A low-cost optimization approach for solving minimum norm linear systems and linear least-squares problems

Debora Cores* Johanna Figueroa †

November, 2021

Abstract

Recently, the authors proposed a low-cost approach, named OPALS (Optimization Approach for Linear Systems) for solving any kind of a consistent linear system regarding the structure, characteristics, and dimension of the coefficient matrix A. The results obtained by this approach for matrices with no structure and with indefinite symmetric part were encouraging when compare with other recent and well-known techniques. In this work, we proposed to extend the OPALS approach for solving the Linear Least-Squares Problem (LLSP) and the Minimum Norm Linear System Problem (MNLSP) using any iterative low-cost gradient-type method, avoiding the construction of the matrices A^TA or AA^T, and taking full advantage of the structure and form of the gradient of the proposed nonlinear objective function in the gradient direction. The combination of those conditions together with the choice of the initial iterate allow us to produce a novel and efficient low-cost numerical scheme for solving both problems. Moreover, the scheme presented in this work can also be used and extended for the weighted minimum norm linear systems and minimum norm linear least-squares problems. We include encouraging numerical results to illustrate the practical behavior of the proposed schemes.

Keywords: Nonlinear convex optimization, gradient-type methods, spectral gradient method, minimum norm solution linear systems, linear least-squares solution.

1 Introduction

The adjustment of a data set by linear mathematical models, that yield systems of the form $Ax = b$, appears in a large number of areas, from the scientific field to the social field, since in many situations a linear behavior is found to be an effective simple approximation to the problem. However, frequently this linear model involves an over-determined linear system of equations with no solutions. In this case, the adjustment is made by solving the minimum square norm of the linear system, which is known as a Linear Least-Squares Problem (LLSP),

$$\min_{x\in R^n} \frac{1}{2}\|Ax-b\|_2^2, \tag{1}$$

where $A \in R^{m\times n}$ y $b \in R^m$. If, on the other hand, the adjustment of the data set to a linear model results in an under-determined system with more than one solution, the most suitable solution corresponds to the minimum norm solution of the linear system of equations, known as the Minimum Norm Linear System Problem (MNLSP),

$$\begin{cases} \min_{x\in R^n} \frac{1}{2}\|x\|_2^2 \\ \text{subject to} \quad Ax = b. \end{cases} \tag{2}$$

*Defense University Center at the Spanish Naval Academy, Plaza de España, 36920 Marín, Pontevedra, Spain. E-mail: cores@cud.uvigo.es

†Departamento de Matemática de la Facultad de Ciencias y Tecnología, Universidad de Carabobo (UC), Valencia 2005, Venezuela. E-mail: jfigueroa1@uc.edu.ve

Furthermore, in some applications, as for example the location of sources in magnetoencephalography, the solution of problem (2) favors the sources closest to the surface, so in this case it is necessary to add weights to the location of the different sources, and then a weighted matrix W must be considered, see e.g., [10,12,14]. The incorporation of the matrix W in problem (2) is a generalization of this problem that will be denote as Weighted Minimum Norm Linear System Problem (WMNLSP). This generalization can be written as,

$$\begin{cases} \min_{x\in R^n} \frac{1}{2}\|Wx\|_2^2 \\ \text{subject to} \quad Ax = b. \end{cases} \tag{3}$$

Clearly, problem (2) is a particular case of the WMNLSP (3), where the matrix $W = I_n$.

In some other applications, where the LLSP has infinite solutions, a minimum norm solution of a linear least-squares problem is the desired solution. Then a Minimum Norm Linear Least-Squares Problem (MNLLSP), where $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$, need to be solved,

$$\begin{cases} \min_{x\in \mathbb{R}^n} \frac{1}{2}\|x\|_2^2 \\ \text{suject to} \quad \|Ax-b\|_2^2. \end{cases} \tag{4}$$

Problem (1), (2), (3) and (4) are different, the first one is a quadratic unconstrained optimization problem and the others are quadratic constrained optimization problems, where the restrictions are the linear equations of an under-determined system, except for Problem (4). In the literature there are different methods for solving the Linear Least-Squares Problem (LLSP) as for example the Normal Equations System (NES), $A^TAx = A^Tb$, the QR factorization of the matrix A, the SVD factorization of the matrix A, see e.g., [6,11,17]. The major disadvantage of solving the NES is the increase in the value of the condition number of the normal equation matrix A^TA, which introduces rounding errors. On the other hand, the classical solution of the Minimum Norm Linear System Problem (MNLSP), consists on solving the second order normal equations, $AA^Ty = b$ and then the solution of the problem is obtained as $x = A^Ty$. This last scheme uses $O(\frac{m^2n}{2} + \frac{m^3}{6})$ floating point operations and has the disadvantage that the formation of the matrix AA^T introduces numerical errors in the approach. Also, this problem can be solved using the QR factorization of the matrix associated with the original system. Moreover, the QR decomposition can be done through the Gram-Schmidt orthogonalization method, whose modified version uses $O(m^2n)$ floating point operations or through the application of Householder reflections that requires $O(m^2n - \frac{m^3}{3})$ floating point operations, see e.g., [6,11,17]. As a consequence, it is of great interest to propose low-cost and efficient numerical schemes for solving large scale LLSP and MNLSP, for which the normal equation matrices do not need to be built and few requirements on the matrix associated to the linear system are needed.

In [7], the authors proposed a non-linear approach, named OPALS (Optimization Approach for Linear Systems) for solving any kind of consistent linear systems regarding the structure, characteristics, and dimension of the matrix A. The results obtained by this technique for matrices with no structure and with indefinite symmetric part were encouraging when compared with other recent and well-known techniques. The approach proposed by the authors consists in solving a convex nonlinear optimization problem to obtain the solution of the linear system. In this work, we propose to extend the OPALS approach for solving any LLSP or MNLSP by using iterative low-cost gradient-type methods that avoids the construction of the matrix A^TA or the matrix AA^T. The extension of the OPALS approach takes advantage of the structure of the gradient of the nonlinear objective function already proposed in [7], as well as the specific choice of the initial iterate. The combination of both conditions allow us to present a new and efficient low-cost scheme for solving both problems. Moreover, the scheme presented in this work can also be extended and used for solving the weighted minimum norm linear system problem (3), and the minimum norm linear least-squares problem (4).

The proposed low-cost gradient strategy not only can be used for solving LLSP, MNLSP, WMNLSP and MNLLSP but also can be applied as an iterative method for finding the solution of some linear and nonlinear well known matrix equations as for example [19–21, 25–30] or for deriving new system identification algorithms such as [22–24].

The rest of the work is organized as follows. In Section 2, we present the OPALS strategy for solving consistent linear systems of equations. Section 3 presents the extended OPALS technique for solving MNLSP, WMNLSP, LLSP and MNLLSP. In Section 4, some numerical results are shown, and they illustrate the behavior and performance of the proposed extended approach. The conclusions, recommendations and future work are presented in Section 5.

2 OPALS Strategy

In [7], the authors present the OPALS strategy, which involves a nonlinear optimization approach for solving a consistent linear system of equations of the form:

$$Ax = b, \tag{5}$$

where $A \in R^{m\times n}$ y $b \in R^m$ con $b \in R(A)$. This nonlinear optimization approach considers the nonlinear convex objective function:

$$f_A^b(x) = \sum_{i=1}^{m} f_i(x) = \sum_{i=1}^{m} (e^{-r_i(x)} + e^{r_i(x)}), \tag{6}$$

where $f_i(x) = e^{-r_i(x)} + e^{r_i(x)}$, $r_i(x) = b_i - A_i x$ is the i-th component of the residual vector, $r(x) = b - Ax$, and A_i is the i-th row of the matrix A. The local minimizers of this objective function correspond to the solution(s) of the linear system (5), for more details about OPALS technique and its convergence see [7]. Therefore, OPALS strategy consists in solving the following nonlinear unconstrained optimization problem:

$$\text{find } x^*, \text{ such that } x^* = arg\left(\min_{x\in R^n} f_A^b(x)\right). \tag{7}$$

The gradient of the objective function f_A^b, denoted by g, has the advantage of being a vector in the row space of matrix A, since it can be written as:

$$g(x) = \nabla f_A^b(x) = A^T E(x) \tag{8}$$

where, $E_i(x) = e^{A_i x - b_i} - e^{b_i - A_i x}$, for $i = 1, 2, \cdots, m$, and

$$E(x) = (E_1(x), E_2(x), \ldots, E_m(x))^T. \tag{9}$$

In the OPALS method [7], the convex optimization problem (7) was solved using the Global Spectral Gradient method (GSG), since it is a simple, effective, low-cost, and globally convergent strategy, see [16]. OPALS scheme was very efficient for consistent linear systems with different type of matrices. Moreover, the use of preconditioning strategies could improve the performance of the method [7]. The results obtained for matrices with indefinite symmetric part are encouraging since many well-known iterative methods cannot be used in this case. Additionally, the method was used for square and rectangular matrices, and it was extended to solve linear systems of equations with convex constraints using the Spectral Projected Gradient method (SPG), see [1,3,4,15]. An important feature when using OPALS for solving linear systems of equations without any constraints is that any fast global low-cost gradient-type method could be used for finding the solution of Problem (7), as for example, the Gradient Method with Adaptive step-sizes proposed by Zhou, Gao and Dai, [18], and its recent variants; see e.g., [8,9].

