

Nicolas Petit

Optimisation

Enseignement spécialisé

<http://cas.ensmp.fr/~petit/>

MINES PARISTECH
ANNÉE SCOLAIRE 2014-2015

TABLE DES MATIÈRES

1. Optimisation continue de dimension finie	1
1.1. Notions fondamentales	1
1.1.1. Existence d'un minimum	2
1.1.2. Conditions nécessaires d'optimalité sur un ouvert	2
1.1.3. Conditions suffisantes d'optimalité sur un ouvert	3
1.1.4. Importance de la convexité	3
1.1.5. Notions avancées	4
1.1.6. Vitesses de convergence	4
1.2. Méthodes numériques de résolution sans contraintes	4
1.2.1. Méthodes sans dérivées	4
1.2.1.1. Méthode de dichotomie	5
1.2.1.2. Section dorée	5
1.2.1.3. Comparaison de l'effort numérique	5
1.2.2. Méthode utilisant la dérivée	5
1.2.2.1. Idée des méthodes de descente	5
1.2.2.2. Gradient à pas optimal	5
1.2.3. Recherche linéaire	7
1.2.3.1. Conditions de Wolfe	7
1.2.3.2. Résultats de convergence numérique	9
1.2.4. Méthode utilisant le Hessien : méthode de Newton	9
1.2.5. Méthode de quasi-Newton	10
1.2.5.1. Approximation de la fonction coût	11
1.2.5.2. Approximation optimale du Hessien	11
1.2.5.3. Calcul itératif de l'inverse de l'approximation du Hessien	12
1.2.5.4. Mise en œuvre	13
1.2.6. Méthode du gradient conjugué	14
1.2.6.1. Fonctions quadratiques	14
1.2.6.2. Application aux fonctions non linéaires	17
1.3. Principes de l'optimisation sous contraintes	18

1.3.1. Contraintes égalités	18
1.3.1.1. Élimination des variables	19
1.3.1.2. Équations adjointes	20
1.3.1.3. Multiplicateurs de Lagrange	20
1.4. Contraintes inégalités	21
1.4.1. Conditions de Karush-Kuhn-Tucker	21
1.4.1.1. Cas des contraintes affines	21
1.4.2. Dualité	23
1.4.3. Algorithme de minimisation sous contraintes utilisant la dualité	25
1.4.4. Méthode de contraintes actives pour les problèmes de programmation quadratique	26
2. Optimisation de trajectoires	29
2.1. Calcul des variations	29
2.1.1. Historique	29
2.1.2. Notions fondamentales	29
2.1.3. Conditions nécessaires d'extrémalité	30
2.2. Optimisation de systèmes dynamiques	32
2.2.1. Problème aux deux bouts	34
2.2.2. Contraintes finales	35
2.2.3. Résolution numérique du problème aux deux bouts	35
2.2.3.1. Calcul du gradient par l'adjoint	36
2.2.4. Principe du minimum	38
2.3. Champs d'extrémales	39
3. Exercices 1	43
3.1. Fonction de Rosenbrock	43
3.2. Gradient et Hessien	43
3.3. Ensemble minimisant	43
3.4. Vitesse de convergence	43
3.5. Conditions de Wolfe	44
3.6. Suite de Fibonacci	44
3.7. Moindres carrés	44
4. Exercices 2	47
4.1. Algorithme de Newton	47
4.2. Algorithme du gradient conjugué	47
4.3. Formule SR1	47
4.4. Invariance d'échelle	48
4.5. Calcul automatique du gradient par parcours de graphe [21]	48
5. Exercices 3	49
5.1. Fonction de mérite	49
5.2. Dual d'une programmation quadratique	49
5.3. Méthode de sous-ensembles actifs	50
5.4. Pénalisation	50

5.5. Plus courte distance à un hyperplan	50
5.6. Tente de cirque	51
6. Exercices 4	53
6.1. Optimisation de plan de vol	53
6.2. Forme d'une chaîne pesante	53
6.3. Investissement optimal	53
7. Exercices et problèmes complémentaires	55
7.1. Minimisation sous contraintes	55
7.2. Loi bi-tangente	55
7.3. Problème de Weber	56
7.4. Une équation HJB	56
7.5. Optimisation géométrique	57
7.6. Régulateur LQR	57
7.7. Programmation dynamique	58
7.8. Remplacement d'équipement	58
7.9. Tarification de billets d'avions	59
7.10. Forme d'une chaîne pesante	60
7.11. Équations d'Euler Lagrange pour les problèmes contraints à un temps variable	60
7.12. Temps minimum pour traverser une rivière à courant parabolique . . .	61
7.13. Calcul des variations	61
7.14. Inégalité de Kantorovich	61
7.15. Dualité	63
7.16. Résultats sur la minimisation d'une fonction elliptique	63
7.17. Une inégalité fonctionnelle	64
7.18. Unicité du temps minimum	65
7.19. Construction d'une autoroute	66
7.20. Programmation dynamique	67
7.21. Sur un théorème de C. Berge	68
7.22. Optimisation sur un espace fonctionnel	69
7.23. Meilleur antécédent par une matrice non inversible	70
7.24. Introduction aux méthodes de points intérieurs	71
7.25. Gestion thermique d'une batterie de véhicule hybride	72
7.26. Polygones inscrits du cercle de longueur maximale	73
7.27. Programmation dynamique	74
7.28. Problème de Dido	74
7.29. Calcul des variations avec saut de l'état	75
7.30. Parallélépipède maximal	76
7.31. Rangement d'un câble	76
7.32. Investissement public	77
7.33. Convexité	77
7.34. Projeté sur une parabole	78
7.35. Existence de minimums	78

7.36. Pénalité intérieure	78
7.37. PSA : profit sharing agreement	80
Bibliographie	83

CHAPITRE 1

OPTIMISATION CONTINUE DE DIMENSION FINIE

1.1. Notions fondamentales

On se place dans \mathbb{R}^n , $n < \infty$ considéré comme un espace vectoriel normé muni de la norme euclidienne notée $\|\cdot\|$. On considérera des fonctions différentiables ou deux fois différentiables $f : \mathbb{R}^n \mapsto \mathbb{R}$. On notera Ω un ouvert de \mathbb{R}^n .

Définition 1. — Soit l'application $\mathbb{R}^n \supset \Omega \ni x \mapsto f(x) \in \mathbb{R}$. On dit que f présente en $x^* \in \Omega$ un minimum (optimum) local, si et seulement s'il existe un voisinage V de x^* tel que $\forall x \in V \cap \Omega$, $f(x^*) \leq f(x)$ (on dit que c'est un minimum strict si la précédente inégalité est stricte). On dit que x est solution locale du problème

$$\min_{x \in \Omega} f(x)$$

On ne considérera que des problèmes de minimisation (calculs de minimum), les problèmes de maximisation étant obtenus en considérant l'opposé de la fonction f .

Définition 2. — On dit que $x^* \in \Omega$ est un minimum (optimum) global de la fonction $\mathbb{R}^n \supset \Omega \ni x \mapsto f(x) \in \mathbb{R}$ si $\forall x \in \Omega$ on a $f(x^*) \leq f(x)$. On dit que c'est un minimum strict global si la précédente inégalité est stricte. On dit que x^* est solution globale du problème

$$\min_{x \in \Omega} f(x)$$

Définition 3. — On note $(\nabla f(x))^T = \frac{\partial f}{\partial x}(x) = \left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n} \right)(x)$, le gradient de f où $x = (x_1, \dots, x_n)$.

Définition 4. — On appelle Hessian de f la matrice symétrique de $\mathcal{M}_n(\mathbb{R})$ $\nabla^2 f(x) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j} \right)(x)$, $i = 1, \dots, n, j = 1, \dots, n$.

Définition 5. — On note $\frac{\partial g}{\partial x}(x) = \left(\frac{\partial g_i}{\partial x_j} \right)(x)$, $i = 1, \dots, m, j = 1, \dots, n$, matrice de $\mathcal{M}_{(m \times n)}(\mathbb{R})$, le Jacobien de $g = (g_1(x), \dots, g_m(x))^T$.

Définition 6. — On dit que x^* est un point stationnaire de f si $\nabla f(x^*) = 0$.

1.1.1. Existence d'un minimum. — Le théorème suivant garantit l'existence d'une solution au problème de recherche de minimum.

Théorème 1 (Weierstrass). — *Si f est une fonction réelle continue sur un compact $K \subset \mathbb{R}^n$ alors le problème de recherche de minimum global*

$$\min_{x \in K} f(x)$$

possède une solution $x^ \in K$.*

Démonstration. — Notons $m = \inf_{x \in K} f(x)$ (éventuellement $m = -\infty$). Par définition, $\forall x \in K, f(x) \geq m$. Soit $(x^k)_{k \in \mathbb{N}}$ suite d'éléments de K telle que $(f(x^k))_{k \in \mathbb{N}}$ converge vers m . K est compact, on peut donc extraire une sous-suite $(x^l)_{l \in \mathcal{I} \subset \mathbb{N}}$ convergeant vers $x^* \in K$. Par continuité de f , on a

$$m = \lim_{k \rightarrow +\infty} f(x^k) = \lim_{l \rightarrow +\infty, l \in \mathcal{I}} f(x^l) = f(x^*)$$

Or $f(x^*) > -\infty$ car f est continue et K compact. En conclusion, il existe $x^* \in K$ tel que $f(x^*) = \min_{x \in K} f(x)$. \square

1.1.2. Conditions nécessaires d'optimalité sur un ouvert. —

Théorème 2. — *Soit Ω un ouvert de \mathbb{R}^n . Une condition nécessaire pour que x^* soit un optimum local de $\mathbb{R}^n \ni x \mapsto f(x) \in \mathbb{R}$ fonction deux fois différentiable est*

$$\{\nabla f(x^*) = 0, \nabla^2 f(x^*) \geq 0\}$$

Démonstration. — $x^* \in \Omega$ ouvert, donc $\forall h \in \mathbb{R}^n, \exists \eta > 0$ tel que $(x^* + \delta h) \in \Omega$ pour $0 \leq \delta \leq \eta$. Pour η suffisamment petit on a alors $f(x^* + \delta h) \geq f(x^*)$ par hypothèse. f est différentiable, on peut utiliser un développement de Taylor

$$f(x^*) + \frac{\partial f}{\partial x}(x^*)\delta h + o(\delta) \geq f(x^*)$$

d'où

$$\frac{\partial f}{\partial x}(x^*)h + o(\delta)/\delta \geq 0$$

Par passage à la limite lorsque δ tend vers zéro on obtient

$$\frac{\partial f}{\partial x}(x^*)h \geq 0 \text{ pour tout } h \in \mathbb{R}^n$$

Nécessairement $\frac{\partial f}{\partial x}(x^*) = 0$. Utilisons maintenant un développement au deuxième ordre

$$f(x^* + \delta h) = f(x^*) + \frac{1}{2}(\delta h)^T \nabla^2 f(x^*) \delta h + o(\delta^2)$$

Dans le même voisinage que précédemment $f(x^* + \delta h) \geq f(x^*)$ implique après passage à la limite que pour tout $h \in \mathbb{R}^n$

$$h^T \nabla^2 f(x^*) h \geq 0$$

et donc que $\nabla^2 f(x^*) \geq 0$. \square

1.1.3. Conditions suffisantes d'optimalité sur un ouvert. —

Théorème 3. — Une condition suffisante pour que x^* soit un optimum local de $\mathbb{R}^n \ni x \mapsto f(x) \in \mathbb{R}$ fonction deux fois différentiable sur Ω ouvert de \mathbb{R}^n est

$$\{\nabla f(x^*) = 0, \nabla^2 f(x^*) > 0\}$$

Démonstration. — évidente par le même développement de Taylor à l'ordre 2. \square

1.1.4. Importance de la convexité. —

Définition 7. — E ensemble de \mathbb{R}^n est convexe si $\forall (x, y) \in E \times E, \forall \lambda \in [0, 1], \lambda x + (1 - \lambda)y \in E$.

Définition 8. — Soit E convexe de \mathbb{R}^n . On dit que l'application $E \ni x \mapsto f(x) \in \mathbb{R}$ est convexe si $\forall (x, y) \in E \times E, \forall \lambda \in [0, 1]$ on a

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$$

Théorème 4. — Soit E convexe de \mathbb{R}^n et $E \ni x \mapsto f(x) \in \mathbb{R}$ application deux fois différentiable. Les propositions suivantes sont équivalentes

1. f est convexe
2. $\forall (x, y) \in E^2, f(y) \geq f(x) + (\nabla f(x))^T(y - x)$
3. $\forall x \in E, \nabla^2 f(x) \geq 0$

Théorème 5. — Pour une application convexe définie sur un ensemble convexe, tout minimum local est global.

Démonstration. — Soit x^* minimum local de f . $\exists \epsilon > 0$ tel que $\forall x$ tel que $\|x - x^*\| < \epsilon, f(x) \geq f(x^*)$. Pour tout $x \in E$ on peut construire, par convexité de $E, E \ni x^\alpha = (1 - \alpha)x^* + \alpha x$, avec $\alpha \in]0, 1]$ tel que $\|x^\alpha - x^*\| < \epsilon$. On alors $f(x^*) \leq f(x^\alpha)$. Or, par convexité de la fonction f on a $f(x^\alpha) \leq (1 - \alpha)f(x^*) + \alpha f(x)$. Après simplification, il vient $f(x^*) \leq f(x)$. \square

Théorème 6. — Soit f une fonction convexe deux fois différentiable. Une condition nécessaire et suffisante pour que $x^* \in \Omega \subset \mathbb{R}^n$ avec Ω convexe soit un optimum global est que

$$\nabla f(x^*) = 0$$

Théorème 7. — Soit $E \subset \mathbb{R}^n$ convexe fermé. Soit $E \ni x \mapsto f(x) \in \mathbb{R}$ continue et convexe telle que pour toute suite $(u^n)_{n \in \mathbb{N}}$ telle que $\lim_{n \rightarrow \infty} \|u^n\| = +\infty$ on a $\lim_{n \rightarrow \infty} f(u^n) = +\infty$. Il existe une solution au problème $\min_{x \in E} f(x)$.

1.1.5. Notions avancées. — On introduit ici la notion de convexité forte qui permettra de prouver la convergence de certains algorithmes de résolution de problèmes d'optimisation.

Définition 9. — Soit $E \subset \mathbb{R}^n$ convexe. On dit que $E \ni x \mapsto f(x) \in \mathbb{R}$ est fortement convexe (α -convexe) si et seulement si $\exists \alpha > 0$ tel que $\forall (x, y) \in E \times E$

$$f\left(\frac{x+y}{2}\right) \leq \frac{f(x)+f(y)}{2} - \frac{\alpha}{8} \|x-y\|^2$$

Théorème 8. — Soit f une application différentiable de Ω dans \mathbb{R} , et $\alpha > 0$. Les propositions suivantes sont équivalentes

- f est α -convexe sur Ω
- $\forall (x, y) \in \Omega^2, f(y) \geq f(x) + (\nabla f(x))^T(y-x) + \frac{\alpha}{2} \|x-y\|^2$
- $\forall (x, y) \in \Omega^2, ((\nabla f(x))^T - (\nabla f(y))^T)(x-y) \geq \alpha \|x-y\|^2$

C'est souvent la dernière écriture qui est la plus utile dans les démonstrations. Pour l'étude des propriétés des algorithmes on caractérise les vitesses de convergence comme suit.

1.1.6. Vitesses de convergence. — Soit une suite de valeurs $(x_k)_{k \in \mathbb{N}}$ convergeant vers x^* . On dira que cette suite converge linéairement, si $\exists r \in]0, 1[$ tel que pour k suffisamment grand

$$\|x_{k+1} - x^*\| \leq r \|x_k - x^*\|$$

On dira que cette suite converge super linéairement, si pour k suffisamment grand

$$\lim_{k \rightarrow \infty} \|x_{k+1} - x^*\| / \|x_k - x^*\| = 0$$

On dira que cette suite converge de façon quadratique si $\exists M > 0$ tel que, pour k suffisamment grand

$$\|x_{k+1} - x^*\| \leq M \|x_k - x^*\|^2$$

1.2. Méthodes numériques de résolution sans contraintes

1.2.1. Méthodes sans dérivées. — Il est possible de résoudre des problèmes (simples) d'optimisation avec peu de connaissance de la fonction. On propose ainsi une méthode n'utilisant que la possibilité d'évaluer la fonction.

Définition 10. — On dit que $f : \mathbb{R} \mapsto \mathbb{R}$ est unimodale sur un intervalle $[A, B]$ si elle admet un minimum x^* et si $\forall x_1 < x_2$ dans $[A, B]$ on a

1. $x_2 \leq x^*$ implique $f(x_1) > f(x_2)$
2. $x_1 \geq x^*$ implique $f(x_1) < f(x_2)$

Autrement dit, f possède un minimum global unique mais n'est ni nécessairement continue ni différentiable partout. Une conséquence de la propriété d'unimodalité est que si on divise l'intervalle $[A, B]$ en 4, il existe toujours un sous-intervalle qui ne contient pas l'optimum.

1.2.1.1. Méthode de dichotomie. — Parmi les 4 intervalles précédents choisis de longueur égales on peut toujours en supprimer 2. L'intervalle de recherche est alors divisé par 2 à chaque itération.

1.2.1.2. Section dorée. — Au lieu de subdiviser en 4 intervalles, on utilise 3 intervalles. À chaque itération on peut exclure l'un des 3 intervalles. De manière à réutiliser les calculs entre deux itérations, il faut choisir une méthode de découpage utilisant le nombre d'or. On vérifie qu'il faut diviser l'intervalle de recherche à l'itération i $[A_i, B_i]$ suivant $[A_i, A_i + (1 - \tau)(B_i - A_i), A_i + \tau(B_i - A_i), B_i]$ où $\tau = (\sqrt{5} - 1)/2$.

1.2.1.3. Comparaison de l'effort numérique. — Les deux méthodes de division en sous-intervalles convergent vers l'unique minimum x^* quel que soit l'intervalle de recherche de départ $[A, B]$. On peut s'intéresser à leur efficacité en terme de nombre d'itérations. Pour obtenir un effort de réduction de l'intervalle de recherche initial de 10^{-2} , 10^{-3} , 10^{-6} , on doit faire le nombre d'appel à la fonction f suivant : 17, 23, 42 pour la méthode de dichotomie et 13, 18, 31 pour la méthode de la section dorée. La méthode de la section dorée est plus efficace mais moins triviale à mettre en œuvre que la méthode de dichotomie qui lui est souvent préférée.

1.2.2. Méthode utilisant la dérivée. — On suppose désormais pouvoir calculer la valeur de la fonction ainsi que son gradient. Partant d'une valeur initiale x^0 , on va mettre à jour une estimation x^k de l'optimum recherché x^* .

1.2.2.1. Idée des méthodes de descente. — Entre deux itérations k et $k + 1$ on fait évoluer l'estimée de l'optimum x^k suivant la formule

$$x^{k+1} = x^k + l^k p^k$$

où $l^k \in \mathbb{R}$ est appelé pas et $p^k \in \mathbb{R}^n$ est une direction. Si par exemple on choisit $p^k = -\nabla f(x^k)$ on obtient alors pour une fonction différentiable

$$f(x^{k+1}) = f(x_k) - l^k \|\nabla f(x^k)\|^2 + o(l^k)$$

On peut donc espérer une décroissance des valeurs de la fonction entre deux itérations, d'où le nom de méthode de descente. On constate qu'une règle simple de pas constant produit souvent une bonne décroissance au début des itérations suivi d'une relative stagnation des valeurs de la fonction coût. En cherchant à améliorer la méthode, on réalise rapidement qu'il est important de bien choisir l^k .

1.2.2.2. Gradient à pas optimal. — L'algorithme suivant possède de très intéressantes propriétés de convergence.

Algorithme 1. — À partir de $x^0 \in \mathbb{R}^n$ quelconque, itérer

$$x^{k+1} = x^k - l^k \nabla f(x^k)$$

où $l^k = \operatorname{argmin}_{l \in \mathbb{R}} f(x^k - l \nabla f(x^k))$.

Théorème 9. — Si f est α -convexe, différentiable et de gradient ∇f Lipschitzien sur tout borné (c.-à-d. $\forall M > 0, \exists C_M > 0$ tel que $\forall (x, y) \in \mathbb{R}^n \times \mathbb{R}^n, \|x\| + \|y\| \leq M$ implique $\|\nabla f(x) - \nabla f(y)\| \leq C_M \|x - y\|$) alors l'algorithme 1 du gradient à pas optimal converge vers l'unique solution x^* du problème d'optimisation $\min_{x \in \mathbb{R}^n} f(x)$.

Démonstration. — La fonction $\mathbb{R} \ni l \mapsto g(l) = f(x^k - l\nabla f(x^k))$ est α -convexe et dérivable sur \mathbb{R} . On suppose qu'à l'itération k , $\nabla f(x^k) \neq 0$. Le problème d'optimisation $\min_{l \in \mathbb{R}} f(x^k - l\nabla f(x^k))$ a une solution unique caractérisée par $\nabla g(l) = 0$. Cette équation s'écrit

$$\nabla f(x^k)^T \nabla f(x^k - l\nabla f(x^k)) = 0$$

On s'aperçoit ainsi que deux directions de descentes successives sont orthogonales. Entre deux itérations on a donc

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

car $l_k \neq 0$. On déduit alors de l' α -convexité de f

$$f(x^k) \geq f(x^{k+1}) + \frac{\alpha}{2} \|x^{k+1} - x^k\|^2$$

Donc

$$(1) \quad \frac{\alpha}{2} \|x^{k+1} - x^k\|^2 \leq f(x^k) - f(x^{k+1})$$

La suite de réels $(f(x^k))_{k \in \mathbb{N}}$ est décroissante, minorée par $f(x^*)$, elle est donc convergente. On en déduit que, par (1),

$$\lim_{k \rightarrow +\infty} \|x^{k+1} - x^k\| = 0$$

D'autre part, $(f(x^k))_{k \in \mathbb{N}}$ est décroissante, minorée par $f(x^*)$, elle est bornée. Il existe donc $M \in \mathbb{R}$ tel que

$$M \geq f(x^k) \geq f(x^*) + \frac{\alpha}{2} \|x^k - x^*\|^2$$

où la dernière inégalité provient de l' α -convexité de f . La suite $(x^k)_{k \in \mathbb{N}}$ est donc bornée. Utilisons maintenant que le gradient ∇f est Lipschitzien sur tout borné. $\exists C_M > 0$ tel que $\|\nabla f(x^k) - \nabla f(x^{k+1})\| \leq C_M \|x^{k+1} - x^k\|$. Par l'orthogonalité entre les directions de descente entre deux pas successifs, il vient alors

$$\left(\|\nabla f(x^k)\|^2 + \|\nabla f(x^{k+1})\|^2 \right)^{1/2} \leq C_M \|x^{k+1} - x^k\|$$

Par passage à la limite on a alors

$$\lim_{k \rightarrow +\infty} \|\nabla f(x^k)\| = 0$$

Enfin, en utilisant une dernière fois l' α -convexité de f , on a

$$\|\nabla f(x^k)\| \|x^k - x^*\| \geq (\nabla f(x^k) - \nabla f(x^*))^T (x^k - x^*) \geq \alpha \|x^k - x^*\|^2$$

D'où

$$\|x^k - x^*\| \leq \frac{1}{\alpha} \|\nabla f(x^k)\|$$

Par passage à la limite on a alors

$$\lim_{k \rightarrow +\infty} x_k = x^*$$

□

L'algorithme du gradient à pas optimal possède donc une intéressante propriété de convergence mais comporte dans chaque itération une recherche de pas optimal. C'est un problème mono-dimensionnel qui peut être traité par les méthodes de dichotomie ou de section dorée. Cette résolution itérée peut être coûteuse en calculs et on lui préférera souvent une des méthodes de recherche linéaire présentée dans la section suivante.

1.2.3. Recherche linéaire. — L'idée générale consiste à proposer une méthode de sélection du pas l^k qui permette de prouver la convergence

$$\lim_{k \rightarrow +\infty} f(x_k) = 0,$$

d'avoir une bonne vitesse de convergence, et qui soit simple à implémenter.

1.2.3.1. Conditions de Wolfe. —

Définition 11. — On appelle condition d'Armijo (de paramètre c_1) sur les itérations $(x^k, p^k, l^k)_{k \in \mathbb{N}}$ l'inéquation

$$(2) \quad f(x^k + l^k p^k) \leq f(x^k) + c_1 l^k \nabla f(x^k)^T p^k$$

Définition 12. — On appelle condition de courbure (de paramètre c_2) sur les itérations $(x^k, p^k, l^k)_{k \in \mathbb{N}}$ l'inéquation

$$(3) \quad \nabla f(x^k + l^k p^k)^T p^k \geq c_2 \nabla f(x^k)^T p^k$$

Dans la suite on utilisera les deux conditions précédentes pour des paramètres $0 < c_1 < c_2 < 1$. La première condition autorise des pas grands mais pas trop (en pratique ceci évite les "rebonds" au cours des itérations). La deuxième condition permet d'éviter les pas trop petits.

Définition 13. — On appelle conditions de Wolfe, les conditions (2) et (3) avec $0 < c_1 < c_2 < 1$.

Théorème 10. — Soit $\mathbb{R}^n \ni x \mapsto f(x) \in \mathbb{R}$ différentiable. Soit p^k direction de descente en x^k telle que f est bornée inférieurement sur la droite $\{x_k + lp^k, l > 0\} \subset \mathbb{R}^n$. Il existe un intervalle $[l_1, l_2]$ tel que tout $l \in [l_1, l_2]$ satisfait les conditions de Wolfe.

Démonstration. — La fonction $\mathbb{R} \ni l \mapsto f(x^k + lp^k)$ est bornée inférieurement donc pour tout $0 < c_1 < 1$, l'ensemble $\{f(x^k) + lc_1 \nabla f(x^k)^T p^k, l > 0\} \subset \mathbb{R}$ contient $f(x^k + lp^k)$. Notons $l' > 0$ la plus petite valeur de l telle que $f(x^k) + lc_1 \nabla f(x^k)^T p^k = f(x^k + lp^k)$. La condition d'Armijo (2) est toujours satisfaite pour $l < l'$. Par le théorème des accroissements finis appliqué à la fonction différentiable $[0, l'] \ni l \mapsto f(x^k + lp^k) \in \mathbb{R}$, $\exists l'' \in [0, l']$ tel que $f(x^k + l'p^k) - f(x^k) = l' \nabla f(x^k + l''p^k)^T p^k$. On a alors

$$\nabla f(x^k + l''p^k)^T p^k = c_1 \nabla f(x^k)^T p^k > c_2 \nabla f(x^k)^T p^k$$

Cette inégalité étant stricte, elle reste large dans un voisinage de l'' ce qui garantit qu'il existe un intervalle non vide autour de l'' dans lequel l satisfait également la condition de courbure (3). \square

On va maintenant chercher à prouver la convergence de l'algorithme de descente utilisant la recherche linéaire qu'on vient de présenter. + cette fin, nous allons utiliser le théorème suivant

Théorème 11 (Zoutendijk). — *Soit une séquence $(x^k, p^k, l^k)_{k \in \mathbb{N}}$ vérifiant $x^{k+1} = x^k + l^k p^k$ satisfaisant les conditions de Wolfe. On définit*

$$\cos \theta^k = \frac{-\nabla f(x^k)^T p^k}{\|\nabla f(x^k)\| \|p^k\|} \geq 0$$

(c.-à-d. le cosinus de l'angle entre la direction de recherche et la plus grande pente). Si f (supposée différentiable) est bornée inférieurement et si ∇f est Lipschitzien alors

$$\sum_{k \in \mathbb{N}} \cos^2 \theta^k \|\nabla f(x^k)\|^2 < +\infty$$

En corollaire $\lim_{k \rightarrow \infty} \cos^2 \theta^k \|\nabla f(x^k)\|^2 = 0$.

