CHAPTER 5

MINIMUM DISTANCE ESTIMATORS

5.1. INTRODUCTION

The practice of obtaining estimators of parameters by minimizing a certain distance between some functions of observations and parameters has long been present in statistics. The classical examples of this method are the Least Square and the minimum Chi Square estimators.

The minimum distance estimation (m.d.e.) method, where one obtains an estimator of a parameter by minimizing some distance between the empirical d.f. and the modeled d.f., was elevated to a general method of estimation by Wolfowitz (1953, 1954, 1957). In these papers he demonstrated that compared to the maximum likelihood estimation method, the m.d.e. method yielded consistent estimators rather cheaply in several problems of varied levels of difficulty.

This methodology saw increasing research activity from the mid–seventy's when many authors demonstrated various robustness properties of certain m.d. estimators. Beran (1977) showed that in the i.i.d. setup the minimum Hellinger distance estimators, obtained by minimizing the Hellinger distance between the modeled parametric density and an empirical density estimate, are asymptotically efficient at the true model and robust against small departures from the model, where the smallness is being measured in terms of the Hellinger metric. Beran (1978) demonstrated the powerfulness of minimum Hellinger distance estimators in the one sample location model by showing that the estimators obtained by minimizing the Hellinger distance between an estimator of the density of the residual and an estimator of the density of the negative residual are qualitatively robust and adaptive for all those symmetric error distributions that have finite Fisher information.

Parr and Schucany (1979) empirically demonstrated that in certain location models several minimum distance estimators (where several comes from the type of distances chosen) are robust. Millar (1981, 1982, 1984) proved local asymptotic minimaxity of a fairly large class of m.d. estimators, using Cramer–Von Mises type distance, in the i.i.d. setup. Donoho and Liu (1988 a, b) demonstrated certain further finite sample robustness properties of a large class of m.d. estimators and certain additional advantages of using Cramer–Von Mises and Hellinger distances. All of these authors restrict their attention to the one sample setup or to the two sample location model. See Parr (1981) for additional bibliography on m.d.e. through 1980.

Little was known till the early 1980's about how to extend the above methodology to one of the most applied models, v.i.z., the multiple linear regression model (1.1.1). Given the above optimality properties in the one- and two- sample location models, it became even more desirable to extend this methodology to this model. Only after realizing that one should use the weighted, rather than the ordinary, empiricals of the residuals to define m.d. estimators was it possible to extend this methodology satisfactorily to the model (1.1.1).