On a property of the Cholesky factorization and its consequences in interior point methods

Csaba Mészáros

WP 98–7

Abstract

The paper concerns the Cholesky factorization of symmetric positive definite matrices resulting from some matrix R by computing RR^T. Our investigation is based on a theorem which shows that "small" diagonal values in the Cholesky factorization indicate rows of R which lie close to the subspace spanned by the preceding rows. A practical, scaling independent technique, based on the above property, is developed for the modified Cholesky factorization. This technique increases the robustness of Cholesky factorizations performed during interior point iterations when the optimization problem is degenerate. Our investigations show also the limitations of interior point methods with the recent implementation technology and floating point arithmetic standard. We present numerical results on the degenerate linear programming problems of NETLIB.

1 Introduction

The impressive progress in the theory and practice of interior point methods (IPM) in the past 15 years raised several questions. The important practical issue, the stability of the computations in IPMs, deserved especially great attention in the literature [10, 11, 12, 13]. For most interior point algorithms the major computational task is to solve systems of linear equations with positive definite matrix, which is done by Cholesky decomposition in practice. One of the most important difficulties for IPMs is the ill–conditioning of these linear systems when the method approaches the optimal solution of the optimization problem. It was observed that in the degenerate case, if the Jacobian of the active constraints of an inequality–constrained optimization problem is not of full row rank at the optimum, the ill–conditioning can brake down the Cholesky factorization. The breakdown of the decomposition usually appears by arising very small (perhaps zero) pivot elements during the factorization process. This problem was recognized by several authors and different solutions were proposed.

One possibility to overcome this situation is performing a modified Cholesky factorization by skipping factors corresponding to small pivots during numerical computations. This can be performed in different ways either by setting the pivot value to a very large number or by setting the elements in the factor to zero [11].

Another possibility is using regularization, i.e., increasing the small pivot values to a "reasonably" large one. In [1] it was described how the effect of the regularization can be taken into account by using quadratic penalty terms. On the other hand, the analysis in [8] shows that small regularization causes small perturbation in the solution of the system which can be decreased by standard iterative refinement techniques.

The common property of these approaches is that they are very sensitive of how the "small" pivot values are identified. Up to now only one approach was discussed in the literature [11], in which the author suggested using a threshold value, relative to the largest diagonal element of the system to be factorized. Unfortunately, this method depends on the scaling of the problem, therefore, it is not reliable in the practice.

In the paper we investigate the Cholesky factorization of a matrix arising from a *source matrix* R by computing RR^T. In section 2 we show that the diagonal values of the Cholesky decomposition of RR^T gives information about the degeneracy of R. This property gives some theoretical support for the modified Cholesky factorization. In section 3 we describe a scaling independent way for identifying small pivot values during the Cholesky decomposition and discuss the consequences of the modified Cholesky factorization in interior point methods. In section 4 we show numerical results on degenerate linear programming problems. Section 5 contains the conclusions of our investigation.

2 On the Cholesky factorization

In this section we concern the Cholesky factorization

$$RR^T = LL^T, \tag{1}$$

where $R \in \Re^{m\times n}$ is of full row rank and L is lower triangular. An important application of this decomposition, such as in the interior point methods, is to compute projections to the linear space of R^T. Usually, the Cholesky factorization is performed by the following standard procedure:

$$\begin{array}{l} \text{For i} = 1, ..., \text{m} \\ \quad L_{ii} = \sqrt{\sum\limits_{k=1}^{n} R_{ik}^2 - \sum\limits_{k=1}^{i-1} L_{ki}^2} \\ \quad \text{For j} = \text{i} + 1, ..., \text{m} \\ \qquad L_{ij} = \dfrac{\sum\limits_{k=1}^{n} R_{ik}R_{jk} - \sum\limits_{k=1}^{i-1} L_{kj}L_{ki}}{L_{ii}}. \end{array} \tag{2}$$

Further on, we will use the following definition [4]:

Definition 1 *Let $r_1, ..., r_k$ be row vectors of the same dimension, and $\mathcal{L}\{r_1, ..., r_{k-1}\}$ the linear subspace spanned by vectors $r_1, ..., r_{k-1}$. The distance of r_k from the linear subspace $\mathcal{L}\{r_1, ..., r_{k-1}\}$ is defined as:*

$$d(r_k, \mathcal{L}\{r_1, ..., r_{k-1}\}) = \min_{p_1,...,p_{k-1}} \left\| r_k - \sum_{i=1}^{k-1} p_i r_i \right\|, \tag{3}$$

where $p_i \in \Re$ and $\|.\|$ denotes the Euclidean norm.

