An improvement of the pivoting strategy in the Bunch and Kaufman decomposition, within Truncated Newton methods

Giovanni Fasano and Massimo Roma

Abstract In this work we consider the solution of large scale (possibly nonconvex) unconstrained optimization problems. We focus on Truncated Newton methods which represent one of the commonest methods to tackle such problems. In particular, we follow the approach detailed in [4], where a modified version of the Bunch and Kaufman decomposition [1] is proposed for solving the Newton equation. Such decomposition is used within SYMMBK routine as proposed by Chandra in [6] (see also [7, 16, 17]) for iteratively solving symmetric possibly indefinite linear systems. The proposal in [4] enabled to overcome a relevant drawback of nonconvex problems, namely the computed search direction might not be gradient–related. Here we propose further extensions of such approach, aiming at improving the pivoting strategy of the Bunch and Kaufman decomposition and enhancing its flexibility.

1 Introduction

Given a real valued function $f : \mathbb{R}^n \longrightarrow \mathbb{R}$, an unconstrained optimization problem consists of determining a local minimizer of f by solving

$$\begin{array}{c} \min f(x) \\ x \in \mathbb{R}^n. \end{array} \tag{1}$$

In particular, we consider problems where *n is large* and the function *f is possibly nonconvex*. Moreover, we assume that both the gradient $\nabla f(x)$ and the Hessian

Giovanni Fasano
Department of Management, University Ca' Foscari of Venice, S. Giobbe, Cannaregio 873 – 30121 Venice, Italy, e-mail: fasano@unive.it

Massimo Roma
Dipartimento di Ingegneria Informatica, Automatica e Gestionale "A. Ruberti", SAPIENZA, Università di Roma, via Ariosto, 25 – 00185 Roma, Italy, e-mail: roma@diag.uniroma1.it

matrix $\nabla^2 f(x)$ exist and are continuous. We do not assume any sparsity pattern on $\nabla^2 f(x)$. The iterative solution of large scale unconstrained optimization problems plays a fundamental role in many and different contexts of applied mathematics. Therefore, it is very important to have at one's disposal an efficient and robust method able to tackle also large scale difficult problems (see also [5, 14]).

As well known, in case the method of choice were a Truncated Newton method, at each iteration h, a search direction p_h and a steplength α_h are determined, so that the current point is updated according to the iterative scheme

$$x_{h+1} = x_h + \alpha_h p_h, \tag{2}$$

being $x_0 \in \mathbb{R}^n$ a given starting point. In the Truncated Newton method, the search direction p_h is often obtained by approximately solving the Newton equation

$$\nabla^2 f(x_h) p = -\nabla f(x_h), \tag{3}$$

by means of a Krylov subspace method. The iterations of the solver (called *inner iterations*) are stopped according to a suited termination criterion, still ensuring a good convergence rate of the method. This is obtained by using a particular trade–off rule between the computational burden required to solve the system (3) and the accuracy with which it is solved. The reader is referred to the seminal paper by S. Nash [18] for a survey on Truncated Newton methods.

Among the Krylov subspace methods, the Conjugate Gradient (CG) algorithm is usually the method of choice, even if it may break down when solving (3) and the matrix $\nabla^2 f(x_h)$ is indefinite. In this case, some alternative strategies have been proposed in literature (see, e.g., [8, 9, 10, 11, 12, 13, 15]).

In [4] the use of the SYMMBK algorithm was proposed as an alternative to the CG method. The SYMMBK algorithm was introduced in [6] and it is based on the Lanczos process, which does not break down in the indefinite case. More precisely, the matrix of the Lanczos vectors is built one column at a time and (after k iterations) the resulting $n \times k$ matrix Q_k has the property that $Q_k^T \nabla^2 f(x_h) Q_k = T_k$, where T_k is tridiagonal. Then, the Bunch and Kaufman decomposition of the tridiagonal matrix T_k is performed, namely $T_k = S_k B_k S_k^T$, where B_k is a block diagonal matrix with 1×1 or 2×2 diagonal blocks, and S_k is a unit lower triangular matrix. At each step k, a suited strategy is adopted for deciding whether a 1×1 or 2×2 diagonal block must be formed, in order to guarantee numerical stability.

