(x^0, z^0)} J$  soit borné. On considère la suite  $((x_k, z_k))_{k \in \mathbb{N}}$  générée par l'algorithme

$$\forall k \in \mathbb{N}, \quad \begin{cases} x_{k+1} \in \underset{x \in \mathcal{X}_1}{\text{argmin}} J(x, z_k) \\ z_{k+1} \in \underset{z \in \mathcal{X}_2}{\text{argmin}} J(x_{k+1}, z) \end{cases}$$

On suppose que la suite  $(x_k)_{k \in \mathbb{N}}$  est convergente. Alors la suite  $((x_k, z_k))_{k \in \mathbb{N}}$  est bornée et admet une sous-suite qui converge. De plus, toute sous-suite convergente de  $((x_k, z_k))_{k \in \mathbb{N}}$  converge vers un point critique de  $J$ .

**DÉMONSTRATION :** Par hypothèse, la suite  $((x_k, z_k))_{k \in \mathbb{N}}$  est bornée. On suppose que la suite  $(x_{k_j}, z_{k_j})_{j \in \mathbb{N}}$  est convergente, de limite  $(x^*, z^*)$ . Alors

$$0 = \lim_{j \rightarrow +\infty} \frac{\partial J}{\partial x}(x_{k_j}, z_{k_j}) = \frac{\partial J}{\partial x}(x^*, z^*)$$

La convergence de la suite des  $x_k$  implique que

$$\lim_{j \rightarrow +\infty} (x_{k_j+1}, z_{k_j}) = (x^*, z^*)$$

Par continuité de  $\nabla J$ , on en déduit que

$$0 = \lim_{j \rightarrow +\infty} \frac{\partial J}{\partial z}(x_{k_j+1}, z_{k_j}) = \frac{\partial J}{\partial z}(x^*, z^*)$$

Autrement dit,  $(x^*, z^*)$  est un point critique de  $J$ . ■

Avant de poursuivre, on va montrer que, si  $J$  est convexe mais non différentiable, l'algorithme BCD peut ne pas converger vers un point critique.

#### CONTRE-EXEMPLE

**Non-convergence vers le minimiseur dans le cas non différentiable.**

On considère le cas où  $\mathcal{X} = \mathcal{Z} = \mathbb{R}$  avec

$$J(x, z) = |x + z| + 2|x - z|$$

Il est aisé de vérifier que, comme combinaison linéaire à coefficients positifs de combinaisons d'une fonction convexe (la valeur absolue) avec une forme linéaire, la fonction  $J$  est convexe. Appliquons l'algorithme BCD à la minimisation de cette fonction. Les itérations s'écrivent

$$\forall k \in \mathbb{N}, \quad \begin{cases} x_{k+1} \in \operatorname{argmin}_{x \in \mathbb{R}} \{|x + z_k| + 2|x - z_k|\} \\ z_{k+1} \in \operatorname{argmin}_{z \in \mathbb{R}} \{|x_{k+1} + z| + 2|x_{k+1} - z|\} \end{cases}$$

On voit qu'il s'agit dans les deux cas de minimiser la fonction strictement convexe, continue, affine par morceaux et coercive  $f_a$ , avec

$$\forall t \in \mathbb{R}, \quad f_a(t) = |t + a| + 2|t - a|$$

Un simple calcul montre que, si  $a \geq 0$ , alors

$$\forall t \in \mathbb{R}, \quad f_a(t) = \begin{cases} -3t + a & \text{si } t \leq -a \\ -t + 3a & \text{si } |t| \leq a \\ 3t - a & \text{si } t \geq a \end{cases}$$

de sorte que son minimiseur se trouve parmi les deux points  $a$  et  $-a$ ; or, puisque  $f_a(a) = 2a$  et que  $f_a(-a) = 4a$ , il s'ensuit que le minimiseur vaut  $a$ . Ainsi, si  $z_0 > 0$ , on démontre par récurrence que, pour tout  $k \in \mathbb{N}$ ,

$$x_{k+1} = z_{k+1} = x_0$$

La suite générée par l'algorithme BCD est donc constante et convergente, de limite  $(x_0, x_0)$ . Or, il est visible que ce point n'est pas un minimiseur de  $J$ , qui atteint son minimum en  $(0, 0)$ . Selon la règle de FERMAT, ce n'est pas non plus un point critique.

En réalité, dans l'exemple qui précède, l'algorithme, lorsqu'il n'est pas correctement initialisé (c'est-à-dire si  $z_0 \neq 0$ ), ne converge pas vers un point critique, mais un point vérifiant la double inclusion

$$0 \in \partial_x J(x, z) \quad \text{et} \quad 0 \in \partial_z J(x, z)$$

Or, on a vu qu'en général, ces points sont différents des points critiques de  $J$ , sauf lorsque le terme couplant les deux variables est continûment différentiable.

Une conséquence immédiate du lemme précédent est la convergence de la suite des itérées vers un point critique lorsque cette suite est convergente. En réalité, une condition importante pour assurer la convergence de l'algorithme BCD est l'unicité des itérées. On peut en effet considérer l'exemple suivant :

## CONTRE-EXEMPLE

**Non-unicité des itérées.** On considère le cas où  $\mathcal{X} = \mathcal{Z} = \mathbb{R}$  avec

$$J(x, z) = \begin{cases} \max \left\{ \min\{1, |x|\}, \min\{1, |z|\} \right\} & \text{si } (x, z) \in [-1; 1] \times [-1; 1] \\ +\infty & \text{sinon} \end{cases}$$

La fonction  $J$  admet visiblement un unique minimum, atteint en  $(0, 0)$  et qui vaut 0. Il est aisé de vérifier que si  $z \in \{-1, 1\}$ , alors  $J(x, z) = 1$  pour tout  $x \in [-1; 1]$ , et vaut  $+\infty$  pour  $x \notin [-1; 1]$  (et on démontre un résultat analogue lorsque  $x \in \{-1, 1\}$ ). On peut donc initialiser l'algorithme BCD avec  $z_0 = 1$ ; dans ce cas, tout le segment  $[-1; 1]$  est éligible pour fournir le point suivant  $x_1$ . Si on choisit  $x_1 = 1$ , alors on peut cette fois choisir n'importe quel point dans  $[-1; 1]$  pour définir le point  $z_1$ . On voit donc qu'il est possible avec l'algorithme BCD de générer la suite constante égale à  $(1, 1)$ . Celle-ci est convergente, mais ne converge pas vers le minimiseur de  $J$ . En revanche, si on avait choisit, avec la même initialisation,  $x_1 = 0$ , puis  $z_1 = 0$ , l'algorithme convergerait en une seule itération avec le minimiseur de  $J$ .

