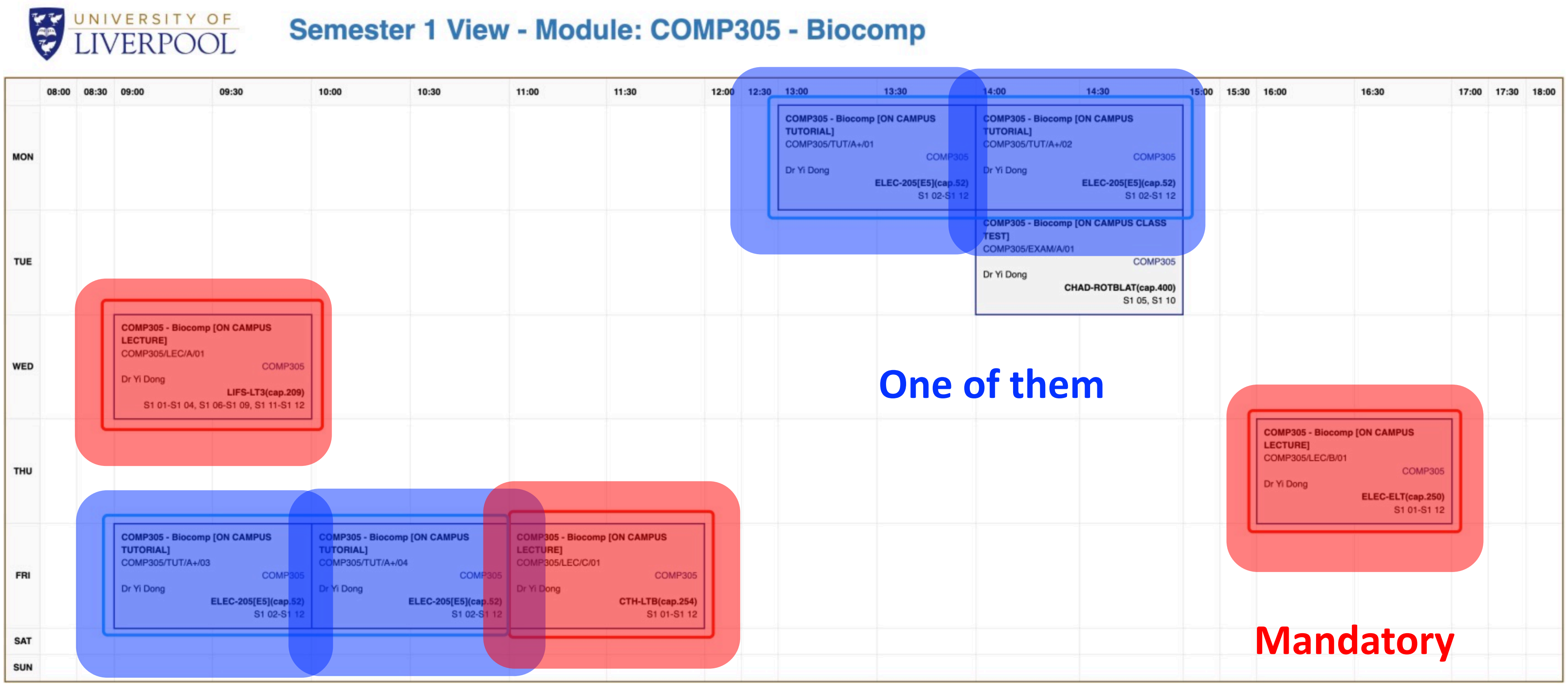
Comp305

Biocomputation

Lecturer: Yi Dong

Comp305 Module Timetable



There will be **26-30** lectures, thee per week. The lecture slides will appear on Canvas. Please use Canvas to access the lecture information. There will be **9** tutorials, one per week.

Lecture/Tutorial Rules

Questions are welcome as soon as they arise, because

1. Questions give feedback to the lecturer;
2. Questions help your understanding;
3. Your questions help your classmates, who might experience difficulties with formulating the same problems/doubts in the form of a question.

Comp305 Part I.

Artificial Neural Networks

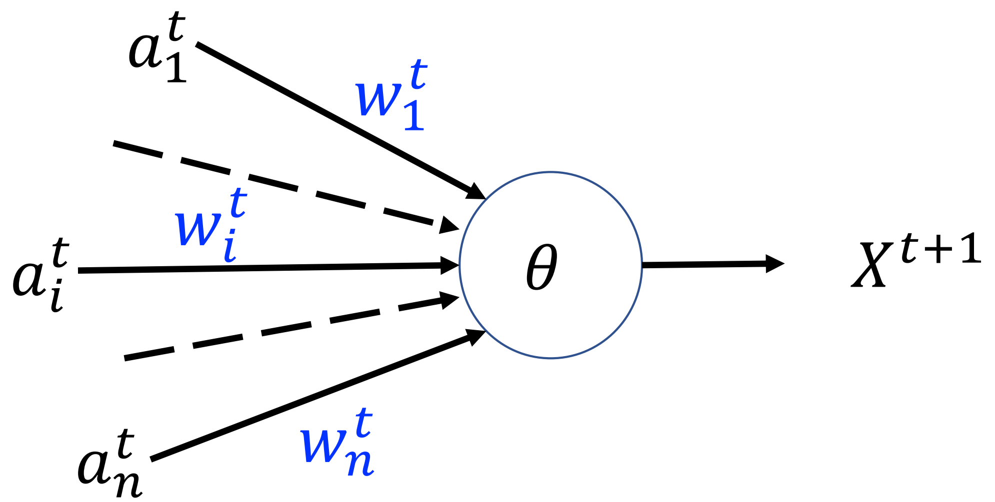
Topic 4.

Normalized Hebb's Rule (Oja's Rule)

Topic of Today’s Lecture

What is Oja’s rule and A Running Example.

Hebb's Rule (1949)



$$w_i^{t+1} = w_i^t + \Delta w_i^t,$$

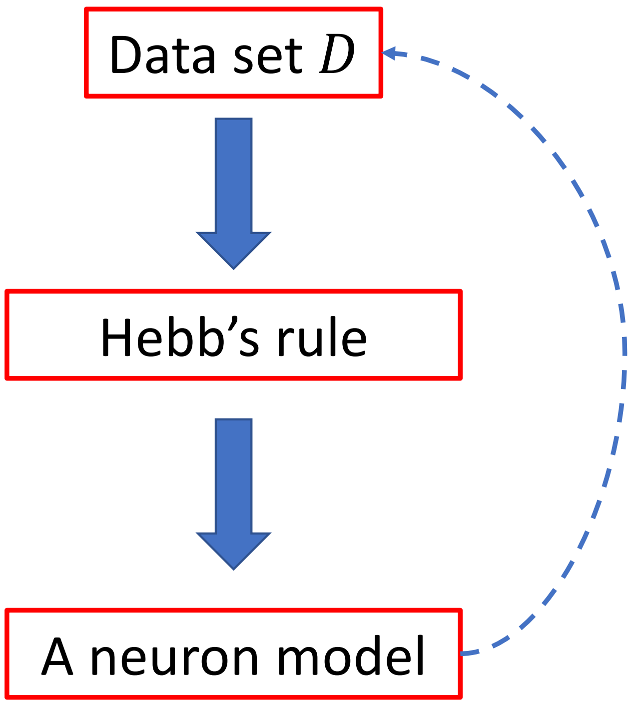
Where

$$\Delta w_i^t = C a_i^t X^{t+1}$$

Algorithm of Hebb's Rule for a Single Neuron

1. Set the neuron threshold value θ and the learning rate C.
2. Set *random initial values* for the weights of connections w_i^t.
3. Give instant input values a_i^t by the input units.
4. Compute the instant state of the neuron $S^t = \sum_i w_i^t a_i^t$
5. Compute the instant output of the neuron X^{t+1}
$$X^{t+1} = g(S^t) = H(S^t - \theta) = \begin{cases} 1, & S^t \geq \theta; \\ 0, & S^t < \theta. \end{cases}$$
6. Compute the instant corrections to the weights of connections $\Delta w_i^t = C a_i^t X^{t+1}$
7. Update the weights of connections $w_i^{t+1} = w_i^t + \Delta w_i^t$
8. **Go to the step 3**.

Meaning behind Hebb's Rule

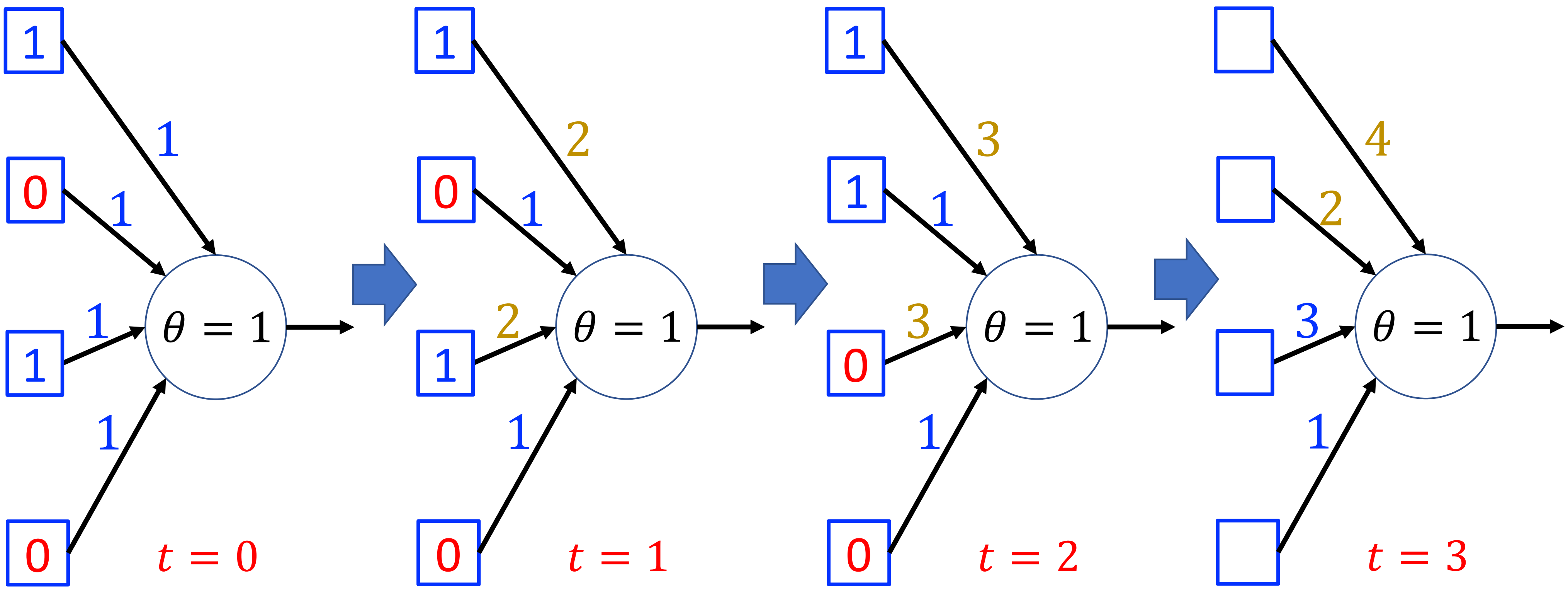


The **weights** of the neuron model after Hebbian learning, reflects **the most important feature of the average $\overline{a}$** of the data set D.

PCA

Principal components analysis

Theory behind Hebb’s Rule



Intuition: If two adjacent neurons always fire together, they should have strong relation (large weight).

Informal Explanation

Let the data set $D = \{\underbrace{[1,0,1,0]}_{\text{Input at } t=0,1}, \underbrace{[1,1,0,0]}_{\text{Input at } t=2}\}$. What is the characteristic of D?

Informal Explanation

Let the data set $D = \{\underbrace{[1,0,1,0]}_{\text{Input at } t\,=\,0,1}, \underbrace{[1,1,0,0]}_{\text{Input at } t\,=\,2}\}$. What is the characteristic of D?

By computing the "average value" of D,

$$\bar{a} = (2\times[1,0,1,0] + [1,1,0,0])/3 = \underline{\left[1, \frac{1}{3}, \frac{2}{3}, 0\right]},$$

we know that

- the first input (component) of the points in D is the most important;
- the third input is the second most important;
- the second is the third most important;
- the fourth input is never fired.

Informal Explanation

Let the data set $D = \{\underline{[1,0,1,0]}, \underline{[1,1,0,0]}\}$. What is the characteristic of D?

Input at $t = 0,1$ Input at $t = 2$

By computing the "average value" of D,

$$\bar{a} = (2\times[1,0,1,0] + [1,1,0,0])/3 = \underline{\left[1, \frac{1}{3}, \frac{2}{3}, 0\right]},$$

we know that

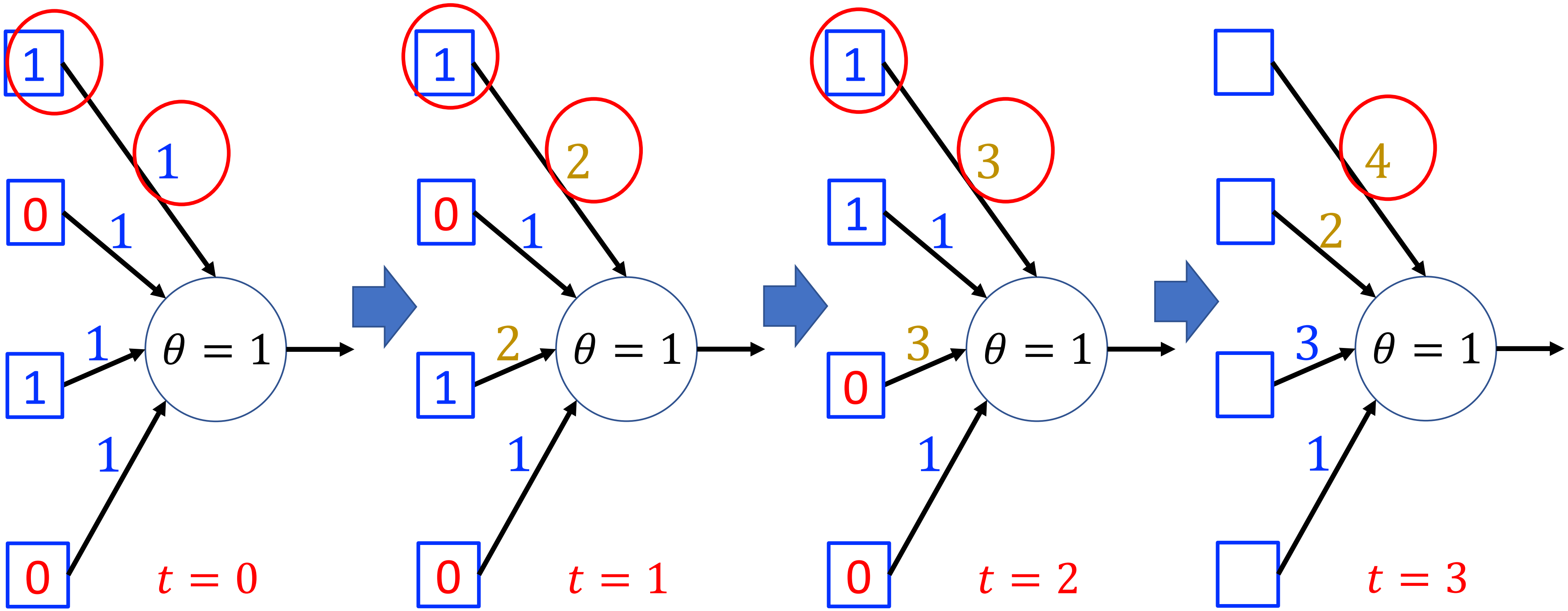
- the first input (component) of the points in D is the most important;
- the third input is the second most important;
- the second is the third most important;
- the fourth input is never fired.

$\boxed{\text{PCA}}$

PCA

Principal components analysis

Meaning behind Hebb's Rule



Intuition: The weights reflect the importance of the inputs. That is, the more important the input is, the larger the corresponding weight is.

