A PARAMETRIC CHARACTERIZATION AND AN 
$\varepsilon$-APPROXIMATION SCHEME FOR THE MINIMIZATION 
OF A QUASICONCAVE PROGRAM

Naoki KATOH

Department of Management Science, Kobe University of Commerce, Tarumi, Kobe, Japan 655

Toshihide IBARAKI

Department of Applied Mathematics and Physics, Faculty of Engineering, Kyoto University, Kyoto, Japan 606

Received April 1985
Revised May 1986

This paper deals with the problem $P$ of minimizing a quasiconcave function over a given feasible region. We first introduce an auxiliary problem $P(\lambda)$ with a parametric vector $\lambda$ such that, for an appropriate $\lambda$, its optimal solution is also optimal to the original problem. Based on this, an approximation scheme for $P$ is developed. If $P(\lambda)$ is polynomially solvable, this becomes a polynomial time approximation scheme. In particular, we show that fully polynomial time approximation schemes can be developed for a large class of stochastic programming problems with 0–1 variables in which cost coefficients are subject to independent normal distributions, if their deterministic versions obtained by replacing cost coefficients by constants have polynomial time algorithms or fully polynomial time approximation schemes (e.g., problems of shortest path, assignment, minimum cut, 0–1 knapsack and minimum directed spanning tree).

1. Introduction and outline of the paper

Maximizing a concave or quasiconcave function over a convex feasible region has been a target of extensive study (see, for example, Avriel [1] and Rockafellar [24]). Contrary to this, minimizing such a function has received less attention (see survey papers by Heising-Goodman [10], Hoffman [11] and McCormick [19] for the case in which the feasible region is described by a set of linear inequalities). In this paper we study the latter subject:

$$P: \text{minimize } z(x) = h(f_1(x), f_2(x), \ldots, f_m(x)).$$  \hspace{1cm} (1)

Here $x$ denotes an $n$-dimensional decision vector, which may be real or integral depending on the cases, and $X$ denotes a feasible region. Functions $f_j$, $j = 1, 2, \ldots, m$, are real-valued functions and $h(u_1, u_2, \ldots, u_m)$ is quasiconcave over a convex set $U$ such that
\[ U \supseteq V, \quad V = \{(f_1(x), f_2(x), \ldots, f_m(x)) \mid x \in X\}. \]  

(Note that \( h \) may not be quasiconcave in \( x \).) We assume throughout the paper that \( m \) is a constant independent of \( n \).

This paper starts with a parametric characterization of \( P \), stating that an optimal solution of the parametric problem \( P(\lambda) \) defined below provides an optimal solution of \( P \), if an appropriate \( \lambda \) is chosen.

\[ P(\lambda): \quad \text{minimize} \sum_{x \in X} \lambda_j f_j(x), \]

where \( \lambda = (\lambda_1, \lambda_2, \ldots, \lambda_m) \) is an \( m \)-dimensional real parameter vector. Thus, solving \( P \) is reduced to finding a \( \lambda = \lambda^* \) with which an optimal solution to \( P(\lambda^*) \) is also optimal to \( P \).

The same characterization for \( m = 2 \) with concave \( h(u_1, u_2) \) has recently been given by Sniedovich [29, 30, 31]. Some special cases have also been reported (e.g., Kataoka [15], Ishii et al. [13] and Ichimori et al. [12] discuss some types of stochastic programs, and Dinkelbach [5] and Jagannathan [14] discuss the fractional program). Similar results are also known for the problem that maximizes a concave or strictly quasiconcave function over a convex set \( X \) (see Geoffrion [7] and Schaible [25]). Based on this characterization, [12, 13] give polynomial time exact algorithms for some stochastic programs in which \( X \) represents the set of spanning trees of a graph. As such \( P(\lambda) \) has only a polynomially bounded number of optimal solutions when \( \lambda \) changes from \(-\infty \) to \( \infty \), the exhaustive search of optimal \( \lambda = \lambda^* \) can yield polynomial time algorithms for \( P \). [16] contains another example, a chance-constrained single machine scheduling problem, for which a polynomial time exact algorithm can be constructed.

However, the number of optimal solutions of \( P(\lambda) \) over the entire range of \( \lambda \) is not polynomially bounded in most cases, e.g., see Carstensen [2]. Therefore, in general, polynomial time algorithms seem to be difficult to develop, and we focus on approximation schemes in this paper. A solution is said to be an \( \varepsilon \)-approximate solution if its relative error is bounded above by \( \varepsilon \). An approximation scheme is an algorithm containing \( \varepsilon > 0 \) as a parameter such that, for any given \( \varepsilon \), it can provide an \( \varepsilon \)-approximate solution. If it runs in time polynomial in the input size of each problem instance, it is a polynomial time approximation scheme. If it is polynomial in both input size and \( 1/\varepsilon \), the scheme is called a fully polynomial time approximation scheme [6, 23].

Our idea is to solve \( P(\lambda) \) only for a polynomially bounded number of \( \lambda \)'s, which are systematically generated so that the relative error of the achieved objective value is within \( \varepsilon \). The required time is

\[ O(p(n)m(2 \log(D/D))^{m-1}/\log^{m-1}(1 + \delta(\varepsilon, m))), \]

where \( p(n) \) is the time required to solve a \( P(\lambda) \) exactly, and \( D \) and \( \delta \) and \( \delta(\varepsilon, m) \) will be specified later. We shall discuss that, if each \( P(\lambda) \) can be polynomially solved,
The minimization of a quasiconcave program

this is in many cases polynomially bounded in the input size of $P$ and sometimes also in $1/\varepsilon$. In addition, we shall show that this approach can be extended to the case in which $P(\lambda)$ has a fully polynomial time approximation scheme.

The plan of this paper is as follows. Section 2 provides several examples of problem $P$ selected from various application fields. These include some types of stochastic programs, an optimization problem associated with the markovian decision process, and a problem encountered in VLSI chip design. Based on the basic concepts explained in Section 3, Section 4 gives key theorems relating $P$ to $P(\lambda)$. Section 5 states some assumptions which are required in the subsequent development. With the partition scheme of parameter space described in Section 6, Section 7 develops an approximation scheme and analyzes its running time, under the assumption that $P(\lambda)$ has a polynomial time exact algorithm. Section 8 modifies it to the case in which $P(\lambda)$ has a fully polynomial time approximation scheme. Section 9 deals with a special case of $m=2$ and shows that one of the assumptions stated in Section 5 can be relaxed. Section 10 argues that objective functions of some problems described in Section 2 satisfy the assumptions fo Section 5. Therefore, as discussed in Section 11, (fully) polynomial time approximation schemes exist for such problems.

2. Some examples of problem $P$

We list here several examples of minimizing quasiconcave objective functions from various application areas. We shall see in this paper that (fully) polynomial time approximation schemes can be developed for some of these problems.

A class of problems can be defined in relation with the following stochastic programming problem with 0–1 variables (e.g., [3,27,33]):

$$
\text{minimize } \sum_{i=1}^{n} c_i x_i
$$

subject to $x = (x_1, x_2, \ldots, x_n) \in X$,

where the cost coefficients $c_i$ are random variables subject to an independent normal distribution $N(m_i, \sigma_i)$ with nonnegative means $m_i$ and variances $\sigma_i^2$, which are assumed to be integers. The following two optimality criteria are often employed [3,15,27,33]:

(a) Chance-constrained minimization. Minimize $I$ such that Prob($\sum_{i=1}^{n} c_i x_i \leq I$) $\geq \alpha$, where $\alpha$ is a given constant satisfying $\frac{1}{2} < \alpha < 1$.

(b) Probability maximization. Maximize Prob($\sum_{i=1}^{n} c_i x_i \leq d$), where $d$ is a given constant satisfying $d > \min_{x \in X} \sum_{i=1}^{n} m_i x_i$.

The stochastic problems with these criteria can be transformed into the following deterministic problems $S_1$ and $S_2$, respectively, as shown in [12,13,15].

$$
S_1: \underset{x \in X}{\text{minimize }} \left\{ \sum_{i=1}^{n} m_i x_i + \beta \left( \sum_{i=1}^{n} \sigma_i^2 x_i \right)^{1/2} \right\},
$$
where $\beta$ is a positive constant determined from $\alpha$, and $\sum v_i^2 x_i > 0$ for any $x \in X$ is assumed in the second case. Letting $h_1(u_1, u_2) = u_1 + \beta u_2$, we see that problem $S_1$ is a special case of $P$. Since maximizing $(d - \sum m_i x_i) / (\sum v_i^2 x_i)^{1/2}$ is equivalent to minimizing $(\sum v_i^2 x_i)^{1/2} / (d - \sum m_i x_i)$ by $d - \sum m_i x_i > 0$ and $\sum v_i^2 x_i > 0$, $S_2$ is also a special case of $P$ with $h_2(u_1, u_2) = \sqrt{u_2} / (d - u_1)$, which is quasiconcave.

Another example in this class is the following variance minimization problem associated with the markovian decision process. Consider a system with a finite set of states $S$. At every discrete time instance $t = 1, 2, \ldots$, one action $k \in K_i$ (where $K_i$ is a given set of actions) is chosen with probability $d_i^k$ if the current state is $i$, which then produces reward $r_i^k$ and causes the state transition to $j$ with probability $p_i^{kj}$. A central issue of the markovian decision processes is to determine $d_i^k$ for all $i$ and $k$ that maximizes the expected reward per unit time over the infinite horizon. It is known [3, 9] that this problem is formulated as the following linear programming problem.

\begin{equation}
\text{LP:} \quad \text{maximize} \quad \sum_{i \in S} \sum_{k \in K_i} r_i^k x_i^k \\
\text{subject to} \quad \sum_{k \in K_i} x_i^k = \sum_{j \in S} \sum_{k \in K_i} p_i^{kj} x_j^k, \quad i \in S, \\
\sum_{i \in S} \sum_{k \in K_i} x_i^k = 1, \\
x_i^k \geq 0, \quad i \in S, \quad k \in K_i.
\end{equation}

Letting $\bar{x}_i^k$ be an optimal solution of LP, optimal $d_i^k$ are given by

\[ d_i^k = \frac{x_i^k}{\sum_{k \in K_i} x_i^k}. \]

As a modification of this problem, we may want to minimize the variance under the constraint that the expected reward is not smaller than a given constant $M$.

\begin{equation}
\text{VP:} \quad \text{minimize} \quad \left\{ \sum_{i \in S} \sum_{k \in K_i} (r_i^k)^2 x_i^k - \left( \sum_{i \in S} \sum_{k \in K_i} r_i^k x_i^k \right)^2 \right\} \\
\text{subject to} \quad \sum_{i \in S} \sum_{k \in K_i} r_i^k x_i^k \geq M.
\end{equation}

Let \[ f_1(x) = \sum_{i \in S} \sum_{k \in K_i} (r_i^k)^2 x_i^k, \quad f_2(x) = \sum_{i \in S} \sum_{k \in K_i} r_i^k x_i^k \]
and $h(u_1, u_2) = u_1 - u_2^2$. This VP is a special case of $P$ since $h$ is quasiconcave.

