

Recherche de pas par Majoration-Minimisation pour les fonctions à barrières logarithmiques

Emilie Chouzenoux

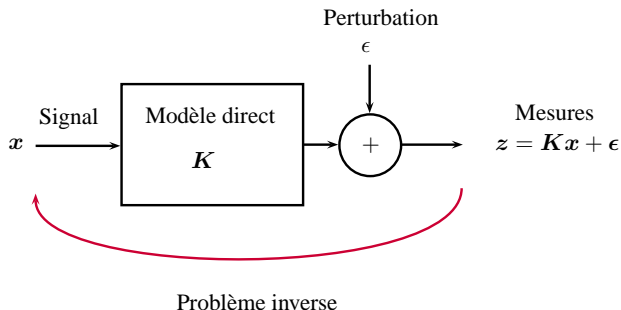
`emilie.chouzenoux@irccyn.ec-nantes.fr`

Institut de Recherche en Communication et Cybernétique de Nantes (FRANCE)

Ecole de Printemps, Porquerolles, 5 mai 2010



Problème inverse



- Rechercher x minimisant un critère $J(x)$
- J : Adéquation aux données + *a priori*
- Minimisation dans \mathbb{R}^N ou sous contraintes.

Algorithmes de Majoration-Minimisation (MM) [Hunter04]

Objectif: Trouver x , minimiseur de J

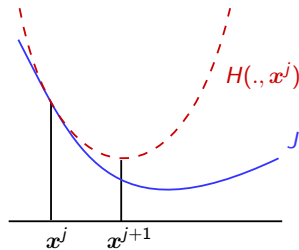
Pour tout y , soit $H(., y)$ une *Approximation Majorante* (AM) de J en y i.e.,

$$H(x, y) \geq J(x), \quad \forall x,$$

$$H(y, y) = J(y)$$

Algorithme MM :

$$x^{j+1} = \arg \min_x H(x, x^j)$$



Approximation Majorante Quadratique (AMQ) :

$$H(\mathbf{x}, \mathbf{y}) = J(\mathbf{y}) + \nabla J(\mathbf{y})^T (\mathbf{x} - \mathbf{y}) + \frac{1}{2} (\mathbf{x} - \mathbf{y})^T \mathbf{A}(\mathbf{y}) (\mathbf{x} - \mathbf{y})$$

avec $\mathbf{A}(\cdot)$ définie positive telle que $H(\cdot, \mathbf{y})$ soit une AM de J en \mathbf{y} .

Recherche de pas MM quadratique [Labat08] :

- Minimisation de $J(\cdot)$ par le schéma $\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{d}_k$
- \mathbf{d}_k direction de descente (Gradient, Newton,...)
- $\alpha_k > 0$ pas obtenu par minimisation de $f(\alpha) = J(\mathbf{x}_k + \alpha \mathbf{d}_k)$

$h(\alpha, \alpha_k^j)$: AMQ de f en α_k^j de courbure

$$\mathbf{a}_k^j = \mathbf{d}_k \mathbf{A}(\mathbf{x}_k + \alpha_k^j \mathbf{d}_k) \mathbf{d}_k$$

Minimisation par MM scalaire:

$$\alpha_k^{j+1} = \arg \min_{\alpha} h(\alpha, \alpha_k^j) = \alpha_k^j - \dot{f}(\alpha_k^j) / \mathbf{a}_k^j.$$

PLAN

- 1 Position du problème
- 2 Recherche de pas par approximation majorante
- 3 Applications
- 4 Conclusion

Fonction barrière

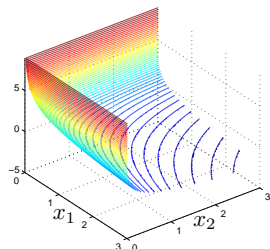
Définition

B strictement convexe est appelée *fonction barrière* associée aux contraintes $\mathbf{x} \in \mathcal{C}$ si B est non bornée à la frontière de \mathcal{C}

Ex : Barrière logarithmique pour les contraintes de positivité

$$B(\mathbf{x}) = - \sum_n \log(x_n)$$

$$\mathcal{C} = \{\mathbf{x} > 0\}$$



$$B(\mathbf{x}) = -\log(x_1) - \log(x_2)$$

⇒ Si un critère contient une fonction barrière, ses minimiseurs vont appartenir au domaine \mathcal{C} .