Dans la proposition suivante, on considère le cas biconvexe et continue, sous l'hypothèse que les itérées sont définies de manière unique (l'algorithme BCD est alors également désigné par le nom ACS ou *Alternate Convex Search* car chaque sous-problème d'optimisation est un problème d'optimisation convexe) :

**Proposition 2** (*Alternate Convex Search*)

Soit  $J : \mathcal{X} \times \mathcal{Z} \rightarrow \mathbb{R} \cup \{+\infty\}$  une fonction continue sur son domaine et **bi-convexe**. On suppose que  $\text{dom } J = \mathcal{C}_1 \times \mathcal{C}_2$  où  $\mathcal{C}_1 \subset \mathcal{X}$  et  $\mathcal{C}_2 \subset \mathcal{Z}$  sont des ensembles convexes et fermés. On suppose qu'il existe  $(x^0, z^0) \in \text{dom } J$  tel que l'ensemble de sous-niveau  $\text{niv}_{\leq J(x^0, z^0)} J$  soit borné, et que pour tout  $z \in \mathcal{Z}$ , la fonction partielle  $x \mapsto J(x, z)$  est **strictement convexe**. On considère la suite  $((x_k, z_k))_{k \in \mathbb{N}}$  générée par l'algorithme

$$\forall k \in \mathbb{N}, \quad \begin{cases} x_{k+1} = \underset{x \in \mathcal{X}}{\text{argmin}} J(x, z_k) \\ z_{k+1} = \underset{z \in \mathcal{Z}}{\text{argmin}} J(x_{k+1}, z) \end{cases}$$

Alors la suite  $((x_k, z_k))_{k \in \mathbb{N}}$  est bornée et admet une sous-suite qui converge. Par ailleurs, la limite  $(x^*, z^*)$  de toute sous-suite convergente de  $((x_k, z_k))_{k \in \mathbb{N}}$  vérifie

$$0 \in \partial_x J(x^*, z^*) \quad \text{et} \quad 0 \in \partial_z J(x^*, z^*)$$

Si, de plus, il existe une décomposition de  $J$  de la forme

$$\forall (x, z) \in \mathcal{X} \times \mathcal{Z}, \quad J(x, z) = f(x) + g(z) + h(x, z)$$

où  $h$  est continûment différentiable autour de  $(x^*, z^*)$ , alors  $(x^*, z^*)$  est un point critique de  $J$ .

**DÉMONSTRATION :** Par hypothèse, la suite  $((x_k, z_k))_{k \in \mathbb{N}}$  est **bornée**. On suppose que la sous-suite  $((x_{k_j}, z_{k_j}))_{j \in \mathbb{N}}$  est convergente, de limite  $(x^*, z^*)$ . Puisque les  $(x_k, z_k)$  sont dans le domaine de  $J$ , et que celui-ci est fermé, il s'ensuit que

$$\lim_{k \rightarrow +\infty} J(x_k, z_k) = \lim_{j \rightarrow +\infty} J(x_{k_j}, z_{k_j}) = J(x^*, z^*)$$

- **La suite des  $\|(x_{k_j+1}, z_{k_j+1}) - (x_{k_j}, z_{k_j})\|$  tend vers 0.** Par l'absurde, supposons que ce n'est pas le cas. Alors, en posant

$$\forall j \in \mathbb{N}, \quad \delta_j = \|(x_{k_j+1}, z_{k_j+1}) - (x_{k_j}, z_{k_j})\|$$

on suppose que la suite  $(\delta_j)_{j \in \mathbb{N}}$  ne converge pas vers 0. Comme elle est par ailleurs bornée (car la suite  $((x_k, z_k))_{k \in \mathbb{N}}$  l'est), elle admet donc une sous-suite  $(\delta_{j_\ell})_{\ell \in \mathbb{N}}$  qui converge vers un certain  $\alpha > 0$ . Pour tout  $\ell \in \mathbb{N}$ , définissons le point

$$s_\ell = \frac{x_{k_{j_\ell}+1} - x_{k_{j_\ell}}}{\delta_{j_\ell}} \in \mathcal{X}$$

Puisque ces points appartiennent à la boule fermée unité, qui est compacte dans  $\mathcal{X}$ , la suite des  $s_\ell$  admet une sous-suite  $(s_{\ell_m})_{m \in \mathbb{N}}$  convergente ; notons  $s^*$  sa limite. Pour tout  $m \in \mathbb{N}$ , posons

$$f_m : x \mapsto J(x, z_{k_{j_{\ell_m}}})$$

Par hypothèse, cette fonction est strictement convexe et atteint son unique minimum en  $x_{k_{j_{\ell_m}}+1}$ . Remarquons que

$$\forall m \in \mathbb{N}, \quad x_{k_{j_{\ell_m}}+1} = x_{k_{j_{\ell_m}}} + \delta_{j_{\ell_m}} s_{\ell_m} \in \mathcal{C}_1$$

Soit  $\lambda \in [0; 1]$ . On a

$$\lambda x_{k_{j_{\ell_m}}+1} + (1 - \lambda) x_{k_{j_{\ell_m}}} = x_{k_{j_{\ell_m}}} + \lambda \delta_{j_{\ell_m}} s_{\ell_m} \in \mathcal{C}_1$$

Par convexité et optimalité, on en déduit que

$$\begin{aligned} f_m(x_{k_{j_{\ell_m}}} + \lambda \delta_{j_{\ell_m}} s_{\ell_m}) &\leq \lambda f_m(x_{k_{j_{\ell_m}}+1}) + (1 - \lambda) f_m(x_{k_{j_{\ell_m}}}) \\ &\leq \lambda f_m(x_{k_{j_{\ell_m}}}) + (1 - \lambda) f_m(x_{k_{j_{\ell_m}}}) \end{aligned}$$

On a donc l'encadrement

$$f_m(x_{k_{j_{\ell_m}}+1}) \leq f_m(x_{k_{j_{\ell_m}}} + \lambda \delta_{j_{\ell_m}} s_{\ell_m}) \leq f_m(x_{k_{j_{\ell_m}}})$$

