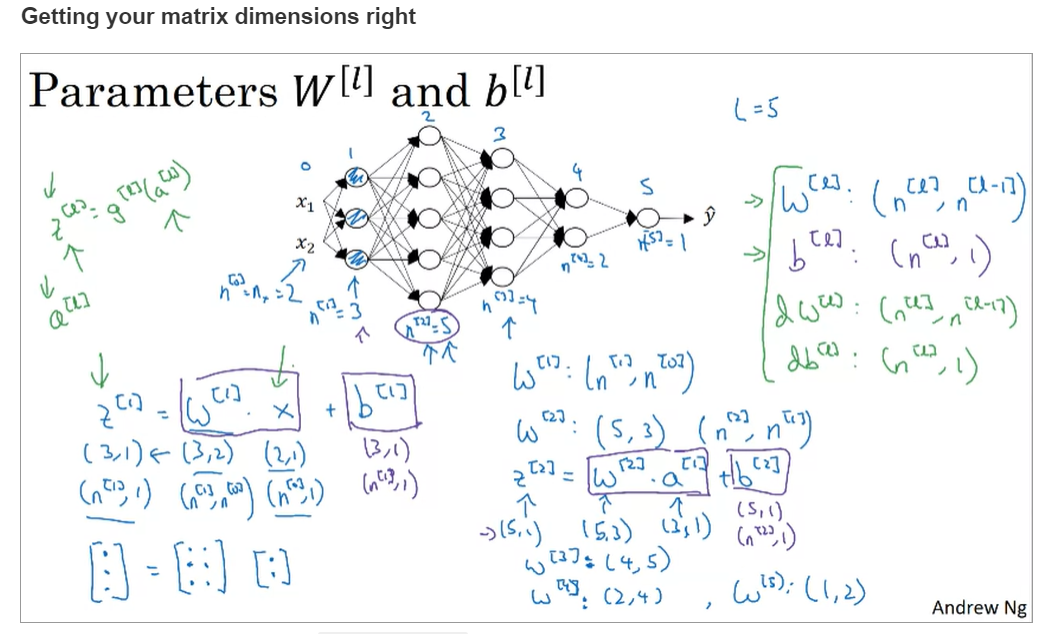
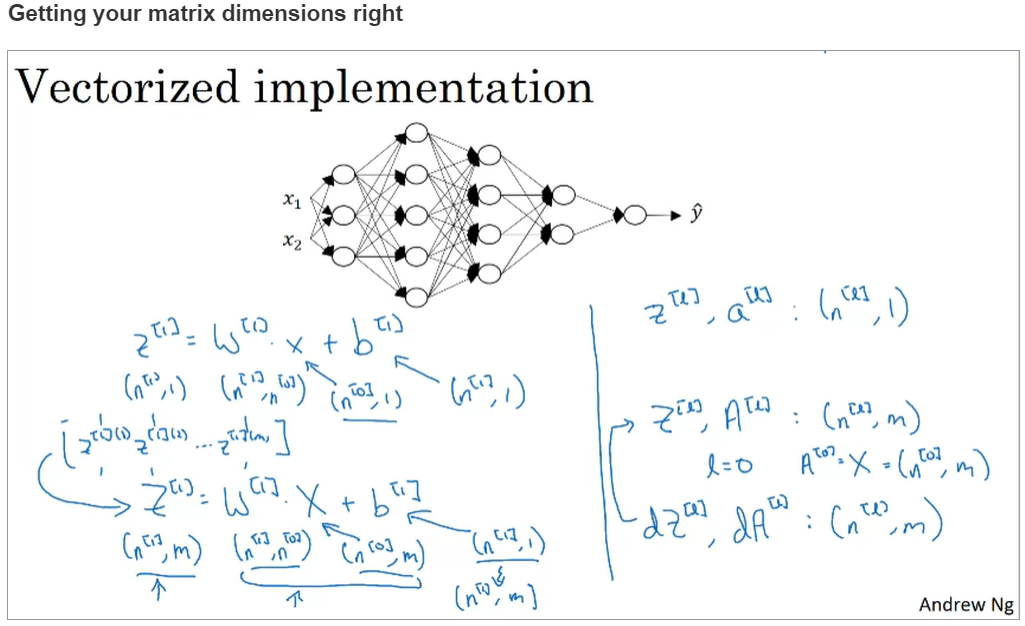
