Subject: Re: fitting mixed gaussians
Posted by Craig Markwardt on Thu, 19 May 2005 00:37:28 GMT
View Forum Message <> Reply to Message

Ben Tupper <btupper@bigelow.org> writes:

> Rob wrote:
>> Thanks for the plug, David. Yep, PAN is just a big GUI wrapper for
>> MPCURVEFIT. In fact the kind of fitting described in the original post
>> is done routinely at our neutron scattering facility where the whole
>> model function can be composed of many Gaussians, Lorentzians,
>> Lognormals, etc.
>>
>
>>>
>>> Just draw and adjust the initial curve parameters on
>>> the data itself, then go fit it with a click of the button.
>>> Of course, it is simply a wrapper for MPFIT. The tutorial
>>> Rob has provided is helpful for getting started.
>>>
> Thanks David and Rob,
>
> Yes I looked at these - the trouble is that I want to perform this on
> literally thousands of images (each image has one object in it).
> Manual seeding is not practical. I think what I am after is described
> here...
>
> http://tinyurl.com/9horr
>
> I have started to translate the MatLab code to IDL - but it is clearly
> over my head.

Long ago I had somebody write to me, asking about how to do this kind of thing. I don't think he ever succeeded.

You are going to have problems like:

1. uniqueness - there are essentially an infinite number of ways to add gaussians to reproduce the data; for example, why not have one gaussian per sample?

2. robustness - the problem is so unconstrained that there is significant potential for screwed up fits.

I would recommend constraining the analysis as much as possible based on your problem domain, for example if you know that peaks must be positive, or the natural widths of the peaks, etc.

One technique would be to find the tallest peak, or perhaps the N tallest peaks in the data, and fit gaussians to those. Then subtract that fit and see what the next highest peaks are, and fit another set of gaussians. You keep doing that until you reach some noise threshold (i.e. the errors on the amplitudes are comparable to the amplitudes). Sometimes this is known as "CLEAN" in image space.

But be prepared for the fitting to fail or give whacked results at least 30% of the time.

Good luck,
Craig

--

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