En revenant à la définition de  $f_m$ , et en vertu de la proposition 1, on a

$$\lim_{m \rightarrow +\infty} f_m(x_{k_{j_{\ell_m}}+1}) = \lim_{m \rightarrow +\infty} f_m(x_{k_{j_{\ell_m}}}) = J(x^*)$$

Ainsi, par encadrement, on a démontré que, pour tout  $\lambda \in [0; 1]$ ,

$$\lim_{m \rightarrow +\infty} f_m(x_{k_{j_{\ell_m}}} + \lambda \delta_{j_{\ell_m}} s_{\ell_m}) = J(x^*)$$

Or, par continuité de  $J$ , on a

$$J(x^*, z^*) = \lim_{m \rightarrow +\infty} f_m(x_{k_{j_{\ell_m}}} + \lambda \delta_{j_{\ell_m}} s_{\ell_m}) = J(x^* + \lambda \alpha s^*, z^*)$$

Comme cette égalité est vraie pour tout  $\lambda \in [0; 1]$ , et que  $\alpha s^* \neq 0$ , elle contredit la stricte convexité de  $x \mapsto J(x, x_{(2)}^*)$ . On en déduit donc que l'hypothèse initiale, à savoir que la suite  $(\delta_j)_{j \in \mathbb{N}}$  ne converge pas vers 0, est fausse.

- **$(x^*, z^*)$  est un point critique partiel de  $J$ .** Une conséquence directe du point précédent est que

$$\lim_{j \rightarrow +\infty} (x_{k_j+1}, z_{k_j}) = (x^*, z^*) = \lim_{j \rightarrow +\infty} (x_{k_j}, z_{k_j})$$

Par continuité et biconvexité de  $J$ , on peut alors utiliser la fermeture des sous-différentiels pour obtenir que

$$0 \in \partial_x J(x^*, z^*) \quad \text{et} \quad 0 \in \partial_z J(x^*, z^*)$$

Le point  $(x^*, z^*)$  est par ailleurs un point critique de  $J$  si

$$\partial_x J(x^*, z^*) \times \partial_z J(x^*, z^*) \subset \partial J(x^*, z^*)$$

ce qui est le cas sous les conditions additionnelles introduites dans l'énoncé de la proposition. ■

## 2.4 Exemples d'applications

Une application historique, qui donne le nom alternatif de méthode GAUSS–SEIDEL non linéaire à l'algorithme BCD, est la résolution itérative d'un système linéaire

$$Ax = b \quad (\mathcal{S})$$

où  $A \in \mathcal{M}_{n,n}(\mathbb{R})$  est une matrice symétrique définie positive et  $b \in \mathbb{R}^n \setminus \{0\}$ . L'idée est de réécrire ce problème sous la forme d'un problème d'optimisation équivalent :

$$\min_{x \in \mathbb{R}^n} \left\{ \frac{1}{2} \langle Ax, x \rangle - \langle b, x \rangle \right\} \quad (\mathcal{P}_{\text{GS}})$$

Le système linéaire  $(\mathcal{S})$  et le problème  $(\mathcal{P}_{\text{GS}})$  partagent le même ensemble de solutions. En effet, il est aisé<sup>2</sup> de vérifier que la fonction objectif de  $(\mathcal{P}_{\text{GS}})$  est une *fonction quadratique généralisée*, qu'elle est fortement convexe si et seulement si  $A$  est définie positive, et que, par ailleurs, si  $A$  est symétrique, alors la règle de FERMAT (qui est une condition nécessaire et suffisante d'optimalité ici) s'écrit

$$Ax = b$$

La fonction objectif est définie sur  $\mathbb{R}^n$ , et on peut naturellement décomposer la variable  $x \in \mathbb{R}^n$  en  $n$  blocs, correspondant aux  $n$  coordonnées de  $x$ . Intéressons-nous à l'expression des fonctions partielles de la fonction objectif  $J$ . Soit  $i \in \llbracket 1; n \rrbracket$ . La  $i$ -ème fonction partielle  $f_i : x_{(i)} \mapsto J(x)$  est de la forme

$$\forall t \in \mathbb{R}, \quad f_i(t) = \frac{1}{2} A_{i,i} t^2 + \sum_{\substack{j=1 \\ j \neq i}}^n A_{j,i} x_{(j)} t - b_i t + \text{constante}$$

Puisque les  $A_{i,i}$  sont strictement positifs (car  $A$  est définie positive), il s'ensuit que la  $i$ -ème fonction partielle est une parabole strictement convexe, donc admet un unique minimiseur  $t_i^*$  donné par

$$t_i^* = \frac{1}{A_{i,i}} \left( b_i - \sum_{\substack{j=1 \\ j \neq i}}^n A_{j,i} x_{(j)} \right) = \frac{1}{A_{i,i}} \left( b_i - \sum_{\substack{j=1 \\ j < i}}^n A_{j,i} x_{(j)} - \sum_{\substack{j=1 \\ j > i}}^n A_{j,i} x_{(j)} \right)$$

Nous pouvons donc maintenant écrire les itérations de l'algorithme BCD appliqué au problème  $(\mathcal{P}_{\text{GS}})$  : pour  $x_0 = ((x_{(i)})_0)_{1 \leq i \leq n} \in \mathbb{R}^n$ , on a pour tout  $k \in \mathbb{N}$

$$\forall i \in \llbracket 1; n \rrbracket, \quad (x_{(i)})_{k+1} = \operatorname{argmin}_{t \in \mathbb{R}} f(x_{k+1}, \dots, (x_{(i-1)})_{k+1}, t, (x_{(i+1)})_k, \dots, (x_{(n)})_k)$$

D'après ce qui précède, on a

$$(x_{(i)})_{k+1} = \frac{1}{A_{i,i}} \left( b_i - \sum_{\substack{j=1 \\ j < i}}^n A_{j,i} (x_{(j)})_{k+1} - \sum_{\substack{j=1 \\ j > i}}^n A_{j,i} (x_{(j)})_k \right)$$