2.1 OPALS-PIV: OPALS with a Particular Initial Vector.

Let $f : R^n \longrightarrow R$ be any continuously differentiable function. Consider any global gradient-type method for solving the following problem:

$$\min_{x \in R^n} f(x), \tag{10}$$

then, the iterations are given by:

$$x_{k+1} = x_k - \alpha_k g_k, \tag{11}$$

where $\alpha_k > 0$ is the step length that guarantees global convergence to a stationary point of problem (10) and $g_k = \nabla f(x_k)$. If the initial vector, x_0, is chosen in the column space of the matrix A^T (denoted by $R(A^T)$), then the sequence of iterates x_k, generated by (11), will remain in the same space, and therefore the solution attained by OPALS will also be in that space. This is a key result and it is presented in the next theorem.

Theorem **1.** *Given a consistent linear system $Cx = d$ where $C \in R^{m \times n}$ and $d \in R^m$. Suppose that any global gradient-type method is applied for minimizing the objective function f_C^d, defined as in (6), and $x_0 \in R(C^T)$ is an initial iterate. Then the sequence of iterates generated for any global gradient-type method of the form (11), belongs to the space $R(C^T)$.*

Proof. The global gradient-type iteration is given by:

$$x_k = x_{k-1} - \alpha_{k-1} g_{k-1}, \tag{12}$$

where, for $k \geq 1$, $g_{k-1} = C^T E(x_{k-1})$, and $E(x_k)$ defined as in equation (9). Therefore,

$$x_k = x_{k-1} - \alpha_{k-1} C^T E(x_{k-1}), \tag{13}$$

applying recursively the equation,

$$x_k = x_0 + C^T \sum_{i=1}^{k} (-\alpha_{i-1}) \, E(x_{i-1}), \tag{14}$$

since $x_0 \in R(C^T)$, the equation can be written as,

$$x_k = C^T \left(\sum_{i=1}^{k} (-\alpha_{i-1}) \, E(x_{i-1}) + p_0 \right), \tag{15}$$

where $x_0 = C^T p_0$, $p_0 \in R^m$. It means that the iterates and the solution generated by any global gradient-type method for minimizing (6) are in $R(C^T)$.

□

Theorem **2.** *Under the assumptions of Theorem 1, any global gradient-type method, with iterates given by equation (11), converges to a solution of the problem:*

$$\begin{cases} \textit{Solve} \quad Cx = d. \\ \textit{subject to:} \quad x \in R(C^T). \end{cases} \tag{16}$$

Proof. Under the conditions of Theorem 1, the system $Cx = d$ is consistent, $d \in R(C)$, and $x_0 \in R(C^T)$. Hence, the global gradient-type method to minimize f_C^d will converge to a vector $x \in R(C^T)$ (the convergence from any x_0 is guarantee for the global behavior of the method). Since, the unique minimizers and stationary points of the objective convex function f_C^d are the solutions of the system $Cx = d$, the gradient-type will converge to the solution of (16). □

Consider the application of OPALS strategy to solve the linear system $Cx = d$, starting with an initial iterate in $R(C^T)$. This particular procedure will be denoted by OPALS-PIV (OPALS with Particular Initial Vector). The purpose of this strategy is to keep the sequence of OPALS iterates generated for solving the system $Cx = d$ in the set $R(C^T)$. To accomplish this objective, the expression of the objective function, the structure of the gradient function, equation (8), as well as the choice of the initial iterate in the set $R(C^T)$, are crucial.

Algorithm 1 OPAPLS-PIV: for solving $Cx = d$, obtaining an approximate solution $x_k \in R(C^T)$.

Input: C, d, x_0, α_0, and all the parameters for the global gradient-type method
Exit: x_k is the obtained approximate solution.
Chose $x_0 \in R(C^T)$.
Apply OPALS until convergence to solve $Cx = d$, obtaining x_k: using any global gradient-type method to minimize $f_C^d(x)$ as in (6), where $g(x) = \nabla f_C^d(x) = C^T E(x)$.

3 Proposed Numerical Strategies

3.1 OPALS-PIV for the solution of MNLSP and WMNLSP

In order to explain the proposed scheme for solving MNLSP, consider the fist order optimality conditions for MNLSP (2),

$$AA^T\mu^* = b \tag{17}$$
$$x^* = A^T\mu^*. \tag{18}$$

It is clear that the Lagrangian Hessian matrix associated to problem (2) is the identity matrix, I_n, of dimension n; see e.g., [2]. Then the second order KKT sufficient conditions are satisfied, that is, for all $\omega \in \mathbb{R}^n$, $\omega \neq 0$, $\omega^T I_n \omega = \omega^T \omega > 0$. Therefore, to find a solution of problem (2) is equivalent to determine a vector $(x^*, \mu^*) \in \mathbb{R}^{n+m}$ that satisfies the first order conditions, equations (17) and (18). Observe that, given $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$, with $b \in R(A)$, the solution of MNLSP is a vector x^* in the column space of the matrix A^T, $x^* = A^T\mu^*$, for some $\mu^* \in \mathbb{R}^m$, such that $Ax^* = b$. So, a solution of MNLSP can be obtained by applying OPALS-PIV, explained in Section 2.1, considering $C = A$, $d = b$ with an initial vector in $R(C^T) = R(A^T)$. From now on we denote this numerical scheme by **OPALS-PIV-MNLS**. Following this procedure, the vector μ^* is computed using (15), by a linear combination of vectors, as follows:

$$\mu_k = \sum_{i=1}^{k} (-\alpha_{i-1})\, E(x_{i-1}) + p_0, \tag{19}$$

where $x_0 = A^T p_0$ and k is the last iteration. The approximation, μ_k, in the case of a unique solution, converges to the unique vector μ^*, and in the case of infinite solutions, it converges to one of them. It is clear that if there exist a solution of problem (2), it must be the unique solution, since it corresponds to the minimization of a strictly convex function on a convex set. So, it does not matter which vector μ^* is being approximated by the OPALS-PIV method, since the vector $x^* = A^T\mu^*$ is unique and hence $x^* \simeq A^T x_k \simeq A^T\mu^*$ converges to the minimum norm solution.

Problem (2) corresponds to the particular case where $W = P^{-1} = I_n$ for the general WMNLS problem:

$$\begin{cases} \text{Find } x^* \text{ such that } x^* = arg\left(\min_{x\in\mathbb{R}^n} \frac{1}{2}\|x\|_W^2\right) \\ \text{suject to:} \quad Ax = b, \end{cases} \tag{20}$$

where, $\|x\|_W = \sqrt{x^T W x}$ and W is a positive definite matrix, allowing problem (20) to be feasible if the linear system is consistent. The general problem (20) becomes relevant for certain applications where the minimum Euclidean standard solution omits information about the real problem [10, 12, 14].

The Lagrangian function associated to problem (20) is given by:

$$L(x, \mu) = \frac{1}{2}x^T W x + \mu^T(b - Ax), \tag{21}$$

where, $\mu \in \mathbb{R}^m$ is the Lagrange multiplier vector. Since the matrix W is positive definite, the local minimizer of problem (20) is a vector, x^*, that satisfies the first order KKT necessary conditions:

$$\nabla L(x^*, \mu^*) = W x^* - A^T \mu^* = 0 \tag{22}$$

$$b - Ax^* = 0. \tag{23}$$

Considering $W = P^{-1}$, where P is a symmetric positive definite matrix, then $P = \Pi\Pi^T$, and equations (22) and (23) can be written as:

$$x^* = \Pi\Pi^T A^T \mu^* \tag{24}$$

$$Ax^* = b. \tag{25}$$

Hence, the solution of problem (20) is a vector x^*, such that $Ax^* = b$ and belongs to the space $R(PA^T) = R(\Pi\Pi^T A^T)$. Once again, we propose to use OPALS-PIV for finding this solution. For that, OPALS-PIV method determines y^* such that $A\Pi y^* = b$, where $y^* \in R(\Pi^T A^T)$, and then $x^* = \Pi y^*$. This approach is presented in the following result.

Theorem **3.** *: Let $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, $Ax = b$ a consistent system, $W = P^{-1}$ a symmetric positive definite matrix such that $P = \Pi\Pi^T$, y^* is computed such that*

$$y^* = arg\left(\min_{y\in\mathbb{R}^n} f^b_{A\Pi}(y)\right), \tag{26}$$

where, $f^b_{A\Pi}$ is defined according to equation (6). Then, if y^* is obtained by OPAPLS-PIV strategy, it follows that $x^* = \Pi y^*$ is the solution of problem (20).

Proof. Since the system $Ax = b$ has a solution and Π is an invertible matrix, then there exists $x = \Pi y$, such that $A\Pi y = b$. Hence, the OPAPLS method guarantees that y is a global minimizer of the function $f^b_{A\Pi}$. So, the application of OPAPLS-PIV strategy to find a minimizer of $f^b_{A\Pi}$ will converge to a vector $y \in R(\Pi^T A^T)$, implying that $y = \Pi^T A^T \mu$ for some $\mu \in \mathbb{R}^m$. Therefore, $A\Pi y = b$ or $Ax = b$ and,

$$x = \Pi y = \Pi\Pi^T A^T \mu = PA^T \mu = W^{-1} A^T \mu,$$

that is, x satisfies conditions (22) and (23) that characterize the solution of problem (20). Thus, $y^* = y$, $x^* = x$ are the solution of the considered problem. □

An important conclusion about this result is that given any $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$, with $b \in R(A)$, the key point to find a numerical scheme for solving problem (20) consists in determining a vector that satisfies conditions (24) and (25), which characterizes the solution. For that, it is also important to observe that these conditions are equivalent to determining a vector, $y^* = \Pi^T A^T \mu$, for some $\mu \in \mathbb{R}^m$, such that $A\Pi y^* = b$ and then construct the vector $x^* = \Pi y^*$. That is, determining a vector y^* in the column space of the matrix $\Pi^T A^T$ that satisfies $A\Pi y^* = b$. According to Theorem 1, we propose to use OPALS-PIV strategy (Section 2.1), considering $C = A\Pi$, $d = b$ and starting with an initial vector $y_0 \in R(C^T) = R(\Pi^T A^T)$ and then the approximate solution is $x = \Pi y_k$, where y_k is the solution obtained by OPALS-PIV. This procedure will be denoted as **OPALS-VIP-WMNLS**.

