ABSTRACT

Partitioned symmetric matrices, in particular the Hessian of the Lagrangian, play a fundamental role in nonlinear optimization. For this type of matrices S.-P. Han and O. Fujiiwara recently presented an inertia theorem under a certain regularity assumption. We prove that this theorem is true without any regularity assumption. Then we consider matrix extensions preserving the sign of the determinant. Such extensions are shown to be related with the positive definiteness of some Schur complement. Under a regularity assumption this shows, from the viewpoint of linear algebra, the equivalence of strong stability in the sense of M. Kojima and strong regularity in the sense of S. M. Robinson. Finally, we discuss the inertia of a typical one-parameter family of symmetric matrices, occurring in various places in optimization (augmented Lagrangians, focal-point theory, etc.).
1. INTRODUCTION

In the excellent and extensive study [19] D. V. Ouelette summarized several results on partitioned matrices of the type

\[
N = \begin{pmatrix}
A & B \\
C & D
\end{pmatrix}.
\]  (1.1)

She used essentially the concept of the (generalized) Schur complement \(S\) of \(A\) in \(N\) defined by \(S = D - CA^{-1}B\), where \(A^{-1}\) is a (generalized) inverse of \(A\). The utility of the Schur complement (also in relation with inertia) had already been emphasized by R. W. Cottle [3].

Matrices of the type (1.1) occur in nonlinear optimization in a special symmetric form as the Hessian of the Lagrangian (see for example [5], [7]):

\[
M = \begin{pmatrix}
A & B \\
B^T & 0
\end{pmatrix}.
\]  (1.2)

In their recent paper [9] S.-P. Han and O. Fujiwara studied the relationship between the inertia of a real symmetric matrix and its inertia restricted to a linear subspace. Recall that the inertia \(\text{In}(K)\) of a real symmetric matrix \(K\) is defined to be the triple \((\rho, \eta, \theta)\), where \(\rho, \eta, \text{ and } \theta\) are the numbers of positive, negative, and zero eigenvalues, respectively, of \(K\) with multiplicities counted. As a consequence they proved a basic inertia theorem [9, Theorem 3.4] for matrices of the type (1.2), relating the inertia of \(M\) and the inertia of the restriction of \(A\) to the linear subspace orthogonal to the columns of the matrix \(B\). Here they assumed the latter restriction to be nonsingular. In Section 2 we prove that their inertia theorem remains valid without any regularity assumption.

In Section 3 we consider the symmetric-skew-symmetric variant \(\tilde{M}\) of \(M\) in (1.2):

\[
\tilde{M} = \begin{pmatrix}
A & B \\
-B^T & 0
\end{pmatrix}.
\]  (1.3)

Using the result of Section 2, we study extensions of \(\tilde{M}\) in (1.3) by adjoining additional columns to the matrix \(B\). In optimization theory this corresponds to taking subsets of the set of gradients of the binding inequality constraints into account. As a consequence of an algebraic result on such extensions we
obtain the equivalence of strong stability of stationary points in the sense of M. Kojima and strong regularity in the sense of S. M. Robinson, under the assumption of the linear independence of the gradients of binding constraint functions. This provides an insight into these two important concepts from an algebraic point of view, rather than from a topological one.

In Section 4 we discuss the one-parameter family \{A + \gamma B, \gamma \in \mathbb{R}\} where A and B are symmetric matrices with B positive semidefinite. Such families play an important role in various areas in optimization, from both a practical and a theoretical point of view, as will become clear.

Throughout the paper all matrices are assumed to be real matrices. Furthermore, orthogonality \( (\perp) \) in \( \mathbb{R}^n \) refers to the standard inner product; i.e., \( x \perp y \) iff \( x^T y = 0 \).

