Optimisation quadratique successive

Dans ce chapitre, nous étudions une famille d'*algorithmes newtoniens*, qui procèdent donc par linéarisation de fonctions, pour résoudre les problèmes d'optimisation avec contraintes d'égalité et d'inégalité. Les algorithmes requièrent la résolution de problèmes d'optimisation quadratique à chaque itération, de problèmes non linéaires donc, mais qui sont plus simples que le problème original. On s'écarte ainsi des méthodes de Newton en résolution de systèmes d'équations différentiables ou de problèmes d'optimisation, décritent au chapitre 10, qui ne requièrent que la résolution d'un système linéaire à chaque itération.

L'approche étudiée est suffisamment générale pour pouvoir considérer les problèmes qui s'écrivent sous la forme suivante

$$(P_{EI}) \qquad \begin{cases} \min\ f(x) \\ c_i(x) = 0, \quad i \in E \\ c_i(x) \leqslant 0, \quad i \in I, \end{cases}$$

où les ensembles d'indices E et I forment une partition de $[1:m] = E \cup I$ ($E \cap I = \varnothing$). Les paramètres x à optimiser sont dans un espace euclidien $\mathbb{E}$ et les fonctions $f : \mathbb{E} \to \mathbb{R}$ et $c_i : \mathbb{E} \to \mathbb{R}$ définissant le critère et les contraintes sont supposées régulières. Les versions quasi-newtoniennes seront aussi étudiées (section 15.4).

Les notations sont celles déjà introduites à la section 4. On rassemble les contraintes d'égalité et d'inégalité en une seule fonction $c : \mathbb{E} \to \mathbb{R}^m$. Si $v \in \mathbb{R}^m$, on note v_E (resp. v_I) le vecteur de $\mathbb{R}^{|E|}$ (resp. $\mathbb{R}^{|I|}$) formé des composantes v_i de v avec $i \in E$ (resp. $i \in I$). Les fonctions définissant les contraintes d'égalité et d'inégalité seront donc notées c_E et c_I, respectivement. À un vecteur $v \in \mathbb{R}^m$, on associe le vecteur $v^\# \in \mathbb{R}^m$, défini par

$$(v^\#)_i = \begin{cases} v_i & \text{si } i \in E \\ v_i^+ & \text{si } i \in I, \end{cases}$$

où $v_i^+ = \max(0, v_i)$. Avec cette notation, les contraintes de (P_{EI}) s'écrivent $c(x)^\# = 0$, qui n'a d'intérêt que par sa compacité, car la fonction $x \mapsto c(x)^\#$ est en général non différentiable.

Si la présence de contraintes d'égalité dans (P_{EI}) rend ce problème plus difficile à résoudre qu'un problème sans contrainte. Le saut de complexité apporté par la présence des contraintes d'inégalité est incomparablement plus important que celui dû aux contraintes d'égalité.

15.1 L'algorithme OQS et sa convergence locale

Nous présentons l'algorithme OQS comme un algorithme de Josephy-Newton sur le système d'optimalité, lequel peut en effet s'écrire comme un problème d'inclusion, et même de complémentarité non linéaire (section 15.1.1). Sa convergence locale peut alors se déduire de celle de l'algorithme de Josephy-Newton (théorème **??**), pourvu que le point stationnaire considéré soit semi-stable et hémi-stable. On montre que c'est effectivement le cas lorsque celui-ci correspond à un minimum local de (P_{EI}) vérifiant les conditions d'optimalité du second ordre semi-fortes (section 15.1.2).

15.1.1 L'algorithme OQS

Par définition, un couple stationnaire de (P_{EI}) est un couple (x, λ) qui vérifie le système d'optimalité (4.33), c'est-à-dire

$$\begin{cases} \nabla f(x) + c'(x)^* \lambda = 0 \\ c_E(x) = 0 \\ 0 \leqslant \lambda \perp c_I(x) \leqslant 0. \end{cases} \tag{15.1}$$

On peut récrire ce système comme l'inclusion ou le problème de complémentarité en $z := (x, \lambda)$ ci-dessous

$$F(z) + \mathrm{N}_K(z) \ni 0 \qquad \text{ou} \qquad K \ni z \perp F(z) \in K^+, \tag{15.2a}$$

avec

$$F(z) = \begin{pmatrix} \nabla f(x) + c'(x)^* \lambda \\ -c(x) \end{pmatrix} \qquad \text{et} \qquad K = \mathbb{E} \times (\mathbb{R}^{m_E} \times \mathbb{R}^{m_I}_+). \tag{15.2b}$$

En effet, alors $K^+ = \{0_{\mathbb{E}}\} \times (\{0_{\mathbb{R}^{m_E}}\} \times \mathbb{R}^{m_I}_+)$ et

- $(x, \lambda) \in K$ s'écrit $\lambda_I \geqslant 0$,
- $F(x, \lambda) \in K^+$ s'écrit $\nabla_x \ell(x, \lambda) = 0$, $c_E(x) = 0$ et $c_I(x) \leqslant 0$,
- $(x, \lambda) \perp F(x, \lambda)$ se ramène alors à la relation de complémentarité $\lambda_I \perp c_I(x)$.

On retrouve donc bien (15.1).

L'algorithme de Josephy-Newton (**??**) appliqué à l'inclusion ou au problème de complémentarité (15.2) consiste à déterminer l'itéré suivant $z_{k+1} := (x_{k+1}, \lambda_{k+1})$ à partir de l'itéré courant $z_k := (x_k, \lambda_k)$ en résolvant l'inclusion linéarisée en $z := (x, \lambda)$ ci-dessous

$$F(z_k) + F'(z_k)(z - z_k) + \mathrm{N}_K(z) \ni 0 \tag{15.3a}$$

ou son problème de complémentarité équivalent

$$K \ni z \perp \big(F(z_k) + F'(z_k)(z - z_k)\big) \in K^+. \tag{15.3b}$$

Voyons comment s'écrit cet algorithme lorsque (F, K) est donné par (15.2b). Observons que

$$F'(x, \lambda) = \begin{pmatrix} L(x, \lambda) & c'(x)^* \\ -c'(x) & 0 \end{pmatrix}, \tag{15.4}$$

où l'on a simplifié l'écriture en introduisant

$$L(x,\lambda) := \nabla^2_{xx}\ell(x,\lambda).$$

La condition $(x,\lambda)\in K$ s'écrit comme précédemment

$$\lambda_I \geqslant 0. \tag{15.5a}$$

La condition $F(z_k)+F'(z_k)(z-z_k)\in K^+$ se traduit par

$$\nabla f(x_k)+c'(x_k)^*\lambda_k+L(x_k,\lambda_k)(x-x_k)+c'(x_k)^*(\lambda-\lambda_k)=0, \tag{15.5b}$$
$$c_E(x_k)+c'_E(x_k)(x-x_k)=0, \tag{15.5c}$$
$$c_I(x_k)+c'_I(x_k)(x-x_k)\leqslant 0. \tag{15.5d}$$

Enfin, la relation d'orthogonaité $(x,\lambda)\perp(F(z_k)+F'(z_k)(z-z_k))$ s'exprime alors par

$$\lambda_I \perp [c_I(x_k)+c'_I(x_k)(x-x_k)]. \tag{15.5e}$$

Si l'on élimine λ_k de (15.5b), ce système peut se récrire

$$\begin{cases} \nabla f(x_k)+L(x_k,\lambda_k)(x-x_k)+c'(x_k)^*\lambda=0, \\ c_E(x_k)+c'_E(x_k)(x-x_k)=0, \\ 0\leqslant \lambda_I \perp [c_I(x_k)+c'_I(x_k)(x-x_k)]\leqslant 0. \end{cases} \tag{15.6}$$

Le point important est maintenant de constater que le système (15.6) est formé des conditions d'optimalité de ce que l'on appelle le *problème quadratique osculateur* de (P_{EI}), à savoir

$$\boxed{\begin{cases} \min_x\ \nabla f(x_k)^\top(x-x_k)+\frac{1}{2}(x-x_k)^\top\nabla^2_{xx}(x_k,\lambda)(x-x_k) \\ c_E(x_k)+c'(x_k)_E(x-x_k)=0 \\ c_I(x_k)+c'(x_k)_I(x-x_k)\leqslant 0. \end{cases}} \tag{15.7}$$

On peut à présent préciser l'itération de l'algorithme OQS local (c'est-à-dire sans technique de globalisation de la convergence).

Algorithme 15.1 (OQS local) Une itération passe de l'itéré courant $(x_k,\lambda_k)\in\mathbb{E}\times\mathbb{R}^m$ à l'itéré suivant (x_{k+1},λ_{k+1}) par les étapes suivantes :

1. *Test d'arrêt* : si le couple (x_k,λ_k) est satisfaisant, on s'arrête.
2. *Nouvel itéré* : On prend comme nouvel itéré (x_{k+1},λ_{k+1}) une solution primale duale du problème quadratique osculateur (15.7).

L'algorithme OQS décompose donc la réslution du problème d'optimisation non-linéaire (P_{EI}) en une suite de problème d'optimisation quadratique, plus imples à résoudre que le problème original.

