

## 为什么使用anaconda以及关于anaconda安装的一些建议

事实上Anaconda 和 Jupyter notebook已成为数据分析的标准环境。

简单来说，Anaconda是包管理器和环境管理器，Jupyter notebook 可以将数据分析的代码、图像和文档全部组合到一个web文档中。

接下来我详细介绍下Anaconda，并在最后给出Jupyter notebook：

### 1.Anaconda是什么？

### 2.如何安装？

### 3. 如何管理包？

### 4.如何管理环境？

### 5.Jupyter notebook如何快速上手？

不过在开始前我需要强调下，下面的步骤你要亲自跟着敲一遍并在自己的电脑上实践。虽然下面你会遇到很多命令，给了谁都记不住的。但是别怕，也别中途放弃，因为你没必要记住命令，因为当你在后面学习数据分析用的多了，自然就记住了，记不住也没关系，学会在哪查找就可以了。你只需要跟着上面步骤操作下，并理解了每一步是干什么的就可以了。后面遇到要做的事情，忘记了回头查这个文档就可以了。

### 1.Anaconda是什么？

Anaconda在英文中是“蟒蛇”的意思

所有你看下面Anaconda的图标就像一个首尾互相咬住的“蟒蛇”。



你可能已经安装了 Python，那么为什么还需要 Anaconda？有以下3个原因：

1) Anaconda 附带了一大批常用数据科学包，它附带了 conda、Python 和 150 多个科学包及其依赖项。因此你可以立即开始处理数据。

2) 管理包

Anaconda 是在 conda（一个包管理器和环境管理器）上发展出来的。

在数据分析中，你会用到很多第三方的包，而conda（包管理器）可以很好的帮助你在计算机上安装和管理这些包，包括安装、卸载和更新包。

### 3) 管理环境

为什么需要管理环境呢？

比如你在A项目中用了 Python 2，而新的项目B老大要求使用Python 3，而同时安装两个Python版本可能会造成许多混乱和错误。这时候 conda就可以帮助你为不同的项目建立不同的运行环境。

还有很多项目使用的包版本不同，比如不同的pandas版本，不可能同时安装两个 Numpy 版本，你要做的应该是，为每个 Numpy 版本创建一个环境，然后项目的对应环境中工作。这时候conda就可以帮你做到。

## 2. 如何安装Anaconda?

Anaconda 可用于多个平台（Windows、Mac OS X 和 Linux）。你可以在下面地址上找到安装程序和安装说明。根据你的操作系统是32位还是64位选择对应的版本下载。

（Anaconda已经不支持Windows XP；同时查看自己电脑是32位还是64位，不要装错了。）

官网地址：<https://www.continuum.io/downloads>\*\*

（如果官网地址网速太慢无法下载。可以从国内清华大学开源软件镜像站进行下载并配置镜像。

直接在下面地址页面拉到最下面找最新版的Anaconda2或Anaconda3)

---

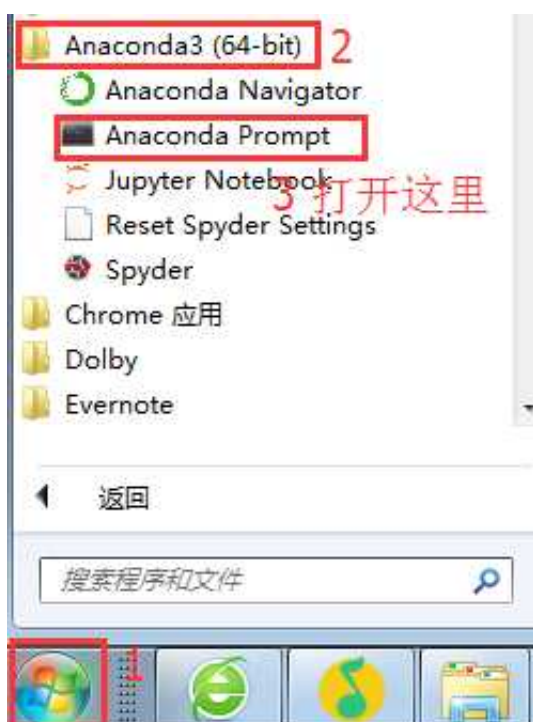
Anaconda 的下载文件比较大（约 500 MB），因为它附带了 Python 中最常用的数据科学包。