En particulier, chacun de ces points satisfait l'équation linéaire suivante :

$$\sum_{\substack{j=1 \\ j < i}}^n A_{i,j} (x_{(j)})_{k+1} + A_{i,i} (x_{(i)})_{k+1} + \sum_{\substack{j=1 \\ j > i}}^n A_{i,j} (x_{(j)})_k = b_i$$

Autrement dit, pour déterminer la  $i$ -ème coordonnée du nouveau vecteur  $x_{k+1}$ , on considère la  $i$ -ème équation du système linéaire étudié et on y fixe la valeur des autres coordonnées  $x_{(j)}$  à leur valeur courante (c'est-à-dire  $x_{(j)} = (x_{(j)})_{k+1}$  si  $j < i$  et  $x_{(j)} = (x_{(j)})_k$

2. Cf. **B2 : Introduction à la méthode des moindres carrés** (3MA261)

si  $j > i$ ), puis on résout cette équation affine. On met ainsi à jour la valeur de  $x_{(i)}$  et on réitère le processus pour la coordonnée suivante  $x_{(i+1)}$ . On reconnaît l'algorithme de GAUSS-SEIDEL pour la résolution d'un système linéaire.

La fonction objectif  $J$  du problème  $(\mathcal{P}_{\text{GS}})$  est continûment différentiable (car polynomiale), et multiconvexe (car convexe) lorsque que  $A$  est (semi-)définie positive. Par ailleurs, chaque fonction partielle est strictement convexe (car les  $A_{i,i}$  sont strictement positifs), donc on peut appliquer une généralisation multiconvexe de la proposition 2, qui assure que toute valeur d'adhérence de la suite générée par cet algorithme est un point critique de  $J$ . Puisque  $J$  est dans ce cas strictement convexe, l'ensemble des points critiques de  $J$  est réduit à l'unique minimiseur de  $J$ , ce qui assure par ailleurs la convergence forte de la suite des itérées vers ce point, qui n'est autre que la solution du système linéaire  $(\mathcal{S})$ .

L'algorithme de GAUSS-SEIDEL n'est *a priori* pas un algorithme très efficace pour résoudre un système linéaire de manière exacte, puisqu'il faut en général une infinité d'itérations pour atteindre la solution (contrairement à la méthode du pivot de GAUSS, par exemple, qui donne une solution en un nombre fini d'itérations). Or, en pratique, on interrompt nécessairement la suite prématurément (donc au bout de  $N$  itérations), ce qui ne donne qu'une approximation de la solution recherchée. L'intérêt de cet algorithme se révèle donc dans les cas où la mise en œuvre d'une résolution exacte n'est pas possible ou déraisonnable, soit parce que le système est trop grand, ou encore parce que la matrice  $A$  est mal conditionnée.

Notons enfin que, pour résoudre de manière approchée le système linéaire  $Ax = b$ , on peut également appliquer la méthode du gradient explicite au problème  $(\mathcal{P}_{\text{GS}})$ , car la fonction objectif  $J$  est  $\|A\|$ -régulière (cf. **B2 : Méthodes du gradient explicite**). Cependant, la convergence de cet algorithme est conditionnée par le choix du pas de temps, qui lui-même dépend du calcul de la norme de  $A$ .

Un second exemple, plus générique, d'application de l'algorithme BCD est celui des estimations jointes. Pour illustrer notre propos, prenons l'exemple de la restauration d'une image numérique  $g$ . Cette image est supposée bruitée et floue ; on peut la modéliser comme la dégradation d'une image idéale  $u$  de la manière suivante :

$$g = K \star u + n$$

où  $K$  est un noyau de convolution modélisant le flou et  $n$  un bruit blanc gaussien. Parmi les méthodes permettant de débruiter ou de rendre nette une image, on a les méthodes dites *variationnelles* qui écrivent ces problèmes comme ceux de la minimisation d'une fonction objectif. Par exemple, pour la restauration de l'image dégradée  $g$ , on peut considérer le problème de la forme

$$\min_{(u,K) \in \mathcal{X} \times \Sigma} \left\{ \frac{\mu}{2} \|v - K \star u\|^2 + \|\nabla u\| + \alpha \|K\|^2 \right\}$$

(on n'entre pas ici dans les détails quant à la pertinence du choix de cette fonction). De manière générale, ce problème est difficile à résoudre ; si on applique l'algorithme BCD, on est amené à résoudre successivement les deux problèmes

$$\min_{u \in \mathcal{X}} \left\{ \frac{\mu}{2} \|v - K \star u\|^2 + \|\nabla u\| \right\}$$

et

$$\min_{K \in \Sigma} \left\{ \frac{\mu}{2} \|v - K \star u\|^2 + \alpha \|K\|^2 \right\}$$



Le premier problème peut être vu comme le débruitage de l'image  $g$  connaissant l'opérateur de flou  $K$ , tandis que le second problème comme la détermination de cet opérateur de flou connaissant l'image nette  $u$ . On a donc scindé la tâche en deux sous-tâches plus abordables. Il est clair qu'une telle procédure est très peu robuste, car dépendant très fortement de l'initialisation (que ce soit de  $K$  ou de  $u$ , à partir de la donnée  $g$ ). Ainsi, la convergence éventuelle de la suite des itérées est très sensible à l'initialisation, avec un risque non négligeable de converger vers un point critique non minimiseur.

De manière générale, en dépit de son apparente simplicité, la convergence de l'algorithme BCD est difficile à établir, d'une part car la stricte convexité d'une des fonctions partielles n'est pas toujours assurée dans les problèmes considérés, et d'autre part car l'ensemble des points critiques n'est pas toujours réduite à un point.

### 3 Descente proximale linéarisée alternée

#### 3.1 Position du problème

On considère le problème  $(\mathcal{P}_{2 \text{ blocs}})$ , dans le cas où la fonction objectif peut être décomposée de la manière suivante :

$$\forall (x, z) \in \mathcal{X} \times \mathcal{Z}, \quad J(x, z) = f(x) + g(z) + h(x, z)$$

avec  $f : \mathcal{X} \rightarrow \mathbb{R} \cup \{+\infty\}$  et  $g : \mathcal{Z} \rightarrow \mathbb{R} \cup \{+\infty\}$  deux fonctions minorées et simples et  $h : \mathcal{X} \times \mathcal{Z} \rightarrow \mathbb{R}$  une fonction continûment différentiable, telles que, pour tout  $(x, z) \in \mathcal{X} \times \mathcal{Z}$ , les fonctions partielles  $x' \mapsto J(x', z)$  et  $z' \mapsto J(x, z')$  sont respectivement  $L_z$  et  $\tilde{L}_x$ -régulières. Cette décomposition n'est pas nécessairement unique.

