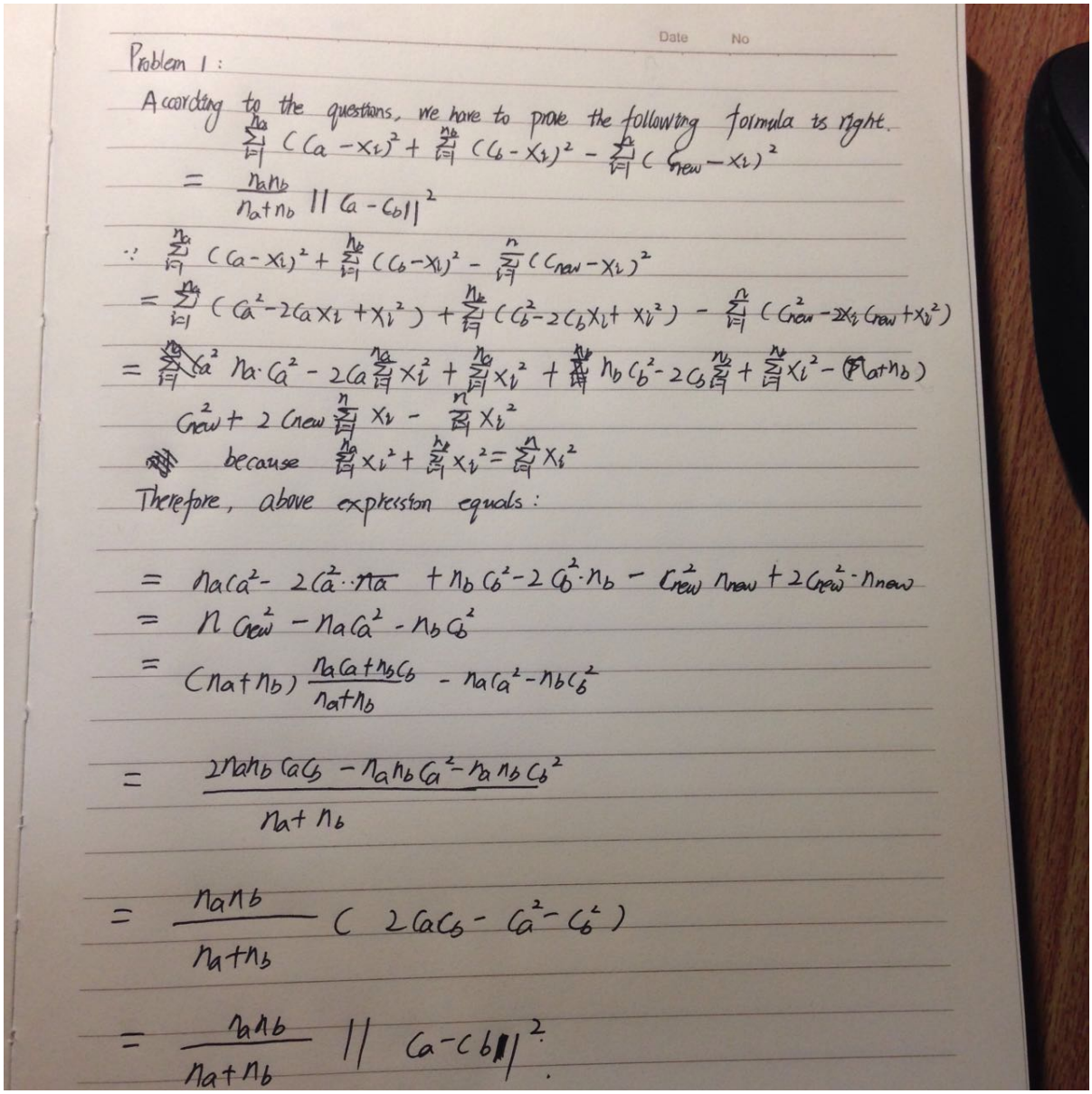
Problem 1



Problem 1:

According to the questions, we have to prove the following formula is right.

$$\sum_{i=1}^{n_a}(C_a - x_i)^2 + \sum_{i=1}^{n_b}(C_b - x_i)^2 - \sum_{i=1}^{n}(C_{new} - x_i)^2 = \frac{n_a n_b}{n_a + n_b}\|C_a - C_b\|^2$$

$$\because \sum_{i=1}^{n_a}(C_a - x_i)^2 + \sum_{i=1}^{n_b}(C_b - x_i)^2 - \sum_{i=1}^{n}(C_{new} - x_i)^2$$

$$= \sum_{i=1}^{n_a}(C_a^2 - 2C_a x_i + x_i^2) + \sum_{i=1}^{n_b}(C_b^2 - 2C_b x_i + x_i^2) - \sum_{i=1}^{n}(C_{new}^2 - 2x_i C_{new} + x_i^2)$$

$$= \sum_{i=1}^{n_a} C_a^2 \; n_a C_a^2 - 2C_a\sum_{i=1}^{n_a} x_i^2 + \sum_{i=1}^{n_a} x_i^2 + \sum_{i=1}^{n_b} n_b C_b^2 - 2C_b\sum_{i=1}^{n_b} + \sum_{i=1}^{n_b} x_i^2 - (n_a + n_b)$$

$$C_{new}^2 + 2C_{new}\sum_{i=1}^{n} x_i - \sum_{i=1}^{n} x_i^2$$

because $\sum_{i=1}^{n_a} x_i^2 + \sum_{i=1}^{n_b} x_i^2 = \sum_{i=1}^{n} x_i^2$

Therefore, above expression equals:

$$= n_a C_a^2 - 2C_a^2 \cdot n_a + n_b C_b^2 - 2C_b^2 \cdot n_b - C_{new}^2 n_{new} + 2C_{new}^2 \cdot n_{new}$$

$$= n C_{new}^2 - n_a C_a^2 - n_b C_b^2$$

$$= (n_a + n_b)\frac{n_a C_a + n_b C_b}{n_a + n_b} - n_a C_a^2 - n_b C_b^2$$

$$= \frac{2n_a n_b C_a C_b - n_a n_b C_a^2 - n_a n_b C_b^2}{n_a + n_b}$$

$$= \frac{n_a n_b}{n_a + n_b}(2C_a C_b - C_a^2 - C_b^2)$$

$$= \frac{n_a n_b}{n_a + n_b}\|C_a - C_b\|^2.$$