

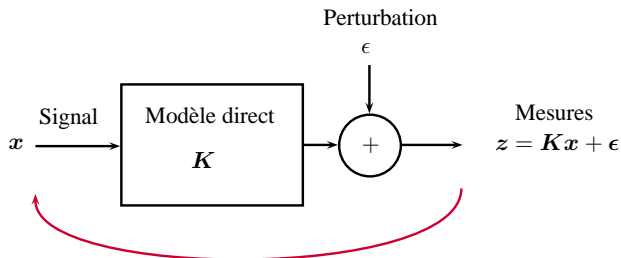
Algorithmes de Majoration-Minimisation pour l'optimisation sous contraintes de positivité

Emilie Chouzenoux - Saïd Moussaoui - Jérôme Idier
`emilie.chouzenoux@irccyn.ec-nantes.fr`

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

Journée GdR ISIS *Contraintes de nonnégalivité en TSI*

Problème inverse



Problème inverse

- Rechercher $x \in \mathbb{R}^N$ minimisant un critère $F(x)$
- F : Adéquation aux données + *a priori*
- Minimisation **sous contraintes de positivité**
- Ex : Reconstruction d'image

Optimisation sous contraintes de positivité

Résolution itérative de

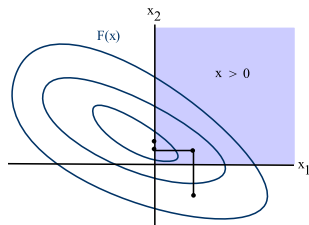
$$\min_{\mathbf{x} \geq 0} F(\mathbf{x}).$$

Deux familles d'algorithmes :

Activation des contraintes

Projection des itérés dans

$$\{\mathbf{x} | x_n \geq 0\}$$

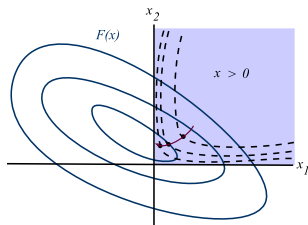


Ex : Minimisation coordonnée par coordonnée

Points intérieurs

Maintien des itérés dans

$$\{\mathbf{x} | x_n > 0\}$$



Ex : Barrière logarithmique

Algorithmes de Majoration-Minimisation (MM) [Hunter04]

Objectif: Trouver \mathbf{x} , minimiseur de F

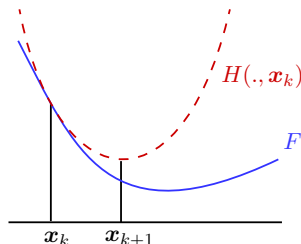
Pour tout \mathbf{y} , soit $H(\cdot, \mathbf{y})$ une *Approximation Majorante* (AM) de F en \mathbf{y} i.e.,

$$H(\mathbf{x}, \mathbf{y}) \geq F(\mathbf{x}), \quad \forall \mathbf{x},$$

$$H(\mathbf{y}, \mathbf{y}) = F(\mathbf{y})$$

Algorithme MM :

$$\mathbf{x}_{k+1} = \arg \min_{\mathbf{x}} H(\mathbf{x}, \mathbf{x}_k)$$



PLAN

- 1 Activation des contraintes
- 2 Points intérieurs
- 3 Minimisation d'un critère avec barrière
- 4 Applications
- 5 Conclusion

PLAN

- 1 Activation des contraintes
- 2 Points intérieurs
- 3 Minimisation d'un critère avec barrière
- 4 Applications
- 5 Conclusion

Algorithme MM contraint

$$\mathbf{x}_{k+1} = \arg \min_{\mathbb{R}^{N+}} H(\mathbf{x}, \mathbf{x}_k).$$

AM séparable = Algorithme SPS [DePierro95]

$$H(\mathbf{x}, \mathbf{x}_k) = \sum_n h_n(x_n, \mathbf{x}_k) \Rightarrow x_{n,k+1} = \arg \min_{\mathbb{R}^+} h_n(x_n, \mathbf{x}_k), \forall n.$$

- Construction de la majorante en utilisant la convexité [Lange95, DePierro95]
- Application aux critères non différentiables [Figueiredo07]
- Algorithmes proximaux [Combettes05, Beck09].

