Regularization and preconditioning of KKT systems arising in nonnegative least-squares problems

Citation for published version:
Bellavia, S, Gondzio, J & Morini, B 2009, 'Regularization and preconditioning of KKT systems arising in nonnegative least-squares problems' Numerical Linear Algebra with Applications, vol. 16, no. 1, pp. 39-61. DOI: 10.1002/nla.610

Digital Object Identifier (DOI):
10.1002/nla.610

Published In:
Numerical Linear Algebra with Applications

Bellavia, S, Gondzio, J & Morini, B 2011, 'Regularization and preconditioning of KKT systems arising in nonnegative least squares problems' Numerical and Linear Algebra with Applications, vol 18, no. 3, pp. 363-385

REGULARIZATION AND PRECONDITIONING OF KKT SYSTEMS ARISING IN NONNEGATIVE LEAST-SQUARES PROBLEMS *

STEFANIA BELLAVIA †, JACEK GONDZIO ‡, AND BENEDETTA MORINI §

Technical Report MS-07-004, August 30th, 2007

Abstract. A regularized Newton-like method for solving nonnegative least-squares problems is proposed and analysed in this paper. A preconditioner for KKT systems arising in the method is introduced and spectral properties of the preconditioned matrix are analysed. A bound on the condition number of the preconditioned matrix is provided. The bound does not depend on the interior-point scaling matrix. Preliminary computational results confirm the effectiveness of the preconditioner and fast convergence of the iterative method established by the analysis performed in this paper.

1. Introduction.

Optimization problems having a least-squares objective function along with simple constraints arise naturally in image processing, data fitting, control problems and intensity modulated radiotherapy problems. In the case where only one-sided bounds apply, there is no lack of generality to consider the Nonnegative Least-Squares (NNLS) problems

$$\min_{x\geq 0} q(x) = \frac{1}{2}\|Ax-b\|_2^2, \tag{1.1}$$

where $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$ are given and $m \geq n$, [5].

We assume that A has full column rank so that the NNLS problem (1.1) is a strictly convex optimization problem and there exists a unique solution x^* for any vector b, [18]. We allow the solution x^* to be degenerate, that is strict complementarity may not hold at x^*. We are concerned with the situation when m and n are large and we expect matrix A to be sparse. We address the issues of the numerical solution of (1.1) and apply the Interior-Point Newton-like method given in [3] to it. ¿From numerical linear algebra perspective this method can be reduced to solving a sequence of (unconstrained) weighted least-squares subproblems. Following [3] the linear systems involved have the following form:

$$\begin{pmatrix} I & AS \\ SA^T & -WE \end{pmatrix} \begin{pmatrix} \tilde{q} \\ \tilde{p} \end{pmatrix} = \begin{pmatrix} d \\ 0 \end{pmatrix}, \tag{1.2}$$

where S, W, E are diagonal matrices satisfying $S^2 + WE = I$. Bellavia et al. [3] solve this system with an iterative method which corresponds to the use of inexact Newton-like method to solve NNLS. The reader interested in the convergence analysis of this method should consult [3, 16] and the references therein.

In this paper we go a step further. The system (1.2) arising in interior point method [3] is potentially very ill-conditioned and therefore difficult for iterative methods. Although preconditioning can help to accelerate the convergence of the iterative method, in some situations it is insufficient to ensure fast convergence. We remedy it in

*The visit of the second author to the University of Florence was partially supported by GNCS/INDAM, Italy.

†Dipartimento di Energetica "S. Stecco", Università di Firenze, via C. Lombroso 6/17, 50134 Firenze, Italia, stefania.bellavia@unifi.it

‡School of Mathematics, The University of Edinburgh, Mayfield Road, Edinburgh EH9 3JZ, Scotland, UK, J.Gondzio@ed.ac.uk

§Dipartimento di Energetica "S. Stecco", Università di Firenze, via C. Lombroso 6/17, 50134 Firenze, Italia, benedetta.morini@unifi.it

this paper. We regularize (1.2) to guarantee that the condition number of the matrix involved is decreased. Next, we design a preconditioner for such regularized system and analyse the spectral properties of the preconditioned matrix.

The diagonal matrix WE in (1.2) displays an undesirable property: certain elements of it may be null and others may be close to one. Inspired by the work of Saunders [22] and encouraging numerical experience reported by Altman and Gondzio [2] we regularize the $(2, 2)$ block in (1.2) and replace element WE with $WE + S^2\Delta$, where Δ is a diagonal matrix. We provide a detailed spectral analysis of the regularized matrix and prove that the condition number of the regularized system is significantly smaller than that of the original system. Following [2] we use a *dynamic* regularization, namely, we do not alter those elements of WE which are bounded away from zero; we change only those which are too close to zero.

Having improved the conditioning of (1.2) we make it easier for iterative methods. In fact, (1.2) is a saddle point problem with a simple structure, [4]. We design an indefinite preconditioner which exploits the partitioning of indices into "large" and "small" elements in the diagonal of WE. The particular form of the regularization yields a relevant advantage: if the partitioning in WE does not change from an iteration to another, then the factorization of the preconditioner is available for the new iteration at no additional cost. Moreover, as the algorithm approaches the optimal solution the partitioning settles down following the splitting of indices into those corresponding to active and inactive constraints. Then, eventually the factorization of the preconditioner does not require computational effort.

We provide the spectral analysis of the preconditioned system and show that the condition number of its matrix is bounded by a number which does not depend on interior-point scaling. We are not aware of any comparable result in the literature. Moreover, the proposed preconditioner allows us to use the short recurrence PPCG method [11] to solve the preconditioned linear system. Thus, we performed also the spectral analysis of the reduced preconditioned normal equation whose eigenvalues determine the convergence of the PPCG method. Our preliminary computational results confirm all theoretical findings. Indeed, we have observed that the augmented system can be solved very efficiently by PPCG method. For problems of medium and large scale (reaching a couple of hundred thousand constraints) iterative methods converge in a low number of iterations.

Regularization does add a perturbation to the Newton system. We prove that it does not worsen the convergence properties of the Newton-like method of Bellavia et al. [3], namely, the inexact regularized Newton-like method still enjoys q-quadratic convergence, even in presence of degenerate solutions.

We remark that our method can be used also to solve regularized least-squares problems:

$$\min_{x\geq 0} q(x) = \frac{1}{2}\|Ax - b\|_2^2 + \mu\|x\|^2,$$

where μ is a strictly positive scalar. In this case, the arising augmented systems are regularized "naturally" and we do not need to introduce the regularization of the (2,2) block. Moreover, A may also be rank deficient. Finally, the method can clearly handle lower and upper bounds on the variables too. We decided to limit our discussion to NNLS problems for sake of simplicity.

The paper is organised as follows. In Section 2, we remind key features of the Newton-like method studied in [3] and justify the need of introducing regularization.

In Section 3, we introduce the regularization and show that it does not worsen local convergence properties of the Newton-like method. In Section 4, we analyse numerical properties of the regularized augmented system, introduce the preconditioner and perform spectral analysis of the preconditioned matrix. We also analyse the reduced normal equations formulation of the problem and for completeness perform spectral analysis of the preconditioned normal equations. In Section 5, we discuss preliminary computational results.

1.1. Notations. We use the subscript k as index for any sequence and for any function f we denote $f(x_k)$ by f_k. The symbol x_i or $(x)_i$ denotes the i-th component of a vector x. The 2-norm is indicated by $\|\cdot\|$.

2. The Newton-like method. In this section we briefly review the Newton-like method proposed in [3]. Then, we study the properties of the linear system arising at each iteration.

The Newton-like method in [3] is applied to the system of nonlinear equations

$$D(x)g(x) = 0, \tag{2.1}$$

where $g(x) = \nabla q(x) = A^T(Ax - b)$, and $D(x) = diag(d_1(x), \ldots, d_n(x))$, $x \geq 0$, has entries of the form

$$d_i(x) = \begin{cases} x_i & \text{if} \quad g_i(x) \geq 0, \\ 1 & \text{otherwise} \end{cases} . \tag{2.2}$$

This system states the Karush-Kuhn-Tucker conditions for problem (1.1), [7].

At kth iteration of the Newton-like method, given $x_k > 0$, one has to solve the following linear system:

$$W_k D_k M_k p = -W_k D_k g_k, \tag{2.3}$$

where, for $x > 0$, the matrices $M(x)$ and $W(x)$ are given by

$$M(x) = A^T A + D(x)^{-1} E(x), \tag{2.4}$$

$$E(x) = diag(e_1(x), \ldots, e_n(x)), \tag{2.5}$$

with

$$e_i(x) = \begin{cases} g_i(x) & \text{if} \quad 0 \leq g_i(x) < x_i^2 \text{ or } g_i(x)^2 > x_i \\ 0 & \text{otherwise} \end{cases} , \tag{2.6}$$

and

$$W(x) = diag(w_1(x), \ldots, w_n(x)), \qquad w_i(x) = \frac{1}{d_i(x) + e_i(x)} . \tag{2.7}$$

We refer to [3] for the motivation to consider such a Newton equation. Here we just mention that this choice of E and W allows to develop fast convergent methods without assuming strict complementarity at x^*.

