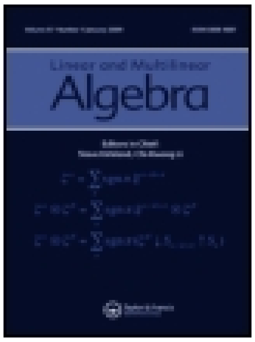

Linear and Multilinear Algebra

ISSN: 0308-1087 (Print) 1563-5139 (Online) Journal homepage: https://www.tandfonline.com/loi/glma20

A note on linearly dependent symmetric matrices

André Pierro de Camargo & Gabriel Haeser

LINEAR AND MULTILINEAR ALGEBRA
https://doi.org/10.1080/03081087.2019.1682495

Taylor & Francis
Taylor & Francis Group

Check for updates

A note on linearly dependent symmetric matrices

André Pierro de Camargo[a] and Gabriel Haeser[b]

[a]Center of Mathematics, Computer Science and Cognition, Federal University of the ABC region, Santo André-SP, Brazil; [b]Department of Applied Mathematics, University of São Paulo, São Paulo-SP, Brazil

ABSTRACT
In this note we present a sufficient condition for an upper bound on the rank of a set $\{A_1, A_2, \ldots, A_m\}$ of real $n \times n$ symmetric matrices to hold. The condition is based on the rank of a matrix defined by vectors in the image of each matrix, namely $A_1x, A_2x, \ldots, A_mx$, $x \in \mathbb{R}^n$ together with an auxiliary matrix $V \in \mathbb{R}^{k\times m}$. The result is inspired by some problems on second-order necessary optimality conditions for constrained optimization with quadratic constraints, and we explain and unify previously known results.

ARTICLE HISTORY
Received 22 June 2018
Accepted 15 October 2019

COMMUNICATED BY
Steve Kirkland

KEYWORDS
Symmetric matrix; linear dependence; second-order optimality conditions

AMS SUBJECT CLASSIFICATIONS
15B57; 15A03

1. Introduction

In this note we present an upper bound for the rank of a set $\{A_1, A_2, \ldots, A_m\}$ of real $n \times n$ symmetric matrices in terms of the maximum rank for $x \in \mathbb{R}^n$ of the $(n + k) \times m$ matrix below:

$$A(x, V) = \begin{bmatrix} A_1x & A_2x & \ldots & A_mx \\ v_{1,1} & v_{1,2} & \ldots & v_{1,m} \\ \vdots & \vdots & & \vdots \\ v_{k,1} & v_{k,2} & \ldots & v_{k,m} \end{bmatrix}, \tag{1}$$

where

$$V = \begin{bmatrix} v_1^T \\ \vdots \\ v_k^T \end{bmatrix} = (v_{ij}) \in \mathbb{R}^{k\times m}$$

is an auxiliary matrix. The rank of $\{A_1, A_2, \ldots, A_m\}$ is defined as the rank of the $n^2 \times m$ matrix

$$[\mathrm{vec}(A_1) \quad \cdots \quad \mathrm{vec}(A_m)],$$

where $\mathrm{vec}(A_i) \in \mathbb{R}^{n^2}$ is the column-wise vectorization of matrix $A_i, i = 1, \ldots, m$. Our main result is the following:

CONTACT André de Camargo andrecamargo.math@gmail.com

Theorem 1.1: *Let k be an integer with $0 \leq k \leq m$ and assume that $V \in \mathbb{R}^{k\times m}$ is such that every $k \times (k+2)$ submatrix of V is full rank. If the rank of $A(x, V)$ is at most $k+1$ for all $x \in \mathbb{R}^n$, then the rank of $\{A_1, \ldots, A_m\}$ is at most $k+1$.*

Our result generalizes and unifies two recent results in the optimization literature (see Section 3 for more information):

Theorem 1.2 (Haeser [1]): *If the matrix $A(x, V)$ has rank at most two for every $x \in \mathbb{R}^n$, where $V = [a, a, \ldots, a] \in \mathbb{R}^{1\times m}, a \neq 0$, then $\{A_1, A_2, \ldots, A_m\}$ has rank at most two.*

Theorem 1.3 (Mascarenhas [2]): *If the matrix $A(x, V)$ has rank at most one for every $x \in \mathbb{R}^n$, where V is an empty $0 \times m$ matrix, then $\{A_1, A_2, \ldots, A_m\}$ has rank at most one.*

The proofs of Theorems 1.2 and 1.3 are based (explicitly or implicitly) on the Spectral Theorem for real symmetric matrices, but the connections between them are not clear. In this note we put Theorems 1.2 and 1.3 under a simple general framework, in the sense that we prove a result for a general $k \times m$ matrix V with non-singular submatrices. In this case, when the rank of $A(x, V)$ is bounded from above by $k+1$, the same upper bound holds for the rank of $\{A_1, \ldots, A_m\}$. Note that since $\text{rank}(A(0, V)) = \text{rank}(V) = k$, we only allow an increase of 1 in the rank of $A(x, V)$ with respect to the rank of $A(0, V)$, similarly to Theorems 1.2 and 1.3.

