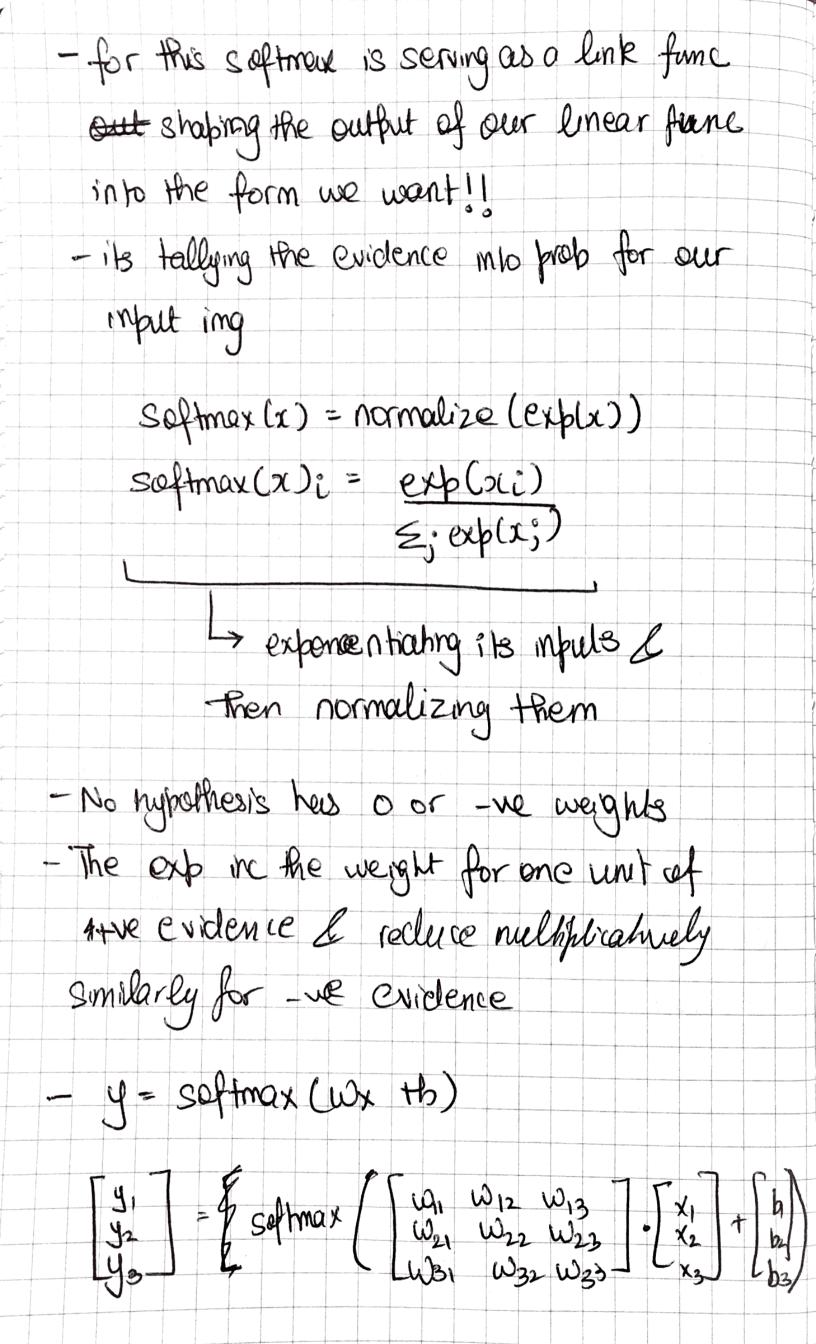
Sunday, 25,2016 - Since softmax turns excidence into prob obst its used as the last layer even in complicated models. - 8000 i = class j = Summing index over all classes oc = input mage



Regression
- Numby comp does computations in C & Hrus cares
about data types. There is overhead in
swortching blu the two though.
- Tensor flow does heavy lifting outside
Python. You describe the graph in by 2
execute outsid of py.
Farning
- Define what it means for the model to be bad?
Gost fine, loss func etc
-minimize the error
- Cross - entrepy
$Hy'(y) = - \ge y'_{i} \log (y_{i})$
y = predicted prob dist
y'= true dist
measure & how inefficient our prediction
is
->

Back Pro	pogation	1			
-> Rove	rse mod	te differ	enhat	ion	
- get	-> Haw	esvi B =	the	derivale	of the
Output	with	respect	ho e	wery sm	ngle node
				Mularal 10	