Parametric Nonlinear Programming For General Cases and Its Application to Some Problems

By

Hisashi MINE, *Masao FUKUSHIMA* and YONG JOON RYANG*

(Received March 31, 1978)

Abstract

This paper deals with a general nonlinear programming problem depending on a scalar parameter. Two algorithms are presented to obtain a parametric optimal solution of the problem by reducing it successively to associated problems which contain a smaller number of variables. The reduction is accomplished by partitioning variables into basic and nonbasic variables, and also by generating a reduced problem from only nonbasic variables. It is shown that both algorithms are essentially equivalent to each other. The finiteness of the algorithms is proved under certain assumptions.

Application of parametric programming to handle some (originally nonparametric) problems is also indicated.

1. Introduction

In a previous paper[11], a method of parametric programming has been examined by the authors for a linearly constrained programming problem with a nonlinear objective function, depending on a scalar parameter. The underlying idea of the Basic Algorithm proposed there is to reduce the problem to a relaxed one by expressing basic variables as functions of nonbasic variables. A striking feature of the reduced parametric problem is that the Kuhn-Tucker conditions for the problem do not involve any Lagrange multiplier explicitly, but only nonbasic variables.

Evidently the fundamental idea of the above method is similar to that of the reduced gradient method for solving nonlinear programs with linear constraints[13]. Therefore, it may be reasonable to try to generalize the method of parametric programming for linear constraints to the case of nonlinear constraints, which also depend on a parameter in a way similar to that in which the reduced gradient

* Department of Applied Mathematics and Physics

method was generalized for nonlinear constraints by Abadie and Carpentier[1]. In fact, as will be shown in the next section, the basic theorems stated in [11, section 2] concerning the relationship between the solutions of the reduced problem and those of the original problem may be clearly generalized. On the basis of such a generalization, in section 3, two algorithms for finding optimal solutions of a parametric nonlinear program are presented, and further computational aspects are discussed. In section 4, four applications of parametric programming are indicated. In section 5, a numerical example is given to illustrate the functioning of this program.

2. The Generalized Reduced Parametric Problem

In this paper, the following parametric nonlinear program is considered: For each $t\in T$, find an n-vector $x(t)$ that

$$\begin{aligned} &\text{minimizes} && f(x,\ t) && (1)\\ &\text{subject to} && g(x,\ t)=0,\ x\geqq 0, \end{aligned}$$

where t is a scalar parameter varying in a compact interval T, and the functions f: $R^n\times T\to R$ and $g=(g_1,\ \ldots,\ g_m)^T$: $R^n\times T\to R^m$ are twice continuously differentiable in x and t. Notice that any nonlinear program involving a scalar parameter may be expessed in the form

$$\begin{aligned} &\text{minimize}_x && f(x,\ t) && (2)\\ &\text{subject to} && g(x,\ t)=0,\ a\leqq x\leqq b, \end{aligned}$$

where the components of a and b are allowed to be $-\infty$ or $+\infty$. It suffices to consider only equality constraints, since any inequality constraint can be made into an equation by introducing a slack variable. The reason why we consider problem (1) rather than problem (2) is merely for simplicity of exposition. It should be noted that the results obtained in this paper may easily be translated into those for problem (2) by suitable modifications.

In this section, we apply the implicit function theorem to problem (1), and obtain some results which may be viewed as an immediate generalization of those given in [11]. For problem (1), we adopt the following *nondegeneracy assumption*: For any $t\in T$ and for any feasible solution x, i.e., $g(x,\ t)=0$ and $x\geqq 0$, there exists a partition of x into y and z, where y is m-dimensional and z is $(n-m)$-dimensional, such that the $m\times m$ matrix $\nabla_y g(y,\ z,\ t)$ is nonsingular and $y>0$.

The components of the vectors y and z are called *basic* and *nonbasic variables*, respectively. With respect to the partition $x=(y,\ z)$, the matrices $\nabla_x f(x,\ t)$ and $\nabla_x g(x,\ t)$ are partitioned as $[\nabla_y f(x,\ t),\ \nabla_z f(x,\ t)]$ and $[\nabla_y g(x,\ t),\ \nabla_z g(x,\ t)]$, respectively.

Given $\bar{t}\in T$, let $\bar{x}$ be a feasible solution of (1). Then by the nondegeneracy assumption, the implicit function theorem assures that we can find a partition $x=(y,z)$

and a twice continuously differentiable function h: $R^{n-m}\times T\to R^m$ implicitly determined by solving the nonlinear equation

$$g(y,\ z,\ t)=0$$

for all $(z,\ t)$ in some neighborhood of $(\bar{z},\bar{t})$, namely,

$$\bar{y}=h(\bar{z},\bar{t})$$

and

$$g(h(z,\ t),\ z,\ t)=0 \tag{3}$$

for all $(z,\ t)$ in the neighborhood of $(\bar{z},\bar{t})$. Consequently, we may define, at least locally, a *generalized reduced parametric problem* as follows: For $t\in T$,

$$\begin{aligned} &\text{minimize}_z && F(z,\ t)\triangleq f(h(z,\ t),\ z,\ t) \\ &\text{subject to} && z\geqq 0, \end{aligned} \tag{4}$$

where h is the implicit function defined by (3). Notice that the objective function F is twice continuously differentiable in x and t, because both f and h have the same property. It is to be noted that more than one such reduced problem may be defined in the neighborhood of $(\bar{x},\bar{t})$, since a possible choice of basic variables is not necessarily unique.

