DEPARTMENT OF EARTH AND PLANETARY SCIENCES

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
CAMBRIDGE, MASSACHUSETTS 02139

54-612

28 July 1975 75-2

TO: Distribution

FROM: R. D. Reasenberg

SUBJECT: The PEP A Priori Facility

In the early part of 1973, R. D. White added to PEP a general-purpose, convenient-to-use a priori facility. The input may contain either the covariance, P_o, or inverse covariance, B_o, matrix in either triangular or diagonal form. It may also contain parameter estimates.

The purpose of the memorandum is to document the equations used and to indicate some possible uses for the facility.

Maximum Likelihood Equations with A Priori Information

Bayes' rule provides a manipulative tool for the analysis of probabilities:

$$[A/B][B] = [A,B] \tag{1}$$

where the three terms are to be read: 1) the probability of A given B; 2) the probability of B; and 3) the joint probability of A and B. (The square brackets will be used throughout for probabilities.)

For a maximum likelihood estimate, the object is to find the values, $\hat{x}$, of the parameters which maximizes the likelihood function

$$P = [x/y,x_o] \tag{2}$$

where the y are the data, x_o is an a priori estimate of the parameters, and

$$
\begin{aligned}
y_i &= H_i(x) + \nu_i \\
\langle \nu_i \rangle &= 0 \\
\langle \nu_i \nu_j \rangle &= R_{ij} \\
\langle x - x_0 \rangle &= 0 \\
\langle (x-x_o)_i (x-x_o)_j \rangle &= (P_o)_{ij}
\end{aligned} \tag{3}
$$

It is not possible to write an explicit expression for $\mathbb{P}$ using Equation (3) without some manipulation. Using Equation (1):

$$
\begin{aligned}
\mathbb{P} &= [y, x_o/x][x]/[y, x_o] \\
&= [y/x][x_o/x][x]/[y, x_o]
\end{aligned} \tag{4}
$$

We consider these four terms separately: The first term is the probability density of the data given the parameters. It is determined by the first three lines of Equation (3), assuming that the noise, ν, is gaussian.

$$
\begin{aligned}
[y/x] &= N_y \exp[-J_y/2] \\
J_y &= (y-H(x))^{\dagger} R^{-1} (y-H(x))
\end{aligned} \tag{5}
$$

where N_y is a normalization

The second term is also determined by Equation (3)

$$
\begin{aligned}
[x_o/x] &= N_x \exp[-J_x/2] \\
J_x &= (x-x_o)^{\dagger} P_o^{-1} (x-x_o)
\end{aligned} \tag{6}
$$

again making a gaussian assumption.

The third and fourth terms are ordinarily ignored, and we shall soon follow this practice. Term 3 represents the probability density when nothing is known. It is thus a constant within the allowed domain of x and zero outside (for some canonical choice of parameters). Clearly then by a change of variables $[x]$ can take on a more interesting form. Failure to include this term when it is not a constant can result in a biased estimator. (The degree of bias depends on the change in $[x]$ on the scale set by the postfit covariance, P, in the region of solution $\hat{x} \approx x$.) Such "phase space" considerations are not useful for the purposes of this memorandum. Similarly, one can ignore the fourth term which represents data-dependent sampling biases and a "phase space" correction to the distribution of the a priori estimate.

Returning to the problem of finding a maximum likelihood estimate, we can now write $\mathbb{P}$ explicitly and differentiate to find its maximum at $x = \hat{x}$ for the linearized problem. Using Equations (5) and (6) in Equation (4)

$$\mathbb{P} = N \exp\,[-J_y/2 - J_x/2] \tag{7}$$

$$\frac{\partial \mathbb{P}}{\partial x} = -\frac{1}{2}\,\mathbb{P}\,\frac{\partial J}{\partial x} = 0 \tag{8}$$

From $\partial J/\partial x\Big|_{x=\hat{x}} = 0$ we get

$$-A^{\dagger}R^{-1}(r - A(\hat{x}-\bar{x})) + P_o^{-1}(\hat{x}-x_o) = 0$$

$$A^{\dagger}R^{-1}A(\hat{x}-\bar{x}) + P_o^{-1}(\hat{x}-\bar{x}) + P_o^{-1}(\bar{x}-x_o) = A^{\dagger}R^{-1}r$$

$$\hat{x}-\bar{x} = [A^{\dagger}R^{-1}A + P_o^{-1}]^{-1}[A^{\dagger}R^{-1}r - P_o^{-1}(\bar{x}-x_o)] \tag{9}$$

In the above expression: r is the prefit residual

$$r = y - H(\bar{x})$$

$\bar{x}$ is the prefit nominal value of x around which the equations are linearized; and A is the sensitivity matrix (also referred to as the "observable partials")

$$A = \partial H(x)/\partial x\Big|_{x=\bar{x}}$$

PEP Implementation

The equations obtained in Section II were used by R. D. White in adding the "*a priori* facility" to PEP. This facility is enabled by setting ICT(44) $\neq$ 0 and providing the needed input control cards on FORTRAN LOGICAL UNIT NUMBER ICT(44). A description of these control cards can be found in PEP subroutine ACMIN.

Uses

In this section classes of applications of the *a priori* facility are discussed. No attempt has been made to create an exhaustive list of such applications and it is expected that PEP users will invent many not included here.

A. The classical use of *a priori* information in the analysis of experiment data is to conveniently include some of the results from one or more other experiments. Thus, for example, when estimating station locations using spacecraft tracking data, ground survey results for the separation of stations in a cluster could be included.

B. By a simple extension, one may include in the parameter set additional parameters which: 1) would be poorly determined by the experiment but 2) are not well enough known *a priori* so as not to pose the threat of biasing the estimates of the quantities of interest. The *a priori* information is then included and a (presumed) better estimate of all of the parameters results.

C. A linear (soft) constraint can be imposed on two or more parameters by including the properly formed "*a priori* information." The B_o matrix is hand formed using one pseudo-observation equation for each constraint needed. By making the pseudo-observable assumed noise small, the constraint can be made as hard as needed.

D. The *a priori* facility can be used to study the sensitivity of an estimate of the parameters x_α to changes in the values assumed for parameters x_β which would not normally be estimated. Such a study requires a reference run and one or more test runs. In the reference run all x_β are given strong *a priori* estimates at their nominal value. In each test run the nominal value of a single element of x_β is changed. From the resulting set of estimates one obtains

$$\frac{\Delta x_\alpha}{\Delta x_\beta} \approx \frac{\partial x_\alpha}{\partial x_\beta}$$

E. The use of the *a priori* facility to help in combining diverse data sets is discussed in the memorandum Distribution/RDR dated 10 Feb 1975.

RDR/jbu

Distribution:

M.E.Ash
L.E.Conrad
C.C.CounselmanIII
R.B.Goldstein
S.A.Gourevitch
L.R.Hazelton,Jr.
G.H.Pettengill
I.I.Shapiro
R.D.White