M-Estimation

- The Model
  - $\min E(q(\mathbf{z},\boldsymbol{\beta}))$
  - $\sum_{i=1,\ldots,n} q(\mathbf{z}_i,\boldsymbol{\beta})/n \rightarrow^p E(q(\mathbf{z},\boldsymbol{\beta}))$
  - $\mathbf{b} = \text{argmin} \sum_{i=1,\ldots,n} q(\mathbf{z}_i,\boldsymbol{\beta})$ (or divided by n)
  - Zero-Score: $\sum_{i=1,\ldots,n} \partial q(\mathbf{z}_i,\mathbf{b})/\partial\boldsymbol{\beta}' = \mathbf{0}$
  - Positive Definite Hessian: $\sum_{i=1,\ldots,n} \partial^2 q(\mathbf{z}_i,\mathbf{b})/\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}')$

NLS, ML as M-Estimation

- Nonlinear Least Squares
  - $y = f(\mathbf{x},\boldsymbol{\beta}) + \boldsymbol{\varepsilon}$, $\mathbf{z} = [y,\mathbf{x}]$
  - $S(\boldsymbol{\beta}) = \boldsymbol{\varepsilon}'\boldsymbol{\varepsilon} = \sum_{i=1,\ldots,n} (y_i - f(\mathbf{x}_i,\boldsymbol{\beta}))^2$
  - $q(y_i,\mathbf{x}_i,\boldsymbol{\beta}) = (y_i - f(\mathbf{x}_i,\boldsymbol{\beta}))^2$
  - $\mathbf{b} = \text{argmin } S(\boldsymbol{\beta})$
- Maximum Log-Likelihood
  - $ll(\boldsymbol{\beta}) = \sum_{i=1,\ldots,n} \ln \text{pdf}(y_i,\mathbf{x}_i,\boldsymbol{\beta})$
  - $q(y_i,\mathbf{x}_i,\boldsymbol{\beta}) = \ln \text{pdf}(y_i,\mathbf{x}_i,\boldsymbol{\beta})$
  - $\mathbf{b} = \text{argmin } -ll(\boldsymbol{\beta})$

The Score Vector

- $\mathbf{s}(\mathbf{z},\boldsymbol{\beta}) = \partial q(\mathbf{z},\boldsymbol{\beta})/\partial\boldsymbol{\beta}'$
  - $E(\mathbf{s}(\mathbf{z},\boldsymbol{\beta})) = E(\partial q(\mathbf{z},\boldsymbol{\beta})/\partial\boldsymbol{\beta}') = \mathbf{0}$
    - $\sum_{i=1,\ldots,n} \mathbf{s}(\mathbf{z}_i,\boldsymbol{\beta})/n \to^p E(\mathbf{s}(\mathbf{z},\boldsymbol{\beta})) = \mathbf{0}$
  - $\boldsymbol{V} = \mathrm{Var}(\mathbf{s}(\mathbf{z},\boldsymbol{\beta})) = E(\mathbf{s}(\mathbf{z},\boldsymbol{\beta})\mathbf{s}(\mathbf{z},\boldsymbol{\beta})')$
    $= E((\partial q(\mathbf{z},\boldsymbol{\beta})/\partial\boldsymbol{\beta}')\ (\partial q(\mathbf{z},\boldsymbol{\beta})/\partial\boldsymbol{\beta})\ )$
    - $\sum_{i=1,\ldots,n} [\mathbf{s}(\mathbf{z}_i,\boldsymbol{\beta})\mathbf{s}(\mathbf{z}_i,\boldsymbol{\beta})']/n \to^P \mathrm{Var}(\mathbf{s}(\mathbf{z},\boldsymbol{\beta}))$
- $\sum_{i=1,\ldots,n} \mathbf{s}(\mathbf{z}_i,\boldsymbol{\beta})/n \to^d N(\mathbf{0},\boldsymbol{V})$

The Hessian Matrix

- $\boldsymbol{H}(\mathbf{z},\boldsymbol{\beta})$
  $= E(\partial \mathbf{s}(\mathbf{z},\boldsymbol{\beta})/\partial\boldsymbol{\beta}) = E(\partial^2 q(\mathbf{z},\boldsymbol{\beta})/\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}')$
  - $\sum_{i=1,\ldots,n} [\partial \mathbf{s}(\mathbf{z}_i,\boldsymbol{\beta})/\partial\boldsymbol{\beta}]/n \rightarrow^p \boldsymbol{H}(\mathbf{z},\boldsymbol{\beta})$
- Information Matrix Equality does not necessary hold in general: **-** $\boldsymbol{H} \neq \boldsymbol{V}$
  - More general than NLS and ML

