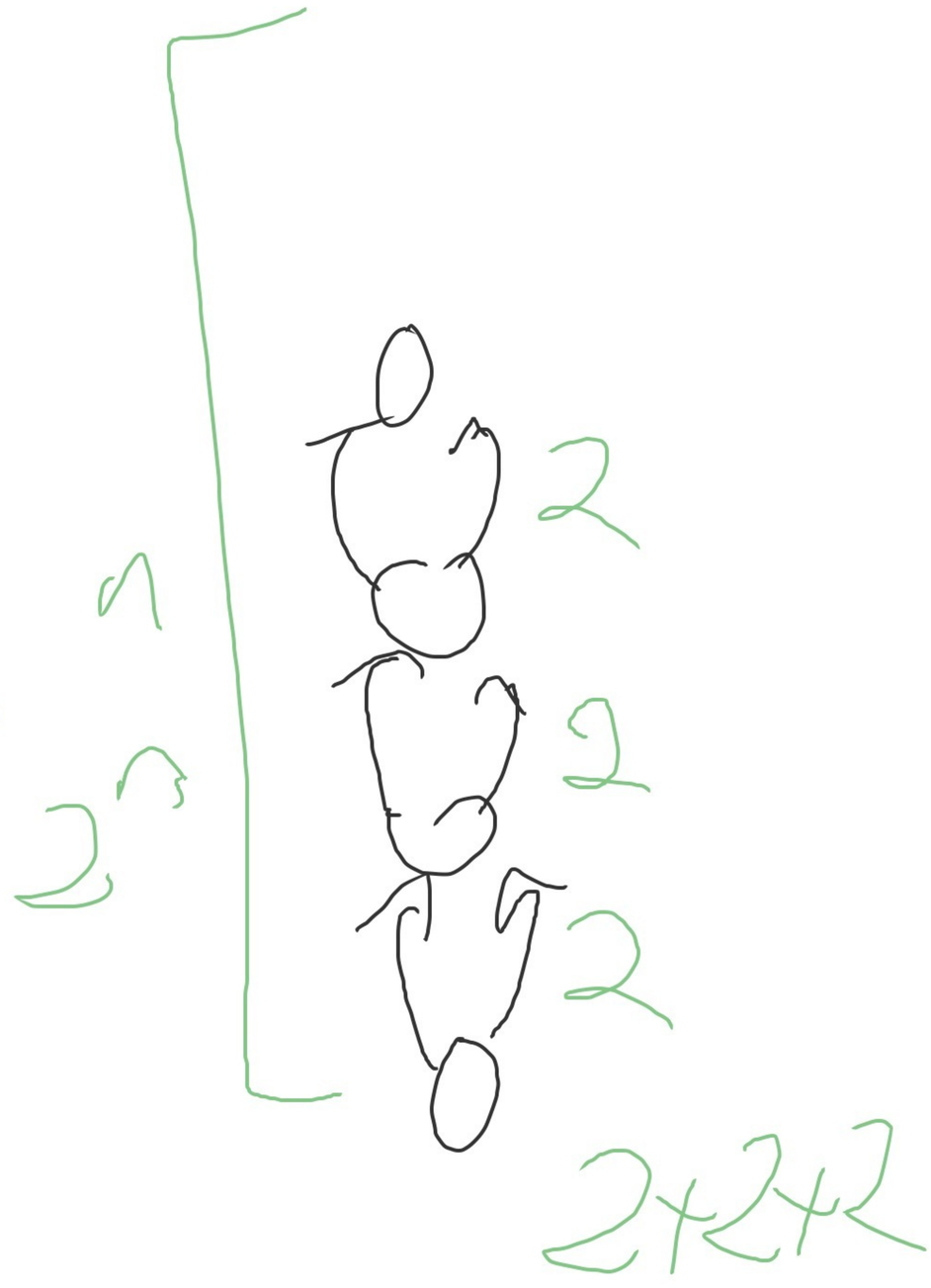
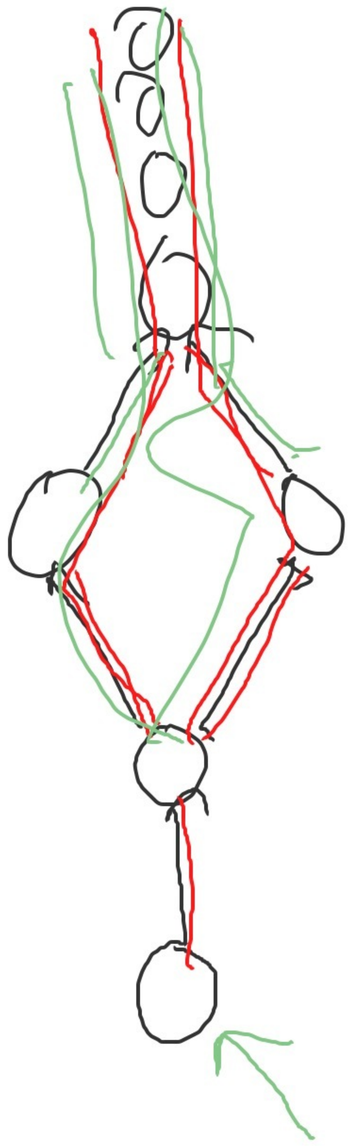
1. determine fanout
2. initialize cotg (gradients reset) set to zero
3. backpropagation backward reverse pass phase sweep