Notons d'ores-et-déjà que, sous ces hypothèses, on a la décomposition suivante du sous-différentiel de  $J$  en  $(x, z) \in \mathcal{X} \times \mathcal{Z}$  :

$$\partial J(x, z) = \left\{ \partial f(x) + \frac{\partial h}{\partial x}(x, z) \right\} \times \left\{ \partial g(z) + \frac{\partial h}{\partial z}(x, z) \right\}$$

En particulier, si  $(x^*, z^*)$  satisfait les inclusions

$$0 \in \partial f(x^*) + \frac{\partial h}{\partial x}(x^*, z^*) \quad \text{et} \quad 0 \in \partial g(z^*) + \frac{\partial h}{\partial z}(x^*, z^*)$$

alors  $(x^*, z^*)$  est un point critique de  $J$ .

#### 3.2 Algorithme PALM

On a vu que, sous certaines conditions, il est possible de résoudre le problème  $(\mathcal{P})_{2 \text{ blocs}}$  en alternant des minimisations partielles. Cependant, les conditions d'application sont très restrictives, et les résultats de convergence relativement faibles. L'idée est donc de remplacer les minimisations partielles par des pas de descentes, appliqués aux fonctions partielles de  $J$ .

Dans l'algorithme PALM (pour *Proximal Alternating Linearized Minimization*), on choisit d'utiliser des pas de descente de gradient explicite-implicite, c'est-à-dire d'appliquer une itération de l'algorithme FBS à la fonction partielle courante. Plus précisément, on considère les itérations suivantes

$$\begin{cases} x_{k+1} \in \text{prox}_{\tau_k f} \left( x_k - \tau_k \frac{\partial h}{\partial x}(x_k, z_k) \right) \\ z_{k+1} \in \text{prox}_{\sigma_k g} \left( z_k - \sigma_k \frac{\partial h}{\partial z}(x_{k+1}, z_k) \right) \end{cases}$$

Notons que les itérées sont définies de manière unique dès que  $f$  et  $g$  sont convexes.

De la même manière que pour l'algorithme FBS, ces itérations peuvent être vues comme des minimisations de fonctions approchées ; ici, la première mise-à-jour comme la minimisation de la fonction

$$x \mapsto f(x) + g(z_k) + h(x_k, z_k) + \left\langle \frac{\partial h}{\partial x}(x_k, z_k), x - x_k \right\rangle + \frac{1}{2\tau_k} \|x - x_k\|^2$$

qui est une approximation autour du point  $x_k$  de la fonction  $x \mapsto J(x, z_k)$ , tandis que la seconde mise-à-jour s'écrit comme la minimisation de la fonction

$$z \mapsto f(x_{k+1}) + g(z) + h(x_{k+1}, z_k) + \sigma_k \left\langle \frac{\partial h}{\partial z}(x_{k+1}, z_k), z - z_k \right\rangle + \frac{1}{2\sigma_k} \|z - z_k\|^2$$

qui est une approximation autour du point  $z_k$  de la fonction  $z \mapsto J(x_{k+1}, z)$ .

Il existe évidemment d'autres manières de remplacer les minimisations partielles. On peut les remplacer par des pas de descentes de gradient explicite, par une itérations de l'algorithme du point proximal, ou encore ne pas appliquer le même traitement (minimisation / descente) aux deux variables.

### 3.3 Propriétés de convergence

Puisque chaque mise-à-jour est un pas de FBS, on peut montrer un premier résultat :

**Proposition 3 (Convergence du critère)**

Soit  $f : \mathcal{X} \rightarrow \mathbb{R} \cup \{+\infty\}$  et  $g : \mathcal{Z} \rightarrow \mathbb{R} \cup \{+\infty\}$  deux fonctions minorées et simples et  $h : \mathcal{X} \times \mathcal{Z} \rightarrow \mathbb{R}$  une fonction continûment différentiable, telles que, pour tout  $(x, z) \in \mathcal{X} \times \mathcal{Z}$ , les fonctions partielles  $x' \mapsto J(x', z)$  et  $z' \mapsto J(x, z')$  sont respectivement  $L_z$  et  $\tilde{L}_x$ -régulières. Posons

$$\forall (x, z) \in \mathcal{X} \times \mathcal{Z}, \quad J(x, z) = f(x) + g(z) + h(x, z)$$

et considérons la suite générée par l'algorithme suivant : soit  $x_0 \in \mathcal{X} \times \mathcal{Z}$  et

$$\forall k \in \mathbb{N}, \quad \begin{cases} x_{k+1} \in \text{prox}_{\tau_k f} \left( x_k - \tau_k \frac{\partial h}{\partial x}(x_k, z_k) \right) \\ z_{k+1} \in \text{prox}_{\sigma_k g} \left( z_k - \sigma_k \frac{\partial h}{\partial z}(x_{k+1}, z_k) \right) \end{cases}$$

On pose  $Q = 2$  (resp.  $\tilde{Q} = 2$ ) si  $f$  (resp.  $g$ ) est convexe, et  $Q = 1$  (resp.  $\tilde{Q} = 1$ ) sinon. Si  $\tau_k \in ]0; Q/L_{z_k}[$  et  $\sigma_k \in ]0; \tilde{Q}/\tilde{L}_{x_{k+1}}[$  pour tout  $k \in \mathbb{N}$ , alors les suites  $(J(x_k, z_k))_{k \in \mathbb{N}}$  et  $(J(x_{k+1}, z_k))_{k \in \mathbb{N}}$  sont décroissantes et convergent vers la même limite. De plus, on a pour tout  $k \in \mathbb{N}$

$$J(x_{k+1}, z_k) \leq J(x_k) - \frac{1}{2} \left( \frac{Q}{\tau_k} - L_{z_k} \right) \|x_{k+1} - x_k\|^2$$

et

$$J(x_{k+1}) \leq J(x_{k+1}, z_k) - \frac{1}{2} \left( \frac{\tilde{Q}}{\sigma_k} - L_{x_{k+1}} \right) \|z_{k+1} - z_k\|^2$$

**REMARQUE :** On voit se profiler dans cette proposition l'un des désavantages de la méthode PALM, à savoir le coûteux calcul des constantes de LIPSCHITZ des dérivées

partielles de  $h$  au point courant. En effet, ce calcul est nécessaire pour choisir correctement les pas de temps  $(\tau_k, \sigma_k)$ .

