Application of Broyden's Method to Reconciliation of Nonlinearly Constrained Data

C. C. David Pai, Gary D. Fisher
ChemShare Corporation
Houston, TX 77251

Process data reconciliation and gross error detection have been the subject of many recent publications, for which Tamhane and Mah (1985) and Mah (1987) have provided a thorough review. Analytical solutions to linear data reconciliation problems can be obtained efficiently, for instance, by matrix projection (Crowe et al., 1983). However, solving general data reconciliation problems with nonlinear constraints invariably requires some type of iterative procedure. In this note, an iterative procedure is developed that makes use of Crowe's matrix projection and for the first time combines a quasi-Newton update with the Gauss-Newton scheme for solving nonlinear data reconciliation problems.

Using the Lagrange method, Britt and Leucke (1973) developed the normal equations for nonlinear parameter estimation. More recently, Stephenson and Shewchuk (1986) and Serth et al. (1987) also based their data reconciliation problems on the normal equations. Although straightforward, the approaches based on solving the system of normal equations suffer from two drawbacks. First, the size of the problem may be unduly large due to the Lagrange multipliers; second, first-order derivatives of the constraints appear in the normal equations and have to be calculated for every iteration.

Crowe (1986) extended his method of matrix projection to bilinear constraints with an iterative procedure for determining one of the variables in the bilinear terms. Unfortunately, this procedure is too specific to be useful for more general cases.

Knepper and Gorman (1980) proposed a Gauss-Newton iterative algorithm and, in order to reduce computational effort, suggested using old derivatives of constraints until the constraints are satisfied (i.e., the constant-direction approach). However, their algorithm is limited to problems with no more constraints than measured variables and the constant-direction approach is characterized by slow convergence. Knepper and Gorman applied the theory of generalized inverses to solve the linearized subproblem. Crowe (private communication, 1987) presented a more general method based on matrix projection (Crowe et al., 1983). The latter not only effectively reduces the problem size but also removes the restriction that no more constraints than measured variables be handled. In this note, a Broyden-type update (Broyden, 1965) is proposed to replace the old derivatives so that the rate of convergence can be improved without repeatedly evaluating the derivatives.

Crowe's Iterative Scheme

A general data reconciliation problem is defined as

$$\min_{x,v} \; (x - \tilde{x})^T \Sigma^{-1} (x - \tilde{x})$$
$$\text{subject to } f(x, v) = 0 \tag{1}$$

where f is an m vector of functions, x is an n vector of measured variables, v is an r vector of unmeasured variables, and Σ is the variance-covariance matrix of measurements $\tilde{x}$, or some weighting matrix. The sizes of these vectors are related by $n + r \geq m > r \geq 0$. Linearizing the constraint functions f around x_k and v_k gives

$$B_k(x - x_k) + P_k(v - v_k) + f(x_k, v_k) = 0 \tag{2}$$

where B_k and P_k denote, respectively, the $(m \times n)$ and $(m \times r)$ matrices of the derivatives of f with respect to x and v evaluated at x_k and v_k.

Following Crowe et al. (1983), the solution (x_k^*, v_k^*) to the linear data reconciliation problem defined by Eq. 2 can be written directly:

$$x_k^* - \tilde{x} = \Sigma B_k^T Y_k H_k^{-1} Y_k^T [-f(x_k, v_k) + B_k(x_k - \tilde{x})] \tag{3}$$

$$v_k^* - v_k = -(P_k^{[1,r]})^{-1} W [f(x_k, v_k) + B_k(x_k^* - x_k)] \tag{4}$$

where Y_k is a matrix whose columns span the null space of P_k^T

that is, $P_k^T Y_k = 0$, W is an $(r \times m)$ matrix defined by

$$W = [I_r | 0] \tag{5}$$

the superscript $[\cdot, \cdot]$ of a matrix denotes the range in rows of a submatrix, and

$$H_k = Y_k^T B_k \Sigma B_k^T Y_k \tag{6}$$

It is assumed that the rank of P_k is r and that $P_k^{[1,r]}$ is a square nonsingular matrix. Y_k can be calculated as suggested by Crowe et al. (1983):

$$Y_k^T = [-P_k^{[r+1,m]}(P_k^{[1,r]})^{-1} | I_{m-r}] \tag{7}$$

For simplicity, define

$$P_k = \begin{bmatrix} x_k^* - x_k \\ v_k^* - v_k \end{bmatrix} \quad z_k = \begin{bmatrix} x_k \\ v_k \end{bmatrix} \tag{8}$$

An iterative algorithm can be easily constructed for solving the nonlinear data reconciliation problem, Eq. 1, as follows:

1. Set $k = 0$ and $x_k = \tilde{x}$. Obtain an initial estimate v_k.
2. Evaluate B_k and P_k.
3. Evaluate Y_k and H_k from Eqs. 7 and 6.
4. Compute p_k from Eqs. 3–5 and 8.
5. Starting with $t_k = 1$, search for $0 \le t_k \le 1$ such that

$$z_{k+1} = z_k + t_k p_k \tag{9}$$

reduces a prespecified penalty function for Eq. 1.