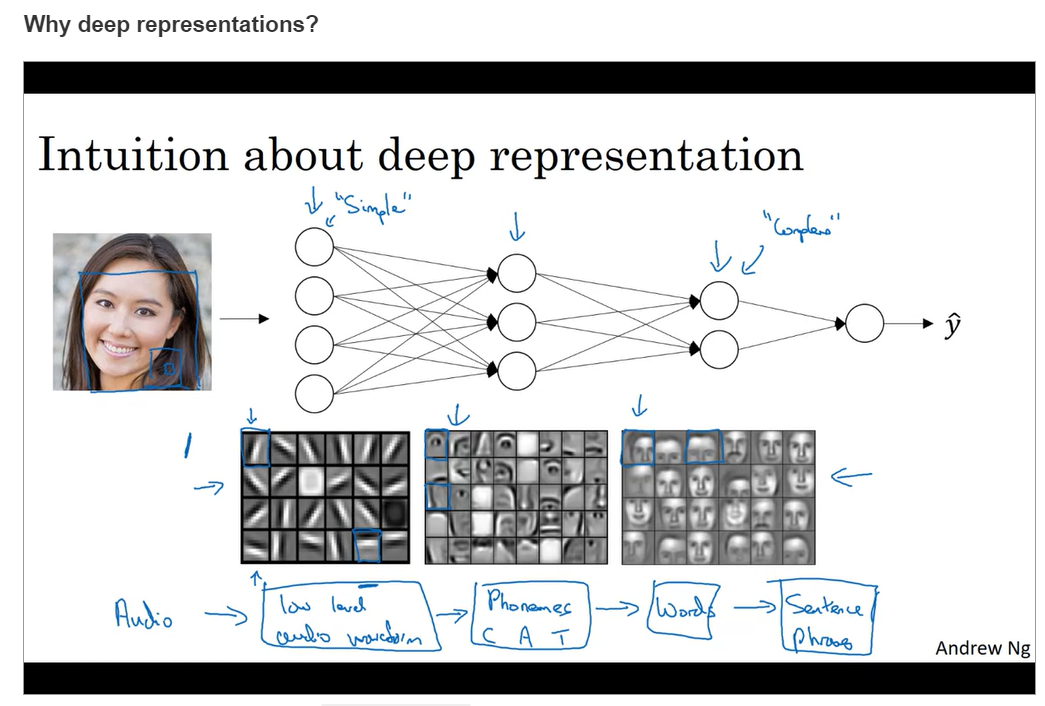