15.1.2 Convergence locale

Définition 15.2 On dit qu'un couple stationnaire (x_*,λ_*) de (P_{EI}) est *semi-stable* (resp. *hémi-stable*) si ce couple est une solution semi-stable (resp. hémi-stable) de l'inclusion dans (15.2a) avec F et K donnés par (15.2b).

Si (x_*, λ_*) vérifie les conditions de KKT, on a

$$F(x_*, \lambda_*) = \begin{pmatrix} 0_{\mathbb{E}} \\ 0_{\mathbb{R}^{m_E}} \\ -c_I(x_*) \end{pmatrix}, \tag{15.8}$$

$$F(x_*, \lambda_*) + F'(x_*, \lambda_*)(d, \mu) = \begin{pmatrix} L_* d + c'(x_*)^* \mu \\ -c'_E(x_*) d \\ -c_I(x_*) - c'_I(x_*) d \end{pmatrix}. \tag{15.9}$$

où l'on a utilisé (15.4) et posé $L_* := L(x_*, \lambda_*) := \nabla^2_{xx} \ell(x_*, \lambda_*)$. La condition $F(x_*, \lambda_*) + F'(x_*, \lambda_*)(d, \mu) + \mathrm{N}_K(x_*, \lambda_*) \ni 0$ s'écrit

$$\begin{gathered} L_* d + c'(x_*)^* \mu = 0, \\ c'_{E \cup I_*^{0+}}(x_*) d = 0, \quad c'_{I_*^{00}}(x_*) d \leqslant 0, \quad c_{I \setminus I_*^0}(x_*) + c'_{I \setminus I_*^0}(x_*) d \leqslant 0. \end{gathered}$$

Le résultat suivant analyse la semi-stabilité d'un couple stationnaire pour cette inclusion.

Proposition 15.3 (semi-stabilité d'une solution primale-duale) *Si $z_* := (x_*, \lambda_*)$ est une solution primale-duale de (P_{EI}), les propriétés suivantes sont équivalentes :*
(i) z_* est une solution primale-duale semi-stable,
(ii) $(d, \mu) = 0$ est l'unique solution du système

$$L_* d + c'(x_*)^* \mu = 0, \tag{15.10a}$$
$$c'_{E \cup I_*^{0+}}(x_*) d = 0, \tag{15.10b}$$
$$\mu_{I_*^{0+}} \geqslant 0, \quad 0 \leqslant \mu_{I_*^{00}} \perp c'_{I_*^{00}}(x_*) d \leqslant 0 \quad \text{et} \quad \mu_{I \setminus I_*^0} = 0. \tag{15.10c}$$

Démonstration. Par l'équivalence $(i) \Leftrightarrow (ii)$ de la proposition **??**, la semi-stabilité de (x_*, λ_*) est équivalente à

$$\begin{gathered} (d, \mu) = 0 \text{ est solution isolée de} \\ L_* d + c'(x_*)^* \mu = 0 \\ c'_E(x_*) d = 0 \\ c_I(x_*) + c'_I(x_*) d \in \mathrm{N}_{\mathbb{R}^{m_I}_+}(\lambda_* + \mu), \end{gathered} \tag{15.11}$$

où l'on a tenu compte du fait que $\nabla_x \ell(x_*, \lambda_*) = 0$ et $c_E(x_*) = 0$. Il revient au même de dire que les *petites* solutions du système ci-dessus sont nulles. Observons que la dernière condition ci-dessus, celle qu'il faut exprimer autrement, s'écrit aussi

$$0 \leqslant \big(\lambda_* + \mu\big) \perp \big(c_I(x_*) + c'_I(x_*) d\big) \leqslant 0. \tag{15.12}$$

Exploitons cette condition.

- Si $i \in I_*^{0+}$, $c_i(x_*) = 0$ et $(\lambda_*)_i > 0$, si bien que $(\lambda_* + \mu)_i > 0$ pour μ petit et la complémentarité dans (15.12) implique que $c'_i(x_*) d = 0$

- Pour les indices dans I_*^{00}, $c_i(x_*) = 0$ et $(\lambda_*)_i = 0$, si bien que (15.12) implique que $0 \leqslant \mu_{I_*^{00}} \perp c'_{I_*^{00}}(x_*)d \leqslant 0$.
- Si $i \in I \setminus I_*^0$, $c_i(x_*) < 0$ et $(\lambda_*)_i = 0$, si bien que pour d suffisamment petit $c_i(x_*)+c'_i(x_*)d < 0$ et donc la complémentarité dans (15.12) implique que $\mu_i = 0$.

On a donc montré qu'une *petite* solution du système dans (15.11) est solution de (15.10). On montre de la même manière qu'une *petite* solution de (15.10) est solution du système dans (15.11). Dès lors (15.11) revient à dire que $(d, \mu) = 0$ est solution isolée de (15.10).

On conclut en observant que l'ensemble des solutions de (15.10) est un cône, si bien qu'il revient au même de dire que $(d, \mu) = 0$ est solution isolée de (15.10) ou que $(d, \mu) = 0$ est l'unique solution de (15.10). □

On rappelle que le cône critique C_* en une solution locale de (P_{EI}) a été défini en (4.45a) et que L_* est une écriture simplifiée pour la hessienne $\nabla^2_{xx}\ell(x_*, \lambda_*)$ du lagrangien ℓ évaluée en $(x_*, \lambda_*) \in \mathbb{E} \times \mathbb{R}^m$.

Proposition 15.4 (semi-stabilité d'un minimum local) *Si x_* est un minimum local de (P_{EI}) et λ_* est un multiplicateur optimal associé, alors les propriétés suivantes sont équivalentes :*
(i) (x_*, λ_*) est semi-stable,
(ii) λ_* est l'unique multiplicateur associé à x_* et les conditions suffisantes du second ordre sont satisfaites, c'est-à-dire, $\forall\, d \in C_* \setminus \{0\}$, on a $\langle L_* d, d\rangle > 0$.

Démonstration. 1) On exploite l'équivalence $(i) \Leftrightarrow (iii)$ de la proposition **??** pour obtenir une autre expression de la semi-stabilité de (x_*, λ_*). Dans ce but, on pose

$$(d, \mu) := z - z_* = (x, \lambda) - (x_*, \lambda_*)$$

et on observe qu'avec F définie en (15.2b), (15.8) et (15.9), on a

$$\langle F'(x_*, \lambda_*)(d, \mu), (d, \mu)\rangle = \langle L_* d, d\rangle,$$
$$\langle F(x_*, \lambda_*), (d, \mu)\rangle = 0 \quad\Longleftrightarrow\quad \mu_{I\setminus I_*^0} = 0.$$

Par l'équivalence $(i) \Leftrightarrow (iii)$ de la proposition **??**, la semi-stabilité de (x_*, λ_*) est équivalente à

$$\text{on a } \langle L_* d, d\rangle > 0 \text{ pour tout } (d, \mu) \in \mathbb{E} \times \mathbb{R}^m \text{ non nul tel que} \tag{15.13a}$$
$$(\mu + \lambda_*)_I \geqslant 0, \quad \mu_{I\setminus I_*^0} = 0, \tag{15.13b}$$
$$L_* d + c'_{E\cup I_*^0}(x_*)^* \mu = 0, \tag{15.13c}$$
$$d \in C_*, \quad c_{I\setminus I_*^0}(x_*) + c'_{I\setminus I_*^0}(x_*)d \leqslant 0. \tag{15.13d}$$

où $C_* := \{d \in \mathbb{E} : c'_{E\cup I_*^{0+}}(x_*)d = 0,\ c'_{I_*^{00}}(x_*)d \leqslant 0\}$ est le cône critique.

2) $[(i) \Rightarrow (ii)]$ Comme (x_*, λ_*) est semi-stable, c'est une solution *isolée* du système d'optimalité 4.33 (proposition **??**). Par ailleurs, l'ensemble des multiplicateurs associé à x_* est convexe, λ_* est nécessairement l'unique multiplicateur associé à x_*.

L'ensemble des multiplicateurs associés à x_* étant borné (c'est un singleton), les conditions de qualification (QC-MF) ont lieu. On déduit alors de l'optimalité locale de x_* que $\langle L_* d, d\rangle \geqslant 0$ pour toute direction critique $d \in C_*$ (théorème 4.48). Donc si la conclusion n'a pas lieu, il existe une direction $d_1 \in C_* \setminus \{0\}$ telle que $\langle L_* d_1, d_1\rangle = 0$. Cette direction d_1 est donc solution de

$$\begin{cases} \min \langle L_* d, d\rangle \\ c'_{E\cup I_*^{0+}}(x_*)d = 0 \\ c'_{I_*^{00}}(x_*)d \leqslant 0, \end{cases}$$

problème dont les contraintes définissent les directions critiques. Les conditions d'optimalité de ce problème, aux contraintes qualifiées par (QC-A), affirment l'existence d'un multiplicateur $\mu_1 \in \mathbb{R}^m$ tel que l'on ait

$$\begin{gathered} (\mu_1)_{I\setminus I_*^0} = 0, \\ L_* d_1 + c'(x_*)^* \mu_1 = 0, \\ c'_{E\cup I_*^{0+}}(x_*)d_1 = 0, \\ 0 \leqslant (\mu_1)_{I_*^{00}} \perp c'_{I_*^{00}}(x_*)d_1 \leqslant 0. \end{gathered}$$

Alors $(d, \mu) = t(d_1, \mu_1)$, avec $t > 0$ suffisamment petit, vérifie (15.13b)-(15.13d) mais pas la conclusion $\langle L_* d, d\rangle > 0$ dans (15.13a). Cette contradiction montre que les CS2 sont vérifiées.

