Journal
of The Franklin Institute
DEVOTED TO SCIENCE AND THE MECHANIC ARTS

Volume 288, Number 6 *December 1969*

Comparison of Some Conjugate Direction Procedures for Function Minimization

by H. W. SORENSON
Department of Aerospace and Mechanical Engineering Sciences
University of California at San Diego, La Jolla, California

ABSTRACT: *Iterative procedures that are quadratically convergent are seeing increasing application to many types of extremization problems. Three of these procedures are analyzed and their underlying theoretical similarities are pointed out. It is shown that the Davidon–Fletcher–Powell, conjugate gradient, and gradient partan methods provide identical results when applied to a quadratic cost function. This occurs because each procedure generates the same set of mutually conjugate vectors and is the basic reason for their quadratic convergence property. A slight modification to the conjugate gradient procedure proposed by Fletcher and Reeves is seen to yield a procedure that exhibits the principal property of the Davidon–Fletcher–Powell method when applied to nonquadratic cost functions. This modified conjugate gradient procedure involves a significantly smaller amount of computation than either the Davidon–Fletcher–Powell or gradient partan methods while retaining the basic properties influencing convergence of the Davidon–Fletcher–Powell procedure.*

1. Introduction

Numerous methods have been devised for finding the minimum (or maximum) of a prescribed function of n variables. For surveys of some of these techniques, one is referred to the papers by Dorn (**1**), Zoutendijk (**2**), Spang (**3**), and Lasdon and Waren (**4**). An important class of procedures that have been developed for finding the minimum of an unconstrained function are those that exhibit quadratic convergence through the construction of a sequence of mutually conjugate directions (**5**). In the following, three of these procedures are discussed and compared. The first of these, originally devised by Davidon (**6**) and later modified by Fletcher and Powell (**7**), receives the most attention. It is the most complicated of the three but has been found in several studies [e.g. (**8**), (**9**), (**10**)] to be especially effective. This method is considered in detail in Section 2 and certain apparent misconceptions regarding the reasons for its performance are discussed.

The second procedure that is considered is the conjugate gradient method of Hestenes and Stiefel (**13**). This method was applied to function minimization problems by Fletcher and Reeves (**8**) and it is discussed in Section 3. Its similarities to the Davidon–Fletcher–Powell (D–F–P) procedure are emphasized and a modified form for the equations is developed. Recently,* a numerical comparison of the conjugate gradient and Davidon methods has been presented by Kelley and Myers (**17**) and a theoretical comparison of the two has been given by Myers (**18**). References (**15**) and (**18**) point out that the D–F–P and conjugate gradient procedures are equivalent for quadratic cost functions when in both cases the initial step is taken in the gradient direction. It is pointed out that Fletcher and Reeves did not use the most appropriate form of the iterative equations. A different choice for the coefficient in the equation used to update the direction vector, although equivalent for a quadratic cost function, should yield performance that is essentially the same as that of the D–F–P procedure for nonquadratic cost functions.

Finally, the gradient partan method of Shah *et al.* (**12**, **5**) is discussed in Section 3.2. Many of these remarks apply directly to a method devised by Powell (**13**) which does not require the function gradient to be evaluated explicitly. The gradient partan method is seen to be equivalent to the conjugate gradient procedure in that the same system of mutually conjugate directions are determined. The former procedure has the disadvantage that twice as many steps must be taken to accomplish the construction.

In the remainder of this section, the basic notation and concepts are stated including a detailed review of the three computational methods and their principal properties. These results are used extensively in the subsequent discussion. For more detailed discussions and explanations, the reader is directed to the original sources.

1.1. *Minimization of a Quadratic Function*

The greater portion of the succeeding discussion will relate to the minimization of a quadratic cost function L having the form

$$L = L_0 + \langle \mathbf{c}, \mathbf{y}\rangle + \tfrac{1}{2}\langle \mathbf{y}, Q\mathbf{y}\rangle, \tag{1}$$

where $\mathbf{y}$ is an n-dimensional vector, $\mathbf{c}$, an n-dimensional vector of constants, and Q, an $(n \times n)$ symmetric positive-definite matrix.

The notation $\langle a, b\rangle$ is used to denote the inner product of two vectors. Since only function minimization is being considered explicitly, this is equivalent to writing

$$\langle \mathbf{x}, \mathbf{y}\rangle \triangleq \mathbf{x}^T \mathbf{y}.$$

Also, the $n \times n$ matrix formed from n-dimensional vectors $\mathbf{x}$ and $\mathbf{y}$ is written as

$$\mathbf{x}\rangle\langle \mathbf{y} = \mathbf{x}\mathbf{y}^T.$$

The dyadic notation is used because the procedures can be considered in

* The author wishes to thank the reviewer who brought these two papers to his attention.

Hilbert space. Antosiewicz and Rheinboldt (**14**) have discussed the extension for conjugate gradients and the D–F–P procedure has been considered in Hilbert space by Horwitz and Sarachik (**15**). The notation hopefully is suggestive of the more general considerations but avoids the details involved in the extension which do not appear to be central to this discussion.

The $\mathbf{y}$ that minimizes L, say $\mathbf{y}^*$, is determined by constructing a sequence $\{\mathbf{y}_i\}$ that converges to $\mathbf{y}^*$ as i becomes large. A procedure for constructing the $\mathbf{y}_i$ is said to be *quadratically convergent* (**5**) if $\mathbf{y}^*$ is determined after n gradient evaluations. The three procedures considered herein all exhibit this property.

Certain conventions and properties associated with the problem of minimizing Eq. 1 will be used frequently. The gradient of L evaluated at the point $\mathbf{y}_i$ is defined as

$$\begin{aligned} \mathbf{g}_i &\triangleq \left[\frac{\partial L}{\partial \mathbf{y}}(\mathbf{y}_i)\right]^T \\ &= Q\mathbf{y}_i + \mathbf{c}. \end{aligned} \tag{2}$$

For Eq. 1, a necessary and sufficient condition for $\mathbf{y}^*$ to minimize L is that

$$\mathbf{g}(\mathbf{y}^*) = 0, \tag{3}$$

so it follows immediately from Eq. 2 that

$$\mathbf{y}^* = -Q^{-1}\mathbf{c}. \tag{4}$$

It is also well known that if one has the point $\mathbf{y}_i$ and the gradient at the point $\mathbf{g}_i$, then the minimum $\mathbf{y}^*$ is found from

$$\mathbf{y}^* = \mathbf{y}_i - Q^{-1}\mathbf{g}_i. \tag{5}$$

From Eqs. 4 and 5 it is clear that if Q^{-1} is known, one can solve directly for the minimizing points. In most problems of practical interest the cost function L is not quadratic so that Eq. 1 is at best an approximation of the cost surface in the vicinity of the minimum. Even then, it is not likely that the approximating Q is known or, if it is, the difficulties involved in finding Q^{-1} are significant enough to warrant seeking alternative methods for finding $\mathbf{y}^*$. As will be discussed below, the conjugate direction methods do provide a means of computing Q^{-1}.

For a general cost function L, the second derivatives $(\partial^2 L/\partial \mathbf{y}^2)$ provide a linear relationship between infinitesimal changes in $\mathbf{y}$ and the resulting changes in the gradient $\mathbf{g}$ given by

$$d\mathbf{g} = \left(\frac{\partial^2 L}{\partial \mathbf{y}^2}\right) d\mathbf{y}. \tag{6}$$

For the quadratic cost function, this linear relationship is true for all changes in $\mathbf{y}$ and $\mathbf{g}$ with Q replacing $(\partial^2 L/\partial \mathbf{y}^2)$. Thus the gradient difference vector $\boldsymbol{\gamma}_i$ is defined as

$$\boldsymbol{\gamma}_i \triangleq \mathbf{g}_{i+1} - \mathbf{g}_i \tag{7}$$

and can be written, analogous to Eq. 6, as

$$\begin{aligned} \boldsymbol{\gamma}_i &= Q(\mathbf{y}_{i+1} - \mathbf{y}_i) \\ &\triangleq Q\Delta\mathbf{y}_i. \end{aligned} \tag{8}$$

The gradient difference vector plays an important role in subsequent discussions.