AM non séparable

- Minimisation coordonnée par coordonnée
= Algorithme PSCD [Erdogan99].

Minimisation coordonnée par coordonnée

Mise à jour séquentielle (Gauss-Seidel) [*Luo92,Sauer93,Bertsekas99*]

$$\begin{aligned}x_{n,k+1} &= \arg \min_{\mathbb{R}^+} F(x_{1,k+1}, \dots, x_{n-1,k+1}, \mathbf{x}_n, x_{n+1,k}, \dots, x_{N,k}), \\x_{m,k+1} &= x_{m,k}, m \neq n.\end{aligned}$$

Mise à jour parallèle (Jacobi) [*Sauer93,Elad07*]

$$\begin{aligned}y_{n,k} &= \arg \min_{\mathbb{R}^+} F(x_{1,k}, \dots, x_{n-1,k}, \mathbf{x}_n, x_{n+1,k}, \dots, x_{N,k}), \forall n, \\x_{k+1} &= \mathbf{x}_k + \alpha_k(\mathbf{y}_k - \mathbf{x}_k)\end{aligned}$$

- Nécessite de savoir calculer les minimiseurs de façon explicite
- Sinon, minimisation approchée par Newton-Raphson ou par MM : Algorithmes CA-NR [*Bouman96*] ou CA-PS [*Zheng00,Sotthivirat02*]
- Permet de traiter un problème non différentiable

PLAN

- 1 Activation des contraintes
- 2 Points intérieurs**
- 3 Minimisation d'un critère avec barrière
- 4 Applications
- 5 Conclusion

Fonction barrière

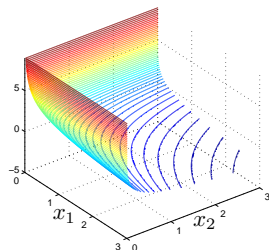
Définition

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

Ex : Barrière logarithmique pour les contraintes de positivité

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

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



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

\Rightarrow Si un critère contient une fonction barrière, ses minimiseurs appartiennent au domaine \mathcal{C} .

Algorithmes de barrière

$$\min_{\mathbf{x} \geq 0} F(\mathbf{x}) \Leftrightarrow \min F(\mathbf{x}) + \mu B(\mathbf{x}), \{\mu\} \rightarrow 0$$

Barrière logarithmique [Fiacco67]

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

- Principalement utilisée en programmation linéaire
- Analyse de convergence [Nesterov94]

Barrière proximale [Eggermont90]

$$B(\mathbf{x}) = \sum_n x_{n,k} \log \frac{x_{n,k}}{x_n} + x_n - x_{n,k}$$

- Il s'agit d'un algorithme MM
- Analyse de convergence [Auslender95, Teboulle97]

Algorithme RL

$$\min_{\mathbf{x} \geq 0} F(\mathbf{x}) = \sum_m [\mathbf{K}\mathbf{x}]_m - z_m \log[\mathbf{K}\mathbf{x}]_m$$

Algorithme RL [Richardson72, Lucy74]:

$$\mathbf{x}_{k+1} = \mathbf{x}_k \left(\mathbf{K}^T \frac{\mathbf{z}}{\mathbf{K}\mathbf{x}_k} \right) \left(\frac{1}{\mathbf{K}^T \mathbf{1}} \right)$$

Nombreuses interprétations

- EM [Shepp82], POCS [Vardi85], barrière proximale [Eggermont98, Chrétien00], gradient séparé [Lanteri02], gradient préconditionné [Kaufman87]...
- Algorithme MM avec approximation majorante séparable [DePierro93]

$$H(\mathbf{x}, \mathbf{x}_k) = \sum_n a_n x_n - b_n(\mathbf{x}_k) \log x_n + c_n(\mathbf{x}_k)$$

