Hello Professor Dixon,

I have questions regarding Home Work assignment 2.

1. Can you elaborate on what is “checkpoints” and what exactly this data structure represents? How does it look (i.e. can you give a short example with actual values)? How the size of “checkpoints” related the size of “samples”?
2. What is rateCurve? Does it replace constant “r” that we used before? Is it to model how rate changes over time? If so, what is the size of this array and how it is aligned with the size of “checkpoints”?
3. I’m confused with how “M” is defined in case we need to generate “samples”. Is it literally the value of the last element in “checkpoints” or is it the size of “checkpoints”. I guess the confusion is from me not getting what the “checkpoints” is. Can you please elaborate?
4. Is “samples” a list or numpy array? The description uses both terms in different places. What is it though?
5. Can you please elaborate on what is array “t” is Is it just integer numbers representing time step (i.e. 1,2,3,…N) or is it an incremental fractions of the “T” (i.e. 1\*(T/N), 2\*(T/N), 3\*(T/N),….T). Can you give a short example with values?
6. We are asked to price “vanilla” options. What is the “vanilla” in the context of the assignment? Is it a Call, a Put or a set of both? If it’s both, in what form we suppose to return in from the function within “TV”?

Hi Oleksandr,

1. Checkpoints is a list of integer of sample counts at which you have to return the running total, a running mean, standard deviation, and estimated error. So you have to iterate through those numbers one by one using for or while loop after under a code block to calculate the running total, a running mean, standard deviation, and estimated error will be passed. Checkpoints data structure is either a NumPy array or a list in which you will pass a number of samples in which your code should be run to calculate all that statistics. Eg, checkpoints = [100 200 300 400 1000000], so first you will calculate those statistics for 100 sample points that are mean of prices using first 100 z's (random normal number of shape (total\_samples,1), then for 200 sample points, and so on. If your sample variable = None, then the last number in the list will be the total number of sample prices.
2. rateCurve: you have to download Treasury Yield Data and that will have treasury yields for 1month, 2m, 3m, 6m, 1yr, 2y, 5 yr, and so on. So if you want to find the rate, r, variable for say 10 months you have to linearly interpolate the rate values between 6m and 1y. You should be able to do that using np.interp function wherein you can pass the number of months (in terms of years) and yield series from Treasury yield dataset. Yes, it replaces the r variable and in this case, you will need to find the estimate r based on Treasure yield dataset.
3. M is either the last element of the checkPoints list if sample == None or it is equal to the values you pass in the sample variable.
4. Samples variable either None or the total number of samples (M). Checkpoints variable is a list or NumPy array.
5. t is an array of timestep variable, t = T/M, where M is the number of paths in the sample matrix. You are right that t is incremental fraction of the “T” (i.e. 1\*(T/N), 2\*(T/N), 3\*(T/N),….T). Eg.: It could look like if M = 100, and T = 1, t = [0.01 0.02 ...1.0], where size of t = (1, 100)
6. I am not 100% sure about this question. I think it is a call option. But, you should definitely refer to the professor for this question during his office hours.

I hope I am able to clear your doubts. I encourage you to join TA office hours if you have other doubts.