On the convergence of iterative schemes for solving a piecewise linear system of equations

Nicolas F. Armijo * Yunier Bello-Cruz† Gabriel Haeser ‡

June 14, 2022

Abstract

This paper is devoted to studying the global and finite convergence of the semi-smooth Newton method for solving a piecewise linear system that arises in cone-constrained quadratic programming problems and absolute value equations. We first provide a negative answer via a counterexample to a conjecture on the global and finite convergence of the Newton iteration for symmetric and positive definite matrices. Additionally, we discuss some surprising features of the semi-smooth Newton iteration in low dimensions and its behavior in higher dimensions. Moreover, we present two iterative schemes inspired by the classical Jacobi and Gauss-Seidel methods for linear systems of equations for finding a solution to the problem. We study sufficient conditions for the convergence of both proposed procedures, which are also sufficient for the existence and uniqueness of solutions to the problem. Lastly, we perform some computational experiments designed to illustrate the behavior (in terms of CPU time) of the proposed iterations versus the semi-smooth Newton method for dense and sparse large-scale problems.

,**Keywords:** Piecewise linear system, quadratic programming, semi-smooth Newton method.

2010 AMS Subject Classification: 90C33, 15A48.

1 Introduction

We consider the following piecewise linear system:

$$x^+ + Tx = b, \tag{1}$$

where, denoting by $\mathbb{R}^{n\times n}$ the set of $n \times n$ matrices with real entries, the data consists of a vector $b \in \mathbb{R}^n \equiv \mathbb{R}^{n\times 1}$ and a nonsingular matrix $T \in \mathbb{R}^{n\times n}$. The variable $x = (x_1\, x_2\, \cdots\, x_n)^T$ is a vector in $\mathbb{R}^n$ and x^+ is the projection of x onto $\mathbb{R}^n_+$, which has the i-th component equal to $(x_i)^+ = \max\{x_i, 0\}$. Some works dealing with problem (1) and its generalizations include [1–7]. Solutions of equation (1) are closely related to at least two important classes of well-known problems, such as the quadratic cone-constrained programming:

$$\begin{array}{ll} \text{minimize} & \frac{1}{2}x^T Q x + q^T x, \\ \text{subject to} & x \in \mathbb{R}^n_+, \end{array} \tag{2}$$

*Department of Applied Mathematics, University of São Paulo, Brazil (e-mail: nfarmijo@ime.usp.br). The author was supported by Fapesp grant 2019/13096-2.

†Northern Illinois University, USA (e-mail: yunierbello@niu.edu). The author was supported in part by R&A Grant from NIU.

‡Department of Applied Mathematics, University of São Paulo, Brazil (e-mail: ghaeser@ime.usp.br). The author was supported by CNPq and Fapesp grant 2018/24293-0.

and the absolute value equation [5, 8]:

$$\hat{T}x - |x| = \hat{b}. \tag{3}$$

Namely, the projection onto $\mathbb{R}^n_+$ of a solution of problem (1) with $T = (Q - \mathrm{Id})^{-1}$ and $b = Tq$ satisfies the linear complementarity problem given by the first order optimality conditions of (2) (see [9] for details), while, on the other hand, for $\hat{T} = -2T - \mathrm{Id}$, and $\hat{b} = -2b$, noting that $x^+ = \frac{x+|x|}{2}$ one can see that problems (1) and (3) are equivalent. Here Id denotes the identity matrix. These relations attest the importance of finding novel and efficient iterative procedures for solving problem (1).

In this paper, we first focus our attention on the *semi-smooth Newton method* for solving problem (1), which consists of specifying a particular generalized Jacobian of F at x in the problem of finding the zeroes of

$$F(x) := x^+ + Tx - b, \qquad x \in \mathbb{R}^n. \tag{4}$$

Namely, starting at the point $x^0 \in \mathbb{R}^n$, the semi-smooth Newton iteration is defined by the following simple equation:

$$\left(P(x^k) + T\right) x^{k+1} = b, \qquad k \in \mathbb{N}, \tag{5}$$

where

$$P(x) := \mathrm{diag}(\mathrm{sgn}(x^+)), \quad x \in \mathbb{R}^n. \tag{6}$$

The above iteration was proposed in [10], which was shown to be globally convergent to a solution of problem (1) under suitable assumptions. It has been extensively studied in the literature for solving generalizations of (1); see, for instance, [2, 5, 6, 11]. We emphasize that the global and linear convergence of (5) has been proved only under restricted assumptions related to the norm of the matrices $(T + P(x))^{-1}$ for all $x \in \mathbb{R}^n$, which come from the study of (5) as a contraction fixed point iteration. A promising and novel approach for establishing finite convergence of (5) was proposed in Theorem 3 of [9] under the assumption that the rows of the matrices $(T + P(x))^{-1}$ for all $x \in \mathbb{R}^n$ have a definite sign. It is worth noting that the number of matrices $P(x)$ with $x \in \mathbb{R}^n$ in (6) is finite (2^n to be precise). Hence, if the semi-smooth Newton method (5) converges, this convergence will occur after finitely many steps. In the pursuit of weaker and verifiable sufficient conditions ensuring convergence of the sequence generated by (5), it was conjectured in [2] that iteration (5) converges after finitely many steps if T is symmetric and positive definite. In this paper, we show that this conjecture is false with a counterexample. However, interestingly, we show that this assumption is enough to guarantee the existence and uniqueness of solutions of problem (1). Moreover, the inverse of $P(x) + T$ for all $x \in \mathbb{R}^n$ always exists, and thus the semi-smooth Newton method (5) is well-defined. We will show that although this method may cycle under this assumption, it can never cycle between only two points. In the second part of this paper, we propose two novel iterative processes inspired by the well-known Jacobi and Gauss-Seidel iterative methods for solving linear systems of equations. The main idea is to consider the Newtonian system (5) and apply a Jacobi or a Gauss-Seidel step at each iteration k. The main advantage of doing so is that the iteration is computed by solving a diagonal or a triangular linear system of equations, which is considerably simpler than solving (5). Also, we are able to present new sufficient conditions for the convergence of these two proposed methods, which are related to the classical diagonal dominance and Sassenfeld's criteria. The existence and uniqueness of the solution of the equation is also proved under both assumptions. We then show with an example that the standard diagonal dominance is not enough to ensure the existence of solutions of problem (1). Lastly, numerical results show that the proposed methods are competitive in terms of CPU time if they are compared with the semi-smooth Newton method (5). The numerical illustration suggests that the proposed iterative methods become more efficient when the dimension is high and the matrix is sparse.