Hebb's Rule (1949)

$$w_i^{t+1} = w_i^t + \Delta w_i^t,$$

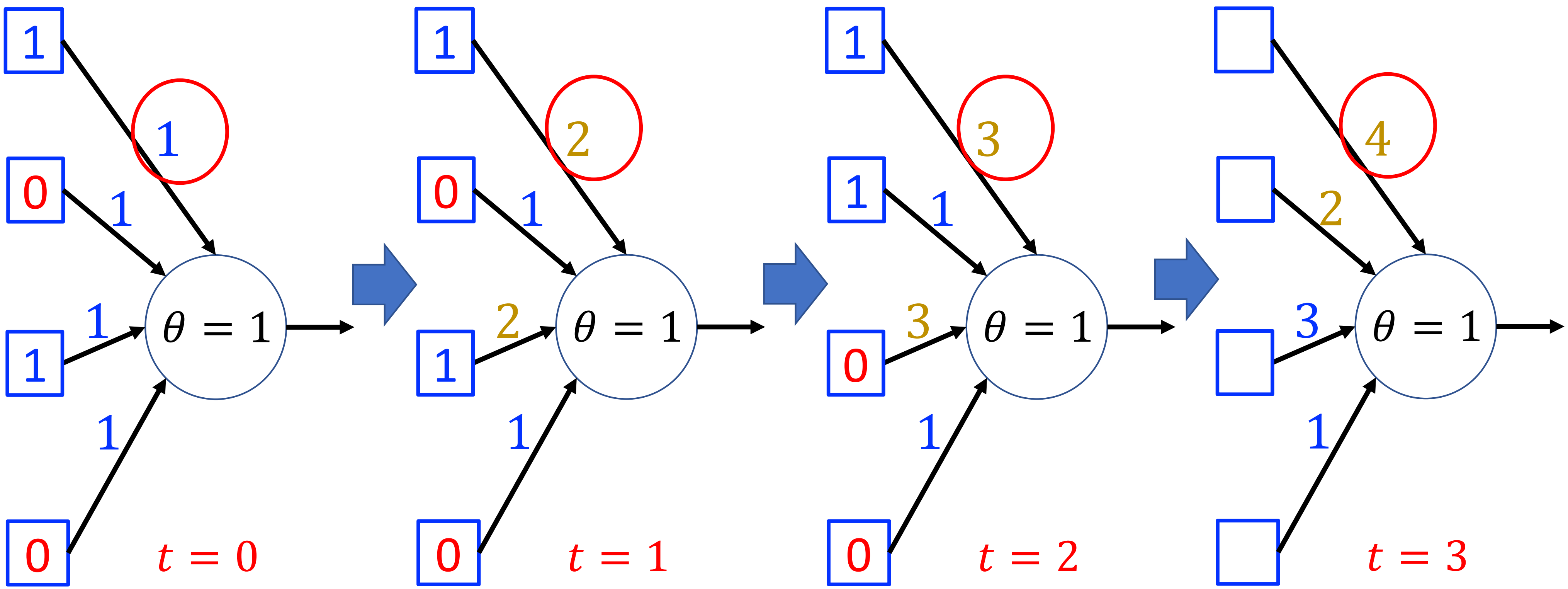
Where

$$\Delta w_i^t = C a_i^t X^{t+1}$$

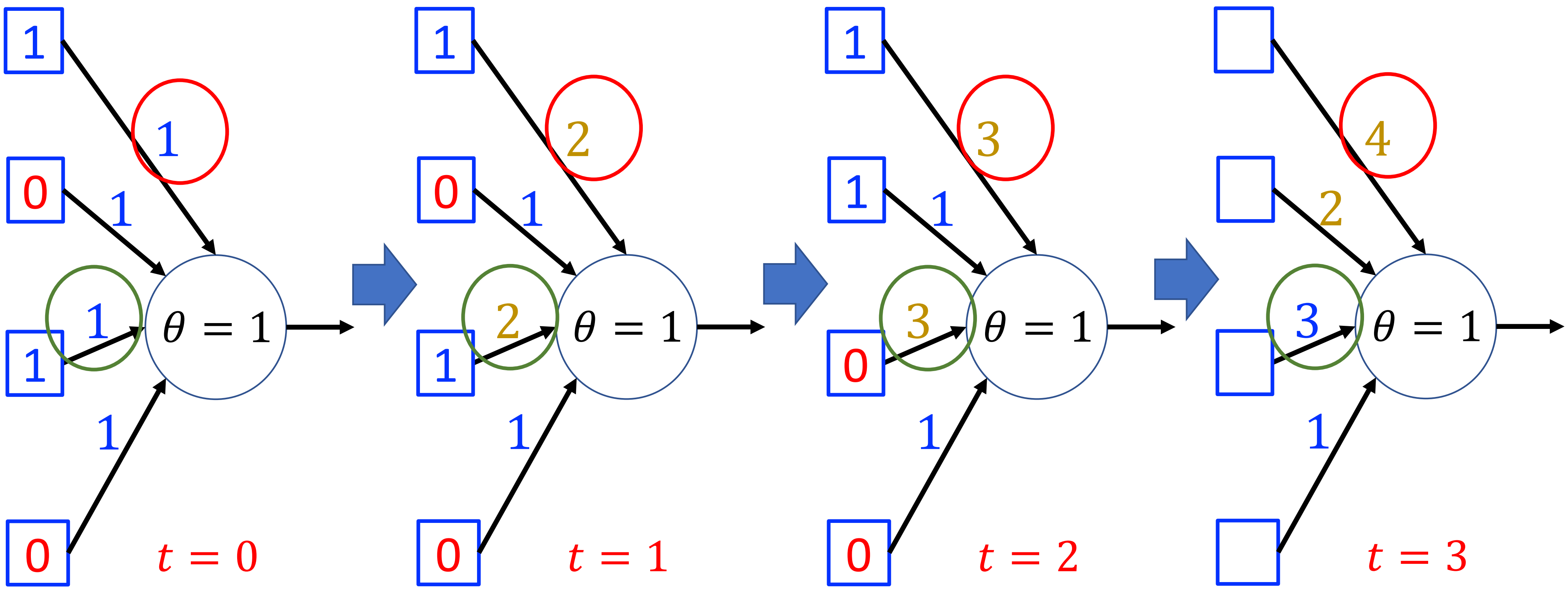
Hebb's original learning rule

has the unfortunate property that it can only monotonously increase synaptic weights, thus washing out the distinctive performance of different neurons, as the connections drive into saturation…

Example: w_1 and w_3



Example: w_1 and w_3



All the weights increase **monotonously**. Finally, each weight will become large enough such that any activated input can fire the neuron alone.

Hebb's Rule (1949)

$$w_i^{t+1} = w_i^t + \Delta w_i^t,$$

Where

$$\Delta w_i^t = C a_i^t X^{t+1}$$

However,

when the Hebbian rule is augmented by a ***normalisation*** rule, *e.g*. keeping constant the total strength of synapses upon a given neuron, it tends to "**sharpen**" a neuron's predisposition "**without a teacher**", causing its firing to become better correlated with a cluster of stimulus patterns.

Normalized Hebb's Rule (Oja's rule)

$$w_i^{t+1} = w_i^t + \Delta w_i^t,$$

Where

$$\Delta w_i^t = C a_i^t X^{t+1}$$

Normalized Hebb's Rule (Oja's rule)

$$w_i^{t+1} = w_i^t + \Delta w_i^t,$$

Where

$$\Delta w_i^t = C a_i^t X^{t+1}$$

Normalization:

$$\|\boldsymbol{w}^t\| = \mathbf{1}$$

- Also known as **Oja's rule**.
- By normalization, the weights will not monotonously increase, but converge after, which reflects the **predisposition to different inputs**. It plays an important role in ***unsupervised learning*** or ***self-organisation***.

Formulation of Hebb's Rule for a single neuron

1. Set the neuron threshold value θ and the learning rate C.
2. Set ***random initial values*** for the weights of connections w_i^t.
3. Give instant input values a_i^t by the input units.
4. Compute the instant state of the neuron $S^t = \sum_i w_i^t a_i^t$
5. Compute the instant output of the neuron X^{t+1}

$$X^{t+1} = g(S^t) = H(S^t - \theta) = \begin{cases} 1, & S^t \geq \theta; \\ 0, & S^t < \theta. \end{cases}$$

6. Compute the instant corrections to the weights $\Delta w_i^t = C a_i^t X^{t+1}$
7. Update the weights of connections $w_i^{t+1} = w_i^t + \Delta w_i^t$
8. Go to the step 3.

Formulation of Normalized Hebb's Rule (Oja's rule)

1. Set the neuron threshold value θ and the learning rate C.
2. Set ***random initial values*** for the weights of connections w_i^t.
3. **Normalization**.
4. Give instant input values a_i^t by the input units.
5. Compute the instant state of the neuron $S^t = \sum_i w_i^t a_i^t$
5. Compute the instant output of the neuron X^{t+1}

$$X^{t+1} = g(S^t) = H(S^t - \theta) = \begin{cases} 1, & S^t \geq \theta; \\ 0, & S^t < \theta. \end{cases}$$

6. Compute the instant corrections to the weights $\Delta w_i^t = C a_i^t X^{t+1}$
7. Update the weights of connections $w_i^{t+1} = w_i^t + \Delta w_i^t$
8. Go to the step 3.

Formulation of Normalized Hebb's Rule

1. Set the neuron threshold value θ and the learning rate C.
2. Set ***random initial values*** for the weights of connections w_i^t.
3. **Normalization.**
   I. Compute the 2-norm of the vector w^t

   $$\|w^t\|_2 = \sqrt{\sum_i \left(w_i^t\right)^2}$$

   II. Normalize the weight of each connection w_i^t

   $$w_i^t = \frac{1}{\|w^t\|_2} w_i^t$$

   III. Check the following **convergence criteria** with a given small positive real δ

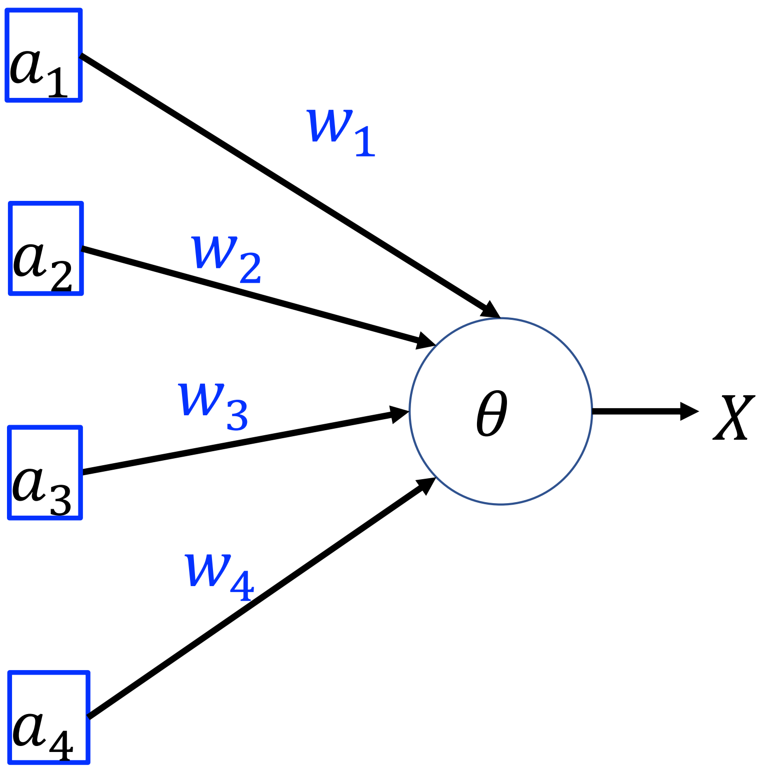
   $$\max_i \left|w_i^t - w_i^{t-1}\right| \le \delta$$

Alternative Version of Normalized Hebb's Rule

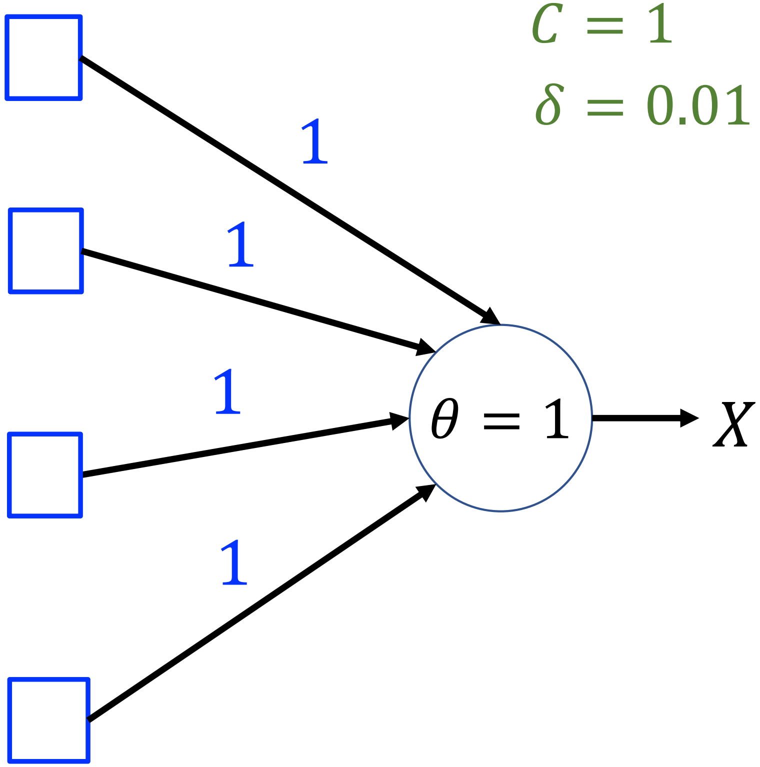
1. Set the neuron threshold value θ and the learning rate C.
2. Set ***random initial values*** for the weights of connections w_i^t.
3. Give instant input values a_i^t by the input units.
4. Compute the instant state of the neuron $S^t = \sum_i w_i^t a_i^t$
5. Compute the instant output of the neuron X^{t+1}
$$X^{t+1} = g(S^t) = H(S^t - \theta) = \begin{cases} 1, & S^t \geq \theta; \\ 0, & S^t < \theta. \end{cases}$$
6. **Compute the instant corrections to the weights $\Delta w_i^t = C\left(a_i^t - w_i^t X^{t+1}\right)X^{t+1}$**
7. **Check the convergence criteria**
8. Update the weights of connections $w_i^{t+1} = w_i^t + \Delta w_i^t$
9. Go to the step 3.

Derived from Taylor expansion at 0 in terms of C and reducing the high-order term when C is small.

A Running Example



A Running Example



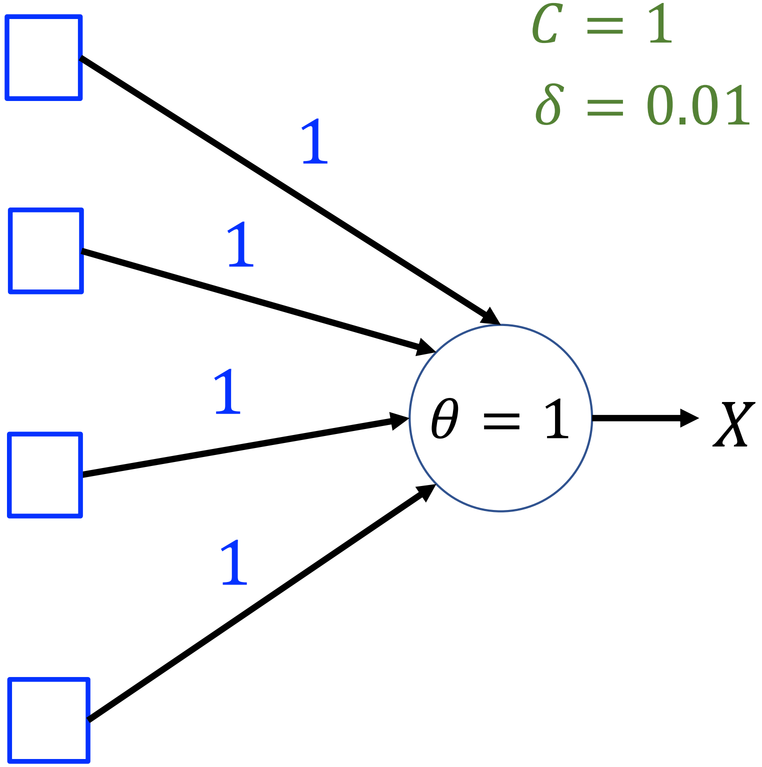
$C = 1$

$\delta = 0.01$

$t = 0$

w_1^0	w_2^0	w_3^0	w_4^0
1	1	1	1

A Running Example



$C = 1$

$\delta = 0.01$

$t = 0$

w_1^0	w_2^0	w_3^0	w_4^0
1	1	1	1

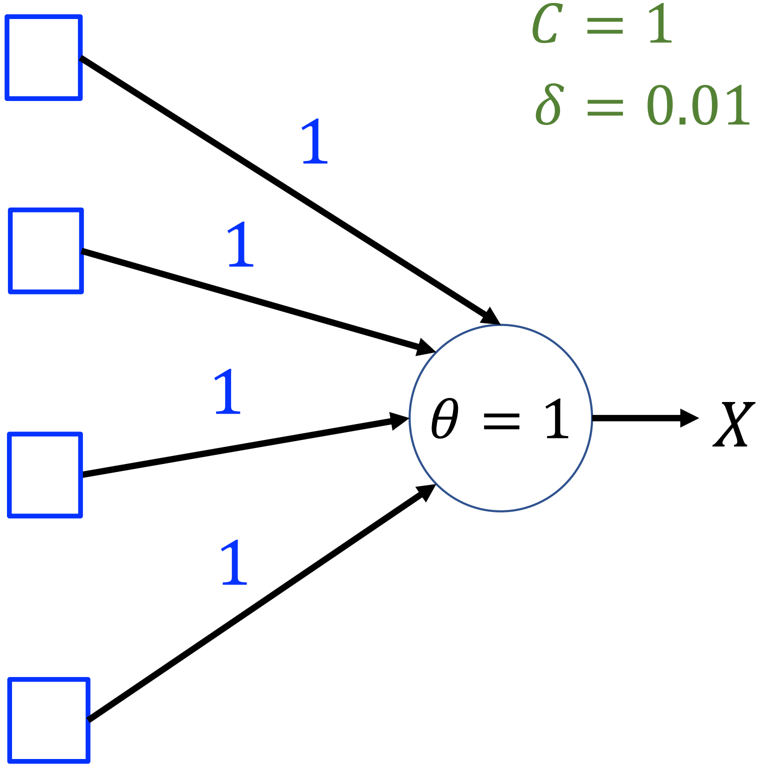
$$\|w^0\|_2 = \sqrt{\sum_i \left(w_i^0\right)^2}$$

$$= \sqrt{(w_1^0)^2 + (w_2^0)^2 + (w_3^0)^2 + (w_4^0)^2}$$

$$= \sqrt{(1)^2 + (1)^2 + (1)^2 + (1)^2}$$

$$= 2$$

A Running Example



$C = 1$ $\quad t = 0$

$\delta = 0.01$

w_1^0	w_2^0	w_3^0	w_4^0
1	1	1	1

$$\|w^0\|_2 = 2$$

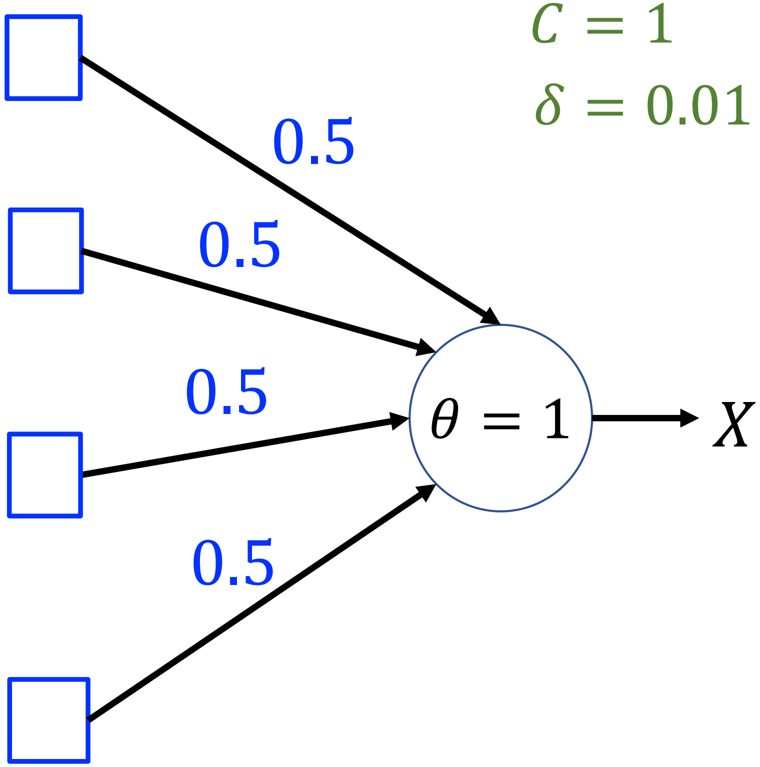
$$w_1^0 = \frac{1}{\|w^0\|_2} w_1^0 = \frac{1}{2} \qquad w_2^0 = \frac{1}{\|w^0\|_2} w_2^0 = \frac{1}{2}$$

$$w_3^0 = \frac{1}{\|w^0\|_2} w_3^0 = \frac{1}{2} \qquad w_4^0 = \frac{1}{\|w^0\|_2} w_4^0 = \frac{1}{2}$$

