Improving an interior-point algorithm for multicommodity flows by quadratic regularizations [1]

Jordi & Castro* Jordi & Cuesta⋆

*Dept. of Statistics and Operations Research, Universitat Politècnica de Catalunya
c. Jordi Girona 1–3, 08034 Barcelona, Catalonia, Spain

⋆Statistics and Operations Research unit, Dept. of Chemical Engineering, Universitat Rovira i Virgili
38-40 Avda. Països Catalans 26, 43007 Tarragona, Catalonia, Spain

Abstract

Some classes of multicommodity flow problems are most efficiently solved by a specialized interior-point method that approaches the normal equations by a combination of Cholesky factorizations and preconditioned conjugate gradient. Its efficiency depends on the spectral radius–in [0,1)–of a certain matrix in the definition of the preconditioner. It was empirically observed that this method was more efficient for quadratic than for linear problems. In this talk we show this is due to a reduction in the above spectral radius by the quadratic term in the objective. Using this fact, we suggest a procedure that solves linear multicommodity flow problems as a sequence of quadratic ones by adding a quadratic regularization. Instead in other regularizations used in interior-point methods, we don't regularize the objective function, but the logarithmic barrier, such that the regularization term tends to zero with the barrier parameter. This barrier is shown to still be self-concordant, which guarantees the convergence and polynomial complexity of the algorithm. Computational experience with some multicommodity instances show the efficiency of this approach. Full details can be found in [5].

Keywords: *interior-point methods, multicommodity network flows, preconditioned conjugate gradient, regularizations, large-scale computational optimization*

1 Extended abstract

The primal block-angular formulation dealt with in this work is

$$
\begin{array}{ll}
\min & \displaystyle\sum_{i=0}^{k} ({c^i}^T x^i + {x^i}^T Q_i x^i) \\
\text{subject to} & \begin{bmatrix} N_1 & & & & \\ & N_2 & & & \\ & & \ddots & & \\ & & & N_k & \\ L_1 & L_2 & \dots & L_k & I \end{bmatrix} \begin{bmatrix} x^1 \\ x^2 \\ \vdots \\ x^k \\ x^0 \end{bmatrix} = \begin{bmatrix} b^1 \\ b^2 \\ \vdots \\ b^k \\ b^0 \end{bmatrix} \\
& 0 \le x^i \le u^i \quad i = 0, \dots, k.
\end{array} \tag{1}
$$

[1] Supported by grant MTM2006-05550 of the Spanish Ministry of Science and Education.

Matrices $N_i \in \mathbb{R}^{m_i \times n_i}$ and $L_i \in \mathbb{R}^{l \times n_i}$, $i = 1, \ldots, k$, respectively define the block-diagonal and linking constraints, k being the number or blocks. Vectors $x^i \in \mathbb{R}^{n_i}, i = 1, \ldots, k$, are the variables for each block. $x^0 \in \mathbb{R}^l$ are the slacks of the linking constraints. $b^i \in \mathbb{R}^{m_i}, i = 1, \ldots, k$, is the right-hand-side vector for each block of constraints, whereas $b^0 \in \mathbb{R}^l$ is for the linking constraints. The upper bounds for each group of variables are defined by $u^i, i = 0, \ldots, k$. This formulation considers the general form of linking constraints $b^0 - u^0 \leq \sum_{i=1}^k L_i x^i \leq b^0$. When N_i is an arc-node incidence matrix, and $L_i = I$ we obtain the standard multicommodity network flows problem. In that case $x^i, b^i, i = 1, \ldots, k$ are respectively the flows and demands per commodity, and block and linking constraints are denoted as network and capacity constraints, respectively. We restrict to the separable case where Q_i, $i = 0, \ldots, k$, are diagonal positive semidefinite matrices.

In [2] a specialized method was introduced by exploiting the structure of A (the constraints matrix of (1)) and Θ (the diagonal scaling matrix of interior-point methods [7]). The normal equations matrix is thus:

$$
\begin{aligned}
A\Theta A^T &= \left[\begin{array}{ccc|c}
N_1\Theta_1 N_1^T & & & N_1\Theta_1 L_1^T \\
& \ddots & & \vdots \\
& & N_k\Theta_k N_k^T & N_k\Theta_k L_k^T \\
\hline
L_1\Theta_1 N_1^T & \ldots & L_k\Theta_k N_k^T & \Theta_0 + \sum_{i=1}^k L_i\Theta_i L_i^T
\end{array}\right] \\
&= \begin{bmatrix} B & C \\ C^T & D \end{bmatrix},
\end{aligned} \tag{2}
$$

$B \in \mathbb{R}^{\tilde{n} \times \tilde{n}}$ ($\tilde{n} = \sum_{i=1}^k n_i$), $C \in \mathbb{R}^{\tilde{n} \times l}$ and $D \in \mathbb{R}^{l \times l}$ being the blocks of $A\Theta A^T$, and Θ_i, $i = 0, \ldots, k$, the submatrices of Θ associated with the $k+1$ groups of variables in (1).

