On Some Properties of Saddle Point Matrices with Vector Blocks

Tadeusz Ostrowski

Institute of Management
The State Vocational University in Gorzów Wlkp.
ul. Chopina 52, 66-400 Gorzów Wlkp., Poland
Tadostr@Interia.pl

Abstract. The paper is concerned on saddle point type problem, often called KKT system in optimization literature, which arises in a wide variety of applications. There are revisited and discussed some properties of saddle point matrices with two vector blocks, connected with the standard quadratic program over the standard simplex, if the matrix block is symmetric, or game theory, if the matrix block is not necessarily symmetric. There are investigated conditions of nonsingularity of these block matrices, and properties related to determinant, rank or inertia as well.

Mathematics Subject Classifications: 90C20, 15A63, 15A24

Keywords: Optimization, Standard Quadratic Program, Saddle Point Matrix

1. Introduction.

The aim of this paper is to revisit and discuss some properties of a block matrix, denoted here by \((A, e)\), and defined as follows:
\((A, e) = \begin{bmatrix} A & e \\ e^T & 0 \end{bmatrix}\),

in which the top-left block is a matrix \(A \in \mathbb{R}^{n \times n}\), the top-right block \(e\) is the column vector with all entries 1, the bottom-left block is \(e^\top\) transpose, and the bottom-right block has the single entry 0. If \(A\) is symmetric, then \((A, e)\) may be interpreted as the bordered Hessian of a standard quadratic program over the standard simplex, and it is called the Karush-Kuhn-Tucker matrix of the program, which is known to have a large spectrum of applications (for a review see, for instance, Bomze, 1998).

In the paper the matrix obtained from \(A\) by replacing the \(i\)-th row with \(e^\top\) is denoted by \(A_i\), and \(A_j\) stands for the matrix of \(A\) with the \(j\)-th column replaced by \(e\), and let \(M/A\) denotes the Schur complement of \(A\) in \(M\).

Consider the linear system of the form given as follows

\[
\begin{bmatrix}
A & e \\
e^T & 0
\end{bmatrix}
\begin{bmatrix}
x \\
-\lambda
\end{bmatrix}
=
\begin{bmatrix}
0 \\
1
\end{bmatrix}
\]

This system of equations arises as the first-order optimality conditions for standard quadratic program over the standard simplex in the \(n\)-dimensional Euclidean space \(\mathbb{R}^n\):

\[
\text{optimize } \{ F(x) = \frac{1}{2} x^T A x \mid x \in \Delta \} \quad (2)
\]

\[
\Delta = \{ x \in \mathbb{R}^n : e^\top x = 1, x_i \geq 0; i = 1, \ldots, n \}
\]

In this case the variable \(\lambda\) represents Lagrange multiplier and any solution \((x^*, \lambda^*)\) of the problem \((2)\) is a saddle point for the Lagrangean

\[
L(x, \lambda) = \frac{1}{2} x^T A x + \lambda (1 - e^\top x) \quad (3)
\]

i.e. \((x^*, \lambda^*)\) is a point that for any \(x \in \mathbb{R}^n, \lambda \in \mathbb{R}\) satisfies

\[
L(x^*, \lambda) \leq L(x^*, \lambda^*) \leq L(x, \lambda^*)
\]

or, equivalently

\[
\min_x \max_\lambda L(x, \lambda) = L(x^*, \lambda^*) = \max_\lambda \min_x L(x, \lambda).
\]

As it is well-known the first order optimality conditions require that the Lagrangean gradients of \((3)\) vanish:
\[ \begin{bmatrix} \nabla_x L(x, \lambda) \\ \nabla_\lambda L(x, \lambda) \end{bmatrix} = \begin{bmatrix} Ax - \lambda e \\ 1 - e^T x \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (4) \]

where \( e^T \) stands now for the Jacobian matrix of the constraint, and the systems (1) and (4) are equivalent.

