Alternative to Least Squares

In the file `dataL1L2.Rdata` are three objects named `Y, x1` and `x2`. Use the `load` command to get these three objects in your workspace. The variable Y is the dependent variable, and x_1 and x_2 are regressors. We wish to fit the following linear model:

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + U.$$

(a) Fit the model, writing a script using the command

```
fit=lm(Y~x1+x2)
summary(fit)
```

Think of some ways to check whether our model fits the data well. Also check the residuals; these you can find in `fit$res`. Try making a histogram (command `hist`), and a qq-plot (`qqnorm`).

We wish to compare the least squares estimator to an alternative which is less sensitive to large outliers. Define the function

$$L_1(\beta) = \sum_{i=1}^{n} |Y_i - \beta_0 - \beta_1 x_{1,i} - \beta_2 x_{2,i}|.$$

We define the L_1-estimator $\hat{\beta}^{(1)}$ as the minimizer of $L_1(\beta)$. To calculate this estimator in `R`, use the following script:

```
L1 = function(b,x1,x2,Y){
sum(abs(Y-b[1]-b[2]*x1-b[3]*x2))
}
par0=fit$coef
par=optim(par0,L1,gr=NULL,x1,x2,Y)
```

Try and understand this script completely (use the `help` command, if necessary).

(b) Calculate the L_1-estimator of β and compare the residuals of the two estimators (least squares and L_1).

(c) We wish to see which estimator performs better. To do this, we will simulate data from a given model, determine the two estimators, and compare them to what they should have been (since the model is given, we know the "true" parameters). Understand and use the following code:

```
X=cbind(1,x1,x2)
M=1000
b1sim=matrix(0,M,3)
b2sim=matrix(0,M,3)
S=sqrt(sum(fit$res^2)/47)
for (i in 1:M){
Usim=rnorm(50)*S
```

```
Ysim= X%*%fit$coef + Usim
fitsim=lm(Ysim~x1+x2)
b2sim[i,]=fitsim$coef
parsim=optim(fitsim$coef,L1,gr=NULL,x1,x2,Ysim)
b1sim[i,]=parsim$par
}
```

Now think of a way to determine which estimator performed the best.

(d) Mike complains that in the above comparison, the least squares estimator had a distinct advantage, since we assumed that the errors U were normally distributed. Prove that the least squares estimator is indeed the maximum likelihood estimator in this case.

(e) We will now assume that U has a $t(3)$ distribution. This distribution has heavier tails than the normal distribution, and we know that the L_1-estimator is less sensitive to large outliers. Repeat the experiment in **(c)**, but give `Usim` a $t(3)$ distribution. Hint: `rt(50,3)` generates 50 draws from the student distribution with 3 degrees of freedom.

(f) Using the model in **(e)**, with the student errors, and fixing the parameters at $\beta = (2, -1, 1)$, determine by simulation the probability of a type I error of the standard null hypothesis testing for the hypothesis $H_0 : \beta_1 = -1$ at a significance of 5%. In the same script, also determine the power of the test for the null hypothesis $H_0 : \beta_1 = -0.7$. Compare these two probabilities with the probabilities we would get using normal errors. Give your conclusions.