Aplikace matematiky

Fridrich Sloboda

A parallel projection method for linear algebraic systems

Aplikace matematiky, Vol. 23 (1978), No. 3, 185–198

Persistent URL: http://dml.cz/dmlcz/103744

A PARALLEL PROJECTION METHOD FOR LINEAR ALGEBRAIC SYSTEMS

FRIDRICH SLOBODA

(Received October 25, 1976)

1. INTRODUCTION

The well-known projection methods of Kaczmarz [10, 3], Cimmino [6, 3] and also the class of projection methods investigated by Householder and Bauer [9] are iterative methods for solving linear algebraic equations and operator equations.

In this paper a direct projection method for linear algebraic systems is described. The method is closely related also to other direct methods, as the elimination method, the orthogonalization method and the method of conjugate directions. The algorithm is in such a form that it may be used for nonlinear problems and also some of its properties can be transferred to nonlinear problems. The method enables us to find the minimum of a function without using derivatives and has the quadratic termination property, i.e., the minimum of a quadratic function is achieved by a finite number of iterations.

2. DESCRIPTION OF THE ALGORITHM

Let us consider the system

$$Ax = b \tag{1}$$

where A is a regular n by n matrix and b is an n-vector. Let us consider the system (1) in the form

$$r_i = b_i - \langle a_i, x \rangle = 0 \quad i = 1, 2, \ldots, n \tag{2}$$

where r_i represent n hyperplanes in the n-dimensional Euclidean space E_n, $\langle a_i, x \rangle = \sum_{j=1}^{n} a_{ij} x_j$, $a_i = (a_{i1}, a_{i2}, \ldots, a_{in})$ is the i-th normal vector of the hyperplane r_i and b_i is the i-th component of the vector b.

Definition. *Let* $x_0^{(0)}, x_0^{(1)}, \ldots, x_0^{(n)}$ *be* $n + 1$ *linearly independent points of the space* E_n. *The algorithm for solution of* (1) *is defined by the recurrent relation*

$$x_i^{(k)} = x_{i-1}^{(k)} + \frac{b_i - \langle a_i, x_{i-1}^{(k)} \rangle}{(a_i, v_{i-1}^{(i)})} v_{i-1}^{(i)} \tag{3}$$

where

$$v_{i-1}^{(i)} = x_{i-1}^{(i)} - x_{i-1}^{(i-1)} \quad i = 1, 2, \ldots, n, \quad k = i, i + 1, \ldots, n .$$

Let $x_0^{(0)}, x_0^{(1)}, \ldots, x_0^{(n)}$ be such points that

$$(a_i, v_{i-1}^{(i)}) \neq 0 \quad i = 1, 2, \ldots, n .$$

Then the following theorem holds:

Theorem 1. *The point* $x_n^{(n)}$ *defined by the algorithm* (3) *is the solution of* (1).

Proof. Let us consider n linearly independent vectors

$$v_0^{(i)} = x_0^{(i)} - x_0^{(0)} \quad i = 1, 2, \ldots, n .$$

Let us project the points $x_0^{(0)}, x_0^{(1)}, \ldots, x_0^{(n)}$ in the direction of the vector $v_0^{(1)}$ into the hyperplane r_1. According to (3) we obtain

$$\begin{aligned} x_1^{(1)} &= x_0^{(1)} + \alpha_0^{(1)} v_0^{(1)} = x_0^{(0)} + \alpha_0^{(0)} v_0^{(1)} \\ x_1^{(2)} &= x_0^{(2)} + \alpha_0^{(2)} v_0^{(1)} \\ &\vdots \\ x_1^{(n)} &= x_0^{(n)} + \alpha_0^{(n)} v_0^{(1)} \end{aligned} \tag{4}$$

where $\alpha_0^{(0)}, \alpha_0^{(1)}, \ldots, \alpha_0^{(n)}$ are real numbers, Let us denote

$$v_1^{(k)} = x_1^{(k)} - x_1^{(1)} \quad k = 2, 3, \ldots, n .$$

Equations (4) yield

$$\begin{aligned} v_1^{(2)} &= x_1^{(2)} - x_1^{(1)} = x_0^{(2)} - x_0^{(0)} + (\alpha_0^{(2)} - \alpha_0^{(0)}) v_0^{(1)} = v_0^{(2)} + (\alpha_0^{(2)} - \alpha_0^{(0)}) v_0^{(1)} \\ v_1^{(3)} &= v_0^{(3)} + (\alpha_0^{(3)} - \alpha_0^{(0)}) v_0^{(1)} \\ &\vdots \\ v_1^{(n)} &= v_0^{(n)} + (\alpha_0^{(n)} - \alpha_0^{(0)}) v_0^{(1)} . \end{aligned} \tag{5}$$

