THE OPTIMIZATION OF A QUADRATIC FUNCTION SUBJECT TO LINEAR CONSTRAINTS

Harry Markowitz[1]

The author discusses a computational technique applicable to the determination of the set of "efficient points" for quadratic programming problems.

1. QUADRATIC PROBLEMS

Suppose that variables $X_1, \ldots, X_N$ are to be chosen subject to linear constraints:

$$\Sigma a_{ij} X_j = b_i, \quad i = 1, \ldots, m_1 \tag{1}$$

$$\Sigma a_{ij} X_j \geq b_i, \quad i = m_1 + 1, \ldots, m \tag{2}$$

$$X_j \geq 0 \quad , \quad j = 1, \ldots, N_1 \tag{3}$$

where $0 \leq m_1 \leq m$, $0 \leq N_1 \leq N$ and the matrix (a_{ij}) $i = 1, \ldots, m_1$ has rank m_1 (otherwise the system is inconsistent or has at least one redundant equation). The payoff is a linear function $R = \Sigma r_j X_j$ whose coefficients r_j are not known at the time the X_j are chosen. The r_j, rather, are random variables with expected values μ_j and covariances σ_{jk} (including variances $\sigma_{jj} = \sigma_j^2$). The expected value of R is

$$E = \Sigma \mu_j X_j . \tag{4}$$

The variance of R is

$$V = \Sigma\Sigma \, \sigma_{jk} X_j X_k . \tag{5}$$

Suppose further that some decision-maker likes expected payoff (E) and dislikes variance of payoff (V). Our problem is to compute for the decision-maker (a) the "efficient combinations" of E and V, i.e., those attainable (E, V) combinations which give minimum V for given E and maximum E for given V (Figure 1); and (b) the points in the X space associated with the efficient E, V combinations, i.e., the set of efficient X's.

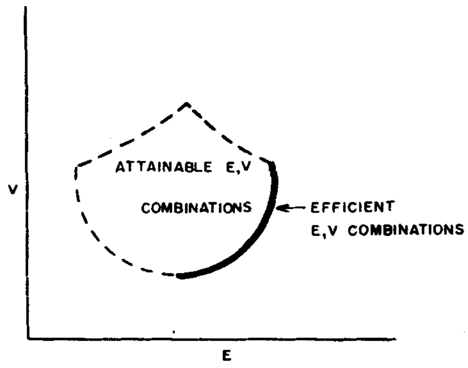

Figure 1

[1]The writer has particularly benefited from discussions with Kenneth Arrow on the subject matter of this paper.

A computing technique is presented in this paper for generating the above efficient sets. An adaptation of this technique can be used for problems of maximizing or minimizing quadratic forms (with the "right" properties) subject to linear constraints.

The practical problem which first suggested the above computing problem was that of selecting a portfolio of securities.[2] Here X_j is the amount invested in the j^{th} security; the μ_j and σ_{jk} are the expected returns and covariances of returns from the various securities. In the simplest case the constraint set is $\Sigma X_j = 1$, $X_j \geq 0$. A problem of very similar structure, analyzed independently by H. S. Houthakker,[3] is that of finding the expenditure on various goods as a function of income for an individual whose utility function is of the form $u = \Sigma a_j X_j + \Sigma\Sigma a_{ij} X_i X_j$. A problem of maximizing a monopolist's quadratic profit function subject to linear constraints is presented by Robert Dorfman.[4] Another problem of this general character is that of maximizing a quadratic likelihood function where there is a priori information concerning the values of parameters to be estimated. Now that reasonably convenient computing procedures exist for such quadratic problems we may be permitted the hope that still other classes of interesting questions can be reduced to this form.

This paper will discuss only minimization problems involving the quadratic form $\Sigma\Sigma \sigma_{ij} X_i X_j$ whose matrix (σ_{ij}) is positive semi-definite. The reader should have no difficulty in extending the results to minimization problems involving $\Sigma \rho_j X_j + \Sigma\Sigma \sigma_{ij} X_i X_j$ where (σ_{ij}) is positive semi-definite or maximization problems where (σ_{ij}) is negative semi-definite.

2. ASSUMPTIONS

According to customary usage we say:

(a) A set of points (S) (in Euclidean n-space) is convex if $X^{(1)} \epsilon$ S and $X^{(2)} \epsilon$ S imply $\lambda X^{(1)} + (1-\lambda) X^{(2)} \epsilon$ S, for any $0 \leq \lambda \leq 1$.

(b) A set is closed if $X_1, \ldots, X_t, \ldots \longrightarrow y$ and $X_1, \ldots, X_t, \ldots \epsilon$ S imply $y \epsilon$ S.

(c) A function f(X) is convex over a set S if $X^{(1)} \epsilon$ S, $X^{(2)} \epsilon$ S and $[\lambda X^{(1)} + (1-\lambda) X^{(2)}] \epsilon$ S imply $f(\lambda X^{(1)} + (1-\lambda) X^{(2)}) \leq \lambda f(X^{(1)}) + (1-\lambda) f(X^{(2)})$ for all $0 \leq \lambda \leq 1$.

(d) A function is strictly convex over a set S if $X^{(1)} \epsilon$ S, $X^{(2)} \epsilon$ S and $[\lambda X^{(1)} + (1-\lambda) X^{(2)}]$ ϵ S imply $f(\lambda X^{(1)} + (1-\lambda) X^{(2)}) < \lambda f(X^{(1)}) + (1-\lambda) f(X^{(2)})$ for all $0 < \lambda < 1$.

The set $(\widetilde{S})$ of points which satisfy constraints (1), (2), and (3) is a closed, convex set. Variance (V) is a positive semi-definite quadratic form, i.e., $\sum_j \sum_k \sigma_{jk} X_j X_k \geq 0$ for all $(X_1, \ldots, X_N)$. It is also convex. The covariance matrix (σ_{jk}) is non-singular if, and only if, V is positive definite $\left\{\text{i.e., } \sum_j \sum_k \sigma_{jk} X_j X_k > 0, \text{ if } (X_1, \ldots, X_N) \neq (0, \ldots, 0)\right\}$, which in turn is true if, and only if, V is strictly convex over the set of all X.[5]

---

[2] Harry Markowitz, "Portfolio Selection," Journal of Finance, 1952.

[3] "La Forme Des Courbes D'Engel," Cahiers du Séminaire d'Econometrie, 1953.

[4] Application of Linear Programming to the Theory of the Firm, University of California Press, 1951.

[5] That $V(X) \geq 0$, for all X, is due to the fact that V is the expected value of a square, $E(r-Er)^2$, and therefore cannot be negative. That $|\sigma_{ij}| \neq 0$ if, and only if, V is positive definite is a corollary of material found, e.g., in Birkhoff and MacLane, A Survey of Modern Algebra, Chapter IX, particularly section 8, pp. 243-247. The implications of positive definiteness and semi-definiteness for convexity may be seen as follows: Let $C = (\sigma_{ij})$. Let X and Y be column vectors; X' and Y' be row vectors. C is symmetric so that $C = C'$ and $X'CY = Y'CX$ for any X,Y. We wish to see the implications of

$$(1)\quad X'CX \geq 0$$

$$(2)\quad X'CX = 0 \Longleftrightarrow X = 0$$

(Continued)

We will assume[19] that $\widetilde{S}$ is not vacuous. We will also assume that V is strictly convex over the set of X's which satisfy the equations

$$\Sigma a_{ij} X_j = b_i , \quad i = 1, \ldots, m_1.$$

This assures[6] us that V takes on a unique minimum over $\widetilde{S}$ and that if $E = E^0$ is attainable in $\widetilde{S}$, then V takes on unique minimum over the set

$$\widetilde{S} \cap \{X \mid \Sigma \mu_j X_j \geq E_0\}.$$

If a function is convex over a set S, it is convex over any subset of S; therefore $|\sigma_{ij}| \neq 0$ implies that V is strictly convex over $\{X \mid \Sigma a_{ij} X_j = b_i, i = 1, \ldots, m_1\}$. This is not a necessary condition, however. Necessary and sufficient conditions on A and $(\sigma_{ij}) = C$ are discussed in the footnote.[7]

5(Continued)
for the difference

$$D = [\lambda X'CX + (1-\lambda) Y'CY] - [(\lambda X' + (1-\lambda)Y') C(\lambda X + (1-\lambda)Y)].$$

Expanding the second term and subtracting we get

$$D = \lambda(1-\lambda) \cdot [X'CX - 2X'CY + Y'CY]$$
$$= \lambda(1-\lambda) \cdot [(X'-Y') C (X-Y)].$$

Assumption (1) implies $D \geq 0$ for all X,Y. Assumptions (1) and (2) imply $D > 0$ if $X \neq Y$. Conversely, if D is positive for all $X \neq Y$, letting $Y = 0$ we find $X'CX > 0$ for all $X \neq 0$.