(1) upon creating a vertex
set fanout to zero

(2) increment fanout
each time process vertex

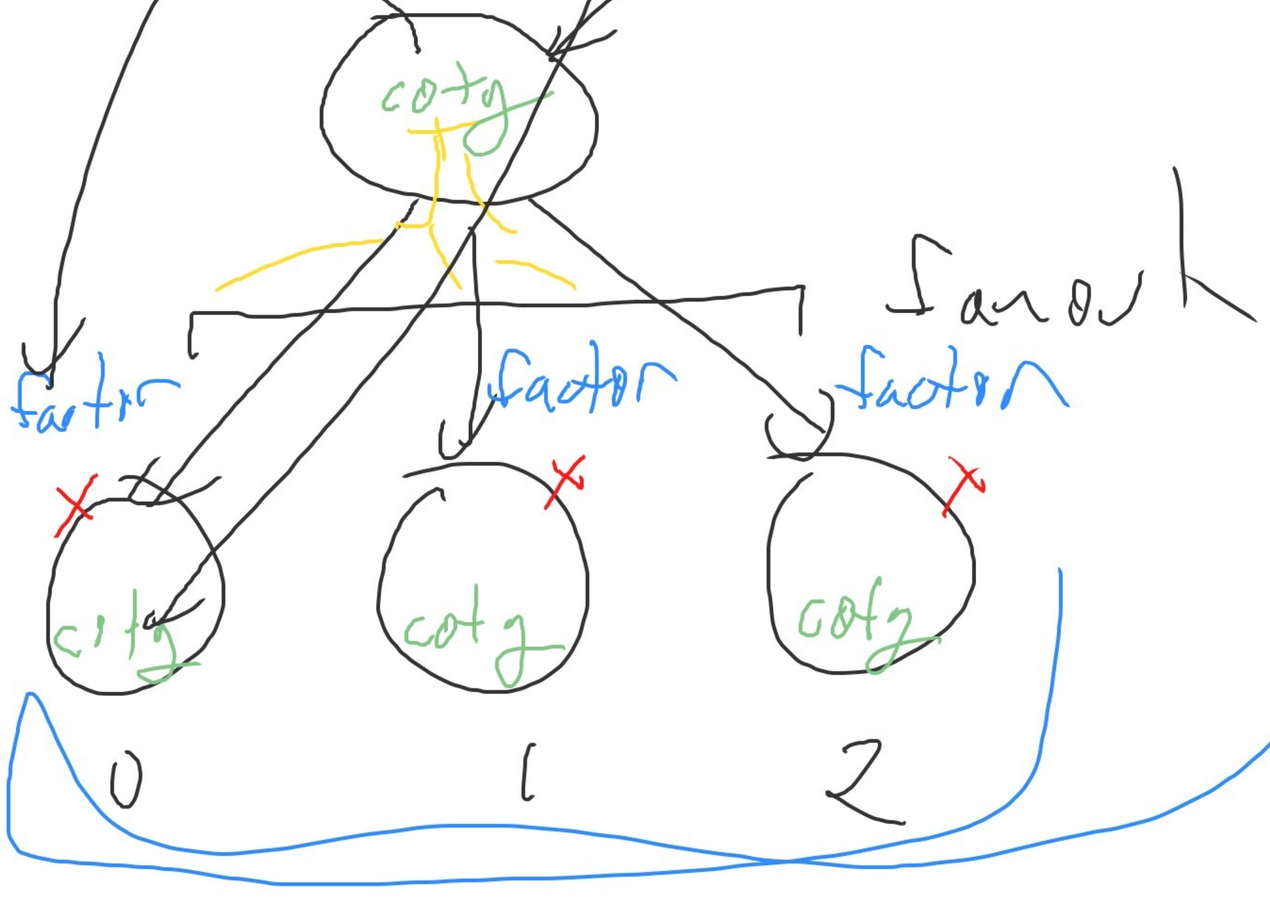
(3) only when incrementing
from zero to one
process children