6. Stop if the penalty function is not significantly reduced. Otherwise go to step 7.
7. Reset $k = k + 1$, then go to step 2.

Updating of Derivatives

In the iterative scheme described above, B_k and P_k need to be reevaluated for each iteration (step 2). In most chemical process applications, however, the derivatives are estimated by finite differences of implicit functions, which are very time-consuming to evaluate. Therefore, the computing time for step 2 will likely be prohibitively large. This prompted Knepper and Gorman (1980) to use old values of derivatives until the constraints are satisfied. However, this approach leads to slow convergence rate and possibly a small radius of convergence.

Broyden (1965) developed a method of updating the Jacobian matrix in order to overcome the disadvantages of Newton's method for solving systems of equations. The idea behind Broyden's method is to use the function residuals at the new point to generate from the old matrix an approximation of the Jacobian at the new point so that repeated calculation of the Jacobian matrix can be avoided. In the following, Broyden's method will be extended to the case of nonlinear data reconciliation as treated by Crowe's iterative scheme above. Before that, the differences between the two cases were worth mentioning. First, while the Jacobian matrix is square, the matrix $[B|P]$ for data reconciliation is generally not. Second, unlike the Jacobian matrix, the matrix $[B|P]$ is not directly related to the adjustments of variables. Nevertheless, both matrices are essential to determining a search direction for the solution.

Let z be defined as

$$z(t) = z_k + t p_k \tag{10}$$

then $f(z)$ becomes a function of t, which is denoted $f_k(t)$. It follows from the chain rule that

$$df_k/dt = [B|P] p_k \tag{11}$$

and from finite differencing around z_{k+1} that

$$f_k(t_k - s_k) \simeq f(z_{k+1}) - s_k(df_k/dt) \tag{12}$$

where s_k is a small number. Eliminating df_k/dt from Eqs. 11 and 12 gives

$$y_k = f(z_{k+1}) - f_k(t_k - s_k) \simeq s_k [B|P] p_k \tag{13}$$

which relates $[B|P]$ to other known quantities. As in method 1 of Broyden (1965), let $[B|P]_{k+1}$ be updated in such a way as to satisfy

$$y_k = f(z_{k+1}) - f_k(t_k - s_k) = s_k [B|P]_{k+1} p_k \tag{14}$$

and

$$[B|P]_{k+1} q_k = [B|P]_k q_k, \ \forall q_k \text{ s.t. } q_k^T p_k = 0 \tag{15}$$

where Eq. 15 is the condition that the change in f predicted by $[B|P]_{k+1}$ in a direction q_k orthogonal to p_k is the same as would be predicted by $[B|P]_k$. It can be verified that the $[B|P]_{k+1}$ satisfying Eqs. 14 and 15 is uniquely given by

$$[B|P]_{k+1} = [B|P]_k + \frac{(y_k - s_k[B|P]_k p_k) p_k^T}{s_k p_k^T p_k} \tag{16}$$

With this updating scheme at hand, Crowe's iterative scheme can now be modified by replacing step 7 above with the following:

7. Generate $[B|P]_{k+1}$ by Eq. 16. Reset $k = k + 1$, then go to step 3.

Therefore, the full matrix is evaluated only once (step 2). The modified scheme is a compromise between the constant-direction approach and repeated computation of derivatives, and to some degree possesses the simplicity of the former and the efficiency of the latter.

The choice of s_k in Eq. 16 has been discussed by Broyden (1965). In practice, s_k is chosen to be the difference of the last two t_k values considered in Step 5; if only one t_k value is considered, then s_k is made equal to t_k, which is unity. The integrity of the update, Eq. 16, depends upon how well the condition of Eq. 14 holds. This is guaranteed by the fact that the norm of p_k becomes smaller and smaller as the solution is approached.

Broyden suggested updating the inverse of the Jacobian matrix through Householder's formula, since the inverse of the Jacobian is directly related to the adjustments of variables by Newton's method. In the present case, however, the adjustments of variables, given by Eqs. 3 and 4, are the solution of a quadratic program. The form of those equations does not provide any incentive for updating the inverse of $[B|P]$.