(5) implies that $v_1^{(2)}, v_1^{(3)}, \ldots, v_1^{(n)}$ are linearly independent vectors, so that $x_1^{(1)}, x_1^{(2)}, \ldots, x_1^{(n)}$ are n linearly independent points of the $(n - 1)$-dimensional hyperplane r_1. Let us project the points $x_1^{(1)}, x_1^{(2)}, \ldots, x_1^{(n)}$ in the direction of the vector $v_1^{(2)}$, which lies in the hyperplane r_1, into the hyperplane r_2. We obtain $n - 1$ linearly independent points $x_2^{(2)}, x_2^{(3)}, \ldots, x_2^{(n)}$ of the $(n - 2)$-dimensional linear subspace $r \cap r_2$. Let us denote

$$v_2^{(k)} = x_2^{(k)} - x_2^{(2)} \quad k = 3, 4, \ldots, n . \tag{6}$$

The vectors $v_2^{(k)}$ satisfy

(7)
$$\begin{aligned} v_2^{(3)} &= x_2^{(3)} - x_2^{(2)} = v_1^{(3)} + \left(\alpha_1^{(3)} - \alpha_1^{(1)}\right) v_1^{(2)} \\ v_2^{(4)} &= v_1^{(4)} + \left(\alpha_1^{(4)} - \alpha_1^{(1)}\right) v_1^{(2)} \\ &\vdots \\ v_2^{(n)} &= v_1^{(n)} + \left(\alpha_1^{(n)} - \alpha_1^{(1)}\right) v_1^{(2)} \end{aligned}$$

where $\alpha_1^{(1)}, \alpha_1^{(2)}, \ldots, \alpha_1^{(n)}$ are real numbers. Again (7) implies that $v_2^{(3)}, v_2^{(4)}, \ldots, v_2^{(n)}$ are linearly independent vectors, so that $x_2^{(2)}, x_2^{(3)}, \ldots, x_2^{(n)}$ are $n - 1$ linearly independent points. According to (6) $v_2^{(k)}$ lies in the linear subspace $r_1 \cap r_2$. By induction we obtain that

$$v_i^{(k)} = x_i^{(k)} - x_i^{(i)} \quad k = i + 1,\ i + 2, \ldots, n$$

are linearly independent vectors and $x_i^{(i)}, x_i^{(i+1)}, \ldots, x_i^{(n)}$ are $n - i + 1$ linearly independent points of the $(n - i)$-dimensional linear subspace $\bigcap_{j=1}^{i} r_j$ and therefore the point $x_n^{(n)} = \bigcap_{j=1}^{n} r_j$ is the solution of (1).

Theorem 2. *The vectors*

$$v_i^{(i+1)} = x_i^{(i+1)} - x_i^{(i)} \quad i = 1, 2, \ldots, n - 1$$

defined by the algorithm (3) *have the following property*:

(8)
$$\left(a_j, v_i^{(i+1)}\right) = 0 \quad i = 1, 2, \ldots, n - 1,\ j = 1, 2, \ldots, i\,.$$

Proof. The vector $v_1^{(2)}$ lies in the hyperplane r_1 and therefore $\left(a_1, v_1^{(2)}\right) = 0$. The vector $v_2^{(3)}$ lies in the linear subspace $r_1 \cap r_2$ and therefore $\left(a_1, v_2^{(3)}\right) = 0 \wedge \wedge \left(a_2, v_2^{(3)}\right) = 0$. Since the vector $v_i^{(i+1)}$ lies in the linear subspace $\bigcap_{j=1}^{i} r_j$ we have

$$\left(a_j, v_i^{(i+1)}\right) = 0 \quad i = 1, 2, \ldots, n - 1,\ j = 1, 2, \ldots, i\,.$$

Theorem 3. *The vectors*

$$v_i^{(k)} = x_i^{(k)} - x_i^{(i)} \quad i = 1, 2, \ldots, n - 1,\ k = i + 1, \ldots, n\,.$$

defined by the algorithm (3) *satisfy the following recurrent relation*:

(9)
$$v_i^{(k)} = v_{i-1}^{(k)} - \frac{\left(a_i, v_{i-1}^{(k)}\right)}{\left(a_i, v_{i-1}^{(i)}\right)} v_{i-1}^{(i)}$$

$$i = 1, 2, \ldots, n - 1,\ k = i + 1, \ldots, n\,.$$

Proof. From (3) we obtain

$$x_i^{(i)} = x_{i-1}^{(i)} + \frac{b_i - \langle a_i, x_{i-1}^{(i)}\rangle}{(a_i, v_{i-1}^{(i)})} v_{i-1}^{(i)} = x_{i-1}^{(i-1)} + \frac{b_i - \langle a_i, x_{i-1}^{(i-1)}\rangle}{(a_i, v_{i-1}^{(i)})} v_{i-1}^{(i)} .$$