As an additional example, we mention here that the problem of chip area minimization encountered in VLSI design. A VLSI chip is composed of a number of rec-
tangular blocks, whose relative positions in a chip are specified in advance. The so-called compaction is one of the useful approaches to achieve the minimum area and has been well studied (see, for example, [17,26,34]). The compaction problem normally assumes that the size (i.e., width and height) of each block is fixed. Another type of problem studied by [21] (see also [18,28]) considers the size of each block to be a decision variable.

In either case, the problem is described as the minimization of chip area \( x_1 x_2 \), where \( x_1 \) and \( x_2 \) are the width and height of the resulting chip, respectively, under the given design constraints. This is again a special case of \( P \), because the objective function \( h(u_1, u_2) = u_1 u_2 \) with \( f_1(x_1) = x_1 \), and \( f_2(x_2) = x_2 \) is quasiconcave.

As we shall see in the subsequent discussion, the above stochastic problems \( S_1 \) and \( S_2 \) have fully polynomial time approximation schemes, if \( P(\lambda) \) is one of such problems as shortest path, assignment, minimum (directed) spanning tree, minimum weight matching, minimum cut and 0–1 knapsack.

In the case of the variance minimization problem of the markovian decision process, \( P(\lambda) \) is a parametric linear programming problem. It is known that the number of distinct optimal solutions generated over the entire range of \( \lambda \) is exponential in the worst case (see [22]). Unfortunately, the theory of this paper does not directly apply to this case, and some modifications are necessary to warrant a polynomial time approximation scheme. Such treatment will be discussed elsewhere.

Finally, \( P(\lambda) \) of the chip minimization problem is the mixed integer program [21], for which the computation of an exact optimal solution or even an \( \varepsilon \)-approximate solution seems quite difficult. At present it does not seem to be possible to develop a polynomial time approximation scheme for this problem.

3. Basic concepts

A function \( h : U (\subset \mathbb{R}^m) \to \mathbb{R} \) is quasiconcave if for any two points \( u^1, u^2 \in \mathbb{R}^m \) with \( u^1 \neq u^2 \) and \( q_1, q_2 \geq 0 \) with \( q_1 + q_2 = 1 \),

\[
h(q_1 u^1 + q_2 u^2) \geq \min\{h(u^1), h(u^2)\}
\]

holds. Concavity obviously implies quasiconcavity. For a set \( S \subset \mathbb{R}^m \), a point \( u \in S \) is an interior point if, for some \( \varepsilon > 0 \), all points \( u' \) satisfying \( \|u' - u\| < \varepsilon \) are contained in \( S \), where \( \| \cdot \| \) denotes the Euclidean norm. The set of interior points of \( S \) is denoted \( \text{int}(S) \). The smallest closed set that contains a set \( S \) is called the closure of \( S \), and is denoted \( \text{cl}(S) \). A point \( u \in \mathbb{R}^m \) is called a boundary point of \( S \) if it belongs to \( \text{cl}(S) \) but is not an interior point of \( S \). The set of boundary points of \( S \) is denoted \( \text{bd}(S) \).

For a convex set \( S \subset \mathbb{R}^m \), a set

\[
H = \{u \in \mathbb{R}^m \mid c(u - u') = 0\}
\]

defined by \( c (\neq 0) \in \mathbb{R}^m \) is the supporting hyperplane of \( S \) at a boundary point \( u' \) if
the closed half space

$$HS = \{ u \in \mathbb{R}^m | c(u-u') \geq 0 \}$$  \hspace{1cm} (10)

contains S, i.e.,

$$c(u-u') \geq 0 \text{ for all } u \in S.$$  \hspace{1cm} (11)

The convex hull $\text{co}(S)$ of a set $S \subset \mathbb{R}^m$ is the smallest convex set that contains S. If S is a finite set, $\text{co}(S)$ is a convex polytope. A point $u$ in a convex polytope $K$ is called a vertex if $u = \mu u' + (1-\mu)u''$, with $u', u'' \in K$ and $0 < \mu < 1$ implies $u = u' = u''$.

**Lemma 3.1** [8,24]. (i) For a finite set $S \subset \mathbb{R}^m$, any point $u \in S$ can be represented by a convex combination of vertices $u_1, u_2, ..., u_p$ of $\text{co}(S)$, i.e.,

$$u = \sum_{k=1}^{p} \mu_k u_k \text{ for some } \mu_k \geq 0 \text{ with } \sum_{k=1}^{p} \mu_k = 1.$$  \hspace{1cm} (i)

(ii) For a finite set $S \subset \mathbb{R}^m$ and a vertex $u$ of $\text{co}(S)$, there exists a supporting hyperplane $H$ of $\text{co}(S)$ at $u$ such that $H \cap \text{co}(S) = \{u\}$.

For a convex set $U$ containing $V$ of (2) and a function $h$ of (1), let the level set $U_\alpha$ for $\alpha \in \mathbb{R}$ be defined by

$$U_\alpha = \{ u \in U | h(u) \geq \alpha \}.$$  \hspace{1cm} (12)

**Lemma 3.2** [1, Chapter 6]. Let $h$ be continuous and quasiconcave.

(i) $U_\alpha$ is a convex set.

(ii) Let $\bar{u} \in \text{bd}(U_\alpha)$. If $h$ is differentiable at $\bar{u}$ and $(\partial h/\partial u_1, ..., \partial h/\partial u_m)|_{u = \bar{u}} \neq 0$, the supporting hyperplane of $U_\alpha$ at $\bar{u}$ is defined by

$$\sum_{j=1}^{m} \frac{\partial h}{\partial u_j} (u_j - \bar{u}_j) = 0.$$  \hspace{1cm} (i)

4. Relationship between $P$ and $P(\lambda)$

We shall discuss in this section three cases, in which $P(\lambda)$ for an appropriate $\lambda$ can solve $P$. The first theorem deals with the case of a finite $V$ (defined in (2)) and states that there exists a parameter $\lambda$ such that any optimal solution of $P(\lambda)$ is also optimal to $P$. The second and third theorems deal with the case in which $V$ is not finite but $h(u)$ is continuous over $U$ and differentiable at any point $u$ in int($U$). Let $x^*$ be an optimal solution of $P$ and let $u^* = (f_1(x^*), ..., f_m(x^*))$. With some additional assumptions on $h$ and $u^*$, the second theorem shows that at least one optimal solution of $P(\lambda)$ for an appropriate $\lambda$ is optimal to $P$. The third theorem strengthens the second theorem, under the additional assumption $u^* \in \text{int}(U)$, and says that any optimal solution of $P(\lambda)$ is optimal to $P$ for an appropriate $\lambda$. 
The minimization of a quasiconcave program

Theorem 4.1. Assume that the set $V$ of (2) is finite. Then there exists a parameter vector $\lambda$ such that any optimal solution of $P(\lambda)$ is optimal to $P$.

Proof. Let

$$V^* = \{(f_1(x), f_2(x), \ldots, f_m(x)) \mid x \text{ is optimal to } P\}. \quad (13)$$

Namely, for any optimal solution $x^*$ of $P$, we have

$$h(u) = z(x^*) \text{ for any } u \in V^*.$$ 

Since $V$ is finite, $\text{co}(V) \subset U$ is a bounded convex polytope, where $U$ is defined in (2). Note that $V^* \subset V \subset \text{co}(V)$ by definition. Assume first that no point in $V^*$ is a vertex of $\text{co}(V)$. Then by Lemma 3.1(i) any $u^* \in V^*$ can be written as a convex combination of vertices $u^1, u^2, \ldots, u^p$ of $\text{co}(V)$:

$$u^* = \sum_{i=1}^{p} \mu_i u^i,$$

where $\mu_i > 0$ for more than one $i$ and $\sum_{i=1}^{p} \mu_i = 1$. Therefore, by the quasiconcavity,

$$h(u^*) = h\left(\sum_{i=1}^{p} \mu_i u^i\right) \geq \min\{h(u^i) \mid i = 1, \ldots, p\} > z(x^*),$$

a contradiction to $u^* \in V^*$. Hence there exists at least one vertex $u^* \in \text{co}(V) \cap V^*$. Then Lemma 3.1(ii) asserts that there exists a supporting hyperplane $H$ of $\text{co}(V)$ at $u^*$ such that $H \cap \text{co}(V) = \{u^*\}$. Now let $\lambda \in \mathbb{R}^m$ define this $H$ (i.e., $c - \lambda$ holds in (9)). This implies that any optimal solution $x^\lambda$ to $P(\lambda)$ (recall that $x^\lambda \in X$ and $(f_1(x^\lambda), f_2(x^\lambda), \ldots, f_m(x^\lambda)) \in V$) satisfies

$$u^* = (f_1(x^\lambda), f_2(x^\lambda), \ldots, f_m(x^\lambda))$$

since otherwise

$$\sum_{j=1}^{m} \lambda_j f_j(x^\lambda) > \sum_{j=1}^{m} \lambda_j u^*_j$$

holds for $x^\lambda \in \text{co}(V)$ by (11) and assumption $H \cap \text{co}(V) = \{u^*\}$, contradicting the optimality of $x^\lambda$. □

Theorem 4.2. Suppose that $h(u)$ is continuous over $U$ and differentiable at any point $u \in \text{int}(U)$. Let $x^*$ be an optimal solution of $P$ and $u^* = (f_1(x^*), f_2(x^*), \ldots, f_m(x^*))$. Note that $u^*$ satisfies either $u^* \in \text{bd}(U)$ or $u^* \in \text{int}(U)$. Assume

$$\text{grad } h(u^*) = (\partial h(u^*)/\partial u_1, \partial h(u^*)/\partial u_2, \ldots, \partial h(u^*)/\partial u_m) \neq 0,$$

in case of $u^* \in \text{int}(U)$. Then there exists a $\lambda$ such that at least one optimal solution of $P(\lambda)$ is optimal to $P$.