- Utilisation en NMF [Lee99, Dhillon05]

Algorithme ISRA

$$\min_{\mathbf{x} \geq 0} F(\mathbf{x}) = \|\mathbf{K}\mathbf{x} - \mathbf{z}\|_2^2$$

Algorithme ISRA [Daube-Witherspoon86]:

$$\mathbf{x}_{k+1} = \mathbf{x}_k \left(\frac{\mathbf{K}^T \mathbf{z}}{\mathbf{K}^T (\mathbf{K} \mathbf{x}_k)} \right)$$

Il s'agit aussi d'un algorithme MM avec approximation majorante séparable quadratique [DePierro87]

$$H(\mathbf{x}, \mathbf{x}_k) = \sum_n a_n(\mathbf{x}_k) x_n^2 + b_n x_n + c_n(\mathbf{x}_k)$$

$$\lim_{x_{n,k} \rightarrow 0} a_n(\mathbf{x}_k) = +\infty$$

Par construction, pour tout k , $\mathbf{x}_k > 0$.

PLAN

- 1 Activation des contraintes
- 2 Points intérieurs
- 3 Minimisation d'un critère avec barrière**
- 4 Applications
- 5 Conclusion

Critères contenant des fonctions barrières

① *Contrainte de positivité :*

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

② *Modèle de bruit Poisson (Ex: Tomographie d'émission)*

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

③ *Pénalisation par maximum d'entropie*

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

Formulation générale du problème

Problèmes d'optimisation de la forme :

$$\min (F(\mathbf{x}) = P(\mathbf{x}) + \mu B(\mathbf{x})), \quad \mu > 0 \quad (1)$$

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

Algorithme itératif

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

- \mathbf{d}_k *direction de descente* (Gradient, Newton, Newton tronqué, gradient conjugué ...)
- $\alpha_k > 0$ pas obtenu par minimisation de $f(\alpha) = F(\mathbf{x}_k + \alpha \mathbf{d}_k)$.

Problématique

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

b non définie pour $\alpha \notin]\alpha_-, \alpha_+[$ s'il existe i tel que

$$\mathbf{c}_i^T (\mathbf{x}_k + \alpha_{\pm} \mathbf{d}_k) + \rho_i = 0.$$

- Restriction de la recherche de pas à $\alpha \in]\alpha_-, \alpha_+[$
- Prise en compte des asymptotes verticales en α_- et α_+
- Méthodes classiques peu efficaces
- Fonctions d'interpolation log-quadratiques [Murray94]

Recherche de pas MM quadratique

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

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

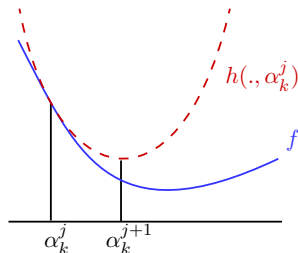
$$h(\alpha, \beta) = f(\beta) + (\alpha - \beta)\dot{f}(\beta) + \frac{1}{2}a_\beta(\alpha - \beta)^2$$

Pas issu de l'AMQ :

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

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

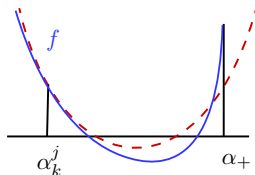
$$\alpha_k = \alpha_k^J$$



Recherche de pas MM non quadratique

La dérivée du terme de barrière est **non bornée** en α_{\pm}

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



Construction d'approximations majorantes de f de la forme

$$h(\alpha, \alpha_k^j) = h_0 + h_1\alpha + h_2\alpha^2 - h_3 \log(h_4 - \alpha)$$

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

Résultats 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

Quel que soit le nombre de sous-itérations MM,

- La première condition de Wolfe est vérifiée
- Le pas est 'suffisamment grand'
- La condition de Zoutendijk est vérifiée

