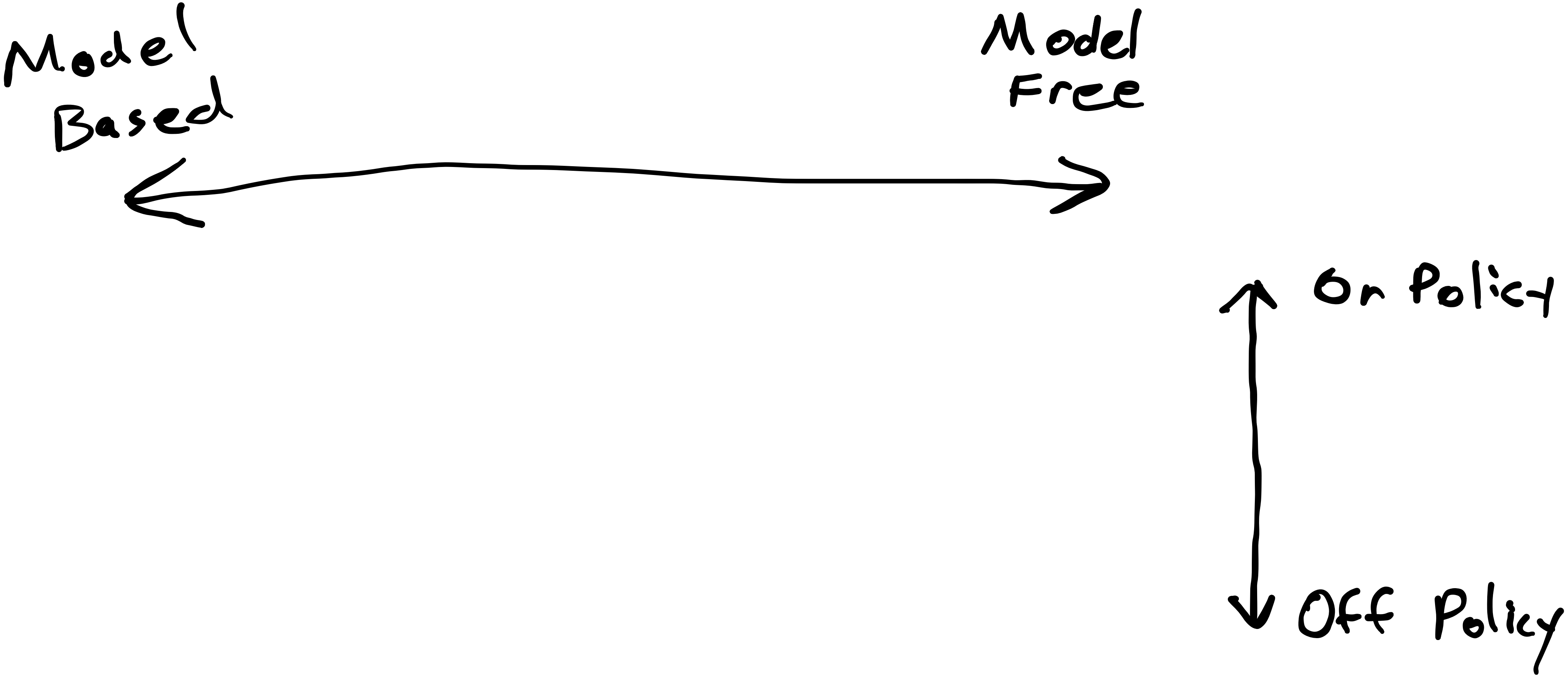
Last Time

Last Time

Model Based $\longleftrightarrow$ Model Free

Last Time



Last Time

Model Based

Model Free

On Policy

Off Policy

MLMBTRL
(learn T, R)

Last Time

Model Based

Model Free

learn Q

learn π

On Policy

Off Policy

MLMBTRL
(learn T,R)

Last Time

Model Based

Model Free

learn Q

learn π

Policy Gradient

On Policy

Off Policy

MLMBRL
(learn T, R)

Last Time

Model Based

Model Free

learn Q

SARSA

learn π

Policy Gradient

On Policy

Off Policy

MLMBTRL
(learn T,R)

Last Time

Model Based

Model Free

learn Q

SARSA

Q-learning

learn π

Policy Gradient

On Policy

Off Policy

MLMBTRL

(learn T,R)

This Time

Challenges in Reinforcement Learning:

- Exploration vs Exploitation
- Credit Assignment
- Generalization

Function Approximation

Function Approximation

$$y \approx f_\theta(x)$$

Function Approximation

input

$$y \approx f_\theta(x)$$

Function Approximation

input

$$y \approx f_\theta(x)$$

output

Function Approximation

input

$$y \approx f_\theta(x)$$

output

parameters

Function Approximation

input

$$y \approx f_\theta(x)$$

output parameters

Previously, Linear:

$$f_\theta(x) = \theta^\top \beta(x)$$

Function Approximation

input

$$y \approx f_\theta(x)$$

output parameters

Previously, Linear:

$$f_\theta(x) = \theta^\top \beta(x)$$

weights features

Function Approximation

input

$$y \approx f_\theta(x)$$

output parameters

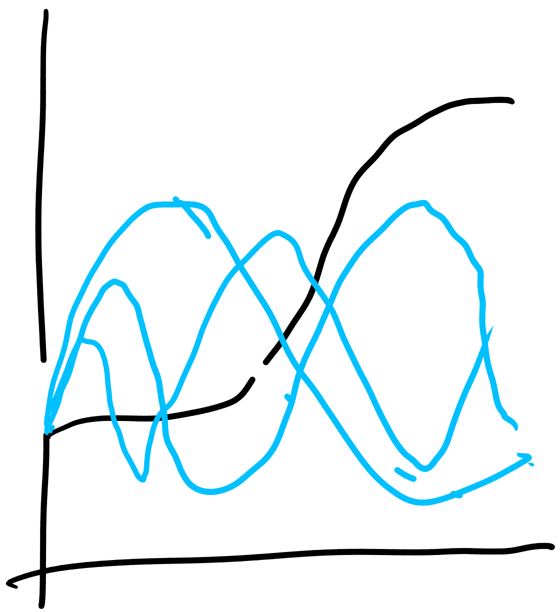
Previously, Linear:

