Cuba.jl Documentation

Release 0.0.4

Mose' Giordano

April 12, 2016

Contents

1	Introduction			
2	Installation 3			
3	Usage 3.1 Mandatory Arguments 3.2 Optional Keywords 3.3 Output	5 5 6 10		
4	Vectorization 1			
5	Examples5.1Vector-valued integrand5.2Integral with non-constant boundaries5.3Complex integrand	13 13 14 14		
6	Performance 1			
7	Related projects	19		
8	License	21		
9	Credits	23		

Introduction

Cuba.jl is a Julia library for multidimensional numerical integration of real-valued functions of real arguments, using different algorithms.

This is just a Julia wrapper around the C Cuba library, version 4.2, by **Thomas Hahn**. All the credits goes to him for the underlying functions, blame me for any problem with the Julia interface. Feel free to report bugs and make suggestions at https://github.com/giordano/Cuba.jl/issues.

Algorithm	Basic integration method	Туре	Variance reduction
	Sobol quasi-random sample	Monte Carlo	
Vegas	or Mersenne Twister	Monte Carlo	importance sampling
	pseudo-random sample		
	or Ranlux pseudo-random	Monte Carlo	
	sample		
	Sobol quasi-random sample	Monte Carlo	globally adaptive subdivision
Suave	or Mersenne Twister	Monte Carlo	
	pseudo-random sample		and importance sampling
	or Ranlux pseudo-random	Monte Carlo	
	sample		
	Korobov quasi-random	Monte Carlo	stratified sampling
	sample		
Divonne	or Sobol quasi-random	Monte Carlo	aided by methods from
	sample		
	or Mersenne Twister	Monte Carlo	numerical optimization
	pseudo-random sample		1
	or Ranlux pseudo-random	Monte Carlo	
	sample		
	or cubature rules	deterministic	
Cuhre	cubature rules	deterministic	globally adaptive subdivi-
			sion

All algorithms provided by Cuba library are supported in Cuba.jl:

For more details on the algorithms see the manual included in Cuba library and available in deps/cuba-julia/cuba.pdf after successful installation of Cuba.jl.

Integration is performed on the *n*-dimensional unit hypercube $[0,1]^n$. If you want to compute an integral over a different set, you have to scale the integrand function in order to have an equivalent integral on $[0,1]^n$. For example, recall that in one dimension

$$\int_{a}^{b} \mathrm{d}x \, f[x] = \int_{0}^{1} \mathrm{d}y \, f[a + (b - a)y](b - a)$$

where the final (b - a) is the one-dimensional version of the Jacobian. This generalizes straightforwardly to more than one dimension. In *Examples* section you can find the computation of a 3-dimensional integral with non-constant boundaries using Cuba.jl.

Note: This package has been tested only on GNU/Linux and OS X systems. Trying to install on Windows will likely fail, please report if you manage to install Cuba.jl on this system.

Installation

Cuba.jl is available for Julia 0.4 and later versions, and can be installed with Julia built-in package manager. In a Julia session run the command

julia> Pkg.add("Cuba")

The build script will download Cuba Library source code and build the Cuba shared object. In order to accomplish this task a C compiler is needed.

You may need to update your package list with Pkg.update() in order to get the latest version of Cuba.jl.

Usage

After installing the package, run

using Cuba or put this command into your Julia script. Cuba.jl provides four functions to integrate, one for each algorithm:

Vegas (*integrand*, *ndim*, *ncomp*[, *keywords*...])

Suave (*integrand*, *ndim*, *ncomp*[, *keywords*...])

Divonne (*integrand*, *ndim*, *ncomp*[, *keywords*...])

Cuhre (*integrand*, *ndim*, *ncomp* , *keywords*...))

Large parts of the following sections are borrowed from Cuba manual. Refer to it for more information on the details.

3.1 Mandatory Arguments

Mandatory arguments of integrator functions are:

- integrand (type: Function): the name of the function to be integrated
- ndim (type: Integer): the number of dimensions of the integral
- ncomp (type: Integer): the number of components of the integrand

Function integrand must be of this type:

Note that xx and ff arguments are passed as pointers, so you have to translate them to Julia objects before actually performing calculations, and finally convert back f into the pointer ff. In section *Examples* you can find concrete example of Cuba.jl use.