1.1 Notations and preliminaries

Next, we quickly present some notations and facts used throughout the paper. The canonical inner product in $\mathbb{R}^n$ will be denoted by $\langle\cdot,\cdot\rangle$ and the induced norm is $\|\cdot\|$. For $x \in \mathbb{R}^n$, $\mathrm{sgn}(x)$ will denote a vector with components equal to 1, 0 or -1 depending on whether the corresponding component of the vector x is positive, zero or negative, respectively. If $a \in \mathbb{R}$ and $x \in \mathbb{R}^n$, then denote $a^+ := \max\{a, 0\}$, $a^- := \max\{-a, 0\}$ and x^+ and x^- the vectors with i-th component equal to $(x_i)^+$ and $(x_i)^-$, respectively, $i = 1, \ldots, n$. Note that x^+ is the projection of x onto the cone $\mathbb{R}^n_+$. The matrix $\mathrm{Id} \in \mathbb{R}^{n\times n}$ denotes the identity matrix. If $x \in \mathbb{R}^n$ then $\mathrm{diag}(x) \in \mathbb{R}^{n\times n}$ will denote a diagonal matrix with (i,i)-th entry equal to x_i, $i = 1, \ldots, n$. Denote $\|M\| := \max\{\|Mx\| \,:\, x \in \mathbb{R}^n,\ \|x\| = 1\}$ for any $M \in \mathbb{R}^{n\times n}$. The following result is well-known and will be needed in the sequel:

Theorem 1.1 (Contraction mapping principle [12], Thm. 8.2.2, page 153)**.** *Let $\Phi : \mathbb{R}^n \to \mathbb{R}^n$. Suppose that there exists $\lambda \in [0,1)$ such that $\|\Phi(y) - \Phi(x)\| \le \lambda\|y - x\|$, for all $x, y \in \mathbb{R}^n$. Then, there exists a unique $\bar{x} \in \mathbb{R}^n$ such that $\Phi(\bar{x}) = \bar{x}$.*

2 The semi-smooth Newton method

In this section, we present and analyze the convergence of the semi-smooth Newton method given by iteration (5) for solving problem (1). In [9, 10], it was shown that the condition $\mathrm{sgn}((x^k)^+) = \mathrm{sgn}((x^{k+1})^+)$ is sufficient to declare that x^{k+1} is a solution of (1). We now show, in fact, that a component-wise version of this stopping criterion holds:

Proposition 2.1 (Component-wise stopping criterion)**.** *Assume that the sequence x^k generated by method* (5) *is well defined. If $\mathrm{sgn}((x_i^{k+1})^+) = \mathrm{sgn}((x_i^k)^+)$ for some $i \in \{1, \ldots, n\}$ and some k, then $F_i(x^{k+1}) = 0$, where F is defined in* (4)*.*

Proof. By the definition of F and x^{k+1}, we have

$$\begin{aligned} F(x^{k+1}) =& (x^{k+1})^+ + Tx^{k+1} - b \\ =& (P(x^{k+1}) + T)x^{k+1} - b \\ =& (P(x^{k+1}) + T)x^{k+1} - (P(x^k) + T)x^{k+1} \\ =& (P(x^{k+1}) - P(x^k))x^{k+1}. \end{aligned}$$

Hence,

$$F_i(x^{k+1}) = (\mathrm{sgn}((x_i^{k+1})^+) - \mathrm{sgn}((x_i^k)^+))(x^{k+1})_i, \tag{7}$$

for any $i \in \{1, \ldots, n\}$, which implies the desired result. □

In our study, a crucial role will be played by diagonal dominance. Let us start by showing that in the most extreme case of diagonal dominance, namely, when the matrix is diagonal, one can list all solutions of the equation and iteration (5) finds a solution in at most two steps.

Proposition 2.2 (Finite convergence for the diagonal case)**.** *Let $b \in \mathbb{R}^n$ and $T \in \mathbb{R}^{n\times n}$ be a diagonal matrix with entries $T = (t_{ii}), i = 1, \ldots, n$ such that $t_{ii} \notin \{0, -1\}$ for all i. Equation* (1) *has no solutions if, and only if, $t_{ii} \in (-1, 0)$ and $b_i < 0$ for some i. If a solution of* (1) *exists, then* (5) *converges in at most two iterations to one of the solutions. In this case, the number of solutions of* (1) *is given by 2^r, where r is the number of indexes i such that $b_i > 0$ and $t_{ii} \in (-1, 0)$.*

Proof. Let $x^0 \in \mathbb{R}^n$ be any starting point. By definition, we have

$$x^1 = (P(x^0) + T)^{-1} b.$$

Now using that T is a diagonal matrix we have the following componentwise expression for x^1:

$$x_i^1 = ((P(x^0) + T)^{-1} b)_i = \begin{cases} \frac{b_i}{t_{ii}}, & x_i^0 \leq 0, \\ \frac{b_i}{1+t_{ii}}, & x_i^0 > 0, \end{cases} \tag{8}$$

for $i = 1, \ldots, n$. For a fixed i, if $t_{ii} > 0$, we have by (8) that $\operatorname{sgn}((x_i^1)^+) = \operatorname{sgn}(b_i^+)$ and since b is fixed we deduce that $\operatorname{sgn}((x_i^2)^+) = \operatorname{sgn}((x_i^1)^+)$. Similarly, if $t_{ii} < -1$, we conclude that $\operatorname{sgn}((x_i^1)^+) = 1 - \operatorname{sgn}(b_i^+)$. Hence, $\operatorname{sgn}((x_i^2)^+) = \operatorname{sgn}((x_i^1)^+)$. Thus, when there is no i such that $t_{ii} \in (-1, 0)$, by Proposition 2.1 we deduce that (5) converges in two steps. In particular one can check that in this case the solution is unique with i-th component equals to

$$\begin{cases} \frac{b_i}{t_{ii}}, & b_i \leq 0, \\ \frac{b_i}{1+t_{ii}}, & b_i > 0, \end{cases}$$

when $t_{ii} > 0$, and

$$\begin{cases} \frac{b_i}{1+t_{ii}}, & b_i \leq 0, \\ \frac{b_i}{t_{ii}}, & b_i > 0, \end{cases}$$

when $t_{ii} < -1$.

Now, it is easy to see from (1) that there is no solution if $b_i < 0$ for some i such that $t_{ii} \in (-1, 0)$. If, however, $b_i \geq 0$ for such i, the solutions for each component-wise equation are given by $\frac{b_i}{t_{ii}}$ and $\frac{b_i}{1+t_{ii}}$, amonting to the desired formula for the number of solutions. Therefore, assuming that the problem has a solution, we conclude that $b_i \geq 0$ for all i such that $t_{ii} \in (-1, 0)$. By (8), it is easy to see that in this case we have $\operatorname{sgn}((x_i^1)^+) = \operatorname{sgn}((x_i^0)^+)$. Using Proposition 2.1 and the computation done previously we conclude that the method converges to a solution in at most two steps. □

An auxiliary result in our analysis follows next.

Proposition 2.3. *Let $x, y \in \mathbb{R}^n$. Then*

i) $(y - x)^T (P(y)x - P(x)y) \leq 0$.

ii) $(y - x)^T (x^+ - y^+) = (y - x)^T (P(x)x - P(y)y) \leq 0$.

Proof. Let $x, y \in \mathbb{R}^n$. Then, we have

$$(y - x)^T (P(x)y - P(y)x) = \sum_{i=1}^{n} (y_i - x_i)(\operatorname{sgn}(y_i^+) x_i - \operatorname{sgn}(x_i^+) y_i).$$

We only have three different cases to be analyzed:

1) if $\operatorname{sgn}(y_i^+) = \operatorname{sgn}(x_i^+) = s$, then $(y_i - x_i)(\operatorname{sgn}(y_i^+) x_i - \operatorname{sgn}(x_i^+) y_i) = -s(y_i - x_i)^2 \leq 0$;

2) if $\operatorname{sgn}(y_i^+) = 1$ and $\operatorname{sgn}(x_i^+) = 0$, then $(y_i - x_i)(\operatorname{sgn}(y_i^+) x_i - \operatorname{sgn}(x_i^+) y_i) = x_i(y_i - x_i) \leq 0$;

3) if $\operatorname{sgn}(y_i^+) = 0$ and $\operatorname{sgn}(x_i^+) = 1$, then $(y_i - x_i)(\operatorname{sgn}(y_i^+) x_i - \operatorname{sgn}(x_i^+) y_i) = -y_i(y_i - x_i) \leq 0$.

