

Méthodes de points intérieurs pour des problèmes de grande taille

Saïd MOUSSAOUI, Jérôme IDIER, Maxime LEGENDRE

Ecole Centrale Nantes, IRCCyN, CNRS UMR 6597, 1, rue de la Noë, BP 92101, F-44321 Nantes Cedex 03, France

prenom.nom@irccyn.ec-nantes.fr

Résumé – L’objectif de cette communication est de fournir un aperçu de quelques contributions récentes en optimisation numérique dans le cadre de la résolution de problèmes inverses de grande taille en traitement du signal et de l’image. On s’intéressera plus particulièrement aux méthodes primales-duales de points intérieurs qui sont connues pour leur convergence rapide au prix d’un coût de calcul important par itération. Les améliorations proposées concernent la réduction du coût de calcul par itération en assurant la convergence des méthodes et conduisant à des algorithmes possédant un potentiel de parallélisation. Cette démarche sera illustrée à travers un exemple de traitement d’images hyperspectrales.

Abstract – The aim of this communication is to present some recent contributions in numerical optimization for the solving of large-scale signal and image processing problems. We will make a focus on primal-dual interior-point methods which are known for their fast convergence rate and their high numerical cost for each iteration. The proposed enhancements concern the development of acceleration strategies allowing to reduce this cost, while maintaining the convergence properties of the method and yielding algorithmic structures suitable for parallel implementation. This strategy is illustrated on an example of hyperspectral image unmixing.

1 Introduction

La résolution de nombreux problèmes en traitement du signal et de l’image fait appel à une étape d’optimisation d’un critère sous des contraintes d’égalité et/ou d’inégalité. Bien que plusieurs méthodes d’optimisation existent dans les domaines de l’analyse numérique et des mathématiques appliquées, le choix de la méthode la plus adaptée ou son adaptation à un problème de traitement du signal ou d’image n’est pas trivial. En effet, ce choix doit tenir compte des propriétés du critère, du type de contraintes et de la dimension du problème. Dans le cadre de problèmes de grande taille, une attention particulière doit être focalisée sur la complexité arithmétique et le besoin mémoire des algorithmes.

Il existe principalement trois grandes familles de méthodes d’optimisation sous contraintes [1] : les méthodes de *pénalité extérieure*, les méthodes de *contraintes actives* et les méthodes de *points-intérieurs*. On s’intéressera particulièrement à cette dernière famille de méthodes. Les méthodes de *points-intérieurs* (qualifiées aussi de méthodes de *pénalité intérieure*) ont la spécificité de garantir le respect des contraintes d’inégalité *strictement* (d’où le qualificatif *intérieurs*).

Le principe d’optimisation par points intérieurs est apparu dans les années cinquante, grâce à la définition de la fonction barrière logarithmique, en 1955, par Frisch [2]. C’est dans le livre de Fiacco et McCormick [3] que le terme de points intérieurs a été introduit. Les travaux de Karmakar [4] en 1984 furent à l’origine de la proposition d’un algorithme à convergence polynomiale, ce qui a ouvert la voie au développement de plusieurs techniques [1], telles que *le suivi de chemin central*, la *barrière logarithmique* et la méthode *primal-duale*.

D’une façon générale, pour la construction des algorithmes

de points-intérieurs, il y a deux points de vue complémentaires qui conduisent au même résultat : *la pénalisation logarithmique* et *la perturbation* des conditions d’optimalité. L’approche primale-duale consiste à estimer conjointement les variables primales (variables d’intérêt) et duales (multiplicateurs de Lagrange) par la résolution d’une séquence de problèmes correspondants à des versions perturbées des conditions d’optimalité de Karush-Kuhn-Tucker (KKT)[5]. Ces méthodes possèdent une vitesse de convergence superlinéaire, voire quadratique [6] au prix d’un coût de calcul important par itération, limitant leur application dans le cas de problèmes de grande taille.

Nous présentons dans cette communication quelques améliorations récentes permettant de réduire ce coût de calcul en exploitant la structure du critère à minimiser et en proposant des adaptations conduisant à une structure algorithmique adaptée pour une implémentation efficace sur des outils de calcul parallèle.