Remark 1.1: Theorem 1.1 also holds for Hermitian matrices $A_1, \ldots, A_m$ and complex matrices V.

Remark 1.2: Theorem 1.1 may not hold if the number of lines of V, or its rank, is less than k. Notice, for example, that the subspace of symmetric 4×4 matrices of the form $\begin{bmatrix} M & 0 \\ 0 & 0 \end{bmatrix}$, where M is 3×3, has dimension 6, but, every basis $\{A_1, A_2, \ldots, A_6\}$ is such that

$$\left\{ \begin{bmatrix} A_1 x \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} A_2 x \\ 2 \\ 0 \\ 0 \end{bmatrix}, \ldots, \begin{bmatrix} A_6 x \\ 6 \\ 0 \\ 0 \end{bmatrix} \right\}$$

has rank at most 4 for every $x \in \mathbb{R}^4$.

Remark 1.3: The converse of Theorem 1.1 is not true. For instance, for $k = 1$,

$$A_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \text{and} \quad A_3 = \alpha A_1 + \beta A_2,$$

we have

$$A\left(\begin{bmatrix} x \\ y \end{bmatrix}, V \right) = \begin{bmatrix} x & y & \alpha x + \beta y \\ y & x & \beta x + \alpha y \\ v_{1,1} & v_{1,2} & v_{1,3} \end{bmatrix}$$

and this matrix has rank 3 if $x^2 + y^2 \neq 0$ and $v_{1,3} - \alpha v_{1,1} - \beta v_{1,2} \neq 0$. Therefore, the hypothesis of Theorem 1.1 in this case are only true for $(v_{1,1}, v_{1,2}, v_{1,3})$ in a set of null measure in the entire set of parameters $\{v_{1,1}, v_{1,2}, v_{1,3}\}$. This example suggests that the hypothesis of Theorem 1.1 are rarely fulfilled when $\{A_1, \dots, A_m\}$ has rank grater than one, but we emphasize that examples with $k > 1$ and rank $k+1$ do exist. In fact, for $k = 2$, take A_1, A_2, A_3 arbitrary symmetric matrices and $A_4 = A_1$. We have that

$$\begin{bmatrix} A_1x & A_2x & A_3x & A_4x \\ a & b & c & a \\ \alpha & \beta & \gamma & \alpha \end{bmatrix},$$

fulfills the hypothesis of Theorem 1.1 for almost all b, c, β, γ when $a \neq 0$ (or $\alpha \neq 0$).

In the next section we prove Theorem 1.1, and in Section 3 we discuss the connections of Theorem 1.1 to the field of quadratically constrained optimization.

2. Proof of Theorem 1.1

The conclusion of Theorem 1.1 is equivalent to stating that every subset $\{A_{i_1}, A_{i_2}, \dots, A_{i_{k+2}}\}$ formed by $k+2$ of the given matrices must be linearly dependent. Thus, it is sufficient to prove Theorem 1.1 for $m = k+2$.

Let $\{u_i = (u_{i,1}, u_{i,2}, \dots, u_{i,m}),\ i = 1, 2\}$ be a basis of the nullspace of V, that is,

$$\langle u_i, v_j \rangle = 0, \quad j = 1, 2, \dots, k,\ i = 1, 2, \tag{2}$$

where $\langle \cdot, \cdot \rangle$ denotes the usual inner product in $\mathbb{R}^m$. Let U_1, U_2 be the symmetric matrices defined by

$$U_i = \sum_{j=1}^{m} u_{i,j} A_j, \quad i = 1, 2. \tag{3}$$

Of course, we can assume that

$$A_i \neq 0, \quad i = 1, 2, \dots, m \quad \text{and} \quad U_i \neq 0, \quad i = 1, 2, \tag{4}$$

otherwise there is nothing to be proved.