2. THE INERTIA THEOREM

Let \( L \) be an \( l \)-dimensional linear subspace of \( \mathbb{R}^n \). An \((n, l)\) matrix \( V \) is called a basis matrix for \( L \) if the columns of \( V \) form a basis for \( L \). For an \((n, n)\) symmetric matrix \( K \) we define the inertia of the restriction of \( K \) to \( L \), denoted by \( \text{In}(K/L) \), to be \( \text{In}(V^T K V) \), where \( V \) is a basis matrix for \( L \). From Sylvester's law of inertia it follows that \( \text{In}(V^T K V) \) does not depend on the choice of the basis matrix \( V \) for the linear subspace \( L \); hence, \( \text{In}(K/L) \) is well defined by \( \text{In}(V^T K V) \). When \( L = \{0\} \), \( \text{In}(K/L) = (0,0,0) \). Moreover, we can uniquely define

\[
\text{sign det}(K/L) := \begin{cases} (-1)^{\eta(K/L)} & \text{if } \partial(K/L) = 0, \\ 0 & \text{otherwise}. \end{cases}
\]

Hence, \( K/L \) is nonsingular [positive definite] if \( \partial(K/L) = 0 \) [\( \partial(K/L) = \eta(K/L) = 0 \)]. Also, for an \((n, p)\) matrix \( W \) we denote

\[
\text{Ker}(W) = \{ \zeta \in \mathbb{R}^p | W\zeta = 0 \}. \tag{2.1}
\]

As was already emphasized in Section 1, the following inertia theorem was proved by S.-P. Han and O. Fujiiwara under the additional assumption that \( A/\text{Ker}(B^T) \) is nonsingular [9].

**Theorem 2.1** (Inertia theorem). *Let \( M \) be a symmetric matrix of the type (1.2), where \( A \) and \( B \) are \((n, n)\) and \((n, m)\) matrices, respectively; let*
In [14], the first draft of this paper, we proved Theorem 2.1 by means of a refinement of the Han-Fujiwara proof [9]. In the meantime we have received several responses to the preprint [14], two of them being of particular importance. Firstly, J.-P. Crouzeix mentioned that his coauthored paper [2] seems to be related with ours [14]. Indeed, from a careful analysis of the presentation of Section 3 in [2] one can easily compose a direct proof of Theorem 2.1. Then, J. Stoer communicated an elegant proof to us, which is slightly different from the ideas in [2, Section 3]. In comparison with [2] and [22], our proof in [14] is a bit clumsy; so we have decided to present Stoer’s proof here.

Proof of Theorem 2.1 [22]. Consider the singular-value decomposition (cf. [1]) of \( B^T \),

\[
U^T B^T V = \begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix},
\]

where \( U, V \) are orthogonal matrices of appropriate dimensions,

\[
\Sigma = \text{diag}(\sigma_1, \ldots, \sigma_k) \quad \text{and} \quad \sigma_1 \geq \cdots \geq \sigma_k > 0.
\]

With \( V = (V_1; V_2) \), \( V_1 \) an \((n, k)\) matrix, and \( \bar{A}_{ij} = V_i^TAV_j \), we obtain

\[
\bar{M} := \begin{pmatrix} V^T & 0 \\ 0 & U^T \end{pmatrix} \begin{pmatrix} A & B \\ B^T & 0 \end{pmatrix} \begin{pmatrix} V & 0 \\ 0 & U \end{pmatrix}
\]

\[
= \begin{pmatrix} \bar{A}_{11} & \bar{A}_{12} & \Sigma & 0 \\ \bar{A}_{21} & \bar{A}_{22} & 0 & 0 \\ \Sigma & 0 & \ddots & \ddots \\ 0 & 0 & \ddots & 0 \end{pmatrix}.
\]

Now, it is not difficult to find an \((n + m, n + m)\) nonsingular matrix \( C \) such
that

\[
C^TMC = \begin{pmatrix}
\bar{A}_{11} & I_k & 0 & 0 \\
I_k & 0 & 0 & 0 \\
\cdots & \cdots & \cdots & \cdots \\
0 & 0 & \bar{A}_{22} & 0 \\
0 & 0 & 0 & 0
\end{pmatrix}_{n-k},
\]

(2.3)

with \(I_k\) the \((k, k)\) identity matrix.