2 Optimisation sous contraintes

On s’intéresse à la résolution d’un problème d’optimisation exprimée sous la forme

$$\begin{aligned} & \underset{\mathbf{x} \in \mathbb{R}^n}{\text{minimiser}} && f_0(\mathbf{x}), \\ & \text{sous contraintes} && f_i(\mathbf{x}) \leq 0 \quad \forall i = 1, \dots, m_i, \\ & && h_j(\mathbf{x}) = 0 \quad \forall j = 1, \dots, m_e. \end{aligned} \quad (1)$$

Ce problème correspond à la minimisation d’une fonction f_0 ($\mathbb{R}^n \mapsto \mathbb{R}$) sous m_e contraintes d’inégalité et m_i contraintes d’égalité, exprimées par les fonctions $\{f_i, i = 1, \dots, m_i\}$ et $\{h_j, j = 1, \dots, m_e\}$, respectivement. Le lagrangien associé à ce *problème primal* s’obtient en augmentant la fonction objectif

f_0 par un terme pondéré par les contraintes selon

$$L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\nu}) = f_0(\mathbf{x}) + \boldsymbol{\lambda}^\top \mathbf{f}(\mathbf{x}) + \boldsymbol{\nu}^\top \mathbf{h}(\mathbf{x}),$$

où $\boldsymbol{\lambda} \in \mathbb{R}^{m_i}$, $\boldsymbol{\nu} \in \mathbb{R}^{m_e}$ sont des vecteurs contenant les *multiplieurs de Lagrange* associés aux contraintes $\mathbf{f} = [f_1, \dots, f_{m_i}]^\top$ et $\mathbf{h} = [h_j, \dots, h_{m_e}]^\top$.

La fonction duale associée au problème d'optimisation initial est définie par

$$g(\boldsymbol{\lambda}, \boldsymbol{\nu}) = \inf_{\mathbf{x} \in \mathcal{D}} L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\nu}), \quad (2)$$

avec $\mathcal{D} = \bigcap_{i=1}^{m_i} \text{dom } f_i \cap \bigcap_{j=1}^{m_e} \text{dom } h_j$. La maximisation de cette fonction duale, sous la contrainte $\boldsymbol{\lambda} \geq 0$, est appelé *problème dual* et conduit au point $d^* = \max_{\boldsymbol{\lambda} \geq 0, \boldsymbol{\nu}} g(\boldsymbol{\lambda}, \boldsymbol{\nu})$. En cas de dualité forte (i.e. existence d'un minimiseur $\mathbf{x}^* \in \mathcal{D}$), une conséquence importante est que $d^* = p^* = \min_{\mathbf{x} \in \mathcal{C}} f_0$, avec $\mathcal{C} = \{\mathbf{x} | \mathbf{f}(\mathbf{x}) \leq 0 \text{ et } \mathbf{h}(\mathbf{x}) = 0\}$.

2.1 Conditions d'optimalité

Les conditions nécessaires pour l'optimalité de la solution $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\nu}^*)$ du problème d'optimisation (1), conditions dites de Karush-Kuhn-Tucker, sont liées à l'annulation du gradient du lagrangien par rapport à \mathbf{x} , la non-négativité des paramètres de Lagrange, le respect des contraintes d'inégalité et d'égalité ainsi que la satisfaction de la *condition de complémentarité*. Cette dernière condition stipule qu'au point optimal chaque paramètre de Lagrange associé à une contrainte d'inégalité vaut zéro sauf si cette contrainte est active. Ces conditions s'expriment sous la forme d'un système d'équations

$$\begin{cases} \nabla f_0(\mathbf{x}^*) + \nabla \mathbf{f}(\mathbf{x}^*) \boldsymbol{\lambda}^* + \nabla \mathbf{h}(\mathbf{x}^*) \boldsymbol{\nu}^* = \mathbf{0}, \\ \mathbf{h}(\mathbf{x}^*) = \mathbf{0}, \\ \mathbf{f}(\mathbf{x}^*) \leq \mathbf{0}, \\ \boldsymbol{\lambda}^* \geq \mathbf{0}, \\ \mathbf{f}(\mathbf{x}^*)^\top \boldsymbol{\lambda}^* = 0, \end{cases} \quad (3)$$

