Steplengths in interior point algorithms of quadratic programming *

Csaba Mészáros †

Computer and Automation Research Institute,
Hungarian Academy of Sciences,
Budapest

Abstract

An approach to determine primal and dual stepsizes in the infeasible–interior–point primal–dual method for convex quadratic problems is presented. The approach reduces the primal and dual infeasibilities in each step and allows different stepsizes. The method is derived by investigating the efficient set of a multiobjective optimization problem. Computational results are also given.

Keywords: interior point methods, quadratic programming, steplength, efficient set

1 Introduction

In the paper we will assume the convex quadratic problem (QP) in the form:

$$\begin{array}{rc} \min & c^T x + \frac{1}{2} x^T Q x, \\ \text{subject to} & Ax = b, \\ & x \geq 0, \end{array} \tag{1}$$

*This work was supported in part by EPSRC grant No. GR/J52655 and Hungarian Research Fund OTKA T-016413.

†H-1518 Budaspest, P.O. BOX 63. Hungary

where $A \in R^{m \times n}$ is of full row rank, $Q \in R^{n \times n}$ is symmetric positive semidefinite and $c, x \in R^n$, $b \in R^m$. The dual of (1) in the Wolfe sense is defined as follows:

$$\begin{array}{rc} \max & b^T y - \frac{1}{2} x^T Q x, \\ \text{subject to} & A^T y + z - Qx = c, \\ & z \geq 0, \end{array} \tag{2}$$

where $z \in R^n$ and $y \in R^m$.

In the paper we focus on one particular aspect of infeasible–interior–point methods, namely on the determination of the steplengths in each iteration. Our main motivation for this investigation is the practice of linear programming (LP) where using different steplengths in the primal and dual space is a standard implementation technique. While it is not supported by theoretical results, the use of different steplengths in linear programming increases the practical efficiency of the infeasible–interior–point primal–dual methods [1].

In LP dual feasibility constraints are independent of primal variables. In QP, however, matrix Q connects the dual feasibility to the primal problem. That is why interior point implementations of quadratic programming are restricted to the use of a common steplength in the primal and dual spaces [2, 9].

We give a simple and computationally cheap procedure to compute different stepsizes in the infeasible–interior–point primal–dual methods of quadratic programming. One variant of our algorithm guarantees that the determined steplengths make at least as good progress in both primal and dual feasibility as the common stepsize, and sometimes performs much better. We study one particular multiobjective optimization problem in which the squared norm of primal and dual infeasibilities are the objective functions with respect to the primal and dual stepsizes under box constraints. We show that our algorithm results in an efficient point of this multiobjective problem.

The paper is organized as follows: Section 2 gives a short review of the infeasible–interior–point primal–dual method for quadratic programming. Section 3 describes our method for determining the steplengths and discusses some relevant questions. Section 4 contains a computational comparison between our suggested method and the traditional technique. We summarize our findings in Section 5.

2 The infeasible–interior–point primal–dual algorithm for quadratic programming

Infeasible–interior–point primal–dual methods are considered most effective approaches for solving large-scale problems. Whereas the results of Monterio and Zhou [8] show the superlinear convergence behavior of these methods, computational practice indicates their usefulness in practice [7, 3].

Primal–dual interior point algorithms are iterative approaches which seek to satisfy the Karush–Kuhn–Tucker type optimality conditions for the primal–dual problem (1-2):

$$\begin{aligned} Ax &= b, && (3)\\ A^T y + z - Qx &= c, && (4)\\ Xz &= 0, && (5)\\ (x,z) &\geq 0, \end{aligned}$$

where $X = diag(x_1, ..., x_n)$. The infeasible–interior–point primal–dual methods generate a sequence of iterates in both the primal and dual spaces

$$(x^k, y^k, z^k) \qquad k = 0, 1, 2, \dots ,$$

which fulfil the strict positivity condition $(x^k, z^k) > 0$, but feasibility (3,4) and complementarity (5) are reached as $k \longrightarrow \infty$.