Normal equations are solved by combining Cholesky factorizations for system with matrix B (actually, k small factorizations for $N_i\Theta_i N_i^T$), and preconditioned conjugate gradient (PCG) for system with the Schur complement $D - C^T B^{-1} C$. In [2] it was proved that, under some conditions, which are guaranteed in our setting, the inverse of $(D - C^T B^{-1} C)$ can be computed as

$$(D - C^T B^{-1} C)^{-1} = \left(\sum_{i=0}^{\infty} (D^{-1}(C^T B^{-1} C))^i \right) D^{-1}. \tag{3}$$

The preconditioner M^{-1}, an approximation of $(D - C^T B^{-1} C)^{-1}$, is thus obtained by truncating the infinite power series at some low-term h.

The effectiveness of the preconditioner depends on the spectral radius of matrix $D^{-1}(C^T B^{-1} C)$, which is always in $[0, 1)$ [2, Theorem 1]. The farther away from 1 is the spectral radius of $D^{-1}(C^T B^{-1} C)$ the better is the quality of the approximation of (3). In practice it was observed that when a quadratic term is present the spectral radius tends to be smaller than that obtained in the simplified linear formulation obtained by removing this quadratic objective term, and the preconditioner become more efficient.

The purpose of this work is twofold. First, to show theoretically that a quadratic term in the objective reduces the spectral radius, significantly improving the overall performance in some classes of instances. Second, to consider the solution of linear multicommodity problem by quadratic multicommodity flows through the addition of a regularization term in the barrier function.

The main result to show the first above goal is the following Theorem:

Theorem 1 *Let A be the constraint matrix of problem (1), with full row rank matrices $N_i \in \mathbb{R}^{m_i \times n_i}$ $i = 1, \ldots, k$, and at least one full row rank matrix $L_i \in \mathbb{R}^{l \times n_i}$, $i = 1, \ldots, k$. Let Θ be a symmetric diagonal matrix, and $B \in \mathbb{R}^{\tilde{n} \times \tilde{n}}$ ($\tilde{n} = \sum_{i=1}^k n_i$), $C \in \mathbb{R}^{\tilde{n} \times l}$ and $D \in \mathbb{R}^{l \times l}$ the submatrices of $A\Theta A^T$ defined in (2). Then, the spectral radius ρ of $D^{-1}(C^T B^{-1} C)$ is bounded by*

$$0 \leq \rho \leq \max_{j \in \{1, \ldots, l\}} \frac{\gamma_j}{\left(\frac{u_j}{v_j}\right)^2 \Theta_{0j} + \gamma_j} < 1, \tag{4}$$

Table 1: Results for a large (10 million variables, 210000 constraints) quadratic multicommodity flows problem instance, with quadratic objective function $\frac{1}{2}x^TQx$, for different $Q = \beta I$

	CPLEX-11		IPM			
β	it.	CPU	it.	PCG	CPU	f^*
0.01	7	29939	10	36	66	-2.6715e+08
0.1	7	31328	9	40	61	-2.6715e+09
1	8	33367	8	38	56	-2.6715e+10
10	9	35220	7	37	51	-2.6715e+11

CPU time in seconds, on Dell PowerEdge 6950 server with four dual core AMD Opteron 8222 3.0 GHZ processors(without exploitation of parallelism capabilities) and 64 GB of RAM

where u is the eigenvector of $D^{-1}(C^TB^{-1}C)$ that has ρ as eigenvalue ; γ_j,$j = 1,\dots,l$, and $V = [V_1,\dots,V_l]$, are respectively the eigenvalues and matrix of eigenvectors of $\sum_{i=1}^{k} L_i\Theta_i {L_i}^T$; $v = V^Tu$; and, abusing of notation, we assume that for $v_j = 0$, $(u_j/v_j)^2 = +\infty$.

By the above theorem it can be proved that the spectral radius tends to 0 when Q_i, $i = 1,\dots,k$, (i.e., the quadratic costs of variables, excluding slacks) tends to infinity. Therefore adding large enough Q_i, $i = 1,\dots,k$, to a linear problem, it is possible to reduce (actually to approach 0) the spectral radius of matrix $D^{-1}(C^TB^{-1}C)$, and thus to improve the quality of the preconditioner. It also explains the good behaviour of the specialized interior-point method in instances of Table 1: since that is a quadratic multicommodity problem, without linear term, $\min \frac{1}{2}x^TQx = \min \frac{1}{2}x^T(\beta Q)x$ for any positive $\beta \in \mathbb{R}$; therefore, the spectral radius is effectively reduced, and PCG solves the Schur complement system very efficiently. Instances of Table 1 correspond to instance CTA-100-100-1000, using four different scaling factors β and $Q = I$ (see [4] for a description of the underlying statistical three-dimensional tabular data protection problem, which is equivalent to a quadratic saturated multicommodity flow problem). The resulting multicommodity problem has 10,000,000 variables, and 210,000 constraints. The specialized interior-point approach was not only much more efficient than CPLEX-11 in terms of CPU time, but also in memory requirements: it needed 1.2 GB of RAM, while CPLEX-11 required 15 GB. Both codes successfully solved the problem, with relative differences in the objective function of about 10^{-11}.