In all of the procedures that are discussed, the search for the minimizing point proceeds iteratively according to

$$\mathbf{y}_{i+1} = \mathbf{y}_i + k_i \mathbf{p}_i, \tag{9}$$

where the k_i is always chosen so that $\mathbf{y}_{i+1}$ minimizes L in the direction $\mathbf{p}_i$ starting from $\mathbf{y}_i$. The $\mathbf{y}_1$ and $\mathbf{p}_1$ are chosen arbitrarily. The manner in which the direction vector $\mathbf{p}_i$ is selected is defined by the particular procedure. The $\mathbf{p}_i$ are restricted to satisfy

$$\langle \mathbf{p}_i, \mathbf{g}_i \rangle < 0 \tag{10}$$

to insure that the cost function is reduced through the choice of k_i. The minimizing choice of k_i results in the satisfaction of the condition

$$\langle \mathbf{p}_i, \mathbf{g}_{i+1} \rangle = 0 \tag{11}$$

for all i. Equations 9, 10, and 11 apply to any cost function, not just Eq. 1. The procedure for determining the desired k_i is an important part of the actual minimization procedure. For this discussion, it shall be assumed that the minimization can be accomplished so that Eq. 11 is always valid. A detailed discussion of procedures for the one-dimensional or linear search can be found in the book by Wilde and Beightler (**5**).

1.2. *Minimizing Quadratic Functions and Solving Linear Equations*

The problem of determining the $\mathbf{y}$ that minimizes Eq. 1 is seen from Eqs. 2–3 to be equivalent to solving the linear equations

$$Q\mathbf{y} = -\mathbf{c}. \tag{12}$$

It was shown by Hestenes and Stiefel (**11**) that this can be accomplished by constructing a set of n mutually conjugate or, equivalently, Q-orthogonal vectors. To be precise, vectors $\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n$ are said to be *mutually conjugate* if for all i and j

$$\langle \mathbf{p}_i, Q\mathbf{p}_j \rangle = 0, \quad i \neq j. \tag{13}$$

This is clearly a generalization of the concept of orthogonal vectors.

The determination of n mutually conjugate vectors allows the solution of Eq. 12 to be constructed immediately. To see this, first note that it is easily shown that a set of n mutually conjugate vectors $\mathbf{p}_i$ in an n-dimensional space are linearly independent. Then, the $\mathbf{p}_i$ span the space and any vector (but particularly $\mathbf{y}^*$) can be written as a linear combination of the $\mathbf{p}_i$.

$$\mathbf{y}^* = \sum_{i=1}^{n} a_i \mathbf{p}_i \tag{14}$$

so that

$$Q\mathbf{y}^* = \sum_{i=1}^{n} a_i Q\mathbf{p}_i = -\mathbf{c}.$$

Then,

$$\begin{aligned}\langle \mathbf{p}_j, Q\mathbf{y}^* \rangle &= \sum_{i=1}^{n} a_i \langle \mathbf{p}_j, Q\mathbf{p}_i \rangle \\ &= a_j \langle \mathbf{p}_j, Q\mathbf{p}_j \rangle\end{aligned}$$

from the mutual conjugacy of the $\mathbf{p}_i$. The coefficients a_j are given by

$$a_j = -\frac{\langle \mathbf{p}_j, \mathbf{c}\rangle}{\langle \mathbf{p}_j, Q\mathbf{p}_j\rangle} \tag{15}$$

and $\mathbf{y}^*$ is

$$\mathbf{y}^* = -\sum_{i=1}^{n} \frac{\langle \mathbf{p}_i, \mathbf{c}\rangle}{\langle \mathbf{p}_i, Q\mathbf{p}_i\rangle}\,\mathbf{p}_i. \tag{16}$$

The inverse Q^{-1} can be constructed from the $\mathbf{p}_i$ by noting that Eq. 16 can be rewritten as

$$\mathbf{y}^* = -\sum_{i=1}^{n}\left[\frac{\mathbf{p}_i\rangle\langle\mathbf{p}_i}{\langle \mathbf{p}_i, Q\mathbf{p}_i\rangle}\right]\mathbf{c}$$

so that

$$Q^{-1} = \sum_{i=1}^{n} \frac{\mathbf{p}_i\rangle\langle\mathbf{p}_i}{\langle \mathbf{p}_i, Q\mathbf{p}_i\rangle}. \tag{17}$$

It is now seen that the determination of n mutually conjugate vectors provides the solution of a system of linear equations or, equivalently, provides the minimum of a quadratic function. Thus, if this set can be constructed as the n directions in the iterative procedure Eq. 9, it follows that the procedure is quadratically convergent. This is characteristic of all of the procedures discussed below.

1.3. *Review of the Computational Procedures*

1.3.1. *The D–F–P Method.* The direction vector $\mathbf{p}_i$ is generated from

$$\mathbf{p}_i = -H_i\,\mathbf{g}_i, \tag{18}$$

where H_i is defined by

$$H_{i+1} = H_i + \frac{\Delta\mathbf{y}_i\rangle\langle\Delta\mathbf{y}_i}{\langle\Delta\mathbf{y}_i, \boldsymbol{\gamma}_i\rangle} - \frac{H_i\boldsymbol{\gamma}_i\rangle\langle H_i\boldsymbol{\gamma}_i}{\langle\boldsymbol{\gamma}_i, H_i\boldsymbol{\gamma}_i\rangle}. \tag{19}$$

The initial deflection matrix H_1 is chosen arbitrarily and must be symmetric and commonly is required to be positive-definite. The implications of a nonnegative-definite choice are discussed below. In the absence of a better choice H_1 is defined as

$$H_1 \triangleq I.$$

For quadratic cost functions, this procedure has the following characteristics:

(1) H_i is symmetric and positive-definite if H_1 is positive-definite.
(2) The $\Delta\mathbf{y}_i$ (and $\mathbf{p}_i$) are mutually conjugate.
(3) $H_{n+1} = Q^{-1}$.

Property (2) implies that the procedure is quadratically convergent and (3) essentially follows from (2) as is shown in Scholium 1 of Section 2. Property (1) insures the stability of the procedure, even for nonquadratic cost functions.

1.3.2. *The Conjugate Gradient Method.* The method described by Fletcher and Reeves (**8**) is taken from the paper by Hestenes and Stiefel (**11**). In this method, the $\mathbf{p}_i$ is generated according to

$$\mathbf{p}_{i+1} = -\mathbf{g}_{i+1} + \beta_i \mathbf{p}_i, \tag{20}$$

where

$$\beta_i = \frac{\langle \mathbf{g}_{i+1}, \mathbf{g}_{i+1} \rangle}{\langle \mathbf{g}_i, \mathbf{g}_i \rangle}. \tag{21}$$

Hestenes and Stiefel suggested an equivalent form for β_i which is

$$\beta_i = \frac{\langle \mathbf{p}_i, Q\mathbf{g}_{i+1} \rangle}{\langle \mathbf{p}_i, Q\mathbf{p}_i \rangle}. \tag{22}$$

The initial point $\mathbf{y}_1$ and the initial direction $\mathbf{p}_1$ are specified arbitrarily. Clearly, $\mathbf{y}_1$ should be selected as the best estimate of the minimum and the $\mathbf{p}_1$ is commonly chosen as

$$\mathbf{p}_1 = -\mathbf{g}_1.$$

The conjugate gradient method has the following properties for quadratic cost functions:

(1) Gradient vectors at different points are orthogonal.

$$\langle \mathbf{g}_i, \mathbf{g}_j \rangle = 0, \quad i \neq j.$$

(2) The $\Delta \mathbf{y}_i$ (and $\mathbf{p}_i$) are mutually conjugate.

(3) The $\mathbf{p}_i$ are orthogonal to subsequent gradient vectors

$$\langle \mathbf{p}_i, \mathbf{g}_j \rangle = 0, \quad i < j.$$

(4) $\langle \mathbf{p}_i, \mathbf{g}_j \rangle = \langle \mathbf{g}_j, \mathbf{g}_j \rangle, \quad i \geqslant j.$

(5) $\langle \mathbf{g}_i, Q\mathbf{p}_i \rangle = \langle \mathbf{p}_i, Q\mathbf{p}_i \rangle.$

(6) $\langle \mathbf{g}_i, Q\mathbf{p}_j \rangle = 0, \quad i \neq j, \quad i \neq j+1.$

1.3.3. *The Gradient Partan Method.* The partan method does not involve the explicit construction of mutually conjugate direction vectors although vectors can be constructed from the direction vectors that are mutually conjugate. This property underlies the convergence properties of the partan method. Shah *et al.* (**12**) discuss a "general partan" and a "gradient partan". The latter has seen more application and the succeeding discussion is restricted to this case.