如果计算机上已经安装了 Python，安装不会对你有任何影响。实际上，脚本和程序使用的默认 Python 是 Anaconda 附带的 Python。

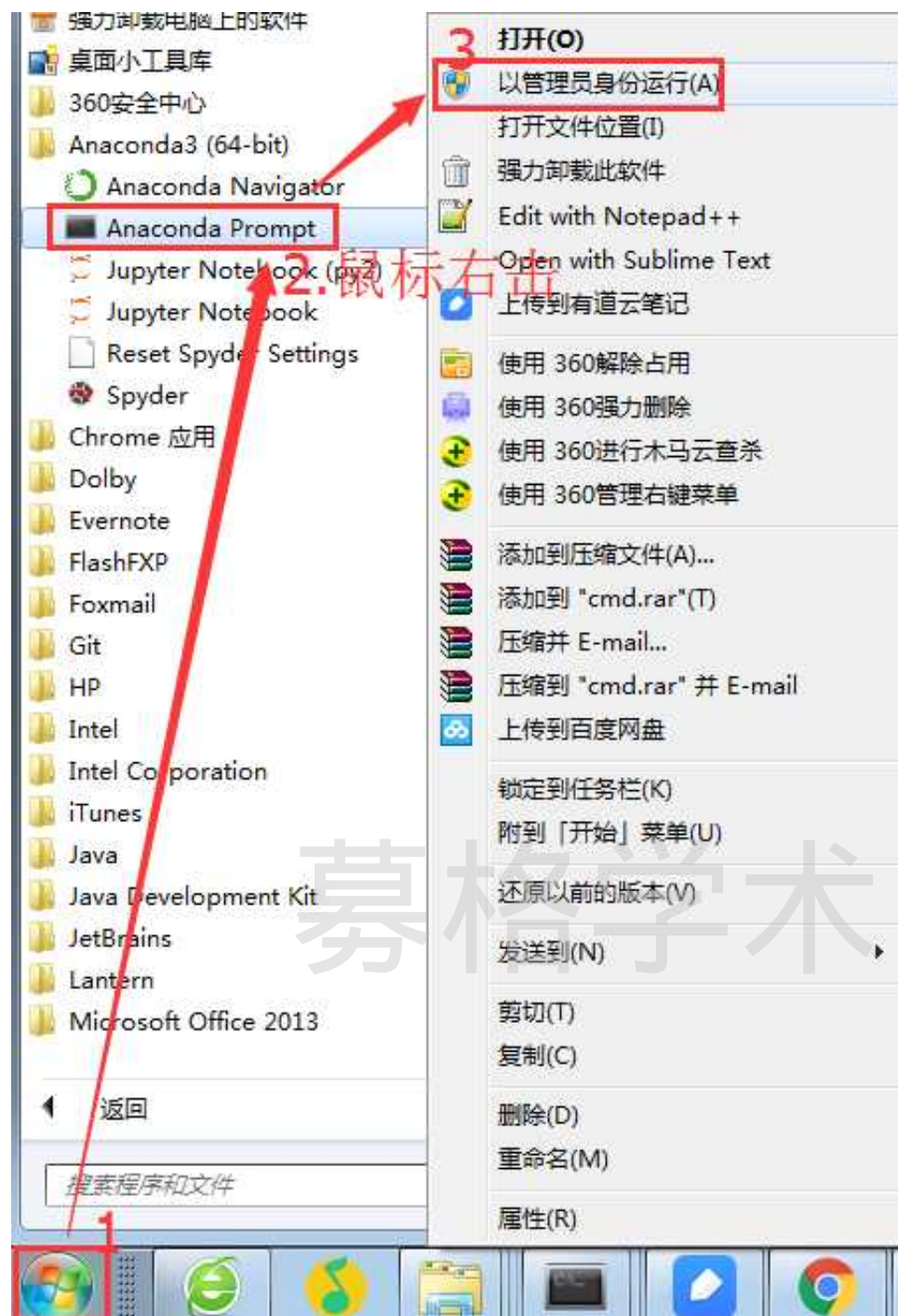
注意：如果你是windows 10系统，注意在安装Anaconda软件的时候，右击安装软件→选择以管理员的身份运行。