In general, the function F is not available in a closed form except in some simple cases like those with linear constraints. Fortunately however, the gradient $\nabla_z F(z,\ t)$ and the Hessian matrix $\nabla^2_{zz}F(z,\ t)$ can be expressed in terms of the gradients and the Hessians of f and g. Specifically, we have

$$\nabla_z F(z,t)=\nabla_z f(h(z,t),z,t)+\nabla_y f(h(z,t),z,t)\nabla_z h(z,t). \tag{5}$$

Differentiating (3) with respect to z yields

$$\nabla_z g(h(z,t),z,t)+\nabla_y g(h(z,t),z,t)\nabla_z h(z,t)=0. \tag{6}$$

Since the matrix $\nabla_y g(h(z,\ t),\ z,\ t)$ is nonsingular, eliminating $\nabla_z h$ from (5) and (6), we obtain

$$\begin{aligned} \nabla_z F(z,t)=&\nabla_z f(h(z,t),z,t) \\ &-\nabla_y f(h(z,t),z,t)[\nabla_y g(h(z,t),z,t)]^{-1}\nabla_z g(h(z,t),z,t) \end{aligned} \tag{7}$$

which is called the *reduced gradient* (with respect to z.)

Let

$$\lambda(x,t)\triangleq\nabla_y f(x,t)[\nabla_y g(x,t)]^{-1} \tag{8}$$

Then (7) may be written as

$$\nabla_z F(z,t)=\nabla_z f(h(z,t),z,t)-\lambda(h(z,t),z,t)\nabla_z g(h(z,t),z,t). \tag{9}$$

Now (8) implies

$$\nabla_y f(h(z,t),z,t)-\lambda(h(z,t),z,t)\nabla_y g(h(z,t),z,t)=0.$$

Differentiating this with respect to z, we have[1]

$$\nabla^2_{y,z}f-\lambda\nabla^2_{y,z}g+[\nabla^2_{y,y}f-\lambda\nabla^2_{y,y}g]\nabla_z h-\nabla_y g^T\frac{d\lambda}{dz}=0. \tag{10}$$

---

[1] $\lambda\nabla^2_{y,y}g=\sum_{i=1}^{m}\lambda_i\nabla^2_{y,y}g_i$, etc.

Similarly, differentiating (9) yields

$$\nabla_{zz}^2 F(z,t)=\nabla_{zz}^2 f-\lambda\nabla_{zz}^2 g+[\nabla_{zy}^2 f-\lambda\nabla_{zy}g]\nabla_z h-\nabla_z g^T\frac{d\lambda}{dz}. \tag{11}$$

Eliminating $\frac{d\lambda}{dz}$ from (11) by using (10), and taking account of (6), we obtain

$$\begin{aligned}\nabla_{zz}^2 F(z,t)&=[\nabla_z h^T, I_{n-m}]\begin{bmatrix}\nabla_{yy}^2 f-\lambda\nabla_{yy}^2 g & \nabla_{yz}^2 f-\lambda\nabla_{yz}^2 g\\ \nabla_{zy}^2 f-\lambda\nabla_{zy}^2 g & \nabla_{zz}^2 f-\lambda\nabla_{zz}^2 g\end{bmatrix}\begin{bmatrix}\nabla_z h\\ I_{n-m}\end{bmatrix}\\ &=\Gamma^T[\nabla_{zz}^2 f-\lambda\nabla_{zz}^2 g]\Gamma,\end{aligned} \tag{12}$$

where

$$\Gamma=\begin{bmatrix}-\nabla_y g^{-1}\nabla_z g\\ I_{n-m}\end{bmatrix}. \tag{13}$$

Notice that all functions in (10)-(13) are evaluated at $(x,\ t)=(h(z,\ t),\ z,\ t)$.

Now review the second-order sufficiency conditions for isolated local minima of problems (1) and (4). A statement of the conditions for more general nonlinear programs can be found elsewhere, for example, in [9, p. 235], and hence, no proof is given here.

For $t\in T$, if an n-vector $x(t)$ satisfies

$$g(x(t),\ t)=0,\ x(t)\geqq 0$$

and if there exists an m-vector λ such that

$$\nabla_x f(x(t),t)-\lambda\nabla_x g(x(t),t)\geqq 0$$
$$[\nabla_x f(x(t),t)-\lambda\nabla_x g(x(t),t)]x(t)=0$$

and

$$s^T[\nabla_{xx}^2 f(x(t),t)-\lambda\nabla_{xx}^2 g(x(t),t)]s>0$$

for all non-zero n-vectors s, such that $\nabla_x g(x(t),t)s=0$ and $s_i=0$ for $i\in\{i;[\nabla_x f(x(t),t)-\lambda\nabla_x g(x(t),t)]_i>0\}$, then $x(t)$ is an isolated local minimum of problem (1).

