Global Convergence Properties of the Splitting Iterations for Solving Linear Least Squares Problems

Felipe Marti-Lopez

Department of NeuroPhysics
Cuban Neuroscience Center
P.O. Box 6880, Havana City, Cuba
optoe@cneuro.edu.cu

Abstract

This paper deals with the splitting iterations for solving linear least squares problems with a certain positive definite splitting matrix which produce globally convergent sequences of monotonically improved approximates. Such sequences monotonically converge to the unique solution of a certain bi-level linear least squares problem, namely to that linear least squares solution, nearest the initial guess in the sense of a certain elliptic distance. As important consequences, the particular analytic expressions of global limit of the so-called Iterated Tikhonov's Regularization and the Proximal Point method sequences in general form are deduced; as well as of the sequences of certain novel approximating versions of the well-known Jacobi and Minimum Norm iteration respectively.

Keywords: Splitting linear stationary iterations, Global convergence, Bi-level linear least squares, Approximations, Proximal Point iterations, Regularization.

1 Introduction

At present, researchers and mathematicians continue investigating general methods for solving linear least squares problems that provide more numerically stable approximates. On one hand, the so-called *exact* methods -including *Singular Value Decomposition*-based methods [5, 7]- provide too sensitive approximates while solving severely ill-conditioned noisy-data problems and, on the other hand, direct *regularization*-wise methods [2, 7] do not provide accurate enough approximates while solving well-conditioned exact-data problems.

Of course, searches are also being focused on iterative methods which could provide sequences of numerically stable approximates and whose convergence properties and voluminous theoretical framework [1, 4, 6, 7, 9] make them attractive for combining the approximating features of both aforementioned groups of methods.

This paper concerns some new convergence properties of the splitting linear stationary iterative methods [1, 6, 9]

$$x^{[k+1]} = Q^{-1} \left(Q - A'A \right) x^{[k]} + Q^{-1} A'b, k = 0, 1, ... \tag{1}$$

for solving linear least squares problems

$$\min_{x \in \mathbf{R}^n} f(x) = \| Ax - b \|_2^2 \tag{2}$$

where $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^{m}$ and m , $n \in \mathbb{N}$; and the splitting matrix is of the general form

$$Q = \frac{1}{2} \left(V + A'A \right) \tag{3}$$

and the matrix (which may depend on A, b or $x^{[0]}$) is positive definite.

Most of the convergence properties of such iterations have been well studied by the splitting iteration theory [1, 6, 9] (particularly, the extensive review by Dax [1] summarizes both the previous and present work with all the indispensable references to the essential results by Keller, Young, Tanabe, Elfving, Dax and many others).

Nevertheless, one important question from the approximation point of view has not been answered yet: which would the limit of the sequence (1)-(3) be for any given system matrix and initial guess?

In the present paper, the answer to this question is provided.

2 Main Convergence Properties of Splitting Iterations

Notice that, whereas V is positive definite, so is Q for any matrix A and, therefore, the matrix $Q' + Q - A'A$ is positive definite too. Hence, according to the results by Dax [1] (theorem 9h and corollary 10),

(a) the sequence (1) converges to a certain minimum norm solution of (2) for any given matrix A and any $x^{[0]} \in Range \left\{ 2 \left(V + A'A \right)^{-1} A'A \right\} \subseteq \mathbb{R}^{n}$ (*local convergence limit*),

© Investigación Operativa 1999

(b) the sequence $\{f(x^{[k]}) \,|k = 0, 1, 2, ...\}$ is always a monotonically decreasing sequence. In other words,

$$f\left(x^{[k+1]}\right) < f\left(x^{[k]}\right) \tag{4}$$

at any $x^{[k+1]} \neq x^{[k]}$, $k = 0, 1, 2, ...$; and for any given A and any $x^{[0]}$ (*monotonicity result*).

It is not too hard to realize that the local convergence limit (a) and the monotonicity result (b) are already appropriate for the sequential approximation of solutions of (2) [4, 7]. Nevertheless, from the approximation theory point of view, any global convergence result would allow not only the use of splitting iterations as a more versatile approximation device, but also a better insight into some important, still open numerical problems such as the system matrix conditionedness handling and the stability and accuracy improvement of linear least squares approximates.