Next, we have (see [2, p. 286] for a nice proof)

\[
\text{In} \begin{pmatrix}
\bar{A}_{11} & I_k \\
I_k & 0
\end{pmatrix} = (k, k, 0).
\]

(2.4)

Since \(\bar{A}_{22} = V_2^TAV_2\) and \(V_2\) is a basis matrix for \(\text{Ker}(B^T)\), it follows that

\[
\text{In}(\bar{A}_{22}) = \text{In}(A/\text{Ker}(B^T)).
\]

(2.5)

Finally, from (2.3), (2.4), (2.5), and Sylvester’s law of inertia we obtain (2.2).

For later reference we mention the following consequences of Theorem 2.1, which have already been proved in other publications.

**Corollary 2.2.** Consider the matrix \(M\) in (1.2) with \(B\) an \((n, m)\) matrix.

(a) [6, 9, Corollary 3.2] The matrix \(M\) is nonsingular if and only if \(A/\text{Ker}(B^T)\) is nonsingular and \(\text{rank}(B) = m\).

(b) [15, p. 110] If \(\text{rank}(B) = m\) then \(\text{sign det}(\tilde{M}) = \text{sign det}(A/\text{Ker}(B^T))\), where

\[
\tilde{M} = \begin{pmatrix}
A & B \\
-B^T & 0
\end{pmatrix}.
\]

3. CERTAIN MATRIX EXTENSIONS AND THE SCHUR COMPLEMENT

Let \(C\) be an \((n, p)\) matrix. For each index set \(J \subset \{1, \ldots, p\}\) we denote by \(C(J)\) the submatrix of \(C\) obtained by deleting from \(C\) exactly those columns whose indices belong to \(\{1, \ldots, p\} \setminus J\). Let \(A\) be an \((n, n)\) symmet-
ric matrix and $B$ an $(n, m)$ matrix. For each $J \subset \{1, \ldots, p\}$ we introduce the $(n + m + |J|, n + m + |J|)$ matrix $\tilde{M}(J)$ ($|\cdot|$ denoting the cardinality):

$$\tilde{M}(J) = \begin{pmatrix} A & B & C(J) \\ -B^T & 0 & 0 \\ -C(J)^T & 0 & 0 \end{pmatrix}. \quad (3.1)$$

Each matrix $\tilde{M}(J)$ can be viewed as an "extension" of the matrix $\tilde{M}(\emptyset)$. The next theorem is the main result of this section.

**Theorem 3.1.** The following two statements are equivalent:

(i) sign $\det \tilde{M}(J)$ is constant and nonvanishing for all $J \subset \{1, \ldots, p\}$,

(ii) $\tilde{M}(\emptyset)$ is nonsingular, and the Schur complement of $\tilde{M}(\emptyset)$ in $\tilde{M}(\{1, \ldots, p\})$ is positive definite.

The proof of Theorem 3.1 is based on the subsequent Lemma 3.2, proved in [4]. Let $L_a$ and $L_b$ be linear subspaces of $\mathbb{R}^n$ such that $L_a \subset L_b$ and $L_a \neq L_b$. According to [4], a finite sequence of linear subspaces $L_1, L_2, \ldots, L_q$ is called a simple chain from $L_a$ to $L_b$ if $L_1 = L_a$, $L_q = L_b$, $L_i \subset L_{i+1}$, and $\dim(L_{i+1}) = \dim(L_i) + 1$ ($i = 1, \ldots, q - 1$).

**Lemma 3.2 [4, Theorem 3.2].** Let $K$ be an $(n, n)$ symmetric and nonsingular matrix, $L$ a proper linear subspace of $\mathbb{R}^n$, and $L^\perp$ the orthogonal complement of $L$. Then $K^{-1}/L^\perp$ is positive definite if and only if there is some simple chain $L_1, \ldots, L_r$ from $L$ to $\mathbb{R}^n$ with $\text{sign} \det(K/L_i)$ constant and nonvanishing for $i = 1, \ldots, r$.

**Remark 3.1.** As remarked in [4], the clause "there is some simple chain..." in Lemma 3.2 can be replaced by "for every simple chain $L_1, \ldots, L_r$ from $L$ to $\mathbb{R}^n$, $\text{sign} \det(K/L_i)$ is constant and nonvanishing for $i = 1, \ldots, r$.

