EFFICIENCY OF LINE SEARCH STRATEGIES IN INTERIOR POINT METHODS FOR LINEARLY CONSTRAINED SIGNAL RESTORATION

Emilie Chouzenoux

Institut Gaspard Monge, CNRS UMR 8049
5, Boulevard Descartes, Champs sur Marne
77454 Marne La Vallée Cedex 2, France

Saïd Moussaoui, Jérôme Idier

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

ABSTRACT

We discuss in this paper the influence of line search on the performances of interior point algorithms applied for constrained signal restoration. Interior point algorithms ensure the fulfillment of the constraints through the minimization of a criterion augmented with a barrier function. However, the presence of the barrier function can slow down the convergence of iterative descent algorithms when general-purpose line search procedures are employed. We recently proposed a line search algorithm, based on a majorization-minimization approach, which allows to handle the singularity introduced by the barrier function. We present here a comparative study of various line search strategies for the resolution of a sparse signal restoration problem with both primal and primal-dual interior point algorithms.

Index Terms— Interior point methods, line search, majorize-minimize algorithm, signal restoration

1. INTRODUCTION

Let us consider the inverse problem of recovering an image or a signal $\boldsymbol{x}^o \in \mathbb{R}^N$ from a set of noisy observations $\boldsymbol{y} \in \mathbb{R}^T$, where the forward model is represented by as

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x}^o + \boldsymbol{\epsilon} \tag{1}$$

with $\boldsymbol{H}$ a known matrix that represents the physical effect of the measurement process and $\boldsymbol{\epsilon}$ an additive noise term representing data acquisition errors and model uncertainties. This simple formalism covers many real situations of signal and image restoration [1]. A usual approach is to recover $\boldsymbol{x}$ by minimizing a composite criterion F, which combines a data-fidelity term and a regularization term, under some linear constraints. For example, for image reconstruction problems, a classical constraint concerns non-negativity of pixel intensities [2]. Linear constraints can also arise in sparse signal reconstruction [3] as illustrated in section 4.

Interior points methods transform the constrained optimization problem into a sequence of unconstrained problems by introducing a *barrier function* in F, which makes the gradient of the augmented criterion unbounded at the boundary of the feasible domain so that its minimizers fulfill the constraints [4, Chap.11].

When an iterative descent method is used for the minimization of the augmented criterion, a search along the line supported by the descent direction must be performed at each iteration. However, the presence of the barrier function may cause the inefficiency of general-purpose line search methods and, thus, the slowdown of the algorithm convergence [5]. We proposed in [6] a majorization-minimization (MM) strategy well suited for stepsize determination when minimizing a criterion which contains a barrier function. The proposed method has been shown to outperform standard line search procedures when applied to the minimization of penalized criteria containing a barrier term, such as Poissonian log-likelihoods [7] and maximum entropy criteria [8]. The aim of this paper is to analyze the impact of the MM line search when it is employed inside interior point algorithms for solving a linearly constrained signal restoration problem and to compare its performances with usual line search approaches.

After introducing the optimization framework in Section 2, we give in Section 3 an overview of the line search methods that have been proposed in the literature to account for the singular behavior of barrier functions. Section 4 analyzes the efficiency of those methods through the resolution of a large-scale signal restoration problem.

2. INTERIOR POINT ALGORITHMS

We consider the constrained optimization problem

$$\min_{\boldsymbol{x}\in\mathbb{R}^N} F(\boldsymbol{x}) \quad \text{s.t} \quad C(\boldsymbol{x}) = \boldsymbol{C}\boldsymbol{x} + \boldsymbol{\rho} \geqslant \boldsymbol{0} \tag{2}$$

where $\boldsymbol{C}$ is a $M \times N$ matrix and $\boldsymbol{\rho} \in \mathbb{R}^M$. We present in the following two main families of interior point methods allowing to solve such problem.

