1967

The Generalized Inverse in Linear Programming: An Intersection Projection Method

L. Duane Pyle

Report Number:
67-013

"The Generalized Inverse in Linear Programming

An Intersection Projection Method"

by L. Duane Pyle

1 August 67
CSD TR 13

"The Generalized Inverse in Linear Programming
An Intersection Projection Method"

by L. Duane Pyle

1. Introduction. This is one of a series of papers dealing with different aspects of the same central theme: how the generalized inverse A^+ [8] of a matrix and related constructs may be used in connection with linear programming to gain greater understanding of underlying mathematical structure and to provide computational techniques for solution.

Much of the material being presented has previously been rather inaccessible, the principal references being Ph.D. theses [1] and [9] and papers given at the 1959 RAND Symposium on Mathematical Programming held in Los Angeles (abstracts were published formally) and at the 1964 International Conference on Mathematical Programming held in London (abstracts were published informally).

In this paper the structure provided by a theorem characterizing duality in terms of orthogonality [10] is used to represent the direct (equalities) form of the linear programming problem together with a problem equivalent to its dual, as a restricted, fixed point problem $Pz = z \geq \theta$ where P is a perpendicular projection matrix. Convergence of a "Kaczmarz-like" intersection projection method for numerical solution of such problems is established. Consideration of computational techniques prompted a more general study leading to the discovery of a closed form for intersection projection matrices based on the Wynn ϵ_n - Algorithem [11]. Further investigations of variations of this form are currently underway [12].

In another paper in this series [2], the work of R. Cline regarding representations of the generalized inverse of a partitioned matrix $[U,V]$ is applied to the classical, multidimensional transportation problem, yielding the closed form of the projection matrix P for such problems. Experience with computer programs employing Cline's results to compute solutions of two dimensional transportation problems using an accelerated intersection projection method is also described in [2].

2. Notation, definitions and supporting theorems. To maintain completeness, certain definitions and theorems will be summarized below. For motivation and other supporting details the reader is referred to reference [10] with which notational agreement has been maintained.

Consider the direct (equalities) form of the linear programming problem (to be abbreviated DLPP):

Definition 2.1: Maximize (x,c) such that $Ax = b$; $x \geq \theta$.

(θ is a vector of zeros.)

Definition 2.2: Let $e^{(1)},\ldots,e^{(n)}$ be any set of orthonormal eigenvectors of $I - A^{+}A$; the set $e^{(1)},\ldots e^{(q)}$ corresponding to eigenvalue $\lambda = 1$, the set $e^{(q+1)},\ldots,e^{(n)}$ corresponding to eigenvalue $\lambda = 0$.

Definition 2.3: Let $\bar{e}^{(1)},\ldots,\bar{e}^{(n)}$ be a specific set of such eigenvectors chosen such that

$$\bar{e}^{(1)} = \frac{1}{\| c^{q} \|} c^{q} \quad \text{where} \quad c^{q} = (I - A^{+}A)c$$

and

$$\bar{e}^{(q+1)} = \frac{1}{\| A^{+}b \|} A^{+}b \quad \text{where} \quad \| z \| = \sqrt{(z,z)} = \left(\sum_{i=1}^{n} z_i^2\right)^{\frac{1}{2}} .$$

As in reference [10] it will be assumed that $||A^{+}b|| \neq \theta$, $||c^{q}|| \neq \theta$ and that $2 \leq q \leq n - 1$ since $q = 0$ implies A is nonsingular and $q = n$ implies $A = 0$. If $q = 1$ the null space of A is a 1-dimensional linear manifold and then solution of the related DLPP is trivial.

<u>Definition 2.4</u>: The vector $\bar{x}$ is said to be feasible for the DLPP if both $\bar{x} \geq \theta$ and $A\bar{x} = b$.

<u>Definition 2.5</u>: The feasible vector $\bar{x}$ is said to be <u>optimal</u> for the DLPP if $(\bar{x},c) \geq (x,c)$ for all feasible x.

Questions relating to existence of solutions are not to be considered; it will be assumed throughout that the set of feasible solutions of the DLPP of Definition 2.1 is non-empty and bounded and thus that both the DLPP and its dual have at least one (finite) optimal solution. A knowledge of standard notions and manipulations concerning vectors and matrices having real elements defined on E^n will be assumed, where E^n is real, Euclidean n-dimensional vector space. The reader is referred to [5] for a detailed treatment. Relevant properties of the (Moore-Penrose-Bjerhammar) generalized inverse are given in [8].

<u>Definition 2.6</u>: Two linear programming problems, I and II, will be said to be <u>equivalent</u> if the set of feasible solutions of I coincides with the set of feasible solutions of II, and if, also, the set of optimal solutions of I coincides with the set of optimal solutions of II.

<u>Definition 2.7</u>:

<u>Problem A</u>: Minimize $(x, -c^q)$ such that $Ex = \beta$, $x \geq \theta$

where $c^q = (I - A^+A)\,c$

$$E = \begin{bmatrix} \longleftarrow e^{(q+1)T} \longrightarrow \\ \vdots \\ \longleftarrow e^{(n)T} \longrightarrow \end{bmatrix} \quad x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \quad \beta = \begin{bmatrix} (A^+b,\ e^{(q+1)}) \\ \vdots \\ (A^+b,\ e^{(n)}) \end{bmatrix} = E\,A^+b$$

<u>Problem B</u>: Minimize (y, A^+b) such that $\bar{E}y = \bar{\beta}$, $y \geq \theta$

where

$$\bar{E} = \begin{bmatrix} \longleftarrow e^{(1)T} \longrightarrow \\ \vdots \\ \longleftarrow e^{(q)T} \longrightarrow \end{bmatrix} \quad y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \quad \bar{\beta} = \begin{bmatrix} (-c^q,\ e^{(1)}) \\ \vdots \\ (-c^q,\ e^{(q)}) \end{bmatrix} = \bar{E}\,(-c^q)$$

<u>Problem A'</u>: Minimize $(\alpha, \bar{\beta})$ such that $\bar{E}^T \alpha + A^+b \geq \theta$

where

$\alpha = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_q \end{bmatrix}$ is unrestricted, $\bar{E}$ and $\bar{\beta}$ defined as in problem B.

