The homogeneous and self-dual model and algorithm for linear optimization. MOSEK Technical report: TR-1-2009.

Erling D. Andersen*

14-April-2009, (Last revised: 02-June-2009)

Abstract

In this white paper we present the homogeneous and self-dual interior point methods which forms the basis for several commercial optimization software packages such as MOSEK.

1 Introduction

The linear optimization problem

$$\begin{array}{ll} \text{min.} & c^T x \\ \text{s.t.} & Ax = b, \\ & x \geq 0 \end{array} \tag{1}$$

may have an optimal solution, be primal infeasible or be dual infeasible for a particular set of data $c \in R^n$, $b \in R^m$, and $A \in R^{m \times n}$. In fact the problem can be both primal and dual infeasible for some data where (1) is denoted dual infeasible if the dual problem

$$\begin{array}{ll} \text{max} & b^T y \\ \text{s.t.} & A^T y + s = c, \\ & s \geq 0 \end{array} \tag{2}$$

corresponding to (1) is infeasible. The vector s is the so-called dual slacks.

2 The homogenous and self dual model

However, most methods for solving (1) assume that the problem has an optimal solution. This is in particular true for interior-point methods. To overcome this

*MOSEK ApS, Frubjergvej 3, Box 16, 2100 Copenhagen, Denmark. Email: e.d.andersen@mosek.com

problem it has been suggested to solve the homogeneous and self-dual model

$$
\begin{array}{lrrrrl}
\min & 0 & & & & \\
\text{s.t.} & & Ax & -b\tau & = & 0, \\
& -A^T y & & +c\tau & \geq & 0, \\
& b^T y & -c^T x & & \geq & 0, \\
& & x \geq 0, & \tau \geq 0 & &
\end{array} \tag{3}
$$

instead of (1). Clearly, (3) is a homogeneous LP and is self-dual which essentially follows from the constraints form a skew-symmetric system. The interpretation of (3) is τ is a homogenizing variable and the constraints represent primal feasibility, dual feasibility, and reversed weak duality.

The homogeneous model (3) was first studied by Goldman and Tucker [2] in 1956 and they proved (3) always has a nontrivial solution (x^*, y^*, τ^*) satisfying

$$
\begin{array}{rclrcl}
x_j^* s_j^* & = & 0, & x_j^* + s_j^* & > & 0, \quad \forall j, \\
\tau^* \kappa^* & = & 0, & \tau^* + \kappa^* & > & 0,
\end{array} \tag{4}
$$

where $s^* := c\tau^* - A^T y^* \geq 0$ and $\kappa^* := b^T y^* - c^T x^* \geq 0$. A solution to (3) satisfying the condition (4) is said to be a strictly complementary solution. Moreover, Goldman and Tucker showed that if $(x^*, \tau^*, y^*, s^*, \kappa^*)$ is any strictly complementary solution, then exactly one of the two following situations occurs:

- $\tau^* > 0$ if and only if (1) has an optimal solution. In this case $(x^*, y^*, s^*)/\tau^*$ is an optimal primal-dual solution to (1).
- $\kappa^* > 0$ if and only if (1) is primal or dual infeasible. In the case $b^T y^* > 0$ ($c^T x^* < 0$) then (1) is primal (dual) infeasible.

The conclusion is that a strictly complementary solution to (3) provides all the information required, because in the case $\tau^* > 0$ then an optimal primal-dual solution to (1) is trivially given by $(x, y, s) = (x^*, y^*, s^*)/\tau^*$. Otherwise, the problem is primal or dual infeasible. Therefore, the main algorithmic idea is to compute a strictly complementary solution to (3) instead of solving (1) directly.

