CC/NUMBER 12
MARCH 21, 1983

This Week's Citation Classic

Marquardt D W. Generalized inverses, ridge regression, biased linear estimation, and nonlinear estimation. *Technometrics* **12**:591-612, 1970.
[Engineering Department, E.I. du Pont de Nemours & Co., Wilmington, DE]

A class of biased statistical estimators employing generalized inverses is introduced. A unifying perspective is established by developing theoretical properties shared with ridge estimators, Bayesian estimators, and nonlinear estimation procedures. An example illustrates the procedures and introduces the term parameter variance inflation factors. [The *Science Citation Index*® (*SCI*®) and the *Social Sciences Citation Index*® (*SSCI*®) indicate that this paper has been cited in over 145 publications since 1970.]

Donald W. Marquardt
Engineering Department
E.I. du Pont de Nemours & Co.
Wilmington, DE 19898

December 23, 1982

"The long title of my 1970 paper highlighted its historical antecedents. Its developments were a synthesis of several separate lines of prior work: (a) my work[1] on algorithms for nonlinear estimation, (b) the growing use of generalized inverses to solve linear problems of less than full rank, and (c) ridge regression, developed by A.E. Hoerl and R.W. Kennard.[2,3] I have discussed in a previous *Citation Classic*[4] the background of my work in nonlinear estimation. My interest in generalized inverses developed from quadratic discrimination problems that had partial-redundancy in the multiple dimensions. I found generalized inverses useful for reducing the effective dimensionality, permitting enhanced and more stable discrimination.

"The theoretical justification for ridge regression first came to my attention when Kennard circulated a long memorandum within DuPont in mid-1963. I referred to the memorandum in my 1970 paper, because it contained virtually all of the theoretical results that eventually appeared in Hoerl and Kennard.[2,3] During the late-1960s, we produced, at DuPont, a comprehensive multiple regression package including least-squares, ridge, and generalized inverse estimators. With this package we explored and demonstrated the practical utility of the ridge and generalized inverse estimators on many ill-conditioned problems.

"Work on my 1970 paper was virtually complete about a year before it was submitted for publication. I felt, however, that the Hoerl and Kennard papers should appear first. Consequently, I employed a variety of methods of persuasion whose net effect may have been to get Hoerl and Kennard to submit their papers about a year earlier than they otherwise would have. My paper appeared shortly after theirs.

"The heavy citation of this work reflects, I believe, two groups of readers, a factor noted also by Hoerl and Kennard.[5] One group is the practitioners who found these biased estimators a useful solution to serious problems they encountered in the use of multiple regression. This paper and the Hoerl and Kennard papers provided a conceptual framework and practical guidance[6,7] for such problems. The other group is the theoreticians. The biased estimators appeared to some theoreticians as a challenge to the philosophical framework that had been built during the preceding decades. The principle of unbiasedness had become enshrined as necessary for 'goodness' in a statistical estimator. These theoreticians had difficulty accommodating mean square error as a more general criterion, and with it the demonstration that the classical unbiased least-squares estimator could be very bad in the presence of commonly encountered ill-conditioning. Gradually it is becoming broadly understood and accepted that the biased estimators have a better batting average for ill-conditioned problems, even though for each such data set there are some possible orientations of the true (but unknown) regression parameter vector for which the biased estimators do not improve on least squares. I believe that my 1970 paper, by showing the parallelism of the ridge theory, the generalized inverse theory, and the Bayesian interpretation via fictitious 'prior' data points, established a perspective from which it is now evident that all good data modeling must involve the explicit imposition of prior information, either in the experimental design (which can assure the adequacy of least-squares model estimation), or in data analysis via biased estimators,[7] or via both routes."

1. **Marquardt D W.** An algorithm for least-squares estimation of nonlinear parameters. *J. Soc. Indust. Appl. Math.* **11**:431-41, 1963.
2. **Hoerl A E & Kennard R W.** Ridge regression: biased estimation for nonorthogonal problems. *Technometrics* **12**:55-67, 1970.
3. ----------------------------------. Ridge regression: applications to nonorthogonal problems. *Technometrics* **12**:69-82, 1970.
4. **Marquardt D W.** Citation Classic. Commentary on *J. Soc. Indust. Appl. Math.* **11**:431-41, 1963. *Current Contents/Engineering, Technology & Applied Sciences* **10**(27):14, 2 July 1979.
5. **Hoerl A E & Kennard R W.** Citation Classic. Commentary on *Technometrics* **12**:55-67, 1970. *Current Contents/Engineering, Technology & Applied Sciences* **13**(35):18, 30 August 1982.
6. **Marquardt D W & Snee R D.** Ridge regression in practice. *Amer. Statist.* **29**:3-20, 1975.
7. **Marquardt D W.** You should standardize the predictor variables in your regression models. *J. Amer. Statist. Assn.* **75**:87-91, 1980.