<u>Problem B'</u>: Minimize (γ, β) such that $E^T \gamma - c^q \geq 0$

where

$\gamma = \begin{bmatrix} \gamma_{q+1} \\ \vdots \\ \gamma_n \end{bmatrix}$ is unrestricted, E, β and c^q defined as in problem A.

Proofs for the following theorems are given in [10]:

<u>Theorem 2.1</u>: Problem A and the DLPP of definition 2.1 are equivalent; problems A and B' are essentially duals as are problems A' and B.

<u>Theorem 2.2</u>: If $\bar{x}$ is optimal for problem A, then $\bar{\alpha} = \bar{E}\,\bar{x}$ is optimal for problem A'. If $\bar{\alpha}$ is optimal for problem A', then $\bar{x} = A^{+}b + \bar{E}^{T}\bar{\alpha}$ is optimal for problem A.

<u>Theorem 2.3</u>: If $\bar{y}$ is optimal for problem B, then $\bar{\gamma} = E\,\bar{y}$ is optimal for problem B'. If $\bar{\gamma}$ is optimal for problem B', then $\bar{y} = -c^{q} + E^{T}\bar{\gamma}$ is optimal for problem B.

<u>Theorem 2.4</u>: If $\hat{x}$ and $\hat{y}$ are feasible for problems A and B, respectively, then $(\hat{x}, \hat{y}) = 0$ if and only if $\hat{x}$ and $\hat{y}$ are optimal.

<u>Theorem 2.5</u>: If $Ax = b$ is solvable and

$$(I - A^{+}A + \bar{e}^{(q+1)}\,\bar{e}^{(q+1)T})\,\tilde{x} = \tilde{x} \neq \theta$$

where $(I - A^{+}A)\,\tilde{x} \neq \tilde{x}$ and $A^{+}b \neq \theta$

then $Ax = b$ where

$$x = \frac{(A^{+}b,\ A^{+}b)}{(\tilde{x},\ A^{+}b)}\,\tilde{x}\,.$$

<u>Theorem 2.6</u>: $I - A^{+}A = \bar{E}^{T}\,\bar{E}$ and $A^{+}A = E^{T}\,E$

3. <u>A restricted fixed point problem</u>. Let the eigenvectors $\bar{e}^{(1)},\ldots,\bar{e}^{(n)}$ be defined as in Section 2 and consider optimal solutions $\hat{x},\hat{y}$ to problems A and B, respectively. By Theorems 2.2, 2.3 and 2.4

$$0 = (\hat{x},\hat{y}) = (A^{+}b + \sum_{i=1}^{q} \hat{\alpha}_i \bar{e}^{(i)}, -c^q + \sum_{i=q+1}^{n} \hat{\gamma}_i \bar{e}^{(i)})$$

$$= (||A^{+}b||\, \bar{e}^{(q+1)} + \sum_{i=1}^{q} \hat{\alpha}_i \bar{e}^{(i)}, -||c^q||\, \bar{e}^{(1)} + \sum_{i=q+1}^{n} \hat{\gamma}_i \bar{e}^{(i)})$$

$$= \hat{\gamma}_{q+1}\, ||A^{+}b|| - \hat{\alpha}_1\, ||c^q||.$$

Thus, for optimal $\hat{x},\hat{y}$, the corresponding $\hat{\alpha}_1$ and $\hat{\gamma}_{q+1}$ must satisfy the relation

$$\hat{\gamma}_{q+1} = \frac{||c^q||}{||A^{+}b||}\, \hat{\alpha}_1$$

Conversely, suppose $\hat{x} = A^{+}b + \sum_{i=1}^{q} \hat{\alpha}_i \bar{e}^{(i)}$ and $\hat{y} = -c^q + \sum_{i=q+1}^{n} \hat{\gamma}_i \bar{e}^{(i)}$

are feasible for problems A and B, respectively, and that

$\hat{\gamma}_{q+1} = \dfrac{||c^q||}{||A^{+}b||}\, \hat{\alpha}_1$. It then follows that $(\hat{x},\hat{y}) = 0$ which, by Theorem 2.4, implies $\hat{x}$ and $\hat{y}$ are optimal for problems A and B, respectively. Since for any feasible $\hat{x},\hat{y}$, $(\hat{x},\bar{e}^{(1)}) = \hat{\alpha}_1$ and $(\hat{y},\bar{e}^{(q+1)}) = \hat{\gamma}_{q+1}$, thus if

$$\hat{\gamma}_{q+1} = (\hat{y}, \bar{e}^{(q+1)}) = \frac{||c^q||}{||A^{+}b||}\, (\hat{x},\bar{e}^{(1)}) = \frac{||c^q||}{||A^{+}b||}\, \hat{\alpha}_1$$

then $\hat{x}$ and $\hat{y}$ are optimal. As a consequence of the above reasoning, it follows that:

<u>Theorem 3.1</u>: Any solution of the following system of $n+2$ equations in $2n+1$ unknowns which is non-negative in its first $2n$ components, provides optimal solutions to problems A and B, respectively:

$$\begin{bmatrix} E & 0 & \theta \\ 0 & \bar{E} & \theta \\ \theta^T & e^{(q+1)T} & -\dfrac{||c^q||}{||A^+b||} \\ \bar{e}^{(1)T} & \theta^T & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ \alpha_1 \end{bmatrix} = \begin{bmatrix} \theta \\ \bar{\beta} \\ 0 \\ 0 \end{bmatrix}$$

where the various submatrices and subvectors have been defined in Section 2.

Proof: The first n equations of the system exhibited, together with the stated non-negativity condition, simply represent a matrix formulation of the feasibility conditions imposed by problems A and B, respectively. The last two equations add the necessary and sufficient conditions of Theorem 2.4 for feasible solutions x and y to problems A and B to be optimal.