Clearly, for $x > 0$, the matrices $D(x)$ and $W(x)$ are invertible and positive definite, while the matrix $E(x)$ is semidefinite positive. Further, the matrix $(W(x)D(x)M(x))^{-1}$ exists and is uniformly bounded for all strictly positive x, see [16, Lemma 2].

For sake of generality, in the sequel we consider the formulation of the method in the context of inexact Newton methods [9]. Thus, the Newton equation takes the form

$$W_k D_k M_k p_k = -W_k D_k g_k + r_k, \tag{2.8}$$

and the residual vector r_k is required to satisfy

$$\|r_k\| \le \eta_k \|W_k D_k g_k\|, \qquad \eta_k \in [0,1). \tag{2.9}$$

The linear system (2.8) can be advantageously formulated as a linear system with symmetric positive definite matrix. To this end, for any $x > 0$, we let

$$S(x) = W(x)^{\frac{1}{2}}\, D(x)^{\frac{1}{2}}, \tag{2.10}$$

$$Z(x) = S(x)M(x)S(x) = S(x)^T A^T A S(x) + W(x)E(x), \tag{2.11}$$

and reformulate (2.8) as the equivalent system

$$Z_k \tilde{p}_k = -S_k g_k + \tilde{r}_k, \tag{2.12}$$

with $\tilde{r}_k = S_k^{-1} r_k$ and $\tilde{p}_k = S_k^{-1} p_k$. This system has nice features. Since the matrix $S(x)$ is invertible for any $x > 0$, then the matrix $Z(x)$ is symmetric positive definite for $x > 0$. Moreover it is remarkable that for $x_k > 0$ the matrix Z_k has uniformly bounded inverse and its conditioning is not worse than that of $W_k D_k M_k$, [3, Lemma 2.1]. Moreover, note that from the definition of S and W it follows

$$S_k^2 + W_k E_k = I. \tag{2.13}$$

The residual control associated to (2.12) can be performed imposing

$$\|\tilde{r}_k\| \le \eta_k \|W_k D_k g_k\|, \qquad \eta_k \in [0,1). \tag{2.14}$$

Since $\|S_k\| \le 1$, this control implies that $r_k = S_k \tilde{r}_k$ satisfies (2.9).

After computing $\tilde{p}_k$ satisfying (2.12) and (2.14), the iterate x_{k+1} can be formed. Specifically, positive iterates are required so that matrix S_k is invertible. Then, $p_k = S_k \tilde{p}_k$ is set and the following vector $\hat{p}_k$ is formed:

$$\hat{p}_k = \max\{\sigma,\, 1 - \|P(x_k + p_k) - x_k\|\,\}\,(P(x_k + p_k) - x_k), \tag{2.15}$$

where $\sigma \in (0,1)$ is close to one, and $P(x) = \max\{0, x\}$ is the projection of x onto the positive orthant. Finally, the new iterate takes the form

$$x_{k+1} = x_k + \hat{p}_k. \tag{2.16}$$

The purpose of this section is to investigate the solution of the Newton equation further. Clearly, the system

$$Z_k \tilde{p}_k = -S_k g_k, \tag{2.17}$$

represents the normal equations for the least-squares problem

$$\min_{\tilde{p} \in \mathbb{R}^n} \|B_k \tilde{p} + h_k\|, \tag{2.18}$$

with

$$B_k = \begin{pmatrix} AS_k \\ W_k^{\frac{1}{2}} E_k^{\frac{1}{2}} \end{pmatrix}, \qquad h_k = \begin{pmatrix} Ax_k - b \\ 0 \end{pmatrix}.$$

The conditioning $\kappa_2(Z_k)$ of the matrix Z_k in the 2-norm, is the square of $\kappa_2(B_k)$. Clearly, if $W_k E_k = 0$ then $S_k = I$ and $\kappa_2(B_k) = \kappa_2(AS_k) = \kappa_2(A)$. Otherwise, letting

$$0 < \sigma_1 \leq \sigma_2 \ldots \leq \sigma_n,$$

be the singular values of AS_k we note that the minimum and maximum eigenvalues λ_{min} and λ_{max} of Z_k satisfy

$$\lambda_{min}(Z_k) \geq \sigma_1^2 + \min_i (w_k e_k)_i \geq \sigma_1^2, \tag{2.19}$$

$$\lambda_{max}(Z_k) \leq \sigma_n^2 + \max_i (w_k e_k)_i \leq \sigma_n^2 + 1. \tag{2.20}$$

Thus, an upper bound on the conditioning of B_k is given by

$$\kappa_2(B_k) \leq \frac{\sqrt{1+\sigma_n^2}}{\sigma_1} \leq \frac{1+\sigma_n}{\sigma_1}. \tag{2.21}$$

To avoid solving the system (2.17), we consider the augmented system approach for the solution of the least-squares problem (2.18). It consists in solving the linear system

$$\begin{pmatrix} I & AS_k \\ S_k A^T & -W_k E_k \end{pmatrix} \begin{pmatrix} \tilde{q}_k \\ \tilde{p}_k \end{pmatrix} = \begin{pmatrix} -(Ax_k - b) \\ 0 \end{pmatrix}. \tag{2.22}$$

Next we investigate if this reformulation has better numerical properties than the normal equations (2.17). To study the spectral properties of the augmented matrix, note that $W_k E_k$ is positive semidefinite and

$$v^T W_k E_k v \geq \delta v^T v, \quad \forall v \in \mathbb{R}^n, \tag{2.23}$$

where $1 > \delta = \min_i (e_k)_i / ((d_k)_i + (e_k)_i)$. Clearly, the scalar δ is null whenever at least one diagonal element $(e_k)_i$ is null. The next lemma provides a bound on the conditioning of the following augmented matrix

$$\mathcal{H}_\delta = \begin{pmatrix} I & AS_k \\ S_k A^T & -W_k E_k \end{pmatrix}. \tag{2.24}$$

LEMMA 2.1. *Let $0 < \sigma_1 \leq \sigma_2 \ldots \leq \sigma_n$ be the singular values of AS_k, $\delta \in [0,1)$ be the scalar given in (2.23), $\mathcal{H}_\delta$ be the augmented matrix in (2.24). Then*

$$\kappa_2(\mathcal{H}_\delta) \leq \frac{\frac{1}{2}(1+\sqrt{1+4\sigma_n^2})}{\min\left\{1, \frac{1}{2}\left(\delta - 1 + \sqrt{(1+\delta)^2 + 4\sigma_1^2}\right)\right\}}. \tag{2.25}$$

Proof. We order the eigenvalues of $\mathcal{H}_\delta$ as

$$\mu_{-n} \leq \mu_{-n+1} \leq \ldots \leq \mu_{-1} \leq 0 \leq \mu_1 \leq \mu_2 \leq \ldots \leq \mu_m.$$

Proceeding as in [23, Lemma 2.1, 2.2] we obtain

$$\mu_{-n} \geq -\sqrt{1+\sigma_n^2}, \tag{2.26}$$

$$\mu_1 \geq 1, \tag{2.27}$$

$$\mu_m \leq \frac{1}{2}\left(1+\sqrt{1+4\sigma_n^2}\right). \tag{2.28}$$

To derive an estimate for μ_{-1}, let $(v_1^T, v_2^T)^T \in \mathbb{R}^{m+n}$ be an eigenvector corresponding to μ_{-1}. Hence, from the definition of $\mathcal{H}_\delta$ we have $(1-\mu_{-1})v_1 = -AS_kv_2$ and $v_2^TS_kA^Tv_1 - v_2^TW_kE_kv_2 = \mu_{-1}v_2^Tv_2$. Consequently

$$(1-\mu_{-1})^{-1}v_2^TS_kA^TAS_kv_2 + v_2^TW_kE_kv_2 = -\mu_{-1}v_2^Tv_2.$$

Since $v_2^TS_kA^TAS_kv_2 \geq \sigma_1^2v_2^Tv_2$ and $v_2^TW_kE_kv_2 \geq \delta v_2^Tv_2$, we get $\sigma_1^2(1-\mu_{-1})^{-1}+\delta \leq -\mu_{-1}$ i.e.

$$\mu_{-1} \leq \frac{1}{2}\left(1-\delta-\sqrt{(1+\delta)^2+4\sigma_1^2}\right). \tag{2.29}$$

Noting that $\kappa_2(\mathcal{H}_\delta) = \max\{|\mu_{-n}|, \mu_m\}/\min\{|\mu_{-1}|, \mu_1\}$, the thesis follows from (2.26), (2.27), (2.28) and (2.29). □

Taking into account (2.23), the bound (2.29) is sharper than that provided in [23]. Also, our results are a generalization of those given in [21].