A Running Example

$C = 1$ $\quad t = 0$

$\delta = 0.01$



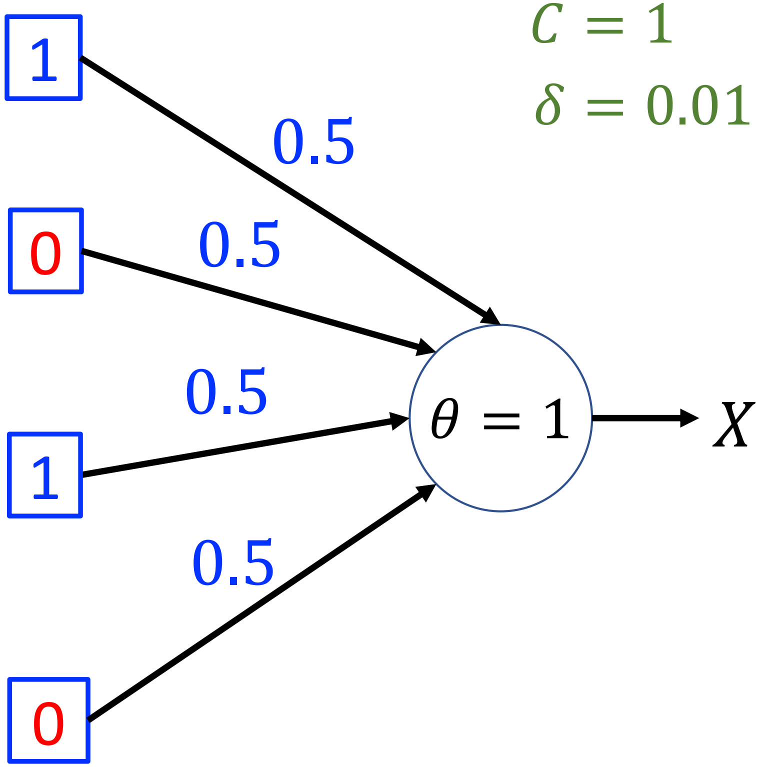
w_1^0	w_2^0	w_3^0	w_4^0
0.5	0.5	0.5	0.5

$\|w^0\|_2 = 2$

$$w_1^0 = \frac{1}{\|w^0\|_2} w_1^0 = \frac{1}{2} \qquad w_2^0 = \frac{1}{\|w^0\|_2} w_2^0 = \frac{1}{2}$$

$$w_3^0 = \frac{1}{\|w^0\|_2} w_3^0 = \frac{1}{2} \qquad w_4^0 = \frac{1}{\|w^0\|_2} w_4^0 = \frac{1}{2}$$

A Running Example



$C = 1$

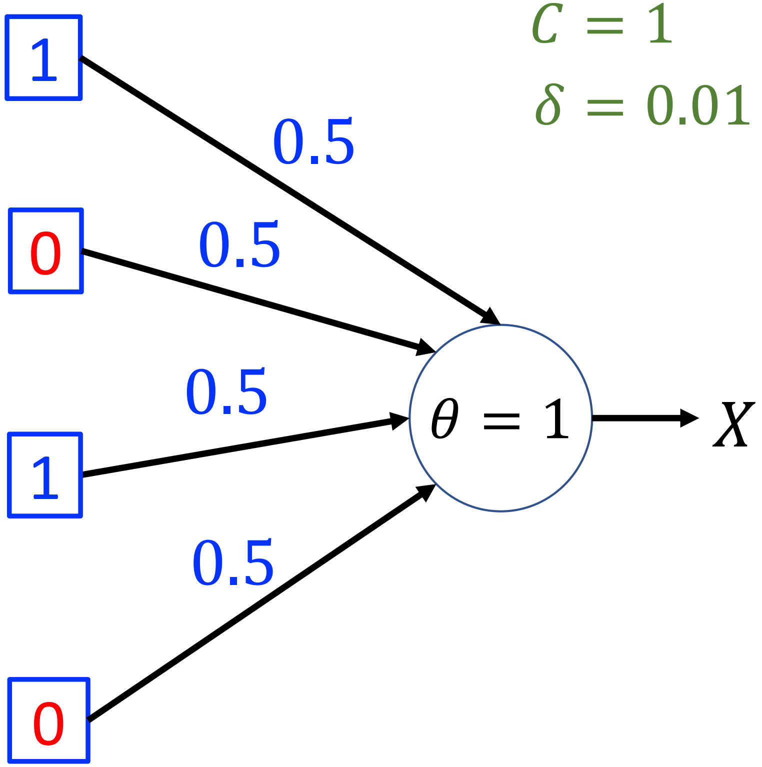
$\delta = 0.01$

$t = 0$

a_1^0	a_2^0	a_3^0	a_4^0
1	0	1	0

w_1^0	w_2^0	w_3^0	w_4^0
0.5	0.5	0.5	0.5

A Running Example



$C = 1$

$\delta = 0.01$

$t = 0$

a_1^0	a_2^0	a_3^0	a_4^0
1	0	1	0

w_1^0	w_2^0	w_3^0	w_4^0
0.5	0.5	0.5	0.5

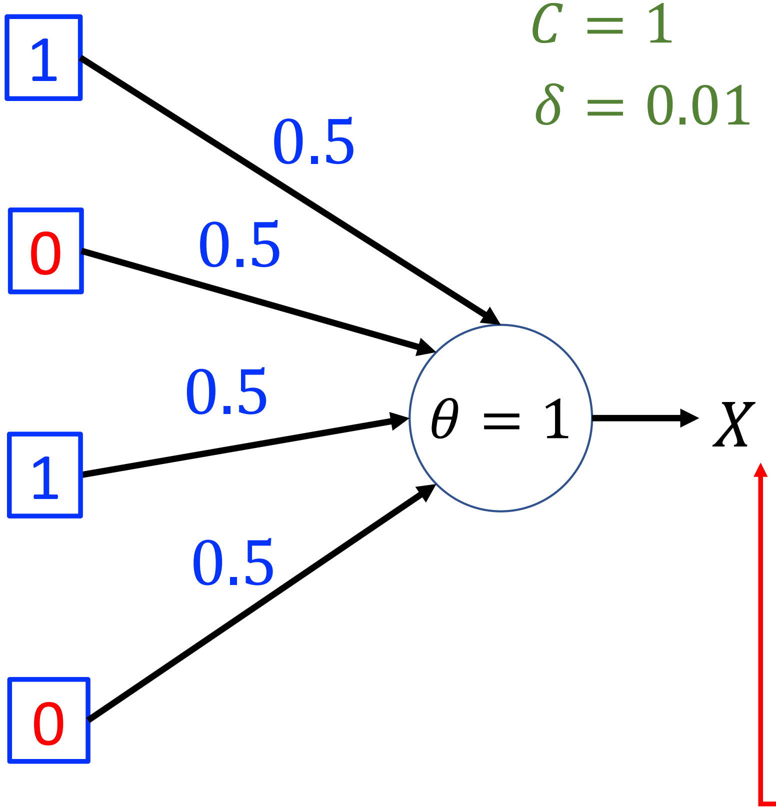
$$S^0 = \sum_{i=1}^{4} w_i^0 a_i^0$$

$$= w_1^0 \times a_1^0 + w_2^0 \times a_2^0 + w_3^0 \times a_3^0 + w_4^0 \times a_4^0$$

$$= 0.5\times 1 + 0.5\times 0 + 0.5\times 1 + 0.5\times 0 = 1 \geq \theta$$

$$X^1 = 1$$

A Running Example



$C = 1$

$\delta = 0.01$

$t = 0$

a_1^0	a_2^0	a_3^0	a_4^0
1	0	1	0

w_1^0	w_2^0	w_3^0	w_4^0
0.5	0.5	0.5	0.5

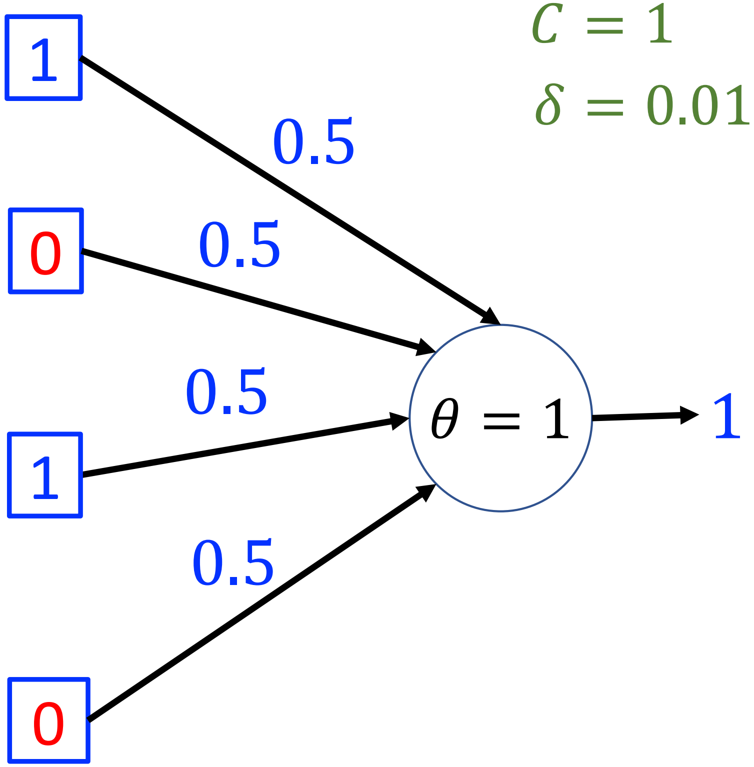
$$S^0 = \sum_{i=1}^{4} w_i^0 a_i^0$$

$$= w_1^0 \times a_1^0 + w_2^0 \times a_2^0 + w_3^0 \times a_3^0 + w_4^0 \times a_4^0$$

$$= 0.5 \times 1 + 0.5 \times 0 + 0.5 \times 1 + 0.5 \times 0 = 1 \geq \theta$$

$$X^1 = 1$$

A Running Example



$C = 1$

$\delta = 0.01$

$t = 0$

a_1^0	a_2^0	a_3^0	a_4^0
1	0	1	0

w_1^0	w_2^0	w_3^0	w_4^0
0.5	0.5	0.5	0.5

$$\Delta w_i^0 = C a_i^0 X^1$$

$$w_i^1 = w_i^0 + \Delta w_i^0$$

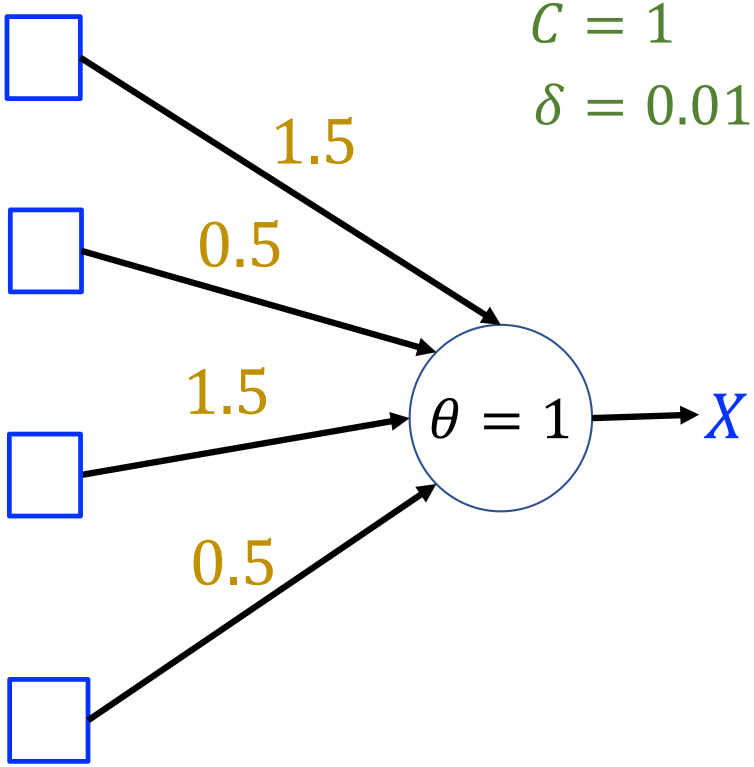
$\Delta w_1^0 = 1\times1\times1 = 1,$ $w_1^1 = w_i^0 + \Delta w_i^0 = 0.5 + 1 = 1.5;$

$\Delta w_2^0 = 1\times0\times1 = 0,$ $w_2^1 = w_2^0 + \Delta w_2^0 = 0.5 + 0 = 0.5;$

$\Delta w_3^0 = 1\times1\times1 = 1,$ $w_3^1 = w_3^0 + \Delta w_3^0 = 0.5 + 1 = 1.5;$

$\Delta w_4^0 = 1\times0\times1 = 0,$ $w_4^1 = w_4^0 + \Delta w_4^0 = 0.5 + 0 = 0.5;$

A Running Example



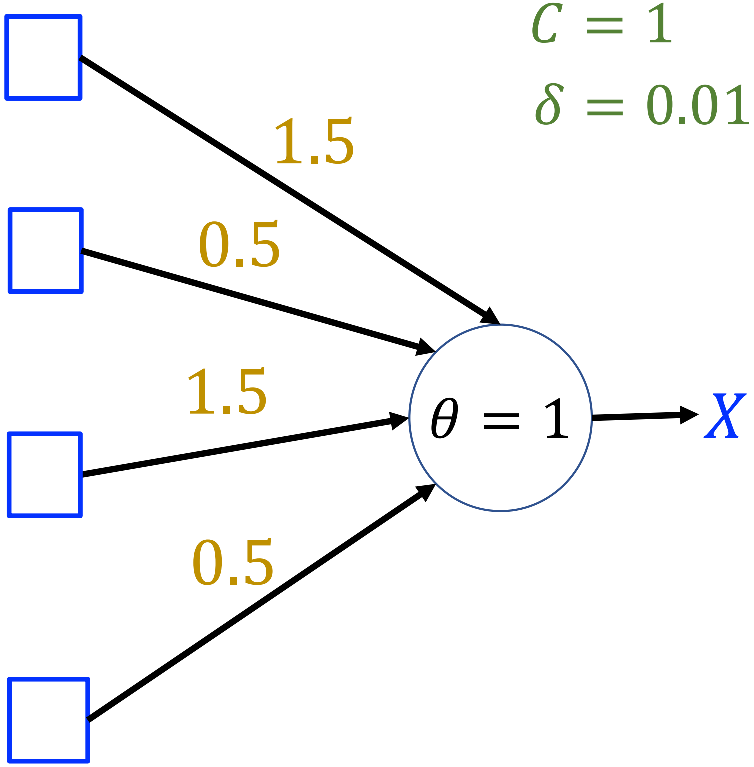
$C = 1$

$\delta = 0.01$

$t = 1$

w_1^1	w_2^1	w_3^1	w_4^1
1.5	0.5	1.5	0.5

A Running Example



$C = 1$ $\quad t = 1$

$\delta = 0.01$

$\boldsymbol{w_1^1}$	$\boldsymbol{w_2^1}$	$\boldsymbol{w_3^1}$	$\boldsymbol{w_4^1}$
1.5	0.5	1.5	0.5

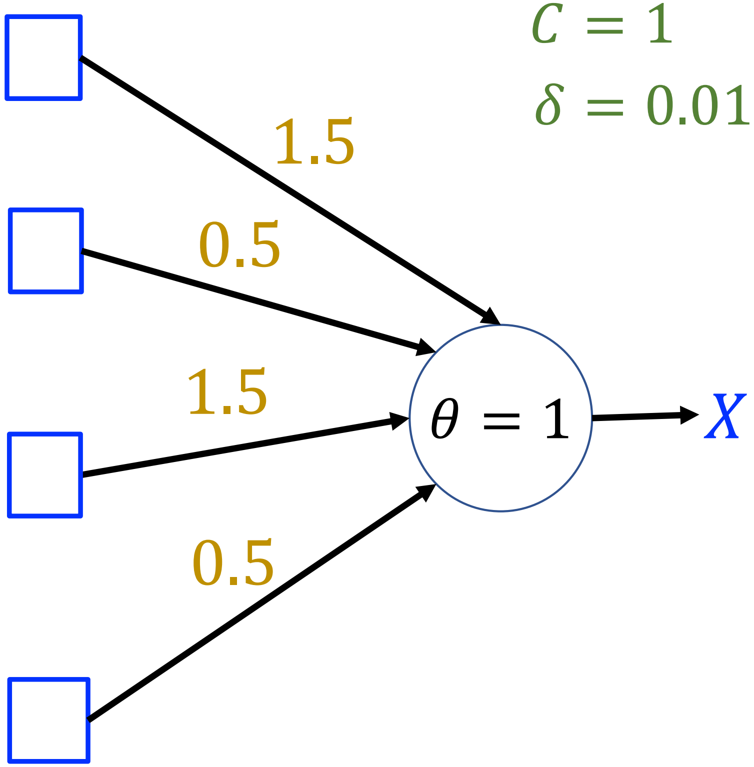
$$\|w^1\|_2 = \sqrt{\sum_i \left(w_i^1\right)^2}$$

$$= \sqrt{(w_1^1)^2 + (w_2^1)^2 + (w_3^1)^2 + (w_4^1)^2}$$

$$= \sqrt{(1.5)^2 + (0.5)^2 + (1.5)^2 + (0.5)^2}$$

$$= \sqrt{5} \approx 2.24$$

A Running Example



$C = 1$

$\delta = 0.01$

$t = 1$

w_1^1	w_2^1	w_3^1	w_4^1
1.5	0.5	1.5	0.5

$\|w^1\|_2 \approx 2.24$

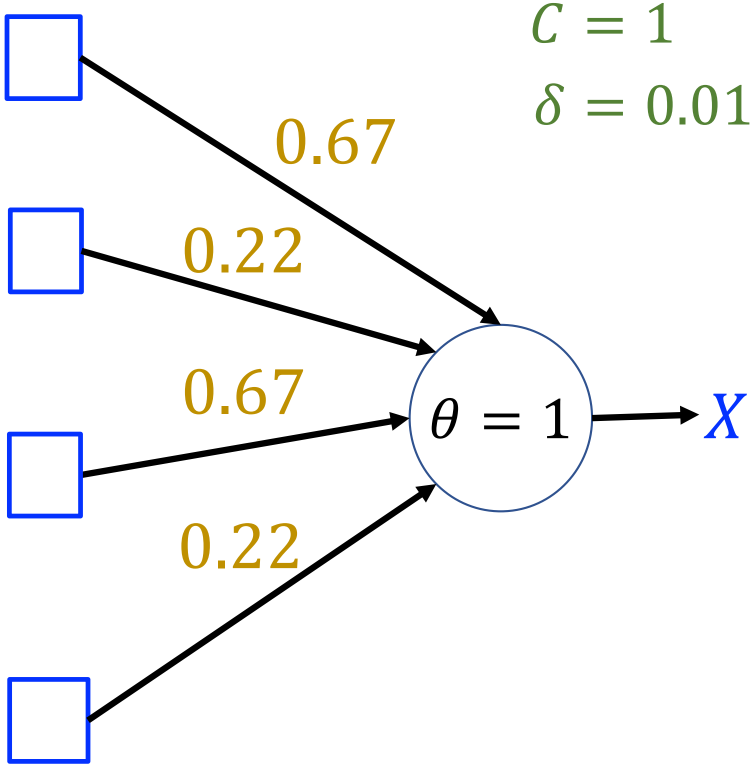
$$w_1^1 = \frac{1}{\|w^1\|_2} w_1^1 = 0.67 \quad w_2^1 = \frac{1}{\|w^1\|_2} w_2^1 = 0.22$$

$$w_3^1 = \frac{1}{\|w^1\|_2} w_3^1 = 0.67 \quad w_4^1 = \frac{1}{\|w^1\|_2} w_4^1 = 0.22$$