In order to avoid carrying along the constant of proportionality $\dfrac{||c^q||}{||A^+b||}$, consider the following linear programming problems:

(3.1) Maximize (x,c) such that $Ax = b$, $x \geq \theta$;

(3.2) Minimize $(x,-c^q)$ such that $Ex = \beta$, $x \geq \theta$;

(3.3) Minimize $(x,-\bar{e}^{(1)})$ such that $Ex = \beta$, $x \geq \theta$.

Letting $x = ||A^+b||\tilde{x}$ in (3.3) gives

(3.4) Minimize $||A^+b||(\tilde{x},-\bar{e}^{(1)})$ such that

$$E\tilde{x} = \frac{1}{||A^+b||}\beta = E\left(\frac{1}{||A^+b||}A^+b\right) = E\bar{e}^{(q+1)}.$$

Finally,

(3.5) Minimize $(\tilde{x}, -\bar{e}^{(1)})$ such that $E\tilde{x} = E\bar{e}^{(q+1)}$, $\tilde{x} \geq \theta$.

That is, if $\tilde{\tilde{x}}$ is an optimal solution of the DLPP (3.5) then $||A^+b||\tilde{\tilde{x}}$ is an optimal solution of (3.1) and optimal values of the functionals are related as follows:

$$\text{Maximum of } (x,c) = (A^+b,c)-||c^q||\ ||A^+b||\ [\text{Minimum of } (\tilde{x},e^{(1)})].$$

Designating (3.5) as Problem $\tilde{A}$, the problem analogous to Problem B has the form

<u>Problem $\tilde{B}$</u>: Minimize $(\tilde{y},\bar{e}^{(q+1)})$ such that $\bar{E}\tilde{y} = \bar{\beta}$, $\tilde{y} \geq \theta$

where $\bar{\beta} = \bar{E}(-\bar{e}^{(1)})$.

Then if $\tilde{x} = \bar{e}^{(q+1)} + \sum_{i=1}^{n} \tilde{\alpha}_i \bar{e}^{(i)}$ and $\tilde{y} = -\bar{e}^{(1)} + \sum_{i=q+1}^{n} \tilde{\gamma}_i \bar{e}^{(i)}$ are optimal for $\tilde{A}$ and $\tilde{B}$, respectively, $0 = (\tilde{x},\tilde{y}) = \tilde{\gamma}_{q+1} - \tilde{\alpha}_1$.

Finally, letting $\tilde{\alpha}_1 = \tilde{\alpha}_1^+ - \tilde{\alpha}_1^-$, where it is required that $\tilde{\alpha}_1^+ \geq 0$, $\tilde{\alpha}_1^- \geq 0$, the system exhibited in Theorem 3.1 assumes the form

$$(3.6)\quad \tilde{\mathcal{A}}\,\nu \equiv \begin{bmatrix} E & 0 & \theta & \theta \\ 0 & \bar{E} & \theta & \theta \\ \theta^T & \bar{e}^{(q+1)T} & -1 & 1 \\ \bar{e}^{(1)T} & \theta^T & -1 & 1 \end{bmatrix} \begin{bmatrix} \tilde{x} \\ \tilde{y} \\ \tilde{\alpha}_1^+ \\ \tilde{\alpha}_1^- \end{bmatrix} \equiv \begin{bmatrix} E\bar{e}^{(q+1)} \\ \bar{E}(-\bar{e}^{(1)}) \\ 0 \\ 0 \end{bmatrix} \equiv b$$

where it is required that

$$\mathcal{X} = \begin{bmatrix} \tilde{x} \\ \tilde{y} \\ \tilde{\alpha}_1^{+} \\ \tilde{\alpha}_1^{-} \end{bmatrix} \geq \theta$$

Any $\mathcal{X} \geq \theta$ such that $\tilde{\mathcal{Q}}\mathcal{X} = \mathcal{b}$ thus yields optimal solutions to problems A and B, after performing the simple modifications described following relation (3.5).

Observe that the row vectors associated with the coefficient array of the first n equations

$$\begin{bmatrix} E & \underline{0} & \theta & \theta \\ 0 & E & \theta & \theta \end{bmatrix} \mathcal{X} = \begin{bmatrix} E\bar{e}^{(q+1)} \\ E(-\bar{e}^{(1)}) \end{bmatrix}$$

are mutually pairwise orthogonal and of length 1. The row vectors of each of the last two equations

$$\begin{bmatrix} \theta^T & \bar{e}^{(q+1)T} & -1 & 1 \\ \bar{e}^{(1)T} & \theta^T & -1 & 1 \end{bmatrix} \mathcal{X} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

are orthogonal to the row vectors of each of the first n equations, although they are not orthogonal to each other. An equivalent system is obtained by replacing the row vector of the (n+2)nd equation by its component orthogonal to the row vector of the (n+1)st equation:

$$\begin{bmatrix} \bar{e}^{(1)} \\ \theta \\ -1 \\ 1 \end{bmatrix}^T - \frac{2}{\sqrt{3}} \begin{bmatrix} \theta \\ \frac{1}{\sqrt{3}}\bar{e}^{(q+1)} \\ -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}^T = \begin{bmatrix} \bar{e}^{(1)} \\ -\frac{2}{3}\bar{e}^{(q+1)} \\ -\frac{1}{3} \\ \frac{1}{3} \end{bmatrix}^T$$

Note that the (n+1)st and (n+2)nd equations have constant terms equal to zero. Upon normalization, the following system, equivalent to (3.6), is obtained:

$$(3.7)\quad \mathcal{A}x \equiv \left[\begin{array}{c|c|c|c} E & 0 & \theta & \theta \\ 0 & \bar{E} & \theta & \theta \\ \theta^{T} & \frac{1}{\sqrt{3}}\,\bar{e}^{(q+1)T} & -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \sqrt{\frac{3}{5}}\,\bar{e}^{(1)T} & -\frac{2}{3}\sqrt{\frac{3}{5}}\,\bar{e}^{(q+1)T} & -\frac{1}{3}\sqrt{\frac{3}{5}} & \frac{1}{3}\sqrt{\frac{3}{5}} \end{array}\right] x = \left[\begin{array}{c} E\bar{e}^{(q+1)} \\ \bar{E}(-\bar{e}^{(1)}) \\ 0 \\ 0 \end{array}\right] \equiv b$$