On the other hand, the test on a pivotal element inside SYMMBK algorithm is uniquely chosen to pursue numerical efficiency and stability, inasmuch as the Bunch and Kaufman decomposition performed by SYMMBK focuses on the growth factor of the matrices resulting from decomposition (see [6]). Thus, some concerns may arise when embedding the SYMMBK algorithm within a Truncated Newton method, where a search direction must be *gradient related*, i.e., eventually *bounded* and of *sufficient descent* (see Definition 1.1 in [4] for a formal statement). The last issue was already addressed in [4], though the modification proposed therein possibly left room to further generalizations. Here we aim at filling the last gap, by proposing an enhancement with respect to [4]. In particular, we are going to propose here an update for the parameters ω and ϕ used in [4], so that they can possibly depend on the gradient vector computed at the current Truncated Newton iteration. More specifically, in the next sections we analyze and discuss the following issues:

- at step k of the Bunch and Kaufman routine, the test $|\delta_k| > \omega\eta\gamma_{k+1}^2$ discussed in [4] represents indeed a test on the curvature along the vector q_k;
- we can replace the quantity ω introduced in [4] with the sequence $\{\omega_k\}$, so that the test $|\delta_k| > \omega\eta\gamma_{k+1}^2$ turns into the test $|\delta_k| > \omega_k\eta\gamma_{k+1}^2$;
- we define a specific expression for the constant ϕ introduced in [4], so that it explicitly depends on the gradient vector currently available from the optimization framework;
- the choice of the sequence $\{\omega_k\}$ and the constant ϕ can partially be steered by the optimization framework, in case any additional knowledge is available which suggests that a better quality of the overall gradient–related direction can be sought.

The paper is organized as follows. In Section 2 some preliminaries on Truncated Newton methods and the Lanczos process are reported; then, the Bunch and Kaufman decomposition, as well as some basics on SYMMBK (see also [7]), are given. In Section 3 we show how to compute a gradient–related direction by using the Bunch and Kaufman decomposition. Finally, Section 4 reports some concluding remarks.

We indicate by $\|\cdot\|$ the Eclidean norm of real vectors and matrices. Moreover, $\lambda_\ell(C)$ and $\kappa(C)$ represent the ℓ-th eigenvalue and the condition number of the real symmetric matrix C, respectively. Finally, e_i is the i-th real unit vector.

2 Preliminary results

Here we report some basic results, including introductory material on the Lanczos process and the Bunch and Kaufman factorization. Some insights on the Lancos process are mandatory, to show how SYMMBK performs an iterative decomposition of the tridiagonal matrix using the Lanczos process. For the sake of brevity, we assume that the reader is familiar with a standard Truncated Newton method which iteratively generates the sequence $\{x_h\}$ in (2). We recall the importance of an efficient truncation criterion for the inner iterations within Truncated Newton methods, as also pointed out in [8, 9, 19], and more recently in [2, 3].

***Assumption* 1**

Let be given the function $f : \mathbb{R}^n \to \mathbb{R}$ in (1), with f twice continuously differentiable. Then, we assume that the sequence $\{x_h\}$ in (2) satisfies $\{x_h\} \subset \Omega$, being $\Omega \subset \mathbb{R}^n$ compact. □

As a very general convergence result for Truncated Newton methods, when convergence to first order stationary points is sought, we give the next proposition.

Proposition 1 *Consider the sequences* $\{x_h\}$ *and* $\{p_h\}$ *in* (2)*, where* $\{x_h\}$ *satisfies Assumption 1 and the search directions are gradient–related. If an Armijo-type linesearch procedure is chosen to select the steplength* α_h *in* (2)*, then*

- $\{f(x_k)\}$ *converges regardless of the choice of the initial iterate* x_0*;*
- *any subsequence of* $\{x_k\}$ *converges to a stationary point of* $f(x)$*.*

Definition 1 Let be given the function $f : \mathbb{R}^n \to \mathbb{R}$ in (1), with f twice continuously differentiable. Consider a vector $d \in \mathbb{R}^n \setminus \{0\}$. Then, the quantity $d^T \nabla^2 f(\bar{x}) d$ is the *normalized curvature* of f at $\bar{x}$, along the direction d.

