COURSEWORK

AFRICAN INSTITUTE FOR MATHEMATICAL SCIENCES

QUANTUM LEAP AFRICA

Foundations of Machine Learning

Instructions

This assignment has only a coding component. The coding part must be done in python.
You are not permitted to use any symbolic manipulation libraries (e.g. `sympy`) or automatic differentiation tools (e.g. `tensorflow`, `pytorch`) for your submitted code (though, of course, you may find these useful for checking your answers). You should not need to import anything other than `numpy` for the submitted code for this assignment.
The writing assignment requires plots, which you can create using any method of your choice. Do not submit the code used to create these plots.
No aspect of your submission may be hand-drawn. You are strongly encouraged to use LaTeX to create the written component.

Do not submit jupyter notebooks and check whether your python code runs.

In summary, you are required to submit a file `mcmc_coursework_<your-name>.py` that implements all the methods for the coding exercises. Send it to `mdeisenroth@aimsammi.org`.

The submission deadline is **November 6, 23:59**.

Gibbs Sampling

We are interested in Gibbs sampling for linear regression with one independent variable. We assume we have paired data (x_i, y_i), $i = 1, \ldots, N$. We wish to find the posterior distributions of the coefficients β_0 (the intercept), β_1 (the gradient), β_2 of a quadratic contribution and of the precision τ, which is the inverse variance. The model can be written as

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \epsilon, \quad \epsilon \sim \mathcal{N}(0, \tau^{-1})$$

where τ is the precision (inverse variance).
The likelihood for this model may be written as the product over N i.i.d. observations is

$$p(y|X, \beta_0, \beta_1, \beta_2, \tau) = \prod_{i=1}^{N} \mathcal{N}(y_i|\beta_0 + \beta_1 x_i + \beta_2 x_i^2, \tau^{-1})$$

We place conjugate priors on $\beta_0, \beta_1, \beta_2$ and τ. We choose

$$\begin{aligned}
\beta_0 &\sim \mathcal{N}(\mu_0, \tau_0^{-1}) \\
\beta_1 &\sim \mathcal{N}(\mu_0, \tau_1^{-1}) \\
\beta_2 &\sim \mathcal{N}(\mu_0, \tau_2^{-1}) \\
\tau &\sim \text{Gamma}(\alpha, \beta)
\end{aligned}$$

Implement a Gibbs sampler for this model.
The general approach to deriving an update for a variable is:

1. Write down the log-joint distribution over all variables
2. Throw away all terms that do not depend on the current sampling variable
3. Pretend this is the density for your variable of interest and all other variables are fixed. What distribution does the log-density remind you of?
4. That is your conditional sampling density.

Task 1 (20 marks)

Implement a sampler for the conditional distribution

$$p(\beta_0|\beta_1, \beta_2, \mu_0, \tau_0, \tau, X, y)$$

Hint: This conditional is Gaussian.

Task 2 (20 marks)

Implement a sampler for the conditional distribution

$$p(\beta_1|\beta_0, \beta_2, \mu_1, \tau_1, \tau, X, y)$$

Hint: This conditional is Gaussian.

Task 3 (20 marks)

Implement a sampler for the conditional distribution

$$p(\beta_2|\beta_0, \beta_1, \mu_2, \tau_2, \tau, X, y)$$

Hint: This conditional is Gaussian.

Task 4 (20 marks)

Implement a sampler for the conditional distribution

$$p(\tau|\beta_0, \beta_1, \beta_2, \alpha, \beta, X, y)$$

Hint: This conditional is Gamma.

Task 5: Gibbs sampler (20 marks)

Implement the Gibbs sampler using the conditional distributions from Tasks 1–4.