2. Properties of \((A, e)\)

The following properties of the block matrix \((A, e)\) can be established, especially whenever it contains a linear combination of \(A\) and \(E = ee^T\). In the whole paper we always assume that \(A, E \in \mathbb{R}^{n \times n}\) and \(n > 1\).

Proposition 1. For any matrix \(A\) the determinant of \((A, e)\) is equal to the sum of all cofactors of \(A\) taken with opposite sign:

\[
\det(A, e) = -\sum_{j=1}^{n} \det A_j^j = -\sum_{i=1}^{n} \det A_i \quad (5)
\]

Proof. To get the first part of these equalities, compute the determinant of \((A, e)\) using Laplace expansion about the last row and next change the order of the rows in the resulting determinants:

\[
\det(A, e) = \det \begin{bmatrix} A_{11} & \ldots & a_{1,k-1} & a_{1,k+1} & \ldots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & \ldots & a_{n,k-1} & a_{n,k+1} & \ldots & a_{nn} \\ 0 & \ldots & 0 & \ldots & 1 \end{bmatrix} = \\
= \sum_{k=1}^{n} (-1)^{n+k} \det \begin{bmatrix} a_{11} & \ldots & a_{1,k-1} & a_{1,k+1} & \ldots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & \ldots & a_{n,k-1} & a_{n,k+1} & \ldots & a_{nn} \end{bmatrix} = \\
= \sum_{k=1}^{n} (-1)^{n+k} (-1)^{n+1-k} \det A^k = -\sum_{k=1}^{n} \det A^k.
\]

The second part can be obtained similarly by expanding \(\det(A, e)\) about the last column and change the order of the columns. ■
Proposition 2. For arbitrary real numbers $\alpha, \beta$ and vectors $v \in \mathbb{R}^n$ the determinant of the saddle point matrix has the following property:

\[
\begin{align*}
det(\alpha A + \beta E, e) &= \alpha^{n-1} \cdot det(A, e) \quad (6a) \\
det(\alpha A + \beta ev^T, e) &= \alpha^{n-1} \cdot det(A, e) \quad (6b) \\
det(\alpha A + \beta ev^T, e) &= \alpha^{n-1} \cdot det(A, e) \quad (6c)
\end{align*}
\]

Proof. To prove (6a) assume that $\alpha \neq 0$; otherwise the determinant on the left side is obviously zero and the proof of (6a) is complete. Using the elementary operations of subtracting the last row multiplied by $\beta$ from each of the other rows, we obtain

\[
\begin{align*}
det(\alpha A + \beta E, e) &= \det \left[ \begin{array}{cc}
\alpha A + \beta E & e \\
e^T & 0
\end{array} \right] = \det \left[ \begin{array}{c}
\alpha A \\
e^T
\end{array} \right]
\end{align*}
\]

Therefore, from multilinearity of the determinant with respect to the rows and columns, respectively, follows

\[
\begin{align*}
det(\alpha A + \beta E, e) &= \alpha^n \det \left[ \begin{array}{c}
A \\
e^T
\end{array} \right] = \alpha^{n-1} \det \left[ \begin{array}{c}
A \\
e^T
\end{array} \right] = \alpha^{n-1} \det(A, e).
\end{align*}
\]

Proofs of the equalities (6a) and (6b) go in similar way. ■

Proposition 3. For any real numbers $\alpha, \beta$ the determinant of the matrix $\alpha A + \beta E$ can be expressed as follows:

\[
\begin{align*}
det(\alpha A + \beta E) &= \alpha^{n-1} [\alpha \cdot \det A - \beta \cdot \det(A, e)] \quad (7)
\end{align*}
\]

Proof. We show at first that for any square matrix $B \in \mathbb{R}^{n \times n}$ the particular equality is true, namely

\[
\begin{align*}
det(B + \beta E) &= \det B - \beta \cdot \det(B, e), \quad (8)
\end{align*}
\]

which trivially holds for $\beta = 0$.

