Technical requirements and an introduction to '.ipynb' notebooks

June 22, 2023

1 BSE Data Science Summer School 2023

2 Coding Bootcamp in Python and R

2.1 Technical requirements and an introduction to `.ipynb` notebooks

Paul Rognon

2.1.1 Technical requirements

The course material will be notebooks (`.ipynb` files).

To open and run an existing `.ipynb` notebooks, **two possible options** are:

- **Google Colab**: the simplest and **preferred option for our course**. Notebooks are run in the cloud. **You need a Google account** (e.g. a BSE account). To start Google Colab, click here. To open an `.ipynb` file with Colab you need to upload it to Colab: click on 'Upload' on the top right corner of the opening window. If you already have a notebook open in Colab you can open an existing `.ipynb` notebook clicking on File > Upload notebook.

OR

- **Jupyter Notebook application**: the notebooks are run on your personal computer. **If you go with that option, you will need to install Python, R and Jupyter Notebook on your personal computer**. You can have the three of them by installing the Anaconda distribution. Alternatively, you can install Python, R and Jupyter Notebook separately. To open an existing `.ipynb` notebook: launch Jupyter Notebook (opens a tab in your web browser), and, on the welcome screen, navigate to your `.ipynb` file and click on it. Alternatively, use the Upload button in the top right corner of the welcome screen.

2.1.2 Notebooks

Shortly, notebooks are documents that combine live code, equations, narrative text and visualizations. In a single document, you can write code, show the code output, write markdown type content, show text and visual outputs. Jupyter/`.ipynb` notebooks are a type of notebooks originally developped for Python and have extension `.ipynb`. The document you are seeing is an example of `.ipynb` notebook converted to a PDF document.

`.ipynb` notebooks are composed of cells that can be text or code:

- in Google Colab, to add a code cell click on `+ Code`, to add a text cell click on `+ Text`.

- in the Jupyter Notebook application, to add a cell which is by default a code cell, click on `+`.

There are two modes to work with cells. An edit mode and a command mode:

- Edit mode (indicated by a green border around the cell in Jupyter Notebook application) can be accessed by double clicking on a cell. It is the default mode when a new cell is created.
- The command mode (indicated by a blue border around the cell in Jupyter Notebook application) can be accessed by pressing `CRTL+m` when you are in edit mode (`Esc` works too in the Jupyter Notebook application). The command mode lets use you shorcuts: press `m` to convert a cell to text type, `b` to create a new cell below for example. To delete a cell, used shorcut `d` in command mode in Google Colab, `dd` in command mode in Jupyter Notebook application.

To execute the code in a code cell: press `CTRL+ENTER` when in edit mode. If your code entails the printing of an output it will show under the cell.

To print the text in a text cell: press `CTRL+ENTER` when in edit mode in Jupyter Notebook application. In Colab, this is no required, just exit the edit mode of the cell and the text will be printed.

In a code cell, everything you write will be interpreted as a Python command, except whatever is after `#` on a line. Anything on a line after `#` will not be executed. `#` is used to introduce comments on your code. It is good pratice to comment your code so it id more readable. A code cell is followed by an out cell that shows the output of the last line of the code in the cell.

In a text cell you can include markdown content: text, hyperlinks, equations…You can for example:
- make a list - add a link: link to google.com
- add formulas: $\hat{\beta} = (X^T X)^{-1} X^T y$

In Colab, to create a new notebook from scracth, two possible options are: - launch Colab, the landing page is a new notebook by default, click cancel on opening window - if you have a notebook already open, do File > New notebook.

In Jupyter Notebook application, to create a new notebook from scracth, two possible options are: - launch Jupyter Notebook app, click on the New button in the top right corner of the welcome screen and select Python 3. - If you already have a notebook open, you can create a new notebook from scracth by doing File > New notebook > Python 3.

2.1.3 Google Colab

In class, we will be working in Google Colab which is a powerful online environment to create, edit and run `.ipynb` notebooks. It is constantly evolving and new features are added. Working with Google Colab is very similar to working with the Jupyter Notebook application. Many of the features for the edition of code and display of output in Google Colab are available on the Jupyter Notebook application too.

Here is an introductory video about Google Colab.

A major advantage of Google Colab is that it executes your code on Google's cloud servers, you leverage Google's hardware (including GPUs and TPUs) instead of your machine.

A big difference with Jupyther Notebook application is that since Colab runs in the cloud, you cannot simply read a file located on your computer by inputing a path to the location on your computer. You need to first upload the file to your Colab environment or to you Google Drive and mount the Drive. We will talk more about it later.

Another consequence of working in the cloud is that the changes you make to a notebook will be saved on the cloud. If you want to save your edited notebook on your computer you need to download a copy of it: File > Download.

Note Bene: While initially developped for Python programming, `.ipynb` notebooks, Google Colab and the Jupyter Notebook application now all support R programming. We will used `.ipynb` notebooks and Google Colab for the `R` part of the course too.