**DÉMONSTRATION :** Laissée au lecteur. Il suffit de suivre la démarche des preuves des corollaires 1 et 2 du module **B4 : Éclatement primal d'opérateurs**.

Notons qu'en particulier, sous les hypothèses de la proposition précédente, on a

$$\lim_{k \rightarrow +\infty} \left( \frac{Q}{\tau_k} - L_{z_k} \right) \|x_{k+1} - x_k\|^2 = 0$$

et

$$\lim_{k \rightarrow +\infty} \left( \frac{\tilde{Q}}{\sigma_k} - L_{x_{k+1}} \right) \|z_{k+1} - z_k\|^2 = 0$$

Pour avoir les convergences des quantités  $\|x_{k+1} - x_k\|$  et  $\|z_{k+1} - z_k\|$  vers 0 (et donc, de la quantité  $\|x_{k+1} - x_k\|$ ), il faut que les quantités positives

$$\frac{Q}{\tau_k} - L_{z_k} \quad \text{et} \quad \frac{\tilde{Q}}{\sigma_k} - L_{x_{k+1}}$$

soient minorées par une quantité strictement négatives. Autrement dit, il faut choisir les pas de temps  $\tau_k$  et  $\sigma_k$  de sorte que

$$\frac{Q}{L_{z_k} + \varepsilon} > \tau_k \quad \text{et} \quad \frac{\tilde{Q}}{L_{x_{k+1}} + \tilde{\varepsilon}} > \sigma_k$$

avec  $\varepsilon > 0$  et  $\tilde{\varepsilon} > 0$  deux réels fixés. Notons que, si les constantes de LIPSCHITZ soient majorées (par  $L$  et  $\tilde{L}$  respectivement), alors il est possible de considérer des pas constants

$$\tau \in \left] 0; \frac{Q}{L} \right[ \quad \text{et} \quad \sigma \in \left] 0; \frac{\tilde{Q}}{\tilde{L}} \right[$$

Un tel choix réduit la complexité de l'algorithme (puisque les pas de temps ne sont calculés qu'une seule fois, avant le début des itérations), mais ne permet plus de tirer avantage, le cas échéant, de propriétés localement plus intéressantes, à savoir une petite constante de LIPSCHITZ, c'est-à-dire un pas de temps adaptatif.

Intéressons-nous maintenant à la convergence des sous-gradients. Commençons par écrire, pour tout  $k \in \mathbb{N}$ , la règle de FERMAT associée aux sous-problèmes d'optimisation :

$$0 \in \partial f(x_{k+1}) + \frac{1}{\tau_k}(x_{k+1} - x_k) + \frac{\partial h}{\partial x}(x_k, z_k)$$

et

$$0 \in \partial g(z_{k+1}) + \frac{1}{\sigma_k}(z_{k+1} - z_k) + \frac{\partial h}{\partial z}(x_{k+1}, z_k)$$

En utilisant la continuité des dérivées partielles de  $h$ , on peut réécrire ces deux inclusions à l'aide des sous-différentiels partiels de  $J$  : on a d'une part

$$\frac{x_k - x_{k+1}}{\tau_k} + \frac{\partial h}{\partial x}(x_{k+1}, z_k) - \frac{\partial h}{\partial x}(x_k, z_k) \in \partial_x J(x_{k+1}, z_k)$$

et d'autre part

$$\frac{z_k - z_{k+1}}{\sigma_k} + \frac{\partial h}{\partial z}(x_{k+1}, z_{k+1}) - \frac{\partial h}{\partial z}(x_{k+1}, z_k) \in \partial_z J(x_{k+1}, z_{k+1})$$

En utilisant le caractère lipschitzien des dérivées partielles de  $h$  et le fait que les quantités  $\|x_{k+1} - x_k\|$  et  $\|z_{k+1} - z_k\|$  tendent vers zéro si les pas de temps sont choisis correctement, on en déduit le résultat suivant :

**Proposition 4** (Convergence des sous-gradients)

Soit  $f : \mathcal{X} \rightarrow \mathbb{R} \cup \{+\infty\}$  et  $g : \mathcal{Z} \rightarrow \mathbb{R} \cup \{+\infty\}$  deux fonctions minorées et simples et  $h : \mathcal{X} \times \mathcal{Z} \rightarrow \mathbb{R}$  une fonction continûment différentiable, telles que, pour tout  $(x, z) \in \mathcal{X} \times \mathcal{Z}$ , les fonctions partielles  $x' \mapsto J(x', z)$  et  $z' \mapsto J(x, z')$  sont respectivement  $L_z$  et  $\tilde{L}_x$ -régulières. Posons

$$\forall (x, z) \in \mathcal{X} \times \mathcal{Z}, \quad J(x, z) = f(x) + g(z) + h(x, z)$$

et considérons la suite générée par l'algorithme suivant : soit  $x_0 \in \mathcal{X} \times \mathcal{Z}$  et

$$\forall k \in \mathbb{N}, \quad \begin{cases} x_{k+1} \in \text{prox}_{\tau_k f} \left( x_k - \tau_k \frac{\partial h}{\partial x}(x_k, z_k) \right) \\ z_{k+1} \in \text{prox}_{\sigma_k g} \left( z_k - \sigma_k \frac{\partial h}{\partial z}(x_{k+1}, z_k) \right) \end{cases}$$

On pose  $Q = 2$  (resp.  $\tilde{Q} = 2$ ) si  $f$  (resp.  $g$ ) est convexe, et  $Q = 1$  (resp.  $\tilde{Q} = 1$ ) sinon. On définit également les vecteurs suivants :

$$p_{k+1} = \frac{x_k - x_{k+1}}{\tau_k} + \frac{\partial h}{\partial x}(x_{k+1}, z_k) - \frac{\partial h}{\partial x}(x_k, z_k)$$

$$\text{et} \quad q_{k+1} = \frac{z_k - z_{k+1}}{\sigma_k} + \frac{\partial h}{\partial z}(x_{k+1}, z_{k+1}) - \frac{\partial h}{\partial z}(x_{k+1}, z_k)$$

