SIAM J. OPTIM.
Vol. 11, No. 1, pp. 243–253

A NOTE ON THE AUGMENTED HESSIAN WHEN THE REDUCED HESSIAN IS SEMIDEFINITE*

KURT M. ANSTREICHER† AND MARGARET H. WRIGHT‡

Abstract. Certain matrix relationships play an important role in optimality conditions and algorithms for nonlinear and semidefinite programming. Let H be an $n \times n$ symmetric matrix, A an $m \times n$ matrix, and Z a basis for the null space of A. (In a typical optimization context, H is the Hessian of a smooth function and A is the Jacobian of a set of constraints.) When the reduced Hessian Z^THZ is positive definite, augmented Lagrangian methods rely on the known existence of a finite $\bar{\rho} \geq 0$ such that, for all $\rho > \bar{\rho}$, the augmented Hessian $H + \rho A^TA$ is positive definite. In this note we analyze the case when Z^THZ is positive semidefinite, i.e., singularity is allowed, and show that the situation is more complicated. In particular, we give a simple necessary and sufficient condition for the existence of a finite $\bar{\rho}$ so that $H + \rho A^TA$ is positive semidefinite for $\rho \geq \bar{\rho}$. A corollary of our result is that if H is nonsingular and indefinite while Z^THZ is positive semidefinite and singular, no such $\bar{\rho}$ exists.

Key words. augmented Hessian, reduced Hessian, inertia, augmented Lagrangian methods

AMS subject classifications. 49D30, 65K05

PII. S1052623499351791

1. Introduction.

Augmented Lagrangian methods, proposed independently in the late 1960s by Hestenes [12] and Powell [18], convert a constrained optimization problem into an unconstrained problem by adding a quadratic penalty term to the Lagrangian function. In contrast to classical quadratic penalty methods, the penalty parameter need not become infinite if the solution of the constrained problem satisfies standard sufficient optimality conditions. This crucial property is a consequence of the following well-known theorem, first quoted by Finsler in 1937 [8].

THEOREM 1.1. *Let H be an $n \times n$ symmetric matrix and A an $m \times n$ matrix of rank m, where $m < n$. Let Z denote a basis for the null space of A. Then Z^THZ is positive definite if and only if there exists a finite $\bar{\rho} \geq 0$ such that, for all $\rho > \bar{\rho}$, $H + \rho A^TA$ is positive definite.*

Proofs can be found in many textbooks; see, for example, [9, 10, 17]. (We also prove this result as part of Theorem 4.2, below.) In the context of constrained optimization, H is the Hessian of a smooth function and A is the Jacobian matrix of a set of constraints. The matrix Z^THZ is usually referred to as the *reduced Hessian*; we shall call $H + \rho A^TA$ the *augmented Hessian*.

In this note we consider the augmented Hessian when Z^THZ is positive semidefinite and thus is allowed to be singular. It is natural to conjecture that in this case there always exists a finite $\bar{\rho}$ such that for all $\rho \geq \bar{\rho}$, $H+\rho A^TA$ is positive semidefinite, but we show by example that this is not true. We also give a precise characterization of when such a finite $\bar{\rho}$ exists. A corollary of our result is that when H is nonsingular and indefinite but Z^THZ is positive semidefinite and singular, no such $\bar{\rho}$ exists.

The results in this paper are closely connected with the theory of augmented Lagrangian methods for constrained nonlinear programming, and are of interest in other

*Received by the editors February 8, 1999; accepted for publication (in revised form) April 7, 2000; published electronically August 24, 2000.
http://www.siam.org/journals/siopt/11-1/35179.html
†Department of Management Sciences, College of Business, The University of Iowa, Iowa City, IA 52242 (kurt-anstreicher@uiowa.edu).
‡Bell Laboratories, Murray Hill, NJ 07974 (mhw@research.bell-labs.com).

areas of optimization as well. In particular, recent work on semidefinite programming (SDP) has made it possible to directly solve optimization problems involving a semidefiniteness condition on a matrix $H(x)$ which is itself a linear function of a vector x of real variables. Such problems are relevant in systems theory, structural optimization, eigenvalue optimization, and combinatorial optimization; see, for example, [2] and [19]. The constraint that $H(x)$ is semidefinite on the null space of a given matrix A can arise in SDP formulations, and in such a case it might be tempting to "reformulate" the constraint as semidefiniteness of $H(x) + \rho A^T A$, where ρ is an additional variable. Our result shows that this reformulation is *not* in general valid. In fact, our interest in the topic arose from an SDP application of this type; see [3].

2. Related work and applications.

In continuous optimization, necessary conditions for optimality typically involve semidefiniteness of certain symmetric matrices. In particular, positive semidefiniteness of $Z^T HZ$ is necessary for existence of a minimizer of the quadratic form $\frac{1}{2}x^T Hx + g^T x$ in the null space of A. Definiteness and semidefiniteness of $Z^T HZ$ are also important in studying generalized convexity of twice-differentiable functions on affine subspaces; see [7].

