2024. Jan. 11th

Computing w_{ij} and θ_i is very troublesome.

↓ We like to find $\boxed{\text{a program}}$ to compute them.

$$E(x_1, x_2, \dots, x_N) = -\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij} x_i x_j + \sum_{i=1}^{N} \theta_i x_i + C$$

The standard form of the energy.

(1) At first we compute C

$$E(0, 0, \dots, 0) = C$$

C is computed just by substituting 0 to all ~~w_{ij} and θ_i, $i=1\sim N, j=1\sim N$.~~ x_i, $i=1\sim$

(2) Then we compute θ_i

$$\theta_i = E(\underset{x_1}{0}, \underset{x_2}{0}, \dots, \underset{x_{i-1}}{0}, \underset{x_i}{1}, \underset{x_{i+1}}{0}, \dots, \underset{x_N}{0}) - C$$

?

Only x_i is 1, other x_j, $j \neq i = 0$

(3) Finally We can compute w_{ij}

$$E(\underset{x_1}{0}, \underset{x_2}{0}, \cdots, \underset{x_{i-1}}{0}, \underset{x_i}{1}, \underset{x_{i+1}}{0}, \cdots, \underset{x_{j-1}}{0}, \underset{x_j}{1}, \underset{x_{j+1}}{0}, \cdots, \underset{x_N}{0}) \quad — \text{☆}$$

Only x_i and x_j take 1

and other x_k, $k \neq i, j$ take 0.

Substitute 1 to x_i and x_j

Sabstitute 0 to $x_{k \neq i, j}$,

No self connection

$w_{kk} = 0$

symmetric connection

$w_{ij} = w_{ji}$

$$= -\frac{1}{2}\Big(\underbrace{w_{12} \cdot \underset{0}{x_1} \cdot \underset{0}{x_2}}_{=0} + \underbrace{w_{13} \cdot \underset{0}{x_1} \cdot \underset{0}{x_3}}_{=0} + \cdots \underbrace{w_{1i} \cdot \underset{0}{x_1} \underset{1}{x_i}}_{=0}$$

$$+ \underbrace{w_{ij}\, \underset{1}{x_i}\, \underset{1}{x_j}}_{=w_{ij}} + \underbrace{w_{ij+1}\, \underset{1}{x_i}\, \underset{0}{x_{j+1}}}_{=0} + \underbrace{\cdots}_{=0} \quad 0$$

Scanned with CamScanner

$$+ \underbrace{w_{ji}\, \underbrace{x_j}_{=1}\, \underbrace{x_i}_{=1}}_{= w_{ji}} + \cdots - \underbrace{(\cdots)}_{=0} \underbrace{(\cdot)}_{=0} + \underbrace{w_{ij}\, x_N x_N}_{=0} + \underbrace{w_{ij}\, x_N x_N}_{=0}$$

$$+ \Big(\underbrace{\theta_1 x_1}_{=0} + \underbrace{\theta_2 x_2}_{=0} + \cdots + \underbrace{\theta_{i-1} x_{i-1}}_{=0} + \underbrace{\theta_i \underbrace{x_i}_{=1}}_{=\theta_i} + \cdots + \underbrace{\theta_{j-1} x_{j-1}}_{=0} + \underbrace{\theta_j x_j}_{\theta_j} + \underbrace{\theta_{j+1} x_{j+1}}_{0} + \cdots + \underbrace{\theta_N x_N}_{0} \Big)$$

$$+ C$$

$$= -\frac{1}{2}\left(w_{ij} + w_{ji} \right) + \theta_i + \theta_j + C \quad (A)$$

symmetric $w_{ij} = w_{ji}$

$$= -w_{ij} + \theta_i + \theta_j + C$$

$$E(0,0,\cdots 0,\underset{\underset{x_i}{=}}{1},0,\cdots,0,\underset{\underset{x_j}{=}}{1},0\cdots 0)$$

$$= -w_{ij} + \theta_i + \theta_j + C$$

We can compute w_{ij} by

$$w_{ij} = \theta_i + \theta_j + C - E(0,0,\cdots 0\,\underset{\underset{x_i}{=}}{1},0\cdots 0,\underset{\underset{x_j}{=}}{1},0\cdots 0)$$

From (1), (2) and (3), we can make a program to compute C, θ_i and w_{ij} $i = 1 \sim N$, $j = 1 \sim N$

(1) $C = E(0,0,\cdots,0)$

(2) $\theta_i = E(0,0,\cdots 0,1,0,\cdots 0) - C$

$x_i = 1$, $x_k = 0$ $k \neq i$

(3) $w_{ij} = \theta_i + \theta_j + C - E(0,0,\cdots,0,1,0,\cdots 0,1,0,\cdots 0)$

$x_i = 1$ $x_j = 1$ $x_k = 0$ $k \neq i,j$

What should we do at first?