Proof. We first consider the case of $u^* \in \text{bd}(U)$. Since $U$ is convex as assumed in
Section 1, there exists a \( \lambda \) such that it defines a supporting hyperplane of \( U \) at \( u^* \) by

\[
\sum_{j=1}^{m} \lambda_j (u_j - u_j^*) = 0.
\]

This means that \( \sum \lambda_j u_j^* \leq \sum \lambda_j u_j \) holds for any \( u \in U \) (see (9) and (11)). Since \( V \subset U \) by (2), \( x^* \) is therefore optimal to \( P(\lambda) \).

Next consider the case of \( u^* \in \text{int}(U) \). We first show that \( u^* \in \text{bd}(U_{z(x^*)}) \). Suppose otherwise. Then there exists a \( d > 0 \) such that

\[
N_d(u^*) \triangleq \{ u \in \mathbb{R}^m \mid ||u - u^*|| < d \} \subset U_{z(x^*)}.
\]

Since \( \text{grad} \, h(u^*) \neq 0 \) and \( h \) is continuous, there exists an \( \alpha > 0 \) such that

\[
u' = u^* - \alpha \text{grad} \, h(u^*) \in N_d(u^*)
\]

and \( h(u') < h(u^*) \). However, any \( u' \in N_d(u^*) \subset U_{z(x^*)} \) must satisfy \( h(u') \geq h(u^*) \) (= \( z(x^*) \)) by the definition of \( U_{z(x^*)} \), a contradiction. Therefore \( u^* \in \text{bd}(U_{z(x^*)}) \).

Now by Lemma 3.2,

\[
\sum_{j=1}^{m} \frac{\partial h(u^*)}{\partial u_j} (u_j - f_j(x^*)) = 0
\]

is a supporting hyperplane of \( U_{z(x^*)} \) at \( u^* \). Thus for \( \lambda = \text{grad} \, h(u^*) \) it is proved in a manner similar to the case of \( u^* \in \text{bd}(U) \) that \( x^* \) is optimal to \( P(\lambda) \). □

**Theorem 4.3.** Assume the conditions of Theorem 4.2, and furthermore assume that \( u^* \in \text{int}(U) \). Then for \( \lambda = \text{grad} \, h(u^*) \) any optimal solution of \( P(\lambda) \) is optimal to \( P \).

**Proof.** By the second half of the proof of Theorem 4.2, it is sufficient to show that, for any \( \tilde{u} = (f_1(\tilde{x}), \ldots, f_m(\tilde{x})) \in V \) satisfying (14) (i.e., \( \tilde{x} \) is optimal to \( P(\lambda) \)), \( \tilde{x} \) is optimal to \( P \), i.e.,

\[
h(\tilde{u}) \triangleq z(\tilde{x}) = h(u^*) = z(x^*)
\]

Since \( \tilde{u} \in V \subset U_{z(x^*)} \) and \( \tilde{u} \notin \text{int}(U_{z(x^*)}) \), it holds that \( \tilde{u} \in \text{bd}(U_{z(x^*)}) \). Suppose \( h(\tilde{u}) > h(u^*) \). By definition, the line segment between \( \tilde{u} \) and \( u^* \) is contained in the supporting hyperplane (14) and \( \text{bd}(U_{z(x^*)}) \). Since \( h \) is continuous and \( u^* \in \text{int}(U) \) by assumption, there exists a point \( u' \) on the line segment such that \( u' \in \text{bd}(U_{z(x^*)}) \cap \text{int}(U) \) and it satisfies \( h(u') > h(u^*) \). Thus for some \( d' > 0 \), \( h(u') > h(u^*) \) holds for all \( u \in N_{d'}(u^*) \triangleq \{ u \in \mathbb{R}^m \mid ||u - u^*|| < d' \} \subset U \). Since \( u' \in \text{bd}(U_{z(x^*)}) \), \( N_{d'}(u^*) \) contains a point \( \tilde{u} \in U_{z(x^*)} \). But \( h(\tilde{u}) > h(u^*) \) contradicts the definition of \( U_{z(x^*)} \), and hence \( h(\tilde{u}) = h(u^*) \). □

Theorem 4.3 has already been observed in [5,12,13,14,15] for some special functions \( h \) with \( m - 2 \). In revising this paper, we were also informed that Sniedovich [32] recently gave the same characterization as Theorem 4.3. Since he assumes the concavity of \( h(u_1, \ldots, u_m) \), our result is more general.
Example 4.1. Let $m = 2$, $n = 3$ and $U = [0, \infty) \times [0, \infty)$. Let

$$h(u_1, u_2) = u_1 + \sqrt{u_2},$$

$$f_1(x) = x_2 + x_3, \quad f_2(x) = 2x_1$$

and

$$X = \{(0, 0, 1, 1), (2, 2, 0), (2, 1, 1), (0, 0, 4)\}.$$ 

Thus

$$h(f_1(x), f_2(x)) = x_2 + x_3 + \sqrt{2x_1}$$

and

$$V = \{(2, 0), (2, 2), (4, 2), (4, 0)\}.$$ 

It is easy to see that $h(u_1, u_2)$ is concave (hence quasiconcave) and $x^* = (0, 1, 1)$ (hence $u^* = (2, 0)$) is an optimal solution of $P$. The convex hull $\text{co}(V)$ of $V$ is illustrated in Fig. 1. A supporting hyperplane $H$ of $\text{co}(V)$ at $u^*$ such that $\text{co}(V) \cap H\{u^*\}$ can be given by any vector $\lambda = (\lambda_1, \lambda_2)$ with $\lambda_1, \lambda_2 > 0$, and $x^*$ is the unique optimal solution of $P(\lambda)$. Thus we can confirm Theorem 4.1 for this example.

Note next that $U_{\min} = U_2$ is given by

$$U_2 = U \cap \{(u_1, u_2) \mid u_1 \geq 0 \text{ and } u_2 \geq (2 - u_1)^2, \text{ or } u_1 \geq 2 \text{ and } u_2 \geq 0\}$$

as illustrated in Fig. 1. Theorem 4.2 also holds for this case, since $h(u)$ is continuous over $U$ and differentiable at any $u \in \text{int}(U)$, and $u^* = (2, 0) \in \text{bd}(U)$. The supporting hyperplane $H'$ of $U$ at $u^*$ is $u_2 = 0$, i.e., given by any $\lambda = (0, \lambda_2)$ with $\lambda_2 > 0$. For this

![Fig. 1. Illustration of the points in $V$ of Example 4.1 (the lightly shaded area denotes $U_{\min} = U_2$).](image-url)
\( x^*(0, 1, 1) \) is an optimal solution of \( P(\lambda) \). But \( x = (0, 0, 4) \), which is not optimal to \( P \), is also optimal to \( P(\lambda) \) since \( (4, 0) \in H' \).

Now consider the same problem with \( X \) replaced by \( X = \{(0, 1, 1)\} \). For this new \( X \), \( V = \{(4, 2), (2, 2), (4, 0)\} \) and the optimal solution of \( P \) is \( x^* = (2, 1, 1) \), i.e., \( u^* = (f_1(x^*), f_2(x^*)) = (2, 2) \) and \( u^* \in \text{int}(U) \). Theorem 4.3 can now be applied and, for \( \lambda = \frac{\partial h(u^*)}{\partial u_1}, \frac{\partial h(u^*)}{\partial u_2} = (1, 1/2\sqrt{2}) \), \( x^* \) is the unique optimal solution of \( P(\lambda) \).

Although Theorems 4.1-4.3 state that \( P(\lambda) \) for an appropriate \( \lambda \) can solve \( P \), such a \( \lambda \) is not known unless \( P \) is solved. A straightforward approach to resolve this dilemma is to solve \( P(\lambda) \) for all \( \lambda \); the one with the minimum \( z(x) \) is an optimal solution of \( P \). As noted in Section 1, this type of approach can sometimes provide polynomial time algorithms. In general, however, the number of solutions generated over the entire range of \( \lambda \) is not polynomially bounded, and it is difficult to develop polynomial time algorithms by this approach.

A notable exception to this observation is the fractional program, i.e., \( m = 2 \) and \( h(f_1(x), f_2(x)) = f_1(x)/f_2(x) \). In this case, an optimal solution \( x^* \) of \( P(\lambda) \) can tell not only whether \( \lambda = \lambda^* \) holds or not (\( \lambda^* \) denotes the \( \lambda \) that solves \( P \)), but also which of \( \lambda^* > \lambda \) and \( \lambda^* < \lambda \) holds if \( \lambda \neq \lambda^* \). Based on this property, Megiddo [20] has shown that polynomial time algorithms exist for a wide class of fractional programs.

5. A class of \( P \) with polynomial approximation scheme

We consider in what follows the class of \( P \) satisfying the following five assumptions.

(A1) \( h(u_1, u_2, \ldots, u_m) \) is nonnegative, continuous and nondecreasing in each \( u_j \). Also \( h \) is differentiable at any \( u \in \text{int}(U) \).

(A2) \( U \) is a hypercube defined by

\[
\{u = (u_1, \ldots, u_m) \mid \text{one of } a_j \leq u_j \leq b_j, a_j < u_j < b_j, a_j < u_j \leq b_j \text{ and } a_j < u_j < b_j \text{ holds for each } j = 1, \ldots, m \}, \tag{15}
\]

where \( a_j \leq b_j \) for all \( j \) and \( a_j \) (resp. \( b_j \)) may be equal to \(-\infty \) (resp. \( \infty \)).

(A3) There exists an optimal solution \( x^* \) of \( P \) such that \( u^* = (f_1(x^*), \ldots, f_m(x^*)) \in \text{int}(U) \) and \( \text{grad } h(u^*) \neq 0 \).

