THE PUBLISHING HOUSE
OF THE ROMANIAN ACADEMY

*PROCEEDINGS OF THE ROMANIAN ACADEMY, Series A,*
*Volume 6, Number 1/2005, pp. 000-000*

# DUALITY FOR MINIMUM MATRIX NORM PROBLEMS

Vasile PREDA[*], Cristinca FULGA[**]

[*] University of Bucharest, Faculty of Mathematics and Informatics, Str. Academiei 14, 010014 Bucharest, Romania
[**] Academy of Economic Sciences, Piata Romana 6, 010174 Bucharest, Romania, E-mail: fulgac@hades.ro
Corresponding author: Vasile PREDA, E-mail: preda@fmi.unibuc.ro

We investigate a minimum matrix norm problem with inequality constraints and its dual. An explicit rule for finding a primal solution from the solution of the dual problem is given. The case with equality constraints is also considered.

*Keywords*: minimum matrix norm solutions, duality, relations between primal and dual solutions.

## 1. PRELIMINARIES

The problem of determining a minimum norm solution arises in various applications, see references [1,2,6,7]. We consider the problem

$$(\text{P1})\quad \min F(X) = \|X\|_p^p \big/ p \text{ , subject to } AX \geq B$$

where $1 < p < \infty$, *A*, *B* are $k \times m, k \times n$ real matrices, $X = (x_{ij})_{\substack{1 \le i \le m \\ 1 \le j \le n}}$ is the $m \times n$ real matrix of unknows. We mention that $\|X\|_p = \left( \sum_{i=1}^{m} \sum_{j=1}^{n} |x_{ij}|^p \right)^{1/p}$ and that $AX \geq B$ means

$$a_i x^j \geq b_{ij}, \forall i = \overline{1,k},\ j = \overline{1,n}$$

where $a_i$ is the *i*th row of *A*, $b^j$ is the *j*th column of $B = (b_{ij})_{\substack{1 \le i \le k \\ 1 \le j \le n}}$, $x^j$ is the *j*th column of *X*.

It is assumed throughout this paper that the feasible region $\Re = \{X \mid AX \geq B\}$ is not empty. Hence the existence and uniqueness of the solution of problem (P1) are ensured by the projection theorem [7]. The dual index of *p* is denoted by *q*, that is $\frac{1}{p} + \frac{1}{q} = 1.$ We will establish that the dual problem of (P1) has the form

$$(\text{D1}) \begin{cases} \max G(y^1,\ldots,y^n) = b^{1T} y^1 + \ldots + b^{nT} y^n - \left\| \widehat{\mathrm{A}}(Y) \right\|_q^q \big/ q \\ y^1 \geq 0,\ldots,y^n \geq 0 \\ y^1,\ldots,y^n \in R^k \end{cases}$$

Reccomended byMarius IOSIFESCU, member of the Romanian Academy

**where** $Y=\begin{pmatrix} y^1 \\ \ldots \\ y^n \end{pmatrix} \in R^{kn}$, with $y^l=\begin{pmatrix} y_{(l-1)m+1} \\ \ldots \\ y_{(l-1)m+k} \end{pmatrix} \in R^k$, $l=\overline{1,n}$ **and** $\widehat{\mathrm{A}}(Y)=\left(A^T y^1 \quad \ldots \quad A^T y^n\right)$**.**We recall that $\left\|\widehat{\mathrm{A}}(Y)\right\|_q=\left(\sum_{l=1}^{n}\sum_{j=1}^{m}\left|c_j^T \cdot y^l\right|^q\right)^{1/q}$.

Next, we will investigate the relations between the two problems, the primal one, (P1), and its dual (D1).