For the vectors $v_i^{(k)}$ we have

$$v_i^{(k)} = x_i^{(k)} - x_i^{(i)} = (x_{i-1}^{(k)} - x_{i-1}^{(i-1)}) - \frac{(a_i, x_{i-1}^{(k)} - x_{i-1}^{(i-1)})}{(a_i, v_{i-1}^{(i)})} v_{i-1}^{(i)} =$$

$$= v_{i-1}^{(k)} - \frac{(a_i, v_{i-1}^{(k)})}{(a_i, v_{i-1}^{(i)})} v_{i-1}^{(i)}$$

$$i = 1, 2, \ldots, n-1, \ k = i+1, \ldots, n .$$

Theorem 4. *Let* $x_0^{(0)} = (0, \ldots, 0)^{\mathrm{T}}$ *and let* $x_0^{(i)}$ *be of the form*

$$x_0^{(i)} = (0, \ldots, t_i, \ldots, 0)^{\mathrm{T}} \quad i = 1, 2, \ldots, n$$

where $t_i = 1$. *Let* A *be a strictly regular matrix, i.e.,*

$$D_i = \begin{vmatrix} a_{11} & \ldots & a_{1i} \\ \vdots & & \\ a_{i1} & \ldots & a_{ii} \end{vmatrix} \neq 0 \quad i = 1, 2, \ldots, n .$$

Then the matrix defined by the columns of the vectors $v_0^{(1)}, v_1^{(2)}, \ldots, v_{n-1}^{(n)}$ *is upper triangular with unit elements on the diagonal and*

$$(a_i, v_{i-1}^{(i)}) \neq 0 \quad i = 1, 2, \ldots, n .$$

Proof. According to the definition

$$v_0^{(i)} = x_0^{(i)} - x_0^{(0)} \quad i = 1, 2, \ldots, n .$$

Since $v_0^{(1)} = (1, 0, \ldots, 0)^{\mathrm{T}}$ and $a_{11} \neq 0$ we have

$$(a_1, v_0^{(1)}) \neq 0 .$$

According to (9) the vector $v_1^{(2)}$ is of the form

$$(10) \qquad v_1^{(2)} = v_0^{(2)} - \gamma_1^{(1)} v_0^{(1)}$$

where $\gamma_1^{(1)}$ is a real number. From (8) we have $(a_1, v_1^{(2)}) = 0$ and according to the assumption $D_2 \neq 0$. Since the vector $v_1^{(2)}$ is of the form (10) we obtain

$$(a_2, v_1^{(2)}) \neq 0 .$$

Using twice the recurrent relation (9) we obtain

$$(11) \qquad v_2^{(3)} = v_0^{(3)} - \gamma_2^{(2)} v_0^{(2)} - \gamma_2^{(1)} v_0^{(1)}$$

where $\gamma_2^{(2)}, \gamma_2^{(1)}$ are real numbers. (8) implies that $(a_j, v_2^{(3)}) = 0, j = 1, 2$. According to the assumption $D_3 \neq 0$. Since $v_2^{(3)}$ is of the form (11) we have

$$(a_3, v_2^{(3)}) \neq 0 .$$

By induction we obtain that the vector $v_i^{(i+1)}$ is of the form

$$v_i^{(i+1)} = v_0^{(i+1)} - \gamma_i^{(i)} v_0^{(i)} - \ldots - \gamma_i^{(1)} v_0^{(1)} . \tag{12}$$

From (8) we have $(a_j, v_i^{(i+1)}) = 0,\ j = 1, 2, \ldots, i$. According to the assumption $D_{i+1} \neq 0$. Since $v_i^{(i+1)}$ is of the form (12) we have

$$(a_i, v_i^{(i+1)}) \neq 0$$

which is the assertion of the theorem.

Comment. The theorem holds also for arbitrary initial vectors $v_0^{(i)} = x_0^{(i)} - x_0^{(0)}$ for which the matrix of their components is upper triangular with unit elements on the diagonal.

3. SOME PROPERTIES OF THE ALGORITHM

Theorem 5. *Let A be a strictly regular symmetric matrix. Let $v_0^{(i)} = x_0^{(i)} - x_0^{(0)}$ be vectors of the form*

$$v_0^{(i)} = (0, \ldots, t_i, \ldots, 0)^{\mathsf{T}} \quad i = 1, 2, \ldots, n$$

where $t_i = 1$. Then

$$(Av_{i-1}^{(i)}, v_{j-1}^{(j)}) = 0 \quad i \neq j .$$

Proof. From (9) we obtain

$$(Av_0^{(1)}, v_1^{(k)}) = (Av_0^{(1)}, v_0^{(k)} - \frac{(a_1, v_0^{(k)})}{(a_1, v_0^{(1)})} v_0^{(1)}) = (Av_0^{(1)}, v_0^{(k)}) - \frac{(a_1, v_0^{(k)})}{(a_1, v_0^{(1)})} (Av_0^{(1)}, v_0^{(1)})$$

$$k = 2, 3, \ldots, n .$$

Let $a_1^{\mathsf{T}} = Av_0^{(1)}$. Then

$$(Av_0^{(1)}, v_1^{(k)}) = 0 \quad k = 2, 3, \ldots, n . \tag{13}$$

For $v_2^{(k)}$ by virtue of (9) and (13) we obtain

$$(Av_1^{(2)}, v_2^{(k)}) = (Av_0^{(2)}, v_2^{(k)}) - \frac{(a_1, v_0^{(2)})}{(a_1, v_0^{(1)})} (Av_0^{(1)}, v_2^{(k)}) \tag{14}$$

