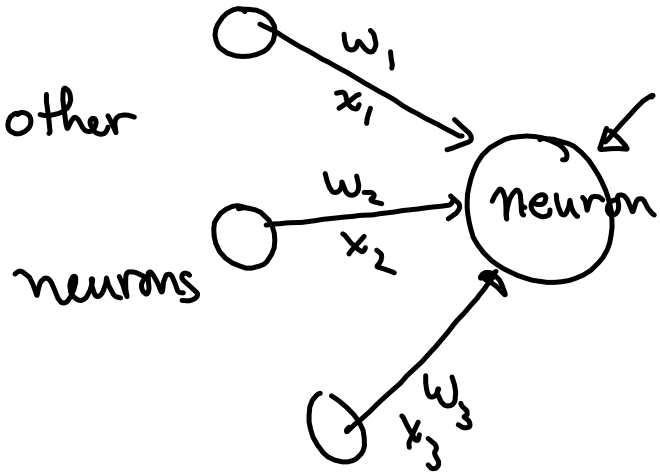
Tuesday April 4, 2017

Neural Networks I

Mathematical Aspects of.

An attempt to mimic the functionality of a zombie brain.

An army of zombies is dangerous.

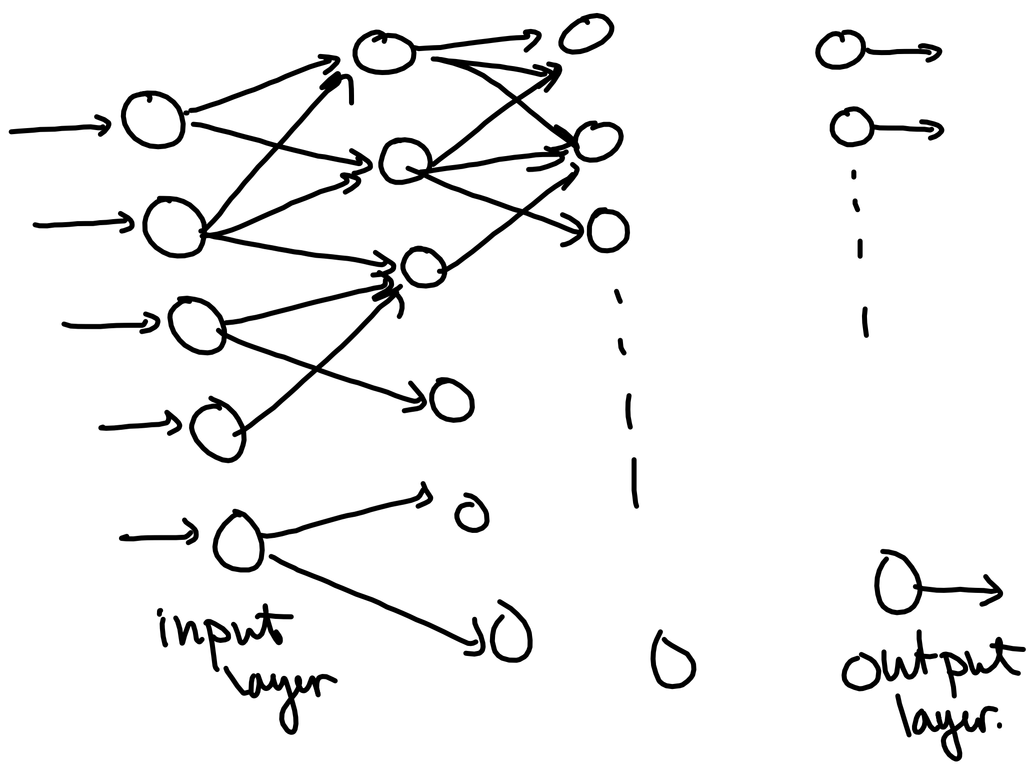



decide if

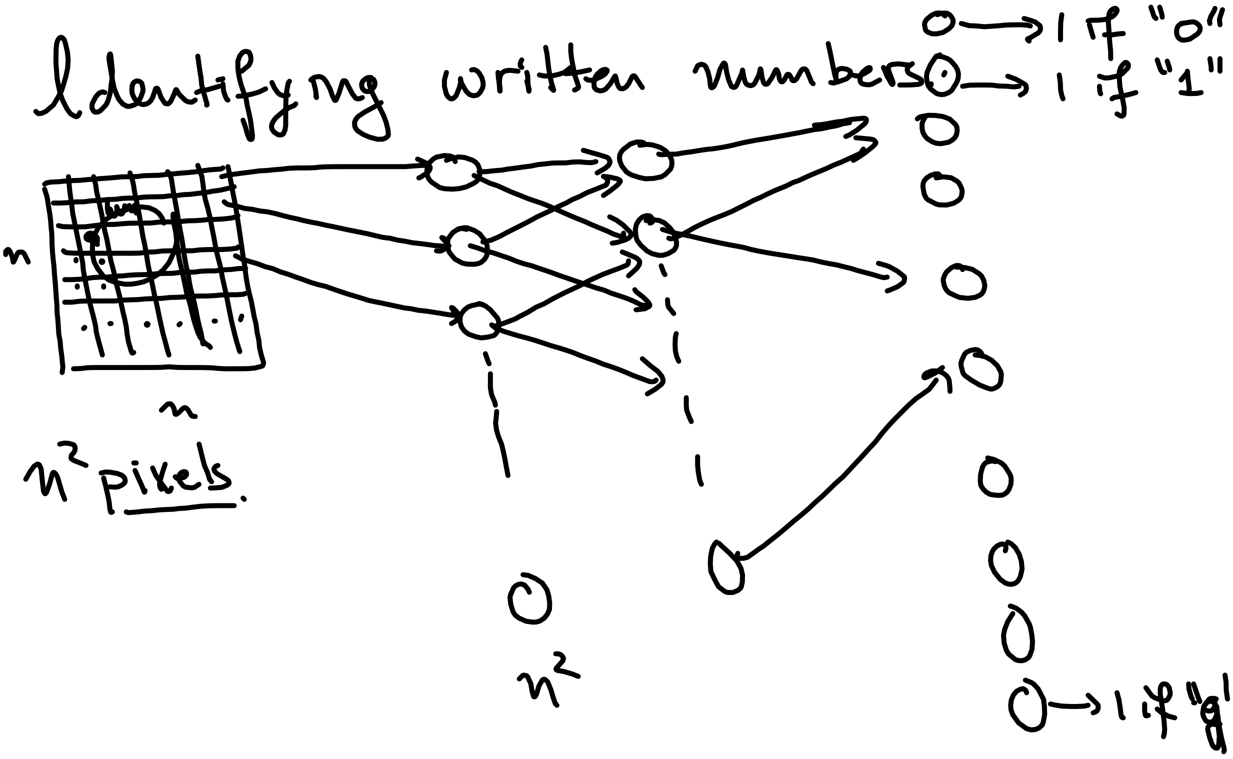
$w_1x_1 + w_2x_2 + w_3x_3$

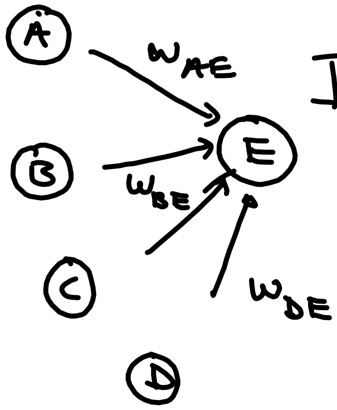
is "important" or not.

If it IS "important" then it "fires".


input
layer
output
layer.

Identifying written numbers

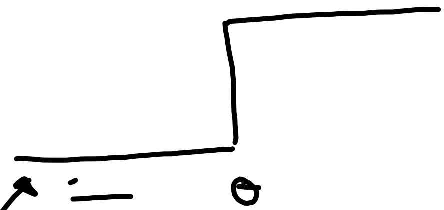





$$I_E = w_{AE} O_A + w_{BE} O_B + w_{CE} O_C + w_{DE} O_D$$

Activation Function.

