# Dual Integer Linear Programs and the Relationship between their Optima

Ron Aharoni

The Technion, Haifa, Israel

Paul Erdös

Hungarian Academy of Sciences

Nathan Linial

Hebrew University, Jerusalem, Israel.

**Abstract**: We consider dual pairs of packing and covering integer linear programs. Best possible bounds are found between their optimal values. Tight inequalities are obtained relating the integral optima and the optimal rational solutions.

## Introduction

Much of the research in discrete optimization deals with max min theorems, such as Hall's Marriage Theorem, Menger's Theorem, Max Flow Min Cut to name only some of the better known ones. Apart from their great aesthetic appeal these results are of fundamental importance in discrete optimization because of the tendency of max min theorems to be coupled with polynomial time algorithms for evaluating the quantities in question.

By now it is widely realized that a good context for studying these results is that of duality theory of mathematical programming.

Most of the *max−min* theorems can be stated in the form that a pair of dual integer linear programs have the same optimum. This is not generally true for integer linear programs and those cases where this phenomenon occurs are of course of great interest and importance. The purpose of the present article is to investigate how much the optima can differ. This is an extremely difficult question for which no simple answer is expected. Here we take a first step towards the resolution of this problem. Let us state the problem in its general form:

Consider the pair of dual integer linear programs

$$\text{(ILP)}\quad \begin{array}{l} \max x\cdot b \\ Ax \le c \\ \textit{integer } x \ge 0 \end{array} \qquad \text{(DILP)}\quad \begin{array}{l} \min y\cdot c \\ yA \ge b \\ \textit{integer } y \ge 0 \end{array}$$

and their rational relaxations.

$$\text{(LP)}\quad \begin{array}{l} \max x\cdot b \\ Ax \le c \\ x \ge 0 \end{array} \qquad \text{(DLP)}\quad \begin{array}{l} \min y\cdot c \\ yA \ge b \\ y \ge 0 \end{array}$$

Assuming that all of these are feasible and bounded three numbers are naturally defined: $z, Z, q$ which are the optima of (ILP),(DILP) and (LP) respectively. By duality theory of linear programming $q$ is also the optimum of DLP. We shall assume throughout that $A$ is an $n \times m$ matrix with nonnegative integer entries,

$$b \in \mathbb{Z}_+^m \quad , \quad c \in \mathbb{Z}_+^n$$

So $z, Z$ are both nonnegative integers. The obvious relationship is

$$Z \geq q \geq z.$$

The theory we have in mind should find further connections between the three parameters which involve only functions of $A, b, c$ which are computable in polynomial time. Assuming $P \neq NP$ it follows that the values of $z, Z$ are not such functions and so only bounds are to be expected, rather than exact results.

An interesting aspect of this problem is based on the integral optima being $NP$- hard to compute whereas the rational optima are computable in polynomial time. A very reasonable thing to do is try to approximate the integral optimum by the rational one. In this context our problem is to seek bounds on the quality of the rational optimum as an approximation for the integral optimum. After some consideration of this general problem it becomes apparent that it involves difficulties of both combinatorial and number theoretical nature. To eliminate the later we restrict our attention to $A$ being a $0-1$ matrix and $b, c$ being the all one vectors $1_m, 1_n$ resp. The matrix $A$ can be thought of as the incidence matrix of a hypergraph $G$ on $n$ vertices having $m$ edges. Then $z$ is the matching number of this hypergraph, namely, the largest number of mutually disjoint edges in $G$. $Z$ is the covering number of $G$, that is , the least size of a set of vertices meeting every edge. The parameter $q$ is commonly referred to as the optimal fractional matching number or optimal fractional covering number of $G$.

$$\begin{array}{ll} z = \max x \cdot 1 & Z = \min x \cdot 1 \\ Ax \leq 1 & x \cdot A \geq 1 \\ \textit{integer } \; x \geq 0 & \textit{integer } \; x \geq 0 \end{array}$$

$$\begin{array}{ll} q = \max x \cdot 1 & q = \min x \cdot 1 \\ Ax \leq 1 & xA \geq 1 \\ x \geq 0 & x \geq 0 \end{array}$$

where $A$ is an $n \times m$ $0-1$ matrix.