$$\begin{bmatrix} \sin \pi x \\ \sin 2\pi x \\ \sin 3\pi x \end{bmatrix}$$

$$f_\theta(x) = \theta^\top \beta(x)$$

weights features

e.g. $\beta_i(x) = \sin(i\,\pi\,x)$



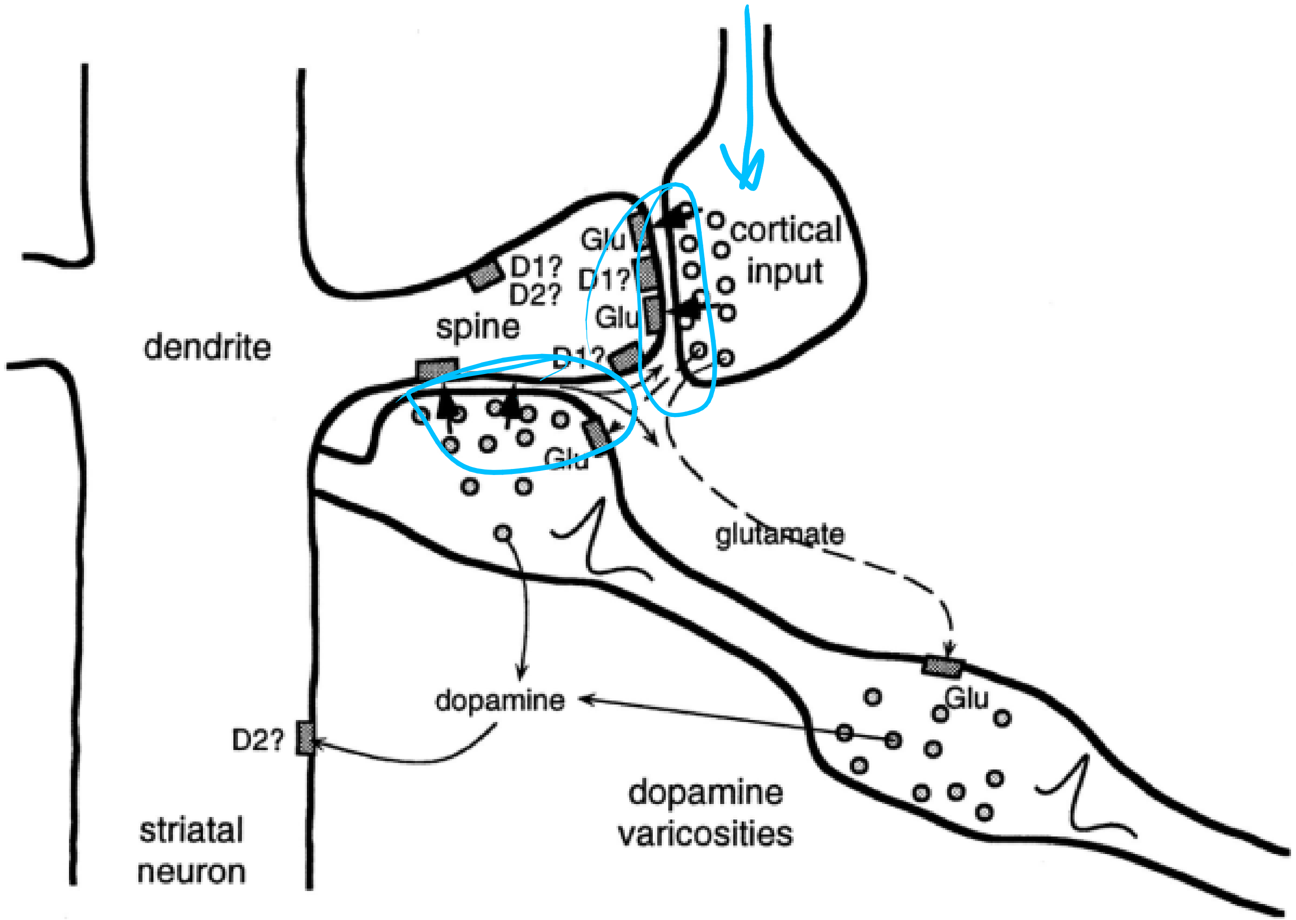
Neural Network

Neural Network

$$h(x) = \sigma(Wx + b)$$

Neural Network

$$h(x) = \sigma(Wx + b)$$



Neural Network

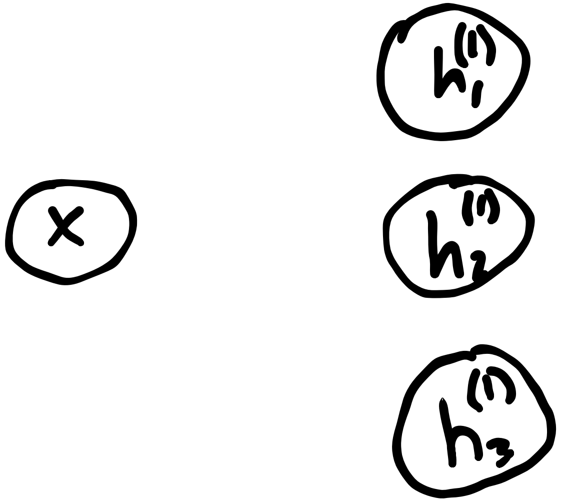
$$h(x) = \sigma(Wx + b)$$

$$f_\theta(x) = h^{(2)}\left(h^{(1)}(x)\right)$$

Neural Network

$$h(x) = \sigma(Wx + b)$$

$$f_\theta(x) = h^{(2)}\left(h^{(1)}(x)\right) = \sigma^{(2)}\left(W^{(2)}\sigma^{(1)}\left(W^{(1)}x + b^{(1)}\right) + b^{(2)}\right)$$

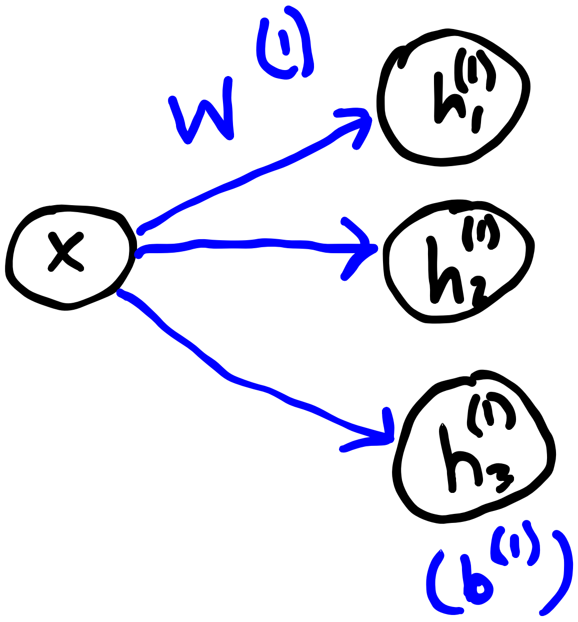


Neural Network

$$h(x) = \sigma(Wx + b)$$

$$f_\theta(x) = h^{(2)}\left(h^{(1)}(x)\right) = \sigma^{(2)}\left(W^{(2)}\sigma^{(1)}\left(W^{(1)}x + b^{(1)}\right) + b^{(2)}\right)$$

$$\theta = (W^{(1)}, b^{(1)}, W^{(2)}, b^{(2)})$$

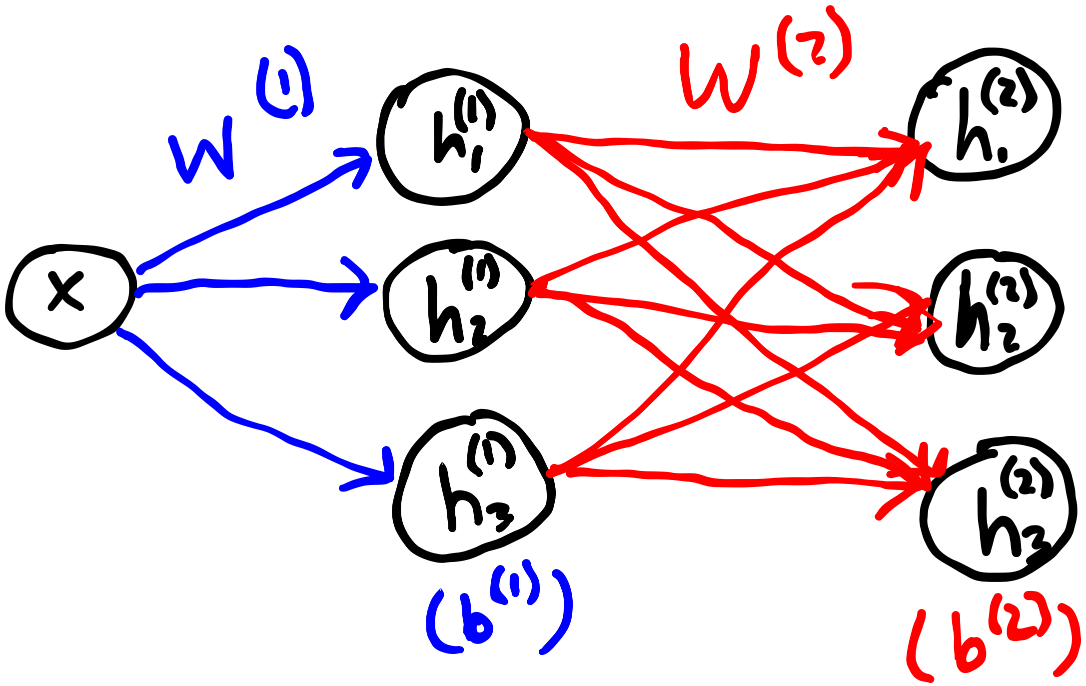


Neural Network

$$h(x) = \sigma(Wx + b)$$

$$f_\theta(x) = h^{(2)}\left(h^{(1)}(x)\right) = \sigma^{(2)}\left(W^{(2)}\sigma^{(1)}\left(W^{(1)}x + b^{(1)}\right) + b^{(2)}\right)$$

$$\theta = (W^{(1)}, b^{(1)}, W^{(2)}, b^{(2)})$$



Neural Network

$$h(x) = \sigma(Wx + b)$$

$$f_\theta(x) = h^{(2)}\left(h^{(1)}(x)\right) = \sigma^{(2)}\left(W^{(2)}\left(\sigma^{(1)}\left(W^{(1)}x + b^{(1)}\right)\right) + b^{(2)}\right)$$