$O_E = \phi(I_E - \theta)$



$\theta$

↑ threshold.

$$\phi(x) = \begin{cases} 1 & \text{if } x>0 \\ 0 & \text{if } x<0 \end{cases}$$

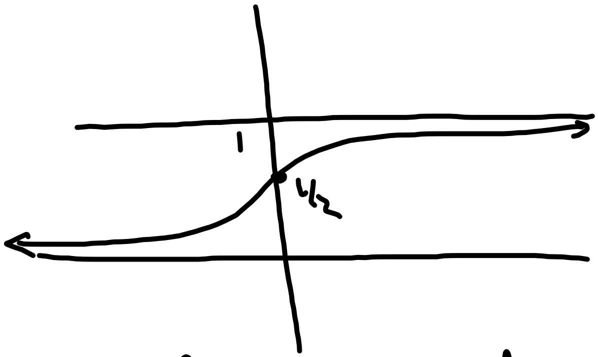
$\phi(x-\theta)$

In practice they use a "smoothed" version of the switch.

Typically $\phi(x) = \frac{1}{1+e^{-x}}$

$$\lim_{x \to \infty} \phi(x) = \frac{1}{1+0} = 1$$

$$\lim_{x \to -\infty} \phi(x) = \frac{1}{1+\infty} = 0$$




$$\phi(x) = \frac{1}{1+e^{-x}}$$

$$\phi(0) = \frac{1}{1+1} = 1/2$$

$$\phi'(x) = \frac{d}{dx}\frac{1}{1+e^{-x}} = \frac{d}{dx}(1+e^{-x})^{-1}$$

$$= (-1)(1+e^{-x})^{-2}(-e^{-x})$$

$$= \frac{e^{-x}}{(1+e^{-x})^2}$$

$$\phi'(x) = \frac{e^{-x}}{(1+e^{-x})^2} \qquad \phi(x) = \frac{1}{1+e^{-x}}$$

Logistic

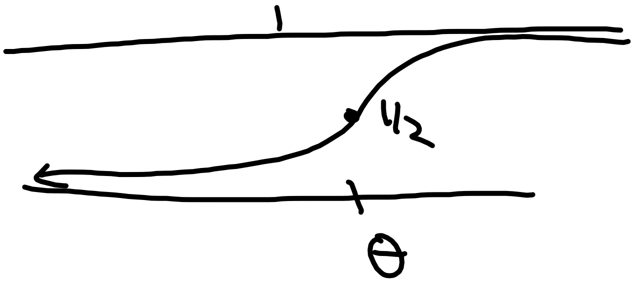
$$= \frac{e^{-x}}{1+e^{-x}} \cdot \underbrace{\left(\frac{1}{1+e^{-x}}\right)}_{\phi(x)}$$

$$\boxed{\begin{array}{c}\phi'(x) = \phi(x)(1-\phi(x)) \\ \phi(0) = 1/2\end{array}}$$