For $t\in T$, if an $(n-m)$-vector $z(t)$ satisfies

$$\begin{cases}\nabla_z F(z(t),t)\geqq 0,\ z(t)\geqq 0\\ \nabla_z F(z(t),t)\,z(t)=0\end{cases} \tag{14}$$

and

$$v^T\nabla_{zz}^2 F(z(t),t)v>0$$

for all non-zero $(n-m)$-vectors v, such that $v_j=0$ for $j\in\{j;\nabla_{z_j}F(z(t),t)>0\}$, then $z(t)$ is an isolated local minimum of problem (4).

The following two theorems state important relationships between the optimality conditions for problems (1) and (4), on which the algorithms in the next section are based. The proofs can be completed by using (7) and (12) in a manner quite analoguous to that for Theorems 1 and 2 in [11], and hence are omitted.

Theorem 1. For $t\in T$, if $x(t)$ satisfies the second-order sufficiency conditions for a local minimum of problem (1), then there exists a partition $x(t)=[y(t),\ z(t)]$ such

that $y(t)>0$ and that $z(t)$ satisfies the second-order sufficiency conditions for a local minimum of the generalized reduced problem (4).

Theorem 2. For $t\in T$ and an $(n-m)$-vector $z(t)$, let h be an implicit function determined by (3) on a neighborhood of $(z(t), t)$. If $z(t)$ satisfies the second-order sufficiency conditions for a local minimum of the generalized reduced problem (4), and if $h(z(t), t)\geqq 0$, then the n-vector $[y(t), z(t)]$ satisfies the second-order sufficiency conditions for a local minimum of the original problem (1), where $y(t)=h(z(t), t)$.

In Theorems 1 and 2, we have tacitly assumed that we know the function h explicitly when handling (14). In practice, however, we need to solve another nonlinear equation (3) separately to have $h(z, t)$. Motivated by this fact, we consider the following system of equations and inequalities in place of (14):

$$\begin{cases} g(y, z, t)=0 \\ H(y, z, t)\geqq 0, \quad z\geqq 0 \\ H(y, z, t)z=0 \end{cases} \tag{15}$$

where, for the partition $x=(y, z)$, the function H: $R^n\times T\to R^{n-m}$ is defined by

$$H(x, t)=\nabla_z f(x, t)-\nabla_y f(x, t)[\nabla_y g(x, t)]^{-1}\nabla_z g(x, t).$$

It is easily verified that

$$\begin{aligned} \nabla_z H(x, t) &= [\nabla_y H(x, t), \nabla_z H(x, t)] \\ &= \begin{bmatrix} -\nabla_y g^{-1}\nabla_z g \\ I_{n-m} \end{bmatrix}^T \begin{bmatrix} \nabla^2_{yy} f-\lambda\nabla^2_{yy} g & \nabla^2_{yz} f-\lambda\nabla^2_{yz} g \\ \nabla^2_{zy} f-\lambda\nabla^2_{zy} g & \nabla^2_{zz} f-\lambda\nabla^2_{zz} g \end{bmatrix} \\ &= \Gamma^T[\nabla^2_{xx} f(x, t)-\lambda(x, t)\nabla^2_{xx} g(x, t)], \end{aligned} \tag{16}$$

where λ and Γ are given by (8) and (13), respectively. Notice that λ and Γ are evaluated at $(x, t)=(y, z, t)$ in (16), while these are evaluated at $(x, t)=(h(z, t), z, t)$ in (12). Theorem 3 shows that (14) may be replaced by (15) in the optimality conditions for the reduced problem.

Theorem 3. For $t\in T$, suppose that an n-vector $x(t)=[y(t), z(t)]$ solves (15). If $y(t)\geqq 0$ and if

$$v^T\Gamma^T[\nabla^2_{xx} f-\lambda\nabla^2_{xx} g]\Gamma v>0 \tag{17}$$

for all v, such that $v_j=0$ for $j\in\{j; H_j(x(t), t)>0\}$, where $\nabla^2_{xx} f$, $\nabla^2_{xx} g$, λ and Γ are evaluated at $(x(t), t)$, then $x(t)$ satisfies the second-order sufficiency conditions for problem (1). Conversely, if $x(t)$ satisfies the second-order sufficiency conditions for problem (1), there exists a partition of x into y and z, such that $y(t)>0$ and that $x(t)=[y(t), z(t)]$ satisfies (15) and (17).

Proof. First note that $y(t)=h(z(t), t)$. Now, the theorem immediately follows from Theorems 1 and 2 and the fact that

$$\nabla_z F(z(t), t)=H(x(t), t)$$

and

$$\nabla^2_{zz} F(z(t), t)=\Gamma^T[\nabla^2_{xx} f(x(t), t)-\lambda(x(t), t)\nabla^2_{xx} g(x(t), t)]\Gamma. \qquad \square$$

3. Methods of Computing Parametric Optimal Solutions

In what follows, we assume the existence and the continuity of $x(t)$ which satisfies the second-order sufficiency conditions for a local minimum of problem (1) for every $t\in T$. Under this assumption, we shall propose two algorithms for solving problem (1) parametrically. Without any loss of generality, let T be the unit interval $[0, 1]$ in the rest of this section.