Conditions characterizing semidefiniteness have been studied in linear algebra and matrix theory for decades; see [5] for a selection of references. In analyzing the quadratic form $x^T Hx$ restricted to the null space of A, [6] shows how the inertia of $Z^T HZ$ is related to that of the bordered matrix

$$\begin{pmatrix} H & A^T \\ A & 0 \end{pmatrix}$$

(usually called the "KKT matrix" in optimization) and various Schur complements. In [5], positive semidefiniteness of $Z^T HZ$ is shown to be equivalent to four different criteria involving (i) the inertia of the bordered matrix, (ii) "augmentability" (that the number of negative eigenvalues of $H - \gamma A^T A$ is exactly equal to $\operatorname{rank}(A)$ for all sufficiently large positive γ), (iii) determinants, and (iv) roots of a polynomial equation.

A complete analysis of the properties of the quadratic form $x^T Hx$ restricted to a general subspace is given in the unifying paper of Maddocks [15], which presents a wide array of inertia theorems that specialize results originally proved in an infinite-dimensional setting [14]. In analyzing stability of KdV multisolitons, where the infinite-dimensional analogues of A and H have the special property that the null space of A contains the null space of H, Lemma 2.3 of [14] gives a sharp lower bound for a penalty parameter that ensures positive semidefiniteness of the augmented Hessian when the reduced Hessian is positive semidefinite.

3. Notation and background.

We consider only real matrices throughout. For a symmetric matrix K we use $\lambda_{\min}(K)$ and $\lambda_{\max}(K)$ to denote the minimal and maximal eigenvalues of K, and $\|K\|$ to denote the spectral norm. For symmetric matrices J and K, $K \succeq J$ means that $K - J$ is positive semidefinite, and $K \succ J$ means that $K - J$ is positive definite. We use $\mathcal{N}(A)$ to denote the null space of a matrix A. When K is symmetric, $\operatorname{In}(K)$ denotes the inertia of K, a triple of nonnegative integers (k_+, k_-, k_0) representing the numbers of positive, negative, and zero eigenvalues of K.

We shall invoke several known properties of symmetric matrices K and J, where

K is sometimes symmetrically partitioned as

$$K = \begin{pmatrix} K_{11} & K_{12} \\ K_{12}^T & K_{22} \end{pmatrix}. \tag{3.1}$$

The following theorem, first proved in [1] (see also [4]), plays a central role in our analysis.

THEOREM 3.1. *The symmetric matrix K, partitioned as in* (3.1)*, is positive semidefinite if and only if the following three conditions hold:*

$$\text{(i) } K_{11} \succeq 0, \qquad \text{(ii) } \mathcal{N}(K_{11}) \subseteq \mathcal{N}(K_{12}^T), \quad \textit{and} \quad \text{(iii) } K_{22} - K_{12}^T K_{11}^{\dagger} K_{12} \succeq 0,$$

where $K_{11}^{\dagger}$ is the Moore–Penrose pseudoinverse of K_{11}.

In addition to Theorem 3.1, we use the following well-known results.

Result 1. $\lambda_{\max}(K)I \succeq K \succeq \lambda_{\min}(K)I$.

Result 2 [13, Observation 7.7.7]. $K \succeq J$ implies $B^T K B \succeq B^T J B$ for any matrix B.

Result 3 [13, Corollary 7.7.4]. If K and J are positive definite, then $K \succeq J$ if and only if $J^{-1} \succeq K^{-1}$.

Result 4 [13, Theorem 7.7.6]. When K is partitioned as in (3.1), K is positive definite if and only if both K_{11} and the Schur complement $K_{22} - K_{12}^T K_{11}^{-1} K_{12}$ are positive definite.

Following common practice in optimization, Z denotes a generic basis for the null space of A, i.e., an $n \times (n-m)$ matrix with full column rank such that $AZ = 0$. Any vector y satisfying $Ay = 0$ can be written as a linear combination of the columns of Z, and the columns of A^T and Z together span all of $\mathcal{R}^n$. If Z_1 and Z_2 are two bases for the null space of A, then $Z_1 = Z_2 F$ for some nonsingular matrix F; this implies, among other things, that the inertia of $Z^T H Z$ does not depend on the choice of Z. For any condition stated in terms of Z, there is an obvious equivalent condition involving A. For example, "$Z^T H Z$ is positive definite" is equivalent to "$x^T H x > 0$ for all x satisfying $Ax = 0$."