$$= \frac{(e^{-x}+1)-1}{1+e^{-x}} \cdot \phi(x) = (1-\phi(x))\,\phi(x)$$

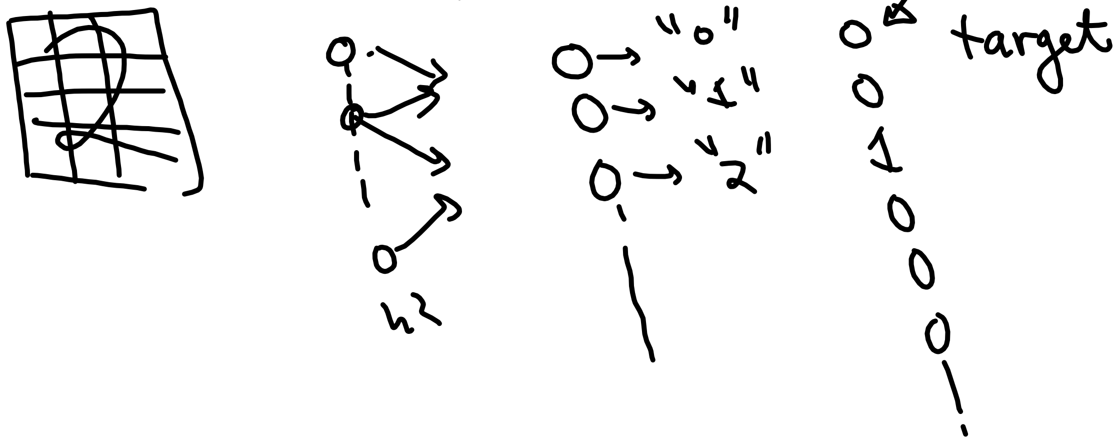
$$\phi'(x) = \phi(x)(1-\phi(x))$$

$$\phi(\theta) = 1/2$$



# Learning

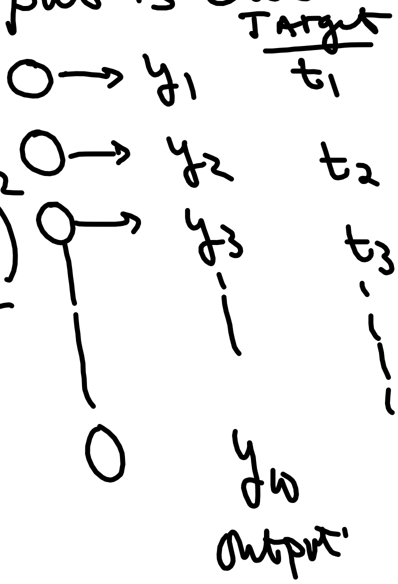
Training set — images where you know the answer.



Want to change the weights
so the output is closer to the target.

Error

$$= \frac{1}{n}\sum_{i=1}^{n}(y_i - t_i)^2$$



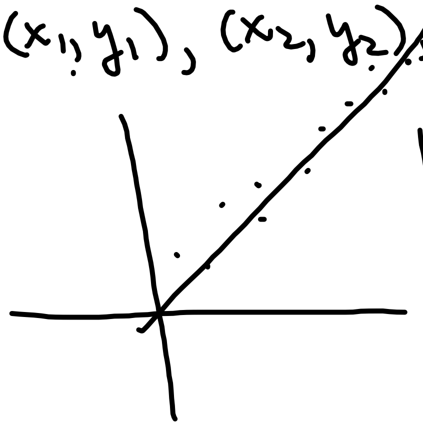
Basic Problem:

Minimize the error ! by adjusting the weights.

A basic mathematical problem.

# Least Squares problem.

<u>Data</u>: $(x_1, y_1), (x_2, y_2), ---, (x_n, y_n)$



line $y = w_1 x + w_2$

$(x_1, y_1)$ $y = w_1 x_1 + w_2$

error for this point is

$$(w_1 x_1 + w_2 - y_1)^2$$

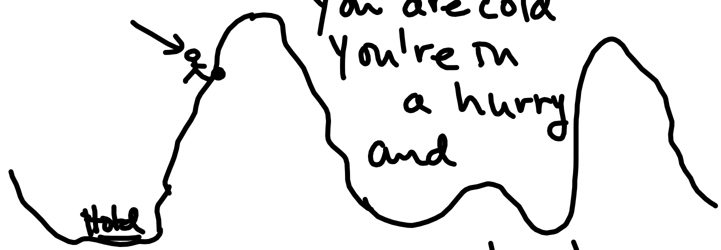
Total error: $E = \sum_{i=1}^{n} (w_1 x_i + w_2 - y_i)^2$

$= F(w_1, w_2)$

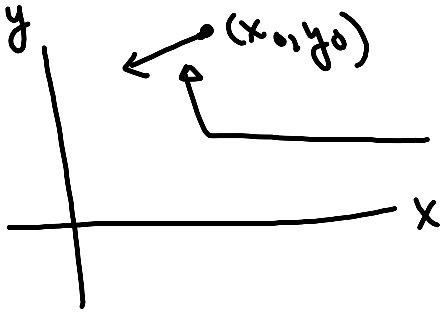
$z = f(x, y)$

height location

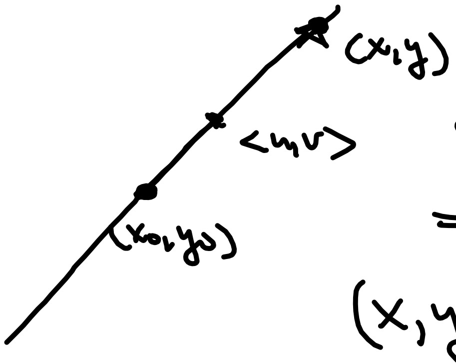
you in a snow storm
You are cold
You're in
a hurry
and




you're hungry
and you're
thirsty.