If $\delta = 0$, noting that $\sqrt{1+4\sigma_n^2} < 1+2\sigma_n$, Lemma 2.1 yields

$$\kappa_2(\mathcal{H}_0) \leq \frac{1+\sigma_n}{\min\left\{1,\ \frac{1}{2}(-1+\sqrt{1+4\sigma_1^2}\ )\right\}}. \tag{2.30}$$

Now, suppose σ_1 is significantly smaller than 1. Since $\frac{1}{2}\left(-1+\sqrt{1+4\sigma_1^2}\right) \simeq \sigma_1^2$, it follows that $\kappa_2(\mathcal{H}_0)$ may be much greater than $\kappa_2(B_k)$ (see (2.21)). On the other hand, if $\delta \neq 0$, the augmented system can be viewed as a regularized system and the regularization can improve the conditioning of the system. To show this fact we proceed as in (2.30) and noting that $\sqrt{(1+\delta)^2+4\sigma_1^2} > 1+\delta$ we get

$$\kappa_2(\mathcal{H}_\delta) \leq \frac{1+\sigma_n}{\min\left\{1,\ \frac{1}{2}(\delta-1+\sqrt{(1+\delta)^2+4\sigma_1^2}\ )\right\}} \leq \frac{1+\sigma_n}{\delta}. \tag{2.31}$$

Therefore, when σ_1 is small and $\delta > \sigma_1$, the condition number of $\kappa_2(\mathcal{H}_\delta)$ may be considerably smaller than $\kappa_2(B_k)$ and than $\kappa_2(\mathcal{H}_0)$.

So far we have assumed that σ_1 is small. If this is not the case, the regularization does not deteriorate $\kappa_2(\mathcal{H}_\delta)$ with respect to $\kappa_2(\mathcal{H}_0)$; e.g. if $\sigma_1 \geq 1$ and $\delta \leq 1/2$ we have $(1+\delta)^2+4\sigma_1^2 \geq (2+\delta)^2$, and the bound on $\kappa_2(\mathcal{H}_\delta)$ becomes

$$\kappa_2(\mathcal{H}) \leq \frac{1+\sigma_n}{\frac{1}{2}+\delta}.$$

This discussion suggests that ensuring $\delta > 0$ is a good strategy in order to overcome the potential ill-conditioning of the augmented system. In particular, the regularization is useful when σ_1 is much smaller than 1 and the scalar δ in (2.23) is such that $\delta > \sigma_1$.

3. The regularized Newton-like method. Here we propose a modification of the Newton-like method given in [3] that gives rise to a regularized augmented system. This way the potential ill-conditioning of the augmented system is avoided.

Since the scalar δ is null if at least one element of matrix E_k given in (2.6) is null, to regularize the system (2.22) we need to modify the (2,2) block of the augmented matrix. To this end, we replace (2.8) with the Newton equation

$$W_k D_k N_k p_k = -W_k D_k g_k + r_k, \tag{3.1}$$

where

$$N_k = A^T A + D_k^{-1} E_k + \Delta_k, \tag{3.2}$$

$$\Delta_k = diag(\delta_{k,1}, \delta_{k,2}, \dots, \delta_{k,n}), \qquad \delta_{k,i} \in [0,\, 1),\ i = 1, \dots, n, \tag{3.3}$$

and r_k satisfies (2.9). From Lemma 2 of [16] it can be easily derived that the matrix $W_k D_k N_k$ is invertible for any $x_k > 0$ and there exists a constant $\bar{C}$ independent of k such that

$$\|(W_k D_k N_k)^{-1}\| < \bar{C}. \tag{3.4}$$

Proceeding as to obtain (2.12), (3.1) can be reformulated as the following symmetric and positive definite system:

$$S_k N_k S_k \tilde{p}_k = -S_k g_k + \tilde{r}_k, \tag{3.5}$$

with $\tilde{p}_k = S_k^{-1} p_k$ and $\tilde{r}_k$ satisfying (2.14).

If $\delta_{k,i}$ is strictly positive whenever $(e_k)_i = 0$, the augmented system is regularized and takes the form

$$\begin{pmatrix} I & AS_k \\ S_k A^T & -C_k \end{pmatrix} \begin{pmatrix} \tilde{q}_k \\ \tilde{p}_k \end{pmatrix} = \begin{pmatrix} -(Ax_k - b) \\ 0 \end{pmatrix}, \tag{3.6}$$

$$C_k = W_k E_k + \Delta_k S_k^2. \tag{3.7}$$

The Newton-like method proposed can be globalized using a simple strategy analogous to the one in [3]. Following such strategy, the new iterate x_{k+1} is required to satisfy

$$\frac{\psi_k(x_{k+1} - x_k)}{\psi_k(p_k^C)} \geq \beta, \qquad \beta \in (0,\, 1), \tag{3.8}$$

where, given $p \in \mathbb{R}^n$, $\psi_k(p)$ is the following quadratic function

$$\psi_k(p) = \frac{1}{2} p^T N_k p + p^T g_k,$$

and the step p_k^C is a constrained scaled Cauchy step which approximates the solution to the problem

$$\text{argmin}\{\psi_k(p) : p = -c_k D_k g_k, c_k > 0, x_k + p > 0\}.$$

In practice, p_k^C is given by

$$p_k^C = -c_k D_k g_k, \tag{3.9}$$

with

$$c_k = \begin{cases} \dfrac{g_k^T D_k g_k}{g_k^T D_k N_k D_k g_k}, & \text{if} \quad x_k - \dfrac{g_k^T D_k g_k}{g_k^T D_k N_k D_k g_k} D_k g_k > 0 \\ \theta \operatorname{argmin}\{l > 0, x_k - l D_k g_k \geq 0\}, & \theta \in (0,1), \text{ otherwise} \end{cases} .$$

(3.10)

In particular, letting p_k be a step satisfying (3.1) and $\hat{p}_k$ be the step defined in (2.15), the new iterate has the form

$$x_{k+1} = x_k + t p_k^C + (1-t)\hat{p}_k.$$

If the point $x_{k+1} = x_k + \hat{p}_k$ satisfies (3.8), t is simply taken equal to zero, otherwise a scalar $t \in (0,1]$ is computed in order to satisfy (3.8). The computation of such t can be performed inexpensively as shown in [3].

The global and local convergence properties of the regularized Newton-like method are proved in the next theorem. To maintain fast convergence, eventually it is necessary to control the forcing term η_k in (2.14) and the entries of Δ_k.

THEOREM 3.1. *Let x_0 be an arbitrary strictly positive initial point.*

i) The sequence x_k generated by the regularized Newton-like method converges to x^*.

ii) If $\|\Delta_k\| \leq \Lambda_1 \|W_k D_k g_k\|$ and $\eta_k \leq \Lambda_2 \|W_k D_k g_k\|$, for some positive Λ_1, Λ_2, and k sufficiently large, then the sequence $\{x_k\}$ converges q-quadratically toward x^*.

Proof.

i) Note that,

$$q(x_k) - q(x_k + p) = -\psi_k(p) + \frac{1}{2} p^T (\Delta_k + D_k^{-1} E_k) p > -\psi_k(p)$$

for any $p \in \mathbb{R}^n$. Then, proceeding as in Theorem 2.2 in [3] we easily get the thesis.

ii) In order to prove quadratic rate of convergence, we estimate the norm of the vector $(x_k + p_k - x^*)$ where p_k solves (3.1) and r_k satisfies (2.9). Subtracting the trivial equality $W_k D_k N_k (x^* - x^*) = -W_k D(x^*) g(x^*)$ from (3.1) with a little algebra we get

$$W_k D_k N_k (x_k + p_k - x^*) = W_k \rho_k, \tag{3.11}$$

where

$$\rho_k = D_k \Delta_k (x_k - x^*) + W_k^{-1} r_k - (D_k - D(x^*)) g(x^*) + E_k (x_k - x^*).$$

Letting $\tilde{\rho}_k$ be the vector such that

$$(\tilde{\rho}_k)_i = \frac{e(x_k)_i (x_k - x^*)_i - (d(x_k)_i - d(x^*)_i) g(x^*)_i}{(d_k)_i + (e_k)_i},$$

and using (2.9) we get

$$\begin{aligned} \|W_k \rho_k\| &\leq \|W_k D_k\| \|\Delta_k\| \|x_k - x^*\| + \|r_k\| + \|\tilde{\rho}_k\| \\ &\leq \|\Delta_k\| \|x_k - x^*\| + \eta_k \|W_k D_k g_k\| + \|\tilde{\rho}_k\|. \end{aligned} \tag{3.12}$$

Then, from (3.4), (3.11) and (3.12) it follows

$$\|x_k + p_k - x^*\| \leq \bar{C}(\|\Delta_k\| \, \|x_k - x^*\| + \eta_k \|W_k D_k g_k\| + \|\tilde{\rho}_k\|). \tag{3.13}$$

Moreover, from the proof of Theorem 4 in [16] it can be derived that there exist constants $C_1 > 0$ and $r_1 > 0$ such that

$$\|\tilde{\rho}_k\| \le C_1 \, \|x_k - x^*\|^2, \tag{3.14}$$

whenever $\|x_k - x^*\| \le r_1$. Finally, from the proof of Lemma 3.2 of [3] it follows that there are $C_2 > 0$ and $r_2 > 0$ such that

$$\|W_k D_k g_k\| \le C_2 \|x_k - x^*\|, \tag{3.15}$$

whenever $\|x_k - x^*\| \le r_2$. Then, (3.14) and (3.15) along with (3.13) yield

$$\|x_k + p_k - x^*\| \le \bar{C} \left(\|\Delta_k\| + C_2 \eta_k + C_1 \|x^* - x_k\| \right) \|x^* - x_k\|, \tag{3.16}$$

whenever $\|x_k - x^*\| \le \min(r_1, r_2)$. Therefore, under the assumptions $\|\Delta_k\| \le \Lambda_1 \|W_k D_k g_k\|$ and $\eta_k \le \Lambda_2 \|W_k D_k g_k\|$ for sufficiently large k, there exist $C > 0$ and $r > 0$ such that

$$\|x_k + p_k - x^*\| \le C \|x^* - x_k\|^2, \tag{3.17}$$

whenever $\|x_k - x^*\| \le r$. With this result at hand, slight modifications in the proofs of Lemma 3.2, Lemma 3.3 and Theorem 3.1 of [3] yield the thesis. □

We underline that Theorem 3.1 holds even if x^* is degenerate and the convergence properties of the method in [3] are not degraded.