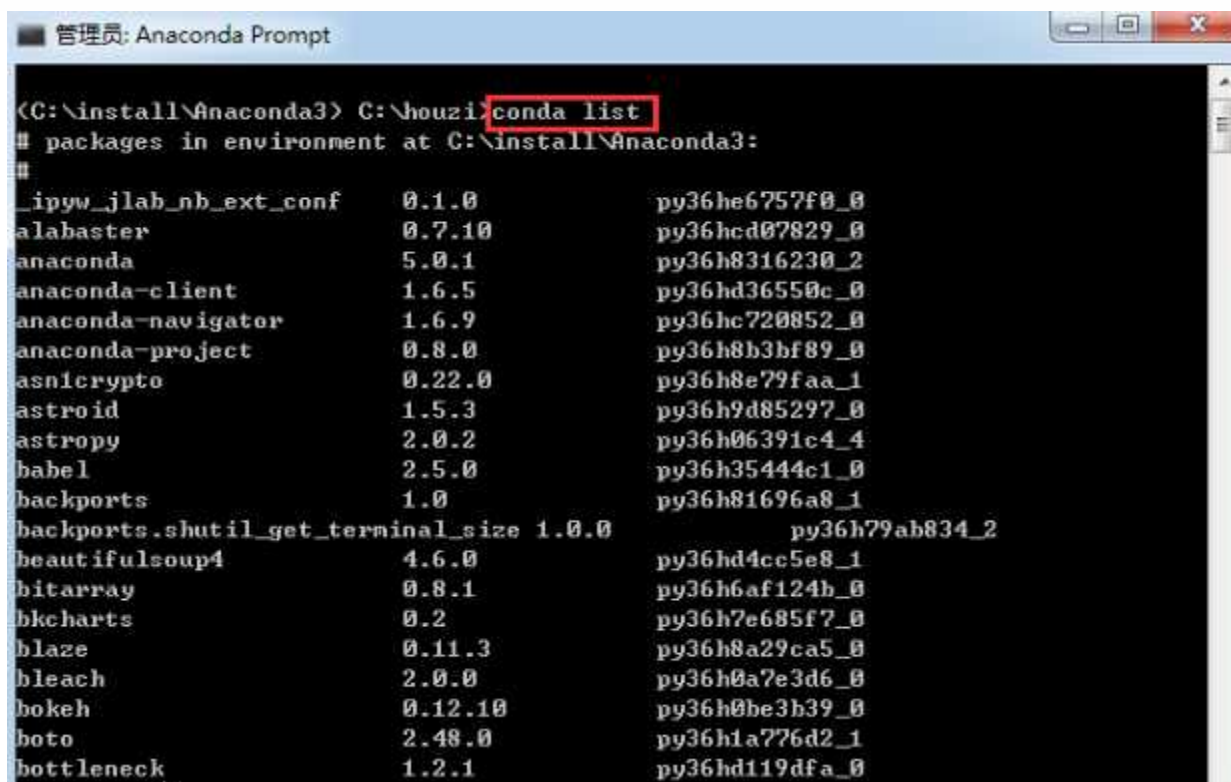


完成安装后，如果你是在windows上操作，按下面图打开 Anaconda Prompt （或者 Mac 下的终端），后面我会将**Anaconda Prompt**统一称为“终端”。



注意：如果你是windows 10系统，按下图操作





```
管理员: Anaconda Prompt
(C:\install\Anaconda3) C:\houzi>conda list
# packages in environment at C:\install\Anaconda3:
#
_ipyw_jlab_nb_ext_conf    0.1.0                py36he6757f0_0
alabaster                 0.7.10              py36hcd07829_0
anaconda                  5.0.1                py36h8316230_2
anaconda-client           1.6.5                py36hd36550c_0
anaconda-navigator        1.6.9                py36hc720852_0
anaconda-project          0.8.0                py36h8b3bf89_0
asn1crypto                0.22.0              py36h8e79faa_1
astroid                   1.5.3                py36h9d85297_0
astropy                   2.0.2                py36h06391c4_4
babel                     2.5.0                py36h35444c1_0
backports                 1.0                  py36h81696a8_1
backports.shutil_get_terminal_size 1.0.0                py36h79ab834_2
beautifulsoup4            4.6.0                py36hd4cc5e8_1
bitarray                  0.8.1                py36h6af124b_0
bkcharts                  0.2                  py36h7e685f7_0
blaze                     0.11.3               py36h8a29ca5_0
bleach                    2.0.0                py36h0a7e3d6_0
bokeh                     0.12.10              py36h0be3b39_0
boto                       2.48.0               py36h1a776d2_1
bottleneck                1.2.1                py36hd119dfa_0
```