On the basis of Theorems 1 and 2, we first state the following algorithm which is a straightforward generalization of the Basic Algorithm presented in [11].

Algorithm 1.

Step 1: Obtain an optimal solution $x(0)$ of problem (1) for $t=0$ by an appropriate method. Choose a sufficiently small number $\gamma\geqq 0$ and set $t'=0$. Go to step 2;

Step 2: Partition x into basic y and nonbasic z, such that the corresponding basis matrix is nonsingular and all components of $y(t')$ have a value greater than γ. Go to step 3.1;

Step 3.1: Choose sets I and J of indices such that

$$\{j;\nabla_{z_j}F(z(t'),t')>0\}\subset J\subset\{j;z_j(t')=0\}$$

and $I=\{1,2,\ldots,n-m\}-J$. Go to step 3.2;

Step 3.2: Obtain the solution $z(t)$ of the system of equations

$$\begin{cases}\nabla_{z_I}F(z,t)=0\\ z_J=0\end{cases}\tag{18}$$

where z_I and z_J are the vectors with components z_i, $i\in I$, and z_j, $j\in J$, respectively, and solve (3) to get $y(t)=h(z(t),t)$, as t increases from t' to $t=t^\circ\triangleq\min\{t^*, t^{**}\}$, where

$$t^*=\sup\{t;y_i(\tau)>\gamma\ \forall i,\text{for all }\tau\text{ such that }t'\leqq\tau\leqq t\leqq 1\}\tag{19}$$

and

$$t^{**}=\sup\{t;\nabla_{z_J}F(z(\tau),\tau)\geqq 0\text{ and }z_I(\tau)\geqq 0\text{ for all }\tau,\text{ such that }t'\leqq\tau\leqq t\leqq 1\}.$$

Go to step 3.3;

Step 3.3: If $t^\circ=1$, terminate. Otherwise, setting $t'=t^\circ$, return to step 2 if $t^\circ=t^*$, and return to step 3.1 if $t^\circ=t^{**}$

Notice that in step 2, we always have $y(t')>0$ by virtue of the nondegeneracy assumption. Thus, it can never happen that $t'=t^*$ in step 3.2. On the other hand, in general, it is not always the case that $t'<t^{**}$ in step 3.2, unless the sets are properly determined in step 3.1. Now the question is how we may find such I and J, especially when the strict complementarity, i.e.,

$$\{j;\nabla_{z_j}F(z(t'),t')>0\}=\{j;z_j(t')=0\}$$

does not hold. However, this problem is basically the same as that of finding an

index set among all *valid* sets in step 3 of Geoffrion's Basic Parametric Procedure [4], where a trial and error technique is described in detail. Hence, we shall not be concerned with that matter in this problem. Thus, it is supposed that adequate index sets I and J are always specified in step 3.1 so that $t'<t^{**}$ holds.

Next, we propose another algorithm for finding a parametric solution of problem (1).

Algorithm 2.

Steps 1 and 2: Same as those in Algorithm 1;

Step 3.1: Choose I and J, such that

$$\{j; H_j(x(t'), t')>0\} \subset J \subset \{j; z_j(t')=0\}$$

and $I=\{1, 2, \ldots, n-m\}-J$. Go to step 3.2;

Step 3.2: Obtain the solution $x(t)=[y(t), z(t)]$ of the system of equations

$$\begin{cases} g(y, z, t)=0 \\ H_I(y, z, t)=0 \\ z_J=0 \end{cases} \tag{20}$$

as t increases from t' to $t=t^\circ \triangleq \min\{t^*, t^{**}\}$, where t^* is defined by (19) and

$$t_{**}=\sup\{t; H_J(x(\tau), \tau)\geqq 0 \text{ and } z_I(\tau)\geqq 0 \text{ for all } \tau, \text{ such that } t'\leqq\tau\leqq t\leqq 1\}$$

where $H=[H_I, H_J]$ and

$$H_I=\nabla_{z_I} f-\nabla_y f[\nabla_y g]^{-1}\nabla_{z_I} g=\nabla_{z_I} f-\lambda\nabla_{z_I} g,$$

$$H_J=\nabla_{z_J} f-\nabla_y f[\nabla_y g]^{-1}\nabla_{z_J} g=\nabla_{z_J} f-\lambda\nabla_{z_J} g.$$

Go to step 3.3;

Step 3.3: Same as that in Algorithm 1.

It follows immediately from Theorem 3 that Algorithms 1 and 2 are essentially equivalent to each other. To summarize the procedures, we determine parametrically the values of nonbasic variables by solving the equations which are derived from the optimality conditions for the generalized reduced problem. We also monitor the basic variables to decide whether the current basis is changed or not.

Both Algorithms 1 and 2 are ideal in the sense that we assume that (18) and (20) are solved for all *continuous* values of the parameter t. In practice, it is usually not possible to solve parametric nonlinear equations continuously except in very simple cases. Although it is often possible, as was indicated in [11] for linearly constrained problems, to derive a differential equation from the system of equations and apply some numerical integration method, we shall content ourselves here with the so called *discretization* approach to the parametric solution of nonlinear equations. Namely, systems of nonlinear equations are solved by some iterative method successively for a finite number of representative values of the parameter.