(A4) There exist the following finite positive lower and upper bounds:

\[
D \geq \max \left\{ \left| \frac{\partial h(u)}{\partial u_j} \right| \left| 1 \leq j \leq m, \frac{\partial h(u)}{\partial u_j} > 0 \text{ and } u \in V \right. \right\},
\]

\[
D \leq \min \left\{ \left| \frac{\partial h(u)}{\partial u_j} \right| \left| 1 \leq j \leq m, \frac{\partial h(u)}{\partial u_j} > 0 \text{ and } u \in V \right. \right\}. \tag{16}
\]

(A5) Given \( \lambda^* = \text{grad } h(u^*) = (\partial h(u^*)/\partial u_1, \ldots, \partial h(u^*)/\partial u_m) \), define the correspond-
The minimization of a quasiconcave program

ing problem for every \( k \in \{1, 2, \ldots, m\} \) and \( \alpha = (\alpha_1, \alpha_2, \ldots, \alpha_m) \) satisfying \( \alpha_j \geq 0 \) for all \( j \) and \( \alpha_k = 0 \):

\[
R(\lambda^*, k, \alpha): \quad \text{maximize} \quad h(u)
\]
\[
\text{subject to} \quad u \in U,
\]
\[
\sum_{j=1}^{m} \lambda_j^* (1 + \alpha_j)(u_j - u_j^*) = 0.
\]

Let \( u(\alpha) \) denote an optimal solution of \( R(\lambda^*, k, \alpha) \). Then there exists an increasing function \( \varphi(\| \alpha \|) \) independent of \( \lambda^* \) and \( k \) such that

\[
\varphi(0) = 0,
\]
\[
\varphi(h(u(\alpha)) - h(u^*)) \leq \varphi(\| \alpha \|) h(u^*).
\]

Since (A1) and (A3) guarantee the conditions of Theorems 4.2 and 4.3, \( u^* \) is always on the boundary of \( U_{h(u^*)} = U_{dx^*} \) as shown in the proof of Theorem 4.2. Also it is shown by Theorem 4.3 that for \( \lambda^* = \text{grad} h(u^*) \)
\[
\sum_{j=1}^{m} \frac{\partial h(u^*)}{\partial u_j} (u_j - f_j(x^*)) = 0
\]

is a supporting hyperplane of \( U_{h(u^*)} \) at \( u^* \). For any \( u \in U \) satisfying (19),

\[
h(u) \leq h(u^*)
\]

holds since (19) is a supporting hyperplane of \( U_{h(u^*)} \). (20) implies that \( u^* \) is an optimal solution of \( R(\lambda^*, k, 0) \) (defined in (A5)). (A5) states that \( h(u(\alpha)) \) is stable with respect to a perturbation \( \alpha \) around \( \lambda^* \). Using assumptions (A1), (A2) and (A5), it will be shown later that

\[
h(u(\alpha)) \geq h(f_1(x^1), \ldots, f_m(x^m))
\]

for any optimal solution \( x^i \) of \( P(\lambda^*) \), where \( \lambda^* = (\lambda_1^*(1 + \alpha_1), \lambda_2^*(1 + \alpha_2), \ldots, \lambda_m^*(1 + \alpha_m)) \). With (18), this implies that \( z(x^i) = h(f_1(x^i), \ldots, f_m(x^i)) \) is also stable with respect to a perturbation around \( \lambda^* \).

The following observation is useful in the subsequent sections to prove the validity of the proposed approximation scheme: The second constraint of (17), \( \sum_{j=1}^{m} \lambda_j^* (1 + \alpha_j)(u_j - u_j^*) = 0 \), is satisfied even if \( \lambda^* \) is replaced by \( a \cdot \lambda^* \) for any constant \( a > 0 \). In other words, \( R(a\lambda^*, k, \alpha) \) is equivalent to \( R(\lambda^*, k, \alpha) \). Since \( \varphi(\| \alpha \|) \) is independent of \( \lambda^* \), (18) remains valid for such \( a \lambda^* \).

Assumptions (A3), (A4) and (A5) are rather messy and not easy to prove in general. (A3) is obvious if \( U \) is an open set and \( (\partial h/\partial u_1, \ldots, \partial h/\partial u_m) \neq 0 \) holds for all \( u \in U \). It will be shown in Section 9 that the condition \( u^* \in \text{int}(U) \) in (A3) can be omitted if \( m = 2 \). Therefore (A3) holds for stochastic programming problems \( S_1 \) and \( S_2 \) given by (4) and (5), since \( m = 2 \) and \( U = [0, \infty) \times [0, \infty) \) for \( S_1 \) and \( U = [0, d) \times [0, \infty) \) for \( S_2 \) can be assumed. If \( V \) is finite, (A4) is trivially satisfied since \( h \) is
nondecreasing and differentiable by (A1), and \((\partial h(u^*)/\partial u_1, \ldots, \partial h(u^*)/\partial u_m) \neq 0\).

As will be shown in Section 10, (A5) holds for typical functions such as

\[ h_1(u_1, u_2) = u_1 + \beta \sqrt{u_2} \quad (\beta > 0) \quad \text{and} \quad h_2(u_1, u_2) = \sqrt{u_2}/(d - u_1) \quad (d > u_1), \]

objective functions of \(S_1\) and \(S_2\) respectively, and some others.

6. Partition of parametric space to guarantee \(\varepsilon\)-approximate solution

Let \(\varepsilon > 0\) be a given constant, and \(z^* > 0\) denote the exact optimum value of \(P\). We shall develop in this and next sections an approximation scheme that always guarantees an approximate solution with its objective value \(\tilde{z}\) satisfying

\[ \frac{\tilde{z} - z^*}{z^*} \leq \varepsilon. \]

Our scheme first partitions the \(\lambda\)-space into a polynomial number of subregions, and choose a polynomial number of \(\lambda\)'s from each subregion, for which optimal solutions \(x^*\) of \(P(\lambda)\) are computed. Among the solutions generated in this way, the one minimizing \(z(x) (= h(f_1(x), \ldots, f_m(x))\) is then selected as an \(\varepsilon\)-approximate solution.

The following lemma is crucial for this purpose.

**Lemma 6.1.** Assume (A1)-(A5) of Section 5. There exists a \(\lambda^* = (\lambda_1^*, \ldots, \lambda_m^*)\) such that any optimal solution of \(P(\lambda^*)\) is optimal to \(P\) and

\[ \lambda_k^* = 1 \quad \text{for some } k, \]

\[ D/D \leq \lambda_j^* \leq D/D \quad \text{for } j \neq k \text{ such that } \frac{\partial h(u^*)}{\partial u_j} > 0. \]

**Proof.** By (A1) and (A3), Theorem 4.3 holds. Then

\[ \lambda = \text{grad } h(u^*) = \left( \frac{\partial h(u^*)}{\partial u_1}, \ldots, \frac{\partial h(u^*)}{\partial u_m} \right) \]

satisfies the lemma assertion except possibly for (21). Then define \(\lambda^* = (\lambda_1^*, \ldots, \lambda_m^*)\) by \(\lambda_k^* = 1\) for a \(k\) with \(\lambda_k > 0\), and \(\lambda_j^* = \lambda_j/\lambda_k\) for \(j \neq k\). This \(\lambda^*\) clearly satisfies (21) by assumption (A4), i.e., (16). □

Lemma 6.1 justifies the search of \(\lambda\) in the \(m-1\) dimensional hypercube

\[ HC^k = \{ \lambda = (\lambda_1, \ldots, \lambda_m) | 0 \leq \lambda_j \leq D/D, j = 1, \ldots, m, \text{ and } \lambda_k = 1 \} \]

for some \(k\). Since such \(k\) is not known in advance, the procedure is applied to all \(k\). The crux of our scheme is how to partition each \(HC^k\) into a polynomial number of subregions. We explain the partition scheme only for \(k = m\), since other cases are similar.
Denote problem $P(\lambda)$ associated with a $\lambda \in HC^m$ by

$$P^m(\lambda): \minimize_{x \in X} \left\{ \sum_{j=1}^{m-1} \lambda_j f_j(x) + f_m(x) \right\}. \tag{23}$$

With

$$A^U = \frac{D}{D}, \quad A^I = 1/A^U, \quad K = \left\lfloor \log(A^U/A^I) / \log(1 + \delta) \right\rfloor = \left\lfloor \log(D/D) / \log(1 + \delta) \right\rfloor, \tag{24}$$

define

$$A^{(0)} = 0, \quad A^{(1)} = A^L, \ldots, A^{(k)} = A^I (1 + \delta)^{k-1}, \ldots, \quad A^{(K)} = A^I (1 + \delta)^{K-1}, \quad A^{(K+1)} = A^U, \tag{25}$$

where $\delta$ is a positive constant determined by function $\varrho$ of (18) as follows:

$$\varrho(\sqrt{m-1} \delta) = \varepsilon. \tag{26}$$

We denote this $\delta$ by $\delta(\varepsilon, m)$ if it is necessary to specify $\varepsilon$ and $m$. These $A^{(k)}$ partition $EC^m$ into $(K + 1)^{m-1}$ meshes:

$$MESH(k_1, k_2, \ldots, k_{m-1}) = \{ \lambda = (\lambda_1, \ldots, \lambda_m) \mid A^{(k_j)} \leq \lambda_j \leq A^{(k_{j+1})}, \quad j = 1, 2, \ldots, m-1, \text{ and } \lambda_m = 1 \}, \tag{27}$$

$$0 \leq k_j \leq K, \quad j = 1, 2, \ldots, m-1.$$

The next lemma shows that the vertices of these meshes provide an $\varepsilon$-approximate solution of $P$.

**Lemma 6.2.** Assume (A1)-(A5), and assume that $u^* = (f_1(x^*), \ldots, f_m(x^*))$ for an optimal solution $x^*$ of $P$ is in $\text{int}(U)$ and satisfies $\partial h(u^*) / \partial u_m > 0$. Assume that

$$\lambda^* = \left( \frac{\partial h(u^*)}{\partial u_1}, \ldots, \frac{\partial h(u^*)}{\partial u_{m-1}}, \frac{\partial h(u^*)}{\partial u_m}, 1 \right)$$

belongs to $MESH(k_1, k_2, \ldots, k_{m-1})$. Then one of its vertices $\bar{\lambda}$ satisfies

$$\frac{z(x^*) - z(x^*)}{z(x^*)} \leq \varepsilon, \tag{28}$$

where $x^*$ is an optimal solution of $P^m(\bar{\lambda})$.