2.1 Matrix Tridiagonalization using the Lanczos process

The Lanczos process [7] is a Krylov-subspace method for tridiagonalizing a symmetric indefinite matrix. Dropping the dependency of the subscript h and setting $A = \nabla^2 f(x_h)$, $b = -\nabla f(x_h)$ in (3), the application of the Lanczos process to the linear system

$$Ad = b \tag{4}$$

yields the orthogonal Lanczos vectors q_i, $i \geq 1$, according with Table 1.

Data: $\varepsilon = \bar{\varepsilon}\|b\|$, with $\bar{\varepsilon} \in (0,1)$. Set $k = 1$, $u_1 = b$, $q_1 = \frac{u_1}{\|u_1\|}$, $\delta_1 = q_1^T A q_1$ and compute the vector $u_2 = Aq_1 - \delta_1 q_1$.

Do
  $k = k + 1$;
  $\gamma_k = \|u_k\|$;
  If $\gamma_k < \varepsilon$ STOP.
  Else set
    $q_k = u_k / \gamma_k$;
    $\delta_k = q_k^T A q_k$;
    $u_{k+1} = Aq_k - \delta_k q_k - \gamma_k q_{k-1}$.
  End If
End Do

Table 1 The Lanczos process for the indefinite linear system (4).

After $k \leq n$ iterations the Lanczos process has generated the unit vectors $q_1, \dots, q_k$ (the *Lanczos vectors*), along with the values $\delta_1, \dots, \delta_k$ and $\gamma_2, \dots, \gamma_k$, so that setting $Q_k = (q_1 \ \cdots \ q_k)$ and defining the nonsingular tridiagonal matrix

$$T_k = \begin{pmatrix} \delta_1 & \gamma_2 & & & \\ \gamma_2 & \delta_2 & \cdot & & \\ & \cdot & \cdot & \cdot & \\ & & \cdot & \delta_{k-1} & \gamma_k \\ & & & \gamma_k & \delta_k \end{pmatrix}, \tag{5}$$

we have

$$AQ_k = Q_k T_k + \gamma_{k+1} q_{k+1} e_k^T; \tag{6}$$
$$Q_k^T A Q_k = T_k; \tag{7}$$
$$Q_k^T Q_k = I; \tag{8}$$
$$Q_k^T q_{k+1} = 0; \tag{9}$$
$$\text{span}\{q_1, q_2, \ldots, q_k\} = \text{span}\left\{u_1, Au_1, \ldots, A^{k-1}u_1\right\}. \tag{10}$$

In the case $\gamma_{k+1} = 0$ in (6), then we have from (6)-(10) (see [6])

$$\begin{cases} T_k y_k = \|b\| e_1 \\ d_k = Q_k y_k \end{cases} \tag{11}$$

which easily allow to first compute the vector y_k (from $T_k y_k = \|b\| e_1$) and then $d_k = Q_k y_k$, being d_k an *approximate* solution of (4). Finally, note that according to Definition 1 the scalar δ_k represents the normalized curvature of the function $\phi(d) = 1/2 d^T A d + b^T d$, along the direction q_k.

3 Our proposal of possible generalizations

As also reported in Section 2, in a Truncated Newton framework the SYMMBK algorithm can be applied for the solution of the indefinite Newton's equation (4), by exploiting the Bunch and Kaufman factorization. Unfortunately, a mere application of the last decomposition possibly provides in (11) a search direction (namely d_k) which might not be gradient–related. Here, we suitably generalize the proposal in [4] and better exploit the SYMMBK algorithm, so that starting from an approximate solution of (3) we can compute a gradient–related direction.

We intend to slightly modify, at step k, the pivoting strategy adopted by the Bunch and Kaufman iterative decomposition $T_k = S_k B_k S_k^T$, when choosing at step k between 1×1 or 2×2 pivot. Note that the matrix B_k is a block diagonal matrix containing 1×1 and/or 2×2 blocks, while S_k is a unit lower triangular matrix. The standard pivoting strategy in SYMMBK (see [6]) consists of performing at step k a 1×1 pivot if $|\delta_k| > \eta \gamma_{k+1}^2$, otherwise a 2×2 pivot is considered (where η is a suited scalar). This strategy has also an interesting geometric interpretation suggested by Definition 1 and summarized in the next result.