As it is known, the residual vector, $r_k - \sum_{i=1}^{k-1} p_i r_i$, is orthogonal to $\mathcal{L}\{r_1, ..., r_{k-1}\}$ and when the vectors $r_1, ..., r_{k-1}$ are linearly independent, then the optimal $p = (p_1, ..., p_{k-1})^T$ is unique and can be obtained as

$$p = \left(R_{k-1}R_{k-1}^T\right)^{-1} R_{k-1}r_k^T,$$

where $R_{k-1} = \left[r_1^T, ..., r_{k-1}^T\right]^T$ (see [4]).

Further on, let R_i be the matrix consisting of the first i row of R, that is $R_i = \left[r_1^T, ..., r_{i-1}^T\right]^T$ and L_i the Cholesky factors of $R_i R_i^T$ for $i = 1, ..., m$. It is easy to see that

$$L_k = \begin{bmatrix} L_{k-1} & \\ c_k^T & \alpha_k \end{bmatrix},$$

where

$$L_{k-1} c_k = R_{k-1} r_k^T \ , \tag{4}$$

and

$$\alpha_k^2 = r_k^T r_k - c_k^T c_k. \tag{5}$$

Theorem 1 *The k-th diagonal value of the Cholesky factorization of RR^T is equal to the distance of vector r_k from the subspace spanned by vectors $r_1, \ldots r_{k-1}$, that is*

$$\alpha_k = d\left(r_k, \mathcal{L}\left\{r_1, \ldots r_{k-1}\right\}\right).$$

Proof. et us take a closer look on the norm of the residual:

$$\begin{aligned} \left\| r_k - \sum_{i=1}^{k-1} p_i r_i \right\|^2 &= \left\| r_k^T - R_{k-1}^T \left(R_{k-1} R_{k-1}^T\right)^{-1} R_{k-1} r_k^T \right\|^2 \\ &= \left\| r_k^T - R_{k-1}^T L_{k-1}^{-T} L_{k-1}^{-1} R_{k-1} r_k^T \right\|^2 \\ &= \left\| r_k^T - R_{k-1}^T L_{k-1}^{-T} c_k \right\|^2 \\ &= \left\| r_k \right\|^2 + \left\langle R_{k-1}^T L_{k-1}^{-T} c_k, R_{k-1}^T L_{k-1}^{-T} c_k \right\rangle + \left\langle R_{k-1}^T L_{k-1}^{-T} c_k, -2 r_k^T \right\rangle \\ &= \left\| r_k \right\|^2 + \left\langle R_{k-1}^T L_{k-1}^{-T} c_k, R_{k-1}^T L_{k-1}^{-T} c_k - 2 r_k^T \right\rangle. \end{aligned}$$

Since $R_{k-1}^T L_{k-1}^{-T} c_k - r_k^T$ is orthogonal to $\mathcal{L}\left\{r_1, \ldots r_{k-1}\right\}$, thus

$$\left\langle R_{k-1}^T L_{k-1}^{-T} c_k, R_{k-1}^T L_{k-1}^{-T} c_k - r_k^T \right\rangle = 0,$$

and

$$\begin{aligned} \left\| r_k - \sum_{i=1}^{k-1} p_i r_i \right\|^2 &= \left\| r_k \right\|^2 - \left\langle R_{k-1}^T L_{k-1}^{-T} c_k, r_k^T \right\rangle \\ &= \left\| r_k \right\|^2 - \left\langle c_k, L_{k-1}^{-1} R_{k-1} r_k^T \right\rangle \\ &= \left\| r_k \right\|^2 - \left\| c_k \right\|^2, \end{aligned}$$