4. A general characterization. To begin, we show by example that the obvious generalization of Theorem 1.1, with "semidefinite" replacing "definite" throughout, is not valid. Consider

$$H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \text{and} \quad A = \begin{pmatrix} 1 & 1 \end{pmatrix}, \quad \text{with} \quad Z \text{ taken as } \begin{pmatrix} 1 \\ -1 \end{pmatrix}. \tag{4.1}$$

Observe that $Z^T H Z = 0$ and so is positive semidefinite; note also that $AHA^T = 0$. The augmented Hessian $H + \rho A^T A$ is

$$H + \rho A^T A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + \rho \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} \rho + 1 & \rho \\ \rho & \rho - 1 \end{pmatrix},$$

with eigenvalues $\rho \pm \sqrt{\rho^2 + 1}$. Hence $H + \rho A^T A$ is indefinite for any finite value of ρ.

In the next two subsections, we develop a general approach that leads to a theorem covering both the positive semidefinite case and the (known) positive definite case.

4.1. Preliminaries. To simplify the analysis, we assume that A has full rank, but analogous results can be obtained without this restriction; see the end of section 4.2. By definition of Z and our assumption that A has full rank, the matrix

$(A^T,\ Z)$ is nonsingular. Let $\widetilde{H}_\rho$ be defined as

$$\widetilde{H}_\rho := \begin{pmatrix} Z^T \\ A \end{pmatrix} \left(H + \rho A^T A\right) \left(Z,\ A^T\right) = \begin{pmatrix} Z^T HZ & Z^T HA^T \\ AHZ & AHA^T + \rho(AA^T)^2 \end{pmatrix}. \tag{4.2}$$

We write the eigensystem of $Z^T HZ$ as

$$Z^T HZ = (V,\ U) \begin{pmatrix} \Phi & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} V^T \\ U^T \end{pmatrix},$$

where $\Phi = V^T Z^T HZV$ is a positive diagonal matrix whose dimension is the number of positive eigenvalues of $Z^T HZ$, $(V,\ U)$ is orthogonal, and $Z^T HZU = 0$. Let n_V and n_U denote the numbers of columns of V and U. If $Z^T HZ$ is positive definite, U is empty and $n_U = 0$; if $Z^T HZ$ is positive semidefinite and singular, then $n_U \geq 1$; and if Φ is empty, then $Z^T HZ$ must be the zero matrix. The columns of U form a basis for the null space of $Z^T HZ$, so that every nonzero vector y satisfying $Z^T HZy = 0$ can be written as $y = Uv$ for some nonzero v.

By analogy with (4.2), we define $\bar{H}_\rho$ as

$$\begin{aligned} \bar{H}_\rho &:= \begin{pmatrix} V^T & 0 \\ U^T & 0 \\ 0 & I \end{pmatrix} \widetilde{H}_\rho \begin{pmatrix} V & U & 0 \\ 0 & 0 & I \end{pmatrix} \\ &= \begin{pmatrix} \Phi & 0 & V^T Z^T HA^T \\ 0 & 0 & U^T Z^T HA^T \\ AHZV & AHZU & AHA^T + \rho(AA^T)^2 \end{pmatrix}. \end{aligned} \tag{4.3}$$

Using Sylvester's Law of Inertia (see, for example, [13, Theorem 4.5.8]) twice, observe that

$$\operatorname{In}(H + \rho A^T A) = \operatorname{In}(\widetilde{H}_\rho) = \operatorname{In}(\bar{H}_\rho). \tag{4.4}$$

We shall obtain conditions under which there is a finite ρ such that the augmented Hessian $H + \rho A^T A$ is positive semidefinite (or positive definite) by examining the structure of $\bar{H}_\rho$. It is clear from the form of the matrix in (4.3) that if $n_U > 0$, then $\bar{H}_\rho$ cannot be positive semidefinite if $AHZU \neq 0$. We show below that when $Z^T HZ$ is positive semidefinite and singular, the condition $AHZU = 0$ is in fact necessary and sufficient for the existence of $\bar{\rho}$ such that $H + \rho A^T A \succeq 0$ for all $\rho \geq \bar{\rho}$. First, however, we describe other conditions that are equivalent to the condition $AHZU = 0$.

LEMMA 4.1. *Let H be an $n \times n$ symmetric matrix and A an $m \times n$ matrix of rank m, where $m < n$. Let Z denote a basis for the null space of A. Assume that $Z^T HZ$ is positive semidefinite and singular, and let U be a matrix whose columns are a basis for $\mathcal{N}(Z^T HZ)$. Then the following three conditions are equivalent:*

(a) $AHZU = 0$

(b) $HZU = 0$

(c) $\mathcal{N}(Z^T HZ) = \mathcal{N}(Z^T H^2 Z)$.

Proof. We first show that (a) and (b) are equivalent. Obviously (b) implies (a). For the converse, assume that t_U is any vector for which $HZUt_U \neq 0$. By definition of U as a basis for the null space of $Z^T HZ$, it must be true that $Z^T HZUt_U = 0$. Since Z is a basis for the null space of A, every nonzero vector ζ such that $Z^T \zeta = 0$ must have the form $\zeta = A^T \zeta_A$ for a nonzero ζ_A. Therefore there is a nonzero vector ξ such that

$HZUt_U = A^T\xi$. Since A has full row rank, AA^T is positive definite, and multiplication by A gives $AHZUt_U = AA^T\xi \neq 0$. It follows that $HZU \neq 0 \implies AHZU \neq 0$.