Ces vecteurs vérifient

$$p_{k+1} \in \partial_x J(x_{k+1}, z_k) \quad \text{et} \quad q_{k+1} \in \partial_z J(x_{k+1}, z_{k+1})$$

Si  $\tau_k \in ]0; Q/(L_{z_k} + \varepsilon)[$  et  $\sigma_k \in ]0; \tilde{Q}/(\tilde{L}_{x_{k+1}} + \varepsilon)[$  pour tout  $k \in \mathbb{N}$  pour  $\varepsilon > 0$  et  $\varepsilon > 0$  donnés, alors les suites  $(p_k)_{k \in \mathbb{N}^*}$  et  $(q_k)_{k \in \mathbb{N}^*}$  convergent vers 0. Par ailleurs, on a pour tout  $k \in \mathbb{N}$

$$\|p_{k+1}\| \leq \left( \frac{1}{\tau_k} + L_{z_k} \right) \|x_k - x_{k+1}\| \quad \text{et} \quad \|q_{k+1}\| \leq \left( \frac{1}{\sigma_k} + \tilde{L}_{x_{k+1}} \right) \|z_k - z_{k+1}\|$$

**DÉMONSTRATION** : L laissée au lecteur.

Notons que les sous-différentiels partiels de  $J$  sont fermés (proposition 22 du module **A2 : Sous-différentiabilité**). Par ailleurs, on a

$$\partial_x J(x_{k+1}, z_{k+1}) = \partial_x J(x_{k+1}, z_k) + \frac{\partial h}{\partial x}(x_{k+1}, z_{k+1}) - \frac{\partial h}{\partial x}(x_{k+1}, z_k)$$

de sorte que, si on pose

$$\tilde{p}_{k+1} = p_{k+1} + \frac{\partial h}{\partial x}(x_{k+1}, z_{k+1}) - \frac{\partial h}{\partial x}(x_{k+1}, z_k)$$

alors on a  $\tilde{p}_{k+1} \in \partial_x J(x_{k+1}, z_{k+1})$ . Ainsi, sous les hypothèses de la proposition précédente, on a

$$\|\tilde{p}_{k+1}\| \leq \left( \frac{1}{\tau_k} + L_{z_k} \right) \|x_k - x_{k+1}\| + \tilde{L}_{x_{k+1}} \|z_k - z_{k+1}\|$$

Ainsi, puisque les hypothèses de régularité sur  $J$  impliquent que

$$\forall k \in \mathbb{N}^*, \quad P_k = (\tilde{p}_k, q_k) \in \partial J(x_k, y_k)$$

et que l'on a  $(a+b)^2 \leq (a+b)^2 + (a-b)^2 = 2(a^2 + b^2)$  pour tout  $(a, b) \in \mathbb{R}^2$ , alors

$$\begin{aligned}
\|P_k\|^2 &= \|\tilde{p}_k\|^2 + \|q_k\|^2 \\
&\leq \left( \left( \frac{1}{\tau_k} + L_{z_k} \right) \|x_k - x_{k+1}\| + \tilde{L}_{x_{k+1}} \|z_k - z_{k+1}\| \right)^2 + \left( \frac{1}{\sigma_k} + \tilde{L}_{x_{k+1}} \right)^2 \|z_k - z_{k+1}\|^2 \\
&\leq 2 \left( \frac{1}{\tau_k} + L_{z_k} \right)^2 \|x_k - x_{k+1}\|^2 + \left( \left( \frac{1}{\sigma_k} + \tilde{L}_{x_{k+1}} \right)^2 + 2\tilde{L}_{x_{k+1}}^2 \right) \|z_k - z_{k+1}\|^2
\end{aligned}$$

Si on suppose que les constantes de LIPSCHITZ  $L_x$  et  $\tilde{L}_z$  sont majorées respectivement par  $L$  et  $\tilde{L}$ , alors on a pour tout  $k \in \mathbb{N}$ ,

$$\frac{Q}{L + \varepsilon} \leq \frac{Q}{L_{z_k} + \varepsilon} \quad \text{et} \quad \frac{\tilde{Q}}{\tilde{L} + \tilde{\varepsilon}} \leq \frac{\tilde{Q}}{\tilde{L}_{x_{k+1}} + \tilde{\varepsilon}}$$

En particulier, on voit qu'il est possible d'imposer le choix de deux bornes inférieures  $\tau > 0$  et  $\sigma > 0$  tels que

$$\forall k \in \mathbb{N}, \quad \tau \leq \tau_k \quad \text{et} \quad \sigma \leq \sigma_k$$

Par conséquent, on peut obtenir la majoration suivante

$$\forall k \in \mathbb{N}^*, \quad \|P_k\| \leq b \|x_{k+1} - x_k\|$$

avec 
$$b = \max \left\{ 2 \left( \frac{1}{\tau} + L \right)^2, \left( \left( \frac{1}{\sigma} + \tilde{L} \right)^2 + 2\tilde{L}^2 \right) \right\}$$

Ainsi, grâce à la proposition 19 du module **A2 : Sous-différentiabilité**, on en déduit que si la suite  $(x_k)_{k \in \mathbb{N}}$  admet une sous-suite convergente, celle-ci converge vers un point critique de  $J$ . Par ailleurs, on a le résultat général dans le cas d'une fonction KL :

**Proposition 5 (Convergence vers un point critique)**

Soit  $f : \mathcal{X} \rightarrow \mathbb{R} \cup \{+\infty\}$  et  $g : \mathcal{Z} \rightarrow \mathbb{R} \cup \{+\infty\}$  deux fonctions minorées et simples et  $h : \mathcal{X} \times \mathcal{Z} \rightarrow \mathbb{R}$  une fonction continûment différentiable, telles que, pour tout  $(x, z) \in \mathcal{X} \times \mathcal{Z}$ , les fonctions partielles  $x' \mapsto J(x', z)$  et  $z' \mapsto J(x, z')$  sont respectivement  $L_z$  et  $\tilde{L}_x$ -régulières. Posons

$$\forall (x, z) \in \mathcal{X} \times \mathcal{Z}, \quad J(x, z) = f(x) + g(z) + h(x, z)$$