which together with (5) proves the theorem. ■

Since the value of $d\left(r_k, \mathcal{L}\left\{r_1, ..., r_{k-1}\right\}\right)$ depends on the scaling of r_k, later on we will use the relative distance of r_k and $\mathcal{L}\left\{r_1, ..., r_{k-1}\right\}$ which can be defined for $r_k \neq 0$ vector as

$$\hat{d}\left(r_k, \mathcal{L}\left\{r_1, ..., r_{k-1}\right\}\right) = \frac{d\left(r_k, \mathcal{L}\left\{r_1, ..., r_{k-1}\right\}\right)}{\|r_k\|}.$$

Note that $\hat{d}\left(r_k, \mathcal{L}\left\{r_1, ..., r_{k-1}\right\}\right)^2$ is the Orchard–Hayes' relative acceptability measure [9] for α_k^2 in (5). Let us observe the following consequence of theorem 1 in the practical factorization where the computations are performed with the finite ϵ_M relative precision:

Corollary 1 *If $\hat{d}\left(r_k, \mathcal{L}\left\{r_1, ..., r_{k-1}\right\}\right) < \sqrt{\epsilon_M}$, then the computed α_k has no significant digits.*

Proof. ince

$$\hat{d}\left(r_k, \mathcal{L}\left\{r_1, ..., r_{k-1}\right\}\right)^2 = \frac{\|r_k\|^2 - \|c_k\|^2}{\|r_k\|^2} < \epsilon_M,$$

the error when computing $\|r_k\|^2 - \|c_k\|^2$ is of the same (or even larger) order of magnitude than the result, therefore, in the computed value of $\sum_{i=1}^{n} R_{ki}^2 - \sum_{i=1}^{k-1} L_{ik}^2$ the round–off error is the dominating factor. ■

The relative precision of today's commonly used double precision arithmetic is $\epsilon_M \approx 10^{-16}$, thus for a matrix R in which $\hat{d}\left(r_k, \mathcal{L}\left\{r_1, ..., r_{k-1}\right\}\right) < 10^{-8}$ the kth and forthcoming computed Cholesky factors of RR^T mostly depend on the previous round–off errors and on the sequence of floating point operations performed in (2), therefore, its use provides unreliable results. In the next section we will investigate this situation in interior point methods.

3 The modified Cholesky factorization and the interior point methods

In this section we introduce a modified version of (2) which handles ill–conditioned cases by skipping the unreliable steps of the factorization. The similar approach was introduced in [11], which differs from ours in the determination of the threshold value of the pivot elements. Our version uses an $\epsilon > 0$ tolerance and an adaptive rule which makes the technique independent of scaling of R in contrast to the method described in [11], where the threshold value is fixed at the beginning of the factorization. Our modified Cholesky decomposition performs the following steps:

$$
\begin{aligned}
&\text{For i} = 1, ..., \text{m} \\
&\quad c_i = \sum_{k=1}^{i-1} L_{ki}^2 \\
&\quad f_i = \sum_{k=1}^{n} R_{ik}^2 \\
&\quad L_{ii} = \begin{cases} \infty & \text{if } (1-\epsilon)\text{f}_\text{i} \leq \text{c}_\text{i}, \\ \sqrt{f_i - c_i} & \text{otherwise} \end{cases} \\
&\quad \text{For j} = \text{i} + 1, ..., \text{m} \\
&\qquad L_{ij} = \frac{\sum_{k=1}^{n} R_{ik}R_{jk} - \sum_{k=1}^{i-1} L_{kj}L_{ki}}{L_{ii}} \quad .
\end{aligned}
\tag{6}
$$

Let us note that the pivot element is computed in the above algorithm in such a way that the negative and positive terms are accumulated separately and added together at the end with the Orchard–Hayes relative tolerance [9]. This idea proved to be numerically quite efficient in other contexts [5].