Starting Procedure: For the first step, let

$$\mathbf{p}_1 = -\mathbf{g}_1 \tag{23}$$

so that

$$\mathbf{y}_2 = \mathbf{y}_1 - k_1 \mathbf{g}_1.$$

Next, choose

$$\mathbf{p}_2 = -\mathbf{g}_2. \tag{24}$$

Then, the fourth point is generated by moving in a direction that is collinear with $(\mathbf{y}_3 - \mathbf{y}_1)$ so that

$$\mathbf{p}_3 = -(\mathbf{y}_3 - \mathbf{y}_1). \tag{25}$$

This is referred to as an acceleration step.

Continuing the Procedure: After determining $\mathbf{y}_4$, the procedure is continued by successively alternating gradient and acceleration steps. Thus,

$$\mathbf{p}_i = -\mathbf{g}_i \quad \text{for } i = 1 \text{ and } i = 2, 4, 6, \ldots, 2n-2 \tag{26a}$$

and

$$\mathbf{p}_i = -(\mathbf{y}_i - \mathbf{y}_{i-3}) \quad \text{for } i = 5, 7, 9, \ldots, 2n-1. \tag{26b}$$

This procedure will reach the minimum of an n-dimensional quadratic surface in no more than $2n$ steps. The $\mathbf{p}_i$ that are generated are *not* mutually conjugate but the following properties are true:

(1) The vectors $(\mathbf{y}_2 - \mathbf{y}_1)$, $(\mathbf{y}_4 - \mathbf{y}_2)$, $(\mathbf{y}_6 - \mathbf{y}_4), \ldots, (\mathbf{y}_{2n} - \mathbf{y}_{2n-2})$ are mutually conjugate.

(2) The points $\mathbf{y}_4, \mathbf{y}_6, \mathbf{y}_8, \ldots, \mathbf{y}_{2n}$ are the minimum for the spaces spanned respectively by $\mathbf{p}_1$ and $\mathbf{g}_2$; $\mathbf{p}_1$, $\mathbf{g}_2$, and $\mathbf{g}_4$; $\ldots, \mathbf{p}_1$, $\mathbf{g}_2$, $\mathbf{g}_4, \ldots, \mathbf{g}_{2n-2}$.

(3) The gradient vectors $\mathbf{g}_1, \mathbf{g}_2, \mathbf{g}_4, \ldots, \mathbf{g}_{2n-2}$ are orthogonal.

2. Some Observations about the D–F–P Procedure

In this section the behavior of the D–F–P procedure relative to quadratic cost functions is emphasized and the implications of these conclusions for nonquadratic cost functions are discussed. The D–F–P procedure has been stated in Section 1.3.1. The fact that the deflection matrix H_i converges to the Q^{-1} has been pointed out in several papers as an important reason for the apparently superior convergence properties of this procedure. This conclusion is reached by noting the similarity of Eq. 9 using Eqs. 18 to 5. A central purpose of this section is to point out that the property that $H_{n+1} = Q^{-1}$ has no direct influence upon the convergence properties and is only a natural consequence of the construction of mutually conjugate vectors.

Note, first, that for quadratic cost functions, the deflection matrix can be written in a more informative manner.

Lemma 1. The deflection matrix H_i as defined in Eq. 19 can be written as

$$H_{i+1} = P_{i+1} + C_{i+1}, \tag{27}$$

where

$$P_{i+1} = P_i - \frac{P_i \gamma_i\rangle\langle P_i \gamma_i}{\langle \gamma_i, P_i \gamma_i\rangle}, \tag{28}$$

$$C_{i+1} = C_i + \frac{\Delta \mathbf{y}_i\rangle\langle \Delta \mathbf{y}_i}{\langle \gamma_i, \Delta \mathbf{y}_i\rangle}, \tag{29}$$

and

$$P_1 \triangleq H_1, \quad C_1 \triangleq 0.$$

Proof. The result can be proved inductively. The matrix H_2 can be written as

$$\begin{aligned} H_2 &= H_1 - \frac{H_1 \gamma_1\rangle\langle H_1 \gamma_1}{\langle \gamma_1, H_1 \gamma_1\rangle} + \frac{\Delta \mathbf{y}_1\rangle\langle \Delta \mathbf{y}_1}{\langle \gamma_1, \Delta \mathbf{y}_1\rangle} \\ &= P_1 - \frac{P_1 \gamma_1\rangle\langle P_1 \gamma_1}{\langle \gamma_1, P_1 \gamma_1\rangle} + C_1 + \frac{\Delta \mathbf{y}_1\rangle\langle \Delta \mathbf{y}_1}{\langle \gamma_1, \Delta \mathbf{y}_1\rangle} \\ &\triangleq P_2 + C_2. \end{aligned}$$

Suppose that Eqs. 27–29 are true for some arbitrary k. Then, for $(k+1)$, the H_{k+1} is

$$\begin{aligned} H_{k+1} &= H_k - \frac{H_k \gamma_k\rangle\langle H_k \gamma_k}{\langle \gamma_k, H_k \gamma_k\rangle} + \frac{\Delta \mathbf{y}_k\rangle\langle \Delta \mathbf{y}_k}{\langle \gamma_k, \Delta y_k\rangle} \\ &= (P_k + C_k) - \frac{H_k \gamma_k\rangle\langle H_k \gamma_k}{\langle \gamma_k, H_k \gamma_k\rangle} + \frac{\Delta \mathbf{y}_k\rangle\langle \Delta \mathbf{y}_k}{\langle \gamma_k, \Delta \mathbf{y}_k\rangle} \\ &= P_k - \frac{H_k \gamma_k\rangle\langle H_k \gamma_k}{\langle \gamma_k, H_k \gamma_k\rangle} + C_{k+1}. \end{aligned}$$

Consider the vector $H_k \gamma_k$,

$$\begin{aligned} H_k \gamma_k &= (P_k + C_k)\gamma_k \\ &= (P_k + C_k) Q \Delta \mathbf{y}_k, \end{aligned}$$

where Eq. 8 has been used. But from the definition of C_k

$$\begin{aligned} C_k Q \Delta \mathbf{y}_k &= \left[C_{k-1} + \frac{\Delta \mathbf{y}_{k-1}\rangle\langle \Delta \mathbf{y}_{k-1}}{\langle \gamma_{k-1}, \Delta \mathbf{y}_{k-1}\rangle}\right] Q \Delta \mathbf{y}_k \\ &= \sum_{i=1}^{k-1} \frac{\Delta \mathbf{y}_i\rangle\langle \Delta \mathbf{y}_i}{\langle \gamma_i, \Delta \mathbf{y}_i\rangle} Q \Delta \mathbf{y}_k. \end{aligned}$$

However, the $\Delta \mathbf{y}_i$ are mutually conjugate so it follows that

$$C_k Q \Delta \mathbf{y}_k = 0$$

and

$$H_k \gamma_k = P_k \gamma_k.$$

This completes the proof since one now has

$$\begin{aligned} H_{k+1} &= P_k - \frac{P_k \gamma_k\rangle\langle P_k \gamma_k}{\langle \gamma_k, P_k \gamma_k\rangle} + C_{k+1} \\ &\triangleq P_{k+1} + C_{k+1}. \quad \text{Q.E.D.} \end{aligned}$$

Equations 27–29 provide for the quadratic case an entirely equivalent description of the deflection matrix H_{i+1}. The matrices P_{i+1} and C_{i+1} play an interesting and seemingly surprising role in the iterative procedure as will be seen below.

Scholium 1. The matrix C_{i+1} has the property after n steps that

$$C_{n+1} = Q^{-1}. \tag{30}$$

Proof. This follows immediately from the definition of C_{i+1}, the mutual conjugacy of the $\Delta \mathbf{y}_i$, and Eq. 17.

The term C_{i+1}, while causing H_{i+1} to be positive-definite, has *no* influence upon the determination of the direction vector $\mathbf{p}_{i+1}$ and for the quadratic case can be eliminated from the calculations.