Démonstration. — D'après la condition de courbure, on a

$$\nabla f(x^{k+1})^T p^k \geq c_2 \nabla f(x^k)^T p^k$$

et donc

$$(\nabla f(x^{k+1}) - \nabla f(x^k))^T p^k \geq (c_2 - 1) \nabla f(x^k)^T p^k$$

Or $x \mapsto \nabla f(x)$ est Lipschitz, donc $\exists L > 0$ tel que

$$\|\nabla f(x^{k+1}) - \nabla f(x^k)\| \leq L \|x^{k+1} - x^k\|$$

Il vient donc $L l^k \|p^k\|^2 \geq (c_2 - 1) \nabla f(x^k)^T p^k$ et donc

$$l^k \geq \frac{(c_2 - 1) \nabla f(x^k)^T p^k}{L \|p^k\|^2} \geq 0$$

D'après la condition d'Armijo, on a alors

$$f(x^k + l^k p^k) \leq f(x^k) + \frac{c_1(c_2 - 1)}{L} \frac{(\nabla f(x^k)^T p^k)^2}{\|p^k\|^2}$$

Une sommation terme à terme donne alors

$$\frac{L}{c_1(c_2 - 1)} \sum_{k \in \mathbb{N}} (f(x^{k+1}) - f(x^k)) \geq \sum_{k \in \mathbb{N}} \cos^2 \theta^k \|\nabla f(x^k)\|^2$$

La suite de réels $(f(x^k))_{k \in \mathbb{N}}$ est décroissante, bornée inférieurement, elle converge. En conclusion

$$\sum_{k \in \mathbb{N}} \cos^2 \theta^k \|\nabla f(x^k)\|^2 < \infty$$

□

On utilise le théorème de Zoutendjik en choisissant $p^k = -\nabla f(x^k)$. On a alors, pour tout $k \in \mathbb{N}$, $\cos \theta^k = 1$. La série

$$\sum_{k \in \mathbb{N}} \|\nabla f(x^k)\|^2$$

converge donc pour tout algorithme de gradient satisfaisant les conditions de Wolfe et on a la convergence

$$\lim_{k \rightarrow \infty} \|\nabla f(x^k)\| = 0$$

1.2.3.2. Résultats de convergence numérique. — Si on applique l'algorithme de gradient avec règles de Wolfe à la fonction $\mathbb{R}^n \ni x \mapsto 1/2x^T Qx - b^T x$ avec Q matrice de $\mathcal{M}_n(\mathbb{R})$ symétrique définie positive et $b \in \mathbb{R}^n$ on peut montrer la convergence

$$(x^{k+1} - x^*)^T Q(x^{k+1} - x^*) \leq \left(\frac{\lambda_n - \lambda_1}{\lambda_n + \lambda_1} \right)^2 (x^k - x^*)^T Q(x^k - x^*)$$

où λ_1 et λ_n sont respectivement la plus petite et la plus grande des valeurs propres de Q . Les performances se dégradent donc avec le conditionnement du problème.

1.2.4. Méthode utilisant le Hessien : méthode de Newton. — Autour de x^k , une approximation de f supposée deux fois différentiable est

$$q(x) = f(x^k) + \nabla f(x^k)^T (x - x^k) + \frac{1}{2} (x - x^k)^T \nabla^2 f(x^k) (x - x^k)$$

Si $\nabla^2 f(x^k) > 0$ alors $\mathbb{R}^n \ni x \mapsto q(x) \in \mathbb{R}$ est strictement convexe. Son minimum est donc atteint en un point x tel que $\nabla q(x) = 0$ c.-à-d.

$$\nabla f(x^k) + \nabla^2 f(x^k)(x - x^k) = 0$$

qui donne $x = x^k - (\nabla^2 f(x^k))^{-1} \nabla f(x^k)$. La méthode de Newton consiste à itérer le calcul précédent. À chaque étape on doit effectuer $O(n^3)$ calculs dont la majeure partie est due à l'inversion du Hessien de f en x^k .

Algorithme 2. — À partir de $x^0 \in \mathbb{R}^n$ quelconque, itérer

$$x^{k+1} = x^k - (\nabla^2 f(x^k))^{-1} \nabla f(x^k)$$

Théorème 12. — Soit $\mathbb{R}^n \ni x \mapsto f(x) \in \mathbb{R}$ deux fois différentiable, possédant un unique minimum global x^* tel que $\nabla^2 f(x^*)$ est définie positive (on note $\lambda > 0$ sa plus petite valeur propre) et tel que $\mathbb{R}^n \ni x \mapsto \nabla^2 f(x) \in \mathcal{M}_n(\mathbb{R})$ est localement Lipschitz au voisinage de x^* (on note C sa constante Lipschitz). L'algorithme 2 de Newton converge quadratiquement vers x^* si on l'initialise en un point x^0 tel que $\|x^0 - x^*\| \leq \frac{2\lambda}{3C}$.

Démonstration. — D'après la formule de Taylor avec reste intégral appliquée à la fonction $t \mapsto \nabla f(x^k + t(x^* - x^k))$, on a, en utilisant $\nabla f(x^*) = 0$,

$$-\nabla f(x^k) - \nabla^2 f(x^k)(x^* - x^k) = \int_0^1 (\nabla^2 f(x^k + t(x^* - x^k)) - \nabla^2 f(x^k)) (x^* - x^k) dt$$

D'autre part, si x^k est suffisamment proche de x^* (c'est vrai pour $k = 0$ et le restera par récurrence par l'équation (7)), la condition de Lipschitz donne qu'il existe $C > 0$ tel que

$$\|\nabla^2 f(x^k + t(x^* - x^k)) - \nabla^2 f(x^k)\| \leq C \|x^* - x^k\| t$$

On peut donc déduire

$$(4) \quad \|\nabla f(x^k) - \nabla^2 f(x^k)(x^* - x^k)\| \leq \frac{C}{2} \|x^* - x^k\|^2$$

L'itération de l'algorithme de Newton donne $x^k = x^{k+1} + \nabla^2 f(x^k)^{-1} \nabla f(x^k)$. Après substitution et simplification dans (4) on a

$$\|\nabla^2 f(x^k)(x^* - x^{k+1})\| \leq \frac{C}{2} \|x^* - x^k\|^2$$

Il s'agit maintenant de faire apparaître $\nabla^2 f(x^*)$ qui est la quantité sur laquelle porte l'hypothèse. Une inégalité triangulaire donne

$$\begin{aligned} \|\nabla^2 f(x^*)(x^* - x^{k+1})\| &\leq \|\nabla^2 f(x^k)(x^* - x^{k+1})\| \\ &\quad + \|(\nabla^2 f(x^*) - \nabla^2 f(x^k))(x^* - x^{k+1})\| \end{aligned}$$

À l'aide des majorations précédentes on a alors

$$(5) \quad \|\nabla^2 f(x^*)(x^* - x^{k+1})\| \leq \frac{C}{2} \|x^* - x^k\|^2 + C \|x^* - x^k\| \|x^{k+1} - x^*\|$$

Notons maintenant $\lambda > 0$ la plus petite valeur propre de $\nabla^2 f(x^*)$. L'inéquation (5) donne alors

$$(6) \quad \lambda \|x^* - x^{k+1}\| \leq \frac{C}{2} \|x^* - x^k\|^2 + C \|x^* - x^k\| \|x^{k+1} - x^*\|$$

On a $\|x^k - x^*\| \leq \frac{2\lambda}{3C}$. Pour tous les indices $p > k$ suivants on encore $\|x^p - x^*\| \leq \frac{2\lambda}{3C}$. En effet on peut tirer de (6) que

$$(\lambda - C \|x^k - x^*\|) \|x^{k+1} - x^*\| \leq \frac{C}{2} \|x^k - x^*\|^2$$

d'où $\|x^{k+1} - x^*\| \leq \frac{2\lambda}{3C}$. Finalement l'inéquation (6) donne donc la convergence quadratique

$$(7) \quad \|x^{k+1} - x^*\| \leq \frac{3C}{2\lambda} \|x^k - x^*\|^2$$

□

1.2.5. Méthode de quasi-Newton. — Cette méthode ne requiert que l'évaluation de la fonction et de son gradient. Elle s'avère en pratique souvent aussi performante que la méthode de Newton dans le calcul de la direction de descente. Au lieu de calculer cette direction par la formule $p^k = -\nabla^2 f(x^k)^{-1} \nabla f(x^k)$, on utilise une formule

$$p^k = -(B^k)^{-1} \nabla f(x^k)$$

où B^k est une matrice symétrique définie positive qu'on va mettre à jour au cours des itérations.

1.2.5.1. *Approximation de la fonction coût.* — À l'itération k on dispose d'une estimation de l'optimum x^k , et on forme la fonction

$$\mathbb{R}^n \ni p \mapsto m^k(p) = f(x^k) + \nabla f(x^k)^T p + \frac{1}{2} p^T B^k p \in \mathbb{R}$$

Le minimum de cette fonction est atteint en $p^k = -(B^k)^{-1} \nabla f(x^k)$ et peut permettre de former une nouvelle estimée de l'optimum

$$x^{k+1} = x^k + l^k p^k$$

où l^k peut être choisi par les conditions de Wolfe (par exemple).

On va chercher à introduire les modifications du gradient entre les itérations (courbure) dans la mise à jour de B^k . Le nouveau modèle autour de l'estimation x^{k+1} est

$$\mathbb{R}^n \ni p \mapsto m^{k+1}(p) = f(x^{k+1}) + \nabla f(x^{k+1})^T p + \frac{1}{2} p^T B^{k+1} p \in \mathbb{R}$$

On va faire coïncider le gradient de m^{k+1} avec celui de f qu'on a évalué aux itérations x^k et x^{k+1} . Par construction, $\nabla m^{k+1}(0) = \nabla f(x^{k+1})$. D'autre part, l'autre gradient recherché $\nabla m^{k+1}(-l^k p^k) = \nabla f(x^{k+1}) - l^k B^{k+1} p^k$. On utilise désormais les notations

$$s^k = x^{k+1} - x^k,$$

$$y^k = \nabla f(x^{k+1}) - \nabla f(x^k)$$

On a donc à résoudre

$$B^{k+1} s^k = y^k$$

Il n'est pas toujours possible de résoudre cette équation dont l'inconnue B^{k+1} est à rechercher symétrique définie positive. Une condition nécessaire et suffisante pour que cette équation possède une solution symétrique définie positive est que

$$(s^k)^T y^k > 0$$

ce qui peut se réécrire $(x^{k+1} - x^k)^T (\nabla f(x^{k+1}) - \nabla f(x^k)) > 0$. On pourra éventuellement chercher à imposer cette condition au moment de la recherche linéaire. Cette condition sera toujours remplie si la fonction f est α -convexe. Les conditions de Wolfe garantissent que cette condition est remplie. En effet on a

$$\nabla f(x^{k+1})^T s^k \geq c_2 \nabla f(x^k)^T s^k$$

d'après la condition de courbure. Ceci implique bien $(s^k)^T y^k \geq (c_2 - 1) l^k \nabla f(x^k)^T p^k > 0$ car on utilise un algorithme de descente.

1.2.5.2. *Approximation optimale du Hessien.* — Notre but était de définir B^{k+1} , on sait qu'on peut toujours le faire par les conditions de Wolfe. On va maintenant rendre ce choix unique. En l'état actuel on a $\frac{n(n+1)}{2}$ coefficients et n équations ainsi que n inégalités (provenant de la contrainte que l'approximation doit être positive définie). On peut montrer que ces inégalités possèdent des solutions. Pour définir de manière unique notre solution, on va chercher à s'approcher le plus possible de B^k au sens d'une certaine norme. Introduisons à cet effet la définition suivante.

Définition 14 (Norme de Frobenius (pondérée)). — Soit $A = (a_{ij})_{i=1\dots n, j=1\dots n}$ une matrice de $\mathcal{M}_n(\mathbb{R})$ et W une matrice définie positive de $\mathcal{M}_n(\mathbb{R})$. On définit $\|A\|_W$ la norme pondérée de Frobenius par l'égalité suivante

$$\|A\|_W = \left\| W^{1/2} A W^{1/2} \right\|_F$$

où $\|A\|_F = \sum_{i=1\dots n, j=1\dots n} a_{ij}^2$

Dans notre cas, nous allons choisir W matrice de pondérations telle que $W y^k = s^k$, par exemple $W^{-1} = \int_0^1 \nabla^2 f(x^k + \tau l^k p^k) d\tau$ qui est la moyenne du Hessien le long du chemin de la dernière itération.

En utilisant cette norme on résout le problème d'optimisation (d'inconnue B) suivant

$$(8) \quad \min_{\substack{B = B^T \\ B s^k = y^k}} \|B - B^k\|_W$$

Proposition 1. — *Le problème d'optimisation (8) possède une unique solution B^* qui est*

$$B^* = (I - \gamma^k y^k (s^k)^T) B^k (I - \gamma^k s^k (y^k)^T) + \gamma^k y^k (y^k)^T$$

où $\gamma^k = 1 / ((y^k)^T s^k)$.

La précédente formule sera utilisée comme mise à jour du Hessien dans l'algorithme DFP (Davidson, Fletcher, Powell).

1.2.5.3. Calcul itératif de l'inverse de l'approximation du Hessien. — La quantité que nous utilisons dans les algorithmes de quasi-Newton est l'inverse du Hessien $(B^k)^{-1}$ et non le Hessien lui-même. Pour calculer cet inverse de manière efficace on peut utiliser la formule issue de la proposition suivante

Proposition 2 (Formule de Sherman-Morrison-Woodbury)

Soient A une matrice de $\mathcal{M}_n(\mathbb{R})$ inversible, U et V deux matrices de $\mathcal{M}_{n,p}(\mathbb{R})$, $1 \leq p \leq n$. On définit $\bar{A} = A + UV^T$. Si \bar{A} est inversible alors

$$(9) \quad \bar{A}^{(-1)} = A^{(-1)} - A^{(-1)} U \left(I_{(p)} + V^T A^{(-1)} U \right)^{(-1)} V^T A^{(-1)}$$

Lorsqu'on met à jour l'approximation du Hessien par la formule de la proposition 1 on a

$$(10) \quad B^{k+1} = (I - \gamma^k y^k (s^k)^T) B^k (I - \gamma^k s^k (y^k)^T) + \gamma^k y^k (y^k)^T$$

Il lui correspond la mise à jour de l'inverse

$$(11) \quad (B^{k+1})^{-1} = (B^k)^{-1} - \frac{1}{(y^k)^T (B^k)^{-1} y^k} \left((B^k)^{-1} y^k (y^k)^T (B^k)^{-1} \right) + \frac{1}{(y^k)^T s^k} s^k (s^k)^T$$

L'équation (11) redonne bien (10). Cette dernière équation fait apparaître une forme factorisée et un terme de mise à jour UV^T en notant $U = (-\frac{1}{a}U_1, \frac{1}{b}U_2)$ et $V = (U_1, U_2)$ avec

$$\begin{aligned} U_1 &= (B^k)^{-1}y^k \\ U_2 &= s^k \\ a &= (y^k)^T (B^k)^{-1}y^k \\ b &= (y^k)^T s^k \end{aligned}$$

En utilisant la formule de Sherman-Morrison-Woodbury (9), il vient alors

$$B^{k+1} = B^k - B^k U (I_{(2)} + V^T B^k U) V^T B^k$$

Après quelques lignes de développement, on obtient

$$\begin{aligned} B^{k+1} &= B^k + (\gamma^k)^2 y^k (s^k)^T B^k s^k (y^k)^T \\ &\quad - \gamma^k y^k (s^k)^T B^k - \gamma^k B^k s^k (y^k)^T + y^k (y^k)^T \gamma^k \\ &= (I - \gamma^k y^k (s^k)^T) B^k (I - \gamma^k s^k (y^k)^T) + \gamma^k y^k (y^k)^T \end{aligned}$$

qui est bien (10).

1.2.5.4. Mise en œuvre. — On retiendra deux formules possibles pour la mise à jour de l'inverse de l'approximation du Hessien : la formule DFP (11) et la formule BFGS (Broyden, Fletcher, Goldfarb, Shanno) suivante

$$(12) \quad (B^{k+1})^{-1} = (I - \gamma^k s^k (y^k)^T) (B^k)^{-1} (I - \gamma^k y^k (s^k)^T) + \gamma^k s^k (s^k)^T$$

qui correspond à la résolution du problème

$$\begin{aligned} \min_{B^{-1} = B^{-T}} \quad & \|B^{-1} - (B^k)^{-1}\|_W \\ B^{-1} y^k &= s^k \end{aligned}$$

où $\gamma^k = 1/((y^k)^T s^k)$.

Algorithme 3 (Algorithme de BFGS). — À partir de $x^0 \in \mathbb{R}^n$ quelconque, et de $H^0 = I(n)$ (d'autres choix de matrice définie positive sont possibles) itérer

$$\begin{aligned} p^k &= -H^k \nabla f(x^k) \\ x^{k+1} &= x^k + l^k p^k, \quad l^k \text{ satisfaisant les conditions de Wolfe} \\ s^k &= x^{k+1} - x^k \\ y^k &= \nabla f(x^{k+1}) - \nabla f(x^k) \\ (B^{k+1})^{-1} &= (I - \gamma^k s^k (y^k)^T) (B^k)^{-1} (I - \gamma^k y^k (s^k)^T) + \gamma^k s^k (s^k)^T \\ H^{k+1} &= (B^{k+1})^{-1} \end{aligned}$$

Théorème 13. — Soit $\mathbb{R}^n \ni x \mapsto f(x) \in \mathbb{R}$ deux fois différentiable telle que la ligne de niveau $\mathcal{L} = \{x/f(x) \leq f(x^0)\}$ est convexe et qu'il existe des constantes positives m et M telles que $\forall z \in \mathbb{R}^n$ et $\forall x \in \mathcal{L}$

$$m \|z\|^2 \leq z^T \nabla^2 f(x) z \leq M \|z\|^2$$

L'algorithme 3 de BFGS converge vers l'unique minimum x^* de f .

Théorème 14. — Avec les hypothèses du théorème précédent, si de plus $\nabla^2 f(x^*)$ est localement Lipschitz alors la convergence de l'algorithme 3 est superlinéaire.

1.2.6. Méthode du gradient conjugué. — L'algorithme que nous présentons dans cette section possède deux intérêts. Il permet de résoudre des systèmes linéaires à coefficients strictement positifs de grande taille et il sert d'algorithme de base pour la résolution itérée de problèmes d'optimisation non linéaire.

1.2.6.1. Fonctions quadratiques. — Résoudre le système linéaire $Ax = b$ où $A \in \mathcal{M}_n(\mathbb{R})$ symétrique positive définie, $x \in \mathbb{R}^n$, $b \in \mathbb{R}^n$ est équivalent à résoudre le problème de minimisation de $\phi(x) = \frac{1}{2}x^T Ax - b^T x$. Ces deux problèmes ont effectivement une même solution unique. En tout point x qui n'est pas cette solution, on note le "reste" $r(x) = Ax - b = \nabla \phi$ qui est non nul.

Considérons maintenant la définition suivante

Définition 15 (Directions conjuguées). — Un ensemble $\{p^0, p^1, \dots, p^l\}$ de vecteurs de \mathbb{R}^n est dit A -conjugué si $\forall i \neq j, (p^i)^T A p^j = 0$.

Une telle famille est libre. Un exemple d'une telle famille est donnée par une famille de vecteurs propres orthogonaux de A .

1.2.6.1.1. Minimisation suivant les directions conjuguées. — Considérons une séquence

$$(13) \quad x^{k+1} = x^k + l^k p^k$$

où les suites $(x^k)_{k \in \mathbb{N}}$ et $(p^k)_{k \in \mathbb{N}}$ sont des suites de vecteurs de \mathbb{R}^n et la suite $(l^k)_{k \in \mathbb{N}}$ est une suite de vecteurs de \mathbb{R} . On a alors

$$\phi(x^{k+1}) = \phi(x^k + l^k p^k) = \phi(x^k) + l^k (x^k)^T A p^k - b^T p^k l^k + \frac{1}{2} (l^k)^2 (p^k)^T A p^k$$

Cette expression définit une fonction convexe de la variable l^k . Son unique minimum est obtenu pour

$$(14) \quad l^k = \frac{b^T p^k - (x^k)^T A p^k}{(p^k)^T A p^k} = - \frac{(r^k)^T p^k}{(p^k)^T A p^k}$$

où $r^k = r(x^k)$.

Théorème 15. — Quel que soit $x^0 \in \mathbb{R}^n$, la séquence générée par (13), où l^k est donné par (14) et où $\{p^0, p^1, \dots, p^{n-1}\}$ est une famille de n directions A -conjuguées, avec $A \in \mathcal{M}_n(\mathbb{R})$ symétrique définie positive, atteint la solution x^* de $Ax = b$ en au plus n étapes.

Démonstration. — Les directions p^0, p^1, \dots, p^{n-1} sont linéairement indépendantes donc elles engendrent \mathbb{R}^n . Il existe donc n réels $\sigma^0, \dots, \sigma^{n-1}$ tels que

$$x^* - x^0 = \sigma^0 p^0 + \dots + \sigma^{n-1} p^{n-1}$$

En multipliant à gauche cette égalité par $(p^k)^T A$ il vient, pour tout $k = 0, \dots, n-1$,

$$\sigma^k = \frac{(p^k)^T A(x^* - x^0)}{(p^k)^T A p^k}$$

Nous allons maintenant montrer que $\sigma^k = l^k$. En k étapes on calcule

$$x^k = x^0 + l^0 p^0 + \dots + l^{k-1} p^{k-1}$$

tel que $(p^k)^T A(x^k - x^0) = 0$. Il vient donc

$$\begin{aligned} (p^k)^T A(x^* - x^0) &= (p^k)^T A(x^* - x^0 - x^k + x^0) \\ &= (p^k)^T A(x^* - x^k) \\ &= (p^k)^T (b - A x^k) = -(p^k)^T r^k \end{aligned}$$

d'où $\sigma^k = -\frac{(p^k)^T r^k}{(p^k)^T A p^k} = l^k$ d'où $x^n = x^*$. Ceci prouve qu'on a trouvé la décomposition de x^* dans la base p^0, p^1, \dots, p^{n-1} en minimisant successivement suivant les directions conjuguées la fonction quadratique $\phi(x) = \frac{1}{2} x^T A x - b^T x$. \square

Théorème 16. — *Quel que soit $x^0 \in \mathbb{R}^n$, la séquence générée par (13), où l^k est donné par (14) et où $\{p^0, p^1, \dots, p^{n-1}\}$ est une famille de n directions A -conjuguées avec $A \in \mathcal{M}_n(\mathbb{R})$ symétrique définie positive, possède les propriétés suivantes*

1. $(r^k)^T p^i = 0$ pour tout $k = 0, \dots, n-1$, et tout $i = 0, \dots, k-1$
2. x^k est le minimum global de $\mathbb{R}^n \ni x \mapsto \phi(x)$ sur l'espace affine $\{x = x^0 + \sum_{j=0}^{k-1} \mu^j p^j, (\mu^0, \dots, \mu^{k-1}) \in \mathbb{R}^k\}$.

Démonstration. — De manière générale, \tilde{x} la combinaison linéaire

$$\tilde{x} = x^0 + \tilde{\alpha}^0 p^0 + \dots + \tilde{\alpha}^{k-1} p^{k-1}$$

procurant la plus petite valeur de $\mathbb{R}^n \ni x \mapsto \phi(x)$ est telle que

$$\nabla \phi(\tilde{x})^T p^i = 0$$

pour tout $i = 0, \dots, k-1$, c.-à-d. $r(\tilde{x})^T p^i = 0$.

Prouvons maintenant le premier point du théorème. Le premier point x^1 calculé par la séquence $k = 1$ est obtenu en calculant

$$l^0 = \operatorname{argmin}_l \phi(x^0 + l p^0)$$

On a donc $\nabla \phi(x^1)^T p^0 = 0$. Raisonnons maintenant par récurrence. Supposons qu'on ait établi pour un certain $k > 1$

$$(r^{k-1})^T p^i = 0, \forall i = 0, \dots, k-2$$

Calculons maintenant $(r^k)^T p^i$ pour tout $i = 0, \dots, k-1$.

$$r^k = A x^k - b = A(x^{k-1} + l^{k-1} p^{k-1}) - b = r^{k-1} + l^{k-1} A p^{k-1}$$

d'où

$$(r^k)^T p^{k-1} = (r^{k-1})^T p^{k-1} + l^{k-1} (p^{k-1})^T A p^{k-1}$$

or $l^{k-1} = -\frac{(r^{k-1})^T p^{k-1}}{(p^{k-1})^T A p^{k-1}}$ d'où $(r^k)^T p^{k-1} = 0$. D'autre part, pour $i = 0, \dots, k-2$,

$$(r^k)^T p^i = (r^{k-1})^T p^i + l^{k-1} (p^{k-1})^T A p^i = 0$$

Ce qui achève de prouver le premier point du théorème.