Therefore $(y-x)^T(P(x)y-P(y)x) \leq 0$. The second statement is proved in a similar way. $\square$

Until the end of this section we assume that the matrix T is symmetric and positive definite. This assumption has been considered before in [2] where the authors conjectured the global and finite convergence of the method under this assumption. To fully address this conjecture, we begin with an existence and uniqueness result to the solution of equation (1) under this assumption.

Proposition 2.4 (Existence and uniqueness of solutions)**.** *If T is a symmetric and positive definite $n \times n$ matrix and $b \in \mathbb{R}^n$, then problem (1) has one and only one solution.*

Proof. Note first that $T + \mathrm{Id}$ is also symmetric and positive definite. Hence, $(T+\mathrm{Id})^{-1}$ exists. For any symmetric matrix M we denote $\lambda_1(M) \leq \cdots \leq \lambda_n(M)$ the ordered eigenvalues of M. The matrix $T+\mathrm{Id}$ also satisfies $\lambda_i(T+\mathrm{Id}) = 1 + \lambda_i(T)$, $i = 1, \ldots, n$. Since $\lambda_1(T) > 0$, we deduce that

$$\begin{aligned}
\|(T+\mathrm{Id})^{-1}\|_2 &= \lambda_n((T+\mathrm{Id})^{-1}) \\
&= \frac{1}{\lambda_1(T+\mathrm{Id})} \\
&= \frac{1}{\lambda_1(T)+1} \\
&< 1.
\end{aligned}$$

Defining the function $\Phi(x) := (T+\mathrm{Id})^{-1}(b-x^-)$ and using the decomposition $x = x^+ - x^-$ we can easily show that the solutions of problem (1) coincide with the fixed points of $\Phi(x)$. So, since the projection onto closed and convex sets is non-expansive, we prove that $\Phi(\cdot)$ is a contraction and then has a unique fixed point. Indeed,

$$\begin{aligned}
\|\Phi(x) - \Phi(y)\|_2 &= \|(T+\mathrm{Id})^{-1}(y^- - x^-)\|_2 \\
&\leq \|(T+\mathrm{Id})^{-1}\|_2 \|(y^- - x^-)\|_2 \\
&\leq \|(T+\mathrm{Id})^{-1}\|_2 \|x-y\|_2 \\
&= \frac{1}{\lambda_1(T)+1}\|x-y\|_2.
\end{aligned}$$

Thus, the result follows from Theorem 1.1. $\square$

Remark 1. *The uniqueness of the solution of Equation (1) when T is positive definite showed in Theorem 2.4 can also be deduced by using ii) of Proposition 2.3. Indeed, if x and y are two solutions, we have by definition*

$$x^+ + Tx = b,$$

$$y^+ + Ty = b.$$

Subtracting those equations and multiplying them by $(y-x)^T$, we get from Proposition 2.3(ii) that

$$(y-x)^T T(y-x) = (y-x)^T(x^+ - y^+) \leq 0.$$

Thus, when T is positive definite it must hold that $x = y$.

Due to the nature of iteration (5) we have that the sequence $(x^k)_{k\in\mathbb{N}}$ has only a finite number of different elements. This happens since the set $S := \{(P(x)+T)^{-1}b\ ;\ x \in \mathbb{R}^n\}$ has at most 2^n different elements and $(x^k)_{k\in\mathbb{N}} \subseteq S$. The conclusion of this observation is that we have only two possible outcomes: the method converges in a finite number of steps or it cycles. Note also that when $\bar{x}$ solves (1), one has $\bar{x} = (P(\bar{x}) + T)^{-1}b$, hence, by (5), one has $x^{k+1} = \bar{x}$ whenever $P(x^k) = P(\bar{x})$. Thus, the problem amounts to finding a point in the same orthant as a solution. Let us first show that the method can only cycle among three or more points.

Theorem 2.1 (Newton does not cycle between two points). *Let $T \in \mathbb{R}^{n\times n}$ be a symmetric and positive definite matrix and $b \in \mathbb{R}^n$, then the sequence generated by the semi-smooth Newton method* (5) *does not cycle between two points.*

Proof. If the sequence generated by (5) cycles between two points x and y, then

$$(P(x) + T)y = b,$$

$$(P(y) + T)x = b.$$

So, we have that $T(y - x) = P(y)x - P(x)y$. Multiplying by $(y - x)^T$, we obtain from i) of Proposition 2.3 that

$$(y - x)^T T(y - x) = (y - x)^T (P(y)x - P(x)y) \leq 0,$$

which implies $x = y$. □

When $n \leq 2$, let us show that the method in fact does not cycle.

Theorem 2.2 (Finite convergence for low dimensions). *Let $b \in \mathbb{R}^n$ and $T \in \mathbb{R}^{n\times n}$ be a symmetric and positive definite matrix. Then, iteration* (5) *has global and finite convergence for $n = 1, 2$.*

Proof. The case $n = 1$ is a direct consequence of Proposition 2.2. When $n = 2$, note that the sequence of Newton iterates has at most four different elements, however, no cycle of size four is possible, as these points are necessarily in four different quadrants of $\mathbb{R}^2$, coinciding in sign with the solution (which necessarily exists due to Proposition 2.4), which implies convergence.

Now, suppose that there is a cycle among three different points x, y, z, different from $\bar{x}$, the unique solution. Clearly, these four points must lie in different quadrants of $\mathbb{R}^2$. Without loss of generality, let us assume that x and y lie in opposite quadrants, in the sense that $P(x) + P(y) = \mathrm{Id}$, and that the points satisfy the following equations:

$$(P(x) + T)y = b,$$

$$(P(y) + T)z = b,$$

$$(P(z) + T)x = b.$$

We have that $T(y - x) = P(z)x - P(x)y$, and therefore, multiplying by $(y - x)^T$, we obtain

$$\begin{aligned}(y - x)^T T(y - x) &= (y - x)^T (P(z)x - P(x)y) \\ &= (y_1 - x_1)(\mathrm{sgn}(z_1^+)x_1 - \mathrm{sgn}(x_1^+)y_1) + (y_2 - x_2)(\mathrm{sgn}(z_2^+)x_2 - \mathrm{sgn}(x_2^+)y_2).\end{aligned}$$

Since z is not in the opposite quadrant of x we only have two options, $\mathrm{sgn}(z_1^+) = \mathrm{sgn}(x_1^+)$ and $\mathrm{sgn}(z_2^+) \neq \mathrm{sgn}(x_2^+)$, or $\mathrm{sgn}(z_1^+) \neq \mathrm{sgn}(x_1^+)$ and $\mathrm{sgn}(z_2^+) = \mathrm{sgn}(x_2^+)$. In both cases it can be checked that $(y - x)^T T(y - x) \leq 0$, which leads to a contradiction. The result now follows from Theorem 2.1. □

In our pursuit of understanding the finite convergence of the semi-smooth Newton method (5), we ran extensive numerical experiments. For symmetric and positive definite matrices $T \in \mathbb{R}^{n\times n}$ (1000 randomly generated problems for each dimension n), we recorded the number of iterations that (5) needed to converge at each dimension. The results are shown below in Table 1, where it is shown the percentage of problems solved for several different problem dimensions and the corresponding iterations.

n	1 iter	2 iter	3 iter	4 iter	5 iter
4	7.2	49.1	35.9	7	0.8
8	0.6	37.7	48.8	11.4	1.5
16	0	16.6	63.1	19.5	0.8
32	0	6	69	24.1	0.9
64	0	1.6	68.5	29.4	0.5
128	0	0.3	60.1	39.1	0.5
256	0	0	57.3	41.9	0.8
512	0	0	50.9	48.8	0.3
1024	0	0	43	56.9	0.1
2048	0	0	36.6	63.2	0.2
4096	0	0	30.6	69.2	0.2

Table 1: Percentage of positive definite problems solved at each iteration for different dimensions.

