Subject: Re: Time series with 75% missing observations
Posted by Craig Markwardt on Wed, 17 Oct 2001 22:59:38 GMT
View Forum Message <> Reply to Message

Joe Means <joe.means@orst.edu> writes:

> I have long data series with many random, missing values. These series
> each have only one or two frequencies. As I read the IDL 5.4
> documentation, time series analysis routines require all values to be
> evenly spaced. The periodicities are not sinusoidal, but such an
> assumption might well find them. How can I find these periodicities?

Greetings!

If your time series are regularly sampled, but have missing values, then it is relatively straightforward to proceed with the FFT.

Theoretically, zeroes do not actually contribute to the FFT power, but in practice you get aliases of the DC power which contaminate all the other frequencies. Contamination is bad.

The solution is to subtract off the mean value of the signal *from the non-missing values* before doing the FFT. Say you have a variable Y which contains the signal, and missing values are set to -1. First you would transform to a new variable, YP, which has the average value subtracted.

```
yp = y*0
wh = where(y NE -1, ct)
if ct EQ 0 then message, 'ERROR: no valid points!'
yp(wh) = y(wh) - avg(y(wh))
```

Note that the original missing values are converted to zeroes, so overall YP should have a zero mean value itself. Then you just do your FFT as normal.

***

The other possibility, is if you have irregularly sampled points. Then you are better off with something like the Lomb Scargle periodogram. It's in Numerical Recipes. I have a crude routine which does this, available by request.

Craig

--
---------- ----------

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

Subject: Re: Time series with 75% missing observations
Posted by Roman Schreiber on Thu, 18 Oct 2001 07:42:20 GMT
View Forum Message <> Reply to Message

Craig Markwardt <craigmnet@cow.physics.wisc.edu> wrote in
news:onpu7lub39.fsf@cow.physics.wisc.edu:

>
> Joe Means <joe.means@orst.edu> writes:
>
>> I have long data series with many random, missing values. These series
>> each have only one or two frequencies. As I read the IDL 5.4
>> documentation, time series analysis routines require all values to be
>> evenly spaced. The periodicities are not sinusoidal, but such an
>> assumption might well find them. How can I find these periodicities?
>
> Greetings!
>
> If your time series are regularly sampled, but have missing values,
> then it is relatively straightforward to proceed with the FFT.
>
> Theoretically, zeroes do not actually contribute to the FFT power, but
> in practice you get aliases of the DC power which contaminate all the
> other frequencies. Contamination is bad.
>
> The solution is to subtract off the mean value of the signal *from the
> non-missing values* before doing the FFT. Say you have a variable Y
> which contains the signal, and missing values are set to -1. First
> you would transform to a new variable, YP, which has the average value
> subtracted.
>
> yp = y*0
> wh = where(y NE -1, ct)
> if ct EQ 0 then message, 'ERROR: no valid points!'
> yp(wh) = y(wh) - avg(y(wh))
>
> Note that the original missing values are converted to zeroes, so
> overall YP should have a zero mean value itself. Then you just do
> your FFT as normal.
>
> ***
>
> The other possibility, is if you have irregularly sampled points.

> Then you are better off with something like the Lomb Scargle
> periodogram. It's in Numerical Recipes. I have a crude routine which
> does this, available by request.
>
> Craig
>
>

Joe,

there is already Lomb periodogram function implemented in IDL 5.4
(LNP_TEST) based on routine fasper you may find in Numerical Recipes.

Best regards

Roman

--
****
Roman Schreiber e-mail: schreibe@ncac.torun.pl
Copernicus Astronomical Centre PAS phone: 0-48-56-6219319
Astroph.Lab I fax: 0-48-56-6219381
ul.Rabianska 8
87-100 Torun POLAND
****

Subject: Re: Time series with 75% missing observations
Posted by Joe Means on Thu, 18 Oct 2001 16:00:41 GMT
View Forum Message <> Reply to Message

Thanks much Craig. This should work great!
Cheers,
Joe

Craig Markwardt wrote:

> Joe Means <joe.means@orst.edu> writes:
>
>> I have long data series with many random, missing values. These series
>> each have only one or two frequencies. As I read the IDL 5.4
>> documentation, time series analysis routines require all values to be
>> evenly spaced. The periodicities are not sinusoidal, but such an
>> assumption might well find them. How can I find these periodicities?
>
>
> Greetings!

>
> If your time series are regularly sampled, but have missing values,
> then it is relatively straightforward to proceed with the FFT.
>
> Theoretically, zeroes do not actually contribute to the FFT power, but
> in practice you get aliases of the DC power which contaminate all the
> other frequencies. Contamination is bad.
>
> The solution is to subtract off the mean value of the signal *from the
> non-missing values* before doing the FFT. Say you have a variable Y
> which contains the signal, and missing values are set to -1. First
> you would transform to a new variable, YP, which has the average value
> subtracted.
>
> yp = y*0
> wh = where(y NE -1, ct)
> if ct EQ 0 then message, 'ERROR: no valid points!'
> yp(wh) = y(wh) - avg(y(wh))
>
> Note that the original missing values are converted to zeroes, so
> overall YP should have a zero mean value itself. Then you just do
> your FFT as normal.
>
> ***
>
> The other possibility, is if you have irregularly sampled points.
> Then you are better off with something like the Lomb Scargle
> periodogram. It's in Numerical Recipes. I have a crude routine which
> does this, available by request.
>
> Craig
>
>