Critères contenant des fonctions barrières

- ① *Tomographie d'émission : Modèle de bruit Poisson*

$$J(\mathbf{x}) = \sum_m [\mathbf{K}\mathbf{x}]_m - z_m \log[\mathbf{K}\mathbf{x}]_m$$

- ② *Contrainte de positivité :*

$$\min_{\mathbf{x} \geq 0} \|\mathbf{K}\mathbf{x} - \mathbf{z}\|^2 \Leftrightarrow \min J(\mathbf{x}) = \|\mathbf{K}\mathbf{x} - \mathbf{z}\|^2 - \mu \sum_n \log(x_n), \{\mu\} \rightarrow 0$$

- ③ *Résonance magnétique nucléaire : Pénalisation par maximum d'entropie*

$$J(\mathbf{x}) = \|\mathbf{K}\mathbf{x} - \mathbf{z}\|^2 + \lambda \sum_n x_n \log x_n$$

Formulation générale du problème

Problèmes d'optimisation de la forme :

$$\min (J(x) = P(x) + \mu B(x)), \quad \mu > 0 \quad (1)$$

- $P(x)$: Possibilité de construire des AM quadratiques
- $B(x) = \sum_{i=1}^l b_i(c_i^T x + \rho_i)$: fonction barrière,
e.g., $b_i(u) = -\log u$ ou $u \log u$

Algorithme itératif

$$x_{k+1} = x_k + \alpha_k d_k, \quad \text{pour } k = 1, \dots, K$$

- d_k *direction de descente* (Gradient, Newton, Newton tronqué, gradient conjugué ...)
- $\alpha_k > 0$ *pas obtenu par minimisation de* $f(\alpha) = J(x_k + \alpha d_k)$.

Problématique

$$f(\alpha) = J(\mathbf{x}_k + \alpha \mathbf{d}_k) = P(\mathbf{x}_k + \alpha \mathbf{d}_k) + \underbrace{\mu B(\mathbf{x}_k + \alpha \mathbf{d}_k)}_{\text{barrière } b(\alpha)}$$

b non définie pour $\alpha > \bar{\alpha}$ si il existe i tel que $\mathbf{c}_i^T(\mathbf{x}_k + \bar{\alpha} \mathbf{d}_k) + \rho_i = 0$.

- Restriction de la recherche de pas à $\alpha \in [0, \bar{\alpha})$
- Prise en compte de l'asymptote verticale en $\bar{\alpha}$
- Méthodes classiques peu efficaces

PLAN

- 1 Position du problème
- 2 Recherche de pas par approximation majorante
- 3 Applications
- 4 Conclusion

Rappel : Recherche de pas MM quadratique

Objectif : trouver α_k minimisant $f(\alpha) = J(\mathbf{x}_k + \alpha \mathbf{d}_k)$

Pour tout β , on construit $h(\cdot, \beta)$ une AMQ de f en β :

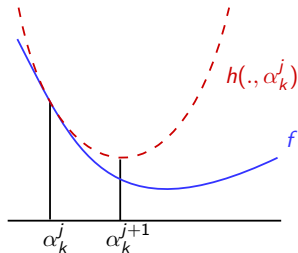
$$h(\alpha, \beta) = f(\beta) + (\alpha - \beta)\dot{f}(\beta) + \frac{1}{2}a_\beta(\alpha - \beta)^2$$

Pas issu de l'AMQ :

$$\alpha_k^{j+1} = \arg \min_{\alpha} h(\alpha, \alpha_k^j), j < J$$

$$= \alpha_k^j - \dot{f}(\alpha_k^j) / a_k^j, j < J$$

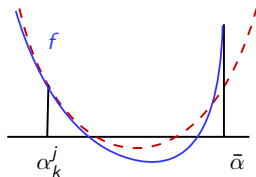
$$\alpha_k = \alpha_k^J$$



Recherche de pas MM non quadratique

La courbure du terme de barrière est **non bornée** en $\bar{\alpha}$

\Rightarrow Il n'existe pas d'AMQ de f .



