OLS, ML og MM i den lineære regressionsmodel

Erik Bennike

20. januar 2005

Denne note udleder OLS-estimatoren i en lineær regressionsmodel på tre forskellige måder, der alle tjener til at illustrere en anvendelse af disse metoder.

1 Den lineære regressionsmodel

Vi betragter en multipel lineær regressionsmodel

$$y_i = x_i'\beta + \epsilon_i, \tag{1}$$

hvor x_i og β er k–dimensionale vektorer. Der er således $k - 1$ forklarende variable i modellen. Tilsvarende skrives i matrix notation:

$$y = X\beta + \epsilon, \tag{2}$$

hvor X er en $n \times k$–dimensional matrix, y og ϵ er n–dimensionale vektorer og β er en k–dimensional vektor. Om der undervejs anvendes matrix-notation eller ej er et spørgsmål om bekvemmelighed. Vi vil løbende antage at de reducerede Gauss-Markov betingelser er opfyldt (opskrevet på matrixform):

$$E[\epsilon|X] = 0 \Rightarrow E[\epsilon X] = 0 \tag{3}$$

$$var[\epsilon|X] = \sigma^2 I_n \tag{4}$$

1.1 Ordinary Least Squares

Termen *OLS* stammer fra den oprindelige udledningsmetode - nemlig *Ordinary Least Squares*, hvilket indikerer at man ønsker at minimere de kvadrerede residualer. Dvs. at vi opstiller følgende minimeringsproblem:

$$\min_{\beta}(y_i - x_i'\beta)^2$$

For at minimere udtrykket udnyttes de k førsteordensbetingelser:

$$\begin{aligned}
2\sum_{i=1}^{n}(y_i - x_i'\widehat{\beta})(-x_i) = 0 \Rightarrow \\
\sum_{i=1}^{n} x_i(y_i - x_i'\widehat{\beta}) = 0 \Rightarrow \\
\sum_{i=1}^{n} x_i x_i'\widehat{\beta} = \sum_{i=1}^{n} x_i y_i \Rightarrow
\end{aligned}$$

$$\widehat{\beta}_{OLS} = \left(\sum_{i=1}^{n} x_i x_i'\right)^{-1} \sum_{i=1}^{n} x_i y_i = (X'X)^{-1}X'Y \tag{5}$$

Bemærk at vi undervejs har antaget at matricen $\sum_{i=1}^{n} x_i x_i' = X'X$ er invertibel. Dette kræver at den er regulær, hvilket igen kræver at der ikke er lineært afhænge søjer i matricen. Hvis der er lineært afhængige søjler i matricen er vi i tilfældet med perfekt multikollinearitet. OLS-estimatoren er således kun defineret under fravær af perfekt multikollinearitet.

Variansen kan estimeres ud fra det faktum, at $\sigma^2 = \mathrm{E}[u^2] \Rightarrow \widehat{\sigma}^2 = n^{-1}\sum_{i=1}^{n}\widehat{\epsilon}_i^2$. Problemet er bare, at denne estimator for variansen viser sig at være biased. Derfor anvendes i stedet estimatoren:

$$\widehat{\sigma}^2 = \frac{1}{n-k}\sum_{i=1}^{n}\widehat{\epsilon}_i^2 \tag{6}$$

1.2 Maximum Likelihood

For at kunne opstille likelihood ligningerne er vi udover antagelserne (3) og (4) nødt til også at specificere en fordeling for ϵ_i. Her vælger vi at antage at fejlleddet er normalfordelt. Dvs. en ekstra antagelse er her, at

$$\epsilon_i | x_i \sim N(0, \sigma^2) \Rightarrow y_i | x_i \sim N(x_i'\beta, \sigma^2)$$

Tæthedsfunktionen for normalfordelingen giver os, at

$$f(y_i | x_i, \beta, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left\{-\frac{1}{2}\frac{(y_i - x_i'\beta)^2}{\sigma^2}\right\},$$

som er sandsynligheden for at observere y_i. Idet observationerne er antaget at være i.i.d. kan vi opskrive sandynligheden for at observere den rent faktisk observerede følge af y_i'er, $\{y_i\}_{i=1}^n$ som

$$f(\{y_i\}_{i=1}^n|\{x_i\}_{i=1}^n, \beta, \sigma^2) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left\{-\frac{1}{2}\frac{(y_i - x_i'\beta)^2}{\sigma^2}\right\}.$$

Dermed kan vi opskrive likelihood funktionen som

$$L(\beta, \sigma^2) = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \cdot \prod_{i=1}^n \exp\left\{-\frac{1}{2}\frac{(y_i - x_i'\beta)^2}{\sigma^2}\right\}.$$

Det er betragteligt lettere at maksimere log-likelihood funktionen:

$$\log L(\beta, \sigma^2) = -\frac{n}{2} \cdot \log(2\pi\sigma^2) - \frac{1}{2} \cdot \sum_{i=1}^n \frac{(y_i - x_i'\beta)^2}{\sigma^2}.$$

Vi vælger derefter estimatorerne for β og σ^2, således at de maksimerer ovenstående udtryk. Førsteordensbetingelserne mht. β er

$$\sum_{i=1}^n \frac{x_i(y_i - x_i'\widehat{\beta})}{\sigma^2} = 0 \Rightarrow$$

$$\sum_{i=1}^n x_i(y_i - x_i'\widehat{\beta}) = \sum_{i=1}^n (x_i y_i - x_i x_i'\widehat{\beta}) = 0 \Rightarrow$$

$$\widehat{\beta}_{ML} = \left(\sum_{i=1}^n x_i x_i'\right)^{-1} \sum_{i=1}^n x_i y_i = (X'X)^{-1}X'Y = \widehat{\beta}_{OLS} \tag{7}$$

Også her må vi således antage fravær af perfekt multikollinearitet, idet matricen $X'X$ ikke må være singulær. Førsteordensbetingelserne mht. σ^2 er

$$-\frac{n}{2\sigma^2} + \frac{1}{2}\frac{(y_i - x_i'\widehat{\beta})^2}{\sigma^4} = 0 \Rightarrow$$

$$\frac{1}{2}\frac{\widehat{\epsilon}_i^2}{\sigma^4} = \frac{n}{2\sigma^2} \Rightarrow$$

$$\widehat{\sigma}^2_{ML} = \frac{1}{n}\sum_{i=1}^n \widehat{\epsilon}_i^2 \tag{8}$$

Denne estimator for variansen er biased, men konsistent, jf. diskussionen ovenfor.

1.3 Method of Moments

I MM metoden tages udgangspunkt i de såkaldte momentbetingelser. Vi benytter antagelsen (3), og opstiller momentbetingelsen for hele populationen på matrixform her:

$$E[X\epsilon] = E[X(y - X\beta)] = 0$$

Den tilsvarende sample momentbetingelse er:

$$g_n(\beta) = \frac{1}{n}X'(y - X\beta) = \frac{1}{n}(X'y - X'X\beta) = 0 \Rightarrow$$

$$X'y = X'X\beta \Rightarrow$$

$$\widehat{\beta}_{MM} = (X'X)^{-1}X'y \tag{9}$$

Som vi såvel OLS som ML må vi således også for MM estimation antage fravær af perfekt multikollinearitet, endnu engang fordi vi må kræve at matricen $X'X$ er regulær.