direction
vector
$\langle u, v \rangle$ ; $\sqrt{u^2+v^2} = 1$




$$\langle x-x_0, y-y_0\rangle$$
$$= t\langle u,v\rangle$$
$$(x,y) = (x_0+tu, y_0+tv)$$

dist between $(x,y)$ and $(x_0,y_0)$

$$= \{(x-x_0)^2+(y-y_0)^2\}^{1/2}$$
$$= \{(tu)^2+(tv)^2\}^{1/2} = |t|\sqrt{u^2+v^2}$$
$$= |t|$$

$$z(t) = f(x, y) = f(x_0 + tu, y_0 + tv)$$

height

rate of change in height is $z'(t)$

$t=0$ gives our location $(x_0, y_0)$

Problem: Find the direction for which $z'(0)$ is as small as possible!

An example: $f(x,y) = Ax + By + C$

$$z(t) = f(x_0 + tu, y_0 + tv)$$

$$= A(x_0 + tu) + B(y_0 + tv) + C$$

$$z'(t) = Au + Bv$$

$$\boxed{z'(0) = Au + Bv}$$

the rates of change of height at our location if we move in the direction $\langle u, v \rangle$.

Problem: find the direction $\langle u, v \rangle$ such that

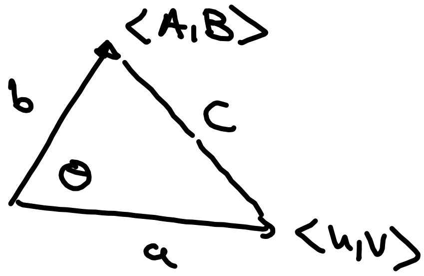
$$z'(0) = Au + Bv$$

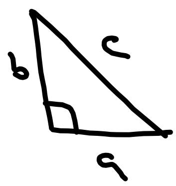
is as small as possible!

# Trigonometry

$Au + Bv$

$\langle u, v \rangle$ ; $\langle A, B \rangle$



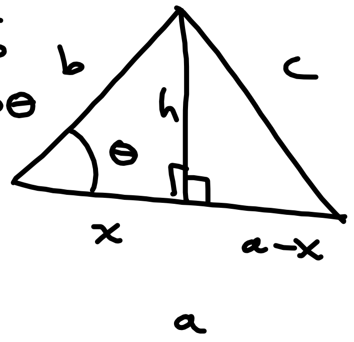



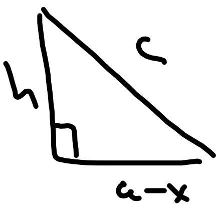

$c^2 = a^2 + b^2$

## Law of Cosines

$\cos\theta = \frac{x}{b}$

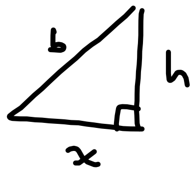
$\Rightarrow x = b\cos\theta$




$h^2 + (a-x)^2 = c^2$

$h^2 = c^2 - (a-x)^2$




$x^2 + h^2 = b^2$

$h^2 = b^2 - x^2$

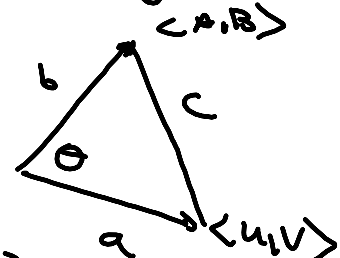
$c^2 - (a-x)^2 = b^2 - x^2$

$c^2 - \{a^2 - 2ax + \cancel{x^2}\} = b^2 - \cancel{x^2}$

$c^2 = a^2 + b^2 - 2ax$

$$\Rightarrow c^2 = a^2 + b^2 - 2ab\cos\theta$$

Law of Cosines




$$b = \sqrt{A^2+B^2}$$

$$a = \sqrt{u^2+v^2}$$

$$= 1$$

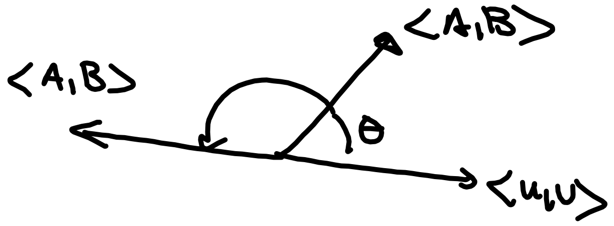
$$c^2 = (u-A)^2 + (v-B)^2$$

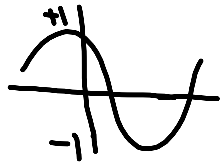
$$(u-A)^2 + (v-B)^2 = 1 + A^2 + B^2 - 2 \times 1 \times \sqrt{A^2+B^2}\cos\theta$$

