A generalization of iteration-free search vectors of ABS methods

Mostafa Khorramizadeh

Department of Mathematical Sciences
Shiraz University of Technology
Shiraz 71555-313, Iran

m.khorrami@sutech.ac.ir

Article history:
Received July 2014
Accepted August 2014
Available online August 2014

Abstract

Recently, we introduced iteration-free search vectors of the ABS methods and showed how they can be used to compute the search directions of primal--dual interior point methods, when the coefficient matrix of the constraints of the linear programming problem is square. Here, we generalize those results for the general case when the coefficient matrix is non-square.

Keywords: Interior point methods, Infeasible interior point methods, Primal--dual algorithms, ABS algorithms, Search direction.

1. Introduction

Assume that $A$ is an $m \times n$ matrix with $\text{rank}(A)=m$. Let $c, x$ and $s$ be $n$-vectors and $b$ be an $m$-vector. Then, the primal linear programming problem [5,7] is defined to be the minimization of the objective function $c^T x$ subject to the functional constraints $Ax = b$ and the non-negativity constraints $x \geq 0$. The dual of this problem is then the maximization of $b^T y$ subject to $A^T y + s = c$ and $s \leq 0$. In the $k$th iteration of primal--dual infeasible interior point algorithms the search direction is computed by solving the following system of linear equations [3,8,9]:

$$
\begin{pmatrix}
0 & A^T & I_n \\
A & 0 & 0 \\
S^k & 0 & X^k
\end{pmatrix}
\begin{pmatrix}
\Delta x^k \\
\Delta \lambda^k \\
\Delta s^k
\end{pmatrix}
= 
\begin{pmatrix}
-r^k_c \\
-r^k_b \\
-r^k_{xs}
\end{pmatrix}
$$

(1)

where $r^k_c$, $r^k_b$ and $r^k_{xs}$ are given by $r^k_b = A x^k - b$, $r^k_c = A^T \lambda^k + s^k - c$, $r^k_{xs} = -X^k S^k e + \sigma^k \mu_k$, and $X^k$ and $S^k$ denote the diagonal matrices whose diagonal elements are the components of the vectors $x^k$.
and $s^k$, respectively, and $\mathbf{1} = (1, \ldots, 1)^T \in \mathbb{R}^n$. Moreover, $\sigma_k \in (0, 1)$ and $\mu_k = \frac{(x^k)^T s^k}{n}$ are centering parameter and duality gap, respectively.

The ABS class of algorithms was first introduced by Abaffy, Broyden and Spedicato [1,2,6] for solving linear systems. For simplicity in notation, assume $\mathbb{R}^{m \times n}$ has full row rank. An ABS method for solving the linear system $Ax = b$, with $b \in \mathbb{R}^m$, starts with an arbitrary initial vector $x_1 \in \mathbb{R}^n$ and an arbitrary nonsingular matrix $H_1 \in \mathbb{R}^{n \times n}$, the so-called Spedicato's parameter. In the $i$th iteration, having computed $x_i$, solution of the first $i - 1$ equations of $Ax = b$, and $H_i$ a matrix with rows generating the null space of the first $i - 1$ rows of $A$, an ABS algorithm computes $x_{i+1}$ as a solution of the first $i$ equations of $Ax = b$ and $H_{i+1}$, with rows generating the null space of the first $i$ rows of $A$ as explained below [1,2]. To compute the search vector, $z_i$ (Broyden’s parameter) is determined so that $z_i^T H_i a_i \neq 0$ and the search vector is set to be $q_i = H_i^T z_i$. Then, the solution is updated by $x_{i+1} = x_i + a_i q_i$, where the step size $a_i$ is given by $a_i = (b_i - a_i^T x_i)/a_i^T q_i$. Next, $H_{i+1}$ is computed so that $H_{i+1} a_j = 0$, $1 \leq j \leq i$. This can be accomplished by updating $H_i$ (the so called Abaffian) by