Pour prouver le second point du théorème il suffit d'expliciter le gradient de l'application $\mathbb{R}^k \ni l \mapsto \phi(x^0 + l^0 p^0 + \dots + l^{k-1} p^{k-1})$. Sa i -ème coordonnée vaut $\nabla \phi(x^k)^T p^i = (r^k)^T p^i = 0$ quel que soit $i = 0, \dots, k-1$. Donc le deuxième point est prouvé. \square

1.2.6.1.2. Calcul des directions conjuguées. — On peut construire une base de vecteurs propres orthogonaux car A est symétrique. Ces vecteurs sont par construction A -conjugués et constituent un choix possible de directions conjuguées. Mais le calcul de ces vecteurs est en général très coûteux en temps de calculs. En fait on peut calculer les directions conjuguées au fur et à mesure qu'on en a besoin par la récurrence suivante

$$p^k = -r^k + \beta^k p^{k-1}$$

qui s'interprète comme l'opposé du gradient au nouveau point altéré par la direction précédente. Pour que p^k et p^{k-1} soient des directions A -conjuguées on choisit

$$(15) \quad \beta^k = \frac{(r^k)^T A p^{k-1}}{(p^{k-1})^T A p^{k-1}}$$

La direction p^k est également A -conjuguée avec les directions p^i pour $i = 0, \dots, k-2$. Pour démontrer ce résultat on utilise la proposition suivante

Proposition 3. — *Avec les notations précédentes, si $x^k \neq x^*$ alors les propriétés suivantes sont vérifiées*

1. $(r^k)^T r^i = 0$, pour tout $i = 0, \dots, k-1$
2. $\text{Vect}(r^0, r^1, \dots, r^k) = \text{Vect}(r^0, Ar^0, \dots, A^k r^0)$ (ce dernier sous espace est dénommé sous-espace de Krylov)
3. $\text{Vect}(p^0, p^1, \dots, p^k) = \text{Vect}(r^0, Ar^0, \dots, A^k r^0)$
4. $(p^k)^T A p^i = 0$ pour tout $i = 0, \dots, k-1$

On peut calculer

$$(p^i)^T A p^k = -(p^i)^T A r^k + \beta^k (p^i)^T A p^{k-1} = -(p^i)^T A r^k$$

On sait d'après le théorème 16 que $(r^k)^T p^i = 0$ pour tout $i = 0, \dots, k-2$. Or d'après les points 2 et 3 de la proposition 3 on sait que

$$A p^i \in A \cdot \text{Vect}(r^0, Ar^0, \dots, A^i r^0) = \text{Vect}(Ar^0, A^2 r^0, \dots, A^{i+1} r^0) \subset \text{Vect}(p^0, p^1, \dots, p^{i+1})$$

Donc il existe un vecteur $(\gamma^0, \dots, \gamma^{i+1}) \in \mathbb{R}^{i+2}$ tel que

$$(p^i)^T A r^k = \sum_{j=0}^{i+1} \gamma^j (p^j)^T r^k = 0$$

quel que soit $i = 0, \dots, k-2$. En conclusion, les directions $\{p^0, \dots, p^k\}$ sont A -conjuguées.

On peut en outre simplifier les expressions (14) et (15) en tirant partie des propriétés de la famille $(r^i)_{i=0, \dots, k-1}$ avec r^k . On obtient alors

$$(16) \quad l^k = \frac{(r^k)^T r^k}{(p^k)^T A p^k}, \quad \beta^{k+1} = \frac{(r^{k+1})^T r^{k+1}}{(r^k)^T r^k}$$

Une fois ces simplifications effectuées on aboutit à l'algorithme suivant

Algorithme 4 (Algorithme du gradient conjugué). — À partir de $x^0 \in \mathbb{R}^n$ quelconque calculer $r^0 = Ax^0 - b$ et $p^0 = -r^0$. Itérer

$$\begin{aligned} l^k &= \frac{(r^k)^T r^k}{(p^k)^T A p^k} \\ x^{k+1} &= x^k + l^k p^k \\ r^{k+1} &= r^k + l^k A p^k \\ \beta^{k+1} &= \frac{\|r^{k+1}\|^2}{\|r^k\|^2} \\ p^{k+1} &= -r^{k+1} + \beta^{k+1} p^k \end{aligned}$$

On peut caractériser la vitesse de convergence de l'algorithme 4.

Proposition 4. — Soit $A \in \mathcal{M}_n(\mathbb{R})$ symétrique définie positive, $K(A) = \frac{\lambda_1}{\lambda_n}$ le rapport entre la plus petite et la plus grande des valeurs propres de A (aussi appelé nombre de conditionnement). Les itérations de l'algorithme 4 du gradient conjugué appliqué à $\mathbb{R}^n \ni x \mapsto \phi(x) = \frac{1}{2}x^T A x - b^T x \in \mathbb{R}$ avec $b \in \mathbb{R}^n$, vérifient l'inégalité

$$\|x^k - x^*\|_A \leq \left(\frac{\sqrt{K(A)} - 1}{\sqrt{K(A)} + 1} \right)^{2k} \|x^0 - x^*\|_A$$

Il est possible d'améliorer cette vitesse de convergence par une technique de préconditionnement utilisant un changement de variable linéaire, on se reportera à [21] pour un exposé de ces techniques.

1.2.6.2. Application aux fonctions non linéaires. — Lorsque la fonction à minimiser n'est pas quadratique, on ne peut pas calculer explicitement le pas optimal le long d'une direction de descente et la notion de "résidu" n'a pas le même sens. On peut néanmoins transposer les idées de l'algorithme du gradient conjugué de la manière suivante

Algorithme 5 (Algorithme de Fletcher-Reeves). — À partir de $x^0 \in \mathbb{R}^n$ quelconque calculer $f^0 = f(x^0)$, $r^0 = \nabla f(x^0)$ et $p^0 = -r^0$. Itérer

1. Choisir l^k par une recherche linéaire dans la direction p^k

2.

$$\begin{aligned}
x^{k+1} &= x^k + l^k p^k \\
r^{k+1} &= \nabla f(x^{k+1}) \\
\beta^{k+1} &= \frac{\|r^{k+1}\|^2}{\|r^k\|^2} \\
p^{k+1} &= -r^{k+1} + \beta^{k+1} p^k
\end{aligned}$$

Tout comme l'algorithme de gradient conjugué, l'implémentation de cet algorithme ne requiert que peu de calculs et peu de mémoire quelle que soit la taille du système. Pour assurer la décroissance entre deux itérations, on utilisera dans la recherche linéaire les conditions de Wolfe (voir définition 13) avec typiquement des paramètres $0 < c_1 < c_2 < 1/2$.

Une amélioration pratique de cet algorithme est

Algorithme 6 (Algorithme de Polak-Ribière). — À partir de $x^0 \in \mathbb{R}^n$ quelconque calculer $f^0 = f(x^0)$, $r^0 = \nabla f(x^0)$ et $p^0 = -r^0$. Itérer choisir l^k par une recherche linéaire dans la direction p^k

$$\begin{aligned}
x^{k+1} &= x^k + l^k p^k \\
r^{k+1} &= \nabla f(x^{k+1}) \\
\beta^{k+1} &= \frac{(r^{k+1})^T (r^{k+1} - r^k)}{\|r^k\|^2} \\
p^{k+1} &= -r^{k+1} + \beta^{k+1} p^k
\end{aligned}$$

Les conditions de Wolfe peuvent également être utilisées mais n'assurent pas la décroissance.

Proposition 5. — Soit $\mathbb{R}^n \ni x \mapsto f(x) \in \mathbb{R}$ deux fois différentiable possédant un unique minimum x^* tel que $\nabla^2 f(x)$ est défini positif. Les itérations des algorithmes 5 et 6 avec recherche linéaire exacte (c.-à-d. que l^k minimise $f(x^k + l^k p^k)$ à chaque itération)

$$\lim_{k \rightarrow +\infty} \frac{\|x^{k+n} - x^*\|}{\|x^k - x^*\|} = 0$$

1.3. Principes de l'optimisation sous contraintes

1.3.1. Contraintes égalités. — On s'intéresse désormais au problème de la minimisation dans \mathbb{R}^{n+m} , $n + m < \infty$ d'une fonction différentiable (fonction coût) $\mathbb{R}^{n+m} \ni (x, u) \mapsto f(x, u) \in \mathbb{R}$ sous la contrainte $c(x, u) = 0$ définie par une fonction différentiable $c : \mathbb{R}^{n+m} \ni (x, u) \mapsto c(x, u) \in \mathbb{R}^n$ localement inversible par rapport à

la variable x . En d'autres termes x est localement déterminé par u par la relation $c(x, u) = 0$:

$$(17) \quad \min_{c(x,u)=0} f(x, u)$$

Une condition suffisante par le théorème des fonctions implicites est que le Jacobien partiel $\frac{\partial c}{\partial x}$ est inversible au voisinage des points tels que $c(x, u) = 0$. On supposera cette condition réalisée.

Ce problème revient à chercher un minimum de f sur l'ensemble $\{(x, u) = c^{-1}(0)\}$ qui est un fermé de \mathbb{R}^{n+m} . Cet ensemble ne contient pas de voisinage de tous ses points. Ainsi, tout déplacement infinitésimal autour d'un des points de son bord ne fournit pas nécessairement de point admissible. Les conditions d'optimalité en seront donc modifiées.

1.3.1.1. Élimination des variables. — Dans cette approche on calcule la variation de la fonction coût en préservant la contrainte. Une variation $(\delta x, \delta u) \in \mathbb{R}^{n+m}$ entraîne une variation

$$\delta f \approx \frac{\partial f}{\partial x} \delta x + \frac{\partial f}{\partial u} \delta u$$

Maintenir la contrainte $c(x + \delta x, u + \delta u) = 0$ impose

$$0 \approx \frac{\partial c}{\partial x} \delta x + \frac{\partial c}{\partial u} \delta u$$

Par hypothèse, $\frac{\partial c}{\partial x}$ est inversible. On peut donc établir

$$\delta f \approx \left(-\frac{\partial f}{\partial x} \left(\frac{\partial c}{\partial x} \right)^{-1} \frac{\partial c}{\partial u} + \frac{\partial f}{\partial u} \right) \delta u$$

et à la limite

$$(18) \quad \left(\frac{\partial f}{\partial u} \right)_{c=0} = -\frac{\partial f}{\partial x} \left(\frac{\partial c}{\partial x} \right)^{-1} \frac{\partial c}{\partial u} + \frac{\partial f}{\partial u}$$

On remarquera que la variation de f en maintenant la contrainte est différente de la variation de f sans maintenir la contrainte (dernier terme de la précédente équation).

On peut étendre la définition 6 à notre cas.

Définition 16. — On dit que (x^*, u^*) est un point stationnaire de f sous la contrainte égalité c si $c(x^*, u^*) = 0$ et $\left(\frac{\partial f}{\partial u} \right)_{c=0} (x^*, u^*) = 0$.

Définition 17. — On dit que $(x^*, u^*) \in \mathbb{R}^n \times \mathbb{R}^m$ est un minimum (optimum) local de f sous la contrainte $c(x, u) = 0$ si $\exists \epsilon > 0$ tel que $f(x^*, u^*) \leq f(x, u)$, $\forall (x, u)$ tel que $\|(x, u) - (x^*, u^*)\| \leq \epsilon$ et $c(x, u) = 0$ (on dit que c'est un minimum local strict si la précédente inégalité est stricte). On dit que (x^*, u^*) est solution locale du problème

$$\min_{c(x,u)=0} f(x, u)$$

1.3.1.2. *Équations adjointes.* — Définissons le Lagrangien du problème (17) en adjoignant les contraintes à la fonction coût :

$$(19) \quad \mathbb{R}^{2n+m} \ni (x, u, \lambda) \mapsto \mathcal{L}(x, u, \lambda) = f(x, u) + \lambda^T c(x, u) \in \mathbb{R}$$

D'après la définition 6, un point stationnaire de \mathcal{L} (sans contraintes) est caractérisé par

$$\frac{\partial \mathcal{L}}{\partial x} = 0, \quad \frac{\partial \mathcal{L}}{\partial u} = 0, \quad \frac{\partial \mathcal{L}}{\partial \lambda} = 0$$

Proposition 6. — Si $(x^*, u^*, \lambda^*) \in \mathbb{R}^{2n+m}$ est un point stationnaire de \mathcal{L} alors (x^*, u^*) est un point stationnaire de f sous la contrainte c .

Démonstration. — D'une part on a $c(x^*, u^*) = 0$. On va montrer que pour toute variation $(\delta x, \delta u) \in \mathbb{R}^{n+m}$ telle que $c(x^* + \delta x, u^* + \delta u) = 0$ au premier ordre on a $f(x^* + \delta x, u^* + \delta u) = f(x^*, u^*)$ au premier ordre. Le calcul suivant permet de conclure

$$\begin{aligned} f(x^* + \delta x, u^* + \delta u) - f(x^*, u^*) &\approx \frac{\partial f}{\partial x}(x^*, u^*) \delta x + \frac{\partial f}{\partial u}(x^*, u^*) \delta u \\ &\approx -\lambda^* \left(\frac{\partial c}{\partial x}(x^*, u^*) \delta x + \frac{\partial c}{\partial u}(x^*, u^*) \delta u \right) \approx 0 \end{aligned}$$

□

1.3.1.3. *Multiplicateurs de Lagrange.* — Le vecteur λ dans l'équation (19) est appelé vecteur des multiplicateurs de Lagrange. Les multiplicateurs de Lagrange apportent beaucoup d'information sur le problème d'optimisation sous contraintes comme en attestent les propriétés suivantes.

Proposition 7. — Si $(x^*, u^*, \lambda^*) \in \mathbb{R}^{2n+m}$ est un point stationnaire de \mathcal{L} , alors

$$(\nabla f)^T(x^*, u^*) = -(\lambda^*)^T (\nabla c)^T(x^*, u^*)$$

Cette proposition relie le gradient du coût et celui de la contrainte à un point stationnaire.

Proposition 8. — Si $(x^*, u^*, \lambda^*) \in \mathbb{R}^{2n+m}$ est un point stationnaire de \mathcal{L} , alors en ce point

$$(20) \quad (\nabla^2 f(x, u))^T(x^*, u^*) = \begin{pmatrix} -\frac{\partial c}{\partial u}^T \left(\frac{\partial c}{\partial x}^T \right)^{-1} & I \end{pmatrix} \begin{pmatrix} \frac{\partial^2 \mathcal{L}}{\partial x^2} & \frac{\partial^2 \mathcal{L}}{\partial x \partial u} \\ \frac{\partial^2 \mathcal{L}}{\partial u \partial x} & \frac{\partial^2 \mathcal{L}}{\partial u^2} \end{pmatrix} \begin{pmatrix} -\left(\frac{\partial c}{\partial x} \right)^{-1} \frac{\partial c}{\partial u} \\ I \end{pmatrix}$$

Cette dernière proposition permet souvent de lever l'ambiguïté sur la nature du point stationnaire. Les multiplicateurs de Lagrange λ^* interviennent dans la matrice centrale du calcul du Hessian (20).

Définissons maintenant un problème proche du problème (17) en utilisant $\delta c \in \mathbb{R}^m$.

$$(21) \quad \min_{c(x, u) = \delta c} f(x, u)$$

On suppose que (17) et (21) possèdent des solutions (x^*, u^*) et (\bar{x}, \bar{u}) respectivement. On note $\frac{\partial f^*}{\partial c}$ le vecteur composé des limites des valeurs $(f(\bar{x}, \bar{u}) - f(x^*, u^*))/\delta c_i$ où $\delta c_i \in \mathbb{R}$ correspond à la i -ème composante de δc .

Proposition 9 (coût marginal). — Si $(x^*, u^*, \lambda^*) \in \mathbb{R}^{2n+m}$ est un point stationnaire de \mathcal{L} , alors

$$\frac{\partial f^*}{\partial c} = (\lambda^*)^T$$

Cette dernière propriété permet de quantifier le coût (marginal) d'une variation de la contrainte (relâchement ou durcissement) sur la valeur de l'optimum.

1.4. Contraintes inégalités

On s'intéresse maintenant au problème de la minimisation dans \mathbb{R}^n , $n < \infty$ d'une fonction différentiable (fonction coût) $\mathbb{R}^n \ni x \mapsto f(x) \in \mathbb{R}$ sous la contrainte $c(x) \leq 0$ définie par une fonction différentiable $\mathbb{R}^n \ni x \mapsto c(x) \in \mathbb{R}^m$, $m < \infty$ (on ne spécifie plus de variable par rapport à laquelle cette fonction sera inversible, et on ne fait pas d'hypothèse sur les entiers n et m).

$$(22) \quad \min_{c(x) \leq 0} f(x)$$

Définition 18. — On dit que $x^* \in \mathbb{R}^n$ est un minimum (optimum) local de f sous la contrainte $c(x) \leq 0$ si $\exists \epsilon > 0$ tel que $f(x^*) \leq f(x)$, $\forall x$ tel que $\|x - x^*\| \leq \epsilon$ et $c(x) \leq 0$ (on dit que c'est un minimum local strict si $\exists \epsilon > 0$ tel que $f(x^*) < f(x)$, $\forall x$ tel que $\|x - x^*\| \leq \epsilon$ et $c(x) \leq 0$). On dit que (x^*) est solution locale du problème

$$\min_{c(x) \leq 0} f(x)$$

1.4.1. Conditions de Karush-Kuhn-Tucker. —

1.4.1.1. Cas des contraintes affines. — Considérons dans un premier temps le cas où le vecteur $c(x)$ est constitué de contraintes affines $\mathbb{R}^n \ni x \mapsto c_i(x) \in \mathbb{R}$, pour $i = 1, \dots, m$ telles que $c(x) \leq 0$ définit un ensemble d'intérieur non vide.

Définition 19. — On appelle cône convexe engendré par $\{v_i, i = 1, \dots, k\}$ un ensemble de vecteurs de \mathbb{R}^n , l'ensemble des combinaisons linéaires à coefficients positifs ou nuls de vecteurs de $\{v_i, i = 1, \dots, k\}$.

Les cônes convexes possèdent d'intéressantes propriétés. On peut établir le résultat suivant

Lemme 1 (Farkas). — Soient $\{a_i, i = 1, \dots, k\}$ une collection de vecteurs de \mathbb{R}^n et g un vecteur de \mathbb{R}^n , les deux propositions suivantes sont équivalentes

1. Il n'existe pas $p \in \mathbb{R}^n$ tel que $\{g^T p < 0 \text{ et } a_i^T p \geq 0 \text{ pour tout } i = 1, \dots, k\}$
2. g appartient au cône convexe \mathcal{C} engendré par $\{a_i, i = 1, \dots, k\}$

Démonstration. — Démontrons que la seconde proposition entraîne la première. Supposons cette seconde proposition vraie. Supposons qu'il existe $\lambda_i \geq 0$, $i = 1, \dots, p$ tels que $g = \sum_{i=1}^p \lambda_i a_i$. S'il existe $h \in \mathbb{R}^n$ tel que $a_i^T h \geq 0$ pour tout $i = 1, \dots, p$, alors on a $g^T h \geq 0$ ce qui est une contradiction.

Démontrons maintenant que la première proposition entraîne la seconde. Supposons cette première proposition vraie. Supposons qu'il n'existe pas $p \in \mathbb{R}^n$ tel que

$$g^T p < 0 \text{ et } a_i^T p \geq 0 \quad \forall i = 1, \dots, p$$

Montrons que nécessairement $g \in \mathcal{C}$. Supposons que ce soit faux. Définissons h la projection de g sur \mathcal{C} comme un vecteur $h \in \mathcal{C}$ qui minimise la distance au carré $f(h) = \|h - g\|^2$. Un tel h existe car il est solution d'un problème de minimisation d'une fonction continue α -convexe sur K convexe fermé. Il n'est pas à l'intérieur de K car on aurait alors $\nabla f(h) = h - g \neq 0$. Il est sur le bord du cône et on a nécessairement $h^T(h - g) = 0$. Posons $p = h - g \in \mathbb{R}^n$. Un calcul donne

$$g^T p = g^T(h - g) = (g - h)^T(h - g) = -\|h - g\|^2$$

Or par hypothèse $g \notin \mathcal{C}$. Donc $h \neq g$ et donc

$$g^T p < 0$$

D'autre part, $h \in \mathcal{C}$ réalise le minimum global de $f(h) = \|h - g\|^2$. Donc, quel que soit $v \in \mathcal{C}$, on a pour tout $\alpha \in [0, 1]$

$$\begin{aligned} \|h - g\|^2 &\leq \|(1 - \alpha)h + \alpha v - g\|^2 \\ &\leq \|h - g\|^2 + \alpha^2 \|v - h\|^2 + 2\alpha(h - g)^T(v - h) \end{aligned}$$

On en déduit que pour tout $\alpha \in [0, 1]$

$$\alpha^2 \|v - h\|^2 + 2\alpha(h - g)^T(v - h) \geq 0$$

Nécessairement on a donc $(h - g)^T(v - h) \geq 0$. En conclusion, pour tout $v \in \mathcal{C}$, on a $p^T v \geq 0$. Cette inégalité est vraie pour le cas particulier des vecteurs a_i , $i = 1, \dots, k$. On a donc trouvé $p \neq 0$ tel que $g^T p < 0$ et $p^T a_i \geq 0$, pour tout $i = 1, \dots, k$. C'est en contradiction avec l'hypothèse et conclut la preuve. \square

La notion de cône convexe est utilisé pour caractériser les solutions de notre problème d'optimisation.

Proposition 10. — *Si x^* est solution du problème (22) alors (sous les hypothèses du paragraphe §1.4.1.1) $\nabla f(x^*)$ est dans le cône convexe engendré par $\{-\nabla c_i(x^*), i \in \mathcal{I}\}$ où $\mathcal{I} = \{i = 1, \dots, m \text{ tel que } c_i(x^*) = 0\}$ (famille des indices des contraintes actives).*

Démonstration. — On va montrer que si $\nabla f(x^*)$ n'appartient pas au cône convexe engendré par $\{-\nabla c_i(x^*), i \in \mathcal{I}\}$ alors x^* n'est pas minimum. D'après le lemme 1 de Farkas, $\exists h \in \mathbb{R}^n$ tel que $\nabla f(x^*)^T h < 0$ et $-\nabla c_i(x^*)^T h \geq 0$ pour tout $i \in \mathcal{I}$. En suivant cette direction pour $\delta \in \mathbb{R}$ il vient

$$\begin{aligned} f(x^* + h\delta) &= f(x^*) + \nabla f(x^*)^T h\delta + o(\delta^2) \\ c_i(x^* + h\delta) &= c_i(x^*) + \nabla c_i(x^*)^T h\delta, \forall i \in \mathcal{I} \end{aligned}$$

Donc pour δ suffisamment petit

$$\begin{aligned} f(x^* + h\delta) &< f(x^*) \\ c_i(x^* + h\delta) &\leq c(x^*) \leq 0, \forall i = 1, \dots, m \end{aligned}$$

Tout voisinage de x^* contient un meilleur point que x^* qui n'est donc pas minimum. \square

Dans le cas où les contraintes $\mathbb{R}^n \ni x \mapsto c_i(x) \in \mathbb{R}$ ne sont pas affines, il peut ne pas exister de direction donnée par le lemme de Farkas utilisable pour trouver un meilleur point que x^* . La courbure des contraintes peut gêner une telle construction. D'autre part, le cône convexe engendré par les contraintes actives peut dégénérer (la frontière peut ressembler à un point de rebroussement).

Néanmoins, il est possible d'étendre le résultat précédent dans le cas général sous une hypothèse supplémentaire souvent peu gênante en pratique.

Théorème 17. — *[Conditions KKT] Considérons un point $x^* \in \mathbb{R}^n$. Notons la famille des indices des contraintes actives en x^* par $\mathcal{I} = \{i \in \{1, \dots, m \text{ tel que } c_i(x^*) = 0\}\}$. Supposons que x^* est un point régulier pour ces contraintes, c.-à-d. que la famille $(\nabla c_i(x^*))_{i \in \mathcal{I}}$ est une famille libre (on dit que les contraintes sont qualifiées). Sous ces hypothèses, si x^* est une solution du problème (22) alors $\nabla f(x^*)$ appartient au cône convexe engendré par $(-\nabla c_i(x^*))_{i \in \mathcal{I}}$. En conséquence, $\exists \lambda_i \geq 0, i = 1, \dots, m$ tels que $\nabla f(x^*) = -\sum_{i=1}^m \lambda_i \nabla c_i(x^*)$ et $\lambda_i c_i(x^*) = 0$ pour $i = 1, \dots, n$.*

Ce résultat (admis, voir [19]) est une condition nécessaire aux points réguliers candidats pour être minimum. On calculera par les multiplicateurs de Lagrange la combinaison linéaire qui permet d'exprimer le gradient du coût en fonction des gradients des contraintes actives et on vérifiera que les multiplicateurs sont bien tous positifs ou nuls.

En pratique on travaillera souvent avec des fonctions convexes et on exploitera le résultat suivant

Proposition 11. — *Soit le problème d'optimisation $\min_{c(x) \leq 0} f(x)$ où les fonctions f et c sont différentiables et convexes. On suppose qu'il existe $x \in \mathbb{R}^n$ tel que $c(x) < 0$. Les conditions KKT du théorème 17 sont nécessaires et suffisantes pour que x^* soit un minimum global.*

1.4.2. Dualité. — Considérons le problème de la minimisation dans X sous ensemble de \mathbb{R}^n , $n < \infty$ d'une fonction différentiable (fonction coût) $X \ni x \mapsto f(x) \in \mathbb{R}$ sous la contrainte $c(x) \leq 0$ définie par une fonction différentiable $\mathbb{R}^n \ni x \mapsto c(x) \in \mathbb{R}^m$

$$(23) \quad \min_{x \in X, c(x) \leq 0} f(x)$$

Considérons alors le Lagrangien associé à ce problème

$$(24) \quad X \times L \ni (x, \lambda) \mapsto \mathcal{L}(x, \lambda) = f(x) + \lambda^T c(x) \in \mathbb{R}$$

où $L \subset \mathbb{R}^m$.

