

Linear Programming with Matrix Variables

B. D. Craven

*Mathematics Department
University of Melbourne
Parkville, Victoria 3052, Australia*
and

B. Mond

*Department of Pure Mathematics
La Trobe University
Bundoora, Victoria 3083, Australia*

Submitted by Hans Schneider

ABSTRACT

Linear programming is formulated with the vector variable replaced by a matrix variable, with the inner product defined using trace of a matrix. The theorems of Motzkin, Farkas (both homogeneous and inhomogeneous forms), and linear-programming duality thus extend to matrix variables. Duality theorems for linear programming over complex spaces, and over quaternion spaces, follow as special cases.

1. INTRODUCTION

Denote by $\mathbb{R}^{m \times r}$ the vector space of all real $m \times r$ matrices, equipped with the usual inner product: $\langle X, Y \rangle = \text{Tr}(X^T Y)$, where Tr denotes the *trace* of a square matrix. Various theorems of mathematical programming, including the Farkas and Motzkin theorems and linear-programming duality, extend to problems where the space \mathbb{R}^n of variables is replaced by $\mathbb{R}^{m \times r}$. Results for programming over complex, or quaternion, spaces follow as special cases.

Any linear functional L on $\mathbf{Y} = \mathbb{R}^{m \times r}$ can be represented by some $H \in \mathbf{Y}$, by

$$(\forall Y \in \mathbf{Y}) \quad L(Y) = \sum_{i=1}^r \sum_{j=1}^m H_{ij} Y_{ij} = \text{Tr}(H^T Y). \quad (1)$$

The isomorphism $\theta: \mathbb{R}^{m \times r} \rightarrow \mathbb{R}^{mr}$ defined by $\theta(Y) = (Y_{11}, \dots, Y_{1r}, Y_{21}, \dots, Y_{mr})$ preserves the trace; thus $\text{Tr}(H^T Y) = (\theta H)^T (\theta Y)$. Correspondingly [10], the inner product $\langle \theta(H), \theta(Y) \rangle = \text{Tr}(H^T Y)$ and $(AXB) = (A \otimes B^T)\theta(X)$, where \otimes denotes Kronecker product. Let S be a closed convex cone in $Y = \mathbb{R}^{m \times r}$. An inequality $AX - K \in S$, where $X \in \mathbb{R}^{n \times r}$, $A \in \mathbb{R}^{m \times n}$, $K \in \mathbb{R}^{m \times r}$, can be expressed as $A^\# \theta(X) - \theta(K) \in \theta(S)$, where $A^\# \in \mathbb{R}^{mr \times nr}$ has components $A_{ij, ut}^\# = \delta_{uj} A_{it}$, where δ_{uj} is the Kronecker delta, the indices i, j label rows of $A^\#$, and t, u label columns of $A^\#$. The (positive) *dual cone* of S is the set S^* of all linear functionals P on Y such that $P(Y) \geq 0$ for each $Y \in S$. It follows from (1) that

$$S^* = \{P \in Y: (\forall Y \in S) \text{Tr}(P^T Y) \geq 0\}. \quad (2)$$

In what follows, I_r (or I) denotes the unit matrix in $\mathbb{R}^{r \times r}$; $\mathbb{R}_+ = [0, \infty)$; symbols $\mathbf{X}, \mathbf{S}, \dots$ denote sets (vector spaces or cones), whereas X, S, \dots denote matrices. Let $\mathbf{U} = \{\lambda I_r: \lambda \in \mathbb{R}_+\}$; then $\mathbf{U}^* = \{Z \in \mathbb{R}^{r \times r}: \text{Tr}(Z) \geq 0\}$. Also let $\mathbf{U}_0 = \{\lambda I_r: \lambda \in \mathbb{R}\}$ and $\mathbf{V}_0 = \{\lambda I_n: \lambda \in \mathbb{R}\}$. The orthogonal complement of a subspace X is denoted by X^\perp . Let $\mathbf{V} = \{\lambda I_s: \lambda \in \mathbb{R}_+\}$.

The *Motzkin alternative theorem* for cone inequalities in finite-dimensional spaces (see [4, Lemma 2] or [5, Theorem 2.5.2]) translates by isomorphism θ to the following theorem for matrix variables.