We shall interchangeably refer to $A$ as a matrix and as a hypergraph on $n$ vertices and $m$ edges. So vertex and row and likewise edge and column will be thought of as identical terms. Thus when we talk about the size of an edge in $A$ we mean the number of 1's in the corresponding column of the matrix.

Three questions suggest themselves immediately

1) Given $n, m, q$ how large can $Z$ be?

2) Given $n, m, z$ how large an $q$ be?

3) Given $n, m, z$ how large can $Z$ be?

The first question has been addressed by Lovasz [Lo] (see also [J]).

**Theorem 1 [Lo]**: Using the above notation

$$q(1 + \log n) \geq Z$$

This result is best possible in that there exists a positive constant $C$ such that for every $q, n$ there is a matrix $A$ for which

$$q(1 + \log n) < CZ.$$

The theorem is also effective in that it is possible to find in time polynomial in $m, n$ a set of not more than

$$q(1 + \log n)$$

vertices which meets all edges.

Our new results concern questions (2) and (3):

**Theorem 2**: With the above notations

$$z \geq \frac{q^2}{n - \frac{f-1}{m} q^2} \geq \frac{q^2}{n}$$

where $f$ is the least cardinality of an edge. This inequality is best possible in the sense that there are infinitely many cases where equality holds. It is effective in that we can find in time polynomial in $m,n$ a collection of mutually disjoint edges having at least

$$\frac{q^2}{n - \frac{f-1}{m} q^2}$$

members. ▪

As for the third question we have:

**Theorem 3**: Define

$$f = f(m,n,z) = \sqrt{nz \log \frac{m}{\sqrt{nz}}}$$

Then

$$(*) \qquad Z \leq \begin{cases} \min(n, 3f) & m \geq e\sqrt{nz} \\ m & m < e\sqrt{nz} \end{cases}$$

It is best possible for $m < e\sqrt{nz}$ and for $m > en$ in that there is a positive constant $\varepsilon > 0$ ($\varepsilon = 1/5$ will do) such that for $m,n,z$ in these ranges we can construct $A_{n \times m}$ for which $z(A) \leq \varepsilon z$ and

$$\varepsilon Z \leq \begin{cases} \min(n, 3f) & m \geq en \\ m & m < e\sqrt{nz} \end{cases}.$$

The theorem is also effective in that we can find in time polynomial in $m,n$ for every $A_{n \times m}$ feasible solutions with values $Y,y$ respectively which satisfy

$$Y \leq \begin{cases} \min(n, 3f(m,n,y)) & \text{for } m \geq e\sqrt{ny} \\ m & \text{for } m > e\sqrt{ny} \end{cases}.$$

**Proof of Theorem 2**: Let $x$ be an optimal solution for (LP), that is $x \cdot 1 = \Sigma x_i = q$, $Ax \leq 1$, $x \geq 0$. Now

$$x^T A^T Ax = (Ax)^T (Ax) \leq 1^T \cdot 1 = n.$$

Let us consider the matrix $A^T A$: it is an $m \times m$ matrix indexed by the edges $A_i$ with the $(i,j)$ entry being $|A_i \cap A_j|$. We may then estimate it as follows :

$$A^T A \geq diag(|A_i| - 1) + J - B$$

where $J$ is the $m \times m$ matrix with all entries equal one and $B$ is 0−1 matrix having 1 at the $(i,j)$ entry iff $A_i$ and $A_j$ are disjoint. Now

$$n \geq x^T A^T Ax \geq x^T \ diag(|A_i| - 1)x + x^T Jx - x^T Bx.$$

We now consider the three terms:

$$x^T diag(|A_i| - 1)x = \sum_{i=1}^{m} (|A_i| - 1)x_i^2 \geq (f-1) \Sigma x_i^2 \geq$$

$$(f-1) \frac{1}{m} (\Sigma x_i)^2 = (f-1) \frac{q^2}{m}$$

We have used the fact that $|A_i| \geq f$ and the last inequality follows from the Cauchy Schwartz inequality.