Next we show that (b) and (c) are equivalent. First, symmetry of H implies that $Z^TH^2Z = Z^TH^THZ$, so that $Z^TH^2Zy = 0$ only if $HZy = 0$, i.e., $\mathcal{N}(Z^TH^2Z) = \mathcal{N}(HZ)$. Thus any y in $\mathcal{N}(Z^TH^2Z)$ also satisfies $Z^THZy = 0$, which means that

$$\mathcal{N}(Z^TH^2Z) \subseteq \mathcal{N}(Z^THZ). \tag{4.5}$$

Suppose now that (b) holds. Then, since any nonzero y satisfying $Z^THZy = 0$ has the form $y = Uv$, it must hold that $HZy = 0$, which implies that $Z^TH^2Zy = 0$ and so $\mathcal{N}(Z^THZ) \subseteq \mathcal{N}(Z^TH^2Z)$. Combining this result with (4.5), (b) implies (c).

On the other hand, assume that (b) does not hold, so that $HZU \neq 0$. Then there is a vector $y \in \mathcal{N}(Z^THZ)$ for which $HZy \neq 0$. It follows that $y^TZ^TH^2Zy \neq 0$, which means that $Z^TH^2Zy \neq 0$, so that $y \notin \mathcal{N}(Z^TH^2Z)$. Thus $\mathcal{N}(Z^THZ) \neq \mathcal{N}(Z^TH^2Z)$. □

4.2. The main theorem. We are now ready for our main result. Condition (c) from Lemma 4.1 appears in the theorem, but condition (a) or (b) from that lemma could be used instead. For completeness we also restate and prove Theorem 1.1.

THEOREM 4.2. *Let H be an $n \times n$ symmetric matrix and A an $m \times n$ matrix of rank m, where $m < n$. Let Z denote a basis for the null space of A.*

(a) *If Z^THZ is positive semidefinite and singular, then there exists a finite $\bar\rho \geq 0$ such that $H + \rho A^TA$ is positive semidefinite for all $\rho \geq \bar\rho$, if and only if $\mathcal{N}(Z^THZ) = \mathcal{N}(Z^TH^2Z)$. In this case, $H + \rho A^TA$ is singular for all ρ.*

(b) *Z^THZ is positive definite if and only if there exists a finite $\bar\rho \geq 0$ such that $H + \rho A^TA$ is positive definite for all $\rho > \bar\rho$.*

Proof. According to (4.4), positive semidefiniteness of the augmented Hessian follows from that of $\bar{H}_\rho$ of (4.3). To determine whether $\bar{H}_\rho$ is positive semidefinite, we need to check the three necessary and sufficient conditions of Theorem 3.1, where

$$K_{11} = \begin{pmatrix} \Phi & 0 \\ 0 & 0 \end{pmatrix}, \quad K_{12}^T = (AHZV, \ \ AHZU), \quad \text{and} \quad K_{22} = AHA^T + \rho(AA^T)^2.$$

Condition (i), that $K_{11} \succeq 0$, is obviously satisfied because Φ is positive definite or empty.

To check condition (ii), we observe that the null space of K_{11} consists of all vectors u satisfying

$$\begin{pmatrix} \Phi & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \text{ so that, since } \Phi \succ 0, \ \mathcal{N}(K_{11}) = \left\{ \begin{pmatrix} 0 \\ u_2 \end{pmatrix} \right\} \quad \text{for any } u_2,$$

where u_1 has dimension n_V and u_2 has dimension n_U. All vectors in the null space of K_{12}^T satisfy

$$(AHZV, \ \ AHZU) \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = AHZVu_1 + AHZUu_2 = 0. \tag{4.6}$$

Clearly, an arbitrary vector $(0, \ u_2)^T$ from the null space of K_{11} will not lie in the null space of K_{12}^T if $AHZU \neq 0$. The condition that $\mathcal{N}(K_{11}) \subseteq \mathcal{N}(K_{12}^T)$ thus holds only if $AHZU = 0$ or if U is empty. Since condition (ii) of Theorem 3.1 fails independently of ρ when $AHZU \neq 0$, we conclude that $H + \rho A^TA$ is not positive semidefinite in this case.

If, on the other hand, $AHZU = 0$, condition (ii) of Theorem 3.1 holds, since any vector $(0,\ u_2)^T$ from the null space of K_{11} satisfies (4.6).