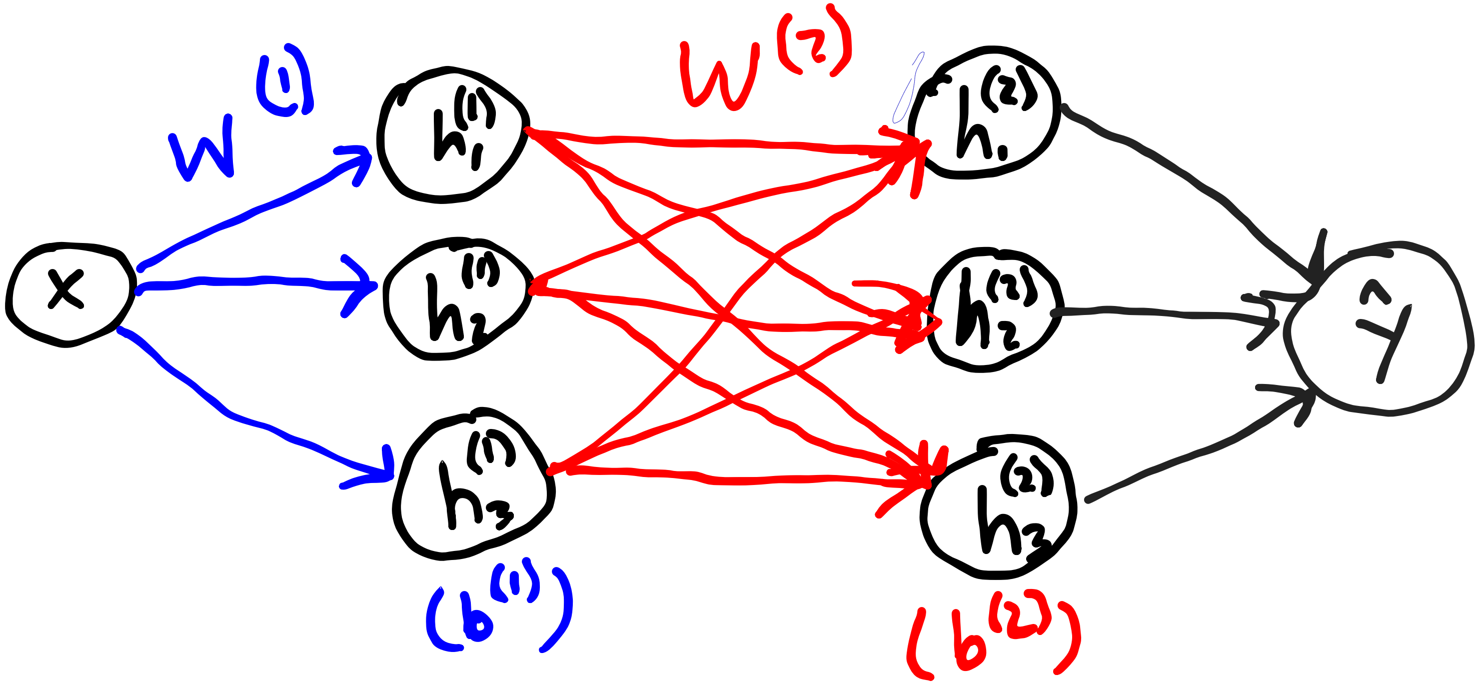
$$\theta = (W^{(1)}, b^{(1)}, W^{(2)}, b^{(2)})$$