3) $[(i) \Leftarrow (ii)]$ Évidemment, si les CS2 ont lieu, x_* est un minimum local de (P_{EI}) (théorème 4.49).

Pour montrer la semi-stabilité de (x_*, λ_*), on montre que (15.13) a lieu. Soit (d, μ) non nul vérifiant (15.13b)-(15.13d). Il suffit de montrer que $d \neq 0$, car alors $d \in C_* \setminus \{0\}$ et la conclusion $\langle L_* d, d\rangle > 0$ de (15.13a) s'obtient par la CS2 supposée dans (ii). Raisonons par l'absurde en supposant que $d = 0$. Alors $\mu \neq 0$ et il est facile de voir que $\lambda_* + \mu$ serait un autre multiplicateur optimal associé à x_*, ce qui contredirait l'unicité supposée de celui-ci. □

Proposition 15.5 (condition suffisante d'hémi-stabilité) *Si x_* est un minimum local de (P_{EI}) et si (x_*, λ_*) est une solution semi-stable du système d'optimalité (4.33), alors (x_*, λ_*) est hémi-stable.*

Démonstration. Il s'agit de montrer que, pour tout $\alpha > 0$ donné, on peut trouver une constante $\beta > 0$, telle que, quel que soit $(x_0, \lambda_0) \in \bar{B}((x_*, \lambda_*), \beta)$, l'inclusion en (x, λ) suivante

$$\begin{pmatrix} \nabla f(x_0) + c'(x_0)^* \lambda_0 \\ -c(x_0) \end{pmatrix} + \begin{pmatrix} L(x_0, \lambda_0) & c'(x_0)^* \\ -c'(x_0) & 0 \end{pmatrix} \begin{pmatrix} x - x_0 \\ \lambda - \lambda_0 \end{pmatrix} + \mathrm{N}_K(x, \lambda) \ni 0$$

a une solution dans $\bar{B}((x_*, \lambda_*), \alpha)$. Cette inclusion est le système d'optimalité du premier ordre du problème quadratique en $x \in \mathbb{E}$ suivant

$$\begin{cases} \min \langle \nabla f(x_0), x - x_0 \rangle + \frac{1}{2}\langle L(x_0, \lambda_0)(x - x_0), x - x_0 \rangle \\ c_E(x_0) + c'_E(x_0)(x - x_0) = 0 \\ c_I(x_0) + c'_I(x_0)(x - x_0) \leqslant 0. \end{cases} \tag{15.14}$$

Celui-ci est une perturbation du problème quadratique que l'on obtient en prenant $(x_0, \lambda_0) = (x_*, \lambda_*)$, à savoir

$$\begin{cases} \min \langle \nabla f(x_*), x - x_* \rangle + \frac{1}{2}\langle L_*(x - x_*), x - x_* \rangle \\ c_E(x_*) + c'_E(x_*)(x - x_*) = 0 \\ c_I(x_*) + c'_I(x_*)(x - x_*) \leqslant 0. \end{cases} \tag{15.15}$$

Observons que le problème non perturbé (15.15) admet (x_*, λ_*) comme solution primale-duale, parce que son système d'optimalité du premier ordre en $x = x_*$ n'est autre que celui de (P_{EI}), vérifié par (x_*, λ_*), et parce que ses conditions d'optimalité du deuxième ordre sont également celles de (P_{EI}), qui sont satisfaites par hypothèses. Par ailleurs, la semi-stabilité entraîne l'unicité du multiplicateur (convexité de l'ensemble des multiplicateurs optimaux et remarque **??**(3)) et donc aussi les conditions de qualification de Mangasarian-Fromovitz (QC-MF) (proposition 4.45). On sait alors [514 ; théorème 4.2 et corollaire 4.3] que l'on peut trouver une solution du système perturbé (15.14) dont l'écart à (x_*, λ_*) est majoré par une constante fois la grandeur de la perturbation. □

Théorème 15.6 (convergence de l'algorithme SQP) *Si f et c sont $C^{2,1}$ dans la voisinage d'un minimum local x_* de (P_{EI}), s'il existe un unique multiplicateur optimal associé à x_*, et si les conditions suffisantes du second ordre sont vérifiées, alors il existe un voisinage V de (x_*, λ_*) tel que si le premier itéré $(x_1, \lambda_1) \in V$,*

1) *l'algorithme SQP peut générer une suite $\{(x_k, \lambda_k)\}$ dans V,*
2) *$\{(x_k, \lambda_k)\}$ converge quadratiquement vers (x_*, λ_*).*

Démonstration. Par la proposition 15.4, l'unicité du multiplicateur optimal et les conditions suffisantes du deuxième ordre, (x_*, λ_*) est une solution semi-stable de (15.2). Par la proposition 15.5, c'est aussi une solution hémi-stable. On peut alors appliquer le théorème **??** qui donne le résultat. □

15.2 L'algorithme local

15.2.1 Un algorithme non convergent

Considérons d'abord le problème avec contraintes d'égalité seulement

$$(P_E) \qquad \begin{cases} \min\ f(x) \\ c(x) = 0, \end{cases}$$

où $f : \mathbb{E} \to \mathbb{R}$ et $c : \mathbb{E} \to \mathbb{R}^m$.

On a vu au chapitre 10 comment on pouvait introduire la méthode de Newton pour résoudre des équations non linéaires (voir (10.3)) et pour minimiser une fonction (voir (10.7)). Pour résoudre le problème (P_E), on pourrait donc être tenté d'obtenir le déplacement d_k en x_k en prenant une approximation quadratique du critère et en linéarisant les contraintes en x_k. Avec cette méthode, on calculerait d_k comme solution du problème quadratique

$$\left\{\begin{array}{l} \min \nabla f(x_k)^\mathsf{T} d + \frac{1}{2} d^\mathsf{T} \nabla^2 f(x_k) d \\ c_E(x_k) + c'_E(x_k) \cdot d = 0 \end{array}\right. \tag{15.16}$$

et on prendrait $x_{k+1} = x_k + d_k$. Attention, **cet algorithme n'est pas nécessairement convergent** ! On connaît des exemples (voir l'exercice 15.1) dans lesquels la solution est répulsive pour cet algorithme : on peut trouver des itérés aussi proches que l'on veut de la solution (mais différents de la solution), tels que l'itéré suivant est plus éloigné de la solution que l'itéré en question. En fait, si on écrit l'algorithme comme un processus de point fixe, $x_{k+1} = \Phi(x_k)$ pour une certaine fonction Φ, l'application Φ n'est pas *strictement contractante* proche de la solution. Hélas, on voit encore parfois cette approche utilisée de nos jours !

Pourtant il doit bien exister une *méthode newtonienne*, c'est-à-dire qui procède par linéarisations, pour résoudre les problèmes avec contraintes ! Les chercheurs en optimisation numérique se sont posés la question bien longtemps et ce n'est qu'au milieu des années 1970, soit 30 ans après l'invention de l'algorithme du simplexe, que la situation s'est éclaircie. Ce qui ne fonctionne pas dans l'approche qui a conduit à (15.16), c'est qu'en traitant séparément les problèmes contradictoires « min $f(x)$ » et « $c(x) = 0$ » (le minimum de f ne vérifie en général pas la contrainte $c(x) = 0$, sinon pourquoi la spécifier), on a mis en place deux algorithmes qui ne se concertent pas et tirent à hue et à dia. Cela paraît à présent évident, la bonne démarche est de résoudre un seul système qui détermine les solutions de (P_E) et cela par un unique algorithme. Ce système est celui formé par les conditions d'optimalité de (P_E) et l'algorithme est celui de Newton. Ce système comporte en effet $n+m$ équations et $n+m$ inconnues (x_*, λ_*), ce qui en fait un candidat convenable pour être résolu par des itérations de Newton. Nous n'allons pas écrire cet algorithme, qui est en fait un cas particulier de celui que nous introduisons à la section suivante.

15.2.2 L'algorithme OQS

La discussion de la section 15.2.1 nous a montré qu'il était judicieux de faire des itérations de Newton sur le système d'optimalité de (P_{EI}). Rappelons (théorème 4.32) que celui-ci détermine les points stationnaires (x_*, λ_*) du problème comme solution du système de Karush, Kuhn et Tucker :

$$\text{(KKT)} \quad \left\{\begin{array}{ll} (a) & \nabla f(x_*) + A(x_*)^\mathsf{T} \lambda_* = 0 \\ (b) & c_E(x_*) = 0, \quad c_I(x_*) \leqslant 0 \\ (c) & (\lambda_*)_I \geqslant 0 \\ (d) & (\lambda_*)_I^\mathsf{T} c_I(x_*) = 0. \end{array}\right. \tag{15.17}$$

Travailler sur ce système d'optimalité n'allait pas de soi et l'on peut dire qu'il a fallu franchir un saut conceptuel pour y arriver.

