

Re: multiple regression and strange results

Source: <http://sci.tech-archive.net/Archive/sci.stat.math/2005-12/msg00351.html>

- *From:* Brett Magill <magillb@xxxxxxxxxxxxxx>
 - *Date:* Fri, 30 Dec 2005 21:44:15 GMT
-

On Fri, 30 Dec 2005 04:22:33 +0000, Brett Magill wrote:

> The discussions were not without contention, however you
> will certainly understand why your "expected signs" are not what you
> expected and why your logic is incorrect.

Just a follow up, and a small contribution to the "expected sign" discussion. If you cannot convince yourself of the error of expecting a particular sign of a regression coefficient in the presence of colinearity (and without justification), this exercise might help. It produces the multiple regression coefficients through a series of simple regressions.

1. Generate some random, correlated data, see attached along with R code.
2. Regress Y on X2 and store the residuals, that is the errors.
3. Regress X1 on X2 and store the residuals.
4. Regress the residuals of step 2 (R1) on the residuals of step 3 (R2).

The coefficient for step 4 for R2 will be the coefficient of X1 for the multiple regression model with X1 and X2. You could do the same for X2.

This is simply said, but seeing it through this exercise is sometimes helpful. The regression coefficients in a multiple regression model are the relationship between the part of the X variable that is not related to all of the other Xs in the model and the part of Y that is not related to all of the other Xs in the model. When there is correlation among the Xs, signs depend on the nature of those relationships.

```
x1 x2 y
[1,] -0.7235478291 0.2944136 -0.23160110
[2,] -0.5627023571 0.4717420 -0.20923889
[3,] 0.0487566530 -0.9819983 -1.01580693
[4,] 0.8784860774 0.6338299 1.59866869
[5,] -0.1853402850 0.2715666 0.19277344
[6,] 0.1930501280 1.6336652 1.35606902
[7,] 0.8164298111 -0.9904565 0.16714954
[8,] -0.0002910502 -1.5243615 -1.84938147
[9,] 0.8595245899 -0.5314447 0.12899101
```

Re: multiple regression and strange results

```
[10,] 0.0309133124 -0.2401435 -0.99943456  
[11,] 1.1208637894 -1.3999358 -0.41084306  
[12,] -0.4334258587 0.7347420 0.09201947  
[13,] 2.0349103930 -0.1455662 2.08735744  
[14,] 0.9497634311 -0.6791138 0.57593595  
[15,] -0.7805548920 1.1496734 0.74200664  
[16,] -0.1257650576 -0.8675856 -1.47340705  
[17,] -0.0526390037 0.1049072 0.09814604  
[18,] 1.0802460320 2.4474141 3.57500599  
[19,] 0.3138742502 0.1648729 1.09174828  
[20,] 0.9024158070 -0.8997844 0.05129468
```

```
> lm(y~x1+x2)
```

Call:

```
lm(formula = y ~ x1 + x2)
```

Coefficients:

(Intercept) x1 x2

```
-0.04982 1.08927 1.04477
```

Residual standard error: 0.3448 on 17 degrees of freedom
Multiple R-Squared: 0.9327, Adjusted R-squared: 0.9248
F-statistic: 117.9 on 2 and 17 DF, p-value: 1.085e-10

```
r1<-residuals(lm(y ~ x2))
```

```
r2<-residuals(lm(x1 ~ x2))
```

```
> lm (r1 ~ r2)
```

Call:

```
lm(formula = r1 ~ r2)
```

Coefficients:

(Intercept) r2

```
1.104e-17 1.089e+00 <--- Equal to 1.08927 for x1 in lm(y~x1+x2).
```

• **References:**

◆ **multiple regression and strange results**

◇ From: joseph Frank

◆ **Re: multiple regression and strange results**

◇ From: Brett Magill

Re: multiple regression and strange results

- Prev by Date: [*How to methodically choose the distribution for a regression equation*](#)
- Next by Date: [*Re: I have a solution manual for casella berger in pdf format!*](#)
- Previous by thread: [*Re: multiple regression and strange results*](#)
- Next by thread: [*Is probability the right tool for everything?*](#)
- Index(es):
 - ◆ [*Date*](#)
 - ◆ [*Thread*](#)