$$k = 3, 4, \ldots, n .$$

From (9) and (13) we obtain

$$(15) \qquad \left(Av_0^{(1)}, v_2^{(k)}\right) = \left(Av_0^{(1)}, v_1^{(k)}\right) - \frac{\left(a_2, v_1^{(k)}\right)}{\left(a_1, v_1^2\right)} \left(Av_0^{(1)}, v_1^{(2)}\right) = 0$$

$$k = 3, 4, \ldots, n$$

$$(16) \qquad \left(Av_0^{(2)}, v_2^{(k)}\right) = \left(Av_0^{(2)}, v_1^{(k)}\right) - \frac{\left(a_2, v_1^{(k)}\right)}{\left(a_2, v_1^{(2)}\right)} \left(Av_0^{(2)}, v_1^{(2)}\right)$$

$$k = 3, 4, \ldots, n\,.$$

Let $a_2^{\mathsf{T}} = Av_0^{(2)}$. According to (14), (15), (16) we have

$$\left(Av_{i-1}^{(i)}, v_2^{(k)}\right) = 0 \quad i = 1, 2, \ k = 3, 4, \ldots, n\,.$$

Let

$$(17) \qquad \left(Av_{i-1}^{(i)}, v_{j-1}^{(k)}\right) = 0 \quad j = 1, 2, \ldots, r, \ i = 1, 2, \ldots, j-1\,.$$

$$k = j, j+1, \ldots, n\,.$$

We shall prove that

$$\left(Av_{i-1}^{(i)}, v_r^{(k)}\right) = 0 \quad i = 1, 2, \ldots, r, \ k = r+1, \ldots, n\,.$$

According to (9), (17) we obtain

$$(18) \qquad \left(Av_{i-1}^{(i)}, v_r^{(k)}\right) = \left(Av_{i-1}^{(i)}, v_{r-1}^{(k)}\right) - \frac{\left(a_r, v_{r-1}^{(k)}\right)}{\left(a_r, v_{r-1}^{(r)}\right)} \left(Av_{i-1}^{(i)}, v_{r-1}^{(r)}\right) = 0$$

$$i = 1, 2, \ldots, r-1, \ k = r+1, \ldots, n\,.$$

It is sufficient to show that

$$(19) \qquad \left(Av_{r-1}^{(r)}, v_r^{(k)}\right) = 0 \quad k = r+1, \ldots, n\,.$$

By means of (9), (18) we have

$$\left(Av_{r-1}^{(r)}, v_r^{(k)}\right) = \left(Av_{r-2}^{(r)}, v_r^{(k)}\right) = \ldots = \left(Av_0^{(r)}, v_r^{(k)}\right) \quad k = r+1, \ldots, n$$

and

$$(20) \qquad \left(Av_0^{(r)}, v_r^{(k)}\right) = \left(Av_0^{(r)}, v_{r-1}^{(k)}\right) - \frac{\left(a_r, v_{r-1}^{(k)}\right)}{\left(a_r, v_{r-1}^{(r)}\right)} \left(Av_0^{(r)}, v_{r-1}^{(r)}\right)$$

$$k = r+1, \ldots, n\,.$$

Let $a_r^{\mathsf{T}} = Av_0^{(r)}$. From (20) we obtain

$$\left(Av_0^{(r)}, v_r^{(k)}\right) = 0 \quad k = r+1, \ldots, n\,.$$

The condition $a_i^{\mathsf{T}} = Av_0^{(i)}$, $i = 1, 2, \ldots, n$ is fullfiled by the unit basic vectors $v_0^{(i)}$ and hence the assertion of the theorem is proved.

Let us denote

$$\alpha_{i-1}^{(k)} = \frac{b_i - \langle a_i, x_{i-1}^{(k)} \rangle}{(a_i, v_{i-1}^{(i)})} \quad i = 1, 2, \ldots, n, \; k = i - 1, i, \ldots, n .$$

For special types of sparse matrices, reduced algorithms can be derived for solution of (1).

Theorem 6. *Let A be a strictly regular, q-diagonal band matrix. Let $x_0^{(0)} = (0, \ldots, 0)^{\mathsf{T}}$ and $x_0^{(k)} = (0, \ldots, t_k, \ldots 0)^{\mathsf{T}}$ where $t_k = 1$. Then*

$$x_i^{(k)} = x_i^{(i)} + x_0^{(k)} \quad k > (q-1)/2 + i \tag{21}$$

$$i = 1, 2, \ldots, n, \; k = i + 1, \ldots, n .$$

Proof. For $i = 1$, according to the algorithm (3), we have

$$x_1^{(1)} = x_0^{(1)} + \alpha_0^{(1)} v_0^{(1)} = x_0^{(0)} + \alpha_0^{(0)} v_0^{(1)} .$$