如果按上面操作后在Anaconda Prompt中都无法使用conda命令，按以下顺序的解决办法来尝试：

1) 检查你是否原来安装过Python，如果安装过请彻底删除Python（同时要删除环境变量）后重装Anaconda

2) 检查自己是否将conda命令添加到了环境变量，操作方法如下：

<https://stackoverflow.com/questions/28612500/why-anaconda-does-not-recognize-conda-command>

如果Anaconda Prompt中可以使用conda命令，接着下面继续操作。

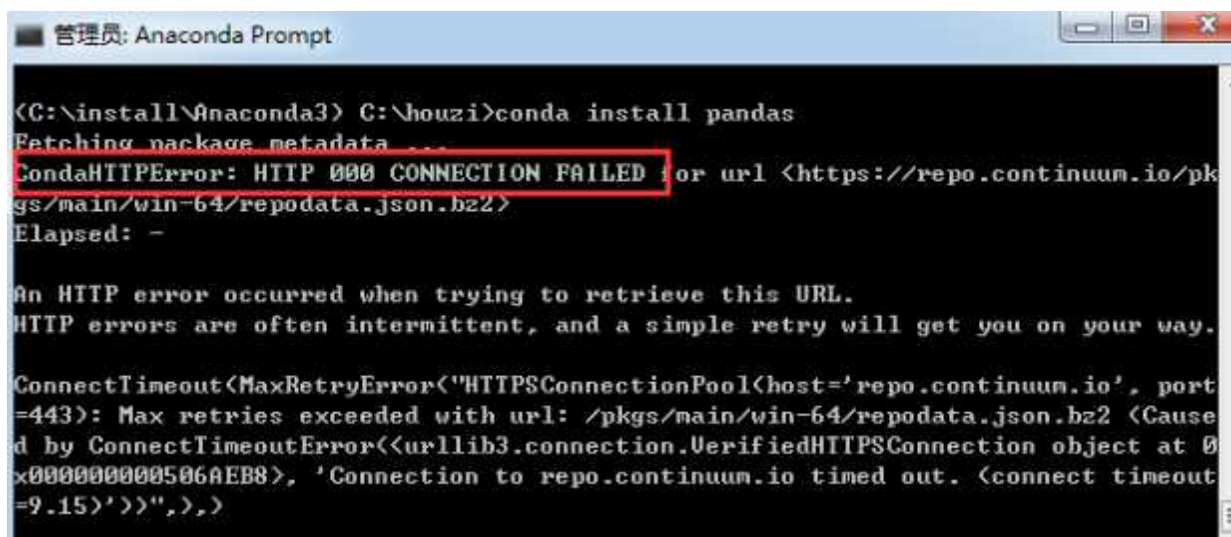
为了避免后面使用报错，你需要先更新下所有包。在终端输入更新所有包的命令：

```
conda upgrade --all
```

并在提示是否更新的时候输入 y（Yes）让更新继续。初次安装下的软件包版本一般都比较老旧，因此提前更新可以避免未来不必要的问题。

如果运行以上命令或者安装包报错：





```
管理员: Anaconda Prompt
(C:\install\Anaconda3) C:\houzi>conda install pandas
Fetching package metadata ...
CondaHTTPError: HTTP 000 CONNECTION FAILED for url <https://repo.continuum.io/pkgs/main/win-64/repo_data.json.bz2>
Elapsed: -

An HTTP error occurred when trying to retrieve this URL.
HTTP errors are often intermittent, and a simple retry will get you on your way.

ConnectTimeout(MaxRetryError("HTTPSConnectionPool(host='repo.continuum.io', port=443): Max retries exceeded with url: /pkgs/main/win-64/repo_data.json.bz2 (Caused by ConnectTimeoutError(<urllib3.connection.VerifiedHTTPSConnection object at 0x000000000506AEB8>, 'Connection to repo.continuum.io timed out. (connect timeout=9.15)>'))",),)
```

