COWLES COMMISSION DISCUSSION PAPER: MATHEMATICS NO. 428

Farkas Lemma, Duality and Lagrangian Multipliers

by William S. Vickrey*

Columbia University

February 14, 1955

I. Farkas Lemma (1901): In order that a homogeneous linear inequality (1) $\sum_r t_r u_r \geq 0$ hold for all values of $\bar{\bar{u}}$ satisfying a system of homogeneous linear inequalities

$$\sum_r b_{sr} u_r \geq 0, \qquad s = 1,2,\ldots\ldots S \tag{2}$$

it is necessary and sufficient that there exist some $\bar{\bar{w}}$ with non-negative elements such that

$$t_r = \sum_r w_s b_{sr}, \qquad s = 1,2\ldots\ldots S. \tag{3}$$

---

* The following proofs were contributed by Prof. Vickrey as part of refereeing STUDIES IN THE ECONOMICS OF TRANSPORTATION. The theorems are not new, but the proofs here presented may perhaps be found more accessible to economists than those often cited from the literature and make it easier for them to establish a firm foundation for linear programming and similar topics.

Proof: The sufficiency is seen immediately by multiplying the inequalities in (2) by the respective v_s and adding:

$$\sum_s v_s \sum_r b_{sr} u_r = \sum_r u_r \sum_s v_s b_{sr} = \sum_r u_r t_r \geq 0.$$

To show the necessity is less trivial. Put U for the set of all $\bar{\bar{u}}$ satisfying (2) (i.e., U is the cone which is the intersection of the half-spaces H_s defined severally, by the individual inequalities of (2)), T for the set of all $\bar{\bar{t}}$ satisfying (1) for all $\bar{\bar{u}}$ in U, (i.e., making acute or right angles with all $\bar{\bar{u}}$ in U), and V for the set of all $\bar{\bar{v}}$ obtained by the use of non-negative w_s in (3) (i.e., the space spanned by the normals $\bar{\bar{b}}_s$ of the half-spaces H_s defining U). From the above, if $\bar{\bar{p}}$ is in V it must be in T. We have to show that if $\bar{\bar{p}}$ is not in V it cannot be in T, i.e., that T and V coincide. Since V is convex (i.e., if v and v' are both in V, then $v'' = v + a\,(v' - v)$ is also in V for all $0 \leq a \leq 1$) and closed (i.e., if for any $\bar{\bar{p}}$, for any positive $\underline{e}$ however small there is always a point $\bar{\bar{v}}$ in V such that $|\bar{\bar{p}} - \bar{\bar{v}}| < e$, then $\bar{\bar{p}}$ is in V), we can for any $\bar{\bar{p}}$ outside V find a halfspace H' containing V but not $\bar{\bar{p}}$, and in particular containing all the normals $\bar{\bar{b}}_s$ to the halfspaces H_s intersecting in U. The normal ray $\bar{\bar{h}}$ defining this halfspace H' will then make non-obtuse angles with all the $\bar{\bar{b}}_s$, and hence will be in all the half-spaces H_s and hence will be in U, and since $\bar{\bar{p}}$ makes an obtuse angle with this ray, $\bar{\bar{p}}$ cannot be in T.

In detail, let $\bar{\bar{z}} = \sum w_s \bar{\bar{b}}_s$ be the element of V closest to $\bar{\bar{p}}$, that is, for which

$$\sum_r (z_r - p_r)^2 \leq \sum_r (v_r - p_r)^2$$

for all $\bar{\bar{v}}$ in V. Such a $\bar{\bar{z}}$ exists by virtue of V being closed. Then we can show that $\sum_r (v_r - z_r)(z_r - p_r) \geq 0$ for all $\bar{\bar{v}}$ in V, (i.e. that $\bar{\bar{v}}\,\bar{\bar{z}}\,\bar{\bar{p}}$ is not an acute angle). For let $\bar{\bar{y}} = \bar{\bar{z}} + a\,(\bar{\bar{v}} - \bar{\bar{z}})$, choosing

$$a = \frac{\sum_r (z_r - v_r)(z_r - p_r)}{\sum_r (z_r - v_r)^2} \quad \text{so that} \quad \sum_r (y_r - p_r)\,(y_r - z_r) = 0 \quad \text{i.e.,}$$