4. Iterative linear algebra. In this section we focus on the solution of the Newton equation (3.1) via the augmented system (3.6) employing an iterative method.

It is interesting to characterize the entries of the matrix $S(x)$ given in (2.10). When the sequence $\{x_k\}$ generated by our method converges to the solution x^*, the entries of S_k corresponding to the active nondegenerate and possibly degenerate components of x^* tend to zero. Therefore there is a splitting of the matrix S_k in two diagonal blocks $(S_k)_1$ and $(S_k)_2$ such that $\lim_{k\to\infty}(S_k)_1 = I$, $\lim_{k\to\infty}(S_k)_2 = 0$.

More generally, given a small positive threshold $\tau \in (0, 1)$, at each iteration we let

$$\mathcal{L}_k = \{i \in \{1, 2, \dots, n\}, \ s.t. \ (s_k^2)_i \ge 1 - \tau\}, \quad n_1 = card(\mathcal{L}_k), \tag{4.1}$$

where $card(\mathcal{L}_k)$ is the cardinality of the set $\mathcal{L}_k$. Then, for simplicity we omit permutations and assume that

$$S_k = \begin{pmatrix} (S_k)_1 & 0 \\ 0 & (S_k)_2 \end{pmatrix}, \tag{4.2}$$

$$\begin{aligned} (S_k)_1 &= diag_{i\in\mathcal{L}_k}\left((s_k)_i\right) \in \mathbb{R}^{n_1\times n_1}, \\ (S_k)_2 &= diag_{i\notin\mathcal{L}_k}\left((s_k)_i\right) \in \mathbb{R}^{(n-n_1)\times(n-n_1)}. \end{aligned} \tag{4.3}$$

Analogously for any diagonal matrix $G \in \mathbb{R}^{n\times n}$ we let $(G)_1 \in \mathbb{R}^{n_1\times n_1}$ be the submatrix formed by the first n_1 rows and n_1 columns and $(G)_2 \in \mathbb{R}^{(n-n_1)\times(n-n_1)}$ be the submatrix formed by the remaining rows and columns. Finally, we consider the partitioning $A = (A_1, A_2)$, $A_1 \in \mathbb{R}^{m\times n_1}$, $A_2 \in \mathbb{R}^{m\times(n-n_1)}$.

Assume the set $\mathcal{L}_k$ is nonempty. Consequently, the augmented system (3.6) takes the form

$$\begin{pmatrix} I & A_1(S_k)_1 & A_2(S_k)_2 \\ (S_k)_1 A_1^T & -(C_k)_1 & 0 \\ (S_k)_2 A_2^T & 0 & -(C_k)_2 \end{pmatrix} \begin{pmatrix} \tilde{q}_k \\ (\tilde{p}_k)_1 \\ (\tilde{p}_k)_2 \end{pmatrix} = \begin{pmatrix} -(Ax_k - b) \\ 0 \\ 0 \end{pmatrix}, \tag{4.4}$$

and eliminating $(\tilde{p}_k)_2$ from the first equation we get

$$(4.5)\qquad \underbrace{\begin{pmatrix} I+Q_k & A_1(S_k)_1 \\ (S_k)_1 A_1^T & -(C_k)_1 \end{pmatrix}}_{\mathcal{A}_k} \begin{pmatrix} \tilde{q}_k \\ (\tilde{p}_k)_1 \end{pmatrix} = \begin{pmatrix} -(Ax_k - b) \\ 0 \end{pmatrix},$$

where $Q_k = A_2(S_k C_k^{-1} S_k)_2 A_2^T$. We note that

$$(4.6)\qquad \mathcal{A}_k = \begin{pmatrix} I & A_1(S_k)_1 \\ (S_k)_1 A_1^T & -(\Delta_k S_k^2)_1 \end{pmatrix} + \begin{pmatrix} Q_k & 0 \\ 0 & -(W_k E_k)_1 \end{pmatrix},$$

and precondition (4.5) with the matrix

$$(4.7)\qquad \mathcal{P}_k = \begin{pmatrix} I & A_1(S_k)_1 \\ (S_k)_1 A_1^T & -(\Delta_k S_k^2)_1 \end{pmatrix}.$$

The preconditioner $\mathcal{P}_k$ has the following features. As $(w_k)_i(e_k)_i + (s_k)_i^2 = 1$ for $i = 1, \dots, n$, by the definition (4.1) we have $\|(W_k E_k)_1\| \le \tau$, $\|(C_k)_2^{-1}\| \le 1/\tau$, $\|Q_k\| \le (1-\tau)\|A_2\|^2/\tau$; further, when $\{x_k\}$ approaches the solution x^* (hence $(S_k)_1 \to I$ and $(S_k)_2 \to 0$) eventually both Q_k and $(W_k E_k)_1$ tend to zero, i.e. $\|\mathcal{P}_k - \mathcal{A}_k\|$ tends to zero.

Let us observe that the factorization of $\mathcal{P}_k$ can be accomplished based on the identity

$$(4.8)\qquad \mathcal{P}_k = \begin{pmatrix} I & 0 \\ 0 & (S_k)_1 \end{pmatrix} \underbrace{\begin{pmatrix} I & A_1 \\ A_1^T & -(\Delta_k)_1 \end{pmatrix}}_{\Pi_k} \begin{pmatrix} I & 0 \\ 0 & (S_k)_1 \end{pmatrix},$$

and factorizing Π_k. If the set $\mathcal{L}_k$ and the matrix Δ_k remain unchanged for a few iterations, the factorization of the matrix Π_k does not have to be updated. In fact, eventually $\mathcal{L}_k$ is expected to settle down as it contains the indices of all the inactive components of x^* and the indices of the degenerate components i such that $(s_k)_i$ tends to one.

The augmented system (4.5) can be solved by iterative methods for indefinite systems, e.g. BiCGSTAB [24], GMRES [20], QMR [15]. The speed of convergence of these methods depends on the spectral properties of the preconditioned matrix $\mathcal{P}_k^{-1}\mathcal{A}_k$ which are provided in the following theorem.

Theorem 4.1. *Let $\mathcal{A}_k$ and $\mathcal{P}_k$ be the matrices given in (4.5) and (4.7). Then at least $m - n + n_1$ eigenvalues of $\mathcal{P}_k^{-1}\mathcal{A}_k$ are unit and the other eigenvalues are positive and of the form*

$$(4.9)\qquad \lambda = 1 + \mu, \quad \mu = \frac{u^T Q_k u + v^T (W_k E_k)_1 v}{u^T u + v^T (\Delta_k S_k^2)_1 v},$$

where $(u^T, v^T)^T$ is an eigenvector associated to λ.

Proof. The eigenvalues and eigenvectors of matrix $\mathcal{P}_k^{-1}\mathcal{A}_k$ satisfy

$$(4.10)\qquad \begin{pmatrix} I+Q_k & A_1(S_k)_1 \\ (S_k)_1 A_1^T & -(C_k)_1 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \lambda \begin{pmatrix} I & A_1(S_k)_1 \\ (S_k)_1 A_1^T & -(\Delta_k S_k^2)_1 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}.$$

If $\lambda = 1$ we get

$$(I + Q_k)u = u$$
$$(C_k)_1 v = (\Delta_k S_k^2)_1 v$$

i.e. u belongs to the null space of Q_k and v belongs to the null space of $(W_k E_k)_1$. As $rank(Q_k) = n - n_1$, it follows that there are at least $m - (n - n_1)$ unit eigenvalues. If $\lambda \neq 1$, denoting $\lambda = 1 + \mu$ we have

$$u^T Q_k u = \mu u^T u + \mu u^T A_1 (S_k)_1 v$$
$$v^T (W_k E_k)_1 v = -\mu v^T (S_k)_1 A_1^T u + \mu v^T (\Delta_k S_k^2)_1 v$$

Then, adding these two equations we obtain (4.9). Since Q_k, $(W_k E_k)_1$ and $\Delta_k S_k^2$ are positive semidefinite we conclude that μ is positive. □

Clearly, if μ is small it means that the eigenvalues of $\mathcal{P}_k^{-1}\mathcal{A}_k$ are clustered around one and fast convergence of Krylov methods can be expected. This is the case when x_k is close to the solution. On the other hand, when x_k is still far away from x^*, the following bounds for μ can be derived.