The (n+2) rows of $\mathcal{A}$ are mutually orthogonal and of length 1. It is easily verified [9] that if the rows of a matrix U are mutually orthogonal and of length 1, that U is a partial isometry and that for any real partial isometry $U^{+} = U^{T}$. Thus $\mathcal{A}^{+} = \mathcal{A}^{T}$. Now, employing Theorems 2.5 and 2.6, it is an exercise in matrix algebra to prove the following restricted fixed point theorem:

Theorem 3.2: Consider the matrix P defined as follows:

(3.8)

$$P=\begin{bmatrix} I-A^{+}A-\frac{3}{5}\bar{e}^{(1)}\bar{e}^{(1)T}+\frac{1}{2}\bar{e}^{(q+1)}\bar{e}^{(q+1)T} & \frac{2}{5}\bar{e}^{(1)}\bar{e}^{(q+1)T}-\frac{1}{2}\bar{e}^{(q+1)}\bar{e}^{(1)T} & \frac{1}{5}\bar{e}^{(1)} & -\frac{1}{5}\bar{e}^{(1)} \\ \frac{2}{5}\bar{e}^{(q+1)}\bar{e}^{(1)T}-\frac{1}{2}\bar{e}^{(1)}\bar{e}^{(q+1)T} & A^{+}A-\frac{3}{5}\bar{e}^{(q+1)}\bar{e}^{(q+1)T}+\frac{1}{2}\bar{e}^{(1)}\bar{e}^{(1)T} & \frac{1}{5}\bar{e}^{(q+1)} & -\frac{1}{5}\bar{e}^{(q+1)} \\ \frac{1}{5}\bar{e}^{(1)T} & \frac{1}{5}\bar{e}^{(q+1)T} & \frac{3}{5} & \frac{2}{5} \\ -\frac{1}{5}\bar{e}^{(1)T} & -\frac{1}{5}\bar{e}^{(q+1)T} & \frac{2}{5} & \frac{3}{5} \end{bmatrix}$$

P is a (2n+2) by (2n+2) projection matrix with the property that if $Pz = z \geq \theta$, $z \neq \theta$ the first n elements of z provide (essentially) an optimal solution for the DLPP of Definition 3.1. More precisely, let

$$v = \begin{bmatrix} v^{(1)} \\ v^{(2)} \\ v^{(3)} \\ v^{(4)} \end{bmatrix} = \frac{2}{(\alpha^{+}, z)} z$$

where $v^{(1)}$, $v^{(2)}$ are n by 1; $v^{(3)}$ and $v^{(4)}$ are single elements,

$$\alpha^{+} = \begin{bmatrix} \bar{e}^{(q+1)} \\ -\bar{e}^{(1)} \\ 0 \\ 0 \end{bmatrix} \qquad \text{and} \quad Pz = z \geq \theta,\ z \neq \theta.$$

Then either (a) $\hat{x} = ||A^+b||\ v^{(1)}$ is an optimal solution for the DLPP (3.1), where Maximum $(x,c) = (A^+b,c) + ||c^q||\ ||A^+b||\ (v^{(3)} - v^{(4)})$, or (b) the set of feasible solutions is unbounded.

Proof: Upon direct substitution, obtain

$$
a^+a = \begin{bmatrix} E^T & 0 & \theta & \sqrt{\frac{3}{5}}\,\bar{e}^{(1)} \\ 0 & \bar{E}^T & \frac{1}{\sqrt{3}}\,\bar{e}^{(q+1)} & -\frac{2}{3}\sqrt{\frac{3}{5}}\,\bar{e}^{(q+1)} \\ \theta^T & \theta^T & -\frac{1}{\sqrt{3}} & -\frac{1}{3}\sqrt{\frac{3}{5}} \\ \theta^T & \theta^T & \frac{1}{\sqrt{3}} & \frac{1}{3}\sqrt{\frac{3}{5}} \end{bmatrix}
\begin{bmatrix} E & 0 & \theta & \theta \\ 0 & \bar{E} & \theta & \theta \\ \theta^T & \frac{1}{\sqrt{3}}\,\bar{e}^{(q+1)T} & -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \sqrt{\frac{3}{5}}\,\bar{e}^{(1)T} & -\frac{2}{3}\sqrt{\frac{3}{5}}\,\bar{e}^{(q+1)T} & -\frac{1}{3}\sqrt{\frac{3}{5}} & \frac{1}{3}\sqrt{\frac{3}{5}} \end{bmatrix}
$$

$$
= \begin{bmatrix} E^TE + \frac{3}{5}\,\bar{e}^{(1)}\bar{e}^{(1)T} & -\frac{2}{5}\,\bar{e}^{(1)}\,\bar{e}^{(q+1)T} & -\frac{1}{5}\,\bar{e}^{(1)} & \frac{1}{5}\,\bar{e}^{(1)} \\ -\frac{2}{5}\,\bar{e}^{(q+1)}\,\bar{e}^{(1)T} & \bar{E}^T\bar{E} + \frac{3}{5}\,\bar{e}^{(q+1)}\,\bar{e}^{(q+1)T} & -\frac{1}{5}\,\bar{e}^{(q+1)} & \frac{1}{5}\,\bar{e}^{(q+1)} \\ -\frac{1}{5}\,\bar{e}^{(1)T} & -\frac{1}{5}\,\bar{e}^{(q+1)T} & \frac{2}{5} & -\frac{2}{5} \\ \frac{1}{5}\,\bar{e}^{(1)T} & \frac{1}{5}\,\bar{e}^{(q+1)T} & -\frac{2}{5} & \frac{2}{5} \end{bmatrix}
$$