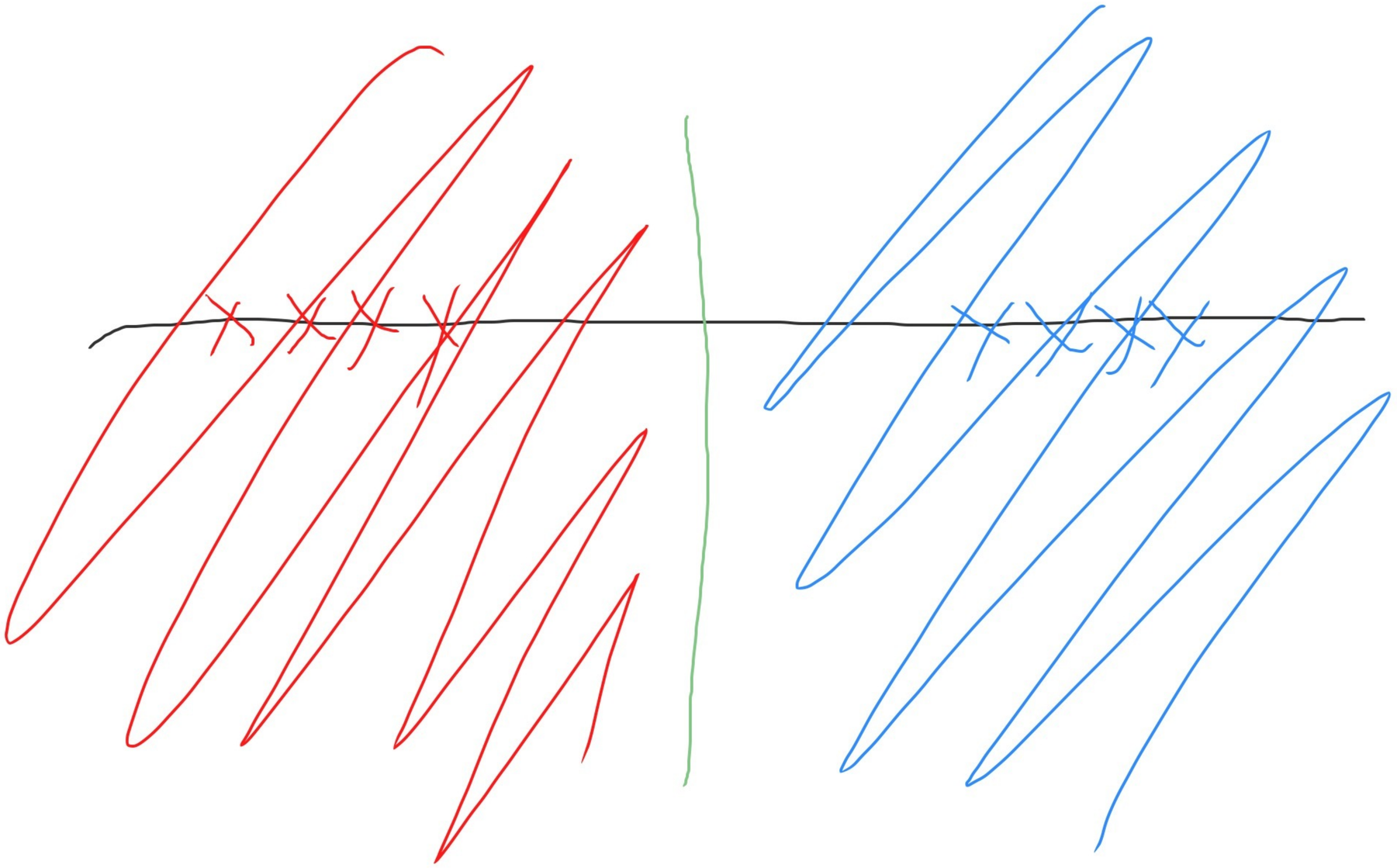
for i = 0, 1, 2

coty += factori × colyi

← only action all these

initialized to zero





parameter

$y = ax + b$

