2(b) Given:

$J_{reg}(\theta) =$	$\frac{\lambda}{2}\left[\sum_{i,j} W_{ij}^{\ 2} + \sum_{i',j'} U_{i'j'}^{\ 2}\right]$	(1)
$J_{full}(\theta) =$	$J(\theta) + J_{reg}(\theta)$	(2)

$\frac{\partial J_{full}}{\partial \mathbf{U}} =$	$\frac{\partial J}{\partial \mathbf{U}} + \frac{\partial J_{reg}}{\partial \mathbf{U}}$	(3)
$\frac{\partial J}{\partial \mathbf{U}} =$	$\mathbf{h}^{\mathbf{T}}\left(\hat{\mathbf{y}} - \mathbf{y}\right)$ from 2(a)	(4)
$\frac{\partial J_{reg}}{\partial \mathbf{U}} =$	$\lambda \mathbf{U}$	
$\frac{\partial J_{full}}{\partial \mathbf{U}} =$	$\mathbf{h}^{\mathbf{T}}\left(\hat{\mathbf{y}} - \mathbf{y}\right) + \lambda \mathbf{U}$	

$\frac{\partial J_{full}}{\partial \mathbf{W}} =$	$\frac{\partial J}{\partial \mathbf{W}} + \frac{\partial J_{reg}}{\partial \mathbf{W}}$	(5)
$\frac{\partial J}{\partial \mathbf{W}} =$	$\mathbf{x}^{(t)\mathbf{T}}\left(\frac{\partial J}{\partial \mathbf{h}} \odot \tanh'(\mathbf{x}^{(t)}\mathbf{W} + \mathbf{b_1})\right)$ from 2(a)	(6)
$\frac{\partial J_{reg}}{\partial \mathbf{W}} =$	$\lambda \mathbf{W}$	(7)
$\frac{\partial J_{full}}{\partial \mathbf{W}} =$	$\mathbf{x}^{(t)\mathbf{T}}\left(\frac{\partial J}{\partial \mathbf{h}} \odot \tanh'(\mathbf{x}^{(t)}\mathbf{W} + \mathbf{b_1})\right) + \lambda \mathbf{W}$	(8)