Since $a_{1k} = 0$ for $k > (q-1)/2 + 1$, we obtain

$$\langle a_1, x_0^{(0)} \rangle = \langle a_1, x_0^{(k)} \rangle = 0 \quad k > (q-1)/2 + 1$$

and hence

$$\alpha_0^{(0)} = \alpha_0^{(k)} \quad k > (q-1)/2 + 1$$

so that according to the algorithm (3) we obtain

$$x_1^{(k)} = x_0^{(k)} + \alpha_0^{(k)} v_0^{(1)} = x_1^{(1)} + x_0^{(k)} \quad k > (q-1)/2 + 1 . \tag{22}$$

For $i = 2$ we get

$$x_2^{(2)} = x_1^{(2)} + \alpha_1^{(2)} v_1^{(2)} = x_1^{(1)} + \alpha_1^{(1)} v_1^{(2)} .$$

According to the assumption $a_{2k} = 0$ for $k > (q-1)/2 + 2$ and by (22) we obtain

$$\langle a_2, x_1^{(1)} \rangle = \langle a_2, x_1^{(k)} \rangle \quad k > (q-1)/2 + 2$$

so that

$$\alpha_1^{(1)} = \alpha_1^{(k)} \quad k > (q-1)/2 + 2 . \tag{23}$$

According to the algorithm (3) and (22), (23) we have

$$x_2^{(k)} = x_1^{(k)} + \alpha_1^{(k)} v_1^{(2)} = x_1^{(1)} + \alpha_1^{(1)} v_1^{(2)} + x_0^{(k)} = x_2^{(2)} + x_0^{(k)}$$

$$k > (q-1)/2 + 2 .$$

By induction the assertion holds for $i = p$. For $i = p + 1$ we have

$$x_{p+1}^{(p+1)} = x_p^{(p+1)} + \alpha_p^{(p+1)} v_p^{(p+1)} = x_p^{(p)} + \alpha_p^{(p)} v_p^{(p+1)} . \tag{24}$$

According to the assumption $a_{p+1k} = 0$ for $k > (q - 1)/2 + p + 1$ and $x_p^{(k)} = x_p^{(p)} + x_0^{(k)}$ for $k > (q - 1)/2 + p + 1$ so that

$$\langle a_{p+1}, x_p^{(p)} \rangle = \langle a_{p+1}, x_p^{(k)} \rangle \quad k > (q - 1)/2 + p + 1$$

and

$$\alpha_p^{(p)} = \alpha_p^{(k)} \quad k > (q - 1)/2 + p + 1 . \tag{25}$$

From (24), (25) and by the induction hypothesis we have

$$x_{p+1}^{(k)} = x_p^{(k)} + \alpha_p^{(k)} v_p^{(p+1)} = x_p^{(p)} + \alpha_p^{(p)} v_p^{(p+1)} + x_0^{(k)} = x_{p+1}^{(p+1)} + x_0^{(k)}$$
$$k > (q - 1)/2 + p + 1 .$$

Similarly we can derive reduced algorithms for other structures of sparse matrices such as block diagonal matrices, bordered block matrices and bordered band matrices. The analysis of the algorithm (3) for 3-diagonal band matrices is given in [18]. The reduced algorithms are also useful for minimization of the corresponding quadratic function. In the case of gradient methods the structure of the matrix has no influence on the structure of the algorithm [14]. The above described reduced algorithms, owing to the simplicity and small storage requirements, may be considered quasi-iterative algorithms for large sparse systems. The point $x_n^{(n)}$ will be considered the initial point $x_0^{(0)}$ for the computation given by the recurrent relation (3). This is the same as to start from $x_0^{(0)} = (0, \ldots, 0)^{\mathsf{T}}$ and to solve

$$Ay = r \tag{26}$$

where $r = b - Ax_n^{(n)}$. The refinement of the solution is then

$$z = x_n^{(n)} + y_n^{(n)}$$

where $y_n^{(n)}$ is the solution of (26).

4. LU DECOMPOSITION OF A

The above described method is related to the elimination method in the sense that the elimination may be viewed as a successive reduction of the number of linearly independent points which lie in the corresponding subspaces. The symmetry of the matrix by the elimination method allows to reduce the number of arithmetic operations. In our case the symmetry has no influence on the number of arithmetic operations, but enables us to generate conjugate vectors and therefore the described method is related also to the methods of conjugate directions.

The method is also related to the elimination method from the viewpoint of LU decomposition of a matrix into the product of a lower triangular matrix and an upper triangular matrix. Let us consider the matrix

$$U = \begin{pmatrix} 1 & u_{12} & \dots & u_{1n} \\ & 1 & \dots & u_{2n} \\ & & \cdot & \\ 0 & & & \cdot \\ & & & 1 \end{pmatrix}$$

where u_{ji} are the components of the vectors $v_{i-1}^{(i)}$ under the conditions of Theorem 4. Then we have

$$(27) \qquad C_1 = AU = \begin{pmatrix} (a_1, v_0^{(1)}) & & & \\ (a_2, v_0^{(1)}) & (a_2, v_1^{(2)}) & & 0 \\ \cdot & & \cdot & \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ (a_n, v_0^{(1)}) & (a_n, v_1^{(2)}) & \dots & (a_n, v_{n-1}^{(n)}) \end{pmatrix}$$

From (27) we obtain

$$(28) \qquad A = C_1 U^{-1} .$$

According to (27) we have

$$\det A = (a_1, v_0^{(1)}) (a_2, v_1^{(2)}) \dots (a_n, v_{n-1}^{(n)}) .$$

The same decomposition is provided by the escalator method [11, 4], the Purcell orthogonalization method [13, 4] and the Fox-Huskey-Wilkinson method of A-orthogonal vectors [5, 4]. The analysis of relationships of various direct methods is given in [4, 7, 17].