3 Global Convergence Properties of the Splitting Iterations for Solving Linear Least Squares Problems with a Certain Positive Definite Splitting Matrix

From (1) and (3), it follows that

$$x^{[k+1]} = (V + A'A)^{-1}(V - A'A)\,x^{[k]} + 2\,(V + A'A)^{-1}A'b = \tag{5}$$

$$= x^{[k]} - (V + A'A)^{-1}\nabla f\left(x^{[k]}\right) =$$

$$= \arg\min_{x\in\mathbb{R}^{\ltimes}}\left\{f(x) + \frac{1}{2}\left(x - x^{[k]}\right)'(V - A'A)\left(x - x^{[k]}\right)\right\}, k = 0, 1, 2, ...,$$

Let L be a nonsingular matrix $L \in \mathbb{R}^{\ltimes\times\ltimes}$ such that

$$V = L'L \tag{6}$$

The *Generalized Singular Value Decomposition* (GSVD) [5, 7] of the matrices A and L provides a certain balanced simultaneous factorization of both matrices that allows the transformation of (5) into the formula of the $(k + 1)$-st term of a certain globally convergent *generalized geometric progression* for any matrix A, because of the special form of the iteration matrix of (5), the positive definiteness of the matrix V and the stationarity of the formula (5).

© Investigación Operativa 1999

In effect, let p be a number $p \in \mathbb{N}$, such that $p = \min\{m, n\}$. After performing the GSVD of A and L, one has

$$A_{m\times n} = L_{A_{m\times p}} \begin{pmatrix} S_A & 0_{p\times(n-p)} \end{pmatrix}_{p\times n} R^{-1}_{n\times n} \tag{7}$$

and

$$L_{n\times n} = L_{L_{n\times p}} \begin{pmatrix} S_L & 0 \\ 0 & I_{(n-p)\times(n-p)} \end{pmatrix}_{n\times n} R^{-1}_{n\times n}, \tag{8}$$

where L_A and L_L are real orthogonal matrices, R is a real nonsingular well-conditioned enough matrix [5, 7], S_A is a real diagonal positive semidefinite matrix, and S_L is a real diagonal positive definite matrix which satisfy the "balancing" equation

$$S'_A S_A + S'_L S_L = I \tag{9}$$

Notice that, the matrix S_A can be assumed to have the structure

$$S_A = \begin{pmatrix} S_{A_+} & 0_{q\times(p-q)} \\ 0_{(p-q)\times q} & 0_{(p-q)\times(p-q)} \end{pmatrix}$$

where $q = rank\,(A)$ and S_{A_+} is a certain diagonal non-zero real matrix.

From the positive definiteness of V and the equality (9) it follows that

$$0 < S'_{A_+} S_{A_+} < 1 \tag{10}$$

the latter inequality meaning that each diagonal entry of the matrix $S'_{A_+} S_{A_+}$ strictly lies between 0 and 1.

So, by making appropriate transformations, from equalities (5) to (9) one obtains

$$x^{[k+1]} = R \begin{pmatrix} I - 2S'_{A_+} S_{A_+} & 0 \\ 0 & I \end{pmatrix} R^{-1} x^{[k]} + 2R \begin{pmatrix} S'_{A_+} & 0 \\ 0 & 0 \end{pmatrix} L'_A b \tag{11}$$

and consequently, by induction,

$$x^{[k+1]} = R \begin{pmatrix} \left[I - 2S'_{A_+} S_{A_+}\right]^{k+1} & 0 \\ 0 & I \end{pmatrix} R^{-1} x^{[0]} + R \begin{pmatrix} \left\{I - \left[I - 2S'_{A_+} S_{A_+}\right]^{k+1}\right\} S^{-1}_{A_+} & 0 \\ 0 & 0 \end{pmatrix} L'_A b \tag{12}$$

The equalities (10) and (12) imply that

$$\lim_{k\to+\infty} x^{[k]} = R \begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix} R^{-1} x^{[0]} + R \begin{pmatrix} S^{-1}_{A_+} & 0 \\ 0 & 0 \end{pmatrix} L'_A b =$$

$$= \arg \min_{A'Ax = A'b} \left(x - x^{[0]}\right)' V \left(x - x^{[0]}\right) =$$

$$= \arg \min_{x \in Arg \min\limits_{y \in \mathbb{R}^{\ltimes}} \|Ay-b\|_2^2} \left(x - x^{[0]}\right)' V \left(x - x^{[0]}\right) = x^* \in \mathbb{R}^{\ltimes} \qquad (13)$$

