RESEARCH IN ORBIT DETERMINATION AND OPTIMIZATION FOR SPACE TRAJECTORIES

(Covering the period February 1972 - September 1975)

S. Pines
H. Kelley

Final Report

prepared for

Lyndon B. Johnson Space Center
Houston, Texas 77058

AMA Report No. 76-2
Contract NAS9-12516

ANALYTICAL MECHANICS ASSOCIATES, INC.
10210 GREENBELT ROAD
SEABROOK, MARYLAND 20801

SUMMARY

This report is a compilation of the research in orbit determination and optimization in space trajectories carried out by AMA, Inc., under contract to Lyndon B. Johnson Space Center, covering the period February 1972 - September 1975.

TABLE OF CONTENTS

INTRODUCTION

Analytical Mechanics Associates, Inc., under contract to the Lyndon B. Johnson Space Center, acted in the capacity of consultants in the areas of orbit determination, optimization techniques and trajectory design for manned space flights. In this capacity, several reports were generated and are included in the text of this final report.

A brief description of each report is included here.

(1) Multi-Spectral Scanners

This report contains a recommended method applying optimization techniques for estimating the most likely source stimulus for a given sequence of multi-spectral signals obtained from a flight scanner passing over an area. The report contains well known statistical discrimination techniques for estimating and refining information from multi-spectral scanners.

(2) Non Singular Earth Gravity Acceleration for Space Shuttle

This report contains a computer routine for computing the earth gravity acceleration designed for space borne computers. The method is both time and computer core efficient.

(3) Drag Acceleration as a Navigation Aid

This report develops a technique for obtaining a more accurate estimate of the vehicle state during reentry blackout from a drag acceleration measurement by including an adaptive atmospheric temperature lapse rate as an additional model parameter.

(4) Roll-Modulated Lifting Entry Optimization

This report develops the optimal technique for a roll-modulated lifting reentry. Thus, it provides valuable insight for obtaining practical reentry guidance laws using roll modulation.

(5) Minimum Variance Linear Estimator for Non-Linear Measurements

This report derives the minimum variance estimator for modifying the Kalman filter using range-rate measurements during highly non-linear flight regimes.

(6) Approximate Matrix Inverses

This report derives a rapid method for obtaining an approximate matrix inverse for use in flight computers when a full inverse is too time consuming and not numerically compelling.

(7) A Variable-Metric Algorithm Employing Linear and Quadratic Penalties

This report develops a variable metric optimization algorithm for accelerated search for optimization problems with linear and non-linear constraints.

AN APPROACH TO VIEWING OF MULTI-SPECTRAL SCANNER DATA

ANALYTICAL MECHANICS ASSOCIATES, INC.
50 JERICHO TURNPIKE
JERICHO, N. Y. 11753

AN APPROACH TO VIEWING OF MULTI-SPECTRAL SCANNER DATA

Consider the problem of presenting multi-spectral data to a viewer via a cathode-ray tube, say a conventional TV or, perhaps, a color TV. Spectral discrimination is to be used in some way to aid distinction between objects of main interest (targets) and other objects (background). Exploitation of the human's pattern-recognition capability is, of course, the main attraction of the viewing approach. Preprocessing of the data as well as the data for ground-truth tracts is assumed on the basis of purely spectral discrimination computations, from which at least approximate models of target and background spectra and probability density distributions are available.

With $\mu_1, ---, \mu_n$ the intensities of signal in the n measurement frequency bands and the probability densities of target and background denoted $f_T(\mu_1, ---, \mu_n)$ and $f_B(\mu_1, ---, \mu_n)$, as per Ref. 1, the ratio of the two

$$\ell = \frac{f_T}{f_B}$$

is the criterion often employed for purely spectral discrimination. Thus, if $\ell \geq C$, the sample is interpreted as a target signal. The threshold C is set to admit a specified percentage of known targets. The approach contemplated is preprocessing, perhaps approximate, to obtain models for f_T and f_B, including numerical values of means and covariances of components, then evaluation of the likelihood ratio ℓ for each data sample and use of it as a signal for a display option.

Mean and covariance for the target probability density distribution are computed from ground truth data, and in many cases a Gaussian representation will suffice. The background sometimes may be approximated by a Gaussian distribution, but more commonly will better by taken as the sum of two or three or several such, each term representing some prominent ingredient. To get means and covariances for the background, the total signal must be processed and those parts clearly target contribution screened out by a coarse criterion such as mean $\pm 2\sigma$. With this deletion, the mean and covariance of the remainder can be computed and the distribution represented as a sum of Gaussian contributions.

Feeding the likelihood ratio to a video display would light up the target areas and leave others dark. If there are two or three targets, a color display could be used. For the purpose of estimating the likelihood ratio for a particular target, the other targets are treated as components of the background. The display should, of course, also function normally, with choice of displaying the picture in any of the bandwidths on option, or perhaps there should be a second display for simultaneous viewing of raw data. In this fashion, the likelihood ratio would not be used for classification according to a black-or-white threshold criterion, but would furnish shades of grey (or green) for display and the viewer, thus assisted, would do the classifying. It is difficult to anticipate whether the approach sketched here might find its best application in preliminary editing of massive amounts of data, perhaps using off-the-shelf density models from previous reductions, or whether it might instead excel for intensive efforts on infrequent difficult cases.

REFERENCE

1. Legault, R.R.; "Multispectral Remote Sensing," lecture notes, University of Michigan, 1968.

HJK 8-72
Analytical Mechanics Associates, Inc.
50 Jericho Turnpike
Jericho, New York 11753

NONSINGULAR EARTH GRAVITY ACCELERATION FOR SPACE SHUTTLE

S. Pines

Report No. 73-17
Contract No. NAS 9-12516
April 1973

ANALYTICAL MECHANICS ASSOCIATES, INC.
50 JERICHO TURNPIKE
JERICHO, N. Y. 11753

INTRODUCTION

The Shuttle computer requires a nonsingular gravity acceleration routine for polar orbits. The following formulation, based on Reference 1, is carried out in detail for an Earth model consisting of J_2, J_3, J_4, C_{22}, and S_{22}.

REFERENCE

1. Pines, S. and Austin, G.; "Gravitational Acceleration of a Point Mass Due to a Rotating Nonspherical Body," Analytical Mechanics Associates, Inc. Report No. 69-12, May 1969.

EQUATIONS

The nonsingular potential in Earth-fixed coordinates is given by

$$\varphi = \frac{\mu}{r} - \frac{\mu}{r}\left(\frac{a}{r}\right)^2 J_2 A_{2,0}(u) - \frac{\mu}{r}\left(\frac{a}{r}\right)^3 J_3 A_{3,0}(u) - \frac{\mu}{r}\left(\frac{a}{r}\right)^4 J_4 A_{4,0}(u)$$

$$+ \frac{\mu}{r}\left(\frac{a}{r}\right)^2 A_{22}(u)\left[C_{22} R_2(s,t) + S_{22} I_2(s,t)\right] \tag{1}$$

where

μ = central body gravity constant

a = Earth radius

$r = (x^2 + y^2 + z^2)^{\frac{1}{2}}$

$s = \frac{x}{r}$

$t = \frac{y}{r}$

$u = \frac{z}{r}$

$$A_{i,j} = \frac{1}{2^i\, i!}\,\frac{d^{i+j}}{du^{i+j}}\,(u^2 - 1)^i \tag{2}$$

$$R_2(s,t) = s^2 - t^2$$

$$I_2(s,t) = 2\,s\,t$$

The acceleration vector in the body-fixed system is given by

$$F = \left(\frac{\partial\varphi}{\partial r} - \frac{s}{r}\frac{\partial\varphi}{\partial s} - \frac{t}{r}\frac{\partial\varphi}{\partial t} - \frac{u}{r}\frac{\partial\varphi}{\partial u}\right)\frac{R}{r} + \frac{1}{r}\frac{\partial\varphi}{\partial s}\hat{i}$$

$$+ \frac{1}{r}\frac{\partial\varphi}{\partial t}\hat{j} + \frac{1}{r}\frac{\partial\varphi}{\partial u}\hat{k} \tag{3}$$

where $\hat{i}$, $\hat{j}$, $\hat{k}$ are unit vectors in the x, y, z directions, respectively.