Note that x and f are both arrays of Cdouble type (alias for Float64), so Cuba.jl can be used to integrate real-valued functions of real arguments. See how to work with complex quantities in the example *Complex integrand*.

The integrand receives nvec samples in x and is supposed to fill the array f with the corresponding integrand values. Note that nvec indicates the actual number of points passed to the integrand here and may be smaller than the nvec given to the integrator (see *Common Keywords* below).

Note: admittedly, this user interface is not REPL-friendly, help on improving it is welcome.

3.2 Optional Keywords

All other arguments required by Cuba integrator routines can be passed as optional keywords. Cuba.jl uses some reasonable default values in order to enable users to invoke integrator functions with a minimal set of arguments. Anyway, if you want to make sure future changes to some default values of keywords will not affect your current script, explicitly specify the value of the keywords.

3.2.1 Common Keywords

These are optional keywords common to all functions:

- userdata (type: Ptr{Void}, default: C_NULL): user data passed to the integrand
- nvec (type: Integer, default: 1): the maximum number of points to be given to the integrand routine in each invocation. Usually this is 1 but if the integrand can profit from e.g. Single Instruction Multiple Data (SIMD) vectorization, a larger value can be chosen. See *Vectorization* section.
- epsrel (type: Real, default: 1e-4), and epsabs (type: Real, default: 1e-12): the requested relative $(\varepsilon_{\text{rel}})$ and absolute $(\varepsilon_{\text{abs}})$ accuracies. The integrator tries to find an estimate \hat{I} for the integral I which for every component c fulfills $|\hat{I}_c I_c| \leq \max(\varepsilon_{\text{abs}}, \varepsilon_{\text{rel}}|I_c|)$.
- flags (type: Integer, default: 0): flags governing the integration:
 - Bits 0 and 1 are taken as the verbosity level, i.e. 0 to 3, unless the CUBAVERBOSE environment variable contains an even higher value (used for debugging).

Level 0 does not print any output, level 1 prints "reasonable" information on the progress of the integration, level 2 also echoes the input parameters, and level 3 further prints the subregion results (if applicable).

- Bit 2 = 0: all sets of samples collected on a subregion during the various iterations or phases contribute to the final result.

Bit 2 = 1, only the last (largest) set of samples is used in the final result.

- (Vegas and Suave only)

Bit 3 = 0, apply additional smoothing to the importance function, this moderately improves convergence for many integrands.

Bit 3 = 1, use the importance function without smoothing, this should be chosen if the integrand has sharp edges.

- Bit 4 = 0, delete the state file (if one is chosen) when the integration terminates successfully.

Bit 4 = 1, retain the state file.

- (Vegas only)

Bit 5 = 0, take the integrator's state from the state file, if one is present.

Bit 5 = 1, reset the integrator's state even if a state file is present, i.e. keep only the grid. Together with Bit 4 this allows a grid adapted by one integration to be used for another integrand.

- Bits 8–31 =: level determines the random-number generator.

To select e.g. last samples only and verbosity level 2, pass 6 = 4 + 2 for the flags.

• seed (type: Integer, default: 0): the seed for the pseudo-random-number generator. The random-number generator is chosen as follows:

seed	level (bits 8–31 of flags)	Generator
zero	N/A	Sobol (quasi-random)
non-zero	zero	Mersenne Twister (pseudo-random)
non-zero	non-zero	Ranlux (pseudo-random)

Ranlux implements Marsaglia and Zaman's 24-bit RCARRY algorithm with generation period p, i.e. for every 24 generated numbers used, another p - 24 are skipped. The luxury level is encoded in level as follows:

- Level 1 (p = 48): very long period, passes the gap test but fails spectral test.
- Level 2 (p = 97): passes all known tests, but theoretically still defective.
- Level 3 (p = 223): any theoretically possible correlations have very small chance of being observed.
- Level 4 (p = 389): highest possible luxury, all 24 bits chaotic.

Levels 5–23 default to 3, values above 24 directly specify the period p. Note that Ranlux's original level 0, (mis)used for selecting Mersenne Twister in Cuba, is equivalent to level = 24.