Since, by Theorem 2.6, $I - A^+A = \bar{E}^T \bar{E}$ and $A^+A = E^T E$,

$$I - \mathcal{A}^+\mathcal{A} = \begin{bmatrix} I - A^+A - \frac{3}{5}\bar{e}^{(1)}\bar{e}^{(1)T} & \frac{2}{5}\bar{e}^{(1)}\bar{e}^{(q+1)T} & \frac{1}{5}\bar{e}^{(1)} & -\frac{1}{5}\bar{e}^{(1)} \\ \frac{2}{5}\bar{e}^{(q+1)}\bar{e}^{(1)T} & A^+A - \frac{3}{5}\bar{e}^{(q+1)}\bar{e}^{(q+1)T} & \frac{1}{5}\bar{e}^{(q+1)} & -\frac{1}{5}\bar{e}^{(q+1)} \\ \frac{1}{5}\bar{e}^{(1)T} & \frac{1}{5}\bar{e}^{(q+1)T} & \frac{3}{5} & \frac{2}{5} \\ -\frac{1}{5}\bar{e}^{(1)T} & -\frac{1}{5}\bar{e}^{(q+1)T} & \frac{2}{5} & \frac{3}{5} \end{bmatrix}$$

Similarly,

$$\mathcal{A}^+ b = \begin{bmatrix} E^T & 0 & \theta & \sqrt{\frac{3}{5}}\,\bar{e}^{(1)} \\ 0 & \bar{E}^T & \frac{1}{\sqrt{3}}\bar{e}^{(q+1)} & -\frac{2}{3}\sqrt{\frac{3}{5}}\,\bar{e}^{(q+1)} \\ \theta^T & \theta^T & -\frac{1}{\sqrt{3}} & -\frac{1}{3}\sqrt{\frac{3}{5}} \\ \theta^T & \theta^T & \frac{1}{\sqrt{3}} & \frac{1}{3}\sqrt{\frac{3}{5}} \end{bmatrix} \begin{bmatrix} E\,\bar{e}^{(q+1)} \\ \bar{E}(-e^{(1)}) \\ 0 \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} E^T E\,\bar{e}^{(q+1)} \\ \bar{E}^T \bar{E}(-\bar{e}^{(1)}) \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} A^+A\,\bar{e}^{(q+1)} \\ (I - A^+A)(-\bar{e}^{(1)}) \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \bar{e}^{(q+1)} \\ -\bar{e}^{(1)} \\ 0 \\ 0 \end{bmatrix}$$

Thus

$$\frac{1}{(\mathcal{A}^+b, \mathcal{A}^+b)} \mathcal{A}^+b(\mathcal{A}^+b)^T = \frac{1}{2} \begin{bmatrix} \bar{e}^{(q+1)}\bar{e}^{(q+1)T} & -\bar{e}^{(q+1)}\bar{e}^{(1)T} & \theta & \theta \\ -\bar{e}^{(1)}e^{(q+1)T} & \bar{e}^{(1)}e^{(1)T} & \theta & \theta \\ \theta^T & \theta^T & 0 & 0 \\ \theta^T & \theta^T & 0 & 0 \end{bmatrix}$$

Combining, obtain $P = I - \mathcal{A}^+\mathcal{A} + \frac{1}{(\mathcal{A}^+b, \mathcal{A}^+b)} \mathcal{A}^+b(\mathcal{A}^+b)^T$ as exhibited in relation (3.8). If $(I-\mathcal{A}^+\mathcal{A})z \nleq z$, Case (a) follows from Theorems 2.5 and 3.1. If $(I-\mathcal{A}^+\mathcal{A})z = z \neq \theta$, $z \geq \theta$, then suppose x is a vector such that $\mathcal{A}x = b$, $x \geq \theta$. Thus, $\tilde{z} = x + \mu z \geq \theta$ for all $\mu > 0$ and $\mathcal{A}\tilde{z} = b$. By Theorem 3.1, the subcomponents of $\tilde{z}$ provide optimal solutions to problems $\tilde{A}$ and $\tilde{B}$, respectively, thus to A and B after the appropriate modifications. Since the elements in $\tilde{z}$ corresponding to positive elements in z may be made arbitrarily large by appropriate choices of μ, the set of feasible solutions to problem A is then unbounded, which is Case (b) (note that this is contrary to the assumptions made in Section 2).

4. <u>An intersection projection method</u>. It is well-known [7] that if P_1 and P_2 are two perpendicular projection matrices with ranges M_1 and M_2, respectively, then

<u>Theorem 4.1</u>: $\lim_{V \to \infty} (P_1P_2)^V = \lim_{V \to \infty} (P_2P_1P_2)^V$ exists and in each case is a perpendicular projection matrix with range $M_1 \cap M_2$.

If $\tilde{x}$ and $\tilde{y}$ are optimal solutions for problems $\tilde{A}$ and $\tilde{B}$ of Section 3, it follows from Theorem 2.4 that the non-zero elements of $\tilde{x}$ and $\tilde{y}$ are interlaced in the sense that if $\tilde{x}_i \neq 0$ then $\tilde{y}_i = 0$ and if $\tilde{y}_i \neq 0$ then $\tilde{x}_i = 0$. Thus, if the indices of the zero elements in $\tilde{x}$ and $\tilde{y}$ are known, the problem of determining $z \neq \theta$ such that $z = Pz$, $z \geq \theta$, is solved by application of Theorem 4.1 with $P_1 = P$ and $P_2 = \tilde{I}$ where $\tilde{I}$ is obtained from the $(2n + 2)$ by $(2n + 2)$ identity matrix by replacing diagonal 1's by 0's in positions corresponding to the zeros in $\tilde{x}$ and $\tilde{y}$. It will be shown that even when the interlace pattern is unknown, a variation of this procedure provides a sequence converging to $z = Pz$, $z \geq \theta$, $z \neq \theta$:

Suppose P is a (real) t by t perpendicular projection matrix (i.e., $P^2 = P^T = P$) having an eigenvector $\xi^{(1)} \geq \theta$ corresponding to the eigenvalue $\lambda = 1$, where it is assumed that $P \neq 0$, $P \neq I$ and $||\xi^{(1)}|| = 1$. It may be remarked that in the case described in the preceding paragraph, it cannot be that $\xi^{(1)} > \theta$ although no use will be made of this property. Let $\{\xi^{(i)}\}$ $(i = 1,\ldots,t)$ be an orthonormal set of eigenvectors of P forming a basis for real, Euclidean t-dimensional vector space, E^t, where $\xi^{(1)},\ldots,\xi^{(g)}$ correspond to $\lambda = 1$, $\xi^{(g+1)},\ldots,\xi^{(t)}$ correspond to $\lambda = 0$.