We have to make a cost function at first, so that the minimum of the cost function can give the solution of the simultaneous equation.

$$E(x_1, x_2, x_3, x_3) = (x_1 - 2x_2 + x_3 - x_4 - 1)^2 + (-x_1 - x_2 + x_3 + 2x_4 - 2)^2 + (2x_1 + x_2 - x_3 + x_4 - 2)^2 + (-2x_1 + 3x_2 - 2x_3 - x_4 + 5)^2$$

This part takes 0 as its minimum. It becomes 0 when the first equation is satisfied.

This cost function is also the energy function of RNN.

By reducing this cost (energy) function and finding its minimum (0), we can solve the simultaneous equation.

Task 4-1

Make a program to compute C, θ_i, $i=1\sim N$, w_{ij}, $i=1\sim N$, $j=1\sim N$ for the energy function of Task 3.

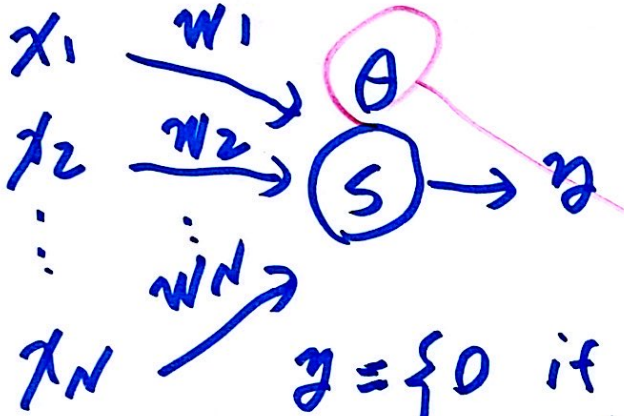
$$
\begin{aligned}
E(x_1, x_2, x_3, x_4) = & (x_1 - 2x_2 + x_3 - x_4 - 1)^2 \\
& + (-x_1 - x_2 + x_3 + 2x_4 - 2)^2 \\
& + (2x_1 + x_2 - x_3 + x_4 - 2)^2 \\
& + (-2x_1 + 3x_2 - 2x_3 - x_4 + 5)^2
\end{aligned}
$$

For this function, we already computed w_{ij}, θ_i, and C by our hand.

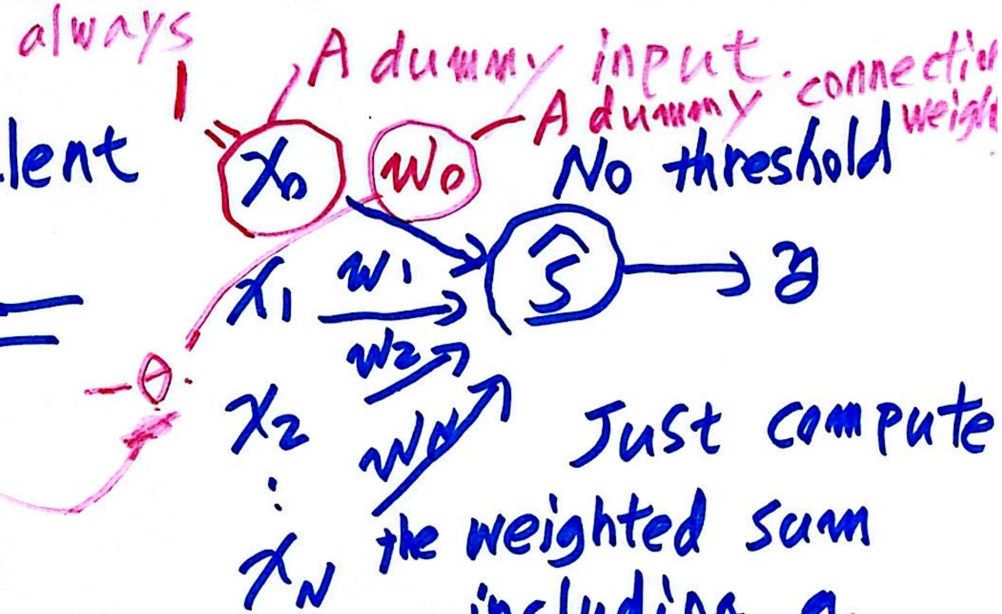
Please check if you can get the same results.