For the specific Earth oblateness coefficients J_2, J_3, J_4, C_{22}, and S_{22}, the partials are given below.

$$\frac{\partial \varphi}{\partial r} = 3 \frac{\mu a^2}{r^4} J_2 A_{2,0}(u) + 4 \frac{\mu a^3}{r^5} J_3 A_{3,0}(u) + 5 \frac{\mu a^4}{r^6} J_4 A_{4,0}(u)$$

$$- \frac{\mu a^2}{r^4} A_{2,2}(u) [C_{22} R_2(s,t) + S_{22} I_2(s,t)]$$

$$\frac{1}{r} \frac{\partial \varphi}{\partial s} = \frac{\mu a^2}{r^4} A_{22}(u) \, 2(S C_{22} + t S_{22}) \tag{4}$$

$$\frac{1}{r} \frac{\partial \varphi}{\partial t} = \frac{\mu a^2}{r^4} A_{22}(u) \, 2(S S_{22} - t C_{22})$$

$$\frac{1}{r} \frac{\partial \varphi}{\partial u} = -\frac{\mu a^2}{r^4} J_2 A_{21}(u) - \frac{\mu a^3}{r^5} J_3 A_{31}(u) - \frac{\mu a^4}{r^6} J_4 A_{41}(u)$$

A simple computer code is given in the next section.

NONSINGULAR EARTH OBLATENESS ROUTINE

OBLAT(R, ACC)

```
      DIMENSION R(3), ACC(3)
      COMMON    EMU, a, J2, J3, J4, C22, S22
      COMMENT   EMU = central mass coefficient
      COMMENT     a = Earth radius

      R2   =  R(1)*R(1) + R(2)*R(2) + R(3)*R(3)
      R    =  SQRT(R2)
      RINV =  1.0/R
      B    =  a*RINV
      C    =  EMU/R2*B
      B2   =  B*C
      B3   =  B*B2
      B4   =  B*B3
      S    =  R(1)*RINV
      T    =  R(2)*RINV
      U    =  R(3)*RINV
      AS   =  B2*6.0*(S*C22 + T*S22)
      AT   =  B2*6.0*(S*S22 - T*C22)
      U2   =  U*U
      U3   =  U*U2
      U4   =  U*U3
      A2   =  (3.0*U2 - 1.0)/2.0
      A21  =  3.0*U
      A3   =  (5.0*U3 - A21)/2.0
      A4   =  (35.0*U4 - 30.0*U2 + 3.0)/8.0
```

```
A31   =   (15.0*U2 - 3.0)/2.0
A41   =   (35.0*U3 - 15.0*U)/2.0
AU    =   - B2*J2*A21 - B3*J3*A31 - B4*J4*A41
AR    =   3.0*B2*J2*A2 + 4.0*B3*J3*A3 + 5.0*B4*J4*A4
            - B2*3.0*(C22*(S*S - T*T) + S22*2.0*S*T)
AR    =   AR - S*AS - T*AT - U*AU
ACC(1) = AR*S + AS*RINV
ACC(2) = AR*T + AT*RINV
ACC(3) = AR*U + AU*RINV
END
```

DRAG ACCELERATION IN A THERMAL-VARIABLE ATMOSPHERE AS A NAVIGATION AID

S. Pines

Report No. 73-22
Contract No. NAS 9-12516
April 1973

ANALYTICAL MECHANICS ASSOCIATES, INC.
50 JERICHO TURNPIKE
JERICHO, N. Y. 11753

NOTATION

p	atmospheric pressure
ρ	atmospheric mass density
T	atmospheric temperature (degrees Kelvin)
m_w	mean molecular weight
q	universal gas constant
h	altitude
R	vehicle position vector
$\dot{R}$	vehicle velocity vector
V_R	vehicle velocity vector relative to the atmosphere
g	acceleration of gravity (assumed constant in the atmosphere)
ΔV	delta velocity readout vector of the accelerometers
Δt	accelerometer readout count time interval
A	effective vehicle area for drag
m	vehicle mass
ω	Earth angular rotation rate
M	Mach number
c	speed of sound in atmosphere
r_E	mean Earth radius
e	eccentricity of Earth

ϕ'	geocentric latitude
ϕ	geodetic latitude
z	polar component of vehicle position vector
γ	ratio of specific heats
r	magnitude of radius vector
V_R	magnitude of relative vehicle velocity

INTRODUCTION

The drag acceleration is presently being considered as an observation type for navigation as a substitute for an altimeter reading during radio blackout in re-entry. The measurement is somewhat degraded in the event the atmospheric density or the aerodynamic drag characteristics become uncertain. This report resolves the density variability by using a model of the atmospheric density from 80 km to 32 km in the form of a layered atmosphere characterized by a sequence of altitude intervals with piece-wise constant temperature lapse rates. The intent is to use this parameterization to enable the navigation filter to determine the vehicle state as well as the atmospheric density variation. In this manner, a more accurate determination of the state and covariance cross-correlation will exist at blackout termination when more effective observations can again be made.

ATMOSPHERIC MODELS

Below 100 km, the Earth's atmosphere seems to obey the perfect gas law for a constant mean molecular weight.

$$\rho = \frac{m_w p}{q T} \tag{1}$$

where

$$\frac{m_w}{q} = \text{constant}$$

Under the assumption of a temperature variation which is piece-wise linear with altitude, we have

$$\begin{aligned} &h_1 \le h \le h_2 \\ &T = T(h_1) + T'(h_1)(h - h_1) \end{aligned} \tag{2}$$

Since the equilibrium pressure variation is given by

$$\frac{dp}{dh} = -g\rho \tag{3}$$

the density between h_1 and h_2 is given by

$$\rho = \rho(h_1)\left[1 + \frac{T'(h_1)}{T(h_1)}(h - h_1)\right]^{-\{[g\, m_w/q\, T'(h_1)] + 1\}} \tag{4}$$

Since the equilibrium atmosphere experiences temperature gradient reversals, there exist altitude regions of 5 km or so during which constant temperature persists. In such intervals, we have

$$T(h_1) = T = T(h_2)$$

$$\rho = \rho(h_1)\, e^{-[m_w\, g/q\, T(h_1)][h-h_1]} \tag{5}$$

The 1962 U.S. Standard Atmosphere (Ref. 1) lists the following nominal values of the temperature lapse rates

TABLE I

Altitude km	Temperature °Kelvin	Lapse Rate °Kelvin/km	Density kg/m^3
32	228.65		1.3225×10^{-2}
		2.8	
47	270.65		1.4275×10^{-3}
		0	
52	270.65		7.5943×10^{-4}
		-2.0	
61	252.65		2.5109×10^{-4}
		-4.0	
79	180.65		2.001×10^{-5}

Using the above four-layer model, it is possible to construct an error model for density variations and require that the filter solve for the parameters during the descent blackout.

DRAG ACCELERATION

The observation of drag acceleration over the short time interval during which the accelerometers read out the delta v count (0.5 seconds) is given by

$$D.A. = \tfrac{1}{2}\rho v_R^2 C_D \frac{A}{m} = \frac{\Delta V \cdot V_R}{v_R \Delta t} \tag{6}$$

While the representation of the observation appears to provide the ability to estimate both position and velocity, the uncertainty in modelling the aerodynamic coefficient C_D as a function of Mach number and angle of attack make this approach undesirable. We will compute the quantity $\frac{1}{2} v_R^2 C_D \frac{A}{m}$ and produce the pseudo-observation of density

$$\rho = \frac{\Delta V \cdot V_R}{v_R^3 \Delta t} \frac{2m}{C_D A} \tag{7}$$

From the knowledge of the vehicle state, we compute

$$\begin{aligned}
V_R &= \dot{R} - \Omega x R \\
v_R &= (V_R \cdot V_R)^{\frac{1}{2}} \\
\Omega &= \hat{k}\,\omega \qquad \text{(Earth's angular rotation vector)} \\
C_D &= C_D(M, \alpha) \quad \text{(drag coefficient as function of Mach number of angle of attack)} \\
M &= \frac{v_R}{c} \\
c &= \left[\gamma \frac{q}{m_w} T(h)\right]^{\frac{1}{2}} \\
T(h) &= T(h_1) + T'(h_1)(h - h_1) \\
\alpha &= \cos^{-1} \frac{\hat{i}_x \cdot V_R}{v_R} \qquad \text{(angle of attack)}
\end{aligned} \tag{8}$$