⇒ Convergence de nombreux algorithmes démontrée :

- Gradient, Newton, Quasi-Newton ... : $\lim_k \|\nabla F(\mathbf{x}_k)\| = 0$
- Gradient conjugué : $\liminf_k \|\nabla F(\mathbf{x}_k)\| = 0$

PLAN

- 1 Activation des contraintes
- 2 Points intérieurs
- 3 Minimisation d'un critère avec barrière
- 4 Applications**
- 5 Conclusion

Illustration en optimisation contrainte

Points intérieurs pour la programmation quadratique

$$\begin{aligned} \min_{\mathbf{x} \in \mathbb{R}^n} F(\mathbf{x}) &= \rho_0 + \mathbf{c}_0^T \mathbf{x} + \frac{1}{2} \mathbf{x}^T \mathbf{C}_0 \mathbf{x} \\ \text{s.c. : } C(\mathbf{x}) &= \mathbf{C} \mathbf{x} + \boldsymbol{\rho} \geq \mathbf{0} \end{aligned}$$

Critère augmenté

$$F_\mu(\mathbf{x}) = F(\mathbf{x}) - \mu \sum_{i=1}^I \log([\mathbf{C} \mathbf{x} + \boldsymbol{\rho}]_i)$$

Points intérieurs: Pour une suite $\{\mu\} \rightarrow 0$,

- Minimiser le critère augmenté F_μ . **Méthode primale**
- Annuler le résidu $r_\mu(\mathbf{x}, \boldsymbol{\lambda})$ afin de résoudre une version perturbée des conditions KKT. **Méthode primale-duale**

Points intérieurs pour la programmation quadratique

Méthode primale [Boyd04]

- 1) Calcul de la direction de Newton \mathbf{d} de F_μ en \mathbf{x}
- 2) Recherche de pas sur $F_\mu(\mathbf{x} + \alpha\mathbf{d})$

Méthode primale-duale [Armand00]

- 1) Calcul des directions de Newton $\mathbf{d}_x, \mathbf{d}_\lambda$ de r_μ en \mathbf{x} et $\boldsymbol{\lambda}$
- 2) Recherche de pas sur $\psi_\mu(\mathbf{x} + \alpha\mathbf{d}_x, \boldsymbol{\lambda} + \alpha\mathbf{d}_\lambda)$, où

$$\psi_\mu(\mathbf{x}, \boldsymbol{\lambda}) = F_\mu(\mathbf{x}) + \boldsymbol{\lambda}^T \mathbf{x} - \mu \sum_{i=1}^m \log(\lambda_i [\mathbf{C}\mathbf{x} + \boldsymbol{\rho}]_i)$$

Résultats expérimentaux

Test : 100 problèmes aléatoires de taille $N = 100$, $I = 100$ [Moré89]

Algorithme primal

Rebroussement	100 iter	0.6 s
Interpolation LQ	50 iter	0.5 s
MM LQ	43.1 iter	0.3 s

Algorithme primal-dual

Rebroussement	9.6 iter	0.06 s
Interpolation LQ	9.7 iter	0.13 s
MM LQ	10.5 iter	0.09 s

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,

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

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

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

Difficultés : Problème inverse de grande taille et très mal posé.

Résonance Magnétique Nucléaire 1D

Reconstruction par maximum d'entropie

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

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

Stratégie d'optimisation

- Algorithme de Newton tronqué
- Préconditionnement basé sur la SVD
- Comparaison de la stratégie MM proposée avec la recherche de pas de Moré & Thuente [*Moré94*] pour deux stratégies d'interpolation.