<u>Definition 4.1</u>: If $x^{(k)}$ is a real vector with elements $x_1^{(k)},\dots,x_t^{(k)}$, then $I^{(k)}$ is that t by t matrix obtained by replacing the "i"th diagonal element of the t by t identity matrix by zero if $x_i^{(k)} < 0$ for $(i = 1,\dots,t)$.

Let $w^{(k)} = I^{(k)}x^{(k)}$ and $x^{(k+1)} = Pw^{(k)}$ for $(k = 0,1,\dots)$, where $x_i^{(0)} = 1$ for $(i = 1,\dots,t)$. With this particular choice of an initial vector, it will be shown that

<u>Theorem 4.2</u>: Either for some $(k = 1,2,\dots)$

(I) $x^{(k)} = Px^{(k)}$, $x^{(k)} \geq \theta$, $x^{(k)} \neq \theta$, or

(II) $\lim\limits_{k \to \infty} \{x^{(k)}\} = z = Pz$, $z \geq \theta$, $z \neq \theta$.

<u>Proof</u>: Since $x_i^{(0)} = 1$ for $(i = 1,\dots,t)$, $I^{(0)} = I$ and $w^{(0)} = x^{(0)}$. Now, $w^{(0)} = \sum\limits_{i=1}^{t} (w^{(0)}, \xi^{(i)})\xi^{(i)}$ where $(w^{(0)},\xi^{(1)}) > 0$, therefore

$x^{(1)} = Pw^{(0)} = P\sum\limits_{i=1}^{t} (w^{(0)},\xi^{(i)})\xi^{(i)} = \sum\limits_{i=1}^{g} (w^{(0)}),\xi^{(i)})\xi^{(i)}$, thus

$(x^{(1)},\xi^{(1)}) = (w^{(0)},\xi^{(1)}) > 0$. It follows that $x^{(1)} \neq \theta$. If $x^{(1)} \geq \theta$ then Case I holds with $k = 1$, since $x^{(1)} = Pw^{(0)} = P^2w^{(0)} = Px^{(1)}$. Otherwise $x^{(1)}$ has at least one element which is strictly negative. Since $\xi^{(1)} \geq \theta$, it follows that $(x^{(1)},\xi^{(1)}) \leq (I^{(1)}x^{(1)},\xi^{(1)}) = (w^{(1)},\xi^{(1)})$, thus $(w^{(1)},\xi^{(1)}) \geq (x^{(1)},\xi^{(1)}) = (w^{(0)},\xi^{(1)}) > 0$, which implies that $w^{(1)} \neq \theta$. Now, $w^{(1)} = \sum\limits_{i=1}^{t} (w^{(1)},\xi^{(i)})\xi^{(i)}$ and $x^{(2)} = Pw^{(1)} = \sum\limits_{i=1}^{g} (w^{(1)},\xi^{(i)})\xi^{(i)}$, thus

$(x^{(2)},\xi^{(1)}) = (w^{(1)},\xi^{(1)}) > 0$ which implies $x^{(2)} \neq \theta$ and, as before, unless Case I holds with $k = 2$, it follows that $(x^{(2)},\xi^{(1)}) \leq (w^{(2)},\xi^{(1)})$, thus $(w^{(2)},\xi^{(1)}) \geq (x^{(2)},\xi^{(1)}) = (w^{(1)},\xi^{(1)}) \geq (x^{(1)},\xi^{(1)}) = (w^{(0)},\xi^{(1)}) > 0$. Continuing in this manner, either Case I holds for some positive integer k, or two infinite sequences $\{x^{(k)}\}, \{w^{(k)}\}$ are generated having the properties $(x^{(k+1)},\xi^{(1)}) = (w^{(k)},\xi^{(1)}) \geq (x^{(k)},\xi^{(1)}) = (w^{(k-1)},\xi^{(1)})$ for $(k=1,2,\ldots)$ where $(w^{(0)},\xi^{(1)}) > 0$, $w^{(k-1)} \neq \theta$, $x^{(k)} \neq \theta$, and $Pw^{(k-1)} \neq w^{(k-1)}$. The last relation follows from the fact that if $x^{(k)} = Pw^{(k-1)} = w^{(k-1)}$, then $x^{(k)} = Pw^{(k-1)} = P^2w^{(k-1)} = Px^{(k)}$, which is Case I. But $Pw^{(k)} \neq w^{(k)}$ for $(k=0,1,\ldots)$ implies that for all such k $w^{(k)} = \sum_{i=1}^{t} (w^{(k)},\xi^{(i)})\xi^{(i)}$,

where $(w^{(k)},\xi^{(i)}) \neq 0$ for at least one value of i in the index set $(i=g+1,\ldots,t)$, from which it follows that

$$||x^{(k+1)}-\xi^{(1)}|| = ||\sum_{i=1}^{g} (w^{(k)},\xi^{(i)})\xi^{(i)}-\xi^{(1)}|| < ||\sum_{i=1}^{t} (w^{(k)},\xi^{(i)})\xi^{(i)}-\xi^{(1)}|| = ||w^{(k)}-\xi^{(1)}||$$

for $(k=0,1,\ldots)$. Since each $x^{(k)}$ possesses at least one negative element,

$$||w^{(k)}-\xi^{(1)}|| = ||I^{(k)}x^{(k)}-\xi^{(1)}|| < ||x^{(k)}-\xi^{(1)}||.$$