A Running Example

$C = 1$ $\quad t = 1$

$\delta = 0.01$



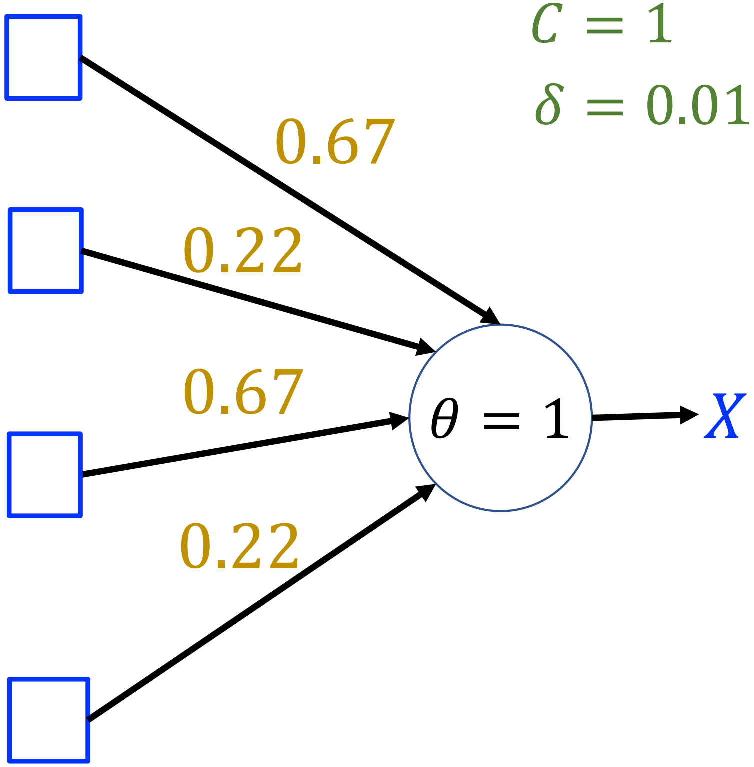
$\boldsymbol{w_1^1}$	$\boldsymbol{w_2^1}$	$\boldsymbol{w_3^1}$	$\boldsymbol{w_4^1}$
0.67	0.22	0.67	0.22

$\|w^1\|_2 \approx 2.24$

$$w_1^1 = \frac{1}{\|w^1\|_2} w_1^1 = 0.67 \quad w_2^1 = \frac{1}{\|w^1\|_2} w_2^1 = 0.22$$

$$w_3^1 = \frac{1}{\|w^1\|_2} w_3^1 = 0.67 \quad w_4^1 = \frac{1}{\|w^1\|_2} w_4^1 = 0.22$$

A Running Example



$C = 1$

$\delta = 0.01$

$t = 1$

w_1^0	w_2^0	w_3^0	w_4^0
0.5	0.5	0.5	0.5

w_1^1	w_2^1	w_3^1	w_4^1
0.67	0.22	0.67	0.22

$\max_i |w_i^1 - w_i^0| \leq \delta$ **?**

$|w_1^1 - w_1^0| = 0.17$

$|w_2^1 - w_2^0| = 0.28$

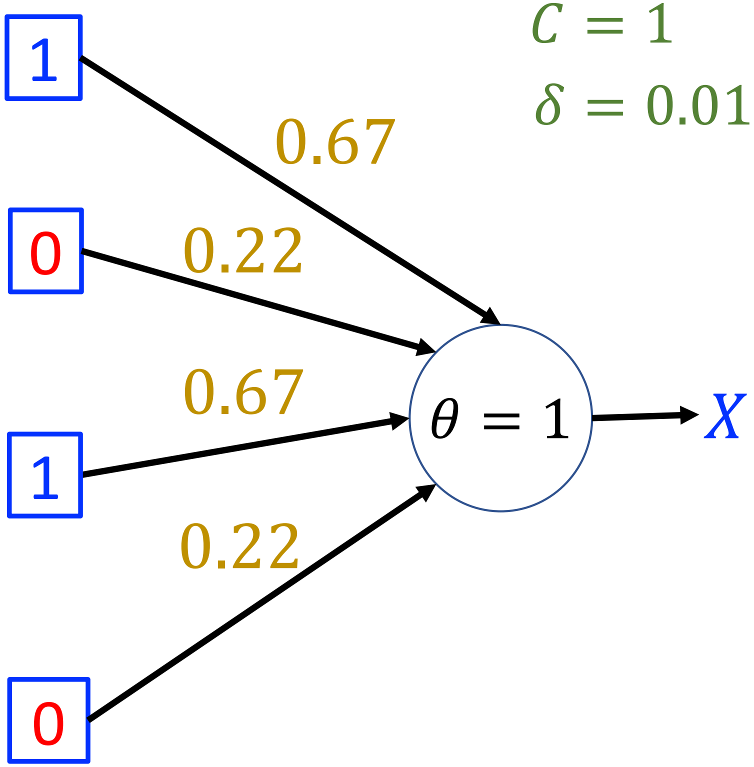
$|w_3^1 - w_3^0| = 0.17$

$|w_4^1 - w_4^0| = 0.28$

$\Rightarrow \max_i |w_i^1 - w_i^0| = 0.28 > \delta$

Convergence criteria is not met. Continue.

A Running Example



$C = 1$

$\delta = 0.01$

$t = 1$

a_1^1	a_2^1	a_3^1	a_4^1
1	0	1	0

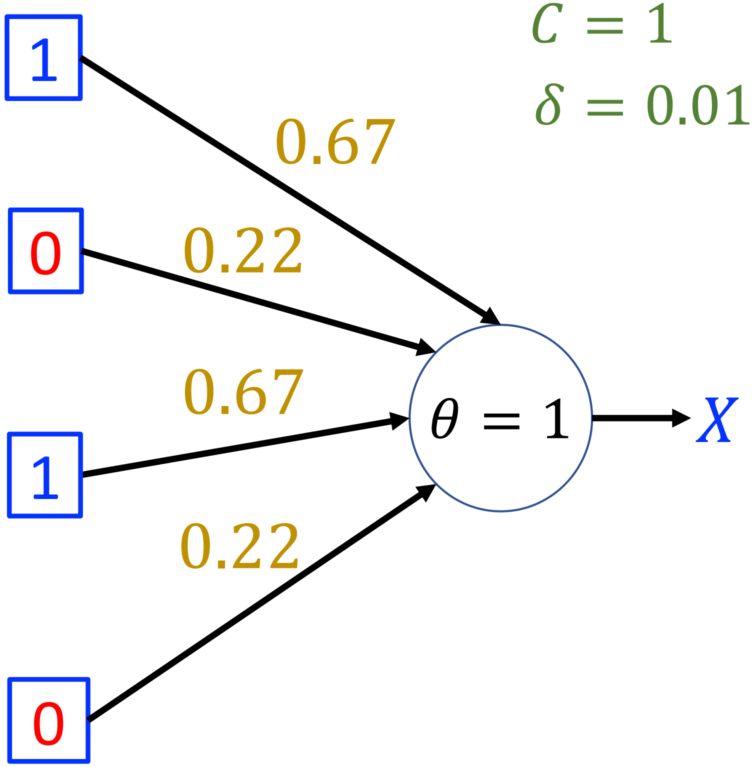
w_1^1	w_2^1	w_3^1	w_4^1
0.67	0.22	0.67	0.22

A Running Example

$C = 1$

$\delta = 0.01$

$t = 1$



a_1^1	a_2^1	a_3^1	a_4^1
1	0	1	0

w_1^1	w_2^1	w_3^1	w_4^1
0.67	0.22	0.67	0.22

$$S^1 = \sum_{i=1}^{4} w_i^1 a_i^1$$

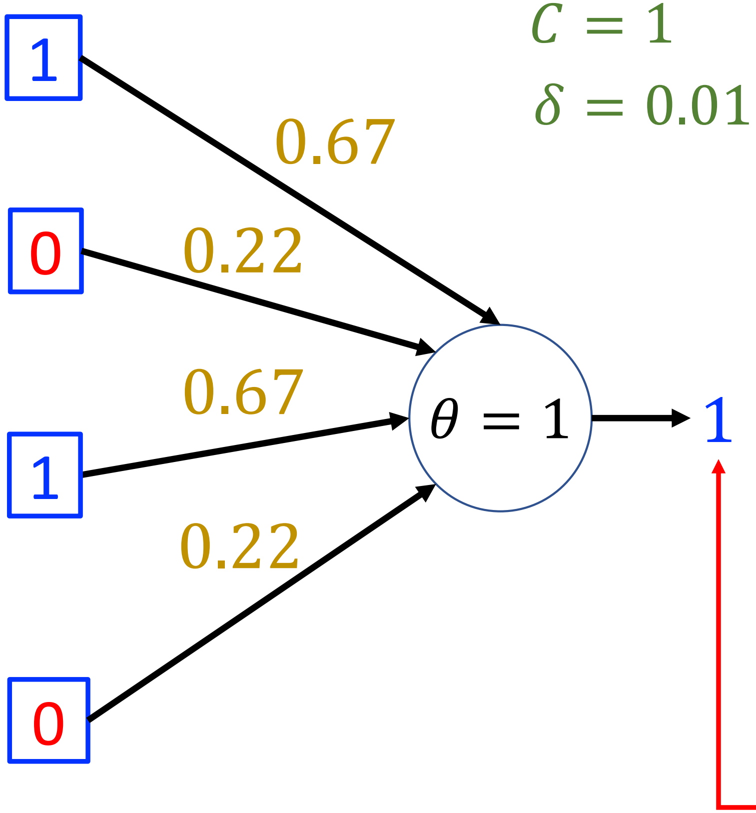
$$= w_1^1 \times a_1^1 + w_2^1 \times a_2^1 + w_3^1 \times a_3^1 + w_4^1 \times a_4^1$$

$$= 0.67 \times 1 + 0.22 \times 0 + 0.67 \times 1 + 0.22 \times 0$$

$$= 1.34 \geq \theta$$

$$X^2 = 1$$

A Running Example



$C = 1$

$\delta = 0.01$

$t = 1$

a_1^1	a_2^1	a_3^1	a_4^1
1	0	1	0

w_1^1	w_2^1	w_3^1	w_4^1
0.67	0.22	0.67	0.22

$$S^1 = \sum_{i=1}^{4} w_i^1 a_i^1$$

$$= w_1^1 \times a_1^1 + w_2^1 \times a_2^1 + w_3^1 \times a_3^1 + w_4^1 \times a_4^1$$

$$= 0.67 \times 1 + 0.22 \times 0 + 0.67 \times 1 + 0.22 \times 0$$

$$= 1.34 \geq \theta$$

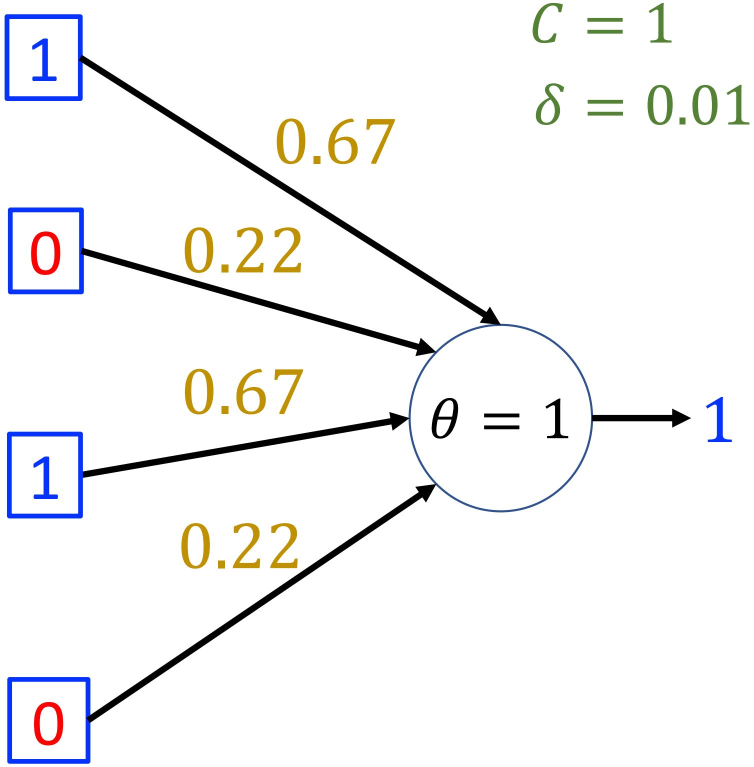
$$X^2 = 1$$

A Running Example

$C = 1$

$\delta = 0.01$

$t = 1$



a_1^1	a_2^1	a_3^1	a_4^1
1	0	1	0

w_1^1	w_2^1	w_3^1	w_4^1
0.67	0.22	0.67	0.22

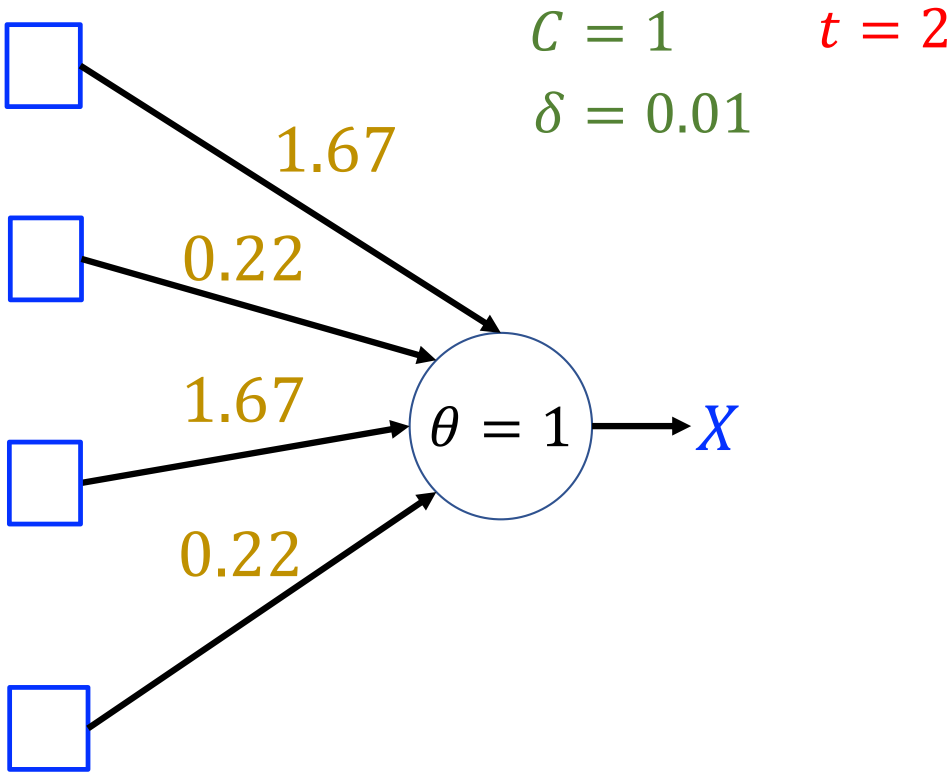
$$\Delta w_i^1 = C a_i^1 X^2$$

$$\Delta w_1^1 = 1 \times 1 \times 1 = 1,$$
$$\Delta w_2^1 = 1 \times 0 \times 1 = 0,$$
$$\Delta w_3^1 = 1 \times 1 \times 1 = 1,$$
$$\Delta w_4^1 = 1 \times 0 \times 1 = 0,$$

$$w_i^2 = w_i^1 + \Delta w_i^1$$

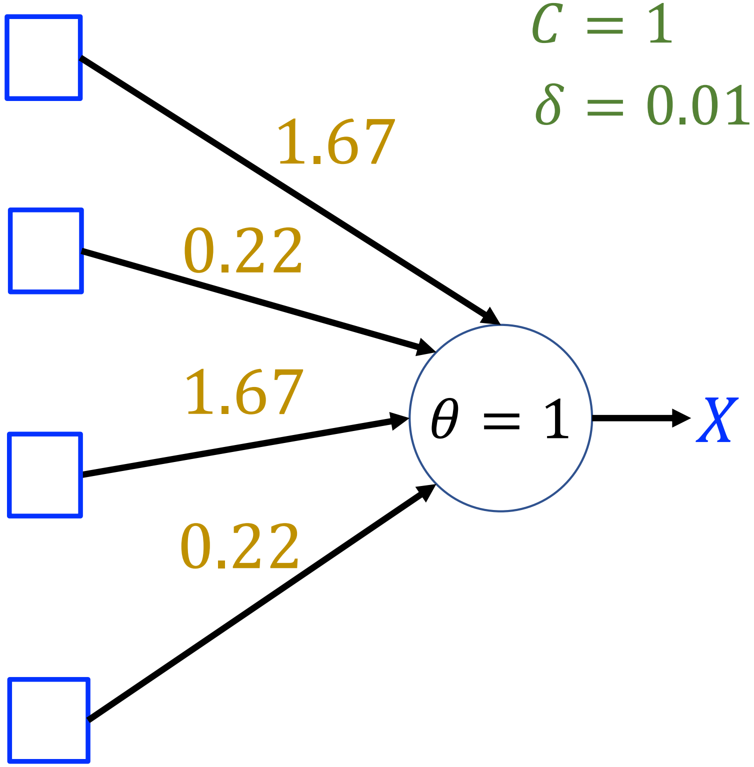
$$w_1^2 = w_i^1 + \Delta w_i^1 = 0.67 + 1 = 1.67;$$
$$w_2^2 = w_2^1 + \Delta w_2^1 = 0.22 + 0 = 0.22;$$
$$w_3^2 = w_3^1 + \Delta w_3^1 = 0.67 + 1 = 1.67;$$
$$w_4^2 = w_4^1 + \Delta w_4^1 = 0.22 + 0 = 0.22;$$

