Inexact Approach for Interior Point Methods for Large Scale Quadratic Optimization

Ghussoun Al-Jeiroudi

Department of Mathematics, Faculty of Science, Damascus University, Syria

Received 11/10/2010
Accepted 07/03/2011

ABSTRACT

Interior point methods are considered the most powerful tools for solving linear, quadratic and nonlinear programming. In each iteration of interior point method (IPM) at least one linear system has to be solved. The main computational effort of IPMs consist in the computational of these linear systems. That drives many researchers to tackle this subject, which make this area of research very active. The issue of finding a preconditioner for this linear system was investigated in many papers.

In this paper, we provide a preconditioner for interior point methods for quadratic programming. This preconditioner makes the system easier to be solved comparing with the direct approach.

The preconditioner used follows the ideas, which is used for linear approach. An explicit null space representation of linear constraints is constructed by using a non-singular basis matrix identified from an estimate of the optimal partition. This is achieved by means of efficient basis matrix factorisation techniques used in implementations of the revised simplex method.

Keywords: Interior Point Methods, Quadratic Programming, Preconditioner, Indefinite System.

طرائق النقاط الداخلية غير المباشرة لحل مسائل الحل الأمثل التربيعية الضخمة

غصون الجيرودي
قسم الرياضيات – كلية العلوم– جامعة دمشق– سورية

تاريخ الإيداع **2010/10/11**
قبل للنشر في **2011/03/07**

الملخّص

تعد طريقة النقاط الداخلية في الوقت الراهن واحدة من أهم الطرائق المستخدمة لحل مسائل الحل الأمثل الخطية، التربيعية واللاخطية. في كل خطوة من خطوات خوارزمية هذه الطريقة نحتاج على الأقل إلى حلّ نظام خطي. ويشغل حل هذا النظام نحو 60-90% من الزمن الكلي لتشغيل البرنامج. مما دفع العديد من الباحثين لدراسة هذا النظام. وجد العديد من الباحثين أن حل هذا النظام بشكل غير مباشر (باستخدام طريقة التقريب بالتتالي) يؤدي إلى تسريع البرنامج بشكل ملحوظ. غير أن الطرائق غير المباشرة الفعالة تحتاج إلى مكيفات. في السنوات الأخيرة قدمت العديد من البحوث مكيفات لطرائق النقاط الداخلية.

في هذه المقالة، قمنا بتصميم مكيف جديد لطريقة النقاط الداخلية وذلك من أجل البرامج التربيعية. وقد قمنا بتوسيع المكيفات المستخدمة في البرامج الخطية، لتشمل البرامج التربيعية. صمم المكيف لأنظمة متناظرة غير معرفة وغير مستقرة. هذا المكيف حول عدم الاستقرار من نقطة ضعف في هذه الأنظمة إلى نقطة قوة. في هذه المقالة أول مرة ينظر إلى النظام الخطي من زاوية جديدة، حيث قمنا بتمديد عدم الاستقرار إلى معظم أجزاء النظام ومن ثم قمنا بتصميم المكيف. اعتمدنا في هذا التصميم على فكرة المتحولات الأساسية التي تستخدم في طريقة السمبلكس. بحيث دمجنا نوعاً ما بين الطريقتين (طريقة النقاط الداخلية وطريقة السمبلكس). فضلاً عما سبق، ذُكرت خواص المكيف الجيد و خلال المقالة قمنا بإثبات أن المكيف المطروح هو مكيف جيد.

الكلمات المفتاحية: طرائق النقاط الداخلية، البرامج التربيعية، المكيفات، الأنظمة غير المعرفة.