6Since equations (1) have rank m_1, m_1 variables and the m_1 equations could be eliminated (as in footnote 7) to leave a system with $N-m_1$ variables and $(m-m_1) + N_1$ inequalities. V is strictly convex in these $N-m_1$ variables and therefore the associated quadratic is positive definite. Let Y be any point in the space of the $N-m_1$ variables satisfying the $(m-m_1) + N_1$ inequalities. The points which satisfy these inequalities and have $V \leq V(Y)$ form a compact, convex set. Since V is continuous, it attains its minimum at least once on this set. Since it is strictly convex, it attains its minimum only once. The same argument applies if the constraint

$$\Sigma \mu_j X_j \geq E^0$$

is added to the $(m-m_1) + N_1$ inequalities.

7Suppose $m_1 \geq 1$. Since $\begin{pmatrix} a_{11} \cdots a_{1N} \\ \vdots \quad \vdots \\ a_{m_1N} \cdots a_{m_1N} \end{pmatrix}$ has rank m_1 we can write, after perhaps relabeling variables,

$$\begin{pmatrix} a_{11} \cdots a_{1m_1} \\ \vdots \quad \vdots \\ a_{m_11} \cdots a_{m_1m_1} \end{pmatrix} \begin{pmatrix} X_1 \\ \vdots \\ X_{m_1} \end{pmatrix} + \begin{pmatrix} a_{1m_1+1} \cdots a_{1N} \\ \vdots \quad \vdots \\ a_{m_1m_1+1} \cdots a_{m_1N} \end{pmatrix} \begin{pmatrix} X_{m_1+1} \\ \vdots \\ X_N \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_{m_1} \end{pmatrix}$$

or $A^{(1)} X^{(1)} + A^{(2)} X^{(2)} = b$ where $A^{(1)}$ is non-singular. We thus have $X^{(1)} = (A^{(1)})^{-1}b - (A^{(1)})^{-1} A^{(2)} X^{(2)}$. We can express V in terms of $X^{(2)}$ by substitution, i.e.,

$$X = \begin{pmatrix} X^{(1)} \\ X^{(2)} \end{pmatrix} = \begin{pmatrix} A^{(1)-1}b \\ 0 \end{pmatrix} - \begin{pmatrix} A^{(1)-1} A^{(2)} \\ I \end{pmatrix} X^{(2)}$$

$$V = X'CX = (X^{(1)'} \ X^{(2)'}) \ C \begin{pmatrix} X^{(1)} \\ X^{(2)} \end{pmatrix}$$

(Continued)

3. THE CRITICAL LINE $\bar{\ell}$

Let us first note the answer to a simpler problem than that of finding all (E, V) efficient points. Suppose we wished to minimize V subject to the equations

$$\Sigma a_{ij} X_j = b_i , \quad i = 1, \ldots, m_1$$

without regard to the inequalities (2) and (3). A necessary condition for a minimum is that $X_1, \ldots, X_N$ be a solution to the Lagrangian equations,

$$\frac{\partial (\Sigma\Sigma\, \sigma_{ij} X_i X_j - 2\, \Sigma\lambda_i\, \Sigma_j a_{ij} X_j)}{\partial X_k} = 0 , \quad k = 1, \ldots, N \tag{6}$$

as well as $\Sigma a_{ij} X_j = b_i$, $i = 1, \ldots, m_1$, i.e.,

$$\sum_{j=1}^{N} \sigma_{kj} X_j + \Sigma(-\lambda_i)\, \alpha_{ik} = 0 , \quad k = 1 \ldots, N \tag{7}$$

$$\Sigma a_{ij} X_j = b_i , \quad i = 1, \ldots, m_1 . \tag{8}$$

Given the assumption that V is strictly convex over $\{X \mid \Sigma a_{ij} X_j = b_i , \ i = 1, \ldots, m_1\}$ it follows that

$$\begin{pmatrix} \sigma_{11} & \cdots & \sigma_{1N} & a_{11} & \cdots & a_{m_1 1} \\ \vdots & & \vdots & \vdots & & \vdots \\ \sigma_{N1} & \cdots & \sigma_{NN} & a_{1N} & \cdots & a_{m_1 N} \\ a_{11} & \cdots & a_{1N} & 0 & \cdots & 0 \\ \vdots & & & \vdots & & \vdots \\ a_{m_1 1} & \cdots & a_{m_1 N} & 0 & \cdots & 0 \end{pmatrix}$$

---

[7](Continued)

$$= [(b' A^{(1)'-1}, 0) - (X^{(2)'} A^{(2)'} A^{(1)'-1}, X^{2'})]\, C \left[\begin{pmatrix} A^{(1)-1} b \\ 0 \end{pmatrix} - \begin{pmatrix} A^{(1)-1} A^{(2)} X^{(2)} \\ X^{(2)} \end{pmatrix}\right]$$

$$= (b' A^{(1)'-1}, 0)\, C \begin{pmatrix} A^{(1)-1} b \\ 0 \end{pmatrix} - 2X^{(2)} (A^{(2)'} A^{(1)'-1}, I)\, C \begin{pmatrix} A^{(1)-1} b \\ 0 \end{pmatrix}$$

$$+ X^{(2)'} \left[(A^{(2)'} A^{(1)-1}, I)\, C \begin{pmatrix} A^{(1)-1} A^{(2)} \\ I \end{pmatrix} \right] X^{(2)} .$$

V is strictly convex for all $X^{(2)}$ if, and only if, the last (i.e., the quadratic) term is strictly convex. This is so if, and only if,

$$(A^{(2)'} A^{(1)-1'}, I)\, C \begin{pmatrix} A^{(1)-1} A^{(2)} \\ I \end{pmatrix}$$

is non-singular.

is non-singular.[8] Since a strictly convex V takes on a unique minimum on a convex set, the unique solution to (7) and (8) is this minimum.

Next consider the problem of minimizing V subject not only to (1) but also to the constraint that $E = E_0$, i.e.,

$$\sum_{j=1}^{N} \mu_j X_j = E_0 . \tag{9}$$

We must distinguish two cases:

<u>Case 1</u>: The row vector $(\mu_1, \ldots, \mu_N)$ can be expressed as a linear combination of the $(a_{i1}, \ldots, a_{iN})$, i.e., there exists $\alpha_1, \ldots, \alpha_{m_1}$ such that

$$(\alpha_1, \ldots, \alpha_{m_1}) \begin{pmatrix} a_{11} & \cdots & a_{1N} \\ \vdots & & \vdots \\ a_{m_1 1} & \cdots & a_{m_1 N} \end{pmatrix} = (\mu_1, \ldots, \mu_N) . \tag{10}$$

<u>Case 2</u>: There does not exist such a linear combination.

As is shown below,[9] in Case 1 only one value of E, say $E = E^*$, is attainable. Therefore if we require $E \neq E^*$ no solution can be found; if we require $E = E^*$, equations (7) and (8) give the minimum. In Case 2 the matrix

$$\begin{pmatrix} \sigma_{11} & \cdots & {}_{1n}a_{11} & \cdots & a_{m_1 N} & \mu_1 \\ \vdots & & \vdots & & & \vdots \\ \sigma_{N1} & \cdots & \sigma_{NN}a_{1N} & \cdots & a_{m_1 N} & \mu_N \\ a_{11} & \cdots & a_{m_1 1} & O \cdots & O & O \\ \vdots & & & \vdots & \vdots & \vdots \\ a_{m_1} & \cdots & a_{m_1 N} & O \cdots & O & O \\ \mu_1 & \cdots & \mu_N & O \cdots & O & O \end{pmatrix}$$

---

[8]If $m_1 = 0$ the statement reduces to one proved in footnote 5. Suppose $m_1 \geq 1$. If $\begin{pmatrix} C & A' \\ A & O \end{pmatrix}$ is singular there is a vector $\begin{pmatrix} Y \\ -\lambda \end{pmatrix} \neq \begin{pmatrix} O \\ O \end{pmatrix}$ such that $\begin{pmatrix} C & A' \\ A & O \end{pmatrix} \begin{pmatrix} Y \\ -\lambda \end{pmatrix} = \begin{pmatrix} O \\ O \end{pmatrix}$ i.e., $\begin{pmatrix} CY \\ AY \end{pmatrix} = \begin{pmatrix} A'\lambda \\ O \end{pmatrix}$. Since the rank of A is m_1 there is no $\lambda \neq O$ such that $A'\lambda = O$, therefore $Y \neq O$ in $\begin{pmatrix} Y \\ -\lambda \end{pmatrix}$ above. $V(Y) = Y'CY = Y'A'\lambda = (AY)'\lambda = O$. Let X be any point in $S = \{X \mid AX = b\}$ where
$X' = (X_1, \ldots, X_N)$
$b' = (b_1, \ldots, b_{m_1})$
$A(X + Y) = AX + O = b$; therefore $X + Y$ is in S
$V(X) = X'CX$
$V(X + Y) = X'CX + 2X'CY = X'CX + 2X'A'\lambda = X'CX + 2b'\lambda$
$V(1/2X + 1/2(X + Y)) = V(X + 1/2\,Y)$
$= X'CX + X'CY$
$= X'CX + b'\lambda$
$= 1/2\,V(X) + 1/2\,V(X + Y)$,
thus contradicting strict convexity.