We claim that there is a subset

$$\mathcal{B} = \{x_1, x_2, \dots, x_n\} \tag{5}$$

of $\mathbb{R}^n$ for which each of the following quantities is non-vanishing:

$$\begin{aligned} &\langle x_i, A_1 x_{i+1} \rangle,\ \langle x_i, A_2 x_{i+1} \rangle, \dots, \langle x_i, A_m x_{i+1} \rangle, \quad i = 1, 2, \dots, n-1, \\ &\langle x_i, U_1 x_{i+1} \rangle,\ \langle x_i, U_2 x_{i+1} \rangle, \quad i = 1, 2, \dots, n-1, \\ &\det[x_1 \quad x_2 \quad \dots \quad x_n]. \end{aligned} \tag{6}$$

Notice that a set fulfilling the above condition is actually a basis of $\mathbb{R}^n$. The existence of such a basis is a direct consequence of the following elementary result in measure theory:

Lemma 2.1 (Caron and Traynor [3]): *The set of zeroes of every non-vanishing polynomial in k real (or complex) variables has measure zero in* $\mathbb{R}^k$ ($\mathbb{C}^k$).

In fact, Lemma 2.1 ensures that the zeroes of the polynomials in $x_1, x_2, \ldots, x_n$ defined by each of the expressions in (6) has measure zero in $\mathbb{R}^{n^2}$ and this proves the existence of such a basis.

Now, consider the vectors

$$w_{i,j} = (\langle x_i, A_1 x_j\rangle,\ \langle x_i, A_2 x_j\rangle, \ldots, \langle x_i, A_m x_j\rangle), \quad i, j \in \{1, 2, \ldots, n\} \tag{7}$$

in $\mathbb{R}^m$, defined in terms of $A_1, A_2, \ldots, A_m$ and the basis $\mathcal{B}$ from (5). We claim that the vectors

$$w_{i,i+1}, v_1, v_2, \ldots, v_k \tag{8}$$

are linearly independent for all $i \in \{1, 2, \ldots, n-1\}$. In fact, by the definition of $\mathcal{B}$,

$$0 \neq \langle x_i, U_\tau x_{i+1}\rangle = \sum_{j=1}^{m} u_{\tau,j}\langle x_i, A_j x_{i+1}\rangle = \langle u_\tau, w_{i,i+1}\rangle, \quad \tau \in \{1, 2\},$$

that is, $w_{i,i+1}$ is not orthogonal to u_1, u_2 and, therefore, it is not a linear combination of $v_1, v_2, \ldots, v_k$.

Since for all x the rank of $A(x, V)$ is at most $k+1$, and $m = k+2$, for each $i \in \{1, 2, \ldots, n\}$, there is a non-vanishing vector $y_i \in \mathbb{R}^m$ such that

$$y_{i,1}A_1 x_i + y_{i,2}A_2 x_i + \cdots + y_{i,m}A_m x_i = 0 \tag{9}$$

and

$$\langle y_i, v_j\rangle = 0, \quad j = 1, 2, \ldots, k. \tag{10}$$

By (9), we obtain

$$\langle y_i, w_{i+1,i}\rangle = 0 \quad \text{and} \quad \langle y_{i+1}, w_{i,i+1}\rangle = 0.$$

Since $w := w_{i,i+1} = w_{i+1,i}$, the relations above and (10) tells us that each of the vectors y_i and y_{i+1} are orthogonal to all the vectors $w, v_1, v_2, \ldots, v_k$. It turns out that y_i and y_{i+1} must be scalar multiples of one another, because by (8) and the assumption $m = k+2$, the orthogonal complement of the space spanned by $w, v_1, v_2, \ldots, v_k$ is one-dimensional. Therefore one can choose $y_i = y_{i+1}$ for all $i = 1, 2, \ldots, n-1$, that is, there is a non-vanishing vector $\alpha = y_i,\ i = 1, 2, \ldots, n$ in $\mathbb{R}^m$ such that

$$\begin{aligned} 0 &= \alpha_1 A_1 x_i + \alpha_2 A_2 x_i + \cdots + \alpha_m A_m x_i \\ &= (\alpha_1 A_1 + \alpha_2 A_2 + \cdots + \alpha_m A_m)x_i, \quad i = 1, 2, \ldots, n, \end{aligned}$$

and this shows that $\alpha_1 A_1 + \alpha_2 A_2 + \cdots + \alpha_m A_m = 0$.

