Nonlinear Least Squares (NLS)

Hüseyin Taştan[1]

[1]Department of Economics
Ph.D. Program - Advanced Econometrics
Yildiz Technical University

November 11, 2013

Nonlinear Least Squares

- So far we considered regression models that are linear in parameters.
- We are able to handle certain types of nonlinearity by transforming regressors and dependent variable in the linear regression model framework.
- However, there many other types of nonlinearity that cannot be handles within this framework. In this case we need to estimate nonlinear regression models.
- In nonlinear regression models conditional expectation of y_i with respect to the information set Ω_i is a nonlinear function of the parameters.
- We assume that no transformation exists to make the model linear in parameters (e.g., natural log transformation).

Nonlinear Least Squares

- A nonlinear regression can be written as
$$y_i = f(\boldsymbol{x}_i, \boldsymbol{\beta}) + u_i, \quad E(u_i) = 0, \quad E(u_i^2) = \sigma^2, \ \forall i. \qquad (1)$$
where
$$\boldsymbol{x}_i = [x_{1i}, x_{2i}, \dots, x_{ki}]^\top, \quad i = 1, \dots, n$$
- Under the assumption that the functional form is correctly specified
$$\mathsf{E}[y_i|\boldsymbol{x}_i] = f(\boldsymbol{x}_i, \boldsymbol{\beta}), \quad i = 1, 2, \dots, n.$$
where $f(\cdot)$ is a twice differentiable function.
- The nonlinear function $f(\boldsymbol{x}_i, \boldsymbol{\beta})$ determines the value of y_i conditional on the information set Ω_i which usually consists of some set of explanatory variables.
- These explanatory variables may include lagged values of the dependent variable (if we have a time series or panel data) and exogenous variables.

Nonlinear Least Squares

- The NLS objective function is
$$Q(\boldsymbol{\beta}) = \frac{1}{2} \sum_{i=1}^{n} (y_i - f(\boldsymbol{x}_i, \boldsymbol{\beta}))^2 \qquad (2)$$
- and the NLS estimator is
$$\hat{\beta} = \arg\min_{\boldsymbol{\beta}} Q(\boldsymbol{\beta}) \qquad (3)$$
- FOC:
$$\frac{\partial Q(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} = -\sum_{i=1}^{n} (y_i - f(\boldsymbol{x}_i, \boldsymbol{\beta})) \frac{\partial f(\boldsymbol{x}_i, \boldsymbol{\beta})}{\partial \boldsymbol{\beta}} = \mathbf{0} \qquad (4)$$

Nonlinear Least Squares

Let the $n \times k$ matrix $\boldsymbol{G}$ be the first derivatives of the nonlinear function whose typical element is

$$g_{ij} = \frac{\partial f(\boldsymbol{x}_i, \boldsymbol{\beta})}{\partial \beta_j}, \quad i = 1, \ldots, n, \quad j = 1, \ldots, k$$

Using this we can write:

$$-\boldsymbol{G}^\top(\boldsymbol{y} - f(\boldsymbol{X}, \boldsymbol{\beta})) = \boldsymbol{0}. \tag{5}$$

Under certain assumptions the NLS estimator has an asymptotic normal distribution:

$$\sqrt{n}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \to N\left(0, \sigma^2 \operatorname*{plim}_{n\to\infty} \left(\frac{1}{n}\boldsymbol{G}^\top \boldsymbol{G}\right)^{-1}\right) \tag{6}$$

where β is the unknown true parameter vector, σ^2 is the unknown true variance of the error term and $\boldsymbol{G}$ is the first derivative matrix evaluated at the true parameter vector.

Nonlinear Least Squares

In practice the covariance matrix can be estimated using

$$\mathsf{V}(\hat{\boldsymbol{\beta}}) = s^2(\hat{\boldsymbol{G}}^\top \hat{\boldsymbol{G}})^{-1} \tag{7}$$

where s^2 is an unbiased estimator of the error variance:

$$s^2 = \frac{1}{n-k} \sum_{t=1}^{n} \left(y_i - f(\boldsymbol{x}_i, \hat{\boldsymbol{\beta}})\right)^2.$$

$\hat{\boldsymbol{G}}$ is the first derivative matrix evaluated at the NLS solution $\hat{\beta}$.

Nonlinear Least Squares

Example (DM,p.212)

Consider the following model:

$$y_i = \beta_1 + \beta_2 x_{i1} + \frac{1}{\beta_2} x_{i2} + u_i$$

Here the nonlinear part is

$$f(\boldsymbol{x}_i, \boldsymbol{\beta}) = \beta_1 + \beta_2 x_{i1} + \frac{1}{\beta_2} x_{i2}$$

Example (The Linear Regression Model with AR(1) Errors (Davidson and MacKinnon, p.212))

Consider the following linear regression model for time series data:

$$y_t = \boldsymbol{X}_t \boldsymbol{\beta} + u_t$$

where u_t follows a first order autoregressive process:

$$u_t = \rho u_{t-1} + \epsilon_t, \quad \epsilon_t \sim \; iid(0, \sigma_\epsilon^2)$$

Substituting $\rho u_{t-1} + \epsilon_t$ for u_t and then replacing u_{t-1} by $y_{t-1} - \boldsymbol{X}_{t-1}\boldsymbol{\beta}$ we obtain

$$y_t = \rho y_{t-1} + \boldsymbol{X}_t \boldsymbol{\beta} - \rho \boldsymbol{X}_{t-1} \boldsymbol{\beta} + \epsilon_t, \quad \epsilon_t \sim \; iid(0, \sigma_\epsilon^2)$$

which is nonlinear in parameters. The model cannot be estimated by OLS. Instead we need to use NLS.