## 2. DUALITY RESULTS

In order to prove that (D1) is the dual of (P1), we need some further notation. Let $c_j$ be the $j$th column of $A$ and let $Z(Y)=\left(z^1(Y) \quad \ldots \quad z^n(Y)\right)$ be a real $m\times n$ matrix, where

$$z^l(Y)=\left(z_{(l-1)m+1}(Y) \quad \ldots \quad z_{(l-1)m+m}(Y)\right)^T \in R^m,\ l=\overline{1,n},$$

and

$$z_{(l-1)m+j}(Y)=\left|c_j^T y^l\right|^{q-1}\cdot \operatorname{sgn}\left(c_j^T y^l\right),\ l=\overline{1,n},\ j=\overline{1,m}.$$

For the component $z_{(l-1)m+j}(Y)$ of the matrix $Z(Y)$ we also use the notation $z_{jl}(Y)$.

The following lemmas are needed in the proof of our main result.

**Lemma 2.1** *We have*

$$\operatorname{grad}\left(\left\|\widehat{\mathrm{A}}(Y)\right\|_q^q \Big/ q\right)=\begin{pmatrix} A\cdot z^1(Y) \\ \ldots \\ A\cdot z^n(Y) \end{pmatrix}\in R^{kn}$$

**Proof.** By direct calculus we have

$$\frac{\partial}{\partial y_{(l-1)m+j}}\left(\left\|\widehat{\mathrm{A}}(Y)\right\|_q^q \Big/ q\right)=\frac{\partial}{\partial y_{(l-1)m+j}}\left(\left(\sum_{l=1}^{n}\sum_{j=1}^{m}\left|c_j^T\cdot y^l\right|^q\right)\Big/ q\right)==\frac{1}{q}\left(\sum_{k=1}^{m} q\cdot a_{jk}\cdot\left|c_k^T y^l\right|^{q-1}\operatorname{sgn}\left(c_k^T y^l\right)\right)=$$
$$=\sum_{k=1}^{m} a_{jk}\cdot z_{(l-1)m+k}=a_j\cdot z^l(Y).$$

**Lemma 2.2** *We have*

$$\sum_{i=1}^{n} y^{iT} A z^i(Y)=\left\|\widehat{\mathrm{A}}(Y)\right\|_q^q$$

**Proof .** We can successively write

$$\sum_{i=1}^{n} y^{iT} A z^i(Y) = \sum_{l=1}^{n}\sum_{j=1}^{k} y_{(l-1)m+j}\cdot a_j \cdot z^l(Y) = \sum_{l=1}^{n}\sum_{j=1}^{k} y_{(l-1)m+j}\cdot \sum_{i=1}^{m} a_{ji}\cdot z_{(l-1)m+i}(Y) =$$
$$= \sum_{l=1}^{n}\sum_{i=1}^{m} z_{(l-1)m+i}(Y)\cdot \sum_{j=1}^{k} y_{(l-1)m+j}\cdot a_{ji} = \sum_{l=1}^{n}\sum_{i=1}^{m} z_{(l-1)m+i}(Y)\cdot c_i^T y^l =$$
$$= \sum_{l=1}^{n}\sum_{i=1}^{m} \left|c_i^T\cdot y^l\right|^{q-1}\cdot \operatorname{sgn}\left(c_i^T y^l\right)\cdot c_i^T y^l = \sum_{l=1}^{n}\sum_{i=1}^{m}\left|c_i^T\cdot y^l\right|^q = \left\|\widehat{\mathrm{A}}(Y)\right\|_q^q$$

**Lemma 2.3.** *We have*

$$\|Z(Y)\|_p^p = \left\|\widehat{\mathrm{A}}(Y)\right\|_q^q$$

**Proof.** Clearly,

$$\|Z(Y)\|_p^p = \sum_{l=1}^{n}\sum_{j=1}^{m}\left|z_{(l-1)m+j}(Y)\right|^p = \sum_{l=1}^{n}\sum_{j=1}^{m}\left|\left|c_j^T\cdot y^l\right|^{q-1}\cdot \operatorname{sgn}\left(c_j^T\cdot y^l\right)\right|^p = \sum_{l=1}^{n}\sum_{j=1}^{m}\left|c_j^T\cdot y^l\right|^q = \left\|\widehat{\mathrm{A}}(Y)\right\|_q^q.$$

Let $g(X) = \left(g_{jl}(X)\right)_{\substack{j=\overline{1,m}\\ l=\overline{1,n}}}$ be a real $m\times n$ real matrix with $g_{jl}(X) = \dfrac{\partial F(X)}{\partial x_{jl}}$.