[9]If $\alpha'A = \mu'$ and $AX = b$ then $E = \mu'X = \alpha'AX = \alpha'b$.

is non-singular,[10] and therefore the equations

(11) $$\Sigma \sigma_{jk} X_k + \Sigma(-\lambda_1) a_{ij} - \lambda_E \mu_j = 0 \quad , \quad j = 1, \ldots, N$$

(12) $$\Sigma a_{ij} X_j = b_i \quad , \quad i = 1, \ldots, m_1$$

(13) $$\Sigma \mu_j X_j = E^0$$

have a unique solution which gives minimum V for the specified E^O. If we let E^O go from $-\infty$ to $+\infty$, the solution to (11), (12), and (13) traces out a line in the (X, λ) space. This line may also be described as the solution to the following $N + m_1$ equations in $N + m_1 + 1$ unknowns:

(14) $$\sum_{j=1}^{N} \sigma_{jk} X_k - \sum_{i=1}^{m_1} \lambda_i a_{ij} - \lambda_E \mu_j = 0 \quad , \quad j = 1, \ldots, N$$

(15) $$\Sigma a_{ij} X_j = b_i \quad , \quad i = 1, \ldots, m_1$$

or

(16) $$\Sigma \sigma_{jk} X_k + \Sigma(-\lambda_j) a_{ij} = \lambda_E \mu_j \quad , \quad j = 1, \ldots, N$$

(17) $$\Sigma \alpha_{ij} X_j = b_i \quad , \quad i = 1, \ldots, m_1$$

$$\text{for } -\infty \le \lambda_E \le +\infty .$$

Since the matrix of equations (16) and (17) is non-singular, given our assumption of strict convexity, they have a solution for every value of λ_E whether we are in Case 1 or Case 2 above. In Case 1 the values of $X_1, \ldots, X_N$ do not change (only the values of the λ's change) as λ_E goes from $-\infty$ to $+\infty$.[11] In Case 2 the X's as well as the λ's change. In Case 2 we can define $\hat{V}(E)$ to be minimum V as a function of $E \cdot 2\lambda_E = d\hat{V}/dE \cdot \hat{V}(E)$ must be strictly convex; therefore, E increases with λ_E. In Section 11 we show that $\hat{V}(E)$ is a parabola.

4. CRITICAL LINES $\ell(\mathcal{S}, \mathcal{G})$

The set of points (X, λ) which satisfy (16) and (17) will be referred to as the <u>critical line</u> $\bar{\ell}$ associated with the subspace

$$\bar{S} = \{X \mid \Sigma a_{ij} X_j = b_i \text{ for } i = 1, \ldots, m_1\}.$$

Critical lines will also be associated with certain other subspaces.

---

[10] Same proof as in footnote 8, using the fact that

$$\begin{pmatrix} A \\ \mu' \end{pmatrix} \text{ has rank } m_1 + 1.$$

[11] For if $\mu = A'\tilde{\lambda}$ and $\begin{pmatrix} C & A' \\ A & O \end{pmatrix} \begin{pmatrix} X^0 \\ \lambda^0 \end{pmatrix} = R + \begin{pmatrix} \mu \\ O \end{pmatrix} \lambda_E^0$ then $\begin{pmatrix} C & A' \\ A & O \end{pmatrix} \begin{pmatrix} X^0 \\ \lambda^0 + \tilde{\lambda}\theta \end{pmatrix}$

$= R + \begin{pmatrix} \mu \\ O \end{pmatrix} \left(\lambda_E^0 + \theta\right).$

Let $X_{j_1}, \ldots, X_{j_J}$ be a subset of variables. Let

$$\Sigma a_{ij} X_j = b_i \quad , \qquad i = i_1, \ldots, i_I$$

be a subset of the constraints (1) and (2) with the inequalities replaced by equalities when $i > m_1$. Let $\mathcal{I}$ be the ordered set of indices $(i_1, \ldots, i_I)$; let $\mathcal{J}$ be the ordered set $(j_1, \ldots, j_J)$. We will be particularly interested in $\mathcal{I}$ and $\mathcal{J}$ of the form

$$(18) \qquad \mathcal{I} = \{1, 2, \ldots, m_1, i_{m_1+1}, \ldots, i_I\} \quad \text{where } I \geq m_1$$

$$(19) \qquad \mathcal{J} = \{j_1, \ldots, j_L, N_1 + 1, \ldots, N\} \quad \text{where } 0 \leq L \leq N_1 .$$

For any indices $\mathcal{I}$ and $\mathcal{J}$ satisfying (18) and (19) we define the submatrix

$$(20) \qquad A_{\mathcal{I}\mathcal{J}} = \begin{pmatrix} a_{i_1 j_1} & \cdots & a_{i_1 j_J} \\ \vdots & & \vdots \\ a_{i_I j_1} & \cdots & a_{i_I j_J} \end{pmatrix}$$

We similarly define subvectors $X_{\mathcal{J}}$, $\lambda_{\mathcal{I}}$ and

$$(21) \qquad B_{\mathcal{I}} = \begin{pmatrix} b_{i_1} \\ \vdots \\ b_{i_I} \end{pmatrix};$$

also submatrices

$$(22) \qquad C_{\mathcal{J}\mathcal{J}} = \begin{pmatrix} \sigma_{j_1 j_1} & \cdots & \sigma_{j_1 j_J} \\ \vdots & & \vdots \\ \sigma_{j_J j_1} & \cdots & \sigma_{j_J j_J} \end{pmatrix}$$

$$(23) \qquad M_{\mathcal{I}\mathcal{J}} = \begin{pmatrix} C_{\mathcal{J}\mathcal{J}} & A'_{\mathcal{I}\mathcal{J}} \\ A_{\mathcal{I}\mathcal{J}} & O \end{pmatrix}$$

If $I = m_1 = 0$, $\mathcal{I}$ is empty. In this case it will sometimes be convenient to think of $A_{\mathcal{I}\mathcal{J}}$ as having no rows and J columns. To every $(\mathcal{I}, \mathcal{J})$ satisfying (18) and (19) we associate a subspace

$$S(\mathcal{I}, \mathcal{J}) = \{X \mid X_j = 0 \text{ for } j \notin \mathcal{J},\ A_{\mathcal{I}\mathcal{J}} X_{\mathcal{J}} = B_{\mathcal{I}}\}.$$

If $A_{\mathcal{I}\mathcal{J}}$ has no rows this reduces to $S(\mathcal{J}) = \{X \mid X_j = 0 \text{ for } j \notin \mathcal{J}\}$. Since $\mathcal{I}$ and $\mathcal{J}$ satisfy (18) and (19), $S(\mathcal{I}, \mathcal{J}) \subset \bar{S}$. Since V is strictly convex over $\bar{S}$, it is strictly convex over $S(\mathcal{I}, \mathcal{J})$.

$A_{\mathcal{I}\mathcal{J}}$ has a rank of I or less. If $A_{\mathcal{I}\mathcal{J}}$ has rank I then the matrix

$$M_{\mathcal{I}\mathcal{J}} = \begin{pmatrix} C_{\mathcal{J}\mathcal{J}} & A_{\mathcal{I}\mathcal{J}}' \\ A_{\mathcal{I}\mathcal{J}} & O \end{pmatrix} \tag{24}$$

is non-singular. (This is a special case of the proposition proved in footnote 8.) If $A_{\mathcal{I}\mathcal{J}}$ has rank less than I, $M_{\mathcal{I}\mathcal{J}}$ is singular, for its last I rows are not independent. For every $(\mathcal{I}, \mathcal{J})$ satisfying (18) and (19) whose $A_{\mathcal{I}\mathcal{J}}$ has rank equal to the number of its rows, we define the <u>critical line</u> $\ell_{\mathcal{I}\mathcal{J}}$ to be the set of points $(X_1, \ldots, X_N, \lambda_1, \ldots, \lambda_m)$ which satisfy

$$\begin{aligned} X_j &= 0 \quad \text{for } j \notin \mathcal{J} \\ \lambda_i &= 0 \quad \text{for } i \notin \mathcal{I} \end{aligned} \tag{25}$$

and

$$\begin{pmatrix} X_{\mathcal{J}} \\ -\lambda_{\mathcal{I}} \end{pmatrix} = M_{\mathcal{I}\mathcal{J}}^{-1} \begin{pmatrix} O \\ B_{\mathcal{I}} \end{pmatrix} + M_{\mathcal{I}\mathcal{J}}^{-1} \begin{pmatrix} \mu_{\mathcal{J}} \\ O \end{pmatrix} \lambda_E .$$

Equations (25) may be written in the form

$$\left.\begin{aligned} X_j &= \alpha_{Xj} + \beta_{Xj} \lambda_E \qquad (26) \\ \lambda_i &= \alpha_{\lambda i} + \beta_{\lambda i} \lambda_E \qquad (27) \end{aligned}\right\} -\infty < \lambda_E < \infty .$$

Equations (26) by themselves are the projection of $\ell(\mathcal{I}, \mathcal{J})$ onto the X-space. As with $\bar{S}$ and $\bar{\ell}$ we have two cases:

(1) Only one value of E is obtainable in $S(\mathcal{I}, \mathcal{J})$ and the X-projection is a point.