We observe that the higher the dimension, the more iterations are needed. Surprisingly, the method performs exceptionally well as the number of iterations grows very slowly with respect to the growth of n. Indeed, for any $n \leq 4096$, no problem required more than five iterations to be solved.

n/%	1 iter	2 iter	3 iter
4	0	98.9	1.1
8	0	100	0
16	0	100	0
32	0	100	0
64	0	100	0
128	0	99.9	0.01
256	0	98.9	1.1
512	0	95	5
1024	0	86.6	13.4
2048	0	74.8	25.2
4096	0	54.9	45.1

Table 2: Numerical experiments for "almost" diagonal positive definite matrices

In the tests illustrated in Table 2, the matrix T is symmetric, positive definite, and "almost" diagonal, i.e., the elements of the diagonal are much greater compared with the off-diagonal entries. The diagonal entries are of the order of thousands, while the off-diagonal elements are smaller than one. Here, as suggested by Proposition 2.2, no problem required more than three iterations to reach the solution.

Although the numerical tests based on random data suggest that the semi-smooth Newton method does not cycle, we were able to find a particular example of problem (1) that shows that the Newton iteration (5) may cycle among three points in $\mathbb{R}^3$ even when T is symmetric and positive definite. This gives us a counterexample to the conjecture raised in [2] on the global and finite convergence of the Newton iteration (5) under this assumption.

Example 1 (Counterexample on the finite convergence). *The semi-smooth Newton method* (5) *fails to converge in the case* $n = 3$ *with the following symmetric and positive definite data*

$$T = \frac{1}{100}\begin{pmatrix} 32 & -26 & 21 \\ -26 & 33 & -23 \\ 21 & -23 & 17 \end{pmatrix} \quad \text{and} \quad b = \frac{1}{100}\begin{pmatrix} 18 \\ -48 \\ 30 \end{pmatrix}. \tag{9}$$

It can be easily checked that the following points x, y and z conform a cycle of the method.

$$x = \begin{pmatrix} \frac{319}{1435} \\ \\ -\frac{1849}{6379} \\ \\ \frac{190}{1191} \end{pmatrix}, \quad y = \begin{pmatrix} -\frac{527}{2978} \\ \\ -\frac{1490}{923} \\ \\ -\frac{81}{2777} \end{pmatrix}, \quad z = \begin{pmatrix} -\frac{306}{95} \\ \\ \frac{18}{95} \\ \\ 6 \end{pmatrix}.$$

Those points satisfy the cycle equations

$$(P(x) + T)y = b,$$
$$(P(y) + T)z = b,$$
$$(P(z) + T)x = b,$$

which proves the statement.

3 Jacobi-Newton and Gauss-Seidel-Newton methods

Based on the well-known Jacobi and Gauss-Seidel methods for solving linear systems, we define and analyze two novel methods, called the Jacobi-Newton method and Gauss-Seidel-Newton method, for solving (1). We start with two definitions related with the classical diagonal dominance and Sassenfeld's criteria.

Definition 3.1 (Strongly diagonal dominance). *Let $T = (t_{ij}) \in \mathbb{R}^{n\times n}$. We say that T is strongly diagonal dominant if*

$$\frac{1}{|t_{ii}|}\left(1 + \sum_{\substack{j=1\\ j\neq i}}^{n} |t_{ij}|\right) < 1, \quad \forall i = 1, \ldots, n.$$

Note that if T is strongly diagonal dominant, then T is diagonal dominant. However, we need this stronger condition to ensure the global convergence of the Jacobi-Newton method, which will be presented later in (10).

We now introduce a weaker condition for T, which is a variation of the classical Sassenfeld's condition.

Definition 3.2 (Strong Sassenfeld's condition)**.** *Let $T = (t_{ij}) \in \mathbb{R}^{n\times n}$. Define β_i and β as follows*

$$\beta_1 := \frac{1}{|t_{11}|}\left(1 + \sum_{j=2}^{n} |t_{1j}|\right),$$

$$\beta_i := \frac{1}{|t_{ii}|}\left(\sum_{j=1}^{i-1} |t_{ij}|\beta_j + \sum_{j=i+1}^{n} |t_{ij}| + 1\right), \quad \forall i = 2, \ldots, n,$$

and

$$\beta := \max_{i=1,\ldots,n.} \beta_i.$$

We say that T satisfies the strong Sassenfeld's condition if $\beta < 1$.

It is easy to see that if T is strongly diagonal dominant, then T satisfies the strong Sassenfeld's condition. However, the converse implication is not true in general. Now we prove the existence and uniqueness of solutions for problem (1) under the strong Sassenfeld's condition. This condition will also be used later to prove the convergence of the Gauss-Seidel-Newton method which we will introduce.

Theorem 3.1 (Existence and uniqueness of solutions under strong Sassenfeld's condition)**.** *Let $b \in \mathbb{R}^n$ and $T \in \mathbb{R}^{n\times n}$. If T satisfies the strong Sassenfeld's condition, then problem* (1) *has a unique solution.*

Proof. To prove the existence and uniqueness of the solution, we will use the contraction mapping principle (Theorem 1.1). Let us first decompose the matrix T as a sum of $L + D + U$ where D is the diagonal of T and L and U are the strictly lower and upper parts of T, respectively. We define a mapping

$$\psi(x) := (D + L)^{-1}(-Ux + b - x^+), \quad \forall x \in \mathbb{R}^n$$

and prove that its fixed points are solutions of problem (1) and that it is a contraction.

To show that the fixed points of $\psi(\cdot)$ are solutions of (1), note that $x = \psi(x)$ implies that $(D + L)x = -Ux + b - x^+$, or equivalently, $(D + L + U)x + x^+ = Tx + x^+ = b$.