$$
\begin{aligned}
C = E(0,0,\cdots,0) &= (0-2\times 0+0-0-1)^2 \\
&\quad + (-0-0+0+2\times 0-2)^2 \\
&\quad + (2\times 0+0-0+0-2)^2 \\
&\quad + (-2\times 0+3\times 0-2\times 0-0+5)^2 \\
&= (-1)^2+(-2)^2+(-2)^2+(5)^2 \\
&= \cancel{35}\ 34
\end{aligned}
$$

$$
\begin{aligned}
\theta_1 &= E(1,0,\cdots,0) - 34 \\
&= (1-2\times 0+0-0-1)^2 \\
&\quad + (-1-0+0+2\times 0-2)^2 \\
&\quad + (2\times 1+0-0+0-2)^2 \\
&\quad + (-2\times 1+3\times 0-2\times 0-0+5)^2 \\
&\quad - 34 \\
&= (0)^2+(-3)^2+(0)^2+(3)^2-34 = -16
\end{aligned}
$$



Equivalent



$$y = \begin{cases} 0 & \text{if } ran() > p \\ 1 & \text{else} \end{cases}$$

$$p = \text{sigmoid}(S - \theta) = \hat{S}$$

$$S = \sum_{i=1}^{N} w_i x_i$$

the same S

Just compute the weighted sum including a dummy input and a dummy connection / weight.

$$\hat{S} = \sum_{i=0}^{N} w_i x_i$$

— start from the dummy.

$$= \underbrace{w_0}_{=-\theta} \underbrace{x_0}_{=1} + \underbrace{\sum_{i=1}^{N} w_i x_i}_{S}$$

$$= S - \theta$$

The energy function can be simple if you use the dummy input and the dummy connection.

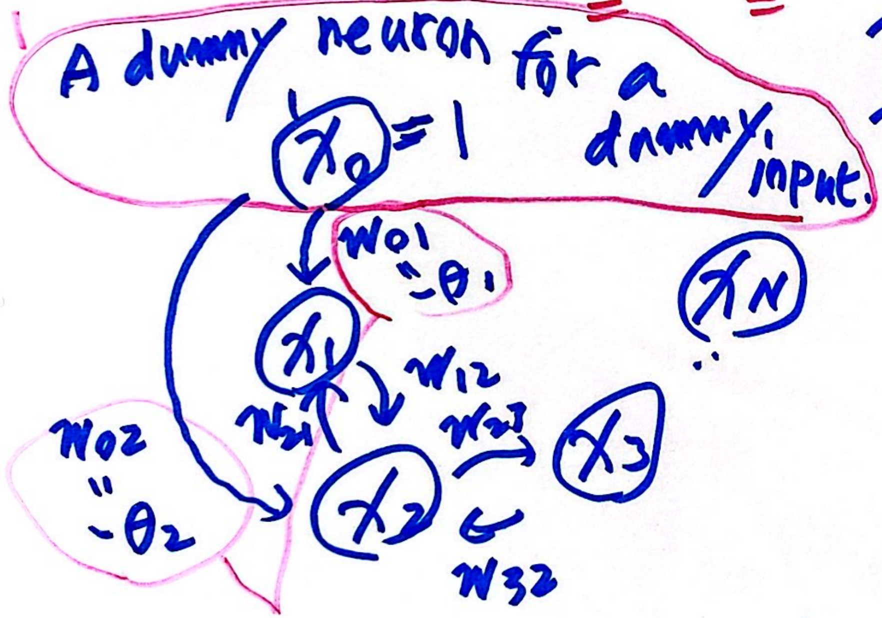
$$E = -\frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} w_{ij} x_i x_j + \sum_{i=1}^{N} \theta_i x_i + C$$

A dummy neuron

$$= -\frac{1}{2} \sum_{i=0}^{N} \sum_{j=0}^{N} w_{ij} x_i x_j + C$$

$$\left.\begin{aligned} w_{00} &= 0 \\ w_{i0} &= w_{0i} = -\theta_i \end{aligned}\right\} \text{symmetric condition}$$

$$x_0 = 1$$



Dummy connections for thresholds

The idea of "the winner takes all" can be used for Eight Queen Problem.

Along the horizontal (column) line (row) (vertical) (diagonal)
only one Queen can exist.

↳ only one neuron x_{ij} can take 1,
and other neurons on the same horizontal (vertical) (diagonal) line must take 0.

same as "the winner takes all".