To satisfy condition (iii) of Theorem 3.1, we need to show that the generalized Schur complement

$$\widetilde{K} = K_{22} - K_{12}^T K_{11}^\dagger K_{12}$$

is positive semidefinite. If Φ is nonempty, $\widetilde{K}$ is given by

$$\widetilde{K} = AHA^T + \rho(AA^T)^2 - (AHZV,\ 0)\begin{pmatrix}\Phi & 0\\ 0 & 0\end{pmatrix}^\dagger \begin{pmatrix}V^TZ^THA^T\\ 0\end{pmatrix} \tag{4.7}$$

$$= AHA^T + \rho(AA^T)^2 - (AHZV,\ 0)\begin{pmatrix}\Phi^{-1} & 0\\ 0 & 0\end{pmatrix}\begin{pmatrix}V^TZ^THA^T\\ 0\end{pmatrix} \tag{4.8}$$

$$= AHA^T + \rho(AA^T)^2 - AHZV\Phi^{-1}V^TZ^THA^T, \tag{4.9}$$

where (4.8) includes the Moore–Penrose pseudoinverse of K_{11}. Let

$$\chi = \lambda_{\min}(H), \quad \alpha = \lambda_{\min}(AA^T), \quad \text{and} \quad \phi = \lambda_{\min}(\Phi),$$

where $\alpha > 0$ because A has independent rows and $\phi > 0$ because Φ is positive definite. Applying Results 1, 2, and 3 to (4.9), we obtain

$$\widetilde{K} \ \succeq\ A\left[(\chi + \rho\alpha)I - \frac{1}{\phi}HZZ^TH\right]A^T, \tag{4.10}$$

where we use the fact that $VV^T \preceq I$ because V has orthonormal columns. It follows that $\widetilde{K} \ \succeq\ 0$ for all $\rho \ge \bar{\rho}$, where

$$\bar{\rho} = \max(0, \tilde{\rho}), \quad \text{with} \quad \tilde{\rho} = \frac{1}{\alpha}\left(\frac{\|HZZ^TH\|}{\phi} - \chi\right). \tag{4.11}$$

When Φ is empty, i.e. Z^THZ is the zero matrix, the last term in (4.7) is zero, and the generalized Schur complement in condition (iii) of Theorem 3.1 is simply $AHA^T + \rho(AA^T)^2$. Using Results 1 and 2, this matrix is positive semidefinite for $\rho \ge \bar{\rho}$, where $\bar{\rho} = \max(0, -\chi/\alpha)$.

Thus, when $AHZU = 0$ and $\rho \ge \bar{\rho}$, we conclude from Theorem 3.1 that $\bar{H}_\rho$ is positive semidefinite and, applying (4.4), that the augmented Hessian is positive semidefinite. Because we have previously shown that $H + \rho A^TA$ cannot be positive semidefinite when $AHZU \ne 0$, the first statement in (a) now follows from Lemma 4.1.

The second statement in (a) follows because, when Z^THZ is positive semidefinite and singular, there must be a nonzero vector y satisfying $Z^THZy = 0$. Since $\mathcal{N}(Z^THZ) = \mathcal{N}(Z^TH^2Z)$, it must also hold that $HZy = 0$, so that

$$(H + \rho A^TA)Zy = HZy = 0,$$

showing that the augmented Hessian is singular for any value of ρ.

To prove (b), note that positive-definiteness of Z^THZ means that V is square and U is empty, so that $K_{11} = \Phi \succ 0$. It follows directly from (4.10) that the Schur complement $K_{22} - K_{12}^TK_{11}^{-1}K_{12}$ satisfies

$$AHA^T + \rho(AA^T)^2 - AHZ\Phi^{-1}Z^THA^T \ \succ\ 0 \quad \text{when} \quad \rho > \bar{\rho},$$

where $\bar{\rho}$ is given in (4.11). Applying Result 4, we obtain that $H + \rho A^TA$ is positive definite for sufficiently large ρ when Z^THZ is positive definite. Finally, if $H + \rho A^TA$ is positive definite, then $y^TZ^T(H + \rho A^TA)Zy = y^TZ^THZy > 0$ for all $y \neq 0$, and therefore Z^THZ is positive definite. □

Although we use properties of $\bar{H}_\rho$ to obtain all the results in Theorem 4.2, there is a simple alternative proof for the "only if" part of (a). It always holds that $\mathcal{N}(Z^TH^2Z) \subseteq \mathcal{N}(Z^THZ)$ (see (4.5) in the proof of Lemma 4.1). Suppose that $\mathcal{N}(Z^THZ)$ contains a vector that does not lie in $\mathcal{N}(Z^TH^2Z)$, i.e., that there is a vector y such that $Z^THZy = 0$ and $HZy \neq 0$. Let L denote the augmented Hessian, $L = H + \rho A^TA$, and let $x = Zy$. Then $Lx = HZy \neq 0$ and $x^TLx = y^TZ^THZy = 0$. But if the symmetric matrix L is positive semidefinite, x^TLx can be zero only if $Lx = 0$ [13, page 400], so that existence of such a y contradicts positive semidefiniteness of L, as desired.

