APPLICATIONS OF MANAGEMENT SCIENCE

A Research Annual

Editor: RANDALL L. SCHULTZ
School of Management and Administration
The University of Texas at Dallas

VOLUME 2 · 1982

JAI PRESS INC.

Greenwich, Connecticut *London, England*

36 Sherwood Place
Greenwich, Connecticut 06836

JAI PRESS INC.
3 Henrietta Street
London WC2E 8LU
England

ISBN NUMBER: 0-89232-258-6

Manufactured in the United States of America

APPLICATIONS OF
MANAGEMENT SCIENCE

Volume 2 · 1982

LIST OF CONTRIBUTORS

Peter Duchessi	Office of Health Systems Management, New York State Department of Health
Salah E. Elmaghraby	Department of Operations Research, North Carolina State University
L.R. Foulds	Department of Operations Research, University of Canterbury, New Zealand
Christopher Gimbrone	Office of Health Systems Management, New York State Department of Health
Edward L. Hannan	Office of Health Systems Management, New York State Department of Health
Michael W. Lawless	School of Business Administration and Economics, California State University, Northridge
Gary L. Lilien	College of Business Administration, Pennsylvania State University
Ronald W. Morrison	Department of Economics, Central Michigan University
A.M. Salem	Department of Operations Research, North Carolina State University

Josef Schmee — Union College and University, Schenectady

C.J. Schumaker, Jr. — Department of Economics, Central Michigan University

Michael R. Summers — Business Administration Division, Pepperdine University

PREFACE

Volume 2 of *Applications of Management Science* continues to present solid research aimed at the resolution of contemporary management problems. The purpose of the series is to show how the utilization of operations research, management science, decision science, and management information system technology can improve decision making in organizations. In addition, the series seeks to demonstrate the problems and opportunities for management science in the area of public policy and at the interface of public and private policy

It has been clear for some time that a considerable gap exists between theoretical developments in management science and realized applications in organizations. On the one hand, this situation has motivated research on the implementation of models and decision support systems; we know more about the behavioral process of implementation and have a better understanding of how to manage the change necessary to bring about implementation. On the other hand, this situation has also brought about a change in the way management science models are developed; we are increasingly concerned with representing problems as they are, even if such representations are less elegant or less tractable than we would wish. However, concern with implementation and with the representativeness of problems does not preclude the highest-quality management science work.

Applications of Management Science is an outlet for original research in management science and is distinguished by its form, its frequency of appearance, and its focus. The series is essentially a research anthology of papers that are substantive and may exceed the length limitations of traditional journal articles. Although any work dealing with the application of management science is appropriate, the "longer form" provides an outlet for papers that is not otherwise available. The series also provides an outlet for papers presented at symposia that are refereed to journal standards. The series includes both theoretical and methodological papers so long as they are extended toward application; in addition, comprehensive review articles are published. Excluded from *Applications of Management Science*, although appropriate for many other journals, are strictly theoretical or methodological developments, such as work on efficient algorithms. Also excluded are papers that do not directly concern decision making in organizations, such as the applied mathematics of sports.

All papers appearing in *Applications of Management Science* are refereed, and I am grateful to those who served as reviewers for each of the manuscripts in Volume 2. My primary debt, of course, is to the authors. By working at the boundary of theory and practice, they have helped to legitimize the mission of management science and to offer direct proof of its applicability to management decision making.

Randall L. Schultz
Series Editor

CONTENTS

OPTIMAL PROJECT COMPRESSION UNDER QUADRATIC COST FUNCTIONS

S.E. Elmaghraby and A.M. Salem

ABSTRACT

We are given a directed acyclic network representing a project, and the cost-time trade-off function for each activity, which is assumed convex and quadratic in the activity duration. It is desired to determine the optimal activity durations that achieve a desired completion date with minimum cost.