Proposition 2 *Let us consider the Bunch and Kaufman decomposition $T_k = S_k B_k S_k^T$ of the tridiagonal matrix T_k in* (11). *Assume that at step k of the Bunch and Kaufman decomposition the generalized test $|\delta_k| > \omega_k \eta (\gamma_{k+1})^2$ is adopted, with $\omega_k > 0$. Then, the normalized curvature δ_k of the function $\phi(d) = 1/2 d^T A d + b^T d$, along the direction q_k, either satisfies $\delta_k \le (\delta_k)_-$ or $\delta_k \ge (\delta_k)_+$, where*

- *for $k = 1$, for any value of $\omega_1 > 0$ the quantities $(\delta_k)_-$ and $(\delta_k)_+$ are bounded away from zero;*

- *for $k \geq 2$, there are values of $\omega_k > 0$ such that $(\delta_k)_-$ and $(\delta_k)_+$ are bounded away from zero.*

Proof: We separately analyze the cases $k = 1$ and $k \geq 2$. When $k = 1$, recalling that $\|q_i\| = 1$ for any $i \geq 1$, the test $|\delta_1| > \omega_1 \eta (\gamma_2)^2$ is equivalent to the inequalities

$$\begin{aligned} \delta_1 &< -\omega_1 \eta \left[\|Aq_1\|^2 - \delta_1^2 \right] \\ \delta_1 &> +\omega_1 \eta \left[\|Aq_1\|^2 - \delta_1^2 \right]. \end{aligned} \tag{12}$$

Recalling that $\eta = (\sqrt{5} - 1)/(2 \max_i |\lambda_i(A)|)$ (see [6]), after an easy computation the first inequality yields

$$(\delta_1)_- = \frac{1 - \left[1 + 4\omega_1^2 \eta^2 \|Aq_1\|^2\right]^{1/2}}{2\omega_1 \eta} \leq \frac{1 - \left[1 + \omega_1^2 (\sqrt{5} - 1)^2 / \kappa^2(A)\right]^{1/2}}{2\omega_1 \eta},$$

$$(\delta_1)_+ = \frac{1 + \left[1 + 4\omega_1^2 \eta^2 \|Aq_1\|^2\right]^{1/2}}{2\omega_1 \eta} \geq \frac{1 + \left[1 + \omega_1^2 (\sqrt{5} - 1)^2 / \kappa^2(A)\right]^{1/2}}{2\omega_1 \eta}.$$

Similarly, by the second inequality we get

$$(\delta_1)_- = \frac{-1 - \left[1 + 4\omega_1^2 \eta^2 \|Aq_1\|^2\right]^{1/2}}{2\omega_1 \eta} \leq \frac{-1 - \left[1 + \omega_1^2 (\sqrt{5} - 1)^2 / \kappa^2(A)\right]^{1/2}}{2\omega \eta},$$

$$(\delta_1)_+ = \frac{-1 + \left[1 + 4\omega_1^2 \eta^2 \|Aq_1\|^2\right]^{1/2}}{2\omega_1 \eta} \geq \frac{-1 + \left[1 + \omega_1^2 (\sqrt{5} - 1)^2 / \kappa^2(A)\right]^{1/2}}{2\omega \eta}.$$

When $k \geq 2$, after some arrangements, the test $|\delta_k| > \omega_k \eta (\gamma_{k+1})^2$ is equivalent to the inequality

$$|\delta_k| > \omega_k \eta \|Aq_k - \delta_k q_k - \gamma_k q_{k-1}\|^2 = \omega_k \eta \left[\|Aq_k\|^2 - \delta_k^2 - \gamma_k^2 \right].$$

Furthermore, an analysis similar to the case $k = 1$ holds, after distinguishing between the subcases $\|Aq_k\| > \gamma_k$ and $\|Aq_k\| \leq \gamma_k$. □

In the next section we show that, for any k, the generalized test $|\delta_k| > \omega_k \eta (\gamma_{k+1})^2$ reported in Proposition 2 allows to use an adapted SYMMBK algorithm for constructing a *gradient–related* direction.