Let $v_{i-1}^{(i)}$ be mutually conjugate vectors. Then we can write

$$(29) \qquad U^{\mathrm{T}} A U = D$$

where U is the matrix defined by the columns of the vectors $v_{i-1}^{(i)}$ and D is the diagonal matrix with the diagonal

$$d = (v_0^{(1)}, Av_0^{(1)}), (v_1^{(2)}, Av_1^{(2)}), \dots, (v_{n-1}^{(n)}, Av_{n-1}^{(n)}) .$$

From (29) we obtain

$$A^{-1} = UD^{-1}U^{\mathrm{T}} .$$

According to Theorem 5 we have

$$(v_{i-1}^{(i)}, Av_{i-1}^{(i)}) = (v_{i-1}^{(i)}, a_i) ,$$

i.e., the inner products in the recurrent relation (3). According to Theorem 4 the matrix U is upper triangular and the number of arithmetic operations of the algorithm (3) for the solution of (1) is as follows:

$$\begin{array}{ll} 1/3n^3 + n^2 - 1/3n & \text{operations of multiplications,} \\ 1/3n^3 + 3/2n^2 - 11/6n & \text{operations of additions,} \\ n & \text{operations of divisions.} \end{array}$$

The total storage requirements are less than $n^2/4 + n + 2$ which is known as the smallest upper bound for storage requirements for a direct method. Input of data is very convenient for the algorithm since single rows of the matrix are required on each iteration. For a q-diagonal band matrix it is necessary to store $(q - 1)/2 + 1$ vectors. The above described method can be applied to the solution of a system with strictly diagonally dominant matrix and symmetric positive definite matrix.

5. APPLICATION

In [14] it is shown how the algorithm (3) can be modified in a suitable way to form an efficient algorithm for minimization. Let us consider the quadratic function

$$f(x) = (Ax, x) - 2(b, x) + c \tag{30}$$

where A is a symmetric positive definite n by n matrix, b is an n-vector and c is a scalar.

General algorithm. Let $x_0^{(0)}, x_0^{(1)}, \ldots, x_0^{(n)}$ be $n + 1$ linearly independent points of the space E_n. Then the algorithm for minimization of (30) is defined as follows:

$$x_i^{(k)} = x_{i-1}^{(k)} + \alpha_{i-1}^{(k)} v_{i-1}^{(i)} \tag{31}$$

where

$$v_{i-1}^{(i)} = x_{i-1}^{(i)} - x_{i-1}^{(i-1)}$$

and $\alpha_{i-1}^{(k)}$ are scalar coefficients such that

$$f(x_{i-1}^{(k)} + \alpha v_{i-1}^{(i)}) = \min! \quad i = 1, 2, \ldots, n, \; k = i, i + 1, \ldots, n .$$
$$\alpha = \alpha_{i-1}^{(k)}$$

It was shown [14] that $v_{i-1}^{(i)}$ are mutually conjugate vectors and at the point $x_n^{(n)}$ the function (30) achieves its minimum, i.e., the algorithm has the quadratic termination property.

Special case. Let $x_0^{(0)} = (0, \ldots, 0)^T$ and $x_0^{(k)} = (0, \ldots, t_k, \ldots, 0)^T$ for $k = 1, 2, \ldots \ldots, n$, where $t_k = 1$. Then $\alpha_{i-1}^{(k)}$ defined by the algorithm (31) satisfies

$$\alpha_{i-1}^{(k)} = \frac{b_i - \langle a_i, x_{i-1}^{(k)} \rangle}{(a_i, v_{i-1}^{(i)})} \quad i = 1, 2, \ldots, n, \; k = i, i + 1, \ldots, n . \tag{32}$$

In [15] a generalization of the above described algorithm for minimization of strictly convex functions is suggested. Let $f : E_n \to E_1$ be a continuously differentiable strictly convex function. Let $x_0^{(0)}, x_0^{(1)}, \ldots, x_0^{(n)}$ be $n + 1$ linearly independent points of the space E_n. Let $x_0^{(0)}$ be an initial point and let $x_0^{(k)} = x_0^{(0)} + v_0^{(k)}, v_0^{(k)} = (0, \ldots \ldots, t_k, \ldots, 0)^T$ where $t_k = \lambda$ is a suitable positive real number. Then the algorithm for minimization of $f(x)$ is defined as follows:

(33) **Algorithm.**

Step (i): For given $x_0^{(0)}, x_0^{(k)} = x_0^{(0)} + v_0^{(k)}$ do the calculation by the recurrent relation

$$x_i^{(k)} = x_{i-1}^{(k)} + \alpha_{i-1}^{(k)} w_{i-1}^{(i)}$$

where

$$w_{i-1}^{(i)} = v_{i-1}^{(i)} / \| v_{i-1}^{(i)} \| , \quad v_{i-1}^{(i)} = x_{i-1}^{(i)} - x_{i-1}^{(i-1)}$$

and $\alpha_{i-1}^{(k)}$ is defined as

$$\underset{\alpha = \alpha_{i-1}^{(k)}}{f(x_{i-1}^{(k)} + \alpha w_{i-1}^{(i)}) = \min!} \quad i = 1, 2, \ldots, n, \; k = i, i + 1, \ldots, n .$$

Step (ii) : Replace $x_0^{(0)}$ by $x_n^{(n)}$ and go to Step (i).

According to the choice of λ which defines the vectors $v_0^{(k)}$ in Step (i), we obtain the following algorithms:

Algorithm I. $\lambda = \min (h, \| x_n^{(n)} - x_0^{(0)} \|)$ where h is a constant, usually $h = 0{\cdot}5(1)$.

Algorithm II: $\lambda = \min (h, | f(x_n^{(n)}) - f(x_0^{(0)}) |^p) , \quad p \in (0, 1\rangle$.

Algorithm III: $\lambda = \min (h, \| f'(x_0^{(0)}) \|)$ where $f'(x) = (\partial f / \partial x_1, \ldots, \partial f / \partial x_n)$.

Let us consider the system

(34) $$\partial f / \partial x_i = f_i(x_1, x_2, \ldots, x_n) = 0 \quad i = 1, 2, \ldots, n .$$

Definition. *The occurrence matrix of the system* (34) *is a Boolean matrix, associated with the system* (34) *as follows:*

An element of the matrix ,s_{ij}, *is either a Boolean* 1 *or* 0 *according to the rule*

$$s_{ij} = \begin{cases} 1 & \text{if the } j\text{-th variable appears in the } i\text{-th equation} \\ 0 & \text{otherwise.} \end{cases}$$

This occurrence matrix influences the structure of the algorithm (33). For the algorithm (33), the following theorem holds [15]:

Theorem 7. *Let $f: E_n \to E_1$ be a continuously differentiable strictly convex function satisfying*

$$\lim_{\|x\| \to \infty} f(x) = +\infty .$$

Let the occurrence matrix corresponding to the function $f(x)$ be a q-diagonal band matrix. Let $x_0^{(0)} \in E_n$ be an arbitrary initial point and let $v_0^{(k)} = x_0^{(k)} - x_0^{(0)}$ have the form

$$v_0^{(k)} = (0, \ldots, t_k, \ldots, 0)^{\mathrm{T}}$$

where $t_k = \lambda$. Then $x_i^{(k)}$ defined by the algorithm (33) fulfils

$$x_i^{(k)} = x_i^{(i)} + v_0^{(k)} \quad k > (q-1)/2 + i$$

$$i = 1, 2, \ldots, n, \; k = i, i+1, \ldots, n .$$

For $q = 1$ we obtain the nonlinear Gauss-Seidel iteration. This theorem enables us to reduce the computational time required for minimization of $f(x)$. For a q-diagonal band matrix it is necessary to store $(q-1)/2 + 1$ vectors. According to Theorem 7 an alternative implementation of the algorithm (33) is described as follows [15]:

(35) **Algorithm.**

Step (0): Define λ and set $x_0^{(k)} = x_0^{(0)} + v_0^{(k)}$ for $k = 1, 2, \ldots, n$; set $i = 1$.

Step (1): Compute

$$x_i^{(i)} = x_{i-1}^{(i-1)} + \alpha_{i-1}^{(i-1)} w_{i-1}^{(i)}$$

where

$$w_{i-1}^{(i)} = v_{i-1}^{(i)} / \|v_{i-1}^{(i)}\| , \quad v_{i-1}^{(i)} = x_{i-1}^{(i)} - x_{i-1}^{(i-1)}$$

and $\alpha_{i-1}^{(i-1)}$ is a scalar coefficient such that

$$f(x_{i-1}^{(i-1)} + \alpha w_{i-1}^{(i)}) = \min!$$

$$\alpha = \alpha_{i-1}^{(i-1)}$$

Step (2): Compute

$$x_{i-1}^{(k)} = x_{i-1}^{(k)} - (x_{i-1}^{(k)} - x_i^{(i)}, w_{i-1}^{(i)})\, w_{i-1}^{(i)} \quad k = i+1, \ldots, n .$$

Step (3): Compute

$$x_i^{(k)} = x_{i-1}^{(k)} + \alpha_{i-1}^{(k)} w_{i-1}^{(i)}$$

where $\alpha_{i-1}^{(k)}$ is a scalar coefficient such that

$$f(x_{i-1}^{(k)} + \alpha w_{i-1}^{(i)}) = \min! \quad k = i+1, \ldots, n .$$

$$\alpha = \alpha_{i-1}^{(k)}$$

Step (4): Set $i = i + 1$. If $i \leqq n$ then go to Step (1), else go to Step (5).