$\bar{\bar{p}}\,\bar{\bar{y}}\,\bar{\bar{z}}$ is a right angle. Then $\bar{\bar{y}}$ is closer to $\bar{\bar{p}}$ than $\bar{\bar{z}}$ is:

$$\sum_r (p_r - y_r)^2 = \sum_r (p_r - z_r)^2 - \sum_r (y_r - z_r)^2 + 2\sum_r (y_r - p_r)(y_r - z_r) \leq \sum_r (p_r - z_r)^2$$

Now if $\sum_r (v_r - z_r)\,(z_r - p_r) \leq 0$, then $a \geq 0$, and either $0 \leq a \leq 1$, in which case $\bar{\bar{y}}$ is in V and closer to $\bar{\bar{p}}$ than $\bar{\bar{z}}$, contrary to construction, or $a > \frac{1}{2}$, in which case

$$2\sum_r (z_r - v_r)\,(z_r - p_r) > \sum_r (z_r - v_r)^2$$

$$\text{or} \qquad \sum_r (p_r - z_r)^2 > \sum_r (p_r - v_r)^2$$

and $\bar{\bar{v}}$ is closer to $\bar{\bar{p}}$ than $\bar{\bar{z}}$ is, contrary to construction.

Thus we have $\sum_r (v_r - z_r)\,(z_r - p_r) \geqq 0$ for all $\bar{\bar{v}}$ in V. Since $\bar{\bar{v}} = 0$ is in V, we have $-\sum_r z_r(z_r - p_r) \geqq 0$ and since $\bar{\bar{v}} = 2\,\bar{\bar{z}}$ is in V, we have $\sum_r z_r\,(z_r - p_r) \geqq 0$. Hence $\sum_r z_r(z_r - p_r) = 0$ and $\sum_r v_r(z_r - p_r) \geqq 0$ for all $\bar{\bar{v}}$ in V (i.e., $\bar{\bar{z}} - \bar{\bar{p}}$ is the required $\bar{\bar{h}}$) and in particular for $\bar{\bar{v}}_s = \bar{\bar{b}}_s = (b_{s1}, b_{s2}, \ldots b_{sr})$, $s = 1 \ldots\ldots r$. $\bar{\bar{u}} = (\bar{\bar{z}} - \bar{\bar{p}})$ therefore satisfies equations (2) and $\bar{\bar{z}} - \bar{\bar{p}}$ is an element of U. Then for $\bar{\bar{p}}$ to be in T would require, from (1), $\sum_r (z_r - p_r) p_r \geq 0$. But $\sum_r (z_r - p_r) p_r = \sum_r z_r (z_r - p_r) - \sum_r (z_r - p_r)^2 < 0$, since $\bar{p} \neq \bar{z}$, so $\bar{p}$ cannot be an element of T. Thus T and V coincide, and the condition is both necessary and sufficient.

2. The Extended Farkas Lemma. This lemma can be extended immediately to the non-homogeneous case by letting the indexes s and r range from 0 to S and 0 to n respectively, and putting $b_{or} = \delta_{or}$ (i.e., $b_{oo} = 1$, $b_{or} = 0$ for $r \neq 0$), so that for $s = 0$, (2) gives $u_o \geq 0$. Let U^o be that part of U for which $u_o = 0$, and $U' = U - U^o$ be that part of U for which $u_o > 0$. Assume first that U' is non-empty.

