Variable-Metric Algorithm Employing Linear and Quadratic Penalties

Henry J. Kelley* and Leon Lefton†
Analytical Mechanics Associates, Inc., Jericho, New York
and
Ivan L. Johnson, Jr.‡
NASA Johnson Space Center, Houston, Texas

A variable-metric algorithm is described which makes use of both linear and quadratic penalty terms for handling nonlinear constraints, and employs both projection and penalty features. Quadratic penalty coefficients are adjusted in a process that attempts to maintain a positive-definite matrix of second partial derivatives of the function, including penalty terms, without generating the large positive eigenvalues that traditionally attend the use of quadratic penalties, which cause zigzagging and slowed convergence. The schemes contemplated make use of inferred second-order properties, not only in terms of the variable metric of DFP (or its relatives), but by estimation of second directional derivatives, by fitting cubics to various functions along directions of search. Some experiments are described with a simple constrained-minimum problem contrived to offer difficulties with methods that use only linear penalties, hence taxing the quadratic-penalty-adjustment procedure.

Introduction

THE arrival of variable-metric parameter optimization, the Davidon-Fletcher-Powell algorithm[1] and its relatives, literally revolutionized numerical optimization in the sixties. Even variational problems, crammed into the mold by sometimes awkward parameterizations, were treated handily by DFP in completion of various sophisticated continuous-control algorithms. The key to success is the superficially first-order character of the technique—only first partial derivatives need be generated explicitly—together with speed and the sureness of convergence accomplished by inference of second-order properties.

However, in most variable-metric applications work, the constraints are treated by the quadratic penalty function, a primitive device well-known to affect convergence rate adversely and to magnify numerical errors. The combination of penalty function and variable metric was explored in a 1966 paper,[2] which included various auxiliary devices to ameliorate the adverse effects of penalty-function approximation. This particular computational procedure has turned out to be a reliable work-horse, and currently is in fairly wide use in day-to-day applications work.

Efforts at adapting variable metrics to the standard alternative scheme for treating constraints, gradient projection, proved straightforward, and immediately tractable only in the case of linear constraints.[3] Variable-metric projection schemes, making selective use of what amount to linear penalty functions, eventually were developed for the case of nonlinear constraints, and proved workable in limited tests.[4,5] This class of variable-metric scheme has seen only limited use in complex applications, however, and is not yet highly developed.

One suspects that current-state-of-the-art schemes are costly and slow in comparison to what is possible. The focus in the following is upon that class of problems in which auxiliary vector-matrix computations are inexpensive in comparison with the generation of function samples and gradients, as, for example, in aerospace trajectory-shaping problems. Use is made both of penalty and projection ideas and various other features of the algorithms of Refs. 1-10; the adjustment of penalty coefficients represents the main innovation, and the bulk of the discussion will be devoted to this.

Variable-Metric Gradient Process with Linear-Plus-Quadratic Penalties

Consider an alternative to the problem of minimizing a function $f(x)$ (x an n-vector) subject to an m-vector equality constraint $g(x)=0$; namely, the minimization of the function $\tilde{f}$, given by

$$\tilde{f}=f+g\lambda+\tfrac{1}{2}gKg^T \tag{1}$$

which contains both linear and quadratic penalty terms. With $\lambda=0$ and the elements of the diagonal matrix $k_{ii}\gg 0$, one has the quadratic penalty scheme[6]; large k values are needed in this approach, not only to insure that the function adopted for minimization actually possesses a minimum near the constrained minimum that is sought, but also to render the magnitudes of the constraint violations $|g_i|$ small at the minimum.

Hestenes' Method of Multipliers[7] employs both linear and quadratic penalty terms, with the quadratic terms viewed as primary; the linear terms, missing in a first major iteration, are introduced as auxiliarites to reduce constraint violations and permit use of somewhat lower quadratic penalty coefficients. The λ vector for each major iteration, which consists of a minimization of $\tilde{f}$, is taken in this algorithm as the value of $\lambda+Kg$ at the end of the preceding major iteration. Of course, any minimization algorithm can be used for the major iterations but, for such unconstrained problems, DFP and its relatives are highly competitive.

The algorithm examined in the following makes use of the preceding form of $\tilde{f}$, including both linear and quadratic penalty terms, However, the viewpoint taken is different from the Method of Multipliers; namely, that the linear terms are primary, and the quadratic ones are supplementary and missing whenever advisable. The λ-vector components will be determined as the projection values every few cycles. The K diagonal elements will be chosen generally, so as to provide the second partial derivative matrix $\tilde{f}_{xx}$ with positive

Received Feb. 27, 1975; revision received Sept. 25, 1975; also presented at the AAS/AIAA Astrodynamics Conference, Nassau, Bahamas, July 28-30, 1975. Supported under Contract NAS 9-12516 with NASA Johnson Space Center, Houston, Tex.