THEOREM 1. *Let $A \in \mathbb{R}^{m \times n}$, $B \in \mathbb{R}^{q \times n}$, $C \in \mathbb{R}^{r \times s}$. Let $\mathbf{S} \subset \mathbb{R}^{m \times r}$ be a convex cone with interior; let $\mathbf{T} \subset \mathbb{R}^{q \times r}$ and $\mathbf{H} \subset \mathbb{R}^{n \times s}$ be closed convex cones; let the cones $B^T(\mathbf{T}^*)$ and $\mathbf{H}^* C^T \equiv \{WC^T: W \in \mathbf{H}^*\}$ be closed. Then exactly one of the following systems has a solution (X or (P, Q, R) , respectively):*

$$(\exists X) \quad AX \in \text{int } \mathbf{S}, \quad BX \in \mathbf{T}, \quad XC \in \mathbf{H}; \quad (3)$$

$$(\exists Q \in \mathbf{T}^*, R \in \mathbf{H}^*, 0 \neq P \in \mathbf{S}^*) \quad P^T A + Q^T B + CR^T = 0. \quad (4)$$

A direct corollary is the following matrix version of Farkas's theorem.

THEOREM 2. *Let \mathbf{X} be a subspace of $\mathbb{R}^{n \times r}$; let $A \in \mathbb{R}^{m \times n}$, $M \in \mathbb{R}^{r \times s}$, $B_0 \in \mathbb{R}^{n \times r}$; let $\mathbf{K} \subset \mathbb{R}^{m \times r}$ and $\mathbf{L} \subset \mathbb{R}^{n \times s}$ be convex cones, such that the cones $A^T(\mathbf{K}^*) + \mathbf{X}^\perp$ and $\mathbf{L}^* M^T + \mathbf{X}^\perp$ are closed in $\mathbb{R}^{n \times r}$. Then*

$$\begin{aligned} & [(AX \in \mathbf{K}, XM \in \mathbf{L}, X \in \mathbf{X}) \Rightarrow \text{Tr}(B_0^T X) \geq 0] \\ & \Leftrightarrow [(\exists C \in \mathbf{K}^*, D \in \mathbf{L}^*, N \in \mathbf{X}^\perp) (B_0 + N)^T = C^T A + MD^T]. \end{aligned} \quad (5)$$

The closed-cone hypothesis is fulfilled if $A^T(\mathbf{K}^)$ and $\mathbf{L}^* M^T$ are closed in \mathbf{X} .*

The cones in Theorems 1 and 2 need *not* be polyhedral. A more restrictive sufficient hypothesis [12] is that $AX_0 \in \text{int} \mathbf{K}$ and $X_0 M \in \text{int} L$ for some $X_0 \in \mathbf{X}$. Ben-Israel has given another matrix Farkas theorem [2]:

$$[AXB=C, X \geq 0] \Leftrightarrow [A^T U B^T \geq 0, \text{Tr}(U^T C) \geq 0], \quad (6)$$

in which the cone considered is, however, a nonnegative orthant, and the subspace \mathbf{X} is the whole space. The system (5) (with $N=0$) is obtainable from (6) by the replacements

$$A^T \rightarrow \begin{bmatrix} A \\ I_n \end{bmatrix}, \quad B^T \rightarrow [I_r, M], \quad C \rightarrow B. \quad (7)$$

Another version of Farkas's theorem, involving the trace inner product, was given in [6], Theorem 11.

The study of minimization problems with matrix variables was suggested to the authors by the matrix quadratic programs in [8] and [11].

2. NONHOMOGENEOUS FARKAS THEOREM FOR MATRIX SPACES

THEOREM 3. *Let $A \in \mathbb{R}^{m \times n}$, $D \in \mathbb{R}^{m \times r}$, $B \in \mathbb{R}^{n \times r}$, $c \in \mathbb{R}$. Let $\mathbf{K} \subset \mathbb{R}^{m \times r}$ be a closed polyhedral cone. Assume that $AX_0 - D \in \mathbf{K}$ for some $X_0 \in \mathbb{R}^{n \times r}$. Then*

$$\begin{aligned} [AX - D \in \mathbf{K} \Rightarrow \text{Tr}(B^T X) \geq c \quad (X \in \mathbb{R}^{n \times r})] \\ \Leftrightarrow [(\exists C \in \mathbf{K}^*) B^T = C^T A, \text{Tr}(D^T C) \geq c]. \quad (8) \end{aligned}$$