3.2 OPALS-PIV for the solution of LLSP.

Given $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$, the LLSP consists in,

$$\text{Finding } x^* \text{ such that } x^* = arg \min_{x\in\mathbb{R}^n} \frac{1}{2}\|b - Ax\|_2^2. \tag{27}$$

It is well-known that $x \in \mathbb{R}^n$ is a solution of problem (27) if and only if it is a solution of the Normal Equation System (NES), $A^TAx = A^Tb$. The NES is always a consistent system since the vector A^Tb belongs to the column space of A^T. For the NES, the matrix A^TA is symmetric, and if A is full rank, A^TA is a positive definite matrix and as a consequence A^TA is invertible and the NES has a unique solution. On the contrary, if A is rank deficient, the matrix A^TA is singular, and then the NES has infinite solutions. The proposed numerical technique for solving LLSP is based on the NES. First, observe that a solution of LLSP can be written as:

$$\begin{cases} \text{Find } y^* \text{ such that } A^Ty^* = \tilde{b} \\ \text{such that:} \quad y^* \in R(A) \end{cases}, \tag{28}$$

where $\tilde{b} = A^Tb$, and the solution set can be written as:

$$SPCM_{A^T}^{\tilde{b}} = \{y : y \in R(A) \text{ y } A^Ty = \tilde{b}\}. \tag{29}$$

Some properties of this set can be summarized in Theorem 4.

Theorem **4.** *The set* $SPCM_{A^T}^{\tilde{b}}$ *is not empty and*

$$SPCM_{A^T}^{\tilde{b}} = A_{SCM_A^b},$$

where

$$\begin{aligned} A_{SCM_A^b} &= \{y = Ax : x \in SCM_A^b\}, \ \ \textit{and} \\ SCM_A^b &= \{x \in \mathbb{R}^n : x = arg \min_{x\in\mathbb{R}^n} \frac{1}{2}\|b - Ax\|_2^2\}. \end{aligned} \tag{30}$$

Proof. Problem (28) is always feasible, since the NES is consistent for all matrices A and for all vectors b. Then, there exits $x^* \in \mathbb{R}^n$ such that,

$$A^TAx^* = A^Tb = \widetilde{b}.$$

So, it is sufficient to take $y^* = Ax^*$, since it satisfies the NES and the restriction of problem (28). Therefore, $y^* \in SPCM_{A^T}^{\tilde{b}}$ and as a consequence $SPCM_{A^T}^{\tilde{b}} \neq \emptyset$.
Now, in order to prove that $SPCM_{A^T}^{\tilde{b}} \subset A_{SCM_A^b}$, let $w \in \mathbb{R}^m$, such that $w \in SPCM_{A^T}^{\tilde{b}}$, then, $A^Tw = \widetilde{b}$, $w = At$, with $t \in \mathbb{R}^n$, and $A^TAt = A^Tb = \widetilde{b}$. So, $A^Tw = A^TAt = \widetilde{b}$, which implies that $t \in SCM_A^b$. So, $w = At \in A_{SCM_A^b}$, and it follows that $SPCM_{A^T}^{\tilde{b}} \subseteq A_{SCM_A^b}$.
To prove that $SPCM_{A^T}^{\tilde{b}} \supset A_{SCM_A^b}$, consider $z \in A_{SCM_A^b}$, that is, $z \in R(A)$, $z = Ax$ with $x \in SCM_A^b$. Therefore, $Az = A^TAx = A^T\widetilde{b}$, implying that $z \in SPCM_{A^T}^{\tilde{b}}$. It follows that $A_{SCM_A^b} \subseteq SPCM_{A^T}^{\tilde{b}}$. □

Moreover, the following theorem states an important property of the set $A_{SCM_A^b}$.

Theorem **5.** *The set* $A_{SCM_A^b} = SPCM_{A^T}^{\tilde{b}}$ *contains a unique element.*

Proof. If the LLSP (27) has a unique solution, that vector will be the unique element of the set SCM_A^b, and as a consequence, $A_{SCM_A^b}$ contains a unique element. On the contrary, if there exist infinite solutions for LLSP, let x_1 and x_2 be any two of them. Hence,

$$A^T A x_1 = A^T b$$

and

$$A^T A x_2 = A^T b$$

and thus,

$$A^T (Ax_1 - Ax_2) = 0,$$

which indicates that,

$$v = Ax_1 - Ax_2 \in N(A^T) = R(A)^{\perp},$$

i.e., $v \perp s$, $\forall s \in R(A)$, but also $v = Ax_1 - Ax_2 = A(x_1 - x_2) \in R(A)$, it follows that $v^T v = \|v\|_2^2 = 0$, implying that v is the null vector. Therefore, $\forall x_1, x_2 \in SCM_A^b$, $Ax_1 = Ax_2$.

This result establishes that even when the set SCM_A^b has infinite elements, the set $A_{SCM_A^b}$ contains only one element, and therefore $SPCM_{A^T}^{\widetilde{b}}$ contains only a unique vector but this vector can be written in many ways as a linear combination of the columns of the matrix A. □

In view of Theorems 4 and 5, we propose an initial strategy for solving LLSP (27), which we name **OPALS-PIV-LLS-v1**. It consists in solving problem (28) to obtain vector y^*, and then compute x^* as the solution of the linear system of equations $Ax = y^*$, which will be the solution of LLSP (27). For the first step of this strategy, problem (28) is solved by using OPALS-PIV, described in Section 2.1, considering $C = A^T$, $d = \widetilde{b}$, and an initial iterate in the space $R(C^T) = R(A)$. For the second step, the strategy OPALS, explained in Section 2 is used for solving $Ax = y^*$. Observe that the proposed scheme avoids the formation of the matrix $A^T A$, as well as the introduction of numerical problems due to an increase in the condition number of the matrix $A^T A$. Then, the computational cost of the proposed procedure is equivalent to the cost of the strategy OPALS for solving two linear system of equations. However, we now present an optimized or simplified version of this procedure that takes in advantage the gradient expression of the objective function as a search direction in the considered gradient-type method, avoiding the second step in the OPALS-PIV-LLS-v1 (initial version). The key idea, in this simplified version consists on observing that when solving problem (28) using OPALS-PIV, the global gradient-type method generates a sequence of iterates of the following form:

$$y_k = A \left(\sum_{i=1}^{k} (-\alpha_{i-1}) E(y_{i-1}) + p_0 \right), \tag{31}$$

where $y_0 = Ap_0$, $p_0 \in \mathbb{R}^n$. Once the stopping criterium is satisfied at the kth iteration, $y_k = Ap_k$ with

$$p_k = \sum_{i=1}^{k} (-\alpha_{i-1}) E(y_{i-1}) + p_0. \tag{32}$$

Therefore, $x = p_k$ is an approximation of the solution of problem (1). The cost of this new simplified version of OPALS-PIV-LLS scheme lies basically in the calculation of the vector $E(y_i)$, with $i = 1, 2, \cdots, k$. However, the computation of each of these vectors is done inside the global gradient-type method when computing the gradient vector, that defines the search direction of the optimization scheme. So, no extra computational cost is required for computing vectors $E(y_i)$, with $i = 1, 2, \cdots, k$. This fact is an important advantage, since in addition to avoiding the solution of the linear system of step 2, only the sum of a previously calculated vector per iteration is required. Hence, the proposed simplified version consists in applying OPALS-PIV scheme for $C = A^T$, $d = \widetilde{b}$ and, starting from any initial iterate in the column space of the matrix $C^T = A$. Then, the solution of the LLSP (1) is computed iteratively in the gradient-type method, which correspond to the accumulated vector (32). This simplified version for solving LLSP will be denoted as **OPALS-PIV-LLSP**. Observe that if a solution of problem (1) from a given initial iterate x_0 is required, OPALS-PIV scheme must be used with an initial vector $y_0 = Ax_0$.

Algorithm 2 OPALS-PIV-LLS: for finding a solution of a linear least-squares problem.

Input: A, b, x_0, α_0, and all the parameters required by the global gradient-type method
Exit: $x_k = p_k$ is the approximate solution to LLSP.
$\widetilde{b} \leftarrow A^T b$.
Apply OPAPLS-PIV until convergence, with $y_0 \in R(A)$, $C = A^T$, $d = \widetilde{b}$ to obtain $y_k = Ap_k$.
Compute p_k as in (32).

3.3 OPALS-PIV for the solution of MNLLSP.

In this section we propose a method for computing the linear least-squares solution of minimum norm. Given $A \in \mathbb{R}^{m\times n}$ y $b \in \mathbb{R}^m$, the MNLLSP can be written as,

$$\begin{cases} \text{Find } x^* \text{ such that } x^* = arg\min_{x\in\mathbb{R}^n} \frac{1}{2}\|x\|_2^2 \\ \text{subject to} \quad x \in SCM_A^b \end{cases} \tag{33}$$

where SCM_A^b is defined as in (30).