COROLLARY 4.1. *Let $\mathcal{A}_k$ and $\mathcal{P}_k$ be the matrices given in (4.6) and (4.7), τ be the scalar in (4.1).*

If the elements $\delta_{k,i}$ in (3.3) are such that $\delta_{k,i} = \delta > 0$ for $i \in \mathcal{L}_k$, then the eigenvalues of $\mathcal{P}_k^{-1}\mathcal{A}_k$ have the form $\lambda = 1 + \mu$ and

$$\mu \leq \frac{\|A_2 (S_k)_2\|^2}{\tau} + \frac{\tau}{\delta(1-\tau)}. \tag{4.11}$$

If the elements $\delta_{k,i}$ in (3.3) are such that $\delta_{k,i} = (w_k)_i (e_k)_i$ for $i \in \mathcal{L}_k$, then the eigenvalues of $\mathcal{P}_k^{-1}\mathcal{A}_k$ have the form $\lambda = 1 + \mu$ and

$$\mu \leq \frac{\|A_2 (S_k)_2\|^2}{\tau} + \frac{1}{1-\tau}. \tag{4.12}$$

Proof. Consider (4.9) and suppose u and v are not null. Then we have

$$\mu \leq \frac{u^T Q_k u}{u^T u} + \frac{v^T (W_k E_k)_1 v}{v^T (\Delta_k S_k^2)_1 v}.$$

Also observe that (4.1) implies

$$\min_{i \in \mathcal{L}_k} (s_k^2)_i \geq 1 - \tau, \quad \|(W_k E_k)_1\| \leq \tau, \quad \|(C_k)_2^{-1}\| \leq \frac{1}{\tau}.$$

Then, when $\delta_{k,i} = \delta > 0$ for $i \in \mathcal{L}_k$, we obtain

$$\mu \leq \|(C_k)_2^{-1}\| \, \|A_2 (S_k)_2\|^2 + \frac{\tau}{\delta(1-\tau)},$$

which yields (4.11).

Letting $\delta_{k,i} = (w_k)_i(e_k)_i$ for $i \in \mathcal{L}_k$, we get

$$\begin{aligned} \mu &\le \|(C_k)_2^{-1}\| \, \|A_2(S_k)_2\|^2 + \frac{\sum_{i\in\mathcal{L}_k} (w_k)_i(e_k)_i v_i^2}{\sum_{i\in\mathcal{L}_k} \delta_{k,i}(s_k^2)_i v_i^2} \\ &\le \frac{\|A_2(S_k)_2\|^2}{\tau} + \frac{\sum_{i\in\mathcal{L}_k} \delta_{k,i} v_i^2}{\sum_{i\in\mathcal{L}_k} (s_k^2)_i \delta_{k,i} v_i^2} \\ &\le \frac{\|A_2(S_k)_2\|^2}{\tau} + \frac{1}{\min_{i\in\mathcal{L}_k} (s_k^2)_i}. \end{aligned} \tag{4.13}$$

Then (4.12) trivially follows from (4.1).

Finally, if either u or v is null the bound (4.9) consists in one of the two terms and the thesis still holds. □

The previous result shows that the choice of the regularization parameters $\delta_{k,i} = \delta > 0$ for $i \in \mathcal{L}_k$ does not provide good properties of the spectrum of $\mathcal{P}_k^{-1}\mathcal{A}_k$ whenever x_k is far from x^*. In fact, to minimize the second term in (4.11), we should fix $\tau = O(\delta)$ but the scalar $\|A_2(S_k)_2\|^2/\tau$ may be large as δ is supposed to be small. On the contrary, letting $\delta_{k,i} = (w_k)_i(e_k)_i$ for $i \in \mathcal{L}_k$ we have a better distribution of the eigenvalues of $\mathcal{P}_k^{-1}\mathcal{A}_k$. Note that for any regularization used it is essential to keep the term $\|A_2(S_k)_2\|^2/\tau$ as small as possible. Hence we advise scaling matrix A at the beginning of the solution process to guarantee that the norm $\|A\|$ is small.

Our regularized augmented system equation (4.5) can be solved by the Projected Preconditioned Conjugate-Gradient (PPCG) method developed in [10, 11]. PPCG provides the vector $\tilde{q}_k$, while the vector $\tilde{p}_k$ is computed by $\tilde{p}_k = (C_k)^{-1}(S_k)A^T\tilde{q}_k$. Solving (4.5) with preconditioner $\mathcal{P}_k$ by PPCG is equivalent to applying Preconditioned Conjugate-Gradient (PCG) to the system

$$\underbrace{(I + Q_k + A_1(S_k C_k^{-1} S_k)_1 A_1^T)}_{\mathcal{F}_k} \tilde{q}_k = -(Ax_k - b), \tag{4.14}$$

using a preconditioner of the form

$$\mathcal{G}_k = I + A_1(\Delta_k)_1^{-1} A_1^T, \tag{4.15}$$

see [12]. Thus we are interested in the distribution of the eigenvalues for matrix $\mathcal{G}_k^{-1}\mathcal{F}_k$.

Theorem 4.2. *Let $\mathcal{F}_k$ and $\mathcal{G}_k$ be the matrices given in (4.14) and (4.15), $\mathcal{L}_k$ be the set given in (4.1). If $\delta_{k,i} = (w_k)_i(e_k)_i$, $i \in \mathcal{L}_k$, then the eigenvalues of $\mathcal{G}_k^{-1}\mathcal{F}_k$ satisfy*

$$1 - \frac{1}{2-\tau} \le \lambda \le 1 + \frac{\|A_2(S_k)_2\|^2}{\tau}. \tag{4.16}$$

Proof. Let λ and $u \in \mathbb{R}^m$ be the eigenvalues and eigenvectors of $\mathcal{G}_k^{-1}\mathcal{F}_k$. Then λ and u satisfy

$$\lambda = \frac{u^T(I + Q_k + A_1(S_k C_k^{-1} S_k)_1 A_1^T)u}{u^T(I + A_1(\Delta_k)_1^{-1} A_1^T)u}.$$

Then, noting that for any $z > 0$ and positive a and b such that $b > a$ we have $(z+a)/(z+b) > a/b$, it follows:

$$\begin{aligned}\lambda &> \frac{u^T A_1 (S_k C_k^{-1} S_k)_1 A_1^T u}{u^T A_1 (\Delta_k)_1^{-1} A_1^T u} \\ &= \frac{((\Delta_k)_1^{-\frac{1}{2}} A_1^T u)^T (S_k)_1 \left((W_k E_k)_1 (\Delta_k)^{-1} + (S_k^2)_1\right)^{-1} (S_k)_1 ((\Delta_k)_1^{-\frac{1}{2}} A_1^T u)}{\|(\Delta_k)_1^{-\frac{1}{2}} A_1^T u\|^2}.\end{aligned}$$

Letting $z = (\Delta_k)^{-\frac{1}{2}} A_1^T u$ we obtain:

$$\lambda \geq \frac{\sum_{i\in\mathcal{L}_k} \dfrac{\delta_{k,i}(s_k^2)_i z_i^2}{(w_k)_i(e_k)_i + \delta_{k,i}(s_k^2)_i}}{\sum_{i\in\mathcal{L}_k} z_i^2} \geq \min_{i\in\mathcal{L}_k} \frac{\delta_{k,i}(s_k^2)_i}{(w_k)_i(e_k)_i + \delta_{k,i}(s_k^2)_i},$$

and for $i \in \mathcal{L}_k$

$$\begin{aligned}\frac{\delta_{k,i}(s_k^2)_i}{(w_k)_i(e_k)_i + \delta_{k,i}(s_k^2)_i} &= 1 - \frac{(w_k)_i(e_k)_i}{(w_k)_i(e_k)_i + \delta_{k,i}(s_k^2)_i} \\ &= 1 - \frac{1}{1 + (s_k^2)_i} \\ &\geq 1 - \frac{1}{2-\tau}.\end{aligned}$$

This yields the lower bound in (4.16).