For i-th horizontal line,
only one x_{ij} takes 1 and other $x_{ik} = 0$ ($k \neq j$)

↳ The energy function should be

$$(x_{i1} + x_{i2} + \cdots + x_{i8} - 1)^2$$

Now, you got a program to compute w_{ij}, θ_i and C, $i=1 \sim N$, $j=1 \sim N$

You can apply RNN to more complex problems.

I like to explain some tips to make your program simple.

So far, we used w_{ij} and θ_i separately.
So we have to make different procedures for w_{ij} and θ_i.

But by using a tip (trick), θ_i can be dealt as a kind of w_{ij}.
↳ We can use a common procedure for w_{ij} and θ_i.

This energy function ~~has~~ takes its minimum 0

when only one of $x_{i1}, x_{i2}, \cdots x_{i8}$ takes 1

and others take 0.

For example:

$$(\underset{x_{i1}}{0} + \underset{x_{i2}}{0} + \cdots \underset{x_{ij-1}}{0} + \underset{x_{ij}}{1} + \underset{x_{ij+1}}{0} \cdots + \underset{x_{i8}}{0} - 1)^2 = \underline{\underline{0}}$$

0 is the minimum because the squared value is always 0 or positive.

The energy function for the horizontal lines $i = 1 \sim 8$

$$E_h(\underbrace{x_{11}, \cdots, x_{88}}_{64}) = \sum_{i=1}^{8} \left(\sum_{j=1}^{8} x_{ij} - 1 \right)^2$$

↳ To avoid horizontal conflicts among eight Queens, this energy function must take 0 (mini

Now we are ready for difficult optimization problems.

Today : Eight Queen Problem

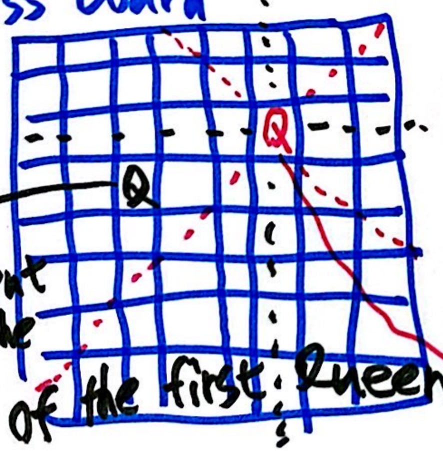
Tomorrow : Travelling Salesman Problem.

Eight Queen Problem (Common Problem for Programming Exercise)

Queen of Chess

Like this, we have to put Eight Queens on the chess board so that they do not conflict with each other.

Chess Board



Put Eight Queens on this Chess Bord avoiding the confliction among them.

The second Queen must be put avoiding the influence of the first Queen.

This Queen has its influence along the horizontal, vertical and diagonal lines.

We need to find the location of Eight Queens.

The location is represented using neurons.

$$\begin{array}{cc} & \begin{array}{ccccc} j=1 & j=2 & j=3 & \cdots & j=8 \end{array} \\ \begin{array}{c} i=1 \\ i=2 \\ \vdots \\ i=8 \end{array} & \underbrace{\left.\begin{array}{ccccc} x_{11} & x_{12} & x_{13} & \cdots & x_{18} \\ x_{21} & x_{22} & x_{23} & \cdots & x_{28} \\ \vdots & & \vdots & & \vdots \\ x_{81} & x_{82} & x_{83} & \cdots & x_{88} \end{array}\right\} 8}_{8} \end{array}$$

We use $8 \times 8 = 64$ neurons.

↓

Represent 8×8 locations (cells) on the chess Board.

x_{ij} ⇒ Using these two indexes, we can specify the position of the cell.

= =

Two indexes to represent position (i, j)