model

$\vec{x}$

$\vec{a} \cdot \vec{x} + b$

≥ 0

≤ 0

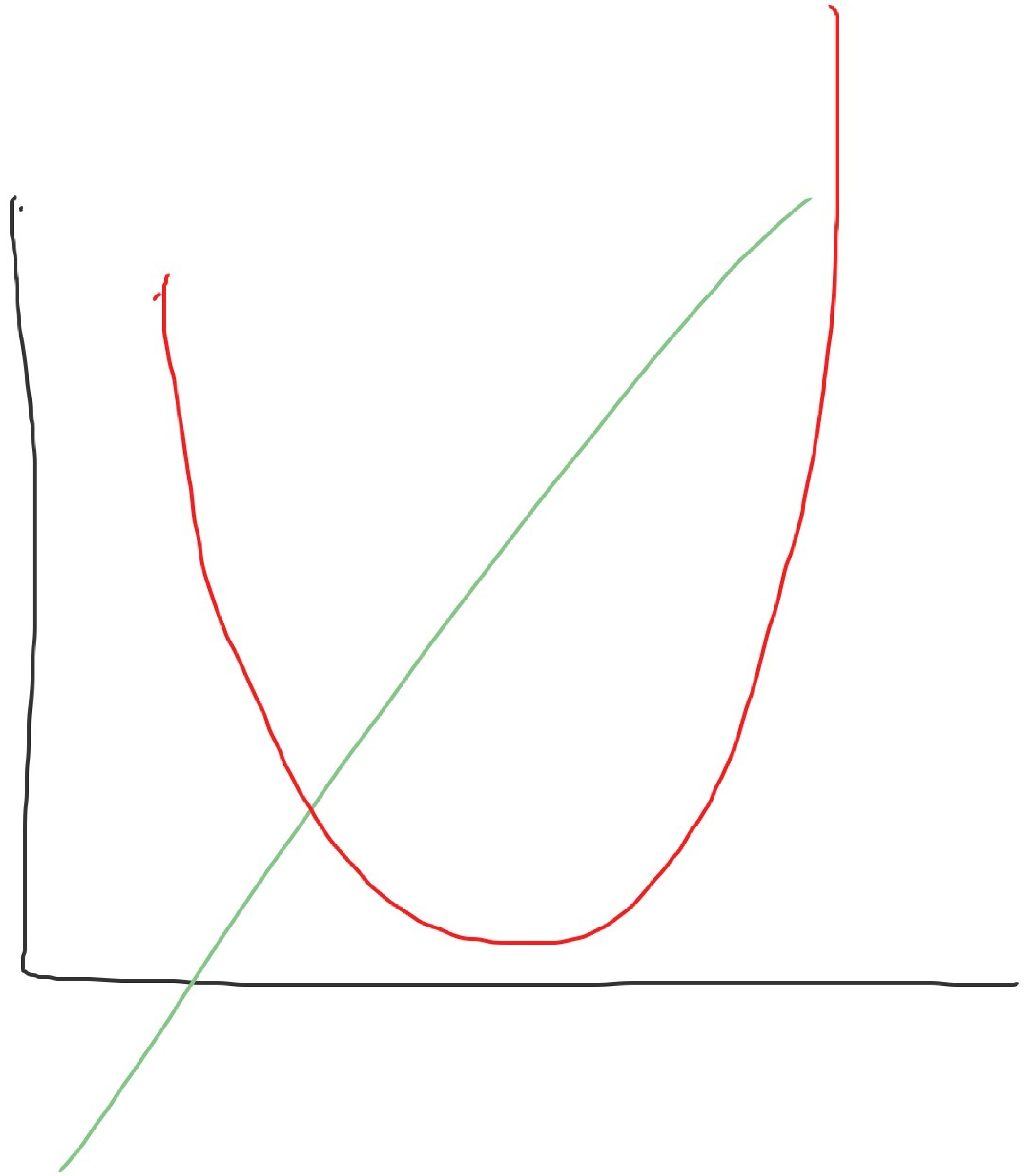
$$M(x; \overbrace{a, b}^{\theta}) \qquad y = M(x; \theta)$$