Concerning the upper bound on λ, first observe that for any vector $v \in \mathbb{R}^{n_1}$, we have $v^T (S_k C_k^{-1} S_k)_1 v \leq v^T (\Delta_k)_1^{-1} v$. Hence

$$\begin{aligned}\lambda &= \frac{u^T (I + Q_k + A_1 (S_k C_k^{-1} S_k)_1 A_1^T) u}{u^T (I + A_1 (\Delta_k)_1^{-1} A_1^T) u} \\ &\leq \frac{u^T (I + Q_k + A_1 (\Delta_k)_1^{-1} A_1^T) u}{u^T (I + A_1 (\Delta_k)_1^{-1} A_1^T) u} \\ &\leq 1 + \frac{u^T Q_k u}{u^T u} \\ &\leq 1 + \|(C_k)_2^{-1}\| \, \|A_2 (S_k)_2\|^2 \\ &\leq 1 + \frac{\|A_2 (S_k)_2\|^2}{\tau}.\end{aligned}$$

□

Let us observe that since the eigenvectors associated with unit eigenvalues satisfy

$$(Q_k + A_1 (S_k C_k^{-1} S_k)_1 A_1^T - A_1 (\Delta_k)_1^{-1} A_1^T) u = 0,$$

the multiplicity of unit eigenvalues is equal to the dimension of the null space of the matrix $Q_k + A_1 (S_k C_k^{-1} S_k)_1 A_1^T - A_1 (\Delta_k)_1^{-1} A_1^T$. Therefore, the existence of unit eigenvalues is not guaranteed.

It is worth comparing two possible systems solved by iterative methods: augmented system (4.5) which involves matrix $\mathcal{A}_k$ preconditioned with $\mathcal{P}_k$ given by

(4.7) and normal equations (4.14) which involve matrix $\mathcal{F}_k$ preconditioned with $\mathcal{G}_k$ given by (4.15). The bounds on the spectra of the preconditioned matrices are given by Corollary 4.1 and Theorem 4.2, respectively. For $\mathcal{P}_k^{-1}\mathcal{A}_k$ (with regularization $\delta_{k,i} = (w_k)_i(e_k)_i$ for $i \in \mathcal{L}_k$) from (4.9) and (4.12) we have

$$1 \le \lambda \le 1 + \frac{\|A_2(S_k)_2\|^2}{\tau} + \frac{1}{1-\tau}, \tag{4.17}$$

and for $\mathcal{G}_k^{-1}\mathcal{F}_k$ we have (4.16). We observe that for a practical choice of τ close to zero the bounds are comparable. However, preconditioning of augmented system offers a slight advantage: first the matrix $\mathcal{P}_k^{-1}\mathcal{A}_k$ is ensured to have a cluster of $m - n + n_1$ eigenvalues at one, second for τ very close to zero the bound of the ratio of the largest to the smallest eigenvalue of the preconditioned matrix is about two times smaller than that for preconditioned normal equations.

We conclude this section considering the limit case where the set $\mathcal{L}_k$ is empty. In this case, the linear system has the form (3.6) where $\|S_k\| \le 1 - \tau$. To use a short-recurrence method, we can apply PCG to the normal system

$$(S_k^T A^T A S_k + C_k)\tilde{p}_k = -S_k A^T (A x_k - b),$$

with preconditioner $S_k A^T A S_k$. The application of the preconditioner can be performed solving a linear system with matrix

$$\begin{pmatrix} I & AS_k \\ S_k A^T & 0 \end{pmatrix}. \tag{4.18}$$

5. Preliminary numerical results. The numerical results were obtained in double precision using MATLAB 7.0 on a Intel Xeon (TM) 3.4 Ghz, 1GB RAM. The threshold parameters were the same as in [3]: β in (3.8) was set to 0.3, θ in (3.10) and σ in (2.15) were set to 0.9995. A successful termination is declared when

$$\left\{ \begin{array}{l} q_{k-1} - q_k < \tau(1 + q_{k-1}), \\ \|x_k - x_{k-1}\| \le \sqrt{\tau}(1 + \|x_k\|) \\ \|P(x_k + g_k) - x_k\| < \tau^{\frac{1}{3}}(1 + \|g_k\|) \end{array} \right. \qquad \text{or} \qquad \|D_k\, g_k\| \le \tau,$$

with $\tau = 10^{-9}$. A failure is declared when the above tests are not satisfied within 100 iterations. All tests were performed letting the initial guess x_0 be the vector of all ones.

First, we intend to investigate the effect of the regularization strategy on the conditioning of the augmented systems (3.6) and on the behaviour of the interior point method. To this end, we monitor the 1-norm condition number of the arising augmented systems via the estimation provided by the MATLAB function `condest` and compare the following two choices of the regularization parameters (3.3)

$$\delta_{k,i} = 0, \;\; i = 1, \ldots, n, \tag{5.1}$$

$$\delta_{k,i} = \left\{ \begin{array}{ll} 0, \;\; \text{if} \;\; i \notin \mathcal{L}_k, & \\ \min\{\max\{10^{-3},\, (w_k)_i(e_k)_i\},\, 10^{-2}\}, & \text{otherwise}. \end{array} \right. \tag{5.2}$$

We remark that with the choice (5.1) the augmented system is not regularized and the method reduces to the interior point method given in [3]. On the other hand,

Test name	m	n	nnz	Set1		Set2	
				σ_1	σ_n	σ_1	σ_n
illc1033	1033	320	4732	$1.135\,10^{-4}$	2.144	$6.801\,10^{-10}$	2.093
illc1850	1850	712	8758	$1.511\,10^{-3}$	2.123	$1.475\,10^{-9}$	1.845
well1033	1033	320	4732	$1.087\,10^{-2}$	1.807	$1.682\,10^{-8}$	1.785
well1850	1850	712	8758	$1.612\,10^{-2}$	1.794	$8.438\,10^{-8}$	1.683

TABLE 5.1
Names, dimensions, nonzero entries, minimum and maximum singular value of the matrices in `Set1` *and* `Set2`

choice (5.2) is in accordance with the results of Corollary 4.1. The safeguards used in the definition of $\delta_{k,i}$ avoid too small and too large regularization parameters.

The experiments are carried out on two sets of problems. The first set (`Set1`) is made up of `illc1033, illc1850, well1033, well1850` problems from the Harwell Boeing collection [13]. These problems are well-conditioned or moderately ill-conditioned and may be degenerate. In the second set (`Set2`), matrices and right hand sides of the tests in `Set1` are scaled by multiplying rows from index $n-1$ to index m by a factor 16^{-5}. This scaling was used in [1] and it gives rise to very ill-conditioned matrices with σ_1 close to zero. In Table 5.1, for each test we report the size of A, the number nnz of nonzero entries of A, the minimum and maximum singular value σ_1 and σ_n of A. Due to the medium scale of these tests we solve the linear systems (3.6) via the MATLAB backslash operator.

Figures 5.1 and 5.2 show the condition number of the augmented system at each iteration of the interior point method applied to tests in `Set1` and `Set2`, with and without regularization. Regarding tests in `Set1`, it is evident that if the problem is moderately ill-conditioned, the regularization produces a reduction, as expected, of the condition number of the augmented system; in fact, in the regularized case the condition numbers of the matrices A and $\mathcal{H}_\delta$ are comparable. On the other hand, when the problem is not ill conditioned the regularization does not affect the condition number of the augmented system, see e.g. `well1033` and `well1850` problems. Finally, the convergence history of the interior point method seems not to be affected by the introduction of the regularization strategy.

Analyzing the results of runs for test problems in `Set2`, we observe that these problems exhibit very ill-conditioned linear systems and the regularization strategy reduces their condition number by several orders of magnitude. The use of regularization in the interior point method has improved significantly the performance of the solution of `illc1033` and `illc1850`. On the other hand a failure occurred in the solution of `well1033` with and without regularization. Such failures are due to the selection of the scaled Cauchy step for a large number of iterations that makes the method extremely slow.

Now, we intend to investigate the effectiveness of our preconditioning technique. We considered problems where the matrix A is the transpose of the matrices in the `LPnetlib` subset of The University of Florida Sparse Matrix Collection [8]. The vector b is set equal to $b=-Ae$, where e is the vector of all ones. From `LPnetlib` collection we discarded the matrices with $m<1000$ and the matrices that are not full rank, getting a total of 56 matrices. When the 1-norm of A exceeded 10^3, we scaled the matrix using a simple row and column scaling scheme.