Résonance Magnétique Nucléaire 1D

MT-Cubic	c_1	c_2	$Iter$	$Tps(s)$
	10^{-3}	0.5	35	10
	10^{-3}	0.9	41	11
MT-LQ	c_1	c_2	$Iter$	$Tps(s)$
	10^{-3}	0.5	35	11
	10^{-3}	0.9	35	11
MM-LQ	J	$Iter$	$Tps(s)$	
	1	36	5	
	2	40	6	
	5	40	7	
	10	40	7	

(c_1, c_2) : Paramètres de Wolfe

J : Nombre de sous-itérations MM

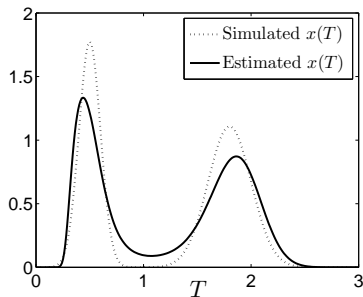


Figure: Reconstruction de données RMN simulées avec $RSB = 25dB$

Résonance Magnétique Nucléaire 2D [Chouzenoux10]

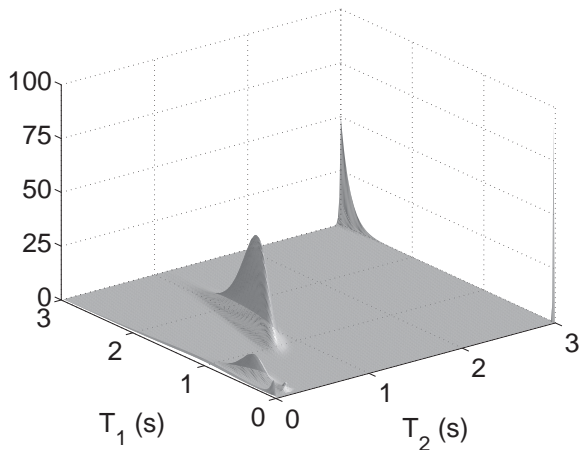


Figure: Reconstruction de données RMN réelles

PLAN

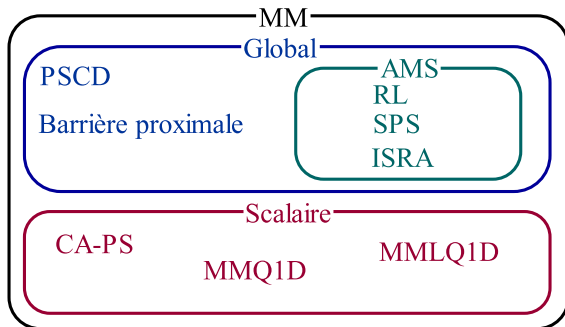
- 1 Activation des contraintes
- 2 Points intérieurs
- 3 Minimisation d'un critère avec barrière
- 4 Applications
- 5 Conclusion

Conclusion

Recherche de pas MM: MMQ1D ou MMLQ1D

- Méthode simple
- Résultats de convergence
- Efficacité pratique

Classification des algorithmes MM pour les contraintes de positivité



Bibliographie (1/6)



D.P. 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.



B. Polyak

Introduction to Optimization.

Optimization Software, Publication Division, 1987.



J.M. Bardsley and C.R. Vogel

A nonnegatively constrained convex programming method for image reconstruction

SIAM journal on Scientific Computing, 25(4):1326–1343, 2003.



D.P. Bertsekas

Projected Newton methods for optimization problems with simple constraints

IEEE Conference on Decision and Control including the Symposium on Adaptive Processes, 20:762–767, 1981.



E.G. Birgin and J.M. Martinez and M. Raydan

Spectral projected gradient methods

In Encyclopedia of Optimization, C.A. Floudas and P. M. Pardalos, Springer US, 3652–3659, 2009.



S. Bonettini and R. Zanella and L. Zanni

A scaled gradient projection method for constrained image deblurring

Inverse Problems, 25(1), 2009.



C.A. Bouman and K. Sauer

A unified approach to statistical tomography using coordinate descent optimization

IEEE Transactions on Image Processing, 5(3):480–492, 1996.