3 The homogenous algorithm

Ye, Todd, and Mizuno [6] suggested to solve (3) by solving the problem

$$
\begin{array}{lrrrrrl}
\min & & & & n^0 z & & \\
\text{s.t.} & & Ax & -b\tau & -\bar{b}z & = & 0, \\
& -A^T y & & +c\tau & +\bar{c}z & \geq & 0, \\
& b^T y & -c^T x & & +\bar{d}z & \geq & 0, \\
& \bar{b}^T y & -\bar{c}^T x & -\bar{d}\tau & & = & -n^0, \\
& & x \geq 0, & \tau \geq 0, & & &
\end{array} \tag{5}
$$

where

$$
\begin{array}{rcl}
\bar{b} & := & Ax^0 - b\tau^0, \\
\bar{c} & := & -c\tau^0 + A^T y^0 + s^0, \\
\bar{d} & := & c^T x^0 - b^T y^0 + \kappa^0, \\
n^0 & := & (x^0)^T s^0 + \tau^0 \kappa^0
\end{array}
$$

and

$$(x^0, \tau^0, y^0, s^0, \kappa^0) = (e, 1, 0, e, 1)$$

(e is a n vector of all ones). It can be proved that the problem (5) always has an optimal solution. Moreover, the optimal value is identical to zero and it is easy to verify that if (x, τ, y, z) is an optimal strictly complementary solution to (5), then (x, τ, y) is a strictly complementary solution to (3). Hence, the problem (5) can solved using any method that generates an optimal strictly complementary solution because the problem always has a solution. Note by construction then $(x, \tau, y, z) = (x^0, \tau^0, y^0, 1)$ is an interior feasible solution to (5). This implies that the problem (1) can be solved by most feasible-interior-point algorithms.

Xu, Hung, and Ye [4] suggest an alternative solution method which is also an interior-point algorithm, but specially adapted to the problem (3). The so-called homogeneous algorithm can be stated as follows:

1. Choose $(x^0, \tau^0, y^0, s^0, \kappa^0)$ such that $(x^0, \tau^0, s^0, \kappa^0) > 0$. Choose $\varepsilon_f, \varepsilon_g > 0$ and $\gamma \in (0, 1)$ and let $\eta := 1 - \gamma$.

2. k := 0.

3. Compute:
$$\begin{array}{rcl} r_p^k &:=& b\tau^k - Ax^k, \\ r_d^k &:=& c\tau^k - A^T y^k - s^k, \\ r_g^k &:=& \kappa^k + c^T x^k - b^T y^k, \\ \mu^k &:=& \frac{(x^k)^T s^k + \tau^k \kappa^k}{n+1}. \end{array}$$

4. If
$$||(r_p^k; r_d^k; r_g^k)|| \leq \varepsilon_f \quad \text{and} \quad \mu^k \leq \varepsilon_g,$$
then terminate.

5. Solve the linear equations
$$\begin{array}{rcl} Ad_x - bd_\tau &=& \eta r_p^k, \\ A^T d_y + d_s - cd_\tau &=& \eta r_d^k, \\ -c^T d_x + b^T d_y - d_\kappa &=& \eta r_g^k, \\ S^k d_x + X^k d_s &=& -X^k s^k + \gamma \mu^k e, \\ \kappa^k d_\tau + \tau^k d_\kappa &=& -\tau^k \kappa^k + \gamma \mu^k \end{array}$$
for $(d_x, d_\tau, d_y, d_s, d_\kappa)$ where $X^k := \text{diag}(x^k)$ and $S^k := \text{diag}(s^k)$.

6. For some $\theta \in (0, 1)$ let α^k be the optimal objective value to
$$\begin{array}{lcccl} \max & \theta\alpha & & & \\ \text{s.t.} & \begin{bmatrix} x^k \\ \tau^k \\ s^k \\ \kappa^k \end{bmatrix} + & \alpha \begin{bmatrix} d_x \\ d_\tau \\ d_s \\ d_\kappa \end{bmatrix} & \geq & 0, \\ & & \alpha & \leq & \theta^{-1}. \end{array}$$

7.
$$\begin{bmatrix} x^{k+1} \\ \tau^{k+1} \\ y^{k+1} \\ s^{k+1} \\ \kappa^{k+1} \end{bmatrix} := \begin{bmatrix} x^{k} \\ \tau^{k} \\ y^{k} \\ s^{k} \\ \kappa^{k} \end{bmatrix} + \alpha^k \begin{bmatrix} d_x \\ d_\tau \\ d_y \\ d_s \\ d_\kappa \end{bmatrix}.$$

8. $k = k + 1$.

