

FE520 Assignment 3

Dan Wang, Zhiyuan Yao

October 2018

1 Practice Numpy (25pt)

Please use Numpy build-in function to solve the problems below, For those problems that I did not cover in class, you can find them in supplement part of my notebook or you can check the official document.

1. Use Numpy to create a zeros array with shape = (3, 4) (here row = 3, column = 4), convert it to float64 data type.
2. Create an array with random number (`np.random.randn`), size =100, count the number whose value smaller than 0.4, and also return the ratio of it.
3. Create an random uniform 'array' with shape = (7, 7), low value = -10, high value = 10. convert the array to an 'new_array', which change the value to largest integer less than or equal to the value if the value in 'array' is smaller than 0, change the value to smallest integer greater than or equal to this value if the 'array' is greater or equal to 0.
4. Generate a normal distribution random number, with $N(1, 2)$, size = 1000, compute its mean, variance, skewness, and kurtosis.
5. Generate a random integer array, with size = 100, return the index of largest and smallest one.
6. Fit a line, $y = mx + c$, through some noisy data-points:
`x = np.array([0, 1, 2, 3])`
`y = np.array([-2.4, 0.5, 1.9, 4.1])`
return m and c by using `lstsq` method.

2 Uniform Distributed Random Number Generator (75pt)

You can not use any random number generators in packages like Numpy, because you are required to build your own generator in this question.

Almost all the random numbers generated by computer are pseudo random numbers, because they are generated by a formula. A good random number generator is

very important for financial simulations, such as Monte-Carlo method.

The goal of this assignment is to let you understand how computer generate random numbers and create pseudo random number generators which can generate 0-1 uniform-distributed random numbers by yourselves.

The probability density function (PDF) of 0-1 **uniform distribution** is

$$f(x) = \begin{cases} 1 & \text{for } x \in [0, 1] \\ 0 & \text{otherwise} \end{cases}$$

It means each point in interval $[0, 1]$ has the same probability to be chosen.

1. For implementation purpose, people use different algorithms to generate pseudo-random numbers. One of the most famous algorithms is called **Linear Congruential Generator**(LCG). This generator can yield a sequence of random numbers $\{X_i\}$. The recurrence relation of this algorithm is:

$$X_{n+1} = (aX_n + c) \bmod M$$

here a , c , M are all parameters. a is called multiplier which satisfies $0 < a < M$, c is called increment which satisfy $0 \leq c < M$, M is the modulus which is also stands for the maximum range of this sequence. You can find more information about the common parameters setting in the link above. There is another parameter X_0 which is so called **seed**. You can generate totally different sequence by different seeds. In the formula above *mod* stands for the modulus operation, for example: $7 \bmod 3 = 1$ due to $7 = 2 \times 3 + 1$.

As you can guess, this generator can generate a sequence of integers within $[0, M)$. To create a uniform distributed sequence, you only need to divide the sequence above by M .

For example, now I have a setting: $X_0 = 1$, and $a = 1103515245$, $M = 2^{32}$, $c = 12345$.

$$X_1 = (a \cdot X_0 + c) \bmod M = 1103527590$$

$$X_2 = (a \cdot X_1 + c) \bmod M = 25248852323$$

...

Then the uniform distributed sequence will be

$$\left\{ \frac{1}{2^{32}}, \frac{1103527590}{2^{32}}, \frac{25248852323}{2^{32}}, \dots \right\} \sim \{2.328 \times 10^{-10}, 0.2569, 0.5879, \dots\}$$

There are some variants of LCG, one of them has such recurrence relation:

$$X_{n+1} = (X_n(X_n + 1)) \bmod M$$

The seed of this generator has to satisfy $X_0 \bmod 4 = 2$.

Now it's your turn to implement it in Python 3.

