
Subject: replace integration by summation
Posted by [fd_luni](#) on Tue, 18 Jun 2013 10:36:02 GMT
[View Forum Message](#) <> [Reply to Message](#)

Hi,
I've been trying to replace the INT_TABULATED by total but the answer is very different and I don't see why.

integration: $A2 = \text{INT_TABULATED}(t[0:i], A1[0:i])$

summation: $A2 = (t[1]-t[0]) * \text{total}(A1, / \text{cumulative})$

I haven't been able to find out what is going wrong because these two commands seems similar to me and I didn't expect to get so wrong results. Can anyone help?

With Thanks
Maria

Subject: Re: replace integration by summation
Posted by [Helder Marchetto](#) on Tue, 18 Jun 2013 10:56:17 GMT
[View Forum Message](#) <> [Reply to Message](#)

On Tuesday, June 18, 2013 12:36:02 PM UTC+2, [fd_...@mail.com](#) wrote:

> Hi,
>
> I've been trying to replace the INT_TABULATED by total but the answer is very different and I don't see why.
>
>
>
> integration: $A2 = \text{INT_TABULATED}(t[0:i], A1[0:i])$
>
>
>
> summation: $A2 = (t[1]-t[0]) * \text{total}(A1, / \text{cumulative})$
>
>
>
> I haven't been able to find out what is going wrong because these two commands seems similar to me and I didn't expect to get so wrong results. Can anyone help?
>
>
>
> With Thanks
>
> Maria

Well, I would suppose that the two types of integration methods are very different in how they integrate. In one case the so-called method of the rectangle rule (summation of array elements with /cumulative) will give very different results from int_tabulated specially when the data is very noisy and you're not sampling enough...

So, I would guess that the problem are the number of points you are using are not enough to describe the noise.

Try playing around with this:

```
x = (findgen(1001)/1000.0)*!pi
y = sin(x)*(RANDOMU(S,1001)*0.5)
print, (x[1]-x[0])*(total(y,/cumulative))[-1], INT_TABULATED(x,y)
```

the less points you use and the higher the noise, the more the results will differ (not tested, it's a guess...)

Cheers,
Helder

Subject: Re: replace integration by summation
Posted by [fd_luni](#) on Tue, 18 Jun 2013 14:13:38 GMT
[View Forum Message](#) <> [Reply to Message](#)

Oh, I see. I had check it and your guess is right. There is any other function that I can use instead of INT_TABULATED?

Subject: Re: replace integration by summation
Posted by [Craig Markwardt](#) on Tue, 18 Jun 2013 16:04:11 GMT
[View Forum Message](#) <> [Reply to Message](#)

On Tuesday, June 18, 2013 10:13:38 AM UTC-4, fd_...@mail.com wrote:
> Oh, I see. I had check it and your guess is right. There is any other function that I can use instead of INT_TABULATED?

To build on what Helder said, INT_TABULATED() and TOTAL() are approximations based on assumptions about the data.

Your TOTAL() method is basically equivalent to the Trapezoid rule, which assumes that your function can be approximated as piecewise linear.

Also, by using (t[1]-t[0]) you are also assuming that your points are exactly regularly spaced. Is that a correct assumption?

INT_TABULATED uses a fifth order Newton-Cotes method, so it basically assumes the tabulated function is a smooth fifth-order polynomial (I think).

If these assumptions are wrong, then you won't get the answer you want. But only you know which assumption is appropriate. You don't really say what is "wrong" about the results you got, so it's hard to judge.

There are other functions like QTRAP, QSIMP, QROMB (and my own QPINT1D), but these methods require you to express your function not as a table, but as an IDL function that can be evaluated at any point. These more advanced functions can be more accurate because they can concentrate on the parts of the function where the integration error is greatest and reduce the error.

IDL doesn't have any other built-in subroutines for integration of tabulated values.