Furthermore, since x^* is the limit of the sequence $\left\{x^{[k]} | k = 0, 1, 2, ...\right\} \subset \mathbb{R}^{\ltimes}$, x^* is a *fixed point* of the map (5). From this and the equality (11) it follows that

$$x^{[k+1]} - x^* = R \begin{pmatrix} I - 2S'_{A_+} S_{A_+} & 0 \\ 0 & I \end{pmatrix} R^{-1} \left(x^{[k]} - x^*\right)$$

and, therefore,

$$x^{[k+1]} - x^* = R \begin{pmatrix} \left[I - 2S'_{A_+} S_{A_+}\right]^{k+1} & 0 \\ 0 & I \end{pmatrix} R^{-1} \left(x^{[0]} - x^*\right) \qquad (14)$$

at any $x^{[k+1]} \neq x^{[k]}$; for $k = 0, 1, 2, ...$ and from any initial guess $x^{[0]}$.

Finally, from (10) it follows that

$$\left\| R^{-1} \left(x^{[k+1]} - x^*\right) \right\| < \left\| R^{-1} \left(x^{[k]} - x^*\right) \right\|$$

at any $x^{[k]} \neq x^*$.

Notice that the intermediate formula (12), being directly derived from (5), is not only its equivalent, but also a much more economic and numerically stable iterative formula for computations than (5).

The result (13), which answers the question addressed at the beginning of this paper, is illustrated for $n = 2$ in the figure 1.

In the figure 2, the convergence of a sequence produced by an iteration process (5) to the true solution of any minimum norm linear least squares problem (dotted plot) is illustrated by the plots of various consecutive iterates (solid lines).

© Investigación Operativa 1999

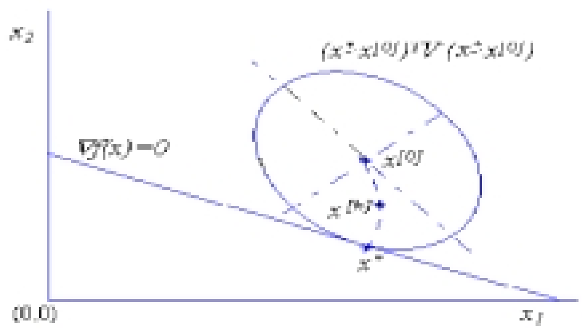

Figure 1.

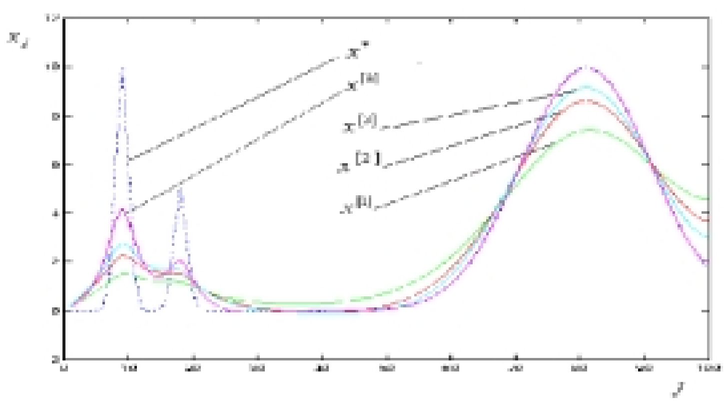

Figure 2.

4 Approximating Splitting Iterations. Some Important Related Cases

Notice that the derived iteration formula (12) produces globally convergent sequences of monotonically improved approximates of solutions of linear least squares problems (2) for any real matrix A and from any initial real guess $x^{[0]}$ as the general formula of an entire class of approximation iterative methods: *Approximating Splitting Iterations.*