Bibliographie (2/6)



R.H. Byrd and P. Lu and J. Nocedal and C. Zhu
A limited memory algorithm for bound constrained optimization
SIAM Journal on Scientific Computing, 16:1190:1208,1995



E. Chouzenoux and S. Moussaoui and J. Idier
A Majorize-Minimize line search algorithm for barrier function optimization
Rapport technique IRCCyN. 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
A paraître dans IEEE Signal Processing. Mars 2010
<http://hal.archives-ouvertes.fr/hal-00455477/fr/>



S. Chrétien and A.O. Hero
Kullback proximal algorithms for maximum-likelihood estimation.
IEEE Transactions on Information Theory, 46(5):1800–1810, 2000.



A.R. Conn and N.I.M. Gould and P.L. Toint
Global convergence of a class of trust region algorithms for optimization with simple bounds
SIAM Journal on Numerical Analysis, 25(2):433–460, 1988.



A.R. De Pierro
On the convergence of the iterative image space reconstruction algorithm for volume ECT
IEEE Transactions on Medical Imaging, 6(2):174–175, 1987.



A.R. De Pierro
On the relation between the ISRA and the EM algorithm for positron emission tomography
IEEE Transactions on Medical Imaging, 12(2):328–333, 1993.

Bibliographie (3/6)



A.R. De Pierro

A modified expectation maximization algorithm for penalized likelihood estimation in emission tomography
IEEE Transactions on Medical Imaging, 14(1):132–137, 1995.



I.S. Dhillon and S. Sra

Generalized nonnegative matrix approximations with Bregman divergences
Advances in Neural Information Processing, 19:283–290, 2005.



P.P.B. Eggermont

Multiplicative iterative algorithms for convex programming
Rapport technique, University of Delaware. Mars 1998.
<http://www.udel.edu/FREC/eggermont/Preprints/emlike.pdf>



P.P.B. Eggermont and V.N. LaRiccia

On EM-like algorithms for minimum distance estimation
Linear Algebra and its Applications, 130:25–42, 1998.



H. Erdogan and J.A. Fessler

Monotonic algorithms for transmission tomography
IEEE Transactions on Medical Imaging, 18(9):801–814, 1999.



C. Févotte and N. Bertin and J.-L. Durrieu

Nonnegative matrix factorization with the Itakura-Saito divergence: With application to music analysis
Neural Computation, 21(3):793–830, 2008.



A.V. Fiacco and G.P. McCormick

The sequential unconstrained minimization technique (SUMT) without parameters
Operations Research, 15(5):820–827, 1967.

Bibliographie (4/6)



W.W. Hager and H. Zhang

Recent advances in bound constrained optimization
In System Modeling and Optimization, 199:67–82, 2006



D.R. Hunter and L. Kenneth

A tutorial on MM algorithms
The American Statistician, 18(1):30–37, 2004



A.N. Iusem

On the convergence properties of the projected gradient method for convex optimization
Computational and Applied Mathematics, 22(1):37–52, 2003.



L. Kaufman

Implementing and accelerating the EM algorithm for Positron Emission Tomography
IEEE Transactions on Medical Imaging, 6(1), 1987.



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.



K. Lange and J.A. Fessler

Globally convergent algorithms for maximum a posteriori transmission tomography
IEEE Transactions on Image Processing, 4(10):1430–1438, 1995.



H. Lanteri and M. Roche and O. Cuevas and C. Aime.

A general method to devise maximum-likelihood signal restoration multiplicative algorithms with non-negativity constraints
Signal Processing, 81:945–974, 2001.

Bibliographie (5/6)



H. Lantéri and M. Roche and C. Aime.

Penalized maximum likelihood image restoration with positivity constraints : multiplicative algorithms
Inverse Problems, 18:1397–1419, 2002.



L. Lucy

An iterative technique for rectification of observed distributions
The Astronomical Journal, 79(6):745–765, 1974.



Z.Q. Luo and P. Tseng

On the convergence of the coordinate descent method for convex differentiable minimization
Journal of Optimization Theory and Applications, 72(1):7–35, 1992.



