Week 3: Problem Solving Tips
Mathematics for Data Science - 1

1 Different forms of equations of a straight line

Forms of equation of straight line	**Representation**
General form	$ax + by + c = 0$
Slope-point form	$(y - y_0) = m(x - x_0)$
Slope-intercept form (y-intercept)	$y = mx + c$
Slope-intercept form (x-intercept)	$y = m(x - d)$
Intercept form	$\frac{x}{a} + \frac{y}{b} = 1$
Two-point form	$(y - y_1) = \frac{(y_2 - y_1)}{(x_2 - x_1)}(x - x_1)$

- Given two straight lines $a_1x + b_1y + c_1 = 0$ and $a_2x + b_2y + c_2 = 0$, where $b_1, b_2 \neq 0$,
  - The lines are parallel to each other, if $a_1 \times b_2 = a_2 \times b_1$.
  - The lines are perpendicular to each other, if $a_1 \times a_2 = -b_1 \times b_2$.
- Distance of a point (x_1, y_1) from the straight line $ax + by + c = 0$ is

$$\frac{|ax_1 + by_1 + c|}{\sqrt{a^2 + b^2}}$$

- Distance between two parallel lines $ax + by + c_1 = 0$ and $ax + by + c_2 = 0$ is

$$\frac{|c_1 - c_2|}{\sqrt{a^2 + b^2}}$$

2 Sum Squared Error (SSE)

- SSE is the sum of the squares of the deviations of the predicted linear model from the actual data set.
- Numerically , if we are given a set of n points (x_i, y_i), such that , $i = 1, 2, 3, ..., n$ and we have a line of fit $y = mx + c$ for fitting these points, then the SSE will be calculated as

$$SSE = \sum_{i=1}^{n} (y_i - mx_i - c)^2$$

- To find the best fitting line (among the given lines) for the given data set, by minimizing SSE.