Final Exam: ELG 5218 Uncertainty Evaluation in Engineering Measurements and Machine Learning, 2022

Instructions

- Total number of points is 125 which means that you do not need to do everything. Maximum number of point that you can get is 100.
- **Submit using Bright Space by Wednesday April 13, 2022 at 3:30pm**
- If there are formulas or text to be written, you can write it in notebook or you can write on paper and then scan it and upload it.
- If you have questions, please send me an email at mbolic@uottawa.ca .

Make sure that you read the statement about the professional integrity and that you write your name, student number and sign below.

I understand the importance of professional integrity in my education and future career in engineering or computer science. I hereby certify that I have done and will do all work on this examination entirely by myself, without outside assistance or the use of unauthorized information sources. Furthermore, I will not provide assistance to others.

Je comprends l'importance de l'intégrité professionnelle dans mes études et ma future carrière en génie ou en informatique. Je certifie par la présente que j'ai fait et ferai tout le travail pour cet examen entièrement par moi-même, sans assistance extérieure ni utilisation de sources d'information non autorisées. De plus, je ne fournirai aucune assistance aux autres.

Name: ___

Student number: __

Signature: __

Problem 1 Short computational problems

Total 40 points

a) You are given a collection of n documents, where the word count of the i-th document is x_i . Assume that the word count is given by an exponential distribution with parameter λ: for a non-negative integer x, $P(wordcount = x|\lambda) = \lambda^x exp^{-\lambda}/x!$ Compute λ such that the likelihood of observing $\{x_1, x_2, \dots, x_n\}$ is maximized.

b) Consider samples $x_1, \dots, x_n$ from a Gaussian random variable with known variance σ^2 and unknown mean μ. We further assume a prior distribution (also Gaussian) over the mean, μ $\sim N(m; s^2)$, with fixed mean m and fixed variance s^2. Thus the only unknown is μ.

1. Calculate the MAP estimate μ_{MAP} . You can state the result without proof. Alternatively, you can compute derivatives of the log posterior, set to zero and solve.
2. Show that as the number of samples n increase, the MAP estimate converges to the maximum likelihood estimate.
3. Suppose n is small and fixed. What does the MAP estimator converge to if we increase the prior variance s^2?
4. Suppose n is small and fixed. What does the MAP estimator converge to if we decrease the prior variance s^2?

c) We are looking at the player shooting basketball free throws. The prior is uniform distribution. We observe the player take three shots, with two of them resulting in scores (success). What is the probability that the player will making the next basket (the fourth one) successfully?

d) Consider the model $y = x_1 \cdot x_2 + x_3 \cdot x_4$ where $x_i \sim N(0, \sigma_i^2)$ for $i = 1, \ldots, 4$. Find Sobol indexes S_i and S_{ij}

Problem 2 Kalman filters, sensor fusion and RNNs

Total 40 points

a) You decide to use Kalman filter to perform sensor fusion. For example, you are tracking a car using two different sensors.

1. For both of these sensors you have the measurement equations that relate the measurements of the sensor with the position of the car. What will change in Kalman filter equations so that you can now include observations from 2 sensors (instead of having observations from only one sensor).
2. Now, assume that you replaced the original sensor 2 with the camera and for the camera you obtain videos and you do not have measurement equation for it. You still want to use Kalman filter for sensor 1. How could you combine information from both sensors?

b) Estimate the unknown value x at the time instances $t = 1h$ and $t = 2h$. Imprecise measurements of y are performed every hour. Prior information include $E[x_0] = 10$ and $\sigma(x_0) = 7$. State space model is shown below:
$y_t = x_t + v_t$ and $v_t \sim N(0, 3^2)$
$x_t = x_{t-1} + w_t$ and $w_t \sim N(0, 0.5^2)$

The observations are $y_1 = 4.8$ and $y_2 = 12.1$.
Your task is to run computation of Kalman filter by hand and to show the values of the mean and variance of the state variable x at time instants $t = 1h$ and $t = 2h$. Comment on the value of Kalman gain.

c) Describe resampling algorithm for particle filters. Is resampling also a kind of Monte Carlo sampling? Why?

d) Explain the problem of vanishing gradients in RNN networks. How is that problem addressed?

Problem 3 Gaussian processes

Total 20 points

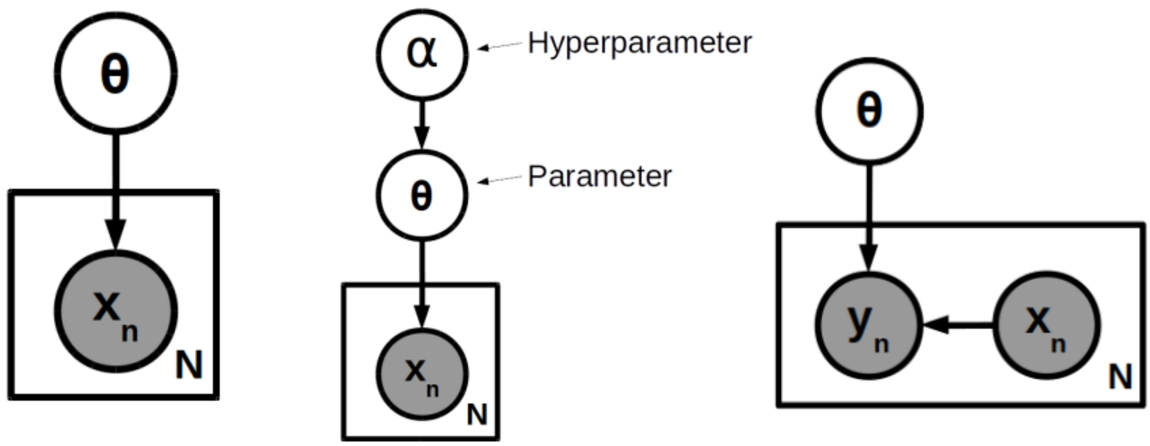
Let's look at the Gaussian process model. We are interested in predicting the response y_* in a test case with inputs x_*. Now, suppose we have just one training data point ($x_1 = 3$ and $y_1 = 4$). The noise-free covariance function is $K(x, x') = 2^{-|x-x'|}$, and the variance of the noise is 1/2. Find the mean and variance of y_* for which the value of the input is $x_* = 5$.