$$x^T Jx = (\Sigma x_i)^2 = q^2$$

To estimate $x^T Bx$ we need an idea of Motzkin and Straus [MS], who consider the following problem:

Let $G = (V,E)$ be a graph, what is the maximum of

$$\sum_{[i,j] \in E} x_i x_j$$

where $x_i$ $(i \in V)$ are nonnegative and $\Sigma x_i = 1$. The answer is that the optimum is attained by picking a largest clique in $G$. If it has size $k$, say, then set $x_i = 1/k$ on its vertices and zero elsewhere . This assignment makes

$$\sum_{[i,j] \in E} x_i x_j = \frac{1}{2}(1 - \frac{1}{k}).$$

Now if $P$ is the adjacency matrix of $G$, then

$$\sum_{[i,j] \in E} x_i x_j = \frac{1}{2} x^T Px,$$

So we conclude that if $P$ is the adjacency matrix of a graph $G$ whose largest clique has $k$ vertices, then for every $x \geq 0$

$$x^T P x \leq (1-\frac{1}{k})(\Sigma x_i)^2$$

In our case $B$ is the adjacency matrix of the graph whose vertices are the edges of our hypergraph, two vertices being adjacent if the corresponding edges are disjoint. Notice that a clique in this graph corresponds to a matching in the hypergraph and so the largest clique has size $z$. Therefore

$$x^T B x \leq (1-\frac{1}{z})(\Sigma x_i)^2 = q^2(1-\frac{1}{z}).$$

We have thus shown

$$n \geq x^T A^T A x \geq q^2(\frac{f-1}{m}+1-(1-\frac{1}{z})) = q^2(\frac{f-1}{m}+\frac{1}{z}).$$

To show the tightness of this inequality let $A$ be the incidence matrix of a finite projective plane. We leave it to the reader to verify that in this case

$$m = n = p^2+p+1,\ z = 1 \quad \text{(every two lines meet)}$$
$$f = p+1,\ q = m/f$$

and so the inequality is in fact an equality.

Notice that equality can hold here only if $m \leq n$; It is quite easy to see that for equality to hold all edges must have size $f$. Now the optimum of (LP) is attained at a vector $x$ with at most $n$ nonzero coordinates. It is therefore possible to omit, as long as $m \geq n$, an edge without changing $q$ and of course $z$ can only decrease. Thus if equality held before then after this step we would be getting a counterexample to our theorem. The idea behind this remark should be accredited to Füredi [Fu]. The point that we would like to make here is that further inequalities should be sought after and our inequality should not be considered as the final answer to our question (2).

Finally let us show that a matching satisfying the inequality can be found in polynomial time. The value of $q$ may be found in time polynomial in $m,n$ using the ellipsoid algorithm, since all input values are 0, 1.

Once the optimal vector $x$ is known, the way to find a large matching is exactly as in [MS]: If $x_i, x_j > 0$ with $A_i \cap A_j \neq \phi$, then consider the vectors $u, v$ defined by

$$u_k = \begin{cases} x_k & k \neq i,j \\ x_i+x_j & k = i \\ 0 & k = j \end{cases} \qquad v_k = \begin{cases} x_k & k \neq i,j \\ 0 & k = i \\ x_i+x_j & k = j \end{cases}$$

It is easily verified that $\Sigma u_i = \Sigma v_i = \Sigma x_i$ and

$$\max(u^T B u,\ v^T B v) \geq x^T B x$$

By continuing in this fashion we will produce a vector $w$ which is supported on a matching and which satisfies $w^T B w \geq x^T B x$, and $\Sigma w_i = \Sigma x_i$.