- D'abord, ces relations forment un système en (x_*, λ_*), pas seulement en x_*. Il faut donc les linéariser par rapport à (x, λ), pas seulement par rapport à x comme dans l'algorithme qui a conduit à (15.16). L'algorithme de Newton qui en résultera sera donc une *méthode primale-duale*, générant à la fois des itérés primaux x_k et duaux λ_k.
- Une autre difficulté provient de la présence d'inégalités dans (15.17). La linéarisation de l'inégalité $F(x) \leqslant 0$ en x, pour un accroissement d, se fera par $F(x) + F'(x) \cdot d \leqslant 0$. On verra par les résultats de convergence locale obtenus qu'il s'agit d'un bon choix.

Soit $(x, \lambda) \in \mathbb{E} \times \mathbb{R}^m$ le point courant auquel on linéarise (15.17) et $(d, \mu) \in \mathbb{E} \times \mathbb{R}^m$ l'accroissement, la correction, que l'on désire apporter au point courant. L'itéré suivant (x_+, λ_+) s'obtiendra donc par

$$x_+ := x + d \qquad \text{et} \qquad \lambda_+ := \lambda + \mu.$$

On obtient le système suivant :

$$\begin{cases} L(x,\lambda)d + A(x)^{\mathsf{T}}\mu = -\nabla_x \ell(x,\lambda) \\ (c(x) + A(x)d)^{\#} = 0 \\ (\lambda + \mu)_I \geqslant 0 \\ (\lambda + \mu)_I^{\mathsf{T}} c_I(x) + \lambda_I^{\mathsf{T}} (A(x)d)_I = 0. \end{cases} \tag{15.18}$$

On a noté $A(x) := c'(x)$ la jacobienne des contraintes et

$$L(x, \lambda) := \nabla^2_{xx} \ell(x, \lambda)$$

la hessienne du lagrangien. Bien que linéaire, le système (15.18) a l'inconvénient d'être très difficile à résoudre. On y trouve en effet des égalités *et* des inégalités. De plus, on ne voit plus le problème d'optimisation dont il est issu. On se rappelle, en effet, que dans le cas de l'optimisation sans contrainte, le pas de Newton pouvait se voir comme un point stationnaire d'un problème d'optimisation quadratique sans contrainte. On aimerait pouvoir faire de même ici.

Dans ce but, on modifie le système (15.18), de manière à ce qu'il soit le système d'optimalité d'un problème quadratique avec contraintes linéaires. Si on ajoute le terme $\mu^{\mathsf{T}} A(x) d$ à la dernière équation, celle-ci ressemble davantage à des conditions de complémentarité (c'est alors un produit de deux facteurs). On observera par ailleurs, que dans le voisinage d'une solution primale-duale, le terme $\mu^{\mathsf{T}} A(x) d$ est formé d'un produit de deux grandeurs très petites, les accroissements μ et d. L'ajout du terme $\mu^{\mathsf{T}} A(x) d$ apporte donc une perturbation insignifiante au système (15.18), qui ne devrait pas modifier les propriétés de convergence rapide de l'algorithme (cette intuition s'avérera correcte). Le système devient alors

$$\begin{cases} L(x,\lambda)d + A(x)^{\mathsf{T}}\mu = -\nabla_x \ell(x,\lambda) \\ (c(x) + A(x)d)^{\#} = 0 \\ (\lambda + \mu)_I \geqslant 0 \\ (\lambda + \mu)_I^{\mathsf{T}} (c(x) + A(x)d)_I = 0. \end{cases} \tag{15.19}$$

On vérifie aisément le fait remarquable suivant : (15.19) est formé des conditions de KKT du problème d'optimisation quadratique

$$\begin{cases} \min_d \ \nabla f(x)^\mathsf{T} d + \frac{1}{2} d^\mathsf{T} L(x,\lambda) d \\ c_E(x) + A_E(x)d = 0 \\ c_I(x) + A_I(x)d \leqslant 0. \end{cases} \tag{15.20}$$

Plus précisément, le lien entre l'accroissement μ de λ et le multiplicateur λ^{PQ} associé aux contraintes de (15.20) s'écrit $\lambda^{\mathrm{PQ}} = \lambda + \mu$. Dès lors la mise à jour du couple (x, λ) se fait par

$$x_+ := x + d \qquad \text{et} \qquad \lambda_+ := \lambda^{\mathrm{PQ}}. \tag{15.21}$$

L'algorithme qui met à jour (x, λ) par ces formules est appelé l'*algorithme OQS* (pour Sequential Quadratic Programming). Le problème (15.20) que l'on doit résoudre à chaque itération pour déterminer $(d, \lambda^{\mathrm{PQ}})$ est appelé le *problème quadratique osculateur* (PQO) au problème (P_{EI}) en (x, λ). On le déduit aisément de (P_{EI}). Ses contraintes sont celles de (P_{EI}), linéarisées en x. Son critère est hybride, avec $\nabla f(x)$ dans la partie linéaire et la hessienne du lagrangien dans la partie quadratique. Ce qui manque au problème quadratique (15.16) est maintenant manifeste : ce problème ne fait pas intervenir la courbure des contraintes (les dérivées secondes de c sont absentes). Dans le problème quadratique osculateur (15.20), cette courbure intervient dans la hessienne du lagrangien, pas par une approximation quadratique des contraintes.

On peut à présent résumer l'algorithme OQS local.

Algorithme 15.7 (OQS local — une itération)

On suppose que l'on dispose au début de l'itération d'un couple $(x, \lambda) \in \mathbb{E} \times \mathbb{R}^m$.

1. *Test d'arrêt* : si (x, λ) vérifie les conditions d'optimalité (15.17), arrêt de l'algorithme.
2. *Déplacement* : calculer $(d, \lambda^{\mathrm{PQ}})$ comme point stationnaire du problème quadratique osculateur (15.20).
3. *Mise à jour des variables* : le nouvel itéré (x_+, λ_+) est donné par (15.21).

Le coût d'une itération est essentiellement dû à la résolution du problème quadratique osculateur. On a donc reporté la *combinatoire* du problème (P_{EI}) sur des problèmes quadratiques, dans lesquels elle est plus facilement prise en compte. Cependant, plusieurs difficultés peuvent se présenter :

- La hessienne du lagrangien n'est pas nécessairement semi-définie positive, si bien que le PQO est NP-ardu (voir le chapitre **??**) et peut présenter des *solutions importunes* (indésirables car trop grandes, nous reviendrons sur cette question à la section 15.2.3). Pour cette raison, les implémentations de cet algorithmes utilisent souvent une modification ou une approximation définie positive M_k de la hessienne de $L(x_k, \lambda_k)$. S'il est réalisable, c'est-à-dire si ses contraintes sont compatibles, le PQO peut alors être résolu en un nombre polynomial d'itérations et n'a plus de solutions indésirables.
- Le PQO peut ne pas être borné (donc ne pas avoir de solution). L'introduction de l'algorithme nous a montré qu'un point stationnaire suffirait, mais il n'y a pas

d'algorithme évident pour trouver un point stationnaire d'un problème quadratique autre qu'un minimum local.

- Les contraintes linéarisées du PQO peuvent être *incompatibles*. Diverses techniques permettent de faire face à cette difficulté : relaxation des contraintes linéarisées dans la globalisation de la convergence par recherche linéaire; la globalisation de la convergence par régions de confiance prend directement en compte cette difficulté.

15.2.3 Convergence locale ▲

Nous présentons ci-dessous le résultat de convergence le plus simple, celui qui se démontre en se ramenant au cas des problèmes avec contraintes d'égalité. Pour un résultat plus fin, nous renvoyons le lecteur à [71, 72, 73].

Commençons donc par considérer le problème où il n'y a que des contraintes d'égalité :

$$(P_E) \qquad \begin{cases} \min\ f(x) \\ c(x) = 0. \end{cases}$$

Dans ce cas, l'algorithme OQS est l'algorithme de Newton (10.2)–(10.3) appliqué au système d'optimalité

$$F(z) = 0, \quad \text{où} \quad z = (x, \lambda) \quad \text{et} \quad F(z) = \begin{pmatrix} \nabla_x \ell(x, \lambda) \\ c(x) \end{pmatrix}.$$

On a noté $(x, \lambda) \mapsto \ell(x, \lambda) = f(x) + \lambda^{\mathsf{T}} c(x)$ le lagrangien associé au problème (P_E). Le résultat de convergence locale est alors un corrolaire immédiat du théorème 10.2, qui requiert une hypothèse de différentiabilité de F et d'inversibilité de $F'(x_*, \lambda_*)$, ce qui conduit à la définition suivante.

Définition 15.8 (point stationnaire régulier) Un point stationnaire (x_*, λ_*) de (P_E) est dit *régulier* si la matrice d'ordre $n + m$

$$\begin{pmatrix} \nabla^2_{xx} \ell(x_*, \lambda_*) & c'(x_*)^* \\ c'(x_*) & 0 \end{pmatrix} \tag{15.22}$$

est inversible.