Problem 4 Short general questions

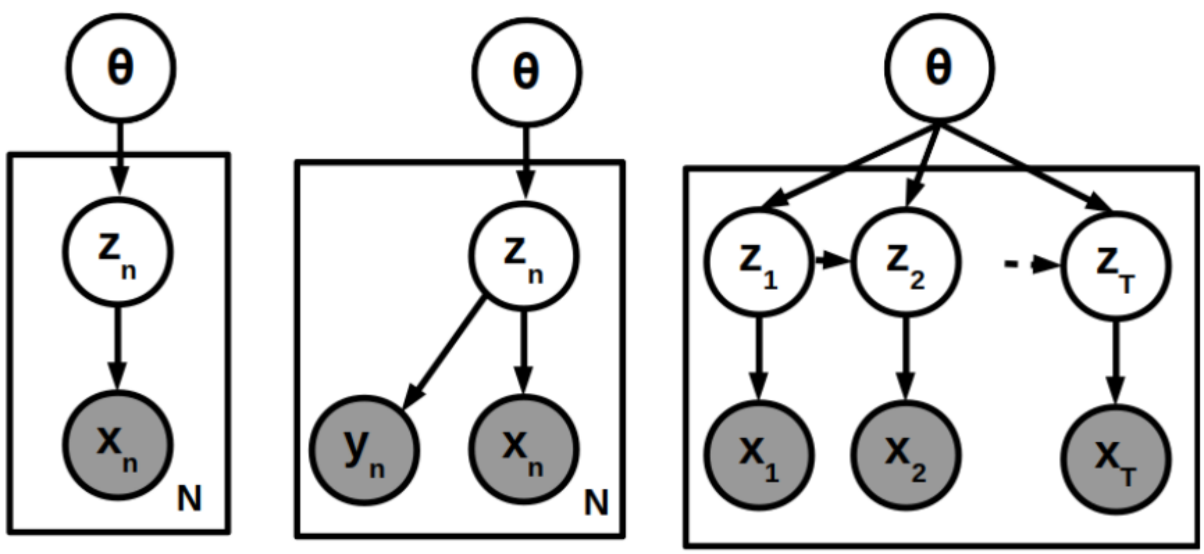
Total 25 points

a) Six grahical models are shown in figures below. Explain the following:

1. What does each model represent? What do the parameters and variables mean?
2. Relate each graph with the model we covered in the class.



a) b) c)



c) d) e)

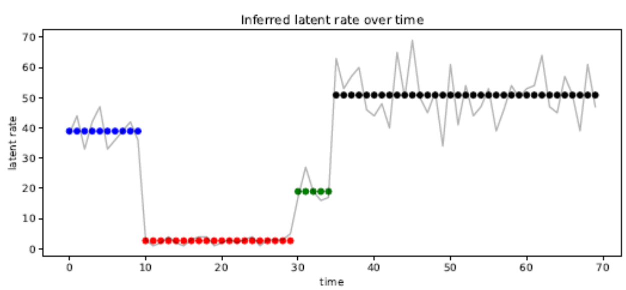
b) You are asked to work on the regression problem. Your data has single dimension in both x and y and you have small number of data points.

1. You are considering Bayesian regression (please note that the regression is linear in parameters but it does not need to be linear in x), Bayesian neural networks and Gaussian processes. List advantages and disadvantages of each approach.
2. Next, you ended up being able to apply known physical model but the model does not fit perfectly. You decided to model residuals using one of the three methods above: Bayesian regression, Bayesian neural networks and Gaussian processes. Which method would you use? How would you train your system?

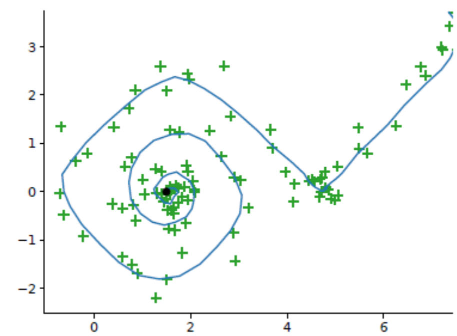
c) Uncertainty propagation is performed using correlated input data.

1. How would that affect computation of uncertainty of the output using perturbation method?
2. How would that affect computation of uncertainty of the output using Monte Carlo method?

d) Explain in your opinion what models are used to generate data shown in following figures.



a)



b)

e) Training of Bayesian neural network is done using variational inference? What is the reparameterization trick and why does it need to be applied?