Définition 20. — On dit que (x^*, λ^*) est un point selle de \mathcal{L} si x^* est un minimum pour $X \ni x \mapsto \mathcal{L}(x, \lambda^*)$ et λ^* est un maximum pour $L \ni \lambda \mapsto \mathcal{L}(x^*, \lambda)$ ou encore si

$$(25) \quad \sup_{\lambda \in L} \mathcal{L}(x^*, \lambda) = \mathcal{L}(x^*, \lambda^*) = \inf_{x \in X} \mathcal{L}(x, \lambda^*)$$

On a aussi pour tout $x \in X$ et pour tout $\lambda \in L$

$$\mathcal{L}(x^*, \lambda) \leq \mathcal{L}(x^*, \lambda^*) \leq \mathcal{L}(x, \lambda^*)$$

D'après le théorème suivant, on peut intervertir l'ordre de la maximisation et de la minimisation

Théorème 18 (Théorème du point selle). — Si (x^*, λ^*) est un point selle de \mathcal{L} sur $X \times L$ alors

$$\sup_{\lambda \in L} \inf_{x \in X} \mathcal{L}(x, \lambda) = \mathcal{L}(x^*, \lambda^*) = \inf_{x \in X} \sup_{\lambda \in L} \mathcal{L}(x, \lambda)$$

Démonstration. — Par définition, on a

$$\mathcal{L}(x, \lambda) \leq \sup_{\lambda \in L} \mathcal{L}(x, \lambda)$$

et donc

$$\inf_{x \in X} \mathcal{L}(x, \lambda) \leq \inf_{x \in X} \sup_{\lambda \in L} \mathcal{L}(x, \lambda)$$

et finalement

$$\sup_{\lambda \in L} \inf_{x \in X} \mathcal{L}(x, \lambda) \leq \inf_{x \in X} \sup_{\lambda \in L} \mathcal{L}(x, \lambda)$$

Utilisons maintenant l'égalité du point selle (25) sur

$$\begin{aligned} \inf_{x \in X} \sup_{\lambda \in L} \mathcal{L}(x, \lambda) &\leq \sup_{\lambda \in L} \mathcal{L}(x^*, \lambda) = \mathcal{L}(x^*, \lambda^*) \\ &= \inf_{x \in X} \mathcal{L}(x, \lambda^*) \\ &\leq \sup_{\lambda \in L} \inf_{x \in X} \mathcal{L}(x, \lambda) \end{aligned}$$

Finalement

$$\sup_{\lambda \in L} \inf_{x \in X} \mathcal{L}(x, \lambda) = \mathcal{L}(x^*, \lambda^*) = \inf_{x \in X} \sup_{\lambda \in L} \mathcal{L}(x, \lambda)$$

□

Lorsque (x^*, λ^*) n'est pas un point selle, l'égalité précédente n'est pas vraie, on n'a que

$$\sup_{\lambda \in L} \inf_{x \in X} \mathcal{L}(x, \lambda) \leq \inf_{x \in X} \sup_{\lambda \in L} \mathcal{L}(x, \lambda)$$

on parle de “saut de dualité” lorsque les deux résultats sont différents.

On appelle problème primal le problème $\min_{x \in X} \max_{\lambda \in L} \mathcal{L}(x, \lambda)$, et problème dual $\max_{\lambda \in L} \min_{x \in X} \mathcal{L}(x, \lambda)$.

Les deux théorèmes suivants permettent de relier notre problème d'optimisation sous contraintes inégalités à l'existence d'un point selle.

Théorème 19 (Optimalité du point selle). — Si (x^*, λ^*) est un point selle de \mathcal{L} sur $X \times (\mathbb{R}^+)^m$, alors x^* est solution du problème (22).

Démonstration. — Par définition, quel que soit $\lambda \in (\mathbb{R}^+)^m$ on a

$$\mathcal{L}(x^*, \lambda) \leq \mathcal{L}(x^*, \lambda^*)$$

Ceci peut s'écrire

$$(26) \quad (\lambda - \lambda^*)^T c(x^*) \leq 0$$

En exprimant cette inéquation pour $\lambda = 0$ et $\lambda = 2\lambda^*$ il vient

$$(\lambda^*)^T c(x^*) = 0$$

Il vient aussi en exprimant le produit scalaire (26) composante par composante

$$c(x^*) \leq 0$$

D'autre part, $\mathcal{L}(x^*, \lambda^*) \leq \mathcal{L}(x, \lambda^*)$ pour tout $x \in X$. On a donc

$$f(x^*) \leq f(x) + (\lambda^*)^T c(x)$$

Donc pour tout $x \in X$ tel que $c(x) \leq 0$ on a $f(x^*) \leq f(x)$ ce qui achève la démonstration. \square

Sans hypothèse sur la régularité des f et de c , on peut énoncer un théorème analogue aux conditions KKT du théorème 17.

Théorème 20 (existence de point selle). — *Supposons $X \ni x \mapsto f(x) \in \mathbb{R}$ et $\mathbb{R}^n \ni x \mapsto c(x) \in \mathbb{R}^m$ convexe. Soit x^* solution du problème (22) tel que les contraintes actives en x^* sont qualifiées, alors il existe un point selle (x^*, λ^*) du Lagrangien (24).*

1.4.3. Algorithme de minimisation sous contraintes utilisant la dualité. —

Les algorithmes de cette section sont des algorithmes de recherche de point selle

Algorithme 7 (Algorithme d'Uzawa). — À partir de $x^0 \in \mathbb{R}^n$, $\lambda^0 \in \mathbb{R}^m$, $\alpha \in \mathbb{R}^+$ quelconques, on note $\mathbb{R}^m \ni \lambda \mapsto P(\lambda) \in (\mathbb{R}^+)^m$ la projection sur $(\mathbb{R}^+)^m$, itérer

$$\begin{aligned} &\text{résoudre } \min_{x \in \mathbb{R}^n} \mathcal{L}(x, \lambda^k), \text{ on note } x^{k+1} \text{ la solution} \\ &\lambda^{k+1} = P(\lambda^k + \alpha c(x^{k+1})) \end{aligned}$$

La seconde partie de la boucle à itérer consiste à effectuer un gradient à pas constant pour maximiser $\mathbb{R}^m \ni \lambda \mapsto \min_{x \in \mathbb{R}^n} \mathcal{L}(x, \lambda)$. Une variante de cet algorithme est

Algorithme 8 (Algorithme d'Arrow-Hurwicz). — À partir de $x^0 \in \mathbb{R}^n$, $\lambda^0 \in \mathbb{R}^m$ quelconques, $\varepsilon \in \mathbb{R}^+$ on note $\mathbb{R}^m \ni \lambda \mapsto P(\lambda) \in (\mathbb{R}^+)^m$ la projection sur $(\mathbb{R}^+)^m$, itérer

$$\begin{aligned} x^{k+1} &= x^k - \varepsilon \left(\nabla f(x^k) + \frac{\partial c}{\partial x}(x^k) \lambda^k \right) \\ \lambda^{k+1} &= P(\lambda^k + \alpha c(x^{k+1})) \end{aligned}$$

1.4.4. Méthode de contraintes actives pour les problèmes de programmation quadratique. — Nous allons maintenant présenter un algorithme efficace (y compris lorsque les dimensions n et m sont grandes) de résolution de problèmes de minimisation d'une fonction quadratique sous contraintes affines.

On va traiter le cas

$$f(x) = \frac{1}{2}x^T Gx + x^T d$$

sous contraintes affines

$$c(x) = Ax - b \leq 0$$

et on cherche à résoudre le problème de minimisation (22). On suppose que G est une matrice symétrique définie positive de $\mathcal{M}_n(\mathbb{R})$ et A une matrice de $\mathcal{M}(m, n)$, avec $m > n$. On notera a_i la i -ème ligne de A et b_i la i -ème composante de $b \in \mathbb{R}^m$.

L'algorithme constitue et met à jour un ensemble de contraintes actives jusqu'à ce que cet ensemble satisfasse les conditions de Karush-Kuhn-Tucker du théorème 17. La mise à jour s'effectue en repérant des directions d'amélioration. On suit ces directions jusqu'à atteindre une nouvelle contrainte bloquante. On élimine si besoin les contraintes, pour pouvoir avancer, en vérifiant le signe de leur multiplicateur de Lagrange. À chaque étape k , on définit W^k ensemble de travail (sous-ensemble) de A_c^k ensemble des contraintes actives au point x^k . On suppose que les gradients (a_i) des contraintes de W^k sont linéairement indépendants.

Algorithme 9 (Algorithme de contraintes actives QP)

À partir de W^0 ensemble de travail de départ, $x^0 \in \mathbb{R}^n$ point où les n contraintes de A constituant W^0 sont actives, itérer

1. Vérifier si x^k est un minimum de $f(x)$ sous les contraintes $a_i^T x - b_i \leq 0$, $i \in W^k$. Si c'est le cas l'algorithme s'achève et x^k est la solution recherchée. Sinon continuer au point 2.
2. Résoudre $\min_{a_i^T x - b_i \leq 0, i \in W^k} f(x)$
 - (a) Calculer une direction p^k solution de $\min_{a_i^T p = 0, i \in W^k} \frac{1}{2}p^T Gp + (Gx^k + d)^T p$
 - (b) si $p^k \neq 0$ noter $\alpha^k = \min \left(1, \min_{i \notin W^k, a_i^T p^k > 0} \frac{b_i - a_i^T x^k}{a_i^T p^k} \right)$
Si $\alpha^k < 1$ il existe un unique j réalisant le précédent minimum. Mettre à jour $W^{k+1} = W^k \cup \{j\}$.
 - (c) si $p^k = 0$ calculer les multiplicateurs de Lagrange $\lambda_i, i \in W^k$ du problème $\min_{a_i^T x - b_i = 0} f(x)$. Si tous ces multiplicateurs de Lagrange sont positifs l'algorithme s'achève et x^k est la solution recherchée. Dans le cas contraire éliminer de W^k la contrainte ayant le plus grand multiplicateur de Lagrange

Proposition 12. — Avec les notations précédentes, l'algorithme 9 de contraintes actives possède les propriétés suivantes

1. On ne revisite jamais exactement l'ensemble W^k .
2. La suite des directions $(p^k)_k \in \mathbb{N}$ revient n -périodiquement à la valeur $p^k = 0$.

Démonstration. — La suite des valeurs $(f(x^k))_k \in \mathbb{N}$ est décroissante. Elle est strictement décroissante aux indices tels que $\alpha^k > 0$. On ne revisite jamais exactement l'ensemble W^k car x^{k+1} est le minimum sous les contraintes d'indices éléments de W^k . La deuxième propriété découle du raisonnement suivant. Si $p^k \neq 0$ le point x^k est mis à jour. Soit le pas α^k utilisé dans cette mise à jour est égal à 1 et on a obtenu un minimiseur donc $p^{k+1} = 0$, soit $\alpha^k < 1$ et on doit rajouter une contrainte à l'ensemble W^k . Au bout de $l < n$ étapes, W^k contient n contraintes linéairement indépendantes. Le problème est complètement contraint et on a donc $p^{k+l} = 0$. \square

Proposition 13. — *Avec les notations précédentes, l'algorithme 9 de contraintes actives converge en un nombre fini d'itérations vers l'unique minimum global du problème (22).*

Démonstration. — Il existe un nombre fini de W^k possibles. L'algorithme parcourt cet ensemble jusqu'à obtenir l'optimum. \square

Cet algorithme est également très utilisé pour résoudre une succession de problèmes où une fonction coût non linéaire est approchée par son développement de Taylor au deuxième ordre et où les contraintes sont approchées par leur développement de Taylor au premier ordre. De tels algorithmes SQP (successive quadratic programming) sont détaillés dans [14, 13].

CHAPITRE 2

OPTIMISATION DE TRAJECTOIRES

2.1. Calcul des variations

2.1.1. Historique. — Au lieu d'un nombre fini de paramètres, on peut considérer qu'il s'agit de l'optimisation d'un nombre infini de valeurs successives de la fonction recherchée.

Ces problèmes sont très anciens. On pourra se référer à [27] pour un exposé complet de leur historique. Le problème de Dido est le calcul du contour de périmètre donné enfermant l'aire maximale. Il date du 9ème siècle avant JC. Galilée (1588) considéra des problèmes célèbres, mais la réelle percée vint de Johann Bernoulli qui mit au défi les plus illustres mathématiciens de son époque (son propre frère Jakob, Leibnitz, Newton). Le problème en question était le Brachistochrone : le calcul de la courbe permettant à un point matériel soumis à la gravité de rallier son point initial à un point d'arrivée donné.

D'autres problèmes importants sont le calcul des géodésiques, les problèmes isopérimétriques, et de manière générale les problèmes avec contraintes.

Les "sciences naturelles" (mécanique, optique, ...) ont été profondément marquées par le calcul des variations. De nombreuses lois de la nature se déduisent de principes variationnels, énonçant que parmi tous les mouvements possibles, celui qui se réalise est un extremum. Cette vision de la nature, indiqua que les réalisations humaines devaient elles aussi être des extrema : Brachistochrone ou problème de Newton.

Une citation célèbre d'Euler est : "il n'y a rien dans le monde qui ne se réalise sans la volonté de minimiser ou maximiser quelque chose".

2.1.2. Notions fondamentales. —

Définition 21 (Fonctionnelle). — Soit X un espace vectoriel normé. On appelle fonctionnelle une application de X dans \mathbb{R} .

Dans ce qui suit nous nous intéressons à l'espace vectoriel normé $X = D$ des fonctions continues à dérivées continues $\mathbb{R} \supset [t_1, t_2] \longrightarrow \mathbb{R}^n$, $n < \infty$ muni de la norme $X \ni u \mapsto \|u\|_D = \max_{t \in [t_1, t_2]} \|u(t)\| + \max_{t \in [t_1, t_2]} \|\dot{u}(t)\|$. On considère les

fonctionnelles du type

$$X \ni u \mapsto J(u) = \int_{t_1}^{t_2} L(u(t), \dot{u}(t), t) dt \in \mathbb{R}$$

où L est une fonction $\mathbb{R}^{2n+1} \ni (u, v, t) \mapsto L(u, v, t) \in \mathbb{R}$ continue possédant des dérivées partielles continues.

On cherchera à minimiser J sous certaines contraintes, en restreignant l'ensemble des fonctions considérées à $X \supset U = \{u \in X, u(t_1) = a, u(t_2) = b\}$ où a et b sont des réels donnés. C'est un ensemble convexe. On cherchera à résoudre le problème suivant

$$(27) \quad \min_{u \in U} J(u)$$

Définition 22. — On dit que u^* est un minimum (local) de J (c.-à-d. une solution du problème (27) s'il existe un voisinage $\mathcal{V}(u^*)$ tel que $J(u^*) \leq J(u)$ pour tout $u \in \mathcal{V}(u^*)$).

Définition 23 (Variation admissible). — On dit que $h \in X$ est une variation admissible au point $u \in U$ si $(u + h) \in U$.

Définition 24 (Différentielle de Gâteaux). — Soient J une fonctionnelle sur X espace vectoriel normé, $u^* \in U$ et $h \in X$ variation admissible. On appelle différentielle de Gâteaux de J au point u^* dans la direction h la limite

$$\delta J(u^*, h) = \lim_{\delta \rightarrow 0} \frac{1}{\delta} (J(u^* + \delta h) - J(u^*))$$

2.1.3. Conditions nécessaires d'extrémalité. — La notion de différentielle de Gâteaux remplace pour les fonctionnelles la notion de différentielle pour les fonctions. Une condition nécessaire pour que $u^* \in U$ soit un minimum est que la différentielle de Gâteaux $\delta J(u^*, h)$ doit être nulle pour toute variation admissible h . On va réécrire cette proposition en une équation exploitable à l'aide du résultat suivant.

Lemme 2 (Lemme de duBois-Reymond). — Soit $\mathbb{R} \supset [t_1, t_2] \ni t \mapsto \phi(t) \in \mathbb{R}$ une fonction continue. Si on a pour tout $\mathbb{R} \supset [t_1, t_2] \ni t \mapsto h(t) \in \mathbb{R}$ telle que $h(t_1) = h(t_2) = 0$

$$\int_{t_1}^{t_2} \phi(t) h(t) dt = 0$$

alors $\phi = 0$.

Démonstration. — Supposons que les hypothèses du lemme sont valides et que $\phi \neq 0$. Sans perte de généralité on peut supposer qu'elle est strictement positive en un point $t \in [t_1, t_2]$. Par continuité, elle est donc strictement positive sur $[t'_1, t'_2] \subset [t_1, t_2]$. Définissons h comme

$$h(t) = (t - t'_1)^2 (t - t'_2)^2 \mathbb{I}_{[t'_1, t'_2]}(t)$$

Avec cette variation, on obtient $\int_{t_1}^{t_2} \phi(t) h(t) dt > 0$. D'où une contradiction et donc $\phi = 0$. \square

L'équation $\delta J(u^*, h) = 0$ pour toute variation admissible se réécrit

$$\int_{t_1}^{t_2} \left(\frac{\partial L}{\partial u}(u^*, \dot{u}^*, t) h(t) + \frac{\partial L}{\partial \dot{u}}(u^*, \dot{u}^*, t) \dot{h}(t) \right) dt = 0$$

Par intégration par parties il vient

$$\int_{t_1}^{t_2} \left(\frac{\partial L}{\partial u}(u^*, \dot{u}^*, t) - \frac{d}{dt} \frac{\partial L}{\partial \dot{u}}(u^*, \dot{u}^*, t) \right) \dot{h}(t) dt + \left[\frac{\partial L}{\partial \dot{u}}(u^*, \dot{u}^*, t) h(t) \right]_{t_1}^{t_2} = 0$$

h est une variation admissible donc $h(t_1) = h(t_2) = 0$. La précédente équation se simplifie

$$\int_{t_1}^{t_2} \left(\frac{\partial L}{\partial u}(u^*, \dot{u}^*, t) - \frac{d}{dt} \frac{\partial L}{\partial \dot{u}}(u^*, \dot{u}^*, t) \right) \dot{h}(t) dt = 0$$

Cette équation doit être vraie pour toute variation admissible h , et le lemme 2 de duBois-Reymond permet de conclure

$$(28) \quad \frac{\partial L}{\partial u}(u^*, \dot{u}^*, t) - \frac{d}{dt} \frac{\partial L}{\partial \dot{u}}(u^*, \dot{u}^*, t) = 0$$

L'équation (28) est appelée équation d'Euler-Lagrange et est une condition nécessaire qui doit être satisfaite par toute solution du problème (27).

Dans de nombreux cas pratique, le Lagrangien ne dépend pas explicitement de t . On pourra alors utiliser la proposition suivante.

Proposition 14. — *Lorsque L ne dépend pas explicitement de t , l'équation d'Euler-Lagrange (28) implique*

$$\frac{d}{dt} \left(L - \dot{u} \frac{\partial L}{\partial \dot{u}} \right) = 0$$

Démonstration. — On calcule

$$\frac{d}{dt} \left(L - \dot{u} \frac{\partial L}{\partial \dot{u}} \right) = \dot{u} \left(\frac{\partial L}{\partial u} - \frac{d}{dt} \frac{\partial L}{\partial \dot{u}} \right) = 0$$

en utilisant (28). □

On peut étendre le calcul des variations aux fonctionnelles du type

$$X \ni u \mapsto J(u) = \int_{t_1}^{t_2} L(u(t), \dot{u}(t), \dots, u^{(n)}(t), t) dt \in \mathbb{R}$$

(en utilisant un espace vectoriel normé X adapté) et on obtient l'équation d'Euler-Poisson

$$\frac{\partial L}{\partial u} - \frac{d}{dt} \frac{\partial L}{\partial \dot{u}} + \frac{d^2}{dt^2} \frac{\partial L}{\partial \ddot{u}} + \dots + (-1)^n \frac{d^n}{dt^n} \frac{\partial L}{\partial u^{(n)}} = 0$$

On peut également appliquer les mêmes règles de calcul ainsi que la formule de Green pour les fonctionnelles

$$u \mapsto J(u) = \int_{t_1}^{t_2} L \left(u(x, y), \frac{\partial}{\partial x} u(x, y), \frac{\partial}{\partial y} u(x, y), x, y \right) dx dy \in \mathbb{R}$$

pour obtenir l'équation d'Ostrogradski

$$\frac{\partial L}{\partial u} - \frac{\partial}{\partial x} \left(\frac{\partial L}{\partial \frac{\partial}{\partial x} u} u(x, y) \right) - \frac{\partial}{\partial y} \left(\frac{\partial L}{\partial \frac{\partial}{\partial y} u} u(x, y) \right) = 0$$

2.2. Optimisation de systèmes dynamiques

On s'intéresse désormais au problème de la minimisation d'une fonctionnelle $X \times U \ni (x, u) \mapsto J(x, u) \in \mathbb{R}$ où $X \times U$ est un espace vectoriel normé. Ici, X est l'espace des fonctions continues à dérivées continues $\mathbb{R} \supset [t_1, t_2] \rightarrow \mathbb{R}^n$, $n < \infty$ muni de la norme $X \ni x \mapsto \|x\|_D = \max_{t \in [t_1, t_2]} \|x(t)\| + \max_{t \in [t_1, t_2]} \|\dot{x}(t)\|$. U est l'espace des fonctions définies sur $[t_1, t_2] \rightarrow \mathbb{R}^m$. Les fonctions x et u sont contraintes par une équation différentielle

$$\dot{x} = f(x, u)$$

où $\mathbb{R}^{n+m} \ni (x, u) \mapsto f(x, u) \in \mathbb{R}^n$ est une fonction de classe \mathcal{C}^1 . En outre on impose la contrainte

$$(29) \quad x(t_1) = x^0 \in \mathbb{R}^n$$

On va chercher une caractérisation des solutions du problème

$$(30) \quad \begin{aligned} \min \quad & J(x, u) \\ \dot{x} = & f(x, u) \\ x(t_1) = & x^0 \end{aligned}$$

On considérera des fonctionnelles du type

$$X \times U \ni (x, u) \mapsto J(x, u) = \varphi(x(t_2), t_2) + \int_{t_1}^{t_2} L(x(t), u(t), t) dt$$

où $\mathbb{R}^n \times \mathbb{R} \ni (x, t) \mapsto \varphi(x, t) \in \mathbb{R}$ est de classe \mathcal{C}^1 et $\mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R} \ni (x, u, t) \mapsto L(x, u, t) \in \mathbb{R}$ est de classe \mathcal{C}^1 . Pour tenir compte de la condition initiale (29), on considère qu'une variation admissible $(\delta x, \delta u) \in X \times U$ doit vérifier $\delta x(t_1) = 0$. Comme dans le cas de l'optimisation continue de dimension finie traité au chapitre 1, on va chercher à résoudre le problème d'optimisation sous contrainte en adjoignant les contraintes à la fonction coût.

On forme alors $X \times U \times M \ni (x, u, \lambda) \mapsto \bar{J}(x, u, \lambda)$ où M est l'espace des fonctions $\mathbb{R} \rightarrow \mathbb{R}^n$ différentiable par

$$\bar{J}(x, u, \lambda) = J(x, u) + \int_{t_1}^{t_2} \lambda(t)^T (f(x, u, t) - \dot{x})(t) dt$$

En notant

$$(31) \quad H(x(t), u(t), \lambda(t), t) = L(x(t), u(t), t) + \lambda(t)^T f(x, u, t)$$

qu'on appellera Hamiltonien du problème (30), il vient

$$(32) \quad \begin{aligned} \bar{J}(x, u, \lambda) = & \varphi(x(t_2), t_2) - \lambda(t_1)^T x(t_1) + \lambda(t_2)^T x(t_2) \\ & + \int_{t_1}^{t_2} \left(H(x(t), u(t), \lambda(t), t) + \dot{\lambda}(t)^T x(t) \right) dt \end{aligned}$$

Proposition 15. — Si $(x^*, u^*, \lambda^*) \in X \times U \times M$ est un point stationnaire de \bar{J} défini par (32) alors (x^*, u^*) est un point stationnaire de J sous les contraintes $\dot{x} = f(x, u, t)$, $x(t_1) = x^0$.

Démonstration. — Nous allons montrer que, autour de (x^*, u^*) , pour toute variation admissible du premier ordre $[t_1, t_2] \ni (\delta x, \delta u)(t) \in \mathbb{R}^n \times \mathbb{R}^m$ telle que $\|(\delta x, \delta u)\|_{X \times U} = o(\delta)$ satisfaisant la contrainte $\dot{x} = f(x, u, t)$, la variation δJ de la fonctionnelle J est du deuxième ordre, c.-à-d. $\delta J = o(\delta)$.

La stationnarité de (x^*, u^*, λ^*) pour \bar{J} est caractérisée par 3 relations. La première est

$$\begin{aligned} & \frac{\partial}{\partial x(t_2)} \varphi(x(t_2), t_2) \delta x(t_2) - \lambda(t_2)^T \delta x(t_2) \\ & + \int_{t_1}^{t_2} \left(\frac{\partial}{\partial x} L(x, u, t) \delta x + \lambda(t)^T \frac{\partial}{\partial x} f(x, u, t) \delta x + \dot{\lambda}(t)^T \delta x \right) dt = o(\delta) \end{aligned}$$

pour toute variation admissible $\mathbb{R} \ni t \mapsto \delta x(t) \in \mathbb{R}^n$. Ce calcul des variations donne

$$(33) \quad \frac{\partial}{\partial x} L(x, u, t) + \lambda(t)^T \frac{\partial}{\partial x} f(x, u, t) + \dot{\lambda}(t)^T = 0$$

$$(34) \quad \lambda^T(t_2) = \frac{\partial}{\partial x(t_2)} \varphi(x(t_2), t_2)$$

On réécrit la condition (33) sous la forme $\dot{\lambda}^T = -\frac{\partial H}{\partial x}$. La seconde condition de stationnarité est que pour toute variation $\mathbb{R} \ni t \mapsto \delta u(t) \in \mathbb{R}^m$ on doit avoir

$$\int_{t_1}^{t_2} \left(\frac{\partial}{\partial u} L(x, u, t) + \lambda(t)^T \frac{\partial}{\partial u} f(x, u, t) \right) \delta u(t) dt = o(\delta)$$

Ce calcul des variations donne

$$(35) \quad \frac{\partial}{\partial u} L(x, u, t) + \lambda(t)^T \frac{\partial}{\partial u} f(x, u, t) = o(\delta)$$

On réécrit cette condition comme $\frac{\partial}{\partial u} H = 0$.

La dernière condition de stationnarité est

$$\int_{t_1}^{t_2} \delta \lambda(t)^T (f(x, u, t) - \dot{x}) dt = 0$$

Elle redonne la contrainte

$$(36) \quad \dot{x} = f(x, u, t)$$

□

Supposons que les équations (33), (34), (35) et (36) soient vérifiées. Calculons la variation de la fonctionnelle J .

$$\begin{aligned}
\delta J &= \frac{\partial}{\partial x(t_2)} \varphi(x(t_2), t_2) \delta x(t_2) + \int_{t_1}^{t_2} \left(\frac{\partial}{\partial x} L(x, u, t) \delta x + \frac{\partial}{\partial u} L(x, u, t) \delta u \right) dt \\
&= \frac{\partial}{\partial x(t_2)} \varphi(x(t_2), t_2) \delta x(t_2) \\
&\quad + \int_{t_1}^{t_2} \left(-\lambda(t)^T \frac{\partial}{\partial x} f(x, u, t) \delta x(t) - \dot{\lambda}(t)^T \delta x(t) - \lambda(t)^T \frac{\partial}{\partial u} f(x, u, t) \delta u(t) \right) dt \\
&= \frac{\partial}{\partial x(t_2)} \varphi(x(t_2), t_2) \delta x(t_2) - \lambda(t_2)^T \delta x(t_2) + \lambda(t_1)^T \delta x(t_1) \\
&\quad + \int_{t_1}^{t_2} \lambda(t)^T \left(\dot{x}(t) - \frac{\partial}{\partial x} f(x, u, t) \delta x(t) - \frac{\partial}{\partial u} f(x, u, t) \delta u(t) \right) dt \\
&= o(\delta)
\end{aligned}$$

2.2.1. Problème aux deux bouts. — Les conditions de stationnarité (33) (34) de la proposition 15 forment, avec les contraintes (36) et (29), le problème “aux deux bouts” suivant

$$(37) \quad \left\{ \begin{array}{l} \dot{x} = f(x, u) \\ x(t_1) = x^0 \\ \dot{\lambda}^T = -\frac{\partial H}{\partial x}(x, u, \lambda) \\ \lambda^T(t_2) = \frac{\partial}{\partial x(t_2)} \varphi(x(t_2), t_2) \\ \text{où } u \text{ est solution de } \frac{\partial H}{\partial u} = 0 \end{array} \right.$$

Le long de l'extrémale on élimine u des équations en résolvant les équations $\frac{\partial}{\partial u} H = 0$. Le problème aux deux bouts a pour seules inconnues les fonctions $\mathbb{R} \supset [t_1, t_2] \ni t \mapsto x(t) \in \mathbb{R}^n$ et $\mathbb{R} \supset [t_1, t_2] \ni t \mapsto \lambda(t) \in \mathbb{R}^n$. Le système d'équations différentielles et de conditions limites (37) ne définit pas un problème de Cauchy à cause de la séparation des conditions de bords aux deux extrémités du domaine. En outre, la condition portant sur λ dépend de l'inconnue x . C'est un problème difficile à résoudre en général.