**Proof of Theorem 3.1.** Replace in (3.1) the entry $-B^T$ by $+B^T$, and denote the resulting matrix by $M^*(J)$. In particular, $\det M^*(J) = (-1)^m \det \tilde{M}(J)$ for every $J \subset \{1, \ldots, p\}$, and then Corollary 2.2(a), (b) implies the equivalence of (i) with

(iii) $\text{sign} \det(M/L(J))$ is constant and nonvanishing for each $J \subset \{1, \ldots, p\}$,
where \( M \) is defined as in (1.2) and

\[
L(J) = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^{n+m} \bigg| C(J)^T x = 0 \right\}.
\]

By means of the subspaces \( L(J) \) we can construct simple chains from \( L := L(\{1, \ldots, p\}) \) to \( \mathbb{R}^{n+m} \). Note that \( M \) is nonsingular iff \( \tilde{M}(\emptyset) \) is nonsingular. But then, it follows from Lemma 3.2 and Remark 3.1 that (iii) is equivalent with

(iv) \( M \) is nonsingular and \( M^{-1}/L^\perp \) is positive definite.

Since \((C^T:0)^T\) with \( C = C(\{1, \ldots, p\}) \) is a basis matrix for the linear subspace \( L^\perp \), the positive definiteness of \( M^{-1}/L^\perp \) means that the matrix \((C^T:0)^T M^{-1} (C^T:0)^T\) is positive definite. But the latter matrix equals the Schur complement of \( \tilde{M} = \tilde{M}(\emptyset) \) in \( \tilde{M}(\{1, \ldots, p\}) \), namely

\[-(C^T:0)^T \tilde{M}^{-1} (C^T:0)^T.\]

So (ii) is equivalent with (iv), and hence (i) is equivalent with (ii).

REMARK 3.2. From the derivation of \( M^{-1} \) (see [13], [17]) it is easy to verify that \( M^{-1}/L \) is positive definite if the matrix \( C^T W(W^T A W)^{-1} W^T C \) is positive definite, where \( W \) is a basis matrix for \( \text{Ker}(B^T) \). Here, the matrix \( W^T A W \) which has to be inverted has dimension \( n - m \).

The ideas in the foregoing are closely related to the concept of strong stability in the sense of Kojima [15] and the concept of strong regularity in the sense of Robinson [20] in nonlinear optimization. Indeed, consider a nonlinear optimization problem of the type

\[
\min \{ f_0(x) \mid f_i(x) = 0, \ i = 1, \ldots, r, \ f_i(x) \leq 0, \ i = r + 1, \ldots, l \}, \quad (3.2)
\]

where \( f_i : \mathbb{R}^n \to \mathbb{R} \) is twice continuously differentiable for \( i = 0, 1, \ldots, l \). The Hessian of the Lagrangian \( L \) of (3.2), \( L(x, u) = f_0(x) + \sum_{i=1}^r u_i f_i(x) \), is of the type (1.2), whereas the Jacobian matrix of the subsequent system of equations (3.3) (used in this form for instance in [15]) is of the type (3.1),

\[
\nabla f_0(x) + \sum_{i \in \{1, \ldots, r\} \cup J} u_i \nabla f_i(x) = 0,
\]

\[-f_i(x) = 0, \quad i \in \{1, \ldots, r\} \cup J \quad (3.3)
\]

with \( J \subset \{r + 1, \ldots, l\} \).
The system (3.3) is closely related to the Karush-Kuhn-Tucker (KKT) relations [5] for (3.2):

\[ \nabla_x L(x, u) = 0, \quad f_i(x) = 0, \quad i = 1, \ldots, r, \]
\[ f_i(x) \leq 0, \quad u_i \geq 0, \quad u_i f_i(x) = 0, \quad i = r+1, \ldots, l. \quad (3.4) \]