Lemma 2. For a quadratic cost function, the direction vector $\mathbf{p}_i$ can be generated from

$$\mathbf{p}_i = -P_i \mathbf{g}_i, \tag{31}$$

where P_i is defined by Eq. 28 because

$$C_i \mathbf{g}_i = 0.$$

Proof. The proof follows inductively. The statement is true for $\mathbf{p}_1$ since C_1 is defined to be zero. Consider the second step

$$\begin{aligned}\mathbf{p}_2 &= -H_2 \mathbf{g}_2 \\ &= -(P_2 + C_2)\mathbf{g}_2.\end{aligned}$$

But

$$\begin{aligned}C_2 \mathbf{g}_2 &= \frac{\Delta\mathbf{y}_1\rangle\langle\Delta\mathbf{y}_1}{\langle\boldsymbol{\gamma}_1, \Delta\mathbf{y}_1\rangle}\mathbf{g}_2 \\ &= \frac{\langle\Delta\mathbf{y}_1, \mathbf{g}_2\rangle}{\langle\boldsymbol{\gamma}_1, \Delta\mathbf{y}_1\rangle}\Delta\mathbf{y}_1.\end{aligned}$$

But the choice of the step size parameter k_1 assures that

$$\langle\Delta\mathbf{y}_1, \mathbf{g}_2\rangle = 0$$

so

$$C_2 \mathbf{g}_2 = 0$$

and

$$\mathbf{p}_2 = -P_2 \mathbf{g}_2.$$

Assume that Eq. 3 is valid for $i = k-1$. Then,

$$\begin{aligned}\mathbf{p}_k &= -H_k \mathbf{g}_k \\ &= -(P_k + C_k)\mathbf{g}_k\end{aligned}$$

and

$$\begin{aligned}C_k \mathbf{g}_k &= \left(C_{k-1} + \frac{\Delta\mathbf{y}_{k-1}\rangle\langle\Delta\mathbf{y}_{k-1}}{\langle\boldsymbol{\gamma}_{k-1}, \Delta\mathbf{y}_{k-1}\rangle}\right)\mathbf{g}_k \\ &= C_{k-1}\mathbf{g}_k\end{aligned}$$

because of the choice of k_{k-1}. Note from Eqs. 7 and 8 that

$$\mathbf{g}_k = \mathbf{g}_{k-1} + Q\Delta\mathbf{y}_{k-1}$$

so that

$$\begin{aligned}C_{k-1}\mathbf{g}_k &= C_{k-1}\mathbf{g}_{k-1} + C_{k-1}Q\Delta\mathbf{y}_{k-1} \\ &= 0,\end{aligned}$$

where $C_{k-1}\mathbf{g}_{k-1}$ vanishes by hypothesis and $C_{k-1}Q\Delta\mathbf{y}_{k-1}$ vanishes because of the mutual conjugacy of the $\Delta\mathbf{y}_i$. Q.E.D.

This result shows that the fact that the $\mathbf{p}_i$ can be used to form Q^{-1} is not directly relevant to the problem of determining the minimum of a quadratic function. Rather, the procedure defines a method for constructing a set of mutually conjugate vectors and it is the properties of these vectors that determines the convergence properties. Thus the matrix P_i is the only matrix of importance in the quadratic case. As discussed below, the P_i is an orthogonal projection operator (**16**).

Lemma 3. Assume that $P_1 = I$. Then, P_i is an orthogonal projection operator and projects any vector onto a subspace orthogonal to the gradient difference vectors $\gamma_1, \gamma_2, \ldots, \gamma_{i-1}$.

Proof. For a linear operator to be an orthogonal projection, it must be self-adjoint and idempotent. Since P_1 is symmetric by assumption, it follows immediately from definition 28 that P_i is symmetric so it is self-adjoint. To prove that P_i is idempotent consider the matrix P_2.

$$P_2^2 = \left(P_1 - \frac{P_1\gamma_1\rangle\langle P_1\gamma_1}{\langle\gamma_1, P_1\gamma_1\rangle}\right)\left(P_1 - \frac{P_1\gamma_1\rangle\langle P_1\gamma_1}{\langle\gamma_1, P_1\gamma_1\rangle}\right)$$
$$= P_1^2 - \frac{P_1\gamma_1\rangle\langle P_1^2\gamma_1}{\langle\gamma_1, P_1\gamma_1\rangle} - \frac{P_1^2\gamma_1\rangle\langle P_1\gamma_1}{\langle\gamma_1, P_1\gamma_1\rangle}$$
$$+ \frac{P_1\gamma_1\rangle\langle P_1\gamma_1, P_1\gamma_1\rangle\langle P_1\gamma_1}{\langle\gamma_1, P_1\gamma_1\rangle}.$$

But

$$P_1^2 = I = P_1$$

so

$$P_2^2 = I - \frac{\gamma_1\rangle\langle\gamma_1}{\langle\gamma_1, \gamma_1\rangle} - \frac{\gamma_1\rangle\langle\gamma_1}{\langle\gamma_1, \gamma_1\rangle} + \frac{\gamma_1\rangle\langle\gamma_1, \gamma_1\rangle\langle\gamma_1}{\langle\gamma_1, \gamma_1\rangle^2}$$
$$= I - \frac{\gamma_1\rangle\langle\gamma_1}{\langle\gamma_1, \gamma_1\rangle}$$
$$\triangleq P_2.$$

Thus, P_2 is an orthogonal projection. Note also that

$$P_2\gamma_1 = \gamma_1 - \frac{\gamma_1\rangle\langle\gamma_1, \gamma_1\rangle}{\langle\gamma_1, \gamma_1\rangle} = 0$$

and that

$$P_2\mathbf{p}_i = \mathbf{p}_i - \frac{\gamma_1\rangle\langle\gamma_1, \mathbf{p}_i\rangle}{\langle\gamma_i, \gamma_i\rangle}$$
$$= \mathbf{p}_i, \quad i = 2, 3, \ldots, n,$$

since

$$\langle\gamma_1, \mathbf{p}_i\rangle = k_1\langle Q\mathbf{p}_1, \mathbf{p}_i\rangle = 0.$$

Assume that P_i is an orthogonal projection and consider P_{i+1}.

$$P_{i+1}^2 = \left(P_i - \frac{P_i\gamma_i\rangle\langle P_i\gamma_i}{\langle\gamma_i, P_i\gamma_i\rangle}\right)\left(P_i - \frac{P_i\gamma_i\rangle\langle P_i\gamma_i}{\langle\gamma_i, P_i\gamma_i\rangle}\right).$$

Since $P_i^2 = P_i$ by assumption, it follows that

$$P_{i+1}^2 = P_{i+1}.$$

Thus, P_{i+1} is an orthogonal projection. It can be written as

$$P_{i+1} = \left(I - \frac{P_i\gamma_i\rangle\langle\gamma_i}{\langle\gamma_i, P_i\gamma_i\rangle}\right)P_i$$
$$= \left(I - \frac{P_i\gamma_i\rangle\langle\gamma_i}{\langle\gamma_i, P_i\gamma_i\rangle}\right)\left(I - \frac{P_{i-1}\gamma_{i-1}\rangle\langle\gamma_{i-1}}{\langle\gamma_{i-1}, P_{i-1}\gamma_{i-1}\rangle}\right)P_{i-1}$$
$$= \left(I - \frac{P_i\gamma_i\rangle\langle\gamma_i}{\langle\gamma_i, P_i\gamma_i\rangle}\right)\cdots\left(I - \frac{P_j\gamma_j\rangle\langle\gamma_j}{\langle\gamma_j, P_j\gamma_j\rangle}\right)P_j, \quad j < i+1.$$

From this property, it follows that

$$P_{i+1}\boldsymbol{\gamma}_j = 0, \quad j = 1, 2, \ldots, i.$$

Also one sees that

$$P_{i+1}\mathbf{p}_j = \mathbf{p}_j, \quad j = i+1, i+2, \ldots, n,$$

since

$$\langle \boldsymbol{\gamma}_i, \mathbf{p}_j \rangle = 0, \quad i \neq j.$$

Thus, P_{i+1} projects any vector onto a space orthogonal to the gradient difference vectors $\boldsymbol{\gamma}_i$ ($i = 1, 2, \ldots, i$) and spanned by $\mathbf{p}_{i+1}, \mathbf{p}_{i+2}, \ldots, \mathbf{p}_k$. Q.E.D.

Scholium 2. The matrix P_{n+1} of Lemma 3 must be the zero matrix.

Proof. This conclusion is an immediate consequence of the fact that P_{n+1} must project vectors on to a subspace orthogonal to n linearly independent gradient difference vectors. This can only be accomplished by the zero matrix.

In Lemma 3, it is assumed that the P_1 is chosen as the identity matrix. As one might expect, the result is only modified slightly for other choices for P_1. Consider two cases.