- (20 pt) Create a file called `generator.py`. In this file, create a class called **LCG** whose instance can generate random numbers by Linear Congruential Generator algorithm. It should have these data attributes: `self.seed`, `self.multiplier`, `self.increment`, `self.modulus` which are assigned in initializer (`__init__()`). The parameters should be passed by argument from outside of the class. This class's instance should also at least have function attributes allow me: 1) get the seed; 2) set the seed; 3) initialize the generator (start from the seed); 4) give me the next random number; 5) give me a sequence (list) of random number, the length should be passed by argument.
 - (20 pt) In `generator.py`, create an inherited class called **SCG** from class **LCG** you have created. You are required to implement the variant of LCG I mentioned [above](#) in this class. It should have the same interfaces (attributes) with **LCG**, but different recurrence relation. Remember the seed of this generator has to satisfy $X_0 \bmod 4 = 2$, so you need to check this condition in initializer, raise error with customized error information if the seed doesn't satisfy this condition.
Hint: To better reuse your code, your class should have a function attribute to specify the recurrence relation. You can only override this function in your inherited class.
 - (Bonus 3pt) Specify `__next__()` or `__iter__()` attributes in your class to make the instance of your class compatible with `iter()` or `next()`, this kind of generator can yield me one random number for each call.
 - A good code style like following good conventions, writing comment and doc string for your code will earn additional bonus.
2. To test the performance of your generator, one way is to use a simple Monte-Carlo method to test it. Consider a square and its inscribed circle in a 2-D rectangular coordinate. The length of square and the diameter of the circle are 2, and the center of them are both origin of this coordinate. See [2](#).

You can populated the points in this 2-D coordinate within this square, and count the number of points which fall in the circle. Then dividing this number by the number of all points will give you an estimate of ratio of area between inscribed circle and its square . Since you have already have theoretical result of this ratio:

$$ratio = \frac{\pi r^2}{4r^2} = \frac{\pi}{4} \sim 0.78539816339$$

where r is the radius of the circle, thus the length of square is $2r$. The closer your result is to this ratio, the better your random generator performs.

Now you need to use Python to do this:

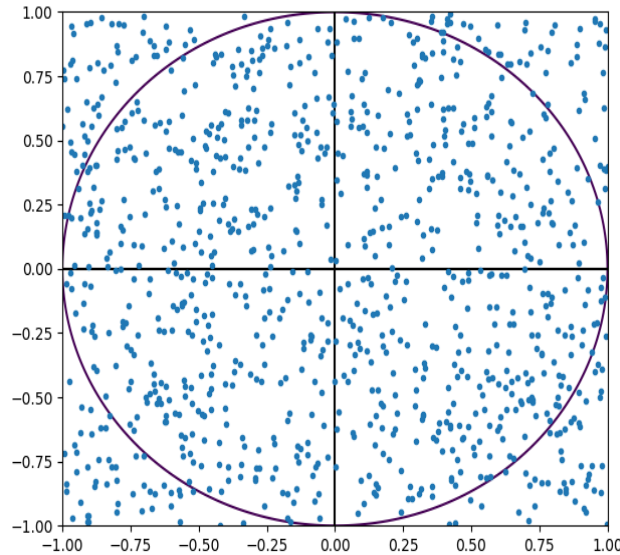


Figure 1: $x^2 + y^2 = 1$, and its circumscribed square. Blue points are uniform distributed random points

- (20 pt) Create a file called `point.py`. Write a class to represent the points in rectangular coordinate. It should have the data attributes that store its x coordinate and y coordinate, and also a function attribute `distance` to calculate its distance from origin point.
 - (15 pt) Create a file called `MCTest.py`. Import the classes you created above (generators and `point`) from Python modules. Populate 10,000,000 random points within this square. To do this, you only need to populate 2×10^7 0-1 uniform distributed random numbers firstly, then re-scale them into $[-1, 1]$, and pair them into 1×10^7 points. Test both generators.
Determine whether points are within the circle. Hint: You can do this by computing its distance from origin, if it is smaller than 1 (the radius), it is within the circle.
Give the estimate of above ratio, and the difference between yours and theoretical one.
 - You are also encouraged to record the time consumed by your program. Hint: use `time` package.
3. (Bonus 5pt) For the task above, use `multiprocessing` package to parallel your code, compare the time consumed by one process and that by multiple processes. Find more information in [official documentation](#).

For this question, you need to submit 3 file: generator.py, point.py and MCTest.py.

3 Bonus Question (10pt)

Finish the problem ([remove duplicates from sorted list](#)) and problem ([best time to buy and sell stock](#)) on Leetcode. Submit a screen-shot of your submission detail where you can see a run time distribution, and submit the file code also.

Submission Requirement:

For all the problems in this assignment you need to design and use Python 3, output and present the results in nice format. Please submit a written report (pdf), where you detail your results and copy your code into an Appendix. You are required to submit a single python file and a brief report. Your grade will be evaluated by combination of report and code. You are strongly encouraged to write comment for your code, because it is a convention to have your code documented all the time. In your python file, you need contain both function and test part of function. Python script must be a '.py' script, Jupyter notebook '.ipynb' is not allowed.

Do NOT copy and paste from others, all homework will be firstly checked by plagiarism detection tool.