Proof. The system

$$\begin{bmatrix} A & -D \\ 0 & I_r \end{bmatrix} \begin{bmatrix} Y \\ \Lambda \end{bmatrix} \in \begin{bmatrix} \mathbf{K} \\ \mathbf{U} \end{bmatrix} \Rightarrow \text{Tr} \left\{ [B^T, -c r^{-1} I_r] \begin{bmatrix} Y \\ \Lambda \end{bmatrix} \right\} \geq 0 \quad (9)$$

is equivalent, by Theorem 2, to the system

$$(\exists C \in \mathbf{K}^*, N \in \mathbf{U}^*) [B^T, -c r^{-1} I_r] = [C^T, N^T] \begin{bmatrix} A & -D \\ 0 & I_r \end{bmatrix}. \quad (10)$$

Substituting for U^* shows that (10) is equivalent to $B^T = C^T A$, $D^T C - cr^{-1} I_r \in U^*$, for some $C \in \mathbf{K}^*$. Hence (9) is equivalent to the right side of (8).

Suppose that the right side of (8) does not hold. Then (9) does not hold; so there exist Y' and Λ' satisfying $AY' - D\Lambda' \in \mathbf{K}$, $\Lambda' \in U$, $\text{Tr}(B^T Y' - cr^{-1} \Lambda') < 0$, and $\Lambda' = \beta I_r$ for some $\beta \geq 0$. If $\beta > 0$, then $X = \beta^{-1} Y'$ satisfies $AX - D \in \mathbf{K}$, $\text{Tr}(B^T X) < c$. If $\beta = 0$, then $AY' \in \mathbf{K}$, $\text{Tr}(B^T Y') < 0$. For $\alpha > 0$,

$$A(X_0 + \alpha Y') - D = (AX_0 - D) + \alpha AY' \in \mathbf{K} + \alpha \mathbf{K} \subset \mathbf{K}; \quad (11)$$

but

$$\text{Tr}\{B^T(X_0 + \alpha Y')\} = \text{Tr}(B^T X_0) + \alpha \text{Tr}(B^T Y') < c$$

for sufficiently large $\alpha > 0$, since $\text{Tr}(B^T Y') < 0$. In either case, the left side of (8) is contradicted. Hence the left side of (8) implies the right side.

Conversely, let the right side of (8) hold. If $AX - D \in \mathbf{K}$, then

$$\text{Tr}(B^T X) = \text{Tr}(C^T A X) \geq \text{Tr}(C^T D) = \text{Tr}(D^T C) \geq c;$$

hence the left side of (8) holds. ■

REMARK. If K is a closed convex cone, not polyhedral, then the result of Theorem 3 remains valid, provided that the convex cone

$$\left[\begin{array}{cc} A^T & 0 \\ -D^T & I_r \end{array} \right] \left[\begin{array}{c} \mathbf{K}^* \\ \mathbf{U}^* \end{array} \right] \quad (12)$$

is assumed closed. Bellman and Fan [1] have given a nonhomogeneous Farkas theorem for positive semidefinite Hermitian matrices only.

4. LINEAR PROGRAMMING FOR MATRIX VARIABLES

Let X be a subspace of $\mathbb{R}^{n \times r}$, $C \in \mathbb{R}^{n \times r}$, $A \in \mathbb{R}^{m \times r}$, $J \in \mathbb{R}^{n \times s}$, $B \in \mathbb{R}^{m \times r}$, $M \in \mathbb{R}^{r \times s}$. Let $\mathbf{K} \subset \mathbb{R}^{m \times r}$ and $\mathbf{L} \subset \mathbb{R}^{n \times s}$ be closed convex cones, not necessarily polyhedral. Consider the pair of programming problems:

$$\text{Minimize}_{X \in X} \text{Tr}(C^T X) \quad \text{subject to} \quad AX - B \in \mathbf{K}, \quad XM - J \in \mathbf{L}; \quad (13)$$

$$\text{Maximize}_{Z, W} \text{Tr}(B^T Z) + \text{Tr}(JW^T)$$

$$\text{subject to} \quad Z \in \mathbf{K}^*, \quad W \in \mathbf{L}^*, \quad C^T = Z^T A + MW^T. \quad (14)$$

