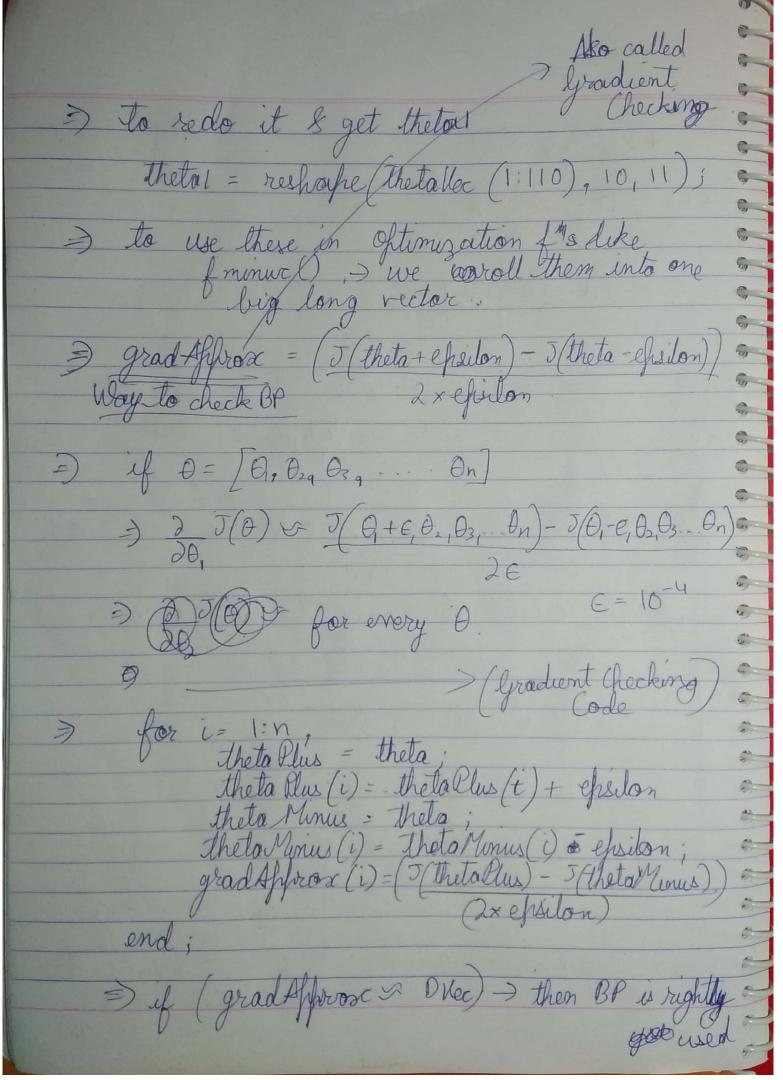
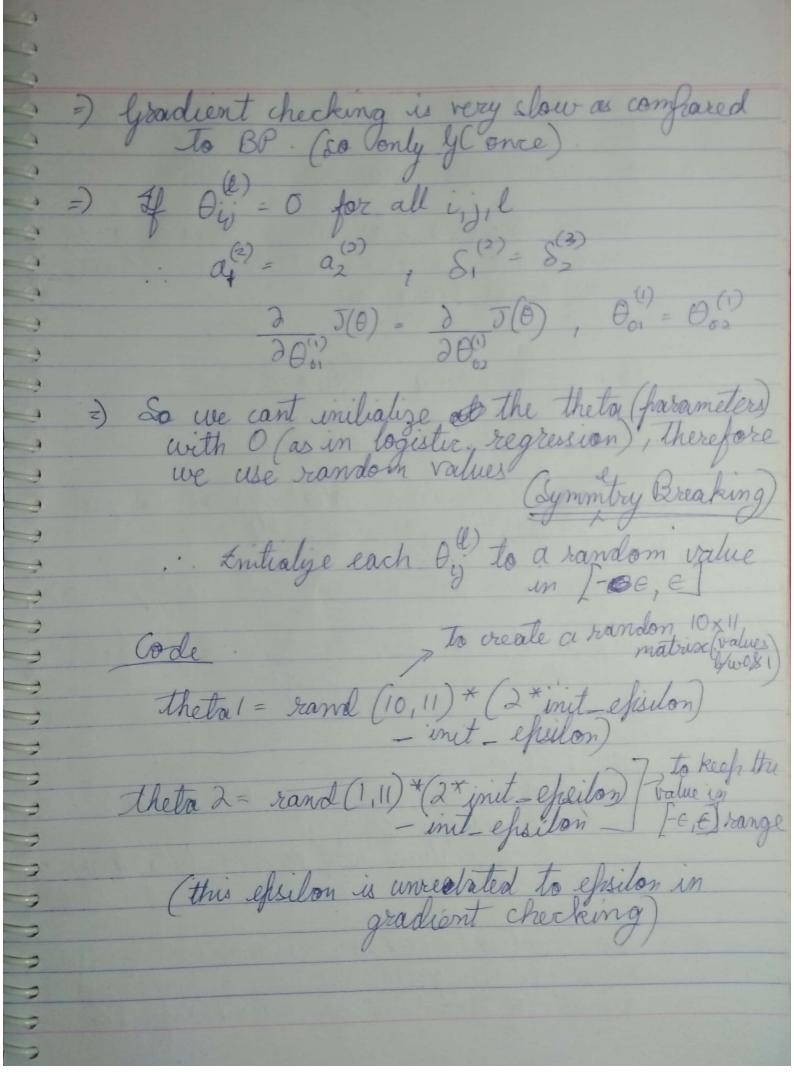
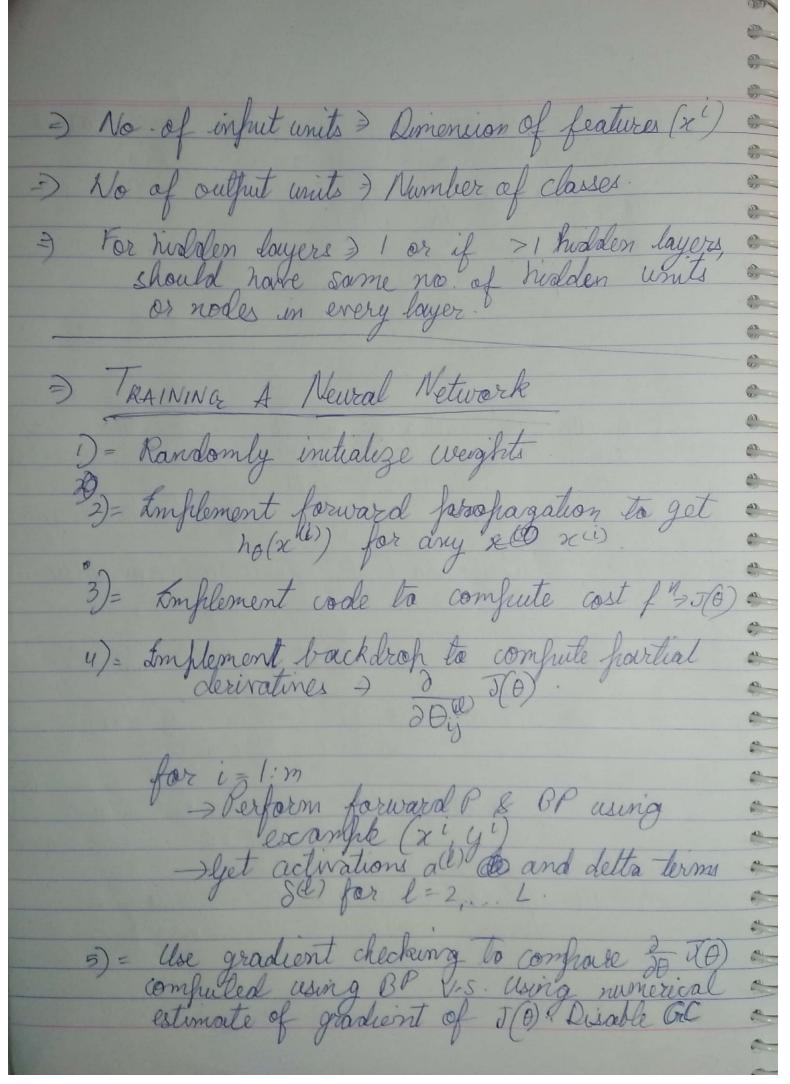


Westorized > \( \( \lambda \text{(1)} = \( \lambda \text{(1)} + \( \lambda \text{(1+1)}, \( \alpha \text{(a} \text{(a} \text{)} \)^T Dij = 1 Dij + 20ij if j 7 0  $=) Oij := \frac{1}{m} \Delta ij \qquad if j = 0$ (b) = Dig deltayed = theta-grad = (m) \*delta;  $\Rightarrow$  Formally,  $\delta^{(e)} = \frac{\partial}{\partial z^{(e)}} cost(i)$  (for  $j \ge 0$ ) where cost(i) = y (i) log ho(x(i) + (1-y(i)) log(ino (x(i))) =)  $S_{2}^{(2)} = \theta_{12}^{(2)} S_{1}^{(3)} + \theta_{22}^{(2)} S_{2}^{(3)}$ S(2) = SO O(3) S(4) ) for example; if S=10, S=10, S=1 ) o'> TR'OX' 02) TR'OX' 0(3)=1X11 ) (0) ) ROX' 00) TROX' (0)=1X11 theta Ver = [thota (:); theta2(:); theta3(:)];
converts these rectors in one long rector







6)= Use gradient descent or advanced oftonization method with BP to try to minumize of (0) as a fin of 0 =) For NN, J(0) is non-convex.