Problem (33) is equivalent to,

$$\begin{cases} \text{Find } x^* \text{ such that } x^* = arg\min_{x\in\mathbb{R}^n} \frac{1}{2}\|x\|_2^2 \\ \text{subject to} \quad Cx = \widetilde{b} \end{cases} \tag{34}$$

where $C = A^T A$ y $\widetilde{b} = A^T b$. It is clear that this last problem correspond to find the solution of minimum norm of a consistent linear system, which have been studied in Section 3.1. In view of the KKT conditions (17) and (18), the solution of Problem (34) must satisfy,

$$Cx^* = \widetilde{b}, \tag{35}$$

$$x^* = C^T \mu^*. \tag{36}$$

where $\mu^* \in \mathbb{R}^n$. However, the computation of the coefficient matrix C must be avoided since it requires additional computational cost and also could introduces some numerical problems. Therefore, we propose an alternative scheme, denoted by OPALS-PIV-MNLLS (OPALS-PIV Linear Least-Squares with Minimum Norm), which consists in solving consecutively the following problems,

$$\begin{cases} \text{Find } z \text{ such that} \quad A^T z = \tilde{b} \\ \text{suject to} \quad z \in R(A) \end{cases} \tag{37}$$

and

$$\begin{cases} \text{Find } x_{MN} \text{ such that } x_{MN} = arg\min_{x\in\mathbb{R}^n} \frac{1}{2}\|x\|_2^2 \\ \text{subject to} \quad Ax = z. \end{cases} \tag{38}$$

Problem (37) is equivalent to Problem (28). In Section 2.1 we presented the OPALS-PIV scheme for obtaining a solution z of this problem. Once z is obtained, Problem (38) is solved using OPALS-PIV-MNLS. Recall that as a consequence of Theorem 4, $SPCM_{A^T}^{\widetilde{b}} = A_{SCM_A^b}$, then the solution of Problem (38) is in the set SCM_A^b. It is also important to mention that when solving Problem (37), a solution of the LLSP is obtained, say x_a, but this solution is not necessarily the minimum norm solution, it is just one of the solutions of the problem.

Therefore, the proposed scheme for solving the MNLLSP (33), named OPALS-PIV-MNLLS, is given in the algorithm 3.

Algorithm 3 OPALS-PIV-MNLLS: for determining the minimum norm solution of a linear least-squares problem.

Input: A, b, x_0, α_0, and all the parameters required for the global gradient-type method
Exit: x_{MN} is approximate solution.

$\widetilde{b} \leftarrow A^T b$.

Apply OPALS-PIV-LLS until convergence, with $z_0 \in R(A)$, $C = A^T$, $d = \widetilde{b}$, to obtain z_k, where $z_k = Ap_k$ and p_k is one of the solutions of the LLSP.

Apply OPAPLS-PIV until convergence, with $x_0 \in R(A^T)$, $C = A$, $d = z_k$, to obtain x_{MN}.

Observe that if the solution of the LLSP is unique, then the set SCM_A^b contains a unique vector, and this vector is the solution of Problem (33). Moreover, the computational cost of OPALS-PIV-MNLLS scheme is equal to the cost of applying OPALS to solve two linear systems.

4 Numerical Results

The four proposed strategies for solving MNLSP, WMNLSP, LLSP, and MNLLSP are base on the use of the OPALS scheme, recently published by Cores and Figueroa [7] for solving linear systems of equations. The key point of the algorithms proposed is the form of the gradient of objective function presented in [7] which makes the gradient search direction of any global gradient-type method a simple and adequate direction for solving these unconstrained and constrained optimization problems. To illustrate the behavior of these new proposed strategies, some test problems are considered. In some of the problems, there exists an explicit exact solution, so it is possible to measure the relative error of the approximation obtained. The results presented in this section were obtained using the Global Spectral Gradient (GSG) method [16], since the step-size is very simple to compute and produces a fast low-cost iterative method. Some recent variations, associated with the use of spectral step lengths (e.g., [8,9,18]) were also considered for our experiments, but we did not observe an improvement as compared to the use of the step length used in the GSG method. In the following section, the GSG method is presented as a particular simple case of the available software for the Spectral Projected Gradient (SPG) method [4], for which the convex set is the whole space.

4.1 The SPG Method

The OPALS scheme can be applied to solve different problems using any gradient-type method. In particular, we consider the Spectral Projected Gradient (SPG) method [4, 16], which is nowadays a well-established nonmonotone numerical scheme for solving large-scale unconstrained or large scale convex constrained optimization problems. In the case of solving convex constrained problems, the projection onto the feasible set can be performed efficiently [4]. The SPG method has been extended to some other constrained optimization settings; see, e.g., [5]. The attractiveness of the SPG method is mainly its simplicity and effectiveness. Moreover, it is globally convergent, i.e., the iterative sequence that it generates converges to stationary points from any initial guess. For more details on the convergence properties of the SPG method see [3, 4, 16].

We now discuss the most important features of the SPG method for solving nonlinear optimization problems of the form:

$$\min_{x \in \Omega} f(x), \tag{39}$$

where Ω is a closed convex set in R^n and $f : R^n \longrightarrow R$ is a function with continuous partial derivatives in an open set that contains Ω. Starting from a given initial $x_0 \in R^n$, the iterations are given by:

$$x_{k+1} = x_k - \alpha_k d_k, \tag{40}$$

where $d_k = P_\Omega(x_k - \lambda_k g_k) - x_k$, $g_k = \bigtriangledown f(x_k)$, P_Ω denotes the projection onto Ω, and λ_k is the spectral choice of step length, given by:

$$\lambda_k = \frac{s_{k-1}^T s_{k-1}}{s_{k-1}^T y_{k-1}}, \tag{41}$$

where $s_{k-1} = x_k - x_{k-1}$ and $y_{k-1} = g_k - g_{k-1}$.

The parameter $\alpha_k > 0$ in (40) is found by a nonmonotone line search to accomplish the following condition:

$$f(x_{k+1}) \leq \max_{0\leq j\leq \min\{k,M-1\}} f(x_{k-j}) + \gamma\alpha_k g_k^T d_k, \tag{42}$$

where M is a nonnegative integer and γ is a small positive number. In this work, we set $M = 10$ and $\gamma = 10^{-4}$. For more details concerning practical issues see [5]. Notice that the SPG method can also be applied to the unconstrained minimization problem (7) by setting $\Omega = R^n$, and in this case the method is known as Global Spectral Gradient (GSG), which is the one used in all the algorithms proposed.

In general, in this work the maximum number of iterations allowed to the GSG method is $miter =$ 20000, and in all the experiments, unless otherwise stated, the value of M for the non-monotone line search is $M = 10$, and $\gamma = 10^{-4}$. Moreover, all the experiments were run in a PC i7 4710MQ with 2.50GHz using Matlab R2013b.

4.2 Results for MNLSP and WMNLSP.

All the experiments in this section focus on finding, between all the solutions of a linear system, the minimum norm solution or the weighted minimum norm solution, using an extended version of OPALS scheme, which was proposed in Section 3.1. First, the proposed schemes OPALS-PIV-MNLS and OPALS-PIV-WNMLS require an initial vector in the column space of A^T. The approximate solution of problem (2) and problem (3) is denoted by x_a, and it was obtained using the GSG method, following the iterative scheme discussed in [4]. The stopping criterion is given by $\frac{\|b-Ax_k\|_2}{\|b\|_2} < 10^{-10}$. The results for both problems are summarized in the first six (6) tables. In those tables, the number of iterations for attaining convergence is denoted by $iter$. Each table has 9 columns with different parameters: m, n, x_0, $iter$, $fcnt$, $tcpu$, $\frac{\|b-Ax_a\|_2}{\|b\|_2}$ $\frac{\|x^*-x_a\|_2}{\|x^*\|_2}$, and $\|x_a\|_2$, where x^* is the exact solution. The first two columns correspond to the row and column dimension of the matrix A respectively; the second column shows the initial iterate used; the column denoted by $fcnt$ represents the number of function evaluations required by the GSG method inside the OPALS-PIV strategy until convergence; column $tcpu$, indicates the CPU, in seconds, required by the method to reach convergence; the last three columns show the value of the stopping criterium at the obtained solution (x_a), the relative error of the approximation, and the Euclidean norm of the approximate solution respectively. In the proposed strategy for solving WMNLS problems, the number of gradient evaluations, denoted by $gcnt$, is always $gcnt = iter + 1$. For that reason, in the tables of this section there is no column with the parameter $gcnt$. Moreover, the symbol ** in any column of a table means that the method did not reach the solution with the parameters used.

4.2.1 Minimum Norm Linear System Results.

For the sake of simplicity, the initial vector is the zero vector of dimension n, denote by ν_0. Also other initial vectors are consider for example: $\nu_i = \frac{A^T v_i}{\|A^T v_i\|_2}$, $i = 1, 2, 3$, where v_1 is an $m-$vector with all the entries equal to 1, v_2 is a random vector with components between zero(0) and one(1), and v_3 is a random vector with all entries in the interval $[-10, 10]$. Here we consider two type of linear systems.