Un point stationnaire (x_*, λ_*) de (P_E) vérifiant les conditions d'optimalité du second ordre du théorème 4.24 et tel que $c'(x_*)$ est surjective est régulier. En effet, il suffit de montrer que la matrice carrée (15.22) est injective pour en conclure qu'elle est inversible. Or si (d, μ) est dans son noyau, alors $d \in \mathcal{N}(c'(x_*))$ et $d^{\mathsf{T}} \nabla^2_{xx} \ell(x_*, \lambda_*) d = 0$. Donc $d = 0$ par les conditions du second ordre. Ensuite $\mu = 0$ par l'injectivité de $c'(x_*)^*$.

Théorème 15.9 (convergence quadratique locale de l'algorithme de Newton pour (P_E)) *Supposons que f et c soient de classe C^2 dans un voisi-*

nage d'un point stationnaire régulier x_* de (P_E), avec multiplicateur associé λ_*. Alors, il existe un voisinage V de (x_*, λ_*) tel que, si le premier itéré $(x_1, \lambda_1) \in V$, l'algorithme de Newton 15.7 est bien défini et génère une suite $\{(x_k, \lambda_k)\}$ convergeant superlinéairement vers (x_*, λ_*). Si f'' et c'' est lipschitzienne dans un voisinage de x_*, la convergence de la suite est quadratique.

Démonstration. On applique le théorème 10.2 avec F définie ci-dessus, $z = (x, \lambda)$ et $z_* = (x_*, \lambda_*)$. La jacobienne $F'(z_*)$ est inversible par la régularité de z_*. La convergence superlinéaire de $\{(x_k, \lambda_k)\}$ vers (x_*, λ_*) a lieu si (x_1, λ_1) est suffisamment proche de (x_*, λ_*). Si f'' et c'' son lipschitzienne proche de x_*, il en est de même de F' proche de z_*, et le convergence quadratique de $\{(x_k, \lambda_k)\}$ s'en suit. □

Une propriété minimale d'un algorithme convergent est de générer un déplacement nul lorsque l'itéré courant est une solution. Cette propriété élémentaire n'a pourtant pas nécessairement lieu lorsque le déplacement de l'algorithme OQS est une solution *arbitraire* du PQO, comme le montre l'exemple suivant.

Exemple 15.10 On cherche à minimiser le logarithme de $(1 + x)$ lorsque x est restreint à l'intervalle $[0, 3]$:

$$\begin{cases} \min_x \log(1 + x) \\ 0 \leqslant x \leqslant 3. \end{cases}$$

Le logarithme a été utilisé de manière à introduire de la non-convexité dans le problème (par la monotonicité du logarithmie, il aurait été équivalent de minimiser $(1+x)$ et le problème serait devenu linéaire). On vérifie aisément que le problème a une unique solution primale-duale $(x_*, \lambda_*) = (0, (1, 0))$, qui satisfait les conditions suffisantes d'optimalité, la complémentarité stricte et la qualification des contraintes (QC-IL). On peut donc arguer qu'il s'agit d'une « bonne » solution. Cependant, le problème quadratique osculateur (15.20) en cette solution s'écrit

$$\begin{cases} \min_d d - \frac{1}{2} d^2 \\ 0 \leqslant d \leqslant 3. \end{cases}$$

Ce PQO a trois points stationnaires primaux-duaux (d, λ) : à savoir un minimum local $(0, (1, 0))$, un maximum $(1, (0, 0))$ et un minimum global $(3, (0, 2))$. Parmi ces points stationnaires, seul le premier est satisfaisant puisqu'il donne un déplacement nul et le multiplicateur optimal. Les deux autres points stationnaires sont dits *importuns*, indésirables. □

Le phénomène qui se produit dans cet exemple ne peut avoir lieu que si $L(x, \lambda)$ n'est pas définie positive. Sinon, le PQO est strictement convexe et a donc une unique solution dès que ses contraintes sont compatibles. Les résultats de convergence de l'algorithme OQS doivent donc faire une hypothèse sur la solution du PQO qui est sélectionnée par l'algorithme à chaque itération. Dans le résultat donné ci-dessous, il est supposé que d est une *solution de norme minimale*.

The next lemma is useful for proving theorem 15.12. We use the notation that is suitable for the present framework. Its proof is proposed in exercise **??**.

Lemme 15.11 (propriété de la surjectivité) *Let A be an $m \times n$ surjective matrix and $\mu \in \mathbb{R}^m$. Then, $\forall \varepsilon > 0$, $\exists \delta > 0$, such that when the $m \times n$ matrix $\tilde{A}$ and the point $\tilde{\mu} \in \mathbb{R}^m$ satisfy $\|\tilde{A} - A\| \leqslant \delta$ and $\|\tilde{A}^\mathsf{T}\tilde{\mu} - A^\mathsf{T}\mu\| \leqslant \delta$, there must hold $\|\tilde{\mu} - \mu\| \leqslant \varepsilon$.*

Théorème 15.12 (convergence quadratique primale-duale de l'algorithme OQS) *Suppose that f and c are of class C^2 in a neighborhood of a stationary point x_* of (P_{EI}), with associated multiplier λ_*. Suppose also that strict complementarity holds and that $(x_*, (\lambda_*)_{E\cup I_*^0})$ is a regular stationary point of the equality constrained problem*

$$\begin{cases} \min_x \ f(x) \\ c_i(x) = 0, \quad \text{for } i \in E \cup I_*^0, \end{cases} \tag{15.23}$$

*in the sense of definition **??**. Consider the OQS algorithm, in which d_k is a minimum-norm stationary point of the osculating quadratic problem (15.20). Then there is a neighborhood V of (x_*, λ_*) such that, if the first iterate $(x_1, \lambda_1) \in V$:*

- *(i) the OQS algorithm is well defined and generates a sequence $\{(x_k, \lambda_k)\}$ that converges superlinearly to (x_*, λ_*);*
- *(ii) the active constraints of the osculating quadratic problem (15.20) are those of problem (P_{EI});*
- *(iii) if, in addition, f and c are of class $C^{2,1}$ in a neighborhood of x_*, the convergence of $\{(x_k, \lambda_k)\}$ is quadratic.*

Démonstration. The idea of the proof is to show that, close to (x_*, λ_*), the selected minimum-norm stationary point of the osculating quadratic problem (15.20) is actually the primal-dual Newton step on (15.23). The result then follows from theorem 15.9.

Suppose that (x, λ) is close enough to (x_*, λ_*). Since $(x_*, (\lambda_*)_{E\cup I_*^0})$ is a regular stationary point of (15.23), $c'_{E\cup I_*^0}(x_*)$ is surjective and the osculating QP associated with (15.23), namely

$$\begin{cases} \min_{\mathrm{d}t} \ \nabla f(x)^\mathsf{T}\mathrm{d}t + \frac{1}{2}\mathrm{d}t^\mathsf{T} L(x,\lambda)\mathrm{d}t \\ c_i(x) + c'_i(x)\cdot \mathrm{d}t = 0, \quad \text{for } i \in E \cup I_*^0, \end{cases} \tag{15.24}$$

has a unique primal-dual stationary point. We denoted it by $(\mathrm{d}t, \tilde{\ell}_{E\cup I_*^0})$ and form with $\tilde{\ell}_{E\cup I_*^0}$ a vector $\tilde{\ell} \in \mathbb{R}^m$, by setting $\tilde{\ell}_i = 0$ for $i \in I \setminus I_*^0$.

Let us show that $(\mathrm{d}t, \tilde{\ell})$ is a stationary point of the osculating quadratic problem (15.20), if $(x, \lambda) := (x_k, \lambda_k)$ is in some neighborhood of (x_*, λ_*); this will imply that, for (x, λ) close to (x_*, λ_*), the OQS algorithm is well defined and that d is small (it is a minimum-norm stationary point and $\mathrm{d}t$ is small by theorem 15.9). We only need to show that

$$c_i(x) + c'_i(x)\cdot \mathrm{d}t \leqslant 0 \text{ for } i \in I \setminus I_*^0, \tag{15.25}$$

$$\lambda_i \geqslant 0 \text{ for } i \in I_*^0. \tag{15.26}$$

From theorem 15.9, $(x + \mathrm{d}t, \tilde{\ell})$ is close to (x_*, λ_*), when (x, λ) is close to (x_*, λ_*), so that $\mathrm{d}t$ is small. Now (15.25) follows when (x, λ) close enough to (x_*, λ_*), since $c_i(x_*) < 0$ for $i \in I \setminus I_*^0$. On the other hand, to prove (15.26), observe that, for $i \in I_*^0$, $(\lambda_*)_i > 0$ by strict complementarity. Therefore $\tilde{\ell}_i \geqslant 0$ since, by theorem 15.9, $\tilde{\ell}$ is close to λ_* when (x, λ) is close to (x_*, λ_*).

