In task 3, we constructed a least squares fit for the Monte Carlo numerical integration error in log-log scale. Below is a brief description about constructing least squares fits using both, linear scale and log-log scale.

Linear least squares regression

Suppose that we wish to fit a linear function $y = a + bx$ to some data $(x_1, y_1), \ldots, (x_n, y_n)$. This yields a (typically overdetermined) linear system of equations

$$\begin{cases} y_1 = a + bx_1 \\ y_2 = a + bx_2 \\ \quad\vdots \\ y_n = a + bx_n \end{cases} \iff \underbrace{\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}}_{=:y} = \underbrace{\begin{bmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}}_{=:A} \begin{bmatrix} a \\ b \end{bmatrix}.$$

We can find a linear fit $y = a + bx$ satisfying

$$\min_{a,b \in \mathbb{R}} \sum_{i=1}^{n} (y_i - a - bx_i)^2$$

by solving a and b from the normal equation

$$A^T A \begin{bmatrix} a \\ b \end{bmatrix} = A^T y.$$

Least squares approximation for power functions

Suppose that we wish to fit $y = ax^b$ to some data $(x_1, y_1), \ldots, (x_n, y_n)$.
The "standard" way is to linearize the model by taking the logarithm on both sides:

$$\begin{cases} y_1 = a x_1^b \\ y_2 = a x_2^b \\ \quad\vdots \\ y_n = a x_n^b \end{cases} \iff \begin{cases} \log y_1 = \log a + b \log x_1 \\ \log y_2 = \log a + b \log x_2 \\ \quad\vdots \\ \log y_n = \log a + b \log x_n \end{cases}$$

This can now be expressed as the system

$$\begin{bmatrix} \log y_1 \\ \vdots \\ \log y_n \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & \log x_1 \\ \vdots & \vdots \\ 1 & \log x_n \end{bmatrix}}_{=:A} \begin{bmatrix} \log a \\ b \end{bmatrix}$$

and we can find the fit by using ordinary linear least squares, i.e., solving the normal equation

$$A^T A \begin{bmatrix} \log a \\ b \end{bmatrix} = A^T \begin{bmatrix} \log y_1 \\ \vdots \\ \log y_n \end{bmatrix}.$$

On error plots

- If the data satisfies a power law $y = ax^b$, then it is typical to represent it using a log-log plot. The rate of decay b will then be the slope of the data represented in log-log scale.
- If the data satisfies an exponential law $y = Ce^{\alpha x}$, then it is typical to represent it using a semi-log plot (i.e., the x-axis is linear, but the y-axis is logarithmic). The slope of a data set which appears linear in semi-log scale is related to α.