$$H_{i+1} = H_i - H_i a_i w_i^T H_i / w_i^T H_i a_i$$

with $w_1 \in \mathbb{R}^n$ (Abaffy’s parameter) satisfying $w_1^T H_i a_i \neq 0$. It can be shown that [1] in an ABS algorithm, we have $s_i = H_i a_i \neq 0$ if and only if $a_i$ is linearly independent of $a_1, a_2, \ldots, a_{i-1}$ (or equivalently, $a_i = 0$ if and only if $a_i$ is linearly dependent on $a_1, a_2, \ldots, a_{i-1}$). The rows of $H_{i+1}$ generate the null space of the first $i$ rows of $A$. If rank($A$) = $n$ then every solution of the first $i$ equations of the system can be rewritten as $x_{i+1} + H_{i+1} s$, for some choice of $s \in \mathbb{R}^n$. Let $x^*$ be the special solution of the linear system $Ax = b$, then there exist a vector $s^* \in \mathbb{R}^n$ such that $x^* = x_{i+1} + H_{i+1} s^*$. Indeed let $r_{i+1} = b - Ax_{i+1}$, $Z = (H_{i+1} a_{i+1}, \ldots, H_{i+1} a_m)^T$ and $d$ be the solution of the linear system $ZZ^Td = (r_{i+1})_m$ where $(r_{i+1})_m$ denotes the last $m-i$ components of the vector $r_{i+1}$. Then, we have $s^* = Z^Td$. If $x_1 = 0$, then the solution of the system is $PT \tau$, where $P = (p_1, \ldots, p_m)$ and $\tau = (\tau_1, \ldots, \tau_m)$.

In Section 2, we describe the ideas of iteration-free search vectors of the ABS algorithm for solving (1). In Section 3, we show how we can use these iteration-free search vectors, to characterize the ABS solution of (1), in case $m < n$. Section 4 is devoted to the concluding remarks.

### 2. Iteration-free search vectors

In the $k$th iteration of primal-dual IIPMs to solve linear optimization problems, the search direction $(\Delta x^k, \Delta \lambda^k, \Delta s^k)$ is computed by solving the linear system (1). We start the ABS algorithm with $x_1 = 0 \in \mathbb{R}^{2n+m}$ and $H_1 = I_{2n+m}$, where $I_{2n+m}$ is the identity matrix. Then, it can be easily verified that for $1 \leq i \leq n$, if we let $z_i = w_i = (0, 0, e_i^T) \in \mathbb{R}^{2n+m}$, where $e_i$ is the $i$th column of the identity matrix $I_n$, then, in the $i$th iteration of the ABS algorithm for solving (1), we have

$$p_i = \begin{pmatrix} 0 \\ 0 \\ e_i \\ \end{pmatrix}, \quad H_{i+1} = \begin{pmatrix} I_n & 0 \\ 0 & I_m - \sum_{j=1}^i \hat{a}_j e_j^T \\ 0 & 0 & I_n - \sum_{j=1}^i e_j e_j^T \\ \end{pmatrix},$$

where $\hat{a}_j$ is the $j$th column of the matrix $A$. Now let $a_i$'s, $1 \leq i \leq m$, denote the $i$th row of the matrix $A$.

By applying the ABS algorithm to the system $A \Delta x^k = -r_b^k$, starting with $H_i = I_n$ and $\bar{x}_i = 0 \in \mathbb{R}^n$, we obtain the parameters $\bar{z}_i \in \mathbb{R}^n$, $\bar{p}_i \in \mathbb{R}^n$ and $H_{i+1} \in \mathbb{R}^{n \times n}$, for $1 \leq i \leq m$ so that $\bar{z}_i^T H_i a_i \neq 0$, $\bar{w}_i^T H_i a_i \neq 0$, $p_i = H_i^T \bar{z}_i$, $H_{i+1} = H_i - H_i a_i \bar{w}_i^T H_i^T / \bar{w}_i^T H_i a_i$. It can be easily verified that for $1 \leq i \leq m$, if we let $z_{n+i} = (\bar{z}_i, 0, 0)^T \in \mathbb{R}^{2n+m}$, $w_{n+i} = (\bar{w}_i, 0, 0)^T \in \mathbb{R}^{2n+m}$, where $\bar{w}_i$ and $\bar{z}_i$ are defined as above, then, in the $(n+i)$th iteration of the ABS algorithm applied to solve (1), we have
\[ p_{n+i} = \begin{pmatrix} \bar{p}_i \\ 0 \\ 0 \end{pmatrix}, \quad H_{n+i+1} = \begin{pmatrix} \bar{H}_{i+1} & 0 & 0 \\ 0 & I_m & -A \end{pmatrix}, \quad (4) \]