A Running Example



w_1^2	w_2^2	w_3^2	w_4^2
1.67	0.22	1.67	0.22

A Running Example



$C = 1$ $\quad t = 2$

$\delta = 0.01$

w_1^2	w_2^2	w_3^2	w_4^2
1.67	0.22	1.67	0.22

$$\|w^2\|_2 = \sqrt{\sum_i (w_i^2)^2}$$

$$= \sqrt{(w_1^2)^2 + (w_2^2)^2 + (w_3^2)^2 + (w_4^2)^2}$$

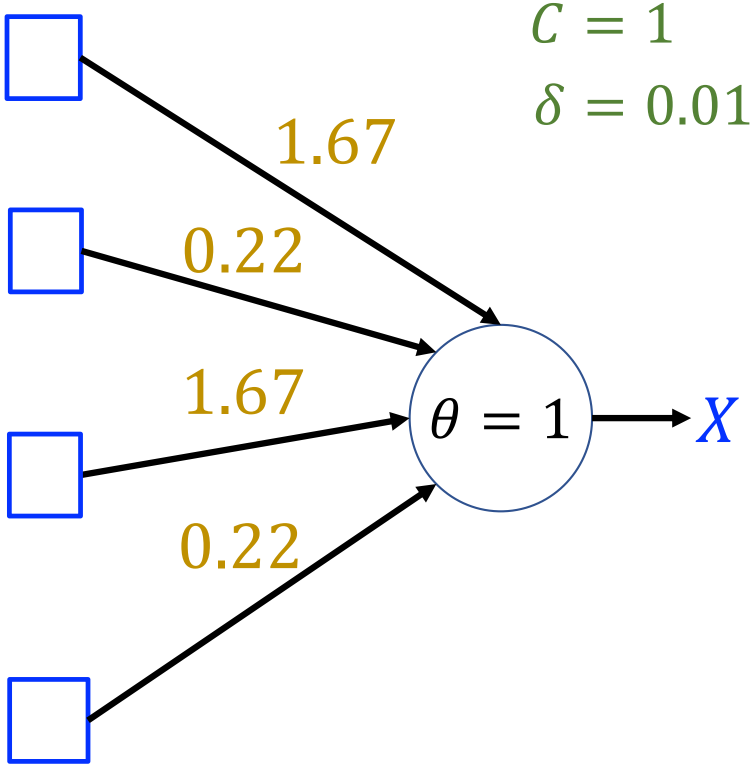
$$= \sqrt{(1.67)^2 + (0.22)^2 + (1.67)^2 + (0.22)^2}$$

$$\approx 2.38$$

A Running Example

$C = 1$ $\quad t = 2$

$\delta = 0.01$



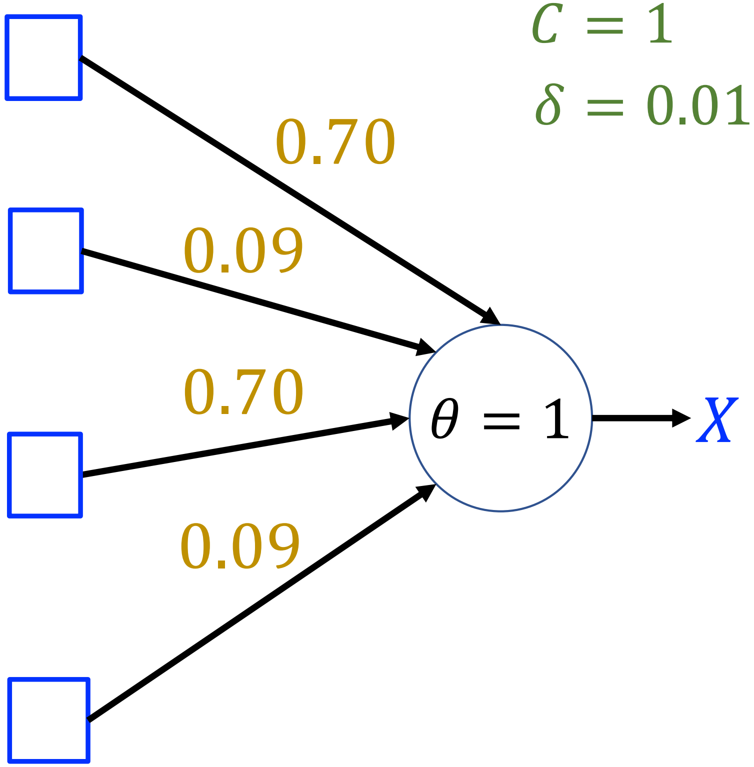
w_1^2	w_2^2	w_3^2	w_4^2
1.67	0.22	1.67	0.22

$$\|w^2\|_2 \approx 2.38$$

$$w_1^2 = \frac{1}{\|w^2\|_2} w_1^2 = 0.70 \qquad w_2^2 = \frac{1}{\|w^2\|_2} w_2^2 = 0.09$$

$$w_3^2 = \frac{1}{\|w^2\|_2} w_3^2 = 0.70 \qquad w_4^2 = \frac{1}{\|w^2\|_2} w_4^2 = 0.09$$

A Running Example



$C = 1$

$\delta = 0.01$

$t = 2$

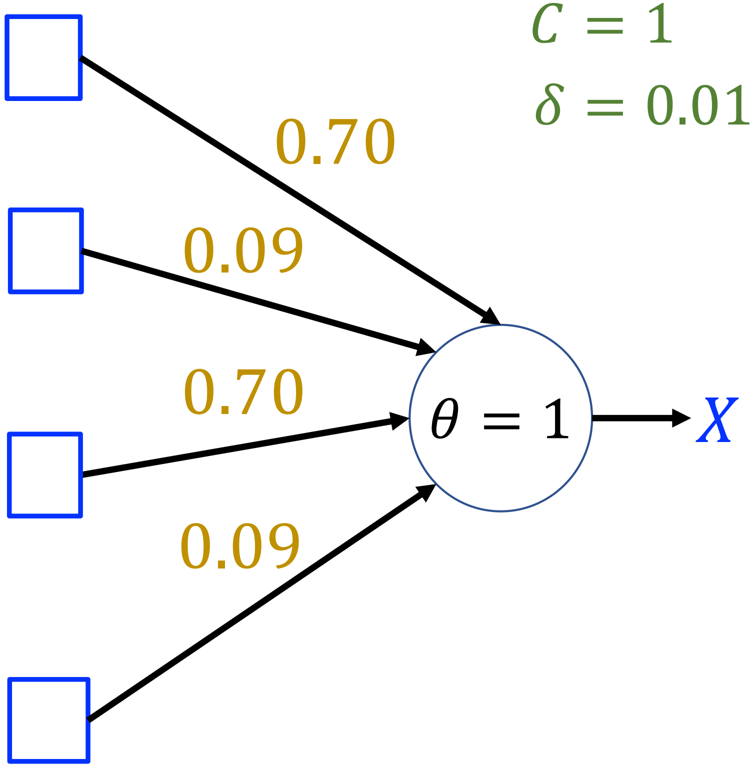
w_1^2	w_2^2	w_3^2	w_4^2
0.70	0.09	0.70	0.09

$\|w^2\|_2 \approx 2.38$

$$w_1^2 = \frac{1}{\|w^2\|_2} w_1^2 = 0.70 \quad w_2^2 = \frac{1}{\|w^2\|_2} w_2^2 = 0.09$$

$$w_3^2 = \frac{1}{\|w^2\|_2} w_3^2 = 0.70 \quad w_4^2 = \frac{1}{\|w^2\|_2} w_4^2 = 0.09$$

A Running Example



$C = 1$

$\delta = 0.01$

$t = 2$

w_1^1	w_2^1	w_3^1	w_4^1
0.67	0.22	0.67	0.22

w_1^2	w_2^2	w_3^2	w_4^2
0.70	0.09	0.70	0.09

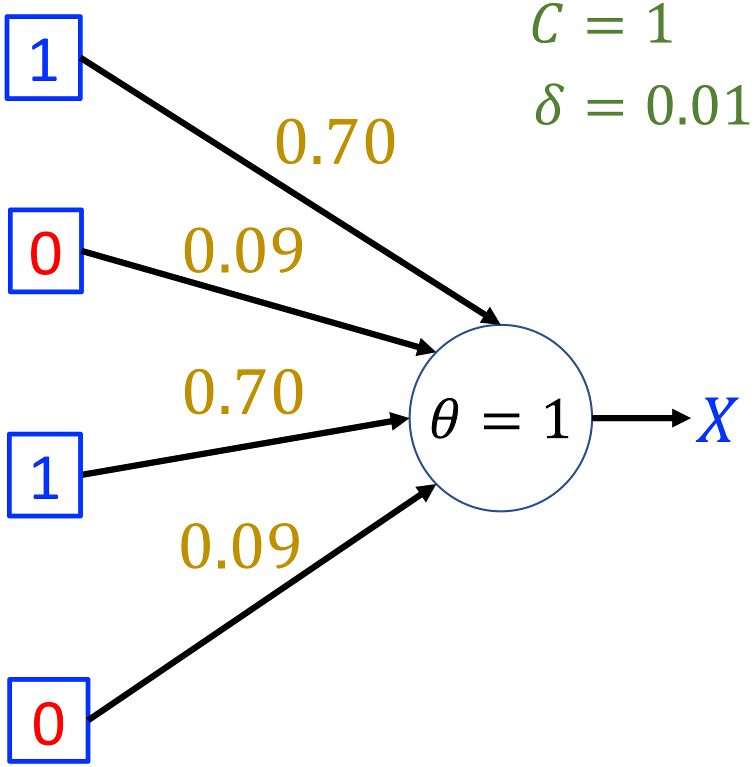
$\max_i |w_i^2 - w_i^1| \le \delta$ **?**

$$|w_1^2 - w_1^1| = 0.03$$
$$|w_2^2 - w_2^1| = 0.13$$
$$|w_3^2 - w_3^1| = 0.03$$
$$|w_4^2 - w_4^1| = 0.13$$

$$\Rightarrow \max_i |w_i^2 - w_i^1| = 0.13 > \delta$$

Long way to go. Continue.

A Running Example



$C = 1$

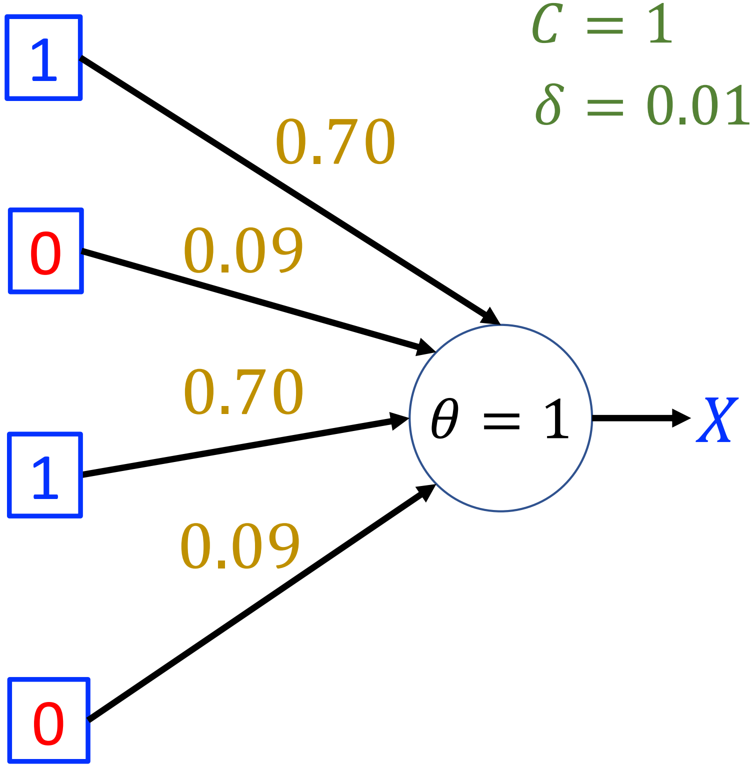
$\delta = 0.01$

$t = 2$

a_1^2	a_2^2	a_3^2	a_4^2
1	0	1	0

w_1^2	w_2^2	w_3^2	w_4^2
0.70	0.09	0.70	0.09

A Running Example



$C = 1$

$\delta = 0.01$

$t = 2$

a_1^2	a_2^2	a_3^2	a_4^2
1	0	1	0

w_1^2	w_2^2	w_3^2	w_4^2
0.70	0.09	0.70	0.09

$$S^2 = \sum_{i=1}^{4} w_i^2 a_i^2$$

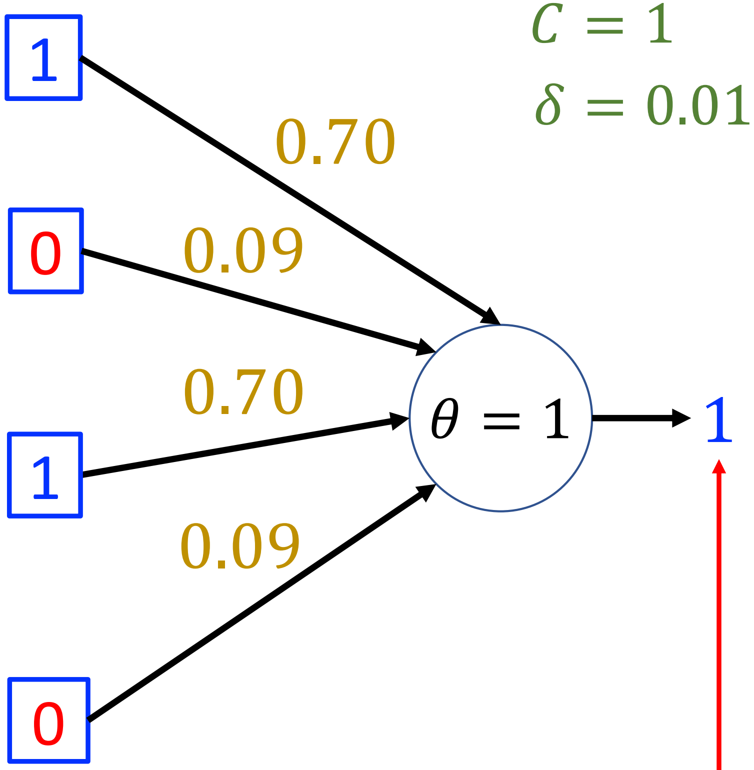
$$= w_1^2 \times a_1^2 + w_2^2 \times a_2^2 + w_3^2 \times a_3^2 + w_4^2 \times a_4^2$$

$$= 0.70 \times 1 + 0.09 \times 0 + 0.70 \times 1 + 0.09 \times 0$$

$$= 1.40 \geq \theta$$

$$X^3 = 1$$

A Running Example



$C = 1$

$\delta = 0.01$

$t = 2$

a_1^2	a_2^2	a_3^2	a_4^2
1	0	1	0

w_1^2	w_2^2	w_3^2	w_4^2
0.70	0.09	0.70	0.09

$$S^2 = \sum_{i=1}^{4} w_i^2 a_i^2$$

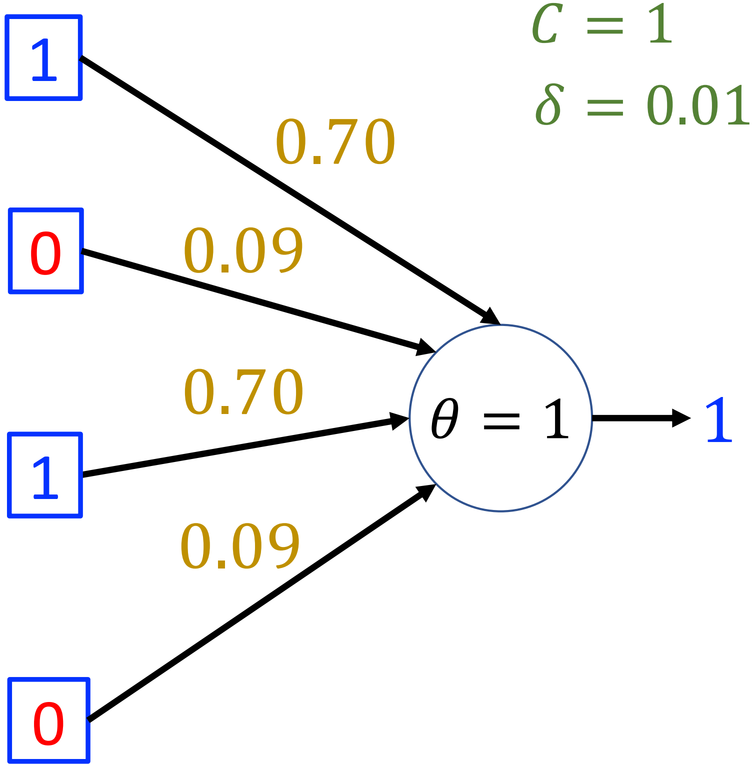
$$= w_1^2 \times a_1^2 + w_2^2 \times a_2^2 + w_3^2 \times a_3^2 + w_4^2 \times a_4^2$$

$$= 0.70 \times 1 + 0.09 \times 0 + 0.70 \times 1 + 0.09 \times 0$$

$$= 1.40 \geq \theta$$

$$X^3 = 1$$

A Running Example



$C = 1$

$\delta = 0.01$

$t = 2$

a_1^2	a_2^2	a_3^2	a_4^2
1	0	1	0

w_1^2	w_2^2	w_3^2	w_4^2
0.70	0.09	0.70	0.09

$$\Delta w_i^2 = C a_i^2 X^3$$

$$w_i^3 = w_i^2 + \Delta w_i^2$$

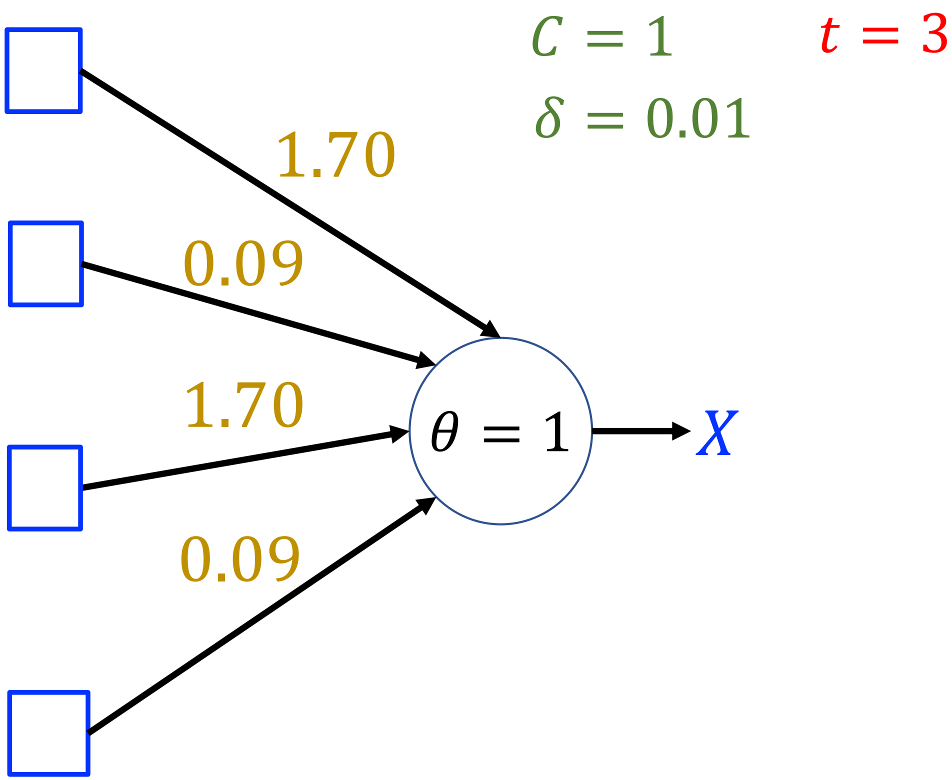
$\Delta w_1^2 = 1 \times 1 \times 1 = 1,$ $\quad w_1^3 = w_i^2 + \Delta w_i^2 = 0.70 + 1 = 1.70;$