Craig Markwardt

Subject: Re: replace integration by summation
Posted by [Phillip Bitzer](#) on Tue, 18 Jun 2013 16:17:49 GMT
[View Forum Message](#) <> [Reply to Message](#)

This may help. It's a little write up about numeric integration, comparing the simplest methods.

<https://docs.google.com/file/d/0BxRn65i6z-Z5T283SzZjZFNOVms/edit?usp=sharing>

The best routine to use depends how the data varies. What is "wrong" with the answers?

Subject: Re: replace integration by summation
Posted by [David Fanning](#) on Tue, 18 Jun 2013 16:37:37 GMT
[View Forum Message](#) <> [Reply to Message](#)

Phillip Bitzer writes:

> This may help. It's a little write up about numeric integration, comparing the simplest methods.
>
> <https://docs.google.com/file/d/0BxRn65i6z-Z5T283SzZjZFNOVms/edit?usp=sharing>

Phillip, that's quite a book you are writing. When can we get our hands on it? :-)

Cheers,

David

--

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

Subject: Re: replace integration by summation
Posted by [Fabzi](#) on Tue, 18 Jun 2013 16:42:58 GMT
[View Forum Message](#) <> [Reply to Message](#)

Hi everyone,

On 06/18/2013 06:17 PM, Phillip Bitzer wrote:

> The best routine to use depends how the data varies. What is "wrong" with the answers?

I will ruthlessly use this thread to ask my question to the atmospheric modellers in the IDL community out here. I am also secretly asking myself how many atmospheric modellers are reading this group ;-)

I am vertically integrating a quantity from an atmospheric model output (in this case: moisture flux) over the atmospheric column. I made some searches and it came out that some people use trapezoidal rule for this, some use the midpoint approximation rectangle rule (http://en.wikipedia.org/wiki/Rectangle_method). There aren't many more options because the data is tabulated (z-coordinates=pressure, values=flux).

To me, the rectangle rule makes more sense from the "gridded point of view" of atmospheric models. Does anyone have a hint or a good reference explaining how this should be "correctly done" in this case? Thanks!

Fab

Subject: Re: replace integration by summation
Posted by [Phillip Bitzer](#) on Tue, 18 Jun 2013 18:46:26 GMT
[View Forum Message](#) <> [Reply to Message](#)

On Tuesday, June 18, 2013 11:37:37 AM UTC-5, David Fanning wrote:

>
> Phillip, that's quite a book you are writing. When can we get our hands on it? :-)
>

I'm hoping to make an "alpha" release relatively soon. Way too many typos at this point, plots that need updating, etc. :-)

As an FYI, these notes are associated with a programming class I teach in the Atmospheric Science department. And since my main research area is lightning, there's a quite a bit on how to use IDL to look at this sort of data.

Subject: Re: replace integration by summation
Posted by [Phillip Bitzer](#) on Tue, 18 Jun 2013 18:47:15 GMT
[View Forum Message](#) <> [Reply to Message](#)

On Tuesday, June 18, 2013 11:42:58 AM UTC-5, Fabien wrote:

> Hi everyone,
>
>
> I will ruthlessly use this thread to ask my question to the atmospheric
>
> modellers in the IDL community out here. I am also secretly asking
>
> myself how many atmospheric modellers are reading this group ;-)
>
>
>
> I am vertically integrating a quantity from an atmospheric model output
>
> (in this case: moisture flux) over the atmospheric column. I made some
>
> searches and it came out that some people use trapezoidal rule for this,
>
> some use the midpoint approximation rectangle rule
>
> (http://en.wikipedia.org/wiki/Rectangle_method). There aren't many more
>
> options because the data is tabulated (z-coordinates=pressure,
>
> values=flux).
>
> To me, the rectangle rule makes more sense from the "gridded point of
>
> view" of atmospheric models. Does anyone have a hint or a good reference
>
> explaining how this should be "correctly done" in this case? Thanks!
>
>
>
> Fab

Well, I wouldn't really consider myself an atmospheric modeler, but I would think it depends on how much the data varies within the grid points. In the link I posted, the implicit underlying function describing the data varies quite a bit in between grid points, so the rectangle method doesn't work

very well. If I had a finer grid (which in calculus is equivalent to letting $\Delta x \rightarrow dx$, an infinitesimal differential), then the rectangle method should work better. On the other hand, the trapezoidal method works fine since it's a pretty good approximation to the underlying function. INT_TABULATED works better since it's the fifth order Newton-Cotes method. I believe the trapezoid method is the first order Newton-Cotes.