- mineval (type: Real, default: 0): the minimum number of integrand evaluations required
- maxeval (type: Real, default: 1000000): the (approximate) maximum number of integrand evaluations allowed
- statefile (type: AbstractString, default: ""): a filename for storing the internal state. To not store the internal state, put "" (empty string, this is the default) or C_NULL (C null pointer).

Cuba can store its entire internal state (i.e. all the information to resume an interrupted integration) in an external file. The state file is updated after every iteration. If, on a subsequent invocation, a Cuba routine finds a file of the specified name, it loads the internal state and continues from the point it left off. Needless to say, using an existing state file with a different integrand generally leads to wrong results.

This feature is useful mainly to define "check-points" in long-running integrations from which the calculation can be restarted.

Once the integration reaches the prescribed accuracy, the state file is removed, unless bit 4 of flags (see above) explicitly requests that it be kept.

• spin (type: Ptr{Void}, default: C_NULL): this is the placeholder for the "spinning cores" pointer. Cuba.jl does not support parallelization, so this keyword should not have a value different from C_NULL.

3.2.2 Vegas-Specific Keywords

These optional keywords can be passed only to Vegas ():

- nstart (type: Integer, default: 1000): the number of integrand evaluations per iteration to start with
- nincrease (type: Integer, default: 500): the increase in the number of integrand evaluations per iteration
- nbatch (type: Integer, default: 1000): the batch size for sampling

Vegas samples points not all at once, but in batches of size nbatch, to avoid excessive memory consumption. 1000 is a reasonable value, though it should not affect performance too much

• gridno (type: Integer, default: 0): the slot in the internal grid table.

It may accelerate convergence to keep the grid accumulated during one integration for the next one, if the integrands are reasonably similar to each other. Vegas maintains an internal table with space for ten grids for this purpose. The slot in this grid is specified by gridno.

If a grid number between 1 and 10 is selected, the grid is not discarded at the end of the integration, but stored in the respective slot of the table for a future invocation. The grid is only re-used if the dimension of the subsequent integration is the same as the one it originates from.

In repeated invocations it may become necessary to flush a slot in memory, in which case the negative of the grid number should be set.

3.2.3 Suave-Specific Keywords

These optional keywords can be passed only to Suave ():

- nnew (type: Integer, default: 1000): the number of new integrand evaluations in each subdivision
- nmin (type: Integer, default: 2): the minimum number of samples a former pass must contribute to a subregion to be considered in that region's compound integral value. Increasing nmin may reduce jumps in the χ^2 value
- flatness (type: Real, default: .25): the type of norm used to compute the fluctuation of a sample. This determines how prominently "outliers", i.e. individual samples with a large fluctuation, figure in the total fluctuation, which in turn determines how a region is split up. As suggested by its name, flatness should be chosen large for "flat" integrands and small for "volatile" integrands with high peaks. Note that since flatness appears in the exponent, one should not use too large values (say, no more than a few hundred) lest terms be truncated internally to prevent overflow.

3.2.4 Divonne-Specific Keywords

These optional keywords can be passed only to Divonne():

• key1 (type: Integer, default: 47): determines sampling in the partitioning phase: key1 = 7,9,11,13 selects the cubature rule of degree key1. Note that the degree-11 rule is available only in 3 dimensions, the degree-13 rule only in 2 dimensions.

For other values of key1, a quasi-random sample of $n_1 = |key1|$ points is used, where the sign of key1 determines the type of sample,

- key1 > 0, use a Korobov quasi-random sample,
- key1 < 0, use a "standard" sample (a Sobol quasi-random sample if seed = 0, otherwise a pseudo-random sample).
- key2 (type: Integer, default: 1): determines sampling in the final integration phase:

key2 = 7, 9, 11, 13 selects the cubature rule of degree key2. Note that the degree-11 rule is available only in 3 dimensions, the degree-13 rule only in 2 dimensions.