**Proof.** First recall that any optimal solution of $P^m(\lambda^*)$ is optimal to $P$ by Theorem 4.3, and that $u^* \in \text{bd}(U_{z(x^*)})$ as shown in the proof of Theorem 4.2. Then

$$\sum_{i=1}^{m-1} \lambda^*_i (u_j - f_j(x^*)) + u_m - f_m(x^*) = 0$$

is the supporting hyperplane of $U_{z(x^*)} (= U_{h(u^*)})$ at $u^*$ by Lemma 3.2(ii). With $\lambda^* \in MESH(k_1, k_2, \ldots, k_{m-1})$, define $\bar{\lambda} = (\lambda_1, \ldots, \lambda_{m-1}, 1)$ by
\[ \bar{\lambda} = A^1(1 + \delta)^j \quad \text{if} \quad \lambda_j^* > 0, \]
\[ \bar{\lambda}_j = 0 \quad \text{if} \quad \lambda_j^* = 0. \]

We shall show that this \( \bar{\lambda} \) satisfies (28).

Note that \( \lambda^* \in \text{MESH}(k_1, k_2, \ldots, k_{m-1}) \) implies
\[ \lambda^*_j \leq \bar{\lambda}_j \leq \lambda^*_j (1 + \delta), \quad j = 1, \ldots, m - 1. \] (29)

Consider assumption (A5) with the \( \alpha \) such that \( \bar{\lambda}_j = \lambda^*_j (1 + \alpha_j) \) for all \( j \). By (29),

\[ 0 \leq \alpha_j \leq \delta \quad \text{for} \quad j = 1, 2, \ldots, m - 1 \quad \text{and} \quad \alpha_m = 0 \]

hold. Then an optimal solution \( u(\alpha) \) of \( R(\lambda^*, m, \alpha) \) satisfies
\[ h(u(\alpha)) - h(u^*) \leq \alpha(1 + \alpha) h(u^*) \leq \alpha h(\sqrt{m-1}) h(u^*). \] (30)

Therefore, it is sufficient to prove that \( z(x^*) \leq h(u(\alpha)) \) since \( h(\sqrt{m-1}) = \varepsilon \) by (26). This \( u(\alpha) \) satisfies (see (17))
\[ \sum_{j=1}^{m-1} \bar{\lambda}_j (u_j(\alpha) - f_j(x^*)) + u_m(\alpha) - f_m(x^*) = 0. \] (31)

Since \( x^* \) is an optimal solution of \( P(\lambda^*) \),
\[ \sum_{j=1}^{m-1} \bar{\lambda}_j f_j(x^*) + f_m(x^*) \leq \sum_{j=1}^{m-1} \bar{\lambda}_j f_j(x^*) + f_m(x^*) \] (32)

holds. (31) and (32) together imply
\[ \sum_{j=1}^{m-1} \bar{\lambda}_j f_j(x^*) + f_m(x^*) \leq \sum_{j=1}^{m-1} \bar{\lambda}_j u_j(\alpha) + u_m(\alpha). \] (33)

We shall show that there exists a vector \( (\bar{\beta}_1, \bar{\beta}_2, \ldots, \bar{\beta}_m) \geq 0 \) such that
\[ \sum_{j=1}^{m-1} \bar{\lambda}_j (f_j(x^*) + \bar{\beta}_j) + f_m(x^*) + \bar{\beta}_m = \sum_{j=1}^{m-1} \bar{\lambda}_j u_j(\alpha) + u_m(\alpha), \] (34)

and
\[ (f_1(x^*) + \bar{\beta}_1, \ldots, f_m(x^*) + \bar{\beta}_m) \in U. \]

Denote
\[ G(\beta_1, \ldots, \beta_m) \triangleq \sum_{j=1}^{m-1} \bar{\lambda}_j (f_j(x^*) + \beta_j) + f_m(x^*) + \beta_m. \]

Let \( \bar{\beta}_j \geq 0 \) satisfy
\[ f_j(x^*) + \bar{\beta}_j = \max \{ u_j(\alpha), f_j(x^*) \}, \quad j = 1, 2, \ldots, m, \]
\[ (f_1(x^*) + \bar{\beta}_1, f_2(x^*) + \bar{\beta}_2, \ldots, f_m(x^*) + \bar{\beta}_m) \in U. \] (35)

By \( u = (f_1(x^*), \ldots, f_m(x^*)) \in U, \ u(\alpha) \in U \) and (A2), such \( \beta_j \) always exist. Then
\[ G(0, \ldots, 0) = \sum_{j=1}^{m-1} \bar{\lambda}_j f_j(x^*) + f_m(x^*) \]
The minimization of a quasiconcave program

\[ \sum_{j=1}^{m-1} \lambda_j u_j(\alpha) + u_m(\alpha) \quad (\text{by (33)}) \]

\[ \leq G(\beta_1, \ldots, \beta_m), \quad (\text{by (35)}). \]

Since \( G(\beta_1, \ldots, \beta_m) \) is continuous in each \( \beta_j \), some \((\beta_1, \ldots, \beta_m)\) with

\[ 0 \leq \beta_j \leq \bar{\beta}_j \quad (36) \]
satisfies (34). Then (35) and (36) imply

\[ (f_1(x^\delta) + \beta_1, f_2(x^\delta) + \beta_2, \ldots, f_m(x^\delta) + \beta_m) \in U. \]

Since \( (f_1(x^\delta) + \beta_1, \ldots, f_m(x^\delta) + \beta_m) \) is feasible to \( R(\lambda^*, m, \alpha) \) by (31) and (34), we have from the optimality of \( u(\alpha) \) that

\[ h(f_1(x^\delta) + \beta_1, \ldots, f_m(x^\delta) + \beta_m) \leq h(u(\alpha)). \quad (37) \]

Consequently

\[ z(x^\delta) = h(f_1(x^\delta), \ldots, f_m(x^\delta)) \]

\[ \leq h(f_1(x^\delta) + \beta_1, \ldots, f_m(x^\delta) + \beta_m) \quad (\text{since } h \text{ is nondecreasing}) \]

\[ \leq h(u(\alpha)) \quad (\text{by (37)}). \]

This completes the proof. \( \square \)

7. An approximation scheme for \( P \)

Based on the results given in the previous sections, an approximation scheme can now be described.

Procedure APPROX

**Input.** Problem \( P \), a given constant \( \varepsilon > 0 \), and upper and lower bounds \( \bar{D} \) and \( D \).

**Output.** An \( \varepsilon \)-approximate solution of \( P \).

**Step 1.** Compute \( A^U, A^L, K, A^{(0)}, A^{(1)}, \ldots, A^{(K+1)} \) and \( \delta \) by (24), (25), and (26).

**Step 2.** For each \( k = 1, 2, \ldots, m \), partition hypercube \( HC^k \) of (22) into \( (K+1)^{m-1} \) meshes by (27), obtain all the vertices of these meshes, and compute \( x^*(k) \) by

\[ z(x^*(k)) = \min \{ z(x^\lambda) \mid \lambda \text{ is a vertex of a mesh in } HC^k \text{ and } x^\lambda \text{ is an optimal solution of } P^k(\lambda) \}. \]

**Step 3.** An \( \varepsilon \)-approximate solution \( \bar{x} \) of \( P \) is then obtained by

\[ z(\bar{x}) = \min_{1 \leq k \leq m} z(x^*(k)). \quad \square \]

**Theorem 7.1.** Assume (A1)–(A5). Then procedure APPROX correctly computes an \( \varepsilon \)-approximate solution of \( P \) in
time, where \( p(n) \) is the time required to compute an optimal solution \( x^* \) of \( P^k(\lambda) \).

**Proof.** First we prove the correctness. By (A3), there exists an optimal solution \( x^* \) of \( P \) such that \( u^* = (f_1(x^*), \ldots, f_m(x^*)) \in \text{int}(U) \) and \( \lambda = \text{grad} h(u^*) \neq 0 \). Together with (A1), Theorem 4.3 then tells that any optimal solution of \( P(\lambda) \) is optimal to \( P \). Since \( h \) is nondecreasing and \( \lambda \neq 0 \), some \( k \) satisfies \( \partial h(u^*)/\partial u_k > 0 \). Assume without loss of generality that \( k = m \). Then by Lemma 6.2 there exists a parameter vector \( \lambda^* \) such that \( \lambda^* \) is a vertex of a mesh in \( HC^m \) and any optimal solution of \( P(\lambda^*) \) is an \( \epsilon \)-approximate solution of \( P \). Since this \( \lambda^* \) is computed in Step 2, the correctness immediately follows.

Now we analyze the running time. The number of different vertices generated for each \( k \) is \((K+2)^{m-1}\). The total number of vertices generated in Step 2 is, therefore,

\[
m(K+2)^{m-1} = O(m2\log(\bar{D}/D))^{m-1}/\log^{m-1}(1 + \delta(\epsilon, m)))
\]

by (24). Since computing \( x^\lambda \) for each vertex \( \lambda \) requires \( O(p(n)) \) time, the total time is given by (38).

**Corollary 7.1.**

(i) If both \( p(n) \) and \( \log(D/D) \) are polynomial in the input size of a problem instance \( P \), procedure APPROX is a polynomial time approximation scheme.

(ii) In addition, if \( \log^{-1}(1 + \delta(\epsilon, m)) \) is a polynomial in \( 1/\epsilon \), procedure APPROX is a fully polynomial time approximation scheme.

**Proof.** Since \( m \) is assumed to be a constant, the corollary follows from definitions.

8. Modification of approximation scheme APPROX

In this section we assume that each \( P^k(\lambda) \) has a polynomial time approximation scheme, instead of a polynomial time exact algorithm. Even in this case, the following modification of APPROX yields a polynomial time approximation scheme:

1. In Step 1, determine \( \delta, K \) and \( A^{(k)} \) of (26), (24) and (25) by using the following \( \epsilon' \) in place of \( \epsilon \).

\[
\epsilon' = \epsilon/(\sqrt{1+\epsilon} + 1).
\]

2. In Step 2, for each \( \lambda \) chosen, compute an \( \epsilon' \)-approximate solution of \( P^k(\lambda) \) (instead of the exact optimal solution) and denote it \( x^\lambda \).