Let us now show that the pair $(d, \lambda^{\mathrm{PQ}}) := (d_k, \lambda_k^{\mathrm{PQ}})$ formed of the selected minimum-norm stationary point of the osculating QP and its associated multiplier is in fact $(\mathrm{d}t, \tilde{\ell})$, if (x, λ) is in some neighborhood of (x_*, λ_*). Since (15.24) has a single stationary point, for (x, λ) close to (x_*, λ_*), we have to show that $(d, \lambda^{\mathrm{PQ}})$ satisfies its optimality conditions. Knowing that $(d, \lambda^{\mathrm{PQ}})$ satisfies the optimality conditions of the osculating quadratic problem (15.20), we just have to show that

$$\lambda_i = 0 \text{ for } i \in I \setminus I_*^0, \tag{15.27}$$

$$c_i(x) + c_i'(x) \cdot d = 0 \text{ for } i \in I_*^0. \tag{15.28}$$

Condition (15.27) is a consequence of the complementarity in (15.20) and of the fact that, for (x, λ) close to (x_*, λ_*) and $i \in I \setminus I_*^0$, $c_i(x) + c_i'(x) \cdot d < 0$ (since $c_i(x_*) < 0$ and d is small). Condition (15.28) results from the complementarity in (15.20) and the fact that $\lambda_i > 0$ for $i \in I_*^0$. For the latter claim, observe indeed that

$$\nabla f(x) + L(x, \lambda) d + \sum_{i \in E \cup I_*^0} \lambda_i \nabla c_i(x) = 0.$$

Since $\nabla f(x) + L(x, \lambda) d$ is close to $\nabla f(x_*)$ and the gradients $\{\nabla c_i(x) : i \in E \cup I_*^0\}$ are close to the linearly independent gradients $\{\nabla c_i(x_*) : i \in E \cup I_*^0\}$, lemma 15.11 implies that $\lambda_{I_*^0}$ is close to $(\lambda_*)_{I_*^0} > 0$.

Because of (15.28) and the fact that $c_i(x) + c_i'(x) \cdot d < 0$ for $i \in I \setminus I_*^0$ (see the reasonning after (15.28)), we have also proven the claim (ii) of the theorem. □

15.3 Globalisation de la convergence par recherche linéaire

On écrit le problème quadratique osculateur en l'itéré (x_k, λ_k) de l'itération k comme suit

$$\begin{cases} \min_d \; g_k^\mathsf{T} d + \frac{1}{2} d^\mathsf{T} M_k d \\ (c_k + A_k d)^\# = 0. \end{cases} \tag{15.29}$$

Nous avons abrégé les écritures en posant

$$g_k = \nabla f(x_k), \quad c_k = c(x_k) \quad \text{et} \quad A_k = A(x_k) = c'(x_k).$$

La matrice M_k est soit $L(x_k, \lambda_k)$ ou une approximation de cette hessienne (version quasi-newtonienne). Un point stationnaire d_k de ce problème vérifie, pour un certain multiplicateur $\lambda_k^{\mathrm{PQ}} \in \mathbb{R}^m$, les conditions d'optimalité suivantes

$$\begin{cases} (a) & g_k + M_k d_k + A_k^\mathsf{T} \lambda_k^{\mathrm{PQ}} = 0 \\ (b) & (c_k + A_k d_k)^\# = 0 \\ (c) & (\lambda_k^{\mathrm{PQ}})_I \geqslant 0 \\ (d) & (\lambda_k^{\mathrm{PQ}})_I^\mathsf{T} (c_k + A_k d_k)_I = 0. \end{cases} \tag{15.30}$$

15.3.1 Fonction de mérite

Afin de globaliser la méthode de Newton, on introduit la fonction de mérite (ou fonction de pénalisation) suivante

$$\Theta_\sigma(x) = f(x) + \sigma\|c(x)\|_{\mathrm{P}}, \tag{15.31}$$

où $\|\cdot\|_{\mathrm{P}}$ est une norme sur $\mathbb{R}^m$ et $\sigma > 0$ est un paramètre. On supposera que la norme $\|\cdot\|_{\mathrm{P}}$ vérifie

$$v \mapsto \|v^\#\|_{\mathrm{P}} \text{ est convexe.} \tag{15.32}$$

C'est le cas des normes ℓ_p définies par (A.5), avec $p \in [1, +\infty]$. On note $\|\cdot\|_{\mathrm{D}}$, la norme duale de $\|\cdot\|_{\mathrm{P}}$ pour le produit scalaire euclidien (voir (A.8)).

15.3.2 Condition de décroissance de la fonction de mérite

Proposition 15.13 (décroissance de Θ_σ le long de d_k) *Si $(d_k, \lambda_k^{\mathrm{PQ}})$ vérifie les conditions d'optimalité (15.30) et si $\|\cdot\|_{\mathrm{P}}$ vérifie (15.32), alors*

$$\Theta'_\sigma(x_k; d_k) \leqslant g_k^\mathsf{T} d_k - \sigma\|c_k^\#\|_{\mathrm{P}} = -d_k^\mathsf{T} M_k d_k + (\lambda_k^{\mathrm{PQ}})^\mathsf{T} c_k - \sigma\|c_k^\#\|_{\mathrm{P}}. \tag{15.33}$$

Si, de plus,

$$\sigma \geqslant \|\lambda_k^{\mathrm{PQ}}\|_{\mathrm{D}}, \tag{15.34}$$

alors

$$\Theta'_\sigma(x_k; d_k) \leqslant -d_k^\mathsf{T} M_k d_k. \tag{15.35}$$

Donc $\Theta'_\sigma(x_k; d_k) < 0$, si $\sigma \geqslant \|\lambda_k^{\mathrm{PQ}}\|_{\mathrm{D}}$, si M_k est définie positive et si x_k n'est pas un point stationnaire de (P_{EI}).

Démonstration. Puisqu'une norme a des dérivées directionnelles (elle est convexe, donc la proposition 3.15 s'applique) et est lipschitzienne (on utilise sa convexité ou l'inégalité triangulaire), la fonction $v \to \|v^\#\|_{\mathrm{P}}$ a des dérivées directionnelles. Pour les calculer, on observe d'abord que, par (15.32) et (15.30)-(b), on a pour $t \in]0, 1[$:

$$\begin{aligned}
\left\|\left[c_k + tA_k d_k\right]^\#\right\|_{\mathrm{P}} &= \left\|\left[(1-t)c_k + t(c_k + A_k d_k)\right]^\#\right\|_{\mathrm{P}} \\
&\leqslant (1-t)\|c_k^\#\|_{\mathrm{P}} + t\left\|\left[c_k + A_k d_k\right]^\#\right\|_{\mathrm{P}} \\
&= (1-t)\|c_k^\#\|_{\mathrm{P}}.
\end{aligned}$$

Dès lors

$$(\|\cdot^\#\|_{\mathrm{P}})'(c_k; A_k d_k) = \lim_{t\downarrow 0} \frac{1}{t}\left(\left\|\left[c_k + tA_k d_k\right]^\#\right\|_{\mathrm{P}} - \|c_k^\#\|_{\mathrm{P}}\right) \leqslant -\|c_k^\#\|_{\mathrm{P}}.$$

Alors, en utilisant successivement (15.30)-(a), (15.30)-(b) et (15.30)-(d), on montre (15.33) :

$$\begin{aligned}\Theta'_\sigma(x_k;d_k) &\leqslant g_k^\mathsf{T} d_k - \sigma\|c_k^\#\|_{\mathrm{P}}\\ &= -d_k^\mathsf{T} M_k d_k - (\lambda_k^{\mathrm{PQ}})^\mathsf{T} A_k d_k - \sigma\|c_k^\#\|_{\mathrm{P}}\\ &= -d_k^\mathsf{T} M_k d_k + (\lambda_k^{\mathrm{PQ}})^\mathsf{T} c_k - \sigma\|c_k^\#\|_{\mathrm{P}}.\end{aligned}$$

Si $\sigma \geqslant \|\lambda_k^{\mathrm{PQ}}\|_{\mathrm{D}}$, en utilisant (15.30)-($c$) et l'inégalité de Cauchy-Schwarz généralisée (A.9), on trouve

$$(\lambda_k^{\mathrm{PQ}})^\mathsf{T} c_k - \sigma\|c_k^\#\|_{\mathrm{P}} \leqslant (\lambda_k^{\mathrm{PQ}})^\mathsf{T} c_k^\# - \sigma\|c_k^\#\|_{\mathrm{P}} \leqslant (\|\lambda_k^{\mathrm{PQ}}\|_{\mathrm{D}} - \sigma)\|c_k^\#\|_{\mathrm{P}} \leqslant 0.$$

Alors, (15.35) se déduit de (15.33). Si $\Theta'_\sigma(x_k;d_k)=0$ et M_k est définie positive, alors $d_k=0$. De (15.30), on déduit la stationnarité de x_k, avec λ_k^{PQ} comme multiplicateur associé. □

Le seuil sur σ dans (15.34) rappelle celui qui fait de Θ_σ une fonction de pénalisation exacte (proposition 13.24). Pour satisfaire l'inégalité (15.34), il va falloir modifier σ à certaines itérations (on n'est pas maître de λ_k^{PQ}). On note alors σ_k la valeur de σ à l'itération k et *avant* de faire de la recherche linéaire (RL) on modifie éventuellement l'ancien σ ($=\sigma_{k-1}$) de manière à avoir

$$\sigma_k \geqslant \|\lambda_k^{\mathrm{PQ}}\|_{\mathrm{D}} + \bar\sigma, \tag{15.36}$$

où $\bar\sigma > 0$ est une « petite » constante (choix heuristique). On utilisera les règles suivantes.