x_1	y_1	$M(x_1; \theta)$
x_2	y_2	$M(x_2; \theta)$
$\vdots$	$\vdots$	$\vdots$
x_n	y_2	$M(x_n; \theta)$
samples	labels	predictions

$$\min \sum_{i=1}^{n} \left(y_i - M(x_i; \theta) \right)^4$$

$M(x, a, b) =$

$a \cdot x + b$



$ax^2 + bx + c$

$x^T A x + b^T x + c$

Newton Raphson

$$\theta \leftarrow \theta - \frac{f(\theta)}{f'(\theta)}$$

$f(\theta) = 0$

↑

g'

Newton

$$\theta \leftarrow \theta - \eta \frac{g'(\theta)}{}$$

$$\arg\min_{\theta} g(\theta)$$

$$\vec{u} + \vec{v}$$

$$k\vec{v}$$

$$\vec{u} - \vec{v} = \vec{u} + (-1)\vec{v}$$

$$\vec{u} \cdot \vec{v}$$

$$\|\vec{u}, \vec{v}\| \propto (\vec{u} - \vec{v}) \cdot (\vec{u} - \vec{v})$$

naive gradient descent

$$x \leftarrow x - \eta f'(x)$$

$$M(\vec{x}; \vec{a}, b) = \vec{a} \cdot \vec{x} + b$$

$$x \leftarrow x - \eta \nabla f(x)$$

gradient loss function

learning rate

[a, b]

[weights, bias]