It is again clear that if the support of $w$ has $p$ members, then

$$w^T B w \leq (1-\frac{1}{p})(\Sigma x_i)^2$$

Thus the matching which supports $w$ has

$$p \geq \frac{q^2}{n - \frac{f-1}{m}q^2}$$

edges and it is clearly found in polynomial time. •

**Proof of Theorem 3:** For the first part we only have to show that $Z \leq 3\sqrt{nz \log \frac{m}{\sqrt{nz}}}$ holds for $m > e\sqrt{nz}$. Given $m,n,z$ we actually construct a set of vertices which meets all edges with cardinality not exceeding

$$3\sqrt{nz \log \frac{m}{\sqrt{nz}}}$$

There are three arguments involved:

a. If "small" edges exist then we consider a maximal matching among them. The set of vertices in the union meets all small edges.

b. The generic step of this construction is a greedy one: we pick a vertex which belongs to the largest number of edges and omit it and the edges containing it.

c. Towards the end of this process we estimate the number of steps by saying that it does not exceed the number of remaining edges.

**Step a**: Assume first that there are some edges of size smaller than

$$r = \sqrt{\frac{n}{z}\log\frac{m}{\sqrt{nz}}} = \frac{f}{z}.$$

Consider a maximal collection of mutually disjoint edges of size $< r$. Such a collection can have at most $z$ members and their union meets all edges of size $< r$. Thus there is a set of

$$zr = \sqrt{nz\log\frac{m}{\sqrt{nz}}} = f$$

vertices which meets all edges of size $< r$.

**Step b**: Now we are allowed to assume that all edges have size $\geq r$. *Therefore there is a vertex which* belongs to a fraction of at least $r/n$ of the edges. We omit this vertex and the edges which contain it. After this is repeated $t$ times the number of remaining edges does not exceed

$$m\left(1-\frac{r}{n}\right)^t < me^{-\frac{rt}{n}}.$$

We repeat this $x$ times where $x$ is defined by

$$(*)\ \ me^{-\frac{rx}{n}} = x$$

**Step c**: Now at most $x$ edges remain and we pick a vertex from each. In steps b and c together we used at most $2x$ vertices to cover all edges of size $\geq r$. We would like to show that $x \leq rz$. Since we have $m = xe^{\frac{rx}{n}}$ and the r.h.s. is an increasing function of $x$ it suffices to show

$$m \geq rz\, e^{\frac{r^2 z}{n}}$$

which easily follows when $m > e\sqrt{nz}$.

We have thus constructed a cover with no more that $3rz$ vertices.

Now we come to showing that our bound is best possible: For $m < e\sqrt{nz}$ we achieve $Z > \frac{1}{3}m$ as follows. Consider first the case where $z = 1$ and $n = \binom{m}{2}$ and let $A$ be the $n \times m$ matrix in which each row has weight 2 and every two rows are distinct. It is easily verified that for this $A$ we have $z(A) = 1$ and $Z(A) = \lceil \frac{m}{2} \rceil$. For other values of $z$ we take $A$ to be the direct sum of $z$ matrices of dimension $\binom{x}{2} \times x$ where $x = m/z$. To cover all the range $m < e\sqrt{nz}$ we may have to either duplicate some of the columns in $A$ or delete some of them but in every case $Z \geq \frac{1}{3}m$ still holds.

In the range $m > en$ the optimality of our bound is shown using a random construction:

Let $r = \sqrt{\frac{n}{a}\log\frac{m}{\sqrt{na}}}$ and consider a random hypergraph on $[n]$ consisting of $m$ edges of size $r$. We want to find the expected values of $z$ and $Z$ for such a hypergraph.