For other values of key2, a quasi-random sample is used, where the sign of key2 determines the type of sample,

- * key 2 > 0, use a Korobov quasi-random sample,
- * key2 < 0, use a "standard" sample (see description of key1 above),
- and $n_2 = |key2|$ determines the number of points,
 - * $n_2 \ge 40$, sample n_2 points,

- * $n_2 < 40$, sample $n_2 n_{need}$ points, where n_{need} is the number of points needed to reach the prescribed accuracy, as estimated by Divonne from the results of the partitioning phase
- key3 (type: Integer, default: 1): sets the strategy for the refinement phase:

key3 = 0, do not treat the subregion any further.

key3 = 1, split the subregion up once more.

Otherwise, the subregion is sampled a third time with key3 specifying the sampling parameters exactly as key2 above.

• maxpass (type: Integer, default: 5): controls the thoroughness of the partitioning phase: The partitioning phase terminates when the estimated total number of integrand evaluations (partitioning plus final integration) does not decrease for maxpass successive iterations.

A decrease in points generally indicates that Divonne discovered new structures of the integrand and was able to find a more effective partitioning. maxpass can be understood as the number of "safety" iterations that are performed before the partition is accepted as final and counting consequently restarts at zero whenever new structures are found.

- border (type: Real, default: 0.): the width of the border of the integration region. Points falling into this border region will not be sampled directly, but will be extrapolated from two samples from the interior. Use a non-zero border if the integrand function cannot produce values directly on the integration boundary
- maxchisq (type: Real, default: 10.): the χ^2 value a single subregion is allowed to have in the final integration phase. Regions which fail this χ^2 test and whose sample averages differ by more than mindeviation move on to the refinement phase.
- mindeviation (type: Real, default: 0.25): a bound, given as the fraction of the requested error of the entire integral, which determines whether it is worthwhile further examining a region that failed the χ^2 test. Only if the two sampling averages obtained for the region differ by more than this bound is the region further treated.
- ngiven (type: Integer, default: 0): the number of points in the xgiven array
- ldxgiven (type: Integer, default: 0): the leading dimension of xgiven, i.e. the offset between one point and the next in memory
- xgiven (type: AbstractArray{Real}, default: zeros(typeof(0.0), ldxgiven, ngiven)): a list of points where the integrand might have peaks. Divonne will consider these points when partitioning the integration region. The idea here is to help the integrator find the extrema of the integrand in the presence of very narrow peaks. Even if only the approximate location of such peaks is known, this can considerably speed up convergence.
- nextra (type: Integer, default: 0): the maximum number of extra points the peak-finder subroutine will return. If nextra is zero, peakfinder is not called and an arbitrary object may be passed in its place, e.g. just 0
- peakfinder (type: Ptr{Void}, default: C_NULL): the peak-finder subroutine

3.2.5 Cuhre-Specific Keyword

This optional keyword can be passed only to Cuhre ():

• key (type: Integer, default: 0): chooses the basic integration rule:

key = 7, 9, 11, 13 selects the cubature rule of degree key. Note that the degree-11 rule is available only in 3 dimensions, the degree-13 rule only in 2 dimensions.

For other values, the default rule is taken, which is the degree-13 rule in 2 dimensions, the degree-11 rule in 3 dimensions, and the degree-9 rule otherwise.

3.3 Output

The integrating functions Vegas (), Suave (), Divonne (), and Cuhre () return the 6-tuple

(integral, error, probability, neval, fail, nregions)

The first three terms of the tuple are arrays with length ncomp, the last three ones are scalars. In particular, if you assign the output of integrator functions to the variable named result, you can access the value of the i-th component of the integral with result[1][i] and the associated error with result[2][i].

- integral (type: Cdouble array with ncomp components): the integral of integrand over the unit hypercube
- error (type: Cdouble array with ncomp components): the presumed absolute error for each component of integral
- probability (type: Cdouble array with ncomp components): the χ^2 -probability (not the χ^2 -value itself!) that error is not a reliable estimate of the true integration error. To judge the reliability of the result expressed through prob, remember that it is the null hypothesis that is tested by the χ^2 test, which is that error *is* a reliable estimate. In statistics, the null hypothesis may be rejected only if prob is fairly close to unity, say prob > .95
- neval (type: Cint): the actual number of integrand evaluations needed
- fail (type: Cint): an error flag:
 - fail = 0, the desired accuracy was reached
 - fail = -1, dimension out of range
 - fail > 0, the accuracy goal was not met within the allowed maximum number of integrand evaluations.
 While Vegas, Suave, and Cubre simply return 1, Divonne can estimate the number of points by which maxeval needs to be increased to reach the desired accuracy and returns this value.
- nregions (type: Cint): the actual number of subregions needed (always 0 in Vegas)

Vectorization

Vectorization means evaluating the integrand function for several points at once. This is also known as Single Instruction Multiple Data (SIMD) paradigm and is different from ordinary parallelization where independent threads are executed concurrently. It is usually possible to employ vectorization on top of parallelization.

