DEPARTMENT OF EARTH AND PLANETARY SCIENCES

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
CAMBRIDGE, MASSACHUSETTS 02139

54-620

24 January 1974

74-01

TO: Distribution

FROM: R. D. Reasenberg

SUBJECT: Convergence indicator in PEP

The ADJUST link of PEP now generates and prints several numbers which are intended to give the user a measure of the size of the adjustment and, when there is more than one adjustment, to indicate the rate of convergence. Historically the first of these was the FRACT vector, the n^{th} element of which is the adjustment of the n^{th} parameter, Δ_n, divided by the corresponding computed standard σ_n. With large numbers of parameters being adjusted, the FRACT vector becomes hard to interpret. Therefore an additional set of quantities is now calculated. These serve to summarize the FRACT vector and include the RMS, average, average absolute, sum square, and average square FRACT for both the entire set of adjusted parameters and for subsets of parameters. (There is a fixed hierarchy of parameter subsets.) In the following paragraphs, a new (to PEP) measure of the estimator step size in introduced, with the expectation that it will be added to PEP.

A metric for the parameter space which gives a "natural" norm or measure of the length of an adjust vector is the inverse covariance matrix

$$N \equiv ||\Delta|| = \sqrt{\Delta^{\dagger} P^{-1} \Delta} \tag{1}$$

This quantity is invariant under nonsingular linear transformations of the parameter set. It is also similar to the familiar RMS FRACT and, in particular, if P is replaced by P*

$$P^*_{ij} = \begin{cases} P_{jj} & i = j \\ 0 & i \neq j \end{cases} \tag{2}$$

$$N^* = \sqrt{\Delta^\dagger (P^*)^{-1} \Delta} \tag{3}$$

the new norm, N^*, is the RSS FRACT. Although N^* is invariant under parameter scaling, it is not invariant under parameter rotations.

The expression for N can be further simplified and better understood by using the basic equations of the least squares fit.

$$\begin{aligned} A &= \partial z/\partial \alpha \\ P^{-1} &= A^\dagger R^{-1} A \\ u &= A^\dagger R^{-1} r \\ \Delta &= Pu \end{aligned} \tag{4}$$

where z is the observable, α is a parameter, R is the data-noise covariance, P is the computed parameter covariance, r is the pre-fit residual vector, u is the right-hand side of the normal equations, and Δ is the parameter adjustment vector. Using the last of these, the norm (Equation 1) takes the form

$$N = \sqrt{\Delta^\dagger u} \tag{5}$$

Thus the quantity of interest is, in principle, quite easy to calculate. It may be useful at a later date to use N as a formal convergence measure with the iterations stopping after N^2 has fallen below a predetermined value.

We next proceed to examine N^2 in connection with the data residuals. From the adjustment and Equation 4, we can find the "predicted residual"

$$\begin{aligned} r_p &= r - A\Delta \\ &= r - A(A^\dagger R^{-1} A)^{-1} A^\dagger R^{-1} r \\ &= (I-F)r \end{aligned} \tag{6}$$

where F is a projection operator.

$$F = A(A^{\dagger}R^{-1}A)^{-1}A^{\dagger}R^{-1}$$

$$= A(A^{\dagger}R^{-1}A)^{-1}A^{\dagger}R^{-1}A(A^{\dagger}R^{-1}A)^{-1}A^{\dagger}R^{-1} = FF \qquad (7)$$

Again using Equation 4

$$N^2 = \Delta^{\dagger}u$$

$$= r^{\dagger}R^{-1}A(A^{\dagger}R^{-1}A)^{-1}A^{\dagger}R^{-1}r$$

$$= r^{\dagger}R^{-1}Fr \qquad (8)$$

The weighted sum squares of the predicted postfit residuals, X_p^2, is found from Equation 6.

$$X_p^2 = r^{\dagger}(I-F)^{\dagger}R^{-1}(I-F)r \qquad (9)$$

To evaluate this we note that

$$F^{\dagger}R^{-1} = R^{-1}A(A^{\dagger}R^{-1}A)^{-1}A^{\dagger}R^{-1} = R^{-1}F \qquad (10)$$

Then

$$(I-F)^{\dagger}R^{-1}(I-F) = R^{-1}(I-F)(I-F) = R^{-1}(I-F)$$

$$X_p^2 = r^{\dagger}R^{-1}r - r^{\dagger}R^{-1}Fr$$

$$= X_o^2 - N^2 \qquad (11)$$

where X_o^2 is the prefit weighted sum of the residuals squared. Thus X_p^2 can be calculated trivially after each iteration without entering the PRDICT link and can be used as a check on both convergence and linearity.

RDR/jbu

Distribution: M. E. Ash
I. I. Shapiro
R. D. White