
Subject: Re: reading old code... why would one do this?

Posted by [Chris\[6\]](#) on Thu, 22 Oct 2009 23:07:05 GMT

[View Forum Message](#) <> [Reply to Message](#)

On Oct 22, 11:36 am, wallabadah <write.to.wpow...@gmail.com> wrote:

> Hi All,
>
> I'm in the process of reviewing and rewriting some old routines I
> inherited... and I've come across the following code snippet:
>
> a = svdfit(x, y, fitting_order + 1, measure_errors = e, \$
> variance = var, chisq = c)
> ; this line added to give results consistent with IDL pre 5.4
> var = var * c
>
> I'm a bit worried about the last line - I don't have access to IDL
> documentation pre version 6.1, so I have two questions...
> - was there some change in the output of svdfit that would require
> such an 'adjustment' of the variance value?
> - if not - maybe the statistics gurus can tell me if variance * chisq
> has some mystical statistical meaning that I'm missing
>
> cheers,
>
> Will.

[begin speculation]

When fitting lines, the error bars (i.e. the variance) of the fitted parameters (slope and intercept) are determined entirely by the error bars in the individual data points. The scatter of those data points about the line is not considered, because the process assumes that the data's error bars accurately describe the amount of scatter. If this assumption is wrong (i.e. the data points have incorrect error bars) this manifests itself as an inappropriately high chi-squared value.

When the error bars are the correct size, the REDUCED chi squared (chi squared / number of data points) should be near one. If the value was 2, for example, that suggests that the data's error bars were too small by a factor of 2-ish, given their intrinsic scatter. Some people use this number to adjust the error bars on the individual data points (make them 2x bigger) which (i think) has the effect of making the fitted parameter error bars bigger by the same factor. It's not a super-kosher approach, since something may be wrong with the data or model assumptions if reduced chi-squared is much different than one.

However, that doesn't really explain what's going on with your code, since the variance is being multiplied by the unreduced chi-squared

value. That just seems wrong. And I don't know what the deal is with the compatibility comment.

[end speculation]

Chris

Subject: Re: reading old code... why would one do this?
Posted by [wallabadah](#) on Thu, 22 Oct 2009 23:47:31 GMT
[View Forum Message](#) <> [Reply to Message](#)

On Oct 23, 10:07 am, Chris <beaum...@ifa.hawaii.edu> wrote:
> On Oct 22, 11:36 am, wallabadah <write.to.wpow...@gmail.com> wrote:
>
>
>
>
>
>
>> Hi All,
>
>> I'm in the process of reviewing and rewriting some old routines I
>> inherited... and I've come across the following code snippet:
>
>> a = svdfit(x, y, fitting_order + 1, measure_errors = e, \$
>> variance = var, chisq = c)
>> ; this line added to give results consistent with IDL pre 5.4
>> var = var * c
>
>> I'm a bit worried about the last line - I don't have access to IDL
>> documentation pre version 6.1, so I have two questions...
>> - was there some change in the output of svdfit that would require
>> such an 'adjustment' of the variance value?
>> - if not - maybe the statistics gurus can tell me if variance * chisq
>> has some mystical statistical meaning that I'm missing
>
>> cheers,
>
>> Will.
>
> [begin speculation]
>
> When fitting lines, the error bars (i.e. the variance) of the fitted
> parameters (slope and intercept) are determined entirely by the error
> bars in the individual data points. The scatter of those data points
> about the line is not considered, because the process assumes that the
> data's error bars accurately describe the amount of scatter. If this
> assumption is wrong (i.e. the data points have incorrect error bars)

> this manifests itself as an inappropriately high chi-squared value.
>
> When the error bars are the correct size, the REDUCED chi squared (chi
> squared / number of data points) should be near one. If the value was
> 2, for example, that suggests that the data's error bars were too
> small by a factor of 2-ish, given their intrinsic scatter. Some people
> use this number to adjust the error bars on the individual data points
> (make them 2x bigger) which (i think) has the effect of making the
> fitted parameter error bars bigger by the same factor. It's not a
> super-kosher approach, since something may be wrong with the data or
> model assumptions if reduced chi-squared is much different than one.
>
> However, that doesn't really explain what's going on with your code,
> since the variance is being multiplied by the unreduced chi-squared
> value. That just seems wrong. And I don't know what the deal is with
> the compatibility comment.
>
> [end speculation]
>
> Chris

Thanks Chris, such speculation is just what I require - not easy to find in any textbook! The errors in the individual data points come from poisson or counting statistics - I don't see any reason to suspect they're incorrect, but I'll have a look at the reduced chi-squared numbers...