Next, for any $x, y \in \mathbb{R}^n$, we define $u := \psi(x)$ and $w := \psi(y)$, and we observe that

$$u - w = \psi(x) - \psi(y) = (D + L)^{-1}(-U(x - y) + y^+ - x^+),$$

which is equivalent to

$$(D + L)(u - w) = -U(x - y) + y^+ - x^+.$$

Hence,

$$u_1 - w_1 = \frac{1}{t_{11}}\left(-\sum_{j=2}^{n} t_{1j}(x_j - y_j) + y_1^+ - x_1^+\right),$$

and after taking the absolute value in both sides, we have

$$\begin{aligned}
|u_1 - w_1| &\leq \frac{1}{|t_{11}|}\left(\sum_{j=2}^{n} |t_{1j}||x_j - y_j| + |x_1^+ - y_1^+|\right) \\
&\leq \frac{1}{|t_{11}|}\left(\sum_{j=2}^{n} |t_{1j}||x_j - y_j| + |x_1 - y_1|\right) \\
&\leq \frac{1}{|t_{11}|}\left(\sum_{j=2}^{n} |t_{1j}| + 1\right) \|x - y\|_\infty \\
&\leq \beta_1 \|x - y\|_\infty,
\end{aligned}$$

using the nonexpasiveness of the projections in the second inequality. By induction let us assume that $|u_i - w_i| \leq \beta_i \|x - y\|_\infty$ for $i = 1, \ldots, p-1$, and we will show that $|u_p - w_p| \leq \beta_p \|x - y\|_\infty$. Analogously,

$$u_p - w_p = \frac{1}{t_{pp}}\left(-\sum_{j=1}^{p-1} t_{pj}(u_j - w_j) - \sum_{j=p+1}^{n} t_{pj}(x_j - y_j) + y_p^+ - x_p^+\right),$$

which implies that

$$\begin{aligned}
|u_p - w_p| &\leq \frac{1}{|t_{pp}|}\left(\sum_{j=1}^{p-1} |t_{pj}|\beta_j + 1 + \sum_{j=p+1}^{n} |t_{pj}|\right) \|x - y\|_\infty \\
&= \beta_p \|x - y\|_\infty,
\end{aligned}$$

where we used the induction assumption and the nonexpasiveness of the projection. Now, we are ready to prove the contraction property of ψ by observing that for all $x, y \in \mathbb{R}^n$, we have

$$\begin{aligned}
\|\psi(x) - \psi(y)\|_\infty &= \|u - w\|_\infty \\
&= \max_{i=1,\ldots,n.} |u_i - w_i| \\
&\leq \beta \|x - y\|_\infty.
\end{aligned}$$

Thus, since $\beta < 1$, we have that $\psi(\cdot)$ is a contraction mapping and therefore has a unique fixed point. □

Clearly, if T is strongly diagonal dominant, the result above is also true, i.e., problem (1) has one and only one solution. An immediate question that arises is if this assumption can be relaxed to the classical diagonal dominance, and the existence and uniqueness of solutions would still hold for problem (1). Unfortunately, that is not true in general, as is shown in the following example.

Example 2 (Diagonal dominance is not sufficient for existence of solutions)**.** *Let $b \in \mathbb{R}^n$ and T a diagonally dominant matrix. Then, it is not the case that problem* (1) *always has a solution. For instance, for the following data*

$$T = \frac{1}{100}\begin{pmatrix} -26 & 16 \\ 23 & -33 \end{pmatrix}, \quad b = \frac{1}{100}\begin{pmatrix} -12 \\ 12 \end{pmatrix}.$$

T is diagonal dominant and the points x, y and z, w conform to two different cycles for the Newton iteration (5) *where*

$$x = \frac{1}{2295}\begin{pmatrix} -498 \\ 582 \end{pmatrix}, \quad y = \frac{1}{1055}\begin{pmatrix} 498 \\ 18 \end{pmatrix},$$

and

$$z = \frac{1}{245}\begin{pmatrix} 102 \\ -18 \end{pmatrix}, \quad w = -\frac{1}{1405}\begin{pmatrix} 102 \\ 582 \end{pmatrix}.$$

Namely, the cycle equations are satisfied: $(P(x)+T)y = b$, $(P(y)+T)x = b$ and $(P(z)+T)w = b$, $(P(w)+T)z = b$. Thus, the method has two cycles of order two and since the points are in different quadrants we have that the equation has no solution (see the discussion before Theorem 2.1).

To define the Jacobi-Newton iteration for solving problem (1), let us recall the decomposition of the matrix T as a sum of $L + D + U$ where D is the diagonal of T and L and U are the strictly lower and upper parts of T, respectively. The iteration is given by:

$$x^{k+1} := -(P(x^k) + D)^{-1}(L + U)x^k + (P(x^k) + D)^{-1}b, \quad k \in \mathbb{N}, \tag{10}$$

or equivalently, x^{k+1} is the solution of the diagonal system

$$(P(x^k) + D)x^{k+1} = -(L + U)x^k + b. \tag{11}$$

First, we prove that if the sequence generated by the Jacobi-Newton iteration (10) converges to $\bar{x} \in \mathbb{R}^n$, then $\bar{x}$ is a solution of problem (1).

Proposition 3.1 (Limit points of the Jacobi-Newton iteration are solutions)**.** *If the sequence $(x^k)_{k\in\mathbb{N}}$ generated by* (10) *converges to $\bar{x}$, then $\bar{x}$ solves problem* (1).

Proof. First, we know that $F(x) = x^+ + Tx - b$ is continuous since x^+ is a continuous piecewise linear function and $Tx - b$ is linear and therefore continuous. Rewriting $F(x^{k+1})$ and using (11), we have

$$\begin{aligned} F(x^{k+1}) &= (P(x^{k+1}) + T)x^{k+1} - b, \\ &= (P(x^{k+1}) + T)x^{k+1} - (P(x^k) + D)x^{k+1} - (L + U)x^k, \\ &= (P(x^{k+1}) + D + L + U)x^{k+1} - (P(x^k) + D)x^{k+1} - (L + U)x^k, \\ &= (P(x^{k+1}) - P(x^k))x^{k+1} + (L + U)(x^{k+1} - x^k), \end{aligned}$$

using (11) in the last equality. Now, the above equality implies that

$$\|F(x^{k+1})\| \leq \|(P(x^{k+1}) - P(x^k))x^{k+1}\| + \|L + U\|\|x^{k+1} - x^k\|.$$

Hence,

$$\begin{aligned} \lim_{k\to\infty} \|F(x^{k+1})\| &\leq \lim_{k\to\infty} \|(P(x^{k+1}) - P(x^k))x^{k+1}\| + \lim_{k\to\infty} \|L + U\|\|x^{k+1} - x^k\| \\ &= \lim_{k\to\infty} \|(P(x^{k+1}) - P(x^k))x^{k+1}\|. \end{aligned}$$

If $\bar{x}$ does not have any component equal to 0, then the result is trivial, since for a certain index of the sequence the sign becomes constant. On the other hand, if $\bar{x}_i = 0$ for some i, we have by the continuity of F that

$$|F_i(\bar{x})| = \lim_{k\to\infty} |F_i(x^{k+1})| \leq \lim_{k\to\infty} |(\text{sgn}((x_i^{k+1})^+) - \text{sgn}((x_i^k)^+))x_i^{k+1}| = 0,$$

where the last equality follows from the fact that the i-th component of x^{k+1} goes to zero and $\operatorname{sgn}((x_i^{k+1})^+) - \operatorname{sgn}((x_i^k)^+) \in \{-1, 0, 1\}$. Thus, all component of $F(\bar{x})$ are zeroes, proving the result. $\square$

Next, we show that T being strongly diagonal dominant is a sufficient condition for the convergence of (10).