Proposition 16 (Conservation de l'Hamiltonien). — *Lorsque L ne dépend pas explicitement de t , les conditions de stationnarité constituant le système d'équations différentielles aux deux bouts (37) impliquent*

$$\frac{d}{dt} H = 0$$

Démonstration. — Un calcul direct donne

$$\frac{d}{dt} H = \frac{\partial H}{\partial x} \dot{x} + 0 + \frac{\partial H}{\partial \lambda} \dot{\lambda} = \frac{\partial H}{\partial x} f(x, u) - \frac{\partial H}{\partial x} f(x, u) = 0$$

□

En pratique on se servira souvent de la conservation de l'Hamiltonien pour éliminer une variable et essayer de résoudre le problème aux deux bouts.

2.2.2. Contraintes finales. — On cherche ici une caractérisation des solutions du problème

$$(38) \quad \begin{aligned} & \min J(x, u) \\ & \dot{x} = f(x, u) \\ & x(t_1) = x^0 \\ & \psi(x(t_2), t_2) = 0 \end{aligned}$$

où $\mathbb{R}^n \times \mathbb{R} \ni (x, t) \mapsto \psi(x, t) \in \mathbb{R}^q$, $1 < q \leq n$ est une fonction de classe \mathcal{C}^1 . Ces contraintes portent sur l'état final et le temps final du système qui est ici fixe, on les nomme contraintes de "rendez-vous".

On considérera des fonctionnelles du type

$$X \times U \ni (x, u) \mapsto J(x, u) = \varphi(x(t_2), t_2) + \int_{t_1}^{t_2} L(x(t), u(t), t) dt$$

où $\mathbb{R}^n \times \mathbb{R} \ni (x, t) \mapsto \varphi(x, t) \in \mathbb{R}$ est de classe \mathcal{C}^1 et $\mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R} \ni (x, u, t) \mapsto L(x, u, t) \in \mathbb{R}$ est de classe \mathcal{C}^1 .

Le problème aux deux bouts correspondant à ce problème d'optimisation est

$$(39) \quad \left\{ \begin{aligned} & \dot{x} = f(x, u) \\ & x(t_1) = x^0 \\ & \dot{\lambda}^T = -\frac{\partial H}{\partial x}(x, u, \lambda) \\ & \lambda^T(t_2) = \frac{\partial}{\partial x(t_2)} \varphi(x(t_2), t_2) + \nu^T \frac{\partial \psi}{\partial x(t_2)} \\ & \psi(x(t_2), t_2) = 0 \\ & \text{où } u \text{ est solution de } \frac{\partial H}{\partial u} = 0 \end{aligned} \right.$$

où $H(x, u, \lambda, t) = L(x, u, t) + \lambda^T f(x, u, t)$ et $\nu \in \mathbb{R}^q$. On obtient ce résultat en adjoignant les contraintes $\dot{x} = f(x, u, t)$ et $\psi(x(t_2), t_2) = 0$ à la fonctionnelle et en explicitant le calcul des variations. On a ainsi

$$(40) \quad \bar{J}(x, u, \lambda, \nu) = J(x, u) + \int_{t_1}^{t_2} \lambda(t)^T (f(x, u, t) - \dot{x})(t) dt + \nu^T \psi(x(t_2), t_2)$$

On peut alors établir la proposition suivante

Proposition 17. — Si $(x^*, u^*, \lambda^*, \nu^*) \in X \times U \times M \times \mathbb{R}^q$ est un point stationnaire de \bar{J} défini par (40) alors (x^*, u^*) est un point stationnaire de J sous les contraintes $\dot{x} = f(x, u, t)$, $x(t_1) = x^0$, $\psi(x(t_2), t_2) = 0$.

2.2.3. Résolution numérique du problème aux deux bouts. —

2.2.3.1. Calcul du gradient par l'adjoint. — On va chercher à étendre les calculs menés en dimension finie établissant la variation de la fonction coût par rapport aux inconnues en satisfaisant les contraintes (18). Dans le cas présent, on revient au problème

$$\begin{aligned} \min \quad & J(x, u) \\ \dot{x} = & f(x, u) \\ x(t_1) = & x^0 \end{aligned}$$

On considérera des fonctionnelles du type

$$X \times U \ni (x, u) \mapsto J(x, u) = \varphi(x(t_2), t_2) + \int_{t_1}^{t_2} L(x(t), u(t), t) dt$$

où $\mathbb{R}^n \times \mathbb{R} \ni (x, t) \mapsto \varphi(x, t) \in \mathbb{R}$ est de classe \mathcal{C}^1 et $\mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R} \ni (x, u, t) \mapsto L(x, u, t) \in \mathbb{R}$ est de classe \mathcal{C}^1 .

On va chercher à calculer la variation de J lorsqu'on fait varier l'inconnue u tout en maintenant la contrainte $\dot{x} = f(x, u)$, $x(t_1) = x^0$. On va noter δJ la variation de la valeur de J engendrée par une telle variation δu . Il vient

$$(41) \quad \delta J = \frac{\partial \varphi}{\partial x}(x(t_2), t_2) \delta x(t_2) + \int_{t_1}^{t_2} \left(\frac{\partial L}{\partial x}(x(t), u(t), t) \delta x(t) + \frac{\partial L}{\partial u} \delta u(t) \right) dt + o(\delta)$$

Dans cette équation δx est liée à δu par la satisfaction de la contrainte $\dot{x} = f(x, u)$ qui donne

$$\dot{\delta x}(t) = \frac{\partial f}{\partial x}(x, u, t) \delta x(t) + \frac{\partial f}{\partial u}(x, u, t) \delta u(t) + o(\delta)$$

Il est possible d'intégrer cette équation linéaire en utilisant la matrice de passage M définie par $\frac{d}{dt} M(t, t_1) = \frac{\partial f}{\partial x}(x, u, t) M(t, t_1)$, $M(t_1, t_1) = I$ ce qui donne alors

$$\delta x(t) = M(t, t_1) \delta x(t_1) + \int_{t_1}^t M(t, \tau) \frac{\partial f}{\partial u}(x, u, \tau) \delta u(\tau) d\tau + o(\delta)$$

On peut donc éliminer δx dans l'équation (41) et obtenir explicitement δJ en fonction de la variation $\mathbb{R} \supset [t_1, t_2] \ni t \mapsto u(t) \in \mathbb{R}^m$. Cette expression est très compliquée à calculer car elle nécessite la résolution complète d'un système d'équations différentielles. En fait, le lemme suivant va nous permettre de simplifier énormément les calculs

Lemme 3 (Lemme de l'adjoint). — *Les solutions du système d'équations différentielles*

$$\begin{cases} \dot{x}(t) = A(t)x(t) + B(t)u(t) \\ \dot{\lambda}(t) = -A^T(t)\lambda(t) + \Gamma(t) \end{cases}$$

où pour tout $t \in [t_1, t_2] \subset \mathbb{R}$, on a $A(t) \in \mathcal{M}_n(\mathbb{R})$, $B(t) \in \mathcal{M}_{(n \times m)}(\mathbb{R})$, $\Gamma(t) \in \mathbb{R}^n$, satisfait l'égalité

$$\lambda^T(t_2)x(t_2) - \lambda^T(t_1)x(t_1) = \int_{t_1}^{t_2} (\lambda^T(t)B(t)u(t) + \Gamma^T(t)x(t)) dt$$

Le système d'équations que nous devons résoudre dans le problème aux deux bouts implique

$$\begin{aligned}\dot{x}(t) &= \frac{\partial f}{\partial x}(x, u, t)\delta x(t) + \frac{\partial f}{\partial u}(x, u, t)\delta u(t) + o(\delta) \\ \dot{\lambda}(t) &= - \left(\frac{\partial f}{\partial x}(x, u, t) \right)^T \lambda(t) - \left(\frac{\partial L}{\partial x}(x, u, t) \right)^T\end{aligned}$$

avec comme conditions limites $\delta x(t_1) = 0$, $\lambda(t_2) = \left(\frac{\partial}{\partial x(t_2)} \varphi(x(t_2), t_2) \right)^T$. Le lemme 3 de l'adjoint donne

$$\delta J = \int_{t_1}^{t_2} \left(\frac{\partial L}{\partial u}(x, u, t) + \lambda^T \frac{\partial f}{\partial u}(x, u, t) \right) \delta u(t) dt + o(\delta) = \int_{t_1}^{t_2} \frac{\partial H}{\partial u}(x, u, \lambda, t) \delta u(t) dt + o(\delta)$$

Par abus de notation on retiendra la formule analogue à (18)

$$(42) \quad \begin{aligned} \left(\frac{\partial J}{\partial u} \right) \dot{x} = f(x, u, t) &= \frac{\partial H}{\partial u} \\ x(t_1) &= x^0 \end{aligned}$$

Autrement dit, la variation de la valeur de J en satisfaisant les contraintes est calculée en formant $\frac{\partial H}{\partial u} = \frac{\partial L}{\partial u}(x, u, t) + \lambda^T \frac{\partial f}{\partial u}(x, u, t)$, évaluée à partir des valeurs de x , u , et de l'adjoint λ . La formule (42) est appelée formule du calcul du gradient par l'adjoint.

Algorithme 10 (Résolution du problème aux deux bouts par le calcul du gradient par l'adjoint)

- À partir de u^0 fonction $\mathbb{R} \supset [t_1, t_2] \mapsto u^0(t) \in \mathbb{R}^m$ quelconque, itérer
 - Résoudre $\dot{x}^k = f(x^k, u^k, t)$, $x^k(t_1) = x^0$ pour $t \in [t_1, t_2]$
 - Résoudre $\dot{\lambda}^k(t) = - \left(\frac{\partial f}{\partial x}(x^k, u^k, t) \right)^T \lambda(t) - \left(\frac{\partial L}{\partial x}(x, u, t) \right)^T$ avec comme condition limite $\lambda^k(t_2) = \left(\frac{\partial}{\partial x(t_2)} \varphi(x^k(t_2), t_2) \right)^T$
 - Calculer $\frac{\partial H^k}{\partial u} = \frac{\partial L}{\partial u}(x^k, u^k, t) + (\lambda^k)^T \frac{\partial f}{\partial u}(x^k, u^k, t)$
 - Mettre à jour $u^{k+1} = u^k - l^k \left(\frac{\partial H^k}{\partial u} \right)$ avec l^k satisfaisant les règles de Wolfe (2) et (3).

En pratique la fonction u^k sera représentée par un vecteur (de coefficients représentant une décomposition dans une base de fonctions) de dimension finie et le problème sera ainsi résolu de manière approchée. Les étapes de résolution des équations différentielles satisfaites par x et λ (équation en temps rétrograde) seront effectuées de manière numérique (par exemple par des schémas de Runge-Kutta [24]) avec une précision finie. On peut adapter l'algorithme 10 pour tenir compte de contraintes finales. Il suffit d'introduire les équations du problème aux deux bouts correspondant et d'ajouter une étape de mise à jour des inconnues ν définies précédemment à la section 2.2.2.

Afin d'obtenir une vitesse de convergence supérieure à celle de l'algorithme 10, on peut utiliser l'algorithme suivant

Algorithme 11 (Algorithme de tir). — À partir de $\lambda^0(t_1) \in \mathbb{R}$ quelconque, itérer

- Calculer formellement u^{k+1} solution de $\frac{\partial H}{\partial u}(x^{k+1}, u^{k+1}, \lambda^{k+1}, t) = 0$ en fonction de $x^{k+1}, \lambda^{k+1}, t$.
- Résoudre le système

$$(43) \quad \begin{cases} \dot{x}^{k+1} = f(x^{k+1}, u^{k+1}, t) \\ \dot{\lambda}^{k+1}(t) = -\frac{\partial f}{\partial x}(x^{k+1}, u^{k+1}, t)\lambda^{k+1}(t) - \left(\frac{\partial L}{\partial x}(x^{k+1}, u^{k+1}, t)\right)^T \end{cases}$$

pour $t \in [t_1, t_2]$ avec comme condition initiale $x^k(t_1) = x^0, \lambda^k(t_1) = \lambda^k(t_1)$.

- Calculer la fonction de sensibilité $F \in \mathcal{M}_n(\mathbb{R})$ définie comme $F = \frac{\partial \lambda(t_2)}{\partial \lambda(t_1)}$
- Mettre à jour $\lambda^{k+1}(t_1) = \lambda^k(t_1) - l^k F^{-1} \left(\lambda(t_2) - \left(\frac{\partial}{\partial x(t_2)} \varphi(x(t_2), t_2) \right)^T \right)$ avec l^k satisfaisant les règles de Wolfe (2) et (3).

Cet algorithme a pour inconnue la seule valeur de l'état adjoint λ à l'instant initial. La commande u est calculée à chaque itération. Une valeur $\lambda(t_1)$ satisfait les conditions de stationnarité si elle fournit par intégration une valeur $\lambda(t_2) = \left(\frac{\partial}{\partial x(t_2)} \varphi(x(t_2), t_2) \right)^T$. Si ce n'est pas le cas, l'algorithme modifie la valeur candidate de l'inconnue pour tenter de satisfaire cette contrainte. Par rapport à l'algorithme 10, on constate en général une convergence plus rapide en pratique lorsque la valeur $\lambda^0(t_1) \in \mathbb{R}$ est proche de la valeur optimale. Si ce n'est pas le cas, l'algorithme 11 souffre souvent de l'instabilité numérique de la résolution simultanée des équations différentielles (43) satisfaites par x et λ . Cette instabilité est liée à l'instabilité du système linéarisé tangent. L'algorithme 11 peut être amélioré de nombreuses façons, on pourra se reporter à [5, 25] pour un exposé sur les méthodes de tirs multiples.

2.2.4. Principe du minimum. — Les calculs des variations menés jusqu'à présent ont consisté à calculer la variation de la valeur d'une fonctionnelle de deux variables x, u lorsque u était libre et x était liée à u par une équation différentielle $\dot{x} = f(x, u, t)$. Maintenant nous allons considérer que la variable u n'est pas libre mais contrainte. Nous regardons maintenant les problèmes de la forme

$$(44) \quad \begin{aligned} \min \quad & J(x, u) \\ \dot{x} = & f(x, u) \\ x(t_1) = & x^0 \\ C(u, t) \leq & 0 \end{aligned}$$

où $\mathbb{R}^m \times \mathbb{R} \ni (u, t) \mapsto C(u, t) \in \mathbb{R}^l, l \in \mathbb{N}$ est une fonction de classe \mathcal{C}^1 . On considérera des fonctionnelles du type

$$X \times U \ni (x, u) \mapsto J(x, u) = \varphi(x(t_2), t_2) + \int_{t_1}^{t_2} L(x(t), u(t), t) dt$$

où $\mathbb{R}^n \times \mathbb{R} \ni (x, t) \mapsto \varphi(x, t) \in \mathbb{R}$ est de classe \mathcal{C}^1 et $\mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R} \ni (x, u, t) \mapsto L(x, u, t) \in \mathbb{R}$ est de classe \mathcal{C}^1 .

La preuve du résultat que nous allons énoncer est difficile. Nous nous contentons d'en suggérer quelques grandes lignes. Comme nous l'avons vu, la variation de la valeur de la fonctionnelle s'écrit en fonction d'une variation δu ,

$$(45) \quad \delta J = \int_{t_1}^{t_2} \frac{\partial H}{\partial u} \delta u(t) dt$$

Autour d'un optimum (x^*, u^*) , il ne doit pas exister de variation admissible, compatible avec les contraintes $C(u, t) \leq 0$ fournissant une variation négative δJ . En reprenant le point de vue utilisé dans la présentation des conditions de Karush, Kuhn et Tucker du théorème 17, on doit avoir que le gradient de chaque contribution ponctuelle dans l'intégrale (45) doit être dans le cône convexe des contraintes actives. Autrement dit, pour tout $t \in [t_1, t_2]$, il existe un vecteur $\mathbb{R} \supset [t_1, t_2] \ni t \mapsto \mu(t) \in \mathbb{R}^l$ dont les composantes vérifient $\mu_i(t) = 0$ si $C_i(x(t), u(t)) < 0$, $\mu_i(t) \geq 0$ si $C_i(x(t), u(t)) = 0$ tel que

$$\frac{\partial H}{\partial u} = -\mu^T \frac{\partial C}{\partial u}$$

Plus formellement le principe du minimum s'énonce

Théorème 21 (Principe du minimum de Pontryagin [23])

Soient $\mathbb{R} \supset [t_1, t_2] \ni t \mapsto (x, u, \lambda)(t) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n$ optimum du problème (44). Pour tout $t \in [t_1, t_2]$, $u(t)$ réalise le minimum de $\mathbb{R}^m \ni u \mapsto H(x(t), u, \lambda(t), t) \in \mathbb{R}$ sous la contrainte $C(u, t) \leq 0$.

2.3. Champs d'extrémales

Il est possible d'exhiber une structure reliant les trajectoires optimales de problèmes d'optimisation voisins. Le but de cette section est de présenter cette structure au travers de l'équation aux dérivées partielles de Hamilton-Jacobi-Bellman.

On considère une fois de plus le problème suivant (d'autres généralisations sont possibles)

$$(46) \quad \begin{aligned} \min \quad & J(x, u) \\ \dot{x} = & f(x, u) \\ x(t_1) = & x^0 \\ \psi(x(t_2), t_2) = & \end{aligned}$$

où $\mathbb{R}^n \times \mathbb{R} \ni (x, t) \mapsto \psi(x, t) \in \mathbb{R}^q$, $1 < q \leq n$ est une fonction de classe \mathcal{C}^1 . On considérera des fonctionnelles du type

$$X \times U \ni (x, u) \mapsto J(x, u) = \varphi(x(t_2), t_2) + \int_{t_1}^{t_2} L(x(t), u(t), t) dt$$

où $\mathbb{R}^n \times \mathbb{R} \ni (x, t) \mapsto \varphi(x, t) \in \mathbb{R}$ est de classe \mathcal{C}^1 et $\mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R} \ni (x, u, t) \mapsto L(x, u, t) \in \mathbb{R}$ est de classe \mathcal{C}^1 .

Définition 25 (Extrémale). — Une extrémale est l'ensemble de \mathbb{R}^n des points $(x(t), t)$, $t \in [t_1, t_2]$ où $\mathbb{R} \supset [t_1, t_2] \ni t \mapsto x(t) \in \mathbb{R}^n$ est solution du problème d'optimisation (46).

Définition 26. — Soit \mathcal{E} une famille d'extrémales obtenue en faisant varier x^0, t_1 . Soit U un ensemble compact de $\mathbb{R}^n \times \mathbb{R}$. On dit que \mathcal{E} est un champs d'extrémales pour U si $\mathcal{E} \supset U$.

En conséquence, il existe une extrémale passant par tous les points de U .

Définition 27 (Fonction de retour optimal). — On appelle fonction de retour optimal $\mathbb{R}^n \times \mathbb{R} \ni (x^0, t_1) \mapsto \mathcal{J}(x^0, t_1)$ (à l'ensemble $\{(x, t) \in \mathbb{R}^n \times \mathbb{R} \text{ t.q. } \psi(x, t) = 0\}$), la fonction ayant pour valeur la valeur optimale pour le problème d'optimisation (46) avec pour condition initiale (x^0, t_1) .

Les extrémales vérifient la propriété suivante.

Proposition 18 (Principe d'optimalité de Bellman). — *Une suite de décisions est optimale si, quel que soit l'état et l'instant considérés sur la trajectoire qui lui est associée, les décisions ultérieures constituent une suite optimale de décisions pour le sous problème dynamique ayant cet état et cet instant comme conditions initiales.*

Grâce à cette propriété, il est possible d'établir une équation aux dérivées partielles satisfaites par \mathcal{J} . On suppose disposer d'un champs d'extrémales \mathcal{E} pour le problème d'optimisation (46). Considérons $(x, t) \in \mathcal{E}$. Il correspond à la fonction de retour optimal évaluée en ce point $J^0(x, t)$ une commande optimale $\mathbb{R} \supset [t, t_2] \ni l \mapsto u^0(l) \in \mathbb{R}^m$ que nous supposons (pour simplifier) unique. Pour toute commande u proche de u^0 on peut calculer une fonction coût à partir de la trajectoire $\mathbb{R} \supset [t_1, t_2] \ni t \mapsto x(t) \in \mathbb{R}^n$ (proche de la trajectoire optimale) issue de $x(t_1) = x^0$. Il vient

$$J(x, u) \geq \mathcal{J}(x^0, t_1)$$

avec

$$(47) \quad \min_u J(x, u) = \mathcal{J}(x^0, t_1)$$

Choisissons comme candidat d'appliquer u une commande différente de la commande optimale u^0 pendant δt puis optimale sur $[t_1 + \delta t, t_2]$. On peut évaluer

$$J(x, u) = \int_{t_1}^{t_1 + \delta t} L(x(t), u(t), t) dt + \mathcal{J}(x(t_1 + \delta t), t_1 + \delta t)$$

Un développement au premier ordre donne

$$\begin{aligned} J(x, u) = & L(x^0, u(t_1), t_1) \delta t + \mathcal{J}(x^0, t_1) + \frac{\partial \mathcal{J}}{\partial x}(x^0, t_1) f(x^0, u(t_1), t_1) \delta t \\ & + \frac{\partial \mathcal{J}}{\partial t}(x^0, t_1) \delta t + o(\delta t) \end{aligned}$$

L'équation (47) donne

$$\mathcal{J}(x^0, t_1) = \mathcal{J}(x^0, t_1) + o(\delta t) + \left(\min_u \left(L(x^0, u(t_1), t_1) + \frac{\partial \mathcal{J}}{\partial x}(x^0, t_1) f(x^0, u(t_1), t_1) \right) + \frac{\partial \mathcal{J}}{\partial t}(x^0, t_1) \right) \delta t$$

Un passage à la limite lorsque $\delta t \rightarrow 0$ donne

$$-\frac{\partial \mathcal{J}}{\partial t} = \min_u \left(L(x, u, t) + \frac{\partial \mathcal{J}}{\partial x} f(x, u, t) \right)$$

On reconnaît ici l'Hamiltonien (31) et on écrit finalement l'équation de Hamilton-Jacobi-Bellman

$$(48) \quad -\frac{\partial \mathcal{J}}{\partial t} = \min_u \left(H(x, u, \frac{\partial \mathcal{J}}{\partial x}, t) \right)$$

Il s'agit d'une équation aux dérivées partielles hyperbolique dont la condition limite est donnée sur l'ensemble $\mathcal{V} = \{(x, t) \in \mathbb{R}^n \times \mathbb{R} \text{ t.q. } \psi(x, t) = 0\}$ par $\mathbb{R}^n \times \mathbb{R} \ni (x, t) \mapsto \mathcal{J}(x, t) = \varphi(x, t)$. On la résoudra en général par une propagation (rétrograde) depuis l'ensemble \mathcal{V} . Cette résolution sera par exemple menée le long d'un réseau quadrillant le plan (lorsque x est à valeur dans \mathbb{R}^2) par la méthode de la programmation dynamique.

CHAPITRE 3

EXERCICES 1

3.1. Fonction de Rosenbrock

Calculer le gradient $\nabla f(x)$ et le Hessien $\nabla^2 f(x)$ de la fonction de Rosenbrock

$$f(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$$

Montrer que $x^* = (1, 1)^T$ est l'unique minimum local de cette fonction et qu'à ce point le Hessien est défini positif.

3.2. Gradient et Hessien

Soit a un vecteur de \mathbb{R}^n et A une matrice $n \times n$ symétrique. Calculer le gradient et le Hessien de $f_1(x) = a^T x$ et $f_2(x) = x^T A x$.

3.3. Ensemble minimisant

Soit f une fonction convexe. Montrer que l'ensemble des minimiseurs globaux de f est convexe.

3.4. Vitesse de convergence

Soit la suite $(x_k)_{k \in \mathbb{N}}$ définie par

$$x_k = \begin{cases} \left(\frac{1}{4}\right)^{2^k}, & \text{si } k \text{ pair} \\ (x_{k-1})/k, & \text{si } k \text{ impair} \end{cases}$$

Cette suite est-elle convergente de type Q-superlinéaire? Est-elle Q-quadratique ou R-quadratique.

Note : on pourra utiliser la suite majorante $\nu_k = (\frac{1}{2})^{2^k}$.

3.5. Conditions de Wolfe

Avec les notations du cours, montrer en utilisant la fonction $f(x) = x^2$ et $x_0 = 1$ que si $0 < c_2 < c_1 < 1$ il peut ne pas exister de pas satisfaisant les conditions de Wolfe.

3.6. Suite de Fibonacci

La méthode de dichotomie présentée en cours n'est pas optimale au sens où pour un nombre fixé d'évaluations de la fonction à minimiser elle n'aboutit pas à l'intervalle réduit de taille minimale. On propose ici une méthode optimale.

On note Δ_k la taille de l'intervalle de recherche à l'itération k , que l'on note $[a_k, d_k]$. On évalue la fonction à minimiser en deux points intermédiaires b_k et c_k ce qui nous donne donc 4 valeurs. En utilisant l'unimodalité on peut éliminer un des deux sous-intervalles $[a_k, b_k]$ ou $[c_k, d_k]$.

- Montrer que si on veut que la longueur de l'intervalle Δ_{k+1} soit indépendante des valeurs de la fonction à minimiser on doit choisir b_k et c_k disposés symétriquement par rapport au milieu de $[a_k, d_k]$.
- Montrer $\Delta_k = \Delta_{k+1} + \Delta_{k+2}$ si on veut n'avoir à recalculer qu'une seule valeur de la fonction à minimiser à chaque itération
- Soit Δ_{N-1} l'intervalle obtenu au bout de N calculs, définissons F_0, F_1, \dots, F_{N-1} par

$$\Delta_k = F_{N-k} \Delta_{N-1}$$

Montrer qu'on a alors

$$F_n = F_{n-1} + F_{n-2}$$

pour $n = 3, \dots, N-1$ et $F_1 = 1$.