1 - Introduction

We are concerned in this paper with the use of the primal-dual interior point method (IPM for short) to solve large-scale quadratic programming problems. The primal-dual method is applied to the primal-dual formulation of the quadratic program

Primal	**Dual**
$\min\ c^T x + \frac{1}{2} x^T Q x$	$\max\ b^T y + \frac{1}{2} x^T Q x$
$\text{s.t.}\quad Ax = b,$	$\text{s.t.}\quad A^T y + s - Qs = c,$
$x \geq 0$	$y \text{ free},\qquad s \geq 0$

where $A \in R^{m\times n}$, $Q \in R^{n\times n}$, $x,s,c \in R^n$ and $y,b \in R^m$. We assume that $m \leq n$ and A has full row rank. The primal-dual algorithm is usually faster and more reliable than the pure primal or pure dual method. The main computational cost of this algorithm is the computation of the primal-dual Newton direction. Applying standard transformations (Andersen et al. 1996 and Wright, 1997) leads to the following linear system that must be solved at each iteration

$$\begin{bmatrix} -(D^{-2} + Q) & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} = \begin{bmatrix} f \\ g \end{bmatrix} \qquad (1)$$

Where $D = X^{1/2} S^{-1/2}$ and X and S are diagonal matrices in $R^{n\times n}$ with the elements of vectors x and s respectively on the diagonal, μ is the average complementarity gap $\mu = x^T s/n$

$$f = c - A^T y + Qx - \mu X^{-1} e \quad \text{and} \quad g = b - Ax.$$

Solving the linear systems which arise from interior point method dominates the computation time, especially when large-scale problems are considered. Furthermore, these linear systems become extremely ill-conditioned as the IPM approaches the solution, which leads to numerical instability in the final iterations of IPM. This ill-conditioning is due to that some elements of D go to zero while the others go to infinity. From the complementarity condition, we know that either x_i or s_i go to zero $\forall$ i $\in$ {1,2,...,n}, see (Wright, 1997). These facts make us concentrate our attention on solving the final iterations of IPM. In this paper we propose a block triangular preconditioner for a modified augmented system and solve the final iterations of IPM iteratively.

Recently many researchers introduce a certain type (class) of preconditioners, the ones which try to guess a ”basis”, a non-singular sub-matrix of A, (Al-Jeiroudi *et al.,* 2009 and 2008, Bergamaschi *et al.,* 2007 and 2004, Chai et al. 2007, Dollar et al. 2006, 2005 and 2004, Keller *et al.,* 2000 and Gill *et al.,* 1992). There has been recently a growing interest in such preconditioners, because these types of preconditioners try to overcome the problem of ill conditioning of the linear system. In (Al-Jeiroudi *et al.,* 2008) the authors propose a successful block triangular preconditioner to solve the augmented system which arises from the linear programming problems. This preconditioner works well at the final iterations of IPM, as they used the ill-conditioning problem as an advantage in designing their preconditioner. In order to get the standard augmented system (1) the term Δs is eliminated. However, in this paper we will eliminate Δx instead of Δs. That leads to spread the matrix D in all sub-matrices of the modified augmented system matrix, which spread the ill-conditioning to all sub-matrices of the augmented system. However, we can overcome this problem by applying a similar strategy to one used in (Al-Jeiroudi *et al.,* 2008) to construct our preconditioner. The paper is organised as follows. In Section 2, we briefly review the primal-dual interior point method for quadratic programming. In Section 3, we introduce the block triangular preconditioner.

2-The interior point method for quadratic programming

It is widely accepted that the primal-dual interior point method is the most efficient variant of interior point method (Andersen *et al.,* 1996 and Oliveira *et al.,* 1997). The usual transformation in interior point methods consists of replacing inequality constraints by the logarithmic barrier. The primal barrier problem becomes:

$$\min \quad c^T x + \frac{1}{2} x^T Q x - \mu \sum_{j=1}^{n} \ln x_j$$

$$\text{s.t.} \qquad Ax = b,$$

Where $\mu \geq 0$ is a barrier parameter. The Lagrangian associated with this problem has the form:

$$L(x, y, \mu) = c^T x + x^T Q x - y^T (Ax - b) - \mu \sum_{j=1}^{n} \ln x_j$$

and the conditions for a stationary point are

$$\nabla_x L(x, y, \mu) = c + Qx - A^T y - \mu X^{-1} e = 0$$

$$\nabla_y L(x, y, \mu) = \quad Ax - b \quad = 0$$

Where $X^{-1} = \operatorname{diag}\{x_1^{-1}, x_2^{-1}, \dots, x_n^{-1}\}$. Denoting $s = \mu X^{-1} e$, i.e. $XSe = \mu e$, where $S = \operatorname{diag}\{s_1, s_2, \dots, s_n\}$ and $e = (1, 1, \dots, 1)^T$, the first order optimality conditions (for the barrier problem) are:

$$\begin{aligned} Ax &= b, \\ A^T y + s - Qx &= c, \\ XSe &= \mu e, \\ (x, s) &> 0 \end{aligned} \qquad (2)$$