Case 1

Let P_1 be an arbitrary, symmetric, positive-definite matrix. Then, the results of Lemma 3 are valid but in a transformed space. The transformation is determined by noting that from the restrictions, P_1 can be expressed as the product of a nonsingular matrix T and its transpose

$$P_1 = TT^T.$$

The matrix T defines a transformation to new variables $\boldsymbol{\eta}$ such that

$$\boldsymbol{\eta} = T^{-1}\mathbf{y}.$$

In terms of these variables, one sees that the iterations to determine the minimizing $\boldsymbol{\eta}$ are described by

$$\boldsymbol{\eta}_{i+1} = \boldsymbol{\eta}_i + k_i \boldsymbol{\pi}_i,$$

where

$$\boldsymbol{\pi}_i = -\Pi_i \mathbf{t}_i,$$

$\mathbf{t}_i$ is the gradient in the transformed space and is related to $\mathbf{g}_i$ by

$$\mathbf{t}_i = T^T \mathbf{g}_i.$$

The deflection matrix Π_i is defined by

$$\Pi_{i+1} = \Pi_i - \frac{\Pi_i \boldsymbol{\tau}_i \rangle \langle \Pi_i \boldsymbol{\tau}_i}{\langle \boldsymbol{\tau}_i, \Pi_i \boldsymbol{\tau}_i \rangle}$$

and the $\boldsymbol{\tau}_i$ is the transformed gradient difference vector.

$$\boldsymbol{\tau}_i = T^T \boldsymbol{\gamma}_i.$$

In these coordinates, the initial value Π_1 is

$$\Pi_1 = I.$$

Thus, the Π_i is a projection matrix in the $\boldsymbol{\eta}$-space rather than the y-space. The Π_i is related to P_i by

$$P_i = T\Pi_i T^T.$$

Case 2

Let P_1 be an orthogonal projection. In this case P_1 is not necessarily positive-definite. It projects any vector onto a subspace spanned by a set of vectors used to define P_1. This set could arise as a system of linear equality constraints. Then, the P_i would define a set of vectors that lead to the minimum of a quadratic cost function subject to a set of linear equality constraints. This aspect will not be treated further here. A more detailed discussion is given by Horwitz and Sarachik (**15**).

In the absence of any information about Q^{-1}, it is reasonable to choose P_1 equal to the identity matrix. At the other extreme, if Q^{-1} were known, then setting $P_1 = Q^{-1}$ would permit the minimum to be determined in one step. In intermediate cases, one would use a "best" estimate of Q^{-1} in order to locate the search in the proximity of the minimum as early as possible in the iterations. Thereafter, the procedure searches the entire space for the minimizing point.

For the application of the D–F–P procedure to nonquadratic cost functions, it is necessary to use the procedure as defined in Section 1.3.1. In this case the direction vectors cannot be considered to be mutually conjugate and the H_i no longer is related to a matrix Q^{-1}. Only in the vicinity of the minimum can the surface be approximated by a quadratic so that the properties of the procedure are not generally valid.

There is one basic property that carries over from the analysis for quadratic cost functions to the nonquadratic case. This property appears to be the basis for the superior performance observed for this procedure when applied to nonquadratic cost functions. As shall be discussed immediately below, the step $\mathbf{p}_i$ is always orthogonal to the gradient difference vector $\boldsymbol{\gamma}_{i-1}$. In the quadratic case, it is also orthogonal to $\boldsymbol{\gamma}_{i-2}, \boldsymbol{\gamma}_{i-3}, \ldots, \boldsymbol{\gamma}_1$. However, in the nonquadratic case, the addition of the term $\Delta\mathbf{y}_i\rangle\langle\Delta\mathbf{y}_i/\langle\Delta\mathbf{y}_i, \boldsymbol{\gamma}_i\rangle$ to H_{i+1} at each step prevents the orthogonality with the earlier gradient difference vectors. Thus, as the minimum is neared and as a quadratic approximation becomes more accurate, the influence of this term tends to vanish and the quadratic convergence properties begin to manifest themselves.

The step-size parameter k_i is chosen so that $\langle\Delta\mathbf{y}_i, \mathbf{g}_{i+1}\rangle = 0$. As a result, the step $\Delta\mathbf{y}_{i+1}$ can be written as

$$\begin{aligned}\Delta\mathbf{y}_{i+1} &= -k_{i+1} H_{i+1}\mathbf{g}_{i+1} \\ &= -k_{i+1}\left(H_i - \frac{H_i\boldsymbol{\gamma}_i\rangle\langle H_i\boldsymbol{\gamma}_i}{\langle\boldsymbol{\gamma}_i, H_i\boldsymbol{\gamma}_i\rangle} + \frac{\Delta\mathbf{y}_i\rangle\langle\Delta\mathbf{y}_i}{\langle\Delta\mathbf{y}_i, \boldsymbol{\gamma}_i\rangle}\right)\mathbf{g}_{i+1} \\ &= -k_{i+1}\left(H_i - \frac{H_i\boldsymbol{\gamma}_i\rangle\langle H_i\boldsymbol{\gamma}_i}{\langle\boldsymbol{\gamma}_i, H_i\boldsymbol{\gamma}_i\rangle}\right)\mathbf{g}_{i+1}. \qquad (32)\end{aligned}$$

Now, consider $\langle \gamma_i, \Delta y_{i+1} \rangle$. It follows from Eq. 32 that

$$\langle \gamma_i, \Delta y_{i+1} \rangle = -k_{i+1}\left(\gamma_i^T H_i - \frac{\langle \gamma_i, H_i \gamma_i \rangle \langle H_i \gamma_i}{\langle \gamma_i, H_i \gamma_i \rangle} \right) g_{i+1}$$
$$= 0. \tag{33}$$

Thus, Δy_{i+1} is orthogonal to γ_i for each i. The γ_j ($j<i$) are generally not orthogonal to Δy_{i+1} although it has been shown to be true for quadratic cost functions. Equation 33 and the quadratic properties near the minimum describe the basic characteristics of the D–F–P procedure when applied to nonquadratic cost functions.

3. Comparison of the Three Procedures

The conjugate gradient procedure devised by Hestenes and Stiefel (**11**) and applied to function minimization problems by Fletcher and Reeves (**8**) is compared in Section 3.1 with the D–F–P procedure. Although apparently dissimilar, these two methods are seen to be basically the same as has been noted previously (**15**, **18**). The basic difference arises because the two methods use different projection operators. Further, it is pointed out that Fletcher and Reeves did not use the most suitable form of the conjugate gradient equations in their investigation. In Section 3.2 the conjugate gradient and gradient partan (**12**) methods are compared with the conclusion that the latter represents a relatively cumbersome means of constructing the results given directly by conjugate gradients.

3.1. *The D–F–P and the Conjugate Gradient Methods*

The conjugate gradient procedure was reviewed in Section 1.3.2. This procedure will be redeveloped in a manner that more clearly exhibits its relation to the D–F–P procedure and which provides a basis for the discussion of Section 3.2.

Consider the following construction of a sequence of mutually conjugate vectors. Choose a point y_1 and consider the initial direction vector $\mathbf{p}_1$. To insure that the cost is decreased, the initial direction is restricted so that

$$\langle \mathbf{p}_1, g_1 \rangle < 0. \tag{34}$$

Clearly, one could choose

$$\mathbf{p}_1 = -g_1 \tag{35}$$

although the choice only must satisfy the constraint 34. To correspond with Section 1.3.2, assume that Eq. 35 is true. Then, y_2 is generated according to Eq. 9 where k_1 is chosen to satisfy Eq. 11.