- Montrer que

$$F_{n+2} = nF_2 + n - 1$$

et que $F_2 \leq 2$.

- Conclure alors que pour réduire le plus les intervalles en N calculs il faut choisir $F_2 = 2$
- La suite obtenue est la suite de Fibonacci. Pour l'utiliser on choisit d'abord un rapport de réduction pour l'intervalle final et on détermine ainsi N . Montrer que pour un rapport de réduction de 1000 il faut choisir $N = 16$. Montrer comment procéder aux itérations en commençant par

$$\Delta_1/\Delta_2 = F_{N-1}/F_{N-2}$$

3.7. Moindres carrés

On admet les deux résultats suivants :

1. si A matrice $p \times n$ est de rang n alors $A^T A$ est inversible
2. une matrice M est symétrique définie positive si et seulement si elle se factorise comme $M = H^T H$ avec H inversible.

On cherche à résoudre le problème dit des “moindres carrés”

$$\min_{x \in \mathbb{R}^n} \|Ax - b\|$$

où A est une matrice $p \times n$, $b \in \mathbb{R}^p$ et en général $p \geq n$.

- Montrer que ce problème est convexe
- On suppose que la norme considérée est la norme euclidienne $\|x\|^2 = x^T x$, calculer le gradient de la fonction à minimiser.
- Quel est alors un point candidat pour être minimum ? Calculer le Hessien en ce point, conclure.
- Reprendre l’expression du minimum lorsque le problème est “pondéré” c’est à dire $\|x\| = x^T Q x$ où Q est une matrice symétrique définie positive.

CHAPITRE 4

EXERCICES 2

4.1. Algorithme de Newton

On considère la fonction

$$\mathbb{R}^2 \ni (x_1, x_2) \mapsto f(x_1, x_2) = \exp(x_1 + x_2 - 2) + (x_1 - x_2)^2$$

Calculer la première itération de l'algorithme de Newton en partant de $(1, 1)^T$ (réponse : $(0.5, 0.5)^T$).

4.2. Algorithme du gradient conjugué

On considère la fonction quadratique

$$\mathbb{R}^2 \ni (x_1, x_2) \mapsto q(x_1, x_2) = x_1^2 + 2x_2^2 + x_1x_2 - x_1 - 2x_2$$

Calculer les deux itérations de l'algorithme du gradient conjugué en partant de $(0, 0)^T$ et montrer qu'on aboutit au minimum global $(2/7, 3/7)^T$.

4.3. Formule SR1

Parmi les méthodes de quasi-Newton, la méthode SR1 est souvent retenue pour sa simplicité et la bonne approximation du Hessien qu'elle procure. Avec les notations du cours elle donne

$$B_{k+1} = B_k + \frac{(y_k - B_k s_k)(y_k - B_k s_k)^T}{(y_k - B_k s_k)^T s_k}$$

Utiliser la formule de Sherman-Morrison-Woodbury pour montrer que la formule de mise à jour de $H_k = B_k^{-1}$ est

$$H_{k+1} = H_k + \frac{(s_k - H_k y_k)(s_k - H_k y_k)^T}{(s_k - H_k y_k)^T y_k}$$

4.4. Invariance d'échelle

Montrer que la formule SR1 précédente, ainsi que la formule de la méthode BFGS sont invariantes par transformation $x = Sz + s$ où S est une matrice inversible et s un vecteur. Pour cela on pourra considérer la fonction $\tilde{f}(z) = f(Sz + s)$, calculer son gradient, son Hessien, et montrer que si on initialise l'approximation du Hessien par $S^T B_0 S$ et qu'on utilise un pas unitaire, on a $x_k = Sz_k + s$ à toutes les itérations.

4.5. Calcul automatique du gradient par parcours de graphe [21]

Soit la fonction

$$\mathbb{R}^2 \times \mathbb{R}^* \ni (x_1, x_2, x_3) \mapsto (x_1 x_2 \sin x_3 + \exp(x_1 x_2)) / x_3$$

En faisant intervenir les variables intermédiaires

$$\begin{aligned} x_4 &= x_1 x_2 \\ x_5 &= \sin x_3 \\ x_6 &= \exp x_4 \\ x_7 &= x_4 x_5 \\ x_8 &= x_6 + x_7 \\ x_9 &= x_8 / x_3 \end{aligned}$$

construire un graphe binaire permettant d'évaluer f au point $(x_1, x_2, x_3) = (1, 2, \pi/2)$.

En utilisant le fait qu'à chaque étape les opérations ne font intervenir que deux variables, montrer comment une propagation en avant puis en arrière des calculs permet de calculer de manière automatique et efficace le gradient de la fonction sans en calculer une expression littérale. C'est une méthode générale très utilisée en pratique car très efficace.

CHAPITRE 5

EXERCICES 3

5.1. Fonction de mérite

Une méthode simple à mettre en oeuvre pour résoudre un problème de minimisation sous contraintes égalités

$$\min_{c(x)=0} f(x)$$

consiste à chercher les solutions de

$$\nabla f(x) + \nabla c(x)^T \lambda = 0, \quad c(x) = 0$$

via la recherche du minimum de la fonction (dite de mérite)

$$M(x, \lambda) = \|\nabla f(x) + \nabla c(x)^T \lambda\|^2 + \|c(x)\|^2$$

En pratique cette méthode permet d'utiliser un code prévu pour la minimisation sans contraintes pour résoudre un problème avec contraintes.

Appliquer cette méthode à

$$\min_{x+y=1} \frac{1}{3}x^3 + \frac{1}{2}y^2$$

et montrer qu'en plus de l'unique minimum local on trouve un maximum local et un autre point parasite (c'est la principale faiblesse connue de cette méthode).

5.2. Dual d'une programmation quadratique

Considérer le problème de programmation quadratique

$$\min_{Ax \geq b} \frac{1}{2}x^T Gx + x^T d$$

où G est symétrique définie positive. Montrer que le problème dual est

$$\max_{Gx+d-A^T\lambda=0, \lambda \geq 0} \frac{1}{2}x^T Gx + x^T d - \lambda^T (Ax - b)$$

qui peut se simplifier en

$$\max_{\lambda \geq 0} -\frac{1}{2}\lambda^T(AG^{-1}A^T)\lambda + \lambda^T(b + AG^{-1}d) - \frac{1}{2}d^TG^{-1}d$$

5.3. Méthode de sous-ensembles actifs

Résoudre par la méthode de sous-ensembles actifs le problème de programmation quadratique $\min q(x)$, $q(x) = (x_1 - 1)^2 + (x_2 - 2.5)^2$ sous les contraintes

$$(49) \quad -x_1 + 2x_2 \leq 2$$

$$(50) \quad x_1 + 2x_2 \leq 6$$

$$(51) \quad x_1 - 2x_2 \leq 2$$

$$(52) \quad -x_1 \leq 0$$

$$(53) \quad -x_2 \leq 0$$

Utiliser comme point de départ $x^0 = (2, 0)^T$ et $W^0 = \{3, 5\}$.

5.4. Pénalisation

Considérer le problème pénalisé à barrière logarithmique

$$(54) \quad \min_{x \in]0,1[} x - \mu \log x - \mu \log(1 - x)$$

provenant du problème

$$(55) \quad \min_{x \geq 0, x \leq 1} x$$

Résoudre (54), vérifier analytiquement que lorsque μ tend vers 0, $x^*(\mu)$, l'optimum obtenu, converge vers la solution de (55).

5.5. Plus courte distance à un hyperplan

La plus courte distance entre un point x_0 et un hyperplan $\{x, Ax = b\}$ où les lignes de A sont linéairement indépendantes, peut s'écrire comme un problème de programmation quadratique

$$\min_{Ax=b} \frac{1}{2}(x - x_0)^T(x - x_0)$$

Montrer que le multiplicateur à l'optimum est

$$\lambda^* = -(AA^T)^{-1}(b - Ax_0)$$

et que la solution est

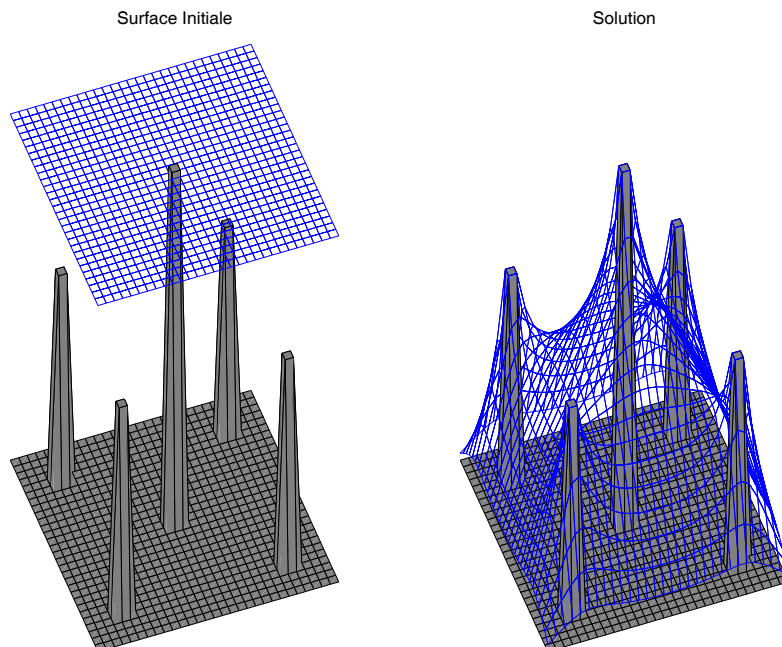
$$x^* = x_0 + A^T(AA^T)^{-1}(b - Ax_0)$$

Montrer que dans le cas où A est un vecteur ligne, la plus petite distance entre x_0 et l'hyperplan vaut

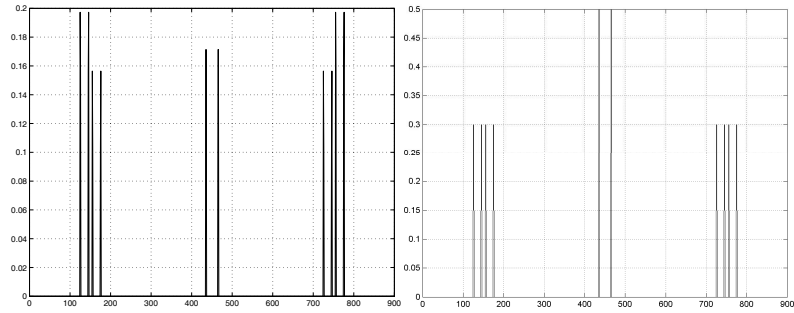
$$\frac{|b - Ax_0|}{\|A\|}$$

5.6. Tente de cirque

Ce problème est issu de l'Optimization Toolbox de Matlab [18]. On considère une tente de cirque comme une membrane élastique posée sur 5 mâts (le mât central étant plus haut que les autres). La forme prise par cette toile de tente minimise une énergie potentielle composée d'un terme de gravité auquel s'oppose un terme d'élasticité. Après avoir utilisé un schéma aux différences pour le gradient de la surface, on obtient un problème de programmation quadratique dont l'inconnue est l'altitude des différents points de la toile. Si on représente la surface de la toile par une grille de points, on désigne par x le vecteur constitué par les lignes de cette grilles mises bout à bout. Ainsi une grille de taille 30×30 donne un vecteur de taille 900. Les contraintes expriment que la toile repose sur les mâts.



Les multiplicateurs de Lagrange et les contraintes correspondantes sont représentés par



Avec les paramètres choisis, on obtient un minimum de 0.44422, de combien faut-il modifier la hauteur du mât central pour faire décroître cette valeur à 0.43767 ?

CHAPITRE 6

EXERCICES 4

6.1. Optimisation de plan de vol

Le pilote d'un avion dont la vitesse V est constante par rapport au vent (de vitesse u) cherche à maximiser l'aire contenue dans la surface fermée tracée au sol par sa trajectoire sur un intervalle $[0, T]$. Les équations de la dynamique sont

$$\begin{aligned}\dot{x} &= V \cos \theta + u, \\ \dot{y} &= V \sin \theta\end{aligned}$$

L'aire à maximiser est

$$A = \int_0^T y \dot{x} dt.$$

Montrer que la courbe recherchée est une ellipse dont on précisera l'excentricité.

6.2. Forme d'une chaîne pesante

Considérons une chaîne de longueur $2L$ suspendue entre deux points de même altitude situés à une distance $2l$ l'un de l'autre ($l \leq L$). Utiliser le calcul des variations pour trouver la forme qui minimise l'énergie potentielle (g la gravité est une constante).

6.3. Investissement optimal

Considérons un investisseur qui possède une somme S à la banque. N'ayant aucun autre revenu que les intérêts qui lui sont versés, il cherche quelle est la meilleure façon (au sens de son agrément) de dépenser son argent sur l'horizon $[0, T]$.

On suppose que son agrément est pour tout temps $E = 2\sqrt{r}$. L'agrément futur compte moins que le présent. La fonction qu'il va maximiser est donc

$$\int_0^T \exp(-\beta t) E(t) dt.$$

Dans le même temps, la banque lui verse des intérêts proportionnels à son capital restant à la banque $x(t)$. L'équation régissant ses économies est donc

$$\dot{x}(t) = \alpha x(t) - r(t).$$

1. Montrer que le problème se ramène à trouver un point stationnaire de

$$J(x) = \int_0^T \exp(-\beta t) 2\sqrt{\alpha x(t) - \dot{x}(t)} dt$$

avec $x(0) = S$, $x(T) = 0$.

2. Écrire l'équation de Euler-Lagrange equation correspondante et montrer que

$$\alpha x(t) - \dot{x}(t) = r(0) \exp(2(\alpha - \beta)t).$$

3. Supposer que $\alpha > \beta > \alpha/2$. Résoudre l'équation précédente. Montrer que la politique optimale de dépense est donnée par

$$x(t) = S \exp(2(\alpha - \beta)t) \frac{1 - \exp((\alpha - 2\beta)(T - t))}{1 - \exp((\alpha - 2\beta)T)}.$$

4. Représenter graphiquement cette fonction. Commenter.

CHAPITRE 7

EXERCICES ET PROBLÈMES COMPLÉMENTAIRES

7.1. Minimisation sous contraintes

En utilisant les conditions de Karush-Kuhn-Tucker, trouver le minimum de

$$J(y_1, y_2) = \left(\frac{y_1}{2} - a\right)^2 + (y_2 - b)^2$$

sous les contraintes

$$\begin{aligned}y_1^2 + y_2^2 &\leq 1 \\ y_1 &\geq 0 \\ y_2 &\geq 0.\end{aligned}$$

Traiter tous les cas possibles pour a et b .

Résoudre le même problème en rajoutant la contrainte $y_1 + y_2 \geq \frac{1}{2}$.

7.2. Loi bi-tangente

Considérons une particule de masse m évoluant dans un plan sur laquelle agit une force d'amplitude ma dont l'orientation (par rapport à l'axe x) est $\beta(t)$. Les équations de la dynamique sont

$$\begin{aligned}\dot{x} &= u \\ \dot{u} &= a \cos \beta \\ \dot{y} &= v \\ \dot{v} &= a \sin \beta\end{aligned}$$

On suppose vouloir minimiser une fonction coût ne comportant qu'un terme final du type $\varphi(x(t_f), y(t_f))$. Montrer qu'alors la commande optimale est du type

$$\tan \beta = \frac{a + bt}{c + dt}.$$

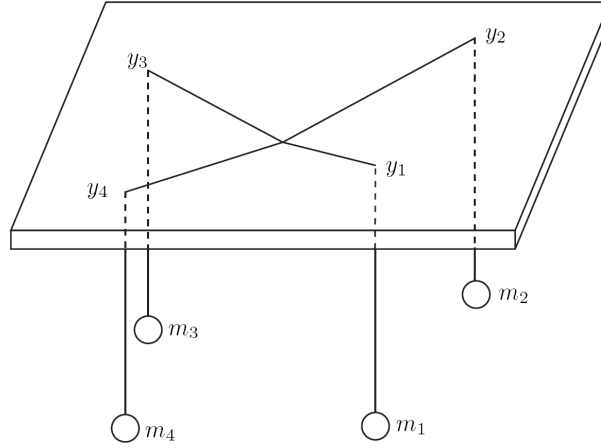


FIGURE 1. Construction de Varignon : planche percée de trous par lesquels passent des ficelles nouées en un point. On attache une masse à l'extrémité libre de chaque ficelle.

7.3. Problème de Weber

Étant donné un ensemble de points dans un plan $\{y_1, \dots, y_m\}$, on cherche le point dont la somme des distances pondérées est minimum. Le problème s'écrit

$$\min_{x \in \mathbb{R}^3} \sum_{i=1}^m m_i \|x - y_i\|$$

où les poids m_i sont des constantes positives.

1. Montrer qu'il existe un minimum global à ce problème et qu'on peut le construire par le schéma représenté Figure 1.
2. Ce minimum est-il unique ?

7.4. Une équation HJB

Considérons le coût

$$J = \int_{t_0}^{\infty} (x^2 + u^2) dt$$

sous la contrainte

$$\dot{x} = -x^4 + u$$

Écrire l'équation HJB associée.

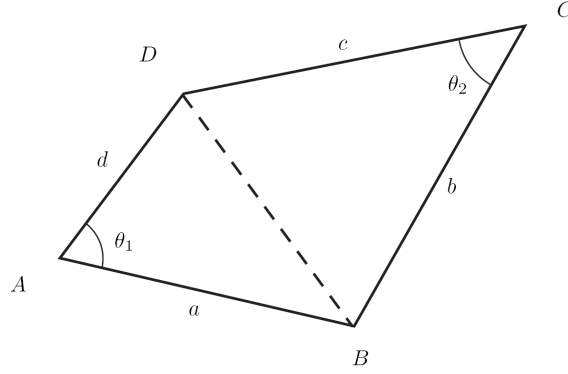


FIGURE 2. Quadrilatère.

7.5. Optimisation géométrique

Considérons le quadrilatère représenté sur la Figure 2 avec les notations telles que définies sur cette figure $((\theta_1, \theta_2) \in]0, \pi[^2, a > 0, b > 0, c > 0, d > 0)$. On cherche la configuration (θ_1, θ_2) fournissant l'aire maximale.

1. Montrer que le problème revient à trouver l'extremum de

$$(\theta_1, \theta_2) \mapsto ad \sin \theta_1 + bc \sin \theta_2$$

sous la contrainte

$$a^2 + d^2 - 2ad \cos \theta_1 = b^2 + c^2 - 2bc \cos \theta_2$$

2. Écrire le Lagrangien associé à ce problème et établir des conditions d'extrémalité.
3. Calculer les valeurs de (θ_1, θ_2) extremum.
4. Quelle figure obtient-on pour $a = b = c = d = 1$? Même question pour $a = d = 1$ et $b = c = 2 + \sqrt{3}$.

7.6. Régulateur LQR

On considère l'optimisation d'un système commandé $\dot{x} = F(t)x + G(t)u$ où F et G sont des matrices de taille $n \times n$. On cherche la commande minimisant le coût

$$J = \frac{1}{2}x^T(T)S_fx(T) + \frac{1}{2} \int_0^T (x^T(t)Ax(t) + u^T(t)Bu(t)) dt$$

où A, B, S_f sont des matrices symétriques définies positives, et T une constante strictement positive.

1. Formuler le problème aux deux bouts associé (où la commande n'apparaît pas).
2. Chercher une solution de ce système dynamique linéaire sous la forme $\lambda(t) = S(t)x(t)$. Quelles sont des conditions suffisantes sur S pour obtenir une solution ?

7.7. Programmation dynamique

On considère le maillage défini sur la Figure 3

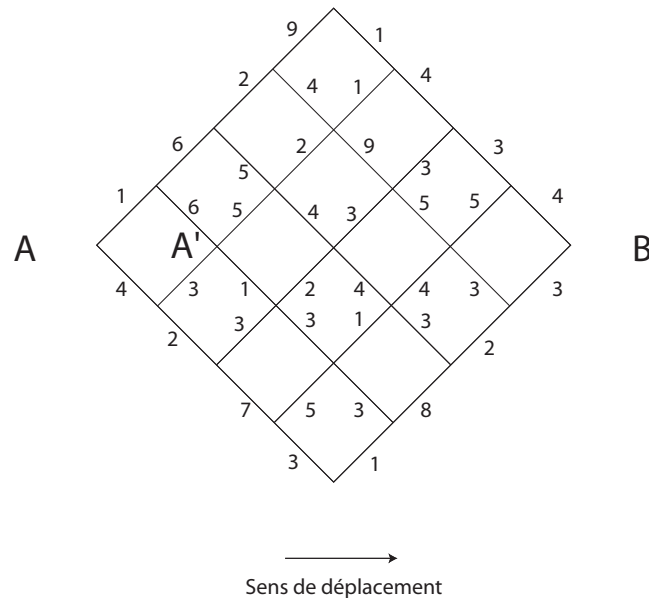


FIGURE 3. Chemins entre A et B.

À chaque segment du maillage est associé un coût. On cherche le meilleur chemin (plus petit coût total) pour aller de A à B , ainsi que de A' à B .

1. Calculer ces chemins par la méthode de la programmation dynamique. On rappelle que cette méthode consiste à construire un champs d'extrémales issues du point final de manière rétrograde.
2. Donner un champs d'extrémales.
3. On inverse maintenant le sens de parcours. Calculer le meilleur chemin de B à A et de B à A' . Quelle différence algorithmique constatez vous dans ce cas ?

7.8. Remplacement d'équipement

Vous devez rester à l'étranger pendant 3 années et vous avez besoin d'un véhicule pour aller travailler. Un véhicule se détériore avec le temps et son coût d'usage est une fonction croissante de son âge (entretien, pannes...). Chaque année vous pouvez conserver votre véhicule ou en acheter un autre neuf et revendre l'ancien. À l'issue des 3 années vous devez vous débarrasser de votre véhicule en le mettant à la casse pour un prix dépendant de son état. On note

- $c(i)$: coût d'usage d'un véhicule pendant 1 période, d'âge i au début de la période,
- $p = 50$: prix d'un nouveau véhicule,

- $t(i)$: valeur marchande d'un véhicule d'âge i au début de la période,
 - $s(i)$: valeur perçue lors de la mise à la casse d'un véhicule d'âge i .
1. À l'instant initial un ami vous donne une voiture qui a 2 ans d'âge. Quelle est la suite de décisions optimale ? Montrer le champs d'extrémales.

i	0	1	2	3	4
$c(i)$	10	13	20	40	70
$t(i)$		32	21	11	5
$s(i)$		25	17	8	0

2. On se place maintenant du point de vue du vendeur de voitures qui peut faire les mêmes calculs de programmation dynamique que vous. Connaissant votre situation, comment peut-il augmenter son profit sans que vous puissiez rien y faire ?

7.9. Tarification de billets d'avions

Lorsque vous achetez un billet d'avion aller retour, vous pouvez bénéficier d'une importante réduction si vous acceptez de rester le samedi soir sur place. Dans cet exercice on va étudier comment, dans un cas très simple, cette ristourne est calculée.

On considère deux populations : les touristes (avec les variables z_n, α_n) et les hommes d'affaires (avec les variables z_b, α_b).

On note z_n (resp. z_b) l'agrément de passer la nuit chez soi le samedi, p_n (resp. p_b) le prix du billet, α_n (resp. α_p) la pénibilité de payer le prix du billet, et t l'agrément du séjour. On note p_n le prix d'un billet aller retour avec séjour sur place le samedi, et $p_b \geq p_n$ le prix d'un billet aller retour sans séjour sur place le samedi.

1. Expliquer le sens des hypothèses (qu'on admettra dans la suite)

$$0 = z_n < z_b, \quad \alpha_n > \alpha_b \geq 0$$

2. Expliquer pourquoi l'agrément total d'un individu s'exprime, suivant les cas, comme

$$u = z - \alpha p + t \quad , \quad \text{où } u = -\alpha p + t$$

3. On souhaite donner l'envie aux touristes et aux hommes d'affaires de voyager. On souhaite choisir une politique tarifaire telle que les touristes aient envie de voyager en passant la nuit de samedi sur place et on souhaite encourager les hommes d'affaire à prendre un billet sans séjour du samedi. Expliquer qu'on aboutit aux relations

$$(56) \quad -\alpha_n p_n + t \geq 0$$

$$(57) \quad z_b - \alpha_b p_b + t \geq 0$$

4. On souhaite que des deux types de voyages le plus intéressant pour les touristes soit le voyage avec séjour du samedi, et que les hommes d'affaires privilégient

les séjours sans séjour du samedi. Montrer qu'on aboutit à

$$(58) \quad -\alpha_n p_n \geq z_n - \alpha_n p_b$$

$$(59) \quad z_b - \alpha_b p_b \geq -\alpha_b p_n$$

5. Étant donnés les flux supposés égaux d'hommes d'affaires et de touristes on souhaite maximiser la fonction

$$p_n/2 + p_b/2$$

Les variables de décisions étant p_n et p_b (la politique tarifaire), écrire le problème d'optimisation sous contraintes.

6. En utilisant les conditions de Karush-Kuhn-Tucker, combien y-a-t-il de cas à considérer pour résoudre ce problème ?
7. Montrer que les conditions (57) et (58) sont redondantes avec (56) et (59). Ré-écrire le problème d'optimisation sous contraintes correspondant.
8. Résoudre ce problème d'optimisation et montrer calculer, à l'optimum, $p_b - p_n$. Interpréter le résultat.
9. Reprendre l'étude si les flux de touristes et d'hommes d'affaires sont différents.
10. On suppose maintenant que la compagnie vend aussi des nuits d'hôtel et que la nuit du samedi soir est facturée c . Comment doit-on modifier l'étude précédente pour maximiser le profit total ?

7.10. Forme d'une chaîne pesante

Considérons une chaîne de longueur $2L$ suspendue entre deux points de même altitude situés à une distance $2l$ l'un de l'autre ($l \leq L$). Utiliser le calcul des variations pour trouver la forme qui minimise l'énergie potentielle (g la gravité est une constante).

7.11. Équations d'Euler Lagrange pour les problèmes contraints à un temps variable

Avec les notations utilisées en cours, déterminer les équations à résoudre pour trouver un extremum du problème suivant

$$\min \phi(x(t_f), t_f) + \int_{t_0}^{t_f} L(x(t), u(t), t) dt$$

sous les contraintes

$$x(t_0) = a \text{ (donné) }, \dot{x}(t) = f(x(t), u(t), t), \psi(x(t_f), t_f) = 0$$

où t_f est une inconnue.

