Homework 4
Derivatives and Quadrature *

10-424/10-624 Bayesian Methods in ML
https://www.cs.cmu.edu/~hchai2/courses/10624

OUT: 03/27/25
DUE: 04/17/25

Instructions

- **Collaboration Policy**: Please read the collaboration policy in the syllabus at https://www.cs.cmu.edu/~hchai2/courses/10624/#Syllabus.
- **Late Submission Policy:** See the late submission policy in the syllabus at https://www.cs.cmu.edu/~hchai2/courses/10624/#Syllabus.
- **Submitting your work:** You will use Gradescope to submit answers to all questions.
  - **Written:** We will provide you with an Overleaf template for you to complete the written portion of this homework. You may also use the raw LaTeX source of this assignment (included in the handout .zip) to typeset your answer. **You must use LaTeX to complete this assignment**; we will not grade any submissions that are not completed using LaTeX and you will be asked to resubmit (with some penalty). You will submit your completed homework as a PDF to Gradescope.
  - **Programming:** This assignment is a written-only assignment. However, for some questions you may find that it is easier and/or beneficial to write some code to help complete your solutions; you may use any programming language and any libraries you wish to complete this assignment (provided you abide by the collaboration policy as outlined above). You are not required to submit any code you write to complete this assignment. **You are required to programmatically generate all figures**.

*Compiled on Thursday 27$^{\text{th}}$ March, 2025 at 04:54

In this homework, you will consider the function

$$f(x) = \exp(-x^2)$$

and its definite integral

$$Z = \int_{-\infty}^{\infty} f(x)\,\mathrm{d}x.$$

It turns out that even this simple function has no elementary antiderivative, so the calculation of Z is not straightforward. However, there is a relatively famous method for computing Z with the trick of considering Z^2 instead, rewriting the resulting $2d$ integral in polar coordinates, and making a convenient substitution. Cutting the punchline, the result is

$$Z = \sqrt{\pi}.$$

For each of the questions below, model f with a zero-mean Gaussian process prior that uses a simple squared exponential kernel,

$$f \sim \mathcal{GP}(f; \mu = 0, k) \text{ where } k(x, x') = \exp(-(x - x')^2),$$

and condition this prior belief on the following dataset $\mathcal{D} = (\mathbf{x}, \mathbf{y})$:

$$\begin{aligned}
\mathbf{x} &= [-3, -2, -1, 0, 1, 2, 3]^\top; \\
\mathbf{f} &= \exp(-\mathbf{x}^2) \\
&= [1.2341 \times 10^{-4}, 1.8316 \times 10^{-2}, 0.36788, 1, 0.36788, 1.8316 \times 10^{-2}, 1.2341 \times 10^{-4}]^\top.
\end{aligned}$$

1 Derivative Observations (30 points)

Consider conditioning a Gaussian process on an observation of the derivative at some location.

1.1. (5 points) Suppose a 1-dimensional function, $f\colon \mathbb{R} \to \mathbb{R}$, follows an arbitrary GP belief: $\mathcal{GP}(f; \mu, K)$. Write down the implied joint distribution between the function's value at an arbitrary point x, $f(x)$, and the value of a derivative at another arbitrary point x', $f'(x) = \frac{\mathrm{d}f}{\mathrm{d}x}|_{x'}$.

1.2. (10 points) Now, instead of an arbitrary GP, suppose f follows the a zero-mean GP with $k(x, x')$ as defined above. Write down the joint distribution from the first part of this question. Evaluate any derivatives or integrals you may encounter; **do not use the expression** $k(\cdot, \cdot)$ **in your answer**, you should explicitly write out the functional forms.

1.3. (5 points) Show that for the model described in the previous part, the function's value at a point x and the value of the derivative *at that point* are independent *a priori*.

1.4. (10 points) Condition this model on the dataset $\mathcal{D}$ defined above as well as the observation that the derivative is 0 at $x = 0$, to derive a posterior GP belief $p\big(f \mid \mathcal{D}, f'(0) = 0\big)$.

Plot the posterior mean and the point-wise 2 standard deviation credible interval for this GP on the interval $x^* \in [-5, 5]$.

2 Bayesian Quadrature (20 points)

Again, model f by a zero-mean GP with kernel $k(x, x')$ as defined above and condition this GP model on the dataset $\mathcal{D}$ defined above. Let

$$Z(a) = \int_{-a}^{a} f(x)\,\mathrm{d}x.$$

2.1. (10 points) Using the posterior GP described above, compute the posterior mean and standard deviation of $Z(3)$.

You should both write out the formula for these quantities *and* provide a numeric answer. Evaluate any integrals in your answer using a numerical integration method of your choice.

2.2. (10 points) Plot the posterior mean and 2 standard deviation credible interval for $Z(a)$ using a dense grid over the range $a \in [0, 5]$; again, you should use numerical integration to compute the relevant intractable quantities.

How does $Z(a)$ compare with the $Z = \sqrt{\pi}$ as a function of a?

3 Collaboration Questions (0 points)

After you have completed all other components of this assignment, report your answers to these questions regarding the collaboration policy. Details of the policy can be found in the syllabus.

3.1. Did you collaborate with anyone on this assignment? If so, list their name or Andrew ID and which problems you worked together on.

3.2. Did you find or come across code that implements any part of this assignment? If so, include full details.