There are likely other folks better suited to give a more "formal" answer....

Subject: Re: replace integration by summation
Posted by [David Fanning](#) on Tue, 18 Jun 2013 18:56:30 GMT
[View Forum Message](#) <> [Reply to Message](#)

Phillip Bitzer writes:

> I'm hoping to make an "alpha" release relatively soon. Way too many typos at this point, plots that need updating, etc. :-)

Let me know if you want a couple of ruthless reviewers. I know a couple of people who don't let you get with ANYTHING! :-)

Cheers,

David

--

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

Subject: Re: replace integration by summation
Posted by [David Fanning](#) on Tue, 18 Jun 2013 18:59:30 GMT
[View Forum Message](#) <> [Reply to Message](#)

David Fanning writes:

> Let me know if you want a couple of ruthless reviewers. I know a couple
> of people who don't let you get with ANYTHING! :-)

In fact, I've already heard from them, pointing out that my sentence here is ungrammatical and missing a word. :-)

Cheers,

David

--

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

Subject: Re: replace integration by summation
Posted by [fd_luni](#) on Wed, 19 Jun 2013 09:01:32 GMT
[View Forum Message](#) <> [Reply to Message](#)

> You don't really say what is "wrong" about the results you got, so it's hard to judge.

I got all zeros.

Maria

Subject: Re: replace integration by summation
Posted by [Paul Van Delst\[1\]](#) on Wed, 19 Jun 2013 12:08:17 GMT
[View Forum Message](#) <> [Reply to Message](#)

Hmm... an all-zero result is typically an indication of user error (it's pretty difficult to get a bunch of numbers to add up to zero).

What about if, instead of
A2= INT_TABULATED(t[0:i], A1[0:i])
you do
A2= INT_TABULATED(t, A1)
?

What do you get?

In your original post you don't use bounds in the TOTAL() example, so may as well do the same in the INT_TABULATED() one.

cheers,

paulv

On 06/19/13 05:01, fd_luni@mail.com wrote:

>> You don't really say what is "wrong" about the results you got, so it's hard to judge.

>
> I got all zeros.
>
> Maria
>

Subject: Re: replace integration by summation
Posted by [fd_luni](#) on Wed, 19 Jun 2013 12:55:37 GMT
[View Forum Message](#) <> [Reply to Message](#)

On Wednesday, 19 June 2013 13:08:17 UTC+1, Paul van Delst wrote:

> Hmm... an all-zero result is typically an indication of user error (it's pretty difficult to get a bunch of numbers to add up to zero). What about if, instead of `A2= INT_TABULATED(t[0:i], A1[0:i])` you do `A2= INT_TABULATED(t, A1)` ? What do you get? In your original post you don't use bounds in the `TOTAL()` example, so may as well do the same in the `INT_TABULATED()` one.

> I actually have a loop
For i=1,n-1 do begin
A1[i]= INT_TABULATED(t[0:i], A2[0:i])
endfor

When I use `A2= INT_TABULATED(t, A1)` I got a single value. I need an array that is why I used
`A1 = (t[1]-t[0])*total(A2,/cumulative)`

I am actually try to avoid the loop and replace it by something else. For this reason I used the `A1 = (t[1]-t[0])*total(A2,/cumulative)`.