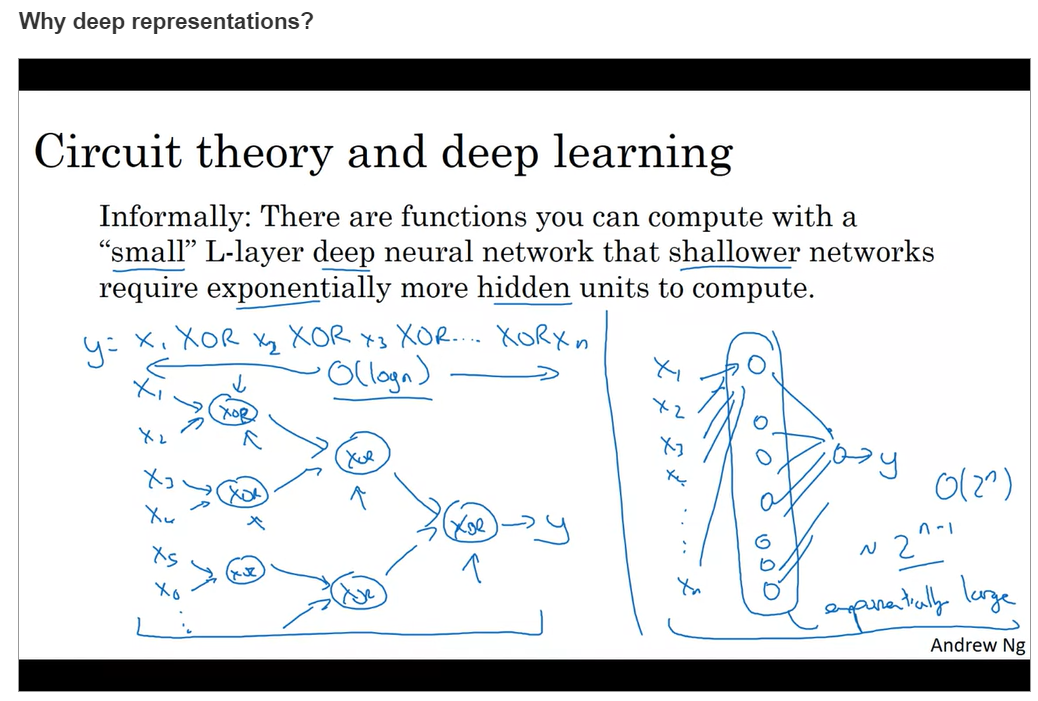
Check dimensions which makes debugging easy