At the new point choose the direction vector $\mathbf{p}_2$ to be orthogonal to the gradient difference vector γ_1. As pointed out in Section 2, this is a property of the direction vectors generated by the D–F–P procedure. To define $\mathbf{p}_2$ further, restrict it to be in the plane defined by $\mathbf{p}_1$ and g_2 and such that $\langle \mathbf{p}_2, g_2 \rangle < 0$. In general, the restriction

$$\langle \mathbf{p}_i, g_i \rangle < 0 \tag{36}$$

must be imposed at each point in order to insure that the procedure is stable. From these considerations, it follows that

$$\mathbf{p}_2 = \beta_0 \mathbf{g}_2 + \beta_1 \mathbf{p}_1. \tag{37}$$

Forming Eq. 36, one has

$$\langle \mathbf{g}_2, \mathbf{p}_2 \rangle = \beta_0 \langle \mathbf{g}_2, \mathbf{g}_2 \rangle + \beta_1 \langle \mathbf{g}_2, \mathbf{p}_1 \rangle = \beta_0 \langle \mathbf{g}_2, \mathbf{g}_2 \rangle.$$

The $\langle \mathbf{g}_2, \mathbf{p}_1 \rangle = 0$ because of the choice of k_1. Equation 36 is satisfied if β_0 is specified as a negative number and, in particular, if one selects

$$\beta_0 = -1.$$

With this value, Eq. 37 reduces to

$$\mathbf{p}_2 = -\mathbf{g}_2 + \beta_1 \mathbf{p}_1,$$

where β_1 is yet to be determined. It is used to satisfy the constraint that $\mathbf{p}_2$ and $\boldsymbol{\gamma}_1$ are orthogonal and is easily found to be

$$\beta_1 = \frac{\langle \boldsymbol{\gamma}_1, \mathbf{g}_2 \rangle}{\langle \boldsymbol{\gamma}_1, \mathbf{p}_1 \rangle}. \tag{38}$$

Since $\mathbf{p}_1 = -\mathbf{g}_1$, it follows that $\langle \mathbf{g}_1, \mathbf{g}_2 \rangle = 0$ so β_1 can be written as

$$\beta_1 = \frac{\langle \mathbf{g}_2, \mathbf{g}_2 \rangle}{\langle \mathbf{g}_1, \mathbf{g}_1 \rangle}.$$

Note that this form corresponds to Eq. 21.

The procedure used to define $\mathbf{p}_2$ is used for all subsequent steps. Then, it follows that the direction vectors are given by

$$\mathbf{p}_{i+1} = -\mathbf{g}_{i+1} + \beta_i \mathbf{p}_i, \tag{39}$$

where

$$\beta_i = \frac{\langle \boldsymbol{\gamma}_i, \mathbf{g}_{i+1} \rangle}{\langle \boldsymbol{\gamma}_i, \mathbf{p}_i \rangle}. \tag{40}$$

The procedure defined in this manner is stable (i.e. the cost is reduced at each step) and $\mathbf{p}_{i+1}$ is orthogonal to $\boldsymbol{\gamma}_i$. Thus, this procedure exhibits two of the most important properties of the D–F–P procedure.

Let us examine Eqs. 39 and 40 relative to the conjugate gradient procedure of Section 1.3.2 for a quadratic cost function. When Eq. 1 describes the cost function, the gradient difference vector is linearly related to $\Delta \mathbf{y}_i$ according to Eq. 8.

In this case, using Eqs. 8 and 9, one sees that Eq. 40 is identical to Eq. 22. This is the alternative form stated by Hestenes and Stiefel and which they found to be less sensitive to round-off errors than Eq. 21. Note, however, that Eq. 40 was derived *without* any assumptions regarding the cost function. Thus, this form for β_i will always cause the direction vector $\mathbf{p}_{i+1}$ to be orthogonal to $\boldsymbol{\gamma}_i$.

Conclusion. When the conjugate gradient procedure is modified so that β_i is computed according to Eq. 40 rather than Eq. 21 or Eq. 22, the steps exhibit the same orthogonality property that the D–F–P procedure exhibits.

That is, the direction vector $\mathbf{p}_{i+1}$ is always orthogonal to $\boldsymbol{\gamma}_i$. As was seen earlier, this really represents the principal characteristic of the D–F–P procedure when applied to nonquadratic cost functions. In both cases, the procedures construct mutually conjugate vectors for quadratic cost surfaces. These conclusions suggest that the conjugate gradient procedure as modified here and the D–F–P procedure should perform similarly for many, if not all, nonquadratic cost surfaces. (Numerical studies that will not be presented here bear out this conclusion.) This is an important conjecture, if true, since the conjugate gradient procedure is computationally simpler.

The modified conjugate gradient procedure and the D–F–P procedure are now seen to have the same basic characteristic. The procedures will next be shown to involve projection operators in the quadratic case. As shown in Section 2, the D–F–P procedure constructs a sequence of mutually conjugate vectors through the use of an orthogonal projection operator. It will now be demonstrated that the modified conjugate gradient procedure implicitly constructs the conjugate vectors with an operator that is a projection but not an orthogonal projection.

Consider the first step. Write $\mathbf{p}_1$ as

$$\mathbf{p}_1 = -I\mathbf{g}_1. \tag{41}$$

If a different choice than $-\mathbf{g}_1$ is made for $\mathbf{p}_1$, the matrix would be a symmetric positive-definite operator G_1 instead of the identity matrix indicated in Eq. 41. This is the equivalent to the choice of a matrix other than the identity matrix for H_1 in Eq. 19. For the second step, $\mathbf{p}_2$ is defined by

$$\begin{aligned}\mathbf{p}_2 &= -\mathbf{g}_2 + \frac{\langle\boldsymbol{\gamma}_1, \mathbf{g}_2\rangle}{\langle\boldsymbol{\gamma}_1, \mathbf{p}_1\rangle}\mathbf{p}_1 \\ &= -\mathbf{g}_2 + \frac{\mathbf{p}_1\rangle\langle\boldsymbol{\gamma}_1}{\langle\boldsymbol{\gamma}_1, \mathbf{p}_1\rangle}\mathbf{g}_2 \\ &= -\left(I - \frac{\mathbf{p}_1\rangle\langle\boldsymbol{\gamma}_1}{\langle\boldsymbol{\gamma}_1, \mathbf{p}_1\rangle}\right)\mathbf{g}_2 \\ &= -G_2\mathbf{g}_2.\end{aligned}$$

The matrix G_2 is a projection since

$$\begin{aligned}G_2^2 &= \left(I - \frac{\mathbf{p}_1\rangle\langle\boldsymbol{\gamma}_1}{\langle\boldsymbol{\gamma}_1, \mathbf{p}_1\rangle}\right)\left(I - \frac{\mathbf{p}_1\rangle\langle\boldsymbol{\gamma}_1}{\langle\boldsymbol{\gamma}_1, \mathbf{p}_1\rangle}\right) = I - \frac{\mathbf{p}_1\rangle\langle\boldsymbol{\gamma}_1}{\langle\boldsymbol{\gamma}_1, \mathbf{p}_1\rangle} \\ &= G_2.\end{aligned}$$

It is not symmetric so G_2 is not an orthogonal projection.

Observe also that

$$G_2\mathbf{p}_1 = 0$$

and that

$$G_2\mathbf{p}_i = \mathbf{p}_i, \quad i = 2, 3, \ldots, n,$$

because of the mutual conjugacy of the $\mathbf{p}_j$. Thus, G_2 projects any vector on to a space spanned by $\mathbf{p}_2, \mathbf{p}_3, \ldots, \mathbf{p}_n$ along the space spanned by $\mathbf{p}_1$. It is also true that

$$\boldsymbol{\gamma}_1^T G_2 = 0$$

and

$$\boldsymbol{\gamma}_i^T G_2 = \boldsymbol{\gamma}_i^T, \quad i = 2, 3, \ldots, n.$$

Consider $\mathbf{p}_3$. It can be written as

$$\mathbf{p}_3 = -\left(I - \frac{\mathbf{p}_2\rangle\langle\boldsymbol{\gamma}_2}{\langle\boldsymbol{\gamma}_2, \mathbf{p}_2\rangle}\right)\mathbf{g}_3.$$

But it is known that $\langle\boldsymbol{\gamma}_1, \mathbf{g}_3\rangle = 0$ so $\mathbf{p}_3$ can be modified to be

$$\begin{aligned}\mathbf{p}_3 &= -\left(I - \frac{\mathbf{p}_1\rangle\langle\boldsymbol{\gamma}_1}{\langle\boldsymbol{\gamma}_1, \mathbf{p}_1\rangle} - \frac{\mathbf{p}_2\rangle\langle\boldsymbol{\gamma}_2}{\langle\boldsymbol{\gamma}_2, \mathbf{p}_2\rangle}\right)\mathbf{g}_3 \\ &= -\left(G_2 - \frac{\mathbf{p}_2\rangle\langle\boldsymbol{\gamma}_2}{\langle\boldsymbol{\gamma}_2, \mathbf{p}_2\rangle}\right)\mathbf{g}_3 \\ &\triangleq -G_3\,\mathbf{g}_3.\end{aligned}$$