The interior point algorithm for quadratic programming applies Newton's method to solve this system of nonlinear equations and gradually reduces the barrier parameter μ to guarantee convergence to the optimal solution of the original problem. The Newton direction is obtained by solving the system of linear equations:

$$A\,\Delta x = b - Ax, \qquad (3)$$

$$A^T \Delta y + \Delta s - Q\Delta x = c - A^T y - s + Qx, \qquad (4)$$

$$S\,\Delta x + X\,\Delta s = \mu e - XSe. \qquad (5)$$

Usually the term Δs is eliminated from the previous equations to obtain the standard augmented system. In this paper, we eliminate Δx instead. That is to have the augmented matrix in a particular form in order to construct our preconditioner (see next section). Eliminating Δx from (5) gives

$$\Delta x = -S^{-1}X\,\Delta s + \mu S^{-1}e - x \qquad (6)$$

Substituting this in (3) and (4) gives the following equations:

$$A\,S^{-1}X\,\Delta s = -b + \mu\,A\,S^{-1}e, \qquad (7)$$

$$(I + Q\,S^{-1}X)\Delta s + A^T\,\Delta y = c - A^Ty - s + \mu Q\,S^{-1}e. \qquad (8)$$

Let us introduce the vectors t_s and t_y such that

$$\Delta s = X^{-1/2}S^{1/2}\,t_s, \qquad \Delta y = t_y.$$

Substituting these in (7) and (8) leads to the following equations:

$$A\,S^{-1/2}X^{1/2}\,t_s = -b + \mu\,A\,S^{-1}e$$

$$(I + X^{1/2}S^{-1/2}Q\,S^{-1/2}X^{1/2})t_s + X^{1/2}S^{-1/2}A^Tt_y = X^{\frac{1}{2}}S^{-\frac{1}{2}}(c - A^Ty - s + \mu Q\,S^{-1}e).$$

That leads to the following indefinite augmented system of linear equations

$$\begin{bmatrix} (I + DQD) & DA^T \\ AD & 0 \end{bmatrix}\begin{bmatrix} t_s \\ t_y \end{bmatrix} = \begin{bmatrix} \bar{f} \\ \bar{g} \end{bmatrix} \qquad \textbf{(9)}$$

where $\bar{f} = X^{1/2}S^{-1/2}(c - A^Ty - s + \mu Q\,S^{-1}e)$ and $\bar{g} = -b + \mu A\,S^{-1}e.$

3- Block triangular preconditioner

We use a similar strategy which is used in (Al-Jeiroudi *et al.*, 2008) for LP to design a block triangular preconditioner for quadratic programming.

In (Al-Jeiroudi *et al.*, 2008) the following reasoning has been used to find the preconditioner for KKT system. From the complementarity condition we know that at the optimum $x_i y_i = 0$, $\forall j \in \{1, 2, \ldots, n\}$.

Primal dual interior point methods usually identify a strong optimal partition near the optimal solution. If at the optimal solution $x_i \to 0$ and $s_i \to \hat{s}_i$, then the corresponding element $D_i \to 0$. If, on the other hand, $x_i \to \hat{x}_i$ and $s_i \to 0$, then the corresponding element $D_i \to \infty$.