A pair \((\bar{x}, \bar{u})\) satisfying (3.4) is called a KKT point for (3.1), whereas \(\bar{x}\) is called a stationary point for (3.1) and \(\bar{u}\) a corresponding Lagrange multiplier. Let \(\bar{x}\) be a feasible point for (3.2), i.e., \(\bar{x}\) satisfies all constraints in (3.2). Denote \(J_0(\bar{x}) = \{ i \in \{r+1, \ldots, l\} | f_i(\bar{x}) = 0 \}\), the index set of binding inequality constraints, and for a given KKT point \((\bar{x}, \bar{u})\) for (3.1) let \(J_+(\bar{x}, \bar{u}) = \{ i \in \{r+1, \ldots, l\} | \bar{u}_i > 0 \} \subseteq J_0(\bar{x})\).

The linear-independence constraint qualification (LICQ) is satisfied at a feasible point \(\bar{x}\) if the vectors \(\nabla f_i(\bar{x}), \quad i \in \{1, \ldots, r\} \cup J_0(\bar{x})\), are linearly independent. Obviously, for a stationary point \(\bar{x}\) of (3.2) satisfying the LICQ, the corresponding vector of multipliers \(\bar{u}\) is unique.

Roughly speaking, the concept of strong stability of Kojima [15] refers locally to the existence of a unique stationary point which depends continuously on the problem data (perturbations of \(f_i\) up to second-order terms, \(i = 0, 1, \ldots, l\)). On the other hand, via the approach of generalized equations, Robinson's concept of strong regularity [20] refers locally to the existence of a unique KKT point which depends Lipschitz-continuously on the problem data. Referring to [15] and [20], equivalent algebraic conditions are now summarized.

For a given KKT point \((\bar{x}, \bar{u})\) for (3.2) we define matrices \(\tilde{M}(J)\) of the type (3.1), where

\[ A = \nabla^2_x L(\bar{x}, \bar{u}), \quad B = (\nabla f_i(\bar{x}), \quad i \in \{1, \ldots, r\} \cup J_+(\bar{x}, \bar{u})) \]
\[ C(J) = (\nabla f_i(\bar{x}), \quad i \in J) \quad \text{with} \quad J \subset J_0(\bar{x}) \setminus J_+(\bar{x}, \bar{u}). \quad (3.5) \]

With this notation and assuming LICQ, it follows that a stationary point \(\bar{x}\) for (3.2) is strongly stable iff (i) in Theorem 3.1 holds (cf. [15, Corollary 4.3]), whereas a KKT point \((x, u)\) is strongly regular iff (ii) in Theorem 3.1 holds (cf. [20, Section 4]).
Hence, assuming the LICQ, strong stability and strong regularity are equivalent by virtue of Theorem 3.1.

4. ON THE INERTIA OF A TYPICAL ONE-PARAMETER FAMILY

Let \( A, B \) be \((n, n)\) symmetric matrices. We consider the following one-parameter family \( P(\gamma) \), \( \gamma \in \mathbb{R} \):

\[
P(\gamma) = A + \gamma B, \quad B \text{ positive semidefinite,}
\]

and we are interested in the inertia of \( P(\gamma) \).

Such a parameter dependent family frequently arises in nonlinear optimization. As an example, \( P(\gamma) \) occurs as the Hessian of an augmented Lagrangian. Here, a constrained optimization problem is replaced by an unconstrained one; the parameter \( \gamma \) plays the role of a penalty parameter related to violated constraints. For a discussion on the subject see [11], [21], and the recent basic paper [10].

A first theorem can be derived using Lemma 4.1 proved in [9]. Associated with an \((n, n)\) matrix \( K \) and a linear subspace \( L \) of \( \mathbb{R}^n \) is the set

\[
K[L] = \{ K\xi | \xi \in L \}.
\]

**LEMMA 4.1** [9, Theorem 2.3]. Let \( K \) be an \((n, n)\) symmetric matrix and \( L \subset \mathbb{R}^n \) a linear subspace. If \( L \cap K[L]^\perp \subset \text{Ker}(K) \) then

\[
\ln(K) = \ln(K/L) + \ln(K/K[L]^\perp) - (0,0,\theta(K/L)). \tag{4.2}
\]