Infeasible–interior–point primal–dual algorithm can be derived by perturbing the complementarity conditions (5) and applying Newton's method to solve the nonlinear system of the first order optimality conditions.

Performing these steps, one can obtain the perturbed Karush-Kuhn-Tucker system of (1–2) as

$$\begin{aligned} Ax &= b,\\ A^T y + z - Qx &= c,\\ XZe &= \mu e, \end{aligned} \qquad (6)$$

where e is the vector having all its coordinates equal to one and $\mu \geq 0$ is the logarithmic barrier parameter.

In one iteration of the primal-dual algorithm one step of Newton's method is applied to the first order optimality conditions (6) with a given μ and then μ is decreased. The algorithm terminates when the infeasibility and the complementarity gap are reduced below predetermined tolerances.

Given an $x, z \in \mathcal{R}^n_+, \quad y \in \mathcal{R}^m$, Newton's direction is obtained by solving the following system of linear equations

$$\begin{bmatrix} A & 0 & 0 \\ -Q & A^T & I \\ Z & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta z \end{bmatrix} = \begin{bmatrix} \xi_b \\ \xi_c \\ \mu e - XZe \end{bmatrix}, \tag{7}$$

where $\xi_b = b - Ax$, and $\xi_c = c - A^T y - z + Qx$. If A is of full row rank then this system has a unique solution.

Once the system (7) has been solved, the maximum allowable stepsizes in the primal space ($\bar{\alpha}_P$) and the dual space ($\bar{\alpha}_D$) are computed such that the nonnegativity of the variables is preserved:

$$\bar{\alpha}_P = \frac{1}{\max\limits_{k=1\ldots n} \left\{1, -\frac{\Delta x_k}{x_k}\right\}}, \tag{8}$$

$$\bar{\alpha}_D = \frac{1}{\max\limits_{k=1\ldots n} \left\{1, -\frac{\Delta z_k}{z_k}\right\}}. \tag{9}$$

To ensure decrease in both the primal and dual infeasibilities the common steplength in the primal and dual spaces is defined as

$$\alpha = \min\left(\bar{\alpha}_P, \bar{\alpha}_D\right). \tag{10}$$

This stepsize is slightly reduced by a factor $0 < \alpha_0 < 1$ to prevent hitting the boundary. Finally, a new iterate is computed as

$$\begin{aligned} x &\leftarrow x + \alpha_0 \alpha \Delta x, \\ y &\leftarrow y + \alpha_0 \alpha \Delta y, \\ z &\leftarrow z + \alpha_0 \alpha \Delta z. \end{aligned}$$

After taking the step, the barrier parameter μ is decreased by a given factor and the process is repeated.

Theoretical results, as well computational practice, show that complementarity (5) should not be approached faster than feasibility (3,4) [4, 5]. Otherwise, the iterates converge close to the nonnegativity boundary, still far away from the feasible region. Another argument for the importance of the rapid reduction of infeasibilities is that the decrease in the complementarity is guaranteed in theory only if the iterates are feasible.

Let us note that damping the largest step as (10) in QP is necessary, otherwise, the decrease in the dual infeasibility is not guaranteed and the algorithm may diverge. It is also noted that in linear programming, where $Q = 0_{n\times n}$ and consequently dual feasibility is not connected to primal variables, the determination of different steplengths can be decoupled. By allowing different stepsizes, the progress in the feasibility is not damped in one of the spaces when only a small step is possible in the other. This speeds up the convergence to feasible points, which increases the efficiency of the algorithm. This behavior was observed even in the first implementations of interior point methods for linear programming and the technique has become a commonly used standard [1, 10].