$\vec{x}_1$ $\|y_1 -$ $\mu(\vec{x}_1,$ $[\vec{\theta}]$ $)\|_1$

$\vdots$ $\vdots$ $\vdots$ $+$ $\vdots$ loss

$\vec{x}_n$ $\|y_n -$ $\mu(\vec{x}_n,$ $[\vec{\theta}]$ $)\|_1$

$\vec{a} \cdot \vec{x} + b$

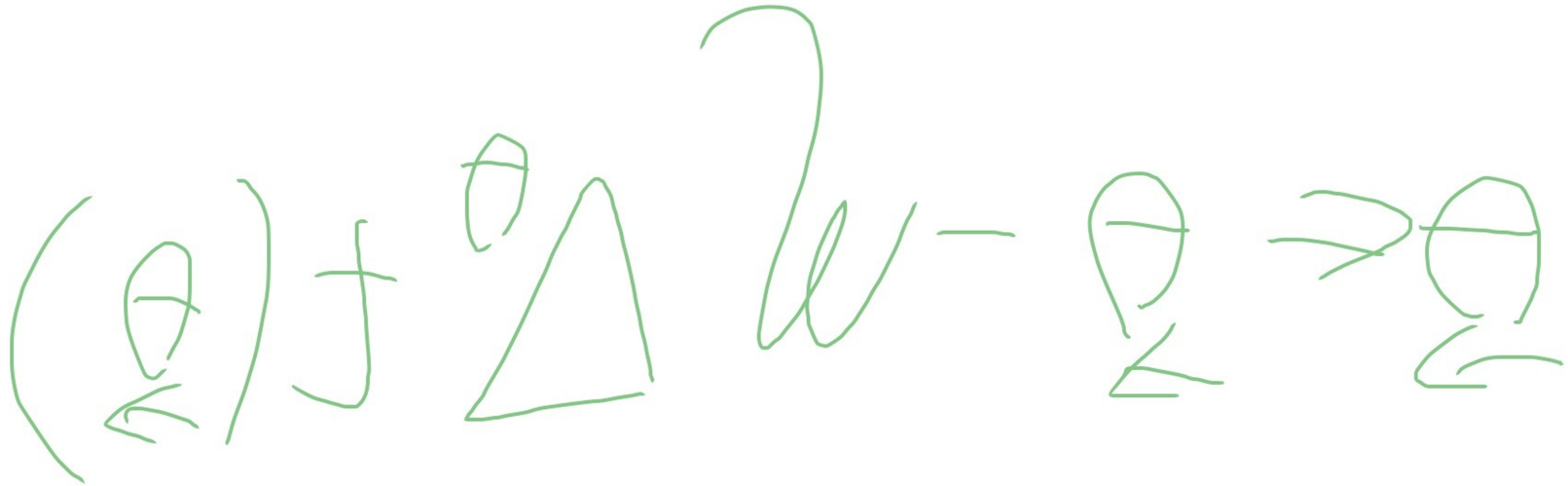
$[a, b]$ θ

samples

label

ground truth

$$f(\theta) = \sum_{i=1}^{n} \|y_i - \mu(\vec{x}_i, \vec{\theta})$$