D.D. Lee and H.S. Seung

Learning the parts of objects with nonnegative matrix factorization
Nature, 401:788–791, 1999.



J.J. Moré and G. Toraldo

Algorithms for bound constrained quadratic programming problems
Numerische Mathematik, 55:377–400, 1989.



J.J. Moré and D.J. Thuente

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



R. Pytlak

An efficient algorithm for large-scale nonlinear programming problems with simple bounds on the variables
SIAM Journal on Optimization, 8(2):532–560, 1998.



W.H. Richardson

Bayesian-based iterative method of image restoration
Journal of the Optical Society of America, 62(1):52–59, 1972.

Bibliographie (6/6)



K. Sauer and C. Bouman

A local update strategy for iterative reconstruction from projections
IEEE Transactions on Image Processing, 41(2):534–548, 1993.



L.A. Shepp and Y. Vardi

Maximum likelihood reconstruction for emission tomography
IEEE Transactions on Medical Imaging, 1(2):113–122, 1982.



S. Soththivirat and J.A. Fessler

Image recovery using partitioned-separable paraboloidal surrogate coordinate ascent algorithms
IEEE Transactions on Image Processing, 11(3):306–317, 2002.



M. Teboulle

Convergence of proximal-like algorithms
SIAM Journal on Optimization, 7(4):1069–1083, 1997.



Y. Vardi and L.A. Shepp and L. Kaufman

A statistical model for positron emission tomography
Journal of American Statistical Association, 80(389):8–37, 1985.



Y. Vardi and L.A. Shepp and L. Kaufman

An active set limited memory BFGS algorithm for large scale bound optimization
Mathematical Methods in Operations Research, 67:443–454, 2008.



J. Zheng and S.S. Saquib and K. Sauer and C.A. Bouman

Parallelizable Bayesian tomography algorithms with rapid, guaranteed convergence
SIAM Journal on Optimization, 14(4):1043–1056, 2000.

Algorithmes de projection

Gradient projeté [Sec.2,Hager06]

$$\begin{aligned}\mathbf{y}_k &= \mathcal{P}^+(\mathbf{x}_k - s_k \nabla J(\mathbf{x}_k)) \\ \mathbf{x}_{k+1} &= \mathbf{x}_k + \alpha_k (\mathbf{y}_k - \mathbf{x}_k)\end{aligned}$$

- Possibilité d'accélérer l'algorithme en préconditionnant le gradient [Berstekas81,Pytlak98], en modifiant la métrique [Bonettini09] ou en utilisant des régions de confiance [Conn88].
- Utilisation du sous-gradient dans le cas non différentiable [Polyak87]
- Calcul des pas (s_k, α_k) par backtracking [Berstekas99,lusem03] ou Barzilai Borwein [Birgin09,Bonettini09].

Activation des contraintes [Sec.3,Hager06]

- Estimation des contraintes actives par gradient projeté suivie d'une minimisation dans le domaine restreint
- Ex : Newton tronqué [More89,Bardsley03] ou quasi-Newton [Byrd95,Xiao08]

Gradient séparé [Lanteri01]

Ecriture du gradient sous la forme :

$$-\nabla J(\mathbf{x}) = U(\mathbf{x}) - V(\mathbf{x}), \quad U(\mathbf{x}), V(\mathbf{x}) \geq 0$$

Algorithme de descente :

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \alpha_k \frac{\mathbf{x}_k}{V(\mathbf{x}_k)} \nabla J(\mathbf{x}_k)$$

où α_k est choisi afin d'avoir $\mathbf{x}_k > 0$.

- Cas particulier : $\alpha_k = 1 \Rightarrow$ schéma multiplicatif.
- Algorithmes RL et ISRA
- Utilisation en NMF [Dhillon05, Fevotte08]
- Accélération en prenant $(\mathbf{x}_k)^{1/m}$ [Lanteri01]