Re: R^2 and beta coefficients in multiple regression

Source: http://sci.tech−archive.net/Archive/sci.stat.math/2008−09/msg00305.html

---

* *From*: hrundle@xxxxxxxxxx
* *Date*: Wed, 24 Sep 2008 07:48:32 −0700 (PDT)

On Sep 24, 3:31 am, Ray Koopman <koop...@xxxxxx> wrote:

> On Sep 23, 9:40 pm, hrun...@xxxxxxxxxx wrote:
>
>> On Sep 23, 11:26 pm, Ray Koopman <koop...@xxxxxx> wrote:
>>
>>> On Sep 23, 1:29 pm, hrun...@xxxxxxxxxx wrote:
>>>
>>>> Hi,
>>>>
>>>> I am trying to work out the relationship
>>>> between the magnitude of the
>>>> vector of standardized regression
>>>> coefficients (beta coefficients) in
>>>> a multiple linear regression framework and
>>>> the coefficient of
>>>> determination (R^2) for the same model.
>>>> Following Bring (1996; Amer.
>>>> Stat. Assoc.), if all variables are
>>>> standardized, R^2 = ||y−hat||^2,
>>>> and ||y−hat||^2 = B1^2 + B2^2 + ... + Bk^2,
>>>> where Bk are the partial
>>>> regression coefficients. This implies that the
>>>> squared magnitude of
>>>> the Beta vector should equal R2. While I can
>>>> confirm this for real
>>>> data in the case of simple linear regression
>>>> (one independent
>>>> variable), it does not seem to work with
>>>> multiple independent
>>>> variables, so I must be doing something
>>>> wrong. Any suggestions would
>>>> be much appreciated.

Best,
Howard

Assisant Professor
Dept. of Biology, University of Ottawa
Ottawa, ON, Canada

When all the variables are standardized, R^2 is guaranteed to equal
B1^2 + B2^2 + ... + Bk^2 only when all the predictors are mutually
uncorrelated. In general, R^2 = r1*B1 + r2*B2 + ... + rk*Bk, where
ri is the correlation of the d.v. with predictor i.

It was my understanding that when there is multicolinearity, the individual betas become unreliable but the explanatory power of the entire model (i.e. R^2), and hence the sum of the squared Bi's, is not affected. For example, when I rerun the model using the principal components of the original variables in place of the variables (hence, they are mutually uncorrelated), R^2 doesn't change, nor does the sum of the squared Bi's. However, I am left with the problem that one still does not equal the other, and hence I am still confused.

It sounds like the program you're using defines the components as an orthonormal transformation of the data (i.e., a rotation, with possible reflection), without rescaling them to unit variance. Then the sum of squares of the regression weights would be unchanged, even though the components are uncorrelated.

Indeed this was the case. Everything's fine when I properly
standardize my principal components. R^2 = sum Bi^2.
Many thanks for your help.
.