**Theorem 8.1.** (i) If each \( P^k(\lambda) \) has a polynomial time approximation scheme, the above modification of APPROX is a polynomial time approximation scheme for \( P \), provided that \( \log(D/D) \) is polynomial in the input size of \( P \).
(ii) In addition, if each $P^k(\lambda)$ has a fully polynomial time approximation scheme, and $\log^{-1}(1 + \delta(\varepsilon', m))$ is polynomial in $1/\varepsilon'$, it becomes a fully polynomial time approximation scheme for $P$.

**Proof.** Let $\bar{x}$ be an $\varepsilon'$-approximate solution of $P$ obtained by the original APPROX with $\varepsilon'>0$, i.e.,

$$z(\bar{x}) - z(x^*) \leq \varepsilon' \cdot z(x^*). \quad (40)$$

Assume that $\bar{x}$ is the exact optimal solution of $P^k(\lambda)$ (i.e., $\bar{x} = x^\bar{x}$), and let $x'$ be an $\varepsilon'$-approximate solution of $P^k(\lambda)$ obtained by the approximation scheme with $\varepsilon'$. Namely

$$z(x') - z(x^*) \leq \varepsilon' \cdot z(x^*). \quad (41)$$

Therefore

$$z(x') - z(x^*) \leq (1 + \varepsilon')z(x^\bar{x}) - z(x^*) \quad (by \ 41))$$

$$\leq (1 + \varepsilon')^2z(x^*) - z(x^*) \quad (by \ 40))$$

$$\leq (2\varepsilon' + (\varepsilon')^2)z(x^*) = \varepsilon z(x^*) \quad (by \ (39)).$$

This shows that the modified APPROX correctly provides an $\varepsilon$-approximate solution of $P$.

Let $p(n, \varepsilon')$ denote the running time of the approximation scheme for $P^k(\lambda)$ with $\varepsilon'>0$. The running time of the modified APPROX is

$$O(p(n, \varepsilon')m(2 \log(D/D'))^{m-1}/\log^{m-1}(1 + \delta(\varepsilon', m))),$$

which is polynomial under the conditions of (i). To prove (ii), assume that $p(n, \varepsilon')$ is polynomial in $n$ and $1/\varepsilon'$. Since $\varepsilon' = \varepsilon/(\sqrt{1 + \varepsilon} + 1)$, noting $\sqrt{\varepsilon} - 1 \leq \sqrt{1 + \varepsilon} \leq 3\sqrt{\varepsilon} - 1$ for $\varepsilon \geq 1$ and $1 < \sqrt{1 + \varepsilon} < \sqrt{2}$ for $0 < \varepsilon < 1$, we have

$$\sqrt{\varepsilon}/3 \leq \varepsilon' \leq \sqrt{\varepsilon} \quad for \ \varepsilon \geq 1,$$

$$\varepsilon/(\sqrt{2} + 1) \leq \varepsilon' \leq \varepsilon/2 \quad for \ 0 \leq \varepsilon \leq 1.$$  \quad (42)

Thus $p(n, \varepsilon')$ and $\log^{-1}(1 + \delta(\varepsilon', m))$ are polynomial in $1/\varepsilon$. This proves (ii).  \quad \square

9. Approximation scheme specialized for $m = 2$

This section deals with a special case of $m = 2$, as it is important in practical applications (e.g., $S_1$ and $S_2$ of (4) and (5)). We show that assumptions (A3), (A5) in Section 5 can be slightly relaxed, and accordingly procedure APPROX is modified. This modification enables us to handle a wider class of $S_1$ and $S_2$, as we shall see in Section 10.

First relax assumption (A5) as follows:
(A5') This is the same as (A5) except that \( m = 2 \) and \( k \) is fixed to \( k = 2 \).

**Lemma 9.1.** Let \( m = 2 \), and assume conditions (A1), ..., (A4) and the above (A5'). Let \( x^* \) be an optimal solution of \( P \) such that \( u^* = (f_1(x^*), f_2(x^*)) \in \text{int}(U) \).

(i) If \( \partial h(u^*)/\partial u_2 > 0 \), take \( \text{MESH}(k_1) \) that contains

\[
\lambda^* = \left( \frac{\partial h(u^*)}{\partial u_1}, \frac{\partial h(u^*)}{\partial u_2}, 1 \right),
\]

i.e., \( A^{(k_1)} \leq \lambda^*_1 \leq A^{(k_1+1)} \) and \( \lambda^*_2 = 1 \). Also define \( \tilde{\lambda} \) by

\[
\tilde{\lambda}_1 = A^{(k_1)} \text{ or } A^{(k_1+1)}, \quad \tilde{\lambda}_2 = 1.
\]

Then any optimal solution \( x^{\tilde{\lambda}} \) of \( P^2(\tilde{\lambda}) \) for one of the above two \( \tilde{\lambda} \)'s is an \( \varepsilon \)-approximate solution of \( P \), i.e.,

\[
\frac{z(x^{\tilde{\lambda}}) - z(x^*)}{z(x^*)} \leq \varepsilon.
\]

(ii) If \( \partial h(u^*)/\partial u_2 = 0 \), \( x^* \) is obtained as an arbitrary optimal solution of the following problem.

\[
P': \quad \text{minimize } f_1(x). \quad \text{For } x \in X.
\]

**Proof.** (i) is a restatement of Lemma 6.2. (ii) is immediate from Theorem 4.3. □

Next we consider the case in which \( u^* = (f_1(x^*), f_2(x^*)) \in \text{bd}(U) \). Modify assumption (A3) as follows:

(A3') There exists an optimal solution \( x^* \) of \( P \) such that \( u^* = (f_1(x^*), ..., f_m(x^*)) \) and \( (\partial h(u^*)/\partial u_1, ..., \partial h(u^*)/\partial u_m) \neq 0 \).

**Lemma 9.2.** Let \( m = 2 \), and assume (A1), (A2), (A4) in Section 5 and (A3'), (A5'). Let \( u^* = (f_1(x^*), f_2(x^*)) \) for an optimal solution \( x^* \) of \( P \), and assume that \( f_j(x), j = 1, 2, \) are nonnegative integer-valued functions bounded above by \( M \):

\[
\max\{f_j(x) | x \in X, j = 1, 2\} \leq M.
\]

Then, if \( u^* \in \text{bd}(U) \), \( x^* \) is obtained as an arbitrary optimal solution of the following parametric problem \( P^2(\lambda_1, 1) \), where \( \lambda_1 \) is a constant satisfying either \( \lambda_1 > M \) or \( 0 < \lambda_1 < 1/M \).

\[
P^2(\lambda_1, 1): \quad \text{minimize } \left\{ \lambda_1 f_1(x) + f_2(x) \right\}. \quad \text{For } x \in X.
\]

**Proof.** By the integrality of \( f_j(x) \) and (45), \( V \) of (2) is a finite set. Thus, \( \text{co}(V) \) is a convex polygon and \( u^* \) is one of its vertices as shown in Fig. 2. Since \( h \) is
The minimization of a quasiconcave program

nondecreasing we can assume that \( u^* \) is either on the left vertical boundary or on the lower horizontal boundary of \( U \). For simplicity, let \( u^* \) be on the left vertical boundary (the other case can be similarly treated). As easily seen from Fig. 2, the supporting hyperplane \( H \) of \( \text{co}(V) \) at \( u^* \) such that \( H \cap \text{co}(V) = \{u^*\} \) is given by

\[
\lambda_1(u_1-u_1^*)+u_2-u_2^*=0
\]

for a sufficiently large \( \lambda_1 > 0 \). Therefore \( x^* \) is an arbitrary optimal solution of \( P^2(A_i, 1) \) for such \( \lambda_1 \).

To derive a bound on \( \lambda_1 \), consider two \( A_i \) and \( A_i' \) with \( A_i' > A_i > M \), and let \( x^{i_1} \) and \( x^{i_1'} \) denote optimal solutions of \( P^2(A_i, 1) \) and \( P^2(A_i', 1) \) respectively. It is known in the theory of parametric programming (Cartensen [2]) that

\[
f_1(x^{i_1}) \geq f_1(x^{i_1'}) \quad (46)
\]

We shall show that \( f_1(x^{i_1}) = f_1(x^{i_1'}) \) and \( f_2(x^{i_1}) = f_2(x^{i_1'}) \).

**Case 1:** \( x^{i_1} \) is optimal to \( P^2(\lambda_1, 1) \). Then

\[
\lambda_1 f_1(x^{i_1}) + f_2(x^{i_1}) = \lambda_1 f_1(x^{i_1'}) + f_2(x^{i_1'}) \quad (47)
\]

holds. If \( f_1(x^{i_1}) \neq f_1(x^{i_1'}) \), we have \( f_1(x^{i_1}) > f_1(x^{i_1'}) \) by (46). Then

\[
\lambda_1 = \frac{f_2(x^{i_1'}) - f_2(x^{i_1})}{f_1(x^{i_1'}) - f_1(x^{i_1})} \leq M
\]
since $f_j(x)$ are nonnegative, integer-valued and $f_j(x) \leq M$ by (45). This is a contradiction, and $f_1(x^{(i)}) = f'_1(x^{(i)})$ (and hence $f_2(x^{(i)}) = f'_2(x^{(i)})$ by (47)) must hold.

Case 2: $x^{(i)}$ is optimal to $P^2(\lambda^*_1, 1)$. This is similar to Case 1.

Case 3: $x^{(i)}$ (resp. $x^{(i)}$) is not optimal to $P^2(\lambda^*_1, 1)$ (resp. $P^2(\lambda^*_1, 1)$). Then

$$\lambda f_1(x^{(i)}) + f_2(x^{(i)}) < \lambda f_1(x^{(i)}) + f_2(x^{(i)})$$

(by $\lambda^*_1 > \lambda$)

$$\leq \lambda f_1(x^{(i)}) + f_2(x^{(i)})$$

(48)

Therefore, there exists a $\lambda$ with $\lambda < \lambda^*_1 < \lambda^*_1$ such that

$$\lambda f_1(x^{(i)}) + f_2(x^{(i)}) = \lambda f_1(x^{(i)}) + f_2(x^{(i)})$$

(49)

Then in a manner similar to Case 1, it follows that

$$\lambda \leq M,$$

a contradiction. This shows that, if $u^*$ is on the left vertical boundary, an arbitrary optimal solution of $P^2(\lambda, 1)$ for $\lambda > M$ is an optimal solution $x^*$ of $P$.