Theorem 3.2 (Global convergence of the Jacobi-Newton method)**.** *Let $b \in \mathbb{R}^n$ and $T \in \mathbb{R}^{n\times n}$. If T is strongly diagonal dominant, then the Jacobi-Newton method* (10) *globally converges to the unique solution of problem* (1)*.*

Proof. Let $\bar{x}$ be the solution of problem (1). Then, $(P(\bar{x}) + D)\bar{x} = -(L+U)\bar{x} + b$, which combined with (11) implies that

$$(P(x^k) + D)(x^{k+1} - \bar{x}) = -(L+U)(x^k - \bar{x}) + P(\bar{x})\bar{x} - P(x^k)\bar{x}.$$

Hence,

$$\begin{aligned}
\|x^{k+1} - \bar{x}\|_\infty &= \| -(P(x^k) + D)^{-1}[(L+U)(x^k - \bar{x}) + (P(\bar{x}) - P(x^k))\bar{x}]\|_\infty \\
&= \max_{i=1,\dots,n} \left| \frac{1}{\operatorname{sgn}((x_i^k)^+) + t_{ii}} \left(-\sum_{\substack{j=1\\ j\neq i}}^n t_{ij}(x_j^k - \bar{x}_j) + \operatorname{sgn}(\bar{x}_i^+)\bar{x}_i - \operatorname{sgn}((x_i^k)^+)\bar{x}_i \right) \right| \\
&\le \max_{i=1,\dots,n} \frac{1}{|\operatorname{sgn}((x_i^k)^+) + t_{ii}|} \left(\sum_{\substack{j=1\\ j\neq i}}^n |t_{ij}||x_j^k - \bar{x}_j| + |\operatorname{sgn}(\bar{x}_i^+)\bar{x}_i - \operatorname{sgn}((x_i^k)^+)\bar{x}_i| \right).
\end{aligned}$$

Note that

$$|\operatorname{sgn}(\bar{x}_i^+)\bar{x}_i - \operatorname{sgn}((x_i^k)^+)\bar{x}_i| = \begin{cases} 0, & \operatorname{sgn}(\bar{x}_i^+) = \operatorname{sgn}((x_i^k)^+) \\ |\bar{x}_i|, & \operatorname{sgn}(\bar{x}_i^+) \neq \operatorname{sgn}((x_i^k)^+). \end{cases}$$

Observe further that $|\operatorname{sgn}(\bar{x}_i^+)\bar{x}_i - \operatorname{sgn}((x_i^k)^+)\bar{x}_i| = |\bar{x}_i| \le |x_i^k - \bar{x}_i|$ when $\operatorname{sgn}(\bar{x}_i^+) \neq \operatorname{sgn}((x_i^k)^+)$ and $|\operatorname{sgn}(\bar{x}_i^+)\bar{x}_i - \operatorname{sgn}((x_i^k)^+)\bar{x}_i| = 0 \le |x_i^k - \bar{x}_i|$ if $\operatorname{sgn}(\bar{x}_i^+) = \operatorname{sgn}((x_i^k)^+)$. Therefore, $|\operatorname{sgn}(\bar{x}_i^+)\bar{x}_i - \operatorname{sgn}((x_i^k)^+)\bar{x}_i| \le |x_i^k - \bar{x}_i|$.

Then,

$$\begin{aligned}
\|x^{k+1} - \bar{x}\|_\infty &\le \max_{i=1,\dots,n} \frac{1}{|\operatorname{sgn}((x_i^k)^+) + t_{ii}|} \left(\sum_{\substack{j=1\\ j\neq i}}^n |t_{ij}||x_j^k - \bar{x}_j| + |x_i^k - \bar{x}_i| \right) \\
&\le \left(\max_{i=1,\dots,n} \frac{1}{|\operatorname{sgn}((x_i^k)^+) + t_{ii}|} \left(\sum_{\substack{j=1\\ j\neq i}}^n |t_{ij}| + 1 \right) \right) \|x^k - \bar{x}\|_\infty \\
&\le \left(\max_{i=1,\dots,n} \frac{1}{|t_{ii}|} \left(\sum_{\substack{j=1\\ j\neq i}}^n |t_{ij}| + 1 \right) \right) \|x^k - \bar{x}\|_\infty.
\end{aligned}$$

The result now follows from the fact that T is strongly diagonal dominant. $\square$

Now we define the Gauss-Seidel-Newton method as follows:

$$x^{k+1} := -(P(x^k) + D + L)^{-1}Ux^k + (P(x^k) + D + L)^{-1}b, \quad k \in \mathbb{N}, \tag{12}$$

or equivalently, x^{k+1} is the solution of the triangular linear system

$$(P(x^k) + D + L)x^{k+1} = -Ux^k + b. \tag{13}$$

Proposition 3.2 (Limit points of the Gauss-Seidel-Newton method are solutions)**.** *If the sequence* $(x^k)_{k\in\mathbb{N}}$ *generated by* (12) *converges to* $\bar{x}$*, then* $\bar{x}$ *solves problem* (1)*.*

Proof. Let $F(x) = x^+ + Tx - b$ and let us rewrite $F(x^{k+1})$ and use (13) to obtain

$$\begin{aligned}
F(x^{k+1}) &= (P(x^{k+1}) + T)x^{k+1} - b, \\
&= (P(x^{k+1}) + T)x^{k+1} - (P(x^k) + D + L)x^{k+1} - Ux^k, \\
&= (P(x^{k+1}) + D + L + U)x^{k+1} - (P(x^k) + D + L)x^{k+1} - Ux^k, \\
&= (P(x^{k+1}) - P(x^k))x^{k+1} + U(x^{k+1} - x^k).
\end{aligned}$$

The result now follows analogously to Proposition 3.1. □

Now, let us show that the strong Sassenfeld's condition gives, besides existence and uniqueness of a solution, a sufficient condition for global convergence of the Gauss-Seidel-Newton method to the solution.

Theorem 3.3 (Global convergence under strong Sassenfeld's condition)**.** *Let* $b \in \mathbb{R}^n$ *and* $T \in \mathbb{R}^{n\times n}$*. If* T *satisfies the strong Sassenfeld's condition, then the Gauss-Seidel-Newton method* (12) *globally converges to the unique solution of* (1)*.*

Proof. Let $\bar{x}$ be the unique solution of (1). From (13), we have

$$(P(x^k) + D - L)(x^{k+1} - \bar{x}) = U(x^k - \bar{x}) + (P(\bar{x}) - P(x^k))\bar{x},$$

which implies

$$x_1^{k+1} - \bar{x}_1 = \frac{1}{\operatorname{sgn}((x_1^k)^+) + t_{11}} \left(-\sum_{j=2}^{n} t_{1j}(x_j^k - \bar{x}_j) + (\operatorname{sgn}(\bar{x}_1^+) - \operatorname{sgn}((x_1^k)^+))\bar{x}_1 \right),$$

then

$$\begin{aligned}
|x_1^{k+1} - \bar{x}_1| &\leq \frac{1}{|\operatorname{sgn}((x_1^k)^+) + t_{11}|} \left(\sum_{j=2}^{n} |t_{1j}||x_j^k - \bar{x}_j| + |\operatorname{sgn}(\bar{x}_1^+) - \operatorname{sgn}((x_1^k)^+)||\bar{x}_1| \right) \\
&\leq \frac{1}{|\operatorname{sgn}((x_1^k)^+) + t_{11}|} \left(\sum_{j=2}^{n} |t_{1j}|\|x^k - \bar{x}\|_\infty + |x_1^k - \bar{x}_1| \right) \\
&\leq \frac{1}{|\operatorname{sgn}((x_1^k)^+) + t_{11}|} \left(1 + \sum_{j=2}^{n} |t_{1j}| \right) \|x^k - \bar{x}\|_\infty \\
&\leq \beta_1 \|x^k - \bar{x}\|_\infty.
\end{aligned}$$