可以用下面的解决办法：

在终端中运行以下命令：

```
conda config --add channels
https://mirrors.tuna.tsinghua.edu.cn/anaconda/pkgs/free/**

conda config --set show_channel_urls yes
```

如果命令行方法添加不上，可以在“notebook工作文件夹”找到文件.condarc，

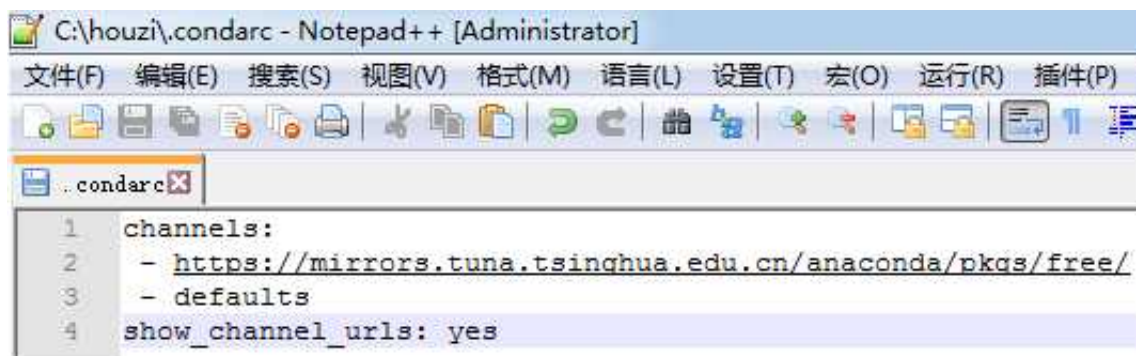
什么是你的“notebook工作文件夹”呢？

默认情况下，是你启动Anaconda Prompt终端中的那个文件夹，比如我电脑上下面是这个文件夹：



打开该文件夹，修改文件内容如下：

（如果没有找到.condarc，用notepad++软件在“notebook工作文件夹”下新建一个该文件即可，并添加以下内容）



为了方便，你可以将下面部分直接复制到上面的文件中：

```
channels:

- https://mirrors.tuna.tsinghua.edu.cn/anaconda/pkg/free/

- defaults

show_channel_urls: yes
```

### 3. 如何管理包?

安装了 Anaconda 之后，就可以很方便的管理包了（安装，卸载，更新）。

#### 1) 安装包

在终端中键入：

```
conda install package_name(这里的packagename就是你要安装的包的名字)
```

例如，要安装 pandas，在终端中输入：

```
conda install pandas
```

你还可以同时安装多个包。类似 `conda install pandas numpy` 的命令会同时安装所有这些包。还可以通过添加版本号（例如 `conda install numpy=1.10`）来指定所需的包版本。

conda 还会自动为你安装依赖项。例如，`scipy` 依赖于 `numpy`，因为它使用并需要 `numpy`。如果你只安装 `scipy` (`conda install scipy`)，则 conda 还会安装 `numpy`（如果尚未安装的话）。

#### 2) 卸载包

在终端中键入：

```
conda remove package_names
```

上面命令中的 `package_names` 是指你要卸载包的名称，例如你想卸载 `pandas` 包：

```
conda remove pandas
```