Regarding the lower bound $\bar{\rho}$ of (4.11), note that if the columns of Z are orthonormal, then $\|HZZ^TH\| \leq \|H\|^2$ because $ZZ^T \preceq I$. In this case it is also easy to see that the value of $\phi = \lambda_{\min}(\Phi)$ is independent of the choice of orthonormal basis Z.

Finally, it is unnecessary to assume that the rows of A are independent. Suppose that A has rank r, with $r < m$. Let $\hat{A}$ consist of r linearly independent rows of A, so that $A^T = (\hat{A}^T,\ R)$, where R is $n \times (m - r)$, and assume that $\rho \geq 0$. Then $H + \rho A^TA = H + \rho\hat{A}^T\hat{A} + \rho RR^T$, which shows that

$$H + \rho\hat{A}^T\hat{A} \succeq 0 \implies H + \rho A^TA \succeq 0 \quad \text{and} \quad H + \rho\hat{A}^T\hat{A} \succ 0 \implies H + \rho A^TA \succ 0.$$

Furthermore, there is an $m \times r$ matrix S of rank r such that $A = S\hat{A}$. Applying Result 1, we have

$$H + \rho A^TA \; = \; H + \rho\hat{A}^TS^TS\hat{A} \; \preceq \; H + \rho\,\|S^TS\|\,\hat{A}^T\hat{A},$$

from which it follows, letting $\hat{\rho} = \rho\,\|S^TS\|$, that

$$H + \rho A^TA \succeq 0 \implies H + \hat{\rho}\,\hat{A}^T\hat{A} \succeq 0 \quad \text{and} \quad H + \rho A^TA \succ 0 \implies H + \hat{\rho}\,\hat{A}^T\hat{A} \succ 0.$$

Since the null spaces of A and $\hat{A}$ coincide, it follows easily that Theorem 4.2 holds without assuming independence of the rows of A.

4.3. Special cases. It is interesting to consider the consequences of Theorem 4.2 in two special cases.

First, assume that H is nonsingular, in which case Z^TH^2Z is positive definite, with an empty null space. Consequently, if Z^THZ is positive semidefinite and singular, then $\mathcal{N}(Z^THZ) \neq \mathcal{N}(Z^TH^2Z)$, and Theorem 4.2(a) shows that the augmented Hessian is not positive semidefinite for any finite ρ. Thus the augmented Hessian can be positive semidefinite only if Z^THZ is positive definite, but in this case Theorem 4.2(b) implies that the augmented Hessian is strictly positive definite for all sufficiently large ρ. Unless H is positive definite, it follows from continuity of the eigenvalues with respect to ρ [13, p. 540] that there must be a positive value of ρ for which the augmented Hessian is positive semidefinite and singular. These observations are summarized in the following corollary.

COROLLARY 4.3. *Let H be a* nonsingular *symmetric $n \times n$ matrix with at least one negative eigenvalue, and let A be an $m \times n$ matrix. Let Z denote a basis for the null space of A. Then*

(a) *if Z^THZ is positive semidefinite and singular, $H + \rho A^TA$ is not positive semidefinite for any finite ρ;*

(b) *if $Z^T HZ$ is positive definite, then $H + \bar\rho A^T A$ is positive semidefinite and singular for some $\bar\rho > 0$, and $H + \rho A^T A$ is positive definite for all $\rho > \bar\rho$.*

A second result involves the case when H itself is positive semidefinite.

LEMMA 4.4. *Let H be a symmetric positive semidefinite $n \times n$ matrix and A an $m \times n$ matrix. Let Z denote a basis for the null space of A. If $Z^T HZ$ is singular, then $\mathcal{N}(Z^T HZ) = \mathcal{N}(Z^T H^2 Z)$.*

Proof. Since H is symmetric and positive semidefinite, it follows immediately that $H + \rho A^T A \succeq 0$ for $\rho \geq 0$ and $Z^T HZ \succeq 0$, and that H has a symmetric square root $H^{1/2}$. Because $Z^T HZ$ is singular, there exists a nonzero vector y such that $Z^T HZy = 0$. For any such y, we have

$$Z^T H^{1/2} H^{1/2} Zy = 0 \implies H^{1/2} Zy = 0 \implies HZy = 0 \implies Z^T H^2 Zy = 0,$$

and therefore $\mathcal{N}(Z^T HZ) \subseteq \mathcal{N}(Z^T H^2 Z)$. As previously observed (see (4.5) in the proof of Lemma 4.1), $\mathcal{N}(Z^T H^2 Z)$ is always a subset of $\mathcal{N}(Z^T HZ)$, so we have $\mathcal{N}(Z^T HZ) = \mathcal{N}(Z^T H^2 Z)$, as required. ☐

4.4. Examples. Example (4.1) shows that the augmented Hessian is not positive semidefinite when H is nonsingular and $Z^T HZ$ is positive semidefinite and singular (case (a) of Corollary 4.3). The condition that $HZU \neq 0$ (which disallows a positive semidefinite augmented Hessian) can also occur when H is singular. Consider

$$H = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \quad \text{and} \quad A = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & -1 \end{pmatrix}, \text{ with } Z \text{ taken as } \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}. \tag{4.12}$$

Then

$$A^T A = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}, \quad Z^T HZ = 0, \quad \text{and} \quad U = 1, \quad \text{so} \quad HZU = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}.$$

The augmented Hessian is given by

$$H + \rho A^T A = \begin{pmatrix} \rho + 1 & 0 & -\rho \\ 0 & \rho & 0 \\ -\rho & 0 & \rho - 1 \end{pmatrix}$$

and has a negative eigenvalue for any finite ρ.