Following the previous section, we model the density as a function of the geodetic altitude, h. The altitude is the height above the vehicle sub-satellite point. The Earth is modelled as an oblate spheroid with eccentricity, e. From Ref. 2, the altitude is given by

$$h = r\cos(\phi - \phi') - r_E(1 - e^2 \sin^2\phi)^{\frac{1}{2}} \tag{9}$$

The geocentric latitude is defined as

$$\sin\phi' = \frac{z}{r} \tag{9a}$$

The geodetic latitude, accurate to twelve digits, is given by

$$\phi = \phi' + a_2(r,e)\sin 2\phi' + a_4(r,e)\sin 4\phi' + a_6(r,e)\sin 6\phi' + a_8(r,e)\sin 8\phi' \tag{9b}$$

where

$$a_2 = \frac{r_E}{r}\frac{1}{1024}(512e^2 + 128e^4 + 60e^6 + 35e^8) + \left(\frac{r_E}{r}\right)^2\frac{1}{32}(e^6 + e^8) - \left(\frac{r_E}{r}\right)^3\frac{3}{256}(4e^6 + 3e^8)$$

$$a_4 = -\frac{r_E}{r}\frac{1}{1024}(64e^4 + 48e^6 + 35e^8) + \left(\frac{r_E}{r}\right)^2\frac{1}{16}(4e^4 + 2e^6 + e^8) + \frac{15e^8}{256}\left(\frac{r_E}{r}\right)^3 - \left(\frac{r_E}{r}\right)^4\frac{e^8}{16} \tag{9c}$$

$$a_6 = \frac{r_E}{r}\frac{3}{1024}(4e^6 + 5e^8) - \frac{3}{32}\left(\frac{r_E}{r}\right)^2(e^6 + e^8) + \frac{35}{768}\left(\frac{r_E}{r}\right)^3(4e^6 + 3e^8)$$

$$a_8 = \frac{e^8}{2048}\left[-5\frac{r_E}{r} + 64\left(\frac{r_E}{r}\right)^2 - 252\left(\frac{r_E}{r}\right)^3 + 320\left(\frac{r_E}{r}\right)^4\right]$$

and

$$e^2 = 2\varepsilon - \varepsilon^2$$

$$\varepsilon = \frac{1}{297.3}$$

The above equations are sufficient to enable one to estimate the density.

PARTIAL DERIVATIVES OF THE DENSITY OBSERVATION

We take as our state

R	vehicle position vector
$\dot{R}$	vehicle velocity vector
$\rho(h_i)$	density at lower altitude of the i^{th} layer
$T'(h_i)$	temperature lapse rate of the i^{th} layer
$T(h_i)$	temperature at the lower altitude of the i^{th} layer

The partials of the density with respect to each component of the state are given by

$$\frac{\partial \rho}{\partial \beta} = \frac{\partial \rho}{\partial h}\frac{\partial h}{\partial \beta} + \frac{\partial \rho}{\partial \rho(h_i)}\frac{\partial \rho(h_i)}{\partial \beta} + \frac{\partial \rho}{\partial T'(h_i)}\frac{\partial T'(h_i)}{\partial \beta} + \frac{\partial \rho}{\partial T(h_i)}\frac{\partial T(h_i)}{\partial \beta} \tag{10}$$

The partials of the density with respect to the altitude, h, are given by

For $T'(h_i) \neq 0$

$$\frac{\partial \rho}{\partial h} = -\left(\frac{g\, m_w}{q\, T'(h_i)} + 1\right)\rho(h_i)\left[1 + \frac{T'(h_i)}{T(h_i)}(h - h_i)\right]^{-\{[g m_w / q T'(h_i)] + 2\}}\frac{T'(h_i)}{T(h_i)} \tag{11a}$$

For $T'(h_i) = 0$

$$\frac{\partial \rho}{\partial h} = -\rho\,\frac{g\, m_w}{q\, T(h_i)} \tag{11b}$$

The partial of the density with respect to $\rho(h_i)$ is given by

$$\frac{\partial \rho}{\partial \rho(h_i)} = \frac{\rho}{\rho(h_i)} \tag{12}$$

The partial of the density with respect to $T'(h_i)$ is given by

For $T'(h_i) \neq 0$

$$\frac{\partial \rho}{\partial T'(h_i)} = \rho\left[\ln\left\{\left(1 + \frac{T'(h_i)(h-h_i)}{T(h_i)}\right)\right\}\left(\frac{g\, m_w}{q\, T'^2(h_i)}\right) - \frac{1+\dfrac{g\, m_w}{q\, T'(h_i)}}{1+\dfrac{T'(h_i)}{T(h_i)}}\ \frac{h-h_i}{T(h_i)}\right] \tag{13a}$$

For $T'(h_i) = 0$

$$\frac{\partial \rho}{\partial T'(h_i)} \tag{13b}$$

The partial of the density with respect to $T(h_i)$ is given by

For $T'(h_i) \neq 0$

$$\frac{\partial \rho}{\partial T(h_i)} = \rho(h_i)\left[\frac{g\, m_w}{q\, T'(h_i)} + 1\right]\left[1 + \frac{T'(h_i)(h-h_i)}{T(h_i)}\right]^{-\{[g\, m_w/q\, T'(h_i)]+2\}} \frac{T'(h_i)(h-h_i)}{T^2(h_i)} \tag{14a}$$

For $T'(h_i) = 0$

$$\frac{\partial \rho}{\partial T(h_i)} = \frac{g\, m_w}{q\, T^2(h_i)}\ \rho \tag{14b}$$

The partial of the altitude, h, with respect to the state is computed on the assumption that $\phi = \phi'$, and is given by

$$\frac{\partial h}{\partial R} = \left[1 - \frac{r_E}{r} \frac{e^2 \sin^2 \phi}{(1-e^2 \sin^2 \phi)^{\frac{1}{2}}}\right] \frac{R}{r} + \frac{r_E}{r} \frac{e^2 \sin^2 \phi}{(1-e^2 \sin^2 \phi)^{\frac{1}{2}}} \hat{k}$$

$$\frac{\partial h}{\partial \dot{R}} = \frac{\partial h}{\partial T'(h_i)} = \frac{\partial h}{\partial T(h_i)} = \frac{\partial h}{\partial \rho(h_i)} = 0 \qquad (15)$$

The partial of $\rho(h_i)$ with respect to the state is

$$\frac{\partial \rho(h_i)}{\partial R} = \frac{\partial \rho(h_i)}{\partial \dot{R}} = \frac{\partial \rho(h_i)}{\partial T'(h_i)} = \frac{\partial \rho(h_i)}{\partial T(h_i)} = 0$$

$$\frac{\partial \rho(h_i)}{\partial \rho(h_i)} = 1 \qquad (16)$$

The partial of the temperature lapse rate with respect to the state is given by

$$\frac{\partial T'(h_i)}{\partial R} = \frac{\partial T'(h_i)}{\partial \dot{R}} = \frac{\partial T'(h_i)}{\partial \rho(h_i)} = \frac{\partial T'(h_i)}{\partial T(h_i)} = 0$$

$$\frac{\partial T'(h_i)}{\partial T'(h_i)} = 1 \qquad (17)$$

The partial of the temperature at h_i, $T(h_i)$, with respect to the state is given by

$$\frac{\partial T(h_i)}{\partial R} = \frac{\partial T(h_i)}{\partial \dot{R}} = \frac{\partial T(h_i)}{\partial \rho(h_i)} = \frac{\partial T(h_i)}{\partial T'(h_i)} = 0 \tag{18}$$

$$\frac{\partial T(h_i)}{\partial T(h_i)} = 1$$

This completes the partials.

Whenever $|T'(h_i)| \leq 0.5^{\circ}/\text{km}$, set $T'(h_i) = 0$.

