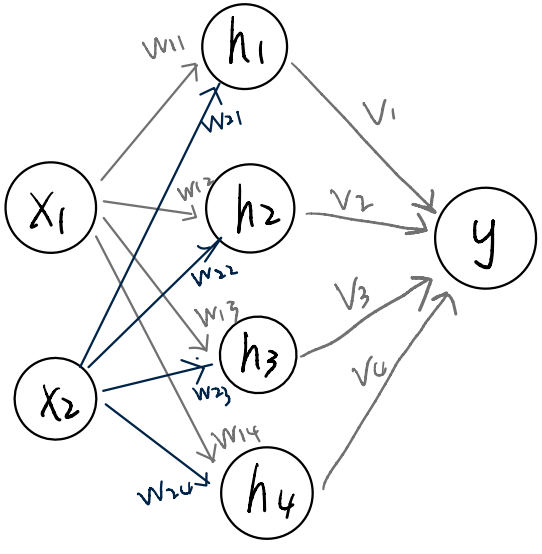
1 Feedforward: Building a ReLU neural net- work

1.



2. Input $\rightarrow$ hidden layers:

$$a_i = w_{1i} x_1 + w_{2i} x_2 + b_i$$

$$h_i = \max(0, a_i)$$

hidden layers:

$$y = v_1 h_1 + v_2 h_2 + v_3 h_3 + C$$

2 Gradient Descent

1. $$\frac{\partial f(x,y)}{\partial x} = -3x^2 - 100(2y^2 - x)$$

$$\frac{\partial f(x,y)}{\partial y} = 800y^3 - 400xy$$

3 Backprop

1. $$\frac{dL}{d\hat{y}} = -\left(\frac{y_i}{\hat{y}_i} - \frac{1-y_i}{1-\hat{y}_i}\right)$$

$$\frac{\partial L}{\partial v_i} = \frac{\partial L}{d\hat{y}} h_i = -\left(\frac{y_i}{\hat{y}_i} - \frac{1-y_i}{1-\hat{y}_i}\right) h_i$$

$$\frac{\partial L}{\partial c} = \frac{\partial L}{\partial \hat{y}} = -\left(\frac{y_i}{\hat{y}_i} - \frac{1-y_i}{1-\hat{y}_i}\right) \quad (h_i > 0)$$

$$\frac{\partial L}{\partial w_{1i}} = \frac{\partial L}{\partial \hat{y}} \cdot v_i x_1 = -\left(\frac{y_i}{\hat{y}_i} - \frac{1-y_i}{1-\hat{y}_i}\right) v_1 x_1$$

$$\frac{\partial L}{\partial w_{2i}} = \frac{\partial L}{\partial \hat{y}} v_2 x_2 = -\left(\frac{y_i}{\hat{y}_i} - \frac{1-y_i}{1-\hat{y}_i}\right) v_2 x_2 \quad (h_i > 0)$$

$$\frac{\partial L}{\partial b_i} = \frac{\partial L}{\partial \hat{y}} v_2$$