First group of experiments: consider the matrix of the linear system, $A \in \mathbb{R}^{m\times n}$, $m \geq n$, where the first column of A is the vector of dimension m with all the entries equal to one (1), the following $n-2$ columns correspond to the first $n-2$ vectors of the ordered canonic base of $\mathbb{R}^m$, and the last column has the first $n-2$ entries equal zero (0) and the value of one (1) in the last two positions. The linear system $Ax = b$ with this kind of matrix will be consistent if $b_{n-1} = b_n = b_{n+1} = \cdots = b_{m-1} = b_m = \beta$, $\beta \in \mathbb{R}$ and, in this case, the system has infinite solutions of the following form:

$$(t, b_1 - t, b_2 - t, \cdots, b_{n-1} - t)^T \quad \text{with} \quad t \in \mathbb{R}$$

Considering $\beta = 5$ and $b_1 = b_2 = \cdots = b_{n-2} = 1$, the exact solution of the system is given by:

$$(t, 1-t, 1-t, \cdots, 5-t)^T, \text{with } t \in \mathbb{R}.$$

Then, the minimum norm solution is reached for $t = \frac{2n+6}{2n} = 1 + \frac{3}{n}$. Therefore,

$$x^* = \left(1 + \frac{3}{n}, \underbrace{\frac{-3}{n}, \cdots, \frac{-3}{n}}_{n-2 \text{ times}}, 4 - \frac{3}{n}\right)^T.$$

Second group of experiments: we consider the linear system $Bx = b$, where $B = A^T$ and the A is the matrix considered in the first experiment. In this case the right hand side vector is given by, $b_1 = n$, $b_i = 1$, $i = 2, \cdots, m-1$, $b_m = n - (m-2) = n - m + 2$. Here $n > m$, and the linear system has infinite solutions. Moreover, the exact minimum norm solution vector is given by :

$$x^* = \left(b_2, \cdots, b_{m-1}, b_m, \underbrace{0, \cdots, 0}_{n-m+1 \text{ times}}\right)^T$$

Table 1 and Table 2 show the results obtained for this first group of experiments and for the second group of experiments, respectively, when using OPAPLS-PIV-MNLS scheme with GSG method.

The complexity and computational cost of the proposed OPAPLS-VIP-MNLS strategy is the same as the one for OPALS. Additionally, in this case we are interested in the behavior of the proposed techniques for small or medium size scale problems.

4.2.2 Weighted Minimum Norm Linear System Results.

As before, since we pretend to study the performance of the proposed OPALS-VIP-WMNLS method for solving the WNMSP, the initial vector must belong to the image space of the matrix $\Pi^T A^T$. For that, and for the sake of simplicity, the initial vector is the null vector of dimension n. The stopping criterium considered is $\frac{\|b - A\Pi x_k\|_2}{\|b\|_2} < 10^{-10}$. We also consider squared weighted matrices P_1, P_2 of dimension n. In particular, P_1 is a $n \times n$ diagonal matrix that satisfies $P_1 = (p_{ij})$, $1 \leq i, j \leq n$, with $p_{ii} = \|A^i\|_2$, where A^i represents the i-th column of A, the matrix associated to the linear constraint $Ax = b$. On the other hand, P_2 is a tridiagonal matrix whose diagonal elements are all equal to the value 8, the sub-diagonal and sup-diagonal elements are all 2, which is obtained from the Matlab command $P_2 = gallery('tridiag', n, 2, 8, 2)$. In the case of the matrix P_1 the factor Π is the diagonal matrix which diagonal elements are $\Pi_{ii} = \sqrt{\|A^i\|_2}$, $i = 1, 2, \cdots, n$. For the matrix P_2, the factor Pi is obtained by performing the Cholesky factorization, using the Matlab command chol. It is important to mention that the product $A\Pi$ is computed once

m	n	x_0	$Iter$	$fcnt$	$tcpu$	$\frac{\|b-Ax_a\|_2}{\|b\|_2}$	$\frac{\|x^*-x_a\|_2}{\|x^*\|_2}$	$\|x_a\|_2$
100	30	ν_0	13	14	1.5600e-02	3.8285e-15	3.2318e-15	4.0865e+00
		ν_1	13	14	0.0000e+00	1.3936e-14	3.9672e-14	4.0865e+00
		ν_2	32	33	0.0000e+00	1.7075e-13	4.8537e-13	4.0865e+00
		ν_3	23	24	1.5600e-02	6.2808e-14	1.7854e-13	4.0865e+00
1000	300	ν_0	13	14	1.5600e-02	1.7421e-15	1.5992e-15	4.1194e+00
		ν_1	13	14	0.0000e+00	8.1713e-14	2.2882e-13	4.1194e+00
		ν_2	33	34	1.5600e-02	5.8340e-14	4.8350e-14	4.1194e+00
		ν_3	56	57	3.1200e-02	1.2675e-13	4.0625e-13	4.1194e+00
3000	1000	ν_0	13	14	3.1200e-02	1.2662e-16	6.7013e-16	4.1220e+00
		ν_1	15	16	1.5600e-02	1.9508e-15	4.4144e-15	4.1220e+00
		ν_2	23	24	3.1200e-02	2.9737e-13	1.6300e-11	4.1220e+00
		ν_3	54	55	6.2400e-02	5.9370e-13	3.1569e-12	4.1220e+00
5000	1500	ν_0	13	14	3.1200e-02	1.6388e-15	1.6584e-15	4.1223e+00
		ν_1	13	14	3.1200e-02	8.8180e-14	2.4662e-13	4.1223e+00
		ν_2	29	30	6.2400e-02	2.9120e-13	1.9733e-11	4.1223e+00
		ν_3	38	39	7.8000e-02	5.5246e-13	3.9991e-11	4.1223e+00
8000	2500	ν_0	18	28	0.13037010	0.0000000e+00	1.8883635e-16	4.1226690e+00
		ν_1	12	22	0.10253520	1.7666766e-16	2.2206812e-16	4.1226690e+00
		ν_2	31	41	0.18726040	5.0752295e-16	4.6170566e-14	4.1226690e+00
		ν_3	43	53	0.24984440	1.3248078e-16	1.2006141e-14	4.1226690e+00
10000	3000	ν_0	19	29	0.2306375	2.8812934e-17	1.4737382e-15	4.1227418e+00
		ν_1	18	28	0.13679430	0.0000000e+00	9.5509405e-16	4.1227418e+00
		ν_2	61	71	0.44585220	5.6717146e-16	5.8068248e-14	4.1227418e+00
		ν_3	25	36	0.20306250	1.7900408e-14	1.8320861e-12	4.1227418e+00
15000	5000	ν_0	12	23	0.17965260	5.3574951e-16	2.3891583e-15	4.1228873e+00
		ν_1	13	24	0.18524900	1.9228797e-16	3.9134374e-15	4.1228873e+00
		ν_2	51	62	0.57203960	3.2555809e-16	4.9711316e-15	4.1228873e+00
		ν_3	34	45	0.37404510	9.1148348e-16	1.1165267e-13	4.1228873e+00
20000	6000	ν_0	19	30	0.32032230	0.0000000e+00	1.2382598e-15	4.1229237e+00
		ν_1	18	29	0.29111350	2.8819682e-17	4.0682958e-16	4.1229237e+00
		ν_2	42	53	0.79564470	2.6948711e-17	5.5079727e-15	4.1229237e+00
		ν_3	76	88	1.3086103	2.1192132e-16	3.0624444e-14	4.1229237e+00
25000	10000	ν_0	13	25	0.28849740	7.1559838e-17	2.6863684e-15	4.1229965e+00
		ν_1	18	29	0.34551170	1.7892822e-16	4.6775982e-15	4.1229965e+00
		ν_2	35	46	0.63221900	4.2315484e-16	6.3215816e-14	4.1229965e+00
		ν_3	29	41	0.54565180	4.7420223e-17	5.1921688e-15	4.1229965e+00

Table 1. Results for the MNLSP with $m > n$, different initial vectors where A and b correspond to the first group of experiments.

before starting the iterative method. The test problems used for studding the performance of OPALS-PIV-WMNLS scheme consists of the same linear systems considered for solving the MNLSP. The obtained results are in Tables from 3, 4, 5, and 6. These tables show the values of same parameters used in the previous tables, except the initial iterate since only one initial vector was considered. Instead of showing the values of $\|x_a\|_2$, the weighted norm of the approximate solution, $\|x_a\|_W$, is illustrated. Moreover, when matrix P_1 is used, the exact solution of the WMNLSP can be explicitly computed, and in this case the tables also show the relative error of the approximation, $\frac{\|x^*-x_a\|_2}{\|x^*\|_2}$.