**Lemma 2.4 .** *We have*

$$g(Z(Y)) = \widehat{\mathrm{A}}(Y)$$

**Proof.** Clearly,

$$g_{jl}(Z(Y)) = \left|z_{jl}(Y)\right|^{p-1}\cdot \operatorname{sgn}\left(z_{jl}(Y)\right) = \left|c_j^T\cdot y^l\right| \operatorname{sgn}\left(c_j^T\cdot y^l\right) = c_j^T y^l \Rightarrow g(Z(Y)) = \widehat{\mathrm{A}}(Y)$$

.

**Theorem 2.5** *Let $X^*$ be the optimal solution of* (P1); *then there exists $Y^*\in R^{nk}$ such that $X^* = Z(Y^*)$ and $Y^*$ is the optimal solution of* (D1). *Conversely, let $Y^*$ be the optimal solution of* (D1); *then $X^* = Z(Y^*)$ is the unique solution of* (P1). *We also have $D(Y^*) = F(X^*)$ and the classical primal-dual inequality $D(Y)\le F(X)$, for any feasible solutions $X, Y$ of* (P1), (D1).

**Proof** Using the Kuhn-Tucker optimality conditions we have that $X^*$ is an optimal solution for (P1) if and only if there exists $Y^*\in R^{nk}$ such that

$$\begin{cases} A\cdot x^{1*}\ge b^1\\ \ldots\\ A\cdot x^{n*}\ge b^n\end{cases};\quad Y^*\ge 0;\quad Y^{*T}\cdot\begin{pmatrix} Ax^{1*}-b^1\\ \ldots\\ Ax^{n*}-b^n\end{pmatrix} = 0;\quad g(X^*) = \hat{A}(Y^*) \tag{2.1}$$

But, by Lemma 2.4, $g(Z(Y^*)) = \hat{A}(Y^*)$**,** hence $X^* = Z(Y^*)$. Substituting this into (2.1) yields

$$\begin{cases} Az^1(Y^*)\ge b^1\\ \quad\ldots\\ Az^n(Y^*)\ge b^n\end{cases};\quad Y^*\ge 0;\quad \begin{cases} y^{1*T}\cdot\left(Az^1(Y^*)-b^1\right) = 0\\ \quad\ldots\\ y^{n*T}\cdot\left(Az^n(Y^*)-b^n\right) = 0\end{cases} \tag{2.2}$$

which are the optimality conditions for (D1). We have proved that if $X^*$ is the optimal solution of (P1), then the corresponding Lagrange vector $Y^*\in R^{nk}$ is the optimal solution of (D1).

Conversely, let $Y^* \in R^{nk}$ be the optimal solution of (D1). Then the optimality conditions (2.2) hold and $X^* = Z(Y^*)$ satisfies (2.1), which proves that $X^*$ is the optimal solution of (P1).

In both cases, by Lemmas 2.1-2.4, we find that $X^*$ and $Y^*$ satisfy

$$D(Y^*) = b^{1T} y^{1*} + \ldots + b^{nT} y^{n*} - \left\| \hat{A}(Y^*) \right\|_q^q \Big/ q =$$

$$= \sum_{i=1}^{n} y^{i*T} \cdot A \cdot z^i(Y^*) - \left\| \hat{A}(Y^*) \right\|_q^q \Big/ q == \left\| Z(Y^*) \right\|_p^p \Big/ p = \left\| X^* \right\|_p^p \Big/ p = F(X^*).$$

Moreover, for all feasible solutions *X*, *Y* of (P1), (D1) we have

$$D(Y) = b^{1T} y^1 + \ldots + b^{nT} y^n - \left\| A^T y^1 + \ldots + A^T y^n \right\|_q^q \Big/ q \le D(Y^*) = F(X^*) \le F(X).$$