The Asymptotic Theory

- $\sqrt{n}(\mathbf{b}-\boldsymbol{\beta}) = \{\sum_{i=1,\ldots,n}[\partial \mathbf{s}(\mathbf{z}_i,\mathbf{b})/\partial\boldsymbol{\beta}]/n\}^{-1}[\sum_{i=1,\ldots,n}\mathbf{s}(\mathbf{z}_i,\mathbf{b})/n]$
- $\sqrt{n}(\mathbf{b}-\boldsymbol{\beta}) \rightarrow^{d} N(\mathbf{0},\boldsymbol{H}^{-1}\boldsymbol{V}\boldsymbol{H}^{-1})$
    - $\boldsymbol{V} = \mathrm{E}(\mathbf{s}(\mathbf{z},\boldsymbol{\beta})\mathbf{s}(\mathbf{z},\boldsymbol{\beta})') = \mathrm{E}[(\partial q(\mathbf{z},\boldsymbol{\beta})/\partial\boldsymbol{\beta})'(\partial q(\mathbf{z},\boldsymbol{\beta})/\partial\boldsymbol{\beta})]$
    - $\boldsymbol{H} = \mathrm{E}(\partial\mathbf{s}(\mathbf{z},\boldsymbol{\beta})/\partial\boldsymbol{\beta}) = \mathrm{E}(\partial^2 q(\mathbf{z},\boldsymbol{\beta})/\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}')$
  - $\sum_{i=1,\ldots,n}[\mathbf{s}(\mathbf{z}_i,\mathbf{b})\mathbf{s}(\mathbf{z}_i,\mathbf{b})']/n \rightarrow^{p} \boldsymbol{V}$
  - $\sum_{i=1,\ldots,n}[\partial\mathbf{s}(\mathbf{z}_i,\mathbf{b})/\partial\boldsymbol{\beta}]/n \rightarrow^{p} \boldsymbol{H}$

Asymptotic Normality

- $\mathbf{b} \to^{p} \boldsymbol{\beta}$
- $\mathbf{b} \sim^{a} N(\boldsymbol{\beta}, \mathrm{Var}(\mathbf{b}))$
- $\mathrm{Var}(\mathbf{b}) = \mathbf{H}^{-1}\mathbf{V}\mathbf{H}^{-1}$
  - $\mathbf{H}$ and $\mathbf{V}$ are evaluated at $\mathbf{b}$:
    - $\mathbf{H} = \sum_{i=1,\ldots,n} [\partial^2 q\,(\mathbf{z}_i, \mathbf{b})/\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}']$
    - $\mathbf{V} = \sum_{i=1,\ldots,n} [\partial q(\mathbf{z}_i, \mathbf{b})/\partial\boldsymbol{\beta}][\partial q(\mathbf{z}_i, \mathbf{b})/\partial\boldsymbol{\beta}']$

Application

- Heteroscedastity Autocorrelation Consistent Variance-Covariance Matrix
  - Non-spherical disturbances in NLS
- Quasi Maximum Likelihood (QML)
  - Misspecified density assumption in ML
  - Information Equality may not hold

Example

- Nonlinear CES Production Function
  $ln(Q) = \beta_1 + \beta_4\, ln[\beta_2 L^{\beta 3}+(1-\beta_2)K^{\beta 3}] + \varepsilon$
  - NLS (homoscedasticy)
  - NLS (robust covariances)
  - ML (normal density)
  - QML (misspecified normality)
- Data: JUDGE.TXT

Example

- Nonlinear CES Production Function

	Parameter Estimate	NLS s.e.	ML s.e.	ML OPG s.e.	ML Robust s.e.
β_1	0.1245	0.0734	0.0737	0.0694	0.0815
β_2	0.3367	0.1020	0.1035	0.1215	0.0895
β_3	-3.0109	2.0022	2.0713	2.0491	2.1403
β_4	-0.3363	0.2343	0.2425	0.2410	0.2496
SSE	1.7611				
Log-Lik	45.942				