3. Connections with quadratically constrained optimization

Both Theorems 1.2 and 1.3 are used in the context of proving the validity of a second-order necessary optimality condition to the nonlinear optimization problem

$$\text{Minimize}\, f(x), \ \text{s.t.}\ g(x) \leq 0,$$

where $f : \mathbb{R}^n \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}^m$ are twice continuously differentiable functions. Throughout this discussion, $x^* \in \mathbb{R}^n$ will be a fixed local solution with $g(x^*) = 0$ and such that the linearized feasible set has a non-empty interior near x^*, in the sense that $J_g(x^*)d < 0$ for some $d \in \mathbb{R}^n$ (here $J_g(x) \in \mathbb{R}^{m\times n}$ is the Jacobian matrix of g at $x \in \mathbb{R}^n$). Under these assumptions, the following first-order necessary optimality condition is well known [4]: the set of Lagrange multipliers at x^*

$$\Lambda(x^*) := \{\lambda \in \mathbb{R}^m : \nabla_x L(x^*, \lambda) = 0, \lambda \geq 0\}$$

is a non-empty, compact, polyhedral set, where

$$(x, \lambda) \mapsto L(x, \lambda) = f(x) + \langle g(x), \lambda\rangle$$

is the usual Lagrangian function. The following second-order necessary optimality condition is also well known [4]:

$$\forall d \in C(x^*),\ \exists \lambda \in \Lambda(x^*) \text{ s.t. } d^T \nabla^2_{xx} L(x^*, \lambda) d \geq 0, \tag{11}$$

where ∇_x and ∇^2_{xx} denote the gradient vector and Hessian matrix with respect to the problem variables, respectively, and $C(x^*) := \{d \in \mathbb{R}^n : \nabla f(x^*)^T d \leq 0, J_g(x^*)d \leq 0\}$ is the critical cone.

In [1], a particular problem with quadratic constraints is considered: the function g is given componentwise by $g_i(x, z) = \frac{1}{2}x^T A_i x - z$, $i = 1, \ldots, m$ and $f(x, z) = z$, where $(x, z) \in \mathbb{R}^n \times \mathbb{R}$ are the problem variables. Note that $J_g(x, z)^T$ is precisely of the form (3), namely, $J_g(x, z)^T = A(x, V)$ where $k = 1$ and $V = [-1, -1, \ldots, -1] \in \mathbb{R}^{1\times m}$.

Hence, whenever $(x^*, z^*) = (0, 0)$ is a local solution, bounding the growth of the rank of $J_g(x, z)$ to at most one in a neighbourhood of the solution (note that $\operatorname{rank}(J_g(x^*, z^*)) = 1$) is sufficient, by Theorem 1.2, to have that $\{A_1, \ldots, A_m\}$ has rank at most two. This allows the application of a theorem of the alternative to arrive at a more informative second-order optimality condition, in the sense that it depends on a single Lagrange multiplier. That is, essentially, the low rank of $\{A_1, \ldots, A_m\}$ allows the exchange of the order of the quantifiers '$\forall d\, \exists\lambda$' to '$\exists\lambda\, \forall d$' in (11). More precisely, let $K \subset C(x^*, z^*)$ be any set defined by the direct sum of a subspace and a ray $\{\alpha v : \alpha \geq 0\}$ for some $v \in \mathbb{R}^n$. Then, it holds that

$$d^T \nabla^2_{xx} L((x^*, z^*), \lambda^*) d \geq 0, \quad \forall d \in K \tag{12}$$

for some fixed Lagrange multiplier $\lambda^* \in \Lambda(x^*, z^*)$, depending on K. The details can be found in [1] and an example is given below:

Let us consider the following optimization problem in the variables $(x, y, z) \in \mathbb{R}^3$:

$$\begin{aligned} \text{Minimize} \quad & z, \\ \text{s.t.} \quad & x^2 - 2xy + y^2 \leq z, \\ & -2x^2 + 2xy + y^2 \leq z, \\ & 4x^2 - 6xy + y^2 \leq z. \end{aligned}$$

A simple calculation shows that a solution is given by $(x^*, y^*, z^*) = (0, 0, 0)$. Computing the transposed Jacobian of the constraints, we arrive at

$$A(x, y, z, V) = \begin{pmatrix} 2x - 2y & -4x + 2y & 8x - 6y \\ -2x + 2y & 2x + 2y & -6x + 2y \\ -1 & -1 & -1 \end{pmatrix}.$$