3 Steplength strategies

Several variants of the infeasible–interior–point primal–dual methods have been developed during the past few years (see [1, 10]). Usually, they differ in the centralization term while the feasibility terms are the same as in (7), therefore, infeasible–interior–point primal–dual methods compute such search directions which satisfy the conditions

$$
\begin{aligned}
A\Delta x &= \xi_b, && (11)\\
A^T\Delta y + \Delta z - Q\Delta x &= \xi_c. && (12)
\end{aligned}
$$

In other words, $(\Delta x, \Delta y, \Delta z)$ are descent directions for $\|b - Ax\|$ and $\|c - A^T y - z + Qx\|$ where $\| \, . \, \|$ denotes the Euclidean norm. As mentioned in the previous section, it is seldom the case that a full step can be made without violating the nonnegativity of x and z. Therefore, stepsizes have to be defined for the Newton direction as (8) and (9). The simplest way to ensure that both primal and dual infeasibilities decrease is to use the common steplength (10) in primal and dual. We call this approach *"simple damping"*. Although simple damping guarantees that both the primal and dual infeasibilities are decreased by factor $(1 - \alpha)$, reducing the larger stepsize to the value of the smaller one is probably not the most efficient approach.

In what follows, we study the behavior of the following two functions:

$$f_P(\alpha_P) = \|b - A(x + \alpha_P \Delta x)\|^2 \tag{13}$$

and

$$f_D(\alpha_P, \alpha_D) = \|c - A^T(y + \alpha_D \Delta y) - (z + \alpha_D \Delta z) + Q(x + \alpha_P \Delta x)\|^2 \tag{14}$$

in the

$$\bar{\alpha}_P \geq \alpha_P \geq 0, \tag{15}$$
$$\bar{\alpha}_D \geq \alpha_D \geq 0 \tag{16}$$

intervals where $\bar{\alpha}_P$, furthermore, $\bar{\alpha}_D$ are defined by (8) and (9). Since the goal is to reduce infeasibilities and the complementarity gap, we want to minimize both f_P and f_D in the interval (15-16). To achieve these goals and determine "optimal" stepsizes a natural idea is to consider the quadratic multiobjective problem as

$$\begin{array}{rc} \min & f_P(\alpha_P), \\ \min & f_D(\alpha_P, \alpha_D), \\ \text{subject to} & \bar{\alpha}_P \geq \alpha_P \geq 0, \\ & \bar{\alpha}_D \geq \alpha_D \geq 0. \end{array} \tag{17}$$

Since the minimization of the two functions at the same time may be impossible, we will investigate the *efficient set* of multiobjective problem (17). By the efficient set of (17) we mean a set of *efficient points* $(\hat{\alpha}_P, \hat{\alpha}_D) \in [0, \bar{\alpha}_P] \times [0, \bar{\alpha}_D]$ with the following properties:
For any $(\alpha_P, \alpha_D) \in [0, \bar{\alpha}_P] \times [0, \bar{\alpha}_D]$

$$\text{if } \left\{ \begin{array}{ccc} f_P(\alpha_P) < f_P(\hat{\alpha}_P) & \Longrightarrow & f_D(\hat{\alpha}_P, \hat{\alpha}_D) < f_D(\alpha_P, \alpha_D), \\ f_D(\alpha_P, \alpha_D) < f_D(\hat{\alpha}_P, \hat{\alpha}_D) & \Longrightarrow & f_P(\hat{\alpha}_P) < f_P(\alpha_P). \end{array} \right.$$

In our study we assume that the iterate is infeasible and the quadratic term influences the dual feasibility, i.e.

$$\xi_b \neq 0, \ \xi_c \neq 0 \text{ and } Q\Delta x \neq 0.$$

Let us note that these assumptions are "automatically" fulfilled in practice, at least because machine precision is finite. Next, we describe some properties of functions (13) and (14).