Numerical Example

Consider the following example:

Constraints

$$
\begin{aligned}
0.5x_1^2 - 0.7x_2 + x_3v_1 + x_2^2v_1v_2 + 2x_3v_3^2 - 255.8 &= 0 \\
x_1 - 2x_2 + 3x_1x_3 - 2x_2v_1 - x_2v_2v_3 + 111.2 &= 0 \\
x_3v_1 - x_1 + 3x_2 + x_1v_2 - x_3(v_3)^{1/2} - 33.57 &= 0 \\
x_4 - x_1 - x_3^2 + v_2 + 3v_3 &= 0 \\
x_5 - 2x_3v_2v_3 &= 0 \\
2x_1 + x_2x_3v_1 + v_2 - v_3 - 126.6 &= 0
\end{aligned}
$$

Measurements

$$\tilde{x} = [4.4, 5.5, 1.7, 1.6, 5]^T \text{ and } \Sigma = I_5.$$

Assumptions

1. $[B|P]_0$ and f_k ($k \geq 0$) are evaluated analytically
2. $v_0 = [8, 0.7, 1.8]^T$
3. Penalty function $= \|f_k\|_2 + 0.1\,(x - \tilde{x})^T\Sigma^{-1}(x - \tilde{x})$

This problem is solved using the above iterative scheme with Broyden's updates and with constant derivatives, respectively. For the latter, the constant-direction approach of Knepper and Gorman (1980) is implemented in effect. Note that their formulae cannot be applied directly to this problem for n is smaller than m. The results are shown in Figure 1 and Table 1. Figure 1 shows the reduction of the penalty function with the number of iterations. Since the constraints are well satisfied at the end of either case, the difference in the final penalty function value is mainly due to the quadratic term of measurement adjustments. The run with constant derivatives fails to reduce this term comparably, even though an extra evaluation of $[B|P]$ is made following iteration 10. Comparisons using three different criteria are summarized in Table 1. Therein $\hat{z}_{exact}$ is the solution obtained using exact derivatives for every iteration. All of the three comparisons indicate that the scheme with Broyden's updates proposed in this work is more efficient than the constant-direction approach.

Table 1. Summary of Comparisons

Criterion	Broyden's Updates	Constant $[B\,|\,P]$
1. Final penalty function	0.0112888	0.0140832
2. Number of f evaluations	21	29*
3. $\|\hat{z} - \hat{z}_{exact}\|_2$**	$6.33 \cdot 10^{-4}$	$6.83 \cdot 10^{-2}$

*One evaluation for each of the eight variables is accounted for the extra gradient evaluation following iteration 10

**$\hat{x}_{exact} = [4.5124, 5.5819, 1.9260, 1.4560, 4.8545]^T$

$\hat{v}_{exact} = [11.070, .61467, 2.0504]^T$

Conclusions

Experimenting with the above iterative scheme on a range of numerical problems leads to the following conclusions:

1. In most cases, one full evaluation of $[B|P]$ initially is sufficient for the algorithm to reach a good solution.

2. The convergence rate is faster than when the constant-direction approach is adopted.

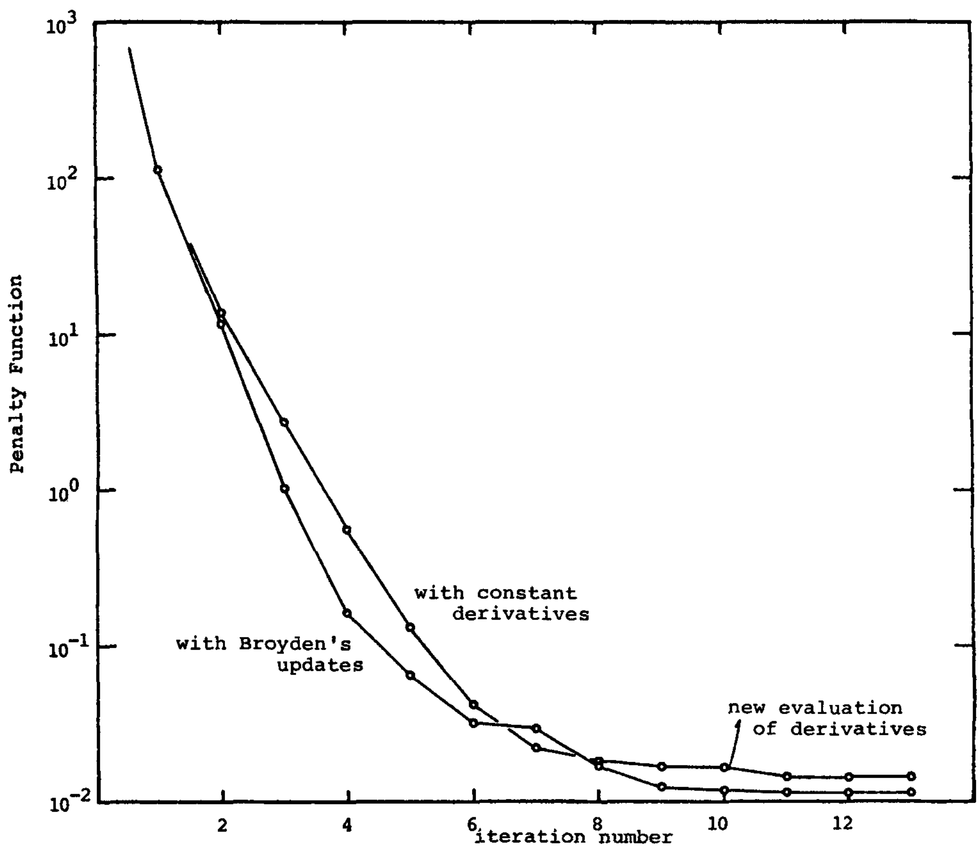

Figure 1. Reduction of penalty function by using Broyden's updates or constant derivatives.

3. If z_0 is far away from the solution, step 5 may fail. When this happens, a second full evaluation of $[B|P]$ is recommended for a restart from z_k. Nevertheless, the total number of full evaluations is mostly less than will be needed for the constant-direction approach.

Data reconciliation has been the emphasis of this note. However, the development here is also important to detection of gross errors. One of the most effective gross error detection methods is the measurement test proposed by Mah and Tamhane (1982). It provides direct identification of gross errors but requires data reconciliation first. Without a reasonably efficient data reconciliation algorithm, it will not be very attractive. Combined with the improved data reconciliation algorithm developed in this note, the measurement test is expected to find wider applications to nonlinear cases.

The update, Eq. 16, has been developed with Crowe's iterative procedure in mind; however, it can also be applied to other iterative procedures that require constraint derivatives.

Notation

$\mathbf{0}$ = zero vector or matrix
B = $(m \times n)$ matrix of derivatives of f with respect to x
f = m-vector of constraint functions
H = matrix, Eq. 6
I = identity matrix
k = iteration counter
m = number of constraints
n = number of measured variables
p = $(n + r)$ vector, Eq. 8
P = $(m \times r)$ matrix of derivatives of f with respect to v
q = $(n + r)$ vector orthogonal to p
r = number of unmeasured variables
s = a small number for differencing
t = a number between 0 and 1, Eq. 9
v = r vector of unmeasured variables
v^* = reconciled values of v corresponding to Eq. 2
W = $(r \times m)$ matrix, Eq. 5
x = n vector of measured variables
x^* = reconciled values of x corresponding to Eq. 2
y = m vector of changes in f along p
Y = matrix defined by $P^T Y = \mathbf{0}$
z = $(n + r)$ vector of all variables
Σ = variance-covariance matrix of $\tilde{x}$ or weighting matrix.

Subscripts

2 = Euclidean norm
k = iteration counter

Superscripts

T = transpose
$[a, b]$ = range in rows for a submatrix
$\tilde{}$ = measured value
$\hat{}$ = adjusted value

Literature Cited

Britt, H. I., and R. H. Leucke, "The Estimation of Parameters in Nonlinear Implicit Models," *Technometrics*, **15**, 233 (1973).

Broyden, C. G., "A Class of Methods for Solving Nonlinear Simultaneous Equations," *Math. Comp.*, **19**, 577 (1965).

Crowe, C. M., "Reconciliation of Process Flow Rates by Matrix Projection. II: The Nonlinear Case," *AIChE J.*, **32**, 616 (1986).

Crowe, C. M., Y. A. G. Campos, and A. Hrymak, "Reconciliation of Process Flow Rates by Matrix Projection. I: Linear Case," *AIChE J.*, **29**, 881 (1983).

Knepper, J. C., and J. W. Gorman, "Statistical Analysis of Constrained Data Sets," *AIChE J.*, **26**, 260 (1980).

Mah, R. S. H., "Data Screening," *Proc. FOCAPO-87*, Park City, UT (1987).

Mah, R. S. H., and A. C. Tamhane, "Detection of Gross Errors in Process Data," *AIChE J.*, **28**, 828 (1982).

Serth, R. W., C. M. Valero, and W. A. Heenan, "Detection of Gross Errors in Nonlinearly Constrained Data: A Case Study," *Chem. Eng. Commun.*, **51**, 89 (1987).

Stephenson, G. R., and C. F. Shewchuk, "Reconciliation of Process Data with Process Simulation," **32**, 247 (1986).

Tamhane, A. C., and R. S. H. Mah, "Data Reconciliation and Gross Error Detection in Chemical Process Networks," *Technometrics*, **27**, 409 (1985).

Manuscript received Sept. 28, 1987, and revision received Nov. 19, 1987.