The probability that two $r$-sets be disjoint is

$$\frac{\binom{n-r}{r}}{\binom{n}{r}} \leq \left(\frac{n-r}{n}\right)^r \leq e^{-r^2/n}$$

and the probability that a set of $z$ edges would be mutually disjoint is

$$\binom{n-r}{r}\binom{n-2r}{n}\cdots\binom{n-(z-1)r}{r}\Big/\binom{n}{r}^{z-1} \leq$$
$$\leq \prod_{t=1}^{z-1}\left(1-\frac{tr}{n}\right)^r \leq \prod_{t=1}^{z-1} e^{-\frac{tr^2}{n}} = \exp\left(\frac{-r^2}{n}\binom{z}{2}\right) \leq$$
$$\leq \exp\left(-\frac{r^2 z^2}{5n/2}\right)$$

the last inequality uses $z \geq 5a \geq 5$.

It follows that the expected number of $z$-matchings does not exceed

$$\binom{m}{z} \exp\left(-\frac{2r^2z^2}{5n}\right).$$

We like to show that for $z \geq 5a$ this number is $< 1/3$ and so

$$Pr(G \text{ has a } 5a\text{-matching}) < 1/3.$$

Extracting $z$-th root and using

$$\binom{m}{z} < \left(\frac{me}{z}\right)^z$$

it suffices to show

$$\frac{em}{z} \exp\left(-\frac{2r^2z}{5n}\right) < \frac{1}{2}$$

for $z \geq 5a$. The lhs decreases with $z$ so it suffices to show it for $z = 5a$. Now

$$\frac{2r^2 \cdot 5a}{5n} = \log \frac{m^2}{na}$$

so we have to show

$$\frac{em}{5a} \cdot \frac{na}{m^2} < \frac{1}{2}$$

which follows since $m > en$.

As for $Z$ the probability that a given set of $Z$ vertices meets all the edge is

$$\frac{\binom{\binom{n}{r} - \binom{n-Z}{r}}{m}}{\binom{\binom{n}{r}}{m}} \leq \exp\left(-m \frac{\binom{n-Z}{r}}{\binom{n}{r}}\right) \leq \exp\left(-m\, e^{-\frac{2rZ}{n}}\right)$$

In the last inequality we used $r, Z \leq \frac{1}{5}n$. Therefore the expected number of $Z$-covers does not exceed

$$\binom{n}{Z} \exp\left(-m e^{-\frac{2rZ}{n}}\right)$$

which we would like to show is less than $1/3$ for

$$Z \geq \frac{1}{5}\sqrt{na \log \frac{m}{\sqrt{na}}}$$

the above function increases with $Z$ so it suffices to check for $Z = \frac{1}{5}\sqrt{na \log \frac{m}{\sqrt{na}}}$. Now

$$\frac{2rZ}{n} = \frac{1}{5} \log \frac{m^2}{na}.$$

We use $\log\binom{n}{Z} \leq Z \log \frac{ne}{Z} \leq Z \log \frac{m}{Z}$ and it remains to verify

$$Z \log \frac{m}{Z} = (m^2an)^{1/5} < -1$$

which certainly follows from

$$2Z \log \frac{m}{Z} < (m^3an)^{1/5}.$$

This reduces, after some simplification to

$$2 \log^{1/2} x \cdot \log \frac{5x}{\sqrt{\log x}} < 5x^{4/5}$$

where $x = ma^{-1/2}n^{-1/2} > e$, and the proof is complete.

Now we turn to the computational aspect of the theorem: To transform the proof into an algorithm notice that we do construct a matching and a cover in the proof. We start by finding a maximal matching among the small sets and then continue to construct a cover by using the vertices covered by that matching plus a cover for the large sets, which we construct greedily. If we know where to switch from the construction of the matching to the construction of the cover our problem is solved. The answer to this is implicit in the above proof: Order the edges of the hypergraph $A_1, \ldots, A_m$ such that their sizes are non decreasing. Define a matching $\{M_i\}$ as follows: $M_1 = A_1$ and $M_j = A_i$ where $A_i$ is disjoint from $\bigcup_{\alpha=1}^{j-1} M_\alpha$

and $i$ is the least index with this property.