Section 1 deals with the special case of continuous derivatives, while Section 2 deals with the general case that permits discontinuities in the derivative at the upper and lower bounds of the activity duration. Efficient algorithms are developed to obtain respective optima. In addition to satisfying the stated objective, the algorithms also yield the optimal cost for all durations between the specified duration and the constrained project duration, which is helpful for the purpose of sensitivity analysis. The

Applications of Management Science, Volume 2, pages 1–39.

ISBN: 0-89232-258-6

algorithms have been programmed on the computer, and computational experience is provided.

1. QUADRATIC COST WITH CONTINUOUS DERIVATIVE

1.1. The Problem

A central problem in the study of deterministic activity networks (DANs) is that of optimal project compression. Simply stated, it runs as follows: Given a project and its associated graph G = (N, A) of N nodes (representing project "events") and A arcs (representing project "activities"[1]) and a function $C_{ij}(y_{ij})$ representing the variation of activity cost with its duration y_{ij}, determine the optimal allocation of specified available funds W to the various activities to terminate the project as early as possible; or, alternatively, determine the minimum funds required to complete the project no later than a specified time T_s. In this latter version, the problem is meaningful only if T_s is less than the "normal" duration of the project; i.e., if T_s is less than the length of the critical path (CP) that would result when each activity is run at its "normal" level (assumed the cheapest). To avoid trivialities, we assume that T_s is feasible; i.e., it is not less than the length of the CP that results when each activity is run at its "crash" duration.

The practical significance of this problem resides in the ability to specify the most efficient utilization of investments in the "speeding up" of the project. Alternatively, it serves to alert the manager to the range of requirements of additional investments should he wish to deviate appreciably from the "normal" flow of work in the project.

The mathematical statement of the problem runs as follows:

$$\text{Minimize} \sum_{(ij)\in A} C_{ij}(y_{ij}) \tag{1}$$

such that the precedence constraints are respected, and the project is completed on or before time T_s. Let t_i denote the time of realization of node **i**. Then if activity (ij) ∈ A, with the arrow in the direction **i** → **j**, we must impose the restriction

$$-t_i + t_j - y_{ij} \geqslant 0, \qquad \forall (ij) \in A. \tag{2}$$

The completion time requirement adds the constraint

$$t_1 - t_n \geqslant -T_s. \tag{3}$$

Here we assume that the "start node" is node **1**, and that the "terminal" node is **n**; whence the set N ≡ {1, 2, …, n}. Finally, the activity duration

y_{ij} is bound from below by a lower limit $\ell_{ij} \geqslant 0$, and from above by an upper limit $u_{ij} > \ell_{ij}$; i.e., $0 \leqslant \ell_{ij} \leqslant y_{ij} \leqslant u_{ij}$. (The only instance in which y_{ij} is permitted to be 0 is in the case of "dummy" activities; see Ref. 4 for a detailed explanation of the utility of these activities.) It is more convenient to rewrite this double inequality as

$$-y_{ij} \geqslant -u_{ij} \quad \text{and} \quad y_{ij} \geqslant \ell_{ij}, \qquad \forall (ij) \in A. \tag{4}$$

The mathematical program (1)–(4) has been extensively studied under the various manifestations of the individual time-cost function C_{ij}: linear (Kelly [8] and Fulkerson [6]), convex decreasing (Berman [1], Clark [2], Lamberson and Hocking [10], and Elmaghraby [3]), concave decreasing (Falk and Horowitz [5]), and discontinuous nonincreasing (Robinson [12]); for a succinct summary of these approaches, see Elmaghraby [4].

Section 1 of this paper is devoted to the case in which $C_{ij}(y_{ij})$ is quadratic decreasing with continuous derivative $C'_{ij}(= \partial C_{ij}/\partial y_{ij})$ in the domain $y_{ij} \in [\ell_{ij}, \infty)$. The case of the discontinuous derivative is the subject of Section 2.

It may appear that the problem is a "straightforward" application of quadratic (in fact, separable) programming, and is thus amenable to resolution by standard approaches. This is indeed true. However, in a vein similar to the linear case, we trust that a specialized algorithm that capitalizes on the special structure of the problem would be at least an order of magnitude more efficient than a general procedure. The remainder of this paper is devoted to the development of such a specialized algorithm.