Construction d'approximations majorantes de f de la forme

$$h(\alpha, \alpha_k^j) = p_0 + p_1\alpha + p_2\alpha^2 - p_3 \log(\bar{\alpha} - \alpha)$$

- Choix des p_i pour les barrières $-\log$ et “entropiques”
- $\arg \min_{\alpha} h(\alpha, \alpha_k^j)$ est la racine d'un polynôme d'ordre 2

Toy Example

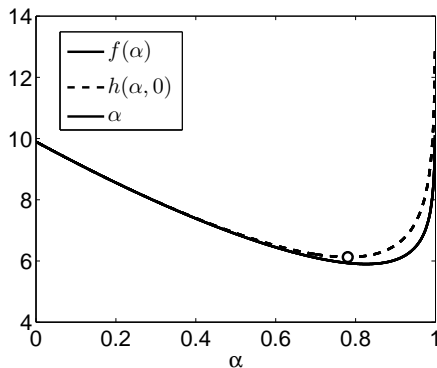
$$f(\alpha) = (\alpha - 5)^2 - \sum_{i=1}^{10} \log(i - \alpha)$$

$$\bar{\alpha} = 1$$

$$m^0 = 2, \gamma^0 = 1.55$$

$$\alpha^1 = 0.7805$$

à comparer avec $\alpha^* = 0.8258$



Analyse de convergence [Chouzenoux09]

Objectif : Etudier la convergence du schéma itératif

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{d}_k, \quad k = 1, \dots, K$$

Propriétés

Quelque soit le nombre de sous-itérations MM,

- Les conditions d'Armijo et de Zoutendijk sont vérifiées
- Le pas est 'suffisamment grand'

⇒ Convergence de nombreux algorithmes démontrée :

Gradient, Gradient conjugué, Newton, Quasi-Newton ...

PLAN

- 1 Position du problème
- 2 Recherche de pas par approximation majorante
- 3 Applications**
- 4 Conclusion

Résonance Magnétique Nucléaire 1D

Modèle direct

$$z(\tau) = \int_{T_{\min}}^{T_{\max}} e^{-\tau/T} x(T) dT$$

T : Temps de relaxation

τ : Temps d'écho

$z(\tau)$: Echo mesuré

$x(T)$: Spectre à estimer

Après discrétisation,

$$z = \mathbf{K}x + \epsilon$$

avec $\mathbf{K} \in \mathbb{R}^{M \times N}$, $z \in \mathbb{R}^M$, $x \in \mathbb{R}^N$

Résonance Magnétique Nucléaire 1D

Modèle direct

$$z(\tau) = \int_{T_{\min}}^{T_{\max}} e^{-\tau/T} x(T) dT$$

T : Temps de relaxation

τ : Temps d'écho

$z(\tau)$: Echo mesuré

$x(T)$: Spectre à estimer

Après discrétisation,

$$z = Kx + \epsilon$$

avec $K \in \mathbb{R}^{M \times N}$, $z \in \mathbb{R}^M$, $x \in \mathbb{R}^N$

Objectif : Estimer x à partir de z sous la contrainte $x \geq 0$

Résonance Magnétique Nucléaire 1D

Reconstruction par maximum d'entropie

$$\min_{\mathbf{x} \geq 0} J(\mathbf{x}) = P(\mathbf{x}) + \mu B(\mathbf{x})$$

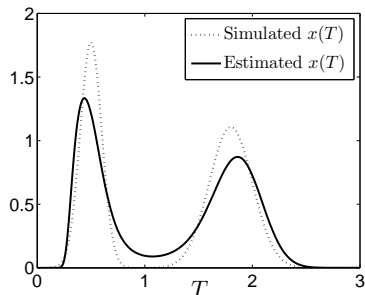
- Adéquation aux données : Moindres carrés $P(\mathbf{x}) = \|\mathbf{K}\mathbf{x} - \mathbf{y}\|^2$
- Régularisation : Entropie de Shannon $B(\mathbf{x}) = \sum_n x_n \ln x_n$
Fonction barrière \Rightarrow Assure la positivité

Stratégie d'optimisation

- Algorithme de Newton tronqué
- Préconditionnement basé sur la SVD
- Comparaison de la recherche de pas de Wolfe [Moré94] avec la stratégie MM proposée.