If a Queen is on a cell reprsented by x_{ij}, then x_{ij} takes the value

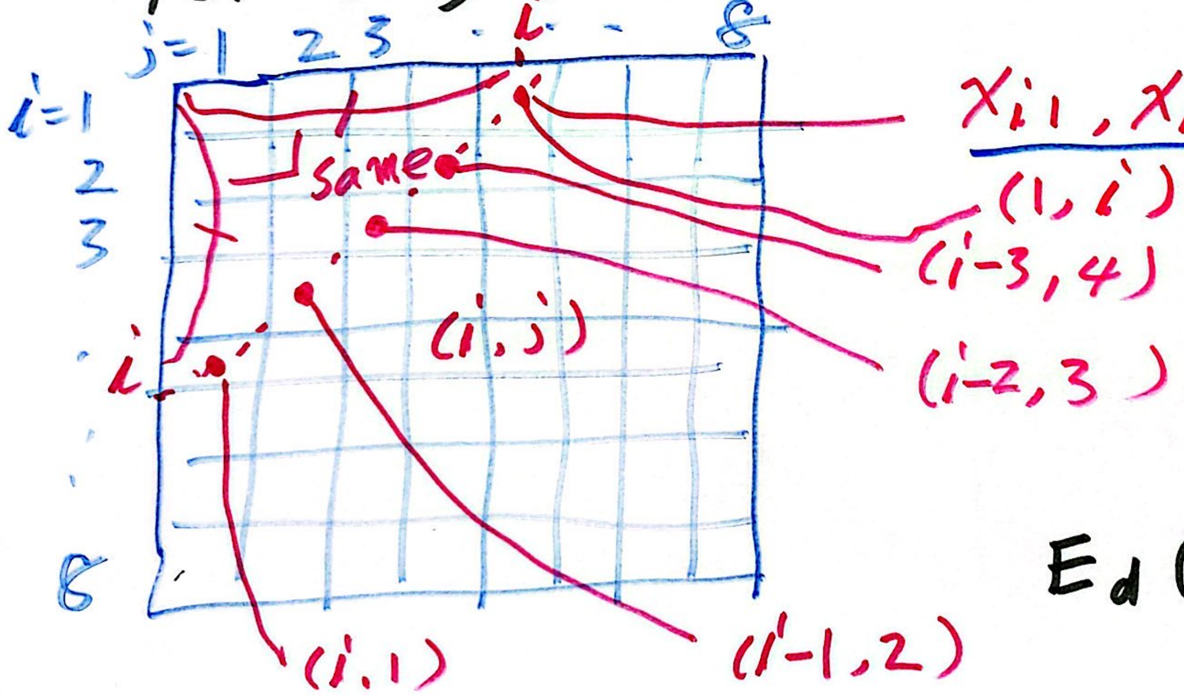
and if there is no Quee

on the cell then x_{ij} takes the value 0.

For vertical conflicts, we can use the following energy function:

$$E_v(\underbrace{x_{11}, \cdots, x_{88}}_{64}) = \sum_{j=1}^{8} \left(\sum_{i=1}^{8} x_{ij} - 1 \right)^2$$

For diagonal conflics.



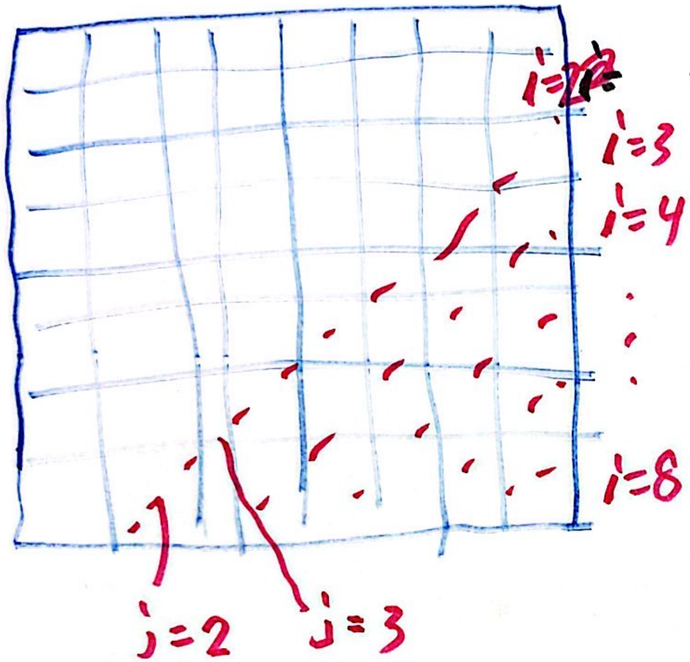
$\underline{x_{i1},\ x_{i-1,2},\ x_{i-2,3},\ \cdots,\ x_{1,i}}$

↓ Among these neur

Only one neuron can take 1 and other neurons take

$$E_d(x_{11}, \cdots, x_{88}) = \sum_{i=1}^{8} \cdots$$

$$E_d(x_{11},\dots x_{88}) = \sum_{i=1}^{8}\left(\sum_{k=1}^{i} x_{i-k+1,k} - 1\right)^2 \rightarrow$$



$$+ \sum_{i=2}^{8}\Big(\rightarrow$$