(2) All values of E are obtainable and the X-projection is a line. This line is the set of X's in $S(\mathcal{I}, \mathcal{J})$ which give minimum V for some E. Let

$$\xi_i = \Sigma a_{ij} X_j - b_i \ , \quad i = 1, \ldots, m \tag{28}$$

$$\eta_j = 1/2 \, \frac{\partial V - 2 \sum_{i=1}^{m} \lambda_i \Sigma a_{ij} X_j - 2\lambda_E \Sigma \mu_j}{\partial X_j} \tag{29}$$

$$= \sum_{j=1}^{N} \sigma_{jk} X_k + \sum_i (-\lambda_i) a_{ij} - \mu_j \lambda_E .$$

Constraints (1) and (2) state that

$$\xi_i = 0 \quad \text{for } i = 1, \ldots, m_1$$

$$\xi_i \geq 0 \quad \text{for } i = m_1 + 1, \ldots, m.$$

Along any critical line we have

$$
\begin{aligned}
X_j &= 0 && \text{for } j \notin \mathcal{J}, \\
\eta_j &= 0 && \text{for } j \in \mathcal{J}, \\
\xi_i &= 0 && \text{for } i \in \mathcal{I}, \\
\lambda_i &= 0 && \text{for } i \notin \mathcal{I}.
\end{aligned}
\tag{30}
$$

Also, from (25), letting m^{st} be the $(s,t)^{th}$ element of $M_{\mathcal{I}\mathcal{J}}^{-1}$, we have

$$
\begin{aligned}
X_{j_s} &= \sum_{h=1}^{I} m^{s,h+J} b_{i_h} + \left(\sum_{h=1}^{J} m^{sh} \mu_{j_h} \right) \lambda_E \\
&= \alpha_{Xj_s} + \beta_{Xj_s} \lambda_E \qquad \text{for } s = 1, \ldots, J.
\end{aligned}
\tag{31}
$$

$$
\begin{aligned}
\lambda_{i_s} &= -\sum_{h=1}^{I} m^{s+J,h+J} b_{i_h} - \left(\sum_{h=1}^{J} m^{s+J,h} \mu_{j_h} \right) \lambda_E \\
&= \alpha_{\lambda i_s} + \beta_{\lambda i_s} \lambda_E \qquad \text{for } s = 1, \ldots, I.
\end{aligned}
\tag{32}
$$

From (28) and (29) we have

$$
\begin{aligned}
\xi_i &= \sum_{h=1}^{J} a_{ij_h} \alpha_{Xj_h} - b_i + \left(\sum_{h=1}^{J} a_{ij_h} \beta_{Xj_h} \right) \lambda_E \\
&= \alpha_{\xi_i} + \beta_{\xi_i} \lambda_E ;
\end{aligned}
\tag{33}
$$

$$
\begin{aligned}
\eta_j &= \left(\sum_{h=1}^{J} \sigma_{jj_h} \alpha_{Xj_h} - \sum a_{i_h j} \alpha_{\lambda i_h} \right) \\
&\quad + \left(\sum_{h=1}^{J} \sigma_{jj_h} \beta_{Xj_h} - \sum_{h=1}^{I} a_{i_h j} \beta_{\lambda i_h} - \mu_j \right) \lambda_E \\
&= \alpha_{\eta j} + \beta_{\eta j} \lambda_E
\end{aligned}
\tag{34}
$$

A corollary of the results of an important paper by Kuhn and Tucker[12] is that a sufficient condition for a point X to give minimum V for a set

$$\widetilde{S} \cap \{X \mid \Sigma \mu_j X_j \geq E_0\}$$

for some E_O is that

$$
(35) \qquad
\begin{aligned}
X_j &\geq 0 && \text{for } j \leq N_1 \text{ and in } \mathcal{J}, \\
\eta_j &\geq 0 && \text{for } j \notin \mathcal{J}, \\
\xi_i &\geq 0 && \text{for } i \notin \mathcal{I}, \\
\lambda_i &\geq 0 && \text{for } i > m_1 \text{ and in } \mathcal{I},
\end{aligned}
$$

and $\lambda_E \geq 0$. If $\lambda_E > 0$, the constraint $E \geq E_0$ is effective; if E_0 were increased, an equally low value of V could not be obtained. If $\lambda_E = 0$, the point gives minimum V in $\widetilde{S}$. In either case the point is efficient.

It will be convenient at times to employ the following relabeling of variables:

$$
(36) \qquad
\begin{aligned}
v_k &= X_k && \text{for } k = 1, \ldots, N_1, \\
v_k &= \eta_{k-N_1} && \text{for } k = N_1 + 1, \ldots, 2N_1, \\
v_k &= \xi_{m_1 + k - 2N_1} && \text{for } k = 2N_1 + 1, \ldots, 2N_1 + m - m_1, \\
v_k &= \lambda_{k + 2m_1 - m - 2N_1} && \text{for } k = 2N_1 + m - m_1 + 1, \ldots, 2N_1 + 2m - 2m_1 .
\end{aligned}
$$

Also

$$(37) \qquad K = 2N_1 + 2m - 2m_1$$

and

$$(38) \qquad \mathcal{K} = \{\text{the set of } k \text{ which identify the variables in equation 30)}\} .$$

Thus on any critical line we have

$$(39) \qquad v_k = 0 \quad \text{for} \quad k \in \mathcal{K}$$

and a point X is efficient if it is a projection of a point on a critical line with

$$(40) \qquad v_k \geq 0 \quad \text{for} \quad k \notin \mathcal{K},$$

or

$$(41) \qquad v_k \geq 0 \quad \text{for} \quad k = 1, \ldots, K.$$

[12] H. W. Kuhn and A. W. Tucker, "Nonlinear Programming," in Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability.

5. INTERSECTIONS OF CRITICAL LINES; NON-DEGENERACY CONDITIONS

In the computing procedure of the next section we move along a critical line until it intersects a plane $v_k = 0$, $k = 1, \ldots, K$. Then either one row and the corresponding column is added to M, or one row and the corresponding column is deleted from M. This raises two questions: (1) under what conditions will the matrix obtained by such additions or deletions be non-singular, and (2) how should the new inverse be obtained? The latter question will not be discussed except to note that the possession of the old inverse is of great value in obtaining the new one.[13]

Concerning the former question, suppose $M_{\mathcal{I}\mathcal{J}}$, with $\mathcal{I}\mathcal{J}$ satisfying (18) and (19), is non-singular and thus defines a critical line ℓ. Suppose ℓ intersects (but is not contained in) the plane $v_k = 0$, $1 \le k \le K$. We distinguish four cases, depending on whether v corresponds to an X, an η, a λ, or a ξ:

1. The deletion of a variable. Suppose ℓ intersects a plane $X_j = 0$, $j = 1, \ldots, N_1$. Suppose that j is deleted from the set $\mathcal{J}$ leaving $\mathcal{J}^*$. Is $M_{\mathcal{I}\mathcal{J}^*}$ non-singular? We may suppose without loss of generality that $j = j_1$ and that $A_{\mathcal{I}\mathcal{J}}$ may therefore be written

$$A_{\mathcal{I}\mathcal{J}} = (\alpha A_{\mathcal{I}\mathcal{J}^*}) \tag{42}$$

where α is the column to be deleted. The matrix $\begin{pmatrix} \alpha & A_{\mathcal{I}\mathcal{J}^*} \\ 1 & 0 \ldots 0 \end{pmatrix}$ has either rank, I or I + 1. If it has rank I, then

$$\begin{pmatrix} C_{\mathcal{J}\mathcal{J}} & A'_{\mathcal{I}\mathcal{J}} & \mu_{j} \\ & & \vdots \\ & & \mu_{j_J} \\ A_{\mathcal{I}\mathcal{J}} & O & 0 \\ & & \vdots \\ & & 0 \\ 1\ 0 \ldots 0 & 0 \ldots 0 & 0 \end{pmatrix} = \widetilde{M} \tag{43}$$

is singular. In this case the equations

$$\widetilde{M} \begin{pmatrix} X_{\mathcal{J}} \\ \lambda_{\mathcal{I}} \\ \lambda_E \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ b_i \\ \vdots \\ b_{i_I} \\ 0 \end{pmatrix} \tag{44}$$

have either no solution or an infinity of solutions. Thus if (44) has one solution, i.e., if ℓ intersects $X_{j_1} = 0$ (but is not contained in it), $\widetilde{M}$ is non-singular and the rank of $\begin{pmatrix} \alpha & A_{\mathcal{I}\mathcal{J}^*} \\ 1 & 0\ 0 \end{pmatrix}$ is I + 1;

13This involves procedures similar to those used in removing a variable from a regression analysis or modifying a basis in linear programming; e.g., see R. A. Fisher, Statistical Methods for Research Workers, p. 164, 10th ed., and R. Dorfman, Activity Analysis of Production and Allocation, p. 358.

nce the rank of $\begin{pmatrix} A_{\mathcal{I}\mathcal{J}^*} \\ 0 \ldots 0 \end{pmatrix}$ is at least I. But the rank of $\begin{pmatrix} A_{\mathcal{I}\mathcal{J}^*} \\ 0 \ldots 0 \end{pmatrix}$ = the rank of $A_{\mathcal{I}\mathcal{J}^*}$. Therefore the rank of $A_{\mathcal{I}\mathcal{J}^*}$ is I and $M_{\mathcal{I}\mathcal{J}^*}$ is non-singular.