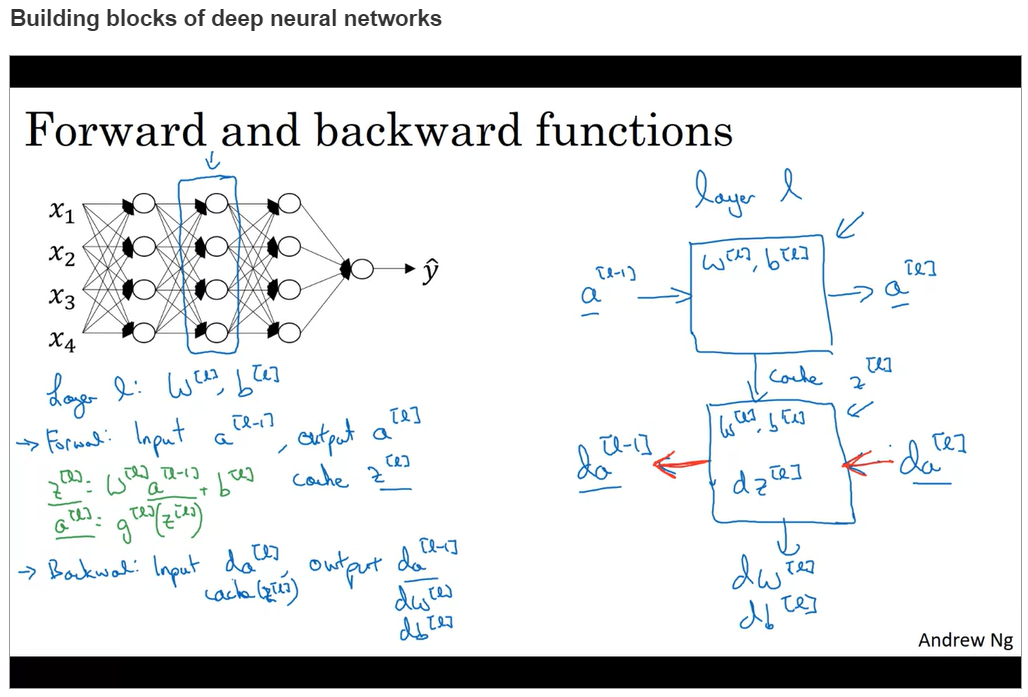


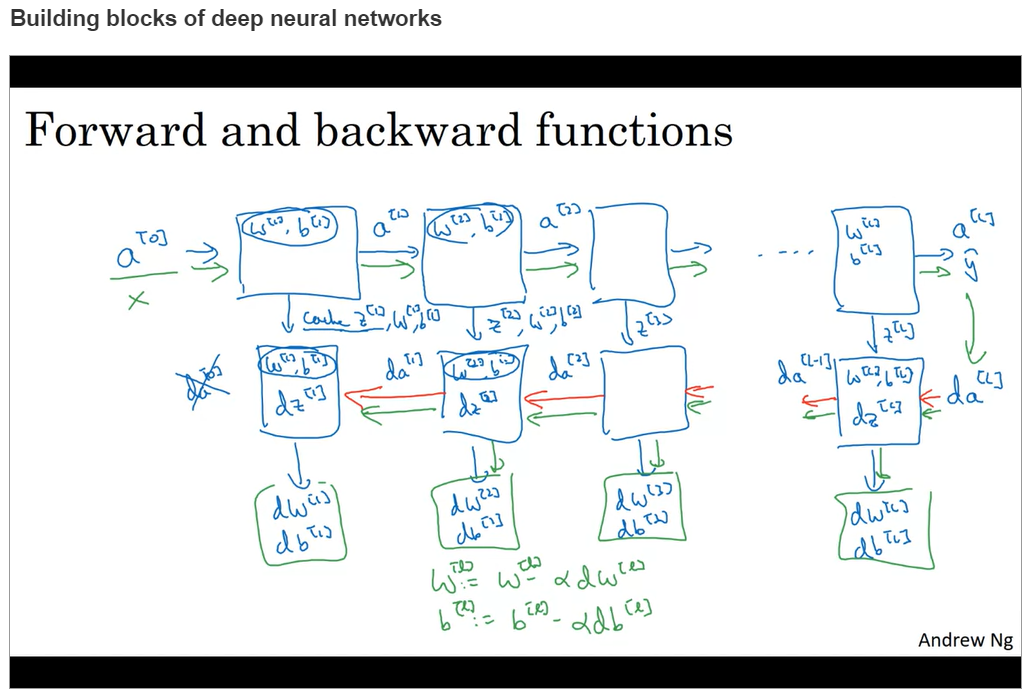
**z[l] = g[l](a[l]) 🡪 a[l] = g[l](z[l])**