$\vec{a} \cdot \vec{x} + b$

x_1 x_2 x_n

a_1 a_n

b

$+$

$\vec{y}$

weight, bias parameters

$\vec{a}, b$

input activation

FC

$\vec{y}$

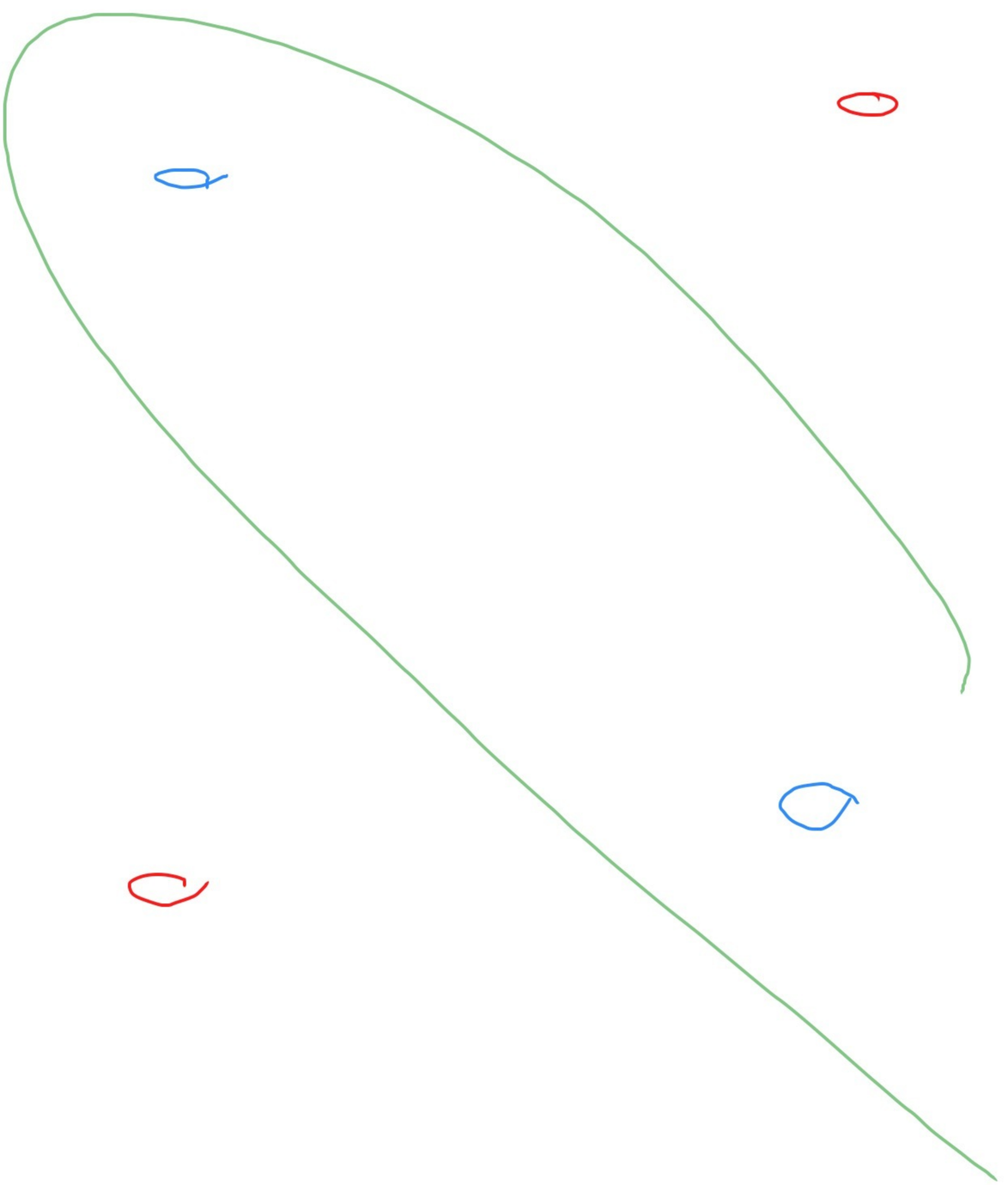
output activation

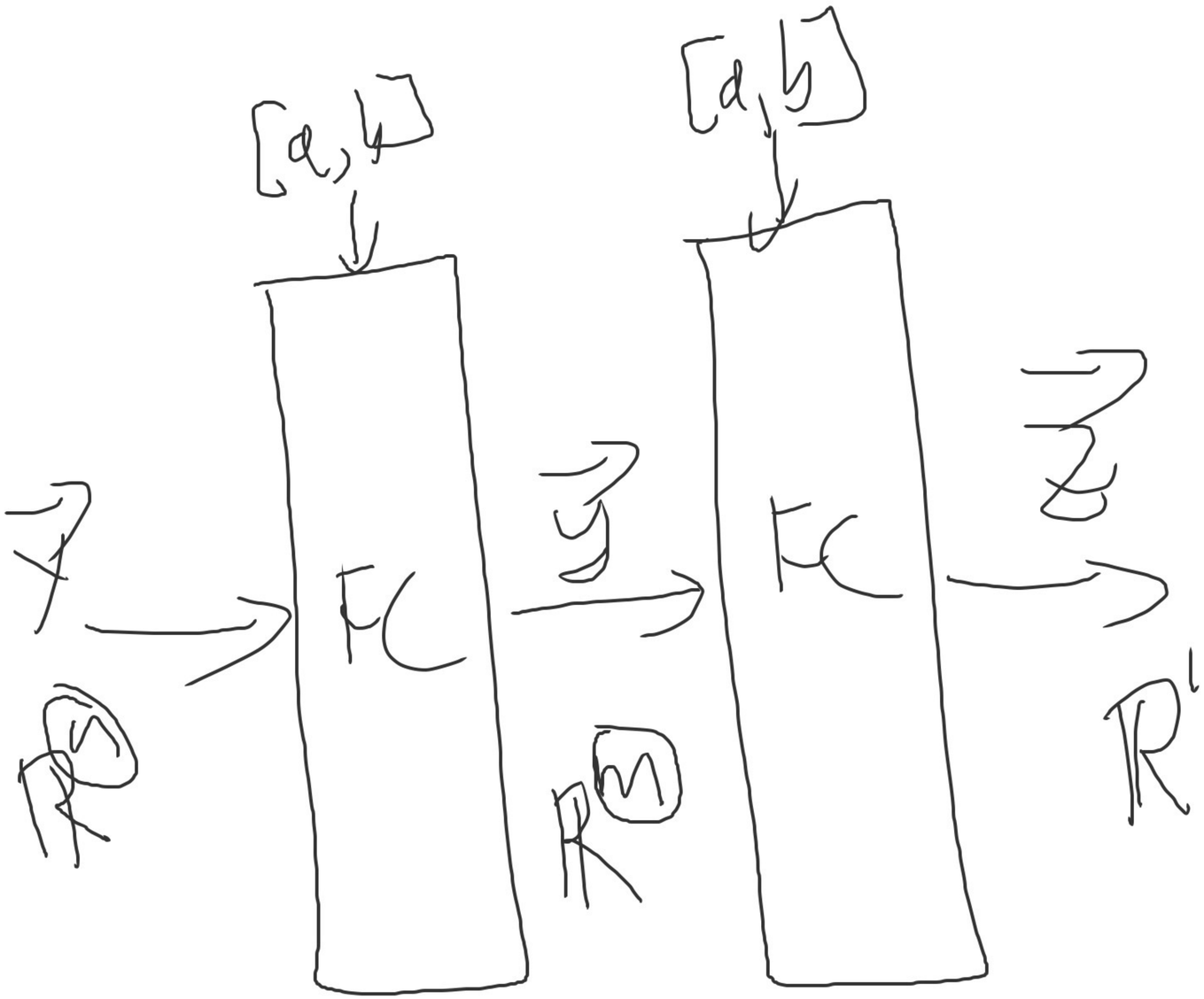
matrix

$\vec{y} = A\vec{x} + \vec{b}$

matrix multiplication

fully-connected




[a,b]
[a,b]
F(
F(
R^n
R^m
R^l