Update Rules for hidden and output layers.
For Regression using mean square error loss.

defining variables used in the derivation:

w_{ij} = weights from input to hidden layer

w_{jk} = weight from hidden layer to output layer.

i = the number of units in the input layer

j = number of units in hidden layer

k = number of units in output layer

z = activation function

Finding w_{jk} and b_{jk}:
weight and bias from hidden to output layer first:

$$\frac{\partial J(w)}{\partial w_{jk}} = \frac{\partial \frac{1}{2} \sum (y-\hat{y})^2}{\partial w_{jk}}$$

$$= (y-\hat{y}) \frac{\partial (y-\hat{y})}{\partial w_{jk}}$$

$$= (y-\hat{y}) \frac{\partial \hat{y}}{\partial w_{jk}}$$

$$= (y-\hat{y}) \frac{\partial \hat{y}}{\partial z_k} \cdot \frac{\partial z_k}{\partial w_{jk}}$$

$$w_{jk} = (y-\hat{y}) [z_k (1-z_k)] \cdot x_j$$

$b_{jk} = (y-\hat{y})[z_k(1-z_k)]$.

Finding the weight and bias from input to hidden layer:

$$w_{ij} = \frac{\partial \frac{1}{2}(y-\hat{y})^2}{\partial w_{ij}}$$

$$= \Sigma(y-\hat{y}) \frac{\partial \hat{y}}{\partial w_{ij}}$$

$$= \Sigma(y-\hat{y}) \frac{\partial \hat{y}}{\partial z_k} \cdot \frac{\partial z_k}{\partial x_j} \cdot \frac{\partial x_j}{\partial z_j} \cdot \frac{\partial z_j}{\partial w_{ij}}$$

$$w_{ij} = \Sigma(y-\hat{y})[z_k(1-z_k)] \cdot w_{jk} \cdot [z_j(1-z_j)] \cdot x_1$$

$$b_{ij} = \Sigma(y-\hat{y})[z_k(1-z_k)] \cdot w_{jk}[z_j(1-z_j)].$$

Regression uses mean square error loss against the log loss used by binary classification.
The update rule for binary uses sigmoid's derivative as its output activation function; but for regression the derivative of the identity function is used in its update rule.

You cannot use the same activation functions for the output layer for both binary and regression