The problems (13) and (14) have linear objective functions, and linear constraints with convex cones. The unconventional restriction of X to a subspace has later application. Linear programs with matrix variables are applicable when a conventional linear program has nr variables, which may relevantly be arranged as a $n \times r$ matrix. For example, in a transportation problem, the constraints are $e_n^T X = B$, $X e_r = J$, $X \geq 0$, where e_r, e_n are columns of ones, and (13) extends this to more general weighting matrices. The following theorem proves duality for (13) and (14).

THEOREM 4. *Let (13) reach a minimum at $X = \bar{X}$; let $Q = A\bar{X} - B$ and $R = \bar{X}M - J$; assume that the cones*

$$\begin{bmatrix} A & Q \\ 0 & I_r \end{bmatrix}^T \begin{bmatrix} \mathbf{K}^* \\ \mathbf{U}^* \end{bmatrix} \quad \text{and} \quad \left\{ \begin{bmatrix} G \\ H \end{bmatrix} \begin{bmatrix} M & 0 \\ R & I_s \end{bmatrix}^T : \begin{bmatrix} G \\ H \end{bmatrix} \in \begin{bmatrix} \mathbf{L}^* \\ \mathbf{V}^* \end{bmatrix} \right\} \quad (15)$$

are closed (in particular, K and L may be polyhedral). Then (14) reaches a maximum at some $Z = \bar{Z}$, $W = \bar{W}$, and

$$\begin{aligned} \text{Tr}(C^T \bar{X}) &= \text{Tr}(B^T \bar{Z}) + \text{Tr}(J \bar{W}^T); \\ C^T - \bar{Z}^T A + M \bar{W}^T &\in \mathbf{X}^\perp, \quad \text{Tr}(\bar{Z}^T (A\bar{X} - B)) = 0, \\ \text{Tr}((\bar{X}M - J)\bar{W}^T) &= 0. \end{aligned} \quad (16)$$

Also $\text{Tr}(C^T X) \geq \text{Tr}(B^T Z) + \text{Tr}(JW^T)$ whenever X is feasible for (13), and Z, W for (14).

Proof. Let X and Z, W satisfy the constraints of (13) and (14), respectively. Then

$$\begin{aligned} \text{Tr}(C^T X) - [\text{Tr}(B^T Z) + \text{Tr}(JW^T)] &= \text{Tr}[Z^T (AX - B)] \\ &\quad + \text{Tr}[(XM - J)W^T] \geq 0. \end{aligned} \quad (17)$$

Now let $Y \in X$, $\Lambda = \lambda I_r \in \mathbf{U}_0$, $\Xi = \xi I_n \in \mathbf{V}_0$ satisfy

$$\begin{aligned} \begin{bmatrix} A & Q \\ 0 & I_r \end{bmatrix} \begin{bmatrix} Y & \Xi \\ \Lambda & 0 \end{bmatrix} &\in \begin{bmatrix} \mathbf{K} & \mathbb{R}^{m \times n} \\ \mathbf{U}_0 & \mathbb{R}^{r \times n} \end{bmatrix}, \\ \begin{bmatrix} Y & \Xi \\ \Lambda & 0 \end{bmatrix} \begin{bmatrix} M & 0 \\ R & I_n \end{bmatrix} &\in \begin{bmatrix} \mathbf{L} & \mathbf{V}_0 \\ \mathbb{R}^{r \times s} & \mathbb{R}^{r \times n} \end{bmatrix}. \end{aligned} \quad (18)$$

If $Y=0$, then $\text{Tr}(C^T Y) \geq 0$. If $Y \neq 0$, then for some $K \in \mathbf{K}$ and $L \in \mathbf{L}$,

$$A(\bar{X} + \alpha Y) - B = (1 - \alpha\lambda)Q + \alpha K \in \mathbf{K},$$

$$(\bar{X} + \alpha Y)M - J = (1 - \alpha\xi)R + \alpha L \in \mathbf{L}$$

whenever $\alpha \in \mathbb{R}_+$ is sufficiently small. Since \bar{X} minimizes (13), $\text{Tr}[C^T(\bar{X} + Y)] \geq \text{Tr}(C^T \bar{X})$. So, in either case, (18) implies that

$$\text{Tr}(C^T Y) = \text{Tr} \left\{ \begin{bmatrix} C^T & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} Y & \Xi \\ \Lambda & 0 \end{bmatrix} \right\} \geq 0.$$