La condition de complémentarité est difficile à résoudre car elle nécessite de connaître au préalable l'ensemble des contraintes actives au point optimal. Pour contourner cette difficulté, l'idée des méthodes de points intérieurs est de relaxer partiellement cette condition en s'appuyant sur un paramètre de perturbation μ et de résoudre des problèmes intermédiaires de sorte à maintenir les solutions intermédiaires à l'intérieur du domaine admissible :

$$\begin{cases} \nabla f_0(\mathbf{x}^*) + \nabla \mathbf{f}(\mathbf{x}^*) \boldsymbol{\lambda}^* + \nabla \mathbf{h}(\mathbf{x}^*) \boldsymbol{\nu}^* = \mathbf{0}, \\ \mathbf{h}(\mathbf{x}^*) = \mathbf{0}, \\ \mathbf{f}(\mathbf{x}^*) < \mathbf{0}, \\ \boldsymbol{\lambda}^* > \mathbf{0}, \\ -\mathbf{F}(\mathbf{x}^*) \boldsymbol{\lambda}^* = \mu \mathbf{1}_{m_i} \end{cases} \quad (4)$$

où $\mathbf{F}(\mathbf{x}) = \text{Diag}(\mathbf{f}(\mathbf{x}))$. Par décroissance du paramètre de perturbation μ vers 0, les solutions convergent vers la solution du système KKT non perturbé, et donc du problème initial.

On montre assez facilement que les conditions KKT perturbées correspondent aux conditions d'optimalité associées à la minimisation du critère

$$f'_\mu(\mathbf{x}) = f_0(\mathbf{x}) - \mu \sum_{i=1}^{m_i} \log(-f_i(\mathbf{x})),$$

sous les contraintes $\mathbf{h}(\mathbf{x}) = \mathbf{0}$. Ainsi, l'ajout de la barrière logarithmique va permettre de garder la solution du problème à l'intérieur du domaine des contraintes.

2.2 Algorithme primal-dual

Afin d'alléger la présentation, on se place dans le cadre de la résolution d'un problème d'optimisation sous contrainte d'inégalités uniquement. L'algorithme primal-dual des points intérieurs consiste à alterner des étapes de résolution du système (4), dit de KKT perturbé, et de décroissance du paramètre de perturbation. Une solution approchée de (4) est obtenue par quelques itérations d'un algorithme de Newton couplées avec une recherche de pas selon le schéma de descente linéaire

$$(\mathbf{x}_{k+1}, \boldsymbol{\lambda}_{k+1}) = (\mathbf{x}_k + \alpha_k \mathbf{d}_k^x, \boldsymbol{\lambda}_k + \alpha_k \mathbf{d}_k^\lambda) \quad (5)$$

avec des directions primale \mathbf{d}_k^x et duale \mathbf{d}_k^λ qui correspondent à un pas de Newton appliqué aux conditions d'égalité du système KKT perturbé [5]. Ces directions sont calculées en résolvant le système linéaire

$$\begin{bmatrix} \nabla^2 L(\mathbf{x}_k, \boldsymbol{\lambda}_k) & \nabla \mathbf{f}(\mathbf{x}_k) \\ -\boldsymbol{\Lambda}_k \nabla \mathbf{f}(\mathbf{x}_k)^\top & -\mathbf{F}(\mathbf{x}_k) \end{bmatrix} \begin{bmatrix} \mathbf{d}_k^x \\ \mathbf{d}_k^\lambda \end{bmatrix} = -\mathbf{r}_{\mu_k}(\mathbf{x}_k, \boldsymbol{\lambda}_k), \quad (6)$$

où $\boldsymbol{\Lambda}_k = \text{Diag}(\boldsymbol{\lambda}_k)$ et $\mathbf{r}_\mu(\mathbf{x}, \boldsymbol{\lambda})$ est le résidu primal-dual,

