ASTR4004/ASTR8004
Astronomical Computing
Lecture 08

Christoph Federrath

19 September 2017

Data processing in IDL

1 Reading data in IDL and making map plots

1. Download 3 HDF5 data files of the world's largest simulation of supersonic turbulence:
   EXTREME_proj_xy_000100
   EXTREME_proj_xy_000200
   EXTREME_proj_xy_000300
   These data are column density (2D) projections of the 3D gas density in the simulations. The 3 files are at 3 different times of the evolution. If you are interested in more details on these data and the simulations, please have a look at Federrath et al. (2016) and references therein.

2. Note that HDF5 is version 5 of the Hierarchical Data Format, which is a powerful data format for scientific data. Many codes use HDF5 as the standard now (it also allows parallel I/O). If you have HDF5 installed on your computer (e.g., for Mac users via `sudo port install hdf5`), you can check the content of HDF5 data files simply by typing `> h5ls [file]`. This and `> h5dump -H [file]` prints useful information about the content and header information. Similar to FITS files, HDF5 is a complex file format that can contain multiple datasets of different type, all with header information.

3. Use the IDL program readhdf5data.pro to read in one of the simulation data files:
   `IDL> readhdf5data, filename, 'dens_proj_xy', coldens`
   Now the content of the dataset '`dens_proj_xy`' is read into the array called '`coldens`'.

4. Use `help`, `reform`, `alog10`, `min`, `max`, `window`, `tvscl`, `bytscl`, and `xloadct` to investigate the data and make images of the gas column density. Display the image on the screen and modify the appearance by loading different colour tables and changing the minimum and maximum of the plotted colour range.

5. However, we also want to add a colour bar and write the image out as a `.eps` file. This can be done with the customised procedure write_eps_wcb.pro provided (note that this is a new version compared to lecture 7, which does the `bytscl` automatically). Use this to make images of the three data files, scaled nicely and with a reasonable colour table and colour bar scaling.

6. Code this all up such that you use procedures and/or functions in order to be able to call these for the different files.

2 Re-binning and beam convolution

1. Lets add some beam smearing in order to simulate the effect of looking at these data through a telescope with a finite beam resolution.

2. First we can use the simple IDL functions `rebin` and `congrid`, which can be used to make lower-resolution versions of the previous images. The `congrid` function works for arbitrary grid interpolations.

3. However, both `rebin` and `congrid` do not simulate Gaussian beam smoothing. Lets implement a Gaussian beam convolution based on the IDL function `gauss_smooth`. The function (or procedure) that does the convolution should take the un-convolved map as an input, as well as the Gaussian beam full-width-half-maximum (FWHM) as an input, and return the beam-smeared image. Use 10 pixels for the FWHM.

4. Write out the beam-smeared image(s) with `write_eps_wcb.pro` as before and compare the non-smoothed with the beam-smoothed images.

3 Statistical functions and PDFs

1. Compute the mean, standard deviation, skewness and kurtosis of the log-normalised column density (intensity),
$$\mathcal{I} = \ln\left(I/\langle I\rangle\right), \tag{1}$$
where $\langle I\rangle$ is the mean column density (intensity) and ln() is the natural logarithm (`alog` in IDL). Use the built-in functions `mean, stddev, skewness, kurtosis, moment`.

2. Write a procedure that writes the PDF locations, the histogram, the PDF (normalised histogram), and the CMF of $\mathcal{I}$ to an ASCII text file.

3. Write a procedure to read these ASCII files and data columns back into memory.

4. Write a procedure to compute the mean, standard deviation, skewness, and kurtosis of the PDFs by summation over the PDF. Compare the statistical moments of the PDF based on the data and intrinsic functions with those based on the summation over the PDF. Is there a difference? Why? How can this be fixed?

4 Averaging data, making plots with error bars, and fitting

1. Average all the three column density PDFs over time to produce a time-averaged PDF with error bars.

2. Produce a plot of the PDF including error bars.

3. Use the MPFIT library to write a procedure to fit the time-averaged PDF with a Gaussian function. Over-plot the fit.

Make use of `stop` in your scripts to turn from automatic script mode to interactive mode. Use `IDL> .cont` to continue the script after `stop`. Search function and procedure help to learn about optional keywords, which control the behaviour of functions/procedures.