We selected $\epsilon = 10^{-15}$ which provides at least one significant digit for all the accepted results of $f_i - c_i$ on standard floating point arithmetic. Our procedure skips all factors, for which the computed diagonal (pivot) value is definitely unreliable. Let us note that the original problem in most of the applications is to compute projections onto the space of R^T and the theorem in section 2 shows that skipping the rows as described does not hurt too much the result because rows of R the distance of which to the subspace spanned by the remaining rows is less than $\sqrt{\epsilon} \approx 3 \cdot 10^{-8}$ are omitted.

In the following part of this section we will consider the linear programming problem and the primal–dual log barrier interior point method for further investigation. It is to be noted, however, that this choice has notational consequences only and one can draw conclusions similar to ours when considering the general nonlinear programming problem or another interior point approach.

Let us consider the linear programming problem as

$$\begin{aligned} &\min\ c^T x \\ Ax &= b, \\ x &\geq 0, \end{aligned} \tag{7}$$

where $x, c \in \Re^n$, $A \in \Re^{m \times n}$ is of full row rank, furthermore, $b \in \Re^m$. The logarithmic barrier problem corresponding to (7) is

$$\begin{aligned} &\min\ c^T x - \mu \sum_{i=1}^{n} \ln\, x_i, \\ &Ax = b, \end{aligned} \tag{8}$$

where μ is a positive scalar barrier parameter. A log barrier interior point method approach the optimal solution of (7) by a sequence of barrier problems (8), while the barrier parameter is decreased to zero. Following the classical introduction of the primal–dual log barrier method, the algorithm can be derived by applying Newton's method to solve the Karush-Kuhn-Tucker system of (8). It is easy to see that the computational task of the resulting method in the kth iteration can be reduced to a solution of a system of linear equations whose matrix is $AD^k A^T$ where D^k is a positive definite diagonal *scaling* matrix and its diagonal values are

$$D^k_{ii} = \frac{\left(x^k_i\right)^2}{\mu}$$

if the kth iterate is perfectly centered. See [1] for the details. Thus, in our case at the kth iteration the source matrix is defined as $R = A\hat{D}^k$, where $\hat{D}^k(\hat{D}^k)^T = D^k$.

In case the set of columns corresponding to variables having positive value at the optimum of (7) is not of full row rank, i.e., if the optimization problem is degenerate, $A\hat{D}^k$ will converge to a singular matrix. In this situation the standard Cholesky factorization will break down because during iterations some row vectors of $A\hat{D}^k$ converge into the subspace spanned by the other row vectors. Our proposed modification of the Cholesky factorization will handle this situation well by skipping the rows in $A\hat{D}^k$ which appear to be linearly dependent form the set of the remaining rows. The analysis in

[11] shows that the obtained search directions are sufficiently accurate in this case.

It is to be noted, however, that the degeneracy is not the only phenomenon which can break down the Cholesky factorization in interior point methods. Let us consider the optimal complementarity partition of problem (7) as (B, N) where B is the set of indices of variables having nonzero value at the optimum and N is its complementer. Furthermore, let A_B denote the matrix consisting of the columns of A whose index belongs to B. It is easy to see that in case A_B is ill–conditioned, $A\hat{D}^k$ will behave in a manner similar to the degenerate case. Let us demonstrate this on the following example:

$$\begin{array}{lccccccccc} \min & & & & & & & x_4 & & \\ & x_1 & & & & & & & = & 1 \\ & & & x_2 & + & x_3 & & & = & 2 \\ & x_1 & + & x_2 & + & (1+10^{-8})x_3 & + & x_4 & = & 3+10^{-8} \\ & & & & & x_1, x_2, x_3, x_4 \geq 0 & & & & \end{array} .$$