From Theorem 2, given that the cones (15) are closed, there follows

$$\begin{bmatrix} C^T & 0 \\ 0 & 0 \end{bmatrix} + P = \begin{bmatrix} \bar{Z}^T & N \\ 0 & 0 \end{bmatrix} \begin{bmatrix} A & Q \\ 0 & I_r \end{bmatrix} + \begin{bmatrix} M & 0 \\ R & I_n \end{bmatrix} \begin{bmatrix} \bar{W}^T & 0 \\ S^T & 0 \end{bmatrix} \quad (20)$$

for some matrices

$$\begin{bmatrix} \bar{Z}^T & N \\ 0 & 0 \end{bmatrix} \in \begin{bmatrix} \mathbf{K} & \mathbb{R}^{m \times n} \\ \mathbf{U}_0 & \mathbb{R}^{r \times n} \end{bmatrix}^* = \begin{bmatrix} \mathbf{K}^* & \mathbf{U}_0^* \\ \{0\} & \{0\} \end{bmatrix}$$

and

$$\begin{bmatrix} \bar{W}^T & 0 \\ S^T & 0 \end{bmatrix} \in \begin{bmatrix} \mathbf{L} & \mathbf{V}_0^* \\ \mathbb{R}^{r \times s} & \mathbb{R}^{r \times n} \end{bmatrix} = \begin{bmatrix} \mathbf{L}^* & \{0\} \\ \mathbf{V}_0^* & \{0\} \end{bmatrix};$$

$$P \in \begin{bmatrix} \mathbf{X} & \mathbf{V}_0 \\ \mathbf{U}_0 & \{0\} \end{bmatrix}^\perp.$$

Since $N \in \mathbf{U}^* \cap (-\mathbf{U})^*$, $\text{Tr}(N) = 0$; similarly $\text{Tr}(S) = 0$. Hence (20) gives $C^T + E = \bar{Z}^T A + M \bar{W}^T$, $E \in \mathbf{X}^\perp$, together with $\text{Tr}(\bar{Z}^T Q) = 0 = \text{Tr}(R \bar{W}^T)$. Therefore

$$\begin{aligned} \text{Tr}(C^T \bar{X}) - [\text{Tr}(B^T \bar{Z}) + \text{Tr}(J \bar{W}^T)] &= \text{Tr}(\bar{Z}^T A \bar{X}) + \text{Tr}(M \bar{W}^T \bar{X}) \\ &\quad - \text{Tr}(B^T \bar{Z}) - \text{Tr}(J \bar{W}^T) \\ &= \text{Tr}[\bar{Z}^T (A \bar{X} - B)] + \text{Tr}[(\bar{X} M - J) \bar{W}^T] = 0, \end{aligned} \quad (21)$$

using $\text{Tr}(M\bar{W}^T\bar{X}) = \text{Tr}(\bar{X}M\bar{W}^T)$. This, with (17), proves that (\bar{Z}, \bar{W}) maximizes (14). ■

5. COMPLEX AND QUATERNION LINEAR PROGRAMMING

A vector in complex space \mathbb{C}^n can be represented isomorphically by a column of n real submatrices

$$\begin{bmatrix} x_j & y_j \\ -y_j & x_j \end{bmatrix} \quad (j=1, 2, \dots, n; \quad x_j, y_j \in \mathbb{R}), \tag{22}$$

and thus by a $2n \times 2$ real matrix. Theorem 5 may be applied, taking \mathbf{X} as the subspace of such matrices. The complex inner product $\langle z, w \rangle = \text{Re}(\bar{z}^T w)$ differs only by a constant factor from $\text{Tr}(Z^T W)$, where $Z, W \in \mathbb{R}^{2n \times 2}$ represent $z, w \in \mathbb{C}^n$. Programming in quaternion spaces may be handled similarly, representing each quaternion $q = a\mathbf{I} + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$ by a real 4×4 matrix

$$\begin{bmatrix} a & b & c & d \\ -b & a & -d & c \\ -c & d & a & b \\ -d & -c & -b & a \end{bmatrix}. \tag{23}$$

If quaternion vectors q, r are thus represented by real matrices Q, R , then the inner product $\frac{1}{4}\text{Tr}(Q^T R)$ represents the scalar part of $\bar{q}r$, where \bar{q} denotes the conjugate of q (thus $a\mathbf{I} - b\mathbf{i} - c\mathbf{j} - d\mathbf{k}$). (The scalar part of q is a .)