Proposition 1: Functions f_P and f_D are convex quadratic. Further, if $\xi_b \neq 0$, then f_P is strictly convex; if $Q\Delta x \neq 0$, then f_D is strictly convex. In the latter cases the steplength one gives the unique minimum.
Proof: It follows from (11) that under the $\xi_b \neq 0$ assumption $f_P(\alpha_P) = 0$ if and only if $\alpha_P = 1$. Furthermore, (12) shows that if $\xi_c \neq 0$ and $Q\Delta x \neq 0$, then $f_D(\alpha_P, \alpha_D) = 0$ if and only if $\alpha_P = 1$ and $\alpha_D = 1$. ■

Proposition 2: The minimum of f_D in the interval (15–16) is unique and if (α_P^*, α_D^*) is the minimizer of f_D in (15–16), then

$$(\alpha_P^*, \alpha_D^*) \in \{(\bar{\alpha}_P, \alpha_D) \mid \bar{\alpha}_D \geq \alpha_D \geq 0\} \cup \{(\alpha_P, \bar{\alpha}_D) \mid \bar{\alpha}_P \geq \alpha_P \geq 0\}.$$

In other words, the minimum is at the boundary of the interval and at least one variable has to be at its upper bound.

Proof: Since $\bar{\alpha}_P \leq 1$ and $\bar{\alpha}_P \leq 1$, the global optimum of f_D is outside the interior of the feasible interval and the proposition follows from the strict convexity of f_D. ■

Furthermore, let (α_P^*, α_D^*) denote the unique minimizer of f_D in the interval given by (15–16). Now, we describe the efficient set of problem (17):

Theorem 1: If $\alpha_D^* < \bar{\alpha}_D$, then (α_P^*, α_D^*) is the only efficient point of (17). If $\alpha_P^* < \bar{\alpha}_P$, then the efficient set of (17) is

$$\{(\alpha_P, \alpha_D^*) \mid \alpha_P^* \leq \alpha_P \leq \bar{\alpha}_P\}.$$

Proof: If $\alpha_D^* < \bar{\alpha}_D$, then (α_P^*, α_D^*) minimizes both f_P and f_D in (15–16) and Proposition 1 shows that (α_P^*, α_D^*) is the unique minimizer of f_D. From Proposition 2 it follows that if $\alpha_P^* < \bar{\alpha}_P$, then $\alpha_D^* = \bar{\alpha}_D$. As a consequence of Proposition 2, the efficient set in this case is a subset of $\{(\alpha_P, \bar{\alpha}_D) \mid \bar{\alpha}_P \geq \alpha_P \geq \alpha_P^*\}$. Let us observe that $f_P(\alpha_P^*) > f_P(\alpha_P)$ if and only if $\alpha_P^* \leq \alpha_P$, from which the theorem follows. ■

Let us note that using the simple damping may result in steplengths which do not belong to the efficient set of (17). This can occur when $\alpha_P^* > \alpha$ or $\alpha_D^* \neq \alpha$.

As a consequence of Proposition 2, α_P^* and α_D^* can be computed by fixing either α_P or α_D to its upper bound and solving one–dimensional quadratic minimizations. Because it is trivial, its technical details are not described here.

We suggest two different choices for the steplengths. One is to use (α_P^*, α_D^*) which has the property that it is from the efficient set of (17) and the step minimizes the dual infeasibility. The other suggested choice is using $\max(\alpha, \alpha_P^*)$ for primal, and α_D^* for dual stepsize. It follows from Theorem 1 that this latter is also from the efficient set of (17). While none of the objectives may be optimal, this choice guarantees that the decrease in both primal and dual infeasibilities is not smaller than that of by the simple damping.