$x \in \mathbb{R}^2$

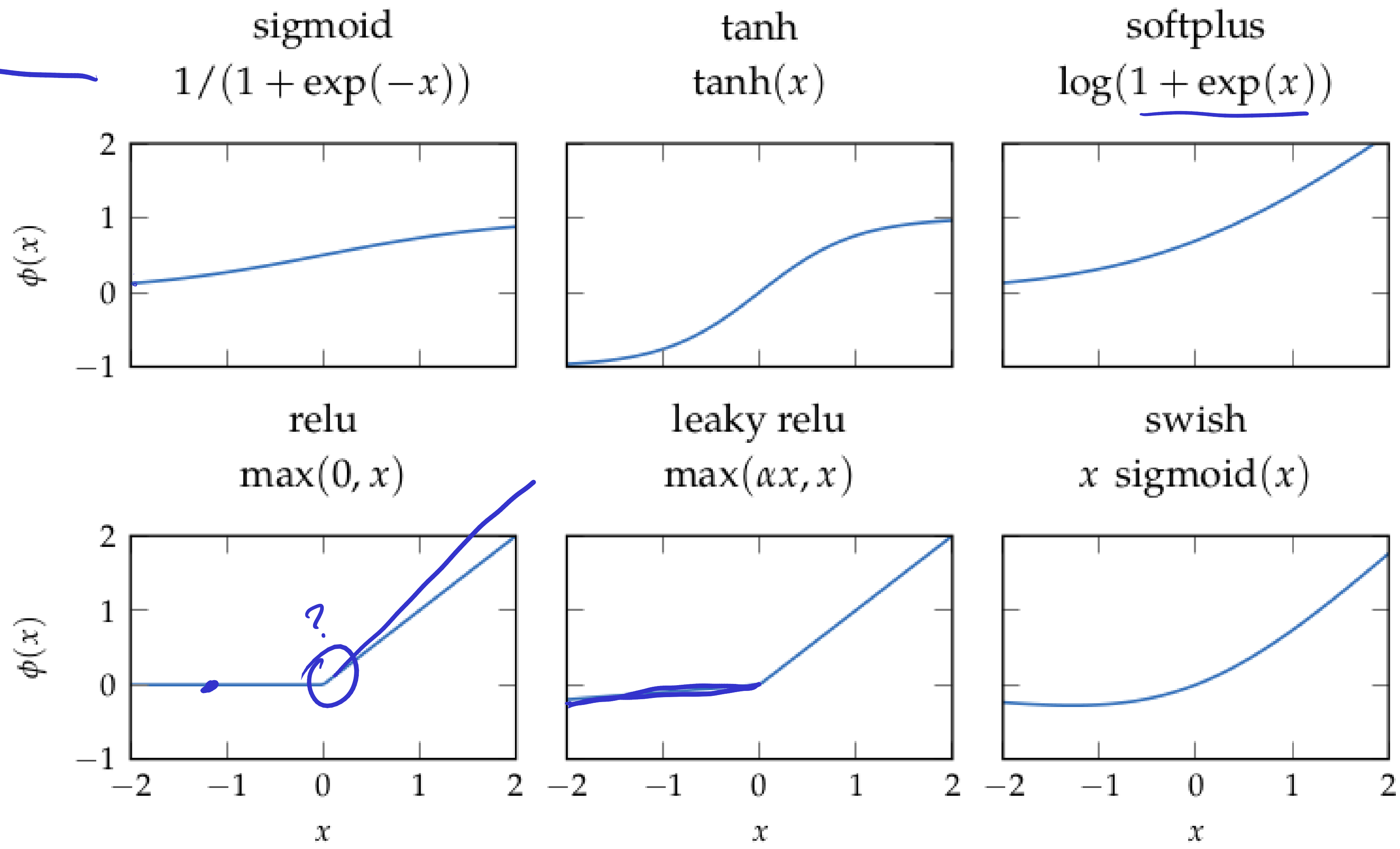
$\hat{y} \in \mathbb{R}^3$

$W^{(1)} \in \mathbb{R}^{3\times 2}$

$b^{(1)} \in \mathbb{R}^3$

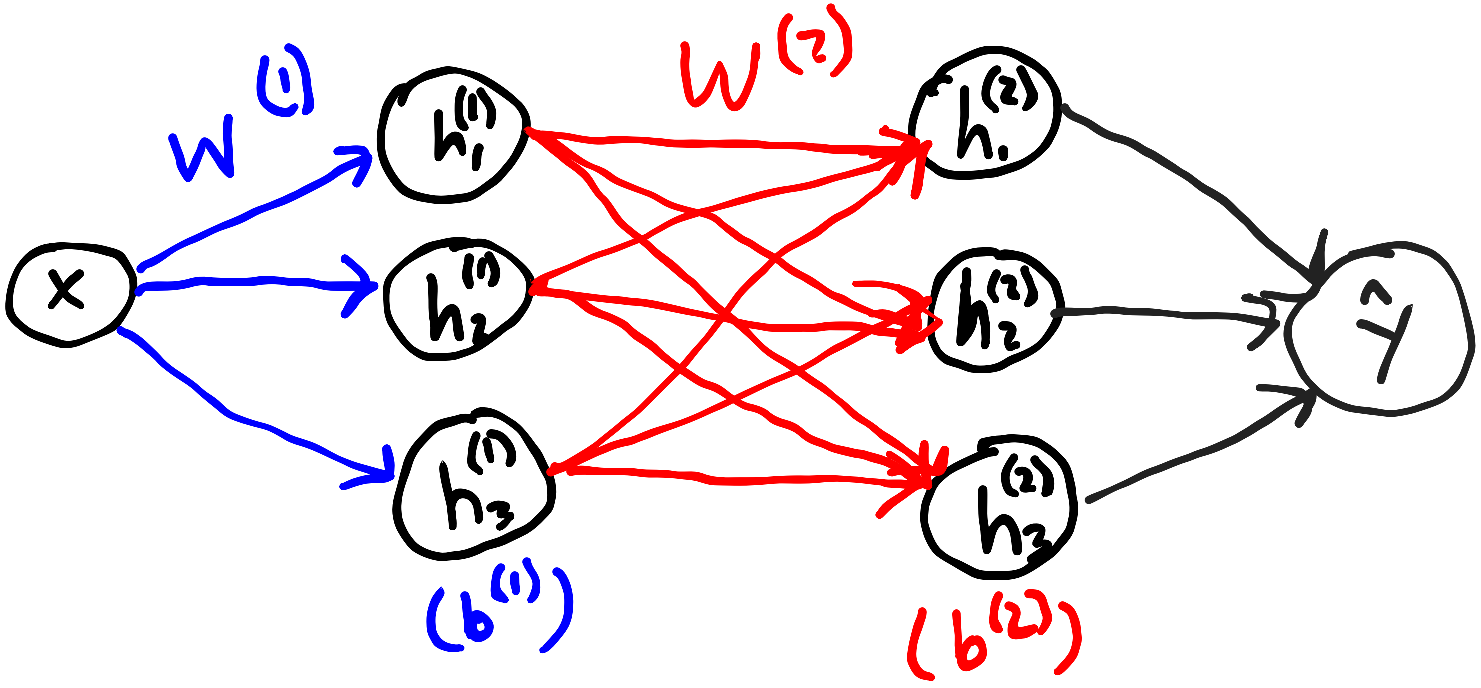


Nonlinearities

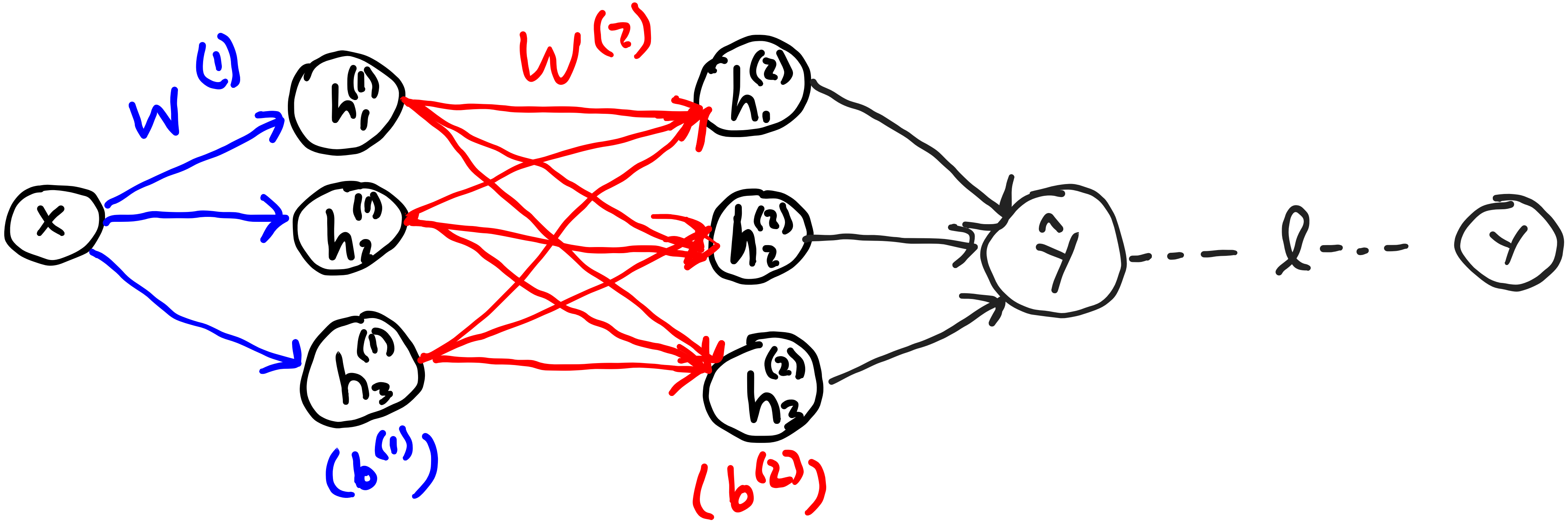


Training

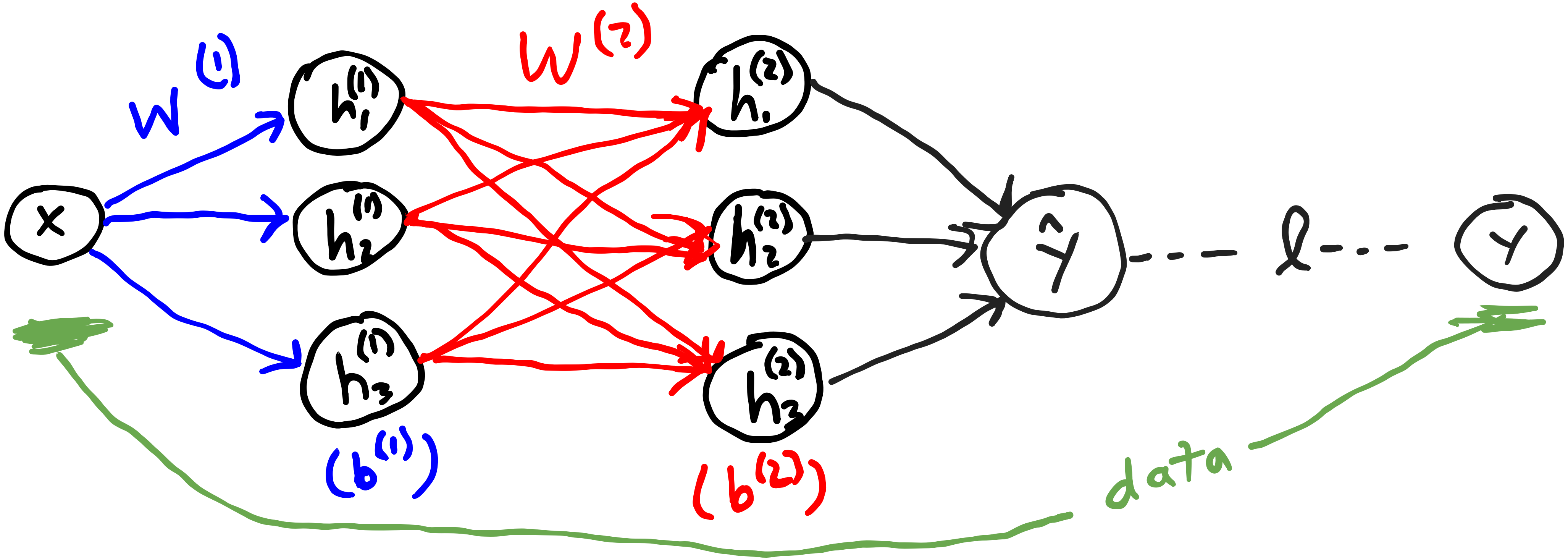
Training