9. goto 3

The following facts can be proved about the algorithm

$$\begin{array}{rcl} r_p^{k+1} & = & (1-(1-\gamma)\alpha^k)r_P^k, \\ r_d^{k+1} & = & (1-(1-\gamma)\alpha^k)r_D^k, \\ r_g^{k+1} & = & (1-(1-\gamma)\alpha^k)r_G^k, \end{array} \tag{6}$$

and

$$\begin{array}{c} ((x^{k+1})^T s^{k+1} + \tau^{k+1}\kappa^{k+1}) \\ = (1-(1-\gamma)\alpha^k)((x^k)^T s^k + \tau^k \kappa^k) \end{array} \tag{7}$$

which shows that the primal residuals (r_p), the dual residuals (r_d), the gap residual (r_g), and the complementary gap ($x^T s + \tau\kappa$) all are reduced strictly if $\alpha^k > 0$ and at the same rate. This shows that $(x^k, \tau^k, y^k, s^k, \kappa^k)$ generated by the algorithm converges towards an optimal solution to (3) (and the termination criteria in step 4 is ultimately reached). In principle the initial point and the stepsize α^k should be chosen such that

$$\min_j (x_j^k s_j^k, \tau^k \kappa^k) \geq \beta \mu^k, \quad \text{for} \quad k = 0, 1, \ldots$$

is satisfied for some $\beta \in (0,1)$ because this guarantees $(x^k, \tau^k, y^k, s^k, \kappa^k)$ converges towards a strictly complementary solution. Finally, it is possible to prove that the algorithm has the complexity $O(n^{3.5}L)$ given an appropriate choice of the starting point and the algorithmic parameters.

4 Termination

Note (6) and (6) implies that that r_p^k, r_d^k, r_g^k, and $((x^k)^T s^k + \tau^k \kappa^k)$ all converge towards zero at exactly the same rate. This implies that feasibility and optimality is reached at the same time. Therefore, if the algorithm is stopped prematurely then solution will neither be feasible nor optimal. Moreover, relaxing ε_g without relaxing ε_f is not likely to have much effect. This can be seen by making the reasonable assumptions that

$$||(r_p^0; r_d^0; r_g^0)|| \approx \mu^0$$

and

$$\varepsilon_g \approx \varepsilon_f.$$

5 Warmstart

It is well known that the simplex algorithm easily can be warmstarted when a sequence of closely related optimization problems has to be solved. This can in many cases cut the computational time significantly. Although there are no guarantees for that. It is also possible warmstart an interior-point algorithm if an initial solution satisfying the conditions in step 4 and

$$||(r_p^0; r_d^0; r_g^0)|| \text{ and } \mu^0$$

are small. Moreover, the initial solution should satisfy

$$\min_j (x_j^0 s_j^0, \tau^0 \kappa^0) \geq \beta \mu^0$$

for a reasonably large β e.g. $\beta = 0.1$. Such an initial solution virtually never known because usually either the primal or dual solution is vastly infeasible. Therefore, in practice it is hard to warmstart an interior-point algorithm with any efficiency gain.

6 Further reading

Further details about the homogeneous algorithm can be seen in [3, 5]. Issues related to implementing the homogeneous algorithm are discussed in [1, 4].

References

[1] E. D. Andersen and K. D. Andersen. The MOSEK interior point optimizer for linear programming: an implementation of the homogeneous algorithm. In J. B. G. Frenk, C. Roos, T. Terlaky, and S. Zhang, editors, *High Performance Optimization Techniques, Proceedings of the HPOPT-II conference*, 1997. forthcoming.

[2] A. J. Goldman and A. W. Tucker. Theory of linear programming. In H. W. Kuhn and A. W. Tucker, editors, *Linear Inequalities and related Systems*, pages 53–97, Princeton, New Jersey, 1956. Princeton University Press.

[3] C. Roos, T. Terlaky, and J. -Ph. Vial. *Theory and algorithms for linear optimization: an interior point approach.* John Wiley and Sons, New York, 1997.

[4] X. Xu, P. -F. Hung, and Y. Ye. A simplified homogeneous and self-dual linear programming algorithm and its implementation. *Annals of Operations Research*, 62:151–171, 1996.

[5] Y. Ye. *Interior point algorithms: theory and analysis.* John Wiley and Sons, New York, 1997.

[6] Y. Ye, M. J. Todd, and S. Mizuno. An $O(\sqrt{n}L)$ - iteration homogeneous and self-dual linear programming algorithm. *Math. Oper. Res.*, 19:53–67, 1994.