Let \( \bar{A} = (\ddot{a}_1, ..., \ddot{a}_n) \), where the \( i \)th column of this matrix is constructed as follows. Let \( |x|^2 = \sum_{i=1}^n x_i^2 \) and \( \ddot{a}_1 = \ddot{a}_1, \ddot{a}_1 = . \) Now, we define \( \ddot{a}_i = \ddot{a}_i - \sum_{j=1}^{i-1} (\dddot{a}_j^T \ddot{a}_i) \ddot{a}_j \), for \( 2 \leq i \leq n \) and \( \ddot{a}_i = \ddot{a}_i / ||\ddot{a}_i|| \), for \( 2 \leq i \leq n \). The vectors \( \ddot{a}_j, 1 \leq j \leq m \), are orthonormal. Using \( \bar{A} \), we define the matrices \( N_j \in \mathbb{R}^{n \times m} \), \( M_j \in \mathbb{R}^{n \times m} \), \( C_j \in \mathbb{R}^{n \times m} \) and \( B^j_k \in \mathbb{R}^{n \times m} \), for \( 1 \leq j \leq m \), according to the following relations:

\[ N_j = -\frac{1}{||\ddot{a}_i||} e_j e_j^T \bar{A}^T, \quad M_j^k = \frac{s^k}{x_j} N_j, \quad C_j = I_m - \sum_{j=1}^i \bar{A} e_j e_j^T \bar{A}^T \quad (5) \]

and

\[ B^j_k = H_{m+1} M_j^k, \quad B^j_k = B^j_{k-1} + B^j_{k-1} AN_j C_{j-1} - \bar{H}_{m+1} M_j^k C_{j-1} \quad (6) \]

In the following theorem, we provide the ABS parameters for the \( (n + m + i) \)th, \( 1 \leq i \leq m \), iteration of the algorithm applied to solve (1). The proof can be found in [4].

**Theorem:** For \( 1 \leq i \leq m \), let

\[ w_{n+m+i} = z_{n+m+i} = \begin{pmatrix} 0 \\ \bar{A} e_i \end{pmatrix} \in \mathbb{R}^{2n+m} \]

Then, in the \((n+m+i)\)th iteration of the ABS algorithm applied to solve (1), we have

\[ p_{n+m+i} = \begin{pmatrix} 0 \\ \bar{A} e_i \\ -A^T \bar{A} e_i \end{pmatrix}, \quad H_{n+m+i+1} = \begin{pmatrix} \bar{H}_{m+1} & B^i_k & -B^i_{k} A \\ 0 & C_i & -C_i A \end{pmatrix}, \quad (7) \]

where the matrices \( C_i \) and \( B^i_k \) are defined by (5) and (6), respectively.

### 2. Computing the search directions

In this section, we provide the search directions of primal-dual IPMs using the search vectors obtained in Section 2 for the Newton system (1) for the general case where, \( m \leq n \). In this case, we first derive an efficient formula to compute \( B^m_k \). Then, using properties of the ABS algorithm, we characterize the solution of system (1) from the solution of the first \( 2n + m \) equations. Consider the case in which \( m \leq n \) and \( B^m_k \) is not zero matrix. For \( 1 \leq i \leq m \), we define

\[ U_i^T = \left( u_i^1, ..., u_i^k \right) \in \mathbb{R}^{m+i}, \quad D^i_k = \left( e_i^1 e_1, ..., e_i^1 e_l \right) \in \mathbb{R}^{n+i} \quad (8) \]

where,

\[ \delta_j = -\frac{1}{||\ddot{a}_i||}, \quad e_j^k = -\frac{s^k \delta_j}{x_j^k}, \quad 1 \leq j \leq i \]

and \( u_i^1 = \bar{A} e_i, \ u_i^j = u_i^{j-1} + \delta_j \bar{A} e_i e_j^T \bar{A}^T u_i^{j-1} \), for \( 1 \leq j \leq i \) and \( u_i^1 = \bar{A} e_i \). The following theorem provides an efficient formula to compute \( B^i_k, 1 \leq i \leq m \).

**Theorem:** Let \( 1 \leq i \leq m \) and the matrices \( D^i_k \) and \( U_i \) be defined as in (8). Then,

\[ B^i_k = \bar{H}_{m+1} D^i_k U_i = \bar{H}_{m+1} \sum_{j=1}^i e_i^k \ddot{e}_j (u_j^i)^T \quad (10) \]

**Proof:** We proceed by induction. For \( i = 1 \), we have

\[ B^1_k = \bar{H}_{m+1} D^1_k U_1 = \bar{H}_{m+1} e_i^1 (u_1^i)^T \]

which is true by the definition of \( B^1_k \) in (6). Suppose that (10) is true up to \( i = 1, 2, ..., t - 1 \). For \( i = t \), from (5), (6) and (9) we have

\[ B^t_k = B^t_k - B^t_{k-1} AN_t C_{t-1} - \bar{H}_{m+1} M^t_k C_{t-1} \quad (11) \]
Using the hypothesis of the induction and (11), (12) and (13), we can write

\[ B_k^T = B_k^{t-1} + \delta_k B_k^{t-1} A e_t e_t^T A^T + \epsilon_k \Pi_{m+1} e_t e_t^T A^T \]