2.1. Primal methods

Primal interior point methods, introduced by Fiacco and McCormick [9], find the solution to the constrained optimization

978-1-4577-0570-0/11/$26.00 ©2011 IEEE

problem (2) by solving a sequence of unconstrained optimization subproblems

$$\min_{\boldsymbol{x}\in\mathbb{R}^N} \left(F_\mu(\boldsymbol{x}) = F(\boldsymbol{x}) + \mu B(\boldsymbol{x})\right) \tag{3}$$

for barrier parameter values μ that decrease to 0. The unconstrained subproblems involve an auxiliary function $B(\boldsymbol{x})$, called *barrier function*, that penalizes closeness to the constraint boundaries. The most widely used auxiliary function is the logarithmic barrier

$$B(\boldsymbol{x}) = -\sum_{i=1}^{M} \log\left(C_i(\boldsymbol{x})\right), \tag{4}$$

with $C_i(\boldsymbol{x}) = [\boldsymbol{C}\boldsymbol{x} + \boldsymbol{\rho}]_i$, which makes the augmented criterion F_μ unbounded at the boundary of the feasible region so that its minimizers $\boldsymbol{x}_\mu^*$ fulfill $C_i(\boldsymbol{x}_\mu^*) > 0$ for all i.

2.2. Primal-dual methods

Primal-dual methods are another class of interior-point methods and have only recently been considered for practical large-scale nonlinear optimization. Unlike the classical primal method, primal-dual methods simultaneously estimate both the primal variables $\boldsymbol{x}$ and dual Lagrange multipliers $\boldsymbol{\lambda}$ associated to the constraints [10]. Exact primal-dual solution $(\boldsymbol{x}_\mu^*, \boldsymbol{\lambda}_\mu^*)$ at a given parameter μ is obtained by solving the following perturbed Karush-Kuhn-Tucker (KKT) equations

$$\begin{cases} \nabla F(\boldsymbol{x}) - \boldsymbol{C}^T\boldsymbol{\lambda} = \boldsymbol{0} \\ \lambda_i C_i(\boldsymbol{x}) = \mu, \quad i = 1, ..., M \end{cases} \tag{5}$$

under the constraint $(C(\boldsymbol{x}), \boldsymbol{\lambda}) \geqslant \boldsymbol{0}$.

2.3. Practical implementation

For large problems, solving (3) or (5) exactly would be prohibitive. In practice, for a given μ, primal or primal-dual subproblems are solved iteratively using Newton's algorithm with a line search strategy [4]. The stepsize value is chosen in such a way as to maintain feasibility and to encourage the convergence of the iterates. One of the most popular choices is to require a sufficient decrease in a merit function that measures the progress toward the solution. A common strategy is to use F_μ as a merit function for primal variables and, when primal-dual method is considered, a separate measure to safeguard the dual variables after the primal step has been taken [2]. The merit function can also include both primal and dual variables [11].

A major inconvenient feature of barrier functions is that they tend to cause inefficiencies of standard line search techniques. Therefore, special-purpose line search procedures have to be designed, which we discuss in the next section.

3. LINE SEARCH FOR BARRIER FUNCTION MINIMIZATION

3.1. Problem statement

The line search problem can be described as follows. Given the current point $\boldsymbol{x}_k$ and a direction $\boldsymbol{d}_k$, find an approximate minimizer α_k of the following single variable function[1]

$$f_\mu(\alpha) = F_\mu(\boldsymbol{x}_k + \alpha\boldsymbol{d}_k). \tag{6}$$

The stepsize α_k is usually obtained by iteratively minimizing $f_\mu(\alpha)$ until some convergence conditions are met [12, Chap.3]. Typically, the strong Wolfe conditions are considered

$$f_\mu(\alpha_k) \leqslant f_\mu(0) + c_1\alpha_k\dot{f}_\mu(0), \tag{7}$$

$$|\dot{f}_\mu(\alpha_k)| \leqslant c_2|\dot{f}_\mu(0)|, \tag{8}$$

where $(c_1, c_2) \in (0, 1)$ are tuning parameters that do not depend on k. There exist several procedures to find an acceptable stepsize: exact minimization of f_μ, backtracking, approximation of f_μ using cubic interpolations [13, 12] or quadratic majorizations [14]. However, the barrier term implies that the derivative of f_μ is unbounded when α is such that $C_i(\boldsymbol{x}_k + \alpha\boldsymbol{d}_k)$ is negative for some i. Since the constraints are linear, this happens when α is outside an interval (α_-, α_+) where

$$\begin{cases} \alpha_- &= \max\limits_{i|[\boldsymbol{C}\boldsymbol{d}_k]_i>0} -\frac{[\boldsymbol{C}\boldsymbol{x}_k+\boldsymbol{\rho}]_i}{[\boldsymbol{C}\boldsymbol{d}_k]_i}, \\ \alpha_+ &= \min\limits_{i|[\boldsymbol{C}\boldsymbol{d}_k]_i<0} -\frac{[\boldsymbol{C}\boldsymbol{x}_k+\boldsymbol{\rho}]_i}{[\boldsymbol{C}\boldsymbol{d}_k]_i}, \end{cases} . \tag{9}$$

Therefore, one must ensure that, during the line search, the stepsize values remain in the interval (α_-, α_+). Moreover, because of the vertical asymptotes at α_- and α_+, standard methods using cubic interpolations or quadratic majorizations are not well suited.