By induction let us assume that $|x_i^{k+1} - \bar{x}_i| \le \beta_i \|x^k - \bar{x}\|_\infty$ for $i = 1, \ldots, p-1$, and we will show that $|x_p^{k+1} - \bar{x}_p| \le \beta_p \|x^k - \bar{x}\|_\infty$. Now, similarly to the proof of Theorem 3.1, we have

$$x_p^{k+1} - \bar{x}_p = \frac{1}{\operatorname{sgn}((x_p^k)^+) + t_{pp}} \left(-\sum_{j=1}^{p-1} t_{pj}(x_j^{k+1} - \bar{x}_j) - \sum_{j=p+1}^{n} t_{pj}(x_j^k - \bar{x}_j) \right.$$
$$\left. + (\operatorname{sgn}(\bar{x}_p^+) - \operatorname{sgn}((x_p^k)^+))\bar{x}_p \right).$$

Taking the absolute value in both sides, we get

$$\begin{aligned} |x_p^{k+1} - \bar{x}_p| \le \frac{1}{|\operatorname{sgn}((x_p^k)^+) + t_{pp}|} & \left(\sum_{j=1}^{p-1} |t_{pj}||x_j^{k+1} - \bar{x}_j| + \sum_{j=p+1}^{n} |t_{pj}||x_j^k - \bar{x}_j| \right. \\ & \left. + |\operatorname{sgn}(\bar{x}_p^+) - \operatorname{sgn}((x_p^k)^+)\bar{x}_p| \right), \\ \le \frac{1}{|\operatorname{sgn}((x_p^k)^+) + t_{pp}|} & \left(\sum_{j=1}^{p-1} |t_{pj}||x_j^{k+1} - \bar{x}_j| + \sum_{j=p+1}^{n} |t_{pj}||x_j^k - \bar{x}_j| + |x_p^k - \bar{x}_p| \right), \\ \le \frac{1}{|\operatorname{sgn}((x_p^k)^+) + t_{pp}|} & \left(\sum_{j=1}^{p-1} |t_{pj}|\beta_j + \sum_{j=p+1}^{n} |t_{pj}| + 1 \right) \|x^k - \bar{x}\|_\infty \\ \le \beta_p \|x^k - \bar{x}\|_\infty. & \end{aligned}$$

Moreover,

$$\begin{aligned} \|x^{k+1} - \bar{x}\|_\infty &= \max_{i=1,\ldots,n.} |x_i^{k+1} - \bar{x}_i| \\ &\le \max_{i=1,\ldots,n.} \beta_i \|x^k - \bar{x}\|_\infty \\ &= \beta \|x^k - \bar{x}\|_\infty, \end{aligned}$$

and the result follows from the fact that $\beta < 1$. □

4 Computational results

To analyze the three methods and see the differences between them in practice, we run several examples of problem (1) applying the Jacobi-Newton, Gauss-Seidel-Newton, and the semi-smooth Newton methods. All codes were implemented in Matlab 9.5.0.944444 (R2018b). We work on two groups of problems for the numerical tests. In the first one, we used dense matrices with different dimensions n equal to 1000, 5000, and 10000. In the second group, we considered a matrix with sparse structure and dimensions 1000, 5000, and 10000. The experiments were run on a 2.3 GHz Intel(R) i5, 16Gb of RAM, and Windows 10 operating system. Next, we describe some details about the implementation.

(1) *Convergence criteria*: We fix the tolerance for the norm of (4) as 10^{-5}. This means that when the 2-norm of $(x^k)^+ + Tx^k - b$ is less than or equal to 10^{-5}, the execution of the algorithm is stopped, and x^k is returned as the solution. The maximum number of iterations was fixed at 1000, but no problem in our test reached it.

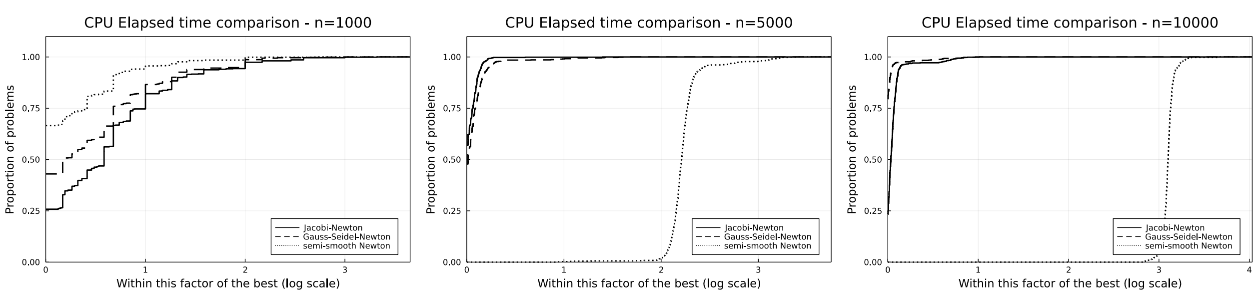
(2) *Generating random problems*: To construct the matrices for the first group of experiments, we used the Matlab routine *rand* to generate a random dense matrix with a predefined dimension with elements between -1 and 1. To achieve convergence of the different methods, we modified the matrices in order to ensure the validity of the strong diagonal dominance condition in Definition 3.1. To do so, we replaced the diagonal entry with 1.001 plus the sum of the off-diagonal elements in absolute value; this ensures the convergence of both Jacobi-Newton and Gauss-Seidel-Newton methods. Finally, for the second group of problems, we evoked *sprand*, a sparse random matrix generator of Matlab, to generate random sparse matrices with entries between -1 and 1 and density 0.3%. The matrices were also similarly modified in order to ensure strong diagonal dominance.

(3) *Solving linear equations*: In each iteration of the semi-smooth Newton, Jacobi-Newton and Gauss-Seidel-Newton methods we need to solve the linear systems (5), (11) and (13) respectively. We used the *backslash* command of Matlab for the dense case since this command uses the diagonality and triangularity structures of the matrices generated by both methods (11) and (13). When solving Newton's linear system for the sparse case, we first use the *symamd* command of Matlab, which makes permutations of rows and columns using the minimum degree algorithm (see [13, 14]) to ensure that the LU-decomposition generates the lowest number of non-zero entries possible.

In order to show the results, we use performance profiles. See [15]. This technique measures and compares the robustness and efficiency of several methods applied to a set of problems using a common indicator, in our case, the CPU time. Let us start by presenting the set of tests with dense matrices.

4.1 Piece-wise linear equation with dense matrices

In the first group of problems we took dense matrices with different dimensions. We show in Figure 1 the results of the performance profiles in our experiments.



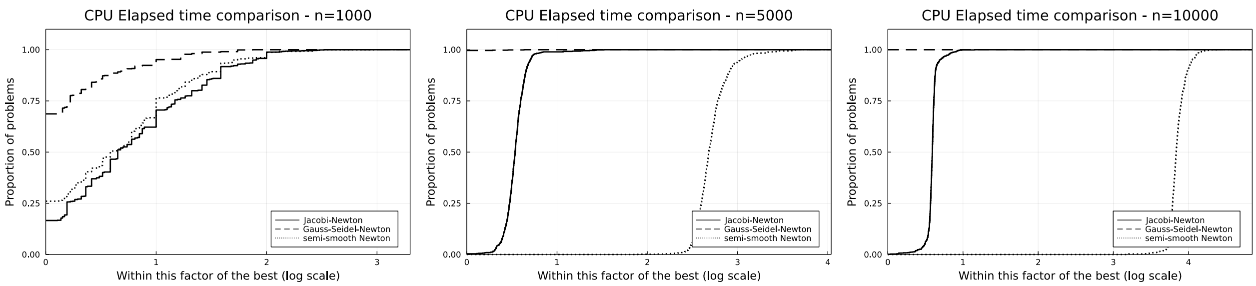
(a) Dense matrices $n = 1000$ (b) Dense matrices $n = 5000$ (c) Dense matrices $n = 10000$

Figure 1: Performance profiles for dense matrices in $\log_2$ scale.