$$\mathbf{r}_\mu(\mathbf{x}, \boldsymbol{\lambda}) = \begin{bmatrix} \nabla f_0(\mathbf{x}) + \nabla \mathbf{f}(\mathbf{x}) \boldsymbol{\lambda} \\ -\mathbf{F}(\mathbf{x}) \boldsymbol{\lambda} - \boldsymbol{\mu} \end{bmatrix} = \begin{bmatrix} \mathbf{r}_\mu^{\text{prim}}(\mathbf{x}, \boldsymbol{\lambda}) \\ \mathbf{r}_\mu^{\text{dual}}(\mathbf{x}, \boldsymbol{\lambda}) \end{bmatrix}.$$

Le pas α_k est déterminé de façon à vérifier les deux contraintes d'inégalité de (4) et à assurer la convergence de l'algorithme de descente en garantissant une décroissance suffisante d'une fonction de mérite primale-duale $\Psi_\mu(\mathbf{x}, \boldsymbol{\lambda})$ liée aux conditions d'optimalité du problème [5, 7]. Une technique de rebroussement avec la règle d'Armijo est suffisante pour cette étape [1]. L'arrêt de l'algorithme de descente est régi par deux conditions liées à la petitesse de la norme des résidus primal et dual

$$\|\mathbf{r}_{\mu_k}^{\text{prim}}(\mathbf{x}_k, \boldsymbol{\lambda}_k)\|_\infty \leq \epsilon_k^{\text{prim}} \text{ et } \|\mathbf{r}_{\mu_k}^{\text{dual}}(\mathbf{x}_k, \boldsymbol{\lambda}_k)\|_1 / m_i \leq \epsilon_k^{\text{dual}}.$$

Une stratégie de choix de ϵ_k^{prim} et ϵ_k^{dual} est décrite dans [8]. Le paramètre de perturbation μ_k est mis à jour selon la règle de μ -criticité

$$\mu_k = -(\theta / m_i) \mathbf{f}(\mathbf{x}_k)^\top \boldsymbol{\lambda}_k, \quad (7)$$

où $\theta \in (0, 1)$. Les itérations de l'algorithme primal-dual sont contrôlées par un test d'arrêt global portant sur la valeur minimale de la perturbation ($\mu_k \leq \mu_{\min}$) et sur la norme du résidu primal-dual ($\|\mathbf{r}_{\mu_k}(\mathbf{x}_k, \boldsymbol{\lambda}_k)\| \leq \epsilon_0$).

Les propriétés de convergence de cette méthode primale-duale des points intérieurs dans le cas de critères fortement convexes sont données dans le théorème 5.1 de [7].

2.3 Cas des problèmes de grande taille

Dans le cadre de problèmes de grande taille, le système (6) ne peut être inversé de façon directe. De plus, celui-ci devient très mal conditionné à l'approche de la convergence de l'algorithme et notamment dès lors qu'une des contraintes est active [9]. Nous décrivons deux stratégies pouvant être adoptées pour surmonter cette difficulté [5, 10].

2.3.1 Résolution tronquée du système primal-dual

L'issue proposée dans [10, 11] consiste à résoudre ce système de façon approchée par une méthode de résolution itérative telle que le gradient conjugué préconditionné (PCG). Cette stratégie se place ainsi dans la famille des méthodes de Newton tronqué [12]. Il est montré dans [10] que cette stratégie conserve les propriétés de convergence de l'algorithme. Cependant, les performances dépendent du choix du préconditionneur et du seuil du résidu relatif de la résolution approchée. De plus, la matrice normale n'étant pas symétrique, l'utilisation d'un algorithme de gradient conjugué dans ce cas ne semble pas être la meilleure option. Une technique telle que le gradient bi-conjugué préconditionné est plus adaptée [13].

