Econometrics HW4

Michael B. Nattinger[*]

April 12, 2021

1 Question 22.1

1.1 Part A

Given the PDF of $e, \exists F(e)$, the corresponding CDF. Note now that $P(y_i \leq y|x_i = x) = P(x_i'\theta + e_i \leq y|x_i = x) = P(e_i \leq y - x_i'\theta|x_i = x)$. From the independence of e, x we have: $P(e_i \leq y - x_i'\theta|x_i = x) = P(e_i \leq y - x_i'\theta) = F(y - x_i\theta)$. By differentiation we find $f(y|x) = f(y - x'\theta)$.

1.2 Part B

We have some functional form for the density of e so we can write our objective function as (negative) log likelihood (negative since we are minimizing). Then, $\rho_i(\theta) = -log(f(y_i - x_i\theta))$, $\psi_i(\theta) = \frac{\partial}{\partial\theta}\rho_i(\theta) = \frac{f'(y_i - x_i'\theta)}{f(y_i - x_i'\theta)}x_i$.

1.3 Part C

$\Omega = E[\psi_i\psi_i'] = E[\left(\frac{f'(y_i - x_i'\theta)}{f(y_i - x_i'\theta)}\right)^2 x_ix_i'] = E[\left(\frac{f'(e_i)}{f(e_i)}\right)^2 x_ix_i'] = E[\left(\frac{f'(e_i)}{f(e_i)}\right)^2]E[x_ix_i']$. As stated in the textbook, for correctly specified MLE $Q = \Omega, V = \Omega^{-1} = E[\left(\frac{f'(e_i)}{f(e_i)}\right)^2]^{-1}E[x_ix_i']^{-1}$

2 Question 23.1

2.1 Part A

Due to the nonlinearity of $exp(\cdot)$, the conditional mean is nonlinear in θ. This would be a nonlinear regression model.

2.2 Part B

We can define $\gamma = exp(\theta)$ and write $Y = \gamma + e$. Then we can estimate $\hat{\gamma} = \bar{Y}_n$, and obtain our estimate $\hat{\theta} = log(\hat{\gamma})$.

2.3 Part C

M-estimators are invariant under reparameterization so this is the same.

[*]I worked on this assignment with my study group: Alex von Hafften, Andrew Smith, and Ryan Mather. I have also discussed problem(s) with Emily Case, Sarah Bass, Katherine Kwok, and Danny Edgel.

3 Question 23.2

No. For even integer λ, $Y^{(\lambda)}$ is not invertible. Therefore one could not (uniquely) express Y as a function of X given parameters.

4 Question 23.7

We can calculate the SE as follows: $s = \sqrt{\frac{\partial m(x,\hat{\theta})}{\partial \theta}\hat{V}\frac{\partial m(x,\hat{\theta})}{\partial \theta'}}$. Then,
$CI = [m(x,\hat{\theta}) - 1.96s, m(x,\hat{\theta}) + 1.96s]$ is an asymptotic 95% confidence interval for $m(x)$.

5 Question 23.8

Source	SS	df	MS			
				Number of obs	=	**338**
Model	**768.35943**	**3**	**256.119809**	R-squared	=	**0.9837**
Residual	**12.713675**	**334**	**.038064896**	Adj R-squared	=	**0.9836**
				Root MSE	=	**.1951023**
Total	**781.0731**	**337**	**2.31772434**	Res. dev.	=	**-149.5618**

lny	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
/beta	**-12.58216**	**.1845737**	**-68.17**	**0.000**	**-12.94523**	**-12.21909**
/nu	**1.042579**	**.007826**	**133.22**	**0.000**	**1.027185**	**1.057974**
/rho	**.4114824**	**.0584956**	**7.03**	**0.000**	**.2964163**	**.5265486**
/alpha	**.3194286**	**.0115813**	**27.58**	**0.000**	**.2966471**	**.3422102**

Parameter beta taken as constant term in model & ANOVA table

$\hat{\sigma} = \frac{1}{1-\hat{\rho}} = 1.70$. The estimated coefficients are almost all really close, but β is not.

```
use "C:\Users\micha\OneDrive\Documents\HOMEWORK\Y1S1\Metrics\Q4\PS4\PSS2017.dta", clear
gen lny = log(EG_total)
rename EC_c_alt x1
rename EC_d_alt x2
drop if missing(lny, x1, x2)
nl (lny = {beta} + ({nu}/{rho})*log({alpha}*x1^{rho} + (1-{alpha})*x2^{rho})), initial(alpha
0.5 beta 1.5 nu 1 rho 0.5)
```

6 Question 24.3

Fix τ and define θ as described in the problem. Note the following:

$$\begin{aligned}
E[\psi(Y-\theta)] &= E[\tau - 1\{Y-\theta<0\}] \\
&= \tau - E[1\{Y<\theta\}] \\
&= \tau - P(Y<\theta) \\
&= 0 \\
\Rightarrow P(Y<\theta) &= \tau,
\end{aligned}$$

so θ is the τ quantile of Y.

7 Question 24.4

7.1 Part A

$E[Y|X] = E[X'\beta + e|X] = X'\beta + E[e|X] = X'\beta$ by the symmetry of the conditional distribution of e. Also by the conditional symmetry, $med[Y|X] = E[Y|X] = X'\beta$.

7.2 Part B

Asymptotically they are the same but under a finite sample they are different in general. This is because they minimize different criterion, and in general although the minimizers of squared error and absolute deviation will converge eventually there is no reason why they should be the same under a finite sample.

