PROBLEM SET 3

General Instructions

Complete the following assignments in a sweave file that shows your code, output, and discussion. Hand in the compiled (pdf) version. You can use this file to get you started. You can work with others, but please hand in your own version. Please report any glitches as soon as you discover them - thanks!

For the theory questions, you may want to work out your answer on paper first, then type up the key steps in LaTeX. You can insert your LaTeXcode directly in in your sweave file, or - if you prefer, in a separate file or files in TeXnicCenter. For full credit, typed answers are required.

1. MLE Bernoulli

Consider the Bernoulli model (also called "Binary response model") where a given random draw $y_i = 1$ with probability π, and $y_i = 0$ with probability $1 - \pi$. Thus, the density for this model can be compactly written as $f(y_i) = \pi^{y_i}(1-\pi)^{1-y_i}$. The moments for this density are given as $E(y_i) = \pi, V(y_i) = \pi(1-\pi)$. Assume you draw a sample of n observations from this distribution.

(1) Derive $lnL(\pi), g(\pi), H(\pi), I(\pi)$.
(2) Derive the ML estimator (call it p), and its asymptotic variance.
(3) Given a sample of four "1"'s and one "0" (so n=5), compute $lnL(\pi)$ in terms of π and numerically.
(4) Now suppose that in the sample of 5 observations, all are "1's". Derive the numerical solution for the ML estimator for this case.

2. MLE Parameterized Exponential

Consider the Exponential density with parameterized mean, i.e

$$f(y_i|\mathbf{x}_i) = \frac{exp\left(-y_i/\mathbf{x}_i'\boldsymbol{\beta}\right)}{\mathbf{x}_i'\boldsymbol{\beta}} \quad \text{with}$$
$$E(y_i|\mathbf{x}_i) = \mathbf{x}_i'\boldsymbol{\beta}, \quad V(y_i|\mathbf{x}_i) = \left(\mathbf{x}_i'\boldsymbol{\beta}\right)^2$$

Assume the sample size is n and $\mathbf{x}_i$ is $kx1$.

(1) Derive $lnL(\boldsymbol{\beta}), g(\boldsymbol{\beta}), H(\boldsymbol{\beta})$ and $I(\boldsymbol{\beta})$
(2) Using the gradient for the entire sample, show that the score identity holds.
(3) Using the gradient for the entire sample, show that the information matrix identity holds.

3. Hypothesis Testing in MLE

Consider the hedonic property value data set from script `mod3s2c` and the log-linear regression version of this model. Estimate this model via MLE, using analytical gradient and Hessian (as in `mod2s1b`). Use the inverted negative Hessian to derive standard errors for all estimates. Capture your MLE results in a nice table.

Choose the following tuner settings for the optimization algorithm:

```
R> cri<-10         #initial setting for convergence criterion
R> cri1<-0.0001     #convergence criterion
R> #(here for the sum of the absolute values of the elements in the gradient)
R> maxiter<-2000   #max. number of allowed iterations
R> stsz<-0.1       #step size, here it seems necessary to keep it on the small side
R> #... try 0.5 and see what happens....
```

(1) Which coefficients are significant at the 1% or 5% level?
(2) Interpret the marginal effects of "lnacres" , "gradeab", and "disthaz" on the dependent variable.
(3) Derive an estimate for the *variance* of the error term, along with a 95% confidence interval.
(4) Perform Wald tests for the following hypotheses. For each hypothesis obtain the test statistic and the corresponding p-value, and state your decision (use script `mod3s2b` for guidance on Wald tests in an MLE context):

   (a) The marginal effects of "age" (β_4), "pkadeq" (β_6), "vacant" (β_7) and "popden" (β_9) are jointly zero.
   (b) The effect of an additional mile away from the airport on the value of a property is LESS OR EQUAL TO 3 times that of an additional mile from the hazardous waste site.
   (*Hint: There are multiple ways to set up the null hypothesis. Choose the one that has a zero on the right hand side of the equation.*)
   (c) The ratio of (additional mile from airport / additional mile from haz. waste site) is NO SMALLER THAN the squared effect of "log of square footage".
   (*Hint: There are multiple ways to set up the null hypothesis. Choose the one that has a zero on the right hand side of the equation.*)