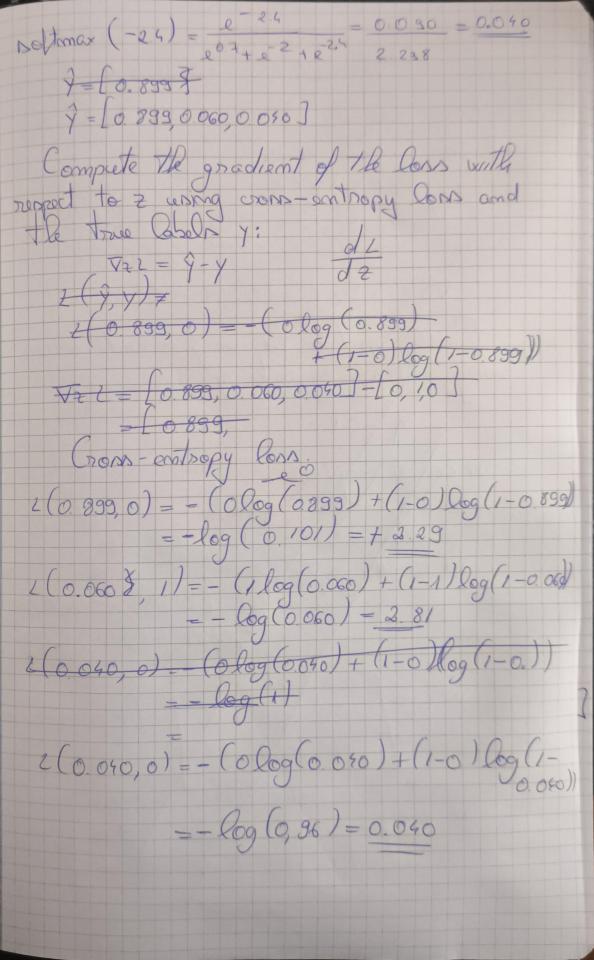
Lunga Andrei - Sebarsten Begin by computing the linear combination 2 Par ach clars 21 = XIII · XI + XII2 · X2 + XII3 · X3 + 51 21=03.1+0.1.3+(-2).0+0.1 21=03+0.3+0.1=0.7 22 = X121 · X + X122 · X2 + X123 · X3 + 52. 22=-0.6.1+(-0.5) 3+2 0+01 72=-0.6-1,5+0.1=-2.0 23 = X/31 · X1 + X/32 · X2 + W33 · X3 + 63 23 = -1.1+(-0,5),3+0,1.0+0,1 23 = -1 -1.5 +0.1 = -25 7 = [0,7, -2.0, -2.4] Apply the softmax function to get the predicted prosessitation ?: (f) x 2mt/301 = ( = 2.0/3 = 0.899 0.135 notimes (-2.0) = 2-2 2.238 = 0.060



Y= [0.899, 0.060, 0.050]-[0,6] V= [0.899, -0.94, 0.050] the gradients with respect compute and bjarren 5: 899, -0.94, 0.040]. [ 130 = (0.899, -2.28-2,82,0 V52 = V2 ( = (0.899, -0.94, 0.090. using a learning Date weights and signer

bc- 5- gx32 7732 = 0.1. [ using a learning rate m:  $m \cdot \sqrt{x/2} = 0.1 \cdot [0.899, -2.82, 0]$   $m \cdot \sqrt{x/2} = [0.0899, -0.282, 0]$  $\times 1 \leftarrow 1 \leftarrow 0.3 \quad 0.1 \leftarrow 2 \quad 7 \leftarrow 10.0899, -0.282, 0 \quad 7 \leftarrow -1 \leftarrow 0.5 \quad 0.1 \quad 9 \leftarrow 0.2101 \quad 0.382 \quad -2 \quad 7 \leftarrow -0.6899 \quad -0.218 \quad 2 \quad 7 \leftarrow -1.0899 \quad -0.218 \quad 2 \leftarrow -1.0899 \quad -0.218 \quad 0.1 \quad 9 \leftarrow -1.0899 \quad -0.218 \quad -0.218 \quad 0.1 \quad 9 \leftarrow -1.0899 \quad -0.218 \quad -0$ m V52 C 0.1. [0,899, -0.94, 0.040] m V52 <- 10.0899, -0.084, 0.004] b <- [0.1,0.1,0.1] - [0.0899, -0.097,0.007] 5 <- (0,0.0101,0.194,0.096)