BELSLEY, D. A. (1984b). Collinearity and forecasting. *Forecasting* **3** 183–196.
BELSLEY, D. A. (1986a). Centering, the constant, first-differencing, and assessing conditioning. In *Model Reliability* (E. Kuh and D. Belsley, eds.). M.I.T. Press, Cambridge.
BELSLEY, D. A. (1986b). Model selection in regression analysis, regression diagnostics and prior knowledge (with discussion). *Internat. J. Forecasting* **2** 41–46.
BELSLEY, D. A. (1986c). Modelling and forecasting reliability. Working paper, Center for Computational Research in Economics and Management Science, M.I.T.
BELSLEY, D. A. and OLDFORD, R. W. (1986). The general problem of ill-conditioning and its role in statistical analysis. *Comput. Statist. Data Anal.* **4** 103–120.
GUNST, R. F. (1983). Regression analysis with multicollinear predictor variables. *Comm. Statist. A—Theory Methods* **12** 2217–2260.
HENDRY, D. F. (1980). Econometrics—alchemy or science? *Economica* **47** 387–406.
LEAMER, E. E. (1978). *Specification Searches.* Wiley, New York.
SIMON, S. D. and LESAGE, J. P. (1986). The impact of collinearity involving the intercept term on the numerical accuracy of regression. Working paper, Dept. Applied Statistics and Operations Research, Bowling Green State Univ.

Comment

Ronald A. Thisted

The statistics profession is fortunate indeed to have such a friend as Professor Stewart. He has repeatedly taken the time and energy to inform statisticians about the relevance of numerical analysis to their day-to-day work, and he has also taken the trouble to understand and to explicate some of our problems from our own point of view. This paper is an example of what numerical analysis can have to say about statistical problems, and it shows that there is a lot that we statisticians can profit from. In particular, Professor Stewart greatly improves our understanding both of collinearity and of one indicator of collinearity—the variance inflation factor.

As is true of most important papers, this one raises as many questions as it answers. I would like to comment on three issues that Professor Stewart only touched on. First, although Stewart would relegate the condition number $\kappa = \| X \| \cdot \| X^\dagger \|$ to the dustbin for statistical purposes, there is an important statistical interpretation which rescues it. Second, Stewart's procedures for using collinearity diagnostics depend upon a measure ι_j of the importance of the jth regressor variable. The notion of relative importance of a regressor is an elusive one, however, particularly when collinearity is present. Finally, I discuss the question of whether statisticians should want collinearity diagnostics at all, and if so, what we should want from them. Where possible, I adopt Stewart's notation. References to equations in his paper are preceded by the letter "S."

Ronald A. Thisted is Associate Professor, Department of Statistics, The University of Chicago, 5734 University Avenue, Chicago, Illinois 60637.

1. THE CONDITION NUMBER

Stewart gives a clear description of the numerical relevance of the condition number κ. In numerical analysis, its primary significance is the inequality (S-3.4), the righthand side of which gives a good indication of the effect of *numerical* errors in the regressors on the regression coefficients themselves. Because the *statistical* errors represented by e in the regression model (S-2.1) are generally much larger in magnitude than the numerical errors resulting from rounding and truncation, the bound from (S-3.4) is often so pessimistic as to be useless. In addition, the condition number is not invariant with respect to rescaling columns of X, so that interpretation of κ is dependent on the way in which X has been scaled. Although Stewart discusses three alternatives for scaling X—equal column scaling of X, scaling X to produce equal column scaling of E, and implicitly, scaling X so that the components of β are roughly equal in size—he finds no single choice compelling.

The condition number of X has an important statistical interpretation in the regression problem which is generally overlooked. Consider an arbitrary linear combination of the estimated regression coefficients, say $\hat{\alpha} = v'\hat{\beta}$. The variance of $\hat{\alpha}$ is given by

$$\text{Var}(\hat{\alpha}) = \sigma^2 v'(X'X)^{-1}v = \sigma^2 \| v'X^\dagger \|^2. \tag{1.1}$$

From this computation it is apparent that the linear combination with smallest variance (subject to the constraint, say, that $\| v \| = 1$) has variance $\sigma^2[\inf(X^\dagger)]^2$. The coefficients v_1 which achieve this minimum value explicitly give the linear combination $\hat{\alpha}_1 = v_1'\beta$ about which the regression data are most