3.2. Damped backtracking strategy

A simple approach is the *damped backtraking* method ([4, 11]) that consists in performing a backtracking procedure initialized with $\alpha^0 = \theta\alpha_+$, where $\theta \in (0, 1)$ is close to one. The stepsize values $\alpha^{j+1} = \tau\alpha^j$ (with typically $\tau = \frac{1}{10}$) are then tried out until the fulfillment of the first Wolfe condition (7). However, since the sequence $\{\alpha^j\}$ is unaffected by the behavior of f_μ, such a strategy may be inefficient in the context of barrier function optimization [5].

3.3. Interpolation-based line search

A second technique is to define the trial steps $\{\alpha^j\}$ from an interpolation procedure that minimizes f_μ within a specified

[1]Although the definition of the merit function may vary between algorithms, for simplicity F_μ and the associated one-variable function f_μ will always denote the generic functions in which a sufficient decrease is sought.

interval of uncertainty. Specific interpolating functions have been proposed to account for the barrier singularity [5]. Murray *et coll.* propose a log-quadratic interpolating function of the form

$$f_\mu(\alpha) \approx f_0 + f_1\alpha + f_2\alpha^2 - \mu\log(f_3 - \alpha) \qquad (10)$$

where the coefficients f_i are chosen to fit f and its derivative at two or three trial points. Interpolating function (10) is then incorporated in a standard line search scheme based on the fulfillment of Wolfe conditions (7)-(8) [15].

3.4. Majorization-based line search

Another strategy is to perform the minimization of f_μ using the Majorization-Minimization (MM) principle [16]. In this procedure, the stepsize results from successive minimizations of majorant functions for f_μ. Function $h(\alpha, \alpha')$ is said to be majorant for $f_\mu(\alpha)$ at α' if for all α, $h(\alpha, \alpha') \geqslant f_\mu(\alpha)$ and $h(\alpha', \alpha') = f_\mu(\alpha')$. In [6], we proposed a majorant function $h(., \alpha^j)$ well suited to the minimization of barrier functions. It is piecewise defined under the following form

$$h(\alpha, \alpha^j) = \begin{cases} f_0^- + f_1^-\alpha + f_2^-\alpha^2 - f_3^-\log(\alpha - \alpha_-) \\ \qquad\qquad \text{for all } \alpha \in (\alpha_-; \alpha^j] \\ f_0^+ + f_1^+\alpha + f_2^+\alpha^2 - f_3^+\log(\alpha_+ - \alpha) \\ \qquad\qquad \text{for all } \alpha \in [\alpha^j; \alpha_+) \end{cases}$$

The initial minimization of $f_\mu(\alpha)$ is then replaced by a sequence of easier subproblems, corresponding to the MM update rule $\alpha^j = \arg\min_\alpha h(\alpha, \alpha^{j-1})$, $j = 1, \ldots, J$, initialized with $\alpha^0 = 0$ and parameterized with its number of iterates $J \geqslant 1$.

We shown in previous experiments that the MM line search method outperforms both damped backtracking and interpolation-based strategies in term of time efficiency, when it is used in a descent algorithm for the minimization of a criterion containing a barrier function [6, 8]. In the next section, we propose to analyze the performances of primal and primal-dual interior point algorithms when the stepsize is computed by our MM line search procedure.