Subject: Re: replace integration by summation
Posted by _____ on Wed, 19 Jun 2013 13:02:57 GMT
[View Forum Message](#) <> [Reply to Message](#)

Den onsdagen den 19:e juni 2013 kl. 14:55:37 UTC+2 skrev [fd_...@mail.com](#):

> On Wednesday, 19 June 2013 13:08:17 UTC+1, Paul van Delst wrote:

>
>> Hmm... an all-zero result is typically an indication of user error (it's pretty difficult to get a bunch of numbers to add up to zero). What about if, instead of `A2= INT_TABULATED(t[0:i], A1[0:i])` you do `A2= INT_TABULATED(t, A1)` ? What do you get? In your original post you don't use bounds in the `TOTAL()` example, so may as well do the same in the `INT_TABULATED()` one.
>
>
>
>> I actually have a loop
>
> For i=1,n-1 do begin
>

```
> A1[i]= INT_TABULATED(t[0:i], A2[0:i])
>
> endfor
>
>
>
> When I use A2= INT_TABULATED(t, A1) I got a single value. I need an array that is why I used
>
>
> A1 = (t[1]-t[0])*total(A2,/cumulative)
>
>
>
> I am actually try to avoid the loop and replace it by something else. For this reason I used the
A1 = (t[1]-t[0])*total(A2,/cumulative).
```

You are switching between calculating A1 as a sum/integral of A2 and A2 as a sum/integral of A1. Is there maybe a typo in your code?

Subject: Re: replace integration by summation
Posted by [fd_luni](#) on Wed, 19 Jun 2013 13:09:15 GMT
[View Forum Message](#) <> [Reply to Message](#)

> You are switching between calculating A1 as a sum/integral of A2 and A2 as a sum/integral of A1. Is there maybe a typo in your code?

No, I wrote it wrong here. In my code is correct

integration: A2= INT_TABULATED(t[0:i], A1[0:i])

summation: A2 = (t[1]-t[0])*total(A1,/cumulative)

Subject: Re: replace integration by summation
Posted by [Craig Markwardt](#) on Wed, 19 Jun 2013 16:29:15 GMT
[View Forum Message](#) <> [Reply to Message](#)

On Wednesday, June 19, 2013 8:55:37 AM UTC-4, fd_...@mail.com wrote:

> On Wednesday, 19 June 2013 13:08:17 UTC+1, Paul van Delst wrote:

>

>> Hmm... an all-zero result is typically an indication of user error (it's pretty difficult to get a bunch of numbers to add up to zero). What about if, instead of A2= INT_TABULATED(t[0:i], A1[0:i]) you do A2= INT_TABULATED(t, A1) ? What do you get? In your original post you don't use bounds in the TOTAL() example, so may as well do the same in the INT_TABULATED() one.

>

>

```
>
>> I actually have a loop
>
> For i=1,n-1 do begin
>
> A1[i]= INT_TABULATED(t[0:i], A2[0:i])
>
> endfor
>
> When I use A2= INT_TABULATED(t, A1) I got a single value. I need an array that is why I used
>
> A1 = (t[1]-t[0])*total(A2,/cumulative)
>
> I am actually try to avoid the loop and replace it by something else. For this reason I used the
A1 = (t[1]-t[0])*total(A2,/cumulative).
```

But when you took Mats's suggestion and computed INT_TABULATED(t,A1), was the single value zero or not?

Craig

Subject: Re: replace integration by summation
Posted by [fd_luni](#) on Wed, 19 Jun 2013 20:05:20 GMT
[View Forum Message](#) <> [Reply to Message](#)

> But when you took Mats's suggestion and computed INT_TABULATED(t,A1), was the single value zero or not?

No it was not a single value zero.

I had two function like this:

```
For i=1,n-1 do begin
A2= INT_TABULATED(t[0:i], A1[0:i])
B2= INT_TABULATED(t[0:i], B1[0:i])
endfor
```

When I replaced the INT_TABULATED by this:

```
A2 = (t[1]-t[0])*total(A1,/cumulative)
B2 = (t[1]-t[0])*total(B1,/cumulative)
```

The function A2 = (t[1]-t[0])*total(A1,/cumulative) gives me completely different values from A2= INT_TABULATED(t[0:i], A1[0:i]). But the function B2 = (t[1]-t[0])*total(B1,/cumulative gives me zeros.