$$\underline{u^2} - 2uA + \cancel{A^2} + \underline{v^2} - 2vB + \cancel{B^2} = \cancel{1} + \cancel{A^2} + \cancel{B^2} - 2\sqrt{A^2+B^2}\cos\theta$$

$$uA + vB = \sqrt{A^2+B^2}\cos\theta$$

$$uA + vB = \sqrt{A^2+B^2}\,\underset{\sim}{\cos\theta}$$




<u>Problem</u>: find $\langle u,v\rangle$ so that $\sqrt{A^2+B^2}\cos\theta$ is as small as possible!

<u>Ans</u>: When $\cos\theta = -1$, $\theta = \pi = 180^\circ$.

$$\langle u,v\rangle = -\frac{\langle A,B\rangle}{\sqrt{A^2+B^2}}$$

<u>Problem</u>: Find the direction $\langle u, v \rangle$
so that the rate of change $z'(0)$
is as small as possible.

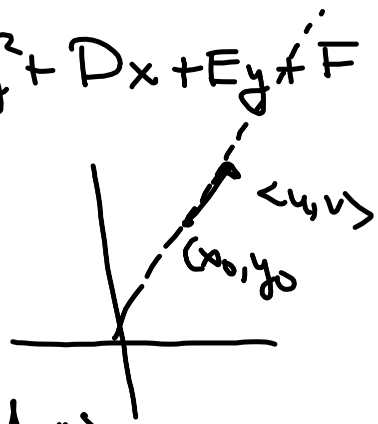
<u>Ans</u> (for $f(x,y) = Ax + By + C$).

is $\langle u, v \rangle = -\dfrac{\langle A, B \rangle}{\sqrt{A^2+B^2}}$

In general?

Try $f(x,y) = Ax^2 + Bxy + Cy^2 + Dx + Ey + F$

$z(t) = f(x_0 + tu, y_0 + tv)$




(1) Find $z'(0)$

(2) Find the direction $\langle u,v\rangle$ so that $z'(0)$ is as small as possible.

$$
\begin{aligned}
z(t) &= f(\underset{x}{x_0+tu}, \underset{y}{y_0+tv}) \\
&= A(x_0+tu)^2 + B(x_0+tu)(y_0+tv) \\
&\quad + C(y_0+tv)^2 + D(x_0+tu) + E(y_0+tv) + F \\
z'(t) &= 2A(x_0+tu)u + B\{u(y_0+tv) + (x_0+tu)v\} \\
&\quad + 2C(y_0+tv)v + Du + Ev + 0 \\
z'(0) &= 2Ax_0u + B\{uy_0 + vx_0\} \\
&\quad + 2Cy_0v + Du + Ev \\
&= A'u + B'v
\end{aligned}
$$

Best direction is

$$\langle u, v \rangle = -\frac{\langle A', B' \rangle}{\sqrt{(A')^2 + (B')^2}}$$

A recipe for $A'$ and $B'$ directly from $f(x,y)$.

$$f(x,y) = Ax + By + C.$$

$$\langle A', B' \rangle = \langle A, B \rangle$$

$$f(x,y) = Ax^2 + Bxy + Cy^2 + Dx + Ey + F$$

$$A' = \underline{2Ax_0} + By_0 + D = \left(\frac{\partial}{\partial x} f\right)(x_0, y_0)$$

$$B' = 2Cy_0 + Bx_0 + E = \left(\frac{\partial f}{\partial y}\right)(x_0, y_0)$$

The best direction in general is ↙ Called the Gradient

$$\langle u,v\rangle = -\frac{\left\langle \frac{\partial f}{\partial x}(x_0,y_0), \frac{\partial f}{\partial y}(x_0,y_0)\right\rangle}{\sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2}}$$

is the direction of
Steepest Descent

<u>Gradient</u> <u>Descent</u>