Cuba.jl cannot automatically vectorize the integrand function, of course, but it does pass (up to) nvec points per integrand call (*Common Keywords*). This value need not correspond to the hardware vector length – computing several points in one call can also make sense e.g. if the computations have significant intermediate results in common.

A note for disambiguation: The nbatch argument of Vegas is related in purpose but not identical to nvec. It internally partitions the sampling done by Vegas but has no bearing on the number of points given to the integrand. On the other hand, it it pointless to choose nvec > nbatch for Vegas.

Examples

5.1 Vector-valued integrand

Consider the integral

$$\int_{\Omega} \boldsymbol{f}(x,y,z) \, \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}z$$

where $\Omega = [0, 1]^3$ and

using Cuba

$$\boldsymbol{f}(x, y, z) = \left(\sin(x)\cos(y)\exp(z), \exp(-(x^2 + y^2 + z^2)), \frac{1}{1 - xyz}\right)$$

This is the Julia script you can use to compute the above integral

```
function integrand(ndim::Cint, xx::Ptr{Cdouble}, ncomp::Cint,
                   ff::Ptr{Cdouble}, userdata::Ptr{Void})
   x = pointer_to_array(xx, (ndim,))
   f = pointer_to_array(ff, (ncomp,))
   f[1] = sin(x[1]) * cos(x[2]) * exp(x[3])
   f[2] = \exp(-(x[1]^{2} + x[2]^{2} + x[3]^{2}))
   f[3] = 1/(1 - x[1] * x[2] * x[3])
   ff = pointer_from_objref(f)
    return Cint(0)::Cint
end
result = Cuhre(integrand, 3, 3, epsabs=1e-12, epsrel=1e-10)
println("Results of Cuba:")
for i=1:3; println("Component $i: ", result[1][i], " ± ", result[2][i]); end
println("Exact results:")
println("Component 1: ", (e-1)*(1-cos(1))*sin(1))
println("Component 2: ", (sqrt(pi)*erf(1)/2)^3)
println("Component 3: ", zeta(3))
```

This is the output

```
Results of Cuba:
Component 1: 0.6646696797813739 \pm 1.0050367631018485e-13
Component 2: 0.4165383858806454 \pm 2.932866749838454e-11
Component 3: 1.2020569031649702 \pm 1.1958522385908214e-10
Exact results:
```

```
Component 1: 0.6646696797813771
Component 2: 0.41653838588663805
Component 3: 1.2020569031595951
```

5.2 Integral with non-constant boundaries

Now consider the integral

$$\int_{-y}^{y} \int_{0}^{z} \int_{0}^{\pi} \cos(x) \sin(y) \exp(z) \,\mathrm{d}x \,\mathrm{d}y \,\mathrm{d}z$$

By applying the substitution rule repeatedly, you can scale the integrand function and get this equivalent integral over the domain $\Omega = [0, 1]^3$

$$\int_{\Omega} 2\pi^3 y z^2 \cos(\pi y z (2x-1)) \sin(\pi y z) \exp(\pi z) \, \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}z$$

that can be computed with Cuba.jl using the following Julia script

This is the output

```
Result of Cuba: 54.98607586826157 ± 5.460606521639899e-9
Exact result: 54.98607586789537
```