$\Delta w_2^2 = 1 \times 0 \times 1 = 0,$ $\quad w_2^3 = w_2^2 + \Delta w_2^2 = 0.09 + 0 = 0.09;$

$\Delta w_3^2 = 1 \times 1 \times 1 = 1,$ $\quad w_3^3 = w_3^2 + \Delta w_3^2 = 0.70 + 1 = 1.70;$

$\Delta w_4^2 = 1 \times 0 \times 1 = 1,$ $\quad w_4^3 = w_4^2 + \Delta w_4^2 = 0.09 + 0 = 0.09;$

A Running Example



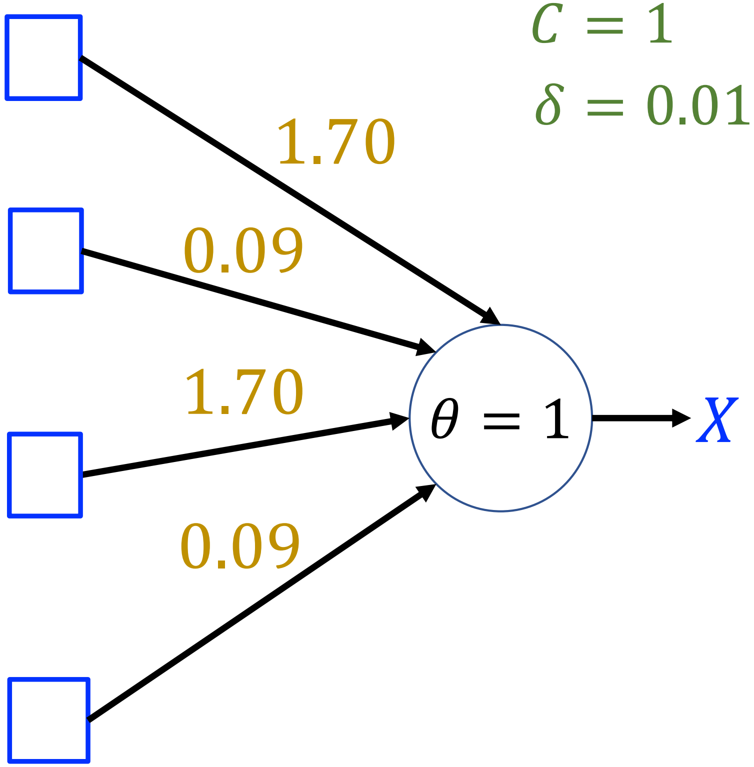
w_1^3	w_2^3	w_3^3	w_4^3
1.70	0.09	1.70	0.09

A Running Example

$C = 1$

$\delta = 0.01$

$t = 3$



w_1^3	w_2^3	w_3^3	w_4^3
1.70	0.09	1.70	0.09

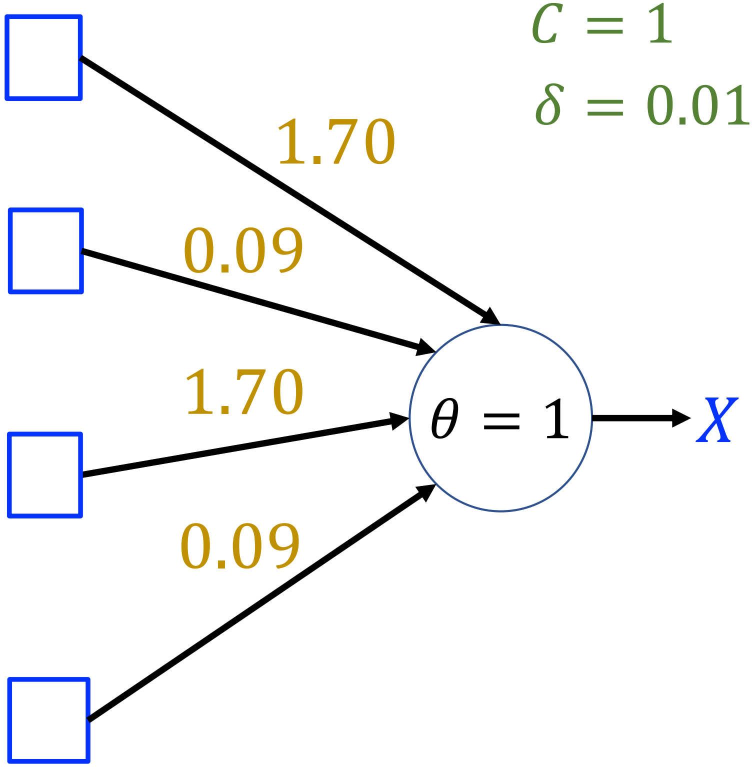
$$\|w^3\|_2 = \sqrt{\sum_i (w_i^3)^2}$$

$$= \sqrt{(w_1^3)^2 + (w_2^3)^2 + (w_3^3)^2 + (w_4^3)^2}$$

$$= \sqrt{(1.70)^2 + (0.09)^2 + (1.70)^2 + (0.09)^2}$$

$$\approx 2.41$$

A Running Example



$C = 1$

$\delta = 0.01$

$t = 3$

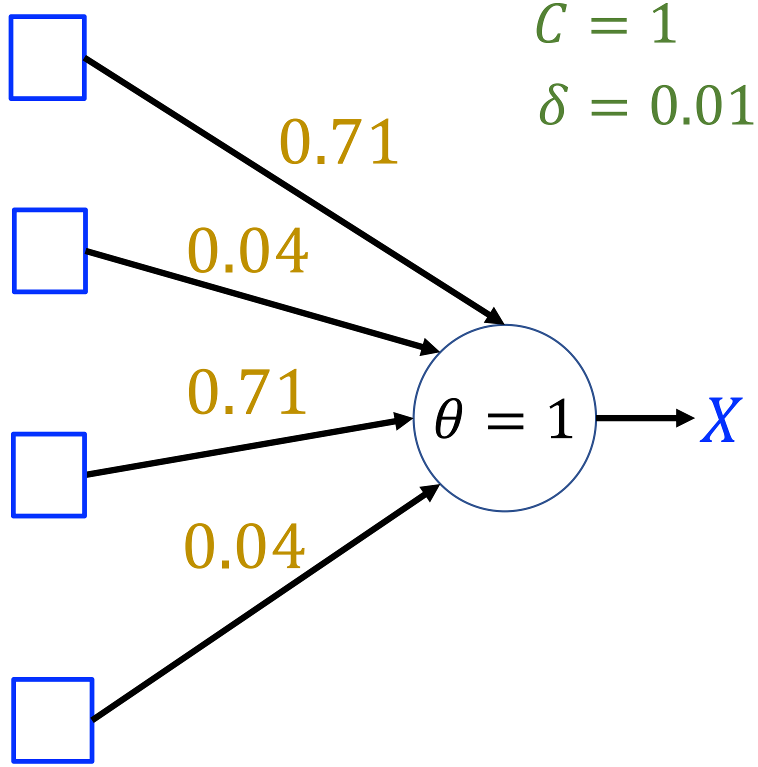
w_1^3	w_2^3	w_3^3	w_4^3
1.70	0.09	1.70	0.09

$\|w^3\|_2 \approx 2.41$

$$w_1^3 = \frac{1}{\|w^3\|_2} w_1^3 = 0.71 \quad w_2^3 = \frac{1}{\|w^3\|_2} w_2^3 = 0.04$$

$$w_3^3 = \frac{1}{\|w^3\|_2} w_3^3 = 0.71 \quad w_4^3 = \frac{1}{\|w^3\|_2} w_4^3 = 0.04$$

A Running Example



$C = 1$

$\delta = 0.01$

$t = 3$

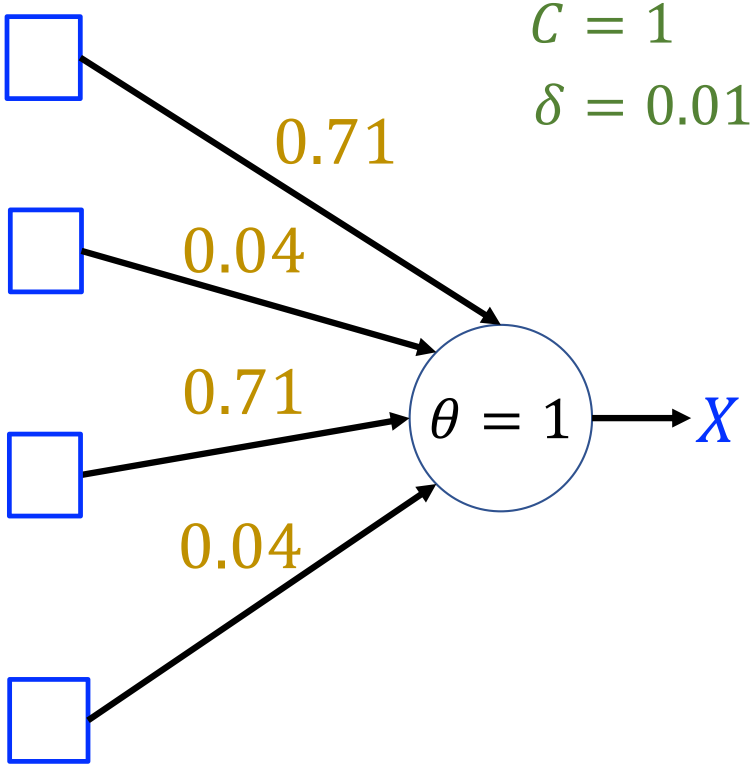
w_1^3	w_2^3	w_3^3	w_4^3
0.71	0.04	0.71	0.04

$\|w^3\|_2 \approx 2.41$

$$w_1^3 = \frac{1}{\|w^3\|_2} w_1^3 = 0.71 \qquad w_2^3 = \frac{1}{\|w^3\|_2} w_2^3 = 0.04$$

$$w_3^3 = \frac{1}{\|w^3\|_2} w_3^3 = 0.71 \qquad w_4^3 = \frac{1}{\|w^3\|_2} w_4^3 = 0.04$$

A Running Example



$C = 1$

$\delta = 0.01$

$t = 3$

w_1^2	w_2^2	w_3^2	w_4^2
0.70	0.09	0.70	0.09

w_1^3	w_2^3	w_3^3	w_4^3
0.71	0.04	0.71	0.04

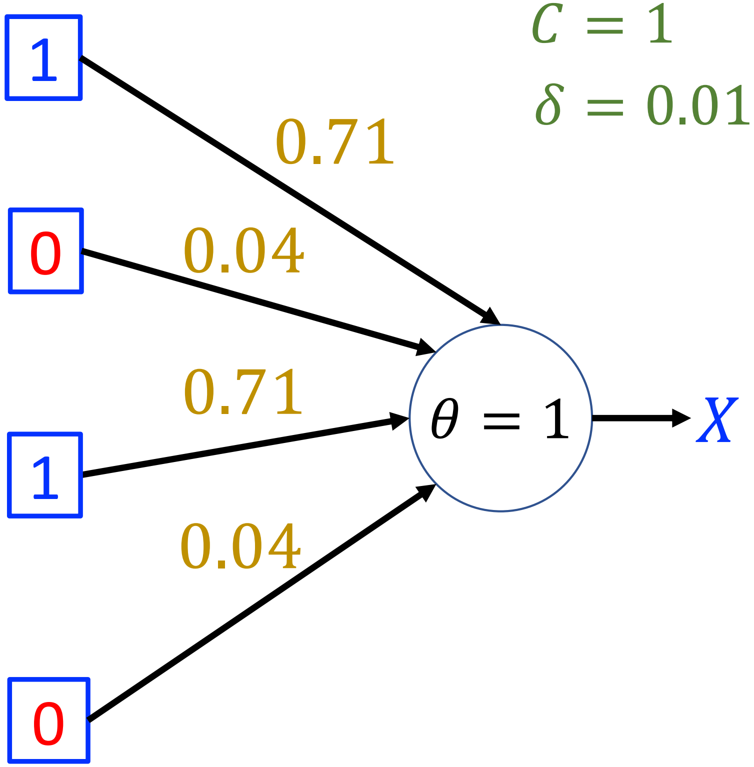
$\max_i |w_i^3 - w_i^2| \leq \delta$ **?**

$$|w_1^3 - w_1^2| = 0.01$$
$$|w_2^3 - w_2^2| = 0.05$$
$$|w_3^3 - w_3^2| = 0.01$$
$$|w_4^3 - w_4^2| = 0.05$$

$\Rightarrow$ $\max_i |w_i^3 - w_i^2| = 0.05 > \delta$

Smaller, but not enough. Continue.

A Running Example



$C = 1$

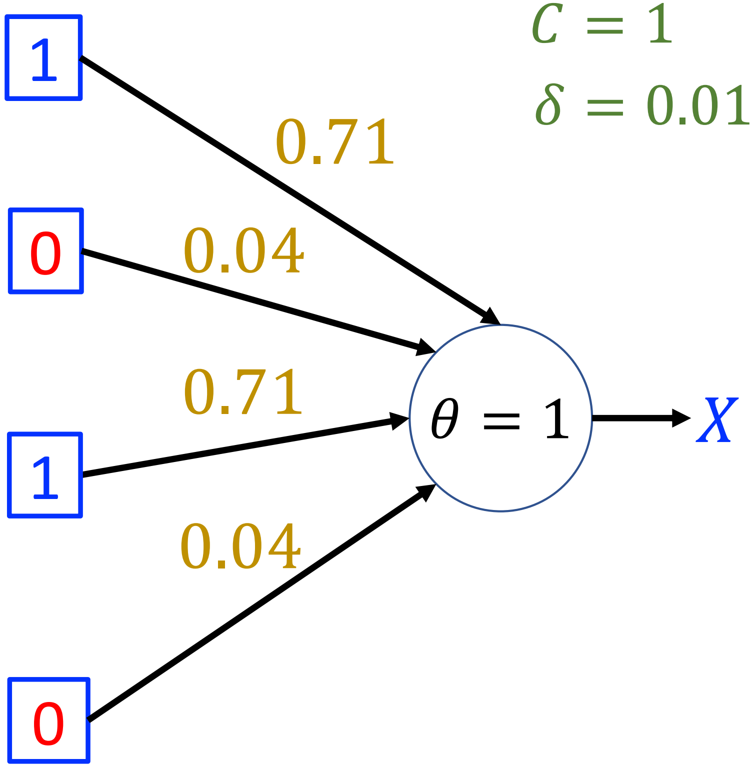
$\delta = 0.01$

$t = 3$

a_1^3	a_2^3	a_3^3	a_4^3
1	0	1	0

w_1^3	w_2^3	w_3^3	w_4^3
0.71	0.04	0.71	0.04

A Running Example



$C = 1$

$\delta = 0.01$

$t = 3$

a_1^3	a_2^3	a_3^3	a_4^3
1	0	1	0

w_1^3	w_2^3	w_3^3	w_4^3
0.71	0.04	0.71	0.04

$$S^3 = \sum_{i=1}^{4} w_i^3 a_i^3$$

$$= w_1^3 \times a_1^3 + w_2^3 \times a_2^3 + w_3^3 \times a_3^3 + w_4^3 \times a_4^3$$

$$= 0.71 \times 1 + 0.04 \times 0 + 0.71 \times 1 + 0.04 \times 0$$

$$= 1.42 \geq \theta$$

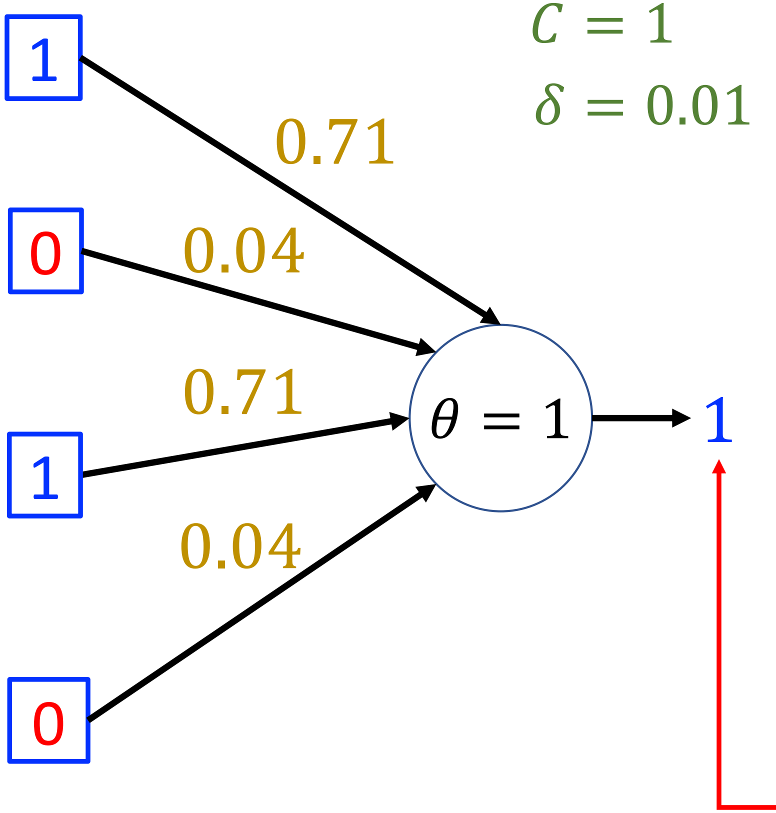
$$X^4 = 1$$

A Running Example

$C = 1$

$\delta = 0.01$

$t = 3$



a_1^3	a_2^3	a_3^3	a_4^3
1	0	1	0

w_1^3	w_2^3	w_3^3	w_4^3
0.71	0.04	0.71	0.04

$$S^3 = \sum_{i=1}^{4} w_i^3 a_i^3$$

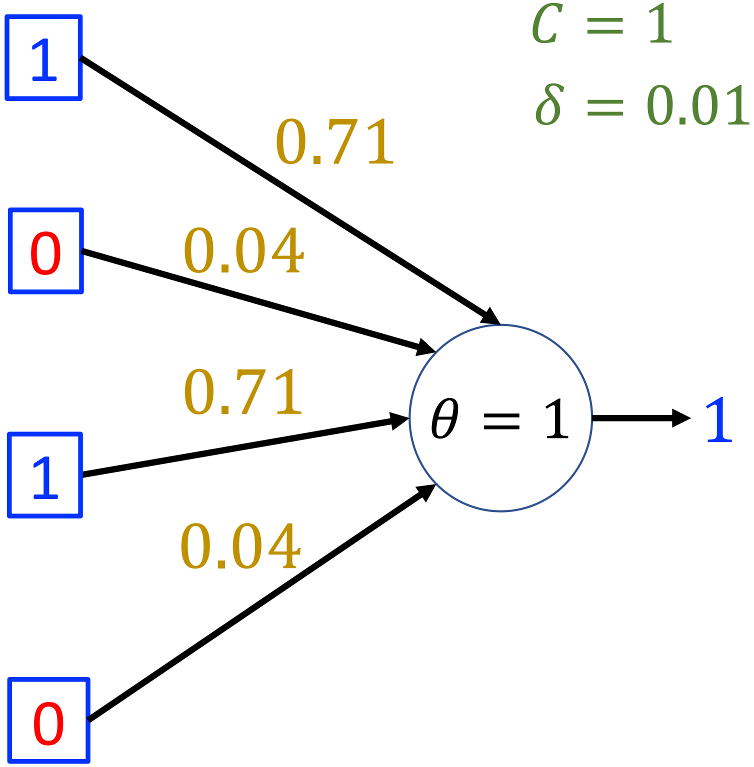
$$= w_1^3 \times a_1^3 + w_2^3 \times a_2^3 + w_3^3 \times a_3^3 + w_4^3 \times a_4^3$$