To demonstrate that the larger steplength does not necessarily have to be damped down to the smaller one, we prepared figures 1 and 2. Figure 1 shows the maximum allowable steplengths (i.e. $\bar{\alpha}_P$ and $\bar{\alpha}_D$) with respect

Figure 1: Maximum steplengths on problem *qffffff80*

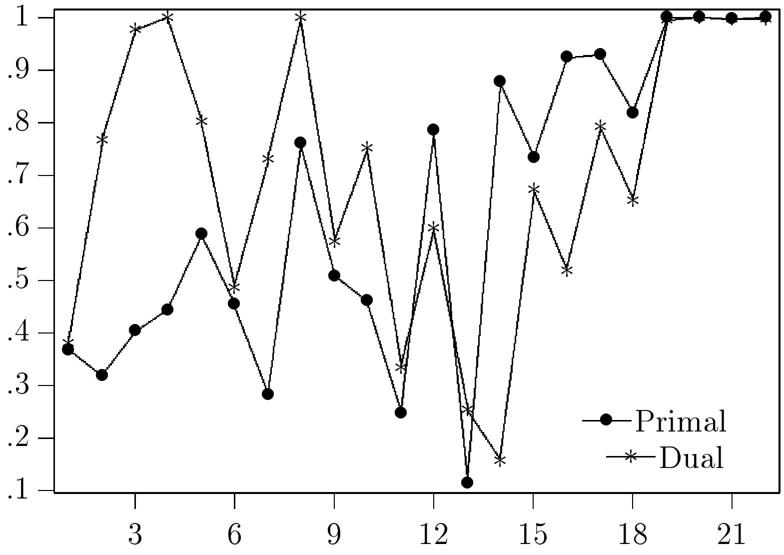
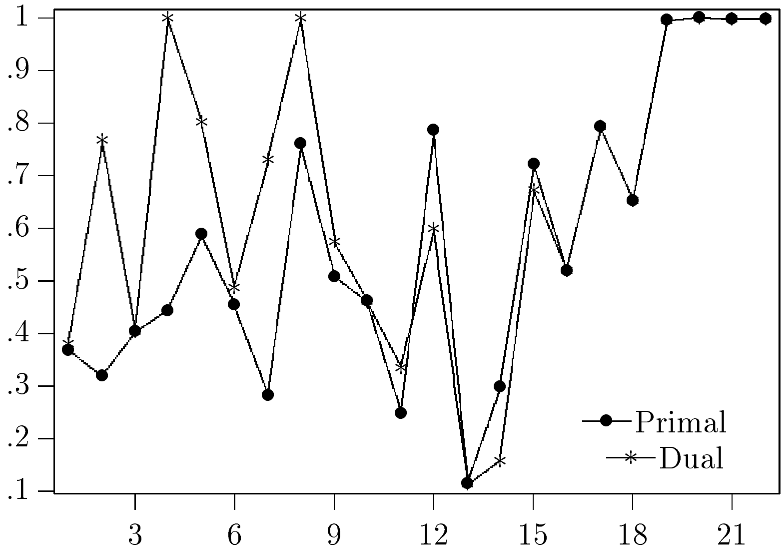


Figure 2: Damped steplengths on problem *qffffff80*



to the iteration count. Simple damping would result in their minimum in each iteration. Figure 2 shows the effect of our second suggested method, i.e. $\max(\alpha, \alpha_P^*)$ and α_D^*. It indicates that different steplengths can be used, mainly at the first stage of the iterations where the iterates are considerably infeasible. Note that the difference between the primal and dual stepsizes decreases and vanishes as feasibility is approached.

As it was pointed out in [4], for the infeasible–interior–point primal–dual algorithm some safeguard techniques regarding the steplength selection have to be included to ensure global convergence. Such safeguard techniques have been proved to be practically importand for linear programming [5]. In our implementation we used the conditions described by Kojima et al. [4]:

$$\begin{aligned} x_j z_j &> \gamma \frac{x^T z}{n}, \;\; 1 \le j \le n, && (18)\\ x^T z &\ge \gamma_P \|Ax - b\| \;\text{ or }\; \|Ax - b\| \le \epsilon_P, && (19)\\ x^T z &\ge \gamma_d \left\|A^T y + z - c - Qx\right\| \;\text{ or }\; \left\|A^T y + z - c - Qx\right\| \le \epsilon_d && (20) \end{aligned}$$