In Figure 1 we compare the robustness and the efficiency of the three methods applied on a set of problems with dimensions 1000 (low), 5000 (mid), and 10000 (high), respectively. The number of problems was fixed at 850 for each dimension. We first see that for the three sets, every problem was solved by the three methods. In the low dimensional test, Newton's method was the most efficient one, being the fastest method for circa 70% of the problems. When the dimension increases, Jacobi-Newton and Gauss-Seidel-Newton are much faster. This difference is accentuated in the highest dimension we tested, where also the Gauss-Seidel variant is now slightly better than Jacobi. In particular, Newton's method took at least four times the time taken by the other methods for almost all mid dimensional problems, while it was at least eight times slower for high dimensional problems. Thus, in our tests, the simplicity of the linear system solved by Gauss-Seidel-Newton and Jacobi-Newton methods (triangular and diagonal, respectively) pays off in comparison with the Newton iteration for mid and high dimensions.

4.2 Piece-wise linear equation with sparse matrices

Our Jacobi and Gauss-Seidel variants of the Newton iterate were devised mainly with large and sparse problems in mind. In Figure 2, we can see the results of our numerical experiments for sparse matrices. Here we also generated 850 problems for each of the dimensions: 1000, 5000, and 10000.



(a) Sparse matrices $n = 1000$ (b) Sparse matrices $n = 5000$ (c) Sparse matrices $n = 10000$

Figure 2: Performance profiles for sparse matrices in $\log_2$ scale.

In these tests, again all problems were solved by all methods, but here, the superiority of the Gauss-Seidel-Newton iterate is already apparent in the low dimensional test, while Jacobi-Newton and Newton behave similarly. The superiority of Gauss-Seidel-Newton is more evident once the dimension increases, being the fastest method for almost all mid and high dimensional problems while Newton becomes considerably slower than both methods. This behavior was already expected and they attest that our Gauss-Seidel variant of Newton's method should be the method of choice for large and sparse problems.

5 Conclusions

In this paper, we considered iterative schemes for solving the piecewise linear equation $x^+ + Tx = b$, where x^+ denotes projection onto the non-negative orthant. This problem appears in solving absolute value equations and minimizing a quadratic function over the non-negative orthant. A semi-smooth Newton method has been proposed for this problem, where the existence and uniqueness of solutions have been studied together with the finite convergence of the method. In [2], the authors conjecture that positive definiteness of T would be sufficient for finite convergence of the semi-smooth Newton method. However, we showed that this assumption is enough only to avoid cycles of size two, in general.

To avoid solving a full linear system of equations at each Newtonian iteration, we proposed Newtonian methods inspired by the classical Jacobi and Gauss-Seidel methods for linear equations, where only a diagonal or triangular linear system is solved at each iteration. The existence and uniqueness of solutions are shown together with global convergence of the methods under stronger variants of the well-known sufficient conditions of convergence for linear systems, namely, diagonal dominance (for the Jacobi iterate) and Sassenfeld's criterion (for the Gauss-Seidel iterate).

Numerical experiments were conducted on random problems to attest that the methods are comparable with the standard Newtonian approach, being considerably faster for large-scale and sparse problems.

For future work, we expect to address the possibility of weakening the sufficient conditions we obtained for the global convergence of the Jacobi-Newton and Gauss-Seidel-Newton iterations. For instance, extensive numerical experiments we run suggests that Gauss-Seidel-Newton converges globally when T is a symmetric and positive definite matrix. Other possibility would be to combine Jacobi and Gauss-Seidel iterates in a SOR-style, which may produce interesting theoretical and numerical results. Additionaly, instead of considering projection onto the non-negative orthant, we expect to address the analogous equations obtained by projecting onto the second-order cone or the semidefinite cone. The situation is more challenging as the projection matrices in those cases do not have such a simple diagonal structure.

References

[1] J. Chen and R. P. Agarwal, "On Newton-type approach for piecewise linear systems," *Linear Algebra and its Applications*, vol. 433, pp. 1463–1471, Dec. 2010.

[2] J.-Y. Bello-Cruz, O. P. Ferreira, S. Z. Németh, and L. F. Prudente, "A semi-smooth Newton method for projection equations and linear complementarity problems with respect to the second order cone," *Linear Algebra Its Appl.*, vol. 513, pp. 160–181, 2017.

[3] A. Griewank, J.-U. Bernt, M. Radons, and T. Streubel, "Solving piecewise linear systems in abs-normal form," *Linear Algebra and its Applications*, vol. 471, pp. 500–530, Apr. 2015.

[4] Z. Sun, L. Wu, and Z. Liu, "A damped semismooth Newton method for the Brugnano–Casulli piecewise linear system," *Bit Numer Math*, vol. 55, pp. 569–589, June 2015.

[5] O. L. Mangasarian, "A generalized Newton method for absolute value equations," *Optim Lett*, vol. 3, pp. 101–108, Jan. 2009.

[6] O. P. Ferreira and S. Z. Németh, "Projection onto simplicial cones by a semi-smooth Newton method," *Optim. Lett.*, vol. 9, no. 4, pp. 731–741, 2015.

[7] J. Barrios, O. P. Ferreira, and S. Z. Németh, "Projection onto simplicial cones by Picard's method," *Linear Algebra Appl.*, vol. 480, pp. 27–43, 2015.

[8] O. L. Mangasarian and R. R. Meyer, "Absolute value equations," *Linear Algebra and its Applications*, vol. 419, pp. 359–367, Dec. 2006.

[9] J. G. Barrios, J. Y. Bello Cruz, O. P. Ferreira, and S. Z. Németh, "A semi-smooth Newton method for a special piecewise linear system with application to positively constrained convex quadratic programming," *J. Comput. Appl. Math.*, vol. 301, pp. 91–100, 2016.

[10] L. Brugnano and V. Casulli, "Iterative solution of piecewise linear systems and applications to flows in porous media," *SIAM J. Sci. Comput.*, vol. 31, no. 3, pp. 1858–1873, 2009.

[11] J. Y. Bello Cruz, O. P. Ferreira, and L. F. Prudente, "On the global convergence of the inexact semi-smooth Newton method for absolute value equation," *Comput. Optim. Appl.*, vol. 65, no. 1, pp. 93–108, 2016.

[12] J. M. Ortega, *Numerical Analysis: A Second Course*. Philadelphia: Society for Industrial and Applied Mathematics, Jan. 1987.

[13] T. A. Davis, *Direct methods for sparse linear systems*. SIAM, 2006.

[14] I. S. Duff, A. M. Erisman, and J. K. Reid, *Direct methods for sparse matrices*. Oxford University Press, 2017.

[15] E. D. Dolan and J. J. Moré, "Benchmarking optimization software with performance profiles," *Mathematical programming*, vol. 91, no. 2, pp. 201–213, 2002.