$$= 0.71 \times 1 + 0.04 \times 0 + 0.71 \times 1 + 0.04 \times 0$$

$$= 1.42 \geq \theta$$

$$X^4 = 1$$

A Running Example



$C = 1$

$\delta = 0.01$

$t = 3$

a_1^3	a_2^3	a_3^3	a_4^3
1	0	1	0

w_1^3	w_2^3	w_3^3	w_4^3
0.71	0.04	0.71	0.04

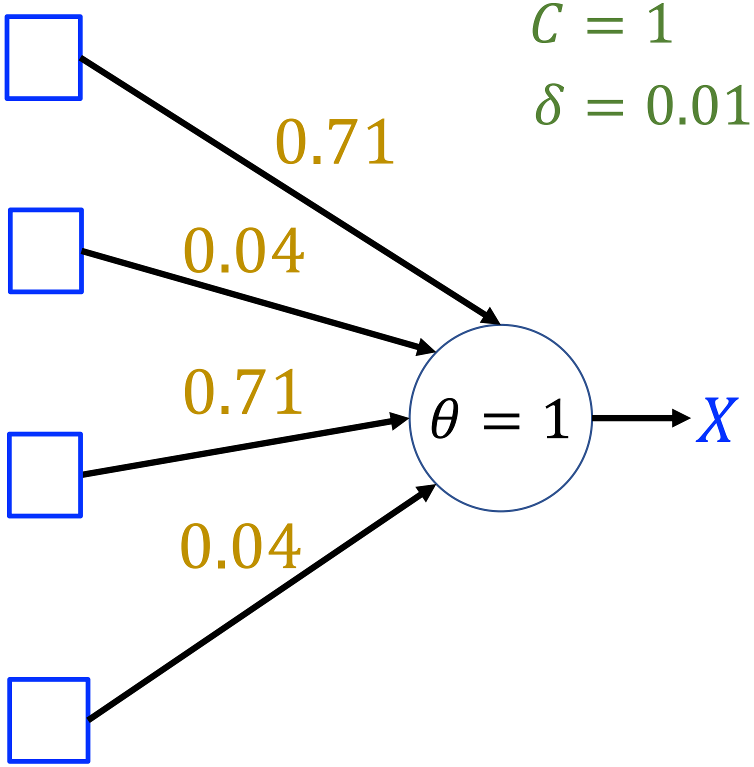
$$\Delta w_i^3 = C a_i^3 X^4$$

$$\Delta w_1^3 = 1\times1\times1 = 1,$$
$$\Delta w_2^3 = 1\times0\times1 = 0,$$
$$\Delta w_3^3 = 1\times1\times1 = 1,$$
$$\Delta w_4^3 = 1\times0\times1 = 1,$$

$$w_i^4 = w_i^3 + \Delta w_i^3$$

$$w_1^4 = w_i^3 + \Delta w_i^3 = 0.71 + 1 = 1.71;$$
$$w_2^4 = w_2^3 + \Delta w_2^3 = 0.04 + 0 = 0.04;$$
$$w_3^4 = w_3^3 + \Delta w_3^3 = 0.71 + 1 = 1.71;$$
$$w_4^4 = w_4^3 + \Delta w_4^3 = 0.04 + 0 = 0.04;$$

A Running Example



$C = 1$

$\delta = 0.01$

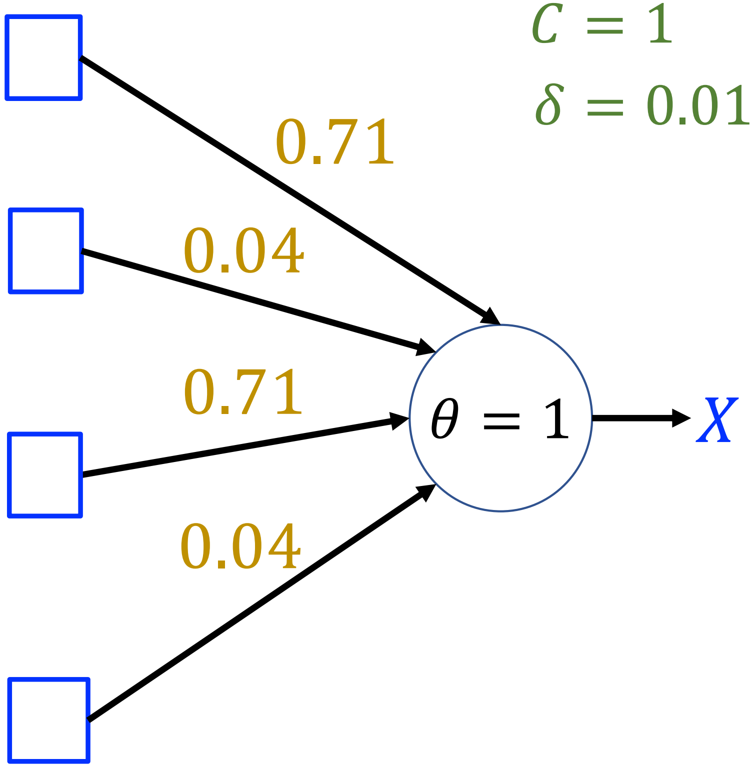
$t = 4$

w_1^4	w_2^4	w_3^4	w_4^4
1.71	0.04	1.71	0.04

A Running Example

$C = 1$ $\quad t = 4$

$\delta = 0.01$



w_1^4	w_2^4	w_3^4	w_4^4
1.71	0.04	1.71	0.04

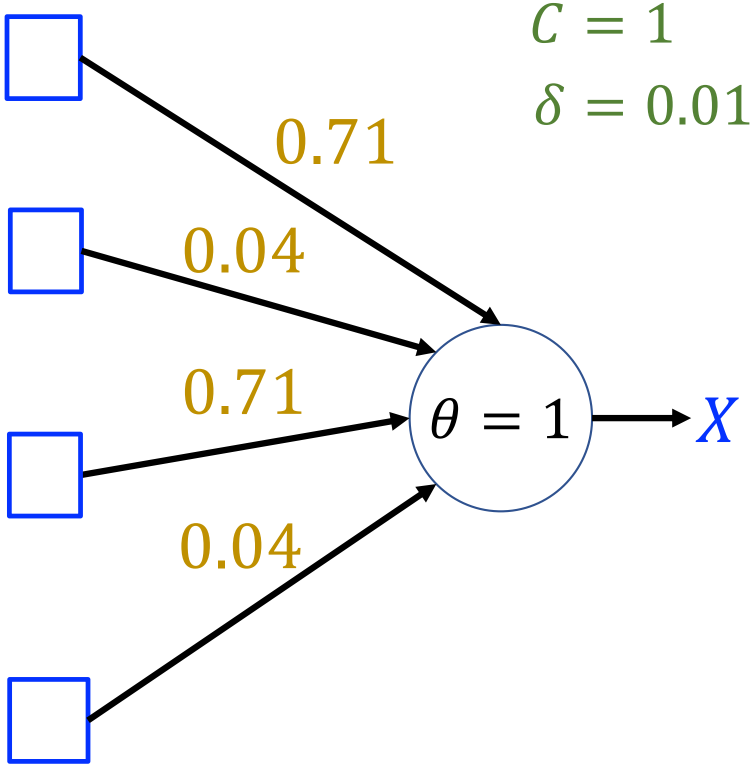
$$\|w^4\|_2 = \sqrt{\sum_i \left(w_i^4\right)^2}$$

$$= \sqrt{(w_1^4)^2 + (w_2^4)^2 + (w_3^4)^2 + (w_4^4)^2}$$

$$= \sqrt{(1.71)^2 + (0.04)^2 + (1.71)^2 + (0.04)^2}$$

$$\approx 2.42$$

A Running Example



$C = 1$

$\delta = 0.01$

$t = 4$

w_1^4	w_2^4	w_3^4	w_4^4
1.71	0.04	1.71	0.04

$\|w^4\|_2 \approx 2.42$

$w_1^4 = \frac{1}{\|w^4\|_2} w_1^4 = 0.71 \quad w_2^4 = \frac{1}{\|w^4\|_2} w_2^4 = 0.02$

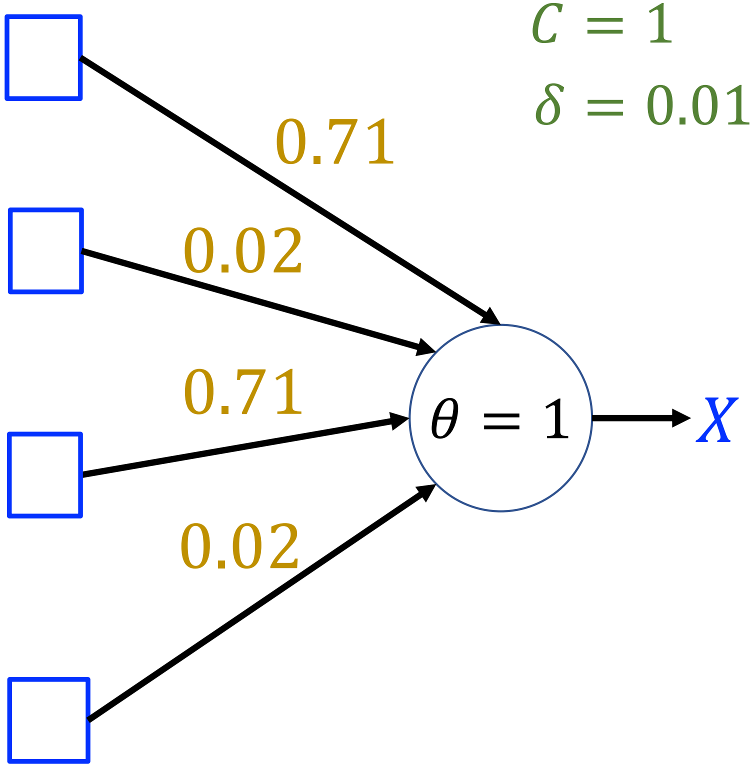
$w_3^4 = \frac{1}{\|w^4\|_2} w_3^4 = 0.71 \quad w_4^4 = \frac{1}{\|w^4\|_2} w_4^4 = 0.02$

A Running Example

$C = 1$

$\delta = 0.01$

$t = 4$



w_1^4	w_2^4	w_3^4	w_4^4
0.71	0.02	0.71	0.02

$\|w^4\|_2 \approx 2.42$

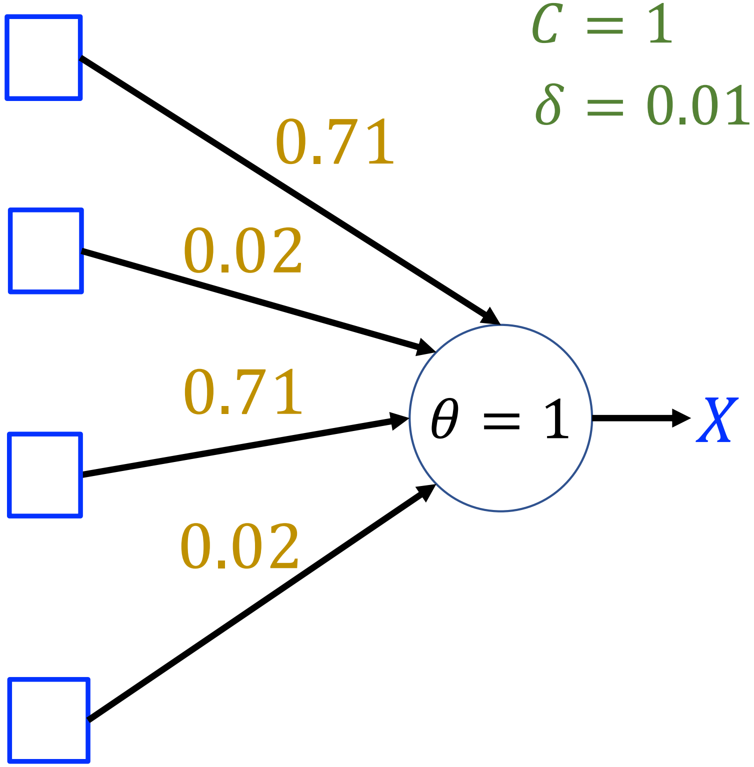
$$w_1^4 = \frac{1}{\|w^4\|_2} w_1^4 = 0.71 \qquad w_2^4 = \frac{1}{\|w^4\|_2} w_2^4 = 0.02$$

$$w_3^4 = \frac{1}{\|w^4\|_2} w_3^4 = 0.71 \qquad w_4^4 = \frac{1}{\|w^4\|_2} w_4^4 = 0.02$$

A Running Example

$C = 1$ $\quad t = 4$

$\delta = 0.01$



w_1^3	w_2^3	w_3^3	w_4^3
0.71	0.04	0.71	0.04
w_1^4	w_2^4	w_3^4	w_4^4
0.71	0.02	0.71	0.02

$\max_i |w_i^4 - w_i^3| \leq \delta$ **?**

$|w_1^4 - w_1^3| = 0.00$

$|w_2^4 - w_2^3| = 0.02$

$|w_3^4 - w_3^3| = 0.00$

$|w_4^4 - w_4^3| = 0.02$

$\Rightarrow \max_i |w_i^4 - w_i^3| = 0.02 > \delta$

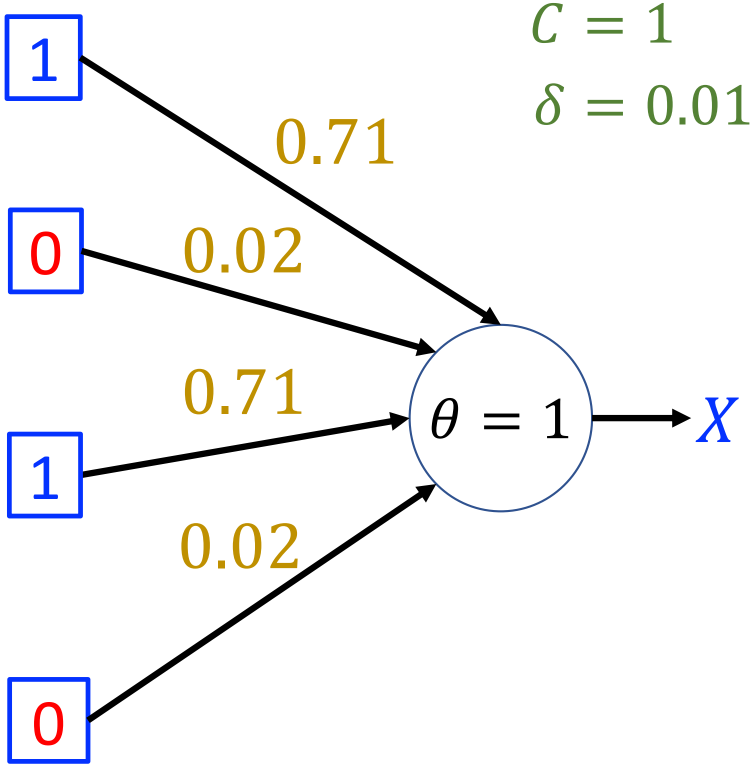
Very close. Continue.