It is easy to see that the unique optimal solution of the above problem is $x = (1, 1, 1, 0)^T$. Let us observe that with perfectly centered iterates $A\hat{D}^k$ converges to

$$\frac{1}{\sqrt{\mu}} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 1 & 1+10^{-8} & 0 \end{bmatrix},$$

thus, at a certain iteration $\hat{d}(r_3, \mathcal{L}\{r_1, r_2\})$ decreases under 10^{-8} (but r_3 do not converge to $\mathcal{L}\{r_1, r_2\}$), which results in breaking down the Cholesky factorization, whereas, the optimization problem is not degenerate. As it was expected, our state of the art interior point optimizer, called BPMPD [6] (with disabled presolve feature) was not able to correctly solve the above problem and stopped at the primal iterate $(1, 10^{-8}, 2-10^{-8}, 10^{-12})$ with primal infeasibility of order 10^{-8}.

We also applied another efficient and freely available interior point implementation, PCx version 1.1 [2], on the above problem. The result provided by PCx (with disabled presolve) was $(1, 10^{-3}, 2-10^{-3}, 10^{-11})$ which is also far from the true optimum of the problem. Let us note that due to the ill–conditioning of A_B, a qualitative better solution cannot be expected not even from simplex-based approaches.

Another situation is if the positive components of the optimal solution are of different orders of magnitude that is if the problem is poorly scaled. This can also break down the Cholesky factorizations in IPMs even if the optimization problem is well–conditioned and non–degenerate. To demonstrate this case let us consider the following example:

$$\begin{array}{lccccccc}
\min & & & & & x_3 & & \\
& x_1 & & & & & = & 10^5 \\
& x_1 & & -x_2 & + & x_3 & = & 10^{-5} \\
& x_1 & + & -x_2 & + & 2x_3 & = & 2\cdot 10^{-5} \\
& & & \multicolumn{3}{c}{x_1, x_2, x_3 \geq 0} & &
\end{array}$$

Note that the matrix of the above problem is well–conditioned and the unique optimum of the problem is $(10^5, 10^5, 10^{-5})$. It is to be observed that with perfectly centered iterates $A\hat{D}^k$ converges to the matrix

$$\frac{1}{\sqrt{\mu}}\begin{bmatrix} 10^5 & 0 & 0 \\ 10^5 & 10^5 & 10^{-5} \\ 10^5 & 10^5 & 2\cdot 10^{-5} \end{bmatrix},$$

thus, $\hat{d}\left(r_3, \mathcal{L}\left\{r_1, r_2\right\}\right)$ can decrease under 10^{-8} during iterations, therefore, the Cholesky factorization may break down. On this problem our solver (without presolve) was not able to compute sufficiently accurate Cholesky decompositions after the 3rd iteration when the value of x_3 decreased under $5\cdot 10^{-5}$. Similarly, the solver PCx reported inaccurate factorizations during the last iterations.

It is to be noted that this latter situation does not result in numerical problems in the factorization process of the algorithms based on the simplex method, because A_B is well-conditioned. The poor scaling of the problem appears in the sensitivity of the optimality tolerances by these approaches.

4 Numerical results

In this section we show our numerical results on degenerate real–life linear programs. We implemented the supernodal loop unrolling variant of the left looking Cholesky factorization [7] with the modification described in the previous section in our primal–dual log barrier solver. In our implementation we store the inverse of the Cholesky factors, thus, in factors skipped we set the corresponding diagonal element to zero.

Let us note that the special case, if the constraint matrix of the LP problem is not of full row rank, the problem is obviously degenerate, but this degeneracy can be detected and removed at the beginning of the algorithm. In our implementation this is done during the computation of the starting point: we compute one factorization with $D^0 = I$ with drop tolerance $\epsilon = 10^{-12}$. All rows, corresponding to dropped factors during the Cholesky decomposition are removed from the optimization problem after checking their feasibility. Let us note that as a result of this procedure, the rows of which distance to the linear space of the remaining rows is less than 10^{-6} are dropped from the problem. During iterations the drop tolerance used is $\epsilon = 10^{-15}$ and rows of the original problem corresponding to skipped factors in the Cholesky decomposition are not removed.