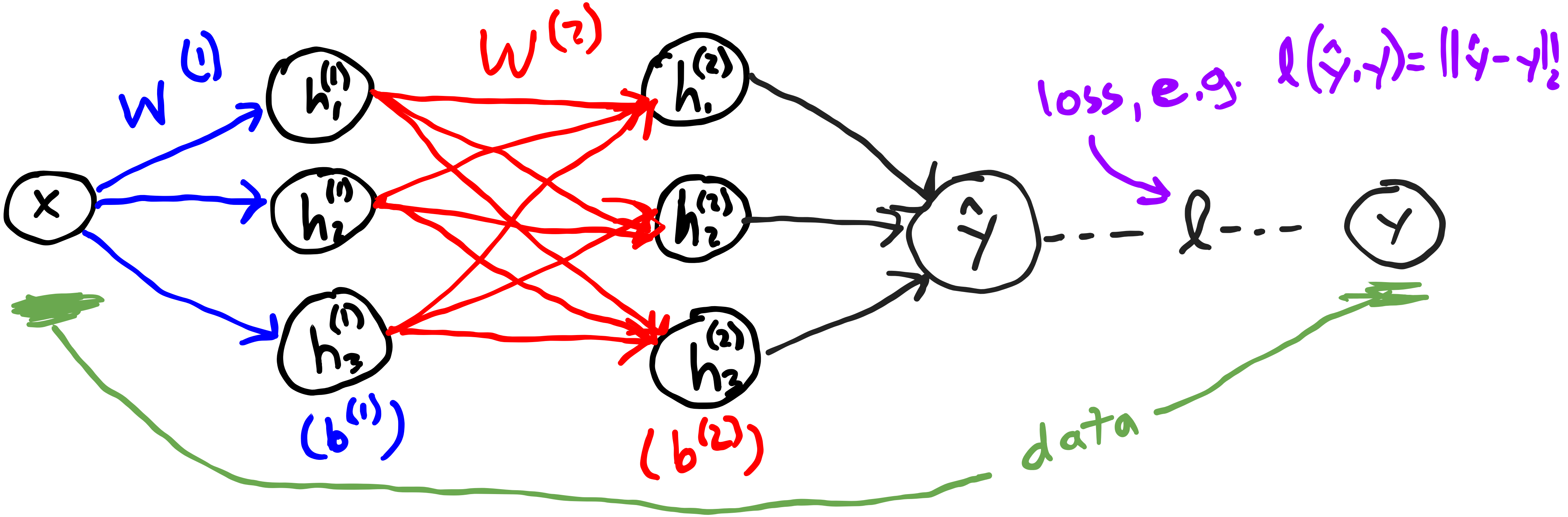
Training



Training

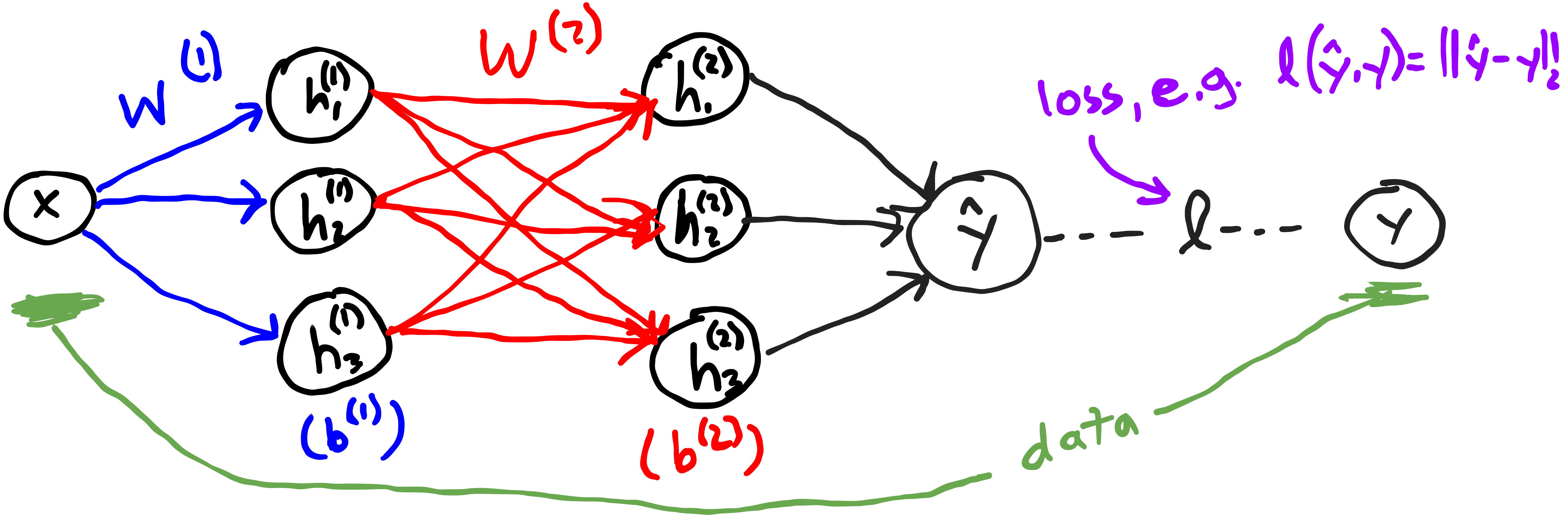


Training



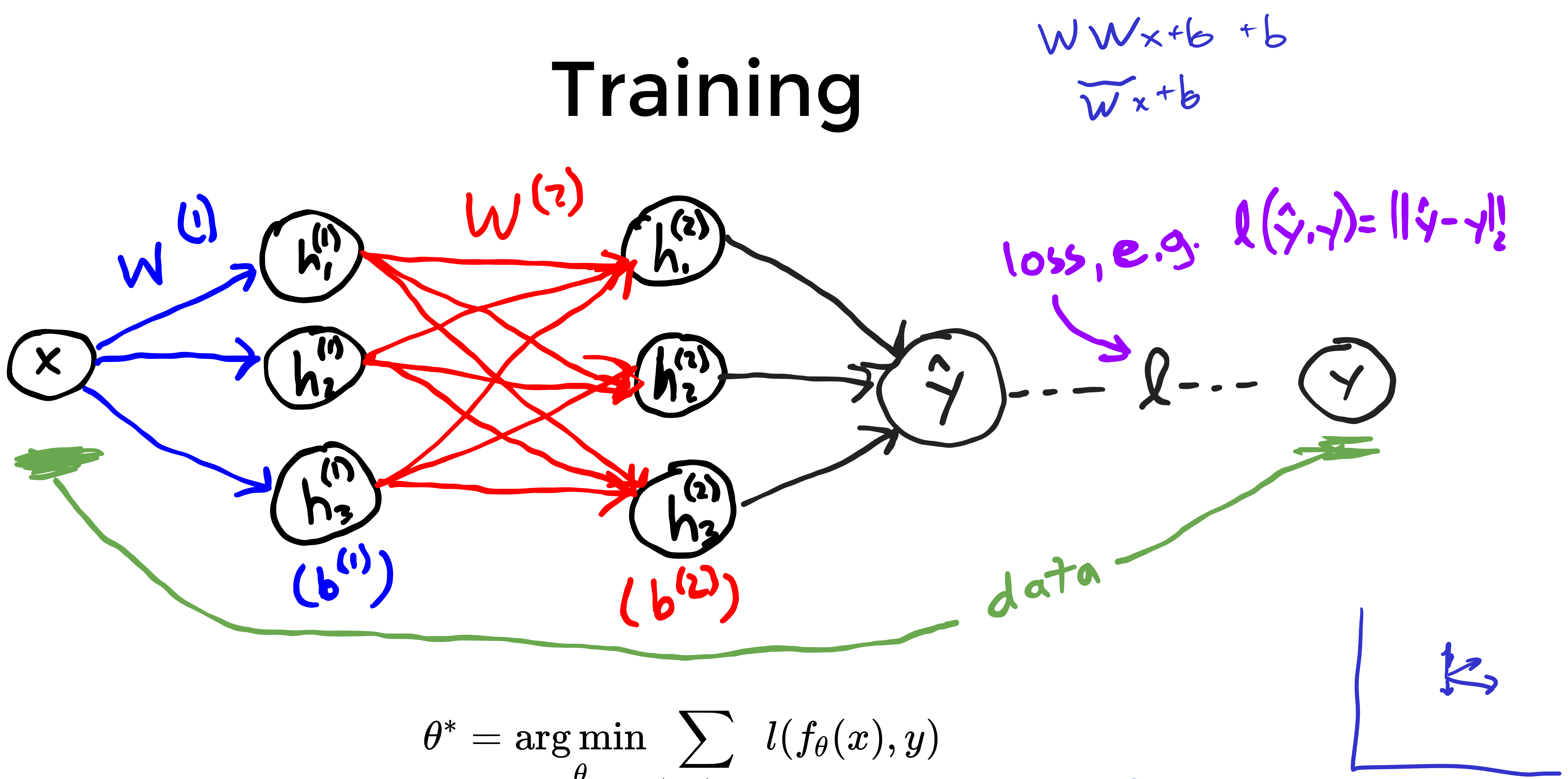
$W^{(1)}$
$W^{(2)}$
x
$h_1^{(1)}$
$h_2^{(1)}$
$h_3^{(1)}$
$(b^{(1)})$
$h_1^{(2)}$
$h_2^{(2)}$
$h_3^{(2)}$
$(b^{(2)})$
$\hat{y}$
loss, e.g. $\ell(\hat{y}, y) = \|\hat{y} - y\|_2$
ℓ
y
data