with the following parameter values:

$$\begin{array}{ll} \gamma = 10^{-3}, & \\ \gamma_P = 10^{-6} \left\|Ax^0 - b\right\|, & \epsilon_p = 10^{-8}(\|b\| + 1), \\ \gamma_P = 10^{-6} \left\|A^T y^0 + z^0 - c - Qx^0\right\|, & \epsilon_d = 10^{-8}(\|c\| + 1), \end{array}$$

where (x^0, y^0, z^0) is the starting point. We selected a starting point satisfying (18) and during iterations the steplengths were reduced until the conditions (18–20) are satisfied. We observed that the most important condition is (18) since in our experiments conditions (19,20) were automatically satisfied whenever (18) was satisfied. We observed that modifying the steplengts to ensure (18–20) was rarely necessary and required only a small reduction in the stepsize.

Let us note that if the iterate is primal and dual feasible, then our procedure automatically selects equal steplengths. Contrary to linear programming, however, in the QP case the complementarity gap does not decrease throughout as the steplength increases since $\Delta x^T \Delta z \neq 0$. If the iterate is primal and dual feasible, we truncate the steplenghts at

$$\tilde{\alpha} = -\frac{(\Delta x^T z + x^T \Delta z)}{2 \Delta x^T \Delta z}, \tag{21}$$

which minimizes

$$(x + \alpha \Delta x)^T (z + \alpha \Delta z) = x^T z + \alpha(\Delta x^T z + x^T \Delta z) + \alpha^2 \Delta x^T \Delta z.$$

It is easy to see that if $\xi_b = 0$ and $\xi_c = 0$, then

$$\Delta x^T \Delta z = \Delta x^T Q \Delta x,$$

and

$$\Delta x^T z + x^T \Delta z = n\mu - x^T z,$$

which shows that $(x + \alpha\Delta x)^T (z + \alpha\Delta z)$ decreases at $\alpha = 0$ if $\mu < \frac{x^T z}{n}$ and has a unique minimum at (21) if $Q\Delta x \neq 0$.

4 Computational results

We demonstrate the effects of the discussed steplength strategies by solving quadratic programming problems from the QP test set [6]. The problems were solved to 10^{-8} relative accuracy on an IBM PC Pentium 200 Mhz machine with 64 MB of memory. In our experiment we compare the iteration counts and the total solution times taken by the different steplength methods. Table 1 shows the results. Its first column contains the name of the problems. The iteration counts and execution times in seconds are given in columns 2-4 and 5-7, respectively. The execution time includes all parts of the solution process, comprising scaling, presolving and postsolving. Columns labelled "damped" refer to the simple damping, whereas "method1" and "method2" denotes our technique with steplengths (α_P^*, α_D^*) and $(\max(\alpha, \alpha_P^*), \alpha_D^*)$, respectively.

The computational results indicate that our method performs better in the infeasible–interior–point primal–dual algorithm than the simple damping. Since our approach results in different steps from the simple damping if the current iterate is "infeasible enough", its performance depends on how fast the critical feasibility level is achieved. As the results suggest, there are practically only minor differences between the two variants of our method.

5 Conclusion

In the paper we suggested a method for computing different primal and dual stepsizes in the infeasible–interior–point primal–dual methods of quadratic programming. Whereas in linear programming the use of different steplengths is straightforward and simple, it makes additional analysis in the quadratic case necessary. As the computational results indicate, our methods outperforms the traditional simple damping approach because it can reduce the infeasibility faster. Our methods turn to be equivalent to the simple damping