2. The deletion of a constraint. Suppose ℓ intersects (but is not contained in) $\lambda_i = 0$ for $i > m_1$. We may assume $i = i_I$ and that

$$A_{\mathcal{I}\mathcal{J}} = \begin{pmatrix} A_{\mathcal{I}^*\mathcal{J}} \\ \alpha' \end{pmatrix}. \tag{45}$$

$A_{\mathcal{I}\mathcal{J}}$ has rank I, $A_{\mathcal{I}^*\mathcal{J}}$ has rank I-1; therefore $M_{\mathcal{I}^*\mathcal{J}}$ is non-singular.

3. Addition of a variable. Continuing the conventions used above, if $A_{\mathcal{I}\mathcal{J}}$ has rank I, so has $A_{\mathcal{I}\mathcal{J}^*} = (A_{\mathcal{I}\mathcal{J}}\ \alpha)$. Therefore $M_{\mathcal{I}\mathcal{J}^*}$ is non-singular.

4. Addition of a constraint. If ℓ intersects but is not contained in the plane $\xi_i = 0$, $i > m_1$ then

$$\begin{pmatrix} C_{\mathcal{J}\mathcal{J}} & A_{\mathcal{I}\mathcal{J}} & \begin{matrix} \mu_{j_1} \\ \vdots \\ \mu_{j_J} \end{matrix} \\ A_{\mathcal{I}\mathcal{J}} & O & O \\ \alpha' & O & O \end{pmatrix} = \widetilde{M} \tag{46}$$

(where α' is the row of coefficients of the new constraint) is non-singular. Therefore $\begin{pmatrix} A_{\mathcal{I}\mathcal{J}} \\ \alpha \end{pmatrix} = A_{\mathcal{I}^*\mathcal{J}}$ has rank I + 1 and $M_{\mathcal{I}^*\mathcal{J}}$ is non-singular.

The tracing out of the efficient set is simplified if certain "accidents" do not occur. These accidents are described in the following "non-degeneracy" conditions. The next section of this paper presents a computing procedure for deriving the set of efficient points when all non-degeneracy conditions hold. In Sections 7-10 these conditions are relaxed.

CONDITION 1. On no critical line do we have

$$v_k = \alpha_{vk} + \beta_{vk}\lambda_E \equiv 0 \quad \text{for } k \notin \mathcal{K}.$$

CONDITION 2. On any given critical line ℓ we do not have

$$\frac{-\alpha_{vk_1}}{\beta_{vk_1}} = \frac{-\alpha_{vk_2}}{\beta_{vk_2}} \quad \text{for any } k_1 \neq k_2 \text{ with } \beta_{vk_1} \neq 0,\ \beta_{vk_2} \neq 0\,.$$

CONDITION 3. E is bounded from above in $\widetilde{S}$.

We will let L_E stand for "the linear programming problem of maximizing E subject to constraints (1), (2), and (3)."

CONDITION 4. L_E has a unique non-degenerate solution.

Condition 4 implies condition 3.

6. THE ALGORITHM UNDER CONDITIONS 1 THROUGH 4

We now assume that conditions 1 through 4 are satisfied. Condition 4 implies that the optimum solution to L_E has:[14]

(a) Exactly m variables X_j and ξ_i are not at their lower extreme;

(b) $A_{\mathcal{I}(1)\mathcal{J}(1)}$ (where $\mathcal{I}^{(1)}$ includes all i with $\xi_i = 0$ and $\mathcal{J}^{(1)}$ includes all j with X_j not at its lower limit) has rank equal to the number of its rows;

(c) There exists "prices" p_i and "profitabilities" δ_j such that

$$p_i > 0 \quad \text{if } \xi_i = 0 \quad \text{for } i = m_1+1, \ldots, m, \tag{47}$$

$$p_i = 0 \quad \text{if } \xi_i > 0 \quad \text{for } i = m_1+1, \ldots, m, \tag{48}$$

$$\delta_j = \sum_i a_{ij}\, p_i + \mu_j, \tag{49}$$

$$\delta_j = 0 \quad \text{for } j = N_1+1, \ldots, N \text{ and for } X_j > 0 \quad j \le N_1, \tag{50}$$

$$\delta_j < 0 \quad \text{for } X_j = 0 \quad j \le N_1 \tag{51}$$

The matrix

$$M_{(1)} = \begin{pmatrix} C_{\mathcal{J}(1)\mathcal{J}(1)} & A'_{\mathcal{I}(1)\mathcal{J}(1)} \\ A_{\mathcal{I}(1)\mathcal{J}(1)} & O \end{pmatrix} \tag{52}$$

is non-singular and thus defines a critical line $\ell^{(1)}$ along which

$$M_{(1)} \begin{pmatrix} X_{\mathcal{J}(1)} \\ -\lambda_{\mathcal{I}(1)} \end{pmatrix} = \begin{pmatrix} O \\ B_{\mathcal{I}(1)} \end{pmatrix} + \begin{pmatrix} \mu_{\mathcal{J}(1)} \\ O \end{pmatrix} \lambda_E \,. \tag{53}$$

Since $-A'_{\mathcal{I}(1)\mathcal{J}(1)} \cdot p_{\mathcal{I}(1)} = \mu_{\mathcal{J}(1)}$, if $X^0_{\mathcal{J}(1)}\ \lambda^0_{\mathcal{I}(1)}$ satisfy (53) for $\lambda_E = 0$, then

$$M_1 \begin{pmatrix} X^0_{\mathcal{J}(1)} \\ -\lambda^0_{\mathcal{I}(1)} - p_{\mathcal{I}(1)}\lambda_E \end{pmatrix} = \begin{pmatrix} O \\ B_{\mathcal{I}(1)} \end{pmatrix} + \begin{pmatrix} \mu_{\mathcal{J}(1)} \\ O \end{pmatrix} \lambda_E \tag{54}$$

for all λ_E. Thus $\ell^{(1)}$ has

[14] The following are corollaries of the basis and pricing theorems of linear programming. See, e.g., George B. Dantzig, Alex Orden, Philip Wolfe, "The Generalized Simplex Method for Minimizing a Linear Form under Linear Inequality Restraints," Pacific Journal of Mathematics, Vol. 5, No. 2, June 1955.

$$(55)\qquad \begin{aligned} X_{\mathcal{J}(1)} &= X^0_{\mathcal{J}(1)} \\ \lambda_{\mathcal{I}(1)} &= \lambda^0_{\mathcal{I}(1)} + p_{\mathcal{I}(1)}\lambda_E \end{aligned}$$

for all λ_E. From (47) it follows that for sufficiently large λ_E

$$\lambda_i > 0 \quad \text{for } i > m_1 \text{ and in } \mathcal{I}\,.$$

$$(56)\qquad \begin{aligned} \eta_j &= \Sigma\sigma_{jh} X^0_h - (\Sigma a_{ij}\lambda_i + \mu_j\lambda_E) \\ &= \Sigma\sigma_{jh} X^0_h - \Sigma a_{ij}\lambda^0_i - (\Sigma a_{ij} p_i + \mu_j)\lambda_E\,. \end{aligned}$$

Equation (51) implies that for sufficiently large $\lambda_E, \eta_j > 0$ for $j \notin \mathcal{J}$. Thus for sufficiently large $\lambda_E, \ell^{(1)}$ satisfies inequalities (40).

