
Subject: Re: Missing data and MPFITFUN
Posted by [Allan Whiteford](#) on Mon, 23 Oct 2006 11:09:51 GMT
[View Forum Message](#) <> [Reply to Message](#)

googlegroups@higginsnet.com wrote:
> I'm trying to fit a function to a data set with many missing data
> points. E.g. (made up data:)
>
> x | y
> 0 2,3,6,7,7
> 1 3,4,5,NaN,8
> 2 2,NaN,5,7,NaN
>
> Having missing data values, or inserting NaN (where NaN =
> !VALUES.F_NAN), trips up MPFITFUN.
>
> Is there a way of saying to IDL "just ignore missing values"?
>

Set the weight of the missing data values to zero when doing the fit.

Thanks,

Allan

Subject: Re: Missing data and MPFITFUN
Posted by [googlegroups](#) on Mon, 23 Oct 2006 13:57:51 GMT
[View Forum Message](#) <> [Reply to Message](#)

Thanks.

I used WHERE to find the values and create the WEIGHTS array.

Subject: Re: Missing data and MPFITFUN
Posted by [David Fanning](#) on Mon, 23 Oct 2006 14:20:28 GMT
[View Forum Message](#) <> [Reply to Message](#)

Allan Whiteford writes:

> Set the weight of the missing data values to zero when doing the fit.

Humm. There must be more to it than that. -(
Consider this test program:

PRO TEST

```

npts = 101
time = Findgen(npts)
signal = 4.3 * Findgen(npts)^2.52
noisySignal = (Randomu(-3L, N_Elements(signal)) - 0.5) * 50000L + $
    signal
noisySignal[ [ 3, 5, 22, 54, 66, 87] ] = !VALUES.F_NAN
Plot, time, noisySignal, PSym=2
model = 'p[0] * x^p[1]'
error = Replicate(100, npts)
params = [1.0,2.0]
weights = Replicate(1.0, npts)
weights[ [ 3, 5, 22, 54, 66, 87] ] = 0.0
fit = MPFITEXPR(model, time, noisySignal, error, params, $
    WEIGHTS=weights)
OPlot, time, fit[0] * Findgen(npts)^fit[1]
END

```

If you comment out the line where I set certain values to !VALUES.F_NAN the program works great. If I leave the line uncommented, it doesn't work so well, even though the WEIGHTS have been set to 0. :-)

Cheers,

David

--

David Fanning, Ph.D.

Fanning Software Consulting, Inc.

Coyote's Guide to IDL Programming: <http://www.dfanning.com/>

Sepore ma de ni thui. ("Perhaps thou speakest truth.")

Subject: Re: Missing data and MPFITFUN

Posted by [Craig Markwardt](#) on Mon, 23 Oct 2006 14:44:28 GMT

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

David Fanning <news@dfanning.com> writes:

> Allan Whiteford writes:

>

>> Set the weight of the missing data values to zero when doing the fit.

>

> Humm. There must be more to it than that. -(

> Consider this test program:

David, the data values should still be finite. Multiplying (VALUE*0) = (NaN*0) is still NaN. Remove the NaNs, include zero-weights, and you should be fine.

You might ask, why not ignore points with NaNs? The answer is, there are too many people who were blithely allowing their data or model to go infinite and then complaining to me when MPFIT didn't work. I prefer to force the user to explicitly set the weight to zero for points they want to ignore, rather than silently ignoring them. Then if non-finite values show up, MPFIT knows it's a problem.

Perhaps I should add a /NAN keyword though.

Craig

```
>
> PRO TEST
> npts = 101
> time = Findgen(npts)
> signal = 4.3 * Findgen(npts)^2.52
> noisySignal = (Randomu(-3L, N_Elements(signal)) - 0.5) * 50000L + $
>     signal
> noisySignal[ [ 3, 5, 22, 54, 66, 87] ] = !VALUES.F_NAN
> Plot, time, noisySignal, PSym=2
> model = 'p[0] * x^p[1]'
> error = Replicate(100, npts)
> params = [1.0,2.0]
> weights = Replicate(1.0, npts)
> weights[ [ 3, 5, 22, 54, 66, 87] ] = 0.0
> fit = MPFITEXPR(model, time, noisySignal, error, params, $
>     WEIGHTS=weights)
> OPlot, time, fit[0] * Findgen(npts)^fit[1]
> END
>
> If you comment out the line where I set certain values
> to !VALUES.F_NAN the program works great. If I leave
> the line uncommented, it doesn't work so well, even
> though the WEIGHTS have been set to 0. :-(
>
> Cheers,
>
> David
>
> --
> David Fanning, Ph.D.
> Fanning Software Consulting, Inc.
> Coyote's Guide to IDL Programming: http://www.dfanning.com/
> Sepore ma de ni thui. ("Perhaps thou speakest truth.")
```

--

Craig Markwardt, Ph.D.
EMAIL: craigm@lheamail.gsfc.nasa.gov

Subject: Re: Missing data and MPFITFUN
Posted by [David Fanning](#) on Mon, 23 Oct 2006 14:54:56 GMT
[View Forum Message](#) <> [Reply to Message](#)

Craig Markwardt writes:

> You might ask, why not ignore points with NaNs? The answer is, there
> are too many people who were blithely allowing their data or model to
> go infinite and then complaining to me when MPFIT didn't work. I
> prefer to force the user to explicitly set the weight to zero for
> points they want to ignore, rather than silently ignoring them. Then
> if non-finite values show up, MPFIT knows it's a problem.
>
> Perhaps I should add a /NAN keyword though.

I've tried educating the masses. I've, uh, mostly given up. :-)

If you are going to add a NAN keyword though (which I think is a good idea), I would also consider setting the weights at that location to 0 yourself. (This is probably what you had in mind.) I'm pretty sure users won't remember to do it.

Thanks for the help!

Cheers,

David

--

David Fanning, Ph.D.
Fanning Software Consulting, Inc.
Coyote's Guide to IDL Programming: <http://www.dfanning.com/>
Sepore ma de ni thui. ("Perhaps thou speakest truth.")

Subject: Re: Missing data and MPFITFUN
Posted by [googlegroups](#) on Tue, 24 Oct 2006 13:51:03 GMT
[View Forum Message](#) <> [Reply to Message](#)

Craig Markwardt wrote:

> Remove the NaNs, include zero-weights, and
> you should be fine.

Does the calculation of the covariance matrix take into account the weighting matrix?

I have missing data points in an array, and so I've filled those data points with zeros, and then created a WEIGHTS matrix which identifies those zeros - to be ignored in the fit.

I don't want my use of zeros instead of missing data to contaminate my results.

Thanks
Dave Higgins

Subject: Re: Missing data and MPFITFUN
Posted by [Craig Markwardt](#) on Sat, 28 Oct 2006 18:57:03 GMT
[View Forum Message](#) <> [Reply to Message](#)

googlegroups@higginsnet.com writes:

> Craig Markwardt wrote:
>
>> Remove the NaNs, include zero-weights, and
>> you should be fine.
>
> Does the calculation of the covariance matrix take into account the
> weighting matrix?

Yes. The calculation of the covariance matrix derives from the same matrix that actually does the fitting.

> I have missing data points in an array, and so I've filled those data
> points with zeros, and then created a WEIGHTS matrix which identifies
> those zeros - to be ignored in the fit.

You did the right thing.

Craig

--

Craig B. Markwardt, Ph.D. EMAIL: craigmnet@REMOVEcow.physics.wisc.edu
Astrophysics, IDL, Finance, Derivatives | Remove "net" for better response