7.12. Temps minimum pour traverser une rivière à courant parabolique

Le problème est de déterminer l'orientation $\theta(t)$ qu'un nageur doit choisir pour minimiser le temps nécessaire pour traverser une rivière dont il connaît le courant. La rivière est définie comme $\{(x, y), -l \leq y \leq l\}$. θ est l'angle entre le vecteur vitesse du nageur et l'axe x .

- Considérer le problème de traverser une rivière d'un point $(x, y) = (0, -l)$ à un point final donné $(0, l)$ où le courant est selon la direction x avec la vitesse

$$u = u_0 \left(1 - \frac{y^2}{h^2}\right)$$

Expliquer comment on peut déterminer $\theta(t)$ et t_f .

- Considérer le problème de traverser une rivière en partant de $(x, y) = (0, -l)$ pour arriver en $y = l$ (la valeur finale de x est libre). Expliquer comment trouver $\theta(t)$ et t_f .

7.13. Calcul des variations

On considère le problème de minimisation de la fonctionnelle

$$J(x) = \int_0^1 (\dot{x}(t))^2 dt$$

où x est une fonction deux fois dérivable définie sur $[0, 1]$ telle que

$$x(0) = \dot{x}(0) = x(1) = 0, \quad \dot{x}(1) = 1$$

1. Former l'équation d'Euler-Poisson correspondante.
2. Calculer x^* extrémale du problème.
3. On veut montrer que x^* fournit effectivement un minimum de la fonctionnelle. Considérer

$$J(x^* + h) = \int_0^1 (\dot{x}^*(t) + \dot{h}(t))^2 dt$$

où h est une variation admissible. En procédant à une double intégration par partie, montrer que

$$J(x^* + h) \geq J(x^*)$$

Conclure.

7.14. Inégalité de Kantorovich

On cherche à montrer le résultat suivant. Soit Q une matrice symétrique définie positive de taille $n \times n$. Quel que soit x , on a

$$q(x) = \frac{(x^T x)^2}{(x^T Q x)(x^T Q^{-1} x)} \geq \frac{4aA}{(a + A)^2}$$

où a et A sont, respectivement, la plus petite et la plus grande des valeurs propres de Q .

1. Dans un premier temps, on considère que Q est une matrice diagonale dont les termes sont

$$0 < a = \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n = A$$

Montrer que

$$q(x) = \frac{(\sum_{i=1,\dots,n} x_i^2)^2}{(\sum_{i=1,\dots,n} \lambda_i x_i^2)(\sum_{i=1,\dots,n} \frac{1}{\lambda_i} x_i^2)}$$

2. Montrer que

$$q(x) = \frac{1}{f(\lambda_1, \dots, \lambda_n, x)g(\lambda_1, \dots, \lambda_n, x)}$$

avec

$$f(\lambda_1, \dots, \lambda_n, x) = \sum_{i=1,\dots,n} \lambda_i \frac{x_i^2}{\sum_{i=1,\dots,n} x_i^2}, \quad g(\lambda_1, \dots, \lambda_n, x) = \sum_{i=1,\dots,n} \frac{1}{\lambda_i} \frac{x_i^2}{\sum_{i=1,\dots,n} x_i^2}$$

3. En interprétant f et g définis précédemment comme les coordonnées d'un barycentre de points de coordonnées $(\lambda_i, 1/\lambda_i)$, représenter dans un plan l'ensemble des points de coordonnées (f, g) lorsque x varie (les λ_i restant constants). Montrer, pour $\lambda_1, \dots, \lambda_n$ fixés, qu'il existe $\alpha(x) \in [0, 1]$ tel que

$$f(\lambda_1, \dots, \lambda_n, x) = (1 - \alpha(x))\lambda_1 + \alpha(x)\lambda_n$$

$$g(\lambda_1, \dots, \lambda_n, x) \leq (1 - \alpha(x))\frac{1}{\lambda_1} + \alpha(x)\frac{1}{\lambda_n}$$

4. Dédurre de ce qui précède que

$$f(\lambda_1, \dots, \lambda_n, x)g(\lambda_1, \dots, \lambda_n, x) \leq \frac{(\lambda_1 + \lambda_n)^2}{4\lambda_1\lambda_n}$$

Conclure. Conclure dans le cas général où Q n'est pas diagonale.

5. On considère le problème de minimisation de la fonction quadratique

$$h(x) = \frac{1}{2}x^T Qx - b^T x$$

où b est un vecteur colonne. Montrer que l'algorithme du gradient à pas optimal engendre les itérations

$$x^{k+1} = x^k - \frac{(g^k)^T g^k}{(g^k)^T Q g^k} g^k$$

avec $g^k = Qx^k - b$.

6. On note

$$E(x) = \frac{1}{2}(x - x^*)^T Q(x - x^*)$$

où x^* est l'unique minimum de h . Calculer $E(x^{k+1})$ en fonction de $E(x^k)$ et de g^k . En utilisant l'inégalité de Kantorovich que vous venez d'établir, montrer que

$$E(x^{k+1}) \leq \left(\frac{A-a}{A+a} \right)^2 E(x^k)$$

7. Que peut-on en déduire sur la convergence de cette méthode de descente en fonction du conditionnement de la matrice Q ?

7.15. Dualité

1. Considérer le problème (très simple)

$$\min_{x \in \mathbb{R}, x \geq 1} |x|$$

Quelle est sa solution ? Former le Lagrangien du problème et considérer la fonction duale

$$f(x) = \sup_{\lambda \geq 0} \lambda(1 - x)$$

Expliciter les valeurs prises par cette fonction pour tout x . En déduire, par dualité, quelle est la solution du problème.

2. Considérer le problème

$$\begin{aligned} \min_{x \in \mathbb{R}^n,} \quad & c^T x \\ & Ax = b, \\ & x \geq 0 \end{aligned}$$

Montrer que son dual est

$$\begin{aligned} \max_{(\lambda, \nu)} \quad & -b^T \nu \\ & A^T \nu - \lambda + c = 0 \\ & \lambda \geq 0 \end{aligned}$$

Préciser les dimensions de λ et ν . Peut-on simplifier cette formulation duale ?

3. Soit A une matrice de taille $m \times n$ et b un vecteur colonne. On cherche $x \in \mathbb{R}^n$ solution du problème

$$\min_{Ax \leq b} c^T x$$

Former le Lagrangien du problème. Par dualité, montrer que ce problème revient à

$$\begin{aligned} \max \quad & -b^T \lambda \\ & A^T \lambda + c = 0 \\ & \lambda \geq 0 \end{aligned}$$

7.16. Résultats sur la minimisation d'une fonction elliptique

On considère une fonction $J : \mathbb{R}^n \rightarrow \mathbb{R}$ qu'on suppose *elliptique*, c'est à dire qu'elle est continûment différentiable, que son gradient ∇J est Lipschitzien avec comme constante M , et qu'il existe $\alpha > 0$ tel que

$$(\nabla J(v) - \nabla J(u))^T (v - u) \geq \alpha \|u - v\|^2$$

pour tout u et v . On s'intéresse à la minimisation de J sur le domaine

$$\mathcal{U} = \{u \in \mathbb{R}^n \text{ tels que } \phi_i(u) \leq 0, \quad 1 \leq i \leq m\}$$

où les fonctions $\phi_i : \mathbb{R}^n \rightarrow \mathbb{R}$ sont convexes.

1. Établir, en utilisant la formule de Taylor avec reste intégral pour évaluer $J(v) - J(u)$, et le fait que J est elliptique, la double inégalité suivante :

$$\nabla J(u)^T(v - u) + \frac{\alpha}{2} \|u - v\|^2 \leq J(v) - J(u) \leq \nabla J(u)^T(v - u) + \frac{M}{2} \|v - u\|^2$$

2. En faisant le lien avec la convexité forte, démontrer l'existence et l'unicité de la solution du problème d'optimisation

$$\min_{u \in \mathcal{U}} J(u)$$

3. On définit le Lagrangien $\mathcal{L}(u, \lambda) = J(u) + \lambda^T \phi(u)$. Montrer que si (u, λ) est point selle du Lagrangien, alors u est solution du problème.

4. On note $p : \mathbb{R} \rightarrow \mathbb{R}_+$ tel que $p(x) = \begin{cases} x & \text{si } x \geq 0 \\ 0 & \text{sinon} \end{cases}$, et $P : \mathbb{R}^m \rightarrow \mathbb{R}^m$ tel

$$\text{que } P(x = (x_1, \dots, x_m)^T) = \begin{pmatrix} p(x_1) \\ \vdots \\ p(x_m) \end{pmatrix} \text{ l'opérateur de } \textit{projection} \text{ sur } (\mathbb{R}_+)^m.$$

Établir, en utilisant les conditions Karush-Kuhn-Tucker, l'équivalence suivante

$$\begin{aligned} & \{\lambda = P(\lambda + \rho \phi(u)), \rho > 0 \text{ fixé}\} \\ \Leftrightarrow & \begin{cases} \lambda \in (\mathbb{R}_+)^m \\ \sum_i \phi_i(u)(\mu_i - \lambda_i) \leq 0 \text{ pour tout } \mu = (\mu_1, \dots, \mu_m)^T \in (\mathbb{R}_+)^m \end{cases} \end{aligned}$$

5. En déduire (en utilisant certaines valeurs bien choisies de μ_i) l'équivalence

$$\begin{aligned} & \{\lambda = P(\lambda + \rho \phi(u)), \rho > 0 \text{ fixé}\} \\ \Leftrightarrow & \{ \lambda \in (\mathbb{R}_+)^m \mid \lambda^T \phi(u) = 0 \} \end{aligned}$$

6. Vérifier que si (u, λ) est point selle du Lagrangien, alors $\lambda = P(\lambda + \rho \phi(u))$ pour un certain ρ , et que

$$\mathcal{L}(u, \lambda) \leq \mathcal{L}(u + \theta(v - u), \lambda)$$

pour tout v et pour tout $\theta \in [0, 1]$. Faire le lien avec l'algorithme d'Uzawa.

7. En déduire, en utilisant un développement de Taylor que

$$\nabla J(u)^T(v - u) + \lambda^T \phi(v) \geq 0$$

pour tout v .

7.17. Une inégalité fonctionnelle

On souhaite établir le résultat suivant. Soit $y : [0, 1] \rightarrow \mathbb{R}$ une fonction régulière telle que $y(0) = 0$. Pour tout $k \in \mathbb{N}$, on a :

$$\int_0^1 (y(x))^{2k} dx \leq C \int_0^1 (y'(x))^{2k} dx$$

avec $C = \frac{1}{2k-1} \left(\frac{2k}{\pi} \sin \frac{\pi}{2k} \right)^{2k}$. Pour cette démonstration, on va utiliser une méthode variationnelle.

1. Soit la fonctionnelle

$$J(y) = \int_0^1 (C(y'(x))^{2k} - (y(x))^{2k}) dx$$

Ecrire l'équation d'Euler-Lagrange associée à l'extrémalité de J . Montrer en particulier que

$$(2k-1)C(y'(x))^{2k} = C' - (y(x))^{2k}, \forall x \in [0, 1]$$

avec C' une constante.

2. On s'intéresse désormais à l'extrémale Y telle que $Y(0) = 0$, $Y(1) = 1$.
 - (a) montrer que

$$1 = ((2k-1)C)^{\frac{1}{2k}} \int_0^1 \frac{dy}{(C' - y(x)^{2k})^{\frac{1}{2k}}}$$

- (b) on rappelle

$$\int_0^1 \frac{dt}{(1 - t^{2k})^{\frac{1}{2k}}} = \frac{\pi}{2k} \left(\sin \frac{\pi}{2k} \right)^{-1}$$

En déduire la valeur de C' .

- (c) montrer que $Y'(1) = 0$.

3. Parmi les extrémales, celle précédemment calculée, Y , réalise effectivement le minimum de J .

- (a) Montrer que $J(Y) = 0$. On pourra pour cela utiliser une intégration par parties en écrivant

$$J(Y) = \int_0^1 (CY'(x)Y'(x)^{2k-1} - Y(x)^{2k}) dx$$

pour faire apparaître l'équation d'Euler-Lagrange.

- (b) En déduire le résultat recherché.

7.18. Unicité du temps minimum

On considère le problème de transition en temps minimum entre deux points x_0 , x_1 pour un système linéaire scalaire (dans un premier temps) $\dot{x} = ax + bu$, avec $a \neq 0$, $b \neq 0$ sous la contrainte $u \in \mathcal{U}$ où \mathcal{U} est un sous ensemble convexe de \mathbb{R} . On cherche à démontrer que la commande réalisant le minimum est unique.

1. On suppose, sans perte de généralité, que $x_1 > x_0$, $a > 0$, $b > 0$ et que $\mathcal{U} = [-\frac{a}{b}x_1 - \epsilon, -\frac{a}{b}x_0 + \epsilon]$, avec $\epsilon > 0$. En considérant la loi horaire

$$x(t) = \begin{cases} x_0, & \text{si } t \leq 0 \\ x_0 + \frac{t}{T}(x_1 - x_0), & \text{si } 0 < t < T \\ x_1 & \text{si } T \leq t \end{cases}$$

où $T > 0$ est un paramètre à choisir, montrer que pour T suffisamment grand, la commande $u(t) = \frac{\dot{x}(t) - ax(t)}{b}$ appartient à \mathcal{U} et permet de rallier x_0 à x_1 .

2. Dédire de ce qui précède que si T fournit une solution admissible, alors $T' > T$ fournit aussi une solution admissible.
3. Considérer un temps t_1 candidat à l'optimalité. On veut montrer que si deux lois horaires $[0, t_1] \ni t \mapsto u_1(t)$ et $[0, t_1] \ni t \mapsto u_2(t)$ sont optimales alors elles coïncident.

- (a) montrer que pour toute loi commande u , on a

$$x(t) = x_0 \exp(at) + \int_0^t \exp(a(t-\tau))bu(\tau)d\tau$$

- (b) en déduire

$$\int_0^{t_1} \exp(-at)bu_1(t)dt = \int_0^{t_1} \exp(-at)bu_2(t)dt$$

- (c) former l'Hamiltonien du problème (on notera λ l'état adjoint). Dédire du principe du minimum de Pontryagin que la ou les solutions optimales sont données par

$$\arg \min \lambda(t)bu(t)$$

- (d) former l'équation différentielle adjointe satisfaite par λ .

- (e) montrer, par le lemme de l'adjoint, que

$$\int_0^{t_1} \lambda(t)bu_1(t)dt = \int_0^{t_1} \lambda(t)bu_2(t)dt$$

- (f) déduire du principe du minimum de Pontryagin que $\lambda(t)u_1(t) \leq \lambda(t)u_2(t)$ et que, d'après l'égalité précédente on a alors

$$\lambda(t)u_1(t) = \lambda(t)u_2(t)$$

presque partout. En déduire sous quelle condition on peut conclure

$$u_1(t) = u_2(t)$$

presque partout.

- (g) dans le cas où x est un vecteur et u un scalaire, proposer une généralisation.

7.19. Construction d'une autoroute

On considère le problème de construire une route à pente limitée sur un relief constitué de collines. On représente le problème dans un plan vertical pour simplifier. On repère $[O, L] \ni l \mapsto z(l)$ l'altitude du relief avant travaux, et $[O, L] \ni l \mapsto x(l)$ l'altitude après les travaux. Les travaux consistent à creuser ou à remblayer le terrain pour définir l'altitude x qui minimise le critère

$$\int_0^L (x(l) - z(l))^2 dl$$

On souhaite que la route ait une pente limitée. On écrit cette contrainte sous la forme

$$\dot{x} = u, \quad |u| \leq a, \quad a > 0 \text{ donnée.}$$

1. Ecrire l'Hamiltonien de ce problème de commande optimale
2. Former le problème aux deux bouts. Montrer que la valeur finale de l'état adjoint λ est $\lambda(L) = 0$.
3. En utilisant le principe de minimum de Pontryagin, montrer que
 - si $\lambda > 0$, $u = -a$
 - si $\lambda = 0$, $u = \dot{z}$
 - si $\lambda < 0$, $u = a$
4. En déduire que la solution optimale consiste en une succession de travaux sur des intervalles $[t_i, t_{i+1}]$ tels que

$$\int_{t_i}^{t_{i+1}} (z(l) - x(l)) dl = 0$$

On ne cherchera pas à calculer les t_i . Interpréter l'égalité précédente.

7.20. Programmation dynamique

On considère un problème de minimisation du coût de trajet entre une destination et tous les points d'un graphe. Le graphe (non orienté) est représenté sur la figure 4. Le coût de trajet entre 2 cases est précisé sur le segment les reliant. A tout instant, il est possible de rester immobile sur une case, avec un coût 0 pendant l'arrêt. Construire

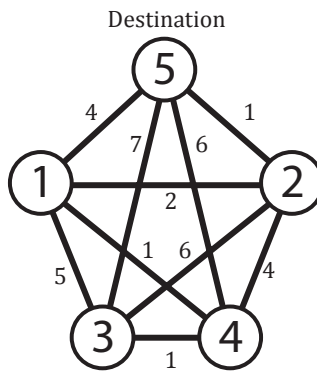
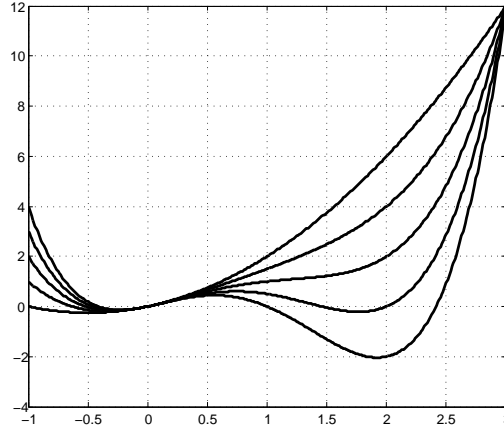


FIGURE 4. Graphe

un champ d'extrémales par la méthode de la programmation dynamique en partant

FIGURE 5. Graphe de la fonction $f(x, \theta)$ pour différentes valeurs de θ .

de la destination. On construira ainsi progressivement les chemins menant en 1, 2, 3, 4, étapes à la destination. D'après cette analyse, quel est le meilleur chemin pour aller de la case 3 à la case 5 ?

7.21. Sur un théorème de C. Berge

On s'intéresse au problème de minimisation (globale)

$$(60) \quad \min_{x \in [a(\theta); b(\theta)]} f(x, \theta)$$

où f est une fonction à valeurs dans \mathbb{R} , continue par rapport à chacune de ses variables et où $a, b : \mathbb{R} \rightarrow \mathbb{R}$ sont continues et vérifient en particulier la propriété suivante : $\forall \theta \in \mathbb{R}, a(\theta) < b(\theta)$ (pour assurer la bonne définition de l'intervalle considéré dans (60)). Le graphe d'une telle fonction f est représentée sur la figure 5.

On définit

$$(61) \quad x^*(\theta) = \operatorname{argmin}_{x \in [a(\theta); b(\theta)]} f(x, \theta)$$

$$(62) \quad f^*(\theta) = f(x, \theta), \quad x \in x^*(\theta)$$

Lorsque deux points donnent la même valeur minimale globale de f , alors l'ensemble $x^*(\theta)$ contient ces deux points. Le but de l'exercice est d'étudier la régularité de la fonction f^* .

1. (a) Justifier que, pour tout $\theta \in \mathbb{R}$, l'ensemble $x^*(\theta)$ est non-vidé.
 (b) Illustrer par un exemple (schéma) le fait que $x^*(\theta)$ peut varier non continûment en θ .
2. Soit $\theta \in \mathbb{R}$. On considère une suite (θ_n) convergente de limite θ et (x_n) avec $x_n \in x^*(\theta_n)$.

- (a) Justifier que l'on puisse considérer une sous-suite convergente de (x_n) , dont on notera x la limite.
- (b) Montrer par l'absurde que $x \in x^*(\theta)$ (on pourra exhiber une suite convergente de limite $\hat{x} \in x^*(\theta)$, dont on justifiera l'existence).
- (c) Conclure que f^* est continue en θ .

7.22. Optimisation sur un espace fonctionnel

Soit un intervalle $[a, b]$, $a < b$. On considère l'espace de Hilbert $H = L_2[a, b]$ constitué des fonctions $y : [a, b] \rightarrow \mathbb{R}^n$ de carré sommable

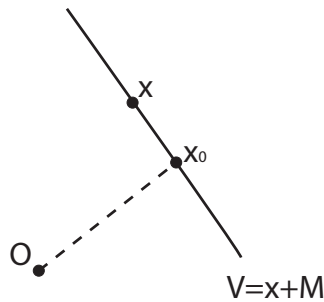
$$\|y\|^2 = \langle y, y \rangle = \int_a^b \|y(t)\|^2 dt < +\infty$$

où on a utilisé la norme Euclidienne de \mathbb{R}^n sous l'intégrale.

On souhaite utiliser le théorème de projection suivant (voir Luenberger, Optimisation by vector spaces methods, Wiley-Interscience 1997) : soit H un Hilbert et M un sous espace fermé de H , alors, à tout $x \in H$ il correspond un unique $x_0 \in V$ où V est l'espace affine $V = x + M$ tel que

$$\|x_0\| \leq \|v\|, \quad \forall v \in V$$

Ce point x_0 est orthogonal à M , on note $x^0 \in M^\perp$.



- On considère $\{y_1, \dots, y_n\}$ un ensemble de n vecteurs (fonctions) indépendants de H et $\{c_1, \dots, c_n\}$ n scalaires (constantes). On note M l'espace engendré par $\{y_i\}_{i=1\dots n}$ et on considère $V = \{x \in H \mid \langle x, y_i \rangle = c_i, \forall i = 1\dots n\}$. Caractériser M^\perp et en déduire que V est un espace affine translaté de M^\perp .
- En appliquant le théorème de projection ci-dessus à V , montrer qu'il existe un unique $x_0 \in V$ solution de

$$\min_{x \in V} \|x\|$$

3. En utilisant le fait que $(M^\perp)^\perp = M$, montrer que $x_0 = \sum_{i=1}^n \beta_i y_i$ où les β_i sont solutions du système linéaire

$$(< y_i, y_j >)_{i=1 \dots n, j=1 \dots n} \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$$

4. Application. On considère le problème suivant

$$\min_{u \in L_2[0,1]} \int_0^1 u^2(t) dt$$

sous les contraintes

$$\dot{x}_1 = x_2, \quad \dot{x}_2 = -x_2 + u, \quad x_1(0) = x_2(0) = 0, \quad x_1(1) = 1, \quad x_2(1) = 0$$

- par intégration, montrer que $x_2(1) = \int_0^1 \exp(t-1)u(t)dt$ et que $x_1(1) = \int_0^1 (1 - \exp(t-1))u(t)dt$
- écrire les contraintes du problème ci-dessus sous la forme $< u, y_1 > = c_1, \quad < u, y_2 > = c_2$.
- en déduire que la solution du problème est $u(t) = \beta_1 y_1(t) + \beta_2 y_2(t)$ où β_1 et β_2 sont solutions d'un système d'équations linéaires qu'on précisera.

7.23. Meilleur antécédent par une matrice non inversible

On cherche le meilleur antécédent d'un vecteur $b \in \mathbb{R}^n$ par une matrice A de taille $n \times n$ non inversible.

1. On considère le problème d'optimisation

$$\min_u \|Au - b\|$$

où $\|\cdot\|$ désigne la norme Euclidienne. Montrer que ce problème est convexe. En déduire que u réalise le minimum de la fonction objectif si et seulement si

$$A^T A u = A^T b$$

2. La matrice A étant non inversible, montrer que $A^T A$ ne l'est pas non plus.
3. Montrer que la matrice $A^T A$ est symétrique positive. On note $\lambda_1, \dots, \lambda_m$ ses m valeurs propres non nulles, avec $m < n$. En déduire que

$$A^T A = \sum_{i=1}^m \lambda_i v_i v_i^T$$

où v_i sont des vecteurs propres orthonormaux de $A^T A$.

4. On considère la matrice

$$P = A^T A \left(\sum_{i=1}^m \frac{1}{\lambda_i} v_i v_i^T \right)$$

Montrer que $\text{rang } P = m$, $P^2 = P$, et que $\text{Im } P = \text{Im } A^T$. En déduire que P est la projection orthogonale sur $\text{Im } A^T$.

5. En déduire que le vecteur suivant

$$u^* = \left(\sum_{i=1}^m \frac{1}{\lambda_i} v_i v_i^T \right) A^T b$$

réalise l'optimum recherché. Montrer que si A est inversible, i.e. $m = n$, alors $Au^* = b$.

7.24. Introduction aux méthodes de points intérieurs

On s'intéresse au problème de programmation quadratique (QP) suivant

$$(63) \quad \begin{cases} \min \frac{1}{2} z^T P z + f^T z \\ Mz \leq b \end{cases}$$

où P est une matrice $n \times n$ symétrique définie positive, M une matrice $m \times n$, $f \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, avec $m > n$.

L'objet de cet exercice est d'exposer la construction d'un algorithme dit de "points intérieurs".

1. Montrer que le problème (63) admet une unique solution optimale.
2. On cherche à réécrire ce problème sous une forme canonique où les inégalités sont juste des relations de positivité des inconnues. Pour cela, on introduit un vecteur s de m inconnues positives ou nulles, et un vecteur y de $2n$ inconnues positives ou nulles. Chaque composante i de z est exprimable sous la forme

$$z_i = y_i - y_{i+n}, \quad y_i \geq 0, \quad y_{i+n} \geq 0$$

Montrer qu'on peut réécrire (63) sous la forme

$$(64) \quad \begin{cases} \min \frac{1}{2} x^T Q x + c^T x \\ Ax = b \\ -x \leq 0 \end{cases}$$

avec

$$x = \begin{pmatrix} y \\ s \end{pmatrix}, \quad Q = \begin{pmatrix} N^T P N & 0_{2n \times m} \\ 0_{m \times 2n} & 0_{m \times m} \end{pmatrix},$$

$$c^T = (f^T N \quad 0_{1 \times m}), \quad A = (MN \quad I_m)$$

où N est une matrice qu'on explicitera.

3. Former le problème dual de (64) (on l'écrira comme un problème de maximisation sous contraintes). Pour cela, on formera le Lagrangien

$$L = \frac{1}{2} x^T Q x + c^T x + \lambda^T (b - Ax) - \mu^T x$$

4. Montrer que les conditions de Karush-Kuhn-Tucker pour le problème général

$$\begin{cases} \min f(x) \\ d(x) = 0, \quad e(x) \leq 0 \end{cases}$$

sont

$$(65) \quad \begin{cases} \nabla f(x^*) + \sum \lambda_i^* \nabla d_i(x^*) + \sum \mu_i^* \nabla e_i(x^*) = 0 \\ d(x^*) = 0 \\ e(x^*) \leq 0 \\ \mu^* \geq 0 \\ \mu_i^* c_i(x^*) = 0 \end{cases}$$

5. Former les conditions (65) pour le problème (64).
 6. On va pénaliser les contraintes $-x \leq 0$ dans le coût en considérant le problème, pour $\varepsilon > 0$,

$$(66) \quad \begin{cases} \min \frac{1}{2} x^T Q x + c^T x - \varepsilon \sum \ln(x_i) \\ Ax = b \end{cases}$$

Montrer que le problème (66) possède une unique solution.