Index categories: Aircraft Performance and Mission Analysis; Guidance and Control.

*Vice President. Associate Fellow AIAA.

†Senior Programer/Analyst.

‡Aerospace Technologist.

definiteness, but without the excessive margin traditionally furnished by large quadratic penalty terms, which hinders convergence. The linear penalty terms provide the means of reducing constraint violations to zero in case the constraints are compatible, i.e., the surfaces defined by $g_i = 0$ have an intersection.

The two algorithms examined employ DFP[1] and its batch-processor DFP modification[8] applied to $\bar{f}$ for major iterations. They bear a resemblance to the Method of Multipliers, differing from it in the determination of λ and k values. They are similar for the first few cycles, during which the diagonal K elements are assigned "moderately large" positive values in the quadratic-penalty-function tradition. The major iteration proceeds by variable metric for n cycles, however, rather than all the way to a minimum.

The general idea of the quadratic-penalty-coefficient selection scheme is control of the eigenvalues of the second-partial-derivative matrix

$$\bar{f}_{xx} = f_{xx} + \sum_{j=1}^{m} \lambda_j g_{j_{xx}} + \sum_{j=1}^{m} k_j (g_j g_{j_{xx}} + g_{j_{xx}} g_{j_x}^T) \tag{2}$$

to produce positive-definiteness and a largest eigenvalue not much exceeding the largest eigenvalue of $f_{xx} + g_{xx}\lambda$ [illegal notation, but suggestive shorthand for the first two terms of Eq. (2)]. One would like this not locally, with λ the projection value, but at the constrained minimum where the projection λ conicides with the Lagrange multiplier vector; however, it would be difficult and expensive enough to calculate the *local* second partials and the largest eigenvalue, and so a less direct and more appropriate approach is taken. The following scheme proposed takes advantage of the fact that there usualy will be a large range of values for the k_i meeting the requirements, the lower limit determined by loss of definiteness and/or excessive constraint violations, and the upper limit related to the largest eigenvalue of $f_{xx} + g_{xx}\lambda$.

During the n cycles of each "batch," or major iteration, second directional derivatives along the n directions of search are estimated for the function $f^* \equiv f + g\lambda^*$, where λ^* is given by

$$\lambda^* = -(g_x^T H_0 g)^{-1} g_x^T H_0 f_x \tag{3}$$

as the gradient-projection value; this varies from cycle to cycle. H_0 is a full-rank $n \times n$ matrix, fixed during a batch. At the constrained minimum sought, the value given by Eq. (3) is equal to the Lagrange multipler vector for stationary $f + g\lambda$; it is independent of the metric H_0 when evaluated at the constrained minimum.

For a step determined by the modified DFP algorithm as

$$\Delta x_i = x_{i+1} - x_i \quad i = 1,\text{---},n \tag{4}$$

the first and second derivatives in the direction are given by

$$f^{*\prime} = \frac{\Delta x^T f_x^*}{|\Delta x|} \tag{5}$$

$$f^{*\prime +} = \frac{\Delta x^T f_x^{*+}}{|\Delta x|} \tag{6}$$

$$f^{*\prime\prime} = \frac{6(f^{+} - f^*)}{|\Delta x|^2} - \frac{2f^{*\prime +} + 4f^{*\prime}}{|\Delta x|} \tag{7}$$

$$f^{*\prime\prime +} = \frac{6(f^* - f^{*+})}{|\Delta x|^2} + \frac{2f^{*\prime} + 4f^{*\prime +}}{|\Delta x|} \tag{8}$$

(Here the + superscript denotes evaluation at the $i+1$ end of a search segment.) The second derivative estimate corresponds to cubic fit to f^* and $f^{*\prime}$ values at the endpoints. In computations carried out with short word length or subject to excessive round-off error, the simple difference-quotient approximation, which is the average of Eqs. (7) and (8), may be preferable.

In the vicinity of the constrained minimum sought, some of the second directional derivatives of the function f^*, which approximates the first two terms of $\bar{f}$, can be expected to be positive as $f_{xx} + g_{xx}\lambda$ possesses at least n-m nonnegative eigenvalues. The largest positive value determined over one or several batches can be adopted as a guide for adjusting the penalty coefficients, as it will fall in the range between zero and the largest eigenvalue.