2.3.2 Calcul approché de la direction primale

Une alternative [8] consiste à effectuer le calcul des directions en deux étapes : la direction primale \mathbf{d}_k^x est d'abord obtenue par inversion du système réduit

$$\mathbf{H}_k \mathbf{d}_k^x = -\mathbf{g}_k \quad (8)$$

avec

$$\begin{cases} \mathbf{g}_k &= \nabla L(\mathbf{x}_k, \boldsymbol{\lambda}_k) - \nabla \mathbf{f}(\mathbf{x}_k) \mathbf{F}(\mathbf{x}_k)^{-1} \boldsymbol{\mu}_k, \\ \mathbf{H}_k &= \nabla^2 L(\mathbf{x}_k, \boldsymbol{\lambda}_k) - \nabla \mathbf{f}(\mathbf{x}_k) \mathbf{F}(\mathbf{x}_k)^{-1} \boldsymbol{\Lambda}_k \nabla \mathbf{f}(\mathbf{x}_k)^t. \end{cases}$$

Ensuite la direction duale \mathbf{d}_k^λ est obtenue selon

$$\mathbf{d}_k^\lambda = -\mathbf{F}(\mathbf{x}_k)^{-1} [\boldsymbol{\mu}_k + \boldsymbol{\Lambda}_k \nabla \mathbf{f}(\mathbf{x}_k)^t \mathbf{d}_k^x] - \boldsymbol{\lambda}_k. \quad (9)$$

Notons que le système réduit s'obtient par substitution de \mathbf{d}_k^λ dans la première équation de (6) par son expression (9) déduite de la seconde partie de ce système.

On peut constater que la matrice \mathbf{H}_k est symétrique et de taille plus petite que la matrice normale du système primal-dual (6). L'algorithme PCG avec une factorisation de Cholesky incomplète a été utilisé dans [13] pour la résolution approchée de ce système. Nous avons par ailleurs proposé dans [14] une approche basée sur les algorithmes de majoration-minimisation (MM). Cette approche exploitant la structure dans la matrice \mathbf{H}_k permet d'assurer la séparabilité des variables et de réaliser une résolution itérative du système en évitant les inversions matricielles et les factorisation de Cholesky. Ces étapes étant coûteuses en temps de calcul et stockage mémoire. Nous proposons d'utiliser cette stratégie MM comme préconditionneur au sein d'un algorithme PCG appliqué à la résolution du système réduit.

2.3.3 Accélération matérielle

Un second volet pour la résolution de problèmes de grande taille est lié à l'utilisation des outils de calcul intensif pour la mise en oeuvre de l'algorithme d'optimisation. Dans le cadre des méthodes de type point-intérieurs l'approximation MM séparable permet de remplacer l'étape d'inversion d'un système de grande taille par la résolution d'un grand nombre de systèmes de taille réduite [14]. Ainsi la résolution de ces systèmes sera répartie sur plusieurs processeurs de calcul parallèle [15].

3 Exemple : estimation des cartes d'abondances en imagerie hyperspectrale

Les différentes approches décrites sont illustrées à travers le problème d'estimation des cartes d'abondances en imagerie hyperspectrale [16, 17]. Une image hyperspectrale consiste en une mesure du spectre de réflectance dans K bandes spectrales pour chacun des N pixels. On suppose que l'image contient P composants dont les spectres $\mathbf{S} = [\mathbf{s}_1, \dots, \mathbf{s}_P] \in \mathbb{R}^{K \times P}$ sont connus. En utilisant le modèle de mélange linéaire, le spectre observé en un pixel n peut s'écrire $\mathbf{y}_n = \mathbf{S} \mathbf{x}_n + \mathbf{e}_n$ avec $\mathbf{y}_n \in \mathbb{R}^K$, $\mathbf{x}_n \in \mathbb{R}^P$ les coefficients de mélanges ou abondances, et $\mathbf{e}_n \in \mathbb{R}^K$ un bruit gaussien i.i.d. et de moyenne nulle.