The derivative of $\det(B + \beta E)$ with respect to $\beta$ gives

\[
\frac{d[\det(B + \beta E)]}{d\beta} = \sum_{i=1}^{n} \det(B + \beta E)_i = \sum_{i=1}^{n} \det B_i = -\det(B, e).
\]

Since then, we have
Saddle Point Matrices

\[ \det(B + \beta E) = -\int \det(B, e) \, d\beta = -\beta \det(B, e) + C. \]

Putting \( \beta = 0 \), we get \( C = \det(B) \), which gives (8).

The general case is now obtained from (8) by replacing \( B \) with \( \alpha A \):

\[ \det(\alpha B + \beta E) = \det(\alpha A) - \beta \det(\alpha A, e) = \]
\[ = \alpha^n \det(A) - \beta \alpha^{n-1} \det(A, e) = \]
\[ = \alpha^{n-1}[\alpha \det(A) - \beta \det(A, e)]. \]

\[ \blacksquare \]

**Proposition 4.** The determinant of \((A, e)\) can be expressed by the determinants of \(A\) and \(A + E\) as follows:

\[ \det(A, e) = \det(A) - \det(A + E) \tag{9} \]

**Proof.** From linearity of the determinant with respect to the last column and the elementary operations of subtracting the last row in the second determinant from each of the other rows, we obtain

\[ \det(A, e) = \det \begin{bmatrix} A & e \\ e^T & 0 \end{bmatrix} = \det \begin{bmatrix} A & 0 \\ e^T & 1 \end{bmatrix} + \det \begin{bmatrix} A & e \\ e^T & -1 \end{bmatrix} = \]
\[ = \det(A) - \det \begin{bmatrix} A & -e \\ e^T & 1 \end{bmatrix} = \det(A) - \det \begin{bmatrix} A + E & 0 \\ e^T & 1 \end{bmatrix} = \det(A) - \det(A + E). \]

It also can be easily seen that equation (9) follows immediately from (7) by substitution \( \alpha = \beta = 1 \).

\[ \blacksquare \]

**Proposition 5.** If \( A \) is nonsingular, then the following equalities hold true:

\[ \det(A, e) = -e^T A^{-1} e \cdot \det(A) = (A, e) / A \cdot \det(A) \tag{10} \]

**Proof.** Equations above follow at once from Banachiewicz inversion formula and from Schur determinant formula or by a well-known formula for the determinant of a block matrix (Gantmacher, 2000). Obviously in considered case

\[ (A, e) / A = -e^T A^{-1} e \]

the Schur complement of \( A \) in \((A, e)\) is the single entry.

\[ \blacksquare \]
Note that nonsingularity of $A$ says nothing about singularity or nonsingularity of the saddle point matrix $(A, e)$ is nonsingular as well. For example if

$$A = \begin{bmatrix} 3 & 2 \\ 2 & 1 \end{bmatrix}$$

then we have

$$e^TA^{-1}e = [1, 1] \begin{bmatrix} 1 & -2 \\ -2 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 0.$$ 

On the other side, for

$$A = \begin{bmatrix} 5 & 2 \\ 2 & 1 \end{bmatrix}$$

we obtain

$$e^TA^{-1}e = [1, 1] \begin{bmatrix} 1 & -2 \\ -2 & 5 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 2.$$ 

**Proposition 6.** If $A$ is nonsingular and for some $\alpha$, $\beta$ the linear combination $\alpha A + \beta E$ is also nonsingular, then

$$\frac{(\alpha A + \beta E, e)}{\alpha - \beta (A, e)} = \frac{(A, e)}{\alpha - \beta (A, e)}$$

*Proof.* From (6a), (7) and Schur formula we obtain

$$\det(\alpha A + \beta E, e) = \alpha^{n-1} \det(A, e) = \alpha^{n-1} \det(\alpha A + \beta E).$$

