现代程序设计第八周作业

主要涉及生成器和迭代器的写法和功能实现

作业要求一

后项需要前项导出，且无法通过列表推导式生成。例如，时间序列中的“随机游走”便是一种满足上述条件的序列数据。其公式为$X_t = \mu + X_{t-1} + w_t$，其中$\mu$为漂移量，$w_t$是满足某种条件的独立同分布的随机变量，这里假设其服从正态分布N(0, σ^2)。本题要求写出实现该功能的迭代器函数。

库的导入

```
import pathlib
import numpy as np
import math
from pathlib import Path
from PIL import Image
from numpy.core.fromnumeric import shape
```

实现random_walk生成器

输入参数μ, X_0, σ^2，N，函数将迭代返回N个随机游走生成的变量。

```
def random_walk(u,X0,sigma_square,N):
    for i in range(N):
        X = X0
        vector = []
        for j in range(10):
            vector.append(X)
            random_value = np.random.normal(0,math.sqrt(sigma_square))
            X = X0 + u + random_value
        yield vector
```

利用zip，实现拼合多个random_walk的生成器

生成一组时间上对齐的多维随机游走序列。

```
def merge_random_walk(generate_vector):
    merge_vector =
list(zip(next(generate_vector),next(generate_vector),next(generate_vector),next(generat
e_vector),next(generate_vector)))
    return merge_vector
```

作业要求二

需要迭代的内容数据量过大，无法一次性加载。例如，在图像相关的深度学习任务中，由于数据总量过大，一次性加载全部数据耗时过长、内存占用过大，因此一般会采用批量加载数据的方法。（注：实际应用中由于需要进行采样等操作，通常数据加载类的实现原理更接近字典，例如pytorch中的Dataset类。）现提供文件FaceImages.zip(http://vis-www.cs.umass.edu/fddb/originalPics.tar.gz，其中包含5000余张人脸图片。要求设计FaceDataset类，实现图片数据的加载。

FaceDataset类

```
class FaceDataset:
    def __init__(self,path):
        path_class = Path(path)
        self.path_generate = path_class.rglob('*')
        self.path_shape()

    def path_shape(self):
        self.path = 0
        while self.path==0:
            path_init = str(next(self.path_generate))
            name = path_init.split('/')[-1]
            if name[0] != '.' and 'jpg' in name:
                self.path = path_init
        print(self.path)
        return self.path

    def image_process(self):
        image = Image.open(self.path)
        image_array = np.array(image)
        print(image_array)
        return image_array

    def __iter__(self):
        return self
```

```
    def __next__(self):
        self.path = self.path_shape()
```

代码测试

```
def main():
    print("hello")
    generate_vector = random_walk(3,10,10,5)
    merge_vector = merge_random_walk(generate_vector)
    print(merge_vector)

    path = '/Users/xihe/Documents/python_mp2021_project/week8_project/originalPics'
    Image_class = FaceDataset(path)
    for i in Image_class:
    #for i in range(1):
        image_array = Image_class.image_process()
        print(image_array)
    #print(next(Image_class.path_generate))
main()
```

测试结果

```
[(10, 10, 10, 10, 10), (12.951214959885244, 10.002295720524087, 18.952504818263765,
15.319606980500152, 10.483114404687608), (10.286402405236215, 7.139978936272824,
15.488042810627652, 14.87877927797158, 11.195889289435941), (13.854072012684567,
17.991358279948855, 16.078049543769772, 15.086476168329105, 14.051762303814927),
(14.900934758479686, 16.752354957309606, 12.892909537076893, 16.255450393689507,
11.157986786285068), (12.672617956451552, 15.696389549284392, 16.99037327104245,
10.329155126800702, 11.961487616675411), (17.388853816227574, 12.91587527134001,
11.350090052507255, 12.083900577168409, 12.165953252951343), (13.879638823460539,
13.41974140011332, 13.06396256923826, 15.58655665573661, 14.547068937837391),
(12.611506441264686, 13.449331072388976, 7.424045663276784, 13.947861600144057,
15.177312642857055), (10.690759714510591, 15.420207389543467, 16.35818309866311,
14.434104642352503, 10.806672120694778)]

/Users/xihe/Documents/python_mp2021_project/week8_project/originalPics/2002/11/03/big/i
mg_512.jpg
[[[ 94  84  74]
  [ 94  84  74]
  [ 94  84  74]
  ...
  [123  96  77]
  [123  96  77]
```

```
 [123  96  77]]

[[ 94  84  74]
 [ 94  84  74]
 [ 94  84  74]
 ...
 [123  96  77]
 [123  96  77]
 [123  96  77]]

[[ 94  84  74]
 [ 94  84  74]
 [ 94  84  74]
 ...
 [123  96  77]
 [123  96  77]
 [123  96  77]]

...

[[177 188 206]
 [178 189 207]
 [180 191 209]
 ...
 [181 194 226]
 [192 207 238]
 [192 207 238]]

[[179 190 208]
 [180 191 209]
 [181 192 210]
 ...
 [182 195 227]
 [193 208 239]
 [193 208 239]]

[[181 193 205]
 [180 192 204]
 [180 192 206]
 ...
 [187 201 227]
 [194 210 235]
 [192 210 234]]]
```