Tables 1 and 2 present the problems of the Netlib and Kennington problem sets [3] in which our modified Cholesky factorization skipped any rows in our interior point implementation. Figures given include the number of rows after presolve, number of dependent rows, number of skipped factors in the last decomposition, final absolute and relative primal infeasibilities and the relative duality gap. The relative primal infeasibility and relative duality gap are defined as

$$\frac{\|Ax - b\|}{\|b\| + 1} \quad \text{and} \quad \frac{\mathrm{b^T y - c^T x}}{|\mathrm{c^T x}| + 1}$$

for an (x, y) where x is an approximate solution of (7) and y is an approximate solution for the dual problem $\left\{ \max\limits_{y \in \Re^m} b^T y \;\middle|\; y^T A \leq c \right\}$. Problems `quap-15` and `ken-18` were not solved, due to the 64 Mbyte memory limitation of our hardware.

The results presented in tables 1 and 2 indicate that our interior point implementation, with the modified Cholesky decomposition which uses our adaptive threshold tolerance, can reliably handle degenerate problems. In all cases our solver was able to achieve the desired 10^{-8} relative duality gap and relative infeasibility level.

Table 1: Results on degenerate Netlib problems

Problem name	No. of rows	Dep. rows	Skip. rows	Abs primal infeas.	Rel. primal infeas.	Rel. duality gap
80bau3b	1965	0	3	4.0E–09	1.0E–14	2.1E–09
bnl2	1059	0	9	1.6E–09	2.5E–12	2.2E–10
boeing1	297	0	7	1.5E–08	4.5E–15	3.0E–11
boeing2	124	0	13	3.2E–08	1.5E–13	3.1E–11
bore3d	30	2	0	1.9E–08	2.9E–13	4.6E–11
capri	137	0	1	5.5E–10	1.3E–13	1.0E–09
cycle	1026	0	159	1.0E–09	1.7E–13	1.7E–11
d6cube	402	0	1	1.8E–10	4.5E–14	7.8E–10
degen2	383	2	52	9.0E–11	2.8E–12	7.2E–11
degen3	1406	2	203	5.6E–08	1.3E–12	6.6E–11
dfl001	3791	13	68	7.7E–09	1.9E–11	7.5E–09
forplan	80	0	12	2.3E–07	2.7E–14	5.8E–10
greenbea	1168	0	30	2.9E–06	9.1E–11	1.4E–09
greenbeb	1171	0	6	4.3E–08	5.6E–11	4.1E–09
lotfi	106	0	1	1.8E–08	9.4E–15	1.7E–11
maros	494	0	13	1.4E–08	7.0E–13	1.0E–10
modszk1	447	0	19	1.5E–11	2.2E–16	1.0E–10
pilot–ja	651	0	1	8.4E–11	1.2E–13	3.7E–09
qap08	894	170	41	1.0E–09	1.8E–10	1.1E–10
qap12	3192	398	141	2.7E–08	9.3E–09	6.0E–10
recipe	62	0	1	1.0E–10	1.4E–12	5.4E–11
scorpion	187	29	0	9.7E–14	6.7E–14	1.4E–10
sctap3	1346	0	1	1.7E–13	1.4E–15	6.6E–11
shell	281	1	0	9.3E–10	9.5E–16	2.2E–09
ship04l	288	0	2	2.6E–10	8.3E–14	5.2E–11
ship12l	610	0	1	6.7E–11	2.5E–13	3.0E–10
sierra	1083	10	5	5.4E–07	7.5E–15	1.2E–10
standata	232	0	32	1.5E–12	3.1E–15	1.3E–10
standgub	232	0	32	1.5E–12	3.1E–15	1.3E–10
standmps	336	0	40	5.8E–11	2.0E–14	1.1E–11
wood1p	170	1	0	4.2E–12	2.7E–12	2.9E–09