NAVIGATION FILTER UPDATE EQUATIONS

Reference 3 contains the update equations for an onboard navigation filter for an 18-element vehicle state vector consisting of the vehicle position vector, R, the vehicle velocity vector, $\dot{R}$, the vehicle gyro tilt error vector, θ, the gyro tilt rate error vector, $\dot{\theta}$, the accelerometer scale factor error vector, k, and the accelerometer bias error vector, b. To this state we add the scalars $\rho(h_i)$, $T'(h_i)$, and $T(h_i)$. Thus the new state is a 21-element state error vector, X. Following Ref. 3, the update equations following each pseudo-density observation are given by

$$X^{+} = X^{-} + C\,P^{T}(P\,C\,P^{T} + Q)^{-1}(\rho_{ob} - \rho_{comp}) \tag{19}$$

where C is the 21×21 covariance matrix of the errors in the estimate of the state vector, P is the 21×1 vector of the partials of the pseudo-observation with respect to the state X, and Q is the observation noise. We have

$$P = \left[\frac{\partial\rho}{\partial R}, \frac{\partial\rho}{\partial \dot{R}}, \frac{\partial\rho}{\partial\theta}, \frac{\partial\rho}{\partial\dot{\theta}}, \frac{\partial\rho}{\partial k}, \frac{\partial\rho}{\partial b}, \frac{\partial\rho}{\partial\rho(h_i)}, \frac{\partial\rho}{\partial T'(h_i)}, \frac{\partial\rho}{\partial T(h_i)}\right] \tag{20}$$

The partials of the density with respect to R, $\dot{R}$, $\rho(h_i)$, $T'(h_i)$, and $T(h_i)$ are given in the previous section, and the remaining partials are all zero. The residual is given by

$$\Delta\rho = \frac{\Delta V \cdot V_R\, 2\,m}{v_R^3\, C_D\, A} - \rho[h, \rho(h_i), T(h_i), T'(h_i)] \tag{21}$$

The value of Q recommended for simulation study is

$$Q = .01\,\rho^2 \tag{22}$$

REFERENCES

1. Anon.; U.S. Standard Atmosphere, 1962, prepared under sponsorship of NASA, USAF, and U.S. Weather Bureau, December 1962.

2. Morrison, J. and Pines, S.; "The Reduction from Geocentric to Geodetic Coordinates," Astronomical Journal, February 1961, pp. 15-16.

3. Lear, W.; "A Prototype Real-Time Navigation Program for Multi-Phase Missions," TRW Systems, Inc. Report No. 17618-6003-TO-00, December 1971.

ROLL-MODULATED LIFTING ENTRY OPTIMIZATION

H.J. Kelley
H.C. Sullivan

ANALYTICAL MECHANICS ASSOCIATES, INC.
50 JERICHO TURNPIKE
JERICHO, N. Y. 11753

Reprinted from AIAA JOURNAL, Vol. 11, No. 7, July 1973, pp 913-915

Roll-Modulated Lifting Entry Optimization

HENRY J. KELLEY*
Analytical Mechanics Associates Inc., Jericho, N.Y.
AND
HENRY C. SULLIVAN†
Lyndon B. Johnson Space Center, Houston, Texas

The equations of lifting entry are examined for fixed angle-of-attack vehicular motion with path control via roll modulation of lift. A complication arising with this is nonconvexity of the hodograph figure, which makes the application of standard variational techniques inadvisable unless the problem is first relaxed, i.e., a related problem is defined with a hodograph figure that is the convex hull of the original. This leads to a new system in new variables that is apparently innocuous in its simplicity; the linear elements of the convex hull, however, are associated with singular extremal subarcs and their attendant difficulties. The singular extremal for minimum-heating symmetric flight with final time and downrange open is simple. Two order-reduction approximations are considered, which may include intervals of two-dimensional motion as subarcs. One of these approximations relegates turning to initial and terminal boundary-layer maneuvers; the other is analogous to the aircraft energy-maneuvering model. Some computations for a space shuttle orbiter configuration are presented.

Nomenclature

D = drag
E = specific energy
g_o = acceleration of gravity
H = variational Hamiltonian
L = lift
Q = total heat load
$\dot{Q}$ = heat rate
r = radius
r_o = radius of the Earth
V = velocity
W = weight
γ = flight path angle to horizontal
Λ = longitude
λ = Lagrange multiplier
μ = bank angle
ζ = relaxation interpolation variable
σ = relaxation control variable
ϕ = latitude
χ = heading angle to south

State Equations

WITH r radius, γ path angle to horizontal, $E \equiv (V^2/2g_o) - (r_o^2/r)$ specific energy, χ heading angle to south, ϕ latitude, Λ longitude, and μ bank angle, the equations of state are

$$\dot{r} = V \sin\gamma \tag{1}$$

$$\dot{E} = -DV/W \tag{2}$$

$$\dot{\gamma} = (g_o L \cos\mu/WV) - (g_o r_o^2/Vr^2)\cos\gamma + (V/r)\cos\gamma \tag{3}$$

$$\dot{\chi} = (g_o L \sin\mu/WV\cos\gamma) - (V/r)\cos\gamma\sin\chi\tan\phi \tag{4}$$

$$\dot{\phi} = -(V/r)\cos\chi\cos\gamma \tag{5}$$

$$\dot{\Lambda} = (V\sin\chi\cos\gamma/r\cos\phi) \tag{6}$$

$$\dot{Q} = \dot{Q}(E, r) \tag{7}$$

The first six equations are particle-dynamics equations of motion for coordinated maneuvering (zero side-force). The last equation is the total heating integral Q in differential form. Lift L and drag D are functions of E and r only; angle of attack is assumed constant. (If trim were to vary with the Mach number, the angle of attack would itself be a function of E and r.) Inequality constraints on dynamic pressure, normal load factor, and local temperatures are in the problem statement.

Roll Modulation

Entry at essentially constant angle of attack has been employed for such vehicles as the Apollo Command Module, with consequent simplification of longitudinal control. The desired vertical component of lift and a desired average out-of-plane component are obtained by bank reversals, square-wave fashion. In the particle-dynamics model, this includes the theoretical possibility of "chattering," since rigid-body rolling dynamics have been neglected. There is design interest in roll modulation for advanced vehicles such as the Earth-orbital shuttle, even though a longitudinal control system will be featured, since design compromises may force a narrow range of trim angle of attack. Thus, constant angle-of-attack operation is of interest as a limiting case for the shuttle entry problem.

Control Relaxation

In the version of the problem with angle of attack controllable within bounds, the figure in hodograph space $(\dot{E}, \dot{\gamma}, \dot{\chi})$ that is traced out by varying the controls α and μ over their complete range (Contensou's "Domain of Maneuverability")[1] is not convex. Operation at points within the figure, which is a paraboloid for lift linear and drag quadratic in α, can be approximated by chattering control operation, square-wave fashion, but cannot actually be attained with piecewise continuous controls. In such circumstances, it is usual to consider instead a related problem with different control variables that attain the convex hull of the hodograph figure; this is the "relaxed" problem.[1,2] The relaxation for the variable angle-of-attack case is sketched in Ref. 3. In the present case of fixed angle of attack, the figure is an ellipse. Relaxation makes the disk within this ellipse attainable.

Relaxation may be accomplished for a general state system of the form

$$\dot{x} = f(x, u, t) \tag{8}$$

Presented as Paper 72-933 at the AIAA/AAS Astrodynamics Specialist Conference, Palo Alto, Calif., September 11-12, 1972; submitted October 6, 1972; revision received March 13, 1973. Research supported in part by Lyndon B. Johnson Space Center under Contract NAS 9-11532.

Index categories: Entry Vehicle Mission Studies and Flight Mechanics; Navigation, Control, and Guidance Theory.

* Vice President. Associate Fellow AIAA.
† Aerospace Technologist. Member AIAA.