7.3 Part C

You would prefer LAD over OLS if you were worried about outliers in your data which may have too much of an influence in finite samples. Otherwise you would use OLS as it is conditionally unbiased and efficient in this context.

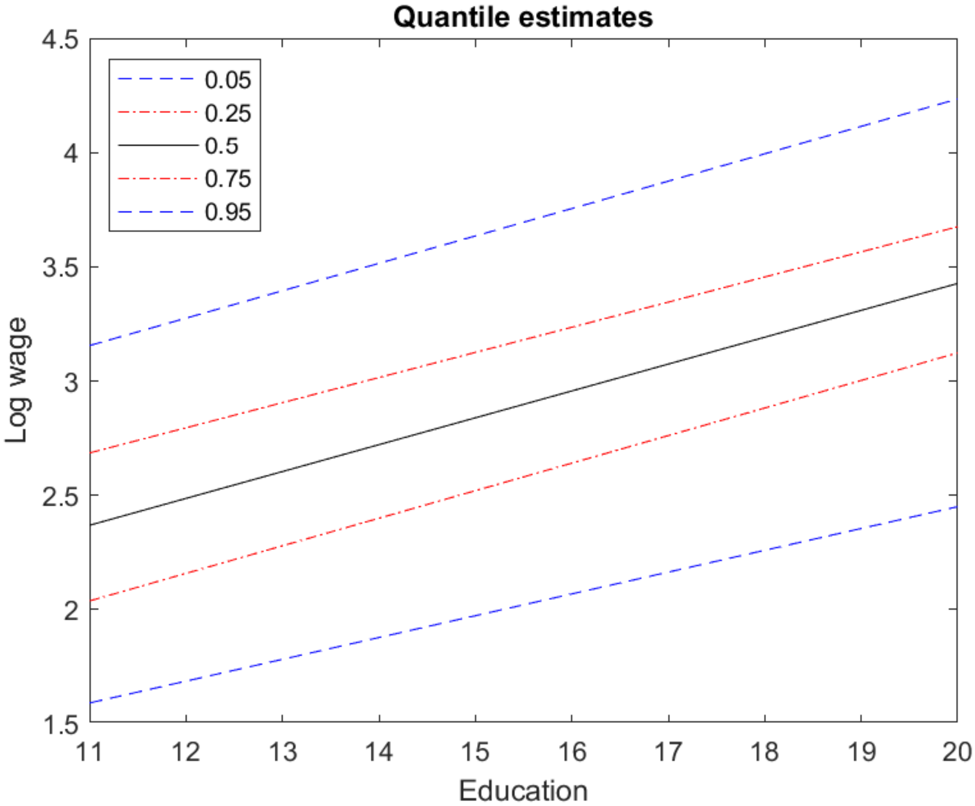
8 Question 24.5

Our colleague is being a bit ridiculous. The OLS estimate $\hat{\beta}_{ols}$ is guaranteed to have a higher R^2 than the LAD estimate $\hat{\beta}_{lad}$ by construction:

$$R^2_{ols} = 1 - \frac{(y - \hat{\beta}_{ols}x)'(y - \hat{\beta}_{ols}x)}{(y - \bar{y})'(y - \bar{y})} \geq 1 - \frac{(y - \hat{\beta}_{lad}x)'(y - \hat{\beta}_{lad}x)}{(y - \bar{y})'(y - \bar{y})} = R^2_{lad}.$$

Comparing R^2 is an unfair test for determining which estimate to trust.

9 Question 24.14



The above figure shows the estimated quantiles of the conditional distributions of log wages, conditioning on education. The results indicate that the conditional distribution is approximately homoskedastic.

```
clear; close all; clc
read = 0;
if read
[x,xt] = xlsread('cps09mar.xlsx','Sheet1');
save 'data'
else
load 'data'
end

female = x(:,2);
hisp = x(:,3);
edu = x(:,4);
ie = edu>=11;
I = logical(female.*hisp.*ie);
earnings = x(:,5);
hours = x(:,6);
week = x(:,7);
lw = log(earnings./(hours.*week));
Y = lw(I,:);
X = edu(I,:);
X = [ones(length(Y),1) X];
q = [0.05 0.25 0.5 0.75 0.95];
b=qr_nattinger(Y,X,q);
grid = (min(X(:,2)):0.1:max(X(:,2)))';
xg = [ones(size(grid)) grid];
Yh = xg*b;
colore = {'b--' 'r-.' 'k-' 'r-.' 'b--'};
figure
for i=1:length(q)
plot(grid,Yh(:,i),colore{i}); hold on
end
hold off
set(gcf,'Color',[1 1 1])
title('Quantile estimates')
xlabel('Education')
ylabel('Log wage')
legend('0.05', '0.25', '0.5', '0.75', '0.95','Location','NorthWest')
cd('pings')
saveas(gcf,'qrfig.png')
cd('..')

function brq = qr_nattinger(y,x,q)
% estimates quantile regression using fminunc
% y: n by 1 vector
% x: n by k matrix
% q: nq by 1 (or 1 by nq) vector
% Michael Nattinger, 2020
nq = length(q);
k = size(x,2);
brq = zeros(k,nq);
ops = optimoptions(@fminunc,'Display','none');
for i=1:nq
obj = @(b) ssew(b,y,x,q(i));
brq(:,i) = fminunc(obj,zeros(k,1),ops);
end
end

function sse = ssew(b,y,x,iq)
sse = iq*sum(abs(y(y>x*b,:) - x(y>x*b,:)*b)) + (1-iq)*sum(abs(y(y<x*b,:) - x(y<x*b,:)*b));
end
```