Training



$$\theta^* = \arg\min_{\theta} \sum_{(x,y)\in\mathcal{D}} l(f_\theta(x), y)$$

Training



$$\theta^* = \arg\min_\theta \sum_{(x,y)\in\mathcal{D}} l(f_\theta(x), y)$$

Software

Stochastic Gradient Descent: $\theta \leftarrow \theta - \alpha \nabla_\theta l(f_\theta(x), y)$

Chain Rule

Chain Rule

$$\frac{\partial \mathbf{f}(\mathbf{g}(\mathbf{h}(\mathbf{x})))}{\partial \mathbf{x}}$$

Chain Rule

$$\frac{\partial \mathbf{f}(\mathbf{g}(\mathbf{h}(\mathbf{x})))}{\partial \mathbf{x}} = \frac{\partial \mathbf{f}(\mathbf{g}(\mathbf{h}))}{\partial \mathbf{h}} \frac{\partial \mathbf{h}(\mathbf{x})}{\partial \mathbf{x}}$$

Chain Rule

$$\frac{\partial \mathbf{f}(\mathbf{g}(\mathbf{h}(\mathbf{x})))}{\partial \mathbf{x}} = \frac{\partial \mathbf{f}(\mathbf{g}(\mathbf{h}))}{\partial \mathbf{h}} \frac{\partial \mathbf{h}(\mathbf{x})}{\partial \mathbf{x}} = \left(\frac{\partial \mathbf{f}(\mathbf{g})}{\partial \mathbf{g}} \frac{\partial \mathbf{g}(\mathbf{h})}{\partial \mathbf{h}} \right) \frac{\partial \mathbf{h}(\mathbf{x})}{\partial \mathbf{x}}$$

Chain Rule

$$\frac{\partial \mathbf{f}(\mathbf{g}(\mathbf{h}(\mathbf{x})))}{\partial \mathbf{x}} = \frac{\partial \mathbf{f}(\mathbf{g}(\mathbf{h}))}{\partial \mathbf{h}} \frac{\partial \mathbf{h}(\mathbf{x})}{\partial \mathbf{x}} = \left(\frac{\partial \mathbf{f}(\mathbf{g})}{\partial \mathbf{g}} \frac{\partial \mathbf{g}(\mathbf{h})}{\partial \mathbf{h}} \right) \frac{\partial \mathbf{h}(\mathbf{x})}{\partial \mathbf{x}}$$

Example:

$$\hat{y} = W^{(2)} \sigma(W^{(1)} x + b^{(1)}) + b^{(2)}$$

Chain Rule

$$\frac{\partial \mathbf{f}(\mathbf{g}(\mathbf{h}(\mathbf{x})))}{\partial \mathbf{x}} = \frac{\partial \mathbf{f}(\mathbf{g}(\mathbf{h}))}{\partial \mathbf{h}} \frac{\partial \mathbf{h}(\mathbf{x})}{\partial \mathbf{x}} = \left(\frac{\partial \mathbf{f}(\mathbf{g})}{\partial \mathbf{g}} \frac{\partial \mathbf{g}(\mathbf{h})}{\partial \mathbf{h}} \right) \frac{\partial \mathbf{h}(\mathbf{x})}{\partial \mathbf{x}}$$

Example:

$$\hat{y} = W^{(2)} \sigma(W^{(1)} x + b^{(1)}) + b^{(2)}$$

$$\frac{\partial l}{\partial W^{(2)}} =$$

Chain Rule

$$\frac{\partial \mathbf{f}(\mathbf{g}(\mathbf{h}(\mathbf{x})))}{\partial \mathbf{x}} = \frac{\partial \mathbf{f}(\mathbf{g}(\mathbf{h}))}{\partial \mathbf{h}} \frac{\partial \mathbf{h}(\mathbf{x})}{\partial \mathbf{x}} = \left(\frac{\partial \mathbf{f}(\mathbf{g})}{\partial \mathbf{g}} \frac{\partial \mathbf{g}(\mathbf{h})}{\partial \mathbf{h}}\right) \frac{\partial \mathbf{h}(\mathbf{x})}{\partial \mathbf{x}}$$

Example:

$$\hat{y} = W^{(2)} \sigma(W^{(1)} x + b^{(1)}) + b^{(2)}$$

$$\frac{\partial l}{\partial W^{(2)}} = \frac{\partial l}{\partial \hat{y}} \left(\frac{\partial \hat{y}}{\partial W^{(2)}}\right)^{\top}$$

Chain Rule

$$\frac{\partial \mathbf{f}(\mathbf{g}(\mathbf{h}(\mathbf{x})))}{\partial \mathbf{x}} = \frac{\partial \mathbf{f}(\mathbf{g}(\mathbf{h}))}{\partial \mathbf{h}} \frac{\partial \mathbf{h}(\mathbf{x})}{\partial \mathbf{x}} = \left(\frac{\partial \mathbf{f}(\mathbf{g})}{\partial \mathbf{g}} \frac{\partial \mathbf{g}(\mathbf{h})}{\partial \mathbf{h}} \right) \frac{\partial \mathbf{h}(\mathbf{x})}{\partial \mathbf{x}}$$

Example:

$$\hat{y} = W^{(2)} \sigma(W^{(1)} x + b^{(1)}) + b^{(2)}$$

$$\frac{\partial l}{\partial W^{(2)}} = \frac{\partial l}{\partial \hat{y}} \left(\frac{\partial \hat{y}}{\partial W^{(2)}} \right)^{\top} = \frac{\partial l}{\partial \hat{y}} \, \sigma \left(W^{(1)} x + b^{(1)} \right)^{\top}$$

Chain Rule

$$\frac{\partial \mathbf{f}(\mathbf{g}(\mathbf{h}(\mathbf{x})))}{\partial \mathbf{x}} = \frac{\partial \mathbf{f}(\mathbf{g}(\mathbf{h}))}{\partial \mathbf{h}} \frac{\partial \mathbf{h}(\mathbf{x})}{\partial \mathbf{x}} = \left(\frac{\partial \mathbf{f}(\mathbf{g})}{\partial \mathbf{g}} \frac{\partial \mathbf{g}(\mathbf{h})}{\partial \mathbf{h}} \right) \frac{\partial \mathbf{h}(\mathbf{x})}{\partial \mathbf{x}}$$

Example:

$$\hat{y} = W^{(2)} \sigma(W^{(1)} x + b^{(1)}) + b^{(2)}$$

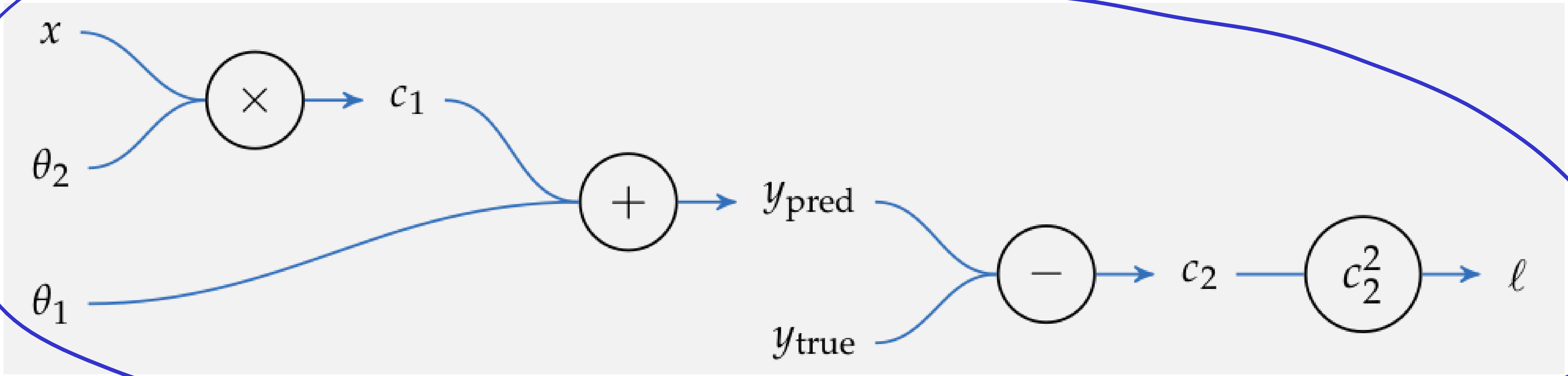
$$\frac{\partial l}{\partial W^{(2)}} = \frac{\partial l}{\partial \hat{y}} \left(\frac{\partial \hat{y}}{\partial W^{(2)}} \right)^\top = \underline{\frac{\partial l}{\partial \hat{y}} \overbrace{\sigma\left(W^{(1)} x + b^{(1)}\right)^\top}^{\text{output of first layer}}}$$