In case $u^*$ is on the lower horizontal boundary, it is similarly shown that an arbitrary optimal solution of $P^2(\lambda, 1)$ for $\lambda < 1/M$ is an optimal solution $x^*$ of $P$. □

Now let procedure APPROX2 be equal to the original APPROX with Steps 2 and 3 being modified as follows:

Step 2. For $k = 2$, partition hypercube $HC^2$ into $(K + 1)^m - 1$ meshes by (27), obtain all the vertices of these meshes, and compute

$$z(x^*(2)) = \min \{z(x^{(i)}) \mid \lambda \text{ is a vertex of a mesh of } HC^2 \text{ and } x^{(i)} \text{ is an optimal solution of } P^2(\lambda)\}.$$

Also compute an optimal solution $x^*(1)$ of problem of $P'$ of (44).

Step 3. Compute optimal solutions $x^{(i)}$ (resp. $x^{(i')}$) of $P^2(\lambda^*_1, 1)$ (resp. $P^2(\lambda^*_1, 1)$) for a $\lambda^*_1$ with $0 < \lambda^*_1 < 1/M$ (resp. $\lambda^*_1$ with $\lambda^*_1 > M$). Then an $\varepsilon$-approximate solution $\bar{x}$ of $P$ is given by

$$z(\bar{x}) = \min \{z(x^*(1)), z(x^*(2)), z(x^{(i)}), z(x^{(i')})\}.$$ □

Theorem 9.1. Let $m = 2$. Assume (A1), (A2), (A4) in Section 5 and (A3'), (A5'). Furthermore, let $f_j(x)$, $j = 1, 2$, be nonnegative, integer valued, and have an $M$ satisfying (45). Then procedure APPROX2 correctly computes an $\varepsilon$-approximate solution of $P$ in

$$O(p(n)\log(D/D)/\log(1 + \delta(\varepsilon, 2)))$$

time.

Proof. Obvious from Theorem 7.1, and Lemmas 9.1 and 9.2. □
10. Assumptions (A1) and (A5) for some functions

We shall show in this section that the following functions \( h_1, h_2, h_3 \) and \( h_4 \) satisfy assumptions (A1) and (A5) (or (A5')). It will be then shown in the next section that problems with these objective functions can have fully polynomial approximation schemes under some additional conditions. In particular, as \( h_1 \) and \( h_2 \) are objective functions of stochastic programming problems \( S_1 \) and \( S_2 \) defined by (4) and (5), respectively, such problems have fully polynomial time approximation schemes in many cases practically important.

\[
\begin{align*}
    h_1(u_1, u_2) &= u_1 + \beta \sqrt{u_2} \\
    h_2(u_1, u_2) &= \sqrt{u_2} / (d - u_1) \\
    h_3(u_1, \ldots, u_m) &= u_1u_2 \cdots u_m \\
    h_4(u_1, \ldots, u_m) &= N - \sum_{j=1}^{m} (q_j - u_j)^2
\end{align*}
\]

for \( U = [0, \infty) \times [0, \infty) \) and \( \beta > 0 \). (50)

For \( U = [0, d) \times [0, \infty) \). (51)

For \( U = (0, \infty) \times \cdots \times (0, \infty) \). (52)

For \( U = [0, q_1] \times \cdots \times [0, q_m] \). (53)

In the last case, \( q_j > 0, j = 1, \ldots, m \) and \( N > 2 \sum_{j=1}^{m} q_j \) are assumed.

As it is obvious that all of \( h_1, h_2, h_3 \) and \( h_4 \) satisfy (A1), only assumption (A5) or (A5') is proved. As discussed in Section 9, \( h_1 \) and \( h_2 \) have \( m = 2 \) and it is sufficient to check (A5') instead of (A5).

(i) \( h_1(u_1, u_2) \): Since \( \partial h_1 / \partial u_1 = 1 \) and \( \partial h_1 / \partial u_2 = \beta / 2 \sqrt{u_2} \), the supporting hyperplane of \( U_{h_1(u^*)} \) at \( u^* \) is

\[
(u_1 - u^*) + \frac{\beta}{2 \sqrt{u_2}} (u_2 - u^*_2) = 0,
\]

with \( \lambda^* = (1, \beta / 2 \sqrt{u_2^*}) \). For \( \alpha = (\alpha_1, 0) \) with \( \alpha_1 \geq 0 \) (as discussed in (A5')), \( R(\lambda^*, 2, \alpha) \) becomes

\[
R(\lambda^*, 2, \alpha): \max h_1(u_1, u_2) = u_1 + \beta \sqrt{u_2} \]

subject to \( (u_1, u_2) \in U \),

\[
(1 + \alpha_1)(u_1 - u_1^*) + \frac{\beta}{2 \sqrt{u_2^*}} (u_2 - u_2^*) = 0.
\]

Instead of an optimal solution \( u(\alpha) \) of \( R(\lambda^*, 2, \alpha) \), we shall obtain an optimal solution \( \hat{u}(\alpha) \) of \( R(\lambda^*, 2, \alpha) \) with condition \( (u_1, u_2) \in U \) dropped. By the method of Lagrangean multipliers, \( \hat{u}(\alpha) \) is given by

\[
\hat{u}_1(\alpha) = u_1^* + \frac{\beta \sqrt{u_2^*}}{2} \left( \frac{1}{1 + \alpha_1} - 1 - \alpha_1 \right),
\]

\[
\hat{u}_2(\alpha) = u_2^*(1 + \alpha_1)^2.
\]

Fig. 3 illustrates the points \( u^* \) and \( \hat{u}(\alpha) \) of (54). Since \( \hat{u}(\alpha) \) is an optimal solution of a relaxation of \( R(\lambda^*, 2, \alpha) \),
\[ h(u_1, u_2) = h(u^*_1, u^*_2) \]
\[ (u_1 - u^*_1) + \frac{\beta}{2\sqrt{u^*_2}} (u_2 - u^*_2) = 0 \]
\[ (1 + \alpha_1)(u_1 - u^*_1) + \frac{\beta}{2\sqrt{u^*_2}} (u_2 - u^*_2) = 0 \quad (\alpha_1 > 0) \]

Fig. 3. Illustration of the points \( u^* \) and \( \hat{u}(\alpha) \) of (54) for \( h_1(u_1, u_2) = u_1 + \beta\sqrt{u_2} \) (in this figure \( \hat{u}(\alpha) = u(\alpha) \) holds).

\[
h_1(u(\alpha)) \leq h_1(\hat{u}(\alpha)) \tag{55}\]

holds. Thus, by \( \alpha_1 \geq 0 \) and \( u^*_1 \leq 0 \),
\[
h_1(u(\alpha)) \leq h_1(\hat{u}(\alpha)) = u^*_1 + \frac{\beta \sqrt{u^*_2}}{2} \left( \frac{1}{1 + \alpha_1} + 1 + \alpha_1 \right) \]
\[
\leq u^*_1 + \beta \sqrt{u^*_2} (1 + \alpha_1) = h_1(u^*) + \alpha_1 \beta \sqrt{u^*_2}. \tag{56}\]

Therefore
\[
h_1(u(\alpha)) - h_1(u^*) \leq \alpha_1 \beta \sqrt{u^*_2} \leq \alpha_1 (u^*_1 + \beta \sqrt{u^*_2}) \quad \text{(by \( u^*_1 \leq 0 \))}
\]
\[
= \alpha_1 \cdot h_1(u^*) = \| \alpha \| \cdot h_1(u^*) \quad \text{(by \( \alpha_2 = 0 \)).} \]

Consequently the \( g \) of (18) can be given by
\[
g_1(\| \alpha \|) = \| \alpha \|. \tag{57}\]

(ii) \( h_2(u_1, u_2) \): The supporting hyperplane of \( U_{h_2(u^*)} \) at \( u^* \) is
\[
\frac{\sqrt{u_2^*}}{(d-u_1^*)^2} (u_1 - u_1^*) + \frac{1}{2\sqrt{u_2^*}(d-u_1^*)} (u_2 - u_2^*) = 0,
\]

since \( \frac{\partial h_2}{\partial u_1} = \sqrt{u_2}/(d-u_1)^2 \) and \( \frac{\partial h_2}{\partial u_2} = 1/2\sqrt{u_2}(d-u_1) \). For \( \alpha = (\alpha_1, 0) \) with \( \alpha_1 \geq 0 \) (assumed in (A5')), \( R(\lambda^*, 2, \alpha) \) now becomes

\[
R(\lambda^*, 2, \alpha): \text{ maximize } \frac{\sqrt{u_2}}{d-u_1} h_2(u_1, u_2) \text{ subject to } (u_1, u_2) \in U,
\]

\[
\frac{\sqrt{u_2^*}}{(d-u_1^*)^2} (1 + \alpha_1)(u_1 - u_1^*) + \frac{1}{2\sqrt{u_2^*}(d-u_1^*)} (u_2 - u_2^*) = 0.
\]

An optimal solution \( \hat{u}(\alpha) \) of \( R(\lambda^*, 2, \alpha) \) with condition \( (u_1, u_2) \in U \) dropped, is similarly given by

\[
\hat{u}_1(\alpha) = 2u_1^* - d + (d - u_1^*) \frac{1}{1 + \alpha_1}, \quad \hat{u}_2(\alpha) = (2(1 + \alpha_1) - 1) u_2^*.
\]

By

\[
h_2(u(\alpha)) \leq h_2(\hat{u}(\alpha)) \quad \text{and} \quad \alpha_1 \geq 0,
\]

we have

\[
h_2(u(\alpha)) \leq h_2(\hat{u}(\alpha)) = \frac{1 + \alpha_1}{\sqrt{2\alpha_1 + 1}} \frac{\sqrt{u_2^*}}{d-u_1^*} = \frac{1 + \alpha_1}{\sqrt{2\alpha_1 + 1}} h_2(u^*) \leq (1 + \alpha_1) h_2(u^*).
\]

Therefore

\[
h_2(u(\alpha)) - h_2(u^*) \leq \alpha_1 h_2(u^*) = \| \alpha \| h_2(u^*) \quad \text{(by } \alpha_2 = 0).\]

Consequently, the following \( \varrho \) satisfies (A5').

\[
\varrho_2(\| \alpha \|) = \| \alpha \|.
\]

(iii) \( h_3(u_1, \ldots, u_m) = u_1 u_2 \cdots u_m \): We omit the details, but note that

\[
\varrho_3(\| \alpha \|) = \left(1 + \frac{\| \alpha \|}{\sqrt{m}}\right)^m - 1
\]

satisfies (A5).

