To get a lower bound on the second term in (A.1), we use the fact that $\inf(R_{**}^{-1}) = \| R_{**} \|^{-1}$, from which it follows that

$$\begin{aligned} \| R_{**}^{-1} t_{*p} \| &\geq \inf(R_{**}^{-1}) \| r_{*p} \| \\ &= \| R_{**} \|^{-1} \| r_{*p} \| \\ &\geq \| R_{**} \|_F^{-1} \| r_{*p} \|. \end{aligned}$$

Since the columns of R_{**} have norm one, $\| R_{**} \|_F^2 = p - 1$ and $\kappa_p^{-2} = \rho_{pp}^2 = 1 - \| r_{*p} \|^2$. Hence

$$\| R_{**}^{-1} r_{*p} \|^2 \geq \frac{1 - \kappa_p^2}{p - 1}. \tag{A.3}$$

Combining (A.1), (A.2), and (A.3) we get

$$(p - 1) \max_{i \neq j} \kappa_i^2 \geq \sum_{i \neq j} \kappa_i^2 \geq p - 1 + \frac{\kappa_p^2 - 1}{p - 1},$$

which is equivalent to (4.4).

REFERENCES

ANDERSON, T. W. (1984). Estimating linear statistical relationships. *Ann. Statist.* **12** 1–45.

BEATON, A. E., RUBIN, D. B. and BARONE, J. L. (1976). The acceptability of regression solutions: another look at computational accuracy. *J. Amer. Statist. Assoc.* **71** 158–168.

BELSLEY, D. A. (1984a). Demeaning conditioning diagnostics through centering (with discussion). *Amer. Statist.* **38** 73–93.

BELSLEY, D. A., KUH, E. and WELSCH, R. E. (1980). *Regression Diagnostics.* Wiley, New York.

DANIEL, C. and WOOD F. S. (1980). *Fitting Equations to Data*, 2nd ed. Wiley, New York.

DAVID, N. A. and STEWART, G. W. (1982). Significance testing in a functional model. Technical Report 1204, Dept. Computer Science, Univ. Maryland.

DAVIES, R. B. and HUTTON, B. (1975). The effects of errors in the independent variables in linear regression. *Biometrika* **62** 383–391.

DONGARRA, J. J., BUNCH, J. R., MOLER, C. B. and STEWART, G. W. (1979). *The LINPACK User's Guide.* SIAM, Philadelphia.

ECKART, G. and YOUNG, G. (1936). The approximation of one matrix by another of lower rank. *Psychometrika* **1** 211–218.

GAUSS, C. F. (1821). Theoria Combinationis Observationum Erroribus Minims Obnoxiae: Pars Prior. In *Werke* 4. Königlichen Gesellschaft der Wissenschaften zu Göttingen, 1880.

GOLUB, G. H., HOFFMAN, A. and STEWART, G. W. (1984). A generalization of Eckhart-Young-Mirsky matrix approximation theorem. To appear in *Linear Algebra and Its Applications.*

GOLUB, G. H. and VAN LOAN, C. F. (1983). *Matrix Computations.* Johns Hopkins, Baltimore, Md.

GOLUB, G. H. and WILKINSON, J. H. (1966). Note on the iterative refinement of least squares solution. *Numer. Math.* **9** 139–148.

HODGES, S. D. and MOORE, P. G. (1972). Data uncertainties and least squares regression. *Appl. Statist.* **21** 185–195.

MARQUARDT, D. W. (1970). Generalized inverses, ridge regression, biased linear estimation, and nonlinear regression. *Technometrics* **12** 591–613.

MIRSKY, L. (1960). Symmetric gauge functions and unitarily invariant norms. *Quart. J. Math.* **11** 50–59.

SEBER, G. A. F. (1977). *Linear Regression Analysis.* Wiley, New York.

STEWART, G. W. (1974). *Introduction to Matrix Computations.* Academic, New York.

STEWART, G. W. (1977). On the perturbation of pseudo-inverses, projections, and linear least squares problems. *SIAM Rev.* **19** 634–666.

STEWART, G. W. (1983). A nonlinear version of Gauss's minimum variance theorem with applications to an errors-in-the variables model. Technical Report TR 1263, Dept. Computer Science, Univ. Maryland.

STEWART, G. W. (1984). Rank degeneracy. *SIAM J. Sci. Statist. Comput.* **5** 403–413.

SWINDEL, B. F. and BOWER, D. R. (1972). Rounding errors in the independent variables in a general model. *Technometrics* **14** 215–218.

TURING, A. M. (1948). Rounding-off errors in matrix processes. *Quart. J. Mech. Appl. Math.* **1** 287–308.

VAN DER SLUIS, A. (1969). Condition numbers and equilibration of matrices. *Numer. Math.* **14** 14–23.

WILKINSON, J. H. (1963). *Rounding Errors in Algebraic Processes.* Prentice-Hall, Englewood Cliffs, N.J.