5.3 Complex integrand

As already explained, Cuba.jl operates on real quantities, so if you want to integrate a complex-valued function of complex arguments you have to treat complex quantities as 2-component arrays or real numbers. For example, the integral

$$\int_0^1 \exp(\mathrm{i}x) \,\mathrm{d}x$$

can be computed with the following Julia script

```
using Cuba
function integrand(ndim::Cint, xx::Ptr{Cdouble}, ncomp::Cint,
                  ff::Ptr{Cdouble}, userdata::Ptr{Void})
   x = pointer_to_array(xx, (ndim,))
   f = pointer_to_array(ff, (ncomp,))
    # Complex integrand
   tmp = exp(im * x[1])
    # Assign to two components of "f" the real
    # and imaginary part of the integrand.
   f[1] = real(tmp)
   f[2] = imag(tmp)
   ff = pointer_from_objref(f)
   return Cint(0)::Cint
end
result = Cuhre(integrand, 1, 2)
println("Result of Cuba: ", result[1][1] + im*result[1][2])
println("Exact result: ", sin(1) + im*(1 - cos(1)))
```

This is the output

Result of Cuba: 0.8414709848078966 + 0.45969769413186035im Exact result: 0.8414709848078965 + 0.45969769413186023im

Performance

Cuba.jl cannot (yet?) take advantage of parallelization capabilities of Cuba Library. Nonetheless, it has performances comparable with (if not slightly better than) equivalent native C or Fortran codes based on Cuba library when CUBACORES environment variable is set to 0 (i.e., multithreading is disabled). The following is the result of running the benchmark present in test directory on a 64-bit GNU/Linux system running Julia 0.4. The C and FORTRAN 77 benchmark codes have been built with GCC 5.3.1.

```
$ CUBACORES=0 julia -e 'cd(Pkg.dir("Cuba")); include("test/benchmark.jl")'
INFO: Performance of Cuba.jl:
  0.332981 seconds (Vegas)
 0.656121 seconds (Suave)
  0.385009 seconds (Divonne)
  0.299737 seconds (Cuhre)
INFO: Performance of Cuba Library in C:
  0.348074 seconds (Vegas)
  0.662016 seconds (Suave)
 0.378092 seconds (Divonne)
 0.303750 seconds (Cuhre)
INFO: Performance of Cuba Library in Fortran:
  0.328000 seconds (Vegas)
  0.660000 seconds (Suave)
  0.364000 seconds (Divonne)
  0.292000 seconds (Cuhre)
```

Of course, native C and Fortran codes making use of Cuba Library outperform Cuba.jl when higher values of CUBACORES are used, for example:

```
$ CUBACORES=1 julia -e 'cd(Pkg.dir("Cuba")); include("test/benchmark.jl")'
INFO: Performance of Cuba.jl:
 0.336093 seconds (Vegas)
  0.663560 seconds (Suave)
  0.391726 seconds (Divonne)
  0.301402 seconds (Cuhre)
INFO: Performance of Cuba Library in C:
 0.122129 seconds (Vegas)
  0.609809 seconds (Suave)
 0.151087 seconds (Divonne)
 0.085541 seconds (Cuhre)
INFO: Performance of Cuba Library in Fortran:
  0.072000 seconds (Vegas)
  0.628000 seconds (Suave)
  0.164000 seconds (Divonne)
  0.096000 seconds (Cuhre)
```

Cuba.jl internally fixes CUBACORES to 0 in order to prevent from forking julia processes that would only slow down calculations eating up the memory, without actually taking advantage of concurrency. Furthemore, without this measure, adding more Julia processes with addprocs () would only make the program segfault.

Related projects

Another Julia package for multidimenensional numerical integration is available: Cubature.jl, by Steven G. Johnson. Differently from Cuba.jl, this works on GNU/Linux, OS X and Windows as well.

License

The Cuba.jl package is licensed under the GNU Lesser General Public License, the same as Cuba library. The original author is Mosè Giordano.

Credits

If you use this library for your work, please credit Thomas Hahn. Citable papers about Cuba Library:

- Hahn, T. 2005, Computer Physics Communications, 168, 78. DOI:10.1016/j.cpc.2005.01.010. arXiv:hep-ph/0404043. Bibcode:2005CoPhC.168...78H.
- Hahn, T. 2015, Journal of Physics Conference Series, 608, 012066. DOI:10.1088/1742-6596/608/1/012066. arXiv:1408.6373. Bibcode:2015JPhCS.608a2066H.

Index

С

Cuhre() (built-in function), 5

D

Divonne() (built-in function), 5

S

Suave() (built-in function), 5

V

Vegas() (built-in function), 5