A Running Example

$C = 1$

$\delta = 0.01$

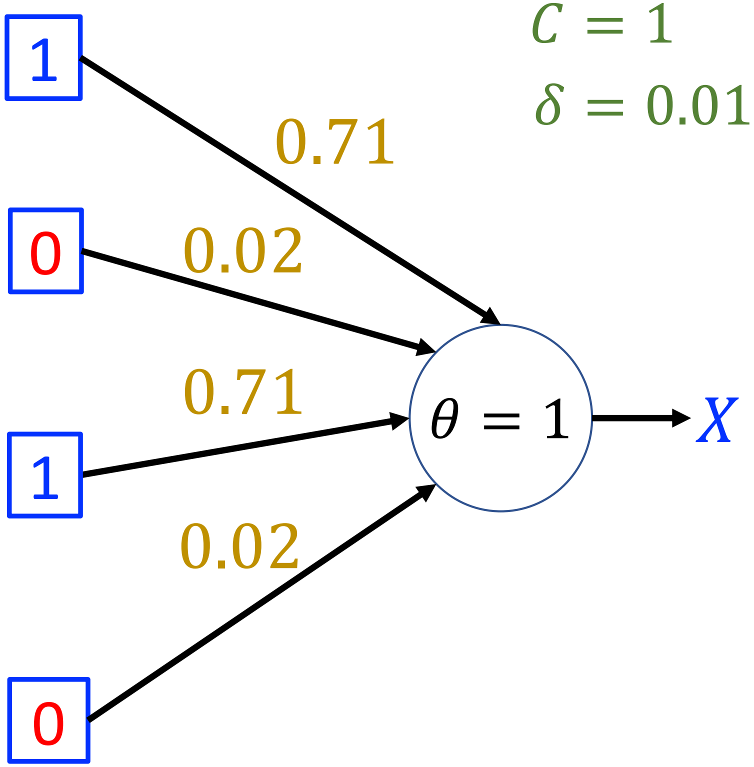
$t = 4$



a_1^4	a_2^4	a_3^4	a_4^4
1	0	1	0

w_1^4	w_2^4	w_3^4	w_4^4
0.71	0.02	0.71	0.02

A Running Example



$C = 1$

$\delta = 0.01$

$t = 4$

a_1^4	a_2^4	a_3^4	a_4^4
1	0	1	0

w_1^4	w_2^4	w_3^4	w_4^4
0.71	0.02	0.71	0.02

$$S^4 = \sum_{i=1}^{4} w_i^4 a_i^4$$

$$= w_1^4 \times a_1^4 + w_2^4 \times a_2^4 + w_3^4 \times a_3^4 + w_4^4 \times a_4^4$$

$$= 0.71 \times 1 + 0.02 \times 0 + 0.71 \times 1 + 0.02 \times 0$$

$$= 1.42 \geq \theta$$

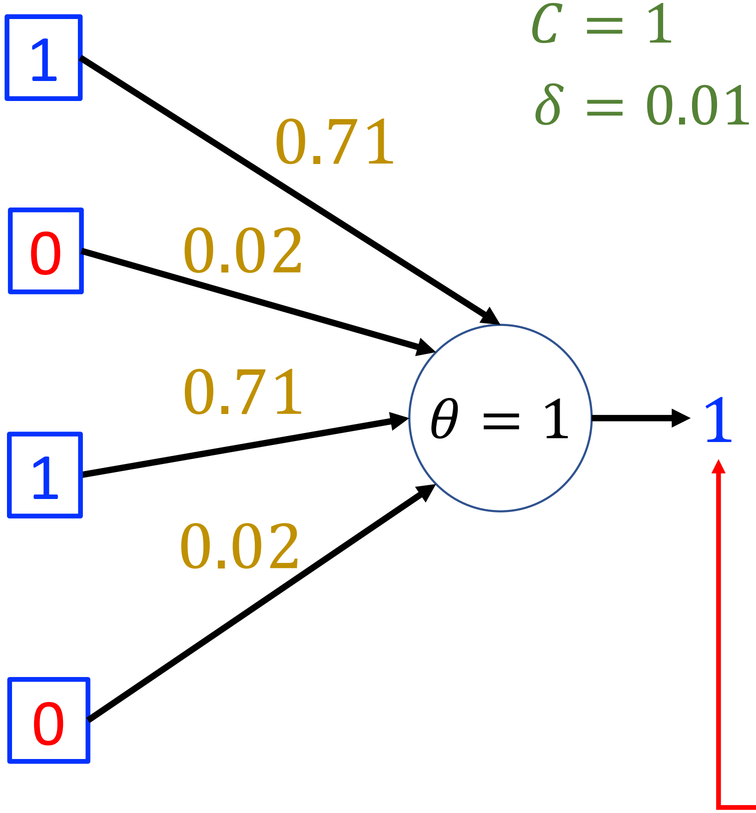
$$X^5 = 1$$

A Running Example

$C = 1$

$\delta = 0.01$

$t = 4$



a_1^4	a_2^4	a_3^4	a_4^4
1	0	1	0

w_1^4	w_2^4	w_3^4	w_4^4
0.71	0.02	0.71	0.02

$$S^4 = \sum_{i=1}^{4} w_i^4 a_i^4$$

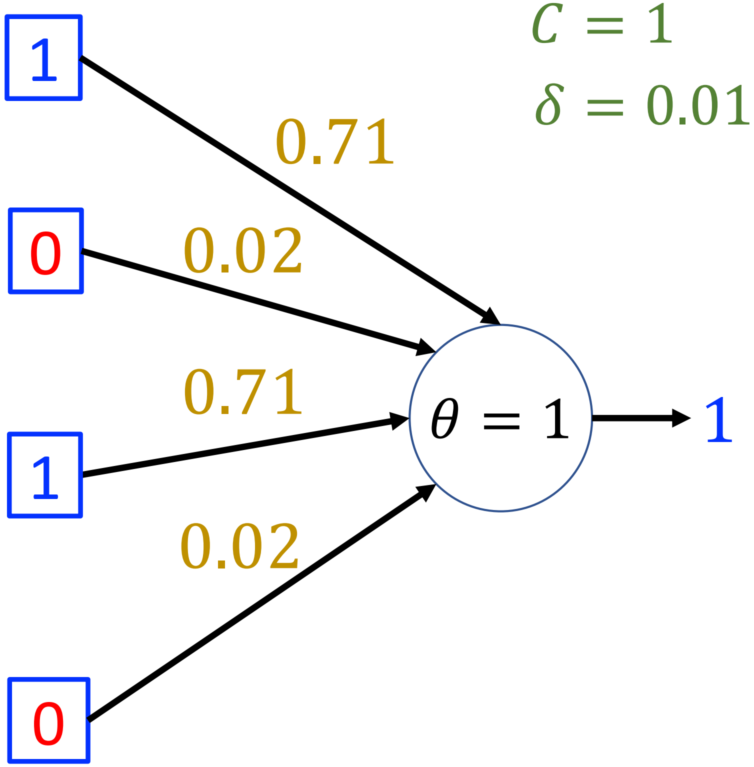
$$= w_1^4 \times a_1^4 + w_2^4 \times a_2^4 + w_3^4 \times a_3^4 + w_4^4 \times a_4^4$$

$$= 0.71 \times 1 + 0.02 \times 0 + 0.71 \times 1 + 0.02 \times 0$$

$$= 1.42 \geq \theta$$

$$X^5 = 1$$

A Running Example



$C = 1$

$\delta = 0.01$

$t = 4$

a_1^4	a_2^4	a_3^4	a_4^4
1	0	1	0

w_1^4	w_2^4	w_3^4	w_4^4
0.71	0.02	0.71	0.02

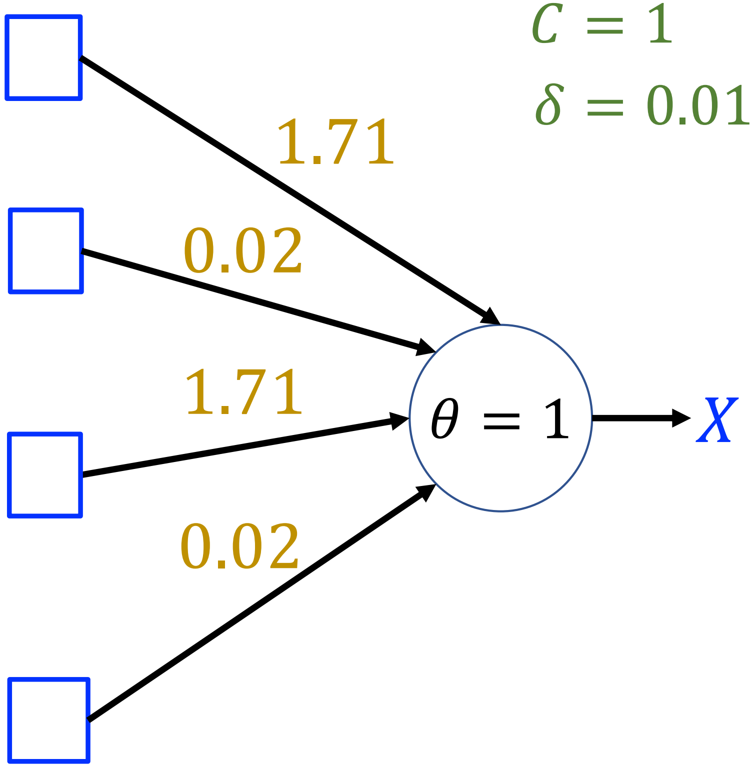
$$\Delta w_i^4 = C a_i^4 X^5$$

$$\Delta w_1^4 = 1\times1\times1 = 1,$$
$$\Delta w_2^4 = 1\times0\times1 = 0,$$
$$\Delta w_3^4 = 1\times1\times1 = 1,$$
$$\Delta w_4^4 = 1\times0\times1 = 1,$$

$$w_i^5 = w_i^4 + \Delta w_i^4$$

$$w_1^5 = w_i^4 + \Delta w_i^4 = 0.71 + 1 = 1.71;$$
$$w_2^5 = w_2^4 + \Delta w_2^4 = 0.02 + 0 = 0.02;$$
$$w_3^5 = w_3^4 + \Delta w_3^4 = 0.71 + 1 = 1.71;$$
$$w_4^5 = w_4^4 + \Delta w_4^4 = 0.02 + 0 = 0.02;$$

A Running Example



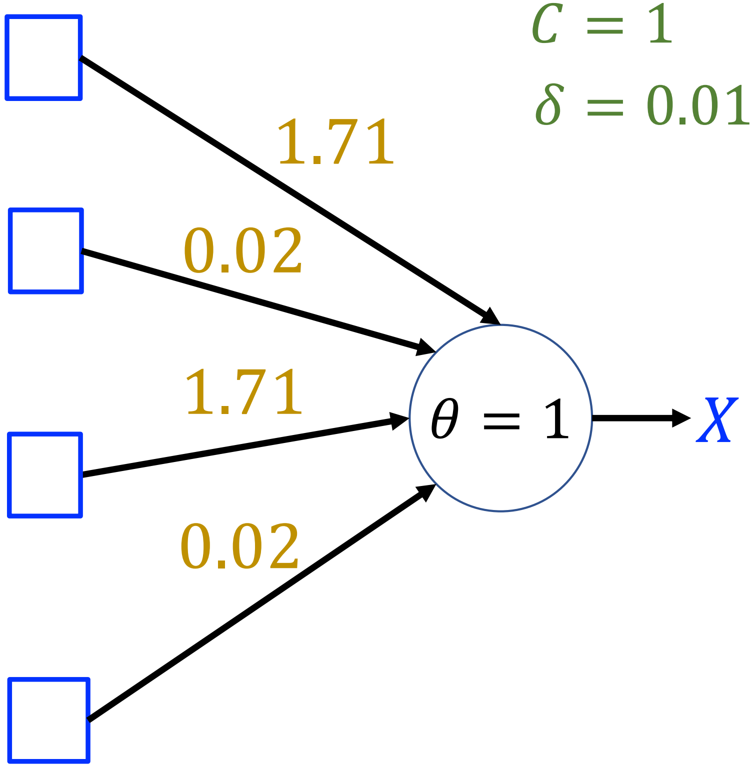
$C = 1$

$\delta = 0.01$

$t = 5$

w_1^5	w_2^5	w_3^5	w_4^5
1.71	0.02	1.71	0.02

A Running Example



$C = 1$

$\delta = 0.01$

$t = 5$

w_1^5	w_2^5	w_3^5	w_4^5
1.71	0.02	1.71	0.02

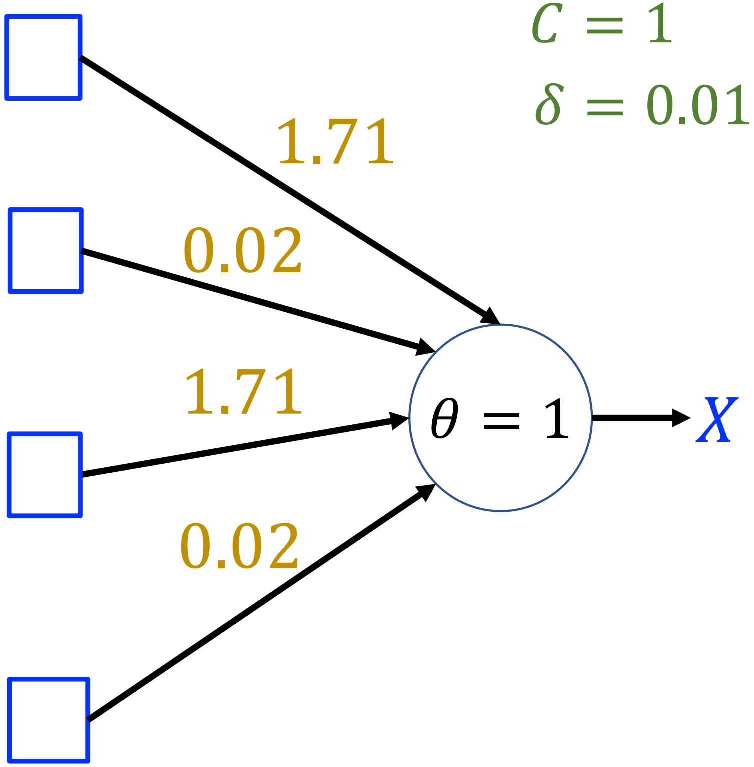
$$\|w^5\|_2 = \sqrt{\sum_i (w_i^5)^2}$$

$$= \sqrt{(w_1^5)^2 + (w_2^5)^2 + (w_3^5)^2 + (w_4^5)^2}$$

$$= \sqrt{(1.71)^2 + (0.02)^2 + (1.71)^2 + (0.02)^2}$$

$$\approx 2.42$$

A Running Example



$C = 1$

$\delta = 0.01$

$t = 5$

w_1^5	w_2^5	w_3^5	w_4^5
1.71	0.02	1.71	0.02

$$\|w^5\|_2 \approx 2.42$$

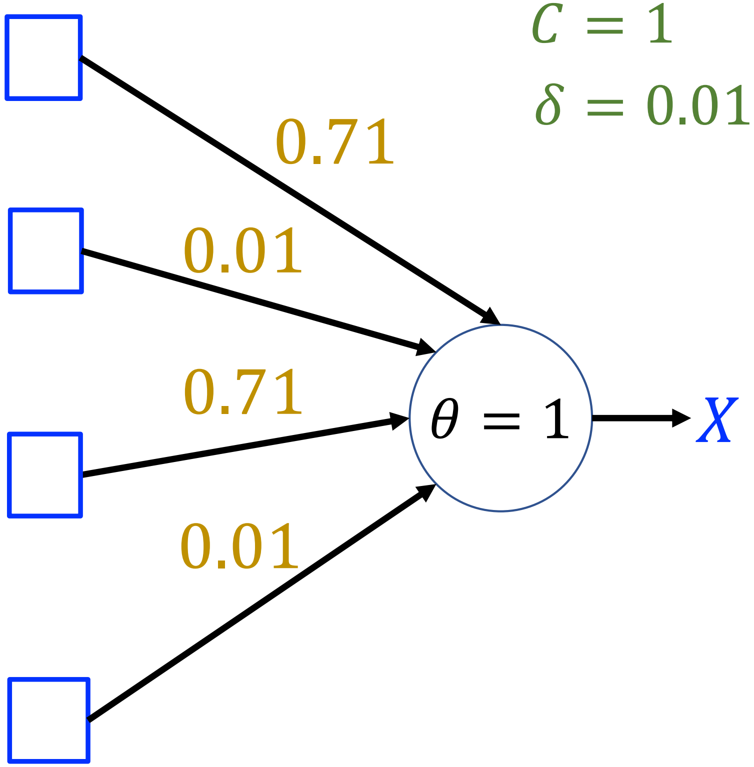
$$w_1^5 = \frac{1}{\|w^5\|_2} w_1^5 = 0.71 \quad w_2^5 = \frac{1}{\|w^5\|_2} w_2^5 = 0.01$$

$$w_3^5 = \frac{1}{\|w^5\|_2} w_3^5 = 0.71 \quad w_4^5 = \frac{1}{\|w^5\|_2} w_4^5 = 0.01$$

A Running Example

$C = 1$ $t = 5$

$\delta = 0.01$



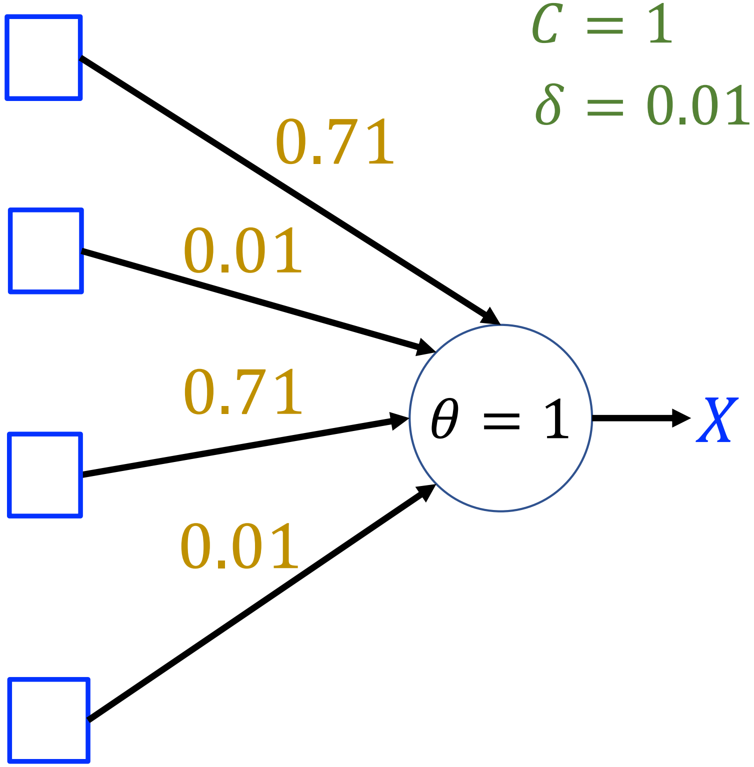
w_1^5	w_2^5	w_3^5	w_4^5
0.71	0.01	0.71	0.01

$\|w^5\|_2 \approx 2.42$

$$w_1^5 = \frac{1}{\|w^5\|_2} w_1^5 = 0.71 \qquad w_2^5 = \frac{1}{\|w^5\|_2} w_2^5 = 0.01$$

$$w_3^5 = \frac{1}{\|w^5\|_2} w_3^5 = 0.71 \qquad w_4^5 = \frac{1}{\|w^5\|_2} w_4^5 = 0.01$$

A Running Example



$C = 1$

$\delta = 0.01$

$t = 5$

w_1^4	w_2^4	w_3^4	w_4^4
0.71	0.02	0.71	0.02

w_1^5	w_2^5	w_3^5	w_4^5
0.71	0.01	0.71	0.01

$$\max_i \left| w_i^5 - w_i^4 \right| \leq \delta \quad ?$$

$$\left| w_1^5 - w_1^4 \right| = 0.00$$
$$\left| w_2^5 - w_2^4 \right| = 0.01$$
$$\left| w_3^5 - w_3^4 \right| = 0.00$$
$$\left| w_4^5 - w_4^4 \right| = 0.01$$

$$\Rightarrow \max_i \left| w_i^5 - w_i^4 \right| = 0.01 \leq \delta$$

Finally. We can STOP!!