Multi-Core Programming
EE 5885 – 04
Dept. of Electrical and Computer Engineering
ASSIGNMENT 4, Points: 30, Due Date: 04/19/2021 by 4:00 PM

Implement the least-squares fitting, a popular linear regression method, as a parallel program using OpenMP. Simple linear regression is a statistical technique that models a linear relationship between two variables. The least-squares fitting uses a set of sample data points to determine the best fit or optimal curve between two variables. The least-square fit when used with the simple linear regression model, the optimal curve is a straight line whose equation is

$$y = mx + b$$

Where,

x is the independent variable
y is the dependent or measured variable
m is the line's slope
b is the line's y-axis intercept point

The least-squares slope and intercept point are calculated using the formulas

$$m = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2}$$

$$b = \frac{\sum_{i=1}^{n} x_i^2 \sum_{i=1}^{n} y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i y_i}{n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2}$$

Where,

n is the number of sample data points
x and y both are of single-precision floating-point data types

Implementation: *Using the provided project template (Assignment4) with the sequential implementation, implement both parts in a single project*

Part 1: *Implement* ***a data-parallel program using only the directive parallel-for with appropriate clauses*** *to compute the individual components of* ***m and b****.*
Part 2: *Implement a parallel program to compute the four components as parallel tasks using the given individual component functions in the project template. Each task should further implement* ***data-parallel programming*** *using only the* ***directive parallel-for with appropriate clauses.***

Submission:

- *Completed project*
- *Collect timing statistics for large values of n. Examples n: 10^3, 10^4, 10^5, 10^6, 10^7 and 10^8*
- *Provide a plot of the* ***execution time vs. n*** *with a trend line and equation.*
- *Provide a plot of the* ***speedup vs. n****.*
- *Identify and explain why an implementation (Parts 1 or 2) has the lowest performance.*