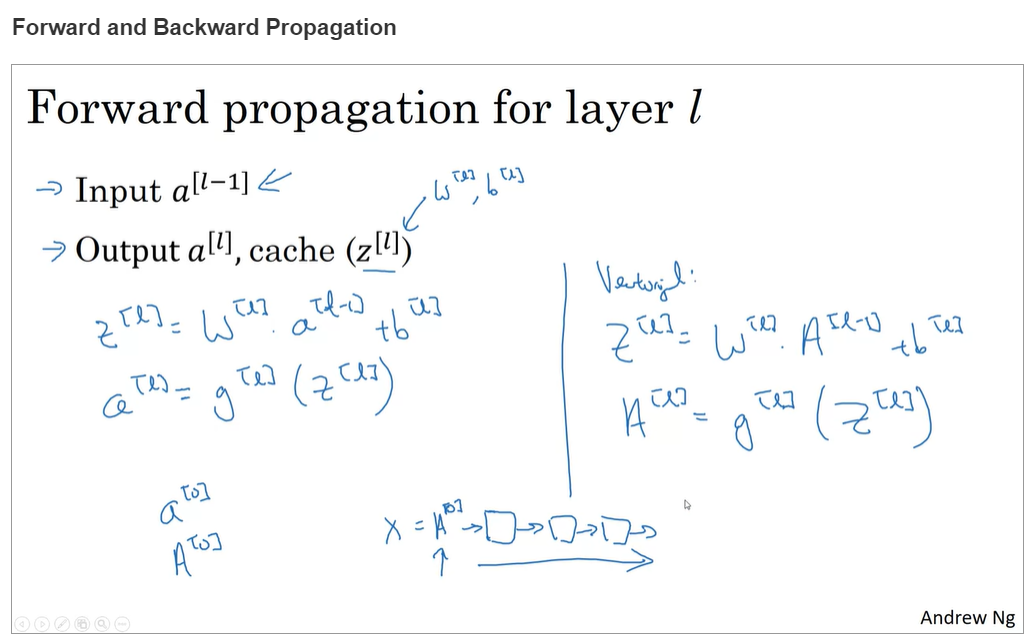


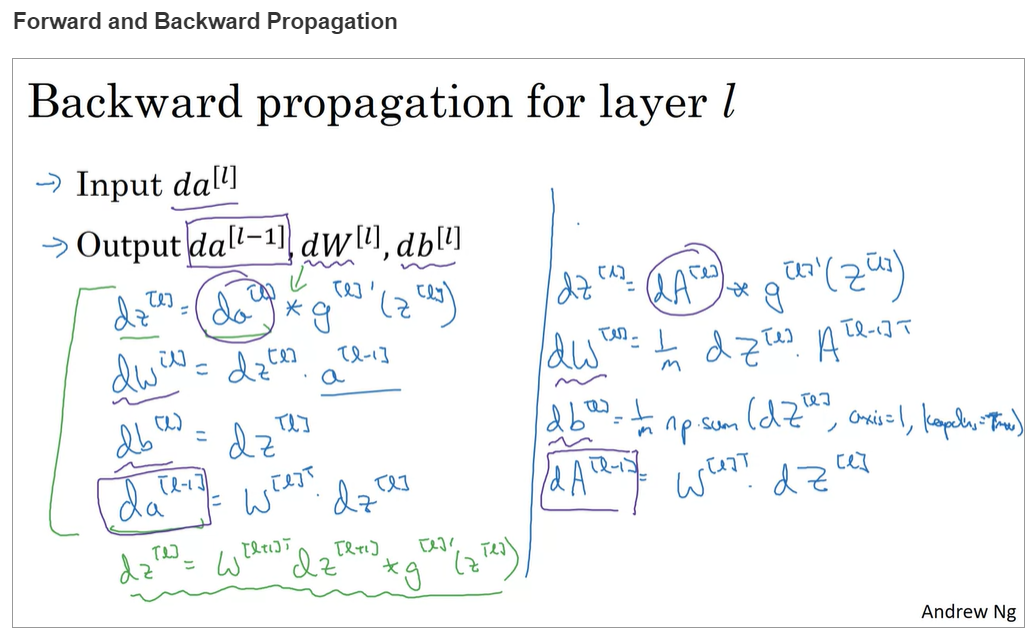




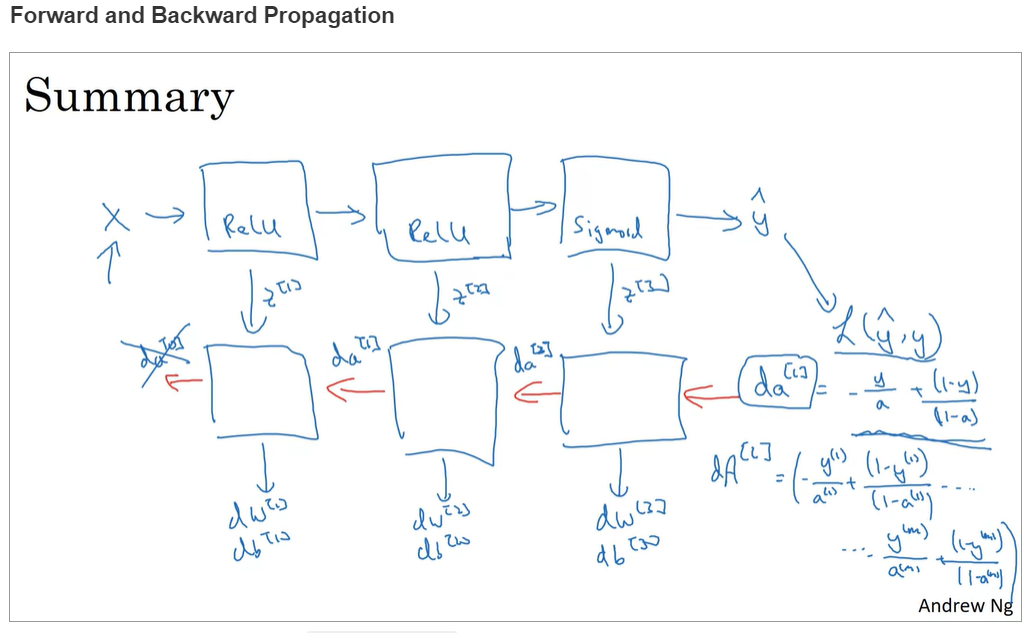


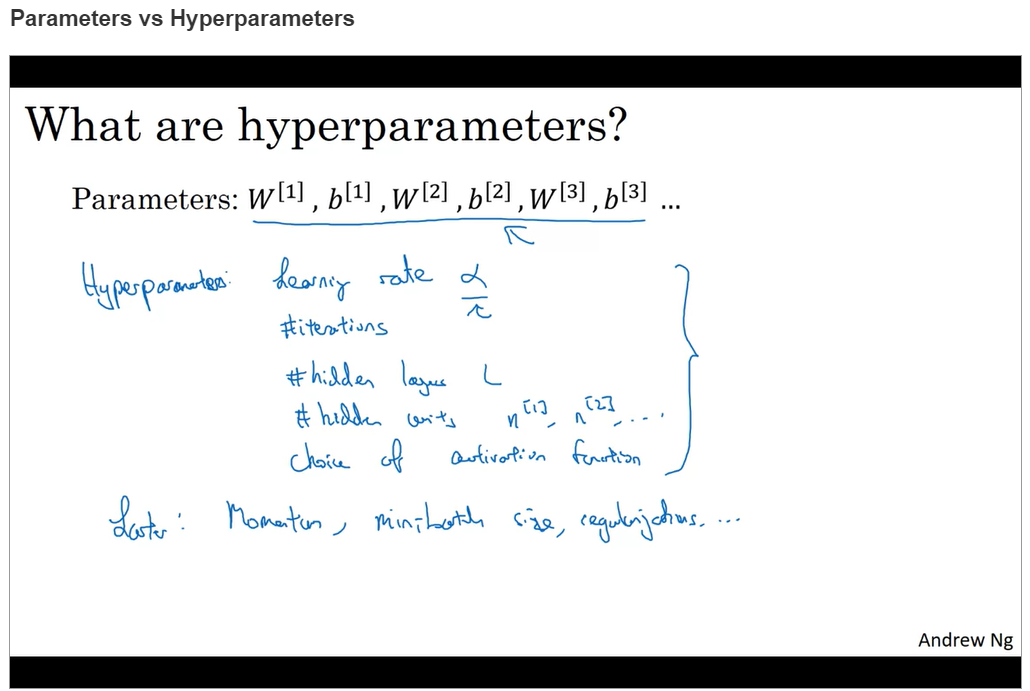


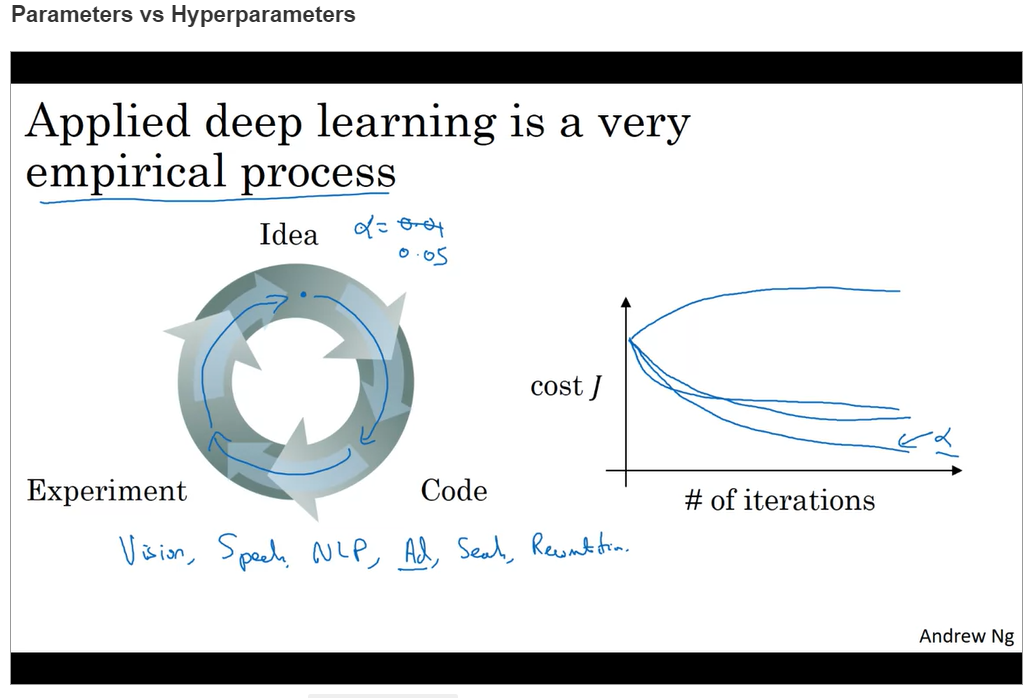


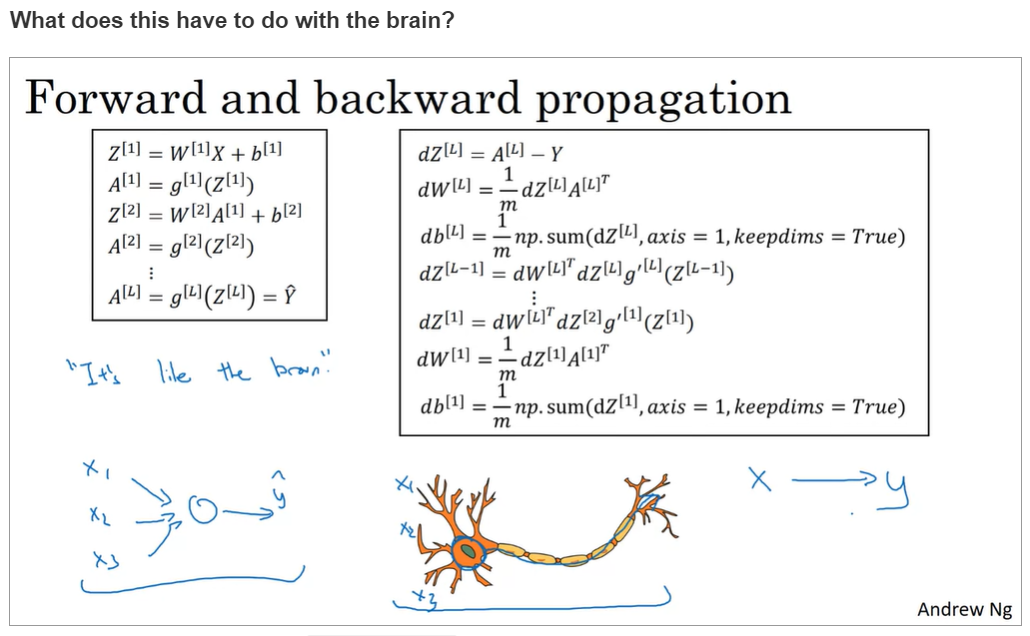


**dw[l] = dz[l].a[l-1] 🡪 dw[l] = dz[l].a[l-1]T**









***dZ*[*L*]=*A*[*L*]−*Y***