Consider first the system (18). It should be noted that $z(t_1)$ and $z(t_2)$ are ex-

pected to be close to each other, provided $|t_1-t_2|$ is sufficiently small, by virtue of the continuity of $z(t)$. Thus, if $z(t_1)$ is known, iteration for solving a system for $t=t_2$ with the starting point $z(t_1)$ may converge to $z(t_2)$. A typical procedure of discretization is as follows: Let $\Delta t>0$ be sufficiently small and $z(t)$ be given for the initial value of t, say t_0. Put $t_p=t_0+p\Delta t$, $p=1, 2, \ldots$. For each p, compute $z(t_p)$ as a limit of Newton's iteration

$$z^{k+1}=\begin{bmatrix} z_I^{k+1} \\ z_J^{k+1} \end{bmatrix}=\begin{bmatrix} z_I^k-[\nabla^2_{z_I z_I}F(z^k, t_p)]^{-1}\nabla_{z_I}F(z^k, t_p) \\ 0 \end{bmatrix} \tag{21}$$

with the initial condition $z^0=z(t_{p-1})$.

Note that at every iteration (21), we must solve

$$g(y, z^k, t_p)=0 \tag{22}$$

for fixed z^k and t_p to obtain $y^k=h(z^k, t_p)$ by using an iterative method in order to evaluate $\nabla^2_{z_I z_I}F(z^k, t_p)$ and $\nabla_{z_I}F(z^k, t_p)$. Therefore, the above procedure is regarded as a two-fold iteration in the sense that each iteration (21) involves another set of iterations for solving (22).

We now turn our attention to the solution of (20). By the continuity of $x(t)$, a procedure similar to that for (18) can be applied. In this case, Newton's iteration corresponding to (21) may be written as

$$\begin{aligned} x^{k+1} &= \begin{bmatrix} y^{k+1} \\ z_I^{k+1} \\ z_J^{k+1} \end{bmatrix} \\ &= \left(\begin{bmatrix} y^k \\ z_I^k \end{bmatrix} - \begin{bmatrix} \nabla_y g(y^k, z^k, t_p) & \nabla_{z_I} g(y^k, z^k, t_p) \\ \nabla_y H_I(y^k, z^k, t_p) & \nabla_{z_I} H_I(y^k, z^k, t_p) \end{bmatrix}^{-1} \begin{bmatrix} g(y^k, z^k, t_p) \\ H_I(y^k, z^k, t_p) \end{bmatrix} \right), \\ & \qquad\qquad 0 \end{aligned} \tag{23}$$

where, of course, the initial condition is $x^0=x(t_{p-1})$.

It is to be noted that the inverse matrices which appear in (21) and (23) exist, provided (y^k, z^k) is sufficiently close to $(y(t_p), z(t_p))$. This is easily verified for (21), since $\nabla^2_{z_I z_I}F(z(t), t)$ is continuous with respect to t, and since the invertibility (or, more precisely, the positive definiteness) of $\nabla^2_{z_I z_I}F(z(t_p), t_p)$ follows from the fact that $z(t_p)$ satisfies the second-order sufficiency conditions for optimality. To show a similar result for (23), it suffices to prove that the matrix

$$\begin{bmatrix} \nabla_y g(y, z, t) & \nabla_{z_I} g(y, z, t) \\ \nabla_y H_I(y, z, t) & \nabla_{z_I} H_I(y, z, t) \end{bmatrix} \tag{24}$$

is nonsingular for $(y, z)=(y(t), z(t))$. Here, we show that this matrix has a non-zero determinant. Since the matrix $\nabla_y g$ is nonsingular by the nondegeneracy assumption, it is easy to see that

$$\det \begin{bmatrix} \nabla_y g & \nabla_{z_I} g \\ \nabla_y H_I & \nabla_{z_I} H_I \end{bmatrix}$$

$$= \det[\nabla_y g] \cdot \det[\nabla_{z_I} H_I - \nabla_y H_I (\nabla_y g)^{-1} \nabla_{z_I} g]$$

$$= \det[\nabla_y g] \cdot \det \left[\begin{bmatrix} -(\nabla_y g)^{-1} \nabla_{z_I} g \\ I_r \end{bmatrix}^T \begin{bmatrix} \nabla^2_{yy} f - \lambda \nabla^2_{yy} g & \nabla^2_{yz_I} f - \lambda \nabla^2_{yz_I} g \\ \nabla^2_{z_I y} f - \lambda \nabla^2_{z_I y} g & \nabla^2_{z_I z_I} f - \lambda \nabla^2_{z_I z_I} g \end{bmatrix} \begin{bmatrix} -(\nabla_y g)^{-1} \nabla_{z_I} g \\ I_r \end{bmatrix} \right]$$

$$= \det[\nabla_y g] \cdot \det[\nabla^2_{z_I z_I} F]$$

for $(y, z) = (y(t), z(t))$, where I_r is the identity matrix of order r, and r is the number of elements in set I. Obviously $\det[\nabla_y g] \neq 0$ and, by the same reason as before, we have $\det[\nabla^2_{z_I z_I} F] \neq 0$. This proves the nonsingularity of (24) for $(y, z) = (y(t), z(t))$.