In fact, the optimal partition is closely related (but not equivalent to) the basic-nonbasic partition in the simplex method. That is due to that simplex method iterations move from vertex to vertex until the optimal solution is found. So the simplex method has exactly m basic variables (variables belong to B) and n-m nonbasic variables (variables belong to N). However, interior point methods approach the optimal solution by moving through the interior of the feasible region. Consequently, interior point methods have m basic variable and n - m nonbasic variables on its limit only. If at the optimal solution $j \in B$, then $x_i \to \hat{x}_i$ and $s_i \to 0$,. hence the corresponding element $D_i \to \infty$. If at the optimal solution $j \in N$, then $x_i \to 0$ and $s_i \to \hat{s}_i$ and $D_i \to 0$. Summing up,

$$D_j \to \begin{cases} \infty, & \text{if } j \in B \\ 0, & \text{if } j \in N \end{cases} \qquad (10)$$

This property of interior point methods is responsible for a number of numerical difficulties. In particular, it causes the linear systems (9) to become very ill-conditioned when an interior point method approaches the optimal solution, see (Andersen et al. 1996). However, it may be used to advantage when constructing a preconditioner for the iterative method, see (Al-Jeiroudi *et al.*, 2008).

We partition the matrices and vectors as has been done in (Al-Jeiroudi *et al.*, 2008):

$$A = [A_B \;\; A_N], \quad D = \begin{bmatrix} D_B & 0 \\ 0 & D_N \end{bmatrix}, \; Q = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix}, \quad x = [x_B \;\; x_N], \; s = [s_B \;\; s_N]$$

according to the partition of {1,2,...,n} into sets B and N. With this notation, from (10) we conclude that $D_N \approx 0$ and $D_B^{-1} \approx 0$. Consequently, the matrix in the augmented system (9) can be approximated as follows:

$$\begin{bmatrix} I + D_B Q_{11} D_B & D_B Q_{12} D_N & D_B A_B^T \\ D_N Q_{21} D_B & I + D_N Q_{22} D_N & D_N A_N^T \\ A_B D_B & A_N D_N & 0 \end{bmatrix} \approx \begin{bmatrix} I + D_B Q_{11} D_B & D_B Q_{12} D_N & D_B A_B^T \\ D_N Q_{21} D_B & I & 0 \\ A_B D_B & 0 & 0 \end{bmatrix}, \quad \textbf{(11)}$$

If the matrix A_B were square and nonsingular then equations (11) would suggest obvious preconditioner for the augmented system. However, there is no guarantee that this is the case. On the contrary, in practical applications it is very unlikely that the matrix A_B corresponding to the optimal partition is square and nonsingular. Moreover, the optimal partition is known only when an IPM approaches the optimal solution, see (Al-Jeiroudi *et al.*, 2008). To construct a preconditioner to (9) with a structure similar to the approximation (11) we need to guess an optimal partition and, additionally, guarantee that the matrix B which approximates A_B is nonsingular. We exploit the difference in magnitude of elements in D to design a preconditioner. We sort the elements of D in non-increasing order: $D_1 \geq D_2 \geq \cdots \geq D_n$. We select the first m linearly independent columns of the matrix A, when permuted according to the order of Dj, and we construct a nonsingular matrix B from these columns. The submatrix of A corresponding to all the remaining columns is denoted by N. Therefore we assume that a partition A = [B, N] is known such that B is nonsingular and the entries D_j corresponding to columns of B are chosen from the largest elements of D, while the entries D_j corresponding to columns of N are chosen from the smallest elements of D. According to this partitioning of A, Q and D (and after a symmetric row and column permutation) the indefinite matrix in (9) can be rewritten in the following form

$$K = \begin{bmatrix} I + D_B Q_{11} D_B & D_B Q_{12} D_N & D_B B^T \\ D_N Q_{21} D_B & I + D_N Q_{22} D_N & D_N N^T \\ B D_B & N D_N & 0 \end{bmatrix} \qquad \textbf{(12)}$$

By construction, the elements of D_N are supposed to be among the smallest elements of D, hence we may assume that D_N≈ 0. The following easily invertible block-triangular matrix

$$P = \begin{bmatrix} I + D_B Q_{11} D_B & D_B Q_{12} D_N & D_B B^T \\ D_N Q_{21} D_B & I & 0 \\ B D_B & 0 & 0 \end{bmatrix} \qquad \textbf{(13)}$$