3.1 Gradient–related directions using SYMMBK

Here we show that, using the results in Proposition 2, within the Bunch and Kaufman decomposition, it is possible to guarantee that a gradient–related direction p_h at x_h

for the optimization problem (1) can be computed, provided that suitable values $\{\omega_k\}$ are used. In this regard, we first recall that to iteratively compute the vector d_k in (11), the Bunch and Kaufman algorithm generates the *intermediate* vectors $\{z_i\}$ (see also [4]), $i \leq k$, such that

- if at step i a 1×1 pivot is performed by the Bunch and Kaufman decomposition, then the vector $z_i = z_{i-1} + \zeta_i w_i$ is generated,
- if at step i a 2×2 pivot is performed by the Bunch and Kaufman decomposition, then the vector $z_i = z_{i-2} + \zeta_{i-1} w_{i-1} + \zeta_i w_i$ is generated,
- when $i = k$ we have $d_k = z_k$,

being the real values $\{\zeta_i\}$ and the vectors $\{w_i\}$ computed using the entries of matrices S_k and B_k. Furthermore, as regards the vectors $\{z_i\}$ we have the next result, which represents a generalization of Proposition 3.1 in [4].

Proposition 3 *Let the matrix A in* (4) *be nonsingular, and let z_i, $i \leq k$, be the directions generated by the SYMMBK algorithm when solving the tridiagonal linear system in* (11)*. Then, for any $0 < \omega_i < 1$, $i \geq 1$, we have that*

- *any direction in the finite sequences $\zeta_1 w_1, \ldots, \zeta_k w_k$ and $z_1, \ldots, z_k$ is bounded;*
- *the vector d_k in* (11) *coincides with z_k (and is bounded).*

Proof: The proof follows guidelines similar to Proposition 3.1 in [4], so that it is omitted. □

From Proposition 3 the real vector $\zeta_i w_i = z_i - z_{i-1}$ (respectively the vector $\zeta_{i-1} w_{i-1} + \zeta_i w_i = z_i - z_{i-2}$) are bounded, and according with the next scheme in Table 2, they can be fruitfully used to compute the search direction p_h, at the outer iteration h of the Truncated Newton method. We remark that the scheme in Table 2 includes a relevant piece of news for the choice of the value ϕ, with respect to the *Reverse Scheme* in [4], as detailed in the proof of Proposition 4.

In the next proposition we show that the direction p_h, obtained by using the scheme in Table 2 at any iterate of the Truncated Newton method, is gradient–related. The next result aims at rephrasing and generalizing the results in Proposition 3.2 of [4], after introducing the sequence $\{\omega_k\}$ in place of the parameter ω and the novel definition for the parameter ϕ in Table 2. Furthermore, the forthcoming result shows that any vector in the sequence $\{p_h\}$ is of *sufficient descent* and eventually is *(uniformly) bounded* by a positive finite constant value.

Proposition 4 *Let Assumption 1 hold. Let us consider Proposition 3 where we set $A = \nabla^2 f(x_h)$ and $b = -\nabla f(x_h)$. Assume the search direction p_h in* (2) *is computed as in Table 2. Then, the direction d_k in* (11) *satisfies $\|d_k\| < \mu$, for any $k \geq 1$, with $\mu > 0$, and p_h is a gradient–related direction.*

Proof: The result surely holds if in Proposition 3 the Lanczos process performs just one iteration, inasmuch as $\gamma_2 < \varepsilon$. On the other hand, in case the Lanczos process has performed at least 2 iterations, the proof follows guidelines similar to those of

Data: Set the initial vector $p_h = z_0 = 0$, along with the parameter $\phi = \bar{\phi}\|b\|$, with $\bar{\phi} > 0$;
Do $i \geq 1$

If at step i of the Bunch and Kaufman decomposition a 1×1 pivot is performed, **then**

If $\nabla f(x_h)^T(\zeta_i w_i) > 0$ **then** set $uu = -\zeta_i w_i$ **else** $uu = \zeta_i w_i$.