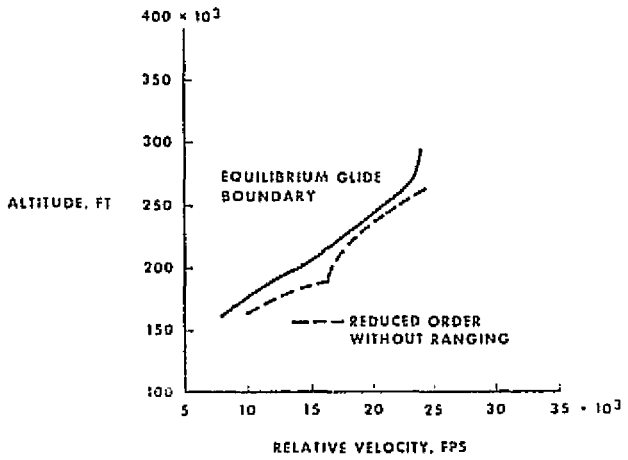

Fig. 1 Minimum-heating trajectory in altitude/velocity chart.

by replacing the system by

$$\dot{x} = f(x, u_1, t) + \zeta[f(x, u_2, t) - f(x, u_1, t)] \tag{9}$$

in which the right members are linearly interpolated between values for control u_1 and control u_2. Here ζ, $0 \leqq \zeta \leqq 1$, is an interpolation parameter. The control variables of the relaxed system are the vectors u_1 and u_2 and the scalar ζ. In the present application, the desired goal of attaining the interior of the ellipse can be accomplished with fewer variables, namely by introducing an additional control variable σ, $0 \leqq \sigma \leqq 1$, multiplicative on L in the $\dot{\gamma}$ and $\dot{\chi}$ state equations

$$\dot{\gamma} = (g_o L\sigma \cos\mu/WV) - (g_o r_o^2/Vr^2)\cos\gamma + (V/r)\cos\gamma \tag{3a}$$

$$\dot{\chi} = (g_o L\sigma \sin\mu/WV\cos\gamma) - (V/r)\cos\gamma \sin\chi \tan\phi \tag{4a}$$

Singular Arcs of the Relaxed Problem

The appearance of the control variable σ linearly in the right members of the state equations indicates the possibility of singular arcs in the solution of optimal entry control problems. This possibility may be investigated by formation of the usual Hamiltonian H, setting $\partial H/\partial\sigma = 0$, and pursuing the consequences.

$$\partial H/\partial\sigma = (g_o L/WV)[\lambda_\gamma \cos\mu + \lambda_\chi(\sin\mu/\cos\gamma)] = 0 \tag{10}$$

$$\partial H/\partial\mu = (g_o L\sigma/WV)[-\lambda_\gamma \sin\mu + \lambda_\chi(\cos\mu/\cos\gamma)] = 0 \tag{11}$$

Left members of Eqs. (10) and (11) must vanish independently. Since these are linearly independent, it follows that both λ_γ and λ_χ are zero along the arc.

The system is already in the canonical form of Ref. 4; thus the variables γ and χ are control-like along singular arcs. A similar result could have been obtained by noting that $\sigma \sin\mu$ and $\sigma\cos\mu$ could be taken as new control variables in the neighborhood of a singular arc for $0 < \sigma < 1$. Desired variations in γ and χ can be realized by varying these, as long as the magnitude of the desired variations is sufficiently small as not to encounter saturation of the σ bounds.

With γ and χ regarded as controls, the problem simplifies to flight in the plane of a great circle. Without loss of generality, take $\mu = \phi = 0$, and $\chi = \pi/2$ for study of this two-dimensional motion, and the state equations become

$$\dot{r} = V\sin\gamma \tag{12}$$

$$\dot{E} = -DV/W \tag{13}$$

$$\dot{\Lambda} = V\cos\gamma/r \tag{14}$$

$$\dot{Q} = \dot{Q}(E, r) \tag{15}$$

In the special case of downrange open (final Λ unspecified for initially equatorial flight), the control variable γ enters only Eq (12) and the variable r becomes control-like along singular arcs as the form with Eq. (12) deleted is again canonical. If final time is open, there is analytical advantage in casting E in the role of independent variable; furthermore, the steady decrease of E makes this interchange feasible for entry applications.

$$dQ/dE = -W\dot{Q}/DV \tag{16}$$

The singular extremal is defined by stationary points of the right member of Eq. (16) regarded as a function of r at various E levels. The generalized Legendre-Clebsch[4] test requires that the stationary value of the right member of Eq. (16) as a function of r be a maximum and $\dot{Q}/DV$ minimum.

Reduced Relaxed Problems

The relaxed problem presents computational difficulties because of singular arcs; approximations are therefore of more than usual interest. Possibilities offered by singular perturbation procedures[5-7] are discussed in the following paragraphs. If nearly symmetric flight were assumed, a singular perturbation approach designating latitude, longitude, and heating as variables of a reduced solution (i.e., solution of a reduced-order approximation problem) would seem attractive. This would relegate turning and altitude transitions to corrective boundary-layer transients near initial and terminal points. Energy is chosen as the independent variable. The reduced problem is of the great-circle type.

The great-circle reduced-order system for the approximation that combines altitude and heading transients takes the form

$$d\Lambda/dE = -W/Dr \tag{17}$$

$$dQ/dE = -W\dot{Q}/DV \tag{18}$$

The order-reduction procedure used is the same one examined and employed in Ref. 6. An upper bound on the control variable r of the reduced problem is furnished by the control bound $\sigma = 1$ of the original problem together with Eq. (3a) and $\dot{\gamma} = \gamma = 0$; a lower bound is provided by state inequalities on panel temperatures and acceleration loads, handled in penalty function approximation in the computations next described. Use of the model given by Eqs. (17) and (18) is limited to problems for which downrange is specified as greater than, or equal to, the downrange-open value; for smaller specified downrange, the singular extremal fails the generalized Legendre-Clebsch test and a zigzag competitor is optimal.

A less drastic approximation using singular perturbations would treat heading as well as latitude, longitude, and heating in a reduced problem. This would idealize only the altitude transients as fast (with respect to energy change) compared to the other transitions. It is the same as aircraft energy approximation.[3,6] Energy approximations have previously been examined for atmospheric entry of a variable angle-of-attack vehicle.[7] No complications arising from the relaxed model are anticipated using this approach. Evidently a solution for the reduced-order fixed angle-of-attack problem consists generally of a turning arc, a great-circle time-open arc, and, if final heading is specified, a terminal turning arc.

Computational Results

Data for a delta-wing space-shuttle orbiter configuration were used for some sample computations with the model of Eqs. (17) and (18). The angle of attack was fixed at 30°. Inequality constraints on normal load factor and numerous panel temperatures were incorporated by using penalty functions.

A minimum of the Hamiltonian consisting of a linear combination of the right members of Eqs. (17) and (18) plus penalties was found by one-dimensional search. With downrange open, the minimum always occurred at the lower bound on altitude furnished by the load factor and temperature constraints (see Fig. 1). With downrange specified at values exceeding the open value, the minimizing altitude was found to be the upper bound value ($\sigma = 1$) during the latter part of the trajectory (see Fig. 2). As range requirements were increased, numerical results indicated the possibility of more than one switch between altitude bounds. The Hamiltonian function for the downrange-specified case of

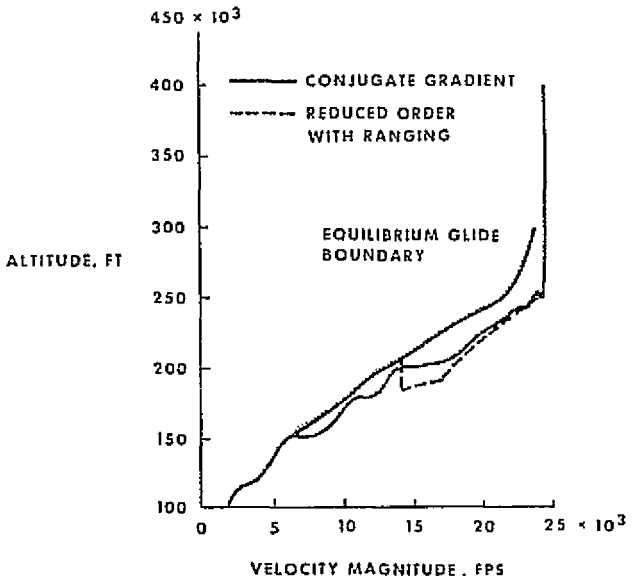

Fig. 2 Minimum-heating downrange-specified trajectories in altitude/velocity chart.