(iv) \( h_4(u_1, \ldots, u_m) = N - \sum_{j=1}^m (q_j - u_j)^2 \): We omit the details, but note that

\[
\varrho_4(\| \alpha \|) = \| \alpha \|^2 + \sqrt{m} \| \alpha \|
\]

satisfies (A5). \( \square \)

Summarizing these results, the next lemma is obtained.

**Lemma 10.1.** The above functions \( h_1 \) and \( h_2 \) satisfy assumptions (A1) and (A5'), and \( h_3 \) and \( h_4 \) satisfy assumptions (A1) and (A5).
11. Development of polynomial time approximation schemes

We have seen in Theorem 7.1, Corollary 7.1, Theorems 8.1 and 9.1 that polynomial time approximation schemes exist if problem $P$ satisfies assumptions (A1)-(A5) (or their modifications) and some additional assumptions. The discussion in Section 10 has shown that assumptions (A1) and (A5) (or (A5')) hold for typical objective functions $h_1, h_2, h_3$ and $h_4$. In this section, therefore, we consider other assumptions, and argue that (fully) polynomial time approximation schemes exist for various problems practically important.

Before proceeding further, recall that assumptions (A2), (A3) and (A4) depend on how $U$ is defined, as well as the objective function $h$. $U$ satisfies (2) and hence depends on set $V$, which corresponds to the feasible region $X$ of $P$.

Lemma 11.1. Assume that $f_j$, $j = 1, 2, \ldots, m$, defined over $X$ are nonnegative integer valued functions with an upper bound $M$:

$$f_j(x) \leq M \text{ for } x \in X, \quad j = 1, 2, \ldots, m.$$  

Furthermore, assume that $\log M$ is a polynomial in the input size of problem $P$. Then $\log(D/Q)$ is polynomial in the input size of $P$ for $h_1, h_2$, $h_3$, and $h_4$ of (50)-(53).

Proof. For simplicity, we consider $h_1$ only because other cases are similar. Since $\partial h_1 / \partial u_1 = 1$ and $\partial h_1 / \partial u_2 = \beta / 2 \sqrt{u_2}$, we have

$$\frac{\beta}{2 \sqrt{M}} \leq \frac{\partial h_1}{\partial u_2} \leq \frac{\beta}{2} \text{ for } u \in V,$$

and hence

$$D = \max \left\{ \frac{\beta}{2}, 1 \right\}, \quad d = \min \left\{ 1, \frac{\beta}{2 \sqrt{M}} \right\}.$$

Thus

$$\min \left\{ \frac{\beta}{2}, 1, \frac{2 \sqrt{M}}{\beta} \right\} \leq \frac{\bar{D}}{D} \leq \max \left\{ \sqrt{M}, \frac{2 \sqrt{M}}{\beta}, \frac{\beta}{2} \right\},$$

i.e., $\log(\bar{D}/D)$ is polynomial in the input size of $P$. \qed

Theorem 11.1. For problems $S_1$ and $S_2$ defined by (4) and (5), assume that $x_i$, $i = 1, 2, \ldots, n$, are 0-1 variables and the corresponding deterministic problem (3) (i.e., $c_i$, $i = 1, \ldots, n$, are given constants) has a polynomial time algorithm. Then both $S_1$ and $S_2$ have fully polynomial time approximation schemes.

Proof. First note that set $U$ for $S_1$ and $S_2$ can be considered as in (50) and (51). Thus (A2) is satisfied. Functions $h_1$ for $S_1$ and $h_2$ for $S_2$ satisfy (A1) and (A5') (Lemma 10.1). Assumption (A3') is also obvious. Since $x_i$ are 0-1 variables,
The minimization of a quasiconcave program

\[ f_1(x) = \sum_{i=1}^{n} m_i x_i \leq \sum_{i=1}^{n} m_i, \]

\[ f_2(x) = \sum_{i=1}^{n} v_i^2 x_i \leq \sum_{i=1}^{n} v_i^2 \]

hold. \( f_1(x) \) and \( f_2(x) \) are nonnegative and integer valued since \( m_i \) and \( v_i^2 \) are assumed to be nonnegative integers in Section 2. Thus letting

\[ M = \max \left\{ \sum_{i=1}^{n} m_i, \sum_{i=1}^{n} v_i^2 \right\}, \]

we see that (A4) is satisfied as shown in the proof of Lemma 11.1, which also shows that \( \log(D/D) \) is a polynomial in the input size. \( P(\lambda) \) for \( S_1 \) and \( S_2 \) is the same form as (3) with \( c_i = \lambda_1 m_i + \lambda_2 v_i^2, \ i = 1, \ldots, n. \) Thus by assumption, \( P(\lambda) \) has a polynomial time algorithm. For both \( S_1 \) and \( S_2, \) we have

\[ \log^{-1}(1 + \varepsilon) = \log^{-1}(1 + \varepsilon) = \varepsilon \]

by (57) and (61). Since

\[ \log^{-1}(1 + \varepsilon) \leq \frac{1 + \varepsilon}{\varepsilon^2} \quad \text{for} \quad 0 < \varepsilon < 1, \]

\[ \log^{-1}(1 + \varepsilon) \leq \log^{-1}2 \quad \text{for} \quad \varepsilon \geq 1, \]

(64)

\[ \log^{-1}(1 + \delta(\varepsilon,2)) = \log^{-1}(1 + \log^{-1}(1 + \varepsilon)) \] (see (26)) is polynomial in \( 1/\varepsilon. \) Therefore, by Theorem 9.1, Algorithm APPROX2 is a fully polynomial time approximation scheme for \( S_1 \) and \( S_2. \)

**Corollary 11.1.** For problems \( S_1 \) and \( S_2, \) if \( x_i, \ i = 1, \ldots, n \) are 0-1 variables and the corresponding deterministic problem (3) has a fully polynomial time approximation scheme, then both \( S_1 \) and \( S_2 \) have fully polynomial time approximation schemes.

**Proof.** Immediate from Theorems 8.1, 9.1 and 11.1. □

It is noted here that a large class of combinatorial optimization problems written in the form of (3) are polynomially solvable or have fully polynomial time approximation schemes, e.g., shortest path, minimum weight perfect matching, minimum directed spanning tree, minimum cut, 0-1 knapsack, and so on. Theorem 11.1 and Corollary 11.1 state that their stochastic versions \( S_1 \) and \( S_2 \) all have fully polynomial time approximation schemes.

**Theorem 11.2.** Assume the conditions in Lemma 11.1 and assumptions (A3) and (A4). Also assume that \( P(\lambda) \) has a polynomial time algorithm. Then problem \( P \) with \( z(x) = h_3(f_1(x), \ldots, f_m(x)) \) or \( z(x) = h_4(f_1(x), \ldots, f_m(x)) \) has a fully polynomial time approximation scheme.
Proof. By assumption and the discussion in Section 10, all of (A1), \ldots, (A5) are satisfied. Thus, by Theorem 7.1, procedure APPROX is an approximation scheme. Since \( \log(D/D) \) is polynomial in the input size of \( P \) by Lemma 11.1 and \( P(\lambda) \) is polynomially solvable, APPROX is a polynomial time approximation scheme.

To complete the proof, we shall show that
\[
\log^{-1}(1 + \delta_3(\epsilon, m)) = \log^{-1}(1 + q_3^{-1}(\epsilon)/\sqrt{m-1})
\]
and
\[
\log^{-1}(1 + \delta_4(\epsilon, m)) = \log^{-1}(1 + q_4^{-1}(\epsilon)/\sqrt{m-1})
\]
are polynomial in \( 1/\epsilon \). Then by Corollary 7.1(ii) APPROX is a fully polynomial time approximation scheme. From (62),
\[
q_3^{-1}(\epsilon) = \sqrt{m}[(\epsilon + 1)^{1/m} - 1]
\]
and hence
\[
\log\left(1 + \frac{q_3^{-1}(\epsilon)}{\sqrt{m-1}}\right) \geq \log\left(1 + \frac{q_3^{-1}(\epsilon)}{\sqrt{m}}\right) = \frac{1}{m} \log(1 + \epsilon).
\]
Thus
\[
\log^{-1}\left(1 + \frac{q_3^{-1}(\epsilon)}{\sqrt{m-1}}\right) \leq m \log^{-1}(1 + \epsilon)
\]
is polynomial in \( 1/\epsilon \) by (64). Next consider \( q_4 \) of (63).
\[
q_4^{-1}(\epsilon) = \frac{2\epsilon}{\sqrt{m + 4\epsilon} + \sqrt{m}}
\]
and hence
\[
\frac{q_4^{-1}(\epsilon)}{\sqrt{m-1}} \geq \frac{1}{\sqrt{m}} \left(\frac{2\epsilon}{\sqrt{m + 4} + \sqrt{m}}\right) \geq \frac{\epsilon}{m + 4} \quad \text{for } 0 < \epsilon \leq 1,
\]
\[
\frac{q_4^{-1}(\epsilon)}{\sqrt{m-1}} \geq -\frac{\sqrt{m + 4}}{2\sqrt{m-1}} \quad \text{for } \epsilon > 1.
\]
This means by (64) that \( \log^{-1}(1 + q_4^{-1}(\epsilon)/\sqrt{m-1}) \) is polynomial in \( 1/\epsilon \).

Acknowledgement

The authors are indebted to Dr. H. Kawai of Osaka Prefectural University for pointing out that the variance minimization problem for the markovian decision process explained in Section 2 belongs to the class of problem \( P \). They also wish to thank Professor H. Mine of Kansai University and Professor T. Hasegawa of Kyoto University for their support and encouragement. This work was partially supported by the Ministry of Education, Science and Culture of Japan under Scientific Research Grant-in-Aid.
References