For linear problems, however, the addition of quadratic terms with large Q_i, $i = 1,\dots,k$, is meaningless, and only small regularizations are used in practice [1, 6]. It is proved in this work that, under some conditions, the bound (4) on the spectral radius for a linear problem is reduced by adding (even small) quadratic costs Q_i, $i = 1,\dots,k$. Since both the bound and the spectral radius approach 1 in the last iterations of the interior-point method, a reduction in the bound also means a reduction in the spectral radius:

Proposition 1 *Let assume the hypotheses of Theorem 1, and consider a linear problem and a quadratic one obtained by adding (likely small) quadratic costs $Q_i > 0$, $i = 1,\dots,k$. Assume $\hat{u}_j/\hat{v}_j \leq u_j/v_j$, $j = 1,\dots,l$, where "hatted" and "non-hatted" terms refer, respectively, to the linear and quadratic problems. Then bound (4) is smaller for the quadratic than for the linear problem.*

The strong assumption $\hat{u}_j/\hat{v}_j \leq u_j/v_j$, $j = 1,\dots,l$ may not be guaranteed for any primal block-angular problem, but it is when L_i is diagonal, i.e., it is for multicommodity flows problems. This means that adding a quadratic regularization we can improve the solution of linear multicommodity network flows problems.

Based on the previous result, we regularize the original multicommodity flow problem. Instead of using regularizations based on proximal-points, like that of [1]

$$B_P(x,\mu) \triangleq c^Tx + \frac{1}{2}(x-\bar{x})^TQ_P(x-\bar{x}) - \mu\left(\sum_{i=1}^{n}\ln x_i - \sum_{i=1}^{n}\ln(u_i - x_i)\right), \tag{5}$$

Table 2: Results with IPM, RIPM and CPLEX11 for difficult multicommodity instances

	IPM			RIPM			CPLEX11	
Instance	it.	PCG	CPU	it.	PCG	CPU	it.	CPU
tripart1	58	1976	1.7	89	721	**1.57**	21	3.99
tripart2	87	4092	17.3	97	1562	**10.2**	25	36.01
tripart3	90	6978	62.4	106	3178	**36.4**	28	138.8
tripart4	133	14660	265	136	4947	**128**	29	1323.2
gridgen1	242	96877	7400	219	5703	**618**	64	12288

CPU time on SUN Fire V20Z server with AMD Opteron 2.46 GHz processor and 8GB RAM

we suggest the alternative one

$$B_Q(x,\mu) \triangleq c^T x + \mu \left(\frac{1}{2} x^T Q x - \sum_{i=1}^{n} \ln x_i - \sum_{i=1}^{n} \ln(u_i - x_i) \right), \tag{6}$$

Q and Q_P being diagonal positive definite matrices. The regularization term of (6) tends to zero with μ. In practice a value of $Q = qI$, $q > 0$, may be used. Moreover, we proved the resulting barrier is still self-concordant, so the interior-point algorithm converges and has polynomial time complexity. In addition, if Q is small enough (i.e., $q_i \leq 1/u_i^2$) then the complexity is the same than for the non-regularized interior-point algorithm; larger Q's increase in theory the number of interior-point iterations. There is then a trade-off between reducing the the number of conjugate gradient iterations (by using large Q's) and the number of interior-point iterations (by using small Q's). The recent computational results provided in the talk shall show that for a rather broad range of Q's the regularized approach is more efficient than the standard one. For instance, Table 2 shows results for some well-known multicommodity instances in the literature [3] with the original specialized interior-point approach (IPM), the regularized version (RIPM) and CPLEX11.

Additional details can be found in [5].

References

[1] A. Altman, J. Gondzio, Regularized symmetric indefinite systems in interior point methods for linear and quadratic optimization. *Optim. Methods Software* 11, 275–302 (1999).

[2] J. Castro, A specialized interior-point algorithm for multicommodity network flows. *SIAM J. Optim.* 10, 852–877 (2000).

[3] J. Castro, Solving difficult multicommodity problems through a specialized interior-point algorithm. *Ann. Oper. Res.* 124, 35–48 (2003).

[4] J. Castro, An interior-point approach for primal block-angular problems. *Comput. Optim. Appl.*, 36, 195–219 (2007).

[5] J. Castro, J. Cuesta, Quadratic regularizations in an interior-point method for primal block-angular problems. Research Report DR2008-07, Dept. of Statistics and Operations Research, Universitat Politècnica de Catalunya (2008). Submitted to *Math. Prog.*, first revision in progress.

[6] M.A. Saunders, J.A. Tomlin, Solving regularized linear programs using barrier methods and KKT systems. Technical Report SOL 96-4. Dept. of Operations Research, Stanford University (1996).

[7] S.J. Wright, Primal-Dual Interior-Point Methods. SIAM, Philadelphia (1996).