The main result given by Theorem 1.2 is that by restricting the rank of $A(x, y, z, V)$ to be at most 2, one necessarily has that the scalars that attest the degeneracy of $A(x, y, z, V)$ can be taken independently of x, y and z. Here, it is clear that the third column is composed by twice the first one minus the second one for all x, y and z. This allows by an extension of Yuan's Lemma [5] to prove that there are Lagrange multipliers $(\lambda_1, \lambda_2, \lambda_3)$, namely, $(\lambda_1, \lambda_2, \lambda_3) = (0, \frac{3}{5}, \frac{2}{5})$, such that the second-order necessary optimality condition below holds:

$$\lambda_1(x^2 - 2xy + y^2) + \lambda_2(-2x^2 + 2xy + y^2) + \lambda_3(4x^2 - 6xy + y^2) \geq 0, \quad \forall (x, y) \in \mathbb{R}^2.$$

The paper [2] uses Theorem 1.3 in a similar fashion without restricting to the quadratic case, namely, the second-order optimality condition depending on a single Lagrange multiplier (12) holds whenever the rank of $J_g(x)$ is at most $\text{rank}(J_g(x^*)) + 1$, as long as the cone K is restricted to be the largest subspace contained in $C(x^*)$, that is, $K = \{d \in \mathbb{R}^n : J_g(x^*)d = 0\}$. This result solved a conjecture formulated in [6], that an increase of one to the rank of the Jacobian matrix would be sufficient to have a second-order necessary optimality condition to hold. This is clearly a stronger result than the usual ones where the Jacobian matrix is assumed to be of full rank [7,8] or of constant rank [9,10]. We refer the reader to [11] for more details on this topic.

Our result is connected with the more general quadratically constrained optimization problem:

$$\text{Minimize} \, f(x, z), \ \text{s.t.} \, g(x, z) \leq 0,$$

where $f : \mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R}$ is twice continuously differentiable and $g : \mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R}^m$ is defined where each g_i, $i = 1, \ldots, m$, is a homogeneous quadratic function with a separable linear term, that is,

$$g_i(x, z) = \tfrac{1}{2} x^T A_i x + w_i^T z.$$

Note that $J_g(x, z)^T = A(x, V)$, where $V = [w_1 \cdots w_m] \in \mathbb{R}^{k \times m}$. Let us assume that every $k \times (k + 2)$ submatrix of V is non-singular. Thus, when the rank of $J_g(x, z)$ increases at most by one with respect to the rank at $(x, z) := (0, 0)$, which is weaker than the usual constant or full rank assumptions, we arrive by Theorem 1.1 that $\{A_1, \ldots, A_m\}$ has rank at most $k + 1$.

Acknowledgments

The authors are grateful to an anonymous referee who carefully read previous versions of this paper and suggested several improvements to produce a much better final result.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

[1] Haeser G. An extension of Yuan's Lemma and its applications in optimization. J Optim Theory Appl. 2017;174(3):641–649.

[2] Mascarenhas WF. A simple canonical form for nonlinear programming problems and its use. J Optim Theory Appl. 2019;181(2):456–469.

[3] Caron RJ, Traynor T. The zero set of a polynomial. WMSR Report 05-03. Windsor, ON, Canada: Department of Mathematics and Statistics, University of Windsor; May 2005. Available from: http://www.uwindsor.ca/math/sites/uwindsor.ca.math/files/05-03.pdf

[4] Bonnans JF, Shapiro A. Pertubation analysis of optimization problems. New York (New York): Springer-Verlag; 2000.

[5] Yuan Y. On a subproblem of trust region algorithms for constrained optimization. Math Program. 1990;47:53–63.

[6] Andreani R, Martínez JM, Schuverdt ML. On second-order optimality conditions for nonlinear programming. Optimization. 2007;56:529–542.

[7] Baccari A, Trad A. On the classical necessary second-order optimality conditions in the presence of equality and inequality constraints. SIAM J Optim. 2005;15(2):394–408.

[8] Behling R, Haeser G, Ramos A, Viana DS. On a conjecture in second-order optimality conditions. J Optim Theory Appl. 2018;176(3):625–633.

[9] Andreani R, Echagüe CE, Schuverdt ML. Constant-rank condition and second-order constraint qualification. J Optim Theory Appl. 2010;146(2):255–266.

[10] Janin R. Directional derivative of the marginal function in nonlinear programming. Math Program Stud. 1984;21:127–138.

[11] Haeser G, Ramos A. On weak constraint qualifications with second-order properties in nonlinear optimization. Optimization online; 2018. Available from: http://www.optimization-online.org/DB_HTML/2018/01/6409.html