2018 - exam

Stochastic Simulation (Universiteit van Amsterdam)

Written exam stochastic simulation.
20 December 2018, 09.00 - 12.00, USC Sporthal 1

This exam has 3 assignments, and each assignment has equal weight.

Assignment 1

Consider a travelling salesman problem with N cities, and a distance matrix **d**, so that d_{ij} is the distance between city i and j. Also assume that standard encoding of a tour as a vector **t**, where t_i is the next destination after city i.

Suppose we want to create a Markov chain, where we jump from one possible tour to a next possible tour. Suppose also that we use the Lin 2-opt transition rule to generate a next tour from the current tour. The transition probability matrix is **P**, so that P_{kl} is the probability to jump from tour k to tour l under the Lin 2-opt rule.

1.1 What is the dimension of the transition probability matrix **P**? [1 point] ✓

For any given k, the transition probabilities P_{kl} can only take two values.

1.2 What are the two possible values, explain. [1 point] ✓

1.3 And what value has P_{kk}? Why? [1 point]

1.4 Prove that $P_{kl} = P_{lk}$. [1 point] ✓

1.5 What does ergodicity of a Markov chain mean? [1 point] ✓

Without further proof we state that the current Markov chain using the transition probability matrix **P** is ergodic, and therefore stationary probabilities $\{\pi_j\}$ exist for this Markov chain.

1.6 Compute $\{\pi_j\}$.[1 point] = 1/N! as each state has the same chance occurring

1.7 Markov chains form the core of the Simulated Annealing (SA) algorithm. Explain in words how Markov chains are used in SA, using the terminology that you learned during classes. Make sure you focus your discussion on special properties of Markov chains that make SA work. [2 point] Metropolis algorithm 1/2

In order to use the Markov chain above for generating a new tour from an existing tour in Simulated Annealing to find the shortest tour it needs to be changed.

1.8 Explain how you would change the Markov chain so that it will produce the shortest tour in a Simulated Annealing algorithm. Also explain why. [2 points]

Assignment 2

Consider a discrete random variable X, with $P(X = i) = ic, i \in \{1,2,3,4,5\}$ and $P(X) = 0$ otherwise.

2.1 Compute c. [1 point] ✓

Consider a binary random variable $Y(X)$, with

$$Y(X) = \begin{cases} 1 & \text{if } 3 \leq X \leq 4, \\ 0 & \text{otherwise.} \end{cases}$$

2.2 Compute $E[Y(X)]$ and $Var[Y(X)]$. [2 points]

Turn page over

Suppose we now want to estimate $\theta = E[Y(X)]$ via a stochastic simulation. To do so, you need to be able to sample Y.

2.3 Propose an algorithm to sample Y using a random number generator that samples from a continuous homogenous random variable U, where $0 \leq U \leq 1$. [2 points] ✓

Suppose you take n samples Y_i using your algorithm from 2.3, and you compute the sample mean $\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$.

2.4 Proof that the sample mean $\bar{Y}$ is an unbiased estimator of θ. [2 points] ✓

2.5 Compute $\mathrm{Var}[\bar{Y}]$.[1 point] ✓

Suppose we want to estimate θ, with 95% confidence, with a relative accuracy of 5%

2.6 How many samples n do we need to take? [2 points] 1/2

<u>Assignment 3 [10 points]</u>

3.1 What is Uncertainty Quantification, and explain the difference between forward and inverse Uncertainty Quantification? [2 points]

3.2 Explain the difference between nonintrusive and intrusive forward Uncertainty Quantification? [2 points]

Consider a situation where you want to perform a non-intrusive forward Uncertainty Quantification on a function $y = f(\mathbf{p})$ where $\mathbf{p}$ is a vector of continuous random input variables and y is a continuous scalar random output variable.

3.3 Describe a Monte Carlo approach to enable this non-intrusive forward Uncertainty Quantification. Clearly indicate which random variable(s) you would estimate in this procedure, and how. [2 points]

Further assume $\mathbf{p} = p_1$, so a 1-dimensional input space, and consider a situation where we want to estimate the uncertainty on $y^0 = f(p_1^0)$, and where p_1^0 has a uniform distribution with a range of $\pm 10\%$ of its mean value.

3.4 Using the Monte Carlo approach for 3.3, you now need to sample from a one-dimensional parameter space. Assuming that you want to estimate the uncertainty with 95% confidence intervals less than some threshold value ε, describe a procedure to determine how many samples you need to take. [2 points]

Finally assume that the model we solve is a one-dimensional diffusion equation, $\frac{\partial y}{\partial t} = D\frac{\partial^2 y}{\partial x^2}$
where y is now a scalar function that depends on time t and position x and where D is the diffusion coefficient. Also assume that the only uncertainty in this problem is an uncertainty in the initial condition $y_0(x) = y(x, t = 0)$

One way to derive an intrusive Uncertainty Quantification method is the perturbation method, where we can write $y(x, t;\ \xi) = \langle y\rangle + \delta y$, where ξ is understood as a random variable to express the uncertainty in $y_0(x)$, $\langle y\rangle$ is the expected value of $y(x, t;\ \xi)$ and δy a perturbation that expresses the uncertainty.

3.5 Formulate an intrusive Uncertainty Quantification for the one-dimensional diffusion equation using the perturbation method. [2 points]