

中学 编程环境配置 Anaconda——Spyder 配置

北京市十一学校 郑子杰、孙工博

2023 年 7 月

本文档是 **Windows 操作系统** 下基于 numpy、matplotlib、scikit-learn、tensorflow、pytorch 等数据科学与机器学习算法库，引导 **中学生** 在自己的电脑上配置入门级别的数据科学与机器学习所需要的 **python 编程环境**。本配置说明文档基于 Anaconda 的库管理程序和 Spyder 编译器（Anaconda 的内置 python 编译器）进行环境配置。分为以下两步：

第 1 步. 安装 Anaconda

第 2 步. 安装数据分析和机器学习所需要的相关库

1. 安装 Anaconda

Anaconda 是一个基于科学计算的 python 库管理程序，在 cmd 中使用 conda 命令替换原有的 pip 命令进行安装，方便进行 python 的库管理。也就是说 Anaconda 一旦安装，就已经帮大家安装好了常用的科学计算库。

1.1 Anaconda 的下载与安装

Anaconda 清华镜像：<https://mirrors.tuna.tsinghua.edu.cn/anaconda/archive/>

如果清华镜像较慢，也可以使用阿里镜像（百度一下）

下载后进行安装，**下载时请务必看最右侧的版本号和更新时间，请下载和安装最新版本！**

为了避免出现问题，请务必安装在系统盘(c 盘)！

不要一直点击下一步，在过程中用户选择“all users”，不要只选择当前用户；

中间需要出现选项框的地方，全选！

1.2 Anaconda 的环境变量配置

安装完成后，需要在**计算机-属性-高级-环境变量** **PATH 环境变量**中新增环境变量

C:\ProgramData\Anaconda3

C:\ProgramData\Anaconda3\Scripts

C:\ProgramData\Anaconda3\Library\bin

C:\ProgramData\Anaconda3\Library\mingw-w64\bin

C:\ProgramData\Anaconda3\Library\usr\bin

- ① 如果电脑之前安装了 python，如果想要保证 Anaconda 内置的 Python 为默认编译器，则需要删除 PATH 中之前的 python 的环境变量；
- ② 如果不想修改默认编译器（例如之前已经安装了 pycharm 等），则不用修改。当然，最安全的方法是单独新开一个虚拟环境，但对于新手非常不友好，对新手而言不推荐。

1.3 更新 Anaconda 版本和 Anaconda 库

由于通过镜像地址下载的 Anaconda 版本和库不一定是最新的，所以 conda 命令有时候会报错，故需要先对 Anaconda 进行更新，具体步骤如下：

不要直接打开 anaconda，而是打开**我的电脑->windows->system32 文件夹**，找到 cmd.exe 程序，**右键管理员运行**。在 cmd 中进行版本更新和库的安装。即在 cmd 中完成以

下①②③三个步骤。

- ① 更换 Anaconda 源，使用阿里云的镜像地址（原地址在墙外，基本下载不下来）

```
conda config --add channels https://mirrors.aliyun.com/anaconda/pkg/free/  
conda config --add channels https://mirrors.aliyun.com/anaconda/pkg/main/
```

参考链接：<https://www.cnblogs.com/lvsling/p/8672404.html>

- ② 更新 Anaconda 自身版本

```
conda update anaconda
```

（时间比较长，需要等待）

备注：使用 Anaconda 时，conda 命令等价于 pip 命令

- ③ 更新 Anaconda 的库

```
conda upgrade --all
```

（两个横线）

2. 安装机器学习相关库

使用 conda 命令安装数据分析和机器学习相关的算法库。

2.1 机器学习算法库

在教学过程中常用的机器学习算法库包括 scikit-learn 库、tensorflow 库和 pytorch 库。

Scikit-Learn 库的优势集成了绝大部分机器学习算法（而不仅仅局限于神经网络），一般情况下两个库配合可以完美的实现机器学习各种算法的组合。其语法格式统一，简单易学。

Tensorflow 的优势是基于特别简单的程序语言就可以完成深度神经网络的构建。适合初学者使用；但是弊端是灵活性差，在科研和业界使用较少。

Pytorch 库是另一个深度神经网络库，其在现今的学术界和业界广泛使用，非常灵活，支持 cuda (GPU、CPU) 混合编程。但是缺点也很明显，对于中学生来讲，pytorch 比 tensorflow 复杂的多，一般中学生难以驾驭如此体量的代码。

同上，需要在 cmd 中安装以上 3 个库。打开我的电脑->windows->system32 文件夹，找到 cmd.exe 程序，右键管理员运行。

- ① 安装 Scikit-Learn:

```
conda install -c anaconda scikit-learn
```

会自动检测与 python 版本对应的 scikit-learn 的库版本。

- ② 安装 tensorflow:

```
conda install tensorflow
```

安装的是 Tensorflow2 版本。

- ③ 安装 pytorch:

对于初学者来说，只需要安装 CPU 版本:

```
conda install pytorch torchvision torchaudio cpuonly -c pytorch
```

如果有 GPU（显卡，也可以尝试安装 GPU 版本，官网上都有具体命令）

2.2 数据处理库

Anaconda 自动集成了常见的数据处理库: 例如 numpy (矩阵表示和运算工具) 和 pandas (与 excel 软件匹配的数据处理工具)。由于我们之前已经更新了所有 Anaconda 自带的库，所以只需要补充安装 Anaconda 没有自动安装的库:

- ① 基本图像处理 (可选)

```
conda install pillow
```

这里特别强调: 在 python 3.x 最常用的图像处理库 pillow; 与之对应的是 python2.x 版本中的 PIL 库。

② 计算机视觉（可选）

openCV 是现在世界上最流行的图像处理和计算机视觉（Computer Vision 库），最开始的 openCV 库仅仅支持 C 语言。后来 openCV 也开发了支持 python 的版本。其比 pillow 要强大很多。但是 openCV 的安装和配置也是最棘手且最容易踩雷的地方。

A. 首先不支持 conda 安装命令，所以请在虚拟环境中使用 pip 命令安装

```
pip install opencv-python
```

如果你发现可以安装，并且在 python 中能够正常导入 import cv2，那么恭喜你，运气真好！

B. 如果非要使用 conda 安装命令，则需要使用指定 menpo 版本的 opencv

```
conda install -c menpo opencv3
```

如果你发现可以安装，并且在 python 中能够正常 import cv2，那么恭喜你，运气还是真好！因为一般情况下会提示 menpo opencv 不支持 python3.6 和 python3.7

C. 最后实在不行只能去手动安装

这个步骤太长了，请参考 <https://blog.csdn.net/mawonly/article/details/87856530>

③ 数据分析可视化（必须）

虽然 pillow 和 openCV 中均有图像输出的 API（接口、类、函数），但是对于科学研究而言，大部分人习惯使用 MATLAB 软件在平面直角坐标系下进行数据的可视化，例如在平面直角坐标系下画函数图像等，所以需要安装 python 中与之对应的库 matplotlib。

```
conda install matplotlib
```

备注：所有想要安装的库都可以采用以上类似的方式进行安装和更新。

最后，在桌面上/开始菜单里启动 Anaconda 自带的编译器 Spyder 就可以进行编程了！



图标如下：

可以尝试运行以下代码看是否安装成功：

```
print("hello word")
```