The second directional derivatives of f^* in directions along the constraint function gradients are not, in general, positive; if they were, in a large enough neighborhood of the constrained minimum, the quadratic penalty terms might be dispensed with. One set of requirements on the quadratic penalty coefficients might be determined from second derivatives of $\bar{f}$ in these directions, by requiring them to be equal at the least to a guideline value.

Carrying this scheme out directly necessitates either special probing operations in the directions of the constraint gradients or the inference of equivalent information from the function samples and gradients computed in the course of minimization iterations. Both have been considered and investigated in a preliminary way, and a combination is recommended for use. An estimate of the latter type for the penalty coefficients k_j is given by the maximum (over one or more batches) of the values k_{ji} given by

$$k_{ji} = \frac{\mathrm{B}_{ji}^2 (cf^{*\prime\prime}_{\mathrm{max}} - \bar{f}_i^{\prime\prime})}{|g_{i+1} g_i^{\prime\prime} + g^{\prime 2}_{i+1}|} \qquad \begin{matrix} i = 1,\text{---},n \\ j = 1,\text{---},m \end{matrix} \tag{9}$$

where

$$\beta_{ji} = \left| \frac{\Delta x_i^T}{|\Delta x_i|} \; \frac{g_{ji_x}}{|g_{ji_x}|} \right| \qquad \begin{matrix} i = 1,\text{---},n \\ j = 1,\text{---},m \end{matrix} \tag{10}$$

Values are to be excluded from consideration when the two terms of the denominator in Eq. (9) are opposite in sign and nearly equal in magnitude; the same is true when β, given by Eq. (10), is smaller than some prescribed value, indicating that the particular step Δx, was nearly in the tangent plane of the constraint whose quadratic penalty coefficient requirement is being estimated. Here $\bar{f} = f + g\bar{\lambda}$, where $\bar{\lambda}$ is the value of the linear penalty λ employed in the function $\bar{f}$ during the particular batch; a prime denotes the first derivative along the direction of the step taken, a double prime the second derivative. The expression (9) was obtained by requiring the k value be large enough to produce $\bar{f}^{\prime\prime}$ equal to the guideline value $cf^{*\prime\prime}_{\mathrm{max}}$ for $\beta = 1$ and remain bounded for smallβ-(inasmuch as the denominator behaves like β^2 for $g = 0$ and β-small). Since it is desired that k estimates err on the high side, the $f^{\prime\prime}$ values used should be the larger of the values at beginning and end of the search segment for $f^{*\prime\prime}$ and the smaller of the two values for $\bar{f}^{\prime\prime}$.

An additional candidate is introduced to cover the frequently-occurring contingency that all β_{ji} are small over one or more batches used in the selection, viz.,

$$k_{j_\ell} = \frac{(cf^{*\prime\prime}_{\mathrm{max}} - f^{*\prime\prime}_{\mathrm{min}})}{\left| g_{j\ell_x} \right|} \tag{11}$$

where ℓ corresponds to the last cycle before k selection, and $f^{*\prime\prime}_{\mathrm{min}}$ is taken as the smallest of the $f^{*\prime\prime}$ values over a chosen number of batches, or zero, whichever is the lesser. The multiplicative constant $c \geq 1$ in the guideline value of $f^{*\prime\prime}$ introduces a measure of conservatism to offset the possibility that none of the candidate values of $f^{*\prime\prime}$ in the maximization determining $f^{*\prime\prime}_{\max}$ is really close to the largest eigenvalue of $f_{xx}+g_{xx}\lambda^*$.

Metric Adjustment

After determining the λ and K elements anew at the beginning of each major iteration, one would like to adjust the variable-metric matrix H to account, at least approximately, for the changes. The corrections are based upon the idea that the H matrix emerging from the preceding major iteration approximates $\bar{f}^{-1}_{xx}$. No correction is made for changes in the sum $\lambda + Kg$ appearing in the second partials (2), since this sum approximates the Lagrange multiplier at the constrained minimum when the g_i are small.

Corrections for k_i changes are done sequentially, using

$$H+\Delta H=H-\left[\frac{\Delta k_i}{1+\Delta k_i g_{ix}^T H g_{ix}}\right] H g_{ix} g_{ix}^T H \tag{12}$$

which accounts for changes in the last term of Eq. (2) via the Schur identity.[2] Each increment Δk_i is limited to some fraction of the original or updated value k_i so as to insure that the denominator of the fraction in parenthesis remains positive and does not nearly vanish.

