Subject: Re: Non-linear data selection
Posted by Joe Daal on Sun, 16 May 2010 00:01:43 GMT
View Forum Message <> Reply to Message

Sorry, I meant to say: ...I need to select all the point ABOVE this
line...

Subject: Re: Non-linear data selection
Posted by David Fanning on Sun, 16 May 2010 00:27:41 GMT
View Forum Message <> Reply to Message

Joe Daal writes:

> There is an interesting problem that I often come across.....
> I have a scattered dataset of x and y, say 1000 elements each. The
> data could be of any shape, for now, let's assume it is scattered
> allover the X-Y window.
>
> There is straight line (y=mx+b, could be some other curve as well)
> that that crosses the plot, I need to select all data points this
> line.
> If it is just a one value for y, the WHERE statement is ok, but how to
> do this?

```
x = Randomu(-3L, 100) * 10
y = Randomu(-5L, 100) * 10
m = 0.5
b = 3.25
yfit = m*x+b
aboveIndices = Where((y GE yfit) EQ 1, count)
Print, StrTrim(count,2) + ' points lie above the line.'

Plot, x, y, /NoData, Background=FSC_Color('white')
OPlot, x, y, PSYM=2, SYMSIZE=0.5, Color=FSC_Color('sky blue')
OPLOT, x[aboveIndices], y[aboveIndices], PSYM=4, Color=FSC_Color('red')
OPlot, x, yfit, LINESTYLE=2
```

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: Non-linear data selection
Posted by Joe Daal on Sun, 16 May 2010 02:18:16 GMT
View Forum Message <> Reply to Message

This is it.
Thanks a lot, David!

Subject: Re: Non-linear data selection
Posted by cgguido on Sun, 16 May 2010 07:18:45 GMT
View Forum Message <> Reply to Message

On May 15, 7:27 pm, David Fanning <n...@dfanning.com> wrote:
> aboveIndices = Where((y GE yfit) EQ 1, count)

David, why didn't you use one of the following?

aboveIndices2 = Where(y GE yfit, count)
aboveIndices3 = Where(y[*] GE yfit[*], count)

I have this vague memory/nightmare that I "need" to do the [*] part in
order not to get crazy results sometimes (though I can't come up with
a crazy example right now) so I always use aboveIndices3...

So *are* there differences among these three lines? Cos they all seem
to work for the example of the original poster. :-|

Thanks,
Gianguido

Subject: Re: Non-linear data selection
Posted by David Fanning on Sun, 16 May 2010 13:16:22 GMT
View Forum Message <> Reply to Message

Gianguido Cianci writes:

>
> On May 15, 7:27 pm, David Fanning <n...@dfanning.com> wrote:
>> aboveIndices = Where((y GE yfit) EQ 1, count)
>
> David, why didn't you use one of the following?
>
> aboveIndices2 = Where(y GE yfit, count)

>   aboveIndices3 = Where(y[*] GE yfit[*], count)
>
> I have this vague memory/nightmare that I "need" to do the [*] part in
> order not to get crazy results sometimes (though I can't come up with
> a crazy example right now) so I always use aboveIndices3...
>
> So *are* there differences among these three lines? Cos they all seem
> to work for the example of the original poster. :-|

The second line (aboveIndices2) is what I originally
wrote in my code, but I changed it to give more clues
to the user about what I was doing. (I was too busy working
on my rock garden yesterday to explain myself.)

The third line is totally unnecessary and, at least with
large arrays, will cost you a big memory penalty:

   http://www.dfanning.com/misc_tips/submemory.html

But, yes, they should all be equivalent in this case.

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.")