1.2. Analytical Results

Since C is quadratic with continuous derivative in the interval $[\ell_{ij}, \infty)$, we may assume it, without any loss of generality, to be of the form

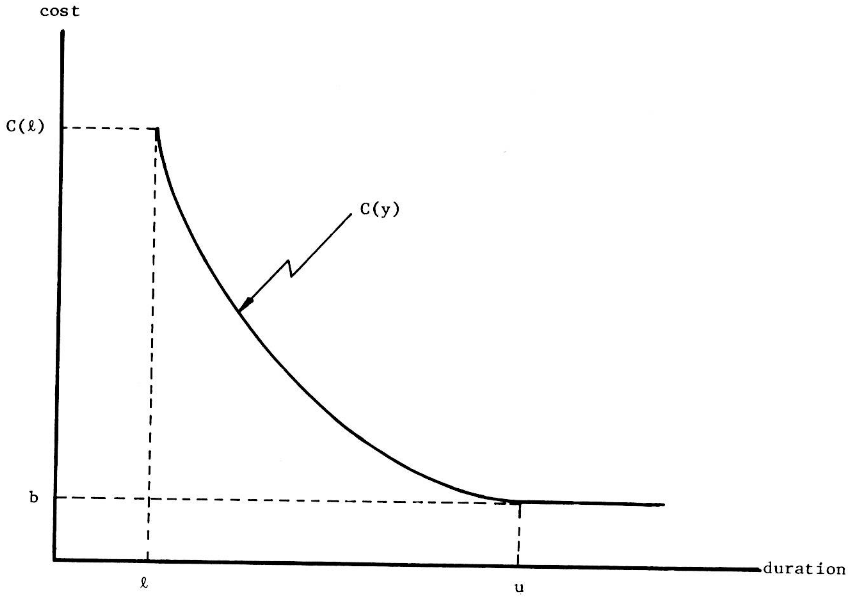
$$C_{ij}(y_{ij}) = b_{ij} + \beta_{ij}(u_{ij} - y_{ij})^2; \qquad \ell_{ij} \leqslant y_{ij} \leqslant u_{ij}. \tag{5}$$

Note that C_{ij} is tangent to the line $C_{ij}(y_{ij}) = b_{ij}$ at $y_{ij} = u_{ij}$, at which point $C'_{ij} = 0$; see Figure 1.

We introduce one mild assumption whose justification is easily established: the specified completion time T_s is such that no activity will be *at* its lower bound; i.e., at the optimum, $y_{ij} > \ell_{ij}$ for all $(ij) \in A$. Note that if ℓ_{ij} is small enough relative to u_{ij}, and T_s is not too "tight," this condition will be automatically satisfied. We refer to it as *Condition L*.

In Ref. 3, the following characterization of the optimum was given. Let the nodes of the network be realized at times $0 = t_1 \leqslant t_{i_2} \leqslant t_{i_3} \leqslant \cdots \leqslant t_n = T_s$, where t_{i_k} corresponds to the kth earliest node. Let $D(\tau)$ denote the sum of the derivatives C'_{ij} of the activities that are "in progress" at time τ. Then the given schedule of activities is optimal if $D(\tau)$ is constant for all $\tau \in [0, T_s]$.

Figure 1. The Quadratic Function $C(y) = b + \beta(u - y)^2$; for $\ell \leq y \leq u$.



There are two remarks to be made about that result. First, it was proved by elementary variational-type arguments. Second, though it characterized the optimum, it gave no procedure for achieving it. The following development pertains to these two remarks.

The function $C_{ij}(y_{ij})$ is evidently convex; add to this that all the constraints are linear, and the conclusion immediately follows that the necessary conditions for optimality of Kuhn and Tucker (K–T) for general nonlinear programming are also sufficient.