An important special case of (P1) occurs when the system $AX \ge B$ is replaced by $AX = B$. In this case the primal problem

$$(\text{P2}) \begin{cases} \min F(X) = \|X\|_p^p \big/ p \\ AX = B \end{cases}$$

can be written as

$$(\text{P3}) \begin{cases} \min F(X) = \|X\|_p^p \big/ p \\ AX \ge B \\ -AX \ge -B. \end{cases}$$

Always under the hypothesis $\{X \in M_{m,n}(R) | AX = B\} \neq \emptyset$, by Theorem 2.5 the dual of (P3) has the form

$$\begin{cases} \max\left( \sum_{i=1}^{n} b^{iT} u^i - \sum_{i=1}^{n} b^{iT} v^i - \left\| \hat{A}(U,V) \right\|_q^q \Big/ q \right) \\ u^i \in R^k, u^i \ge 0, \ i = \overline{1,n} \\ v^i \in R^k, v^i \ge 0, \ i = \overline{1,n} \end{cases}$$

where $\hat{A}(U,V) = \begin{pmatrix} A^T u^1 & \ldots & A^T u^n \\ A^T v^1 & \ldots & A^T v^n \end{pmatrix}$ is a real $2m \times n$ matrix, $u^i, v^i \in R^k, i = \overline{1,n}.$ Substituting $u^i - v^i = y^i, i = \overline{1,n}$ yields the following result.

**Corollary 2.6** *The dual of the minimum norm problem* (P2) *is*

$$\begin{cases} \max\left( \sum_{i=1}^{n} b^{iT} y^i - \left\| \hat{A}(Y) \right\|_q^q \Big/ q \right). \\ y^i \in R^k, i = \overline{1,n}. \end{cases}$$

*If* $X^*$ *is the optimal solution of* (P2), *then there exists* $Y^* \in R^{nk}$ *such that* $X^* = Z(Y^*)$ *and* $Y^*$ *is the optimal solution of* (D2). *Conversely, if* $Y^*$ *is the optimal solution of* (D2) *then* $X^* = Z(Y^*)$ *is the unique solution of* (P2). *We also have* $D(Y^*) = F(X^*)$ *and the classical primal-dual inequality* $D(Y) \le F(X)$ *for any feasible solutions* $X, Y$ *of* (P2), (D2).

## REFERENCES

1. CADZOW, J. A., *A finite algorithm for the minimum* $l_\infty$*-solution to a system of consistent linear equations,* SIAM Numerical Analysis, **10**, pp. 607-617, 1973.

2. DAX A., *An extended Kaczmarz's method for* $l_p$*-minimum norm solutions***,** Technical Report, Hydrological Service of Israel, 1992.

3. DAX A., *On minimum norm solutions***,** Journal of Optimization Theory and Applications, **76**, no. 1, pp. 183-193, 1993.

4. FULGA C., *An extension of the minimum norm problem*, Mathematical Reports, **7** (**57**), pp. 45-50, 2005.

5. FULGA C., *On minimum norm problems*, Analele Universitatii Bucuresti Mat., **52**, no. 2, pp. 175-180, 2003.

6. HERMAN G. T., LENT A., *A family of iterative quadratic optimization algorithms for pairs of inequalities with application in diagnostic radiology*, Mathematical Programming Study, **9**, pp. 15-29, 1978.

7. LUENBERGER D.G., *Optimization by Vector Space Methods*, Wiley, New York, 1969.

8. PREDA V., *Optimality conditions for a class of mathematical programming problems***,** Bull. Soc. Sci. Math. Roumanie, **34**, (**82**), pp. 343-350, 1990.

9. PREDA V., *Nonlinear programming and matrix game equivalence***,** J. Austral. Math. Soc. Ser. B**, 35**, pp. 429-438, 1994.

10. ROCKAFELLAR R.T., *Convex Analysis* , Princeton University Press , Princeton, N.J., 1970.

*Received January 24, 2005*