4.3 Linear Least-Squares Results.

Recall that OPALS-PIV-LLS strategy, for solving LLSP does not require the formation of the NES matrix. Moreover, it only requires for evaluating the objective function, at each iteration, the product of the matrix A with a vector, and also the product of the matrix A^T with a vector in the gradient

m	n	x_0	$Iter$	$fcnt$	$tcpu$	$\frac{\|b-Ax_a\|_2}{\|b\|_2}$	$\frac{\|x^*-x_a\|_2}{\|x^*\|_2}$	$\|x_a\|_2$
30	100	ν_0	13	14	0.0000e+00	3.1510e-14	1.0638e-13	1.0000e+01
		ν_1	12	13	1.5600e-02	9.0605e-16	1.1766e-15	1.0000e+01
		ν_2	35	36	1.5600e-02	3.8990e-13	3.8046e-13	1.0000e+01
		ν_3	34	35	0.0000e+00	3.5728e-14	4.9354e-14	1.0000e+01
300	1000	ν_0	13	14	1.5600e-02	2.6388e-14	3.3516e-14	3.1622e+01
		ν_1	12	13	1.5600e-02	1.3204e-13	4.1183e-13	3.1622e+01
		ν_2	33	34	1.5600e-02	7.5853e-16	1.7796e-14	3.1622e+01
		ν_3	31	32	1.5600e-02	8.0097e-13	7.7898e-13	3.1622e+01
1000	3000	ν_0	13	14	1.5600e-02	6.0888e-14	6.2716e-14	5.4772e+01
		ν_1	13	14	1.5600e-02	1.2276e-13	3.7709e-13	5.4772e+01
		ν_2	32	33	3.1200e-02	6.7038e-13	2.0939e-12	5.4772e+01
		ν_3	44	45	6.2400e-02	3.0770e-13	4.6845e-13	5.4772e+01
1500	5000	ν_0	13	14	1.5600e-02	8.0486e-14	1.2832e-13	7.0710e+01
		ν_1	13	14	3.1200e-02	5.1082e-13	1.5339e-12	7.0710e+01
		ν_2	309	358	5.3040e-01	8.6974e-13	7.4983e-11	7.0710e+01
		ν_3	45	46	9.3600e-02	6.2134e-14	1.2380e-13	7.0710e+01
2500	8000	ν_0	11	12	1.2152640e-01	9.7493427e-13	2.1221113e-12	8.9442719e+01
		ν_1	11	12	4.5751800e-02	2.5717382e-12	3.2531053e-12	8.9442719e+01
		ν_2	27	28	8.4170200e-02	1.1591306e-12	1.2537761e-10	8.9442719e+01
		ν_3	25	26	8.9099000e-02	7.6678548e-10	8.3238657e-08	8.9442719e+01
3000	10000	ν_0	11	12	5.6505300e-02	6.1549382e-13	1.1579992e-12	1.0000000e+02
		ν_1	11	12	4.2244200e-02	8.6839444e-13	2.5419991e-12	1.0000000e+02
		ν_2	38	39	1.4767350e-01	6.9227202e-11	3.9807391e-10	1.0000000e+02
		ν_3	27	28	1.1274530e-01	1.9574944e-12	2.3890390e-10	1.0000000e+02
5000	15000	ν_0	11	12	7.0253400e-02	1.5742247e-13	3.7348587e-13	1.2247449e+02
		ν_1	11	12	6.5067900e-02	4.5257548e-12	5.4820649e-12	1.2247449e+02
		ν_2	48	49	2.8185110e-01	5.0235387e-10	7.3661801e-08	1.2247449e+02
		ν_3	43	44	2.5246970e-01	5.1915206e-10	7.6381172e-08	1.2247449e+02
5000	15000	ν_0	11	12	1.0047330e-01	8.3716680e-12	1.3610633e-11	1.4142136e+02
		ν_1	11	12	9.0094300e-02	5.9951532e-12	8.3089889e-12	1.4142136e+02
		ν_2	49	50	3.9964650e-01	6.3541482e-11	7.7935858e-09	1.4142136e+02
		ν_3	27	28	2.1690910e-01	7.8896212e-10	1.3620306e-07	1.4142136e+02
10000	25000	ν_0	11	12	1.3616540e-01	1.1532595e-12	2.8543158e-12	1.5811388e+02
		ν_1	11	12	1.0922550e-01	1.5669467e-12	3.3109246e-12	1.5811388e+02
		ν_2	36	37	3.5230340e-01	1.6939842e-10	3.1094572e-08	1.5811388e+02
		ν_3	28	29	2.9100510e-01	2.3754857e-09	4.3785518e-07	1.5811388e+02

Table 2. Results for the MNLSP with $m < n$, different initial vectors, where A and b belong to second group of experiments.

m	n	$Iter$	$fcnt$	$tcpu$	$\frac{\|b-Ax_a\|_2}{\|b\|_2}$	$\frac{\|x^*-x_a\|_2}{\|x^*\|_2}$	$\|x_a\|_W$
100	30	15	16	4.0000e-02	7.2023e-13	2.0885e-12	1.4514e+02
1000	300	12	13	5.0000e-02	1.6858e-15	1.4539e-15	4.5548e+02
3000	1000	13	14	1.3000e-01	2.4483e-14	6.4208e-14	7.7065e+02
5000	3000	13	14	5.0000e-02	3.3377e-14	9.3380e-14	1.0175e+03
8000	2500	18	32	1.0746980e-01	2.9655004e-17	5.1489505e-16	4.7632379e-01
10000	3000	20	34	1.5889860e-01	1.9620088e-16	6.4962254e-16	4.4856302e-01
15000	5000	18	33	2.2729600e-01	2.3471345e-16	8.0936254e-16	4.1005947e-01
20000	6000	18	33	2.9848580e-01	1.7613310e-16	1.0545313e-15	3.7720813e-01
25000	10000	32	64	6.6113930e-01	1.0733976e-16	2.1206982e-15	3.7007463e-01

Table 3. Results WMNLSP with matrix P_1, $m > n$, first group of experiments.

function evaluation. Therefore, a perfect candidate method for comparing the results of the proposed

m	n	$Iter$	$fcnt$	$tcpu$	$\frac{\|b-Ax^*\|_2}{\|b\|_2}$	$\frac{\|x^*-x_a\|_2}{\|x^*\|_2}$	$\|x_a\|_W$
30	100	12	13	3.0000e-02	1.3958e-12	1.3637e-12	1.4142e+02
300	1000	12	13	5.0000e-02	2.6727e-12	2.6010e-12	1.4142e+03
1000	3000	12	13	4.0000e-02	1.0863e-12	1.1076e-12	4.2426e+03
1500	5000	12	13	1.5000e-01	2.8182e-12	2.7484e-12	7.0710e+03
2500	8000	11	12	5.9731200e-02	1.1430515e-13	2.5465293e-13	7.5212062e+01
3000	10000	11	12	5.4017400e-02	4.3395913e-12	7.1003252e-12	8.4089642e+01
5000	15000	11	12	7.8048000e-02	3.8341013e-12	5.7787006e-12	1.0298836e+02
6000	20000	11	12	1.0484260e-01	1.0016320e-12	1.8466689e-12	1.1892071e+02
10000	25000	11	12	2.2365690e-01	6.2811630e-12	1.0825289e-11	1.3295740e+02

Table 4. Results for the WMNLSP with P_1, $m < n$, second group of experiments.

m	n	$Iter$	$fcnt$	$tcpu$	$\frac{\|b-Ax_a\|_2}{\|b\|_2}$	$\|x_a\|_W$
100	30	78	79	5.0000e-02	7.1883e-11	1.3267e+02
1000	300	142	143	1.4000e-01	9.5767e-11	1.3568e+02
3000	1000	132	133	3.5000e-01	5.2619e-11	1.3590e+02
5000	1500	100	101	4.1000e-01	7.0083e-11	1.3593e+02
8000	2500	474	539	2.4486937e+00	6.9490385e-15	1.5090537e+00
10000	3000	235	249	1.7771266e+00	5.6633358e-15	1.5090716e+00
15000	5000	618	637	6.4605905e+00	4.9487146e-15	1.5091075e+00
20000	6000	364	380	5.1702983e+00	4.6487864e-15	1.5091165e+00
25000	10000	240	259	4.3030063e+00	7.7781882e-16	1.5091345e+00

Table 5. Results for the WMNLSP with P_2, $m > n$, first group of experiments.

m	n	$Iter$	$fcnt$	$tcpu$	$\frac{\|b-Ax_a\|_2}{\|b\|_2}$	$\|x_a\|_W$
30	100	61	62	2.0000e-02	6.0106e-11	1.1968e+03
300	1000	102	103	8.0000e-02	8.9766e-11	1.1996e+04
1000	300	90	91	1.7000e-01	5.4585e-11	3.5996e+04
1500	5000	175	176	4.6000e-01	6.5365e-11	5.9996e+04
2500	8000	115	116	3.7820560e-01	1.0068954e-09	2.5820499e+01
3000	1000	297	298	1.3644234e+00	1.2428650e-09	2.8868059e+01
5000	15000	108	109	6.9112950e-01	1.9313049e-10	3.5355785e+01
6000	20000	831	841	7.0320280e-01	1.2110297e-09	4.0825215e+01
10000	25000	199	200	0.022844922e+02	0.2.9415e-09	4.5643891e+01

Table 6. Results for the WMNLSP with P_2, $m < n$, second group of experiments.

scheme is the Conjugate Gradient method applied to the NES (Conjugate Gradient Normal Residual, CGNR), which is a well-known iterative scheme in the literature; see, e.g., [17]. In this section,we present the results for solving LLSP, $\min_{x\in R^n} \frac{1}{2}\|Ax-b\|_2^2$, where the coefficient matrix is generated in Matlab with the command $A = sprand(m, n, \rho, \frac{1}{\kappa})$. This command allows to obtain a matrix A of size $m \times n$ with approximately $m \cdot n \cdot \rho$ non zero entries uniformly distributed in the interval $[0, 1]$, with condition number $\|A\|\|A^\dagger\| = \kappa$. Here, the right hand side vector is $b = (1, 1, \cdots, 1)^T$. In order to make a fair comparison between the results obtained by CGRN method and by OPALS-PIV-LLS scheme, the initial starting point and the stopping criterium for both methods are the same. As before, the initial iterate is the null vector, since it is in the space $R(A)$. Table 7 shows the label given for each matrix A considered and also the corresponding values for the parameters m, n, ρ y κ for each matrix. The obtained results are shown in the Tables 8 and 9. The first column of these tables indicates the label of the matrix used, the following two columns correspond to the results obtained with CGNR, the fourth and fifth columns corresponds to the results for the proposed OPALS-PIV-LLS scheme, and the last column represents the relative error of the solution obatined with OPALS-PIV-LLS with respect to the solution obtained by

CGNR, $\frac{\|x_{CG}-x_a\|_2}{\|x_{CG}\|_2}$, where x_{CG} is the CGNR estimate and x_a is the OPALS-PIV-LLS approximation. In these tables only the number of iterations, denoted by iter, and the norm of the residual for each method, denoted by $\|r_{CG}\|_2$ and $\|r_a\|_2$ are shown. The difference between the results presented in Table 8 and Table 9 is the size of the matrices considered, $10^4 \times 10^3$, and $3 \cdot 10^4 \times 3 \cdot 10^3$ respectively.