Let the Lagrange multipliers be f_{ij} for (2), q for (3), and g_{ij} and h_{ij} for (4). These K–T conditions are [stated as equalities because the y terms are unrestricted in sign in the program (1)–(4)]—

Stationarity conditions:

$$C'_{ij} - f_{ij} - g_{ij} + h_{ij} = 0, \qquad (ij) \in A. \tag{6.1}$$

Flow conditions:

$$\sum_j (f_{ji} - f_{ij}) = \begin{cases} -q & \text{for } i = 1 \\ 0 & \text{for } i \neq 1, n \\ q & \text{for } i = n; \qquad (ij) \text{ and } (ji) \in A. \end{cases} \tag{6.2}$$

Complementary slackness conditions:

$$(-t_i - y_i + t_j)f_{ij} = 0 \tag{6.3}$$

$$(t_1 - t_n + T_s)q = 0 \tag{6.4}$$

$$(-y_{ij} + u_{ij})g_{ij} = 0 \tag{6.5}$$

$$(y_{ij} - \ell_{ij})h_{ij} = 0 \tag{6.6}$$

$$f_{ij},\ g_{ij},\ h_{ij}, \text{ and } q \geq 0. \tag{6.7}$$

We are interested only in activities for which $u > \ell$; since $u = \ell$ implies that the duration is a constant equal to ℓ, in contradiction to Condition L. At the optimum we know, by Condition L, that $\ell_{ij} < y^*_{ij} \leq u_{ij}$, for all $(ij) \in A$. First, consider the activities of durations $y^*_{ij} = u_{ij}$: (6.6) $\Rightarrow$ $h^*_{ij} = 0$, and we know that $C'_{ij}(u_{ij}) = 0$. Consequently, (6.1) $\Rightarrow$ $-f^*_{ij} - g^*_{ij} = 0$, which, by condition (6.7), implies that $f^*_{ij} = 0 = g^*_{ij}$. We therefore conclude that such activities do not contribute to the sum of (6.2).

Next consider the activities of durations $y^*_{ij} \in (\ell_{ij}, u_{ij})$: (6.5) and (6.6) $\Rightarrow$ $g^*_{ij} = 0 = h^*_{ij}$; then (6.1) $\Rightarrow$ $f^*_{ij} = C'_{ij}(y^*_{ij})$. Finally, (6.2) $\Rightarrow$ $D(\tau) =$ constant ($= q^*$) for all $\tau \in [0, T_s]$, which is precisely the condition of Elmaghraby [3]. This provides an alternate and more direct proof of that result, albeit not elementary.

We turn now to the problem of algorithmic procedures to achieve the optimum.

Assertion 1.1: The necessary and sufficient conditions for optimality stated above are equivalent to the conditions

$$\sum_j (\beta_{ji}d_{ji} - \beta_{ij}d_{ij}) = \begin{cases} -a & \text{for } i = 1 \\ 0 & \text{for } i \neq 1, n \\ a & \text{for } i = n, \end{cases} \tag{7}$$

where d_{ij} is the reduction in activity (ij), and a is some constant (actually $= q^*/2$). *Proof:* Just substitute for the value of $C'_{ij}(y_{ij})$ in (6.2). ■

There are two pertinent remarks at this point. First, it is evident that one need not deal with the *total* reduction in duration d_{ij}, but that it is sufficient to write (7) for the incremental reductions $\{d^{(r)}_{ij}\}$ carried out in iteration r. Second, it is equally evident that Equations (7) have the form of "conservation of flow" equations (in the normalized variables $\partial_{ij} = d_{ij}/\beta_{ij}$), in which one equation is redundant. In the sequel we shall consider the first equation (corresponding to node $i = 1$) to be the redundant one, and eliminate it from consideration, which is equivalent to putting $t_1 = 0$.