$$W^{(2)} \leftarrow W^{(2)} - \alpha \frac{\partial l}{\partial W^{(2)}}$$

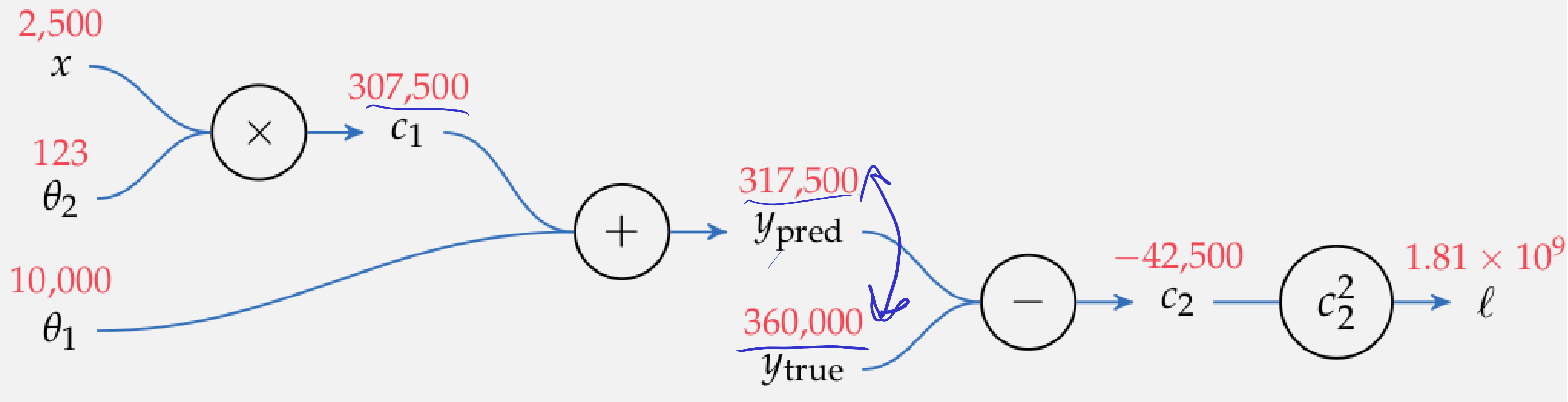
Backprop

Backprop

$\hat{y} = \theta_2 x + \theta_1$

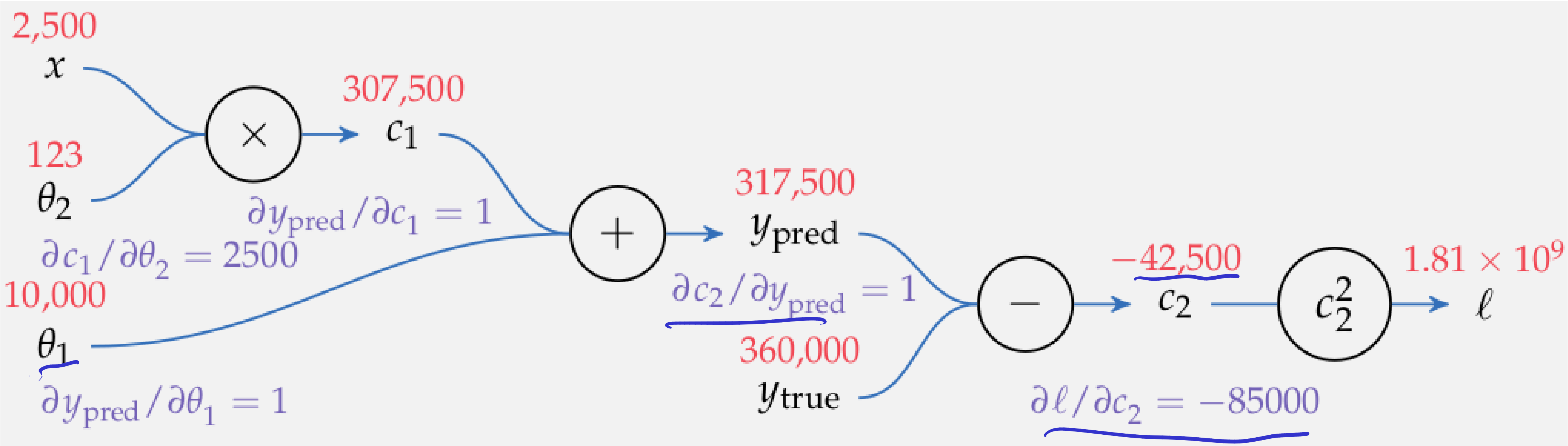


Backprop



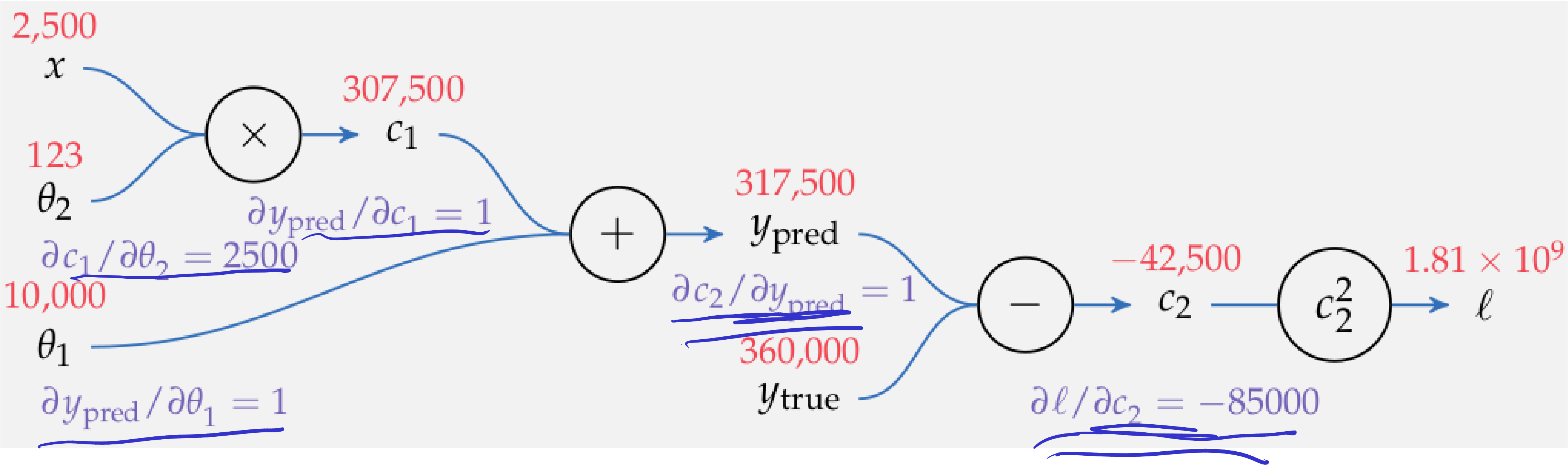
2,500
x
123
θ_2
10,000
θ_1
307,500
c_1
317,500
y_{pred}
360,000
y_{true}
−42,500
c_2
c_2^2
1.81×10^9
ℓ

Backprop



2,500
x
123
θ_2
307,500
c_1
$\partial y_{\text{pred}} / \partial c_1 = 1$
$\partial c_1 / \partial \theta_2 = 2500$
10,000
θ_1
$\partial y_{\text{pred}} / \partial \theta_1 = 1$
317,500
y_{pred}
$\partial c_2 / \partial y_{\text{pred}} = 1$
360,000
y_{true}
$-42,500$
c_2
c_2^2
1.81×10^9
ℓ
$\partial \ell / \partial c_2 = -85000$