		$A = gallery('sprand', m, n, \rho, \kappa)$									
		$\rho = 10^{-2}$									
		κ									
m	n	10.41	17.59	19.59	22.24	23.04	23.09	26.36	33.38	33.74	45.98
10^4	10^3	ls1	ls2	ls3	ls4	ls5	ls6	ls7	ls8	ls9	ls10
$3 \cdot 10^4$	$3 \cdot 10^3$	ls21	ls22	ls23	ls24	ls25	ls26	ls27	ls28	ls29	ls30
10^5	10^4	ls41	ls42	ls43	ls44	ls45	ls46	ls47	ls48	ls49	ls50
		κ									
m	n	48.82	59.48	62.17	79.74	82.00	83.55	88.18	88.23	91.95	96.57
10^4	10^3	ls11	ls12	ls13	ls14	ls15	ls16	ls17	ls18	ls19	ls20
$3 \cdot 10^4$	$3 \cdot 10^3$	ls31	ls32	ls33	ls34	ls35	ls36	ls37	ls38	ls39	ls40
10^5	10^4	ls51	ls52	ls53	ls54	ls55	ls56	ls57	ls58	ls59	ls60

Table 7. Description of matrices with $\rho \cdot m \cdot n$ non null elements, and $10 < \kappa = \frac{\sigma_{max}}{\sigma_{min}} < 100$.

	CGNR		OPAPLS-PIV-LLS		
Matrix	*iter*	$\|r_{CG}\|_2$	*Iter*	$\|r_a\|_2$	$\frac{\|x_{CG}-x_a\|_2}{\|x_{CG}\|_2}$
ls1	134	57.825977	174	57.825977	1.9827e-09
ls2	216	56.925560	281	56.925560	3.5367e-09
ls3	243	58.460664	279	58.460664	1.1040e-09
ls4	268	56.825352	355	56.825352	1.5424e-09
ls5	282	58.467847	401	58.467847	3.2136e-10
ls6	284	57.325820	363	57.325820	2.3127e-09
ls7	320	56.678822	461	56.678822	4.5401e-09
ls8	383	57.165163	432	57.165163	1.0448e-08
ls9	408	57.771934	583	57.771934	8.1123e-09
ls10	510	54.814349	809	54.814349	5.6093e-09
ls11	477	56.936138	573	56.936138	9.7758e-09
ls12	675	55.872893	772	55.872893	1.4147e-08
ls13	559	58.554064	1094	5.8554064	2.8313e-08
ls14	915	56.638526	1325	56.638526	2.0994e-08
ls15	841	56.511553	1638	56.511553	6.9053e-09
ls16	881	56.485183	1307	56.485183	9.4688e-09
ls17	976	56.879774	1803	56.879774	4.0053e-08
ls18	839	57.336381	1218	57.336381	4.5694e-08
ls19	768	56.341644	1766	56.341644	5.1352e-08
ls20	833	56.942767	1972	56.942767	5.5243e-08

Table 8. Results for CGNR and OPAPLS-PIV-LLS, for matrices $10^4 \times 10^3$ of Table 7.

From Tables 8 and 9, it is clear that CGNR requires less iterations (from 20 to 40 per cent) for reaching the solution of the Linear Least Square Problem than OPALS-PIV-LLS scheme. We also observe that both low-cost methods reach almost the same precision and both approximations agree in at least 8 correct decimal digits. So, we can say that the OPALS-PIV-LLS technique is comparable with the CGNR method.

	CGNR		OPAPLS-PIV-LLS		
Matrix	$iter$	$\|r_{CG}\|_2$	$Iter$	$\|r_a\|_2$	$\frac{\|x_{CG}-x_a\|_2}{\|x_{CG}\|_2}$
ls21	138	112.29075	151	112.29075	1.4255e-09
ls22	228	111.58973	293	111.58973	4.2813e-09
ls23	253	113.03557	323	113.03557	3.9916e-09
ls24	284	111.11409	399	111.11409	3.0192e-09
ls25	295	112.88712	382	112.88712	6.4262e-09
ls26	294	112.45437	384	112.45437	3.6093e-09
ls27	334	111.72804	425	111.72804	7.2979e-09
ls28	419	114.43380	525	114.43380	2.4048e-09
ls29	426	112.70322	600	112.70322	4.7802e-09
ls30	557	112.55676	608	112.55676	1.1774e-08
ls31	602	112.43395	841	112.43395	1.5341e-09
ls32	731	112.43007	1134	112.43007	7.0111e-10
ls33	737	114.13975	1060	114.13975	1.4778e-08
ls34	951	112.06163	1025	112.06163	1.0665e-08
ls35	974	112.84651	1481	112.84651	4.1758e-09
ls36	1019	112.23497	1168	112.23497	3.1873e-08
ls37	1040	113.23721	1592	113.23721	1.6282e-08
ls38	1034	112.23063	1475	112.23063	3.6730e-08
ls39	1084	113.08014	2124	113.08014	3.2339e-08
ls40	1150	113.84818	1302	113.84818	3.2582e-08

Table 9. Results for CGNR and OPAPLS-PIV-LLS, for matrices $3 \cdot 10^4 \times 3 \cdot 10^3$ of Table 7.

4.4 Linear Least-Squares of Minimum Norm Results

In this section we show the performance of the OPALS-PIV-LLSMN scheme by solving two small problems. This strategy requires the solution of two(2) linear systems in a consecutive way with OPALS-PIV: first solving Problem (37) and then solving Problem (38), which requires the solution of the previous one as the right hand side vector. We denote the solution of first problem by x_{LLS} and the solution of the second one by x_{LLSMN}. As we can observe in Algorithm (3), the first iterative procedure requires an initial vector in $R(A)$ and the second an initial vector in $R(A^T)$. For the sake of simplicity the initial iterate for solving the second problem is the null vector and for the first problem different initial vectors are considered. Since OPALS-PIV is used for solving two different problems, in the tables we denote by $iter_i$, $fcnt_i$ y $tcpu_i$ the number of iterations for convergence, the total number of function evaluations and the CPU time (in seconds) until convergence for procedure $i = 1, 2$. Here $i = 1$ correspond to OPALS-PIV for solving Problem (37) and $i = 2$ corresponds to OPALS-PIV for solving Problem (38). Moreover, the norm of the solutions of both problems are shown in the tables in the columns with the label $\|x_{LLS}\|_2$ and $\|x_{LLSMN}\|_2$. The stopping criterium for the first process ($i = 1$) is,

$$\frac{\|\tilde{b} - A^T z_k\|_2}{\|\tilde{b}\|_2} < tol = 10^{-10},$$

and for the second process ($i = 2$) is,

$$\frac{\|z - Ax_k\|_2}{\|z\|_2} < tol = 10^{-10}.$$

The first example is given by: $A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \\ 1 & 2 \end{pmatrix}$ y $b = \begin{pmatrix} 3 \\ 0 \\ 3 \end{pmatrix}$. In this example the coefficient matrix, A, is rank deficient, so the LLSP has infinite solutions,

$$SCM_A^b = \{(t_1, t_2)^T \in \mathbb{R}^2 : t_1 + 2t_2 = 1\},$$

and the minimum norm solution is the vector,

$$x^* = (\frac{1}{5}, \frac{2}{5})^T,$$

which Euclidean norm is $\frac{\sqrt{5}}{5}$. The results for solving this first example, starting the first procedure with vectors, $v_0 = (0,0)^T$, $v_1 = (5,2)^T$ y $v_2 = (2,0)^T$, are given in Table 10, and the solution of the first LLSP, x_{LLS}, for each initial vector is illustrated in Table 11.

x_0	$iter_1$	$fcnt_1$	$tcpu_1$	$\|x_{LLS}\|_2$	$iter_2$	$fcnt_2$	$tcpu_2$	$\|x_{LLSMN}\|_2$
v_0	14	15	0.0000e+00	4.7839e-01	5	6	0.0000e+00	4.4721e-01
v_1	135	136	0.0000e+00	5.3283e+00	5	6	0.0000e+00	4.4721e-01
v_2	14	15	0.0000e+00	2.0091e+00	5	6	0.0000e+00	4.4721e-01

Table 10. Results for first Example.

v_0	v_1	v_2
x_{LLS}	x_{LLS}	x_{LLS}
4.80547e-02	4.94900e+00	1.95194e+00
4.75972e-01	-1.97450e+00	-4.75972e-01

Table 11. Solution vector x_{LLS} obtained for the first OPALS-PIV process for different initial vectors, first example.