Let $\lambda_E^{(1)}$ be the largest value of λ_E at which $\ell^{(1)}$ intersects a plane $\eta_j = 0$ for $j \notin \mathcal{J}$ or $\lambda_i = 0$ for $i \in \mathcal{I}$. (The X and ξ do not vary along $\ell^{(1)}$.) If $\lambda_E^{(1)} \le 0$ then X^0 gives minimum V as well as maximum E. Suppose $\lambda_E^{(1)} > 0$. Non-degeneracy condition 2 implies $\ell^{(1)}$ intersects only one plane $\eta_j = 0$ or $\lambda_i = 0$ at $\lambda_E^{(1)}$. In the former case we add j to $\mathcal{J}$; in the latter case we delete i from $\mathcal{I}$, to form $\mathcal{I}^{(2)}$, $\mathcal{J}^{(2)}$. The new matrix $M_{(2)} = M_{\mathcal{I}(2)\mathcal{J}(2)}$ is non-singular and defines a critical line $\ell^{(2)}$. Suppose for a moment that it was $\eta_{j_0} = 0$ which $\ell^{(1)}$ intersected at $\lambda_E^{(1)}$. On $\ell^{(2)}$ we have at $\lambda_E = \lambda_E^{(1)}$:

$$(57)\qquad \begin{aligned} &\lambda_i > 0 \quad \text{for } i > m_1 \text{ and } \in \mathcal{I}^{(2)}, \\ &\eta_i > 0 \quad \text{for } j \notin \mathcal{J}^{(2)}, \\ &\xi_i > 0 \quad \text{for } i \notin \mathcal{I}, \\ &\text{and } \begin{cases} X_{j_0} = 0 \\ X_j > 0 \quad \text{for all other } j \le N_1 \text{ and } \in \mathcal{J}^{(2)}\,. \end{cases} \end{aligned}$$

As always $X_{j_0} = a + b\lambda_E$ along $\ell^{(2)}$. Non-degeneracy condition 1 assures $b \ne 0$. If $b < 0$ the projection of $\ell^{(2)}$ would be efficient for $\lambda^* \ge \lambda_E \ge \lambda_E^{(1)}$ where $\lambda^* > \lambda_E^{(1)}$. This is impossible.[15] Therefore $b > 0$ and $\ell^{(2)}$ is efficient for $\lambda_E^{(1)} \ge \lambda_E \ge \lambda_E^{(2)}$ where $\lambda_E^{(1)} > \lambda_E^{(2)}$. Similar remarks would apply if $\ell^{(1)}$ first intersected $\lambda_{i_0} = 0$ and i_0 was deleted from $\mathcal{I}$.

---

[15]Since $b \ne 0$ the X-projection of the critical line is a line rather than a point. Along this line E increases with λ_E. If $\lambda_E > \lambda_E^{(1)}$ were feasible, then $E > E^{(1)} = \max E$ would be obtainable, which is impossible.

$\lambda_E^{(2)}$ is the highest value of $\lambda_E < \lambda_E^{(1)}$ at which $\ell^{(2)}$ intersects a plane $v_k = 0$ for $k = 1, \ldots, K$. If this is an η_j we again add a j to $\mathcal{J}$. If it is a λ_i we delete i from $\mathcal{I}$; if a ξ_i, we add i to $\mathcal{I}$; if an X_j, we delete j from $\mathcal{J}$. We form $M_{(3)}$ and $\ell_{(3)}$ accordingly and find $\lambda^{(3)} < \lambda^{(2)}$. This process is repeated until $\lambda_E = 0$ is reached. At each step (s) $M_{(s)}$ is non-singular and if v_{k_s} is the new variable (η, X, λ, or ξ) which is no longer constrained to be zero we have at $\lambda^{(s-1)}$.

$$v_k > 0 \quad \text{for } k \neq k_s \text{ and } \notin \mathcal{K}, \tag{58}$$

$$v_{k_s} = 0 .$$

By condition 1, $b_{vk_s} \neq 0$ along $\ell^{(s)}$. We argue below[16] that we cannot have $b_{vk_s} < 0$. So $b_{vk_s} > 0$ and $\ell^{(s)}$ is efficient for $\lambda_E^{(s-1)} \geq \lambda_E \geq \lambda_E^{(s)}$ where $\lambda_E^{(s-1)} > \lambda_E^{(s)}$. Since there are only a finite number of critical lines, and each can satisfy inequalities (40) for only one segment, $\lambda_E = 0$ is reached in a finite number of steps.

7. THE ALGORITHM UNDER CONDITIONS 3 AND 4

Let us now drop non-degeneracy conditions 1 and 2 but still assume conditions 3 and 4. We will use techniques analogous to the degeneracy-avoiding techniques of linear programming.[17]

For every number ε we define a new problem $P(\varepsilon)$ as follows:

$$\text{minimize } V(\varepsilon) = \Sigma\Sigma\sigma_{ij} X_i X_j + \Sigma \varepsilon^j X_j$$

subject to

$$\Sigma a_{ij} X_j = b_i + \varepsilon^{N+i} \quad , \quad i = 1, \ldots, m_1 , \tag{59}$$

$$\Sigma a_{ij} X_j \geqq b_i + \varepsilon^{N+i} \quad , \quad i = m_1 + 1, \ldots, m , \tag{60}$$

$$X_j \geq 0 \quad , \quad j = 1, \ldots, N_1 . \tag{61}$$

---

[16] If v_{k_s} is an X_j or ξ_i, $b_{vk_s} < 0$ implies that there are two distinct points which minimize V for some $E > E^{(s-1)}$, which is impossible. This argument also applies if v_k is a λ_i or η_j unless the X-projection of the new critical line is a point. In the latter case we note (from the Kuhn and Tucker conditions) that an efficient point gives minimum $Q(\lambda_E) = V - \lambda_E E$ subject to (1), (2), (3). For fixed λ_E, $Q(\lambda_E)$ has a unique minimum. If $v_k < 0$ then two distinct points give minimum $Q(\lambda_E)$ for some $\lambda_E > \lambda_E^{(s-1)}$.

[17] In linear programming these techniques are generally not needed in practice. In quadratic programming arbitrary selection of $v_k = 0$ with $b_{vk} < 0$ to go into $\mathcal{K}$ may (or may not) prove adequate. In any case, the degeneracy-handling techniques are available if needed. See George Dantzig, "Application of the Simplex Method to a Transportation Problem," Activity Analysis of Production and Allocation, Tjalling C. Koopmans, ed.; A. Charnes, "Optimality and Degeneracy in Linear Programming," Econometrica, Vol. 20, No. 2, April, 1952, p. 160; and Dantzig, Orden, and Wolfe, Op. Cit.

391368 O - 56 - 9

For sufficiently small ε the unique, optimal basis of L_E is feasible and, since it still satisfies the pricing conditions, is optimal.

As we will see shortly for sufficiently small ε, $P(\varepsilon)$ satisfies non-degeneracy conditions (1) and (2). We will also see that for a sufficiently small ε^*, the sequence of indices $(\mathfrak{I}\,\mathfrak{J})^s$ associated with the critical lines $\ell^{(s)}$, until $\lambda_E = 0$ is reached, is the same for all $P(\varepsilon)$ for $\varepsilon^* \geq \varepsilon > 0$. If we change indices $(\mathfrak{I}, \mathfrak{J})$ in the same sequence as $P(\varepsilon)$ for small ε, if we let λ_E decrease along any critical line when it can without violating $v_k \geq 0$, until we reach $\lambda_E = 0$, then:

(a) We will pass through a finite number of index sets each associated with a non-singular $M_{\mathfrak{I}\mathfrak{J}}$, before we reach $\lambda_E = 0$.

(b) Since $v_k \geq 0$ is maintained we have the desired solution to the original problem.

Along any critical line of $P(\varepsilon)$ we have

$$\begin{pmatrix} X_{\mathfrak{J}} \\ -\lambda_{\mathfrak{I}} \end{pmatrix} = M_{\mathfrak{I}\mathfrak{J}}^{-1} \begin{pmatrix} O \\ B_{\mathfrak{I}} \end{pmatrix} + M_{\mathfrak{I}\mathfrak{J}}^{-1} \begin{pmatrix} \mu_{\mathfrak{J}} \\ O \end{pmatrix} \lambda_E + M_{\mathfrak{I}\mathfrak{J}}^{-1} \begin{pmatrix} \varepsilon^{j_1} \\ \vdots \\ \varepsilon^{N+i_I} \end{pmatrix} \tag{62}$$

or

$$X_{j_s} = \alpha_{Xj_s} + \beta_{Xj_s} \lambda_E + \sum_{h=1}^{I+J} m^{sh} \varepsilon^{f(h)} \tag{63}$$

where $f(1) = j_1,\ f(2) = j_2, \ldots, f(I+J) = N + i_I$

or

$$X_{j_s} = \alpha_{Xj_s} + \beta_{Xj_s} \lambda_E + p_{Xj_s}(\varepsilon) . \tag{64}$$

Similarly

$$\lambda_{i_s} = \alpha_{\lambda i_s} + \beta_{\lambda i_s} \lambda_E - \sum_{h=1}^{I+J} m^{J+s,h} \varepsilon^{f(h)}$$

$$= \alpha_{\lambda i_s} + \beta_{\lambda i_s} \lambda_E + p_{\lambda i_s}(\varepsilon) .$$

$$\xi_i = \alpha_{\xi_i} + \beta_{\xi_i} \lambda_E + \sum_{s=1}^{J} a_{ij_s} p_{Xj_s}(\varepsilon) + \varepsilon^{N+i} = \alpha_{\xi_i} + \beta_{\xi_i} \lambda_E + p_{\xi_i}(\varepsilon) \ . \tag{65}$$

$$\eta_j = \alpha_{\eta_j} + \beta_{\eta_j} \lambda_E + \sum_{s=1}^{J} \sigma_{jj_s} p_{Xj_s}(\varepsilon) - \sum_{s=1}^{I} a_{i_s j} p_{\lambda i_s}(\varepsilon) + \varepsilon^j \tag{66}$$

$$= \alpha_{\eta_j} + \beta_{\eta_j} \lambda_E + p_{\eta_j}(\varepsilon) .$$

Consider the polynomials:

(67) $$p_{Xj}(\varepsilon) \quad \text{for} \quad j \in \mathcal{J},$$

(68) $$p_{\lambda i}(\varepsilon) \quad \text{for} \quad i \in \mathcal{I},$$

(69) $$p_{\xi i}(\varepsilon) \quad \text{for} \quad i \notin \mathcal{I},$$

(70) $$p_{\eta j}(\varepsilon) \quad \text{for} \quad j \notin \mathcal{J}.$$

None of the polynomials listed above have all zero coefficients, and no two have proportional coefficients. For each polynomial of (69) and (70) has a term with a coefficient of 1 which every other polynomial has with a coefficient of zero. This leaves only the possibilities that some polynomial of (67) or (68) has all zero coefficients or two of these polynomials have proportional coefficients. Both these possibilities imply that M^{-1} is singular and therefore are impossible.