Table 1: Comparison of different steplength strategies

Problem	Iterations			Solution time		
name	dampf	method 1	method2	dampf	method 1	method 2
q25fv47	20	18	18	28.32	26.27	26.32
qetamacr	19	15	15	3.79	3.18	3.08
qfffff80	28	22	21	3.73	2.96	2.97
qgrow15	26	20	19	2.14	1.75	1.70
qgrow22	29	22	22	3.02	2.42	2.42
qisrael	22	18	17	2.26	1.93	1.75
qpilotno	27	23	22	11.53	9.89	9.45
qscfxm2	30	25	23	2.53	2.15	1.98
qscfxm3	32	26	24	3.68	3.08	2.97
qscrs8	25	18	18	1.59	1.21	1.21
qscsd8	17	11	11	2.47	1.87	1.81
qsctap2	21	14	13	3.24	2.30	2.31
qsctap3	22	14	14	4.67	3.13	3.24
qshell	28	23	23	2.41	2.03	2.14
qship04l	16	10	10	1.42	1.10	1.05
qship08l	14	10	10	15.27	10.60	11.42
qship12l	16	12	12	14.06	11.10	11.10
qsierra	21	16	15	3.02	2.48	2.36
qstair	21	14	15	3.13	2.31	2.36
cvxqp1_m	14	9	9	22.63	16.80	16.75
cvxqp2_m	15	10	10	10.66	7.97	7.97
cvxqp3_m	14	10	10	45.64	35.70	35.43
hues-mod	26	19	20	7.25	5.60	5.88
huestis	26	19	20	7.36	5.60	5.93
mosarqp1	12	8	8	2.09	1.59	1.59
mosarqp2	13	9	9	1.65	1.26	1.26
stcqp1	9	8	8	1.54	1.54	1.54
stcqp2	9	9	9	2.46	2.47	2.47

if the iterate is feasible (or close feasible). Its effectiveness depends on the infeasibility of the iterates as well.

References

[1] E.D. Andersen, J. Gondzio, Cs. Mészáros, and X. Xu. Implementation of interior point methods for large scale linear programs. In T. Terlaky, editor, *Interior point methods of mathematical programming*, pages 189–252. Kluwer Academic Publishers, 1996.

[2] M.G. Breitfeld and D.F. Shanno. Preliminary computational experience with modified log-barrier functions for large-scale nonlinear programming. In W.W. Hager, D.W. Hearn, and P.M. Pardalos, editors, *Large scale optimization: Sate of the art*, pages 45–66. Kluwer Academic Publishers, London, 1994.

[3] T.J. Carpenter, I.J. Lustig, J.M. Mulvey, and D.F. Shanno. Higher order predictor-corrector interior point methods with application to quadratic objectives. *SIAM Journal on Optim.*, 3:696–725, 1993.

[4] M. Kojima, N. Megiddo, and S. Mizuno. A primal-dual infeasible-interior-point algorithm for linear programming. *Math. Programming*, 61:263–280, 1993.

[5] I. J. Lustig, R. E. Marsten, and D. F. Shanno. Computational experience with a globally convergent primal-dual predictor-corrector algorithm for linear programming. *Math. Programming*, 66(1):123–135, 1994.

[6] I. Maros and Cs. Mészáros. A repository of convex quadratic programming problems. Technical Report DOC 97/6, Imperial College, Department of Computing, London, 1997. submitted to Optimization Methods and Software, the paper and the test problems are available from ftp://ftp.sztaki.hu/pub/oplab/QPSET.

[7] R. D. C. Monteiro and I. Adler. Interior path following primal-dual algorithms. Part II: Convex quadratic programming. *Math. Programming*, 44:43–66, 1989.

[8] R.D.C. Monteiro and F. Zhou. On superlinear convergence of infeasible–interior–point algorithms for linearly constrained convex programs. *Computational Optimization and Applications*, (8):245–262, 1997.

[9] R. Vanderbei. LOQO user's manual. Technical Report SOR-96-07, Princeton University, School of Engineering and Applied Science, Dept. of Civil Engineering and Op.Res., 1996.

[10] S. J. Wright. *Primal–dual interior point methods.* SIAM, T57.74.W75, University Science Center, Philadelphia, 1997.