Subject: Re: replace integration by summation
Posted by [Phillip Bitzer](#) on Wed, 19 Jun 2013 21:04:32 GMT
[View Forum Message](#) <> [Reply to Message](#)

OK, let's go with a simple example.

```
Let:
x= FINDGEN(4)
y=x^2
```

Your loop over INT_TABULATED will give the area under the curve for $x=0 \rightarrow x[i]$; this means

```
tab = FLTARR(4)
FOR i=1, 3 DO tab[i] = INT_TABULATED(x[0:i], y[0:i])
```

yields

```
print, tab
      0.00000   0.500000   2.75556   9.10000
```

You can check these aren't quite the correct answers, *for the integration*, given the (known) underlying function, but are fairly close. The discrepancy is caused by the rather coarse grid of $dx=1$.

For example, the integral of x^2 between 0 and 2 ($i=2$) is analytically $8/3=2.67$. This method is off by about 5%.

The method of using the cumulative total is *not* the area under the curve, i.e., it's not integration. In this example,

```
tot = total(y, /cum)*(x[1]-x[0])
print, tot
      0.00000   1.00000   5.00000   14.0000
```

Clearly, this is not doing the same thing as integration. In this case, for $i=2$ you are finding the area of a rectangle 5 tics tall and 1 tics wide. This is not the same as the integration of x^2 between 0 and 2.

The underlying answer to your question is these two methods should not give the same answer - they are different operations.

Subject: Re: replace integration by summation
Posted by [Paul Van Delst\[1\]](#) on Wed, 19 Jun 2013 21:20:48 GMT
[View Forum Message](#) <> [Reply to Message](#)

D'oh... mistakenly hit reply instead of followup. Sorry. Stoopid tbird.

On 06/19/13 16:05, fd_luni@mail.com wrote:

>> But when you took Mats's suggestion and computed

>> INT_TABULATED(t,A1), was the single value zero or not?

>

> No it was not a single value zero.

>

> I had two function like this: For i=1,n-1 do begin A2=

> INT_TABULATED(t[0:i], A1[0:i]) B2= INT_TABULATED(t[0:i], B1[0:i])

> endfor

>

> When I replaced the INT_TABULATED by this: A2 =

> (t[1]-t[0])*total(A1,/cumulative) B2 =

> (t[1]-t[0])*total(B1,/cumulative)

>

> The function A2 = (t[1]-t[0])*total(A1,/cumulative) gives me

> completely different values from A2= INT_TABULATED(t[0:i], A1[0:i]).

That's fair enough. They should be different (how much depends on the data - see Phillip Bitzer's nice example.)

> But the function B2 = (t[1]-t[0])*total(B1,/cumulative gives me

> zeros.

Well, I would posit that either

a) $t[1] = t[0]$?

b) B1 is full of zeroes?

cheers,

paulv

Subject: Re: replace integration by summation

Posted by [Craig Markwardt](#) on Thu, 20 Jun 2013 04:06:05 GMT

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

On Wednesday, June 19, 2013 4:05:20 PM UTC-4, fd_...@mail.com wrote:

>> But when you took Mats's suggestion and computed INT_TABULATED(t,A1), was the single value zero or not?

>

>

>

> No it was not a single value zero.

>

> I had two function like this:

> For i=1,n-1 do begin

> A2= INT_TABULATED(t[0:i], A1[0:i])

> B2= INT_TABULATED(t[0:i], B1[0:i])

> endfor

Problem 1: A2 and B2 should be arrays.

>
> When I replaced the INT_TABULATED by this:
> A2 = (t[1]-t[0])*total(A1,/cumulative)
> B2 = (t[1]-t[0])*total(B1,/cumulative)
>
>
> The function A2 = (t[1]-t[0])*total(A1,/cumulative)
>
> gives me completely different values from A2= INT_TABULATED(t[0:i], A1[0:i]). But the function
B2 = (t[1]-t[0])*total(B1,/cumulative gives me zeros.