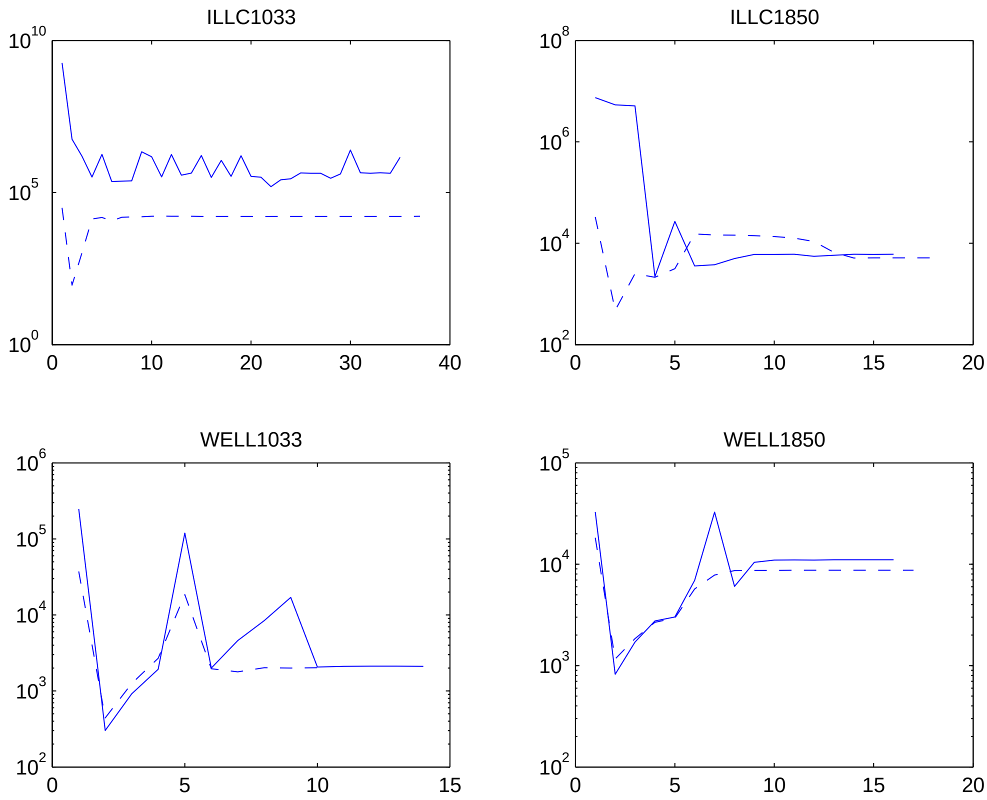

FIG. 5.1. *Condition number of the augmented systems arising in Set 1 with regularization (dashed line) and without regularization (solid line)*

We use the iterative solver PPCG [11] and we set to 100 the maximal number of PPCG iterations. If PPCG was not able to satisfy the stopping criterion within 100 iterations, the algorithm employed the last computed iterate. Regarding the choice of the stopping tolerance for PPCG, a comment is needed. PPCG is an iterative procedure that yields the solution of the indefinite system (4.5) via a CG procedure for the symmetric and positive system (4.14) preconditioned by $\mathcal{G}_k$ given in (4.15). Then, it provides a step $\tilde{p}_k$ such that

$$I + Q_k + A_1(S_k C_k^{-1} S_k)_1 A_1^T \tilde{q}_k = -(Ax_k - b) + \bar{r}_k, \tag{5.3}$$

and monitors the norm of preconditioned residual $\mathcal{G}_k^{-1/2}\bar{r}_k$. Letting

$$\eta_k = \max(500\epsilon_m, \min(10^{-1}, 10^{-2}\|W_k D_k g_k\|),$$

we stop PPCG when $\|\mathcal{G}_k^{-1}\bar{r}_k\|$ drops below *tol* given by:

$$tol = \max(10^{-7}, \frac{\eta_k \|W_k D_k g_k\|}{\|A^T S_k\|_1}).$$

With this adaptive choice of *tol*, the linear systems are solved with a low accuracy when the current iterate is far from the solution, while the accuracy increases as the solution is approached. This choice allows to solve with a sufficient accuracy also the

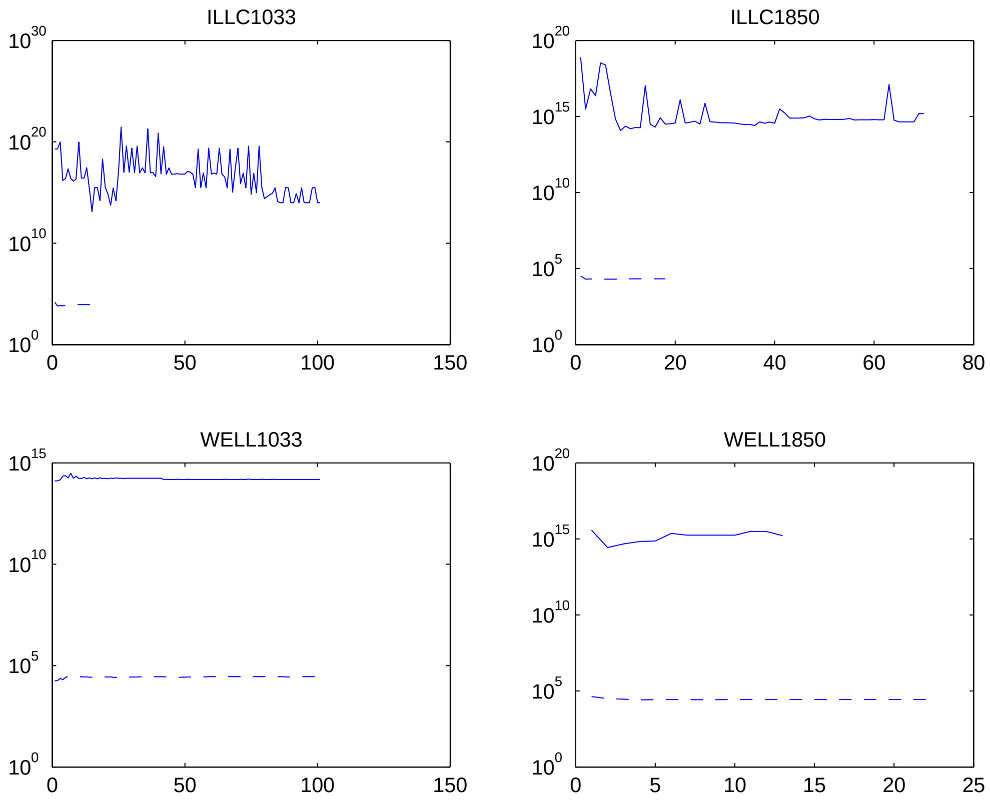

FIG. 5.2. *Condition number of the augmented systems arising in Set 2 with regularization (dashed line) and without regularization (solid line)*

linear system (3.5). Indeed, the unpreconditioned residual $\tilde{r}_k$ in (3.5) is related to the unpreconditioned residual $\bar{r}_k$ in (5.3) as follows:

$$\bar{r}_k = S_k^T A^T \tilde{r}_k;$$

Then, with this choice of *tol* we enforce condition (2.14) with an η_k sufficiently small to ensure quadratic convergence (see Theorem 3.1). Simpler choices of *tol* were not satisfying; for example: $tol = 10^{-3}$ for all k, yields a very inaccurate solution of (3.5) that precludes the convergence of the method. On the other hand, $tol = 10^{-7}$ for all k, yields an exceedingly accurate solution in the first iterations, and this requires many PPCG iterations as typically the preconditioner is not very efficient at the first iterations of the method.

When the set $\mathcal{L}_k$ is empty we use the CG method applied to the normal equations system without any preconditioner.

The regularization parameters δ_k have been chosen according to the rule (5.2). Moreover, to avoid preconditioner updates and factorizations, at iteration $k+1$ we freeze the set $\mathcal{L}_k$ and the vector δ_k if at kth iteration PPCG has succeeded within 30 iterations and the following condition holds:

$$|card(\mathcal{L}_{k+1}) - card(\mathcal{L}_k)| \leq 10.$$

Table 5.2 collects the results of the interior point method on the performed runs.

We report the problem name, the size of A, the number nnz of nonzero elements of A, the number Nit of nonlinear iterations, the overall number Lit of PPCG iterations, the overall number Pf of preconditioner factorizations, the average number $ALit$ of PPCG iterations and the average cardinality Avc of the set $\mathcal{L}_k$. On a total of 56 tests we have 5 failures in solving the problem within 100 nonlinear iterations; these runs are denoted by the symbol $**$.

More insight into these failures, first we note that the linear algebra phase is effectively solved for all problems. Failures in problems `lp_scsd8` and `lp_stocfor3` are recovered allowing up to 300 nonlinear iterations. Specifically, `lp_scsd8` is solved with $Nit = 201$, $Lit = 1704$, $Pf = 8$, $ALit = 8$, $Avc = 366$; `lp_stocfor3` is solved with $Nit = 212$, $Lit = 1956$, $Pf = 67$, $ALit = 9$ and $Avc = 7260$. In the solution of problems `lp_d2q06c` and `lpi_klein3` the progress of the method is very slow as the scaled Cauchy step is taken to update the iterates. On the contrary, in problem `lp_ganges` the projected Newton step is taken at each iterate but the procedure fails to converge in a reasonable number of iteration; we ascribe this failure to the use of the inexact approach as the Newton-like method with direct solver converges in 8 iterations.

We conclude with some comments on these results.

- The interior point method is robust and typically requires a low number of nonlinear iterations. On a total of 56 test problems we have 5 failures.
- The 8 problems where $Avr(n_1)$ is null are such that the solution is the null vector. In practice, for these problems we noted that S_k is very small for all $k \geq 0$. Therefore, as $\Delta_k = 0$, we have $S_k^T A^T A S_k + C_k \simeq I$ and the convergence of the linear solver is very fast. So we did not employ the preconditioner (4.18).
- There are noticeable savings in the number of preconditioner factorizations needed. Focusing on the 43 successfully solved test examples where the preconditioner was used, for 29 of them we avoided to update and factorize the preconditioner in at least 30% of the nonlinear iterations performed.
- For 32 out of 43 problems, n_1 is smaller than $n/2$; that is we have to solve augmented systems of considerably smaller dimension.
- The average number $ALit$ of PPCG iterations is quite low. In case of 40 out of 51 problems $ALit$ does not exceed 40.