The second Example, consists in constructing $A = vw^T$, where v and w are vectors with all the components equal to one(1) of dimension 100 and 30, respectively. Clearly, this matrix is also rank deficient. The right hand side vector, b, has all the entries equal zero(0), except to the first one that it is equal to 100. In this example, the solution set of the LLSP is,

$$SCM_A^b = \{t \in \mathbb{R}^{30} : \sum_{i=1}^{30} t_i = 1\},$$

and the minimum norm solution corresponds to

$$x^* = \left(\underbrace{\frac{1}{30}, \frac{1}{30}, \cdots, \frac{1}{30}}_{30 \text{ times}} \right)^T,$$

whose Euclidean norm is equal $\frac{\sqrt{30}}{30}$. In this case, the first OPALS-PIV procedure starts with v_0, v_1, v_2, where v_0 is the null vector, v_1 has all the entries equal zero(0), except the first one that is equal to 0.5, and v_2 has all the entries equal zero(0), except the first one that is equal to 2. Tables 12 shows the results of Algorithm (2) for solving this second example. Table 13 illustrates the vector x_{LLS} obtained for each initial vector when the solution was obtained by the GSG method.

x_0	$iter_1$	$fcnt_1$	$tcpu_1$	$\|x_{LLS}\|$	$iter_2$	$fcnt_2$	$tcpu_2$	$\|x_{LLSMN}\|$
v_0	1	2	0.0000e+00	1.8257e-01	5	6	0.0000e+00	1.8257e-01
v_1	1	3	0.0000e+00	5.2440e-01	5	6	0.0000e+00	1.8257e-01
v_2	1	2	0.0000e+00	1.9748e+00	5	6	0.0000e+00	1.8257e-01

Table 12. Results for the second example.

	x_{LLS}
v_0	(3.3333e-02, ...,3.3333e-02)^T
v_1	(5.1666e-01,1.6666e-02,...,1.6666e-02)^T
v_2	(-3.3333e-02, ..., -3.3333e-02)^T

Table 13. Solution vector x_{LLS} obtained for the first OPALS-PIV process for different initial vectors, second example.

Observe that the proposed OPALS-PIV-LLSMN scheme could obtain the minimum norm solution of the linear least-squares problem at the first step (solving the LLSP), but there is not guarantee that this will occur. So, if the second problem uses as initial vector the one obtained as the solution of the first one, the proposed strategy finishes solving only the first problem. Observe that this will happen in the second example with the initial vector v_0, if we had used as initial guess of the second problem the solution of the first problem, and so the value of $iter_2$ would had been 0 instead of 5.

5 Conclusions

In this work we present three different numerical approaches for solving MNLSP, WMNLSP and LLSP, respectively. All these techniques are based on the recently developed non linear optimization scheme OPALS [7], for solving any linear systems of equations. These approaches could be considered as extensions of the OPALS scheme where any gradient-type method, e.g., [3,8,9,18], could be used. The strategies are based on the suitable combination of the objective function used by OPALS, the structure of its gradient, and the specialized choice of the initial vector. The choice of the initial iterate in the row space of the associated matrix guarantees that all iterates remains in the same space, converging to the solution of the considered problem. The obtained numerical results for solving MNLSP and WMNLSP show that the strategies reach the solution of all the problems requiring low CPU time and a good precision. Finally, the numerical results for solving the LLSP indicate that the proposed strategy is competitive with the well known Conjugate Gradient method for Linear least-squares problems (CGNR), reaching the solution in all the examples tested and with very good precision. For solving LLSP, we note that the proposed strategy works directly on the matrix A associated to the linear system, avoiding the cost and excessive rounding errors associated with the construction of either AA^T or A^TA. Finally, it would be interesting to extend this approach to solve linear least-squares problems with linear constraints.

References

[1] J. Barzilai and J. M. Borwein (1988). *Two point step size gradient methods.* IMA, J.Numer. Anal., vol. 8, pp. 141–148.

[2] D. Bertsekas (1999). *Nonlinear Programming.* Athena Scientific, Boston, USA.

[3] E. G. Birgin, J.M. Martínez and M. Raydan (2000). *Nonmonotone Spectral Projected Gradient methods on convexs sets.* SIAM, J. OPTIM., vol.10(4), pp. 1196–1211.

[4] E. G. Birgin, J. M. Martínez and M. Raydan (2001). *Algorithm 813: SPG - Software for convex-constrained optimization.* ACM Transactions on Mathematical Software, vol. 27, pp. 340–349.

[5] E. G. Birgin, J. M. Martínez and M. Raydan (2014). *Spectral Projected Gradient methods: Review and Perspectives.* Journal of Statistical Sofware, vol. 60(3), pp. 1–21.

[6] A. Björck (1996). *Numerical methods for least square problems.* SIAM, Philadelphia, USA.

[7] D. Cores and J. Figueroa, (2017). *A convex optimization approach for solving large scale linear systems*, Bull. Comput. Appl. Math., Vol. 5(1), pp. 51–74.

[8] Di Serafino D., Ruggiero V., Toraldo G., and Zanni L. (2018), *On the steplength selection in gradient methods for unconstrained optimization*, Appl. Math. Comput., Vol. 318, pp. 176–195.

[9] G. Frassoldati, L. Zanni and G. Zanghirati, (2008). *New adaptive stepsize selections in gradient Methods*. Journal of Industrial and Management Optimization, vol. 4(2), pp. 299–312.

[10] I. F. Gorodnitsky and B. D. Rao (1993). *A recursive weighted minimum norm algorithm : analysis and applications*, ICASSP 1993.

[11] P. C. Hansen, V. Pereya and G. Scherer (2013). *Least Squares Data Fitting with Applications*. Johns Hopkins University Press, Maryland, USA.

[12] S. Iwaki and S. Ueno (1998). *Weighted minimum-norm source estimation of magnetoencephalography utilizing the temporal information of measured data*. Journal of Applied Physics, vol. 83(11), pp. 6441–6443.

[13] W. La Cruz and M. Raydan (2008). *Residual iterative schemes for large-scale nonsymmetric positive definite linear systems*. Computacional and applied mathematics, vol. 27, pp. 151–173.

[14] R. Leahy, J. Moscher (1997). *MEG-Based imaging of focal neural cirrent sources*. IEEE transactions on medical imaging, vol. 16(3), pp. 338–348.

[15] M. Raydan (1993). *On the Barzilai and Borwein choice o steplegth for the gradient method*. IMA J. Numer. Anal., vol. 13, pp. 321–326.

[16] M. Raydan (1997). *The Barzilai and Borwein gradient method for the large-scale unconstrained minimization problem*. SIAM J. Opt., vol.7, pp. 26–33.

[17] Y. Saad (2003). *Iterative methods for sparse linear systems*. SIAM, Philadelphia, USA.

[18] B. Zhou, L. Gao and Y. H. Dai (2006). *Gradient methods with adaptive step-sizes*. Comput. Optim. Appl., vol. 35, pp. 69–86.

[19] F. Ding and H. Zhang (2014). *Gradient-based iterative algorithm for a class of the coupled matrix equations related to control systems*. IET Control Theory and Applications, vol. 15(8), pp. 1588–1595.

[20] H. Zhang (2015). *Reduced-rank gradient-based algorithms for generalized coupled Sylvester matrix equations and its applications*. Computers and Mathematics with Applications, vol. 70(8), pp. 2049–2062.

[21] H. Zhang and F. Ding (2016). *Iterative algorithms for X+ATX-1A=I by using the hierarchical identification principle*. Journal of the Franklin Institute, vol. 353(5), pp.1132–1146.

[22] H. Zhang and H. Yin (2019). *Refinements of the Hadamard and Cauchy Schwarz inequalities with two inequalities of the principal angles*. J. Math. Inequal., vol. 13, pp. 423–435.

[23] H. Zhang (2019). *Quasi gradient-based inversion-free iterative algorithm for solving a class of the nonlinear matrix equations*. Computers and Mathematics with Applications, vol. 77(5), pp. 1233–1244.

[24] H. Zhang and Wan, L. (2020). *Zeroing neural network methods for solving the Yang-Baxter-like matrix equation*. Neurocomputing, vol. 383, pp. 409–418.

[25] F. Ding and T. Chen (2005). *Gradient based iterative algorithms for solving a class of matrix equations.* IEEE Transactions on Automatic Control, vol. 50(8), pp. 1216–1221.

[26] F. Ding and T. Chen (2006). *On iterative solutions of general coupled matrix equations.* SIAM Journal on Control and Optimization, vol. 44(6), pp. 2269–2284.

[27] L. Xie, Y. Liu and H. Yang (2010). *Gradient based and least squares based iterative algorithms for matrix equations* $AXB+ CXTD= F$. Applied Mathematics and Computation, vol. 217(5), pp. 2191–2199.

[28] F. Ding, P. X. Liu and J. Ding (2008). *Iterative solutions of the generalized Sylvester matrix equations by using the hierarchical identification principle.* Applied Mathematics and Computation, vol. 197(1), pp. 41–50.

[29] L. Xie, J. Ding and F. Ding (2009). *Gradient based iterative solutions for general linear matrix equations.* Computers and Mathematics with Applications, vol. 58(7), pp.1441–1448.

[30] J. Ding, Y. Liu and F. Ding (2010). *Iterative solutions to matrix equations of the form* $AiXBi= Fi$. Computers and Mathematics with Applications, vol. 59(11), 3500–3507.