Let $t$ be the index for which

$$t\,|A_t| \exp(|A_t|^2 t/m) \leq m < (t+1)\,|A_{t+1}| \exp(|A_{t+1}|^2 (t+1)/m)$$

The case where no such $t$ exists will be considered later.

Our construction proceeds as follows: We consider the matching $M_1, \ldots, M_t$ and a cover which consists of $\bigcup_1^t M_i$ and the cover for the family

$\{A_j \mid A_j \cap (\bigcup_{i=1}^{t} M_i) = \phi\}$ which is produced by the greedy algorithm. The same calculation which was used to prove the upper bound on $Z$ applies now with $z$ replaced by $t$.

We come back to the case where no such $t$ exists. If already

$$|A_1| \exp(|A_1|^2/n) > m$$

then we apply the greedy algorithm for cover immediately and get the upper bound for $Z$ valid already with $z = 1$.

At the other extreme, if

$$k\,|A_k| \exp(k\,|A_k|^2/n) < m$$

and $(M_1, \ldots, M_k)$ is a maximal matching then consider $\bigcup_1^k M_i$ as a cover the upper bound holds. •

A somewhat bothering point is that we do not have the best result yet for $O(\sqrt{nz}) < m < O(n)$. At the moment our upper and lower bounds for $Z$ are still not of the same order of magnitude. It should be pointed out that only a slight modification is required to cover the range $m > \varepsilon n$ rather than $m > en$ and $m < \varepsilon^{-1}\sqrt{nz}$ rather than $m < e\sqrt{nz}$, the constants that will come in will depend on $\varepsilon$. This has not been done because we hope to be able to give the best bound in this range in the full version of this article.

**Further directions for research:**

As we said in the introduction this work should only be considered as a first step towards a better understanding of the relationship between integral and rational optima. Once the coefficients cease being $0-1$ a host of number theoretical issues come in which we did not yet touch upon. The following questions immediately arise.

1) Solve the problem which we considered in this article under the assumption that all $|a_{ij}|, |b_i|, |c_i|$ are bounded by some constant $M$.

2) The case of totally unimodular suggests that some positive results could follow from an assumption that all minors in $A$ have a bounded determinant. Is this true?

3) As an aside of the previous question, the following problem seems very intriguing: What is the computational complexity of the followng problem: Given an integral matrix $A$ and an integer $k$, decide whether all minors in $A$ have a determinant not exceeding $k$ in absolute value.

   The case $k = 1$ is the case of total unimodularity where a polynomial time algorithm was given by Seymour [S]. The general case seems open.

4) In the course of our investigation the following class of problems came up: We have found very tight bounds on $z, Z$ if $A$ is a random $0-1$ matrix. What can be said about $q$ then? The answer depends of course on the probability distribution from which $A$ is drawn. Making the appropriate choices interesting applications can result from an answer to this question.

5) The common approach to solving integer programming problems is via introducing cuts. These are inequalities which must hold due to the integrality of the solution. It would be interesting to investigate the improvement of the approximation of the rational relaxations as more cuts are introduced.

## References

[Fu] Z. Füredi, Maximum degree and fractional matchings in uniform hypergraphs, Combinatorica, 1 (1981), 155-162.

[J] D. Johnson, Approximation Algorithms for combinatorial problems, J. Computer System Sci. 9 (1974), 26-278.

[L] L. Lovasz, On the ratio of optimal and fractional covers, Discrete Math. 13 (1975), 383-390.

[MS] T.S. Motzkin, E. G. Straus, Maxim for graphs and a new proof of a theorem of Turan, Canad. J. Math. 17 (1965), 533-540.

[S] P. Seymour, Decomposition of regular matroids, J. Combinatorial Th., ser. B. 28 (1980), 305-359.