Résonance Magnétique Nucléaire 1D

MT	c_1	c_2	<i>Iter</i>	<i>Tps(s)</i>
	10^{-3}	0.5	34	12
	10^{-3}	0.9	42	13
MM	10^{-3}	0.99	71	20
	<i>J</i>		<i>Iter</i>	<i>Tps(s)</i>
	1		36	8
	2		40	9
	5		40	10
10		40	14	



(c_1, c_2) : Paramètres de Wolfe
J: Nombre de sous-itérations MM

Figure: Reconstruction RMN de données simulées avec $RSB = 25dB$

Points intérieurs pour la programmation quadratique

Problème quadratique contraint

$$\begin{aligned} \min_{\mathbf{x} \in \mathbb{R}^N} J_0(\mathbf{x}) &= \rho_0 + \mathbf{c}_0^T \mathbf{x} + \frac{1}{2} \mathbf{x}^T \mathbf{C}_0 \mathbf{x} \\ \text{s.t. : } C_i(\mathbf{x}) &= \mathbf{c}_i^T \mathbf{x} + \rho_i \geq 0, \quad 1 \leq i \leq l \end{aligned}$$

Critère augmenté

$$J_\mu(\mathbf{x}) = J_0(\mathbf{x}) - \mu \sum_{i=1}^l \log(C_i(\mathbf{x}))$$

Points intérieurs: Résoudre $\arg \min J_\mu$ pour une suite $\{\mu\} \rightarrow 0$.

Points intérieurs pour la programmation quadratique

Algorithme de points intérieurs [Nesterov94]

- 1) Fixer $\mu = 1$, des tolérances ϵ, τ et \mathbf{x} tel que $C_i(\mathbf{x}) > 0$
- 2) **TANTQUE** $|\mathbf{d}^T \nabla J_\mu| > \tau$,
Calculer la direction de Newton \mathbf{d} de J_μ en \mathbf{x}
Déterminer le pas α
Mettre à jour $\mathbf{x} \leftarrow \mathbf{x} + \alpha \mathbf{d}$
- 3) **SI** $\mu < \epsilon$, **FIN**
SINON Diminuer μ et **RETOUR** à l'étape 2.

Points intérieurs pour la programmation quadratique

Algorithme de points intérieurs [Nesterov94]

- 1) Fixer $\mu = 1$, des tolérances ϵ, τ et \mathbf{x} tel que $C_i(\mathbf{x}) > 0$
- 2) **TANTQUE** $|\mathbf{d}^T \nabla J_\mu| > \tau$,
Calculer la direction de Newton \mathbf{d} de J_μ en \mathbf{x}
Déterminer le pas α
Mettre à jour $\mathbf{x} \leftarrow \mathbf{x} + \alpha \mathbf{d}$
- 3) **SI** $\mu < \epsilon$, **FIN**
SINON Diminuer μ et **RETOUR** à l'étape 2.

Résultats

100 problèmes aléatoires de taille $N = 400$, $l = 200$

Rebroussement	151 iter	13.6 s
Méthode de [Nesterov94]	310 iter	19.6 s
Méthode proposée	75.9 iter	5 s

PLAN

- 1 Position du problème
- 2 Recherche de pas par approximation majorante
- 3 Applications
- 4 Conclusion**

Conclusion

Principale contribution

Recherche de pas MM pour les critères contenant une fonction barrière

- Méthode simple
- Résultats de convergence
- Efficacité pratique

Application en reconstruction RMN 2D [*Chouzenoux2010*]

Perspectives

- Contraintes non linéaires
- Autres fonctions barrières
- AM multi-dimensionnelle

Bibliographie



P. D. Bertsekas
Nonlinear Programming.
Athena Scientific, 1999.



Y. Nesterov and A. Nemirovskii
Interior point polynomial algorithms in convex programming.
Studies in Applied and Numerical Mathematics, 1994.



E. Chouzenoux and S. Moussaoui and J. Idier
A Majorize-Minimize line search algorithm for barrier function optimization
Soumis à Computational Optimization and Applications. Novembre 2009
<http://hal.archives-ouvertes.fr/hal-00362304/fr/>



E. Chouzenoux and S. Moussaoui and J. Idier and F. Mariette
Efficient maximum entropy reconstruction of nuclear magnetic resonance T1-T2 spectra
Soumis à IEEE Signal Processing. Mars 2010
<http://hal.archives-ouvertes.fr/hal-00455477/fr/>



C. Labat and J. Idier.
Convergence of conjugate gradient methods with a closed-form stepsize formula.
Journal of Optimization Theory and Applications, 136(1):43–60, 2008.



J. J. Moré and D. J. Thuente.
Line search algorithms with guaranteed sufficient decrease.
ACM Transactions on Mathematical Software, 20(3):286–307, 1994.