Table 2: Results on degenerate Kennington problems

Problem name	No. of rows	Dep. rows	Skip. rows	Abs primal infeas.	Rel. primal infeas.	Rel. duality gap
cre–a	2760	5	60	8.1E–10	1.8E–13	1.4E–10
cre–b	4995	8	37	7.4E–08	6.4E–12	3.3E–09
cre–c	2168	5	17	7.1E–10	2.2E–13	1.1E–10
cre–d	3791	8	19	7.9E–09	1.4E–12	1.1E–09
ken–07	896	49	0	9.5E–12	3.3E–16	3.5E–10
ken–11	6314	121	0	1.5E–11	7.8E–17	4.3E–10
ken–13	13141	169	2	1.6E–11	5.9E–16	1.1E–10
pds–02	1615	11	0	1.1E–07	2.5E–12	8.9E–11
pds–06	5056	11	0	2.3E–10	1.4E–15	2.1E–09
pds–10	8410	11	31	2.4E–08	1.0E–13	2.6E–11

5 Conclusions

In the paper we investigated the Cholesky factorization of symmetric positive definite systems which arise from a matrix R by computing RR^T. A theorem, which shows that the kth diagonal element of the Cholesky decomposition is equal to the distance of the kth row of R and the linear space spanned by the preceding rows, gives a relationship between the Orchard-Hayes relative tolerance when computing the pivot elements of the factorization and the relative distance of rows of R and the subspace of the preceding rows. The theorem shows that the numerical breakdown of the standard Cholesky factorization can occur if some rows of R are closely linearly dependent from the remaining set of rows. A modification of the standard factorization was proposed, which can handle rank deficient and ill–conditioned cases.

Since the Cholesky decomposition is perhaps the most important operation in recent interior point implementations, the practice of IPMs is a natural place for the application of our results. We showed that the modified Cholesky factorization can reliably handle matrices arising from the application of interior point methods on degenerate optimization problems. Unfortunately, the degeneracy is not the only source of the breakdown of the Cholesky decomposition in IPMs. We pointed out that non–degenerate problems with well-conditioned constraint set can result in numerical troubles in the Cholesky factorization of interior point methods. We showed situations which are impossible to handle with the modifications of the standard decomposition scheme. We demonstrated that in these cases recent interior point implementations are not able to solve very small optimization problems either. Techniques capable of successfully handling such situations require further research.

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] J. Czyzyk, S. Mehrotra, and S. J. Wright. PCx user guide. Technical Report OTC 96/01, Optimization Technology Center, Argonne National Laboratory and Northwestern University, 1996.

[3] D. M. Gay. Electronic mail distribution of linear programming test problems. *COAL Newsletter*, 13:10–12, 1985.

[4] G.H. Golub and C.F. Van Loan. *Matrix Computations*. The John Hopkins University Press, Baltimore and London, 2 edition, 1989.

[5] I. Maros and Cs. Mészáros. A numerically exact implementation of the simplex method. *Annals of Operations Research*, 58:3–17, 1995.

[6] Cs. Mészáros. *The Efficient Implementation of Interior Point Methods for Linear Programming and their Applications*. PhD thesis, Eötvös Loránd University of Sciences, 1996.

[7] Cs. Mészáros. Fast Cholesky factorization for interior point methods of linear programming. *Computers & Mathematics with Applications*, 31(4/5):49–54, 1996.

[8] Cs. Mészáros. On free variables in interior point methods. *Optimization Methods and Software*, 9:121–139, 1997.

[9] W. Orchard-Hays. *Advanced Linear Programming Computing Techniques*. McGraw Hill, 1968.

[10] M. H. Wright. Some properties of the Hessian of the logarithmic barrier function. *Math. Programming*, 67:265–295, 1994.

[11] S. J. Wright. Modified Cholesky factorization in interior–point algorithms for linear programming. Technical Report P600-0596, Mathematics and Computer Science Division, Argonne National Laboratory, 1996.

[12] S. J. Wright. Stability of linear algebra computations in interior-point methods for linear programming. *SIAM J. Matrix Analysis and Applications*, 18:191–222, 1997.

[13] S. J. Wright. Effects of finite–precision arithmetic on interior–point methods for nonlinear programming. Technical Report P705-0198, Mathematics and Computer Science Division, Argonne National Laboratory, 1998.