© Investigación Operativa 1999

Indeed, let D be a diagonal matrix $D \in \mathbb{R}^{\ltimes \times \ltimes}$, such that

$$D = Diagonal\,(A'A)$$

One can easily verify that if $\lambda \in \mathbb{R}$, $\lambda > -\frac{\alpha^{\min}}{2}$, where α is the least eigenvalue of $2D - A'A$, then the iteration formula of a certain *approximating* version of the well-known *Jacobi Iteration* formula [1, 9]

$$x^{[k+1]} = \left(I - (D + \lambda I)^{-1} A'A\right) x^{[k]} + (D + \lambda I)^{-1} A'b =$$

$$= \arg \min_{x \in \mathbb{R}^{\ltimes}} \left\{ f(x) + \left(x - x^{[k]}\right)' (D - A'A + \lambda I) \left(x - x^{[k]}\right) \right\} \tag{15}$$

$$k = 0, 1, 2, ...$$

can be derived from (5) by taking to be V there the positive definite matrix

$$V = 2D - A'A + 2\lambda I$$

Therefore,

$$\lim_{k \to +\infty} x^{[k]} = \arg \min_{x \in Arg \min\limits_{y \in \mathbb{R}^{\ltimes}} \|Ay - b\|_2^2} \left(x - x^{[0]}\right)' (2D - A'A + 2\lambda I) \left(x - x^{[0]}\right) \tag{16}$$

One can also verify that, if $\lambda \in \mathbb{R}$, $\lambda > 0$, $x^{[0]} = 0$ and P is any nonsingular matrix $P \in \mathbb{R}^{\ltimes \times \ltimes}$, then a certain *approximating Minimum Norm* iteration formula in general form

$$x^{[k+1]} = \left(I - 2\,(\lambda P'P + A'A)^{-1} A'A\right) x^{[k]} + 2\,(\lambda P'P + A'A)^{-1} A'b =$$

$$= \arg \min_{x \in \mathbb{R}^{\ltimes}} \left\{ f(x) + \frac{1}{2} \left(x - x^{[k]}\right)' (\lambda P'P - A'A) \left(x - x^{[k]}\right) \right\}, k = 0, 1, 2, ... \tag{17}$$

can be derived from (5) by taking V to be there the positive definite matrix

$$V = \lambda P'P \tag{18}$$

Therefore,

$$\lim_{k \to +\infty} x^{[k]} = \arg \min_{x \in Arg \min\limits_{y \in \mathbb{R}^{\ltimes}} \|Ay - b\|_2^2} \lambda \left(x - x^{[0]}\right)' P'P \left(x - x^{[0]}\right) \tag{19}$$

© Investigación Operativa 1999

Notice that if $P = I$, then the right-hand side of (19) becomes the expression of the well-known *Minimum Euclidean Norm* solution [4] of any problem (2).

Furthermore, if $\lambda \in \mathbb{R}$, $\lambda > 0$, $x^{[0]} = 0$ and P is a matrix $P \in \mathbb{R}^{\ltimes \times \ltimes}$ such that

$$Null\,(P) \cap Null\,(A) = \{0\}$$

then the formula of the so-called *Iterated Tikhonov's Regularization* method in general form [2, 7],

$$x^{[k+1]} = \arg \min_{x \in \mathbb{R}^{\ltimes}} \left\{ f\,(x) + \lambda \left\| P\left(x - x^{[k]}\right) \right\|_2^2 \right\} =$$

$$= \left(\lambda P'P + A'A\right)^{-1} \left(\lambda P'P\right) x^{[k]} + \left(\lambda P'P + A'A\right)^{-1} A'b, k = 0, 1, 2, ... \tag{20}$$

can be derived from (5) by taking V to be there the positive definite matrix

$$V = 2\lambda P'P + A'A \tag{21}$$

Therefore,

$$\lim_{k \to +\infty} x^{[k]} = \arg \min_{x \in Arg \min\limits_{y \in \mathbb{R}^{\ltimes}} \|Ay - b\|_2^2} x' \left(2\lambda P'P + A'A\right) x.$$

Notice that, from (13) and (21) it follows that the latter result implies the *global convergence* of any sequence (20). Indeed, for any initial guess $x^{[0]} \in \mathbb{R}^{\ltimes}$

$$\lim_{k \to +\infty} x^{[k]} = \arg \min_{x \in Arg \min\limits_{y \in \mathbb{R}^{\ltimes}} \|Ay - b\|_2^2} \left(x - x^{[0]}\right)' \left(2\lambda P'P + A'A\right) \left(x - x^{[0]}\right) \tag{22}$$