An instance in which $HZU = 0$ but H is not positive semidefinite is

$$H = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \quad \text{and} \quad A = \begin{pmatrix} 0 & 0 & 1 \end{pmatrix}, \text{ with } Z \text{ taken as } \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}. \tag{4.13}$$

Then

$$A^T A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad Z^T HZ = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \text{and} \quad U = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

The augmented Hessian is

$$H + \rho A^T A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \rho - 1 \end{pmatrix} \tag{4.14}$$

and is obviously positive semidefinite for $\rho \geq 1$. Note that, as indicated in Theorem 4.2(a), the augmented Hessian (4.14) is always singular.

To illustrate situation (b) of Corollary 4.3, where H is nonsingular but not positive definite and $Z^T HZ$ is positive definite, consider H from (4.1) with a different A:

$$H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \text{and} \quad A = \begin{pmatrix} 0 & 1, \end{pmatrix} \quad \text{with} \quad Z \text{ taken as } \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Then $Z^T HZ = 1$ and the augmented Hessian is

$$H + \rho A^T A = \begin{pmatrix} 1 & 0 \\ 0 & \rho - 1 \end{pmatrix},$$

so that the augmented Hessian is positive semidefinite and singular for $\rho = 1$ and positive definite for $\rho > 1$.

4.5. A limiting property. A final general property is suggested by examples (4.1) and (4.12), where the smallest eigenvalue of $H + \rho A^T A$ is negative for any finite ρ but converges to zero as $\rho \to \infty$. In the next lemma we show that this is always the case when $HZU \neq 0$ and $Z^T HZ$ is positive semidefinite and singular.

LEMMA 4.5. *Let H be an $n \times n$ symmetric matrix and A an $m \times n$ matrix. Let Z denote a basis for the null space of A. Assume that $Z^T HZ$ is positive semidefinite and singular and that $HZU \neq 0$, where U is a matrix whose columns are a basis for the null space of $Z^T HZ$. Then $\lambda_{\min}(H + \rho A^T A) \to 0$ as $\rho \to \infty$.*

Proof. Let $\{\rho_k\}$ be a sequence of monotonically increasing positive scalars with $\rho_k \to \infty$. Since $HZU \neq 0$, we know from Theorem 4.2 that $H + \rho A^T A$ has a negative eigenvalue for any finite ρ, so that $\lambda_{\min}(H + \rho_k A^T A) < 0$ for all k. If $\lambda_{\min}(H + \rho_k A^T A)$ does not converge to zero as $\rho_k \to \infty$, then there must exist a scalar $\beta > 0$ and vectors x_k, with $\|x_k\| = 1$, such that for all k,

$$x_k^T (H + \rho_k A^T A) x_k = x_k^T H x_k + \rho_k \|A x_k\|^2 \leq -\beta. \tag{4.15}$$

Letting $\bar{x}$ denote any accumulation point of $\{x_k\}$, (4.15) cannot hold as $\rho_k \to \infty$ unless $A\bar{x} = 0$ and $\bar{x}^T H \bar{x} < 0$, which contradicts the assumption that $Z^T HZ$ is positive semidefinite. ☐

5. Relationship to optimality conditions. Consider the equality-constrained quadratic program (QP) of minimizing $\frac{1}{2} x^T H x + x^T g$ subject to $Ax = b$, where A has full rank. Let x_A satisfy $AA^T x_A = b$. As discussed in [11], if $Z^T HZ$ is positive semidefinite and singular, then the QP has weak minimizers if and only if the linear system

$$Z^T HZ x_Z = -Z^T g - Z^T H A^T x_A \tag{5.1}$$

is compatible. It is interesting that there is no direct correspondence between compatibility of (5.1) and positive semidefiniteness of the augmented Hessian.