Step (5): Replace $x_0^{(0)}$ by $x_n^{(n)}$ and go to Step (0).

In Step (2) we define the orthogonal projections of the point $x_i^{(i)}$ defined in Step (1), on the corresponding parallel directions using the fact that the vectors $w_{i-1}^{(i)}$ are normalized. In Step (3) we consider the linear minimizations on the corresponding parallel directions using the result of Step (2). Step (2) may be considered a predictor and Step (3) a corrector of the local minimizers.

In [16] a modification of the above described algorithm is suggested. The algorithm for band occurrence matrices requires $0(n)$ function evaluations for determination of n linearly independent vectors. The conjugate gradient methods require $0(n^2)$ function evaluations in each iteration and the structure of the occurrence matrix has no influence on the total number of function evaluations.

The other methods of this type [1, 2, 12] require n^2 linear minimizations per iteration and the structure of the occurrence matrix has no influence on the total number of linear minimizations. The structure of algorithms [1, 12] is purely sequential. The above described algorithm is suitable for implementation on a parallel machine. The minimizations on parallel directions are independent of the computational point of view and can be calculated simultaneously. The structure of the occurrence matrix influences the number of processors. Each processor of a parallel machine has to store one vector.

Acknowledgement. I am greatly indebted to Dr. J. Mikloško CSc., Dr. G. Podhájecký and to Professor Dr. M. Fiedler DrSc., for their advice and criticism during the preparation of this paper.

References

[1] *R. P. Brent:* Algorithms for minimization without derivatives, Prentice-Hall, Englewood Cliffs, New Jersey, (1973).

[2] *D. Chazan, W. L. Miranker:* A nongradient and parallel algorithm for unconstrained minimization, SIAM J. Control, 2 (1970), 207—217.

[3] *E. Durand:* Solution numérique des equations algebraiques II, Masson, Paris, (1961).

[4] *D. K. Faddeev, V. N. Faddeeva:* Computational methods of linear algebra, Fizmatgiz, Moscow, (1960), (Russian).

[5] *L. Fox, H. D. Huskey, J. D. Wilkinson:* Notes on the solution of algebraic linear simutlaneous equations, Quart. J. Mech. Appl. Math., 1 (1948), 149—173.

[6] *N. Gastinel:* Analyse numérique linéaire, Hermann, Paris, (1966).

[7] *D. Goldfarb:* Modification methods for inverting matrices and solving systems of linear algebraic equations, Math. of Comp., 26 (1972), 829—852.

[8] *M. R. Hestenes, E. Stiefel:* The method of conjugate gradients for solving linear systems, J. Res. Nat. Bur. Standards, 49 (1952), 409—436.

[9] *A. S. Householder, F. L. Bauer:* On certain iterative methods for solving linear systems, Numer. Math., 2 (1960), 55—59.

[10] *S. Kaczmarz:* Angenäherte Auflösung von Systemen linearen Gleichungen, Bull. Acad. Polon. Sciences et Lettres, A, (1937), 355—357.
[11] *J. Morris:* An escalator process for the solution of linear simultaneous equations, Philos. Mag., 37 (1946), 106—120.
[12] *M. J. D. Powell:* An efficient method for finding minimum of a function of several variables without calculating derivatives, Comp. J., 7 (1964), 155—162.
[13] *E. W. Purcell:* The vector method for solving simultaneous linear equations, J. Math. and Phys., 32 (1954), 180—183.
[14] *F. Sloboda:* Parallel method of conjugate directions for minimization, Apl. mat., 20 (1975), 436—446.
[15] *F. Sloboda:* Nonlinear iterative methods and parallel computation, Apl. mat., 21 (1976), 252—262.
[16] *F. Sloboda:* A conjugate directions method and its application, Proc. of the 8th IFIP Conference on Optimization Techniques, Würzburg, (1977), to appear in Springer Verlag.
[17] *G. Stewart:* Conjugate direction methods for solving systems of linear equations, Numer. Math., 21 (1973), 285—297.
[18] *P. Václavík:* Parallel algorithms for solving 3-diagonal systems of linear equations (Slovak), Thesis, Fac. of Sc., Komenský Univ., Bratislava (1974).

Súhrn

PARALELNÁ PROJEKČNÁ METÓDA PRE LINEÁRNE ALGEBRAICKÉ SYSTÉMY

FRIDRICH SLOBODA

V článku je popísaná priama projekčná metóda pre riešenie systému lineárnych algebraických rovníc. Algoritmus je ekvivalentný algoritmu pre minimalizáciu odpovedajúcej kvadratickej funkcii a možno ho zobecniť pre minimalizáciu ostro konvexných funkcií.

Author's address: Dr. *Fridrich Sloboda*, Ústav technickej kybernetiky SAV, Dúbravská cesta 1, 809 31 Bratislava.