I asked this before: Is your T array regularly sampled or irregularly sampled? You are assuming that it is regularly sampled. If that assumption is wrong, you will get very different answers!

As Paul said, the only way for B2 to be all zeroes if t[1]-t[0] is zero or B1 is all zeroes to be begin with.

Craig

Subject: Re: replace integration by summation
Posted by [fd_luni](#) on Thu, 20 Jun 2013 10:12:42 GMT
[View Forum Message](#) <> [Reply to Message](#)

Oh, I see. Thanks Phillip, your explanation it was really good and clear.

> I asked this before: Is your T array regularly sampled or irregularly sampled?
No, is irregularly sampled, so that is why I got such different answers. Thanks Craig I didn't notice this.

There is any other way to avoid the loop in INT_TABULATED function? It seems that my idea to use cumulative it doesn't work.

With Thanks
M

Subject: Re: replace integration by summation
Posted by [fd_luni](#) on Thu, 20 Jun 2013 11:10:56 GMT
[View Forum Message](#) <> [Reply to Message](#)

Oh, I see. Thanks Phillip, your explanation it was really good and clear.

> I asked this before: Is your T array regularly sampled or irregularly sampled?
Yes, it is regularly sampled.

There is any other way to avoid the loop in INT_TABULATED function? It seems that my idea to use cumulative it doesn't work.

With Thanks
M

Subject: Re: replace integration by summation
Posted by [Paul Van Delst\[1\]](#) on Thu, 20 Jun 2013 14:11:46 GMT
[View Forum Message](#) <> [Reply to Message](#)

On 06/20/13 00:06, Craig Markwardt wrote:

> On Wednesday, June 19, 2013 4:05:20 PM UTC-4, fd_...@mail.com wrote:

>>> But when you took Mats's suggestion and computed

>>> INT_TABULATED(t,A1), was the single value zero or not?

>>

>>

>>

>> No it was not a single value zero.

>>

>> I had two function like this: For i=1,n-1 do begin A2=

>> INT_TABULATED(t[0:i], A1[0:i]) B2= INT_TABULATED(t[0:i], B1[0:i])

>> endfor

>

> Problem 1: A2 and B2 should be arrays.

>

>>

>> When I replaced the INT_TABULATED by this: A2 =

>> (t[1]-t[0])*total(A1,/cumulative) B2 =

>> (t[1]-t[0])*total(B1,/cumulative)

>>

>>

>> The function A2 = (t[1]-t[0])*total(A1,/cumulative)

>>

>> gives me completely different values from A2= INT_TABULATED(t[0:i],

>> A1[0:i]). But the function B2 = (t[1]-t[0])*total(B1,/cumulative

>> gives me zeros.

>

> I asked this before: Is your T array regularly sampled or irregularly

> sampled? You are assuming that it is regularly sampled. If that

> assumption is wrong, you will get very different answers!

>

> As Paul said, the only way for B2 to be all zeroes if t[1]-t[0] is

> zero or B1 is all zeroes to be begin with.

From an email reply Maria stated: "Yes, B1 is full of zeroes."

Problem solved (on our end at least :o)

cheers,

paulv

Subject: Re: replace integration by summation
Posted by [Phillip Bitzer](#) on Thu, 20 Jun 2013 16:28:53 GMT
[View Forum Message](#) <> [Reply to Message](#)

> There is any other way to avoid the loop in INT_TABULATED function? It seems that my idea to use cumulative it doesn't work.

Well, it didn't work the way you've implemented it. Take a look at those notes I posted. Your answer is there, assuming the trapezoidal (or rectangle) method works for your data. Look at eq (5.17) and the associated code below. Add a /CUMULATIVE to that command, and this should be what you're looking for.

Again, these methods of integration are not the same as used by INT_TABULATED and will not (in general) give you the same answers.