Backprop

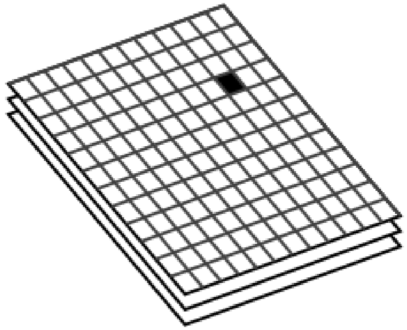


$$\frac{\partial \ell}{\partial \theta_1} = \frac{\partial \ell}{\partial c_2}\frac{\partial c_2}{\partial y_{\text{pred}}}\frac{\partial y_{\text{pred}}}{\partial \theta_1} = -85{,}000 \cdot 1 \cdot 1 = -85{,}000$$

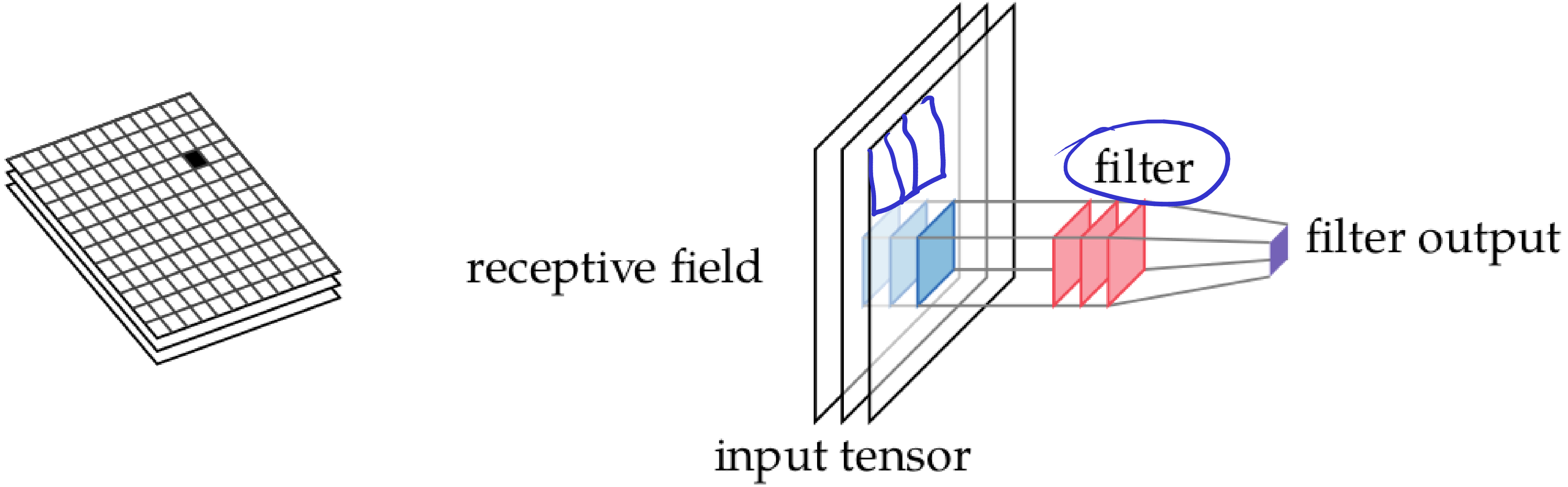
$$\frac{\partial \ell}{\partial \theta_2} = \frac{\partial \ell}{\partial c_2}\frac{\partial c_2}{\partial y_{\text{pred}}}\frac{\partial y_{\text{pred}}}{\partial c_1}\frac{\partial c_1}{\partial \theta_2} = -85{,}000 \cdot 1 \cdot 1 \cdot 2500 = -2.125 \times 10^8$$

On Your Radar: ConvNets

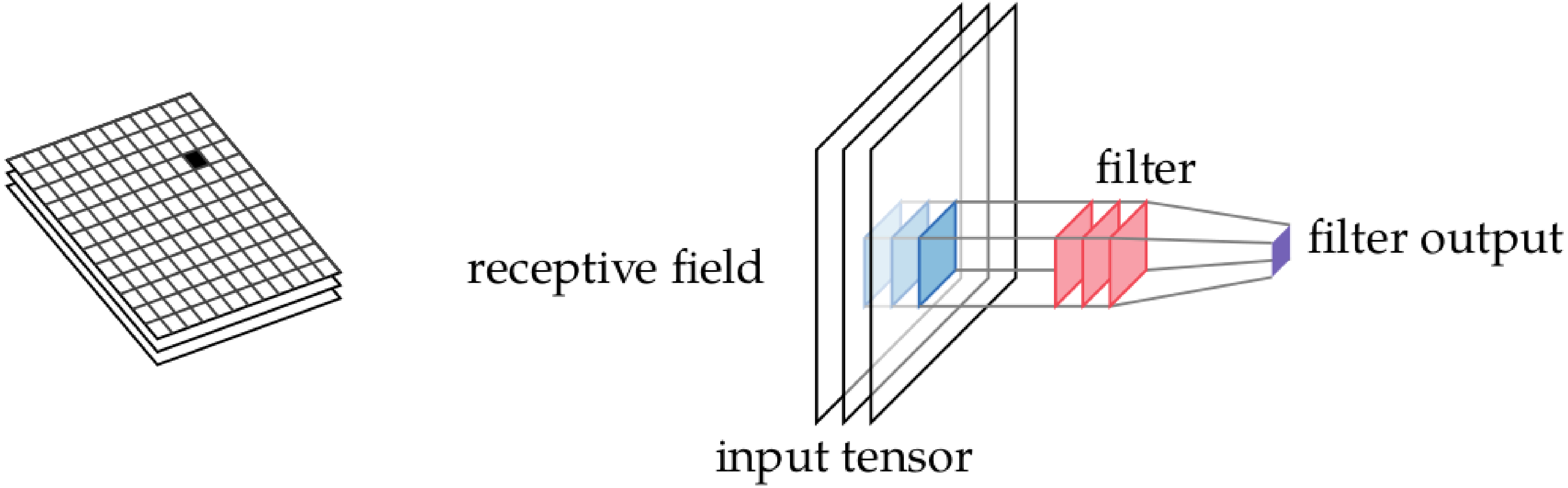
On Your Radar: ConvNets

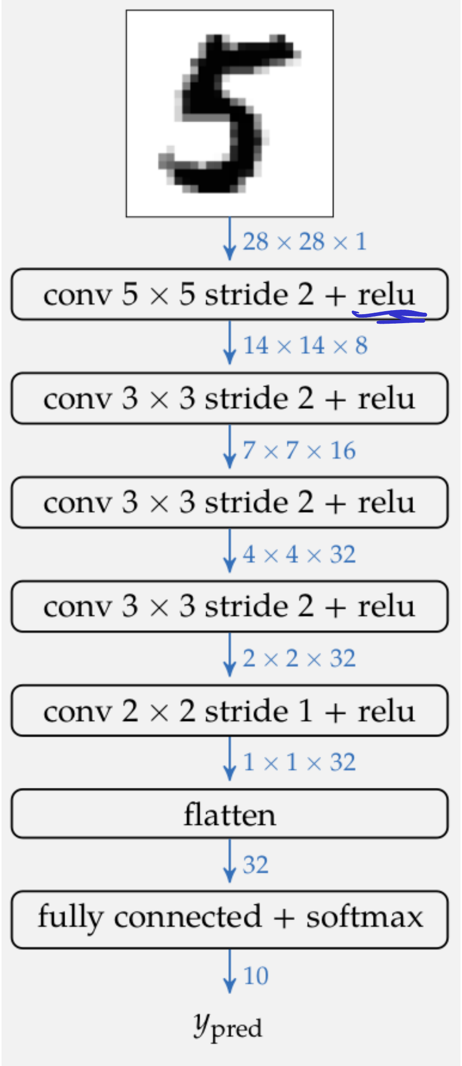


On Your Radar: ConvNets



On Your Radar: ConvNets





On Your Radar: Regularization

On Your Radar: Regularization

$$\arg\min_{\theta} \sum_{(x,y)\in\mathbf{D}} \ell(f_{\theta}(x), y) - \beta\|\theta\|^2$$

On Your Radar: Regularization

$$\arg\min_{\theta} \sum_{(x,y)\in \mathbf{D}} \ell(f_{\theta}(x), y) - \beta \|\theta\|^2$$

e.g. Batch norm, dropout, resnets