ASTRO 410 — Mid-term Exam

DUE: 03/31/2021

IMPORTANT: Solutions to the mid-term exam should include a write up in pdf format (using Latex or Word) detailing your answers to the questions and results (e.g., plots or tables), all your programs and all output files generated by your programs, as well as a README file that lists all the files included, and instructions to compile and run your programs. All the files should be submitted to CANVAS as a single tarred, gzipped (tgz) file named midterm-lastname.tgz. The file when expanded should yield a directory named midterm-lastname. **Note: submission after the due date will not be accepted for this exam.**

Report Requirements:
The assignment report is a scientific report which should include the following sections:
1. Describe the equations and algorithms to solve the assignment;
2. Answer the questions in the assignment;
3. Present the results with figures and tables (if applicable).

1. [100 points] Use Markov Chain Monte Carlo (MCMC) and Metropolis sampling algorithm to fit the data midterm.dat, which is the same as in Homework 2.

The data can be modeled as a Gaussian function, $G(x, \mu, \alpha_D, A)$. Use the MCMC to estimate its parameters, μ, α_D, and A, where

$$G(x, \mu, \alpha_D, A) = \frac{A}{\alpha_D}\sqrt{\frac{\ln 2}{\pi}}e^{\frac{-(\ln 2)(x-\mu)^2}{\alpha_D^2}}. \tag{1}$$

(1) [25 points] [**Grading Instructions: The report must include the steps and equations of the derivations. There are 4 main equations in the derivations, Bayes theorem (5 points), probability of the data (provided below), posterior probability of the model parameters (15 points), posterior probability expressed in likelyhood of the data set (5 points).**]
Assume each data point is distributed normally with mean given by the model, $P(y_i) = \frac{1}{\sqrt{2\pi}\sigma_i}e^{-\left(\frac{y_i-m_i}{\sqrt{2}\sigma_i}\right)^2}$, where $m_i = G(x_i, \mu, \alpha_D, A)$, and σ_i is the standard deviation of the data points due to the measurement uncertainties. Derive the joint posterior distribution of parameters μ, α_D, and A using the Bayes theorem.

(2) [35 points] [**Grading Instructions: The report must include the steps of the MCMC algorithm. There are 3 main steps in the algorithm /code: the**

proposal function for the new move using the Metropolis algorithm (15 points), the acceptance ratio (10 points), the initial parameters (5 points).]
Construct a Markov Chain to sample from the posterior, assuming flat prior for μ, α_D and A, and using a uniform proposal distribution. Write a computer program to Implement the Markov Chain.

(3) [30 points] [**Grading Instructions: The report must include at least 3 MC chains of each of the three model parameters (μ, α_D, A), and for each MC chain there are at least 10000 iterations (5 points for the MC chains for each of the 3 parameters). The plots must be included in the report (5 points for plots of each parameter).**]
Run the MCMC program and experiment with different starting values of the parameters and the width of the proposal distribution function, and choose the most appropriate ones that result in a well-mixed Markov Chain. Plot μ, α_D, and A along the Markov Chain history, and determine the burn-in period. Plot the distribution functions of μ, α_D, and A using the Markov Chain after discarding the burn-in period.

(4) [10 points] [**Grading Instructions: The report must include the plots of each of the 3 model parameters (μ, α_D, A) (5 points for each parameter). The plots must include the MCMC history (parameter value as a function of the number of iterations), and the probability distribution of the parameter.**]
Show the best fitting parameters and estimate the mean and standard deviation of (μ, α_D, A). Plot the data and the best fit curve. Compare them with the best-fit parameters and errors derived from the non-linear least squares fit in homework 2.