Now if $\sum_r t_r u_r$ is negative for any $\bar{u}^o$ in U^o, it must also be negative in a sufficiently small neighborhood of $\bar{u}^o$. This neighborhood must contain elements of U', for since U is convex, if $\bar{u}'$ is an element of U', $\bar{p} = \bar{u}^o + a(\bar{u}' - \bar{u}^o)$ is an element of U' for all $a > 0$. Therefore if (1) holds throughout U', it must also hold throughout U; the converse is trivially true. We can therefore substitute $u_o > 0$ for $u_o \geq 0$ in the equations (2) which will now define U' instead of U, and Farkas' lemma will still hold; if then we divide through by u_o and put $v_i = u_i/u_o$, application of the Farkas lemma yields the following extended lemma:

(4) $$\sum_{r=1}^{n} t_r v_r \geq -t_o \qquad \text{holds for all } \bar{v} \text{ satisfying}$$

(5) $$\sum_{r=1}^{n} b_{sr} v_r \geq -b_{so} \qquad s = 1,2,\ldots.S$$

if and only if, for some $\bar{w}$ with $w_i \geq 0$,

$$(6) \qquad t_r = \sum_{s=1}^{S} b_{sr} w_s, \qquad r = 1,2,\ldots.n, \quad \text{and}$$

$$(7) \qquad t_o = \sum_{s=1}^{S} b_{so} w_s + v_o, \quad \text{i.e.,} \quad t_o \geq \sum_{s=1}^{S} b_{so} w_s.$$

On the other hand if U' is empty, then $-u_o \geq 0$ for all $\bar{\bar{u}}$ in U, and we can use this as the (1) of Farkas lemma to show the existence of an $\bar{\bar{h}}$ with non-negative h_i such that

$$\sum_s h_s b_{sr} = 0, \quad r = 1,2,\ldots.n$$

and $\sum_s h_s b_{so} = -1$; applying these h_s to the equations (5) and adding yields $0 \geqq 1$; i.e., there are no $\bar{\bar{v}}$ satisfying (5), and the theorem expressed in (4, 5, 6, 7) holds trivially.

3. The Duality Theorem

The duality theorem states that

$$(8) \qquad \sum_r t_r u_r$$

has a minimum value m with respect to all $\bar{\bar{u}}$ satisfying the constraints:

$$(9) \qquad \sum_r b_{sr} u_r \geqq c_s, \qquad s = 1,2,\ldots.S$$

$$(10) \qquad u_r \geqq 0 \qquad r = 1,2,3,\ldots..R$$

if and only if

$$(11) \qquad \sum_s c_s v_s$$

has a maximum value M with respect to all $\bar{\bar{v}}$ satisfying the constraints:

$$(12) \qquad \sum_s b_{sr} v_s \leq t_r, \qquad r = 1,2,\ldots.R, \quad \text{and}$$

(13) $$v_s \geqq 0, \qquad s = 1,2,\ldots S;$$

moreover, if m and M exist, they are equal, and if $\bar{\bar{u}}^o$ and $\bar{\bar{v}}^o$ are minimizing and maximizing vectors, respectively,

(14) $$\sum_s b_{sr} v_s^o < t_r \qquad \text{implies} \quad u_r^o = 0, \quad \text{and}$$

(15) $$\sum_r b_{sr}\, u_r^o > c_s \qquad \text{implies} \quad v_s^o = 0.$$

This theorem follows directly from the extended Farkas lemma if we note that now (9) and (10) provide $S + R$ constraints to take the place of (5), so that in terms of the previous notation the summation of (6) extends over $s = 1,\ldots S + R$, and the former $\bar{\bar{w}}$ has $S + R$ elements which we will now designate $(w_1,\ldots w_s,\ldots w_S, h_1,\ldots h_r,\ldots h_R)$. Thus (6) now becomes $t_r = h_r + \sum_{s=1}^{S} b_{sr} w_s$; $r = 1,\ldots R$, for some non-negative $\bar{\bar{h}}$ and $\bar{\bar{w}}$. This is obviously equivalent to the requirement that

(16) $$t_r \geqq \sum_{s=1}^{S} b_{sr} w_s, \qquad r = 1,\ldots R, \quad \text{and}$$

$$w_s \geqq 0, \qquad s = 1,\ldots S.$$

But these are precisely the set of constraints of (12, 13) with appropriate substitution of v for w, and conversely a similar substitution shows that the existence of a $\bar{\bar{u}}$ satisfying the constraints (9, 10) is necessary and sufficient for the existence of an upper bound for the maximand of (11). As the admissible $\bar{\bar{v}}$ form a closed set, the existence of an upper bound implies the existence of a maximum, and similarly for the minimum.