It follows that G_3 is a projection operator because $G_3^2 = G_3$. Again, G_3 is not symmetric so it is not an orthogonal projection. One shows from the properties of the $\mathbf{p}_i$ that

$$G_3\,\mathbf{p}_i = \begin{cases} 0, & i = 1, 2, \\ \mathbf{p}_i, & i = 3, 4, \ldots, n, \end{cases}$$

and

$$\boldsymbol{\gamma}_i^T G_3 = \begin{cases} 0, & i = 1, 2, \\ \boldsymbol{\gamma}_i^T, & i = 3, 4, \ldots, n. \end{cases}$$

These results generalize directly so that the $\mathbf{p}_i$ are generated according to

$$\mathbf{p}_i = -G_i\,\mathbf{g}_i, \tag{42}$$

where

$$G_{i+1} = G_i - \frac{G_i\,\mathbf{g}_i\rangle\langle\boldsymbol{\gamma}_i}{\langle\boldsymbol{\gamma}_i, G_i\,\mathbf{g}_i\rangle}, \tag{43}$$

where G_i projects any vector onto the space spanned by $\mathbf{p}_i, \mathbf{p}_{i+1}, \ldots, \mathbf{p}_n$ along the space spanned by $\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_{i+1}$. An immediate consequence is that G_{n+1} must vanish.

The matrix G_i is similar to the P_i of the D–F–P procedure in that it is a projection operator but unlike it in the sense that it is not an orthogonal projection. Thus, the two procedures are very similar but appear to generate different sets of conjugate directions. As seen immediately below, the directions are collinear and differ only in magnitude.

Lemma 4. If the initial direction $\mathbf{p}_1$ is the same for the D–F–P and conjugate gradient procedures and is chosen as $-\mathbf{g}_i$, then $P_i\,\mathbf{g}_i$ and $G_i\,\mathbf{g}_i$ are collinear for all i.

Proof. It shall be shown by induction that

$$G_i P_i = P_i \quad \text{and} \quad P_i G_i = G_i. \tag{44}$$

First, consider H_2 and G_2. By hypothesis,

$$P_1 = I = G_1$$

so

$$\begin{aligned} H_2 G_2 &= \left(I - \frac{\gamma_1\rangle\langle\gamma_1}{\langle\gamma_1, \gamma_1\rangle}\right)\left(I - \frac{\mathbf{p}_1\rangle\langle\gamma_1}{\langle\gamma_1, \mathbf{p}_1\rangle}\right) \\ &= I - \frac{\gamma_1\rangle\langle\gamma_1}{\langle\gamma_1, \gamma_1\rangle} - \frac{\mathbf{p}_1\rangle\langle\gamma_1}{\langle\gamma_1, \mathbf{p}_1\rangle} + \frac{\gamma_1\rangle\langle\gamma_1, \mathbf{p}_1\rangle\langle\gamma_1}{\langle\gamma_1, \gamma_1\rangle\langle\gamma_1, \mathbf{p}_1\rangle} \\ &= I - \frac{\mathbf{p}_1\rangle\langle\gamma_1}{\langle\gamma_1, \mathbf{p}_1\rangle} \\ &= G_2. \end{aligned}$$

Similarly

$$\begin{aligned} G_2 H_2 &= \left(I - \frac{\mathbf{p}_1\rangle\langle\gamma_1}{\langle\gamma_1, \mathbf{p}_1\rangle}\right)\left(I - \frac{\gamma_1\rangle\langle\gamma_1}{\langle\gamma_1, \gamma_1\rangle}\right). \\ &= H_2. \end{aligned}$$

Suppose that for some k

$$P_k G_k = G_k \quad \text{and} \quad G_k P_k = P_k$$

then

$$\begin{aligned} G_{k+1} P_{k+1} &= \left(G_k - \frac{\mathbf{p}_k\rangle\langle\gamma_k}{\langle\gamma_k, \mathbf{p}_k\rangle}\right)\left(P_k - \frac{P_k\gamma_k\rangle\langle P_k\gamma_k}{\langle\gamma_k, P_k\gamma_k\rangle}\right) \\ &= G_k P_k - \frac{G_k P_k\gamma_k\rangle\langle P_k\gamma_k}{\langle\gamma_k, P_k\gamma_k\rangle} - \frac{\mathbf{p}_k\rangle\langle P_k\gamma_k}{\langle\gamma_k, \mathbf{p}_k\rangle} \\ &\quad + \frac{\mathbf{p}_k\rangle\langle\gamma_k, P_k\gamma_k\rangle\langle P_k\gamma_k}{\langle\gamma_k, \mathbf{p}_k\rangle\langle\gamma_k, P_k\gamma_k\rangle} \\ &= P_{k+1} \end{aligned}$$

since $G_k P_k = P_k$ by hypothesis.

Similarly, it follows that

$$\begin{aligned} P_{k+1} G_{k+1} &= \left(P_k - \frac{P_k\gamma_k\rangle\langle P_k\gamma_k}{\langle\gamma_k, P_k\gamma_k\rangle}\right)\left(G_k - \frac{\mathbf{p}_k\rangle\langle\gamma_k}{\langle\gamma_k, \mathbf{p}_k\rangle}\right) \\ &= G_{k+1}. \end{aligned}$$

This shows that Eq. 44 is valid and the collinearity of $P_i \mathbf{g}_i$ and $G_i \mathbf{g}_i$ is an immediate consequence. Q.E.D.

The magnitude of the direction vectors generated by the two procedures is of no importance since the step-size k_i is always chosen to go to the minimum in the given direction. As a result of Lemma 4, one sees that if the first point and initial direction are the same for each procedure, then corresponding points in the search for the minimum of a quadratic cost function are identical. While this is true for quadratic cost functions, the algorithms do yield different results for nonquadratic cost functions even though they share basic properties.

3.2. *Gradient Partan and Conjugate Gradient Methods*

In the gradient partan procedure, one alternately takes a gradient step and an acceleration step. Points in the procedure with an even-numbered subscript (i.e. $\mathbf{y}_2, \mathbf{y}_4, \ldots, \mathbf{y}_{2n-2}$) are associated with gradient steps described by Eq. 26a. At odd-numbered points (i.e. $\mathbf{y}_3, \mathbf{y}_5, \ldots, \mathbf{y}_{2n-1}$) an acceleration step is taken as defined by Eq. 26b. In the following, it will be shown that the gradient step is unnecessary and that the procedure is equivalent to the conjugate gradient procedure for quadratic cost functions.

Consider the first step to be in the gradient direction. Then,

$$\begin{aligned}\mathbf{y}_2 &= \mathbf{y}_1 + k_1 \mathbf{p}_1 \\ &= \mathbf{y}_1 - k_1 \mathbf{g}_1.\end{aligned}$$

In the gradient partan, the next step is also in the gradient direction so that

$$\begin{aligned}\mathbf{y}_3 &= \mathbf{y}_2 + k_2 \mathbf{p}_2 \\ &= \mathbf{y}_2 - k_2 \mathbf{g}_2.\end{aligned}$$

The acceleration step is

$$\mathbf{y}_4 = \mathbf{y}_3 - k_3(\mathbf{y}_3 - \mathbf{y}_1)$$

and $\mathbf{y}_4$ is the minimum of the cost function in the plane spanned by $\mathbf{g}_1$ and $\mathbf{g}_2$. This is manifested by the orthogonality of the gradient $\mathbf{g}_4$ to $\mathbf{g}_1$ and $\mathbf{g}_2$. Thus, the projection of the gradient $\mathbf{g}_4$ onto this plane is zero so $\mathbf{y}_4$ is the minimum in this plane.

But the point $\mathbf{y}_3$ of the conjugate gradient procedure has the same properties. Recall from Eq. 37 that $\mathbf{p}_2$ is restricted to the plane defined by $\mathbf{g}_2$ and $\mathbf{g}_1$ so that $\mathbf{y}_3$ must be in this plane. A property of the resulting step is that $\mathbf{g}_3$ is orthogonal to $\mathbf{g}_2$ and $\mathbf{g}_1$ so $\mathbf{y}_3$ is the minimum in this plane. This implies that $\mathbf{y}_3$ of the conjugate gradient procedure and the $\mathbf{y}_4$ of the gradient partan are identical.