4. APPLICATION TO SPARSE SIGNAL RECONSTRUCTION

Let us consider the signal processing problem of recovering a sparse spike train sequence $\boldsymbol{x} \in \mathbb{R}^N$ from an observation vector $\boldsymbol{y} \in \mathbb{R}^M$ which results from the noisy convolution of $\boldsymbol{x}$ with a filter $\boldsymbol{h}$ of length L and a white centered Gaussian additive noise. The ℓ_1 norm is a suited regularization function to account for the sparseness of $\boldsymbol{x}$, which leads to the following optimization problem

$$\min_{\boldsymbol{x}\in\mathbb{R}^N} \|\boldsymbol{y} - \boldsymbol{h} \star \boldsymbol{x}\|_2^2 + \eta\,\|\boldsymbol{x}\|_1 . \qquad (11)$$

			Primal		Primal-Dual	
DB	c_1		K	T	K	T
	0.5		502	1104	**9**	**5.5**
	0.2		169	381	**9**	**5.5**
	0.1		151	365	**9**	**5.5**
	0.01		144	358	**9**	**5.5**
INTERP	c_1	c_2	K	T	K	T
	10^{-1}	0.5	66	174	9	5.7
	10^{-1}	0.9	78	180	9	5.8
	10^{-1}	0.99	86	216	9	6
	10^{-2}	0.5	**67**	**175**	**9**	**5.5**
	10^{-2}	0.9	81	181	9	5.8
	10^{-2}	0.99	91	202	9	6
MM	J		K	T	K	T
	1		73	185	10	6.6
	2		**62**	**168**	**9**	**5.3**
	5		**62**	**173**	**9**	**5.5**
	10		**60**	**175**	9	5.9

Table 1. Comparison between different line search strategies for both primal and primal-dual interior point algorithms, onto a sparse signal reconstruction problem. K denotes the sum of inner iterations and T the time before convergence, with tolerance parameter $\mu_{\min} = 10^{-8}$.

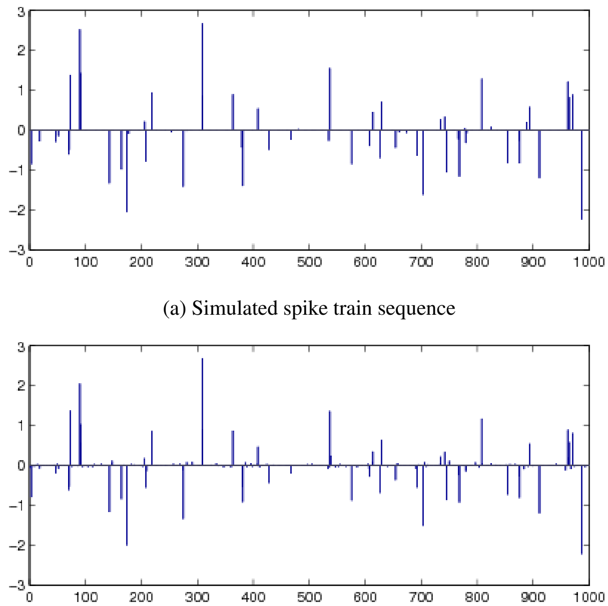
To tackle the non differentiability of the ℓ_1 norm, problem (11) is reformulated as a quadratic programming problem [3]

$$\min_{\boldsymbol{x},\boldsymbol{u}} F(\boldsymbol{x}, \boldsymbol{u}) = \|\boldsymbol{y} - \boldsymbol{h} \star \boldsymbol{x}\|_2^2 + \lambda \sum_i u_i$$
$$\text{subject to } -u_i \leqslant x_i \leqslant u_i, \quad i = 1, \ldots, N. \qquad (12)$$

The purpose of this study is to compare several line search procedures in interior point methods for solving (12). To this end, the primal interior point algorithm [4, Alg.11.1] and the primal-dual interior point algorithm NOPTIQ from [11] are considered. For the latter algorithm, we use the Matlab code from P. Carbonetto available at `www.cs.ubc.ca/~pcarbo/convexprog.html`.

In our experiment, $M = 1020$, $L = 20$, $N = 1000$, the spike train sequence is simulated from a Bernoulli-Gaussian distribution with parameter $\beta = 0.06$, and the signal to noise ratio is 13dB. The regularization parameter is set to $\eta = 0.01$ to get the best result in terms of similarity between the simulated and the estimated signals (in the sense of quadratic error). The two signals are illustrated in Figure 1.

The barrier parameter μ is initially set to 1 and both primal and primal-dual algorithms are stopped when $\mu \leqslant \mu_{\min}$. Table 1 summarizes the computational results in terms of iteration number K and computation time T on an Intel Core 7 2.67 GHz, 4 GB RAM. The performances of the following line search strategies are reported: the damped backtracking line search (DB) with $\theta = 0.99$ and Wolfe parameter c_1, the Wolfe line search [13] with parameters c_1 and c_2 asso-



(a) Simulated spike train sequence

(b) Reconstruction with similarity error 4%

Fig. 1. Sparse signal reconstruction with interior point method.

ciated with the interpolation scheme (10) (INTERP), and the Majorize-Minimize (MM) line search with parameter J.