Set $p_h = p_h + uu$.

If at step i of the Bunch and Kaufman decomposition a 2×2 pivot is performed, **then** set

$$\tilde{\zeta}_{i-1} = \begin{cases} sgn(\zeta_{i-1}) \max\{|\zeta_{i-1}|, \phi\} & i = 2 \\ \zeta_{i-1} & i > 2. \end{cases}$$

If $\nabla f(x_h)^T(\tilde{\zeta}_{i-1} w_{i-1}) > 0$ **then** set $uu = -\tilde{\zeta}_{i-1} w_{i-1}$ **else** $uu = \tilde{\zeta}_{i-1} w_{i-1}$.

If $\nabla f(x_h)^T(\zeta_i w_i) > 0$ **then** set $vv = -\zeta_i w_i$ **else** $vv = \zeta_i w_i$.

Set $p_h = p_h + uu + vv$.

End Do

Table 2 Computing a gradient–related search direction with SYMMBK algorithm.

Proposition 3.2 in [4], so that we only report the differences. For the case in which a the Lanczos process performs a 2×2 pivot, by Table 2 we obtain

$$\begin{aligned} \nabla f(x_h)^T p_h &= -sgn[\nabla f(x_h)^T(\tilde{\zeta}_1 q_1)] \nabla f(x_h)^T(\tilde{\zeta}_1 q_1) \\ &\quad + \sum_{i>2} -sgn\left[\nabla f(x_h)^T(\zeta_i w_i)\right] \nabla f(x_h)^T(\zeta_i w_i) \\ &\leq -|\tilde{\zeta}_1| \|\nabla f(x_h)\| \;\leq\; -\phi \|\nabla f(x_h)\|^2 \;=\; \bar{\phi} \|\nabla f(x_h)\|^3, \end{aligned}$$

showing that p_h is gradient related.

As regards the property of boundedness for the vectors d_k in (11) and p_h, for any $h \geq 1$, we first have $d_k = Q_k y_k$, so that $\|d_k\| = \|y_k\| \leq \|T_k^{-1}\| \cdot \|\nabla f(x_h)\| \leq \|S_k^{-T} B_k^{-1} S_k^{-1}\| \|\nabla f(x_h)\| \leq \|S_k^{-1}\|^2 \cdot \|B_k^{-1}\| \cdot \|\nabla f(x_h)\|$. Now, following the analysis of Proposition 3.2 in [4] we can similarly prove that $\|S_k^{-1}\| \leq \beta$, where

$$\begin{aligned} \beta = \left(\frac{m_1}{\bar{\omega}\eta\varepsilon}\right) + \Bigg[\frac{k - m_1}{2} \max\Bigg(4 \max_{\ell}\{|\lambda_\ell(\nabla^2 f(x_h))|\}\left(\frac{1}{\varepsilon} + \tilde{\omega}\eta\right), \\ \frac{16}{\varepsilon^2 \xi} \max_{\ell}\{|\lambda_\ell(\nabla^2 f(x_h))|\}^2\Bigg)\Bigg] + k, \end{aligned}$$

being

$$\bar{\omega} = \min_{i \text{ is } 1\times1 \text{ pivot step}} \{\omega_i\}$$

and

$$\tilde{\omega} = \max_{i \text{ is } 2\times2 \text{ pivot step}} \{\omega_i\}.$$

In addition, we also need to provide a suitable bound for the diagonal blocks of B_k^{-1}. From the proof of Proposition 3 and the compactness of Ω we can easily obtain the next results:

- for 1×1 pivot: δ_i^{-1} is a diagonal block of B_k^{-1} and $|\delta_i|^{-1} \leq 1/\bar{\omega}\eta\varepsilon^2$,
- for 2×2 pivot: the reader can refer to the proof of Proposition 3.2 in [4].