Denote by $\langle \cdot, \cdot \rangle$ the inner product, for either complex or quaternion n -space. Denote by \mathbf{W} either \mathbb{C} , or quaternion space; let $M: \mathbf{W}^n \rightarrow \mathbf{W}^n$ be a linear map; let $b, c \in \mathbf{W}^n$; let $S \subset \mathbf{W}^m$ be a closed convex cone; denote by \mathbf{H} the orthogonal complement of \mathbb{R} in \mathbf{W} . Consider the pair of linear programs (considered as either in complex or in quaternion space)

$$\text{Minimize } \langle c, x \rangle \quad \text{subject to } Mx - b \in S; \tag{24}$$

$x \in \mathbf{W}^n$

$$\text{Maximize } \langle b, z \rangle \quad \text{subject to } z \in S^*, \quad M^T z = c. \tag{25}$$

$z \in \mathbf{W}^m$

THEOREM 5. *If x and z satisfy the constraints of (24) and (25), respectively, then $\langle c, x \rangle \geq \langle b, z \rangle$. If (24) reaches a minimum at $x = \bar{x}$, and the*

convex cone

$$\begin{bmatrix} M & M\bar{x}-b \\ 0 & I \end{bmatrix}^T \begin{bmatrix} \mathbf{S}^* \\ \mathbf{H} \end{bmatrix} \quad (26)$$

is closed, then (25) reaches a maximum at some \bar{z} , where $\langle c, \bar{x} \rangle = \langle b, \bar{z} \rangle$.

Proof. Apply Theorem 4. ■

For linear programming in complex spaces, with vector variables, see [9] and [7].

The authors thank two referees for several references and amendments.

REFERENCES

- 1 R. Bellman and Ky Fan, On systems of linear inequalities in Hermitian matrix variables, in *Proceedings of the Symposium on Pure Mathematics 7* (V. Klee, Ed.), Amer. Math. Soc., 1963, pp. 1–11.
- 2 A. Ben-Israel, On decompositions of matrix spaces with applications to matrix equations, *Atti Accad. Naz. Lincei Mem. Cl. Sci. Fis. Mat. Natur. Sez. I* 44 (2):122–127 (1968).
- 3 A. Berman and A. Ben-Israel, More on linear inequalities with applications to matrix theory, *J. Math. Anal. Appl* 33:482–495 (1971).
- 4 B. D. Craven, Lagrangean conditions and quasiduality, *Bull. Austral. Math. Soc.* 16:325–339 (1977).
- 5 B. D. Craven, *Mathematical Programming and Control Theory*, Chapman & Hall, London, 1978.
- 6 B. D. Craven and J. J. Koliha, Generalizations of Farkas's theorem, *SIAM J. Math. Anal.* 8:983–997 (1977).
- 7 B. D. Craven and B. Mond, On duality in complex linear programming, *J. Austral. Math. Soc.* 16:172–175 (1973).
- 8 D. G. Kabe, A note on a quadratic programming problem, *J. Indust. Math. Soc.* 23:61–66 (1973).
- 9 N. Levinson, Linear programming in complex space, *J. Math. Anal. Appl.* 14:44–62 (1966).
- 10 M. Marcus and H. Minc, *A Survey of Matrix Theory and Matrix Inequalities*, Prindle, Weber, & Schmidt, Boston, 1964.
- 11 P. Scobey and D. G. Kabe, Vector quadratic programming problems and inequality constrained least squares estimation, *J. Indust. Math. Soc.* 28:37–49 (1978).
- 12 V. A. Sposito and H. T. David, A note on Farkas' theorem over cone domains, *SIAM J. Appl. Math.* 22:356–358 (1972).

Received 27 July 1979; revised 26 July 1980