**THEOREM 4.2.** Let \( P(\gamma) \) be defined as in (4.1) with \( k = \text{rank}(B) \) and \( L = \text{Ker}(B) \). If \( L \cap A[L]^\perp \subset \text{Ker}(A) \) then

\[
\ln(P(\gamma)) = \ln(P(\gamma)/A[L]^\perp) + \ln(A/L) - (0,0,\theta(A/L)). \tag{4.3}
\]

**Proof.** Note that \( P(\gamma)[L] = (A + \gamma B)[L] = A[L] \). Next, if \( L \cap A[L]^\perp \subset \text{Ker}(A) \), then \( L \cap A[L]^\perp \subset \text{Ker}(A + \gamma B) \). Hence, \( L \cap (P(\gamma)[L])^\perp \subset \text{Ker}(A + \gamma B) \).
Ker\( (P(\gamma)) \), and we can apply Lemma 4.1. Then (4.3) follows from (4.2) and the additional observation that \( \text{In}(A/L) = \text{In}(A + \gamma B/L) \).

In [10] a further investigation is made under the assumption \( L \cap A[L] \perp \subset \text{Ker}(A) \). If the latter assumption is not fulfilled, then the analysis of the inertia for \( P(\gamma) \) is more complicated. However, using the following two theorems, the inertia can be obtained recursively. These theorems are based on the result in Section 2 and the analysis made in [12, p. 192 ff.]

Let \( K \) be an \( (n, n) \) symmetric matrix and \( L \) a linear subspace of \( \mathbb{R}^n \). We say that \( K/L \) vanishes if \( \text{In}(K/L) = (0, 0, \dim L) \). Moreover, \( K/L \) is said to be singular if \( \mathfrak{g}(K/L) \neq 0 \).

**Theorem 4.3.** Let \( P(\gamma) \) be defined as in (4.1) with \( \text{rank}(B) = k \). Suppose that \( A/\text{Ker}(B) \) vanishes. Let \( W \) and \( V \) be basis matrices for \( \text{Ker}(B) \) and \( \text{Ker}(B)\perp \), respectively. Let \( r = \text{rank}(W^\perp A V) \). Then,

\[
\text{In}(P(\gamma)) = \text{In}(V^TAV + \gamma V^T BV/\text{Ker}(W^TAV)) + (r, r, n - k - r),
\]

where \( V^TBV \) is positive definite.

**Proof.** The matrix \((V; W)\) is a nonsingular \((n, n)\) matrix, and a straightforward calculation shows

\[
(V; W)^T(A + \gamma B)(V; W) = \begin{pmatrix}
V^TAV + \gamma V^T BV & V^TAW \\
W^TAV & 0
\end{pmatrix}.
\]

Now, (4.4) follows from (4.5), Theorem 2.1, and Sylvester’s law.

**Theorem 4.4.** Let \( P(\gamma) \) be defined as in (4.1) with \( \text{rank}(B) = k \). Suppose that \( A/\text{Ker}(B) \) is singular, but nonvanishing. Then,

\[
\text{In}(P(\gamma)) = \text{In}(V^TAV + \gamma V^T BV) + \text{In}((WU)^TA(WU)),
\]

where \( V^TBV \) is positive semidefinite. The matrices \( V, W, U \) in (4.6) are specified as follows. The matrix \( W \) is a basis matrix for \( \text{Ker}(B) \). Let \( s = \text{rank}(W^TAV) \); \( U \) is an \((n - k, s)\) basis matrix for \( \text{Ker}(W^TAW)\perp \). Finally, \( V \) is an \((n, n - s)\) basis matrix for \( \text{Ker}(AWU)^T \).
Proof. From the very definition it follows that the matrix $(V; WU)$ is nonsingular. Then, by calculation, we see

$$(V; WU)^T(A + \gamma B)(V; WU) = \begin{pmatrix} V^TAV + \gamma V^TBV & 0 \\ 0 & (WU)^TA(WU) \end{pmatrix}.$$

(4.7)

Now, (4.6) follows from (4.7) and Sylvester's law.