When $AHZ = 0$, we know from Lemma 4.1 and Theorem 4.2 that $H + \rho A^T A$ is positive semidefinite for sufficiently large finite ρ. However, the system (5.1), which reduces in this case to $Z^T HZ x_Z = -Z^T g$, need not be compatible. In example (4.13), the QP is

$$\begin{aligned} &\text{minimize } \tfrac{1}{2}(x_1^2 - x_3^2) + x_1 g_1 + x_2 g_2 + x_3 g_3 \\ &\text{subject to } x_3 = b. \end{aligned}$$

Using Z from (4.13), we have

$$Z^T g = \begin{pmatrix} g_1 \\ g_2 \end{pmatrix} \quad \text{and} \quad Z^T H Z = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

Thus (5.1) is compatible if and only if $g_2 = 0$, which means that x_2 does not appear in the objective function. If $g_2 \neq 0$, the system (5.1) is not compatible and no weak minimizers exist, although $H + \rho A^T A$ is positive semidefinite for sufficiently large ρ.

When $AHZU \neq 0$, there is no finite ρ for which $H + \rho A^T A$ is positive semidefinite. Even so, there are associated QPs with weak minimizers. The QP problem associated with (4.1) is

$$\begin{array}{rl} \text{minimize} & \frac{1}{2}(x_1^2 - x_2^2) + g_1 x_1 + g_2 x_2 \\ \text{subject to} & x_1 + x_2 = b. \end{array}$$

Since $Z^T HZ = 0$, the system (5.1) is compatible when $Z^T g + Z^T H A^T x_{\scriptscriptstyle A} = 0$, for example, when

$$g = \begin{pmatrix} -1 \\ 0 \end{pmatrix} \quad \text{and} \quad b = 1,$$

and the QP has weak minimizers of the form $(\frac{1}{2} + \beta, \frac{1}{2} - \beta)^T$ for any scalar β.

Acknowledgments. We are very grateful to John Maddocks for his enlightening discussion of many aspects of the problem and for bringing [14], [15], and [16] to our attention. We also thank an anonymous referee for helpful suggestions and for making us aware of the Finsler reference [8].

REFERENCES

[1] A. Albert, *Conditions for positive and nonnegative definiteness in terms of pseudoinverses*, SIAM J. Appl. Math., 17 (1969), pp. 434–440.

[2] F. Alizadeh, *Interior point methods in semidefinite programming with applications to combinatorial optimization*, SIAM J. Optim., 5 (1995), pp. 13–51.

[3] K. M. Anstreicher, *Eigenvalue bounds versus semidefinite relaxations for the quadratic assignment problem*, SIAM J. Optim., 11 (2000), pp. 254–265.

[4] D. Carlson, E. Haynsworth, and T. Markham, *A generalization of the Schur complement by means of the Moore–Penrose inverse*, SIAM J. Appl. Math., 26 (1974), pp. 169–175.

[5] Y. Chabrillac and J.-P. Crouzeix, *Definiteness and semidefiniteness of quadratic forms revisited*, Linear Algebra Appl., 63 (1984), pp. 283–292.

[6] R. W. Cottle, *Manifestations of the Schur complement*, Linear Algebra Appl., 8 (1974), pp. 189–211.

[7] J.-P. Crouzeix, J. A. Ferland, and S. Schaible, *Generalized convexity on affine subspaces with an application to potential functions*, Math. Programming, 56 (1992), pp. 223–232.

[8] P. Finsler, *Über das Vorkommen definiter und semidefiniter Formen und Scharen quadratischer Formen*, Commentarii Mathematici Helvetica, 9 (1937), pp. 188–192.

[9] R. Fletcher, *Practical Methods of Optimization*, 2nd ed., John Wiley and Sons, Chichester, UK, 1987.

[10] P. E. Gill, W. Murray, and M. H. Wright, *Practical Optimization*, Academic Press, London, New York, 1981.

[11] N. I. M. Gould, *On practical conditions for the existence and uniqueness of solutions to the general equality quadratic programming problem*, Math. Programming, 32 (1985), pp. 90–99.

[12] M. R. Hestenes, *Multiplier and gradient methods*, J. Optim. Theory Appl., 4 (1969), pp. 303–320.

[13] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, UK, 1985.

[14] J. H. Maddocks, *Restricted quadratic forms and their application to bifurcation and stability in constrained variational principles*, SIAM J. Math. Anal., 16 (1985), pp. 47–68. *Corrigendum*, SIAM J. Math. Anal, 19 (1988), pp. 1256–1257.
[15] J. H. Maddocks, *Restricted quadratic forms, inertia theorems, and the Schur complement*, Linear Algebra Appl., 108 (1988), pp. 1–36.
[16] J. H. Maddocks and R. L. Sachs, *On the stability of KdV multi-solitons*, Comm. Pure Appl. Math., 46 (1993), pp. 867–901.
[17] S. G. Nash and A. Sofer, *Linear and Nonlinear Programming*, McGraw-Hill, New York, 1996.
[18] M. J. D. Powell, *A method for nonlinear constraints in minimization problems*, in Optimization, R. Fletcher, ed., Academic Press, London, New York, 1969, pp. 283–298.
[19] L. Vandenberghe and S. Boyd, *Semidefinite programming*, SIAM Rev., 38 (1996), pp. 49–95.