As a preliminary to the following results, we introduce a few concepts and notations. Define the "critical subnetwork" (CSN) to be the set of longest paths in the network and let K denote its cardinality; i.e., K = |CSN|. Clearly, the CSN changes from iteration to iteration (since it is augmented by new paths that become critical as the project's duration is shortened), and we may sometimes resort to the notation $CSN^{(r)}$ to denote the CSN at a particular iteration, and set $K^{(r)} = |CSN^{(r)}|$. The networks of concern to us are directed and *acyclic,* i.e., contain no loops (in the usual sense). Still, we wish to characterize two simple paths (or subpaths) that have the same start and terminal nodes; we shall refer to them as *loops,* it being understood that "going around the loop" must, necessarily, move in the direction of the arrows from start to terminal and *against* the arrows from terminal to start.

Every subgraph of G which contains N − 1 arcs and has no loops is a *spanning tree*. Any tree of G has A − N + 1 chords; and each chord of a tree is contained in a unique loop which contains no other chords of the tree. The set of A − N + 1 loops corresponding to the chords of the tree is called a *fundamental set of loops,* because any other loop may be obtained as a linear combination of the elements of this fundamental set.

Assertion 1.2: If the CSN contains K arcs, there shall be K simultaneous linear equations relating the values of the individual (incremental) reductions $\{d_{ij}\}$ to the constant **a** at the rth iteration. *Proof:* Consider a particular iteration. If the CSN has K arcs and m nodes, we must have $K \geq m - 1$. Assertion 1, applied to CSN, would result in m − 1 (independent) equations. Additionally, there are K − m + 1 fundamental loops (see Ref. 11, p. 397), yielding that many linear equations in the d_{ij} terms (each loop equates the total amount of reduction in the critical activities in the two branches of the loop to maintain their criticality) which are *independent.* The sum is K independent equations, as asserted. ■

For ease of reference, we denote the resulting matrix of coefficients by B, and represent the set of equations in the concise form

$$BD = a\mathbf{e}_{m-1}, \tag{8}$$

where $B^{(r)}$ is a $K^{(r)} \times K^{(r)}$ square matrix; $D^{(r)}$ is a $K^{(r)} \times 1$ vector of (incremental) reductions $\{d_{ij}^{(r)}\}$; $a^{(r)}$ is a scalar; and $\mathbf{e}_{m(r)-1}$ is a $K^{(r)} \times 1$ vector of zeros except in position $m^{(r)} - 1$ where it has a 1.

The (unique) solution of (8) gives the values of $d_{ij}^{(r)}$ in terms of $a^{(r)}$, which we write as

$$d_{ij}^{(r)} = v_{ij}^{(r)} a^{(r)}. \tag{9}$$

It is easy to see that v_{ij} is given by the (ij)*th* entry in column m − 1 of the inverse of B. The matrix B itself is easily determined from any *arborescence* on the CSN (that is, a tree rooted at node **1** such that there is a unique directed simple path from node **1** to each node **j** in the CSN), and its corresponding chords, together with the system of Equations (7). Therefore, there remains only the determination of the value of $a^{(r)}$ to completely determine the necessary reduction in each activity in the current $CSN^{(r)}$.

Slight reflection reveals that $a^{(r)}$ measures the current reduction in the value of t_n, the total project duration.[2] Evidently, such reduction is bounded by two eventualities: either a noncritical path becomes critical (necessitating the redefinition of the CSN) or the specified project duration T_s is achieved. Let a_1 and a_2 denote these two bounds on the value of **a**, respectively.[3] Clearly, $a = \min(a_1, a_2)$. We concentrate on the determination of a_1, since a_2 is trivially obtained from the difference between the current value of t_n and T_s.

The basic idea for the determination of the value of the bound a_1 runs as follows. At the r*th* iteration, the arcs of G are partitioned into two mutually exclusive subsets: CSN and NCSN (for *noncritical subnetwork*). Note that while CSN is connected and contains nodes **1** and **n**, the graph NCSN need not be connected and may *not* contain nodes **1** and **n**. Some nodes may be repeated between CSN and NCSN, since criticality (and noncriticality) is defined relative to the arcs (= activities).