Finally, if P is any nonsingular matrix $P \in \mathbb{R}^{\ltimes \times \ltimes}$, then the iteration formula of the well-known *Proximal Point* method [8] in general form for solving (2),

$$x^{[k+1]} = \arg \min_{x \in \mathbb{R}^{\ltimes}} \left\{ f\,(x) + \frac{1}{2} \left\| P\left(x - x^{[k]}\right) \right\|_2^2 \right\} =$$

$$= \left(\frac{1}{2} P'P + A'A\right)^{-1} \left(\frac{1}{2} P'P\right) x^{[k]} + \left(\frac{1}{2} P'P + A'A\right)^{-1} A'b, k = 0, 1, 2, ...$$

can be derived from (5) by taking V to be there the positive definite matrix

$$V = P'P + A'A$$

Therefore,

$$\lim_{k \to +\infty} x^{[k]} = \arg \min_{x \in Arg \min\limits_{y \in \mathbb{R}^{\ltimes}} \|Ay - b\|_2^2} \lambda \left(x - x^{[0]}\right)' \left(P'P + A'A\right) \left(x - x^{[0]}\right) \tag{23}$$

5 Conclusions

In this paper, it was proved that *the iteration formula (12), derived from the splitting linear stationary iteration formula (5), produces globally convergent sequences of monotonically improved approximates of solutions of linear least squares problems (2) for any real matrix A and from any initial real guess $x^{[0]}$ as the general formula of an entire class of certain genuine Successive Approximations methods for solving that kind of problems: Approximating Splitting Iterations.*

Moreover, for each matrix V, such sequences converge to the unique solution of a certain *bi-level linear least squares* problem, namely to that solution of (2), nearest the initial guess $x^{[0]}$ in the sense of the elliptic distance

$$d(x,y) = \sqrt{(x-y)' V (x-y)}$$

As important consequences, the particular analytic expressions of global limit of the so-called Iterated Tikhonov's Regularization (22) and Proximal Point (23) method sequences in general form were deduced for any matrix A; as well as of the sequence of certain novel approximating versions of the well-known Jacobi (16) and the Minimum Norm (19) iterations.

Acknowledgements

The author wishes to thank Professors Luis Martí López, Alfredo Iusem, Juan A. Gómez, Tommy Elfving and Juan C. Jiménez for the encouraging support and all their useful comments and suggestions; and Francisco Saavedra, Jesús Nuñez, Angela Pérez (Ms.) and Josefina Pedroso (Ms.) for the valuable help.

He is also grateful to the anonymous referee for his constructive criticism.

This research was entirely supported on funds under the neurophysical project for Neuroscience P-33-97 "New methods for Brain Electrical Tomography" of National Center for Scientific Research (CNIC), Higher Education Ministry of Cuba.

References

[1] DAX, A., "The convergence of linear stationary iterative processes for solving singular unstructured systems of linear equations". SIAM Reviews, vol. 32, no. 4 (1990), p.p. 611-635.

[2] ELDÉN L., "Algorithms for the regularization of ill-conditioned least squares problems". BIT, 17 (1977), p.p. 134-145.

© Investigación Operativa 1999

[3] FADDEEV D. K. , V. N. FADDEEVA, Computational methods of Linear Algebra. Moscow, MIR Publishers, 1963.

[4] GOLUB G. H., "Numerical methods for solving least squares problems". Numerisches Mathematik, 7 (1965), p.p. 206-216.

[5] GOLUB, G. H. and C. F. VAN LOAN, Matrix Computations (second edition). Baltimore, Johns Hopkins Publishing Corporation, 1989.

[6] HANKE M. , M. NEUMANN,. "Preconditionings and splittings for rectangular systems". Numerische Mathematik, 57 (1990), p.p. 85-95.

[7] HANSEN P. CH., "Regularization Tools: A MatLab package for analysis and solution of discrete ill-posed problems". Numerical Algorithms , 6 (1994) p.p. 1-35.

[8] ROCKAFELLAR, R. T., "Monotone operators and proximal point algorithm". SIAM Journal of Control and Optimization, no. 14 (1976), p.p. 877-898.

[9] STOER J., R. BULIRSCH, Introduction to Numerical Analysis. Second edition. Springer-Verlag, New York-Berlin, 1993.

© Investigación Operativa 1999