We have adopted Newton's formula (21) and (23) to solve (18) and (20), respectively. This is for expository purposes only. Obviously, it is also possible to apply other methods such as the conjugate gradient or the quasi-Newton methods.

Notice that step 3.2 of Algorithm 2 requires the solution of $m+r$ equations (20) in $m+r$ unknowns y and z_I, while that of Algorithm 1 needs the solution of r equations (18) in r unknowns z_I. However, at each step of the iteration one should solve a system (22) of m equations in m unknowns y, where r may vary but does not exceed $n-m$. Although the efficiency of the algorithms might entirely depend upon the structure of equations (18) and (20), as for the execution time, the latter seems preferable within our limited numerical experiments.

It is an important problem whether each step of an algorithm will be executed only a finite number of times before termination. For Algorithms 1 and 2, this problem consists of two parts, namely, one is the finiteness of the number of points t^* at which the basis is updated. The other is that of the number of points t^{**} at which I and J are revised. Observe that these questions may be simultaneously answered for either of the Algorithms, since their actions are essentially identical. For the first question, we provide the following.

Theorem 4. Assume a "strong" nondegeneracy: There exists a scalar $\bar{\gamma} > 0$, such that for any feasible x and t, a partition $x = (y, z)$ exists such that $\nabla_y g(y, z, t)$ is nonsingular and $y_i > \bar{\gamma}$ for all i. If either of Algorithms 1 and 2 with any γ, such that $0 < \gamma \leq \bar{\gamma}$ is applied to problem (1) whose solution exists and is continuous with respect to t, then the number of points t^* at which the basis change occurs is finite in T.

Proof. The theorem can be proved in a similar manner to the one for [11, Theorem 3], and hence is ommitted here.

Another important question concerns the finiteness of the number of points t^{**} where sets I and J are revised. A sufficient condition for the finiteness may be the analyticy of the functions f and g on some open set containing $X\times T$, where X is the feasible region of problem (1). Noting that function F defined in (4) also becomes analytic by the implicit function theorem, the proof may be completed in a way quite analogous to that of [4, Theorem 3].

4. Application

In this section, we show four examples of applying parametric programming to different kinds of problems, that is, stochastic programming, fractional programming, bicriterion mathematical programming and a deformation method for nonlinear programs. It is shown that a parametric programming problem, to which Algorithm 1 or 2 may be applied, plays an important role in each of them.

4.1. Stochastic Programming Let a stochastic programming problem be

$$\begin{aligned} &\text{minimize}_x && f(x, a)\\ &\text{subject to} && g(x, a)=0,\ x\geqq 0, \end{aligned}$$

where a is a random variable with some known distribution.

The "wait-and see" problem associated with this problem may be written as

$$E\min_x\{f(x, a)\,|\,g(x, a)=0, x\geqq 0\}, \tag{25}$$

where E denotes the expectation taken with respect to the distribution of a. This problem and related ones are studied elsewhere, for example, [10]. Obviously, problem (25) involves a parametric nonlinear programming problem of the form (1). Algorithm 1 and 2 may be efficient in particular if the random variable has a continuous distribution.

4.2. Fractional Programming A general nonlinear fractional programming prob lem is

$$\begin{aligned} &\text{minimize} && N(x)/D(x)\\ &\text{subject to} && x\in S\triangleq\{x;g(x)=0, x\geqq 0\}, \end{aligned} \tag{26}$$

where it is assumed that $D(x)>0$ for all $x\in S$. Dinkelbach [2] proves an interesting fact that

$$q_0=N(x^\circ)/D(x^\circ)=\min\{N(x)/D(x)\,|\,x\in S\}$$

if and only if

$$N(x^\circ)-q_0D(x^\circ)=\min\{N(x)-q_0D(x)\,|\,x\in S\}=0.$$

Thus, associated with (26) is the following parametric programming problem:

$$\text{minimize}_x\ N(x)-qD(x)\quad\text{subject to } x\in S, \tag{27}$$

and problem (26) reduces to that of finding the value of parameter q for which the minimal value of (27) is exactly zero. Specifically, this is accomplished by solving (27) parametrically for various values of q as it increases (or decreases) until $q=q_0$

is encountered, for which the minimal value becomes zero.

4.3. Bicriterion Mathematical Programming A bicriterion nonlinear programming problem may formally be stated as

$$\begin{aligned} &\text{minimize} \begin{bmatrix} f_1(x) \\ f_2(x) \end{bmatrix} \\ &\text{subject to} \quad g(x)=0, \quad x \geqq 0, \end{aligned} \tag{28}$$

where f_1 and f_2 are real-valued functions. A feasible point x° is said to be *efficient* [5] (noninferior, nondominated, or Pareto optimal), if there does not exist a feasible point x, such that $f_i(x) \leqq f_i(x^\circ)$, $i=1, 2$, with at least one strict inequality. Efficient points play a central role in analyzing vector minimization problems, and it is well known [6] that if x° is an optimal solution of the following *scalarized* problem for some $0<\lambda<1$

$$\begin{aligned} &\text{minimize}_x \quad \lambda f_1(x) + (1-\lambda) f_2(x) \\ &\text{subject to} \quad g(x)=0, \quad x \geqq 0, \end{aligned} \tag{29}$$

then x° is (properly) efficient in problem (28). Moreover, it can be shown that, under a certain convexity assumption, any efficient x° solves problem (29) for some $0 \leqq \lambda \leqq 1$ [5, Lemma 2]. Therefore, finding efficient solutions of (28) may be reduced to the parametric programming problem (29).