- Première itération. On prend

  $$\bar\sigma = \max(\sqrt{\texttt{eps}}, \|\lambda_1\|_{\mathrm{D}}/100) \quad \text{et} \quad \sigma_1 = \|\lambda_1\|_{\mathrm{D}} + \bar\sigma,$$

  où `eps` est l'epsilon-machine.
- Itérations suivantes. On augmente éventuellement σ de manière à réaliser (15.36), mais on peut aussi le faire décroître s'il s'avère être « beaucoup » trop grand (afin d'éviter la troncature du pas).

  si $\sigma_{k-1} < \|\lambda_k^{\mathrm{PQ}}\|_{\mathrm{D}} + \bar\sigma$;
  alors $\sigma_k = \max(1.5\,\sigma_{k-1}, \|\lambda_k^{\mathrm{PQ}}\|_{\mathrm{D}} + \bar\sigma)$;
  sinon
      si $\sigma_{k-1} > 1.1(\|\lambda_k^{\mathrm{PQ}}\|_{\mathrm{D}} + \bar\sigma)$;
      alors $\sigma_k = (\sigma_{k-1} + \|\lambda_k^{\mathrm{PQ}}\|_{\mathrm{D}} + \bar\sigma)/2$;
      sinon $\sigma_k = \sigma_{k-1}$.

Plus rien ne garantit la convergence « théorique » de l'algorithme si l'on fait décroître σ_k comme dans la règle ci-dessus, mais en pratique une telle heuristique peut améliorer grandement l'efficacité de l'algorithme. On ne s'en prive donc pas. Aux analystes d'en trouver une qui n'empêche pas la convergence !

15.4 Versions quasi-newtoniennes ▲

Dans les méthodes de quasi-Newton, on remplace la hessienne du lagrangien L qui intervient dans le problème quadratique osculateur (15.20) par une matrice M, symétrique *définie positive*, mise à jour par l'algorithme. Dans le cadre de OQS, cette approche a au moins deux intérêts : il ne faut pas calculer de dérivées secondes et le problème quadratique (PQ) est toujours « mieux » posé (il a au plus une solution). Bien que la convergence soit plus lente qu'avec la méthode de Newton (elle n'est plus que superlinéaire) et malgré les difficultés conceptuelles rencontrées par cette approche (voir plus loin), le second avantage cité ci-dessous fait que c'est essentiellement l'approche utilisée dans les codes commerciaux ou de recherche.

La version quasi-newtonienne de OQS génère donc une suite primale-duale $\{(x_k, \lambda_k)\}$ et une suite de matrices symétriques définies positives $\{M_k\}$ de la manière suivante. À l'étape k, on calcule d'abord *la* solution primale-duale $(d_k, \lambda_k^{\text{PQ}})$ du problème quadratique osculateur (en espérant qu'elle existe...)

$$\begin{cases} \min_{d\in\mathbb{E}}\ g_k^\mathsf{T} d + \frac{1}{2} d^\mathsf{T} M_k d \\ (c_k + A_k d)^\# = 0, \end{cases} \tag{15.37}$$

dans lequel M_k joue le rôle de $L_k = L(x_k, \lambda_k) = \nabla^2_{xx}\ell(x_k, \lambda_k)$, la hessienne du lagrangien ℓ, $A_E(x_k)$ et $A_I(x_k)$ sont les jacobiennes des contraintes d'égalité c_E et c_I. Ensuite on prend $x_{k+1} = x_k + d_k$, $\lambda_{k+1} = \lambda_k^{\text{PQ}}$ et on met à jour M_k par la formule de BFGS (c'est la formule la plus utilisée).

La formule de BFGS s'écrit :

$$M_{k+1} = M_k - \frac{M_k \delta_k \delta_k^\mathsf{T} M_k}{\delta_k^\mathsf{T} M_k \delta_k} + \frac{\gamma_k \gamma_k^\mathsf{T}}{\gamma_k^\mathsf{T} \delta_k}.$$

Les vecteurs γ_k et δ_k sont déterminés de manière à forcer M_{k+1} à être définie positive (on suppose que M_k l'est) et à être proche de L_{k+1}, ce qui, dans certains cas, peut être contradictoire. Pour cela, on prend pour δ_k le déplacement en x :

$$\delta_k = x_{k+1} - x_k.$$

Le vecteur γ_k devrait idéalement être la variation du gradient du lagrangien

$$\gamma_k^\ell = \nabla_x \ell(x_{k+1}, \lambda_{k+1}) - \nabla_x \ell(x_k, \lambda_{k+1}).$$

Mais, pour conserver la définie positivité de M_k, on doit avoir $\gamma_k^\mathsf{T}\delta_k > 0$, ce qui n'est pas garanti avec $\gamma_k = \gamma_k^\ell$. Si bien que l'on utilisera la *correction de Powell* qui consiste à prendre

$$\gamma_k = \theta_k \gamma_k^\ell + (1 - \theta_k) M_k \delta_k,$$

où θ_k est pris maximal dans $]0, 1]$ de manière à avoir $\gamma_k^\mathsf{T}\delta_k \geqslant 0.2\, \delta_k^\mathsf{T} M_k \delta_k$. On trouve

$$\theta_k = \begin{cases} 0.8\, \dfrac{\delta_k^\mathsf{T} M_k \delta_k}{\delta_k^\mathsf{T} M_k \delta_k - (\gamma_k^\ell)^\mathsf{T}\delta_k} & \text{si } (\gamma_k^\ell)^\mathsf{T}\delta_k < 0.2 \delta_k^\mathsf{T} M_k \delta_k \\ 1 & \text{sinon.} \end{cases}$$

Il reste à spécifier la matrice initiale. On prendra $M_1 = I$ (matrice identité) à la première itération. Mais, après le calcul de x_2 et avant le calcul de M_2, on modifie la valeur de M_1 en ηI, où η a une valeur reflétant l'échelle du problème (ou la « valeur moyenne » de L_1) :

$$\eta = \frac{\|\gamma_1\|_2^2}{\gamma_1^\mathsf{T}\delta_1}.$$

Il faut en effet attendre que la première itération soit terminée pour évaluer cette grandeur. Puis on calcule M_2 par la formule de BFGS.

15.5 Le diable se cache dans les détails ▲

Si l'on veut qu'un algorithme fonctionne bien, il faut soigner les « détails », qui n'en sont pas en réalité, mais qui ont été passé sous silence dans les développements précédents. De nombreuses questions d'apparence anodine doivent en effet être traitées avec le plus grand soin si l'on veut avoir une méthode numérique efficace, donnant de bons résultats sur des bans d'essai de problèmes-tests.

15.5.1 Incompatibilité des contraintes

Un problème d'optimisation non linéaire peut avoir des contraintes incompatibles, donc ne pas avoir de points admissibles, auquel cas le problème n'a évidemment pas de solution. Il est souhaitable toutefois que l'algorithme puisse détecter une telle situation. Dans l'algorithme OQS, cette situation se manifestera par l'intermédiaire du problème quadratique osculateur (PQO) qui pourra lui aussi présenter des contraintes linéaires incompatibles. Le lien entre l'incompatibilité des contraintes du problème non linéaire et celle des PQOs est toutefois complexe à étudier et on se contentera ici de présenter les méthodes qui ont été proposées pour traiter les PQOs inconsistants.

Méthodes possibles :

- La *pénalisation exacte des contraintes linéarisées* [211 ; 1982] consiste à remplacer le PQO par sa version pénalisée exacte suivante

$$\begin{cases} \min_d \nabla f(x)^\mathsf{T} d + \frac{1}{2} d^\mathsf{T} H d + \sigma \, \|(c(x) + A(x)d)^\#\|_1 \\ \|d\|_\infty \leqslant \Delta, \end{cases} \tag{15.38}$$

  dans laquelle $H \simeq \nabla^2_{xx}\ell(x,\lambda)$, $\sigma > 0$ est pris assez grand et la contrainte joue le rôle de région de confiance. Ce problème a toujours une solution, mais il est *non lisse* ; il peut se récrire sous la forme d'un problème quadratique standard, pourvu que l'on introduise des variables auxiliaires, en faisant passer le terme normé du critère en contrainte (exercice 1.8). Si $H \succ 0$, le direction d est de descente pour la fonction de mérite (15.31).

  – Avantage de l'approche : robustesse.

  – Inconvénient de l'approche : pas de solveur disponible pour résoudre un problème sous la forme (15.38); la transformation de (15.38) en problème quadratique standard introduit des variables auxiliaires.

- Le *mode élastique* [250 ; 2002] consiste à relaxer le PQO comme suit

$$\begin{cases} \min_{(d,v,w)} \nabla f(x)^\mathsf{T} d + \frac{1}{2} d^\mathsf{T} H d + \sigma \sum_{i \in E \cup I} (v_i + w_i) \\ c_E(x) + A_E(x)d - v_E + w_E = 0 \\ l \leqslant c_I(x) + A_I(x)d - v_I + w_I \leqslant u \\ v_I \geqslant 0 \\ w_I \geqslant 0, \end{cases}$$

  dans lequel $v = (v_E, v_I)$ et $w = (w_E, w_I)$ assurent la compatibilité des nouvelles contraintes. Par le terme de pénalisation linéaire $\sigma \sum_i (v_i + w_i)$, on essaye d'annuler (v,w) pour que, si possible, d soit une solution du PQO initial.