Is a good approximation to K. Hence P is an attractive preconditioner for K. The preconditioner inverse is

$$P^{-1} = \begin{bmatrix} 0 & 0 & {D_B}^{-1} B^{-1} \\ 0 & I & -D_N Q_{21} B^{-1} \\ B^{-T} {D_B}^{-1} & -B^{-T} Q_{12} D_N & B^{-T}(-{D_B}^{-2} - Q_{11} + Q_{12} {D_N}^2 Q_{21}) B^{-1} \end{bmatrix}.$$

Naturally there are further requirements that a successful preconditioner should satisfy: it should be easily invertible and it should capture the numerical properties of (12). P is easily invertible because it is block-triangular with non-singular diagonal blocks BD_B , I and $D_B \ B^T$. In the next section, we give an explicit formulae for the solution of equations with the preconditioner (13).

4- Solving equations with P

The matrix (13) is block triangular and its diagonal blocks BD_B , I and $D_B \ B^T$and are invertible.

Let $d = [\, d_B, d_N, d_y]$ and $r = [\, r_B, r_N, r_y]$ and consider the system of equations

$$\begin{bmatrix} I + D_B Q_{11} D_B & D_B Q_{12} D_N & D_B B^T \\ D_N Q_{21} D_B & I & 0 \\ BD_B & 0 & 0 \end{bmatrix} \begin{bmatrix} d_B \\ d_N \\ d_y \end{bmatrix} = \begin{bmatrix} r_B \\ r_N \\ r_y \end{bmatrix} \qquad (14)$$

The solution of (14) can easily be computed by exploiting the block-triangular structure of the matrix:

$$BD_B d_B = r_y \ \Rightarrow \ d_B = D_B^{-1} B^{-1} \, r_y$$

$$d_N + D_N Q_{21} D_B d_B = r_N \ \Rightarrow \ d_N = \ r_N - D_N \, Q_{21} \, D_B d_B$$

$$D_B B^T d_y + D_B Q_{12} D_N d_N + (I + D_B Q_{11} D_B) d_B = r_B \ \Rightarrow \ d_y = \ B^{-T} \, D_B^{-1} (r_B - D_B \, Q_{12} \, D_N d_N - (I + \ D_B Q_{11} \, D_B) d_B).$$

The operation $d = P^{-1} r$ involves solving two equations (one with B and one with B^T) and a couple of matrix-vector multiplications. These operations will be performed at every iteration of the iterative solver procedure hence they should be implemented in the most efficient way. The issues of choosing a well-conditioned basis matrix B with sparse factored inverse are addressed in (Al-Jeiroudi *et al.*, 2008).

5 - Analysing the Preconditioner P

At each iteration of interior point method for quadratic programming, the linear system (9) is required to be solved. This system can be solved using iterative solver preconditioned with the preconditioner (13). Searching for good preconditioner is an essential step in solving any linear system iteratively. In this section we address what makes (13) a good preconditioner. In order to get good preconditioner, the preconditioner should satisfy the following conditions: Firstly, the preconditioner should be well approximated to original matrix. Secondly, it should be easy to solve an equation with the preconditioner, see (Bjorck, 1996).

Section 3 shows that the preconditioner P is a good approximation to the original matrix K, as we construct P by eliminating the elements of D which are close to zero. Section 4 shows that it is easy to solve an equation with the preconditioner P. In the iterative solver, we require to solve an equation with P. This approach however, does not require to solve any equation with the original matrix K. In the following Table we provide numerical results to compare solving an equation with P to an equation with K. In practice, the system with K is usually reduced to smaller system (its size is m by m). This system is called by normal equation, its matrix is $A\,(D^{-2}+Q)^{-1}A^{T}$. In Section 4, we mentioned that to solve a system with P, we require to solve a system with B and a system with B^{T}, so we need to factorise B only. Cholesky factorization is used to solve the normal equation and the LU factorization is used to solve a system with B.