Hocking and Shepard [8] propose an alternative parametric approach to the bicriterion mathematical program (28). They consider the following parametric programming problem:

$$\begin{aligned} &\text{minimize}_x \quad f_1(x) \\ &\text{subject to} \quad f_2(x) \leqq \beta \\ &\qquad\qquad g(x)=0, \quad x \geqq 0, \end{aligned} \tag{30}$$

in place of (29), and prove that if x° is efficient in (28), then x° also solves (30) for some β. Clearly, (30) as well as (29) belongs to a class of problems of the form (1).

4.4. A Deformation Method Suppose that the following nonlinear programming problem is to be solved:

$$\begin{aligned} &\text{minimize} \quad f_1(x) \\ &\text{subject to} \quad g_1(x)=0, \quad x \geqq 0. \end{aligned} \tag{31}$$

On the other hand, we suppose that the vector x° is known to be an optimal solution of another problem

$$\begin{aligned} &\text{minimize} \quad f_0(x) \\ &\text{subject to} \quad g_0(x)=0, \quad x \geqq 0, \end{aligned} \tag{32}$$

where g_0 has the same dimension as g_1. In practice, the solution of (32) may be considerably easier to obtain than that of (31).

Consider now a parametric nonlinear program in which problems (31) and (32)

are imbedded:

$$\begin{aligned} &\text{minimize}_x \quad tf_1(x)+(1-t)f_0(x) \\ &\text{subject to} \quad tg_1(x)+(1-t)g_0(x)=0, \quad x\geqq 0. \end{aligned} \tag{33}$$

It is easily verified that (31) and (32) are identified with (33) for $t=0$ and $t=1$, respectively. Therefore, an optimal solution of (31) may be obtained by solving (33) parametrically as t increases until $t=1$ with the initial condition $x=x^\circ$ for $t=0$.

5. Numerical Experience

An example is solved by the deformation method (4. 4) which incorporates Algorithms 1 and 2. The objective function and the constraints of (33) are transformed as follows:

$$\begin{aligned} &\text{minimize} \quad (1-e^{-t})f_1(x)+e^{-t}f_0(x) \\ &\text{subject to} \quad (1-e^{-t})g_1(x)+e^{-t}g_0(x)=0, \quad a\leqq x\leqq b \end{aligned} \tag{34}$$

where the components of a, b are allowed to be $-\infty$ or $+\infty$. It is evident that (33) and (34) are equivalent by taking $t\in[0, 1]$ for (33) and $t\in[0,\infty]$ for (34).

In our experiments, we use $f_0(x)=\|x-x_0\|^2$ and $g_0(x)=g_1(x)+r$ as an initial problem (32) where x_0 and r are vectors, such that $g_1(x_0)=-r$ and $a\leqq x\leqq b$. Using (34), the computer experiments are performed on the well-known Rosen-Suzuki Test Problem:

minimize

$$f(x)=x_1^2+x_2^2+2x_3^2-x_4^2-5x_1-5x_2-21x_3+7x_4$$

subject to

$$g_1(x)=x_1^2+x_2^2+x_3^2+x_4^2+x_1-x_2+x_3-x_4+x_5-8=0$$
$$g_2(x)=x_1^2+2x_2^2+x_3^2+2x_4^2-x_1-x_4+x_6-10=0$$
$$g_3(x)=2x_1^2+x_2^2+x_3^2+2x_1-x_2-x_4+x_7-5=0$$
$$x_5\geqq 0, \quad x_6\geqq 0, \quad x_7\geqq 0$$

where every inequality constraint has been transformed into an equality constraint by introducing slack variables x_5, x_6, x_7.

The optimal solution is $x^*=(0, 1, 2, -1)$ with $f(x^*)=-44$. These computations were given, using double precision, and carried out on the FACOM M 190 computer of Kyoto University Computation Center. The results of the calculations are summarized in the following Tables 1 and 2, with the same initial estimates as given.

