A Necessary and Sufficient Condition for Intersecting a Translated Hyperplane and a Ball

M. El Hallabi

May 1995

TR95-14
A Necessary and Sufficient Condition for Intersecting a Translated Hyperplane and a Ball\textsuperscript{1}

M. El Hallabi\textsuperscript{2}

May 15, 1995

Abstract

In this note we derive a necessary and sufficient condition that ensures the intersection of a translated hyperplane in $\mathbb{R}^n$ and a ball of a given radius, centered at the origin. As application, we use this result to derive a feasible region for the local model subproblem of Vardi’s formulation to solve an equality constrained minimization problem. Also, we relate the translating parameter to the nonzero singular values of the constraint gradient.

Key Words: Translated Linearized Constraints, Translated Hyperplane, QR Decomposition, Singular Value Decomposition, Constrained Optimization, Equality Constrained, Trust-Region.

AMS subject classifications. 65K05, 49D37

\textsuperscript{1}This work was performed while the author was visiting the Computational and Applied Mathematics Department and the Center for Research on Parallel Computation, Rice University, Houston, Texas.

\textsuperscript{2}Currently at Ecole Hassania des Travaux Publics, B.P. 8108 Route d’El Jadida, Km.7, Oasis, Casablanca, Morocco.
Introduction

In this paper, using a theorem of the alternative due to Dax [2], we derive a necessary and sufficient condition for translating, using some parameter, a given hyperplane so that it intersects a ball of radius $\Delta$ centered at the origin. In Section 2, we define our translating parameter. In Section 3, we apply this result to the local model subproblem of Vardi's approach in [8] to solve an equality constrained minimization problem.

2 The translating parameter

Consider the following subset of $\mathbb{R}^n$

$$\mathcal{F}_2(\alpha, \Delta) = \left\{ s \in \mathbb{R}^n \mid \alpha b + A^T s = 0, \|s\|_2 \leq \Delta \right\}$$

where $A$ is a matrix in $\mathbb{R}^n \times \mathbb{R}^m$, $b \in \mathbb{R}^m$, $\Delta > 0$, and $\alpha \neq 0$. In general this subset may be empty. Our goal is to find the values of the parameter $\alpha$ such the related subset

$$\mathcal{F}_2(\alpha, \Delta) = \left\{ s \in \mathbb{R}^n \mid \alpha b + A^T s = 0, \|s\|_2 \leq \Delta \right\}$$

is not empty.

To obtain our result we need the two following technical lemmas.

**Lemma 2.1.** Assume $b \neq 0$, $\Delta > 0$ and that the linear system $A^T s + b = 0$ is consistent. Let $\sigma$ be the smallest positive singular value of $A$. Then

$$|\alpha| \leq \Delta \frac{\sigma}{\|b\|_2},$$

if and only if the matrix

$$M = A^T A - \left( \frac{\alpha}{\Delta} \right)^2 b b^T$$

is positive semi-definite.

**Proof.** Let $A = U \sum V^T$ be the singular value decomposition of $A$, (see Golub and Van Loan [3]), with

$$\sum = \begin{pmatrix} \sum_1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \sum_1 = \text{diag}(\sigma_1, \ldots, \sigma_r)$$

$$V = (V_1, V_2), \quad U = (U_1, U_2),$$

where $U$ and $V$ are orthogonal matrices, and $\sigma_1 \geq \cdots \geq \sigma_r > 0$ are the positive singular values of $A$. First, we show that

$$V_2^T b = 0.$$

Let $s$ be any solution of the linear system $A^T s + b = 0$. We have

$$V^T b + \sum U^T s = 0,$$

or equivalently

$$V_1^T b + \sum_1 U_1^T s = 0.$$
and
\[ V_2^T b = 0 \]
i.e. (2.5). Now we prove that the matrix in (2.3) is positive semi-definite. We have
\[
A^T A - \frac{\sigma}{\Delta} bb^T = V \sum_{i=1}^r V_i^T V_i - \frac{\sigma}{\Delta} bb^T = V \left( \sum_{i=1}^r V_i^T V_i - \frac{\sigma}{\Delta} V_i^T b (V_i^T b)^T \right) V^T.
\]
On the other hand, using (2.5), we obtain
\[
A^T A - \frac{\sigma}{\Delta} bb^T = \begin{pmatrix}
(\sum_1)^2 & 0 \\
0 & 0 
\end{pmatrix} - \frac{\sigma}{\Delta} \begin{pmatrix}
V_1^T b \\
0
\end{pmatrix} \begin{pmatrix}
V_1^T b \\
0
\end{pmatrix}^T
= \begin{pmatrix}
H_1 & 0 \\
0 & 0
\end{pmatrix},
\]
where
(2.7)
\[
H_1 = (\sum_1)^2 - \frac{\sigma}{\Delta} (V_1^T b) (V_1^T b)^T.
\]
Therefore, the matrix \( M \) is positive semi-definite if and only if the matrix \( H_1 \) is positive semi-definite. But the matrix \( H_1 \) results from shifting the matrix \((\sum_1)^2\) by a rank one matrix. Therefore, if the \( \ell_2 \)-norm of the rank-one-matrix is less than \( (\sigma)^2 \), i.e. (2.8)
\[
\frac{\sigma}{\Delta} \| V_1^T b \|_2 \leq (\sigma)^2
\]
which, by (2.5), is equivalent to (2.2), then the matrix \( H_1 \) is positive semi-definite. \( \Box \)

**Lemma 2.2.** Assume that the hypotheses of Lemma 2.1 hold. Then the matrix \( M \) in (2.3) is positive semi-definite if and only if the subset \( F_2(\alpha, \Delta) \) defined in (2.1) is not empty.