\[ = \bar{H}_{m+1}(\sum_{j=1}^{t-1} \delta_k e_j (u_j^{t-1})^T + \epsilon_k \bar{H}_{m+1} e_t e_t^T A^T + \delta_k \sum_{j=1}^{t-1} \delta_k e_j (u_j^{t-1})^T A e_t e_t^T A^T) \]

\[ = \bar{H}_{m+1}(\sum_{j=1}^{t} \delta_k e_j (u_j^T)^T + \sum_{j=1}^{t} \delta_k e_j (u_j^T)^T + \delta_k (u_j^{t-1})^T A e_t e_t^T A^T) \]

\[ = \bar{H}_{m+1} \sum_{j=1}^{t} \delta_k e_j (u_j^T)^T = \bar{H}_{m+1} D_k^T U_t \]

This completes the induction.

It is worth mentioning that the matrices \(U_t\) and \(\bar{H}_{m+1}\) need to be computed only once in the first iteration of the IIPMs, and \(D_k\) is a diagonal matrix. Here, using properties of the ABS algorithm, we construct the solution of system (1) for the case \(m \leq n\). Let \(S_{k_{(n-m)}}^k = \sum_{j=m+1}^n \delta_k e_j e_j^T A^T - \epsilon_k \Pi_{m+1} e_t e_t^T A^T \).

\((\bar{Z}^k)^T = \bar{H}_{m+1} S_{k_{(n-m)}}^k - B_k^T A X_{(n-m)}^k\), where \(\bar{e}_i\) is the \(i\)th column of the identity matrix \(I_{n-m}\). Assume that

\[ (Z^k)^T = \begin{pmatrix} H_{2n+1} \left( \begin{array}{c} S_{m+1} e_{m+1} \\ 0 \\ x_{m+1} e_{m+1} \end{array} \right), ..., H_{2n+1} \left( \begin{array}{c} s_k e_n \\ 0 \\ x_k e_n \end{array} \right) \end{pmatrix} + (\bar{Z}^k)^T \]

We note that

\[ Z^k(\bar{Z}^k)^T = (Z^k, 0, 0) \begin{pmatrix} 0 \\ 0 \end{pmatrix} Z^k(\bar{Z}^k)^T. \]

The residual vector of system (1) in the solution of the first \(2m + n\) equations is:

\[ r^k = \begin{pmatrix} -r_{c}^k \\ -r_{b}^k \\ -r_{xx}^k \end{pmatrix} - \begin{pmatrix} 0 & A^T & I_n \\ A & 0 & 0 \\ S_k^0 & X_k & 0 \end{pmatrix} \begin{pmatrix} P \lambda_k \\ \bar{A} \beta_k \\ -r_{xx}^k - A^T \bar{A} \beta_k \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} - S_k P \lambda_k + X_k r_c^k + X_k A^T \bar{A} \beta_k, \]

where the last equality obtains from the fact that \((\lambda_k, \beta_k)^T A^T, \lambda_k, \beta_k)^T\) is the solution of the first \(2m + n\) equations. Let \(r_n^k_{(n-m)}\) denotes the last \(n - m\) components of the vector \(r^k\), i.e., \(r_n^k_{(n-m)} = \begin{pmatrix} -r_{xx}^k - S_k P \lambda_k + X_k r_c^k + X_k A^T \bar{A} \beta_k \end{pmatrix}_{n-m}\). Now, let \(d_k\) satisfies \(Z^k(\bar{Z}^k)^T d_k = r_n^k_{(n-m)}\). Thus, using properties of the ABS algorithms, the solution of (1) is as follows:

\[ \begin{pmatrix} \Delta x^k \\ \Delta \lambda^k \\ \Delta s_k \end{pmatrix} = \begin{pmatrix} P \lambda_k \\ -\bar{A} \beta_k \\ -r_{xx}^k - A^T \bar{A} \beta_k \end{pmatrix} + \begin{pmatrix} H_{m+1}^T \left( B_m^T \right)^T 0 0 \\ 0 0 \left( B_m^T \right)^T 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \end{pmatrix} + \begin{pmatrix} P \lambda_k + H_{m+1}^T (Z^k)^T d_k \\ -\bar{A} \beta_k + (B_m^T)^T (\bar{Z}^k)^T d_k \\ -r_{xx}^k - A^T \bar{A} \beta_k - A^T (B_m^T)^T (\bar{Z}^k)^T d_k \end{pmatrix} \]
2. Conclusions

We generalized iteration free search vectors of the ABS algorithms. Then we used these iteration free search vectors to characterize the solution of the Newton systems of primal-dual infeasible interior point methods.

ACKNOWLEDGEMENTS

The author thanks Shiraz university of Technology for supporting this work.

References