□

4 Conclusions

In this paper we have considered efficient Truncated Newton methods for large scale unconstrained optimization problems, where the effective use of a modified Bunch and Kaufman decomposition within the SYMMBK algorithm is considered. We slightly modified the test performed at each iteration of the Bunch and Kaufman decomposition, using the guidelines in [4], so that a more general framework with respect to the last paper is obtained. In particular, we were able to prove that the numerical efficiency of the SYMMBK routine can be suitably coupled with some mild arrangements on the Bunch and Kaufman decomposition, so that the computed search direction for the optimization framework is gradient–related.
We are persuaded that further extensions can be studied, in the case the Truncated Newton method in hand also claims for the global convergence to limit points which satisfy both first and second order necessary optimality conditions. As well known, the accomplishment of the last result needs an accurate analysis of the normalized curvature of the Hessian matrix at any iterate, along any nonzero vector.

References

1. Bunch, J., Kaufman, L.: Some stable methods for calculating inertia and solving symmetric linear equations. Mathematics of Computations **31**, 163–179 (1977)
2. Caliciotti, A., Fasano, G., Nash, S.G., Roma, M.: An adaptive truncation criterion, for linesearch-based truncated Newton methods in large scale nonconvex optimization. Operations Research Letters **46**, 7–12 (2018)
3. Caliciotti, A., Fasano, G., Nash, S.G., Roma, M.: Data and performance profiles applying an adaptive truncation criterion, within linesearch-based truncated Newton methods, in large scale nonconvex optimization. Data in Brief **17**, 246–255 (2018)
4. Caliciotti, A., Fasano, G., Potra, F., Roma, M.: Issues on the use of a modified Bunch and Kaufman decomposition for large scale Newton's equation. Computational Optimization and Applications **77**, 627–651 (2020)
5. Caliciotti, A., Fasano, G., Roma, M.: Preconditioned nonlinear conjugate gradient methods based on a modified secant equation. Applied Mathematics and Computation **318**, 196–214 (2018)
6. Chandra, R.: Conjugate gradient methods for partial differential equations. Ph.D. thesis, Yale University, New Haven (1978). Research Report 129
7. Conn, A.R., Gould, N.I.M., Toint, P.L.: Trust–region methods. MPS–SIAM Series on Optimization, Philadelphia, PA (2000)

8. Dembo, R., Eisenstat, S., Steihaug, T.: Inexact Newton methods. SIAM Journal on Numerical Analysis **19**, 400–408 (1982)
9. Dembo, R., Steihaug, T.: Truncated-Newton algorithms for large-scale unconstrained optimization. Mathematical Programming **26**, 190–212 (1983)
10. Fasano, G.: Planar–conjugate gradient algorithm for large–scale unconstrained optimization, Part 1: Theory. Journal of Optimization Theory and Applications **125**, 523–541 (2005)
11. Fasano, G.: Planar–conjugate gradient algorithm for large–scale unconstrained optimization, Part 2: Application. Journal of Optimization Theory and Applications **125**, 543–558 (2005)
12. Fasano, G.: Lanczos-conjugate gradient method and pseudoinverse computation, in unconstrained optimization. Journal of Optimization Theory and Applications **132**, 267–285 (2006)
13. Fasano, G., Roma, M.: Iterative computation of negative curvature directions in large scale optimization. Computational Optimization and Applications **38**, 81–104 (2007)
14. Fasano, G., Roma, M.: Preconditioning Newton–Krylov methods in nonconvex large scale optimization. Computational Optimization and Applications **56**, 253–290 (2013)
15. Grippo, L., Lampariello, F., Lucidi, S.: A truncated Newton method with nonmonotone linesearch for unconstrained optimization. Journal of Optimization Theory and Applications **60**, 401–419 (1989)
16. HSL: A collection of Fortran codes for large scale scientific computation. Http://www.hsl.rl.ac.uk/
17. Marcia, R.: On solving sparse symmetric linear systems whose definiteness is unknown. Applied Numerical Mathematics **58**(4), 449–458 (2008)
18. Nash, S.: A survey of truncated-Newton methods. Journal of Computational and Applied Mathematics **124**, 45–59 (2000)
19. Nash, S., Sofer, A.: Assessing a search direction within a truncated-Newton method. Operations Research Letters **9**, 219–221 (1990)