**dW^{[L]} = \frac{1}{m}dZ^{[L]}{A^{[L-1]^{T}}}*dW*[*L*]=*m*1​*dZ*[*L*]*A*[*L*−1]*T***

**db^{[L]} = \frac{1}{m} np.sum(dZ^{[L]}, axis=1, keepdims=True)*db*[*L*]=*m*1​*np*.*sum*(*dZ*[*L*],*axis*=1,*keepdims*=*True*)**

**dZ^{[L-1]} = W^{[L]^{T}}dZ^{[L]} \* g'^{[L-1]}(Z^{[L-1]})*dZ*[*L*−1]=*W*[*L*]*TdZ*[*L*]∗*g*′[*L*−1](*Z*[*L*−1])**

**Note that \* denotes element-wise multiplication)**

**\vdots⋮**

**dZ^{[1]} = W^{[2]}dZ^{[2]} \* g'^{[1]}(Z^{[1]})*dZ*[1]=*W*[2]*dZ*[2]∗*g*′[1](*Z*[1])**

**dW^{[1]} = \frac{1}{m}dZ^{[1]}A^{[0]^{T}}*dW*[1]=*m*1​*dZ*[1]*A*[0]*T***

**Note that A^{[0]^{T}}*A*[0]*T* is another way to denote the input features, which is also written as X^{T}*XT***

**db^{[1]} = \frac{1}{m} np.sum(dZ^{[1]}, axis=1, keepdims=True)*db*[1]=*m*1​*np*.*sum*(*dZ*[1],*axis*=1,*keepdims*=*True*)**