7. Former le Lagrangien associé à (66) et montrer que les conditions de Karush-Kuhn-Tucker sont

$$(67) \quad \begin{cases} Ax = b \\ -Qx + A^T \lambda + s = c \\ x_i s_i = \varepsilon \\ x_i \geq 0 \\ s_i \geq 0 \end{cases}$$

8. Montrer comment résoudre les conditions (67) par une méthode de Newton avec projection. Former pour cela le Jacobien de la fonction à annuler. Former la première itération de cet algorithme en exhibant le calcul de la direction et la procédure de projection.
 9. Comment atteindre la solution de (65) ?
 10. Une méthode itérative consiste à considérer $\epsilon = \sigma \frac{x^T s}{2n+m}$, $\sigma \in]0, 1[$.
 Comment peut-on interpréter ce choix ?

7.25. Gestion thermique d'une batterie de véhicule hybride

Dans un véhicule hybride, une question importante est celle du refroidissement de la batterie qui perd de son efficacité (et de sa disponibilité) si elle est trop chaude. On va chercher à établir une politique optimale de refroidissement sur un horizon de temps donné $T > 0$ (et une route connue à l'avance). Soit θ_b la température interne

de la batterie, θ la température de son échangeur thermique, u son courant de sortie, et ξ sa charge. En tenant compte des échanges thermiques, de la consommation et de l'échauffement par effet Joule, on a les équations suivantes

$$\frac{d}{dt}\theta_b = \frac{a}{2}u^2 - b\theta, \quad \frac{d}{dt}\theta = cu - k\theta, \quad \frac{d}{dt}\xi = D(t) - u$$

où a, b, c, k sont des coefficients positifs (connus) et $D(t)$ est une fonction connue du temps. Le système est initialisé au temps 0 aux valeurs

$$\theta_b(0), \quad \theta(0), \quad \xi(0)$$

1. On cherche à minimiser la température finale $\theta_b(T)$ sous la contrainte de maintien de la charge $\xi(T) = \xi(0)$. Montrer qu'on peut formuler ce problème sous la forme

$$\begin{cases} \min & \int_0^T L(\theta, u) dt \\ \frac{d}{dt}\theta = cu - k\theta \\ \frac{d}{dt}\xi = D(t) - u \\ \xi(T) = \xi(0) \\ \theta(0), \xi(0) \text{ donnés} \end{cases}$$

où on précisera $L(\theta, u)$.

2. Former l'Hamiltonien associé au problème et les équations différentielles du problème aux deux bouts.
3. Intégrer, au moyen de deux constantes qu'on notera α et β les équations adjointes.
4. Donner l'expression de la commande optimale au cours du temps en fonction de α et β .
5. Expliciter une première relation entre α et β provenant de la contrainte $\xi(T) = \xi(0)$.
6. Trouver la valeur optimale de α en utilisant la relation précédente et en invoquant la stationnarité de la fonction coût à l'optimum.

7.26. Polygones inscrits du cercle de longueur maximale

Soit $f : I \rightarrow \mathbb{R}$ avec I intervalle de \mathbb{R} . On suppose que f est convexe sur I , et on considère $\{x_i, i = 1, \dots, n\}$ une famille d'éléments de I , n entier $n > 1$.

1. Montrer (par exemple par récurrence) que

$$f(x_1) + \dots + f(x_n) \geq nf\left(\frac{x_1 + \dots + x_n}{n}\right)$$

On a égalité si et seulement si tous les x_i sont égaux.

2. En utilisant l'inégalité précédente pour $f(x) = -\sin x$ et $I = [0, \pi]$, montrer que parmi les polygones à n côtés inscrits dans un cercle donné, les polygones réguliers ont le plus grand périmètre.

7.27. Programmation dynamique

On se donne une collection d'arcs orientés reliant des points P_1, P_2, P_3, P_4, P_5 . Tous les points sont pas connectés directement avec les autres. Les arcs sont orientés, c'est à dire qu'on peut aller de P_1 à P_2 directement, mais que le retour n'existe pas forcément. Les longueurs (coûts) des arcs sont rangés dans une matrice

$$M^1 = \begin{bmatrix} 0 & 3 & 7 & +\infty & -3 \\ +\infty & 0 & +\infty & 1 & 6 \\ +\infty & 4 & 0 & +\infty & +\infty \\ 3 & +\infty & -4 & 0 & +\infty \\ +\infty & +\infty & +\infty & 5 & 0 \end{bmatrix}$$

Ainsi, l'arc reliant P_1 à P_2 a une longueur 3, alors qu'il n'y a pas d'arc (on note que la longueur est infinie) reliant P_1 à P_4 .

On cherche à établir une carte des plus courts chemins entre chaque point du graphe. Pour résoudre ce problème on va utiliser la programmation dynamique.

1. Dessiner tous les arcs du graphe reliant les points P_1, P_2, P_3, P_4, P_5 . De combien d'arc le graphe est-il ainsi constitué ?
2. On note M^k la matrice dont chaque coefficient $M_{i,j}^k$ vaut la longueur totale de la séquence optimale de k arcs (ou moins) reliant P_i à P_j . En utilisant le principe d'optimalité de Bellman, montrer que

$$M_{i,j}^{k+1} = \min_{p=1,\dots,5} (M_{i,p}^1 + M_{p,j}^k)$$

3. En déduire quel est le coût de la meilleure séquence pour aller de P_1 à P_3 .

7.28. Problème de Dido

On veut montrer que parmi les courbes fermées régulières du plan (simples : ne faisant pas de boucles) de longueur donnée, celles qui ont une aire intérieure maximale sont des cercles⁽¹⁾.

On note \mathcal{C} une courbe dans le plan, fermée simple et régulière, et on la suppose paramétrée par son abscisse curviligne, en notant L sa longueur :

$$c : [0, L] \ni s \mapsto c(s) \in \mathbb{R}^2$$

Cette courbe (dite de Jordan) sépare le plan en une partie intérieure \mathcal{I} et une partie extérieure.

1. La longueur L est solution de l'équation intégrale

$$L = \int_0^L \left\| \frac{dc}{ds} \right\| ds$$

Vérifiez votre formule en calculant le périmètre d'un cercle de rayon r .

1. c'est la solution du célèbre problème de Dido de l'Antiquité

2. L'aire A de la zone \mathcal{I} est

$$A = \frac{1}{2} \int_0^L \det \left[c(s), \frac{dc}{ds}(s) \right] ds$$

Utiliser cette formule pour calculer l'aire d'un disque de rayon r .

3. On considère une variation autour de c , en utilisant $\nu(s)$ le vecteur normal (extérieur) à la courbe \mathcal{C} au point courant $c(s)$, et f une fonction régulière L -périodique à valeur dans \mathbb{R} . Ceci permet de construire, pour tout $t \in \mathbb{R}$,

$$v(t, s) = c(s) + tf(s)\nu(s)$$

Faire un dessin représentant c et v , mettant en évidence que lorsque $|t|$ est suffisamment petit, $s \mapsto v(t, s)$ définit encore une courbe fermée simple.

4. On note L_f et A_f la longueur et l'aire intérieure associées à $s \mapsto v(t, s)$. Montrer que la dérivée à $t = 0$ de $t \mapsto \frac{L_f(t)^2}{A_f(t)}$ s'annule si et seulement si

$$\frac{d}{dt} L_f(0) = \frac{L}{2A} \frac{d}{dt} A_f(0)$$

5. Ré-écrire cette condition en utilisant le fait que

$$\frac{d}{dt} L_f(0) = - \int_0^L k(s) f(s) ds, \quad \frac{d}{dt} A_f(0) = - \int_0^L f(s) ds$$

où k désigne la courbure de \mathcal{C} . On demande d'obtenir une condition que doit vérifier $k(s)$ pour que \mathcal{C} réalise un extremum de $\frac{L^2}{A}$.

6. Conclure par le lemme de duBois-Reymond que \mathcal{C} est nécessairement un cercle.

7.29. Calcul des variations avec saut de l'état

On considère un problème de commande optimale où le coût à minimiser est

$$J = \int_0^1 L(x(t), u(t)) dt + \varphi(x(1))$$

La particularité du problème est que x subit un saut en un instant $\tau \in]0, 1[$ défini implicitement par la condition (on suppose τ unique)

$$\phi(x(\tau^-)) = 0$$

A l'instant τ , l'état subit un saut de la forme

$$x(\tau^+) = x(\tau^-) - \Delta$$

où Δ est un vecteur de paramètres donné. Sur l'intervalle $[0, \tau[$, l'état satisfait une équation différentielle $\dot{x} = f_1(x, u)$, et sur l'intervalle $]\tau, 1]$, il satisfait $\dot{x} = f_2(x, u)$. La condition initiale $x(0) = x^0$ est imposée.

1. On demande d'établir le problème aux deux bouts correspondant. Pour cela, on pourra adjoindre les contraintes, et faire intervenir deux intégrales pour distinguer le saut. On montrera que l'état adjoint subit lui aussi un saut.

2. Sans détailler les calculs, expliquer comment ce formalisme permet de traiter le problème suivant : lors de l'optimisation de trajectoire d'une fusée, on cherche à se débarrasser des protections thermiques (la coiffe) lorsqu'en sortant de l'atmosphère les frottements aérodynamiques deviennent suffisamment faibles.

7.30. Parallélépipède maximal

Soit un ellipsoïde défini par l'équation

$$c(x, y, z) = x^2/a^2 + y^2/b^2 + z^2/c^2 - 1 = 0$$

avec a, b, c des scalaires non nuls. On cherche le parallélépipède rectangle dont les arêtes sont parallèles aux axes ayant un volume maximal en restant contenu dans l'ellipsoïde. On note x, y, z les demi-longueurs de ses arêtes, si bien que son volume est

$$f(x, y, z) = 8xyz$$

1. Écrire les conditions de stationnarité pour ce problème sous contrainte. On pourra noter λ le multiplicateur associé à la contrainte c .
2. En utilisant que le volume du parallélépipède solution est de volume non nul, montrer que λ est non nul.
3. Montrer que le volume optimal s'exprime directement en fonction de λ .
4. Éliminer λ des équations pour obtenir $x = a/\sqrt{3}$. En déduire que le volume optimal est $\frac{8abc}{3\sqrt{3}}$.

7.31. Rangement d'un câble

On dispose d'un câble *semi-rigide* de longueur L qu'on souhaite étaler sur le sol entre 2 points à connecter situés à une distance $L/2$. Le câble est donc trop long.

Pour ne pas trop encombrer, on va chercher à trouver une disposition au sol qui ne prend pas trop de place tout en respectant les contraintes mécaniques du câble. Pour déterminer cette disposition, on va résoudre un problème d'optimisation de trajectoires.

On considère l'axe x reliant le point de départ et le point d'arrivée. On complète le repère avec un axe orthogonal y et on résoud le problème suivant

$$\min \frac{K}{2}(x(T) - L/2)^2 + \int_0^T y(\tau)^2 d\tau$$

sous les contraintes

$$\dot{x} = \cos \theta$$

$$\dot{y} = \sin \theta$$

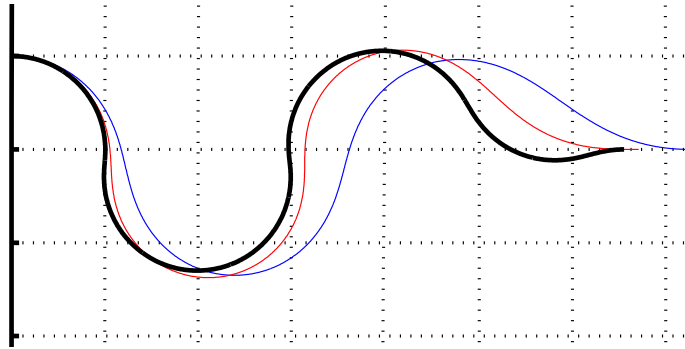
$$\dot{\theta} = u$$

$$u \in [-\alpha, \alpha]$$

$$x(0) = 0, y(0) = 0, \theta(0) = 0, y(T) = 0, \theta(T) = 0$$

avec K et α des paramètres positifs.

1. Faire un schéma du problème et esquisser la solution.
2. Former le problème aux 2 bouts correspondant.
3. Comment doit-on choisir T pour prendre en compte la longueur donnée du câble ?
4. Quel paramètre du problème permet de prendre en compte la rigidité du câble ?
5. Quel paramètre permet d'assurer que $x(T) \approx L/2$?



7.32. Investissement public

Lors d'un investissement public, deux options s'offrent en général aux décideurs :

1. investir dans un endroit possédant déjà une forte activité économique
2. investir dans une zone défavorisée.

Considérant que chaque territoire du pays se voit attribuer un budget d'investissement économique $x(z)$ où z désigne le territoire, on note $f(x)$ la rentabilité de l'investissement (fonction supposée continue et différentiable avec $f(0) = 0$). Le décideur qui doit prendre une mesure à l'échelle nationale cherche à maximiser la rentabilité totale

$$\int_{\text{pays}} f(x(z)) dz$$

Si on considère qu'un investissement est plus rentable sur un territoire pauvrement doté, on décrit

$$f(x) \geq f(y) \text{ implique } f'(x) \leq f'(y)$$

Montrer à cette hypothèse politique correspond au fait que f est concave et croissante.

7.33. Convexité

Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction quadratique strictement convexe. Soit x, y_1, \dots, y_n des vecteurs de \mathbb{R}^n .

1. Montrer que la fonction $\mathbb{R}^n \ni (\lambda_1, \dots, \lambda_n) \mapsto f(x + \sum_{i=1}^n \lambda_i y_i)$ est également quadratique convexe.
2. Donner l'expression de son Hessien.
3. A quelle condition est-elle strictement convexe ?

7.34. Projeté sur une parabole

Soit la parabole d'équation $5y = (x - 1)^2$. On cherche le point de la parabole de distance Euclidienne minimale avec le point de coordonnées $(1, 2)$.

1. Définir un problème de minimisation sous contraintes correspondant.
2. Trouver les points satisfaisant les conditions de Lagrange.
3. Déterminer la solution.
4. Procéder par substitution directe de y en fonction de x et formuler un problème de minimisation de la seule variable x . Retrouver ainsi le résultat de la question précédente.

7.35. Existence de minimums

On considère les trois problèmes suivants. Expliquer lesquels possèdent des solutions.

$$(68) \quad \min x_1 + x_2 \text{ sous les contraintes } x_1^2 + x_2^2 = 1, \ 0 \leq x_1 \leq \frac{1}{2}, \ 0 \leq x_2$$

$$(69) \quad \min x_1 + 2x_2 \text{ sous les contraintes } x_1^2 + 2x_2^2 \leq 1, \ x_1 + x_2 = 10$$

$$(70) \quad \min x_1 \text{ sous les contraintes } x_1 + x_2 = 1$$

7.36. Pénalité intérieure

On s'intéresse ici à une classe de méthodes permettant de résoudre un problème de minimisation sous contraintes comme limite d'une séquence de problèmes sans contraintes. On note

$$\min_x J(x), \quad \text{sous les contraintes } g_i(x) \leq 0, \ i = 1 \dots m$$

où on suppose que les fonctions J et g_i sont régulières de $\mathbb{R}^n \rightarrow \mathbb{R}$, et que les contraintes définissent un ensemble compact non vide S , contenant une unique solution globale x^* , non isolée dans S . On veut résoudre le problème en résolvant une séquence de problèmes sans contraintes dits "pénalisés" du type

$$\min_{x \in \mathbb{R}^n} J(x) + \epsilon \sum_{i=1}^m \gamma(g_i(x)), \quad \epsilon > 0$$

où γ est une fonction positive, régulière sur $] -\infty, 0[$, telle que $\lim_{x \rightarrow 0-} \gamma(x) = +\infty$. Par convention, $\gamma(x) = +\infty$, pour $x \geq 0$. On considère une suite strictement décroissante (ϵ_k) de réels strictement positifs et on note, pour tout k

$$J_k(x) = J(x) + \epsilon_k p(x) \triangleq J(x) + \epsilon_k \sum_{i=1}^m \gamma(g_i(x))$$

On suppose que J_k admet un minimum sur \mathbb{R}^n en un point unique noté x_k .

1. Montrer que pour tout k , x_k est dans l'intérieur de S .
2. Montrer que pour tout k on a

$$(71) \quad J_k(x_k) \leq J_k(x_{k+1})$$

$$(72) \quad J_{k+1}(x_{k+1}) \leq J_{k+1}(x_k)$$

3. Dédurre, par une combinaison linéaire des lignes qui précèdent, que

$$(73) \quad J(x_{k+1}) \leq J(x_k)$$

4. Montrer, en utilisant (71) et la relation précédente que

$$p(x_k) \leq p(x_{k+1})$$

5. Montrer, en utilisant (72) que

$$(74) \quad J_{k+1}(x_{k+1}) \leq J_k(x_k)$$

6. Montrer que pour tout $\delta > 0$, il existe $x^\delta \in S$ tel que $J(x^\delta) < J(x^*) + \delta/2$
7. On choisit un entier l tel que

$$\epsilon_l p(x^\delta) < \delta/2$$

Montrer, en utilisant (74), qu'on a, pour tout $k > l$

$$J_k(x_k) \leq J_l(x^\delta)$$

8. En déduire que

$$J(x^*) \leq J(x_k) < J(x^*) + \delta$$

pour tout $k > l$ et que donc

$$\lim_{k \rightarrow \infty} J(x_k) = J(x^*)$$

9. Montrer enfin que (x_k) converge vers x^* .

La séquence des problèmes non contraints mais pénalisés génère une suite qui converge donc vers la solution x^* du problème contraint.

7.37. PSA : profit sharing agreement

Le PSA est un type de contrat fréquemment utilisé dans le domaine des ressources naturelles, notamment dans l'industrie pétrolière. Un PSA définit pour la durée de vie d'un réservoir (typiquement 25 ans) la part du pétrole produit qui va revenir, à chaque instant (année), à l'exploitant du gisement et au pays propriétaire du gisement, respectivement. Deux données importantes sont à prendre en compte dans le cas d'un réservoir pétrolier : *i*) si on produit beaucoup au début, on risque de "fatiguer" le réservoir, ce qui entraînera un épuisement rapide des réserves ultérieures *ii*) un exploitant cherchera naturellement à rentabiliser le plus vite possible les investissements lourds qu'il a réalisés lors de l'équipement du réservoir (forage, cimenterie du puits, instrumentation, riser, etc...). Ces deux effets sont antagonistes.

On considère que le problème peut-être décrit au moyen de 2 variables d'état : r les réserves exploitables dans le réservoir, p le profit cumulé de l'exploitant. Ces deux variables satisfont les équations différentielles

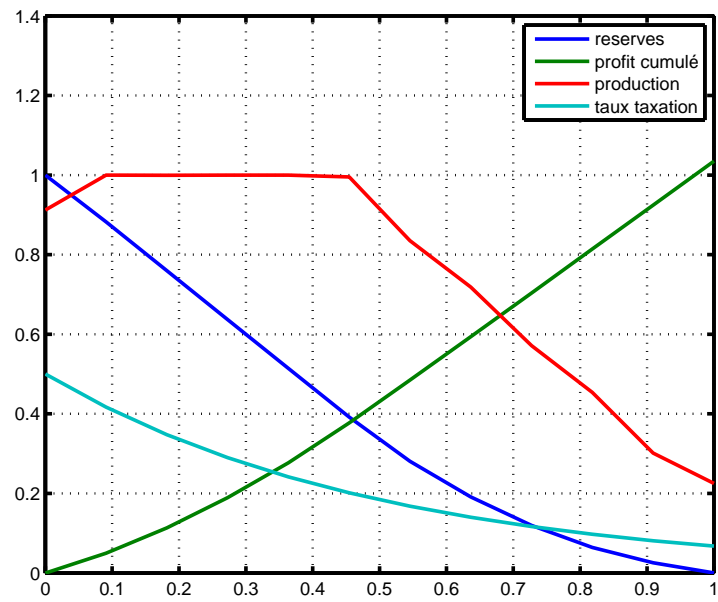
$$\begin{aligned}\dot{r} &= -uh(u) \\ \dot{p} &= uv + \epsilon p\end{aligned}$$

où $u \in [0, 1]$ désigne la production instantanée, h est une fonction d'épuisement positive croissante, ϵ définit les intérêts du capital. Le contrat PSA définit la fonction $v(t) \in [0, 1]$ qui est la fraction de la production revenant à l'exploitant. Le reste revient au pays propriétaire du gisement.

1. On considère le point de vue de l'exploitant. Il souhaite maximiser son profit cumulé $p(T)$, où $T = 25$ ans. Le réservoir est initialement plein $r(0) = 1$. Il est vide en fin d'exploitation $r(T) = 0$. Le cumul du profit est initialement nul $p(0) = 0$. Former le problème d'optimisation sous contraintes. Former l'Hamiltonien du système.
2. On considère $\epsilon = 0$ et v constante. Montrer que les états adjoints sont constants. Montrer, qu'il existe une production u constante satisfaisant le principe du minimum de Pontryaguine. Interpréter ce résultat.
3. Former le problème aux deux bouts dans le cas général.
4. On considère maintenant le point de vue du pays propriétaire du gisement. Son intérêt est de maximiser la quantité totale de pétrole qui lui revient produite au cours de toute la durée de vie du réservoir (jusqu'à épuisement $r(T) = 0$). Montrer que le critère à minimiser est

$$\int_0^T u(t)(v(t) - 1)dt$$

5. Former le problème aux deux bouts correspondant.



BIBLIOGRAPHIE

- [1] A. D. Aleksandrov, A. N. Kolmogorov, and M. A. Lavrent'ev. *Mathematics. Its contents, methods and meaning*. The M. I. T. Press, 1969.
- [2] G. Allaire. Analyse numérique et optimisation. cours de l'école polytechnique. Technical report, École Polytechnique, 2004.
- [3] D. P. Bertsekas. *Nonlinear programming*. Athena Scientific, 2 edition, 1999.
- [4] D. P. Bertsekas. *Dynamic Programming and Optimal Control*, volume 1, 2. Athena Scientific, 2 edition, 2001.
- [5] F. Bonnans. The shooting algorithm for optimal control problems : A review of some theoretical and numerical aspects. Notes de cours, <http://www-rocq.inria.fr/sydoco/cours/tutorial.html>, INRIA, 2002.
- [6] S. Boyd and L. Vandenberghe. *Convex optimization*. Cambridge University Press, 2004.
- [7] A. E. Bryson. *Dynamic Optimization*. Addison-Wesley, 1999.
- [8] A. E. Bryson and Y. C. Ho. *Applied Optimal Control*. Ginn and Company, 1969.
- [9] R. Courant and D. Hilbert. *Methods of Mathematical Physics*, volume 2. Interscience, 1962.
- [10] J.-Ch. Culioli. *Introduction à l'optimisation*. Ellipses, 1994.
- [11] S. E. Dreyfus and A. M. Law. *The art and theory of dynamic programming*. Academic Press, 1977.
- [12] A. R. Forsyth. *Calculus of variations*. Cambridge at the University Press, 1927.
- [13] P. E. Gill, W. Murray, and M. A. Saunders. *Large-scale SQP Methods and their Application in Trajectory Optimization*, chapter 1. Birkhäuser Verlag, 1994.

- [14] P. E. Gill, W. Murray, M. A. Saunders, and M. A. Wright. *User's Guide for NP-SOL 5.0 : A Fortran Package for Nonlinear Programming*. Systems Optimization Laboratory, Stanford University, Stanford, CA 94305, 1998.
- [15] J.-B. Hiriart-Urruty. *Optimisation et analyse convexe*. Presses Universitaires de France, 1998.
- [16] L. P. Lebedev and M. J. Cloud. *The calculus of variations and functional analysis*, volume 12 of *Series on stability, vibration and control of systems*. World Scientific Publishing Co., 2003.
- [17] D. G. Luenberger. *Optimization by vector spaces methods*. Wiley, 1969.
- [18] Mathworks. Optimization toolbox. User's guide, <http://www.mathworks.com>, The Math Works Inc., 2002.
- [19] M. Minoux. *Programmation Mathématique. Théorie et algorithmes*, volume 1, 2. Dunod, 1983.
- [20] J. J. Moré and S. J. Wright. *Optimization Software guide*. Frontiers in Applied Mathematics. SIAM, 1993.
- [21] J. Nocedal and S. J. Wright. *Numerical Optimization*. Springer Series in Operations Research. Springer, 1999.
- [22] E. Polak. *Optimization - Algorithms and Consistent Approximations*. Springer-Verlag, 1998.
- [23] L. Pontryagin and et al. *Théorie Mathématique de Processus Optimaux*. Editions Mir, Moscou, 1974.
- [24] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling. *Numerical Recipes : The Art of Scientific Computing*. New York :Cambridge University Press, 1986.
- [25] S. M. Roberts and J. S. Shipman. *Two-point boundary value problems : shooting methods*. American Elsevier, 1972.
- [26] H. J. Sussmann and J. C. Willems. 300 years of optimal control : from the brachystochrone to the maximum principle. *IEEE Control Systems Magazine*, pages 32–44, 1997. disponible sur <http://www.math.rutgers.edu/~sussmann>.
- [27] V. M. Tikhomirov. *Stories about Maxima and Minima*. Mathematical World. Volume 1. American Mathematical Society, 1990.

Notes. — On pourra se référer à [20] pour un exposé de certaines librairies et de certains logiciels existants pour la résolution des problèmes d'optimisation. On pourra se reporter à [3] pour tout détails portant sur le chapitre 1 du cours. En ce qui concerne l'analyse et l'optimization convexe on pourra se reporter à [15] et [6]. On trouvera de plus amples détails sur les algorithmes de descente en dimension finie dans [22]. Pour le chapitre 2 on pourra trouver tous les détails concernant l'optimisation de fonctionnelles dans [17] et les différentes conditions de stationnarité dans [8, 7, 12]. On trouvera les détails sur l'histoire du calcul des variations dans [26, 1]. Pour les équations variationnelles d'ordre élevé on pourra se reporter à [9]. Pour une présentation en lien avec l'analyse fonctionnelle, on se reportera à [16]. Pour une vue d'ensemble on pourra se reporter à [2]. Pour une présentation des techniques de programmation dynamique on pourra consulter [4, 11]. L'ouvrage [10] contient une présentation très complète de l'optimisation.