Table 1. Computation Results by Algorithm 1

	starting point (opt. sol. of (32))	optimum value (opt. sol. of (31))	starting point (opt. sol. of (32))	optimum vlaue (opt. sol. of (31))	starting point (opt. sol. of (32))	optimum value (opt. sol. of (31))
x_1	0.1	0.390426D−05	0.5	−0.440028D−06	4.	0.112255D−05
x_2	1.2	0.999967D+00	1.5	0.999995D+00	4.	0.999991D+00
x_3	1.8	0.200001D+01	1.5	0.200000D+01	4.	0.200000D+01
x_4	-0.5	−0.100000D+01	−0.7	−0.100000D+01	4.	−0.100000D+01
x_5	1.86	0.0	1.56	0.0	4.	0.0
x_6	2.96	0.100000D+01	1.82	0.100000D+01	4.	0.100000D+01
x_7	0.8	0.0	0.8	0.0	4.	0.0
f	0.0	−0.440000D+02	0.0	−0.440000D+02	0.0	−0.440000D+02
CPU time (sec)	3.957		3.944		4.551	

Table 2. Computation Results by Algorithm 2

	starting point (opt. sol. of (32))	optimum value (opt. sol. of (31))	starting point (opt. sol. of (32))	optimum value (opt. sol. of (31))	starting point (opt. sol. of (32))	optimum value (opt. sol. of (31))
x_1	0.1	0.100593D−10	0.5	−0.819380D−12	4.	−0.432060D−10
x_2	1.2	0.100000D+01	1.5	0.100000D+01	4.	0.100000D+01
x_3	1.8	0.200000D+01	1.5	0.200000D+01	4.	0.200000D+01
x_4	−0.5	−0.100000D+01	−0.7	−0.100000D+01	4.	−0.100000D+01
x_5	1.86	0.0	1.56	0.0	4.	0.0
x_6	2.97	0.100000D+01	1.82	0.100000D+01	4.	0.100000D+01
x_7	0.8	0.0	0.8	0.0	4.	0.0
f	0.0	−0.440000D+02	0.0	−0.440000D+02	0.0	−0.440000D+02
CPU time (sec)	2.760		4.090		4.977	

6. Concluding Remarks

It has been assumed that the optimal solution $x(t)$ of problem (1) is continuous with respect to t. In fact, the continuity of $x(t)$ is an indispensable condition to be satisfied when Algorithm 1 or 2 is applied. Detailed discussions for continuity properties of the solutions of parametric programs can be found elsewhere, e. g., [3, 7, 12].

Theorems 1 and 2 relate the optimal solution of problem (1) to the optimality conditions for the reduced problem involving only nonbasic variables. Theorem 3 gives other conditions which may take the place of those optimality conditions in terms of not only nonbasic variables but also basic ones. It should be noted that (14) and (15) do not contain any Lagrange multiplier, and this leads to a considerable reduction of the number of variables compared with the ordinary Kuhn-Tucker conditions for the original problem (1). Furthermore, the second-order sufficiency conditions ensure that the systems of equations (18) and (20) have nonsingular Jacobians at the solutions, and that the optimal solutions are locally unique as well. The nonsingularity of a Jacobian is a standard assumption in many iterative methods such as Newton's method for solving a system of nonlinear equations.

References

1) J. Abadie and J. Carpentier, "Generalization of the Wolfe Reduced Gradient Method to the Case of Nonlinear Constraints," in Fletcher, R. (ed.), *Optimization*, Academic Press, New York, 1969, pp. 37-47.
2) W. Dinkelbach, "On Nonlinear Fractional Programming," *Management Science*, Vol. 13 (1967), pp. 492-498.
3) J. P. Evans and F. J. Gould, "Stability in Nonlinear Programming," *Operations Research*, Vol. 18 (1970), pp. 107-118.
4) A. M. Geoffrion, "Strictly Concave Parametric Programming, Part I: Basic Theory," *Management Science*, Vol. 13 (1966), pp. 244-253; Part II: Additional Theory and Computational Considerations," *Management Science*, Vol. 13 (1967), pp. 359-370.
5) ————, "Solving Bicriterion Mathematical Programs," *Operations Research*, Vol. 15 (1967), pp. 39-54.
6) ————, "Proper Efficiency and the Theory of Vector Maximization," *Journal of Mathematical Analysis and Applications*, Vol. 22 (1968), pp. 618-630.
7. D. J. Hartfiel and G. L. Curry, "Concerning the Solution Vector for Parametric Programming Problems," *SIAM Journal on Applied Mathematics*, Vol. 24 (1974), pp. 294-296.
8) R. R. Hocking and R. L. Shepard, "Parametric Soution of a Class of Nonconvex Programs, "*Operations Research*, Vol. 19 (1971), pp. 1742-1747.
9) D. G. Luenberger, *Introduction to Linear and Nonlinear Programming*, Addison Wesley, Reading, Mass., 1973.
10) O. L. Mangasarian, "Nonlinear Programming Problems with Stochastic Objective Functions," *Management Science*, Vol. 10 (1964), pp. 353-359.
11) H. Mine, M. Fukushima, and Y. J. Ryang, "Parametric Nonlinear Programming for Linear Constraints," *Kyung-Pan Ron-Chong*, Vol. 2 (1977) pp. 135-147.
12. S. M. Robinson, "Perturbed Kuhn-Tucker Points and Rates of Convergence for a Class of Nonlinear-Programming Algorithms," *Mathematical Programming*, Vol. 7 (1974), pp. 1-16.
13) P. Wolfe, "Methods of Nonlinear Programming," in Graves, R. L. and Wolfe P. (eds.), *Recent Advances in Mathematical Programming*, McGraw-Hill, New York, 1963, pp. 67-86.