
Subject: Re: how to extend a vector from size A to size B, (A<B) without "damage" the data in A

Posted by [Craig Markwardt](#) on Thu, 08 Mar 2012 05:37:41 GMT

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

On Wednesday, March 7, 2012 4:51:20 PM UTC-5, Tito wrote:

> Hi all!

>

> I have the fowling problem:

> I have a array vector A with size [970] elements lets say, and another that is 'B' with [1024] elements. now... both contain spectra and I want to do cross-correlation between the two vectors. This will be possible if A is 1024 element vector.

>

> I am new in IDL and shoot and try several thing here like:

> A = interpol(A,1024)

> and I saw also someone suggested:

> A = [A, intarr(n_elements(B)-n_elements(A))] , but doesn't give me what I want.

>

> A is synthetic spectra (mask) and I just want to increase the resolution without losing the quality of the data(the relative distances between the lines)

>

> A = interpol(A,1024), actually works, but I see very obvious differences in the 970 and the result spectra.

I think you need to define your problem a little more clearly.

Either your A spectrum is sampled on a different grid than B, or it is simply truncated. Either way you have to make an assumption about the quality of your data.

Sampled on a different grid means that the spacing between measurement points in A was different by about 5.3%, or it could also mean uneven sampling of A. Your best bet in this case is to do some form of interpolation, to put them on the same grid. The assumption you would need to make here is that the values vary smoothly from point to point, so that the missing data can be represented by its interpolant. Yes, there could be aliasing artifacts, but I think those would be minimized by doing sinc-interpolation, maybe cubic spline is OK. Both `spl_init/spl_interp` and `interpolate()` offer smoothly varying interpolation functions.

If instead A is *truncated* then things are easier. You need to pad your A to the full length. The assumption you make here is that the missing data do not contain useful information, so there is no harm in padding the series. It's probably best to pad with the mean value of your series instead of zero like you indicated above.

Best wishes,
Craig