REFERENCES

[1] M. Arioli, I.S. Duff, P.P.M. de Rijk, *On the augmented system approach to sparse least-squares problems*, Numer. Math., 55 (1989) pp. 667-684.

[2] A. Altman, J. Gondzio, *Regularized symmetric indefinite systems in interior point methods for linear and quadratic optimization*, Optimization Methods and Software, 11-12 (1999) pp. 275-302.

[3] S. Bellavia, M. Macconi, B. Morini, *An interior Newton-like method for nonnegative least-squares problems with degenerate solution*, Numer. Linear Algebra Appl., 13 (2006) pp. 825-846.

[4] M. Benzi and G. H. Golub and J. Liesen *Numerical solution of saddle point problems*, Acta Numerica, 14 (2005) pp. 1-137.

[5] A. Björck, *Numerical methods for least-squares problems*, SIAM, Philadelphia, 1996.

[6] S.S. Chen, D.L. Donoho, M.A. Saunders *Atomic decomposition by basis pursuit*, SIAM Review 43 (2001) pp. 129-159.

[7] T.F. Coleman, Y. Li, *An interior trust-region approach for nonlinear minimization subject to bounds*, SIAM J. Optim., 6 (1996) pp. 418-445.

[8] T.A. Davis, The University of Florida sparse matrix collection NA Digest, vol. 92, no. 42,

Test name	m	n	nnz	Nit	Lit	Pf	$ALit$	Avc
lp_80bau3b	12061	2262	23264	12	459	8	38	546
lp_bnl2	4486	2324	14996	10	566	8	57	747
lp_czprob	3562	929	10708	6	7	0	1	0
lp_d2q06c	5831	2171	33081	**	**	**	**	**
lp_degen3	2604	1503	25432	21	1033	16	49	586
lp_dfl001	12230	6071	35632	10	417	7	42	829
lp_fffff800	1028	524	6401	100	2384	94	24	69
lp_finnis	1064	497	2760	15	588	10	39	149
lp_fit2d	10524	25	129042	6	6	0	1	0
lp_fit2p	13525	3000	50284	6	6	0	1	0
lp_ganges	1309	1706	6937	**	**	**	**	**
lp_gfrd_pnc	1160	616	2445	7	106	5	15	317
lp_ken_07	3602	2426	8404	12	122	6	10	2273
lp_ken_11	21349	14694	49058	14	255	6	18	14185
lp_ken_13	42659	28632	97246	12	172	8	14	27662
lp_ken_18	154699	105127	358171	11	156	7	14	101180
lp_maros	1966	846	10137	13	500	7	38	323
lp_maros_r7	9408	3136	144848	6	6	0	1	0
lp_osa_07	25067	1118	144812	6	6	0	1	0
lp_osa_14	54797	2337	317097	6	6	0	1	0
lp_osa_30	104374	4350	604488	6	6	0	1	0
lp_osa_60	243246	10280	1408073	6	6	0	1	0
lp_pds_02	7716	2953	16571	10	121	4	12	2709
lp_pds_06	29351	9881	63220	10	192	6	19	9037
lp_pds_10	49932	16558	107605	10	215	9	22	15118
lp_perold	1506	625	6148	34	1732	21	51	247
lp_pilot	4860	1441	44375	13	760	10	58	501
lp_pilot4	1123	410	5264	15	597	8	40	98
lp_pilot87	6680	2030	74949	42	3839	40	91	914
lp_pilot_ja	2267	940	14977	62	4154	50	67	461
lp_pilot_we	2928	722	9265	12	662	9	55	359
lp_pilotnov	2446	975	13331	44	2483	38	56	452
lp_qap8	1632	912	7296	7	11	3	2	11
lp_qap12	8856	3192	38304	7	11	3	2	17
lp_qap15	22275	6330	94950	7	11	3	2	21
lp_scfxm2	1200	660	5469	76	2769	59	36	158
lp_scfxm3	1800	990	8206	84	3048	68	36	237
lp_scrs8	1275	490	3288	10	329	6	33	215
lp_scsd6	1350	147	4316	9	151	5	17	117
lp_scsd8	2750	397	8584	**	**	**	**	**
lp_sctap2	2500	1090	7334	8	68	4	9	10
lp_sctap3	3340	1480	9734	11	439	7	40	46
lp_shell	1777	536	3558	7	68	1	10	528
lp_sierra	2735	1227	8001	12	240	4	20	457
lp_standata	1274	359	3230	7	107	5	15	67
lp_standmps	1274	467	3878	10	216	4	22	94
lp_stocfor2	3045	2157	9357	18	319	8	18	874
lp_stocfor3	23541	16675	72721	**	**	**	**	**
lp_truss	8806	1000	27836	21	395	6	19	942
lp_wood1p	2595	244	70216	22	656	14	30	204
lp_woodw	8418	1098	37487	11	486	6	44	234
lpi_bgindy	10880	2671	66266	7	19	3	3	23
lpi_ceria3d	4400	3576	21178	9	94	2	10	287
lpi_klein3	1082	994	13101	**	**	**	**	**
lpi_cplex1	5224	3005	10947	12	266	9	22	1660
lpi_pilot4i	1123	410	5264	14	638	8	46	95

Table 5.2
Summary of the results for matrices from `LPnetib`

October 16, 1994, NA Digest, vol. 96, no. 28, July 23, 1996, and NA Digest, vol. 97, no. 23, June 7, 1997, `http://www.cise.ufl.edu/research/sparse/matrices`.

[9] R.S. Dembo, S.C. Eisenstat, T. Steihaug, *Inexact Newton methods*, SIAM J. Numer. Anal., 19 (1982) pp. 400-408.

[10] S. Dollar, *Iterative linear algebra for constrained optimization*, Oxford University Computing Laboratory, 2005.

[11] H.S. Dollar, N.I.M. Gould, W.H.A. Schilders, A.J. Wathen *Implicit-Factorization Preconditioning and Iterative Solvers for Regularized Saddle-Point Systems* SIAM. J. Matrix Anal. and Appl. 28 (2006).

[12] H.S. Dollar, N.I.M. Gould, W.H.A. Schilders, A.J. Wathen *Using constraint preconditioners with regularized saddle-point problems*, Computational Optimization and Applications 36 (2007) pp. 249-270.

[13] I. Duff, R. Grimes, J. Lewis, *Sparse Matrix Test Problems*, ACM Trans. Math. Softw., 15 (1989) pp. 1-14.

[14] R. Fletcher, *Conjugate gradient methods for indefinite systems*, Numerical Analysis Dundee 1975, G. Watson ed.,Springer-Verlag, 1976, pp. 73-89.

[15] R. Freund, N. Nachtigal, *Software for simplified Lanczos and QMR algorithms*, Appl. Numer. Math., 19 (1995) pp. 319-341.

[16] M.Heinkenschloss, M. Ulbrich, S. Ulbrich, *Superlinear and quadratic convergence of affine-scaling interior-point Newton methods for problems with simple bounds without strict complementarity assumptions*, Math. Program., 86 (1999) pp. 615-635.

[17] C.L. Lawson, R.J. Hanson, *Solving least-squares problems*, Prentice Hall, 1974.

[18] P. Lotstedt, *Perturbation bounds for the linear least-squares problem subject to linear inequality constraints*, BIT, 23 (1983) pp. 500-519.

[19] C.C. Paige, M.A. Saunders *LSQR: an algorithm for sparse linear equations and sparse least-squares*, ACM Trans. Math. Softw., 8 (1982) pp. 43-71.

[20] Y. Saad, M.H. Schultz, *GMRES: A generalized minimal residual algorithm for solving non-symmetric linear systems*, SIAM J. Sci. Statist. Comput., 7 (1986) pp. 856-869.

[21] M. Saunders, *Solution of sparse rectangular systems using LSQR and CRAIG*, BIT, 35 (1995) pp. 588-604.

[22] M. Saunders, *Cholesky-based methods for sparse least-squares: the benefits of regularization*, in L. Adams and L. Nazareth (eds.) Linear and Nonlinear Conjugate Gradient-Related Methods, SIAM, Philadelphia, 1996, pp. 92–100.

[23] D. Silvester, A. Wathen, *Fast Iterative Solution of Stabilised Stokes Systems Part II: Using General Block Preconditioners*, SIAM J. Numer. Anal. 31 (1994) pp. 1352-1367.

[24] H.A. van der Vorst, *Bi-CGSTAB: a fast and smoothly converging variant of Bi-CG for the solution of nonlinear systems*, SIAM J. Sci. Statist. Comput., 13 (1992) pp. 631-644.