The sequence of non-negative numbers $\{||w^{(k)}-\xi^{(1)}||\}$ is strictly monotone decreasing, therefore $\lim_{k\to\infty} \{||w^{(k)}-\xi^{(1)}||\}$ exists. Denote this limit by the value d. If $d = 0$ then $\lim_{k\to\infty} \{w^{(k)}\} = \lim_{k\to\infty} \{x^{(k)}\} = \xi^{(1)}$ and Case II holds with $z = \xi^{(1)}$. If $d > 0$, then there must exist a subsequence $\{\hat{w}^{(k)}\}$ such that $\lim_{k\to\infty} \{\hat{w}^{(k)}\} = v^{(0)} \geq \theta$, where $||v^{(0)}-\xi^{(1)}|| = d$. There are then two possibilities:

(i) $v^{(0)} = \theta$

(ii) $v^{(0)} \neq \theta$

(i) If $v^{(0)} = \theta$, then $\lim_{k\to\infty} (\hat{w}^{(k)},\xi^{(i)}) = 0$ for $(i=1,\ldots,t)$. In particular $\lim_{k\to\infty} (\hat{w}^{(k)},\xi^{(1)}) = 0$. But it was previously established that the sequence $\{(w^{(k)},\xi^{(1)})\}$ is monotone non-decreasing and bounded by the positive quantity $(w^{(0)},\xi^{(1)})$. The assumption that $v^{(0)} = \theta$ thus leads to a contradiction.

(ii) If $v^{(0)} \neq \theta$, then either $Pv^{(0)} = v^{(0)}$, in which event Case II holds with $z = v^{(0)}$, or the entire process may be repeated with $x^{(0)}$ replaced by $v^{(0)}$. Since $v^{(0)} \geq \theta$, $I^{(0)} = I$ and $w^{(0)} = v^{(0)}$, thus $Pv^{(0)} \neq v^{(0)}$ implies $v^{(1)} = Pw^{(0)} = Pv^{(0)} \neq w^{(0)}$, hence

$$||v^{(1)} - \xi^{(1)}|| < ||w^{(0)} - \xi^{(1)}|| = ||v^{(0)} - \xi^{(1)}|| = d.$$

Now consider the sequence $\{\hat{x}^{(k)}\}$ where $\hat{x}^{(k+1)} = P\hat{w}^{(k)}$. The sequence $\{\hat{w}^{(k)}\}$ converges to $v^{(0)}$. The sequence $\{\hat{x}^{(k)}\}$ is a subsequence of the sequence $\{x^{(k)}\}$, and thus $d \leq ||\hat{x}^{(k+1)} - \xi^{(1)}|| < ||\hat{x}^{(k)} - \xi^{(1)}||$ for $(k=0,1,\ldots)$. Let $d_1 = ||v^{(1)} - \xi^{(1)}||$. It has been shown above that $d_1 < d$. Now $\lim\limits_{k \to \infty} \{\hat{x}^{(k)}\} = \lim\limits_{k \to \infty} \{P\hat{w}^{(k-1)}\} = P \lim\limits_{k \to \infty} \{\hat{w}^{(k-1)}\} = Pv^{(0)} = Pw^{(0)} = v^{(1)}$, thus $\lim\limits_{k \to \infty} \{||\hat{x}^{(k)} - \xi^{(1)}||\} = ||v^{(1)} - \xi^{(1)}|| = d_1 < d$. But $\lim\limits_{k \to \infty} \{||\hat{x}^{(k)} - \xi^{(1)}||\} = d$, which provides a contradiction.

In summary it has been shown that either Case I holds or the sequence $\{w^{(k)}\}$ converges to $\xi^{(1)}$ thus satisfying Case II, or the sequence $\{w^{(k)}\}$ possesses a subsequence $\{\hat{w}^{(k)}\}$ which converges to a vector $v^{(0)}$ which has the properties required in Case II. But the appropriate portions of the previous development may now be repeated with $v^{(0)}$ in place of $\xi^{(1)}$, in which case it follows that $\lim\limits_{k \to \infty} \{||w^{(k)} - v^{(o)}||\} = \lim\limits_{k \to \infty} \{||\hat{w}^{(k)} - v^{(0)}||\} = 0$, and thus that $\lim\limits_{k \to \infty} \{w^{(k)}\} = v^{(0)}$, where $v^{(0)} = Pv^{(0)}$, $v^{(0)} \geq \theta$, $v^{(0)} \neq \theta$, which completes the proof of the theorem.

5. Numerical example and concluding remarks. To illustrate the results given in Section 4, consider the following linear programming problem:

Maximize (x,c) where $Ax = b$, $x \geq \theta$

and

$$A = \begin{bmatrix} 1 & 1 & 2 \\ 1 & -1 & 1 \end{bmatrix} \quad b = \begin{bmatrix} 10 \\ 7 \end{bmatrix} \quad c = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} .$$

Then

$$A^+ = \frac{1}{14} \begin{bmatrix} 1 & 4 \\ 5 & -8 \\ 4 & 2 \end{bmatrix} \quad A^+b = \frac{1}{7} \begin{bmatrix} 19 \\ -3 \\ 27 \end{bmatrix} \quad c^q = \frac{1}{14} \begin{bmatrix} 3 \\ 1 \\ -2 \end{bmatrix}$$

$$\bar{e}^{(1)} = \begin{bmatrix} 0.8018 \\ 0.2673 \\ -0.5345 \end{bmatrix} \quad \bar{e}^{(q+1)} = \begin{bmatrix} 0.5731 \\ -0.0905 \\ 0.8145 \end{bmatrix}$$

$$I - A^+A = \frac{1}{14} \begin{bmatrix} 9 & 3 & -6 \\ 3 & 1 & -2 \\ -6 & -2 & 4 \end{bmatrix}$$