WOODS, H., STEINOUR, H. H. and STARKE, H. R. (1932). Effect of composition of Portland cement on heat evolved during hardening. *Indust. Engrg. Chem.* **24** 1207–1214.

Comment

Donald W. Marquardt

Statisticians and numerical analysts owe a large debt of gratitude to Dr. Stewart for his demonstration and lucid exposition of the mathematical connection between the condition number and the parameter variance inflation factors. In doing so, he has also clarified the reasons why the condition number is not really helpful in the multiple regression context, nor in many other contexts. The insights he provides in this paper are important for all statisticians, because collinearity problems occur in many statistical contexts, including multiple linear regression, nonlinear regression, unbalanced analysis of variance, and estimation from inverse integral transform models. In this brief commentary I have selected three facets of Dr. Stewart's paper for discussion.

Donald W. Marquardt is Consultant Manager, Applied Statistics Group, Engineering Service Division, E. I. Du Pont De Nemours & Company, Wilmington, Delaware 19898.

HISTORY

The presentation by Stewart would leave the impression that when during the 1960s, I selected the name "variance inflation factors," the connection between parameter variance inflation and collinearity was not well understood. For example, Stewart claims (Section 4) "variance inflation factors and multiple correlations were not introduced to analyze collinearity in regression models, and their names show it." Quite the opposite. I chose the name to emphasize the critical *effect* which all practicing statisticians would understand, rather than a name that would emphasize the *cause* (i.e., collinearity), which was not yet so widely understood. However, I was fully aware of the general algebraic connection between the effect and its cause. Thus, in my 1970 paper (page 606) I emphasized that "the inflation factors depend on the partial correlation of each X with the other Xs." I also showed the algebraic relationship between the inflation factor and the correlation between the Xs for a simple example with only two Xs, and noted that "in problems larger than 2×2 the variance inflation factors are not usually equal for all parameters," and that "in larger problems attention is focused on the largest parameter variance inflation factor." I am sure that Cuthbert Daniel also understood the algebraic relationship.

Moreover, the usefulness of the eigenvalues and eigenvectors of $(X^TX)^{-1}$ for diagnosing the detailed structure of the collinearities was well understood, as described in the 1970 example (pages 606 and 607).

NOMENCLATURE

Stewart's collinearity indices are simply the square roots of the corresponding variance inflation factors. It is not clear to me whether giving a new name to the square root of a VIF is a help or a hindrance to understanding. There is a long and precisely analogous history of using the term "standard error" for the square root of the corresponding "variance." Given the continuing necessity for dealing with statistical quantities on both the scale of the observable and the scale of the observable squared, there may be a place for a new term. Clearly, the essential intellectual content is identical for both terms. However, with Stewart's proposed name we have the situation where we create the misleading impression that the variance inflation factors measure one thing, whereas the collinearity indices measure something else. Can we count on software producers to always display both quantities and both labels, and their close relationship? I think not. I would prefer for the square roots a name that focuses on the effect and is self-defining. That would be to name them parameter "standard error inflation factors."

CENTERING OF PREDICTOR VARIABLES

Stewart correctly notes that although the variance inflation factors (and their square roots) are invariant with scale factor changes in the columns of X, they are not invariant with changes of origin of the predictor variables. He points out the ability of centering to remove what I have called "nonessential ill-conditioning, thus reducing the variance inflation in the coefficient estimates" (Marquardt and Snee, 1975, page 3).

Stewart also discusses the numerical example from Belsley (1984). As Stewart notes, the diagnostic results from this example should give one pause about the model proposed by Belsley for the data. I fully agree with Stewart that "when there is a constant term in the model, the model should be centered before the importance of the remaining variables is assessed" and the "centering simply shows the variable for what it is." An analysis of the importance, and the statistical-inferential basis of centering is given in an extended discussion of Belsley's paper (cf. Snee and Marquardt, 1984) and in an extended discussion of an earlier paper (cf. Marquardt, 1980).

SUMMARY

The present paper by Stewart summarizes a conceptual breakthrough relating variance inflation factors to the condition number. The variance inflation factors (or their square roots) are the measures of choice for assessing the structure of the predictor variables in a data set when estimating the parameters of a specified linear model in a relevant domain.

ADDITIONAL REFERENCES

MARQUARDT, D. W. (1980). You should standardize the predictor variables in your regression models. Discussion of "A critique of some ridge regression methods" by G. Smith and F. Campbell. *J. Amer. Statist. Assoc.* **75** 87–91.

MARQUARDT, D. W. and SNEE, R. D. (1975). Ridge regression in practice. *Amer. Statist.* **29** 3–20.

SNEE, R. D. and MARQUARDT, D. W. (1984). Collinearity diagnostics depend on the domain of prediction, the model, and the data. Discussion of "Demeaning conditioning diagnostics through centering" by D. A. Belsley. *Amer. Statist.* **38** 83–87.