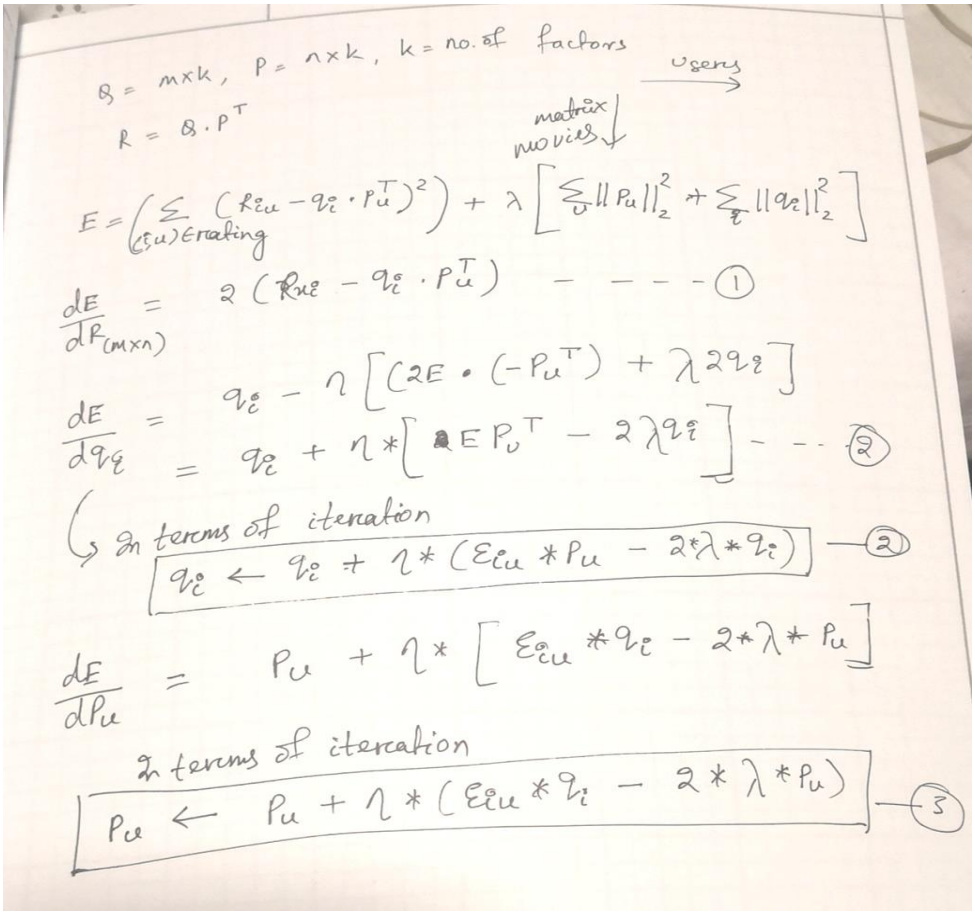
1. **The equations for *εiu*. Update equations in the Stochastic Gradient Descent algorithm**

Ans:-

The equations include the following given derivative steps:-

$$\varepsilon_{xi} = 2(r_{xi} - q_i . p_x)$$
$$q_i \leftarrow q_i + \mu_1 (\varepsilon_{xi}\, p_x - 2\lambda_2\, q_i)$$
$$p_x \leftarrow p_x + \mu_2 (\varepsilon_{xi}\, q_i - 2\lambda_1\, p_x)$$



2. **What was the lowest error you got? What was the value of *η*?**

1. **For learning_rate = 0.01 #*η*,**
40th iteration error =80622 #lowest

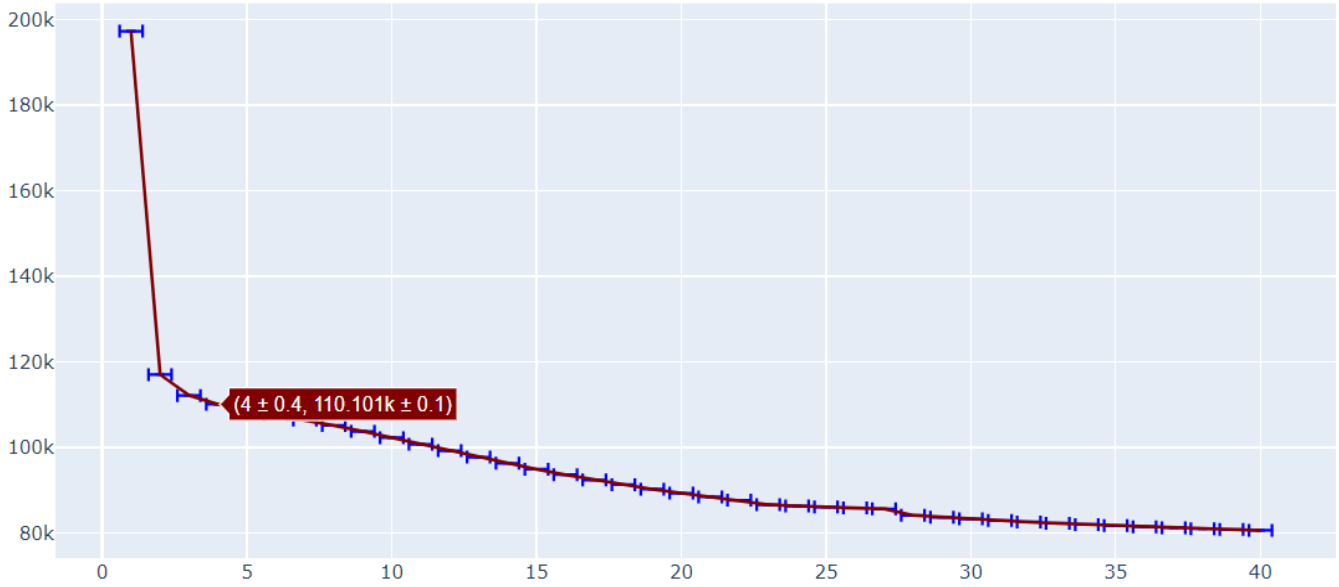
2. For learning_rate = 0.02 #*η*, 40th error =84102
3. For learning_rate = 0.03 #*η*, 40th error =89872
4. For learning_rate = 0.03 #*η*, 40th error =120098
5. For learning_rate = 0.001 #*η*, 40th error =88102

3. **For the best η, plot of E vs. number of iterations. Make sure your graph has a y-axis so that we can read the value of E.**

For learning_rate = 0.01 #η, 40th iteration **error =80622** #lowest

X-axis = number of Interation and

Y = Errors after execution



Inference:

1. In 1st iteration the error = 197282.6
2. 40th iteration Error = 80622.64
3. With increasing iteration the error value is decreasing
4. **Iteration start time -** `08:46:47`
   40th iteration completion time:- 16:11:49
   Time taken = 07:65:0

As weights are updated after each iteration within individual rows so the error optimizati on happens. I think tuning the hyper parameter like “learning rate is quite a challenge and time taking ”