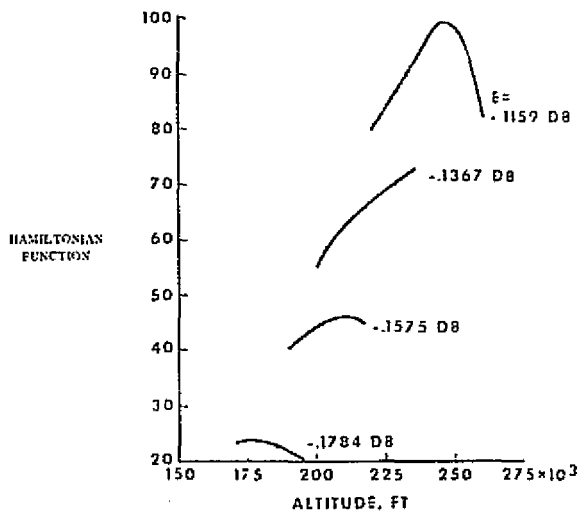

Fig. 3 Hamiltonian vs altitude at several energy levels for downrange-specified case.

Fig. 2 is sketched vs altitude in Fig. 3 for several energy values. The sign of the second derivative H_{rr} would seem to indicate nonconvexity and a need for further relaxation. However, one recalls that r in the role of control variable is not the real thing but the result of an order-reduction approximation amounting to assumed instantaneous vertical dynamics. This implies that weak as well as strong minima should be considered; hence, that transitions determined according to absolute minimum H, as in Fig. 2, are somewhat arbitrary.[6] Boundary-layer transition fairings at discontinuities in r, as in Ref. 6, are needed for consistency in degree of approximation of the control, but they contribute nothing to the performance index in this approximation.

When heating was heavily weighted compared to down-ranging, values of σ were found to be below unity indicating a need for roll modulation in two-dimensional flight. However, solutions with downrange heavily weighted ride the upper bound $\sigma = 1$ at low energies and at near-orbital energies.

Results obtained by a conjugate gradient method that used a particle-dynamics model are shown for comparison. The cross-range was specified at a somewhat challenging value of about 1300 nm. The conjugate gradient formulation did not employ a relaxed model and was unsuitable for nearly symmetric flight cases. It exhibited poor convergence that was, perhaps, attributable to the absence of convexity. Nonetheless, the result of Fig. 2 seems of interest for the qualitative similarity of the altitude history with the great-circle model. This was obtained using as a first guess a trajectory which had been forced to follow the lower bound representing temperature limit approximately. The comparisons suggest that the idealization of early heading and altitude transitions followed by altitude control based mainly on heating and down-ranging may warrant further investigation. A separate treatment of the initial transition as a boundary layer in which altitude, path angle, and heading motions are fast (with respect to E changes) could be carried out along the lines of that for aircraft altitude transitions in Ref. 6.

Conclusion

Attention has been directed to relaxation and its consequences for the fixed angle-of-attack atmospheric entry problem. Two reduced-order approximations for the resulting system of equations have been briefly examined and appear to warrant additional study.

References

[1] Contensou, P., "Etude Théorique des Trajectoires Optimales dans un Champ de Gravitation. Application au Cas d'un Centre d'Attraction Unique," *Astronautica Acta*, Vol. VIII, Fasc. 2–3, 1962, pp. 134–150.

[2] Warga, J., "Relaxed Variational Problems," *Journal of Mathematical Analysis and Applications*, Vol. 4, 1962, pp. 111–127.

[3] Kelley, H. J. and Edelbaum, T. N., "Energy Climbs, Energy Turns and Asymptotic Expansions," *Journal of Aircraft*, Vol. 7, No. 1, 1970, pp. 93–95.

[4] Kelley, H. J., Kopp, R. E., and Moyer, H. G., "Singular Extremals," *Topics in Optimization*, edited by G. Leitmann, Academic Press, New York, 1967.

[5] Wasow, W., *Asymptotic Expansions for Ordinary Differential Equations*, Interscience, New York, 1965.

[6] Kelley, H. J., "Aircraft Maneuver Optimization by Reduced-Order Approximation," *Controls and Dynamic Systems: Advances in Theory and Applications*, Vol. X, edited by C. T. Leondes, Academic Press, New York, 1973.

[7] Krenkel, R., Kelley, H. J., O'Dwyer, W., and Hinz, H., "Euler Equations for 3-D Reentry in Energy Approximation," RN-304, June 1971, Grumman Aerospace Corp., Bethpage, N.Y.

MINIMUM VARIANCE LINEAR ESTIMATOR FOR NONLINEAR MEASUREMENTS

Samuel Pines

Report No. 73-42
Contract NAS 9-12516
October 1973

ANALYTICAL MECHANICS ASSOCIATES, INC.
50 JERICHO TURNPIKE
JERICHO, N. Y. 11753

Let $\hat{x}_o$ be the best estimate of the vector state, and x the true state vector. Then the vector error in the best estimate is given by

$$e_o = x - \hat{x}_o \tag{1}$$

Let y be a scalar measurement which is a nonlinear (quadratic) function of the vector state x, contaminated by white noise. Then, the true measurement is given by

$$y(x, \eta) = y(\hat{x}_o) + y_x^T e_o + \tfrac{1}{2} e_o^T y_{xx} e_o + \eta \tag{2}$$

We seek a minimum variance linear estimate of $\hat{x}$ of the form

$$\hat{x} - \hat{x}_o = K[y(x, \eta) - y(\hat{x}_o)] \tag{3}$$

where K is the linear vector gain. The expected change in the error in the estimate is given by

$$e = e_o - K[y(x, \eta) - y(\hat{x}_o)] \tag{4}$$

The variance of e is given by

$$E(e, e^T) = E(e_o, e_o^T) - K\,E(\Delta y, e_o^T) - E(e_o, \Delta y^T)K^T + K\,E(\Delta y, \Delta y^T)K^T \tag{5}$$

The minimum variance estimate of e over all linear K gain vectors is given by

$$K_{mv} = \frac{E(\Delta y, e_o^T)}{E(\Delta y, \Delta y^T)} \tag{6}$$

If we now make the assumption that

$$E(e_o, \eta) = E(e_o, e_o^T y_{xx} e_o) = E(\eta, e_o^T y_{xx} e_o) = 0 \tag{7}$$

and that

$$E(e_o, e_o^T) = P$$
$$E(\eta, \eta^T) = Q$$

we have

$$K = \frac{P y_x}{y_x^T P y_x + Q + \frac{1}{4} E(e_o^T y_{xx} e_o, e_o^T y_{xx} e_o)} \tag{8}$$

From matrix algebra, we have, for any vector ℓ

$$\ell^T y_{xx} (\ell \ell^T) y_{xx} \ell = \text{trace} [y_{xx} (\ell \ell^T) y_{xx} (\ell \ell^T)] \tag{9}$$

We seek the expected value of the trace of the square of the matrix A, where

$$A = y_{xx} (e_o e_o^T) \tag{10}$$

Given any $n \times n$ matrix A, the trace of its square is given by

$$\text{trace } A^2 = \sum_{i=1}^{N} \sum_{j=1}^{N} a_{ij} a_{ji} \tag{11}$$

Every element of the matrix A is given by

$$a_{ij} = \sum_{k=1}^{N} y_{ik} e_{0k} e_{0j} \tag{12}$$

where y_{ik} is the ik^{th} element of y_{xx}.