In practice, most problems are sparse, so we save the problems in sparse form. The following Table shows the saving in term of computer storage comparing (the number of nonzero element in the off-diagonal matrix of the Cholesky factorization for the normal equation's matrix) to (the number of nonzero element in the LU factorization for the matrix B).

problem	**The problem size**			**nnz for Cholesky Factorization of** $A\,(D^{-2}+Q)^{-1}A^{T}$	**nnz for LU Factorization of B**	**Saving %**
	m	**n**	**nnz of A**			
problem1	27	79	129	107	53	50.5
Problem2	402	1456	8450	4096	970	76.3
Problem3	402	1456	5810	3252	1094	66.4
Problem4	616	1092	3467	2658	1798	32.4
Problem5	2953	7335	21252	40288	14711	63.5

Note: we denote by nnz the number of nonzero elements.

We notice that, we save more than 50% of the storage in most cases. The matrix Q in all previous problems are diagonal matrices. In problem1, if we take Q a non-diagonal matrix the number of nonzero elements in Cholesky factorization will become 298, while the number of nonzero elements of LU factorization of B will still the same. In this case will save 82% of the saving storage.

6 - Conclusions

In this paper, we provide a preconditioner for the linear system which arise from IPM for quadratic programming. All the researches which have been made in this area study the linear system (1). However, we suggest to solve the linear system (9). Nobody has driven or used this linear system before. The matrix D appears in most terms of the system (9), which makes the system unstable, that is because some elements of D go to zero and the other go to infinity. We use this instability as an advantage to design our preconditioner. Hence we replace the elements, which go to zero, with zeros in the preconditioner. The result is a nice block triangular matrix. The constraint matrix A is partitioned into [B, N], where B is m by m non-singular matrix. B is produced from m linearly independent columns of A, which correspond to small D.

REFERENCES المراجع

Al-Jeiroudi, G. and Gondzio, J. (2009). Convergence analysis of inexact infeasible interior point method for linear optimization. Journal on Optimization Theory and Applications, V.247, pp. 141:231.

Al-Jeiroudi, G. Gondzio, J. and Hall, J. (2008). Preconditioning indefinite systems in interior point methods for large scale linear optimization. Optimization Methods and Software, V.23, pp.345–363.

Andersen, E. Gondzio, J. Meszaros, C. and Xu, X. (1996). Implementation of interior point methods for large scale linear programming. In T. Terlaky, editor, Interior Point Methods in Mathematical Programming, Kluwer Academic Publishers, pp. 189–252.

Bergamaschi, L. Gondzio, J. Venturin, M. and Zilli, G. (2007). Inexact constraint preconditioners for linear systems arising in interior point methods. Computational Optimization and Applications, V.36, pp.137–147.

Bergamaschi, L. Gondzio, J. and Zilli, G. (2004). Preconditioning indefinite systems in interior point methods for optimization. Computational Optimization and Applications, V.28, pp.149–171.

Bjorck, A. (1996). Numerical Methods for Least squares problems, SIAM, Philadelphia.

Chai J. S. and Toh, K. C. (2007). Preconditioning and iterative solution of symmetric indefinite linear systems arising from interior point methods for linear programming. Computational Optimization and Applications, V.36, pp.221–247.

Dollar, H. Gould, N. and Wathen, A. (2004). On implicit-factorization constraint preconditioners. Technical Report RAL-TR-2004-036, Rutherford Appleton Laboratory, England.

Dollar H. and Wathen, A. (2005). Approximate factorization constraint preconditioners for saddle point matrices. SIAM Journal on Scientific Computing, V.27, pp.1555–1572.

Dollar, H. Gould, N. Schilders, W. and Wathen, A. (2006). Implicit-factorization preconditioning and iterative solvers for regularized saddle-point systems. SIAM Journal on Matrix and Applications, V.28, pp.170–189.

Gill, P. Murray, W. Ponceleon, D. and Saunders, M. (1992). Preconditioners for indefinite systems arising in optimization. SIAM Journal on Matrix Analysis and Applications, V.13, pp. 292–311.

Keller, C. Gould, N. and Wathen, A. (2000). Constraint preconditioning for indefinite linear systems. SIAM Journal on Matrix Analysis and Applications, V.21, pp.1300–1317.

Oliveira A. and Sorensen, D. (1997). A new class of preconditioners for large-scale linear systems from interior point methods for linear programming. Linear Algebra and its Applications, V.394, pp.1–24.

Wright, S. (1997). Primal-Dual Interior-Point Methods. SIAM, Philadelphia.