3.1 Formulation du problème d'optimisation

L'estimation des cartes d'abondances consiste à estimer en chaque pixel les abondances \mathbf{x}_n connaissant \mathbf{y}_n et \mathbf{S} en résolvant le problème de moindres carrés contraint suivant :

$$\begin{aligned} & \underset{\mathbf{x} \in \mathbb{R}^{NP}}{\text{minimiser}} & f_0(\mathbf{x}) &= \sum_{n=1}^N \|\mathbf{y}_n - \mathbf{S} \mathbf{x}_n\|_2^2 + \beta R(\mathbf{x}), \\ & \text{sous contraintes} & \mathbf{x}_n &\geq \mathbf{0} \quad \forall n = 1, \dots, N, \\ & & \mathbf{1}_P^t \mathbf{x}_n &= 1 \quad \forall n = 1, \dots, N. \end{aligned} \quad (10)$$

où $\mathbf{x} = [\mathbf{x}_1^t, \dots, \mathbf{x}_N^t]^t \in \mathbb{R}^{NP}$ est la concaténation des vecteurs d'abondances de tous les pixels, $\beta \in \mathbb{R}$, et $R(\mathbf{x})$ est un terme de pénalisation assurant la régularité spatiale des abondances et s'écrit $R(\mathbf{x}) = \sum_{\{i,j\} \in \epsilon} \|\mathbf{x}_i - \mathbf{x}_j\|_2^2$ où ϵ représente l'ensemble des pixels voisins verticalement et horizontalement.

3.2 Résultats

Les performances des différentes approches sont évaluées sur des données simulées à partir de spectres \mathbf{S} issus de la bibliothèque spectrale USGS (mesurés dans $K = 224$ bandes spectrales) et de cartes d'abondances \mathbf{x} contenant des motifs de forme gaussienne. Chaque pixel est généré par le mélange linéaire $\mathbf{y}_n = \mathbf{S} \mathbf{x}_n$. Un terme de bruit est ajouté pour obtenir un rapport signal à bruit de 10 dB.

L'estimation des cartes d'abondances est réalisée pour chaque image simulée en résolvant le problème (10) par une méthode primale-duale de points intérieurs. Cinq variantes de l'algorithme sont testées, chacune se distinguant par le choix de la méthode de Newton tronquée pour de calcul de la direction

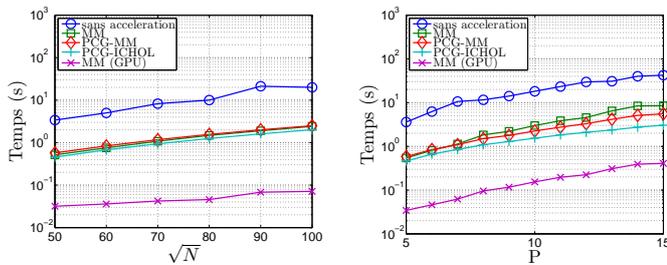


FIGURE 1 – Temps de calcul des différentes approches proposées pour différentes tailles de données : à gauche pour différents nombre de pixels et $P = 5$, à droite pour différents nombres de composants et $N = 50^2$. Valeurs moyennes pour 100 simulations de Monte Carlo.

primale : inversion directe (sans accélération), approche MM, approche MM sur GPU, approche PCG avec préconditionneur issu de l’approche MM, approche PCG avec préconditionneur issu d’une factorisation de Cholesky incomplète.

La figure 1 illustre les temps de calcul des 5 variantes pour différentes tailles de données. Les calculs sont effectués sur une station de travail Dell Precision T7400 équipée d’un processeur Intel Xeon X5472, 16 GO de mémoire RAM, et une carte graphique Tesla C1060 contenant 240 unités de calcul et 4 GO de mémoire RAM. Le paramètre de régularisation β est fixé à 1 pour tous les tests. Dans les conditions testées, les 3 alternatives pour le calcul de la direction primale sur CPU offrent un facteur d’accélération variant entre 4.8 et 14.6 par rapport à la version sans accélération. Ce facteur augmente avec la taille des données. L’implémentation GPU de l’approche MM offre un facteur d’accélération variant entre 15.8 et 33.9 par rapport à l’approche MM sur CPU, ou entre 102.5 et 312.7 par rapport à la version sans accélération sur CPU.