We seek the expected value of

$$E \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{k=1}^{N} \sum_{\ell=1}^{N} y_{ik}\, y_{j\ell}\, e_{0k}\, e_{0j}\, e_{0\ell}\, e_{0i} \tag{13}$$

For a normally-distributed random variable with zero mean, we have

$$E(e_{0k}\, e_{0j}\, e_{0\ell}\, e_{0i}) = E(e_{0k}\, e_{0j})\, E(e_{0\ell}\, e_{0i}) + E(e_{0k}\, e_{0\ell})\, E(e_{0j}\, e_{0i}) + E(e_{0k}\, e_{0i})\, E(e_{0j}\, e_{0\ell}) \tag{14}$$

Let

$$p_{ij} = E(e_{0i}\, e_{0j}) \tag{15}$$

Then Eq. (13) becomes

$$\sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{k=1}^{N} \sum_{\ell=1}^{N} y_{ik}\, y_{j\ell} (p_{kj}\, p_{\ell i} + p_{k\ell}\, p_{ji} + p_{ki}\, p_{j\ell}) \tag{16}$$

If we define the matrix C to be the expected value of the matrix A

$$C = E(y_{xx}\, e_o\, e_o^T) = y_{xx}\, P \tag{17}$$

Equation (17) may be written as

$$\sum_{i=1}^{N} \sum_{j=1}^{N} C_{ij}\, C_{ji} + C_{ji}\, C_{ij} + C_{ii}\, C_{jj}) \tag{18}$$

from which we obtain

$$\Delta Q = \frac{1}{4} E \,\text{trace}\,(A^2) = \frac{1}{2} \,\text{trace}\,(C^2) + \frac{1}{4} (\text{trace}\, C)^2 \tag{19}$$

For the purposes of computation, we recommend that we first compute the nonzero elements of $C = y_{xx} P$ and then form ΔQ.

$$\Delta Q = \frac{1}{2} \sum_{i=1}^{N} C_{ii}^2 + \sum_{i=1}^{N} \sum_{j=i+1}^{N} C_{ij} C_{ji} + \frac{1}{4} \left[\sum_{i=1}^{N} (C_{ii}) \right]^2 \tag{20}$$

Example

Let y depend only upon position (e.g., in a range measurement, in DML, or in angle measurement). Then y_{xx} is an upper 3×3 and the trace of the expected value of $\frac{1}{4} A^2$ is given by

Let

$$y_{xx} P = C\,(3 \times 3) = \begin{bmatrix} C_{11} & C_{12} & C_{13} \\ C_{21} & C_{22} & C_{23} \\ C_{31} & C_{32} & C_{33} \end{bmatrix} \tag{21}$$

$$E\left(\frac{1}{4} \text{ trace } A^2\right) = \frac{1}{2} (C_{11}^2 + C_{22}^2 + C_{33}^2) + (C_{12} C_{21} + C_{13} C_{31} + C_{23} C_{32})$$

$$+ \frac{1}{4} (C_{11} + C_{22} + C_{33})^2 \tag{22}$$

APPROXIMATE MATRIX INVERSES

S. Pines

Report No. 73-47
Contract No. NAS 9-12516
November 1973

ANALYTICAL MECHANICS ASSOCIATES, INC.
50 JERICHO TURNPIKE
JERICHO, N. Y. 11753

SUMMARY

This report derives a procedure for a rapid determination of an approximate matrix inverse for use on real-time on-board computer control systems.

INTRODUCTION

On-board computers often require matrix inversions during real-time computer control system computations when the cycle time does not allow for a precise solution. The procedure outlined here yields an approximate solution which can be executed in considerably less computer time.

DERIVATION OF THE APPROXIMATION

Let A be an nxn matrix, and x and y be nx1 vectors. We seek an approximate solution of the problem: Given A and y, find x, when

$$Ax = y \tag{1}$$

Let B be the matrix formed of the diagonal elements of A.

$$b_{ij} = \delta_{ij} a_{ij} \tag{2a}$$

Then

$$A = B B^{-1} A \tag{2b}$$

The solution of Eq. (1) is given by

$$x = (B^{-1} A)^{-1} B^{-1} y \tag{3}$$

Let

$$C = B^{-1} A$$

and

$$z' = B^{-1} y \tag{4}$$

The matric, C, satisfies its own characteristic equation

$$\sum_{i=0}^{n} \alpha_i C^{n-i} = 0 \qquad (\alpha_0 = 1) \tag{5}$$

It follows from Eq. (5) that the entire n-space is annihilated. Thus,

$$\sum_{i=0}^{n} \alpha_i C^{n-i} z = 0 \tag{6}$$

and

$$\sum_{i=1}^{n} \alpha_i C^{n-i} z = - C^n z \tag{7}$$

Equation (7) can be used to determine the n unknown coefficients, α_i. Once these are determined, we have, after multiplication by C^{-1},

$$C^{-1} z = - \frac{1}{\alpha_n} \left\{ \sum_{i=0}^{n-1} \alpha_i C^{n-1-i} z \right\} \tag{8}$$

To realize a saving in computer time, we require an approximate characteristic equation for the matrix C, of much lower degree than n. Thus, we are led to find the p, β_i, scalars which minimize the length of the vector, ℓ.

$$\sum_{i=0}^{p} \beta_i C^{n-i} z = \ell \quad , \qquad p \ll n \tag{9}$$

The β's are given by the least squares solution for $\ell = 0$, with $\beta_0 = 1$. We form the Gram-Schmidt upper triangular decomposition of the $n \times p+1$ matrix of the vectors $z, Cz, \ldots, C^p z$. We have

$$\underset{(n \times p+1)}{(z, Cz, C^2 z, \ldots, C^p z)} = \underset{(n \times p+1)}{(\theta)} \underset{}{(T)} \tag{10}$$

where (θ) is an orthogonal matrix $(n \times p+1)$ and (T) is upper triangular matrix $(p+1 \times p+1)$.

The solution for β_i is given by

$$\left(T_{(p+1 \times p+1)}\right) \begin{bmatrix} \beta_p \\ \beta_{p-1} \\ \vdots \\ \beta_1 \\ 1 \end{bmatrix}_{(p+1 \times 1)} = 0 \tag{11}$$

Transposing the $p+1^{st}$ vector of (T), we have, for β_i

$$\begin{bmatrix} \beta_p \\ \beta_{p-1} \\ \vdots \\ \beta_1 \end{bmatrix} = -\left(T\ (p \times p)\right)^{-1} \begin{bmatrix} t_{1(p+1)} \\ t_{2(p+1)} \\ \vdots \\ t_{p(p+1)} \end{bmatrix} \tag{12}$$

The first few solutions are provided below.

Zero Order

$$x = z. \tag{13}$$

good only if $\dfrac{a_{ij}}{\sqrt{a_{ii}\, a_{jj}}} \ll 1$.

First Order

$$x = \frac{t_{11}}{t_{12}}\, z \tag{14}$$

Second Order

$$x = \frac{t_{23}\, t_{11}}{t_{12}\, t_{23} - t_{13}\, t_{22}}\, z + \frac{t_{11}\, t_{22}}{t_{12}\, t_{23} - t_{13}\, t_{22}}\, Cz \tag{15}$$

The Gram-Schmidt coefficients are listed below:

$$t_{ij} = \theta_j^T\, C^{i-1}\, z \qquad j < i$$

$$t_{ii} = \left[\left(C^{i-1} z - \sum_{j=1}^{i-1} t_{ij}\, \theta_j\right)^T \left(C^{i-1} z - \sum_{j=1}^{i-1} t_{ij}\, \theta_j\right)\right]^{1/2}$$

$$t_{11} = (z^T z)^{1/2}$$

$$\theta_i = \frac{1}{t_{ii}} \left\{ C^{i-1} z - \sum_{j=1}^{i-1} t_{ij}\, \theta_j \right\}$$

To obtain the upper triangular inverse of T, let S be the inverse of T, then the nonzero upper triangular elements of S are given by

$$s_{ii} = \frac{1}{t_{ii}}$$

$$s_{ij} = -\frac{1}{t_{jj}} \left(\sum_{k=1}^{j-1} s_{ik}\, t_{kj} \right) \qquad j > i$$

From the above, it follows that the β_i coefficients are given by

$$\beta_{p+1-i} = s_{i,p+1} \qquad i = 1, 2, \ldots, p$$

Since we are interested only in the ratio $\frac{\beta_i}{\beta_p}$, the $t_{p+1,p+1}$ coefficient need not be generated and may be arbitrarily set equal to unity.

A VARIABLE-METRIC ALGORITHM EMPLOYING LINEAR AND QUADRATIC PENALTIES

H. J. Kelley
I. L. Johnson, Jr.

ANALYTICAL MECHANICS ASSOCIATES, INC.
50 JERICHO TURNPIKE
JERICHO, N. Y. 11753

A VARIABLE-METRIC ALGORITHM EMPLOYING LINEAR AND QUADRATIC PENALTIES*

Henry J. Kelley+
Leon Lefton‡
Analytical Mechanics Associates, Inc., Jericho, N.Y.

and

Ivan L. Johnson, Jr.++
NASA Johnson Space Center, Houston, Texas

ABSTRACT

A variable-metric algorithm is described that 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 which attempts to maintain a positive-definite matrix of second partial derivatives of the function including penalty terms without generating the large positive eigenvalues traditionally attending 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.