- Certaines approches algorithmiques de résolution du PQO, comme celle du lagrangien augmenté (section 13.4), calculent une solution de

$$\begin{cases} \min_d \; \nabla f(x)^\mathsf{T} d + \frac{1}{2}\, d^\mathsf{T} H d \\ c_E(x) + A_E(x) d + \bar{s}_E(x) = 0 \\ l \leqslant c_I(x) + A_I(x) d + \bar{s}_I(x) \leqslant u, \end{cases}$$

  où $\bar{s}(x) = (\bar{s}_E(x), \bar{s}_I(x))$ est le vecteur de norme euclidienne minimale rendant les contraintes du PQO compatibles [122 ; 2016]. La direction d ainsi calculée est de descente pour Θ_σ, pourvu que σ soit assez grand et $H \succcurlyeq 0$.

15.5.2 Troncature du pas : l'effet Maratos

15.5.3 Problèmes de commande optimale

Du point de vue de l'optimisation, un problème de commande optimale discrétisé se présente sous la forme

$$(P_{SI}) \qquad \begin{cases} \min \; f(x) \\ c_S(x) = 0 \\ c_I(x) \leqslant 0. \end{cases}$$

Il ressemble très fort au problème (P_{EI}), si ce n'est que la fonction $c_S : \mathbb{R} \to \mathbb{R}^{m_S}$, qui remplace la fonction c_E, a des propriétés particulières. Il peut d'ailleurs y avoir en plus des contraintes d'égalité additionnelles $c_E(x) = 0$ sans ces propriétés particulières; nous les avons omises pour alléger l'exposé. Voici la structure apportée par l'équation $c_S(x) = 0$, dite *équation d'état*.

La variable $x = (y, u)$ est partitionnée en *variable d'état* $y \in \mathbb{R}^{m_S}$ et en *variable de commande* $u \in \mathbb{R}^{n-m_S}$. On partitionne de la même manière la jacobienne de c_S :

$$c'_S(x) = \begin{pmatrix} B(x) & N(x) \end{pmatrix},$$

avec une matrice $B(x)$ carrée d'ordre m_S. L'hypothèse-clé est de supposer que $B(x)$ est inversible en tout point rencontré, en particulier en la solution (et donc dans son voisinage). Cette hypothèse permet, dans certaines approches algorithmiques, de représenter l'état comme fonction implicite de la commande et donc d'éliminer l'état du problème, ce qui peut représenter une réduction importante de la dimension du problème (lorsque $n - m_S \ll n$). Ce n'est cependant pas ce point de vue que nous allons présenter ici. Notre but est de montrer que l'algorithme OQS, dans une version dite *réduite*, permet d'avoir le même gain en dimension et donc en nombre d'opérations, tout en évitant la nécessité souvent coûteuse d'avoir des itérés qui satisfont l'équation d'état. *L'approche est surtout avantageuse lorsqu'il n'y a que des contraintes d'inégalité sur la commande*, autrement dit lorsque $\partial c_I / \partial y \equiv 0$.

Le problème quadratique osculateur (PQO) associé à (P_{SI}) en (x_k, λ_k) s'écrit

$$\begin{cases} \min_d \; g_k^\mathsf{T} d + \frac{1}{2} d^\mathsf{T} L_k d \\ c_S(x_k) + c'_S(x_k) d = 0 \\ c_I(x_k) + c'_I(x_k) d \leqslant 0. \end{cases} \tag{15.39}$$

L'hypothèse d'inversibilité de $B_k = B(x_k)$ permet d'introduire un inverse à droite A_k^- de $A_k \equiv c'_S(x_k)$ et une matrice Z_k^- dont les colonnes forment une base du noyau de A_k :

$$A_k^- = \begin{pmatrix} B_k^{-1} \\ 0 \end{pmatrix} \quad \text{et} \quad Z_k^- = \begin{pmatrix} -B_k^{-1} N_k \\ I_{n-m} \end{pmatrix}.$$

Ces matrices ne peuvent en général pas être calculées explicitement, mais on s'autorise à les appliquer à un vecteur, ce qui requiert à chaque fois la résolution d'un système linéaire. Alors toute solution de l'équation d'état linéarisée est de la forme

$$d_k = r_k + Z_k^- h_k,$$

où

$$r_k = -A_k^- c_S(x_k) \in \mathbb{E}$$

est un pas de restauration de l'équation d'état à commande fixée et $h_k \in \mathbb{R}^{n-m_S}$ est à déterminer. Si on reporte cette structure de d_k dans (15.39), le PQO devient

$$\begin{cases} \min_h \; (g_k + L_k r_k)^{\mathsf{T}} Z_k^- h + \frac{1}{2} h^{\mathsf{T}} Z_k^{-\mathsf{T}} L_k Z_k^- h \\ c_I(x_k) + c_I'(x_k) r_k + c_I'(x_k) Z_k^- h \leqslant 0. \end{cases}$$

Ce problème se simplifie considérablement si l'on peut faire disparaître la matrice Z_k^-. Dans ce but, on suit les étapes suivantes :

- on approche $Z_k^{-\mathsf{T}} L_k Z_k^-$ par une matrice M_k générée par une technique quasi-newtonienne,
- ne pouvant plus calculer le terme $L_k r_k$ dans la partie linéaire du critère (car L_k n'est ni calculé ni approché), on le néglige,
- il faut par ailleurs supposer que la matrice $c_I'(x_k) Z_k^-$ est simple à calculer.

Les deux premières étapes définissent ce que l'on appelle les *méthodes de quasi-Newton réduites*. L'abandon du terme $L_k r_k$ fait perdre la convergence superlinéaire, mais on garde toutefois la *convergence superlinéaire en* 2 *pas*, c'est-à-dire

$$\frac{\|x_{k+2} - x_*\|}{\|x_k - x_*\|} \to 0.$$

Le troisième point ci-dessus sera certainement satisfait s'il y a peu de contraintes d'inégalité ou si celles-ci portent uniquement sur les variables de commande (alors $c_I'(x_k) Z_k^- = (\partial c_I / \partial u)(x_k)$). Dans ce dernier cas, le PQO devient particulièrement simple

$$\begin{cases} \min_h \; g_k^{\mathsf{T}} Z_k^- h + \frac{1}{2} h^{\mathsf{T}} M_k h \\ c_I(x_k) + c_I'(x_k) r + (\partial c_I / \partial u)(x_k) h \leqslant 0. \end{cases}$$

Le vecteur $Z_k^{-\mathsf{T}} g_k = -N_k^{\mathsf{T}} B_k^{-\mathsf{T}} \nabla_y f(x_k) + \nabla_u f(x_k)$ est appelé le *gradient réduit*. On y reconnaît l'*état adjoint* $B_k^{-\mathsf{T}} \nabla_y f(x_k)$.

Notes

La section **??** ne fait que guigner sur les inclusions fonctionnelles en espérant inciter le lecteur à davantage s'intéresser davantage à ce vaste sujet [173].

On a mis longtemps à mettre au point l'algorithme de Newton pour résoudre les problèmes d'optimisation avec contraintes d'égalité et d'inégalité. Ce n'est, en effet, qu'au milieu des années 1970 que cette recherche a abouti, soit près de 30 ans après l'invention de l'algorithme du simplexe (chapitre 17). C'est d'ailleurs ce dernier que l'on a d'abord essayé de généraliser à l'optimisation non linéaire [621], mais la voie n'était pas directe. Les numériciens ont ensuite développé les méthodes de pénalisation (chapitre 13), en particulier l'approche par lagrangien augmenté. La prise de conscience de l'existence d'une méthode newtonienne directe n'est venue qu'ensuite [500, 302, 303, 491].

Pour un état de l'art sur l'algorithme OQS, on pourra consulter par exemple la section 5 de [272 ; 2005], la partie III de [73 ; 2006], Fletcher [214 ; 2010] et l'ouvrage technique et approfondi d'Izmailov et Solodov [330 ; 2014]. Voir aussi Izmailov, Kurennoy et Solodov [329 ; 2012]. La dérivation du résultat de convergence locale de l'algorithme OQS (section 15.1) à partir de celui de Josephy-Newton est reprise de Bonnans [72 ; 1994]. Pour l'établissement de conditions kantorovitchéennes de convergence, on pourra consulter Argyros et Hilout [19 ; 2010].

Exercices

15.1. *Non convergence locale de l'algorithme* (15.16). Écrire l'algorithme dont les itérés sont calculés par la récurrence $x_{k+1} = x_k + d_k$, avec d_k solution de (15.16), sous la forme $x_{k+1} = \Phi(x_k)$, où $\Phi : \mathbb{E} \to \mathbb{E}$. On supposera que x_k est voisin d'une solution x_*, que $c'(x_*)$ est surjective et que $\nabla^2 f(x_*)$ est inversible. Montrez que le spectre de $\Phi'(x_*)$ n'est pas dans la boule unité ouverte pour l'exemple en $x = (x_1, x_2) \in \mathbb{R}^2$ suivant :

$$\begin{cases} \min_x \; -a x_1^2 + 2x_2 \\ x_1^2 + x_2^2 = 1, \end{cases}$$

dans lequel $a \in \,]0, 1[$.