Indeed, putting $k = -t_o$ and $c_s = - b_{so}$ in (4) and (7) shows that $\sum_r t_r u_r \geqq k$, for all admissible $\bar{\bar{u}}$ (i.e., those satisfying (9,10)), for just those $\underline{\underline{k}}$ satisfying $k \leq \sum_s c_s v_s$ for some admissible $\bar{\bar{v}}$. The maximum possible of such k, i.e., $k = M = \sum_s c_s v_s^o$ thus furnishes the greatest lower bound and hence the minimum for $\sum_r t_r u_r$;

$$m = \sum_r t_r u_r^o = \sum_s c_s v_s^o = M. \tag{17}$$

Further, if now we use (12) in (17) we get

$$\sum_s c_s v_s^o = \sum_r t_r u_r^o \geqq \sum_s \sum_r v_s b_{sr} u_r^o$$

or

$$\sum_s v_s^o \left(\sum_r b_{sr} u_r^o - c_s\right) \leq 0. \tag{18}$$

But since from (9, 10) both factors in these products are non-negative, we must have

$$v_s^o \left(\sum_r b_{sr} u_r^o - c_s\right) = 0, \qquad s = 1, 2, \ldots S \tag{19}$$

yielding the results of (14),(15) being derived similarly.

4. <u>Lagrangian Multiplier Rules</u>

From this duality theorem two different Lagrangian multiplier rules can be devised that are of some significance. The more significant of these transforms the constrained maximum problem into an unconstrained orthogonal saddle point problem. Form

$$F(\bar{\bar{u}},\bar{\bar{v}}) = \sum_r t_r u_r + \sum_s c_s v_s - \sum_s \sum_r b_{sr} u_r v_s,$$

where now the v_i are introduced as Lagrangian multipliers. If we put $\bar{\bar{u}}^o$ for the solution of the minimum problem of (8, 9, 10), and $\bar{\bar{v}}^o$ for its dual, then we have immediately that

$$F(\bar{\bar{u}}^o, \bar{\bar{v}}^o) = \sum_r t_r u_r^o = \sum_s c_s v_s^o = \sum_s \sum_r b_{sr} u_r^o v_s^o = m = M.$$

$$(20) \quad \begin{cases} F(\bar{\bar{u}},\bar{\bar{v}}^o) = \sum_s c_s v_s^o + \sum_r u_r (t_r - \sum_s b_{sr} v_s^o) \geqq F(\bar{\bar{u}}^o,\bar{\bar{v}}^o) \\ \text{and similarly } F(\bar{\bar{u}}^o,\bar{\bar{v}}^o) \geqq F(\bar{\bar{u}}^o,\bar{\bar{v}}) \end{cases}$$

for all non-negative $\bar{\bar{u}}$ and $\bar{\bar{v}}$ so that $(\bar{\bar{u}}^o, \bar{\bar{v}}^o)$ is an orthogonal saddle point of F over the positive orthant. Conversely, if $(\bar{\bar{u}}', \bar{\bar{v}}')$ is a saddle point of F satisfying (20), it is easy to see that $\bar{\bar{u}}'$ is a minimum of (8, 9, 10) and $\bar{\bar{v}}'$ a maximum of the dual problem (11, 12, 13).

A somewhat less significant transformation of the minimizing problem occurs if we put

$$G(\bar{\bar{u}},\bar{\bar{v}}) = \sum_r t_r u_r - \sum_s v_s [c_s - \sum_r b_{sr} u_r]$$

In this case the solution of (8, 9, 10) can be obtained by maximizing G, with respect to both $\bar{\bar{u}}$ and $\bar{\bar{v}}$, not over the entire positive orthant, but only over that part where $v_s = 0$ whenever $c_s - \sum_r b_{sr} u_r < 0$. In effect the minimum point so found is a point where G is non-decreasing in the positive orthant for changes in either $\bar{\bar{v}}$ or $\bar{\bar{u}}$ separately, but not necessarily for changes in both together, e.g. for increases in v_i coupled with changes in $\bar{\bar{u}}$ that make $c_i - \sum_r b_{ir} u_r$ positive.