$$||A^+b|| = 4.7359$$

$$||c^q|| = 0.2673$$

$$(A^+b,c) = 16\tfrac{4}{7}$$

$$P = I - A^+A + \frac{1}{(A^+b, A^+b)} A^+b\,(A^+b)^T$$

$$= \begin{bmatrix} 0.4214 & 0.0597 & 0.0620 & -0.0460 & -0.1056 & 0.4144 & 0.1604 & -0.1604 \\ 0.0597 & 0.0327 & -0.0941 & 0.0976 & 0.0024 & 0.0629 & 0.0535 & -0.0535 \\ 0.0620 & -0.0941 & 0.4460 & -0.4491 & -0.0895 & 0.0435 & -0.1069 & 0.1069 \\ -0.0460 & 0.0976 & -0.4491 & 0.4816 & -0.0760 & -0.0658 & 0.1146 & -0.1146 \\ -0.1056 & 0.0024 & -0.0895 & -0.0760 & 0.9594 & 0.1156 & -0.0181 & 0.0181 \\ 0.4144 & 0.0629 & 0.0435 & -0.0658 & 0.1156 & 0.4592 & 0.1629 & -0.1629 \\ 0.1604 & 0.0535 & -0.1069 & 0.1146 & -0.0181 & 0.1629 & 0.6000 & 0.4000 \\ -0.1604 & -0.0535 & 0.1069 & -0.1146 & 0.0181 & -0.1629 & 0.4000 & 0.6000 \end{bmatrix}$$

Application of the iteration defined in Section 4

To the example yields the vector

$$z \equiv x^{(651)} = \begin{bmatrix} 0.8968 \\ 0.1583 \\ 0.0000 \\ 0.0000 \\ 0.0000 \\ 0.9347 \\ 1.3807 \\ 0.6194 \end{bmatrix}$$

Now $(\mathcal{A}^+\ell, z) = 0.9992$ where

$$\mathcal{A}^+\ell = \begin{bmatrix} \uparrow \\ \bar{e}^{q+1} \\ \downarrow \\ \uparrow \\ -\bar{e}^{(1)} \\ \downarrow \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0.5731 \\ -0.0905 \\ 0.8145 \\ -0.8018 \\ -0.2673 \\ 0.5345 \\ 0.0000 \\ 0.0000 \end{bmatrix}$$

Thus $$\hat{x} = \frac{2||A^+b||}{(\mathcal{A}^+\ell, z)} \begin{bmatrix} 0.8968 \\ 0.1583 \\ 0.0000 \end{bmatrix} = \begin{bmatrix} 8.5011 \\ 1.5006 \\ 0.0000 \end{bmatrix}$$

and $(A^+b,c) + ||c^q||\ ||A^+b||(v^{(3)} - v^{(4)}) = 18.5004$

By way of comparison, the exact optimal solution for the example is

$$x = \begin{bmatrix} 8.5 \\ 1.5 \\ 0.0 \end{bmatrix} \quad \text{with maximum } (x,c) = 18.5.$$

Remarks:

(1) The iterative method employed in solving the example required a large number of repetitions. This method is an example of what might appropriately be called "Kaczmarz-like" iterations [6]. Recently the problem of accelerating the rather regular, albeit slow, convergence of such methods has been studied (see, for example [3], [11] and [12]. In the numerical example given above, Aitken's δ^2 process was applied to the individual elements of the vector sequence, subject to the condition that acceleration was employed only after successive differences of corresponding elements were monotonically decreasing. The result was that the 651 iterations required for "convergence" without acceleration were reduced to 11, using the same convergence criterion. In another small but less trivial problem (n=8), 1274 iterations were reduced to 155 employing the same approach. A rather intriguing possibility which invites further investigation is that the approximate nature of the "interlace" pattern in a pair of optimal solutions $\tilde{x},\tilde{y}$ becomes fairly well established after far fewer iterations than are required for convergence. Should this prove to be a general property, the method described might prove useful in determining near-optimal solutions which may then be used in initiating a standard simplex solution.

(2) For certain problems with special structure, such as the multi-dimensional transportation problem [4], a closed form for A^+ may be used in implementing the intersection projection method [2]. For problems with no special structure, Kaczmarz-like iterations may be used in computing A^+b, $(I-A^+A)c$ and the subcomponents of the vectors $Pw^{(k)}$ using the techniques described in [11], in which case sparsity plays a significant, yet not an essential, role.

6. Bibliography

[1] Cline, R. E., "Representations for the generalized inverse of matrices with applications in linear programming", Doc. Th., Purdue U. (1967).

[2] Cline, R. E., and Pyle, L. D., "On the solution of structured LP problems using generalized inverse", Inter. Sym. on Math. Prog., Princeton U., (1967).

[3] Dyer, J., "Acceleration of the convergence of the Kaczmarz method and iterated homogeneous transformations", Doc. Th., Univ. of Calif., Los Angeles, (1965).

[4] Haley, K. B., "The multi-index problem", Ops. Res. 11, 3, (1963).

[5] Halmos, P. R., "Finite dimensional vector spaces", 2nd ed. Princeton U. Press (1958).

[6] Kaczmarz, S., "Angenäherte Auflösung von Systemen linearer Gleichungen, Bull. Internat. Acad. Polon. Sci., A 16-17, (1937).

[7] Nakano, H., "Spectral theory in the Hilbert space", Japan Soc. for the Promotion of Science, (1953).

[8] Penrose, R., "A generalized inverse for matrices", Proc. Cambridge Philos. Soc., 51, (1955).

[9] Pyle, L. D., "The generalized inverse in linear programming", Doc. Th., Purdue U., (1960).

[10] — "The generalized inverse in linear programming—Basic Theory", Computer Sciences Department Technical Rpt. 1, Purdue U., (1965).

[11] — "A generalized inverse ϵ-algorithm for constructing intersection projection matrices, with applications", Num. Math. 10, (1967).

[12] Pyle, L. D., and Smith, D. K., "Iterative solution of large, sparse linear systems", Inter. Sym. on Math. Prog., Princeton U., (1967).