On the other side, by (10), (7) and (10) again, we have

$$\det(\alpha A + \beta E, e) = (\alpha A + \beta E, e)/(\alpha A + \beta E) \det(\alpha A + \beta E) =$$

$$= (\alpha A + \beta E, e)/(\alpha A + \beta E) \alpha^{n-1} [\alpha \det A - \beta \det(A, e)] =$$

$$= (\alpha A + \beta E, e)/(\alpha A + \beta E) \alpha^{n-1} [\alpha \det A - \beta (A, e)/A \det A],$$

and by comparing both sides we obtain (11). \hfill \blacksquare

**Proposition 7.** If $A$ is symmetric and nonsingular, then

$$(A, e) \sim \begin{bmatrix} A & 0 \\ 0 & -e^T A^{-1} e \end{bmatrix}$$

where, by (10), the determinants of both congruent matrices are the same.
If both $A$ and $(A, e)$ are nonsingular, then

\[
(A, e)^{-1} = \frac{1}{e^T A^{-1} e} \begin{bmatrix}
    e^T A^{-1} e & -A^{-1} e A^{-1} \\
    -A^{-1} e^T A^{-1} & 1
\end{bmatrix}
\]  

(13)

**Proof.** The congruency follows from the decomposition below (see also Higham and Cheng, 1998; Tian and Takane, 2005)

\[
\begin{bmatrix}
    A & e \\
    e^T & 0
\end{bmatrix} = \begin{bmatrix}
    I & 0 \\
    e^T A^{-1} & 1
\end{bmatrix} \begin{bmatrix}
    A & e \\
    0 & -e^T A^{-1} e
\end{bmatrix} \begin{bmatrix}
    I & A^{-1} e \\
    0 & 1
\end{bmatrix}
\]  

(14)

Assume now that $A$, $(A, e)$ are nonsingular. Therefore we have

\[
(A, e)^{-1} = \begin{bmatrix}
    I & A^{-1} e \\
    0 & 1
\end{bmatrix} \begin{bmatrix}
    A & 0 \\
    0 & -e^T A^{-1} e
\end{bmatrix}^{-1} \begin{bmatrix}
    I & 0 \\
    e^T A^{-1} & 1
\end{bmatrix} = \begin{bmatrix}
    I & -A^{-1} e \\
    0 & 1
\end{bmatrix} \begin{bmatrix}
    A^{-1} & 0 \\
    0 & (-e^T A^{-1} e)^{-1}
\end{bmatrix} \begin{bmatrix}
    I & 0 \\
    -e^T A^{-1} & 1
\end{bmatrix} = \frac{1}{e^T A^{-1} e} \begin{bmatrix}
    e^T A^{-1} e A^{-1} - A^{-1} e A^{-1} & A^{-1} e \\
    e^T A^{-1} & -1
\end{bmatrix}.
\]

\[\blacksquare\]

**Proposition 8.** If the saddle point matrix $(A, e)$ is nonsingular, then the following holds:

\[
\text{rank}(A, e) = n + 1 \implies \text{rank} \begin{bmatrix}
    A \\
    e^T
\end{bmatrix} = n
\]

(15)

**Proof by contradiction.** Assume that \(\text{rank} \begin{bmatrix}
    A \\
    e^T
\end{bmatrix} < n\). In that case there exists a nonzero vector \(x \in \mathbb{R}^n\) such that

\[
\begin{bmatrix}
    A \\
    e^T
\end{bmatrix} x = \begin{bmatrix}
    Ax \\
    e^T x
\end{bmatrix} = \begin{bmatrix}
    0 \\
    0
\end{bmatrix}.
\]

Therefore putting

\[
v = \begin{bmatrix}
    x \\
    0
\end{bmatrix} \in \mathbb{R}^{n+1}
\]

we get since then

\[
(A, e)v = \begin{bmatrix}
    A & e \\
    e^T & 0
\end{bmatrix} \begin{bmatrix}
    x \\
    0
\end{bmatrix} = \begin{bmatrix}
    Ax \\
    e^T x
\end{bmatrix} = \begin{bmatrix}
    0 \\
    0
\end{bmatrix}, \text{ which is a contradiction.} \quad \blacksquare
\]
3. Concluding remarks.

