10 OCT 76

STATISTICAL DECISION THEORY - IX LINEAR MODELS AND DECISIONS

THE GENERAL LINEAR HYPOTHESIS

OBSERVE RANDOM VECTOR $\underline{Y}$ OF DIMENSION M

$E(\underline{Y})$ ASSUMED TO LIE IN SUBSPACE OF DIMENSION $N < M$ SPANNED BY COLUMNS OF KNOWN X

THAT IS, $E(\underline{Y}) = X\underline{\beta}$ FOR SOME VECTOR OF LINEAR COMBINATION COEFFICIENTS OF DIMENSION N

ESTIMATION

$E(\underline{Y}) = X\underline{\beta} \Rightarrow \underline{Y} = X\underline{\beta} + \underline{\varepsilon}$, $\underline{\varepsilon}$ A VECTOR (OF DIM M) OF DEVIATIONS FROM THE MEAN (ERROR VECTOR)

GAME $\langle \Theta, A, L \rangle = \langle \{\underline{\beta}\}, \{\text{SELECT A } \underline{\beta}\}, L \rangle$

A SELECTION (ESTIMATION) OF $\underline{\beta}$, DESIGNATED $\hat{\underline{\beta}}$, INDUCES $\hat{\underline{Y}} = X\hat{\underline{\beta}}$ AND $\hat{\underline{\varepsilon}} = \underline{Y} - \hat{\underline{Y}}$

THE STANDARD ESTIMATION CRITERION FOR SELECTING $\hat{\underline{\beta}}$ IN THE GENERAL LINEAR HYPOTHESIS IS TO SELECT $\hat{\underline{\beta}}$ SO THAT THE LENGTH OF THE ERROR VECTOR, $|\hat{\underline{\varepsilon}}|$, IS A MINIMUM (MINIMUM ERROR)

WITH THE STANDARD EUCLIDEAN INNER PRODUCT, $|\hat{\underline{\varepsilon}}| = \sqrt{\hat{\underline{\varepsilon}}'\hat{\underline{\varepsilon}}} = \sqrt{\sum_i \hat{\varepsilon}_i^2}$; THUS, TO MINIMIZE THE LENGTH OF THE ERROR VECTOR IS TO MINIMIZE THE SUM OF THE SQUARES OF THE INDIVIDUAL ERRORS.

FROM THIS DERIVES THE NAME LEAST SQUARES

THE LEAST SQUARES ESTIMATE $\hat{\underline{\beta}}$ IS LINEAR IN $\underline{Y}$, UNBIASED, AND HAS MINIMUM VARIANCE AMONG ALL OTHER LINEAR UNBIASED ESTIMATES OF $\underline{\beta}$

SIMILARLY, $\underline{P}'\hat{\underline{\beta}}$ IS AN MVU LINEAR ESTIMATE OF $\underline{P}'\underline{\beta}$ FOR ANY VECTOR $\underline{P} \ni \exists$ A LINEAR FUNCTION OF $\underline{Y}$ WITH EXPECTATION $\underline{P}'\underline{\beta}$, OR, EQUIVALENTLY, IF $\underline{P} \in \text{SPAN}(X')$ — NOTE THAT, IF X IS OF FULL RANK, THEN ANY VECTOR $\underline{P}$ OF LENGTH N (FOR $\underline{\beta}$ OF LENGTH N) IS ACCEPTABLE — ALLOWS ESTIMATION OF LINEAR COMBINATIONS OF THE β_i's

FOR THE ASSUMPTION $\underline{Y} \sim N(X\underline{\beta}, \sigma^2 I)$, THE LEAST SQUARES ESTIMATE OF $\underline{\beta}$ IS ALSO THE MAXIMUM LIKLIHOOD ESTIMATE (AND LIKEWISE FOR $\underline{P}'\hat{\underline{\beta}}$)

FOR $\underline{Y} \sim N(X\underline{\beta}, \Sigma)$, THE MAXIMUM LIKLIHOOD ESTIMATE FOR $\underline{\beta}$ WILL, IN GENERAL, BE EQUIVALENT TO A MINIMUM LENGTH ERROR VECTOR ESTIMATE FOR SOME NON-EUCLIDEAN INNER PRODUCT

HYPOTHESIS TESTING

ASSUME $\underline{Y} \sim N(X\underline{\beta}, \sigma^2 I)$ — LIKLIHOOD RATIO TESTS

TESTS OF $H_0: \beta_{K+1} = \ldots = \beta_N = 0$ AGAINST THEIR NEGATIONS (i.e., $H_0: E(\underline{Y}) \in \text{SPAN}\{\text{FIRST K VECTORS OF } X\}$, RATHER THAN THE SPAN OF ALL OF X) YIELD F STATISTICS

TESTS OF $H_0: \beta_i = A_i$ OR $H_0: \underline{P}'\underline{\beta} = c$ YIELD t TESTS (WITH APPROPRIATE VARIANCE ESTIMATORS).