Example (CES Production Function)

A flexible form for the production function is the constant elasticity of substitution (CES):

$$y = \alpha \left(\delta L^{-\gamma} + (1-\delta)K^{-\gamma}\right)^{-\lambda/\gamma}$$

where y is output, K is capital and L is labor, and $\alpha > 0$, $0 < \delta < 1$, $\gamma \geq -1$. Elasticity of substitution is given by

$$s = \frac{d\log(K/L)}{d\log(MP_L/MP_K)} = \frac{1}{1+\gamma} \geq 0$$

where MP_j is the marginal product of inputs, $j = K, L$. Taking natural logs and adding an error term we obtain a nonlinear model:

$$\log(y) = \log(\alpha) - \frac{\lambda}{\gamma}\log\left(\delta L^{-\gamma} + (1-\delta)K^{-\gamma}\right) + u$$

An alternative to this model is translog production function which can be estimated by OLS.

NLS Application 1: Consumption Function

- This example is from Greene (p.191)
- Consider the following nonlinear consumption function:

$$C = \alpha + \beta Y^{\gamma} + u$$

- If we assume that γ the model becomes linear in parameters.
- We will use income and consumption data for the US economy 1950Q1-2000Q4
- STATA: TableF5-2.dta

NLS Application 1: Consumption Function

```
. nl (realcons = {alpha=400} + {beta=1}*realdpi^{gamma=1} )
(obs = 204)

      Source |       SS       df       MS
-------------+------------------------------           Number of obs =       204
       Model |  432036555      2   216018277           R-squared     =    0.9988
    Residual | 504403.222    201  2509.46877           Adj R-squared =    0.9988
-------------+------------------------------           Root MSE      =   50.0946
       Total |  432540958    203  2130743.64           Res. dev.     =  2172.781

------------------------------------------------------------------------------
    realcons |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      /alpha |   458.7991   22.50139    20.39   0.000     414.4301    503.1682
       /beta |    .100852   .0109104     9.24   0.000     .0793386    .1223655
      /gamma |   1.244828   .0120548   103.26   0.000     1.221057    1.268598
------------------------------------------------------------------------------
  Parameter alpha taken as constant term in model & ANOVA table

. matrix list e(V)

symmetric e(V)[3,3]
                 alpha:       beta:      gamma:
                 _cons       _cons       _cons
alpha:_cons   506.31259
 beta:_cons   -.2345446   .00011904
gamma:_cons   .25756744  -.00013149   .00014532
```


NLS Application 2: CES Production Function

- CES production function

$$\log(y) = \log(\alpha) - \frac{\lambda}{\gamma}\log\left(\delta L^{-\gamma} + (1-\delta)K^{-\gamma}\right) + u$$

- Elasticity of substitution

$$s = \frac{1}{1+\gamma} \geq 0$$

- STATA: production.dta

Example: CES Production Function

```
nl (logQ = log({alpha=1}) - ({lambda=1}/{gamma=1})*
         log( {delta=0.5}*L^(-1*{gamma}) + (1-{delta})*K^(-1*{gamma}) ) )

      Source |       SS       df       MS
-------------+------------------------------         Number of obs =        100
       Model |  91.1542478     3  30.3847493         R-squared     =     0.7564
    Residual |  29.3565583    96  .305797482         Adj R-squared =     0.7488
-------------+------------------------------         Root MSE      =   .5529896
       Total |  120.510806    99  1.21728087         Res. dev.     =   161.2223

------------------------------------------------------------------------------
        logQ |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      /alpha |   44.32413    4.51106     9.83   0.000     35.36975    53.27851
     /lambda |   .9888227   .0632858    15.62   0.000     .8632014    1.114444
      /gamma |   1.432666   .5570501     2.57   0.012     .3269304    2.538402
      /delta |   .5193183   .0552948     9.39   0.000     .4095589    .6290776
------------------------------------------------------------------------------
  Parameter alpha taken as constant term in model & ANOVA table
```


Example: CES Production Function

```
. // Elasticity of substitution
. nlcom (1 / (1+_b[/gamma]))

       _nl_1:  1 / (1+_b[/gamma])

------------------------------------------------------------------------------
        logQ |      Coef.   Std. Err.      t    P>|t|     [95% Conf.
-------------+----------------------------------------------------------------
       _nl_1 |   .4110716   .0941303     4.37   0.000     .2242246
------------------------------------------------------------------------------
```