---

* 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, Texas.

+ Vice President

‡ Senior Programmer/Analyst

++ Aerospace Technologist

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

INTRODUCTION

The arrival of variable-metric parameter optimization, the Davidon-Fletcher-Powell algorithm (Ref. 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 competition with 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.

But 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 (Ref. 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 is currently 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 (Ref. 3). Variable-metric projection schemes, making selective use of what amount to linear penalty functions, were eventually developed for the case of nonlinear constraints and proved workable in limited tests (Refs. 4 and 5). This class of variable-metric scheme has only seen 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 compared 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 of both 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.

A 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

$$\tilde{f} = f + g\lambda + \frac{1}{2}\, g\ K\ g^T \tag{1}$$

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 (Ref. 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 sought, but also to render the magnitudes of the constraint violations, $|g_i|$, small at the minimum.

Hestenes' Method of Multipliers (Ref. 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 auxiliaries 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 form of $\tilde{f}$ above 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, the quadratic ones supplementary and missing whenever advisable. The λ-vector components will be determined as the projection values every few cycles, and the K diagonal elements chosen generally so as to provide the second partial derivative matrix $\tilde{f}_{xx}$ with positive 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 (Ref. 1) and its batch-processor DFP modification (Ref. 8) applied to $\tilde{f}$ for major iterations. They bear a resemblance to the Method of Multipliers, differing from it in the determination of λ and k values. It is 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

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

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 (2)]. One would like this not locally, with λ the projection value, but at the constrained minimum where the projection λ coincides with the Lagrange multiplier vector; however, it would be difficult and expensive enough to calculate the local

second partials and the largest eigenvalue, so a less direct and more approximate approach is taken. The scheme proposed as follows takes advantage of the fact that there will usually 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

$$\lambda^* = -\left(g_x^T H_o g_x\right)^{-1} g_x^T H_o f_x \tag{3}$$

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

For a step determined by the modified DFP algorithm as

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

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 $f^{*'}$ 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 (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 $\tilde{f}$, can be expected to be positive as $f_{xx} + g_{xx}\lambda$ possesses at least n-m non-negative 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 $\tilde{f}$ 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_{j_i} given by

$$k_{j_i} = \frac{\beta_{j_i}^2 (c f^{*''}_{\max} - \bar{f}_i'')}{|g_{i+1} g_i'' + g_{i+1}'^2|} \quad , \quad \begin{matrix} i=1, \text{---}, n \\ j=1, \text{---}, m \end{matrix} \qquad (9)$$

where

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

Values are to be excluded from consideration when the two terms of the denominator in (9) are opposite in sign and nearly equal in magnitude; likewise when β given by (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 $\tilde{f}''$ equal to the guideline value $c f^{*''}_{max}$ for $\beta = 1$ and remain bounded for small β (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'' values used should be the larger of the values at beginning and end of the search segment for $f^{*''}$ and the smaller of the two values for $\bar{f}''$.

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

$$k_{j_\ell} = \frac{(c f^{*''}_{max} - f^{*''}_{min})}{\left| g_{j_\ell x} \right|} \tag{11}$$

where ℓ corresponds to the last cycle before k selection and $f^{*''}_{min}$ is taken as the smallest of the $f^{*''}$ values over a chosen number of batches, or zero, whichever is the lesser. The multiplicative constant $c \geq 1$ in the guideline value of $f^{*''}$ introduces a measure of conservatism to offset the possibility that none of the candidate values of $f^{*''}$ in the maximization determining $f^{*''}_{max}$ 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 $\tilde{f}_{xx}^{-1}$. No correction is made for changes in the sum $\lambda + Kg$ appearing in the second partials (2) as 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

$$H + \Delta H = H - \left(\frac{\Delta k_i}{1 + \Delta k_i g_{i_x}^T H g_{i_x}}\right) H g_{i_x} g_{i_x}^T H \tag{12}$$

which accounts for changes in the last term of eq. (2) via the Schur identity (Ref. 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:

$$f = x_1 + a_1 x_1^2 + a_2 x_1^3 \tag{13}$$

$$g = x_1 - b_1 x_2^2 - b_2 x_3^2 - b_3 x_2^4 \tag{14}$$

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 cross-section 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 slightly 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 as 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, as $f + g\lambda$ 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 $a_2 > 0$ large enough 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 are such as to offer 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 (9) and (11) to designate the guideline value of $f^{*''}$ was taken as unity in the comparison.

COMPUTATIONAL COMPARISON

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 in the accompanying table 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 the table are for sequential and batch-processor versions of the algorithm described in the proceding. 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 (Ref. 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.

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

POSSIBLE IMPROVEMENTS IN PENALTY-COEFFICIENT-ADJUSTMENT PROCEDURES

Two features intended to aid the process of adjusting linear and quadratic penalty coefficients will be described very briefly. These have been explored computationally and found to produce reasonable results, but evaluated insufficiently to permit overall judgment on their merits.

Early values of the linear penalty coefficients, the components of λ, by projection, tend to be far off the mark, which recommends use of a zero value during the first batch. A short first batch suggests itself, m cycles instead of n, mainly to reduce the magnitudes of the constraint violations. It might be hoped that a better metric would also be obtained as well, better at least in the subspace defined by the constraint gradient vectors. An obvious modification of the batch metric update formula of Ref. 8 is required.

Another 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 these two features to date, it has been found that they generally enhance 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.

CONCLUDING REMARKS

The results are thought to indicate promise for the class of algorithm combining linear and quadratic penalty adjustment with variable-metric optimization. More extensive testing is obviously needed, including the large class of problems in which the quadratic terms can safely be adjusted 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 $x = \hat{x}$.

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

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

$$\hat{g}_x(x - \hat{x}) = 0 \tag{A 2}$$

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

$$\lambda = -(\hat{g}_x^T H_o \hat{g}_x)^{-1} \hat{g}_x^T H_o (f_x + g_x\hat{\lambda}) \tag{A 3}$$

A linear search is made in the direction

$$\Delta x = -\alpha H (f_x + g_x\hat{\lambda} + \hat{g}_x\lambda) \tag{A 4}$$

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 using the gradient of $f + g\hat{\lambda}$:

$$H + \Delta H = H - \frac{H(\Delta f_x + \Delta g_x\hat{\lambda})(\Delta f_x + \Delta g_x\hat{\lambda})^T H}{(\Delta f_x + \Delta g_x\hat{\lambda})^T H (\Delta f_x + \Delta g_x\hat{\lambda})} + \frac{\Delta x\, \Delta x^T}{\Delta x^T (\Delta f_x + \Delta g_x\hat{\lambda})} \tag{A 5}$$

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

$$\Delta x = -\alpha H (f_x + g_x \lambda) \tag{B 1}$$

$$\lambda = -(g_x^T H g_x)^{-1} g_x^T H f_x \tag{B 2}$$

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

$$H + \Delta H = H + \frac{\Delta x \, \Delta x^T}{\Delta x^T (\Delta f_x + \Delta g_x \lambda)} - \frac{H(\Delta f_x + \Delta g_x \lambda)(\Delta f_x + \Delta g_x \lambda)^T H}{(\Delta f_x + \Delta g_x \lambda)^T H (\Delta f_x + \Delta g_x \lambda)} \tag{B 3}$$

which is the DFP formula applied to the linear combination $f + g\lambda$, hence guarantees that H remains positive definite. Constraint restorations are carried out after each optimization cycle (Ref. 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," A.I.Ch.E. 61st National Meeting, Houston, Texas, February 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. Rosen, J.B. and Kreuser, J.; "A Gradient Projection Algorithm for Nonlinear Constraints," presented at the Conference on Numerical Methods for Nonlinear Optimization, University of Dundee, Scotland, June 28-July 1, 1971.

6. Kelley, H.J.; "Method of Gradients," Chapter 6 of Optimization Techniques, G. Leitmann, ed., Academic Press, 1962.

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, No. 11, November 1970.

9. Johnson, I. L. and Kamm, J. L.; "Accelerating One-Dimensional Searches," AIAA Journal, Vol. 11, No. 5, May 1973.

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