**Proof.** By Theorem 1 of the alternative of Dax [2], the subset \( F_2(\alpha, \Delta) \) is not empty if and only if
(2.9)
\[
\Delta \| Ay \|_2 \geq \alpha b^T y \quad \forall y \in \mathbb{R}^n.
\]
On the other hand, the matrix \( M \) in (2.3) is positive semi-definite if and only if
\[
(\Delta)^2 y^T A^T A y - \alpha^2 y^T bb^T y \geq 0
\]
holds for all \( y \in \mathbb{R}^m \), or
(2.10)
\[
\Delta \| Ay \|_2 \geq \alpha b^T y \quad \forall y \in \mathbb{R}^n
\]
which is equivalent to (2.9). \( \Box \)

Now, we derive a necessary and sufficient condition on the parameter \( \alpha \) such that the subset (2.1) is not empty.

**Theorem 2.1.** Assume that \( b \neq 0 \), \( \Delta > 0 \) and that the linear system
(2.11)
\[
b + A^T s = 0,
\]
is consistent. Let $\sigma$ be the smallest positive singular value of $A$. Then

$$|a| \leq \Delta \frac{\sigma}{\|b\|_2},$$

holds if and only if then the subset

$$\mathcal{F}_2(\alpha, \Delta) = \left\{ s \in \mathbb{R}^n \mid \alpha b + A^T s = 0, \quad \|s\|_2 \leq \Delta \right\}$$

is not empty.

Proof. The proof follows obviously from Lemmas 2.1 and 2.2. $\square$

Theorem 2.1 generalizes the result in [1] where $A$ is assumed to have full rank.

3 Application

Consider the following minimization problem

$$\begin{align*}
(EQCP) & \equiv \left\{ \begin{array}{l}
\text{minimize} \quad f(x) \\
\text{subject to} \quad h_i(x) = 0, \quad i = 1 \cdots m,
\end{array} \right.
\end{align*}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ and $h_i : \mathbb{R}^n \to \mathbb{R}, \quad i = 1 \cdots m < n$, are continuously differentiable. To solve (EQCP), SQP algorithms generate sequences $\{x_k\}$ by setting $x_{k+1} = x_k + s_k$, where $s_k$ is obtained as the solution of the local model subproblem

$$\begin{align*}
(QP) & \equiv \left\{ \begin{array}{l}
\text{minimize} \quad c_k^T s + \frac{1}{2} s^T B_k s \\
\text{subject to} \quad h(x_k) + \nabla h(x_k)^T s = 0.
\end{array} \right.
\end{align*}$$

In (QP), $B_k$ is an approximation of the Hessian of the Lagrangian, and $c_k$ represents either the gradient of the objective function of (EQCP) or the gradient of the Lagrangian.

The problem of global convergence has been given much consideration recently. Because the trust-region strategy had proven to be a very successful tool for designing globally convergent algorithms for unconstrained optimization (e.g. Powell [6] and [7]), it was quite natural to extend this strategy to constrained optimization.

In Vardi [8], the author uses a local model subproblem of the form

$$\begin{align*}
(RTRQP) & \equiv \left\{ \begin{array}{l}
\text{minimize} \quad q_k(s) = \nabla c_k^T s + \frac{1}{2} s^T B_k s \\
\text{subject to} \quad \alpha_k h(x_k) + \nabla h(x_k)^T s = 0, \quad \|s\|_2 \leq \Delta_k,
\end{array} \right.
\end{align*}$$

where the parameter $\alpha_k \in (0, 1]$ is chosen such that the feasible region, i.e.

$$\begin{align*}
\mathcal{F}_2(\alpha_k, \Delta_k) = \left\{ s \in \mathbb{R}^n \mid \alpha_k h(x_k) + \nabla h(x_k)^T s = 0, \quad \|s\|_2 \leq \Delta_k \right\}
\end{align*}$$

is not empty. But there was no clear way of choosing such a parameter.

In this paper, we derived a way of computing, a translating parameter $\alpha_k$ such that the constraints of subproblem (RTRQP) are consistent. Then the trial step can be obtained as a solution of subproblem (RTRQP).
The formulation of Vardi [8] along with the choice of $\alpha$ suggested by Theorem 2.1 is used in El Hallabi [4]. The author replace the $\ell_2$-norm in subproblem (RTRQP) by an arbitrary $\ell_p$-norm and uses

$$\alpha^*_k = \min \left( 1, \Delta_k \frac{\sqrt{2}}{2} \frac{\sigma_k}{\|h(x_k)\|_2} \right)$$

where $\sigma_k$ is an estimation of the smallest singular value of $\nabla h(x_k)$, so that the translated hyper plan of linearized constraints intersects the $\ell_2$-ball of radius $\frac{\sqrt{2}}{2} \Delta_k$. The use of the constant $\frac{\sqrt{2}}{2}$ is used to force the translated hyperplane to intersect any $\ell_p$-ball of radius $\Delta_k$. In particular if the $\ell_\infty$-norm is used, then the subproblem (RTRQP) can be formulated as a sequential linear or sequential quadratic programming problem.

Also, assuming $\Delta_k$ remains bounded away from zero, we obtain from (3.5) that there exist a neighborhood of the subset \( \{x \in \mathbb{R}^n | h(x) = 0\} \), say $\mathcal{N}$, depending on of $\nabla h(x)$, such that

$$\alpha_k = 1$$

whenever $x_k \in \mathcal{N}$. The size of such a neighborhood depends on the size of the smallest nonzero singular value of $\nabla h(x_k)$ and on the size of the lower bound of $\Delta_k$. This gives more insight on the condition of restarting the trust region radius at each iteration introduced in El Hallabi and Tapia [5], i.e. $\Delta_k \geq \Delta_{\min}$.

References