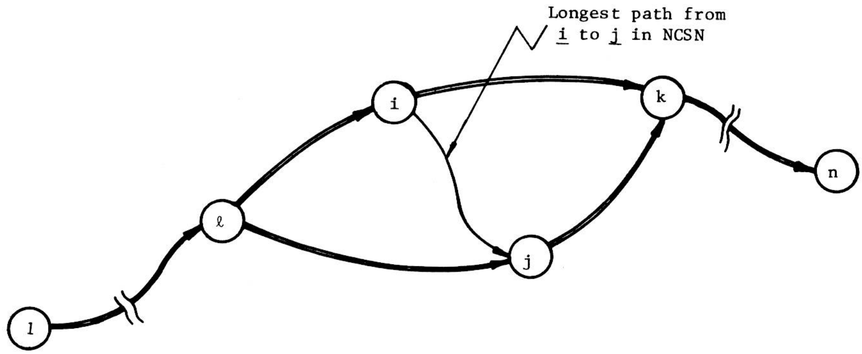
Consider any pair of nodes **i** and **j**, $i < j$, in CSN; either there exists a path in CSN directed from **i** to **j**, or none exists. Consider the same pair of nodes from the point of view of NCSN; we have two possibilities:

1. Either node, or both, are missing from NCSN. Then this pair will have no influence on the value of a_1.
2. Both nodes are present in NCSN. Then either there exists no path (in NCSN) from **i** to **j**, in which case this pair of nodes will also have no influence on the value of a_1; or there exists at least one path (in NCSN) from **i** to **j**. In the latter case, determine the duration of the longest path (in NCSN) from **i** to **j**.

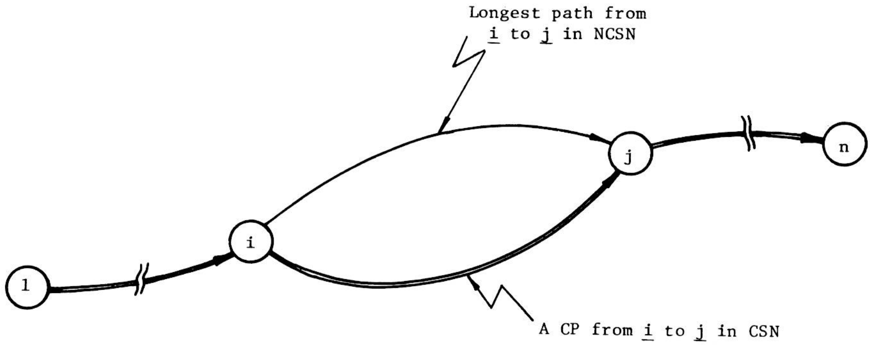
Our analysis thus far leads us to one of two situations depicted schematically in Figure 2:

(a) No path exists between **i** and **j** in CSN, but a path exists in NCSN. Then, by the structure of CSN, there must exist two nodes ℓ and **k**, with $1 \leq \ell < i < j < k \leq n$, such that there are *two* simple critical paths from ℓ to **k**, one containing node **i** and

Figure 2. The Two Possible Eventualities in the Determination of a_1.



(a) No path exists between i and j in CSN (a cross-over exists)



(b) A path exists between i and j in CSN (a loop exists)

the other containing node **j**, with the noncritical path (in NCSN) from **i** to **j** "crossing over" as shown in Figure 2(a).

(b) A path exists between **i** and **j** in CSN and a path exists in NCSN. These two paths must form a loop, as depicted in Figure 2(b).

To determine the value of a_1 under either eventuality, denoted by $a_1(ij)$ since it depends on the pair **i** and **j**, we reason as follows:

Denote any path (or subpath) in CSN between nodes **i** and **j** by $\pi(ij)$, and the longest path in NCSN between the same two nodes by P(ij). Clearly, if the length of path $\pi(ij)$ is reduced by an amount $\delta > 0$, it must be true that $\delta = a_1\Sigma_{\omega\in\pi(ij)} v_\omega$ for some value of a_1, where the summation is taken over all activities $\{\omega\}$ that lie on the path $\pi(ij)$. It