Test Problem

The problem used for experiments employed a cubic in one variable, x_1, for f, and a quartic of the following form for the single constraint function g:

$$f=x_1+a_1x_2^2+a_2x_1^3 \tag{13}$$

$$g=x_1-b_1x_2^2-b_2x_3^2-b_3x_2^4 \tag{14}$$

In the simplest case, used for functional checks of computer programming, $a_1=a_2=b_3=0$, $b_1>0$, $b_2>0$, the constraint surface is a paraboloid of elliptic crosssection, and the minimum of the linear function f is attained at the origin. The constraint function nonlinearity is an essential feature of the well-defined constrained minimum. If a slighty negative value of a_1 is introduced, one already has a problem for which no minimum of $f+g\lambda$ exists at the constrained minimum, since the Hessian matrix is indefinite and, accordingly, quadratic penalty terms are essential. This is not quite enough complexity for algorithm development, evaluation, and comparison, however, since $f+g\lambda$ is then quadratic and the variable-metric projection schemes have too easy a time of it. Hence, use of $a_2\neq 0$ and $b_3\neq 0$ is attractive. It should be noted that a large enough $a_2>0$ precludes the appearance of minima other than at the origin. The numerical values of the coefficients used in the computational experiments were $a_1=-10^{-2}$, $a_2=10^{-3}$, $b_1=1$, $b_2=10^2$, $b_3=10^{-1}$; these offer a modest challenge.

The starting point for the numerical computations of the example was $x_1=10$, $x_2=5$, and $x_3=10$. The multiplicative constant c used in Eqs. (9) and (11) to designate the guideline value of $f^{*\prime\prime}$ was taken as unity in the comparison.

Table 1 Convergence comparison

Algorithm	Quadratic penalty coefficient k	Number of cycles to convergence
DFP	10^3	105
DFP	10^2	58
DFP	10	24
Linear-quadratic penalty/sequential	variable	27
Linear-quadratic penalty/batch	variable	21
Rosen-Kreuser (modified)	projection	64
Kelley-Speyer	projection	72

Computational Comparison

In order to afford a basis for comparison, DFP was run on the example, with the quadratic penalty coefficient fixed at several values and zero linear penalty coefficient. The first three entries on Table 1 present these results for quadratic penalty coefficients of 10^3, 10^2, and 10. At the minima found, the constraint $g=0$ was not satisfied owing to the absence of linear penalty terms, 'boundary shifting,' or any other palliative. The violations were found to be excessive for $k\leq 10$.

The next two entries in Table 1 are for sequential and batch-processor versions of the algorithm described in the preceding. The batch version was unaccountably better than the sequential; usually, the sequential is slightly better with fixed quadratic penalty when an accurate linear search is performed.[8] The accelerated search of Ref. 9 was employed in all of the computations presently reported, with a tight tolerance employed for termination. The quadratic-penalty adjustment procedure was restrained from changing the coefficient more than an order of magnitude on any single adjustment. The scheme brought the coefficient down into the range $10^1\leq k \leq 10$, a favorable range when the linear term is present to avert large constraint violations.

The last two entries correspond to the variable-metric projection algorithms of Refs. 5 and 4, respectively, the latter slightly modified. These are reviewed in Appendices A and B for the reader's convenience.

An obvious temptation is smoothing of λ values used on successive batches by weighted averaging, heavily weighting the new projection value when there is an indication of accelerating convergence, as by drastic shrinkage in the magnitude of the projected gradient vector during the batch just completed. The motivation for avoiding unduly large fluctuations in linear penalty coefficients, of course, is that changing the function being minimized taxes the machinery for inferring the metric and the various second derivatives.

In the limited trials of this feature, to date, it has been found that it generally enhances the smoothness and "surefootedness" of the algorithm, although at slight expense in convergence speed. The use of higher values of the constant $c\geq 1$, say 2, or even 10, has a similar effect.

Conclusions

The results are thought to indicate promise for the class of algorithm combining linear and quadratic penalty adjustment with variable-metric oprimization. More extensive testing obviously is needed, including the large class of problems in which the quadratic terms can be adjusted safely downward to zero.