If det(A, e) ≠ 0, then the rank of A is either n or n - 1:

(rank(A, e) = n + 1)  =>  (n - 1 ≤ rank A ≤ n)  \hspace{1cm} (16)

The converse is not true. However, if A is both symmetric and positive definite, then it immediately follows from (10) that det(A, e) ≠ 0 (see also Benzi, Golub and Liesen, 2005). The following three examples give some illustrations of (16)

1. \[ A = \begin{bmatrix} a & b \\ ca & cb \end{bmatrix} \text{ (} a \neq b, c \neq 1\text{): rank}(A, e) = 3 \text{ and rank} A = 1 \]

2. \[ A = \begin{bmatrix} a & a \\ a & a \end{bmatrix} \text{ (} a \neq 0\text{): rank}(A, e) = 2 \text{ and rank} A = 1 \]

3. \[ A = \begin{bmatrix} a & 0 \\ 3a & 2a \end{bmatrix} \text{ (} a \neq 0\text{): rank}(A, e) = \text{rank} A = 2. \]

Moreover, positive definite (A, e) has one negative eigenvalue and n positive ones (Boyd and Vandenberghe, 2004). So it can be written

\[ A > 0 \Rightarrow \det(A, e) \neq 0 \text{ and Inertia}(A, e) = (n, 1, 0) \hspace{1cm} (17) \]

The explanation of (17) is that if the matrix A is symmetric, then the block matrix (A, e), being a Hessian at a saddle point of an appropriate Lagrange function, is also symmetric but inevitably indefinite and it must have at least one positive and one negative eigenvalue (see Horn and Johnson, 1991; Gould, 1995; Griewank and Walther, 2002; Forsgren, 2002).

Equality (7) implies, that if we put \( \alpha = \det(A, e) \), \( \beta = \det A \), then a matrix given as a linear combination of the form

\[ \det(A, e) \cdot A + \det A \cdot E \]

is singular for any matrix A.

There are two more factorizations of the saddle point matrix (A, e):

\[ (A, e) = \begin{bmatrix} I & A^{-1}e \\ e & -e^T A^{-1} e \end{bmatrix} \begin{bmatrix} A & 0 \\ e & -e^T A^{-1} e \end{bmatrix} \hspace{1cm} (18) \]

\[ (A, e) = \begin{bmatrix} I & 0 \\ e^T A^{-1} & 1 \end{bmatrix} \begin{bmatrix} A & e \\ 0 & -e^T A^{-1} e \end{bmatrix} \hspace{1cm} (19) \]
All decompositions given in (14), (18) and (19) are equivalent and they also imply that if $A$ is nonsingular then the following equivalence (see also Benzi, Golub and Liesen, 2005) holds:

$$\det(A, e) \neq 0 \iff e^T A^{-1} e \neq 0$$

The assumption that the matrix $A$ is nonsingular may appear to be quite restrictive, since it is known that singular $A$ appears in many applications (see for instance Haber and Ascher, 2001).

However if $A \in \mathbb{R}^{n \times n}$ is singular and rank $A = n-1$ then one can replace $A$ by nonsingular matrix $\alpha A + \beta E$ for some $\alpha, \beta \in \mathbb{R}$, and which does not change the optimizer (see also Bomze and de Klerk, 2002). Furthermore, letting $\alpha = 1, \beta = -\lambda$ one can get the matrix associated with so called dual formulation of copositive program (for details see Bomze et al., 2000).

Otherwise, one can use augmented Lagrangean techniques to replace (1) with an equivalent system in which the left-up block is nonsingular and has the same solution (see Golub and Greif, 2003).

References


Received: August 7, 2006