Econ 512
Fall 2018

Homework 3 – Nonlinear Optimization
Due 10/10/2018

In 1969, the popular magazine *Psychology Today* published a 101-question survey on affairs. Professor Ray Fair (1978) extracted a sample of 601 observations on men and women who are currently married for the first time and analyzed their responses to a question about extramarital affairs. He used the tobit model as his estimation framework for this study. The dependent variable is a count of the number of affairs which suggests that a standard Poisson model may be a better choice. Download the data set hw3.mat, and estimate the parameters by the methods of nonlinear least squares and maximum likelihood using different algorithms.

Data description:

y - count data: number of affairs in the past year

$\mathbf{x}$ - constant term=1, age, number of years married, religiousness (scale $1-5$), occupation (scale $1-7$), self-rating of marriage (scale $1-5$)

The data generating assumptions for the Poisson model, where j= number of affairs, are:

$$\begin{aligned}
\Pr\left[y_i = j\right] &= \frac{e^{-\lambda_i}\lambda_i^j}{j!} \\
\log \lambda_i &= \mathbf{x}_i'\beta \\
E\left(y_i \left| x_i\right.\right) &= e^{\mathbf{x}_i{}'\beta}
\end{aligned}$$

for some $\beta = \left(\beta_0, \beta_1, \beta_2, \beta_3, \beta_4, \beta_5\right)'$.

The log-likelihood function is:

$$\begin{aligned}
\ln L &= \sum_{i=1}^{n} \ln f\left(y_i \left| x_i\right., \beta\right) \\
&= \sum_{i=1}^{n} \ln \frac{e^{-\lambda_i}\lambda_i^j}{j!} \\
&= \sum_{i=1}^{n} \left[-\lambda_i + y_i \ln \lambda_i - \ln j!\right] \\
&= \sum_{i=1}^{n} \left[-e^{\mathbf{x}_i{}'\beta} + y_i x_i'\beta - \ln y_i!\right]
\end{aligned}$$

The residual sum of squares is:

$$S\left(\beta\right) = \sum_{i=1}^{n}\left(y_i - e^{\beta' x_i}\right)^2$$

1. Estimate the parameter vector β using the maximum likelihood estimator computed via the Nelder-Mead simplex method.

Please check my matlab codes for the algorith. The MLE estimator is [2.5924, -0.0333, 0.1162, -0.3572, 0.0783, -0.4131] when starting value[1,1,1,1,1,1] and the Nelder-Mead simplex method are used.

2. Estimate the parameter vector β using the maximum likelihood estimator computed via a quasi-Newton optimization method, report which method you choose.

I use Broyden's method to solve for MLE. First, I find the FOC:

$$\sum_{i=1}^{n}\{-x_i e^{x_i'\beta} + y_i x_i\} = 0 \tag{1}$$

The Jacobian of equation (1) is:

$$\sum_{i=1}^{n}\{-x_i x_i' e^{x_i'\beta}\}$$

The estimator generated using the estimator in the above question as initial value, is [2.5339, -0.0323, 0.1157, -0.3540, 0.0798, -0.4094].

3. Estimate the parameter vector β using nonlinear least squares estimator computed using the command `lsqnonlin`. What computation method are you using?

I use the trust-region-reflective option in solver `lsqnonlin` and the inital value [2.5, -0.0, 0.1, -0.4, 0.1, -0.5]. The NLS estimator got is [2.5122, -0.0384, 0.1141, -0.2799, 0.0677, -0.3697].

4. Estimate the parameter vector β using the nonlinear least squares estimator computed using the Nelder-Mead simplex method.

The estimator is [2.5339 -0.0323 0.1157 -0.3540 0.0798 -0.4094] in this case.

5. Test all four approaches with regard to the choice of initial values. Roughly rank them in order of robustness and time to convergence. Submit a short writeup summarizing your results.

I select 100 starting points and test the four approaches.

For the first method, each time the maximum number of function evaluations is exceeded. So it is not robust. The average time of reaching the maximum number of function evaluations is 0.2788 second.
For the second method, number of convergence is 4, when we select the Broyden iteration number to be 100. And the average time to convergence is 0.024 second.
For the third method, number of convergence is 91. And the average time to convergence is 0.0699 second.
The fourth approach is the same as the first approach when specifying the starting value.
Therefore, the most robust method is Nonlinear least sqaure when using the lsqnonlin solver. And conditional on convergence, the Broyden's method would be faster.