Appendix A: the Rosen-Kreuser Projection Algorithm

After restoration of constraints, the gradients of f and g and the projection multiplier vector are calculated and designated $\hat{f}_x$, $\hat{g}_x$, $\hat{\lambda}$ at this batch-reference point $x=\hat{x}$:

$$\hat{\lambda}=-(\hat{g}_x^T H_0 \hat{g}_x)^{-1}\hat{g}_x^T H_0 \hat{f}_x \tag{A1}$$

The algorithm proceeds to minimize $f+g\hat{\lambda}$ subject to the linear constraint

$$\hat{g}_x(x-\hat{x}))=0 \tag{A2}$$

by projecting the gradient of $f+g\hat{\lambda}$ upon this constraint at each step. The projection multiplier needed is

$$\lambda=-(\hat{g}_x^T H_0 \hat{g}_x)^{-1}\hat{g}_x^T H_0 (f_x+g_x\hat{\lambda}) \tag{A3}$$

A linear search is made in the direction

$$\Delta x=-\alpha H(f_x+g_x\hat{\lambda}+\hat{g}_x\lambda) \tag{A4}$$

to a one-dimensional minimum. On the first cycle $\lambda=0$, but not subsequently, except in special cases such as linearly constrained problems. The metric H is updated sequentially by the DFP formula, evaluated by use of the gradient of $f+g\hat{\lambda}$:

$$H+\Delta H=H-\frac{H(\Delta f_x+\Delta g_x\hat{\lambda})(\Delta f_x-\Delta g_x\hat{\lambda})^T H}{(\Delta f_x+\Delta g_x\hat{\lambda})^T H(\Delta f_x+\Delta g_x\hat{\lambda})}+\frac{\Delta x\Delta x^T}{\Delta x^T(\Delta f_x+\Delta g_x\hat{\lambda})} \tag{A5}$$

After n-m cycles, H attains its limiting value for a quadratic f, linear g model; hence n-m is a natural batch size. It would seem generally more efficient to restore and relinearize after each n-m cycles of DFP than to run to a minimum of $f+g\hat{\lambda}$ as proposed in Ref. 5. In fact, this feature encountered difficulty in the numerical computations reported, and relinearization each n-m cycles was the modification actually used.

Appendix B: the Kelley-Speyer Projection Algorithm

The accelerated gradient projection process of Ref. 4 employs the formulas

$$\Delta x=-\alpha H(f_x+g_x\lambda) \tag{B1}$$

$$\lambda=-(g_x^T H g_x)^{-1} g_x^T H f \tag{B2}$$

with λ recalculated every optimization cycle, which successfully terminates on a one-dimensional minimum of $f+g\lambda$, with H updated by

$$H+\Delta H=H+\frac{\Delta x\Delta x^T}{\Delta x^T(\Delta f_x+\Delta g_x\lambda)}-\frac{H(\Delta f_x+\Delta g_x\lambda)(\Delta f_x+\Delta g_x\lambda)^T H}{(\Delta f_x+\Delta g_x\lambda)^T H(\Delta f_x+\Delta g_x\lambda)} \tag{B3}$$

which is the DFP formula applied to the linear combination $f+g\lambda$; hence, this guarantees that H remains positive definite. Constraint restorations are carried out after each optimization cycle.[10]

References

[1] Fletcher, R. and Powell, M.J.D., "A Rapidly Convergent Descent Method for Minimization," *Computer Journal,* July 1963.

[2] Kelley, H.J., Denham, W.F., Johnson, I.L., and Wheatley, P.O., "An Accelerated Gradient Method for Parameter Optimization with Nonlinear Constraints," *Journal of the Astronautical Sciences,* Vol. 13, 1966, pp. 166-169.

[3] Goldfarb, D. and Lapidus, L., "A Conjugate Gradient Method for Nonlinear Programming," American Institute of Chemical Engineers 61st National Meeting, Houston, Tex., Feb. 1967; also, "Extension of Davidon's Variable Metric Method to Maximization Under Linear Inequality and Equality Constraints," *SIAM Journal of Applied Mathematics,* July 1969.

[4] Kelley, H.J. and Speyer, J.L., "Accelerated Gradient Projection," Colloquium on Optimization, Nice, France, June 29-July 5, 1969. Proceedings published as *Lecture Notes in Mathematics No. 132,* Springer-Verlag, Berlin, 1970.

[5] Kelley, H.J., "Method of Gradients," *Optimization Techniques,*- edited by G. Leitmann, Academic Press, New York, 1962, Chap. 6.

[7] Hestenes, M.R., "Multiplier and Gradient Methods," *Journal of Optimization, Theory and Applications,* July 1969.

[8] Kelley, H.J., Myers, G.E., and Johnson, I.L., "An Improved Conjugate Direction Minimization Procedure," *AIAA Journal,* Vol. 8, Nov. 1970.

[9] Johnson, I.L. and Kamm, J.L., "Accelerating One-Dimensional Searches," *AIAA Journal,* Vol. 11, May 1973.

[10] Myers, G.E., "Numerical Experience with Accelerated Gradient Projection," presented t the Conference on Numerical Methods for Nonlinear Optimization, University of Dundee, Scotland, June 28-July 1, 1971.