Since $p_{v_k}(\varepsilon)$ has only a finite number of roots, for ε sufficiently small

$$p_{v_k}(\varepsilon) \neq 0 \qquad \text{for } k \notin \mathcal{K}.$$

Thus

(71) $$v_k = \alpha_{vk} + \beta_{vk}\lambda_E + p_{vk}(\varepsilon) \qquad -\infty < \lambda_E < \infty$$

cannot be identically zero for $k \notin \mathcal{K}$. The critical line intersects the plane $v_{k_1} = 0$ at

(72) $$\lambda_E' = \frac{-\alpha_{vk_1}}{\beta_{vk_1}} - \frac{p_{vk_1}(\varepsilon)}{\beta_{vk_1}}$$

and the plane

$$v_{k_2} = 0 \text{ at}$$

(73) $$\lambda_E'' = \frac{-\alpha_{vk_2}}{\beta_{vk_2}} - \frac{p_{vk_2}(\varepsilon)}{\beta_{vk_2}}.$$

If, say,

(74) $$\frac{-\alpha_{vk_1}}{\beta_{vk_1}} > \frac{-\alpha_{vk_2}}{\beta_{vk_1}}$$

then for sufficiently small ε

$$\lambda'_E > \lambda''_E .$$

On the other hand, since

$$p_{vk_1}(\varepsilon) - p_{vk_2}(\varepsilon) = 0 \qquad k_1, k_2 \not\in \mathcal{K}$$

has a finite number of solutions, for sufficiently small ε

$$\lambda'_E \neq \lambda''_E$$

even if

$$\frac{-\alpha_{vk_1}}{\beta_{vk_1}} = \frac{-\alpha_{vk_2}}{\beta_{vk_2}}. \tag{75}$$

As $\varepsilon \longrightarrow 0$ the smallest power of ε dominates; i.e., if, say, $\dfrac{-p_{vk_1}(\varepsilon)}{\beta_{vk_1}}$ has an algebraically larger coefficient for the first power of ε, then $\lambda' > \lambda''$ as $\varepsilon \longrightarrow 0$. If both have the same coefficient of ε, then it is the coefficients of ε^2 that decide. And so on.

Since there are a finite number of critical lines and a finite number of planes $v_k = 0$, there is a single ε^* such that for $\varepsilon^* \geq \varepsilon > 0$.

$P(\varepsilon)$ satisfies non-degeneracy conditions (1) and (2); and the order of the index sets $(\mathcal{I}\,\mathcal{J})^s$ is the same for all such ε.

The m^{st} are needed for other purposes and are thus available for resolving degeneracy problems. The other coefficients of $p_{v_k}(\varepsilon)$ can be computed when needed.

8. THE ALGORITHM WHEN L_E IS DEGENERATE BUT UNIQUE

Suppose that the solution to L_E is degenerate in that one or more of the basis variables X_j or ξ_i is "accidentally" zero, but is unique in that $\delta_j < 0$ for all X_j not in the basis and $p_i > 0$ for all ξ_i not in the basis.

The constraints of L_E may be written as a system of equations including the ξ_i as variables:

$$B \begin{pmatrix} X \\ \xi \end{pmatrix} = b . \tag{76}$$

If $\widetilde{B}$ is the submatrix of optimal basis vectors and if $X_{\mathcal{J}}$ and $\xi_{\bar{\mathcal{I}}}$ are the optimal basis variables, then the optimal solution is given by

$$\begin{pmatrix} X^0_{\mathcal{J}} \\ \xi^0_{\bar{\mathcal{I}}} \end{pmatrix} = \widetilde{B}^{-1} b \tag{77}$$

while all other variables are zero. After we solve L_E we may modify it, forming $L_E(\varepsilon)$ as follows:

$$B \begin{pmatrix} X \\ \xi \end{pmatrix} = b + \widetilde{B} \begin{pmatrix} \varepsilon \\ \varepsilon \\ \vdots \\ \varepsilon \end{pmatrix} \quad \text{for } \varepsilon > 0 \tag{78}$$

$$= b + \varepsilon \begin{pmatrix} r_1 \\ \vdots \\ r_m \end{pmatrix}$$

where r_i is the sum of the i^{th} row of $\widetilde{B}$.

Then

$$\begin{pmatrix} X_{\mathfrak{J}}(\varepsilon) \\ \\ \xi_{\bar{\mathfrak{J}}}(\varepsilon) \end{pmatrix} = \widetilde{B}^{-1} b + \begin{pmatrix} \varepsilon \\ \vdots \\ \varepsilon \end{pmatrix}. \tag{79}$$

Thus the original optimal basis is still feasible and therefore uniquely optimal (since it still satisfies the pricing relationships). Also for $\varepsilon > 0$

$$X_j(\varepsilon) > 0 \qquad \text{for } j \in \mathfrak{J},$$

$$\xi_i(\varepsilon) > 0 \qquad \text{for } i \in \bar{\mathfrak{J}},$$

and

$$\begin{pmatrix} X_{\mathfrak{J}}(\varepsilon) \\ \xi_{\bar{\mathfrak{J}}}(\varepsilon) \end{pmatrix} \longrightarrow \begin{pmatrix} X^0_{\mathfrak{J}} \\ \xi^0_{\bar{\mathfrak{J}}} \end{pmatrix}$$

as $\varepsilon \longrightarrow 0$.

The procedures of the last section which apply when L_E has a unique and non-degenerate solution apply with essentially no modification if L_E has a unique but possibly degenerate solution, if we let $P(\varepsilon)$ be

$$\min V = \Sigma\Sigma\, \sigma_{ij} X_i X_j + \Sigma\, \varepsilon^{j+1} X_j$$

subject to

$$\begin{aligned} &\Sigma\, a_{ij} X_j = b_i + r_i\, \varepsilon + \varepsilon^{N+i+1} \quad \text{for } i = 1, \dots, m_1, \\ &\Sigma\, a_{ij} X_j \geqq b_i + r_i\, \varepsilon + \varepsilon^{N+i+1} \quad \text{for } i = m_1+1, \dots, m. \end{aligned} \tag{80}$$

The solution to $L_E(\varepsilon)$ is non-degenerate for sufficiently small ε. Along any critical line we now have

$$
\begin{aligned}
X_{j_s} &= \alpha_{Xj_s} + \beta_{Xj_s}\lambda_E + \varepsilon\, p_{Xj_s}(\varepsilon) + \left(\sum_{h=1}^{I} m^{s,h+J} r_{i_h}\right)\varepsilon \\
&= \alpha_{Xj_s} + \beta_{Xj_s}\lambda_E + q_{Xj_s}(\varepsilon)\,;
\end{aligned} \tag{81}
$$

$$
\begin{aligned}
\lambda_{i_s} &= \alpha_{\lambda i_s} + \beta_{\lambda i_s}\lambda_E + \varepsilon\, p_{\lambda i_s}(\varepsilon) - \left(\sum_{h=1}^{I} m^{s+J,h+J} r_{i_h}\right)\varepsilon \\
&= \alpha_{\lambda i_s} + \beta_{\lambda i_s}\lambda_E + q_{\lambda i_s}(\varepsilon)\,;
\end{aligned} \tag{82}
$$

$$
\begin{aligned}
\xi_i &= \alpha_{\xi i} + \beta_{\xi i}\lambda_E + \sum_{s=1}^{J} a_{ij_s} q_{Xj_s}(\varepsilon) + \varepsilon^{N+i+1} \\
&= \alpha_{\xi i} + \beta_{\xi i}\lambda_E + q_{\xi i}(\varepsilon)\,;
\end{aligned} \tag{83}
$$

$$
\begin{aligned}
\eta_j &= \alpha_{\eta j} + \beta_{\eta j}\lambda_E + \sum_{s=1}^{J} \sigma_{jj_s} q_{Xj_s}(\varepsilon) - \sum_{s=1}^{I} a_{i_s j} q_{\lambda i_s}(\varepsilon) + \varepsilon^{j+1} \\
&= \alpha_{\eta j} + \beta_{\eta j}\lambda_E + q_{\eta j}(\varepsilon)
\end{aligned} \tag{84}
$$

where the $p_v(\varepsilon)$ are as defined in (63) through (66). Since no $p_v(\varepsilon)$ can have zero coefficients and no two can have proportional coefficients, the same is true of the $q_v(\varepsilon)$.