The gradient partan procedure requires a step away from $\mathbf{y}_4$ in the gradient direction $(-\mathbf{g}_4)$

$$\mathbf{y}_5 = \mathbf{y}_4 - k_4 \mathbf{g}_4$$

and then an acceleration step

$$\mathbf{y}_6 = \mathbf{y}_5 - k_5(\mathbf{y}_5 - \mathbf{y}_2)$$

reaches the minimum of the cost function in the subspace spanned by $\mathbf{g}_1$, $\mathbf{g}_2$, and $\mathbf{g}_4$. This is manifested by the orthogonality of $\mathbf{g}_6$ with $\mathbf{g}_4$, $\mathbf{g}_2$, and $\mathbf{g}_1$.

The point $\mathbf{y}_4$ generated by the conjugate gradient algorithm has the property that it is in the plane spanned by $\mathbf{g}_3$ and $\mathbf{p}_2$. But this is the same plane that contains $\mathbf{y}_6$ of the gradient partan since it is in the plane formed by $\mathbf{g}_4$ and $(\mathbf{y}_4 - \mathbf{y}_2)$. These vectors are identical with the $\mathbf{g}_3$ and $\mathbf{p}_2$ of the conjugate gradient procedure. The $\mathbf{y}_4$ of the conjugate gradient procedure is the minimum of the cost function spanned by $\mathbf{g}_1$, $\mathbf{g}_2$, and $\mathbf{g}_3$ since $\mathbf{g}_4$ is orthogonal to all these gradient vectors. Thus, the $\mathbf{y}_6$ of the gradient partan and $\mathbf{y}_4$ of conjugate gradients is the same point.

These arguments are easily extended to any point with the conclusion that a point $\mathbf{y}_k$ of the conjugate gradient procedure is identical to the point $\mathbf{y}_{2k-2}$ of the gradient partan ($k = 2, 3, 4, \ldots, n+1$). Note that the gradients at these points are orthogonal to the gradient vectors at preceding points in these sequences. This equivalence of the two procedures casts considerable doubt on the usefulness of the partan procedure since it requires twice as many one-dimensional minimizations as the conjugate directions method. It is difficult to see how this extra computation could be beneficial since the conjugate gradient procedure is so simple to implement.

4. Conclusions

The Davidon–Fletcher–Powell procedure has been shown to be a method for constructing a set of conjugate directions (**11**). The fact that the deflection matrix H_i converges to Q^{-1} (i.e. Q is the matrix of the quadratic cost function) after n steps does not directly influence the iterative procedure. In fact, only a portion of the deflection matrix influences the determination of the direction vector at each step. This portion, designated as P_i in Section 2, is an orthogonal projection operator and vanishes after n steps.

The D–F–P procedure is effective in the nonquadratic case because it always computes a direction p_{i+1} so that it is orthogonal to the gradient difference vector $\boldsymbol{\gamma}_i$. This policy generally causes steps to be taken that explore new regions of the minimization space. The addition of the term $\Delta\mathbf{y}_i\rangle\langle\Delta\mathbf{y}_i/\langle\Delta\mathbf{y}_i,\boldsymbol{\gamma}_i\rangle$ to the deflection matrix H_{i+1} at each step serves to relax the restriction that $\mathbf{p}_{i+1}$ be orthogonal to $\boldsymbol{\gamma}_{i-1}, \boldsymbol{\gamma}_{i-2}, \ldots, \boldsymbol{\gamma}_1$ that would result otherwise. This insures that steps are not restricted to unnecessarily small subspaces.

The basic conjugate gradient method proposed by Fletcher and Reeves appears to be in an inappropriate form for application to nonquadratic cost functions. This is the cause of their needing to restart the procedure after $(n+1)$ steps in order to prevent the direction vectors from becoming too similar, thereby reducing the convergence rate. A different formula for the β_i [see Eqs. 39 and 40] is proposed which will cause the direction vector $\mathbf{p}_{i+1}$ to be orthogonal to $\boldsymbol{\gamma}_i$. Thus, the modified conjugate gradient procedure exhibits the outstanding characteristics of the D–F–P procedure.

The modified conjugate gradient procedure implicitly involves a projection operator in the way in which it constructs the set of mutually conjugate direction vectors. In this sense it is similar to the D–F–P procedure. The procedures are different because the conjugate gradient operator is not an orthogonal projection whereas the D–F–P procedure does use an orthogonal projection operator. However, the two procedures generate collinear direction vectors so that identical results are obtained when the procedures are applied to a quadratic cost function.

The theoretical similarities of the modified conjugate gradient and the D–F–P procedures lead to the conjecture that the two procedures will yield similar performance when applied to nonquadratic cost functions. This appears to be borne out by numerical studies that will be reported at another time. If the performance is actually comparable, then the simpler form of the modified conjugate gradient procedure would seem to make it the more desirable of the two.

The gradient partan procedure has been seen to be theoretically equivalent to the conjugate gradient procedure for quadratic cost functions. However, it requires twice as many linear searches as the latter. Since the linear search generally involves a significant amount of computation in any application, the gradient partan procedure would seem to be generally inferior to the D–F–P and modified conjugate gradient procedures.

Acknowledgements

The author wishes to acknowledge many interesting and stimulating conversations with J. E. Prussing regarding these matters and to thank Dr. Prussing for his comments and suggestions.

References

(1) W. S. Dorn, "Nonlinear Programming—A Survey", *Man. Sci.*, Vol. 9, pp. 171–208, Jan. 1963.
(2) G. Zoutendijk, "Nonlinear Programming; A Numerical Survey", *J. SIAM Control,* Vol. 4, No. 1, pp. 194–210, 1966.
(3) H. A. Spang, "A Review of Minimization Techniques for Nonlinear Functions", *SIAM Review,* Vol. 4, No. 4, pp. 343–365, Oct. 1962.
(4) L. S. Lasdon and A. D. Waren, "Mathematical Programming for Optimal Design", *Electro-Technology,* pp. 55–70, Nov. 1967.
(5) D. J. Wilde and C. S. Beightler, "Foundations of Optimization", Englewood Cliffs, N.J., Prentice-Hall, 1967.
(6) W. C. Davidon, "Variable Metric Method for Minimization", Argonne National Laboratory Report ANL-5990, Dec. 1959.
(7) R. Fletcher and M. J. D. Powell, "A Rapidly Convergent Descent Method for Minimization", *Computer J.*, Vol. 6, No. 2, pp. 163–168, July 1963.
(8) R. Fletcher and C. M. Reeves, "Function Minimization by Conjugate Gradients", *Computer J.*, Vol. 7, pp. 149–154, 1964.
(9) M. J. Box, "A Comparison of Several Current Optimization Methods", *Computer J.*, Vol. 9, pp. 67–77, 1966.
(10) B. H. Paiewonsky and P. J. Woodrow, "A Study of Time-optimal Rendezvous in Three Dimensions", Vol. I, Air Force Flight Dynamics Laboratory Report AFFDL-TR-65-20, Jan. 1965.
(11) M. R. Hestenes and E. Stiefel, "Methods of Conjugate Gradients for Solving Linear Systems", *J. Res. N.B.S.*, Vol. 49, pp. 409–436.
(12) B. V. Shah, R. J. Buehler, and O. Kempthorne, "Some Algorithms for Minimizing a Function of Several Variables", *J. SIAM,* Vol. 12, No. 1, pp. 74–92, March 1964.
(13) M. J. D. Powell, "An Efficient Method for Finding the Minimum of a Function of Several Variables without Calculating Derivations", *Computer J.*, Vol. 7, pp. 155–162, 1964.

(**14**) H. A. Antosiewicz and W. C. Rheinboldt, "Numerical Analysis and Functional Analysis", Chap. 14 in *Survey of Numerical Analysis* (edited by J. Todd), New York, McGraw-Hill, 1962.

(**15**) L. B. Horwitz and P. E. Sarachik, "Davidon's Method in Hilbert Space", *J. SIAM*, Vol. 16, No. 4, pp. 676–695, July 1968.

(**16**) L. A. Zadeh and C. A. Desoer, "Linear System Theory: The State Space Approach", New York, McGraw-Hill, 1963.

(**17**) H. J. Kelley and G. E. Myers, "Conjugate Direction Methods for Parameter Optimization", *18th International Astronautical Congress*, Belgrade, Yugoslavia, 1967.

(**18**) G. E. Myers, "Properties of the Conjugate-gradient and Davidon Methods", *J. Opt. Theory Appl.*, Vol. 2, No. 4, pp. 209–219, July 1968.