4 Conclusion

Bien qu’étant une méthode de second ordre, la méthode primale-duale de points intérieurs peut être employée pour résoudre des problèmes d’optimisation de grande taille grâce à l’intégration d’une étape de type Newton tronqué au sein de l’algorithme. Conjugée à l’exploitation de la structure du problème, cette amélioration permet non seulement de diminuer le coût de calcul tout en garantissant la convergence, mais elle facilite aussi une implémentation parallèle.

Références

[1] J. Nocedal et S. J. Wright, *Numerical Optimization*, Springer-Verlag, New York, NY, USA, 1999.
 [2] K. R. Frisch, « The logarithmic potential method of convex programming », Memorandum, University Institute of Economics, Oslo, Norway, may 1955.
 [3] A. V. Fiacco et G. P. McCormick, *Nonlinear programming : Sequential unconstrained minimization techniques*, John Wiley and Sons, Inc., 1968.

[4] N. K. Karmarkar, « A new polynomial-time algorithm for linear programming », *Combinatorica*, vol. 4, pp. 373–395, 1984.
 [5] A. Forsgren, P. Gill et M. Wright, « Interior methods for nonlinear optimization », *SIAM Review*, vol. 44, n°4, pp. 525–597, 2002.
 [6] H. Yamashita et H. Yabe, « Superlinear and quadratic convergence of some primal-dual interior point methods for constrained optimization », *Mathematical Programming*, vol. 75, pp. 377–397, 1996.
 [7] P. Armand, J. C. Gilbert et S. Jan-Jégou, « A feasible BFGS interior point algorithm for solving strongly convex minimization problems », *SIAM Journal on Optimization*, vol. 11, pp. 199–222, 2000.
 [8] A. Conn, N. Gould et P. L. Toint, « A primal-dual algorithm for minimizing a nonconvex function subject to bounds and nonlinear constraints », in *Nonlinear Optimization and Applications*, G. Di Pillo et F. Giannesi, Eds. Kluwer Academic Publishers, 2 edition, 1996.
 [9] M. H. Wright, « Ill-conditioning and computational error in interior methods for nonlinear programming », *SIAM Journal on Optimization*, vol. 9, n°1, pp. 84–111, 1998.
 [10] P. Armand, J. Benoist et J.-P. Dussault, « Local path-following property of inexact interior methods in nonlinear programming », *Computational Optimization and Applications*, vol. 52, n°1, pp. 209–238, 2012.
 [11] S. Bonettini, E. Galligani et V. Ruggiero, « Inner solvers for interior point methods for large scale nonlinear programming », *Computational Optimization and Applications*, vol. 37, n°1, pp. 1–34, mai 2007.
 [12] R. S. Dembo et T. Steihaug, « Truncated-newton methods algorithms for large scale unconstrained optimization », *Mathematical Programming*, vol. 26, pp. 190–212, 1983.
 [13] E. Chouzenoux, S. Moussaoui, M. Legendre et J. Idier, « Algorithme primal-dual de points intérieurs pour l’estimation pénalisée des cartes d’abondances en imagerie hyperspectrale », *Traitement du Signal*, vol. 30, n°1-2, pp. 35–59, 2013.
 [14] M. Legendre, S. Moussaoui, E. Chouzenoux et J. Idier, « Primal-dual interior-point optimization based on majorization-minimization for edge-preserving spectral unmixing », in *Proc. of IEEE International Conference on Image Processing*, Paris, France, octobre 2014.
 [15] M. Legendre, S. Moussaoui, F. Schmidt et J. Idier, « Parallel implementation of a primal-dual interior-point optimization method for fast abundance maps estimation », in *Proc. of IEEE Workshop on Hyperspectral Image and Signal Processing : Evolution in Remote Sensing (WHISPERS)*, Gainesville, Etats-Unis, 2013.
 [16] J. M. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, D. Qian, P. Gader et J. Chanussot, « Hyperspectral unmixing overview : Geometrical, statistical, and sparse regression-based approaches », *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 5, n°2, pp. 354–379, April 2012.
 [17] E. Chouzenoux, M. Legendre, S. Moussaoui et J. Idier, « Fast constrained least squares spectral unmixing using primal-dual interior-point optimization », *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, n°1, pp. 59–69, 2014.