9. THE ALGORITHM WHEN L_E IS NOT UNIQUE

A non-degenerate optimal solution to L_E is unique if, and only if,

$$
\begin{aligned}
&\delta_j < 0 \text{ for } X_j \text{ not in the basis} \\
&p_i > 0 \text{ for } \xi_i \text{ not in the basis.}
\end{aligned} \tag{85}
$$

If L_E has a degenerate solution and (85) does not hold, then either the solution is not unique or else only the optimal basis is not unique. If L_E does not have a unique solution we must find the point $\overline{X}$ which gives minimum V for $E = \overline{E} = \max E$. If only the optimal basis of L_E is not unique we still must decide on the $\mathcal{I}\mathcal{J}$ of our first critical line. Both these problems will be resolved in the same manner. Our procedure may be considered as a special case of either approach 1 or approach 4 for minimizing a quadratic subject to linear constraints described in Section 12.

Let us create a new linear programming problem $L_F(\varepsilon)$ by adding a constraint to and modifying the form to be maximized in $L_E(\varepsilon)$. The equation we add is

$$
\sum_{j=1}^{N} \mu_j X_j - \xi_E = \overline{E} + \left(\sum_{j\in\mathcal{J}} \mu_j - 1\right)\varepsilon\,. \tag{86}
$$

If we add ξ_E to the optimum basis variables of $L_E(\varepsilon)$ we have a feasible basis corresponding to a solution with

$$
\begin{aligned}
X_j &= X_j^0 + \varepsilon \qquad && j \in \mathcal{J}, \\
\xi_i &= \xi_i^0 + \varepsilon \qquad && i \in \bar{\mathcal{I}}, \\
\xi_E &= \varepsilon .
\end{aligned}
\tag{87}
$$

Next let us replace the objective function $E = \Sigma\ \mu X$ with a new one

$$F = \Sigma\, \nu_j X_j \tag{88}$$

such that the solution in (87) is the unique optimum of $L_F(\varepsilon)$. This may be done easily by assigning any values $p_i > 0$ to $i \notin \bar{\mathcal{I}}$, $p_i = 0$ for $i \in \bar{\mathcal{I}}$ as well as $p_E = 0$. Then choose any set for ν_j so that $\delta_j = 0$ for $j \in \mathcal{J}$ and $\delta_j < 0$ for $j \notin \mathcal{J}$. Since $L_F(\varepsilon)$ has a unique non-degenerate solution we may use methods already described to trace out the set of points which give minimum $V(\varepsilon)$ for given F until $\lambda_F = 0$. If only a few bases are feasible for $L_F(\varepsilon)$, i.e., if not too many bases are optimal for $L_E(\varepsilon)$, $\lambda_F = 0$ will be reached quickly. At $\lambda_F = 0$ either $\lambda_E = 0$ or $\lambda_E > 0$. In the former case we have arrived at a point with minimum V <u>and</u> maximum E. In the latter case we have $\bar{X}$ and are ready to trace out the set of efficient X's. From this point on we let $\lambda_F = 0$, i.e., we ignore F completely. Since at $\lambda_F = 0$, $v_k > 0$ for all $k \notin \mathcal{K}$ we may reduce λ_E until we intersect a plane $v_k = 0$ and continue as in Section 8.

10. THE ALGORITHM, WHEN CONDITION 3 DOES NOT HOLD

If E is unbounded procedure 4 of Section 12 can be used to find the point $\underline{X}$ with minimum V. The efficient set can then be traced out in the direction of increasing λ_E. Since there are only a finite number of critical lines and each critical line is efficient at most once the efficient set is traced out in a finite number of steps.

11. THE SET OF EFFICIENT E, V COMBINATIONS

Once the set of efficient X's is found the set of efficient E, V combinations can be obtained easily. The critical line of a subspace in which more than one value of E is obtainable may be expressed as the solution to

$$\Sigma\, \sigma_{jk} X_k + \Sigma(-\lambda_i) a_{ij} + (-\lambda_E)\,\mu_j = 0, \qquad j \in \mathcal{J}, \tag{89}$$

$$\Sigma\, a_{ij} X_j = b_i, \qquad i \in \mathcal{I}, \tag{90}$$

$$\Sigma\, \mu_j X_j = E \tag{91}$$

$$\text{for } -\infty < E < +\infty .$$

If we let N^{-1} be the inverse of the matrix N in (89), (90), (91) we have

(92)
$$\begin{pmatrix} X \\ -\lambda \end{pmatrix} = N^{-1} \begin{pmatrix} 0 \\ b \\ E \end{pmatrix}$$

(93)
$$V = (X', -\lambda') \begin{pmatrix} C & O \\ O & O \end{pmatrix} \begin{pmatrix} X \\ -\lambda \end{pmatrix} = (0', b', E) N^{-1} C N^{-1} \begin{pmatrix} 0 \\ b \\ E \end{pmatrix}$$

from which it follows that along any such critical line V and E are related by a formula of the form

(94)
$$V = a + bE + CE^2 .$$

Thus the set of efficient E, V combinations is piecewise parabolic. We know, or can easily get, the values of E and dV/dE at the end points of each of the pieces. We can also evaluate V at $\overline{X}$.[18] Knowing V at one value of E and $dV/dE = b + 2CE$ at two values of E we can solve for the a, b, and c in (94) for the segment from $\overline{E}$ to $\overline{E} - \varepsilon_1$. Having a, b, and c we can evaluate V at $\overline{E} - \varepsilon_1$ by means of (94). This provides us with the value of V at one value of E on the segment which is efficient from $E - \varepsilon_1$ to $E - \varepsilon_2$. This, combined with the values of dV/dE at two values of E, allows us to obtain the a, b, c of (94) for this next segment—and so on until we trace out the set of E, V combinations.

12. MINIMIZING A QUADRATIC

One of the "by-products" of the calculation of efficient sets is the point at which V is a minimum, i.e., where $\lambda_E = 0$. The computing procedures described in Sections 6 through 10 are analogous to the simplex method of linear programming (as contrasted with the "gradient methods" that have been suggested for both linear and non-linear programming). Both the procedure described in the preceding section—considered as a way of getting to min V—and the simplex method require a finite number of iterations, each iteration typically taking a "jump" to a new point which is superior to the old. Each iteration makes use of the inverse of a matrix which is a "slight" modification of the matrix of the previous iteration. The success of the simplex method in linear programming suggests that it may be desirable to use a variant of the "critical line" method in the quadratic case.

Our problem then is to minimize a quadratic

$$V = \Sigma \Sigma \sigma_{ij} X_i X_j$$

subject to constraints (1), (2) and (3). We wish to translate this into a problem of tracing out an efficient set. This may be done in several ways.

1. An arbitrary set of μ_j can be selected and the efficient set traced out until $\lambda_E = 0$. The μ_j should be selected so that the "artificial" E has a unique maximum.

2. An equality, say,

$$\Sigma a_{1j} X_j = b_1$$

[18] If $\overline{X}$ does not exist we can evaluate V at $\underline{X}$ and use the same process "in reverse."

can be eliminated from (1). E can be defined as

$$E = \Sigma a_{1j} X_j$$

and the critical set traced out until $E = b_1$. If $E = b_1$ is reached before $\lambda_E = 0$ the computing procedures of the last section must be continued into the region of $\lambda_E < 0$. While the points thus generated will not be efficient—for they do not give max E for given V—they do give min V for given E. In particular, they will arrive at the point of min V for

$$E = \Sigma a_{1j} X_j = b_1 .$$

3. An inequality, say,

$$\Sigma a_{mj} X_j \geq b_m$$

can be eliminated from (2). E can be defined as

$$E = \Sigma a_{mj} X_j .$$

The efficient set is traced out until either $E = b_m$ or else $\lambda_E = 0$. If the former happens first the constraint is effective; if the latter happens first the constraint is ineffective. In either case, the point associated with the first of these to occur gives min V subject to (1), (2), and (3).

4. An initial guess $X_1^0, \ldots, X_N^0$ which satisfies (1), (2), and (3) can be made and μ_j defined so that, given these μ_j, X^0 is efficient. The efficient set can then be traced out until $\lambda_E = 0$. To choose μ_j so that X^0 is efficient, choose arbitrary positive values of λ_i ($i \in \mathcal{S}$) and λ_E. Then choose μ_j so that

$$\eta_j = 0 \quad \text{for } X_j \text{ not at its lower bound,}$$

$$\eta_j > 0 \quad \text{for } X_j \text{ at its lower bound.}$$

If X^0 is in the same subspace as the optimal solution, the latter is reached in one iteration.

---

[19]It has been found recently that the strict convexity assumption can be relaxed. The procedures described in this paper apply, without modification, to the homogeneous quadratic whenever (σ_{ij}) is positive semi-definite. The non-homogeneous quadratic requires a slight modification of procedure for the general semi-definite (σ_{ij}). (Footnote added in proof.)

* * *