#### 3) 更新包

在终端中键入：

```
conda update package_name
```

如果想更新环境中的所有包（这样做常常很有用），使用：

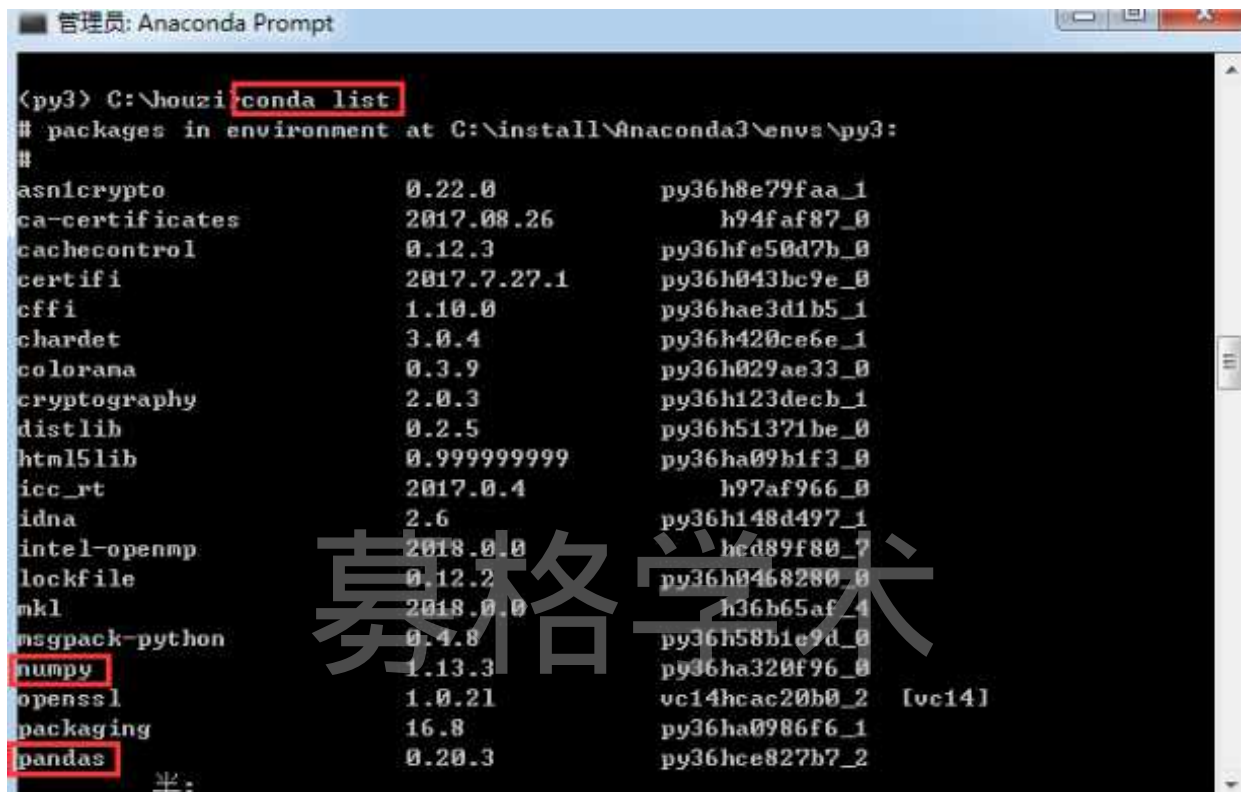
```
conda update --all
```

#### 4) 列出已安装的包

命令：

```
conda list
```

例如我已经成功安装了numpy和pandas这两个常用的包。



```
管理员: Anaconda Prompt
(py3) C:\houzi>conda list
# packages in environment at C:\install\Anaconda3\envs\py3:
#
asn1crypto                0.22.0                py36h8e79faa_1
ca-certificates            2017.08.26             h94faf87_0
cachecontrol               0.12.3                py36hfe50d7b_0
certifi                    2017.7.27.1           py36h043bc9e_0
cffi                       1.10.0                py36hae3d1b5_1
chardet                    3.0.4                 py36h420ce6e_1
colorama                   0.3.9                 py36h029ae33_0
cryptography               2.0.3                 py36h123dec1_1
distlib                    0.2.5                 py36h51371be_0
html5lib                   0.999999999           py36ha09b1f3_0
icc_rt                     2017.0.4               h97af966_0
idna                       2.6                   py36h148d497_1
intel-openmp               2018.0.0              hcd89f80_7
lockfile                   0.12.2                py36h0468280_0
mkl                        2018.0.0              h36b65af_4
msgpack-python             0.4.8                 py36h58b1e9d_0
numpy                       1.13.3                py36ha320f96_0
openssl                    1.0.2j                vc14hcac20b0_2 [vc14]
packaging                  16.8                  py36ha0986f6_1
pandas                     0.20.3                py36hce827b7_2
```

如果不知道要查找的包的确切名称，可以尝试使用 `conda search search_term` 进行搜索。例如，我知道我想安装numpy，但我不清楚确切的包名称。我可以这样尝试：

```
conda search num
```

最后重新再强调下，不要被上面的命令吓到。虽然上述命令多，给了谁都记不住的。用的多了自然记住了。你只需要跟着上面步骤操作下，并理解了每一步是干什么的就可以了。后面遇到要做的事情，忘记了回头查这个文档就可以了。

conda的官方文档：<https://conda.io/docs/user-guide/tasks/index.html>\*\*