

Python Note

tonguste@gmail.com

February 21, 2017

Summary

| | | |
|---|---|---|
| 1 | python max function using key and lambda expression | 1 |
| 2 | List indices must be integers, not list | 2 |
| 3 | numpy exchange between 1d array and 2d array | 2 |
| 4 | IPython with Pyspark | 3 |
| | List of Listings | 3 |
| | List of Listings | 3 |

1 python max function using key and lambda expression

By default in Python 2 key compares items based on a set of rules based on the type of the objects(for example a string is always greater than an integer). To modify the object before comparison or to compare based on a particular attribute/index you've to use the key argument.

Example 1:

```
1 >>>lis = [1,2,3,'4','5','14']
2 >>>max(lis)
3 WRONG!!!
4 >>>max(lis , key=lambda x: int(x))
5 # change the value to be order into interge
6 >>>lis = [(1,'w'),(2,'r'),(4,'g')]
7 >>>max(lis , key = lambda x:x[1])
8
```

Listing 1: lambda and max key

2 List indices must be integers, not list

List in Python must be integers, such as

```

1      a = [1,2,3,4]
2      a[1]      #will return 2
3      a[[1,0]] # it wont echo 2 1 as expected.
4      # via numpy
5      import numpy as np
6      a = np.array(a)
7      a[[1,0]] # Will return 2,1
8

```

However, array in **numpy** can be indiced by a list.

3 numpy exchange between 1d array and 2d array

Sllicing the i^{th} column in numpy is totally different with MatLab.

```

1      import numpy as np
2      c = np.arange(16).reshape((4,4))
3      # [[ 0,  1,  2,  3],
4      #    [ 4,  5,  6,  7],
5      #    [ 8,  9, 10, 11],
6      #    [12, 13, 14, 15]]
7      # if you want extract the 1 column, namely, [[1],[5],[9],
8      # [13]], you cannot use c[:, 1] or c[:,1] which doesn't return
9      # a vector
10     c[:, 1]
11     # array([1,  5,  9, 13])
12     c[:]
13     # array([4,  5,  6,  7]), surprisely!
14     # If you want change the some some column of c, you can do
15     # as:
16     c[:,1] = np.zeros((4,1))
17     # Oops...
18     # Traceback (most recent call last):
19     #   File "<stdin>", line 1, in <module>
20     # TypeError: data type not understood
21
22     c[:, 1] = np.zeros((4,1)).shape((1,4))

```

```

20 # Wrong again!!
21 np.zeros((4,1)).shape((4,1))
22 # array([[ 0.,  0.,  0.,  0.]]) A matrix NOT vector
23 # right way:
24 c[:, 1] = np.zeros((4,1)).shape(4)
25 # Okay, np.array.shape() (without tuple) will return a
    vector
26 #array([[ 0,  0,  2,  3],
27 #       [ 4,  0,  6,  7],
28 #       [ 8,  0, 10, 11],
29 #       [12,  0, 14, 15]])
30

```

4 IPython with Pyspark

We can use IPYTHON when lanuching the Pyspark console.

```

1 $ipython=1 IPYTHON_OPTS="--pylab" $SPARK_HOME/bin/pyspark
2

```

List of Listings