According to Table 1, the primal algorithm performs better when the stepsize is obtained with the MM search. The best performances in terms of time efficiency are obtained when J is larger or equal to 2. When dealing with INTERP and DB strategies, the better results are obtained for small c_1 and high c_2, which indicates that the best stepsize strategy for primal interior point methods corresponds to a very rough minimization of the barrier function. On this particular problem, very similar performances of the primal-dual interior point algorithm have been obtained with the different stepsize strategies.

5. CONCLUSION

In this work, we have considered the resolution of linearly constrained signal reconstruction problems with interior-point methods. In [6], we proposed a simple MM line search method for barrier function optimization. In a sparse signal restoration application, we showed that this approach outperforms state-of-the-art line search methods, in term of convergence speed, as far as the primal interior point algorithm is concerned. In contrast, the choice of the line search procedure has very little impact on the performances of the primal-dual algorithm. Such a conclusion meets that of [17, Sec.20.6.2] in the context of linear programming.

6. REFERENCES

[1] Guy Demoment, "Image reconstruction and restoration: Overview of common estimation structure and problems," *IEEE Trans. Acoust. Speech, Signal Processing*, vol. 37, no. 12, pp. 2024–2036, December 1989.

[2] C. A. Johnson, J. Seidel, and A. Sofer, "Interior-point methodology for 3-D PET reconstruction," *IEEE Trans. Medical Imaging*, vol. 19, no. 4, April 2000.

[3] S. Chen, D.L. Donoho, and M. Saunders, "Atomic decompositions by basis pursuit," *SIAM J. Sci. Comput.*, vol. 20, pp. 30–61, 1998.

[4] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, New York, 1st edition, 2004.

[5] W. Murray and M. H. Wright, "Line search procedures for the logarithmic barrier function," *SIAM J. Optimization*, vol. 4, no. 2, pp. 229–246, 1994.

[6] E. Chouzenoux, S. Moussaoui, and J. Idier, "A Majorize-Minimize line search algorithm for barrier functions," Tech. Rep., IRCCyN, November 2009.

[7] E. Chouzenoux, S. Moussaoui, and J. Idier, "A majorize-minimize line search algorithm for barrier function optimization," in *Proc. 17th EUSIPCO*, Glasgow, UK, August 2009, pp. 1379–1393.

[8] E. Chouzenoux, S. Moussaoui, J. Idier, and F. Mariette, "Efficient maximum entropy reconstruction of nuclear magnetic resonance T1-T2 spectra," *IEEE Trans. Signal Processing*, vol. 58, no. 12, pp. 6040–6051, December 2010.

[9] A. V. Fiacco and G. P. McCormick, "The sequential unconstrained minimization technique (SUMT) without parameters," vol. 15, no. 5, pp. 820–827, September 1967.

[10] S. J. Wright, *Primal-dual interior-point methods*, SIAM, Philadelphia, PA, 1st edition, 1997.

[11] P. Armand, J. C. Gilbert, and S. Jan-Jégou, "A feasible BFGS interior point algorithm for solving strongly convex minimization problems," *SIAM J. Optimization*, vol. 11, pp. 199–222, 2000.

[12] J. Nocedal and S. J. Wright, *Numerical Optimization*, Springer-Verlag, New York, NY, 1999.

[13] J. J. Moré and D. J. Thuente, "Line search algorithms with guaranteed sufficient decrease," *ACM Trans. on Math. Soft.*, vol. 20, no. 3, pp. 286–307, 1994.

[14] C. Labat and J. Idier, "Convergence of conjugate gradient methods with a closed-form stepsize formula," *J. Optim. Theory Appl.*, vol. 136, no. 1, pp. 43–60, January 2008.

[15] S. G. Nash and A. Sofer, "A barrier method for large-scale constrained optimization," *ORSA J. Comput.*, vol. 5, no. 41, 1993.

[16] D. R. Hunter and Kenneth L., "A tutorial on MM algorithms," *Amer. Statist.*, vol. 58, no. 1, pp. 30–37, February 2004.

[17] C. Roos, T. Terlaky, and J.P. Vial, *Interior Point Methods for Linear Optimization*, Springer-Verlag, New York, NY, 2nd edition, 2006.