On suppose que  $J$  est continue sur son domaine, supposé fermé. Considérons la suite générée par l'algorithme suivant : soit  $x_0 \in \mathcal{X} \times \mathcal{Z}$  et

$$\forall k \in \mathbb{N}, \quad \begin{cases} x_{k+1} \in \text{prox}_{\tau_k f} \left( x_k - \tau_k \frac{\partial h}{\partial x}(x_k, z_k) \right) \\ z_{k+1} \in \text{prox}_{\sigma_k g} \left( z_k - \sigma_k \frac{\partial h}{\partial z}(x_{k+1}, z_k) \right) \end{cases}$$

On pose  $Q = 2$  (resp.  $\tilde{Q} = 2$ ) si  $f$  (resp.  $g$ ) est convexe, et  $Q = 1$  (resp.  $\tilde{Q} = 1$ ) sinon. Soit  $x^*$  un point d'adhérence de la suite  $(x_k)_{k \in \mathbb{N}}$ . On suppose que  $J$  satisfait la propriété KL en  $x^*$ . On suppose que

$$\forall (x, z) \in \mathcal{X} \times \mathcal{Z}, \quad L_z \leq L \quad \text{et} \quad \tilde{L}_x \leq \tilde{L}$$

Soit 
$$0 < \tau \leq \frac{Q}{L + \varepsilon} \quad \text{et} \quad 0 < \sigma \leq \frac{\tilde{Q}}{\tilde{L} + \tilde{\varepsilon}}$$

Si  $\tau_k \in ]\tau; Q/(L_{z_k} + \varepsilon)[$  et  $\sigma_k \in ]\sigma; \tilde{Q}/(\tilde{L}_{x_{k+1}} + \tilde{\varepsilon})[$  pour tout  $k \in \mathbb{N}$  pour  $\varepsilon > 0$  et  $\tilde{\varepsilon} > 0$  donnés, alors la suite  $(x_k)_{k \in \mathbb{N}}$  converge vers  $x^*$ . De plus,  $x^*$  est un point critique de  $J$ .

**DÉMONSTRATION** : Laissée au lecteur. Il s'agit d'une application du théorème d'ATTOUCH, BOLTE & SVAITER (théorème 2 du module **A1 : Éléments de topologie**).

### 3.4 Exemple d'applications

On considère ici un problème bien connu, qui est celui de la factorisation de matrices positives. Il s'agit de déterminer une factorisation de la matrice  $A$  à coefficients positifs

$$A = XY$$

où  $X \in \mathcal{M}_{m,r}(\mathbb{R}^+)$  et  $Y \in \mathcal{M}_{r,n}(\mathbb{R}^+)$ . La connaissance de cette factorisation peut être utile dans de nombreuses applications. En particulier, on peut parfois supposer que, parmi les factorisations possibles de ce type, l'une d'elles implique une matrice parcimonieuse, c'est-à-dire avec un petit nombre de coefficients non nuls. Le problème de la factorisation peut alors s'exprimer comme un problème d'optimisation :

$$\min_{\substack{X \in \mathcal{M}_{m,r}(\mathbb{R}) \\ Y \in \mathcal{M}_{r,n}(\mathbb{R})}} \left\{ \frac{1}{2} \|A - XY\|_F^2 + \chi_{\{X \in \mathcal{M}_{m,r}(\mathbb{R}^+) \mid \|X\|_0 \leq s\}}(X) + \chi_{\mathcal{M}_{m,r}(\mathbb{R}^+)}(Y) \right\}$$

Ici, la norme de FROBENIUS est définie pour toute matrice  $A = (a_{i,j})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$  par

$$\|A\|_F = \sqrt{\sum_{i=1}^m \sum_{j=1}^n a_{i,j}^2}$$

tandis que la norme  $\ell_0$  est la norme de comptage, définit par

$$\|A\|_0 = \text{card} \{a_{i,j} \neq 0\}$$

Ainsi, les matrices  $X$  satisfaisant  $\|X\|_0 \leq s$  sont celles qui possèdent au plus  $s \in \mathbb{N}$  coefficients non nuls.

On peut aisément vérifier que, si on pose pour tout  $X, Y$

$$h(X, Y) = \frac{1}{2} \|A - XY\|_F^2$$

alors la fonction  $h$  est différentiable, et ses dérivées partielles sont lipschitziennes ; plus précisément, en adaptant les notations de la section précédente, on a

$$L_Y = \|Y^t Y\|_F \quad \text{et} \quad \tilde{L}_X = \|X^t X\|_F$$

De plus, si on pose

$$f(X) = \chi_{\{X \in \mathcal{M}_{m,r}(\mathbb{R}^+) \mid \|X\|_0 \leq s\}}(X) \quad \text{et} \quad g(Y) = \chi_{\mathcal{M}_{m,r}(\mathbb{R}^+)}(Y)$$

alors les fonctions  $f$  et  $g$  sont simples ; en effet, l'opérateur proximal d'une indicatrice d'ensemble étant la projection orthogonale sur cet ensemble, il est immédiat que  $\text{prox}_g$  est calculable (il suffit de remplacer tous les coefficients négatifs de la matrice  $Y$  par zéro). L'autre projection (sur les matrices à au plus  $s$  coefficients non nuls et positifs) est également facilement calculables. Il suffit d'appliquer la projection sur les matrices à coefficients positifs puis de retenir les  $s$  coefficients les plus grands parmi ceux qui restent, les autres étant ramenés à 0. On voit donc que cette projection n'est pas unique (il suffit de songer à la matrice avec 1 pour tous les coefficients).

### Pour aller plus loin

**Minimisation et descentes alternées.** Les stratégies de minimisation partielle (alternée) ou de descentes alternées (et leurs équivalents en maximisation) sont utilisées massivement dans la conception des algorithmes en optimisation numérique. On peut citer les exemples déjà vus de l'éclatement de DYKSTRA (étudié dans le module **B4 : Éclatement primal d'opérateurs**), de l'algorithme de CHAMBOLLE-POCK et l'ADMM (introduits dans le module **B5 : Éclatement primal-dual**).

**Méthode de GAUSS-SEIDEL** L'étude de la convergence de cette méthode est à retrouver dans le **PROBLÈME 1 : Optimisation convexe et systèmes linéaires** (3MA261). La théorie sur les fonctions quadratiques généralisées est développée quant à elle dans **B3 : Introduction à la méthode des moindres carrés** (3MA261).