The inertia $\text{In}(P(\gamma))$ is analyzed according to the following cases:

(a) If $B = 0$, then $\text{In}(P(\gamma)) = \text{In}(A)$, independently of $\gamma$.

(b) If $B \neq 0$ and $A/Ker(B)$ nonsingular, then (4.3) applies with $\delta(A/L) = 0$.

(c) If $B \neq 0$ and $A/Ker(B)$ vanishes, then (4.4) applies.

(d) If $B \neq 0$ and $A/Ker(B)$ is singular but nonvanishing, we can apply Theorem 4.4 and proceed with $V^TAV$ and $V^TBV$ replacing $A$ and $B$, respectively, until one of the former cases occurs.

Another occurrence of the family (4.1) is in focal-point theory. Let $\mathcal{M} \subset \mathbb{R}^n$ be a smooth manifold of dimension less than $n$. Choose $\bar{y} \in \mathbb{R}^n \setminus \text{closure}(\mathcal{M})$, and consider the "distance" function $f_{\bar{y}}(x) = \langle x - \bar{y}, x - \bar{y} \rangle$, where $\langle \cdot , \cdot \rangle$ is some inner product on $\mathbb{R}^n$. Note that $\nabla^2 f_{\bar{y}}$ is positive definite and independent of $\bar{y}$. Suppose that $\bar{x} \in \mathcal{M}$ is a stationary (= critical) point for $f_{\bar{y}}$. Then $\bar{x}$ is also a stationary point for $f_{\bar{y}|\mathcal{M}}$ for all $\bar{y}$ on the straight line $L$ through $\bar{x}$ and $\bar{y}$. Moreover, it is well known (see [12, Theorem 4.3.1], [18, Lemma 6.9]) that $\bar{x}$ is a local minimum for $f_{\bar{y}|\mathcal{M}}$ if $\bar{y} \in L$ and $\bar{y}$ sufficiently near $\bar{x}$. In fact, the one-parameter family (4.1) shows up as the Hessian for the restricted function $f_{\bar{y}|\mathcal{M}}$ and takes the form $B + \gamma(\bar{y})A$, where $B$ is positive definite (i.e., it is the restriction of $\nabla^2 f_{\bar{y}}$ to the tangent space of $\mathcal{M}$ at $\bar{x}$) and $A$ is related to the curvature of $\mathcal{M}$ at $\bar{x}$. Moreover, $\gamma(\bar{y})$ tends to zero as $\bar{y}$ tends to $\bar{x}$. The points $\bar{y}$ on $L$ for which $B + \gamma(\bar{y})A$ is singular are called focal points. Replacement of $f_{\bar{y}}$ by means of the (usual) distance function $\langle x - \bar{y}, x - \bar{y} \rangle^{1/2}$ has no essential influence (cf. [12, Remark 4.3.1]). However, norms with flat faces on the unit sphere (such as $\|x - \bar{y}\|_\infty := \max_i |x_i - \bar{y}_i|$) can give rise to one-parameter families with $B$ positive semi-definite. Such phenomena occur in the theory of Chebyshev approximation (see [12, Chapter 4]). In particular, although the point $\bar{y}$ to which the distance is taken is arbitrarily close to the manifold $\mathcal{M}$, the stationary point $x$ under consideration need not to be a local minimum.
On the level of positive (semi)definiteness, one-parameter families of the type (4.1) subject to polyhedral cones (pointed at the origin) are studied in the work of D. H. Martin and D. H. Jacobson [16]. Within this context we also mention the interesting approach by S.-P. Han and O. L. Mangasarian in [8].

We would like to thank O. Fujiwara for sending the paper [9] by which this research was initiated, and O. L. Mangasarian for drawing his attention to Ouelette’s work. Furthermore, we are indebted to J.-P. Crouzeix and J. Stoer for their stimulating contributions. Last but not least, we are grateful to the referees; their detailed remarks considerably improved the first version of the manuscript.

REFERENCES


22 J. Stoer, Personal communication, 1986.

Received 18 February 1986; revised 12 November 1986